International Journal of Educational Research 121 (2023) 102215

Contents lists available at ScienceDirect

International Journal of Educational Research

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

The big-fish-little-pond effect and overclaiming

John Jerrim [a,c,*], Philip D Parker [b], Nikki Shure [c]

[a] UCL Social Research Institute, University College London, 20 Bedford Way London, WC1H 0AL, UK
[b] Australian Catholic University, Australia
[c] UCL Social Research Institute, University College London, UK

ARTICLE INFO

Keywords:
PISA
Overclaiming
Big-fish-little-pond effect
Relative achievement

ABSTRACT

Using the OECD's Programme for International Student Assessment (PISA) study, we investigate whether students' *relative* ability in mathematics (in comparison to their school peers) is linked to their tendency to overclaim. Although the estimated effect size is modest (around 0.1 standard deviations) we find empirical support that being a big fish in a small pond is linked to overclaiming, with this robust to different analytic approaches and model specifications. Thus, being one of the highest academic achievers within a school may push young people's beliefs in their own abilities too far, straying into overconfidence and making claims about their knowledge and skills that they cannot justify.

1. Introduction

The Big-Fish-Little-Pond (BFLP) effect is a widely known psychological phenomena. It postulates that young people use their peers as a frame of reference to make inferences about their own academic abilities. Thus, one's ***relative*** ability in a subject in comparison to their school or class peers contributes to their own academic self-concept. The magnitude and direction of the association between one's relative academic position and their self-concept is driven by two countervailing forces – assimilation effects and contrast effects (Bless & Burger, 2016). While the former (assimilation effects) drives a young person's self-concept towards those of their peers over time, the latter (contrast effects) pushes it further away. Empirically, contrast effects – young people tending to become more confident when they are higher achieving than their peers – tend to be the larger of the two (Wouters et al., 2013). This in-turn means that being one of the higher academic achievers within a school boosts a young person's academic self-belief, over and above their absolute level of academic achievement.

This phenomenon has become one of the most widely studied – and empirically validated – psychological constructs over the last half century. Starting with the pioneering work of Davis (1966), Marsh and Parker (1984) and Franks (1986) it has become well-established that young people's self-efficacy is not only linked to their own academic abilities, but also their academic abilities relative to those of their school peers. Decades of subsequent research – see Fang et al. (2018) for a review - has illustrated the

Using the OECD's Programme for International Student Assessment (PISA) study, we investigate whether students' *relative* ability in mathematics (in comparison to their school peers) is linked to their tendency to overclaim. Although the estimated effect size is modest (around 0.1 standard deviations) we find empirical support that being a big fish in a small pond is linked to overclaiming, with this robust to different analytic approaches and model specifications. Thus, being one of the highest academic achievers within a school may push young people's beliefs in their own abilities too far, straying into overconfidence and making claims about their knowledge and skills that they cannot justify.

* Corresponding author.
E-mail address: J.Jerrim@ucl.ac.uk (J. Jerrim).

https://doi.org/10.1016/j.ijer.2023.102215
Received 31 March 2023; Received in revised form 25 May 2023; Accepted 8 July 2023
Available online 18 July 2023
0883-0355/

robustness of this pattern across different cultural settings (Marsh & Hau, 2003; Nagengast & Marsh, 2011), ages and school stages (Becker & Neumann, 2016), longitudinal versus cross-sectional analyses, and methodological approach (Marsh et al., 2008). It has more recently been a topic explored by academics from other social science disciplines, including economists (e.g. Denning et al., 2018; Elsner & Isphording, 2017; Murphy & Weinhardt, 2020), whose focus on young people's "rank position" within school has largely confirmed findings from the psychological literature (Delaney & Devereux, 2022). Yet both psychologists and economics have further extended work on BFLP and rank position effects to other outcomes, finding that relative academic position in school is linked to university access and completion, earnings during early adulthood and mental health (Denning et al., 2018; Kiessling & Norris, 2020).

More recently, a smaller psychological literature has emerged into the phenomena of "overclaiming". In other words, the tendency for certain individuals to claim they know more about a topic than they really – or could possibly – do. This has typically been measured by asking respondents about their knowledge of a list of concepts, consumer products or facts within a social survey. Most of the items included within the list are real, but some are fake (sometimes known as "foils"). It is how participants respond to these fake items – the extent that they claim knowledge of things that do not really exist - that determine their position on an overall overclaiming scale.

Several studies within this literature have considered exactly what such "overclaiming" measures. Paulhus et al. (2003, p899) argue that overclaiming is a "*trait-like tendency to self-enhance*", demonstrating it to be correlated with self-esteem, ego-resiliency, and self-rated adjustment. However, Goecke et al. (2020) suggest four different perspectives that may be taken about overclaiming – that it (a) reflects self-enhancement tendencies, (b) is the result of a cognitive bias, (c) that it proxies cognitive abilities and (d) is a sign of creative engagement. They go on to argue that "*overclaiming as a phenomenon… requires more than self-enhancement motivation*" (Goecke et al., 2020, p1). Nevertheless, the general view that has emerged is that "*the overclaiming technique has proven itself as an efficient and robust method for indexing self-enhancement*" (Paulhus & Dubois, 2014, p979).

This has then led to several studies considering how overclaiming is linked to other psychological constructs, as well as cognitive skills. For instance, Keller et al. (2021) investigated whether overclaiming is linked to "dark triad personality traits" - e.g. narcissism, psychopathy, and Machiavellianism – finding no evidence that this is the case. Dunlop et al. (2020) relate overclaiming to "faking" in the context of job applications, finding evidence that applicants are most likely to overclaim when asked about job-relevant content. Rather, the same authors find that overclaiming is linked to "openness" (Dunlop et al., 2017). A recent study by Salvi et al. (2022) linked overclaiming to insightfulness, arguing that individuals with more insight are less likely to overclaim. Ackerman and Ellingsen (2014) explore the association between overclaiming and vocabulary abilities, concluding that "*higher ability individuals tended to claim more knowledge, contradicting the assertions that low-ability individuals are more likely than higher ability individuals to overstate their capabilities*" (Ackerman & Ellingsen, 2014, p225).

The inclusion of overclaiming questions in the 2012 edition of the Programme for International Assessment (PISA) has led to further interest into overclaiming amongst school-aged children. This study asked 15-year-olds across the world their knowledge about a series of mathematical constructs, three of which were fake. Various studies have also considered the measurement properties of the overclaiming questions included in PISA 2012 (e.g. Jerrim et al., 2023; Jin et al., 2023; Muszyński et al., 2021) arguing that students responses to the three foil items reflects a combination of their "*knowledge accuracy and their knowledge bias*" (Jin et al., 2023, p230). Yang et al. (2019) have used the PISA 2012 data to explore how overclaiming is linked to PISA scores and student engagement in the United States, finding that the least engaged students are the least likely to overclaim. Likewise, Fell et al. (2019) explore how overclaiming at the country level (as measured using data from PISA 2012) is linked with the extent that the population tolerates the violation of rules, finding a positive correlation.

However, a notable gap is that no existing study has attempted to bring the literature on BFLP effects and overclaiming together. That is, no previous work has empirically investigated whether one's relative academic position within their school is related to their propensity to "overclaim" – i.e. whether Big-Fish-Little-Pond effects can be observed for this important and interesting psychological outcome. This is despite clear theoretical reasons to suspect this might be the case; if a young person is the best at something in their school, they may start to believe they know everything possible about a topic – more than they really do.

The primary contribution of this paper is to hence make this addition to these two currently distinct literatures, being the first study to empirically investigate the link between relative academic position in school (i.e. BFLP effects) and the propensity for teenagers to overclaim. We do so by using international data from the OECD's Programme for International Student Assessment (PISA) on seven predominantly English-speaking countries. The advantages of using PISA data are its large sample size and the rich measures on skills and overclaiming. Since PISA has a school-based survey design, we observe a set of school peers for each participating individual, which allows us to compare their performance relative to their peers and compute a measure of relative performance. We then construct an index of overclaiming using self-reported ability measures on a range of mathematical constructs where, importantly, three fake constructs were included.

We answer the following research questions: is there a relationship between relative performance on PISA within a school and the propensity to overclaim? To do this we estimate a series of linear regression models where we regress our overclaiming index on a young person's relative performance on PISA within their school, controlling for a range of rich demographic and other control variables. These regression results confirm the existence of a moderately sized BFLP effect for overclaiming (∼0.1 SD).

