Education Research and Perspectives
An international journal

Volume 41, 2014, Pages 217-231


# Identifying Threshold Concepts in First-Year Statistics

R. Nazim Khan†

*The University of Western Australia*

Every discipline has a limited number of fundamental concepts. Discipline experts will identify some of these as threshold concepts. An understanding of threshold concepts is expected to provide an insight into the discipline and enhance learning and understanding. In this paper I will investigate threshold concepts for first-year business statistics, which is a compulsory unit for all business majors in the School of Business at The University of Western Australia. The investigation will be based on analysis of examination scripts. As this is an initial investigation, I will focus on identifying threshold concepts in one topic only.

## Introduction

Threshold concepts were introduced by Meyer and Land in 2003 (Myer and Land, 2003), based on their experiences in economics education. They followed this with another paper (Myer and Land, 2005) in which they discussed epistemological aspects of threshold concepts. Meyer and Land (Myer and Land, 2005) identified five key characteristics of threshold concepts.

1. Transformative: once it is understood it leads to a much improved understanding and perception of the subject.
2. Probably irreversible: the concept and ensuing change in perspective is unlikely to be quickly or easily forgotten.
3. Integrative: it leads to an integration of ideas within the subject.
4. Expansive: it has boundaries with other threshold concepts, and in certain instances with other disciplines.

† Address for correspondence: Dr. R. Nazim Khan, School of Mathematics and Statistics, The University of Western Australia (M019), 35 Stirling Highway, Crawley, Western Australia, 6009. Email: Nazim.Khan@uwa.edu.au.

5. Troublesome: students have difficulty grasping or understanding the concept.

Since then threshold concepts have found resonance in higher education and the literature has grown rapidly. Several other researchers have further investigated threshold concepts. Cousin (Cousin, 2010) explained that threshold concepts requires the academic to go more deeply into his/her own discipline with the aim of determining the best way of teaching and learning it. Barradell and Kennedy-Jones (Barradell and Kennedy-Jones, 2013) discussed the integration of threshold concepts with student learning and the curriculum.

Threshold concepts are important due to their impact on teaching and learning. A common issue in any course is the volume of material to teach. Too much content may encourage a superficial approach to learning by the students (Ramsden, 2003). An important consideration is that more often high content volume is a result of what the teacher wants to include and not what the student needs to know. Threshold concepts offer a rigorous approach to streamlining what is taught, providing value for both teachers and students. The definition of threshold concepts implies a focus on key aspects, devoting more time and resources to these. Cousin (Cousin, 2006) refers to this as a `less is more' approach to designing a curriculum. The lack of breadth is compensated for by insights gained through deeper understanding of the threshold concepts, allowing students to extend their learning themselves. Finally, a student-focused approach to teaching is generally considered to be more effective for learning (Prosser and Trigwell, 1999). Threshold concepts require teachers to reflect deeply on three aspects: what is to be taught, why it is to be taught and how and when it is to be taught (Barradell, 2013). In addition, the salient, underlying and important links between topics is also revealed in this approach and this can guide curriculum design.

Very little research has been conducted on threshold concepts in mathematics and statistics. This is perhaps somewhat surprising as mathematics and statistics are based so fundamentally around concepts. Cope and Byrne (Cope and Byrne, 2006) used the example

of confidence intervals to illustrate teaching and learning of threshold concepts. Bulmer et al. (Bulmer et al., 2007) considered threshold concepts in statistics for first year biology students. Their approach was to ask students and tutors to list the concepts they found difficult. Dunne et al. (Dunne et al., 2003) considered randomness or variation, sampling, the central limit theorem, linear regression, Bayes' theorem, interval estimation and hypothesis testing as suitable candidates for threshold concepts. They invoked the collective past experience of statistics teachers for this selection of topics. They also conducted an internet based study of first-year business statistics students in which they asked five questions to assess innovations in their teaching. The first four questions were on simple statistical concepts, while the fifth asked for two lists, one consisting of topics the students found easy and the other topics that students found difficult. Their survey showed that chi-square test was the topic most respondents found difficult.

