



Threshold concepts and metalearning capacity

Jan H.F. Meyer, Sophie C. Ward and Paul Latreille

Abstract

This study operationalises the empowering concept of *metalearning* in the specific context of engagement with a *threshold concept*. An experience of metalearning was constituted in two parts. First students' awareness of themselves as learners is prompted by, and focuses on, a *learning profile* that is generated online through the completion of the Reflections on Learning Inventory (RoLI). Second, students are given an opportunity to interpret their respective profiles and write a short and undirected reflective account of their interpretation. The second part of the experience focuses not only on students' awareness but also on their capacity to control their future learning on the basis of their heightened awareness.

An initial metalearning experience was provided early in a microeconomics module encouraging students to reflect on their learning in the context of their prior learning of microeconomics. A second metalearning experience was provided later in the module when the RoLI response context was shifted specifically to *the learning of the threshold concept* in question.

This metalearning experience intervention yielded matched sets of quantitative data (the RoLI responses used to generate the learning profiles) and qualitative data (the reflective accounts based on the interpretation of the learning profiles). Data analysed here emanate from a research project involving three UK universities, two of which are represented in the present study with one foregrounded in terms of a relatively large sample of c. 300 participating students.

The findings indicate that a metalearning experience can be successfully constituted and encapsulated within the learning of a threshold concept for a majority of students. There is however variation in the success (or not) of the metalearning experience, the detail of which reveals much about the dynamics of changed or changing metalearning capacity in relation to the threshold concept considered.

Introduction

The focus of the present study is on *metalearning*, a concept that is situated within the more general conceptual framework of metacognition after the work of Flavell (1976). The concept of metalearning as defined by Biggs (1985) encapsulates two complementary features of (variation in) deep level, self-regulated, learning capacity; namely an *awareness* of, and *control* over, self as learner in some specified context – a capacity that is also conceptually associated with deep-level learning outcomes.¹

Student S253, second essay, on *awareness*:

My only concern is raised by [a small increase in] ... 'detail related process' [an observable² that captures variation in an aspect of learning in which there is a focus on the detail of what is being learned]. The only thing I can point out is that I have always been a bit obsessed with the small details so I can imagine it's going to take me a little longer than a month to get over this.'

Student S278, second essay, on *control*:

By changing my learning methods after the initial [metalearning experience] through reflection, my approach to learning has been positively changed. I am able to better understand the content and learning outcomes, which not only helps to boost confidence, but also makes learning more enjoyable and productive. Less time is spent in undertaking ineffective tasks such as *memorising before understanding*.

The research described here is grounded in the experience of three earlier exploratory studies by Meyer and Shanahan (2004), Meyer *et al.* (2006), and Ward, Meyer and Shanahan (2006), that have explored the degree to which first-year students of microeconomics are able to develop metalearning capacity in that general learning context over a period of one semester (13 weeks) at the beginning of their first-year of university study.

Two questions are addressed here. First, can students studying microeconomics evidence their capacity to gain from a 'metalearning experience' over a shorter period, which in this case averaged 34 days? Second, can they evidence capacity to gain from a metalearning experience defined more tightly by a *threshold concept*?³ In this respect the present study differs significantly from the earlier exploratory studies referred to above.

The findings reported here come from a completed research project involving three participating UK universities, each of which contributed data related to a

specific threshold concept in economics. The two largest datasets (representing two of these universities) respectively represent a metalearning experience involving a specific threshold concept, and these data form the basis of the present study. The first threshold concept, that of the *ISLM model*, is referred to as a *discipline* threshold concept by Davies and Mangan (2008) while the second, that of *elasticity*, is referred to as a *procedural* threshold concept. The distinction being made here by Davies and Mangan is essentially that the 'ISLM model' depicts the interaction between markets, while 'elasticity' (and other threshold concepts) enable such modelling in procedural terms in the 'ISLM model' and more generally.

Students in each university were invited, on two occasions, to (a) generate a self-reported and contextualised *learning profile* of themselves (via completion of an online learning inventory), (b) interpret this profile with the aid of a non-evaluative guide written for this purpose and, (c) write a short (500 word) reflective essay on their profile as self-interpreted. Sandwiched in between these two 'A: first' and 'B: second' metalearning experiences was the learning engagement of the designated threshold concept. The first metalearning experience was contextualised in terms of the learning in general of the microeconomics content up to that point, and the second experience was contextualised in terms of the ensuing learning engagement of a designated threshold concept. It is emphasised that, although the focus here is on economics threshold concepts, the architecture of the metalearning experience as described here is transferable to threshold concepts in other disciplines.

Data collection

Two data sets were gathered at each university: A *quantitative* dataset comprises the aggregated data (the raw inventory input data) generated by the students' completion of the RoLI inventory during the first and the second metalearning experiences. These matched sets of raw data generate the first and second learning profiles. Within sample size limitations these data can be analysed using standard multivariate statistical procedures to explore any underlying patterns of difference (between the first and second self-reported experiences) that might be interpreted as evidence of the presence and development (or not) of metalearning capacity.

A second *qualitative* dataset comprises, for each student, a *matched* set of two learning profiles, each respectively connected to a corresponding short reflective essay. In writing their essays students were deliberately not directed or constrained in terms of what they should write about. All that they were required to do on each occasion was to attempt to interpret their profile (using the non-evaluative guide) and comment on what they thought it told them about themselves and what, if anything, they might like to do as a consequence (giving reasons if possible).

It is acknowledged at the outset that these datasets represent students who have persisted in their studies up to the point of completed data gathering, and the findings reported here are strictly within this constraint. Resultant analyses of these datasets may therefore be biased in that they do not reflect all the students who entered the module(s) in question but dropped out before data gathering commenced, nor those who, for whatever reason, either failed to persist in contributing data, or failed to complete the module(s) in question.

