

ข้อผิดพลาดที่ควรระวังในการสร้าง ข้อสอบปรนัย

ผู้ช่วยศาสตราจารย์ นายแพทย์เชิดศักดิ์ โสมเกียรติ

ภาควิชาสัตวศาสตร์, คณะแพทยศาสตร์ศรีราชพยาบาล, มหาวิทยาลัยมหิดล, กรุงเทพมหานคร ๑๐๓๐๐.

ข้อผิดพลาดที่ควรระวังในการสร้างข้อสอบปรนัย

ข้อสอบปรนัย (multiple-choice question) เป็นรูปแบบการประเมินผลที่นิยมใช้กันอย่างแพร่หลาย ในวงการแพทยศาสตรศึกษา ข้อสอบชนิดนี้เป็นที่ชื่นชอบของนักศึกษาผู้เข้าสอบจำนวนมากเนื่องจากมีคำตอบให้เลือก หากไม่มีความรู้ก็สามารถเดาได้ ซึ่งต่างไปจากข้อสอบประเภทอัตนัยซึ่งผู้สอบต้องเขียนคำตอบจากความคิดของตนเอง^๑ ดังนั้นข้อสอบปรนัยจึงเป็นข้อสอบที่ผู้สอบทำได้ง่าย แต่ในทางตรงข้ามข้อสอบปรนัยเป็นข้อสอบที่สร้างปัญหาให้กับอาจารย์ผู้สร้างข้อสอบไม่น้อย เนื่องจากในกระบวนการเขียนข้อสอบปรนัยแต่ละข้อนั้นต้องใช้ทักษะอย่างมาก ต้องใช้ทั้งศาสตร์และศิลป์ และบ่อยครั้งอาจารย์ผู้สร้างข้อสอบก็ถูกขอให้ทำการปรับแก้ข้อสอบเนื่องจากคณะกรรมการพิจารณาข้อสอบมีความเห็นว่ารายละเอียดในข้อสอบไม่เหมาะสม มีการศึกษาวิจัยพบว่าคุณภาพของข้อสอบปรนัยที่พัฒนาขึ้นในโรงเรียนแพทย์หลายแห่งนั้นไม่ผู้ดีนัก มีข้อสอบที่มีลักษณะไม่เหมาะสมอยู่จำนวนไม่น้อย^{๒-๔} ข้อสอบปรนัยที่มีลักษณะไม่เหมาะสมเหล่านี้ส่งผลเสียต่อการสอบได้หลายประการ เช่น ทำให้ข้อสอบยากขึ้น สร้างความสับสนให้ผู้สอบ ทำให้ผู้สอบบางกลุ่มเสียเปรียบ และทำให้การตัดสินผลสอบผิดพลาด เป็นต้น^{๕-๘} ดังนั้นการออกข้อสอบปรนัยที่มีคุณภาพดีจึงเป็นงานที่มีความสำคัญและท้าทายความสามารถ

