

การสร้างข้อสอบปรนัย

ผู้ช่วยศาสตราจารย์ นายแพทย์เชิดศักดิ์ โสมณรัตน์

ภาควิชาศัลยศาสตร์, คณะแพทยศาสตร์ศิริราชพยาบาล, มหาวิทยาลัยมหิดล, กรุงเทพมหานคร ๑๐๗๑๐๐.

ข้อสอบปรนัย (multiple-choice question) เป็นรูปแบบการประเมินผลที่นิยมใช้กันอย่างแพร่หลายในวงการแพทยศาสตรศึกษาเนื่องจากคุณสมบัติที่ดีหลายประการด้วยกัน ได้แก่ ประสิทธิภาพในการประเมินความรู้ปริมาณมากในเวลาอันสั้น ผลการประเมินที่ไม่มีผลกระทบจากความรู้สึกส่วนตัวของผู้ตรวจให้คะแนน คะแนนที่มีความเที่ยงสูง รวมถึงผลการวิจัยจำนวนมากที่สนับสนุนความถูกต้องของผลการประเมินด้วยข้อสอบปรนัย^{๑-๓} ข้อสอบปรนัยที่พัฒนาขึ้นอย่างดีนั้นสามารถวัดความรู้ได้ทั้งระดับการจดจำ การทำความเข้าใจ และการประยุกต์ความรู้ไปใช้ในการดูแลคนไข้^{๓-๕} อย่างไรก็ตาม การศึกษาวิจัยเกี่ยวกับคุณภาพของข้อสอบปรนัยที่พัฒนาขึ้นใช้ในโรงเรียนแพทย์หลายแห่งพบว่าข้อสอบจำนวนไม่น้อยมีลักษณะที่ไม่เหมาะสม^{๓-๖} ข้อสอบปรนัยที่ถูกพัฒนาขึ้นอย่างไม่ถูกหลักการนั้นส่งผลเสียหลายอย่าง เช่น ทำให้ข้อสอบยากขึ้นโดยไม่จำเป็น ทำให้ผู้สอบเกิดความสับสน ทำให้ผู้สอบบางกลุ่มเสียเปรียบผู้สอบคนอื่น ทำให้การตัดสินผลสอบผิดพลาด เป็นต้น^{๖-๗} ดังนั้นการออกข้อสอบปรนัยที่ดี วางอยู่บนหลักการที่ถูกต้องจึงมีความสำคัญมากในการควบคุมคุณภาพการศึกษาในโรงเรียนแพทย์ บทความนี้จะจึงถูกเขียนขึ้นเพื่อเป็นการรวบรวมหลักการพื้นฐานในการออกข้อสอบปรนัยที่ได้รับการยอมรับกันทั่วไปในวงการวัดและประเมินผล ผู้มีพันธะหวังว่าข้อแนะนำต่าง ๆ ที่ได้นำเสนอในบทความนี้จะเป็แนวทางที่เป็นประโยชน์ในการพัฒนาข้อสอบปรนัยที่มีคุณภาพให้ผู้อ่านไม่มากก็น้อย

รูปแบบพื้นฐานของข้อสอบปรนัย

ข้อสอบปรนัยคือข้อสอบชนิดที่มีคำถามแล้วมีตัวเลือกให้ผู้สอบเลือกตัวเลือกที่เหมาะสมเพื่อตอบคำถามดังกล่าว ข้อสอบปรนัยสามารถแบ่งออกได้เป็น ๒ รูปแบบ^๘ ได้แก่

๑. ข้อสอบถูกผิด (True/false item)

ในข้อสอบประเภทนี้จะมีข้อความให้ผู้สอบพิจารณาว่าถูกหรือผิด ในยุคแรกข้อสอบเหล่านี้แต่ละข้อจะแยกเป็นอิสระจากกัน ผู้สอบตัดสินใจว่าข้อความแต่ละข้อถูกหรือผิดโดยไม่เกี่ยวข้องกับข้อความในข้ออื่น ต่อมาผู้พัฒนาข้อสอบเป็นชุดของข้อความ (multiple true/false หรือ K-type item) โดยในแต่ละข้อจะมีสี่ข้อความ ผู้สอบต้องพิจารณาว่าแต่ละข้อความถูกหรือผิด แล้วทำการเลือกตัวเลือกที่บรรยายจำนวนข้อความที่ถูกต้องได้อย่างเหมาะสม (เช่น ตอบ ก. เมื่อข้อความที่ ๑, ๒, และ ๓ ถูกต้อง, ตอบ ข. เมื่อข้อความที่ ๑ และ ๓ ถูกต้อง ฯลฯ)

ข้อสอบชนิดถูกผิดนี้เคยเป็นที่นิยมมากในการแพทยศาสตรศึกษาอยู่ระยะหนึ่งเนื่องจากสามารถทดสอบความรู้ได้ปริมาณมาก แต่ข้อสอบชนิดนี้มีข้อจำกัดที่สำคัญคือสามารถใช้ได้เฉพาะกับเนื้อหาที่มีความถูกต้องชัดเจนเท่านั้น ซึ่งการตัดสินใจทางการแพทย์ส่วนมากไม่เป็นเช่นนั้น การตัดสินใจในการวินิจฉัย การตรวจค้นเพิ่มเติม หรือการรักษาผู้ป่วยส่วนใหญ่นั้นแพทย์ตัดสินใจเลือกระหว่างทางเลือกที่แตกต่างกันสามสี่อย่างซึ่งทุกทางเลือกมีความเป็นไปได้ มีส่วนถูก หรือมีความเหมาะสมในบางด้าน



แต่ก็มีความไม่เหมาะสมในด้านอื่นด้วย เช่นการเลือกใช้ยาในผู้ป่วยที่มีการติดเชื้อ นักศึกษาแพทย์มักคิดว่าควรใช้ยาปฏิชีวนะ ซึ่งยาปฏิชีวนะหลายชนิดก็รักษาการติดเชื้อชนิดนั้น ๆ ได้ แต่นักศึกษาต้องเลือกกระหว่างยาที่ล้นใช้ได้ในการรักษานั้นว่ายาใดที่มีประสิทธิภาพสูงสุด เหมาะสมที่สุดกับชนิดของเชื้อก่อโรคที่พบป่วยในการติดเชื้อนั้น มีผลข้างเคียงน้อยที่สุด และราคาเหมาะสมด้วย ซึ่งในสถานการณ์นี้ข้อสอบชนิดถูกผิดจะนำมาใช้ได้ยาก ด้วยเหตุนี้ทำให้ข้อสอบชนิดถูกผิดไม่เป็นที่นิยมกันมากนักในปัจจุบัน