The remainder of the present study is presented in four sections. The first provides a selected commentary on the quantitative analyses of the largest dataset from the first university. The second further illuminates this quantitative analysis in terms of a summary of the findings from the corresponding qualitative analyses, and also findings yielded by a smaller dataset from the second university for which outcomes measures were also available. In the third section a summary of students' reflections on their metalearning experience is presented, and we finally conclude with a discussion and indications for future work.

The quantitative analyses

In order to generate a self-reported learning profile as part of a metalearning experience each student completed the Reflections on Learning Inventory (RoLI) via an online portal (www.RoLIsps.com) especially designed for this purpose. The psychometric development of the RoLI is summarised in Meyer (2004) and its domain, in the generic version used in the present study, is defined in terms of 16 observables, each represented by a subscale of five items scored in terms of a mixed metric that includes a Likert-type response format. It is the subscale scores that are the discrete conceptual components of the learning profile and that constitute the data for the quantitative analyses.

Location differences

The two matched sets of response data (n=354 individual students) comprise 16 derived subscale scores representing the 16 observables in the RoLI domain. The first set (referred to as the A dataset) comprises responses from students within an economics module just prior to the introduction of the learning segment containing the procedural threshold concept of *elasticity*. In completing the RoLI students were asked to respond in terms of their experiences of learning economics within the module up to that point. The second set of responses (referred to as the B dataset) was obtained after the concept of 'elasticity' and its application had been dealt with, and the same students were asked to respond specifically in terms of their learning engagement with this concept.

A multivariate analysis of variance using the general linear modelling procedure in SAS reveals an overall statistically significant effect on the subscale mean scores attributable to change(s) in the A/B response status. (All the relevant test statistics Wilks' Lambda = 0.88, Pillai's Trace = 0.12, Hotelling-Lawley Trace = 0.14, and Roy's Greatest Root = 0.14 exhibit p values < 0.0001, further details not presented.)

There is therefore multivariate justification for examining the univariate responses via one-way analyses of variance. Thus revealed in Columns 3 and 4 of Table 1 is a conceptually meaningful, and perhaps consequential, change reflected in all 16 RoLI observables, with statistically significant effects within the A/B response status on 15 of them as obtained by one-way ANOVA of A versus B. *In terms of mean scores* there is thus a first impression of overall change between the first and second sets of responses, and a second impression of the extent to which this change is significantly evidenced at the discrete univariate level.

The magnitude and the conceptually appropriate direction of these specific univariate changes is crucial. The matched pair structure linking A and B is therefore invoked to determine 16 measures of change for each case (individual student) so as to examine these changes via an analysis of paired comparisons that takes full and appropriate account of the correlation between the first and second responses. The interest here is on the *difference* (B–A) between the A and B responses, but expressed as a percentage change of A. Calculated in this manner, the *percentage change* scores pre-empt any concerns about normality assumptions for integer valued scores. The paired comparisons analysis determines whether these percentage change scores are significantly different from zero for each of the 16 observables. The results of these paired comparisons exhibit a conceptually appropriate and clearly *differentiated* pattern in Columns 5 and 6 in Table 1.

The RoLI domain (Meyer, 2004) captures variation in four observables (*inter alia*) that have been previously argued to be conceptually associated with metalearning capacity (Meyer *et al.*, 2006). The first, *knowing about learning* (Kal) captures variation in an aspect of learning engagement that is conceptually directly associated with metalearning awareness (but not necessarily control)⁴. A second related subscale *knowledge objects* (Kob) captures variation in another aspect of metalearning capacity. Two further complementary subscales, *seeing things differently* (Sdi), and *relating ideas* (Rid), respectively capture variation in a conception of learning and a deep-level learning process.

It is clear from Table 1 that the percentage score changes across this ensemble of four metalearning observables *are all positive and statistically significant*. These changes furthermore occur in the presence of significant positive changes in two additional observables representing deep-level learning processes (*memorising*

Table 1: Statistical dynamics of increased or increasing metalearning capacity (n=354)

Sym- bol	Observables (RoLI subscales)	Cols. 3,4		Cols. 5,6		Cols. 7–10			Cols. 11–14			
		One-way ANOVA ^a	p values, sign ^b	Paired comparisons whole sample	% change ^c from A	p values n=354	F1	F2	F3	Paired comparisons by Gender	Paired comparisons by English first language	Y: n=269 F: n=134 M: n=220 N: n=85
Kal	Knowing about learning	†	+	2.63	0.001	67	07	-02			2.74 †	2.30 ns
Kob	Knowledge objects	§	+	12.84	<0.0001	65	07	-12				
Sdi	Seeing things differently	§	+	6.90	<0.0001	65	-10	02				
Rid	Relating ideas	§	+	5.71	<0.0001	58	-09	04				
Mwu	Memorising with understanding	‡	+	4.99	<0.0001	49	08	13				
Mau	Memorising after understanding	‡	+	7.02	<0.0001	39	-00	07				
Lbe	Learning by example	ns	+	5.68	0.0005	25	08	13			6.67 †	2.55 ns
Fra	Fragmentation	†	-	-1.95	ns	-34	33	-03				
Fac	Learning is fact based	§	-	-6.99	<0.0001	03	81	01				
Kdf	Knowledge discrete and factual	†	-	-3.39	0.0201	03	75	-04			-5.67 †	0.35 ns
Mbu	Memorise before understanding	‡	-	-5.31	0.0037	04	43	18			-3.50 ns	-8.27 †
Dut	Learning experienced as a duty	†	-	-4.90	0.0037	09	41	01			-5.83 †	-3.38 ns
Drp	Detail related process	‡	-	-5.12	<0.0001	-16	28	01				-6.13 †
Rau	Repeating aids understanding	†	-	-2.77	0.0162	-00	-01	95			-2.35 ns	-3.44 †
Rer	Rereading texts	†	+	3.08	0.0002	27	03	49				-2.69 †
Mar	Memorising as rehearsal	†	-	-3.96	0.0094	-13	32	44			-2.67 ns	-6.07 †
Inter-factor correlations												
F1												
F2												
-0.08												
0.07												
0.44												