การสร้างข้อสอบปรนัยที่มีคุณภาพดีนั้นควรเริ่มต้นจากการมีองค์ความรู้พื้นฐานในการสร้างข้อสอบแล้ว เกิดการฝึกฝนทักษะ สังเกตประสบการณ์ในการออกข้อสอบ จนเกิดความชำนาญ ปัญหาที่พบบ่อยในโรงเรียนแพทย์หลายแห่งคือมีอาจารย์จำนวนไม่น้อยที่ได้รับมอบหมายให้ออกข้อสอบปรนัย โดยไม่ได้มีการพัฒนาองค์ความรู้พื้นฐานที่เหมาะสมก่อน ซึ่งเป็นเหตุให้มีข้อสอบปรนัยที่มีลักษณะไม่เหมาะสมตามหลักการออกข้อสอบปะปนมาในข้อสอบที่ให้นักศึกษาแพทย์และแพทย์ประจำบ้านทำอยู่บ้าง ผู้นิพนธ์จึงเห็นความสำคัญของการเผยแพร่องค์ความรู้พื้นฐานของการออกข้อสอบปรนัย องค์ความรู้พื้นฐานในการสร้างข้อสอบปรนัยนั้นมีสองส่วน ส่วนแรกเป็นหลัก การของการสร้างข้อสอบทั่วไปซึ่งได้มีผู้รวบรวมเป็นข้อแนะนำตีพิมพ์ในตำราและวารสารทางวิชาการอยู่บ้าง^{๑,๕-๘} ส่วนที่สองเป็นข้อผิดพลาดในการสร้างข้อสอบที่อาจารย์ผู้ออกข้อสอบพึงหลีกเลี่ยง ในบทความนี้ผู้นิพนธ์จะมุ่งเน้นในส่วนที่สองนี้ โดยจะรวบรวมข้อผิดพลาดในการสร้างข้อสอบปรนัย ที่อาจเป็นดั่งบ่วงไปให้ผู้สอบที่ไม่มีความรู้ในเรื่องที่ทำการทดสอบสามารถเลือกคำตอบที่ถูกต้องได้ ดังนั้นการที่อาจารย์ผู้ออกข้อสอบทราบถึงสิ่งเหล่านี้และหลีกเลี่ยงเสียจะส่งผลให้ข้อสอบปรนัยที่สร้างขึ้นสามารถใช้วัดองค์ความรู้ทางการแพทย์ได้จริง โดยปราศจากปัจจัยรบกวนจากการสังเกตพบสิ่งบอกรับคำตอบ

ข้อสอบปรนัยที่กล่าวถึงในบทความนี้มุ่งประเด็นไปที่ข้อสอบปรนัยชนิดเลือกคำตอบที่ถูกที่สุด (one best response) เป็นสำคัญ เนื่องจากเป็นข้อสอบที่ใช้กันแพร่หลายมากที่สุดในการวัดผลการศึกษาในโรงเรียนแพทย์ไทยปัจจุบัน ในข้อสอบชนิดนี้แต่ละข้อจะมีโจทย์ (stem) ตามด้วยตัวเลือก (options) จำนวน ๔-๕ ตัวเลือก ผู้สอบต้องเลือกคำตอบที่ถูกที่สุดเพียงคำตอบเดียวจากตัวเลือกเหล่านี้ ตัวเลือกอื่น ๆ ที่ไม่ใช่คำตอบเรียกว่าตัวลวง (distractors)

ในบทความนี้ผู้เขียนขอนำเสนอข้อผิดพลาดในการออกข้อสอบ ๗ กลุ่มด้วยกัน ได้แก่ (๑) ข้อผิดพลาดในไวยากรณ์, (๒) การใบ้คำตอบด้วยหลักตรรกะ, (๓) การใช้คำคุณศัพท์บอกระดับของความแน่ชัด, (๔) ความยาวของตัวเลือก, (๕) การใช้คำซ้ำในโจทย์และตัวเลือก, (๖) การเข้าพวกของคำหรือข้อความที่ปรากฏในตัวเลือก, และ (๗) การบอกใบ้คำตอบโดยโจทย์ข้ออื่น

๑. ข้อผิดพลาดในไวยากรณ์

ตัวเลือกทุกตัวต้องสามารถตอบโจทย์ได้อย่างถูกต้องตามหลักไวยากรณ์ บ่อยครั้งอาจารย์ผู้ออกข้อสอบมุ่งความสนใจไปที่คำตอบที่ถูก และให้ความสนใจกับตัวลวงน้อยไปจนทำให้ตัวลวงผิดหลักไวยากรณ์ โดยมักพบบ่อยในข้อสอบที่เป็นภาษาอังกฤษ ข้อผิดพลาดที่พบได้บ่อยเช่น ความไม่เข้ากันของ article (A, An, The) กับคำนามที่ตามหลัง, คำนามกับกริยาที่ไม่เข้ากันในเชิงเอกพจน์หรือพหูพจน์, การเติมคำในประโยคที่เว้นว่างไว้สำหรับเติมคำนามแต่ตัวลวงเป็นกริยาหรือเป็นคำนามในลักษณะที่ไม่เข้ากับรูปประโยค เป็นต้น

ตัวอย่างที่ ๑. A 70-year-old woman was brought in an emergency room with alteration of consciousness. Her vital signs were stable, but her Glasgow coma score was E1V1M3. After endotracheal intubation, the next step is to provide intravenous administration of ...

