Mahidol University
Faculty of Medicine Siriraj Hospital

หน่วยพัฒนาแพทยศาสตรศึกษาและวิจัยการศึกษา
คณะแพทยศาสตร์ศิริราชพยาบาล มหาวิทยาลัยมหิดล
ขอเรียนเชิญอาจารย์ เข้าร่วมโครงการอบรม เรื่อง

MCQ การพัฒนาข้อสอบปรนัย เพื่อประเมินความรู้ทางการแพทย์

Multiple Choice Question

ออกข้อสอบกันทุกปี คิดว่าดีแล้วหรือยัง?

ระหว่างวันที่ 22-23 มิถุนายน 2560
ณ ห้องบรรยาย 3A01 ชั้น 3A
อาคารศรีสวรินทิรา คณะแพทยศาสตร์ศิริราชพยาบาล

เอกสารประกอบการอบรม

si-merd.com
MERD
ลงทะเบียน Online

สอบถามเพิ่มเติม
คุณภัทรพร/ คุณสุวรรณี
โทร. 02-419-9978/ 02-419-6637
E-mail: merd.project@gmail.com

กำหนดการ

โครงการอบรมเชิงปฏิบัติ เรื่อง "การพัฒนาข้อสอบปรนัยเพื่อประเมินความรู้ทางการแพทย์"

ระหว่างวันที่ 22-23 มิถุนายน พ.ศ. 2560

ณ ห้องบรรยาย 3A01 ชั้น 3A และห้องปฏิบัติการพระอาจวิทยาคม ชั้น 5 อาคารศรีสวรินทิรา คณะแพทยศาสตร์ศิริราชพยาบาล

วันพฤหัสบดีที่ 22 มิถุนายน พ.ศ. 2560		**วิทยากรหลัก**	**วิทยากรร่วม**
08.30 - 09.00 น.	**ลงทะเบียนภาคเช้า**	**ห้องบรรยาย 3A01 อาคารศรีสวรินทิรา ชั้น 3A**	
09.00 – 09.15 น.	หลักการพื้นฐานของข้อสอบปรนัย	รศ.ดร. นพ.เชิดศักดิ์ ไอรมณีรัตน์	รศ. นพ.สุพจน์ พงศ์ประสบชัย
09.15 – 09.30 น.	การวางแผนและจัดทำตารางกำหนดเนื้อหาข้อสอบ	รศ.ดร. นพ.เชิดศักดิ์ ไอรมณีรัตน์	รศ. นพ.วันชัย เดชสมฤทธิ์ฤทัย อ.ดร. นพ.ยอดยิ่ง แดงประไพ
09.30 – 10.30 น	การสร้างโจทย์และตัวเลือกข้อสอบปรนัย	รศ.ดร. นพ.เชิดศักดิ์ ไอรมณีรัตน์	ผศ. นพ.สุประพัฒน์ สนใจพาณิชย์ ผศ. พญ.อนัญญา พงษ์ไพบูลย์
10.45 – 12.00 น.	ข้อผิดพลาดที่พบบ่อยในการสร้างข้อสอบปรนัย	รศ. นพ.สุพจน์ พงศ์ประสบชัย รศ. พญ.พรพรรณ กู้มานะชัย	รศ.ดร. นพ.เชิดศักดิ์ ไอรมณีรัตน์
12.00 - 13.00 น.	*พักรับประทานอาหารกลางวัน*		
13.00 – 16.00 น.	การพิจารณาข้อสอบ	อ.ดร. นพ.ยอดยิ่ง แดงประไพ ผศ. พญ.อนัญญา พงษ์ไพบูลย์	**กลุ่มที่ 1** ห้องปฏิบัติการฯ ชั้น 5
		รศ. นพ.วันชัย เดชสมฤทธิ์ฤทัย ผศ. นพ.สุประพัฒน์ สนใจพาณิชย์	**กลุ่มที่ 2** ห้องปฏิบัติการฯ ชั้น 5
		รศ. นพ.สุพจน์ พงศ์ประสบชัย รศ. พญ.พรพรรณ กู้มานะชัย	**กลุ่มที่ 3** ห้องปฏิบัติการฯ ชั้น 5
		รศ. นพ.อรรถพล ใจชื่น ผศ. นพ. ตรีภพ เลิศบรรณพงษ์	**กลุ่มที่ 4** ห้องปฏิบัติการฯ ชั้น 5
		รศ. นพ.รุ่งนิรันดร์ ประดิษฐสุวรรณ ผศ. นพ. ยิ่งยง ชินธรรมมิตร์	**กลุ่มที่ 5** ห้องปฏิบัติการฯ ชั้น 5
		รศ.ดร. นพ.เชิดศักดิ์ ไอรมณีรัตน์ อ. นพ. อนุภพ จิตต์เมือง	**กลุ่มที่ 6** ห้องปฏิบัติการฯ ชั้น 5
วันศุกร์ที่ 23 มิถุนายน พ.ศ. 2560		**วิทยากรหลัก**	**วิทยากรร่วม**
08.30 – 09.00 น.	**ลงทะเบียนภาคเช้า**	**ห้องบรรยาย 3A01 อาคารศรีสวรินทิรา ชั้น 3A**	
09.00 – 09.30 น	ปัญหาในการสร้างข้อสอบปรนัย	รศ.ดร. นพ.เชิดศักดิ์ ไอรมณีรัตน์	รศ. นพ.วันชัย เดชสมฤทธิ์ฤทัย รศ. พญ.พรพรรณ กู้มานะชัย ผศ. พญ.อนัญญา พงษ์ไพบูลย์
09.30 – 10.45 น.	การผลิตข้อสอบและการจัดสอบ	ผศ. พญ.อนัญญา พงษ์ไพบูลย์ คุณสุนันท์ มีเทศ คุณทศพร มาสวัสดิ์	รศ.ดร. นพ.เชิดศักดิ์ ไอรมณีรัตน์
11.00 – 12.00 น	การตรวจและวิเคราะห์ข้อสอบปรนัย	รศ.ดร. นพ.เชิดศักดิ์ ไอรมณีรัตน์ คุณสุนันท์ มีเทศ	
12.00 – 13.00 น.	*พักรับประทานอาหารกลางวัน*		
13.00 – 14.30 น.	การตั้งเกณฑ์ผ่านการสอบ	รศ.ดร. นพ.เชิดศักดิ์ ไอรมณีรัตน์	
14.45 – 15.30 น.	การจัดทำคลังข้อสอบ	ผศ. พญ.อนัญญา พงษ์ไพบูลย์ คุณทศพร มาสวัสดิ์	รศ.ดร. นพ.เชิดศักดิ์ ไอรมณีรัตน์
15.30 – 16.00 น	สรุป	รศ.ดร. นพ.เชิดศักดิ์ ไอรมณีรัตน์	

หมายเหตุ: กำหนดการอาจมีการเปลี่ยนแปลงตามความเหมาะสม

รายชื่อผู้ร่วมอบรม

ณ ห้องบรรยาย 3A01 ชั้น 3A และห้องปฏิบัติการพระอาจวิทยาคม ชั้น 5 อาคารศรีสวรินทิรา คณะแพทยศาสตร์ศิริราชพยาบาล

กลุ่มที่ 1 (อาจารย์ประจำกลุ่ม : อ.ดร. นพ.ยอดยิ่ง แดงประไพ และผศ. พญ.อนัญญา พงษ์ไพบูลย์)

ลำดับ	คำนำหน้า	ชื่อ	สกุล	สังกัด	หน่วยงาน/ภาควิชา
1	ผศ. ดร.	เตือนจิต	คำพิทักษ์	คณะแพทยศาสตร์ มหาวิทยาลัยขอนแก่น	สาขา ชีวเคมี
2	อ. ดร. นพ.	จตุรงค์	เสวตานนท์	คณะแพทยศาสตร์ศิริราชพยาบาล	ภาควิชาจุลชีววิทยา
3	รศ. พญ.	จุฬาพรรณ	อิ้งจะนิล	คณะแพทยศาสตร์ มหาวิทยาลัยขอนแก่น	สาขาจุลชีววิทยา
4	ผศ. พญ.	กาญจนา	อมรพิเชษฐ์กูล	คณะแพทยศาสตร์ศิริราชพยาบาล	ภาควิชาพยาธิวิทยา
5	ผศ. พญ.	พนิตตา	สิทธินามสุวรรณ	คณะแพทยศาสตร์ศิริราชพยาบาล	ภาควิชาพยาธิวิทยา
6	ผศ. นพ.	ถิรพล	บุญญาอรุณเนตร	คณะแพทยศาสตร์ศิริราชพยาบาล	ภาควิชาพยาธิวิทยา
7	รศ. พญ.	ตุ้มทิพย์	แสงรุจิ	คณะแพทยศาสตร์ศิริราชพยาบาล	ภาควิชาพยาธิวิทยา
8	นาย	ชัยวัชร์	อภิวาทนสิริ	คณะแพทยศาสตร์ มหาวิทยาลัยขอนแก่น	สาขา วิชาพยาธิวิทยา
9	ผศ. พญ.	สุดารัตน์	ปิโยพีระพงศ์	คณะแพทยศาสตร์ศิริราชพยาบาล	ภาควิชาพยาธิวิทยาคลินิก
10	พญ.	ขวัญอนงค์	อรรถเกตุถาวร	คณะแพทยศาสตร์วชิรพยาบาล มหาวิทยาลัยนวมินทราธิราช	ภาควิชาพยาธิวิทยากายวิภาค
11	อ.	บุตรโต	นามบุญเรือง	คณะแพทยศาสตร์ มหาวิทยาลัยสยาม	สาขา กายวิภาคศาสตร์
12	นางสาว	เฉลิมขวัญ	รุ่งสว่าง	คณะแพทยศาสตร์ มหาวิทยาลัยสยาม	สาขาวิชา วิทยาศาสตร์การแพทย์
13	ผศ. ดร.	สุเมธ	วจนรจนา	คณะแพทยศาสตร์วชิรพยาบาล มหาวิทยาลัยนวมินทราธิราช	ภาควิชาวิทยาศาสตร์การแพทย์พื้นฐาน
14	ภญ.ผศ.	อารีวรรณ	เชี่ยวชาญวัฒนา	คณะเภสัชศาสตร์ มหาวิทยาลัยขอนแก่น	สาขาวิชาเภสัชศาสตร์สังคมและการบริหาร
15	รศ. ดร. ทพญ.	ศรัญญา	ตันเจริญ	คณะทันตแพทยศาสตร์ มหาวิทยาลัยมหิดล	สาขาวิชาเภสัชวิทยา
16	ดร.	พูลสิทธิ์	หิรัญสาย	สำนักวิชาสหเวชศาสตร์ มหาวิทยาลัยวลัยลักษณ์	สาขาเทคนิคการแพทย์

กลุ่มที่ 2 (อาจารย์ประจำกลุ่ม : รศ. นพ.วันชัย เดชสมฤทธิ์ฤทัย และผศ. นพ.สุประพัฒน์ สนใจพาณิชย์)

ลำดับ	คำนำหน้า	ชื่อ	สกุล	สังกัด	หน่วยงาน/ภาควิชา
1	พญ.	สุดารัตน์	เอื้อศิริวรรณ	คณะแพทยศาสตร์วชิรพยาบาล มหาวิทยาลัยนวมินทราธิราช	ภาควิชากุมารเวชศาสตร์
2	ผศ. นพ.	ยุทธพงศ์	พุทธรักษา	คณะแพทยศาสตร์ มหาวิทยาลัยนเรศวร	ภาควิชากุมารเวชศาสตร์
3	นพ.	พนิต	ทักขิญเสถียร	ศูนย์การแพทย์สมเด็จพระเทพรัตนราชสุดา ฯ สยามบรมราชกุมารี มหาวิทยาลัยศรีนครินทรวิโรฒ	กุมารเวชศาสตร์
4	ผศ.พญ.	ยิ่งวรรณ	เจริญยิ่ง	คณะแพทยศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ	กุมารเวชศาสตร์
5	พญ.	สรัสวดี	เถลิงศก	คณะแพทยศาสตร์ มหาวิทยาลัยพะเยา	กุมารเวชศาสตร์
6	ผศ.นพ.	สมบูรณ์	จันทร์สกุลพร	คณะแพทยศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ	กุมารเวชศาสตร์
7	นพ.	ประวิทย์	เจตนชัย	สถาบันสุขภาพเด็กแห่งชาติมหาราชินี	สาขาโรคปอด
8	พญ.	กัญฑิมาศ	สิทธิกูล	สถาบันสุขภาพเด็กแห่งชาติมหาราชินี	สาขาโรคปอด
9	พญ.	นิจวรรณ	เกิดเจริญ	คณะแพทยศาสตร์วชิรพยาบาล มหาวิทยาลัยนวมินทราธิราช	ภาควิชาจิตเวชศาสตร์
10	ผศ. พญ.	ดาวชมพู	นาคะวิโร	คณะแพทยศาสตร์โรงพยาบาลรามาธิบดี มหาวิทยาลัยมหิดล	ภาควิชาจิตเวชศาสตร์
11	นาย	ชัยวัฒน์	อินไชยา	วิทยาลัยพยาบาลบรมราชชนนีนครพนม มหาวิทยาลัยนครพนม	กลุ่มวิชาการพยาบาลสุขภาพจิตและจิตเวชศาสตร์
12	ผศ.ดร.	วิรมณ	กาสีวงศ์	วิทยาลัยพยาบาลบรมราชชนนีนครพนม มหาวิทยาลัยนครพนม	กลุ่มวิชาการพยาบาลสุขภาพจิตและจิตเวชศาสตร์
13	นางสาว	วราพร	มีแก้ว	วิทยาลัยพยาบาลบรมราชชนนีนครพนม มหาวิทยาลัยนครพนม	กลุ่มวิชาการพยาบาลสุขภาพจิตและจิตเวชศาสตร์
14	นางสาว	ศิริภรณ์	เหมะธุลิน	วิทยาลัยพยาบาลบรมราชชนนีนครพนม มหาวิทยาลัยนครพนม	กลุ่มวิชาการพยาบาลสุขภาพจิตและจิตเวชศาสตร์
15	นาย	อนันตศักดิ์	พันธ์พุฒ	วิทยาลัยพยาบาลบรมราชชนนีนครพนม มหาวิทยาลัยนครพนม	กลุ่มวิชาการพยาบาลสุขภาพจิตและจิตเวชศาสตร์

กลุ่มที่ 3 (อาจารย์ประจำกลุ่ม : รศ. นพ.สุพจน์ พงศ์ประสบชัย และรศ. พญ.พรพรรณ กู้มานะชัย)

ลำดับ	คำนำหน้า	ชื่อ	สกุล	สังกัด	หน่วยงาน/ภาควิชา
1	ผศ. พญ.	ศิริพร	ลิมป์วิริยะกุล	คณะแพทยศาสตร์ศิริราชพยาบาล	ภาควิชาโสต นาสิก ลาริงซ์วิทยา
2	พญ.	ปรัตถา	เกิดศิริ	โรงพยาบาลเจริญกรุงประชารักษ์	สาขา โสต ศอ นาสิก
3	รศ. ร.อ. พญ.	ภิญนิตา	ตันธุวนิตย์	คณะแพทยศาสตร์ศิริราชพยาบาล	ภาควิชาจักษุวิทยา
4	รศ. พญ.	อติพร	ตวงทอง	คณะแพทยศาสตร์ศิริราชพยาบาล	ภาควิชาจักษุวิทยา
5	นาง	ณภัทร	ฤทธิ์ล้ำเลิศ	วิทยาลัยวิทยาศาสตร์การแพทย์เจ้าฟ้าจุฬาภรณ์ ราชวิทยาลัยจุฬาภรณ์	รังสีเทคนิค
6	รศ. พญ.	พรพิมพ์	กอแพร่พงศ์	คณะแพทยศาสตร์ศิริราชพยาบาล	ภาควิชารังสีวิทยา
7	นาง	สุทธิรักษ์	ตั้งเรืองเกียรติ	คณะแพทยศาสตร์ ราชวิทยาลัยจุฬาภรณ์	หลักสูตรรังสีเทคนิค
8	อ. ทพ.	คมสัน	ลาภาอุตย์	คณะทันตแพทยศาสตร์ มหาวิทยาลัยมหิดล	ภาควิชาทันตกรรมชุมชน
9	ผศ.ดร. ทพญ.	วรางคณา	ชิดช่วงชัย	คณะทันตแพทยศาสตร์ มหาวิทยาลัยมหิดล	ภาควิชาเวชศาสตร์ช่องปากและปริทันตวิทยา
10	นางสาว	กันต์ชมขวัญ	กัณฐ์กรกช	โรงพยาบาลมหาราชนครศรีธรรมราช	พยาบาลวิชาชีพชำนาญการ
11	นาง	วิสัย	คะตา	วิทยาลัยพยาบาลบรมราชชนนีนครพนม มหาวิทยาลัยนครพนม	กลุ่มวิชาการพยาบาลสุขภาพจิตและจิตเวชศาสตร์
12	นางสาว	เนาวรัตน์	เสนาไชย	วิทยาลัยพยาบาลบรมราชชนนีนครพนม มหาวิทยาลัยนครพนม	กลุ่มวิชาการพยาบาลสุขภาพจิตและจิตเวชศาสตร์
13	นาง	ละอองดาว	คำชาตา	วิทยาลัยพยาบาลบรมราชชนนีนครพนม มหาวิทยาลัยนครพนม	กลุ่มวิชาการพยาบาลสุขภาพจิตและจิตเวชศาสตร์
14	นาย	สมเกียรติ	บุญคง	วิทยาลัยพยาบาลบรมราชชนนีนครพนม มหาวิทยาลัยนครพนม	กลุ่มวิชาการพยาบาลสุขภาพจิตและจิตเวชศาสตร์
15	นางสาว	ชลินดา	คำศรีพล	วิทยาลัยพยาบาลบรมราชชนนีนครพนม มหาวิทยาลัยนครพนม	กลุ่มวิชาการพยาบาลสุขภาพจิตและจิตเวชศาสตร์
16	นาง	สุชาดา	ปราบมีชัย	วิทยาลัยพยาบาลบรมราชชนนีนครพนม มหาวิทยาลัยนครพนม	กลุ่มวิชาการพยาบาลสุขภาพจิตและจิตเวชศาสตร์

กลุ่มที่ 4 (อาจารย์ประจำกลุ่ม : รศ. นพ.อรรถพล ใจชื่น และผศ. นพ. ตรีภพ เลิศบรรณพงษ์)

ลำดับ	คำนำหน้า	ชื่อ	สกุล	สังกัด	หน่วยงาน/ภาควิชา
1	พญ.	จิตสุภา	คุณาเศรษฐ	โรงพยาบาลตำรวจ	กลุ่มงานสูตินรเวชกรรม
2	พญ.	ปัทมา	ทองดี	สำนักวิชาแพทยศาสตร์ มหาวิทยาลัยเทคโนโลยีสุรนารี	สูติศาสตร์และนรีเวชวิทยา
3	พญ.	หนึ่งฤทัย	แซ่เอียบ	คณะแพทยศาสตร์ มหาวิทยาลัยสงขลานครินทร์	สูติศาสตร์และนรีเวชวิทยา
4	รศ. พญ.	ธันยารัตน์	วงศ์วนานุรักษ์	คณะแพทยศาสตร์ศิริราชพยาบาล	ภาควิชาสูติศาสตร์-นรีเวชวิทยา
5	รศ. นพ.	พฤหัส	จันทร์ประภาพ	คณะแพทยศาสตร์ศิริราชพยาบาล	ภาควิชาสูติศาสตร์-นรีเวชวิทยา
6	อ. พญ.	กุลวรา	กิตติสาเรศ	คณะแพทยศาสตร์ศิริราชพยาบาล	ภาควิชาเวชศาสตร์การธนาคารเลือด
7	นาง	ฉันทนา	เคหะ	คณะแพทยศาสตร์ศิริราชพยาบาล	ตรวจโรคแพทย์เวร-ฉุกเฉิน
8	นางสาว	วิลาวัลย์	ชุมนุมดวง	คณะแพทยศาสตร์ศิริราชพยาบาล	ตรวจโรคแพทย์เวร-ฉุกเฉิน
9	นาย	ศุภกิจ	สุวรรณไตรย์	คณะแพทยศาสตร์ศิริราชพยาบาล	สถานการแพทย์แผนไทยประยุกต์
10	นางสาว	ดอกไม้	วิวรรธมงคล	คณะแพทยศาสตร์ศิริราชพยาบาล	สถานการแพทย์แผนไทยประยุกต์
11	นางสาว	กนกอร	วัฒนพรชัย	คณะแพทยศาสตร์ศิริราชพยาบาล	สถานการแพทย์แผนไทยประยุกต์
12	นาย	ปรัชญาวุฒิ	รอดสัมฤทธิ์	คณะแพทยศาสตร์ศิริราชพยาบาล	สถานการแพทย์แผนไทยประยุกต์
13	นางสาว	ผ่องพิมล	ดาศรี	คณะแพทยศาสตร์ศิริราชพยาบาล	สถานการแพทย์แผนไทยประยุกต์
14	นาย	ประสพพร	พันธ์เพ็ง	คณะแพทยศาสตร์ศิริราชพยาบาล	สถานการแพทย์แผนไทยประยุกต์
15	นาง	ฉันทณัฏฐ์	ทิพยเจริญธัม	คณะแพทยศาสตร์ศิริราชพยาบาล	สถานการแพทย์แผนไทยประยุกต์
16	นางสาว	นาริชาติ	ชื่นสกุล	คณะแพทยศาสตร์ศิริราชพยาบาล	สถานการแพทย์แผนไทยประยุกต์

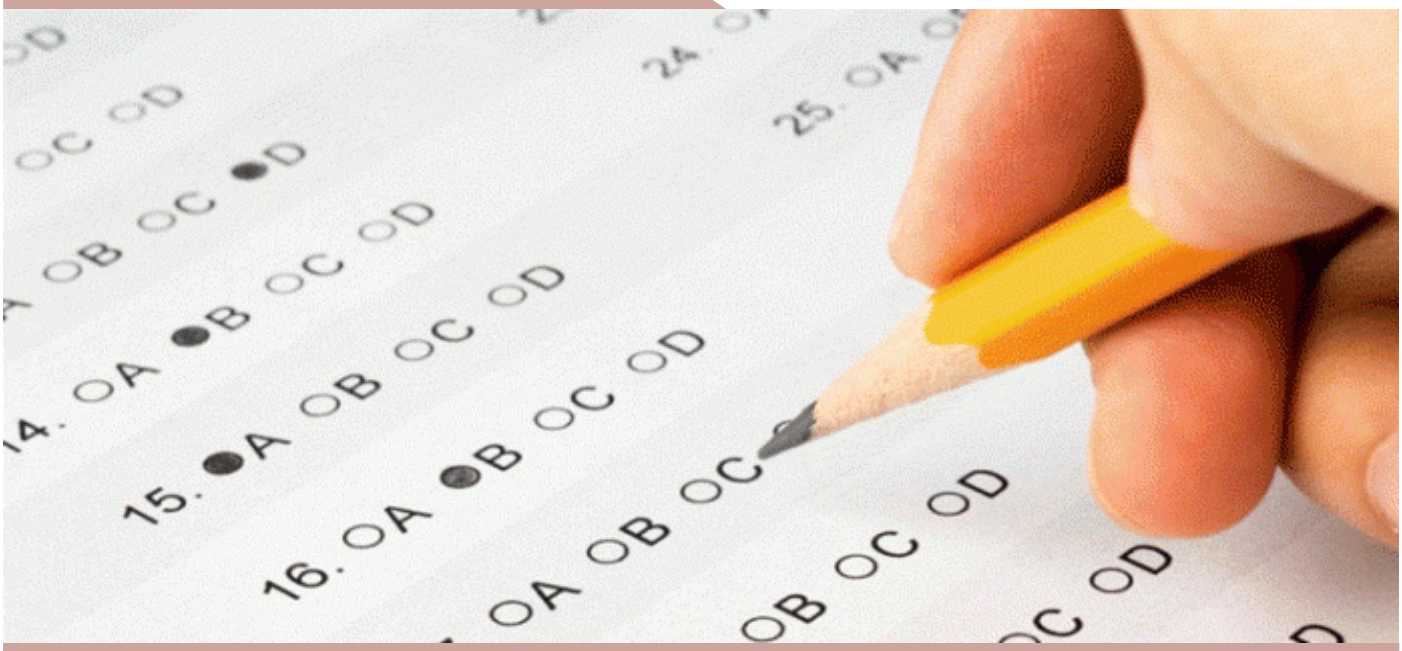
กลุ่มที่ 5 (อาจารย์ประจำกลุ่ม : รศ. นพ.รุ่งนิรันดร์ ประดิษฐสุวรรณ และผศ. นพ. ยิ่งยง ชินธรรมมิตร์)

