找到下一個鑷子超人：
從情境判斷到處置推理的專業養成

高雄長庚 王慧婷醫師

2025 中山大學 AMEE 分享會

分享大綱

- **PCW 6: Contextualising and Evaluating Situational Judgement Tests for Use in Selection, Assessment and Development**

  情境判斷測驗在甄選、評估與發展中的情境化與評估

- **PCW 7: Management reasoning**

**Work Psychology Group**
Thinking differently

Contextualising and Evaluating Situational Judgement Tests for Use in Selection, Development and Assessment 情境判斷測驗在甄選、評估與發展中的情境化與評估

Jordan Buxton, Senior Consultant Psychologist, UK

■ **Work Psychology Group (WPG):**

- 英國組織心理學顧問公司
- 專注於人員甄選與評估（**selection & assessment**）
- **SJT** 設計與實施
- 跨專業的評估方案（包括醫教、企業等）
- 幫助機構在招聘、培訓與專業發展中提升公平性與效度
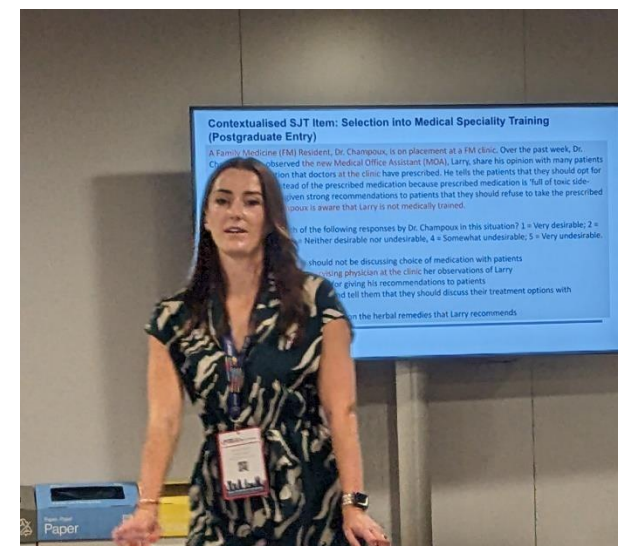
情境判斷測驗 課程大綱

- 情境判斷測驗 (SJT) 介紹

- SJT 現有實證證據/應用案例

- SJT 設計方法：量身打造 VS.在地化

- SJT 題目發展注意事項

- 小組活動：SJT 題目撰寫與情境化練習

- AI如何影響SJT發展

什麼是情境判斷測驗 -1
(Situational Judgement Tests, SJTs)?

- 測量方法：**評估學習者或人員在與角色相關情境中的判斷與決策**

- 測量重點：**在於非學術性的能力、行為與價值觀，如：誠信、同理心、抗壓力、團隊合作**

- 設計方式：**呈現具挑戰性、真實性或高風險的臨床/專業情境**

- 作答要求：**考生需對可能的反應或行動進行判斷與選擇**

- 評分依據：**依照預設標準答案（expert key）或共識進行評分**

SJT 命題示範

- **請從以下選項中，選出最適切做法**

表 1 「空服員人際實用智能量表」例題

類別	情境	選項
起降危安	一段從香港返回臺灣的航程，飛機開始進入降落的階段，空服員宋蕙喬瞧見了一位老伯伯急忙地往廁所的方向前去。然而飛機正在滑行，仍需一段時間才能降落於地。試問宋蕙喬應該如何處置呢？	1. 陪同老伯伯乘客一起去，並注意他的安危。 2. 口頭制止老伯伯前往，但他堅持的話，仍然默許。 3. 藉故移轉注意力，合理地忽略老伯伯上廁所的行動。 4. 按照規定阻止老伯伯前往廁所，以避免失職受懲。

SJT 命題示範

• 請將以下做法依適切性排序：1 代表最合適，5 代表最不合適

Example SJT Item - Selection into UK Pharmacy postgraduate training

You are in a hurry to deliver an urgent discharge prescription to a ward. While you are on the way, a patient approaches you and tells you that they are lost and asks for your help to find a clinic where they are scheduled for an appointment. You know the location of the clinic, but the prescription needs to be delivered urgently.

