The Psychometric Properties of the Invitational School Survey (ISS): An Australian Study

Kenneth H. Smith, Ph.D.
Australian Catholic University

John Barnard
EPEC Pty Ltd

This study provides psychometric data on the Inviting School Survey (Purkey & Fuller, 1995) using a rating scale analysis within the framework of the Rasch measurement philosophy (Bond & Fox, 2001; Rasch, 1980). The Inviting School Survey's factor structure and internal consistency are examined and compared with the Invitational Education Model (Purkey, 1978; Purkey & Novak. 1996; Purkey & Schimdt, 1996; Purkey & Stanley, 1991). The Invitational School Survey (ISS) is based on five areas (5 "Ps"): People, Places, Policies, Processes, and Programs. The ISS purports to assess these qualities of the global school climate. Suggestions for further development and refinement of the Inviting School Survey are presented.

Introduction

Based on perceptual psychology tenets, Invitational theory was developed to provide a model of practice to promote people to realise their potential in all areas of worthwhile endeavours.

Perceptual psychology postulates that each person creates their own reality through their perceptions of what they believe to be real (Combs, Richards, & Richards, 1988; Combs & Gonzales, 1994; Kelly, 1955, 1963; Jourard, 1971). Furthermore, a person's behaviour is contingent on how an individual perceives and interprets his/her experiences. Purkey and Novak (1996) identified three assumptions of the percep-

tual psychology approach that are relevant to the present study:

1. Behaviour is based on perceptions. Individuals behave according to their subjective perception of the environment (internal and external).
2. Perceptions are learned. One's interpretation of the environment is learned and therefore can be unlearned given new information and new experiences. This particular assumption embraces the idea that a change in perception will bring about a change in behaviour.
3. Perceptions can be reflected upon. Being aware of one's past and present perception and being able to go beyond them allows for further development and understanding of oneself, others, and the world.

From the perspective of perceptual psychology, it is clear that to understand an individual's behaviour we need to know how that individual perceives and interprets his/her life experiences. An individual's personal interpretation or frame of reference is more important than "objective reality" because an individual responds to their perception of reality and not to reality itself (Purkey & Schmidt, 1996; Seligman, 1991).

Invitational theory focuses on five environmental areas (5 "Ps") that support or hinder an individual's success or failure. The areas identified involve people, places, policies, programs, and processes (Purkey & Lehr, 1996; Purkey & Novak, 1996).

"People" assesses respect, caring, and the honouring of diversity and refers to the positive or "inviting" influence of the teachers and support staff in the school. "Places" relates to the physical aspects of the school. "Policies" refers to the procedures, codes, and rules (written and unwritten), used to regulate the ongoing functions of individuals and organisations. "Programs" refers to the curriculum for students to de-

8

velop both academically, physically, and socially inviting environment. "Processes" refers to such issues as cooperative spirit, democratic activities, values, and attitudes of students, teachers, administrators, and support staff.

Aims

The Inviting School Survey (Purkey & Fuller, 1995) was developed to assess how teachers, administrators, support staff, and students perceive the level of "invitation" in schools across the five areas (5 "Ps") identified in the Invitational Education paradigm. However, there has been limited investigation in the instrument's psychometric integrity and properties.

The main aim of this study was to address this void by providing psychometric data on the Inviting School Survey using a rating scale analysis within the framework of the Rasch measurement philosophy (Bond & Fox, 2001; Rasch, 1980).

If the instrument is to be useful to administrators and counsellors who work in the school setting the present instrument needs to be reduced in number of items without loosing too much of its psychometric properties. An instrument of 50-items would be efficient and yet still effective in assessing the global school environment. Presently, the 100-item instrument takes too much time for students, particularly at the lower grade levels to complete.

As such the goal of this study was to produce the rating scale that yields the highest quality measure for the constructs under study. In order to achieve this goal the following tasks were undertaken:
1. Analysis of participants' responses.
2. Verification of the five subscales and the total measure empirically.

