Scandinavian Journal of Educational Research
Vol. 50, No. 5, November 2006, pp. 483–502

Routledge
Taylor & Francis Group

Young Children's Number Sense in China and Finland

Pirjo Aunio[a]*, Markku Niemivirta[a], Jarkko Hautamäki[a], Johannes E. H. Van Luit[b], Jiannong Shi[c] and Meiling Zhang[c]
[a]*University of Helsinki, Finland*; [b]*Utrecht University, The Netherlands*; [c]*Chinese Academy of Sciences, People's Republic of China*

This study examined the influence of nationality, age and gender on Chinese (N=130) and Finnish (N=203) pre-schoolers' number sense. Two highly correlated aspects of number sense were extracted: one reflecting the children's ability to organise and compare quantities (i.e. relational skills), and another pertaining to their ability to operate with number-word sequence (i.e. counting skills). The results showed a significant age-related gain in both aspects of number sense, whereas no gender differences were found. With respect to counting skills, the Chinese children outperformed the Finnish children irrespective of age, whereas in relation to relational skills, this was true only among the older children. Differences in language, teaching and cultural ethos are considered as alternative explanations for the findings.

Keywords: *Chinese; Finnish; Number Sense; Pre-schoolers*

Young Children's Number Sense

Children's number sense refers to their abilities to operate with quantities and number-word systems. One prominent theoretical view examines young children's number sense in terms of general and specific numerical skills. In their model of the development of young children's general and specific numerical skills, Case and his colleagues (Case & Okamoto, 1996; Griffin & Case, 1998) focus on the central conceptual structure for whole numbers. It is a cognitive structure that permits a child to interpret the world of quantity and numbers in increasingly sophisticated ways, to acquire new knowledge in this domain, and to solve the range of problems that it presents (Griffin, 2003). The central conceptual structure of number develops through two stages in early childhood. In the predimensional period, roughly at 4 years of age, children have two separate mathematical schemas: the global quantity schema that permits them to answer questions about "more" and "less", and the

*Corresponding author. Faculty of Behavioural Sciences, P.O. Box 26 (Teollisuuskatu 23), FIN-00014 University of Helsinki, Finland. Email: pirjo.aunio@helsinki.fi

ISSN 0031-3831 (print)/ISSN 1470-1170 (online)/06/050483-20

DOI: 10.1080/00313830600953576

initial counting schema that permits them to state how many objects are in a set. An important early central numerical structure is the mental number line, emerging at around 6 years (i.e. the unidimensional stage), in which the two above-mentioned schemas are merged. This mental number line consists of knowledge of written numerals, knowledge of number words, the ability to point to objects when counting, and knowledge of cardinal set values. Within the central conceptual concept of number at each developmental stage, a reciprocal relationship exists between the general (e.g. the categorisation used in counting situations) and the specific (e.g. knowledge of number words) numerical skills (Case, 1996). The development of number sense is thus a combination of progress in general and specific skills. In this process, associative and conceptual learning feed on each other in a reciprocal and dynamic way. Together these feedback loops form a hierarchical learning loop. When a child learns a specific skill, such as the ability to compare the numerosity of two sets, this learning has an impact on more general comparison skills. The same support mechanism works from general to specific skills as well. As a result of the activity within the learning loop, the rate of learning in low-exposure situations is faster than one might otherwise expect since it is mediated by general understanding and insights acquired in high-exposure situations.

The model by Case and his colleagues suggests that specific numerical skills build on general numerical skills and are mostly affected by social, environmental and cultural factors. It seems that the development of general skills, in contrast, is more dependent on maturation than on direct influence (e.g. teaching). Case also suggests that general skills may also be influenced by factors other than maturation, such as instruction, which would be secondary in nature and occur through the mechanisms of a hierarchical learning loop.

The Present Study

Cultural and Age-Related Differences in Number Sense

As pre-schoolers'[1] number sense already reflects skills (e.g. understanding number words) that are influenced by the social environment, differences in mathematical performance between young children from different cultures become of interest. Cross-cultural comparisons of pre-schoolers' mathematical skills have shown that the mathematical performance of Asian children is better than that of their non-Asian peers, for example, young Chinese children consistently outperform their western peers in abstract and object counting, concrete and mental addition and subtraction, and in the use of sophisticated strategies in mathematical problem solving (Ginsburg, Choi, Lopez, Netly, & Chi, 1997; Huntsinger, Jose, Liaw, & Ching, 1997; Miller, Smith, Zhu, & Zhang, 1995; Zhou, Cheng, Mottram, & Rosenblum, 1999; see Song & Ginsburg, 1987 for contrasting results). Differences in language (cf. Fuson & Kwon, 1992; Zhou & Boehm, 2001), teaching (Stevenson, Lee, & Graham, 1993) and cultural ethos (Campbell & Xue, 2001; Caplan, Choy, &

Whitmore, 1992; Jose, Huntsinger, Huntsinger, & Liaw, 2000; Stevenson et al., 1993; Tuss & Zimmer, 1995) have been considered as factors underlying Asian children's superior mathematical performance. However, since cross-cultural comparisons of mathematical performance have mainly focused on specific skills and particularly on children aged 6 years or above, our knowledge of the cross-cultural differences in young children's combined general and specific numerical skills is limited. Given that the numeracy skills develop from early infancy onwards, and advances in learning are enormous in early childhood, it is surprising how little interest the mathematical abilities of children under 6 years old have raised. Thus, our first aim was to examine the influence of nationality and age on young children's number sense.

Gender Differences

Regarding gender differences, the research shows that girls and boys possess identical primary numerical abilities (e.g. Dehaene, 1997; Nunes & Bryant, 1996), yet some differences favouring boys has been observed in elementary school-children's problem-solving and geometric performance (Dowker, 1998; Hall, Davis, Bolen, & Chia, 1999; Ma, 1995). Studies also demonstrate that, in their first school year, boys and girls may use different strategies for solving mathematical problems, but there is no difference in the level of performance (see Carr & Jessup, 1997). However, gender differences in general and specific numerical skills have attracted little attention (Torbeyns et al., 2002; Van de Rijt et al., 2003). Thus, another aim of our study was to examine such differences in young children's number sense.

The Assessment of Young Children's Number Sense

The Utrecht Early Numeracy Test (ENT; Van Luit, Van de Rijt, & Pennings, 1994) was used in this study to assess young children's number sense. Based on a developmental perspective, the ENT taps several aspects of young children's numerical and non-numerical knowledge of quantity. The test includes eight separate scales for assessing concepts of comparison, classification, one-to-one correspondence, seriation, the use of number words, structured counting, resultative counting, and general understanding of numbers (see Appendix B). Although the ENT is assumed to yield a common unidimensional measure of children's number sense (Van de Rijt, Van Luit, & Pennings, 1999), the practice of reporting results deviates from this. For example, Schopman, Van Luit, and Van de Rijt (1996) reported two domains, mathematical prerequisites (i.e. items assessing concepts of comparison, classification, one-to-one correspondence, and seriation) and counting skills (i.e. items assessing the use of number words, structured counting, resultative counting, and general understanding of numbers), whereas Van Luit and Schopman (2000) reported three domains, math prerequisites, counting skills, and general knowledge of numbers (see Torbeyns, 1999;Tzouriadou, Barbas, & Bonti, 2002 for yet another use of the test). From a theoretical point of view, the first four subscales of the instrument (concepts of comparison, classification, one-to-one correspondence, seriation), undoubtedly refer to the logical principles often identified as the key factors underlying children's understanding of quantities and relations (Piaget, 1966). Based on the type of skill these tasks measure (i.e. relational skills, or general numerical skills in Case et al.'s terminology), we will call them relational tasks. The rest of the scales (the use of number words, structured counting, resultative counting, and general understanding of numbers) focus more explicitly on the use and understanding of numbers (Fuson, 1988; Gelman & Gallistel, 1978; i.e. counting skills, or specific numerical skills in Case et al.'s terminology), thus the term counting tasks. As it seemed plausible that the ENT tapped two different but highly related aspects of young children's number sense, we explicitly tested this possibility before conducting any substantive comparisons.

