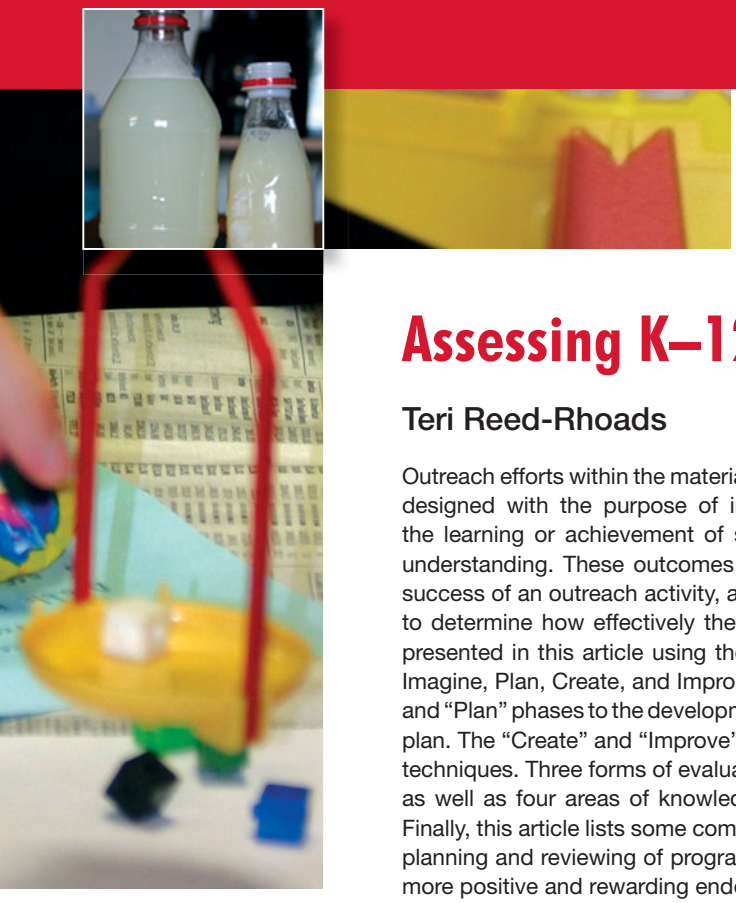


Assessing K–12 outreach

Teri Reed-Rhoads

Outreach efforts within the materials science community include learning goals or outcomes designed with the purpose of increasing awareness or interest in the field, improving the learning or achievement of subject matter, or even supporting a type of literacy or understanding. These outcomes need to be defined and measured. To demonstrate the success of an outreach activity, an appropriate assessment scheme must be implemented to determine how effectively the goals have been met. Assessment of K–12 outreach is presented in this article using the framework of a K–12 engineering design cycle of Ask, Imagine, Plan, Create, and Improve. Links will be drawn from among the "Ask," "Imagine," and "Plan" phases to the development of objectives and the establishment of the assessment plan. The "Create" and "Improve" phases will be connected to data collection and analysis techniques. Three forms of evaluation are discussed—formative, summative, and interim—as well as four areas of knowledge—cognitive, affective, behavioral, and metacognitive. Finally, this article lists some common pitfalls to avoid when considering these issues in the planning and reviewing of programs that will make assessment of K–12 outreach efforts a more positive and rewarding endeavor.

Introduction

Appropriate assessment is an essential component of all education outreach programs. While it can be fun and rewarding for researchers to engage with K–12 students and teachers and share their knowledge and personal passion for the engineering and science of materials, experience has taught us that what intuitively seems valuable often does not achieve the anticipated goals. As researchers, we often start with a hypothesis and follow with a controlled experiment. This helps to assess the validity of our original projections. In education, we use assessment to measure the hypothesis or outcomes of outreach programs in an experimental manner. Through carefully planned and implemented assessments, which validate the efficacy of programs, researchers can maintain or gain further support and/or interest not only from funding agencies but also from various K–12 constituents. This article discusses the assessment process using a framework that is already known to many engineers and scientists: the design process. Including this article first in this series of articles on outreach illustrates the importance of planning the assessment process upfront before a program is implemented.

