Adaptation of the 36-Month Ages and Stages Questionnaire in Taiwan: Results From a Preliminary Study

HUEI-LING AGNES TSAI
National Pingtung University of Science and Technology

MEGAN M. MCCLELLAND & CLARA PRATT
Oregon State University

JANE SQUIRES
University of Oregon

Identification of children with developmental disabilities is the first critical step in providing early intervention services. Currently, only 20% of Taiwanese children who could potentially benefit from early intervention have been identified. One possible reason for this low identification rate is the lack of a culturally appropriate, developmental screening instrument for Taiwanese children. A multi-pronged pilot study was conducted to explore the reliability and validity of scores using a Chinese translation of the 36-month version of the Ages and Stages Questionnaire (ASQ) with a sample of Taiwanese children. Results indicated that the translated Chinese version of the 36-month ASQ demonstrates initial cultural appropriateness, and reliability and validity of scores when used with a sample of 3-year-old Taiwanese children. The potential viability of the 36-month ASQ as a developmental screening instrument to identify Taiwanese children in need of early intervention services and the need for further validation research are discussed.

In the United States, early intervention services for children with developmental delays promote children’s optimal health and development (American Academy of Pediatrics, 2002; Florian, 1995; Guralnick, 1997; Ramey & Ramey, 1998). Recently, there have been legislative efforts to stimulate the development of early intervention services in Taiwan. In 1993, the Taiwanese national government mandated in the Revised Child Welfare Law that local governments provide early intervention services for children with developmental delays (Liao, Lee, Wu, Lin, & Kau, 1999). In 1997, the Disability Welfare Law was revised, requiring the Taiwanese health service system to provide early identification, early medical intervention, and health insurance services to individuals with disabilities (Liao et al., 1999). In that same year, a new special education law required that all 3-year-old children with disabilities attend preschool by 2003 (Kang, Lovett, & Haring, 2002).

Despite these legislative mandates, early intervention services have not reached all potentially eligible children. Based on World Health Organization population estimates, about 166,000 Taiwanese children under 6 years of age are likely to have developmental delays, yet only 30,000 children have been registered for services (Children’s Bureau, 2004). Thus, over 80% of Taiwanese young children with potential developmental delays have not been identified effectively and
registered for appropriate early intervention services.

Many factors contribute to this low rate of identification and service, including negative cultural attitudes toward disabilities (Kang, Lovett, & Haring, 2002; Tsai & Rosenkoetter, 2003), limited early intervention services, and the absence of a reliable and valid developmental screening instrument for Taiwanese children. For those pediatricians who conduct developmental screenings during well-baby visits, the most often used measures are the Denver Developmental Screening Test (Denver II; Frankenburg, Dodds, Archer, Shapiro, & Bresnick, 1990), the Chinese Children Development Inventory (CCDI; Shu, Su, & Shiao, 1978), and professional knowledge (Wann, 1993). Norms for the CCDI, however, were established over 25 years ago (Shu et al., 1978), the validity and reliability of scores for Taiwanese children on the Denver II have not been established, and professional knowledge alone does not result in accurate developmental assessment (Glascoe, 2000a; Glascoe, 2000b; Glascoe, 1997; Glascoe, Martin, & Humphrey, 1990). Without standardized procedures, professionals such as pediatricians and preschool teachers might overlook subtle differences in developmental domains leading to delays in referral. A systematic and cost-effective developmental screening system is needed that utilizes assessment tools with established reliability and validity of scores.

Theoretical perspectives such as ecological systems theory (Bronfenbrenner & Morris, 1998) and dynamic systems theory (Thelen, 2000) stress the importance of cultural contexts in children’s development. Research indicates that culture can influence parenting beliefs and practices in ways that cause children’s development to differ from culture to culture (Cheng & Tsai, 2002; Solomons & Solomons, 1975; Super, 1976). Thus, it is questionable to adapt a developmental screening tool without first establishing its validity and reliability in a new cultural setting, such as the current practice of using the Denver II in Taiwan. The present study examined the reliability and validity of scores of a sample of Taiwanese children on the 36-month Ages and Stages Questionnaire (ASQ). The ASQ is a developmental screening tool that is used widely and successfully in the U.S. and other nations, to screen children for developmental delays.