Note: (a) For the one-way ANOVAs, the significance of the effect of the A/B response status on each observable score is summarised as § ($p < 0.0001$), † ($0.0001 \leq p \leq 0.001$), ‡ ($0.001 < p \leq 0.05$) or ns ($p > 0.05$). (b) The sign indicates an increase (+) or decrease (-) in the mean difference score (relative to A) that is used to (c) calculate the percentage change score (relative to A) for the paired comparisons. (d) All factor loadings rounded to two decimal places and multiplied by 100.

with, and after, understanding: Mwu and Mau). In contrast, there are significant negative changes (lower mean B vs A scores) on a range of observables conceptually expected to *inhibit* (in terms of high scores) deep-level learning engagement (and therefore also metalearning capacity); notably, *fact based learning* (Fac), *knowledge discrete and factual* (Kdf), *memorise before understanding* (Mbu), *detail related process* (Drp), and *memorising as rehearsal* (Mar). Although not significant, the small decrease in the *fragmentation* (Fra) mean score is conceptually consistent and, more importantly, the obvious contribution of this observable is noted further on in (factor) structural terms. That is, the favourable change in metalearning capacity has been augmented by a favourable change in students' deep-learning processes.

There is thus a third impression of change; in this case in terms of a conceptually appropriate and differentiated statistical discernibility of non-zero percentage changes between the A and B responses. *Viewed collectively*, the three impressions regarding mean score differences are conceptually consistent and encouraging insofar as they are conjoint with the conjecture that what has thus far been observed is an overall increased or increasing level of metalearning capacity. But there is, in the analysis presented thus far, no insight into the underlying structure or internal coherence of the evidenced changes.

Structural differences

The next stage of the analysis therefore begins with the observation that the covariance matrices for the A and B datasets are significantly dissimilar ($\chi^2=249.76$, $df\ 136$, $p<0.0001$). There is thus justification *prima facie* for independent analyses of the two variance/covariance structures to elicit contrasts, for example, via separate exploratory factor analyses. Such separate analyses would typically be expected to reveal A versus B contrasts in terms of factor composition and emphasis in factor loadings, but would not reveal the common underlying structure of the observed changes within responses.

The interest here thus lies in the *dimensionality of structural change* occurring between the A and B datasets. This question of dimensionality is addressed directly by factor analysing the differences between the A and B scores, relative to A (that is, B–A). The factor extraction method used is maximum likelihood, under oblique rotation, using squared multiple correlations as the initial communality estimates. A consideration of the scree plot (not presented), and the eigenvalue > 1 factor extraction criterion (first four eigenvalues 4.47, 3.45, 0.98, 0.60), unambiguously indicates the extraction of at most three common factors, as presented for maximum interpretability in columns 7–10 of Table 1.

The factor solution is readily interpretable in conceptual terms. Factor 1, which is virtually independent of the other two factors, is of immediate interest because it represents a proxy for (variation in) metalearning capacity. It emphasises, in terms of the first four highest loadings, the aforementioned ensemble of metalearning observables (Kal, Kob, Sdi, Rdi) augmented here with two additional deep-level learning processes (Mwu, Mau). Given the positive changes in the mean scores on these key observables that have already been observed, the structural cohesion of Factor 1 presents further consistent evidence of (a particular dimension of variation in) increased or increasing metalearning capacity.

In terms of the four highest loadings, Factor 2 represents a dimension of variation defined by fact based learning (Fac), a discrete and factual conception of knowledge (Kdf), memorising before understanding (Mbu), and the culturally sensitive observable of learning experienced as a duty (Dut). There is also a lesser and unambiguous emphasis in terms of the two learning pathologies of *fragmentation* (Fra) and *detail related process* (Drp). Factor 3 emphasises three contrasting aspects of memorising and repetition as a dimension (of variation) that is moderately correlated with Factor 2 ($r = 0.44$).

A more parsimonious two factor solution (not presented) retains the composition of Factor 1 in Table 1, as well as the status (but not the rank order) of the ensemble of metalearning observables in terms of the four highest loadings. The second factor is virtually independent of the first ($r = -0.01$) and combines Factors 2 and 3 within Table 1 with a different set of loadings.

Of further interest is a comparison of correlation structure for the first metalearning ensemble over time. This comparison is presented in Table 2 in which the first A set of correlation coefficients appear above the diagonal in italics and the second B set below it in bold. It is clear from this table that the linear association between the four subscale responses is stronger in the second set of responses. To explore this observation further, the subscales may be treated as four conceptually discrete, but interrelated, constituents of a posited second order unidimensional *metalearning capacity* scale and, as such, there is an interest in the *internal consistency* exhibited by this scale. The alpha coefficients (measures of scale reliability or internal consistency) and average inter-subscale correlations (of responses) for the two scales, are presented in Table 3. It is clear from Table 3 that the A/B responses to the metalearning capacity scale exhibit an increased or increasing internal consistency over time. What is being observed here from a metalearning perspective is an increase in the internal cohesion, in particular, between scores on the prime proxy measure (Kal) and the scores on the other three associated observables. This increase is again consistent with earlier change observations and their conjectured interpretation.

The intervening period

There is variation in the length of time for students between the A and B metalearning experiences and this variation (mean period is 34.2 days, standard deviation 9.20) provides an opportunity to test whether the results reflect students' familiarity with the RoLI. The technical question is whether there is a linear association between the period, and the percentage change relative to A, of the metalearning capacity scale scores and also, for the sake of completeness, the same scale augmented by the two additional process observables of Mwu and Mau (that is, a scale comprising the defining features of Factor 1 in Table 1). The correlation in terms of the metalearning capacity scale scores is practically zero and of no significance ($r=0.07, p=0.1774$). And similarly in respect of the augmented metalearning capacity scale ($r=0.08, p=0.1212$). These results indicate that positive changes in students' profiles on these observables were not simply an artefact of increased familiarity with the RoLI.

