Data Gathering Dilemmas: Postgraduate Experiences

Kay Fielden
School of Computing and Information Technology
Unitec New Zealand
kfielden@unitec.ac.nz

ABSTRACT

In order to complete the Masters in Computing (MComp) degree at Unitec students are required to complete either a thesis (120 credit points) or a dissertation (60 credit points). For a thesis or dissertation students are required to conduct a research project in which they gather primary data in an appropriate manner. This paper addresses dilemmas faced by MComp students in gathering primary data for their research project. Challenges that have emerged for MComp students are then presented. Following this ways of turning these challenges into learning tools to enhance postgraduate education and postgraduate supervision are discussed. Finally the ways in which these challenges add to the general pool of supervisory practices is addressed.

Keywords

Data gathering, Postgraduate education, Research

1. INTRODUCTION

The issues addressed in this paper are the dilemmas faced by postgraduate students in the Masters in Computing (MComp) program during the data gathering process required for the research component of the degree (thesis or dissertation). The main research question posed for this study explored the dilemmas faced by postgraduate students as they gather their research data. Findings from this add to the learning experience of future postgraduate students as well as adding to the general knowledge pool of supervisory practices.

A useful theoretical framework based on Habermas (1996) and extended by Fielden (2004) situates data gathering in the world of business, often expressed in academic terms and administered by delegated governmental authority from academic administration (Figure 1). Information for this paper was amassed by: surveying postgraduate students on their experiences in gathering their research data; observing postgraduate seminars; postgraduate supervision; and sharing best practice with other postgraduate supervisors. These results have been compared to literature results and these findings are presented in this paper.

2. THE RESEARCH COMPONENT OF THE MASTERS IN COMPUTING

The research component of the MComp degree program takes the form of either a thesis (120 credit points) or a dissertation (60 credit points). In order to complete a research project in the MComp students are required to gather primary data in an appropriate manner.

2.1 Learning Outcomes for Thesis or Dissertation

Learning Outcomes (LO) include the ability to conduct applied research in the field of computing. In order to achieve this postgraduate students are required to:

- Carry out a supervised research project (LO1);
- Demonstrate appropriate time management and research planning (LO2);
- Carry out a review of the relevant literature (LO3);
- Collect data using appropriate methods (LO4);
- Analyse and interpret data (LO5); and
- Draw conclusions from the data and make recommendations (LO6).

A necessary requirement for any research project is that primary data must be gathered by the postgraduate student. This necessary requirement
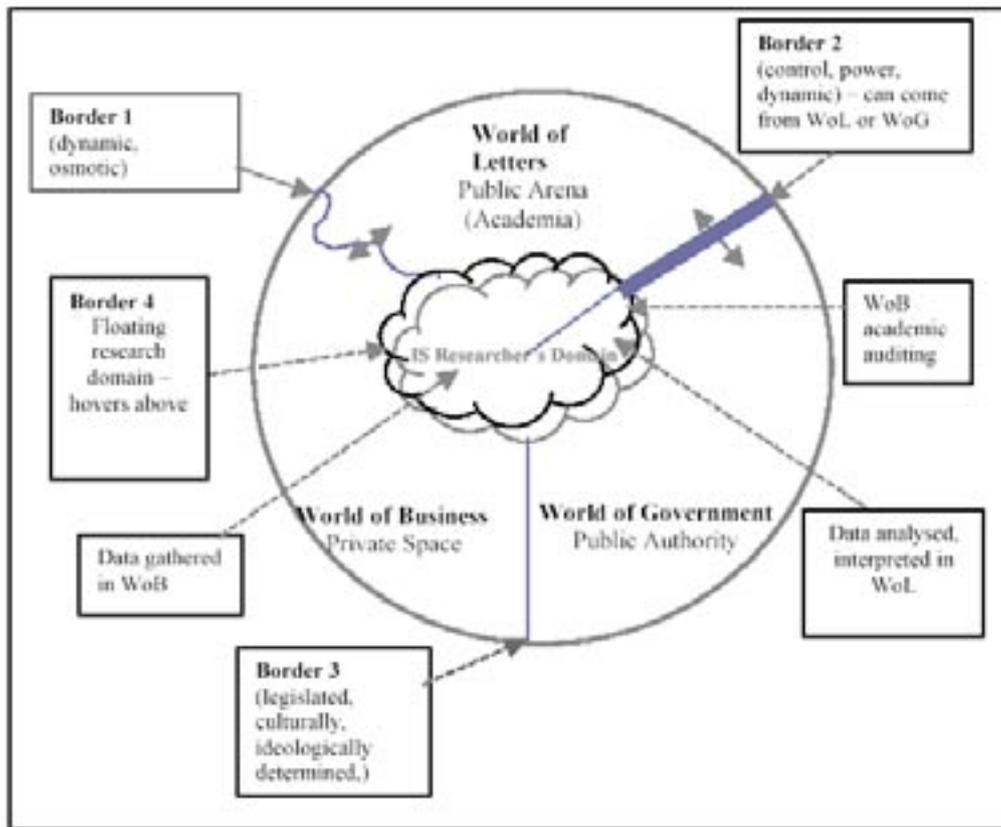
Border 1
(dynamic,
osmotic)