The Ages and Stages Questionnaires

The ASQ is a series of parent-report, age-graded developmental screening instruments. The ASQ screening system is composed of 19 questionnaires (Squires, Potter, & Bricker, 1999) that can be completed by parents when children are 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 27, 30, 33, 36, 42, 48, 54, and 60 months of age. Typically, the ASQ is completed by parents and thus, also can serve to enhance parents’ knowledge about their child’s development (Diamond & Squires, 1993; Squires, 1996; Squires & Bricker, 1991).

Each ASQ questionnaire consists of 30 items with six items in each of five developmental domains: communication, gross motor, fine motor, problem solving, and personal-social. Test items are written in simple and straightforward language, asking parents to assess their child’s current behavior. For example, one item from the gross motor domain of the 36-month ASQ asks, “Does your child jump with both feet leaving the floor at the same time?” An item from the personal-social domain asks, “Does your child take turns by waiting while another child or adult takes a turn?” Parents are asked to check yes, sometimes, or not yet on the item according to their everyday observation of the child’s performance. For many items, simple illustrations are provided to assist parents and other caregivers in understanding the behavior of interest. Unlike the Denver II (Frankenburg et al., 1990), the ASQ does not require specialized training to administer, which is a powerful, practical advantage. The major requirement is that the parent answering the questionnaires be able to read at a fourth to sixth grade reading level.

The reliability and validity of ASQ scores have been well established in the U.S.
(Bricker & Squires, 1989a; Bricker & Squires, 1989b; Bricker, Squires, Kaminski, & Mounts, 1988; Dale, Bates, Reznick, & Morisset, 1989; Squires, 2000; Squires & Bricker, 1991; Squires, Katzev, & Jenkins, 2002). According to the ASQ technical manual (Squires et al., 1999), interobserver reliability, measured as percentage agreement between classifications based on the questionnaires completed by 112 parents and those completed by 2 professional examiners, was 94%. Test-retest reliability completed by 175 parents at 2 to 3 weeks intervals was 94%. Validity of ASQ scores has been established using the Revised Gesell and Armatruda Developmental and Neurological Examination (Knobloch, Stevens, & Malone, 1980), Bayley Scales of Infant Development (Bayley, 1969), McCarthy Scales of Children’s Abilities (McCarthy, 1972), Stanford Binet (Thorndike, Hagen, & Sattler, 1985), and Battelle Developmental Inventory (Newborg, Stock, Wnek, Guidubaldi, & Svinicki, 1987). Concurrent validity as reported in percent agreement between questionnaires and standardized assessments ranged from 76% for the 4 month ASQ to 91% for the 36 month ASQ, with 84% overall agreement (N = 1613; Squires et al., 1999).

Cutoff points establish the ASQ scores on which children are identified as below the “typically developing” range and should be recommended for more intensive follow-up assessment. A score that is 2 SD below the mean is recommended as the principal cutoff point for identification of possible developmental delays (Squires et al., 1999). In addition, the ASQ scoring method allows for omission of culturally inappropriate items so that a child will not be penalized because of an inappropriate test item (Squires et al., 1999).

The ASQ is available in English, Spanish, Norwegian, French, and Korean (Squires et al., 1999). Several cross culture adaptations of the ASQ have been undertaken and these translations have been found to be reliable in Norwegian and French (Janson, 2002; Janson & Squires, 2004; Dionne, Squires, Leclerc, & Peloquin, 2006). Presently, the validity and reliability of ASQ scores are being established in Korea and Canada.

Although Squires et al. (1999) stress the cultural adaptability of the ASQ, the U.S. normative sample (N = 1287) included few Asian/Pacific Islanders (0.4%). Given that more than a dozen countries (e.g., Japan, Korea, India, Philippines, China, and Taiwan) with diverse cultures encompass the category of Asian/Pacific Islander, a normative U.S. sample of 5 children is inadequate to generalize to all Asian and Pacific island cultures. Research and recent developmental theories argue that because of the differences in Asian/Pacific Islander cultures, it is inappropriate to use the ASQ in any country until it has been standardized and scores have been tested for validity and reliability in that country (McLean, Bailey, & Wolery, 1996; Vazir, Lansdown, Naidu, Vidyasagar, & Reddy, 1994).

