

Research Article

Cross Cultural Gender Differences in Social-emotional Competence of Young Children: Comparisons with Brazil, China, South Korea, and the United States

Chieh-Yu Chen

Early Intervention Program, University of Oregon, USA

Jane Squires

Early Intervention Program, University of Oregon, USA

Kay H. Heo

Department of Early Childhood Education, Chongshin University, South Korea

Xiaoyan Bian

Shanghai Maternal and Child Health Center, China

Ching-I Chen

School of Lifespan Development and Educational Sciences, Kent State University, USA

Alberto Filgueiras

Departamento de Fundamentos em Psicologia, Rio de Janeiro State University, Brazil

Huichao Xie

Early Intervention Program, University of Oregon, USA

Kimberly Murphy

Early Intervention Program, University of Oregon, USA

Jill Dolata

Institute on Development and Disability, Oregon Health and Science University, USA

Jesus Landeira-Fernandez

Department of Psychology, Pontifical Catholic University of Rio de Janeiro, Brazil

ABSTRACT

Background: The current study investigated differences in gender-based expectations across cultures by examining whether items in Ages & Stages Questionnaires: Social-Emotional (ASQ:SE), a social-emotional screening instrument, functioned differently for boys and girls 3 to 66 months old. A total of 74,776 ASQ:SE questionnaires were collected in Brazil, South Korea, China and the United States (U.S.).

Methods: The analyses were conducted using differential item functioning (DIF) by country for each ASQ:SE age interval (i.e., 6, 12, 18, 24, 30, 36, 48, 60 months).

Results: Results showed that eight items with gender DIF were present in two or three countries simultaneously, and 19 items with gender DIF were detected in a single country.

Conclusions: With an equal level of social-emotional competence, girls had higher scores than boys, indicating more problem behaviors on items reflecting internalizing behaviors; and boys had higher scores on items measuring externalizing problem behaviors. Additionally, in all four countries, differences in item functioning associated with gender increased as children grew older. Interpretations of ASQ:SE results should consider that families' cultural beliefs may have differing effects according to the child's gender.

MeSH Headings/Keywords: Social-emotional competence, Gender difference, Cross cultural, Ages & Stages Questionnaires: Social-Emotional (ASQ:SE), Differential item functioning (DIF)

Introduction

The critical nature of social-emotional competence in young children has been emphasized in recent years. Prevalence studies suggest that 13% to 25% of young children have at least mild emotional and behavioral problems [1-3]. Growing evidence indicates that social-emotional problems in young

children can have long-term effects [4-6] and problems that go untreated can become serious and costly to treat as children age [7,8]. In addition, social-emotional competence is crucial for school readiness for young children and serves as an essential foundation for learning skills that are important to academic performance in later years [9-11]. Given the importance of

social-emotional competence, using technically sound screening instruments for the early identification of social-emotional problems is critical, as well as providing timely intervention once problems are detected [7,12,13]. Screening tests with solid technical adequacy help identify children who may be at risk for social-emotional delays and assist in early identification and referral for needed services. Thus, investigating the properties of screening tools (e.g., psychometric properties, item functioning, and bias for specific groups) is necessary to assure accuracy, especially when tests are translated and used across cultures.

The Ages & Stages Questionnaires: Social-Emotional (ASQ:SE) [14], a tool with adequate psychometric properties for screening social-emotional problems [12,13,15,16], has been widely used in the Head Start and Early Head Start programs across the United States [17,18] as well as internationally. The ASQ:SE has been translated into several language versions and its psychometric properties and cultural appropriateness have been studied, such as in South Korea [19,20] and China [21]. Other social-emotional assessments such as the Infant-Toddler Social and Emotional Assessment [22], Social Competence and Behavior Evaluation [23], and Devereux Early Childhood Assessment [24] have been studied and found to have solid psychometric properties. However, the ASQ:SE is the only social-emotional assessment designated for the entire birth through age five age span. In addition, its world-wide usage provides an opportunity to examine potential cross-cultural differences in young children from birth through age five.

Human development has been thought of as an interactive process, with biological and cultural factors interacting [25,26]. Culture strongly affects human attitudes, beliefs, values and behaviors [8], including parents' perceptions of appropriate social-emotional performance and gender roles [27]. Previous studies investigating cross-cultural gender differences have mainly focused on school-age children/adolescents rather than infants and young children. In one preschool study, Chen and colleagues [28] found few gender differences on behavioral inhibition in Chinese and Canadian toddlers, suggesting that socialization in early childhood might play an essential role in gender differences in social-emotional disabilities that surface in later childhood. Furthermore, international studies addressing gender differences in young children's social-emotional competence have focused most often on a single cultural group or region [29-34], rather than comparisons across cultures.

The aim of this study was to examine social-emotional assessment items with potential gender differences across cultures. Specifically, item functioning analyses on a social-emotional screening assessment, the ASQ:SE, were used to investigate gender differences in social-emotional competence across four diverse countries – United States, Brazil, China, and Korea. If items with gender difference were identified across countries, inherent cultural similarity and/or underlying bio-cultural aspects might be an explanation. If items with gender differences were identified in one country only, they might be attributed to cultural practices and attitudes in that country only.

In each country, parents or caregivers completed the ASQ:SE on their young children (3 to 66 months). Differential Item Functioning (DIF) was then employed to examine gender differences at the item level to determine if and how

ASQ:SE items functioned differently between two groups (i.e., boys and girls). Two research questions were asked: (a) Are there differences in item functioning based on gender across countries?; (b) Do ASQ:SE items function differently for boys and girls in their respective countries?

Methods

Participants: Participants were recruited in the United States, Brazil, China, and Korea. A total of 74,776 infants and young children (39,605 boys and 35,171 girls) in the age range of 3 to 66 months were included (Table 1). The demographic information is presented in (Table 2).