Rank in order the following actions in response to this situation (1 = Most appropriate; 5 = Least appropriate).

A   Give the patient directions to the clinic
B   Walk the patient to the clinic
C   Suggest that the patient asks for help at the hospital reception desk
D   Ask a colleague in the corridor to take the patient to the clinic
E   Direct the patient to a map of the hospital on the wall

你正急著要把一張緊急的出院處方送到病房。途中，有一位病人走向你，告訴你他迷路了，並請你幫忙帶他去一個預約好的門診。你知道門診的位置，但這張處方也需要緊急送達...

什麼是情境判斷測驗-2
(Situational Judgement Tests, SJTs)?

- SJT 是一種測量方法，而不是單一的測驗形式
- ✓考生收穫：提供一個真實的角色與情境預覽
- ✓可以有不同設計：題目內容、測驗長度、作答形式、作答指示
- ✓可能有不同評分方式：計分規則、與其他評估的權重設定
- ✓情境呈現：**文字、圖像/音訊、影片、互動模擬案例、 VR/AR等**

SJT 現有實證證據

Selection Method	Reliability	Validity	Candidate acceptability	Promotes widening access?
Academic records	High	High	High	Low
Structured Interviews/MMIs	Moderate to high	Moderate to high	High	Moderate
Situational Judgement Tests	High	High	Moderate to high	High
Aptitude testing	High	Various	Moderate	Moderate
Personality Tests	High	Moderate	Low to moderate	N/A
Traditional Interviews	Low	Low	High	Low
Personal statements	Low	Low	High	Low
References	Low	Low	High	Low

Patterson, et al, (2016). How effective are selection methods in medical education and training? A systematic review. Medical Education. 50(1), pp. 36-60.

SJT 現有應用案例-1

Using Situational Judgement Tests for Selection

Case study example	Approx N	Sector	Purpose	Location	Additional Comments
Selection to UK Medical/Dental School	N=50,000	**Medicine & Dentistry**	Undergraduate	UK / ANZ	Used alongside CAT
Entry to Speciality Training	N=18,000	**Medicine**	Specialty training	UK	Test used for different purposes across approx. 15 specialties
Postgraduate entry to Speciality Training	ENG N=3,600 FRA N=500	**Medicine**	Specialty training	Canada	Translated into French version
Selection into Pharmacy Training	N=2,500	**Pharmacy**	Postgraduate	UK	Test centre and online version
Entry into Veterinary School	N=2,000	**Veterinary Medicine**	Postgraduate	USA	Contextualised items
Selection into Junior Barrister Training	N=130	**Pupillage / Legal sector**	Postgraduate	UK	Administered online and unproctored
Selection into Medical School	N=1,000	**Medicine**	Undergraduate	Singapore	
Graduate Entry into Local Government	N=3,000	**Local Government**	Postgraduate	UK	Untimed

SJT 現有應用案例-2

Using Situational Judgement Scenarios for Development

Case study example	Sector	Location	Additional Comments
Development tool used across Pharmacy Degree	**Pharmacy**	Australia / Malaysia	Candidates receive a feedback report including attribute scores, contextualised feedback and reflective questions to support development
Development tool and supporting workshops for Foundation Year 2 doctors	**Medicine**	UK	Interactive online tool with video scenarios. Follow-up workshops using SJS' to provide additional development opportunities.
Development of Professionalism in Dental Students	**Dentistry**	UK	SJS tool was standard set, to ensure all students had the minimum level of professionalism before entering practice.
Development tool for Pharmacy Students	**Pharmacy**	US	SJS tools developed at 2 different pharmacy schools in the US, to support in ensuring students had the professional skills required to begin placements and inform targeted support (if below the threshold).
Development of resilience across professions	**Multiple (Veterinary, Medicine, the Arts)**	UK	Use of SJS' set-in real-life situations to support individuals develop resilience resources in a workshop setting.

SJT 設計方法：量身打造 VS.在地化

- **量身打造**

Bespoke SJT
A bespoke SJT is designed from scratch, specifically for the organisation.

