

PAPER • OPEN ACCESS

Coherently Connecting Computer and Information Literacy Classroom and Accountability Assessments

To cite this article: Emily Pey-Tee Oon and W P Fisher Jr 2018 *J. Phys.: Conf. Ser.* **1065** 022010

View the [article online](#) for updates and enhancements.



IOP | ebooks™

Bringing you innovative digital publishing with leading voices to create your essential collection of books in STEM research.

Start exploring the collection - download the first chapter of every title for free.

Coherently Connecting Computer and Information Literacy Classroom and Accountability Assessments

Emily Pey-Tee Oon¹, W P Fisher Jr²

¹Faculty of Education, University of Macau, Taipa, Macau SAR, China

² Research Associate, BEAR Center, Graduate School of Education, University of California, Berkeley, Berkeley California 94702 USA

¹E-mail: peyteoon@umac.mo

Abstract. Can classroom assessment be improved by providing teachers with tools useful in more coherently connecting formative assessment in the classroom and international accountability assessments? Though formative assessment is widely recognized for its value in improving educational outcomes, teachers' efforts in putting it to use are plagued with problems that can be addressed and solved, though not without difficulty. Effective implementation of formative assessment and teachers' capacities for learning from their own and their colleagues' experiences with it are typically stymied in education by a lack of coherence between classroom and accountability assessments. Over 2,000 students at 118 Hong Kong schools participating in the 2013 International Computer and Information Literacy Study provided responses to 62 assessment items. Applying a probabilistic model implementing Rasch measurement theory, the published ICILS item calibrations were reproduced by the Hong Kong sample (Pearson's $r=0.86$; disattenuated, 1.00). Plans for an expanded computer and information literacy item bank supporting connections between classroom and international assessment are described.

1. Introduction

Information, Media and Technology (IMT) is an area of skills students need to succeed in 21st century work and life [1]. Computer and Information Literacy (CIL), an aspect of IMT, is a key skill set enabling information handling, communication, and collaboration in digital society [2,3]. However important CIL may be, concerns remain as to how it is assessed, with more attention to measurement-based formative integrations of assessment and instruction being called for [3-5]. In addition, demands on teachers for evidence of learning outcomes are made in the context of accountability assessments that are incoherently disconnected from classroom assessments; teachers are then held responsible for results not systematically connected to the learning process [6-8]. We aim to resolve these comparability problems to some degree by adapting the International Computer and Information Literacy Study (ICILS) [3] for formative use.

2. Theoretical framework

2.1 Assessing CIL

The ICILS, administered in 2013, was the first international comparative study assessing CIL among secondary school students. Used in 21 countries, the CIL in ICILS is defined as "individual's ability to



use computers to investigate, create, and communicate in order to participate effectively at home, at school, in the workplace, and in the community” [9].

The CIL for each aspect of the two strands was measured using three different question types: (a) Information-based response tasks – a non-interactive representation of a computer-based problem made use of technology to record students’ responses; (b) Skill tasks – an interactive simulation of software applications required students to complete tasks either linearly (e.g. Open a file) or non-linearly (e.g. Select, copy, and paste image); and (c) Authoring tasks – Use of authentic software applications to modify and create information. Measures produced from responses to the questions from the two strands were found to be unidimensional. The Hong Kong data were evaluated for unidimensionality before further investigations were undertaken.

2.2 Scientific measurement theory

The paradigmatic differences between statistical and scientific models have been the subject of intense debate for decades; a 1967 article by Meehl [10] defined issues remaining unresolved today. Statistical models are used to describe data and maximize the multivariate variance explained at the group level. Measurement models, in contrast, test hypotheses concerning the existence of a quantitative unit of individual comparison, with the goal of replacing statistical significance with substantive significance.

A substantively significant scientific model intended to quantify educational outcomes ought to take a mathematical form that explicitly states these assumptions and makes them testable. For instance,

$$\ln[P_{nij} / (P_{nij-1})] = B_n - D_{ij}$$

says that the log-odds of observed success on CIL tasks for student n on item i at partial credit score j is equal to the difference between the estimate B of person n 's CIL ability and the CIL difficulty estimate D of item i at rating j [11-12].

