International Education Journal

Volume 1 Number 2
July 2000

Published by

Shannon Research Press
Adelaide, South Australia
ISSN 1443-1475
http://iej.cjb.net

International Education Journal

Volume 1, Number 2, July 2000

Welcome to the International Education Journal. This is a broadly based journal encompassing research and review articles. Education is interpreted in a wide manner and includes human development, learning, school education, formal and informal education, tertiary and vocational education, industry training and lifelong learning.

Online - While there will be a hard copy version of the journal it is intended that it will be accessed in almost all instances via the web.

Peer Review - All articles appearing in this journal have been subjected to a peer-review process.

Published by:
Shannon Research Press,
13 Endeavour Court,
North Haven, SA 5018,
Australia

ISSN 1443-1475

Designed by Katherine Dix

Printed in Adelaide
December 2000

CONTENTS

FREE Online Subscription
Subscription is available as a free service in online format only. Either complete the subscription form at: http://iej.cjb.net or email the onlineeditor@iej.cjb.net. The reader will be notified of the latest issue and contents.

Publication Frequency
It is intended that the journal be published three times per year. However given the electronic characteristics of online publishing it is not necessary to wait until a full journal is assembled prior to publishing. Articles once accepted in appropriate form may be published without delay. It should be noted however that in electronic format pagination may vary according to the parameters of the downloading computer. Alternatively the journal article can be downloaded in PDF format, in which case the original pagination will be preserved.

Aims of IEJ

The aim of the International Education Journal is to publish articles that possess one or more of the following characteristics or qualities.

1. The presentation of substantive findings which are of importance for policy and practice beyond the country in which the research was conducted.
2. The integration of education with academic disciplines such as anthropology, demography, economics, history, law, linguistics, philosophy, political science, psychology and sociology, or examines educational issues from the perspective of the disciplines or investigates issues at the interface between education and one or more of these disciplines.
3. The examination of educational issues from a cross-cultural or indigenous people's perspective.
4. The evaluation of educational policy or programs or the use of information technology of cross-national interest and significance.
5. The employment of advanced research methods and measurement procedures which are clearly explained.
6. The presentation of empirically or analytically based investigations of theory, models or conceptual frame works in the field of education.
7. The syntheses of research findings from comparative and cross-national studies in education.

Invitation

Contributors - Authors are invited to submit material to this journal. As a general guide articles should be not more than 5000 words. The journal may publish longer works such as theses or disserations as occassional mongraphs. In every instance material must be received in publisher ready format. Full details of publication style and other requirements are set out under the heading Author Information.

Publishers - If you wish to have books or educational software reviewed and published in the International Education Journal please contact the editors.

Reviewers - If you are interested in contributing to the journal by writing a review article (500-1000 words) please contact the authors indicating your areas of interest. We look forward to hearing from you.

Author Information

IEJ welcomes practical and research manuscripts that focus on educational issues, provide a clear purpose and depth of discussion, and are presented in a straight-forward style accessible to an international audience.

The initial submission:

- should be prepared using the downloadable Microsoft WORD97 template document (iejbase.doc), complete with layout styles, images and tables correctly positioned in the text
- images that are imported or inserted into the text from an external source as figures should be included as separate graphic files (either GIF or JPG format)
- should be forwarded electronically, either by:

  e-mail to: onlineeditor@iej.cjb.net

  or,

  on a 1.4MB 3.5 inch floppy diskette to: IEJ Online Editor, 13 Endeavour Court, North Haven, SA 5018.

The editors reserve the right to return poorly edited or improperly formatted manuscripts. Hard-copy only submissions cannot be accepted or returned.

The manuscript should:

- show the article's title, authors' names, affiliation, and email addresses (these will be included on the online page), and postal address and telephone numbers (these will be removed from the final online page)
- abstract should not exeed 150 words
- paper should not exceed 5,000 words (20,000 for occasional papers)
- include an abstract of about150 words
- include 5 keywords describing article
- be in final form ready for reviewing
- conform to the APA Reference System

General Principles for layout:

- Papers should be prepared using the template *iejbase.doc* with Microsoft WORD on Mac or PC
- Authors with WORD 98 or later and a basic understanding of web page construction should also save their paper in web format (*.*html*) incorporating any hypertext features.
- Use end-notes, with one set of numbers running through the whole article.
- Use italics (not underline) for emphasis.
- Do not create additional headers and footers in the template document *iejbase.doc*, just modify appropriately.
- Do not use page numbers in text (web pages do not follow conventional page numbering systems).

If you have any questions regarding layout and presentation, please contact the IEJ Online Editor by email on: onlineeditor@iej.cjb.net

FUIIE

The publishers have arranged for the Flinders University Institute of International Education to produce and host an electronic mirror site containing all material in the Journal, accessed free of charge. There are no restrictions on downloading or printing single copies. The electronic version also provides readers with a search tool.

The Course Experience Questionnaire as an Institutional Performance Indicator

David D Curtis
Flinders University, School of Education *david.curtis@flinders.edu.au*

John P Keeves
Flinders University, School of Education *john.keeves@flinders.edu.au*

Data from the 1996 Course Experience Questionnaire (CEQ) were analysed using the Rasch measurement model. This analysis indicates that 17 of the 25 CEQ items fit a unitary scale that measures course quality as perceived by graduates. Graduates are located on the interval measurement scale produced in the Rasch analysis. The interval nature of the scale renders the graduates' scores amenable to analyses that are not wisely employed using ordered raw CEQ scores. Analysis of variance indicates that variations in graduates' responses are attributable to field of study and institutional factors. In order to compare universities, corrections are made for the course mix of each institution to produce expected institutional scores. These are compared with observed institutional scores to determine those universities that have performed above, at, or below expectation. (Individual institutions are not identified in this analysis).

Important issues relating to the educational and statistical significance of the findings have emerged. The data collected through the CEQ do not represent a simple random sample of all graduates. Instead, the data model is a hierarchical one, with individual graduates nested within courses, which are nested within institutions. This requires analysis using multilevel analytical tools. Conventional analyses substantially underestimate the standard errors of aggregated measures (such as institutional means) and therefore report institutional differences as significant when they are not. The implications of the measurement and analytical problems for policy decisions over the distribution of funding among institutions and among courses within institutions are discussed.

The Course Experience Questionnaire

The Course Experience Questionnaire (CEQ) is a survey instrument of 25 items that is posted to recent graduates of all Australian universities. It seeks to establish graduates' perceptions of the quality of the courses that they have completed and its results are used to compare courses and institutions. Twenty-four items are statements representing views about five main aspects (clear goals, good teaching, appropriate assessment, generic skills, and appropriate workload) of the courses that graduates have recently completed. There is also a single summary item. Graduates are asked to express their opinions about their courses by indicating the extent of their agreement with these propositions by selecting one of the five response options from *strongly disagree* to *strongly agree.*

In 1996, the CEQ was distributed to 137,603 graduates of whom 93,967 replied (Johnson, 1997). For the purposes of calibration, we used only those responses from graduates of undergraduate programs, and then we used only those who had completed all 25 CEQ items. This left us with the returns of 51,631 individuals.

We are concerned that the methods commonly employed for the analysis of CEQ data are less than optimal. In the methods of analysis being used for the CEQ data, each item response option is coded as: *strongly disagree*, -100; *disagree*, -50; *neutral*, 0; *agree*, +50; and *strongly agree*, +100. These coded response options are subject to conventional statistical analyses with means and standard deviations being computed for items and for sub-sets of items and are the basis of comparisons between courses. An assumption implicit in these forms of analysis is that the data are interval and that a particular response option indicates the same level on the underlying trait for all items. Our concern is that graduates' responses are ordinal and therefore should be analysed differently. The current analyses produce useful information and the large numbers of cases involved in the survey probably mean that alternative analyses may not produce a substantially different picture. However, if policy decisions are to be based on the results of the survey or if the survey instrument is to be modified over time while still permitting comparisons, we contend that the alternative and superior analytical techniques that are available should be used.

The short-comings of present analytical methods have been identified (Karmel, Aungles, & Andrews, 1998). They employed an unnecessarily complex method of analysis to generate indicators of institutional performance, but they were obliged to use the raw ordinal data of the CEQ. Below, we demonstrate that the Rasch measurement model produces an interval measure that can be manipulated much more readily. Locating items on an interval scale permits items to be substituted over time and thus permits comparisons over time. We go on to show that CEQ data are only adequately represented using a hierarchical data model and that this should be acknowledged in analyses of the data.

Approaches to measurement

Measurement is conducted to describe phenomena with a greater degree of precision than mere indicators can provide and to facilitate comparisons over time or between cases. If the CEQ is to be used as a ***measure*** of perceived course quality, and if that measure of quality is to be compared over time or across institutions or courses, then it must acknowledge the assumptions that underpin measurement and conform to accepted criteria for measurement. Weiss and Yoes (1991, pp 72-74) identify four requirements made of measures. They are:

- if a respondent holds a certain attitude, s/he will respond honestly to an item which taps that attitude;
- the choices that respondents make among response options indicate the strength of the underlying attitude trait that they hold;
- the responses that participants make to particular items are not influenced by the presence of other items in the instrument (local independence);
- the pattern of responses to items will conform to a probability function.

The first three of these requirements have long been characteristic of measurement. However, the fourth requirement reflects the need to understand the basis of survey measurement and provides a foundation for estimating the precision and reliability of the measures that are derived from such observation.

Wright and Masters (1981) relate seven criteria for true measurement. They are:

- each item should function as intended;
- each item can be positioned on a common scale;
- the scale should be an interval one;
- each person can be located along the same common scale used for items;
- the responses should form a valid response pattern for each item;
- estimates of precision must be available for all scale measures; and
- each item should retain its meaning and function across individuals and groups.

While these requirements and criteria apply to all genuine measurement, in the Rasch measurement model, these requirements have particular salience. In most forms of measurement it is difficult to be sure that all items contribute to a common scale. Some of the methods for ensuring this lack sensitivity, and this criterion has not been as rigorously enforced as it might have been. The use of a common scale is important since it may be necessary to use alternative items in parallel versions of an instrument, especially if it must be administered to a group on several occasions, and different items must be comparable on that measurement scale. It is also desirable that the scale be at least an interval one so that differences among subjects can be compared meaningfully and change over time validly assessed.

The Rasch measurement model meets these criteria for measurement. It requires a particular response probability function form that for each item depends upon only two variables: the strength of the individual's affect on the trait being measured and the trait threshold required to accept that item. This gives the Rasch model some unique measurement characteristics (Wright & Masters, 1982). Items that are influenced significantly by external factors will not fit the function. Those that do fit reveal a threshold (or series of thresholds for ordered response items) and they can be used to place individuals on the scale for the underlying trait.

Analyses of the 1996 CEQ

Both exploratory and confirmatory factor analyses have been undertaken to examine the factor structure of the CEQ items. Data have also been analysed using the Rasch measurement model. This has enabled us to identify items that fit a common measurement scale, compute item thresholds, and compute scale scores for individuals on the underlying trait.

We do not report detailed results of these analyses since they are reported in detail in Curtis (1999) and they are similar to those reported in Waugh (1998). In summary, from the exploratory factor analysis, we found that there is evidence for five factors that do correspond with the five sub-scales identified in the CEQ. The confirmatory factor analysis indicates that there is a single underlying factor and that the five separate factors are nested within it. We take this to indicate that there is an underlying 'perception of course quality' factor and that it is expressed to varying extents in the five components that have been identified as elements of course quality.

Rasch analysis indicates that eight of the 25 items misfit the measurement model, but that the 17 remaining items do form a coherent measure. The items that were used in the 1996 CEQ are shown in Appendix 1. Table 1 shows fitting and non-fitting items by their CEQ sub-scale. This suggests that the overall perception of quality is most strongly influenced by three of the five factors – good teaching, generic skills, and clear goals.

Table 1: The CEQ sub-scale structure and Rasch item fits

	GTS	GSS	CGS	AAS	AWS	OA
Fitting	7	2				
items	15	5	1			
	17	10	6			
	18	11	13	12		
	20	22	24	19	14	
Non fitting	3	9		8	4	25
items				16	21	
					23	

From the Rasch calibration, individuals were assigned a score on the scale formed for the underlying trait from the 17 fitting items. The interval scale that results from Rasch scaling is measured in logits. These units result from the probability function used to model responses. As

an interval scale it has an arbitrary origin (like commonly used temperature scales) but, since it has units that arise from the logistic model, it can be linearly transformed into a scale with a more obvious meaning. We have transformed the scale to a mean of 500 with each logit being re-scaled to 100 units. This scale has been tentatively called the ***Graduate Satisfaction Index*** (GSI) and we have used this as the basis of further analyses.

Institutional comparisons

Since the GSI is an interval scale, we could compare institutions by simply taking the mean GSI of all graduates for each institution. However, institutions are far from homogeneous and vary on several dimensions: some have a high proportion of graduates from technology and applied science courses, others have a greater proportion of graduates from the humanities and social sciences. In addition, it seems that institutions differ on individual characteristics such as age and gender of graduates. If, as it seems from existing analyses of CEQ data (Johnson, 1997), there are differences on CEQ scores among course types, such differences should be taken into account before any attempt is made to compare institutions on CEQ scores. This was the primary purpose of the analyses undertaken by Karmel at al. (1998).

Influence of course type

In order to ascertain whether there are differences among course types, we undertook a two way analysis of variance using individual GSI scores as the criterion measure and both course type and institution as categorical variables. We classified courses into nine broad fields of study: Agricultural Sciences; Architecture; Humanities and Social Sciences; Business; Education; Engineering; Medical Sciences; Law; and Mathematics and Science. Karmel et al. (1998) used ten categories in their analysis. They separated Veterinary Sciences which we included under Agricultural Sciences. We are quite sure that within these categories there are substantial differences between individual courses. For example, within the Medical Sciences we expect differences to be apparent between MB, BS awards, the various nursing awards, and other courses such as medical radiations, physiotherapy, and speech pathology. Our purpose was to test whether there were differences among broad course types. For other purposes a finer grained analysis would be warranted.

The analysis of variance indicated strong main effects for broad field of study and institution and some interactions. In Table 2 the national mean GSI score for each of the nine broad fields of study are shown.

Table 2: National GSI means by Broad Field of Study

	AgSci	Arch	HSS	Bus	Educ	Eng	Med	Law	MaSci
N	766	1004	12196	10332	6172	3257	8145	1894	7725
Mean	511.50	480.90	527.06	485.71	507.08	476.14	483.09	493.80	502.30
St Dev	89.26	82.34	104.02	75.11	92.87	73.21	81.28	89.84	83.70

Institutional GSI scores

Since there are differences among broad fields of study, it is apparent that simply taking the mean GSI for all graduates of an institution would bias the institutional score in favour of those universities with high proportions of humanities graduates and against those with high proportions of engineering graduates. Thus we employed a method of correcting for the course mix of institutions in developing an institutional GSI score. We took the national average GSI score for each broad field of study and the proportion of graduates in that broad field of study to produce a

weighted expected mean GSI for each institution. This is the expected institutional mean GSI score, assuming that its courses are performing at the national average. We then computed the actual GSI mean for each institution as the simple mean GSI for all graduates and tabulated the difference between actual and expected GSI means. These data are shown in rank order in Table 3. Note that we have used the 57 separate institutional codes from the raw CEQ data and that individual institutions are not identified. The data are also shown graphically in Figure 1.

Table 3: Actual – Expected mean GSI scores for institutions in rank order

Rank	Actual-Expected GSI	Rank	Actual-Expected GSI	Rank	Actual-Expected GSI	Rank	Actual-Expected GSI
1	124.13	16	12.67	31	0.14	46	-9.17
2	74.73	17	9.43	32	0.02	47	-9.86
3	54.77	18	8.44	33	-0.62	48	-10.46
4	50.08	19	8.20	34	-1.13	49	-13.26
5	30.05	20	7.83	35	-3.48	50	-14.09
6	24.16	21	7.71	36	-3.91	51	-14.91
7	23.67	22	6.55	37	-4.19	52	-15.09
8	23.25	23	6.06	38	-4.47	53	-16.43
9	22.19	24	5.88	39	-5.34	54	-17.43
10	16.33	25	4.28	40	-5.54	55	-17.62
11	15.61	26	4.11	41	-6.61	56	-32.95
12	15.12	27	2.91	42	-6.68	57	-46.53
13	14.48	28	1.56	43	-7.09		
14	14.21	29	0.94	44	-7.29		
15	12.96	30	0.78	45	-7.63		

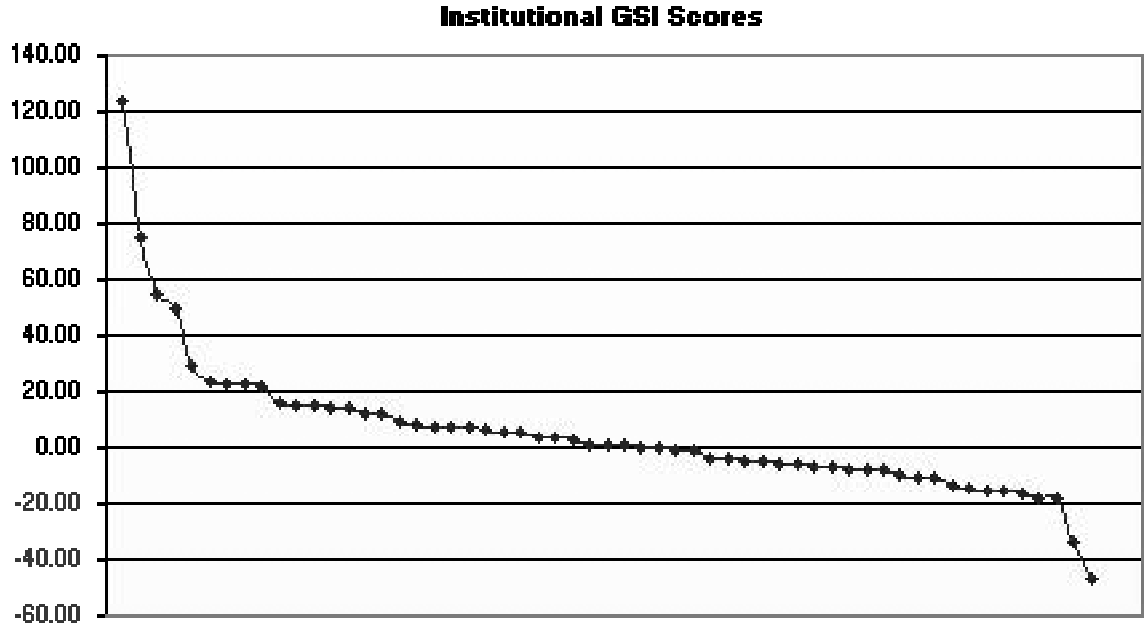


Figure 1: Differences between Actual and Expected mean GSI scores for institutions

Categories of institutions

On the basis of the differences between expected and actual mean GSI scores, we could classify institutions as performing above, at, or below expectation. A question arises about the points at which it is possible to discriminate the top and bottom groups from the middle one. An obvious technique is to calculate a confidence interval about the mean using its standard error. This would yield a band about 5 units on either side of the mean with 14 institutions in that band, leaving 24

in the top group and 19 in the bottom group. However, we have two concerns about this approach. The first is that the sample that we have is about 50 per cent of the 1996 population, and under these circumstances we should apply a finite population correction to the estimate of the precision. In this case, the confidence interval would be reduced to 2.5 units either side of zero. There is, however, scope for debate about whether the 1996 graduate cohort is a unique population or is a subset of the population of all graduates from Australian universities over time. To some extent this matter is only resolved by establishing the purposes of the analyses. If comparisons are to be made over time to ascertain whether there is an improvement in graduates' perceptions of the quality of university teaching, there is a case for regarding each annual cohort as a discrete population. If we are interested only in validating a measure of graduates' perceptions of quality, then there is a case for regarding the 1996 group as one intact convenience sample of a much larger population.

Of greater concern are sampling errors. We do not have a genuine random sample from a homogeneous population: instead we have an intact sample from a stratified heterogeneous population of graduates from different courses and from different institutions. We calculated the design effect at 3.6, which results in a confidence interval of between 10 and 20 units either side of zero, depending upon whether one chooses to apply a finite population correction. With a confidence interval of 20 units, nine institutions would lie in the above expectation group and two below. The two low performing institutions are small specialised institutions whose graduates are placed within the Humanities and Social Sciences group where the expected GSI score is very high. Their specialised courses may well be ones that are rated at the low end of the range, but are disadvantaged by being compared with all other Humanities and Social Sciences courses.

The sampling problems referred to above present a problem that must be addressed. The stratified nature of the sample however suggests a solution to the problem of comparing unlike institutions. We now turn to that matter.

The hierarchical nature of the population

The graduates who completed the CEQ differ on individual characteristics such as age and sex, have undertaken courses which also differ in their characteristics, and have graduated from institutions which have distinct histories and missions. It is worth noting, as Meyler (1997) did, that students' judgements of subjects are influenced by the size of the subject, whether the subject is compulsory, and whether the subject is quantitative. Since the current CEQ items invite aggregated judgements about subjects, rather than the totality of graduates' course experiences, we might expect to see substantial differences in the ratings of different types of courses. The stratified nature of the population suggests that we have a sample of graduates that should be considered to have three levels – the individual, the course, and the institution. Given that this is the case, the forms of analysis that are reported above and that have been used by most others who have researched this area are not appropriate and that analytical tools that do recognise the hierarchical nature of the sample should be used. One such tool is HLM (Hierarchical Linear Modelling) (Bryk, Raudenbush, & Congdon, 1996).

We undertook a series of analyses to see if we could identify differences between institutions when individual and course level variables are separated. In order to make these analyses tractable, we confined our attention to the three South Australian universities and used a subset of data of graduates from those institutions. Under this three level model, it is argued that the score of any individual is the result of an institutional component, a course component, an individual component, and error terms that account for unexplained variation at each level.

At the level of individuals we used sex, age, non-English speaking background (NESB) status, employment status, and mode of study as explanatory variables. The regression equation for this relationship is shown below.

$$GSI = P0 + P1.Sex + P2.Age + P3.Nesb + P4.Emp + P5.Mode + E$$

That is, an individual's GSI score can be understood as depending upon his or her own characteristics of sex, age, NESB status, employment status, and mode of study, and that there is an intercept term (P0) that is a result of variables which operate at the course and institutional levels.

At course and institutional levels we did not use explanatory variables. We included only proxy categorical (dummy) variables to separate the influences of the different broad fields of study and the different universities. In estimating the parameters in the following equations, with data from N sources, N-1 parameters can be estimated so one parameter (for a course type or institution) must be omitted from the estimation. (In the equations below, the omitted parameters are shown in brackets). However, an individual must be a member of one, and only one, category so the omitted parameter must be the complement of the sum of the estimated ones. The course level regression equation was therefore:

$$P0 = B00 + B01.AgSci + B02.Arch + B03.HSS + B04.Bus + B05.Educ + B06.Eng + B07.Eng + B08.Law\ [+ B09.MaSci] + R0$$

Thus the intercept term used in the first level equation (P0) is the result of the particular course type that the individual completed, an error term to represent unexplained variance (R0), and an intercept term (B00) that reflects variation at the third or institutional level.

For the institutional level, the regression equation was:

$$B00 = G000 + G001.Flin + G002.Adel\ [+ G003.UniSA] + U0$$

It should be noted that the three levels of the model are related through intercept terms. At the individual level, there is an intercept P0, and it is the criterion variable of the course level equation. Its intercept term, B00, is the criterion variable in the third level equation. In that equation, the parameters of interest to us are the coefficients of the categorical variables for each of the institutions. Those parameters tell us about the relative standings of the three institutions when the hierarchical nature of the sample is modelled and when variables at the individual and course levels are taken into account. Table 4 shows a summary of the results of the hierarchical analyses completed using HLM.

From the hierarchical analyses, we found that mode of study was not significant and has been dropped from the model. Sex was only marginally significant, but we have chosen to leave it in the model as it assists in explaining some of the features that emerge from the analyses. To estimate the GSI of any individual the separate regression equations with their estimated parameters can be combined. For any individual, the institutional score is taken and added to it is the broad field of study score, and then the individual characteristic variable scores.

$$GSI = Inst + BFStud + 3.81\ Sex + 0.92\ Age - 6.43\ Nesb - 8.73\ Emp + E + R0 + U0$$

It is instructive to compare course and institutional means found from multilevel analysis with those found from earlier methods.

