



Geographical Education That Matters – A Critical Discussion of Consequential Validity in Assessment of School Geography

Chew Hung Chang and Tricia Seow

National Institute of Education, Nanyang Technological University, Singapore

Abstract

The word *test* comes to mind when a person, who is unacquainted with education discourses, reads about assessment issues. Beyond issues of reliability and validity in designing measurement constructs, assessment for school geography must result in better geographical learning. In other words, there must be *consequential validity* so that the way teachers collect information about students is aligned to the goals of improving learning. While geographical educators agree that finding out if someone has learnt what you intend for them to learn goes beyond performance in pen and paper examinations, school geography intends for children to learn beyond geographical knowledge. In fact, geographical educators are interested in evaluating if our students are better in developing skills that will help them be actively engaged and contributing citizens of the world that they are living in.

Introduction: Consequential Validity in Assessment

Assessment is important in teaching and learning because teachers need feedback on their practice and on students' learning (Voltz, Sims, & Nelson, 2010). Assessment also serves a diagnostic function, allowing teachers to identify the learning needs of different students, so as to better differentiate assessment for various student profiles. Finally, assessment serves as a means by which student progress (and teacher performance) may be reported to stakeholders including school administrators and parents. Given these different uses, it is important that assessment is reliable and valid.

While reliability in assessment refers to consistent replication of results when using a test, validity is concerned with the claims or inferences made from the test results. In other words, validity refers to how well assessment results are used in describing performance or inherent attributes of learners. Content validity pertains to the individual's "performance on a defined universe of tasks" (Shepard, 1993 p. 408), while criterion-

related validity refers to test types and the predictive accuracy for performance. Construct validity is needed "when making inferences about unseen traits such as intelligence or anxiety" (Shepard, 1993 p. 409).

Stobart (2001) notes that discussions of assessment validity in the literature include whether the concept of validity should include the issue of consequences of assessment. For instance, Messick (1995) argues the validity of a test should also consider "potential and actual social consequences of applied testing" beyond the content, criterion and construct validity that statisticians refer to. He argues that all assessment has positive or negative social consequences. While standardised testing may improve student learning and motivation, and allow students to have roughly equal access to classroom content, the negative aspects include merely teaching students to pass. While content, construct and criterion-related validity are key concerns in test interpretation, the use of test results has social consequences. This position is supported by other researchers (e.g. Crooks, Kane, & Cohen, 1996; Linn, 1997; Stobart, 2001), while other researchers like Popham (1997) and Mehrens (1997) have disagreed on the grounds that the social consequences of assessment go beyond the responsibility of the test setters and should be separated from validity arguments.

In general, discussions on consequential validity focus on the impacts on students of the assessment process itself. For example, Crooks et al. (1996) identify threats to consequential validity as the non-achievement of positive consequences and the occurrence of negative impacts such as poor student motivation and assessment anxiety. Stobart (2001), building on Crooks et al. (1996), suggests that this leads to the need for a review of inappropriate standards, as well as an evaluation of aspects of the assessment process itself, including the conditions of assessment, the nature of the assessment tasks, the weighting of different aspects of tasks, and scoring criterion.

In this paper, we move beyond the social implications of the assessment process itself to explore the concept of consequential validity in terms of whether assessment in geography helps meet the desired positive consequences of geographical education. In order to do this, however, we first examine what the aims of geographical education might be.

What should we be assessing in geography?

Geography as a discipline and school subject: what matters?

A consideration of the consequential validity of assessment in geography should begin with the question of what is considered core to geography as a discipline or school subject. Unsurprisingly, there is some divergence in the literature on this issue. For instance, Brooks, Qian, and Salinas-Silva (2017) suggest that how geography is understood may vary across space. Uhlenwinkel, Béneker, Bladh, Tani, and Lambert (2017) also note a propensity for teachers, across different European contexts, to have varying understandings of geography as a natural science or as a social studies subject. This may be due to the ways in which school subjects are organised across national contexts (Brooks et al. 2017; Uhlenwinkel et al., 2017). The attitudinal and behavioural goals of geographical education have also been debated. For example, Lambert and Balderstone (2000) argue that geography teachers cannot ignore the moral and ethical dimensions of their work. However, research has suggested a split between teachers who feel a responsibility to advocate for environmental attitudes and values (Ballantyne, 1999; Grace & Sharp, 2000), and those who are reluctant to do so (Tomlins & Froud, 1994; Cross, 1998; Cotton, 2006).

