

บทปริทัศน์

การสร้างข้อสอบปรนัยแบบเลือกตอบ

วัลลี สัตยาชัย

บทคัดย่อ

ข้อสอบปรนัยแบบเลือกตอบ เป็นข้อสอบที่นิยมใช้กันแพร่หลาย เพราะสามารถวัดความรู้ได้ดี และครอบคลุมเนื้อหาได้มาก อย่างไรก็ตาม ข้อสอบชนิดนี้ก็มีทั้งจุดเด่นและจุดอ่อนเช่นเดียวกับเครื่องมือวัดผลชนิดอื่นๆ การสร้างข้อสอบปรนัยแบบเลือกตอบให้มีคุณภาพ ต้องรู้หลักการในการสร้างที่ถูกต้องทั้งการสร้างโจทย์ คำถาม และชุดคำตอบที่ให้เลือก บทความนี้กล่าวถึง หลักการสร้างข้อสอบปรนัยแบบเลือกตอบ จุดเด่นและจุดอ่อนของข้อสอบ ตลอดจนการสร้างข้อสอบปรนัยแบบเลือกตอบในวิชาวิทยาศาสตร์การแพทย์พื้นฐาน และตัวอย่างของข้อสอบ

คำสำคัญ: ข้อสอบปรนัยแบบเลือกตอบ

วันที่รับบทความ: ๒๖ กันยายน ๒๕๕๕

วันที่อนุญาตให้ตีพิมพ์: ๒๕ ตุลาคม ๒๕๕๕

ข้อสอบปรนัย (objective examination) มีอยู่หลายแบบด้วยกัน ได้แก่ ข้อสอบแบบเลือกตอบ (multiple choice question) ข้อสอบแบบถูกผิด (true/false) และข้อสอบแบบจับคู่ (matching) แต่ที่นิยมใช้ในการวัดผลมากที่สุดในปัจจุบันคือ ข้อสอบแบบเลือกตอบ (multiple choice question) ประเภทที่มีคำตอบที่ถูกที่สุดเพียงคำตอบเดียว (one best answer) เหตุผลที่ multiple choice question หรือที่เรียกกันสั้นๆ ว่า MCQ เป็นที่นิยมใช้กันแพร่หลายในการวัดผลตั้งแต่อดีตจนถึงปัจจุบัน ก็เพราะเป็นเครื่องมือที่วัดความรู้ (cognitive domain) ได้ดี สามารถครอบคลุมเนื้อหาได้มาก (high content validity) การตรวจให้คะแนนทำได้ง่ายและรวดเร็ว เหมาะสำหรับการวัดผลที่ผู้เข้าสอบมีจำนวนมาก แต่อย่างไรก็ตาม เครื่องมือวัดผลทุกชนิดมีทั้งจุดเด่นและจุดอ่อน ซึ่งจะได้นำกล่าวถึงหัวข้อนี้ต่อไปในภายหลัง

หลักการในการสร้างข้อสอบแบบ MCQ^{๑-๔}

๑. ข้อสอบต้องสามารถวัดผลการเรียนรู้ที่สำคัญ สิ่งที่บรรจุในข้อสอบควรมีความสำคัญ และเกี่ยวข้องกับผลการเรียนรู้โดยตรงในแต่ละวิชา

๒. ข้อสอบต้องวัดในสิ่งที่ตรงกับวัตถุประสงค์ของแต่ละวิชา และเป็นไปตามเกณฑ์มาตรฐานของสาขาวิชาชีพ

๓. ข้อสอบควรมีความยากง่ายเหมาะสมกับระดับของผู้ถูกทดสอบ (โดยทั่วไปจะมีข้อสอบที่ง่ายประมาณร้อยละ ๒๕ ปานกลางร้อยละ ๕๐ และยากร้อยละ ๒๕)

๔. ข้อสอบแต่ละข้อควรเป็นอิสระต่อกัน ไม่ควรให้ผู้เรียนต้องรู้คำตอบที่ถูกต้องของข้อแรกก่อนจะตอบคำตอบที่ถูกต้องของข้อต่อไปได้