- A. lumbar puncture
- B. computerized scan of the brain
- C. glucose with Thiamine
- D. Sodium bicarbonate

ในตัวอย่างที่ ๑ นี้โจทย์ให้ผู้สอบเลือกตัวเลือกไปเติมในช่องว่าง ซึ่งสิ่งที่เติมลงในช่องว่างได้นั้นต้องเป็นยาที่สามารถให้ทางหลอดเลือดดำได้ ผู้สอบที่มีทักษะการทำข้อสอบดีจะตัดตัวเลือก A และ B ได้โดยไม่ต้องใช้ความรู้ทางการแพทย์

ตัวอย่างที่ ๒. Which organism is the cause of syphilis?

- A. *Neisseria gonorrhoea*
- B. *Chlamydia trachomatis* and *Giardia lamblia*
- C. *Treponema pallidum*
- D. *Ureaplasma urealyticum* and *Mycoplasma genitalium*

ในตัวอย่างที่ ๒ นี้โจทย์ถามหาเชื้อก่อโรค โดยใช้รูปประโยคถามหาคำตอบที่เป็นเอกพจน์ ดังนั้นคำตอบที่ต้องยอมมีเชื้อก่อโรคตัวเดียว ผู้สอบที่มีทักษะการทำข้อสอบดีจะตัดตัวเลือก B และ D ได้โดยไม่ต้องใช้ความรู้ทางการแพทย์

๒. การใบ้คำตอบด้วยหลักตรรกะ

ในการเขียนตัวเลือก อาจารย์ผู้ออกข้อสอบต้องระมัดระวังไม่ให้ผู้สอบสามารถตัดตัวเลือกได้ด้วยหลักตรรกศาสตร์ เนื่องจากผู้สอบที่มีทักษะการทำข้อสอบดีจะสามารถพิจารณาความเป็นไปได้ของตัวเลือกต่าง ๆ และตัดตัวลวงที่ไม่มีทางเป็นไปได้ตามหลักของเหตุและผลออกไปได้โดยไม่ต้องอาศัยความรู้เรื่องที่อาจารย์ตั้งเป้าหมายว่าจะทดสอบ

ตัวอย่างที่ ๓. ภาวะไส้เลื่อนบริเวณขาหนีบ (inguinal hernia)

- A. พบในผู้ชายบ่อยกว่าผู้หญิง
- B. พบในผู้หญิงบ่อยกว่าผู้ชาย
- C. พบเกิดขึ้นในผู้หญิงและผู้ชายในอัตราเท่ากัน
- D. พบบ่อยในผู้ที่มีเศรษฐกิจฐานะยากจน
- E. พบในผู้ที่มีภูมิลาเนาในทวีปเอเชีย มากกว่าผู้ที่มีภูมิลาเนาในทวีปยุโรป

ในตัวอย่างที่ ๓ นี้อาจารย์ผู้ออกข้อสอบต้องการวัดความรู้เรื่องอุบัติการณ์ของไส้เลื่อนขาหนีบ แต่หาก



พักรณาตามหลักตรรกศาสตร์แล้ว ตัวเลอก A, B, และ C เพียงสามตัวเลอกก็ครอบคลุทสิ่งที่เป็นไปได้ทั้งหมดแล้ว (เนือจกาทนุษย์มีสองเพศ ภาวะได้เลือนนี้หากไม่มีอัตรการเกิดเท่ากันในสองเพศแล้วก็ต้องมีเพศใดเป็นมากกว่าอีกเพศหนึ่ง) ดังนั้นผู้สอบที่มีทักษะการทำข้อสอบดีสามารถตัดตัวเลอก D และ E ได้โดยไม่ต้องมีความรู้เรื่องได้เลือนเลย

