Personality and Individual Differences 50 (2011) 1222–1226

Contents lists available at ScienceDirect

Personality and Individual Differences

journal homepage: www.elsevier.com/locate/paid

Factor structure and sex differences on the Wechsler Preschool and Primary Scale of Intelligence in China, Japan and United States

Jianghong Liu [a,*], Richard Lynn [b]

[a] *School of Nursing and School of Medicine, 418 Curie Blvd., Claire M. Fagin Hall, University of Pennsylvania, Philadelphia, PA 19104-6096, USA*
[b] *University of Ulster, Coleraine, Northern Ireland BT52 1SA, UK*

ARTICLE INFO

Article history:
Received 4 November 2010
Received in revised form 11 February 2011
Accepted 17 February 2011

Keywords:
IQ
Sex differences
WPPSI
China
Japan
United States

ABSTRACT

This study presents data on the factor structure of the Wechsler Preschool and Primary Scale of Intelligence (WPPSI) and sex and cultural differences in WPPSI test scores among 5- and 6-year-olds from China, Japan, and the United States. Results show the presence of a verbal and nonverbal factor structure across all three countries. Sex differences on the 10 subtests were generally consistent, with a male advantage on a subtest of spatial abilities (Mazes). Males in the Chinese sample obtained significantly higher Full Scale IQ scores than females and had lower variability in their test scores. These observations were not present in the Japan and United States samples. Mean Full Scale IQ score in the Chinese sample was 104.1, representing a 4-point increase from 1988 to 2004.

1. Introduction

The study of human intelligence is a topic of global interest and has yielded insights from researchers across the United States, Greece, China, Japan, the Netherlands, Scotland, Israel, Spain, Scandinavia, United Kingdom, Germany, and other countries (Andersson, Sonnander, & Sommerfelt, 1998; Colom, García, Juan-Espinosa, & Abad, 2002; Goldbeck, Daseking, Hellwig-Brida, Waldmann, & Petermann, 2010; Hattori & Lynn, 1997; Lynn, 1998; Lynn & Dai, 1993; Lynn & Mulhern, 1991; Lynn, Wilberg-Neidhardt, & Margraf-Stiksrud, 2005). Understanding intelligence is important in the study of many aspects of human development, for example, recent research as shown that IQ serves as a mediating factor between early health risk factors and behavioral outcomes in children and adolescents (Liu, in press; Liu, Raine, Venables, & Mednick, 2004). The Wechsler battery of intelligence tests provides one of the most highly respected and widely used intelligence assessments in the United States (Watkins, Campbell, Nieberding, & Hallmark, 1995) and has been used extensively in international studies. However, there have been few studies of cross-cultural comparisons of intelligence using the Wechsler tests. The few investigations detailing cross-cultural views on intelligence have generally focused on comparisons of Western countries, such as that of the United States with France, countries within Europe, and the United States with Australia (Dockal, 2006; Grob et al., 2008; Kamieniecki & Lynd-Stevenson, 2002; Kaufman, Kaufman, Kaufman, & Simon, 1996). There have been three studies comparing Wechsler intelligence test scores between children of Western and Eastern countries, China and Japan with the United States (Dai & Song, 1995; Mann, Sasanuma, Sakuma, & Masaki, 1990; Wang, Zhang, & Lin, 1992) but these have not compared the factor structures.

There has also been little cross-cultural research on sex differences in general intelligence and specific abilities such as verbal, spatial, etc., measured by the Wechsler tests. Some authors argue that no or only minimal differences in general cognitive performance exist between males and females (e.g., Brody, 1992; Halpern & LaMay, 2000), while others have reported higher general intelligence of approximately 4 IQ points in males among adults (Irwing & Lynn, 2005; Lynn, 1994, 1999; Lynn & Irwing, 2004).

