Cross-Cultural Equivalence and Psychometric Properties of the Traditional Chinese Version of the Inviting School Survey-Revised

Citation

Kenneth H. Smith
Australian Catholic University

Abstract

The Inviting School Survey-Revised (ISS-R) was adapted and translated into Traditional Chinese (ISS-RC), using a five-step process, based on international test administration guidelines, involving judgmental, logical, and empirical methods. Both versions were administered to a convenience sample of Chinese-English fluent Hong Kong school community members (administrators, teachers, students, parents, and support personnel). A series of repeated measures ANOVAs revealed equivalence between the two versions, ISS-R and ISS-RC Total Scale and Subscales (variances, subscale correlations, internal consistency) other than the Program Subscale overall mean difference. Item analyses, utilizing repeated measures ANOVAs, revealed significant differences in 11 of the 50 scale items. Suggestions for further development and refinement of the Chinese Invitational School Survey (ISS-RC) are presented. Additionally, recommendations for future research and application of the ISS-R and ISS-RC are provided

The purpose of the present study is twofold; (1) to describe the processes and procedures used to adapt and translate the Inviting School Survey-Revised (ISS-R) into Traditional Chinese (ISS-RC) for use in Hong Kong and mainland China and (2) to assess the psychometric equivalence of the two versions.

There is rapidly growing evidence that school climate is one of the most important contributors to student achievement, success, and psychological wellbeing (Fan, Williams, & Corkin, 2011; Zullig, Koopman, Patton, & Ubbes, 2010; Cohen, McCabe, Michelli, & Pickeral, 2009). School climate also heavily influences healthy development as well as effective risk prevention, positive youth development, and increased teacher and student retention (Cohen et al., 2009; Huebner & Diener, 2008).

School climate may be defined in a number of ways. This study viewed school climate as that which reflects the perceptions of the social, emotional, and academic experiences of school life by students, administrators, teachers, parents, and the wider community. That is, school climate reflects a subjective experience in the school (Cohen, 2006; Freiberg, 1999).

In order to make informed decisions regarding school development it is important for a school administrator to be aware of how the school community perceives the school (school climate) and, to have access to a reliable and valid instrument that purports to measure school climate.

Such an instrument has existed since the early 1990s, the Inviting School Survey-ISS (Purkey & Fuller, 1995; Purkey & Schmidt, 1990). The ISS was designed to assess the total school climate and the five environmental areas as outlined by Invitational Education theory: People, Places, Policies, Programs, and Processes (Purkey & Novak, 1996, 2008; Purkey & Schmidt, 1990). The ISS is a 100-item, Likert-type, hand-scored instrument, utilized by few schools and there are no psychometrics, such as norms, reliability and validity indices, supporting the instrument.

To address this shortcoming, Smith (2005) significantly revised the 100-item instrument to be a 50-item, on-line, computer scored instrument titled, the Inviting School Survey-Revised (ISS-R). The ISS-R provides school communities a "user-friendly", theoretical-grounded, empirical-based instrument that assists in evaluating their schools for future development as the ISS-R identifies areas of strength and weakness in a school's climate.

Following the revision, the ISS-R has been utilized Australia, New Zealand, North America, Asia, and Africa. In 2006, 18 schools (596) participants completed the ISS-R.

Corresponding Author:

Kenneth H. Smith, Ph.D. is an associate professor at the Australian Catholic University in Melbourne, Australia.

Ken.Smith@acu.edu.au

I am particularly grateful to Kai Hung (Peter) Wong, Chief Curriculum Development Officer, Education Bureau, Hong Kong, and So Chan So Ming (Clio), Principal, Creative Primary School, Kowloon for their contribution in making this research project possible.

In 2010, 78 schools, over 6,000 participants have used the ISS-R (Smith, 2012).

In 2010, as a result of the huge increase in use of the ISS-R, particularly in Hong Kong and mainland China it was decided to adapt and translate the ISS-R into Traditional Chinese.