Table 4: Summary results of the hierarchical analysis of data from the three South Australian universities

Fixed Effect	Coefficient	Standard Error	T-ratio	Sig
Level 3 effects				
FLIN	484.25	13.04	4.58	**
ADEL	472.72	5.56	8.66	**
UNISA	461.14	5.12	7.15	**
Level 2 effects				
AGSCI, B01	-0.50	14.58	-0.03	
ARCH, B02	-3.11	15.89	-0.20	
HSS, B03	26.60	11.84	2.25	**
BUS, B04	-14.32	12.17	-1.18	
EDUC, B05	25.68	12.82	2.00	*
ENG, B06	-18.53	13.50	-1.37	
MED, B07	-9.79	12.17	-0.80	
LAW, B08	0.50	14.58	0.03	
MASCI, B09	6.88	12.04	0.57	
Level 1 effects				
SEX, P1	4.77	3.29	1.45	
AGE, P2	1.15	0.18	6.35	**
NESB, P3	-8.04	3.99	-2.02	**
UNEMP, P4	-10.91	5.35	-2.04	**

An '*' in the Significance column indicates p<0.10 while '**' indicates p<0.05

Course type performances

When raw means of graduates from each of the nine broad fields of study are computed, no allowance is made for the characteristics of graduates from those courses. For example, graduates of engineering courses are younger than those of education awards and more of them are males. In the multilevel analysis, it is apparent that younger graduates make harsher judgements of course quality than do older ones, and males tend to make harsher judgements than do females. Table 5 shows the deviation of the raw mean from 500 (the overall mean of all graduates) and the course intercept from the multilevel analysis. While for some course types there is very little difference between the two measures of perceived course quality, for others eg, Architecture there are substantial differences. Architecture graduates are predominantly male, younger than other graduates, and experience greater difficulty in finding employment. Each of these factors is associated with significantly lower judgements of course quality. By not separating these factors, Architecture courses are perceived to rate poorly by comparison with others. For Education graduates, there is a substantial difference between the raw score deviation and the HLM intercept. This is attributed to the low representation of NESB persons and the difficulty gradates experience in finding satisfactory employment as many find only part time and short term contract work. However, when the influence of individuals' characteristics is removed, the influence of the type of course on graduates' perceptions of course quality is shown to be little different from the overall mean. We argue that when institutions are comparing course types with each other, it would be more sensible to use a measure that has greater meaning and that has extracted from it influences other than those due to the course itself.

Table 5: A comparison of Broad Field of Study means (expressed as deviations from the overall mean) with intercepts from multilevel analysis

	AgSci	Arch	HSS	Bus	Educ	Eng	Med	Law	MaSci
Dev from o'all mean	11.50	-19.10	27.06	-14.29	7.08	-23.86	-16.91	-6.20	2.30
HLM dev	-0.50	-3.11	26.60	-14.32	25.68	-18.53	-9.79	0.50	6.88

Institutional performance measures

It has been argued earlier in this paper that institutions can be compared using graduates' raw GSI scores, but we have argued that failing to correct for course mix biases the measure. In the multilevel analysis, we have found that there are individual graduate characteristics that influence their judgements about courses. In order to unpack the results of the multilevel analysis, it is instructive to examine the measures that are available. These measures are presented in Table 6.

Table 6: A comparison of alternative measures of institutional performance derived from the GSI

Institution	Raw mean GSI	Expected GSI	Difference (raw-expected)	Multilevel intercept
Flinders	526.86	503.61	23.25	484.25
Adelaide	507.91	503.63	4.28	472.72
University of SA	502.10	496.04	6.06	461.14

First, it should be noted that under the analyses described above, the three South Australian universities perform at or above the national average. Indeed Flinders University performs well above it. Both the University of Adelaide and the University of South Australia perform slightly better than expected, but not significantly so. On the measure corrected for course mix, Flinders performs 19 points ahead of Adelaide. However, when corrected for individual graduate characteristics, its lead over Adelaide is reduced to about 12 points. This is because Flinders graduates are almost six years older than Adelaide graduates, it has a greater proportion of women graduates, and a lower proportion of NESB graduates. When Adelaide graduates' characteristics are considered, it has a lead of 11 points over the University of South Australia.

Summary and conclusion

We have shown that using the Rasch measurement model, it is possible to identify items that fit a coherent scale and to convert the ordinal ratings of graduates on CEQ items to an interval measure of perceived course quality. This measure has been re-scaled to produce what we have called the *Graduate Satisfaction Index* (GSI). We have shown that, while it is possible to use raw GSI mean to compare institutions, this produces biased ratings because of differences in the ratings of different types of courses and because of differences in institutional course profiles. We have also shown that it is possible to correct for the influence of course type to generate a more satisfactory measure of institutional performance. However, it is clear that the problem of measuring course quality is a multilevel one and that it is necessary to examine factors in a multilevel model. In doing this, we have found that there are individual graduate characteristics that influence the judgements made about the courses that graduates have just completed. These characteristics vary among courses as well as institutions, and following multilevel analysis, we have shown that the influence of individual characteristics can be separated to develop better comparative measures of both courses and institutions, and we have done this for the three South Australian universities.

Multilevel analysis has permitted influences of variables that were previously confounded to be disaggregated. For example, in earlier studies (eg Johnson, 1997), it was reported that employment status at the time of completing the CEQ did not influence graduates' perceptions of their courses. By separating effects at individual and course levels, we have been able to show that employment status is significant. In the past, its influence has been masked by course type because of different rates of graduate employment from different courses.

Multilevel analysis has also enabled reliable estimates of institutional effects to be established. It is desirable that institutions now consider their relative positions and begin to explore factors that may explain these estimates. It is quite possible that decisions made within institutions on the

allocation of funds to libraries and other student services or that expenditures on teaching and research activities influence graduates' judgments of their courses. Johnson and Keeves (2000) have begun to do just these forms of analysis and are able to show that these and other decisions do influence graduates' perceptions.

We do not suggest that such detailed analyses should be routinely undertaken. However, if analyses like these are done, we could identify salient factors (and their parameters) at each of the three levels of the model and they could be used to correct 'raw' course and institutional measures of course quality as perceived by graduates. If significant policy decisions such as resource allocations are to be based upon instruments like the CEQ, we suggest that better analytical techniques such as those that have been reported in this study should be employed so that those policy decisions are more soundly based.

References

Bryk, A., Raudenbush, S., & Congdon, R. (1996). HLM for Windows. Hierarchical linear and nonlinear modeling with HLM/2L and HLM/3L. (Version 4) [Multilevel analysis software]. Chicago: Scientific Software International.

Curtis, D. D. (1999). *The 1996 Course Experience Questionnaire: A Re-Analysis.* Unpublished Ed. D. dissertation, The Flinders University of South Australia, Adelaide.

Johnson, T. (1997). *The 1996 Course Experience Questionnaire: a report prepared for the Graduate Careers Council of Australia* . Parkville: Graduate Careers Council of Australia.

Johnson, T. G., & Keeves, J. P. (2000). Spending on the selling of wisdom. *Issues in Educational Research, 10*(1). (In press)

Karmel, T., Aungles, P., & Andrews, L. (1998, 30 September). *Presentation of the Course Experience Questionnaire (CEQ).* Paper presented at the Course Experience Questionnaire Symposium 1998, University of New South Wales.

Meyler, M. (1997, 8 -11 July). *What do SET surveys really measure? And why does it matter?* Paper presented at the HERDSA'97 Advancing International Perspectives conference, Adelaide.

Waugh, R. F. (1998). The Course Experience Questionnaire: a Rasch measurement model analysis. *Higher Education Research and Development, 17*(1), 45-64.

Weiss, D. J., & Yoes, M. E. (1991). Item response theory. In R. K. Hambleton & J. N. Zaal (Eds.), *Advances in educational and psychological testing: theory and applications* (pp. 69-95). Boston: Kluwer Academic Publishers.

Wright, B. D., & Masters, G. (1981). *The measurement of knowledge and attitude* (Research Memorandum 30). Chicago: University of Chicago, Department of Education, Statistical Laboratory.

Wright, B. D., & Masters, G. N. (1982). *Rating scale analysis*. Chicago: MESA Press.

Appendix 1: Items used in the 1996 CEQ

Item	Sub-scale	Item statement
1	CGS	It was always easy to know the standard of work expected.
2	GSS	The course developed my problem solving skills.
3	GTS	The teaching staff of this course motivated me to do my best work.
4 *	AWS	The workload was too heavy.
5	GSS	The course sharpened my analytic skills.
6	CGS	I usually had a clear idea of where I was going and what was expected of me in this course.
7	GTS	The staff put a lot of time into commenting on my work.
8 *	AAS	To do well in this course all you really needed was a good memory.
9	GSS	The course helped me develop my ability to work as a team member.
10	GSS	As a result of my course, I feel confident about tackling unfamiliar problems.
11	GSS	The course improved my skills in written communication.
12 *	AAS	The staff seemed more interested in testing what I had memorised than what I had understood.
13	CGS	It was often hard to discover what was expected of me in this course.
14	AWS	I was generally given enough time to understand the things I had to learn.
15	GTS	The staff made a real effort to understand difficulties I might be having with my work.
16 *	AAS	Feedback on my work was usually provided only as marks or grades.
17	GTS	The teaching staff normally gave me helpful feedback on how I was going.
18	GTS	My lecturers were extremely good at explaining things.
19 *	AAS	Too many staff asked me questions just about facts.
20	GTS	The teaching staff worked hard to make their subjects interesting.
21 *	AWS	There was a lot of pressure on me to do well in this course.
22	GSS	My course helped me to develop the ability to plan my own work.
23 *	AWS	The sheer volume of work to be got through in this course meant it couldn't all be thoroughly comprehended.
24	CGS	The staff made it clear right from the start what they expected from students.
25		Overall, I was satisfied with the quality of this course.

* indicates a reversed item.

Distance No Longer a Barrier: Using the internet as a survey tool in educational research

Katherine L Dix
Flinders University, School of Education *katherine.dix@flinders.edu.au*

Jonathan Anderson
Flinders University, School of Education *jonathan.anderson@flinders.edu.au*

The existence of the World Wide Web clearly provides new horizons for educational research. In particular, one aspect still in its infancy, is the use of the Web to access individuals as research subjects, which is emerging as a major, new research tool. This paper discusses the processes involved in developing online surveys and how these may be administered to participants in a research study however dispersed they may be, provided there is access to the Internet. Although there are distinct advantages in using the Internet as the interface between researcher and researched, there are difficulties too, and these are considered as a practical guide to undertaking data collection via the Web.

Introduction

At one time or another, we have all been the survey-taker, filling out long and often complex paper surveys to provide information to the surveyor. As educational researchers, we also understand the time and resources that are necessary to process survey data. Developments in Web technology now offer an opportunity to reduce many complications normally associated with administering surveys. Such opportunities, however, are often counter balanced by concerns arising, in part, from the researcher's lack of understanding of this new research tool and its impact on the task of data collection in educational research.

Literature in the field of online research predominantly addresses the commercial market in terms of customer surveys, and the Internet market in terms of user questionnaires. Each targeted audience requires the development of its own unique online environment to maximise survey effectiveness. Developing online surveys for use in educational research is no exception. Although the art of designing an effective instrument is critical, it is not considered here. Instead, this paper addresses the promises, together with the pitfalls, of developing and conducting online surveys for educational research, based upon the authors' experience in conducting surveys where participants are often situated in remote locations. In particular, this paper details the form survey that uses hypertext markup language (HTML) and common-gateway-interface (CGI) programming to construct, format, and administer surveys to participants, wherever they might be located online. Within this domain, general issues of methodology are offered as a practical guide to undertaking data collection via the Web.

Potential versus current reality

Current literature suggests that web-based surveys are not the solution for every research project, even though the technological capabilities exist (Farmer 1998). There will always be a need for interviews, for telephone surveys, and for traditional pencil and paper surveys in situations where

Web technology is not appropriate or available. If a self-administered survey is appropriate, then its adoption in cyber form in preference to the paper variety will depend largely upon the respondents. Clearly, respondents need to have easy access to the Internet, both in terms of hardware and knowledge. The implementation of Government initiatives like the Learning Technologies Project (DETE 1999) in South Australia and similar projects in other States (Education Queensland 1999; Education Victoria 1998; Education Western Australia 1998; NSW DET 1997) provide equipment and training support, reducing the access void and increasing the viability of online research in the education sector.

Beyond the basic issue of access, the decision to adopt an online format requires additional criteria to be considered. These are most effectively discussed by addressing four questions.

1. What can be done that couldn't be done before?
2. What can be done better than before?
3. What no longer needs to be done?
4. What are the hazards?

What can be done that couldn't be done before?

One of the key benefits of an online environment is time. Instantaneous electronic distribution of survey materials and subsequent electronic return of completed surveys give the fastest possible opportunity for responses to be collected and automatically compiled into a database. Time losses incurred from survey distribution by hand or mail and the manual entry of responses to form a database are no longer necessary.

An added benefit of an online survey allows responses to questionnaires to be pre-coded as words, symbols or numbers. For example, if a survey employs a three-point Likert scale of disagree, uncertain, and agree, the transmitted data may be coded as 1, 2, or 3 respectively. Further, if certain items are negatively worded to avoid response bias, then responses to similar three-point Likert scales can be reverse scored (3, 2, or 1). Any inaccuracies resulting from manual data entry are thus removed.

Yet another benefit is the increased degree of flexibility afforded in design and presentation. Respondent-friendly online surveys can be developed for targeted audiences by using varying backgrounds, colourful graphics, and effective layouts, thereby heightening interest and increasing subjects' motivation to complete a survey.

What can be done better than before?

A factor significant in any form of research is cost. For a given sample size, an online survey is the least expensive research methodology and constitutes one of the most attractive aspects to going online. Costs typically incurred in traditional paper surveys include outgoing and return postage, stationary materials like paper, envelopes, and printing, and the expense of manual data entry, all of which increase with larger sample sizes. By contrast, the minimal one-off cost to employ a programmer (if you choose not to do it yourself) is independent of sample size and can be an insignificant cost in large studies.

An additional improvement, particularly evident in complex surveys, is the facility of the Internet to be adaptive. Web-languages like Java, VBScript and ActiveX make adaptive surveys possible where subjects can be directed to particular items according to how they have responded to previous items. In a paper-based environment, instructions directing participants, on the basis of certain prior responses, to skip to particular items complicate the questionnaire process and can cause confusion and frustration. Very complex surveys often require interviewers to guide the

participant. However, in cyberspace the transition from one item to another non-adjacent item is seamless, with the result that participants are not even aware that particular items have been skipped.

What no longer needs to be done?

Removing the apparent complexity of some surveys by using interactive and adaptive techniques may remove the need and cost of an interviewer. The flexibility offered by advances in the language of the Web means that complex surveys, requiring one-to-one interviewing, can be simplified to a point suitable for self-administration.

One of the more onerous and time-consuming tasks associated with survey-based research is the manual entering of questionnaire responses into data files for statistical analysis by computer. An online environment virtually removes the need for such data entry and, more importantly, removes any possibility of typographical errors occurring due to lapses in concentration or fatigue.

What are the hazards?

In addition to the positive aspects associated with online surveys, there are a number of weaknesses to consider. First, who is answering the survey? As with all unsupervised mail-out studies, it is difficult to ensure that the desired person actually answers the survey. The procedure of notifying potential participants by email is as reliable a process as mail. Providing the survey site is not promoted beyond the targeted audience, the chance of random Internet surfers finding an online survey site is minimal. Besides, site access can be restricted by employing identification numbers and password protection.

In contrast to unwanted access, a second disadvantage is one of access limitation. Quite simply, not everyone has access to Internet technology. However, of further concern is the possibility of inadvertent biasing of a study by selecting participants (even if they are a random sample) on the basis of Internet access.

A third and frustrating concern regards the fickle nature of technology. The Internet as a technology is far from perfect and continues to be under intensive development. If systems are overloaded or resource conflicts arise, computers and servers can sometimes 'crash', resulting in a loss of data, perhaps without detection. Consultation of server log-files is one measure of site traffic, independent of the responses received. The difference between the number of visits and the number of responses gained provides some indication of data loss. Small discrepancies can be expected due to participants viewing the survey but completing it at a later date. However, large differences should not be ignored and may require an additional email to respondents fielding any access difficulties.

Finally, an issue common to all new research methodologies, is their divergence from the mainstream. Since Internet surveys have developed over quite a short period and are still in a phase of intensive development, the existence of comprehensive research in this field is relatively scarce (Pilypas 1997). With more experience and comparative analysis of online surveys versus paper-based versions, online surveys should reach the same level of acceptance as other methodologies (DSS Research 1998).

Reflection upon these four questions would suggest that as the Internet improves both in terms of access and reliability, and as the methodology of online research becomes better known, the choice of conducting surveys online may be increasingly viable. Related to the adoption of Internet technology is the further decision regarding the type of online format.

Online alternatives

Although this paper is primarily concerned with standard HTML form surveys, a brief introduction is given to a range of alternative methods. Discussed in order of increasing complexity, these include email surveys, standard HTML form surveys, adaptive HTML surveys, and downloadable interactive survey applications.

Email surveys

Of the various online alternatives, email surveys are often the fastest and simplest, offering wide audience-access (anyone with an email account) and requiring relatively little set-up time (Bowers 1998). The availability of off-the-shelf email survey software packages supports even novice users in constructing and conducting email surveys. As with any email message, however, the survey form is limited to simple text format with minimal control over page layout and interactivity (Sheehan and Hoy 1999).

Standard HTML form surveys

HTML form surveys offer more flexibility than do email surveys, but require substantially more time to design and construct. In essence, the former are just web pages with a form area containing questions, response fields, and a submit button. As with many educational surveys, the standard HTML survey requires participants to complete the same survey answering all questions in a linear style. Since no complex programming is necessary, adapting a paper-based survey is only marginally more difficult than creating a document in a word processor, thus making the technique accessible to most educational researchers.

Adaptive HTML surveys

The adaptive questionnaire extends from the standard HTML form by utilising skip routines and branching. Adaptive or inferential questions are based upon responses to previously asked questions (Pitkow and Recker 1994). Accordingly, participants are not asked to respond to all possible questions but proceed through the survey by skipping sections according to responses given to previous items. In a paper-based environment, this is highly visible, but in cyberspace the transition is seamless. Understandably, the complexity of design and programming places adaptive HTML surveys in the realms of professional programmers with greater associated expenses.

Interactive survey applications

Offering a similar level of flexibility and complexity as adaptive HTML surveys, the downloadable survey 'application' incorporates the questionnaire into an executable file that respondents download to their own computer (Bowers 1998). The construction, flow, and content of the program are limited only by the creativity of the survey researcher. The resulting survey should be highly interesting and enjoyable. The main disadvantages, however, include the cost and time required for sophisticated programming, distribution and access to the downloadable file and respondents' concerns both in their ability to download files and the fear of such files containing viruses.

For educational research purposes, of these four alternatives, the standard Web page survey supports current educational survey design and offers the greatest flexibility for minimal programming and time requirements.

Conducting a standard online survey

Many function specific survey-building packages are available for purchase from the Internet or from commercial software suppliers. Alternatively, there are numerous online survey providers, some of which offer limited but free survey hosting and data analysis, such as *workwork.com, interscore.com, or formsite.com* (Ullman 1999). In educational research, however, these options are usually quite expensive or too narrow and restrictive to meet particular purposes. Creating a standard online survey using a general web-page-maker minimises costs, yet allows the researcher complete design and administrative control. Although some level of technical competence is necessary, the process of conducting a standard online survey is not difficult.

There are five basic stages to conducting a Web survey, each of which is addressed in detail below.
1. Planning and preparing
2. Going online
3. Testing
4. Promoting
5. Collecting the data.

Planning and preparing

Plan the survey by first creating a paper-based prototype, complete with questions, possible responses, scoring of the responses, and general layout. A word processor such as *Microsoft Word* serves this purpose well with the additional capability of being able to save the file in HTML format (available in *Word97* or later). The subsequent file *(*.html)* can then be edited in Word or with a general web-page-maker such as *Claris HomePage* by placing the survey questions within a form. Simply, a form area is the cyber equivalent of an envelope, detailing where and how the information is to be sent. The basic range of form components is presented in Figure 1.

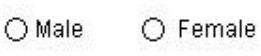
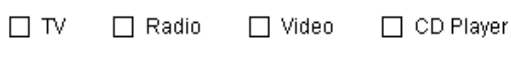
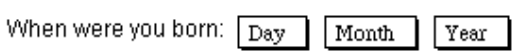
Icon	Name	Description	Example
	Radio button	Radio buttons allow your audience to select exactly one choice from a list of options. Only one radio button in a list can be selected at a time.	
	Check box	You can add a single check box to a form or multiple check boxes to a list of choices in a form. Check boxes allow the viewer to select as many choices as they like from the list.	
	Drop-Down menu	Drop-down or Pop-up menus and scrolling lists allow your audience to choose from a list of items. A drop-down menu shows only one item in the list until someone clicks on it. A scrolling list shows more than one item, and your audience can scroll to see the remaining items in the list.	
	Text Field	Text fields can contain only one line of text, can be 1 to 500 characters wide, and can only be resized horizontally. These fields are generally used for shorter information such as a name or email address.	
	Text Area	Text areas allow your audience to enter multiple lines of text, which they can scroll through using the scrollbars.	
	Submit button	You need to include a submit button for your audience to send their information. Each time the submit-button is pushed, the information filled in the form area is sent to the CGI script located on the server. Until the information is sent participants can alter any of their responses.	

Figure 1. Form components used in an Online survey together with examples

Although many design and methodological issues concerning paper-based surveys apply to the online equivalent, there are additional elements that ensure an effective and respondent-friendly online survey. As stated above, respondent-friendly design refers to the construction of online questionnaires that:

> increases the likelihood that sampled individuals will respond to the survey request, and that they will do so accurately, i.e., by answering each question in the manner intended by the surveyor. (Dillman, Tortora and Bowker 1998b:3)

Accordingly, the survey should have an equal chance of being accessed by each participant (regardless of their level of computer skill or sophistication of hardware and software) with all items presented in a way that can be easily understood and answered.

To ensure the construction of a respondent-friendly survey, some useful guidelines should be followed.

An introduction page. Introduce the online survey with a welcome screen that is motivational, emphasizes the ease of responding, and instructs the participant on the general actions necessary to complete the survey. The page may also request an identification field such as a name or ID number.

Ensuring equivalent access. Dramatic advances in software and hardware allow increasingly complex web surveys to be designed. Although it may be tempting to add fancy qualities that the viewer will enjoy, research comparing plain and fancy survey design comes down in favour of plain design (Dillman, Tortora, Conradt and Bowker 1998a). Most importantly, graphic-rich and audio-rich designs require greater transmission time, greater browser power, as well as increasing the chances of overloading and crashing participants' computers. Such features should be kept to a minimum and only used to assist understanding. Secondly, surveys using an advanced language such as Java will not function on all browsers thus limiting accessibility, possibly by as much as 50 per cent of the intended audience (Spain 1998). Applying these findings within the limits of what the computer, browser, and transmission line to respondents can manage, results in a simple accessible survey that is quick to download and utilises less memory.

Page Format. Currently, the two common formats for Online Surveys are 'single page' and 'one question per page'. Ultimately, the choice of layout depends on the structure of the survey, whether it is linear or branching, and the maturity of the target audience. For example, the one question per page format may be highly suitable for younger children as it offers fewer distractions but, by the same token, may be inappropriate for older respondents since it requires more page-turning.

Traditional principles of layout. The recommended approach is to present each item in a conventional format similar to that normally used in traditional paper-based surveys. Clearly, surveys that list questions without numbers, centre questions on the screen, have insufficient spacing around questions, or place yes and no response buttons on opposing sides of the question, are unconventional and challenging for participants to complete. By contrast, a survey that provides questions that are numbered and justified in a well-spaced table, with the placement of the response buttons to the right hand side of the screen closest to the scroll bar (thus minimising mouse movement and control), will appear logical and be easier to complete.

The first question. Begin the main survey with a question that is fully visible on the first screen, and is likely to be easily comprehended and answered by all or most participants. The first

question and the process of responding tend to define for the participant whether the questionnaire is easy or difficult to complete.

Question specific user instructions. Provide item specific instruction as part of each item where the action is to be taken, rather than in a separate section at the beginning of the survey. From a cognitive perspective, there is considerable likelihood that detailed instructions placed at the beginning of a survey will be forgotten by the time each action is taken.

Avoid required responses. Unlike paper-based surveys where questions can generally be answered in any order or not at all, online surveys can require participants to answer each question before being able to proceed to the next. Although this has been promoted as an advantage, its use should generally be avoided for survey questions. Circumstances may arise where a respondent has legitimate reasons for objecting to answering a question, or may be unable to provide an answer. Far better that only one question is missed than that a whole survey is not completed. An exception to this general principle regards certain identification fields. Items such as name, age and gender may be crucial to subsequent statistical analysis of the data, and accordingly should require a response (paper-based surveys are disadvantaged in this respect). Such requirements are achieved by reminding participants that particular items (e.g. name or gender) have been omitted.

More choices than size of screen permits. When the number of response choices exceeds the size limits of the screen, consider using double or triple columns to display the choices. Avoid forcing the participant to scroll down to see the alternative choices, at the same time removing the question from easy view.

Bigger is not better. There is nothing worse than having to scroll across a screen to read a question. Placing questions in a table that is formatted to span 90 per cent (for example) of the viewer's screen provides a simple yet effective solution to the wide screen viewing problem.

Progress monitor. Use simple graphic symbols or words that indicate the progress of the participant through a survey. Estimating progress through a paper-based survey is far easier than judging completion progress in an equivalent online survey. Therefore, it is necessary to include on each new screen a progress monitor that may be as simple as a 'percentage completed' message or star rating system.

Going online

Once the design of the survey appears to be complete, it can be uploaded to the server where it will have its own unique Web or URL address allowing the survey to be accessed by the research participants. How this occurs will vary depending on server security and software used to deliver the survey form.

