

Measuring student satisfaction from the Student Outcomes Survey

Peter Fieger

NATIONAL CENTRE FOR
VOCATIONAL EDUCATION RESEARCH

NATIONAL CENTRE FOR
VOCATIONAL EDUCATION RESEARCH
TECHNICAL PAPER





Australian Government

**Department of Industry, Innovation
Science, Research and Tertiary Education**

Measuring student satisfaction from the Student Outcomes Survey

Peter Fieger
National Centre for Vocational Education Research

NATIONAL CENTRE FOR VOCATIONAL
EDUCATION RESEARCH

TECHNICAL PAPER

The views and opinions expressed in this document are those of the author/project team and do not necessarily reflect the views of the Australian Government or state and territory governments.

© Commonwealth of Australia, 2012



With the exception of the Commonwealth Coat of Arms, the Department's logo, any material protected by a trade mark and where otherwise noted all material presented in this document is provided under a Creative Commons Attribution 3.0 Australia
<<http://creativecommons.org/licenses/by/3.0/au>> licence.

The details of the relevant licence conditions are available on the Creative Commons website (accessible using the links provided) as is the full legal code for the CC BY 3.0 AU licence
<<http://creativecommons.org/licenses/by/3.0/legalcode>>.

The Creative Commons licence conditions do not apply to all logos, graphic design, artwork and photographs. Requests and enquiries concerning other reproduction and rights should be directed to the National Centre for Vocational Education Research (NCVER).

This document should be attributed as Fieger, P 2012, *Measuring student satisfaction from the Student Outcomes Survey*, NCVER, Adelaide.

This work has been produced by NCVER on behalf of the Australian Government, and state and territory governments, with funding provided through the Department of Industry, Innovation, Science, Research and Tertiary Education.

The views and opinions expressed in this document are those of NCVER and do not necessarily reflect the views of the Australian Government or state and territory governments.

ISBN 978 1 922056 06 1
TD/TNC 108.04

Published by NCVER
ABN 87 007 967 311

Level 11, 33 King William Street, Adelaide, SA 5000
PO Box 8288 Station Arcade, Adelaide SA 5000, Australia

ph +61 8 8230 8400 fax +61 8 8212 3436

About the research

Measuring student satisfaction from the Student Outcomes Survey

Peter Fieger, National Centre for Vocational Education Research

The Student Outcomes Survey is an annual national survey of vocational education and training (VET) students. Since 1995, participants have been asked to rate their satisfaction with different aspects of their training, grouped under three main themes: teaching, assessment, and generic skills and learning experiences. While the composition of the bank of satisfaction questions has remained fairly constant over time and the suitability of the three overarching satisfaction categories has been validated statistically on several occasions, little progress has been made on creating summary measures that encapsulate the three main themes of student satisfaction. Such summary measures would be much more useful to researchers than responses to the bank of 19 satisfaction questions, which are very detailed. This paper compares three methods of creating a composite score and evaluates their statistical veracity.

Key messages

- The grouping of satisfaction questions into themes of teaching, assessment, and generic skills and learning experiences remains statistically valid in the current Student Outcomes Survey.
- A composite score for questions under these three main themes is needed to facilitate post-survey analytical studies.
- We review and compare three different methods of creating summary measures in respect of their utility. These methods are Rasch analysis, weighted means and simple means.
- We find that all three methods yield similar results and so recommend using the simple means method to create the summary measures.

Tom Karmel
Managing Director, NCVER

Contents

Tables and figures	6
Introduction	7
Satisfaction themes	9
Comparison of composite measures	12
Rasch analysis	12
Simple averages	13
Weighted averages	13
Evaluation/best fit	13
Conclusion	16
References	17

Tables and figures

Tables

1	Eigenvalues of the correlation matrix (abridged)	9
2	Factor loadings after transformation using varimax rotation	11
3	Descriptive statistics and coefficients of reliability	12
4	Descriptive statistics of composite scores	14
5	Comparison teaching composite scores	14
6	Comparison assessment composite scores	14
7	Comparison generic skills and learning composite scores	14

Figures

1	Student satisfaction items in the Student Outcomes Survey	8
2	Eigenvalues based on parallel analysis	10

Introduction

The Student Outcomes Survey is an annual national survey of vocational education and training (VET) students. The survey aims to gather information on students, including their employment situation, their reasons for undertaking the training, the relevance of their training to their employment, any further study aspirations, reasons for not undertaking further training and satisfaction with their training experience. The survey is aimed at students who have completed a qualification (graduates) or who successfully completed part of a course and then leave the VET system (module completers).

