

まえがき

『テスト作成ハンドブック (*Handbook of Test Development*)』は知識、技能、能力などを測るテストを作成するに当たって、体系的で包括的な情報源となることを目指して書かれている。この『ハンドブック』は21世紀に入ってからのテスト開発の現況を伝えるものである。歴史的に見て、学問としての教育・心理測定学の中で、テスト作成に関する議論や実際面には科学的な関心がわずかしか払われてこなかった。研究や著作の多くはテストの統計学的な議論の方に多くの焦点が合わされてきた。それでも、テストの開発と学力テストや能力テストにまつわる多くの議論や実際面は、どのテストプログラムにおいても中心的で注意深く考えるべき重要な役割を占めていることに変わりはない。

教育・心理測定学の分野はこの10年間で劇的に変化した。コンピュータを利用したテスト、コンピュータによる適応型テスト、項目応答理論、一般化可能性理論、妥当性についての統一的理解などは、多くのテストプログラムの中に見られる光景である。教育・心理測定学の理論や実践、あるいはテストの実施法にこうした大きな変化があるにしても、高い品質のテストを開発しようとする活動は依然として変わりなく続けられている。ただ、その多くは文書化されないままである。この『ハンドブック』はテストの作成に影響を与える問題を研究している作成者や研究者に、健全なテスト実践に関する情報を役に立つ形で用意しようとして作られたものである。

この『ハンドブック』の目的は包括的で、一貫した、そして学術的で、かつ実際的な議論を提供することである。そこには、作成者の意図するテスト得点の解釈が十分支持されるような妥当性の証拠を持つテストを作るにはどうしたらよいかといった議論と、そのための健全な実施に向けてのすべてが扱われている。学力テストや能力テストの開発現況に詳しい著名な学者や実務家がこの本の各章に力を割いてくださった。各章にはテストの作成に当たって十分に支持され、実際的な助言として信頼のできる有益な情報源が提供されている。

テスト作成の技と科学は、テスト機関で業務を通じて学習されるのが普通である(いわゆるOJT)。それはきちんとした教育測定理論に基づく標準的なやり方だとして、マスターから初心者に伝えられている。大学院プログラムでは教育測定のコースで教えられ、またテスト作成法についてのセミナーで教えられることもある。しかし、大学院の学生は実際のテスト作成についてはわずかな洞察しか得られないか、深い実習経験も経ずに教育測定の博士コースを終わることが多い。心理測定理論はテストの開発活動すべてに必要な基礎となるものであるが、理論と実際とのつながりは必ずしも明らかになっていない。この『ハンドブック』はそうしたテスト開発のすべての段階において、テスト開発者の包括的な指針となるものを提供している。そこには、テストの計画段階から始まって、テスト内容の定義、質問項目や問題作成、その制作、実施、採点、報告といったすべての過程が含まれる。

米国教育研究学会、米国心理学会、全米教育測定協議会によって編集された『教育・心理検査法のスタンダード (*Standards for Educational and Psychological Testing*)』(AERA, APA, & NCME, 1999)は、この本で議論されているすべてのテスト作成の背景になるものである。各章の著者は自分の議論を支えるものとして、しばしばこのスタンダードを参照していることが分かるであろう。ここで述べられるテスト作成の諸局面において、ある特定の妥当性の証拠について述べるときにはこの『テストスタンダード』について言及している。

(一部)