It is the view of the authors, however, that despite these differences, there is some congruence in general understandings of what the goals of geographical education might be. For instance, the International Charter on Geographical Education (International Geographical Union Commission on Geographical Education [IGU-CGE], 1992, 2016) provides a basis for this discussion. Bourke and Lane (2017) identified a number of key themes in both the 1992 and 2016 Charters. These include an explication of why a geographical education is beneficial and essential to the development of a person because it “helps people to understand and appreciate how places and landscapes are formed, how people and environments interact, the consequences that arise from our everyday spatial decisions, and Earth’s diverse and interconnected mosaic of cultures and societies” (IGU-CGE, 2016, p. 5). The authors also noted a

discourse of concern around the environment and an emphasis on how a geographical education is important to addressing this.

Geography is therefore a vital subject and resource for 21st century citizens living in a tightly interconnected world. It enables us to face questions of what it means to live sustainably in this world. Geographically educated individuals understand human relationships and their responsibilities to both the natural environment and to others. Geographical education helps people to learn how to exist harmoniously with all living species (IGU-CGE, 2016, p. 5).

The 1992 Charter also identified key conceptual knowledge and skills (1992, pp. 1.7–1.8), as well as ways of questioning and thinking, to be developed within a geographical education. These ideas have been further elaborated upon in the discussions around *Powerful Knowledge in Geography* (Stoltman, Lidstone, & Kidman, 2015; Lambert, Graves, & Slater, 2016, Maude, 2018), geographical thinking (see edited volume by Brooks et al., 2017) and *GeoCapabilities* (<http://www.geocapabilities.org>).

The discussions within geography appear well aligned to the Delors (1998) report to the United Nations Educational, Scientific and Cultural Organization, which included four pillars as key concepts to developing education for the twenty-first century. Delors (1998, p. 97) argues for an integrated approach to formal education based on the four pillars of learning including:

1. learning to know – a broad general knowledge but also depth in a few subjects;
2. learning to do – to acquire not only occupational skills but also the competence to deal with many situations;
3. learning to be – to develop one’s personality and to be able to act with growing autonomy, judgment and personal responsibility;
4. learning to live together – by developing an understanding of other people and an appreciation of interdependence.

In light of the discussions on what the aim of a geographical education might be, we suggest that assessment in geography has consequential validity if it allows teachers to infer if students have truly learnt geography through knowing, doing and being. We argue that assessments in geography that indicate someone has learnt about environmental problems, but which do not help to determine the social consequences of the student’s learning, may be seen as lacking consequential validity.

Defining the learning outcomes for assessing consequential validity in geography

Describing assessment consequential validity based on determining the interpretation and use of evidence of learning to know, do, be and live together may not provide practical guidelines to geographical educators. The Trends in International Geography Assessment Study (TIGAS) 2023 group (www.tigas2023.com) has been working on developing international geography assessment that meets the standards of the Trends in International Mathematics and Science Study

(TIMSS) eighth grade assessment mode, in order to introduce geography for international assessment in 2023. We refer to this work (of which the first author is a member) in order to provide a broad overview of how consequential validity (as described in the preceding section) might be incorporated into assessment design.

Table 1 below illustrates Krathwohl's (2002) revision of Bloom's taxonomy of knowledge domains and cognitive processes.