However, the server may contain server-specific software and server-sided (CGI format) form handling software along with a folder containing the entire survey (see Figure 2). The location of the folder on the server will determine the Web address of the survey. Accessing the address from another computer using a browser will call up the specified survey on the server. When the questions are completed and the submit button is pressed, the survey data are uploaded to the server. Form handling software collects the data and saves these to a file, usually including details about the submitting computer and the date and time at which it was sent – if all goes according to plan.

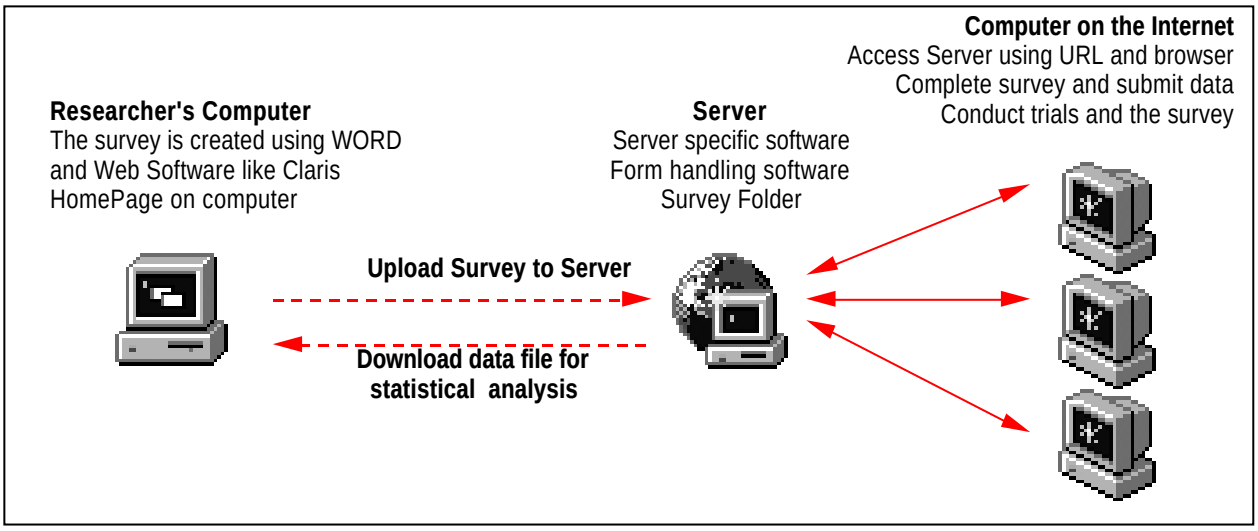


Figure 2. Representation of the online survey process

Testing the survey

Prior to conducting a survey, issues regarding usability, functionality, access and stability need to be addressed (Gould, Gurevich and Pagerey 1998). The survey should look appropriate for its anticipated audience and should function properly on the computer on which it was created. Unfortunately though, when accessed on a different platform, using a different browser, the layout frequently alters due to different screen widths, among other things. The resulting survey may therefore appear quite different making it difficult to read and complete. Overcoming variable layout usually requires a tightening of specifications, such that a table width set on automatic may need to be set to a percentage or pixel width.

Since familiarity breeds acceptance, the survey requires scrutiny and testing by others. Trials should be conducted on a test group composed of participants who are representative of the sample group. Little is gained, for example, if only colleagues trial a survey designed for young school students.

Promoting the survey

Promoting the survey is generally achieved by informing the participants in a study by email. The content of the email should include a brief introductory statement about the survey and its purpose. Next, the location of the survey is required, by providing the site URL address. In many email browsers the address becomes a hypertext link from which the participant can directly access the site.

In addition to the survey address, the email sent to participants should include the researcher's name and contact details, along with an invitation for comment or support. Although surveys are not new, conducting them online is new. As a consequence, it is vital that the participant feels confident and able to complete the survey. For most, the novelty factor alone is sufficient (Zukerberg, Nichols and Tedesco 1999). However, for some technophobes, merely providing an email address does not ensure their participation, as in the case of an English speaking person unable to complete a survey written in Chinese – they will need additional support.

Collecting and analysing the data

The survey is online, participants have been notified, they complete the form, and submit their responses, but where do the data go? The data are sent to a file located with the survey on the

server (the computer that makes the survey available to participants over the web). In addition to the responses, the data file contains records of the time, date and each participant's computer location. Depending on the platform (Mac or PC) and the software used, the data may be saved in a variety of formats, the most common and portable of which is a tab-delimited text file. The subsequent text file can be viewed using a word processor or imported into a statistics package, like *SPSS* or *Excel*, ready for immediate analysis.

Some Sample Surveys

To clarify some of the features of online survey forms described above, this section contains samples from current online surveys with which the authors are involved. Figure 3 illustrates the use of radio buttons in a language and an information technology survey respectively, where respondents are asked to select one choice from a list of options.

Figure 3. The use of radio buttons in items from an online survey of language and computer knowledge respectively

Check boxes are commonly used in surveys where respondents are invited to make one or more responses to particular items. Figure 4 illustrates the use of check boxes in an online mathematics survey.

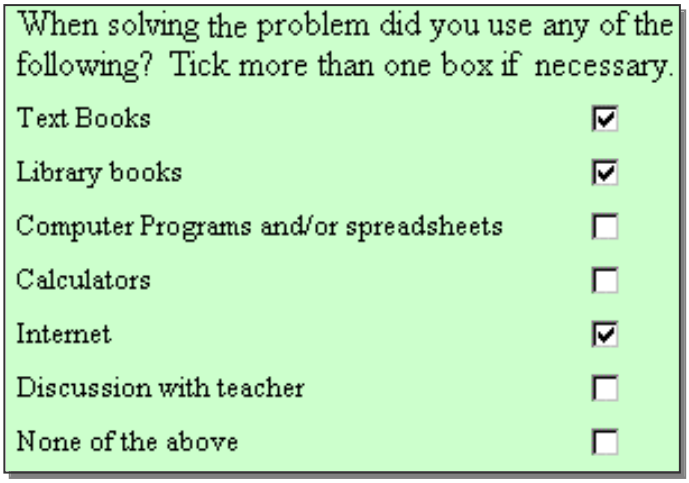


Figure 4. The use of check boxes in an item from an online mathematics survey

Drop-down menus and text fields are illustrated in Figure 5. The first item is a passage of text from an online cloze reading comprehension test where students are asked to replace each blank by entering at the keyboard the one word they think best fits. The second text passage shows an online maze reading comprehension test where students, by contrast to what was required in the cloze test, click on each blank and choose the word they think best fits from a number of alternative responses.

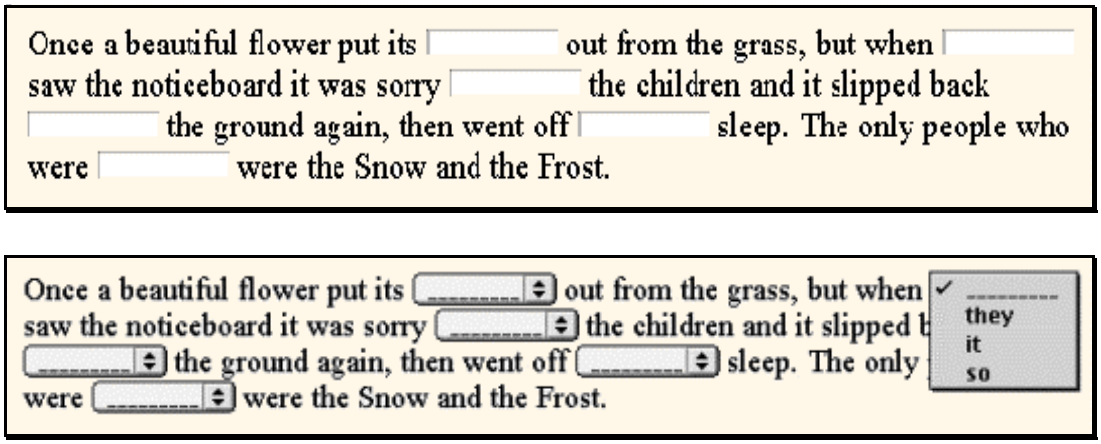


Figure 5. The use of text fields and drop-down menus in cloze and maze procedure items from an online survey of reading comprehension

Conclusion

This paper provides an overview of the possibilities offered by online surveys in educational research, with a particular focus on the standard Web page survey. Although online research methodology of the kind described is still very much in a developmental phase, clearly it has much to offer to the education research community. In an environment of increased funding cuts, educational researchers need to explore avenues that achieve equivalent outcomes at reduced costs. Use of the Internet to conduct online surveys provides an avenue that can only become more effective with time. As with surveys of the pencil and paper variety, certain problems can arise in the administration of online surveys. The researcher needs to be aware of these and to take appropriate cautions. However, the overwhelming advantage of online surveys over their pencil and paper counterparts is that when subjects complete a questionnaire online, their responses are automatically entered into a database for subsequent data analysis. Not only are online surveys more efficient, they are less prone to error.

Online surveys are not going to be appropriate for every situation where information is sought by eliciting subjects' views or knowledge about some issue or topic. However, where a sample of subjects is dispersed throughout a country, or even across countries, and if they have access to the Internet, online surveys provide a practical solution to efficient data collection.

References

Batagelj Z. and Vehovar, V. (1998) Technical and Methodological Issues in WWW Surveys. Paper presented at AAPOR '98. *Software and Methods for Conducting Internet Surveys,* St. Louis. [Online] http://www.ris.org/ris98/stlouis/index.html [24 May 2000]

Bowers D.K. (1998) FAQ's on online research. *Marketing Research,* 10 (4), 45-48.

DETE (1999) *Learning futures:DECStech 1997-2001 making connections.* Department of Education, Training and Employment. [Online] http://www.decstech.nexus.edu.au/ learn.htm [24 May 2000]

Dillman, D.A., Tortora, R.D. and Bowker, D (1998a) Influence of Plain versus Fancy design on response rates for web surveys. *Proceedings of Survey Method Section*. Annual Meeting of the American Statistical Association, Texas. [Online] http://survey.sesrc. wsu.edu/dillman/papers.htm [24 May 2000]

Dillman, D.A., Tortora, R.D. and Bowker, D. (1998b) Principles for constructing Web Surveys. *Social and Economic Sciences Research Centre Technical Report 98-50*, Pullman; Washington. [Online] http://surveys.over.net/method/litercro.html [24 May 2000]

DSS Research (1998) *Online Research: Good, bad and the ugly*. [Online] http://www.dssresearch.com/library/general/online.htm [24 May 2000]

Education Queensland (1999) *Schooling 2001: Summary of schooling 2001 project.* [Online] http://education.qld.gov.au/tal/2001/abo_sum.htm [24 May 2000]

Education Victoria (1998) *Learning Technologies in Victorian Schools, 1998-2001*. [Online] http://www.sofweb.vic.edu.au/It/pdfs/Itis.pdf [24 May 2000]

Education Western Australia (1998) *Technology 2000: A framework for implementation of learning technologies in WA Government schools.* [Online] http://www.eddept.wa. edu.au/t2000/l_t.htm [24 May 2000]

Farmer, T. (1998) Using the Internet for primary research data collection. *Market Research Library*. [Online] http://www.researchinfo.com/library/infotek/index.shtml [24 May 2000]

Gould, E.W., Gurevich, M. and Pagerey, P.D. (1998) Conducting surveys over the World Wide Web. *STC's 45th Annual Conference Proceedings*. 294-297. [Online] http://surveys.over.net/method/litercro.html [24 May 2000]

NSW DET (1997) *Making the Net returns worthwhile.* [Online] http://www.dse.nsw.edu.au/ staff/F2.0/tilt/about/projects.htm [24 May 2000]

Pilypas, H. (1997). *The Use of the Computer as a Tool for Testing Reading Comprehension*. Unpublished BEd Honours Thesis, School of Education, Flinders University, Adelaide. [Online] http://wwwed.sturt.flinders.edu.au/edweb/onpub/theses/index.htm [21 May 2000]

Pitkow, J.E. and Recker, M.M. (1994) *Using the Web as a survey tool: Results from the second WWW user survey*. [Online] http://www.cc.gatech.edu/gvu/user_surveys/ survey-09-1994/html-paper/ [24 May 2000]

Sheehan, K.B. and Hoy, M.G. (1999) Using Email to survey internet users in the United States: Methodology and Assessment. *JCMC*, 4 (3), [Online] http://209.130.1.169/ jcmc/vol4/issue3/sheehan.html [24 May 2000]

Spain, S.W. (1998) *Top-10 Web Survey issues and how to address them*. [Online] http://www.researchinfo.com/library/top_10_web.shtml [24 May 2000]

Ullman, E. (1999) Survey Says! *Small Business Computing and Communications*. [Online] http://www.smalloffice.com/sbc/071999/fea04.htm [16 Nov 1999]

Zukerberg, A., Nichols, E. and Tedesco, H. (1999) Designing surveys for the next millennium: Internet questionnaire design issues. Paper presented at the 1999 *AAPOR Conference*, Florida. [Online] http://surveys.over.net/method/ litercro.html [24 May 2000]

Resolving binary responses to the Visual Arts Attitude Scale with the Hyperbolic Cosine Model

Joanna Touloumtzoglou
Flinders University, School of Education *joanna.touloumtzoglou@flinders.edu.au*

Recent studies have shown the appropriateness of unfolding models in the analysis of binary disagree-agree responses as opposed to the use of traditional cumulative measurement models. An adaptation of Cohen's (1941) Scale of Attitude towards Aesthetic Value was used for the purpose of attitude measurement toward the visual arts. The study aimed to: (i) examine reliability of the modified scale and consistency of the items; (ii) clarify and facilitate interpretation of results obtained from the Thurstone-type scale; and (iii) improve the quality of attitude measurement. Pupils (n = 131) from two South Australian secondary schools completed two equivalent forms that comprised the Visual Arts Attitude Scale on a single occasion. The study demonstrated the usefulness of the HCM in analysing binary responses and in explaining the disagree responses in terms of their constituent elements. The manifest disagree responses were resolved into two separate latent response curves, which characterised the directions of disagree responses. These were unfolded to correspond to the data, thus providing a description of the effect that is expected as a function of the distances of persons from statements.

Key Words: Attitude Measurement, Binary Coded Data, Hyperbolic Cosine Method, Unfolding, Single Peaked Response Function

The Development of the Unfolding Model

Recent studies have shown the appropriateness of unfolding models in the analysis of binary coded data (Roberts, 1995; van Schuur & Kiers, 1994). The use of binary disagree-agree responses is traced back to Thurstone (1928, 1931). Andrich and Styles (1998, p. 454) argue that Thurstone's rationale for measuring attitudes can be summarised by two features:

(i) statements ranging from positive to ambivalent to negative affect, with respect to a construct being measured, are located numerically on an attitude continuum; and

(ii) persons are located on the continuum according to the statements they endorse.

Under Thurstone's process referred to as the pair comparisons design, the locations of statements on a scale are quantified by requiring judges to compare the selected statements in pairs, and to indicate which statement of the pair presents greater intensity of the construct. In order to analyse data derived from the pair comparisons design, Thurstone (1928) formulated the law of comparative judgement (LCJ) which is a probabilistic expression wherein a person declares that one statement has a stronger affective value than another with respect to the attitude being measured, independent of a person's location on the continuum. Thurstone used the normal distribution for the latent error and formulated a number of different cases for the LCJ depending on the constraints imposed on this error (Andrich & Styles, 1998).

One of the easiest and most common cases used is that in which the error distribution for every statement is identical. This case, according to Andrich (1978a) is equivalent to the logistic expression:

$$\Pr\{X_{ij}=1\} = \exp(\delta_i-\delta_j)[1 + \exp(\delta_i-\delta_j)] \quad (1)$$

where $Pr\{X_{ij}=1\}$ indicates the probability that statement i is considered to have a greater affective value than statement j; and δ_i and δ_j are the locations of statements i and j, respectively.

In order to estimate the location of statements from the LCJ, it was possible to examine the consistency of responses to the requirements of the model, ensuring they were located on a single continuum. Location of a person on the attitude continuum was calculated by Thurstone, based on an individual's responses to the scaled statements. This method implied a single peaked response function characterising the probability that the person agreed with a statement, with the peak occurring where the person's ideal point was at the location of the statement. Thurstone's attitude measurement procedure is used to analyse the binary responses following the ideal point process (Coombs, 1964) wherein, a person is assumed to agree with an attitude item to an extent that the item content satisfactorily represents the person's own opinion. With this perspective, it is appropriate to analyse binary responses with the single-peaked response function implemented by the unfolding model. Coombs (1964) was the first to formalise analyses of binary responses using the term unfolding, and from which the location of the items and respondents can be identified simultaneously.

In order to avoid Thurstone's time consuming and onerous approach, other formal stochastic models have been developed following the work of Likert (1932), that advanced a procedure to measure attitudes where a cumulative mechanism was presupposed. Under this procedure, the probability of an agree response (or strongly agree) increases as a function of the location of that person. Unlike the Thurstone scales, where the responses need to be unfolded to find their locations, person measures are obtained from summing the responses across statements, and by applying the correlational procedures of classical test theory. However, the method does not require that statements be located on an attitude continuum (Andrich & Styles, 1998). Traditionally, binary response data are analysed with the use of cumulative measurement models. Likert's (1932) procedure increased the number of different response categories, but did not require scaling of the statements. He also included an undecided or neutral category between disagree and agree. Other researchers have shown that this category does not necessarily operate as a middle category (Bock & Jones, 1968; Dubois & Burns, 1975; Andrich, de Jong & Sheridan, 1997). By using traditional checks of statements, such as item-total correlations and factor analyses, Likert failed to check the internal consistency of the statements (Andrich & Styles, 1998). Nevertheless, the total score of persons across statements was taken as a measure of attitude, which implied a cumulative mechanism, such that the greater the location of a person, the greater the total score.

In Thurstone's method, two people can have the same number of statements to which they agree, but their respective attitude values differ because these values depend on the type of statements with which individuals agree. In contrast, in Likert's method (1932), the greater the number of statements agreed to, the greater the person's attitude toward the construct being measured. Therefore, Likert needed to convert the single peaked unfolding Thurstone method into a cumulative method by reversing certain scores (Andrich & Styles, 1998). Consequently, the sum of the scores across statements could then provide an index of a person's attitude. Statements that reflected an ambivalent attitude toward the relevant construct could not be included in the Likert scale, since scoring would not be clear (whether to score as a negative or a positive). For this reason, statements that reflect an ambivalent attitude have no place in a Likert questionnaire (Andrich & Styles, 1998). Hence, by using the simpler Likert procedure, the location of the items

on the scale which is central to Thurstone scaling is lost. Furthermore, van Schuur and Kiers (1994) and van Schuur and Kruijtbosch (1995) argued that a factor analysis of responses to scales, which included positive and negative worded statements, gave the impression that positive and negative statements with respect to a construct are on separate dimensions, whereas when analysed according to an unfolding model, the items are located at different parts of a single continuum. Since traditional checks of statements, total correlations, and factor analyses cannot possibly test for the internal consistency of items scaled from an ideal point perspective, the necessity of analysing binary response scales with an explicit probabilistic model is mandatory.

In this study binary graded responses to attitude items were collected for the purpose of attitude measurement toward the visual arts from two secondary public schools. The instrument piloted was an adaptation of Cohen's (1941) *Scale of Attitude toward Aesthetic Value.* The purpose of the study was to examine the structure and reliability of the modified scale and the item consistency. In order to clarify and facilitate the interpretation of results obtained from the Thurstone-type scale, the dimensionality of the measuring instrument was explored, thus ascertaining the quality of attitude measurement to be included in subsequent multilevel and multitrait analyses of data. Because of the aforementioned limitations of cumulative deterministic methods of analysis the study considered the use of the probabilistic Hyperdolic Cosine Method of unfolding.

The Hyperbolic Cosine Model for Unfolding Single Stimulus Responses

The probabilistic Hyperbolic Cosine Model (HCM) for dichotomous responses of each person to each item, assumes a single peaked item response function for the items, in which the probability of a correct response decreases as the distance between a person's trait level and the item's location, increases in either direction above or below the items. Data obtained from direct responses to a dichotomous attitude item are analysed to provide estimates of item location parameters. The unfolding method is developed from two assumptions:

(i) the respondent tends to agree with the item located closest to his/her position on a unidimensional latent attitude continuum; and

(ii) the individual may choose a disagree response category from one of two possible locations on the scale of attitude. Thus a respondent might disagree with a statement when holding a very positive or a very negative attitude.

If a statement is located far below the respondent's position on the attitude continuum, then the statement is more negative than the respondent's attitude and thus the person disagrees from above the item (DA), since the person is at an attitude level higher up the attitude scale than the statement. If however, the item is above the person's location on the attitude continuum, then the item is more favourable than the individual's attitude and thus the person disagrees from below the item (DB). In this case, the unfolding model postulates two subjective responses for each observable disagree response on a rating scale. Figure 1 gives a graphical representation of the three categories. Here, the single disagree response is composed of the sum of two latent responses, DA and DB. In addition to these two responses, there is also the single response (A-agree), which implies that the respondent's location is close to that of the statement.

By Andrich and Luo (1993) the HCM model takes the form:

$$\Pr\{x_{ni}=1\} = \exp(\theta_i)/[\exp(\theta_i)+2\cosh(\beta_n-\delta_i)]$$

$$\Pr\{x_{ni}=0\} = 2\cosh(\beta_n-\delta_i)/[\exp(\theta_i)+2\cosh(\beta_n-\delta_i)] \qquad (2)$$

where x = 1 is a positive response; x = 0 is a negative response; β_n is the location of person n; δ_i is the location of item i; θ_i is the unit of item i; and cosh(u) = [exp(-u)+exp(u)]/2 is the hyperbolic cosine function.

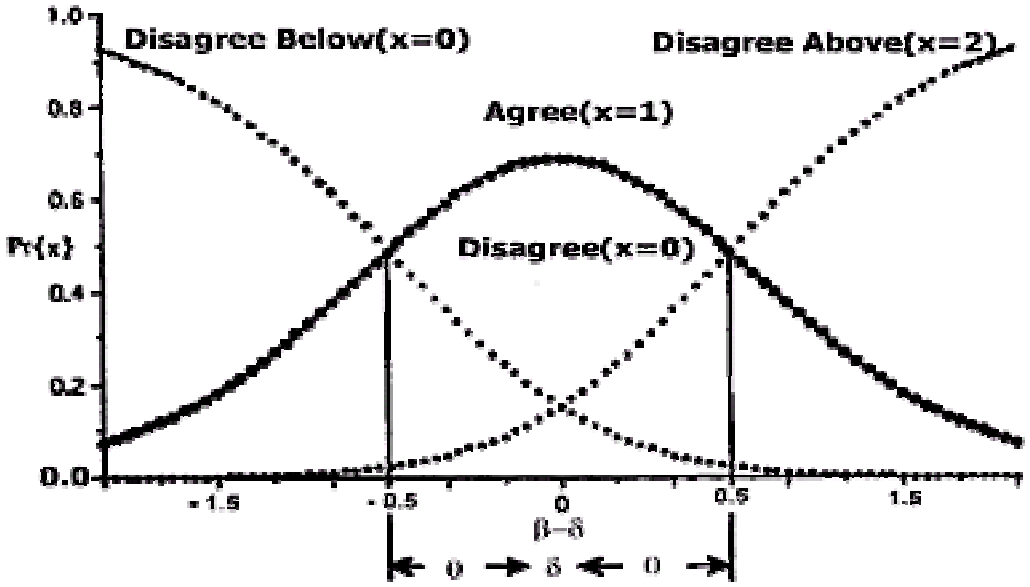


Figure 1. Response functions of the HCM, including the resolved disagree response

The model used to construct the HCM in the case of three ordered categories (where DB = 0, DA = 2 and A = 1) as shown by Andrich and Styles (1998), takes the explicit form:

$$\Pr\{Y_{ni}=0\} = 1/\eta_{ni}$$

$$\Pr\{Y_{ni}=1\} = \exp[\theta_i+(\beta_n-\delta_i)]/\eta_{ni}$$

$$\Pr\{Y_{ni}=2\} = \exp[2(\beta_n-\delta_i)]/\eta_{ni} \quad (3)$$

where β_n and δ_i are the locations of person n and statement i respectively; θ_i reflects the distance from the statement location δ_i where the DB and DA response functions intersect with the A response function; and η_{ni} is a normalising factor that is the sum of the numerators and ensures that the probabilities of the responses sum to one.

Method

Sample, measures and procedure

A total of 131 students (Year 8 to Year 11) from two South Australian secondary schools participated in the study, which was a trial for a more extensive investigation. Binary graded disagree-agree responses to attitude items were obtained for the purpose of piloting the Visual Arts Attitude Scale adapted from Cohen's (1941) Scale of Attitude toward Aesthetic Value. The attitude scale consisted of two equivalent forms each comprised of 20 statements. Form A is presented in Table 6 and Form B in Table 10. They provided two sets of scores that were analysed for an estimate of reliability. Statements included in the original two forms had been scaled by the paired comparison process and were placed in decreasing intensity of affect (from extreme positive to extreme negative). The equivalent-forms reliability method is a traditional method that estimates reliability based on the administration of two forms that may be considered

equivalent. In this method, the two forms provide two sets of scores that are correlated to obtain an estimate of reliability. An average of the intercorrelations among the forms may be taken as the estimate of reliability. Provided both forms actually measure the same attitude, the correlation coefficient will reflect consistency of measurement. The purpose of this pilot study was to examine the reliability of the modified scale and consistency of items within the scale. Both forms of the attitude scale were administered in the classroom on a single occasion. The scale was not timed, thus allowing ample time for reading and understanding statements. Respondents were requested to agree or disagree with each statement.