The rest of this paper is structured as follows. In Section 2 we delve deeper into the theoretical considerations behind why we might expect there to be a BFLP effect on overclaiming. In Section 3 we discuss the data used. In Section 4 we discuss how we construct the key measures used in this paper and our empirical analysis. Results are presented in Section 5, with conclusions in Section 6.

2. Why might students' achievement relative to their school peers be associated with the propensity to overclaim?

Big-Fish-Little-Pond effects – the link between relative achievement compared to school peers and future outcomes – are thought to be the result of two countervailing forces. These are "assimilation" and "contrast" effects. The former (assimilation) refers to an acculturation process, where a member of a group (e.g. a student within a school) becomes more like other members of the group over time. This is because they will develop joint shared interests, will tend to act like others to "fit in" (Phalet & Baysu, 2020)– i.e. will experience social group pressures (Makarova, 2019) – and be exposed to the same school resources and culture (e.g. disciplinary climate) that will, over time, tend to make the group more and more alike.

Contrast effects, on the other hand, work in the opposite direction. Individuals will tend to judge themselves and their skills and attributes to those of other members of the group, which will lead them to develop different views, feelings and opinions from their peers. For instance, in the classic example of self-efficacy, contrast effects lead young people to evaluate their academic skills relative to the average of their school or class peer group. Those who are amongst the highest achievers within the group then evaluate themselves more highly than the lower-achieving members of the group, and thus develop higher levels of self-efficacy. As assimilation and contrast effects work in the opposite directions, it becomes an empirical question as to whether the aggregate impact is positive or negative. Most previous research into BFLP and school "rank position" effects suggest that contrast effects tend to be the dominant of the two - hence the consensus view that it is beneficial to be a big academic fish within a smaller school pond.

Why, then, might one expect there to be a link between relative achievement and the tendency for young people to "overclaim" (to declare they know more about a topic than they really do)? We suggest four potential mechanisms.

First, it is possible that contrast effects push young people's self-belief in their skills too far. This could be driven by social comparison bias such as the "better-than-average effect" (Larrick et al. 2007). In particular, teenagers may make comparisons to their school peers, with those above the average developing a view that they are superior to others. Social comparison effects such as the better-than-average effects have previously been reported amongst secondary school children (Kuyper et al., 2011). Being the top academic achiever within their group (school) could thus lead to teenagers becoming *overconfident* in their skills and abilities, leading them to believe that they know more about a topic than they could possibly do. In essence, BFLP effects could push young people's self-efficacy up too far.

Second, relatedly, being a high-achiever within a low-performing school could lead to the development of other psychological or attitudinal traits. For instance, previous work has linked overclaiming to narcissism (see Grosz et al. (2017) for an overview), with the potential for teenagers to become arrogant "know it alls" when they excel at something relative to their peers. Previous research into BFLP effects have also suggested that they are moderated by personality factors such as narcissism (Jonkmann et al., 2012). This could either lead comparatively high-achievers to assume they know everything there is to know about a certain issue, or to become careless/complacent when asked about their skills - e.g. simply ticking that they "know everything" about the issue when asked (see Yates et al., 1997, for a discussion of this literature).

Third is the role of expectations. Being amongst the highest achievers within their school, young people may feel that others (teachers, parents, peers) expect them to know about certain issues. Indeed, as noted by Mueller and Winsor (2016, p241), previous research has suggested that "*high ability students may suffer from unrealistically high social (*i.e.*, parental, teacher, peer) performance expectations, and exaggerated social pressures to excel*". Young people may then internalise these expectations, with previous research finding that this then increases their academic self-efficacy levels (Yamamoto & Holloway, 2010). This could lead teenagers to report that they know about certain topics because they think others will expect them to. This, in many ways, reflects a social desirability to their responses, reporting the option they think others would expect of them (Krumpal, 2013).

Finally, perhaps such a relationship could exist because of the low self-efficacy of relatively low-achieving students. This group may assume they know very little about a subject because their school peers are so much more able than they are. Indeed, recent research from England suggests that students in lower-ability groups are less confident in their mathematics skills than their peers in higher-ability groups, with the difference growing as they progress through school (Francis et al., 2020). This would, in-turn, make low-achieving pupils within the school very unlikely to overclaim. Although this would run counter to the Dunning-Kruger effect (Mahmood, 2016) – which postulates that it is the least able who are most likely to overestimate their abilities relative to others – it has recently been argued that this is driven by statistical artifacts (Magnus & Peresetsky, 2022). On the other hand, average and high achievers within the same school may be willing to take more risks and venture they may know *something* about a topic, even if they are not completely sure. This would be consistent with recent research that suggests higher-achievers are more willing to offer a guess at a question, even when not sure of the answer (Iriberri & Rey-Biel, 2021).

Importantly, each of these potential mechanisms suggest that relatively high achievers within a school are more likely to exaggerate their knowledge than relatively low achievers (over and above their actual academic abilities). We thus empirically test the direction and strength of this association between relative achievement and overclaiming by asking:

Research Question. Is there an association between relative academic position within a school and the propensity for young people to "overclaim"?

3. Data

The data we use are drawn from the 2012 round of the Programme for International Student Assessment (PISA) (OECD, 2014). This round of data is chosen as the most recent where the relevant information to construct our key variables is available. PISA is a cross-national study of 15-year-olds that provides internationally comparable data on mathematics, science and reading achievement. We focus on seven predominantly English-speaking countries (Australia, Canada, Ireland, New Zealand, Scotland, United Kingdom -

Table 1
The distribution of responses to the mathematics constructs (row percentages).

Constructs	Never heard of it	Heard of it once or twice	Heard of it a few times	Heard of it often	Know it well - understand the concept
Exponential function	45	16	16	12	11
Divisor	30	20	18	15	18
Quadratic function	21	13	17	21	28
*Proper number**	*16*	17	22	23	22
Linear equation	11	9	13	24	42
Vectors	36	18	18	14	13
Complex number	22	21	23	19	15
Rational number	17	16	20	22	25
Radicals	31	20	19	15	15
*Subjunctive scaling**	*64*	16	11	6	3
Polygon	13	9	13	20	44
*Declarative fraction**	*59*	18	13	7	4
Congruent figure	37	15	14	13	21
Cosine	36	8	8	13	34
Arithmetic mean	42	15	13	12	18
Probability	5	5	8	19	62

Notes: Figures refer to row percentages such that each row sums to 100. Constructs in italics with asterisk are the three fake ones. Source: Authors' calculations using PISA 2012 data (OECD, 2014).

England, Northern Ireland, and Wales- and United States), due to uncertainty over how the fake mathematics constructs we use would have been interpreted in other languages (further details provided below). Each of these countries draw nationally representative samples via a two-stage sample design, with schools first selected with probability proportional to size and then around 35 students randomly drawn within each school. Response rates in most of our countries were reasonably high (averaging 92% for schools[1] and 85% for students). Final sample sizes were higher in Canada, Australia and the UK[2] than other countries due to oversampling of certain geographic regions – see Appendix A for further details. We apply student senate weights to ensure each country contributes equally to our results.

One unique feature of PISA 2012 is that students were randomly assigned to complete one of three separate questionnaire forms. Only two (form A and form C) include the measure used to construct our overclaiming index. After creating any relevant school level variables (including students relative achievement compared to their school peers) we restrict the analytic sample to students who completed forms A and C. Given students were randomly assigned to forms, the part of the sample we drop can be considered Missing Completely At Random (MCAR) and not introduce bias into our results. The final analytical sample is 40,429 students. See Appendix A for further details.

3.1. Measurement of overclaiming

As part of the background questionnaire students were asked the following question:

"Thinking about mathematical concepts: how familiar are you with the following terms?"

To which there were five response options: never heard of it, heard of it once or twice, heard of it a few times, heard of it often, know it well – understand the concept. This was accompanied by the following 16 mathematical constructs:

1 Exponential function
2 Divisor
3 Quadratic function
4 *Proper number**
5 Linear equation
6 Vectors
7 Complex number
8 Rational number
9 Radicals
10 *Subjunctive scaling**
11 Polygon
12 *Declarative fraction**
13 Congruent figure
14 Cosine
15 Arithmetic mean

[1] After replacement schools included.