Border 2
(control, power,
dynamic) – can come
from WoL or WoG

World of
Letters
Public Arena
(Academia)

Border 4
Floating
research
domain –
hovers above

IS Researcher's Domain

WoB
academic
auditing

World of Business
Private Space

World of Government
Public Authority

Data gathered
in WoB

Data analysed,
interpreted in
WoL

Border 3
(legislated,
culturally,
ideologically
determined,)

Figure 1 (Fielden, 2004) Research World (based on Habermas's (1996) Spheres of Activity)

is listed explicitly in LO4 and it is assumed that data gathering has taken place in LO5 and LO6. Implicit also in the learning outcomes is the fact the postgraduate research is a governed activity (LO1). Hidden beneath these learning outcomes is the plethora of challenges faced by students as they gather their research data.

3. DATA GATHERING METHODS

Appropriate data gathering methods employed by MComp students include: surveys and questionnaires, interviews, focus groups, observations, experimentation, and data mining (both real and virtual).

Each one of these methods has its own set of challenges and dilemmas associated with it.

4. LITERATURE REVIEW

Much has been written about how postgraduate students should be supervised (Australia, 2003), roles of supervisors (Bartlett & Mercer, 2001; Craswell, 1996; McMichael, 1993), roles of postgraduate students (Heath, 2002), tracking postgraduate supervision (Edwards *et al.*, 1995) institutional rules (Melrose, 1999) and quality of supervision (Russell, 1994; Soliman, 1999).

The literature on data gathering dilemmas is rich with ethical considerations (Australia, 2003; Bouffard & Little, 2004; Knobel & Lankshear, 2004; Milne, 1996) that face postgraduate students as they embark on their first research project. However, there appears to be a paucity of literature on dilemmas faced by students as they collect primary research data (Bouffard & Little, 2004; Knobel & Lankshear, 2004). Bouffard and Little (2004) embark on the usual list of 'ethically appropriate data collection' (surveys, questionnaires, interviews, focus groups, observations, test and data reviews). Bouffard *et al* (2004) also list the advantages and disadvantages of each data collection method in the standard manner.

Knobel and Lankshear (2004) consider the ethical dimensions of data collection from both a positive (trustworthiness, reciprocity) and a negative point of view (intrusion, consent, disclosure, 'lurking' in public spaces both real and virtual, privacy and dignity). Knobel *et al* also discuss trustworthiness of online data and new dimensions introduced by collecting data from public internet sites. Knobel *et al* also consider risks both calculated and unwanted, and implications arising from landing in difficult situations.

Nowhere can I find a discussion based on the dilemmas faced by postgraduate students as they gather research data. Supervision experience has led me to believe that most likely there will be challenges emerging during the data collection phase. These have been related to logistics, amount of data – too little or too much, timing of data collection and personality issues that arise between researcher and participant. Data collection issues have also arisen because the data collection method was relatively new – such as automated data collection from newsgroup sites or web-crawling.

5. A THEORETICAL FRAMEWORK

In figure 1 it can be seen that the researcher's domain floats above and hovers between academia, business and government. Data is normally gathered in the world of business after permissions have been obtained from the delegated authority of academic postgraduate administration. Experience gained both as a supervisor and a researcher suggest that many of the data gathering dilemmas that emerge arise on the borders between the disparate worlds of academia and business. Each world has its own use of language, culture, working environment and legislated procedures. Unless the postgraduate student is collecting data within her/his own world of work little is know about the chosen research domain.

6. RESEARCH QUESTIONS FOR THIS PAPER

In order to learn from the dilemmas faced by postgraduate students as they gather their data the following research questions were posed:

• What challenges have presented themselves to students as they gather their research data?