Statistical analyses

The following statistical analyses were performed on SPSS version 8.0, QUEST (Adams & Khoo, 1993) and RUMMFOLDss (Andrich & Luo, 1998). The basis of RUMMFOLDss lies in the principles of the Rasch unidimensional measurement model (RUMM) for three ordered categories, which gives the Hyperbolic Cosine Model (HCM) for unfolding. In the Rasch model for a simple dichotomous item (agree-disagree), an individual's (n) response to item (i) can be expressed as a function of the person's ability (β_n) on the trait being tested as well as item difficulty (δ_i). The probabilistic HCM for the dichotomous responses ('1' = agree, '0' = disagree) of each respondent to each item is single peaked. The HCM resolves the disagree responses into two components (Andrich & Luo, 1993). An individual may regard the location of an item as being below himself or herself, or an individual may perceive himself or herself as being lower than the item. An individual gives an agree response to an item when he or she perceives him or herself to be close to the location of the item. Thus with the aid of RUMMFOLDss, responses are examined using the Rasch model for three categories (Andrich, 1982). Figure 1 contains the probabilistic curves of the HCM (x = 0, 1) and the Rasch model for three ordered categories (y = 0, 1, 2) (Andrich & Luo, 1998). The HCM equation is expressed by Andrich and Luo (1993) as:

$$P\{x_{ni}=1\} = \cosh\rho_i/[\cosh\rho_i+\cosh(\beta_n-\delta_i)]$$

$$P\{x_{ni}=0\} = \cosh(\beta_n-\delta_i)/[\cosh\rho_i+\cosh(\beta_n-\delta_i)] \quad (4)$$

where the parameter ρ_i *characterises the latitude of acceptance of the item/statement. The relationship between* ρ_i *and* θ_i *is* $\exp\theta_i = 2\cosh\rho_i$.

The model deals here with dichotomised responses, and since the HCM is single-peaked and symmetric about $\beta_n-\delta_i = 0$, for each of the possible values there are two corresponding positions on the latent trait continuum. Therefore, a sign analysis is performed where the program identifies a pair of items that have the minimum correlation coefficient, so that the signs of correlation coefficients of items with either one of these two items, is chosen as their initial sign. There are four major cycles performed using the Joint Maximum Likelihood (JML) principle, with a constraint that the mean of item locations are equal to zero (0). These four cycles are defined as follows:

(i) Cycle 1 - Parameters are estimated with the additional constraint that all item units have the same value θ_0;

(ii) Cycle 2 - Parameters are estimated with item units allowed to vary, but with their mean value constrained to be equal to θ_0;

(iii) Cycle 3 - Parameter estimates of Cycle 1 are corrected because of inconsistencies in the estimates of parameters generated from the JML procedure; and

(iv) Cycle 4 - Parameter estimates of Cycle 2 are corrected for the same reasons stated in Cycle 3.

From the location estimates for each person and their associated standard errors, an index of (person) separation is constructed according to the expression:

$$\textit{Separation Index} = (V[\beta]-V[e])/V[\beta] \quad (5)$$

where V[e] is the average variance of the errors from estimates of β *among persons. It can be interpreted as Cronbach's alpha reliability coefficient, that is, as indicators of the items being responded to consistently by subjects or whether the set of items are measuring several different incoherent attributes. In particular, V[e] is a characteristic of a set of persons in the context of a particular set of items.*

Individuals are ordered according to the estimates of their locations on the continuum and grouped into three class intervals. The number of observed agree (A) responses to each statement by individuals in each class interval are then compared with the expected number according to HCM. This comparison is formalised as an approximate chi-square statistic for each item i, on G-1 degrees of freedom, according to the following equation:

$$\chi_i^2 = \sum_{g=1}^{G} [\Sigma x_{ni} - E(\Sigma x_{ni})]^2 / V[\Sigma x_{ni}] \quad (6)$$

where g = 1, 2.…G is the number of class intervals; Σx_{ni}, (n g) is the total number of individuals in class interval g who agreed to statement i; $E[\Sigma x_{ni}]$, (n g) is the expected number of A responses according to the model; and $V[\Sigma x_{ni}]$, (n g) is the variance of expected number of A responses. Since the model is single-peaked, if responses are consistent with the model and if individuals are located around the location of the statement, then they too will be single-peaked (Andrich & Styles, 1998).

Results

Exploratory Factor Analysis of the Visual Arts Attitude Scale (Forms A and B)

Initial analyses of data were conducted with principal component analysis followed by varimax rotation, in order to determine the factor structure of the Visual Arts Attitude Scale. A number of factors resulted from the principal components analysis that did not coincide with the hypothesised underlying unidimensionality of the instrument, but rather confirmed the argument stated by van Schuur and Kiers (1994) and van Schuur and Kruijtbosch (1995), that a factor analysis of responses to scales gave the impression that negative and positive statements with respect to a construct located on separate dimensions.

From the principal components analysis there were six components extracted with eigenvalues greater than unity. Since some items had substantial loadings on more than one factor, a varimax rotation was performed to elucidate the results. Tables 1 and 2 contain respectively the factor loadings (Form A) for the initial principal components analysis which resulted in the extraction of six factors, and from the varimax rotation of the three strongest factors.

The results from the two analyses presented in Tables 1 and 2 showed contrasting factor structures as several items that previously had recorded the highest factor loadings on a specific component, following the varimax rotation showed the highest factor loadings on an alternate

component. For example, Item 12 initially had the highest factor loading (-0.63) on Factor 3, while following varimax rotation its highest factor loading was 0.58 on Factor 1. Similarly, Item 15 had its highest loading (-0.52) on Factor 1 and following varimax rotation its highest loading (0.70) was on Factor 2. Items 17, 18, 19 and 20 had high negative loadings (-0.62, -0.61, -0.63 and -0.57 respectively) on Factor 1, while following varimax rotation their highest positive loadings (0.72, 0.69, 0.64 and 0.67 respectively) were on Factor 2. Finally, Items 3, 8 and 9 had their highest factor loadings (0.67, 0.64 and 0.68 respectively) on Component 1 whereas following the varimax rotation their highest factor loadings (0.51, 0.75 and 0.83 respectively) were on Component 3.

Table 1. Principal Components Analysis of Visual Arts Attitude Scale Form A

Item No.	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
13	-0.33	0.46	0.23	0.23	0.22	0.52*
11	0.13	0.43	0.17	-0.69*	-0.23	-0.12
12	0.21	0.31	-0.63*	-0.22	0.32	0.02
16	-0.47	0.66*	0.27	-0.10	-0.07	-0.04
14	-0.29	0.58*	-0.07	-0.22	-0.48	0.01
01	0.71*	0.44	-0.24	0.07	0.11	-0.30
15	-0.52*	0.43	0.29	0.03	0.51*	-0.13
17	-0.62*	0.45	0.06	-0.27	-0.28	0.11
18	-0.61*	0.46	-0.07	0.35	0.08	0.11
19	-0.63*	0.38	-0.07	-0.26	0.32	-0.22
02	0.75*	0.21	-0.03	-0.01	0.14	0.35
20	-0.57*	0.48	-0.13	0.26	0.00	0.14
03	0.67*	0.14	0.22	0.30	-0.16	0.22
04	0.82*	0.15	-0.11	0.13	0.01	-0.08
05	0.75*	0.37	-0.34	0.04	0.18	-0.11
06	0.78*	0.38	0.12	0.18	-0.03	0.04
07	0.61*	0.39	-0.21	-0.19	-0.13	-0.24
08	0.64*	0.11	0.51	-0.22	0.35	0.28
09	0.68*	0.07	0.59	-0.04	0.00	0.00
10	0.44*	0.43	0.03	0.31	-0.29	0.40
Total	7.02	3.16	1.57	1.23	1.18	1.02
%Variance	35.08	15.81	7.83	6.15	5.89	5.09

* Highest loadings of items contributing to each factor

Item No.	Factor 1	Factor 2	Factor 3
01	0.85*	-0.07	0.14
02	0.67*	-0.24	0.33
03	0.47	-0.20	0.51*
04	0.71*	-0.35	0.28
05	0.88*	-0.18	0.06
06	0.72*	-0.08	0.49
07	0.75*	-0.02	0.19
08	0.30	-0.15	0.75*
09	0.28	-0.19	0.83*
10	0.55*	0.12	0.26
11	0.27	0.32	0.23
12	0.58*	0.00	-0.44
13	-0.05	0.60*	0.08
14	0.16	0.62*	-0.15
15	-0.24	0.70*	0.05
16	-0.05	0.84*	0.06
17	-0.20	0.72*	-0.20
18	-0.14	0.69*	-0.31
19	-0.20	0.64*	-0.33
20	-0.07	0.67*	-0.34
Total	4.99	3.37	1.96
%Variance	24.97	16.88	9.82

Table 2.
Varimax rotation for Visual Arts Attitude Scale Form A with limited number of factors extracted

* Highest loadings of items contributing to each factor

A principal components analysis was also undertaken on data collected from Form B of the Visual Arts Attitude Scale. From the analysis, five components were extracted with eigenvalues greater than unity. Examination of the factor loadings of each item, showed that further analysis using varimax rotation was necessary to clarify the results. Table 3 contains the factor loadings for the initial principal components.

Table 3. Factor analysis of responses to Visual Arts Attitude Scale Form B

Item No.	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
08	-0.09	0.34	0.19	0.43	-0.53*
17	0.26	0.11	0.45	-0.06	0.57*
16	0.42	0.30	0.11	0.51*	-0.40
03	0.25	0.45	-0.68*	0.30	0.04
13	0.37	0.61*	0.28	0.07	0.12
09	0.35	0.51*	0.07	-0.52	-0.15
10	0.13	0.47*	-0.40	-0.30	-0.11
01	-0.07	0.44*	0.03	-0.08	0.38
15	0.66*	0.53	0.21	-0.01	0.13
14	0.62*	0.16	0.37	0.00	-0.10
18	0.74*	0.31	0.09	0.33	-0.08
19	0.78*	0.04	-0.39	0.25	-0.13
02	-0.55*	0.54	-0.10	0.34	0.01
20	0.73*	0.01	-0.35	0.16	-0.12
11	0.59*	0.30	-0.11	-0.45	-0.06
12	0.59*	0.07	0.32	-0.28	-0.19
04	-0.67*	0.50	0.02	0.30	-0.01
05	-0.58*	0.50	0.21	-0.03	-0.15
06	-0.74*	0.48	-0.08	0.08	0.04
07	-0.60*	0.36	0.08	0.20	-0.19
Total	5.51	3.53	1.95	1.65	1.13
%Variance	27.57	17.67	1.95	1.65	1.13

* Highest loadings of items contributing to each factor

Table 4 contains factor loadings resulting from the varimax rotation. The factor loadings demonstrated that Items 10, 11, 12, 14, 15, 16, 17, 18 and 8 contributed to a different factor from the one observed prior to varimax rotation. For example, Item 10 presented the highest loading (0.60) on Factor 3 as opposed to the highest loading (0.47) on Factor 2 presented prior to varimax rotation. Consequently, results extracted from the exploratory factor analysis did not facilitate interpretation of data obtained from the administration of the Visual Arts Attitude Scale (Forms A and B). In view of the results obtained by factor analyses conducted on responses from both Forms A and B, it appeared that there were more factors than desired, as the ideal instrument would need to conform to a unidimensional model for measuring a single attribute based on the analysis of individual responses. Therefore, in order to facilitate further understanding and interpretation of the data, a Rasch analysis was undertaken.

Rasch analysis of the Visual Arts Attitude Scale (Form A and Form B)

The Rasch measurement model is employed to produce a scale on which all items in an instrument and all respondents can be placed, and embodies the assumption that all responses are described by a response model. Particular interest is given to the item fit through the examination of the infit mean square (MNSQ) values which are calculated. Figure 2, displays the input file ‘att30.inp’ used in the Rasch analysis. Key codes were inserted so that the program identified items coded ‘1’ as having a positive attitude toward the visual arts and items coded ‘0’ as having a negative attitude.

Item No.	Factor 1	Factor 2	Factor 3
01	0.18	-0.05	0.71*
03	-0.08	0.11	0.84*
10	0.09	0.18	0.60*
09	0.06	0.57*	0.24
12	-0.32	0.57*	-0.15
13	0.16	0.74*	0.12
14	-0.14	0.81*	0.25
15	-0.14	0.81*	0.25
16	-0.11	0.50*	0.12
17	-0.02	0.45*	-0.30
18	-0.35	0.70*	0.18
04	0.74*	-0.18	0.03
05	0.79*	0.02	-0.00
06	0.83*	-0.23	0.21
07	0.70*	-0.13	0.02
08	0.32*	0.24	-0.00
11	0.48*	0.46	0.04
02	0.72*	-0.07	0.29
19	-0.66*	0.34	0.47
20	-0.62*	0.32	0.41
Total	4.37	3.47	2.26
%Variance	21.85	17.34	11.29

Table 4.
Factor loadings following a varimax rotation on Visual Arts Attitude Scale Form B

* Highest loadings of items contributing to each factor

```
title Attitude scales 1999
data_file c:\progra~1\spss\attitud3.dat
codes 01
set with=132!page
format id 1-3 items 7-136
key 11111111111111111111111110000000000000000000-
    11111111111111111111111110000000000000000000-
    11111111110000000000111111110000000000
scale 1-45 ! vas-a
scale 46-90 ! vas-b
scale 91-110 ! va-a
scale 111-130 ! va-b
estimate ! iter=100;scale=vas-a,vas-b,va-a,va-b>>att300.out
show ! scale=vas-a,vas-b,va-a,va-b>>att311.out
show items !scale=vas-a,vas-b,va-a,va-b>>att322.out
show case !scale=vas-a,vas-b,va-a,va-b>>attitud333.out
itanal !scale=vas-a,vas-b,va-a,va-b>>att344.out
correlate vas-a,vas-b,va-a,va-b>>att355.out
quit
```

Figure 2. Input file for Visual Arts Attitude Scales

However, in using this coding system it is difficult to determine the exact location of central items that involve an ambivalent response toward the visual arts. From this system, it is presumed that individuals who agree with the statements in items coded ‘1’ and disagree with statements in items coded ‘0’, are characterised as having demonstrated a positive attitude towards the visual arts. Moreover, students who disagree with the statements in items coded ‘1’ and agree with statements in items coded ‘0’, are subsequently characterised as having demonstrated a negative attitude toward the visual arts. Table 5 contains the item estimates for Forms A and B of the Attitude Scale.

Table 5. Item estimates* for Form A and Form B of the Visual Arts Attitude Scale (N = 131 Probability Level = 0.50)

Form A								
ITEM No.	SCORE	MAXSCR	THRSH	SE	INFT MNSQ	OUTFT MNSQ	INFT t	OUTFT t
01	38	111	1.00	0.23	1.02	0.97	0.20	-0.10
02	54	123	0.53	0.21	0.76	0.70	-3.00	-1.80
03	54	120	0.52	0.22	0.85	0.75	-1.80	-1.40
04	40	117	1.05	0.23	0.80	0.69	-1.90	-1.60
05	35	106	1.06	0.24	0.92	0.77	-0.60	-1.00
06	34	98	0.93	0.25	1.13	1.01	1.10	0.10
07	33	107	1.22	0.25	0.94	0.97	-0.40	0.00
08	78	123	-0.47	0.22	0.99	0.91	0.00	-0.30
09	78	117	-0.66	0.23	0.85	0.66	-1.50	-1.40
10	55	111	0.20	0.22	**1.47**	2.06	5.00	4.40
11	89	119	-1.13	0.25	**1.45**	1.61	3.10	1.70
12	84	111	-1.33	0.25	**1.53**	3.70	3.40	5.00
13	37	103	0.87	0.24	1.11	1.13	1.00	0.70
14	63	98	-0.60	0.25	1.13	1.22	1.20	0.90
15	62	110	-0.13	0.23	0.93	0.79	-0.80	-1.00
16	69	109	-0.53	0.23	0.87	0.80	-1.40	-0.80
17	74	105	-0.90	0.25	0.83	0.66	-1.40	-1.20
18	75	114	-0.61	0.23	0.82	0.68	-1.80	-1.40
19	63	117	-0.01	0.22	0.79	0.66	-2.70	-2.00
20	84	116	-1.02	0.25	0.87	0.70	-1.10	-1.00
Mean			0.00		1.00	1.07	-0.20	-0.10
SD			0.84		0.24	0.71	2.10	1.90

Form B								
ITEM No.	SCORE	MAXSCR	THRSH	SE	INFT MNSQ	OUTFT MNSQ	INFT t	OUTFT t
01	31	124	1.61	0.23	0.98	1.11	-0.10	0.50
02	56	121	0.46	0.20	0.90	0.86	-1.30	-1.00
03	25	111	1.75	0.25	1.11	1.24	0.80	0.90
04	42	106	0.82	0.22	0.75	0.67	-2.80	-2.00
05	73	102	-0.73	0.24	0.87	0.71	-1.20	-1.30
06	66	106	-0.28	0.22	0.75	0.65	-3.00	-2.20
07	64	107	-0.16	0.22	0.88	0.85	-1.50	-0.80
08	42	90	0.35	0.23	1.40	1.43	4.00	2.30
09	51	109	0.43	0.21	1.64	2.16	6.60	5.60
10	83	117	-0.73	0.22	1.35	1.59	3.10	2.40
11	60	107	0.00	0.22	1.31	1.31	3.40	1.80
12	51	120	0.64	0.21	0.83	0.86	-2.20	-0.90
13	71	116	-0.23	0.21	1.17	1.30	2.00	1.60
14	64	106	-0.19	0.22	0.85	0.79	-1.70	-1.20
15	79	116	-0.56	0.22	0.79	0.68	-2.30	-1.80
16	64	100	-0.38	0.23	0.95	1.07	-0.50	0.40
17	76	113	-0.48	0.22	1.02	1.16	0.20	0.80
18	88	121	-0.80	0.22	0.81	0.68	-2.00	-1.60
19	88	120	-0.85	0.23	0.77	0.61	-2.20	-1.90
20	84	119	-0.69	0.22	0.82	0.66	-1.90	-1.80
Mean			0.00		1.00	1.02	-0.10	0.00
SD			0.76		0.25	0.40	2.60	2.00

* where SCORE = raw score; MAXSCR = maximum possible score; THRSH = item difficulty estimate; SE = standard error; and INFT (OUTFT) MNSQ = infit (oufit) mean squares. Bold indicates very high infit MNSQ values, a misfitting of the central items.

From the results it was observed that for Form A of the Visual Arts Attitude Scale Items 10, 11, and 12 presented very high infit MNSQ values of 1.47, 1.45, and 1.53 respectively, while for Form B, Items 8, 9, 10, and 11 presented infit MNSQ values of 1.40, 1.64, 1.35 and 1.31 respectively. A clearer view of these misfitting items is shown in Figures 3(a) and 3(b) where the infit MNSQ values are plotted for each item. These items fall outside the range of acceptance, or close to the margins of the established range (between 0.77 and 1.30). For both forms, misfitting items appear in the centre of the scale, indicating a difficulty in resolving negative responses related to the degree of affect expressed in these ambivalent central statements. This effect is a consequence of the structure of the scales, since the scales are designed to contain items indicating extremely positive or extremely negative attitudes on either end of the scale, and items with an ambivalent position in the centre. As a result, students with very positive attitudes could agree strongly with items on the top end of the scale and disagree strongly with items on the bottom end of the scale. However, students agreement or disagreement with central items of the scale, would be difficult to define as to whether students disagreed because of a stronger positive approach or because of a stronger negative approach toward the statement contained in the item (Andrich, 1996). For this reason, it was important to identify students whose negative responses to the central items were due to an overall positive attitude toward the visual arts, or students whose negative responses to central items were due to an overall negative attitude toward the visual arts. In other words, careful attention was necessary for interpreting negative responses given to items located in the centre of the Visual Arts Attitude Scale. A method of resolving the analysis and interpretation of such data was obtained from the Hyperbolic Cosine Model (Andrich & Luo, 1993) for unfolding single stimulus responses. Consequently, an analysis using the unfolding model could result in proving that the components generated from exploratory factor analyses and the inconsistency present in items demonstrated through Rasch analyses of equivalent forms of the Visual Arts Attitude Scale, were just another expression of describing items located at different parts of a single attitude continuum.

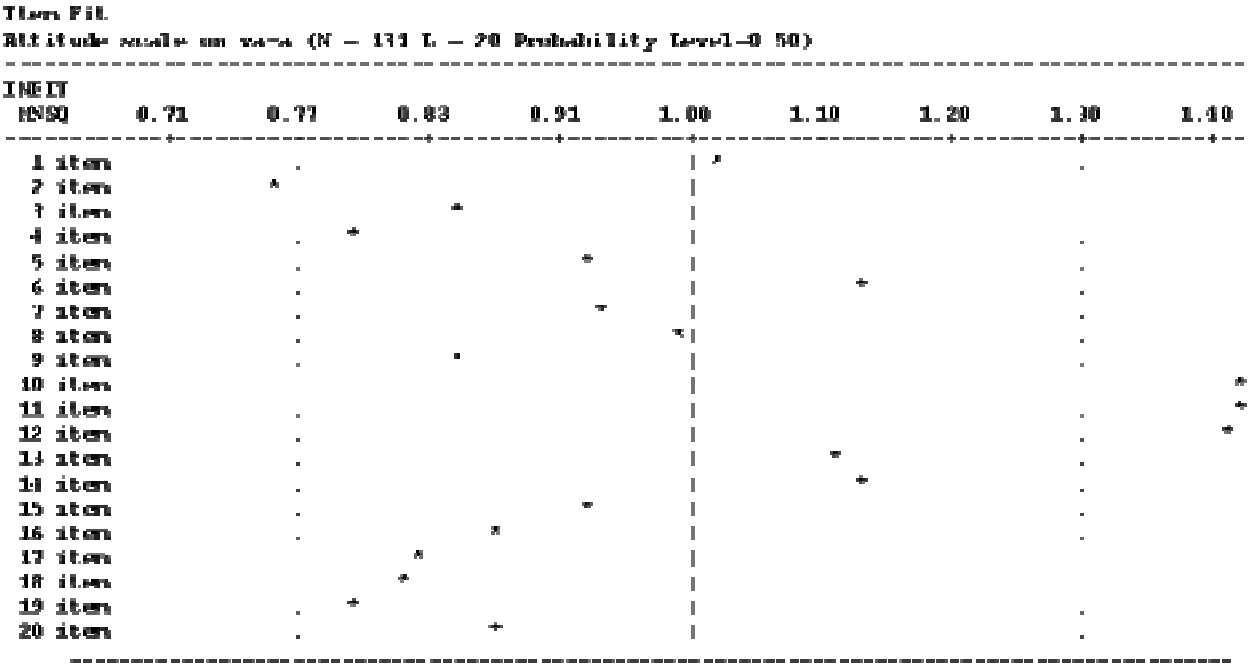


Figure 3(a). Plot of item fit for Form A of the Visual Arts Attitude Scale

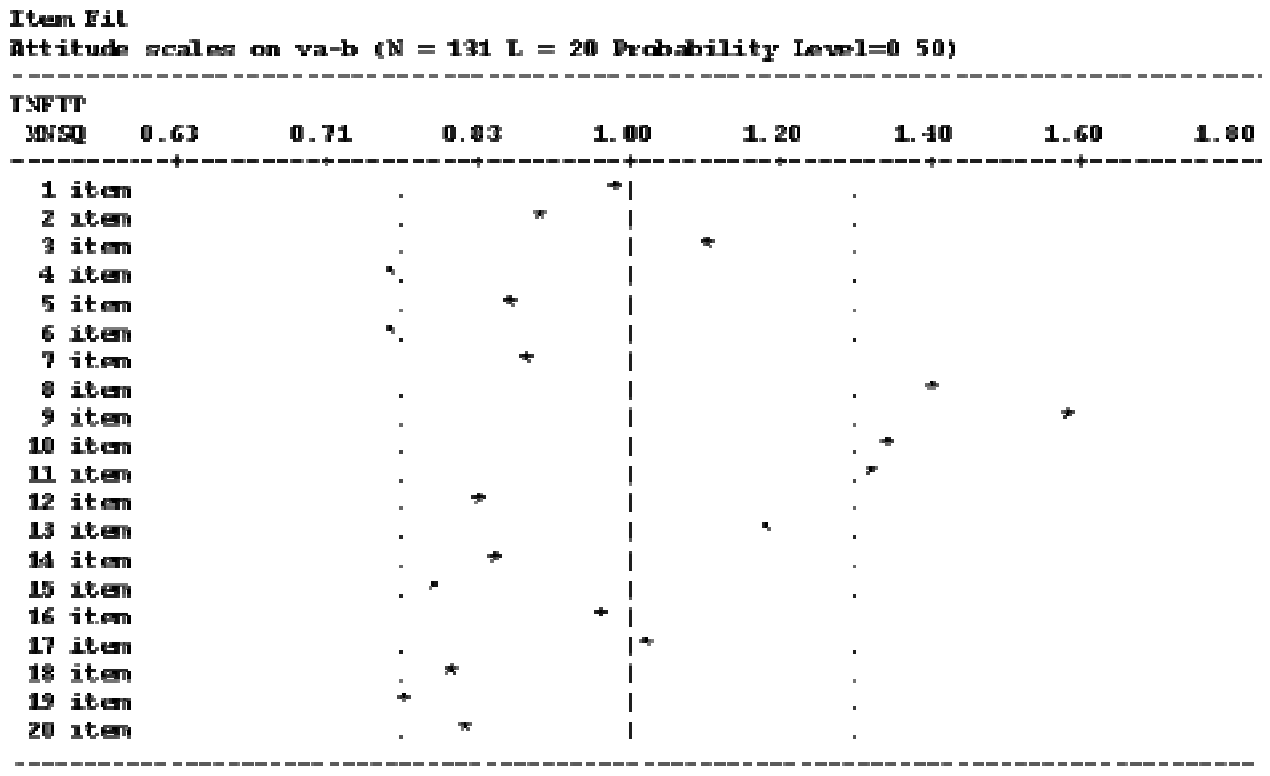


Figure 3(b). Plot of item fit for Form B of the Visual Arts Attitude Scale

HCM analysis of the Visual Arts Attitude Scale (Form A and B)

Since the HCM is single-peaked and symmetric about $\beta_n - \delta_i = 0$, then for each of the possible values, there are two corresponding positions on the latent trait continuum. The assignment of a sign to each item is needed in order to initiate the estimation process. The sign analysis performed by RUMMFOLDss identifies pairs of items with the minimum correlation coefficient, so that the signs of correlation coefficients of items with either one of these items is selected as their initial sign. In this study, the sign analysis identified the lowest correlation between pairs of items being Item 2 and Item 19 (-0.53), with which the correlations between items were exported into the input card. Item 2 was chosen as the initial sign, representing the extreme positive attitude. Table 6 contains the correlation coefficients of the items with Item 2. The correlation matrix of the ranked items resulted in Item 2 and Item 19 being the two extreme opposite statements on the scale.