Designing assessment

Engineering is most often defined using some reference to design and/or a design cycle. In a recent National Academy of Engineering report on K–12 Engineering Education, engineering was defined as "the process of designing the human-made world."[1] The engineering design process typically involves a relatively small number of steps that engineers use to solve problems and is very similar to the scientific method. Given that engineers and scientists are already familiar with these processes, we will use the design cycle here as foundational knowledge to describe the assessment process.

Though there are numerous design cycles, one often used in K–12 outreach is the following: Ask, Imagine, Plan, Create, and Improve. This process is displayed in **Figure 1** and has been used successfully in *Engineering is Elementary*,[2] a project that fosters engineering and technological literacy among children, and *TeachEngineering*,[3] a digital library that provides teacher-tested, standards-based engineering content for K–12 teachers to use in science and mathematics classrooms. Note that at any step, it may be necessary to go back one or two steps or repeat a step; therefore, the design cycle is not a linear system. In fact, as the image of the design cycle displayed in Figure 1 illustrates, it is circular, and circles can be imposed on top of circles as steps are reiterated, demonstrating the need for critical thinking and creativity in engineering.

The National Academy of Engineering report, on "Changing the Conversation: Messages for Improving Public Understanding of Engineering,"[4] highlights that the messages in outreach programs can have a positive impact on a young person's decisions to pursue a career in, and the public perception of, engineering.

Teri Reed-Rhoads, Purdue University; trhoads@purdue.edu
DOI: 10.1557/mrs.2011.62

© 2011 Materials Research Society
https://doi.org/10.1557/mrs.2011.62 Published online by Cambridge University Press

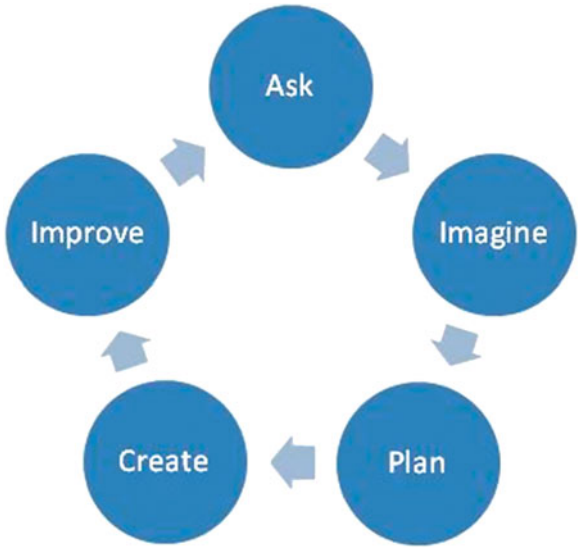


Figure 1. Design Cycle.[3] The Ask, Imagine, Plan, Create, and Improve process has been used successfully in the outreach efforts, *Engineering is Elementary*[2] and *TeachEngineering*.[3] Students are asked to use the steps of the engineering design process as they solve problems. The process is cyclical and may begin at and return to any step. Reprinted with permission from Reference 3. ©2010, TeachEngineering.

Given this potential impact, outreach programs should include carefully scripted messages and carefully designed measurements to determine whether the messages have had the desired impact.

In the following sections, we discuss the use of the design cycle, step-by-step, to support the planning and implementation of the K–12 outreach assessment. The initial questions are those used in *Engineering is Elementary* to teach primary students the design process.[2] These are then followed by the equivalent and expanded questions associated with designing assessment. The sidebar by Keith Bowman provides an example of the implementation of this process for assessing modeling self-efficacy in a K–12 materials outreach project.

Step 1: Ask

- What is the problem?
- What are the constraints?
- What have others done?[2]

This phase of the assessment process is often referred to in educational literature as defining the desired goals and objectives of an activity. Goals are broad-based statements of knowledge, which the activity is designed to support. Objectives are statements concerning the intermediate knowledge necessary to attain the goals. Outcomes are external displays that support that a given goal or objective has been reached. The pressing questions, rephrased for outreach, follow.

Who is the audience of the outreach? Who is the audience of the assessment results? What are the outreach goals? What are the possible outcomes of the proposed innovation/intervention? Which of the possible outcomes will be the focus of the assessment plan? Is the list of outcomes both manageable and actionable? Are the outcomes confounded? What are the constraints from regulations such as human subjects' research? What have others done who have asked similar questions?

The questions mentioned guide the outreach scientists through the process of defining what it is that they hope to achieve through the outreach activity. Much like research without a guiding hypothesis, outreach without purpose serves no cause. To define the desired goals and objectives, researchers first need to understand who their audience is and what they hope to achieve with that audience.