Custom Content: Scenarios are created from scratch to reflect the organisation's unique values, culture, and job roles.
客製化內容

High Face Validity: Candidates perceive the test as highly relevant and realistic, increasing engagement.
高表面效度(題目可高度評估到想要的概念)

Exclusive Use: Content is bespoke for, and owned by that organisation.
專屬機構使用

SME Involvement: Significant input from subject matter experts and psychometricians in the development, increasing the validity of the SJT.
需要專家與心理計量專家大量投入

Validation Required: Needs piloting and statistical validation before operational use.
試測與統計效度分析

Increased Design Options: Greater flexibility in design of item types.
題目類型/設計高靈活度

SJT 設計方法：量身打造 VS.在地化

- 在地化：從既有、已驗證題庫中取題，由領域專家調整符合情境脈絡

Contextualised SJT

A contextualised SJT is content that is taken from a bank of pre-validated SJT items, which are contextualised to a specific context using Subject Matter Expert input.

Quick Deployment: Can be implemented rapidly due to existing validated item banks. 快速實施

Budget-Friendly: Offer a lower resource and cost alternative to bespoke SJTs. 預算友善

Broad Competency Coverage: Measures general workplace behaviours and competencies. 可測量職場一般具備行為與能力
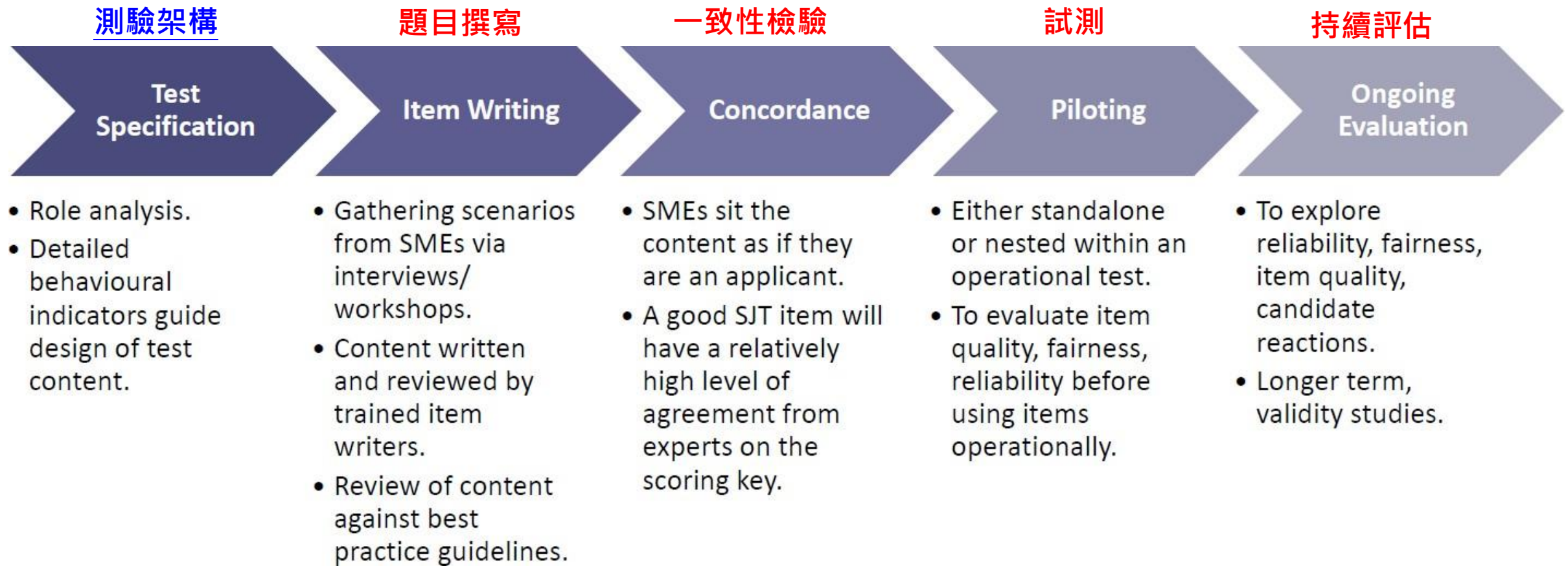
Scalable: Easily amended to be used across multiple roles or departments. 具可擴展性

