Paper presented at the ARTIST Roundtable Conference on Assessment in Statistics
held at Lawrence University, August 1-4, 2004

Assessment: viewing through rose coloured glasses?[i]

by
Anne Porter and David Griffiths
School of Mathematics and Applied Statistics
ARTIST Roundtable Conference August 4-8, 2004, Appleton

Assessment can be the droll, time-consuming task through which we grade students' performance or it can be an illuminating activity revealing how students think and interpret. It can lead us to changes in both pedagogy and curriculum.

This paper is based on the experience of teaching in a tertiary preparation Mathematics and Statistics program. The pedagogical approach adopted was that of activity based learning. Students were to construct their knowledge based on the activities in which they participated. Students were asked to complete five pieces of homework for which they were given full marks if the work was submitted. Essentially they were asked to reflect upon the activity they had completed in class and to exemplify it with another example. They were encouraged to identify their weaknesses rather than to disguise them. Feedback to students was designed to clarify the statistical ideas as discussed by students. At the end of term they were to illustrate all the major ideas in a piece of work that was graded on quality and to sit a final examination.

Never would we have believed that adults participating in the one activity could have so many different interpretations of their experience and of the feedback that was given to them. This paper will discuss the insights gained and the subsequent impact on developing explicit frameworks for learning and the implications for teaching tertiary level students.

Assessment is just another teaching strategy, or so it was thought until, an activity based statistics program with reflective home-works, opened a Pandora's box.

Assessment is one mechanism through which we can gain insight into how students construe what it is we teach. It is a mechanism through which we can gain insight into the student experience and quality of student learning. Through this insight we can potentially develop better learning environments, although, as will be explained later, they are not necessarily more comfortable. Change in assessment practices is only one of the many mechanisms that can lead to an improvement in the quality of teaching and learning.

In evaluating innovation in the context of improving student experience and learning, it is important to recognise that the innovation is part of a context. In this paper the focus is upon assessment but the context is much broader. The effectiveness of one innovation is in the context of there being an underlying philosophical approach to teaching and learning, curricula and associated content and sequencing, the presence or absence of an emphasis on meta-cognitive aspects of learning, teaching practices, management systems and these days online learning resources, (Porter, 2002; Porter et al 2003). Assessment written about in this paper was implemented and analysed in the context of there being a study directed at improving statistical education. The key mechanism for improvement was deemed to be reflective practice (Schon, 1983; 1991), and assessment was one means of providing evidence upon which to reflect. Reflective practice was defined as:

> …a process for up-ending what was thought known. It became a strategy for looking at practices and student learning from different vantage points. Changes made improved learning outcomes only if the reflection directed the appropriate path and, as I reflected, only if student voices were heard. When changes do not result in an improvement, the changed practice in turn becomes problematic and reflection may suggest another strategy for improvement. (Porter, 2001, p. 334).

The outcomes that can be examined when improving teaching and learning can be various. They may be attitudinal changes or changes in achievement levels, or in retention, critical thinking and a host of other outcomes.

Education is situated in an economically and politically constrained environment. Time and cost constraints affect what assessment can be given and marked. Policies to ensure that work is not plagiarised, or completed in a supervised environment at a defined time may also limit what can be asked of students. The type of assessment given to small classes of students may not be feasible when dealing with large classes. Assessment used in a tertiary entrance program may not be considered appropriate at a tertiary level.

The students participating were completing a subject embedded in a tertiary entrance program, Gateway. This program included three strands: Language and Literacy, Statistical and Mathematical Literacy, and, University Learning (critical thinking, history, psychology, intercultural awareness and library skills). On completion of the Gateway program, a student's aggregate mark was used to produce a ranking equivalent for entry to the University of Wollongong, Australia, in competition with School –leavers and

other applicants. The students were typically from disadvantaged social, economic and educational backgrounds. Many had non-English speaking or aboriginal backgrounds, little recent schooling and many had poor experiences in Mathematics. The students could be described as emotionally raw as they grappled with the demands of the course. Two of the key aims of the Gateway programs were to encourage students to accept responsibility for their own learning within a community of learners, and to prepare students for entry to University. Data for this paper were collected over five sessional intakes of students. There were typically two or three classes per session with approximately 25 students in each class, with in excess of 300 students participating and more than 1500 marked assessment tasks.