๒. ข้อสอบเลือกคำตอบที่ถูกที่สุด (one best response item)

ในข้อสอบประเภทนี้จะมีคำถามแล้วตามด้วยตัวเลือกจำนวนหนึ่งให้ผู้สอบเลือกตัวเลือกที่เหมาะสมที่สุดเป็นคำตอบ ข้อสอบประเภทนี้ที่เป็นที่นิยมกันมากที่สุดคือข้อสอบที่มีตัวเลือก ๔-๕ ตัวเลือก (A-type) แต่นอกจากข้อสอบมาตรฐานนี้แล้วก็มีผู้ใช้ข้อสอบประเภทที่มีลักษณะเป็นการจับคู่ (extended matching item) โดยให้ผู้สอบเลือกตัวเลือกที่เหมาะสม (จากตัวเลือกจำนวนมาก ๘-๒๐ ตัวเลือก) ไปจับคู่กับโจทย์ (stem) ซึ่งมีหลายข้อ เช่นจับคู่ระหว่างคำบรรยายอาการของผู้ป่วยจำนวน ๕-๑๐ ราย กับการวินิจฉัยโรคที่เหมาะสม จำนวน ๑๕ โรค เป็นต้น

เนื่องจากข้อสอบชนิดที่มีใช้กันแพร่หลายในวงการแพทยศาสตรศึกษาในประเทศไทยในปัจจุบันคือข้อสอบประเภทที่มีตัวเลือก ๔-๕ ตัวเลือก (A-type) ผู้นิพนธ์จะขอเน้นหลักการสำหรับการออกข้อสอบประเภทนี้เป็นสำคัญ

องค์ประกอบของข้อสอบปรนัยชนิดเลือกคำตอบที่ถูกที่สุด

ข้อสอบปรนัยแต่ละข้อมีส่วนประกอบสำคัญ ๒ ส่วนด้วยกันคือ

๑. โจทย์ (stem) เป็นข้อมูลของโรค หรือภาวะ หรือผู้ป่วยตามด้วยคำถาม หรือเว้นช่องว่างสำหรับเติมคำ หรือข้อความที่เหมาะสมลงไป

๒. ตัวเลือก (options) คือคำ หรือข้อความที่

ผู้ออกข้อสอบนำเสนอตามหลังจากโจทย์เพื่อให้ผู้สอบเลือกไปใช้ตอบคำถาม หรือเติมลงในช่องว่างในโจทย์

๒.๑ ตัวเลือกที่ถูกต้อง (correct option) เป็นคำตอบที่ถูกต้องมีเพียงตัวเลือกเดียวต่อข้อสอบข้อหนึ่ง

๒.๒ ตัวลวง (distractors) เป็นคำตอบที่ผิด หรือไม่เหมาะสม มีไว้ลวงให้ผู้สอบที่ไม่มีความรู้ หรือมีความเข้าใจไม่ถูกต้องในเนื้อหาที่นำมาออกข้อสอบเลือกตอบ ตัวลวงไม่จำเป็นต้องเป็นคำตอบที่ผิดชัดเจนเสมอไป ตัวลวงที่ดีมักมีส่วนถูกบ้าง แต่มีระดับของความถูกต้องเหมาะสมน้อยกว่าคำตอบที่ถูก

ข้อแนะนำพื้นฐานของการเขียนข้อสอบปรนัย

มีผู้เชี่ยวชาญทางการประเมินผลให้ข้อแนะนำจำนวนมากในการเขียนข้อสอบปรนัย เคยมีผู้รวบรวมไว้ถึง ๔๓ ข้อ^{๒-๔} ในที่นี้ผู้นิพนธ์ขอนำเสนอเฉพาะข้อแนะนำที่ได้รับการยอมรับอย่างกว้างขวางและสามารถประยุกต์ใช้ได้ชัดเจนในการพัฒนาข้อสอบทางการแพทย์ โดยจะทำการจัดหมวดหมู่ของข้อแนะนำเหล่านี้ออกเป็น ๔ กลุ่มด้วยกัน ได้แก่ (๑) เนื้อหาข้อสอบ, (๒) การจัดรูปแบบข้อสอบ, (๓) การเขียนโจทย์, และ (๔) การเขียนตัวเลือก

๑. เนื้อหาข้อสอบ

๑.๑ ข้อสอบหนึ่งข้อควรมุ่งเน้นประเมินความรู้เพียงเรื่องเดียว

ก่อนเริ่มเขียนข้อสอบอาจารย์ผู้ออกข้อสอบควรตั้งวัตถุประสงค์ให้ชัดเจนว่าต้องการประเมินความรู้ของผู้สอบในเรื่องใด และเขียนโจทย์เพื่อตอบสนองวัตถุประสงค์ดังกล่าวเท่านั้น เนื่องจากเนื้อหาวิชาทางการแพทย์มีมาก อาจารย์แต่ละท่านเมื่อทำการสอนไปแล้วจึงอยากจะทดสอบความรู้ในหลายเรื่องที่ได้สอนไป แต่กลับมีข้อจำกัดในการออกข้อสอบ ทำให้อาจารย์จำนวนไม่น้อยเขียนข้อสอบหนึ่งข้อถามทั้งเรื่องการวินิจฉัยโรค การตรวจค้นเพิ่มเติม การรักษาโรค และ ภาวะแทรกซ้อนของโรค ไปพร้อมกัน ลักษณะข้อสอบเช่นนี้ไม่ควรใช้ เพราะมักซับซ้อนเกินไป เมื่อผู้สอบตอบข้อสอบผิด ก็ไม่สามารถวินิจฉัยได้ว่าผู้สอบขาดความรู้ ความเข้าใจในเรื่องใด

๑.๒ หลีกเลี่ยงการถามความรู้ในรายละเอียดปลีกย่อยที่ไม่มีที่ใช้ทางคลินิก (trivial content)



