

The Efficacy of Inquiry-Based Instruction in Science: a Comparative Analysis of Six Countries Using PISA 2015

Mary Oliver¹  · Andrew McConney² · Amanda Woods-McConney²

Published online: 26 December 2019

© The Author(s) 2019

Abstract

This study is a comparative analysis of 15-year-old students' scientific literacy, and its association with the instructional strategies that students experience, across six OECD countries that participated in PISA 2015. Across the six countries, the study investigates the efficacy of inquiry-based instruction in science in contrast with two other instructional approaches to teaching secondary science: adaptive and teacher-directed teaching. The analysis shows that students who reported experiencing high frequencies of inquiry strategies in their classrooms consistently evidenced lower levels of scientific literacy across the six countries. Benchmark analysis also showed, common to all six countries, a strongly positive association between the frequency of teacher-directed and adaptive teaching strategies and students' scientific literacy. Additionally, the study disaggregates PISA's composite variable representing inquiry-based instruction and shows that different components of inquiry are differentially associated with students' scientific literacy. We discuss the implications of these analyses for science teacher educators, science teachers, and educational policy makers. In doing so, we add nuance to our understanding of the efficacy of inquiry-based instruction in science, suggesting that some components, as conceptualised and assessed in PISA, seem to suggest greater attention and use, and others more moderated use.

Keywords Inquiry-based instruction · Scientific literacy · Comparative analysis · PISA · Secondary analysis

✉ Mary Oliver
mary.oliver@nottingham.ac.uk

¹ School of Education, University of Nottingham, Nottingham, UK

² School of Education, Murdoch University, Perth, Western Australia

Introduction

Scientific literacy has been maintained as a keystone goal of science education policy and practice. In part, this may be because of widespread agreement that a scientifically literate citizenry is well advantaged in making decisions about its health and wellbeing at personal, social, and ecological levels (McConney et al. 2011, OECD 2018). One contemporary interpretation of scientific literacy is that used in the Organisation for Economic Co-operation and Development's (OECD) Programme for International Student Assessment (PISA). In PISA (OECD 2016b), scientific literacy is defined as:

...the ability to engage with science-related issues, and with the ideas of science, as a reflective citizen. A scientifically literate person is willing to engage in reasoned discourse about science and technology, which requires the competencies to explain phenomena scientifically, evaluate and design scientific enquiry, and interpret data and evidence scientifically (p. 28)

Of course, also reflected in PISA's conceptualisation, is the second, equally important purpose of science education: the enhancement of students' interest, motivations, and engagement in science (Woods-McConney et al. 2013b; Thomson et al. 2013).

Perennially, for science educators and educational policy makers, an important question is how best, from a pedagogical perspective, to achieve these keystone purposes with students in schools. Particularly, it would seem critically important for teachers to be aware of, and to use effectively, instructional approaches to teaching science anchored to strong research evidence, rather than ideology. Equally, science teacher educators are responsible for teaching prospective teachers evidence-based pedagogy for teaching and learning science. A variety of approaches, and competing evidentiary claims, however, can make it challenging for science teacher educators and prospective and newly practicing teachers to identify, assess, choose, and effectively use evidence-based instructional approaches.

This study is a comparative examination of the association between three distinct instructional approaches to teaching science and students' scientific literacy, as conceptualised and assessed in PISA (OECD 2018). Specifically, it is about the efficacy of inquiry-based instruction in science, in comparison with two other distinguishable approaches to teaching secondary science: adaptive instruction and teacher-directed instruction. The study builds on previous analyses of relationships between inquiry-based teaching and learning and students' scientific literacy and engagement in science (Cairns and Areepattamannil 2017; McConney et al. 2014). In our previous research, we examined PISA 2006 data for Australia, Canada, and New Zealand. What emerged from these analyses raised important questions about longstanding doctrine in science education research and practice. Consistently across the three countries, students who reported experiencing high levels of inquiry-based teaching and learning in their science classrooms also evidenced scientific literacy lower on average than their respective country means. The corollary was also true; students who reported experiencing low levels of inquiry-based teaching in their science classrooms consistently evidenced scientific literacy on average higher than their respective country means (McConney et al. 2014). We also noted, consistent with much previous literature, that higher than average levels of inquiry-based teaching and learning in students' science classrooms were consistently and positively associated with higher than average student interest and engagement in science (McConney et al. 2014).

As recently summarised by Sjøberg

The PISA student questionnaire includes a series of questions....to students about the teaching methods and classroom practices that characterize their school experiences. When looking for possible relationships between these variables and the PISA scores, many of the results are surprising and should be given attention. The most intriguing aspect of the results is that they run contrary to current advice from science educators as well as “accepted wisdom” among policymakers and curriculum specialists on what constitutes good instruction (2016, p. 123).

In the current study, we used the latest publically available PISA data that again has science as its focus, to expand our previous analysis of the association between students’ scientific literacy and instructional approach to six Anglophone Western democracies, namely Australia, Canada, Ireland, New Zealand, the UK, and the USA. We chose these six because their schools arguably have had substantial exposure to inquiry-based teaching and learning as preferred pedagogy in school science and because they share broadly similar systems of comprehensive secondary schooling, similar socio-cultural roots, and similar economic and government systems.

In the current study, we ask two research questions:

1. To what extent is the variability in high school students’ scientific literacy associated with inquiry-based instruction in science as compared with two other distinct approaches to teaching and learning science that students experience in science classrooms? Does this vary by country?
2. In relationships between students’ scientific literacy and inquiry-based instruction, are differences evident when inquiry is represented as an index (i.e., a composite variable made up of several items) versus when it is represented by individual items? Does this vary by country?

To answer the first question, we used secondary analysis of publically available data from PISA 2015 for the six countries and described the patterning observed between students’ scientific literacy and the extent to which they reported having experienced inquiry-based, teacher-directed, and adaptive pedagogies. In answering the second question, and given that PISA 2015 included the composite variable IBTEACH as a measure of the extent to which students experience inquiry-based teaching in their science classrooms, we examined the extent to which individual components of inquiry are positively (or negatively) associated with students’ scientific literacy. To accomplish this, we disaggregated IBTEACH into its item-level components and examined associations between students’ scientific literacy and the nine individual items that comprise the composite variable.

Literature Review

In science classrooms, particularly those across the western world, it is apparent that inquiry-based teaching and learning is well embedded. It is also apparent that the science education research literature in support of inquiry-based teaching and learning is substantial and enduring (e.g., Lee & Songer 2003; Shymansky et al. 1990). Yip (2001) described inquiry-based

pedagogy as “a teaching strategy that fosters creativity, autonomy, intellectual scepticism, active participation and interaction of students” (p. 114). UK Government reports attributed positive outcomes to inquiry, noting that in schools that showed improvement in science, an important factor facilitating students’ progress was more “practical science lessons and the development of the skills of scientific enquiry” (Ofsted 2011, p. 6). Analysing TIMSS 2015 data for Norway, researchers have also shown non-linear relationships between inquiry-based instruction and science achievement (Teig et al. 2018) and suggested that inquiry instruction “seems to be beneficial for students’ achievement” (p. 28). In the UK, the Learned Societies advocate investment in practicals and problem solving in science in schools (Royal Society 2014) to “reflect the nature of scientific enquiry” (p. 49) as an integral part of learning about, through, and of science. This is consistent with the view that laboratories are an integral part of teaching and learning science and “often seen as a prerequisite for quality science teaching” (Sjøberg 2018, p. 198).

Thus, science educators across the world routinely promote the benefits of inquiry-based instruction as “best practice” or the “method of choice” for fostering students’ interest and understanding in science (Rennie 2010). Australia, for example, explicitly emphasises *Science Inquiry Skills* as one of three interrelated strands of science in the structure of the curriculum. Along with *Science Understanding* (including Biology, Chemistry, Physics, and Earth and Space Sciences) and *Science as a Human Endeavour*, *Science Inquiry Skills* “involves identifying and posing questions; planning, conducting and reflecting on investigations; processing, analysing and interpreting evidence; and communicating findings” (Australian Curriculum, Assessment and Reporting Authority (ACARA) 2019).

Inquiry-based teaching in science has similarly become part of the lexicon across Western Europe, receiving policy support, funding, and promotion from the European Union (EU) following the publication of the Rocard (2007) report. Rocard recommended that “introduction of inquiry-based approaches in schools ... should be actively promoted” (p. 17) and the report gave rise to a concerted EU commitment to inquiry-based learning projects such as INQUIRE, MASCIL, PRIMAS, and SAILS. The various projects’ webpages reinforce the position that there is general agreement within the science education community about the effectiveness of inquiry-based pedagogical practices, with the shared understanding that inquiry-based pedagogy is typified when students are asked to pose questions, plan, and investigate. The value and necessity of inquiry-based science education (IBSE) in the current education context are captured by Harlen (2013), who emphasised the need for inquiry-based science education:

...the value of IBSE is not a matter that can be decided by empirical evidence, but it is a value judgement that the competences, understanding, interest and attitudes that are its aims are worthwhile and indeed are necessary in a modern education (2013, p. 4).

