Operationalising a Threshold Concept in Economics: A Pilot Study Using Multiple Choice Questions on Opportunity Cost

Martin P. Shanahan and Gigi Foster
University of South Australia

Jan H. F. Meyer
University of Durham and University of South Australia

Abstract

This paper addresses the emerging educational framework that envisions threshold concepts as mediators of learning outcomes. While the threshold concepts framework is highly appealing on a theoretical level, few researchers have attempted to measure threshold concept acquisition empirically. Achieving this would open a new arena for exploration and debate in the threshold concepts field, and provide potential results to inform teaching practice. We begin the process of operationalising threshold concepts in economics by attempting to measure students' grasp of the threshold concept of opportunity cost in an introductory economics class. We suggest two potential measures and correlate them with an array of ex ante and ex post variables, including students' expectations of success, prior misconceptions about economics and the work of economists, and actual success in the course. Results cast new light onto the factors that influence the acquisition of threshold concepts, the relationship between threshold concept acquisition and final learning outcomes, and the empirical viability of threshold concepts generally.

Introduction

The notion of threshold concepts has recently been identified as an important lens through which to view discipline-specific ways of transformative thinking (Meyer

and Land, 2003). One metaphor for the idea of the threshold concept is that it is a portal, which, when crossed by the learner, grants access to a previously inaccessible or transformed way of thinking. Meyer and Land (2003) have argued that a threshold concept is likely to be: (i) integrative, in that it exposes the previously hidden interrelatedness of phenomena; (ii) transformative, so that once understood, it produces a significant shift in the perception of the subject; (iii) potentially irreversible, such that once acquired it is likely to permanently alter the student's perspective, potentially to the extent of transfiguring the identity of the student; and (iv) potentially troublesome. More recently, within the discipline of economics, the idea of opportunity cost has been identified as a threshold concept (Shanahan and Meyer, 2006).

The present paper examines variation in students' acquisition of opportunity cost at an individual level. It uses a relatively simple adaptation of a common assessment methodology – multiple choice questions – to examine whether variation in students' acquisition of opportunity cost, as a threshold concept, can be detected and evaluated through such questions. Our aim is to advance the debate on threshold concepts by offering the first empirical evaluation of students' acquisition of opportunity cost. We also exploit a detailed data set to explore the correlation of a variety of student-level factors and apparent academic success with the acquisition of the threshold concept of opportunity cost.

Theoretical foundation: the threshold concept of opportunity cost

Within the discipline of economics, one threshold concept is opportunity cost. This concept, which attempts to shift students' thinking to consider the value of the next best (rejected) alternative when making a choice, is generally taught in the first weeks of most introductory courses. The transformative effect of this concept is outlined in Meyer and Land (2003: 414–415):

> …opportunity cost captures the idea that choices can be compared, and that every choice (including not choosing) means rejecting alternatives. A student who has a good grasp of this concept has moved a long way toward breaking out of a framework of thinking that sees choices as predetermined, or unchangeable. They have also moved toward seeing [at least] 'two sides' of every choice, and in looking beyond immediate consequences, and even just monetary 'costs' towards a more abstract way of thinking [about human behaviour].

Thus to quote Eatwell *et al.* (1998: 719):

> Opportunity cost, the value placed on the rejected option by the chooser, is the obstacle to choice; it is that which must be considered, evaluated and

> ultimately rejected before the preferred option is chosen. Opportunity cost in any particular choice is, of course, influenced by prior choices that have been made, but with respect to this choice itself, opportunity cost is *choice-influencing* rather than *choice-influenced* (emphasis in original).

Thus, if 'accepted' by the individual student as a valid way of interpreting the world, it *fundamentally changes their way of thinking about their own choices, as well as serving as a tool to interpret the choices made by others*.

There has been only limited work that tries to link theory on threshold concepts to concrete measures. One attempt to operationalise the acquisition of understanding about opportunity cost as a threshold concept appears in Shanahan and Meyer (2006). That paper explores the variation among students in opportunity cost understanding in light of the notion of 'trouble- some knowledge', after Perkins (1999), as applied to threshold concepts. The present paper builds on this earlier work in exploring variation in students' acquisition of understanding of opportunity cost via analyses of responses to multiple choice questions. We take the step from theory to empirics by exploiting answers to conventional multiple choice questions to measure the acquisition of understanding about opportunity cost.

It should be noted that 'opportunity cost' as a threshold concept is not generally recognised as an independent entity within the analytical discourse of economics. On the contrary, Davies and Mangan (2007) explicate the existence of a web of interconnected threshold concepts within economics. It remains true, however, that 'opportunity cost' initially 'comes into view' for students as a discrete concept. Davies and Mangan (2006) furthermore argue that, within economics, there may exist personal concepts (economically-oriented perspectives on everyday life) and procedural concepts (ways of practising or articulating economics) that surround the discipline-based threshold concept of opportunity cost. Variation in students' liminality (the term coined in Meyer and Land (2006) to describe variation in the 'threshold state' of acquiring a given discipline-based threshold concept) may be dependent upon their acquisition of these other types of concepts. Failure to articulate an understanding of opportunity cost may therefore be because the necessary supporting concepts (either personal or procedural) have not been sufficiently acquired by students to enable them to acquire, or express, a discipline-based concept. This insight is also potentially consistent with previous work that has revealed a correlation between variation in economic misconceptions and measured student outcomes (Shanahan and Meyer 2003). To the extent that there is a negative association between students' acquisition of threshold concepts and their economic misconceptions, a greater understanding of students' liminality in

threshold concepts may enhance our understanding of the factors underlying students' acquisition of economic misconceptions.

With this caveat in mind, we base our key measures of threshold concept acquisition on information that is often available to teachers in introductory microeconomics: students' answers to multiple choice questions on opportunity cost that appear as part of the normal assessment in the course. At this initial stage of assessing threshold concepts, it was not possible to create, evaluate and apply questions solely designed to detect variation in students' state of liminality (*a la* Meyer and Land, 2006); nor is it yet possible to measure directly students' acquisition of procedural or personal concepts relevant to opportunity cost. Our work represents a first attempt to develop mechanisms for measuring a student's grasp of opportunity cost from the perspective of examining it as a threshold concept.

One criticism of using multiple choice questions in such research is that they are generally not sufficiently sensitive to detect small differences in the degree of acquisition of a concept such as opportunity cost. Two responses follow. First, the sophistication and quality of multiple choice questions is determined by the ingenuity of the assessor. While it is difficult to construct multiple choice questions that assess deeper levels of in-depth understanding (and by implication, acquisition of threshold concepts), it can be done (e.g. Buckles and Siegfried, 2006).[1] Of note, earlier attempts to detect these reasons using students' written work have been hampered by some students' skills in written English (Shanahan and Meyer, 2006). Multiple choice questions at least partly remove this barrier, as they do not require students to compose answers, but merely to recognise the correct alternative.

Second, and importantly, multiple choice questions remain a frequently-used method of assessment in economics, particularly in large courses (Becker and Watts, 2001). Developing multiple choice questions that can detect students' acquisition of threshold concepts can therefore be an important and powerful way of making the theory of conceptual thresholds accessible to teachers, so that ultimately, students in university courses may benefit.

In the present paper, we rescore students' responses to a selection of pre-existing multiple choice questions that are used in normal course assessment and are relevant to opportunity cost (see Appendix A). We then use levels and differences in students' rescored responses to these questions to measure the stock and growth, respectively, of opportunity cost understanding. Our rescoring procedure is an attempt to create measures of threshold concept acquisition that are sensitised to different degrees of student comprehension (liminality) of the basic concept, through the creation of what we hope is meaningful variation. We explore the conditional and unconditional

associations of our measures with students' background characteristics, expectations of success in the course, prior misconceptions about economics and the work of economists, and actual success in the course.

