

國立臺中師範學院進修暨推廣部九十四年度
教育測驗統計研究所國民小學教師在職進修教學碩士學位班（暑期班）招生考試試題

測驗與評量

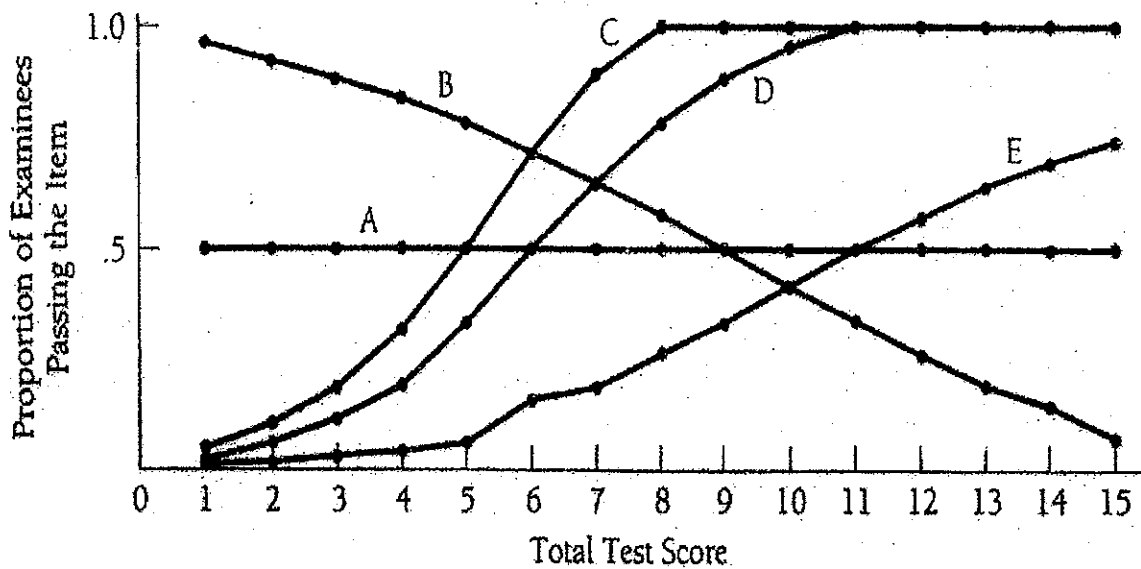
一、選擇題(單選)，每題 2.5 分，20 題共 50 分

1. 以下何者是依據完形心理學理論的「封閉」原理所設計的題型？
 - (1) 配合題
 - (2) 選擇題
 - (3) 克漏字題
 - (4) 填充題
2. 關於對同一學科下多種評量方式分數進行加權組合的原則，下列何者錯誤？
 - (1) 客觀計分測驗的加權量應該高於主觀計分測驗的加權量
 - (2) 若使用經過效度驗證的標準化測驗，應占有較大的比重
 - (3) 期中與期末考的加權量應該高於隨堂考或單元測驗
 - (4) 龐大的小組作業應該比個人作業有較高的加權量
3. 在認知領域的教育目標中，下列何者可以再細分成轉譯、解釋與外推三種？
 - (1) 分析
 - (2) 綜合
 - (3) 理解
 - (4) 應用
4. 閱卷者先評閱全部試卷的某一題，待該題全部都評閱之後，再評閱下一題，如此可以避免下列哪一現象？
 - (1) 月暈效應
 - (2) 聚斂效度
 - (3) 重疊效應
 - (4) 複合效應
5. 使用複本法來估計測驗信度時，它的主要誤差來源是什麼？
 - (1) 內容取樣
 - (2) 評分者誤差
 - (3) 機遇誤差
 - (4) 時間取樣

