MEASUREMENT UNCERTAINTY AS A THRESHOLD CONCEPT IN PHYSICS

Anna Wilson[a], Gerlese Åkerlind [a], Paul Francis[a], Les Kirkup[b], Jo McKenzie [b], Darren Pearce[c], Manjula D. Sharma[d]

Presenting author: Anna Wilson (anna.wilson@anu.edu.au)
[a] Australian National University, Canberra ACT 0200, Australia
[b] University of Technology Sydney, Sydney NSW 2007, Australia
[c] Queensland University of Technology, Brisbane Qld 4000, Australia
[d] School of Physics, The University of Sydney, Sydney NSW 2006, Australia

KEYWORDS: uncertainty, threshold concepts, curriculum design

ABSTRACT

We report on the initial findings of a study aimed at developing ways to address threshold concepts in the design of undergraduate curricula, involving academics in two disciplines (physics and law) from four Australian universities
The present paper compares two different processes by which physics academics identified and characterised a candidate threshold concept, measurement uncertainty, using student interviews and their own experiences as teachers.

Proceedings of the 16th UniServe Science Annual Conference, University of Sydney, Sept 29th to Oct 1st, 2010, pages 98-103, ISBN Number 978-0-9808597-1-3

INTRODUCTION

It has been suggested that, within each discipline, there are a limited number of concepts that are 'threshold' in nature, so-called because they act as 'conceptual gateways' to disciplinary ways of thinking. Such concepts are proposed to form a subset of fundamental (or key) discipline concepts, distinguished by five criteria: that such concepts are (1) transformative, (2) integrative, (3) probably irreversible, (4) frequently boundary-defining and (5) potentially troublesome to learn (Meyer & Land, 2006).

Because of their gateway nature, threshold concepts are considered to play a key role in the development of students' disciplinary ways of thinking. Students who gain understanding of a threshold concept obtain "a transformed internal view of subject matter, subject landscape or even world view" (Meyer & Land, 2005 p.373), leading not only to new ways of understanding a subject area but a shift in the learner's sense of identity. Students who fail to grasp a threshold concept find their learning path blocked, with no means to proceed.

However, the very transformative and integrative nature of these concepts can make them troublesome to learn (Perkins, 2006). Incomplete understanding or misunderstanding of threshold concepts is likely to have long-lasting implications for students' learning in the subject area, including their ability to apply their learning in new and unfamiliar contexts. Such incomplete learning creates a path-blocking effect and a subsequent push to rote learning (Davies, 2006). Such rote-learning may explain why students may be able to apply discipline methods well enough to pass an exam, but may not be able to adapt their learning to a new context or setting and may not acquire discipline ways of thinking. Teachers of physics (and other science disciplines) may recognise such instances in, for example, students who can apply the equations that express Newton's law, but without letting go of a fundamentally Aristotelian view of the relationships between force and motion.

Threshold concepts may thus provide a particularly valuable focus for curriculum design attention. However, the exploitation of such concepts as curriculum foci is not unproblematic. Although the idea of threshold concepts has proved to have widespread appeal to teachers in higher education, and Meyer and Land (2005) note that academics are quick to suggest threshold concepts in their own disciplines, it is not necessarily easy for academics to distinguish threshold concepts from key concepts. As yet, no clear strategies have been devised to assist teachers in making this distinction, nor in helping them identify what it is about a particular threshold concept that makes it troublesome for students. The development of such strategies is the focus of this paper.

Davies (2006) suggests that threshold concepts may be identified and characterised through comparison of the reflections of multiple experts, and comparison of the concepts of experts and students. Another approach is illustrated in two recent studies of threshold concepts in disciplines closely related to physics. Park and Light (2009) used semi-structured interviews to categorise types and levels of student understanding of atomic structure, aiming to better understand the possible threshold nature of that concept. In an analysis of interviews with undergraduate engineering students, Scheja and Pettersson (2010) found compelling evidence of the threshold nature of both 'limit' and 'integral' in mathematics. Such studies raise the potential value of student interviews to explicate threshold concepts. However, this approach is resource intensive, and so not to be undertaken lightly.

