2005

Measuring preparedness to teach with ICT

Patricia A. Forster
Edith Cowan University, p.firster@ecu.edu.au

Vaille M. Dawson
Edith Cowan University, v.dawson@ecu.edu.au

Doug Reid
University of Wollongong, dougr@uow.edu.au

Measuring preparedness to teach with ICT

Abstract

In this paper we discuss the development and implementation of a questionnaire that measures preparedness to teach secondary school science with information and communication technologies (IeT). The questionnaire was designed for the purpose of evaluating the effectiveness of instruction in a science education unit in a Graduate Diploma teacher education course. Rasch analysis of the pre- and post-unit responses indicated domains of expertise for which students perceived they had improved their knowledge during the unit, and domains for which they perceived they were less knowledgeable after the unit than before it. We discuss students' responses in relation to the unit, and report the technical decisions that we 'made as part of the analysis. The questionnaire could be adapted easily to suit preservice teacher education in disciplines other than science.

Keywords

ict, teach, measuring, preparedness

Disciplines

Engineering | Science and Technology Studies

Publication Details

Forster, P. A., Dawson, V. M. & Reid, D. (2005). Measuring preparedness to teach with ICT. Australasian Journal of Educational Technology, 21 (1), 1-18.

Australasian Journal of
Educational Technology

2005, 21(1), 1-18

Measuring preparedness to teach with ICT

Patricia A. Forster, Vaille M. Dawson and Doug Reid
Edith Cowan University

In this paper we discuss the development and implementation of a questionnaire that measures preparedness to teach secondary school science with information and communication technologies (ICT). The questionnaire was designed for the purpose of evaluating the effectiveness of instruction in a science education unit in a Graduate Diploma teacher education course. Rasch analysis of the pre- and post-unit responses indicated domains of expertise for which students perceived they had improved their knowledge during the unit, and domains for which they perceived they were less knowledgeable after the unit than before it. We discuss students' responses in relation to the unit, and report the technical decisions that we made as part of the analysis. The questionnaire could be adapted easily to suit preservice teacher education in disciplines other than science.

Introduction

Since the 1970s various information and communication technologies (ICT) have been appropriated for teaching and learning, including digital cameras, projection technologies, and the world wide web. Education authorities, internationally, are promoting the use of ICT. For example, an expected outcome in the UK is that:

> ...pupils become increasingly independent users of ICT tools and information sources. They have a better understanding of how ICT can help their work in other subjects and develop their ability to judge when and how to use ICT and where it has limitations. They think about the quality and reliability of information, and access and combine increasing amounts of information. They become more focused, efficient and rigorous in their use of ICT, and carry out a range of increasingly complex tasks. (Department for Education and Employment, 1999, p. 20)

In Western Australia it is expected that:

> ...students choose between or integrate various technologies for a purpose. They adapt familiar or existing technologies to meet the demands of new

> tasks or situations. As confident and capable users of a wide range of technological applications and processes, they critically appreciate the consequences of technological innovation. (Curriculum Council, 1998, p. 22)

Further, a recent report by the British Educational Communications and Technology Agency BECTA (2003) identifies that teachers must be trained in the technical and pedagogical aspects of ICT. This paper is part of an action research study addressing such training. The context is preservice teacher education in science.

The study commenced in 2002 with planning, implementing and critically assessing instruction on teaching and learning with ICT, in a Graduate Diploma science education unit for preservice secondary teachers, which extended over eight weeks. Student learning outcomes were encouraging. A survey of students at the end of the unit indicated that 97% (37/38) would use ICT in their teaching if the resources were available, while 76% (29/38) of students agreed that they had better ICT skills as a result of studying the unit. A significant component of the unit was an authentic assessment where students were required to produce an ICT resource suitable for use in lower secondary science (see Dawson & Reid, 2003). Ninety percent (34/38) of the students indicated they thought they had improved their ICT skills through completing the assignment.

The second phase of the study took place in 2003 with a new student cohort (n=42) in the science education unit. Consistent with the action research paradigm (Carson & Sumara, 1997), the purpose of the second phase was building on what had been learnt in the first phase in order to develop further the ICT curriculum. As part of the second phase, a questionnaire was designed to measure *students' perceptions of preparedness to teach science with ICT*. The questionnaire was administered early in the unit and again after instruction and project work on ICT use had concluded. The principal purpose for the questionnaire was that comparison of the responses from the two implementations would assist unit evaluation: the comparison would allow us to identify the domains in which students perceived they were more prepared to teach with ICT after completing the unit, and the domains in which they perceived they were less prepared. A second purpose was that the first implementation would raise students' awareness of the scope of ICT use.

The main topic of this paper is the questionnaire. First, we place its design in context by describing the theoretical views on learning that guided instruction and the scope of the ICT curriculum. Sections on construction of the questionnaire and the analysis of responses using Rasch modelling follow. The Rasch model is suited to analysing a single trait, which in the case of the questionnaire was 'preparedness to teach science with ICT'.

Finally, we discuss the Rasch outputs in relation to the conduct of the unit, and discuss future refinement of the questionnaire and development of the unit.

