International Journal of Science Education

ISSN: 0950-0693 (Print) 1464-5289 (Online) Journal homepage: https://www.tandfonline.com/loi/tsed20

Forms of inquiry-based science instruction and their relations with learning outcomes: evidence from high and low-performing education systems

Anindito Aditomo & Eckhard Klieme

To cite this article: Anindito Aditomo & Eckhard Klieme (2020): Forms of inquiry-based science instruction and their relations with learning outcomes: evidence from high and low-performing education systems, International Journal of Science Education, DOI: 10.1080/09500693.2020.1716093

To link to this article: https://doi.org/10.1080/09500693.2020.1716093

INTERNATIONAL JOURNAL OF SCIENCE EDUCATION
https://doi.org/10.1080/09500693.2020.1716093

Forms of inquiry-based science instruction and their relations with learning outcomes: evidence from high and low-performing education systems

Anindito Aditomo [a,b] and Eckhard Klieme [b]

[a]Faculty of Psychology, University of Surabaya, Surabaya, Indonesia; [b]Leibniz Institute for Research and Information in Education, Frankfurt am Main, Germany

ABSTRACT

Inquiry-based science instruction is widely advocated, but studies based on international large-scale assessments often show inquiry to be negatively associated with achievement. We re-examine this issue by examining whether the association between inquiry and learning depends upon the provision of teacher guidance. Participants were 151,721 students from 5089 schools from 10 highest and 10 lowest science performers in PISA 2015. Multi-group confirmatory factor analyses found that measurement invariance cannot be established, suggesting substantial regional variation in the pattern of inquiry-based instruction. Nonetheless, exploratory factor analyses indicated that at the conceptual level, many regions exhibit a pattern which contrasted between 'Guided inquiry' and 'Independent inquiry'. Results of structural equation modelling showed that inquiry is positively associated with outcomes when it incorporates teacher guidance, and negatively when it doesn't. However, the strength of the positive associations is stronger in regions where guided inquiry is measured with fewer items referring to student-centred activities. These findings are in line with current theories regarding the importance of scaffolding in learning from inquiry. This study suggests that it would be misguided to use PISA findings to support arguments to scale back inquiry and other constructivist approaches to teaching science.

ARTICLE HISTORY
Received 15 January 2019
Accepted 10 January 2020

KEYWORDS
Science achievement; scientific literacy; epistemic belief; intrinsic motivation; inquiry-based instruction; large-scale assessment of learning; measurement invariance

Instructional approaches broadly described as 'inquiry-based' are considered to be essential for developing students' scientific literacy (Engeln, Mikelskis-seifert, & Euler, 2014). Studies based on international large-scale assessments (ILSA), however, often indicate inquiry to be associated with lower science achievement (Cairns & Areepattamannil, 2017; Chi, Liu, Wang, & Won Han, 2018; Grabau & Ma, 2017). Given the influence which ILSA can have on educational policy (Grek, 2009), such findings have raised concerns among advocates of inquiry-based instruction in science (Sjøberg, 2018; Zhang, 2016).

CONTACT Anindito Aditomo aditomo@staff.ubaya.ac.id, aditomo@dipf.de Room 10-04, Rostocker Strasse 6, Frankfurt am Main 60323, Germany

Supplemental data for this article can be accessed at https://doi.org/10.1080/09500693.2020.1716093.

We re-examine this issue by taking a more nuanced look at the inquiry-based instruction measure in the Program for International Student Assessment (PISA) 2015. Unlike prior studies, we explicitly test the assumption of comparability (measurement invariance) across regions, and show that patterns of instruction are better characterised as being region-specific rather than universal. Building upon the variation of regional patterns, we also test the prediction that inquiry-based instruction is positively associated with learning outcomes when it involves some form of teacher guidance. In the following sections, we first summarise theoretical perspectives which link inquiry-based instruction and student learning. We then consider prior studies of inquiry-based science instruction based on ILSA and note areas of limitations which this study seeks to remedy.

Inquiry and science learning

Inquiry-based instruction involves engaging students in formulating questions, collecting and analysing data, and reasoning and arguing about what the results mean (Barron & Darling-Hammond, 2008; Hmelo-silver, Duncan, & Chinn, 2007). Inquiry prompts active knowledge construction and thereby facilitates deeper learning (Barron & Darling-Hammond, 2008). It has become especially prominent in science education as part of the shift towards the 'practice turn' which recognises scientific practices as a central organising theme for teaching and learning (National Research Council, 2012).

The 'practice turn' is underpinned by socio-cultural theories with their central metaphor of learning as participation (Forman, 2018; Sfard, 1998). In this view, learning is the process of becoming a member of a community of practice. It is less about acquiring and having knowledge, and more about becoming able to participate in activities which are valued by a community, communicate using the language of that community, and act in ways which conform the community's norms. Science learning thus means participating in authentic scientific practices, i.e. constructing models/theories which explain some aspect of the natural world and arguing for their value and validity (Osborne, Simon, Christodoulou, & Howell-richardson, 2013; Windschitl, Thompson, & Braaten, 2008). Consequently, inquiry-based science instruction should weave together the conceptual, epistemic, and social dimensions of scientific practice (Duschl, 2008).

This conception of inquiry-based instruction highlights the non-cognitive dimensions of learning. Becoming competent participants of scientific practice involves changing beliefs about the nature of science, and thus good inquiry-based instruction should not only develop conceptual understanding, but also more mature epistemic beliefs (Sandoval, Greene, & Bråten, 2016). Such beliefs include, for instance, an understanding that scientific knowledge is subject to revision, and that knowledge is based on empirical evidence whose meaning is influenced by the models/theories which scientists employ (Duschl, 2007; Pluta, Chinn, & Duncan, 2011). In addition, inquiry may provide students more autonomy (e.g. in formulating questions and choosing how to address them) and opportunities for meaningful interactions and positive relationships. Thus, authentic inquiry has the potential to improve students' intrinsic motivation through the fulfilment of basic psychological needs (Ryan & Deci, 2000).

Critics charge that inquiry is unstructured and impose irrelevant cognitive which impede learning (Kirschner, Sweller, & Clark, 2006). However, inquiry-based instruction does not necessarily be unstructured. Guided forms of inquiry incorporate various scaffolds to guide learners' meaning-making process. For instance, expert guidance can be embedded as 'just-in-time' mini-lectures; tasks can be sequenced to reduce cognitive load; and tools can be designed to model or make salient disciplinary strategies (Hmelo-silver et al., 2007). Indeed, experimental studies have shown that innovative inquiry-based interventions are superior to conventional science teaching (Furtak, Seidel, Iverson, & Briggs, 2016; Lazonder & Harmsen, 2016).

Prior ILSA studies

ILSA such as PISA and TIMSS include measures designed to assess instructional practices, including inquiry-based ones. Accordingly, secondary analyses of ILSA data related to inquiry have been published. Some treated the PISA inquiry scale as a single index and found that higher frequencies of inquiry activities were related with lower science literacy (Cairns & Areepattamannil, 2017; Chi et al., 2018). A study using the TIMSS 2007 data also reported that 'student-oriented instruction', which reflect planning and conducting observations and investigations, was negatively associated with science achievement (Liou & Jessie Ho, 2018). Other studies found that different dimensions of the inquiry measure are related differently to outcomes. For example, science literacy was positively related with 'student investigation' activities, but negatively with 'hands-on' activities in the USA sample of PISA 2006 (Grabau & Ma, 2017). For the Qatar sample from the same data, however, both the hands-on and student investigation dimensions of inquiry were negatively related with science literacy (Areepattamannil, 2012).

Collectively these studies have contributed to the empirical base related to inquiry-oriented instruction as practiced in nationally representative schools in many countries/regions. A number of important limitations need to be noted, however. Methodologically, these studies assume that the same pattern of instructional practice exist across regions and can be measured using the same instrument. Given the possibility that instructional practices are region-specific, or that the measures function differently across the regions, measurement invariance needs to be explicitly tested (Wu, Li, & Zumbo, 2007). Conceptually, previous studies have not attempted to provide theoretically grounded explanations regarding the association between inquiry and learning. Without a theoretical account, findings such as differential relations between distinct dimensions of inquiry and learning outcomes are difficult to interpret. We argue that it is possible to propose and test a theoretical explanation regarding the relationship between inquiry and learning using ILSA data.

Current study

Given the evidence from experimental studies about the efficacy of inquiry-based instruction, negative associations between inquiry and learning/achievement found in ILSA studies call for an explanation. One possibility is simply that the positive experimental evidence reflects the effects of innovative programmes in selected school/classroom settings, whereas findings from ILSA studies reflect inquiry activities as practiced in the 'regular' or typical school. This conjecture is supported by research which shows that successful enactment of inquiry-based curricula requires extensive training and support (Fitzgerald, Danaia, & McKinnon, 2017; Fogleman, McNeill, & Krajcik, 2011), which is unlikely to be available for teachers in the 'average' school.

Another possible explanation is related to the level of guidance/structure. Teacher guidance can be seen as a form of structure necessary to facilitate learning, especially in complex activities such as scientific inquiry (Hmelo-silver et al., 2007; Schmidt, Loyens, & Paas, 2007). Without adequate guidance, learners may be overwhelmed by unessential features of the activity and fail to construct meaningful knowledge (Kirschner et al., 2006). Measures of inquiry-based instruction in PISA, specifically, include items which refer to student-independent activities (e.g. 'Students are asked to do an investigation to test ideas'), as well as those which refer to teacher guidance (e.g. 'The teacher explains how science ideas can be applied'). Thus, it may be possible to use PISA to test whether the association between inquiry and outcomes depends on teacher guidance.

These conjectures can be tested by comparing guided and unguided forms of inquiry. A recent study found that PISA's inquiry-based instruction scale form two separate dimensions, one reflecting teacher-guided interactive instruction and the other reflecting unguided inquiry (Lau & Lam, 2017). However, these authors utilised only 6 of the 9 available items, thereby further narrowing the scope of the construct. Furthermore, while the authors analysed data from high-performing regions, they did not examine whether the measurement model was statistically invariant/equivalent across the regions. Thus, the existence of alternative measurement models (reflecting different instructional patterns across the regions) cannot be ruled out.

We build upon and extend Lau and Lam's (2017) study in several ways. First, we expand the generality of the findings by analysing high and low-performing regions. It would be important to examine whether teacher guidance and inquiry can be effectively implemented by teachers in low-performing regions, where teacher competence is generally lower and school resources are more limited (Scheerens, 2001). Second, we explicitly test whether the single-factor structure (reflecting PISA's original design) and Lau and Lam's (2017) two-factor measurement models of inquiry-based instruction are statistically invariant across the selected regions. Third, we examine how inquiry relates to intrinsic motivation and epistemic beliefs as non-cognitive outcomes. We formulate the following research questions to structure our analysis and presentation of results:

(1) (a) Can the same forms of inquiry-based instructional practices be observed across high and low-performing regions, and (b) if not, what regional forms could be identified?

To answer this question, we tested the measurement invariance of a two-dimensional model which distinguishes between teacher-guided instruction and unguided inquiry based on Lau and Lam's (2017) study, and compared it to a unidimensional model of inquiry as intended by the PISA questionnaire designers (Müller, Prenzel, Seidel, Schiepe-tiska, & Kjærnsli, 2016).

(2) (a) How do the different forms of instruction relate to learning outcomes, and (b) how consistent is the relationship across various regions?

For this question, we employed structural equation modelling which incorporated instructional practices to predict three learning outcomes: science literacy, intrinsic motivation, and epistemic beliefs. We expect learning outcomes to be related positively with forms of instruction which incorporate teacher guidance, and negatively with ones which do not.

Method

Sample and data

We examined nationally representative samples of 15-year-old students from 10 highest and 10 lowest performing regions in PISA 2015 (Table 1). For each region, PISA adopted a two-stage stratified sampling strategy in which randomly sampled schools and then 15-year-old students from each school (OECD, 2016). The total sample is composed of more than 150,000 students from 5089 schools. Participating students completed cognitive tests in science, math, and reading, as well as a background questionnaire which includes experiences of science instruction.

Instructional practice measures

Inquiry-based instruction

We utilised nine items intended to measure inquiry-based instructional practices. The first two items refer explicitly to teacher guidance ('The teacher explains how science ideas can be applied' and 'The teacher clearly explains the relevance of science concepts to our lives'). One item ('Students are given the opportunity to explain their ideas') reflected a student-centred activity but does not refer to inquiry. The remainder (six items) explicitly refer to activities related to different aspects of inquiry (designing and conducting experiments, interpreting data, debating/arguing about science investigations, see Table 3).