๕. การสร้างข้อสอบ ควรให้ ตัวโจทย์หรือคำถามยาวแต่คำตอบสั้น และแต่ละข้อควรให้ใช้เวลาในการตอบประมาณ ๑ นาที โดยทั่วไปจึงมักให้เวลาในการทำข้อสอบ ๕๐ ข้อต่อ ๖๐ นาที เพื่อให้มีเวลาเหลือสำหรับการทบทวน

๖. ควรเรียงลำดับข้อสอบจากข้อที่ง่ายไปหาข้อที่ยาก จำนวนข้อสอบไม่ควรต่ำกว่า ๕๐ ข้อ เพื่อให้ข้อสอบมีความเที่ยง (reliability) สูง

นิยามคำศัพท์ที่ใช้ในการออกข้อสอบ MCQ

Item หมายถึง ข้อสอบแต่ละข้อ

Stem หมายถึง ตัวโจทย์

Lead-in หมายถึง คำถาม

Answer set / Option หมายถึง ชุดคำตอบที่ให้เลือก

Key / Answer หมายถึง คำตอบที่ถูกต้อง

Distractor / Foil หมายถึง ตัวลวง

การสร้างโจทย์ (stem)

๑. โจทย์ควรเป็นข้อความที่สั้นที่สุดที่สามารถอ่านแล้วเข้าใจ ใช้คำให้น้อยที่สุดเท่าที่จำเป็น โดยใส่ข้อความเฉพาะที่สำคัญที่เกี่ยวข้องกับคำถาม ภาษาที่ใช้ควรชัดเจน ไม่กำกวม ไม่วกวนหรือเยิ่นเย้อ และไม่ใช้คำย่อที่ไม่เป็นมาตรฐานสากล เช่น ใช้คำย่อ standard ว่า STD ซึ่งบางคนอาจไปแปลความหมายว่า sexual transmitted disease

ตัวอย่าง

โจทย์ : ชายไทยอายุ ๒๕ ปี โสัด อาชีพค้าขาย ปรกติแข็งแรงดี ไม่มีโรคประจำตัว มาร.พ.ด้วยเรื่องปวดท้องน้อยด้านขวา ๑ วัน ไม่มีไข้ ไม่มีตัวเหลืองตาเหลือง ไม่มีปวดศีรษะ ไม่มีคลื่นไส้อาเจียน ปัสสาวะอุจจาระปรกติ ไม่ได้รักษาและไม่ได้กินยาอะไรมาก่อน

P.E. T 38° C, PR 88/min, RR 20/min, BP 100/60 mmHg

GA : looked ill, no pallor, no jaundice

HEENT : no pale conjunctivae, no icteric sclera

Chest : heart and lungs are normal

Abdomen : no distention, decreased bowel sound, tenderness at RLQ with positive rebound tenderness, no hepatosplenomegaly

คำถาม : จงให้การวินิจฉัยโรคที่เป็นไปได้มากที่สุด จะเห็นว่าข้อสอบข้อนี้มิใช่โจทย์ที่ยาวมาก และข้อมูลหลายส่วนสามารถตัดทิ้งได้โดยไม่มีผลกระทบบต่อการตอบคำถาม ซึ่งเมื่อแก้ไขแล้วจะได้ตัวโจทย์ใหม่ที่สั้นและกระชับดังนี้ :

โจทย์ : ชายไทยอายุ ๒๕ ปี มาร.พ.ด้วยเรื่องปวดท้องน้อยด้านขวา ๑ วัน ไม่อาเจียน ปัสสาวะอุจจาระปรกติ P.E.: T 38° C, looked ill, tenderness at RLQ of abdomen with positive rebound tenderness and decreased bowel sound

๒. ควรหลีกเลี่ยงการใช้โจทย์ที่มีลักษณะเป็นหลาย dimension หรือเป็นลักษณะของ multiple true-false

ตัวอย่าง

Which statement is true ?