The subject and students were the focus of a grounded theory with teacher-as-researcher (Cookson; 1987; Doig, 1994; Gurney, 1989; Mitchell, 1985; Zeichner, 1995) doctoral study into how to improve statistical education (Porter, 2001). The subject taught changed in name from *Mathematics and Statistical Literacy* to *Learning Mathematics and Statistics* as pedagogical emphasis shifted from the discipline to how to learn the discipline. Assessment, within a reflective practitioner framework, played a major role in revealing both what students were learning and the nature of the learning experience. Indeed for the first author whose teaching was the subject of these experiments in assessment, they provided the most useful insight into how to improve student learning of statistics that she has ever encountered.

The teaching of the statistics component was based on an interpretation (possibly a misinterpretation) of constructivist theory that students need to construct their own knowledge. Each week, students participated in classroom activities designed to reveal the nature of variation. Exercises were conducted on measurement, sampling, designing and comparing. All involved aspects of data exploration analysis and decision making.

The assessment that was so revealing involved students reflecting upon a statistics activity undertaken in class. Students completed three to five reflective homeworks, depending on session of intake. Students were given full marks for submission of the homework irrespective of content. This represented 10% of the total mark for the subject. However, the amount of feedback given to students so that they could improve their work

and submit a high quality final paper was in direct relationship to the amount of written material they presented in the homework.

The small reflective homeworks led to a major piece of writing that was graded for quality (30%) and in this they were prompted as follows:

> Throughout the course we have been identifying, using and reflecting upon a number of statistical concepts and ideas. Weave a story around as many of these concepts as you can to show the place of statistical concepts in building knowledge. (Porter, 2001, p. 235)

At the end of session they were to complete a final mathematics and statistics examination (30%) and this included drawing a concept map of major statistical ideas. Assessment of mathematical skills, through skills test and the final examination formed the other 30% of the assessment schedule.

The lecturer provided to all students not only individual feedback but generic feedback of outcomes expected and selected examples of marked student papers. In so doing, she needed to maintain the dignity of the students. She developed an analytical approach in which she tried to ascertain *how students were diverted from writing what was considered an appropriate answer*, (Porter, 2001, p. 267). Over several iterations of the subject it became apparent that reflection is not necessarily easy for adults. Students' written reflections yielded diverse interpretations of the statistical lectures and activities.

Reflective homework – examples from Porter (2001)

Exhibits 1 and 2 are two examples of reflective homeworks. They were completed after an exercise in which students were to determine the questions they would need to have answered in order to decide if there was a difference in sealife before and after a breakwater was built. They were asked to think back over the exercise, identify the major ideas, exemplify them in another situation and to write them up as homework.

Student feedback included both a commentary on individual work together with exemplification of how the question had been addressed by the class, together with the teachers expectations as to what would be covered. The process of maintaining students'

integrity by making meta-level comments on student reflective homeworks revealed issues about statistical understanding and also issues related to student learning.

Several homeworks such as in Exhibit 1 focussed only on the level of generating questions and perhaps applying those questions to an area of their everyday life. They omitted the explicit identification of the statistical/research process.

Exhibit 1 Student homework - questions but no process

Student homework and teacher commentary
Students' homework
I have realised now that I use statistics in my work everyday. When an invoice comes in I have to ask many questions before I can pay it. The questions I need to ask are: Has an order been done for the goods? Does the invoice match the order? Have the goods been received? All of the answers have to be collated and then the invoice can be paid or not paid depending upon the answers.
Teacher's feedback
This homework can be used to illustrate the notion of levels of meaning that were illustrated in the previous week's work. Quite correctly the writer has tapped into the idea that there are questions that must be answered in order to make decisions. The student has applied this to an everyday situation as requested. However, the seashell exercise went beyond asking questions to classifying those questions into distinct categories. It is the application of each of these types of questions (eg measurement, design, sampling, ethics etc) that is the real guts of the question.

There was evidence in this and other homeworks that students might not be able to unpack questions in order to identify the issues which needed to be addressed. To undertake the homework, students needed firstly to identify the major ideas or issues that were covered in the unit. This reconstruction could initially be chronological but it ultimately needed to be a logical reorganisation of ideas. Regrouping ideas in other than chronological order to provide a coherent and meaningful structure is a non trivial task that may require several attempts. Then students needed to generalise the ideas to issues found in everyday life. Ideally, the example chosen incorporated as many of the statistical issues as was possible, rather than having a different example for each different idea encountered in class.

