Multi-dimensional Structure of a (Use of) Mathematics Self Efficacy Instrument

Abstract

An instrument was built to measure self-efficacy (s.e.) of 16-17 year old students in relation to their use of mathematics, a 'soft' learning outcome measure that is expected to provide important information when contrasting subgroups following a 'mathematics' and a 'use of mathematics' programme. Analysis revealed significant DIF between the two subgroups. Further, multi-dimensional analysis suggests that 'pure (P)' and 'applied (A)' scores might better be reported separately (in addition to the overall 'maths (M)' s.e. score). Furthermore the subgroup score means $\overline{P}$ and $\overline{A}$ are significantly different in the expected direction (i.e. the use of maths group is significantly more confident on the Applied dimension and vice versa), while the relation of M (A, P) is nearly invariant across subgroups.

1. Objectives

This proposal builds on a larger study aiming to understand how cultures of learning and teaching can support learners in ways that help widen and extend participation in mathematically demanding courses in Further and Higher Education; this involves contrasting a traditional 'mathematics' programme with a new 'use of mathematics' programme in the U.K. (more details provided in the full paper). The particular interest in this paper is self-efficacy, given that social science literature on widening participation suggests that a positive disposition towards a subject is crucial to continuing to study a subject (Bandura and Cervone, 1983; Bandura and Locke, 2003).

Therefore, the purpose of the proposed paper is to report on the development of an instrument to measure mathematical self-efficacy of college students studying 'traditional mathematics' and 'use of mathematics' programs. We will focus on the measurement issues emerging during the validation process with regard to the dimensionality of the construct; the full paper will address validity issues in the comparative context more generally (e.g. dealing with the issue of differential performance among the two subgroups being studied)..

2. Measurement of 'mathematics self-efficacy'- towards a theoretical framework

This study draws on literature on self-efficacy and particularly the mathematically related work and instruments.

The self-efficacy construct was initially described and contextualised by Bandura who distinguished two cognitive dimensions in this construct, i.e. personal self-efficacy and outcome expectancy. Self-efficacy (s.e.) beliefs *"involve peoples' capabilities to organise and execute courses of action required to produce given attainments"* and perceived self-efficacy *"is a judgement of one's ability to organise and execute given types of performances…"* (Bandura, 1997, p. 3). As such, perceived s.e. measures self-belief only in relation to specific activities, for example one's self-belief in ones ability to use mathematics effectively in the future.

Various instruments have been constructed and validated that measure perceived s.e. in relation to mathematics by (a) undergraduates (Hackett & Betz, 1989; Betz & Hackett, 1983; Krantzler & Pajares, 1997; Pajares & Miller, 1994; Hall and Ponton, 2002) and (b) High School (year 7 and 10, Plessis, 2003).

Our conceptual framework

Taking an individual's self efficacy to be their belief in their capability to successfully complete an identified range of actions in a given field, we devised an instrument that will measure students' self efficacy in their use (or application) of mathematics. In order to define the construct 'use of mathematics' we utilised the concept of general mathematical competences which had proved useful at this level of the curriculum (as developed and used by Williams, Wake, & Jervis, 1999).

The construct 'general mathematical competence' (g.m.c.) was devised as a result of an analysis of the mathematical curriculum required by science students following pre-vocational courses. The authors asked not only 'what mathematics do science students use' but also how do they organise their use of mathematics and they resulted in 7 g.m.c.s which cover a wide range of mathematical content and skills (Williams, Wake and Jervis, 1999). These competences were used for the construction of our self-efficacy instrument, as will be explained next.

3. Methods and Techniques: The development of the instrument

As already stated, items (mathematical tasks) were constructed based on the seven mathematical competences listed below (with a code in parentheses used for item description):

- Costing a project (CP)

- Handling experimental data graphically (G)
- Interpreting large data sets (D)
- Using mathematical diagrams (including plans or scale drawings) (MD)
- Using models of direct proportion (P)
- Using formulae (F)
- Measuring (M)

In addition, however, it was thought that 'use of mathematics' should include some purely symbolic mathematical items (e.g. solving an equation in x). So six such items were added (and coded with A).

The items were categorized into three different levels:

Level 1: immediately post-compulsory study (i.e. in the UK, GCSE);
Level 2: towards/at the end of the first year of pre-university optional mathematics study (called 'AS' in the UK system);
Level 3: post 'AS' study.

As usual in s.e. studies, items were presented to the students in the form of a Likert type scale where they were asked to choose the level of their confidence in solving them (but, it was stressed, without the need for solving them). An example of a pure mathematics item (Level 2) is presented below:

C16
Solving a linear and a quadratic equation simultaneously when there are two different solutions, such as:

not confident at all	Not very confident	fairly confident	Very confident
1	2	3	4

Find the two points of intersection of the straight line, $y + x = 5$, with the parabola, $y = x^2 - 2x + 1$.