A. DHF is found in preschool more than school age.
B. Intracranial bleeding is common complication of childhood ITP.

C. Fever is the major criteria of acute rheumatic fever.

D. Adenovirus is common cause of hemorrhagic cystitis.

E. Gum hypertrophy is side effect of phenobarb.

จะเห็นว่า ตัวโจทย์หรือคำถามเป็นข้อความที่สั้นมาก แต่ชุดคำตอบนั้นแต่ละข้อยาวและไม่มีความเกี่ยวข้องกันเลย แตกต่างไปกันคนละทิศทาง ทำให้ข้อสอบข้อนี้ไม่สามารถวัดความรู้ของผู้เข้าสอบได้ว่ารู้หรือไม่รู้ในเรื่องใด เพราะอาจจะรู้ข้อที่ถูกต้องเพียงข้อเดียวก็ได้

การสร้างคำถาม (lead-in)^{๒๔}

๑. ควรเป็นประโยคคำถามที่สมบูรณ์ และเป็นคำถามที่สามารถทำให้ออกคำตอบได้ โดยไม่ต้องเห็นชุดคำตอบก่อนจึงทราบว่าการถามอะไร

ตัวอย่าง : The presence of large amounts of protein in the urine :

จะเห็นว่า เมื่ออ่านแต่โจทย์ จะยังไม่สามารถคิดคำตอบได้ เพราะไม่ทราบว่าผู้ถามต้องการจะถามอะไรเนื่องจากไม่ใช่เป็นประโยคคำถามที่สมบูรณ์ ดังนั้น จึงควรเปลี่ยนเป็น :

What does a large amount of protein in the urine indicate ?

๒. ควรเขียนให้อยู่ในรูปของข้อความบอกเล่า (positive form) เพราะจะวัดผลการเรียนรู้ที่สำคัญได้ดีกว่าข้อความปฏิเสธ (negative form) และหลีกเลี่ยงการใช้คำถามปฏิเสธซ้อนปฏิเสธ (double negative) เช่น all of the followings are incorrect except

กรณีที่มีความจำเป็นต้องใช้คำถามที่เป็นปฏิเสธในการวัดผลการเรียนรู้ที่สำคัญ ควรเน้นในคำนั้นโดยการขีดเส้นใต้ หรือพิมพ์ตัวใหญ่ หรือจัดไว้ตอนท้ายของประโยค โดยถ้ามีข้อสอบที่เป็นลักษณะคำถามปฏิเสธมากกว่า ๑ ข้อ ควรใช้การเน้นคำ หรือการใช้คำที่เป็นปฏิเสธในรูปแบบเดียวกัน ไม่ปะปนกันไปหมด เช่น ใช้คำว่า except ในข้อสอบที่เป็นปฏิเสธทุกข้อแทนที่จะใช้ wrong บ้าง incorrect บ้าง the least likely บ้าง หรือ except บ้าง นอกจากนี้ก็ควรจัดข้อสอบประเภทนี้รวมกลุ่มไว้ด้วยกันเพื่อป้องกันความสับสนของผู้สอบ

๓. ควรถามเพียงประเด็นใดประเด็นหนึ่ง เช่น แทนการถามสองประเด็นว่า What is the most likely diagnosis and management ? ควรเลือกถามเพียงประเด็นใดประเด็นหนึ่งที่คิดว่าสำคัญกว่า เพื่อที่จะสามารถวัดได้ว่าผู้สอบรู้หรือไม่รู้ในประเด็นนั้น

การสร้างชุดคำตอบ (answer set)^{๒๕}

๑. ตัวเลือกที่เป็นคำตอบที่ถูกต้อง (key) ควร มีคำตอบเดียวเท่านั้น และควรชัดเจนถูกต้องสมบูรณ์ที่สุด ผู้เชี่ยวชาญมีความเห็นตรงกันไม่ควรให้มีปัญหาสงสัยในคำตอบนั้น

๒. ตัวเลือกต่างๆ ควรสอดคล้องกับโจทย์ ทั้งไวยากรณ์และรูปแบบ

๓. ตัวเลือกที่ถูกต้องและตัวลวง (distractor) ต่างๆ ควรมีความละม้ายคล้ายคลึงกันมากที่สุดทั้งไวยากรณ์และรูปแบบ ตลอดจนความยาว เพื่อไม่ให้เป็นการบอกร่องรอยคำตอบที่ถูกต้องให้ผู้เรียนสามารถเดาคำตอบได้