องค์ความรู้ทางการแพทย์นั้นมีปริมาณมาก ไม่มีผู้ใดที่จดจำเนื้อหาที่มีในตำรา หรือวารสารทางการแพทย์ได้ทั้งหมด แม้ว่าองค์ความรู้หลายเรื่องมีความน่าสนใจ แต่มีประโยชน์ในการประยุกต์ใช้ทางคลินิกค่อนข้างน้อย องค์ความรู้ดังกล่าวจัดเป็นรายละเอียดปลีกย่อย (trivial content) ซึ่งไม่แนะนำให้ทำการทดสอบ สิ่งที่เราควรทำการประเมินคือความสามารถในการประยุกต์ใช้ความรู้ในทางคลินิก (application of knowledge) ไม่แนะนำการทดสอบวัดความสามารถในการจดจำเป็นหลัก อย่างไรก็ตามที่แนะนำให้ออกข้อสอบที่เน้นการประยุกต์ใช้ความรู้ ไม่ได้หมายความว่าความจำเป็นที่ผู้ป่วยนั้นต้องใช้ความจำเลย ตรงกันข้ามการจดจำเนื้อหาเป็นพื้นฐานที่สำคัญในการแก้ปัญหาทางคลินิก ผู้สอบย่อมต้องจำเนื้อหาได้บ้าง จึงจะประยุกต์องค์ความรู้ดังกล่าวไปแก้โจทย์ปัญหาที่นำเสนอได้

๑.๓ หลักเลียงการถามความรู้ในเรื่องที่ยังมีความขัดแย้งกันในแนวทางปฏิบัติ (controversy)

ความรู้ทางการแพทย์ในหลายหัวข้อยังเป็นเรื่องที่ยังมีผู้เชี่ยวชาญมีความเห็นแตกต่างกัน ผู้ป่วยรายเดียวกันไปพบแพทย์สองคนอาจได้รับการรักษาที่แตกต่างกันซึ่งวิธีการรักษาทั้งสองวิธีก็มีความจำเป็นสนับสนุนด้วยกันทั้งคู่ อย่างนี้จัดว่ายังมีความขัดแย้ง (controversy) ในเรื่องดังกล่าวอยู่ เนื้อหาในลักษณะนี้ไม่ควรนำมาออกสอบด้วยข้อสอบปรนัย เนื่องจากในขณะที่ทำข้อสอบอยู่นั้น ผู้สอบไม่มีทางรู้ได้เลยว่าอาจารย์ผู้ออกข้อสอบอ้างอิงจากตำราหรือบทความวิชาการใด เนื้อหาที่ยังมีความขัดแย้ง ที่ผู้เชี่ยวชาญจากต่างสถาบันมีแนวทางในการปฏิบัติที่ต่างกันนั้นแนะนำให้ข้อสอบในรูปแบบอื่นในการทดสอบเช่นข้อสอบอัตนัย เป็นต้น

๑.๔ หลักเลียงการลอกประโยคหรือข้อความจากตำราโดยตรง

ดังได้กล่าวแล้วว่าข้อสอบที่ดีควรมุ่งเน้นการประเมินความเข้าใจ หรือ การประยุกต์ใช้ความรู้ ไม่ควรออกข้อสอบที่ประเมินความสามารถในการจำรายละเอียดปลีกย่อย การออกข้อสอบโดยวิธีการเปิดตำราแล้วคัดลอกประโยคจากตำราโดยตรงมักจะลงเอยด้วยข้อสอบที่ทดสอบความจำว่าผู้สอบท่องเนื้อหาในตำราตรงส่วนนั้นได้หรือไม่

ข้อสอบที่ดีควรได้จากการดูผู้ป่วย โจทย์ที่ดีควรเป็นปัญหาของผู้ป่วยที่พบในการทำงานนั่นเอง ตัวเลือกก็ได้จากข้อผิดพลาดที่นักศึกษาหรือแพทย์ประจำบ้านมักปฏิบัติกับผู้ป่วยแล้วทำให้ผลการรักษาไม่ดีขึ้นเอง

๑.๕ หลักเลียงการนำเสนอข้อสอบที่ประเมินความรู้ในเรื่องเดียวกันสองข้อในข้อสอบชุดเดียวกัน

เนื่องจากเนื้อหาวิชาที่ต้องทำการประเมินในการสอบแต่ละครั้งนั้นมีมาก ดังนั้นองค์ความรู้ในแต่ละเรื่องแต่ละโรคจึงมักมีสัดส่วนของข้อสอบที่จะออกได้เพียงหนึ่งหรือสองข้อเท่านั้น การที่อาจารย์ออกข้อสอบในเรื่องหรือโรคเดียวกันซ้ำสองข้อในชุดข้อสอบเดียวกันจึงมักเป็นการลดโอกาสในการประเมินความรู้เรื่องอื่นซึ่งก็มีความสำคัญเช่นกัน การออกข้อสอบที่ได้นั้นควรต้องครอบคลุมวัตถุประสงค์การเรียนรู้ตามที่กำหนดในหลักสูตร หรือในเกณฑ์มาตรฐานผู้ประกอบวิชาชีพเวชกรรมอย่างสมดุล การที่จะบรรลุเป้าหมายดังกล่าวได้นั้นต้องเริ่มต้นจากการกำหนดสัดส่วนข้อสอบสร้างเป็นตารางกำหนดจำนวนข้อสอบ (table of specification) เมื่ออาจารย์ได้รับมอบหมายให้ออกข้อสอบควรต้องตรวจสอบให้ชัดเจนว่าเนื้อหาที่ต้องออกข้อสอบนั้นอยู่ในส่วนใดของตารางดังกล่าว การออกข้อสอบซ้ำซ้อนในเนื้อหาเรื่องเดียวกันเป็นสัญญาณบอกว่าอาจไม่ได้สร้างข้อสอบตามข้อกำหนดในตาราง นอกจากนี้การมีโจทย์สองข้อประเมินความรู้เรื่องเดียวกันมีความเป็นไปได้สูงที่เนื้อหาในข้อสอบข้อหนึ่งอาจบอกคำตอบในข้อสอบอีกข้อหนึ่งได้

๒. การจัดรูปแบบข้อสอบ

๒.๑ เลือกใช้คำศัพท์หรือรูปแบบประโยคที่ง่ายต่อการทำความเข้าใจ

อาจารย์ผู้ออกข้อสอบต้องระลึกไว้เสมอว่าข้อสอบที่อาจารย์ออกเพื่อใช้ในการประเมินผลนักศึกษาแพทย์หรือแพทย์ประจำบ้านนั้นมีวัตถุประสงค์เพื่อทดสอบความรู้ทางการแพทย์เป็นสำคัญ มิใช่การประเมินความรู้ทางภาษาศาสตร์ ดังนั้นการเขียนข้อสอบของอาจารย์ควรเลือกใช้รูปแบบประโยคที่ง่ายต่อการทำความเข้าใจ อย่าเขียนประโยคซับซ้อนที่มีความยาวประโยคหลายบรรทัด มุ่งเน้นให้ภาษาเป็นสื่อในการนำเสนอความคิดของอาจารย์ผู้ออกข้อสอบไปยังผู้สอบ อย่าให้