๓. การใช้คำคุณศัพท์บอกระดับของความแนชัด

อาจารย์ผู้ออกข้อสอบพึงระมัดระวังการใช้คำคุณศัพท์ที่บงบอกถึงความแนชัดของข้อความ ซึ่งจะมีหลายระดับ โดยทั่วไปแล้วคำคุณศัพท์ที่แสดงความแนชัดมาก แสดงความม่ใจมาก (เช่น always, never) มักไม่ถูกลง เนือจกาทนทางการแพทย์นั้นมีความไม่แนนอนเกิดขึ้นเป็นประจำ ข้อความที่บอกเล่าถึงสิ่งทีอาจเป็นไปได้ โดยไม่ชี้ชัดลงไปว่าต้องเกิดขึ้นแนนอน (เช่น may, might, can, could) มักเป็นข้อความที่ถูกลง

ตัวอย่างที่ ๔. Which of the following statements is true regarding the etiology of an inguinal hernia?

- A. Some connective tissue diseases may increase the incidence of inguinal hernia.
- B. Patients with Marfan syndrome always developed inguinal hernia.
- C. MRI scan of pelvis is the only reliable investigation for detection of groin hernia.
- D. Persistent lifting of heavy weights inevitably leads to the development of groin hernia.

ในตัวอย่างที่ ๔ นี้ผู้สอบต้องเลือกลงข้อความเกี่ยวกับได้เลือนขาหนีบที่ถูกลงหนึ่งข้อความ หากสังเกตดูทั้งสี่ข้อความมีการใช้คำคุณศัพท์บอกระดับของความแนชัดของข้อความ ได้แก่ may (ตัวเลอก A), always (ตัวเลอก B), the only (ตัวเลอก C), inevitably (ตัวเลอก D) ซึ่งจะเห็นว่าตัวเลอก B, C, และ D เป็นข้อความที่แสดงความแนชัดว่าต้องเป็นแน ต้องใช่แนนอน ไม่มีทางเลียงได้ ข้อความทำนองนี้มีโอกาสสูงทีจะผิด ในทางตรงข้ามตัวเลอก A เป็นข้อความบอกรว่ามีโอกาสเป็นไปได้โดยไม่ชี้ชัดว่าต้องเกิด

ผู้สอบที่มีทักษะการทำข้อสอบดีจะตัดตัวเลอก B, C, และ D ได้โดยไม่ต้องอาศัยความรู้ทางการแพทย์เลย

๔. ความยาวของตัวเลอก

มีการตั้งข้อสังเกตว่าอาจารย์แพทย์มักชอบสอนและอธิบายแม้กระทั่งในการสอบอาจารย์แพทย์หลายท่านก็ติดนิสัยรักการสอนนี้มาด้วย ทำให้อาจารย์มักเขียนตัวเลอกที่ถูกลงที่มีคำอธิบายประกอบอย่างครบถ้วนทำให้ตัวเลอกที่ถูกลงมีความยาวมากกว่าตัวลงวงักศึกษาผู้เข้าสอบจำนวนไม่น้อยรู้ถึงความจริงข้อนี้และมักเลือกลงตัวเลอกที่มีความยาวมากที่สุด หากเขาไม่สามารถหาคำตอบได้ด้วยความรู้ทางการแพทย์ทีเขามี

ตัวอย่างที่ ๕. ผู้หญิงอายุ ๒๘ ปี แต่งงานมานาน ๑ ปี ยังไม่มีบุตร คุณกำเนิดโดยการกินยาคุมเป็นประจำ สังเกตว่าตนเองนำหนักตัวเพิ่มขึ้นหลังจากกินยาคุมมาขอคำแนะนำเรื่องการคุมกำเนิด ท่านจะแนะนำอย่างไร

