ELECTRONIC ASSESSMENT IN HIGHER EDUCATION

Roelien Brink
University of Johannesburg, South Africa
rbrink@uj.ac.za

Geoffrey Lautenbach
Department of Mathematics, Science, Technology & Computer Education
University of Johannesburg, South Africa
geoffl@uj.ac.za

Abstract

Assessment is an important cornerstone of education. "Assessment is central to the learning process and is a crucial aspect of teaching. It is the most significant factor that influences student learning" (UniSA Learning Connection, 2008). A world trend in staying abreast of the latest developments in the field of ICT has led to an increased demand for electronic assessment in education circles. The critical need and responsibility for higher education to stay on par with the latest techniques regarding assessment subsequently led the University of Johannesburg (UJ) to implement electronic assessment in some departments in 2004. Several challenges led to this exploration into the use of one e-assessment tool within the University.

Key words: *Assessment, Electronic Assessment, E-assessment, Higher Education*

Introduction and background to the inquiry

With technological advances and the increased expectations on individuals, the use of technology in the teaching and learning process has become essential. Computer technology is continuously advancing and has impacted on teaching and learning (Shelly, Cashman, Waggoner & Waggoner, 2000). Tertiary Institutions worldwide have accordingly realised the value of electronic assessment (e-assessment). Electronic assessment can be defined as a method of using information technology for any assessment-related activities (Graff, 2008). Personnel shortages and the need for enhanced throughput both add to the rationale to move towards this new method of assessment. Traditional paper-and-pencil testing is now seen as obsolete and outdated compared with the latest techniques in teaching, learning, and assessment (Parshall, Spray, Kalohn & Davey, 2002). Being computer literate has become a necessity not by choice, but due to the requirements and standards set by industry where there is a need for trained and digitally literate personnel. This has evolved into an increased demand for electronic assessment for training and education (Pangali, 2003).

According to Benson (2003) the principles of assessment do not change in an electronic environment. E-assessment is underpinned by the same principles of validity, flexibility and fairness, and also uses the same strategies as traditional assessment methods (Booth, Clayton, Hyde, Hartcher & Hungar, 2002). According to Hricko and Howell (2006) one of the important considerations for effective e-assessment is to ensure that the tool incorporates these elements, fits the mode of delivery, and legitimately measures the desired outcome. Therefore, both the lecturer and student can obtain maximum benefit from the electronic assessment process, if there is a good relationship between learning targets, instruction and assessment.

Electronic assessment is utilised at many universities in South Africa. The demand has also been fuelled by a growing number of students and a decrease in allotted class times. Universities using e-assessment products like *Skills Assessment Manager* (SAM - a Web-based application measuring proficiency in *MSOffice* applications, including *MSWord*, *MSExcel*, *MSPowerPoint* and *MSAccess*. *SAM* can also measure user skills in *Windows 2000*, *Windows XP* and Internet usage) and *Electric Paper* (an automated testing system for the International Computer Driving Licence, with immediate and accurate evaluation, in a self-contained system consisting of software simulations, requiring no other software applications to run it) include Wits University, Nelson Mandela Metropolitan University, and UNISA.

The sub-department End User Computing (EUC) at the University of Johannesburg implemented CompAssess as an e-assessment tool in 2004. CompAssess is a tool that allows the student to work in a simulated environment for *MSWord*, *MSExcel*, *MSPowerPoint,* and *MSAccess*. CompAssess facilitates the creation of customised assessments for any of the software applications mentioned above. It allows for the selection and customisation of generic built-in tests and the specification of assessment parameters such as time, question weighting, and passing grades. Input of student details and the exporting and printing of reports are included (Masterskill CompAssess, 2006). The software logs and records the last choice made as the final answer. When submitted, the assessment is marked and the system informs the lecturer of the submission.