Our approach is clearly limited by the original intent of the questions on which we base our measures of opportunity cost acquisition. For this reason, we anticipate that our final results, while based on variation in student responses, will not be as sensitive as we would like to the type of variation in liminality required to assess individual student's progress in acquiring a threshold concept. Nonetheless, our study represents a first pass at empirically operationalising the notion of threshold concepts in a manner that could make the notion accessible and relevant to teaching practice. Our efforts are intended to highlight the most pressing requirements for such a translation of theory into practice to be successful, and to give rise to suggestions for future work.

Empirical design

The University of South Australia (UniSA) is a public university with approximately 30,000 students and 2000 staff, whose main campuses are located in Adelaide, South Australia. The Division of Business represents approximately one-third of the university. Within this division, all students undertake a core of business-related courses in their first year before studying more specialised programmes. One course (of eight) taught to full-time students in their first year is Economic Principles. This course is taken by most students in their first semester, although it can also be taken in the second year or later in some programmes. In 2005 it was designed as a first-year microeconomics course whose principal text was an Australian adaptation of the introductory American textbook by McConnell.

Face-to-face delivery of this course consists of two hours of lectures and a 1.5 hour tutorial per week. To better service the large number of students enrolled, the same two-hour lecture is presented twice in a given week, and students are assigned to one of those two lectures, as well as to one tutorial. Approximately 700 students typically enroll in the course in the first semester of every academic year, of whom approximately 85% are Australian and 15% are international students. Small group work is emphasised in tutorials, with typically 2–4 people per group and the total number of students per tutorial around 25. During the semester in which we collected our data, there were 30 tutorial groups, taught by 11 different tutors, and somewhere between 400 and 700 students observed to be participating in the course at any given point in time. A more detailed discussion on attendance and attrition appears later.

There are several components of formative and summative assessment in the course. Individually, students must respond (for credit toward their final course mark) to a series of statements referred to as the 'Reflections on Learning Inventory (RoLI)' which is designed to help students become aware of and improve their metalearning capacity (Shanahan and Meyer, 2003). One element of this inventory includes three scales, each based on five statements taken from previous students, about what they think economics is, their perceptions about the role of market forces in determining prices, and what they believe economists do. These statements are designed to detect variation in students' perceptions of these three economics-specific conceptual dimensions. Students respond to each statement on a five-point Likert scale and their responses are tallied. These responses form the basis of three measures in this paper: 'economic misconceptions,' 'market forces,' and 'perceptions of economists'. (A more detailed analysis of these scales, their meaning and their statistical reliability is presented in Shanahan and Meyer (2001a) and Meyer and Shanahan (2002).) A number of other questions regarding students' background and preferences are also included on the RoLI.

At the time they complete the RoLI, students also must individually complete an on-line questionnaire, which consists of 40 multiple choice questions assessing students' understanding of economics: the 'Test of Economic Literacy' (or TEL). Versions of this test have been used elsewhere to measure economics understanding (Beck and Krumm, 1991; McKenna, 1994; Whitehead and Halil, 1991).[2] Of particular concern with the use of this instrument for our purposes is that the TEL, while it has construct validity in assessing core concepts, was not constructed to assess students' grasp of threshold concepts. Two responses to this concern follow. First, given the content- and discipline-specific context of first-year economics, there is likely to be a good deal of overlap between the theoretical properties of opportunity cost as a threshold concept and the content of the TEL questions regarding opportunity cost. As a first approximation, students' responses to the TEL questions on this concept, as these questions are so basic and general, are likely to capture some of the variation in students' acquisition of opportunity cost as a threshold concept. Second, questions designed specifically to detect dimensions of threshold concept acquisition for individual threshold concepts remain to be written. Given the measurement error resulting from the fact that the questions posed in the TEL may not map perfectly into the theoretical definition of what it means to have acquired the threshold concept of opportunity cost, our results may be biased towards zero. As our study is a pilot project, and as no alternative presently exists to measure students' acquisition of the concept of opportunity cost in the field, we believe it is still valuable to proceed using a non-purpose-built instrument such as the TEL.

The RoLI and TEL are both completed in the first two weeks and again in the final two weeks of the 13-week semester. There are two additional pieces of continuous assessment that students complete individually (take-home assignments) and one method of groupwork- based continuous assessment (weekly quizzes in tutorials). The final piece of assessment is an end-of-semester, closed-book examination.

Data

In our data, we observe the following variables. First, students' scores on both iterations of the TEL are available, broken down by question. As we discuss below, we use the overall score on the TEL; a partial score that excludes those questions relating to opportunity cost; and a partial score including only those questions that relate to opportunity cost.

From both iterations of the RoLI, we observe the following: students' responses regarding prior misconceptions of economics and economists' work; the strength of their views about markets; how well they expect to do in the course; whether English is their first language; whether any family members went to university; the highest level of mathematics they have undertaken previously; whether they have studied economics elsewhere; and whether they like mathematics.

From course records, we observe students' tutorial quiz answers for five of the 13 weeks of the semester; their final exam scores (broken down by section); and the final percentage they earned in the course. The last is a weighted sum of their scores on all the assessment items, as follows: RoLI completion, 5% (2.5% towards students' marks were given upon completion of each iteration of the RoLI, regardless of the answers provided); tutorial quizzes, 15%; individual assignment 1, 10%; individual assignment 2, 10%; and final exam, 60%. We also have information on the tutor, the size of tutorial, the day of the week and time of day that the tutorial was held, and which of the two lectures the student attended.

We also observe extensive student-level background information that is drawn from UniSA's student record files, including birth date, sex, citizenship status, method of admission to university (e.g. direct from high school, transfer from another institution, etc.), course load, part-time or full-time student status, home state, and university programme (e.g. accountancy, management, international business, etc.).[3]

Column 1 of Table 1 (Panel A) shows the number of students in successively restricted subsamples of our data set. Given the fluidity of student numbers in the first few weeks of the semester caused primarily by late or initially erroneous enrolments, we use two early enrolment measures as indicators of initial student

Table 1 Sample sizes

Criterion[a]	**(1)**	**(2)**
Panel A	N	Percent represented in main analysis sample
Took first Test of Economic Literacy (TEL)	778	55%
and enrolled at March 7	608	70%
and attended week 3 tutorial	497	75%
and attended week 6 tutorial	422	83%
and attended week 8 tutorial	361	86%
and attended week 9 tutorial	308	91%
and attended week 12 tutorial	263	94%
and took second Test of Economic Literacy (TEL)	248	100%
and sat final exam	244	100%
Sat final exam (but skipped second TEL)[b]	256	95%
Panel B	N	Percent represented in main analysis sample
Took first Test of Economic Literary (TEL)	778	55%
and was enrolled at March 7	608	70%
and attended any tutorial	593	72%
and took second Test of Economic Literacy (TEL)	430	99%
and sat final exam	414	99%
Sat final exam (but skipped second TEL)[c]	473	87%

[a] Sample selection criteria are applied successively moving down the column in each panel to generate the sample sizes given in Column 1 of each panel. Each entry in Column 2 shows the percent of students from the subsample shown in that row who are represented in our main analysis sample (the sample used to create the first two columns of Table 5).

[b] Students in this row were present at all points in the course up to and including the week 12 tutorial, but need not have taken the second TEL.