ลำดับ	คำนำหน้า	ชื่อ	สกุล	สังกัด	หน่วยงาน/ภาควิชา
1	พญ.	พรทิพย์	นิ่มขุนทด	สำนักวิชาแพทยศาสตร์ มหาวิทยาลัยเทคโนโลยีสุรนารี	อายุรศาสตร์
2	อ. พญ.	จิตรวดี	หอพิบูลสุข	สำนักวิชาแพทยศาสตร์ มหาวิทยาลัยเทคโนโลยีสุรนารี	สาขา อายุรศาสตร์
3	พญ.	ปริฉัตร	เอื้ออารีวงศา	คณะแพทยศาสตร์ มหาวิทยาลัยสงขลานครินทร์	ภาควิชาอายุรศาสตร์
4	รศ. พญ.	ธนันดา	ตระการวนิช	คณะแพทยศาสตร์วชิรพยาบาล มหาวิทยาลัยนวมินทราธิราช	อายุรศาสตร์
5	รศ. พญ.	สุภัทศรี	เศรษฐสินธุ์	คณะแพทยศาสตร์วชิรพยาบาล มหาวิทยาลัยนวมินทราธิราช	อายุรศาสตร์
6	นพ.	ธนรร	งามวิชชุกร	คณะแพทยศาสตร์วชิรพยาบาล มหาวิทยาลัยนวมินทราธิราช	อายุรศาสตร์
7	นพ.	สิรวัฒน์	ศรีฉัตราภิมุข	คณะแพทยศาสตร์โรงพยาบาลรามาธิบดี มหาวิทยาลัยมหิดล	สาขา อายุรศาสตร์
8	นพ.	พีระพัชร	ไทยสยาม	โรงพยาบาลมหาราชนครศรีธรรมราช	สาขาวิชาโรคติดเชื้อ กลุ่มงานอายุรกรรม
9	นพ.	ปณิธาน	วัจนาคมกุล	คณะแพทยศาสตร์ มหาวิทยาลัยสงขลานครินทร์	เวชศาสตร์ครอบครัวและเวชศาสตร์ป้องกัน
10	พ.ต.อ.หญิง	พัชรีวรรณ	วิเศษสุมน	โรงพยาบาลตำรวจ	กลุ่มงานเวชศาสตร์ครอบครัว
11	นพ.	ประเสริฐ	ประกายรุ้งทอง	โรงพยาบาลพระนั่งเกล้า	กลุ่มงานเวชกรรมสังคม
12	อ.	สุมิตรา	เมืองขวา	คณะแพทยศาสตร์วชิรพยาบาล มหาวิทยาลัยนวมินทราธิราช	ภาควิชาเวชศาสตร์เขตเมือง
13	นางสาว	วรรณภา	เพชรแก้วมณี	คณะแพทยศาสตร์ศิริราชพยาบาล	โรงเรียนผู้ช่วยพยาบาล
14	นาง	ลัดดาวัลย์	ปิยะทรงสุทธิ์	คณะแพทยศาสตร์ศิริราชพยาบาล	โรงเรียนผู้ช่วยพยาบาล
15	นาง	สุกัญญา	ศรีสวัสดิ์	คณะแพทยศาสตร์ มหาวิทยาลัยสงขลานครินทร์	หน่วยทะเบียนและประเมินผล
16	นางสาว	พรพิมล	ปิลันธนาภรณ์	คณะแพทยศาสตร์ มหาวิทยาลัยสงขลานครินทร์	หน่วยทะเบียนและประเมินผล

กลุ่มที่ 6 (อาจารย์ประจำกลุ่ม : รศ.ดร. นพ.เชิดศักดิ์ ไอรมณีรัตน์ และอ. นพ. อนุภพ จิตต์เมือง)

ลำดับ	คำนำหน้า	ชื่อ	สกุล	สังกัด	หน่วยงาน/ภาควิชา
1	ผศ. นพ.	สมบูรณ์	ทรัพย์วงศ์เจริญ	วิทยาลัยแพทยศาสตร์ มหาวิทยาลัยรังสิต	ภาควิชาศัลยศาสตร์
2	ผศ. นพ.	ไพโรจน์	อาจแย้มสรวล	คณะแพทยศาสตร์ศิริราชพยาบาล	ภาควิชาศัลยศาสตร์
3	อ. นพ.	วิษณุ	โลห์สิริวัฒน์	คณะแพทยศาสตร์ศิริราชพยาบาล	ภาควิชาศัลยศาสตร์
4	พญ.	ปฐมา	นาคเจือ	โรงพยาบาลเจริญกรุงประชารักษ์	สาขาวิชาศัลยศาสตร์
5	พญ.	พิชามญชุ์	คณิตานุพงษ์	คณะแพทยศาสตร์ มหาวิทยาลัยสงขลานครินทร์	ศัลยศาสตร์ออร์โธปิดิกส์
6	นพ.	คณินทร์	เอี่ยมธนาภรณ์	คณะแพทยศาสตร์ มหาวิทยาลัยสงขลานครินทร์	ศัลยศาสตร์ออร์โธปิดิกส์
7	นพ.	ภูมิบาล	เวศย์พิริยะกุล	คณะแพทยศาสตร์วชิรพยาบาล มหาวิทยาลัยนวมินทราธิราช	ศัลยศาสตร์ออร์โธปิดิกส์
8	นพ.	สุธี	ทวีพันธุ์สานต์	คณะแพทยศาสตร์วชิรพยาบาล มหาวิทยาลัยนวมินทราธิราช	ออร์โธปิดิกส์
9	นพ.	กิจชัย	ลักษมีอรุโณทัย	คณะแพทยศาสตร์วชิรพยาบาล มหาวิทยาลัยนวมินทราธิราช	ศัลยศาสตร์กระดูกและข้อ
10	นางสาว	ชวนพิศ	บุญเกิด	คณะสหเวชศาสตร์ มหาวิทยาลัยธรรมศาสตร์	สาขา กายภาพบำบัด
11	พญ.	อุทัยวรรณ	เล็กยิ่งยง	โรงพยาบาลตำรวจ	เวชศาสตร์ฟื้นฟู
12	พญ.	อังคณา	เทพเลิศบุญ	โรงพยาบาลเจริญกรุงประชารักษ์	กลุ่มงานเวชกรรมฟื้นฟู
13	พญ.	วันวิชนี	เจริญผล	โรงพยาบาลตำรวจ	สาขาวิสัญญี
14	อ. นพ.	ชูพงศ์	ล้วนศรีติสกุล	คณะแพทยศาสตร์ศิริราชพยาบาล	ภาควิชาวิสัญญีวิทยา
15	อ. นพ.	ศุภชัย	พูลสรรพสิทธิ์	คณะแพทยศาสตร์ศิริราชพยาบาล	ภาควิชาวิสัญญีวิทยา
16	รศ. นพ.	วิสูตร	ฟองศิริไพบูลย์	คณะแพทยศาสตร์ศิริราชพยาบาล	ภาควิชานิติเวชศาสตร์

เอกสารประกอบการอบรม



22 June 2017

22 June 2017

หัวข้อ : หลักการพื้นฐานของข้อสอบปรนัย

หลักการพื้นฐานของข้อสอบปรนัย

รศ.นพ. เชิดศักดิ์ ไอรมณีรัตน์
ภาควิชาศัลยศาสตร์ คณะแพทยศาสตร์ศิริราชพยาบาล
มหาวิทยาลัยมหิดล

Multiple-Choice Questions

- Selected Response Exam
  - True/False
    - Simple True/False items
    - Multiple true/false items (K-type)
  - One best response
    - Standard MCQ
    - Extended matching items

Multiple-Choice Questions

- Advantages
  - Objective scoring
  - High internal consistency reliability
  - Strong research evidence to support its validity
  - Efficiency in testing and scoring

10

Multiple-Choice Questions

- Limitations
  - Cueing of correct answer
  - Random guessing
  - Testing of trivial knowledge
  - Difficulty of development of good MCQ items
  - Focus only in cognitive abilities, not good for assessing psychomotor skills or attitudes

11

MCQ in Thai Medical Education

- Medical school admission
- Classroom tests
- Comprehensive exam
- National licensing exam steps 1, 2
- Postgraduate exam
  - Basic science exam
  - Board exam

Twelve Steps for Test Development

1. Overall plan
2. Content definition
3. Test specification
4. Item development
5. Test assembly
6. Test production
7. Test administration
8. Scoring
9. Passing score
10. Reporting results
11. Item banking
12. Technical report

Downing SM, Haladyna TM. Handbook of test development. Mahwah, NJ: Lawrence Earlbaum Associates 2006.

Cherdsak.ira@mahidol.ac.th

22 June 2017

หัวข้อ : การวางแผนและจัดทำตารางกำหนดเนื้อหาข้อสอบ

การวางแผนและ
จัดทำตารางกำหนดเนื้อหาข้อสอบ

รศ.นพ. เชิดศักดิ์ ไอรมณีรัตน์
ภาควิชาศัลยศาสตร์ คณะแพทยศาสตร์ศิริราชพยาบาล
มหาวิทยาลัยมหิดล

Overall Plan

- What construct is to be measured?
- What test format will be used?
- What is the test administration modality?
- When will the test be given? (timeline)
- What is a quality control process?

A Construct

- A latent trait: A personal quality that we want to measure but cannot do so directly
  - Examples: mathematic intelligence, knowledge in physiology, etc.

Test Administration Modality

1.) Paper-based test
2.) Computer-based test

Twelve Steps for Test Development

1. Overall plan
2. Content definition
3. Test specification
4. Item development
5. Test assembly
6. Test production
7. Test administration
8. Scoring
9. Passing score
10. Reporting results
11. Item banking
12. Technical report

Downing SM, Haladyna TM. Handbook of test development. Mahwah, NJ: Lawrence Earlbaum Associates 2006.

Content Definition

- What content is to be tested?
- Human judgment

Categorization of the Test Items

1. Nature of the content
2. Nature of learning

Test specification

Cognitive Hierarchy

- Knowledge (รู้)
- Comprehension (เข้าใจ)
- Application (ประยุกต์ใช้)
- Analysis (วิเคราะห์)
- Synthesis (สังเคราะห์)
- Evaluation (ประเมินค่า)

Test specification

A Simplified Cognitive Hierarchy

- **Recall** (ความจำ)
- **Comprehension** (ความเข้าใจ)
- **Application** (การประยุกต์ใช้)

Test specification

Test Specification Table

- Methods: vary in rigor, depending on the purpose of the test, and consequences
  1. High-stakes exam:
     - Practice analysis
  2. Medium-stakes exam:
     - A committee of content experts
  3. Low-stakes exam:
     - Instructor's informal judgment

Practice Analysis

- A type of study to identify the responsibilities of practicing individuals in a particular profession
- Purpose: identify knowledge, skills, and attitudes required for the profession

Questions & Comments

22 June 2017

หัวข้อ : การสร้างโจทย์และตัวเลือกข้อสอบปรนัย

การสร้างโจทย์และตัวเลือกข้อสอบ

รศ.นพ. เชิดศักดิ์ ไอรมณีรัตน์
ภาควิชาศัลยศาสตร์ คณะแพทยศาสตร์ศิริราชพยาบาล
มหาวิทยาลัยมหิดล

Twelve Steps for Test Development

1. Overall plan
2. Content definition
3. Test specification
4. Item development
5. Test assembly
6. Test production
7. Test administration
8. Scoring
9. Passing score
10. Reporting results
11. Item banking
12. Technical report

Downing SM, Haladyna TM. Handbook of test development. Mahwah, NJ: Lawrence Earlbaum Associates 2006.

A Good MCQ Item

1. Content
2. Structure

Guidelines for MCQ items

- Content guidelines
- Format guidelines
- Stem guidelines
- Option guidelines

40

Content Guidelines

- Focus on a single idea for each item
- Avoid trivial content
- Avoid opinion-based items
- Avoid direct quotes from textbooks
- Keep item content independent from one another

41

Format Guidelines

- Simplify vocabulary and sentence structures
- Avoid presenting unrelated information, minimize reading time
- Proofread each item for correct grammar, punctuation, and spelling

43

Stem Guidelines

- Make the question as clear as possible
- Avoid using negative words (not, except)
- Place the main idea of an item in the stem, not in options

45

Option Guidelines

- Develop as many effective options as you can
- Vary the location of the correct answers
- Keep options independent
- Keep options homogeneous
- Keep the length of options about the same
- Avoid “none of above” or “all of above”
- Avoid giving clues

47

Questions & Comments

“I've missed more than 9000 shots in my career. I've lost almost 300 games. 26 times, I've been trusted to take the game winning shot and missed. I've failed over and over and over again in my life. And that is why I succeed.”

Michael Jordan

การสร้างข้อสอบปรนัย

ผู้ช่วยศาสตราจารย์ นายแพทย์เชิดศักดิ์ ไอรมณีรัตน์
ภาควิชาศัลยศาสตร์, คณะแพทยศาสตร์ศิริราชพยาบาล, มหาวิทยาลัยมหิดล, กรุงเทพมหานคร ๑๐๗๐๐.

ข้อสอบปรนัย (multiple-choice question) เป็นรูปแบบการประเมินผลที่นิยมใช้กันอย่างแพร่หลายในวงการแพทยศาสตรศึกษาเนื่องด้วยคุณสมบัติที่ดีหลายประการด้วยกัน ได้แก่ ประสิทธิภาพในการประเมินความรู้ปริมาณมากในเวลาอันสั้น ผลการประเมินที่ไม่มีผลกระทบจากความรู้สึกส่วนตัวของผู้ตรวจให้คะแนน คะแนนที่มีความเที่ยงสูง รวมถึงผลการวิจัยจำนวนมากที่สนับสนุนความถูกต้องของผลการประเมินด้วยข้อสอบปรนัย[๑-๒] ข้อสอบปรนัยที่พัฒนาขึ้นอย่างดีนั้นสามารถวัดความรู้ได้ทั้งระดับการจดจำ การทำความเข้าใจ และการประยุกต์ความรู้ไปใช้ในการดูแลคนไข้[๓-๕] อย่างไรก็ดีผลการศึกษาวิจัยเกี่ยวกับคุณภาพของข้อสอบปรนัยที่พัฒนาขึ้นใช้ในโรงเรียนแพทย์หลายแห่งพบว่าข้อสอบจำนวนไม่น้อยมีลักษณะที่ไม่เหมาะสม[๕-๖] ข้อสอบปรนัยที่ถูกพัฒนาขึ้นอย่างไม่ถูกหลักการนั้นส่งผลเสียหลายอย่าง เช่นทำให้ข้อสอบยากขึ้นโดยไม่จำเป็น ทำให้ผู้สอบเกิดความสับสน ทำให้ผู้สอบบางกลุ่มเสียเปรียบผู้สอบคนอื่น ทำให้การตัดสินผลสอบผิดพลาด เป็นต้น[๖-๗] ดังนั้นการออกข้อสอบปรนัยที่ดี วางอยู่บนหลักการที่ถูกต้องจึงมีความสำคัญมากในการควบคุมคุณภาพการศึกษาในโรงเรียนแพทย์ บทความนี้จึงถูกเขียนขึ้นเพื่อเป็นการรวบรวมหลักการพื้นฐานในการออกข้อสอบปรนัยที่ได้รับการยอมรับกันทั่วไปในวงการวัดและประเมินผล ผู้นิพนธ์หวังว่าข้อแนะนำต่าง ๆ ที่ได้นำเสนอในบทความนี้จะเป็นแนวทางที่เป็นประโยชน์ในการพัฒนาข้อสอบปรนัยที่มีคุณภาพให้ผู้อ่านไม่มากก็น้อย

รูปแบบพื้นฐานของข้อสอบปรนัย

ข้อสอบปรนัยคือข้อสอบชนิดที่มีคำถามแล้วมีตัวเลือกให้ผู้สอบเลือกตัวเลือกที่เหมาะสมเพื่อตอบคำถามดังกล่าว ข้อสอบปรนัยสามารถแบ่งออกได้เป็น ๒ รูปแบบ[๘] ได้แก่

๑. ข้อสอบถูกผิด (True/false item)

ในข้อสอบประเภทนี้จะมีข้อความให้ผู้สอบพิจารณาว่าถูกหรือผิด ในยุคแรกข้อสอบเหล่านี้แต่ละข้อจะแยกเป็นอิสระจากกัน ผู้สอบตัดสินใจว่าข้อความแต่ละข้อถูกหรือผิดโดยไม่เกี่ยวข้องกับข้อความในข้ออื่น ต่อมามีผู้พัฒนาข้อสอบเป็นชุดของข้อความ (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)

ข้อสอบชนิดถูกผิดนี้เคยเป็นที่นิยมมากในวงการแพทยศาสตรศึกษาอยู่ระยะหนึ่งเนื่องจากสามารถทดสอบความรู้ได้ปริมาณมาก แต่ข้อสอบชนิดนี้มีข้อจำกัดที่สำคัญคือสามารถใช้ได้เฉพาะกับเนื้อหาที่มีความถูกผิดชัดเจนเท่านั้น ซึ่งการตัดสินใจทางการแพทย์ส่วนมากไม่เป็นเช่นนั้น การตัดสินใจในการวินิจฉัย การตรวจค้นเพิ่มเติม หรือการรักษาผู้ป่วยส่วนใหญ่นั้นแพทย์ตัดสินใจเลือกระหว่างทางเลือกที่แตกต่างกันสามสี่อย่างซึ่งทุกทางเลือกมีความเป็นไปได้ มีส่วนถูก หรือมีความเหมาะสมในบางด้าน

แต่ก็มีความไม่เหมาะสมในด้านอื่นด้วย เช่นการเลือกใช้ยาในผู้ป่วยที่มีการติดเชื้อ นักศึกษาแพทย์มักรู้ว่าควรใช้ยาปฏิชีวนะ ซึ่งยาปฏิชีวนะหลายชนิดก็รักษาการติดเชื้อชนิดนั้น ๆ ได้ แต่นักศึกษาต้องเลือกระหว่างยาที่ล้วนใช้ได้ในการรักษานั้นว่ายาใดที่มีประสิทธิภาพสูงสุด เหมาะสมที่สุดกับชนิดของเชื้อก่อโรคที่พบบ่อยในการติดเชื้อนั้น 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 หรือภาวะหรือผู้ป่วยตามด้วยคำถาม หรือเว้นช่องว่างสำหรับเติมคำหรือข้อความที่เหมาะสมลงไป

๒. ตัวเลือก (options) คือคำ หรือข้อความที่ผู้ออกข้อสอบนำเสนอตามหลังจากโจทย์เพื่อให้ผู้สอบเลือกไปใช้ตอบคำถาม หรือเติมลงในช่องว่างในโจทย์

๒.๑ ตัวเลือกที่ถูกต้อง (correct option) เป็นคำตอบที่ถูกต้องมีเพียงตัวเลือกเดียวต่อข้อสอบข้อหนึ่ง

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

ข้อแนะนำพื้นฐานของการเขียนข้อสอบปรนัย

มีผู้เชี่ยวชาญทางการประเมินผลให้ข้อแนะนำจำนวนมากในการเขียนข้อสอบปรนัย เคยมีผู้รวบรวมไว้ถึง ๔๓ ข้อ[๒,๙] ในที่นี้ผู้นิพนธ์ขอนำเสนอเฉพาะข้อแนะนำที่ได้รับการยอมรับอย่างกว้างขวางและสามารถประยุกต์ใช้ได้ชัดเจนในการพัฒนาข้อสอบทางการแพทย์ โดยจะทำการจัดหมวดหมู่ของข้อแนะนำเหล่านี้ออกเป็น ๔ กลุ่มด้วยกัน ได้แก่ (๑) เนื้อหาข้อสอบ, (๒) การจัดรูปแบบข้อสอบ, (๓) การเขียนโจทย์, และ (๔) การเขียนตัวเลือก

๑. เนื้อหาข้อสอบ

๑.๑ ข้อสอบหนึ่งข้อควรมุ่งเน้นประเมินความรู้เพียงเรื่องเดียว

ก่อนเริ่มเขียนข้อสอบอาจารย์ผู้ออกข้อสอบควรตั้งวัตถุประสงค์ให้ชัดเจนว่าต้องการประเมินความรู้ของผู้สอบในเรื่องใด และเขียนโจทย์เพื่อตอบสนองวัตถุประสงค์ดังกล่าวเท่านั้น เนื่องจากเนื้อหาวิชาทางการแพทย์มีมาก อาจารย์แต่ละท่านเมื่อทำการสอนไปแล้วจึงอยากจะทดสอบความรู้ในหลายเรื่องที่ตนได้สอนไป แต่กลับมีโควต้าจำกัดในการออกข้อสอบ 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 สิ่งที่ควรทำการประเมินคือความสามารถในการประยุกต์ใช้ความรู้ในทางคลินิก (application of knowledge) ไม่แนะนำการทดสอบวัดความสามารถในการจดจำเป็นหลัก อย่างไรก็ตามการที่แนะนำให้ออกข้อสอบที่เน้นการประยุกต์ใช้ความรู้ ไม่ได้หมายความว่าการแก้ปัญหาผู้ป่วยนั้นไม่ต้องใช้ความจำเลย ตรงกันข้ามการจดจำเนื้อหาเป็นพื้นฐานที่สำคัญในการแก้ปัญหาทางคลินิก ผู้สอบย่อมต้องจำเนื้อหาได้บ้าง จึงจะประยุกต์องค์ความรู้ดังกล่าวไปแก้โจทย์ปัญหาที่นำเสนอได้

๑.๓ หลีกเลี่ยงการถามความรู้ในเรื่องที่ยังมีความขัดแย้งกันในแนวทางปฏิบัติ (controversy)

ความรู้ทางการแพทย์ในหลายหัวข้อยังเป็นเรื่องที่ผู้เชี่ยวชาญยังมีความเห็นแตกต่างกัน ผู้ป่วยรายเดียวกันไปพบแพทย์สองคนอาจได้รับการรักษาที่แตกต่างกันซึ่งวิธีการรักษาทั้งสองวิธีก็มีงานวิจัยสนับสนุนด้วยกันทั้งคู่ อย่างนี้จัดว่ายังคงมีความขัดแย้ง (controversy) ในเรื่องดังกล่าวอยู่ เนื้อหาในลักษณะนี้ไม่ควรนำมาออกสอบด้วยข้อสอบปรนัย เนื่องจากในขณะที่ทำข้อสอบอยู่นั้น ผู้สอบไม่มีทางรู้ได้เลยว่าอาจารย์ผู้ออกข้อสอบอ้างอิงจากตำราหรือบทความวิชาการใด เนื้อหาที่ยังมีความขัดแย้ง ที่ผู้เชี่ยวชาญจากต่างสถาบันมีแนวทางในการปฏิบัติที่ต่างกันนี้แนะนำให้ใช้ข้อสอบในรูปแบบอื่นในการทดสอบเช่น ข้อสอบอัตนัย เป็นต้น

๑.๔ หลีกเลี่ยงการลอกประโยคหรือข้อความจากตำราโดยตรง

ดังได้กล่าวแล้วว่าข้อสอบที่ดีควรมุ่งเน้นการประเมินความเข้าใจ หรือ การประยุกต์ใช้ความรู้ ไม่ควรออกข้อสอบที่ประเมินความสามารถในการจำรายละเอียดปลีกย่อย การออกข้อสอบโดยวิธีการเปิดตำราแล้วคัดลอกประโยคจากตำราโดยตรงมักจะลงเอยด้วยข้อสอบที่ทดสอบความจำว่าผู้สอบท่องเนื้อหาในตำราตรงส่วนนั้นได้หรือไม่ ข้อสอบที่ดีควรได้จากการดูผู้ป่วย โจทย์ที่ดีควรเป็นปัญหาของผู้ป่วยที่พบในการทำงานนั่นเอง ตัวเลือกก็ได้จากข้อผิดพลาดที่นักศึกษาหรือแพทย์ประจำบ้านมักปฏิบัติกับผู้ป่วยแล้วทำให้ผลการรักษาไม่ดีนั่นเอง

