

This article was downloaded by: [Stanford University]

On: 27 September 2008

Access details: Access Details: [subscription number 776101540]

Publisher Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Applied Measurement in Education

Publication details, including instructions for authors and subscription information:

<http://www.informaworld.com/smpp/title~content=t775653631>

Guest Editor's Introduction

Richard J. Shavelson ^a

^a Stanford University,

Online Publication Date: 01 October 2008

To cite this Article Shavelson, Richard J.(2008)'Guest Editor's Introduction',Applied Measurement in Education,21:4,293 — 294

To link to this Article: DOI: 10.1080/08957340802347613

URL: <http://dx.doi.org/10.1080/08957340802347613>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

GUEST EDITOR'S INTRODUCTION

Today, formative assessment is high on policymakers' and educators' lists of instructional tools for reforming education. The logic seems impeccable and goes as follows. Formative assessment should be carried out by teachers in the service of student learning because such assessment makes clear learning goals, indexes where students are with respect to those goals, and, consequently, reveals a gap between what is and what is desired. Moreover, the logic goes, formative assessment would require teachers to use their knowledge of "the gap" to provide timely feedback to students as to how they might close that gap. In science classes, for example, argumentation provides an important feedback mechanism: Students should, science educators assume, be asked to give (often conflicting) explanations of the natural world (e.g., why things sink or float) and to justify their views based on empirical evidence. In this way, so the reasoning goes, students receive timely feedback from one another and the teacher, eliminating untenable explanations along the way. By embedding assessments that elicit students' explanations—formally within a unit, as part of a lesson plan, or as "on-the-fly" teachable moments occur—teachers would take this opportunity to close the gap in student understanding. As a consequence students' learning would be expected to improve. Finally, this logic is supported by Paul Black and Dylan Wiliam's now famous review of formative-assessment studies that found, if carried out well, formative assessment would be expected to produce moderate or large learning gains compared to "regular" instruction.

The research on formative assessment reported in this special issue addresses a "big idea": for a relatively small investment—embedding conceptually coherent formative assessments in a nationally used science curriculum—a big learning payoff would be realized.

Although influenced by the formative-assessment logic, we did not arrive at this big idea immediately. The research reported here grew out of extensive prior experience working with science teachers as we embedded performance assessments in their science programs, and out of the research literature, such as that reported in Black and Wiliam.

With respect to teaching practice, we noted long ago, and well before the current formative assessment movement, that teachers did not take full advantage of inquiry

(kit-based) science programs. One possible reason was that the programs were so broad multiple goals might be reached with the materials. A second possible reason was that teachers were unsure of which goals were to be achieved. Some combination of the two is most likely. We reasoned that by embedding performance assessments in their kit-based curricula, we might offer teachers guideposts to the instructional goals of the unit. Moreover, teachers would have additional information about students' performance where teachers initially *observed* and did not immediately *intervene* as the assessments produced "teachable moments." Only once student performance played out would teachers then focus on gaps and inquiry science teaching would be guided toward reasonable ends.

With respect to the research literature, we found the Black and Wiliam review (and others) compelling. We reached this conclusion perhaps somewhat optimistically, while realizing at the same time that much of the research found in such reviews was carried out in the laboratory with internal-validity controls whereas the research findings were generalized to the messiness of classrooms with teachers implementing "the treatment," so-called external-validity concerns.

Nevertheless, and with considerable euphoria, colleagues at the Stanford Education Assessment Laboratory and the Curriculum Research & Development Group (CRDG) at the University of Hawaii entered into a *romance*—of assessment developer and curriculum developer—to provide embedded formative assessments in a nationally used curriculum developed at CRDG: *Foundational Approaches in Science Teaching*. In the study we asked: (a) What does it take to build and embed conceptually coherent formative assessments in a curriculum? (b) Do teachers implement formative assessments as intended after training? (c) What is the impact of formative assessments on students' learning, motivation, and conceptual change? (d) What might be learned from the collaboration of curriculum developers and assessment developers so that in the future, such romances may learn from our experience.

I encourage you to dip into the pages of this *special issue* for tentative answers to these questions, and to a whole host of new questions that emerged from the research. After five years of work, our euphoria devolved into a reality that formative assessment, like so many other education reforms, has a long way to go before it can be wielded masterfully by a majority of teachers to positive ends. This is not to discourage the formative assessment practice and research agenda. We do provide evidence that when used as intended, formative assessment might very well be a productive instructional tool. Rather, the *special issue* is intended to be a sobering call to the task ahead.

Richard J. Shavelson
Stanford University