

Rising to challenges of combining qualitative and quantitative research

Vreda Pieterse

Department of Computer Science
University of Pretoria
Pretoria, 0002
+27 12 420 3652
+27 12 362 5188
E-mail: vpierse@cs.up.ac.za

Inge P Sonnekus

Department of Further Teacher Education
University of South Africa
P O Box 392, UNISA, 0003
+27 12 429 4089
+27 12 429 4922
E-mail: sonneip@unisa.ac.za

Abstract

This paper reports on the methodological challenges encountered during a research project employing a combination of quantitative and qualitative approaches. The worldwide shortage of educators for information technology prompted the researchers to explore the personal qualities desired for Computer Educators who are Destined to Achieve Results (CEDARs) and to develop an assessment instrument to identify a CEDAR, which could be used in the recruitment of potential CEDARs. The overall research project is divided into three stages. This paper only reports on the progress made in stage one.

The research design of stage one follows a three-phase format in which data gathering is done by a questionnaire containing closed items accruing quantitative data and open-ended items accruing qualitative data. A strategy of stepwise refinement is applied during the three-phase progressive project. The first sample was comprised by staff in the Department of Computer Science at the University of Pretoria. Based on the findings of this phase, the questionnaire was subsequently refined and presented to delegates at a national computer educators' conference. The third phase, still in process, will involve a sample of computer science educators located by means of the Internet.

1. Orientation

This paper reports on the methodological challenges we encountered while we exploring the issue of the lack of computer educators. The worldwide shortage of information technology educators and a 42% drop in computer science majors between 1990 and 1995 (Leong 1998:67) have led to an exploration of the problem by various scholars. Although various solutions to the problem have been suggested, we maintain that attracting the right people to the job can solve the problem. In this context 'the right people' are Computer Educators who are Destined to Achieve Results (CEDARs).

As this identified task is not trivial, we used stepwise refinement, which is a technique used to view the problem at hand in terms of smaller subproblems (Brookshear 1994:147). The task of addressing the shortage of computer educators has been broken down into the following subproblems:

1. Ascertain which personal qualities are essential for CEDARs.
2. Develop an assessment instrument that can identify a CEDAR.
3. Implement the assessment instrument to attract identified CEDARs to become computer educators.

Stepwise refinement needs to be applied to each of these identified problems in turn recursively until the subproblems identified by the process are small enough to manage. We have divided the first of the subproblems identified above into a three-phase progressive project and have completed the first two phases. We are currently seeking a forum for the third phase.

2. Aim

This paper describes the challenges we faced when combining quantitative and qualitative methodologies during the first two phases.

3. Method

The literature review revealed little information about the personal qualities of computer educators. Thus, we embarked on the project without prior knowledge regarding the personal qualities of computer educators. The nature of the project pre-empted combining quantitative and qualitative methodologies.

Hathaway (1995) uses the term empirical-analytical to describe the paradigm underlying quantitative research, and interpretive for the paradigm underlying qualitative research. Schulze (2003) argues that educational researchers require both modes of inquiry to advance their understanding of teaching, learning and other human phenomena. Mouton and Marais (1990:169-170) regard such a bridge between approaches as necessary, since a single approach cannot succeed in encompassing human beings in their full complexity. In our project we use questionnaires that combine these two research methodologies in order to grasp the breadth, depth, and richness of computer educators.

Quantitative

The questionnaires used to gather data include certain questions requiring fixed responses also known as closed questions. These items encourage participation because the questions are quick and easy to answer, allow the results to be revealed immediately, and stimulate respondents' thinking with regard to subsequent participation on the qualitative level. The progression of each quantitative questionnaire incorporates the qualitative feedback obtained from respondents in the previous phase.

Qualitative

Open questions gathering qualitative data allow the respondents to reveal their insights once they have completed the quantitative component of the questionnaire. These responses are analysed to ascertain further essential personal qualities of CEDARs and resulting categories are incorporated into the quantitative items in the questionnaire used in the next phase.