ภาษาเป็นอุปสรรคในการสื่อสาร การจะเลือกใช้ภาษาใดในการเขียนข้อสอบนั้นให้พิจารณาตามข้อกำหนดขององค์กรหรือหน่วยงานที่ควบคุมการสอบที่อาจารย์ส่งข้อสอบไปให้ใช้ ข้อสอบที่ใช้ในระดับการศึกษาหลักสูตรแพทยศาสตรบัณฑิตทั้งในระดับคณะ หรือข้อสอบที่ใช้ในการสอบระดับประเทศในปัจจุบันยังนิยมใช้ข้อสอบที่เขียนด้วยภาษาไทยโดยมีการใช้ศัพท์เทคนิคเป็นภาษาอังกฤษเหมือนดังภาษาที่แพทย์ใช้สื่อสารกันในการทำงานปกติ ส่วนข้อสอบในระดับหลังปริญญามีหลายการสอบที่ภาควิชา หรือราชวิทยาลัยที่เกี่ยวข้องกำหนดให้ใช้ภาษาอังกฤษทั้งหมด ก่อนที่อาจารย์จะสร้างข้อสอบต้องมีการศึกษาข้อกำหนดของแต่ละการสอบให้ดี

๒.๒ หลักเลี่ยงการนำเสนอข้อมูลที่ไม่เกี่ยวข้องกับการแก้ปัญหาของโจทย์ข้อนั้น

โจทย์แต่ละข้อควรเขียนให้กระชับ ไม่ยาวเยิ่นเย้อโดยไม่จำเป็น นำเสนอเฉพาะข้อมูลที่จำเป็นในการแก้ปัญหาโจทย์ดังกล่าว อาจารย์บางท่านนำเสนอข้อมูลเยอะมากในโจทย์หนึ่งข้อ บางครั้งข้อสอบข้อหนึ่งมีความยาวถึงครึ่งหน้า โดยให้เหตุผลว่าเป็นเหมือนสถานการณ์จริงที่แพทย์ต้องตัดสินใจบนข้อมูลทางคลินิกปริมาณมาก แพทย์ต้องพิจารณาเองว่าข้อมูลใดสำคัญกับการแก้ปัญหาโจทย์ข้อนั้น ๆ แต่อาจารย์ก็ต้องไม่ลืมว่าเวลาที่ผู้สอบมีในการทำข้อสอบแต่ละข้อนั้นมีจำกัด ในการสอบทางการแพทย์ในประเทศไทยส่วนใหญ่ผู้สอบจะมีเวลาราว ๑ นาทีในการทำข้อสอบ ๑ ข้อ หากเนื้อหาโจทย์ข้อใดมีความยาวมาก ผู้สอบจำนวนไม่น้อยจะเลือกที่จะข้ามข้อสอบข้อนั้นไปก่อนด้วยเกรงว่าจะเสียเวลาอ่านและคิดแก้ปัญหาในข้อนั้นนานเกินไปทำให้ทำข้อสอบไม่ทัน ดังนั้นหากอาจารย์ต้องการให้ข้อสอบที่อาจารย์เขียนขึ้นมานั้นได้ถูกใช้จริง และผู้เข้าสอบได้คิดแก้ปัญหาจริงในการสอบ ไม่ถูกอ่านข้ามไป อาจารย์ควรเขียนข้อสอบให้กระชับ ไม่นำเสนอข้อมูลที่ไม่เกี่ยวข้องกับการแก้ปัญหา

๒.๓ จัดให้มีการตรวจสอบเนื้อหา คำศัพท์ และรูปประโยคที่ใช้ในข้อสอบแต่ละข้อก่อนนำไปใช้

ถึงแม้ว่าอาจารย์ผู้เขียนข้อสอบจะได้มีการอ่านทวนสิ่งที่ตนเองเขียนแล้วเข้าใจเนื้อหาได้ดีและคิดว่าข้อสอบอยู่ในรูปแบบที่สามารถนำไปใช้ได้แล้ว ก็ไม่ควร

นำข้อสอบข้อนั้นไปใช้สอบเลย ควรให้มีคณะกรรมการข้อสอบซึ่งประกอบไปด้วยอาจารย์หลายท่านช่วยกันตรวจสอบและพิจารณาปรับแก้ข้อสอบทุกข้อก่อนนำไปใช้จริงเสมอ เนื่องจากผู้เขียนข้อสอบย่อมเข้าใจสิ่งที่ตนเขียน แต่เมื่อผู้อื่นอ่านแล้วอาจพบว่ามีเนื้อหาที่กำกวมหรือเข้าใจโจทย์ต่างออกไปได้ การปรับแก้เนื้อหาที่มีความกำกวม หรือเฉลยซึ่งอาจารย์บางท่านอาจไม่เห็นด้วยให้ได้ข้อสอบที่มีความชัดเจน และอาจารย์ทุกท่านยอมรับในค่าเฉลี่ยได้ก่อนจะนำข้อสอบไปทำการสอบจริงย่อมเป็นสิ่งที่ดีกว่าการตรวจพบปัญหาหลังจากสอบเสร็จแล้วซึ่งต้องมาตัดสินใจกันอีกว่าจะทำอย่างไรกับการคิดคะแนนของข้อสอบข้อดังกล่าว

๓. การเขียนโจทย์

๓.๑ เขียนโจทย์ให้มีความชัดเจน ผู้สอบทุกคนอ่านแล้วมีความเข้าใจตรงกัน

ข้อแนะนำนี้อาจดูเหมือนตรงไปตรงมา แต่กลับเป็นปัญหาที่พบบ่อยมากในการพัฒนาข้อสอบปรนัย ประเด็นสำคัญคือโจทย์ที่ดีนั้นต้องมีความสมบูรณ์ในตัวเองโดยไม่ต้องอาศัยตัวเลือก โจทย์ข้อสอบที่ดีนั้นเมื่ออ่านโจทย์เสร็จแล้ว หากผู้สอบมีความรู้ในเรื่องที่ทำการประเมินนั้น เขาจะบอกคำตอบได้โดยไม่ต้องอ่านตัวเลือกเลย ดังนั้นเมื่ออาจารย์เขียนข้อสอบเสร็จแล้วแนะนำให้ลองปิดตัวเลือกแล้วอ่านเฉพาะโจทย์ดู หากอาจารย์อ่านแล้วบอกได้ว่าโจทย์ถามอะไรและบอกได้ว่าควรตอบอะไรโดยไม่ต้องอ่านตัวเลือกจัดว่าข้อสอบข้อดังกล่าวมีโจทย์ที่มีความชัดเจน