Hypotheses

On the basis of our review of prior research, we formulated two sets of general hypotheses, one reflecting the measurement and the other reflecting substantive issues.

H1a: A model of two correlated factors of number sense describes the data better than a one-factor model.
H1b: The same measurement model holds for both Chinese and Finnish samples.
H2a: There is a significant gain in children's test scores as a function of age, and it is similar in both nationalities.
H2b: Chinese children outperform Finnish children in counting tasks (i.e. specific numerical skills), but there are no differences in relational tasks (i.e. general numerical skills).

Method

Participants

One hundred and thirty Chinese children (64 boys and 66 girls) from two pre-schools and one primary school in Beijing, and 203 Finnish children (95 boys and 108 girls) from nine pre-schools and one primary school in Helsinki participated in the study. The age of the children ranged from 55 months (4 years 7 months) to 90 months (7 years 6 months), the mean age being 73 months (*SD* 10.42) in China and 72 months (*SD* 8.06) in Finland. The participants were chosen from primary and pre-schools that had already been collaborating with university or academy researchers. In China, the schools' principals gave permission for children of the appropriate age to take part, while in Finland the parents authorised the children's participation.

Of children under 7 years old approximately 33% in Finland (Ministry of Social Affairs and Health, 2000) and 54% in Helsinki (City of Helsinki Social Services Department, 2004) attend public pre-schools. Primary schooling begins at the age of 7 years. The structured teaching of mathematical skills to young children in pre-school is not common practice. If the children are taught any mathematical skills, it is done informally, and quite unnoticed. In primary school, mathematics teaching is formal, and mathematics textbooks are the basic teaching materials. On average, children in their first school year have three mathematics lessons per week of 45 minutes each. Ninety-three percent of the Finnish children in our sample were in pre-schools.

In Mainland China primary school begins at 6 years of age, and at least in Beijing, most children aged 3 to 5 years attend public pre-schools. The pre-school teachers support the young children's number sense by following certain general guidelines (e.g. Zeng, 1995) that cover four developmental areas: physical, cognitive, linguistic and socio-behavioural. For example, the area of cognitive development includes physical knowledge, matching, classification, seriation, temporal ordering, common relations, cause and effect, conservation, and numbers. However, there are no statements indicating how often and how much the children need to practise these skills. Mathematics teaching in primary school is formal teaching with mathematics textbooks. Pupils in their first and second years have an average of three mathematics lessons per week, of 40 minutes each.

Instrument

The children's number sense was assessed using the Utrecht Early Numeracy Test (Van Luit et al., 1994) form A, which includes 40 items underlying the eight previously mentioned aspects of young children's number sense. The test is given individually and takes about 30 minutes for a child to complete. The child is given one point for a correct answer and zero for a wrong answer, the maximum being 40 points (e.g. Van de Rijt et al., 1999).

The test had previously been translated from Dutch into Finnish and was used in a study with 252 children demonstrating the usefulness and psychometric adequacy of the Finnish scale (Kautonen, 1999). For the present purposes, the ENT was translated from Finnish into Chinese. The back-translation procedure was used to confirm the linguistic similarity of the test in both languages. Consulting several experts in the field further ensured the cultural suitability.

Procedure

The data collection took place in autumn 1999. Trained experimenters administered the test to individual children in their own schools, usually in a separate quiet room with chairs and a table. After spending a few minutes establishing a rapport with the child, the experimenter provided the test materials (pictures, cubes, paper and pencil) according to the instructions for each task. Nine pre-school teachers and one research assistant in Helsinki, and four psychology students in Beijing conducted the test sessions.

Data analysis

A key assumption in testing for mean differences is that the measurement of the underlying construct is equivalent across the groups (cf. Van de Vijver & Tanzer, 1997). A solid method for assessing construct comparability, or measurement equivalence, across different groups is multiple-group mean and covariance structures (MACS) analysis (Little, 1997). MACS analysis extends the standard structural equation modelling (SEM) techniques (see Kline, 1998) by utilising information on observed mean structures in addition to the usual variance-covariance information. The strength of MACS analysis comes from the possibility to simultaneously fit factor models with mean structures into different groups and thereby assess group differences on disattenuated latent means and covariances. However, it generally requires relatively large samples and the procedure becomes progressively more complex when the number of groups to be compared increases. Thus, when a study with a relatively small sample involves many grouping variables, and when the effects of continuous variables and interactions are also of interest, MACS analysis becomes less practical. In such a situation, a plausible alternative is the so-called MIMIC (multiple causes, multiple indicators) modelling approach (cf. Muthén, 1989). A MIMIC model allows for simultaneous confirmatory factor analysis and, unlike MACS analysis, regression of the factor scores on several covariates without dividing the sample into numerous groups. MIMIC modelling also enables the examination of group differences in observed variables while controlling for group differences in the latent variable. This unique aspect permits the explicit consideration of possible sources of item bias or differential item functioning (DIF). In other words, the implicit assumption in the MIMIC model of strictly invariant covariance structures across different background variables can be relaxed by introducing direct effects from predictors to observed variables. A sort of "top-down" approach to invariance testing is thus utilised instead of the "bottom-up" approach (i.e. the hierarchical inclusion of parameter constraints in a model) typically used in MACS analysis (Horn & McArdle, 1992; Vandenberg & Lance, 2000).

Given the limited sample size and the combination of background variables included, the MIMIC modelling approach seemed most appropriate for the present purposes. Accordingly, the analyses were conducted in multiple stages. First, we compared the alternative one- and two-factor models across the nationalities. In doing this we sought to determine which of the two competing models fit the data better. On the basis of the results of the first stage, we fitted a series of MIMIC models for each covariate at a time. The purpose of this was to examine measurement invariance across each background variable, as well as to check for possible sources of item bias. Finally, we fitted a main effect model, in which nationality, gender and age were simultaneously included as covariates, and a full model, which included all two-way interaction effects in addition to the main effects. An illustration of the hypothesised final stage with the various aspects of MIMIC modelling is given in Figure 1.