๔. ควรหลีกเลี่ยงคำพูดที่เป็นเครื่องช่วยให้ผู้เรียนเลือกคำตอบที่ถูกต้อง หรือกำจัดตัวลวงที่ผิดออกไปได้ง่าย เช่น

- มีคำที่คล้ายกันในโจทย์และคำตอบที่ถูกต้อง

- ใช้คำในคำตอบที่ถูกต้องเหมือนกับคำในตำราเรียน

- เขียนคำตอบที่ถูกต้องยาวหรือสั้นต่างจากตัวเลือกอื่น ๆ

- บรรจุคำ *always, never* ลงในตัวเลือก เพราะคำเหล่านี้โดยทั่วไปเป็นคำที่ใช้ในข้อความที่ผิด ทำให้ผู้เรียนเดาได้ง่าย

- สร้างชุดคำตอบที่มีตัวเลือกซึ่งเป็นไปได้เพียงสองข้อเท่านั้น ส่วนตัวเลือกอื่นๆ ไม่ใกล้เคียงหรือเป็นไปได้ ซึ่งจะทำให้ผู้เรียนสามารถเดาคำตอบได้ถูกต้องร้อยละ ๕๐

๕. กรณีที่ชุดคำตอบเป็นตัวเลข ควรเรียงข้อตามลำดับตัวเลขจากน้อยไปมาก เช่น ๑๕ ๒๐ ๒๕ ๓๐ ๓๕

๖. การสร้างตัวลวงที่มีประสิทธิภาพ ควรใช้คำหรือข้อความที่ผู้เรียนมักเข้าใจผิดบ่อยๆ หรือเป็นข้อบกพร่อง หรือเป็นข้อผิดพลาดของผู้เรียนเสมอ และตัวลวงที่ใช้ไม่ควรเป็นข้อความที่ตรงข้ามโดยตรงกับคำตอบที่ถูกต้อง เช่น คำตอบที่ถูกต้องคือ hyponatremia แต่ใช้ตัวลวงเป็น hypernatremia

๗. หลีกเลี่ยงการใช้คำว่า “ถูกหมดทุกข้อ” (all of the above) เป็นตัวเลือก เพราะจะทำให้เดาคำตอบได้ง่ายขึ้น และถ้าไม่จำเป็นจริงๆ ก็ไม่ควรใช้คำว่า “ผิดหมดทุกข้อ” (none of the above) ในตัวเลือกเช่นกัน

๘. ควรให้รหัสคำตอบที่ถูกต้องอยู่ในลักษณะแบบคละกัน เพื่อป้องกันไม่ให้ผู้เรียนเดาหรือจั่วรหัสของคำตอบที่ถูกต้องได้ เช่น ไม่ใช่ข้อ D เป็นคำตอบที่ถูกต้องตลอด

ระดับของข้อสอบ

การสร้างข้อสอบ MCQ แบ่งเป็น ๓ ระดับ^๑ คือ

๑. ระดับความจำ (recall) เป็นข้อสอบที่ใช้ทดสอบความจำ จะเป็นคำถามตรงๆ ไม่มีโจทย์ นำมาก่อน เช่น What is the most common cause of anemia ?

๒. ระดับการแปลผล (interpretation) เป็นข้อสอบที่มีโจทย์ร่วมกับคำถามประเภทให้แปลผลหรือวินิจฉัยโรค เช่น

A 5-year-old child presents with epistaxis. He has a history of common cold last week. Physical examination reveals petechiae on extremities, bloody discharge from nose, no pallor, no hepatosplenomegaly.

Which of the following is the most likely diagnosis ?

๓. ระดับการแก้ปัญหา (problem solving) เป็น

ข้อสอบที่มีโจทย์ร่วมกับคำถามประเภทให้ตัดสินใจแก้ปัญหา เช่น

A 2-year-old child is brought to the emergency room because of high fever. During examination he suddenly develops generalized tonic and clonic convulsion.

Which of the following is the most appropriate initial management ?