[2] As Scotland participates in PISA as a separate sub-national entity, we treat Scotland and the "rest of the UK" as two separate countries.

Table 2
Crosstabulation between relative mathematics achievement within school and absolute mathematics achievement within their country.

		Within-school achievement tercile		
		Bottom third	Middle third	Top third
Mathematics achievement across whole country tercile	Bottom third	74	20	2
	Middle third	23	57	20
	Top third	2	24	79
	Total	100	100	100

Notes: Figures refer to column percentages. For example, 74% of students who are a low achiever within their school (bottom third within school) are also a low achiever when compared across the whole country. In contrast, 2% of individuals who are low achievers within their school are high achievers when compared across their country as a whole. Analysis based upon first plausible value only. Source: Authors' calculations using PISA 2012 data (OECD, 2014).

16 Probability

Note, however, that three of these constructs (items 4, 10 and 12) are fake; students were asked about their familiarity with some mathematics concepts that do not exist. This follows a similar structure as the Over-Claiming Questionnaire (OCQ) used by Paulhus et al. (2003)) These three fake items are used to measure young people's propensity to overclaim. Specifically, a partial credit item-response theory (IRT) model – a generalisation of the dichotomous Rasch model applied to ordinal response scales - is estimated using these three measures, with higher values indicating a greater propensity for young people to overclaim. We standardise this measure within each country to mean zero and standard deviation one, meaning our estimates can be interpreted in terms of effect sizes. We also perform an analysis of each of the fake questions separately – i.e. whether there is an association between relative achievement and selecting one of the top two categories (*heard of it often* or *know it well – understand the concept*). The distribution of responses to these questions sample can be found in Table 1. In Appendix E we present alternative estimates adjusting our overclaiming measure for students' responses to the 13 real mathematics concepts in the spirit of the approach used by Paulhus et al. (2003)). Similarly, in Appendix H we provide alternative estimates using summative scores (rather than IRT) to create the overclaiming scale and standardising the scale across countries instead. Neither of these alternative approaches change our substantive results.

3.2. Mathematics achievement

As part of PISA, students sit a two-hour mathematics, reading and science test. In 2012, mathematics was the core domain, meaning most test questions focused on this subject. An item-response theory model was used by the survey organisers to convert student responses to the test questions into five "plausible values" (five separate estimates of students' mathematics achievement). We follow recommended practise, running all our estimates five times (once using each plausible value) and then averaging the results following "Rubin's Rules". See OECD (2009) for further details. The PISA 2012 technical report (OECD, 2014: Tables 12.5 and 12.11) note that the reliability of the mathematics scale is high (around 0.85–0.90).

3.3. Relative mathematics achievement

Our primary covariate is students' *relative* mathematics achievement – i.e., how they performed in the PISA mathematics test *relative* to their school peers. We measure this in two ways.

First, for each student *i* we calculate the difference between their mathematics score (plausible value) and the average mathematics score across all their school peers who took the PISA test[3]:

$$Rel_{ij} = M_{ij} - u_{j-i}$$

Where:

Rel_{ij} The relative achievement of student *i* in comparison to their peers in school *j*.

M_{ij} The mathematics score (plausible value) of student *i* in school *j*.

u_{j-i} The average mathematics score (plausible value) of the students' peers in school *j* (with student *i* excluded from this average).

This relative measure (Rel_{ij}) is then standardised to mean zero and standard deviation one within each country. The process outlined above is followed for each of the five plausible values. In Appendix D we present alternative estimates where we focus on the link between school average achievement and overclaiming, controlling for students' own achievement scores. This is similar to the approach used in previous studies into BFLP effects and students' self-efficacy (Marsh et al., 2008) and produces similar results.

Second, within each school, we divide students into terciles. This captures whether each student is in the top, middle or bottom third of mathematics achievement *within their school*. Again, we do this for each plausible value. This measure will hence capture

3 These averages exclude the students' own score as is standard in the peer effects literature (Sacerdote, 2011). See Appendix D for robustness checks around this approach.

potential non-linearities in the relationship between relative achievement and overclaiming, while also providing a simple way of investigating differences between the lowest, highest, and average achieving students within schools. Table 2 illustrates how these relative terciles (whether students are in the top, middle or bottom third of mathematics achievers within their school) relate to absolute terciles of achievement (whether students are in the top, middle or third of mathematics achievers within their country).

We note two important issues with this measure. The first is that PISA includes a random sample of students within each school rather than all students. As noted by Micklewright et al. (2012) this means there will be classical measurement error in peer comparison variables such as ours, which is likely to lead to a downward bias in one's results. Indeed, according to Micklewright et al. (2012) the estimated effect "*is biased downwards by about one third*" in a sample design such as PISA. The second (related) issue is that there will also be random error in young people's relative achievement due to less than perfect reliability of the mathematics assessment, which is also likely to attenuate the results. Hence, the key implication of these issues is that it will mean our estimates are likely to be conservative, providing a lower-bound on the link between teenagers' relative mathematics achievement and the extent they overclaim.

4. Methodology

Our primary analysis is based on the following ordinary least squares (OLS) regression model:

$$O_{ijk} = \alpha + \beta.Rel_{ijk} + \pi.A_{ijk} + \gamma.B_{ijk} + \delta.S_{jk} + \varphi.Fam_{ijk} + u_k + \varepsilon_{ij} \quad (1)$$

Where:

O_{ijk} The overclaiming scale constructed from the three fake mathematics constructs, standardised to mean zero and standard deviation one.

Rel_{ijk} The mathematics achievement of student *i* relative to their peers in school *j*.

A_{ijk} The PISA reading, mathematics, and science test scores (plausible values) of student *i*.

B_{ijk} A set of basic student *i* (e.g. socio-economic status, gender, year group) and school *j* (e.g. student-teacher ratios, public/private school) controls.

S_{jk} The percent of male students and the average socio-economic status of students in the school.

Fam_{ijk} Student *i*'s self-reported familiarity with actual mathematics constructs (i.e. the real mathematics constructs listed in Section 2).

u_k Country fixed effects. This refers to a dummy variable being included in the model for each country, thus controlling for time invariant country-level differences in the outcome (overclaiming).

ε_{ij} Random error term. The clustering of students within schools is accounted for via the application of the PISA Balanced-Repeated-Replication weights.

Where *i* denotes the student, *j* the school, and *k* the country.

Note that to account for the hierarchical nature of the data – with students nested within schools – we follow recommended practise in applying the Balanced Repeated Replication (BRR) weights that are supplied as part of the public use PISA database (OECD, 2009; OECD, 2014; Jerrim et al., 2017). This is implemented via the 'repest' STATA package developed by analysts from the OECD who lead the PISA study (Avvisati & Keslair, 2019). The application of these weights fully account for the complex PISA survey design. Nevertheless, in Appendix I we present an alternative set of estimates using a different analytic approach, estimating a hierarchical linear model instead. This leads to virtually no change to our substantive results.

The parameter of interest from the model presented in Eq. (1) is β. This captures the effect size change in the overclaiming scale for each standard deviation increase in students' relative mathematics achievement (in comparison to their school peers). We hypothesise that β is positive – that when teenagers are higher achieving than their school peers, they are more likely to overclaim about their mathematics skills.

Four specifications of this model are estimated. The bivariate association between overclaiming (O_{ijk}) and students' relative achievement (Rel_{ijk}) is first estimated without any controls. This provides a baseline against which results from more detailed specifications can be compared. Students' absolute mathematics achievement – i.e. their scores on the PISA test – and a set of other student and school controls are added in specification M1.[4] The β estimates from this specification thus reveal whether students who are high achieving compared to their school peers continue to be more likely to overclaim about their skills, conditional on their *actual* mathematics abilities and a set of other background factors that could potentially confound our results. In model M2 we further control for two other aspects of students' school peers – most notably their school peer group's socioeconomic status and gender composition. This will reveal whether the link between relative achievement and overclaiming could be due to these other aspects of students' school peer group. Finally, specification M3 further controls for students' self-reported familiarity with actual mathematics constructs. This will address whether there remains an independent association between relative achievement and overclaiming, over and above any impact relative achievement may have on how confident young people are in their mathematics knowledge in general.

All models are estimated using the pooled data across all seven countries (while controlling for country fixed effects). This maximises sample size – particularly the number of schools ($n = 2689$) – given that our variable of interest (Rel_{ijk}) is measured at the school-

[4] Students scores on the PISA science and reading test are included as controls in the model from specification M1 onwards as well. See Appendix F for further robustness checks around the inclusion of the plausible values.