A/2

Figure 1: Example of a pure mathematics item

Thirty (30) items were devised overall (10 in each level) and administered to the pilot sample of the study (together with other instruments not relevant here). 3 different versions of the questionnaire were administered to optimise the use of the sample, with a set of link items ensuring that an IRT model could potentially provide a good measurement model. In our case Rasch rating scale (including checks for Differential Item Functioning - DIF) and multifacet measurement and multidimensional logistic modelling are employed (Bond and Fox, 2001; Linacre, 2003; Briggs and Wilson, 2003).

4. Data sources – Sampling

Data come from 341 students, from 23 different further education institutions in UK. Some of these colleges teach both the traditional AS 'Mathematics' (Trad) course as well as AS Use of Mathematics (UoM). The distribution of the students according to their gender and their course is presented in Table 1:

Table 1: Distribution (frequencies) of students according to gender and course

	Maths Course		
Gender	**AS UoM**	**AS Trad**	Total
Male	144	55	199
Female	70	44	114
Missing value			1
			314

27 GCSE students (i.e. students who have not yet started an AS course) (15 male, 12 female) were additionally involved in the study.

5. Results

Initial Rasch analysis showed acceptable fit of almost all the items suggesting that they could constitute a scale, i.e. they might measure students' perceived 'self efficacy in AS mathematics' (named 'MSE'). Infit and outfit mean-square fit statistics were examined and three possible misfit items were identified (more details will be provided in the paper).

Additionally, equation of the item parameter estimates for the two groups (Figure 2, with the lines indicating the 95% confidence intervals) and DIF analysis (Table 2) revealed significant difference between the two subgroups performances (and some of these were potentially interpretable as due to bias). In Figure 2, items circled at the top are those shown to be easier for the UoM students to report higher levels of confidence. Items that have statistically significant DIF value (Linacre, 2003) are highlighted in Table 2.

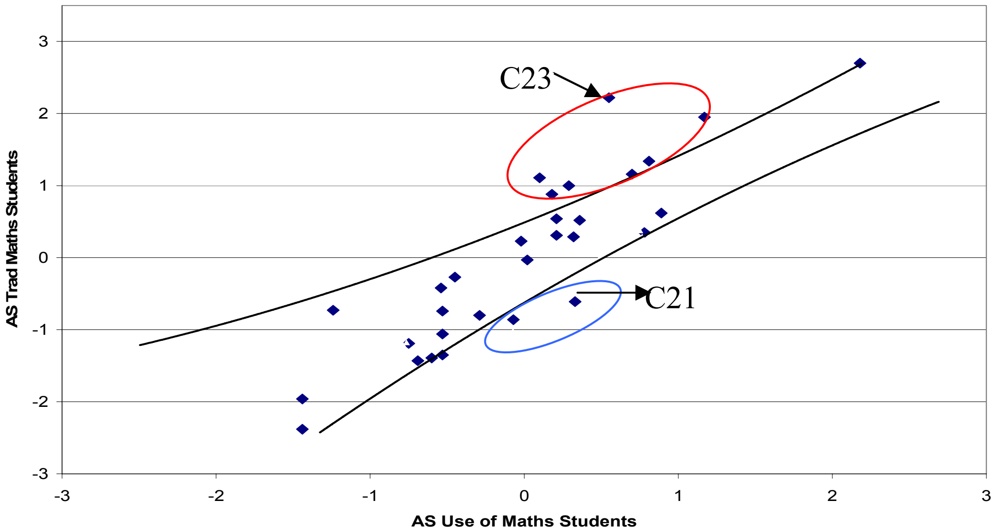

Figure 2: Item estimates (i.e. logits) for the two student groups (with the 95% confidence intervals)