Research Goals and Hypotheses
This is the first study to examine the cultural appropriateness and psychometric properties of ASQ scores in a sample of Taiwanese children. The study investigated the preliminary cultural appropriateness and initial reliability and validity of scores from the 36-month ASQ for a sample of 3-year-old children in Taiwan. Three research goals were addressed: (a) to examine whether the 36-month ASQ appeared to be culturally appropriate for 3-year-old Taiwanese children, (b) to test the initial reliability and internal consistency of scores on the 36-month ASQ with this sample, and (c) to check for the initial content and criterion validity of the 36-month ASQ in Taiwan. The 36-month ASQ was selected because this questionnaire has the highest sensitivity rate of all ASQs (Squires et al., 1999). In addition, the 36-month ASQ was chosen because most Taiwanese children enter preschool at 3 years of age. Entrance to preschool is an ideal time to screen children for developmental problems and to provide timely early intervention services as Taiwanese law requires.

Tsai, McClelland, Pratt, & Squires
METHOD

This pilot study used a cross-sectional sample of convenience design to investigate the translated Chinese 36-month ASQ. Specifically, preliminary cultural appropriateness, inter-rater reliability, internal consistency, criterion validity, and content validity were examined. The initial cultural appropriateness of the 36-month ASQ was reviewed by a panel of Taiwanese child development experts and participating parents and preschool teachers of 3-year-old Taiwanese children. Preliminary inter-rater reliability was examined by comparing parents’ and preschool teachers’ 36-month ASQ ratings of the same child. Internal consistency was investigated by examining alpha coefficients from both parents’ and teachers’ ASQ ratings. Initial criterion validity was examined by using a group of children with diagnosed developmental delays to test if the 36-month ASQ scores correctly identified these children. Finally, content validity was examined by a panel of Taiwanese child development experts.

Sampling

Children with no known developmental delays. After discussion with the ASQ author, it was agreed that the 36-month ASQ would be appropriate to give to children aged 34 to 38 months; therefore children aged 34 to 38 months with no known developmental delays were recruited from preschools across Taiwan. A total of 35 preschools were contacted and all agreed to participate. Each preschool was provided with questionnaire packets for parents of enrolled children whose age fell between 34 to 38 months. Parents who agreed to participate returned their completed questionnaire packet to the preschools. One preschool did not distribute the questionnaires; thus, the total number of participating preschools was 34.

A total of 155 pairs of questionnaire packets were distributed to parents and teachers of eligible children; 133 packets were returned by parents; and 155 packets were returned by teachers. Of the 133 packets, data from 32 children were dropped from the study because their ages did not fall between 34 and 38 months. The parent response rate for the final sample of children with no known developmental delays (n = 101) was 82%; the teacher response rate was 100%. These data were used to compare ratings of the same children by parents and teachers and to examine the psychometric properties of the instrument.

Children with known developmental delays. Twenty-three children between 34 and 38 months were identified from a list of children with disabilities who were registered with the Early Intervention Center of the Social Welfare Bureau in Pingtung, Taiwan. Parents of these children were sent a letter and questionnaire packet and asked to complete and return by mail to the researcher. Of the 23 parents contacted, 11 parents returned a completed packet. The resulting response rate for parents of children with known developmental delay was approximately 50%. Data from these children were used to study criterion validity to determine whether the 36-month ASQ could correctly identify children with known delays. Data from these children also were included in analyses of psychometric properties of the instrument.

Study Sample

Data on the 101 children with no known developmental delays and the 11 children (4 girls and 7 boys) with developmental delays were combined into a sample of 112 Taiwanese children ages 34 to 38 months (see Table 1). Data on the 11 children with known delays were collected only from their parents because the majority of children with developmental delays were not in preschool.

Child characteristics. Of 112 participating Taiwanese children, 51 (46%) were females and 61 (55%) males. The majority of the children (n = 85) were from the southern part of Taiwan. Most children in the present study (n = 80, 71%) were of Chinese descent.

Parent characteristics. Mothers’ and fathers’ average age were 32 and 35 years.
old, respectively. The majority of children (66%) lived in dual-parents homes; 20% lived with both parents and grandparents of either side. Average education levels were approximately 14 years for both fathers and mothers (see Table 1). Over 66% (n = 74) of mothers worked outside the home either full-time (63%) or part-time (3%); 13% (n = 14) were unemployed. Over 80% (n = 89) of the fathers were employed full-time; one father was unemployed.

