

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？

2. 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率為多少？
3. 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率為多少？
4. 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率為多少？
5. 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為何？
6. 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為何？

二、研究假設

為回答上述之研究問題，本研究之研究假設如下：

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具有改變。
2. 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具有改變。
3. 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具有改變。

4. 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具有改變。
5. 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的校準後，所呈現的真分數與學力測驗的原始分數不同。
6. 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的校準後，所呈現的真分數與學力測驗的原始分數不同。

第三節 名詞解釋

茲將本研究所涉及之重要名詞解釋如下：

一、測驗等化

測驗等化是利用統計過程，來調整不同測驗間的分數，使得不同測驗間的分數可以彼此交換，但這些不同測驗的內容與難度是類似的(Kolen & Brennan, 1995)。

等化可分為水平等化與垂直等化，水平測驗是指兩個或兩個以上測量某一特質的測驗，其考生的能力分配與試題難度相似，其不同測驗間原始分數互相轉換的過程。如國中基本學力測驗，可在一年內實施兩次，這兩次測驗經過等化的程序後，其測驗成績是可以互相比較。垂直測驗是指兩個或兩個以上測量某一特質的測驗，其考生的能力分配與試題難度皆不同，其原始分數互相轉換

的過程，即測驗內容程度高低的等化。如美國加州成就測驗。

二、國中基本學力測驗

國中基本學力測驗發展概述

國中基本學力測驗之緣起係伴隨著高中多元入學理念而來，自民國八十三年第七次全國教育會議後，教育部於八十四年提出「中華民國教育報告書」，即揭示高中教育遠景之一為建立多元入學管道，導引國中教學正常發展。至八十五年十二月行政院教育改革審議委員會提出教育改革總諮議報告書中，亦建議推動多元入學制度。教育部遂於八十七年七月正式公佈高級中學多元入學方案，同年九月公布高級職業學校多元入學方案，同時委託臺灣師大心測中心，以三年時間研究與發展基本學力測驗，並自九十學年度起正式實施(教育部，2003；鄭同僚、張原禎，2001)。

國民中學學生基本學力測驗的緣起及施行理念

民國八十五年教改總諮議報告書中，對入學政策的主張是「推動多元入學制度」，對評量的積極主張則是「基礎學科能力的評量」與「其他項目表現的評量」。國中基本學力測驗就是在前述的前提下所形成的，其目的在評量國民中學學生能力表現及其發展潛能。其主要目標則是在維持制度公平的前提下，消除入學考試對於國民中學教育的不利影響，進而充分發展學生的潛能。因此基本學力測驗的命題方向，偏重在對學生未來的學習與生活有所幫助之基礎的、核

心的、重要的知識與能力；命題概念是完整的、周延的，而非偏狹的、殘缺的。至於測驗難度則是以國中學生平均能力為主，因此過去以「過度學習」和「機械式練習」來提升考試分數的學習方法，將無法在本測驗中獲得預期的效果。希望在這樣的制度與測驗方式之下，能夠還給學生一個理想、正常的學習空間。

國中基本學力測驗的特色(國民中學學生基本學力測驗推動工作委員會，2003)

1. 標準化的

國中基本學力測驗是結合測驗專業與各學科領域所發展出來的標準化測驗。所謂標準化測驗可以從兩個方向來看：第一個部分是測驗發展程序的標準化，這是指測驗發展的過程完全依據標準的測驗編製程序來建立，從各學科領域雙向細目表的擬定、試題的編製與取樣、預試、分析試題以及信、效度的建立，完全依照標準化的測驗編製程序來進行，其主要目的是使測驗題目能有較好的品質。第二部分是施測與計分的標準化，包括固定而且明確的施測程序與指導語、計分標準、分數的計算方式等，其目的是降低測量的誤差，增加測驗分數的客觀性。

2. 可比較的

基本學力測驗使用測驗等化技術所提供的測驗分數，對所有的考生來說，即使不同考生參加不同次別的基本學力測驗，其結果仍然是客觀、公平而且可以互相比較的。

3. 一年多試、一試多用的

基本學力測驗一年舉辦兩次，學生可依自己的意願選擇報考兩次或報考其中一次，測驗分數亦可擇優使用，避免一試定終身之憾。測驗分數可供高中、高職、五專多元入學參採之用。

4. 能力導向的

基本學力測驗以能力為導向，旨在評量學生的基本學力，評量的重點是學生能夠帶著走的能力。

5. 對教學有良性影響的

基本學力測驗的命題理念與實施方式，能夠對國中教育發揮正面的影響，讓老師正常教學、學生快樂學習。