Further analyses of location and structural differences are presented in Appendix 1. These analysis, by gender and the status of English as a first language, provide a fuller account of the dynamics of change already presented thus far and also indicate directions for future work.

Table 2: Matched correlations (n=354)

Subscale	Kal	Kob	Sdi	Rid
Kal		0.36	0.46	0.49
Kob	0.51		0.44	0.38
Sdi	0.62	0.52		0.43
Rid	0.58	0.43	0.55	

Note: A set in upper triangle, B set in lower triangle.

Table 3: Comparison of metalearning capacity scale properties (n=354)

	A dataset	B dataset
Scale mean score	58.1	62.3
Std. dev.	9.0	10.1
Alpha	.75	.82
Inter-subscale correlation (mean)	.43	.54

Interim discussion

There is consistent and compelling statistical evidence that, for the sample in question, exposure to an opportunity for students to reflect on themselves as learners twice over a relatively short period results in self-reported increased or increasing levels of metalearning capacity. This capacity is evidenced, in particular, in terms of percentage change scores on six proxy observables, four of which constitute a metalearning ensemble that has also been referred to in earlier work. In terms of this ensemble the findings of the present study are directly comparable with earlier reported findings, derived from the same protocols and also carried out (in an Australian University in a microeconomics response context) but with a longer intervening period. Tables 4 and 5, adapted from the study by Meyer *et al.* 2006 (pp. 254–256) exhibit a pattern of change that is reassuringly similar to the already discussed patterns reflected in Tables 2 and 3.

There is however a major difference between the present and the earlier study referred to above in that the response context of the second exposure to the metalearning exercise is explicitly grounded in the learning engagement of the threshold concept of *elasticity*. Thus, we interpret the location and structural parameters as evidence of the dynamics of change associated *with the learning of a particular concept*, relative to the immediate prior experience of learning within the

Table 4: Matched correlations (n=524)

Subscale	Kal	Kob	Sdi	Rid
Kal		0.42	0.50	0.37
Kob	0.43		0.34	0.34
Sdi	0.60	0.46		0.46
Rid	0.50	0.44	0.53	

Note: First set in upper triangle, second set in lower triangle.

Table 5: Comparison of scale properties (n=524)

	Set 1	Set 2
Scale mean score	60.06	61.1
Std. dev.	8.0	9.1
Alpha	.71	.78
Inter-subscale correlation (mean)	.41	.49

module, rather than the dynamics of change within the module as a whole. Within this focus of the learning engagement of a threshold concept the statistical analyses are unambiguous in their interpretation. Clearly signalled are the multivariate dynamics (in aggregated data) of the metalearning experience in location and structural terms. What the statistical analyses do not, and cannot reveal, are the contextualised details of these dynamics as experienced by the individual students themselves. Some of these details are revealed in the qualitative data that are next considered.

The qualitative analyses

In seeking evidence of metalearning capacity (in terms of both the awareness and control aspects) an obvious difficulty presents itself in relation to the categorisation and comparison of unstructured data. Experience from the already mentioned previous studies has confirmed that inter-individual variation exhibited in the content of *the essays* largely precludes the construction of neat and unambiguous categorisations of description. Although students are invited to write their reflective essays as a personal interpretation of their learning profiles, they are deliberately undirected in this task in terms of a specific prescribed focus for this reflective component of the metalearning experience. This undirected focus, and the resultant uncontrollable variation that reflects it, was intended to ensure that any evidence of metalearning capacity in the essays would arise spontaneously within the locus of students' self-referencing interpretation of their learning profiles, rather than be artificially imposed by prescription. However, based on earlier work (Ward, Meyer and Shanahan, 2006) there was a realistic assumption that, notwithstanding this uncontrollable variation, this evidence would embody discernable metalearning aspects of awareness and control.

The analytical approach adopted conceptually categorised *the profiles* (the primary sources of variation), rather than the essays, and used these categorisations (and supporting evidence in the profiles themselves) to interpret and thematise the reflective essay responses to them. Towards this end, and viewing the profiles as *preference structures* after the work of Meyer (1991), a simple two-way categorisation protocol was devised. Five learning-inhibiting observables are embedded in each profile.⁵ Any one of these, if emphasised in terms of its relative ranking, signifies the foregrounded influence of what is referred to as an interference condition in that student's self-reported learning engagement. If thus emphasised, these *interference conditions* generally signify aspects of learning engagement that are likely to inhibit or indeed preclude the understanding of whatever is being contextualised in the response. This signification applies to the learning of subject content to date (the first response) or the learning of a

threshold concept (in the intervening period up to the time of the second response). In a simple adaptation of the earlier work by Meyer (1996) a profile was categorised as 'good' in the absence of any interference conditions in the five most highly emphasised (ranked) observables in it. A profile was categorised as 'problematic' if it contained one or more of these interference conditions in the five most highly emphasised learning observables.

A summary of the detailed analyses of the matched profiles and correspondingly matched essays (n=302; not all the students for whom matched quantitative data was captured submitted matched essays) is now presented. As background, a recurring theme within the essays is student references to 'study plans' – *expressions of intent* that collectively cover a range of activities. Some students provide a detailed account of their intended actions, and provide a rationale for these actions. For example, some students express the intention to talk with classmates about an economic principle in order to clarify their thoughts about this principle, or express the intention to spend more time in the library in order to read books in order to broaden their understanding of an economic principle. Some students also express an intention to adopt multiple strategies to assist learning, such as using the internet, questioning tutors and using 'mind maps'. For some students, 'study plans' are literally plans for study; for example, creating a revision timetable. Others mention 'study plans' without identifying a specific course of action or a rationale for action, simply announcing instead an intention to work harder, or concentrate more in class.