The Inviting School Survey-Revised (ISS-R)

The Inviting School Survey-Revised (Smith, 2005), based upon the tenets of Invitational Education, is designed to empirically identify areas in a school that are inviting and dis-inviting. Invitational Education theory is a theory that is strongly grounded on well-established psychological paradigms such as Perceptual Psychology (Combs, Richards, & Richards, 1976; Combs, 1962), Cognitive-Behavior (Meichenbaum, 1974, 1977; Ellis, 1962, 1970), and Self-Concept (Purkey, 1970; Jourard, 1968; Rogers, 1969).

The ISS-R is a 50-item, five-point Likert-type scale based on the 100-item Inviting School Survey-ISS (Smith & Bernard, 2004; Purkey & Fuller, 1995; Purkey & Schmidt, 1990). Both the original ISS and the ISS-R are designed to be completed by students (age 8-9 and older), parents, teachers, school administrators, and others associated with the school, such as counselors, psychologists, and social workers.

The Inviting School Survey-Revised (ISS-R) is a theoretical-based instrument, designed to assess the invitational qualities of the total school climate and the five environmental domains of People, Programs, Processes, Policies, and Places, as outlined in Invitational Education theory (Purkey & Novak, 1996). For a comprehensive description of the 5P's go to Purkey and Novak, *Fundamentals of invitational education*, 2008.

The *ISS-R* is designed for electronic, self-administration through the IAIE website. Individuals completing the ISS-R are asked to respond to all items ranging from 1, 'Strongly Disagree' to 5, 'Strongly Agree' (0, 'Not Applicable' is treated as missing, if a question is not relevant to the participant's school context).

The ISS-R is a theoretical five-factor model, where factors pertain to 1. People (16 items), 2. Program (7 items), 3. Process (8 items), 4. Policy (7 items), 5. Place (12 items) and the Total scale comprised of the 50 items.

The validity of the ISS-R has been empirically documented and its internal consistency reliability has been reported to range from .86 to .88 for the Total scale (Smith, 2005).

For further details of the Inviting School Survey-Revised (ISS-R) refer to Smith (2012), *Manual for the Inviting Survey (ISS-R): A survey for measuring the invitational qualities (I.Q.) of the total school climate.*

Method

Procedure

The following are the five steps used to adapt and translate the ISS-R into Traditional Chinese. As stated previously, the ISS-R is comprised of a brief instruction statement and 50 randomized, positively-worded items requiring a Likert-type response on a five point. Three additional demographic questions provided information about the respondents' position, gender, and age (students only).

The ISS-R was translated and adapted into Traditional Chinese (ISS-RC) using the guidelines set down by the International Test Commission (ITP, 2010), the American Psychological Association (1999) and contributing suggestions from Sousa and Rojjanasrirat, (2010). These guidelines address concerns about test context, test development and adaptation, administration procedures, and score interpretations. The guidelines require that test adaptation and translation procedures should use systematic judgmental evidence at every step.

Stage 1. Two bilingual individuals translated (forward translation) the English version, the ISS-R, into a Traditional Chinese version (ISS-RC). Certain words, such as 'everyone' were changed to be more meaningful and appropriate within the Chinese context.

Stage 2. Following the forward translation of the ISS-R, a blind back-translation was completed, in which an independent bilingual expert, different from the two translators used in Step 1, converted the translated instrument back into the original English language without having seen the original instrument.

Stage 3. A panel of two bilingual individuals, experts in the field of invitational education, examined both versions in terms of consistency, grammar, and structure. Minor discrepancies between the original and translated instruments were identified. Revisions were made in the translated Chinese measure until the two forms appeared to be equivalent. After reaching a consensus in relation to the consistency of the forward and backward translations of the ISS-R, a Traditional Chinese draft version of the ISS-R was produced.

Stage 4. As an additional confirmation of the equivalence of the two ISS versions an external company, Second Language Testing Inc. (SLTI), Rockville, Maryland, judged the equivalence of the two versions. SLTI reviewed the Chinese translation of the ISS-R and suggested some minor revisions. The report from this company identified no discrepancies between the 50 items. However, the

demographic questions asked at the beginning of the instrument were found to be non-equivalent and suggestions for change were submitted.

Stage 5. Implementing the changes as suggested by SLTI, a panel of 3 individuals, conversant in invitational education, compared the English and Chinese versions. After reaching a consensus that the two versions were equivalent, the final Chinese version of the ISS-R was produced. English and Chinese versions of the Inviting School Survey-Revised can be found in the appendix.