Combined

In the first phase of the project, the questionnaire comprised items, which were either quantitative or qualitative. We combined the data to constitute a generic description of a CEDAR. In the second phase we integrated the qualitative aspects into the quantitative items. An open option was added to each closed item to provide an opportunity to the respondents to express their opinion. This opportunity will be enhanced in the third phase of the research with the addition of two-way communication by e-mail. Conversely, quantitative techniques were infused into the representation of the qualitative findings. The frequency with which opinions were expressed in each category was noted and reflected in our representation of the qualitative results.

4. Research design

The research design emanated from a practical need to identify the personal qualities of CEDARs. Data gathering was done by means of questionnaires. In the first phase of the research aimed at addressing the first subproblem, local computer educators were invited to participate by responding to the initial questionnaire. Thereafter, the opportunity to present the results of the first phase of the project at a national computer educators' conference, allowed the researchers to gather further data with a second refined questionnaire. The progressive nature of the research project dawned on the researchers when the opportunity for international exposure through the Internet emerged. Thus, the research design was truly qualitative and oriented to action research as it used the given situation without disturbing the fabric of the context and rose to opportunities while addressing threats as challenges.

5. Validity, Reliability and Ethical Measures

Validity of the findings was addressed by several strategies. Firstly, the author of this paper is a registered educational psychologist with experience in assessing people. The co-author has more than ten years' experience in lecturing computer science at

secondary and tertiary level. The professional experience of computer science educators/lecturers was tapped into by means of the questionnaires. Secondly, in order to eliminate bias in the results, responses to the open ended items were analysed with the aid of an academic peer, thus ensuring inter coder reliability.

Moreover, all participation was voluntary and anonymous. In the first phase respondents were invited by means of an email message to complete the first questionnaire. No results from previous phases were available when respondents answered the questionnaires used in the first two phases. Therefore, we are convinced that the respondents responded without any pre-set ideas other than their own observation and experience.

6. Sample

The challenge was to find suitable samples to use as the project moved through its progressive phases. The selection of each sample is discussed below. The sampling method is a combination of purposeful and comprehensive sampling (McMillan & Schumacher 1997:397-398).

First phase

During the first phase, the sample was selected from one location. The staff on the e-mail list of the School of Information Technology of the University of Pretoria, Pretoria, South Africa was invited to participate. This list comprises 73 academic members of which 34 members responded to the questionnaire within one week. Participation was completely voluntary and anonymous.

As the list comprised academic and administrative staff and did not require respondents to indicate their function, we were unsure who had responded. However, we presumed that the content of the questionnaire would inspire only the academic fraternity to respond. This uncertainty was rectified in the following phases as dedicated computer educator mailing lists were used.

Second phase

For the second phase, the sample represented a national spread of respondents. All members on the e-mail list of the Southern African Computer Lecturers' Association (SACLA) were invited to participate. This ensured that the sample was purposeful and comprehensive. This list has 235 members of which 72 attended the 2003 SACLA conference. A total of 35 conference delegates and 31 other members on the list completed the questionnaire.

Third phase

For the envisaged third phase, an international IT Training group will be approached. Their members will be invited to complete the final questionnaire online.

7. Compiling the questionnaires

Having found no description of the personal qualities of computer educators, neither ideal nor typical, in the literature review, we were confronted with the challenge to generate items for a questionnaire from scratch. The questions that were used in each phase were carefully created to serve the aim of each questionnaire as the project moved through its progressive phases. The aim of each questionnaire and the process that was followed to compile each is discussed below.

First phase

The aim of this phase was to expand the description of a typical CEDAR as seen by us by determining the views of a number of current computer educators at one location.

Nine questions were selected. It was decided to keep the scope of the questionnaire small in an attempt to maximise participation. From own experience we know that CEDARs will more likely be willing to participate if it is clear that responding to a questionnaire will not require much time.