There is more consensus regarding gender-specific abilities on specific cognitive factors. It has become generally accepted that females obtain higher average scores on tests of some verbal abilities, including word fluency and verbal memory, while males perform better on tests of spatial and mathematical reasoning abilities (Halpern & LaMay, 2000), such as the Mazes subtest of the Wechsler Intelligence Scale for Children (Fairweather & Butterworth, 1977; Kaiser & Reynolds, 1985). Females have also demonstrated better performance in tasks requiring rapid access to long-term memory, processing speed involving memory skills, and symbol comparisons (Kaiser & Reynolds, 1985; Mann et al.,

* Corresponding author. Tel.: +1 215 898 8293; fax: +1 215 573 7492.
E-mail address: jhliu@nursing.upenn.edu (J. Liu).

0191-8869/$ - see front matter Published by Elsevier Ltd.
doi:10.1016/j.paid.2011.02.013

1990). However, with conflicting reports from studies on general intelligence, the literature can benefit from further clarification in the area of sex differences.

One of the most widely used assessments for measuring intelligence in young children is the Wechsler Preschool and Primary Scale of Intelligence (WPPSI), constructed in the United States by Wechsler (1967) for children ages 3–7 years. It has been translated and adapted in other nations such as the Third UK Edition, the Korean Wechsler Preschool and Primary Scale of Intelligence, and the Chinese version (Choi, Kwak, & Park, 1996; Gong & Dai, 1986, 1988; McKeown, 2003). The WPPSI has a two-factor model, which distinguishes between Verbal and Performance factors of intelligence (Ottem, 2003). The two-factor model has been shown to support grouping of individual subtests into Verbal and Performance components and has been conceptually adequate for the entire age range of the WPPSI (LoBello & Gülgöz, 1991). However, recent research has examined the fit of a four-factor model of the third edition of the WPPSI (Guo, Aveyard, & Dai, 2009; Keith, Fine, Taub, Reynolds, & Kranzler, 2006; Ottem, 2003). Examination of new factor models may improve our understanding of how intelligence test scores map onto specific areas of cognitive abilities, such as verbal versus nonverbal skills, performance skills, and knowledge-dependent skills. These important nuances may enhance our overall understanding of what is meant by "intelligence" and how that is manifest differently across different groups (e.g., by culture, by age, by sex, etc.).

Studies on sex and cultural differences on the WPPSI are lacking. Though there are a number of studies that have assessed sex differences in intelligence, few studies have done so in pre-school aged children, and even fewer specifically in Chinese populations. One study investigated intelligence in Chinese children using the WPSSI and found significant relationships between intelligence score and parental education, parental occupation, child caretakers, and location of education (Wang & Oakland, 1995). A second study used the WPPSI to report normative data on intelligence in Shanghai, China (Zhu, Lu, & Tang, 1984). The literature can benefit from additional research on Eastern–Western cultural comparisons as well as gender-specific comparisons on WPPSI outcomes.

The present study aims to address gaps in the intelligence literature in several ways. First, it presents data on the factor structure and cross-cultural comparisons between Chinese, Japanese, and American populations using the WPPSI, which, to our knowledge, has never been carried out before. Second, this study adds to the literature by including comparisons of WPPSI subtest scores by sex, not only within each cultural group but across all three groups. Third, this study presents data on intelligence within the Chinese population, which has only been reported hitherto by Wang and Oakland (1995).

2. Method

2.1. Sample

The Chinese sample was obtained from the Jintan Child Study. The sample is a preschool cohort of 1656 children consisting of 24.3% of all children in this age range in the Jintan city region, which includes the city center (Jianshe preschool), suburbs (Huacheng preschool), and the surrounding rural area (Xuebu and Huashan preschools). These four pre-schools were selected to represent the geographic, social, and economic profile of the region. The city of Jintan is located in the south-east of China, approximately 50 miles south of Nanjing and 120 miles north of Shanghai. The cohort consists of 55.5% boys and 44.5% girls. Among the original group of 1656 children, complete data are available for 1331 (728 boys and 603 girls). The children were between ages 5.0 and 6.11 years with a mean age of 4.65 ± 0.87 years. Further detailed information on the subjects, recruitment, and setting are given in Liu et al. (2010). Institutional Review Board approval was obtained from the University of Pennsylvania and the ethical committee for research at Jintan Hospital in China.