For many students, the reconstruction of what took place in class was the first and only phase in writing. Some students did not write about the major statistical issues, for example the student who after an exercise comparing seashells exemplified with another environmental theme, not the statistical ideas. Students were often seen to be diverted by the easy concepts, explaining what they felt they understood. Others were diverted by the hard issues trying to unravel the meaning in what was difficult for the. Students often had to learn to provide a good coverage of the statistical issues encountered no matter whether they were hard or easy. They had to recognise that learning did not have to be difficult for it to have taken place. Students who expected they that they would struggle when learning statistics, did not always recognise that they had learned if it came easily from their experiences.

The homeworks also suggested that students could make 'powerful use' of their statistical ideas (refer Exhibit 2). This student was able to implicitly demonstrate the use of the statistical process, without reference to the original exercise, but through reference to an example of his own creation. The student was able to identify and use the many ideas that had been elicited in the exercise. In his analysis of the situation, he had fully pursued the implications of operationalising his definitions in various forms, and understood the implications of his sampling and design plans. He was also able to convey this without needing to refer to the previous class example; he had successfully transferred the ideas to a different topic.

Exhibit 2 Student homework - transfer of ideas

Student homework	
I held the view that there were more job vacancies advertised on Wednesday and Saturdays in the Illawarra Mercury.	
1.	I counted the number of vacancies in Saturday's and Wednesday's Mercury. Though initially the figures for these two days was higher (Saturday being the highest) than those placed in Tuesdays, Thursdays and Fridays edition this was before other variables were taken into consideration.
2.	When counting the advertisements, the ones that were run for more than one day eg an advertisement ran for two or three consecutive days was not counted on the second or third day of publication.
3.	Categorising. The positions advertised were placed into categories determined by the hours or length employed. 1. Permanent (35 hours per week or more) 2. Casual (less than 24 hours per week) 3. Temporary (less than two months)

Student homework

4. Seasonal (less than three months)

4. What type of employment was advertised eg professional, semi-professional, skilled and unskilled. Was there any relationship to the days they appeared?

5. What was newspaper circulation on particular days such as Wednesdays and Saturdays opposed to other days of the week. If so were there more readers of advertisements?

6. What percentage of employment adds appeared that only paid commission (are they real jobs?)

7. What adds appeared consistently week in week out? eg letter box deliveries.

8. What percentage of jobs was placed by employment agencies as a draw card (the jobs may not exist?)

9. What percentage of jobs advertised paid no commission or no wage eg acting. (that is were not really jobs)

10. What percentage of jobs that appeared in the paper had already been filled before the paper had been published.

Using a standard count of positions vacant does not necessarily give an accurate answer to the question but by including factors previously stated and allowing for deviations the initial statement may not necessarily be true. Even if this survey had been undertaken it would still be inaccurate as it would have only been taken over a 7 day, seven publication issue, which would discount factors such as seasonality. Thus it would not be a conclusive or true measure unless it was undertaken for a longer period of time.

This homework was illustrative of good statistical thinking yet, it was generated by a pre-tertiary student. The questions reflected the sorts of issues that many senior students and researchers often fail to address.

For all homework there was evidence that many students attempt to recall ideas rather than trying to reconstruct them from the task description or exercise. For example the student who came to see me, disturbed because he could not remember, but who when asked to reconstruct ideas from the exercise could cover a blackboard in statistical issues which had arisen, or Mai who for her first homework wrote:

> No notes for reference and no memory of the particulars to draw upon.

Many adult learners found reflection difficult. Max is one of many students who deliberately sought out a text and echoed a list of terms rather than actually reflecting upon the class activities. He wrote:

> I borrowed basic statistics 3rd edition (Spatz and Johnston) and learnt about descriptive and inferential stats, parameters, population, sample, subsample, variables, quantitative variables, qualitative variables, nominal scales, ordinal, interval and ratio scales, as well as dependent, independent, extraneous variables and lower and upper limits, (qualitative variable which use the nominal scale). These are now part of my memory yet, I still do not understand their place. JIGSAW. (Max)

Seeking out a textbook is not an uncommon response. Several other students found a text and used it for their homework. It is quite common to see student responses indicating what they should be looking for in a graph or plot (the theory) rather than indicating what it is they see (the data) when they look at a plot.

Also evident from the reflections on the class activities was that students writing reflected different levels of meaning. For example Lyn provided a list of terms encountered for her first homework.