Table 3
The association between relative academic achievement compared to school peers and the propensity to overclaim.

	M0		M1		M2		M3	
	Effect size	SE	Effect size	SE	Effect size	SE	Effect size	SE
Relative achievement in school								
One SD increase	0.106*	0.007	0.102*	0.011	0.092*	0.015	0.079*	0.014
N	40,429		40,429		40,429		40,429	
R-squared	0.01		0.06		0.06		0.42	
Controls								
Achievement scores	–		Y		Y		Y	
Gender	–		Y		Y		Y	
Socio-economic status	–		Y		Y		Y	
Immigrant status	–		Y		Y		Y	
grade	–		Y		Y		Y	
Private school	–		Y		Y		Y	
Student:teacher ratio	–		Y		Y		Y	
Country fixed effects	–		Y		Y		Y	
% male school peers	–		–		Y		Y	
School peers SES	–		–		Y		Y	
Self-reported familiarity with real maths concepts	–		–		–		Y	

Notes: Data pooled across Australia, Canada, Ireland, New Zealand, United States, Scotland, and the rest of the UK. Estimates refer to the change in the propensity for individuals to overclaim associated with a one standard deviation increase in mathematics achievement relative to school peers. SE refers to standard errors. Senate and BRR weights have been applied, with each country given equal weight in the analysis.
* indicates statistical significance at the five percent level. Source: Authors' calculations using PISA 2012 data (OECD, 2014).

Table 4
The association between relative academic achievement group (top/middle/bottom tercile within school) and the propensity to overclaim.

	M0		M1		M2		M3	
	Beta	SE	Beta	SE	Beta	SE	Beta	SE
Relative achievement in school (ref: bottom tercile)								
Middle tercile	0.133*	0.022	0.094*	0.023	0.072*	0.023	0.066*	0.019
Top tercile	0.230*	0.019	0.125*	0.025	0.080*	0.029	0.079*	0.029
N	40,429		40,429		40,429		40,428	
R-squared	0.01		0.06		0.06		0.42	
Controls								
Achievement scores	–		Y		Y		Y	
Gender	–		Y		Y		Y	
Socio-economic status	–		Y		Y		Y	
Immigrant status	–		Y		Y		Y	
grade	–		Y		Y		Y	
Private school	–		Y		Y		Y	
Student:teacher ratio	–		Y		Y		Y	
Country fixed effects	–		Y		Y		Y	
% male school peers	–		–		Y		Y	
School peers SES	–		–		Y		Y	
Self-reported familiarity with real maths concepts	–		–		–		Y	

Notes: Data pooled across Australia, Canada, Ireland, New Zealand, United States, Scotland and the rest of the UK. Estimates refer to the difference in the propensity for individuals to overclaim compared to the reference group (being in the lowest achieving third within their school). SE refers to standard errors. Senate and BRR weights have been applied, with each country given equal weight in the analysis.
* indicates statistical significance at the five percent level. Source: Authors' calculations using PISA 2012 data (OECD, 2014).

level. Separate estimates by country are provided in Appendix B. The STATA 'repest' command (Avvisati & Keslair, 2019) is used to ensure all technical aspects of the PISA test and survey design (including the clustering of students within schools) are handled appropriately.

Two alternative versions of the OLS model presented in Eq. (1) are then estimated. First, students' relative achievement in comparison to their school peers (Rel_{ijk}) is measured in a different way. Specifically, it is entered as a set of dummy variables, capturing whether each student is in the top, middle or bottom third of the mathematics achievement distribution within their school. Second, we re-estimate the model for each of the three fake mathematics constructs using logistic regression. In particular, each of the fake items are recoded into binary format, coded zero for the bottom three categories ("*never heard of it*", "*heard of it once or twice*" and "*heard of it a few times*") and coded one for the top two categories ("*heard of it often*" and "*know it well, understand the concept*"). For each item we then estimate our penultimate model specification (M2) to investigate whether students' achievement relative to their school peers is associated with overclaiming across each of the three measures.

Table 5
The association between relative academic achievement and the odds of claiming strong knowledge of the fake mathematics constructs.

Fake item	Odds ratio	Lower CI	Upper CI
Proper number	1.21	1.12	1.30
Subjunctive scaling	1.02	0.91	1.14
Declarative fraction	1.18	1.06	1.32

Notes: Estimates based upon a set of logistic regression models (one for each question). The outcome has been coded as zero if the student said that they had "*never heard of it*", "*heard of it once or twice*" or "*heard of it a few times*" and coded one if they said they had "*heard of it often*" or "*know it well, understand the concept*". Estimates refer to the odds ratio along with the lower and upper confidence interval. These refer to the change in the odds associated with a one standard deviation increase in mathematics achievement relative to their school peers. Estimates based upon model specification M2 (see Table 4 for a full list of controls). Source: Authors' calculations using PISA 2012 data (OECD, 2014).

5. Results

Table 3 presents results from our regression models. Estimates refer to the effect size increase in the overclaiming scale per each standard deviation increase in students' mathematics achievement relative to their school peers. Recall these estimates pool the data across countries with each given equal weight. Appendix B provides a country-by-country breakdown (with sample sizes in Appendix C).

Model M0 provides the bivariate association. This reveals how each standard deviation increase in students' relative mathematics achievement is associated with a 0.11 standard deviation increase in the overclaiming scale. In other words, consistent with our hypothesis, teenagers who are higher achieving than their school peers are more likely to make exaggerated claims about their knowledge. Results from model M1 investigate whether this finding holds once a range of demographic and school background factors are controlled, including students' reading, science, and mathematics PISA scores. There is little evidence that these factors confound our results; the estimated association between relative achievement and overclaiming remains broadly stable at 0.1 standard deviation.

Model M2 adds additional controls for school-level socio-economic status and the percent of male students; two factors that, at the individual level, are key predictors of overclaiming (Jerrim et al., 2023). Again, the inclusion of these controls makes almost no difference to our results – the association remains of the same magnitude (0.1 standard deviation) and statistically significant.

Finally, we investigate whether the link between relative achievement and overclaiming is simply a reflection of young people's confidence in their knowledge about real mathematics constructs. This does not seem to be the case. Even once this additional factor has been controlled the association between relative achievement and overclaiming remains of similar magnitude (0.08 standard deviations) and statistically significant. This is despite the R^2 of the model becoming relatively high, with 43% of the variation in our overclaiming outcome explained.

Table 4 repeats this analysis using an alternative measure of young people's relative academic achievement. Specifically, we now compare students who are (a) around the average for their school and (b) who are amongst the highest achieving third within their school to (c) those who are the weakest third within their school (reference group). The bivariate association in M0 suggests that those in the middle (0.13 standard deviations) and top (0.23 standard deviations) third of the relative achievement distribution are more likely to overclaim than their peers in the bottom third. The magnitude of these differences declines once controls are added in M1 and M2, though remain non-trivial. Interestingly, however, the difference between those in the bottom third and the middle third (0.076 in model M2) is notably bigger than the difference between the middle and top third (0.085 – 0.076 = 0.009 in model M2). This suggests it may be that the least able students within a school are much less likely to exaggerate their knowledge, rather than the most able being particularly likely to exaggerate theirs. Finally, we once again see that controlling for students' self-reported knowledge of real mathematics constructs does not substantively change our results.

To conclude, Table 5 presents estimates from our logistic regression models, where we investigate the link between relative achievement and each fake mathematics construct in turn. Estimates refer to odds ratios, with values greater than one indicating that students who are higher achieving than their school peers are more likely to say that they have either "heard of it often" or "know it well [and] understand the concept". Estimates are based on specification M2.

For two of the three questions ("proper number" and "declarative fraction") there is a moderate association that is statistically significant at the 5% level; each standard deviation increase in relative achievement is associated with around a 1.2 increase in the odds that students' claim to have good knowledge of the fake construct. The exception is "subjunctive scaling" where, although the odds ratio is above one, the magnitude is smaller (1.05) and does not reach statistical significance. Nevertheless, the results reported in Table 5 are largely consistent with those using the "overclaiming" scale (Tables 3 and 4).