This paper is on identifying some threshold concepts in STAT1520, a first-year unit in Business Statistics at The University of Western Australia (UWA). This is a compulsory unit for all majors in the Business School, with an enrolment of around 1100 over two semesters. The unit is multi-disciplinary, as the current structure of degree programmes at UWA also admits some science and engineering-science students into this unit. Typical issues with this unit are a lack of student motivation, low class attendance and a general lack of engagement. Around ten percent of students enter this unit with the lowest level of mathematics requirement, and this adds another dimension to the teaching.

Several innovations in teaching this unit have been introduced over the last few semesters, described in Khan (Khan, 2013), which have improved student engagement and participation. I was motivated to investigate threshold concepts with a view to further improving student learning, and making this unit the best possible student experience.

This paper is organised as follows. The method is outlined in the next section. In Section Three I discuss student responses to the selected examination question that highlight student difficulties with

this question and enable the identification of related threshold concepts. Design of course materials based on the identified threshold concepts will be covered in Section Four. I will conclude in Section Five by discussing the consequences of the findings for teaching and learning, and indicate further research.

## Method

While educators consider threshold concepts to be important, there are no agreed methods of identifying them. However, there is general agreement that identifying threshold concepts involves several dimensions. Barradell (Barradell, 2013) contains a good summary of methods used to identify threshold concepts. Methods used by researchers include phenomenographic interviews of students (Davies and Mangan, 2007), questionnaires, student surveys, short answer problems and review of examination scripts (Davies and Mangan, 2007; Holloway et al., 2009). Davies and Mangan (Davies and Mangan, 2007) applied the above methods along with semi-structured interviews of students to identify threshold concepts in Economics. Male and Baillie (Male and Baillie, 2011) used interviews with academics (lecturers and tutors, who are discipline experts), focus groups with students, and workshops with academics and students, separately and together, to identify threshold concepts in Engineering.

Our interest is in identifying threshold concepts in first-year statistics units. While the investigation is based on a business statistics class, the findings are expected to be applicable to general first-year statistics courses since the topics they cover are common and generic in nature.

The approach for this research will have the following steps.

1. Consultation with discipline experts around Australia and New Zealand. Also included will be selected international experts.
2. Analysis of student examination scripts with a view to identifying concepts that students struggle with. Examination scripts for the past six semesters are available.

3. Cognitive interviews of students centred on formative assessments, with a view to understanding their reasoning and misunderstandings. Such interviews will occur during semester.
4. Workshop with tutors.
5. Once the threshold concepts are identified based on the above steps, the entire course will be re-designed, with these concepts at the core.
6. Steps 2 to 5 will be repeated for the re-designed curriculum.

This is the initial study to gain experience and insight into the process before embarking on the larger research project. For the purpose of this paper, a threshold concept is considered to be one that students find troublesome. It is envisaged that concepts identified by this criteria will also include some of the other characteristics of threshold concepts discussed above. Davies and Mangan (Davies and Mangan, 2007) note that the first three characteristics of threshold concepts are interwoven. A concept that integrates ideas within a subject is likely to be transformative. A transformative concept that enhances and deepens understanding is necessarily irreversible. Since there is considerable scaffolding and linking between topics in a unit, threshold concepts in one topic often have boundaries with those in other topics. Such concepts are therefore integrative. Concepts that enhance understanding are necessarily troublesome. This is especially true of mathematics and statistics, which employs abstraction to extend existing knowledge and develop new ideas and concepts. Rodger and Turpin (Rodger and Turpin, 2011) also based their investigation of threshold concepts in occupational therapy on troublesome knowledge, but found that their identified threshold concepts possessed all five characteristics of Meyer and Land (Myer and Land, 2005). Cartensen and Bernhard (Cartensen and Bernhard, 2008) also consider troublesome as the most important characteristic in identifying threshold concepts. The original paper by Meyer and Land (Meyer and Land, 2003) did place a strong emphasis on troublesome knowledge.

Barradell (Barrdell, 2013) contains a good discussion on characteristics of threshold concepts. In particular, it is not clear how

many of the five characteristics a concept should possess to be considered as a threshold concept. Male and Baillie (Male and Ballie, 2011) discuss threshold capability as a complementary concept. Irvine and Carmichael (Irvine and Carmichael, 2009) found that very few identified threshold concepts possessed all the five characteristics of Meyer and Land (Myer and Land, 2005).