Table 2: Results of DIF analysis for the MSE (unidimensional) construct

Items	Bias measure				DIF size	t-test	d.f.
	Trad	SE	UoM	SE			
C1	-0.1	0.32	0.22	0.46	-0.32	-0.57106	66
C2	-0.09	0.25	0.24	0.4	-0.33	-0.6996	66
C3	-0.18	0.25	0.44	0.38	-0.62	-1.36305	67
C4	-0.1	0.26	0.25	0.4	-0.35	-0.73364	67
C5	-0.25	0.36	0.48	0.46	-0.73	-1.24974	66
C6	-0.05	0.11	0.1	0.15	-0.15	-0.8064	304
C7	-0.19	0.27	0.45	0.4	-0.64	-1.32616	67
C8	-0.18	0.13	0.33	0.16	-0.51	-2.47386	309
C9	-0.04	0.21	0.12	0.37	-0.16	-0.37608	67
C10	-0.05	0.14	0.09	0.19	-0.14	-0.5932	173
C11	-0.1	0.15	0.18	0.19	-0.28	-1.15667	174
C12	0.01	0.1	-0.01	0.15	0.02	0.11094	308
C13	-0.01	0.25	0.02	0.4	-0.03	-0.0636	67
C14	0.04	0.11	-0.08	0.15	0.12	0.645124	306
C15	0.11	0.1	-0.23	0.15	0.34	1.885981	302
C16	-0.32	0.13	0.54	0.16	-0.86	-4.17161	294
C17	0.06	0.1	-0.13	0.15	0.19	1.05393	304
C18	0.17	0.25	-0.52	0.47	0.69	1.296132	65
C19	-0.19	0.11	0.38	0.15	-0.57	-3.06434	302
C20	0.01	0.11	-0.01	0.16	0.02	0.103005	298
C21	-0.44	0.16	0.72	0.19	-1.16	**-4.66998**	166
C22	0.26	0.14	-0.48	0.19	0.74	3.135481	166
C23	0.46	0.14	-0.86	0.19	1.32	**5.59302**	165
C24	0.16	0.14	-0.29	0.19	0.45	1.906711	166
C25	-0.16	0.28	0.38	0.41	-0.54	-1.08764	67
C26	-0.01	0.15	0.02	0.21	-0.03	-0.11625	166
C27	0.09	0.1	-0.19	0.15	0.28	1.553161	300
C28	0.15	0.14	-0.28	0.2	0.43	1.761349	165
C29	0.15	0.14	-0.28	0.19	0.43	1.821968	165
C30	0.03	0.15	-0.06	0.21	0.09	0.348743	163

Thus, we find some items being significantly 'harder' for the traditional 'mathematics' group and some others harder for the 'use of mathematics' group. What is more, removal of the items with most DIF from the instrument does not solve the problem (more details in the full paper). Given these findings, we hypothesized that there might be a latent sub-structure of the instrument that may be better modelled by a two-dimensional model.

Using a ConQuest (Wu, Adams and Wilson, 1998) multi-dimensional analysis (after Briggs and Wilson, 2003) items were scored on a 'Pure' demand and an 'Applied' demand on the basis of judgement about the nature of the 'demand' of each item (i.e. items with no significant pure maths or applied maths demand, as judged by experts, were allocated to 'P' or 'A' respectively while those with some of each were allocated to both dimensions).

Following Briggs & Wilson (2003), we evaluate the comparison of the unidimensional and multi-dimensional models via a chi-square test on the deviance statistic (Table 3). As they state *"because the multidimensional approach is hierarchically related to the unidimensional approach, the model fit can be compared relative to the change in the deviance value, where the difference in deviance between the two models is approximately distributed as a chi-square"* (p. 95) with 4 degrees (37-34) of freedom, in our case.

Table 3: Comparison of the Unidimensional and Multi-dimensional models

MODEL	**Deviance**	**Number of parameters**
Unidimensional	11852.106	33
Multidimensional	11924.786	37

According to the above table, the difference in deviance between the two models is 72.68, suggesting that the multidimensional model fits the data significantly better than the unidimensional model.

We infer that the two dimensional substructure of the instrument suggests that 'pure (P)' and 'applied (A)' scores might need to be reported separately (in addition to the overall 'maths (MSE)' score), as the 2-dimensional analysis reveals a significant improvement over the one dimensional model and additionally the correlation of the two latent dimensions is low ($r=-0.22$, $p<0.05$).

Furthermore, the subgroup score means of PSE and ASE are significantly different in the expected direction, as shown in Table 4. Particularly, the scores of 'Use of maths' group are significantly greater on the Applied dimension and the scores of the 'Traditional maths' group are significantly greater on the Pure dimension.

Table 4: Means for the Trad and UoM students' group for the three s.e. measures and their comparisons via independent samples t-test

S.E. Measures	Groups: mean (SD)		t (d.f)	significance
	AS Trad (N=212)	**AS UoM (N=99)**		
Maths (MSE)	1.263 (1.06)	1.027 (0.85)	1.938 (309)	0.054
'Pure' (PSE)	0.196 (1.02)	-0.398 (1.04)	4.755 (309)	<0.001
'Applied' (ASE)	-0.167 (1.03)	0.162 (1.01)	-2.627 (309)	<0.01

Multiple regression with response variable the one-dimensional construct *Mathematics s.e.* (MSE) and with the two explanatory variables (the latent dimensions PSE and ASE) is summarised in Table 5. The regression equation of the model is also reported afterwards.