Benchmarking Possible: Ability to compare scores across candidates and organisations. 可進行不同群體比較

Contextualised: Content is contextualised to local context to reflect the role or organisational culture.
情境化，以符合在地脈絡，反映角色或組織文化

SJT 設計方法：量身打造制定流程

測驗架構　　　　題目撰寫　　　　一致性檢驗　　　　試測　　　　持續評估

Test Specification	Item Writing	Concordance	Piloting	Ongoing Evaluation
• Role analysis. • Detailed behavioural indicators guide design of test content.	• Gathering scenarios from SMEs via interviews/ workshops. • Content written and reviewed by trained item writers. • Review of content against best practice guidelines.	• SMEs sit the content as if they are an applicant. • A good SJT item will have a relatively high level of agreement from experts on the scoring key.	• Either standalone or nested within an operational test. • To evaluate item quality, fairness, reliability before using items operationally.	• To explore reliability, fairness, item quality, candidate reactions. • Longer term, validity studies.

Patterson, F., Zibarras, L., & Ashworth, V. (2016). Situational judgement tests in medical education and training: Research, theory and practice: AMEE Guide No. 100. Medical teacher, 38(1), 3-17.

如何設計測驗架構(Test Specification)

測驗目的

Test Purpose

E.g. As an assessment tool, used alongside other methods or to rank applicants

Who is to be assessed?

測驗核心內容

Test Content

To assess non-academic attributes relevant to the role (*e.g. empathy*)