๓.๒ เรียบเรียงเนื้อหาให้ใจความสำคัญของข้อสอบอยู่ในโจทย์

เนื่องจากข้อสอบปรนัยมีตัวเลือกที่อาจารย์ต้องสร้างขึ้นหลายตัวเลือก บางครั้งอาจารย์ผู้พัฒนาข้อสอบอาจเผลอเรอเอาใจความสำคัญไปใส่ไว้ในตัวเลือกซึ่งทำให้เนื้อหาในโจทย์ขาดสาระสำคัญ อ่านโจทย์แล้วไม่เข้าใจว่าผู้ออกข้อสอบต้องการถามความรู้เรื่องอะไร ตัวอย่างข้อสอบที่ไม่เป็นไปตามข้อแนะนำนี้คือข้อสอบที่ถามว่า ข้อใดต่อไปนี้ถูกต้อง หรือข้อใดต่อไปนี้ไม่ถูกต้อง แล้วเขียนรายละเอียดเกี่ยวกับโรค หรือการรักษาบางอย่างในตัวเลือกแต่ละข้อ ข้อสอบในลักษณะนี้มักทำให้

ผู้สอบต้องอ่านข้อสอบย้อนไปมาหลายรอบกว่าจะเข้าใจจุดประสงค์ของข้อสอบ แล้วจึงตัดสินใจเลือกคำตอบโดยทั่วไปแนะนำให้อาจารย์นำเสนอรายละเอียดต่าง ๆ ไว้ในตัวโจทย์ให้มากที่สุด ส่วนตัวเลือกเขียนเป็นคำหรือข้อความสั้น ๆ

๓.๓ หลักเขียนการเขียนโจทย์ที่มีรูปประกอบเป็นเชิงปฏิเสธ

โจทย์ที่ดีไม่ควรอยู่ในประโยคเชิงปฏิเสธ เช่นถามถึงสิ่งที่เป็นข้อยกเว้น สิ่งที่ไม่ควรปฏิบัติ สิ่งที่ไม่พบน้อยที่สุด หรือสิ่งที่ไม่แน่นอนถึงเป็นต้น งานวิจัยส่วนใหญ่พบว่าข้อสอบที่มีโจทย์ในรูปแบบปฏิเสธเหล่านี้มีระดับความยากง่ายไม่ต่างจากข้อสอบอื่น ๆ แต่งานวิจัยบางชิ้นพบว่าข้อสอบที่มีโจทย์ในรูปแบบปฏิเสธมีความยากมากกว่าข้อสอบอื่นชัดเจนโดยเฉพาะในข้อสอบวัดความรู้ระดับสูง^{๓๓} แต่ผู้เชี่ยวชาญในการประเมินผลส่วนใหญ่มีความเห็นพ้องกันว่าข้อสอบประเภทนี้สามารถสร้างความสับสนให้กับผู้สอบได้ จึงไม่แนะนำให้ใช้ แต่หากอาจารย์ผู้ออกข้อสอบมีความจำเป็นต้องใช้ข้อสอบที่มีการใช้คำปฏิเสธในโจทย์ แนะนำให้พิมพ์คำปฏิเสธให้เด่นชัด โดยใช้ตัวหนาและขีดเส้นใต้เพื่อให้ผู้สอบเห็นชัด^{๓๔}

๔. การเขียนตัวเลือก

๔.๑ เขียนตัวเลือกที่มีประสิทธิภาพให้มีจำนวนมากที่สุดเท่าที่เหมาะสมกับบริบท

เรื่องจำนวนตัวเลือกที่เหมาะสมนี้เป็นเรื่องที่ผู้เชี่ยวชาญด้านการประเมินผลจำนวนมากสนใจ มีงานวิจัยเกี่ยวกับเรื่องจำนวนตัวเลือกที่เหมาะสมในข้อสอบปรนัยอยู่มากมาย^{๓๕} อาจารย์ผู้ออกข้อสอบส่วนมากจะคุ้นเคยกับข้อสอบปรนัยชนิดที่มีห้าตัวเลือก บ่อยครั้งที่อาจารย์ออกข้อสอบแล้วนึกตัวเลือกได้เพียงสามหรือสี่ตัว จึงเกิดคำถามว่าจำเป็นต้องมีตัวเลือกครบห้าตัวเลือกหรือไม่ งานวิจัยบางชิ้นพบว่าการลดจำนวนตัวเลือกลงทำให้ข้อสอบง่ายขึ้น^{๓๖-๓๘} แต่งานวิจัยบางชิ้นพบว่าการลดจำนวนตัวเลือกลงทำให้ได้ข้อสอบยากขึ้น^{๓๙-๔๐} ผู้เชี่ยวชาญในการประเมินผลเสนอว่าข้อสอบปรนัยที่มีตัวเลือกเพียงสามตัวเลือกก็สามารถทดสอบความรู้ได้อย่างมีประสิทธิภาพ^{๔๑, ๔๒, ๔๓} แต่มีอาจารย์จำนวนไม่น้อยที่ไม่สบายใจที่มีตัวเลือกในข้อสอบแต่ละข้อน้อยกว่าห้าตัว

เลือกด้วยกังวลว่าจะทำให้มีโอกาสสูงที่ผู้สอบที่ไม่มีความรู้จะเดาสุ่มได้คำตอบที่ถูกต้อง แต่จากข้อมูลที่ปรากฏในปัจจุบันพบว่าผู้สอบในการสอบในระดับสูงนั้นพฤติกรรมเดาสุ่มโดยที่ผู้สอบปราศจากความรู้ไม่น่าจะมีพบนาน้อยมาก ผู้สอบส่วนใหญ่มักพอมีความรู้บ้างและสามารถตัดตัวเลือกที่ไม่สมเหตุสมผลอย่างชัดเจนได้^{๔๔} ในการศึกษาข้อสอบปรนัยส่วนใหญ่พบตัวเลือกที่ไม่ทำงานเป็นจำนวนไม่น้อย^{๔๕} ข้อมูลที่ได้จากการวิเคราะห์ข้อสอบปรนัยที่ใช้ในทางแพทยศาสตรศึกษาในประเทศไทยหลายครั้งก็สอดคล้องกับงานวิจัยในต่างประเทศที่พบว่าข้อสอบส่วนใหญ่ก็มีตัวเลือกที่ทำงานจริงราวสามหรือสี่ตัวเลือก มีข้อสอบน้อยข้อสอบที่ตัวเลือกทั้งห้าตัวเลือกทำงานอย่างมีประสิทธิภาพ