The module EUC is a practical computer module, which assesses the competency of students in a software application environment. EUC is a practical module and it is required from students to display their ability in using a computer. Students demonstrate their capacity in the execution of specific tasks on a computer. Instructions are executed in a software driven environment which guides and tests students in a *Microsoft Office* environment. According to Haywood (2000) the concept of authentic assessment relates to the idea of competency and performance-based learning. Stiggins (1987) states that performance assessments call upon the examinee to demonstrate specific skills and competencies, that is, to apply the skills and knowledge they have mastered. The module EUC gives students the opportunity to apply skills in a realistic simulated electronic assessment environment.

Methods

Both qualitative and quantitative methods were used in this inquiry, with the qualitative data collected first. According to Creswell (2003), a sequential exploratory design is conducted in two phases, the first phase involving qualitative methods, where after a quantitative phase is developed to test or generalise initial results. According to Creswell (2003) if qualitative data are collected first, the intent is to explore the topic with participants at sites. Creswell also states that the second phase explores understanding of data collected from a large number of people. In this inquiry the qualitative data were collected and then analysed to identify variables to be explored in the second, quantitative phase of the inquiry. According to Bergman (2008) in exploratory design the participants in the first phase of data collection are typically not the same as those in the second phase. The purpose of the second phase, in this inquiry the quantitative phase, is to generalise the results to the population. Both qualitative and quantitative phases carried the same weight and emphasis.

Qualitative research aims to understand the phenomena in context-specific settings, such as a "real world setting where the researcher does not attempt to manipulate the phenomenon of interest" (Patton, 2001). Both permanent and part-time EUC lecturers from the sub-department EUC and one technical support staff member from the University were included in two focus group interviews. This phase was conducted with the researcher aiming to explore the engagement of lecturers using an e-assessment tool. This took place in the lecturers' natural setting, as qualitative researchers study phenomena in natural settings to make sense of phenomena (Creswell, 1998; Henning, Van Rensburg & Smit, 2004). Two focus group interviews allowed participants to share their feelings, opinions, experiences, behaviours and meanings (Denzin, 2001; Rubin & Rubin, 1995). This was also a useful way of gathering large volumes of data quickly (Marshall & Rossman 1999). The first focus group consisted of three lecturers and one technical support specialist who was directly involved with the administration of the e-assessment tool. The second focus group consisted of lecturers involved with the actual assessment of learners in the classroom environment. From this qualitative phase, variables were identified for further quantitative analysis.

Quantitative methods are based on numerical measurements of specific aspects of phenomena (Thomas, 2003). A quantitative phase was used in this inquiry utilising a structured questionnaire compiled using variables that were identified during the qualitative phase of the inquiry. In this phase of the inquiry the data collection was conducted sequentially within an exploratory design from 330 EUC students to explore the correlation between variables identified in the first phase of qualitative data collection. The questionnaire used for this purpose was compiled in collaboration with the statistical department at the University. The intent was to pose questions that could secure responses from students which could corroborate or negate lecturers' perceptions of e-assessment. In this process, students completed the questionnaire, structured in three categories namely background information, previous assessment experience, and e-assessment experience during EUC studies at the University.

The first section of the questionnaire was the demographic category to secure background information from the students. This was to check the representativeness of the sample and to enable the researcher to make statistical comparisons (Gray, Williamson, Karp & Dalphin, 2007). An example of the type of background information in this section was the "*home language of students"*. Through analysis of this background information the researcher can, for example, determine if the students are comfortable being assessed in English. The second part of the questionnaire included contingency questions where the researcher determined students' previous experiences with an e-assessment tool. Contingency questions are used to identify a subgroup for further questioning (Gray et al., 2007). An example of this section is when students need to answer "yes" or "no" to a question. If they answer in the affirmative they will complete a particular subset questions (see appendix H) and ignore others.