[c] Students in this row need not have taken the second TEL.

numbers: the number of students who completed the first iteration of the TEL/RoLI, and the number of students who were listed when we downloaded, from university records, a snapshot of all students enrolled on 7 March (the beginning of week 2 of the course). The amount of attrition throughout the semester from active participation in the course is high. Of the 778 students who complete the first TEL, only 244 appear at all key points that we observe during the semester, i.e. are

Table 2 Summary statistics: continuous variables

Variable[a]	Mean	Standard Deviation	Min	Max
OC difference: TEL	0.10	1.42	–5	4
OC level: Exam	2.79	0.97	0	4
OC level: First TEL	3.58	1.29	0	5
OC level: Last TEL	3.68	1.23	0	5
Initial full TEL score	24.60	6.35	7	37
Final percent in course	51.64	11.94	9	85
Raw exam score	32.92	12.80	0	66
Expectations of success	2.50	0.78	0	4
Attendance at tutes	4.39	0.84	1	5
Age	22.10	5.62	17.47	58.27

a All summary statistics are reported for the main analysis sample of 431 students. The variable *OC difference: TEL* is the difference between a student's scored responses to the questions regarding opportunity cost in the last versus first iteration of the TEL; *OC level: Exam* is the student's combined scored responses to the questions on the final exam regarding opportunity cost; *OC level: First TEL* is the student's combined scored responses to the questions on the first iteration of the TEL regarding opportunity cost; and *OC level: Last TEL* is the student's combined scored responses to the questions on the second iteration of the TEL regarding opportunity cost. For most variables, the sample size is 431. For the variables *Final Percent, Raw Exam Score*, and *OC level: Exam*, the sample size is 426.

enrolled in the second week of the course, participate in all five observed tutorials, complete the second TEL, and sit the final exam. Panel B of Table 1, constructed in the same way as Panel A but measuring tutorial participation as attendance at any of the five tutorials we observe, shows that just under half of the students who begin the course fail to appear at all of these selected key points throughout the semester.

Tables 2 and 3 show univariate statistics for the key continuous and indicator variables we use in the paper. In addition to means and variances, Table 2 shows the minimum and maximum values of a number of variables used in the study. The variable 'OC difference: TEL' (shown in row 1, and defined as the difference in opportunity cost understanding between the start and end of the course, as measured via the two iterations of the TEL) also shows the effect of the rescaling of 'almost right' and 'very wrong' answers to multiple choice questions, as discussed in Appendix A. Table 2 also reveals that, on average, students passed both the initial full TEL (with an average score of 24.60 out of a possible 40) and the course (with 51.64 out of 100), though not the final exam (on which the average pre-mediated

Table 3 Summary statistics: indicator variables

Variable[a]	**Percent where *var* = 1 (`yes')**
Has anyone in your family previously gone to uni?	66.36%
Is English your first language?	69.84%
Have you studied economics elsewhere?	42.00%
Do you like mathematics?	52.67%
Did you take Specialist Mathematics?[b]	3.02%
International student	17.17%
Monday tutorial	73.78%
Morning tutorial	60.09%
Female	52.67%
Self-selected into small groups	50.58%

[a] All percentages are reported across the main analysis sample of 431 students. The first five variables are taken from the initial RoLI; the remainder are gathered from course records and university data banks. The variable *Self-selected into small groups* is 1 if the student was in a tutorial where small workgroups were chosen by students, and 0 if the student was in a tutorial where small workgroups were randomly assigned. For all variables, the sample size is 431.

[b] This refers to the final year of high school. Students may also take other mathematics courses, all of which are less advanced than Specialist Mathematics.

score was only 33 out of a possible 80 points). On average, students attended four of the five tutorials where their quiz scores were collected for assessment, and on average, students were about 22 years of age.

Table 3 reflects the equity and access mission of the university, with around one third of students reporting to be the first in the their families to attend university, and also roughly one third with a language other than English as their mother tongue. While most students had not previously studied economics, about half claimed to 'like' mathematics (although only 3% had done the highest level of maths possible at school). The large percentage of students in the sample attending Monday tutorials is, in large part, a reflection of the distribution of tutorials through the week, rather than a genuine oversampling of this subgroup.[4]

Proposed measures of threshold concept acquisition

We are fortunate that three questions on the TEL, and two questions on the final exam in this course, are designed to measure (more or less directly) a student's comprehension of the concept of opportunity cost. (See Appendix A for each exact question and its possible answers.) In order to increase the sensitivity with which

these three questions revealed students' understanding, we assigned scores to each alternative response, rather than marking alternatives as simply right or wrong. The rescoring methodology we used is documented in Appendix A.

Our key outcome variables are constructed from these ingredients and are measured in both levels and differences. The first outcome measure we use, and the one on which we focus most heavily in this paper, is the difference between the sum of students' coded responses to the three opportunity cost questions on the TEL taken at the end of the course and the sum of their coded responses to the same questions on the TEL taken at the start of the course. Given the coding algorithm described in Appendix A, this outcome measure has a potential range of –5 to +5 (but an actual range of –5 to +4 in the data, as can be seen from Table 2). We also use, as a secondary outcome measure, the final level of students' opportunity cost understanding, as measured by the sum of students' coded responses to the two questions on opportunity cost included in the final examination.[5]

We also use several variables which reflect students' degree of 'official' success in the course. The first is students' final percentage in the course; the second is students' raw score in the final examination; and the third is an array of three variables corresponding to students' performance on each of the three sections of the final examination.

Because some students did not take the final exam or did not complete either or both of the two iterations of the TEL, as evident from the declining numbers in the first column of Table 1, we do not have both outcome measures for all students in the sample. The second two columns of Table 1 show what percentage of the subsample of students in each row is represented in our main analysis sample (which is the sample of 431 students for whom we observe all variables necessary to run the regressions reported in Columns 1 and 2 of Table 5).[6]

The vast majority of students we cannot use in our main analyses are excluded due to missing outcomes rather than missing covariates. Two-sample t-tests on the characteristics of students who took both the first iteration of the TEL and the final exam, compared to the sample of students who took the first iteration of the TEL but failed to take the final exam, indicate not surprisingly that those students who persevere (and stay in our sample) are more committed to university. Exam takers are disproportionately full-time students, have on average higher course loads, attend class far more regularly, tend disproportionately to have come to university through traditional means (following from high school) rather than through other pathways, and are younger than those who drop out of the sample between the first TEL and the final examination. They are also disproportionately female. Regressing an indicator for whether a student took the final exam against those

Table 4 Correlations: sample size = 431

	A	B	C	D	E	F	G	H	I
A OC difference: TEL[a]	1								
B Final percent in course	0.01	1							
C Raw exam score	–0.00	0.96*	1						
D OC level: Exam	0.03	0.45*	0.48*	1					
E OC level: First TEL	–0.59*	0.26*	0.26*	0.24*	1				
F Initial full TEL score	–0.10*	0.36*	0.39*	0.30*	0.59*	1			
G Attendance at tutes	–0.06	0.31*	0.26*	0.11*	0.08	0.00	1		
H Expectations of success	–0.10*	0.23*	0.21*	0.11*	0.13*	0.20*	0.09	1	
I OC level: Last TEL	0.53*	0.28*	0.27*	0.28*	0.37*	0.51*	0.02	0.02	1

[a] Correlations significant at the 95% level are starred. OC difference: TEL is the outcome measure used in the regressions reported in the first two columns of Table 5. See text or previous tables for further variable descriptions.

student-level observable characteristics available for the bulk of the sample, using 774 of the 778 students who took the first TEL, shows that the following characteristics are significantly related to perseverance: original TEL score (positively); having taken specialist mathematics (negatively); having at least one family member who previously attended university (positively); course load (positively); South Australian prime residence (positively); and international student status (positively). Overall, this analysis indicates to us that the sample of students we are able to analyse is composed disproportionately of those who are prepared for and/or committed to success in the course.