The M*plus* program (Muthén & Muthén, 1998–2001) was used for all of the analyses. Since our data on number sense were based on binary variables (i.e. dichotomous test items), robust weighted-least square estimates (WLSMV; Muthén, duToit, & Spisic, 1997) were used. The advantage of the WLSMV estimation implemented in M*plus* is that it not only provides correct parameter estimates for analysis of categorical variables as well as mean- and variance-adjusted chi-square test statistics, it also works well with limited sample sizes. A disadvantage is that the typical procedures for comparing alternative non-nested models, such as comparing the values of information criteria (e.g. AIC or BIC), are not applicable. Thus, we followed general guidelines for model comparison by examining the overall model fit, the meaningfulness of the parameter estimates, and model parsimony when

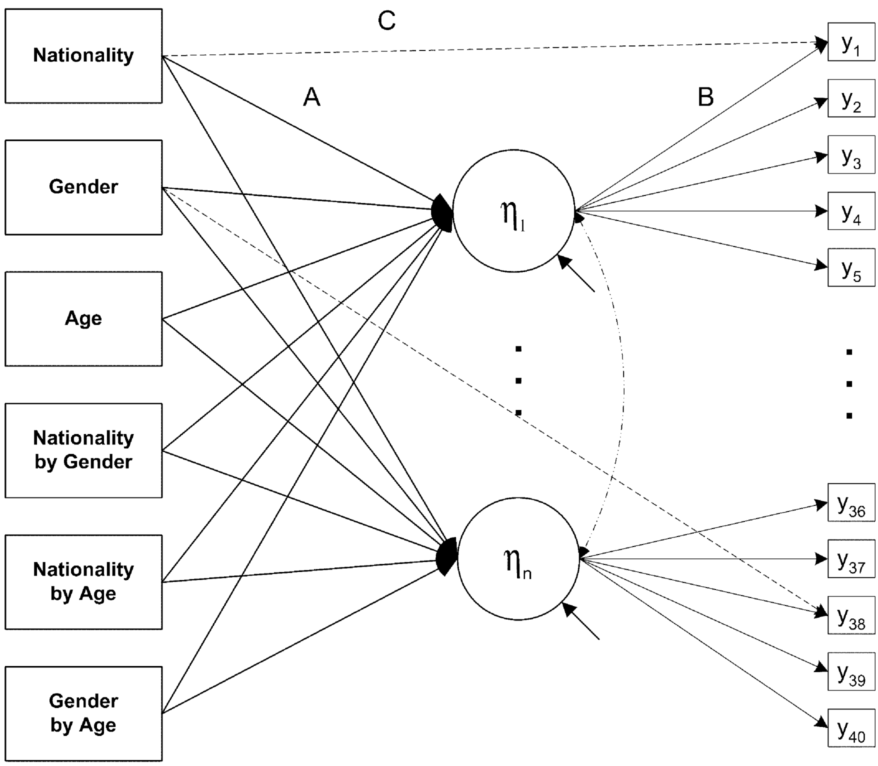

Figure 1. Path diagram for a hypothetical MIMIC model. A refers to regression estimates, which, in the case of a dichotomous covariate (e.g. nationality and gender), reflect mean differences between the given groups. B refers to factor loadings, and C illustrates the direct independent effect of a predictor (covariate) on an observed dependent variable (i.e. individual items)

deciding which of the alternative models described the data most adequately. In assessing the overall model fit, we followed the recent recommendations of Hu and Bentler (1999), who state that values around .95 for the Comparative Fit Index (CFI; Bentler, 1990) and values around .06 for the Root Mean Square of Error Approximation (RMSEA; Steiger, 1990) are indications of good fit.

Results

First, before comparing the alternative factor models, we examined the descriptive statistics for the individual items within both nationalities and across gender (see Appendix A). The purpose of this preliminary data screening was to improve the quality of the data by identifying items that were either too easy or too difficult. None of the items appeared to be too difficult (i.e. having the probability of a correct response equal to or below .05 in any group), whereas several items seemed too easy (i.e. having the probability of a correct response equal to or above .95 in any group). Accordingly, items 1, 3, 5, 6, and 7 were found to be too easy in both samples and were thus excluded. Furthermore, since the response probability for item 11 was above .95 for the Finnish children and the girls, and nearly .95 for the Chinese children (.93) and the boys (.94), it was excluded as well.

Next, we examined the assumed latent structure underlying children's ENT scores. We conducted comparative MACS analyses in order to test the construct equivalence across the two nationalities. In the one-factor solution, all items were set to load on a single latent factor, while in the two-factor solution, items 2, 4, 8–10 and 12–20 were set to load on the first factor (i.e. relational tasks) and items 21–40 on the second factor (i.e. counting tasks). For identification and scaling purposes, one item for each latent factor was fixed at one, and in the two-factor solution, the latent factors were allowed to correlate. All key parameters, factor loadings, item thresholds, and factor variances were constrained to be equal across the two nationalities (i.e. fully invariant models). The indices for the model fit are summarised in Table 1.

Although, in a strict sense, the chi-square statistics failed to indicate a good fit, other indices suggested that both models described the data well. The two-factor model seemed to fit the data slightly better than the one-factor model, even though the disattenuated correlation between the two factors was very high. Despite this strong dependency between the factors, we deemed that focusing on the one-factor

Table 1. Model fit indices for alternative CFAs

Model	χ^2 (df)	p	CFI	RMSEA	r_{F1F2}
I-factor model (fully invariant)	196.325 (131)	.000	.96	.055	
II-factor model (fully invariant)	184.868 (131)	.014	.97	.050	.93

Note. CFI=Comparative Fit Index; RMSEA=Root Mean Square of Error Approximation; r_{F1F2}=Between-Factor Correlation model only would potentially mask noteworthy covariate effects. Thus, we decided to run all subsequent analyses on both one- and two-factor solutions, although the results based on the two-factor model are emphasised and will be reported in more detail.

In the next stage, we further examined measurement invariance using the MIMIC modelling approach. Thus, one- and two-factor models were specified for each covariate at a time (cf. Figure 1). As in the previous stage, one item for each factor was set at one, and in the two-factor model, the factors were set to correlate. The direct effects of the covariates on individual items were constrained to zero, which permitted the estimation of first-order derivatives. First-order derivatives could then be used as a sort of modification index flagging for undue constraints.

As shown in Table 2, the base model results for each covariate indicated that the fit of the two-factor model was better than that of the one-factor model. The direct effects of nationality on the latent factors showed significant mean differences favouring the Chinese children. In other words, the Chinese children obtained higher scores on both the relational tasks, γ=.260 (z=2.671, p<.01) and the counting tasks, γ=.732 (z=6.081, p<.001). An examination of the derivatives revealed relatively large values for items 9, 20 and 31. Consequently, adding the direct effects of nationality to these three items resulted in a slightly improved model fit, $\chi 2$ (143)=210.119, p=.0002; CFI=.97; RMSEA=.038. The direct effects were also significant: κ=−.466 (z=−2.951, p<.01) for item 9, κ=.410 (z=2.966, p<.01) for item 20, and κ=.616 (z=4.207, p<.001) for item 31, respectively, although their inclusion did not markedly influence group differences in latent means. All in all, nationality explained 3% and 13% of the variance in the relational and counting

Table 2. Model fit indices for alternative MIMIC models

Model	χ^2 (df)	p	CFI	RMSEA	r_{F1F2}
Models with single covariates					
Nationality as covariate					
I-factor model	240.690 (143)	.000	.95	.045	
II-factor model	219.382 (143)	.000	.96	.040	.88
Modified model	210.119 (143)	.000	.97	.038	.88
Gender as covariate					
I-factor model	219.728 (143)	.000	.97	.040	
II-factor model	203.433 (144)	.001	.97	.035	.88
Age (continuous) as covariate					
I-factor model	230.286 (142)	.000	.93	.043	
II-factor model	210.810 (142)	.000	.95	.038	.88
Models with main effects and interaction effects					
Main effects					
I-factor model	230.756 (151)	.000	.93	.040	
II-factor model	214.407 (151)	.001	.95	.036	.88
Main effects & two-way interactions					
I-factor model	237.414 (154)	.000	.92	.038	
II-factor model	222.399 (160)	.001	.94	.034	.88

tasks, respectively. The interpretation of the independent direct effects is that, given the performance level on the relational tasks, the Finnish children were more likely to answer item 9 correctly, and the Chinese children were more likely to answer item 20 correctly, while given the performance level on the counting tasks, the Chinese children were more likely to answer item 31 correctly. To conclude, despite the minor bias detected in three items, the necessary condition for meaningful group comparison—partial measurement invariance—was clearly achieved.