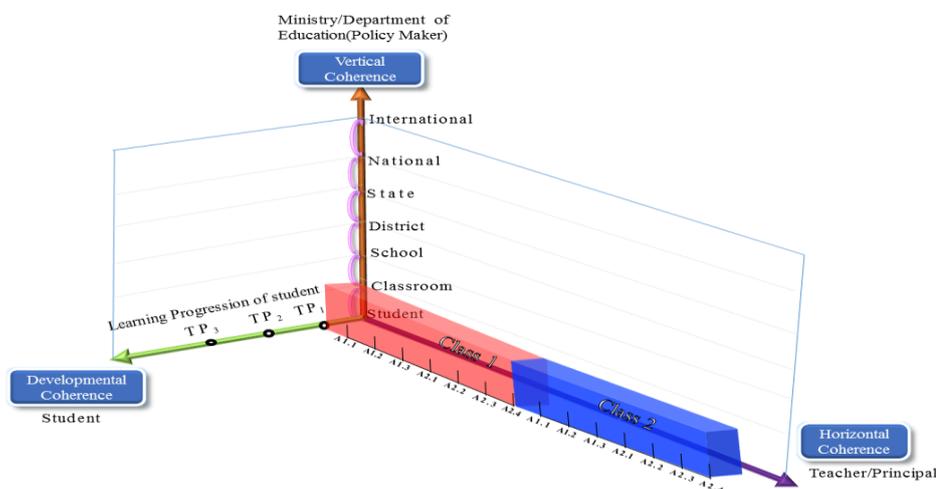


Figure 1. Developmental, horizontal, and vertical coherence (TP=Time Point; A1.1-A2.4 are ICILS aspects)

3. Research Objectives

Our goals are to conceptualize and demonstrate a CIL assessment system prototype capable of coherently connecting formative and accountability applications. Realization of these goals is facilitated by a model of developmental, horizontal, and vertical coherence across assessments (Figure 1). Research in progress will provide an example to teachers on how to evaluate learning progressions of students relative to national CIL curriculum goals, individually and in groups, by mapping their CIL abilities onto a progress chart representing the overall, coherent construct.

4. Methods

4.1 Sample

This study used data from the 2013 ICILS. A total of 150 schools were randomly sampled in Hong Kong by IEA using a procedure in which the sampling probability was proportional to school size [3]. Sampled schools refusing participation were replaced by schools closest in size and other characteristics. Twenty secondary two students were then randomly selected from each participating school. A total of 118 (79%) of the sampled Hong Kong schools, with 2089 students, participated. As the required participation rate was 85%, Hong Kong results were categorized in the final ICILS report under the heading of “countries not meeting sample requirement,” alongside Denmark, the Netherlands, and Switzerland [3].

4.2 Instrument

Students answered computer-administered questions posed in two of four 30-minute test modules. Each module followed a similar sequence starting from short questions followed by a larger task. The four modules comprise of a total of 62 items unified by theme across seven different aspects of CIL. Nine items were scored in three categories as incorrect (0), partially correct (1), and fully correct (2).

