

# **Assessing Employability Skills: the case of problem-solving**

*David D Curtis*

*Centre for Lifelong Learning and Development  
and School of Education, Flinders University*

## **Abstract**

Over the past two years, renewed interest in employability skills has emerged. The lack of effective and efficient assessment and reporting approaches have been identified as impediments to the successful implementation of past employability skills schemes. In order to address these deficiencies, a new approach to the assessment of problem-solving, involving both a new assessment tool and a novel method of administration, has been developed. The approach is argued to be authentic, being based on tasks that participants undertake routinely rather than being developed specifically for assessment. The method is performance-based and, consistent with assessment practices in the VET sector, it is criterion-referenced.

The paper describes the development of the Problem-Solving Assessment tool, including its theoretical basis, and the method of its administration. It also presents the results of a study undertaken to validate the tool. The results of the analysis indicate that several bands of performance can be recognised and suggest that the performance bands may provide a basis for profiling individuals' employability skills. These profiles may be matched to the needs of industry and to the requirements of occupation types. Implications for the assessment of other employability skills using similar tools and methods are outlined.

## **Introduction**

The project reported here was conceived on two main bases. First, after reviewing many generic skills schemes, considerable agreement about the major constituents is apparent. Among the generic skills that are almost universally accepted are communication, teamwork and problem-solving. The general agreement about these, and some other skills, suggests that there is a broadly endorsed tacit theory of human abilities. However, what is lacking in tacit theory is a detailed and explicit account that has predictive value. In the problem-solving domain there is a substantial body of explicit theory that can be applied to problem-solving as a generic skill. In this study, alternative theoretical accounts of problem-solving were evaluated and sets of dispositions and behaviours were identified that are associated with effective problem-solving performance. These have been used as a basis for identifying both major problem-solving processes and specific indicators that those processes are being employed by individuals.

Second, the position has been taken that reported achievement should be based on assessment that leads to the estimation of ability of a known and satisfactory precision. That is, ability should not simply be asserted qualitatively, but that for reported achievement to be useful to individuals, institutions, systems and employers, it must be based on assessments that lead to the measurement of the ability in question. This position is taken to be as valid for generic employability skills as for vocationally specific skills.

In addition to the requirements outlined above, assessment tasks should provide opportunities for individuals to enhance the skills being tested and assessment should

be authentic and efficient. In the paper the development and implementation of a tool to support the development of problem-solving and to assess its performance are described.

## **Literature Review**

### *Theoretical bases for problem-solving*

Greeno, Collins and Resnick (1996) outlined three major theoretical stances on learning and problem-solving that they term associationist/behaviourist/empiricist, cognitive/rationalist view, and pragmatist/situative/sociohistoric. Behaviourist views are of limited value in predicting future problem-solving performance as they invoke learned responses to known situations. The dynamic work environment of advanced economies demands that workers be adaptable to novel circumstances, and an ability to perform only in known situations is of limited value. The situative paradigm provides powerful explanations of performance in domains within which individuals develop specific knowledge and strategies that are closely bound to the context in which problems are encountered. The examples of performance that are explained well by this model (Lave, 1988; Scribner, 1986) demonstrate just how adaptive individuals are in developing specialised skills. However, these models do not provide the predictive function that is required in assessing the potential of individuals in future novel problem tasks. Here, cognitive explanations of problem-solving performance appear to be more productive.

Cognitive accounts of problem-solving include both performance and metacognitive components. Many authors have provided descriptions of problem-solving processes. Polya (1957) described four major processes in problem-solving: understand the problem; devise a plan; carry out the plan; and look back. Bransford and Stein (1984, p.12) listed five processes: identify problems; define and represent them with precision; explore possible strategies; act on these strategies; and look back and evaluate the effects on your activities. Mayer and Wittrock (1996, p.50) included: assessing the requirements of the problem; constructing a solution plan; selecting an appropriate solution strategy; monitoring progress toward the goal; and modifying the solution plan when necessary. These are but a few of the many descriptions of problem-solving processes that are present in the very extensive literature on this topic.

The major processes that were selected for inclusion in the assessment of problem-solving in this project were:

- apprehending that a problem exists and elaborating or defining the problem;
- planning an approach to the problem including selecting strategies;
- carrying out the plan;
- monitoring progress towards the goal; and
- looking back at the progress of the solution attempt.

### *Past approaches to the assessment of generic skills*

A variety of approaches have been taken to the assessment of generic skills. These include:

- holistic judgments by teachers;
- portfolios created by students;
- assessment based on work experience; and
- assessment using purpose-developed instruments.