This literature raises two potential ways to clarify the troublesome and path-blocking nature of threshold concepts – a resource intensive approach involving interviews with students, and a less resource intensive approach based on the prior experiences of teachers with student misconceptions in their own fields. This paper provides a comparison of the two approaches in one particular case, in which a group of experienced academics, all teaching in first year physics, were brought together to identify a threshold concept in physics and collaborate on targeted curriculum development.

SELECTING A THRESHOLD CONCEPT

Initially, five physicists from four Australian universities, chosen to sample different university types and different geographical regions, were brought together in a one-day meeting facilitated by educational developers. The physicists were introduced to threshold concepts through the five criteria outlined above, which were illustrated through a concrete, everyday example (ethics). They were then asked to brainstorm key discipline concepts without reference to their possible threshold nature, and subsequently to identify candidates for threshold concept status using Meyer and Land's criteria.

They identified several concepts as fundamental to the discipline, listed in Table 1, where they have been loosely grouped into different categories. We note that more than one third of the concepts (all of those in the two categories related to modelling and observation) are not exclusively characteristic of physics – that is, they are not terms that can be associated primarily with the disciplinary language – and yet they were clearly perceived as crucial elements of the discipline. From this list, the academics selected the concept of *uncertainty* as the potential threshold concept to be used in this project. This was subsequently clarified to be *measurement uncertainty*, to exclude the concept of inherent uncertainty and incompatible observables in quantum mechanics. An analysis of measurement uncertainty in terms of Meyer and Land's criteria suggests it is a suitable candidate.

Table 1: Key concepts emerging from the physics brainstorm

Key concept	Grouping
Field, flux, force, momentum, entropy, impulse, energy, potential, temperature, induction, acceleration, wave-particle duality, conservation laws, space-time, gravity, relativity, equilibrium	Terms used to name key discipline concepts or concept clusters
First principles, diagrams, modelling, vectors, frames of reference, idealisation-reality	Terms related to modelling or the tools used in modelling
Significance, approximation, orders of magnitude, uncertainty, measurement	Terms connected with the act of observation or measurement

1. *Transformative* There are several symptoms of students' behaviour which suggest that a good understanding of the role of measurement uncertainty results in a transformation in a student's thinking. Before a student has grasped the role of uncertainty in measurement, they see the outcome of an experiment as a single number (Buffler, Allie & Lubben, 2001). This means that comparisons are made between the values x_1 and x_2 rather than say $x_1 \pm \sigma_1$ and $x_2 \pm \sigma_2$ (where σ_1 and σ_2 are measures of uncertainty on x_1 and x_2 respectively). Uncertainty is seen as a mistake – something to be eliminated or remedied, or which indicates an experiment has been performed incorrectly. When graphing data, lines are drawn to connect points rather than to show a trend. Once the threshold has been crossed, students experience a radically revised view of many aspects of measurement, including factors contributing to experimental design, the limitations on experiments and inferences from data, and indeed the very nature of experimental results. Uncertainty is seen as an intrinsic part of the result of a measurement and as essential in assessing the quality of the outcomes of an

experiment; its target magnitude becomes something that is an important criterion in the design of an experiment; and extrapolations/interpretations are made taking uncertainties into account.

2. *Integrative* The concept of measurement uncertainty integrates a range of concepts and skills in a way that makes more meaning out of the whole. Concepts such as random and systematic error, calibration, repetition, hypothesis testing, significance, tolerances, populations and samples, experimental design, the limits of what can be discovered, interpolation and extrapolation, modelling, approximation and more are brought together in a complex cluster to form a key element of scientific method.

3. *Irreversible* Once a student has grasped the role of uncertainty, their views of the interdependency of theory, experiment and data are irreversibly changed. Their ways of reading data change so that, for example, they distinguish between scatter and pattern, and they recognise all data as contestable.