Learning theory and the scope of the ICT curriculum

A constructivist epistemology (Dawson & Taylor, 1998; Taylor, 1998; Tytler, 2002) underpinned the selection and implementation of learning activities for the science education unit overall. In particular, open ended activities were included which allowed students of varying ability and experience to achieve success and advance their knowledge. Small group and whole group discussions were facilitated, in which pedagogy and learning outcomes were discussed critically. Authentic assessments were put in place, where the work was not solely for the purpose of grading but students could use the products in their future teaching. It is noted that students who enrol for the science education unit have diverse backgrounds. They have different undergraduate degrees and workplace experiences. Some are mature age and are changing career.

Three other principles underpinned the design of the curriculum addressing the use of ICT. First was that learning how to use technology for technology's sake is not productive: instead, technology should make available knowledge or content that enhances students' learning (Leamnson, 2001). Second, tertiary educators need to help preservice teachers develop a clear vision about how they can facilitate student learning with computers (Wang, 2002). Third, one of the ways this clear vision can be formed is by modelling appropriate ICT skills in tertiary education: "as preservice teachers see technology modelled and as they are provided with more opportunities to use technology in the classroom setting, high anxiety levels stemming from negative attitudes toward computers will be lessened" (Pope, Hare & Howard, 2002, p. 201).

The actual initiatives taken in regards to ICT were as follows. Unlimited access to a laptop computer, a data projector and a digital camera was secured for teaching purposes, for the duration of the unit. This enabled the instructor (second author) to model the use of ICT. Two consecutive three hour sessions were used to familiarise students with ICT resources that are used in secondary school science including: interactive applets and simulations; the interactive CDs *Biotechnology Online* (Biotechnology Australia, 2001) and *The Ultimate Human Body* (Dorling Kindersley, 1996); electronic portfolios and self paced online modules; web sites for students and teachers; electronic textbooks including Heinemann's *Science for Western Australia* series (Cahill & Spence, 1999); data loggers for temperature, humidity and heart rate; video resources; the digital camera, data projector and laptop, and Lego *Dacta*. The second session was

conducted in a computer laboratory. During these sessions, students were able to 'play' with the equipment and CDs and peer coaching was encouraged. The instructor and a technical assistant (third author) were available for advice. Students were made aware that they could use the equipment in situ for science education related activities, and could borrow the software resources for their final teaching practice which commenced after the conclusion of the unit.

Subsequently, students were asked to discuss with each other and as a whole group the advantages, disadvantages and issues associated with using ICT resources in secondary science. Advantages that were voiced were that: ICT can be useful for one to one interactive learning, remediation, revision, visual learners and lateral thinkers. Disadvantages related to technical issues, and access within schools. Other issues that were discussed were that: ICT use must add to educational outcomes for students; use of ICT must suit teachers' and students' comfort levels; and content of ICT resources must suit students' interests and ability levels. In addition, the instructor raised the issue of copyright in relation to ICT resources. Students were provided with the Year 11 and 12 science syllabus documents for Western Australia on a CD. They were told that copyright permission had been obtained for reproduction of the documents and that copyright requirements need to be addressed with any reproduction and use of web resources in the classroom.

For assessment of the ICT component of the unit, each student was asked to develop a curriculum resource that would require lower secondary science students to use ICT and they were given the option of having their resource placed on a CD that would be made available to all students enrolled in the unit. It was expected that students would develop their own ICT skills through completing the assessment requirements and, although they were required to individually produce a resource, they were encouraged to work collaboratively and seek technical assistance. Because of the diversity of technical skills, they were given a choice of presentation and submission formats. For example, the ICT resource could comprise a web quest worksheet, a web site, an online assessment or an interactive *Powerpoint*. The assignment could be submitted online, by email attachment, CD, floppy disc, hard copy or a combination. Students were asked to: provide instructions on how the resource could be used; discuss the advantages and disadvantages when compared to traditional resources; and describe how their resource was cognisant of current theories of teaching and learning.

The assignment contributed 20% towards the unit mark. The allocation of marks was based primarily on creativity, accuracy of content and pedagogy rather than ICT skills, so as not to discourage and disadvantage

students with limited computing skills. Students were provided with an outcomes based rubric at the outset that was used to assess their resource. The ICT curriculum and the resources that were submitted are discussed in more detail in Dawson, Forster & Reid (2004).

The questionnaire

Construction of the questionnaire for measuring students' preparedness to teach science with ICT commenced with writing a bank of items and designing suitable Likert scales. Then, items were chosen from the bank, the interpretative validity of them was checked, and the instrument was finalised.

Drawing on the experience of the first phase of the study and our other experience in teaching with ICT, we created a bank of 72 questionnaire items. They addressed: *skills* to do with *Word* documents, spreadsheets, *Powerpoint*, the world wide web, digital cameras and other equipment; and *knowledge* of ICT resources, pedagogy with ICT, and schemes for assessing students' performance and for assessing the merit of resources. Items were intentionally made short and addressed one topic, to minimise ambiguity. Negatively worded items were avoided, as they are more open to multiple interpretations (Peck, 2000), responses to them are likely to be anomalous (Taylor, Fraser, & Fisher, 1997), and responses can favour the lower end of Likert scales (Andrich & van Schoubroek, 1989).