The last section of the questionnaire consisted of filtered and Likert scale type questions. A filtered question is posed to determine which respondents have sufficient information on an issue that can provide meaningful opinions (Gray et al., 2007). An example of this type of question is when the researcher poses a question, for example, "were students introduced by way of a tutorial before doing the electronic assessment?" During this process, students only need to reply "yes" or "no". The last number of questions posed will be based on a Likert scale format, where the students select the category which they believe best reflects their perspective on the statement. Likert scales are presented with a series of statements where the respondents indicate their levels of agreement or disagreement (Bryman & Cramer, 2003). One question, for example, was posed to determine whether the terminology used in the electronic assessment environment was similar to the terminology used in class.

Data Analysis

Henning et al. (2004) state that the true test of a competent qualitative researcher comes in the analysis of the data. This process requires analytical craftsmanship and the ability to capture an understanding of the data writing. Merriam (1998), states that data analysis is the process of making sense out of the data, by evaluating, consolidating, reducing and interpreting what the participants said. It relates to the researcher securing information from diverse sources as identified by way of accepted research principles. Interviews were recorded and copies stored as audio files on CD-ROMS for archiving purposes. Transcriptions were perused several times before beginning the process of analysis (Creswell, 2002). According to Henning et al. (2004) one needs to first read through the entire text in order to get a global impression of the content. After this, data were categorised as per figure 1 into themes or categories.

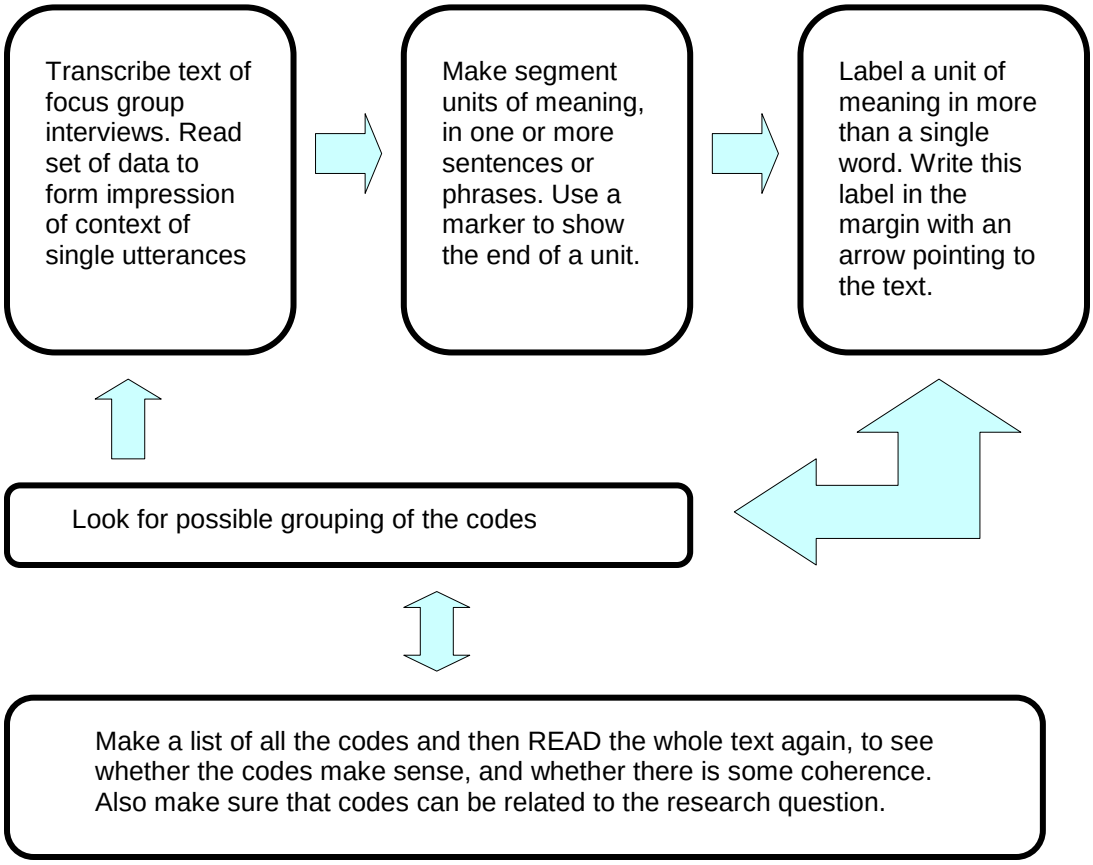
Transcribe text of focus group interviews. Read set of data to form impression of context of single utterances
Make segment units of meaning, in one or more sentences or phrases. Use a marker to show the end of a unit.
Label a unit of meaning in more than a single word. Write this label in the margin with an arrow pointing to the text.
Look for possible grouping of the codes
Make a list of all the codes and then READ the whole text again, to see whether the codes make sense, and whether there is some coherence. Also make sure that codes can be related to the research question.