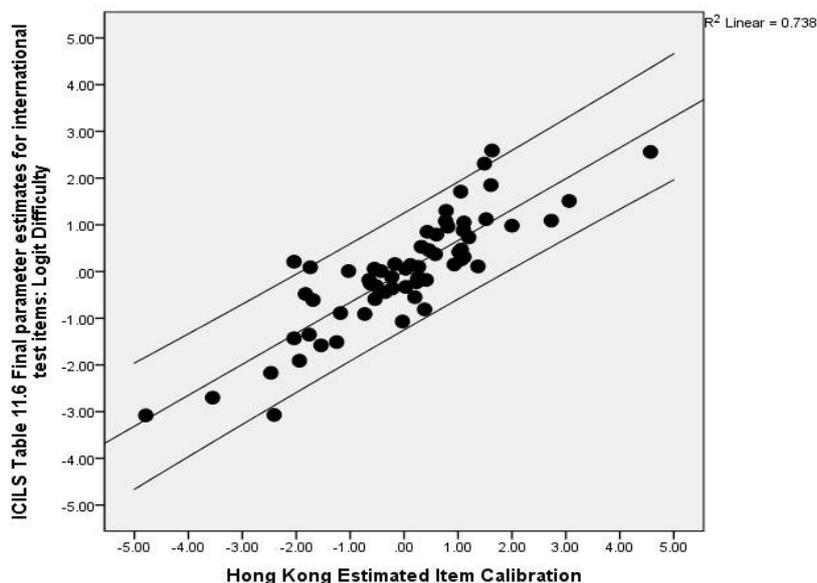


Figure 2. ICILS published item calibrations vs Hong Kong only estimates

5. Results

The item difficulties estimated from the Hong Kong data alone correlated 0.86 (disattenuated, 1.00) with the published ICILS values (Figure 2). Modelled measurement separation reliability was 0.84. ICILS reports that the overall international mean measure was 500, with a standard deviation of 100, and that the Hong Kong mean measure was 509 [3]. No Hong Kong standard deviation was given, though a graph can be read as showing the 5th, 25th, 75th, and 95th percentiles as being at about 325, 440, 575, and 640, respectively. Rescaling the logit measures in our analysis reproduced the mean of 509, with a standard deviation of 97, and the respective 5th to 95th percentiles at 322, 449, 580, and 645. ICILS reports the Hong Kong standard error at 7.4, but our estimates put it at about 39. Measures estimated from the Hong Kong sample are then consistent with the international measures and can be interpreted in the same frame of reference. Because the ICILS content evaluates learning incorporated in the Hong Kong CIL curriculum, teachers in Hong Kong could, then, use the ICILS items in formative assessments scaled to the same unit of measurement as reported in the extensive analyses published by the University of Hong Kong [3].

For instance, a student with a measure of 508, near the overall Hong Kong mean of 509, could, like the following examinee (entry order number 1564) in the Hong Kong data, exhibit this pattern of 13 partially or fully correct (1 and 2) and 14 incorrect (0) responses, shown here in difficulty order:

1564 + 1 11 0 121 1 100 1 0 200 00 1110 0000 0

Teachers with interpretive guidelines showing the item content could use this feedback on student performance to individualize lesson plans capitalizing on strengths and targeting special areas of weakness (the incorrect response to the fourth item from the left).

6. Discussion

Formative feedback revealing progress in learning over time can be presented in ways coherently connecting classroom and accountability assessments. At the individual student level, responses to the 62 ICILS items could be mapped into the developmental, horizontal and vertical forms of coherence shown in Figure 1. The coherent borrowing of information across different assessments allows teachers to record progress in learning over different time (developmental coherence), across classes (horizontal coherence), and in relation to accountability demands (vertical coherence). Meaningful comparisons of these kinds will enable teachers to maximize outcomes while also advancing their own learning.

7. Conclusion

Teachers need to know that coherent representations of assessment and instruction exist. Psychometric models provide mathematical formulations of learning progressions that enable experimental evidence-based tests of the independence of student learning from problem difficulty. Where the state of the art in education assumes that student abilities and curricular challenges float in uncontrollable states of flux, decades of research results show that these parameters can be brought into meaningful relation. Making this explicit is a necessary step in the direction of creating a credible context in which teachers might confidently rely on test data to improve student learning.

8. References

- [1] Partnership for 21st century Learning 2017 *Framework for 21st century Learning*. www.p21.org
- [2] Griffin P, McGaw B, and Care E 2012 *Assessment and teaching of 21st century skills* (Dordrecht: Springer)
- [3] Law N, Yuen J, and Lee Y 2015 *E-Learning Pedagogy and School Leadership Practices to Improve Hong Kong Students Computer and Information Literacy: Findings from ICILS 2013 and Beyond* (Hong Kong: University of Hong Kong)
- [4] Wilson M, Scalise K, and Gochyyev P 2015 *Thinking Skills and Creativity* **18** 65-80
- [5] Siddiq F, Gochyyev P, and Wilson M 2017 *Computers & Education* **169** 11-37
- [6] Gorin J, and Mislevy R 2013 *Inherent measurement challenges in the next generation science standards for both formative and summative assessment* (Princeton, NJ, USA: ETS)
- [7] Wilson M 2004 *Towards coherence between classroom assessment and accountability* (Chicago: University of Chicago Press)
- [8] Ladd H F 2017 *J Policy Analysis and Management* **36** 461-469
- [9] Fraillon J, Schulz W, and Ainley J 2013 *International Computer and Information Literacy Study assessment framework* (Amsterdam, the Netherlands: IEA)
- [10] Meehl P E 1967 *Philosophy of Science* **34** 103-115
- [11] Rasch G 1980 *Probabilistic models for some intelligence and attainment tests* (Chicago: University of Chicago Press)
- [12] Andrich D 1978 *Psychometrika* **43** 561-573