We intended using a single Likert scale but found it was impossible to design one that suited the measurement of competence in regards to skills *and* knowledge in the various domains. So, we decided on different scales for skills and knowledge. For skill items, we chose *no experience, novice, competent,* and *highly competent*. Our rationale for variation in the wording was that it would encourage more deliberate responses than categories like *no competence, some competence, competent,* and *highly competent*. The variation was in accordance with design strategies reported in Wildy, Forster, Louden, and Wallace (2004). The scale for knowledge items was *none, one, two,* and *more than two,* which indicated the number of times students could demonstrate particular types of knowledge.

After deciding the scales, 36 items were selected from the item bank for use in the questionnaire. They were chosen to cover ICT use that would be treated explicitly in the science education unit. We grouped the items under four major headings: Skills, Resources, Pedagogy and Critique, and used subheadings to distinguish skill and resource types. The subheadings for Skills were World Wide Web, *Powerpoint,* and Computer related equipment. Subheadings for Resources were World Wide Web and CDs.

The headings and subheadings were included in the questionnaire to aid respondents' interpretation of individual items. Extracts from the questionnaire are provided below. The full set of items is provided later in the paper (Table 1).

Skills

Please tick **ONE** box for each item.

WORLD WIDE WEB	No experience	Novice	Competent	Highly competent
I know how to				
1. -search the web.	☐	☐	☐	☐
2. -create a web page with text.	☐	☐	☐	☐

Resources

WORLD WIDE WEB	None	One	Two	More than two
I can name web sites with				
3. -graphics for learning science.	☐	☐	☐	☐
4. -information for preparing my teaching.	☐	☐	☐	☐

For interpretative validity, we asked three Graduate Diploma students to complete the questionnaire, separately, and to identify ambiguities and items they were uncertain about. As a result we reworded two items and deleted two. The finalised questionnaire comprised 34 items and was implemented in the first workshop of the science education unit, before any instruction on teaching and learning with ICT in science, and eight weeks later, at the end of the unit and after project work had been submitted. It was completed anonymously except that we asked students to include their student number so that we could match pre-unit and post-unit responses. We guaranteed the numbers would not be matched with names. To encourage genuine responses, we guaranteed also that the analysis of the responses would take place after grading for the unit had been completed. We acknowledge that any gain in knowledge implied by comparison of pre- and post-unit responses could have been due to students' experience outside the unit (e.g., in other education units) as well as to experience within it.

Analysis of responses

Rasch modelling

The Rasch model for ordered response categories (Andrich, de Jong & Sheridan, 1997) and the rack and stack approach described by Wright (1996, 2003) were used to analyse the data, with the RUMM2010 software (Andrich, Sheridan & Luo, 2000). The Rasch model locates persons and items on a scale, where locations reflect person ability (i.e., the extent to which persons possess the trait being examined) and the difficulty of items (high difficulty is synonymous with few 'highly competent' responses).

More precisely, the difficulty value for an item equals the natural logarithm of the odds of being successful on the item (Cavanagh & Romanoski, 2004). Being successful is responding in the top half of the Likert categories. If a person and item have the same location (ie., the same value on the scale), then the probability that he or she was successful on the item is 50%. The probability that persons were successful on items below them on the scale is greater than 50%, and the probability they were successful on items above them is less than 50%.

Racking and stacking (Wright, 2003) allows the treatment of pre-test and post-test data. With racking, the items in the pre-test and post-test are considered to be different items. The rationale is that the same students are involved in each test but their knowledge about the items changes - in other words, the items change for them. The purpose of racking is the examination of the difficulty of individual items, in other words, the tendency of students to respond at the left hand end or right hand end of the Likert scale on each item in the pre-test and on each item in the post-test.

Stacking (Wright, 2003) involves treating students who completed the pre-test and post-test as different students, each of whom answered the same items. The rationale for stacking is that students change as a result of a course of instruction, so they are effectively different students. The purpose of stacking is the examination of the ability of individual students across all items in the pre-test and across all items in the post-test.

A number of checks are instituted with Rasch modelling, to establish that data are suited to the model. In the first instance, we checked (a) the thresholds and (b) discrimination between students for each item in the racked data and, later, we conducted the same checks for the items in the stacked data. Thresholds indicate the abilities of persons for whom adjacent scores on the Likert scale are equally likely. The first threshold for

an item is the ability of persons for whom scoring 0 and 1 is equally likely; the second threshold is the ability of persons for whom scoring 1 and 2 is equally likely; and so on for the third and other thresholds. Reversed thresholds, where the thresholds are out of order (eg., the third threshold is between the first and second thresholds) indicate the Likert scale does not work well for the item. The thresholds should indicate a monotonic (one direction) response process (Hagquist & Andrich, 2004). Reversed thresholds can be addressed by rescoring items to combine adjacent categories of response, or problematic items can be deleted from the analysis.