In the UK, policy makers and the Office for Standards in Education, Children’s Services and Skills (Ofsted) have encouraged the use of *practical work* in school science, associating lack of challenge with “poor opportunities for pupils to plan, carry out and evaluate investigations independently” (Ofsted, 2013, p. 10). *Maintaining Curiosity*, the 2013 Ofsted report, explains that teachers perhaps lack confidence or “understanding of the purpose of scientific enquiry and of the value of constructing activities that lead pupils to discover the scientific ideas’ themselves (p. 10), making a case for more time for enquiry” (p. 44). Additionally, and importantly, the US National Research Council (2012) conceptualised a three-dimensional model of science education that reflects the work of scientists and engineers (p. 45) with an

implicit focus on inquiry to “help students make sense of phenomena” (Roseman et al. 2017, p. 118).

Inquiry-based teaching and learning, however, continue to span a wide range of meanings and strategies, from collaborative small group work to discovery learning, hands-on or practical work, and the nature of science. In science education, inquiry has often been seen as including “student-centered interactions, student investigations and hands-on activities, and focus on models or applications in science” (Areepattamannil 2012, p. 135). More broadly, and consistent with previous work, our view of inquiry-based teaching and learning in science reflects pedagogies in which “students may be responsible for naming the scientific question under investigation, designing investigations to research their questions and interpreting findings from investigations” a description initially provided by Nadelson, Williams, and Turner in their Campbell Collaboration systematic review (2011, p. 1).

Additionally, models of inquiry-based instruction in science have included aspects of both “the *doing* (practices) of inquiry and learning about the *nature of scientific inquiry*” (Crawford 2014, p. 517). For example, Minner et al. (2010) identified particular aspects of inquiry-based learning that supported students’ conceptual understanding, suggesting that this goal was more likely to be achieved through the use of “teaching strategies that actively engage students in the learning process through scientific investigations” with emphasis on “active thinking and drawing conclusions from the data” (p. 474). Authors like Cairns and Areepattamannil (2017) draw a further distinction between inquiry-based teaching and inquiry-based learning, with the latter, as a “learning process engaged in by students that ... follows the process of scientific inquiry” (p. 4). Moreover, an emphasis on “science as argument and explanation” [rather than] ‘science as exploration and experiment’ reflects a shift on the part of researchers to identify those contexts in which inquiry can be effective (Kawalkar & Vijapurkar 2011, p. 2005).

Inquiry in science classrooms has also been represented as *doing practical work* (Ofsted, 2013; Rennie 2010) which is seen as a “vital element of learning science, helping pupils to develop enquiry skills and gain scientific knowledge” (Wellcome 2017, p. 2). Osborne has presented a rationale for practical work for both the “demonstration of a phenomenon” and what the “experience of that means to engage in the whole experience of empirical inquiry” (Osborne 2015, p. 21). For example, materials produced in *Toward High School Biology* (THSB) use inquiry-based approaches to learning aligned with “three science practices (recording observation, making scientific predictions, and making evidence-based claims)” (Roseman et al. 2017, p. 113). The argument for a more explicit approach to teaching the content and nature of science, rather than learning science through “doing science” (Hodson 2014, p. 2535), has thus been balanced against the necessity of “doing science in a critical and supportive learning environment” (p. 2552).

Despite its ubiquity in practice and policy across the western world, however, the efficacy of inquiry-based approaches for fostering scientific literacy has been increasingly scrutinised (McConney et al. 2014 ; Cairns and Areepattamannil 2017; Hattie 2009; Kirschner et al. 2006; Klahr 2013). Many of these studies have become possible in part because of the advent of large-scale international assessments like PISA and Trends in Mathematics and Science Study (TIMSS). For example, using PISA 2006, our previous research (McConney et al. 2014) showed that students reporting high levels of inquiry-based instruction in their science classrooms performed less well in science, on average, in comparison with their peers who reported lower levels of inquiry. Similarly, using PISA, Jiang and McComas (2015) found that students’ highest science achievement is evident when they are involved in conducting

activities and drawing conclusions from data, rather than in what is considered to be “higher level” inquiry activities such as designing the investigation or raising their own questions. In seeking to understand the association between inquiry-based teaching and Qatari students’ achievement and interest, using PISA data, Areepattamannil (2012) reported that students who experienced more frequent inquiry had higher than average levels of interest but lower achievement scores, as we had also shown (McConney et al. 2014). Interestingly, in a large study of the top ten performing countries in PISA, *investigation* was found to be “negatively associated with performance” (Lau & Lam 2017, p. 2142).

In making the case for greater levels of teacher-guided instruction, Kirschner et al. (2006) concluded “minimally guided instruction is less effective and less efficient than instructional approaches that place a strong emphasis on guidance of the student learning process” (p. 75). This view was supported in a meta-analysis of inquiry-based learning reporting “larger effect sizes were associated with more specific types of guidance” (Lazonder and Harmsen 2016, p. 704). Additionally, using a framework to distinguish between the “cognitive features of the activity and degree of guidance given to students”, Furtak et al. (2012, p. 300) concluded that epistemic inquiry had the strongest positive effect compared with other forms of inquiry, namely, procedural and social. Additionally, “studies involving teacher-led activities had mean effect sizes about 0.40 larger than those with student-led conditions” (p. 300).

As explained above, PISA’s view of scientific literacy is tightly aligned with understandings and expectations of three core competencies for students: “Explain phenomena scientifically”; “Evaluate and design scientific enquiry”; and “Interpret data and evidence scientifically” (OECD 2018, p. 72). From this, it becomes clear that inquiry is an integral part of PISA’s conceptualisation and operationalisation of scientific literacy (OECD 2018; Sjøberg 2018). What this short review illustrates, however, is that despite the ubiquity of inquiry in both practice and educational policy, several important questions remain about what essential strategies comprise inquiry and importantly their empirical relationships to the goal of scientific literacy for all.

Method

Building on our previous work (McConney et al. 2014), the current study examines associations between students’ scientific literacy scores, and students’ reports on the frequency of various teaching strategies used in their classrooms, in six Anglophone countries. We used secondary analysis of the Organisation for Economic Co-operation and Development’s (OECD) PISA 2015 data for Australia, Canada, Ireland, New Zealand, the UK, and the USA. We purposively chose these six because they arguably have been among the countries most exposed, and perhaps most receptive, to the well-established view of science education researchers about the efficacy of inquiry-based teaching and learning in science. Additionally, the six countries share broadly similar systems of comprehensive secondary schooling and similar socio-cultural histories of English colonisation and post-colonial development. Furthermore, the six share similar levels of economic development (all are considered highly developed) and all are among the top 20 countries on the United Nations Development Programme (UNDP) 2016 Human Development Report’s (released in March 2017) list of countries with very high human development.

Datasets PISA is an international standardised assessment of the literacy of 15-year-old students in reading, mathematics, and science conducted on a 3-year cycle that began in

2000. Each round of PISA assesses all three subjects and focuses in depth on one of the three; in 2015, PISA's focus, for the second time, was science. The OECD's overarching intent for PISA is to support further development of countries' educational systems toward facilitating knowledge and skills necessary for participation in highly developed economies (OECD 2004, 2007). Different from other international assessments, rather than assessing students using a particular national or curriculum-based measure, PISA surveys have been intentionally decoupled from specific school or country curricula; the assessments are purposely based on holistic descriptions of discipline-specific literacies that refer to "students' capacity to apply knowledge and skills in key subjects, and to analyse, reason and communicate effectively as they identify, interpret and solve problems in a variety of situations" (OECD 2016a, b, p. 25).

It is also important to note that PISA uses a two-stage sampling process in which schools are sampled first and then students sampled within participating schools. This means that sampling weights are associated with each student because students and schools in any particular country may not have the same probability of selection, and some groups are over-sampled to allow national reporting priorities to be met (OECD 2009). This approach to sampling has the potential to increase the standard errors of population estimates. In this study therefore, and consistent with PISA's recommendation, both descriptive and inferential statistics have been produced using a Balanced Repeated Replication (BRR) procedure (Fay variant) with 80 replication estimates to generate unbiased standard errors that take account of clustering in the samples (OECD 2009). All statistics are produced using the International Database (IDB) Analyser, an application developed by the International Association for the Evaluation of Educational Achievement (IEA) that can be used to analyse most major large-scale assessment surveys, including those conducted by the OECD. Additionally, when possible, we retrieved descriptive statistics directly from the publically available primary analysis of PISA 2015 conducted by the (OECD 2016b, p. 25).