Each of these approaches has value in particular situations. For example, holistic judgements have been shown to work well in senior secondary schools where teachers are able to make frequent observations of students' performances (McCurry & Bryce, 1997; National Industry Education Forum, 2000). However, in other situations, such as VET and higher education courses where such close observation does not occur, the method is less likely to be relevant. Portfolio assessment is a very useful method in that it requires individuals to assemble a substantial body of evidence for the emergence of their skill sets (Feast, 2000; National Industry Education Forum, 2000; Reynolds, 1996). Because of the detail contained in portfolios, they are not readily condensed into a form that is amenable to rapid evaluation. Where this has been attempted, some doubt has been expressed about the reliability of this form of assessment (Troper & Smith, 1997). Workplace assessment is potentially valuable (National Industry Education Forum, 2000; Queensland Department of Education, 1997). For employees in a workplace and who want their employability skills assessed, it has the advantage of being an authentic exercise. For students who are undertaking work experience, the value of assessment in the workplace depends very much on the context of the workplace and on the knowledge and skill of workplace assessors and this may compromise the generality of this form of assessment (Robertson, Harford, Strickland, Simons, & Harris, 2000).

Several forms of instrumental assessment have been trialed. The Graduate Skills Assessment (GSA) includes multiple choice items for problem-solving, critical thinking and reasoning, and interpersonal understanding (Australian Council for Educational Research, 2001). GSA results correlate significantly with tertiary entrance scores, suggesting that performance on this instrument is related to general ability. Further, performance on different components of the GSA are observed among students taking particular courses. Engineering students perform well on the problem-solving component while nursing students do well on the interpersonal skill section. Herl et al. (1999) used an alternative approach in which several tasks were set and a series of questions were asked arising from each task. This approach has been criticised by Shavelson, Gao and Baxter (1993). In a review of many similar assessments, they found that the particular task used was the major contributor to performance variance. This confounds the measurement of individual achievement and, in order to obtain a reliable estimate of ability, many tasks are required. This would place a very heavy assessment load on both students and teachers.

Thus, while several assessment options are available and each has been shown to work well in particular situations, no single approach is able to meet all assessment requirements. It would appear that holistic judgement by teachers would not work well in the VET sector as a whole, although there may be situations in which it is applicable. Work experience portfolios are used and could be extended to include employability skills, although the comments by Robertson et al. (2000) suggest that they would not lead to comparable judgements. Instrumental assessment has a particular role, and an investigation of the extension of the GSA to the VET sector would appear to have great merit. However, there does appear to be scope for the development of an assessment approach that combines the authenticity of in-context assessment with the reliability of a structured instrumental approach.

#### *Assessment as measurement*

Following concerns about the status of measurement in the social sciences, Stevens in 1946, proclaimed that measurement was the "assignment of numerals to objects or events according to a rule." Michell (1997) has shown that such assignment is a

necessary but insufficient basis for true measurement as it does not require additivity. Such assignment may produce an ordered sequence, but not an interval one. Wright and Masters (1982, p.3) argued that measurement requires:

- The reduction of experiences to a *one dimensional* abstraction;
- More or less* comparisons among people and items;
- The idea of a *linear magnitude* inherent in positioning objects along a line; and
- A unit determined by a *process* which can be repeated without modification over the range of the variable.

Harwell and Gatti (2001) have argued that the application of item response theory (IRT) is essential to convert the ordered observations that arise from the application of survey instruments to true measures.

A purpose in the project was to develop an instrument that provides a sound framework within which consistent judgements can be made by teachers and that leads to reported scores of known precision. The development of an instrument on a theoretically sound basis and the analysis of achievement data collected using it are described.

## **Research Method**

### *Development of the Problem-Solving Assessment tool*

As outlined in the review of literature, most past efforts based on instrumental approaches to assessing problem-solving have focused on individuals' performance on selected problem-solving tasks. However, the tasks chosen have been shown to contribute a substantial component of performance variability and therefore to mask the individual ability contribution to performance variability. Since the purpose of the problem-solving assessment is to identify individual ability, approaches in which this is contaminated by other factors are compromised.

In past efforts to assess problem-solving in a componential, rather than holistic, way separate scoring rubrics were developed for each task (Herl et al., 1999; Shavelson et al., 1993). If this approach were to be taken in the VET context within Australia, the load on assessors would be excessive. Each training package has many units of competency and each unit has many tasks. The process of developing separate rubrics for this number of tasks and then of providing professional development to ensure that they were used consistently would be onerous at system and provider levels and for individual assessors. Thus, in this project an intention was to develop either a single instrument or a very small number of generally applicable instruments.