Griffin and Care (2014) suggest that developmental taxonomies such as this provide generic levels of

Table 1: Krathwohl's revision of Bloom's taxonomy

| | Cognitive Processes | | | | | |
|--------------------------|---------------------|------------|-------|---------|----------|--------|
| The knowledge dimensions | Remember | Understand | Apply | Analyse | Evaluate | Create |
| Factual | | | | | | |
| Conceptual | | | | | | |
| Procedural | | | | | | |
| Metacognitive | | | | | | |

(Source: Krathwohl, 2002, p. 216)

complexity and sophistication that can be used to classify and interpret task requirements and student task responses. Indeed, Krathwohl's categorisation of knowledge domains and cognitive processes indicates the need to consider both the curriculum content as well as cognitive processes for Geography in designing assessment. Teachers specify criteria to be evaluated which enables students to demonstrate the performance of those outcomes (Cohen, 1995).

The TIGAS group has referred extensively to the 1992 and 2016 Charters in order to delineate the learning outcomes that can guide assessment in geography, in terms of both the knowledge domains and cognitive processes outlined above. The knowledge domain includes knowledge of different issues such as globalisation, urbanisation, climate change, sustainable development, and food security, across spatial, social and cultural contexts so that children will be able to face questions of what it means to live sustainably in an interconnected world. Geography's key concepts are also key components of the knowledge domain for assessment. These include:

1. location and distribution,
2. place,

3. spatial interaction,
4. region, and
5. people-environment relationships.

These concepts are infused and learnt across topics about the earth's physical structure and physical environments, as well as human environments. There is also a consideration for the changing human environments across social, cultural and economic systems. Ultimately, geography students need to demonstrate an understanding of the interactions between humans and their environment in assessment with consequential validity.

The cognitive skills that geography students need to be able to demonstrate can be derived from the 1992 Charter and include:

1. identifying questions and issues,
2. collecting and structuring information,
3. processing, interpreting, and evaluating data,
4. developing generalisations,
5. making judgements,
6. making decisions,
7. solving problems, and
8. working co-operatively.

These cognitive skills can be simplified according to Krathwohl's terms of:

1. remember,
2. understand,
3. apply,
4. analyse,
5. evaluate, and
6. create.

However, these cognitive skills cannot be considered independently from the type of information and data that geography students have to work with including:

1. maps,
2. diagrams,
3. tables,
4. graphs,
5. pictures,
6. symbolic data,
7. quantitative data, and
8. verbal information.

(IGU-CGE, 1992)

These ways of thinking about learning geography provide a frame of reference in examining the notion of consequential validity. The TIGAS group described assessment specifications based on the content and cognitive domains, as well as a geographical practices domain. These domains map back to the topics that are common across geography curricula, cognitive learning outcomes that range from simple factual recall to hypothesis testing, the use of geographical concepts like space and place, together with resources such as maps.

This work is currently being prepared for publication, but it is important to note that the notion of consequential validity is implicit in this categorisation of test items, where evidence on how well students can solve geographical problems, that have social consequences, is included. The intentional extension beyond the content and cognitive domains to include geographical practice is aligned with the Charters' aim to develop "[g]eographically educated individuals [who can] understand human relationships and their responsibilities to both the natural environment and to others" (IGU-CGE, 2016, p. 5).

Having outlined what consequential validity might look like in geography, we move on to address the notion of consequential validity at three levels.

1. Is the assessment literature in the field of geography and environmental education of social consequence?

2. How can consequential validity in assessment be built into the geography curriculum for a country?
3. What do assessment items that have consequential validity look like?

Is the assessment literature in geography of social consequence?

Chang and Aman (2017) in a recent publication analysed all article titles published in four prominent geographical and environmental education journals between the period 2010 to 2017: *Environmental Education Research*; *International Research in Geographical and Environmental Education*; *The Journal of Environmental Education*; and *Journal of Geography*. The findings show that the published research articles contribute to achieving some of the action plan items from the International Charter on Geographical Education. In particular, there was a good spread of research published on assessing knowledge, skills and attitudes. There was also a number of assessment articles on environmental issues, but there was practically no research on issues of reliability and validity.

