

教育科學研究期刊 第五十六卷第一期

2011 年，56 (1)，1-32

大型資料庫國小四年級自然科學學習成就評量 標準設定結果之效度評估

謝進昌

國家教育研究院
測驗及評量研究中心
助理研究員

林陳涌

國立臺灣師範大學
生命科學系
教授

謝名娟

國家教育研究院
測驗及評量研究中心
助理研究員

陳清溪

國家教育研究院
院長室
主任秘書

林世華

國立臺灣師範大學
教育心理與輔導學系
副教授

謝佩蓉

國家教育研究院
測驗及評量研究中心
博士後研究員

摘要

學習成就評量資料庫建置目的之一，在於描述學生於各學科的學習成就表現概況；而學習成就表現的描述，即牽涉到標準設定議題。隨著臺灣學生學習成就評量資料庫對於國小四年級自然科之評量架構調整，致使沿用的表現標準已多不符合實務運用，而興起重新修訂之舉。本研究主要目的在於評估 2009 年自然科學學習成就評量國小四年級標準設定結果之效度，以 Kane (1994, 2001) 的效度評估證據為依據，藉由效度的過程證據、內部證據以及外部證據，作為標準設定結果之效度證據來源。結果發現，從各項效度證據來看，國小四年級自然科標準設定可由標準設定成員的組成及其回應、設定結果的一致性，以及外推論於數學成就的效度外部證據得到支持。本研究最後提出數項建議，供未來標準設定研究者參考。

關鍵字：效度評估、書籤標定法、標準設定

通訊作者：謝佩蓉，E-mail: pjh@mail.naer.edu.tw

收稿日期：2010/09/29；修正日期：2011/01/23、2011/02/26；接受日期：2011/03/15。

Validation of the Standard Setting Procedure for a Large Scale 4th Grade Science Assessment

Jin-Chang Hsieh

Research Center for Testing and
Assessment,
National Academy for Educational
Research
Assistant Research Fellow

Chen-Yung Lin

Department of Life Science,
National Taiwan Normal University
Professor

Ming-Chuan Hsieh

Research Center for Testing and
Assessment,
National Academy for Educational
Research
Assistant Research Fellow

Chin-Hsi Chen

President Office,
National Academy for Educational
Research
Secretary General

Sieh-Hwa Lin

Department of Educational Psychology and
Counseling,
National Taiwan Normal University
Associate Professor

Pei-Jung Hsieh

Research Center for Testing and
Assessment,
National Academy for Educational
Research
Post-Doctoral Fellow

Abstract

One of the main purposes for constructing an academic achievement database is to understand the students' academic performance for each content area. Standard setting is the main issue in the academic achievement. Due to the adjustment of the test blueprint for Taiwan Assessment of Student Achievement 4th grade science, the previous performance level description is outmoded for practice. Therefore, modifying the standards is necessary. The purpose of this study is to validate a science standard setting procedure designed for 2009 4th grade science students. Based on Kane's (1994, 2001) framework for validation, this study collected procedural evidence, internal evidence, as well as external evidence using mathematics academic performance. Several suggestions are provided in the conclusion for further studies.

Keywords: validation, Bookmark method, standard setting