ด้วยข้อมูลจากการศึกษาต่าง ๆ ข้อแนะนำในการออกข้อสอบปรนัยในปัจจุบันคือให้อาจารย์เขียนจำนวนตัวเลือกมากที่สุดที่มีความเหมาะสมกับเนื้อหาโจทย์ ไม่จำเป็นต้องเขียนตัวเลือก ๕ ตัวเลือกเสมอไป เนื่องจากตัวเลือกที่ห้าที่เขียนขึ้นเพื่อเติมเต็มโดยไม่สมเหตุสมผลนั้นมักไม่ค่อยมีคนเลือก หากเนื้อหาที่อาจารย์นำมาสอบมีตัวเลือกที่เหมาะสมเพียงสามหรือสี่ตัวเลือกก็เขียนจำนวนตัวเลือกเพียงสามหรือสี่ตัวเลือก^{๔๖} แต่อย่างไรก็ตามให้อาจารย์ศึกษาข้อกำหนดของแต่ละการสอบที่อาจารย์เกี่ยวข้องด้วย เนื่องจากนโยบายของแต่ละการสอบแตกต่างกันไป องค์การที่จัดสอบทางแพทยศาสตรศึกษาจำนวนไม่น้อยยังคงตั้งข้อกำหนดให้ใช้ข้อสอบ ๕ ตัวเลือกเสมอ ซึ่งหากอาจารย์ไม่ทำตามข้อกำหนดดังกล่าวข้อสอบที่ออกไปอาจไม่ได้รับการพิจารณาได้

๔.๒ จัดให้ตัวเลือกที่ถูกต้องมีการกระจายตำแหน่งไปให้มีจำนวนพอ ๆ กันในทุกตัวเลือก

ข้อแนะนำนี้มีวัตถุประสงค์เพื่อป้องกันไม่ให้ผู้สอบที่ตอบแบบเดาสุ่มแบบเลือกตัวเลือกเดียวกันทั้งหมดสอบผ่านได้ด้วยความบังเอิญ หากอาจารย์สร้างข้อสอบที่มีสี่ตัวเลือก เป็น ก ข ค ง อาจารย์ก็ต้องกระจายให้ตัวเลือกที่ถูกมีทั้งข้อ ก ข ค และ ง ในสัดส่วนที่ใกล้เคียงกัน

๔.๓ เขียนตัวเลือกแต่ละข้อให้เป็นอิสระ ไม่ขึ้นต่อกัน



ในการเขียนตัวเลือกของข้อสอบแต่ละข้อ อาจารย์ต้องระมัดระวังให้ตัวเลือกแต่ละตัวเลือกไม่มีความซ้ำซ้อนกัน เช่นตัวเลือก ก เป็นยาในกลุ่มยอยของตัวเลือก ข ตัวเลือก ก เป็นช่วงอายุ ๒ – ๑๐ ปี ตัวเลือก ข เป็นช่วงอายุ ๕ – ๑๑ ปี เป็นต้น การเขียนตัวเลือกที่ซ้ำซ้อนกันนี้ หากเกี่ยวเนื่องกับตัวเลือกที่ถูกต้องอาจมีผู้สอบแย้งว่ามีตัวเลือกที่ถูกต้องมากกว่าหนึ่งตัวเลือก หากตัวเลือกที่ซ้ำซ้อนกันนี้ไม่เกี่ยวกับคำตอบที่ถูก ก็จะทำให้ผู้สอบบางส่วนสามารถตัดตัวเลือกบางตัวเลือกได้โดยไม่ต้องมีความรู้ทางการแพทย์ในเรื่องดังกล่าวได้

๔.๔ เขียนตัวเลือกให้ทุกตัวเลือกมีความเป็นเนื้อเดียวกัน (homogeneous)

การเขียนตัวเลือกให้มีความเป็นเนื้อเดียวกันนั้นหมายถึง ตัวเลือกแต่ละตัวมีรูปร่างหน้าตาและรายละเอียดไปในทิศทางหรือเรื่องราวเดียวกัน หรือเป็นของกลุ่มเดียวกัน การเป็นเนื้อเดียวกันนี้ครอบคลุมตั้งแต่รูปร่างหน้าตา (ตัวเลือกทุกตัวเป็นภาษาแบบเดียวกัน หากตัวเลือกตัวหนึ่งเป็นคำ ตัวเลือกอื่น ๆ ก็ควรเป็นคำ ไม่ใช่วลี หรือประโยค, ตัวเลือกหนึ่งเป็นคำนาม ตัวเลือกอื่นก็เป็นคำนามเหมือนกัน ไม่ใช่กิริยา หรือคำคุณศัพท์) และเนื้อหา (โจทย์ถามการรักษา ตัวเลือกทุกตัวก็เป็นการรักษา ไม่ใช่บางตัวเป็นการตรวจค้นเพิ่มเติม, ตัวเลือกหนึ่งเป็นยาปฏิชีวนะ ตัวเลือกอื่น ๆ ก็น่าจะเป็นยาปฏิชีวนะเช่นกัน ไม่ใช่ยาเคมีบำบัด หรือยาต้านเชื้อรา) การที่มีตัวเลือกที่ไม่เข้าพวก ไม่มีความเป็นเนื้อเดียวกันกับตัวเลือกอื่นเป็นคำบอกใบ้ในการตัดตัวเลือกที่ผู้สอบนิยมใช้มาก ดังนั้นอาจารย์ผู้ออกข้อสอบควรหลีกเลี่ยง

ในบางบริบทของการดูแลรักษาผู้ป่วย สิ่งแพทย์ต้องตัดสินใจเลือกอาจมีทั้งการเลือกที่จะให้การรักษาเลยหรือจะส่งตรวจค้นเพิ่มเติมก่อน ในกรณีนี้อาจารย์สามารถเขียนตัวเลือกที่มีการรักษาและการตรวจเพิ่มเติมปะปนกันได้ แต่การเขียนรูปประโยคคำถามต้องไม่เป็นการบอกใบ้ว่าจะไปทิศทางใด แต่ต้องเลือกใช้คำถามที่เป็นกลาง เช่น ท่านจะปฏิบัติต่อผู้ป่วยอย่างไร, ท่านจะดำเนินการอย่างไรต่อไป เป็นต้น

