Learning and Individual Differences 49 (2016) 305–313

Contents lists available at ScienceDirect

Learning and Individual Differences

journal homepage: www.elsevier.com/locate/lindif

Differential predictors of under-confidence and over-confidence for mathematics and science students in England

Richard Sheldrake

UCL Institute of Education, University College London, 20 Bedford Way, London, WC1H 0AL, England, United Kingdom

ARTICLE INFO

Article history:
Received 22 August 2015
Received in revised form 31 January 2016
Accepted 19 May 2016

Keywords:
Self-concept
Confidence bias
Calibration
Self-enhancement
TIMSS

ABSTRACT

An enhanced understanding of what predicts students' confidence, and what predicts specific cases of under-confidence or over-confidence, benefits educational practices and motivational theories. For secondary-school students in England, confidence expressed as self-concept was most strongly predicted by (intrinsic) interest, perceived encouragement (praise), and subject-comparisons for mathematics, and by praise, interest, and peer-comparisons for science, controlling for achievement and various other factors. The students' reported subject-comparisons, peer-comparisons, anxiety, interest, and (extrinsic) utility differentially predicted the self-concept beliefs of under-confident, accurate, and over-confident students in various ways. For example, for mathematics, higher utility predicted higher self-concept when over-confident (but not when under-confident). For science, lower subject-comparisons (science thought to be harder than any other subject) predicted lower self-concept when under-confident (but not when over-confident). Understanding what predicts someone's self-concept when they are under-confident or over-confident may help these confidence biases to be corrected by educators or even by the students themselves.

1. Introduction

Students' confidence has associated with many aspects of education (OECD, 2015), including motivations to learn (Jiang, Song, Lee, & Bong, 2014), interest in particular subjects (Viljaranta, Tolvanen, Aunola, & Nurmi, 2014), and with choices of what subjects to study (Regan & DeWitt, 2015). However, students' confidence does not necessarily correspond to their actual achievement: some students may be under-confident, with lower confidence than would be expected given their achievement, while others may be over-confident. Since students' confidence influences their motivations and their choices, under-confidence may be limiting or detrimental (Bouffard & Narciss, 2011).

It remains important to understand how students' confidence may be influenced. Students may theoretically form their confidence by considering their own achievement, but they may also receive praise or criticism, compare themselves against their peers, be interested in or anxious about their studies, or be subjected to various other potential influences or factors (Bong & Skaalvik, 2003). Prior research has often explored what predicts students' confidence (e.g. Bong & Skaalvik, 2003; Usher & Pajares, 2008b), but not what might associate with or predict specific cases of under-confidence or over-confidence.

An enhanced understanding can lead to practical benefits: over-confidence or under-confidence could potentially be amended via teachers or wider interventions. Addressing under-confidence may help to (indirectly) increase the numbers of students who chose to study non-compulsory mathematics or science, which remains an important concern for various countries including England (EACEA, 2011; The Royal Society, 2014).

Accordingly, the research presented here explored what factors associated with students' confidence in mathematics and science, and whether any such factors could be associated with under-confidence or over-confidence. The research considered the Trends in International Mathematics and Science Study (TIMSS) of 2011, from the International Association for the Evaluation of Educational Achievement (IEA), and focused on England in order to increase contextualised understanding and relevance to national teaching and policy.

While some studies have undertaken between-country comparisons in under-confidence and over-confidence (e.g. Chen & Zimmerman, 2007; Morony, Kleitman, Lee, & Stankov, 2013; Stankov & Lee, 2014), sometimes defining country-wide or cultural features to help consider any differences (e.g. Chiu & Klassen, 2010; Stankov, 2010), within-country studies remain important. Most students likely select courses within their home country, where under-confidence compared to other students within that country has the most relevance. Nevertheless, compared to other countries, students from across the United Kingdom (which encompasses the country of England, together with Scotland, Wales, and Northern Ireland) have exhibited relatively-accurate confidence, on average, broadly similar to other European countries (e.g. Chiu & Klassen, 2010). England may then provide an informative baseline for other international research or potential comparisons.

E-mail address: rsheldrake@ioe.ac.uk.

http://dx.doi.org/10.1016/j.lindif.2016.05.009

1.1. Students' confidence and potential influences

Within educational, psychological, and other research into individual differences, students' confidence, or their various beliefs in their own abilities or achievement, has often been conceptualised into separate 'self-concept' and 'self-efficacy' beliefs, usually specific to particular academic subjects (Bong & Clark, 1999; Bong & Skaalvik, 2003). Self-concept considers someone's current confidence in their abilities or achievement, integrating their historic and current experiences (e.g. particular achievement grades or accomplishments) and evaluative beliefs (e.g. subjective beliefs of 'doing well' or 'being good' at a subject). Alternately, self-efficacy considers someone's confidence in their future capacities or future achievement, covering contextualised and evaluative beliefs (e.g. perceived confidence in gaining a particular future grade or in completing particular future tasks).

Students' self-efficacy beliefs have been theorised to be influenced by four factors (Bandura, 1997): 'mastery experiences' (successfully completing tasks or not, or gaining particular grades or results); 'vicarious experiences' (seeing others succeed); 'social persuasions' (such as feedback or comments from teachers or other people); and 'physiological states' (such as physical and emotional responses such as anxiety). From these, mastery experiences have generally been the most predictive, while the predictive associations of the others have varied across studies (Britner, 2008; Britner & Pajares, 2006; Usher & Pajares, 2008b; Usher & Pajares, 2009).

Students' self-concept beliefs have been theorised to be influenced by students' mastery experiences, self-comparisons over time, self-comparisons across subjects, comparisons with their peers, and potentially various other factors (Bong & Skaalvik, 2003). Much research has focused on peer-comparisons (Marsh & Parker, 1984; Marsh, Abduljabbar, et al., 2015) and subject-comparisons (Marsh, 1986; Marsh, Lüdtke, et al., 2015), often inferred through specific statistical approaches but also through directly seeking students' views (Huguet et al., 2009; Thijs, Verkuyten, & Helmond, 2010). The effects of peer-comparisons, for example, have been proposed to be large (e.g. Nagengast & Marsh, 2011), but have generally only been considered in isolation. While some research has attempted to explore peer-comparisons and subject-comparisons concurrently, results have varied (Chiu, 2012; Pinxten et al., 2015).

1.2. Motivational benefits or detriments, and biases in students' confidence

Within education, high confidence may be beneficial. Higher self-efficacy has associated with higher motivations to learn and master academic work (mastery goals or goal orientations) (Jiang et al., 2014; Phillips & Gully, 1997), persistence (Multon, Brown, & Lent, 1991), and self-regulation for learning (Usher & Pajares, 2008a; Zimmerman & Schunk, 2011). Higher self-concept has also sometimes been associated with higher subsequent interest (Viljaranta et al., 2014) and higher subsequent achievement (Huang, 2011; Marsh & Martin, 2011; Seaton, Parker, Marsh, Craven, & Yeung, 2014), over and above the effect of prior achievement.