Figure 1 Coding from texts (Henning et al., 2004)

Quantitative analysis of the survey data was done bythe statistical services department, *Statcon*. *Statcon* used the *Statistical Package for Social Sciences (SPSS)* programme to process the questionnaire. We initially requested a factor analysis in order to group together related variables, as one combined variable or factor. Mentioned variables related to specific group were then compiled in such a way as to form a combined variable. Bryman and Cramer (2003) state that a "factor analysis is a set of techniques for determining the extent to which variables that are related can be grouped together so that they can be treated as one combined variable or factor, rather than as a series of separate variables." Sapsford and Jupp (2006) support this by stating that a factor analysis is a technique based on assessing the correlation between variables. Simplification of data is secured by way of identification of groups of variables which are connected and may then be regarded as part of a single factor.

In this inquiry an exploratory, factor analysis was used to indicate how many groups or dimensions were found within this questionnaire. According to Bryman et al. (2003) exploratory factor analysis is a more formal method to determine how many groups there are in quantitative data analysis. In order to determine the extent to which the variables are correlated to each other we used Cronbach's alpha (α). According to Sapsford et al. (2006), Cronbach's alpha (α) is a statistic calculated to assess the extent to which items in a scale are correlated with each other. These items should be highly correlated if they all measure the same thing. Cronbach's alpha (α) will generally increase when the correlation between the items increase. This means that the coefficient is also called the internal consistency reliability of the test (Schmitt, 1996). Cronbach's alpha (α) assessed the extent to which the items on a subscale were measuring the same thing. It allowed us to discard some of the items to produce a more unidimensional subscale with less internal error variance (Sapsford et al., 2006). Cronbach's alpha (α) was, therefore, used to test the consistency and reliability of the unidimensional subscale of the dimensions as found in the questionnaires provided to students as part of the quantitative analysis process.

All quantitative data were analysed with the SPSS software, displayed as tables, charts and numerical statistical measurements which were then interpreted further.

Findings

The qualitative data referred to in this inquiry have been thoroughly interpreted and categorised based on their relevancy (Sapsford & Jupp, 2006). The following three categories related to lecturers' perceptions of student engagement with the electronic assessment tool emerged from the qualitative analysis of the two focus group interviews: *issues related to language*, *issues related to the electronic assessment system*, and *issues related to teaching and learning with the electronic assessment tool*.

In *issues related to language*, lecturers were concerned with the language and terminology used in questions and in the phrasing of questions. Students, however, experienced difficulty understanding scope, and context of the questions. Students' concerns are in contrast to those of the lecturers. Lecturers felt that the language used in the electronic assessment tool is too difficult for the students to understand. In contrast, 79.40% of students are of the opinion that the language used in electronic assessments was understandable. 56.90% of the students agreed that the terminology used in class and in the electronic assessment environment are the same.

The fact that students come from countries all over the African continent also has an impact on their engagement with the tool. These students attend school in their home tongue and at times do not receive training in English. This was also corroborated when lecturers mentioned that they had to explain certain concepts to students but found that after several attempts they failed to convey information in such a way that students understood.