United States: A total of 15,110 young children (8,748 boys and 6,362 girls) in the U.S. sample were collected between 2007 and 2014 as part of a national normative study of the ASQ:SE second edition [35]. Recruitment of families was strategically conducted so that the sample roughly reflected the U.S. 2010 census data on variables including race/ethnicity, mother's education level, and family income. Recruitment materials included notices in online publications and parenting newsletters. Depending on how parents were recruited, they completed a paper-pencil version or completed an on-line version of the ASQ:SE. The distribution of ethnicity was White (62.9%); African American (6.7%); Hispanic (6.3%); mixed race/ethnicity (5.7%); Asian (3.7%); Native American (0.7%); Pacific Islander (0.2%); Hawaiian (0.1%); other race/ethnicity (0.4%); don't know(0.3%); and missing (13.1%).

South Korea: A sample of 2,568 young children (1,359 boys and 1,209 girls) was recruited in Korea between 2008 and 2009 for a national normative study of the Korean ASQ:SE. Recruitment included sending flyers to childcare centers, preschools, pediatric offices, children's play centers, and posting on web pages. Children were recruited nationally in a variety of early childhood settings and were stratified according to the Korean population census related to gender and family income. In terms of parental education level, the sample included about 15% more than the national average of higher education graduates. Detailed information on recruitment and research process was provided in the Korea ASQ:SE psychometric study [20].

Brazil: A total of 54,570 children (28,228 boys and 26,342 girls) were recruited in Brazil in 2011 as part of an evaluation of child development in public childcare centers, kindergarten and preschools in Rio de Janeiro and surrounding areas. The cross-cultural adaptation of ASQ:SE followed the same procedures as the Ages & Stages Questionnaires - Brazil (ASQ-BR) [36]. Paper-pencil questionnaires were delivered to all public child childcare centers and preschools in Rio de Janeiro, and teachers were asked to answer a Brazilian-adapted version of the ASQ:SE. If there were items that the teachers were unable to observe in the childcare centers, they interviewed parents. Due to data insufficiency, children at the 6 month interval (3-9 months old) were not included in the analysis; only results for children between 9 and 66 months were analyzed. Demographic information regarding family income and mother's education was not collected in the project. However, general information collected on this Rio de Janeiro population includes that personal income levels can be categorized in four levels: extreme

poverty (8.2%, < R\$120), low income (20.7%, R\$120-R\$300), mid income (40.9%, R\$300-R\$1,200), high income (30.2%, >R\$1,200) (37). The majority of the participating families were from first three categories; few came from high income families. The ethnicity distribution of Rio de Janeiro was Caucasian (51.2%), Black (11.5%), and Brown (36.5%) [37]. The educational level of mothers was 11.9% no education (illiteracy), 44.1% elementary school, 30.7% high school and 13.1% higher education [38].

China: A total of 2,528 children (1,270 boys and 1,258 girls) were collected in China between 2010 and 2011 as part of a national normative study of the ASQ:SE. Parents completed the ASQ:SE in pediatric clinics with assistance of health professionals. Recruitment of families closely reflected China's 2010 census data on variables including regions and children gender, and roughly on race and family income regarding the urban/rural area. The distribution of ethnicity was Han (94.7%) and other (5.3%). Other included more than ten different races.

Measure

The ASQ:SE, the primary outcome measure, is a screening instrument for detecting potential problems in social-emotional competence during the first 3 to 66 months of life. It includes a series of 8 questionnaires for different age intervals (i.e., 6, 12, 18, 24, 30, 36, 48, and 60 months) with the total scored items in each interval ranging from 18 – 32. An example of ASQ:SE items is shown in (Figure 1). ASQ:SE items ask about young children's social-emotional competence or behavioral problems; for example, *Does your child look at you when you talk to him?* Each item has three response options that allow parents or primary caregivers to observe their children and indicate the frequency of their children's social-emotional skills or problems (i.e., "Often or Always," "Sometimes," or "Rarely or Never"). Answers received numeric values reflecting competence (0 points) or problem behaviors (5 or 10 points). A fourth response option allowed parents or primary caregivers to indicate if the behavior is a concern; however, the "concern" option (with an associated value of 5 points) was not included in this analysis because it is not part of the Likert scoring scale, but a separate response option. Thus the separate "concern" option operated as an independent response option and was not appropriate for the categorical scale used in the DIF analysis. For this study, children in each country were assessed with an ASQ:SE translated and adapted to their native language (i.e., U.S. – English; Brazil – Portuguese; China – Chinese; and Korea – Korean)

Analysis

Unidimensionality: One of the assumptions of DIF analysis is to evaluate the unidimensionality of the test, to determine whether items target the same latent trait. If the latent trait consists of multiple dimensions, the DIF examination has to be conducted for each dimension. Unidimensionality was tested by two processes of factor analyses. Data from each ASQ:SE interval of the U.S. sample was randomly divided into two equal samples--a derivation sample (i.e., for exploratory factor analysis; EFA) and replication sample (i.e., for confirmatory factor analysis; CFA). EFA was first employed to evaluate

Figure 1: ASQ:SE item sample. From "Ages and Stages Questionnaires: Social-Emotional, 24 months interval" by J. Squires, D. Bricker, and E. Twombly, 2002, *The Ages & Stages Questionnaires: Social-Emotional*, p. 3. Copyright 2002 by Paul H. Brookes.

the unidimensionality of each of the eight age intervals. The maximum likelihood method was used for factor extraction, and the number of factor extractions was evaluated by the minimum average partial (MAP) test [39] and Hull method [40]. The weighted oblimin rotation [41] was performed for factor rotation. The CFA was then conducted for each interval based on the result of EFA. The indices were reported to evaluate the model fit, including: Chi-squared (χ^2) values, degrees of freedom (df), the goodness-of-fit index (GFI), the root mean square error of approximation (RMSEA), and the root mean square residual (RMR). Due to the large sample size, χ^2 was expected to be high and the p value significant. Furthermore, to be considered as having a good model fit, the criteria of the indicators were expected: GFI > 0.9, RMR < .05, RMASE < 0.8. EFA was conducted using FACTOR [42] and CFA was operated using AMOS 18.0 [43].