Just as with nationality, the two-factor base model with gender as a covariate fit the data better than the one-factor model (Table 2). However, no gender differences emerged in the latent means, nor did the derivative values suggest sources of item bias. The two-factor model also described the data better than the one-factor model when age was included as a (continuous) covariate. In this case, no item bias was detected either, which supports the assumption of measurement invariance across both gender and age. However, clear age effects were found on the latent means. An examination of the latent means demonstrated a gain in scores on the relational tasks and an even stronger gain in scores on the counting tasks as a function of age: $\gamma=.638$, ($z=4.991$, $p<.001$) for the relational tasks, and $\gamma=1.023$ ($z=11.720$, $p<.001$) for the counting tasks, respectively. Age explained 45% of the variance in relational tasks and 43% of the variance in the counting tasks.

In order to examine the joint influence of the background variables on the children's number sense, we first fitted a main effect model in which all of the background variables were included as covariates and the independent direct effects were specified according to the results obtained in the previous stage, and we then added another model incorporating all of the two-way interactions. As expected, the two-factor main effect model fit the data better than the one-factor model (see Table 2). Both the direct effects of the covariates on the latent factors and the independent direct effects on the individual items were similar to those found earlier. The model explained 50% of the variance in the relational tasks and 57% of that in the counting tasks.

The final two-factor model with both main and interaction effects included fitted the data rather well. The fact that the fit indices were slightly below the recommended cut-off values—except for RMSEA, which indicated a good fit—was most likely due to penalising for model complexity. Other than that, there were no definite reasons to reject the model. In fact, the full model clearly increased the explanatory power. There was a significant increase in the amount of variance explained in the relational tasks (the change in R^2 was .09, which is statistically significant, $F=22.917$, $p<.001$), and the effect size (f^2) for the interaction effects over and above the main effects was .20, which, according to Cohen (1988), is medium in magnitude. The parameter estimates are summarised in Table 3, and an illustration of the final model is given in Figure 2.

The increase in the explained variance in the relational tasks was mainly due to a significant nationality by age interaction effect ($\gamma=0.492$; $z=3.613$, $p<.001$). For a detailed illustration of the interaction, we plotted the model-based factor score

Table 3. Parameter estimates for the final MIMIC model with main effects and interaction effects

	Standardized loadings				
Item	Factor 1	Factor 2	Threshold	Residual Variance	R-Squared
2	.72		−1.44	.69	.52
4	.47		−1.53	.90	.22
8	.48		−1.05	.89	.23
9	.66		−1.19	.77	.41
10	.47		−.83	.89	.22
12	.72		−1.06	.69	.52
13	.73		−1.23	.68	.53
14	.58		−.57	.83	.34
15	.77		−.42	.62	.60
16	.76		−.63	.64	.58
17	.73		−.46	.68	.54
18	.75		−.71	.65	.57
19	.72		−.52	.69	.52
20	.56		.18	.81	.45
21		.95	−.60	.22	.89
22		.80	−1.07	.56	.65
23		.91	−.74	.33	.83
24		.77	.22	.62	.59
25		.89	.23	.38	.80
26		.88	−1.10	.40	.78
27		.59	−.88	.82	.34
28		.73	.02	.68	.53
29		.77	.01	.62	.59
30		.84	.40	.49	.71
31		.58	−.09	.67	.72
32		.81	−.23	.54	.66
33		.74	−.62	.66	.55
34		.47	.33	.89	.23
35		.72	.51	.68	.52
36		.54	−.45	.85	.30
37		.66	−.81	.75	.43
38		.57	−.53	.83	.32
39		.36	−.10	.94	.13
40		.68	−.35	.73	.46
Variance	.31	.78			
R-Squared	.58	.57			
r_{F1F2}		.88			

estimates against the children's age by nationality (see Figure 3). As can be seen, the starting point of the score level is slightly higher for the Finnish children, but the increase in performance score over age is steeper for the Chinese children, thus resulting in the detected interaction effect. In other words, compared to the Finnish children, the Chinese children's level of relational skills as a function of age was increasingly higher. At this point, it is worth mentioning that the one-factor solution resulted in a similar interaction effect, although weaker in magnitude. However, as it was based on the overall score, it did not reveal the fact that the nationality by age

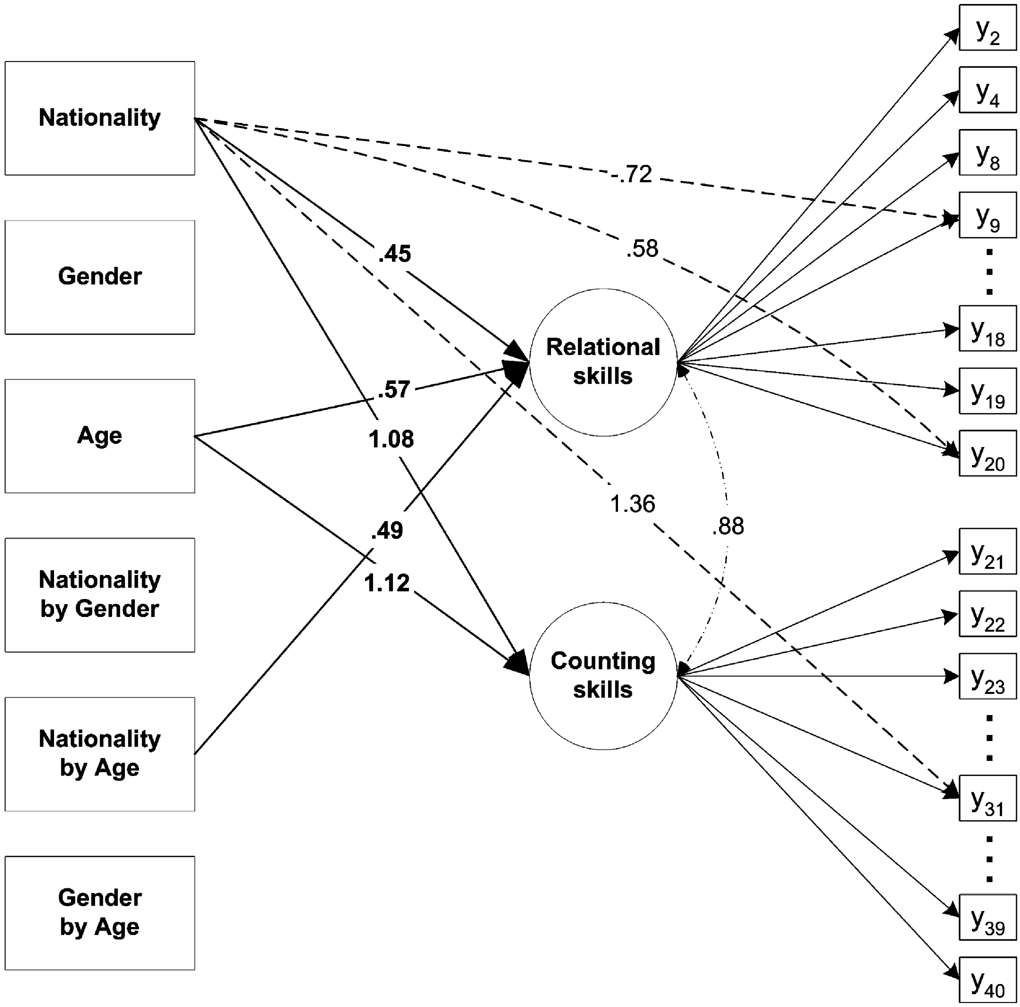

Figure 2. Path diagram for the final MIMIC model with both main effects and interaction effects. The dashed arrows represent the independent direct effects of predictors on individual items. Only significant effects (p<.05) are included