The Problem-Solving Assessment (PSA) was designed to assess the use of problem-solving processes directly, as these processes are thought to be important in the emergence of expertise within a domain and also to be transferable between tasks within and possibly between domains. The development of the initial version of the instrument followed four stages:

- the identification of a coherent theoretically sound construct;
- the identification of major component processes;
- the identification of indicators of those processes; and
- the establishment of levels of performance on each indicator.

This approach is an elaboration of the first steps of the instrument development process described by Gable and Wolf (1993, see pp. 237-245). In the method used in this study, particular attention was paid to transforming theoretical constructs into

operational processes and to the identification of indicators of the processes and performance levels on those indicators. These are crucial steps in promoting achievement and assessing and reporting performance.

The Problem-Solving Assessment tool is shown in Appendix 4 of the report *The Authentic Performance-Based Assessment of Problem-Solving* (Curtis & Denton, 2002, pp.69-70). The assessment tool includes five major processes. In all, 21 indicators are used to assess performance and the number of performance levels varies from two for some indicators to a maximum of four. The performance levels are based on the Biggs and Collis (1982) SOLO taxonomy. The number of performance levels for each indicator was decided on the basis of the number of levels that could be discriminated readily by lecturers.

#### *Administration of the Problem-Solving Assessment tool*

The procedure for assessing students' problem-solving performances is:

- selection of the technical assessment task that is to be presented for problem-solving assessment;
- completion by the learner of the technical activity or task;
- a self-assessment by the learner using the Problem-Solving Assessment;
- presenting evidence for facilitator validation, including discussion between student and facilitator;
- judgement (validation) by facilitator;
- recording of result for later reporting and certification.

The assessment of problem-solving is based on an assignment that has been presented for technical assessment. The grade for the technical assessment is recorded. Then the same assessment activity is re-presented for problem-solving assessment. Re-presenting the activity for problem-solving assessment shortly after the technical assessment is regarded as necessary in ensuring that the assessment of problem-solving is authentic. In the pilot study, 25 assessment tasks were identified by staff and recommended as productive activities in which students could demonstrate their abilities. In the study, students were free to choose other tasks that they may have preferred.

Two key elements of the process are self-assessment by learners in order to collate evidence for the claimed problem-solving performance level and validation by the lecturer (facilitator). These are now described.

#### ***Self Assessment***

The assessment of problem-solving in this project was intended to form an iterative process through which students learned what was meant by problem-solving so that they could improve the process in successive assessments. Thus the assessment procedure was intended to be a very overt and explicit process of both learning and assessment.

Students and staff used the same form of the Problem-Solving Assessment tool. This was done in order to ensure that the process was a completely open one in which all assessment criteria were clearly laid out for the students. They were aware of exactly how their lecturer would conduct that part of the assessment and what evidence they would be seeking. Indeed, the reverse side of the assessment form includes brief scoring instructions and sets of questions that relate to each indicator used in the assessment. The questions are designed to focus attention on the evidence that might

be expected to support the levels of performance that have been suggested for that indicator.

### ***Lecturer Validation***

Lecturer validation comprised two elements. First, students were required to present their self-assessment forms and to present or point to evidence to support their self assessment. The second element of the process was the judgement, by the lecturer, of the problem-solving performance based upon the evidence presented by the student.

Once the lecturer has made her/his judgement, the assessment is discussed by the lecturer and the student. The purpose of this discussion is to draw attention to aspects of the assessment in which the student has either not presented evidence or has misinterpreted either the evidence or the criteria. The purpose of this discussion is clearly instructional with the aim of enhancing the student's understanding of both problem-solving and the process of assembling evidence against specified criteria.

The performance level that is recorded in the student record system is determined by the lecturer. It is essential that the processes that underlie any formal certificate awarded by the training provider are credible and robust and that they would, if required, withstand an external review such as a quality audit.

### ***Pilot and Validation Studies***

The Problem-Solving Assessment tool was trialed in a pilot study involving 33 students, all of whom were enrolled in programs offered through Electronics and Information Technology at Torrens Valley Institute of TAFE. In a subsequent validation study, a further 48 students enrolled in the Certificate IV in Workplace Assessment and Training participated.

Two sets of data were collected in the studies. All lecturer responses recorded on the instrument were entered into a data file and analysed using both conventional and advanced measurement techniques. In addition, comments were sought from both staff and students involved in the pilot study in an anonymous feedback survey.