๔.๕ เขียนตัวเลือกแต่ละข้อให้มีความยาวพอ ๆ กัน

จากการสังเกตข้อสอบปรนัยจำนวนมากจะพบว่าตัวเลือกที่ถูกต้องมักมีความยาวมากกว่าตัวเลือกอื่น ซึ่งข้อสังเกตนี้ผู้สอบจำนวนไม่น้อยก็ทราบดี และผู้สอบส่วนมากเมื่อไม่ทราบคำตอบก็มักเลือกตัวเลือกที่มีความยาวมากที่สุด ดังนั้นอาจารย์ผู้ออกข้อสอบควรระมัดระวังไม่ให้ตัวเลือกตัวใดตัวหนึ่งมีความยาวแตกต่างไปจากตัวเลือกอื่นชัดเจน เพราะจะทำให้ผู้สอบเดาคำตอบที่ถูกได้ง่าย

๔.๖ หลีกเลี่ยงการใช้ตัวเลือก “ถูกทุกข้อ” หรือ “ไม่มีข้อใดถูก”

ตัวเลือก “ถูกทุกข้อ” เป็นตัวเลือกที่ผู้เชี่ยวชาญในการประเมินผลส่วนใหญ่เห็นสอดคล้องกันว่าไม่ควรใช้เนื่องจากมักช่วยใบ้ตัวเลือกที่ถูกต้องให้กับผู้สอบ ทำให้ผู้สอบส่วนหนึ่งตอบถูกโดยไม่ต้องอาศัยองค์ความรู้ที่สมบูรณ์ในเรื่องที่ทดสอบ งานวิจัยพบว่าข้อสอบที่มีตัวเลือกชนิดนี้จะมีผลให้ค่าความเที่ยงของคะแนนสอบลดลง^{๑๑} จึงแนะนำให้หลีกเลี่ยงการใช้

ตัวเลือก “ไม่มีข้อใดถูก” เป็นประเด็นที่ผู้เชี่ยวชาญในการประเมินผลยังคงถกเถียงกันอยู่บ้าง ผู้เชี่ยวชาญบางส่วนเห็นว่าไม่ควรใช้ตัวเลือกประเภทนี้ แต่ผู้เชี่ยวชาญบางส่วนให้ความเห็นว่าสามารถใช้ได้ในบางกรณี^{๑๒} เหตุผลที่ตัวเลือกชนิดนี้เป็นปัญหาคือการใช้ตัวเลือกนี้มักสร้างความลำบากใจให้กับผู้สอบในการเลือกคำตอบที่ถูกในกรณีที่ตัวเลือกแต่ละตัวเลือกไม่ถูกหรือผิดชัดเจน เพราะผู้สอบจะต้องทำการเปรียบเทียบตัวเลือกที่น่าเสนอในข้อสอบกับทางเลือกอื่น ๆ ที่เขานึกได้^{๑๓} หากโจทย์ถามว่ายาใดที่ควรให้แก่ผู้ป่วย แล้วมีชื่อยาสี่ชนิด และมีตัวเลือก “ไม่มีข้อใดถูก” นอกจากที่ผู้สอบต้องนึกว่าในบรรดายาที่ปรากฏในตัวเลือกนั้นเหมาะสมหรือไม่แล้วเขายังนึกต่อไปอีกว่ามียาอื่นใดที่สามารถให้ในผู้ป่วยรายนี้ได้อีก หากเขานึกออกว่ามียาอื่นที่น่าจะเหมาะสมกับผู้ป่วยมากกว่ายาในตัวเลือก (ด้วยเหตุผลที่อาจแตกต่างไปจากที่อาจารย์ผู้ออกข้อสอบคิด) เขาก็จะเลือก “ไม่มีข้อใดถูก”

การใช้ตัวเลือก “ไม่มีข้อใดถูก” จะยังเป็นปัญหามากขึ้นในข้อสอบที่ถามถึงสิ่งที่ไม่ควรทำ เช่นยาใดไม่ควรใช้ในผู้ป่วย ซึ่งนอกจากยาที่น่าเสนอในตัวเลือกแล้วย่อมมียาชนิดอื่นอีกมากมายในบัญชียาที่ไม่เหมาะสม ซึ่งไม่มี

ทางที่ใครจะรู้ได้ว่าการที่ผู้สอบเลือกตอบ “ไม่มีข้อใดถูก” นั้นเขาคิดถึงยาใด และยานั้นไม่เหมาะสมมากไปกว่ายาที่มีอยู่ในตัวเลือกหรือไม่ งานวิจัยทั้งหมดที่ศึกษาถึงตัวเลือกชนิดนี้ได้ข้อสรุปที่ตรงกันว่าข้อสอบที่ใช้ตัวเลือกประเภทนี้เพิ่มระดับความยากให้ข้อสอบ^{๑๑} โดยทั่วไปแล้วจึงไม่แนะนำให้ใช้ตัวเลือกประเภทนี้ในการสอบทางแพทยศาสตรศึกษาซึ่งทางเลือกสำหรับสถานการณ์ที่น่าเสนอมิได้มากและการตัดสินใจเลือกคำตอบต้องอาศัยการเปรียบเทียบข้อดีข้อเสียของแต่ละตัวเลือก

สรุป

ในบทความนี้ผู้นิพนธ์ได้กล่าวถึงข้อแนะนำขั้นพื้นฐานในการพัฒนาข้อสอบปรนัยชนิดเลือกคำตอบที่ถูกที่สุดโดยสรุปข้อแนะนำเหล่านี้ออกเป็นสี่กลุ่มด้วยกัน ได้แก่ (๑) เนื้อหาข้อสอบ, (๒) การจัดรูปแบบข้อสอบ, (๓) การเขียนโจทย์, และ (๔) การเขียนตัวเลือก ผู้นิพนธ์หวังว่าข้อแนะนำเหล่านี้คงพอเป็นแนวทางสำหรับอาจารย์แพทย์ในการพัฒนาข้อสอบปรนัยที่มีคุณภาพเพื่อใช้ในการประเมินนักศึกษาแพทย์และแพทย์ประจำบ้านได้บ้าง อย่างไรก็ตามบทความนี้เป็นกรกล่าวถึงข้อแนะนำเบื้องต้นเท่านั้น ยังมีข้อแนะนำอื่น ๆ ที่ผู้นิพนธ์ไม่ได้นำมารวบรวมไว้ในบทความนี้เพื่อต้องการทำให้เนื้อหากระชับโดยข้อแนะนำอื่น ๆ ที่ผู้นิพนธ์ไม่ได้กล่าวถึงนี้พบว่าเป็นปัญหาน้อยในการออกข้อสอบทางการแพทย์ หรือเป็นข้อแนะนำที่ไม่ได้รับการสนับสนุนอย่างกว้างขวางจากผู้เชี่ยวชาญทางการวัดและประเมินผล หากผู้อ่านสนใจรายละเอียดของข้อแนะนำอื่น ๆ ที่มีผู้กล่าวไว้สามารถศึกษาเพิ่มเติมได้จากเอกสารอ้างอิงที่แสดงไว้ท้ายบทความ