Unconditional relationships among variables

Before turning to our regression results, it is worthwhile to consider some of the raw correlations evident in our data. Table 4 presents pairwise correlations amongst a set of key outcome and control variables for all students in our main analysis sample.

The high correlation between the raw exam score and final percent mark in the course is as expected, as are most of the other strong, positive correlations amongst success measures in the course. There is a significant negative correlation between scores on the opportunity cost questions of the first TEL ('OC level: First TEL') and the difference in scores on the opportunity cost questions across the two TELs ('OC difference: TEL'), and similarly between scores on the opportunity cost questions on the last TEL ('OC level: Last TEL') and the difference in scores on the opportunity cost questions across the two TELs. These two negative correlations are both expected,

given the fact that the 'OC difference: TEL' variable is a simple difference of two bounded levels.[7] Finally, attending tutorials and having higher expectations of success in the course are both associated with higher levels of final performance in the course (although they are *not* associated with higher opportunity cost acquisition, as measured by the change in opportunity cost understanding over the course of the semester).

Conditional estimation results

With some understanding of the raw variances and covariances present in our variables, we now turn to the prediction of several outcomes: the growth in students' opportunity cost understanding; the endpoint of students' opportunity cost understanding; and students' success in the course generally.

Table 5 Regression results: predicting changes and levels of opportunity cost understanding

Variable[a]	**Outcome: OC difference: TEL**		**Outcome: OC level: Exam**	**Outcome: OC level: Last TEL**
	(1)	(2)	(3)	(4)
OC level: First TEL	**–0.87****	**–0.86****	–	–
	(0.05)	(0.05)		
OC difference: TEL	–	–	**0.05**	**0.49****
			(0.03)	(0.03)
First TEL (or portion)	**0.08****	**0.09****	**0.05****	**0.11****
	(0.01)	(0.01)	(0.01)	(0.01)
Miscon. – Economics	**–0.03**	**0.02**	**–0.03**	**–0.01**
	(0.02)	(0.03)	(0.02)	(0.02)
Percept. – Economists	**–0.01**	**–0.04**	**0.05***	**–0.01**
	(0.02)	(0.03)	(0.03)	(0.02)
Market forces	**0.01**	**0.01**	**–0.03**	**0.02**
	(0.03)	(0.03)	(0.02)	(0.02)
Took specialist maths	**0.67**	**0.66**	**0.49**	**0.75***
	(0.36)	(0.36)	(0.33)	(0.30)
Family at Uni	**–0.01**	**–0.00**	**–0.01**	**0.08**
	(0.11)	(0.11)	(0.10)	(0.09)
Enjoy maths	**0.26***	**0.27***	**–0.14**	**0.06**
	(0.12)	(0.12)	(0.10)	(0.09)
Expectations	**–0.14**	**–0.13**	**0.11**	**–0.06**
	(0.07)	(0.07)	(0.07)	(0.05)
English first lang.	**–0.21**	**–0.19**	**–0.01**	**–0.22**
	(0.16)	(0.16)	(0.13)	(0.12)

Table 5 (continued) Regression results: predicting changes and levels of opportunity cost understanding

Variable[a]	**Outcome: OC difference: TEL**		**Outcome: OC level: Exam**	**Outcome: OC level: Last TEL**
Took econ. elsewhere	**–0.07**	**–0.08**	**–0.07**	**–0.06**
	(0.12)	(0.12)	(0.09)	(0.09)
Age	**0.01**	**0.01**	**0.01**	**0.01**
	(0.01)	(0.01)	(0.01)	(0.01)
Female	**0.16**	**0.16**	**–0.11**	**0.16**
	(0.11)	(0.11)	(0.10)	(0.09)
International student	**–0.76***	**–0.76***	**0.00**	**–0.73****
	(0.34)	(0.34)	(0.25)	(0.25)
N	431	431	415	415
R^2	0.5455	0.5436	0.2209	0.6582

[a] Robust standard errors are in parentheses; results significant at the (1%) 5% level are double-starred. Column 1 uses students' total scores on the misconception/perception variables, whereas Column 2 uses their absolute scores. In all remaining regressions, absolute scores are used. In regressions controlling for the variable *OC level: First TEL*, only the responses to questions on the first TEL that are *not* used to construct *OC level: First TEL* are summed and used in Row 2. The following other variables are included in each specification: initial understanding of opportunity cost, as measured using responses from the first iteration of the TEL (in regressions where the change in opportunity cost understanding is not included); part-time student status; residency status; course load; term of entry; number of tutorials attended; which lecture the student attended; indicator variables for South Australian residence, public school attendance, Year 12 entry status, quarter of birth and academic programme; and indicator variables for attendance at Monday and morning tutorials, and for having different tutors.

Linear regression is used to detect the impact of the independent variables on the proxies for threshold concept acquisition and course success. While linear regression has significant advantages over other statistical techniques in its simplicity and ease of interpretation, and while it allows us to examine the effect on our outcomes of each independent variable in a *ceteris paribus* fashion, the validity of assessing the effects it estimates is dependent upon assumptions that are worthy of addressing. Appendix B reports the results of tests for normality, multicollinearity, omitted variables and outliers. None of these test results indicate that linear regression or the standard statistical testing procedures on the estimates obtained thereby are inappropriate. Furthermore, as linear regression is often used as a first-pass modelling technique, even when the outcomes they are predicting are not perfectly continuous or unlimited in range, and given that this is a pilot study, the use of such a standard technique appears reasonable.

Columns 1 and 2 of Table 5 show the results of regressions predicting the growth in opportunity cost understanding over the course of the semester. The sole difference between these first two columns is the way in which students' misconceptions of economics and perceptions of economists are parameterised.

It may appear from Row 1 of Table 5 that students' scores on the opportunity cost questions from the first iteration of the TEL strongly predict acquisition of the concept. However, recalling the analysis of correlations in the previous section, this apparently strong effect should be seen as an artifact of the mechanical relationship between the dependent variable and this particular independent variable. We chose to leave the variable in the model, despite this concern, because it was important to eliminate the effect of students' first scores before interpreting the estimated effects of other potential predictors of threshold concept acquisition. The estimated effects of other variables in the model in the specification should be pure of any bias relating to their association with students' first scores on the opportunity cost questions (or, indeed, any other questions) on the first TEL.

As can be seen from the results in the third row of Columns 1 and 2 of Table 5, a student's score on the first TEL is an important predictor of threshold concept acquisition. We take this as evidence that prior learning significantly determines levels of acquisition of this threshold concept. Having taken economics elsewhere, however, does not predict opportunity cost acquisition, despite the potential interpretation of previous coursework in the area as an indicator for prior knowledge. Perhaps, prior knowledge (at least in economics) may be better measured by the TEL than by a simple 0–1 measure capturing the student's academic history in the discipline.

From the significant estimated effects (in Columns 1 and 2) of enjoying mathematics on threshold concept acquisition, it is possible to infer that liking maths may make students ready or willing to intellectually engage with economic threshold concepts. This may be a discipline-specific result, as economics is more maths-based than other social sciences.

There is a significant and large estimated effect of international student status on threshold concept acquisition and, even more intriguing, this effect is independent of whether or not English is a student's mother tongue. We interpret this as evidence that culture, even more than starting communication skills, is a determinant of the acquisition of opportunity cost. It may be that the mental reorientation required for full absorption of this concept is more difficult for students who come from cultures whose values or norms are perceived (by the student) as being at odds with the concept itself.

