Discussion and Future Research

As we discussed at the beginning of this book, traditional assessments are often too simplified, abstract, and decontextualized to suit current and future education needs. We need new assessments that measure what students actually can *do* with the knowledge and skills obtained inside and outside of school. Digital games can provide meaningful assessment environments by providing students with problems that require the application of various competencies. We also presented an assessment methodology that enables us to develop tasks in digital games using the principles of ECD. These tasks are designed to elicit specific performance data, which are then statistically linked to our focal competencies.

The first and most important step of this research project will be the determination of the validity of our stealth assessments. We will also be examining any learning of conceptual physics that ensues from *Newton's Playground* gameplay. If, in fact, the stealth assessments accurately estimate the focal competencies relative to existing measures, and learning does occur after four hours of gameplay, then the next step is to examine scalability.

That is, what are the costs and benefits of recycling ECD-based models in different games to assess the same kinds of competencies? These issues are being studied in our research project described herein. If we find that our stealth assessment methodology yields valid and reliable information, and is scalable, we plan to make the process as well as models broadly available to the community so that the work will continue and grow.

The research can expand in a number of general directions. First, we (and/or others) can explore the development of stealth assessments for other competencies that have been shown to play crucial roles in academic (and life) success (e.g., communication skills, computational thinking, empathy, civic engagement, problem-solving skill, and teamwork). Second, we can look at the development of stealth assessments relating to content that is directly aligned with the common core standards (e.g., mathematics modeling, probability, or reading comprehension). Third, we can push the bounds of our stealth assessments relative to implementing the models in additional digital games as well as other digital learning environments to determine the range of environments that may employ the same competency and evidence models, for a scalable, cost-effective, and engaging solution to the assessment of complex competencies. And fourth, we can examine any added value of including exploratory, data-mining methods to stealth assessment's more theoretically driven approach relative to the quality of the assessment.

Regarding future research related to learning, stealth assessment has the potential to be quite useful for diagnostic purposes due to the fine-grained analysis of student behavior in situated contexts. In addition, real-time information about player competency states can be useful to support learning through hints and feedback as well as the dynamic matching of game difficulty level to player ability (e.g., providing more challenging problems for those with high levels of various skills). Regarding the example used in this report, the indicators linked to the agents of force and motion can serve as the basis for diagnoses. If a student created a lever that did not successfully solve a problem that could have been solved via a lever, for instance, the indicators would inform the most likely reason(s) why. That is, the lever may have failed given the wrong mass of an object that was used on one side of the lever, because the fulcrum was positioned inaccurately, and/or because the size/length of the lever was too short or too long. Those data (mass, position, and length) are calculated as part of the stealth assessment.

Specific future research in the area of stealth assessment includes working with middle-school teachers to embed *Newton's Playground* into the physical science curriculum. This will involve linking Newtonian physics formalizations (e.g., $F = ma$) to relevant *Newton's Playground* problems for instructional support. Teachers can also design their own levels in *Newton's Playground* to highlight physics concepts that could benefit from more hands-on experience and support. Additional scaling of the game includes adding more levels to game, especially interactions among Newton's laws of motion; creating more physics content, like principles of collision; examining predictive validity of the game relative to future science courses taken and grades received therein; using the indicators associated with the four agents of force and motion to infer misconceptions for diagnostic and support purposes; and expanding the platform of *Newton's Playground* from computer- to browser-based gameplay.

In the more distant future, we can foresee dynamic and unobtrusive assessments being used in classrooms as well as outside of school. The data from these assessments may be aggregated into rich and valid profiles of students, reducing (or removing) the need for the teach-stop-test model that has governed classroom instruction for too long. We can also imagine representations of "academic success" to go beyond letter grades. Just what does a C in algebra substantively mean?

We are excited that researchers are starting to use digital games for learning and assessment. We think stealth assessment is one way to maximize the positive impact that digital games can have on students.

This is a section of doi:10.7551/mitpress/9589.001.0001

Stealth Assessment

Measuring and Supporting Learning in Video Games

By: Valerie Shute, Matthew Ventura

Citation:

Stealth Assessment: Measuring and Supporting Learning in Video Games
By: Valerie Shute, Matthew Ventura
DOI: 10.7551/mitpress/9589.001.0001
ISBN (electronic): 9780262315210
Publisher: The MIT Press
Published: 2013

The MIT Press

MIT Press books may be purchased at special quantity discounts for business or sales promotional use. For information, please email special_sales@mitpress.mit.edu or write to Special Sales Department, The MIT Press, 55 Hayward Street, Cambridge, MA 02142.

This book was set in Stone Serif and Stone Sans by the MIT Press. Printed and bound in the United States of America.

Library of Congress Cataloging-in-Publication Data
Shute, Valerie J. (Valerie Jean), 1953– , author.
Stealth assessment : measuring and supporting learning in video games / Valerie Shute and Matthew Ventura.
pages cm. — (The John D. and Catherine T. MacArthur Foundation reports on digital media and learning)
Includes bibliographical references.
ISBN 978-0-262-51881-9 (pbk. : alk. paper)
1. Educational tests and measurements. 2. Video games. I. Ventura, Matthew, author. II. Title.
LB3051.S518 2013
371.26—dc23
2012038217

10 9 8 7 6 5 4 3 2 1