A potential significant difference between education research and materials research is the inclusion and measurement of humans. Since outreach deals with intervention or experimentation that involves humans, the research is called human subject research and, in the United States, is governed by both state and federal guidelines.[5] Other nations have similar requirements (See http://www.hhs.gov/ohrp/international/intlcompilation/hspcompilation-v20101130.pdf for a compilation of human subjects protection in 101 countries).[6] Outreach scientists need to be aware of human subject requirements within their country and their institution and carefully adhere to those requirements. In general, human subject requirements are designed to reduce any risk on the individual that could result from participating in a data collection process. Assessment plans may require approval from your Institutional Review Board (IRB), so understanding their requirements early in the planning process is imperative.

Step 2: Imagine

- What are some solutions?
- Brainstorm ideas.
- Choose the best one.[2]

The second step is the fun step—"Imagine." What are the potential assessment plans that could support this outreach effort? Typical questions asked during this phase include the following: What will provide the most persuasive evidence for your intended audience? How much information is manageable to collect? What has already been gathered by other sources or for other reasons, and can you get access? Do the data to be collected map directly to the objectives identified as important? Are these the most direct measures? Are you reinventing the wheel—do methods or instruments, such as questionnaires, conceptual inventories, or surveys, already exist that others have created for the same objective? What type of program are you planning?

In order to respond to these questions, the outreach scientist needs to understand the different forms of assessment and the different methods that are available to support the various forms of assessment. Activities included within this step are searching the literature in order to develop a better understanding of the assessment process and to determine whether there are existing instruments that will support your assessment needs (see the

https://doi.org/10.1557/mrs.2011.62 Published online by Cambridge University Press

An example of assessing Materials Science Engineering activities in K–12

Keith Bowman (Purdue University)

Model Eliciting Activities (MEAs) are open-ended, real-world, client-driven, and process-oriented problems[21,22] that are part of an international movement of mathematical modeling initiatives in learning. MEAs in areas such as engineering design, nanotechnology, microstructure, or mechanics find ways to integrate material science-related activities into mathematics and/or science lessons.[23–25] An example of one MEA follows, along with the related outcomes.

Beginning at about 80°C (180°F), typical 20-oz polyethylene terephthalate soda bottles begin to shrink back toward the dimensions they possessed before the stretch blow-molding process inflated them to their current size. This simultaneously produces crystallization that prevents the loss of carbonation. The two containers in the figure show the effect of (carefully) shrinking a bottle in an oven.

The shrunken bottles can be used to teach mathematical concepts at the sixth grade level of how to predict one characteristic (volume) of an object from another (dimensions).[26,27] Student teams focus on how the linear dimensions change and how that is related to the volume change by plotting the bottle volumes measured using a liquid versus the diameter or the height of the bottle, and then developing a mathematical model of the relationship.

A potential cognitive domain learning outcome would be whether students can confirm a rule regarding the relationship between linear dimensions and volume. A potential affective domain outcome would be to examine student self-efficacy using an instruments such as the Engineering Modeling Self-Efficacy (EMSE) scale.[28] A potential behavioral domain outcome would be to ask teachers to indicate the percent of students who repeated the modeling methodology in solving future problems as an indication of transferability of the skills taught.



Typical 20-oz polyethylene terephthalate soda bottle (with labels removed) before (left) and after (right) shrinking in an oven at 80°C (180°F). The bottles are filled with a dilute solution for greater visibility. The bottle shrinks to its original dimensions before the stretch blow-molding process inflated it to the current size. Photograph by Xu Nie.

sidebar by Bowman). Most outreach scientists use this literature review to stimulate discussions among the research team and brainstorm an assessment direction.

There are two, and some are recently advocating for three, types of assessments. The two traditional forms are called formative, referring to within program assessment or assessments that collect data that can be used to make adjustments as the program is being implemented. Summative assessment refers to a comprehensive evaluation of whether the program objectives have been met. The summative assessment is designed to provide compelling evidence to the funders and potential funders that the outreach program has been effective.

Formative assessments are intermediate in that they take place as the program is being implemented, and they provide brief tests of the effectiveness of small components of the program. Formative assessments are designed to allow the outreach scientist to adapt their program as it is being implemented in

https://doi.org/10.1557/mrs.2011.62 Published online by Cambridge University Press

a manner that makes the program more effective toward the attainment of goals (as defined in the previous step, "Ask").