2.2. The test

The children were tested with the Chinese version of the Wechsler Preschool and Primary Scale of Intelligence (WPPSI). The test was constructed in the United States by Wechsler (1967) to assess the intelligence of children aged from 3 to 7 years of age. The WPPSI was standardized in China in 1988 (Gong & Dai, 1986, 1988) and has been shown to have good reliability in Chinese children (Gong & Dai, 1986, 1988; Yang, Liu, & Townes, 1994; Zhu et al., 1984).

The WPPSI consists of 10 subtests. Five of these make up a Verbal IQ (Information, Comprehension, Arithmetic, Vocabulary and Similarities) and five a Performance IQ (Object Assembly, Geometric Design, Block Design, Mazes and Picture Completion). The Verbal IQ and Performance IQs are combined to give a Full-Scale IQ, which is defined as the average of all cognitive abilities and is widely recognized as a good measure of general intelligence.

2.3. Procedure

The IQ data were collected between spring 2005 and spring 2007 for three groups of children when they were in their last year in preschool (equivalent of Western kindergarten). The test was administered by two research assistants (RAs), and overseen by a clinical psychologist whose training is in the area of cognitive methodology and who is an expert on IQ testing at Nanjing Brain Hospital. The two RAs are bachelor-prepared pediatric nurses from the Jintan Hospital who received an intensive three-week training course at the cognitive testing laboratory of Nanjing Brain Hospital. The training course includes four components: learning theory, performing IQ testing on children, exam-taking, and pilot testing. A pilot IQ test was performed on 32 five-year-old children in the sample to determine the reliability of testing prior to conducting a large scale IQ test. In keeping with Bracken's (1987) criteria for testing reliability (DeThorne & Schaefer, 2004; Gyurke, Marmor, & Melrose, 2000), two approaches to reliability have been taken: (1) Test–retest reliability (across time) within three weeks. The correlation between the test–retest was computed as adequate ($r = 0.87$, $p < .001$); and (2) the inter-rater reliability (across two examiners) was tested by the correlation between assessments of the two raters ($r = 0.91$, $p < .001$). We concluded that IQ test procedures are adequately reliable before we conducted the large-sample testing. Children were assessed in a quiet room at their preschool.

The data for Japan are derived from the Japanese standardization sample of the WPPSI ($n = 599$: 300 boys and 299 girls, ages 4.0–6.11 years, mean age 5.5 years) given by Hattori (2000, p. 26). The data for the United States are derived from the American standardization sample of the WPPSI ($n = 1199$: 600 boys and 599 girls, ages 4.0–6.11 years, mean age 5.5 years) given by Kaiser and Reynolds (1985) and the WPPSI manual (Wechsler, 1967).

3. Results

In order to be consistent with approaches utilized in previous factor analyses on the WPPSI (Hollenbeck & Kaufman, 1973; LoBello & Gülgöz, 1991), the data for the three countries were first analyzed by principal components. This showed that in all three countries there were two factors with eigenvalues greater than unity, showing the presence of two significant factors. To identify these, the principal components results were analyzed by varimax rotation with Kaiser normalization. The results are given in Table 1. This shows that in all three countries, the first factor can be identified as verbal ability with high loadings for Information, Vocabulary, Similarities and Comprehension. The second factor can be identified as non-verbal or perceptual/spatial ability with high loadings for Animal House, Picture Completion, Mazes, Geometric Design and Block Design.

To examine the consistency of the factor structure in the three countries, correlations were computed between the factor loadings. For factor 1, these are as follows: China–United States, $r = 0.987$; China–Japan, $r = 0.947$; Japan–United States, $r = 0.967$. For factor 2, these are as follows: China–United States, $r = 0.929$; China–Japan, $r = 0.978$; Japan–United States, $r = 0.930$. All these correlations are statistically significant at the 0.01 level and very high, showing almost identical factor structures in the three samples.

