Medical Science Educator
https://doi.org/10.1007/s40670-024-02006-y

MONOGRAPH

# Rapid Blueprinting: An Efficient Method for Designing Content of Assessments

Raja G. Subhiyah[1] · Amanda L. Clauser[1] · David F. Martin[2]

Accepted: 8 February 2024


## Abstract

**Problem** Many assessments in medical education involve measuring proficiency in a content area. Thus, proper content development (blueprinting) of tests in this field is of primary importance. Prior efforts to conduct content review as part of assessment development have been time- and resource-intensive, relying on practice analysis and then on linking methods. This monograph explores a "rapid, cost-effective" approach to blueprinting that allows efficient assessment development with rigor. Our investigation seeks to explore an efficient and effective alternate method for creating a content design (blueprint) for medical credentialing and evaluation examinations by focusing directly on assessment requirements.

**Approach** We employed a two-phase process to propose a *rapid blueprinting* method. Phase 1 involved a 1-day direct meeting of content experts/practitioners. Phase 2 involved a corroboration survey sent to a wider group of content experts/practitioners. The rapid blueprinting method was applied to developing eleven blueprints (five for medical specialty certification; five for health professions certification; and one for in-training assessment).

**Outcomes** The methods we used resulted in effective, well-balanced, operational examinations that successfully implemented the resulting blueprints in item writing assignments and test development. Assessments resulting from the use of the rapid blueprinting method also generated psychometrically sound inferences from the scores. For example, the assessments resulting from this methodology of test construction had KR-20 reliability coefficients ranging from .87 to .92.

**Next Steps** This approach leveraged the effectiveness and feasibility of the rapid blueprinting method and demonstrated successful examination designs (blueprints) that are cost- and time-effective. The rapid blueprinting method may be explored for further implementation in local assessment settings beyond medical credentialing examinations.

**Keywords** Assessment development · Blueprinting · Content validity · Test design

Broadly, a credentialing test (licensure, specialty certification) or assessment aims to ensure that a candidate has the knowledge and skills necessary to practice competently and safely [1, 2]. Thus, the knowledge assessed by the test is of primary concern because the inference made from the test score should reflect the adequacy of the candidate's practice-relevant knowledge and/or skills [3, 4]. Further, designing the scope and breadth of an assessment should reflect the appropriate content necessary to ensure that those who pass possess the essential knowledge and/or skills to move forward in the credentialing process.

It is generally accepted that practice analyses, followed by a linking or expert review activity, are the "gold standard" methods used in the medical field to determine the scope and depth of tested content for most certification and licensing tests [5–8]. Developing a content outline from data gathered from a practice analysis study is a traditional approach with strong historical and theoretical underpinnings. However, this approach is expensive and time-consuming. Challenges associated with such resource-intensive content review also apply to local settings where assessments are developed and administered at medical schools and residency programs. Moreover, this approach gathers data that, though potentially important to understanding practice, does not necessarily directly relate to the material that is relevant to effective practice and testable in the necessary format (e.g., multiple

✉ Raja G. Subhiyah
rsubhiyah@nbme.org

1 National Board of Medical Examiners, 3750 Market Street, Philadelphia, PA 19104, USA

2 American Board of Orthopaedic Surgery, 400 Silver Cedar Court, Chapel Hill, NC 27514, USA

choice questions) [9]. Finally, the questions asked of the content experts (e.g., what skills and knowledge are needed to perform your job) do not directly address the testing situation and have to be modified and extracted to address a test design in an additional step. Many of the required skills to do a job are not readily testable in an efficient or cost-effective format.

A test blueprint describes the document outlining test content that is relevant and testable. The *blueprint* is a part of a broader test specifications document used to describe the examination's format, purpose, administration conditions, scoring, and feedback among other things [1]. The blueprint organizes the content areas covered by the test and defines the proportion (weight) of each area in terms of items or tasks. In our view, a blueprint serves three purposes. First, it gives an educator an outline of important topics to be included in teaching, forming the basis for a curriculum. Second, it provides an outline of necessary knowledge and skills for students who are preparing for the examination and working towards credentialing. And third, it allows the examination to be constructed in a consistent form across different administrations. While blueprints are critical in test development, the literature in medical education is lacking in efficient approaches for creating them. In this document, we seek to begin addressing this issue.