The two-way categorisation applied to both the first and second profiles permits a four-way subgrouping of the students and their associated qualitative data, and these are discussed one at a time. In doing so, a distinction is made between those students who report either an 'improvement in their learning', and those who 'claim that the exercise has been of benefit'. The former make no mention of the RoLI while the latter make direct reference to the RoLI, 'the inventory' or 'this exercise'.

'Good' first and second profile subgroup (n = 103)

This is the largest subgroup with 39% of the students in it reporting an improvement in their learning engagement, and 19% claiming, in their second essay, that the RoLI metalearning experience (simply referred to hereinafter as 'the experience') has been of benefit. Students in this subgroup express confidence in their learning and claim that the experience has heightened their *awareness* of self as learner. In the second essay they appear to have taken *control* of their learning to improve upon their initial learning profile. However, some members of this group are dismissive of the experience, and display a level of complacency that is absent from the other groups. Student S121, for example, did not think that his learning

engagement had been accurately reflected and furthermore had this to say in his second essay:

Fragmentation score increased from 10 to 12. I rarely use the topic Elasticity's [sic] and have it memorised in a pocket of my brain.

'Problematic' first and second profile subgroup (n = 85)

Notwithstanding the conceptual contrast between this second subgroup and the first, 28% of the students report an improvement in their learning engagement, and 20% claim, in their second essay, that the experience has been of benefit. However students in this subgroup describe their *study plans* (intentions) in their second essays, rather than reflect upon any increased level of control over their learning. Although these students develop an awareness of self as learner and recognise the need to change, they do not appear to have been able (at the time of writing their essays) to apply what they have learned about themselves as learners. In spite of their self-reported improvement in learning engagement, these students position themselves as travellers on a 'journey' to improvement, with the destination some way off.

'Good' first, 'problematic' second profile subgroup (n = 28)

None of the students in this smallest of the subgroups report an improvement in their learning style, but 19% nevertheless claim, in their second essay, that the experience has been of benefit. Students in this category generally appear to change from being confident about their learning to being anxious and/or depressed.

My inventory has taken a u-turn...in the short run it looks dismal and problematic...I feel my inventory has varied due to numerous identifiable reasons. Firstly, the subject 'elasticity' is statistically orientated and uses formulas to create one definite answer. Therefore whilst studying this topic, I found it extremely difficult to alter my perception of some of the statements such as 'learning is fact based' because 'elasticity' is fact based. (Student S40, second essay)

Students also feel that the experience has increased their *awareness* of self as learner. However, they find themselves unable to take control *over their learning of the procedural concept of elasticity* because they encounter difficulty in comprehending the concept, and regress to inappropriate strategies, such as memorising without understanding.

The first and biggest anomaly is the 'memorising before understanding' observable [up from 17, ranked 7th to 21, ranked 2nd]. This shows that

when I am reciting information on elasticity rather than actually trying to understand the method of the subject I am putting too much emphasis into memorising the material. This could be due to the fact that the subject has a lot of equations in it and rather than trying to find the meaning behind the equation I've just tried to memorise them. (Student S270, second essay)

Although these students produce 'good' first profiles in a general context, their lack of inclination to mention learning in another specific and demanding (threshold) context may reflect, for some, an absence of versatility in responding to the learning demands placed before them.

In their first essays, their focus is also generally narrower than that of the students in the other three subgroups.

I was quite surprised that... 'knowing about learning' actually decreased by 3.0 [from 24, ranked 1st to 21, ranked 9th]. This may have decreased as at the moment in elasticity I am finding it quite difficult to relate the topic to other economics topics in great detail which shows maybe I am not completely comfortable with the topic. Overall I think my learning styles have worsened even though I thought I was changing my learning styles to improve! (Student S49, second essay)

'Problematic' first and 'good' second profile subgroup (n = 86)

Students in this subgroup generally appear to have been rewarded the most with 66% of them reporting an improvement in their learning engagement, and 31% of them claiming, in their second essay, that the experience has been of benefit. Students in this subgroup generally demonstrate the most sophisticated level of awareness of self as learner.

Completing the [RoLI] inventory has helped me recognise my own learning preferences. It has helped me to identify my strengths and weaknesses and areas to develop. It has helped me to redirect my approaches... (Student S284, second essay)

While these students exhibit a tendency to make study plans in their first essay, this tendency diminishes in their second essay. Instead of making more plans about how to improve their learning, students in this subgroup tend to reflect upon how the development of their awareness of self as learner has *enabled them to take control* over their learning of (the procedural concept) of *elasticity*. These students identify, in particular, that they now avoid rote learning and no longer view learning as fact based. Students in this category are less critical of RoLI, and have the most satisfaction with their second learning profile.

A minority of students, some 8% of the 302 students in the sample, volunteered critical comments.

Six claimed in their first essay, and two in their second essay, that their RoLI profile was 'not accurate.' 'Not accurate' is interpreted here as a perception of lack of recognition of self 'in the profile':

S268 Essay two: 'I do think the questionnaire is difficult to understand and that the questions catch you out so therefore I don't personally think that all the outcomes are fully applicable to my real methods of learning.'

No students were critical in challenging the contextual value or rationale the RoLI in their first essay, but six were in their second essay:

S180 Essay two: 'If you study something which is not best suited to your abilities and tastes then I believe that your learning style will be irrelevant.'

Four students claimed in their first essay, and five in their second essay, that they may have provided 'false answers' to RoLI. 'False answers' may be defined as answers that the students believed would create a desirable profile, rather than honestly reflect their learning engagement:

S121 Essay two: 'I subconsciously knew what the profile was looking for and adjusted my answer to satisfy the criteria.'

One student claimed in the first essay, and one student in the second essay, that they were 'confused' by RoLI and its interpretation:

S199 Essay one: 'After completing the questionnaire on the RoLI website and creating my own learning profile I am a little confused as to how to interpret the profile.'