ปัจจุบัน มักนิยมแบ่งเป็นเพียง ๒ ระดับ คือ ระดับความจำ (recall) และระดับประยุกต์ความรู้ (application of knowledge) ซึ่งรวมทั้งระดับแปลผลและการแก้ปัญหา

จุดเด่นและจุดอ่อนของข้อสอบ MCQ^๒

เครื่องมือวัดผลทุกชนิด มีทั้งจุดเด่นและจุดอ่อน ข้อสอบ MCQ ก็เช่นกัน

จุดเด่นของข้อสอบ MCQ

๑. ครอบคลุมเนื้อหาได้มาก

๒. เป็นข้อสอบที่มีความเที่ยงหรือความเชื่อมั่น (reliability) สูง

๓. ตรวจให้คะแนนได้ง่าย ซึ่งในปัจจุบันสามารถใช้เครื่องตรวจและวิเคราะห์ข้อสอบได้เลย จึงเป็นการสะดวกแม้จะมีผู้สอบจำนวนมาก

จุดอ่อนของข้อสอบ MCQ

๑. ใช้เวลานานในการสร้างข้อสอบที่ดี

๒. วัดได้เฉพาะความรู้

๓. วัดความสามารถในการแก้ปัญหาได้จำกัด และไม่สามารถวัดความสามารถในการหาข้อมูลได้

๔. มีคำตอบให้เลือก จึงสามารถเดาได้ โดยเฉพาะถ้าสร้างข้อสอบไม่ดี จะสามารถเดาได้ง่าย

การสร้างข้อสอบปรนัยแบบเลือกตอบสำหรับวิชาวิทยาศาสตร์การแพทย์พื้นฐาน^๓

ปัจจุบัน การสอบวิชาวิทยาศาสตร์การแพทย์พื้นฐาน ได้มีการปรับปรุงเพื่อวัดผลว่าผู้สอบสามารถนำความรู้ทางวิทยาศาสตร์การแพทย์พื้นฐานไปประยุกต์ใช้ทางคลินิกได้หรือไม่ ดังนั้น การสร้างข้อสอบปรนัยแบบเลือกตอบสำหรับ

วิชาวิทยาศาสตร์การแพทย์พื้นฐานส่วนใหญ่จึงควรสร้างให้สามารถวัดระดับความเข้าใจ การแปลผล และการแก้ปัญหาได้ โดยการสร้างข้อสอบที่มีโจทย์เป็นผลการทดสอบหรือข้อมูลของผู้ป่วยซึ่งไม่ควรจะยาวเกินไป แล้วตั้งคำถามที่เกี่ยวกับการทดสอบความรู้ทางวิทยาศาสตร์การแพทย์พื้นฐาน ไม่มุ่งเน้นทางด้านการดูแลรักษาผู้ป่วย

การสร้างต้นแบบโจทย์ที่เป็นข้อมูลผู้ป่วย (item template)^๑

ประกอบด้วยข้อมูลส่วนต่างๆ ดังนี้

๑. อายุ เพศ
๒. สถานที่ (ถ้ามีความจำเป็นต้องบอก)
๓. อาการและอาการแสดงสำคัญๆ ที่เกี่ยวข้อง
๔. ระยะเวลาของอาการ
๕. การตรวจร่างกาย
๖. ผลการตรวจทางห้องปฏิบัติการ (ถ้าจำเป็นต้องมี)
๗. การรักษาที่ผ่านมา (ถ้าจำเป็นต้องมี)

ตัวอย่าง item template

๑. **โจทย์ :** ข้อมูลทั่วไปของผู้ป่วย + ชนิดและตำแหน่งที่ได้รับบาดเจ็บ
คำถาม : อวัยวะใดของผู้ป่วยที่ได้รับผลกระทบมากที่สุด
หรือ : อะไรเป็นกลไกของยาที่ทำให้เกิดอาการดังกล่าว
๒. **โจทย์ :** ข้อมูลทั่วไปของผู้ป่วย + อาการ / อาการแสดง / ผลการตรวจทางห้องปฏิบัติการ + ประวัติการได้รับยา
คำถาม : ยาชนิดใดที่เป็นสาเหตุมากที่สุด
หรือ : อะไรเป็นกลไกของยาที่ทำให้เกิดอาการดังกล่าว
๓. **โจทย์ :** ข้อมูลทั่วไปของผู้ป่วย + อาการ / อาการแสดง
คำถาม : อะไรเป็นกลไกในการเกิดอาการดังกล่าว
หรือ : เชื้อก่อโรคใดที่เป็นไปได้มากที่สุด
หรือ : สารพิษชนิดใดที่เป็นสาเหตุได้มากที่สุด
๔. **โจทย์ :** ข้อมูลทั่วไปของผู้ป่วย + อาการผิดปกติ
คำถาม : ผลการตรวจทางห้องปฏิบัติการในข้อใดที่คาดว่าจะพบ