The assessment of student satisfaction with their training consists of 19 individual questions and one summary question (see figure 1). The teaching and learning questions are based on questions asked in the Higher Education Course Experience Survey, and the generic skills and learning experience questions are based on questions developed by Western Australia as part of the VET student survey (Bontempo & Morgan 2001). These questions occupy a significant portion of the questionnaire (20 out of 56 questions). To date the focus has been on reporting only the overall satisfaction item. Use of the individual satisfaction questions has been limited, mainly due to their specificity, narrow scope and number of measures.

The individual satisfaction questions are grouped under three themes: teaching, assessment, and generic skills and learning experiences. While there has been some initial statistical validation of these three groupings, no significant recent analysis has been undertaken, and no summary measure of the constituent questions has been devised.

It is the purpose of this paper to validate statistically the grouping of the satisfaction questions in the context of current surveys and to develop a summary measure for each of the three themes to make the data more accessible. We use principal component analysis to identify the underlying dimensions of the 19 satisfaction items and group the questions accordingly. Cronbach's alpha scores are calculated to assess the internal consistency of the resulting groups.

We then use three different approaches to derive composite scores to represent the groups created: Rasch analysis, weighted composite averages and straight averages.¹ Finally, we determine the extent to which the newly established composite scores differ and which ones would be most useful in future research and reporting.

¹ Further explanation of these methods is found on pages 11 and 12.

Figure 1 Student satisfaction items in the Student Outcomes Survey

	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree	Not applicable
Teaching						
1 My instructors had a thorough knowledge of the subject content	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2 My instructors provided opportunities to ask questions	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3 My instructors treated me with respect	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4 My instructors understood my learning needs	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5 My instructors communicated the subject content effectively	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6 My instructors made the subject as interesting as possible	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Assessment						
7 I knew how I was going to be assessed	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8 The way I was assessed was a fair test of my skills	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9 I was assessed at appropriate intervals	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10 I received useful feedback on my assessment	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
11 The assessment was a good test of what I was taught	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Generic skills and learning experiences						
12 My training developed my problem-solving skills	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
13 My training helped me develop my ability to work as a team member	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
14 My training improved my skills in written communication	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
15 My training helped me to develop the ability to plan my own work	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
16 As a result of my training, I feel more confident about tackling unfamiliar problems	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
17 My training has made me more confident about my ability to learn	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
18 As a result of my training, I am more positive about achieving my goals	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
19 My training has helped me think about new opportunities in life	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Overall satisfaction with the training						
<i>How would you rate, on average, your satisfaction with the overall quality of the training?</i>						
20 Overall, I was satisfied with the quality of this training	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Source: NCVET Student Outcomes Survey 2010 questionnaire.

Satisfaction themes

The bank of satisfaction questions in the Student Outcomes Survey was based on questions developed for use in the Higher Education Course Experience Survey and the Western Australian State Student Survey. The initial statistical validation of the satisfaction questions in the TAFE setting was undertaken by the Western Australian Department of Education and Training. (For more information on the history of the satisfaction questions see Bontempo & Morgan [2001] and Sevastos [2001].) Western Australia used this bank of questions in 2003 and a modified version became a constituent part of the current national Student Outcomes Survey in 2004.

While there have been several evaluations of the categorisation of the satisfaction questions into the three main themes, and these have provided a statistical basis for question groupings over the history of the survey (Morgan & Bontempo 2003), there has been scant progress towards creating summary measures beyond the initial categorisation into the three current themes.

Our investigations are based on the results of the 2009 survey. This represents the most recent large sample year (the Student Outcomes Survey is run with an augmented sample in alternating years). Our analysis was then duplicated for validation purposes with 2007 and 2008 data, yielding similar results.

Data were prepared by combining module completers and graduates. While the individual satisfaction means of these two groups differed significantly, in respect of this analysis, we find that module completers and graduates display similar response patterns.

Using principal component analysis, we can identify the underlying dimensions of the 19 satisfaction items and group the questions accordingly. The Eigenvalues of the correlation matrix of the initial weighted principal component analysis are shown in table 1.

