

第一章 緒論

第一節 研究動機與目的

回顧我國數十年的教育評量實務，例如高中職、五專及大學聯考，乃至民國九十年開始舉辦的國中學力測驗，這些考試決定大部份國中生進入高中職，及高中生進入某所大學的標準。學生努力了三年，就是為了畢其功於此一役——聯考，此對於考生造成很大的壓力。而其他不可抗拒的因素，如考試當時學生的身心狀況皆可能影響學生作答的成績。此種一年一次的大會考相當不符合人性。假如此種決定性的考試能改為一年多試，不僅能減少考生的壓力，且讓考生有更多的機會來測試自己的能力，此種考試才能更公平、更真實地測出考生的能力。

自從民國九十年起所實施的國民中學學生基本學力測驗，其目的是要評量學生在經過國中三年後所應具備的基礎的、核心的、重要的知識與能力，學生測驗所得的分數將取代過去的高中職及五專聯招的分數作為申請進入高中、高職或五專的依據。國民中學學生基本學力測驗之取材，是以學生學習及生活經驗為主(國民中學學生基本學力測驗推動工作委員會，民90年)，其進行方式不同於以往的考試形式，是採取事先命題，經過學科、測驗專家多次修題、審題，再進行預試後，經試題分析取得試題的參數及相關訊息，然後再將合乎標準的優良試題納入題庫之中。在入闈時根據事前公布的測驗目標，從題庫中選取試

題，組成正式題本進行測驗。其目的在於提昇測驗品質、降低測量的誤差，以增加測驗的信度、效度。國民中學學生基本學力測驗與過去的聯考有一很大的不同，是一年舉辦兩次，測驗分數是利用現代測驗理論 IRT 所建立的量尺分數，其測驗結果是客觀、公平的，且兩次測驗的結果是利用測驗等化技術得到的，彼此是可以互相比較的。因此，學生可以選擇報考兩次，或只選擇其中一次報考，然後選擇較優的成績作用申請入學用，以避免一年一試來減輕考生的壓力。面對如何將相同或不同的考生，於第一次學測或第二次學測的測驗結果，校準於同一量尺上作直接比較的問題，正是測驗等化所要解決的問題，也是一般參加國中基本學力測驗學生、家長及老師，所一致關心和重視的問題。

例如，美國 ETS 在全世界所辦的 TOFEL、GRE 和 GMAT 等考試有多試的複本測驗，可以在一年內實施多次測驗，不同受試者可接受不同複本測驗，這些複本測驗經過等化後，測驗的成績即可互相比較。

已有很多關於測驗等化方法的文章及書廣為發表及討論，其詳細內容擬於文獻探討中討論之。因利用試題反應測驗理論(IRT)所建立之等化測驗，已為現代測驗界所廣泛應用，而其中因方法不同所做的等化結果亦不盡相同，本研究擬利用 IRT 之 1-PL、3-PL 的模型，來探討國中學力測驗—數學科兩次測驗間的量尺與等化結果，再與目前的學力測驗等化結果做比較及討論。另外擬應用 IRT 真分數等化法，來探討其與學力測驗原始分數的差異。因此，本研究之主要目的有下列六項：

1. 探討民國 90 年第一次學測，數學科利用 1-PL 模式的校準後，所呈現的量尺分數與學力測驗的量尺分數改變情形。
2. 探討民國 90 年第二次學測，數學科利用 1-PL 模式的校準後，所呈現的量尺分數與學力測驗的量尺分數改變情形。
3. 探討民國 90 年第一次學測，數學科利用 3-PL 模式的校準後，所呈現的量尺分數與學力測驗的量尺分數改變情形。
4. 探討民國 90 年第二次學測，數學科利用 3-PL 模式的校準後，所呈現的量尺分數與學力測驗的量尺分數改變情形。
5. 探討民國 90 年第一次學測，數學科利用 1-PL 與 3PL 模式校準後之題目難度估計值，應用 IRT 真分數法所呈現的真分數分佈情形。
6. 探討民國 90 年第二次學測，數學科利用 1-PL 與 3PL 模式校準後之題目難度估計值，應用 IRT 真分數法所呈現的真分數分佈情形。

第二節 研究問題與假設

一、研究問題

根據上述之研究動機與目的，本研究擬探討的問題如下：

1. 民國 90 年第一次學測，數學科利用 1-PL 模式的校準後，所呈現的量尺分數與學力測驗的量尺分數改變率為多少？