Table 1. Regions and sample size.

No	Region	N of schools	N of students
High-performing			
1	B-S-J-G (China)	268	9841
2	Canada	759	20,058
3	Chinese Taipei	214	7708
4	Estonia	206	5587
5	Finland	168	5882
6	Hong Kong	138	5359
7	Japan	198	6647
8	Macao	45	4476
9	Singapore	177	6115
10	Vietnam	188	5826
Low-performing			
11	Algeria	161	5519
12	Brazil	841	23,141
13	Dominican Republic	194	4740
14	Indonesia	236	6513
15	Jordan	250	7267
16	Kosovo	224	4826
17	Lebanon	270	4546
18	Macedonia	106	5324
19	Peru	281	6971
20	Tunisia	165	5375
	TOTAL	5089	151,721

Transmissionist instruction

We also utilised four items measuring 'transmissionist instruction', i.e. a traditional teacher-centred mode in which content is delivered from teacher to students (example item: 'The teacher demonstrates an idea'). Incorporating this variable as a predictor in the models provides a benchmark to assist interpretation about the magnitude of associations between inquiry-based instruction and learning outcomes.

Learning outcomes

PISA's science literacy score was used as the cognitive learning outcome variable in this study. The science literacy test measured students' ability to explain phenomena scientifically, evaluate and design scientific investigations, and interpret data and evidence. The test content is defined not by curriculum content, but rather by contexts and problems for which science concepts can be fruitfully applied. Due to time constraints, each student completed only part of the test and an IRT technique was used to derive 10 plausible values as estimates of students' science literacy.

Epistemic belief and intrinsic motivation (enjoyment of science) were examined as affective outcomes. Epistemic belief refers to personal views about the empirical basis and the evolving nature of scientific knowledge. Enjoyment of science refers to intrinsic motivation or the drive to learn science for the sake of the activity itself (Ryan & Deci, 2000).

Covariates

Several variables which are known to correlate with academic achievement were used as co-variates: gender, immigrant status (whether one is an immigrant), mother tongue (whether is a native speaker of the test language), economic-social-cultural status (ESCS), and science self-efficacy. ESCS in PISA was a composite index which reflected parental education level and occupational status, cultural-educational resources at home, and overall family wealth. Science self-efficacy refers to subjective judgment about one's ability to perform actions and achieve certain goals related to science, e.g. to explain scientific phenomena and interpret data from scientific investigations. This variable was represented by an IRT-scaled score provided by the OECD.

Analysis

Data analyses were performed using Mplus v.8 (Muthén & Muthén, 2017). All models were estimated using the robust maximum likelihood (MLR) estimator, which is robust against deviations from normal distributions and is also suitable for ordinal variables with at least four response categories (Scherer, Nilsen, & Jansen, 2016). Bias introduced by the two-stage stratified sampling was addressed by incorporating the final student weight variable (W_FSTUWT). Missing values were replaced using the full-information likelihood procedure in Mplus. The TYPE = COMPLEX setting in the ANALYSIS option in Mplus was used to correct for standard errors due to the clustered nature of the data (students nested within schools).[1]

Dimensionality and measurement invariance

Two models were examined for their invariance across regions. The first is a model which combines all nine inquiry practice items in a single factor, representing the original PISA design. The second is a model based on Lau and Lam (2017) which separates between an 'interactive application' factor (three items which did not refer to inquiry) and an 'inquiry' factor (six items which explicitly refer to inquiry activities). For each hypothesised structure, multi-group confirmatory factor analyses (MG-CFA) were implemented using the CONFIGURAL-METRIC-SCALAR setting within the MODEL option. The CONFIGURAL model constrained the factor structure but allowing item-factor loadings to vary. In the METRIC model, both factor structure and loadings were constrained to be equal across regions. Finally, in the SCALAR model, item intercepts were also constrained to be equal across regions (Muthén & Muthén, 2017).

The comparative fit index (CFI) and root mean square error approximation (RMSEA) were used to evaluate overall goodness-of-fit. Models were considered to have good fit if CFI was at least .95 and RMSEA not more than .08 (Rutkowski & Svetina, 2014). In the case of non-invariance, exploratory factor analyses (EFA) were performed separately for each region to identify instructional patterns in a bottom-up manner.

Instruction and learning outcomes

To examine how instructional practices were related to learning, we added science literacy, enjoyment of science, and epistemic belief as outcome variables onto the measurement models identified in the previous step. All 10 science literacy plausible values were

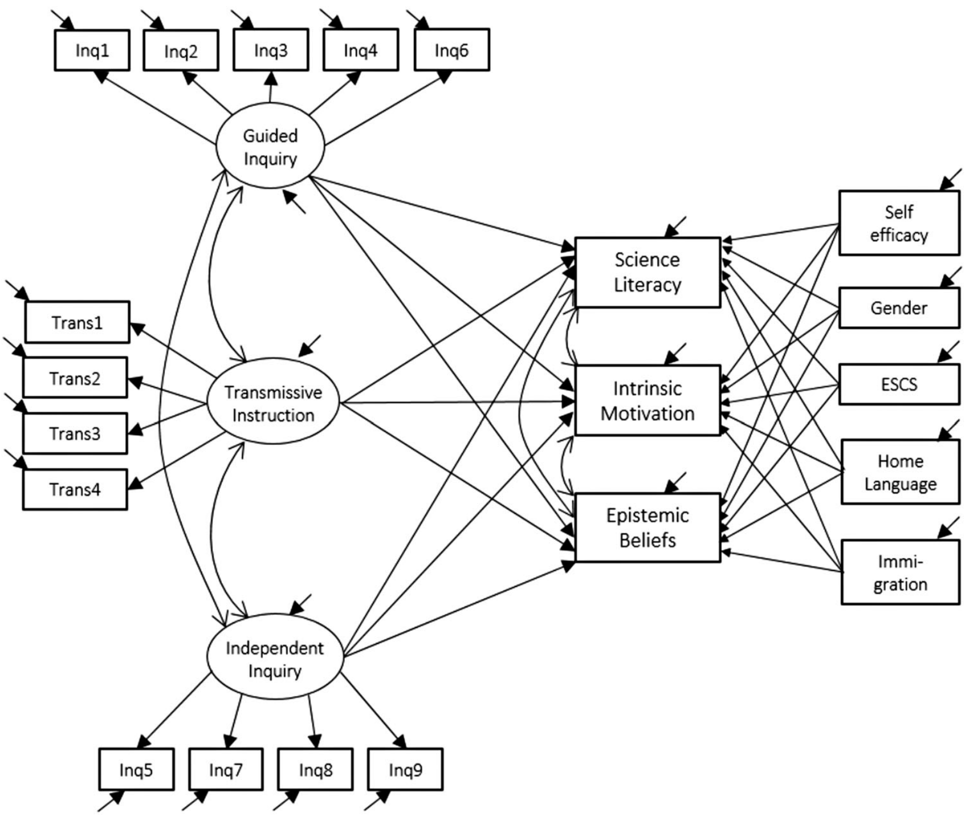

Figure 1. Measurement and structural model for Finland.

incorporated using the TYPE = IMPUTATION setting. In addition, ESCS, gender, immigrant status, language spoken at home, and science self-efficacy were included as co-variates. This amounted to structural equation modelling in which responses to the instructional practice items were included in the models as observed indicators which formed certain instructional practices as latent scores. As an illustration, Figure 1 displays the SEM model for Finland. The item content, see Table 3.

Results

Dimensionality and measurement invariance (RQ 1a)

Fit indices from the MG-CFA indicate that the 2-factor structure (teacher-guided instruction vs. unguided inquiry) outperformed the 1-factor structures (see Table 2). However, the poor fit of this model at the configural level suggests that the basic factor structure is not universal. Rather, there is substantial variation in the pattern of inquiry-based science instruction across the 20 regions. Uncovering these regional patterns of instruction requires an exploratory approach, the results of which are reported next.

Regional forms and patterns of instruction (RQ 1b)

Given the lack of invariance, EFA were performed for each region to generate insights about regional forms and patterns of instruction. EFA results were then used to inform the construction and testing of a measurement model for each region. To test model fit, CFA was used for regions with low cross loadings (<0.2 on non-target factors), while Exploratory SEM (which allows items to cross load onto factors other than its target) was used for other regions. The factor loadings and fit indices are displayed in Tables 3–5.

Note that 'instructional form' refers to the formation of items which load together in a factor.[2] Meanwhile, we use 'instructional pattern' to refer to the combination of instructional forms which characterise the factor structure of a region. We first present findings regarding instructional *forms*, before commenting on instructional patterns, which were observed in the data.

Although no universal measurement model could be established, four instructional forms could be observed across the regions. Each instructional form is characterised by a certain combination of item loadings, and they could be arranged according to their degree of teacher guidance (Figure 2). While the specific items which compose an instructional form may vary across regions, their combination reflects the same conceptual meaning.

Thus, starting from the most student-centred, 'Independent Inquiry' is an instructional form composed only of items referring to inquiry activities, without any of the teacher

Table 2. Measurement invariance test results.

		Configural		*Metric*		*Scalar*	
Invariance level		CFI	RMSEA	CFI	RMSEA	CFI	RMSEA
Number of hypotesised factors	1 factor	0.882	0.090 (0.090–0.091)	0.868	0.084 (0.083–0.085)	0.735	0.108 (0.108–0.109)
	2 factors	0.923	0.074 (0.073–0.075)	0.910	0.072 (0.071–0.072)	0.777	0.103 (0.102–0.104)

Table 3. CFA measurement models for 'Guided vs Independent Inquiry' pattern.

	Finland		Macao		Singapore		Vietnam		Algeria		Kosovo		Lebanon		Tunisia	
Items	Guided Inq.	Indep. Inq.	Guided Inq.	Indep Inq.	Guided Inq.	Indep. Inq.	Guided Inq.	Indep Inq.	Guided Inq.	Indep. Inq.	Guided Inq.	Indep Inq.	Guided Inq.	Indep. Inq.	Guided Inq.	Indep Inq.
inq1. The teacher explains < school science > idea can be applied	0.702		0.565		0.675		0.469		0.462		0.478		0.559		0.552	
inq2. The teacher clearly explains relevance < broad science > concepts to our lives.	0.573		0.514		0.642		0.507		0.505		0.495		0.534		0.604	
inq3. Students are given opportunities to explain their ideas.	0.466		0.502		0.557		0.474		0.433		0.377		0.468		0.544	
inq4. Students are asked to draw conclusions from an experiment they have conducted.	0.696		0.704		0.655		0.593		0.486			0.813	0.587		0.611	
inq5. Students are required to argue about science questions.		0.729	0.754			0.687	0.626		0.576		0.613		0.572		0.691	
inq6. Students spend time in the laboratory doing practical experiments.	0.563		0.691			0.581		0.463	0.566			0.588		0.414		0.597
inq7. Students are allowed to design their own experiments.		0.671		0.717		0.700		0.654		0.698		0.737		0.643		0.680
inq8. Students are asked to do an investigation to test ideas.		0.697		0.758		0.747		0.663	0.578		0.611			0.627		0.704
inq9. There is a class debate about investigations.		0.812		0.802		0.774		0.649		0.549	0.520			0.574		0.698
Fit indices	*CFI = 0.960* *RMSEA = 0.064*		*CFI = 0.962* *RMSEA = 0.058*		*CFI = 0.954* *RMSEA = 0.072*		*CFI = 0.951* *RMSEA = 0.052*		*CFI = 0.956* *RMSEA = 0.044*		*CFI = 0.950* *RMSEA = 0.047*		*CFI = 0.961* *RMSEA = 0.030*		*CFI = 0.952* *RMSEA = 0.062*	

Table 4. ESEM measurement models for 'Guided vs independent inquiry' pattern.