Another article by Lane and Bourke (2017) systematically reviewed over 700 articles on assessment in geographical education. While they concluded that more needs to be done to clarify the essential geographical knowledge and skills students should develop, they also called for more work to examine "the types and formats of assessment instruments that will provide valid and reliable measures" (Lane & Bourke, 2017). The fact that both papers point to a need to discuss reliability and validity issues is not coincidental, and indicates an area of work that is much needed.

The findings from these two papers affirm the framework for discussing consequential validity advanced in this paper. Both papers identified that formative assessment, or assessment for learning, is one of the key themes of research in geographical education. Moreover, assessment can be integrated within instruction for learning (Hagstrom, 2006) as there is vast potential in formative assessments in the form of class quizzes, reflection papers, posters or project work within the classroom (Chang, 2014). These practices will "inform the process before, during and after teaching has occurred" (Voltz et al., 2010, p. 116). Indeed, formative assessment can encourage testing beyond learning to know to learning to do, be and live together. High stakes standardised testing usually drives instruction in the classroom (Voltz et al., p. 114) and there is a pursuit of head knowledge at the expense of learning to do and learning to be.

In addition, Lane and Bourke (2017) also found only five out of the 700 papers related to geospatial thinking, of which four were in secondary or university settings and one in a primary setting. Chang and Aman (2017) also reported on six papers with the term *sustainability*. These topical themes are aligned to the cognitive as well as content domains that geographical education endeavours to achieve. This is unsurprising as “students require increasing international competence in order to ensure effective cooperation on a broad range of economic, political, cultural and environmental issues in a shrinking world” (IGU-CGE, 2016, p. 3). There is also an additional statement that says the Charter is supportive of the principle set out in the UN Sustainable Development Goals (IGU-CGE, 2016, p. 1).

In sum, the international level analyses presented by these two papers indicate a dearth of empirical work that describes research on issues of consequential validity. While the authors are confident that students around the world are assessed on key issues like their understanding of global warming, their ability to discuss the positive and negative impact of climate change on local agricultural practices, and even the use of maps and photographs to demonstrate these impacts, the research indicates that much more needs to be done to tackle the issue of consequential validity in the field of geographical education.

How can consequential validity in assessment be built into the geography curriculum?

In considering how assessment could support geographical education that is of social consequence – learning to know, do, be and live together – there has been an interesting development in Singapore where levels-grading and field-based geographical investigation has become a key component of the high school geography subject over the last decade. The Singapore example used here will illustrate how consequential validity can be a part of curriculum design, at the outset.

Incorporating consequential validity through changing the assessment type

High school geography in Singapore has adopted since 2007 a levels marking approach for part of the national written examination paper (Singapore Examinations and Assessment Board, 2010). The main rationale for the change was to allow assessors to collect evidence of students’ abilities to discuss and evaluate geographical problems beyond just describing or explaining them. In the past, students were required to answer 40

multiple-choice questions and four structured essay questions at the national examination for Geography. Marks were awarded for each point in the structured essay questions. The change in the curriculum, which is twinned with the change in examination format, encourages students to engage “the challenges of an increasingly globalised world . . . [and] to promote critical and creative thinking skills, and to nurture problem-solving and independent learning abilities in students” (Sellan, Chong, & Tay, 2006).

By 2014, Singapore introduced another change to the high school geography curriculum which required students to answer questions about how field-based geographical investigations can be conducted, how information and data can be collected, organised, analysed and presented, and what they can conclude based on their findings (Singapore Examinations and Assessment Board, 2014). This takes the national level examination a step towards assessment that has social consequence as the investigation framework was designed based on Roberts’ (2003) cycle of enquiry. Without a shift in the curriculum that requires students to conduct field investigation, and an accompanying change in the national assessment mode and form, students would continue only to learn to know. In fact, National Institute of Education researchers (Seow, Irvine, & Chang., 2018) found that this change in curriculum assessment helped teachers to induct novice practitioners (i.e. students) into geographical disciplinary ways of knowing and doing. Field-based inquiry featured the habitual elements of a signature pedagogy that inducts disciplinary novices “to *think*, to *perform*, and to *act with integrity*” like disciplinary practitioners (Shulman, 2005, p. 52).