Sex differences (means and standard deviations) on the WPPSI in the three countries are shown in Table 2. This gives first the data for China boys and girls (means and standard deviations), followed by analysis of variance *F* values and the statistical significance of the values. The three columns on the right give the sex differences expressed as *ds* (standard deviation units). The correlations between the sex differences expressed as *ds* in the three countries are as follows: China–United States, $r = .793$, $p = .006$; China–Japan, $r = .453$, $p = .189$ (NS); Japan–United States, $r = .685$, $p = .029$. The correlations show that the sex differences in the subtests are fairly although not completely consistent in the three countries.

4. Discussion

There are five points of interest in the results. First, the factor structures of the WPPSI in the three countries are generally consistent. The data given in Table 1 show that in all the three countries the WPPSI contains a verbal ability and a non-verbal or perceptual/spatial ability factor. There are, however, three tests that do not load cleanly on one of the two factors in all three countries. The first is Arithmetic, which in all of the three countries loads about equally on both factors. The verbal loading of Arithmetic might be due to the fact that performance requires good verbal ability to understand questions and the instructions given by the test administrator. The second inconsistency is Animal House, which is cleanly loaded on Factor 2 in China and the United States, but has only a marginally higher loading on Factor 2 in Japan (0.37 compared with 0.30 on Factor 1). The third inconsistency is Picture Completion, which is cleanly loaded on Factor 2 in Japan and the United States, but has only a marginally higher loading on Factor 2 in China (0.51 compared with 0.46 on Factor 1). As both Animal House and Picture Completion are non-verbal tests, the most likely explanation for these inconsistencies lies in differences in the verbal instructions given in the three languages. For instance, if the verbal instructions in Animal House given in the Japanese are more cognitively demanding than in Chinese and English, the test would acquire a higher loading on the verbal Factor 1 and a lower loading on the non-verbal Factor 2, i.e. it would be measuring both verbal and non-verbal factors about equally.

Second, the WPPSI traditionally follows a two-factor model, which distinguishes between Verbal and Performance factors of intelligence (Ottem, 2003), which is consistent with our findings. The two-factor model has been shown to support the grouping of individual subtests into Verbal and Performance components and has been conceptually adequate for the entire age range of the WPPSI (LoBello & Gülgöz, 1991). Previous factor analyses have confirmed this grouping (Gyurke, Stone, & Beyer, 1990; Hollenbeck & Kaufman, 1973; O'Grady, 1990; Ramanaiah & Adams, 1979). Recent research examining a four-factor model for the WPPSI utilized the third version of the test (Guo et al.,

Table 1
Loadings of WPPSI subtests on two factors.

Test	China		Japan		USA	
	Factor 1	Factor 2	Factor 1	Factor 2	Factor 1	Factor 2
Information	.74	.30	.78	.16	.74	.35
Vocabulary	.81	.12	.77	.03	.71	.29
Arithmetic	.45	.56	.49	.45	.58	.50
Similarities	.66	.06	.68	.18	.77	.15
Comprehension	.79	.14	.72	.09	.78	.24
Animal house	.26	.56	.30	.37	.29	.60
Picture completion	.46	.51	.37	.53	.36	.60
Mazes	.10	.77	.10	.69	.14	.76
Geometric design	.04	.74	.02	.75	.19	.77
Block design	.09	.62	.13	.72	.29	.69

Table 2
Sex differences (means and SDs) on the WPPSI in China, United States and Japan.