This investigation is based on analysis of examination papers. Davies and Mangan (Davies and Mangan, 2007) also analyse examination scripts to identify threshold concepts. The rationale is that the examination questions reflect what the lecturer wants students to understand. Additionally, they also provide evidence of whether understanding of particular concepts translates to a "more wide ranging and integrated" answer to the question (Davies and Mangan, 2007).

In this study, final examination papers for semester 2, 2012, were analysed. The paper had three sections. Section One consisted of ten multiple choice questions, and these were omitted for this investigation. This left four questions each in Section Two and Section Three for consideration. Concepts that students found troublesome were identified as follows.

1. The three questions with the lowest mean percentage marks were identified.
2. Histograms of marks (as a percentage) for each question were produced. The three questions with the most right skewed histograms indicated a preponderance of low marks for the corresponding questions, with fewer marks at the higher end.
3. Finally, it is possible that lower marks correspond to the weaker students. This would cast doubt on whether the question is troublesome. To investigate this issue, scatterplots of the marks for the questions selected in the above two steps against the total examination mark were produced.

In addition, examination scripts were perused to identify common types of misunderstandings or problems.

## Identifying threshold concepts

### *Analysis of examination marks*

The three lowest mean marks were 36%, 39% and 44%, corresponding respectively to Section Two Question 4, and Section Three Questions 3 and 4. An inspection of the histograms of the marks indicated that the same three questions had the most right skewed distributions. Table 1 summarises the information relating to these questions.

*Table 1. Information on questions with the lowest mean marks*

| | **S2 Q4** | **S3 Q3** | **S3 Q4** |
|---|---|---|---|
| **Mean (%)** | 36 | 39 | 44 |
| **Topic** | Simple Regression | Multiple Regression | Chi-square test |

Question 4 in Section One was a short question worth only 4 marks examining one aspect of simple linear regression. As such it will not be considered any further in this investigation. The other two were longer questions worth 20 and 10 marks respectively, covering each of the topics in depth. Scatterplots of marks for these questions against the total marks for the examination indicated that of 271 students who obtained at least 60% in the examination, 50 scored at most 50% in S3 Q3. The corresponding number for S3 Q4 was 113. For this reason S3 Q4 was selected for further investigation. The scatterplot also indicates that this question was difficult for both weak and good students, identifying it as troublesome. Interestingly, this topic of chi square test was also identified by Dunne et al. (Dunne et al., 2003).

One issue that needs to be considered is that this is the last question in the paper and students may not have had time to attempt it. However, perusal of examination scripts showed that this was not the case. There were a total of 580 scripts of which 110 scored a mark of 0 for this question—these students did not attempt the question. As a comparison, 85 students had not attempted S3 Q3 and 240 students

had not attempted S2 Q4. This observation provides some confidence that shortage of time was not a reason for students not completing the chi square question.

Figure 1 shows a barplot of the marks for this question. Only 47% of students obtained a pass mark (50%) in this question, and only 3 students obtained full marks.