Emphasising the behavioural and attitudinal domains in assessment

We also argue for the need to purposefully incorporate behavioural and attitudinal domains into assessment. Consider the following two quotations that come from students in Singapore about what they have learnt in the geography classroom, and how it has shaped their environmental behaviour:

“I don’t think most people would bring home what they actually discussed about. And some people would forget about it. Some people actually take down notes to study for the exams. I think after the exams, everybody would just (pause) yeah, forget about it.”

“Actually, I think exams are very effective of making us remember things. But erm . . . but (if) you remember, do you do it? I don’t, you know. I remember, I

know everything. I know things that I'm supposed to remember, but I don't . . . (long pause).”
(Chang, 2014)

While one can argue that teachers should be encouraging independent and critical thinking in their classrooms (Jickling, 1992; Aldrich-Moodie & Kwang, 1997), and that advocating for values education and social change for the environment (as suggested by Fien, 1993; Huckle, 1985; Morgan, 2012) is anti-educational (Williams, 2008), it is our contention that not providing any opportunity to examine the behavioural and attitudinal learning outcomes would be worse. For instance, some studies have found that teachers are uncomfortable and reluctant to advocate for the environment in their lessons (Tomlins & Froud, 1994; Cross, 1998), but are more comfortable in presenting a neutral position while discussing a range of different viewpoints about environmental topics (Cotton, 2006). In contrast, some teachers feel responsible for promoting environmental attitudes and behaviours (Ballantyne, 1999; Grace & Sharp, 2000). In fact, Ho and Seow (2015) have found that the same geography curriculum in Singapore is interpreted and enacted differently by teachers based on their beliefs about the purpose of geographical education.

Further, teachers will select different assessment based on where they stand on the issues. Teachers who want to focus on geographical knowledge will choose to test students' ability to apply geographical theories to make sense of the dynamic processes they observe in the field. Those who prefer to emphasise geographical skills may design assessment around students' field-based procedural knowledge. On the other hand, teachers who want to focus on behavioural or attitudinal dimensions may seek to evaluate the types of actions that students choose to take to tackle problems in the field, and pay close attention to the reasons guiding these actions.

Regardless of the ethical stances adopted by the teacher (to just present options versus promoting action), assessment has consequential validity if it encourages students to think about their own behaviours and attitudes. This could be through the national level curriculum and assessment design or the deliberate inclusion of behavioural and attitudinal aspects of learning outcomes. Although the Singapore example will be different from geographical education contexts in other regions and states, these innovations in assessment align with the vision of the 2016 Charter, where school geography plays a critical role in preparing young people to engage in the global issues of their time.

What do assessment items that have consequential validity look like?

To take the example of assessment in school geography in Singapore further, two examples of authentic assessment items are taken from grade 8 classrooms for discussion on how the items may or may not have consequential validity.

A question on the topic of human-environment interaction is shown in Example 1.

Study the photograph, which shows water pollution along the river, one of the negative consequences of slums and squatter settlements.



(By Jonathan McIntosh [CC BY 2.0 (<https://creativecommons.org/licenses/by/2.0/>)], from Wikimedia Commons)

- Define 'pollution'.
- Explain how slum and squatter settlements lead to water pollution such as the one shown in the photograph.
- Other than the pollution shown in the photograph, list and explain another form of pollution resulted by slum and squatter settlements.

Example 1: Sample Geography Content Domain Question

Example 1 shows how a question on human-environment interaction appears. The answers to Part (a) and Part (b) can be learned/memorised from the textbook. It is a retrieval of information (recall) and there is not much thinking or reasoning involved in the process. However, Part (c) requires the student to infer another possible form of pollution. For this question on photograph interpretation, students are required to know (e.g., recall or describe), apply (e.g., compare, interpret and relate) or reason (e.g., analyse, evaluate and draw conclusions).