Essentially, the motivational benefits of high confidence appear to be clear, as assumed within social-cognitive theory (Bandura, 1997): high confidence may be beneficial in motivating individuals to surpass their normal performance and overcome initial barriers through persistence or other strategies; but low confidence may mean that some actions are not even attempted. Social-cognitive theory specifically considered self-efficacy beliefs, however; self-concept evolved from general psychological measures and was not originally integrated within wider theory (Shavelson, Hubner, & Stanton, 1976). Nevertheless, subsequent applications of social-cognitive theory, such as the expectancy-value model of students' choices (e.g. Eccles, 2009), have assumed that confidence has a motivational role regardless of whether it is expressed as self-efficacy and/or as self-concept.

However, it remains unclear whether any motivational benefits of high confidence occur regardless of whether someone is accurate in their beliefs (they have correspondingly high achievement) or whether someone is over-confident (they have lower achievement than would be expected given their high beliefs). Someone may also be under-confident, and show lower beliefs than would be expected given their achievement, and it is unclear whether this is always detrimental or limiting. For example, studies of secondary-school students have often associated accurately-evaluated confidence (not being over-confident or under-confident) with higher performance (Chen, 2003; Chen & Zimmerman, 2007; Möller & Pohlmann, 2010; Pajares & Graham, 1999), but have also associated over-confidence with lower performance and under-confidence with higher performance (Chiu & Klassen, 2010). Younger students who consistently over-estimated their abilities over time have reported higher self-esteem than all other students, and generally performed higher than others (Bouffard, Vezeau, Roy, & Lengelé, 2011). Over-confidence has also associated with higher persistence and mastery goals than accuracy and under-confidence (Gonida & Leondari, 2011). In England, over-confidence associated with higher (intrinsic) interest in mathematics and (extrinsic) utility of mathematics at Year 8, while accuracy associated with higher affective responses and intentions to study mathematics further at Year 10 (Sheldrake, Mujtaba, & Reiss, 2014). While it appears uncertain whether over-confident or accurately-evaluated beliefs are the most beneficial, it is often inferred that under-confidence may be detrimental or limiting in various ways (Bouffard & Narciss, 2011).

1.3. Research aims and hypotheses

Diverse research has been undertaken into what predicts students' confidence (whether self-efficacy or self-concept) but often without considering consistent or extensive sets of potential predictors. It remains unclear whether theorised predictors of self-efficacy also predict self-concept, and vice versa, and (especially for self-concept) which factors have the highest predictive associations (or 'effects') when controlling for other factors. Additionally, less research has considered confidence biases and these studies have generally considered differences in students' reported attitudes. It remains unclear whether any factors predictively associate with either under-confidence or over-confidence.

Accordingly, the aims of this study were to: (1) identify what best predicted students' confidence, operationalised as self-concept beliefs, for students in England; and (2) identify what predicted self-concept for cases of under-confidence, accuracy, and over-confidence, and to consider any differences across these cases.

The following hypotheses were made.

Hypothesis 1A. Results would broadly follow those seen when predicting self-efficacy (e.g. Usher & Pajares, 2008b): students' achievement (representing 'mastery experiences') would have a relatively-higher predictive association with self-concept while anxiety, praise, subject-comparisons, and peer-comparisons, would have relatively-lower associations, when controlling for these and other factors. However, subject-comparisons and peer-comparisons could potentially show higher effects (e.g. Marsh, Abduljabbar, et al., 2015; Marsh, Lüdtke, et al., 2015).

Hypothesis 1B. Interest and utility would have moderate predictive associations with self-concept, even when controlling for other factors. Various associations between interest, confidence, and achievement have been seen in prior research (e.g. Viljaranta et al., 2014) and motivational factors such as interest have been theorised to reciprocally associate with confidence (e.g. Eccles, 2009).

Hypothesis 2A. Peer-comparisons/subject-comparisons (i.e. finding science/mathematics harder/easier than other subjects and/or students) would relatively-symmetrically associate with both under-confidence and over-confidence. Research has emphasised peer-comparisons predicting lower self-concept in some contexts (e.g. Nagengast & Marsh, 2011), perhaps entailing under-confidence, but such results would conversely predict higher self-concept in other contexts, perhaps entailing over-confidence.

Hypothesis 2B. Interest and utility would more-strongly predict self-concept beliefs for over-confident students. For example, given close theorised links between factors (e.g. Eccles, 2009), someone may think that they are good at science through (perhaps inadvertently) focusing on their interest or enjoyment, or their wider goals, rather than reflecting on their own achievement, and so become over-confident. Given less prior research in the area, however, it was difficult to form further specific hypotheses.

2. Methods

2.1. Features of TIMSS

TIMSS surveyed 3842 'Grade 8' students (Year 9, on average aged 14.2 years) in England in 2011. TIMSS sampled schools (via strata) and then entire classes of students (Foy, Arora, & Stanco, 2013; Martin & Mullis, 2013). Sampling-weights were applied in order to increase generalisation to the wider population of students across England. For brevity, some of the following methodological aspects are elaborated in the supplementary material (Appendix 1).

2.2. Considered items/factors

A measure of self-concept was defined and separated from its various potential influences (Table 1). Self-concept does not necessarily encompass affective aspects such as interest (Arens, Seeshing Yeung, Craven, & Hasselhorn, 2011), and anxiety and praise are usually conceptualised as potential sources or influences rather than as measures of someone's confidence (Bandura, 1997; Bong & Skaalvik, 2003).

Further factors were defined (following: Foy et al., 2013; Martin & Mullis, 2013) covering the students' perceived interest in (or intrinsic value of) mathematics/science (e.g. 'I learn many interesting things in maths/science'), perceived utility (or extrinsic value) of mathematics/science (e.g. 'I need to do well in maths/science to get the job I want'), and their perceptions of their lessons and/or teachers (e.g. 'My teacher is easy to understand' in the context of mathematics/science lessons). These factors can be contextualised within motivational theories, such as the expectancy-value model of social-cognitive theory (Eccles, 2009) or self-determination theory (Deci & Ryan, 1985), and have been similarly measured in prior mathematics and science research (Bøe & Henriksen, 2015; Wang & Degol, 2013).

Table 1
Items measuring self-concept and theorised influences for the subjects of mathematics and science.

Item	Factor
I usually do well in [subject]	Self-concept
[Subject] is not one of my strengths	Self-concept
I learn things quickly in [subject]	Self-concept
I am good at working out difficult [subject] problems	Self-concept
[Subject] makes me confused and nervous	Anxiety
[Subject] is more difficult for me than for many of my classmates	Peer-comparison
[Subject] is harder for me than any other subject	Subject-comparison
My teacher thinks I can do well in [subject] classes with difficult materials	Teacher encouragement
My teacher tells me I am good at [subject]	Teacher encouragement

The students' reported gender was also considered, given prior research in mathematics and science (Bøe & Henriksen, 2015). Further background items/factors were considered in preliminary analysis (Appendix 1) but were ultimately not substantially/significantly predictive of students' mathematics or science self-concept beliefs when considered with the theorised influences (e.g. anxiety and peer-comparisons) and motivational factors (e.g. interest and utility). For brevity, these background items/factors were omitted from the final analysis.

For brevity and consistency, the analysis only considered the students' own reports. The considered items/factors were all at the student-level, and class-level/school-level aggregate or contextual factors were not considered.