9

3. Calculate statistics that provide guidance in assessing how the response categories and statements are functioning.
4. Create an interpretable measure.

Method

Participants

A total of 539 students from 3 Catholic secondary schools (Years 8-12) completed the 100-item instrument. Two schools were single-sex (males) from the Melbourne metropolitan area while the third school was a co-ed school from a major Victoria rural area.

Of the 539 instruments completed, 46 (8.5%) had 10 or more missing data or were inappropriately completed and were eliminated from further analysis.

The final sample consisted of 434 male students and 59 female students (493 total participants) with a mean age of 16.03 years (Range = 13yrs – 19yrs; SD = 1.25).

Instrument

The Inviting School Survey (ISS) is a self-report 100-item instrument based on the Invitational Education model (Purkey & Fuller, 1995). All items pertain to school environment and were designed to assess the qualities of the total school environment in the areas of people, places, policies, programs, and processes (Refer to Appendix 1). Respondents report their agreement with each item using a five-point response (Likert) scale where 1 = Strongly Agree, 2 = Agree, 3 = Undecided, 4 = Disagree, and 5 = Strongly Disagree.

The ISS consists of five subscales representing the degree to which schools are inviting in the five environmental areas as outlined in Invitational Education theory: 30 people items (e.g. "The principal involves students in the decision-making process"), 10 program items (e.g. "There is a student health program in this school"); 20 process items (e.g. "Grades are assigned by means of fair and comprehensive assessment of work and effort"); 20 policy items (e.g. "Few, if any, students fail in this school"); 20 place items (e.g. "Soap and towels are available in student toilets").

There are 27 items that are negatively stated such that a high score reflects a negative invitation (9 policy items, 1 program item, 6 process items, 7 people items; 4 place items).

For this study, American words were replaced with Australian words (e.g. 'cafeteria' changed to 'tuck shop', 'restrooms' changed to 'toilets', etc.).

Procedure

Participants voluntarily completed the instrument during their regularly scheduled homeroom program and assured of anonymity. Between 30-45 minutes was required to complete the 100-item instrument.

Response sheets were reviewed and screened for completeness, and answer sheets that had more than 10 missing items or were inappropriately completed (n = 46, 8.5%) were eliminated from further analyses. Missing values were replaced with the mean for that particular domain.

Data was "cleaned" and all items (n = 27) that were phrased in the opposite direction were reversed coded so that the codes allocated had the same meaning.

Rasch Measurement Model[1]

11

Often Likert-scale[2] data is analysed by only allocating "scores" to certain responses and then adding the statement scores to yield an overall "score." For example, on a 5-point scale a 1 is given to a "strongly disagree" response; a 2 to a "disagree"; a 3 to a "neutral"; a 4 to an "agree" and a 5 to a "strongly agree" response. A person's "score" is then computed by adding the individual statement scores. If 20 statements are given, a maximum score of 100 can be achieved and the closer the score is to 100, the more the person generally agrees with whatever is being measured. However, by merely adding the statement scores at least two assumptions are made, namely (i) that the differences between the response categories are constant and the same, and (ii) that each statement contributes equally to the construct being measured. In other words it is assumed that the difference between Strongly Disagree and Disagree is the same as the difference between Disagree and Neutral, and so on. It is further assumed that all the categories have the same meaning over all the statements. Although counts and means are useful statistics for "eyeball inspection", they are less useful for further examination of the underlying structure of the scale. Such counts and standard methods disregard the subjective nature of the data.

Modern Test Theory[3] models in general and Rasch measurement in particular do not make these assumptions. In Rasch models the probability of a person of a certain "ability" achieving a certain score on an item of a certain difficulty is estimated from the data instead of calculating the difficulty of an item as the proportion of people who answered the item correct and expressing a person's performance in terms of how many items were answered correctly. Rasch models thus propose a relationship between person ability and item difficulty and express the relationship as the probability of a certain response. The more able the person, the higher the person's chance to answer the item correctly. If it is known how a person has performed on other items, an estimate of

12

his/her ability can be obtained, and if it is known how other persons have performed on an item, an estimate of how difficult the item is can be obtained. The chances of a correct response are therefore a function of the difference between the person's ability and the difficulty of the item.