	Canada		B-S-J-G (China)		Hong Kong		Brazil		Dominican		Indonesia		Jordan		Peru		
Items	Guided Inq.	Indep. Inq.	Guided Inq.	Indep Inq.	Guided Inq.	Indep. Inq.	Guided Inq.	Indep Inq.	Guided Inq.	Indep. Inq.	Guided Inq.	Indep Inq.	Guided Inq.	Indep. Inq.	Guided Inq.	Indep. Inq. 1	Indep. Inq. 2
inq1. The teacher explains < school science > idea can be applied	**0.923**	−0.139	**0.855**	−0.097	**0.97**	−0.213	**0.464**	0.275	**0.586**	0.039	**0.322**	0.175	**0.633**	−0.027	**0.75**	0.066	−0.119
inq2. The teacher clearly explains relevance < broad science > concepts to our lives.	**0.616**	0.134	**0.519**	0.238	**0.86**	−0.082	**0.916**	−0.081	**0.748**	−0.155	**0.369**	0.086	**0.784**	−0.206	**0.775**	−0.204	0.143
inq3. Students are given opportunities to explain their ideas.	**0.470**	0.147	**0.581**	0.035	**0.606**	0.032	0.056	**0.631**	**0.513**	0.016	**0.529**	0.035	**0.381**	0.237	**0.496**	0.077	0.000
inq4. Students are asked to draw conclusions from an experiment they have conducted.	**0.318**	**0.371**	**0.350**	**0.493**	**0.579**	0.237	−0.018	**0.770**	**0.385**	**0.471**	**0.315**	**0.351**	0.234	**0.513**	0.169	**0.725**	0.027
inq5. Students are required to argue about science questions.	0.096	**0.658**	0.217	**0.613**	0.169	**0.654**	0.133	**0.669**	**0.519**	0.228	**0.806**	−0.122	0.24	**0.368**	**0.481**	0.277	0.021
inq6. Students spend time in the laboratory doing practical experiments.	0.081	**0.637**	0.112	**0.659**	**0.488**	0.286	−0.215	**0.654**	0.068	**0.526**	0.009	**0.456**	−0.118	**0.867**	−0.004	**0.519**	0.188
inq7. Students are allowed to design their own experiments.	−0.086	**0.793**	−0.002	**0.821**	0.012	**0.747**	−0.012	**0.696**	**0.376**	**0.369**	0.006	**0.624**	**0.361**	**0.373**	−0.071	0.176	**0.628**
inq8. Students are asked to do an investigation to test ideas.	0.091	**0.684**	−0.006	**0.801**	**0.495**	**0.389**	**0.371**	**0.392**	**0.687**	−0.055	0.099	**0.558**	**0.52**	0.209	**0.391**	−0.036	**0.431**
inq9. There is a class debate about investigations.	−0.082	**0.893**	−0.169	**0.917**	0.008	**0.865**	0.148	**0.592**	**0.687**	0.013	−0.046	**0.608**	**0.65**	0.103	0.130	−0.037	**0.704**
Fit indices	*CFI = 0.968 RMSEA = 0.051*		*CFI = 0.979 RMSEA = 0.056*		*CFI = 0.974 RMSEA = 0.061*		*CFI = 0.965 RMSEA = 0.051*		*CFI = 0.964 RMSEA = 0.054*		*CFI = 0.972 RMSEA = 0.038*		*CFI = 0.971 RMSEA = 0.051*		*CFI = 0.983 RMSEA = 0.055*		

Table 5. Measurement models for other instructional patterns.

	Taipei		Estonia		Japan			Macedonia	
Items	Interactive Conceptual Instruction	Indep. Inquiry	Interactive Conceptual Instruction	Indep. Inquiry	Interactive Conceptual Instruction	Indep. Inquiry 1	Indep. Inquiry 2	Teacher-centred Conceptual Instruction	Indep. Inquiry
inq1. The teacher explains < school science > idea can be applied	0.821		0.740		0.793			**0.385**	0.284
inq2. The teacher clearly explains relevance < broad science > concepts to our lives.	0.834		0.732		0.745			**0.764**	−0.008
inq3. Students are given opportunities to explain their ideas.	0.532		0.529		0.488			0.22	**0.306**
inq4. Students are asked to draw conclusions from an experiment they have conducted.		0.838		0.686			0.892	−0.016	**0.695**
inq5. Students are required to argue about science questions.		0.831		0.656		0.723		0.036	**0.62**
inq6. Students spend time in the laboratory doing practical experiments.		0.744		0.581			0.724	−0.216	**0.738**
inq7. Students are allowed to design their own experiments.		0.730		0.682		0.743		−0.049	**0.697**
inq8. Students are asked to do an investigation to test ideas.		0.809		0.647		0.820		0.226	**0.482**
inq9. There is a class debate about investigations.		0.794		0.723		0.832		0.104	**0.566**
Fit indices	*CFI = 0.977 RMSEA = 0.054*		*CFI = 0.957 RMSEA = 0.057*		*CFI = 0.950 RMSEA = 0.057*			*CFI = 0.964 RMSEA = 0.035*	

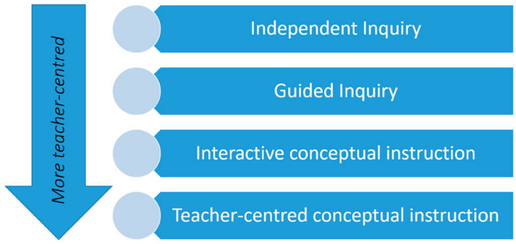

Figure 2. Forms of science instruction.

guidance items. Some variation of Independent Inquiry could be observed in all regions. Note that Independent Inquiry refers to instructional forms which centre on student-led inquiry activities but do not involve explicit conceptual teaching/content exposition from teachers. While it does not involve explicit conceptual teaching, we cannot rule out the possibility that some other forms of guidance are provided within the context of Independent Inquiry. For example, teachers may help structure students' group interactions or provide feedback regarding students' experimental designs.

The second instructional form, 'Guided Inquiry' combines the two teacher guidance items with at least one item referring to an inquiry activity. This form of inquiry was found in 16 of the 20 regions. The last two instructional forms involve some kind of teacher guidance, but without referring to any inquiry activities. Thus, 'Interactive Conceptual Instruction' is characterised by a combination of the two teacher guidance items with the one student-centred non-inquiry item ('Students are given opportunities to explain their ideas'). Interactive Conceptual Instruction was observed in Taipei, Estonia, and Japan. Last and the most teacher-centred, 'Teacher-centred Conceptual Instruction' is characterised simply by the two teacher guidance items. This instructional form was observed only in Macedonia.

Looking at the level of instructional pattern, 16 of the 20 regions contrasted between the two forms of inquiry, i.e. 'Independent Inquiry' and 'Guided Inquiry' (Tables 3 and 4). Meanwhile, the instructional patterns in the remaining four regions combined of 'Independent Inquiry' with either 'Interactive Conceptual Instruction' or 'Teacher-centred Conceptual Instruction'.

Associations with outcomes (RQ 2a and 2b)

Results from structural models predicting learning outcomes are displayed in Tables 6–8 (note on p values: $^{***}p < 0.001$; $^{**}\ p < 0.01$; $^{*}p < 0.05$). We use the instructional pattern in 16 regions to test the hypothesis regarding the importance of teacher guidance for learning from inquiry. Regression results largely supported this hypothesis: guided inquiry was positively associated with science achievement in all 16 regions, with enjoyment in 15 regions, and with epistemic beliefs in 13 regions.

We use all 20 regions to test whether independent inquiry is negatively associated with learning outcomes. Again, the regression results by and large support this hypothesis: independent inquiry was found to be negatively associated with achievement in 18 regions, with enjoyment in 17 regions, and with epistemic beliefs also in 17 regions. The notable exception was Japan, where two forms of independent inquiry could be

Table 6. Standardised estimates (standard errors) from SEM models predicting literacy (cognitive achievement).

Regions/predictors	Guided Inquiry	Independent Inquiry	Transmission. Instruction	Science self-efficacy	Econ-social-cultural status
Finland	0.507*** (0.038)	−0.627*** (0.032)	−0.010 (0.021)	0.225*** (0.014)	0.208*** (0.013)
Macao[C]	0.299*** (0.040)	−0.501*** (0.040)	0.072*** (0.018)	0.202*** (0.017)	0.151*** (0.032)
Vietnam	0.408*** (0.048)	−0.525*** (0.045)	0.103*** (0.028)	0.198*** (0.018)	0.277*** (0.030)
Singapore	0.495*** (0.048)	−0.546*** (0.042)	0.044* (0.021)	0.215*** (0.016)	0.322*** (0.018)
Algeria	0.291*** (0.058)	−0.418*** (0.054)	0.048*** (0.023)	0.016 (0.018)	0.122*** (0.036)
Kosovo	0.159** (0.060)	−0.366*** (0.055)	0.190*** (0.019)	0.052** (0.018)	0.207*** (0.022)
Lebanon	0.301*** (0.061)	−0.543*** (0.058)	0.126*** (0.033)	0.166*** (0.027)	0.298*** (0.033)
Tunisia	0.886*** (0.129)	−1.111*** (0.120)	0.034 (0.032)	0.081*** (0.017)	0.300*** (0.030)
Canada	0.321*** (0.019)	−0.482*** (0.015)	0.035* (0.015)	0.238*** (0.010)	0.223*** (0.010)
Beijing	0.595*** (0.030)	−0.556*** (0.023)	0.040* (0.019)	0.104*** (0.013)	0.381*** (0.020)
Hong Kong	0.326*** (0.032)	−0.492*** (0.031)	0.109*** (0.022)	0.149*** (0.016)	0.183*** (0.019)
Brazil	0.222*** (0.035)	−0.417*** (0.028)	0.134*** (0.015)	0.100*** (0.013)	0.315*** (0.018)
Dominica	0.149** (0.051)	−0.512*** (0.051)	0.081** (0.028)	0.001 (0.018)	0.350*** (0.025)
Indonesia	0.275*** (0.042)	−0.392*** (0.040)	0.006 (0.020)	0.021 (0.015)	0.421*** (0.026)
Jordan	0.150** (0.045)	−0.364*** (0.045)	0.159*** (0.019)	0.132*** (0.015)	0.295*** (0.017)
Peru	0.347*** (0.043)	−0.026 (0.039) and −0.569*** (0.035)	0.086*** (0.020)	0.045** (0.015)	0.382*** (0.020)
Regions/predictors	Interactive conceptual inst.	Independent inquiry	Transmission. Instruction	Science self-efficacy	Econ.-social-cultural status
Estonia	0.561*** (0.045)	−0.754*** (0.041)	−0.025 (0.019)	0.154*** (0.018)	0.219*** (0.018)
Taipei	0.485*** (0.025)	−0.541*** (0.027)	0.045** (0.013)	0.189*** (0.013)	0.261*** (0.016)
Macedonia	0.367*** (0.082)	−0.446*** (0.073)	0.073*** (0.021)	0.153*** (0.024)	0.232*** (0.026)
Japan	0.130*** (0.031)	−0.483*** (0.032) and 0.254*** (0.038)	0.082*** (0.021)	0.182*** (0.012)	0.244*** (0.015)

Table 7. Standardised estimates (standard errors) from SEM models predicting enjoyment.

Regions/predictors	Guided Inquiry	Independent Inquiry	Transmission. Instruction	Science self-efficacy	Econ-social-cultural status
Finland	0.252*** (0.034)	−0.125*** (0.032)	0.157*** (0.022)	0.303*** (0.016)	0.076*** (0.014)
Macao	0.137*** (0.038)	−0.088** (0.029)	0.121*** (0.021)	0.318*** (0.019)	0.027 (0.019)
Vietnam	0.028 (0.042)	−0.060 (0.038)	0.193*** (0.031)	0.203*** (0.019)	0.022 (0.020)
Singapore	0.387*** (0.051)	−0.245*** (0.039)	0.094*** (0.023)	0.319*** (0.019)	0.051*** (0.014)
Algeria	0.286*** (0.045)	−0.234*** (0.048)	0.214*** (0.022)	0.077*** (0.020)	−0.026 (0.015)
Kosovo	0.169*** (0.040)	−0.081** (0.037)	0.173*** (0.019)	0.042 (0.023)	0.018 (0.018)
Lebanon	0.213*** (0.053)	−0.189*** (0.051)	0.278*** (0.031)	0.174*** (0.027)	0.075** (0.024)
Tunisia	0.563*** (0.108)	−0.498*** (0.102)	0.167*** (0.030)	0.107*** (0.023)	0.019 (0.015)
Canada	0.276*** (0.022)	−0.161*** (0.017)	0.115*** (0.016)	0.326*** (0.013)	0.066*** (0.010)
Beijing	0.204*** (0.029)	−0.026 (0.026)	0.122*** (0.018)	0.252*** (0.018)	0.086*** (0.015)
Hong Kong	0.241*** (0.032)	−0.144*** (0.028)	0.179*** (0.030)	0.342*** (0.020)	0.026 (0.015)
Brazil	0.227*** (0.025)	−0.061* (0.024)	0.158*** (0.016)	0.188*** (0.016)	0.074*** (0.013)
Dominica	0.076* (0.032)	−0.065 (0.039)	0.233*** (0.024)	0.101*** (0.026)	−0.029 (0.019)
Indonesia	0.203*** (0.038)	−0.104** (0.035)	0.148*** (0.023)	0.110*** (0.016)	−0.019 (0.019)
Jordan	0.204*** (0.036)	−0.210*** (0.034)	0.288*** (0.022)	0.208*** (0.021)	0.006 (0.017)
Peru	0.150*** (0.036)	−0.080** (0.031) and −0.062* (0.029)	0.162*** (0.023)	0.178*** (0.018)	−0.017 (0.016)
Regions/predictors	Interactive conceptual inst.	Independent inquiry	Transmission. Instruction	Science self-efficacy	Econ.-social-cultural status
Estonia	0.343*** (0.037)	−0.280*** (0.039)	0.118*** (0.018)	0.202*** (0.022)	0.084*** (0.017)
Taipei	0.135*** (0.021)	−0.006 (0.019)	0.083*** (0.013)	0.330*** (0.012)	0.075*** (0.012)
Macedonia	0.292*** (0.055)	−0.217*** (0.055)	0.204*** (0.018)	0.098*** (0.027)	−0.017 (0.021)
Japan	0.244*** (0.027)	−0.149*** (0.024) and 0.014 (0.025)	0.102*** (0.020)	0.314*** (0.015)	0.071*** (0.014)

Table 8. Standardised estimates (standard errors) from SEM models predicting epistemic beliefs.