After completion of the four cycles of the estimation, item locations, item units and associated standard errors per cycle were presented in the output. Table 7 contains the estimates of item units, with the four columns each corresponding to a different cycle of estimation. The first column, 'theta1', contains the estimates for item units, where all item units have the same value 3.06, or, all items have the same width of acceptance. The item unit value is 2.74 in the column 'theta3' which includes a correction of equal value to all item units for JML inconsistencies which stretch out the estimates of the parameters. In columns 'theta2' and 'theta4', the item units differ in value as they are allowed to float, as would be observed in the case of the Partial Credit Model (Masters, 1982), in which the item width varies. These item unit values are associated with corresponding standard errors for each cycle. Thetas can be regarded as proxy performance or attitude level parameters, since large thetas can be used to distinguish between items and individuals. Positive $+\theta$ is the boundary set for the top end on the attitude continuum, while

negative $-\theta$ sets the boundary for the bottom end on the attitude continuum. From the results observed in column 'theta4', Items 9 and 12 appear to have the largest latitude of acceptance with theta values of 3.64 and 4.00 respectively, after allowing for correction. These items may be regarded as having good discrimination indices. Standard errors which in conjunction with a person's location can also function as a discrimination index. These estimates are also provided by RUMMFOLDss. For example, if a person has a very negative attitude toward the visual arts, yet presents a high standard error, this shows an inconsistency in that person's behaviour.

Table 6. Correlation coefficient exported with Item 2 (Form A)

Item No.	Statements	Correlation coefficient
01	It is in the visual artistic experiences of life that I find my greatest satisfaction.	0.39
02	I have a great interest in matters of the visual arts.	1.00
03	I am interested in anything in which I can see a visual artistic quality.	0.42
04	Attendance at a visual arts exhibition gives me inspiration.	0.46
05	I believe that the pursuit of visual artistic interests increase one's satisfaction in living.	0.50
06	I believe that visual arts promote desirable relationships between nations	0.17
07	I am attracted to individuals that pursue visual artistic interests.	0.49
08	I believe that everyone should have a little training in the visual arts.	0.37
09	I am in favour of visual arts exhibitions for they do no harm to anyone.	0.33
10	I believe that the teaching of the visual arts subject is O.K., but the type of person now teaching it fails to 'get it across'.	0.08
11	Visual arts do not interest me now, but I expect that sometime I shall find time to pursue them actively.	-0.07
12	I would be willing to give money to support visual arts enterprises if it were not for the 'highbrow' atmosphere surrounding them.	0.18
13	Practical considerations should come first, visual arts second.	-0.26
14	I believe that individuals engaged in purely visual arts occupations are parasites on society.	-0.14
15	I do not believe I would receive any benefit from lectures concerning visual arts subjects.	-0.34
16	I do not care for visual artists because their interests seem to me to be more emotional than rational.	-0.31
17	I see no reason for the government to spend money on aesthetic objects and activities.	-0.28
18	I have no desire to join or have anything to do with any organisation developed to visual arts activities.	-0.33
19	I see very little worth while in visual artistic interests.	**-0.53**
20	Visual arts education is nonsense.	-0.27

Bold indicates the smallest correlation coefficient found between Item 2 and Item 19.

A generalised χ^2 test of fit is performed for the fourth cycle, with different item units with correction, including three person-groups or class intervals because of the single-peaked structure of the HCM function. The three class intervals correspond to three person-groups, ordered according to the estimates of their locations on the continuum and represent three graded responses. The number of observed responses to each statement by individuals in each class interval, are compared to the expected number of responses according to the HCM, resulting in the calculation of an approximate χ^2 statistic for each item. The reason for selecting different unit values for analysis of the fit statistic, lies in the nature of the instrument and context of the items. Attitude items express diverse emotive intensity for each person, as they elicit highly subjective responses. As a consequence, the latitude of acceptance is allowed to float and vary according to the sample, and is not restricted after implementation of a single unit value for all items. Table 8

shows the observed and expected agreements and disagreements for three class intervals of an extreme positive statement (Item 1), a fairly negative statement (Item 14) and an extreme negative statement (Item 20). Figures 3, 4 and 5 provide the location and the theta value (unit) for Items 1, 14 and 20.

Table 7. Estimates of item units for each corresponding cycle of estimation (Form A)

Item No.	theta1	theta2	theta3	theta4	StdEr1	StdEr2	StdEr3	StdEr4
01	3.06	2.55	2.74	2.26	0.24	0.24	0.23	0.23
02	3.06	2.72	2.74	2.38	0.21	0.22	0.21	0.21
03	3.06	2.73	2.74	2.40	0.21	0.22	0.21	0.21
04	3.06	2.31	2.74	2.02	0.23	0.23	0.23	0.22
05	3.06	2.44	2.74	2.13	0.23	0.24	0.23	0.23
06	3.06	2.53	2.74	2.22	0.24	0.23	0.23	0.23
07	3.06	2.15	2.74	1.84	0.23	0.23	0.22	0.22
08	3.06	3.40	2.74	3.04	0.21	0.21	0.21	0.21
09	3.06	**4.02**	2.74	**3.64**	0.21	0.23	0.21	0.22
10	3.06	2.56	2.74	2.24	0.21	0.20	0.21	0.20
11	3.06	3.85	2.74	3.54	0.23	0.24	0.22	0.24
12	3.06	**4.22**	2.74	**4.00**	0.24	0.28	0.22	0.27
13	3.06	2.16	2.74	1.85	0.22	0.21	0.22	0.21
14	3.06	3.33	2.74	3.03	0.22	0.22	0.21	0.22
15	3.06	3.18	2.74	2.84	0.21	0.21	0.21	0.21
16	3.06	3.22	2.74	2.91	0.22	0.21	0.21	0.21
17	3.06	3.59	2.74	3.29	0.23	0.23	0.22	0.23
18	3.06	3.53	2.74	3.19	0.21	0.22	0.21	0.21
19	3.06	2.92	2.74	2.59	0.21	0.20	0.21	0.20
20	3.06	3.76	2.74	3.43	0.22	0.23	0.21	0.23

Bold indicates items with the largest latitude of acceptance

Table 8. Expected and observed responses for Item 1, Item 14 and Item 20

Item/	**Class Interval**						χ^2	**p**
statement	1		2		3		**(2, N = 131)**	
PA1	0	1	0	1	0	1		
Obs.	37.00	3.00	20.00	17.00	10.00	21.00		
Exp.	33.95	6.05	22.30	14.70	12.34	18.66		
χ^2	1.946		0.610		1.042		3.60	0.1
NA14	0	1	0	1	0	1		
Obs.	16.00	18.00	20.00	10.00	27.00	4.00		
Exp.	14.15	19.85	19.43	10.57	30.06	0.94		
χ^2	0.442		0.048		10.923		11.41	0.0
NA20	0	1	0	1	0	1		
Obs.	21.00	15.00	24.00	13.00	40.00	0.00		
Exp.	17.22	18.78	29.22	7.78	39.37	0.63		
χ^2	1.78		4.49		0.66		6.94	0.0

Note. NA = negative worded attitude statement; PA = positive worded attitude statement; Obs. = observed responses; Exp. = expected responses

In Table 8, the expected and observed number of responses (score) and the chi-square value for each item are displayed. Item 1 has 37 observed negative responses and 3 observed positive responses from the first class interval, the group with an extremely negative attitude. This item also has 20 observed negative responses and 17 positive responses from the second class interval, the group with an ambivalent attitude toward the visual arts. Finally, the third interval, representing individuals with an extreme positive attitude toward the visual arts gave 10 negative responses and 21 positive responses to Item 1. The number of observed responses does not differ significantly from the expected responses, an indicator of fit of this item to the expected or ideal model. Figure 4 contains the plot of the average value for each group and the related observed

proportion of positive responses for Item 1. Observed proportions are close to the curve indicating that the item fits the model. The rectangle presents the latitude of acceptance of the item, and it is evident that Item 1 is an extremely positive statement. From Figure 4, Item 1 location is 2.28 with a unit value of $\theta = 2.28$, which demonstrates it is a highly positive statement placed close to the top of the attitude continuum. It has a chi-squared value of 3.60 for two degrees of freedom and fits the model well.

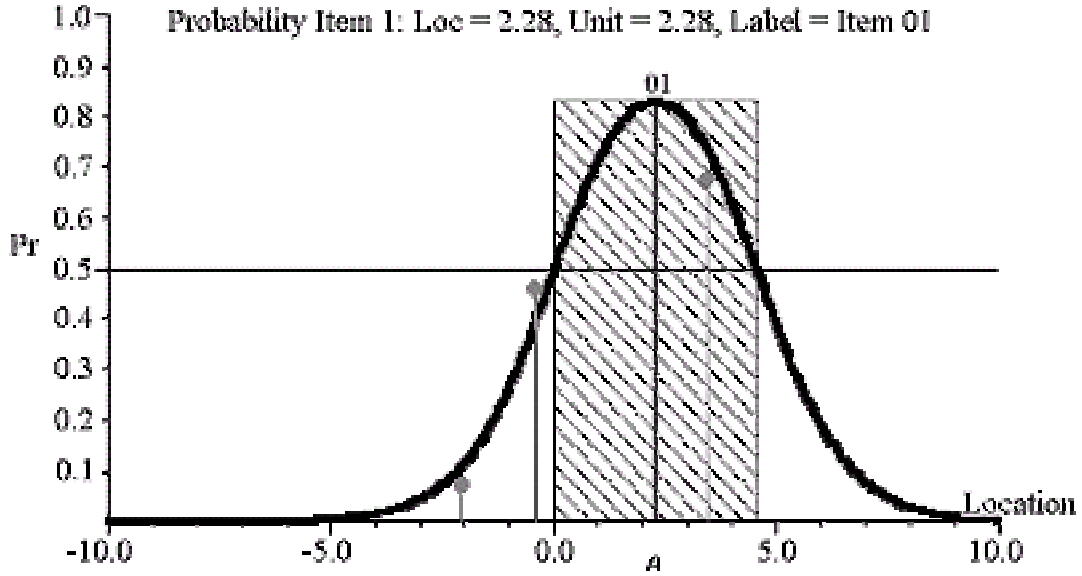


Figure 4. Observed responses, latitude of acceptance and location for Item 1 (Form A)

Figure 5 shows a plot of the average value for each group and the related observed proportion of positive responses for Item 14. Observed proportions for the third group are not close to the curve, indicating that this item does not fit the model. The rectangle presents the latitude of acceptance of the item, indicating that Item 14 is a negative statement.

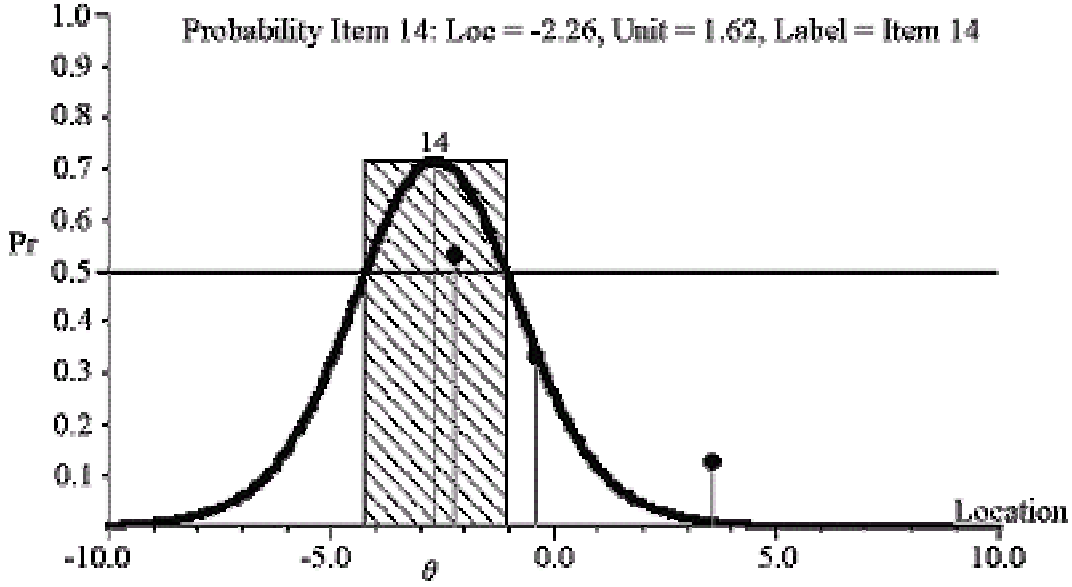


Figure 5. Observed responses, latitude of acceptance and location for Item 14 (Form A)

Figure 6 contains a plot of the average value for each group and the related observed proportion of positive responses for Item 20. Observed proportions for the first and second group are not close

to the curve indicating that this item does not fit the model. The rectangle presents the latitude of acceptance of the item and shows that Item 20 is a negative statement.

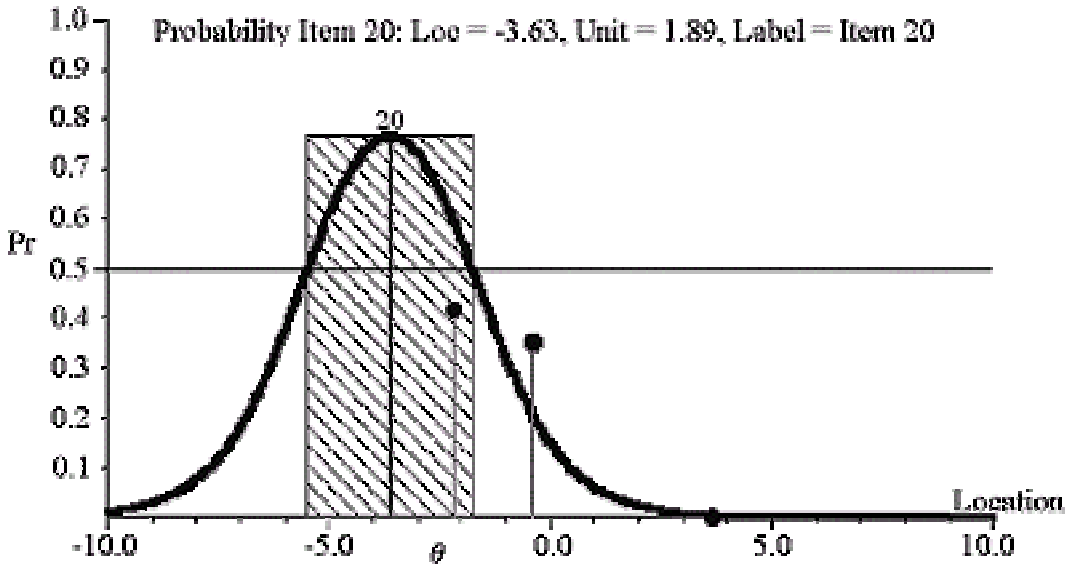


Figure 6. Observed responses, latitude of acceptance and location for item 20 (Form A).

For the three items presented above, the response patterns across the class intervals are monotonic. As the size of the difference between the locations of individuals and of items increases, the probability of an agree response decreases.

Table 9 contains a summary of the goodness of fit test while listing the items in location order. It is also interesting to note that the original order of items has been changed to reflect their location on the attitude continuum. From an analysis using a three class interval model, two statements, Item 14 and Item 20, with large fit statistics showed probability values of less than 0.05 thus misfitting the model. Item 14 has a single peaked structure, but it over discriminated relative to other statements such that the third interval has a lower number of observed disagreements than expected as well as a higher number of observed agreements than expected. Thus the item has 16 observed negative responses as opposed to 16 expected, and 18 observed positive responses as opposed to 20 expected from the first class interval, the group with the extremely negative attitude. The same item has 20 observed negative responses as opposed to 19 expected, and 10 observed and expected positive responses from the second class interval, the group with an ambivalent attitude toward the visual arts. However, the third interval, representing individuals with an extreme positive attitude toward the visual arts, gave 27 negative responses as opposed to the expected 30, and 4 positive responses as opposed to the expected 1. The latter inconsistency caused an inflated chi-squared value of 10.92 with a significant probability value of 0.00, indicating a significant difference between expected and observed responses in the third interval, and verifying a problem with the fit. The item location was -2.66 with a unit value of 1.62, demonstrating that it is a fairly negative statement placed closer to the lower end of the attitude continuum.

Item 20 was another item found to have similar problems. This item had 21 observed negative responses as opposed to 17 expected, and 15 observed positive responses as opposed to 19 expected from the first class interval. The same item had 24 observed negative responses as opposed to 29 expected, and 13 positive responses as opposed to 8 expected from the second class interval. Finally, the third interval gave 40 observed negative responses as opposed to 39 expected and 0 positive responses as expected. The second class interval presented a statistically

significant difference between the number of observed and the number of expected responses. The item location was -3.63 and had a unit value of 1.89, indicating that it was a highly negative statement placed close to the lower end of the attitude continuum. An inflated chi-squared value of 6.94 and the significant probability value of 0.03, demonstrated that Item 20 was sensitive at this extreme. From the chi-squared and probability values it was evident that all other items apart from Item 14 and Item 20 fitted the HCM. In order to investigate further the misfitting items, an adjustment was made in the model by increasing the number of class intervals from three to four (Table 9). It was observed that the chi square values for Items 14 and 20 did not appear inflated after this adjustment, but rather, all items presented a good fit.

Table 9. Summary of goodness of fit test for Form A with three and four class interval models

Item No.	Statements	Locat.	Unit	Three class interval			Four class interval		
				χ^2	Deg.	Sig.	χ^2	Deg.	Sig.
17	I see no reason for the government to spend money on aesthetic objects and activities.	-4.40	2.68	1.62	2	0.37	1.93	3	0.59
18	I have no desire to join or have anything to do with any organisation developed to visual arts activities	-4.16	2.85	2.53	2	0.22	1.17	3	0.76
16	I do not care for visual artists because their interests seem to me to be more emotional than rational.	-4.14	2.84	1.52	2	0.37	0.64	3	0.89
20	Visual arts education is nonsense.	-3.63	1.89	6.94	2	**0.03**	2.52	3	**0.47**
19	I see very little worth while in visual artistic interests.	-3.17	2.65	0.74	2	0.61	4.13	3	0.25
15	I do not believe I would receive any benefit from lectures concerning visual arts subjects.	-3.00	2.42	0.08	2	1.0	0.94	3	0.82
14	I believe that individuals engaged in purely visual arts occupations are parasites on society.	-2.66	1.62	11.41	2	**0.00**	4.11	3	**0.25**
13	Practical considerations should come first, visual arts second.	-2.55	3.20	2.79	2	0.22	4.58	3	0.21
11	Visual arts do not interest me now, but I expect that sometime I shall find time to pursue them actively.	0.14	0.36	2.15	2	0.37	4.20	3	0.24
10	I believe that the teaching of the visual arts subject is O.K., but the type of person now teaching it fails to 'get it across'.	0.86	2.25	0.67	2	0.61	1.86	3	0.60
12	I would be willing to give money to support visual arts enterprises if it were not for the 'highbrow' atmosphere surrounding them.	1.22	0.81	1.43	2	0.61	1.19	3	0.76
06	I believe the visual arts promote desirable relationships between nations.	1.76	2.02	0.93	2	0.61	1.00	3	0.80
01	It is in the visual artistic experiences of life that I find my greatest satisfaction.	2.28	2.28	3.60	2	0.14	2.63	3	0.45
05	I believe that the pursuit of visual artistic interests increase one's satisfaction in living.	2.30	2.24	3.08	2	0.22	4.62	3	0.20
07	I am attracted to individuals that pursue visual artistic interests.	2.65	2.30	2.11	2	0.37	3.85	3	0.28
04	Attendance at a visual arts exhibition gives me inspiration.	2.87	2.60	3.28	2	0.22	6.83	3	0.08
08	I believe that everyone should have a little training in the visual arts.	3.18	4.55	1.04	2	0.61	1.28	3	0.73
03	I am interested in anything in which I can see a visual artistic quality.	3.21	3.44	1.06	2	0.61	2.24	3	0.53
09	I am in favour of visual arts exhibitions for they do no harm to anyone.	3.47	5.07	1.91	2	0.37	2.28	3	0.52
02	I have a great interest in matters of the visual arts.	3.78	3.86	1.73	2	0.37	4.53	3	0.21
Total				50.62	38	0.07	56.52	57	0.49

Bold indicates misfitting items for Form A for three class intervals and the same items improvement in fit after adjustment was made by increasing the number of class intervals from three to four.

Figure 7 shows a plot of the statements' locations and respondents' frequency. An obvious gap was observed in the continuum, with items being polarised at either extreme corresponding to the negative and positive statements. Andrich (1996) has argued that when statements from Likert type questionnaires are located on a continuum, based on Thurstone's principles, they tend to be located at extremes with a gap in the middle, whereas Thurstone-type statements are required to be located more or less uniformly on the continuum. This contrasts with the present findings, where Thurstone-type positively worded statements were located uniformly across the continuum while negatively worded statements formed a gap between 0.00 and –3.00 logits. This indicated that there was an uneven distribution of items in the negative extreme. However, the locations of individuals on the same continuum were primarily between the locations of positive and negative statements. A slightly larger number of individuals was located among the negative statements than those among positive statements, suggesting that there were more individuals with a negative attitude than a positive attitude toward the visual arts.

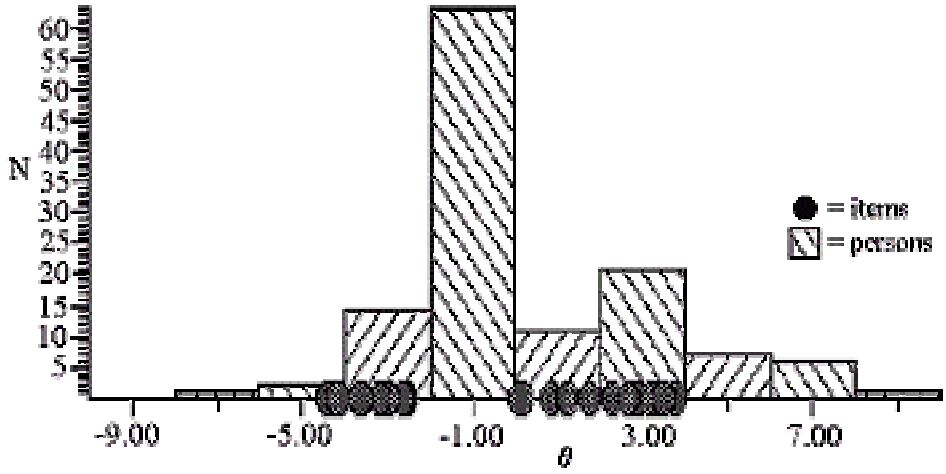


Figure 7. Frequency plot of persons with item locations for Form A (different units with correction)

The same procedure was repeated during the analysis of the data originating from Form B of the Visual Arts Attitude Scale. A sign analysis identified Item 4 and Item 12 as the two extreme opposite items on the attitude continuum with a correlation coefficient of -0.49. The correlations of items were exported with Item 4, being the extreme positive statement. Table 10 contains the correlation coefficients of the items with Item 4.