Preschool teacher characteristics. As shown in Table 1, all 55 preschool teachers who participated were female, with ages ranging from 22 to 51 years (average of 32 years). The average number of years of experience in early childhood settings was eight years, and over half (56%) had 14 years or more of education. Eighteen teachers (33%) reported that they held degrees in child care and early childhood education. Teachers reported a range in their classroom’s teacher-child ratio from 1 : 5 to 1 : 26 (M = 1 : 11).

**Table 1**

| Demographic Characteristics of Participant Families and Preschool Teachers |
|-----------------------------|---|---|---|---|
| Child’s age in months       | 36.3 | 34–38 | 1.4 | 112 |
| Mother’s age                | 31.7 | 23–41 | 4.4 | 75a |
| Mother’s years of education | 14.1 | 06–18 | 2.3 | 81a |
| Father’s age                | 35.4 | 25–47 | 4.3 | 74a |
| Father’s years of education | 14.2 | 09–18 | 2.5 | 80a |
| Teacher’s age               | 32.3 | 22–51 | 7.5 | 55  |
| Teacher’s years of education| 13.9 | 12–16 | 1.5 | 55  |
| Teacher’s years of experience in early childhood | 7.9 | 01–27 | 5.6 | 55  |
| Preschool adult/child ratio | 11.0 | 05–26 | 5.5 | 55  |

*Note.* The total number of parent participants (including parents of children with and without known developmental delays) was 112. The total number of teachers was 55. *a* = Different Ns are due to missing data.

In the 36-month ASQ was compared with the original 36-month ASQ for each item and no discrepancies were found except for the term “gross motor” which was back translated to “big movement.”

**ASQ review by an expert panel.** Five Taiwanese child development experts reviewed the initial cultural appropriateness and content validity of the translated 36-month ASQ. Each of the five experts was given an item by item yes/no checklist of the 36-month ASQ along with a copy of the Chinese 36-month ASQ. The checklist asked experts to rate the age and cultural appropriateness of items. Comments were requested if there were disagreements on cultural appropriateness of any of the items.

The experts were identified based on their area of expertise. This five member expert panel included two leading Taiwanese researchers and three health professionals. One researcher held doctoral degrees in counseling psychology and language pathology and worked as an associate professor in the department of special education in a southern university. The other researcher was a professor from a separate university who had expertise in measurement and child psychology. The three health professionals worked in two hospitals and included a pediatrician, occupational therapist, and physical therapist. Both the occupational therapist and physical therapist worked with children in early intervention in the hospitals. All expert

*Tsai, McClelland, Pratt, & Squires* 217
panel members held terminal professional degrees in their fields. These experts were chosen to evaluate the potential cultural and developmental appropriateness of the Chinese 36-month ASQ, based on their training and experience working with Taiwanese children and families.

**Questionnaire packets.** Questionnaire packets were developed for parents and preschool teachers. Packets included an invitational letter, the Chinese 36-month ASQ, a demographic questionnaire, and a satisfaction survey. For the children with no known developmental delays, parents and teachers were asked to rate the same child using the 36-month ASQ. Although parents rated only their child, teachers rated between one and five children. The ASQ satisfaction survey included two yes/no questions asking parents and teachers to rate the appropriateness of the ASQ for the targeted child and to identify any inappropriate items from the translated Chinese questionnaire. The demographic questionnaire asked participants to provide background information including their age, gender, education, occupation, and employment status. In addition, parents were asked to report on the child’s past and current primary care providers. The majority (57%) of parents who completed the ASQ indicated that they were the child’s primary care provider before the child entered preschool at 3 years of age. Questionnaire packets for teachers asked additional questions about the characteristics of the child’s preschool such as the teacher-child ratio.

**RESULTS**

**Cultural Appropriateness**

To obtain initial information on the cultural appropriateness of the translated Chinese 36-month ASQ in Taiwan, data were collected from Taiwanese child development experts, parents, and preschool teachers. Four of the five Taiwanese child development experts verified the initial appropriateness of the translated Chinese 36-month ASQ. One expert, a pediatrician, pointed out that a picture on the fine motor subscale item showed a left hand image holding a pair of scissors. He suggested that a right hand image would be more appropriate because of the traditional Chinese preference for children to use their right hands. For other items, all experts agreed that the Chinese 36-month ASQ was culturally appropriate for Taiwanese children.