2.2.1. Calculating factor scores/estimates

Some potential influences on self-concept (e.g. anxiety) were unavoidably covered through single or dual items (Table 1). All other factors were confirmed with single-factor structures (via confirmatory factor analysis) and acceptable indicators of reliability (Cronbach's α coefficients; Table 2). Items/factors were re-coded and/or calculated so that high magnitudes indicated a positive belief or experience (e.g. doing well, being interested, and the absence of anxiety).

Measures of students' self-concept, interest, utility, and lesson/teacher perceptions were estimated via one-parameter-logistic partial-credit item-response models (de Ayala, 2009; Rabe-Hesketh, Skrondal, & Pickles, 2004; Zheng & Rabe-Hesketh, 2007). Conceptually, this technique was akin to using structural equation modelling or factor analysis to estimate 'latent' factor-scores. Nevertheless, correlations and preliminary analysis highlighted that these 'latent' partial-credit-model factor-scores operated virtually identically to alternate factor-scores formed through 'observed' simple-averages of the relevant items (Appendix 1). Preliminary sensitivity analysis confirmed that the fundamental results and conclusions occurred regardless of how the factors were calculated.

2.2.2. Achievement estimates (task scores) and multiple imputation

The TIMSS questionnaires included numerous mathematics and science achievement tasks, which covered curricula areas from the majority of participating countries (Mullis, Martin, Ruddock, O'Sullivan, & Preuschoff, 2009); performance in TIMSS should then be relatively representative of students' achievement in classroom tests and other national examinations. Students' achievement (or task score) was measured via the five 'plausible-values' from the IEA (i.e. estimates from the IEA's item-response models that also inferred across any missing-by-design tasks; Martin & Mullis, 2013). The plausible-values were analytically handled via 'multiple imputation' software features, following guidelines to combine parameter estimates (Rubin, 1987).

2.3. Estimating confidence biases

Indicators of confidence bias (also referred to as calibration bias, or the degree of under-confidence through accuracy through to over-confidence) were calculated via the 'self-criteria residual' or 'regression-residual' approach (e.g. Gonida & Leondari, 2011): self-concept was predicted by achievement through regression models; the students' confidence bias was therefore shown by the regression-residual or the difference between their reported self-concept and their predicted self-concept, given the students' own particular achievement and the entire sample of students. This provided a 'relative' within-country indicator of confidence bias rather than an 'absolute' indicator, which would only be measurable when comparing confidence explicitly paired to achievement (e.g. someone's self-efficacy expressed as an expected score compared with their actual score).

Groups were then created via standardising the indicators (via z-scores formed across England): below $-.5$ was classified as 'under-confident'; between $-.5$ and $+.5$ as 'accurate' (one standard deviation range); and above $+.5$ as 'over-confident'. The number of students per

Table 2
Factor reliabilities (Cronbach's alpha coefficients).

Factor	Items	Mathematics	Science
Self-concept	4	.840	.858
Task score	Varies	.813	.766
Teacher encouragement	2	.728	.804
Interest (intrinsic value)	5	.877	.896
Utility (extrinsic value)	6	.774	.870
Lesson perception	5	.779	.804

Notes: Task score reliability was calculated as the mean Cronbach's alpha (α) coefficient across the rotated blocks of tasks; the equivalent reliability coefficients listed in IEA documentation (Martin & Mullis, 2013) show the median Cronbach's alpha coefficients.

group slightly varied across the indicators formed from the different plausible-values. Therefore, 'consistent' groups were formed, aggregating those students who were assigned to the same group (under-confident, accurate, or over-confident) across all five indicators. The stable student numbers for these resulting groups then allowed multiple imputation techniques to be applied (Rubin, 1987).

2.4. Predicting self-concept beliefs

Students' self-concept beliefs were predicted using the students' gender, achievement, theorised influences (reported subject-comparisons, peer-comparison, anxiety, and encouragement from teachers), and reported interest, utility, and lesson/teacher perceptions.

Missing values were minimal for these items/factors (i.e. a maximum of 2% missing per item/factor). Analysis only considered cases without any missing values (i.e. via 'listwise deletion', the default approach in most statistical software), and the majority of cases remained even with this approach (around 97%).

Various approaches were explored in preliminary sensitivity analysis, including ordinary-least-squares regression, multi-level (also called mixed or hierarchical) predictive modelling using different software (SPSS and STATA, each with different approaches to handling sampling-weights), multi-level modelling using different structures (i.e. variable-intercepts per-class and/or per-school), and different sampling-weight re-scaling approaches (Rabe-Hesketh & Skrondal, 2006; StataCorp, 2013). Estimated parameters were sufficiently similar across different approaches, excepting for ordinary-least-squares regression being slightly more lenient concerning significance. Most self-concept variance occurred at the student-level and school-level and less occurred at the class-level; when predicting self-concept for all students in England, essentially no class-level variance remained unexplained.

Multi-level predictive modelling was selected for the final analysis to account for students being clustered within groups (e.g. schools): students may be relatively-similar within groups (e.g. through sharing the same teachers, school environment, and geographical location); without accounting for such potential similarities, for example when using ordinary-least-squares regression, estimated standard errors (and hence p-values) may appear overly-low (Snijders & Bosker, 2012).

Given the preliminary analysis, and for efficiency, the final multi-level models used variable-intercepts per-school (not per-class and per-school, given little residual class-level variance in more-complex models), and used nationally re-scaled sample-weighting (i.e. IEA 'house-weighting'; Foy et al., 2013). Explained/unexplained variance was calculated as proportional reductions compared to models with no predictors (Snijders & Bosker, 2012). Effect sizes were calculated to represent Cohen's d when comparing predicted outcomes (self-concept) for students one standard deviation below and one standard deviation above the mean of the predicting item/factor (Tymms, 2004).

A rigorous criterion for significance ($p < .01$) was applied, given that further statistical techniques advised by the IEA were unavailable (i.e. jack-knife replication to further increase the precision of any estimated standard errors; Foy et al., 2013).

2.5. Predicting under-confident, accurately-evaluated, and over-confident self-concept beliefs

The students' self-concept beliefs were also predicted separately for each confidence bias group. Any differences across the groups (i.e. in coefficient magnitude and/or significance) could be plausibly (but indirectly) attributed to the confidence bias itself.

Significant differences across groups were identified through additional interactional models: two groups were modelled together (e.g. under-confident students and accurately-evaluating students); the various items/factors were included as predictors (as in the separate models), together with a group-membership indicator (e.g. accurately-evaluating = 1) and with the interactions between the group-membership indicator and the predictors. The significance associated with the interaction terms then highlighted differences in coefficient magnitude across the two groups.

3. Results

The correlations between the students' achievement (task scores) and their self-concept beliefs were relatively modest (mathematics: $R = .454$, $p < .001$; science: $R = .306$, $p < .001$). The imperfect correlation may reflect that: students form their beliefs in reference to diverse measures of achievement (e.g. examination grades and homework) that are only approximated by the IEA tasks; and/or students' beliefs are influenced by factors other than achievement; and/or some students are variously under-confident or over-confident (when compared to the IEA tasks at least). Accordingly, and unavoidably, there is some uncertainty and imprecision in any estimate of confidence biases, although this does not necessarily make the process invalid or without benefit.