Participants

A convenience sample of 67 participants from a number of Hong Kong educational institutions, fluent in English and Chinese, completed both versions of the 50-item Inviting School Survey (ISS-R and the ISS-RC) at one sitting. Of the 67 paired-instruments completed, four participants (6%) were eliminated from further analysis because of excessive missing items (greater than 10). The final sample, as depicted in Table 1, consisted of 43 females (68%) and 20 males (32%) in positions of administrators (16, 25%), teachers (25, 40%), counselors (2, 3%), students (6, 10%), parents (11, 17%), and others (3, 5%).

Table 1. Number of Participants by School Community Status

Participant	Females		Males		Total	
	N	%	N	%	N	%
Administrator	9	21	7	35	16	25
Teacher	15	35	10	50	25	40
Counselor	2	1	0	0	2	3
Parent	8	19	3	15	11	17
Student	6	14	0	0	6	10
Other	3	7	0	0	3	5
Total	43	100	20	100	63	100

Data Preparation and Analysis

All data analyses were conducted using PASW Statistics Version 18 (SPSS, 2009) and significant levels for the analyses were at $\alpha < .05$. However, application of the Bonferroni correction statistic, which would restrict the acceptable alpha level for interpretation to alpha of .05 divided by the number of tests used in each of the Total scale and the five subscales, was employed, .05/6 = .008; while for the 50-item-difference analyses .05/50 = .001.

Initial inspection of the data indicated some incomplete data. Of the 67 participants, four participants had more than ten missing items and were removed from further analysis. The remaining participants' missing items were replaced by the participant's subscale item mean (Tabachnick & Fidell, 2007).

To investigate subscale relationships, Pearson's correlation ISS Cronbach's alpha (α) and Fisher-Bonett test (Kim & Feldt, 2008) were utilized respectively.

A series of repeated measures ANOVAs were conducted to determine equivalence of the Total scale, the five subscales and the 50 items between the two versions (ISS-R and ISS-RC).

Results

Internal Consistency and Homogeneity

The Pearson's r correlation coefficients, ranged from .58 to .95 for the English version (ISS-R) scale measures and .63 to .96 for the Chinese version (ISS-RC). The ISS-R Total scale and the 5 subscales correlations are presented in Table 2. All correlations were statistically significant at $p < .001$.

Table 2. Correlations of the Inviting School Survey (ISS-R) Total Scale and Subscales by Version

Scale	People	Program	Process	Policy	Place	Total
People	-----	.76	.85	.83	.73	.95
Program	.80	-----	.78	.63	.72	.86
Process	.88	.74	-----	.81	.68	.91
Policy	80	.67	.80	-----	.58	.86
Place	.76	.73	.75	.63	-----	.85
Total	.96	.86	.93	.85	.87	-----

Note. Upper diagonal is English version (ISS-R); Lower diagonal is Chinese version (ISS-RC)
N = 63; All correlations are significant ($p < .001$).

As depicted in Table 3 all of the ISS-R reliability coefficients, for both versions, were greater than .70 suggesting that these measures demonstrated acceptable levels of reliability (Nunnally & Bernstein, 1994).

Based on the Fisher-Bonett test of equivalence (Kim & Feldt, 2008), the paired-scale reliability coefficients (Cronbach's α) were statistically equivalent ($p > .05$) between the ISS-R and the ISS-RC.

Table 3. Inviting School Survey (ISS-R) Total Scale and Subscales Coefficient Alpha by Version

Scale	Number of Items	English Version	Chinese Version
People	16	.92	.92
Program	7	.79	.76
Process	8	.80	.81
Policy	7	.77	.79
Place	12	.84	.84
Total	50	.96	.96

Note. N = 63; No statistical significant alpha differences, based on the Fisher-Bonett test.

Reliability of the measures was also evaluated by examining inter-item correlations. As shown in Table 4, the inter-item correlation mean for each scale ranged from .31 for the Chinese version Program subscale to .44 for both versions' People subscale. All of the inter-item correlations were statistically significant ($p < .01$).