Regions/predictors	Guided Inquiry	Independent Inquiry	Transmission. Instruction	Science self-efficacy	Econ-social-cultural status
Finland	0.312*** (0.042)	−0.317*** (0.039)	0.050* (0.025)	0.176*** (0.020)	0.133*** (0.014)
Macao	0.126*** (0.036)	−0.129*** (0.029)	0.144*** (0.022)	0.181*** (0.018)	0.050** (0.019)
Vietnam	0.040 (0.043)	−0.151*** (0.040)	0.231*** (0.026)	0.113*** (0.022)	0.139*** (0.018)
Singapore	0.410*** (0.049)	−0.347*** (0.041)	0.064** (0.023)	0.202*** (0.020)	0.073*** (0.015)
Algeria	0.086 (0.044)	−0.108* (0.042)	0.165*** (0.023)	0.022 (0.028)	0.031 (0.016)
Kosovo	0.193*** (0.039)	−0.170*** (0.036)	0.137*** (0.022)	−0.041 (0.026)	0.076*** (0.020)
Lebanon	0.337*** (0.060)	−0.375*** (0.059)	0.154*** (0.033)	0.116*** (0.034)	0.098*** (0.031)
Tunisia	0.320*** (0.095)	−0.343*** (0.089)	0.134*** (0.029)	0.054* (0.023)	0.078*** (0.019)
Canada	0.230*** (0.024)	−0.231*** (0.018)	0.097*** (0.018)	0.175*** (0.013)	0.105*** (0.010)
Beijing	0.220*** (0.029)	−0.146*** (0.025)	0.074*** (0.018)	0.191*** (0.022)	0.141*** (0.015)
Hong Kong	0.277*** (0.035)	−0.257*** (0.033)	0.152*** (0.025)	0.193*** (0.025)	0.074*** (0.017)
Brazil	0.160*** (0.026)	−0.162*** (0.025)	0.169*** (0.018)	0.097*** (0.016)	0.101*** (0.014)
Dominica	0.066* (0.033)	−0.164*** (0.038)	0.150*** (0.030)	0.015 (0.021)	0.045* (0.020)
Indonesia	0.031 (0.034)	−0.007 (0.034)	0.102*** (0.023)	0.029 (0.021)	0.103*** (0.017)
Jordan	0.200*** (0.036)	−0.255*** (0.032)	0.223*** (0.023)	0.125*** (0.021)	0.088*** (0.016)
Peru	0.136*** (0.038)	−0.015 (0.028) and −0.184*** (0.032)	0.134*** (0.021)	0.057** (0.019)	0.135*** (0.014)
Regions/predictors	Interactive conceptual inst.	Independent inquiry	Transmission. Instruction	Science self-efficacy	Econ.-social-cultural status
Estonia	0.393*** (0.037)	−0.423*** (0.039)	0.074** (0.021)	0.080*** (0.023)	0.094*** (0.017)
Taipei	0.287*** (0.023)	−0.246*** (0.022)	0.078*** (0.015)	0.186*** (0.016)	0.118*** (0.013)
Macedonia	0.241*** (0.063)	−0.237*** (0.058)	0.192*** (0.020)	0.040 (0.025)	0.114*** (0.018)
Japan	0.182*** (0.030)	−0.219*** (0.027) and 0.051* (0.025)	0.144*** (0.021)	0.214*** (0.019)	0.129*** (0.015)

observed, one of which was positively associated with achievement and epistemic beliefs.

Additional analysis[3] found that the positive associations between Guided Inquiry and achievement (science literacy) were weaker in regions were Guided Inquiry was measured using more inquiry-specific items. In other words, Guided Inquiry which includes more student-centred inquiry activities seems to be less strongly associated with achievement, compared to Guided Inquiry which includes fewer inquiry activities. To illustrate, in the Dominican Republic, where Guided Inquiry included five inquiry items, the effect size (predicting literacy) was 0.149. Meanwhile, the corresponding effect sizes were 0.291 in Algeria (where Guided Inquiry included 4 inquiry-specific items), 0.299 in Macao (with 3 inquiry-specific items), 0.408 in Vietnam (with 2 inquiry-specific items), and 0.495 in Singapore (with 1 inquiry-specific item). Further analysis at the item level showed that all inquiry items were associated with higher science literacy in all countries except Vietnam (where the correlations were positive but not significant).

Discussion

The current study takes a more nuanced look at the measurement of inquiry-based instruction and how it relates to learning outcomes in the highest and lowest performing regions of PISA 2015. Prior studies have assumed, without explicitly testing, that the structure of inquiry-based instruction is equivalent across regions. Our examination found little support for this assumption of measurement invariance. That is, the nine items designed to assess inquiry-based instruction in PISA do not form the same universal structure. Rather, the analysis revealed regional patterns of instruction. Nonetheless, at a higher level of abstraction, our analysis also suggests that in many regions, a distinction between 'Guided Inquiry' and 'Independent Inquiry' can be found. Both forms of instruction involve the use of inquiry activities. The difference between them is that Guided Inquiry combines inquiry activities with teachers' explanations about how science concepts can be applied.

The contrast between Guided and Independent Inquiry allowed us to test the conjecture that, when coupled with teacher guidance, inquiry can be associated with better learning outcomes. Results of the structural equation modelling provide strong support for this conjecture for all types of outcomes examined. Guided Inquiry was positively associated with scores on science achievement test in all the 16 regions where this form of instruction was observed. Positive associations between Guided Inquiry and affective outcomes (intrinsic motivation and epistemic beliefs) could also be found in the majority of the 16 regions.

Conversely, Independent Inquiry was found to be negatively associated with learning outcomes in 19 of the 20 regions where this form of instruction could be observed. As previously noted, the exception was Japan, where one form of Independent Inquiry was positively associated with science achievement and epistemic belief. This suggests that, in some contexts, inquiry can be effective even without teacher guidance.

This study brings PISA-based findings in line with mainstream theories of learning and empirical findings in science education (Furtak et al., 2016; Lazonder & Harmsen, 2016). As elaborated in the Introduction, cognitive and socio-constructivist theories emphasise the importance of scaffolding for learning from inquiry (Hmelo-silver et al., 2007).

Accordingly, this study found that inquiry activities were associated positively with learning outcomes when they are coupled with teachers' conceptual guidance. Further lending support for this conclusion is the observation that the positive association between Guided Inquiry and achievement seemed to be weaker in regions where Guided Inquiry was more student-centred (measured using more items which reflect student-centred inquiry activities).

More importantly, our findings not only confirm theoretical predictions. They also serve as evidence that teachers teaching in the 'average' or typical school are able to provide guidance which makes inquiry meaningful and effective. Moreover, this was true not only for high-performing education systems, but also for low-performing ones. In other words, Guided Inquiry seems to be effective (and more so than traditional instruction) even when implemented in schools with more limited resources and in education systems where teacher and teaching quality are generally poor (Aslam et al., 2016; Scheerens, 2001).

A critical reader might question whether the statistically significant associations found in this study are also practically meaningful. This question needs to be addressed especially because this study utilised large sample sizes which increase the possibility of Type I error ('false positives'). The meaning of effect sizes is difficult to judge in absolute terms. One way to judge the practical significance of this study's findings is by comparing it to effect sizes found in prior studies. The associations between inquiry-based instruction and learning outcomes found in most regions in this study are comparable to the effects sizes summarised in meta-analyses of educational effectiveness studies (Kyriakides, Christoforou, & Charalambous, 2013; Scheerens, Luyten, Steen, & Luyten-De Thouars, 2007). In this metric, the effects observed in the present study can be considered as moderate.

Another way to gauge magnitude of the observed effects is through comparisons with the effect sizes other predictors of learning outcomes. Using this metric, effect sizes of Guided Inquiry on achievement are larger and more consistently positive than the effect sizes of Transmissionist Instruction, especially when looking at cognitive outcome. In addition, the effect sizes of Guided Inquiry on achievement are also often larger than, or at least comparable to, the effects of science self-efficacy as well as family economic-socio-cultural background. Thus, we argue that the magnitude of associations observed in the current study can be considered meaningful.

These findings are significant because ILSA of learning such as PISA can exert significant influence on educational policy (Berliner, 2015; Grek, 2009). While its cross-sectional design prevents causal inferences to be made, findings from ILSA are perceived to have strong external validity because they are based on nationally representative samples of schools and students. With regard to science teaching, analysis based on ILSA data often shows inquiry to be negatively associated with science achievement (Cairns & Areepattamannil, 2017; Chi et al., 2018). This finding can lead to the suggestion that teachers who teach in the 'typical' school may not have the capacity or support required to implement inquiry effectively. Thus, some have voiced concerns that the desire to climb the 'PISA ladder' may prompt policy makers to discourage the use of inquiry-based instruction (Sjøberg, 2018). The current study, however, suggests that prior negative associations between inquiry-based instruction and learning were likely due to the conflation between guided and unguided forms of inquiry. In accordance with the mainstream view in science education, this study finds that when combined with teacher guidance, inquiry is positively associated with cognitive and affective learning outcomes. Furthermore, in almost all regions the positive effects of guided inquiry on learning were found to be larger than transmissionist instruction. Thus, it would be misguided to use PISA findings to support arguments which favour explicit/direct forms of instruction over constructivist approaches such as inquiry.

Another point worth discussing is that the instructional forms observed in this study may not reflect the kind of authentic inquiry advocated by science educators (Chinn & Malhotra, 2002). In none of the regions did all nine items intended to measure inquiry-based instruction form a single dimension. In other words, according to the students in our sample, science teachers tend to employ only limited sets of inquiry activities. In a sense, this finding is unsurprising. Interweaving the empirical, epistemic, social, and conceptual dimensions to enact authentic inquiry is no easy feat (Harris & Rooks, 2010). Nonetheless, from a practice point of view, the findings here suggest that it doesn't really matter which aspect of inquiry are implemented. It matters little whether a teacher asks students to design and conduct lab-based experiments, or another provides empirical data for students to discuss and debate. Rather, what matters for student learning is whether teachers are actively involved to help students make sense of and conceptualise their inquiry activity.

On the issue of measurement, another limitation of this study stems from the fact that PISA's inquiry scale was not designed to measure guidance in the context of inquiry. More specifically, while the scale included items reflecting teacher conceptual guidance, it does not measure other forms of guidance or structure which may be incorporated to scaffold students' inquiry. Consequently, the forms of 'Independent Inquiry' identified in this study need to be interpreted with caution, as they may include forms of guidance which are not measured by the scale. This points to the need to design instruments which explicitly assess the use of scaffolding and guidance in the context of inquiry-based instruction.

Last, we note a general limitation of findings based on ILSA which stems from the cross-sectional nature of the data. In examining the relations between instruction and learning outcomes, it is difficult to ascertain the direction of causality. It is possible that teachers tend to refrain from providing conceptual guidance to students who – at the outset of instruction – exhibit low interest, efficacy, motivation, and/or achievement. In other words, students' initial motivation and achievement maybe the driving force for which type of instruction teachers employ. In this study, we have attempted to address this problem by including important determinants of achievement as co-variates, including students' science self-efficacy and socio-economic background. While this may partially mitigate the issue, we recognise that no strong causal inference can be made with regard to inquiry-based instruction and learning outcomes. Future studies should strive to include measures of prior achievement, ideally within a longitudinal design, to address this issue.