Test	China: Boys	China: Girls	China: F	China: Sig.	China:*d*	Japan:*d*	US:*d*
Information	15.63 (2.60)	15.13 (2.69)	11.571	0.001	0.19***	−0.06	0.05
Vocabulary	19.76 (5.91)	18.63 (5.66)	12.512	0.000	0.20***	−0.06	0.05
Arithmetic	15.83 (2.51)	15.49 (2.30)	6.786	0.009	0.14**	0.05	−0.09
Similarities	12.72 (3.57)	13.11 (5.34)	2.579	0.109	−0.09	−0.02	−0.1
Comprehension	18.31 (3.76)	17.39 (4.09)	17.982	0.000	0.23***	0.03	0.01
Animal house	43.59 (9.40)	43.24 (9.48)	0.439	0.508	0.04	−.36***	−.31*
Picture completion	14.61 (3.54)	14.20 (3.63)	4.318	0.038	0.11*	.21**	0.01
Mazes	18.55 (5.10)	16.62 (5.97)	40.284	0.000	0.35***	.33***	0.23
Geometric design	17.02 (4.50)	17.19 (4.64)	0.411	0.522	−0.04	−0.05	−0.18
Block design	14.35 (3.75)	13.98 (4.10)	2.951	0.086	0.09	0.22	−0.12
Verbal IQ	104.80 (14.78)	102.94 (14.86)	5.171	0.023	0.13*	−0.01	−0.02
Performance IQ	104.81 (15.10)	103.15 (15.16)	4.010	0.045	0.11*	0.11	−0.1
Full-scale IQ	105.09 (14.19)	103.10 (14.55)	6.310	0.012	0.14*	0.06	−0.06

* Difference significant at the 0.05 level (2-tailed).
** Difference is significant at the 0.01 level (2-tailed).
*** Difference is significant at the 0.001 level (2-tailed).

2009; Keith et al., 2006; Ottem, 2003), while this study utilized the original WPPSI, perhaps explaining why the four-factor structure did not provide the best fit for our data. Our findings further support those of Wechsler's original hypothesis that verbal and performance components are the major underlying dimensions of intelligence. Additionally, a two-factor structure suggests that the verbal and performance components are most widely expressive of intelligence, as shown through this cross-cultural investigation.

Third, the sex differences in the Verbal, Performance and Full Scale IQs were negligible in the Japanese and American samples, whereas in the Chinese sample males obtained significantly higher IQs than Chinese females (i.e. by .14*d* or 2.1 IQ points on the Full Scale IQ). This may be due to the traditional preference for males that still exists in China today despite Westernization during the late 20th century. This male-preference is prevalent today even among younger and highly educated Chinese women (Loo, Luo, Su, Presson, & Li, 2009). Because of this, Chinese boys may receive more cognitive enrichment, including early educational exposure and nutritional advantages, during the prenatal and early childhood periods, as previous studies have shown that early nutritional factors have long-term effects on children's cognitive development (Liu, Raine, Venables, & Mednick, 2003).

Fourth, males performed better than females on the Mazes subtest in all three countries. This is a test of spatial ability, and this finding is consistent with those of other researchers who have found greater spatial task scores among males. For instance, a male advantage in spatial ability among adults and school-age children above the age of 14 years was well-established in the meta-analysis carried out by Voyer, Voyer, and Bryden (1995). However, as Geiser, Lehmann, and Eid (2008) have written, "there is still uncertainty as to the age when sex differences emerge" (p. 556). The present results suggest that a male advantage on spatial ability may be consistently present as young as ages of 4 and 5 years in China, Japan, and the United States.

Fifth, in the Chinese data, the scores obtained by females in seven of the 10 subtests and in the Verbal, Performance, and Full Scale IQs have larger standard deviations and therefore greater variability than the scores of males. The standard deviation of the Full Scale IQ was 14.19 for boys and 14.55 for girls. These results are contrary to the frequent contention that males have greater variability of IQs than females. This has been asserted since the early years of the 20th Century, when it was proposed by Ellis (1904), restated by Thorndike (1910) and Terman (1916) and later reaffirmed by Eysenck (1981) and Hedges and Nowell (1995), and more recently by Deary, Irwing, Der, and Bates (2007). The greater variability in IQs of females in the present Chinese data suggests that a greater variance in male IQ may not be a universal phenomenon.

Sixth, the Full Scale IQ of the Chinese sample was 104.1. This is in relation to a mean of 100 for the standardization sample obtained in year 1988. The higher mean obtained by the present sample indicates that the Chinese IQ has increased by approximately 4 IQ points over the years 1988–2004, representing a gain of 2.5 IQ points per decade. This is similar to the increases in intelligence that have been reported in Japan (Lynn, 1982) and in many other countries (Flynn, 1987). The rate of increase in IQs in the United States from 1932–2001 has been approximately 3.1 IQ points per decade, using the Wechsler and Binet intelligence tests (Flynn, 2007). This is the first report of a secular increase in intelligence in China. A number of explanations have been proposed for these increases including improvements in education, increased test sophistication, greater cognitive stimulation from more complex environments, improvements in child rearing, improvements in nutrition, and a reduction of inbreeding (Lynn, 2009).