Discrimination between students is checked visually using item characteristic curves (ICCs). The ICC for an item displays the responses predicted by the model for persons across the range of abilities, and the mean actual responses on the item when persons are grouped into class intervals. The discrepancies between predicted scores and mean actual scores indicate how well data fit a model.

Fit is checked using chi-square probability values. Generally, probabilities < 0.05 imply poor fit, however, probability depends on sample size and the criterion of 0.05 is most appropriate for data sets with 50-250 points (Linacre, 2003). Moreover, a problem with deleting poor fit items is that probabilities associated with the remaining items change and may indicate misfit. Therefore, following the method of Hagquist and Andrich (2004), we used ICCs and chi-square probabilities for detecting poor fit, but rejected items only if there were problems with the wording of them or possible peculiar interpretation by the students.

Two final statistics that measure reliability were checked for the racked and stacked data: the separation index, which indicates how well item thresholds cover the range of person abilities; and the chi-square probability for the item-trait interaction, which indicates how well students agree on difficulties of items across the scale and whether the assumption of unidimensionality (analysis of a single trait) is reasonable (Cavanagh & Romanoski, 2004).

The Rasch analysis

Thirty-seven students of the cohort of 42 completed both the pre-unit and post-unit implementations of questionnaire. Three of the 37 did not complete the last page in the pre-test and another student ticked repeatedly the same Likert category down the pages of the post-test. Hence, data for 33 students were suitable for analysis.

We scored students' responses (no experience 0, novice 1, competent 2 and highly competent 3; none 0, one 1, two 2, more than two 3), then entered the pre-test and post-test data into the RUMM software, in racked format (as though 33 students had answered 68 different items), and applied the single factor, polytomous unrestricted Rasch model to the data. Multiple items displayed reversed thresholds and, in large part, they were not corresponding pre-test and post-test items. Others (eg., Andrich et al., 1997; Wright, 1996) have also observed that Likert scales can work differently in two implementations of an instrument so that, not only do responses shift along the Likert scales, but the structure of responses changes. The cause is that students' interpretation of items changes fundamentally over time.

We addressed the reversal problem by rescoring the problematic items as 0, 0, 1, and 2. Multiple items still displayed reverse thresholds, so we tried scores of 0, 1, 1, and 2. The scoring worked well for all except three items. Inspection of the response patterns on the three items showed responses were at the ends of the scales and not in the middle. Hence, we rescored the three items as 0, 0, 1, and 1. Finally, we checked that corresponding items in the pre-test and post-test were scored the same way, so that results on them could be compared.

We used two class intervals only for the analysis of the ICCs because the data set was small (n=33). Visual inspection of the ICCs indicated that the most grossly misfitting item was in the pre-test (Item 23). This item had the lowest chi-square probability of all items (p=0.005). However, (a) discrimination was in the correct direction (actual mean scores for students in the first class interval were lower than the actual mean scores for students in the second class interval), (b) data for the corresponding item in the post-test fitted the model well (p=0.646), and (c) inspection of the wording did not indicate any reason to reject the item. Therefore it was retained. Three other items had chi-square probabilities of less than 0.05 (p= 0.021, 0.022 and 0.042) and were retained on the same basis. The separation index for the final set of racked data was 0.962, and the chi-square probability for item-trait interaction was 0.464, which indicated high reliability and mapping onto a single trait.

The analysis of stacked data, where the students in the pre-test and post-test were treated as (2x33) different students who had each answered the same (34) items, proceeded as follows. The scoring of responses for racked data was retained, so responses on most items were scored 0, 1, 1, 2. No items showed reversed thresholds, and discrimination between groups was appropriate for all items, according to ICCs with three class intervals and the chi-square probabilities, which were greater than 0.05 for all items. Three class intervals instead of two were chosen because of the larger number of data points (66 instead of 33). The separation index was 0.939

and the chi-square probability for item-trait interaction was 0.803, which indicated high reliability.

Interpretation of results: Racked data

The issue of interest with the racked data was the location of 68 items, 34 from the pre-unit implementation and 34 from the post-unit implementation. The item locations are shown in Table 1, together with changes in location (post-test location - pre-test location).

The items with low negative locations attracted high numbers of responses at the right hand end of the Likert scales. Negative changes in location indicated that students as a group knew more at the end of the unit than at the beginning about the topic addressed by the that item (whether due to their experience in the unit or outside it). For example, pre-test Item 1 (I know how to search the web) was located at –2.68, so attracted a relatively high number of 'highly competent' responses. The same item in the post-test was located at -3.50, so attracted an increased number of responses at the 'highly competent' end of the scale. The change in location was negative (-0.82) and indicated that, overall, students' competence in searching the web was greater at the conclusion of the unit than before it. The distribution of responses for the item 'I know how to search the web' is shown in Figure 1.