6. Conclusions

The Big-Fish-Little-Pond effect is a widely studied phenomena. Almost half a century of work has established that young people who are higher achieving than their school peers have higher levels of academic self-efficacy – over and above their actual academic skill. But might being a high achiever relative to one's school peers lead to a young person becoming overconfident in their abilities, leading them to assume they know more than they really do? Despite growing interest into the issue of overclaiming, no previous research has considered whether it too is influenced by students' relative academic abilities. This paper has contributed international evidence on this matter.

Our results suggest that there is indeed a link between students' mathematics achievement relative to their school peers and the propensity to overclaim. Specifically, a one standard deviation increase in relative mathematics achievement is associated with approximately a 0.1 standard deviation increase in the likelihood of overclaiming, with estimates stable across various model specifications. Further exploration has suggested it is those in the bottom third of the mathematics achievement distribution within their school who are less likely to overclaim than their peers in the middle and top third (between whom there is little difference). Together, these findings provide clear evidence that BFLP effects can be observed for overclaiming.

Our work does have some limitations. First, the PISA data are cross-sectional, focusing on young people's relative achievement and chances of overclaiming at one point in time. Future work may extend our findings using longitudinal data, investigating change in students' relative achievement and change in the tendency to overclaim, leading to stronger evidence of cause and effect. Such work may also reveal when this relationship starts to occur (e.g. is it apparent in both primary and secondary schools)? Second, PISA is a sample survey, meaning not all students within each school take part. As noted by Micklewright et al. (2012) this leads to a "classical" errors-in-variables problem in our measure of interest (relative achievement in school) likely leading to a downward bias in our estimates. Third, our overclaiming measure is based on only three questions. Ideally more would be used to capture the phenomena of overclaiming as precisely as possible. Fourth, our focus has been upon mathematics; we do not know the extent that our findings generalise to other academic areas. Future work may seek to explore whether relative achievement is linked to overclaiming in some subjects (e.g. mathematics) but not others (e.g. art). Fifth, PISA samples only a limited number of schools per country. As our measure of interest is essentially a school-level variable, sample sizes are too small to produce and compare individual country results (at least more detailed model specifications that include other school-level controls). Finally, our findings are based on a pool of primarily English-speaking countries, due to concerns about how the fake mathematics constructs would be translated and interpreted in other languages. Future studies may wish to explore the generalisability of findings to other language groups.

Despite these limitations, we produce robust evidence that there is a BFLP effect for overclaiming. This has implications for how parents, teachers, and other adults interact with young people. While there may be situations in which overclaiming has some benefits (e.g. working in a sales job), in general it is not seen as positive. Recent research has found that individuals who overclaim are more susceptible to falling victim to others' overclaiming and to other types of misinformation (Littrell et al., 2021). In the era of fake news, educating critical thinkers who can discern when someone is overclaiming seems more important than ever. Thus, while self-efficacy should be nurtured, overclaiming should be nipped in the bud. Educators may play an important role here. As argued by Covington and Beery (1976) what young people need from teachers is *realistic* feedback about their abilities – otherwise, they may strive for outcomes that are unsuitable or they are unable to obtain. This includes teachers helping students who are at the top of their class – the relative high achievers – to realise that they still have many things to learn. Indeed, our results suggest that targeting realistic feedback to the highest relative achievers within each school may be important to help young people stay grounded and to prevent their self-belief in their abilities being raised too far.

Appendix A. Sample sizes by country

	School response rate (post replacement)	Student response rate	Full sample	Form A and C only	Baseline analytic sample
Australia	98	87	14,481	9645	9229
Canada	93	81	21,544	14,371	13,840
Ireland	99	84	9714	6478	6253
New Zealand	89	85	5016	3347	3265
Scotland	99	83	4291	2872	2762
UK (England/Wales/NI)	89	86	2945	1956	1900
United States	77	89	4978	3313	3180
Average / Total	92	85	62,969	41,982	40,429

Notes: The difference between the "full sample" and "form A and C only" is that those students who were randomly assigned to complete questionnaire form B have been dropped. The baseline analytic sample refers to where complete data is available on students' relative achievement and the overclaiming scale. See Appendix C for a country-by-country breakdown of sample sizes. Source: Authors' calculations using PISA 2012 data (OECD, 2014).

Appendix B. Country-by-country estimates of the association between a one standard deviation increase in relative achievement and the overclaiming scale

	M0		M1		M2		M3	
	Effect size	SE	Effect size	SE	Effect size	SE	Effect size	SE
Australia	0.11	0.01	0.06	0.02	0.04	0.03	0.08	0.02

(*continued on next page*)

(*continued*)

	M0		M1		M2		M3	
	Effect size	SE	Effect size	SE	Effect size	SE	Effect size	SE
Canada	0.01	0.02	0.11	0.04	0.14	0.04	0.12	0.04
UK (England, Wales, NI)	0.11	0.02	0.09	0.03	0.06	0.05	0.03	0.04
Ireland	0.12	0.02	0.08	0.05	0.13	0.08	0.19	0.07
New Zealand	0.16	0.02	0.10	0.04	0.12	0.06	0.07	0.06
Scotland	0.22	0.02	0.20	0.06	0.20	0.08	0.11	0.06
United States	0.00	0.02	0.14	0.04	0.12	0.07	0.05	0.06
Average	0.10	–	0.11	–	0.12	–	0.09	

Notes: Estimates refer to the change in the propensity for individuals to overclaim associated with a one standard deviation increase in mathematics achievement relative to school peers. SE refers to standard errors. Senate and BRR weights have been applied. See Table 4 for further details about the controls included in each model. Source: Authors' calculations using PISA 2012 data (OECD, 2014).

Appendix C. Sample sizes by country

Country	N
Australia	9229
Canada	13,840
UK (England, Wales, NI)	6253
Ireland	3265
New Zealand	2762
Scotland	1900
United States	3180
Total	40,429

Notes: Source: Authors' calculations using PISA 2012 data (OECD, 2014).

Appendix D. Alternative estimates using school-average (rather than relative) achievement

In the main body of the paper, we define our variable of interest as young people's academic achievement relative to their school peers. We measure this as follows:

$$Rel_{ij} = M_{ij} - u_{j-i}$$

Where:

Rel_{ij} The relative achievement of student *i* in comparison to their peers in school *j*.

Table D1
The association between school average mathematics achievement and the propensity to overclaim.

	M0		M1		M2		M3	
	Effect size	SE	Effect size	SE	Effect size	SE	Effect size	SE
School average achievement								
One SD increase	0.018*	0.006	−0.065*	0.007	−0.061*	0.009	−0.053*	0.009
N	40,438		40,438		40,438		40,437	
R-squared	0.00		0.06		0.06		0.42	
Controls								
Achievement scores	–		Y		Y		Y	
Gender	–		Y		Y		Y	
Socio-economic status	–		Y		Y		Y	
Immigrant status	–		Y		Y		Y	
grade	–		Y		Y		Y	
Private school	–		Y		Y		Y	
Student:teacher ratio	–		Y		Y		Y	
Country fixed effects	–		Y		Y		Y	
% male school peers	–		–		Y		Y	
School peers SES	–		–		Y		Y	
Self-reported familiarity with real maths concepts	–		–		–		Y	

Notes: Data pooled across Australia, Canada, Ireland, New Zealand, United States, Scotland, and the rest of the UK. Estimates refer to the change in the propensity for individuals to overclaim associated with a one standard deviation increase in the average mathematics achievement of the school. SE refers to standard errors. Senate and BRR weights have been applied, with each country given equal weight in the analysis.

* indicates statistical significance at the five percent level. Source: Authors' calculations using PISA 2012 data (OECD, 2014).

M_{ij} The mathematics score (plausible value) of student *i* in school *j*.

u_{j-i} The average mathematics score (plausible value) of the students' peers in school *j* (with student *i* excluded from this average).

There are however alternative approaches. For instance, previous research by educational psychologists into Big Fish Little Pond effects have simply included the scores of young people and the group (e.g. school) average in the same model. The parameter of interest is then how group (school) average achievement is related to the outcome.

In this appendix we present alternative estimates following this approach to test the robustness of our results. Specifically, we estimate models of the form:

$$O_{ijk} = \alpha + \beta.AV_{jk} + \pi.A_{ijk} + \gamma.B_{ijk} + \delta.S_{jk}\varphi.Fam_{ijk} + u_k + \varepsilon_{ij} \tag{2}$$

Where:

AV_{jk} The *average* mathematics achievement of students in the school (e.g. the mean of the first plausible across all students within school *j*). This has then been standardised to mean zero and standard deviation one.