Conflict of interest statement

No conflict of interest has been declared by the authors.

Acknowledgments

This study is supported, in part, by NIH/NIEHS K01-ES015 877 and R01 NIH/NIEHS R01-ES018858 to the first author.

References

Andersson, H. W., Sonnander, K., & Sommerfelt, K. (1998). Gender and its contribution to the prediction of cognitive abilities at 5 years. *Scandinavian Journal of Psychology, 39*, 267–274.

Bracken, B. A. (1987). Limitations of preschool instruments and standards for minimal levels of technical adequacy. *Journal of Psychoeducational Assessment, 5*, 313–326.

Brody, N. (1992). *Intelligence* (2nd ed.). San Diego, CA: Academic Press.

Choi, H. P., Kwak, K. J., & Park, K. B. (1996). The development of Korean version of WPPSI: The standardization study (1). *Korean Journal of Developmental Psychology, 9*, 60–70.

Colom, R., García, L. F., Juan-Espinosa, M., & Abad, F. J. (2002). Null sex differences in general intelligence. Evidence from the WAIS-III. *The Spanish Journal of Psychology, 5*, 29–35.

Dai, X. Y., & Song, Z. Z. (1995). Comparing inter sub-test scatter of the Wechsler intelligence scales between Chinese and Americans. *Chinese Journal of Clinical Psychology, 3*, 147–150.

Deary, I. J., Irwing, P., Der, G., & Bates, T. C. (2007). Brother-sister differences in the g factor in intelligence. analysis of full, opposite-sex siblings from the NLSY 1970. *Intelligence, 35*, 451–456.

DeThorne, L. S., & Schaefer, B. A. (2004). A guide to currently-used child nonverbal IQ measures. *American Journal of Speech-Language Pathology, 13*, 275–290.

Dočkal, V. (2006). Intercultural differences in intelligence in the mirror of European WISC-III standardizations. *Studia Psychologica, 48*, 213–228.

Ellis, H. (1904). *Man and woman: A study of human secondary sexual characteristics*. London: Walter Scott.

Eysenck, H. J. (1981). *Intelligence. Battle for the mind*. London: Pan.

Fairweather, H., & Butterworth, G. (1977). The WPPSI at four years: A sex difference in verbal-performance discrepancies. *British Journal of Educational Psychology, 47*(1), 85–90.

Flynn, J. R. (1987). Massive IQ gains in 14 nations: What IQ tests really measure. *Psychological Bulletin, 101*, 171–191.

Flynn, J. R. (2007). *What is intelligence? Beyond the Flynn effect*. New York, NY: Cambridge University Press.

Geiser, C., Lehmann, W., & Eid, M. (2008). A note on sex differences in mental rotation in different age groups. *Intelligence, 36*, 556–563.

Goldbeck, L., Daseking, M., Hellwig-Brida, S., Waldmann, H. C., & Petermann, F. (2010). Sex differences on the German Wechsler intelligence test for children (WISC-IV). *Journal of Individual Differences, 31*(1), 22–28.

Gong, Y. X., & Dai, X. Y. (1986). China-Wechsler younger children scale of intelligence (C-WYCSI). *Psychological Science, 2*, 23–30.

Gong, Y. X., & Dai, X. Y. (1988). China–Wechsler younger children scale of intelligence. *Acta Psychologica Sinica, 20*, 364–376.

Grob, A., Petermann, F., Lipsius, M., Costan-Dorigon, J., Petermann, U., & Daseking, M. (2008). Differences in Swiss and German children's intelligence as measured by the HAWIK-IV. *Swiss Journal of Psychology, 67*(2), 113–118.

Guo, B. L., Aveyard, P., & Dai, X. Y. (2009). The Chinese intelligence scale for young children: Testing factor structure and measurement invariance using the framework of the Wechsler intelligence tests. *Educational and Psychological Measurement, 69*(3), 459–474.