Differential item functioning (DIF): The current study examined whether ASQ:SE items functioned differently for girls and boys within and across participating countries. DIF measures whether there are differences between two groups with the same level of latent trait (e.g., social-emotional competence), by modeling the likelihood that items receive scores indicating the target competence [44]. The ASQ:SE sample of each country was divided into different classification groups (i.e., by gender) and then stratified as slices by ability measure (i.e., social emotional competence). "For each slice, a cross-tabulation is constructed for each pair of person classifications against each scored response level. An odds-ratio is computed from the cross-tab [45]". This study provided information on whether parents/primary caregivers responded to items differently based on their child's gender. The items with DIF were those that functioned differently for girls and boys.

Data were analyzed by using the Mantel [46] approach to detect DIF; the Educational Testing Service (ETS) standard for evaluation of DIF levels was employed: level C = moderate to large [$|DIF| \geq 0.64$ logits, $\text{prob}(|DIF| \leq 0.43 \text{ logits}) \leq .05$], level B = slight to moderate [$|DIF| \geq 0.43$ logits, $\text{prob}(|DIF| = 0 \text{ logits}) \leq .05$], and level A = negligible (e.g., other than level B and C) [47]. The analyses were performed using Winstep version 3.91.0 [48].

Results

EFA and CFA were conducted to examine the unidimensionality of the ASQ:SE. The purpose of confirming unidimensionality is to assure ASQ:SE items are measuring the same latent trait which the DIF examination assumed.

EFA results using Velicer's MAP test suggested that one factor should be retained for the 6m, 12m, 30m intervals, and the Hull method indicated a single factor structure for 18m, 24m, 30m, 36m, 48m and 60m intervals. Therefore, a one-factor model was examined by CFA for each interval. CFA results presented acceptable models after adjusting covariance. The original and the adjusted models are presented in (Table 3). Based on the criteria, most intervals fit the one-factor model structure, though the GFI values in 30m, 36m, 48m, 60m were close to .90. Based on the one-factor structure model testing result, DIF analysis was conducted for all items in an interval without separation.

A total of 74,776 dyads (caregiver/parent and child) were included in the DIF analyses as summarized in (Table 1). A DIF analysis was conducted by age interval to detect the gender differences for each country. Eight items were detected with gender DIF across two or more countries (Table 4). For individual country results, five items with gender DIF were detected in the U.S. sample; six items were found in the Korea sample; seven items were found in the China sample; and one item was found in the Brazil sample (Table 5).

Discussion

When examining gender DIF items detected in individual country samples, a trend was found—compared to older age intervals, there were fewer items with DIF in the younger ages. For example, in the older age intervals of the U.S. sample (30 months to 60 months), there were 13 DIF items, whereas in the younger age intervals (6 months to 24 months) there were none. The increasing number of DIF items with age might be because children's social-emotional behaviors become more complex and multifaceted as they grow [49], resulting in greater variability in behaviors demonstrated by older children. Another possible explanation is a change in parental expectations for behaviors related to children's gender and age. The finding of fewer items with gender differences in infant and toddler items across cultures is consistent with previous research [28,50]. In contrast, more differences were found at preschooler age [51-53].

With equal levels of social-emotional competence, girls were more likely to be reported as exhibiting internalizing problematic behaviors than boys, including on items asking about clinging (“Does your child cling to you more than you expect?”) and hours spent sleeping (“Does your child sleep at least 8 hours in a 24-hour period?”). On the other hand, boys tended to be rated by parents as having a higher probability of potential problems with social interactions and externalizing behaviors (i.e., “Does your child try to hurt other children, adults, or animals?”). Our findings are similar to previous studies in Dutch and China that found boys overall had more behavioral problems and peer difficulties than girls [52,53].

Items with gender DIF across countries: Eight items showed DIF across multiple countries (Table 4), possibly due to differing cultural perspectives. For example, results on the item, “Does your child cling to you more than you expect?” indicated that girls in the U.S., Korea and Brazil were rated as more likely to cling to their parents than boys with equal levels of social-emotional competence. This finding might suggest that girls were more emotionally dependent or had more need for contact

Table 1: Descriptive Statistics for ASQ:SE Intervals Scores by Country (N = 74,776).

ASQ:SE Interval	Boys			Girls		
	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>
Brazil						
6m	--	--	--	--	--	--
12m	184	5.40	4.08	152	5.47	4.61
18m	1383	9.20	4.52	1228	8.63	4.23
24m	2030	8.58	4.25	1746	8.04	4.37
30m	2647	10.11	5.49	2417	9.08	4.73
36m	4198	10.28	5.87	3781	8.93	4.99
48m	6495	7.48	6.62	5978	5.82	5.51
60m	11291	8.96	7.33	11040	6.80	6.09
China						
6m	164	3.31	2.86	158	3.63	3.10
12m	167	4.34	3.42	163	4.76	3.18
18m	164	4.81	3.41	156	5.30	4.23
24m	156	5.79	3.97	161	4.99	3.76
30m	161	7.31	5.30	152	7.85	5.43
36m	151	7.04	4.73	156	7.64	5.08
48m	157	8.31	5.43	157	7.66	4.85
60m	150	7.03	4.80	155	6.81	4.64
Korea						
6m	84	3.67	2.51	58	3.90	2.85
12m	179	4.91	3.26	172	4.95	3.13
18m	136	5.57	3.81	129	5.14	3.63
24m	155	5.18	3.73	138	5.18	3.73
30m	104	6.73	4.11	102	6.25	4.73
36m	151	8.30	5.33	140	7.23	4.20
48m	234	7.30	5.24	218	6.32	4.16
60m	316	7.34	5.33	252	6.35	4.46
USA						
6m	1010	4.20	3.28	912	4.03	3.21
12m	1214	5.40	4.20	932	4.59	3.68
18m	1232	7.27	5.70	814	6.47	5.72
24m	1039	8.05	6.38	655	6.52	5.79
30m	826	11.12	7.99	585	8.93	7.11
36m	1206	12.74	8.84	864	9.41	7.42
48m	1388	11.77	8.78	1020	8.48	7.58
60m	833	11.40	8.83	580	9.55	8.47
Total	39,605	--	--	35,171	--	--

Note. The item scores were converted from 0, 5, 10 points into 0, 1, 2 points for the later DIF analysis. *M* and *SD* were calculated from the total scores of the intervals.