It should be noted that Rasch measures are traditionally expressed as "person abilities" and "item difficulties" on an interval scale—opposed to raw scores on an ordinal scale. A logit is the unit of measurement that results when the Rasch model is used to transform raw scores to log odds ratios on the logit scale[4]. The value of 0.00 logits is usually allocated to the mean of the item difficulty estimates and typically estimated values vary between -3 and 3 logits where negative values indicate estimates below the mean and positive values indicate estimates above the mean. An ability or difficulty measure is obtained by converting a raw score percentage into odds of success. For example, a raw score of 30% correct converts to –0.85 logits and a raw score of 80% correct converts to 1.39 logits.

In the simple Rasch model where questions (items) are scored dichotomously (either right or wrong) ability estimates of persons and difficulties estimates of items are usually derived from the analysis. However, when Likert scales are used and the items are statements, a more complex model such as a Rating Scale model is commonly used. Statements are not scored as correct or incorrect, and therefore the term "ability" will mean "agreeability" in this context, i.e. a higher estimate would indicate more in agreement with the statements. A person with a high agreeability estimate is more likely to endorse the statements than a person with a lower agreeability estimate.

Also, in addition to establishing the position of each item (statement) on the scale, the pattern of the categories in the items is also established—where categories are labels such as "strongly disagree", "agree", etc. The relative position of

13

each category of each item is estimated. The "difficulty" of a statement gives an indication of how difficult it is to endorse the statement. But, since each statement has different categories, the likelihood of choosing one option rather than another should also be explored. The points where the likelihood changes from the one category to the next are called thresholds. Since there are 5 categories in the statements, there are 4 thresholds for each statement.

Associated with estimates are errors that indicate the amount of imprecision associated with each estimate. In addition to the error, fit statistics are also usually computed. Where an aberrant response pattern is identified for a particular item across all the persons in the group, this may indicate that the item is flawed in some way, or that it does not tap the same ability as the others in the set, or, in certain systematic inconsistencies in the responses of identifiable subgroups are observed, that the item is biased. Likewise, a lack of fit for an individual person indicates that the model is an inappropriate means of describing the behaviour of the student on that set of items. Where most persons have responded largely in accordance with the model's expectations, misfit of an individual person can be attributed to anomalous test taking behaviour of some kind. Whatever the underlying cause, a response vector which is inconsistent with an otherwise well-fitting model may indicate that the test, though possibly functioning well for the group as a whole, has failed to provide an appropriate measure of the relevant ability for that particular person.

Four fit statistics were produced for each estimate, namely INFIT Mean Square, OUTFIT Mean Square, INFIT t and OUTFIT t. An estimate of infit is obtained by weighing up the residuals (discrepancies between predicted and observed data) near the central point to the residuals at the extremes which are weighted down by using the expression for the variance as the weighting variable. The outfit statistic is summed over persons, which are assumed to be independent

14

estimates. Since it examines the residuals for an item across persons, and is more sensitive to outliers, the infit statistic is usually preferred. Mean square values have an expected value of 1 and individual values above or below this show greater variation (values above 1) or less variation (values less than 1) than might normally be expected. As a rule of thumb values in the range 0.75 to 1.30 are considered acceptable, but for larger samples the range is usually computed in terms of the standard deviation. Fit values in terms of the t-distribution will vary around a mean of zero and will be positive (if observed values show greater variation) or negative (if observed values show less variation). Values outside the range –2 to 2 are said to indicate significant departure from the expectations of the model.

Results

Participant Analysis

Following Rasch analysis, 24 of the 493 participants were identified as misfitting (Infit Mean Squares of 1.9 or more). These participants were not involved in any further analysis leaving 469 final total participants. Details of the participants are found in Table 1.