With all other variables as defined as under Eq. (1) in the main text.

The parameter of interest is β, capturing how school average achievement is related to overclaiming, over and above individual student-level achievement and a range of other controls. Our results are presented in Appendix Table D1.

In Model M0 there is a small, positive relationship between school average achievement and our overclaiming scale. The sign changes direction however in model M1 when a set of basic controls are added to the model – including individual-level student achievement scores. This is consistent with previous research into BFLP effects on student self-efficacy, suggesting that school-average achievement has a negative relationship with the propensity of students to overclaim. It is also consistent with the findings we present in the main body of the paper – overclaiming is more common amongst higher-achieving students within a school. The magnitude of the effect is around a 0.07 decrease in our overclaiming scale for each standard deviation increase in school-average achievement. The addition of extra controls in model specifications M2 and M3 leads to little change in the results or magnitude of parameter estimates.

Appendix E. Alternative estimates using an alternative measure of overclaiming

In the main body of our paper, we use what Paulhus et al. (2003)) call the "false-alarm rate" to measure overclaiming – i.e. the propensity of students to report that they know of the "fake" mathematics concepts. Indeed, Paulhus et al. (2003)) note how this is "*an appealing choice*" since it provides "*the most face-valid operationalization of over-claiming*". A possible limitation of this approach to measuring overclaiming, however, is that it is highly correlated with what Paulhus et al. (2003)) call the "hit rate" – the extent that students claim to have knowledge of real mathematics concepts. This in turn could mean that our measure of overclaiming is confounded by student reports of their knowledge of real mathematics constructs.

In the main body of the paper, we tackle this issue via including a control for students' self-reported knowledge of the real mathematics concepts in model M3. The inclusion of this control leads to little change in the parameter estimates or our substantive results.

To provide further evidence on this point, in this appendix we present alternative estimates constructing our measure of over-

Table E1
Alternative estimates of the association between relative academic achievement compared to school peers and the propensity to overclaim.

	M0		M1		M2		M3	
	Effect size	SE	Effect size	SE	Effect size	SE	Effect size	SE
Relative achievement in school								
One SD increase	−0.31*	0.007	0.093*	0.012	0.073*	0.017	0.079*	0.014
N	40,428		40,428		40,428		40,428	
R-squared	0.12		0.22		0.22		0.29	
Controls								
Achievement scores	–		Y		Y		Y	
Gender	–		Y		Y		Y	
Socio-economic status	–		Y		Y		Y	
Immigrant status	–		Y		Y		Y	
Grade	–		Y		Y		Y	
Private school	–		Y		Y		Y	
Student:teacher ratio	–		Y		Y		Y	
Country fixed effects	–		Y		Y		Y	
% male school peers	–		–		Y		Y	
School peers SES	–		–		Y		Y	
Self-reported familiarity with real maths concepts	–		–		–		Y	

Notes: Data pooled across Australia, Canada, Ireland, New Zealand, United States, Scotland, and the rest of the UK. Estimates refer to the change in the propensity for individuals to overclaim associated with a one standard deviation increase in mathematics achievement relative to school peers. SE refers to standard errors. Senate and BRR weights have been applied, with each country given equal weight in the analysis.

* indicates statistical significance at the five percent level. Source: Authors' calculations using PISA 2012 data (OECD, 2014).

Table F1
Alternative estimates of the association between relative academic achievement compared to school peers and the propensity to overclaim. Controlling for all five plausible values in a single model.

	M0		M1		M2		M3	
	Effect size	SE	Effect size	SE	Effect size	SE	Effect size	SE
Relative achievement in school								
One SD increase	0.106*	0.007	0.114*	0.013	0.119*	0.018	0.072*	0.017
Controls								
Achievement scores	–		Y		Y		Y	
Gender	–		Y		Y		Y	
Socio-economic status	–		Y		Y		Y	
Immigrant status	–		Y		Y		Y	
grade	–		Y		Y		Y	
Private school	–		Y		Y		Y	
Student:teacher ratio	–		Y		Y		Y	
Country fixed effects	–		Y		Y		Y	
% male school peers	–		–		Y		Y	
School peers SES	–		–		Y		Y	
Self-reported familiarity with real maths concepts	–		–		–		Y	

Notes: Data pooled across Australia, Canada, Ireland, New Zealand, United States, Scotland, and the rest of the UK. Estimates refer to the change in the propensity for individuals to overclaim associated with a one standard deviation increase in mathematics achievement relative to school peers. SE refers to standard errors. Senate and BRR weights have been applied, with each country given equal weight in the analysis.

* indicates statistical significance at the five percent level. Source: Authors' calculations using PISA 2012 data (OECD, 2014).

claiming a different way. Specifically, we adjust our overclaiming index to account for how students responded to the real mathematics concepts, so that:

$$Overclaiming\_Adjusted_{ij} = Overclaiming\_Raw_{ij} - Accurate\_Reports_{ij}$$

Where:

$Overclaiming\_Adjusted_{ij}$ The "adjusted" overclaiming scale used in this appendix, where students' responses to the fake concepts are adjusted to account for their self-reported knowledge of the real concepts.

$Overclaiming\_Raw_{ij}$ The overclaiming scale based on students' responses to the fake mathematics concepts, used in the main body of the text.

$Accurate\_Reports_{ij}$ A scale based upon students' responses to the real mathematics concepts.

In Appendix Table E1 we replicate our main analytic models using this alternative measure. Overall, the results are consistent with those presented in Table 4 in the main text. Once students' individual achievement has been controlled (in models M1, M2 and M3), we see a positive association between relative achievement and the propensity to overclaim. In particular, a one standard deviation increase in mathematics achievement relative to school peers is associated with around a 0.07–0.09 standard deviation increase in our adjusted overclaiming scale. Our results thus seem to be robust to using this alternative way of constructing the overclaiming scale.

Appendix F. Alternative estimates controlling for all five plausible values for individual student achievement simultaneously

In the main body of the paper, we follow recommended practice in the use of the PISA plausible values, estimating each model five times (once using each plausible values) and then combining the results. Such an approach may not however fully control for the measurement error in students test scores – arguably controlling for all five plausible values (for individual student achievement) simultaneously in a single model might be a better approach. Appendix Table F1 hence presents alternative estimates, where all five plausible values for individual student achievement (in each of mathematics, reading and science) are controlled in the same model. This leads to little change to our parameter estimates of interest or our substantive results.

Appendix G. Why including school fixed effects is inappropriate in this context

In the economics literature of academic rank – which shares many conceptual similarities with the literature on Big Fish Little Pond effects – it is common for school-cohort or class-cohort fixed effects to be included in the analytic model (e.g. Murphy & Weinhardt, 2020). However, such studies typically use administrative data – with very large sample sizes and information available on all individuals within a school – allowing a focus on within-school rank position. We do not believe that including school fixed effects in our models using the PISA data would be a sensible approach.

To understand why, we have estimated the following regression model, including the same controls as listed under M1 in the main body of the paper:

$$Rel_{ijk} = \alpha + \pi.A_{ijk} + \gamma.B_{ijk} + u_{j\ or\ k} + \epsilon_{ij} \quad (3)$$

Table H1
The association between relative academic achievement compared to school. Alternative estimates using summative scores.

	M0		M1		M2		M3	
	Effect size	SE	Effect size	SE	Effect size	SE	Effect size	SE
Relative achievement in school								
One SD increase	0.099*	0.007	0.099*	0.012	0.092*	0.016	0.078*	0.015
N	39,774		39,774		39,774		39,774	
R-squared	0.01		0.08		0.08		0.44	
Controls								
Achievement scores	–		Y		Y		Y	
Gender	–		Y		Y		Y	
Socio-economic status	–		Y		Y		Y	
Immigrant status	–		Y		Y		Y	
grade	–		Y		Y		Y	
Private school	–		Y		Y		Y	
Student:teacher ratio	–		Y		Y		Y	
Country fixed effects	–		Y		Y		Y	
% male school peers	–		–		Y		Y	
School peers SES	–		–		Y		Y	
Self-reported familiarity with real maths concepts	–		–		–		Y	

Notes: Data pooled across Australia, Canada, Ireland, New Zealand, United States, Scotland, and the rest of the UK. Estimates refer to the change in the propensity for individuals to overclaim associated with a one standard deviation increase in mathematics achievement relative to school peers. SE refers to standard errors. Senate and BRR weights have been applied, with each country given equal weight in the analysis.
* indicates statistical significance at the five percent level. Source: Authors' calculations using PISA 2012 data (OECD, 2014). Number of observations is 39,774 rather than 40,429 due to missing information on some of the individual overclaiming items.