Summative assessment is usually over a longer period—a semester, a year, or an overall project. Summative assessment examines the final impact of the program and the extent to which a program has reached its stated goals. At the summative level, comparisons to other programs or initiatives can also be made with the purpose of providing a more persuasive result. As this discussion seeks to illustrate, formative and summative assessments should be designed to complement each other. Formative moves the program toward the desired goals, while summative measures the extent to which those goals have been reached.

Recently, educators have advocated for a third type of assessment called interim. This refers to a more comprehensive mid-term assessment than formative in that it may cross multiple classrooms, schools, or even districts.[7] At a minimum, in the development of an assessment plan, formative and summative assessments should be considered.

Once observable and measurable outcomes of the proposed outreach activity are determined, the next step is to decide which of the outcomes will be the focus of the assessment plan. This narrowing process returns us to the "Ask" phase—what are the goals or outcomes that we hope to achieve? Two to five goals/outcomes are typically considered attainable.

There are four categories of outcomes that capture the majority of outreach goals. Examples of several such outcomes with respect to a student activity are provided in the sidebar by Bowman.

1. Cognitive—What will the outreach participants know at the conclusion of the activity with respect to the subject area? Common methods for measuring cognitive knowledge are the use of concept inventories, written or oral exams, or project summaries. For more information on concept inventories in science, technology, engineering, and mathematics areas, see www.ciHUB.org.[8]
2. Affective—What are the participants' attitudes and beliefs with respect to the content area? Measurement in this domain can be completed through surveys or interviews and focus on theories such as self-efficacy.[9,10]
3. Behavioral—How are participants' behaviors impacted by the outreach activity? Measurement in this domain often involves observations or self-reporting.
4. Metacognitive—Is the participant aware and able to monitor their own learning within the content area? Measurement in this domain is often completed through self-reflection, self-planning, or surveying.[11]

Step 3: Plan

- Draw a diagram.
- Make lists of materials needed.[2]

Planning is hashing out the details of the assessment. Who is going to gather the data? How will the data be analyzed and processed? Will new instruments need to be written and validated? What is the criterion for distinguishing between those who have and have not met the outcomes? Is it a curricular change in a classroom? Is there an afterschool or weekend activity offering? Are these criteria persuasive or informative? Will these criteria lead to adoption or change? Are quantitative or qualitative methods best suited to address the individual outcomes, or is a combined or mixed-methods approach preferred?

Many assessments within educational research are "qualitative" in nature. Qualitative refers to research methods that result in thick, rich data, such as interviews and open-ended surveys. Qualitative data lend themselves to the investigation of new situations or situations in which there is little prior data. Outreach activities are typically unique in that they have a specific focus on content or activities in which there is little or no prior research. Qualitative methods are often an appropriate format for the collection of formative data.

Quantitative methods result in numerical analyses or numbers. Scientists and engineers are often more familiar, and thus more comfortable, with these forms of assessment. The "gold standard" in educational research is the fully randomized experimental design with both treatment and control. However, this can rarely be achieved in educational research; randomly selecting students to not receive a promising treatment is inappropriate from a learning standpoint. The quasi-experimental designs, non-random experiments that seek to control or explain the impact of external factors, and triangulation, or multiple measures of the same outcome, have been used to address the lack of random design. Quantitative and quasi-experimental designs are often used to support summative assessments.[12] For those interested in pursuing a further understanding of the different types of appropriate assessment designs, the Coalition for Evidence Based Policy offers further insight into how to match an assessment method to the goals and resources of the program.[13]

Another consideration in the design of an assessment process is whether the data that are needed are direct or indirect. For example, due to concerns regarding the use of human subjects in the United States and many other countries, it is often easier to gather data through teachers about students. Students are usually minors, requiring consent from their parents or guardians. Teachers can provide their own consent, greatly reducing the demands of the human subject process. Many outreach scientists design their programs to collect data from teachers, thus reducing the human subject demand. However, this should only be done when compelling evidence can be collected from the alternative group with respect to the original goal.