Table 2: Demographic Characteristics for Families Using the ASQ:SE by Country.

United States	N	%	Korea	N	%	China	N	%
Income (Family; USD)			Income (Family; KRW)			Income - City (Person; CNY)		
0-12,000	1,131	7.5%	0-1,000,000	12	0.5%	≤6,000	231	9.1%
12,001-24,000	2,382	15.8%	1,000,000-1,500,000	67	2.6%	6,000-12,000	389	15.4%
24,001-40,000	1,942	12.9%	1,500,000-2,000,000	139	5.4%	12,000-20,000	401	15.9%
Over 40,000	8,086	53.5%	2,000,000-3,000,000	589	22.9%	≥20,000	249	9.8%
Don't know	1,083	7.2%	3,000,000-4,000,000	476	18.5%	Missing	6	0.2%
Missing	486	3.2%	Over 4000000	661	25.7%	Income - Rural (Person)		
			Missing	624	24.3%	≤2,000	250	9.9%
Mother's education			Mother's education			2,000-5,000	567	22.4%
College degree or above	7,424	49.1%	College degree or above	1,754	68.3%	5,000-8,000	203	8.0%
Associate degree	1,839	12.2%	Associate degree	158	6.2%	≥8,000	231	9.1%
High school graduate	2,941	19.5%	High school graduate	553	21.5%	Missing	1	0.0%
Less than high school	496	3.3%	Less than high school	5	0.2%	Mother's education		
Don't know	408	2.7%	Missing	98	3.8%	College degree or above	177	7.0%
Missing	2,002	13.2%				Associate degree	321	12.7%
						Senior high School	482	19.1%
						Junior high School	1,202	47.5%
						Elementary School	310	12.3%
						No education	32	1.3%
						Missing	4	0.2%
Total	15,110	100%	Total	2,568	100%	Total	2,528	100%

Note. Brazilian demographic information was explained in the participants section.

Table 3: Indicators of the Confirmatory Factor Analysis for the Original and the Adjusted Model of the ASQ:SE in the U.S. Sampl.

Interval	Model	χ^2	df	GFI	RMR	RMSEA
6m	Original	778.77	135	0.90	0.01	0.07
	Adjusted	530.11	132	0.93	0.01	0.06
12m	Original	942.77	189	0.90	0.02	0.06
	Adjusted	777.93	187	0.92	0.02	0.06
18m	Original	1630.13	275	0.85	0.02	0.07
	Adjusted	1061.67	269	0.90	0.02	0.06
24m	Original	1061.44	275	0.88	0.02	0.06
	Adjusted	872.59	271	0.91	0.02	0.05
30m	Original	1299.34	350	0.85	0.02	0.07
	Adjusted	1019.76	344	0.89	0.02	0.06
36m	Original	2545.42	405	0.82	0.02	0.08
	Adjusted	1724.20	398	0.88	0.02	0.06
48m	Original	2942.59	464	0.82	0.02	0.07
	Adjusted	2112.21	458	0.87	0.02	0.06
60m	Original	1845.72	464	0.81	0.02	0.07
	Adjusted	1457.14	458	0.85	0.02	0.06

Note. χ^2 are all statistically significant ($p < .001$.)

than boys. However, this item did not function differently for girls and boys in the Chinese sample. Due to the one-child policy in China, parents might expect clinging or they may be very protective of or attentive to their only child [54]. Additionally, in Chinese cultures, parents may consider childrearing to include significant involvement in the care and discipline of their children [55] so that “clinging” might be expected. The findings related to the “clinging” item also help explain why boys were more likely to be rated with higher scores on the item, “Does your baby look for you when a stranger approaches?,” since boys may be more emotionally independent than girls. Therefore they may reference their parents less frequently when they start to develop stranger anxiety around 12 to 18 months (this item expects children should look for their parents when a stranger approaches).

Around four or five years of age, young children begin to realize differences between boys and girls, and they start to explore their own bodies and use language to test social rules regarding gender and sexuality [56]. Parents appeared to be more sensitive to any interest girls had in sexual language or activity, as DIF appeared in the U.S., Korea, and China (i.e., “Does your child show an unusual interest or knowledge of sexual language and activity?”), with girls rated with more problems, except caregivers in Brazil. These findings might be due to observation opportunities, with caregivers in daycare center having limited chances to observe interest in sexual language, etc. compared with parents at home. Another item detected as having DIF was “Does your child try to hurt other children, adults, or animals?”. This item suggested boys were

Table 4: ASQ:SE Items with Gender DIF Detected in Two or Three Countries.