Table 5: Regression model for MSE

	Coefficient B	**s.e.**	**t**	**p**
(Constant)	1.239	0.10	128.48	<0.001
'Applied' - ASE	0.874	0.09	92.48	<0.001
'Pure' - PSE	0.786	0.09	83.58	<0.001

$F(2,336) = 6108$, $p<0.001$, $R^2=0.973$ (Adjusted $R^2=0.973$)

MSE=0.874*ASE+0.786*PSE + 1.239

We also find that the relation of MSE with the two latent dimensions (i.e. ASE and PSE) is nearly invariant to the two subgroups (UoM and Trad students). This is justified by the two separate regression models presented in Table 6 that do not seem to differ from the model above, that was estimated using the whole sample.

Table 6: Separate regression models by the two student groups

Model for the group of 'Traditional' math students
MSE = 0.88*ASE + 0.766*PSE + 1.260 *$F(2,309) = 3165.6$, $p<0.001$, $R^2=0.968$ (Adjusted $R^2=0.968$)*
Model for the group of 'UoM' students
MSE = 0.887 * AMSE + 0.832 * PMSE + 1.216 *$F(2,96) = 1866.8$, $p<0.001$, $R^2=0.975$ (Adjusted $R^2=0.974$)*

6. Educational and scientific importance of the study

We conclude that mathematical self-efficacy (s.e.) can be measured by the instrument to provide an appropriate measure of the target students' disposition towards (actually self-efficacy with regard to their) 'use of their AS mathematics'. The subgroup DIF and the multidimensional substructure indicate that there may be two relevant, underlying dimensions that suggest separate s.e. constructs of 'Use of mathematics' and 'mathematics' self-efficacy might better be reported separately. The full paper will also report how these results are mediated by student ability/prior attainment.

References

Bandura, A. (1997). *Self-efficacy: the exercise of control.* NY: Freeman.

Bandura, A., & Cervone, D. (1983). Self-Evaluative and Self-Efficacy Mechanisms Governing Motivational Effects of Goal Systems. *Journal of Personality and Social Psychology, 45*(5), 1017-1028.

Bandura, A., & Locke, E. A. (2003). Negative Self-Efficacy and Goal Effects Revisited. *Journal of Applied Psychology, 88*(1), 87 – 99.

Betz, N.E., & Hackett, G. (1983). The relationship of mathematics self-efficacy expectations to the selection of science-based college majors. *Journal of Vocational Behaviour, 23*(3), 329-345.

Bond, T. G., & Fox, C. M. (2001). Applying the Rasch model: fundamental measurement in the human sciences. Lawrence Erlbaum assoc: Mahwah, NJ.

Briggs, D. C., & Wilson, M. (2003). An introduction to Multi-dimensional Measurement using Rasch Models. *Journal of Applied Measurement, 4*(1), 87-100.

Hackett, G., & Betz, N.E. (1989). An exploration of the mathematics self-efficacy/mathematics performance correspondence. JRME 20(3), 261-273.

Hall, M., & Ponton, M. (2002). A comparative Analysis of Mathematics Self-Efficacy of Developmental and Non-developmental Freshman Mathematics Students. Presented at the *2002 Meeting of Louisiana/Mississippi Section of the Mathematics Association of America*.

Kranzler, J.H., & Pajares, F. (1997). An exploratory factor analysis of the mathematics self-efficacy scale – revised (MSES-R). *Measurement and evaluation in counselling and development, 29*(4), 215-228.

Lent, R.W., Lopez, F.G., & Bieschke, K.J. (1993). Predicitng mathematics-related choice and success behaviours. *Journal of Vocational Behavior, 42*(2), 223-236.

Linacre, J. M. (2003). *A user's guide to FACETS: Rasch-Model Computer programs* [software manual]. Chigago: Winsteps.com.

OECD (2003). *The PISA 2003 Assessment Framework: Mathematics, reading, science and problem solving knowledge and skills.* Paris: OECD.

Pajares, F., & Miller, M.D. (1994). Role of self-efficacy and self-concept beliefs in mathematical problem solving: a path analysis. *Journal of Educational Psychology,* 86 (2), 193-203.

Plessis, J. K. (2003). An exploration and comparison of parents' and children's attitudes to mathemaics in English medium schools in South Wales. PhD thesis, University of Swansea, Swansea.

Williams, J.S., Wake G.W. & Jervis, A. (1999). General mathematical competence in vocational education. In C. Hoyles, C. Morgan & G. Woodhouse (Eds.) *Mathematics Education for the 21st Century*. London: Falmer Press.

Wu, M. L., Adams, R., & Wilson, M. (1998). ACER ConQuest: Generalised Item Response Modelling Software. Melbourne: Australian Council for Educational Research.