Gyurke, J. S., Marmor, D. S., & Melrose, S. E. (2000). The assessment of preschool children with the Wechsler preschool and Primary Scale of Intelligence – revised. In A. B. Bruce (Ed.), *The psychoeducational assessment of preschool children (3rd ed.)* (pp. 57–75). Needham Heights, MA: Allyn & Bacon.

Gyurke, J. S., Stone, B. J., & Beyer, M. (1990). A confirmatory factor analysis of the WPPSI-R. *Journal of Psychoeducational Assessment, 8*(1), 15–21.

Halpern, D. F., & LaMay, M. L. (2000). The smarter sex: A critical review of sex differences in intelligence. *Educational Psychology Review, 12*(2), 229–246.

Hattori, K. (2000). Sex differences in intelligence and its evolutionary implications. PhD thesis, University of Ulster.

Hattori, K., & Lynn, R. (1997). Male–female differences on the Japanese WAIS-R. *Personality and Individual Differences, 23*(3), 531–533.

Hedges, L. V., & Nowell, A. (1995). Sex differences in mental test scores, variability, and numbers of high-scoring individuals. *Science, 269*, 41–45.

Hollenbeck, G. P., & Kaufman, A. S. (1973). Factor analysis of the Wechsler Preschool and Primary Scale of Intelligence (WPPSI). *Journal of Clinical Psychology, 29*(1), 41–45.

Irwing, P., & Lynn, R. (2005). Sex differences in means and variability on the progressive matrices in university students: A meta-analysis. *British Journal of Psychology, 96*, 505–524.

Kaiser, S. M., & Reynolds, C. R. (1985). Sex differences on the Wechsler Preschool and Primary Scale of Intelligence. *Personality and Individual Differences, 6*, 405–407.

Kamieniecki, G. W., & Lynd-Stevenson, R. M. (2002). Is it appropriate to use United States norms to assess the ''intelligence'' of Australian children? *Australian Journal of Psychology, 54*(2), 67–78.

Kaufman, A. S., Kaufman, J. C., Kaufman, N. L., & Simon, M. (1996). The verbal and nonverbal intelligence of American vs. French children at ages 6 to 16 years. *Research in the Schools, 3*(1), 23–33.

Keith, T. Z., Fine, J. G., Taub, G. E., Reynolds, M. R., & Kranzler, J. H. (2006). Higher order, multisample, confirmatory factor analysis of the Wechsler intelligence scale for children–fourth edition: What does it measure? *School Psychology Review, 35*, 108–127.

Liu, J. (in press). Early health risk factors for violence: Conceptualization, review of the evidence, and implications. *Aggression and Violent Behavior*. doi:10.1016/j.avb.2010.12.003.

Liu, J., McCauley, L. A., Zhao, Y., Zhang, H., & Pinto-Martin, J.Jintan Cohort Study Group. (2010). Cohort Profile: The China Jintan child cohort study. *International Journal of Epidemiology, 39*, 668–674.

Liu, J., Raine, A., Venables, P., & Mednick, S. A. (2003). Malnutrition at age 3 years and lower cognitive ability at age 11: A prospective longitudinal study. *Archives of Pediatrics and Adolescent Medicine, 157*, 593–600.

Liu, J., Raine, A., Venables, P., & Mednick, S. A. (2004). Malnutrition at age 3 years predisposes to externalizing behavior problems at ages 8, 11 and 17 years. *American Journal of Psychiatry, 161*, 2005–2013.

LoBello, S. G., & Gülgöz, S. (1991). Factor analysis of the Wechsler Preschool and Primary Scale of Intelligence – Revised. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 3*, 130–132.

Loo, K. K., Luo, X. Y., Su, H., Presson, A., & Li, Y. (2009). Dreams of tigers and flowers: Child gender predictions and preferences in urban mainland Chinese sample during pregnancy. *Women and Health, 49*, 50–65.

Lynn, R. (1982). IQ in Japan and the United States shows a growing disparity. *Nature, 297*, 222–223.