To challenge candidates

Setting or context relevant to role

Scenario set at the appropriate level

題型與作答方式

Item Types and Response Format

E.g. Ranking, rating, choose3

Testing what candidates *should* do, not *would* do

測驗長度與計分

Test Length and Scoring

Number of items per test paper

Length of time allowed to complete a test paper

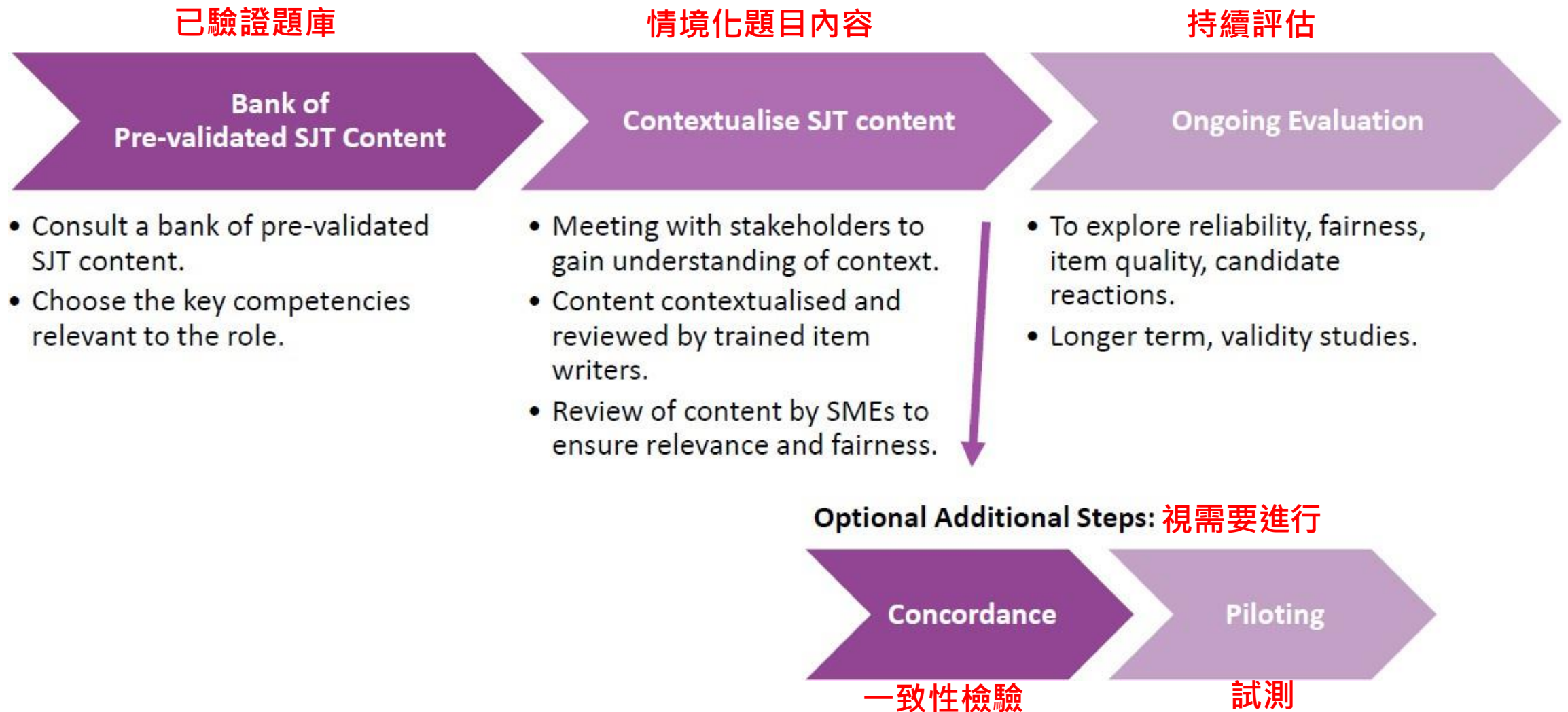**測驗管理與執行**

Test Management and Administration

Paper based vs computer delivered

Standardised administration and security of materials

Proctored vs unproctored

可及性

Accessibility

Reasonable adjustments

To ensure all candidates have opportunity to demonstrate their skills

15

SJT 設計方法：在地化制定流程

已驗證題庫   **情境化題目內容**   **持續評估**

Bank of Pre-validated SJT Content

- Consult a bank of pre-validated SJT content.
- Choose the key competencies relevant to the role.

Contextualise SJT content

- Meeting with stakeholders to gain understanding of context.
- Content contextualised and reviewed by trained item writers.
- Review of content by SMEs to ensure relevance and fairness.

Ongoing Evaluation

- To explore reliability, fairness, item quality, candidate reactions.
- Longer term, validity studies.

Optional Additional Steps: 視需要進行

Concordance

一致性檢驗

Piloting

試測

Patterson, F., Zibarras, L., & Ashworth, V. (2016). Situational judgement tests in medical education and training: Research, theory and practice: AMEE Guide No. 100. Medical teacher, 38(1), 3-17.

16

SJT 命題發展注意事項-1

- ## 不同題型可在同份測驗中混合使用

  ✓ 職涯早期適用「低認知負荷」
  題型

  ✓ 不同題型認知負荷由低至高：
  - 最佳/最差題 (Most/Least)

  - 評分題(Rating)

  - 多重選擇 (Multiple Choice, MCQ)

  - 排序題(Ranking)

選出「最合適」與「最不合適」

Most / Least

3 to 4 'standalone' responses to each scenario

Select the most appropriate and least appropriate

評分 (對4-6個選項單獨評分)

Rating

4 to 6 'standalone' responses to each scenario

Rate each response separately

多重選擇(8選3)

Multiple choice

Choose 3 correct options of 8 possible options

Appropriate when multiple actions done together generate a good response to the situation

排序 (5個選項進行排序)

Ranking

Rank 5 options in response to scenario

Appropriate when all the options are variations on a theme – i.e. you can't combine them

SJT 命題發展注意事項-2

- 命題難度具高度調整性，依需求設計，適用學生到高階主管等不同角色

✓設定合適的難度是設計關鍵

✓題目難度主要因素：
- 題型
- 選項設計
- 情境複雜度
- 角色層級

作答形式

Response format

Ranking and MCQ more cognitively loaded than Rating and Most/Least

干擾選項品質

Distractor quality

Distractors that are more closely related to 'correct' responses typically lead to a more difficult overall question