For brevity, descriptive statistics are provided as supplementary material. Using regression-residuals to identify confidence biases ensured that the groups did not differ, on average, in their achievement but differed in their self-concept beliefs (i.e. reflecting the different degrees of confidence bias; Appendix 2). Under-confident students reported lower, and over-confident students reported higher, than accurately-evaluating students for the considered items/factors, including the students' interest and utility for mathematics and science (Appendix 2).

The correlations between items/factors for mathematics (Appendix 3) and science (Appendix 4) highlighted that the highest associations occurred between the students' interest and self-concept, perhaps highlighting that these may be more closely-related than previously assumed.

3.1. Predicting self-concept beliefs

For students across England (Table 3), the various theorised influences and the motivational factors were indeed predictive of the students' self-concept beliefs, over and above their achievement.

For mathematics, students' self-concept beliefs were most strongly predicted by their interest in mathematics, their perceived encouragement from teachers, and by their reported subject-comparisons

Table 3
Predicting subject-specific self-concept beliefs using subject-specific factors for all students in England.

Predictor (item/factor)	Mathematics				Science			
	Est.	SE	Sig.	Effect	Est.	SE	Sig.	Effect
Gender (being male)	**.339**	**.038**	**<.001**	**.173**	**.335**	**.039**	**<.001**	**.152**
Task score (PVs)	**.004**	**<.001**	**<.001**	**.389**	**.002**	**<.001**	**<.001**	**.159**
Subject-comparison	**.362**	**.024**	**<.001**	**.395**	**.278**	**.028**	**<.001**	**.238**
Peer-comparison	**.374**	**.025**	**<.001**	**.350**	**.452**	**.029**	**<.001**	**.366**
Anxiety (absence of)	**.175**	**.024**	**<.001**	**.170**	**.180**	**.028**	**<.001**	**.145**
Teacher encouragement	**.621**	**.028**	**<.001**	**.483**	**.838**	**.031**	**<.001**	**.610**
Interest (intrinsic value)	**.241**	**.014**	**<.001**	**.484**	**.260**	**.013**	**<.001**	**.572**
Utility (extrinsic value)	**.166**	**.017**	**<.001**	**.206**	**.105**	**.013**	**<.001**	**.171**
Lesson perception	−.025	.021	.242	−.030	.055	.022	.012	.065
Intercept	**−6.444**	**.165**	**<.001**	**NA**	**−6.207**	**.207**	**<.001**	**NA**
Explained variance	68.3%				72.3%			
Unexplained variance, school level	1.4%				1.3%			
Unexplained variance, residual	30.3%				26.5%			

Notes: The mathematic parameters show the effect of mathematics task score, mathematics subject-comparisons, mathematics interest, etc., when predicting mathematics self-concept. The science parameters show the effects of science task score, science subject-comparisons, science interest, etc., when predicting science self-concept. Estimated coefficients (Est.), standard errors (SE), significance (p-values; Sig.), and effect sizes (Effect) are shown. Significant predictors (at least $p < .01$) have been highlighted in bold. The factors were calculated via partial-credit-models. Analysis was undertaken with SPSS via linear mixed/multi-level models, multiple-imputation, and (nationally-rescaled) sample-weighting (i.e. 'house-weights').

(mathematics being harder/easier than other subjects). For science, students' self-concept beliefs were most strongly predicted by their perceived encouragement from teachers, their interest in science, and by their reported peer-comparisons (science thought to be harder/easier for the student than for their classmates).

Notable proportions of variance were explained by the (student-level) predictors. Little unexplained variance remained at the school-level, suggesting that the inclusion of aggregate or contextual school-level factors may not necessarily be helpful. The greater portion of unexplained variance occurred at the residual level (i.e. the student-level), nevertheless highlighting that further (unknown) factors are likely relevant.

3.2. Predicting under-confident, accurately-evaluated, and over-confident self-concept beliefs

When self-concept was predicted for the under-confident, accurately-evaluating, and over-confident students, various across-group differences in coefficient magnitude (at $p < .01$ via the additional interaction models) and/or significance were apparent.

For mathematics (Table 4) and for science (Table 5), the self-concept beliefs of students who accurately-evaluated were (unsurprisingly) primarily predicted by their achievement; by definition and the applied approach, accurate students were those with beliefs that were closely associated with their achievement. The group formation may have allowed achievement and self-concept to more-easily associate within-group compared to across all students. Nevertheless, informative results can be seen when considering the items/factors other than achievement.

For mathematics (Table 4), compared to accurate students, the self-concept beliefs of under-confident students were predicted more by their perceived peer-comparisons, perceived encouragement from their teachers, and interest in mathematics. Alternately, compared to accurate students, the self-concept beliefs of over-confident students were predicted more by their perceived peer-comparisons, teacher encouragement, and (extrinsic) utility of mathematics.

For science (Table 5), compared to accurate students, the self-concept beliefs of under-confident students were predicted more by their perceived peer-comparisons, perceived teacher encouragement, and by their interest in science (a similar pattern to mathematics). Compared to accurate students, the self-concept beliefs of over-confident students were predicted more by the (absence of) anxiety, perceived teacher encouragement, and by their interest in science.

Differences across the under-confident and the over-confident groups highlighted that, for mathematics and for science, the self-concept beliefs of under-confident students were predicted more by their interest. For mathematics, the self-concept beliefs of over-confident

Table 4
Predicting mathematics self-concept beliefs across confidence bias groups: under-confident, accurate, and over-confident groups.

Predictor (item/factor)	Under-confident (U)				Accurate (A)				Over-confident (O)			
	Est.	SE	Sig.	Effect	Est.	SE	Sig.	Effect	Est.	SE	Sig.	Effect
Gender (being male)	.086	.055	.116	.068	.056	.025	.025	.063	**.162**	**.060**	**.007**	**.122**
MAT task score (PVs)	^{UA}**.007**	**<.001**	**<.001**	**1.097**	^{UA,AO}**.009**	**<.001**	**<.001**	**1.642**	^{AO}**.007**	**<.001**	**<.001**	**1.051**
MAT subject-comparison	.057	.031	.064	.096	.054	.021	.014	.110	**.134**	**.046**	**.004**	**.165**
MAT peer-comparison	^{UA}**.166**	**.034**	**<.001**	**.237**	^{UA,AO}**.058**	**.017**	**.001**	**.104**	^{AO}**.175**	**.038**	**<.001**	**.235**
MAT anxiety (absence of)	**.111**	**.032**	**.001**	**.167**	.020	.018	.254	.039	.086	.041	.035	.114
MAT teacher encouragement	^{UA}**.303**	**.043**	**<.001**	**.361**	^{UA,AO}.041	.021	.057	.058	^{AO}**.312**	**.045**	**<.001**	**.340**
MAT interest (intrinsic value)	^{UA,UO}**.129**	**.020**	**<.001**	**.361**	^{UA}**.053**	**.014**	**.002**	**.193**	^{UO}**.053**	**.020**	**.008**	**.157**
MAT utility (extrinsic value)	^{UO}.051	.024	.032	.102	^{AO}.006	.016	.707	.016	^{UO,AO}**.173**	**.028**	**<.001**	**.293**
MAT lesson perception	.021	.031	.500	.038	−.006	.020	.766	−.014	−.011	.030	.709	−.022
Intercept	**−7.168**	**.225**	**<.001**	**NA**	**−5.007**	**.177**	**<.001**	**NA**	**−3.768**	**.225**	**<.001**	**NA**
Explained variance	60.9%				81.9%				57.4%			
Unexplained variance, school level	2.6%				.9%				2.3%			
Unexplained variance, residual	36.5%				17.2%				40.3%			
Group percentage of all students	32.6%				35.3%				32.1%			