- A. ให้เปลี่ยนไปใช้การใส่ห่วงอนามัย
- B. ให้ใช้ถุงยางอนามัย
- C. ให้กินยาคุมกำเนิดต่อไปได้เนือจกามีการศึกษาแล้วว่ายาคุมกำเนิดชนิดกินไม่ส่งผลให้เกิดการเพิ่มขึ้นของนำหนักตัว
- D. ให้รับประทานยาลดความอ้วน

ในตัวอย่างที่ ๕ นี้จะสังเกตเห็นว่าตัวเลอก C มีการอธิบายเหตุผลประกอบส่งผลให้มีความยาวมากกว่าตัวเลอกอื่นชัดเจน ลักษณะเช่นนี้จะเป็นการบอกใบ้ให้นักศึกษาเลือกลงตัวเลอกนี้

๕. การใช้คำซ้ำในใจหทยและตัวเลอก

การใช้คำเดียวกัน หรือคำทีมีความหมายเหมือนกันในใจหทยและตัวเลอก มักเป็นการบอกใบ้ว่าตัวเลอกดังกล่าวเป็นตัวเลอกที่ถูกลง

ตัวอย่างที่ ๖. Which of the following statements is true regarding sacular theory of indirect inguinal hernia formation?

- A. An increased intra-abdominal pressure is the cause of inguinal hernia.
- B. A developmental diverticulum associated with a patent processus vaginalis is the cause of inguinal hernia.

C. All persons with a persistent processus vaginalis will develop an inguinal hernia.

D. A direct inguinal hernia is caused by the weakness of the posterior inguinal wall.

ในตัวอย่างที่ ๖ นี้ โจทย์ถามถึง saccular theory ซึ่งหากแปลความหมายก็น่าจะเป็นเรื่องที่เกี่ยวข้องกับถุง (sac) ผู้สอบที่มีทักษะการทำข้อสอบดีจะหาตัวเลือกที่มีคำที่มีความหมายเกี่ยวกับถุง แล้วเลือกตัวเลือกดังกล่าวทันที ซึ่งในที่นี่จะพบคำว่า diverticulum ซึ่งมีความหมายว่าถุงในข้อ B การที่มีคำที่มีความหมายซ้ำกันเช่นนี้เป็นตัวบ่งชี้คำตอบที่อาจารย์ผู้ออกข้อสอบต้องตรวจตราให้ดีก่อนนำข้อสอบไปใช้

๖. การเข้าพวของคำ หรือข้อความที่ปรากฏในตัวเลือก

ข้อสอบจำนวนไม่น้อยนำเสนอรายการของหลายอย่างในตัวเลือก (เช่น ชื่อการตรวจค้นเพิ่มเติม ชื่อโรค ชื่อยา ฯลฯ) มีผู้เชี่ยวชาญในการประเมินผลตั้งข้อสังเกตว่าในข้อสอบเหล่านี้ตัวเลือกที่ถูกต้องมักมีลักษณะเข้าพวกับตัวเลือกอื่นมากที่สุด หากเป็นรายการของตัวเลือกที่ถูกก็คือข้อที่มีจำนวนรายการซ้ำกับตัวเลือกอื่นมากที่สุด ดังนั้นในการนำเสนอตัวเลือกอาจารย์ผู้ออกข้อสอบพึงระมัดระวังอย่าให้ตัวเลือกที่ถูกต้องมีลักษณะที่เข้าพวได้อย่างชัดเจน พยายามทำตัวหลงอื่นให้มีลักษณะเข้าพวให้ใกล้เคียงกับตัวเลือกที่ถูกต้อง

ตัวอย่างที่ ๗. โรคที่แพทย์วินิจฉัยผิดว่าเป็นได้ตั้งอันดับที่สี่เรียงลำดับจากมากไปน้อยคือ

- A. acute mesenteric lymphadenitis, pelvic inflammatory disease, twisted ovarian cyst
- B. acute mesenteric lymphadenitis, Meckel diverticulitis, acute cholecystitis
- C. Meckel diverticulitis, twisted ovarian cyst, sigmoid diverticulitis
- D. pelvic inflammatory disease, acute gastroenteritis, right ureteric calculi