• How can these challenges add to the learning experience of future postgraduate students?

• How can postgraduate educators improve the way in which they supervise?

• How can these challenges add to the general knowledge pool of supervisory practices?

7. RESEARCH METHOD

7.1 Data Collection for this Paper

Multiple data collection methods were utilised for this exercise of enquiry into data gathering dilemmas encountered by postgraduate students when they conduct their own research. As postgraduate educator, supervisor and examiner for a number of years it has become evident that patterns have emerged in the dilemmas facing students as they gather data. This knowledge has been passed on to both postgraduate students and supervisors via supervision meetings, postgraduate seminars and informal discussions. This paper details more formal data collection so that a record can be passed to the postgraduate community.

7.1.1 Data Collection Method 1 - Research Journal

Just as all postgraduate students are encouraged to maintain a research journal, I have kept a research journal of any and all issues pertaining to research for a number of years. Mining this longitudinal record of observations, reflections and insights was a valuable source of data to gain a better understanding of my own point of view on dilemmas faced gathering data. In gathering data for one of my own recent projects I realised – only after the interview had been conducted – that the way in which the interview had been arranged had placed the participant in a position of power for the interview. The literature suggests that it is the researcher that has the power in conducting research interviews. In this case it was important to the participant that he/she retains ascendancy in the manner in which data was gathered. Postgraduate students have also reported such incidents as they have gathered research data.

7.1.2 Data Collection Method 2 – Shared Practice

Within the MComp program postgraduate research seminars have been conducted regularly for the past few years. These seminars provide postgraduate students and supervisors with the opportunity to share practice and experience in collecting data. From a recent seminar conducted exclusively on data gathering the many issues were aired by postgraduate students. These

are discussed below in the appropriate analysis section (6.2).

7.1.2 Data Collection Method 3 – Survey

10 students engaged in postgraduate research in the MComp program answered a brief survey that asked if any problems had arisen for them in the data collection process. The data collection process was divided in to before data was collected (ethics approval), data collection methods used, and data analysis. Issues arising for students before data collection included: sample size (either too many or too few), disability and/or health status of participants, and age of participants.

7.2 Analysis

The findings have been amalgamated from the three data gathering methods so that the pooled experiences of the postgraduate community can be analysed. Each data gathering method utilised by postgraduate students has been summarised. The richest data set emerging is from those postgraduate students that used interviews to gather their data. This is not surprising as those students who interview are more likely to have better developed people skills, enjoying the depth of interaction provided by this means of data collection.

7.2.1 Interviews

Dilemmas encountered in conducting interviews crossing the academic border into the world of business included the following sets of dilemmas.

7.2.1.2 Making contact with participant issues

A number of students expressed an initial difficulty contacting people within the required categories.

7.2.1.3 Participant related issues

Some participants seemed to regard taking part in the research a chore. One student commented on the participant's mood on the day of the interview. Sometimes it was not known whether the person would be in a suitable emotional state to give sensible answers. There was also a dilemma posed by the proximity of other activities within the organisation to interview times. If the interview of a participant coincided with key meetings, visits or deadlines within an organisation this had undue effects on the quality of the data obtained.

7.2.1.4 Organisational related issues

In some cases institutional support for research was given but it was then discovered that this did not equate to individual participant commitment. Participant resentment was also encountered. If the agreed time for the interview was decided by someone other than the participant, for instance a manager, then it was evident that there was some ill feeling present at the interview. If the participant was not well – and still insisted that the interview go ahead - this also affected the quality of the data collected. If the participant was not in agreement with the research being conducted and agreed to the interview for what ever reason was presented within the organisation this also affected the quality of the data collected.

7.2.1.5 Time related issues

Many students discovered that agreeing on a suitable interview time with participants was an issue. Students are required to conduct their research in a timely fashion according to an agreed research plan. Organisations on the other hand have their own agendas to meet. One student also discovered that the planned order of interviews may not be the way people in an organisation present themselves. This in turn made some questions redundant and required other questions instead. This also impacts on the level of knowledge the researcher has of the organisation being researched.

7.2.1.6 Location related issues

Geographically dispersed participants raised their own set of issues. These included: travelling to a distant participant only to find that the participant was not present at work that day, the participant was ill, the were more important issues that the organisation was required to deal with that day or even that the participant had forgotten about the research interview.