6. 設計「雙向細目表」作為測驗的命題藍圖，主要可以提高測驗的：
- (1) 表面效度
 - (2) 預測效度
 - (3) 內容效度
 - (4) 建構效度
7. CEEB 分數又稱為什麼？
- (1) T 分數
 - (2) ETS 分數
 - (3) GRE 分數
 - (4) AGCT 分數
8. 在標準參照測驗的信度研究中，通常是以下列何者作為其信度指標？
- (1) K-R 公式
 - (2) alpha 係數
 - (3) 肯德爾和諧係數
 - (4) 百分比一致性
9. 下列何種分數是經過常態化轉換的衍生分數？
- (1) AGCT 分數
 - (2) T 分數
 - (3) 離差智商
 - (4) 標準九
10. 要解釋同一人在兩個測驗之間，其標準分數的差異時，不需要符合哪個條件？
- (1) 這兩個標準分數是採用同一種量尺
 - (2) 建立常模時，兩測驗使用相同的常模樣本
 - (3) 兩個分數之間差異要達統計上的顯著水準
 - (4) 這兩個測驗的試題數要相同
11. 若將某一群參加考試的受試者的分數全部都加一個常數，則以下的測驗指標何者會產生改變？
- (1) 信度
 - (2) 變異數
 - (3) 鑑別度
 - (4) 平均數
 - (5) 結構效度
12. 實作評量(performance assessment)的批改，常需要兩位或兩位以上的評分者，若需要對這類評量方式的信度進行檢測，最好是利用以下何種指標？
- (1) Cronbach α
 - (2) 折半信度
 - (3) 庫李公式
 - (4) 概化理論
 - (5) 鑑別度指數

13. 結構效度(construct validity)的檢測方法中，如因素分析(factor analysis)，研究者是利用試題間的何種性質，來達到對該測驗效度的檢視？
- (1) 不變性(unchanged)
 - (2) 可信性(trustful)
 - (3) 有效性(efficient)
 - (4) 相依性(relationship)
 - (5) 廣泛性(generalizable)
14. 若大學入學考試主要目的是在預測受試者在大學就讀時的表現，並檢測受試者能否適應未來大學各個學系中的要求，則大學入學考試應依以下何種測驗的設計原理，比較適合？
- (1) 興趣測驗(interest test)
 - (2) 人格測驗(personality test)
 - (3) 態度測驗(attitude test)
 - (4) 性向測驗(apptitude test)
 - (5) 成就測驗(achievement test)
15. 以下數個測驗中，其測驗使用者在對所得測驗分數做解釋時，所必須依據的分數解釋架構將比較相近？
- (1) 全民英檢、大學聯考
 - (2) 全民英檢、智力測驗
 - (3) 智力測驗、大學聯考
 - (4) 智力測驗、期中考試
 - (5) 全民英檢、期末考試
16. 在古典測驗理論的架構下，某測驗施測者將對兩群受試者實施同一套測驗，若其他的條件皆相同，但A群受試者間的真實能力的變異程度(variation)較B群高，則對此兩群受試者的測驗結果的信度預估描述，相對之下，下列何者比較正確？
- (1) A 測驗結果的信度較低
 - (2) B 測驗結果的信度較低
 - (3) 兩測驗結果的信度會相等
 - (4) 兩測驗結果的信度沒有必然關係
 - (5) 視效度大小的情況而定
17. 以下何種評量方式，不適合被納入多元化評量方法中？
- (1) 紙筆評量(paper & pencil test)
 - (2) 實作評量(performance assessment)
 - (3) 檔案評量(portfolio assessment)
 - (4) 闖關評量
 - (5) 以上均可納入多元化評量

18. 下圖摘錄自 Allen & Yen (1979)，請問圖中五個試題的鑑別度，何者較優良？



- (1) 第 A 題
- (2) 第 B 題
- (3) 第 C 題
- (4) 第 D 題
- (5) 第 E 題

19. 學生 A 在某智力測驗甲式中所獲得的分數為 130 分，此甲式智力測驗的常模資料為平均數 100，標準差 15，該測驗的乙式常模資料為平均數 105，標準差 20，請問該學生在乙式常模中，原始分數應該較接近下列何者？

- (1) 140
- (2) 130
- (3) 120
- (4) 110
- (5) 100

20. 以下是對適性化(adaptive)測驗的相關描述，下列組合中，相對之下何者比較正確？

- 甲：魏氏(Wechsler)智力測驗是一種適性化的測驗
- 乙：適性化測驗的測驗題數並不一定
- 丙：對受試者而言，適性化測驗比較省時
- 丁：瑞文氏(Raven)智力測驗是一種適性化的測驗

- (1) 甲、乙
- (2) 乙、丙
- (3) 甲、乙、丙
- (4) 乙、丙、丁
- (5) 甲、乙、丙、丁

二、閱讀選擇題，每題 2.5 分，20 題共 50 分

第一部份 (10 題)