Items	Group	6m	12m	18m	24m	30m	36m	48m	60m
Cling more than expect	USA					0.44**	0.48**	--	--
	Korea					0.53*	--	--	0.46**
	Brazil					0.43***	0.54***	0.63***	0.73***
	China					--	--	--	--
Unusual interest of sexual language and activity	USA						--	--	0.60*
	Korea						--	--	0.52*
	Brazil						--	--	--
	China						--	--	0.79*
Try to hurt others	USA		--	--	--	--	--	--	--
	Korea		-0.51*	--	--	--	-0.80*	-0.48*	--
	Brazil		--	--	--	--	--	--	--
	China		--	--	--	--	--	--	-0.79*
Like to hear stories and sing songs	USA			--	--	-0.53**			
	Korea			-1.41**	--	-2.07**			
	Brazil			--	--	--			
	China			--	--	--			
Hurt self	USA			--	--	--	--	--	--
	Korea			--	--	-1.38*	--	--	--
	Brazil			--	--	--	--	--	--
	China			0.52*	--	--	--	--	--
Like to be hugged	USA			--	--	--	--	--	--
	Korea			--	--	--	--	--	--
	Brazil			--	--	--	--	-0.45***	--
	China			-0.55*	--	-0.59*	--	--	--
Too friendly to strangers	USA				--	--	0.49*	0.49**	0.57***
	Korea				--	--	--	--	--
	Brazil				--	--	--	0.44***	--
	China				--	--	--	--	--
Look you when a stranger approaches	USA		--	--					
	Korea		-0.92***	-0.61*					
	Brazil		-0.54**	--					
	China		--	--					

Note1. Values are logits.

Note2. DIF level: -- = negligible, normal font = slight to moderate ($|DIF| \geq 0.43$ logits), bold font = moderate to large ($|DIF| \geq 0.64$ logits); blank = the item is not tested in the interval; positive values = girls are the preference group; negative values = boys are the preference group.

* $p < .05$. ** $p < .01$. *** $p < .001$.

more likely to be rated as demonstrating aggressive behaviors than girls in Korea and China. Traditionally, boys are considered to have more externalizing problems such as hyperactivity and aggressive behaviors, including rough-and-tumble play [57]. Parents/caregivers in the U.S. seem to have more tolerance or even encouragement for this type of physical play than in Asian cultures [57, 58].

It is unclear why the item “Does your child hurt himself on purpose?” was identified as a DIF item since no relevant cultural explanations can be hypothesized. Furthermore, the direction of gender differences was opposite in Korea and China, leading to the thought that more information may be needed on the definition of self-injurious behaviors in these two cultures. Random error of measurement might be a potential cause rather than actual behavioral differences.

Patterns of DIF items within countries: The items with gender DIF detected in each country may be worth further discussion (Table 5). In the U.S. sample, responses of items in the social communication area such as, “Can your child name a friend?”, “Does your child like to hear stories and sing songs?” and “Does your child use words to describe his feelings and the feelings of others?” indicated that boys were more likely to be rated with more problems behaviors than girls. The item, “Does your child sleep at least 8 hours in a 24-hour period?” showed a consistent DIF across age intervals in the U.S. only, indicating that girls were more likely to be rated with sleeping difficulties by parents across ages.

In the China sample, the item “Does your child stay away from dangerous things?” was identified as a DIF item, with boys reported as having more problems. This finding may be related to a more restrictive and controlling parenting style in China, [59, 60] which might affect parents’ perceptions in defining dangerous

Table 5: ASQ:SE Items with Gender DIF Detected in Single Country.

Items	6m	12m	18m	24m	30m	36m	48m	60m
The United States								
Sleep x hours	--	--	--	--	--	0.53***	--	0.43*
Change of Activity					--	--	--	-0.44***
Describe feelings						-0.62***	--	--
Name a friend						-0.57***	--	--
Look pointed			--	--	-0.51**			
Korea								
Destroy things					--	--	--	-0.64*
Explore places			--	--	--	--	0.46*	--
Look when you talk	--	--	--	--	--	-0.96*	--	--
Cry for long time	--	--	--	--	0.56*	--	--	--
Remain upset			0.53*	--				
Constipated	--	--	0.51*	--				
China								
Eating problems	--	--	--	--	--	--	--	0.59*
Away dangerous					--	--	--	-0.86*
Settle after activity					--	0.45*	--	--
Enjoy mealtimes	--	--	--	--	0.57*	--	--	--
Play alongside others					0.83**			
Calm down	--	--	--	-0.46*	--	--	--	--
Interested in around		--	-0.51*	--	--	--	--	--
Brazil								
Smile at you	--	0.65*	--	--				

Note1. Values are logits.

Note2. DIF level: -- = negligible, normal font = slight to moderate ($|DIF| \geq 0.43$ logits), bold font = moderate to large ($|DIF| \geq 0.64$ logits); blank = the item is not tested in the interval; positive values = girls are the preference group; negative values = boys are the preference group.

* $p < .05$. ** $p < .01$. *** $p < .001$.

things. Meanwhile, boys are commonly considered more active than girls, which may be why Chinese parents in this sample rated boys higher on this item. In addition, the DIF item “Does your child like to be hugged or cuddled?” might reflect a typical cultural phenomenon that hugging is uncommon for showing friendliness or for greeting purposes in China, especially for males.

In the Korea sample, “Does your child try to hurt other children, adults, or animals?” and “Does your child destroy or damage things on purpose?” showed moderate to large DIF, with boys as the preference group with higher scores. This might indicate that Korean parents have a lower tolerance for boys’ aggressive behaviors as a result of parenting attitudes that value stricter control and conformity of behavior [61]. On the other hand, Korean parents were more likely to rate girls higher on internalizing problem behaviors such as “When you leave, does your child remain upset and cry for more than an hour?” “Does your child get constipated or have diarrhea?” and “Does your child cry, scream, or have tantrums for long periods of time?” The DIF results for “Does your child explore new places, such as a park or a friend’s home?” were similar to the findings of a previously conducted review [62] that concluded that the Korean culture encourages boys to explore their environments more actively than girls.

The Brazilian sample has the least gender DIF items compared to the other cultures. This result is consistent with the cross-country research of Williams and Best [63] that concluded that Brazilian

children had atypically low gender stereotype scores compared to children from 25 countries. However, most of the questionnaires were answered by daycare teachers, which might have resulted in differences in ratings compared with other countries in which parents were most common.