Item Analysis

One of the aims of the study was to determine whether the 100-item instrument could be shortened without compromising its psychometric properties. The focus of the item analysis was therefore to identify misfitting items in sequential calibrations, remove the identified item(s) and repeat the computations. The infit mean square statistic was used as the criterion for uni-dimensionality and to investigate whether the subgroups of items "hang together" which is also a check of validity.

- In the first round analysis Question 67 was deleted as a result of non-fit statistics.

15

- In the second round analysis Questions 12, 24, 34, 45, 72, and 80 were then removed.

Table 1

<u>Participants' age by gender by year level</u>

Year	<u>N</u>	Mean	SD	Min	Max
Year 8					
Males	14	13.57	.51	13	14
Females	21	13.62	.50	13	14
Year 9					
Males	9	14.33	.50	14	15
Females	12	14.33	.49	14	15
Year 10					
Males	146	15.27	.46	14	16
Females	10	15.50	.53	15	16
Year 11					
Males	102	16.30	.46	16	17
Females	7	16.29	.49	16	17
Year 12					
Males	139	17.39	.52	17	19
Females	9	17.67	.50	17	18
Total					
Males	410	16.16	1.16	13	19
Females	59	15.02	1.54	13	18
Total	469	16.02	1.27	13	19

- In the third round analysis Questions 18 and 38 were removed.
- In the fourth round analysis Questions 15, 48, and 70 were removed.
- In the fifth round analysis Questions 6, 9, 27, 29, 41, 54, 56, 59, 61, 77, and 82 were removed. Question 47 also had a fit outside the acceptable range but was left in because of the small number of items in the program subscale.
- In the final round of analysis Questions 1, 8, 10, 20, 23, 32, 47, 52, 65, 66, 68, 73, 78, 79, 88, 97, and 99 were deleted as a result of non-fit statistics. Question 74 should have been deleted but was retained to ensure at least 7 items in the Program subscale was achieved.

The final instrument had 60 items with at least 7 items per subscale (see Tables 2 and 3 for detailed descriptive statistics). The overall reliability of 0.92 for the 100-item instrument was reduced to 0.88 for the 60-item instrument.

Discussion

These results have shown that reducing the present 100-item ISS to 60 items does not compromise its reliability significantly. In fact it may be advisable to reduce the ISS to 50-items since the 'People' subscale still has 20 items. One could take the best 10 'People' items and still have a very reliable instrument.

It is important to have an instrument in which all items are performing adequately. However, previous research has show that the domain, Programs, is problematic (Shoffner & Vacc, 1999). That is, this particular domain may be subsumed under the other four domains. This study has also shown that few 'Programs' items have strong psychometric properties. It is suggested that more specific reliable items need to be writ-

ten so as to have a comprehensive instrument to measure this specific subscale area.

Regarding the other domains, it is suggested that no more than 10 items be allocated to each domain. In order for the ISS to be used more extensively it is highly recommended that no more than 50 items be used. As shown by this study, reducing the number of items has not reduced the reliability of the instrument significantly. A shorter version of the ISS would lend itself to be used more often by school personnel to assess school culture as perceived by the major stake-holders: students, teachers, parents, and administrators.

Table 2
Invitational School Survey (ISS) Chronbach's
Coefficient Alphas

Number of Items	People	Places	Programs	Polices	Processes	Total
100	.81	.71	.54	.61	.68	.92
	30 Items	20 Items	10 Items	20 Items	20 Items	100 Items
60	.77	.66	.48	.50	.48	.88
	20 Items	13 Items	7 Items	11 Items	9 Items	60 Items

n = 469

In order to gain a comprehensive understanding of the psy-chometrics of the ISS it is suggested that further research, using the Rasch measurement model (Bond a& Fox, 2001) and classical test measurement principles (e.g. Factor Analy-sis, Structural Equation Modelling), be undertaken. Additionally,

future psychometric studies of the Inviting School Survey need to examine the stability of the instrument across age, gender, country, and other school environment demographics.