Finally, misconceptions of economics, perceptions of economists and level of market understanding do not seem to matter to students' acquisition of this threshold concept. Indeed, no other variables in this regression are significant than those discussed above.

Columns 3 and 4 of Table 5 show results of predicting the *level* (as opposed to growth) of opportunity cost understanding at the end of the course, as measured, respectively, by the final exam questions on opportunity cost, and the final TEL questions on opportunity cost. While these models are not predicting acquisition of the concept (but rather, students' level of understanding about it at the close of the course), they do allow us to compare two alternative measures of understanding: one derived from questions on opportunity cost from the final exam, and the other from the final iteration of the TEL.

It is of interest that very few explanatory variables are significant in Column 3 of table 5. This may indicate, in part, the presence of a fair amount of random variation present in the dependent variable particularly given that it is a function of only two questions, and moreover questions that were answered in an examination environment. For example, students who had in fact acquired the concept may have lost focus, become bored, or felt the need to move on, when confronted with those two particular questions. This speaks to the desirability of operationalising any threshold concept using as many overlapping variables as possible.

In the final column, it again should be noted that the difference in opportunity cost understanding is mechanically related to the final level of this understanding – one reason why estimating this equation may be seen as questionable by some readers. Again, however, our goal was to control both for initial level of understanding (effectively, prior knowledge as approximated by this variable) and acquisition of the concept over the course of the semester, so that the effects of the remaining independent variables could be interpreted as pure from any bias caused by their correlation with these variables.[8]

Column 4 shows that mathematics seems to be important in determining the final level of opportunity cost understanding, as measured on the TEL, but here – unlike in Columns 1 and 3 – it is whether the student took high-level mathematics, rather than whether they liked maths, that seems to matter. This is consistent with earlier work by Shanahan and Meyer (2001b). International student status again comes up negative and highly significant, while the remaining variables (including English language) are insignificant.

Finally, none of the independent variables that are within the direct control of the lecturer or tutor appear to affect opportunity cost acquisition or understanding

(see the footnote to Table 5 for a full list of control variables). In particular, controls for tutorial attendance, lecturer, tutor, Monday tutorial and morning tutorial that are included in all specifications (though not shown) are estimated to have no effect on opportunity cost acquisition although as a group they do significantly impact on measured success in the course. Only those attributes that students arrive at university in possession of already – in particular, mathematics history or preference, international student status, and prior knowledge of the discipline – seem relevant in explaining why some students acquire the concept of opportunity cost and others do not. This seems to indicate that what happens in the classroom is not nearly as important to the acquisition of threshold concepts as the characteristics and knowledge which students bring to the table.[9]

As we are also interested in determining whether threshold concept acquisition catalyses measured course performance, we next run five regressions predicting each of the following overall success measures: student' final percentage in the

Table 6 Regression results: predicting final course performance

Variable[a]	**Outcome Final percent**	**Outcome: Raw exam**	**Outcome: Exam A**	**Outcome: Exam B**	**Outcome: Exam C**
	(1)	(2)	(3)	(4)	(5)
OC difference: TEL	**0.70***	**0.72**	**0.28**	**0.11**	**0.33***
	(0.33)	(0.34)	(0.14)	(0.14)	(0.16)
First TEL	**0.71****	**0.77****	**0.31****	**0.21****	**0.25****
	(0.09)	(0.10)	(0.04)	(0.04)	(0.05)
Miscon. – Economics	**–0.01**	**–0.06**	**0.05**	**–0.13**	**0.02**
	(0.27)	(0.30)	(0.11)	(0.11)	(0.14)
Percept. – Economists	**0.76****	**0.81****	**0.29***	**0.22**	**0.30***
	(0.25)	(0.28)	(0.13)	(0.10)	(0.14)
Market forces	**–0.52**	**–0.55**	**–0.31***	**–0.04**	**–0.21**
	(0.28)	(0.29)	(0.12)	(0.11)	(0.14)
Took specialist maths	**2.93**	**3.23**	**1.57**	**2.32**	**–.65**
	(4.64)	(4.76)	(1.84)	(1.92)	(2.29)
Family at Uni	**0.91**	**0.97**	**–0.01**	**0.14**	**0.84**
	(1.03)	(1.12)	(0.44)	(0.45)	(0.55)
Enjoy maths	**0.29**	**0.30**	**0.32**	**–0.17**	**0.16**
	(1.11)	(1.20)	(0.50)	(0.47)	(0.57)
Expectations	**1.44***	**1.25**	**0.23**	**0.29**	**0.73***
	(0.71)	(0.74)	(0.31)	(0.30)	(0.34)
English first lang.	**0.44**	**2.56**	**0.77**	**0.42**	**1.36**
	(1.79)	(1.79)	(0.72)	(0.68)	(0.77)

Table 6 (continued) Regression results: predicting final course performance

Variable[a]	Outcome Final percent	Outcome: Raw exam	Outcome: Exam A	Outcome: Exam B	Outcome: Exam C
Took econ. elsewhere	**–1.75**	**–2.12**	**–1.22***	**–0.92***	**0.03**
	(1.16)	(1.24)	(0.50)	(0.47)	(0.57)
Age	**0.04**	**0.07**	**–0.00**	**0.04**	**0.03**
	(0.11)	(0.11)	(0.04)	(0.05)	(0.06)
Female	**2.20***	**1.74**	**0.06**	**0.84**	**0.85**
	(1.03)	(1.08)	(0.44)	(0.43)	(0.53)
International student	**2.03**	**5.30**	**3.56****	**–0.74**	**2.48**
	(3.29)	(3.00)	(1.36)	(1.29)	(1.36)
N	415	415	415	415	415
R^2	0.3965	0.3948	0.3931	0.2823	0.2995

[a] Robust standard errors are in parentheses; results significant at the 5% level are starred; those significant at the 1% level are double-starred. In all regressions, students' absolute scores on the two misconceptions variables are used. `Final percent' is students' final percentage in the course; `Raw Exam' is students' raw examination score. See text for more details on the nature of each part of the final exam. The following array of additional variables is included in each specification: part-time student status; residency status; course load; term of entry; number of tutorials attended; which lecture the student attended; indicator variables for South Australian residence, whether the student attended a public school, whether student admitted through the Year 12 entry programme, quarter of birth and academic programme; and indicator variables for attendance at Monday and morning tutorials, and for having different tutors.

course; raw scores on the final examination; and scores on each of the three sections of the final examination (A, B, and C).[10] In these regressions, we include both standard predictors of success and the measure of opportunity cost acquisition that was itself predicted in Columns 1 and 2 of Table 5. The results of these regressions are shown in Table 6.

Row 1 of Table 6 reveals that opportunity cost acquisition is statistically related to final percentage in the course, and has a strong showing in the exam equations as well (with the exception of Part B of the exam). We take this as tentative evidence that threshold concept acquisition may be not only theoretically valuable, but useful to students in a practical sense. Therefore, efforts to assist students in crossing learning thresholds by helping them acquire transformative concepts may result not only in students with more profound knowledge, but in students who are more successful on paper.

Despite their lower threshold concept acquisition, as estimated in Table 5, international students are estimated to do better on the exam overall than non-international students (a result driven by international students' higher performance on Part A of the exam). Given that Part A is designed to assess basic core concepts taught in the course using multiple-choice questions, this result suggests that English language ability may be a driver of student success and that multiple choice questions remove or diminish language barriers for international students. It may also be suggestive that these same students are less comfortable with questions requiring an argued response. Further work is necessary to isolate the reason(s) for this finding.