interaction was truly present only for the relational skills, thus masking the more valid differences.

Discussion

The purpose of this study was to examine the influence of nationality, age and gender on Chinese and Finnish children's number sense. In the following, we will discuss our findings in terms of the hypotheses we set, beginning with the issue of measurement comparability.

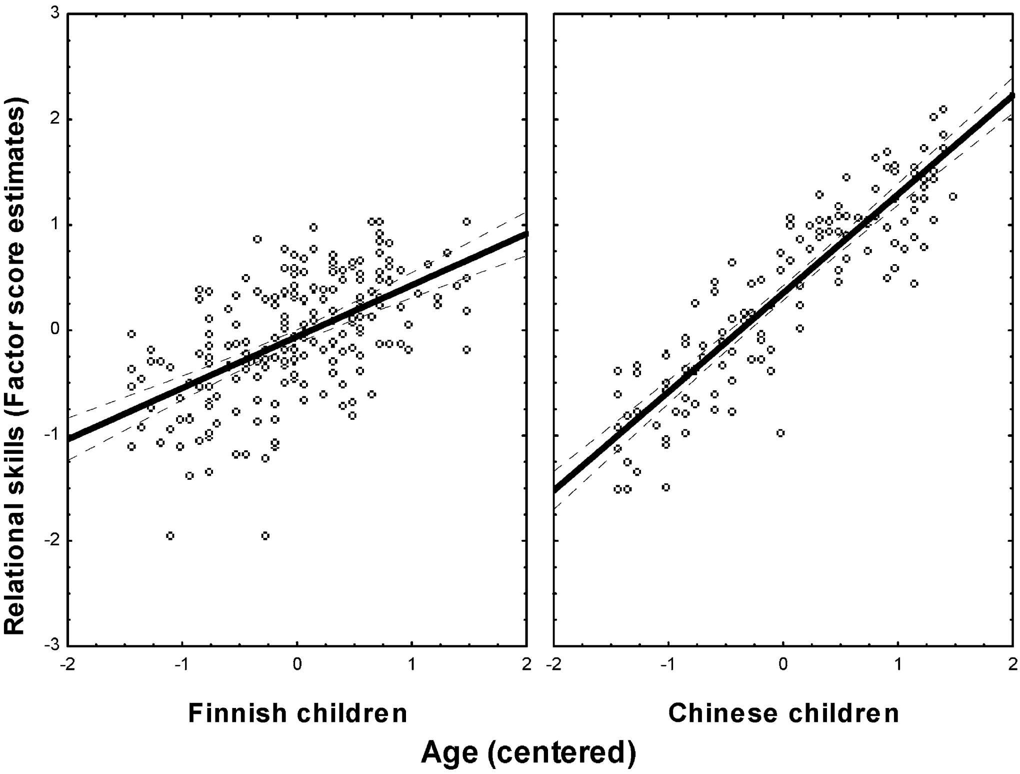

Figure 3. The Chinese and Finnish children's level of relational skills as a function of age

Our hypotheses concerning the measurement of children's number sense stated that a bi-dimensional structure would underlie their test scores, and that a similar structure would hold for both the Chinese and the Finnish children. Using the MIMIC modelling approach, we found that a two-factor model explained the data somewhat better than a single-factor model. This is not to say that the ENT measures two different mathematical ability domains, but rather to suggest that the test seems to capture variation in two specific and strongly interrelated aspects of one general latent construct. Although this, in a sense, goes against the assumption of a unitary number-sense construct, our view seems reasonable given the theoretical perspectives underlying the ENT item construction. As noted earlier, from a developmental perspective, the items clearly reflect two different aspects of the number-sense construct: the first 20 items measure children's ability to organise and compare given information, while the final 20 items measure their ability to use and operate with number-word sequence. According to the view adopted in this study, the former would reflect general numerical skills (i.e. relational skills), and the latter specific numerical skills (i.e. counting skills). Our substantive findings agree with this interpretation.

In support of our hypothesis, both samples showed a systematic increase in relational and counting skills as a function of age. This demonstrates a clear developmental effect in young children's number sense. However, the slightly dissimilar pattern of differences we found between the two nationalities suggests other sources of variation as well. In terms of the two different aspects of number sense mentioned above, it would seem that counting skills, on which systematic age-independent differences favouring the Chinese children were found, are influenced by different factors than relational skills, which showed differences only among the older children. This patterning of differences clearly supports the validity and utility of distinguishing the two forms of number sense. A more important question, however, is why the patterning of differences between the Chinese and the Finnish children emerged, and what sort of factors might underlie such differences.

First of all, our findings imply that the natures of relational and counting skills are somewhat different. In line with the ideas of Case et al. (e.g. Case & Okamoto, 1996) we could state that, since relational skills reflect general numerical abilities, they are less influenced by direct teaching or language differences than counting skills: as a reflection of specific numerical abilities, the latter are in part cultural products. Thus, unlike relational skills, counting skills rely on the use of a culturally based symbolic system (e.g. ten-base system), which, as such, is an explicit object of teaching. Now, in light of the fact that Chinese children begin their primary school earlier than their Finnish peers, and that mathematics is a more explicit part of Chinese pre-school curricula, it would seem that Chinese children's exposure to more systematic teaching might be one key source of explanation of the considerable differences counting skills in the two groups.

Language is another possible explanation. As argued by some scholars (e.g. Fuson & Kwon, 1992), the systematic number words in the Chinese language facilitate children's understanding of numbers, counting and the underlying ten-base structure, all of which make basic computation faster and more accurate. The Finnish number-word sequence, in contrast, is unsystematic, thus linking the acquisition of counting skills with rote learning and making it more prone to errors.

The interesting question here is this. If, as we argued, children's performance on relational tasks reflects their general numerical skills, which for the most part are uninfluenced by cultural factors, why did the Chinese children's performance level gradually surpass that of the Finnish children as a function of age? Although the present data do not provide direct support for our assumption, we suspect that Chinese children's relative gain in relational skills as a function of age is an indirect and secondary result of the systematic teaching of counting skills. This interpretation is supported by the notion of the hierarchical learning loop (Case, 1996), according to which the development in high-exposure specific skill situations accelerates the learning rate in low-exposure specific skill situations and in general numerical understanding. As to the age-related differences, given Ginsburg's (1997) claim that the development of informal mathematical skills is universal, but that the pace of development/learning may vary, we would maintain that Finnish children's number sense seems to develop at a slower pace than that of their Chinese peers. Naturally, this interpretation is speculative at best, and needs to be backed up with a thorough longitudinal study.