Participating parents and preschool teachers also endorsed the cultural appropriateness of ASQ items. Of 49 teachers and 112 parents who answered the question about cultural appropriateness for the targeted child, only one teacher and one parent answered no. The parent’s written comment suggested the rating should be a Likert-type scale instead of yes or no. The teacher’s written comment questioned why there were so many items on a screening test, rather than addressing cultural appropriateness.

**Preliminary Reliability of the Chinese 36-Month ASQ**

**Inter-rater reliability.** Inter-rater reliability of the Chinese 36-month ASQ was evaluated by comparing the teacher’s and parent’s subscale scores for the same child. All subscale correlations were statistically significant \( p < .001 \) and ranged from .45 to .78. Paired sample \( t \) tests revealed no statistically significant differences between the parents’ and teachers’ mean scores on four of the five 36-month ASQ subscales (see Table 2). There was a statistically significant difference in parents and teachers’ mean scores on the personal-social subscale, with parents rating their children higher, \( t(95) = 4.28, p < .001; d = .38 \).

Using 2 SD below the mean of this Taiwanese sample as the cutoff point, 22 children with no known developmental delays were identified by parents and teachers as in need of further evaluation. Fourteen children received these scores from parents and 11 children received these scores from their preschool teachers. Of the 22 children, only 3 (14%) were identified by both parents and teachers as in need of further developmental evaluation.
The preliminary internal consistency of the Chinese 36-month ASQ was examined using Cronbach’s alpha coefficients for total scale scores and the five subscale scores for parent and teacher data. Results indicated that the parent version of the Chinese 36-month ASQ was internally consistent with a .96 standardized Cronbach’s alpha coefficient for total scale scores. Item deletion did not improve alpha coefficients for total scale scores. The communication subscale had the highest internal consistency (α = .91) and the personal-social subscale had the lowest (α = .83) of the five subscales. Alpha coefficients of the five subscales were not significantly improved with the deletion of items. Overall, Cronbach’s alpha coefficients exceeded .80 for the whole scale and for each subscale, suggesting that the parent-completed Chinese 36-month ASQ was internally consistent.

Similar to the findings reported above, the standardized Cronbach’s alpha coefficient for the teacher completed 36-month ASQ scale was .96 for total scale scores, and deletion of items did not improve this standardized alpha coefficient. The highest subscale alpha coefficient was for the communication subscale (α = .88) and the lowest alpha level was for the personal-social subscale (α = .77). Only the problem solving subscale alpha coefficient improved slightly (from .79 to .81) if a subscale item (asking “what is this?” with an incomplete picture of a person) was deleted.

**Initial Validity of the Chinese 36-Month ASQ**

**Criterion validity.** The 36-month ASQ scores of 11 3-year-old Taiwanese children with known developmental problems were examined to assess the ability of the Chinese 36-month ASQ to identify children with known developmental delays. As suggested in the ASQ User’s Guide (Squires et al., 1999), 2 SD below the mean was the principle cutoff point for identification of possible developmental delays. Using 2 SD below the mean of this Taiwanese sample as the cutoff point, all children with known delays were correctly identified. All but one child had at least three subscale mean scores below the cutoff points (see Table 3). Most of these children demonstrated significant develop-
mental delays in many areas of functioning as rated by their parents.

**Content validity.** Preliminary data on content validity of the Chinese 36-month ASQ was assessed using data from the five Taiwanese child developmental experts. An item-by-item yes/no checklist of the 36-month ASQ with a copy of the Chinese 36-month ASQ was given to each of the experts. The checklist asked experts to rate the developmental and cultural appropriateness of items. Comments were requested if there were disagreements on any of the items. Two themes emerged from the experts’ comments. First, some subscale names did not match the content of subscale test items, particularly in the problem-solving and personal-social subscales. Specifically, the problem-solving domain included two questions related to auditory sequential memory (e.g., repeating two numbers in the same order), which did not fit with the other subscale items. In the personal-social domain, the experts noted that two subscale items (feed self with spoon; put on jacket or coat) were assessing adaptive motor skills rather than personal-social skills because no social interaction was involved, and another item (pushing cart backwards or around an object) assessed a motor, not a personal-social skill. In addition, the experts commented that one personal-social subscale item (knows gender of self) was more of a cognitive skill than a personal-social skill.