A few other results from Table 6 are worth mentioning, although they do not bear directly on the notion of threshold concepts. One interesting result, seen in the fourth row of this table, is that students' perceptions of what economists do are powerfully related to their final percentage in the course and to their raw exam score, and are also significantly related to their separate scores on Parts A and C of the exam. A related result is the negative estimated effect of the understanding of market forces on students' performance in Part A of the exam. Both of these variables may, in fact, be reflecting important aspects of students' preparation for learning in economics. Having a tacit understanding of what economists do may indicate a deeper acquisition of some of the personal thresholds necessary to do well in economics. Similarly, the variable 'market forces,' designed in the RoLI to reveal quite simplistic understandings of market forces, may pick up students with an overly simplistic view of how markets and economies work, which may in turn bode poorly for more nuanced learning. These issues are discussed in more detail in Shanahan and Meyer (2001a).

Students who have taken economics elsewhere are estimated to do worse in Parts A and B of the exam – a result that may seem surprising. However, this may indicate an important difference between the ways in which economics is taught in high school or technical colleges, and the way in which it is taught at university; and/or, it may reflect a too-casual attitude on the part of students who think they will do well due to previous exposure to the discipline, and as a consequence do not prepare for the exam as thoroughly as others.

Finally, students' expectations of success in the course are related to their performance in Part C of the exam, and to their final percentage in the course but not to performance in either Part A or Part B of the exam. This may be due to the fact that these students are more prepared to acquire the more comprehensive learning in the course that helps them succeed on the portions of course assessment that reward holistic knowledge (in particular, essays). Note, however,

that expectations had no role in determining the acquisition of opportunity cost. Perhaps breadth, rather than (necessarily) depth, of learning is what these students are particularly ready to accept. Alternatively, expectations of success may be correlated with underlying characteristics, such as writing ability, which are best reflected in essay questions.

Finally, it is worth commenting on the difference between focusing on opportunity cost measures and traditional measures of success. Taken together, Tables 4, 5 and 6 first suggest that there is a positive association between examination success and a student's level of understanding of the concept of opportunity cost at both the start and end of the course.

Further, we can observe how levels and changes in students' articulation of opportunity cost (measured via multiple choice questions) are affected by prior knowledge, just as are more traditional measures of success (e.g. examination results). To the extent that the examination results reveal students' acquisition of fundamental concepts in economics, these results are both expected and pleasing. Nonetheless, it is also true that to date, traditional methods of assessment have not explicitly aimed to assess threshold concept acquisition, but instead often demand a broader ability to absorb, reflect upon and apply such concepts in new contexts.

This necessarily makes the statistical link between threshold concepts and traditional success measures comparatively weak (and helps to explain why many variables which are estimated to impact on exam success are not estimated to impact on threshold concept acquisition, or even are estimated to impact on it in the opposite direction – e.g. international student status). This also suggests a fruitful area for future research: an exploration of the correlation or qualitative congruence between traditional measures of student learning and students' acquisition of threshold concepts.

Conclusion

The present paper has attempted to operationalise the notion of threshold concepts, by explicitly measuring students' acquisition of one particular concept using readily available multiple choice question responses. Our results indicate that the acquisition of the threshold concept of opportunity cost is determined more by students' background and preferences than by what happens in the classroom. Prior experience and prior knowledge matter. We further provide evidence that students' final performance in the course is weakly related to their acquisition of this key threshold concept during the course of the semester. The weakness of this relationship may provide some evidence for the proposition put forth by Davies

and Mangan (2006) that a single discipline-based threshold concept is nested within other concepts and that students must have progressed through these to demonstrate progress through the discipline- based concept. More disturbing, these results are consistent with a view outlined in Shanahan and Meyer (2006) that current assessment practices may be more inclined to select students whose prior conceptions are compatible with the discipline, rather than to assist in reorienting students' misconceptions over the course of a semester. It may also be consistent with the view that one semester is insufficient time to alter such misconceptions.

Our results are tentative, primarily due to the way in which we construct our measures of opportunity cost acquisition and this highlights the need for better measures. The limitations of using an instrument designed initially to examine core concepts, and here reapplied to threshold concepts, are acknowledged. While we do detect significant variation in students' abilities to acquire the concept of opportunity cost, the variation in our constructed measures is likely to contain measurement error. We show that using more conventional learning outcome measures, standard achievement predictors are estimated to have significant effects – whereas many do not when using our measures of the growth in opportunity cost understanding. We hope that these results motivate other researchers to put more effort into the development of tractable, reliable measures of student acquisition of threshold concepts, and thus better map the patterns in students' passage through basic portals in economics.

Our initial work in this area suggests to us that any attempt to design multi-layered and empirically verifiable questions capable of detecting variation in liminality will prove challenging and is likely to be contested. There are a number of reasons for this, not the least of which includes practical issues, such as variation in students' language abilities; the intent of individual courses (are service courses aimed at bringing students 'through' thresholds or only 'up to them'?); curriculum content pressures (the need to introduce a certain range of concepts in a set time); and the acceptance of the profession as to the efficacy of such approaches in revealing student comprehension.

However, using a simple technique where we exploit student answers to multiple choice questions that are relevant to the particular threshold concept we consider, we are able to make some headway in bringing threshold concepts to the empirical arena. Our hope is that our efforts and similar explorations of the drivers of threshold concept acquisition, and the nature of the relationship between such acquisition and final academic success, will eventually allow educationalists working in this area to develop strategic educational policy implications designed to maximise student learning.

Appendix A: Opportunity cost questions – coding methodology

This appendix describes the way we coded answers to the questions on opportunity cost from the final examination and the Test of Economic Literacy (TEL). We used students' resultant scores to construct our measures of the degree of absorption of the concept, both in levels and changes. For all but one opportunity cost question, the best alternative was scored as 2, but alternative incorrect answers could be scored as either 1 (incorrect, but the response was not completely wrong) or 0 (the response given was very wrong). For the remaining question, which was less directly related to opportunity cost, the best alternative was assigned a score of 1, and the rest 0.

Two questions asked in the final exam that related to opportunity cost were used in constructing one of our measures of the stock of threshold concept knowledge. The questions and the scores we assigned each of their alternative answers are as follows.

The rationale for the law of increasing opportunity costs revolves around the fact that:

(a) directing resources away from consumer goods involves a sacrifice of consumer satisfaction
(b) the production of one good can only be undertaken by sacrificing the production of others
(c) resources can be switched easily from the production of one good to the production of another
(d) there is full employment of some resources but underemployment of others
(e) resources are not perfectly adaptable between alternative uses.

The correct answer here is (e). The alternative responses were scored as follows: a = 0, b = 1, c = 0, d = 0, e = 2. This question involves an extension of the concept of opportunity cost, as students must combine their knowledge of that threshold concept with the concept of frictions and associated stickiness of resource allocation. However, as the second-best answer would reveal understanding of opportunity cost by itself, albeit without the additional appreciation of frictions, we believed it was worthy of one point.

You have carefully calculated your weekly budget, and you have not more than $85 available to spend on groceries this week. At the checkout your total purchases add up to $82.80. At that moment your three year-old daughter asks you to buy her a chocolate which costs $2.50. Which of the following statements would you use if you wished to deny her request while at the same time teaching her about opportunity cost?