A view from learning outcomes

In further seeking to explore any association between metalearning capacity and learning outcomes (where these were available) a problem again arises as to how to compare qualitative unstructured data with quantitative outcome measures expressed as percentages. In this case the outcome measures were used to group both the profiles and the essays and, *within* these outcome groupings, an attempt was made to thematise variation in pattern and structure.

The analysis of qualitative data contributed by the second university in relation to the *ISLM model* (representing a *discipline* threshold concept), although based on a relatively small sample, yields an unmistakable impression of pattern in the data.

This pattern was obtained and interpreted within a conceptually driven partitioning of the data according to the already introduced two-way categorisation.

The number of students involved is small and caution must be exercised in terms of inferring anything beyond what has been observed. There is evidence that, for the self-selecting subsample in question (those who chose to answer an optional examination question on the threshold concept), a metalearning experience can be constituted within the learning context of the ISLM model. There is however a range of variation reflected in this experience and its possible consequential effects. A conservative interpretation of the data supports a conclusion that an understanding of the threshold concept is closely associated with two categories of 'good' second profiles and respectively matched 'good' second reflective essay. Simply put, a 'good' second profile supported by a reflective essay exhibiting some metalearning capacity is a necessary but not sufficient condition for relatively high achievement in the understanding of the threshold concept as evidenced in examination marks. The association in respect of failure is less distinct; failure to demonstrate an understanding is linked to a range of both 'good' and 'poor' profiles. And there is here a tantalising glimpse of the relationship between metalearning capacity and learning outcomes that future research needs to address.

Students' reflections on the experience

Although not originally part of the project research agenda, anecdotal conversations between the third author and several of the participating students at the first university suggested that there might be value in pursuing a follow-up exercise involving semi-structured interviews with students to explore with them their perceptions of the experience they had been through.

Students were invited by e-mail to contribute, on a voluntary basis, to the follow-up exercise and 22 accepted. Just over half were males, with four mature students⁶ and a number of international students, the majority from China. It is acknowledged that this sample is self-selected and, by its nature, likely to comprise those who found their experience of interest or benefit.

One of the key findings in the interviews was that most of the students felt that the profiles *were accurate depictions of their learning*, and were able to recognise themselves and their approaches, both generally and in terms of specific aspects of their learning. For some this acuity came as something of a surprise, since they were sceptical of the ability of the exercise to generate useful insights, and a number initially undertook the exercise solely or primarily for the purpose of gaining credit (a small incentive). Several commented that it confirmed things they recognised but had never thought about previously or at least consciously.

The interviews quickly revealed an important difference between native English speakers and other (especially Chinese) students, something that was explored more explicitly in later interviews. In particular, fluent English speakers were more likely to identify subtle changes of nuance in the RoLI items arising from small variations in phrasing, and to vary their responses accordingly. Even so, international students typically still found the profiles accurate, suggesting that RoLI has diagnostic power even when some linguistic subtlety may be lost.

Nearly all students reported the exercise as having been of benefit, several commenting on this at the outset and entirely without prompting. Among the main benefits was a greater awareness of their learning: 'Now I know how I learn I know what to do' (K). Perhaps the most extreme case was student R, who lamented that:

This could have been done so many years ago. And it could have helped me through my A levels, 'cause I did struggle. I spent two years struggling with work. And if I had've done this [pause] I think it would have helped a long time ago. [Pause] There was never a time when they sat down with us and said "These are inactive ways of learning; these are good ways of learning" ... I've realised what I've done wrong for so many years and actually now being [pause] learning how to change it... It's changed the way I learn for the rest of my life.

Most other students had also made changes to the way they learned, some substantially and others more incrementally. Reading more extensively was a common change, as was an awareness of the need to understand something *before* memorising it and to link ideas. Where they had done so between the two rounds of the metalearning experience many had seen improvements to their profiles, despite the short timescale involved. The impact was not restricted to their learning in economics, *but extended across their modules*, with students recognising that the insights gained, and the greater awareness that most articulated, were more broadly applicable. The impact, as for student R, was also perceived as being of longer term benefit. A majority of students also agreed that the exercise *had enhanced their self-confidence*. For some this was because the profiles confirmed that they were 'on the right track' while, for others, it was the result of making changes to improve the way they learn, or indeed a mixture of the two. A number of students also said the exercise had made them feel *more 'in control' of their learning*, although many, and in particular international students, found this concept (of control) difficult to define.

In summary, overall the responses reveal almost universal approbation for the exercise. Most telling perhaps is that the vast majority would welcome the opportunity to repeat the exercise again, even without credit!

Concluding discussion

Overall there is every indication that a metalearning experience can be encapsulated within a relatively narrow time span during which students engage the learning of a threshold concept.

In the first [RoLI] exercise, I recognised the fact that I should steer clear of using repetition as a way of learning as this does not prove effective for me. I have attempted to do this whilst learning about 'elasticity'. I feel this has proved effective and my learning has become more efficient as a result of this. I have made a conscious effort to learn material through examining different perspectives as well as studying related topics and looking at the bigger picture. I have done this through further reading and research. Through the [RoLI] exercise, it also came to my attention that the most effective way of learning for me was to understand a topic fully before committing it to memory. (Student S190, second essay)

This is in itself an important finding of benefit for the economics education community for two reasons:

First, it confirms that it is possible (even with relatively large student numbers) *as a normal part of academic practice* to create an opportunity for students to voluntarily participate in a metalearning experience that will either benefit them in varying degrees or unveil a range of personal learning issues that invite a response in terms of learning support. Variation in perceived benefit (or not) by students themselves invites reflexive responses in terms of the pedagogy of threshold concepts that can only enhance the student learning experience. It is clear from the qualitative analysis that a majority of students self-reportedly do benefit from their metalearning experience. And also reflected here for the first time in statistical terms are insights into the multivariate dynamics of this process.