๑.๕ หลีกเลี่ยงการนำเสนอข้อสอบที่ประเมินความรู้ในเรื่องเดียวกันสองข้อในข้อสอบชุดเดียวกัน

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

๒. การจัดรูปแบบข้อสอบ

๒.๑ เลือกใช้คำศัพท์หรือรูปประโยคที่ง่ายต่อการทำความเข้าใจ

อาจารย์ผู้ออกข้อสอบต้องระลึกไว้เสมอว่าข้อสอบที่อาจารย์ออกเพื่อใช้ในการประเมินผลนักศึกษาแพทย์หรือแพทย์ประจำบ้านนั้นมีวัตถุประสงค์เพื่อทดสอบความรู้ทางการแพทย์เป็นสำคัญ มิใช่การประเมินความรู้ทางภาษาศาสตร์ ดังนั้นการเขียนข้อสอบของอาจารย์ควรเลือกใช้รูปแบบประโยคที่ง่ายต่อการทำความเข้าใจ อย่าเขียนประโยคซับซ้อนที่มีความยาวประโยคละหลายบรรทัด มุ่งเน้นให้ภาษาเป็นสื่อในการนำเสนอความคิดของอาจารย์ผู้ออกข้อสอบไปยังผู้สอบ อย่าให้

ภาษาเป็นอุปสรรคในการสื่อสาร การจะเลือกใช้ภาษาใดในการเขียนข้อสอบนั้นให้พิจารณาตามข้อกำหนดขององค์กรหรือหน่วยงานที่ควบคุมการสอบที่อาจารย์ส่งข้อสอบไปให้ใช้ ข้อสอบที่ใช้ในระดับการศึกษาหลักสูตรแพทยศาสตรบัณฑิตทั้งในระดับคณะ หรือข้อสอบที่ใช้ในการสอบระดับประเทศในปัจจุบันยังนิยมใช้ข้อสอบที่เขียนด้วยภาษาไทยโดยมีการใช้ศัพท์เทคนิคเป็นภาษาอังกฤษเหมือนดังภาษาที่แพทย์ใช้สื่อสารกันในการทำงานปกติ ส่วนข้อสอบในระดับหลังปริญญามีหลายการสอบที่ภาควิชา หรือราชวิทยาลัยที่เกี่ยวข้องกำหนดให้ใช้ภาษาอังกฤษทั้งหมด ก่อนที่อาจารย์จะสร้างข้อสอบต้องมีการศึกษาข้อกำหนดของแต่ละการสอบให้ดี

๒.๒ หลีกเลี่ยงการนำเสนอข้อมูลที่ไม่เกี่ยวข้องกับการแก้ปัญหาของโจทย์ข้อนั้น

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 ก็ไม่ควรนำข้อสอบข้อนั้นไปใช้สอบเลย ควรให้มีคณะกรรมการข้อสอบซึ่งประกอบไปด้วยอาจารย์หลายท่านช่วยกันตรวจสอบและพิจารณาปรับแก้ข้อสอบทุกข้อก่อนนำไปใช้จริงเสมอ เนื่องจากผู้เขียนข้อสอบย่อมเข้าใจสิ่งที่ตนเขียนเสมอ แต่เมื่อผู้อื่นอ่านแล้วอาจพบว่ามีเนื้อหาที่กำกวมหรือเข้าใจโจทย์ต่างออกไปได้ การปรับแก้เนื้อหาที่มีความกำกวม หรือเฉลยซึ่งอาจารย์บางท่านอาจไม่เห็นด้วยให้ได้ข้อสอบที่มีความชัดเจน และอาจารย์ทุกท่านยอมรับในคำเฉลยได้ก่อนจะนำข้อสอบไปทำการสอบจริงย่อมเป็นสิ่งที่ดีกว่าการตรวจพบปัญหาหลังจากสอบเสร็จแล้วซึ่งต้องมาตัดสินใจกันอีกว่าจะทำอย่างไรกับการคิดคะแนนของข้อสอบข้อดังกล่าว

๓. การเขียนโจทย์

๓.๑ เขียนโจทย์ให้มีความชัดเจน ผู้สอบทุกคนอ่านแล้วมีความเข้าใจตรงกัน

ข้อแนะนำนี้อาจดูเหมือนตรงไปตรงมา แต่กลับเป็นปัญหาที่พบบ่อยมากในการพัฒนาข้อสอบปรนัย ประเด็นสำคัญคือโจทย์ที่ดีนั้นต้องมีความสมบูรณ์ในตัวเองโดยไม่ต้องอาศัยตัวเลือก โจทย์ข้อสอบที่ดีนั้นเมื่ออ่านโจทย์เสร็จแล้ว หากผู้สอบมีความรู้ในเรื่องที่ทำการประเมินนั้น เขาจะบอกคำตอบได้โดยไม่จำเป็นต้องอ่านตัวเลือกเลย ดังนั้นเมื่ออาจารย์เขียนข้อสอบเสร็จแล้วแนะนำให้ลองปิดตัวเลือกแล้วอ่านเฉพาะโจทย์ดู หากอาจารย์อ่านแล้วบอกได้ว่าโจทย์ถามอะไรและบอกได้ว่าควรตอบอะไรโดยไม่ต้องอ่านตัวเลือกจัดว่าข้อสอบข้อดังกล่าวมีโจทย์ที่มีความชัดเจน

๓.๒ เรียบเรียงเนื้อหาให้ใจความสำคัญของข้อสอบอยู่ในโจทย์

เนื่องจากข้อสอบปรนัยมีตัวเลือกที่อาจารย์ต้องสร้างขึ้นหลายตัวเลือก บางครั้งอาจารย์ผู้พัฒนาข้อสอบอาจเผลอเรอนำเอาใจความสำคัญไปใส่ไว้ในตัวเลือก ซึ่งทำให้เนื้อหาในโจทย์ขาดสาระสำคัญ อ่านโจทย์แล้วไม่เข้าใจว่าผู้ออกข้อสอบต้องการถามความรู้เรื่องอะไร ตัวอย่างข้อสอบที่ไม่เป็นไปตามข้อแนะนำนี้คือข้อสอบที่ถามว่า ข้อใดต่อไปนี้ถูกต้อง หรือข้อใดต่อไปนี้ไม่ถูกต้อง แล้วเขียนรายละเอียดเกี่ยวกับโรค หรือการรักษาบางอย่างในตัวเลือกแต่ละข้อ ข้อสอบในลักษณะนี้มักทำให้

ผู้สอบต้องอ่านข้อสอบย้อนไปมาหลายรอบกว่าจะเข้าใจจุดประสงค์ของข้อสอบ แล้วจึงตัดสินใจเลือกคำตอบ โดยทั่วไปแนะนำให้อาจารย์นำเสนอรายละเอียดต่าง ๆ ไว้ในตัวโจทย์ให้มากที่สุด ส่วนตัวเลือกเขียนเป็นคำหรือข้อความสั้น ๆ

๓.๓ หลีกเลี่ยงการเขียนโจทย์ที่มีรูปประโยคเป็นเชิงปฏิเสธ

โจทย์ที่ดีไม่ควรอยู่ในประโยคเชิงปฏิเสธ เช่น ถามถึงสิ่งที่เป็นข้อยกเว้น สิ่งที่ไม่ควรปฏิบัติ สิ่งที่พบน้อยที่สุด หรือสิ่งที่ไม่น่านึกถึงเป็นต้น งานวิจัยส่วนใหญ่พบว่าข้อสอบที่มีโจทย์ในรูปแบบปฏิเสธเหล่านี้มีระดับความยากง่ายไม่ต่างจากข้อสอบอื่น ๆ แต่งานวิจัยบางชิ้นพบว่าข้อสอบที่มีโจทย์ในรูปแบบปฏิเสธมีความยากมากกว่าข้อสอบอื่นชัดเจนโดยเฉพาะในข้อสอบวัดความรู้ระดับสูง[๑๐-๑๒] แต่ผู้เชี่ยวชาญในการประเมินผลส่วนใหญ่มีความเห็นพ้องกันว่าข้อสอบประเภทนี้สามารถสร้างความสับสนให้กับผู้สอบได้ จึงไม่แนะนำให้ใช้ แต่หากอาจารย์ผู้ออกข้อสอบมีความจำเป็นต้องใช้ข้อสอบที่มีการใช้คำปฏิเสธในโจทย์ แนะนำให้พิมพ์คำปฏิเสธให้เด่นชัด โดยใช้ตัวหนาและขีดเส้นใต้เพื่อให้ผู้สอบเห็นชัด[๑๐]

๔. การเขียนตัวเลือก

๔.๑ เขียนตัวเลือกที่มีประสิทธิภาพให้มีจำนวนมากที่สุดเท่าที่เหมาะสมกับบริบท

เรื่องจำนวนตัวเลือกที่เหมาะสมนี้เป็นเรื่องที่ผู้เชี่ยวชาญด้านการประเมินผลจำนวนมากสนใจ มีงานวิจัยเกี่ยวกับเรื่องจำนวนตัวเลือกที่เหมาะสมในข้อสอบปรนัยอยู่มากมาย[๙] อาจารย์ผู้ออกข้อสอบส่วนมากจะคุ้นเคยกับข้อสอบปรนัยชนิดที่มีห้าตัวเลือก บ่อยครั้งที่อาจารย์ออกข้อสอบแล้วนึกตัวเลือกได้เพียงสามหรือสี่ตัว จึงเกิดคำถามว่าจำเป็นต้องมีตัวเลือกครบห้าตัวเลือกหรือไม่ งานวิจัยบางชิ้นพบว่าการลดจำนวนตัวเลือกลงทำให้ข้อสอบง่ายขึ้น[๑๓-๑๔] แต่งานวิจัยบางชิ้นพบว่าการลดจำนวนตัวเลือกลงทำให้ได้ข้อสอบยากขึ้น[๑๕-๑๖] ผู้เชี่ยวชาญในการประเมินผลเสนอว่าข้อสอบปรนัยที่มีตัวเลือกเพียงสามตัวเลือกก็สามารถทดสอบความรู้ได้อย่างมีประสิทธิภาพ[๑, ๙-๑๐, ๑๗] แต่มีอาจารย์จำนวนไม่น้อยที่ไม่สบายใจที่มีตัวเลือกในข้อสอบแต่ละข้อน้อยกว่าห้าตัวเลือกด้วยกังวลว่าจะทำให้มีโอกาสสูงที่ผู้สอบที่ไม่มีความรู้จะเดาสุ่มได้คำตอบที่ถูกต้อง แต่จากข้อมูลที่ปรากฏในปัจจุบันพบว่าผู้สอบในการสอบในระดับสูงนั้นพฤติกรรมการเดาสุ่มโดยที่ผู้สอบปราศจากความรู้นั้นน่าจะมีบทบาทน้อยมาก ผู้สอบส่วนใหญ่มักพอมีความรู้บ้างและสามารถตัดตัวเลือกที่ไม่สมเหตุสมผลอย่างชัดเจนได้[๑] ในการศึกษาข้อสอบปรนัยส่วนใหญ่พบตัวเลือกที่ไม่ทำงานเป็นจำนวนไม่น้อย[๑๙] ข้อมูลที่ได้จากการวิเคราะห์ข้อสอบปรนัยที่ใช้ในทางแพทยศาสตรศึกษาในประเทศไทยหลายครั้งก็สอดคล้องกับงานวิจัยในต่างประเทศที่พบว่าข้อสอบส่วนใหญ่มักมีตัวเลือกที่ทำงานจริงราวสามหรือสี่ตัวเลือก มีข้อสอบน้อยข้อมากที่ตัวเลือกทั้งห้าตัวเลือกทำงานอย่างมีประสิทธิภาพ

ด้วยข้อมูลจากการศึกษาต่าง ๆ ข้อแนะนำในการออกข้อสอบปรนัยในปัจจุบันคือให้อาจารย์เขียนจำนวนตัวเลือกมากที่สุดที่มีความเหมาะสมกับเนื้อหาโจทย์ ไม่จำเป็นต้องเขียนตัวเลือก ๕ ตัวเลือกเสมอไป เนื่องจากตัวเลือกที่ห้าที่เขียนขึ้นเพื่อเติมเต็มโดยไม่สมเหตุสมผลนั้นมักไม่ค่อยมีคนเลือก หากเนื้อหาที่อาจารย์นำมาสอบมีตัวเลือกที่เหมาะสมเพียงสามหรือสี่ตัวเลือกก็เขียนจำนวนตัวเลือกเพียงสามหรือสี่ตัวเลือก[๑๐] แต่อย่างไรก็ตามให้อาจารย์ศึกษาข้อกำหนดของแต่ละการสอบที่อาจารย์เกี่ยวข้องด้วย เนื่องจากนโยบายของแต่ละการสอบแตกต่างกันไป องค์กรที่จัดสอบทางแพทยศาสตรศึกษาจำนวนไม่น้อยยังคงตั้งข้อกำหนดให้ใช้ข้อสอบ 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 เป็นยากลุ่มย่อยของตัวเลือก 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 ก็น่าจะเป็นยาปฏิชีวนะเช่นกันไม่ใช่ยาเคมีบำบัด หรือยาต้านเชื้อรา) การที่มีตัวเลือกที่ไม่เข้าพวก ไม่มีความเป็นเนื้อเดียวกันกับตัวเลือกอื่นเป็นคำบอกใบ้ในการตัดตัวเลือกที่ผู้สอบนิยมใช้มาก ดังนั้นอาจารย์ผู้ออกข้อสอบควรหลีกเลี่ยง

ในบางบริบทของการดูแลรักษาผู้ป่วย สิ่งที่แพทย์ต้องตัดสินใจเลือกอาจมีทั้งการเลือกที่จะให้การรักษาเลยหรือจะส่งตรวจค้นเพิ่มเติมก่อน 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[๑๐] จึงแนะนำให้หลีกเลี่ยงการใช้

ตัวเลือก “ไม่มีข้อใดถูก” เป็นประเด็นที่ผู้เชี่ยวชาญในการประเมินผลยังคงถกเถียงกันอยู่บ้าง ผู้เชี่ยวชาญบางส่วนเห็นว่าไม่ควรใช้ตัวเลือกประเภทนี้ แต่ผู้เชี่ยวชาญบางส่วนให้ความเห็นว่าสามารถใช้ได้ในบางกรณี[๑๐] เหตุผลที่ตัวเลือกชนิดนี้เป็นปัญหาคือการใช้ตัวเลือกนี้มักสร้างความลำบากใจให้กับผู้สอบในการเลือกคำตอบที่ถูกในกรณีที่ตัวเลือกแต่ละตัวเลือกไม่ถูกหรือผิดชัดเจน เพราะผู้สอบจะต้องทำการเปรียบเทียบตัวเลือกที่นำเสนอในข้อสอบกับทางเลือกอื่น ๆ ที่เขานึกได้[๙] หากโจทย์ถามว่ายาใดที่ควรให้แก่ผู้ป่วย แล้วมีชื่อยาสี่ชนิด และมีตัวเลือก “ไม่มีข้อใดถูก” นอกจากที่ผู้สอบต้องนึกว่าในบรรดายา ที่ปรากฏในตัวเลือกนั้นเหมาะสมหรือไม่แล้วเขายังนึกต่อไปอีกว่ามียาอื่นใดที่สามารถให้ในผู้ป่วยรายนี้ได้อีก หากเขานึกออกว่ามียาอื่นที่น่าจะเหมาะสมกับผู้ป่วยมากกว่ายาในตัวเลือก (ด้วยเหตุผลที่อาจแตกต่างไปจากที่อาจารย์ผู้ออกข้อสอบคิด) เขาก็จะเลือก “ไม่มีข้อใดถูก”

การใช้ตัวเลือก “ไม่มีข้อใดถูก” จะยิ่งเป็นปัญหามากขึ้นในข้อสอบที่ถามถึงสิ่งที่ไม่ควรทำ เช่นยาใดไม่ควรใช้ในผู้ป่วย ซึ่งนอกจากยาที่นำเสนอในตัวเลือกแล้วย่อมมียาชนิดอื่นอีกมากมายในบัญชียาที่ไม่เหมาะสม ซึ่งไม่มี

ทางที่ใครจะรู้ได้ว่าการที่ผู้สอบเลือกตอบ “ไม่มีข้อใดถูก” นั้นเขาคิดถึงยาใด และยานั้นไม่เหมาะสมมากไปกว่ายาที่มีอยู่ในตัวเลือกหรือไม่ งานวิจัยทั้งหมดที่ศึกษาถึงตัวเลือกชนิดนี้ได้ข้อสรุปที่ตรงกันว่าข้อสอบที่ใช้ตัวเลือกประเภทนี้เพิ่มระดับความยากให้ข้อสอบ[๑๖] โดยทั่วไปแล้วจึงไม่แนะนำให้ใช้ตัวเลือกประเภทนี้ในการสอบทางแพทยศาสตรศึกษาซึ่งทางเลือกสำหรับสถานการณ์ที่นำเสนอมีได้มากและการตัดสินใจเลือกคำตอบต้องอาศัยการเปรียบเทียบข้อดีข้อเสียของแต่ละตัวเลือก

สรุป

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

มีข้อควรพิจารณาในการประยุกต์ใช้ข้อแนะนำเหล่านี้ในการพัฒนาข้อสอบที่ผู้นิพนธ์ขอกล่าวถึงประการหนึ่งคือ แม้ว่าข้อแนะนำที่กล่าวถึงเหล่านี้หลายข้อมีการศึกษาวิจัยสนับสนุนที่ชัดเจน แต่สิ่งเหล่านี้ก็เป็นเพียงข้อแนะนำว่าผู้ออกข้อสอบควรปฏิบัติ ไม่ใช่กฎเกณฑ์ตายตัว การเขียนข้อสอบปรนัยนั้นเป็นงานที่ต้องอาศัยทั้งศาสตร์และศิลปะผสมผสานกันอย่างเหมาะสม หาใช่สูตรคณิตศาสตร์ที่ไม่มีข้อยกเว้น ผู้นิพนธ์ไม่คาดหวังให้อาจารย์ผู้พัฒนาข้อสอบยึดข้อแนะนำเหล่านี้เสมือนกฎเกณฑ์ตายตัวที่ต้องทำตามในทุกกรณี หากแต่ต้องการให้อาจารย์ใช้เป็นแนวทางในการสร้างข้อสอบ ในบางบริบทผู้ออกข้อสอบอาจเลือกที่จะไม่ปฏิบัติตามข้อแนะนำบางประการได้บ้าง แต่การที่จะไม่ปฏิบัติตามข้อแนะนำเหล่านี้นั้นต้องมีเหตุผลที่เหมาะสม และควรทำไม่บ่อยนัก ยกตัวอย่างเช่นข้อแนะนำว่า โจทย์ไม่ควรเขียนถามข้อยกเว้น จะพบได้ว่ามีบางบริบทที่การรู้ข้อยกเว้น หรือข้อห้ามปฏิบัติก็เป็นองค์ความรู้ที่สำคัญในการดูแลรักษาผู้ป่วย ดังนั้นในบริบทที่เหมาะสมผู้นิพนธ์เองก็เห็นด้วยว่าอาจเขียนโจทย์ที่ถามข้อยกเว้นได้ แต่อย่างไรก็ตามการจะไม่ปฏิบัติตามข้อแนะนำนี้ต้องไม่ทำบ่อยจนเกินจำเป็น หากออกข้อสอบ ๑๐๐ ข้อ จะมีข้อสอบที่ถามข้อยกเว้น ปะปนมาบ้าง 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 medial education. Dordrecht: Kluwer Academic Pubishers, 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 Assoicates, 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.

22 June 2017

หัวข้อ : ข้อผิดพลาดที่พบบ่อยในการสร้างข้อสอบปรนัย

เวชบันทึกศิริราช บทความทั่วไป

ข้อผิดพลาดที่ควรระวังในการสร้างข้อสอบปรนัย

ผู้ช่วยศาสตราจารย์ นายแพทย์เชิดศักดิ์ ไอรมณีรัตน์
ภาควิชาศัลยศาสตร์, คณะแพทยศาสตร์ศิริราชพยาบาล, มหาวิทยาลัยมหิดล, กรุงเทพมหานคร ๑๐๗๐๐.

ข้อผิดพลาดที่ควรระวังในการสร้างข้อสอบปรนัย

ข้อสอบปรนัย (multiple-choice question) เป็นรูปแบบการประเมินผลที่นิยมใช้กันอย่างแพร่หลาย ในวงการแพทยศาสตรศึกษา ข้อสอบชนิดนี้เป็นที่ชื่นชอบของนักศึกษาผู้เข้าสอบจำนวนมากเนื่องจากมีคำตอบให้เลือก หากไม่มีความรู้ก็สามารถเดาได้ ซึ่งต่างไปจากข้อสอบประเภทอัตนัยซึ่งผู้สอบต้องเขียนคำตอบจากความคิดของตนเอง[๑] ดังนั้นข้อสอบปรนัยจึงเป็นข้อสอบที่ผู้สอบทำได้ง่าย แต่ในทางตรงข้ามข้อสอบปรนัยเป็นข้อสอบที่สร้างปัญหาให้กับอาจารย์ผู้สร้างข้อสอบไม่น้อย เนื่องจากในกระบวนการเขียนข้อสอบปรนัยแต่ละข้อนั้นต้องใช้ทักษะอย่างมาก ต้องใช้ทั้งศาสตร์และศิลป์ และบ่อยครั้งอาจารย์ผู้สร้างข้อสอบก็ถูกขอให้ทำการปรับแก้ข้อสอบเนื่องจากคณะกรรมการพิจารณาข้อสอบมีความเห็นว่ารายละเอียดในข้อสอบไม่เหมาะสม มีการศึกษาวิจัยพบว่าคุณภาพของข้อสอบปรนัยที่พัฒนาขึ้นในโรงเรียนแพทย์หลายแห่งนั้นไม่สู้ดีนัก มีข้อสอบที่มีลักษณะไม่เหมาะสมอยู่จำนวนไม่น้อย[๒-๓] ข้อสอบปรนัยที่มีลักษณะไม่เหมาะสมเหล่านี้ส่งผลเสียต่อการสอบได้หลายประการ เช่น ทำให้ข้อสอบยากขึ้น สร้างความสับสนให้ผู้สอบ ทำให้ผู้สอบบางกลุ่มเสียเปรียบ และทำให้การตัดสินผลสอบผิดพลาด เป็นต้น[๓-๕] ดังนั้นการออกข้อสอบปรนัยที่มีคุณภาพดีจึงเป็นงานที่มีความสำคัญและท้าทายความสามารถ

การสร้างข้อสอบปรนัยที่มีคุณภาพดีนั้นควรเริ่มต้นจากการมีองค์ความรู้พื้นฐานในการสร้างข้อสอบแล้วเกิดการฝึกฝนทักษะ สั่งสมประสบการณ์ในการออกข้อสอบจนเกิดความชำนาญ ปัญหาที่พบบ่อยในโรงเรียนแพทย์หลายแห่งคือมีอาจารย์จำนวนไม่น้อยที่ได้รับมอบหมายให้ออกข้อสอบปรนัย โดยไม่ได้มีการพัฒนาองค์ความรู้พื้นฐานที่เหมาะสมก่อน ซึ่งเป็นเหตุให้มีข้อสอบปรนัยที่มีลักษณะไม่เหมาะสมตามหลักการออกข้อสอบปะปนมาในข้อสอบที่ให้นักศึกษาแพทย์และแพทย์ประจำบ้านทำอยู่บ้าง ผู้นิพนธ์จึงเห็นความสำคัญของการเผยแพร่องค์ความรู้พื้นฐานของการออกข้อสอบปรนัย องค์ความรู้พื้นฐานในการสร้างข้อสอบปรนัยนั้นมีสองส่วน ส่วนแรกเป็นหลักการของการสร้างข้อสอบทั่วไปซึ่งได้มีผู้รวบรวมเป็นข้อแนะนำตีพิมพ์ในตำราและวารสารทางวิชาการอยู่บ้าง[๑,๕-๗] ส่วนที่สองเป็นข้อผิดพลาดในการสร้างข้อสอบที่อาจารย์ผู้ออกข้อสอบพึงหลีกเลี่ยง ในบทความนี้ผู้นิพนธ์จะมุ่งเน้นในส่วนที่สองนี้โดยจะรวบรวมข้อผิดพลาดในการสร้างข้อสอบปรนัย ที่อาจเป็นตัวบอกใบ้ให้ผู้สอบที่ไม่มีความรู้ในเรื่องที่ทำการทดสอบสามารถเลือกคำตอบที่ถูกต้องได้ ดังนั้นการที่อาจารย์ผู้ออกข้อสอบทราบถึงสิ่งเหล่านี้และหลีกเลี่ยงเสียจะส่งผลให้ข้อสอบปรนัยที่สร้างขึ้นสามารถใช้วัดองค์ความรู้ทางการแพทย์ได้จริง โดยปราศจากปัจจัยรบกวนจากการสังเกตพบสิ่งบอกใบ้คำตอบ