มีข้อควรพิจารณาในการประยุกต์ใช้ข้อแนะนำเหล่านี้ในการพัฒนาข้อสอบที่ผู้นิพนธ์กล่าวถึงประการหนึ่งคือ แม้ว่าข้อแนะนำที่กล่าวถึงเหล่านี้หลายข้อมีการศึกษาวิจัยสนับสนุนที่ชัดเจน แต่สิ่งเหล่านี้ก็เป็นเพียงข้อแนะนำว่าผู้ออกข้อสอบควรปฏิบัติ ไม่ใช่กฎเกณฑ์ตายตัว การเขียนข้อสอบปรนัยนั้นเป็นงานที่ต้องอาศัยทั้งศาสตร์และศิลปะผสมผสานกันอย่างเหมาะสม

หาใช้สูตรคณิตศาสตร์ที่ไม่มีข้อยกเว้น ผู้นิพนธ์ไม่คาดหวังให้อาจารย์ผู้พัฒนาข้อสอบยึดข้อแนะนำเหล่านี้เสมือนกฎเกณฑ์ตายตัวที่ต้องทำตามในทุกกรณี หากแต่ต้องการให้อาจารย์ใช้เป็นแนวทางในการสร้างข้อสอบ ในบางบริบทผู้ออกข้อสอบอาจเลือกที่จะไม่ปฏิบัติตามข้อแนะนำบางประการได้บ้าง แต่การที่จะไม่ปฏิบัติตามข้อแนะนำเหล่านี้จำเป็นต้องมีเหตุผลที่เหมาะสม และควรทำไม่บ่อยนัก ยกตัวอย่างเช่นข้อแนะนำว่า โจทย์ไม่ควรเขียนถามข้อยกเว้น จะพบได้ว่ามีบางบริบทที่การรู้ข้อยกเว้น หรือข้อห้ามปฏิบัติก็เป็นองค์ความรู้ที่สำคัญในการดูแลรักษาผู้ป่วย ดังนั้นในบริบทที่เหมาะสมผู้นิพนธ์เองก็เห็นด้วยว่าอาจเขียนโจทย์ที่ถามข้อยกเว้นได้ แต่อย่างไรก็ตามการจะไม่ปฏิบัติตามข้อแนะนำนี้ต้องไม่ทำบ่อยจนเกินไป จำเป็น หากออกข้อสอบ ๑๐๐ ข้อ จะมีข้อสอบที่ถามข้อยกเว้น ประมาณมาบ้าง ๒-๓ ข้อ ย่อมเป็นสิ่งที่พอยอมรับได้ แต่หากในชุดข้อสอบมีข้อสอบถึงร้อยละ ๒๐ - ๓๐ ที่โจทย์เขียนในรูปประโยคปฏิเสธ ถามสิ่งที่ไม่ควรปฏิบัติ หรือสิ่งที่ไม่ถูกต้อง อย่างนี้ย่อมจัดว่าละเลยแนวทางในการพัฒนาข้อสอบอย่างไม่เหมาะสม ซึ่งย่อมส่งผลให้คุณภาพของข้อสอบด้อยลงอย่างชัดเจน

เอกสารอ้างอิง

1. Downing SM. Assessment of knowledge with written test forms. In: Norman GR, van der Vleuten C, Newble DI, editors. International handbook of research in medical education. Dordrecht: Kluwer Academic Publishers, 2002:647 - 72.
2. Haladyna TM, Downing SM. A taxonomy of multiple-choice item-writing rules. Appl Meas Educ 1989;2:37-50.
3. Haladyna TM. Developing and validating multiple-choice test items, 3rd ed. Mahwah, NJ: Lawrence Erlbaum Associates, 2004.
4. Maatsch JL, Huang RR, Downing SM, Munger BS. The predictive validity of test formats and a psychometric theory of clinical competence. The 23rd Conference on Research in Medical Education. Washington, DC: Association of American Medical Colleges, 1984.
5. Jozefowicz RF, Koeppen BM, Case S, Galbraith R, Swanson D, Glew RH. The quality of in-house medical school examinations. Acad Med 2002;77(2):156-61.
6. 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.



7. 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.
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.
9. Haladyna TM, Downing SM. Validity of a taxonomy of multiple-choice item-writing rules. *Appl Meas Educ* 1989;2(1):51-78.
10. Haladyna TM, Downing SM, Rodriguez MC. A review of multiple-choice item-writing guidelines for classroom assessment. *Appl Meas Educ* 2002;15:309-34.
11. Downing SM, Dawson-Saunders B, Case SM, Powell RD. The psychometric effects of negative stems, unfocused questions, and heterogeneous options on NBME Part I and Part II item characteristics. the annual meeting of the National Council on Measurement in Education. Chicago, IL, 1991.
12. Tamir P. Positive and negative multiple choice items: How different are they? *Stud Educ Eval* 1993;19:311-25.
13. Rogers WT, Harley D. An empirical comparison of three- and four-choice items and tests: Susceptibility to testwiseness and internal consistency reliability. *Educ Psychol Meas* 1999;59:234-47.
14. Sidick JT, Barrett GV, Doverspike D. Three-alternative multiple choices tests: An attractive option. *Pers Psychol* 1994;47:829-35.
15. Cizek GJ, Rachor RE. Nonfunctioning options: A closer look. The annual meeting of the American Educational Research Association. San Francisco, CA, 1995.
16. Crehan KD, Haladyna TM, Brewer BW. Use of an inclusive option and the optimal number of options for multiple-choice items. *Educ Psychol Meas* 1993;53:241-7.
17. Lord FM. Optimal number of choices per item. *J Educ Meas* 1977; 14:33-8.
18. Haladyna TM, Downing SM. How many options is enough for a multiple-choice item? *Educ Psychol Meas* 1993;53:999-1010.