Our interpretations are restricted by some limitations of the study. First of all, the generally small sample sizes prevented us from using more comprehensive analysis, thus excluding the possibility of focusing on some more detailed substantive issues such as within-nationality interaction effects. More importantly, however, our samples were not nationally representative and were thus prone to different sources of bias. For example, we only had two urban pre-schools represented in the Chinese sample, which—due to the considerable differences between rural and urban life in China—possibly resulted in somewhat biased outcomes. Had the sample included rural children, who rarely even have access to pre-school education, we would have expected somewhat different results.

We were also aware of major language and instructional differences, but we cannot state with full confidence that these indications of cultural differences were in fact the factors responsible for the differences we found. Therefore, future studies should involve more explicit indicators of cultural differences. Not only should we then include more comprehensive and representative samples and control for the most obvious sources of variation, like differences in exposure to pre-school education or parental efforts of teaching about numbers (LeFevre, Clarke, & Stringer, 2002) within the samples, we should also consider focusing on countries and cultures that exhibit certain "controllable" cultural aspects in a convenient manner. For example, comparing Finnish children with Chinese-speaking and English-speaking Chinese children (e.g. from Singapore or Hong Kong) would certainly allow for more direct interpretations of the influence of language vs. cultural background on young children's mathematical skills. Moreover, in order to better understand how the "culture is put into practice", we should examine the concrete processes and practices that take place in real-world settings. Observational studies as such, or in combination with inclusive interventions, are virtually a necessity for future advancement in this field.

Acknowledgements

This study was supported by the grant from Helsinki University's Research Funds. We are grateful to research assistants and early childhood educators in two countries for help in data collection. We also would like to thank Dr Jin He for helping us with the translation procedure.

Note

1. We use the term pre-school education to refer to the early childhood education of children from 1 to 6 years old. To be specific, pre-school children in Mainland China are aged between 1 and 5 and in Finland between 1 and 6.

References

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*, 238–246.

Cambell, J., & Xue, Q. (2001). Cognitive arithmetic across cultures. *Journal of Experimental Psychology: General*, *130*, 299–315.

Caplan, N., Choy, M., & Whitmore, J. (1992, February). Indochinese refugee families and academic achievement. *Scientific American*, *266*, 36–42.

Carr, M., & Jessup, D. (1997). Gender differences in first grade mathematics strategy use: Social and metacognitive influences. *Journal of Educational Psychology*, *89*, 318–328.

Case, R. (1996). Reconceptualizing the nature of children's conceptual structures and their development in middle childhood. In R. Case & Y. Okamoto (Eds.), *The role of central conceptual structures in the development of children's thought* (pp. 1–26). Monographs of the Society for Research in Child Development, 61 (1–2), Serial#246.

Case, R., & Okamoto, Y. (Eds). (1996). *The role of central conceptual structures in the development of children's thought.* Monographs of the Society for Research in Child Development, 61 (1–2), Serial #246.

City of Helsinki Social Services Department. (2004). *Lasten päivähoidon tilastotietoja vuodelta 2003* [Statistics for children's day care in 2003; in Finnish]. Reports for planning and monitoring, 6. Helsinki, Finland.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed). New York: Academic Press.

Dehaene, S. (1997). *The number sense: How the mind creates mathematics*. New York: Oxford University Press.

Dowker, A. (1998). Individual differences in normal arithmetical development. In C. Donlan (Ed.), *The development of mathematical skills* (pp. 275–301). Hove, UK: Psychology Press.

Fuson, K. (1988). *Children's counting and concept of number*. New York: Springer Verlag.

Fuson, K., & Kwon, Y. (1992). Learning addition and subtraction: Effects of number words and other cultural tools. In J. Bideaud, C. Meljac & J-P. Fisher (Eds.), *Pathways to number: Children's developing numerical abilities* (pp. 283–306). Hillsdale, NJ: Erlbaum.

Gelman, R., & Gallistel, C. (1978). *The child's understanding of number*. Cambridge, UK: Harvard University Press.

Ginsburg, H. (1997). Mathematics learning disabilities: A view from developmental psychology. *Journal of Learning Disabilities*, *30*, 20–33.

Ginsburg, H., Choi, Y., Lopez, L., Netly, R., & Chi, C. (1997). Happy birthday to you: Early mathematical thinking of Asian, South American and U.S. children. In T. Nunes & P. Bryant (Eds.), *Learning and teaching mathematics: An international perspective* (pp. 163–207). East Sussex, UK: Taylor & Francis.

Griffin, S. (2003). The development of math competence in the preschool and early school years. In J. M. Royer (Ed.), *Mathematical cognition* (pp. 1–32). Greenwich, CT: Information age publishing.

Griffin, S., & Case, R. (1998). Re-thinking the primary school math curriculum: An approach based on cognitive science. *Issues in Education*, *4*, 1–51.

Hall, C., Davis, N., Bolen, L., & Chia, R. (1999). Gender and racial differences in mathematical performance. *The Journal of Social Psychology*, *139*, 677–689.

Horn, J. L., & McArdle, J. J. (1992). A practical and theoretical guide to measurement invariance in aging research. *Experimental Aging Research*, *18*(3–4), 117–144.

Hu, L.-t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1–55.

Huntsinger, C., Jose, P., Liaw, F-R., & Ching, W-D. (1997). Cultural differences in early mathematics learning: A comparison of Euro-American, Chinese-American, and Taiwan-Chinese families. *International Journal of Behavioral Development*, *21*, 371–388.

Jose, P., Huntsinger, C., Huntsinger, P., & Liaw, F-R. (2000). Parental values and practices relevant to young children's social development in Taiwan and the United States. *Journal of Cross-Cultural Psychology*, *31*, 677–702.

Kautonen, P. (1999). *Lasten matemaattiset taidot esikoulussa ja ensimmäisellä luokalla. Seuranta Utrechtin lukukäsitetestillä* [Children's mathematical abilities in preschool and primary one: A follow-up study with Early Numeracy Test; in Finnish]. Unpublished master's thesis. Joensuu, Finland: University of Joensuu, Department of Psychology.

Kline, R. B. (1998). *Principles and practice of structural equation modeling*. New York: Guilford.

LeFevre, J-A., Clarke, T., & Stringer, A. P. (2002). Influences of language and parental involvement on the development of counting skills: comparisons of French- and English-speaking Canadian children. *Early Childhood Development and Care*, *172*, 283–300.

Little, T. D. (1997). Mean and covariance structures (macs) analyses of cross-cultural data: Practical and theoretical issues. *Multivariate Behavioral Research*, *32*, 53–76.

Ma, X. (1995). Gender differences in mathematics achievement between Canadian and Asian education systems. *Journal of Educational Research*, *89*, 118–128.

Miller, K., Smith, C., Zhu, J., & Zhang, H. (1995). Preschool origins of cross-national differences in mathematical competence: The role of number-naming system. *Psychological Science*, *6*, 56–60.

Ministry of Social Affairs and Health. (2000). Päivähoidon kuntatason hallinnon vaihtoehtoja [Municipal possibilities for children's day care; in Finnish]. Working group memorandums 2000:15.