Second, the experts felt that some repetitive items were included within a single subscale domain. Specifically, jumping was assessed by two questions in the gross motor domain: one question asked about jumping with both feet leaving the floor at the same time, and the second question assessed a more advanced skill, jumping forward at least six inches. The experts questioned the redundancy of these items in a screening test. Results from the five experts suggested possible problems regarding the content appropriateness of the gross motor, problem-solving, and personal-social subscales of the ASQ itself, rather than of the Chinese version. Despite these issues, there was overall agreement about the content validity of the 36-month ASQ for Taiwanese children.

**DISCUSSION**

The present study was an initial investigation of the cultural appropriateness, and reliability and validity of scores using a Chinese

### Table 3

*Cutoff Scores and ASQ Scores of 11 Taiwanese Children with Known Developmental Delays*

<table>
<thead>
<tr>
<th>Subscale</th>
<th>Communication</th>
<th>Gross Motor</th>
<th>Fine Motor</th>
<th>Problem solving</th>
<th>Personal-social</th>
</tr>
</thead>
<tbody>
<tr>
<td>Child ID 1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>35</td>
<td>5</td>
<td>10</td>
<td>25</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>25</td>
<td>10</td>
<td>30</td>
<td>20</td>
</tr>
<tr>
<td>4</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>40</td>
<td>30</td>
</tr>
<tr>
<td>5</td>
<td>40</td>
<td>50</td>
<td>20</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>6</td>
<td>30</td>
<td>40</td>
<td>15</td>
<td>15</td>
<td>50</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>0</td>
<td>25</td>
<td>0</td>
<td>20</td>
<td>5</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>50</td>
<td>15</td>
<td>20</td>
<td>40</td>
</tr>
<tr>
<td>11</td>
<td>0</td>
<td>30</td>
<td>0</td>
<td>30</td>
<td>10</td>
</tr>
</tbody>
</table>

*Note. Means and standard deviations for the overall Taiwan sample were: communication, 55(8); gross motor, 54(10); fine motor, 48(14); problem solving, 51(11); personal-social, 54(7). One or more scores at or below the subscale cutoff score indicates the need for further developmental evaluation (Squires et al., 1999).*
version of the 36-month ASQ in Taiwan. Initial results from child development experts, parents, and preschool teachers indicated that the Chinese 36-month ASQ showed cultural appropriateness and adequate reliability and validity of scores when used with a sample of 3-year-old Taiwanese children. This study, although preliminary, extends the cross-cultural research on screening tools and assessment for young children.

Cultural Appropriateness of the Chinese 36-Month ASQ

Although some researchers have argued that culture can influence parenting beliefs and practices leading to differences in children’s development (Cheng & Tsai, 2002; Solomons & Solomons, 1975; Super, 1976), results of the present study support previous research findings that behaviors targeted by the ASQ are not culturally dependent (Janson & Squires, 2004; Dionne, Squires, Leclerc, & Peloquin, 2006). The panel of Taiwanese experts and preschool teachers and parents of 3-year-old children supported the preliminary cultural appropriateness of the translated Chinese 36-month ASQ. The only concern regarding cultural appropriateness raised by a pediatrician serving as an expert reviewer was the use of a left-hand image holding a pair of scissors. Although this comment reflects a traditional Chinese right-hand preference, in current Chinese society, few parents openly object to left-handedness. These results need further validation but suggest that the ASQ could have important practical implications as a low-cost developmental screening tool in Taiwan.

Preliminary Reliability of the Chinese 36-Month ASQ

The inter-rater reliability and internal consistency analyses in this pilot study provided preliminary support for the psychometric qualities of the Chinese 36-month ASQ. Inter-rater reliability of scores was demonstrated by comparing parents’ and teachers’ ASQ ratings. Taiwanese parents and preschool teachers evaluated the same child similarly with the exception of ratings on the children’s personal-social subscale scores. One possible explanation for this difference is contextual differences between home and school that resulted in different expectations and observations regarding young children’s social behaviors. For example, parents and teachers might have different expectations about how well a child waits for a turn. This behavior might be easier to achieve at home than at school. More than 70% of the children in the sample lived in a household with only one or two children, whereas in the preschool settings, the average adult to child ratio is 1 : 11. A child might have to wait longer for his or her turn; thus, teachers might be more likely to see turn-taking behavior as more difficult for a child.