Notes: Estimated coefficients (Est.), standard errors (SE), significance (p-values; Sig.), and effect sizes (Effect) are shown. Significant predictors (at least $p < .01$) have been highlighted in bold. The factors were calculated via partial-credit-models. Analysis was undertaken with SPSS via linear mixed/multi-level models, multiple-imputation, and (nationally-rescaled) sample-weighting (i.e. 'house-weights'). Significant differences in coefficient magnitude across the groups (at least $p < .01$, via additional paired-group interaction models) have been highlighted in superscript.

Table 5
Predicting science self-concept beliefs across confidence bias groups: under-confident, accurate, and over-confident groups.

Predictor (item/factor)	Under-confident (U)				Accurate (A)				Over-confident (O)			
	Est.	SE	Sig.	Effect	Est.	SE	Sig.	Effect	Est.	SE	Sig.	Effect
Gender (being male)	.107	.053	.043	.088	**.091**	**.028**	**.002**	**.120**	.154	.060	.010	.129
SCI task score (PVs)	[UA] **.004**	**<.001**	**<.001**	**.576**	[UA,AO] **.006**	**<.001**	**<.001**	**1.436**	[AO] **.004**	**<.001**	**<.001**	**.727**
SCI subject-comparison	[UO] **.133**	**.032**	**<.001**	**.215**	**.069**	**.023**	**.003**	**.137**	[UO] −.038	.045	.398	−.052
SCI peer-comparison	[UA] **.302**	**.035**	**<.001**	**.428**	[UA] **.085**	**.022**	**<.001**	**.166**	**.175**	**.046**	**<.001**	**.235**
SCI anxiety (absence of)	[UO] .066	.032	.040	.101	[AO] .036	.022	.104	.071	[UO,AO] **.233**	**.051**	**<.001**	**.290**
SCI teacher encouragement	[UA] **.388**	**.041**	**<.001**	**.460**	[UA,AO] **.145**	**.024**	**<.001**	**.237**	[AO] **.386**	**.049**	**<.001**	**.418**
SCI interest (intrinsic value)	[UA,UO] **.178**	**.019**	**<.001**	**.546**	[UA,AO] **.032**	**.011**	**.004**	**.159**	[UO,AO] **.094**	**.019**	**<.001**	**.320**
SCI utility (extrinsic value)	.042	.017	.014	.119	**.043**	**.011**	**<.001**	**.177**	**.073**	**.021**	**<.001**	**.194**
SCI lesson perception	.048	.030	.115	.089	.010	.018	.554	.029	.010	.029	.729	.020
Intercept	**−6.421**	**.261**	**<.001**	**NA**	**−4.236**	**.129**	**<.001**	**NA**	**−2.611**	**.280**	**<.001**	**NA**
Explained variance	55.8%				66.8%				44.9%			
Unexplained variance, school level	2.3%				.7%				2.9%			
Unexplained variance, residual	41.9%				32.5%				52.3%			
Group percentage of all students	32.4%				35.4%				32.2%			

Notes: Estimated coefficients (Est.), standard errors (SE), significance (*p*-values; Sig.), and effect sizes (Effect) are shown. Significant predictors (at least $p < .01$) have been highlighted in bold. The factors were calculated via partial-credit-models. Analysis was undertaken with SPSS via linear mixed/multi-level models, multiple-imputation, and (nationally-rescaled) sample-weighting (i.e. 'house-weights'). Significant differences in coefficient magnitude across the groups (at least $p < .01$, via additional paired-group interaction models) have been highlighted in superscript.

students were predicted more by their perceived utility of mathematics (compared to under-confident students). For science, the self-concept beliefs of under-confident students were predicted more by their subject-comparisons (compared to over-confident students, where subject-comparisons were non-significant). Additionally, for science, the self-concept beliefs of over-confident students were predicted more by their (absence of) anxiety when compared to the other groups.

4. Discussion

Students' confidence, measured as their self-concept beliefs, was predicted using various factors, integrating those from self-efficacy research (Usher & Pajares, 2008b), self-concept research (Bong & Skaalvik, 2003), and the expectancy-value model of social-cognitive theory (Eccles, 2009; Wang & Degol, 2013). Students' self-concept was strongly predicted by interest and praise, controlling for the students' achievement, which helps extend theoretical assumptions and operational models of key predictors. Additionally, the self-concept beliefs of students with different confidence biases were predicted by different factors in different ways. Such results offer increased insight into confidence biases, extending earlier research that has seldom applied predictive modelling (e.g. Sheldrake, Mujtaba, & Reiss, 2014).

4.1. Predicting self-concept beliefs

Across England, for mathematics and for science, students' self-concept beliefs were most strongly predicted by their interest and perceived praise (encouragement from teachers), while controlling for their achievement and further factors, and were least strongly predicted by anxiety.

Contrary to Hypothesis 1A, and differing from the magnitudes seen when predicting self-efficacy beliefs (Usher & Pajares, 2008b), the predictive association (effect size) of praise was higher than expected. Expressions of self-concept may inherently involve more subjectivity than expressions of self-efficacy (Bong & Clark, 1999; Bong & Skaalvik, 2003); it seems plausible to infer that students may consider different sources of information (such as praise) to a greater extent in order to evaluate whether they are 'good' or 'doing well' at mathematics or science.

The predictive associations between the students' self-concept beliefs and their reported peer-comparisons and subject-comparisons were relatively low in comparison to other modelled factors, contrary to Hypothesis 1A. Prior research has considered such effects in isolation via specific models (e.g. Marsh, Abduljabbar, et al., 2015; Marsh, Lüdtke, et al., 2015); for example, in England, large effects of peer-comparisons have been inferred through highlighting that students with the same achievement in different schools have reported different self-concept beliefs, but without controlling for any other factors (Nagengast & Marsh, 2011). Effect sizes may be inflated without controlling for further/mediating factors, and/or it may be difficult to directly compare research using different methodologies.

The students' interest was strongly predictive of their self-concept beliefs, controlling for other factors; utility was less-strongly predictive but still significant. This was as hypothesised (Hypothesis 1B), although the effect of interest was higher than expected. These results broadly cohere with earlier research highlighting various potential links between interest, confidence, and achievement (Köller, Baumert, & Schnabel, 2001; Marsh, Trautwein, Lüdtke, Köller, & Baumert, 2005; Viljaranta et al., 2014). Nevertheless, interest has perhaps been overlooked as a key predictor of self-concept.