7.2.1.7 Research rigour issues

Discrepancies between information sheet and data gathering method also posed problems for student researchers. In one case the interview was taped however the statement regarding taping was omitted from information sheet. This was a major

concern for one student who was researching within her own institution. She had gained full ethics approval within the organisation, only to find out when she arrived to conduct an interview that the participant was also a member of the organisational ethics committee that approved this particular research project. None of the other interviewees had commented on the omission in the information sheet. This participant however was wearing multiple hats – employee, participant and member of the ethics committee. The postgraduate student was also wearing multiple hats – postgraduate student as well as employee. The interview commenced with the student, ethics committee member reprimand – not the researcher/participant query.

7.2.1.8 Researcher related issues

Postgraduate students who talked too much in interviews realised that their behaviour affected the data collected. Talking too much, prompting the participant, putting words in their mouths – and changing interview technique all posed dilemmas for student researchers. One postgraduate student in particular was adamant that she/he would conduct interviews in a very different manner - questions would be asked much more carefully, topics would be more focussed and the primary task was to listen.

This same postgraduate student discovered that it was important to seek clarification on answers given. This was a useful ploy for bringing participants back on track and it as also useful to clear up any misconceptions that may have arisen between participant and interviewer.

Another observation on interviewing made was that it was very easy to make assumptions about what was said. The researcher could assume that the participant meant one thing that on reflection and after listening to the tape again turned out to be quite a different issue.

One student commented that he/she made the assumption that everyone within one organisation would have a common view on what was required. This student was most surprised to discover that there were multiple points of view within the organisation about the interview questions being asked.

Another student commented that a pilot study would have been a good idea to trial interview questions.

One student who gathered data with interviews commented that we are all wise after the event. The MComp is a learning experience and all students stated that they would adopt different and more informed interview strategies next time.

7.2.1.9 Data analysis related issues for interviews

A common cry about transcribing interview data was that 'transcription is a nightmare. It is slow and tedious - and very useful to do'. On the other hand one student stated that she would have lost a lot if she had employed someone else to transcribe the data.

An observation made a number of times was 'all interviews are not equal". The same set of questions does not elicit the same set of responses from different participants. Difficulties therefore arise - in comparing results. It is essential in evaluating interview data that the context must always be given.

Common reflections after data had been gathered was that doing the research again would mean that it would be done differently.

There were also some confidentiality issues after data collection when one participant wanted to know who had access to the transcripts.

7.2.3. Surveys

Surveys are also a common data collection method in the MComp. There are a different set of dilemmas that arise in conducting surveys.

7.2.3.1 Making contact with participant issues

A major problem in gathering data by survey is gaining participation regardless of website, email or posted submission. One student contacted survey participants before the survey posted (but the student believed that this should have been planned better). Students also raised the issue of gaining emotional involvement from participants before survey questions are asked.

7.2.3.1 Cost

One student commented on the cost of distributing the survey and getting the survey back again (and this must be funded by the student).

43

Table 1: Data Collection Dilemmas in Domains/Across Borders

Research Method	Academia	Border 1 Academia/ Business	Business	Knowledge gained from Experience
Interviews	1	8	3	9
Surveys	2	11	0	6
Focus Groups	1	0	2	2
Automated	5	0	0	5
Website	0	0	0	1
Documentation	0	0	1	0

7.2.3.2 Sample issues

Sample issues that arose were sample size too small, obtaining a balance between survey size and survey usefulness; and misinterpretation of the questions asked in the survey.

7.2.3.3 Research rigour issues

Postgraduate students also discovered the importance of research rigour. After the survey had been conducted the realisation that the survey should have been trialled initially was made and that such a trial should extend beyond friends and family. One student stated categorically that he must 'ENTER THE DATA AS IT COMES IN' and not set returned surveys aside until later.

7.2.3.4 Survey returns

Dilemmas that arose with survey returns were: surveys returned unanswered; response rate depending on how extensively the organisation was being surveyed by other researchers; survey results informing a second data collection method of interviewing a small number of participants; the usefulness of a pilot study; and that when a statistically significant quantitative study is being conducted the statistics tests to be used need to be considered when the survey form is designed.

7.2.3.5 Organisational issues

Issues that arose were: the way organisations structure themselves may not be aligned with the way in which the survey is structured; questions must be written in an unambiguous way and the questions need to relate to the context of the organisation; and obtaining buy-in from tutor participants to administer survey was a problem. There was resentment about 'wasted class time', and inevitable resistance to administering the survey in class.