What seems to be lacking in this example is consequential validity. A student who is able to define pollution, explain the contribution of

waste production through domestic activities in slums, and even extend the knowledge to noise pollution may or may not be someone who has attitudinal or behaviour learning outcomes associated with the empathy for people living in these adverse urban conditions. Indeed, evidence of the students' learning to live together aspect of

Delors' four pillars of education may not be easily obtained from this assessment item.

In the next example, students are asked four questions based on two information sources. One is a photograph of traffic congestion and the other is a description of the negative impact of traffic congestion.

Study Figs. 2A and 2B which show traffic congestion in Delhi.



(By NOMAD [CC BY 2.0 (<https://creativecommons.org/>)], via Wikimedia Commons)

Fig 2A: Stress, pollution, fatigue: How traffic jams affect your health.

Late Tuesday night, thousands of commuters and motorists were caught in a gridlock* on the roads during rush hour. The traffic jam was so bad that many commuters spent hours on the road just to get home. Some managed to reach their destination past midnight. Gridlocks are part of the daily grind here in Delhi.

Fig. 2B

* Gridlock – a type of traffic jam where continuous queues of vehicles block an entire

network of intersecting streets, bringing traffic in all directions to a complete standstill.

1. With the help of Fig. 2A and what you have learnt, describe the characteristics of traffic congestion.
2. Explain how inadequate transport infrastructure and poor provision of public transport services can lead to traffic congestion.
3. With reference to Fig. 2B and what you have learnt, explain how traffic congestion can affect people psychologically.
4. "Imposing road pricing is the best way to reduce traffic congestion." How far do you agree with this statement? Explain your answer using relevant examples.

Example 2: Sample Geography Cognitive Domain Question

Example 2 shows how students are required to demonstrate that they apply and provide reasons for their answers. Part 1 is a knowing question which requires students to recall what they have learnt from textbooks and describe the characteristics of traffic congestion. Students rely on recall of facts and information obtained from the textbook. However, Part 2 requires students

to apply what they have learnt and explain how inadequate transport infrastructure contributes to traffic congestion. Part 3 requires more thinking from students as they are required to use their own experience or information gathered from other resources, including the internet and newspapers, to explain the connection between traffic congestion and human psychology in the

transport context. Part 4 is a high-level thinking question which requires the student to analyse, relate and provide a reason for their conclusion. In essence, students who can answer this part will provide evidence that they have formed an opinion about the issue and furnish justification for their opinion. While we do not need students to develop any emotional response to the situation, we would like them to be able to critically analyse, and form an opinion based on sound reasoning. In particular, students will have to use examples to explain how having a road pricing system may or may not alleviate the traffic problem posed in this question. In this way, the item connects more with the social consequences of the traffic problem by engaging the student with the task of evaluating a possible mitigation strategy set against the issues that affect others' lives. In some sense, the student has to go beyond learning to know, do and be to also learn what it means to live together.

These two examples provided in this discussion are clearly not exhaustive but they do indicate that with the framework of considering consequential validity beyond cognitive, construct and criterion-related validity, assessment can help teachers collect evidence on what students have learnt in school geography.

Conclusion

At the core of the discussion in this paper, the authors are interested to find out if students have learnt what we have intended for them to learn. We have also sought to provide clarity on what we think students should be learning in geographical education, by drawing on the International Charter on Geographical Education (IGU-CGE 1992, 2016), as well as such major developments in geographical education research as Powerful Knowledge. Geographical education matters, as the documents, eloquently argue. As such, it is important that the ways in which we assess geographical learning also matter. We advocate for the use of the concept of consequential validity to move the discussion on assessment beyond content, criterion-related or construct validity, towards considerations of the social significance of a geographical education. This is critical in helping geographical educators design assessment that will help improve student learning in geography beyond learning to know to learning to do, be and live with each other.