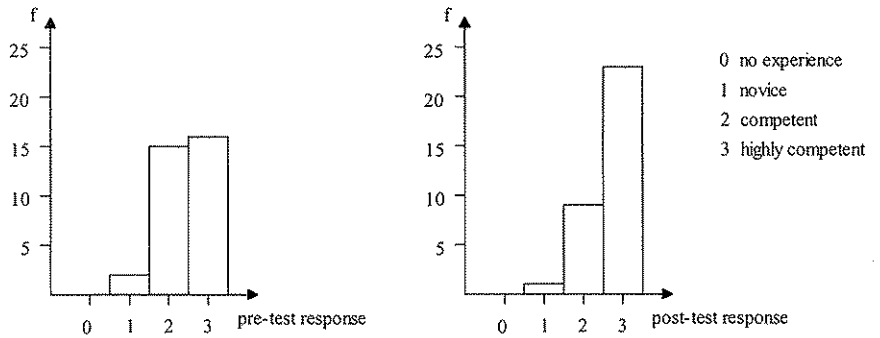

Figure 1: Pre and post test responses on 'I know how to search the web'.

Items with high locations (see Table 1) attracted relatively low numbers of responses at the right end of the scales, and positive changes in location between corresponding items indicated students as a group perceived their competence/. knowledge was less at the end of the unit than at the beginning, in regards to the topics addressed by the items. For example, pre-test Item 4 (I know how to import video into a web page) was located

at 2.85 and the same item in the post-test was located at 4.78. The change in location was 1.93. So, students' responses indicated that, as a group, they started knowing little about importing video into a web page, and perceived they knew less by the end of the unit. Their awareness of limited knowledge was raised over the duration of the unit and the unit apparently did not successfully address knowledge in this domain.

Table 1: Location of pre-test and post-test items, and changes in location

Item	Pre-test location	Post-test location	Change in location	Item description
Skills				
1	-2.68	-3.50	-0.82	I know how to search the web.
2	2.28	1.15	-1.13	I know how to create a web page with text.
3	1.88	1.02	-0.85	I know how to import graphics into a web page.
4	2.85	4.78	1.93	I know how to import video into a web page.
5	3.99	1.92	-2.07	I know how to create an animated icon.
6	1.58	0.56	-1.02	I know how to determine copyright restrictions to using web sites in the classroom.
7	0.16	-1.11	-1.26	I know how to prepare a power-point demonstration with text.
8	0.84	-0.02	-0.86	I know how to prepare a power-point demonstration with sound.
9	0.62	-1.03	-1.64	I know how to import graphics into power-point.
10	1.30	0.96	-0.34	I know how to import video into power-point.
11	1.12	1.16	0.04	I know how to connect a computer to a data projector.
12	-0.20	-0.29	-0.10	I know how to take pictures with a digital camera.
13	0.82	0.62	-0.20	I know how to load still pictures from a digital camera onto a computer.
14	1.60	1.31	-0.29	I know how to load video from a digital camera onto a computer.
15	2.10	2.01	-0.09	I know how to use data loggers to measure temperature.
Resources				
16	0.27	-1.16	-1.43	I know of web sites with graphics for learning science.
17	-0.89	-4.14	-3.25	I know of web sites with information for preparing my teaching.
18	-0.26	-1.93	-1.67	I know of web sites with information for students project work.
19	2.28	0.36	-1.93	I know web sites with marking keys for assessments.
20	0.83	-0.07	-0.90	I know of web sites that demonstrate experiments.
21	-0.60	-2.14	-1.54	I know of web sites that students could use for learning science content.
22	1.69	0.65	-1.04	I know of web sites that allow students to contribute data.
23	1.11	-0.53	-1.64	I can name series of textbooks that are available on CD.
24	1.26	-0.28	-1.53	I can name specialised CDs that could assist teaching science.
Pedagogy				
25	-1.82	-1.96	-0.14	I can name topics that would be suitable for research projects using the web.
26	-1.34	-1.97	-0.63	I can name topics where computer and non-computer activities could be integrated.
27	0.88	-0.30	-1.18	I can name ways in which I could assess students' web-based work.
28	-1.12	-4.94	-3.82	I can name problems that can occur in having students use the web.
29	0.18	-0.12	-0.30	I can name ways in which boys and girls reactions to using computers are likely to differ.
Critique				
30	0.80	-0.90	-1.70	I can name criteria on how to judge if interactive web resources would be good for learning.
31	-0.60	-2.03	-1.44	I can name reasons why computer learning does not suit some students.
32	-0.85	-1.39	-0.54	I can name benefits to students from working in pairs on computer applications.
33	-1.45	-4.78	-3.33	I can name disadvantages of having students work together on computer applications.
34	0.85	-1.42	-2.27	I can name ways in which I could assess the effectiveness of my web-based teaching.

The change in location for most items was negative (see Table 1), implying gains in knowledge/competence. The result is highlighted in Figure 2, which shows the plot of post-test location against pre-test location for all items. Points below the line belong to items for which the change in location was negative. Points close to the line belong to items for which the

change in location was small, implying only marginal changes in knowledge.