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

ในบทความนี้ผู้นิพนธ์ขอนำเสนอข้อผิดพลาดในการออกข้อสอบ ๗ กลุ่มด้วยกัน ได้แก่ (๑) ข้อผิดพลาดในไวยากรณ์, (๒) การใบ้คำตอบด้วยหลักตรรกะ, (๓) การใช้คำคุณศัพท์บอกระดับของความแน่ชัด, (๔) ความยาวของตัวเลือก, (๕) การใช้คำซ้ำในโจทย์และตัวเลือก, (๖) การเข้าพวกของคำ หรือข้อความที่ปรากฏในตัวเลือก, และ (๗) การบอกใบ้คำตอบโดยโจทย์ข้ออื่น

๑. ข้อผิดพลาดในไวยากรณ์

ตัวเลือกทุกตัวต้องสามารถตอบโจทย์ได้อย่างถูกต้องตามหลักไวยากรณ์ บ่อยครั้งอาจารย์ผู้ออกข้อสอบมุ่งความสนใจไปที่คำตอบที่ถูก และให้ความสนใจกับตัวลวงน้อยไปจนทำให้ตัวลวงผิดหลักไวยากรณ์[๕] โดยมักพบบ่อยในข้อสอบที่เป็นภาษาอังกฤษ ข้อผิดพลาดที่พบได้บ่อยเช่น ความไม่เข้ากันของ article (A, An, The) กับคำนามที่ตามหลัง, คำนามกับกิริยาที่ไม่เข้ากันในเชิงเอกพจน์หรือพหูพจน์, การเติมคำในประโยคที่เว้นว่างไว้สำหรับเติมคำนามแต่ตัวลวงเป็นกิริยาหรือเป็นคำนามในลักษณะที่ไม่เข้ากับรูปประโยค 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 gonorrheae*

B. *Chlamydia trachomatis and Giardia lamblia*

C. *Treponema pallidum*

D. *Ureaplasma urealyticum and Mycoplasma genitalium*

ในตัวอย่างที่ ๒ นี้โจทย์ถามหาเชื้อก่อโรค โดยใช้รูปประโยคถามหาคำตอบที่เป็นเอกพจน์ ดังนั้นคำตอบที่ถูกต้องย่อมมีเชื้อก่อโรคตัวเดียว ผู้สอบที่มีทักษะการทำข้อสอบดีจะตัดตัวเลือก B และ D ได้โดยไม่ต้องใช้ความรู้ทางการแพทย์

๒. การใบ้คำตอบด้วยหลักตรรกะ

ในการเขียนตัวเลือก อาจารย์ผู้ออกข้อสอบต้องระมัดระวังไม่ให้ผู้สอบสามารถตัดตัวเลือกได้ด้วยหลักตรรกศาสตร์ เนื่องจากผู้สอบที่มีทักษะการทำข้อสอบดีจะสามารถพิจารณาความเป็นไปได้ของตัวเลือกต่าง ๆ และตัดตัวลวงที่ไม่มีทางเป็นไปได้ตามหลักของเหตุและผลออกไปได้โดยไม่ต้องอาศัยความรู้เรื่องที่อาจารย์ตั้งเป้าหมายว่าจะทดสอบ

ตัวอย่างที่ ๓. ภาวะไส้เลื่อนบริเวณขาหนีบ (inguinal hernia) …….

A. พบในผู้ชายบ่อยกว่าผู้หญิง

B. พบในผู้หญิงบ่อยกว่าผู้ชาย

C. พบเกิดขึ้นในผู้หญิงและผู้ชายในอัตราเท่ากัน

D. พบบ่อยในผู้ที่มีเศรษฐานะยากจน

E. พบในผู้ที่มีภูมิลำเนาในทวีปเอเชีย มากกว่าผู้ที่มีภูมิลำเนาในทวีปยุโรป

ในตัวอย่างที่ ๓ นี้อาจารย์ผู้ออกข้อสอบต้องการวัดความรู้เรื่องอุบัติการณ์ของไส้เลื่อนขาหนีบ แต่หาก

พิจารณาตามหลักตรรกศาสตร์แล้ว ตัวเลือก A, B, และ C เพียงสามตัวเลือกก็ครอบคลุมสิ่งที่เป็นไปได้ทั้งหมดแล้ว (เนื่องจากมนุษย์มีสองเพศ ภาวะไส้เลื่อนนี้หากไม่มีอัตราการเกิดเท่ากันในสองเพศแล้วก็ต้องมีเพศใดเป็นมากกว่าอีกเพศหนึ่ง) ดังนั้นผู้สอบที่มีทักษะการทำข้อสอบดีสามารถตัดตัวเลือก D และ E ได้โดยไม่ต้องมีความรู้เรื่องไส้เลื่อนเลย

๓. การใช้คำคุณศัพท์บอกระดับของความแน่ชัด

อาจารย์ผู้ออกข้อสอบพึงระมัดระวังการใช้คำคุณศัพท์ที่บ่งบอกถึงความแน่ชัดของข้อความ ซึ่งจะมีหลายระดับ โดยทั่วไปแล้วคำคุณศัพท์ที่แสดงความแน่ชัดมาก แสดงความมั่นใจมาก (เช่น always, never) มักไม่ถูกต้อง เนื่องจากในทางการแพทย์นั้นมีความไม่แน่นอนเกิดขึ้นเป็นประจำ ข้อความที่บอกเล่าถึงสิ่งที่อาจเป็นไปได้โดยไม่ชี้ชัดลงไปว่าต้องเกิดขึ้นแน่นอน (เช่น may, might, can, could) มักเป็นข้อความที่ถูก

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 นี้ผู้สอบต้องเลือกข้อความเกี่ยวกับไส้เลื่อนขาหนีบที่ถูกต้องหนึ่งข้อความ หากสังเกตดูทั้งสี่ข้อความมีการใช้คำคุณศัพท์บอกความแน่ชัดของข้อความ ได้แก่ may (ตัวเลือก A), always (ตัวเลือก B), the only (ตัวเลือก C), inevitably (ตัวเลือก D) ซึ่งจะเห็นว่าตัวเลือก B, C, และ D เป็นข้อความที่แสดงความแน่ชัดว่าต้องเป็นแน่ ต้องใช่แน่นอน ไม่มีทางเลี่ยงได้ ข้อความทำนองนี้มีโอกาสสูงที่จะผิด ในทางตรงข้ามตัวเลือก A เป็นข้อความบอกว่ามีโอกาสเป็นได้โดยไม่ชี้ชัดว่าต้องเกิด ผู้สอบที่มีทักษะการทำข้อสอบดีจะตัดตัวเลือก B, C, และ D ได้โดยไม่ต้องอาศัยความรู้ทางการแพทย์เลย

๔. ความยาวของตัวเลือก

มีการตั้งข้อสังเกตว่าอาจารย์แพทย์มักชอบสอนและอธิบายแม้กระทั่งในการสอบอาจารย์แพทย์หลายท่านก็ติดนิสัยรักการสอนนี้มาด้วย ทำให้อาจารย์มักเขียนตัวเลือกที่ถูกต้องที่มีคำอธิบายประกอบอย่างครบถ้วนทำให้ตัวเลือกที่ถูกมักมีความยาวมากกว่าตัวลวง นักศึกษาผู้เข้าสอบจำนวนไม่น้อยรู้ถึงความจริงข้อนี้และมักเลือกตัวเลือกที่มีความยาวมากที่สุด หากเขาไม่สามารถหาคำตอบได้ด้วยความรู้ทางการแพทย์ที่เขามี

ตัวอย่างที่ ๕. ผู้หญิงอายุ ๒๘ ปี แต่งงานมานาน ๑ ปี ยังไม่มีบุตร คุมกำเนิดโดยการกินยาคุมเป็นประจำ สังเกตว่าตนเองน้ำหนักตัวเพิ่มขึ้นหลังจากกินยาคุมมาขอคำแนะนำเรื่องการคุมกำเนิด ท่านจะแนะนำอย่างไร

A. ให้เปลี่ยนไปใช้การใส่ห่วงอนามัย

B. ให้ใช้ถุงยางอนามัย

C. ให้กินยาคุมกำเนิดต่อได้เนื่องจากมีการศึกษาแล้วว่ายาคุมกำเนิดชนิดกินไม่ส่งผลให้เกิดการเพิ่มขึ้นของน้ำหนักตัว

D. ให้รับประทานยาลดความอ้วน

ในตัวอย่างที่ ๕ นี้จะสังเกตเห็นว่าตัวเลือก C มีการอธิบายเหตุผลประกอบส่งผลให้มีความยาวมากกว่าตัวเลือกอื่นชัดเจน ลักษณะเช่นนี้จะเป็นการบอกใบ้ให้นักศึกษาเลือกตัวเลือกนี้

๕. การใช้คำซ้ำในโจทย์และตัวเลือก

การใช้คำเดียวกัน หรือคำที่มีความหมายเหมือนกันในโจทย์และตัวเลือก มักเป็นการบอกใบ้ว่าตัวเลือกดังกล่าวเป็นตัวเลือกที่ถูกต้อง

ตัวอย่างที่ ๖. Which of the following statements is true regarding saccular 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 นี้โจทย์ถามถึง saccular theory ซึ่งหากแปลความหมายก็น่าจะเป็นเรื่องที่เกี่ยวข้องกับถุง (sac) ผู้สอบที่มีทักษะการทำข้อสอบดีจะหาตัวเลือกที่มีคำที่มีความหมายเกี่ยวกับถุง แล้วเลือกตัวเลือกดังกล่าวทันที ซึ่งในที่นี้จะพบคำว่า diverticulum ซึ่งมีความหมายว่าถุงในข้อ B การที่มีคำที่มีความหมายซ้ำกันเช่นนี้เป็นตัวบอกใบ้คำตอบที่อาจารย์ผู้ออกข้อสอบต้องตรวจตราให้ดีก่อนนำข้อสอบไปใช้

๖. การเข้าพวกของคำ หรือข้อความที่ปรากฏในตัวเลือก

ข้อสอบจำนวนไม่น้อยนำเสนอรายการของหลายอย่างในตัวเลือก (เช่น ชื่อการตรวจค้นเพิ่มเติม ชื่อโรค ชื่อยา ฯลฯ) มีผู้เชี่ยวชาญในการประเมินผลตั้งข้อสังเกตว่าในข้อสอบเหล่านี้ตัวเลือกที่ถูกต้องมักมีลักษณะเข้าพวกกับตัวเลือกอื่นมากที่สุด หากเป็นรายการของตัวเลือกที่ถูกก็คือข้อที่มีจำนวนรายการซ้ำกับตัวเลือกอื่นมากที่สุด ดังนั้นในการนำเสนอตัวเลือกอาจารย์ผู้ออกข้อสอบพึงระมัดระวังอย่าให้ตัวเลือกที่ถูกต้องมีลักษณะที่เข้าพวกได้อย่างชัดเจน พยายามทำตัวลวงอื่นให้มีลักษณะเข้าพวกให้ใกล้เคียงกับตัวเลือกที่ถูกต้อง

ตัวอย่างที่ ๗. โรคที่แพทย์วินิจฉัยผิดว่าเป็นไส้ติ่งอักเสบบ่อยที่สุดเรียงลำดับจากมากไปน้อยคือ

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 นี้โจทย์ถามชื่อโรค ตัวเลือกแสดงรายการชื่อโรค ตัวเลือกละสามโรค หากนับจำนวนของคำซ้ำจะพบว่าโรคที่กล่าวถึงบ่อยที่สุดคือ acute mesenteric lymphadenitis, pelvic inflammatory disease, twisted ovarian cyst, และ Meckel diverticulitis (กล่าวถึงโรคละ ๒ ครั้ง) ส่วนโรคที่เหลือกล่าวถึงโรคละครั้งเดียว ดังนั้นตัวเลือกที่มีพวกมากที่สุดคือตัวเลือก A ซึ่งเป็นคำตอบที่ถูกต้อง

การเข้าพวกของตัวเลือกที่ถูกนั้น 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 นี้โจทย์ถามถึงยาที่ควรให้กับผู้ป่วย ในตัวเลือกสี่ตัวเลือกนี้มียากินเพียงข้อเดียว (A) ที่เหลือเป็นยาฉีดสองขนานควบกัน ดังนั้นตัวเลือกข้อ A ไม่เข้าพวก จะถูกตัดทิ้งได้โดยง่าย ในบรรดายาฉีดจะเห็นว่ามียาต้านเชื้อราที่ไม่เข้าพวก (ตัวเลือก C) ดังนั้นจะเหลือตัวเลือกที่นักศึกษาต้องคิดเลือกจริง ๆ เพียงตัวเลือก B กับ D ซึ่งหากดูกลุ่มยาก็จะพบว่ายากลุ่ม Cephalosporin เข้าพวกมากที่สุด ทำให้ผู้สอบที่มีทักษะการทำข้อสอบดีสามารถเลือกคำตอบที่ถูกต้อง (ตัวเลือก B) ได้โดยไม่ต้องมีความรู้เรื่องการรักษาผู้ป่วย febrile neutropenia

๗. การบอกใบ้คำตอบโดยโจทย์ข้ออื่น

ข้อผิดพลาดนี้เป็นข้อผิดพลาดที่ตัวผู้เขียนข้อสอบไม่ค่อยรู้ แต่ผู้ที่จะตรวจพบข้อผิดพลาดนี้คืออาจารย์ผู้เลือกข้อสอบไปใช้ เนื่องจากในการสอบแต่ละครั้งใช้ข้อสอบจำนวนมาก หากเลือกข้อสอบโดยไม่ระมัดระวังอาจมีข้อสอบสองข้อที่ถามเกี่ยวกับโรคหรือกลุ่มอาการเดียวกัน ซึ่งข้อมูลจากโจทย์ในข้อหนึ่งอาจเป็นตัวบอกใบ้คำตอบของข้อสอบอีกข้อได้ 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 ซึ่งผู้สอบที่มีทักษะการทำข้อสอบดีสามารถย้อนกลับไปอ่านโจทย์ในข้อก่อนหน้านี้แล้วได้ข้อมูลว่าผู้ป่วยที่นำเสนอว่าเป็น febrile neutropenia มีไข้ขึ้น ๑๔ วันหลังได้ยาเคมีบำบัด ก็สามารถตอบข้อสอบข้อนี้ถูกได้โดยง่าย

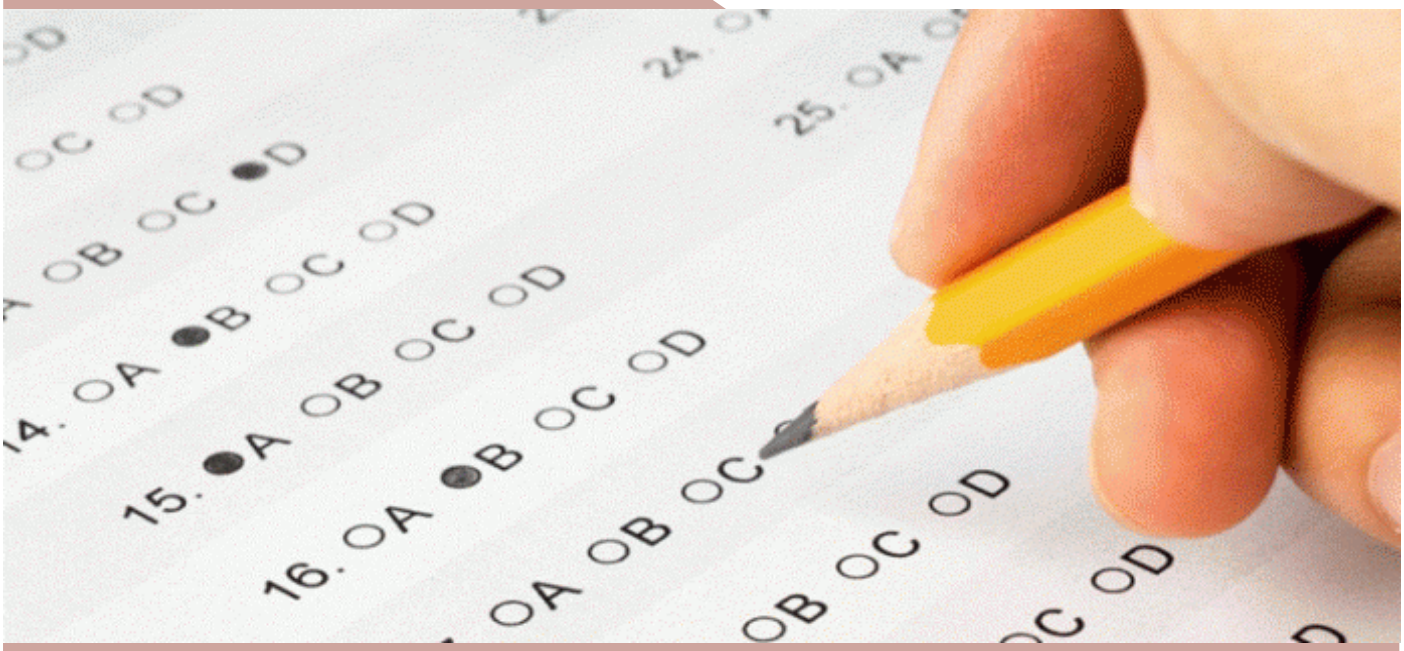
สรุป

ผู้นิพนธ์ได้รวบรวมข้อผิดพลาดในการสร้างข้อสอบปรนัยที่ผู้สอบอาจใช้เป็นแนวทางในการเลือกคำตอบที่ถูกได้โดยไม่ต้องอาศัยความรู้ทางการแพทย์ที่อาจารย์ต้องการประเมินผล โดยเรียบเรียงเป็นเจ็ดกลุ่มข้อผิดพลาดด้วยกัน ผู้อ่านทุกท่านพึงตระหนักว่าสิ่งเหล่านี้ไม่ใช่หลักการทางวิทยาศาสตร์ที่ชัดเจนดังกฎทางคณิตศาสตร์หรือฟิสิกส์ หากแต่เป็นการรวบรวมข้อสังเกตและคำแนะนำของผู้เชี่ยวชาญทางการวัดและประเมินผล จึงเป็นเพียงแนวทางเบื้องต้นในการพิจารณาตรวจสอบเนื้อหาของข้อสอบเท่านั้น การประยุกต์ใช้องค์ความรู้นี้คงต้องอาศัยศิลปะพอสมควรเพื่อที่จะได้ข้อสอบที่ดี สามารถวัดองค์ความรู้ทางการแพทย์ของนักศึกษาหรือแพทย์ประจำบ้านที่เข้าสอบได้ตามวัตถุประสงค์ของการสอบ

เอกสารอ้างอิง

1. Haladyna TM. Developing and validating multiple-choice test items, 3rd ed. Mahwah, NJ: Lawrence Erlbaum Assoicates, 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.

เอกสารประกอบการอบรม



23 June 2017

23 June 2017

หัวข้อ : ปัญหาในการสร้างข้อสอบปรนัย

ปัญหาในการสร้างข้อสอบปรนัย

รศ.นพ. เชิดศักดิ์ ไอรมณีรัตน์
ภาควิชาศัลยศาสตร์ คณะแพทยศาสตร์ศิริราชพยาบาล
มหาวิทยาลัยมหิดล

1

Common Problems

1. Objectives
2. Stem
3. Options

2

Objectives

- No clear objective
- Curriculum mapping
- Recall/Comprehension/Application

3

Stem

- Multiple ideas
- Confusing information
- Too much information
- Text vs photo
- Grammatical errors

4

Options

- Inappropriate cues
- Inadequate number of options
- Non-functioning options

5

Questions & Comments

Cherdsak.ira@mahidol.ac.th

6

23 June 2017

หัวข้อ : การผลิตข้อสอบและการจัดสอบ

การผลิตข้อสอบ

Test assembly and production

ผศ. พญ. อนัญญา พงษ์ไพบูลย์
คุณสุนันท์ มีเทศ
คุณทศพร มาสวัสดิ์

ประเด็นที่ต้องพิจารณา

1) Content coverage --- test specification table
- ไม่นำเนื้อหาเรื่องเดียวกันมาสอบในข้อสอบชุดเดียวกัน
- ไม่ให้เนื้อหาของข้อสอบข้อหนึ่งบอกคำตอบของข้อสอบอื่นในชุดเดียวกัน

Test _____

2) Formatting : maximize readability & minimize cognitive burden
- ตัวสะกด การเว้นวรรคตอน การขึ้นบรรทัดหรือหน้าใหม่
- รูปแบบหน้ากระดาษ การเรียงตัวเลือก การจัดรูปประกอบ

3) User interface

Test _____

รูปแบบการสอบ

1) Paper-based test
   - Single test form
   - Parallel test form
2) Computer-based test
   - CAT: Computerized adaptive test
   - CFT: Computerized fixed test
   - LOFT: Linear-on-the fly test

Test production

What are the key issues in the test production step?

1. Security
2. Content representation
3. Speed
4. Readability

A. 1,2 **B. 1,3**
C. 2,4 **D. 1,4** **E. 1,2,3**

Test production: _____

Human error is the most likely source of security breaches.

- Electronic transmission
- Printing of test books
- Secure shipping
- Secure destruction of excess printed materials

} Paper-based test

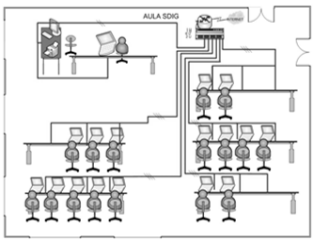
Test production: _____

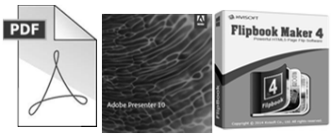
- Final review
  - Pre-printing test form
  - Final presented format *** (CBT)
- Random sampling of final printed test books

การสอบด้วย Computer-based test

Facility :

- Examination room
- Computer
- Software & network





ประสบการณ์การใช้งาน software สำเร็จรูปต่างๆ

- ไม่ตอบโจทย์ ความต้องการของผู้ใช้
- ไม่สามารถเพิ่ม features ต่างๆได้
- บาง software มีราคาแพง

จึงมีแนวคิด พัฒนา software ขึ้นใช้เอง

คุณลักษณะ software ที่ต้องการพัฒนา ...

- บันทึกคำตอบไว้ทั้งที่ เครื่องผู้สอบ และ server (ไฟฟ้าดับ – คำตอบยังคงอยู่)
- เริ่มและสิ้นสุดการสอบพร้อมกัน โดยการควบคุมผ่าน server

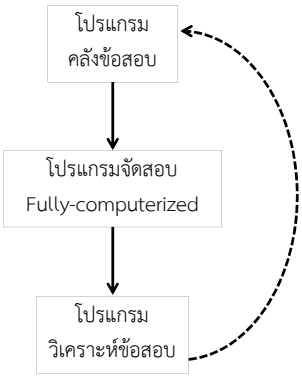
คุณลักษณะ software ที่ต้องการพัฒนา ...

User interface

- เลือกทำข้อสอบที่ต้องการ
- แสดงสถานะข้อสอบที่ทำ/ยังไม่ได้ทำ
- ใส่รูปภาพ/VDO
- ค้นหาคำ
- เปลี่ยนสีพื้นหลังและตัวอักษร, highlight คำ

คุณลักษณะ software ที่ต้องการพัฒนา ...