๕. **โจทย์ :** ข้อมูลทั่วไปของผู้ป่วย + โรคที่ทำให้เสียชีวิต

คำถาม : จงบอกความผิดปกติที่พบทางพยาธิวิทยา

ตัวอย่างข้อสอบวิชาวิทยาศาสตร์การแพทย์พื้นฐาน

๑. หลังจากกินเลี้ยงงานรับน้องใหม่ประมาณ ๖

ชั่วโมง นักศึกษาหลายคนมีอาการปวดท้องอย่างรุนแรง อาเจียน และต่อมาก็ถ่ายอุจจาระเป็นน้ำ

ท่านคิดว่าเชื้อ bacteria ใดในอาหารที่เป็นสาเหตุได้มากที่สุด ?

- A. Vibrio cholera
- B. Eschericia coli
- C. Streptococcus fecalis
- D. Staphylococcus aureus
- E. Salmonella typhimurium

คำตอบ D

๒. หญิงอายุ ๖๒ ปี มีบุตร ๕ คน มีอาการปัสสาวะเล็ดเวลาไอ จาม หรือหัวเราะมาก

แพทย์ควรแนะนำให้หญิงคนนี้บริหารกล้ามเนื้อในข้อใด

- A. Detrusor
- B. Obturator internus
- C. Piriformis
- D. Rectus abdominis
- E. Urogenital diaphragm

คำตอบ E

๓. เด็กชายอายุ ๓ ปี ๖ เดือน ได้รับการวินิจฉัยว่าเป็น mental retardation และตรวจพบว่ามี hepatosplenomegaly ผล liver biopsy พบมี glucocerebroside จำนวนมาก

ผู้ป่วยรายนี้ขาด enzymes ชนิดใด

- A. Hormone-sensitive lipase
- B. Lipoprotein lipase
- C. Lysosomal hydrolase
- D. Sphingolipid synthase
- E. Tissue phospholipase

คำตอบ C

เอกสารอ้างอิง

๑. Huxham GJ, Naeraa N. Bloom's taxonomy reflected in the response pattern to MCQ items. *Med Educ* 1980;14:23-6.
๒. Haladyna TM, Downing SM. Comparative review: Validity of a Taxonomy of Multiple-choice Item-Writing Rules. *Appl Measurement in Educ* 1989;2:51-78.
๓. Haladyna TM. Writing multiple choice items. Chicago CAT Inc; 2003.
๔. Haladyna TM. Developing and validating multiple-choice test items, 3rd ed. Mahwah, NJ: Lawrence Erlbaum Associates; 2004.
๕. Case SM, Jacob LC. Multiple choice items. In: Chase CI, Jacobs LC, eds. Developing and using test effectively. A guide for faculty. San Francisco: Jossey-Bass Publishers, 1992:51-82.
๖. United Medical Licensing Examination (USMLE) Step 1. Content description and general information 2008.
๗. United Medical Licensing Examination (USMLE) Step 1. Sample items 2008.

Abstract

Multiple choice question

Wallee Satayasai

Department of Pediatrics, Faculty of Medicine, Thammasat University

Multiple choice question is the most popular test in education because it is the good tool for knowledge measurement and can cover a lot of contents. Like other tests, multiple choice question has both strengths and weaknesses. Construction of good multiple choice question requires the knowledge of principle of writing stem, lead-in, and option. This article demonstrates how to construct good multiple choice question, strengths and weaknesses of this test, including samples of this test in basic medical sciences.

Key words: Multiple choice question