ในตัวอย่างที่ ๗ นี้ โจทย์ถามชื่อโรค ตัวเลือกแสดงรายการชื่อโรค ตัวเลือกละสามโรค หากนับจำนวนของคำซ้ำจะพบว่าโรคที่กล่าวถึงบ่อยที่สุดคือ acute

mesenteric lymphadenitis, pelvic inflammatory disease, twisted ovarian cyst, และ Meckel diverticulitis (กล่าวถึงโรคละ ๒ ครั้ง) ส่วนโรคที่เหลือกกล่าวถึงโรคละครั้งเดียว ดังนั้นตัวเลือกที่มีพวมากที่สุดคือตัวเลือก A ซึ่งเป็นคำตอบที่ถูกต้อง

การเข้าพวของตัวเลือกที่ถูกต้องนั้น ไม่จำเป็นต้องเป็นลักษณะของการมีจำนวน หรือความถี่ของคำมากที่สุดเพียงเท่านั้น อาจหมายรวมถึงการมีรูปร่างลักษณะ หรือความหมายคล้ายคลึงกันได้ด้วย

ตัวอย่างที่ ๘. ชายอายุ ๕๕ ปีเป็นมะเร็งเม็ดเลือดขาว หลังได้รับยาเคมีบำบัด ๑๔ วันมีไข้สูง ได้รับการวินิจฉัยเป็น febrile neutropenia การรักษาในข้อใดเหมาะสมที่สุด

- A. Amoxycillin PO
- B. Ceftazidime IV + Amikacin IV
- C. Amphotericin B-IV + Ceftazidime IV
- D. Cloxacillin IV + Metronidazole IV

ในตัวอย่างที่ ๘ นี้ โจทย์ถามถึงยาที่ควรให้กับผู้ป่วย ในตัวเลือกสี่ตัวเลือกนี้มียาเกินเพียงข้อเดียว (A) ที่เหลือเป็นยาฉีดสองขนานควบกัน ดังนั้นตัวเลือกข้อ A ไม่เข้าพว จะถูกตัดทิ้งได้โดยง่าย ในบรรดา ยาฉีดจะเห็นว่ามียาด้านเชื้อราที่ไม่เข้าพว (ตัวเลือก C) ดังนั้นจะเหลือตัวเลือกที่นักศึกษาต้องคิดเลือกจริง ๆ เพียงตัวเลือก B กับ D ซึ่งหากดูกลุ่มยา ก็จะพบว่ายาในกลุ่ม Cephalosporin เข้าพวมากที่สุด ทำให้ผู้สอบที่มีทักษะการทำข้อสอบดีสามารถเลือกคำตอบที่ถูกต้อง (ตัวเลือก B) ได้โดยไม่ต้องมีความรู้เรื่องการรักษาผู้ป่วย febrile neutropenia

๗. การบอไปคำตอบโดยโจทย์ข้ออื่น

ข้อผิดพลาดนี้เป็นข้อผิดพลาดที่ตัวผู้เขียนข้อสอบไม่ค่อยรู้ แต่ผู้ที่จะตรวจพบข้อผิดพลาดนี้คืออาจารย์ผู้เลือกข้อสอบไปใช้ เนื่องจากในการสอบแต่ละครั้งใช้ข้อสอบจำนวนมาก หากเลือกข้อสอบโดยไม่ระมัดระวังอาจมีข้อสอบสองข้อที่ถามเกี่ยวกับโรคหรือกลุ่มอาการเดียวกัน ซึ่งข้อมูลจากโจทย์ในข้อหนึ่งอาจเป็นตัวบอไปคำตอบของข้อสอบอีกข้อได้ ดังนั้นเมื่อทำการเลือกข้อสอบเสร็จแล้วจัดหน้ากระดาษเข้ารูปเล่มข้อสอบแล้วอาจารย์ควรอ่านข้อสอบฉบับสมบูรณ์นี้อีกหนึ่งหรือสองรอบก่อนส่ง