Five categories of characteristics were identified using our description of a typical CEDAR, which was developed through our experience in dealing with CEDARs. Each category was represented by a closed question with four alternative options. The alternatives allowed the respondent to choose the specific personal quality (as defined by the question), which applies, to him/her. The alternatives range from a definite yes, to an absolute no. An even number of alternatives was used to avoid the possibility of opting to remain neutral. These closed questions were included for two reasons: firstly, to stimulate the respondent's thinking and secondly, to attempt to discover laws concerning CEDAR behaviour.

Three open-ended questions were included to facilitate input from the respondents. These inputs were needed to generate further items and gain more insight. Originally we considered having only one open question since we had only one aim for it, namely to identify other characteristics of CEDARs. However, we decided to include two leading questions before the main question. The leading questions were included to assist the respondent shift from a typical analytical mindset to a more qualitative mode.

One closed question was chosen to conclude the questionnaire. The question was formulated to determine if the respondent sees himself/herself as a CEDAR. It could, at a later stage, be used to determine if the respondents' conception of themselves coincided with the information given in the other answers.

Second phase

The aim of this phase was to expand the description of a typical CEDAR as seen by a number of practising computer educators at one location to include the view of computer educators on a national level.

Twenty questions were finally selected. Nine questions from the first questionnaire were retained. The last open question was rephrased to avoid the possibility of generating answers referring to the ideal CEDAR while we are interested in the

characteristics of people who are actually lecturing IT subjects. No new open questions were added because we felt confident that the current open questions were adequate. Three questions were included to gain more information about the respondent's experience as an IT lecturer and to determine if he/she was attending the SACLA conference.

The eleven categories of characteristics that were identified on the basis of the descriptions of a typical CEDAR given by the respondents of the first questionnaire were considered more closely. The categories Altruistic, Cheerful and Dedicated were not used for the new questions as we had already included questions to cover these traits in the first questionnaire that were transferred to the second questionnaire. Two of the categories were of particular interest, namely Patient and Organised, as their exact opposites were also present, namely Impatient and Disorganised. To address these four categories, a direct question was included where the respondent had to indicate to what extent he/she might become impatient. To be organised can refer to many different levels of organisation, thus we chose to include two questions to cover this aspect: a question relating to the physical environment dealing with the neatness of the respondent's workplace and a question concerning the respondent's ability to avoid work-overload. Each of the remaining five categories, namely Cognitive, Free spirited, Perfectionist and Anti-social, were used to design a new question to allow the respondent to indicate the degree to which he/she complied with the implied characteristic.

To complete the even number of 20 questions, a question about marking papers was included. From the analysis of the responses to the first questionnaire, this was the most prominent burden experienced by the respondents. It is our view that it is not the marking as such that is problematic, but the volume with which computer educators must cope.

As in the first questionnaire, the options presented by the closed questions allow the respondent to choose the specific personal quality (as defined in the question) that applies to him/her. The alternatives range from a definite yes to an absolute no. An even number of options was again used to avoid the possibility of neutrality. A difference from the first questionnaire was that all closed questions were accompanied by an open alternative to give respondents more freedom to express their ideas.

Third Phase

The aim of this phase will be to expand the description of a typical CEDAR as seen by computer educators on a national level to include the view which computer educators hold internationally.

The experience and knowledge gained during the second phase and the subsequent data analysis will be used to adapt the questionnaire used during this phase for use in the third phase.

The progressive compilation of the final questionnaire is visualised in Figure 1:

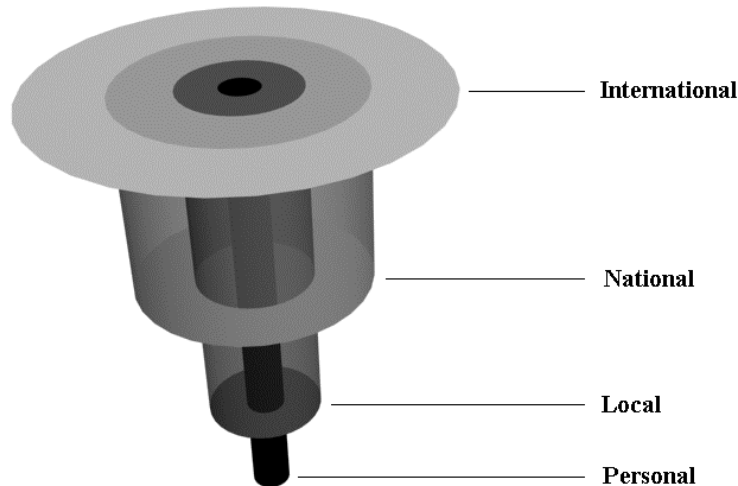


Figure 1: The Progressive Compilation Of The Final Questionnaire

The original personal thinking was expanded in three phases moving upwards from the personal through the local and national levels to the international level. Each new level validates the insights of the previous level and is expanded to include new insights that were gained on the new level.

8. Data collection

We used online questionnaires to collect data in all phases. The challenge was to find suitable tools to deliver the questionnaire in each phase of the project. The characteristics of each tool are discussed below. A further challenge was to motivate the respondents to complete the questionnaire.

First Phase

We used version 1.6 of the open-source web-based survey application called phpESP (which stands for php Easy Survey Package) to create our survey. This survey application can be downloaded from <http://phpesp.sourceforge.net>.

The version of PhpESP that we used enables users to create complex and advanced surveys. It offers a wide variety of options for survey questions, from simple yes/no to multiple choice to scales to open-ended fill-ins. PhpESP provides a good amount of results analysis that can be viewed by the survey manager as they arrive. Results from multiple surveys can be cross-tabulated to view trends, and results can also be exported as comma-delimited text for use in Excel and other analysis tools.

The survey we created was hosted on the server of the Computer Science Department of the University of Pretoria. The IT lecturers were invited to participate using the following e-mail message on a local mailing list:

Subject: Why do we do this to ourselves?

Content:

My paper with the topic "Why do we do this to ourselves?" has been accepted to be presented at SACLA. WE will discuss reasons for not teaching and reasons for teaching Computer Science.

Please contribute to my research effort by answering a few questions online at <http://phugeet.cs.up.ac.za/phpESP/public/survey.php?name=ITLectUP>

It will take only about 5 to 10 minutes of your time.

Thanking You in Advance
Vreda Pieterse

The title of the paper was formulated to arouse the curiosity of people thus encouraging them to read the message. The message was deliberately short and succinct. Own experience has shown that a CEDAR will scan the message quickly and discard it within the first five seconds if he/she does not find it interesting or relevant. Although the message was sent via a mailing list, the message was written in a colloquial style to encourage the receiver to regard his/her participation in the survey as a personal favour, especially for those who know the sender.

The nature of this research differs radically from most computer research projects encountered by the respondents. This may have caused certain respondents to regard the exercise as less serious research, which may have made them more relaxed and honest. We regarded this as an advantage and therefore did not try to persuade them to think otherwise.

Second Phase

Initially we planned to use only the online questionnaire. However, due to logistical issues, this was not possible and we had to face the challenge of disseminating the questionnaire both online and on paper. For the paper version an invitation to participate in the questionnaire printed on brightly coloured paper was distributed at the opening session of the SACLA conference. The paper version of the questionnaire was attached to the invitation.

For the online version, an invitation to participate in the questionnaire was e-mailed to the SACLA mailing list, which included all the delegates to the conference. A convenient link to the questionnaire was included in the e-mail message. After the conference a follow-up call for participation was e-mailed using the SACLA mailing list and using the following e-mail message:

Subject: Why are we doing IT to ourselves?

Content:

We want to thank all those SACLA delegates who answered the 20 questions on our print questionnaire or on line. We found interesting material that is discussed in our paper entitled: "Why are we doing IT to ourselves?" that was presented at SACLA. The revised paper that includes the results from the questionnaire completed by you at the conference will soon be made available.

If you have not yet participated in this survey, you are still welcome to do so by doing the online survey at:

<http://phugeet.cs.up.ac.za/phpESP/public/survey.php?name=CEDAR1>

This will only take about 5 to 10 minutes.