4. *Bounded or boundary-making* Of the five criteria, the idea that threshold concepts help to define the boundaries of the discipline to which they belong is often the most difficult to interpret, and indeed the physics academics in this project decided to discard it as a criterion. Measurement uncertainty certainly does not serve to demarcate the discipline boundaries of physics from other sciences, instead comprising a cluster of ideas, capabilities and concepts shared by statistics, the sciences and many social science disciplines. It could, however, be seen as one of the ‘boundary’ concepts of quantitative, empirical studies.

5. *Troublesome* There is no doubt that uncertainty is frequently a troublesome concept for students to grasp. The mathematical formalism is non-trivial; the idea of quantifying something that by definition you are unsure of and cannot directly measure is deeply challenging; and learning to ‘read’ data is something that takes practice. In addition, the challenge to the idea of a ‘true’ or ‘exact’ value is often at odds with the definite language of theory (and hence lectures and textbooks). The realisation that data (upon which theories depend) are inherently uncertain, and that the process of measurement is imperfect, leads students naturally and compellingly to question the basis of physical knowledge. This can be deeply unsettling for students who crave clarity and certainty.

A key reason why the academics chose measurement uncertainty as the subject of study (rather than previously identified threshold concepts such as gravity or entropy) was the importance they felt it should be accorded in the development of disciplinary thinking. An understanding of measurement uncertainty – that is, an understanding of how to identify different sources of uncertainty, quantify their effects, take those effects into account in planning experiments, analysing data and making logical inferences from those data, and an appreciation of the consequences of uncertainty – is one of the core characteristics and capabilities of an effective scientist. This is particularly true in physics, where many experiments are aimed at making precise quantitative measurements, and many theoretical predictions are expressed as numbers. However, traditional physics curricula often relegate uncertainties to the realm of the lab alone, or perhaps augment lab experiences with one or two supporting lectures or tutorials. Unfortunately, these activities rarely seem to engage students’ deeper learning. Indeed, for many students, the process of identifying, quantifying and propagating uncertainties is a tedious and occasionally mystifying chore, and a distraction from the real business of getting a result. More explicit pedagogical interventions may be required to provide students with structured opportunities to acquire this threshold concept.

UNCOVERING STUDENT CONCEPTIONS OF UNCERTAINTY

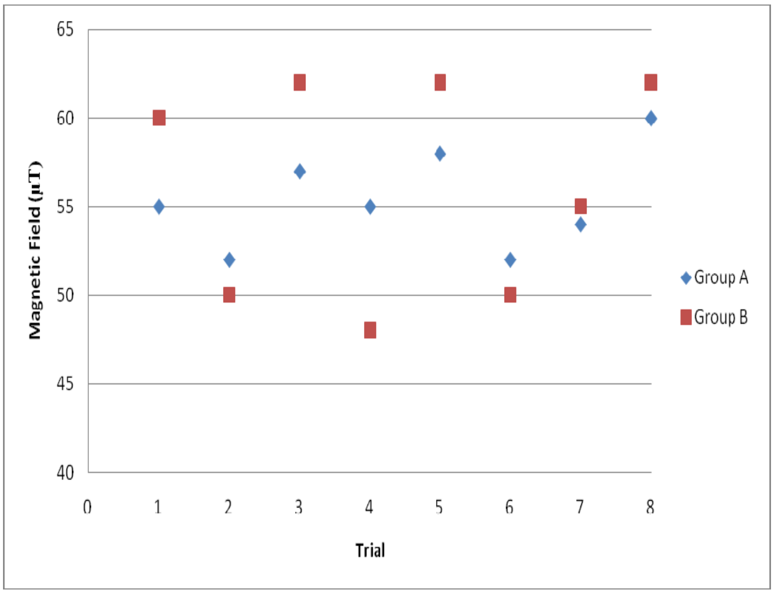
In line with Davies’s suggestion, the lecturers first tried to use their own experiences and expertise to describe common student (mis)understandings of uncertainty. Their expectations, emerging from discussions at the initial one-day meeting, are shown in Table 2.