Conclusions and implications

In closing, we conclude that suggestions to discourage science teachers from utilising inquiry activities seem to be misguided. When examined in more detail, ILSA data yield findings consistent with mainstream theory and experimental studies. In this study, we show that student-centred inquiry activities, in and of themselves, are not the culprit of low motivation or achievement in science classes. On the contrary, inquiry activities tend to be associated with higher outcomes when coupled with teachers' active involvement to help students make sense or conceptualise their experiences. Significantly, this study suggests that the 'average' teacher teaching in the 'average' school is capable of providing the type of conceptual guidance needed to facilitate productive science learning.

For future research, one implication arising from our findings is that researchers should pay careful attention to issues of measurement invariance when examining instructional practices in ILSA data. Another implication is that future ILSA would benefit from developing measures specifically designed to assess the quality of teacher guidance in inquiry. With regard to practical implications, teacher-guided inquiry activities seem to be an essential component of instruction when the goal is to develop students' scientific literacy. Even lower performing education systems would benefit from encouraging teachers to couple one or another inquiry activity with conceptual explanations. Also, to the extent that some teachers view student-centred teaching as equating to letting students on their own course without providing structure and guidance, policy documents and teacher training should counter such misconceptions.

Notes

1. The research questions in this study deal with relationships of variables at a single (student) level, and thus preclude the need for multilevel modelling (Stapleton, McNeish, & Yang, 2016). The TYPE=COMPLEX approach was preferred because it handles the clustered/non-independent observations while requiring substantially less computational time compared to multilevel modelling (Muthen & Satorra, 1995).
2. As further evidence of the validity of the distinction between Guided and Independent Inquiry, we examined the correlations between both inquiry forms and three teaching variables: Transmissionist Instruction, Adaptive Instruction, and Emotional Support. Both forms of inquiry were positively correlated with the other teaching variables in the vast majority of the regions. More importantly, correlations were stronger with Guided Inquiry compared to Independent Inquiry in all regions. We argue that this is strong evidence supporting our interpretation of the conceptual distinction between the two forms of inquiry. On average, Guided Inquiry correlated with Transmissionist Instruction at 0.38 (range −0.02 to 0.59), with Adaptive Instruction at 0.46 (range 0.32 to 0.60), and with Emotion Support at 0.59 (range 0.47 to 0.73). Meanwhile, on average Independent Inquiry correlated with Transmissionist Instruction at 0.21 (range −0.16 to 0.49), with Adaptive Instruction at 0.27 (range 0.07 to 0.41), and with Emotional Support at 0.34 (range 0.17 to 0.50). See the Online Appendix for complete results.
3. We thank an anonymous reviewer for pointing this out by analysing the correlation between number of inquiry-specific items and Guided Inquiry effect size (across regions).

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the Alexander von Humboldt-Stiftung.

ORCID

Anindito Aditomo http://orcid.org/0000-0003-3711-3773
Eckhard Klieme http://orcid.org/0000-0003-0728-4950

References

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. doi:10.1080/00220671.2010.533717

Aslam, M., Rawal, S., Kingdon, G., Moon, B., Banerji, R., Das, S., … Sharma, S. K. (2016). *Reforms to increase teacher effectiveness in developing countries*. London: EPPI-Centre, Social Science Research Unit, UCL Institute of Education, University College London.

Barron, B., & Darling-Hammond, L. (2008). *Teaching for meaningful learning: A review of research on inquiry-based and cooperative learning*. San Rafael, CA: George Lucas Educational Foundation.

Berliner, D. C. (2015). The many facets of PISA. *Teachers College Record*, *117*(1), 1–20.

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. doi:10.1007/s11165-017-9639-x

Chi, S., Liu, X., Wang, Z., & Won Han, S. (2018). Moderation of the effects of scientific inquiry activities on low SES students' PISA 2015 science achievement by school teacher support and disciplinary climate in science classroom across gender. *International Journal of Science Education*, 1–21. doi:10.1080/09500693.2018.1476742

Chinn, C. A., & Malhotra, B. A. (2002). Epistemologically authentic inquiry in schools: A theoretical framework for evaluating inquiry tasks. *Science Education*, *86*, 175–218. doi:10.1002/sce.10001

Duschl, R. A. (2007). Quality argumentation and epistemic criteria. In S. Erduran & M. P. Jimenez-Aleixandre (Eds.), *Argumentation in science education: Perspectives from classroom-based research* (pp. 159–175). Dordrecht: Springer.

Duschl, R. A. (2008). Science education in three-part harmony: Balancing conceptual, epistemic, and social learning goals. *Review of Research in Education*, *32*, 268–291. doi:10.3102/0091732X07309371

Engeln, K., Mikelskis-seifert, S., & Euler, M. (2014). Inquiry-based mathematics and science education across Europe: A synopsis of various approaches and their potentials. In C. Bruguière (Ed.), *Topics and Trends in current science education: 9th ESERA conference selected contributions* (pp. 229–242). Dordrecht: Springer Science. doi:10.1007/978-94-007-7281-6

Fitzgerald, M., Danaia, L., & McKinnon, D. H. (2017). Barriers inhibiting inquiry-based science teaching and potential solutions: Perceptions of positively inclined early adopters. *Research in Science Education*, 1–24. doi:10.1007/s11165-017-9623-5

Fogleman, J., McNeill, K. L., & Krajcik, J. (2011). Examining the effect of teachers' adaptations of a middle school science inquiry-oriented curriculum unit on student learning. *Journal of Research in Science Teaching*, *48*(2), 149–169. doi:10.1002/tea.20399

Forman, E. A. (2018). The practice turn in learning theory and science education. In David W. Kritt (Ed.), *Constructivist education in an age of accountability* (pp. 97–104). Cham, Switzerland: Palgrave Macmillan. doi:10.1007/978-3-319-66050-9_5

Furtak, E. M., Seidel, T., Iverson, H., & Briggs, D. C. (2016). Experimental and quasi-experimental studies of inquiry-based science teaching: A meta-analysis. *Review of Educational Research*, *82*(3), 300–329. doi:10.3102/0034654312457206

Grabau, L. J., & Ma, X. (2017). Science engagement and science achievement in the context of science instruction: A multilevel analysis of U.S. students and schools. *International Journal of Science Education*, *39*(8), 1045–1068. doi:10.1080/09500693.2017.1313468

Grek, S. (2009). Governing by numbers: The PISA "effect" in Europe. *Journal of Education Policy*, *24*(1), 23–37. doi:10.1080/02680930802412669

Harris, C. J., & Rooks, D. L. (2010). Managing inquiry-based science: Challenges in enacting complex science instruction in elementary and middle school classrooms. *Journal of Science Teacher Education*, *21*(2), 227–240. doi:10.1007/s10972-009-9172-5

Hmelo-silver, C. E., Duncan, R. G., & Chinn, C. A. (2007). Scaffolding and achievement in problem-based and inquiry learning: A response to Kirschner, Sweller, and Clark (2006). *Educational Psychologist*, *42*(2), 99–107.

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. doi:10.1207/s15326985ep4102

Kyriakides, L., Christoforou, C., & Charalambous, C. Y. (2013). What matters for student learning outcomes: A meta-analysis of studies exploring factors of effective teaching. *Teaching and Teacher Education*, *36*, 143–152. doi:10.1016/j.tate.2013.07.010

Lau, K., & Lam, T. Y. (2017). Instructional practices and science performance of 10 top-performing regions in PISA 2015. *International Journal of Science Education*, 1–22. doi: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. doi:10.3102/0034654315627366

Liou, P.-Y., & Jessie Ho, H.-N. (2018). Relationships among instructional practices, students' motivational beliefs and science achievement in Taiwan using hierarchical linear modelling. *Research Papers in Education*, *33*(1), 73. doi:10.1080/02671522.2016.1236832

Müller, K., Prenzel, M., Seidel, T., Schiepe-tiska, A., & Kjærnsli, M. (2016). Science teaching and learning in schools: Theoretical and empirical foundations for investigating classroom-level processes. In S. Kuger, E. Klieme, N. Jude, & D. Kaplan (Eds.), *Assessing contexts of learning: An international Perspective* (pp. 423–446). Cham, Switzerland: Springer International. doi:10.1007/978-3-319-45357-6

Muthén, L. K., & Muthén, B. O. (2017). *Mplus user's guide* (8th ed.). Los Angeles, CA: Muthén & Muthén.

Muthen, B. O., & Satorra, A. (1995). Complex sample data in structural equation modeling. *Sociological Methodology*, *25*, 267–316. doi:10.1016/j.acthis.2013.11.004

National Research Council. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. Washington, DC: The National Academies.

OECD. (2016). *PISA 2015 results (Volume I): Excellence and equity in education (Vol. I)*. Paris: Author.

Osborne, J., Simon, S., Christodoulou, A., & Howell-richardson, C. (2013). Learning to argue: A study of four schools and their attempt to develop the use of argumentation as a common instructional practice and its impact on students. *Journal of Research in Science Teaching*, *50*(3), 315–347. doi:10.1002/tea.21073

Pluta, W. J., Chinn, C. A., & Duncan, R. G. (2011). Learners' epistemic criteria for good scientific models. *Journal of Research in Science Teaching*, *48*(5), 486–511.

Rutkowski, L., & Svetina, D. (2014). Assessing the hypothesis of measurement invariance in the context of large-scale international surveys. *Educational and Psychological Measurement*, *74*(1), 31–57. doi:10.1177/0013164413498257

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1), 68–78. doi:10.1037/0003-066X.5501.68

Sandoval, W. A., Greene, J. A., & Bråten, I. (2016). Understanding and promoting thinking about knowledge origins, issues, and future directions of research on epistemic cognition. *Review of Research in Education*, *40*(1), 457–496.

Scheerens, J. (2001). Monitoring school effectiveness in developing countries. *School Effectiveness and School Improvement*, *12*(4), 359–384. doi:10.1076/sesi.12.4.359.3447

Scheerens, J., Luyten, H., Steen, R., & Luyten-De Thouars, Y. (2007). *Review and meta-analyses of school and teaching effectiveness*. Retrieved from https://www.utwente.nl/nl/bms/omd/Medewerkers/artikelen/scheerens/rev-meta-analysis2007.PDF

Scherer, R., Nilsen, T., & Jansen, M. (2016). Evaluating individual students' perceptions of instructional quality: An investigation of their factor structure, measurement invariance, and relations to educational outcomes. *Frontiers in Psychology*, *7*(February), 1–16. doi:10.3389/fpsyg.2016.00110

Schmidt, H. G., Loyens, S. M. M., & Paas, F. (2007). Problem-based learning is compatible with human cognitive architecture: Commentary on Kirschner, Sweller, and Clark (2006). *Educational Psychologist*, *42*(2), 91–97.

Sfard, A. (1998). On two metaphors for learning and the dangers of choosing just one. *Educational Researcher*, *27*(2), 4–13.

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. doi:10.5617/nordina.6185

Stapleton, L. M., McNeish, D. M., & Yang, J. S. (2016). Multilevel and single-level models for measured and latent variables when data are clustered. *Educational Psychologist*, *51*(3–4), 317–330. doi:10.1080/00461520.2016.1207178

Windschitl, M., Thompson, J., & Braaten, M. (2008). Beyond the scientific method: Model-based inquiry as a new paradigm of preference for school science investigations. *Science Education*, 1–27. doi:10.1002/sce.20259

Wu, A. D., Li, Z., & Zumbo, B. D. (2007). Decoding the meaning of factorial invariance and updating the practice of multi-group confirmatory factor analysis: A demonstration with TIMSS data. *Practical Assessment, Research & Evaluation*, *12*(3), 1–26.

Zhang, L. (2016). Is inquiry-based science teaching worth the effort? *Science & Education*, *25*(7–8), 897–915. doi:10.1007/s11191-016-9856-0

International Journal of Science Education

Browse this journal

- Latest articles
- Current issue
- List of issues
- Special issues
- Open access articles
- Most read articles
- Most cited articles

International Journal of Science Education

ISSN: 0950-0693 (Print) 1464-5289 (Online) Journal homepage: https://www.tandfonline.com/loi/tsed20

Celebrating the life of John Kenward Gilbert

Rosária Justi

To cite this article: Rosária Justi (2020) Celebrating the life of John Kenward Gilbert, International Journal of Science Education, 42:4, 493-503, DOI: 10.1080/09500693.2020.1740819

To link to this article: https://doi.org/10.1080/09500693.2020.1740819

INTERNATIONAL JOURNAL OF SCIENCE EDUCATION
2020, VOL. 42, NO. 4, 493–503
https://doi.org/10.1080/09500693.2020.1740819

EDITORIAL

Celebrating the life of John Kenward Gilbert

Introductory notes

When John was invited to organise his selected works in a book to the Routledge series World Library of Educationalists, he wrote in the Preface that the invitation to compose 'what amounts to a "professional autobiography of ideas"' was 'a most disconcerting experience'. He thought that 'to address the task efficiently, one would need to be both many years away from all relevant facts and to be emotionally detached from them' (Gilbert, 2005a, p. 1).