Pen and paper examinations remain popular in many places, and we have suggested, in this paper, that there is room for innovation in summative assessment modes for geography by providing examples from the Singapore context. However, alternative, but not mutually exclusive methods of gathering evidence to measure

students' geographical learning, need to be considered, and researched.

“Assessment is today's means of understanding how to modify tomorrow's instruction” (Tomlinson, 2014, p. 17). Indeed, assessment should be considered as an integral part of the curriculum process, for while it helps us gather evidence to find out what students have learnt or how well we have taught them, it also helps us design instruction that is aligned to the intended outcomes, of learning to know, do, be and live together, through the school geography curriculum.

References

- Aldrich-Moodie, B. and Kwong, J. 1997. *Environmental Education*, London: Institute of Economic Affairs.
- Ballantyne, R. (1999). Teaching environmental concepts, attitudes and behaviour through geography education: findings of an international survey. *International Research in Geographical and Environmental Education*, 8, 40–58.
- Bourke, T., & Lane, R. (2017). A Comparison of the International Charters on Geographical Education. *Journal of Geography*, 117, 174–180.
- Brooks, C., Qian, G., & Salinas-Silva, V. (2017). What next for Geography Education? A perspective from the International Geographical Union–Commission for Geography Education. *J-Reading-Journal of Research and Didactics in Geography*, 1(6), pp. 5–15.
- Chang, C. H. (2014). *Climate change education: Knowing, doing and being*. Abingdon: Routledge.
- Chang, C. H., & Aman, M. F. (2017). The International Charter on Geographical Education – a reflection on published research articles on assessment. *J-Reading: Journal of Research and Didactics in Geography*, 2(6).
- Cohen, P. (1995). Designing performance assessment tasks. *Education Update*, 37(6), 1–5.
- Cotton, D. R. E. (2006). Implementing curriculum guidance on environmental education: The importance of teachers' beliefs. *Journal of Curriculum Studies*, 38, 67–83.
- Crooks, T. J., Kane, M. T., & Cohen, A. S. (1996). Threats to the valid use of assessments. *Assessment in Education: Principles, Policy & Practice*, 3, 265–286.