Table 2: Physics teachers’ expectations regarding student conceptions of uncertainty

Stage 1	No conception of uncertainty; no thought of it in relation to experimental outcomes
Stage 2	Uncertainty is seen as mistakes, errors
Stage 3	Uncertainty is seen as a means of quantifying how wrong you are
Stage 4	Uncertainty is understood as something that must be planned for
Stage 5	Uncertainty is a comprehensible, modelable, quantifiable, communicable result

The next step in the project was to investigate students' actual understandings of measurement uncertainty. Through a series of web-based discussions and video conferences, the team used the outcomes of their discussions at the initial meeting and existing literature on teaching measurement uncertainty, particularly the work of Andy Buffler (Buffler et al., 2001, 2008), to design two scenarios to present to students for discussion in interviews. Scenario A (Figure 1) used a physics context (measuring the earth's gravitational field) but only minimal description of the experiment, while scenario B (Figure 2) used a daily-life context (mobile phone battery lifetimes) and more experimental detail.

You go to a Magnetic Observatory where scientists are making sensitive measurements of the Earth's magnetic field and they wish to compare these measurements with theories about the composition of the Earth.
You go into a laboratory where two groups of scientists (group A and group B) are each busy with their own experiment to measure the magnetic field in the laboratory on that day. The table and graph below show the data gathered by each group.

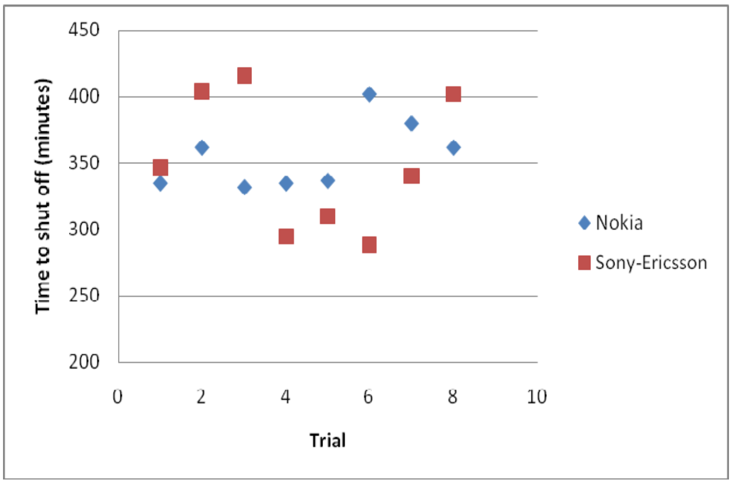


Trial	Group A Magnetic field (µT)	Group B Magnetic field (µT)
1	55	60
2	52	50
3	57	62
4	55	48
5	58	62
6	52	50
7	54	55
8	60	62
average	55.4	56.1

Figure 1: Scenario A, Magnetic Field Measurements

Semi-structured interviews were carried out with 24 first year students (6 randomly-selected from each institution). During the interviews, Scenario A was presented first at two institutions, while Scenario B was presented first at the other two. In both cases, students were initially asked to comment generally on the data. With Scenario A, students were asked to comment on whether one group's measurements were "better' than the others; with Scenario B, students were asked to say which brand they would recommend to a friend. They were also asked if there was a true value of battery life or B field. Finally, they were asked for examples of situations where uncertainty might matter. At all stages, the students were asked to explain their answers, and interviewers followed up on concepts and difficulties they raised.

Battery life is a big factor that customers take into account when deciding which mobile phone to purchase.
A consumer group wishes to advise potential purchasers of competing Sony-Ericsson and Nokia phones about the battery life of phones available from each manufacturer.
The consumer group devise a standard test in which each phone is initially charged fully. For each phone, the display brightness is set to 50% and backlight is illuminated for 10s. used
A landline is called with each phone. The call continues until the battery drains, and the phone shuts itself off.
The consumer group test the Sony Ericsson K610i and the Nokia N72.
The time for each phone to shut off is shown in the table and graph below (each mobile phone is recharged between trials).