Homogeneity was examined using the Pearson's item-to-total scale correlations found in Table 4. The correlations were between the acceptable range .30 and .70, meeting the necessary criteria for internal consistency (Ferketich, 1991). All of the item-to-total correlations were found to be statistically significant ($p < .01$).

Analyses of Scales

As depicted in Table 5 the descriptive statistics for both versions, ISS-R and ISS-RC, showed that the participants scored similarly on the scales' measures. Examination of the data (skewness and kurtosis) identified that the measures were normally distributed and the variances were equivalent, thus adhering to the assumptions underlying a repeated measures ANOVA (Tabachnick & Fidell, 2007). A series of repeated measures ANOVAs were conducted to evaluate the overall mean score of the Total scale and the five subscales between the ISS versions.

Table 4. Scale Item Means, Inter-Item Correlations Means, and Item-to-Total Scale Correlations Means by Version (N = 63)

Scale	English Version	Chinese Version
Program (16 items)		
Item Mean	4.21	4.19
Inter-Item Correlation Mean	.44	.44
Item-to-Total Scale Correlation Mean	.69	.69
Program (7 items)		
Item Mean	4.18	4.31
Inter-Item Correlation Mean	.34	.31
Item-to-Total Scale Correlation Mean	.66	.64
Process (8 items)		
Item Mean	4.14	4.11
Inter-Item Correlation Mean	.34	.34
Item-to-Total Scale Correlation Mean	.64	.65
Policy (7 items)		
Item Mean	4.22	4.25
Inter-Item Correlation Mean	.35	.38
Item-to-Total Scale Correlation Mean	.66	.68
Place (12 items)		
Item Mean	4.20	4.20
Inter-Item Correlation Mean	.32	.33
Item-to-Total Scale Correlation Mean	.45	.47
Total (50 items)		
Item Mean	4.19	4.21
Inter-Item Correlation Mean	.34	.35
Item-to-Total Scale Correlation Mean	.59	.60

Table 5. Inviting School Survey (ISS-R) Total Scale and Subscales Descriptive Statistics by Version

Scale	English Version	Chinese Version
People (16-80)		
Mean	67.42	67.08
SD	6.96	7.14
Skewness	.09	.11
Kurtosis	-1.05	-1.06
Program (7-35)		
Mean	29.27	30.16
SD	3.32	2.77
Skewness	-.12	.06
Kurtosis	-.39	-.80
Process (8-40)		
Mean	33.11	32.88
SD	3.62	3.57
Skewness	.04	.23
Kurtosis	-.91	-.97
Policy (7-35)		
Mean	29.53	29.76
SD	3.29	3.17
Skewness	-.13	-.06
Kurtosis	-.77	-.79
Place (12-60)		
Mean	50.38	50.41
SD	5.20	4.92
Skewness	.00	.08
Kurtosis	-.88	-1.14
Total (50 items)		
Mean	209.71	210.29
SD	20.03	19.53
Skewness	.18	.24
Kurtosis	-1.00	-.97

Note. N = 63.

The results of the repeated measures ANOVAs, as shown in Table 6, indicated a significant overall mean difference between the two versions' subscale, Program, F (1, 62 = 14.25, $p < .001$. The strength of the difference between the two versions of the Program subscale, as assessed by partial eta-squared (η_p^2), was large, with the subscale factor, version, accounting for 19% of the variance (Cohen, 1988).

Table 6. Inviting School Survey (ISS-R) Total and Subscales Repeated Measures Analysis of Variance Summary for the English and Chinese Versions

Scale	F[a]	p	η_p^2	Power	Mean Difference	95% CI
People	1.62	.208	.03	.24	.3421	-0.196 to 0.880
Program	14.25	.000	.19	.96	-.8942	-1.368 to -0.421[b]
Process	1.75	.191	.03	.25	-.2376	-0.497 to 0.022
Policy	3.36	.072	.05	.44	.2343	-0.120 to 0.589
Place	0.01	.927	.00	.04	-.0249	-0.565 to 0.515
Total	0.52	.475	.01	.12	-.5802	-2.194 to 1.034

[a] $df = 1, 62$.
[b] Indicates that the 95% CI does not contain zero, and therefore the difference in means is significant.