Muthén, B. (1989). Latent variable modeling in heterogeneous populations. *Psychometrika*, *54*(4), 557–585.

Muthén, B., du Toit, S. H. C., & Spisic, D. (1997). *Robust inference using weighted least squares and quadratic estimating equations in latent variable modeling with categorical and continuous outcomes.* Technical report.

Muthén, L. K., & Muthén, B. O. (1998–2001). *Mplus user's guide*. Los Angeles: Author.

Nunes, T., & Bryant, P. (1996). *Children doing mathematics*. Oxford, England: Blackwell.

Piaget, J. (1966). *Mathematical epistemology and psychology*. Dordrecht, The Netherlands: Reidell.

Schopman, E. A. M., Van Luit, J. E. H., & Van de Rijt, B. A. M. (1996). Learning and transfer of preparatory arithmetic strategies among young children with developmental lag. *Journal of Cognitive Education*, *5*, 117–131.

Song, M-J., & Ginsburg, H. (1987). The development of informal and formal mathematical thinking in Korean and U.S. children. *Child Development*, *58*, 1286–1296.

Steiger, J. H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research*, *25*, 173–180.

Stevenson, H., Lee, S., & Graham, T. (1993). Chinese and Japanese kindergartens: Case study in comparative research. In B. Spodek (Ed.), *Handbook of research on the education of young children* (pp. 519–535). New York: Macmillan.

Torbeyns, J. (1999). *Vlaamse normering van de Utrechtse Getalbegrip Toets* [The construction of the norms for the Early Numeracy Test for Flanders; in Dutch]. Unpublished master's thesis. Leuven, Belgium: University of Leuven.

Torbeyns, J., Van de Noortgate, W., Ghesquière, P., Verschaffel, L., Van de Rijt, B. A. M., & Van Luit, J. E. H. (2002). The development of early mathematical competence of 5- to 7-year-old children—A comparison between Flanders and The Netherlands. *Educational Research and Evaluation*, *8*, 249–275.

Tuss, P., & Zimmer, J. (1995). Causal attributions of underachieving fourth-grade students in China, Japan and the United States. *Journal of Cross-Cultural Psychology*, *26*, 408–426.

Tzouriadou, M., Barbas, G., & Bonti, E. (2002). Socio-cultural environment as a factor of differentation in mathematical reasoning. *Psychology*, *9*, 281–294.

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, *3*(1), 4–69.

Van de Rijt, B. A. M., Godfrey, R., Aubrey, C., Van Luit, J. E. H., Ghesquiere, P., & Torbeyns, J. L., et al. (2003). The development of early numeracy in Europe. *Journal of Early Childhood Research*, *1*, 155–180.

Van de Rijt, B. A. M., Van Luit, J. E. H., & Pennings, A. H. (1999). The construction of the Utrecht early mathematical competence scale. *Educational and Psychological Measurement*, *59*, 289–309.

Van de Vijver, F., & Tanzer, N. K. (1997). Bias and equivalence in cross-cultural assessment: An overview. *European Review of Applied Psychology*, *47*(4), 263–280.

Van Luit, J. E. H., & Schopman, E. A. M. (2000). Improving early numeracy of young children with special educational needs. *Remedial and Special Education*, *21*, 27–40.

Van Luit, J. E. H., Van de Rijt, B. A. M., & Pennings, A. H. (1994). *Utrechtse Gatalbegrip Toets, UGT* [Utrecht Test of Number Sense; in Dutch]. Doetinchem, The Netherlands: Graviant.

Wynn, K. (1998). Numerical competence in infants. In C. Donlan (Ed.), *The development of mathematical competence* (pp. 3–25). Hove, UK: Psychology Press.

Zeng, G. (1995). *Youer Fazhan Kecheng Ji Huodong Jianyi* [Young children's developmental courses and suggestions to their activities; in Chinese]. Beijing, China: Beijing Normal University.

Zhou, Z., & Boehm, A. (2001). American and Chinese children's knowledge of basic relational concepts. *School Psychology International*, *22*, 5–21.

Zhou, Z., Cheng, C., Mottram, L., & Rosenblum, S. (1999, April). *Understanding informal and formal mathematical abilities in Mainland Chinese and Chinese-American children.* Paper presented at the Biennial Meeting of the Society for Research in Child Development, Albuquerque, US.