Table 11 contains the theta estimates of parameters resulting from the four estimation cycles, after which the item locations and item units were calculated. From these results, it appears that Items 5 and 6 have the largest latitude of acceptance with theta values of 4.44 and 4.10 respectively. These items have good discrimination indices as they have a broad bandwidth.

Table 12 shows the fit statistics for Form B. Overall, the items in this form have a good fit, although Item 6 presents a marginally high chi-squared value of 5.85 ($p = 0.049$).

Figure 8 contains the frequency plot of persons with item locations for Form B. This frequency plot revealed that gaps were located amongst the statements between +1.00 and +3.00 logits, 0.00 and –1.00 logit and –2.00 and –3.00 logits. This is a disturbing result considering the expected uniformity found in Thurstone-type item-distributions. In addition, contrary to observations made for data collected with Form A, the majority of persons (according to the statements endorsed from Form B) are located among the statements reflecting a positive attitude toward the visual arts and fewer persons are located among more negative statements. Finally, the equivalent

forms reliability coefficients were estimated, which indicated that with this sample of students, the two forms were operating in an equivalent way to the extent of a correlation of 0.76 between Form A and Form B and very similar Cronbach alpha coefficients of 0.87 and 0.85 for Forms A and B respectively.

Table 10. Correlation coefficients exported with Item 4 (Form B)

Item No.	Statements	Correlation coefficient
01	I find more satisfaction in artistic pursuits than in anything else.	0.37
02	I like art works because they give me genuine pleasure.	0.55
03	I believe that the great leaders of the world come from the ranks of those individuals who are artistically inclined.	0.28
04	Appreciation of the visual arts aids in making my life happier.	1.00
05	I believe that artistically sensitive people are fine people.	0.40
06	I believe that artistic pursuits are satisfying.	0.48
07	Artistic interests are not essential but make for happy existence.	0.49
08	I believe in the value of artistic interests but I do not like the stilted way in which the ideas on this subject are presented to me.	-0.05
09	I believe in the value of artistic interests but I seldom take time to pursue them.	-0.32
10	I go to such things as art exhibitions, installations etc., occasionally, but I have no strong liking for them.	0.07
11	Sometimes I believe that artistic interests are necessary and sometimes I doubt it.	-0.16
12	The visual arts do not play an especially large part in my life.	**-0.49**
13	I can enjoy the beauty of such things as paintings, sculpture and prints only occasionally, for I feel they are impractical.	-0.08
14	I have no interest in visual aesthetic objects (such as fine paintings and pottery) because I do not understand their technical aspects.	-0.32
15	I find the life of people pursuing artistic interests too slow and uninteresting.	-0.22
16	The 'highbrow' attitude of individuals having a great deal of artistic interest is quite distasteful.	-0.10
17	I believe that artistic interests are rarely genuine and sincere.	-0.08
18	It is hard for me to understand how anybody can be stupid enough to concentrate all his/her energies on artistic activities.	-0.24
19	Education in artistic things is a waste of public funds.	-0.39
20	The pursuit of visual arts activities is a sheer waste of time.	-0.36

Bold indicates the smallest correlation coefficient found between Item 4 and Item 12.

Table 11. Theta estimate parameters for Form B

Item No.	theta1	theta2	theta3	theta4	StdEr1	StdEr2	StdEr3	StdEr4
01	2.56	1.97	2.28	0.60	0.23	0.23	0.23	0.22
02	2.56	3.12	2.28	2.88	0.21	0.22	0.21	0.23
03	2.56	0.78	2.28	0.39	0.24	0.22	0.24	0.22
04	2.56	3.12	2.28	2.63	0.21	0.22	0.21	0.23
05	2.56	**4.24**	2.28	**4.44**	0.22	0.24	0.22	0.25
06	2.56	**4.13**	2.28	**4.10**	0.22	0.23	0.21	0.24
07	2.56	3.78	2.28	3.85	0.22	0.23	0.21	0.23
08	2.56	1.53	2.28	1.69	0.20	0.20	0.20	0.20
09	2.56	2.14	2.28	2.04	0.20	0.21	0.20	0.21
10	2.56	0.59	2.28	0.75	0.22	0.21	0.22	0.21
11	2.56	1.36	2.28	1.50	0.20	0.20	0.20	0.20
12	2.56	3.71	2.28	3.33	0.20	0.22	0.20	0.23
13	2.56	1.20	2.28	1.30	0.20	0.20	0.20	0.20
14	2.56	3.43	2.28	2.45	0.21	0.22	0.21	0.22
15	2.56	2.27	2.28	1.94	0.22	0.22	0.22	0.23
16	2.56	2.38	2.28	2.03	0.21	0.22	0.21	0.22
17	2.56	1.70	2.28	1.60	0.21	0.22	0.21	0.22
18	2.56	2.85	2.28	2.48	0.23	0.24	0.24	0.24
19	2.56	3.95	2.28	3.00	0.26	0.25	0.25	0.25
20	2.56	2.92	2.28	2.54	0.23	0.23	0.23	0.24

Bold indicates items with the largest latitude of acceptance

Table 12. Summary of goodness of fit test for Form B with three and four class interval models

Item No.	Statements	Loc.	Unit	Three class interval		
				χ^2	Deg.	Sig.
19	Education in artistic things is a waste of public funds.	-4.10	3.00	1.44	2	0.61
18	It is hard for me to understand how anybody can be stupid enough to concentrate all his/her energies on artistic activities.	-3.42	2.48	0.12	2	1.00
20	The pursuit of visual arts activities is a sheer waste of time.	-3.29	2.54	3.51	2	0.14
15	I find the life of people pursuing artistic interests too slow and uninteresting	-2.29	1.94	2.86	2	0.22
14	I have no interest in visual aesthetic objects (such as fine paintings and pottery) because I do not understand their technical aspects	-2.25	2.45	0.83	2	0.61
16	The 'highbrow' attitude of individuals having a great deal of artistic interest is quite distasteful.	-2.11	2.03	4.74	2	0.08
12	The visual arts do not play an especially large part in my life.	-1.96	3.33	0.05	2	1.00
17	I believe that artistic interests are rarely genuine and sincere.	-1.54	1.60	2.01	2	0.37
09	I believe in the value of artistic interests but I seldom take time to pursue them.	-1.01	2.04	3.77	2	0.14
13	I can enjoy the beauty of such things as paintings, sculpture and prints only occasionally, for I feel they are impractical	0.02	1.30	0.19	2	1.00
08	I believe in the value of artistic interests but I do not like the stilted way in which the ideas on this subject are presented to me.	0.27	1.69	2.04	2	0.37
11	Sometimes I believe that artistic interests are necessary and sometimes I doubt it.	0.45	1.50	0.53	2	0.61
10	I go to such things as art exhibitions, installations etc., occasionally, but I have no strong liking for them	0.87	0.75	0.96	2	0.61
03	I believe that the great leaders of the world come from the ranks of those individuals who are artistically inclined	1.05	0.39	0.26	2	1.00
01	I find more satisfaction in artistic pursuits than in anything else.	1.53	0.60	0.51	2	0.61
02	I like art works because they give me genuine pleasure.	3.41	2.88	4.56	2	0.08
05	I believe that artistically sensitive people are fine people.	3.42	4.44	1.58	2	0.37
07	Artistic interests are not essential but make for happy existence	3.57	3.85	2.93	2	0.22
04	Appreciation of the visual arts aids in making my life happier.	3.65	2.63	0.80	2	0.61
06	I believe that artistic pursuits are satisfying.	3.75	4.10	5.85	2	0.05
Total				39.54	38	0.32

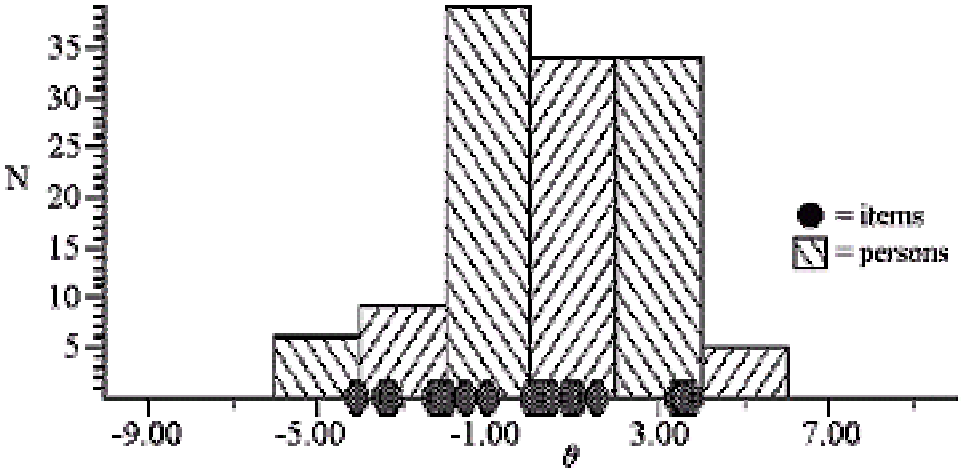


Figure 8. Frequency plot of persons with item locations for Form B (different units with correction)

Discussion

Binary graded disagree-agree responses to attitude items for the purpose of attitude measurement toward the visual arts, were collected and examined using three methods of analysis. The first method used was exploratory factor analysis in an attempt to determine the factor structure and provide a framework for understanding the network of relationships among measures included in the scale. The exploratory factor analysis did not provide adequate information to facilitate interpretation of data. Conversely, such traditional cumulative methods of analysis of responses to Thurstone scales have proved to be misleading, giving the impression of statements with respect to a construct are on separate dimensions (Van Schuur & Kiers, 1994; Van Schuur & Kruijtbosch, 1995). The second method used was Rasch analysis of the Attitude Scale, particularly employing the partial credit model. Results obtained from the analyses implied an additional difficulty in explaining non-fitting items located in the middle of the attitude continuum. Observed non-fitting items appeared in the centre of the scale, indicating a difficulty in discriminating between positive and negative attitudes towards the visual arts. Because of the necessity of identifying and interpreting students' negative responses to central items on the continuum, a third analysis was conducted using the Hyperbolic Cosine Model (Andrich & Luo, 1993) for unfolding single stimulus responses.

The HCM calculated responses to follow the Rasch model for three categories. Thus, the manifest dichotomous response (Agree/Disagree) was resolved into three latent responses – Disagree Below (DB) because a person was below the statement, Disagree Above (DA) because the person was above the statement and Agree (A) because the person was close to the statement. In this HCM analysis, an unfolding model was applied to the latent responses, in which a direction was implied. Two response curves which characterised the directions of the Disagree response, were unfolded to correspond to the data. So that the dichotomous observed data is unfolded by the HCM, in order to explain the Disagree response in terms of its constituent elements. The HCM described an effect that was expected as a function of the distances of persons from the statements. Furthermore, from the location estimate of each person and the associated standard error, an index of (person) separation was constructed, which is an indicator of items being answered either inconsistently by respondents, or of items measuring different incoherent attributes. From the results obtained using the HCM, items that did not fit this model were detected, while the location of each item and each person on the attitude continuum was clearly interpreted. Item locations and unit values were of particular interest in demonstrating a lack of uniformity in certain areas of the distribution of statements along the continuum. However, the results did provide person location on the attitude continuum, hence an accurate determination of an individuals' degree of affect toward the visual arts was derived, which may be used for further analyses in multilevel and multivariate studies. The information obtained supported the underlying unidimensionality of the measurement model employed and was particularly useful in the selection of items that in fact fitted the model and had good discrimination indices. Furthermore, the item location could be accurately estimated and thus statements could be put in location order on the scale.

Single-peaked response functions have been rarely used in substantive research although modern computing algorithms have overcome the problems of time-consuming and laborious traditional analysis of single-peaked response data. This study has demonstrated how a single-peaked response function can be employed for the development of a scale of attitude measurement toward the visual arts. Nevertheless, single-peaked response functions have further implications for the investigation of stage developmental constructs (Inhelder & Piaget 1958), cognitive stages and the

application of models in psychological, social and educational development and measurement (Leik & Matthews, 1968; Birnbaum, 1968; Coombs & Smith, 1973; Andrich & Styles, 1998).

References

Adams, R. J & Khoo S. T. (1993) *QUEST: The Interactive Test Analysis System.* Hawthorn, Victoria: Australian Council For Educational Research.

Andrich, D. (1978) Relationships between the Thurstone and Rasch approaches to item scaling. *Applied Psychological Measurement*. 2, 449-460.

Andrich, D. (1982). An extension of the Rasch model for ratings providing both location and dispersion parameters. *Psychometrika.* 47, 105-113.

Andrich, D. (1995). Distinctive and Incompatible Properties of the two common classes of the IRT Models for graded responses. *Applied Psychological Measurement,* 19(1), 101-119.

Andrich, D. (1996). A hyperbolic cosine latent trait model for unfolding polytomous responses: Reconciling Thurstone and Likert methodologies. *British Journal of Mathematical and Statistical Psychology*. 49, 347-365.

Andrich, D., & Luo, G. (1993). A hyperbolic cosine latent trait model for unfolding dichotomous single-stimulus responses. *Applied Psychological Measurement.* 17, 253-276.

Andrich, D., de Jong, J. H. A., & Sheridan, B.E. (1997). Diagnostic opportunities with the Rasch model for ordered response categories. In J. Rost & R.Langeheine (Eds.). *Applications of latent trait and latent class models in the social sciences* (pp. 58-68). Munster, Germany: Waxmann Verlag.

Andrich, D., & Styles, I.M. (1998). The structural relationships between attitude and behaviour statements from the unfolding perspective. *Psychological Methods.* 3(4), 454-469.

Bock, R.D. (1972). Estimating item parameters and latent ability when responses are scored in two or more nominal categories. *Psychometrika.* 37, 29-51.

Bock, R.D., & Jones L.V. (1968). *The measurement and prediction of judgement and choice.* San Francisco: Holden Day.

Birnbaum, A. (1968). Some latent trait models and their use in inferring an examinee's ability. In F.M. Lord & M.R. Novick, *Statistical Theories of Mental Test Scores* (pp. 397-479). Reading MA: Addison-Wesley.

Cohen, J.A. (1941). A scale for the measurement of attitude toward the aesthetic value. *J.Psychol.*, 12, 75-79.

Coombs, C. H. (1964). A *Theory of Data.* New York: Wiley.

Coombs, C.H., & Smith, J.E. (1973). On the detection of structure in attitudes and developmental process. *Psychological Review,* 80, 337-351.

Donoghue, J. R. (1994). An examination of the IRT information of polytomously scored reading items under the generalised partial credit model. *Journal of Educational Measurement.* 41, 295-311.

Dubois, B., & Burns, J.A. (1975). An analysis of the meaning of the question mark response category in attitude scales. *Educational and Psychological Measurement.* 35, 869-884.

Guttman, L. (1950). The basis for scalogram analysis. In S.A. Stouffer, L. Guttman, E. Suchman, P.E. Lazarsfeld, S.A. Star, & J.A. Gardner (Eds.). *Measurement and prediction* (pp. 60-69). Princeton NJ: Princeton University Press.

Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence. London*: Routledge & Kegan Paul.

Leik, R.K., & Matthews, M. (1968). A scale for the development processes. *American Sociological Review,* 33, 62-75.

Likert, R. (1932). *A technique for the measurement of attitudes.* Archives of Psychology: No 140.

Masters, G.N. (1982). A Rasch model for partial credit scoring, *Psychometrika.* 47, 149-174.

Roberts, J.S. (1995). *Item response theory approaches to attitude measurement.* (Doctoral dissertation, University of South Carolina, Columbia, 1995). Dissertation Abstracts International. 56, 7089B.

Roberts, J.S., Laughlin, J.E.M & Wedell, D.H. (1999). Validity issues in the Likert and Thurstone approaches to attitude measurement. *Educational and Psychological Measurement,* 59 (2), 211-233.

Silance, E.B., & Remmers, H.H. (1934). An experimental generalised master scale: A scale to measure attitudes toward any school subject. *Purdue Univ. Stud. High. Educ.*, XXVI, 35, 84-88.

Thissen, D.M. (1976). Information in wrong responses to the Raven's Progressive Matrices. *Journal of Educational Measurement.* 13, 201-214.

Thurstone, L.L. (1928). Attitudes can be measured. *The American Journal of Sociology*. 33, 529-554.

Thurstone, L.L. (1928). The measurement of social attitudes. *Journal of Abnormal and Social Psychology*. 26, 249-269.

Van Schuur, W. H., Kiers, H.A.L. (1994). Why factor analysis is often the wrong model for analysing bipolar concepts and what model to use instead. *Applied Psychological Measurement.* 18, 97-110.

Van Schuur, W.H., & Kruijtbosch, M. (1995). Measuring subjective well being: Unfolding the Bradburn Affect Balance Scale. *Social Indicators Research.* 36, 49-74.

The Chinese Value Survey: An interpretation of value scales and consideration of some preliminary results

Barbara Marshall Matthews
Flinders University, School of Education *bobbie.matthews@flinders.edu.au*

The Chinese Value Survey (CVS) was developed by Bond and his colleagues as a complement to survey instruments constructed by research workers such as Rokeach (1973) and Schwartz (1992). It was designed to be used with people living in geographical regions where Eastern life values are pre-eminent. Preliminary studies have been carried out using this instrument with university students from ethnically Chinese backgrounds studying in three Australian universities. Data were examined using principal components analysis rotated to orthogonal structure. Initial results indicate that of the 40 values measured, 39 neatly form four factors, which are renamed to suit their content. This analysis makes interpretation of the values held by students from an ethnically Chinese background more accessible.

Background literature

The Chinese Value Survey (henceforth referred to as the CVS) was developed by Bond and a group of research workers known collectively as the Chinese Culture Connection in response to their perceived need to measure and evaluate cultural values within the setting of a Chinese social value system that is derived from the Confucian ethos. Value perceptions from those who were ethnically Chinese which were obtained by Bond and his colleagues seemed very different from those obtained in previous studies. This is especially evident in Chinese students' responses in Western value studies by Rokeach (1973) and subsequently Schwartz (1992,1994), when compared to these alternative instruments measuring Eastern life values (Bond, 1983; Bond, 1991; Bond and Cheung, 1983; Bond, Cheung, and Wan, 1982; Bond and Forgas, 1984; Bond, Leung and Schwartz,1992; Bond and Pang, 1991; Chan, 1988; Feast and Churchman, 1997; Leung and Bond, 1989; Slay, 1999a and Slay, 1999b). Interpretations of student responses appear to lack validity when related to Western values. Consequently, Bond and his colleagues have set about designing a questionnaire that is non-Western in nature. The CVS is the result. It contains a decidedly Chinese cultural bias not previously assessed in other Western value surveys.

The underlying human universals measured by Bond and his co-workers do not fit with any extant survey evaluating Western value responses. The values measured by Bond and his colleagues remain universal in nature, but they also include some values which are uniquely Confucian (Hofstede, 1991). Examples are respect for tradition, humility, filial piety, and protecting one's face. This is not to say such values do not exist in Western culture, but rather their implied importance is not as great as in Eastern culture. In Western culture the search is for truth in life values, people from Eastern cultures search for virtue which comes from the teachings central to an understanding of Confucian teachings (Hofstede, 1991).

The Confucian ethos has tended towards the formation of a dedicated and motivated as well as educated population which manifests a responsible and enhanced sense of commitment to its ideals and institutions as well as to the identity and moral fibre of its organisation with the resultant establishment of desirable outcomes (Chang et al.,1997; Chen, Liu and Ennis, 1997; Chinese Culture Connection, 1987; Chiu, 1989 and Ralston et al.,1992). The CVS has been

instrumental in enabling the validity of these constructs to be tested and confirmed. Previously these value constructs lay buried under Western perceptions of their lack of importance in Chinese culture (Bond & King, 1985).

In the CVS consideration is focussed on four dimensions or factors and the components of these factors. This is necessary when it is found that the Western value responses to the Rokeach and Schwartz value questionnaires are markedly different from the responses of students from ethnically Chinese backgrounds among whom Eastern life values are pre-eminent. Separate consideration is needed to be given to those values which are related to the Eastern value system. This examination involved life values empirically designated as oriental or Chinese. The uniqueness of these values emerges as a separate but integral system. Hence the Chinese value system co-exists compatibly with the Western way of thinking but does not become subsumed in that way of thinking. Thus both the Western and Eastern values co-exist without conflict. Therefore, it is those values where there is a divergence of opinion as to their importance within a given society, that are considered in detail. Where overlap with Western thinking, in particular, occurs, little consideration is given as detailed analysis is extant in the literature. Further, as the Chinese value system has been previously invisible using Western instruments of value measurement, greater analytic consideration and thought have been given to this latest set of data from the CVS (Bond, 1991).

The nature of these differences warrants consideration. In the People's Republic of China (henceforth referred to as PRC), a system of *guanxi* or using connections to obtain something, is integral to the very existence of society itself. This is often viewed by Western eyes as a form of corruption but, in fact, it forms a vital network of social relationships essential in a population of 1.3 billion people. Such an enormity of people is beyond comprehension in a small Australian city until it is realised that without such a nexus of intermingling networks, survival would be impossible. Moderation or following the middle way is another value leading to harmony and stability. These values support the collective rather than the individualistic way of thought and strengthen the construct of unity within the extended family unit. The particular value that has raised the greatest controversy because of its overtly contradictory nature is non-competitiveness which is classified here as part of the Confucian ethos. Many perceive ethnic Chinese as one of the most competitive of all peoples. This may, in fact, be true. However, self and other perceptions seem to be very different. It is considered in very poor taste among Chinese people to condone aggressive competition. Such aggression is integral to sport and attainment of goals in a Western individualistic society, as is contributing to the sustenance of the family unit in Eastern societies. Therefore, such outward competition and rivalry is considered a negative influence and manifestly rejected. Yet in business it would appear to be an acceptable course of action. Once again it may well be Western to evaluate an Eastern homeostatic mechanism in this way.

Methods of Analysis

The sample surveyed in this investigation has been derived from students studying at the three universities in South Australia. All members of the sample are sojourners to Australia most of whom plan to return to their home countries on completion of their studies. Their age and length of time in Australia varies and will be analysed as factors in relation to the other data when collection of information is completed late in the year 2000. These data are the initial part of a longitudinal study and represent a departure from previously reported analyses.

Students were asked to respond to the CVS using a Likert type scale of 1 to 9. One (1) represented a value which was '*of no importance to me'* and nine (9) '*of supreme importance to me*.' Students were asked to respond as a pen and paper exercise to each listed value using an

answer sheet. Future data are being collected using the internet and a web site. The questionnaires are presented in both English and Chinese and located in an easily accessible web site which the students are able to locate with a normally configured computer. (A copy of the complete questionnaire may be found in the Appendix.)

The data were analysed using principal component analysis which was rotated orthogonally using SPSS (Version 8). The results obtained are presented in Table 1.

Table 1. CVS factor values and description (N=151)

CVS-I- INTEGRITY &TOLERANCE DEVELOPMENT OF SELF	**CVS-II—CONFUCIAN ETHOS** RELATIONSHIPS WITH OTHERS
1. Filial piety (0.50)	6. Loyalty to superiors (0.60)
2. Industry –working hard (0.56)	16. Benevolent authority (0.63)
3. Tolerance—of others (0.52)	17. Non-competitiveness (0.50)
4. Harmony—with others (0.70)	22. Keeping oneself disinterested and pure (0.43)
5. Humbleness (0.42)	33. Contentedness with one's position in life(0.63)
9. Kindness (forgiveness, compassion (0.67)	34. Being conservative (0.58)
10. Knowledge (education) (0.62)	35. Protecting your 'face' (0.53)
13. Self-cultivation (0.57)	37. Chastity in women (0.55)
15. Sense of righteousness (0.48)	38. Having few desires (0.66)
18. Personal steadiness & stability (0.61)	39. Respect for tradition (0.62)
21. Sincerity (0.74)	
24. Persistence (perseverance) (0.73)	
25. Patience(0.77)	
28. Adaptability (0.60)	
29. Prudence (carefulness) (0.40)	
30. Trustworthiness (0.65)	
32. Courtesy (0.62)	
VARIANCE ACCOUNTED FOR BEFORE ROTATION	
8.70	28.70
VARIANCE ACCOUNTED FOR AFTER ROTATION	
11.91	18.91
RELIABILITY	
0.82	0.91

CVS-III-LOYALTY TO IDEALS & HUMANITY SOCIAL RESPONSIBILITY	**CVS-IV-MODERATION & MORAL DISCIPLINE** WORLDLY WISDOM
7. Observation of rites & social rituals (0.42)	26. Repayment of good or evil of others(0.53)
8. Reciprocation of greetings, favours (0.53)	27. Sense of cultural superiority (0.67)
11. Solidarity with others (0.65)	40. Wealth (0.69)
12. Moderation-following the middle way (0.47)	
14. Ordering relationships by status (0.64)	
19. Resistance to corruption (0.54)	
20. Patriotism (0.63)	
23. Thrift (0.50)	
31. Having a sense of shame (0.43)	
VARIANCE ACCOUNTED FOR BEFORE ROTATION	
5.05	4.52
VARIANCE ACCOUNTED FOR AFTER ROTATION	
11.00	5.28
RELIABILITY	
0.82	0.57

Results

This factor structure was tested using principal components analysis. There were 10 Eigenvalues that were greater than one which accounted for 66 per cent of the variance. A scree test was performed to estimate the number of distinct factors. Three factors were clearly evident from the plot and a fourth seemed possible. A varimax rotation of four factors was performed and 39 out

of the 40 values had loadings of greater than 0.40. All loadings were positive. The value statement, 'a close, intimate friend' was excluded from further analysis as its greatest loading on the four rotated factors was found to be less than 0.30.