在試題反應理論 (Item Response Theory, 簡稱 IRT) 中，以答對機率 $P_i(\theta_i)$ 來表示受試者能力 θ_i 以及試題 i 之特性間的關係。依據採用試題參數個數的不同，可分為單參數、二參數以及三參數 Logistic 模式，分別表示如下：

單參數 Logistic 模式 (1PL)

$$P_i(\theta) = \frac{\exp(\theta - b_i)}{1 + \exp(\theta - b_i)}$$

二參數 Logistic 模式 (2PL)

$$P_i(\theta) = \frac{\exp(a_i(\theta - b_i))}{1 + \exp(a_i(\theta - b_i))}$$

三參數 Logistic 模式 (3PL)

$$P_i(\theta) = c_i + (1 - c_i) \cdot \frac{\exp(a_i(\theta - b_i))}{1 + \exp(a_i(\theta - b_i))}$$

四參數 logistic 模式 (4PL)

$$P_i(\theta) = c_i + (1 - c_i - d_i) \cdot \frac{\exp(a_i(\theta - b_i))}{1 + \exp(a_i(\theta - b_i))}$$

其中， a_i, b_i, c_i, d_i 分別表第 i 題的鑑別度參數、難度參數、猜測度參數以及天井參數。

上述的模式均適用在二元計分 (非對即錯) 的資料中，但也有適用於多元計分的模式，在此舉出以下兩種模式：

k 點計分 (0 到 $k-1$ 分) 評分量尺模式 (Rating Scale Model, 簡稱 RSM)

$$P_{nix} = \frac{\exp\left(\sum_{j=0}^x (\theta_n - (b_i + \tau_j))\right)}{1 + \sum_{r=1}^{k-1} \exp\left(\sum_{j=1}^r (\theta_n - (b_i + \tau_j))\right)}$$

其中， τ 是指兩個選項間的閾值 (Threshold)。

請根據以上敘述回答下列各題：

21. 在 3PL 模式中，各試題的試題特徵曲線 (Item Characteristic Curve, ICC) 下列何者是相同的？

- (1) 能力值為負無限大者的答對機率
- (2) ICC 在 θ 量尺上的位置
- (3) ICC 的切線斜率
- (4) 能力值為正無限大者的答對機率

22. 當 RSM 模式退化為二元計分且 τ 皆為 0 時，較接近下列何種模式？

- (1) 1PL
- (2) 2PL
- (3) 3PL
- (4) 4PL

23. 一般所謂的 Rasch 模式是指下列何者？

- (1) 1PL
- (2) 2PL
- (3) 3PL
- (4) RSM

24. 對於以填充題為主的測驗而言，較適合用哪個模式來分析？

- (1) 1PL
- (2) 2PL
- (3) 3PL
- (4) RSM

25. 下列哪個模式屬於所謂的 Rasch 家族 (Rasch family) ？

- (1) 2PL
- (2) 3PL
- (3) 4PL
- (4) RSM

26. RSM 模式較適合用來分析何種題型的資料？

- (1) 數學科的應用題
- (2) 國語科的作文題
- (3) 李克特式量表
- (4) 魏氏智力量表

27. 4PL 模式較適合用來分析何種題型的資料？

- (1) 數學科的應用題
- (2) 李克特式量表
- (3) 國語科的作文題
- (4) 魏氏智力量表

28. 對於七點計分的試題而言，請問評分量尺模式會有幾個閾值？

- (1) 5
- (2) 6
- (3) 7
- (4) 8

29. 對一份有 10 題二元計分試題的測驗而言，採用 RSM 模式跟 3PL 模式，所估計的試題參數個數會相差幾個？
- (1) 10
 - (2) 20
 - (3) 30
 - (4) 40
30. 假設有一資料中受試者有 N 位，試題有 n 題，如以 2PL 模式來分析資料，有幾個參數需要估計？
- (1) $N+n$
 - (2) $2N+n$
 - (3) $2N+2n$
 - (4) $N+2n$

第二部份 (10 題)