- สามารถเชื่อมโยงกับโปรแกรมอื่นๆ
  - โปรแกรมคลังข้อสอบ
  - โปรแกรมวิเคราะห์ข้อสอบ
- Import ข้อสอบจากคลังข้อสอบเพื่อมาจัดสอบ
- Export ผลการสอบไปยังโปรแกรมวิเคราะห์ข้อสอบ



การจัดสอบ

Test administration

The most public and visible aspect of testing

The most vulnerable for _____

Test administration

What are the key issues in the test administration?

1. Security
2. Speed
3. Standardization
4. Accuracy
5. Identification

A. 1,2,3 **B. 1,2,4**

C. 1,3,4 **D. 1,3,5** **E. 1,4,5**

Test administration: _____

- คุณภาพของชุดข้อสอบ
- สถานที่สอบ : โต๊ะเก้าอี้ แสงสว่าง อุณหภูมิ เสียงรบกวน
- ข้อปฏิบัติในการสอบ
  - เข้าห้องสอบ : เวลาสอบ อุปกรณ์ที่อนุญาตให้นำเข้าสอบ
  - ขณะสอบ
  - สอบเสร็จ
- การดำเนินการในกรณีพิเศษ : ป่วย หรือ ผู้พิการ

Test administration: _____

- **Low-stakes local examination**
  ผู้คุมสอบมักรู้จักผู้เข้าสอบเป็นอย่างดี
  - ตรวจสอบการเขียนชื่อ สกุล เลขที่ ให้ไม่สลับคน
- **High-stakes national examination**
  ตรวจสอบผู้เข้าสอบ(ที่ปรากฏตัว) ว่าเป็นบุคคลเดียวกันกับผู้มีสิทธิ์สอบ (ผู้สมัครสอบ)
  - ตรวจสอบบัตรประจำตัว ให้ตรงกับหน้าผู้เข้าสอบ
  - การลงลายมือชื่อทั้งที่ชุดข้อสอบและกระดาษคำตอบ
  - การประทับรอยนิ้วมือ

Test administration: _____

Test proctoring การคุมสอบ

- Protocol ระเบียบปฏิบัติในการคุมสอบที่ชัดเจน
- Training การอบรมบุคลากรที่เกี่ยวข้อง
- Organization การจัดระบบ แนะนำให้มี chief proctor ในแต่ละศูนย์สอบทำหน้าที่บริหารจัดการ

ประเด็นที่ต้องพิจารณา

- ตรวจสอบอุปกรณ์ที่ไม่อนุญาตให้นำเข้าห้องสอบ
- พฤติกรรมที่ต้องจับตามอง
- การอนุญาตให้เข้าห้องน้ำ
- การเข้าสอบช้า หรือการออกจากห้องสอบก่อนหมดเวลา
- การเริ่มทำข้อสอบก่อน หรือ การไม่หยุดทำข้อสอบเมื่อหมดเวลา
- การจดหรือบันทึกเนื้อหาข้อสอบออกจากห้องสอบ

การจัดสอบด้วยระบบ computerized

ประเด็นที่ต้องพิจารณา

- ความพร้อมของอุปกรณ์ hardware, software และ network

23 June 2017

หัวข้อ : การตรวจและวิเคราะห์ข้อสอบปรนัย

การตรวจและวิเคราะห์ข้อสอบ

สุนันท์ มีเทศ

ขั้นตอนก่อนที่จะตรวจกระดาษคำตอบ

1. ประธานรายวิชาส่งเฉลยข้อสอบให้ฝ่ายการศึกษา
2. ก่อนนำกระดาษคำตอบเข้าเครื่องตรวจ เจ้าหน้าที่จะทำการตรวจกระดาษคำตอบว่าเป็นรหัสเดียวกันหรือไม่ และไม่มีสิ่งแปลกปลอมสอดแทรกอยู่ในกระดาษคำตอบ เช่น เศษยางลบ เศษกระดาษ เศษสติกเกอร์ ฯลฯ เพราะถ้ามีสิ่งพวกนี้ติดกับกระดาษคำตอบจะทำให้ติดและไม่สามารถตรวจต่อไปได้

ขั้นตอนการตรวจ

มีขั้นตอนต่างๆ ดังนี้

1. เปิดโปรแกรม Scantools Plus ซึ่งเป็นโปรแกรมกระดาษคำตอบ
2. เลื่อนเมาส์ ไปคลิกที่ menu New จากนั้นตั้งชื่อวิชาที่ทำการตรวจกระดาษคำตอบ

ขั้นตอนการตรวจกระดาษคำตอบ (ต่อ)

3. ในการตรวจกระดาษคำตอบจะใช้เครื่อง Opscan8 ในการตรวจกระดาษคำตอบ จะวางเฉลยที่จะใช้ตรวจก่อนและตามด้วยกระดาษคำตอบของนักศึกษา จะทำการ scan กระดาษคำตอบไปเรื่อยๆ จนกว่ากระดาษคำตอบจะหมด ตรวจกระดาษคำตอบอย่างน้อย 2 ครั้ง จากเครื่องตรวจกระดาษคำตอบ 2 เครื่อง

ขั้นตอนการตรวจกระดาษคำตอบ (ต่อ)

4. จะนำผลที่ได้จากการ scan กระดาษคำตอบมาตรวจสอบด้วยโปรแกรม excel ว่าข้อมูลตรงกันหรือไม่
5. เมื่อตรวจสอบข้อมูลเรียบร้อย นำข้อมูลที่ได้เข้าโปรแกรมวิเคราะห์ข้อสอบ

ขั้นตอนการวิเคราะห์ข้อสอบ

มีขั้นตอนต่างๆ ดังนี้

1. เข้าเมนูข้อมูลการจัดสอบ เพื่อใส่ข้อมูลในโปรแกรมวิเคราะห์ข้อสอบให้ครบถ้วน
2. เข้าเมนูเก็บข้อมูลเพื่อทำการประมวลผล
3. เข้าเมนูเรียกดูวิเคราะห์ข้อสอบ เพื่อดูผลคะแนนและผลการวิเคราะห์ข้อสอบ

MCQ Item Analysis

Cherdsak Iramaneerat
Department of Surgery
Faculty of Medicine Siriraj Hospital
Mahidol University

Item Analysis

- A group of statistical analyses having two characteristics:
  - The data consist of actual responses of test takers to individual test items
  - The primary purpose is to gain information about the items (rather than about test takers)

Livingston SA. Item analysis. In: Downing SM, Haladyna TM. Handbook of test development. Mahwah, NJ: LEA, 2006, p. 421-444.

Outline

- Item statistics
- Test statistics
- Applications
- Limitations

Two Parts of Item Analysis

- Item statistics
  - Item difficulty
  - Item discrimination
  - Distractor functionality
- Test statistics
  - Internal consistency reliability
  - Standard deviation and mean
  - Average difficulty
  - Average discrimination

Item Difficulty

- Proportion of examinees answering an item correctly (p)

C = number of examinees with a correct answer

I = number of examinees with incorrect answers

- Ideal: 0.45 – 0.75
- Good: 0.76 – 0.91
- Acceptable: 0.25 – 0.44
- Problematic: < 0.24 or > 0.91

Item Discrimination

- The ability of an item to discriminate high scorers from low scorers
- Point-biserial correlation (r)

Mp	=	Mean score of examinees with a correct answer
Mq	=	Mean score of examinees with incorrect answers
SD	=	Standard deviation of test scores
p	=	Proportion of examinees with a correct answer
q	=	Proportion of examinees with incorrect answers

Point-Biserial Correlation

- The correlation between an item score with the total score
  - **Range: -1.0 – 1.0**
  - **Point-biserial of an item should be positive**
    - Ideal: 0.20 or higher
    - Acceptable: 0.1 – 0.19
    - Problematic: < 0

Distractor Functionality

A functioning distractor is an incorrect option that:

1. Is chosen by at least 5 percent of examinees
2. Has a negative point-biserial correlation with the total score

11

Example 1

Number 148		Correct answer = 2			
P-VALUE = 0.65	PT BISERIAL =0.1				Total number of examinees
DISTRACTOR	1	2	3	4	5
N OF PEOPLE	4	158	17	58	5	242
MEAN SCORE	77.25	84.81	81.35	83.86	76.6	
P-VALUE	0.02	0.65	0.07	0.24	0.02	
PT BISERIAL	-0.09	0.1	-0.07	-0.01	-0.11	

12

Siriraj Hospital's IA report

No. : 1		p Value : 0.64		r_{pbi} : 0.23					
A		B		C		* D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
0.02	6.98	-0.18	5.08	-0.17	8.57	0.23	63.81	-0.07	15.56

MCQ item analysis

Item Analysis and Option Analysis
Faculty of Medicine Siriraj Hospital
Mahidol University

No. : 1 p Value : 0.64 r_{pbi} : 0.23

A		B		C		* D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
0.02	6.98	-0.18	5.08	-0.17	8.57	0.23	63.81	-0.07	15.56

No. : 2 p Value : 0.34 r_{pbi} : 0.19

A		B		C		D		* E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
0.01	4.76	-0.02	25.40	-0.19	10.76	-0.06	34.76	0.19	33.97

No. : 3 p Value : 0.56 r_{pbi} : 0.35

A		B		* C		D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
-0.03	8.89	-0.26	23.17	0.35	55.87	-0.05	3.17	-0.16	8.89

No. : 4 p Value : 0.50 r_{pbi} : 0.33

A		* B		C		D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
-0.15	1.90	0.33	50.48	-0.15	4.13	-0.18	10.48	-0.13	33.02

No. : 5 p Value : 0.24 r_{pbi} : 0.06

A		B		C		* D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
0.06	3.49	-0.08	53.02	0.05	12.08	0.04	23.81	0.02	7.32

No. : 6 p Value : 0.53 r_{pbi} : 0.20

A		B		* C		D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
0.18	23.17	-0.11	3.81	0.20	53.33	0.02	5.40	0.02	14.29

MCQ item analysis

Reliability

- Consistency of test scores
  - If we test the students again, will they get the same scores?
  - Range: 0 – 1
  - High values: highly consistent test scores

20

Internal Consistency Reliability

- Consistency of test scores: If we test the students again, will they get the same scores?
- In MCQ exam, one commonly reported index of reliability is Cronbach's Alpha

$$\alpha = \frac{n}{n-1}\left(1 - \frac{\sum_{i=1}^{n} \sigma_i^2}{\sigma_t^2}\right)$$

n = number of testlets
σ_t^2 = score variance of total scores
σ_i^2 = score variance of the i^{th} testlet

How Much is Enough?

- Depends on test scores uses
  - –High-stakes exam: 0.9 or higher
  - –Medium-stakes exam: 0.80 – 0.89
  - –Low-stakes exam: 0.70 – 0.79

22

Improving Reliability

- Increase the number of test items
- Adjust item difficulty to obtain larger spread of test scores
- Adjust testing conditions to eliminate interruptions, noise, and other disrupting factors
- Eliminate subjectivity in scoring

23

Mean and Standard Deviation

- Effective instruction => All students can do the test well.
  - —High mean scores
  - —Low standard deviation
- High standard deviation: Wide range of students' scores
  - —Some students can solve the problems in the tests, while some students cannot do.
- Too difficult test => Most students fail to get correct answers.
  - —Low mean scores
  - —Low standard deviation

Average Difficulty

- Average of p values of all items on the test
- Small group of students:
  - —Difficult to interpret
  - —Depends on the ability distribution of students
- Large group of students:
  - —Assume a fair sampling of students
  - —Indicates the average difficulty of the whole test

Average Discrimination

- Average point-biserial correlation of the whole test
- Indicates how good the items on the test can differentiate high scorers from low scorers.
- High values generally indicate a good test.
- Effective instruction: All students can do well on the test.
  - —A low value does not necessarily indicate bad items.

Applications

1. Posttest score adjustment
2. Item revision
3. Item pool management
4. Improvement of instruction

Limitations

1. Sample dependency
2. Reliability is the property of test scores, not test items.
3. Numbers are there to serve us, not the other way around.

Questions & Comments

Cherdsaklramaneerat@gmail.com

"We all need people who will give us feedback. That's how we improve."

Bill Gates

MCQ item analysis

การวิเคราะห์ข้อสอบปรนัย

อาจารย์ นายแพทย์เชิดศักดิ์ ไอรมณีรัตน์
ภาควิชาศัลยศาสตร์, คณะแพทยศาสตร์ศิริราชพยาบาล, มหาวิทยาลัยมหิดล, กรุงเทพมหานคร ๑๐๗๐๐.

การวิเคราะห์ข้อสอบปรนัย (Item analysis) เป็นการใช้วิธีการทางสถิติเพื่อวิเคราะห์คำตอบที่ผู้สอบตอบข้อสอบปรนัยในการสอบครั้งหนึ่ง เพื่อประเมินว่าข้อสอบที่นำมาใช้ในการสอบครั้งนั้นมีคุณสมบัติอย่างไร ทำงานได้ตามที่ต้องการหรือไม่ มีระดับความยากง่ายของข้อสอบเหมาะสมหรือไม่ มีข้อบกพร่องหรือไม่ และควรได้รับการปรับปรุงแก้ไขอย่างไร[๑] การวิเคราะห์ข้อสอบเป็นศาสตร์ที่ได้รับการพัฒนาอย่างต่อเนื่องมาเป็นเวลานาน มีเทคนิคและวิธีการต่าง ๆ มากมายที่ผู้วิเคราะห์สามารถใช้เพื่อบอกคุณสมบัติของข้อสอบแต่ละข้อ ตั้งแต่วิธีการง่าย ๆ ไปจนถึงวิธีการที่มีความซับซ้อนมาก โดยแต่ละเทคนิคการวิเคราะห์ก็มีจุดประสงค์แตกต่างกันไป ตั้งแต่การบอกระดับความยากง่าย การบอกถึงความสามารถในการแยกผู้สอบที่เก่งออกจากผู้สอบที่ไม่เก่ง ไปจนถึงเทคนิคขั้นสูงที่สามารถบอกได้ว่าข้อสอบมีความลำเอียงต่อผู้สอบเพศใดเพศหนึ่ง หรือผู้สอบจากสถาบันใดสถาบันหนึ่งเป็นพิเศษหรือไม่ มีการเดาข้อสอบมากน้อยเพียงใด ผู้สอบรู้ข้อสอบมาก่อนเข้าสอบหรือไม่ หรือมีความน่าจะเป็นมากน้อยเพียงใดที่ผู้สอบลอกคำตอบ ในบทความนี้ผู้นิพนธ์ไม่ได้ตั้งเป้าประสงค์ที่จะรวบรวมและอภิปรายเทคนิคการวิเคราะห์ข้อสอบทุกวิธีที่มีใช้อยู่ในปัจจุบัน แต่ต้องการเพียงนำเสนอความรู้พื้นฐานที่เกี่ยวกับการวิเคราะห์ข้อสอบและอธิบายถึงวิธีการวิเคราะห์ข้อสอบที่นิยมใช้กันในทางแพทยศาสตรศึกษา โดยเฉพาะในประเทศไทย โดยประสงค์ให้อาจารย์ผู้อ่านสามารถนำเอาความรู้ที่ได้จากบทความนี้ไปใช้แปลผลการวิเคราะห์ข้อสอบที่ตนเกี่ยวข้อง และดำเนินการปรับปรุงคุณภาพของข้อสอบได้อย่างเหมาะสม

ความรู้พื้นฐานเกี่ยวกับข้อสอบปรนัย

ก่อนที่จะกล่าวถึงรายละเอียดในการวิเคราะห์ข้อสอบ ผู้นิพนธ์ก็จะขอทบทวนความรู้พื้นฐานเกี่ยวกับข้อสอบปรนัยก่อน โดยทั่วไปข้อสอบปรนัยแต่ละข้อมีส่วนประกอบสำคัญ ๒ ส่วนด้วยกันคือ

๑. โจทย์ (stem) เป็นข้อมูลของโรค หรือภาวะหรือผู้ป่วยตามด้วยคำถาม หรือเว้นช่องว่างสำหรับเติมคำหรือข้อความที่เหมาะสมลงไป

๒. ตัวเลือก (options) คือคำ หรือข้อความที่ผู้ออกข้อสอบนำเสนอตามหลังจากโจทย์เพื่อให้ผู้สอบเลือกไปใช้ตอบคำถาม หรือเติมลงในช่องว่างในโจทย์

๒.๑ ตัวเลือกที่ถูกต้อง (correct option) เป็นคำตอบที่ถูกต้องมีเพียงตัวเลือกเดียวต่อข้อสอบข้อหนึ่ง

๒.๒ ตัวลวง (distractors) เป็นคำตอบที่ผิด มีไว้ลวงให้ผู้สอบที่ไม่มีความรู้ หรือมีความเข้าใจไม่ถูกต้องในเนื้อหาที่นำมาออกข้อสอบเลือกตอบ ข้อสอบที่ใช้ในคณะแพทยศาสตร์ศิริราชพยาบาล และที่ใช้ทั่วไปในการสอบของนักศึกษาแพทย์ และแพทย์ประจำบ้านในประเทศไทย นิยมจัดให้มีตัวลวง ๔ ตัวต่อข้อสอบ ๑ ข้อ

ทฤษฎีที่ใช้ในการวิเคราะห์ข้อสอบ

ทฤษฎีที่ใช้ในการวิเคราะห์ข้อสอบในปัจจุบันนั้นมี ๒ ทฤษฎีด้วยกัน ได้แก่ทฤษฎีการสอบแบบดั้งเดิม

(classical test theory) และทฤษฎีการตอบสนองต่อข้อสอบ (item response theory) ทฤษฎีการสอบแบบดั้งเดิมนั้นเป็นทฤษฎีที่ได้ถูกพัฒนาขึ้นตั้งแต่ตอนต้นของศตวรรษที่ ๒๐ โดยมีการรวบรวมเป็นตำราในครั้งแรกตั้งแต่ปี ค.ศ. ๑๙๒๑ โดย William Brown และ Godfrey H Thomson[๒] หลังจากนั้นทฤษฎีนี้ก็ได้รับการใช้อย่างแพร่หลายในการวิเคราะห์ข้อสอบและได้รับการพัฒนาอย่างต่อเนื่อง ทฤษฎีการสอบแบบดั้งเดิมนี้วางรากฐานอยู่บนสมมติฐานว่าคะแนนสอบที่ได้มานั้นประกอบไปด้วยคะแนนที่แท้จริง (true score) กับความผิดพลาดจากการวัด (error) ซึ่งสมมติฐานดังกล่าวต่อมาพบว่ามีข้อจำกัดหลายประการด้วยกัน ในราว ค.ศ. ๑๙๗๐ จึงได้มีความพยายามพัฒนาทฤษฎีที่ใช้ในการวิเคราะห์ข้อสอบแบบใหม่ขึ้นซึ่งใช้หลักการของความน่าจะเป็นมาวิเคราะห์ข้อสอบ ทำให้สามารถแยกผลการวิเคราะห์ข้อสอบแต่ละข้อเป็นอิสระจากข้อสอบข้ออื่นในการสอบเดียวกัน ทฤษฎีใหม่นี้เรียกว่าทฤษฎีการตอบสนองต่อข้อสอบ (item response theory) ทฤษฎีใหม่นี้มีข้อได้เปรียบกว่าทฤษฎีเดิมหลายประการด้วยกัน ได้แก่ ความสามารถในการปรับตัวเข้ากับสถานการณ์ต่าง ๆ (flexibility) ความมีประสิทธิภาพในการใช้ข้อมูล (efficiency) และความสามารถในการวิเคราะห์ถึงคุณภาพของข้อสอบ และผู้สอบโดยละเอียด (in-depth analysis)[๓] จึงเป็นเหตุให้ทฤษฎีการตอบสนองต่อข้อสอบนี้ได้รับความนิยมอย่างกว้างขวางตั้งแต่ในค.ศ. ๑๙๘๐ ในปัจจุบันการสอบต่าง ๆ ได้ถูกวิเคราะห์ด้วยทฤษฎีการตอบสนองต่อข้อสอบนี้มากขึ้นเรื่อย ๆ

เนื่องจากการวิเคราะห์ข้อสอบในวงการแพทยศาสตรศึกษาในประเทศไทยทั้งหมดในปัจจุบันยังใช้เทคนิคต่าง ๆ ตามทฤษฎีการสอบแบบดั้งเดิมอยู่ ดังนั้นผู้นิพนธ์จะขอกล่าวถึงเทคนิคการวิเคราะห์ข้อสอบตามทฤษฎีการสอบแบบดั้งเดิมเท่านั้น เพราะจะเป็นสิ่งที่อาจารย์แพทย์ทุกท่านจะได้พบและใช้งานเป็นประจำ

การวิเคราะห์ข้อสอบตามทฤษฎีการสอบแบบดั้งเดิม

การวิเคราะห์ข้อสอบตามทฤษฎีการสอบแบบดั้งเดิมนี้ประกอบไปด้วย ๒ ส่วนใหญ่ ๆ คือ (๑) การวิเคราะห์ข้อสอบรายข้อ (item analysis) และ (๒) การวิเคราะห์ข้อสอบโดยรวม (test analysis)

๑. การวิเคราะห์ข้อสอบรายข้อ (item analysis)

การวิเคราะห์ข้อสอบแต่ละข้อให้อาจารย์พิจารณา ๓ ปัจจัย คือ

๑.๑ ความยากง่ายของข้อสอบ (item difficulty, p)

ความยากง่ายของข้อสอบวัดโดยใช้ค่า p ซึ่งย่อมาจาก proportion of examinees answering items correctly (สัดส่วนของผู้สอบที่ตอบข้อสอบข้อนั้นถูก) ซึ่งหาได้จากการนำจำนวนผู้สอบที่ตอบข้อสอบข้อนั้นถูกต้องหารด้วยจำนวนผู้สอบที่ตอบข้อสอบข้อนั้นทั้งหมด หากข้อสอบข้อนั้นเป็นข้อสอบที่ง่ายผู้สอบทุกคนตอบถูก ค่า p ก็จะเป็น ๑ หากไม่มีผู้สอบคนใดตอบถูกเลยข้อสอบข้อนั้นก็จะมีค่า p เป็น ๐ หากมีคนตอบถูก ๗๐% ข้อสอบข้อนั้นก็มีค่า p เท่ากับ ๐.๗ ข้อสอบที่ดีมากจะมีค่า p อยู่ในช่วง ๐.๔๕ – ๐.๗๕, ข้อสอบที่ดีจะมีค่า p อยู่ในช่วง ๐.๗๖ – ๐.๙๑, ข้อสอบที่พอใช้ได้มีค่า p อยู่ในช่วง ๐.๒๕ – ๐.๔๔, ข้อสอบที่มีค่า p ต่ำกว่า ๐.๒๕ เป็นข้อสอบที่ยากเกินไป และข้อสอบที่มีค่า p สูงกว่า ๐.๙๑ เป็นข้อสอบที่ง่ายเกินไป[๔-๖]

๑.๒ ความสามารถในการจำแนกผู้สอบตามระดับความสามารถ (item discrimination, r)