(a) 'I don't care if you stamp your feet or cry, you can't always get your way'
(b) 'We need the money to buy vegetables'
(c) 'You can have chocolate when you go and see Granny and Grandpa'
(d) 'That chocolate is too expensive compared to other chocolates'
(e) 'You've got your own money in your piggy bank'

The correct answer here is (b). The alternative responses were scored as follows: a = 0, b = 2, c = 0, d = 1, e = 0. The concept of opportunity cost is here juxtaposed with several other concepts, not all of them economic. The economics-relevant ideas that appear include opportunity cost, discounting and operation within a budget constraint. To choose the best answer, students must separate the notion of being on a production (or, in this case, consumption) possibility frontier from the concept of opportunity cost. However, as there is a connection between the two concepts, we decided to award one point for grasping the notion of budgets and least-cost choice.

Three questions from the TEL were also used to reveal a student's understanding of opportunity cost. Because these questions were posed both at the start and at the end of the course, we are able to use responses to these questions to construct measures of both the stock of students' knowledge at both time periods and the change in their knowledge across the course of the semester. The questions and the methodology we used to assign scores to each of their alternative answers are as follows. Note that these questions are standard in the discipline. In order not to compromise their future usability, only the scoring method and not the correct answers are given here.

Question 1. The opportunity cost of a new public high school is the

(a) money cost of hiring teachers for the new school
(b) cost of constructing the new school at a later date
(c) change in the annual tax rate to pay for the new school
(d) other goods and services that must be given up for the new school.

The best answer here was allocated a score of 2. The next-best answer was allocated a score of 1, and the remaining two answers a score of zero each.

When a nation's human and material resources are fully employed, more of any one product

(a) can be produced only if private enterprise does so rather than the government
(b) can be produced only if there is less production of other goods
(c) can be produced only if there is a general decrease in prices
(d) cannot be produced.

The best answer here was allocated a score of 2. The next-best answer was allocated a score of 1 and the remaining two answers a score of zero each.

If Britain has a comparative advantage over France in the production of cars, then

(a) the opportunity cost of producing cars is Britain is lower than in France
(b) the opportunity cost of producing cars in Britain is higher than in France
(c) there are no gains from specialisation and trade in cars between Britain and France
(d) only Britain will gain from specialisation and trade in cars between Britain and France

The correct answer here was allocated a score of 1, and the remaining three answers a score of zero each.

Appendix B: Use of regression

This appendix provides the results of various testing procedures aimed at gauging whether linear regression and the conventional testing we apply to our regression results are appropriate in our data context.

First, we show histograms of our five primary outcome variables, labelled 'OC difference: TEL,' 'OC level: Exam,' 'OC level: Last TEL,' 'Final Percent' and 'Raw Exam' in the paper, with overlaid normal curves in all graphs. While these variables (and three out of five of their residuals after estimation) do not pass strict statistical tests for normality, we feel that using a more complicated estimation procedure, such as tobit, Poisson regression, quantile regression, or kernel estimation would have serious implications for ease of interpretation given our goals and our target audience. Further, lack of normality does not imply that the coefficient estimates we obtain are unbiased or inconsistent, and our standard errors are still valid asymptotically, even if our dependent variables are not exactly normally distributed (for further information, see Chapter 5 of Wooldridge (2006)). This is the primary reason that even variables that do not pass tests for strict normality are frequently regressed linearly in applied social science.

Finally, we also check for omitted variables (using the Ramsey RESET test) and outliers. We find no evidence of outliers, and eight out of nine of our reported regressions pass the Ramsey RESET test at the 5% level. Variance inflation factors are less than 10 for all independent variables in every regression and all standard errors reported in this paper are robust to heteroskedasticity of unknown form.