In the current analysis, the repositioning of value statements is necessary because of the factorial structure arrived at in the analysis. The foci have changed and the position of the values has altered. In light of the realigned content, new and more definitive names for the factors are justified.

The scales are named as follows:CVS I: Integrity and Tolerance, CVS II: Confucian Ethos, CVS III: Loyalty to Ideals and Humanity and CVS IV: Moderation and Moral Discipline. CVS I has 17 value statements, CVS II has 10, CVS III has nine and CVS IV has three value statements.

The reliabilities (Kaiser and Caffrey, 1965) found are respectively: CVS I (Integrity and Tolerance): 0.82, CVS II (Confucian Ethos): 0.91, CVS III (Loyalty to Ideals and Humanity): 0.82 and CVS IV (Moderation and Moral Discipline): 0.57.

In Bond's writing (Bond, 1991) the scale that has caused the greatest discussion is the Confucian Ethos. It differs most from the Western conceptualisation of values. The Confucian values represent cornerstones of Chinese morality. Filial piety, for instance, can be traced directly to Confucius' thinking and is referred to in *Five Cardinal Relations.* This value is still an essential contemporary construct, engendering parental respect as well as respect for tradition, honouring one's ancestors and financial support of parents. The reciprocation of greetings and favours and protecting of 'face' are also still very important in a society which operates from an inner core of close relationships which, in turn, are surrounded by more casual but still significant personal links (Bond, 1991). This is analogous to a series of concentric circles like the layers of an egg with the family represented by the yolk.

The realignment of the value statements is necessary following the factor analyses conducted in this study. Aside from the Confucian factor, the largest number of values is associated with the concept of Integrity and Tolerance. It includes value statements which promote the development of the social person. Social stabilisation acts as the umbrella concept. Therefore, values such as knowledge (education), self cultivation, personal steadiness and stability, patience and adaptability are included. The third factor: Loyalty to Ideals and Humanity in the individual's contact with others is seen as the embodiment of social responsibility. This factor promotes moderation and following the middle way, the ordering of relationships by status, the observation of rites and rituals, resistance to corruption and a solidarity with others. The latter three values reinforce a collectivistic approach to harmonious group maintenance and living in a peace-seeking manner. Acceptance of this grouping of values shifts the meaning of the factor itself toward the importance of group cohesion because of the heightened sense of social responsibility.

The final factor represents both moderation as well as moral discipline. Moderation, including wealth and a sense of cultural superiority promote an aura of worldly wisdom as does the construct of repayment of good or evil to others. Self-discipline is reflected in the self-discipline and implied self-control needed to maintain these values positively.

These factors vary in name and composition from those enumerated by Bond and his colleagues in their original evaluation of the CVS by the Chinese Culture Connection (CCC, 1987). The CCC had identified the scales as follows: CVS I: Integration, CVS II: Confucian Work Dynamism, CVS III: Human-Heartedness and CVS IV: Moral Discipline. CVS I was seen to reflect values which focus on social stability, strong family bonding and chastity in women, all of which indicate the importance of family. Filial piety and patriotism although negatively correlated, were included

in this factor. CVS II contained values which all reflected the Confucian work ethic. In contrast to this hierarchical dynamism were four negatively correlated values: reciprocation of good and evil in others, personal steadiness, protection of face and respect for tradition. These latter values advocated checks and distractions at the personal, interpersonal and social levels (Chinese Culture Connection, 1987). CVS III contained the value statements which embraced ideas suggesting gentleness and compassion as contrasted with a sterner, legalistic approach to life. In CVS IV the positively loaded values reflected a moral restraint, in contrast to the negatively loaded values, adaptability and prudence which implied a position lacking such self-control. In Bond's study moderation was seen to represent a firm and disciplined stance rather than the flexibility normally associated with 'following the middle way' (Chinese Culture Connection, 1987).

Discussion

Comparative Data Analysis

Comparing students in PRC and Taiwan, Yuan & Shen (1998) found many shared values and concerns for the world in which they lived. They found traditional Confucian values still extant in rural societies which today represent 75-80 per cent of the population in PRC compared with Taiwan. However, those young people living in urbanised areas considered these values somewhat antiquated but still acceptable to guide and mould the preferences of young people. Although embracing the new and modern values from Western culture, older, more stable values still pervade in the moral education and training of the young. Once again, the CVS enables a greater understanding of these perceptions in the minds of young Chinese as the CVS measures values relevant to their thinking rather than their evaluation of Western values. It has given credence to the concept of the *tao* or 'the way' which has led, guided and still transcends Chinese culture today. The feudalistic social order espoused by Confucius (Ralston et al., 1992) has merely been replaced by the Communist social order. Therefore, Confucian values remain intact today in spite of, or despite, attempts to discredit his teachings. These values operating in a massive unwieldy bureaucracy have survived and even flourished as they promote the well being of the basic unit of the family and work place alike. The work unit is considered to be both a link and a control over the working men and women in PRC today. Basic Confucian ideals and values have not changed, only altered or adapted to fit the newer model of the political state.

Comparing results in different cultures and their relationship to Confucian values

Bond (CCC, 1987 and Bond, 1988) in the analysis of the data collected was looking for transcultural trends. He analysed data at the collective or country level rather than the individual level with both discriminant analysis and factor analysis with very high factor loadings which he compared to trends in gross national product as a measure of socio-economic status. Such methods of analysis are vulnerable to aggregation bias. In the present study, inter student comparisons are undertaken and a measure of individual change over time is also envisioned. The students used in the sample are sojourners, as already noted, and plan to return to their home country after finishing their studies in Australia. Many have also indicated a desire to do further study after completion of their present degree. Some have even expressed the hope of gaining work experience after finishing their studies but none has contemplated migration at this point in time. It remains to be seen if these ideas will change over time. The results obtained, as with most preliminary study analyses, have raised further issues for consideration and it is hoped that additional questions may satisfactorily resolve some of the issues raised.

Chiu et al. (1998) studied Chinese business people with regard to conflict behaviour, aggression and the acceptance of traditional Confucian values. They found, as has been found in personal

dealings with Chinese teachers and students, that conflict is avoided if at all possible. The Chinese tend to avoid assertive or competitive behaviour as it produces conflict and a lack of harmony and is socially unacceptable. Sublimation of personal goals to the collective whole is helpful in promoting values such as harmony, humility, courtesy, patience, obedience and modesty, so important in the development of self and maintenance of the group collectivist ethic evident in inter-student relationships, especially in a Chinese university situation when six to eight students, on the average share, a single dormitory room. These values have formed the basis of the Confucian ethos which even today, despite political upheaval and social change, has emerged as the dominant cognitive construct in ethnically Chinese societies. The self control and strong self-effacement and discipline required to avoid conflict are difficult for the Western mind set to comprehend much less put into practice (Chiu et al., 1998). Retention of traditional beliefs and cultural values as primal cohesive factors are considered to be essential in countries following collectivistic patterns of life.

It was interesting to note that when interviewing the sample in this study, most students preferred to live in groups with people from the same country or culture rather than alone even though they preferred to study alone rather than in a group. From the demographic questionnaire it was found that 73 per cent preferred to study alone and 49 per cent preferred to study at home where they cannot readily confer with friends or colleagues compared to 40 per cent in the library where they can confer with others. Only 24 per cent have family here in Adelaide but 54 per cent speak their home language when away from their university settings and another five per cent speak a language other than English when outside of the study environment. Chan (1988) found that students adapted to the Australian way of life have values and aspirations closer to those of Australian students than to other, more recently arrived Chinese students. Migrants appear willing to forego their own cultural values in order to avoid conflict in their anxiety to assimilate. This is not found among sojourners where Confucian values and concomitant academic success appear to be a more important mediator of behaviour (Chan 1988). The students would appear to identify with the Chinese rather than the Western ethos even more than Asian migrants and refugees to this country. This is best understood in light of a common cultural heritage even though the actual country of origin may be different.

In consideration of the value factors already mentioned, allusion to Confucian heritage has been made. An elaboration of the dominating cultural structure may be helpful here. In traditional Chinese society, the parent-child, family and neighbour relationships are built on blood and marriage (Chiu 1989). In the past, families lived in the same area of a town or city and formed tight networks that were rarely broken. Traditionally it is difficult to separate self and group identity and the collective is preferred to the individual in parent-family-neighbour interaction. Friendship and work co-operation are voluntary in group formation in that members are not required to relinquish individual identity but co-operation is expected in order to attain social and functional goals. Over the past 30 years modernisation has been perceived as a weakening influence on traditional family relationships. It has also raised the importance of voluntary associations on interpersonal behaviour, making co-operation and self-reliance more important than belonging to involuntary groups like the family where people define their social identity. Such changes tend to enhance the concept of Chinese collectivism as being highly focussed whereby some expectations are enforced and others are not. In Hong Kong and the other ‘little dragons’ countries of Japan, Korea, Macau Malaysia, Singapore, Taiwan, , as well as PRC, there has been rapid economic development and industrialisation which have tended to reduce the significance of the collective when compared with individual behaviour. These external changes have led to changes in the socialisation process and, in turn, they have produced significant changes in the endorsement and inculcation of traditional Chinese values (Chiu, 1989). There are now many nuclear families living

in large cities and even greater evidence of change with the one child family policy practised out of necessity in PRC. These changes have, in turn, led to the weakening of traditional famialism and the growing importance of the individual and the way in which cultural values are operationalised. This gives a clearer understanding of a personal-cultural identity, an important component of which is a person' self-concept. However, the major component of self-concept is the individual's social group membership which determines social identity.

Bond and King (1985) depict Hong Kong, ethnically 98 per cent Chinese, as it emerged from cultural ruralism and colonisation by the British to its status today as a modern industrialised arm of PRC. For the Chinese, 'China was the world' *(tian xia).* Even today it is often referred to as 'the Middle Kingdom' *(zhong guo)*, the implication being that China is the centre of the world. In the past, educated Chinese have seen no need to go beyond its borders as they, in limited exploration, have found no other civilisation better developed than their own. Any invading idea that might improve their lifestyle has been simply adopted (Bond & King, 1985). The foreigners living there are called in Cantonese *gui lao (*and in Mandarin *yang guizi)* or foreign devils, to be tolerated and worked with but never assimilated. This extends to the Hong Kong and other Chinese interpretations of Westernisation and modernisation. It is modernisation which involves technology, behaviour and material progress whereas Westernisation involves values, thinking or traditions originating in the West. In this way, Hong Kong Chinese can preserve their interior Chinese domain as culturally inviolate by rejecting Westernisation while simultaneously embracing modernisation. This distinction allows modern Chinese to adopt what they wish from the West while still preserving essentially Confucian values such as filial piety, thrift and respect for the individual's teacher (Bond & King, 1985). The bifurcation created by these definitions has allowed technical modernisation to be applauded while condemnation of 'spiritual pollution' of the culture continues. In conversations with local Hong Kong Chinese prior to 1 July 1997, none seemed apprehensive. After all, they said, 'we are all Chinese and that is all that matters.'

Different results using similar measurement instruments at one point in time

Guan and Dodder (1998) have conducted a comparative study looking at Chinese students in American universities and Chinese students in PRC. They have considered the impact of cross-cultural contact on values and identity. Guan and Dodder (1998) have measured values using the CVS and individual interviews to assess the importance of cross-cultural contact. They have found that two principal forces, length of time away from their home culture and age are important influences on change. Two years would seem to be the time barrier for change. Within two years the students tend to retain those values which enable them to maintain personal cultural identity within a Western ambience. However, Guan and Dodder (1998) found that the longer the students were away from PRC, the greater the change in values. Two hypotheses have been advanced for this study. A paraphrasing of these will assist understanding of the investigations being conducted. The first states that overseas Chinese students (sojourners) in Australia staying longer view cultural values as less important than newer arrivals. Secondly, it is hypothesised that Chinese students in Australia tend to maintain certain Chinese values which may help them to sustain their cultural identity in a new cultural milieu.

Guan and Dodder (1998) also used principal components analysis and then rotated the factors and identified value statements with loadings greater than 0.40 orthogonally. Only 18 of their value statements had high enough loadings whereas 39 out of the 40 had high loadings in this study. Guan and Dodder also rearranged and renamed the value structure as was done here for the sake of clarity of meaning and contextual cohesion. It would be of interest to see if length of stay changed the position and/or values obtained for the variables measured here in Australia. The study cited was conducted in the United States and PRC and measured at a single point in time whereas the

present study is a longitudinal one, in order to measure change over time. It remains to be seen if those values which are peculiarly Chinese, those representing what is here called the Confucian ethos, change greatly with the passage of time and whether the change is sufficient to alter the value structure described above.

References

Anderson, J.F. & Berdie, D. R. (1974). *Questionnaire: Design and Use*. Metuchen: The Scarecrow Press.

Banks, J.A. (1993). The canon debate, knowledge construction and multicultural education. *Education Researcher*, 22(5), 4-14.

Barker, M. (1990). Cross-cultural communication at Australian universities: issues faced by learners, ethnic Chinese and Australian students, (p 259). In Hedrick, C. & Holten, R. (eds.) *Cross-cultural Communication and Professional Education*. Adelaide: Centre for Multicultural Studies

Bond, M. H. (1983) How language variation affects inter-cultural differentiation of values by Hong Kong bilinguals. *Journal of Language and Social Psychology*, 2(1), 57-67.

Bond, M. H. (1988). Finding dimensions of individual variation in multicultural studies of values: the Rokeach and Chinese value surveys. *Journal of Personality and Social Psychology,* 55(6), 1009-1115.

Bond, M. H. (1991) .Chinese values and health: a cultural level examination. *Psychology and Health: an International Journal*, 5, 137-152

Bond, M. H. & Cheung, T. S. (1983). College students spontaneous self-concept: the effect of culture among respondents in Hong Kong, Japan and the United States. *Journal of Cross-cultural Psychology*, 14 (2), 153-171.

Bond, M. H. Cheung, K. & Wan. K. C. (1982). How does cultural collectivism operate: the impact of task and maintenance on reward distribution. *Journal of Cross-cultural Psychology,* 13(2), 186-200.

Bond, M. H. & Forgas, J. P. (1984). Linking perception to behaviour intention across cultures: the role of cultural collectivism. *Journal of Cross-cultural Psychology*, 15(3), 337-352.

Bond, M. H. & King, A. Y. C. (1985). Coping with the threat of westernisation in Hong Kong. *International Journal of Intercultural Relations*, 9, 351-364.

Bond, M. H., Leung K. & Schwartz, S (1992). Explaining chores in procedural and distributive justice across cultures. *International Journal of Psychology*, 27(2), 211-225.

Bond, M. H. & Pang, M. K. (1991). Trusting to the *Tao*: Chinese values and the re-centering of psychology. *Bulletin of the Hong Kong Psychological Society, 26-27*, 5-27.

Chan, H. (1988). Chinese students: adaptation, life satisfaction and academic achievement. In Hanks P. & Perry A. (eds.) *The Chinese in Australia* (pp. 25-36). Clayton: Centre for Migrant and Intercultural Studies, Monash University.

Chang, W.N., Wing, K. W., Teo, G. & Fam, A.(1997). The motivation to achieve in Singapore: in search of a core construct. *Psychology and Individual Differences*, 23(5), 885-895.

Chen, A., Liu, Z. & Ennis, C. (1997). Universality and uniqueness of teacher educational values orientation: a cross-cultural comparison between the United States and China. *Journal of Research and Development in Education,* 30(3), 135-143.

Chinese Culture Connection, (1987). Chinese values and the search for culture-free dimensions of culture. *Journal of Cross-cultural Psychology*, 18(2), 143-164.

Chiu, C. Y.(1989). Normative expectations of social behaviour and concern for members of the collective in Chinese society. *Journal of Psychology*, 124(1), 103-111.

Chiu, R. K., Wong, N. M. & Kosinski, F. A.. (1998). Confucian value and conflict behaviour of Asian managers: a comparison of two countries. *Social Behaviour and Personality,* 26(1), 11-22.

Exum, H.A. & Lau, E. Y. (1988). Counseling style: preference of Chinese college students. *Journal of Multicultural Counseling and Development*, 16, 84-92.

Feast, V. & Churchman, D. (1997). The CHC student success story: a case study in one Australian university. In *Learning and Teaching in Higher Education: Advancing International Perspectives.* Papers presented at the 23rd Annual Conference of the Higher Education Research and Development Society of Australasia, (pp. 250-256). Campbelltown: HERDSA.

Guan, J. & Dodder, R. (1998). Instrument validity of cross-cultural studies: a further search for culture-free dimensions. Poster Presentation at the 93rd Annual Meeting of the American Sociological Association. San Francisco, 31-8-1998.

Guan, J.& Dodder, R. (in press). The impact of cross-cultural contact on value and identity:a comparative study of Chinese students in China and in the United States. *Journal of Cross-cultural Psychology*.

Hofstede, G. (1991). Empirical models of cultural differences, (pp. 4-20). In Bleichrodt N. & P.D. Drenth et al. (eds.) *Contemporary Issues in Cross-cultural Psychology*. Amsterdam: Swets & Zeitlinger.

Jones, A. P., Rozelle, R. M. & Chang, W. N. (1990). Perceived punishment and reward values of supervisor actions in a Chinese sample. *Psychological Studies*, 35(1), 1-10.

Kaiser, H. F. & Caffrey, J. (1965). Alpha factor analysis. *Psychometrika,* 30(1), 1-14.

Keeves, J. P. (ed.) (1997). *Educational Research Methodology and Measurement: An International Handbook*. (2nd edition). Oxford: Pergamon.

Leung, K. & Bond, M. H.. (1989). On the empirical identification of dimensions for cross-cultural comparisons. *Journal of Cross-cultural Psychology,* 20(2), 133-151.

Man, L., et al. (1998). Cross cultural differences in self-reported decision-making style and confidence. *International Journal of Psychology,* 33(5), 325-335.

Mukherjee, K. (1990). Asian women of Australia: what future in the promised multi-cultural society. In Hedrick, C. & Holten, R. (eds.) *Cross-cultural Communication and Professional Education,* (pp 278-284). Adelaide: Centre for Multicultural Studies.

Oppenheim, A. N. (1992). *Questionnaire Design, Interviewing and Attitude Measurement*. London: Printer Publishers.

Ralston, D. et al. (1992). Eastern values: a comparison of managers in the United States, Hong Kong and the People's Republic of China. *Journal of Applied Psychology*, 77(5), 664-671.

Schwartz, S. H. (1992). Universals in the content and structure of values: theoretical advances and empirical tests in 20 countries. In Zouma, M. (ed). *Advances in Experimental Social Psychology*, 25, 1-65. Orlando: Academic Press.

Schwartz, S. H. (1994). Are there universal aspects in the structure and contents of human values? Human values and social issues: current understanding and implications for the future. *Journal of Social Issues*, 50(4), 19-45.

Slay, J. (1999a). Education online: cross cultural reflection on the Chinese situation. http://ausweb.scu.edu.au/papers/slay1/paper.html.

Slay, J. (1999b). The nature of native Chinese culture and science education. Paper presented at The National Association for Science Research in Science Teaching (NARST). Boston, 28-31 March 1999. http://www.narst.org/conference/slay/slay/html.

Wright, D. B.(1997). *Understanding Statistics: An Introduction for the Social Sciences*. London: Sage.

Yuan, B. & Chen, J. P. (1998). Moral values held by early adolescents in Taiwan and Mainland China. *Journal of Moral Education,* 27(2), 191-207.

Zhang, J. P. & Collis, B. (1995) A comparison of teaching models in the west and in China. *Journal of Instrumental Science and Technology*, 1 (1), 1-10

Appendix

THE CHINESE VALUE SURVEY

[Adapted from the Chinese Culture Connection, 1987. Chinese values and the search for culture-free dimensions of culture. *Journal of Cross-cultural Psychology*, 18(2):143-164.]

The aim of this study is to find out what matters are important or unimportant to people. You will find below a list of 40 items. Please indicate how important to you is each of the 40 items.

To express your opinions, imagine an Importance Scale that varies from 1 to a maximum of 9. (1) stands for "of no importance to me at all, and (9) stands for "of supreme importance to me." In other words, the larger the number, the greater will be the degree of importance to you. Give one number (either 1,2,3,4,5,6,7,8 or 9) to each item below, in the brackets provided to express the importance of that item to you personally.

You can concentrate better by asking yourself the following question when you rate an item: "How important is this item to me personally?" Repeat the same question when you rate the next item, and so on. THANK YOU.

( ) 1. Filial piety (Obedience to parents, respect for parents, honouring ancestors, financial support of parents)	( ) 21. Sincerity
( ) 2. Industry (Working hard)	( ) 22. Keeping oneself disinterested and pure
( ) 3. Tolerance of others	( ) 23. Thrift
( ) 4. Harmony with others	( ) 24. Persistence (Perseverance)
( ) 5. Humbleness	( ) 25. Patience
( ) 6. Loyalty to superiors	( ) 26. Repayment of both the good and the evil that another person has caused you
( ) 7. Observation of rites and rituals	( ) 27. A sense of cultural superiority
( ) 8. Reciprocation of greetings and favours, gifts	( ) 28. Adaptability
( ) 9. Kindness (Forgiveness, compassion)	( ) 29. Prudence (Carefulness)
( ) 10. Knowledge (Education)	( ) 30. Trustworthiness
( ) 11. Solidarity with others	( ) 31. Having a sense of shame
( ) 12. Moderation, following the middle way	( ) 32. Courtesy
( ) 13. Self-cultivation	( ) 33. Contentedness with one's position in life
( ) 14. Ordering relationships by status and observing this order	( ) 34. Being conservative
( ) 15. Sense of righteousness	( ) 35. Protecting your "face"
( ) 16. Benevolent authority	( ) 36. A close, intimate friend
( ) 17. Non-competitiveness	( ) 37. Chastity in women
( ) 18. Personal steadiness and stability	( ) 38. Having few desires
( ) 19. Resistance to corruption	( ) 39. Respect for tradition
( ) 20. Patriotism	( ) 40. Wealth

The work/study dilemma: A pilot study

Steven Barrett
Flinders University of South Australia *steven.barrett@unisa.edu.au*

Over the past decade Australian universities have experienced rapid growth and increased diversity of the student population. Increased diversity has been achieved by increased enrolments of students from groups in Australian society that have little or no experience of university study. Many of these students are working either full-time or part-time while pursuing their university studies. This combination of work and study poses many problems, both for individual students and the universities they attend.

This study has three aims: (a) to conduct a large pilot study of a survey instrument designed to collect information about the work/study dilemma; (b) to investigate methods of improving response rates for surveys that are conducted across the Internet; (c) to provide policy makers with some information about the strategies students employ to resolve the work/study dilemma.

The study concludes that recent attempts by Australian universities to increase the flexibility of course delivery in order to resolve the work/study dilemma is to some extent misplaced.

Internet survey, diversity, attrition, flexible delivery

Introduction

Since 1998 the number of Australian university students has increased markedly. Furthermore, the diversity of the student population has also increased. In part, this is due to the former Colleges of Advanced Education gaining university status. Increased diversity is also in part, a result of policies specifically designed to increase the participation of students from groups in Australian society that have little or no tradition of university study. Unfortunately, students from these social groups tend to be less well prepared for university studies. Consequently, their level of performance is generally lower than that of traditional groups of students (Ramsay, Tranter, Sumner and Barrett 1996). In particular, high rates of attrition for these students are becoming a cause for concern among policy makers. The origins of these problems appear to lie in the inability of many students from non-traditional backgrounds to resolve the so-called work/study dilemma. Hence, a major cause of attrition is the inability of many students to devote sufficient time to their studies, while continuing to meet their work, family and social responsibilities. The response of the university sector in general has been to increase the flexibility of course offerings. However, this decision has been made in something of an information vacuum regarding the expectations that students have about a university education.

The primary aim of this study was to pilot a survey instrument to collect information about how students address the work/study dilemma. As a result of the prohibitive associated with printing, posting and processing paper-based questionnaires it was decided to administer this survey over the Internet. However, for many people the Internet is a new medium, both for administering and responding to questionnaires. The challenges faced in the use of this new medium include the

development of strategies that lead to acceptable response rates. Although there is a wealth of information about boosting response rates of postal surveys, the relative newness of on-line surveys results in a lack of equivalent knowledge for Internet based surveys. Hence, the second aim of this study is to investigate ways of increasing the response rate of surveys conducted over the Internet. The third aim of this study is to provide policy makers with some understanding about how students suggest that the work/study dilemma might be resolved.

Perceptions of attrition

Attrition can be viewed as either a negative or a positive outcome of an educational experience. Highly prestigious institutions may argue that high rates of attrition are an inevitable consequence of maintaining the competitive academic conditions upon which their reputations are based (Lenning, Beal and Sauer 1980). In this context, high attrition rates may not be perceived as a problem, but rather as a form of quality assurance and hence as a positive educational outcome for the institution. Alternatively, some people may enrol in a course with the intention of completing only a selection of subjects to reach specific personal or professional goals. In such cases, withdrawal also can be viewed as a positive educational outcome.