ความสามารถในการจำแนกผู้สอบ หมายถึง ความสามารถของข้อสอบข้อหนึ่ง ๆ ในการแยกผู้สอบที่ทำคะแนนได้ดี ออกจากผู้สอบที่ทำคะแนนได้ไม่ดี ข้อสอบที่มีความสามารถในการแยกแยะได้ดีนั้นผู้สอบที่ตอบข้อสอบข้อนั้นถูกมักจะได้คะแนนสูง และผู้สอบที่ตอบข้อสอบข้อนั้นผิดมักจะได้คะแนนต่ำ ดัชนีที่ใช้วัดความสามารถในการจำแนกผู้สอบที่ใช้กันมากที่สุดในปัจจุบันคือค่า point-biserial correlation ซึ่งนิยมใช้อักษรย่อเป็น r_{pbis}[๗] ซึ่งสามารถคำนวณได้จากสูตรต่อไปนี้[๗]

$$r = \frac{M_p - M_q}{SD}\sqrt{pq}$$

เมื่อ Mp = คะแนนรวมเฉลี่ยของผู้สอบที่ตอบข้อสอบถูก

Mq = คะแนนรวมเฉลี่ยของผู้สอบที่ตอบข้อสอบผิด

SD = ค่าเบี่ยงเบนมาตรฐาน (standard deviation) ของคะแนนสอบ

p = สัดส่วนของผู้สอบที่ตอบข้อสอบถูกต่อผู้สอบทั้งหมด

q = สัดส่วนของผู้สอบที่ตอบข้อสอบผิดต่อผู้สอบทั้งหมด

ค่า point-biserial correlation ที่คำนวณได้นี้มีค่าอยู่ในช่วง -๑ ถึง ๑ โดยค่าที่ติดลบหมายถึง ข้อสอบข้อนั้นผู้ที่ตอบถูกมักสอบได้คะแนนรวมต่ำ แต่ผู้ที่ตอบผิดมักสอบได้คะแนนรวมสูง ในทางตรงข้าม หากค่า point-biserial ยิ่งสูง แสดงถึงข้อสอบที่มีความสามารถในการแยกแยะดี ผู้ที่ตอบข้อสอบข้อนั้นถูกมักทำคะแนนรวมได้สูง ข้อสอบที่ดีควรมีค่า point-biserial สูงกว่า ๐.๒๐, ข้อสอบที่พอใช้ได้ควรมีค่า point-biserial อยู่ในช่วง ๐.๑ – ๐.๑๙, ข้อสอบที่มีค่า point-biserial ต่ำกว่า ๐.๑ เป็นข้อสอบที่ไม่สู้ดีนัก โดยเฉพาะอย่างยิ่งข้อสอบที่มีค่า point-biserial ต่ำกว่า ๐ ไม่ควรนำมาคิดคะแนน[๔-๖] (โดยทั่วไปแล้วข้อสอบที่มีค่า point-biserial ติดลบ ให้สงสัยว่าจะเฉลยผิด)

๑.๓ ประสิทธิภาพของตัวลวง (distractor functionality)

ตัวลวงที่มีประสิทธิภาพนั้นมีคุณสมบัติ ๒ ประการคือ[๓]

(๑) มีผู้สอบเลือกตัวลวงนั้นไม่ต่ำกว่าร้อยละ ๕ ของจำนวนผู้สอบทั้งหมด

(๒) มีค่า point-biserial correlation ของตัวลวงนั้นเป็นลบ กล่าวคือตัวลวงที่ดีจะลวงให้ผู้สอบที่มีความรู้ไม่ดี (มีคะแนนต่ำ) มาเลือก แต่ไม่ลวงให้ผู้สอบที่มีความรู้ดี (มีคะแนนสูง) มาเลือก หากตัวลวงใดมีค่า point-biserial correlation เป็นบวก ให้ทบทวนข้อสอบข้อนั้นดูว่าอาจจะเฉลยผิดหรือมีคำตอบที่ถูกต้องมากกว่า ๑ ตัวเลือก

ตัวลวงใดที่มีผู้สอบเลือกน้อย หรือลวงให้ผู้ที่มีความรู้ดีมาเลือกจัดเป็นตัวลวงที่ไม่ดี สมควรพิจารณาตัดทิ้งหรือปรับเปลี่ยน

๒. การวิเคราะห์ข้อสอบโดยรวม (test analysis)

การวิเคราะห์ข้อสอบโดยรวมเป็นการพิจารณาว่าเมื่อข้อสอบทั้งชุดทำงานร่วมกันแล้วผลสอบที่ได้ออกมาเป็นอย่างไร มีระดับความยากง่ายเป็นอย่างไร มีการกระจายตัวของคะแนนเป็นอย่างไร มีความน่าเชื่อถือของคะแนนสอบมากน้อยเพียงใด ดัชนีต่าง ๆ ที่ต้องพิจารณาได้แก่

๒.๑ ความเที่ยงตรงของคะแนนสอบ (internal consistency reliability)

การประเมินความเที่ยงตรงของคะแนนสอบเป็นการตรวจสอบว่าคะแนนที่ได้ออกมานั้นมีความน่าเชื่อถือเพียงใด เป็นการตอบคำถามว่าหากนำผู้สอบมาสอบใหม่ในสภาวการณ์เดิม ด้วยข้อสอบที่มีระดับความยากง่ายเท่าเดิม และผู้สอบมีความรู้เท่าเดิมไม่ได้ไปศึกษาหาความรู้เพิ่มเติม จะได้คะแนนสอบเท่าเดิมหรือไม่[๙-๑๑]

ดัชนีชี้วัดความเที่ยงตรงของคะแนนสอบที่นิยมใช้ในการรายงานผลสอบด้วยข้อสอบปรนัยคือ ค่าสัมประสิทธิ์ อัลฟ่า (Coefficient Alpha) ซึ่งสามารถคำนวณได้จากสูตร[๑๑]

$$\alpha = \frac{n}{n\text{-}1}\left(1\text{-}\frac{\sum \sigma^2_{\chi_i}}{\sigma^2_{\chi}}\right)$$

เมื่อ α = สัมประสิทธิ์ อัลฟ่า (Coefficient Alpha)

n = จำนวนชุดย่อยของข้อสอบที่ทำการแบ่งออกเพื่อหาความเที่ยง

σ^2_{χ} = การกระจายตัว (variance) ของคะแนนรวม

$\sigma^2_{\chi_i}$ = การกระจายตัว (variance) ของคะแนนข้อสอบย่อยชุดที่ i

ค่าสัมประสิทธิ์อัลฟ่านี้มีค่าอยู่ในช่วง ๐ – ๑ ค่าต่ำแสดงว่าคะแนนที่ได้มีความเชื่อถือได้น้อย ไม่แตกต่างไปจากการเดาสุ่ม ค่าสูงแสดงว่าคะแนนที่ได้นั้นมีความน่าเชื่อถือมาก หากทำการทดสอบซ้ำคะแนนที่ได้ก็จะใกล้เคียงเดิม โดยทั่วไประดับของความเที่ยงตรง

ของคะแนนสอบที่ยอมรับได้นั้นขึ้นกับว่าต้องการนำเอาคะแนนสอบไปใช้ทำอะไร หากการตัดสินผลสอบนั้นมีความสำคัญมาก (high-stakes examination) เช่น การตัดสินผลสอบขอรับใบประกอบวิชาชีพเวชกรรม หรือประกาศนียบัตรแพทย์ผู้เชี่ยวชาญเฉพาะสาขา มักต้องการคะแนนสอบที่มีค่าสัมประสิทธิ์ อัลฟ่า ไม่ต่ำกว่า ๐.๙ หากการตัดสินผลสอบนั้นมีความสำคัญปานกลาง (medium-stakes examination) เช่นการสอบลงกองการสอบเลื่อนชั้นเรียน มักต้องการคะแนนสอบที่มีค่าสัมประสิทธิ์ อัลฟ่า อยู่ในช่วง ๐.๘ – ๐.๘๙ หากการตัดสินผลสอบนั้นมีความสำคัญน้อย (low-stakes examination) เช่นการสอบย่อยในชั้นเรียน การสอบแบบ formative assessment มักต้องการคะแนนสอบที่มีค่าสัมประสิทธิ์ อัลฟ่า อยู่ในช่วง ๐.๗ - ๐.๗๙[๑๒]

ประเด็นสำคัญที่ต้องพิจารณาคือเมื่อได้คะแนนสอบที่มีค่าสัมประสิทธิ์ อัลฟ่าต่ำ จะต้องดำเนินการอย่างไรเพื่อพัฒนาให้การสอบครั้งต่อไปไม่ประสบปัญหาเรื่องความไม่น่าเชื่อถือของคะแนนสอบอีก ปัจจัยหลักที่จะช่วยเพิ่มความเที่ยงตรงของคะแนนสอบปรนัยมี ๓ ปัจจัยด้วยกัน[๙] คือ

(๑) เพิ่มจำนวนข้อสอบให้มากขึ้น ยิ่งมีข้อสอบมากข้อคะแนนที่ได้ก็จะมีความเที่ยงตรงเพิ่มมากขึ้น

(๒) ปรับให้ข้อสอบมีการคละกันของข้อสอบที่ยากและง่ายอย่างเหมาะสม เพื่อปรับให้คะแนนมีการกระจายตัวมากขึ้น หากข้อสอบทั้งชุดประกอบไปด้วยข้อสอบที่ง่ายหมด ผู้สอบเกือบทั้งหมดได้คะแนนสูงมาก จะทำให้มีความแตกต่างของคะแนนน้อย โอกาสที่จะแยกแยะผู้สอบที่มีความรู้ดีออกจากผู้ที่มีความรู้ปานกลาง หรือไม่สู้ดีได้อย่างมั่นใจก็เป็นไปได้น้อย ดังนั้นหากอาจารย์ปรับให้มีการคละกันของข้อสอบยากและง่ายอย่างเหมาะสม ก็จะทำให้ผู้สอบมีระดับคะแนนแตกต่างกันมาก ค่าสัมประสิทธิ์อัลฟ่าก็จะสูงขึ้นด้วย

(๓) ปรับสภาวะแวดล้อมของการสอบให้เหมาะสม กำจัดสิ่งรบกวนสมาธิของผู้สอบให้มากที่สุด เช่น เสียงรบกวน แสงไฟที่ไม่เพียงพอ หรือไฟที่ติด ๆ ดับ ๆ เป็นต้น

๒.๒ การกระจายตัวของคะแนน และคะแนนเฉลี่ย (standard deviation and mean score)

การตรวจดูลักษณะพื้นฐานของคะแนนสอบนี้จะช่วยบอกได้คร่าว ๆ ว่าการเรียนการสอนมีประสิทธิภาพเพียงใด หากอาจารย์สอนได้ดี นักเรียนทั้งชั้นเรียนเข้าใจเนื้อหาดี คะแนนสอบที่ได้ออกมาก็ควรจะไม่กระจายตัวมากนัก (คะแนนเกาะกลุ่มกัน) และคะแนนเฉลี่ยก็ควรจะค่อนข้างสูงเมื่อเทียบกับนักเรียนรุ่นอื่น ๆ หากคะแนนสอบของนักเรียนมีการกระจายตัวมากผิดปกติ แสดงว่าอาจมีปัญหาบางประการในการเรียนการสอนทำให้นักเรียนบางคนมีความรู้ความเข้าใจดี แต่มีนักเรียนบางกลุ่มที่ไม่ค่อยรู้เรื่อง[๙]

๒.๓ ค่าความยากง่ายเฉลี่ยของข้อสอบ (average difficulty)

จากการวิเคราะห์ข้อสอบรายข้อ เราได้ค่าความยากง่ายของข้อสอบแต่ละข้อ (p) เมื่อนำค่า p ของข้อสอบทุกข้อมาหาค่าเฉลี่ย เราก็จะได้ค่าความยากง่ายของข้อสอบทั้งชุด ค่าที่ได้มานี้ใช้เป็นดัชนีชี้วัดว่าข้อสอบทั้งชุดโดยรวมแล้วมีระดับความยากง่ายเป็นอย่างไร หากผู้สอบเป็นนักศึกษากลุ่มใหญ่พอที่เราจะตั้งสมมติฐานว่าระดับความสามารถมีการกระจายตัวอย่างเหมาะสม และไม่ต่างจากระดับความสามารถเฉลี่ยของกลุ่มผู้สอบปีก่อน ๆ เราก็สามารถนำค่าความยากง่ายของข้อสอบทั้งชุดนี้มาเทียบได้ว่าข้อสอบที่นำมาใช้ในปีนี้ยาก หรือง่ายกว่าข้อสอบปีก่อน ๆ ซึ่งอาจารย์อาจนำข้อมูลนี้มาใช้พิจารณาปรับเกณฑ์การตัดเกรดด้วยว่าต้องมีการปรับระดับคะแนนที่ได้เกรดต่าง ๆ หรือไม่ อย่างไร

๒.๔ ค่าความสามารถในการแยกแยะผู้สอบเฉลี่ย (average discrimination)

การนำค่า point-biserial correlation ของข้อสอบทั้งชุดมาหาค่าเฉลี่ย เป็นการบอกคร่าว ๆ ว่าโดยรวมแล้วข้อสอบชุดนี้มีความสามารถในการแยกแยะผู้สอบตามระดับความสามารถเพียงใด ยิ่งได้ค่าสูงก็ยิ่งดี แต่มีข้อควรระวังในการแปลผลในกรณีที่การเรียนการสอนเป็นไปได้ดี และผู้สอบทั้งหมด หรือเกือบทั้งหมดทำคะแนนได้สูง ค่า point-biserial correlation เฉลี่ยของข้อสอบทั้งชุดจะไม่สูงแต่ไม่ได้แปลว่าข้อสอบที่ใช้มีคุณภาพไม่ดี[๙]

การนำผลการวิเคราะห์ข้อสอบไปใช้

ผลการวิเคราะห์ข้อสอบด้วยดัชนีชี้วัดต่าง ๆ ดังกล่าวข้างต้นสามารถนำไปใช้ประโยชน์ได้หลายประการ เช่น

๑. ใช้เป็นประโยชน์ในการปรับแก้คะแนนสอบ

จากผลการวิเคราะห์ข้อสอบจะช่วยชี้แนะให้เราทราบว่าข้อสอบข้อใดน่าจะเฉลยผิด ข้อสอบข้อใดน่าจะมีคำตอบที่ถูกมากกว่า ๑ ตัวเลือก ข้อสอบข้อใดน่าจะมีปัญหาเช่น มีความคลุมเครือในคำถาม หรือตัวเลือกมีความซ้ำซ้อนกัน หรือเนื้อหาของข้อสอบอยู่นอกเหนือไปจากสิ่งที่สอนนักเรียน เป็นต้น ข้อสอบที่มีปัญหาเหล่านี้ต้องได้รับการประเมินโดยคณะกรรมการตรวจข้อสอบซึ่งประกอบไปด้วยอาจารย์ผู้มีความรู้ความชำนาญในเนื้อหาวิชาที่ทำการสอบว่าจะดำเนินการอย่างไรกับการคิดคะแนน หากปัญหาที่พบมีความรุนแรงไม่มากจนทำให้การตัดสินใจเลือกคำตอบที่ถูกต้องเปลี่ยนไป คณะกรรมการอาจพิจารณาคิดคะแนนของข้อสอบข้อนั้นตามปกติ หากข้อสอบเฉลยผิดคณะกรรมการสามารถพิจารณาแก้คำตอบแล้วทำการตรวจให้คะแนนข้อสอบข้อนั้นใหม่ หากข้อสอบข้อใดมีคำตอบที่เหมาะสม ๒ ข้อ คณะกรรมการอาจพิจารณาให้ผู้สอบที่ตอบข้อใดข้อหนึ่งใน ๒ ข้อดังกล่าวได้คะแนนในข้อนั้น หากข้อสอบนั้นมีความคลุมเครือมากจนไม่สามารถตัดสินใจเลือกคำตอบที่เหมาะสมได้ คณะกรรมการสามารถตัดข้อสอบข้อนั้นออกจากการคิดคะแนน และปรับคะแนนเกณฑ์ผ่านลดลงตามความเหมาะสม

๒. ใช้เป็นประโยชน์ในการปรับปรุงคุณภาพข้อสอบ

ภายหลังจากการรายงานคะแนนสอบเป็นที่เรียบร้อยแล้ว คณะกรรมการสอบสามารถนำผลการวิเคราะห์ข้อสอบแต่ละข้อมาพิจารณาโดยละเอียดเพื่อดูว่าข้อสอบข้อใดสมควรได้รับการปรับปรุงแก้ไข ข้อสอบที่พบว่ายากเกินไปอาจเกิดจากโจทย์คำถามมีความคลุมเครือ ต้องทำการปรับแก้ให้โจทย์ชัดเจนขึ้น หรือเพิ่มเติมข้อมูลบางประการเข้าไปเพื่อให้การวินิจฉัยชัดเจนขึ้น ข้อสอบที่พบว่าง่ายเกินไปอาจพิจารณาปรับให้ยากขึ้นโดยการแก้โจทย์หรือตัวเลือก ข้อสอบที่มีค่า point-biserial ต่ำมักเกิดจากโจทย์ที่คลุมเครือ สร้างความสับสนให้ผู้สอบ สมควรได้รับการปรับโจทย์คำถามใหม่

นอกจากนี้อาจารย์ยังต้องพิจารณาถึงการทำงานของตัวเลือกด้วย ปัญหาที่พบบ่อยมากในการวิเคราะห์ข้อสอบปรนัยคือมีตัวลวงจำนวนมากที่ไม่ทำงาน (มีผู้สอบเลือกน้อยมาก หรือลวงเฉพาะผู้ที่มีความรู้ดีให้มาเลือก) จากการศึกษาวิจัยข้อสอบปรนัยจำนวนมากพบว่าข้อสอบส่วนใหญ่มักมีตัวเลือกที่ทำงานจริงเพียง ๓ ตัวเลือกเท่านั้น[๕] ตัวเลือกที่เหลือเป็นตัวเลือกที่ไม่มีประโยชน์ พิมพ์ลงมาในข้อสอบก็เป็นการเปลืองเนื้อที่หน้ากระดาษ และเสียเวลาอ่านโดยใช่เหตุ อาจารย์ควรพิจารณาตัดตัวลวงที่ไม่ทำงานออกเสียหรือเปลี่ยนเป็นตัวลวงอื่นที่น่าจะมีประสิทธิภาพมากขึ้น

๓. ใช้เป็นประโยชน์ในการบริหารคลังข้อสอบ

ข้อสอบแต่ละข้อนั้นได้มาด้วยความยากลำบาก อาจารย์แต่ละท่านต้องใช้เวลาและความคิดอย่างมากเพื่อพัฒนาข้อสอบที่ดีขึ้นมาใช้ ดังนั้นเมื่อนำข้อสอบมาใช้แล้วผลการวิเคราะห์ข้อสอบแสดงว่าข้อสอบข้อใดเป็นข้อสอบที่ดี มีระดับความยากง่ายเหมาะสม มีความสามารถในการจำแนกผู้สอบที่ดีก็ควรจะพิจารณาเลือกเก็บข้อสอบดังกล่าวไว้ในคลังข้อสอบเพื่อที่จะได้นำกลับมาใช้ใหม่ในอนาคต ในการเก็บข้อสอบเข้าในคลังข้อสอบก็ต้องมีการแนบข้อมูลเกี่ยวกับประวัติการใช้งานและผลการวิเคราะห์ข้อสอบในแต่ละครั้งไว้คู่กันด้วย เพื่อที่จะได้เป็นประโยชน์ในการเลือกข้อสอบมาใช้งาน หากอาจารย์ต้องการข้อสอบที่มีระดับความยากง่าย หรือความสามารถในการจำแนกผู้สอบมากน้อยเพียงใดจะได้ดึงเอาข้อสอบที่มีคุณลักษณะตามต้องการออกมาใช้ได้ตามต้องการ

๔. ใช้เป็นประโยชน์ในการพัฒนาคุณภาพการสอน

การพิจารณาผลการวิเคราะห์ข้อสอบโดยละเอียดในหัวข้อที่อาจารย์ท่านใดท่านหนึ่งรับผิดชอบ

ในการสอนนักเรียนหรือแพทย์ประจำบ้านอยู่นั้นจะทำให้ได้ข้อมูลที่เป็นประโยชน์ในการพัฒนาการเรียนการสอนได้ กล่าวคืออาจารย์สามารถตรวจสอบดูได้ว่านักเรียนหรือแพทย์ประจำบ้านมีความเข้าใจที่ถูกต้องในเรื่องดังกล่าวหรือไม่ ประเด็นใดที่มีผู้เข้าใจผิดอยู่มากก็สมควรที่อาจารย์จะทำการเน้นย้ำในบรรดานักเรียนหรือแพทย์ประจำบ้านในการสอนครั้งต่อ ๆ ไปเพื่อแก้ไขความเข้าใจผิดดังกล่าว ประเด็นใดที่นักเรียนหรือแพทย์ประจำบ้านมีความเข้าใจดีมากอยู่แล้ว อาจารย์อาจไม่ต้องใช้เวลามากนักในการสอนเรื่องดังกล่าว แต่เอาเวลามาใช้สอนในเรื่องที่นักเรียนหรือแพทย์ประจำบ้านยังไม่ค่อยเข้าใจให้มากขึ้นได้

ข้อจำกัดของการวิเคราะห์ข้อสอบ

ถึงแม้ว่าการวิเคราะห์ข้อสอบด้วยวิธีการที่ได้อธิบายมาข้างต้นจะให้ข้อมูลที่เป็นประโยชน์หลายอย่างด้วยกัน แต่เนื่องจากวิธีการวิเคราะห์เหล่านี้เป็นเทคนิคที่วางรากฐานอยู่บนทฤษฎีการสอบแบบดั้งเดิม (classical test theory) ซึ่งมีข้อจำกัดหลายประการด้วยกัน ในการนำค่าต่าง ๆ ที่ได้จากการวิเคราะห์ข้อสอบไปใช้นั้น อาจารย์ควรคำนึงถึงข้อจำกัดของผลการวิเคราะห์ด้วย ในที่นี้จะกล่าวถึงเฉพาะข้อจำกัดในการแปลผลการวิเคราะห์ขั้นพื้นฐานเท่านั้นเนื่องจากเป็นการแปลผลที่ใช้กันทั่วไปในวงการแพทยศาสตรศึกษา ข้อจำกัดในการนำผลการวิเคราะห์ไปประยุกต์ในงานวิจัยทางจิตวิทยาการศึกษายังมีอีกหลายประการที่ผู้นิพนธ์ขอไม่นำมากล่าวในที่นี้ เนื่องจากมีความซับซ้อนและไม่มีที่ใช้ในวงการแพทยศาสตรศึกษาในประเทศไทยในปัจจุบัน

พื้นฐานสำคัญที่เป็นข้อจำกัดของผลการวิเคราะห์ข้อสอบด้วยทฤษฎีการสอบแบบดั้งเดิมคือค่าต่าง ๆ ที่ได้มาจากการวิเคราะห์นั้นขึ้นอยู่กับกลุ่มตัวอย่างที่ใช้ในการเก็บข้อมูล[๑๓,๑๔] หากได้ข้อมูลมาจากกลุ่มตัวอย่างที่มีขนาดใหญ่พอและมีการกระจายตัวของระดับความสามารถของผู้สอบที่เหมาะสม ค่าต่าง ๆ ที่ได้ (p, r, coefficient alpha) จะค่อนข้างเที่ยงตรง ปัญหาที่สำคัญในการวิเคราะห์ข้อสอบในโรงเรียนแพทย์คือการสอบจำนวนมากจัดในนักศึกษากลุ่มเล็ก และนักศึกษาแต่ละกลุ่มก็มีการกระจายตัวของระดับความสามารถแตกต่างกัน นักศึกษาบางกลุ่มมีความสามารถสูงกว่านักศึกษากลุ่มอื่น ดังนั้นผลการวิเคราะห์ข้อสอบไม่ว่าจะเป็นค่า p, r, coefficient alpha, mean, หรือ standard deviation อาจจะเปลี่ยนแปลงไปในแต่ละกลุ่มของนักศึกษา ดังนั้นการนำผลการวิเคราะห์ข้อสอบไปใช้ในทางปฏิบัติจึงมีข้อควรระวังดังต่อไปนี้

การพิจารณาว่าข้อสอบยากหรือง่ายโดยใช้ค่า p นั้นเป็นค่าที่ไม่คงที่ ขึ้นอยู่กับกลุ่มผู้สอบ หากนำข้อสอบข้อหนึ่งไปไปใช้กับนักเรียนกลุ่มที่มีความรู้ดี นักเรียนส่วนใหญ่จะทำข้อสอบได้ถูกต้องทำให้ค่า p สูง แต่เมื่อนำข้อสอบข้อเดิมไปใช้กับนักเรียกลุ่มที่ความรู้ไม่ดีนัก สัดส่วนของนักเรียนที่ทำข้อสอบข้อเดียวกันได้ถูกต้องจะลดลงทำให้ค่า p ลดลง นอกจากนี้ในข้อสอบที่เน้นการท่องจำที่เคยใช้แล้ว เมื่อนำกลับมาใช้ใหม่ในนักเรียนกลุ่มใหม่ อาจมีนักเรียนจำนวนหนึ่งที่สามารถตอบข้อสอบถูกได้เนื่องจากรู้ข้อสอบมาก่อนก็จะทำให้ค่า p สูงขึ้นกว่าเดิมได้