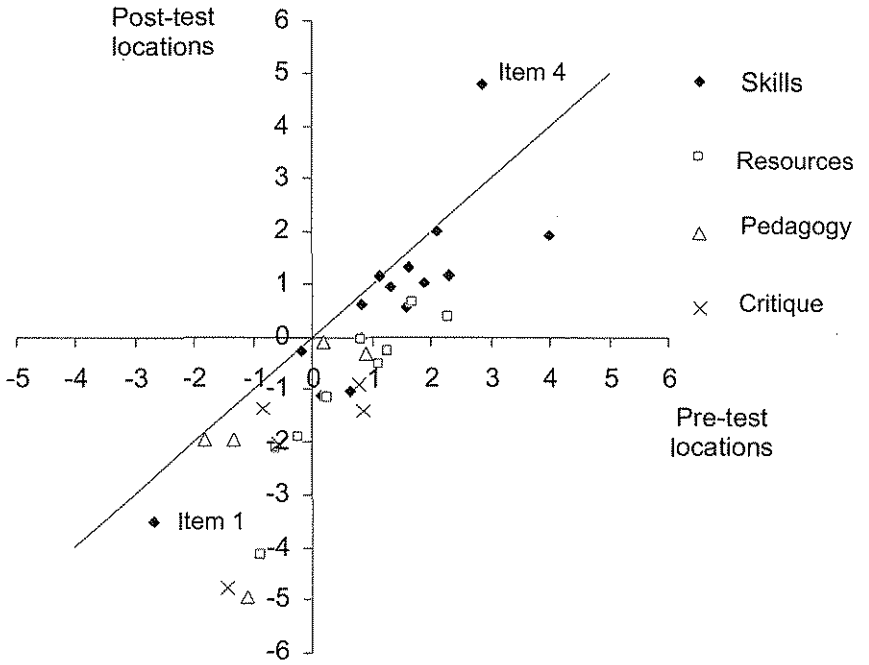

Figure 2: Pre-test and post-test locations for all items, together with the line $y=x$.

The post-test results, in particular, had implications for the conduct of the unit and future units for the student cohort. We divided the items into three groups according to post-test locations - the eleven items with highest locations (the points that were highest in the vertical direction on the graph), the eleven items with the lowest locations, and a middle group of 12 items.

High post-test locations implied relatively low levels of knowledge after the unit. Most items in this group belonged to the skill category. They were skill Items 2-6, 10-11, 13-15 and resource Item 22 (see Table 1). The ways the relevant topics were addressed in the course were interrogated. In particular, the instructor for the course was concerned about the results for Items 6 and 11. On several occasions, she had described and demonstrated how to determine copyright restrictions (Item 6) and each student signed a

copyright release form for their own assessment resource. In future, students will be given written guidelines for determining copyright restrictions and asked to discuss the ramifications of not obtaining a copyright clearance. In relation to Item 11, the instructor used the data projector and computer during most sessions of the unit. In addition, she specifically demonstrated the set up procedures. Several groups of students borrowed the data projector for presentations in other units. However, an implication of the questionnaire result for Item 11 was that students, in general, should be directed to practise and be given support in connecting the data projector with a laptop. The results for Item 15 indicated wider practice with data loggers was needed for expertise to develop. However, it was decided to delete treatment of data loggers from the unit in the future because substantial time would be required for comprehensive treatment and would be at the expense of other important topics. Regarding Item 22, it was realised that web sites for students to contribute data had been mentioned, but had not been demonstrated, and they were last on the list of web addresses given to students to explore. They will be highlighted more in the future.

The results to do with techniques for creating web pages (Items 2-5), *Powerpoint* (Item 10), and digital cameras (Items 13-14) were not a major concern, for the purpose of the ICT intervention was not to increase ICT skills *per se*, and students can enrol in an 'Introduction to Computing' unit for the development of such skills. Rather, the ICT focus in the unit was enhancing knowledge about ICT pedagogy and critical awareness of ICT resources within the context of the science education. However, the assessment will be modified for future use, to include guidelines on basic ICT skills, and technical assistance will be more widely available in relation to the assessment. Only those students who chose in their authentic assessment to construct a web page or a *Powerpoint* resource, and use a digital camera were likely to have developed the associated skills.

The eleven items with lowest locations, which indicated relatively high levels of knowledge/competence after the unit, came from all categories. They were skills Item 1; resources Items 17-18, 21; pedagogy Items 25-26, 28; and critique Items 31-34. These items pointed to domains where students' knowledge/competence more adequately prepared them for teaching science with ICT and, by implication, pointed to domains of knowledge that were treated adequately by the unit. Hence, the questionnaire results indicated that the emphasis on pedagogy and critique in the unit had impacted favourably on students' learning.

The remaining items (with middle locations) indicated skills, resources, and aspects of pedagogy and critique that could be given more emphasis in future years in the science education unit, for the potential benefit of the

students. They were skills Items 7-9, 12; resource Items 16, 19, 23-24; pedagogy Items 27, 29; and critique Item 30.

Interpretation of results: Stacked data

The issue of interest with the stacked data was the pre-test and post-test locations of each student. The locations are graphed in Figure 3. High locations indicate students chose relatively high numbers of responses in the right half of the Likert scales. Therefore, high locations signify relatively high preparedness to teach science with ICT (as judged by the students). Points above the line $y=x$ indicate that post-test locations were greater than pre-test locations. Hence, the majority of students chose more responses to the right on the Likert scales in the post-test than in the pre-test, which implies they felt more prepared to teach science with ICT after the unit than before it.