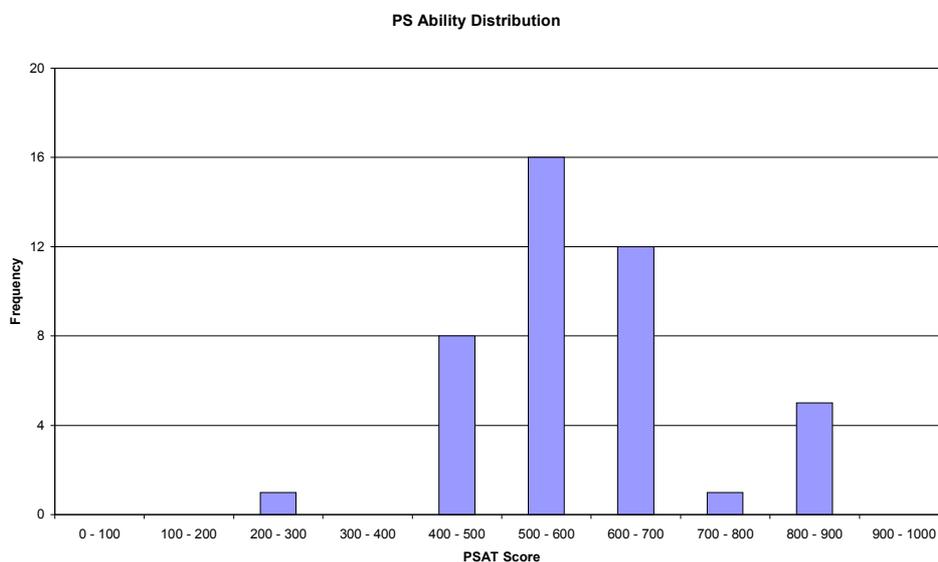
### **Key Findings**

Data arising from the administration of the PSA tool were subject to factor analysis and classical item analysis. Despite some concerns about skewness in the performance indicators, the Cronbach alpha for the scale was 0.80. This is regarded as indicating that the scale is a coherent instrument. However, given the argument advanced earlier in this paper, that valid measurement is a requirement for effective reporting of valid achievement assessment, the Rasch measurement model was used in the main analysis of the data set collected. This analysis was conducted using Quest (Adams & Khoo, 1999). Rasch measurement produces a scale on which the average location of items defines a zero point. Persons who perform below this level receive negative scores and those who perform above this level achieve positive scores. The scale has the important attribute of being an interval one, and this enables linear transformations without distorting the meaning conveyed by the scale.

In the Rasch analysis, two items were found not to fit the measurement model well and were removed from further analyses. The remaining items produced a scale with good measurement properties. An index of the reliability of items was 0.84 and an index of person reliability (equivalent to alpha) was 0.81. The range of item

thresholds – the point on the trait scale at which a person is likely to move from a lower to a higher performance category – varied from -3.1 to +3.3 units. This range would appear to provide a sound basis for the measurement of individual ability over a useful range of the underlying trait – Problem-Solving ability. The range of person abilities varied from -3 to +6 scale units. However, most participants were rated quite favourably and were clustered in the positive ranges of the scale. The skewness of responses is believed to be a consequence of the requirement for voluntary participation in the study, as a result of which there was a tendency for more able students to participate.

The scores achieved by participants were linearly transformed to a mean of 500 and a standard deviation of 100. The distribution transformed scores is shown in Figure 1. The measurement properties of the instrument suggest that four overall performance levels can be recognised in the sample of respondents (Wright & Masters, 1982). However, the tendency for lower achieving students to be under-represented in the sample may have reduced the number of discriminable levels of performance.



**Figure 1: Distribution of individuals' transformed problem-solving ability scores**

## Conclusions

Despite the fact that eighteen different tasks were the subject of problem-solving assessments, the coherence of the instrument suggests that the use of the framework instrument developed on the basis of a theory of problem-solving provides a sound basis for the assessment of problem-solving ability.

The use of the Rasch measurement model has enabled student performances to be reported on a linear scale that should provide a basis for the valid comparison of performance over time and can thus be used to track improvements in student achievement in this important employability skills through their courses.

The results of the validation study suggest that the PSA tool that was developed in one domain (electronics and information technology) could be applied in a quite disparate domain. What remains to be shown is that the PSA can be applied in other domains.

A comparable approach has been taken in developing a second assessment and development tool for Information Literacy. Other tools are under development using a similar approach. If the instruments can be shown to work in a variety of settings and programs, it would be possible to use them in encouraging the development of a range of employability skills and in assessing and reporting on performance across that range of skills. If this is achieved, the profiling of employability skills and then matching individual profiles to industry requirements becomes technically valid and feasible.