情境困境複雜程度

Complexity of dilemma in the scenario

概念/職責層級是否符合

Concepts/responsibilities familiar to specific level

SJT 命題發展注意事項-3

- 情境內容和選項設計是每一題命題的關鍵

情境內容

Developing Effective Scenario Content

- ✓ Relevant to the role 與角色相關
- ✓ Scenarios should not require clinical or specific procedural/legal knowledge
- ✓ Relevant to one of the target selection/development attributes
- ✓ Relatively short but add detail/specifics if necessary

簡潔為主，必要時可補充細節

選項設計

Developing Effective Item Responses

- ✓ Actions to address the scenario 對應並回應情境
- ✓ Keep responses simple & short where possible 精簡避免冗長
- ✓ Avoid implausible options 避免不合理/不可能選項
- ✓ Include realistic 'best thing to do' 至少有一個最適合做法
- ✓ Avoid ambiguity, e.g. 'speak to your supervisor' 避免模糊/籠統選項

SJT 命題檢核表

- **如何確保SJT題目的設計品質**

Is it set in the correct context?	題目是否設定在合適的情境脈絡？
Is it pitched at appropriate level?	題目是否符合考生的專業層級？
Is the item type appropriate?	作答形式（排序、評分、選擇）是否適切？
Are the responses plausible?	題目選項是否合乎邏輯？
Is there sufficient detail?	是否提供必要資訊，讓考生能正確判斷？
Language and style	避免過度複雜、冗長或不必要的術語？
Ambiguity	選項或情境是否有模糊不清的地方？

AI 如何影響 SJT 發展 -1

- **AI可協助SJT的命題發展、量性分析與施測管理**

協助發展測驗

Assessment Development
- Construct identification and development 確認測驗構念
- Automated item generation
- New item types

協助生成題目，開發選項

協助心理計量分析

Psychometric Methodology
- Item parameter calibration
- Generation of scoring keys
- Automated scoring

協助心理計量模型分析
建立評分標準
自動化評分

防止舞弊

Assessment Administration
- Flexible administration and authentic responses
- Prevention of cheating
- Detection of cheating

隨機出題
偵測異常行為
分析作答模式

Hao, J., von Davier, A. A., Yaneva, V., Lottridge, S., von Davier, M. & Harris, D. J. (2024). Transforming assessment: The impacts and implications of large language models and generative AI. Educational Measurement: Issues and Practice. 43(2), pp. 16-29

AI 如何影響 SJT 發展-2

- **AI也可能嚴重影響測驗的公平性與信效度**

- **透過以下實務步驟來降低 AI 帶來的風險：**

命題設計	測驗施行方式	確保測驗系統穩健	提升透明度與考生評估
Design of Assessments (e.g. content creation, random ordering of questions, questions that require problem solving)	**Delivery of Assessments** (e.g. different proctoring, time-limits)	**Ensuring Assessment Systems are Robust** (e.g. multiple assessments across methods, AI detection software, review response timing).	**Transparency and Candidate Evaluation** (e.g. clear messages about AI's involvement and human touch points, clear guidelines, honesty statements).
增加問題解決能力題目		多元測驗方式 AI偵測 分析作答反應時間	清楚說明AI介入程度 提供明確指引 誠信聲明

Hao, J., von Davier, A. A., Yaneva, V., Lottridge, S., von Davier, M. & Harris, D. J. (2024). Transforming assessment: The impacts and implications of large language models and generative AI. Educational Measurement: Issues and Practice. 43(2), pp. 16-29

Summary

- **SJT 價值**
✓測量**非學術的關鍵屬性**（如專業判斷、同理心），具有**高度預測效度**，適用於不同職位與層級

- **應用具多元性**
✓用於甄選，也能支持專業發展與課程評估，可依需求調整情境與屬性
✓**可作為教學工具，透過情境測驗與討論**，幫助學員將理論轉化為實務

- **發展方法**

特點	量身打造(Bespoke SJT)	在地化(Contextualized SJT)
內容來源	從零開始設計	從現有題庫選取，依照脈絡調整
表面效度	高度相關、貼近情境	中等相關，依調整程度而定
使用範圍	專屬於該組織，不適合外部使用	可跨單位或跨組織使用，較具擴展性
專家參與	需要大量專家，建立信效度	需專家確認在地適用性，投入相對較少
驗證需求	需重新進行試測與統計驗證	依原題庫基礎，驗證需求較低
成本	高成本、需較長時間	相對快速、具成本效益