Variables In addition to being assessed on scientific literacy as defined by PISA's conceptual framework (OECD 2016a), participating students also respond to a short questionnaire about "themselves, their homes, and their schools and learning experiences" (OECD 2018, p. 3). Several items from the background questionnaire (e.g., parents' education, parents' occupations, home possessions, number of books, and other educational resources available in the home) are combined to form a student-level index representing socioeconomic status. In PISA, this variable is named the index of economic, social, and cultural status (ESCS) and is standardised to a mean of zero and a standard deviation of one (OECD 2016a).

As in previous rounds, PISA 2015 also included surveys of teaching and learning strategies experienced by 15-year-olds in their science classrooms. Specifically, PISA asks students how often several learning/teaching activities happened in their science classrooms. These items were used to create several composite variables, including indices of inquiry-based instruction (IBTEACH), adaptive instruction (ADINST), and teacher-directed instruction (TDTEACH). The individual items comprising these indices ask students to indicate using a four-point scale ("in all lessons"; "in most lessons"; "in some lessons"; "never or hardly ever"), the frequency with which they experience various learning and teaching activities. For all indices, higher values indicate that the activities happened more frequently in science lessons (OECD 2016b). In answering the research questions posed for this study, we used the three indices representing distinct teaching/learning approaches (IBTEACH, ADINST, TDTEACH), categorical variables such as country, and PISA's composite measure of student socioeconomic status (ESCS), as a covariate to statistically "level the playing field" across countries.

The composite variable “inquiry-based instruction” includes questions about experimentation and hands-on activities as well as developing conceptual understanding of scientific ideas. PISA constructed its index of inquiry-based instruction (IBTEACH) from students’ responses to nine survey items about the frequency with which they experienced specific activities. These included the following: (1) students are given opportunities to explain their ideas; (2) students spend time in the laboratory doing practical experiments; (3) students are required to argue about science questions; (4) students are asked to draw conclusions from an experiment they have conducted; (5) the teacher explains how a science idea can be applied to different phenomena; (6) students are allowed to design their own experiments; (7) there is a class debate about investigations; (8) the teacher clearly explains the relevance of science concepts; and (9) students are asked to do an investigation to test ideas (OECD 2016a, b).

For the composite variable “adaptive instruction”, students were asked how frequently their teacher adapts the lessons based on students’ needs. PISA constructed its index of adaptive instruction (ADINST) from students’ reports on three survey items about teacher activities in science classrooms. These included the following: (1) the teacher adapts the lesson to my class’s needs and knowledge; (2) the teacher provides individual help when a student has difficulties understanding a topic or task; and (3) the teacher changes the structure of the lesson on a topic that most students find difficult to understand. Taking these items together, this index of adaptive instruction could also be characterised as “differentiated instruction” (OECD 2016a, b).

Lastly, for “teacher-directed instruction”, students were asked about the frequency of activities such as “*the teacher explains scientific ideas*” to determine the amount of teacher direction in the lessons. PISA constructed its index of teacher-directed instruction (TDTEACH) from students’ reports about how often four activities happened in science classes: (1) the teacher explains scientific ideas ; (2) a whole class discussion takes place with the teacher; (3) the teacher discusses our questions; and (4) the teacher demonstrates an idea (OECD 2016a, b).

Analysis In addition to descriptive benchmark analyses, we used multivariate regression analysis via the IDB Analyzer, and accounting for students’ socioeconomic status (SES), we examined the direction and relative size of the effect on scientific literacy for each teaching approach, while controlling for the other two approaches. Further, as indicated by research question 2, we also examined the nine component aspects of inquiry to ask whether the association seen between inquiry as a composite index variable and students’ scientific literacy might appear differently if inquiry were represented by its individual component parts. In other words, we used the data to determine whether individual aspects of inquiry (for example, designing investigations, doing practical experiments, drawing conclusions) are differentially associated with scientific literacy, as measured by PISA.

Findings

Our purpose in this study is to examine empirically the association between students’ scientific literacy as measured in PISA and inquiry-based instruction in science, in comparison with two other distinguishable instructional approaches to teaching secondary science, thereby extending our previous research (McConney et al. 2014). Additionally, we further asked whether differences are evident in associations between students’ scientific literacy and inquiry-based

instruction when inquiry is represented as a *composite index* versus when it is represented by *individual questionnaire items*.

The scientific literacy of students representing the six countries included in this study, ordered by their relative performance in PISA 2015 (Canada being the highest performer in this group of countries), is shown in Table 1. Table 1 also provides the number of students who participated in each country and various measures of the variability around each country’s scientific literacy mean. For example, among 32 OECD countries, Canada’s rank in scientific literacy could potentially range between third and fourth. Showing considerably more variability, students’ average scientific literacy for the UK would place it between 6th and 13th among OECD countries. Secondary school students in the USA evidenced both the lowest scientific literacy mean and the largest variability in scientific literacy among the six countries examined.

As described above, in this study, we used three PISA indices representing contrasting teaching and learning approaches (IBTEACH, ADINST, TDTEACH). All three of these composite variables are scaled to a mean of zero and a standard deviation of one. The three indices represent students’ aggregated reports of the frequency with which they experience classroom activities that comprise each pedagogical approach. Table 2 provides the means and standard errors for each composite variable, by country.

Although having notably different scientific literacy means, students from the USA and Canada report similar frequencies of inquiry-based teaching and learning (IBTEACH) in science, *considerably above* the international average. Australian and New Zealander students, in contrast, report inquiry-based activities *moderately* above the international average, and students in Ireland and the UK report experiencing inquiry-based teaching activities in their secondary science classrooms essentially *equal to the scaled international mean*.

A similar pattern is apparent for teacher-directed teaching and learning in science (TDTEACH). Students in Canada and the USA report the highest frequencies, on average, of teacher-directed activities in science, both considerably above the international average (0.37 for Canada and 0.32 for the USA). Students in Australia and New Zealand report slightly more modest frequencies of teacher-directed activities in science (0.29 for New Zealand and 0.27 for Australia), and students in the UK and Ireland report teacher-directed activities in their science classrooms at the international mean.

The patterning of country groupings changes for students’ experiences of activities consistent with adaptive instruction (ADINST). Secondary students in Canada, New Zealand, the USA, and Australia report mean levels of adaptive instruction well above

Table 1 Scientific literacy in PISA 2015 for six Anglophone countries (Australia, Canada, Ireland, New Zealand, UK, and USA)

Country	No. of cases	Mean	SE	95% confidence interval	Range of ranks among OECD countries	
					Upper rank	Lower rank
Canada	20,058	528	2.08	524–532	3	4
New Zealand	4520	513	2.38	509–518	5	9
Australia	14,530	510	1.54	507–513	6	11
UK	14,157	509	2.56	504–514	6	13
Ireland	5741	503	2.39	498–507	11	18
USA	5712	496	3.18	490–502	15	25

Table adapted from OECD, 2016, PISA 2015 Results (Vol. I)

Table 2 Student-reported science teaching activity means and standard deviations for six Anglophone countries (Australia, Canada, Ireland, New Zealand, UK, USA)*

	IBTEACH (inquiry-based)			TDTEACH (teacher directed)			ADINST (adaptive instruction)		
	Mean	SE	SD	Mean	SE	SD	Mean	SE	SD
CAN	0.27	0.01	0.97	0.37	0.01	1.06	0.26	0.02	1.01
NZL	0.16	0.02	0.86	0.29	0.02	0.98	0.25	0.02	0.93
AUS	0.18	0.01	0.84	0.27	0.01	0.99	0.20	0.01	0.95
GBR	-0.01	0.01	0.84	0.09	0.01	0.94	0.15	0.02	0.97
IRL	0.01	0.02	0.80	-0.02	0.02	0.93	-0.02	0.02	0.95
USA	0.34	0.03	1.04	0.32	0.02	1.07	0.24	0.02	1.01

Countries ordered according to science literacy mean performance

the international mean (in all cases, about one quarter of a standard deviation above). Students in the UK report a slightly more modest occurrence of adaptive instruction, but still well above the international mean. Irish students report adaptive instruction at a level consistent with the international mean.