請先閱讀以下的參考資料，然後根據其內容與相關的測驗與評量概念，回答下面的選擇題。
參考資料第一篇：

中華民國八十九年三月三十日台(八九)國字第八九〇三九二一九號函正式公告

國民教育階段的課程設計應以學生為主體，以生活經驗為重心，培養現代國民所需的基本能力。

一、瞭解自我與發展潛能

充分瞭解自己的身體、能力、情緒、需求與個性，愛護自我，養成自省、自律的習慣、樂觀進取的態度及良好的品德；並能表現個人特質，積極開發自己的潛能，形成正確的價值觀。

二、欣賞、表現與創新

培養感受、想像、鑑賞、審美、表現與創造的能力，具有積極創新的精神，表現自我特質，提升日常生活的品質。

三、生涯規劃與終身學習

積極運用社會資源與個人潛能，使其適性發展，建立人生方向，並因應社會與環境變遷，培養終身學習的能力。

四、表達、溝通與分享

有效利用各種符號(例如語言、文字、聲音、動作、圖像或藝術等)和工具(例如各種媒體、科技等)，表達個人的思想或觀念、情感，善於傾聽與他人溝通，並能與他人分享不同的見解或資訊。

五、尊重、關懷與團隊合作

具有民主素養，包容不同意見，平等對待他人與各族群；尊重生命，積極主動關懷社會、環境與自然，並遵守法治與團體規範，發揮團隊合作的精神。

六、文化學習與國際瞭解

認識並尊重不同族群文化，瞭解與欣賞本國及世界各地歷史文化，並體認世界為一整體的地球村，培養相互依賴、互信互助的世界觀。

七、規劃、組織與實踐

具備規劃、組織的能力，且能在日常生活中實踐，增強手腦並用、群策群力的做事方

法，與積極服務人群與國家。

八、運用科技與資訊

正確、安全和有效地利用科技，蒐集、分析、研判、整合與運用資訊，提升學習效率與生活品質。

九、主動探索與研究

激發好奇心及觀察力，主動探索和發現問題，並積極運用所學的知能於生活中。

十、獨立思考與解決問題

養成獨立思考及反省的能力與習慣，有系統地研判問題，並能有效解決問題和衝突。

參考資料第二篇：

國民中小學學生成績評量準則

九十年三月廿九日台(九〇)參字第九〇〇四二五七八號令發布

第一條

本準則依國民教育法施行細則第二十一條第一項規定訂定之。

第二條

國民中小學學生成績評量旨在了解學生學習情形，激發學生多元潛能，促進學生適性發展，肯定個別學習成就，並作為教師教學改進及學生學習輔導之依據。

第三條

國民中小學學生成績評量應依學習領域及日常生活表現，分別評量之；其評量範圍如下：

一、學習領域評量：依能力指標、學生努力程度、進步情形，兼顧認知、技能、情意等層面，並重視各領域學習結果之分析。

二、日常生活表現評量：學生出席情形、獎懲、日常行為表現、團體活動表現、公共服務及校外特殊表現等。

第四條

國民中小學學生成績評量應本適性化、多元化之原則，兼顧形成性評量、總結性評量，必要時應實施診斷性評量及安置性評量。

第五條

國民中小學學生成績評量，分定期評量及平時評量二種；定期評量之次數，由直轄市、縣(市)主管教育行政機關定之。

第六條

國民中小學學生成績評量，應視學生身心發展及個別差異，依各學習領域內容及活動性質，採取筆試、口試、表演、實作、作業、報告、資料蒐集整、鑑賞、晤談、實踐等適當之多元評量方式，並得視實際需要，參酌學生自評、同儕互評辦理之。前項評量方式，由任課教師依教學計畫在學期初向學生及家長說明，並負責評量。

第七條

國民中小學學生成績評量紀錄應兼顧文字描述及量化紀錄。文字描述應依評量內涵與結果詳加說明，並提供具體建議。量化紀錄得以分數計之，至學期末應轉換為甲、乙、丙、丁、戊五等第方式紀錄。