In *issues related to the e-assessment system*, both lecturers and students experienced major challenges in their engagement with the electronic assessment system. The system challenges lecturers by not allowing them to import assessments from multiple campuses. As the EUC offices are located on different campuses at the University, it is at times necessary to import assessments from one campus to another.

The students experienced the time allowed for assessments to be insufficient. The slow pace of the system negatively affected all persons utilizing system. The poor performance rate of system (including delayed responses and "hang time" impacted negatively on students in that they were having difficulty completing exams in the allotted time period.

Moreover, lecturers had difficulty in viewing results. The system also displays results achieved by students immediately upon submission of the tests and this functionality resulted in instances where students could benefit unfairly. It was found that the functionality to view results can affect fellow student's performance. To address this possible unfair practice certain administrative instructions were activated on the system to deactivate viewing functionality. Both lecturers and students were concerned with the security of the system in terms of access to results during assessment sessions. From the quantitative data analysis, 68.1% of the students were found to be concerned with privacy issues.

Issues related to teaching and learning with the e-assessment tool also emerged from the qualitative data. Electronic assessment is a new concept at the University and lecturers are of the opinion that it is important for students to be exposed to a standard way of assessment. *Lecturers have specific teaching styles and the questioning used in the electronic assessment may not be aligned with the way they teach*. Lecturers need training on the use of the electronic assessment tool and it is also important for them to address challenges during assessment in class.

Lecturers believe that if they have knowledge of the assessment system, it will put them in full control of their teaching and assessment. This capacity will also allow lecturers the ability to solve their own challenges on the system and appear more professional from a students' perspective. Some lecturers even requested accredited training on the electronic assessment system.

On the other hand, the following three dimensions related to students' engagement with an electronic assessment tool emerged from the quantitative analysis of the questionnaire: *language and usability of electronic assessment, fairness of electronic assessment, and preparation for electronic assessment.* The dimension which relates to the *preparation for electronic assessment* incorporates three variables which were grouped together as one combined variable: *tutorial assists students to understand the electronic assessment environment; the tutorial prepared students adequately for the electronic assessment session;* and *the assessment provides a true reflection of ability at the time.* Students are of the opinion that the preparation for the electronic assessment was sufficient. This is in direct contrast to the lecturers' comments on the qualitative analysis as contained in interviews. The lecturers stated that, if students are allowed access to system before hand, they will be more successful with assessment results.

During the quantitative analysis, a four point Likert scale was used to determine the level of agreement from the students with the variables during the qualitative data process. The following four point Likert scale was used in this inquiry: 1: Strongly Disagree (SD); 2: Disagree (D), 3: Agree (A) and 4: Strongly Agree (SA). The average value for each dimension within the quantitative component was calculated. These values were used to group variables based on their value. Values relate to a scale of one to four. To calculate the average of each dimension from the quantitative analysis the following formula was used:

$$\text{Average per Dimension} = \frac{\text{Mean}}{\text{Number of variables}}$$

Figure 2 below, shows the first dimension of the quantitative analysis *language and usability of electronic assessment.* Five items were grouped together as one combined variable, the sum of the means being 14.20. Therefore the average for this dimension was 2.84. This relates to the number which is closer to 3 on the Likert scale. Therefore it can be said that the majority of the students agree that the *language and usability of the electronic assessment tool* is understandable.