Today, I feel I am in the very same situation. I hope this text may make those who shared moments with him remember them smiling, and those who have not had such an opportunity to wish they could have met him.

Personal and professional births and initial developments

John was born on 27th April 1940, in London, just some months before the city started being bombed. More than once he told me that, due to having been a Londoner child during the II World War, he had been invited to participate in some psychological studies about whether and how the war affected children's emotions. He was proud to have always been a point outside the data curves that showed kids with psychological traumas, to have found a way to become a happy single child at those difficult times.

In the Preface above, John recalled regularly visiting the Science Museum and the Dome of Discovery as a child. The artefacts he saw in both of them, as well as the work of a physics school teacher who established clear relationships between phenomena and the abstract ideas that explained them, increased his interest in both science and 'teaching to promote thinking' (Gilbert, 2005a, p. 2) – two issues that guided his professional career.

Some years later, in 1962, he got a B.Sc. (Hons.) Chemistry from the University of Leicester, which was followed by a D.Phil. Chemistry from the University of Sussex in 1965. But rather than get a university chemistry post, he became an Assistant Master for chemistry at the King's School, in Rochester, where he taught for three years.

In 1968, John was granted a Postgraduate Certificate in Education (Science) from the University of London, and, in the same year, he became the Head of Chemistry at a large comprehensive school: the Banbury School (now the Wykham Park Academy), in Oxfordshire. There, he 'started to think beyond the confines of the immediate classroom and laboratory' (Gilbert, 2005a, pp. 2–3) and had the opportunity to teach from the Nuffield 'A' Level Chemistry. Due to his experience, John was asked to give a lecture on the Nuffield 'A' Level schemes to members of the Association for Science Education. He acknowledged this was one of the first times he was convinced that, if approached from a coherent intellectual basis, science education could be an exciting and fun experience for both students and teachers. It seems that the years John spent as a chemistry teacher became a seedbed for the development of future ideas on children's ideas.

First steps as a prominent researcher

From such enthusiastic and embryonic ideas in mind, John started his university career in 1972, at the University of Keele, as a Lecturer in Science Education (a period of time when he was also a Professional Tutor at Shrewsbury School, in Shropshire). Soon he moved to the University of Surrey, in Guildford, when he worked as Senior Lecturer and Reader in Science Education from 1974 to 1988. More than once, he told me that Surrey was the university where he most enjoyed working because there was a pleasant and favourable environment to think, learn, and exchange ideas.

At the Institute of Education Technology of the University of Surrey, John participated in the creation of an innovative course which combined physics (or chemistry) with education and granted both a degree of the University and a teaching certificate. The course was created in conjunction with Bulmershe College of Higher Education, in Reading, and was motivated by worries about 'the future of science in schools and universities' (Gilbert, 1975, p. 516) – something that attracted John's attention in distinct stages of his academic career, leading to distinct initiatives.

It was at Surrey that John started his ample circle of international collaborations – something that he highly valued. John's first international collaborator, Roger Osborne, who he recognised as 'one of the leading lights of his generation' (Gilbert, 2009a, p. 322), was one of the main ones in this entire career, as well as one of his best friends. When replying the questions that resulted in his contribution to Peter Fensham's (2004) book *Defining an Identity: The Evolution of Science Education as a Field of Research*, John revealed that, for years after Roger returned to Hamilton after a short stay in Guilford in 1979–1980, they were in touch almost every week to discuss ideas. Since the mid-1970s, the importance of identifying and considering students' conceptions in science education had been the topic of studies conducted by A. Champagne, J. Clement, R. Driver, R. Gunstone, J. Novak, J. Nussbaum, R. White (among many others). In a not very well-known paper (Gilbert, 1977), John informed that at the University of Surrey they had been interested in such matters for some time. As a result, in 1976, they started a tentative work "done jointly with Visiting Staff", with "a number of aims: to investigate the usefulness of different question types for diagnosis, to look for patterns of performance within higher education and for trends between school and higher education, to explore types of result analysis and presentation likely to be of greatest use to practising teachers, and to seek a formula for future diagnostic procedures". (p. 166)

It is likely that as a consequence of this project, or as part of it, the fruitful collaboration with Roger Osborne had started. Together, they made a significant methodological contribution to the area by producing an original, simple, and powerful tool for collecting data on such conceptions: the Interview about Instances (Osborne & Gilbert, 1980). Assuming the importance of students' ideas, as well as limitations of instruments being used to collect data at that time (including the self-evolving questionnaire, that had been produced some years earlier (Bridge & Gilbert, 1977)), they concluded that:

"We could just simple ask them, but what could we use as stimulus (…) We came to stick figures, on the basis that they were less contextualised, (…) Then we sat down and thought of a variety of situations in which the concept of interest could or could not have application, because I remembered Dudley Herron had used instances and non-instances of concepts in one of his chemistry studies." Two single questions were asked about each figure: "Is this an example of C (the concept)? Why?" (Fensham, 2004, p. 124)

Due to the simplicity to be produced, replicated, and administered, as well as to the results obtained, hundreds of studies were conducted using the Interview about Instance on a series of scientific concepts (Fensham, 2004). In one of them, John and colleagues identified a set of challenging concerns:

is it invariably desirable or feasible to change all students towards the consensus scientific viewpoint?; what implications are there for class composition and syllabus construction?; what changes in examination techniques do they imply?; how do patterns of student conception relate to the historical development of a subject? The future of this field of interlocked research and development seems likely to be a busy one. (Gilbert et al., 1982, p. 66)

The future showed they were absolutely right. The research programme on students' conceptions of science concepts was certainly one of the most successful in science education due not only to the amount of empirical studies developed – summarised in a series of reviews and books (e.g. Driver (1983); Driver et al., 1985; Gilbert et al., 1982; Gilbert & Watts, 1983) –, but also to the emergence of a new field of research related to how students learn, and to the formation of a whole generation of researchers who led the area for the next decades.

Moving ahead

At the same time John was involved with the research briefly described above and with the supervision of his first PhD students, he became interested in the parallels between the processes of science and science teaching. It seems such interest was first expressed in a paper, also written in collaboration with Roger Osborne, in which they discussed 'the types and uses of models found in science and science teaching', explored 'the contention that the misuse of models in science teaching can lead to misunderstandings by students of both models and their embodied concepts' (Gilbert & Osborne, 1980, p. 3), and raised a series of questions to guide further investigations. However, it was only some years later that the topic 'model' was focused on in John's studies.

In 1988, he became Professor of Science Education at the University of Reading, where he worked until his official retirement in 2005, when he was bestowed the title of Professor Emeritus. There, together with Carol Boulter, he ran the Centre for Models in Science and

Technology: Research in Education (CMISTRE), an international collaborative venture that brought together people who have an interest in models (including analogies) and modelling. From 1994, I had the privilege of becoming a member of the CMISTRE, one of the most remarkable experiences I have had during my Ph.D. course. For short stays or for attending the regular seminars where new ideas were democratically discussed, John invited scholars from the UK, Australia, Brazil, Israel, New Zealand, South Africa, and the USA. One of the seminal publications of the CMISTRE was the book *Developing Models in Science Education* (Gilbert & Boulter, 2000). It expressed ideas developed by the members of CMISTRE at that time, grouped in three sections focused, respectively, on the nature and significance of models, the development of mental models, and the teaching and learning of consensus models. Mainly due to John's comprehensive view on knowledge that he spread among the members of the Group, the book was based on ideas from disciplines like philosophy, history, sociology and language of science, and psychology of science teaching and learning. One example of the integration of ideas from distinct disciplines was the concept of hybrid model, initially published in one of the papers originated from my Ph.D. thesis (Justi & Gilbert, 1999) and discussed in one of the chapters of the book (Justi, 2000). It clearly illustrated John's view that, on the one hand, a significant idea should be obvious and clearly expressed, whilst on the other, it should make people think about and from it. In his academic life, several ideas had such characteristics (like the ones that based the previously discussed Interviews about Instances).

From the initial studies reported in that book, as well as from John's belief that science education must be more authentic (that is, 'as closely alike the conduct of science per se as is possible under the current conditions of mass education' (Gilbert, 2004, p. 116)), his interest in models and analogies advanced resulting in a new research programme focussed on modelling. Some of his previous ideas, like those concerning thought experiments – approached initially when he was at Surrey (Helm et al., 1985; Helm & Gilbert, 1985) and detailed later (Gilbert & Reiner, 2000; Reiner & Gilbert, 2000) – were crucial in that new enterprise.

In the first research project I coordinated after my Doctorate, John participated as a researcher and, as he always used to do, he made this an opportunity of mutual learning and production of knowledge. In the context of that project, when we started analysing the ideas expressed by teachers from distinct educational levels about models and modelling, we felt the need to deeply understand the meaning of modelling in science. This led us to study the philosophy of science, and the history of the development of some scientific ideas, as well as John Clement's (1989) ideas on modelling in science education – all of which inspired and informed our own ideas. In a well-known paper in which we published some of the results of that project (Justi & Gilbert, 2002), we proposed the first version of our Model of Modelling, a diagrammatic representation of how we understood the process. In the following years, that Model supported many empirical studies conducted in Brazilian regular classrooms that aimed at increasing the authenticity of science teaching through modelling-based teaching (MBT). All of them were discussed in the book that both brought together research we conducted during 15 years and presented our new studies and countless discussions mainly occurred from 2012 to 2015. Such discussions also resulted in the production of the new version of the Model of Modelling (Gilbert & Justi, 2016).

This book (which since its launch has been one of the top 25% best-selling books published by Springer) also shows how we managed to broaden our ideas and analysis of MBT situations by discussing issues concerning the contributions of MBT to a more authentic science education, the role of argumentation in MBT, the contributions of visualisation to MBT, analogies and analogical reasoning in MBT, the learning about science through MBT, learning progressions during MBT, and the education of teachers to facilitate MBT. In the last chapter, we also focus on challenges and novel perspectives, most of which have been addressed in the studies conducted in the last years. The discussion of some of them was interrupted by his unexpected death, last 9th February.

Two of the topics discussed in the book showed how he tried to think outside the box by approaching a given subject from distinct and innovative perspectives. One topic is visualisation, about which he wrote papers and individual chapters (e.g. Cheng & Gilbert, 2015; Gilbert, 2009b) and edited three books (Eilam & Gilbert, 2014; Gilbert, 2005b; Gilbert et al., 2008). In the introductory chapter of the first of these books, he explained that the emergence of visualisation as a focus of research could be related to two factors: the increasing

> "emphasis being placed on introducing students, at all levels of the education system, to the nature and processes of science"; and "the ready availability of powerful computers with which models, especially dynamic models and simulations, can be displayed and manipulates in a virtual format" (Gilbert, 2005c)

This justified the attempt of bringing together computer software specialists, scientists, and educationalists drawing on the insights from science, education, and cognitive psychology, in order to disseminate their ideas and promote the formation of links between them – which is also promoted in the second book. On the other hand, the third book discusses how science teachers use visual representations in diverse ways (mainly by using different diagrams, simulations and slow-motions), and in culturally diverse classrooms, as well as the place of visualisation in informal science education.

Thinking on teachers' development

The second topic discussed in one of the chapters of our book that had permeated John's previous projects and publications is teachers' development. A book published 20 years before (Bell & Gilbert, 1996), based on the findings of a three-year research project, presented and illustrated a model that integrates teachers' personal, professional, and social development. In Bell's view, the book is a significant contribution because it continues

> the debate about constructivist views of learning as applied to teacher education, moving it forward from personal into social constructivism, including what it means to be a science teacher on a collective basis. (Fensham, 2004, p. 110)

The teachers' development model proposed in this book also based John's additional reflections on the topic in a more recent chapter (Gilbert, 2010), where he also discussed the challenges of becoming an effective science teacher; approaches to successful professional development; and good practices in the organisation of teacher development activities.