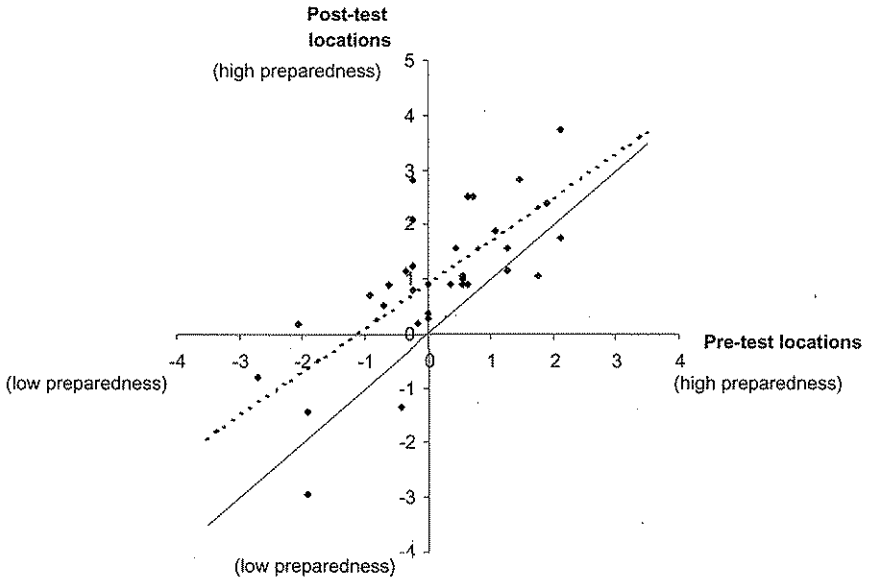

Figure 3: Pre-test and post-test person locations, indicating high and low preparedness to teach science with ICT, together with the least squares regression line for the data (dotted) and the line $y=x$ (solid).

Points on the graph (Figure 3) that are close to the line $y=x$ and above it indicate marginal gains, whereas points distant from the line indicate greater gains. Furthermore, judging by the gap between the least squares regression line and the line $y=x$, gains for students with low pre-test

locations tended to be greater than gains for students with higher pre-test locations. Therefore, students who started the unit with relatively low levels of preparedness to teach with ICT may have gained more from the unit than some of the students who started with higher levels of preparedness. Even so, the points in the third quadrant indicate low post-test locations, and indicate students for whom special attention may be warranted in future units, in regards to moving them towards using ICT for teaching science.

Conclusion

In regards to the instrument, finding a single Likert scale that would address skills *and* knowledge to do with preparedness to teach science with ICT proved impossible. As a result, we included two scales in the questionnaire. Rasch analysis of the response data after rescoring (to 0, 1, 1, 2) indicated the items and scales were directed sufficiently at a single trait ('preparedness to teach science with ICT').

Rescoring of responses was necessary because initial analysis revealed widespread discrepancies in the patterns of response. The results suggested that four categories of response were too many. Collapsing the middle two categories allowed more satisfactory analysis. Therefore, for future implementations of the questionnaire we will use three categories of response on each Likert scale (no experience, some experience, competent - for skill items; and none, one or two, more than two - for knowledge items).

We recommend the Rasch modelling and graphical display of the outputs as appropriate means for unit evaluation. The scatter plot of pre-test and post-test item locations facilitated quick discernment of domains of increased knowledge/competence (as indicated by students' responses) and domains for greater attention in the future, in the science education unit. The scatter plot of pre-test, post-test person locations facilitated analysis of individuals' increased preparedness to teach science with ICT (as indicated by their responses) and highlighted the individuals who could benefit from more assistance. The Rasch model addresses many limitations of traditional models. In particular, it does not assume a linear relationship between scores on a Likert scale and competence/knowledge.

However, we recognise that the questionnaire measured students' perceptions of the extent of their knowledge and competence, and did not elucidate the extent to which they could demonstrate knowledge and competence. The anomaly of students as a group indicating decreased competence in regards to importing video into a web page highlights that

perceptions were being measured. Nevertheless, questionnaire results alerted us to possible domains of low expertise and knowledge that did not become apparent during the conduct of the unit or in the formal assessment. Consequently, treatment of some aspects of ICT will be modified for future conduct of the unit. The changes include more explicit treatment of copyright and use of the data projector, and documentation will be provided on basic techniques and more assistance will be available on techniques for project work.

The outcomes of the questionnaire also indicated that students from the whole range of ICT backgrounds felt more prepared to teach with ICT after the unit than before it, and therefore offered evidence that the principles underpinning instruction were sound. Hence, the science education unit will continue to include and expand on opportunities for the instructor to model the use of ICT, and for students to (a) increase their awareness of ICT resources through using them and (b) increase their understanding of effective ICT pedagogy through discussion and project work.

Acknowledgement

The study of which this paper is part was supported by a grant from Edith Cowan University.

References

Andrich, D., de Jong, J. H. A. L. & 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. 59-70). New York, NY: Waxmann.