第八條

國民中小學學生成績評量紀錄，每學期至少應以書面通知家長及學生一次；其次數、方式、內容，由直轄市、縣(市)主管教育行政機關定之。

第九條

國民中小學學生修業期滿，成績及格，由學校發給畢業證書。

第十條

國民中小學學生成績評量結果及紀錄，應本保密及維護學生權益原則，非經學校、家長及學生本人同意，不得提供作為非教育之用。

第十一條

國民中小學學生各項成績評量相關表冊，由直轄市、縣（市）主管教育行政機關定之。

第十二條

直轄市、縣（市）主管教育行政機關應依本準則訂定國民中小學學生（含身心障礙學生）成績評量辦法。前項辦法應衡酌身心障礙學生之學習優勢管道，彈性調整其評量方式。非地方政府所屬之國民中小學學校學生，其成績評量得參照所在地主管教育行政機關所定之成績評量辦法辦理之。

第十三條

國民中小學實施九年一貫課程前，其學生成績評量辦法，由直轄市、縣（市）主管教育行政機關參照國民中小學學生成績考查辦法、國民中學學生成績考查辦法定之。

第十四條

本準則自中華民國九十年八月一日施行。

請先閱讀以上的參考資料，然後根據其內容與相關的測驗與評量概念，回答下面的選擇題。

31. 依照參考資料的法規，國民中小學生成績評量方式應為下列何者？
- (1) 紙筆測驗
 - (2) 多元評量
 - (3) 基本學力測驗
 - (4) 段考、小考、作業的平均成績，再加上教師評語
32. 九年一貫課程強調培養學生的十大基本能力，請問十大基本能力最適合用下列哪一種評量方式？
- (1) 基本學力測驗
 - (2) 紙筆測驗
 - (3) 段考、小考、作業的平均成績，再加上教師評語
 - (4) 多元評量
33. 請問國中基本學力測驗是屬於下列何者？
- (1) 安置性評量
 - (2) 診斷性評量
 - (3) 總結性評量
 - (4) 形成性評量
34. 九年一貫課程強調培養學生的十大基本能力，請問第一種基本能力（瞭解自我與發展潛能）最適合用下列哪一種評量方式？
- (1) 實作評量
 - (2) 非紙筆測驗
 - (3) 學生自評量表
 - (4) 檔案評量

35. 有關評量辦法的主管機關，下列何者是最不符合參考資料所述？
- (1) 國民中小學學生成績評量準則是由教育廳訂定。
 - (2) 國民中小學成績評量辦法一直是由各縣、市教育局主管。
 - (3) 實施九年一貫課程以後，國民中小學成績評量辦法不再由直轄市、縣(市)主管教育行政機關訂定。
 - (4) 國民教育法和其施行細則都是由教育部訂定。
36. 九年一貫課程強調培養學生的十大基本能力，請問第二種基本能力(欣賞、表現與創新)最適合用下列哪一種評量方式？
- (1) 檔案評量
 - (2) 非紙筆測驗
 - (3) 心理測驗
 - (4) 闖關遊戲
37. 就法令的位階而言，若由高往低排，請問下列何者最符合參考資料的內容？
- (1) 實施要點 → 準則 → 辦法 → 注意事項
 - (2) 法 → 施行細則 → 準則 → 辦法
 - (3) 辦法 → 準則 → 施行細則 → 法
 - (4) 施行細則 → 準則 → 辦法 → 注意事項
38. 九年一貫課程強調培養學生的十大基本能力，請問第四種基本能力(表達、溝通與分享)最不適合用下列哪一種評量方式？
- (1) 作文
 - (2) 表演
 - (3) 筆試
 - (4) 演講
39. 依法，國民中小學應該進行同儕互評。請問有關同儕互評的敘述，下列何者最符合學理與實情？
- (1) 每位學生自我評量，可以促進學生反省的能力。
 - (2) 學生判斷能力仍未成熟，因此其評分之效度一向很低。
 - (3) 學生的道德不好，所以他們常藉評分的机会，夾怨報復。
 - (4) 學生之間互相評分，雖然容易進行，但是卻不容易統計分數。
40. 九年一貫課程強調培養學生的十大基本能力，而且也實施多年，並且遭到社會大眾的許多批評。當然，亦有些政府官員和御用學者出來辯駁。在此，姑且不論誰是誰非。不過，從參考資料可以得知，其最大的缺陷是下列何者？
- (1) 建構式數學大幅降低學生的數學程度。
 - (2) 九年一貫課程的市售教材並未包含十大基本能力的評量材料。
 - (3) 國中基本學力測驗的難度太低，造成學生不再用功讀書。
 - (4) 大部分的國中小教師仍不認同九年一貫課程的理念。