Second, in terms of opportunities for extending the scholarship of economics education, present findings signal the presence of *new research agendas* for both qualitative and quantitative studies of student metalearning at the level of threshold concepts. The challenges thus presented contrast with the research on just 'learning' of the past 40 years that has focused almost exclusively on student learning within general, and to a lesser degree, discipline- or subject-specific response contexts.

Appendix 1. Summary of additional subgroup location and structural differences

Location differences

Presented here is a more detailed exploration of overall effects on the RoLI observables *within* the A and B datasets. For *both* the A and the B datasets there are overall significant effects attributable to gender (males: M; females: F), and English first language status (yes: Y, no: N), but not for the status of having studied economics at high school. The four relevant multivariate test statistics (Wilks' Lambda, Pillai's Trace, Hotelling-Lawley Trace and Roy's Greatest Root) however exhibit different levels of significance across the datasets for *gender* (for the A dataset statistics *all four p values* = 0.0003, for the B dataset statistics all four *p values* = 0.0115), but not for English as a first language (for both the A and the B datasets all *p values* for both sets of four statistics < 0.0001). Further details of these statistics, and the statistics for contrasts between the non-significant status of having studied *economics at school*, are not presented. The interest here, as earlier for the whole sample, lies in the percentage changes of the difference scores relative to A, and as a percentage of, A by gender and English language status respectively.

The results of the respective pairwise comparisons are summarised in columns 11 to 14 of Table 1. These results can be compared to the corresponding results for the undifferentiated sample (in columns 5 and 6), and row entries (for each observable, in each of the two respective analyses) and are only tabulated for results that suggest additional conceptual insights *within* the undifferentiated analysis.

The first six observables that define Factor 1 are conceptually the most important and are of immediate interest. For these observables there are no tabulated entries for the *gender* subgroups because, notwithstanding differences in *p* for each gender subgroup in comparison to the undifferentiated analysis, all the percentage change scores for these six observables remain significantly non-zero at, or above, the 5% level. The interpretation is that gender status does not materially affect the conceptual significance attached to the earlier presented observed differences for the undifferentiated dataset. The same interpretation holds for the subgroup for whom English has first language status but, for *knowing about learning* (Kal), the 5% significance level in this particular case is marginal. And while the corresponding shift in the percentage change score for Kal is also positive for the subgroup for whom English is not the first language, the change is not significantly non-zero.

Further gender contrasts emerge for some of the remaining observables. Males exhibit larger (negative and significant non-zero shifts) on Kdf and Dut than females for whom the corresponding shifts are not significant. Additionally the

females *change sign* for the change in Kdf. But for Mbu, Rau, and Mar it is the females who exhibit larger and significant (negative) shifts compared to the males. Thus there are subtle contrasts in gender-based levels of awareness in increased or increasing metalearning capacity that do not materially alter the substantive findings of the present study but which future work might nevertheless explore.

All the tabulated row entries for the English not first language subgroup exhibit a similar pattern of same sign, and smaller (but not significant) percentage change scores compared to the English first language status subgroup. The comparative failure of the tabulated changes to exhibit significance for the English not first language subgroup may in part be reflecting the influence of the smaller sample size but there is nevertheless a suggestion here of a generally lower level of conceptually desirable change that future research needs to address.

Structural differences

A finer grained exploration of (covariance) structural differences is also presented:

- (a) For the A dataset: a significant difference attributable to English as a first language (yes; 269, no; 85) (chisq=202.19, df 136, $p=0.0002$), no significant difference attributable to having studied economics at high school (yes; 92, no; 262) (chisq=160.72, df 136, $p=0.0726$), no significant difference attributable to gender (females; 134, males; 220) (chisq=125.86, df 136, $p=0.7224$).
- (b) For the B dataset: a significant difference attributable to English as a first language (yes; 269, no; 85) (chisq=214.19, df 136, $p<0.0001$), a near significant difference attributable to having studied economics at high school (yes; 92, no; 262) (chisq=162.89, df 136, $p=0.0578$), no significant difference attributable to gender (females; 134, males; 220) (chisq=153.35, df 136, $p=0.1468$).

Although the overall covariance structure differs significantly between the A and B datasets, there is a comparable similarity of pattern in significant differences (or not) in covariance structure exhibited by subgroups *within* each of the (A and B) datasets. Of some note is that the significance levels for all three of the above one-way tests in the B dataset are *lower* than for the corresponding tests in the A dataset. Lower *p*-values in B may suggest corresponding increases for evidence in favour of those differences.

For comparative purposes separate exploratory factor analyses (not presented) were performed on the change scores (B-A) based on the categorical status of *English as a first language*. For the sake of parsimony, two-factor solutions were extracted to facilitate a direct and simple A vs B comparison and also with the earlier mentioned two-factor solution for the undifferentiated sample. The first of

these analyses, for the non-English first language speaking subsample (n=85), is barely credible with a case to observable ratio of approximately 5:1, but the solution nevertheless reflects a pattern similar to that obtained for the English first language subsample (n=269). Both these solutions furthermore exhibit a structural composition similar to that of the two-factor solution for the undifferentiated sample, but with different emphases in terms of factor loadings.

Of interest however is a gender comparison of the attributes of the metalearning capacity subscale as presented in Table 6. For the male subgroup there is unambiguous evidence of an increase in alpha and in the corresponding inter-subscale correlation (mean). An examination of the A and B correlation matrices (not presented) further confirms that all six correlation coefficients for the B responses are higher than those for the A responses and this observation is consistent with the contents of Table 2 for the undifferentiated responses. The female subgroup (Table 6) also evidences an increase in alpha and inter-subscale correlation (mean), but the values are by inspection lower than those for the male subgroup. An inspection of the A and B correlation matrices also confirms that with one exception the correlation coefficients for the B responses are higher than those for the A responses.