On a theoretical level, the results help suggest why self-concept appears to have a motivational role. In previous research (e.g. Huang, 2011), higher self-concept has sometimes associated with higher future achievement even when controlling for prior achievement. The results presented above highlight that, when controlling for achievement and other factors, students with higher interest were predicted to express higher self-concept beliefs. Higher interest may explain any higher future achievement (e.g. Köller et al., 2001), perhaps through the students applying more effort, engagement, or persistence in their studies.

4.2. Predicting under-confident, accurately-evaluated, and over-confident self-concept beliefs

The results suggested that students with different confidence biases form their self-concept beliefs differently: predictors of self-concept varied in magnitude and/or significance across the groups. For example, for science, lower reported subject-comparisons (thinking science to be harder than other subjects) predicted lower science self-concept beliefs for under-confident students, controlling for their achievement and other factors; however, subject-comparisons were non-significant for over-confident students.

Contrary to Hypothesis 2A, the students' relative-comparisons (reported peer-comparisons and subject-comparisons) more-strongly predicted self-concept beliefs for under-confident science students than for other students. However, partly as hypothesised (Hypothesis 2A), peer-comparisons were equally-predictive of under-confident and over-confident self-concept beliefs in mathematics, which coheres with implications from prior research (e.g.

Marsh, Abduljabbar, et al., 2015). Nevertheless, peer-comparisons were less-strongly predictive for accurately-evaluating students than other students, which highlights that much remains unclear regarding any 'universal' effects of students' relative-comparisons.

Relative-comparisons may influence different students in different ways and perhaps operate differently across subjects, which may help extend and inform wider research (Marsh, Lüdtke, et al., 2015; Pohlmann & Möller, 2009). The results suggested some degree of asymmetry: science subject-comparisons predicted under-confident science self-concept beliefs, but not over-confident beliefs. Conversely, mathematics subject-comparisons only predicted over-confident mathematics self-concept beliefs. Nevertheless, various methodologies consider 'subject-comparisons' differently (e.g. as reported beliefs or as inferred from structural equation modelling), and it perhaps remains unclear what phenomenological processes occur or what reasoning students actually follow.

As hypothesised (Hypothesis 2B), the students' reported (extrinsic) utility of mathematics had a stronger predictive association with the self-concept beliefs of over-confident students, when compared to other students. However, group differences were not significant for science, although the pattern of coefficients broadly followed those seen for mathematics. Given the uncertainty, future research may need to clarify the area.

Higher (extrinsic) utility of mathematics (agreement with items such as 'I need to do well in mathematics to get into the university of my choice' and 'I need to do well in mathematics to get the job I want'), controlling for achievement, predicted higher over-confident self-concept beliefs; yet utility was non-significant for under-confident and for accurately-evaluating students. Problematically, some students may believe that they are 'doing well' partially because they think that they 'need to do well' in order to meet their future goals. Increased confidence may be motivationally beneficial, following social-cognitive theory (Bandura, 1997), but it would be (ultimately) detrimental if the students were sufficiently over-confident as to lack the achievement necessary to meet their goals.

Surprisingly, contrary to Hypothesis 2B, for mathematics and for science, the students' interest was more-strongly predictive of their self-concept beliefs for under-confident students when compared to other students. Further research may be necessary into whether increasing interest for some students may help address under-confidence.

4.3. Limitations and implications to subsequent research

Fundamentally, the various results described above cannot conclusively explain confidence biases. The IEA collected TIMSS data at a single time point, and the analytical models therefore only considered associations between concurrently-reported items/factors; it cannot therefore be concluded that peer-comparisons, subject-comparisons, anxiety, and other factors are necessarily temporal or causal antecedents to self-concept beliefs and/or particular confidence biases. Additionally, statistical association does not entail that phenomenological processes occur, which is perhaps under-emphasised in self-concept research. Following self-efficacy research (e.g. Butz & Usher, 2015), qualitative studies into self-concept beliefs may prove informative.

Confidence biases and groups can be explored and defined in various ways. Accordingly, the presented results are plausible (given the specific methods applied) but not definitive. Confidence biases can be explored via considering single measures of confidence and achievement via large-scale studies such as TIMSS (as presented here) or the Programme for International Student Assessment (PISA; e.g. Chiu & Klassen, 2010), but also via linking achievement tasks with confidence ratings (e.g. Chen & Zimmerman, 2007; Sheldrake, Mujtaba & Reiss, 2014). Results may potentially vary across methods, as different methods may provide different insights.

A number of predictors were unavoidably measured through single items. While single items may indeed be acceptable indicators (Gogol et al., 2014), their use increases dependence on the exact phrasing used. Future research likely needs to explore the area with more extensive item sets.

4.4. Conclusions and educational implications

For both mathematics and science, students' self-concept beliefs were most strongly predicted by their interest and perceived praise, over and above their achievement and other factors. On average, expressions of higher self-concept may reflect, partially, expressions of higher interest, which may help explain why self-concept (when considered alone in prior research) has appeared to be motivationally beneficial (e.g. Huang, 2011).

Various factors differentially predicted the self-concept beliefs of under-confident, accurate, and over-confident students in various ways. For example, for mathematics, utility predicted self-concept when students were over-confident but not when under-confident; for science, subject-comparisons predicted self-concept when under-confident but not when over-confident.

Understanding what predicts someone's self-concept when they are considered to be under-confident or over-confident may potentially help these confidence biases to be corrected by educators or even by the students themselves. Relative comparisons (across academic subjects and/or peers) may be somewhat unavoidable in educational systems, but educators could perhaps emphasise self-reflection or help highlight that students may be under-estimating themselves in some cases. Educators may also need to be aware that, potentially problematically, some students may believe that they are 'doing well' partially because they think that they 'need to do well' in order to meet their future goals. While higher confidence may be motivationally beneficial, it would be ultimately detrimental if the students were sufficiently over-confident as to lack the achievement necessary to meet their goals.

Acknowledgements

This work was supported by the Economic and Social Research Council. Please contact the author regarding access to the underlying research materials if required.

Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.lindif.2016.05.009.

References

Arens, A. K., Seeshing Yeung, A., Craven, R., & Hasselhorn, M. (2011). The twofold multidimensionality of academic self-concept: Domain specificity and separation between competence and affect components. *Journal of Educational Psychology*, *103*(4), 970–981. http://dx.doi.org/10.1037/a0025047.

Bandura, A. (1997). *Self-efficacy: The exercise of control.* New York: Freeman.

Bøe, M. V., & Henriksen, E. K. (2015). Expectancy-value perspectives on choice of science and technology education in late-modern societies. In E. K. Henriksen, J. Dillon, & J. Ryder (Eds.), *Understanding student participation and choice in science and technology education* (pp. 17–29). Dordrecht: Springer. http://dx.doi.org/10.1007/978-94-007-7793-4_2.

Bong, M., & Clark, R. (1999). Comparison between self-concept and self-efficacy in academic motivation research. *Educational Psychologist*, *34*(3), 139–153. http://dx.doi.org/10.1207/s15326985ep3403_1.

Bong, M., & Skaalvik, E. (2003). Academic self-concept and self-efficacy: How different are they really? *Educational Psychology Review*, *15*(1), 1–40. http://dx.doi.org/10.1023/A:1021302408382.