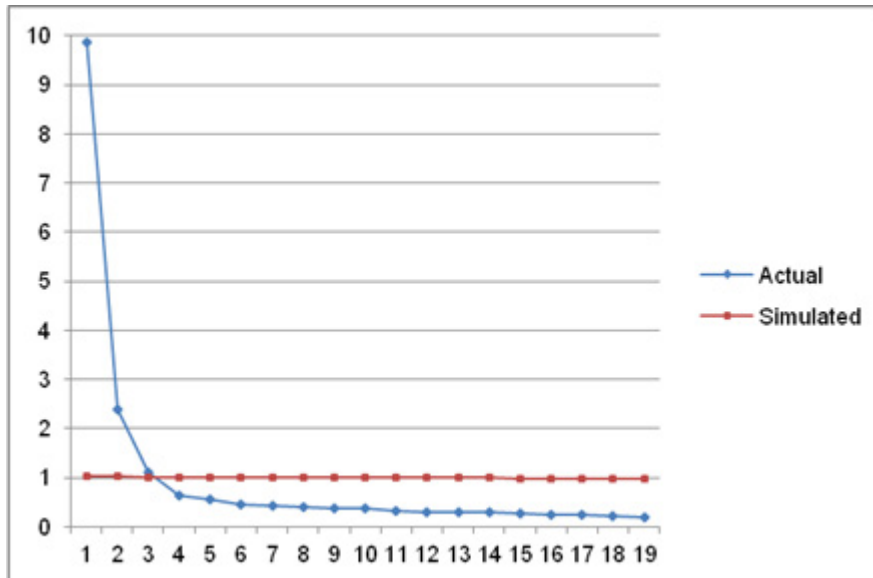
Table 1 Eigenvalues of the correlation matrix (abridged)

	Eigenvalue	Difference	Proportion	Cumulative
1	9.8397	7.4394	0.5179	0.5179
2	2.4004	1.2989	0.1263	0.6442
3	1.1014	0.4719	0.058	0.7022
4	0.6295	0.0816	0.0331	0.7353
5	0.5478	0.0841	0.0288	0.7641
...				
18	0.2337	0.0456	0.0123	0.9901
19	0.1881		0.0099	1

Note: Rows 6—17 are omitted but can be supplied upon request.

While there are various ways of assessing the number of factors that ideally should be retained, we applied Horns parallel analysis that uses a Monte Carlo-based simulation to compare the observed Eigenvalues with those obtained from uncorrelated normal variables. The visual inspection of the resulting graph (figure 2) indicates that three components should be retained. These three extracted components account for about 70% of the variance in the 19 satisfaction items.

Figure 2 Eigenvalues based on parallel analysis



The factor pattern resulting from the three retained factors was then transformed via varimax rotation (table 2). It is very apparent that each single question unambiguously correlates with one particular factor (shaded in table) and that the resulting three groups correspond to the three thematic question groups from the survey. For example, those questions (numbered 1 to 6) that correlate with factor 2 correspond to the teaching block, those (numbered 7 to 11) correlating with factor 3, correspond to the assessment block, and those (numbered 12 to 19) correlating with factor 1, correspond to the generic skills and learning experience block of questions.

We further tested the reliability of the three question groups by means of Cronbach's coefficient of reliability (table 3). All three groups represent excellent internal consistency as evidenced by a very high Cronbach's alpha statistic. None of the 'alpha if deleted' values exceeds the overall alpha score, which further documents the high reliability of the selected satisfaction groupings.

Based on the results of the principal component analysis and the review of the Cronbach's alpha scores, we conclude that the grouping of the satisfaction items into the themes of teaching, assessment, and generic skills and learning experiences in the Student Outcomes Survey is statistically justified.

Table 2 Factor loadings after transformation using varimax rotation

Question	Factor 1	Factor 2	Factor 3
1 My instructors had a thorough knowledge of the subject content	0.1898	0.7597	0.2226
2 My instructors provided opportunities to ask questions	0.1699	0.7929	0.2434
3 My instructors treated me with respect	0.1790	0.7829	0.2373
4 My instructors understood my learning needs	0.2794	0.7378	0.3180
5 My instructors communicated the subject content effectively	0.2442	0.7817	0.2980
6 My instructors made the subject as interesting as possible	0.2838	0.7181	0.2836
7 I knew how I was going to be assessed	0.1673	0.2132	0.7560
8 The way I was assessed was a fair test of my skills	0.2557	0.3426	0.7650
9 I was assessed at appropriate intervals	0.2437	0.3295	0.7623
10 I received useful feedback on my assessment	0.3012	0.3626	0.6523
11 The assessment was a good test of what I was taught	0.3296	0.3905	0.6843
12 My training developed my problem-solving skills	0.7314	0.2280	0.2539
13 My training helped me develop my ability to work as a team member	0.7583	0.2128	0.1924
14 My training improved my skills in written communication	0.7716	0.1170	0.1916
15 My training helped me to develop the ability to plan my own work	0.8085	0.1551	0.1943
16 As a result of my training, I feel more confident about tackling unfamiliar problems	0.8111	0.2257	0.1851
17 My training has made me more confident about my ability to learn	0.8235	0.2243	0.1866
18 As a result of my training, I am more positive about achieving my own goals	0.8174	0.2317	0.1865
19 My training has helped me think about new opportunities in life	0.7496	0.1995	0.1591

Note: Shading indicates the question highly correlates with one particular factor.