Thank you in anticipation

Vreda Pieterse and Inge Sonnekus

Third phase

The challenge is to motivate strangers to respond. We addressed this issue by inviting them to see the results of the questionnaire in real time and by offering a prize.

At this stage, we realised that the tool used in the previous two phases was no longer suitable. Although we could see the results, the respondents could not view them. We could not find a tool that can display results in real time and support open options to closed questions simultaneously. This unique requirement for the tool represented an immense challenge. Consequently, we have developed a custom-made survey tool to serve our purpose for the third phase. This tool facilitates the following:

1. The possibility to combine all closed questions with open options allowing respondents to add comments even if they have chosen one of the options.
2. Immediate graphical feedback on the closed questions to all respondents after they have answered the questions.
3. The possibility to initiate an e-mail conversation between the respondent and the researcher to clarify obscurities that may arise.
4. The possibility for a respondent to alter his/her responses to questions after they have been submitted the first time and the updating of the corresponding statistics that would be affected by the change. This will allow respondents to continue with the questionnaire should they, for some reason, be unable to finish it at one sitting.

9. Data Analysis

We consider the turnover time from the completion of the data collection to the production of the final results crucial. Having the results ready for respondents soon after they completed the questionnaire is part of our strategy to encourage participation. The challenge was to find ways to accelerate the process of analysis in all phases.

First phase

In an attempt to discover general principles concerning CEDAR behaviour, the frequency of the different answers to the closed questions was determined. Due to the relatively small size of the data sample, the results are not significant. The quality of the responses to the open-ended questions, however, illustrated that the aim of this type of question, namely to stimulate the respondents' thinking, was achieved.

The following bottom-up strategy of data analysis similar to the method used by Johnson and Christensen (2002:426-431) was adopted for the analysis of the responses to the open ended questions:

1. Each answer was studied and divided into meaningful units.
2. Each identified unit was assigned a category code and was described in a full sentence on a master list.
3. The master list was then reapplied to new answers and expanded to include new categories. Some descriptions were altered as the need arose.
4. In order to address inter coder and intra coder reliability, the above mentioned strategy was followed individually and then discussed with each other before the final result was agreed upon by the authors.

The frequency with which opinions were expressed was noted to identify prominent patterns in the data sample. The data analysis resulted in the emergence of interesting facts and personal views, which may prove significant using a larger data sample.

Second phase

During this phase we were faced with a logistical problem that prevented us from using our prepared online questionnaire. The challenge was to find a viable alternative to gather data. The implementation of a print version of the questionnaire introduced yet another challenge to find a way to combine the data that was gathered on paper with the electronic data gathered through the online version of the same questionnaire without losing too much time.

For the closed questions the following time consuming work was done at the conference:

1. We created a spreadsheet containing a separate worksheet for each closed question. Each worksheet had a framework to total the incidence of each of the answers, and draw histograms of the results.
2. Macros were written to import the data gathered in the online questionnaire into these spreadsheets.
3. A document containing graphs linked to the above spreadsheets, showing the quantitative results was created.

When faced with the challenge of having the data gathered on paper included in the results, we considered the option of omitting it. We were, however, determined to uphold the promise we had made to the conference delegates and decided to rise to the challenge of including this data in the results revealed at the conference. To achieve this goal we did the following while the delegates were completing the questionnaire on paper:

1. An extra worksheet was added to the spreadsheet.
2. The cells in the new sheet were linked to the rest of the spreadsheet in such a manner that the answers of the respondents could be typed in on the single sheet in adjacent cells while the corresponding data that was used to create the histograms were automatically updated.

When the respondents had completed the survey, the above-mentioned documents and tools were implemented to generate the resulting handout. The printable format of the results was generated within 22 minutes from the time we started processing the answered questionnaires. The generated graphs were used to identify prominent patterns. This document can be viewed at http://www.cs.up.ac.za/~vpieterse/pub/SACLA_Graphs1.htm

The same method of analysis was followed with the answers to the open-ended questions, as in the first phase. Due to the time consuming nature of qualitative analysis, they were processed after the conference and were therefore not included in the handout that was prepared for the presentation at the conference. Information gathered from the open options of closed questions was also interpreted. Information that was obtained through informal conversations at the conference with respondents was also incorporated to produce the results and draw the conclusion shown in this paper.