Appendix A. Descriptive Statistics for the Items in the Samples

	Whole sample						Chinese sample						Finnish sample					
	All		Boys		Girls		All		Boys		Girls		All		Boys		Girls	
Item	*M*	*(SD)*	*M*	*(SD)*	*M*	*(SD)*	*M*	*(SD)*	*M*	*(SD)*	*M*	*(SD)*	*M*	*(SD)*	*M*	*(SD)*	*M*	*(SD)*
1	.96	(.20)	.97	(.18)	.95	(.22)	.95	(.21)	.97	(.18)	.94	(.24)	.96	(.20)	.97	(.18)	.95	(.21)
2	.92	(.26)	.94	(.24)	.91	(.28)	.94	(.24)	.92	(.27)	.95	(.21)	.92	(.28)	.95	(.22)	.89	(.32)
3	.99	(.00)	.99	(.00)	.99	(.11)	1.00	(.00)	1.00	(.00)	1.00	(.00)	.99	(.12)	.99	(.10)	.98	(.14)
4	.88	(.33)	.81	(.39)	.94	(.24)	.88	(.32)	.83	(.38)	.94	(.24)	.87	(.34)	.80	(.40)	.94	(.25)
5	.96	(.19)	.96	(.19)	.97	(.18)	.95	(.21)	.92	(.27)	.98	(.12)	.97	(.17)	.99	(.10)	.95	(.21)
6	.99	(.11)	.99	(.00)	.98	(.13)	1.00	(.00)	1.00	(.00)	1.00	(.00)	.98	(.14)	.99	(.10)	.97	(.17)
7	.96	(.19)	.97	(.18)	.96	(.20)	.98	(.15)	.95	(.21)	1.00	(.00)	.96	(.21)	.98	(.14)	.94	(.25)
8	.86	(.35)	.85	(.36)	.86	(.35)	.86	(.35)	.84	(.37)	.88	(.33)	.85	(.36)	.85	(.36)	.85	(.36)
9	.83	(.38)	.81	(.39)	.84	(.36)	.78	(.41)	.78	(.42)	.79	(.41)	.86	(.35)	.83	(.38)	.88	(.33)
10	.71	(.46)	.65	(.48)	.75	(.43)	.70	(.46)	.67	(.47)	.73	(.45)	.71	(.46)	.64	(.48)	.77	(.42)
11	.95	(.21)	.94	(.23)	.96	(.20)	.93	(.25)	.94	(.24)	.92	(.27)	.97	(.18)	.95	(.22)	.98	(.14)
12	.81	(.39)	.79	(.41)	.82	(.38)	.82	(.38)	.83	(.38)	.82	(.39)	.80	(.40)	.77	(.42)	.82	(.38)
13	.86	(.35)	.84	(.37)	.87	(.33)	.88	(.32)	.88	(.33)	.89	(.31)	.84	(.37)	.81	(.39)	.86	(.35)
14	.76	(.42)	.79	(.41)	.75	(.44)	.84	(.37)	.88	(.33)	.80	(.40)	.72	(.45)	.73	(.45)	.71	(.45)
15	.65	(.48)	.63	(.48)	.67	(.47)	.71	(.47)	.69	(.47)	.73	(.45)	.62	(.49)	.59	(.49)	.64	(.48)
16	.76	(.43)	.76	(.43)	.76	(.43)	.87	(.34)	.86	(.35)	.88	(.33)	.69	(.46)	.69	(.46)	.69	(.46)
17	.71	(.45)	.70	(.46)	.72	(.45)	.83	(.38)	.84	(.37)	.82	(.39)	.64	(.48)	.61	(.49)	.66	(.48)
18	.79	(.41)	.78	(.42)	.80	(.40)	.85	(.36)	.80	(.41)	.89	(.31)	.75	(.43)	.77	(.42)	.74	(.44)
19	.71	(.46)	.71	(.45)	.71	(.46)	.73	(.45)	.72	(.45)	.74	(.44)	.69	(.46)	.71	(.46)	.69	(.47)
20	.56	(.50)	.57	(.50)	.55	(.50)	.70	(.46)	.66	(.48)	.74	(.44)	.47	(.50)	.52	(.50)	.43	(.50)
21	.80	(.40)	.81	(.40)	.79	(.41)	.93	(.25)	.92	(.27)	.94	(.24)	.71	(.46)	.73	(.45)	.69	(.46)
22	.84	(.36)	.84	(.37)	.84	(.36)	.90	(.30)	.89	(.31)	.91	(.29)	.81	(.40)	.81	(.39)	.81	(.40)
23	.79	(.41)	.77	(.42)	.81	(.39)	.92	(.27)	.92	(.27)	.92	(.27)	.71	(.46)	.67	(.47)	.74	(.44)
24	.55	(.50)	.54	(.50)	.55	(.50)	.75	(.43)	.73	(.44)	.77	(.42)	.41	(.49)	.41	(.49)	.42	(.50)
25	.54	(.50)	.53	(.50)	.54	(.50)	.69	(.46)	.64	(.48)	.74	(.44)	.44	(.50)	.46	(.50)	.42	(.50)
26	.86	(.35)	.86	(.35)	.87	(.34)	.94	(.24)	.92	(.27)	.95	(.21)	.81	(.39)	.81	(.39)	.81	(.39)
27	.80	(.40)	.79	(.41)	.82	(.39)	.85	(.36)	.83	(.38)	.86	(.34)	.78	(.42)	.77	(.42)	.79	(.41)
28	.56	(.50)	.56	(.50)	.57	(.50)	.68	(.47)	.67	(.47)	.70	(.46)	.49	(.50)	.48	(.50)	.49	(.50)
29	.59	(.50)	.61	(.49)	.57	(.50)	.72	(.45)	.73	(.45)	.70	(.46)	.51	(.50)	.53	(.50)	.49	(.50)
30	.53	(.50)	.56	(.50)	.51	(.50)	.74	(.44)	.73	(.45)	.74	(.44)	.40	(.49)	.44	(.50)	.36	(.48)
31	.69	(.46)	.70	(.46)	.68	(.47)	.90	(.30)	.88	(.33)	.92	(.27)	.55	(.50)	.58	(.50)	.53	(.50)
32	.64	(.48)	.64	(.48)	.64	(.48)	.74	(.44)	.73	(.45)	.74	(.44)	.57	(.50)	.57	(.50)	.57	(.50)
33	.72	(.45)	.69	(.46)	.75	(.43)	.78	(.41)	.73	(.45)	.83	(.38)	.68	(.47)	.66	(.48)	.70	(.46)
34	.41	(.49)	.40	(.49)	.43	(.50)	.50	(.50)	.48	(.50)	.52	(.50)	.36	(.48)	.34	(.48)	.38	(.49)
35	.41	(.49)	.38	(.49)	.43	(.50)	.55	(.50)	.50	(.50)	.61	(.49)	.31	(.46)	.29	(.46)	.32	(.47)
36	.75	(.44)	.75	(.44)	.75	(.44)	.88	(.33)	.86	(.35)	.89	(.31)	.67	(.47)	.67	(.47)	.66	(.48)
37	.81	(.39)	.83	(.38)	.80	(.40)	.86	(.35)	.86	(.35)	.86	(.35)	.78	(.41)	.81	(.39)	.76	(.43)
38	.66	(.48)	.63	(.48)	.68	(.47)	.72	(.45)	.73	(.45)	.71	(.46)	.62	(.49)	.56	(.50)	.67	(.47)
39	.62	(.49)	.63	(.48)	.61	(.49)	.70	(.46)	.67	(.47)	.73	(.45)	.57	(.50)	.60	(.49)	.54	(.50)
40	.67	(.47)	.67	(.47)	.67	(.47)	.75	(.43)	.72	(.45)	.79	(.41)	.62	(.49)	.64	(.48)	.60	(.49)

Appendix B. A Description of the ENT Item Groups According to Van de Rijt et al. (1999) and Van Luit et al. (1994)

(1) Comparison (Items 1–5). This aspect is about the use of concepts in making comparisons between two non-equivalent cardinal, ordinal and measurement situations. The child has to demonstrate an understanding of concepts in drawings of order relations. *A sample item (4): Here are some Indians. Can you point out the Indian who has fewer feathers than the one you see here?*

(2) Classification (Items 6–10). These tasks require the grouping of objects in a class on the basis of one or more features. *A sample item (6): Look at these squares. Can you point out the square with five blocks but no triangles?*

(3) Making Correspondence (Items 11–15). This group includes tasks that measure children's understanding about one-to-one relationships of simultaneously presented objects. Overt and covert indicating tasks (e.g., moving blocks, drawing lines, pointing) are necessary in responding to the one-to-one correspondence items. *A sample item (12): (The child has 15 blocks) The administrator shows a drawing representing two dice with showing 5 and 6. Then the administrator asks: Can you put as many blocks on the table as are shown on the dice here?*

(4) Seriation (Items 16–20). This aspect refers to dealing with discrete and ordered entities. *A sample item (19): (The child has a paper and pencil). Here are some dogs. Each dog is going to fetch a stick. The big dog is going to fetch a big stick and the small dog is going to fetch a small stick. Can you draw lines from all of the dogs to the sticks that they fetch?*

(5) Using Number Words (Items 21–25). These tasks are about the ability to use number words in the number-word sequence up to 20. Number words must be produced forwards and backwards. *A sample item (23): Count further from 9 to 15*

(6) Synchronous and Shortened Counting (Items 26–30). This aspect refers to the counting of objects in organised and unorganised arrangements by pointing. *A sample item (28): The administrator puts 20 blocks on the table in an unorganised manner. The child is required to count the blocks. The child is allowed to point to the blocks with his/her finger or to move them.*

(7) Resultative counting (Items 31–35). These tasks require accurate counting and last-word response: pointing is not allowed. Most questions are of the kind: How many Xs are there? *A sample item (33): The administrator puts 15 blocks on the table in three rows of five with some space between them, and asks: How many blocks are there? The child is not allowed to point to the blocks with his/her finger or to move them.*

(8) General Knowledge of Numbers (Items 36–40). This aspect refers to the application of numeracy in daily life situations, which are represented in drawings. *A sample item (38): The administrator points to a picture of eight chickens and says: A farmer has eight chickens. He buys two more. The administrator then points to the picture with two chickens and continues: How many chickens does the farmer have now? Show the square with the right answer. The administrator points to the row of squares at the bottom of the paper.*