Table 3 Descriptive statistics and coefficients of reliability

Question	N	Mean	Std dev.	Alpha if deleted	Alpha score
1	103 997	4.461	0.750	0.9074	0.9151
2	103 939	4.487	0.731	0.9018	
3	103 744	4.504	0.748	0.9030	
4	103 293	4.257	0.869	0.8997	
5	103 607	4.272	0.856	0.8950	
6	103 040	4.165	0.930	0.9035	
7	102 602	4.197	0.838	0.8909	0.8916
8	102 491	4.248	0.810	0.8587	
9	101 224	4.218	0.813	0.8623	
10	101 634	4.068	0.974	0.8775	
11	101 995	4.194	0.850	0.8631	
12	100 029	3.886	0.896	0.9304	0.9363
13	98 254	3.879	0.948	0.9301	
14	96 099	3.653	1.013	0.9313	
15	98 356	3.859	0.941	0.9274	
16	100 749	3.962	0.914	0.9257	
17	101 472	4.009	0.912	0.9249	
18	101 193	4.000	0.920	0.9251	
19	100 372	4.037	0.937	0.9319	

Comparison of composite measures

It seems reasonable to speculate that the narrow scope of the individual satisfaction questions, along with the number of questions, has discouraged their use in research. It is therefore desirable to have a composite score or summary measure for each of the three themes that encapsulates the data collected. This should be done by capturing the core information contained in the individual questions, while retaining as much information as possible. The result should be three individual scores representing teaching, assessment, and generic skills and learning experiences.

Rasch analysis

Rasch analysis is a variant of item response theory and is used chiefly to analyse test scores or attitudes that are represented by Likert-type scales. The Rasch measurement model is used to evaluate the fit of items to their intended scales and to generate individual scores and estimate the precision of those scores on an interval scale. The method also provides diagnostic information about the items and responses to them. Under item response theory, a set of items is assumed to reflect an underlying trait (such as satisfaction, teaching, assessment and learning) and responses to items are taken to indicate how strong individuals are on that trait and how easy or difficult it is to agree with an item reflecting that trait.

In this paper, we are using the Rasch scores created by Curtis (2010). This work also contains a more detailed description of the method used to derive them.

Simple averages

As a second measure, we created a composite score for each of the three themes by calculating straightforward averages for each individual. These mean scores were created even when individual responses to satisfaction questions were missing; for example, if the response to a question is missing the measure is calculated on the average of the remaining questions. This method thus maximises the use of the available data while, at the same time, using the fewest administrative and computational resources.

Weighted averages

When using the above simple average scores, it can be argued that not all individual items contribute to the composite score to the same extent. It is useful to create a measure that accounts for the varying contributions of individual responses to the overall score. To create such a measure, we estimate factor scores for the three identified dimensions. The scores have a mean of zero and a standard deviation of one, and represent the three themes of teaching, assessment, and generic skills and learning experiences. We then regress the constituent satisfaction scores onto the factor scores, with the aim of determining the strength of association of individual questions to the composite score. The resulting beta standardised regression coefficient provides a measure of the strength of the contribution to the composite score. The composite scores are calculated as:

$$\text{Teaching}_{\text{weighted}} = Q_1 * W_{q1} + Q_2 * W_{q2} + Q_3 * W_{q3} + Q_4 * W_{q4} + Q_5 * W_{q5} + Q_6 * W_{q6}$$

with weights derived by:

$$W_{qi} = \frac{B_i}{\sum_{i=1}^n B_i}$$

The result represents the weighted average score for teaching satisfaction that has the same metric as the simple average score. The composite scores for assessment satisfaction and generic skills and learning experiences are created using analogous procedures. One disadvantage of this method is that when a response for an individual satisfaction question is missing, a meaningful weighted composite score cannot be calculated unless the missing response is imputed. Since response data for individual questions are only rarely missing (if satisfaction responses are missing they are usually missing for the entire respondent record), this issue is considered to be a negligible problem.

Evaluation/best fit

As a result of the application of the above methodologies, we now have available three different sets of composite scores for the three themes. The basic descriptive statistics of the three summary measures can be found in table 4.