To answer research question 1, in addition to the descriptive statistics associated with scientific literacy performance and three contrasting teaching approaches, we conducted benchmark analysis of inquiry-based, adaptive, and teacher-directed instructions for scientific literacy performance groupings across the six countries. These analyses were produced with the IDB Analyser using a balanced repeated replication (BRR) procedure (Fay variant) with 80 replications (OECD 2009). Benchmarks reflect PISA's differentiated levels of scientific literacy (OECD 2016a); for 2015, there were 8 levels, but for our purposes, the two at either end of the scientific literacy distribution were collapsed into one to achieve more robust numbers of students represented at every level. The figures depicting the benchmark analyses therefore use 6 benchmarks (levels) of scientific literacy.

As shown in Fig. 1, for the six Anglophone countries in this study, students at the lower levels of scientific literacy are consistently those who tend to report the highest frequencies of inquiry-based activities in their respective countries. The corollary is also generally true. In all countries (except Australia), student groups performing at the highest levels of scientific literacy are those who also report low levels of inquiry in their science classrooms. Generally, for these six countries, the patterning evident from benchmark analysis indicates a *negative association* between the frequency of inquiry-based activities (taken as a whole) and students' scientific literacy.

By contrast, as depicted in Fig. 2, for the six countries in this study, students at the lower levels of scientific literacy are consistently those who report the lowest frequencies of teacher-directed activities in science classrooms. Furthermore, for all six countries, student groups evidencing the highest levels of scientific literacy are those who also report high levels of teacher-directed teaching and learning in their science classrooms. Consistently, for these six countries, the patterning evident from benchmark analysis suggests a *strong positive association* between the frequency of teacher-directed activities and students' scientific literacy.

Similarly, as depicted in Fig. 3, students who performed at the lower levels of scientific literacy are consistently those who report the lowest frequencies of adaptive instruction in science classrooms (except for New Zealand). Additionally, for all six countries, student groups evidencing the highest levels of scientific literacy are those who also report the highest

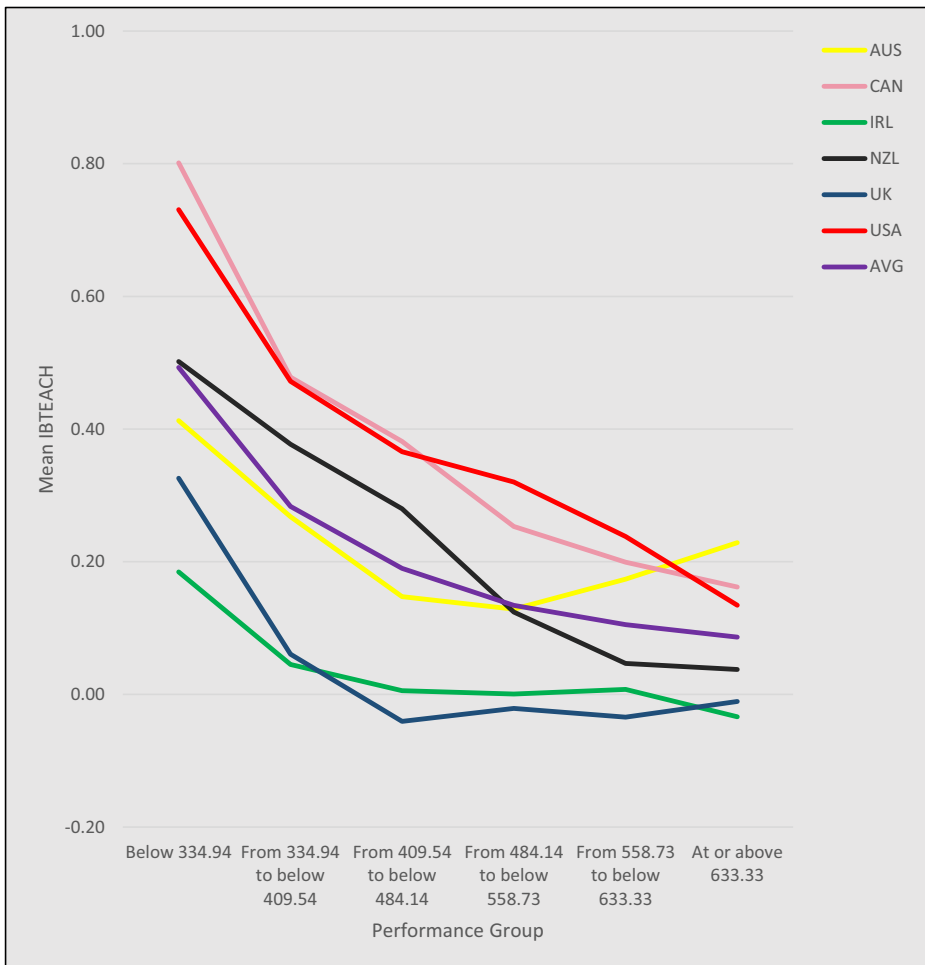


Fig. 1 Mean levels of inquiry-based instructional activities at six science literacy performance benchmarks for students in six countries in PISA 2015

levels of adaptive teacher instruction in their science classrooms. Consistently, for these six countries, the patterning evident from benchmark analysis also suggests a relatively *strong positive association* between the frequency of adaptive instruction and students’ scientific literacy.

To complete our answer to research question 1, we conducted multivariate regression analysis, again using IEA’s IDB Analyser. In the regression analysis, student scientific literacy, as measured by the cognitive component of PISA, served as the dependent (criterion) variable. We included student-level socioeconomic status (SES, ESCS in PISA) as a covariate to level the statistical playing field by controlling for SES since we know that this factor typically accounts for a substantial portion of the variability in scientific literacy (e.g., Woods-McConney et al. 2013a, b). Additionally, we used the three composite indices provided by PISA as independent (predictor) variables, in a simultaneous solution, to observe the effect of each in the context of the other two, as we can be confident that science teachers would

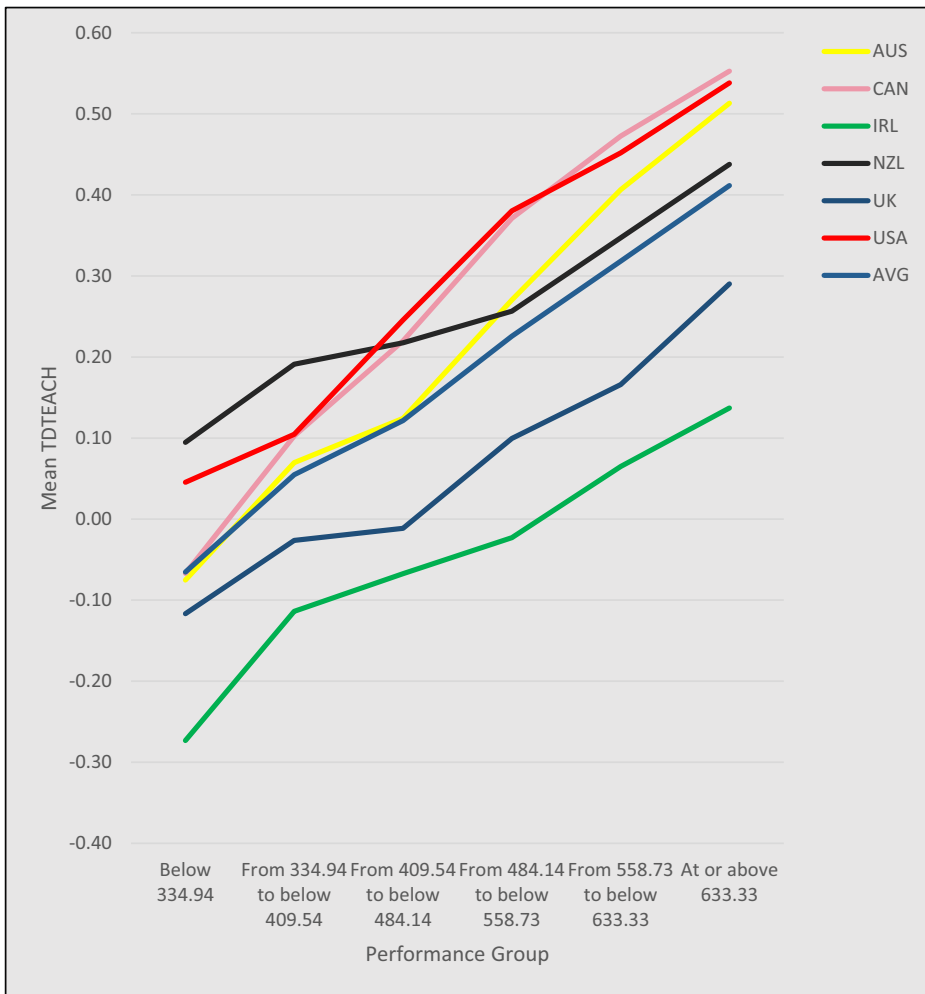


Fig. 2 Mean levels of teacher-directed instructional activities at six science literacy performance benchmarks for students in six countries in PISA 2015

typically not use one approach exclusively. Table 3 provides the results of this multivariate analysis.