Trial	Nokia time to shut off (minutes)	Sony-Ericsson time to shut off (minutes)
1	335	346
2	362	404
3	332	416
4	335	295
5	337	310
6	402	288
7	380	340
8	362	402
average	355.6	350.1

Figure 2: Scenario B, Mobile phone battery lifetimes

PRELIMINARY RESULTS

The interviews were analysed phenomenographically (Åkerlind, 2005), with the aim of identifying variation in the ways in which students understand the concept of measurement uncertainty. This involved four physics lecturers and two educational developers experienced in phenomenographic analysis comparing and contrasting each of the interview transcripts in a search for key similarities and differences. The similarities enabled identification of common aspects of students' understanding of uncertainty, whilst the differences highlighted variation in what some students noticed about uncertainty that others were unaware of. This analysis suggests that there are three distinct aspects to students' understandings of uncertainty – a **pattern**-recognition element that allows students to distinguish between trends, noise and potential anomalies; a **formal**, procedural understanding that allows them to quantify and combine different elements of uncertainty; and a "**meaning**" element that invests uncertainty with a communicable meaning that has implications beyond the given data. A sophisticated understanding of uncertainty involved the integration of all three aspects, whilst a less sophisticated understanding emphasised only one or two of these aspects.

Examples of student comments that focus on the pattern aspect include *"Group A's data is more constant than Group B's," "it goes up and then down and then up again. That is the first thing I noticed." "… it looks like a parabola … actually I take back the parabola thing. It looks a lot more like a sine wave."* Some students explicitly made a connection between scatter and error, e.g. *"It seems like there is a lot of error because it is not close. There is a wide change here… if there was less error the data would be closer together …"*

The difficulty that many students had in quantifying uncertainty (the formal aspect) is evident from responses such as *"This is the worst part about physics, it's working out that absolute or relative error, yeah, what is the error?…It's either half the smallest measurement that you're using…. Or else you do all these weird, complicated equations…," "…that is how you calculate uncertainty by doing the* *equation of maximum value over the number of trials,"* and again, *"Minus one something from something divided by something?"*

An awareness of the meaning of uncertainty was most often evident in the students' suggestions for causes of scatter or for situations where uncertainty might matter, such as this response from a student doing biomedical sciences: *"…if I go on to do medicine, and I give someone morphine, two milligrams of morphine, you can say make you nice and happy, and the syringe is 10 mills plus or minus half a mill, that half a mill is going to matter. Because if I give them too little, it mightn't do anything to them; and if I give them too much, it might just kill them,"* and in this more prosaic but highly practical suggestion, *"They should have uncertainty in their bus timetables."*

The interviews suggest that it is the integration of these three elements into a coherent, interacting whole that characterises the threshold aspect of coming to understand measurement uncertainty. Most students recognised differences in spread in the data, and most were willing and able to suggest experimental and/or environmental factors contributing to the scatter and to differences between data sets. Most were also reluctant to imagine a 'true value,' instead using the term as a kind of short hand for an accepted or published value that they might compare results to. However, very few students were able to quantify their sense of the scatter, or employed disciplinary language such as mean or standard deviation when describing the data. This lack of integration is somewhat different from the stages in student understanding proposed by the physics academics, suggesting that the less resource intensive way of identifying threshold concepts, while potentially powerful, may miss some of the ways in which threshold concepts present barriers to conceptual change and development of disciplinary ways of thinking. Thus although both processes for analysing threshold concepts were valuable, in this study, the student interviews provided additional insights to guide subsequent pedagogical interventions.

The next stage of the project (not described here) uses the understanding of student conceptions that resulted from the interviews to design different curriculum interventions aimed at improving student learning of the threshold concept. A common 'post-test' was implemented in each institution to assess the effectiveness of the interventions. Further details of the interview analysis and the results of the pedagogical interventions will form the focus of future papers and will be available at http://www.thresholdvariation.edu.au from late 2010.

ACKNOWLEDGEMENTS

Support for this publication has been provided by the Australian Learning and Teaching Council Ltd, an initiative of the Australian Government Department of Education, Employment and Workplace Relations. The views expressed in this publication do not necessarily reflect the views of the Australian Learning and Teaching Council.