Management Reasoning (MR)
臨床處置推理工作坊

高雄長庚

王慧婷醫師

工作坊目標

1. 什麼是 MR
2. 優化 MR 發展的關鍵做法
3. 能即刻上手的 MR 教學工具（腳本、暫停、公平性反思）
4. 會設計 MR 的評量與補救（coaching + 刻意練習）

MR 與 DR 的區分

- Clinical Reasoning
  - Processing and assimilating of preexisting knowledge & current patient-specific information to support medical practice
- Diagnostic Reasoning
  - Classification task that assigns meaningful labels to constellations of data (s/sx, results) to make diagnoses
- Management Reasoning
  - Prioritization task about testing, treatment, & allocation of resources

MR 的定義、特性、與文獻

- **定義 Definition（Cook 2018）**
  - *英文*: *"The process of making decisions about patient management, including choices about treatment, follow-up visits, further testing, and allocation of limited resources."*
  - *中文*:**「針對病人處置所做的決策過程，包括治療、追蹤、進一步檢查以及有限資源的分配。」**
- **特性**
  - 行動導向與風險承擔
  - 情境性（situated）與分散性（distributed）：受人員、場域、工具共同影響
  - 動態監測與調整：非一次到位，需持續微調

Why Management Reasoning?

- Most real-world decisions are management decisions, not diagnostic ones

- High risk — requires doing something to the patient

- Learners receive limited structured guidance on how to reason through management

- The shift from "knower" to decision-maker defines professional growth (especially in a world of AI)

Situated reasoning

學習者回饋：MR 如何長成？

- Often contrasted with diagnostic reasoning as:
  - More open-ended, less certain, and less satisfying
  - Focused on action, contextual fit, and feasibility
- Over time, learners recognize that effective MR is about making defensible choices amid ambiguity and multiple acceptable options.
- Growth occurs through situated experience, not didactics. Specifically, exposure to:
  - Real-world complexity
  - Evolving illness trajectories
  - Contextual constraints (e.g., team dynamics, patient preferences)
- Progression from rule-based decisions to adaptive expertise, weighing tradeoffs
- Critical role of feedback, role modeling, verbalization of reasoning, and guided autonomy

常見障礙（Barriers）

- **時間壓力**：臨床流程過快，無暇 verbalize reasoning
- **文化因素**：亞洲場域常有「不質疑上級」現象
- **心理安全不足**：害怕顯得能力不足，不敢問、不敢承諾
- **教師差異大**：不同主治醫師偏好不一，學員混亂
- **課程缺口**：診斷推理常有課程，處置推理較少被結構化教導
- **評量不足**：多停在「知識考核」，缺乏針對 MR 的明確評量工具

Learner 支援框架：四大面向

- **Experiential learning**：透過真實臨床與多樣化案例，讓學員在安全情境中實際參與並承擔處置決策

- **Reflective practice**：引導學員持續追問『為什麼』並反思自己的處置選擇與結果，從錯誤中成長。

- **Clinical environment**：營造逐步放手的學習環境，鼓勵表達推理並在多變情境中練習自主決策。

- **Structured teaching**：以結構化框架和案例教學，幫助學員在不確定性中培養彈性與適應性處置能力

三大工具

- **Management Script Template（處置腳本模板）**
  - 檢驗｜影像｜處置/操作｜專科會診｜藥物/監測｜去向/追蹤
  - 訓練路徑：**What → Why → How**
- **Management Pause（處置暫停）**
  - 在不確定/高風險節點刻意停下來反思
  - 問題：為何選？壞處？替代？病人觀點？
- **Equity Reflection（公平性反思）**
  - 是否偏離標準？為何？如何排除障礙？
  - 若是你/家人會希望什麼？可否補強？