As for teachers' development, a particular important project was coordinated by John and Matthew Newberry: the Cams Hill Science Consortium. It started in 2001 by involving teachers from six secondary schools in a collaborative classroom-based action research, a network that, by 2007, had expanded to teachers from over 30 primary and secondary schools in South East England (Gilbert & Newberry, 2007). From John and Matthew's initial ideas that models and modelling have a great potential to engage students in science lessons, issues concerning models and modelling were introduced, developed, and discussed during meetings. After each meeting, the teachers applied the discussed ideas in their classes and prepared a report of the outcomes to be presented and discussed in the next meeting. When commenting about this project, John always emphasised that (i) the production and discussion with the teachers of the representation for increasing levels of understanding required by the British National Curriculum and based on the distinct approaches to learning about models and modelling[1] were so interesting; and (ii) the outcomes of the project in terms of most teachers' engagement and level of reflection about their actions and what had happened in their classes. After so many years working at universities, returning to schools, even as a collaborative researcher, was a relevant experience in terms of giving him feedback on the application of many of the ideas he developed in collaboration with distinct colleagues throughout his career. In his words:

> I propose to ignore it (the retirement age of 65). The future looks bright, for I am now working ever-more intensively with Matthew Newberry and the teachers of the Cams Hill Science Consortium, who are conducting action research into the significance of 'models and modelling' for all aspects of the school science curriculum. (…) It would be wonderful, at the close of a career of 40 years, to be able to help science teachers of England regain some sense of professional self-determination after many years in the wilderness of the 'Stalinist command economy' created by the educational policies of successive UK governments since 1988. There are glimmers of hope. (Gilbert, 2005a , p. 4)

Acting in some other areas

John was fascinated by chemical ideas, as well as the particularities and challenges involved in teaching and learning chemistry. At the National Association for Research in Science Teaching (NARST) conference held in St. Louis in 2001, conversations among a group of chemical educators from different nationalities and with distinct experiences in terms of teaching and research resulted in the decision of editing a book on chemical education from the research perspective – then a missing topic in the literature. The book was published some years later (Gilbert et al., 2003).

At the interface of the research on models and modelling, on visualisation, and on chemistry education, John also dedicated special attention to the difficulties faced by students (and teachers) when dealing with the three types of representation of chemical knowledge: macro, sub-micro, and symbolic ones (Johnstone, 1982). Besides having discussed such issues in the context of papers mainly based on the above mentioned research, John co-edited, with David Treagust, a book focused on multiple representations in Chemistry (Gilbert & Treagust, 2009). The great reception of the book among the chemical education community made John think that the knowledge and the teaching and learning of the other major sciences (Physics and Biology) should be approached from the same perspective. As the editor of the series in which the book on multiple representations in Chemistry was published, he went to great lengths to find editors for the books on multiple representations in Biology and Physics (published in 2013 and 2017, respectively).

The interplay of two areas in a book was also found in Gilbert and Stocklmayer (2013). Both of them – science communication and the relations between science and technology education – have been addressed in John's previous enterprises or publications. At the University of Reading, he had created a course on science communication which he ran for some years attracting a huge number of students. Nowadays science communication can be viewed as a scientific practice that involves many distinct groups (e.g. scientists themselves, mediators, funding agencies, the general public) that try to communicate to each other through several modes and communication vehicles that not always are proper to communicate a given message to a given audience. The discussion of these and other related topics in John's course on science communication from his experience of being a good listener and communicator, and from his knowledge on both models of representation and people may have been the main causes of the success of the course. On the other hand, he always claimed that technology (rather than science) was the main focus of interest of the general public (including most of the students). Therefore, communication supported by evidence-based information involving technology education has to reach the general public. But how, if both are relatively new areas? That is the gap that this book tried to fill by providing an overview of the major issues concerning science and technology communication, an introduction to the research-based literature of the area, and suggestions for activities that may be explored by readers.

Finally, John's last published book (Rennie et al., 2019) addressed a topic he had been interested in for many years: adult and lifelong learning in science and technology. Like the central topic, the structure of this book is also different from all the others. Based on the analysis of case studies written by adults who learnt 'the science and technology they need to know in order to deal with issues in their everyday lives' (p. viii), the authors provide a research-based exploration of adults self-learning and tools to support adults' learning experiences.

Taking other positions

As John expressed in a previous mentioned quotation, he proposed to ignore the retirement age of 65 years. In the following year of his official retirement from the University of Reading, he started a Visiting Professorship at King's College London and, more recently, from 2017, he was an Honorary Fellow at the Australian National University. Apart from these official positions, he continued studying, participating in research projects with some colleagues, attending conferences (mainly the ESERA ones), writing papers and chapters, and editing books.

Due to his leadership in the area of models and modelling and his knowledge of the absence of seminal publications in the area, in 2003, after the book on Chemical Education had been published, John proposed to Springer the creation of the series of books *Models and Modelling in Science Education*. The aims of the series were related to issues he viewed as essential to the area: to draw together reports of research and evaluated innovations from throughout the world, so as to provide a global perspective on the field; to draw together research in the field that is conducted within diverse academic specialisms e.g. history and philosophy of science, cognitive science, the separate science disciplines, to provide an integrated whole; and to produce overviews of work in major sub-sectors of the field e.g. role in the curriculum, teaching methods, implications for teacher education. The first book published was the one on Visualisation, edited by him (Gilbert, 2005b). Until his death, John continued to be the series editor, dealing with proposals, helping book editors to produce relevant volumes. At the end of 2019, the 12th book of the series was published (Upmeier zu Belzen et al., 2019).

John was also invited by Routledge to edit the four volumes of the series *Major Themes in Education related to Science Education* (Gilbert, 2006). As requested by the title of the series, some of the most common important issues being debated in the area are addressed from distinct perspectives in the four volumes composed by 74 papers: 'Science, Education and the Formal Curriculum', 'Science Education and Assessment in the Formal Curriculum', 'Teaching and Learning in Science Education' and 'Conceptual and Teacher Development in Science Education'. By selecting such papers, John aimed at both providing students 'with an effective entry into the literature on complex themes', and supporting 'researchers in identifying important topics for enquiry' (v 1, p. 2). By having this later aim in mind, he tried to select papers whose authors were not only

> "from anglophone, industrially developed countries". Moreover, "Any lessons drawn from the articles included must be subject to the process of analogy to see if the topics addressed, the methods used, and the conclusions reached, are relevant in any particular national context. Unless this is done, there is a real risk of 'cultural imperialism' as one country's concerns are imported into another where they may be of marginal relevance." (v. 1, p. 2)

Only a world citizen who was really committed to promote science education that could make difference in people's life would think from such a perspective.

John was known as THE Editor of the *International Journal of Science Education* (IJSE), a post he occupied from 1991 to 2017! These were 26 years of dedication to improve the quality of the journal, to make it effectively international, to make it a vehicle of education for authors, reviewers and associate editors. John noted that the IJSE provided him

> with an opportunity to support science education at world level and especially to provide professional development for new and/or poorly resourced researchers. More selfishly, it enables me to keep abreast of trends in the field at global level. (Gilbert, 2005, p. 4)

From the discussions we had concerning editorship and difficult decisions, I (and I would say all the other associate editors who had the same kind of discussions with him) learnt a lot not only about science education or criteria to analyse manuscripts, but also about how to help authors to produce better papers.

Being awarded

In 2001, John received the NARST 'Distinguished Contributions to Science Education Through Research'. In his typical way of being, he said he 'was greatly honoured, and even more surprised to be given the annual award' (Gilbert, 2005, p. 4). I remember that, on the award day, he was wearing a special suit and had a large smile on his face and eyes (which I was fortunate to register in a photo), but kept it secret until his name was announced.

A different tribute, but I think as important as the NARST award, was a surprise ceremony funded by Taylor & Francis, the publishing of the IJSE, during the 2017 ESERA Conference in Dublin, in order to celebrate John's 29 years of dedication to the journal (since he had also been an associate editor from 1988 to 1991).

Thinking from a different perspective

John was knowledgeable and experienced and had an amazing amount of energy and clear thoughts about the directions for future research. He always tried to analyse what was being discussed from different perspectives and to ask hard and unexpected questions that others would avoid. In doing so, he advanced our thinking, he taught us that to face simple or complex situations with an open mind (and heart) and without prejudice against a given idea or approach may always be a way to reach a good result.

John respected and supported researchers of all nationalities and creeds (whether they were novices or experienced ones) and was always ready to introduce people to other people when this could result in the generation of an active synergy in research. He was also always willing to write reference letters with eagerness, objective and fairness, but maintaining a pleasant attitude. There aren't many people who can combine these qualities. This is one of the reasons John was special for many people.

John was an excellent, charismatic and inspiring mentor both in academic and personal life of his students (and, sometimes, his colleagues). He always listened to what was being said or asked trying to identify the relevant points to be emphasised in criticism or advice and leaving less relevant points out of focus. Maybe due to thinking broadly, to analysing facts and situations from distinct perspectives, or even as a kind of inheritance of having been a happy child (and a happy man), John had also a unique sense of humour, many times expressed in sincere smiles. And his smiles were special when directed at Julie, his beloved wife and company for more than 50 years. Being together with them in conference places (as many of those who may read this text know, Julie almost always accompanied him at conferences) or in their house, it was so sweet to see how they worried about each other (even in terms of ordinary things); how they took care of each other; how they supported each other; how, even being so different in some senses, they built their lives together.

It was my pleasure and great privilege to have met John, to have had him as my Ph.D. supervisor my main academic collaborator and inspiration for the last 25 years and, mainly, as a friend. So, to finish this text, I would like to thank you, John, for being such a special person. We, your friends, will miss you so much …

Note

1. Learning a curricular model, learning to use a model, learning to revise a model, learning to reconstruct a model, and producing learning to construct a model *de novo*, i.e. learning to modelling (Gilbert, 2004; Justi & Gilbert, 2002).

Acknowledgements

I am grateful to Jan van Driel and Gail Jones, for having invited me to produce this obituary, a difficult but essential task in trying to overcome the sadness of the current moment; and to many colleagues all over the world, for having sent messages expressing details of their personal relationships with John and emphasising some of his characteristics. Such messages became valuable contributions to the organisation of the countless ideas I had in my mind before writing this text.

Disclosure statement

No potential conflict of interest was reported by the author(s).

References

Bell, B. F., & Gilbert, J. K. (1996). *Teacher development: A model from science education*. Falmer.
Bridge, W. A., & Gilbert, J. K. (1977). The self-evolving questionnaire item. *Programmed Learning and Educational Technology*, *14*(2), 176–179. https://doi.org/10.1080/1355800770140211
Cheng, M. W., & Gilbert, J. K. (2015). Students' visualization of diagrams Representing the Human Circulatory system: The use of spatial isomorphism and representational conventions. *International Journal of Science Education*, *37*(1), 136–161. https://doi.org/10.1080/09500693.2014.969359
Clement, J. J. (1989). Learning via model construction and criticism: Protocol evidence on sources of creativity in science. In J. A. Glover, R. R. Ronning, & C. R. Reynolds (Eds.), *Handbook of creativity* (pp. 341–381). Plenum.
Driver, R. (1983). *The pupil as scientist*. Open University Press.
Driver, R., Guesne, E., & Tiberghien, A. (1985). *Children's ideas in science*. Open University Press.
Eilam, B., & Gilbert, J. K. (eds.). (2014). *Science teachers' use of visual representations*. Springer.
Fensham, P. J. (2004). *Defining an identity: The evolution of science education as a field of research*. Kluwer.
Gilbert, J. K. (1975). For the good of science. *Physics Bulletin*, *26*(12), 516. https://doi.org/10.1088/0031-9112/26/12/004
Gilbert, J. K. (1977). The study of student misunderstandings in the physical sciences. *Research in Science Education*, *7*(1), 165–171. https://doi.org/10.1007/BF02643123
Gilbert, J. K. (2004). Models and modelling: Routes to a more authentic science education. *International Journal of Science and Mathematics Education*, *2*(2), 115–130. https://doi.org/10.1007/s10763-004-3186-4
Gilbert, J. K. (2005a). *Constructing worlds through science education*. Routledge.
Gilbert, J. K. (2005b). *Visualization in science education*. Springer.
Gilbert, J. K. (2005c). Introduction. In J. K. Gilbert (Ed.), *Visualization in science education* (pp. 1–6). Springer.
Gilbert, J. K. (ed.). (2006). *Science education: Major themes in education*, Vols. 1–4. Routledge.
Gilbert, J. K. (2009a). Roger Osborne (1940-1985), University of Waikato, New Zealand. *Cultural Studies of Science Education*, *4*(2), 315–322. https://doi.org/10.1007/s11422-008-9165-2
Gilbert, J. K. (2009b). Visualization: A metacognitive skill in science and science education. In J. K. Gilbert, & D. F. Treagust (Eds.), *Multiple representations in chemical education* (pp. 9–28). Springer.
Gilbert, J. K. (2010). Supporting the development of effective science teachers. In J. Osborne, & J. Dillon (Eds.), *Good practice in science teaching: What research has to say* (pp. 274–300). Open University Press.
Gilbert, J. K., & Boulter, C. J. (eds.). (2000). *Developing models in science education*. Kluwer.
Gilbert, J. K., de Jong, O., Justi, R., Treagust, D. F., & van Driel, J. H. (eds.). (2003). *Chemical education: Towards research-based practice*. Kluwer.
Gilbert, J. K., & Justi, R. (2016). *Modelling-based teaching in science education*. Springer International Publishing.
Gilbert, J. K., & Newberry, M. (2007). The characteristics of the gifted and exceptionally able in science. In K. Taber (Ed.), *Science education for gifted learners* (pp. 15–31). Routledge.
Gilbert, J. K., Osborne, R. J., & Fensham, P. J. (1982). Children's science and its consequences for teaching. *Science Education*, *66*(4), 623–633. https://doi.org/10.1002/sce.3730660412
Gilbert, J. K., & Osborne, R. J. (1980). The use of models in science and science teaching. *European Journal of Science Education*, *2*(1), 3–13. https://doi.org/10.1080/0140528800020103
Gilbert, J., & Reiner, M. (2000). Thought experiments in science education: Potential and current realization. *International Journal of Science Education*, *22*(3), 265–283. https://doi.org/10.1080/095006900289877
Gilbert, J. K., Reiner, M., & Nakhleh, M. (eds.). (2008). *Visualization: Theory and practice in science education*. Springer.
Gilbert, J. K., & Stocklmayer, S. (eds.). (2013). *Communication and engagement with science education: Issues and dilemmas*. Routledge.
Gilbert, J. K., & Treagust, D. F. (eds.). (2009). *Multiple representations in chemical education*. Springer.
Gilbert, J., & Watts, M. (1983). Concepts, misconceptions and alternative conceptions: Changing perspectives in science education. *Studies in Science Education*, *10*(1), 61–98. https://doi.org/10.1080/03057268308559905
Gilbert, J. K., Watts, D. M., & Osborne, R. J. (1982). Students' conceptions of ideas in mechanics. *Physics Education*, *17*(2), 62–66. https://doi.org/10.1088/0031-9120/17/2/309

