

Chapter 16

Automatic Item Generation

Mark Gierl

University of Alberta, Canada

Hollis Lai

University of Alberta, Canada

Xinxin Zhang

University of Alberta, Canada

ABSTRACT

Changes to the design and development of educational tests are resulting in the unprecedented demand for a large supply of content-specific test items. One way to address this growing demand is with automatic item generation. Automatic item generation is the process of using models to create test items with the aid of computer technology. The purpose of this chapter is to describe and illustrate a method for generating test items. The method is also illustrated using an example from the medical health sciences.

INTRODUCTION

As the importance of technology in society continues to increase, countries require skilled workers who can produce new ideas, make new products, and provide new services. The ability to create these ideas, products, and services will be determined by the effectiveness of our educational programs. Education provide students with the knowledge and skills required to think, reason, communicate, and collaborate in a world that is shaped by knowledge services, information, and communication technologies (e.g., Binkley, Erstad, Herman, Raizen, Ripley, Miller-Ricci, & Rumble, 2012; Darling-Hammond, 2014). Educational testing has an important role to play in helping students acquire these foundational skills. Educational tests, once developed almost exclusively to satisfy demands for accountability and outcomes-based summative testing, are now expected to provide teachers and students with timely, detailed, formative feedback to directly support teaching and learning. To meet these teaching and learning directives, formative principles are beginning to guide our educational testing practices. Formative principles can include any assessment-related activities that yield constant and specific feedback to modify teaching and improve learning, including administering tests more frequently (Black & Wiliam, 1998, 2010). But

DOI: 10.4018/978-1-5225-7365-4.ch016

when testing occurs frequently, more test items are required. These additional test items must be created efficiently and economically while maintaining a high standard of quality. Fortunately, this requirement for frequent and timely educational testing coincides with the dramatic changes occurring in instructional technology. Developers of local, national, and international educational tests are now implementing computerized tests at an extraordinary rate (Beller, 2013). Computerized testing offers many important benefits to support and promote key principles in formative assessment. Computers permit testing on-demand thereby allowing students to take the test at any time during instruction; items on computerized tests are scored immediately thereby providing students with instant feedback; computerized tests permit continuous administration thereby allowing students to have more choices about when they write their exams. In short, computers are helping infuse formative principles into our testing practices to support teaching and learning.

Despite these important benefits, the advent of computerized testing has also raised formidable challenges, particularly in the area of test item development. Educators must have access to large numbers of diverse, high-quality test items to implement computerized testing because items are continuously administered to students. Hundreds of items are needed to develop the test item banks necessary for computerized testing. Unfortunately, test items, as they are currently created, are time consuming and expensive to develop because each individual item is written, initially, by a content specialist and, then, reviewed, edited, and revised by groups of content specialists (Gierl & Lai, 2016a; Rudner, 2010). Hence, item development has been identified as one of the most important problems that must be solved before we can fully migrate to computerized testing because large numbers of high-quality, content-specific, test items are required (Haladyna & Rodriguez, 2013; Webb, Gibson, & Forkosh-Baruch, 2013).

One promising test item development method that may be used to address this challenge is with *automatic item generation* (AIG) (Gierl & Haladyna, 2013; Irvine & Kyllonen, 2002). AIG is a relatively new but rapidly evolving research area where cognitive and psychometric modeling practices guide the production of tests that include items generated with the aid of computer technology. Research on AIG has adopted different perspectives, including the use of natural language processing and rule-based artificial intelligence (e.g., Gütl, Lankmayr, Weinhofer, & Höfler, 2011; Moser, Gütl, & Lui, 2012), frame-semantic representations (e.g., Cubric & Toasic, 2010; Higgins, Futagi, & Deane, 2005), schema theory (e.g., Singley & Bennett, 2002), and semantic web-rule language (Zoumpatianos, Papasalouros, & Kotis, 2011). The purpose of this chapter is to describe and illustrate the most practical method for generating test items, which is template based. By template-based AIG, we mean methods that draw on *item models* to guide the generative process. Gierl and Lai (2013, 2016a, 2016b) developed a three-step process for template-based AIG. In step 1, content specialists create a cognitive model for AIG. A cognitive model is a representation that highlights the knowledge, skills, and content required to generate new test items. In step 2, an item model is developed to specify where the cognitive model content is placed in each generated item. An item model is a template that highlights the variables in a test item that can be manipulated to produce new items. In step 3, computer algorithms place the cognitive content into the item model. With this process, hundreds of items can be created from a single item model. The purpose of this chapter is to describe how AIG can be used to generate test items using the selected-response (i.e., multiple-choice) format. To ensure our description is both concrete and practical, we illustrate template-based item generation using an example from the complex problem-solving domain of the medical health sciences. The chapter is concluded by describing two directions for future research.

10 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/automatic-item-generation/212812

Related Content

Public Policy Reforms: A Scholarly Perspective on Education 5.0 Primary and Secondary Education in Zimbabwe

Cleophas Gwakwara and Eric Blanco Niyitunga (2024). *International Journal of Technology-Enhanced Education* (pp. 1-18).

www.irma-international.org/article/public-policy-reforms/338364

Leadership in Global Open, Online, and Distance Learning

Ebba Ossiannilsson (2017). *Handbook of Research on Transformative Digital Content and Learning Technologies* (pp. 345-373).

www.irma-international.org/chapter/leadership-in-global-open-online-and-distance-learning/174061

Investigating the Effects of Gamification and Ludicization on Learning Achievement and Motivation: An Empirical Study Employing Kahoot! and Habitica

Qi Zhang (2023). *International Journal of Technology-Enhanced Education* (pp. 1-19).

www.irma-international.org/article/investigating-the-effects-of-gamification-and-ludicization-on-learning-achievement-and-motivation/326127

The Relationship of Classroom Behavior and Income Inequality to Literacy in Early Childhood

Ross Glen Chandler Nunamaker and William Arthur Mosier (2023). *Research Anthology on Early Childhood Development and School Transition in the Digital Era* (pp. 698-726).

www.irma-international.org/chapter/the-relationship-of-classroom-behavior-and-income-inequality-to-literacy-in-early-childhood/315706

Developing Education in the Digital Paradigm

Irena V. Robert (2023). *The Impact of Digitalization in a Changing Educational Environment* (pp. 70-78).

www.irma-international.org/chapter/developing-education-in-the-digital-paradigm/330882