7.2.3.6 Researcher issues

One student commented: "I would be a better organised human being if I did this again'.
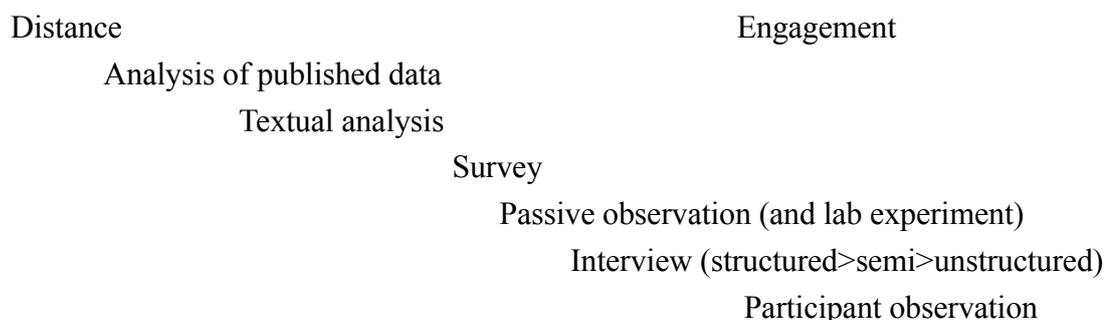
7.2.4 Focus Groups

Fewer students used focus groups as a data gathering method. Some of the issues faced in gathering data in this way included:

7.2.4.1 Participant issues

There was a problem for one student in finding participants to take part. This student endeavoured to organise a focus group of senior executives within one organisation only to find that it was too difficult to arrange a common meeting time. Another student had no response to 30 letters posted out for people to participate. Only a personal appeal and personal delivery of information worked. She/he finally managed to arrange for 7 people to take part in the focus group. One student commented that it was difficult to facilitate a focus group and to take notes at the same time. She/he overcame this problem by employing an observer to take notes during the focus group. The quality of recording equipment is very important when conducting a focus group and needs to be good enough to capture group discussion. If the equipment used does not capture all the conversation in the focus group valuable data may be lost.

7.2.5 Automated Data Collection, Data Mining

A smaller number of students decided to use automated data collection for their research

Distance   Engagement
            Analysis of published data
                    Textual analysis
                            Survey
                     Passive observation (and lab experiment)
                            Interview (structured>semi>unstructured)
                            Participant observation

Figure 2 Distance and engagement between researcher and subject with different data gathering methods. Based on (Nandhakumar & Jones, 1997)

project.

Issues that arise include: large sample size. For automated data collection the problem of having too much data arises. Most statistical packages have an upper limit to the sample size that can be processed; 'cleaning the data'. Data mining exercises are as good as the rules used to gather the data. If it is discovered that garbage has been collected either because the original data was not 'clean', or the rules used to collect the data did not use the right discriminators, then the researcher must spend time with the data to obtain accurate data. Problems therefore arise with discarded responses that could not be coded;

Regular backup is required for automated data collection. Automated data collection requires all the procedures of sound information management; and only those postgraduate students with the appropriate level of programming skills are able to make use of automated data collection. Programs are required to be written and tested.

7.2.6 Website collection

Some students gathered data from existing website according to a predefined set of criteria. Problems that arose included: formulating the categories to gather data from websites designed in very different ways; and rapid broadband access was required to download websites.

7.2.7 Organisational Documents

Whilst there are many students that gather data from organisational documents problems associated with this method of data collection did not appear often. One issue mentioned was that some participants were not happy to hand over documents to the researcher.

7.3 Data Analysis (for this project)

When the postgraduate student responses were analysed to see within which domain(Figure 1) the dilemmas were situated it was found that for data gathered by interviewing most of the responses were either within the academic/ business border or situated within the researcher's own experiential domain (Table 1) that floats above the research environment. When surveys were considered it can be seen from Table 1 that most responses lie within the academic/ business border. No responses lie solely within the business arena. For focus groups the few results obtained indicate that knowledge gained from experience within the researcher's domain is an issue. Issues also arise within the business world when focus groups are conducted. When automated data collection and web crawling are considered there are no data gathering dilemmas arising either within the academic/business border or in the business world. The only issue with documentation as a data source lies within business world.