Bouffard, T., & Narciss, S. (2011). Benefits and risks of positive biases in self-evaluation of academic competence: Introduction. *International Journal of Educational Research*, *50*(4), 205–208. http://dx.doi.org/10.1016/j.ijer.2011.08.001.

Bouffard, T., Vezeau, C., Roy, M., & Lengelé, A. (2011). Stability of biases in self-evaluation and relations to well-being among elementary school children. *International Journal of Educational Research*, *50*(4), 221–229. http://dx.doi.org/10.1016/j.ijer.2011.08.003.

Britner, S. (2008). Motivation in high school science students: A comparison of gender differences in life, physical, and earth science classes. *Journal of Research in Science Teaching*, *45*(8), 955–970. http://dx.doi.org/10.1002/tea.20249.

Britner, S., & Pajares, F. (2006). Sources of science self-efficacy beliefs of middle school students. *Journal of Research in Science Teaching*, *43*(5), 485–499. http://dx.doi.org/10.1002/tea.20131.

Butz, A., & Usher, E. (2015). Salient sources of early adolescents' self-efficacy in two domains. *Contemporary Educational Psychology*, *42*, 49–61. http://dx.doi.org/10.1016/j.cedpsych.2015.04.001.

Chen, P. (2003). Exploring the accuracy and predictability of the self-efficacy beliefs of seventh-grade mathematics students. *Learning and Individual Differences*, *14*(1), 79–92. http://dx.doi.org/10.1016/j.lindif.2003.08.003.

Chen, P., & Zimmerman, B. (2007). A cross-national comparison study on the accuracy of self-efficacy beliefs of middle-school mathematics students. *The Journal of Experimental Education*, *75*(3), 221–244. http://dx.doi.org/10.3200/JEXE.75.3.221-244.

Chiu, M. -S. (2012). The internal/external frame of reference model, big-fish-little-pond effect, and combined model for mathematics and science. *Journal of Educational Psychology*, *104*(1), 87–107. http://dx.doi.org/10.1037/a0025734.

Chiu, M. M., & Klassen, R. (2010). Relations of mathematics self-concept and its calibration with mathematics achievement: Cultural differences among fifteen-year-olds in 34 countries. *Learning and Instruction*, *20*(1), 2–17. http://dx.doi.org/10.1016/j.learninstruc.2008.11.002.

de Ayala, R. J. (2009). *The theory and practice of item response theory*. New York: The Guilford Press.

Deci, E., & Ryan, R. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum.

EACEA [Education, Audiovisual, and Culture Executive Agency] (2011y). *Science education in Europe: National policies, practices and research*. Brussels: European Commission. http://dx.doi.org/10.2797/7170.

Eccles, J. (2009). Who am I and what am I going to do with my life? Personal and collective identities as motivators of action. *Educational Psychologist*, *44*(2), 78–89. http://dx.doi.org/10.1080/00461520902832368.

Foy, P., Arora, A., & Stanco, G. (Eds.). (2013). *TIMSS 2011 user guide for the international database*. Chestnut Hill, Massachusetts: TIMSS & PIRLS International Study Centre.

Gogol, K., Brunner, M., Goetz, T., Martin, R., Ugen, S., Keller, U., ... Preckel, F. (2014). “My questionnaire is too long!” The assessments of motivational-affective constructs with three-item and single-item measures. *Contemporary Educational Psychology*, *39*(3), 188–205. http://dx.doi.org/10.1016/j.cedpsych.2014.04.002.

Gonida, E., & Leondari, A. (2011). Patterns of motivation among adolescents with biased and accurate self-efficacy beliefs. *International Journal of Educational Research*, *50*(4), 209–220. http://dx.doi.org/10.1016/j.ijer.2011.08.002.

Huang, C. (2011). Self-concept and academic achievement: A meta-analysis of longitudinal relations. *Journal of School Psychology*, *49*(5), 505–528. http://dx.doi.org/10.1016/j.jsp.2011.07.001.

Huguet, P., Dumas, F., Marsh, H., Régner, I., Wheeler, L., Suls, J., ... Nezlek, J. (2009). Clarifying the role of social comparison in the big-fish–little-pond effect (BFLPE): An integrative study. *Journal of Personality and Social Psychology*, *97*(1), 156–170. http://dx.doi.org/10.1037/a0015558.

Jiang, Y., Song, J., Lee, M., & Bong, M. (2014). Self-efficacy and achievement goals as motivational links between perceived contexts and achievement. *Educational Psychology*, *34*(1), 92–117. http://dx.doi.org/10.1080/01443410.2013.863831.

Köller, O., Baumert, J., & Schnabel, K. (2001). Does interest matter: The relationship between academic interest and achievement in mathematics. *Journal for Research in Mathematics Education*, *32*(5), 448–470. http://dx.doi.org/10.2307/749801.

Marsh, H. (1986). Verbal and math self-concepts: An internal/external frame of reference model. *American Educational Research Journal*, *23*(1), 129–149. http://dx.doi.org/10.3102/00028312023001129.

Marsh, H., & Martin, A. (2011). Academic self-concept and academic achievement: Relations and causal ordering. *British Journal of Educational Psychology*, *81*(1), 59–77. http://dx.doi.org/10.1348/000709910X503501.

Marsh, H., & Parker, J. (1984). Determinants of student self-concept: Is it better to be a relatively large fish in a small pond even if you don't learn to swim as well? *Journal of Personality and Social Psychology*, *47*(1), 213–231. http://dx.doi.org/10.1037/0022-3514.47.1.213.

Marsh, H., Abduljabbar, A. S., Morin, A., Parker, P., Abdelfattah, F., Nagengast, B., & Abu-Hilal, M. (2015a). The big-fish-little-pond effect: Generalizability of social comparison processes over two age cohorts from Western, Asian, and Middle Eastern Islamic countries. *Journal of Educational Psychology*, *107*(1), 258–271. http://dx.doi.org/10.1037/a0037485.

Marsh, H., Lüdtke, O., Nagengast, B., Trautwein, U., Abduljabbar, A. S., Abdelfattah, F., & Jansen, M. (2015b). Dimensional Comparison Theory: Paradoxical relations between self-beliefs and achievements in multiple domains. *Learning and Instruction*, *35*, 16–32. http://dx.doi.org/10.1016/j.learninstruc.2014.08.005.

Marsh, H., Trautwein, U., Lüdtke, O., Köller, O., & Baumert, J. (2005). Academic self-concept, interest, grades, and standardized test scores: Reciprocal effects models of causal ordering. *Child Development*, *76*(2), 397–416. http://dx.doi.org/10.1111/j.1467-8624.2005.00853.x.

Martin, M., & Mullis, I. (Eds.). (2013). *Methods and procedures in TIMSS and PIRLS 2011*. Chestnut Hill, Massachusetts: TIMSS & PIRLS International Study Centre Retrieved from http://timssandpirls.bc.edu/methods/

Möller, J., & Pohlmann, B. (2010). Achievement differences and self-concept differences: Stronger associations for above or below average students? *British Journal of Educational Psychology*, *80*(3), 435–450. http://dx.doi.org/10.1348/000709909X485234.

Morony, S., Kleitman, S., Lee, Y. P., & Stankov, L. (2013). Predicting achievement: Confidence vs self-efficacy, anxiety, and self-concept in Confucian and European countries. *International Journal of Educational Research*, *58*, 79–96. http://dx.doi.org/10.1016/j.ijer.2012.11.002.