As shown in Table 3, student-level socioeconomic status (ESCS) consistently accounts for a substantial proportion of the variance in students’ scientific literacy across the six countries in this study. Specifically, a one unit increase in student ESCS is associated with increases in scientific literacy of between 31 (Canada, USA) and 43 (New Zealand) PISA score points, on average. Also consistently, and in the context of the other composite variables representing distinguishable pedagogical approaches to teaching and learning science, both teacher-directed and adaptive instruction showed positive, albeit moderately strong, associations with scientific literacy, *net of students’ socioeconomic status*. For example, the regression coefficients associated with teacher-directed instruction ranged between 9.75 (Ireland) and 11.72 (Australia). In other words, a one-unit increase in teacher-directed activities is associated with a 12-point increase in scientific literacy for Australian students, on average. Similarly, a one-unit increase

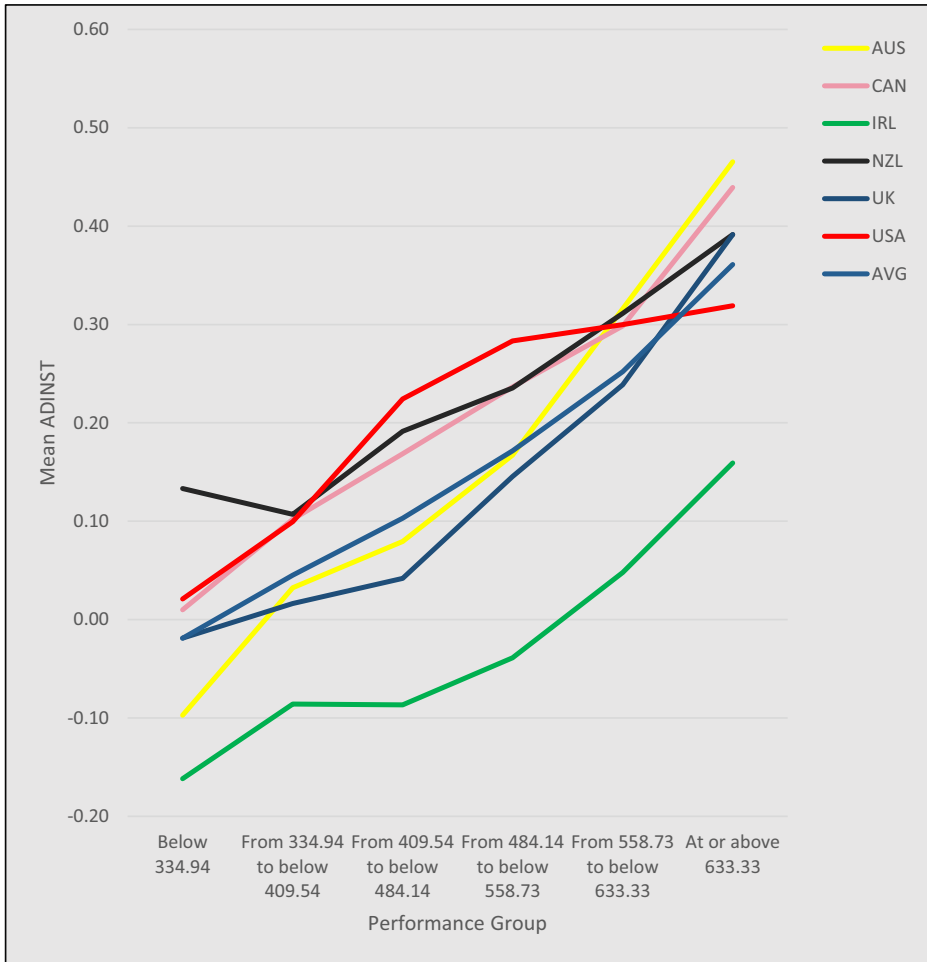


Fig. 3 Mean levels of adaptive instruction in science at six literacy performance benchmarks for students in six countries in PISA 2015

in adaptive instruction would suggest a 13-point net increase in scientific literacy for UK students, on average.

Table 3 Multivariate (simultaneous solution) regression coefficients for science literacy on three approaches to teaching science in PISA 2015

	Constant (a)	ESCS (b1)	IBTEACH (b2)	TDTEACH (b3)	ADINST (b4)
AUS	508.47	38.62	- 15.13	11.72	11.83
CAN	518.03	31.42	- 17.08	10.46	7.56
IRL	504.67	35.53	- 13.44	9.75	6.64
NZL	519.12	42.81	- 25.68	10.74	9.77
GBR	508.62	36.24	- 15.49	8.07	12.65
USA	499.94	31.98	- 15.83	11.57	4.31

$$\hat{Y} \text{ science literacy} = a + b1ESCS + b2IBTEACH + b3TDTEACH + b4ADINST$$

All regression coefficients in Table 3 are statistically significant ($p < 0.05$)

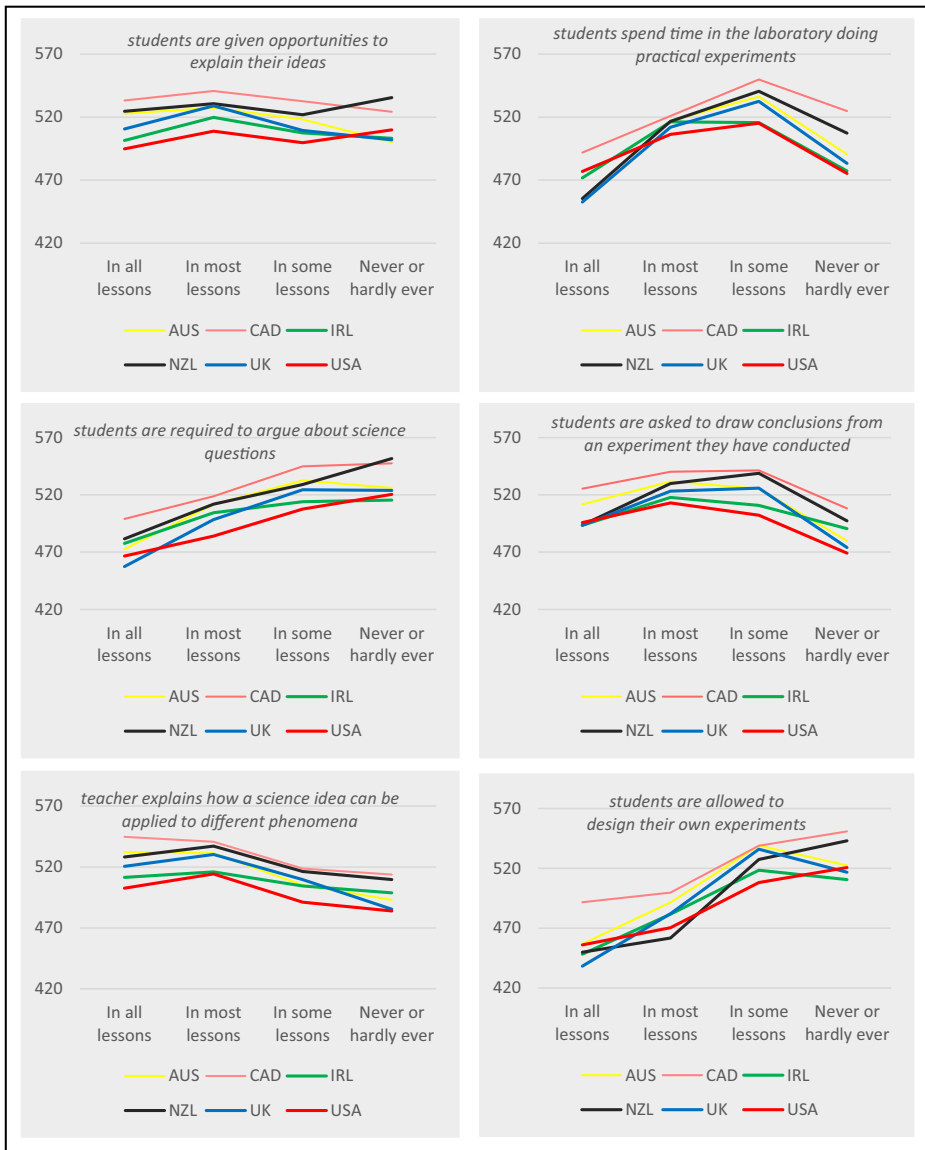


Fig. 4 Associations between students’ scientific literacy and the frequency with which they experience various inquiry-based instructional strategies across six countries participating in PISA 2015. In PISA 2015, IBTEACH comprises the following: (1) students are given opportunities to explain their ideas; (2) students spend time in the laboratory doing practical experiments; (3) students are required to argue about science questions; (4) students are asked to draw conclusions from an experiment they have conducted; (5) the teacher explains how a science idea can be applied to different phenomena; (6) students are allowed to design their own experiments; (7) there is a class debate about investigations; (8) the teacher clearly explains the relevance of science concepts; and (9) students are asked to do an investigation to test ideas (OECD 2016a, b)

