

Chapter XIII

A Case Study of Authentic Assessment

David A. Eubanks
Coker College, USA

ABSTRACT

This chapter describes Coker College's subjective performance assessment program to rate student thinking and communication skills. It uses a discussion of the epistemology of assessment to motivate an emphasis on direct observation by experts as the basis for "authentic" assessment for complex learning outcomes. Results from Coker College's experience are given and discussed in the context of this approach. The purpose of this chapter is to give a philosophical framework and practical methods that can help institutions assess liberal arts learning outcomes. Such assessments can provide information crucial to improving programs and pedagogy and form the basis for institutional effectiveness reports to stakeholders.

INTRODUCTION

Because this is a case study, we will be interested in implementations and details and uses of data. But these are more valuable if we know the rationale behind the assessment methods, so we begin with an examination of the measurement of learning outcomes. The questions we address are in turn theoretical and practical: What can we know, and how can we know it? How does work in practice?

Answers to these questions are explored in detail in the following sections:

- The Nature of Assessment
- Measurement and Reality
- Measurement and Probability
- Subjective Measurement
- Authentic Assessment in Practice
- Defining Outcomes
- Gathering Data

- Institutional uses of Assessment Data
- Program uses of Assessment Data
- FACS in the Classroom
- Reliability and Validity
- CONCLUSION
- Advantages
- Challenges
- Conclusion

It can be the case that an assessment program is put into place without asking these questions, but this may invite confusion when the results arrive.

There are many options open to an institution wishing to assess general education or liberal arts skills like thinking and communication. Standardized testing and portfolio assessment are two examples at opposite ends of a spectrum of possibilities. In this chapter we will derive a kind of performance testing.

Performance testing includes such varied procedures as observing a foreign-language student having a conversation in the foreign language, requiring science students to conduct a real experiment, asking students to work together as a group and observing the interaction, and giving problems that have no answer or more than one correct answer and observing a student's approach. (Kaplan & Saccuzzo, 2004, p. 631)

The description above makes it clear that the observer is crucial to this form of assessment. At Coker College, course instructors use normal class work to find opportunities for observing performance. As we shall see, there is a trade-off between this hands-on approach and objectivity (and hence reliability), but the results will support this decision.

We will first seek an understanding of what measurement is, and in the process reject overly deterministic models like fill-in-the-bubble tests for the purpose of assessing complex outcomes. We are not alone in this: "The 'new' assessment currently being called for by the National Educa-

tion Association and others is performance testing." (Kaplan & Saccuzzo, 2004, p. 631).

True performance testing tries to maximize reliability by carefully controlling testing conditions. See (Harris, 2002, pg. 81) for a list of testing components including "[...] specification of the steps or actions required to perform the test, typically in the order in which they are to be performed." This will prove to be a bit too restrictive for our agenda of assessing educational outcomes that are not the results of a step-by-step process. We will use the term "authentic assessment" instead, although in the literature the two terms are sometimes used interchangeably. In our case "authentic assessment" will depend less on rigid specifications and more on subjective judgment by experts.

THE NATURE OF ASSESSMENT

Observation is crucial to assessment, and we begin there, with the idea of perception and its subjectivity. An example often used to illustrate the subjective nature of perception is the sense of color. People agree through consensus that there is such a thing as the color red, and the idea of redness is therefore useful in the communication of information. We cannot directly know how someone else perceives color, but a nearly universal agreement among observers creates meaning. The degree of this agreement varies, however, depending on what is being observed. Judgments about traits like thinking or communication skills will likely have less agreement among observers than if they were looking at a color swatch. This does not mean that thinking and communicating are immune to observation or that we cannot usefully assess them in an educational setting, but it does mean that the results will be more subject to debate. Therefore, the reliability of such assessments will necessarily be lower than if we had, say, weighed the students to find their masses.

16 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/case-study-authentic-assessment/19673

Related Content

Monitoring of Road Deicers in a Retention Pond

Rémi Suaire, Ivana Durickovic, Marie-Odile Simonnotand Mario Marchetti (2013). *International Journal of Measurement Technologies and Instrumentation Engineering* (pp. 39-47).

www.irma-international.org/article/monitoring-of-road-deicers-in-a-retention-pond/89753

Application of Statistical Quality Control in Testing Labs for Estimation of Machine Interference

Nitin K. Mandavgade, Santosh B. Jajuand Ramesh R. Lakhe (2019). *Strategic Applications of Measurement Technologies and Instrumentation* (pp. 77-98).

www.irma-international.org/chapter/application-of-statistical-quality-control-in-testing-labs-for-estimation-of-machine-interference/217765

Dynamic Characteristics of Tiny Ultrasonic Linear Actuators

Hermes Hernándezand Ramiro Velázquez (2011). *International Journal of Measurement Technologies and Instrumentation Engineering* (pp. 12-23).

www.irma-international.org/article/dynamic-characteristics-tiny-ultrasonic-linear/53358

An Architectural and Evaluative Review of Implicit and Explicit SIP Overload Handling

Marco Happenhofer, Joachim Fabini, Christoph Eggerand Michael Hirschbichler (2013). *Advanced Instrument Engineering: Measurement, Calibration, and Design* (pp. 257-273).

www.irma-international.org/chapter/architectural-evaluative-review-implicit-explicit/78184

Using Messicks Framework to Validate Assessment Tasks in Online Environments: A Course in Writing Effectively for UNHCR

Valerie Ruheand Bruno D. Zumbo (2006). *Online Assessment, Measurement and Evaluation: Emerging Practices* (pp. 203-226).

www.irma-international.org/chapter/using-messicks-framework-validate-assessment/27707