8. IMPLICATIONS

When the results obtained by analysing student responses to data gathering dilemmas for this study are compared to (Nandhakumar & Jones, 1997) distance/engagement with the research process scale, it can be seen that participant observation and interviews rank the highest for engagement. This correlates well with the results

obtained as interviews particularly score highest in experiential learning then followed by surveys. For postgraduate students the highest level of interaction, and therefore the highest incidence of dilemmas occur when data is gathered with interviews. These interactions take place within the world of business or within the border between academia and business. Textual analysis involves little or no interaction with participants and this is reflected in the results obtained. Such analysis – by automated data collection, data mining or website analysis involves little or no interaction with participants and the postgraduate student does not enter engage with the world of business.

9. THE RESEARCH QUESTIONS ADDRESSED

This paper has addressed multiple challenges that present themselves to students as they gather their research data. It appears that the closer the engagement with the participant the more likely the dilemmas are to be found within the world of business in which the research is situated or within the complex border that exists between academia and business. Presenting this collection of dilemmas faced by postgraduate students formally adds to future students' learning and provides a further teaching aid for postgraduate educators. These results also add to the general knowledge pool for postgraduate education and student learning resources.

10. CONCLUSION

Postgraduate students in the MComp program encounter a wide cross section of problems and dilemmas as they gather primary data for their thesis or dissertation. Information about data gathering dilemmas was collected from multiple sources and analysed according to a theoretical framework based on Habermas (1996) worlds of activities and extended by Fielden (2004). This theoretical framework situated data gathering in the world of business, often expressed in academic terms and administered by delegated governmental authority from academic administration (Figure 1).

Information for this paper were amassed by: surveying postgraduate students on their experiences in gathering their research data; observing postgraduate seminars; postgraduate

supervision; and sharing best practice with other postgraduate supervisors. These results were been compared to literature results and the findings presented.

The findings from this paper add to the general knowledge pool for supervisory practices and postgraduate research education.

REFERENCES

University of Western Sydney (2003). Draft Code of Postgraduate Supervision, Australia.

Bartlett, A., & Mercer, G. (2001). *Postgraduate Research Supervision: Transforming Relations*. New York: Peter Lang Publishing, Inc.

Bouffard, S., & Little, P. M. D. (2004). Detangling Data Collection: Methods for Gathering Data. *Harvard Family Research Project, August* (5) gse.harvard. edu/hfrp/projects/afterschool/resources/snapshot5. html.

Craswell, G. (1996). This unfathomable thing called supervision: negotiating working relationships with supervisors. *Quality in Postgraduate Research" Is it happening? Conference, 18-19 April, Adelaide*.

Edwards, H., Aspland, T., O'Leary, J., Ryan, Y., Southey, G., & Timms, P. (1995). *Tracking Postgraduate Supervision*. Brisbane, Queensland: Queensland University of Technology.

Fielden, K. (2004). Dilemmas, disconnects and distortions: boundary crossings explored in conducting research in computing. Keynote address at 2004 NACCQ Post Graduate Research Symposium. In B. Cusack (Ed.), *Proceedings of the NACCQ2004 Post-graduate Symposium* (pp. 1-10).

Heath, T. (2002). A Quantitative Analysis of PhD Students' Views of Supervision. *Higher Education Research & Development, 21*(1), 41-61.

Knobel, M., & Lankshear, C. (2004, April 15). *Dilemmas of the datasphere: Issues impacting information gathering in physical and virtual sites*. Paper presented at the American Education Research Association Annual Meeting, San Dieo, www. geocities.com/c.lankshear/datasphere. html?200525 accessed 25 February 2005.

McMichael, P. (1993). Starting up as supervisors: the perceptions of postgraduate supervision in Australia. *Studies in Higher Education, 18*(1), 15-27.

Melrose, M. (1999). Postgraduate supervision in a New Zealand Polytechnic: staff development, dilemmas and evaluative questions. *Herdsa Conference: Cornerstones*.

Milne, K. S. (1996). Quality Management at Massey University. *Quality in Postgraduate Research" Is it happening? Conference, 18-19 April, Adelaide*.

Nandhakumar, & Jones. (1997). Distance and engagement in IS research. *Information Systems Journal, 7*(2), 109-131.

Russell, A. (1994). Quality in postgraduate research: Some findings from a three-Universities project in South Australia. *presented at Quality in postgraduate research conference, April 7-8, Adelaide*.

Soliman, I. (1999). *Postgraduate Supervision*. Armidale: University of New England.