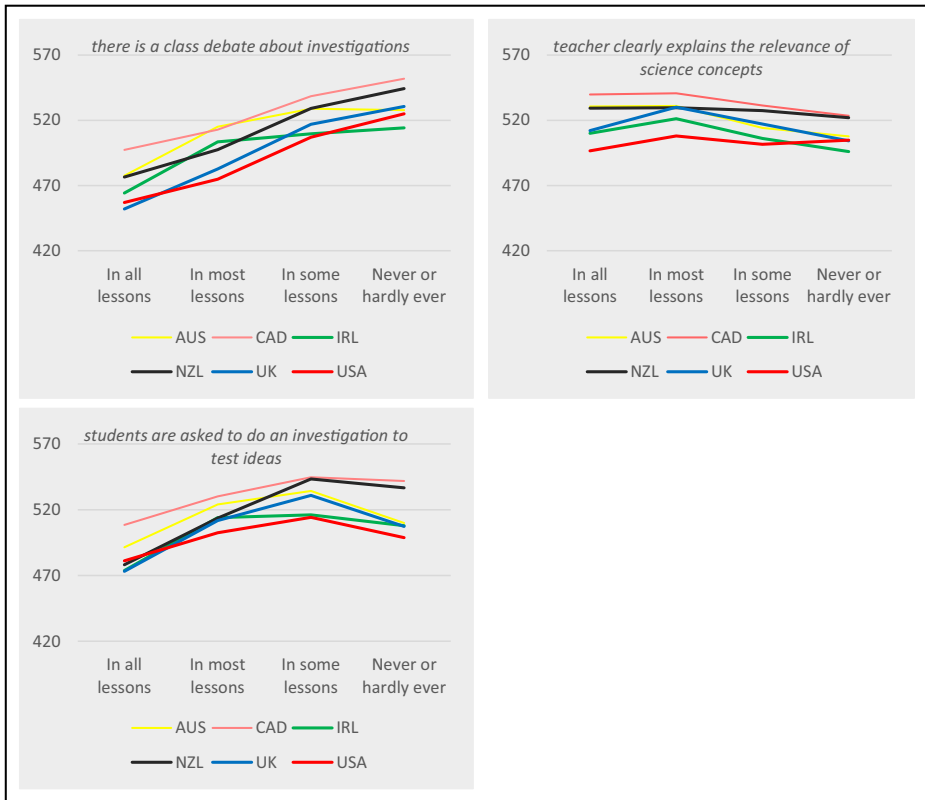


Fig. 4 (continued)

Quite different in both direction and magnitude, however, were the regression coefficients associated with inquiry-based instruction in science. As seen in Table 3, across the six countries, regression coefficients associated with inquiry are consistently negative and larger than coefficients associated with teacher-directed or adaptive instruction. For example, a one-unit increase in inquiry-based activities is associated with net decreases on average of 26 PISA score points for students in New Zealand, and 15 score points for students in the UK, respectively.

The regression coefficients generated for each composite variable and representing a distinct instructional approach to teaching science appear to support the patterning observed in the benchmark analyses. Controlling for the variance in students' scientific literacy associated with student SES, teacher-directed and adaptive instruction are positively associated, albeit modestly, with scientific literacy across the six countries. On the other hand, inquiry-based instruction taken as a whole is strongly and negatively associated with students' scientific literacy at similar levels of magnitude in all six countries. Based on these findings, the nature of the association between inquiry-based instruction (represented as a composite index) and secondary students' scientific literacy seems clear for these six countries; it is negative, and in magnitude ranges between two-fifths (Ireland) and four-fifths (New Zealand) of a school year's learning in science (estimating that 30 score points in PISA equals about one school year's learning) (Thomson et al. 2016).

Nevertheless, given recent research suggesting that inquiry comprises several distinguishable conceptual and pedagogical aspects (Furtak et al. 2012; Hodson 2014) that have been shown to be differentially effective in learning and teaching science (Capps and Crawford 2013; Furtak et al. 2012), we were prompted to ask one further question. Specifically, in research question 2, we asked whether the association seen between inquiry-based instruction as a composite index variable and students' scientific literacy might appear differently if inquiry were represented by its individual component parts.

In PISA 2015, inquiry-based instruction is comprised of nine items. As shown in Fig. 4, we provide our analysis of the nine items for the six countries included. For each item that makes up inquiry-based instruction, Fig. 4 represents the association between mean scientific literacy and the frequency with which students spend time in their science classrooms engaged in various teaching and learning activities. Immediately apparent from this analysis of the nine items that make up PISA's composite index IBTEACH is that not all items have the same types of association with students' scientific literacy, at least for the six countries included in this analysis.

At least three types of association seem evident between students' scientific literacy and the frequency with which they experience various inquiry-based instructional activities. First, there are a couple of items for which the rate at which students experience the instructional strategy appears *unrelated* to students' scientific literacy. An example of this type of relationship is the item that asks students the frequency at which they are *given opportunities to explain their ideas* (Fig. 4, top left). For this item, the frequency with which students experience opportunities to explain their ideas in science seems to have no association with students' scientific literacy. A second type of association can be characterised as a negative, linear relationship between the frequency with which students experience the instructional activity and their scientific literacy. Examples of this type of association are seen for items that ask students about the frequency with which they experience *class debates in science* (Fig. 4, bottom left), or the frequency with which they are *required to argue about science questions* (Fig. 4, top left). For both of these instructional strategies, higher frequencies of the strategy are seemingly associated with lower levels of scientific literacy.

A third type of association, perhaps the most interesting for science teacher educators and science teachers, is also evident. This type can be characterised as non-linear (curvilinear). Two examples are the items that ask students about the frequency with which they experience *spending time in the laboratory doing practical experiments* (in our view, a sine qua non of inquiry-based science education) and the item asking how often students experience *drawing conclusions from an experiment they have conducted*, a widely recognised critical aspect of inquiry-oriented learning and teaching in science (Furtak et al. 2012; McConney et al. 2014; Minner et al. 2010). In the first case (Fig. 4, top right), for each country, the highest level of students' scientific literacy is associated with spending time doing practical experiments in *some lessons*, rather than in most or all lessons. Similarly, in drawing conclusions from an experiment they have conducted (Fig. 4, middle right), in all six countries, higher levels of scientific literacy are associated with students engaged this activity in *some lessons* (Canada, New Zealand, UK) or in *most lessons* (Australia, Ireland, USA) rather than in all lessons or never. This more nuanced, non-linear patterning was remarkably consistent across the six countries included in this analysis.

In conjunction with the descriptive statistics of Tables 1 and 2 and the regression analyses presented in Table 3, these item-level analyses of the nine items comprising PISA's composite index representing inquiry-based instruction provide the beginnings of an answer to research

question 2. It seems evident that a different, more nuanced picture emerges from these item-wise analyses, in contrast with the analysis of inquiry-based instruction as a composite index variable (IBTEACH). This more nuanced picture suggests that some aspects of what is typically considered “inquiry-based” instruction are more effective than others with regard to association with scientific literacy. It also seems clear, at least for some of the nine items, that the association between scientific literacy and the frequency at which students experience particular instructional strategies is better thought of as non-linear. For some items, like doing practical experiments, more is not necessarily, or always, better in terms of students’ scientific literacy.

Caveat

Prior to discussing our findings, we readily acknowledge that we have no insight into the *quality* of students’ pedagogical experiences in science classrooms in the six countries we examined here, or into teachers’ beliefs and practices around using inquiry-based instruction. IBTEACH, ADINST, and TDTEACH in PISA 2015 are composite index variables that reflect *frequency*, not *quality*, and are based on students’ reports of their classroom experiences. We also cannot be sure that students’ efforts in PISA are uniform or stable across countries (Hopfenbeck et al. 2018) compared with their motivation or effort in high-stakes domestic assessments. In gaining additional insights into these issues, classroom observations could be used to investigate students’ and teachers’ experiences or video studies could be used to verify students’ and teachers’ reports about various instructional approaches. Nonetheless, the consistency of the patterns we have observed across these six countries and the consistency of this secondary analysis with those reported in the primary analysis (OECD 2016a, b) deserve further investigation and reflection, in our view. We have previously explored the relationships between scientific literacy, engagement in science, and inquiry-based teaching and learning using PISA 2006 data (McConney et al. 2014). Additional analyses might include more complex multilevel models that involve more variables and reflect the nested structure of students, classrooms, and schools. Further, Jerrim, Oliver and Sims (2019) developed and tested mathematical models in exploring relationships between PISA scores and attainment. In this, we hope to advance the discussion around evidence-informed pedagogy to help science teachers, science teacher educators, and science education policy makers understand the nuanced complexity of teaching and learning science and achieving scientific literacy.