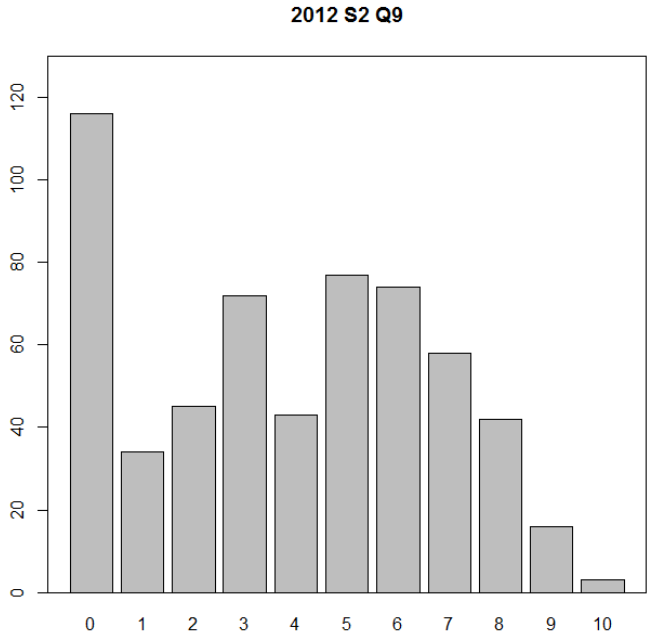


*Figure 1. A barplot of the marks for Question 4, Section 3*

### *Student Responses*

The question covered chi-square test for independence of categorical variables based on observed frequencies and is reproduced in Figure 2 below. The four entries in bold in the Excel output had been deleted and were required to be calculated in order to determine the pattern in orders.

**Question 4 [10 marks]** The owner of a five star restaurant serving Continental style entrées and desserts wants to learn more about the patterns of demand for these items. Data on 630 orders were analysed and the relevant output is given below. Write a report no longer than a page to the manager stating the key findings.

| | | | | | | |
|---|---|---|---|---|---|---|
| | | Chi-square test statistic = | 26.6491 | | Number of: | |
| | | p-value = | 0.0016 | | rows = | 4 |
| | | | | | columns = | 4 |
| | | | | | | |
| | | *Actual frequencies* | | | | |
| | | | **Entrée** | | | |
| | | Beef | Poultry | Fish | Pasta | Totals |
| **Dessert** | Ice Cream | 13 | 8 | 12 | 47 | 80 |
| | Cake | 98 | 12 | 29 | 145 | 284 |
| | Fruit | 8 | 10 | 6 | 26 | 50 |
| | None | 124 | 98 | 149 | 412 | 783 |
| | Totals | 243 | 128 | 196 | 630 | 1197 |
| | | | | | | |
| | | *Expected frequencies* | | | | |
| | | | **Variable B** | | | |
| | | Beef | Poultry | Fish | Pasta | |
| **Variable A** | Ice Cream | 16.2406 | 8.5547 | 13.0994 | 42.1053 | |
| | Cake | **57.6541** | **30.3693** | 46.5029 | 149.4737 | |
| | Fruit | 10.1504 | 5.3467 | 8.1871 | 26.3158 | |
| | None | 158.9549 | 83.7293 | 128.2105 | 412.1053 | |
| | Totals | 243 | 128 | 196 | 630 | |
| NOTE: Expected frequencies should not be less than 5.0 | | | | | | |
| | | | | | | |
| | | *Chi-square calculations* | | | | |
| | | | **Entrée** | | | |
| | | Beef | Poultry | Fish | Pasta | |
| **Dessert** | Ice Cream | 0.6466 | 0.0360 | 0.0923 | 0.5690 | |
| | Cake | **28.2337** | **11.1109** | 6.5878 | 0.1339 | |
| | Fruit | 0.4556 | 4.0498 | 0.5843 | 0.0038 | |
| | None | 7.6867 | 2.4323 | 3.3710 | 0.0000 | |

*Figure 2. The chi-square examination question*

For this question examination scripts were analysed to identify the types of errors made by students. The question examined the following five concepts.

1. Stating the hypotheses of interest (1 mark). This is specific to the topic being tested. All students who attempted this question got this part correct. Hence this is not a threshold concept.

2. Stating the conclusion of the hypothesis test given the p-value (1 mark). This concept is common to all hypothesis testing situations, so it cannot be considered to be a threshold concept specific to this topic.
3. Calculating the two missing expected frequencies (2 marks). The calculations are based on the assumption that the entrée and dessert ordered by a customer are independent. Of the students who attempted this question, 32% did not get full marks for this part.
4. Calculating the two missing chi-square values, which are based on the previous calculations and a simple formula (2 marks). Of the students who attempted this question, 58% did not get full marks for this part.
5. Determining the form of the trend or association in the data, with justification (4 marks). This part requires inspection of the chi-square values, and then comparison of the corresponding observed and expected frequencies. Only 0.6% of students who attempted this question obtained full marks for this part.

## The Threshold Concepts

Based on the above analysis, the following two threshold concepts are identified for this topic.