Mullis, I., Martin, M., Ruddock, G., O'Sullivan, C., & Preuschoff, C. (2009). *TIMSS 2011 assessment frameworks*. Chestnut Hill, Massachusetts: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.

Multon, K., Brown, S., & Lent, R. (1991). Relation of self-efficacy beliefs to academic outcomes: A meta-analytic investigation. *Journal of Counseling Psychology*, *38*(1), 30–38. http://dx.doi.org/10.1037/0022-0167.38.1.30.

Nagengast, B., & Marsh, H. (2011). The negative effect of school-average ability on science self-concept in the UK, the UK countries and the world: The Big-Fish-Little-Pond-Effect for PISA 2006. *Educational Psychology*, *31*(5), 629–659. http://dx.doi.org/10.1080/01443410.2011.586416.

OECD (2015). *The ABC of gender equality in education: Aptitude, behaviour, confidence*. Paris: OECD Publishing. http://dx.doi.org/10.1787/9789264229945-en.

Pajares, F., & Graham, L. (1999). Self-efficacy, motivation constructs, and mathematics performance of entering middle school students. *Contemporary Educational Psychology*, *24*(2), 124–139. http://dx.doi.org/10.1006/ceps.1998.0991.

Phillips, J., & Gully, S. (1997). Role of goal orientation, ability, need for achievement, and locus of control in the self-efficacy and goal-setting process. *Journal of Applied Psychology*, *82*(5), 792–802. http://dx.doi.org/10.1037/0021-9010.82.5.792.

Pinxten, M., Wouters, S., Preckel, F., Niepel, C., De Fraine, B., & Verschueren, K. (2015). The formation of academic self-concept in elementary education: A unifying model for external and internal comparisons. *Contemporary Educational Psychology*, *41*, 124–132. http://dx.doi.org/10.1016/j.cedpsych.2014.12.003.

Pohlmann, B., & Möller, J. (2009). On the benefit of dimensional comparisons. *Journal of Educational Psychology*, *101*(1), 248–258. http://dx.doi.org/10.1037/a0013151.

Rabe-Hesketh, S., & Skrondal, A. (2006). Multilevel modelling of complex survey data. *Journal of the Royal Statistical Society, Series A*, *169*(4), 805–827. http://dx.doi.org/10.1111/j.1467-985X.2006.00426.x.

Rabe-Hesketh, S., Skrondal, A., & Pickles, A. (2004). *GLLAMM manual*. Berkeley, California: University of California, Berkeley Retrieved June 24, 2014, from http://biostats.bepress.com/ucbbiostat/paper160

Regan, E., & DeWitt, J. (2015). Attitudes, interest and factors influencing STEM enrolment behaviour: An overview of relevant literature. In E. K. Henriksen, J. Dillon, & J. Ryder (Eds.), *Understanding student participation and choice in science and technology education* (pp. 63–88). Dordrecht: Springer. http://dx.doi.org/10.1007/978-94-007-7793-4_5.

Rubin, D. (1987). *Multiple imputation for nonresponse in surveys*. New York: John Wiley & Sons. http://dx.doi.org/10.1002/9780470316696.

Seaton, M., Parker, P., Marsh, H., Craven, R., & Yeung, A. S. (2014). The reciprocal relations between self-concept, motivation and achievement: Juxtaposing academic self-concept and achievement goal orientations for mathematics success. *Educational Psychology*, *34*(1), 49–72. http://dx.doi.org/10.1080/01443410.2013.825232.

Shavelson, R., Hubner, J., & Stanton, G. (1976). Self-concept: Validation of construct interpretations. *Review of Educational Research*, *46*(3), 407–441. http://dx.doi.org/10.2307/1170010.

Snijders, T., & Bosker, R. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling* (2nd ed.). London: SAGE Publications.

Sheldrake, R., Mujtaba, T., & Reiss, M. J. (2014). Calibration of self-evaluations of mathematical ability for students in England aged 13 and 15, and their intentions to study non-compulsory mathematics after age 16. *International Journal of Educational Research*, *64*, 49–61. http://dx.doi.org/10.1016/j.ijer.2013.10.008.

Stankov, L. (2010). Unforgiving Confucian culture: A breeding ground for high academic achievement, test anxiety and self-doubt? *Learning and Individual Differences*, *20*(6), 555–563. http://dx.doi.org/10.1016/j.lindif.2010.05.003.

Stankov, L., & Lee, J. (2014). Overconfidence across world regions. *Journal of Cross-Cultural Psychology*, *45*(5), 821–837. http://dx.doi.org/10.1177/0022022114527345.

StataCorp (2013). *STATA multilevel mixed-effects reference manual release 13*. College Station, Texas: Stata Press Retrieved July 1, 2015, from http://www.stata.com/manuals13/me.pdf

The Royal Society (2014). *Vision for science and mathematics education*. London: The Royal Society.

Thijs, J., Verkuyten, M., & Helmond, P. (2010). A further examination of the Big-Fish–Little-Pond Effect: Perceived position in class, class size, and gender comparisons. *Sociology of Education*, *83*(4), 333–345. http://dx.doi.org/10.1177/0038040710383521.

Tymms, P. (2004). Effect sizes in multilevel models. In I. Schagen, & K. Elliot (Eds.), *But what does it mean? The use of effect sizes in educational research* (pp. 55–66). Slough: National Foundation for Educational Research.

Usher, E., & Pajares, F. (2008a). Self-efficacy for self-regulated learning: A validation study. *Educational and Psychological Measurement*, *68*(2), 443–463. http://dx.doi.org/10.1177/0013164407308475.

Usher, E., & Pajares, F. (2008b). Sources of self-efficacy in school: Critical review of the literature and future directions. *Review of Educational Research*, *78*(4), 751–796. http://dx.doi.org/10.3102/0034654308321456.

Usher, E., & Pajares, F. (2009). Sources of self-efficacy in mathematics: A validation study. *Contemporary Educational Psychology*, *34*(1), 89–101. http://dx.doi.org/10.1016/j.cedpsych.2008.09.002.

Viljaranta, J., Tolvanen, A., Aunola, K., & Nurmi, J. -E. (2014). The developmental dynamics between interest, self-concept of ability, and academic performance. *Scandinavian Journal of Educational Research*, *58*(6), 734–756. http://dx.doi.org/10.1080/00313831.2014.904419.

Wang, M. -T., & Degol, J. (2013). Motivational pathways to STEM career choices: Using expectancy-value perspective to understand individual and gender differences in STEM fields. *Developmental Review*, *33*(4), 304–340. http://dx.doi.org/10.1016/j.dr.2013.08.001.

Zheng, X., & Rabe-Hesketh, S. (2007). Estimating parameters of dichotomous and ordinal item response models with gllamm. *The Stata Journal*, *7*(3), 313–333.

Zimmerman, B., & Schunk, D. (2011). Self-regulated learning and performance: An introduction and an overview. In B. Zimmerman, & D. Schunk (Eds.), *Handbook of self-regulation of learning and performance* (pp. 1–12). Abingdon: Routledge.