Yet another consideration in the planning process at the K–12 level is alignment with standards. Many countries have national standards that guide the pre-college educational system, such as the United Kingdom's National Curriculum or the Australian Qualification Network.[14,15] In the United States, each state and often each district has its own set of standards. Goals that can be mapped to pre-college standards are more likely to receive classroom support.

https://doi.org/10.1557/mrs.2011.62 Published online by Cambridge University Press

Step 4: Create

- Follow your plan and create it.
- Test it.[2]

The "create" phase consists of developing the final plan, implementing it, and possibly pilot testing. This step in the assessment process is directly aligned with formative assessment, or assessments designed to provide feedback for the improvement of a plan. This step includes both judging whether the goals have been met as well as whether adjustments should be made for improvement. Piloting is often included before a full scale roll out, ensuring there are minimum glitches prior to full implementation. Writing an early version of the report that summarizes the results of these efforts helps guide the understanding of the data analysis and reporting issues that may arise and ensures that the appropriate data are collected. The results of this phase may send the outreach scientist back to any of the previous phases to fine tune or totally rework previous steps.

Step 5: Improve

- Talk about what works, what does not, and what could work better.
- Modify your design to make it better.
- Test it.[2]

A major consideration of this step is whether the plan will be used again. Is it the intention of the outreach scientist to implement the outreach activity a second or a third time? If yes, then information should be gathered and analyzed for improvement of the program. If not, then the outreach scientist may wish to report the results to the broader community—What worked? What did not work? By reflecting on the plan and refining this plan based on feedback from other team members or members of your audience, your assessment plan can be validated and improved. This step can also send you back to any previous step to modify and improve the activity and its assessment.

Potential pitfalls

Following the cycle for assessment just described does not guarantee success. Here, we list common pitfalls that, if considered upfront, can be avoided.

- Over-dependence on self-reporting can result in less persuasive results. Surveying subjects tends to be easier than behavioral observations or cognitive measurements, therefore, many researchers tend to rely on self-report instruments. The purpose of assessment is not to collect the data that are easiest to acquire but rather to collect the data that are most compelling with respect to the intended outcomes.
- Collecting too many data without a plan for how it will be used. The purpose of assessment is not to collect an overwhelming amount of data, but rather to collect persuasive and compelling data. Over-assessment can also result in fatigue in your subjects.
- Too many objectives. The principle of keeping it simple is usually the correct approach for outreach scientists, subjects, and reporting audiences.
- The "not-invented-here" mindset can lead to time and effort being spent when an instrument/technique/survey already exists. The many instruments that already exist should be researched and considered for use before a new instrument is created.[8,16] See, for example, the sidebar by Krause and Purzer.
- The opposite can also be true. This is the "should-have-been-invented-here-because-others-don't-really-relate" result. Make sure the objective of previously created and tried assessments are understood, and make sure they apply.
- Misinterpretation of cause and effect in assessment. In most outreach settings, the results are correlational in nature, since observational or quasi-experimental techniques are the ones most used. Only truly randomized designed experiments can support cause and effect interpretations.
- Attend to the human subjects' requirements of your country. Failing to address human subjects' concerns can be against the law.

Using existing instruments

Steven Krause (Purdue University) and Senay Purzer (Arizona State University)

Depending on the K–12 outreach objectives, an existing instrument may be the best choice. For example, for outreach activities that are designed to improve teachers' perspectives and familiarity with respect to engineering, the design, engineering, and technology (DET) instrument may be appropriate.[17] This instrument has been found to be valid and reliable in K–12 settings in the United States. Since this tool focuses on teacher attitudes (affective measurement), we refer to this tool as being part of the affective domain.

An existing cognitive instrument in materials education is the Materials Concept Inventory.[18] Concept inventories have been expanding in use since the original concept inventory, the Force Concept Inventory,[19,20] was developed in physics. These instruments are typically multiple choice instruments where each of the distracters in the potential answers highlights a misconception or misunderstanding.

https://doi.org/10.1557/mrs.2011.62 Published online by Cambridge University Press

Summary

An engineering design cycle can be used to guide the development and implementation of assessment of K–12 outreach activities. Given the similarities of the engineering design cycle to the scientific method, this framework may also benefit scientists who seek to develop an assessment of an outreach program. The process begins with clear statements of goals, objectives, and outcomes. This is followed by the development and implementation of the data collection plan, which is carefully aligned with the given objectives. Throughout this process, formative feedback is acquired, potentially leading to the necessity of returning to earlier steps to incorporate improvements.