Andrich, D., Sheridan, B. & Luo, G. (2000). *RUMM2010: A Windows interactive program for analysing data with Rasch unidimensional models for measurement*. Perth Western Australia: RUMM Laboratory.

Andrich, D. & van Schoubroek, L. (1989). The general health questionnaire: A psychometric analysis using latent trait theory. *Psychological Medicine*, 19, 469-485.

Biotechnology Australia (2001). *Biotechnology Online School Resource*. Canberra, ACT: Commonwealth of Australia. [verified 23 Oct 2004] http://www.biotechnology.gov.au/biotechnologyOnline/Resource/Resource.htm

British Educational Communications and Technology Agency (2003). *What research says about using ICT in science. Report to the BECTA ICT Research network*. [15 Feb 2004, verified 23 Oct 2004] http://www.becta.org.uk/research/ictrn/

Cahill, J. & Spence, R. (1999). *Heinemann science for Western Australia 2: An outcomes based approach*. Post Melbourne, Vic: Reed International.

Carson, T. R. & Sumara, D. (1997). *Action research as a living practice.* New York, NY: Peter Lang.

Cavanagh, R. F. & Romanoski, J. T. (2004). Application of the Rasch model to develop a measure of classroom information and communication technology learning culture. Paper presented at the Second International Rasch Conference, Murdoch University, Western Australia, January.

Curriculum Council of WA (1998). *Curriculum Framework for Kindergarten to Year 12 Education in Western Australia,* Perth, Western Australia: Curriculum Council of WA.

Dawson, V. M. & Reid, D. (2003). Preparing preservice science teachers for ICT rich secondary classrooms. Paper presented at the 34th annual conference of the Australasian Science Education Research Association (ASERA), Melbourne, July.

Dawson, V., Forster, P. & Reid, D. (2004). Evaluating preservice teachers' ICT pedagogy skills. Paper presented at the annual conference of the Australasian Science Education Research Association (ASERA), Armidale, New South Wales, July.

Dawson, V. & Taylor, P. C. (1998). Establishing open and critical discourses in the science classroom: Reflecting on initial difficulties. *Research in Science Education,* 28(3), 317-336.

Department for Education and Employment (1999). *Information and communication technology: National curriculum for England.* London: Department for Education and Employment and Qualifications and Curriculum Authority.

Dorling Kindersley (1996). *The Ultimate Human Body 2.0.* London: Author.

Hagquist, C. & Andrich, D. (2004). Is the sense of coherence-instrument applicable on adolescents? A latent trait analysis using Rasch-modelling. *Personality and Individual Differences,* 36, 955-968.

Leamnson, R. (2001). Does technology present a new way of learning? *Educational Technology & Society,* 4(1), 75-79. [verified 5 Oct 2004] http://www.ifets.info/journals/4_1/leamnson.html

Linacre, J. M. (2003). Size vs. significance: Standardized chi-square fit statistic. *Rasch Measurement Transactions,* 17(1), 918. [2 Feb 2004, verified 23 Oct 2004] http://www.rasch.org/rmt/rmt171n.htm

Peck, R. (2000). *Ethnicity: A study of participation, aspirations and achievement in education.* Unpublished doctoral thesis, Murdoch University, Western Australia.

Pope, M., Hare, D. & Howard, E. (2002). Technology integration: Closing the gap between what preservice teachers are taught to do and what they can do. *Journal of Technology and Teacher Education,* 10(2), 191-203.

Taylor, P. C. (1998). Constructivism: Value added. In B. Fraser & K. Tobin (Eds), *The international book of science education* (pp. 1111-1123). Dordrecht: Kluwer Academic.

Taylor, P. C., Fraser, B. J. & Fisher, D. L. (1997). Monitoring constructivist learning environments. *International Journal of Educational Research*, 27, 293-302.

Tytler, R. (2002). Teaching for understanding in science: Student conceptions, research, and changing views of science. *Australian Science Teachers' Journal*, 48(3), 14-21.

Wang, Y.-M. (2002). When technology meets beliefs: Preservice teachers' perception of the teacher's role in the classroom with computers. *Journal of Research on Technology in Education*, 35(1), 150-161.

Wildy, H., Forster, P., Louden, W. & Wallace, J. (2004). The international study of leadership in education: Monitoring decision making by school leaders. *Journal of Educational Administration*, 42(4), 416-430.

Wright, B. D. (1996). Time 1 to time 2 comparison. *Rasch Measurement Transactions*, 10(1), 478-479. [2 Feb 2004, verified 23 Oct 2004] http://www.rasch.org/rmt/rmt101f.htm

Wright, B. D. (2003). Rack and stack: Time 1 vs. time 2. *Rasch Measurement Transactions*, 17(1), 905-906. [2 Feb 2004, verified 23 Oct 2004] http://www.rasch.org/rmt/rmt171a.htm

Patricia A. Forster, Vaille M. Dawson and Doug Reid
School of Education, Edith Cowan University, Perth, Western Australia
p.forster@ecu.edu.au, v.dawson@ecu.edu.au, d.reid@ecu.edu.au
http://www.ecu.edu.au/ses/educ/