Participants reported significantly higher Program subscale scores on the Chinese version (M = 30.16, SD = 7.14) than the English version (M = 29.27, SD = 2.77). The mean difference between the versions of the Program subscale was .89 (95% CI = -1.37, -.42). All other scale measures between the versions were found to be non-significant ($p > .05$).

As depicted in Table 7, when identical scores were added to scores that deviated by only one point difference, it was found that on the People subscale, 75% of the participants fell in this category; Program subscale- 67%; Process- 84%; Policy- 84% Place- 76%; and for the Total Scale- 48%.

Analyses of Scale Items

Participants' responses on each of the 50 items on the English version (ISS-R) scale were compared to their responses on the same items of the Chinese version (ISS-RC). Identical responses on the 50 items ranged from 21 (42%) to 50 (100%). However, when identical responses were added to responses that deviated by only 1 Likert-scale point, it was found that for all 50 items the range was 41

Table 7. Number of Total Scale and Subscale Score Version Differences (N = 63)

Scale	N	%	Range
People (16-80)			-4 to 10
No Difference	29	46	
1-Point Difference	18	29	
More than 1-Point Difference	16	25	
Program (7-35)			
No Difference	32	51	
1-Point Difference	10	16	
More than 1-Point Difference	21	33	
Process (8-40)			-3 to 6
No Difference	37	59	
1-Point Difference	16	25	
More than 1-Point Difference	10	16	
Policy (7-35)			-3 to 2
No Difference	37	59	
1-Point Difference	16	25	
More than 1-Point Difference	10	16	
Place (12-60)			-8 to 6
No Difference	27	43	
1-Point Difference	21	33	
More than 1-Point Difference	15	24	
Total (50 items)			-16 to 27
No Difference	23	37	
1-Point Difference	7	11	
More than 1-Point Difference	33	52	

(82%) to 50 (100%). Of the 63 participants, 60% (N = 38) responded either identically on both versions or selected the next closest response alternative on the scale while 16 (25%) responded the same for 49 items, 2 (3%) the same for 48 items, 3 (5%) the same for 47 items, and the remaining 4 (6%) participants ranged from 41 to 46 identical responses.

Table 8 reports the summary of significant item differences extracted from the series of repeated measures ANOVAs between the original 50 ISS-R items (English) and the Chinese (ISS-RC) translation of these items.

Table 8 Inviting School Survey Repeated Measures Analysis of Variance Summary for Significant Item Differences

Scale	F[a]	p	η_p^2	Power	Mean Difference	95% CI
People						
Q36	11.62	.001	.15	.90	-.1474	-0.236 to -0.059[b]
Q39	4.37	.041	.07	.54	.0794	0.005 to 0.204[b]
Program						
Q10	11.05	.001	.15	.90	-.2857	-0.458 to -0.114[b]
Q17	7.19	.009	.10	.75	-.1905	0.332 to -0.048[b]
Q23	4.54	.037	.07	.55	.1614	0.010 to 0.313[b]
Q38	6.80	.011	.10	.73	-.1429	-0.252 to -0.033[b]
Q46	9.46	.003	.13	.86	-.2460	-0.406 to -0.086[b]
Process						
Q7	5.88	.018	.09	.66	.1701	0.030 to 0.310[b]
Place						
Q8	4.89	.031	.07	.58	-.0693	-0.132 to -0.007[b]
Q16	4.20	.045	.06	.52	.1270	0.003 to 0.251[b]
Q49	9.04	.004	.13	.84	-.3333	-0.555 to -0.112[b]

[a] $df = 1, 62$.
[b] Indicates that the 95% CI does not contain zero, and therefore the difference in means is significant.

Note. Individual item statements:
Q7 . Grades are assigned by means of fair and comprehensive assessment of work and effort.
Q8. The air smells fresh in this school.
Q10. There is a wellness (health) program in this school.
Q16. The restrooms in this school are clean and properly maintained.
Q17. School programs involve out of school experience.
Q23. Good health practices are encouraged in this school.
Q36. People in this school try to stop vandalism when they see it happening.
Q38. The school sponsors extracurricular activities apart from sports.
Q39. Teachers appear to enjoy life.
Q46. Mini courses are available to students.
Q49. The lighting in this school is more than adequate.