Table 3
Invitational School Survey (ISS) Means and
Standard Deviations

Number of Items	People	Places	Programs	Polices	Processes	Total
Mean 100	2.93	3.02	2.62	2.93	2.95	2.92
SD	.42	.43	.45	.36	.38	.35
Mean 60	3.06	3.11	2.69	2.92	2.91	2.98
SD	.47	.50	.50	.43	.45	.38

n = 469

Notes:
1. *Rasch measurement* is a philosophy of measurement in which a person's probability of responding correctly to a question is expressed as a function of the ability of the person and the difficulty of the question. In more advanced models the probability that a person will agree with a statement can be estimated – see Likert-scale.
2. A Likert-scale is a rating scale consisting of two parts, namely a declarative statement and a list of response categories, usually ranging from "Strongly Agree" to "Strongly Disagree". Such scales are popular in attitude questionnaires, etc.
3. *Modern Test Theory* is a generic term used to distinguish Rash measurement and Item Response Theory from Classical Test Theory (e.g. Factor Analysis). Models of Modern Test Theory are probabilistic rather than deterministic.
4. Just like metres is a unit for measuring distances or degree Celsius is a unit for measuring temperature.

References

Bond, T., & Fox, C. (2001). *Applying the Rasch model: Fundamental measurement in the human sciences.* Mahwah, NJ: Erlbaum.

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

Purkey, W. (1978). *Inviting school success: A self-concept approach to teaching and learning.* Belmont, CA: Wadsworth.

Purkey, W. & Lehr, J. (1996) *Creating SAFE schools through invitational education.* Unpublished manuscript, The University of North Carolina at Greensboro: The Alliance for Invitational Education, Greensboro, NC.

Purkey, W. & Stanley, P. (1991). *Invitational teaching, learning and living.* Washington, DC: National Educational Association Professional Library, National Education Association.

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

Purkey, W., & Schmidt, J. (1996). *Invitational counseling: A self-concept approach to professional practice.* Pacific Grove, CA: Brooks/Cole.

Rasch, G. (1980). *Probabilistic models for some intelligence and attainment tests.* Chicago: University of Chicago Press.

Shoffner, M. & Vacc, N. (1999). Psychometric analysis of the inviting school safety survey. *Measurement and Evaluation in Counseling and Development, 32,* 66-72.

SPSS for Windows (Version 11) [Computer software]. (2001). Chicago, IL: SPSS.

Appendix 1: Invitational School Survey (ISS)

* Denotes negatively worded statement. ~~denotes deleted item.~~

1. ~~Rules in this school are fairly administered.~~
*2. Teachers are unwilling to help students who have special problems.
3. People in this school have ample time to go to the bathroom.
*4. Furniture is unpleasant and uncomfortable.
5. Everyone is encouraged to participate in athletic programs.
6. ~~School policy provides for assistance for those students who need it.~~
7. Students work cooperatively with one another.
8. ~~Teachers express appreciation for students' presence in their classes.~~
9. ~~Custodians take pride in keeping the school as clean as possible.~~
10. ~~Special efforts are made to recognize the cultural contributions of minority groups.~~
11. The principal involves everyone in the decision-making process.
12. ~~Soap and towels are available in student restrooms.~~
13. Everyone in this school takes responsibility for keeping it clean.
14. The air smells fresh in this school.
*15. ~~Bathroom time is strictly scheduled into the school day.~~
*16. Teachers in this school show a lack of respect for students.
17. Few, if any, students fail in this school.
*18. ~~Tardiness is a problem in this school.~~
19. Students have the opportunity to talk to one another during class activities.
20. ~~Students are pleased when they are called upon.~~
*21. Teachers are difficult to talk with.