Discussion

PISA 2015 asked students about their classroom experiences with respect to inquiry-based, teacher-directed, and adaptive instruction in science. The intention of this study is twofold: first, to examine associations between the variation in students’ scientific literacy and the instructional approaches students experience in learning science, and second, to examine the relationship between students’ scientific literacy and disaggregated aspects of inquiry-based instruction as operationalised by PISA. To achieve these intentions, we examined relationships between students’ scientific literacy and inquiry-based, teacher-directed, and adaptive instructional approaches across the six countries included. In so doing, we made no assumptions or

judgments about the quality (how well teachers in each of the six countries use different strategies) of the three instructional approaches that students report experiencing.

We share PISA's view that "what happens inside the classroom is crucial for students' learning" (OECD 2016b, p. 228) and note that across these six countries, there are striking commonalities in the performance of students associated with the three instructional approaches. With regard to mean scores for scientific literacy, this group of countries comprises a range including top performing (Canada), above average (Australia, Ireland, New Zealand, UK), and average (USA). Across the six, on average, students who report experiencing high levels of teacher-directed and adaptive instruction in science achieve more strongly in scientific literacy compared with students reporting lower levels of these two approaches. On the other hand, students experiencing high levels of inquiry-based instruction in their science classes typically show lower levels of scientific literacy as compared with within-country peers reporting lower frequencies of inquiry-based instructional activities. We observed these relationships generally across all six countries. Further, this finding is consistent with previous research, including our own (McConney et al 2014; Cairns and Areepattamannil 2017; Jiang and McComas 2015). In our view, the consistency of this patterning warrants close examination.

While inquiry-based instruction holds different meanings for different stakeholders, PISA's composite variable (IBTEACH) reflects a relatively broad spectrum of pedagogical strategies arguably associated with the approach. Considering the relationship between inquiry-based instruction and achievement in science, others have argued that certain aspects of inquiry might be more (or less) effective (Capps and Crawford 2013; Furtak et al. 2012; Hodson 2014; Lau and Lam 2017; Osborne 2015). PISA 2015 thus provided another opportunity to interrogate empirically these views by considering separately the associations between scientific literacy and different aspects of inquiry. To do this, we disaggregated PISA's composite variable (IBTEACH) and mapped the frequency of students experiencing each component item against their scientific literacy.

The components of inquiry, operationalised by PISA, include opportunities for students to explain their ideas, spend time in the laboratory doing practical experiments, having class debates about science questions, drawing conclusions from experiments, and, designing their own experiments to test their ideas. The analysis presented here reveals a complex, often non-linear, pattern of associations between students' scientific literacy and component aspects of inquiry (Fig. 4). For example, the association between students spending time in the laboratory designing their own experiments and scientific literacy could be perceived as negative, particularly in Canada and New Zealand, where students who report this activity *never* or *hardly ever* have higher scientific literacy scores on average than students who report this activity *in all lessons*. Finer examination, however, reveals that students who reported spending time in the laboratory doing experiments in *some lessons* are those who have the highest scores in scientific literacy for 5 of the 6 countries (Ireland being the exception, where students reporting the activity in *most lessons* are those with the highest scientific literacy, on average) (see Fig. 4). Similarly, regarding students drawing conclusions based on an experiment they conducted, those who reported doing this in *some* or *most lessons* have stronger scientific literacy on average compared with peers who report doing this either *in all lessons* or *never* or *hardly ever* (see Fig. 4). This finding shows little variation across countries. In New Zealand, for example, students with higher scientific literacy are those reporting drawing conclusions from an experiment *in some lessons*. In the other Anglophone countries, higher scientific literacy is associated with students reporting this activity *in most lessons*.

In PISA 2015, another aspect of inquiry-based instruction was the frequency with which students are asked *to do an investigation to test ideas*. On average, the strongest scientific literacy performance is associated with students who reported they experienced this *in some lessons*. Students reporting the activity *in most* or *all lessons* achieved considerably lower scientific literacy, on average. This finding would appear to question advice encouraging practical, investigative work in science, which some have argued helps students learn about the processes as well as the concepts of science. Physical action during investigative work may well increase students' cognitive load and "constrain the learners from thinking about the problem" (Zhang 2018a, p. 5). This is especially so when an investigation is an *open inquiry* wherein students manipulate experimental materials but are not provided with answers (Zhang 2018b), an approach that more authentically replicates scientists' work. By contrast, students in a *direct instruction* group who saw a demonstration of the same experiment "performed the best in gaining the class content knowledge and reasoning with the content" (Zhang 2018a, p. 6). This finding is consistent with similar research reported by Klahr and Nigam (2004) almost 15 years ago.

That the item-level patterning that we have depicted in Fig. 4 is generally consistent across all countries should raise questions for both researchers and policy makers. For example, is doing frequent practical work associated mainly with particular cohorts of students or types of school? Are students in disadvantaged schools or streamed classes more likely to experience a hands-on approach to teaching and learning in science? Do more academically able students experience a more examination-focused curriculum compared with their peers in classes tailored more toward vocational education? We cannot know the answers to these questions from these data. We do know that previous research has shown that aspects of inquiry that support students' science learning tend to be cognitive rather than behavioural or procedural (Furtak et al. 2012). Cognitive strategies require students to link their practical investigative findings to science concepts, thus privileging evidence-informed conclusions rather than focusing mainly on the skills of planning investigations and gathering data (Osborne 2015). However, the analyses shared in Fig. 4 seem to suggest a different relationship for some items within IBTEACH. Could it be that social factors, like requiring public argument or debate in science class, moderate the typically positive effects of cognitively focused inquiry pedagogy? These are questions for future research.

These finer-grained distinctions regarding the optimal frequencies at which various aspects of inquiry-based instruction in secondary school science become observably effective must also be important and of interest for teachers and teacher educators. For example, doing practical work in *every lesson* or *very rarely* is unlikely to support the development of students' scientific literacy. Importantly, these findings show that "inquiry" is not only multifaceted, but also its relationship with scientific literacy varies according to the particular strategy being examined, and is often best conceptualised as non-linear. An important message from this research is that teacher educators and policy makers hold responsibility to support the development of pre- and in-service teachers using carefully developed evidence that informs recommended practice. In our view, this commitment to a finer-grained examination of pedagogical strategies applies as much to teacher-directed and adaptive instructional strategies as it does to inquiry-based teaching, as described here.

To be clear, we do not advocate one instructional approach over others in teaching and learning science. Inquiry is currently a much-favoured pedagogy and has been shown effective in supporting students' engagement in science (McConney, et al. 2014; Sjøberg 2018). Additionally, undertaking and developing experimental work are important and can support