Limitations

The findings of this study should be interpreted with caution due to several limitations related to the population and sample selection in participating countries. First, the questionnaires were completed by different caregivers and in different environments in participating countries. In Brazil, teachers were the most common raters. In China, parents answered the ASQ:SE but typically completed the assessment in pediatric clinics with help of health professionals, rather than in the home environment. Second, the U.S. and Brazil samples consisted of heterogeneous, multicultural populations (e.g., African American, Hispanic American, Brazilian Pardo, or Brazilian Black), and each subculture might have very different belief systems. China and Korea by contrast, are more homogeneous (e.g., Han Chinese and Koreans), reflecting more widely held practices and beliefs across the general population.

Finally, the data analyzed in the current study were from existing data sets, which had differing data collection methods. Administration, incentives for participants (e.g., gift cards), and the time span for data collection all varied from country to country. Nevertheless, the DIF statistical techniques that were used should be appropriate for

inspecting the gender differences in social-emotional competence of young children across these countries despite these differences.

Conclusion

The purpose of this study was to explore gender differences in young children's social-emotional competence across countries by using a DIF analysis. Through the completion of a social-emotional screening assessment, parents rated their perceptions of behaviors across gender groups from their own cultural perspective. We propose a visual representation, seen in (Figure 2), on how cultural beliefs may result in gender DIF on the ASQ:SE. Culture is assumed to have a significant impact on parenting style (e.g., authoritarian parenting, indulgent parenting), parents' perceptions of child behavior (e.g., beliefs, toleration of behavioral problems) [64], and also on children's social-emotional competence development (e.g., self-concept, self-image) [65,66]. Furthermore, the development of children's social-emotion competence is affected by parenting styles [67,68]. If there are differing expectations related to raising boys and girls, differences in expectations for social-emotional competence may manifest. Nevertheless, the interpretation of children's social-emotional/behavioral competence is filtered through parents' perception, as they completed the ASQ:SE. Therefore, in (Figure 2), the solid line starting from boy and girl's social-emotional competence becomes a dotted line to the ASQ:SE scores, reflecting the process of parental perception. Through these perceptions and a cultural lens, parental assessments of their children may differ from their actual performance.

Interpretation of parent-completed questionnaires should always consider parental perceptions and expectations in addition to biological gender-based difference, parenting style and/or the cultural phenomena. The findings of this study will further researchers' understanding of diverse parental expectations of gender in four different countries, and may contribute to the existing body of literature on cross-country comparison studies.

Future studies should consider adding more cultural groups to the analyses, which may make DIF patterns more obvious and

provide more cultural clues to explain the appearance of gender DIF. Second, comparing the gender DIF results from other social-emotional assessment tools (e.g., the Infant-Toddler Social and Emotional Assessment) [22] may provide supplemental information and also avoid the potential bias from only using one tool. Exploring gender difference across cultures may assist in interpreting screening results from an international perspective as well as aid in early identification of social-emotional delays across differing cultures and countries. The results can also be used to improve screening tools with item revisions for items with gender DIF across cultures. These revisions will help make assessment tools more accurate and culturally competent for a global population.

Conflict of Interest

Jane Squires acknowledges a conflict of interest as ASQ:SE author who receives royalties for publication of the ASQ:SE. The other authors have indicated they have no financial relationships relevant to this article to disclose and no potential conflicts of interest to disclose.

Acknowledgement

We acknowledge the many families and children that participated in our research studies.

REFERENCES

1. Campbell SB. Behavior problems in preschool children: a review of recent research. *Journal of Child Psychology and Psychiatry and Allied Disciplines* 1995; 36: 113-149.
2. Jellinek MS, Murphy J, Little M, Pagano ME, Comer DM, et al. Use of the pediatric symptom checklist to screen for psychosocial problems in pediatric primary care: A national feasibility study. *Archives of Pediatrics & Adolescent Medicine* 1999; 153: 254-260.
3. Lavigne JV, Gibbons RD, Christoffel KK, Arend R, Rosenbaum D, et al. Prevalence rates and correlates of psychiatric disorders among preschool children. *Journal of the American Academy of Child & Adolescent Psychiatry* 1996; 35: 204-214.
4. Feil EG, Walker HM, Severson HH. The early screening project for young children with behavior problems. *Journal of Emotional and Behavioral Disorders* 1995; 3: 194-202.
5. Lavigne JV, Arend R, Rosenbaum D, Binns HJ, Christoffel KK, et al. Psychiatric disorders with onset in the preschool years: II. Correlates and predictors of stable case status. *Journal of the American Academy of Child & Adolescent Psychiatry* 1998; 37: 1255-1261.
6. Shaw DS, Gilliom M, Ingoldsby EM, Nagin DS. Trajectories leading to school-age conduct problems. *Developmental Psychology* 2003; 39: 189-200.
7. Bricker D, Davis M-S, Squires J. Mental health screening in young children. *Infants and Young Children* 2004; 17: 129-144.
8. Slonim M. Children, Culture, and Ethnicity: Evaluating and Understanding the Impact. Garland Publication: New York, 1991.
9. Denham SA. Social-emotional competence as support for