A statistical significant difference ($p < .05$) was found between the means of 11 (22%) of the 50 items. However, application of the Bonferroni correction statistic, which would restrict the acceptable alpha level for interpretation to $p < .001$ (alpha of .05 divided by 50 tests = .001), suggests that these mean differences may have been due to chance for 9 of the 11 items. Only questions 10 and 36 were significant when taking into account the Bonferroni correction statistic ($p < .001$).

The strength of the different versions, the independent variable, accounting for the variance between the scale items, the dependent variables, as assessed by partial eta-squared (η_p^2), ranged from a medium effect of 6% (Q16) to a large effect of 15% (Q10, Q36) according to Cohen (1988). However, power estimates greater than .80 were found in only four of the questionable items (Q10, Q36, Q46, and Q49).

Discussion

Following the implementation of the translation procedures, a small-scale field testing of 63 Hong Kong participants was undertaken to determine the equivalence of the original Inviting School Survey (ISS-R) and the adapted/translated Chinese version (ISS-RC).

Examining reliability alphas, inter-item correlation means, and item-to-Total Scale correlation means revealed no significant differences suggesting that both instruments were equivalent in terms of construct validity, internal consistency, and homogeneity.

However, a few differences were found between the ISS-R and the ISS-RC measurement scales and individual items.

Analyses of the five subscales and the Total scale revealed that the Program subscale was significantly higher for the Chinese version than for the English version. There were no other significant scale differences. Even when taking into account the Bonferroni correction statistic the probability was significant ($p < .001$). Additionally, the effect size, as assessed by partial eta-squared, showed that approximately 19% of the variance between the two versions is explained. This statistical effect is quite large according to Cohen (1988).

Examining the 50 item differences between the two measurement versions, it was found that 11 items differed significantly, with five of the items coming from the Program subscale, two items from the People subscale, one item from the Process subscale, three items from the Place subscale, and no items from the Policy subscale. However, the application of the Bonferroni correction statistic, which makes the acceptable alpha more stringent, suggests that these differences may have been due to chance for all but two of the items (Q10 and Q36). Alternatively, these differences in the items between the two versions may reflect simple semantic differences or more important underlying cultural frameworks. As such, it is recommended that further translation/adaptation meetings be undertaken with the aim of identifying if any cultural or language differences exist for the particular questionable items.

Limitations

One limitation of the present study was the use of a small sample composed from one major China centre, Hong Kong. Additional research needs to be conducted with a larger and more diverse Chinese population, not only in Hong Kong but also mainland China, as these results may not be nationally representative or generalizable.

A second limitation of the present study, that may have had an effect on the results, is if respondents did not fully understand all of the words and concepts contained within the survey, particularly the English version.

Further Directions

Although the results are promising, the ISS-RC should still be considered in its initial stages and requires further investigations. Future studies should conduct analyses of the scales' psychometrics with a large diverse population to ensure generalizability. In particular, confirmatory factor analyses (Bentler, 2006) should be used to determine the equivalence of the factor structures for both instruments.

Conclusions

Healthy psychological and physical wellbeing cannot be achieved if a school environment/climate is dis-inviting. The ISS-R and the ISS-RC are able to give data-based and measurable evidence in supporting school administrators when evaluating school improvement programs. However, these instruments do not constitute a comprehensive assessment system. The ISS-R and the ISS-RC should supplement other school improvement measures such as document analyses, interviews, focus groups, etc. in order to make informed decisions regarding implementing changes within the school that will influence perceptions of the invitational qualities of the school by the relevant school community members.

References

American Psychological Association (1999). *Standards for educational and psychological testing*. Washington, DC: Author.

Bentler, P. (2006*). EQS structural equations model program manual*. Encino, CA: Multivariate Software.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Cohen, J. (2006). Social, emotional, ethical and academic education: Creating a climate for learning, participation in democracy and well-being. *Harvard Educational Review*, *76*(2), 201-237.

Cohen, J., McCabe, E., Michelli, N., & Pickeral, T. (2009). School climate: Research, policy, practice, and teacher education. *Teachers College Record*, *111*(1), 180-213.