Where:

Rel_{ijk}	The mathematics achievement of student *i* relative to their peers in school *j*.
A_{ijk}	The PISA reading, mathematics, and science test scores (plausible values) of student *i*. We follow recommended practice in their use, estimating models separately for each set of plausible values and then averaging the results.
B_{ijk}	A set of basic student *i* (e.g. socio-economic status, gender, year group) and school *j* (e.g. student-teacher ratios, public/private school) level controls.
Fam_{ijk}	Student *i*'s self-reported familiarity with actual mathematics constructs (i.e. the real mathematics constructs listed in Section 2).
$u_{j\ or\ k}$	Fixed-effects at either the country-level (k) or school-level (j)
ϵ_{ij}	Random error term. The clustering of students within schools is accounted for via the application of the PISA Balanced-Repeated-Replication (BRR) weights.
i	Student i
j	School j.
k	Country k.

We estimate two separate specifications of this model – one including fixed effects at the country level (u_k) and one included fixed effects at the school-level (u_j). We are interested in comparing the R^2 from these two models; how much variation in our covariate of interest (students' relative achievement compared to their school peers) is explained by the controls, including the fixed effect terms. The value 1 - R^2 will then reveal how much variation in our key covariate there is left to exploit.

When the model includes the country fixed effects, the R^2 equals 0.76. Although this is high, there is clearly still sufficient variation in our key covariate of interest (Rel_{ijk}) upon which to base our analysis.

Contrast this with a model that also includes school fixed effects. The R^2 increases to 0.998. Almost all the variation in our covariate of interest has been explained by the controls. There is, in other words, almost no variation left in the data for us to exploit.

There are two reasons why we believe this to be a problem.

First, there would clearly be quite a severe collinearity issue between our variable of interest and our controls (including the school fixed effects). This would lead to the well-known problems of very large standard errors and instable coefficient estimates.

Second, as noted by Jennings et al. (2020), "*regression models with high-dimensional fixed effects can exacerbate measurement error bias and increase the likelihood of false positives*". This would likely be a particular problem in our setting, given how we have noted that both students' relative and absolute achievement will be measured with some error – and that the error in these two variables will be correlated with one another. Hence, as Jennings et al. (2020) show, high-dimensional fixed effects (e.g. school fixed effects in our setting) are likely to lead to spurious results.

Appendix H. Alternative estimates using summative scores

In the main body of the paper we construct the overclaiming index using a partial credit item-response theory model, with the scales then standardised within each country. This appendix presents alternative estimates where the raw question scores (ranging

Table I1
The association between relative academic achievement compared to school. Alternative estimates using a hierarchical linear model.

	M0		M1		M2		M3	
	Effect size	SE	Effect size	SE	Effect size	SE	Effect size	SE
Relative achievement in school								
One SD increase	0.105*	0.008	0.100*	0.014	0.089*	0.017	0.081*	0.015
N	40,429		40,429		40,429		40,428	
Controls								
Achievement scores	–		Y		Y		Y	
Gender	–		Y		Y		Y	
Socio-economic status	–		Y		Y		Y	
Immigrant status	–		Y		Y		Y	
grade	–		Y		Y		Y	
Private school	–		Y		Y		Y	
Student:teacher ratio	–		Y		Y		Y	
Country fixed effects	–		Y		Y		Y	
% male school peers	–		–		Y		Y	
School peers SES	–		–		Y		Y	
Self-reported familiarity with real maths concepts	–		–		–		Y	

Notes: Data pooled across Australia, Canada, Ireland, New Zealand, United States, Scotland, and the rest of the UK. Estimates refer to the change in the propensity for individuals to overclaim associated with a one standard deviation increase in mathematics achievement relative to school peers. SE refers to standard errors. Senate weights have been applied, with each country given equal weight in the analysis.
* indicates statistical significance at the five percent level. Source: Authors' calculations using PISA 2012 data (OECD, 2014).

from 1 to 5) are simply added together instead, with the scale then standardised across countries to mean 0 and standard deviation 1. (Table H1)

Appendix I. Alternative estimates using a hierarchical linear model (HLM)

In the main body of the paper we follow recommended practise and account for the complex PISA survey design – including the nesting of pupils within schools – via the application of the supplied Balanced Repeated Replication (BRR) weights. In Appendix Table I1 below, we illustrate how the same substantive results are returned if a hierarchical linear model – sometimes called a multilevel or random effects model – is estimated instead.

References

Ackerman, P. L., & Ellingsen, V. J. (2014). Vocabulary overclaiming — A complete approach: Ability, personality, self-concept correlates, and gender differences. *Intelligence, 46*, 216–227. https://doi.org/10.1016/j.intell.2014.07.003

Avvisati, F., & Keslair, F. (2019). *REPEST: Stata module to run estimations with weighted replicate samples and plausible values*. Boston College Department of Economics. https://ideas.repec.org/c/boc/bocode/s457918.html.

Becker, M., & Neumann, M. (2016). Context-related changes in academic self concept development: On the long-term persistence of big-fish-little-pond effects. *Learning and Instruction, 45*, 31–39. https://doi.org/10.1016/j.learninstruc.2016.06.003

Bless, H., & Burger, A. M (2016). Assimilation and contrast in social priming. *Current Opinion in Psychology, 12*, 26–31. https://doi.org/10.1016/j.copsyc.2016.04.018

Covington, M. V, & Beery, R. G (1976). *Self-worth and school learning (pp. viii, 166)*. Holt, Rinehart & Winston.

Delaney, J., & Devereux, P. (2022). Rank effects in education: What do we know so far? IZA discussion paper, 15128. doi:10.2139/ssrn.4114687.

Denning, J. T., Murphy, R., & Weinhardt, F. (2018). Class rank and long-run outcomes. *The Review of Economics and Statistics*, 1–45. https://doi.org/10.1162/rest_a_01125

Dunlop, P. D., Bourdage, J. S., de Vries, R. E., Hilbig, B. E., Zettler, I., & Ludeke, S. G (2017). Openness to (reporting) experiences that one never had: Overclaiming as an outcome of the knowledge accumulated through a proclivity for cognitive and aesthetic exploration. *Journal of Personality and Social Psychology, 113*(5), 810–834. https://doi.org/10.1037/pspp0000110

Dunlop, P. D., Bourdage, J. S., de Vries, R. E., McNeill, I. M., Jorritsma, K., Orchard, M., et al. (2020). Liar! Liar!(when stakes are higher): Understanding how the overclaiming technique can be used to measure faking in personnel selection. *Journal of Applied Psychology, 105*(8), 784–799. https://doi.org/10.1037/apl0000463

Elsner, B., & Isphording, I. E (2017). A big fish in a small pond: Ability rank and human capital investment. *Journal of Labor Economics, 35*(3), 787–828. https://doi.org/10.1086/690714

Fang, J., Huang, X., Zhang, M., Huang, F., Li, Z., & Yuan, Q. (2018). The big-fish-little-pond effect on academic self-concept: A meta-analysis. *Frontiers in Psychology, 9*, 1569. https://doi.org/10.3389/fpsyg.2018.01569

Fell, C. B., König, C. J., Jung, S., Sorg, D., & Ziegler, M. (2019). Are country level prevalences of rule violations associated with knowledge overclaiming among students? *International Journal of Psychology, 54*, 17–22. https://doi.org/10.1002/ijop.12441

Francis, B., Craig, N., Hodgen, J., Taylor, B., Tereshchenko, A., Connolly, P., et al. (2020). The impact of tracking by attainment on pupil self-confidence over time: Demonstrating the accumulative impact of self-fulfilling prophecy. *British Journal of Sociology of Education, 41*(5), 626–642. https://doi.org/10.1080/01425692.2020.1763162

Goecke, B., Weiss, S., Steger, D., Schroeders, U., & Wilhelm, O. (2020). Testing competing claims about overclaiming. *Intelligence, 81*, Article 101470. https://doi.org/10.1016/j.intell.2020.101470