Although the differences in parent and teacher subscale ratings of individual children were minimal, parents and teachers did not identify the same children as needing referral for further assessment of developmental delays. Because the 22 children who were identified by their parents or teachers did not receive further evaluation, it is unclear if this disagreement was due to problems with the scale (e.g., over referrals or false positives, under referrals or false negatives), or merely reflected a difference in expectations between parents and teachers.

Further examination of teacher and parent scores for the 22 children indicated that although parent and teacher scores fell on different sides of the cut-off score on a particular subscale, differences between total parent and teacher ratings were minimal. For example, in 12 of the 22 children (54%), parent and teacher overall ratings were within 10 points of each other, and in 18 of the 22 children (82%), overall ratings were within 15 points of each other. In general, teachers’ subscale ratings were lower than parents’ ratings for these children. Thus, although there was low agreement between parents and teachers on specific subscales, overall scores for parents and teachers were not that different for the majority of children who were identified as needing further evaluation.

The ASQ was designed for parents or primary care providers to complete. Previous
research has found low to moderate agreement in teacher and parent ratings of children’s behavior (Treuting & Elliot, 1997). When rating behavior, teachers often compare children to each other, whereas parents are more likely to rate their child independently. It is possible that the different expectations and observations for children’s behavior are responsible for this disagreement. Some scholars also suggest that parents’ and teachers’ perspectives, although different, are equally valuable to obtain a complete picture of the child (Suen, Logan, Neisworth, & Bagnato, 1995; Bailey, Siemonsson, Buysse, Smith, 1993). This research supports the use of generalizability analyses as an effective method to examine inter-rater reliability of the ASQ.

Results from parent and teacher ratings demonstrated internal consistency ($\alpha = .96$) in this pilot study, with only one teacher-rated subscale alpha level (personal-social) falling below the desired .80. This finding is similar to that reported for a U.S. sample (Squires et al., 1999) in which the personal-social subscale of the 36-month ASQ also had the lowest standardized alpha coefficient ($\alpha = .55$). Concerns about the personal-social subscale also were raised by the Taiwanese child development experts, who argued that the personal-social subscale seemed to be testing two different areas of child development: personal adaptive behaviors (such as feed self with spoon) and social interaction (such as turn-taking). If different domains are assessed within a subscale, it is not surprising that alpha coefficients indicate relatively lower levels of internal consistency.

A recent Canadian study of the French ASQ (Dionne, Squires, Leclerc, & Peloquin, 2006) also found that for three of four ASQs (42, 48, 54, and 60 month), the personal-social subscales had lower alpha coefficients compared to other ASQ subscales. The Chinese 36-month ASQ personal-social subscale’s somewhat low internal consistency might result from items that sample different developmental domains included in the subscale. High alpha coefficients for the subscales and total scale, however, are important initial findings and suggest that the ASQ might be useful in Taiwan.

**Validity of the Chinese 36-Month ASQ**

In general, results supported the initial content and criterion validity of scores of the Chinese 36-month ASQ in a small sample of Taiwanese children. The Chinese 36-month ASQ correctly identified all children with developmental delays in this pilot study. These results suggest that the Chinese 36-month ASQ demonstrated initial criterion validity in its ability to identify children with known developmental delays.

Findings regarding the criterion validity of the Chinese 36-month ASQ must be interpreted with caution. This small sample might not be representative of the larger population of Taiwanese children with developmental delays. In addition, although criterion validity can be achieved by comparing children’s results on the ASQ with an existing condition (a diagnosed disability), parents’ views about development might be influenced by their child’s diagnosed disability. Parents of children who already have a disability might be responding to the ASQ using this diagnosis as a reference point. Further, because of the low identification rate of children with developmental delays in Taiwan, children who are registered with the social welfare services are usually demonstrate severe delays. In the present study, 10 of 11 children with diagnosed developmental delays showed weaknesses on multiple developmental domains. It is unclear if the Chinese 36-month ASQ will identify less severe developmental variations among Taiwanese children. Future research should include discriminate function analyses to examine the ability of the Chinese 36-month ASQ to distinguish Taiwanese children with and without developmental delays. For example, future research could select children from inclusion settings to further examine criterion validity of the 36-month ASQ because children who are able to participate in inclusive preschools might demonstrate less severe developmental delays. In addition, children at high risk for developmental delays such as those with low-

222

*JEI, 2006, 28:3*
birth weight or of low socioeconomic status might be another population to include in further examination of this issue.