Item Statistics

	Mean	Std. Deviation	N
The language used in the Electronic Assessment is understandable	3.00	.797	313
The terminology used in Electronic Assessment is similar to the terminology used in class	2.57	.958	313
The use of electronic assessment is more beneficial than the traditional pen and	3.14	1.029	313

paper assessment			
The Electronic Assessment and the work done in class are aligned	2.70	.779	313
The Electronic Assessment environment is user friendly	2.78	.857	313

Scale Statistics			
Mean	**Variance**	**Std. Deviation**	**N of Items**
14.20	8.668	2.944	5

Figure 2 Items and scale statistics of language and usability of electronic assessment

In figure 3 below, the second dimension of the quantitative analysis *fairness of the electronic assessment tool*, two items were grouped together as one combined variable. The mean of the variables are 4.72 and the number of items was two. Therefore, the average for this dimension is 2.36. This relates to the number which is closer to 2 on the Likert scale. Therefore it can be said that the majority of the students disagree that the electronic assessment tool is fair.

Item Statistics			
	Mean	**Std. Deviation**	**N**
The time allocated for questions was sufficient	2.04	.960	315
There were sufficient guidelines on how to complete the assessment	2.68	.857	315

Scale Statistics			
Mean	**Variance**	**Std. Deviation**	**N of Items**
4.72	2.070	1.439	2

Figure 3 Fairness of electronic assessment

In figure 4 below, the last dimension of the quantitative analysis *preparation for electronic assessment*, three items were grouped together as one combined variable. The mean of the variables is 8.09 and the numbers of items was three. Therefore the average for this dimension is 2.69. This relates to the number which is closer to 3 on the Likert scale. Therefore it can be said that the majority of students also feel that the electronic assessment results were a true reflection of their ability at the time.

Item Statistics			
	Mean	**Std. Deviation**	**N**
The tutorial helped me understand the Electronic Assessment environment.	2.77	.929	318
The tutorial prepared me adequately for the Electronic Assessment session.	2.64	.869	318
The assessment result gave a true reflection of my abilities at the time.	2.68	.898	318

Scale Statistics			
Mean	**Variance**	**Std. Deviation**	**N of Items**
8.09	4.531	2.129	3

Figure 4 Preparation for electronic assessment

The lecturers' concerns from the qualitative analysis were, therefore, largely confirmed by the quantitative analysis process. These concerns will be dealt with in more detail below.

Conclusion

With the implementation of the CompAssess system at the University, it became apparent that several "gaps" exist within the system. It had become critical to identify and implement a system that would allow more time to be spent with an increasing number of students at the University. Students expect more time to be allocated to them for class presentations, facilitation and assessment. The use of the electronic assessment tool CompAssess has significantly contributed to alleviating some of this pressure. As with all systems, CompAssess needs to be continuously evaluated and corrective steps taken with system developers. Lecturers should receive comprehensive training on the system. These trained accredited staff with specific administrative rights should be allowed to access all administrative functions on system and be able to address and correct challenges as they arise or are reported by students undertaking assessments. Lecturers also need to be able to have access to the system in order to prepare and run initial assessment sessions to test system and actually undertake actual assessment exam to verify system credibility and performance.

Another requirement of lecturers is the necessity to secure authorised access to the system in order to take control of setup and preparation settings on system. This access is also required to resolve system failures or system crashes whenever they occur. It is important that lecturers have greater knowledge of the system so that they can address challenges and explain them to students who require explanations of system issues.

Students should be allowed access to the system before assessments take place. There should be practical sessions where students are exposed to the system. Students should be able to take an actual preparation assessment testing session to prepare for assessment to be completed for marks.

The security of an electronic assessment system is a priority. Unauthorised access to an electronic assessment system or observation of the electronic assessment system content should not be allowed

This inquiry is relevant to developments in technology and processes related to the Information Technology (IT) landscape. Educational Technologies have developed at such a fast pace that it has become critical to keep up to date to optimally benefit from

these advances. This paper provides a contextualised overview of challenges facing both lecturers and students at a University engaging with a single electronic assessment tool.

List of sources

Benson, A.D. (2003). Assessing participant learning in online environment. New Directions for adult and continuing education, no 100, Winter 2003. Wiley Periodicals, Inc.

Bergman, M.M. (2008). Advances in mixed methods research. London: Sage Publications.