Combs, A. (Ed.) (1962). *Perceiving, behaving, becoming: A new focus for education*. Washington, D.C.: Association for Supervision and Curriculum Development.

Combs, A., Richards, A., & Richards, F. (1976). *Perceptual psychology: A humanistic approach to the study of person*. New York, NY: Harper & Row.

Ellis, A. (1962). *Reason and emotion in psychotherapy*. New York, NY: Lyle Stuart.

Ellis, A. (1970). *The essence of rational psychotherapy*. New York, NY: Institute for Rational Living.

Fan, W., Williams, C., & Corkin, D. (2011). A multilevel analysis of student perceptions of school climate: The effect of social and academic risk factors. *Psychology in the schools*, *48*(6), 632-647.

Ferketich, S. (1991). Aspects of item analysis. *Research in Nursing and Health, 14*(2), 165-168.

Freiberg, H. (Ed.). (1999). *School climate: Measuring, improving and sustaining healthy learning environments*. Philadelphia, PA: Falmer.

Huebner, E., & Diener, C. (2008). Research on life satisfaction of children and youth: Implications for the delivery of school-related services. In M. Eid & R. Larsen (Eds.), *The science of subjective well-being* (pp. 376-392). New York, NY: Guilford.

International Test Commission (2010). *International Test Commission guidelines for translating and adapting Tests*. http://www.intestcom.org.

Jourard, S. (1968). *Disclosing man to himself*. Princeton, NJ: Van Nostrand.

Kim, S., & Feldt, L. (2008). *Journal of Educational Measurement*, *45*, 179-193.

Meichenbaum, D. (1974). *Cognitive behavior modification*. Morristown, NJ: Plenum.

Meichenbaum, D. (1977). *Cognitive behavior modification: An integrated approach*. New York, NY: Plenum.

Nunnally, J., & Bernstein, I. (1994). *Psychometric theory* (3rd ed.). New York, NY: McGraw-Hill.

Purkey, W. (1970). *Self concept and school achievement*. Englewood Cliffs, NJ: Prentice-Hall.

Purkey, W., & Fuller J. (1995). *The Inviting School Survey (ISS) user's manual*. Greensboro, NC: The University of North Carolina at Greensboro.

Purkey, W., & Novak, J. (1996). *Inviting school success: A self-concept approach to teaching and learning* (3rd ed.). Belmont, CA: Wadsworth.

Purkey, W.; Novak, J. (2008). *Fundamentals of invitational education*. Kennesaw, GA: International Alliance for Invitational Education.

Purkey, W., & Schmidt, J. (1990). *Invitational learning for counseling and development*. Ann Arbor, MI: Eric/Caps Clearinghouse: The University of Michigan.

Rogers, C. (1969). *Freedom to learn*. Columbus, OH: Merrill.

Smith, K. (2005). The Inviting School Survey-Revised (ISS-R): A survey for measuring the invitational qualities (I.Q.) of the total school climate. *Journal of Invitational Theory and Practice, 11*, 35-53.

Smith, K. (2012). *Manual for the Inviting Survey (ISS-R): A survey for measuring the invitational qualities (I.Q.) of the total school climate (2nd ed.)*. Kennesaw, GA: International Alliance for Invitational Education. (In Press)

Smith, K., & Bernard, J. (2004). The psychometric properties of the Inviting School Survey (ISS): An Australian study. *Journal of Invitational Theory and Practice, 10*, 7-25.

Sousa, V., & Rojjanasrirat, W. (2010). Translation, adaptation and validation of instruments or scales for use in cross-cultural health care research: A clear and user friendly guideline. *Journal of Evaluation in Clinical Practice*, *17*(2), 268-274.

SPSS (2009). PASW Statistics (Version 18) [Computer software]. Chicago, IL: Author.

Tabachnick, B., & Fidell, L. (2007). *Using multivariate statistics* (5th ed.). Boston, MA: Allyn & Bacon.

Zullig, K., Koopman, T., Patton, J., & Ubbes, V. (2010). School climate: Historical review, instrument development, & school assessment. *Journal of Psychoeducational Assessment*, *28*, 139-152.