1. Calculation of expected frequencies. This calculation is based on the concepts of independence of random variables and expectations, both covered earlier in the unit.
2. Determining the form of the trend or association in the data. The trends are determined by first identifying the largest chi-square values, and then comparing the expected and actual frequencies for the corresponding cells.

The first of these concepts is one that students typically find difficult. The concept of independence is taught early in the course in the context of joint probability distributions for discrete random variables. It is expected that students would have forgotten some of this material or not understood it in the first place. Also, independence here is in the context of observed frequencies, and the link between the two contexts may not be clear.

Determining the form of any association between the variables is a higher idea and many students are expected to find this difficult.

The formula for the calculation of the chi-square statistic is not difficult and can be motivated by other more intuitive ideas. Similar formulae have been used in other contexts, and better links can be made between them. However, a simple formula cannot be considered to be a threshold concept.

## Design of Course Materials

While each of these concepts is difficult in its own right, an added confounding factor is that this is the last topic taught. Also, this is treated as a stand-alone topic and does not follow in a logical sequence from the immediate previous topics. However, this topic covers one of the most common types of data that businesses deal with. Therefore it is important to impart to business students a better understanding of the concepts involved.

Currently, joint distributions are taught within the larger topic of random variables. This topic also contains discussion of binomial and Poisson random variables, and hypothesis tests for the binomial proportion (based on sample sizes less than 30) and the Poisson mean. Continuous random variables are also covered within this section. As mentioned earlier, chi square tests are the last topic taught, some six weeks later.

In light of the identified threshold concepts, course materials were redesigned as follows.

1. The correct placement of chi-square test of independence is immediately after discussion of joint random variables. This requires some re-ordering of material. The idea of calculating expected frequencies can be covered more properly when discussing joint distributions of random variables. It can be motivated more easily in light of the idea of independence of random variables.
2. The large chapter on random variables has been divided into three. The first covers concepts of random variables, binomial and Poisson distributions and associated hypotheses tests. This is followed by a chapter on continuous

random variables. Since the chi square is a continuous distribution, hypotheses tests based on this distribution requires the ideas of continuous distributions. Joint distributions are covered next in a separate chapter. Finally, chi square test for contingency tables follows in a chapter dedicated to it. Thus all the necessary ideas for the chi square test are covered in immediately preceding chapters. The chi square tests are given appropriate importance by still being in a separate chapter, but linked closely with preceding material.

3. The concept of investigating patterns of association can perhaps be better motivated and illustrated in extreme cases. Thus, where the expected and actual frequencies are very different, a clearer pattern will be observed. Appropriate examples of this in familiar and intuitive contexts have been developed.

The above redesign of the course materials will be trialled in semester 2 of 2014.

## Discussion and Conclusions

Identifying threshold concepts in a discipline has major implications for teaching and curriculum design. I attempted to identify threshold concepts for chi-square tests of independence in a first-year business statistics unit. Two threshold concepts have been identified for this topic. While these concepts were based on the criteria that students find them troublesome, they are also concepts that are integrative, as they merge ideas from other parts of the course, specifically joint distributions. At the same time they are expansive as they have boundaries with the concepts of independence and joint distributions. In particular, when independence in the joint distribution context is discussed, the problem of testing for independence based on observed data can be pre-empted. In addition, the identified threshold concepts are transformative, as their understanding is expected to lead to a deeper understanding of the topic. Finally, the concepts are probably irreversible, as once the ideas of independence and identifying association are understood they are unlikely to be easily forgotten.

These threshold concepts have been used to inform teaching. In particular, course materials have been re-designed, enabling integration and merging of different sections of the unit by creating links between them. A different breakdown and more appropriate order of topics has also been suggested by the identified threshold concepts. This in itself is quite revealing and perhaps the real benefit of this research. As such, this exercise has been transformative and integrative for the lecturer as well! As far as I am aware, no other statistics course teaches chi-square tests in the context of and immediately after independence of jointly distributed discrete random variables. In particular, most textbooks also treat this as a separate stand-alone topic.

The method used to identify threshold concepts in this paper is the analysis of examination scripts for one semester only. Similar analysis of other examination scripts should be conducted to determine if the results are consistent. The methods section has listed several steps in the identification of threshold concepts. With the experience of this study, a larger project with the aim of identifying threshold concepts for the whole unit will be undertaken.