Booth, R., Clayton, B., Hyde, P., Hartcher, R. & Hungar, S. (2002). Creative, quality online assessment. Progress report to FLAG, July 2002. In Australian Flexible Learning Framework. Assessment and Online Teaching. Retrieved from http://www.flexiblelearning.net.au/guides/assessment.pdf. (Verified 29 January 2008).

Bryman, A. & Cramer, D. (2003). Advanced quantitative data analysis. Philadelphia: McGraw Hill Open University Press.

Cashman, T.J., Waggoner, G.A., Waggoner, W.C. & Shelly, G.B. (2000). Discovering computers. Boyd & Fraser Pub Co.

Creswell, J.W. (1998). Qualitative inquiry and research design: Choosing among five traditions. Thousand Oaks: Sage Publications.

Creswell, J.W. (2002). Educational research: planning, conducting and evaluating quantitative and qualitative research. New Jersey: Merrill Prentice Hall.

Creswell, J.W. (2003). Research design: Qualitative, quantitative and mixed methods approaches. (2nd ed.). Thousand Oaks: Sage Publications.

Denzin, N.K. (2001). The reflexive interview and performative social life: Qualitative research. In Berg, B.L. 2004. Qualitative research methods for the social sciences. Boston: Pearson Education Inc.

Graff, M. (2004). Cognitive style and attitudes towards using online learning and assessment methods. Electronic Journal of e-Learning. Vol. 2, Issue. 2, December 2004. ISSN: 1479-4403.

Gray, P.S., Williamson, J.B., Karp, D.A. & Dalphin, J.R. (2007). The research imagination: An introduction to qualitative and quantitative methods. New York: Cambridge University Press.

Henning, E. van Rensburg, W. & Smit, B. (2004). Finding your way in qualitative research. Pretoria: Van Schaik.

Heywood, J. (2000). Assessment in higher education. London: Jessica Kingsley Publishers.

Hricko, M. & Howel, S.L. (2006). Online assessment and measurement: Foundations and challenges. Hershey, PA: Information Science Publishing..

Marshall, C. & Rossman, G.B. (1999). Designing qualitative research. (3rd ed.). London: Sage Publications.

Masterskill. (2006). CompAssess. Retrieved from http://www.masterskill.com/products/ASSESSMENTTOOL/compassess/tabid/824/Default.aspx. (Verified 20 April 2007).

Merriam, S.B. (1998). Qualitative research and case study applications in education. San Francisco: Jossey-Bass.

Pangali, C. (2003). Using technology to measure learning. Chief Learning Officer; November 2003, Vol. 2 Issue 7, p36, 5p, 3c. Article.

Parshall, C.G., Spray, J.A., Kalohn, J.C. & Davey, T. (2002). Practical considerations in computer-based testing. Springer.

Patton, M.Q. (2001). Qualitative evaluation and research methods. (3rd ed.). Thousand Oaks, CA: Sage Publications, Inc. In Golafshani, N. (2003). Understanding reliability and validity in qualitative research. The qualitative report. Vol. 8, No. 4, December 2003, p. 597 – 607. Journal.

Rubin, J. & Rubin, S. (1995). Qualitative interviewing: The art of hearing data. London: Sage publications Thousand Oaks.

Sapsford, R. & Jupp, V. (2006). Data collection and analysis. (2nd ed.). London: Sage Publications.

Schmitt, N. (1996). Uses and abuses of coefficient alpha. Psychological assessment. Vol. 8. No. 4, p. 350 - 353.

Stiggins, R.J. (1987). The design and development of performance assessment. Educational Measurement: Issues and Practice, 6, p.33-42.

Thomas, R.M. (2003). Blending qualitative & quantitative research methods in theses and dissertations. Thousand Oaks, California: Corwin Press.

UniSA Learning Connection. (2008) What is assessment? Retrieved from http://www.unisanet.unisa.edu.au/LearningConnection/staff/assessment. (Verified 18 January 2008).