Table 6: Gender comparison of metalearning capacity scale properties

	Males (n=220)		Females (n=134)	
	A dataset	B dataset	A dataset	B dataset
Scale mean score	57.3	61.5	59.5	63.4
Std. dev.	8.94	10.29	8.85	9.71
Alpha	.74	.84	.75	.79
Inter-subscale correlation (mean)	.42	.56	.43	.49

Acknowledgements

The work reported here was carried out with the support of a JISC Del II funded project on 'Developing metalearning capacity in economics'. The genesis of this project owes much to the vision and support of John Sloman. The enthusiastic support of Perihan Yavash representing the second participating university, and that of Tim Dunne who contributed specialist statistical insights, is gratefully acknowledged.

Notes

- 1 The editorial by Meyer and Norton (2004), in introducing a special edition of *Innovations in Education and Teaching International* devoted to 'metalearning in higher education', provides a perspective on various methodologies to help students develop their metalearning capacity.
- 2 The term 'observable' is used in preference to the term 'variable' throughout this paper to refer to a conceptually discrete aspect of learning that has been externalised by students and 'observed' in a sense of estimation rather than precise measurement.
- 3 The theoretical framework of threshold concepts basically distinguishes transformative concepts that lead to a new and previously inaccessible way of thinking about something. The seminal work by Davies and Mangan (2008) provides a definitive perspective of threshold concepts within the discipline of economics.
- 4 The Kal observable captures variation in *knowing* when learning has occurred through an experience of acquiring personal meaning, being able to inter-relate further what one already knows, and making sense of what others say. The *control* aspect of learning is implicit in deep level processing as a manifestation of self regulation.
- 5 Detail related process (Drp), fragmentation (Fra), memorising as rehearsal (Mar), fact based learning (Fac), knowledge discreet and factual (Kdf).
- 6 Defined as 21 or over, at the point at which they started their current degree scheme.

References

- Biggs, J. (1985). The role of metalearning in study processes. *British Journal of Educational Psychology*, 55, 185–212.
- Davies, P. and Mangan, J. (2008). Embedding threshold concepts: from theory to pedagogical principles to learning activities. In R. Land, J.H.F. Meyer and J. Smith (eds), *Threshold concepts within the disciplines*. Sense publishers: Rotterdam and Taipei (pp. 37–50).
- Flavell, J.H. (1976). Metacognitive aspects of problem solving. In L. Resnick (ed) *The nature of intelligence*. Mahwah, New Jersey: Laurence Erlbaum Associates.
- Meyer, J.H.F. (1991). Study Orchestration: The manifestation, interpretation and consequences of contextualised approaches to studying. *Higher Education*, 22, 297–316.
- Meyer, J.H.F. (1996). Some aspects of the individual-difference modelling of causal attribution, *Higher Education*, special edition on student learning, 31, 51–71.
- Meyer, J.H.F. (2004). An introduction to the RoLI. *Innovations in Education and Teaching International*, 41 (4) 491–497.
- Meyer, J.H.F. and Norton, L.S. (2004). Metalearning in higher education. *Innovations in Education and Teaching International*, 41 (4) 387–390. (Introduction to the special edition.)
- Meyer, J.H.F. and Shanahan, M.P. (2004). Developing metalearning capacity in students – Actionable theory and practical lessons learned in first-year economics. *Innovations in Education and Teaching International* (Special issue: Metalearning in Higher Education), 41 (4) 443–458.
- Meyer, J.H.F., Shanahan, M., Norton, L.S., Walters, D., Ward, S. and Hewertson, H. (2006). Developing students' metalearning capacity: A grounded assessment framework. In C.

Rust (ed.), *Improving Student Learning 13 – Improving Student Learning Through Assessment*. OCSLD, Oxford Brookes University, Oxford, 248–266.

Ward, S.C., Meyer, J.H.F. and Shanahan, M.P. (2006). Awareness of self as learner and its contribution to learning outcomes in first-year economics. Paper presented at the 9th EARLI JURE Conference, Tartu, Estonia, June 30 – July 4.

Contact details

Prof. Jan H.F. Meyer
Director, Centre for Learning, Teaching and Research in Higher Education
University of Durham
School of Education
Leazes Road
Durham DH1 1TA
UK

Email: j.h.f.meyer@durham.ac.uk

Animal Spirits: How Human Psychology Drives the Economy, and Why it Matters for Global Capitalism



by George A. Akerlof and Robert J. Shiller, Princeton University Press, 2009.
264 pages. ISBN 978-0-691-14233-3

Animal Spirits: How Human Psychology Drives the Economy, and Why It Matters for Global Capitalism provides an interesting and timely critique of traditional economic theory and discusses applications of behavioural economics to various economics crises, including the current US economic crisis. In the first part of the book, Akerlof and Shiller describe animal spirits as ‘noneconomic motivations’ comprised of aspects of confidence, fairness, corruption, money illusion, and ‘stories we tell ourselves,’ and outline why the presence of these behaviours leads to failures in market capitalism. They contend that, because human behavior is influenced by animal spirits, humans do not behave rationally; therefore, economic theories fail to predict human behavior accurately. The second part of the book illustrates how these irrational animal spirits affect economic decisions by answering eight questions that deal with various economic failures.

Animal Spirits is provocative. The authors provide clear explanations and extensive documentation so that a general audience with an interest in economics can stay engaged. Given current levels of economic upheaval, this book should find wide appeal. The evidence that human psychology drives the economy is persuasive and provides the reader with a different perspective on the economic crises of our day. The authors provide a broader view of economics than the ‘cold economic calculus that economists think should be at the root of all economic behavior’ (p. 104). This book is important and useful in understanding some of the limitations of the economic theory of rational expectations. Incorporating animal spirits into macroeconomic theory is an improvement because it provides a more realistic depiction of our world.