22

22. School policy permits and encourages freedom of expression of students, faculty, parents and administrators.
23. People in this school laugh a lot.
*24.   Observations indicate that space is cluttered and otherwise misused.
25. The school grounds are clean and well-maintained.
26. People in this school find ways to serve the surrounding community.
27. There are many living green plants inside this school.
*28.   Teachers take little or no time to talk with students about their out-of-class activities.
29. Teachers and principals work cooperatively in this school.
30. Teachers are generally prepared for class.
31. The restrooms in this school are clean and properly maintained.
32. Students like to visit the school library.
33. Teachers exhibit a sense of humor.
34. The lunch program at this school is a pleasant addition to the school day.
35. Grades are assigned by means of fair and comprehensive assessment of work and effort.
36. There is a school wellness program in this school.
*37.   People in this school are impolite to one another.
38. The library is open before and after school.
39. The principal's or headmaster's office is attractive.
40. Teachers work to encourage students' self-confidence.
41. Teachers expect high academic performance from students.
42. Signs posted in and around this school are positively worded.
43. School programs involve out of school experience.
44. Bulletin boards are attractive and up-to-date.
*45.   The cafeteria food is unappetizing.
*46.   Trash is left on school buses.
47. Provisions are made for students of varying needs.

48. Everyone in this school has a say in deciding school rules.
49. All telephone calls to this school are answered promptly and politely.
50. The principal treats people as though they are responsible.
51. Everyone arrives on time for school.
52. Creative thinking is encouraged in this school.
53. Space is available for student independent study.
54. Student discipline is approached from a positive standpoint.
55. Fire alarm instructions are well posted and seem reasonable.
56. Music is played in gym classes during indoor exercise periods.
57. The messages and notes sent home are positive.
*58.   Teachers show insensitivity to the feelings of students.
59. Teachers discuss planning and student process in teams.
60. Students work cooperatively with each other.
61. Teachers maintain clear and reasonable work standards.
62. Classrooms offer a variety of furniture arrangements.
63. People in this school want to be here.
*64.   People often feel unwelcome when they enter the school facility.
*65.   Communicating directly with this school is a difficult and time-consuming task.
*66.   Much of this school's correspondence is negative in tone.
*67.   Corporal punishment is used to punish students.
*68.   Parents feel they are not welcome in this school.
69. People in this school try to stop vandalism when they see it happening.
70. Salad bar/salad/fresh fruit choices are available in the cafeteria.
71. Clocks and water fountains are in good repair.

*72.   The school intercom (P.A. System) interrupts classroom learning.
*73.   The cafeteria is an unpleasant place to eat lunch.
74. Good health practices are encouraged in this school.
*75.   A high percentage of students fail in this school.
76. Teachers appear to enjoy life.
77. The school administrators show a strong interest in making this school inviting.
78. Teachers use a variety of methods to help students learn.
*79.   Teachers demonstrate a lack of enthusiasm about their work.
80. The principal of this school knows the names of many students.
81. Interruptions to classroom academic activities are kept to a minimum in this school.
82. People in this school succeed in doing what is expected of them.
83. School pride is evident among students.
84. Teachers share out-of-class experiences with students.
*85.   This school's policy provides for guidance in academic matters and athletic activities only.
*86.   Only a select few in this school are involved in making decisions.
87. Daily attendance by students, staff and faculty is high.
88. Grass, evergreens, shrubs around the school are well-kept.
89. There are comfortable chairs for visitors.
90. Nutritious and health-promoting refreshments are served at school meetings.
91. Teachers spend time after school with those who need extra help.
92. The lighting in this school is more than adequate.
*93.   People are ignored when they enter offices in this school.
94. Classes get started quickly.

95. The school sponsors extracurricular activities beyond sports.
96. Mini courses are available to students.
97. ~~People in this school feel free to disagree with one another.~~
*98.   School buses sometimes leave without waiting for students.
*99.   ~~People are discouraged from beginning new projects in this school.~~
*100. The grading practices in this school are unfair.

Kenneth H, Smith, PhD, MAPS, Head of School, Trescothick School of Education (Victoria) and Faculty of Education, Australian Catholic University, 115 Victoria Parade, Fitzroy, Victoria, Australia, 3065. Correspondence about this article may be sent to: k.smith@patrick.acu.edu.au