In our endeavor to develop content for several new testing programs, we started looking for alternative methods that were effective and more cost- and time-efficient than practice analysis studies. We borrowed ideas from the field of psychology and education [7, 8, 10] to develop a direct method that was relatively inexpensive, efficient, and relied on the expertise of practitioners. Here, we describe our experience in developing and using this alternate methodology which we named: *rapid blueprinting*. The thrust for developing this system was motivated by our need for a method that is faster and less expensive than a practice analysis study and linking activity for creating a blueprint for a credentialing examination. The goal was to narrow the activities but still achieve this result.

## Approach

In the main, the rapid blueprinting method differs from the classic practice analysis in two major ways:

1. Focus: Instead of asking about skills needed to perform a job, and then extracting from that what needs to be included in a test, we directly ask the content experts, all of whom should be practitioners, what should be on the test to certify the practitioner? Restrict your list to what is readily measurable on the test, e.g., by multiple choice questions (MCQ).
2. Efficiency: Focusing directly on the testing process eliminates extra steps, time, and costs needed to transform the initial skills into practical, testable items.

The method was developed with two stages: development and confirmation. An effective development stage depends on selecting an appropriate group of subject matter experts. To obtain adequate diversity of experiences and training as well as manageability to facilitate communication and consensus, we recommend a panel of 7 to 15 content experts who are active practitioners in the targeted medical field (e.g., orthopedic spine surgeons, echocardiographers, cardiologists, health and wellness coaches). Representative sampling of these experts is essential to developing an acceptable test blueprint, as this group should represent the breadth of knowledge and experience in actually performing the tasks required for practice. Diversity in practice pattern, background, age, and geographic distribution are important in this step of the process.

The primary focus of this group is on the question of what needs to be tested in order to ensure that those passing the test have the knowledge and skills necessary for safe and effective practice. Here, we assume that panel members, being practitioners themselves, know what knowledge and/or skills are required and can separate important knowledge that can be tested from that unsuited to the format of the test (e.g., MCQ).

The following steps are then taken to complete the development stage of the method:

1. Prepare background materials and references relevant to the target field, the aims of the testing program, and the inferences desired from the scores. This could include curriculum outlines from training programs, relevant data for practice activities, professional development offerings, textbooks, review course outlines, and previous blueprints.
2. Facilitate a consensus building exercise with the subject matter expert panel:
    a. Review background materials, the objectives of the test, and desired inferences from the scores.
    b. Define the major content areas that should be included in the test in conjunction with the aims and inferences desired.
    c. Define the weight (percentage of items) of each major content area within the examination.
    d. Within each major area, define specific concepts that need to be tested. These specific concepts should be sufficiently detailed so each can be tested by one or more test items. In addition, concepts that should not be tested are defined.

e. Review and finalize the resulting blueprint, emphasizing areas of agreement and compromise within the expert group.

The resulting document is then a draft of the test blueprint that describes the content that should be tested and the proportion (relative weight) of each area within that content domain.

The second (confirmation) stage consists of sending this draft blueprint out to a broader group (10 to 50) of content expert practitioners for feedback and refinement. These confirmation experts are asked to review the blueprint draft and:

1. Make general comments on the content and the major content areas.
2. Suggest new weights for the major areas or agree with the ones in the draft.
3. Review the specific areas for perceived omissions and/or suggested deletions.

Results of the survey are consolidated and routed back to the original panel, who reviews this feedback and refines the draft accordingly. This final document is made publicly available to students and educators and serves as the blueprint for that examination.

## Outcomes

We developed eleven blueprints using this method. Five of the examinations were for medical specialty certification, five were for health professions certification, and one was for an in-training test. Each of these blueprints was used to develop content (test questions) and subsequent examinations that were administered and used to inform credentialing processes. Each of these examinations mirrored its blueprint in content and in the proportions thereof. This study was approved by the National Board of Medical Examiners Internal Review Board (IRB) Process and determined to be exempt, as non-human subject research.