## References

- Adams, R. J., & Khoo, S. T. (1999). Quest: the interactive test analysis system (Version PISA) [Statistical analysis software]. Melbourne: Australian Council for Educational Research.
- Australian Council for Educational Research. (2001). *Graduate Skills Assessment. Summary Report. GSA Exit 2000*. Camberwell, Vic.: ACER.
- Biggs, J. B., & Collis, K. F. (1982). *Evaluating the quality of learning: the SOLO taxonomy (Structure of the Observed Learning Outcome)*. New York: Academic Press.
- Bransford, J. D., & Stein, B. S. (1984). *The ideal problem solver: a guide for improving thinking, learning, and creativity*. New York: W H Freeman.
- Curtis, D. D., & Denton, R. (2002). *The authentic performance-based assessment of problem-solving*. Adelaide: NCVET.
- Feast, V. (2000). *Student perceptions of the importance and value of a Graduate Quality framework in a tertiary environment*. Unpublished Doctor of Education Dissertation, Flinders University, Adelaide.
- Gable, R. K., & Wolf, M. B. (1993). *Instrument development in the affective domain. Measuring attitudes and values in corporate and school settings* (2nd ed.). Boston: Kluwer Academic Publishers.
- Greeno, J. G., Collins, A. M., & Resnick, L. B. (1996). Cognition and learning. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 15-46). New York: Macmillan.
- Harwell, M. R., & Gatti, G. G. (2001). Rescaling ordinal data to interval data in educational research. *Review of Educational Research*, 71(1), 105-131.
- Herl, H. E., O'Neil, H. F., Chung, G. K. W. K., Bianchi, C., Wang, S.-L., Mayer, R., Lee, C. Y., Choi, A., Suen, T., & Tu, A. (1999). *Final report for validation of problem-solving measures* (CSE Technical Report 501). Los Angeles: Center for the Study of Evaluation and National Centre for Research in Evaluation, Standards, and Student Testing.
- Lave, J. (1988). *Cognition in practice: Mind, mathematics and culture in everyday life*. Cambridge, N.Y.: Cambridge University Press.
- Mayer, R. E., & Wittrock, M. C. (1996). Problem-solving transfer. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 47-62). New York: Macmillan.
- McCurry, D., & Bryce, J. (1997). *The school-based key competencies levels assessment project. Final report*. Canberra: DEETYA.
- Michell, J. (1997). Quantitative science and the definition of measurement in psychology. *British Journal of Psychology*, 88, 355-383.
- National Industry Education Forum. (2000). *The Key Competencies portfolio approach - a kit*. Canberra: Department of Education, Training and Youth Affairs.
- Polya, G. (1957). *How to solve it. A new aspect of mathematical method* (2nd ed.). Harmondsworth: Penguin.
- Queensland Department of Education. (1997). *Assessing and reporting the key competencies of students of post-compulsory age through 'work experience'*. Canberra: Department of Education, Employment, Training and Youth Affairs.

- Reynolds, C. (1996). *Business, industry, key competencies, and portfolios*. Canberra: Department of Education, Employment, Training and Youth Affairs.
- Robertson, I., Harford, M., Strickland, A., Simons, M., & Harris, R. (2000, 19 February 2001). *Learning and assessment issues in apprenticeships and traineeships* [Conference paper]. AVETRA. Retrieved 07 June, 2001, from the World Wide Web:
- Scribner, S. (1986). Thinking in action: some characteristics of practical thought. In R. J. Sternberg & R. K. Wagner (Eds.), *Practical intelligence. Nature and origins of competence in the everyday world* (pp. 13-30). Cambridge: Cambridge University Press.
- Shavelson, R. J., Gao, X., & Baxter, G. P. (1993). *Sampling variability of performance assessments* (CSE Technical Report 361). Los Angeles: University of California, CRESST.
- Troper, J., & Smith, C. (1997). Workplace readiness portfolios. In H. F. O'Neil (Ed.), *Workforce readiness. Competencies and assessment* (pp. 357-382). Mahwah, NJ: Lawrence Erlbaum.
- Wright, B. D., & Masters, G. N. (1982). *Rating scale analysis*. Chicago: MESA Press.

### **Acknowledgment**

The research reported in this paper was supported under a grant from the National VET Research and Evaluation Committee (NR2007) administered by the National Centre for Vocational Education Research.

The cooperation of staff and students in the Electronics and Information Technology program at the Torrens Valley Institute of TAFE is gratefully acknowledged.