Table 4 Descriptive statistics of composite scores

Variable	Method	N	Mean	Std dev.	Sum.	Min.	Max.
Teaching	Rasch scores	90 111	3.432	2.377	309 229	-4.85	6.27
	Means	90 486	4.354	0.687	393 946	1	5
	Weighted means	87 605	4.402	0.664	385 597	1	5
Assessment	Rasch scores	88 728	2.742	2.285	243 327	-4.77	5.93
	Means	89 556	4.184	0.717	374 745	1	5
	Weighted means	86 095	4.203	0.704	361 870	1	5
Generic skills and learning experiences	Rasch scores	87 443	2.326	2.460	203 431	-6.09	6.89
	Means	89 910	3.915	0.773	352 017	1	5
	Weighted means	79 268	3.889	0.785	308 293	1	5

While the means and weighted means scores appear fairly similar, the mean and variation of Rasch scores are different. We therefore calculate correlations and Cronbach's alpha to determine commonalities between the different methods and their reliability (tables 5 to 7).

Table 5 Comparison teaching composite scores

Calculation method	Rasch scores	Means	Weighted means
Rasch scores	1	0.9571	0.9442
Means	0.9571	1	0.9928
Weighted means	0.9442	0.9928	1
<hr/>			
Cronbach's alpha	Raw	0.7744	
	Standardised	0.9879	

Table 6 Comparison assessment composite scores

Calculation method	Rasch scores	Means	Weighted means
Rasch scores	1	0.9633	0.9473
Means	0.9633	1	0.9809
Weighted means	0.9473	0.9809	1
<hr/>			
Cronbach's alpha	Raw	0.8029	
	Standardised	0.9876	

Table 7 Comparison generic skills and learning composite scores

Calculation method	Rasch scores	Means	Weighted means
Rasch scores	1	0.9727	0.9711
Means	0.9727	1	0.9978
Weighted means	0.9711	0.9978	1
<hr/>			
Cronbach's alpha	Raw	0.8157	
	Standardised	0.9934	

The main finding here is that correlations between the three methods are exceptionally high, with minimum correlations of 0.94 between Rasch scores and the weighted means method in the teaching and assessment themes (tables 5 and 6) and reaching almost one between means and weighted means methods in the generic skills and learning experiences theme (table 7).

Cronbach's raw alpha scores encompassing the three aggregation methods are 0.77 for teaching, 0.80 for assessment, and 0.82 for generic skills and learning. The values suggest a very high degree of inter-item correlation.² Cronbach's standardised alpha scores can be interpreted as an indicator of inter-item covariance. In the three themes of teaching, assessment, and generic skills and learning experiences, the standardised values are all around 0.99. This suggests a very similar distribution of Rasch scores, means, and weighted means. Taken together, Cronbach's raw and standardised scores indicate strong internal consistency and uni-dimensionality between Rasch, means, and weighted means scores, and this is the case for all three groups under consideration. As a result, all three aggregation methods yield comparable results and can be used interchangeably for analysis purposes.

² Values in excess of 0.7 are normally considered to signal very strong reliability.

Conclusion

This paper provides a statistical foundation for the grouping of the satisfaction questions in the Student Outcomes Survey into three coherent categories. Results of the principal component analysis show this grouping is statistically valid.

The second aim of the paper was to create summary measures that encapsulate the three main themes of student satisfaction to aid future research and reporting. To achieve this, three different quantitative methods were devised, evaluated and compared. While all three methods each have a distinct scoring technique, as far as the measurement of the core outcome for each category is concerned, the statistical outcome differed very little.

So which method should be used?

Given that all three methods yield very similar results and that Rasch analysis and weighted means analysis each require explicit preparation of the data, it is reasonable to rely on simple average scores for the three components. This will minimise the required effort and the potential for error among users of the data.

We thus recommend, for analytical purposes, that simple satisfaction means be used for each of the three themes. This methodology can easily be applied retrospectively to historical data and applied to future survey results with minimal effort.

References

- Bontempo, J & Morgan, P 2001, *2001 TAFE Student Survey validation – measuring student perceptions of teaching, assessment and learning*, [The Department of Education, Western Australia].
- Curtis, D 2010, *Evaluating institutional performance indicators in VET*, NCVER, Adelaide.
- Morgan, P & Bontempo, J 2003, 'Validating measures of student satisfaction with learning, teaching, assessment in the Western Australian Student Outcomes Survey', VET Policy and Corporate Communications, Department of Education and Training, Western Australia.
- Sevastos, P 2001, *Validation and cross validation of the TAFE student satisfaction scales*, Curtin University of Technology, Perth.



National Centre for Vocational Education Research Ltd
Level 11, 33 King William Street, Adelaide, South Australia
PO Box 8288, Station Arcade, SA 5000 Australia
Telephone +61 8 8230 8400 Facsimile +61 8 8212 3436
Website www.ncver.edu.au Email ncver@ncver.edu.au