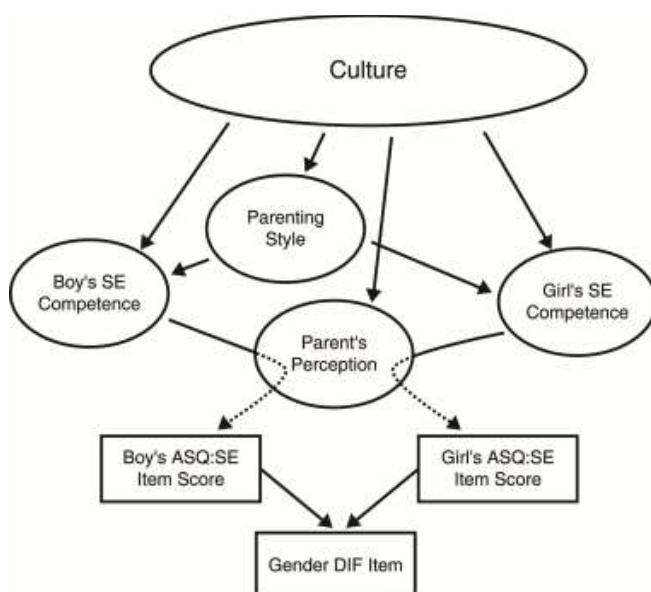


Figure 2: The effect of culture on ASQ: SE items with gender difference

- school readiness: What is it and how do we assess it? *Early Education and Development* 2006; 17: 57-89.
10. Jones SM, Bouffard SM. Social and emotional learning in schools: From programs to strategies. *Social Policy Report* 2012; 26: 1-22.
 11. Zins JE, Bloodworth MR, Weissberg RP, Walberg HJ. The scientific base linking social and emotional learning to school success. *Journal of Educational and Psychological Consultation*. 2007; 17: 191-210.
 12. Briggs RD, Stettler EM, Silver EJ, Schrag RDA, Nayak M, et al. Social-emotional screening for infants and toddlers in primary care. *Pediatrics* 2012; 129: e377-384.
 13. Briggs-Gowan MJ, Carter AS, Irwin JR, Wachtel K, Cicchetti DV, et al. The Brief Infant-Toddler Social and Emotional Assessment: screening for social-emotional problems and delays in competence. *Journal of Pediatric Psychology* 2004; 29: 143-155.
 14. Squires J, Bricker D, Twombly E. The Ages & Stages Questionnaires: Social-Emotional. A Parent-completed, Child-monitoring System for Social-emotional Behaviors. Baltimore: Paul H. Brookes, 2002.
 15. Carter AS, Briggs-Gowan MJ, Davis NO. Assessment of young children's social-emotional development and psychopathology: Recent advances and recommendations for practice. *Journal of Child Psychology and Psychiatry* 2004; 45: 109-134.
 16. Marks K, LaRosa A. Understanding developmental-behavioral screening measures. *Pediatrics in Review* 2012; 33: 448-458.
 17. Baggett KM, Warlen L, Hamilton JL, Roberts JL, Staker M, et al. Screening infant mental health indicators: An Early Head Start initiative. *Infants and Young Children* 2007; 20: 300-310.
 18. Beeber LS, Holditch-Davis D, Perreira K, Schwartz TA, Lewis V, et al. Short-term in-home intervention reduces depressive symptoms in Early Head Start Latina mothers of infants and toddlers. *Research in Nursing and Health* 2010; 33: 60-76.
 19. Heo KH, Lee JS, Squires J. Korea - The Ages and Stages Questionnaires: Social-Emotional User's Guide. Seoul: Seoul Rehabilitation Center, 2012.
 20. Heo KH, Squires J. Cultural adaptation of a parent completed social emotional screening instrument for young children: Ages and Stages Questionnaire Social-Emotional. *Early Human Development* 2012; 88: 151-158.
 21. Bian X, Wang R, Chen C-I. Ages and Stages Questionnaires: Social Emotional-Simplified Chinese Version. Shanghai, China: Shanghai Scientific and Technical Publishers, 2013.
 22. Carter AS, Briggs-Gowan MJ. Infant-toddler Social and Emotional Assessment. San Antonio, TX: Harcourt, 2006.
 23. LaFreniere PJ, Dumas JE. Social Competence and Behavior Evaluation in Children Ages 3 to 6 Years: The Short Form (SCBE-30). *Psychological Assessment* 1996; 8: 369-377.
 24. LeBuffe P, Naglieri J. The Devereux Early Childhood Assessment Clinical Form (DECA-C): A Measure of Behaviors Related to Risk and Resilience in Preschool Children. Kaplan Press: Lewisville, NC, 2003.
 25. Cole M. Cultural Psychology: A Once and Future Discipline. Harvard University Press: Cambridge, MA, 1998.
 26. Keller H. Cultures of Infancy. Psychology Press: New York, 2007.
 27. Hetherington EM, Parke RD, Locke VO. Child Psychology: A Contemporary Viewpoint. McGraw-Hill; 1999.
 28. Chen X, Hastings PD, Rubin KH, Chen H, Cen G, et al. Child-rearing attitudes and behavioral inhibition in Chinese and Canadian toddlers: a cross-cultural study. *Developmental Psychology* 1998; 34: 677.
 29. Carter AS, Briggs-Gowan MJ, Jones SM, Little TD. The infant-toddler social and emotional assessment (ITSEA): Factor structure, reliability, and validity. *Journal of Abnormal Child Psychology* 2003; 31: 495-514.
 30. Fantuzzo J, Grim S, Mordell M, McDermott P, Miller L, et al. A multivariate analysis of the revised Conners' Teacher Rating Scale with low-income, urban preschool children. *Journal of Abnormal Child Psychology* 2001; 29: 141-152.
 31. Kim M-H, Moon H. Infants' social-emotional adjustment within a childcare context of Korea. *Asia Pacific Journal of Education* 2011; 31: 487.
 32. Liu J, Cheng H, Leung PW. The application of the Preschool Child Behavior Checklist and the Caregiver-Teacher Report Form to Mainland Chinese children: Syndrome structure, gender differences, country effects, and inter-informant agreement. *Journal of Abnormal Child Psychology* 2011; 39: 251-264.
 33. Winsler A, Wallace GL. Behavior problems and social skills in preschool children: Parent-teacher agreement and relations with classroom observations. *Early Education and Development* 2002; 13: 41-58.
 34. Zsolnai A, Lesznyak M, Kasik L. Pre-school children's aggressive and pro-social behaviours in stressful situations. *Early Child Development and Care* 2012; 182: 1503-1522.
 35. Squires J, Bricker D, Twombly E. Ages & Stages Questionnaires: Social-Emotional Second Edition. Paul H. Brookes Publishing: Baltimore, MD, 2015.
 36. Filgueiras A, Landeira-Fernandez J. Políticas públicas na primeira infância: a importância do investimento público adequado e da avaliação global do desenvolvimento [Early childhood public policies: the importance of public investment on a global development assessment]. In: Magalhães AS, editor. *Perspectivas Contemporâneas da Teoria e da Prática em Psicologia [Contemporary Perspectives of Theory and Practice in Psychology]*. Brazil: Prospectiva Publishers, 2014.
 37. Guimarães S. *Censo 2010: The City of Rio de Janeiro*, 2012.
 38. National Education Council. Educação Brasileira: Indicadores e Desafios [Brazilian Education: Indexes and Challenges], 2013.