ไปพิมพ์ ซึ่งการอ่านทวนในขั้นตอนนี้อาจทำให้ตรวจพบข้อสอบที่มีเนื้อหาซ้ำซ้อนกันได้

ตัวอย่างที่ ๙. ผู้ป่วย febrile neutropenia มักมีไข้ขึ้นหลังจากได้รับยาเคมีบำบัดเป็นเวลากี่วัน

- A. 2 - 4 วัน
- B. 3 - 5 วัน
- C. 5 - 7 วัน
- D. 10 - 14 วัน

ในตัวอย่างที่ ๙ นี้อาจารย์ผู้ออกข้อสอบต้องการวัดความรู้ของผู้สอบเรื่อง febrile neutropenia ซึ่งเนื้อหาไปซ้ำซ้อนกับโจทย์ในตัวอย่างที่ ๘ ซึ่งผู้สอบที่มีทักษะการทำข้อสอบดีสามารถย้อนกลับไปอ่านโจทย์ในข้อก่อนหน้านี้อแล้วได้ข้อมูลว่าผู้ป่วยที่น่าเสนอว่าเป็น febrile neutropenia มีไข้ขึ้น ๑๔ วันหลังได้ยาเคมีบำบัด ก็สามารถตอบข้อสอบข้อนี้ถูกได้โดยง่าย

สรุป

ผู้นิพนธ์ได้รวบรวมข้อผิดพลาดในการสร้างข้อสอบปรนัยที่ผู้สอบอาจใช้เป็นแนวทางในการเลือกคำตอบที่ถูกต้องโดยไม่ต้องอาศัยความรู้ทางการแพทย์ที่อาจารย์ต้องการประเมินผล โดยเรียงเรียงเป็นเจ็ดกลุ่มข้อผิดพลาดด้วยกัน ผู้อ่านทุกท่านพึงตระหนักว่าสิ่งเหล่านี้ไม่ใช่หลักการทางวิทยาศาสตร์ที่ชัดเจนดังกฎทางคณิตศาสตร์หรือฟิสิกส์ หากแต่เป็นการรวบรวมข้อสังเกต

และคำแนะนำของผู้เชี่ยวชาญทางการวัดและประเมินผล จึงเป็นเพียงแนวทางเบื้องต้นในการพิจารณาตรวจสอบเนื้อหาของข้อสอบเท่านั้น การประยุกต์ใช้องค์ความรู้นี้คงต้องอาศัยศิลปะพอสมควรเพื่อที่จะได้ข้อสอบที่ดีสามารถวัดองค์ความรู้ทางการแพทย์ของนักศึกษาหรือแพทย์ประจำบ้านที่เข้าสอบได้ตามวัตถุประสงค์ของการสอบ

เอกสารอ้างอิง

1. Haladyna TM. Developing and validating multiple-choice test items, 3rd ed. Mahwah, NJ: Lawrence Erlbaum Associates, 2004.
2. Jozefowicz RF, Koeppen BM, Case S, Galbraith R, Swanson D, Glew RH. The quality of in-house medical school examinations. Acad Med. 2002;77:156-61.
3. Tarrant M, Ware J. Impact of item-writing flaws in multiple-choice questions on student achievement in high-stakes nursing assessments. Med Educ. 2008;42:198-206.
4. Downing SM. The effects of violating standard item writing principles on tests and students: The consequences of using flawed test items on achievement examinations in medical education. Adv Health Sci Educ Theory Pract. 2005;10:133-43.
5. Haladyna TM, Downing SM. A taxonomy of multiple-choice item-writing rules. Appl Meas Educ. 1989;2:37-50.
6. Haladyna TM, Downing SM. Validity of a taxonomy of multiple-choice item-writing rules. Appl Meas Educ. 1989;2:51-78.
7. Haladyna TM, Downing SM, Rodriguez MC. A review of multiple-choice item-writing guidelines for classroom assessment. Appl Meas Educ. 2002;15:309-34.
8. Case SM, Swanson D. Constructing written test questions for the basic and clinical sciences, 3rd ed. Philadelphia, PA: National Board of Medical Examiners, 2002.