- Cross, R. T. (1998). Teachers' Views about What to do about Sustainable Development. *Environmental Education Research*, 4, 41–52.
- Delors, J. (1998). *Learning: The treasure within – report to UNESCO of the International Commission on Education for the twenty-first century*. Paris: Unesco.
- Fien, J. (1993). *Education for the environment: critical curriculum theorising and environmental education*. Geelong, Deakin University Press.
- Grace, M., & Sharp, J. (2000). Exploring the actual and potential rhetoric-reality gaps in environmental education and their implications for pre-service teacher training. *Environmental Education Research*, 6, 331–345.
- Griffin, P., & Care, E. (Eds.). (2014). *Assessment and teaching of 21st century skills: Methods and approach*. Dordrecht: Springer.
- Hagstrom, F. (2006). Formative learning and assessment. *Communication Disorders Quarterly*, 28, 24–36.
- Halstead, J. M. & Taylor, M. J. (1996). *Values in education and education in values*. London: Falmer Press.
- Ho, L. C., & Seow, T. (2015). Teaching controversial issues in geography: Climate change education in Singaporean schools. *Theory & Research in Social Education*, 43, 314–344.
- Huckle, J. (1985). Values education through geography: a radical critique. In D. Boardman (Ed.), *New directions in geographical education* (pp. 1–13). London: Falmer Press.
- International Geographical Union Commission on Geographical Education [IGU-CGE]. (1992). *International charter on geographical education*. Retrieved from <http://www.igu-cge.org/wp-content/uploads/2018/02/1.-English.pdf>
- International Geographical Union Commission on Geographical Education [IGU-CGE]. (2016). *International charter on geographical education*. Retrieved from http://www.igu-cge.org/wp-content/uploads/2018/02/IGU_2016_def.pdf
- Jickling, B. (1992). Why I don't want my children to be educated for sustainable development. *The Journal of Environmental Education*, 23(4), 5–8.
- Krathwohl, D. (2002). A revision of Bloom's Taxonomy: An overview. *Theory into Practice*, 41, 212–218.
- Lambert, D., & Balderstone, D. (2000). *Learning to teach geography in the secondary school: a companion to school experience*. London: Routledge.
- Lambert, D. M., Graves, N., & Slater, F. (2016). Geography as powerful knowledge. *International Research in Geographical and Environmental Education*, 25(3), 189–194.
- Lane, R., & Bourke, T. (2017). Assessment in geography education: a systematic review. *International Research in Geographical and Environmental Education*, online first publication. doi.org/10.1080/10382046.2017.1385348
- Linn, R. L. (1997). Evaluating the validity of assessments: The consequences of use. *Educational Measurement: Issues and Practice*, 16(2), 14–16.
- Maude, A. (2018). Geography and powerful knowledge: a contribution to the debate. *International Research in Geographical and Environmental Education*, 27(2), 179–190.
- Mehrens, W. A. (1997). The consequences of consequential validity. *Educational Measurement: Issues and Practice*, 16(2), 16–18.
- Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50, 741–749.
- Morgan, A. (2012). Morality and geography education. In G. Butt (Ed.), *Geography, Education and the Future* (pp. 187–205). London: Continuum.
- Popham, W. J. (1997). Consequential validity: Right concern-wrong concept. *Educational Measurement: Issues and Practice*, 16(2), 9–13.
- Roberts, M. (2003). *Learning through enquiry: Making sense of geography in the key stage 3 classroom*. Sheffield: Geographical Association.
- Sellan, R., Chong, K., & Tay, C. (2006). Assessment shifts in the Singapore education system. Retrieved from Singapore Examinations and Assessment Branch at <https://www.seab.gov.sg/>
- Seow, T., Irvine, K., & Chang, J. (2018, March). *The Sustainability Learning Lab: Enhancing geographical inquiry in the field and classrooms*. Presentation at the Humanities Research Seminar.
- Shepard, L. A. (1993). Chapter 9: Evaluating test validity. *Review of research in education*, 19, 405–450.

- Shulman, L. S. (2005). Signature pedagogies in the professions, *Daedalus* 134(3), 52–59.
- Singapore School Examinations and Assessment Board. (2010). *GCE O-Level Syllabus Page*. Retrieved from Singapore Examinations and Assessment Board.
- Singapore Examinations and Assessment Board. (2014). *GCE O-Level Syllabus 2236*. Retrieved from Singapore Examinations and Assessment Board: https://www.seab.gov.sg/content/syllabus/olevel/2018Syllabus/2236_2018.pdf
- Stobart, G. (2001). The validity of national curriculum assessment. *British Journal of Educational Studies*, 49(1), 26–39.
- Stoltman, J., Lidstone, J., & Kidman, G. (2015). Powerful knowledge in geography: IRGEE editors interview Professor David Lambert, London Institute of Education. *International Research in Geographical and Environmental Education*, 24, 1–5.
- Tomlins, B., & Froud, K. (1994). *Environmental education teaching approaches and students' attitudes*. Slough: National Foundation for Educational Research.
- Tomlinson, C. (2014). *The differentiated classroom: Responding to the needs of all learners* (2nd edition). Alexandria, VA: Association for Supervision and Curriculum Development.
- Voltz, D., Sims, M., & Nelson, B. (2010). *Connecting teachers, students and standards: Strategies for success in diverse and inclusive classrooms*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Uhlenwinkel, A., Béneker, T., Bladh, G., Tani, S., & Lambert, D. (2017). GeoCapabilities and curriculum leadership: balancing the priorities of aim-based and knowledge-led curriculum thinking in schools. *International Research in Geographical and Environmental Education*, 26, 327–341.
- Williams, A. (2008). *The enemies of progress: The dangers of sustainability*. Exeter: Societas Imprint Academic.