Lynn, R. (1994). Sex differences in intelligence and brain size: A paradox resolved. *Personality and Individual Differences, 17*(2), 257–271.

Lynn, R. (1998). Sex differences in intelligence. Data from a Scottish standardization of the WAIS-R. *Personality and Individual Differences, 24*(2), 289–290.

Lynn, R. (1999). Sex differences in intelligence and brain size: A developmental theory. *Intelligence, 27*(1), 1–12.

Lynn, R. (2009). What has caused the Flynn effect? Secular increases in the Development Quotients of infants. *Intelligence, 37*(1), 16–24.

Lynn, R., & Dai, X. Y. (1993). Sex differences on the Chinese standardization sample of the WAIS-R. *The Journal of Genetic Psychology, 154*, 459–463.

Lynn, R., & Irwing, P. (2004). Sex differences on the Progressive Matrices: A meta-analysis. *Intelligence, 32*, 481–498.

Lynn, R., & Mulhern, G. (1991). A comparison of sex differences on the Scottish and American standardisation samples of the WISC-R. *Personality and Individual Differences, 12*, 1179–1182.

Lynn, R., Wilberg-Neidhardt, S., & Margraf-Stiksrud, J. (2005). Sex differences in general knowledge in German and Northern Irish university students. *Sexualities, Evolution and Gender, 7*(3), 277–285.

Mann, V. A., Sasanuma, S., Sakuma, N., & Masaki, S. (1990). Sex differences in cognitive abilities: A cross-cultural perspective. *Neuropsychologia, 28*, 1063–1077.

McKeown, P. (2003). United Kingdom. In J. Georgas, Weiss, L. G. van de Vijver, J. R. Fons, & D. H. Saklofske (Eds.), *Culture and children's intelligence. Cross-cultural analysis of the WISC-III* (pp. 77–88). San Diego, CA: Academic Press.

O'Grady, K. E. (1990). A confirmatory maximum likelihood factor analysis of the WPPSI. *Personality and Individual Differences, 11*, 135–140.

Ottem, E. (2003). Confirmatory factor analysis of the WPPSI, WPPSI-R, and the WISC-R: Evaluation of a model based on knowledge-dependent and processing-dependent subtests. *Journal of Psychoeducational Assessment, 21*, 3–15.

Ramanaiah, N. V., & Adams, M. L. (1979). Confirmatory factor analysis of the WAIS and the WPPSI. *Psychological Reports, 45*, 351–355.

Terman, L. M. (1916). *The measurement of intelligence*. New York: Houghton Mifflin.

Thorndike, E. L. (1910). *Educational psychology*. New York: Houghton Mifflin.

Voyer, D., Voyer, S., & Bryden, M. P. (1995). Magnitude of sex differences in spatial ability: A meta-analysis and consideration of critical variables. *Psychological Bulletin, 117*, 250–270.

Wang, F. F., & Oakland, T. (1995). Qualities related to intelligence among Chinese children between ages 4 and 6. *School Psychology International, 16*, 59–65.

Wang, X., Zhang, H., & Lin, C. T. (1992). Factor analysis of the WISC–R and the WISC–CR: A cross-age comparison between American and Chinese samples. *Psychological Science, 5*, 11–16.

Watkins, C. E., Campbell, V. L., Nieberding, R., & Hallmark, R. (1995). Contemporary practice of psychological assessment by clinical psychologists. *Professional Psychology: Research and Practice, 26*, 54–60.

Wechsler, D. (1967). *Wechsler preschool and Primary Scale of Intelligence*. New York: Psychological Corporation.

Yang, L. L., Liu, M. L., & Townes, B. D. (1994). Neuropsychological and behavioral status of Chinese children with acyanotic congenital heart disease. *International Journal of Neuroscience, 74*, 109–115.

Zhu, Y. M., Lu, S. Y., & Tang, C. H. (1984). The employment of the Wechsler pre-school and Primary Scale of Intelligence in urban Shanghai. *Information on Psychology Science, 5*, 22–29 [in Chinese].