Third phase

For the closed questions the web-application used to host the on-line questionnaire will total the incidence of each of the answers, and draw histograms of the results that will be shown immediately to the respondents.

Our experience while conducting the questionnaire at the SACLA conference and analysing the information of the open options appended to the closed questions resulted in the inclusion of the facility to initiate an e-mail conversation between a respondent and the researchers to clarify obscurities that may arise. After any obscurities are clarified, the respondents will be able to alter their responses. Such alterations will have an immediate effect on the histograms.

We will follow the same method of analysis with the answers to the open questions, as in the previous phases. Where deemed necessary, further clarification of responses can be obtained through informal conversations with respondents using e-mail.

10. Results

Each phase produced both quantitative and qualitative data. Although the quantitative data were easily shown in histograms, the challenge was to represent the content of the qualitative items comprehensibly. A further challenge was to amalgamate them sensibly into a useful description of the CEDAR.

First phase

The quantitative results were shown in histograms. Figure 2 shows an example.

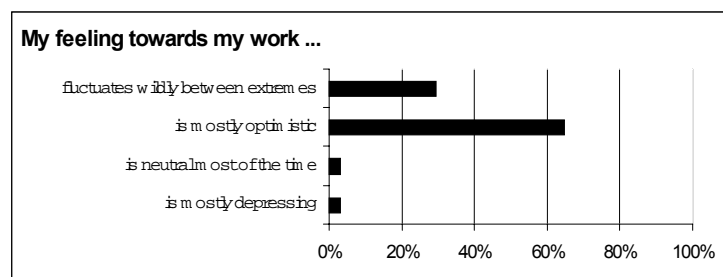


Figure 2: Example Of First Phase Quantitative Results

Table 1 shows the categories in the master list of personal qualities. Some examples of raw data are also shown.

Table 1: First Phase Qualitative Results

Description	Examples of raw data (verbatim)
Cognitive	Thirst for more knowledge, wide range of interests, always on the forefront of new developments, logical.
Dedicated	Persistence, devotion to duty, enthusiasm. Someone who has passion for what they teach.
Impatient	Impatient.
Patient	Loads of patience.
Altruistic	Like sharing knowledge and experience. The ability to convey (using imagery, etc.) highly abstract concepts in ways that really help the students. Able to capture a student's attention
Free spirited	Idealism, rather than realism. Completely NUTS. Creative. Innovative. Have difficulty expressing concepts (since they are intuitive).
Perfectionist	Attention to small details. Properly following and implementing set rules.
Cheerful	Positive attitude, ... less stressful than the commercial world..., ... but always fun. Friendly.
Organised	An expert in time management.
Anti-social	Not very sportive. No social life.
Disorganised	Always stressed up to meet deadlines.

At this stage, we concluded that the emergence of contradictory categories could be attributed to the fact that some respondents described the ideal CEDAR while others described the situation as it is. This constituted a challenge to avoid such confusion in the second phase.

The combined interpretation of both quantitative and qualitative results gave rise to the following description of the personal qualities of a CEDAR:

CEDARs are individuals who are keen to learn new technological and challenging things. They are thrilled to share their knowledge with others who can become

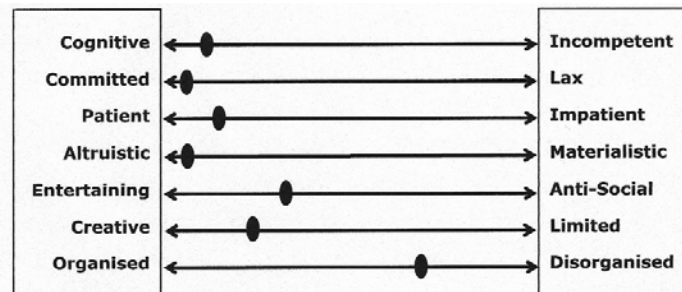
equally excited about it. They are excessively passionate and dedicated to their work and are willing to work long hours at personal expense to achieve certain goals.