39. Velicer WF. Determining the number of components from the matrix of partial correlations. *Psychometrika* 1976; 41: 321-327.
40. Lorenzo-Seva U, Timmerman ME, Kiers HA. The Hull method for selecting the number of common factors. *Multivariate Behavioral Research* 2011; 46: 340-364.
41. Lorenzo-Seva U. The weighted oblimin rotation. *Psychometrika* 2000; 65: 301-318.
42. Lorenzo-Seva U, Ferrando PJ. FACTOR: A computer program to fit the exploratory factor analysis model. *Behavior Research Methods* 2006; 38: 88-91.
43. Arbuckle JL. Amos (Version 18.0) Chicago: SPSS, 2009.
44. Embretson SE, Reise S. Item Response Theory for Psychologists. Lawrence Erlbaum Associates: Mahwah, 2000.
45. Linacre JM. Winsteps Rasch Measurement Computer Program User's Guide. Winsteps.com: Beaverton, 2014.
46. Mantel N. Chi-square tests with one degree of freedom; extensions of the Mantel-Haenszel procedure. *Journal of the American Statistical Association* 1963; 58: 690-700.
47. Zwick R, Thayer DT, Lewis C. An Empirical Bayes Approach to Mantel-Haenszel DIF Analysis. *Journal of Educational Measurement* 1999; 36: 1-28.
48. Linacre JM. Winsteps (Version 3.91.0). Beaverton, 2015.
49. Sroufe LA. Emotional Development: The Organization of Emotional Life in the Early Years. Cambridge University: Press England, 1997.
50. Briggs-Gowan MJ, Carter AS, Skuban EM, Horwitz SM. Prevalence of social-emotional and behavioral problems in a community sample of 1-and 2-year-old children. *Journal of the American Academy of Child & Adolescent Psychiatry* 2001; 40: 811-819.
51. Wichstrom L, Berg-Nielsen TS, Angold A, Egger HL, Solheim E, et al. Prevalence of psychiatric disorders in preschoolers. *Journal of Child Psychology and Psychiatry* 2012; 53: 695-705.
52. Ringoot AP, Jansen PW, Steenweg-de Graaff J, Measelle JR, van der Ende J, et al. Young children's self-reported emotional, behavioral, and peer problems: The Berkeley Puppet Interview. *Psychological Assessment* 2013; 25: 1273.
53. Chen Q, Jiang Y. Social competence and behavior problems in Chinese preschoolers. *Early Education and Development* 2002; 13: 171-186.
54. Xue ZJ. The only-child psychological problems and treatments. *Education for Chinese After-school* 2010; 9: 14.
55. Chao RK. Beyond parental control and authoritarian parenting style: Understanding Chinese parenting through the cultural notion of training. *Child Development* 1994; 65: 1111-1119.
56. Coleman H, Charles G. Sexual development and behavior in children, 2009.
57. Smith PK, Boulton M. Rough-and-tumble play, aggression and dominance: Perception and behaviour in children's encounters. *Human Development* 1990; 33: 271-282.
58. Jose PE, Huntsinger CS, Huntsinger PR, Liaw F-R. Parental values and practices relevant to young children's social development in Taiwan and the United States. *Journal of Cross-Cultural Psychology* 2000; 31: 677-702.
59. Chiu L-H. Child-rearing attitudes of Chinese, Chinese-American, and Anglo-American mothers. *International Journal of Psychology* 1987; 22: 409-419.
60. Lin CYC, Fu VR. A comparison of child rearing practices among Chinese, immigrant Chinese, and Caucasian American parents. *Child Development* 1990; 61: 429-433.
61. Vinden PG. Parenting attitudes and children's understanding of mind: a comparison of Korean American and Anglo-American families. *Cognitive Development* 2001; 16: 793-809.
62. Yi SH. Transformation of child socialization in Korean culture. *Early Child Development and Care* 1993; 85: 17-24.
63. Williams JE, Best DL. Measuring Sex Stereotypes: A Multination Study. Sage Publications, 1990.
64. Okagaki L, Sternberg RJ. Parental beliefs and children's school performance. *Child Development* 1993; 64: 36-56.
65. Cousins SD. Culture and self-perception in Japan and the United States. *Journal of Personality and Social Psychology* 1989; 56: 124.
66. Wang Q. Culture effects on adults' earliest childhood recollection and self-description: implications for the relation between memory and the self. *Journal of Personality and Social Psychology* 2001; 81: 220-233.
67. Aunola K, Nurmi JE. The role of parenting styles in children's problem behavior. *Child Development* 2005; 76: 1144-1159.
68. Chang L, Schwartz D, Dodge KA, McBride-Chang C. Harsh parenting in relation to child emotion regulation and aggression. *Journal of Family Psychology* 2003; 17: 598-606.

ADDRESS FOR CORRESPONDENCE

Jane Squires, Professor, Early Intervention Program, University of Oregon, 901 E. 18th Avenue, Eugene, USA; Tel: 541.346.0807; E-mail: jsquires@uoregon.edu

Submitted Oct 08, 2015

Accepted Nov 05, 2015