- frequency
- variables
- trimming mean
- pie chart
- validity of variables
- common denominator (Lyn)

Another student's list communicated a greater awareness of what the student of concepts by grouping them revealing, two levels of meaning in the terms as follows:

Centre
    mean, mode, average
Distribution
    pie chart, cumulative frequency
Comparison
    table, chart (David)

These levels of meaning can easily be extended to a third (definitions of the statistical terms, mean and mode) and to a fourth, which is *how to do it*!, and even to a fifth: how to interpret what they have calculated. A constant dilemma is how to get students to address the multiple layers of meaning when given an assessment task.

Perhaps one of the most astounding revelations from this approach to assessment was that the system was essentially not abused. On only two occasions did different students submit their homework making no attempt to answer the questions, that is they submitted only their name and the assignment title to gain full marks. Perhaps they recognised that they needed to get feedback in order to be able to produce their major assignment. Students did move from short one line, one paragraph pieces of work to comprehensive, and high quality output by the time they produced their major assignment. It was striking with this assessment package that students who did not survive the course could not

submit assessment that they were not comfortable with. Indeed it became evident that for many students to survive the rigours of study that they needed learn to play the system to their advantage. They needed to learn to ask for help and take help where it was available. They needed to learn that when life's events were difficult to deal with, that on occasion they needed to simply comply with the system. In this instance students needed to be able to submit assessment that was less than perfect. By contrast Universities tend to develop rules and regulations to circumvent those students who play the system too well and play little heed to encouraging students to use the system to their advantage.

Most students could readily talk about ideas but had great difficulty in writing about them. When asked to write their reflections they became anxious.

Feedback

The feedback on the homework involved taking each student assignment and identifying how explanations of statistical concepts could be improved, refined, extended or exemplified. Concepts omitted were listed. Suggestions as to how their writing could be improved were also included. There was a deliberate attempt to remove the "good" and "bad" commentary, to not create praise as a motivator but to establish a classroom wherein improvement comes from critical and constructive commentary. The rationale for such feedback as provided to students was that all work at a pre-tertiary level could be improved upon. Getting things wrong, or being able to be less than perfect is a key aspect of learning environments. All assignments whether of poor or high distinction quality were extensively commented. The next step for improvement was clearly identified. The more work the student gave the more comments would be used to extend their thinking and writing so that they were better prepared for the final assignment. Students were told that if their paper was not covered in comments that they should feel short-changed, for a lot of improvement can be made between a first level paper and that provided by a third year.

Lecturers often wonder if students have understood feedback given to them on their assignments. Do they ever listen? These thoughts led Porter (2001) to investigate the

impact on students of the feedback given to them. She asked three questions of her students.

- How did you interpret the feedback?
- Did the feedback make sense to you?
- What was your reaction to the feedback?

Not surprisingly perhaps there is considerable variation in how students construe feedback on their assessment. The early comments on feedback revealed the following.

- For some the written feedback was insufficient for them to understand.

> I have trouble understanding the comments. It is a bit clearer after discussion but I feel more practice would help me most. Perhaps an in class activity would be beneficial so I can reflect on what I am doing right or wrong.

- For some feedback showing how to extend their work is like a reprimand.

> I apparently don't know what is required of me. I'm not doing it right. I am not picking up what I am supposed to do. I don't seem to have a concept of Statistics at all.
>
> Does no tick mean that it was aweful? How can I change my way of thinking – only practice. My writing relates to letters, reports etc as is evident from my homework. Would an appointment with you help please?

- Some interpreted feedback as indicating they did not understand, while for others their lack of understanding was confirmed.

> My reaction to my paper was "well I knew I was doing it incorrectly in the first place – and this just confirmed it!"

- Some recognised that their submissions were incomplete or needed to be expanded and justified with reasons

> Thought I had discussed method. After comparing last weeks to this weeks I think I may have developed further. Unsure what 'go on' precisely means. Did this comment actually mean ideas as to how I would have gone about my procedure.

- Some focussed on working out what the teacher wanted in writing, rather than identifying an approach to writing that best explained the concepts they encountered

> I wasn't upset by your comments put on my paper. I was a little disappointed though as I did spend a lot of time trying to work out what was expected of me and how I was going to do it. I was half expecting I had done it wrong because I wasn't sure what you meant by essay, lab report, whether to do it in an essay form like I did, or go under headings in point form like you suggested; I also didn't know whether you wanted us to do diagrams or not. But I'll know next time.