The dominant view of attrition however, focuses on its perceived negative effects. First, such effects are identified as lowering the self-esteem and self-confidence of students. Even temporary withdrawal may adversely affect the confidence of people and have serious implications for any subsequent study or career path that they may wish to pursue (Lam 1984). Second, attrition is seen as causing a social loss as these people do not achieve their potential. Their talents are wasted, and society does not benefit by their further education. Third, attrition can be considered a waste of institutional resources, which have increasing opportunity costs in a period of declining real per student funding. Fourth, attrition can damage the reputations of courses and institutions by bringing into question the relevance of their courses, the quality of teaching and the adequacy of student support services. Finally, attrition can compound problems associated with falling enrolments leading institutions to experience difficulties with planning and budgeting (Ewell 1984: Price, Harte and Cole 1991).

Regardless of whether attrition is seen from a positive or negative perspective, it is commonly investigated in terms of course and institutional loss. As such, two attrition-related matters require further discussion. These have been referred to as the ‘stop out’ and the ‘goal fulfilment’ issues (Ewell 1984). The crux of the stop out problem is that university students are displaying increasingly complex patterns of enrolment. A conventional, but outmoded, view assumes that young people complete Year 12, enrol in an undergraduate degree the following year at a university and graduate on time after three to six years of uninterrupted full-time study. However, the Australian student population is becoming increasingly diverse and hence students are becoming more flexible in their pathways through and between post-compulsory education, employment and training. Combining full-time study with periods of part-time study and paid work is increasingly common among students. Furthermore, intermission from studies in order to travel, undertake paid work or for other personal reasons are also becoming more common. The majority of students may graduate eventually, but not necessarily be from the university in which they were originally enroled. Such students may appear as institutional attrition, but they are not lost to the system and their temporary withdrawal can be viewed positively.

The second issue, commonly referred to as goal fulfilment, is more complex. Some students may withdraw because they feel that their studies are not helping them to attain their goals. This can be perceived as a negative reason for withdrawing. However, other students withdraw for more positive goal-related reasons, as mentioned earlier. A student may have no intention of ever

finishing a degree program having enroled in selected subjects for personal or professional interest, to assist career progression or perhaps to gain entrance to another university (Roberts 1984). Consequently, some students withdraw because they have met their goals, while others redefine their goals or identify other means of achieving them. Unfortunately, the hopes, expectations and aspirations of the Australian student population are not fully understood by university planners and administrators. The assumption that, despite the increased diversity of the Australian student body, that all university students have one goal, to complete a bachelor's degree is increasingly less likely to be valid. If indeed it ever was.

Studies into student performance and attrition

Student success

There is little reported research into student success and attrition that specifically focuses on students from non-traditional backgrounds. The limited Australian literature in the area falls into two main categories. The first attempts to develop regression or econometric models of student performance and progress through courses. These studies usually concentrate on the pre-entry characteristics of students and inform debates concerning appropriate selection criteria for university courses. The second group of studies attempt to explain pass and retention rates. The findings of both types of studies are relevant in the development of a policy framework and initiatives aimed at improving student performance and progression rates.

West (1985) investigated the effects of three pre-entry characteristics on the performance of students who entered Monash University directly from secondary school in 1975, 1980 and 1982. The characteristics were, type of school attended, father's occupation and the student's country of birth. The measures of performance used were, grade point average, pass rate as measured by the percentage of subjects in which students attained a pass or better and credit rate as measured by the percentage of subjects in which students attained a credit or better. A least squares regression model and a logistic regression model were employed for these analyses. West concluded that students who undertook most of their secondary education in government schools performed better at the end of their first year of University than students with the same selection score from independent schools. Neither father's occupation nor country of birth were significant predictors of university performance over each of the three years investigated.

Smyth, Knuiman, Thornett and Kiiveri (1990) investigated the performance of 3,734 first-year university students who enroled at the University of Western Australia between 1977 and 1980. This project focused on both individual subject and overall performance. Three statistical models were developed, each incorporating personal data, comprised of variables relating to the performance of students in the final year of secondary school and variables relating to performance in their first year of university study. Among their findings they argued that prediction models of first year performance needed to be more sophisticated and incorporate pre-entry student characteristics in addition to those employed in their study.

Lewis (1994) analysed the results of 10,482 commencing undergraduate students who enroled at the University of Wollongong between 1990 and 1993. The study assessed the performance of students who were admitted to the University by means of its access and equity schemes, in order to determine whether the performance of these groups was comparable to that of other students. Two measures were employed. The first, the mean aggregate mark, was analysed using multivariate regression analysis. The second, whether or not students had passed 75 per cent or more of the subjects in which they were enroled, was analysed using a logistic regression model. The study concluded that female students, students who had attended government schools and

older students who were not school leavers, performed significantly better than the university-wide average. Conversely, the performance of students from non-English speaking backgrounds and those of indigenous students were significantly below the university average.

Killen (1994) reviewed a number of research studies into teaching and learning in higher education and concluded that a number of factors could influence student success, as measured by pass rates. The main factors identified by Killen in these studies were the motivation of students, their approach to studying, and their cultural expectations, all characteristics of students rather than institution or the higher education system. Killen then attempted to isolate the factors that were associated with student success by conducting interviews with a sample of students and lecturers at the University of Newcastle. As a result, four groups of factors were identified that significantly affected student performance. The first two factors were internal to students, self-motivation and effective study techniques. The two factors external to students were family support and enthusiastic lecturers. This study represented a shift from a focus on student deficits to a consideration of institutional factors by arguing that institutional factors could play a role in improving pass and attrition rates.

Student withdrawal

Since the early1980s, a relatively small number of studies have attempted to explain attrition in Australian universities. The National Institute of Labour Studies conducted an investigation into the attrition of first year students in all South Australian institutions of higher education during 1985 (Power, Robertson and Baker 1987). Two groups of students who did not proceed to the second year of their courses were identified, those who passed and withdrew and those who failed and withdrew. Students in these two sub-groups had one essential characteristic in common. They concluded that their course no longer matched their educational goals or interests. Consequently, they lacked the required level of commitment to continue. The study identified a relationship between reduced attrition rates and the implementation of appropriate strategies and programs both to generate student commitment to their course at entry and then to strengthen systematically and develop this commitment during their initial year at university.

Price, Harte and Cole (1991) conducted a study of student attrition at the Northern Territory University, focusing on comparative attrition rates across faculties, stage of course, attendance status and previous educational background, as well as on the reported reasons for of attrition. The population for this study was all students who were enroled in three successive years (1988, 1989 and 1990) who did not re-enrol in the following year. Over 2,000 students were surveyed with a response rate of 23 per cent. The study identified a particular group of students who had a higher tendency to withdraw. They were part-time students who had left school more than five years prior to enroling, had matriculated at high school, were engaged in full-time employment and were enroled in the first stage of their course.

Sharma and Burgess (1994) surveyed 855 undergraduate students (a response rate of 43 per cent) who withdrew from Swinburne University of Technology during 1993. This study identified 11 factors that students saw as important considerations in their decision to discontinue studies. In order of importance, lack of motivation, employment related matters, study workload, failure in examinations, financial situation, dissatisfaction with the learning environment and enrolment in the wrong course were identified by at least 30 per cent of respondents as a significant consideration for withdrawing. It was argued that there was a strong inverse relationship between student success and attrition. Therefore, action taken to improve student pass rates would reduce attrition. Sharma and Burgess concluded that a range of largely extraneous factors were important

determinants of attrition, including financial status, student information, employment conditions and personal or family situation.

The studies into student attrition and performance reviewed here acknowledge that these issues are becoming increasingly serious in Australian universities and require remedial action. Furthermore, these problems are especially acute for first year students. These studies suggest that these problems are cause by two sets of factors, that that are internal to students and those that are external. These studies recommend that universities implement two sets of strategies to reduce attrition and increase student performance. Strategies designed to address internal factors could include, *inter alia*, assisting students to;

- make more informed decisions about university study;
- become more highly motivated and committed to their course of study;
- develop university literacy skills; and
- develop the time management and other life skills required to manage the competing demands on their time.

Strategies designed to address external factors could include, *inter alia*, assisting students to;

- improving the enthusiasm of lecturers;
- implementing programs designed to monitor and attenuate attrition; and
- providing a teaching and learning environment that is empathetic to the needs of diverse student population.

These studies do not recommend increased flexibility of subject and course delivery through the increased use of distance education teaching methodologies and the Internet. It would therefore appear that the decision of many Australian universities to embrace flexible delivery as the panacea for the sector's woes was not based on a consideration of this body of literature.

Methods

The initial aim of this study was to pilot the work/study dilemma survey instrument. This was achieved by surveying every undergraduate student enroled in the Division of Business and Enterprise at the University of South Australia during semester 2, 1999, some 4,365 students in total. The usual mailout of questionnaires plus two reminders was considered prohibitively expensive. Furthermore, when large surveys achieve a reasonable response rate the costs of data entry also become a matter of concern. In order to minimise expense, a survey instrument that could be answered on the Internet was developed. Participant responses collected by the on-line instrument were automatically compiled to form the database and all measures were taken to ensure confidentiality.

The address of the Work/Study website was emailed to all undergraduate students of the Division of Business and Enterprise at their university email account addresses. This message took the form of a covering letter seeking student involvement and supplying the information required by the University's Ethics Committee. Two follow up email messages were also sent to participants. The website contained a secure database of the names, identity numbers and dates of birth of all of students who were invited to participate in the study. This database served a number of purposes. First, the names of the students were inserted into the survey instrument so that it included a personalised greeting. Second, the student identity number and date of birth provided a degree of security for the system. Finally, this database controlled access to the website. Once participants had submitted a completed questionnaire they were refused permission to access the site a second time. Consequently, people in the sample were prohibited from completing more than one questionnaire. The data collected were analysed in order to develop some insights into the nature

of the work/study dilemma. SPSS^X, version 9 (Norusis 2000), was used to prepare the cross tabulations and the cluster analysis that formed the basis of the following discussion.

Results

A total of 383 valid questionnaires were submitted to the database. Since the survey was emailed to 4,365 students this represented a response rate of 8.8 per cent. The address of the survey website was sent to the email accounts that the University provided to its students. However, the University's information technology services unit indicated that only 30 per cent of students accessed their university email accounts. Consequently, students who did not have access to the Internet at home or at work found it very difficult to access their email accounts. Consequently, it could be argued that only 1,310, or 30 per cent of 4,365, students read the emails that were sent to them regarding the survey. Hence, this lower figure should be regarded as the sample size. In which case, the response rate would be 29.3 per cent.

The 383 respondents were fairly representative of the student population of the Division, with the notable exception of the gender balance. There were 230 female and 153 male respondents to the survey. Males and females were fairly equally represented in the student body of the Division. The over representation of females would seem to indicate that the work/study dilemma was of more serious concern for women than it was for men. However, it might also reflect increased access by women to the Internet or that women made better use of email. Furthermore, access to and confidence with information technology may play a significant role in academic success and progress. Consequently, these sources of bias needed to be investigated. Table 1 shows the age distribution of the respondents. The youngest respondent was only 17 years of age and the oldest 51 years old, with an average age of 23.6 years. The age distribution of the respondents appeared to be representative of the student population of the Division. Table 2 shows the pathways that the respondents used to gain entry to the University. This table confirmed the view that the respondents to the survey reflected the general population of the Division.

Table 1: Respondents by Gender and Age.

Age group	Females	Males	Total	Per cent
Under 21	108	65	173	45
21 to 25	65	48	113	30
25 to 30	34	14	48	13
31 to 35	8	9	17	4
36 to 40	7	11	18	4
41 to 45	6	5	11	3
Over 45	2	1	3	1
Total	230	153	383	100

Table 2: Respondents by Entry Pathway.

Entry pathway	Females	Males	Total	Per cent
School leaver	130	89	219	58
Mature age test	39	29	68	18
Tertiary transfer	13	3	17	5
Previous TAFE studies	27	13	40	11
Recognition of prior learning	9	6	15	3
Special access award	3	0	3	1
Special entry applicant	1	4	5	1
Other	7	6	13	3
Total	229	151	380	100

Table 3 shows the courses that the respondents were enroled in when the survey was conducted. The table also indicates that this group of students reflected the general population of the Division. Courses listed as "Other" in Table 3 refers to courses offered by Schools outside of the Division of Business and Enterprise. An over representation of commerce students responded to the survey instrument. In the main, these students were studying accounting and were required to undertake a high proportion of their classes in the computer pools. Hence, they had much better access to the University's computing facilities and consequently made much better use of their university email accounts.

Table 3: Respondents by Course of Enrolment.

Course	Females	Males	Total	Per cent
Commerce	69	33	102	29
Management	44	29	73	21
Marketing	22	16	38	11
International business	26	11	37	11
Information systems	15	18	33	9
Administrative management	24	6	30	9
Banking and finance	7	6	13	4
Property	7	4	11	3
Other	16	29	45	13
Total	230	152	352	100

Table 4: Source of Income by Sex

Source of income	Females	Males	Total	Per cent
Full-time employment	53	28	81	21
Casual employment	43	19	62	16
Youth Allowance	31	29	60	16
Part-time employment	34	14	48	13
Family Support	25	22	47	12
Austudy	27	18	45	12
Casual employment	4	11	15	4
Private income	3	2	5	1
Other	9	7	16	4
Total	229	150	379	100

Table 4 indicates that the vast majority of students were employed. Just over a quarter of students were in receipt of a government benefit, either Austudy or Youth Allowance. Nevertheless, the majority of these students worked on either a casual or part-time basis in order to supplement their benefits. Only those 47 respondents who stated that their sole source of income was family support fitted the typical student model. In all likelihood, these students were probably international students and the survey instrument needs modification to identify this group of students for separate analysis in future studies.

Of more importance is the amount of time and energy that students devoted to paid employment. The relationship between the sex of the respondents and the number of hours they worked is shown in Table 5. On average, the respondents to the survey undertook 23.6 hours of paid work per week outside of the home. This figure was slightly lower for females, 23.5 hours per week, than it was for men, 23.9 hours per week. More interesting, was the high proportion of students who could be considered to be full-time employees. Thirty hours of paid employment outside of the home per week is commonly considered to be full-time employment. In which case, Table 5 shows that 21 per cent of respondents who worked were employed on a full-time basis. The concept of time budgets and time management skills is an increasingly important issue for many students. Especially given the relationship between time on task and student performance (Kokkinn Head Feast and Barrett 1998). However, information about the time students spent meeting their family, sporting or social commitments, which was likely to be considerable, was not

collected during this study. Nor were any questions about how much time respondents devoted to their studies was asked. These important questions should be followed up in a later survey.

Table 5: Hours Worked by Sex

Number of hours worked	Females	Males	Total	Per cent
10 or less	37	18	55	22
More than 10 and less than 20	46	31	77	31
More than 20 and less than 30	17	12	29	12
More than 30 and less than 40	35	17	52	21
More than 40 and less than 50	16	8	24	10
More than 50 and less than 60	4	3	7	3
More than 60 and less than 70	0	1	1	0
Total	155	90	245	100

Table 6: Income by Sex

Weekly income	Females	Males	Total	Per cent
Less than $50	10	9	18	6
More than $50 and less than $100	36	19	55	18
More than $100 and less than $150	31	20	51	16
More than $150 and less than $200	30	25	55	18
More than $200 and less than $300	23	16	39	13
More than $300 and less than $400	15	6	21	7
More than $400 and less than $500	9	11	20	6
More than $500 and less than $1000	29	13	42	14
More than $1000 and less than $2000	3	5	8	3
Total	186	124	310	100

The relationship between sex and income is presented in Table 6, which shows that not only did students devote significant amounts of time to undertaking paid employment, but also these respondents were also quite well paid. The average weekly income of employed respondents was $295.38. The figure was a little lower for females ($288.63) than it was for males ($305.52). Interestingly, the maximum income was reported to be $1800 per week. Clearly, the preceding tables present a picture of a diversified student population, which reflects the macro level changes that have occurred in the Australian higher education system since 1988.

Table 7: Preferred Class Times by Sex

Preferred class times	Females	Males	Total	Per cent
Weekdays 9 am to 5 pm	57	41	98	88
Weekdays 5 pm to 9 pm	4	4	8	7
Weekends	4	2	6	5
Total	65	47	112	100

Table 8: Preferred Teaching Mode by Sex

Preferred Teaching Mode	Females	Males	Total	Per cent
Offered internally on campus	50	35	85	83
Offered externally	5	5	10	10
Offered via the Internet	5	3	8	8
Total	60	43	103	100

The third aim of this study was to develop some understanding of the work/study dilemma. The University of South Australia has identified the problems that students encounter combining their studies with work and other commitments as a major cause of student attrition and poor performance (Ramsay, Sumner, Tranter and Barrett 1996). The response of the University has been to increase the flexibility of program delivery. This usually means increased use of distance education modes of delivery and increased use of the Internet. As this has been a major thrust of curriculum development over the past few years it was decided to gauge the opinions of students

about the University's increased reliance of these non-traditional forms of subject delivery. The responses to these two questions are reported in Tables 7 and 8. These tables demonstrate that despite the diversity of the respondents they share two common characteristics. Firstly, 88 per cent of the respondents wanted to attend classes during normal business hours (Table 7), between 9 am and 5 pm on Monday to Friday. Secondly, 83 per cent of the respondents revealed a strong preference to study on-campus in the internal mode (Table 8). However, these results need to be interpreted with some caution due to the much lower response rates on these two questions, compared to the response rates for the questions reported in Tables 1 to 6, which were among the first questions on the survey instrument.

Results of the Cluster Analysis

In order to understand better the attitudes of students, a series of cluster analyses were undertaken to identify whether there were any distinct groups of respondents (Table 9). The cluster analysis identified the presence of two distinct groups of students. Surprisingly, given the literature reviewed above, the sex of students did not emerge as a characteristic of these groups. Group 1 accounted for the majority of students, 84 per cent, in the sample. This group of 190 people was characterised by young students, who live at home with their parents and earn a small income as a result of casual employment. Group 2, which accounted for 37 people in the sample, is almost the exact opposite to Group 1. It is characterised by older people, in well-paid full-time employment, who lived at home with their partners and children.

Cluster analysis is essentially an iterative process. Once the analysis identified the two groups discussed above the cluster analysis was extended to investigate whether a third or even a fourth cluster of respondents existed. Further analysis identified a third group of six people as a distinct cluster, which was comprised of a small number of the highest income earners in Group 2. In effect, it was a sub-group of the second group. Given the size of this sub-group and the characteristics of its members it was concluded that only two distinct clusters were present.

Table 9: Results of the Cluster Analysis

Respondent characteristics	Cluster 1	Cluster 2
Age	22.1	32.3
Employment	Casual employment	Full-time employment
Weekly income	$237.42	$908.75
Hours of work per week	20	42
Length of enrolment	2 years	3 years
Entry pathway	School leaver	Previous TAFE studies
Enrolment status	Full-time	Part-time
Mode of study	Internal student	Internal student
Parental responsibilities	No	Yes
Domestic arrangements	Live at home with parents	Live at home with partner
Number of respondents	190 (84 %)	37 (16 %)

Despite the marked differences between the members of Group 1 and Group 2, the two groups of students had one thing in common. They wanted to study on campus during normal business hours, as shown in Tables 7 and 8. Not surprisingly, they also displayed a strong preference for traditional forms of subject delivery. However, given the obvious differences in the characteristics of the members of these two groups, the reasons underlying these preferences were probably quite different. For example, young school leaver students might value highly the social interactions that took place with other young people on campus. Young students might also come to university with strong expectations that university is an extension of secondary school and they would attend classes and be taught by teachers. Conversely, older students might feel the need to interact with other students in order to gain support, reduce social isolation and develop business

networks. They may also have seen study as an integral part of their employment. Hence, they held the view that classes should be taken during business hours. Clearly, the survey instrument needs further modification to include a number of items that unpack the information shown in Tables 7 and 8 and explore the expectations and aspirations of students. It might also be necessary to conduct a number of focus groups with participants to further refine the questionnaire items.

Conclusions

If this pilot study is to be expanded then the response rate will need to be markedly increased. Moreover, sources of bias need to be identified and understood. Hence, several strategies should be employed. Firstly, the number of items in the survey instrument must be reduced and refined to improve its focus. Secondly, students must be encouraged to visit the survey website by a variety of means other than sending messages to their university email accounts, which 70 per cent do not read. Hence, the cooperation of people such as course coordinators, subject conveners, school office staff and other administrators, is vital. Thirdly, the Internet is increasingly being used for both teaching and administration. Courses, subjects, staff and even students increasingly have their own website. Furthermore, student administration, such as checking enrolment details, amending information held on student information databases, receiving provisional as well as final results and tutorial allocation are increasingly being undertaken across the Internet. Consequently, hypertext links to the survey website could be added to the administration and teaching websites students should visit. Fourth, approximately 30 per cent of students subscribe to an Internet service provider. Many lecturers at the University of South Australia have responded to this trend by setting up webserve lists for their subjects as a medium to distribute teaching materials and to communicate directly with their students. Hence, the address of the survey website could be posted to these webserve lists. Finally, students are rational and do respond to inducements. Hence, the personal information database could be used to pick prizewinners.

The survey indicates that students use a wide variety of pathways into and through their university studies. They combine study with periods of part-time and full-time work and step out of their studies for travel or other personal reasons. These observations may also indicate changed patterns of workforce entry. The conventional view is that students complete secondary school, complete their university studies and then enter the workforce after graduation. However, for business students this does not appear to be the case. University studies and workforce entry increasingly seem to be occurring in tandem. However, this may differ for students in other courses and in other universities. Hence, the survey instrument needs to include questions that provide greater insights about the pathways that students follow through their studies and unpack the process by which entry to the workforce is undertaken.

Despite the increased diversity of the Australian student population and the complexity of the pathways through their studies, the respondents expressed a strong desire to study on campus with their classes conducted during normal business hours. This preference for traditional methods of subject delivery is clearly at odds with recent trends at the University of South Australia towards increased flexibility of course offerings. The survey found that a high proportion of students were employed on a casual or part-time basis, especially in the growing service sector. Students appear to be using the increased flexibility that has emerged in the labour market, as a result of economic restructuring and recent changes to workplace relations to provide the flexibility required to overcome the work/study dilemma. If this is the case then the University's flexible delivery agenda appears to be misplaced. Hence, the University needs to understand these issues if it is to improve the performance of students from non-traditional backgrounds.

The findings of the survey indicated the existence of two distinct clusters of students, which have very different demographic characteristics. Consequently, different reasons probably underlie the preferences for teaching traditional methods. The expectations of these two groups of students need to be fully understood if appropriate policies to improve student performance are to be implemented. The next step in this project should be to conduct focus group discussions with respondents to unpack the issues and refine the survey instrument.

REFERENCES

Ewell, P.T. (1984) *Conducting Student Retention Studies.* Boulder: National Center for Higher Education Management Systems.

Killen, R. (1994) Differences between students' and lecturers' perceptions of factors influencing success at university, *Higher Education Research and Development*, 13 (2), 199-211.

Kokkinn, B. Head, M., Feast, V and Barrett, S.R.F. (1998) Transforming the teaching of economics: Embedding tertiary literacy, paper presented at the HERDSA Conference98, Auckland, July 1998.

Lam, Y.L.J. (1984) Predicting dropouts of university freshmen: A logistic regression analysis, *Journal of Educational Administration*, 22 (1), 74-82.

Lenning, O.T., Beal, P.E. and Sauer, K. (1980) *Retention and Attrition: Evidence for Action and Research.* Boulder: National Center for Higher Education Management Systems.

Lewis, D.E. (1994) *The Performance at University of Equity Groups and Students Admitted Via Alternative Modes of Entry*. Canberra: Australian Government Publishing Service.

Norusis, M.J.(2000) *SPSS 9.0 Guide to Data Analysis*, Chicago: Statistical Package for the Social Sciences.

Power, C., Robertson, F. and Baker, M. (1987) *Success in Higher Education*. Canberra: Australian Government Publishing Service.

Price, D., Harte, J. and Cole, M. (1991) *Student Progression in Higher Education: A Study of Attrition at Northern Territory University*. Canberra: Australian Government Publishing Service.

Ramsay, E., Tranter, D., Sumner, R. and Barrett, S.R.F. (1996) *Outcomes of a University's Flexible Admissions Policies*, Evaluations and Investigations Program, Department of Employment, Education, Training and Youth Affairs, Report Number 96/8, Canberra: Australian Government Publishing Service.

Roberts, D. (1984) Ways and means of reducing early student drop-out rates, *Distance Education*, 5 (1), 50-71.

Sharma, R. and Burgess, Z. (1994) Student discontinuation at a technological higher education institution, *Journal of Tertiary Education Administration*, 16, 225-238.

Smyth, G.K., Knuiman., M.W., Thornett, M.L. and Kiiveri, H. (1990) Using the EM algorithm to predict first year university performance, *Australian Journal of Education*, 34, 204-224.

West, L.H.T. (1985) Differential prediction of first year university performance for students from different social backgrounds, *Australian Journal of Education*, 29 (2), 175-187.