การพิจารณาว่าข้อสอบมีความสามารถในการแยกแยะผู้สอบได้ดีเพียงใดโดยใช้ค่า r ก็ประสบปัญหาในลักษณะเดียวกัน กล่าวคือค่า r นั้นขึ้นกับกลุ่มตัวอย่างของผู้สอบ หากกลุ่มผู้สอบมีระดับความรู้ที่ใกล้เคียงกัน มีคะแนนค่อนข้างเกาะกลุ่มกัน เมื่อคิดค่า r ก็จะได้ต่ำ แต่หากใช้ข้อสอบข้อเดิมในกลุ่มผู้สอบที่มาจากหลายสถาบัน มีความแตกต่างกันของระดับความรู้อย่างมาก ก็จะได้ค่า r สูง

ค่าสัมประสิทธิ์อัลฟ่า เป็นค่าที่มีความเฉพาะเจาะจงกับการสอบของนักเรียนกลุ่มใดกลุ่มหนึ่งเท่านั้น หาใช่เป็นคุณสมบัติติดตัวข้อสอบแต่ละข้อไม่ หากข้อสอบชุดหนึ่งทำการสอบกับนักเรียนกลุ่มหนึ่งแล้วพบว่าคะแนนสอบที่ได้มานั้นมีค่าสัมประสิทธิ์อัลฟ่าสูงในระดับที่ต้องการก็ไม่ได้เป็นตัวรับประกันว่าหากนำข้อสอบชุดเดิมนั้นไปทำการสอบกับนักเรียนกลุ่มอื่นจะได้ค่าสัมประสิทธิ์อัลฟ่าที่สูงเช่นเดียวกัน นอกจากนี้ค่าสัมประสิทธิ์อัลฟ่าที่สูงไม่ได้เป็นตัวบอกถึงคุณภาพของข้อสอบรายข้อแต่อย่างใด

ค่าสัมประสิทธิ์อัลฟ่าที่สูงช่วยบอกแค่เพียงว่า

คะแนนสอบในข้อสอบข้อหนึ่งมีความผันแปรไปในทิศทางเดียวกันกับคะแนนสอบในข้อสอบข้ออื่นในการสอบชุดเดียวกัน นั่นคือในข้อสอบชุดที่มีค่าสัมประสิทธิ์อัลฟ่าสูงก็อาจประกอบไปด้วยข้อสอบที่ดี และข้อสอบที่ไม่ดีรวมกันอยู่ ต้องไปตรวจสอบดัชนีชี้วัดคุณภาพของข้อสอบตัวอื่น ๆ ในแต่ละข้ออีกครั้ง

ข้อควรจำในการวิเคราะห์ข้อสอบที่ผู้นิพนธ์ขอย้ำในตอนท้ายของบทความนี้ก็คือค่าดัชนีชี้วัดคุณภาพต่าง ๆ ของข้อสอบที่กล่าวมาทั้งหมดนี้เป็นเพียงตัวช่วยให้อาจารย์เข้าใจข้อสอบดีขึ้นและช่วยแนะแนวทางในการพัฒนาปรับปรุงข้อสอบให้ดีขึ้น ดัชนีเหล่านี้ไม่ใช่คำตัดสินหรือตัวชี้ชะตาของข้อสอบ ไม่มีดัชนีใดที่ได้จากการวิเคราะห์ข้อสอบจะมาทดแทนดุลยพินิจของอาจารย์ไปได้ ดัชนีคุณภาพของข้อสอบไม่ว่าจะคำนวณมาด้วยวิธีการที่ถูกต้องแล้วก็ตามก็เป็นเพียงตัวเลขที่สามารถเกิดความผิดพลาดในการแปลผลได้ดังเช่นการแปลผลการวิเคราะห์ทางสถิติต่าง ๆ บทบาทของอาจารย์ในการวิเคราะห์ข้อสอบคงไม่ใช่การยึดถือตัวเลขดัชนีต่าง ๆ เป็นกฎตายตัว หากแต่ใช้ดัชนีเหล่านี้ช่วยเป็นแนวทางในการพิจารณาข้อสอบ หากดัชนีตัวใดแนะว่าข้อสอบอาจมีปัญหา อาจารย์ก็นำข้อสอบนั้นมาพิจารณากันโดยคณะกรรมการข้อสอบ หากหลังจากการพิจารณาโดยถี่ถ้วนแล้วอาจารย์คิดว่าข้อสอบข้อนั้นเหมาะสมแล้ว ไม่ควรทำการปรับแก้เนื้อหา อาจารย์ก็ยืนยันไปว่าไม่แก้ไข อาจารย์คงไม่ตัดสินการรักษาผู้ป่วยโดยใช้ผลเลือดตัวใดตัวหนึ่งเป็นเกณฑ์โดยไม่พิจารณาอาการและอาการแสดงของผู้ป่วยร่วมด้วย ฉันใดก็ฉันนั้น อาจารย์ไม่ควรตัดสินชะตากรรมของข้อสอบโดยใช้เพียงค่า p หรือ r โดยไม่พิจารณาความเหมาะสมของเนื้อหาโจทย์และตัวเลือกต่าง ๆ ในข้อสอบข้อนั้น

เอกสารอ้างอิง

๑. Livingston SA. Item analysis. In: Downing SM, Haladyna TM, eds. Handbook of test development. Mahwah, NJ: Lawrence Erlbaum Associates; 2006:421-41.
๒. Brown W, Thomson GH. The essentials of mental measurement, 2nd ed. Cambridge, England: University Press; 1921.
๓. Yen WM, Fitzpatrick AR. Item response theory. In: Brennan RL, ed. Educational measurement, 4th ed. Westport, CT: Praeger Publishers; 2006:111-53.
๔. Haladyna TM. Writing test items to evaluate higher order thinking. Boston, MA: Allyn and Bacon; 1997.
๕. Haladyna TM. Writing multiple choice items. Chicago, IL: CAT Inc.; 2003.
๖. Haladyna TM. Developing and validating multiple-choice test items, 3rd ed. Mahwah, NJ: Lawrence Erlbaum Associates; 2004.
๗. Aleamoni LM, Spencer RE. A comparison of biserial discrimination, point biserial discrimination, and difficulty indices in item analysis data. Educ Psychol Meas 1969;29:353–8.
๘. Haladyna TM, Downing SM. How many options is enough for a multiple-choice test item? Educ Psychol Meas 1993;53:999–1010.
๙. Gronlund NE. Assessment of student achievement, 7th ed. Boston: Allyn & Bacon, 2003.
๑๐. Linn RL, Miller MD. Measurement and assessment in teaching, 9th ed. Upper Saddle River, NJ: Prentice Hall, 2004.
๑๑. Haertel EH. Reliability. In: Brennan RL, editor. Educational measurement, 4th ed. Westport, CT: Praeger Publishers; 2006:65-110.
๑๒. Downing SM. Reliability: On the reproducibility of assessment data. Med Educ 2004;38:1006–12.
๑๓. Embretson SE, Reise SP. Item response theory for psychologists. Mahwah, NJ: Lawrence Erlbaum Associates, 2000.
๑๔. Smith EV. Evidence for the reliability of measures and validity of measure interpretation: A Rasch measurement perspective. In: Smith EV. Smith RM, eds. Introduction to Rasch measurement: Theory, models, and applications. Maple Grove, MN: JAM Press, 2004:93–112

โปรแกรมวิเคราะห์ข้อสอบ

รุ่น 2.0

การสอบ : SIID 521 (Basic Sciences)

วันที่ : 22 ธันวาคม 2555

จำนวนข้อสอบ = 120

จำนวนผู้เข้าสอบ = 244

Difficulty Index --> p-value (proportion of students answer item correctly)

$$\text{p-Value} = \frac{\text{number of students answer correctly}}{\text{total number of students answer that item}}$$

Discrimination Index --> D or r-value --> Point-biserial correlation coefficient (r^{pbi})

==========

SCORE STATISTICS

Mean = **68.152** S.D. = **11.915**

Mode = **65** (freq = **14**)

Max = **94** Min = **28**

DIFFICULTY INDEX (p value)

Average (p-bar) = **0.566** Max p = **0.990** Min p = **0.010**

DISCRIMINATION INDEX (D or r value)

Average (D-bar) = **0.244** Max D = **0.680** Min D = **-0.180**

RELIABILITY COEFFICIENT (rtt) = **0.847**
(Kuder-Richardson formula 20)

STANDARD ERROR OF MEASUREMENT (SEM) = **4.655**
(S.D. x SQR(1-rtt))

Item Analysis and Option Analysis

Faculty of Medicine Siriraj Hospital
Mahidol University

No. : 1 p Value : 0.55 r_{pbi} : 0.37

A		B		* C		D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
-0.24	21.31	-0.10	13.52	0.37	54.92	-0.16	6.15	-0.07	4.10

No. : 2 p Value : 0.74 r_{pbi} : 0.00

A		B		C		* D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
0.02	5.33	0.07	11.48	-0.02	1.23	0.00	74.18	-0.09	7.79

No. : 3 p Value : 0.84 r_{pbi} : 0.25

A		* B		C		D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
-0.22	14.34	0.25	84.43	0.01	0.41	0.00	0.00	-0.12	0.41

No. : 4 p Value : 0.68 r_{pbi} : 0.43

A		B		* C		D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
-0.26	8.20	-0.09	8.20	0.43	68.03	-0.06	1.64	-0.29	13.93

No. : 5 p Value : 0.92 r_{pbi} : 0.26

A		B		* C		D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
-0.16	4.10	-0.07	0.41	0.26	91.80	-0.16	2.87	-0.08	0.82

No. : 6 p Value : 0.75 r_{pbi} : 0.30

* A		B		C		D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
0.30	74.59	-0.03	13.93	-0.22	2.87	-0.24	3.69	-0.17	4.92

No. : 7 p Value : 0.99 r_{pbi} : 0.06

A		B		C		D		* E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
-0.06	0.82	0.00	0.00	0.00	0.00	0.00	0.00	0.06	99.18

No. : 8 p Value : 0.70 r_{pbi} : 0.53

* A		B		C		D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
0.53	70.49	-0.13	1.23	-0.21	5.74	-0.38	17.21	-0.17	5.33

No. : 9 p Value : 0.63 r_{pbi} : 0.19

A		B		C		D		* E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
0.00	0.41	0.00	0.00	0.01	2.05	-0.19	34.43	0.19	63.11

No. : 10 p Value : 0.90 r_{pbi} : 0.25

* A		B		C		D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
0.25	90.16	-0.09	0.41	-0.22	9.02	-0.08	0.41	0.00	0.00

No. : 11 p Value : 0.54 r_{pbi} : 0.48

A		B		C		* D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
-0.44	31.97	-0.09	4.51	-0.05	8.61	0.48	53.69	-0.06	1.23

No. : 12 p Value : 0.55 r_{pbi} : 0.47

A		* B		C		D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
-0.27	28.28	0.47	54.92	0.00	0.00	-0.24	11.07	-0.16	5.74

No. : 13 p Value : 0.81 r_{pbi} : 0.32

A		B		* C		D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
-0.23	5.33	-0.16	9.84	0.32	81.15	-0.13	3.28	-0.06	0.41

No. : 14 p Value : 0.45 r_{pbi} : 0.39

A		B		C		D		* E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
-0.22	34.84	-0.09	1.64	-0.17	11.89	-0.08	6.15	0.39	45.49

No. : 15 p Value : 0.73 r_{pbi} : 0.32

A		* B		C		D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
-0.24	2.46	0.32	72.95	-0.17	2.05	-0.17	21.72	-0.07	0.41

No. : 16 p Value : 0.09 r_{pbi} : -0.03

A		B		C		D		* E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
-0.14	11.89	0.15	70.08	-0.18	3.28	0.08	5.74	-0.03	8.61

No. : 17 p Value : 0.36 r_{pbi} : 0.13

A		B		* C		D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
-0.05	4.10	0.06	22.13	0.13	35.66	-0.07	9.43	-0.12	28.69

No. : 18 p Value : 0.83 r_{pbi} : 0.06

* A		B		C		D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
0.06	82.79	0.01	0.82	-0.05	2.05	-0.10	4.92	0.01	9.43

Item Analysis and Option Analysis

Faculty of Medicine Siriraj Hospital
Mahidol University

No. : 19 p Value : 0.25 r_{pbi} : 0.04

A		B		C		* D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
-0.10	51.23	0.04	13.11	0.00	0.00	0.04	24.59	0.05	11.07

No. : 20 p Value : 0.36 r_{pbi} : 0.55

A		* B		C		D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
-0.21	22.54	0.55	35.66	-0.12	2.46	-0.25	34.43	-0.19	4.92

No. : 21 p Value : 0.81 r_{pbi} : 0.20

* A		B		C		D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
0.20	80.74	-0.07	3.69	-0.13	11.89	-0.05	1.64	-0.11	2.05

No. : 22 p Value : 0.46 r_{pbi} : 0.47

* A		B		C		D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
0.47	45.90	-0.14	6.15	-0.11	4.92	-0.18	17.21	-0.24	25.82

No. : 23 p Value : 0.00 r_{pbi} : -0.06

A		B		* C		D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
-0.03	0.41	0.00	0.41	-0.06	0.41	-0.14	4.10	0.16	94.26

No. : 24 p Value : 0.64 r_{pbi} : 0.40

A		B		* C		D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
-0.08	5.33	-0.16	9.43	0.40	64.34	-0.20	9.02	-0.21	11.89

No. : 25 p Value : 0.61 r_{pbi} : 0.40

A		B		C		* D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
-0.15	2.87	-0.10	13.11	-0.23	14.34	0.40	60.66	-0.19	9.02

No. : 26 p Value : 0.70 r_{pbi} : 0.47

A		B		C		D		* E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
-0.15	7.38	-0.22	9.84	-0.26	7.79	-0.18	5.33	0.47	69.67

No. : 27 p Value : 0.51 r_{pbi} : 0.35

A		* B		C		D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
-0.15	9.02	0.35	50.82	-0.26	25.82	-0.05	5.33	-0.02	9.02

No. : 28 p Value : 0.50 r_{pbi} : 0.17

* A		B		C		D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
0.17	49.59	-0.17	20.49	-0.03	4.51	-0.04	15.98	0.01	9.43

No. : 29 p Value : 0.75 r_{pbi} : 0.17

A		B		C		D		* E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
-0.09	14.34	-0.16	3.28	-0.01	2.87	-0.06	4.92	0.17	74.59

No. : 30 p Value : 0.58 r_{pbi} : 0.37

A		B		* C		D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
-0.22	6.15	-0.30	31.15	0.37	57.79	0.05	4.92	0.00	0.00

No. : 31 p Value : 0.86 r_{pbi} : 0.28

* A		B		C		D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
0.28	86.07	-0.05	2.05	-0.21	9.43	-0.10	1.23	-0.17	1.23

No. : 32 p Value : 0.88 r_{pbi} : 0.32

A		B		* C		D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
-0.30	8.20	-0.16	2.87	0.32	87.70	0.03	1.23	0.00	0.00

No. : 33 p Value : 0.44 r_{pbi} : 0.37

A		* B		C		D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
0.09	4.92	0.37	44.26	-0.41	45.08	0.01	2.46	-0.03	3.28

No. : 34 p Value : 0.73 r_{pbi} : 0.25

* A		B		C		D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
0.25	72.54	-0.22	9.02	-0.15	6.15	-0.05	1.23	-0.02	11.07

No. : 35 p Value : 0.45 r_{pbi} : 0.42

A		B		C		D		* E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
0.06	9.02	-0.18	12.30	-0.38	18.44	-0.06	15.16	0.42	45.08

No. : 36 p Value : 0.68 r_{pbi} : 0.35

A		B		* C		D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
-0.15	4.51	-0.29	16.39	0.35	68.03	-0.04	6.97	-0.07	4.10

Item Analysis and Option Analysis

Faculty of Medicine Siriraj Hospital
Mahidol University

No. : 37 p Value : 0.29 r_{pbi} : -0.02

A		B		C		D		* E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
-0.05	2.05	0.22	52.05	-0.14	7.38	-0.20	9.84	-0.02	28.69

No. : 38 p Value : 0.75 r_{pbi} : 0.11

* A		B		C		D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
0.11	74.59	-0.11	22.95	-0.14	0.82	0.08	0.82	0.08	0.82

No. : 39 p Value : 0.51 r_{pbi} : 0.23

A		B		* C		D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
-0.02	10.25	-0.21	27.46	0.23	51.23	-0.07	9.02	0.09	1.64

No. : 40 p Value : 0.21 r_{pbi} : 0.13

A		* B		C		D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
0.00	40.57	0.13	20.90	0.00	4.51	0.07	17.62	-0.21	16.39

No. : 41 p Value : 0.42 r_{pbi} : -0.03

A		B		C		* D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
0.02	7.38	0.07	43.03	-0.02	0.41	-0.03	41.80	-0.10	7.38

No. : 42 p Value : 0.79 r_{pbi} : 0.33

A		* B		C		D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
-0.22	5.33	0.33	79.10	-0.20	4.92	-0.02	2.87	-0.15	7.79

No. : 43 p Value : 0.81 r_{pbi} : 0.37

* A		B		C		D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
0.37	80.74	-0.33	14.75	0.01	0.82	-0.14	2.05	-0.07	1.64

No. : 44 p Value : 0.56 r_{pbi} : 0.34

A		B		* C		D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
-0.14	1.64	-0.18	6.56	0.34	55.74	-0.22	20.08	-0.05	15.98

No. : 45 p Value : 0.86 r_{pbi} : 0.39

A		B		C		D		* E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
-0.16	2.05	-0.11	0.82	-0.04	1.23	-0.33	9.84	0.39	86.07

No. : 46 p Value : 0.81 r_{pbi} : 0.31

A		* B		C		D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
-0.19	10.66	0.31	80.74	-0.09	2.87	-0.15	1.64	-0.15	4.10

No. : 47 p Value : 0.93 r_{pbi} : 0.26

A		* B		C		D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
-0.14	2.46	0.26	93.44	-0.01	0.82	-0.17	1.64	-0.15	1.64

No. : 48 p Value : 0.07 r_{pbi} : -0.20

A		B		C		* D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
-0.20	12.70	-0.08	4.51	-0.18	2.87	-0.20	6.56	0.37	73.36

No. : 49 p Value : 0.95 r_{pbi} : 0.21

A		B		C		* D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
0.00	0.00	0.00	0.00	-0.21	4.92	0.21	95.08	0.00	0.00

No. : 50 p Value : 0.83 r_{pbi} : 0.24

A		B		* C		D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
0.00	0.00	0.00	0.00	0.24	83.20	-0.23	15.98	-0.09	0.82

No. : 51 p Value : 0.76 r_{pbi} : 0.26

* A		B		C		D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
0.26	76.23	-0.14	2.87	-0.04	2.46	0.07	0.41	-0.23	18.03

No. : 52 p Value : 0.70 r_{pbi} : 0.24

A		B		C		* D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
-0.15	0.82	-0.21	11.89	0.01	12.70	0.25	70.08	-0.16	4.51

No. : 53 p Value : 0.51 r_{pbi} : 0.31

A		* B		C		D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
0.02	4.51	0.31	50.82	-0.07	2.05	-0.07	2.87	-0.28	39.75

No. : 54 p Value : 0.37 r_{pbi} : 0.28

A		* B		C		D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
-0.07	9.43	0.28	36.89	-0.19	13.52	-0.09	16.80	-0.04	23.36

Item Analysis and Option Analysis

Faculty of Medicine Siriraj Hospital
Mahidol University

No. : 55 p Value : 0.71 r_{pbi} : 0.25

A		B		C		* D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
-0.18	2.87	-0.20	14.75	-0.08	5.74	0.25	70.90	0.01	5.74

No. : 56 p Value : 0.81 r_{pbi} : 0.29

A		* B		C		D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
-0.02	1.23	0.29	81.15	-0.15	7.38	-0.10	4.92	-0.22	5.33

No. : 57 p Value : 0.26 r_{pbi} : 0.19

A		B		C		* D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
-0.08	6.15	-0.17	29.51	-0.01	15.57	0.19	26.23	0.03	22.54

No. : 58 p Value : 0.66 r_{pbi} : 0.29

A		B		C		* D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
-0.16	25.00	-0.14	2.46	-0.22	0.41	0.29	65.98	-0.14	6.15

No. : 59 p Value : 0.73 r_{pbi} : 0.36

A		B		C		* D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
-0.13	0.82	-0.25	19.67	-0.26	5.33	0.36	73.36	0.10	0.82

No. : 60 p Value : 0.93 r_{pbi} : 0.28

A		B		C		D		* E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
0.00	0.00	-0.13	4.10	-0.27	2.87	-0.03	0.41	0.28	92.62

No. : 61 p Value : 0.89 r_{pbi} : 0.26

A		B		C		D		* E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
0.05	0.41	-0.30	2.46	-0.13	5.74	-0.06	2.46	0.26	88.93

No. : 62 p Value : 0.89 r_{pbi} : 0.38

A		B		C		* D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
-0.32	7.38	-0.09	0.82	-0.17	3.28	0.38	88.52	0.00	0.00

No. : 63 p Value : 0.69 r_{pbi} : 0.05

A		B		C		* D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
0.00	0.00	-0.12	1.64	-0.02	29.51	0.05	68.85	0.00	0.00

No. : 64 p Value : 0.81 r_{pbi} : 0.20

A		B		* C		D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
-0.09	0.82	0.05	2.46	0.20	80.74	-0.16	11.89	-0.10	3.69

No. : 65 p Value : 0.68 r_{pbi} : 0.10

A		B		* C		D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
-0.06	9.43	-0.15	1.64	0.10	68.44	-0.04	1.23	-0.01	19.26

No. : 66 p Value : 0.55 r_{pbi} : 0.32

A		B		* C		D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
-0.22	23.36	-0.08	11.48	0.32	54.92	-0.11	6.15	-0.07	4.10

No. : 67 p Value : 0.45 r_{pbi} : 0.29

A		B		C		* D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
-0.20	26.64	-0.07	17.62	-0.05	1.23	0.29	45.49	-0.06	8.61

No. : 68 p Value : 0.28 r_{pbi} : -0.03

A		B		* C		D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
0.02	14.34	0.07	1.64	-0.03	27.87	0.06	10.25	-0.04	45.90

No. : 69 p Value : 0.39 r_{pbi} : 0.37

A		B		C		* D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
-0.05	23.77	-0.07	13.93	-0.22	0.41	0.37	38.93	-0.28	22.95

No. : 70 p Value : 0.25 r_{pbi} : 0.13

A		* B		C		D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
-0.02	7.79	0.13	24.59	-0.10	1.64	0.06	10.66	-0.10	54.92

No. : 71 p Value : 0.80 r_{pbi} : 0.09

* A		B		C		D		E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
0.09	80.33	-0.03	1.64	-0.13	3.28	0.00	5.74	-0.03	9.02

No. : 72 p Value : 0.65 r_{pbi} : 0.37

A		B		C		D		* E	
r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%	r_{pbi}	%
-0.25	6.97	-0.05	6.56	-0.23	20.08	-0.05	1.23	0.37	65.16

23 June 2017

หัวข้อ : การตั้งเกณฑ์ผ่านการสอบ

www.si-merd.com June 2017

การตั้งเกณฑ์ผ่านการสอบ

นพ.เชิดศักดิ์ ไอรมณีรัตน์
ภาควิชาศัลยศาสตร์
คณะแพทยศาสตร์ศิริราชพยาบาล

Twelve Steps for Test Development

1. Overall plan
2. Content definition
3. Test specification
4. Item development
5. Test assembly
6. Test production
7. Test administration
8. Scoring
9. Passing score
10. Reporting results
11. Item banking
12. Technical report

Downing SM, Haladyna TM. Handbook of test development. Mahwah, NJ: Lawrence Earlbaum Associates 2006.