All but one of the blueprints we developed were used to guide item writing assignments, develop item pool classification systems, assemble the tests, and define score reporting categories in operational accreditation programs. Each test, except the Health System Science Examination, was administered via one of three modes: on computers at Prometric© Centers, on computer via the web or by pencil and paper.

During the development meetings, the content experts were each asked to outline their own scheme for the blueprint. These were posted for discussion, and members discussed and debated the different ideas. After discussion and (occasionally hot) debate, the panel reached a consensus blueprint that was acceptable to all members. While the panelists initially disagreed on many points, all invariably agreed on the final document. Most programs reached an agreement within 4 h; two programs took more time than that to reach an agreement. All resulting blueprint drafts had to have the consensus of the panel.

Each examination underwent rigorous quality control procedures at all stages of development, administration, and scoring. We computed scores using a Rasch model system, and pass/fail decisions were made after a standard setting procedure (modified Angoff). The resulting decisions were implemented and applied.

We observed that the content properties of these tests were in keeping with industry standards. The psychometric characteristics of the tests were also found to be adequate and similar to other medical credentialing tests developed by using full practice analyses. For example, both sets of tests had an acceptable K-R20 reliability coefficient range of 0.87–0.92. The tests using rapid blueprinting were generally well-targeted and had mean percent correct scores that ranged from 68 to 80% with a standard deviation around 9–10. Fail rates in the certification programs ranged from 10 to 28%, but these were largely the results of content-based standard setting studies.

The administrators and governing boards controlling examinations developed by rapid blueprinting received positive feedback from training programs about the content of the examinations. While most programs were very positive about the blueprints, our records show that roughly 5% of program personnel had minor disagreements, mostly concerning the weights of some content areas. We do not have records of any major disagreement. Six of the examinations are currently in operational use. Four examinations were terminated after one or two successful administrations for reasons unrelated to blueprints or content. One examination is still under development.

In general, we found positive feedback from the committees who developed the blueprints and from the boards overseeing the examinations. Further, the comments received from the survey participants generally showed solid agreement with the draft sent for evaluation. Over all the programs, there were fewer than 5% of disagreement statements from the respondents to the surveys. Most disagreements were minor adjustments to the weights of content areas. In only one program, three important disagreements were resolved by discussions with the developing panel.

Facilitating additional ownership of the examination content by those practicing in the field has the added benefit of building faith in the examination program. By utilizing national organizations as sources for the participants in the process, significant gains were made in overall acceptance of the examinations as reasonable measures of knowledge in each field. In addition, the mechanics of the examination-building process were made known to influential individuals, which further enhanced the reputation of the examinations and the credentialing process.

Finally, we would like to give an idea of the monetary and time savings that can be achieved by using the rapid blueprinting method. We asked the different programs if they had obtained cost estimates for performing a practice analysis from a vendor. The estimates ranged from approximately $60,000 to $80,000 and required 4 to 6 months to complete. In contrast, all rapid blueprint procedures we performed cost less than $30,000 and took an average of 2 months to complete.

## Next Steps

The rapid blueprinting approach can serve as a potential solution for conducting content review and creating blueprints beyond credentialing and licensing purposes. At more local contexts, medical schools and residency programs (and consortiums of institutions) may leverage their existing resources to apply this technique for generating more rigorous blueprinting processes, allowing for more cost-effective and feasible assessment development. In addition, the context for rapid blueprinting may extend beyond written tests as explored in this study; it may be modified for the implementation in performance-based assessments, including objective structured clinical examination and other assessments that rely on a blueprinting process.

The method does have potentially meaningful limitations. While a single subject matter expert meeting has existed as a practice analysis method for a long time [8, 9], this method is subject to bias introduced by one or more panel members who share their opinions more forcefully than others, swaying the eventual blueprint to better represent the area(s) that they deem most important. This method attempts to correct for that potential bias through the invitation of a representative panel and the confirmation step wherein the blueprint draft is reviewed and edited by individuals who did not participate in the original meeting. This is imperfect, of course, as those in the profession with less perceived power may not feel comfortable expressing their concerns either by speaking up in the meeting or through a survey tool. However, we found that most individuals in the process took great pride in this process and offered meaningful, uninhibited input.