Figure 1. Histogram of outcome 1

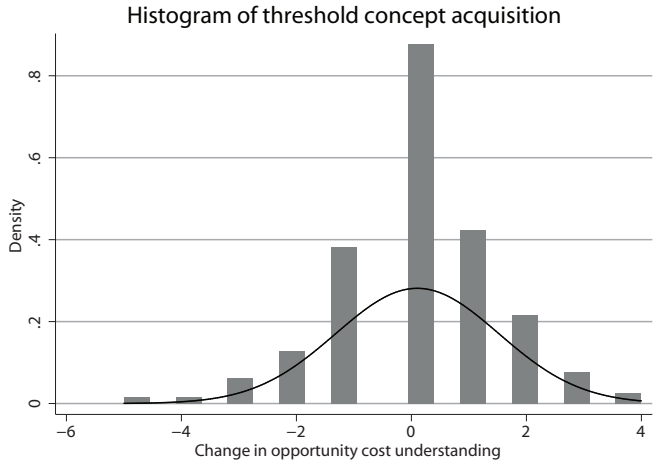


Figure 2. Histogram of outcome 2

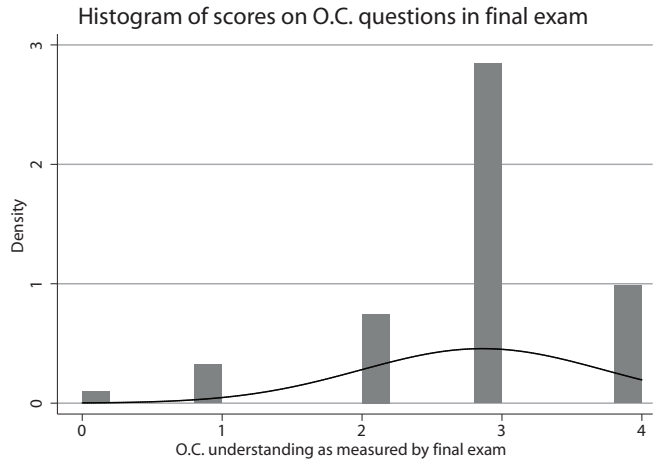


Figure 3. Histogram of outcome 3

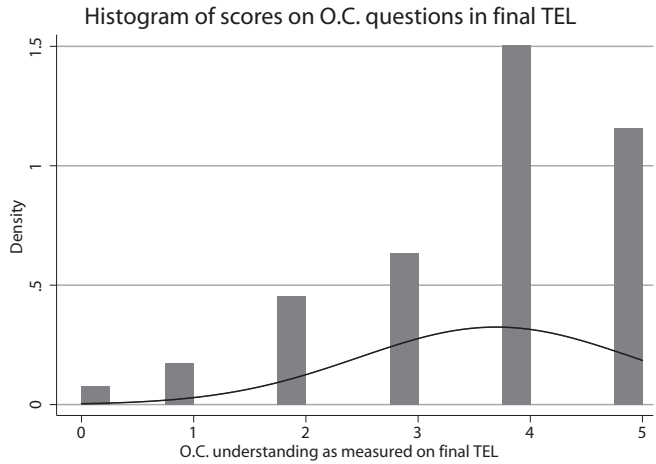


Figure 4. Histogram of outcome 4

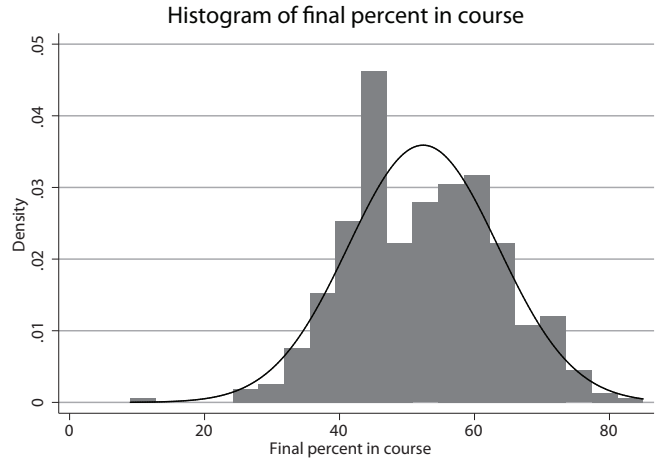
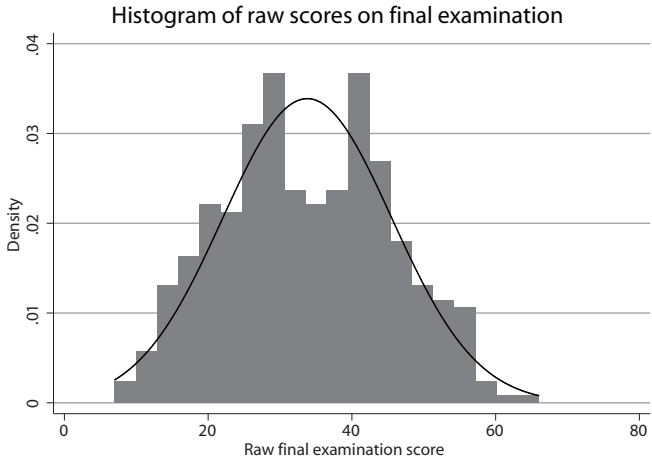


Figure 5. Histogram of outcome 5



References

Beck, K. and Krumm, V. (1991) 'Economic literacy in German speaking countries and the United States: First steps to a comparative study', *Economia*, 1(1), 17–23.

Becker, W. E. and Watts, M. (2001) 'Teaching methods in US undergraduate economics courses', *Journal of Economic Education*, 32(3), 269–280.

Buckles, S. and Siegfried, J. J. (2006) 'Using multiple choice questions to evaluate in-depth learning of economics', *Journal of Economic Education*, 37(1), 48–57.

Davies, P. and Mangan, J. (2007). 'Recognising threshold concepts: An exploration of different approaches', *Studies in Higher Education* 32 (6). Forthcoming.

Eatwell, J., Milgate, M. and Newman, P. (1998) *The New Palgrave: A Dictionary of Economics*, MacMillan Reference, London.

McKenna, K. (1994) 'High school economics in Australia', in W. Walstad (ed.), *An International Perspective on Economic Education*, Kluwer Academic Publishers, 219–232.

Meyer, J. H. F. and Land, R. (2003) 'Threshold concepts and troublesome knowledge: linkages to thinking and practising within the disciplines', in C. Rust (ed.), *Improving Student Learning: Ten Years On*, Oxford, 431–443.

Meyer, J. H. F. and Land, R. (2006) 'Threshold concepts and troublesome knowledge: Issues of liminality', in J. H. F. Meyer and R. Land (eds), *Overcoming Barriers to Student Understanding: Threshold concepts and troublesome knowledge*, Routledge, London, 19–32.

Meyer, J. H. F. and Shanahan, M. P. (2002) 'On variations in conceptions of "price" in economics', *Higher Education*, 43, 203–225.

Perkins, D. (1999) 'The many faces of constructivism', *Educational Leadership*, 57(3), 6–11.

Shanahan, M. P. and Meyer, J. H. F. (2001a) 'A student learning inventory for economics based on the students' experience of learning literature', *Journal of Economic Education*, 32(3), 259–268.

Shanahan, M. P. and Meyer, J. H. F. (2001b) 'A triangulated approach to the modelling of learning outcomes in first-year economics', *Higher Education Research and Development*, 20(2), 127–145.

Shanahan, M. P. and Meyer, J. H. F. (2003) 'Measuring and responding to variation in aspects of students' economic conceptions and learning engagement in economics', *International Review of Economics Education*, 1, 9–35.

Shanahan, M. P. and Meyer, J. H. F. (2006) 'The troublesome nature of a threshold concept in Economics', in J. H. F. Meyer and R. Land (eds), *Overcoming Barriers to Student Understanding: Threshold concepts and troublesome knowledge*, Routledge, London, 100–114.

Whitehead, D. J. and Halil, T. (1991) 'Economic literacy in the United Kingdom and the United States: A comparative study', *Journal of Economic Education*, 22(2), 101–110.

Wooldridge, J. M. (2006). *Introductory Econometrics: A Modern Approach*, Thomson South-Western, Mason, Ohio.

Notes

1 Ideally, multiple choice questions designed to assess variation in the acquisition of threshold concepts should be statistically robust and theoretically sound. To help achieve this, the alternatives in any question should be built and graded on empirically valid and consistent grounds, while the context of such alternatives should also permit the examiner to detect possible underlying reasons why an incorrect alternative has been selected. For example, the alternatives could be based upon categories of `troublesome knowledge' (Perkins, 1999).

2 We are grateful to Professor Bill Walstad for advice and assistance in using this instrument.

3 Most control variables are coded into indicator variables (e.g. 'Have you studied economics elsewhere?' = 0 (no) or 1 (yes)) for inclusion in the descriptive analysis and regressions reported in the present paper. Student age is, however, included as a continuous variable, as is TEL score.

4 Note that as our data include all students who participated in the course at the various points of observation reported in Table 1, the only possible source of over- or under-sampling that we face is students' own non- attendance at tutorials or non-participation in other course activities.

5 A third outcome measure we consider is students' scores on the questions appearing on the final iteration of the TEL that are relevant to opportunity cost. However, as this measure in mechanically correlated with our measure of growth in opportunity cost knowledge, we do not focus on this measure (although we do show one regression to predict it, in Column 4 of Table 5).

6 A comprehensive list of the covariates for which we require non-missing observations is provided is the footnote to Table 5.

7 Intuitively, the first of these correlations can be thought of as due to the fact that scoring highly on the initial TEL leaves well-prepared students little room for

improvement. Similarly, students scoring poorly on the second TEL are more likely to have a lower difference score, as there is a minimum threshold to their first TEL score.

[8] We thank an anonymous referee for pointing out the need to justify our decision in this regard.

[9] It is also possible that the TEL, as an instrument for measuring threshold concept acquisition, is in some sense invalid or unstable. This highlights the pilot nature of this work. In the absence of a better instrument, we believe that the reliability and validity of the TEL as a measure of economics knowledge make it a good first step in operationalising the detection of variation in threshold concept acquisition in economics.

[10] Part A of the exam was comprised of 30 multiple choice questions; Part B was made up of 8 true/false questions also requiring a short written explanation; and Part C gave students a choice of writing an extended answer to one of two questions. Both questions in Part C presented students with a scenario and asked them to use the economic concepts taught in the course to 'solve,' or respond, to the scenario.

Acknowledgements

The authors wish to thank Ken Adams, George Bredon, Pauline Halchuk, Anna Hurlimann, Ron Donato, Gerald McBride, Mahin Mahjabeen, Nicky Kassoudakis, Zakir Machmud, Jerry Shen, Marni Mead, Steve Reid, Senada Nukic, Mark Emerson, Mark Nading, Allan Tabor, Garry McLeod, Judy Szekeres, Tania Kanellos, Sarah-Jane Branford, Matthew Giro, and participants at the 11th Biennial Conference of the European Association for Research on Learning and Instruction, held in Nicosia, Cyprus, 23–27 August 2005. Shanahan and Foster also acknowledge the support of a UniSA Division of Business research grant, as well as ongoing support from UniSA's Centre for Regulation and Market Analysis.

Contact details

Martin P. Shanahan (Corresponding author)
University of South Australia
School of Commerce
Centre for Regulation and Market Analysis
North Terrace
Adelaide 5000
Australia

email: martin.shanahan@unisa.edu.au