Helm, H., & Gilbert, J. K. (1985). Thought experiments and physics education – part 1. *Physics Education*, *20*(3), 124–131. https://doi.org/10.1088/0031-9120/20/3/311

Helm, H., Gilbert, J. K., & Watts, D. M. (1985). Thought experiments and physics education – part 2. *Physics Education*, *20*(5), 211–217. https://doi.org/10.1088/0031-9120/20/5/003

Johnstone, A. H. (1982). Macro- and micro-chemistry. *School Science Review*, *64*, 377–379.

Justi, R. (2000). Teaching with historical models. In J. K. Gilbert, & C. J. Boulter (Eds.), *Developing models in science education* (pp. 209–226). Kluwer.

Justi, R., & Gilbert, J. K. (1999). A cause of ahistorical science teaching: Use of hybrid models. *Science Education*, *83*(2), 163–177. https://doi.org/10.1002/(SICI)1098-237X(199903)83:2<163::AID-SCE5>3.0.CO;2-I

Justi, R., & Gilbert, J. K. (2002). Science teachers' knowledge about and attitudes towards the use of models and modelling in learning science. *International Journal of Science Education*, *24*(12), 1273–1292. https://doi.org/10.1080/09500690210163198

Osborne, R. J., & Gilbert, J. K. (1980). A technique for exploring students' views of the world. *Physics Education*, *15*(6), 376–379. https://doi.org/10.1088/0031-9120/15/6/312

Reiner, M., & Gilbert, J. (2000). Epistemological resources for thought experimentation in science learning. *International Journal of Science Education*, *22*(5), 489–506. https://doi.org/10.1080/095006900289741

Rennie, L. J., Stocklmayer, S. M., & Gilbert, J. K. (2019). *Supporting self-directed learning in science and technology beyond the school years*. Routledge.

Upmeier zu Belzen, A., Krüger, D., & Van Driel, J. (2019). *Towards a competence-based view on models and modeling in science education*. Springer International Publishing.

Rosária Justi
Chemistry Education, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil
rjusti@ufmg.br

International Journal of Science Education, Volume 42, Issue 4 (2020)

Editorial

Editorial

Celebrating the life of John Kenward Gilbert

Rosária Justi

Pages: 493-503

Published online: 31 Mar 2020

Articles

Article

Forms of inquiry-based science instruction and their relations with learning outcomes: evidence from high and low-performing education systems

Anindito Aditomo & Eckhard Klieme

Pages: 504-525

Published online: 25 Jan 2020

Article

Bringing design practices to chemistry classrooms: studying teachers' pedagogical ideas in the context of a professional learning community

Hanna Stammes, Ineke Henze, Erik Barendsen & Marc de Vries

Pages: 526-546

Published online: 30 Jan 2020

Article

Identifying predictors of teachers' intention and willingness to teach about cancer by using direct and belief-based measures in the context of the theory of planned behaviour

Benedikt Heuckmann, Marcus Hammann & Roman Asshoff

Pages: 547-575

Published online: 06 Feb 2020

Article

Characterising pre-service secondary science teachers' noticing of different forms of evidence of student thinking

Dirac Sze Him Lam & Kennedy Kam Ho Chan

Pages: 576-597

Published online: 30 Jan 2020

Article

How informal science learning experience influences students' science performance: a cross-cultural study based on PISA 2015

Xing Tang & Danhui Zhang

Pages: 598-616

Published online: 06 Feb 2020

Article

Teachers' ability to construct arguments, but not their perceived self-efficacy of teaching, predicts their ability to evaluate arguments

Evangelia Lytzerinou & Kalypso Iordanou

Pages: 617-634

Published online: 09 Feb 2020

Article

Low-cost emerging technologies as a tool to support informal environmental education in children from vulnerable public schools of

southern Chile

Dominique Alò, Andrea Castillo, Paula Marín Vial & Horacio Samaniego

Pages: 635-655

Published online: 18 Feb 2020

Article

The meaning of ‘experiment’ in the intended chemistry curriculum in China: the changes over the period from 1952 to 2018

Bing Wei & Yue Chen

Pages: 656-674

Published online: 06 Feb 2020

Note

Note

John Gilbert – A Publisher’s note

Ian White

Page: 675

Published online: 31 Mar 2020

Explore articles

Explore the most recently published articles

Teachers and pre-service teachers' scientific competencies: a methodological development for a systematic review

Marta Torrijos-Muelas et al.

Article | Published online: 9 Apr 2023

Life stories of graduate students in Chile and the United States: influences on becoming a scientist from childhood to adulthood

Marta A. Silva Fernández et al.

Article | Published online: 4 Apr 2023

DAST or VoSAL? Adaptation of the VoSAL to Turkish and a comparison between the instruments

Menşure Alkış Küçükaydın et al.

Article | Published online: 30 Mar 2023

Dedicating, faking, and surviving: disclosing tensions in how three women university students negotiate collectively celebrated norms across European contexts

K. P. Günter et al.

Article | Published online: 22 Mar 2023

View all latest articles

International Journal of Science Education

Ready to submit?

Start a new manuscript submission or continue a submission in progress

Go to submission site

Submission information

- Instructions for authors
- Editorial policies

Editing services

- Editing services site

About this journal

- Journal metrics
- Aims & scope
- Journal information
- Special issue information
- Society information
- Editorial board
- News & call for papers

Journal metrics

Usage

- **590K** annual downloads/views

Citation metrics

- **2.518 (2021)** Impact Factor
- **Q2** Impact Factor Best Quartile
- **2.776 (2021)** 5 year IF
- **4.1 (2021)** CiteScore (Scopus)
- **Q1 (2021)** CiteScore Best Quartile
- **1.791 (2021)** SNIP
- **1.151 (2021)** SJR

Understanding and using journal metrics

Journal metrics can be a useful tool for readers, as well as for authors who are deciding where to submit their next manuscript for publication. However, any one metric only tells a part of the story of a journal's quality and impact. Each metric has its limitations which means that it should never be considered in isolation, and metrics should be used to support and not replace qualitative review.

We strongly recommend that you always use a number of metrics, alongside other qualitative factors such as a journal's aims & scope, its readership, and a review of past content published in the journal. In addition, a single article should always be assessed on its own merits and never based on the metrics of the journal it was published in.

For more details, please read the Author Services guide to understanding journal metrics.

Journal metrics in brief

Usage and acceptance rate data above are for the last full calendar year and are updated annually in February. Speed data is updated every six months, based on the prior six months. Citation metrics are updated annually mid-year. Please note that some journals do not display all of the following metrics (find out why).

- **Usage:** the total number of times articles in the journal were viewed by users of Taylor & Francis Online in the previous calendar year, rounded to the nearest thousand.

Citation Metrics

- **Impact Factor*:** the average number of citations received by articles published in the journal within a two-year window. Only journals in the Clarivate Science Citation Index Expanded (SCIE) or Social Sciences Citation Index (SSCI) have an Impact Factor.

- **Impact Factor Best Quartile*:** the journal's highest subject category ranking in the Journal Citation Reports. Q1 = 25% of journals with the highest Impact Factors.

- **5 Year Impact Factor*:** the average number of citations received by articles in the journal within a five-year window.

- **CiteScore (Scopus)†:** the average number of citations received by articles in the journal over a four-year period.

- **CiteScore Best Quartile†:** the journal's highest CiteScore ranking in a Scopus subject category. Q1 = 25% of journals with the highest CiteScores.

- **SNIP (Source Normalized Impact per Paper):** the number of citations per paper in the journal, divided by citation potential in the field.

- **SJR (Scimago Journal Rank):** Average number of (weighted) citations in one year, divided by the number of articles published in the journal in the previous three years.

Speed/acceptance

- **From submission to first decision:** the average (median) number of days for a manuscript submitted to the journal to receive a first decision. Based on manuscripts receiving a first decision in the last six months.

- **From submission to first post-review decision:** the average (median) number of days for a manuscript submitted to the journal to receive a first decision if it is sent out for peer review. Based on manuscripts receiving a post-review first decision in the last six months.

- **From acceptance to online publication:** the average (median) number of days from acceptance of a manuscript to online publication of the Version of Record. Based on articles published in the last six months.

- **Acceptance rate:** articles accepted for publication by the journal in the previous calendar year as percentage of all papers receiving a final decision.

For more details on the data above, please read the Author Services guide to understanding journal metrics.

On Mon, Aug 12, 2019 at 12:04 PM International Journal of Science Education <onbehalfof@manuscriptcentral.com> wrote:

@@12-Aug-2019@@

Dear Dr. Aditomo:

I am writing to you in regard of Manuscript ID TSED-2019-0023-A.R1 entitled "Forms of inquiry-based science instruction a learning outcomes: Evidence from high and low-performing education systems" which you submitted to the International Jo Education. The comments of the reviewers are included at the bottom of this letter.

The reviewers require a major revision of your manuscript to be made before it can be considered further. Therefore, I inv reviewer' comments and substantially revise your manuscript.

To revise your manuscript, log into https://mc.manuscriptcentral.com/ijse and enter your Author Centre, where you will find listed under "Manuscripts with Decisions." Under "Actions," click on "Create a Revision." Your manuscript number has bee revision.

You will be unable to make your revisions on the originally submitted version of the manuscript. Instead, revise your manu processing program and save it on your computer. Please also highlight the changes to your manuscript within the docum changes mode in MS Word or by using bold or coloured text.

Once the revised manuscript is prepared, you can upload it and submit it through your Author Centre.

When submitting your revised manuscript, you are required to respond to the comments made by the reviewers in the spa

International Journal of Science Education - Decision on Manuscript ID TS 2019-0023-A.R1

Inbox ×

O

International Journal of Science Education <onbehalfof@manuscriptcentral.com> Mon, Aug 12, 2019, 5:0

to aditomo, me

@@12-Aug-2019@@

Dear Dr. Aditomo:

I am writing to you in regard of Manuscript ID TSED-2019-0023-A.R1 entitled "Forms of inquiry-based science instruction and learning outcomes: Evidence from high and low-performing education systems" which you submitted to the International Jou Education. The comments of the reviewers are included at the bottom of this letter.

The reviewers require a major revision of your manuscript to be made before it can be considered further. Therefore, I invite reviewer' comments and substantially revise your manuscript.

To revise your manuscript, log into https://mc.manuscriptcentral.com/ijse and enter your Author Centre, where you will find yo under "Manuscripts with Decisions." Under "Actions," click on "Create a Revision." Your manuscript number has been appe revision.

You will be unable to make your revisions on the originally submitted version of the manuscript. Instead, revise your manusc