案例討論

- **Case 1**：67 歲女，疑 UTI、強烈要求抗生素
  - 工具：**Pause** → 釐清焦慮、風險、替代、共享決策
- **Case 2**：T2DM + 肺炎；PGY-1 想因成本改口服
  - 工具：**Equity Reflection** → 經濟/可近性、風險、惡化門檻
- **Case 3**：疑 CAP；學生只答「廣效抗生素」
  - 工具：**Script Template** → 具體化藥物、監測、去向
-

Three Major Categories of MR Assessment

- **Non-workplace-based assessments**
  - Clinical vignettes, Concept maps, Extended matching questions, Key features exam, MCQ, Short essays, Oral exams, Script concordance test, Patient management problems
- **Simulation-based assessments**
  - OSCEs, Technology-enhanced simulation
- **Workplace-based assessments (WBAs)**
  - Direct observation, Case presentations, Chart-stimulated recall, Think-aloud protocols, Global ratings, Micro-analytics
- **Key point**: No single method is sufficient → need multiple complementary approaches

Common Behavioral Struggles

- Unable to prioritize tasks (e.g., taking a long history while patient in shock)
- Failure to commit / indecisive
- Failure to escalate concerns (missed deterioration, drug interaction)
- No trajectory or contingency plan (no monitoring indicators or Plan B)
- Poor interprofessional coordination (not engaging nurses, RT, pharmacists)
- Knowledge gaps + lack of confidence → hesitation, rigidity

**Diagnose the deficit → 分類 → simulation → 臨床工作**

Remediation 策略（Coaching + Deliberate Practice）

- **Diagnose the deficit ➔ 分類 ➔ simulation ➔ 臨床工作**
- **流程**：設定明確目標 → 學員嘗試 → 教師觀察 → 即時回饋 → 重複循環
- **工具化輔助**
  - Script Template：What → Why → How
  - Benefit–Harm List：每項處置都需列出好處與壞處（含成本/可近性/健保給付）
  - Think-Aloud + Frequent Stops：強迫 verbalization，顯示推理過程
- **評量連動**
  - OSCE/CSR：檢核學員是否展現 **承諾 + 理由 + 監測/追蹤 + 共享決策**

Take Home Message

- **處置推理（MR）不是要找到唯一的標準答案**
- 重點在於：
- **多種合理選擇** → 評量的是思考過程與理由，而不是唯一正解
- **工具簡單好上手**：處置腳本（Script Template）、處置暫停（Pause）、公平性反思（Equity Reflection）
- **臨床容易落地**：可融入交班、OSCE、模擬訓練、EPA 評量
- **補救靠教練式練習**：明確目標 × 刻意練習 × 即時回饋
- **最終目標**：培養學員能在不確定與限制中，做出有理有據、以病人為中心的臨床決策

感謝聆聽 敬請
指導！
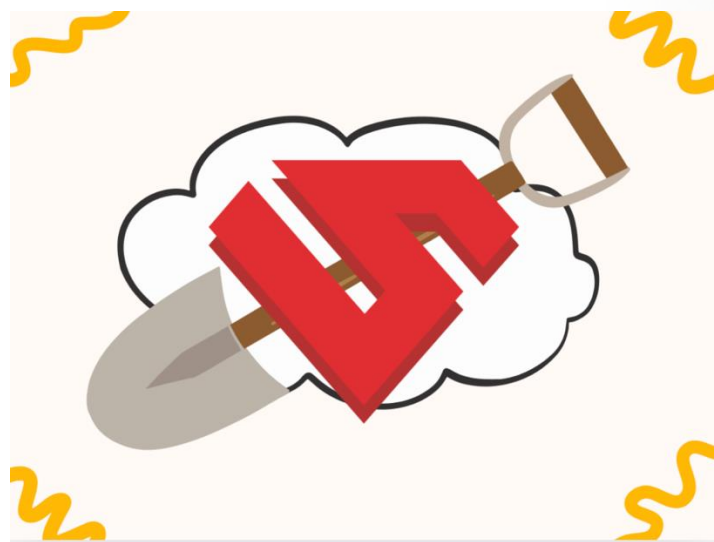
高長王慧婷醫師
gardinea1983@gmail.com