students' skill acquisition, learning, and interest in science (Sjøberg 2018). Further, with appropriate guidance from the teacher, inquiry-based instruction has been shown to support students' science achievement (Minner et al. 2010). Nevertheless, there have also been critics of the approach (Kirschner et al. 2006). In our view, rather than a question of *whether* to implement inquiry-based pedagogy, the question may be better framed as *how often* a teacher might use inquiry-based instruction, and *for what purposes?* Just like Goldilocks, there may be a level of use that is not too much and not too little, but just right. To simply assume a position about inquiry as an advocate or critic seems unwise and ultimately a disservice to both science education research and the teaching profession.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Andrew McConney, Mary Oliver, Amanda Woods-McConney, Renato Schibeci, (2011) Bridging the Gap? A Comparative, Retrospective Analysis of Science Literacy and Interest in Science for Indigenous and Non-Indigenous Australian Students. *International Journal of Science Education* 33 (14):2017-2035
- Areepattamannil, S. (2012). Effects of inquiry-based science instruction on science achievement and interest in science: evidence from Qatar. *The Journal of Educational Research*, 105(2), 134–146.
- Australian Curriculum, Assessment and Reporting Authority (ACARA). (2019). The three interrelated strands of science. Retrieved from <https://www.australiancurriculum.edu.au/f-10-curriculum/science/structure/>.
- Author (n.d.-a).
- Author (n.d.-b).
- Authors (n.d.).
- Cairns, D., & Areepattamannil, S. (2017). Exploring the relations of inquiry-based teaching to science achievement and dispositions in 54 countries. *Research in Science Education*, 1-23. <https://doi.org/10.1007/s11165-017-9639-x>.
- Capps, D. K., & Crawford, B. A. (2013). Inquiry-based professional development: what does it take to support teachers in learning about inquiry and nature of science? *International Journal of Science Education*, 1–32. <https://doi.org/10.1080/09500693.2012.760209>.
- Crawford, B. A. (2014). From inquiry to scientific practices in the science classroom. In N. G. Lederman & S. K. Abell (Eds.), *Handbook of research on science education* (Vol. II, pp. 515–541). New York: Routledge.
- Furtak, E. M., Seidel, T., Iverson, H., & Briggs, D. C. (2012). Experimental and quasi-experimental studies of inquiry-based science teaching. *Review of Educational Research*, 82(3), 300–329. <https://doi.org/10.3102/0034654312457206>.
- Harlen, W. (2013). Assessment and inquiry-based science education: issues in policy and practice. Trieste: Global Network of Science Academies (IAP) Science Education Programme.
- Hattie, J. (2009). *Visible learning: a synthesis of over 800 meta-analyses relating to achievement* London. UK: Routledge.
- Hodson, D. (2014). Learning science, learning about science, doing science: different goals demand different learning methods. *International Journal of Science Education*, 36(15), 2534–2553. <https://doi.org/10.1080/09500693.2014.899722>.
- Hopfenbeck, T. N., Lenkeit, J., El Masri, Y., Cantrell, K., Ryan, J., & Baird, J.-A. (2018). Lessons learned from PISA: a systematic review of peer-reviewed articles on the programme for international student assessment. *Scandinavian Journal of Educational Research*, 62(3), 333–353. <https://doi.org/10.1080/00313831.2016.1258726>.
- Jiang, F., & McComas, W. F. (2015). The effects of inquiry teaching on student science achievement and attitudes: evidence from propensity score analysis of PISA data. *International Journal of Science Education*, 37(3), 554–576. <https://doi.org/10.1080/09500693.2014.1000426>.
- John Jerrim, Mary Oliver, Sam Sims, (2019) The relationship between inquiry-based teaching and students' achievement. New evidence from a longitudinal PISA study in England. *Learning and Instruction* 61:35-44

- Kawalkar, A., & Vijapurkar, J. (2011). Scaffolding science talk: the role of teachers' questions in the inquiry classroom. *International Journal of Science Education*, 35(12), 2004–2027. <https://doi.org/10.1080/09500693.2011.604684>.
- Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: an analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, 41(2), 75–86.
- Klahr, D. (2013). What do we mean? On the importance of not abandoning scientific rigor when talking about science education. *Proceedings of the National Academy of Sciences*, 110(Supplement 3), 14075–14080. <https://doi.org/10.1073/pnas.1212738110>.
- Klahr, D., & Nigam, M. (2004). The equivalence of learning paths in early science instruction. *Psychological Science*, 15(10), 661–667.
- Lau, K.-C., & Lam, T. Y.-P. (2017). Instructional practices and science performance of 10 top-performing regions in PISA 2015. *International Journal of Science Education*, 39(15), 2128–2149. <https://doi.org/10.1080/09500693.2017.1387947>.
- Lazonder, A. W., & Harmsen, R. (2016). Meta-analysis of inquiry-based learning: effects of guidance. *Review of Educational Research*, 86(3), 681–718.
- Minner, D. D., Levy, A. J., & Century, J. (2010). Inquiry-based science instruction—what is it and does it matter? Results from a research synthesis years 1984 to 2002. *Journal of Research in Science Teaching*, 47(4), 474–496. <https://doi.org/10.1002/tea.20347>.
- Nadelson, L., Williams, S., & Turner, H. (2011). *Influence of inquiry-based science interventions on middle school students' cognitive, behavioral, and affective outcomes*. The Campbell Corporation.
- National Research Council. (2012). *A framework for K-12 science education: practices, crosscutting concepts, and core ideas*. Committee on a conceptual framework for new K-12 science education standards. Board on science education, division of behavioral and social sciences and education. Washington: The National Academies Press.
- OECD. (2004). *Learning for Tomorrow's world: first results from PISA 2003*. Paris: OCED.
- OECD. (2007). *PISA 2006: science competencies for tomorrow's world*. Paris: OECD.
- OECD. (2009). PISA 2009 results: what students know and can do. In *Student performance in reading, mathematics and science*. Paris: Author.
- OECD. (2016a). *PISA 2015 results (volume I): excellence and equity in education*. Paris: PISA, OECD Publishing. <https://doi.org/10.1787/9789264266490-en>.
- OECD. (2016b). *PISA 2015 results (volume II): policies and practices for successful schools*. Paris: PISA, OECD Publishing. <https://doi.org/10.1787/9789264267510-en>.
- OECD. (2018). *PISA 2015 results in focus*. Paris: PISA, OECD Publishing <http://www.oecd.org/pisa/pisa-2015-results-in-focus.pdf>.
- Office for Standards in Education, Children's Services and Skills. (2013). *Maintaining curiosity: a survey into science education in schools*. Manchester, 2013.
- Osborne, J. (2015). Practical work in science: misunderstood and badly used? *School Science Review*, 96(357), 16–24.
- Rennie, L. J. (2010). *Evaluation of the science by doing stage one professional learning approach 2010*. Australian Academy of Science: Canberra.
- Rocard, M. (2007). Science education NOW: a renewed pedagogy for the future of Europe, Brussels: EurEopean Commission. Retrieved from: http://ec.europa.eu/research/science-society/document_library/pdf_06/report-rocard-onscience-education_en.pdf (2.06.2015).
- Roseman, J. E., Hermann-Abell, C. F., & Koppal, M. (2017). Designing for the next generation science standards: educative curriculum materials and measures of teacher knowledge. *Journal of Science Teacher Education*, 28(1), 111–141. <https://doi.org/10.1080/1046560X.2016.1277598>.
- Sjøberg, S. (2016). OECD, PISA, and globalization: the influence of the international assessment regime *Education Policy Perils. Tackling the Tough Issues*. (pp. 102–133): Routledge.
- Sjøberg, S. (2018). The power and paradoxes of PISA: should inquiry-based science education be sacrificed to climb on the rankings? *Nordic Studies in Science Education*, 14(2), 186–202.
- Teig, N., Scherer, R., & Nilsen, T. (2018). More isn't always better: the curvilinear relationship between inquiry-based teaching and student achievement in science. *Learning and Instruction*, 56, 20–29. <https://doi.org/10.1016/j.learninstruc.2018.02.006>.
- Thomson, S., Hillman, K., & De Bortoli, L. (2013). *A teacher's guide to PISA scientific literacy*. Camberwell: Australian Council for Educational Research Ltd.
- Thomson, S., De Bortoli, L., & Underwood, C. (2016). *PISA 2015: a first look at Australia's results*. Camberwell: Australian Council for Educational Research Ltd.

- Woods-McConney, A., Oliver, M., McConney, A., Maor, D., & Schibeci, R. (2013a). Science Engagement and Literacy: A Retrospective Analysis for Indigenous and Non-Indigenous Students in Aotearoa New Zealand and Australia. *Research in Science Education*, 43(1): 233–252. <https://doi.org/10.1007/s11165-011-9265-y>
- Woods-McConney, A., Oliver, M. C., McConney, A., Schibeci, R., & Maor, D. (2013b). Science Engagement and Literacy: A retrospective analysis for students in Canada and Australia. *International Journal of Science Education*, 36(10), 1588–1608. <https://doi.org/10.1080/09500693.2013.871658>
- United Nations Development Programme (UNDP). (2016). Human development report 2016: human development for everyone. Retrieved from: <http://hdr.undp.org/en/2016-report/download>
- Wellcome (2017). A review of Ofsted inspection reports: science. Retrieved from: <https://wellcome.ac.uk/sites/default/files/review-of-ofsted-inspection-reports-2017.pdf>
- Zhang, L. (2018a). “Hands-on” plus “inquiry”? Effects of withholding answers coupled with physical manipulations on students’ learning of energy-related science concepts. *Learning and Instruction*. <https://doi.org/10.1016/j.learninstruc.2018.01.001>.
- Zhang, L. (2018b). Withholding answers during hands-on scientific investigations? Comparing effects on developing students’ scientific knowledge, reasoning, and application. *International Journal of Science Education*, 40(2), 1–11 Published online 26 Jan, 2018. <https://doi.org/10.1080/09500693.2018.1429692>.

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.