REFERENCES

Åkerlind, G.S. (2005) Variation and commonality in phenomenographic research methods, HERD 24, 321-334.

Buffler, A., Allie, S. and Lubben, F. (2001) The development of first year physics students' ideas about measurement in terms of point and set paradigms. Int. J. Sci. Ed. 23, 1137-1146

Buffler, A., Allie, S. and Lubben, F. (2008) Teaching measurement and uncertainty the GUM way. The Physics Teacher 46, 539-543

Davies, P. (2006) Threshold concepts: how can we recognize them? In J.H.F. Meyer and R. Land (eds.), Overcoming barriers to student understanding: Threshold concepts and troublesome knowledge (pp 70-84). New York: Routledge

Meyer, J.H.F. and Land, R. (2005) Threshold concepts and troublesome knowledge (2): Epistemological considerations and a conceptual framework for teaching and learning. Higher Ed. 49, 373-388

Meyer, J.H.F. and Land, R. (2006) Overcoming barriers to student understanding: Threshold concepts and troublesome knowledge. New York: Routledge

Park, E.J., and Light, G. (2009) Identifying Atomic Structure as a Threshold Concept: Student mental models and troublesomeness, Int. J. Sci. Ed, 31, 233-258

Perkins, D. Constructivism and troublesome knowledge (2006). In: Meyer, J. H. F. & Land, R. (eds). Overcoming Barriers to Student Understanding: Threshold Concepts and Troublesome Knowledge. Abingdon: Routledge, 33–47.

Scheja, M. and Pettersson, K. (2010) Transformation and contextualisation: conceptualising students' conceptual understandings of threshold concepts in calculus. Higher Ed. 59, 221-241

PROCEEDINGS OF THE 16TH UNISERVE SCIENCE ANNUAL CONFERENCE
SEPTEMBER 29TH- OCTOBER 1ST 2010
THE UNIVERSITY OF SYDNEY

CREATING ACTIVE MINDS IN OUR SCIENCE AND MATHEMATICS STUDENTS

Published by: UniServe Science, Carslaw Building (F07), The University of Sydney, NSW 2006
http://sydney.edu.au/science/uniserve_science/
ISBN: 978-0-9808597-0-6

PREFACE

Dr Alexandra Hugman has changed careers and will not join us in 2010. We wish her the best in the new direction she is pursuing. We continue her vision of including science and mathematics teacher educators in the Uniserve Science Conference. We look forward to the continued presence of Associate Professors Mary Peat and Ian Johnston.

With this year's Proceedings we begin a new tradition: that of Open Source publishing. Furthermore, the submission, reviewing and publication processes have all been handled by an online system. We would like to thank all involved for their patience as we have worked through teething problems. The system allows incredible benefits in terms of tracking and we look forward to a smoother flow in 2011.

A series of papers from the 2009 UniServe Science Conference have been published in the International Journal of Innovation in Science and Mathematics Education (IJISME) http://escholarship.usyd.edu.au/journals/index.php/CAL. We invite authors and all attendees to seriously consider publishing their work with IJISME.

This Proceeding owes its existence to the Editorial and Review Panel listed below who volunteer their own time and expertise to help improve the quality of the publication.

EDITORIAL AND REVIEW PANEL

The Proceedings of the 16th UniServe Science Conference contains three types of papers.

- **Full Refereed Papers** which have been peer reviewed by two independent experts and satisfy the Australian DEST E1 category.
- **Full Written Papers (non-refereed)** which have been subject to editorial assessment and satisfy the Australian DEST E2 category.
- **Abstracts (extract of paper)** which have been subject to editorial assessment and satisfy the Australian DEST E3 category.

We look forward to seeing you at the 16th UniServe Science Conference.

Editor-in-Chief
Associate Professor Manjula Sharma

TABLE OF CONTENTS

KEYNOTE PRESENTATIONS (REFEREED)

REFEREED PAPERS

NON-REFEREED PAPERS

ABSTRACTS