- Others simply saw the potential for feedback to expand and clarify their ideas rather than interpreting the comments to mean that it was wrong.

> Glad for the feedback. I felt like I was working blind. Hopefully this weeks will be better. I think that I have a long way to go! I have included lots of variability in this next paper, and the sampling as well. Showed me where I needed to expand and to see where the different processes do fit in. Made things a bit clearer to me.

In a situation where there was no grade, and comments made to extend the work of all, the reactions were initially dominated by negative expressions of disappointment and confusion. It is extremely disconcerting for students homework to be treated in this manner – no grades, no ticks, but as the students also conceded failure on their early papers would not have been a comfortable experience either. However the intention was to create a situation where students wanted feedback for the sake of learning not to get a particular grade.

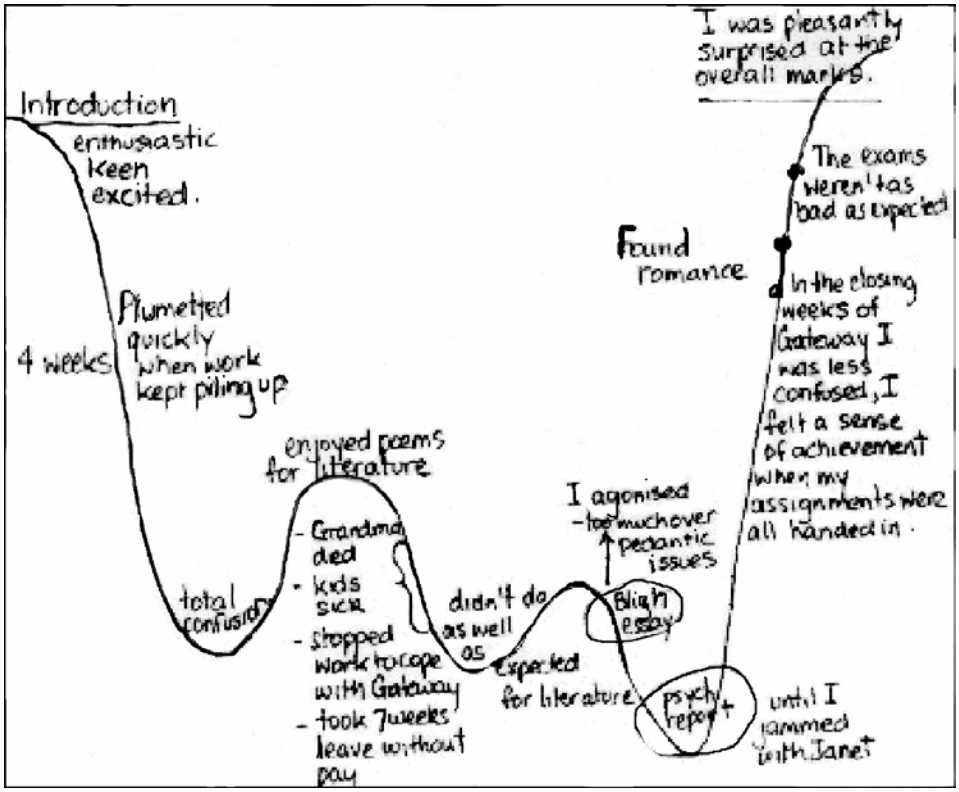
There is an interesting connection between statistics and the experiences of learning as revealed through teaching this subject. Statistics is a discipline which is about variation and uncertainty. Decisions and conclusions are made but attached to them is a varying amount of uncertainty. Often as consultant statisticians we make decisions which are the best we can under the circumstances and this leaves an uncomfortable feeling. The struggle for students learning statistics in this manner is predominantly one to do with feeling uncertain, "is this right?" Rather than remove situations that would lead to uncomfortable feelings typically associated with being uncertain in the learning environment the subject increasingly shifted in orientation to be a subject explicitly described and taught as learning statistics. In so doing both achievements in statistics and student attitudes toward statistics were enhanced. As one student commented:

> I think I was lost at times because I wanted to be spoon-fed but now I realise that I had to tackle statistics the way I knew how and keep persevering with it. …I'm so glad I was pushed to think and was not told the things I wanted to be told. I don't think I have conquered the uncertainty but I think I have become aware of it and I am not as concerned about it any more.