While this research has focused on business statistics, the findings are expected to be relevant to other first-year statistics courses since they cover similar topics and concepts, and have students with similar backgrounds. In addition, the research method is applicable to mathematics courses as well.

## Acknowledgements

The Faculty Academy for the Scholarship of Education (FASE) workshops organised by Caroline Baillie and conducted by Helen Wildy and Elaine Chapman were the inspiration behind this research. All three deserve my heartfelt thanks. I thank Caroline Baillie for her comments, which greatly improved this paper. Finally, I thank two anonymous referees for their helpful comments.

## References

[1] Barradell, S. (2013). The identification of threshold concepts: a review of theoretical complexities and methodological challenges. *Higher Education* 65, 265-276.

[2] Barradell, S. and Kennedy-Jones, M. (2013). Threshold concepts, student learning and curriculum: making the connections between theory and practice. *Innovations in Education and Teaching International*.
[3] Bulmer, M., O'Brien, M. and Price, S. (2007). Troublesome concepts in statistics: a student perspective on what they are and how to learn them. Uniserve teaching and learning proceedings. http://science.uniserve.edu.au/pubs/procs/2007/06.pdf
[4] Cartensen, A. K. and Bernhard, J. (2008). Threshold concepts and the key to the portal of understanding. In Land, R., Meyer, J. H. F. and Smith, J. (Eds). *Threshold concepts within the disciplines*. 143-154. Rotterdam: Sense Publishers.
[5] Cope, C.J. and Byrne, G. (2006). Improving teaching and learning about threshold concepts: the example of confidence intervals. Symposium on Threshold Concepts, University of Strathclyde, Glasgow, August 29 - September 1.
[6] Cousin, G. (2006). An introduction to threshold concepts. *Planet No. 17, December*. http://neillthew.typepad.com/files/threshold-concepts-1.pdf
[7] Cousin, G. (2010). Neither teacher-centred not student-centred: threshold concepts and research partnerships. *Journal of learning and development in higher education* 2.
[8] Davies, P. and Mangan, J. (2007). Threshold concepts and the integration of understanding economics. *Studies in Higher Education* 32(6), 711-726.
[9] Dunne, T., Low, T. and Ardington, C. (2003). Exploring threshold concepts in basic statistics, using the internet. IASE/ISI Satellite. Retrieved 4 July 2014 from https://www.stat.auckland.ac.nz/~iase/publications/6/Dunne.pdf
[10] Holloway, M., Alpay, E. and Bull, A. (2009). A quantitative approach to identifying threshold concepts in engineering education. Retrieved 4 July 2014 from http://www.heacademy.ac.uk/assets/documents/subjects/engineering/EE2010/101_GP_Holloway.pdf
[11] Irvine, N. and Carmichael, P. (2009). Threshold concepts: A point of focus for practioner research. *Active Learning in Higher Education*. 10(2), 103-119.

[12] Khan, R. N. (2013). Teaching first-year business statistics three ways. *Lighthouse Delta 2013: The 9th Delta Conference on teaching and learning of undergraduate mathematics and statistics*, 24-29 November 2013, Kiama, Australia.
[13] Male, S. A. and Baillie, C. A. (2011). Threshold capabilities: an emerging methodology to locate curricula thresholds. Paper presented at Research in Engineering Education Symposium, Madrid, Spain.
[14] Meyer, J.H.F. and Land, R. (2003). Threshold concepts and troublesome knowledge 1: linkages to the ways of thinking and practising within the disciplines. In *Improving student learning—ten years on.* C. Rust (Ed), OCSLD, Oxford.
[15] 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 Education* 49, 373-388.
[16] Prosser, M. and Trigwell, K. (1999). *Understanding learning and teaching: The experience in higher education*. Buckingham: SRHE ad Open University Press.
[17] Ramsden, P. (2003). *Learning to teach in higher education*. Routledge: London.
[18] Rodger, S. and Turpin, M. (2011). Using threshold concepts to transform entry level curricula. Paper presented at Higher Education Research and Development Society of Australasia, Inc Conference, Gold Coast, Australia.