References

1. L. Katehi, G. Pearson, M. Feder, Eds., *Engineering in K-12 Education: Understanding the Status and Improving the Prospects. Committee on K-12 Engineering Education* (The National Academies Press, Washington, DC, 2009), pp. 27.
2. *Engineering is Elementary*, Museum of Science, Boston; www.mos.org/eie/index.php(accessed July 5, 2010).
3. *TeachEngineering*; www.teachengineering.com/index.php (accessed July 5, 2010).
4. D.P. Giddens, *Changing the Conversation: Messages for Improving Public Understanding of Engineering. Committee on Public Understanding of Engineering Messages* (The National Academies Press, Washington, DC, 2008).
5. Title 45 Code of Federal Regulations Part 46.102(f); www.hhs.gov/ohrp/humansubjects/guidance/45cfr46.htm#46.102 (accessed July 5, 2010).
6. International Compilation of Human Research Protections; http://www.hhs.gov/ohrp/international/intlcompilation/hspcompilation-v20101130.pdf (accessed February 17, 2011).
7. M. Perie, S. Marion, B. Gong, *Educ. Meas.: Issues Pract.* **28** (3), 5 (2009).
8. Concept Inventory Hub; www.ciHUB.org (accessed September 30, 2010).
9. A. Bandura, "Self-Efficacy: The Exercise of Control" (W.H. Freeman and Company, New York, 1997).
10. *Information on Self-Efficacy*; www.des.emory.edu/mfp/self-efficacy.html (accessed September 30, 2010).
11. J. Flavell, *Am. Psychol.* **34** (10), 906 (1979).
12. M. Gall, J. Gall, W. Borg, *Educational Research: An Introduction, 8th Edition* (Pearson, Boston, MA, 2007).
13. Coalition for Evidence Based Policy; http://coalition4evidence.org/wordpress/ (accessed February 17, 2011).
14. National Curriculum; http://curriculum.qca.org.uk (accessed September 30, 2010).
15. Australian Qualifications *Framework*; www.aqf.edu.au (accessed September 30, 2010).
16. Resources—Other Assessment Web sites; www.flaguide.org/resource/websites.php (accessed September 30, 2010).
17. S. Yasar, D. Baker, S. Robinson-Kurpius, S. Krause, C. Roberts, *J. Eng. Educ.* 205 (July 2006).
18. S. Krause, J. Decker, J. Niska, T. Alford, R. Griffin, Annu. Conf. Expo. *(Am. Soc. Eng. Educ.)* 732 (2003).
19. D. Hestenese, I. Halloun, *Phys. Teach.* **33**, 502 (1992).
20. D. Hestenes, M. Wells, G. Swackhamer, *Phys. Teach.* **30** (March), 141 (1992).
21. G. Kaiser, B. Sriraman, *Int. J. Math. Educ.* **38**(3), 302 (2006).
22. H. Diefes-Dux, T. Moore, J. Zawojewski, P. Imbrie, D. Follman, *Conf. Proc. Front. Educ.* (2004).
23. J. Zawojewski, H. Diefes-Dux, K. Bowman, Eds., *Models and Modeling in Engineering Education: Designing Experiences for All Students* (Sense Publishers, Rotterdam/Taipei, 2008).
24. M. Lambert, H. Diefes-Dux, M. Beck, D. Duncan, E. Oware, R. Nemeth, *Conf. Proc. Front. Educ.* (2007).
25. H. Diefes-Dux, P. Imbrie, K. Haghighi, G. Lee, S. Wereley, P. Wankat, in *Proc. of the International Conference on Engineering Education and Research*, Olomouc, Czech Republic, 27–30 June 2004.
26. C. Smith, M. Wiser, C. Anderson, J. Krajcik, *Measurement* **14** (1–2), 1 (2006).
27. D. Thompson, M. Battista, S. Mayberry, C. Yeatts, J. Zawojewski, *Navigating through Problem Solving and Reasoning in Grade 6* (National Council of Teachers of Mathematics, Reston, VA, 2009).
28. T. Yildirim, M. Besterfield-Sacre, L. Shuman. *Annu. Conf. Expo. (Am. Soc. Eng. Educ.)* (2010). □



https://doi.org/10.1557/mrs.2011.62 Published online by Cambridge University Press