This method for developing a test blueprint should not be confused with a method for practice analysis or a principled test design approach. Whereas both focus in different ways on the knowledge and skills required for effective practice [6, 8], this method is solely focused on creating a test blueprint and not on a rich understanding of the role and responsibilities of a practitioner. This approach, then, is well suited to well-defined areas in the health professions such as an established medical specialty. As such, representing the breadth of practice in a relatively narrow area of specialization among a group of 7–10 is much more feasible than in a very broad area of practice. It is also possible, in a very specific area, to potentially poll representatives from all training programs in the confirmation stage. This leads to a blueprint that represents the practice area and breadth of testable knowledge thereof instead of a blueprint that represents the individuals selected to participate in the exercise.

As noted earlier, this method results in a high-level test blueprint representing the knowledge areas that are testable within a well-defined area of practice. A practice analysis would provide much broader results describing the breadth of practice and the knowledge, skills, and behaviors needed to be successful, safe, and effective within it. This activity is by its nature different from a curricular review or an alignment activity where assessment and curricular content are compared. One of the benefits of a blueprinting process is that, like practice analysis, it exists independent of any training program curricula or performance measures. For example, medical curricula may focus on anatomy; in practice, even within a relatively narrow specialty area, some practitioners may need to apply detailed knowledge of anatomy to their day-to-day work while others may not. A blueprinting activity brings together subject matter experts familiar with the breadth and scope of practice to arrive at a compromise position to represent the knowledge needed in a specialty area. In addition, it should be expected that the scope of many assessments used to assess fitness for specialty certification are not only assessing the curriculum of the specialists' training programs but instead add value by assessing professional skills, behaviors, and judgment honed through experience.

For the cases we utilized here, the efficiency in cost and time seems to have outweighed the limitations in scope of this method. We feel that this method for developing blueprints has a place in the practical implementation and development of credentialing examinations. The method would have advantages especially for small, well-defined, programs where cost and time efficiency are important.

## Declarations

**Ethical Approval** This study was approved by the National Board of Medical Examiners. No human subjects were involved in the study.

**Conflict of Interest** The authors declare no competing interests.

## References

1. American Educational Research Association, Association AP, Council N, on Measurement in Education (AERA, APA, & NCME). Standards for educational and psychological testing. Washington, DC: American Educational Research Association; 2014.

2. Raymond M, Neustel S. Determining the content of credentialing examinations. Handbook of Test Development; 2006. pp. 181–224.
3. Kane M. Model-based practice analysis and test specifications. Appl Measur Educ. 1997;10(1):5–18.
4. Perie M, Huff K. Determining content and cognitive demand for achievement tests. Handbook of Test Development; 2015: pp. 135–159.
5. Goldstein IL, Zedeck S, Schneider B. An exploration of the job analysis-content validity process. In: Schmitt N, Borman WC, editors. Personnel selection in organizations. San Francisco, CA: Jossey-Bass; 1993.
6. Knapp JE, Knapp LG. Practice analysis: building the foundation for validity. In: Impara JC, editor. Buros-Nebraska series on measurement and testing. Buros Institute of Mental Measurements: Lincoln, NE; 1995. p. 93–116.
7. Morgeson FP, Campion MA. A framework of potential sources of inaccuracy in job analysis. In: Wilson M, Harvey R, Alliger G, Bennett W, editors. The handbook of work analysis: the methods, systems, applications, and science of work measurement in organizations. New York: Psychology Press/Taylor and Francis Group; 2004.
8. Newman LS, Slaughter RC, Taranath SN. The selection and use of rating scales in task surveys: a review of current job analysis practice. Paper presented at the meeting of the National Council on Measurement in Education, Montreal, Canada; 1999.
9. Raymond MR. A practical guide to practice analysis for credentialing examinations. Educ Meas Issues Pract. 2002;21:25–37.
10. Hendrickson A, Ewing M, Kaliski P. Evidence-centered design: recommendations for implementation and practice. J Appl Test Technol. 2013;1(1):1–27.