We recognised the following limitations of this phase:

1. The study was conducted in only one location and may therefore be specific to this location.
2. Although 46.6% participation was achieved, we infer that the characteristics and insight of those who did not participate might be crucially significant.
3. The conclusion is possibly flawed as some respondents described the ideal CEDAR while others described the situation as it is. These two approaches can account for contradictions in the findings. Questions in the second phase were accordingly specifically focused on the real situation.
4. The description can easily be amended to describe a successful educator in other areas with only subject specific changes. It can therefore not be seen as a description of the unique personal qualities of a CEDAR.

Second phase

The quantitative results were shown in histograms identical to those used in the first phase.



The analysis of the qualitative responses culminated in the identification of seven categories of personal qualities, and their antitheses. The challenge was to show these categories combined with their antitheses. We chose to use a Likert scale format to indicate the extent to which each personal quality or its antithesis is present as shown in Figure 3:

Figure 3: Categories and Antitheses of Personal Qualities of Computer Educators

Combining the above results, the following short description of a CEDAR was compiled:

CEDARs are brainy, committed, patient and altruistic people who have a tendency to be entertaining and creative yet disorganised.

The following limitations will be challenged in the third phase:

1. The study was conducted in South Africa and may therefore be specific to this country.
2. The percentage return rate on the survey that was conducted online was only 13.2%. The low response can be attributed to the fact that respondents might not have considered this research relevant to computer science. Wilson (1996) feels

"that neither market nor academic research will achieve good response rates if the research is not seen to be directly relevant to people's lives" (in Mann & Stewart 2000:28).

3. Respondents queried items, yet could not clarify their doubts.
4. The qualitative results seem to duplicate the quantitative results: possibly the respondents were sensitised by the content of the closed questions. This constitutes a challenge to avoid such sensitising in the third phase.

11. Conclusion

The greatest challenge that we were confronted with was to conceptualise the problem at hand, namely the shortage of computer educators. We have risen to it by applying the technique of stepwise refinement. This well-known computer science problem solving technique is uniquely applied to this social science problem. However, other challenges surfaced during the three phase progressive project that we generated to address the first subproblem. This paper pinpoints the challenges we encountered and explains how we rose to them. The use of both quantitative and qualitative research methods was necessary due to the nature of the IT world in which our problem is embedded. In the same manner, some of the challenges we encountered were uniquely related to IT discipline.

During the course of the research, we discovered not only problems, but also new opportunities, which we had not envisaged – these were also seen as challenges and discussed in this paper.

12. References

- Brookshear, J.G. 1994. *Computer science: An overview*. (4th edition.) California: The Benjamin/Cummings Publishing Co., Inc.
- Hathaway, R.S. 1995. "Assumptions underlying quantitative and qualitative research: Implications for institutional research." *Research in Higher Education* 36(5): 535-562.
- Johnson, B. & Christensen, L. 2002. *Educational research. Quantitative and qualitative approaches*. Boston: Allyn & Bacon.
- Leong, Kathy Chin. 1998. "Feds take a Crack at IT Crunch." *Internet Week* 698:67 <http://0-web5.epnet.com.oasis.unisa.ac.za/elivery.asp> Accessed 2003/07/09.
- McMillan, J.H. & Schumacher, S. 1997. *Research in education: a conceptual framework*. (4th edition.) New York: Longman.
- Mann, C. & Stewart, F. 2000. *Internet communication and qualitative research*. London: Sage.
- Mouton, J., & Marais, H.C. 1990. *Basic concepts in the methodology of the social sciences*. Pretoria: Human Sciences Research Council.
- Schulze, S. 2003. "Bridging the Grand Canyon: Views on the combination of quantitative and qualitative research approaches." (Inaugural lecture, Faculty of Education, University of South Africa, Pretoria).