Chapter 6
Computer-Adaptive Sequential Testing

Richard M. Luecht & Ronald J. Nungester
University of North Carolina at Greensboro, USA
National Board of Medical Examiners, USA

1. Introduction

This chapter describes a framework for the large scale production and administration of computerized tests called *computer-adaptive sequential testing* or CAST (Luecht, Nungester & Hadadi, 1996; Luecht, 1997; Luecht & Nungester, 1998). CAST integrates test design, test assembly, test administration, and data management components in a comprehensive manner intended to support the mass production of secure, high quality, parallel test forms over time. The framework is a modular approach to testing which makes use of modern psychometric and computer technologies.

CAST was originally conceived as a test design methodology for developing computer-adaptive[1] versions of the United States Medical Licensing Examination™ (USMLE™) Steps (Federation of State Medical Boards and National Board of Medical Examiners, 1999). The USMLE Steps are high stakes examinations used to evaluate the medical knowledge of candidate physicians as part of the medical licensure process in the U.S.. Although the Steps are primarily mastery tests, used for making pass/fail licensure decisions, total test scores and discipline-based subscores are also reported. Therefore, an adaptive-testing component is also attractive to economize on test length, while maximizing score precision.

The CAST framework has been successfully used in two empirical computerized field trials for the USMLE Step 1 and 2 examinations (NBME, 1996, 1997; Case, Luecht and Swanson, 1998; Luecht, Nungester, Swanson and Hadadi, 1998; Swanson, Luecht, Gessaroli, and Nungester, 1998). In addition, CAST has undergone rather extensive simulation research with positive outcomes (Luecht, Nungester and Hadadi, 1996; Luecht and Nungester, 1998).

In terms of USMLE, what CAST offers is a comprehensive means to develop mastery tests, with adaptive capabilities. CAST also helps

[1] The computerized versions of the USMLE™ Steps were introduced in 1999 as non-adaptive tests. CAST versions are slated for implementation in the near future.

W.J. van der Linden and C.A.W. Glas (eds.),
Computerized Adaptive Testing: Theory and Practice, 117–128.

guarantee that every test form can satisfy the extensive and complex content requirements needed to represent the spectrum of medical knowledge covered by each of the USMLE Steps. Within that more general context, the CAST framework was recognized as a useful way to design many different types of high stakes computerized tests, where the quality assurance of individual test forms is essential.

The basic test design model for CAST is a multistage test, with adaptive testing capabilities at the level of subtests or testlets (e.g. Adema, 1990; Sheehan and Lewis, 1992; Luecht, Nungester and Hadadi, 1996). That basic design can be engineered to produce a specific type of adaptive test and then be used as a template for creating future test forms. As part of the template, CAST incorporates explicit targeting of measurement precision where it is most needed and simultaneously contends with potentially complex content requirements, item exposure issues, and other constraints on test construction. The CAST test forms can also be preconstructed, allowing test developers to assure quality before releasing the tests. This quality assurance capability is viewed as essential for testing programs like the USMLE Steps, where content validity and test form quality are mandatory for EVERY test form.

2. The Basic CAST Framework

On the surface, the CAST framework presents as a multistage test, with adaptive capabilities. In that respect, CAST is hardly a new concept (see, for example, Cronbach and Gleser, 1965; Lord, 1971, 1980). However, CAST formalizes a set of statistical targets and all of the other test specifications into a template that can be used in conjunction with automated test assembly to manufacture in large scale adaptive test forms with the desired parallel statistical and content characteristics. CAST also allows the test forms to be created before test administration takes place. This capability provides some important advantages for interventions in terms of quality assurance and heading off security risks due to overexposed materials.

Figure 1 displays a simple CAST configuration. A single instance of this configuration is referred to as a *panel.* When all of the statistical targets and categorical content constraints are added to the panel, it becomes a template for test assembly. Each instance of the CAST panel can be treated as an independent test administration entity that can be uniquely assigned to an examinee an used for tracking responses conducting data parity checks, etc..

The panel shown in Figure 1 is made up of seven *modules,* labeled 1M, 2E, 2M, 2H, 3E, 3M, and 3H (top to bottom and left to right).