**Limitations and Future Research**

Despite the importance of this pilot work, the study had a number of limitations. First, the non-randomized sampling method resulted in a majority of the sample being recruited from southern Taiwan and thus, results might not represent the perspectives of parents and teachers throughout Taiwan. The sample included some children from other areas of Taiwan and preliminary analyses showed no statistically significant differences in children’s ASQ scores for children living in southern Taiwan compared to children living in other areas of Taiwan after controlling for mothers’ education level. Although children in southern Taiwan might be different from children in other parts of the country, Taiwan is a fairly homogeneous society. Nevertheless, future research should address these concerns by including a broader sample of participants from more diverse areas of Taiwan and a broader range of educational and socioeconomic backgrounds.

Second, although a panel of experts viewed the Chinese 36-month ASQ as culturally appropriate, it is possible that the experts’ training and background reflect western values and norms. In addition, all five experts worked in southern Taiwan, and their perspectives might not represent the norms and values in other sections of Taiwan. Although the experts in this study were all trained in Taiwan and had extensive experience working with children and families in Taiwan, future research should include experts from more diverse areas of Taiwan.

Third, this study did not examine the sensitivity and specificity of the Chinese 36-month ASQ scores, which are important indices for screening instruments. Results revealed some discrepancy in the identification of children needing further evaluation between parents’ and teachers’ ASQ ratings. It is not clear if this discrepancy is a result of parents and teachers having different expectations or if the Chinese 36-month ASQ demonstrates sensitivity and specificity problems. Future research should include developmental diagnostic assessments to examine these issues in more detail.

Fourth, although the 36-month ASQ correctly identified the 11 Taiwanese children with developmental delays, only a small number of parents of children with developmental problems responded to the study and results should be seen as preliminary. It is unclear why so few parents participated, although there is often a significant bias against children who have developmental problems in Taiwan. Parents might not have wanted to call attention to their child’s diagnosis. Future research should include confirmatory factor analyses to further explore the content and construct validity of the scale.

Finally, this study did not investigate contextual factors such as family structure, child care arrangement, and parents’ socioeconomic status, education, or ethnicity. These factors influence children’s development and are important to examine.

**Implications for Practice**

This study contributed to the growing literature on the cross cultural relevance of the ASQ by examining the initial cultural appropriateness, and reliability and validity of scores for the Chinese translation of the 36-month ASQ with a sample of Taiwanese children. Based on research findings that simply translating a measure into another language does not equal a psychometrically sound version of that measure, careful validation is necessary for an assessment instrument to be accurate for children from different cultural backgrounds (Cheng & Tsai, 2002; Dionne, Squires, Leclerc, & Peloquin, 2006; Janson & Squires, 2004; Vazir et al., 1994). This pilot study supported the preliminary reliability and validity of scores on the Chinese 36-month ASQ and

*Tsai, McClelland, Pratt, & Squires*
suggests that further comprehensive examination and refinement are needed. In particular, content validity issues raised by the Taiwanese child development experts need further investigation, as do issues related to the ability of the Chinese 36-month ASQ to identify children with mild developmental delays.

The present study has important implications for early intervention practice in Taiwan. Using the World Health Organization and Taiwan Children’s Bureau (2004) figures, it is estimated that the majority (80%) of Taiwanese children under 6 years of age who are developmentally compromized have not been identified and registered for early intervention services. Clearly, the current child-find system in Taiwan does not adequately identify children with developmental delays. There is a need for an effective and culturally appropriate child developmental screening instrument. The results of the Chinese 36-month ASQ, although preliminary, demonstrate the initial psychometric properties of this scale. Future investigation is warranted to validate the scale further and establish normative data for Taiwanese children. This information can provide practitioners with a culturally appropriate, developmental screening instrument for early identification of children needing early intervention services.

REFERENCES


224

JEI, 2006, 28:3


Correspondence concerning this article should be addressed to Huei-Ling Agnes Tsai, Department of Child Care, National Pingtung University of Science and Technology, No. 1, Shuey Fu Road, Nei Pu, Pingtung, Taiwan, Republic of China. E-mail: agnes@mail.npust.edu.tw