Grosz, M. P., Lösch, T., & Back, M. D. (2017). The narcissism-overclaiming link revisited. *Journal of Research in Personality, 70*, 134–138. https://doi.org/10.1016/j.jrp.2017.05.006

Iriberri, N., & Rey-Biel, P. (2021). Brave boys and play-it-safe girls: Gender differences in willingness to guess in a large scale natural field experiment. *European Economic Review, 131*, Article 103603. https://doi.org/10.1016/j.euroecorev.2020.103603
Jennings, J. N., Kim, J. M., Lee, J. A., & Taylor, D. (2020). Measurement error and bias in causal models in accounting research. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3731197
Jerrim, J., Lopez-Agudo, L. A., Marcenaro-Gutierrez, O. D., & Shure, N. (2017). What Happens When Econometrics and Psychometrics Collide? An Example Using PISA Data. *Economics of Education Review, 61*, 51–58.
Jerrim, J., Parker, P., & Shure, N. (2023). Overclaiming. An international investigation using PISA data. Assessment in Education. *Principles, Policy & Practise*.
Jin, K. Y., Paulhus, D. L., & Shih, C. L. (2023). A new approach to desirable responding: Multidimensional item response model of overclaiming data. *Applied Psychological Measurement, 47*(3), 221–236. https://doi.org/10.1177/01466216231151704
Jonkmann, K., Becker, M., Marsh, H. W., Lüdtke, O., & Trautwein, U. (2012). Personality traits moderate the big-fish–little-pond effect of academic self-concept. *Learning and Individual Differences, 22*(6), 736–746. https://doi.org/10.1016/j.lindif.2012.07.020
Keller, L., Bieleke, M., Koppe, K. M., & Gollwitzer, P. M (2021). Overclaiming is not related to dark triad personality traits or stated and revealed risk preferences. *PloS one, 16*(8), Article e0255207. https://doi.org/10.1371/journal.pone.0255207
Krumpal, I. (2013). Determinants of social desirability bias in sensitive surveys: A literature review. *Quality & Quantity, 47*, 2025–2047. https://doi.org/10.1007/s11135-011-9640-9
Kuyper, H., Dijkstra, P., Buunk, A. P., & van der Werf, M. P. C. (2011). Social comparisons in the classroom: An investigation of the better than average effect among secondary school children. *Journal of School Psychology, 49*(1), 25–53. https://doi.org/10.1016/j.jsp.2010.10.002
Larrick, R. P., Burson, K. A., & Soll, J. B (2007). Social comparison and confidence: When thinking you're better than average predicts overconfidence (and when it does not). *Organizational Behavior and Human Decision Processes, 102*(1), 76–94. https://doi.org/10.1016/j.obhdp.2006.10.002
Littrell, S., Risko, E. F., & Fugelsang, J. A (2021). 'You can't bullshit a bullshitter' (or can you?): Bullshitting frequency predicts receptivity to various types of misleading information. *British Journal of Social Psychology, 60*(4), 1484–1505. https://doi.org/10.1111/bjso.12447
Magnus, J. R, & Peresetsky, A. A (2022). A statistical explanation of the Dunning-Kruger effect. *Frontiers in Psychology, 13*. https://doi.org/10.3389/fpsyg.2022.840180
Mahmood, K. (2016). Do people overestimate their information literacy skills? A systematic review of empirical evidence on the Dunning-Kruger effect. *Communications in Information Literacy, 10*(2), 199–213. https://doi.org/10.15760/comminfolit.2016.10.2.24
Makarova, E. (2019). Acculturation and school adjustment of minority students: School and family-related factors. In *Intercultural education, 30* pp. 445–447). Taylor & Francis.
Marsh, H. W, & Hau, K. T. (2003). Big-Fish—Little-Pond effect on academic self-concept: A cross-cultural (26-country) test of the negative effects of academically selective schools. *American Psychologist, 58*(5), 364–376. https://doi.org/10.1037/0003-066X.58.5.364
Marsh, H. W, & Parker, J. W (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. https://doi.org/10.1037/0022-3514.47.1.213
Marsh, H. W., Seaton, M., Trautwein, U., Lüdtke, O., Hau, K. T., O'Mara, A. J., et al. (2008). The Big-fish–little-pond-effect stands up to critical scrutiny: Implications for theory, methodology, and future research. *Educational Psychology Review, 20*(3), 319–350. https://doi.org/10.1007/s10648-008-9075-6
Micklewright, J., Schnepf, S. V., & Silva, P. N. (2012). Peer effects and measurement error: The impact of sampling variation in school survey data (evidence from PISA). *Economics of Education Review, 31*(6), 1136–1142. https://doi.org/10.1016/j.econedurev.2012.07.015
Mueller, C. E., & Winsor, D. L. (2016). Math and verbal academic self-concept: Subject-specificity across four distinctive groups of high ability adolescents. *Learning and Individual Differences, 50*, 240–245. https://doi.org/10.1016/j.lindif.2016.08.002
Murphy, R., & Weinhardt, F. (2020). Top of the class: The importance of ordinal rank. *The Review of Economic Studies, 87*(6), 2777–2826. https://doi.org/10.1093/restud/rdaa020
Muszyński, M., Pokropek, A., & Tomasz, Ż. (2021). Structural validity of overclaiming scores: Analysing PISA 2012 data. *Psychological Test and Assessment Modeling, 63*(1), 119–145. https://www.psychologie-aktuell.com/fileadmin/download/ptam/1-2021/Seiten_aus_PTAM_2021-1_ebook_6.pdf.
Nagengast, B., & Marsh, H. W (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), 629–656. https://doi.org/10.1080/01443410.2011.586416
OECD. (2009). *PISA data analysis manual: Spss and sas* (2nd ed.). OECD Publishing.
OECD. (2014). PISA 2012 Technical Report. https://www.oecd.org/pisa/pisaproducts/PISA-2012-technical-report-final.pdf.
Paulhus, D. L., & Dubois, P. J. (2014). Application of the overclaiming technique to scholastic assessment. *Educational and Psychological Measurement, 74*(6), 975–990. https://doi.org/10.1177/0013164414536184
Paulhus, D. L., Harms, P. D., Bruce, M. N., & Lysy, D. C. (2003). The over-claiming technique: Measuring self-enhancement independent of ability. *Journal of Personality and Social Psychology, 84*(4), 890–904. https://doi.org/10.1037/0022-3514.84.4.890
Phalet, K., & Baysu, G. (2020). Fitting in: How the intergroup context shapes minority acculturation and achievement. *European Review of Social Psychology, 31*(1), 1–39. https://doi.org/10.1080/10463283.2020.1711627
Sacerdote, B. (2011). Peer effects in education: How might they work, how big are they and how much do we know thus far?. In *Handbook of the economics of education, 3* pp. 249–277) Elsevier. https://doi.org/10.1016/B978-0-444-53429-3.00004-1
Salvi, C., Barr, N., Dunsmoor, J. E., & Grafman, J. (2022). Insight problem solving ability predicts reduced susceptibility to fake news, bullshit, and overclaiming. *Thinking & Reasoning*. https://doi.org/10.1080/13546783.2022.2146191
Wouters, S., Colpin, H., Van Damme, J., De Laet, S., & Verschueren, K. (2013). Early adolescents' academic self-concept formation: Do classmates or friends matter most? *Learning and Individual Differences, 27*, 193–200. https://doi.org/10.1016/j.lindif.2013.09.002
Yamamoto, Y., & Holloway, S. D. (2010). Parental expectations and children's academic performance in sociocultural context. *Educational Psychology Review, 22*, 189–214. https://doi.org/10.1007/s10648-010-9121-z
Yang, Z., Barnard-Brak, L., & Lan, W. Y. (2019). Examining the association of over-claiming with mathematics achievement. *Learning and Individual Differences, 70*, 30–38. https://doi.org/10.1016/j.lindif.2019.01.004
Yates, J. F., Lee, J. W., & Bush, J. G. G. (1997). General knowledge overconfidence: Cross-national variations, response style, and "reality". *Organizational Behavior and Human Decision Processes, 70*(2), 87–94. https://doi.org/10.1006/obhd.1997.2696

John Jerrim is a professor of Education and Social Statistics at UCL.

Philip D Parker is a professor of Psychology at Australian Catholic University.

Nikki Shure is an associate professor of Economics at UCL.