It also became evident that students in each of the classes travelled similar pathways. The patterns of anxiety, uncertainty and empowerment could be identified over the session. Students typically peaked in anxiety about the fourth week just after the first piece of feedback was received and the second piece of work was due. Through knowing the characteristic pathways that students followed in terms of their affect (refer Exhibit 5 for an example) typically evident through their experiences trying to write it was possible to

build a framework around the discipline content and processes that addressed metacognitive aspects of learning.

Exhibit 3 A Students' journey through Gateway



Conclusion

Kelly (1955) describes one of the central ideas in constructive alternativism.

> … there are many workable alternative ways for one to construe one's world. The theory itself starts with the basic assumption, or postulate, that a persons' processes are psychologically channelized [sic] by ways in which he anticipates events. That is to say human behaviour may be viewed as basically anticipatory rather than reactive, and that new avenues of behaviour open themselves to a person when he reconstrues the course of events surrounding him (p.3).

Our interpretation of this is to view each of us as seeing the world through our own coloured glasses, these glasses representing our construct system. While there was much diversity in what sense students made of the activities in class it also became apparent over hundreds of assignments that there were patterns repeated by students. As the focus shifted toward students learning to learn statistics rather than simply learning the discipline attention was explicitly drawn to characteristic ways students had responded. For example there is an issue for students to confront if they do not naturally look for

overall patterns or if they focus only on detail of their own contributions. Statisticians need to be able to look both at specific data points and to identify overall patterns.

The approach to assessment crystallised into one in which all within session assessment was process oriented, with students encouraged to illuminate what they do not know rather than disguise it so that feedback could clarify misinterpretation. Students needed to convey the decision points they had reached, what they were doing, why they were doing it and how they could carry out the various tasks and why they had interpreted in the manner they did. Hence, there was less concern with marking the accuracy of interpretation, for example aptly describing a distribution of data as skewed in shape, than there was with ensuring that the student realised that shape was an important feature of data that needs examination. They needed to express the logic that led them their interpretation so that this interpretation could then be discussed. The final end of session assessment required both process and reasonable interpretation.

The combination of activity based teaching, wherein students were to reflect and write about what they had experienced has led to the conclusion that "knowledge is actively constructed by the learner no matter what the pedagogy" Porter (2001). The key to improving teaching and through implication students learning is for the teacher to engage in reflective practice and in so doing to know how students have constructed their knowledge. Understanding how students construct their knowledge enables the teacher to change, rearrange or indeed retain aspects of the learning environment. In the process of reflecting to improve statistical education the form of assessment is critical.

References

Cookson, P. W. (1987). Closing the rift between scholarship and practice: the need to revitalize educational research. *Educational Policy*, 1, 321-331.

Doig, S. (1994). *Pinning the tail on the donkey: the placement of teacher voice in educational research*. Paper presented at the Annual Meeting of the Australasian Associateion of Research In Education, Newcastle.

Gurney, M. (1989). Implementor or innovator? A teacher's challenge to the retrictive paradigm of traditional research. In P. Lomax (Ed.), *The Management of Change* (pp. 13-51). Clevedon, UK: Multilingual Matters.

Kelly, G. A. (1955) *The Psychology of Personal Constructs: Clinical diagnosis and Psychotherapy.* (Vol 2). New York: Routledge.

Mitchell, P. (1985) A teacher's view of Educational Research. In M. Shipman (Ed.) *Educational Research: Principles,Policies andPractices* (pp 81-96). London: Falmer Press.

Porter, A. L. (2001). *Improving Statistical education through the experience of reflective practice.* Doctoral Dissertation, University of Wollongong, Australia. Available online at http://www.math.uow.edu.au/people/ap_refs.pdf.

Porter, A. L (2002). Improving teaching, improving learning: a framework for contextualising statistical education research. Invited Paper, Australian Mathematical Society, Newcastle.

Porter, A. L., Griffiths, D. G., and Hedberg, J. (2003). From classroom to online teaching: experiences in improving statistics education. *Journal of Applied Mathematics and Decision Sciences*, Vol 7, No 2, 2003, 65-73.

Zeichner, K. M.(1995). Beyond the divide of teacher research and academic research. *Teachers and Teaching: Theory and Practice*, 1(2), 153-172.

Schon, D. A. (1983). *The Reflective Practitioner: How Professionals Think in Action*. New York: Basic Books.

Schon (1991). *The reflective turn*. New York: Teachers College Press

[i] Much of the work in this paper has been reported in a doctoral dissertation (Porter, 2001)