2

Standard

- A score that is set to be a boundary between those who perform well enough on the test (pass) from those who do not (fail).
- Standard = cutpoint

Objectives

- เมื่อสิ้นสุดการบรรยายแล้ว ผู้เข้าอบรมสามารถ
  - บอกถึงความสำคัญของการตั้งเกณฑ์ผ่านได้ถูกต้อง
  - บอกถึงขั้นตอนของการตั้งเกณฑ์ผ่านได้ถูกต้อง
  - ยกตัวอย่างวิธีการตั้งเกณฑ์ผ่านได้อย่างน้อยสามวิธี
  - จัดทำเกณฑ์ผ่านการสอบ MCQ ด้วยวิธีการ modified Angoff method ในการสอบที่ตนเกี่ยวข้องได้อย่างเหมาะสม

Outline

- Basic concepts
- Steps in setting standards
  - The type of standard
  - The method
  - Selecting judges
  - Standard setting meeting
  - Calculate the standards
  - Checking the standards

Basic Concepts

- A standard is an answer to the question, “How much is enough?”
- The classification of examinees into two groups can result in two types of wrong decisions
  - False positive: Passing an examinee who should fail the exam
  - False negative: Failing an examinee who should pass the exam

Cherdsak.ira@mahidol.ac.th

1

Judgment

1. Made by qualified judges
2. Meaningful to the persons who are making the decision
3. Made in a way that takes into account the purpose of the test

Steps in Setting Standards

1. Deciding on the type of standard
2. Deciding on the method for setting standards
3. Selecting judges
4. Holding the standard setting meeting
5. Calculating the standards
6. Checking the standards after test

1. Types of Standards

- Absolute standard
- Relative standard

Absolute Standard

- The standard is fixed, based on specific criteria of performance, but may undergo periodic re-evaluation of the standard
- Strengths
  - A standard is known in advance
  - A stable performance level is required to pass the examination => content-related standard
  - Provide clear feedback to examinees
  - Nobody has to fail the exam if their knowledge/skills is adequate for the purpose of the exam.
  - Promote a collaborative learning environment.

Relative Standard

- The standard is set in reference to the group of examinees. The resulting standard may be reasonable providing a representative heterogeneous group.
- Strengths
  - The failure rate is stable, which in someway is easy for curriculum management

2.Methods for Setting Standards

1. Test-centered methods
2. Examinee-centered methods
3. Compromised methods

Test-Centered Methods

- The judges set standards by reviewing the test items and provide judgments regarding the "just adequate" level of performance on these items.
  - Angoff's method
  - Nedelsky's method
  - Ebel's method

Modified Angoff's Method

- The judgment
  - The probability that a borderline examinee would answer the test item correctly
- The passing score
  - The sum of all the probability of correct answers for all items on the exam

Modified Angoff's Method (2)

Item	Probability
1	0.8
2	0.6
3	0.4
4	0.5
5	0.5
Passing score	2.8

Nedelsky's Method

- The judgment
  - How many options a borderline examinee can eliminate from choosing in an item
- The passing score
  - The probability of correct answer for an item = 1/(the number of options not eliminated)
  - The passing score of the test = the sum of all the probability of correct answers of all items on the test

Nedelsky's Method (2)

Item	A	B	C	D	E	Not eliminated	Probability
1			X	X	X	2	1/2 =0.50
2	X	X				3	1/3 = 0.33
3	X					4	1/4 = 0.25
4	X		X	X		2	1/2 = 0.50
5	X				X	3	1/3 = 0.33
Passing score							1.91

Ebel's Method

- The judgment
  - What is the level of difficulty of an item?
    - Easy/Medium/difficult
  - What is the level of importance of that content in clinical practice?
    - Essential/Important/Acceptable/Questionable
  - The probability that a borderline examinee will answer an item in each category correctly
- The passing score
  - The sum of all the probability of correct answers for all items on the exam

Ebel's Method (2)

	Easy	Medium	Difficult
Essential	0.95	0.85	0.80
Important	0.90	0.75	0.60
Acceptable	0.85	0.60	0.40
Questionable	0.55	0.45	0.35

Ebel's Method (3)

Item	Difficulty	Importance	Probability
1	Easy	Essential	0.95
2	Easy	Importance	0.90
3	Difficult	Essential	0.80
4	Difficult	Acceptable	0.40
5	Medium	Acceptable	0.60
Passing score			3.65

2.Methods for Setting Standards

1. Test-centered methods
2. Examinee-centered methods
3. Compromised methods

Examinee-Centered Methods

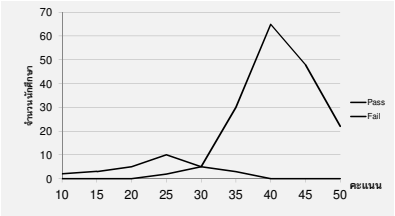
- The judges set a standard by reviewing the overall performance of examinees and determine who should pass and who should fail. The scores of examinees are reviewed and the passing score is set based on these judgments
  - Borderline-group method
  - Contrasting-groups method

Borderline-Group Method

- The judgment
  - Identify examinees who are "borderline"
- The passing score
  - The median score of this "borderline group"

Contrasting-Groups Method

- The judgment
  - Identify examinees who should "pass" and those who should "fail"
- The passing score



2.Methods for Setting Standards

1. Test-centered methods
2. Examinee-centered methods
3. Compromised methods

Compromised Method

- Combining relative and absolute standard setting methods
  - Hofstee method

Hofstee Method

- The judgment
  - Minimum failure rate
  - Maximum failure rate
  - Minimum passing score
  - Maximum passing score
- The passing score
  - The intersection of test scores curve with diagonal line drawn from upper left to lower right corner

3. Selecting Judges

- The number of judges
- The qualification of judges

4. Standard Setting Meeting

- Discussion of the purpose of the test, the characteristics of examinees, and the nature of competence.
- Explanation of the method and practice before the real standard setting procedure.

Practice

- A modified Angoff standard setting method

1. Defining the characteristics of borderline examinees
2. Estimating the performance of borderline candidates for each item
3. Calculating the passing standard for the exam

Borderline Examinees

- Examinees who have an equal chance of passing or failing the exam
- Individual members of standard-setting committee should think about few candidates they have known who are marginally competent.
- Independent proposal of definitions
- Group consensus on the definition
- This exercise: An extern who has minimal ability to provide safe medical service

Item Review and Rating

- Independent review of an item and provide a rating
  1. How many individuals in a group of 100 borderline examinees will answer the item correctly?
  2. What is the probability that one borderline examinee will answer this item correctly?
- Collection of independent ratings of all items
- Group discussion of the ratings

Calculating the Passing Score

- If the ratings differ less than 20%, the average of the rating for that item is used.
- If the ratings differ by 20% or more, group discussion is carried out, and the members of the committee can modify their ratings.
- Passing score for the exam = the sum of mean ratings of all items on the exam

5. Calculating Standard

- Outliers
- Errors of the cutpoint

Do we have to care about error?

- True score theory
  - Each student has a true score, a hypothetical value representing a score free of error.
  - If we test a student repeatedly, the average of the obtained scores would approximate the true score, with a standard deviation of SEM.

SEM

$$SEM = SD\sqrt{(1-r)}$$

SD = standard deviation
r = internal consistency reliability

↑SD (more spread of score): higher SEM
↑r (more accurate measures): smaller SEM

What should we do with students with an SEM around cut score?

- False positive: Passing students who should have fail the examination
- False negative: Failing students who should have pass the examination

6. Checking Standard

- Stakeholders' acceptance of the results
- Relationship with other markers of competence
- Prediction of future performance

Summary

- Steps in setting up a standard

1. Deciding on the type of standard
2. Deciding on the method for setting standards
3. Selecting judges
4. Holding the standard setting meeting
5. Calculating the standards
6. Checking the standards after test

"It does not matter how slowly you go, as long as you do not stop."

Confucius

40

Iramaneerat C. Passing standard: Part I [Thai]. Medical Education Pamphlet 2006; 2(1): 3.

วิธีการตั้งเกณฑ์สอบผ่าน (passing standard) (ตอนที่ 1)

เชิดศักดิ์ ไอรมณีรัตน์

เกณฑ์สอบผ่าน (passing standard) คือคะแนนสอบที่น้อยที่สุดที่คณาจารย์ยินยอมให้นักเรียนสามารถสอบผ่าน นักเรียนที่สอบได้คะแนนน้อยกว่าเกณฑ์สอบผ่านจะถูกตัดสินว่าสอบตก การตั้งเกณฑ์สอบผ่านจัดเป็นขั้นตอนที่มีความสำคัญมากในการจัดสอบ แต่กลับไม่ได้รับความสนใจเท่าที่ควรในการวัดผลทางแพทยศาสตรศึกษาจำนวนมาก ในบทความนี้ผมขอเสนอเกร็ดความรู้เกี่ยวกับวิธีการตั้งเกณฑ์สอบผ่าน ผมหวังว่าอาจารย์ผู้อ่านจะสามารถนำเกร็ดความรู้นี้ไปใช้พัฒนาคุณภาพของการตั้งเกณฑ์สอบผ่านได้ไม่มากก็น้อยครับ

เกณฑ์สอบผ่านในทางแพทยศาสตรศึกษาจัดว่ามีความสำคัญมากเนื่องจากเกณฑ์สอบผ่านเป็นการแสดงออกถึงมาตรฐานของวิชาชีพที่อาจารย์ยอมรับ เกณฑ์สอบผ่านที่ดีต้องได้รับการตั้งขึ้นโดยใช้ดุลยพินิจของคณาจารย์ผู้เชี่ยวชาญในสาขาวิชานั้นๆเพื่อรักษามาตรฐานการประกอบวิชาชีพเพื่อให้สังคมได้รับบริการทางการแพทย์ที่มีคุณภาพ ในขณะเดียวกันกับให้ความเป็นธรรมกับนักเรียนผู้สอบ เนื่องจากเกณฑ์สอบผ่านเป็นการแสดงออกถึง "ความยอมรับได้" ในดุลยพินิจของคณาจารย์ผู้เชี่ยวชาญ จึงไม่มีวิธีการทางวิทยาศาสตร์ใดที่จะตัดสินว่าเกณฑ์ที่ตั้งขึ้นนั้นถูกหรือผิด สิ่งที่สำคัญที่สุดในการตั้งเกณฑ์สอบผ่านหาใช่ "ตัวเลข" คะแนนที่จะใช้ตัดสินได้ตก หากแต่เป็น "กระบวนการ" ให้ได้มาซึ่งเกณฑ์ดังกล่าว เกณฑ์สอบผ่านที่ตั้งขึ้นโดยใช้อาจารย์ 1 ท่านเลือกตัวเลข 1 ตัวเลขขึ้นมาโดยไม่ได้พิจารณาถึงข้อสอบหรือนักเรียนผู้สอบ เป็นวิธีการตั้งเกณฑ์ที่ล่อแหลมต่อการถูกวิจารณ์ (และประท้วง) โดยผู้ที่ไม่พอใจในผลสอบ วิธีการตั้งเกณฑ์สอบผ่านที่ดีนั้นต้องมีหลักการและเหตุผลประกอบ และผ่านดุลยพินิจของคณาจารย์ จำนวนของอาจารย์ผู้เชี่ยวชาญที่ต้องใช้ในการตั้งเกณฑ์นั้นขึ้นกับความสำคัญของการสอบนั้นๆ ในการสอบที่มีความสำคัญสูงเช่นการสอบวุฒิบัตรแพทย์ผู้เชี่ยวชาญ แนะนำให้ใช้คณาจารย์อย่างน้อย 6 – 8 ท่าน ในการตั้งเกณฑ์ แต่หากเป็นการสอบเล็กๆ เช่น การทดสอบหลังการสอนกลุ่มย่อย อาจใช้อาจารย์เพียง 1 ท่านก็ได้

การตั้งเกณฑ์สอบผ่านมี 2 ชนิดคือ การตัดสินแบบอิงเกณฑ์ (criterion-referenced standard, absolute standard) และการตัดสินแบบอิงกลุ่ม (norm-referenced standard, relative standard) การตัดสินแบบอิงเกณฑ์ เป็นการตั้งว่า <u>คะแนน</u>เท่าไร จึงจัดว่าผ่านการสอบ ในทางตรงข้าม การตัดสินแบบอิงกลุ่ม เป็นการตั้งว่า จะให้ <u>นักเรียน</u>จำนวนเท่าไร ผ่านการสอบ การตัดสินแบบอิงเกณฑ์นั้นเหมาะกับการสอบเพื่อวัดว่าผู้สอบมีความรู้ความสามารถในด้านใดด้านหนึ่งเพียงพอหรือไม่ ส่วนการสอบแบบอิงกลุ่มนั้นเหมาะสำหรับการสอบแข่งขันเพื่อเข้าศึกษาต่อ หรือ ทำงาน ในสถาบันที่มีตำแหน่งที่จะรับได้จำกัด เช่น การสอบเข้าโรงเรียนแพทย์ หรือ การสอบคัดเลือกแพทย์ประจำบ้าน การสอบส่วนใหญ่ในทางแพทยศาสตรศึกษานั้นเหมาะกับการตัดสินแบบอิงเกณฑ์ หากผู้สอบทุกคนมีความสามารถเพียงพอก็ไม่จำเป็นต้องมีผู้สอบตก การใช้การตัดสินแบบอิงกลุ่มเพื่อวัดความรู้ความสามารถในสถานการณ์อื่นนอกจากการสอบคัดเลือกนั้นเป็นการส่งเสริมให้นักเรียนเกิดความแข่งขันกัน (แทนที่จะช่วยกันเรียน) โดยไม่จำเป็น

เนื่องจากการสอบทางแพทยศาสตรศึกษาแทบทั้งหมดเหมาะกับการตั้งเกณฑ์สอบผ่านแบบอิงเกณฑ์ ผมจะขอขยายความวิธีการตั้งเกณฑ์สอบผ่านแบบอิงเกณฑ์ที่สำคัญและใช้บ่อย 2 วิธีใหญ่ๆ คือ 1. การตั้งเกณฑ์โดยพิจารณาข้อสอบ และ 2. การตั้งเกณฑ์โดยพิจารณาจากผู้สอบ ในบทความตอนต่อไปครับ

Iramaneerat C. Passing standard: Part II [Thai]. Medical Education Pamphlet 2006; 2(2): 2.

วิธีการตั้งเกณฑ์สอบผ่าน (passing standard) (ตอนที่ 2)

เชิดศักดิ์ ไอรมณีรัตน์

ในบทความนี้ผมจะขอแนะนำวิธีการตั้งเกณฑ์สอบผ่านโดยพิจารณาตัวข้อสอบที่ใช้สอบ วิธีการตั้งเกณฑ์ผ่านแบบนี้เหมาะสำหรับการสอบ multiple-choice questions ซึ่งอาจารย์ผู้ตั้งเกณฑ์ผ่านสามารถประเมินความน่าจะเป็นของการตอบข้อสอบแต่ละข้อถูกได้ การตั้งเกณฑ์ผ่านแบบนี้ประกอบด้วย 3 ขั้นตอนหลักคือ

1. ระบุลักษณะของนักเรียน"คาบเส้น" (borderline examinees): นักเรียนในกลุ่มคาบเส้นนี้คือนักเรียนที่มีความรู้ความสามารถอยู่ระหว่าง "ยอมรับได้" กับ "ยอมรับไม่ได้" นักเรียนกลุ่มนี้มีความรู้ไม่มากพอที่อาจารย์จะตัดสินให้สอบผ่านได้อย่างสบายใจ แต่ก็มีความรู้ไม่น้อยจนอาจารย์จะตัดสินให้สอบตกได้โดยไม่มีข้อสงสัย คณะกรรมการตั้งเกณฑ์สอบผ่านต้องระบุลักษณะของนักเรียนในกลุ่มคาบเส้นนี้อย่างชัดเจนว่า ในเนื้อหาวิชาที่ทำการสอบ นักเรียนกลุ่มนี้ควรมีความรู้ในเรื่องใด และไม่มีความรู้ในเรื่องใด ขั้นตอนนี้อาจทำได้ง่ายขึ้นหากอาจารย์แต่ละท่านนึกภาพของนักเรียนจริงที่อาจารย์เคยรู้จักที่สมควรถูกจัดให้อยู่ในกลุ่มนักเรียนคาบเส้น แล้วบรรยายลักษณะของนักเรียนคนนั้นๆ ว่าทำอะไรได้ และทำอะไรไม่ได้ รู้เรื่องอะไรบ้าง ไม่รู้เรื่องอะไรบ้าง
2. ให้กรรมการแต่ละท่านพิจารณาข้อสอบแต่ละข้อ และตัดสินว่านักเรียนคาบเส้นน่าจะมีโอกาสตอบข้อสอบถูกมากน้อยเพียงใด ขั้นตอนนี้สามารถทำได้หลายวิธีด้วยกัน ผมขอยกตัวอย่างวิธีที่เป็นที่แพร่หลายมาก 2 วิธีด้วยกัน คือ
   2.1. Angoff's method: ให้อาจารย์ระบุว่าหากนักเรียนคาบเส้น 100 คนทำข้อสอบข้อนั้น จะมีนักเรียนกี่คนที่ตอบข้อสอบข้อนั้นถูก (หรือความน่าจะเป็นที่นักเรียนคาบเส้นตอบข้อสอบข้อนั้นถูก)
   2.2. Ebel's method: ให้อาจารย์สร้างตารางแยกประเภทข้อสอบตามความสำคัญของเนื้อหาและตามความยากง่ายของข้อสอบและระบุว่าในช้อสอบแต่ละกลุ่ม หากนักเรียนคาบเส้น 100 คนทำข้อสอบจะมีนักเรียนกี่คนที่ตอบถูก หลังจากนั้นให้อาจารย์พิจารณาข้อสอบแต่ละข้อแล้วจัดประเภทเข้าในกลุ่ม ตัวอย่างเช่น

ความยากง่าย / ความสำคัญ	ง่าย	ปานกลาง	ยาก
สำคัญมาก	95%	85%	80%
สำคัญพอควร	90%	75%	60%
สำคัญน้อย	80%	55%	35%
สำคัญน้อยมาก	50%	30%	20%

3. ทำการคิดเกณฑ์สอบผ่านสำหรับข้อสอบนั้น

   3.1. Angoff's method เกณฑ์ผ่านคือผลรวมของความน่าจะเป็นของการตอบข้อสอบแต่ละข้อถูก

Item	1	2	3	4	5	Passing score
Probability	0.95	0.85	0.30	0.40	0.70	3.20

   3.2. Ebel's method เกณฑ์ผ่านคือผลรวมของ (จำนวนข้อสอบในแต่ละกลุ่ม x ความน่าจะเป็นของการตอบข้อสอบถูกสำหรับข้อสอบในกลุ่มนั้น)จากข้อสอบทั้ง 12 กลุ่ม

ความยากง่าย / ความสำคัญ	ง่าย (24 ข้อ)	ปานกลาง (15 ข้อ)	ยาก (11 ข้อ)
สำคัญมาก (15 ข้อ)	95% x 5	85% x 5	80% x 5
สำคัญพอควร (20 ข้อ)	90% x 10	75% x 7	60% x 3
สำคัญน้อย (10 ข้อ)	80% x 5	55% x 3	35% x 2
สำคัญน้อยมาก (5 ข้อ)	50% x 4	30% x 0	20% x 1
Passing score	37.6		

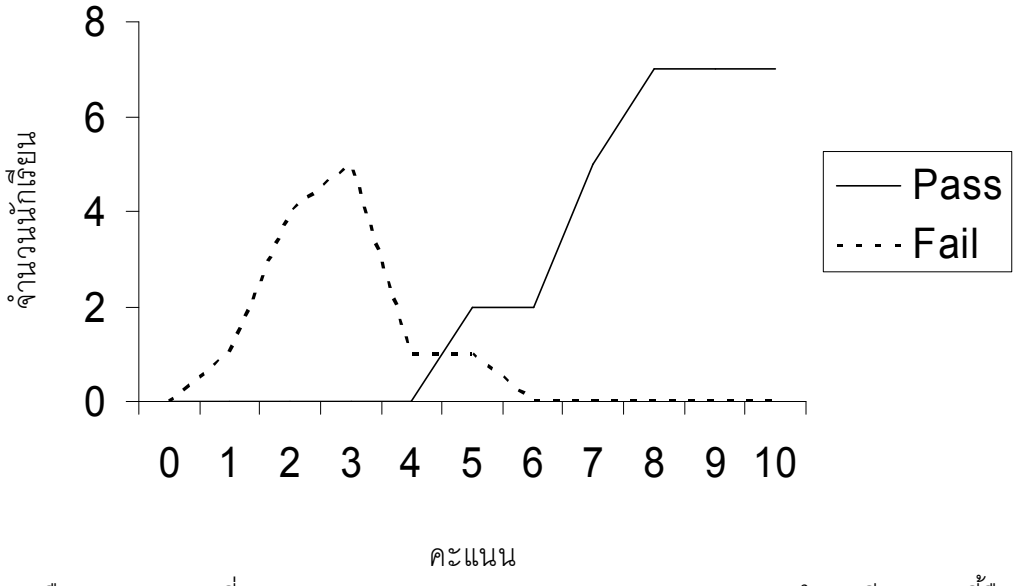
Iramaneerat C. Passing standard: Part III [Thai]. Medical Education Pamphlet 2006; 2(3): 1.

วิธีการตั้งเกณฑ์สอบผ่าน (passing standard) (ตอนที่ 3)

เชิดศักดิ์ ไอรมณีรัตน์

ในบทความนี้ผมจะขอแนะนำวิธีการตั้งเกณฑ์สอบผ่านโดยพิจารณานักเรียนผู้สอบ วิธีการตั้งเกณฑ์ผ่านแบบนี้เหมาะสำหรับการสอบวัดทักษะ การสอบสัมภาษณ์ หรือการประเมินการปฏิบัติงาน ซึ่งมักตัดสินการสอบผ่านโดยดูจากความสามารถของผู้สอบโดยรวมได้ง่ายกว่าดูจากคะแนนที่ได้ในหัวข้อประเมินแต่ละข้อ วิธีการตั้งเกณฑ์ผ่านลักษณะนี้ที่ใช้บ่อยมีด้วยกัน 2 วิธีคือ

1. Borderline-group method: การตั้งเกณฑ์ผ่านวิธีนี้เริ่มจากให้คณะกรรมการสอบประชุมตกลงกันก่อนถึงลักษณะของผู้สอบที่อยู่ในกลุ่มคาบเส้น (ผู้สอบที่มีความรู้ไม่มากพอที่อาจารย์จะให้สอบผ่านได้อย่างสบายใจ แต่ก็มีความรู้ไม่น้อยจนอาจารย์สามารถตัดสินให้สอบตกได้โดยไม่มีข้อสงสัย) หลังจากนั้นอาจารย์พิจารณาความสามารถโดยรวมของผู้สอบแต่ละคน (โดยไม่ทราบคะแนนที่ผู้สอบคนนั้นได้รับ) แล้วระบุว่าผู้สอบคนใดจัดว่ามีความสามารถอยู่ในเกณฑ์ "คาบเส้น" เมื่อระบุว่าผู้สอบคนใดบ้างจัดว่ามีความสามารถคาบเส้นแล้วให้ตั้งเกณฑ์สอบผ่านที่คะแนน median ของผู้สอบกลุ่มนี้ (ไม่แนะนำให้ใช้ค่าเฉลี่ย (mean) เนื่องจากเกณฑ์ผ่านจะเบี่ยงเบนได้มากหากมีคะแนนที่สูงหรือต่ำมากเข้ามาร่วมในการคำนวณ)
2. Contrasting groups method: การตั้งเกณฑ์ผ่านวิธีนี้เริ่มจากการระบุลักษณะของผู้สอบที่ควรสอบผ่าน และ ผู้ที่ควรสอบตก หลังจากนั้นให้อาจารย์พิจารณาความสามารถของผู้สอบทีละคน (โดยไม่ทราบคะแนนที่ผู้สอบคนนั้นได้รับ) แล้วระบุว่าผู้สอบคนนั้นควรอยู่ในกลุ่ม "สอบผ่าน" หรือ "สอบตก" หลังจากนั้นให้ทำการวาดกราฟแสดงความสัมพันธ์ระหว่างจำนวนนักเรียนที่ถูกจัดให้สอบผ่าน และ สอบตก กับคะแนนที่นักเรียนได้รับ ดังตัวอย่างข้างล่าง



เกณฑ์ผ่านคือคะแนน ณ จุดที่ false positive และ false negative passing เท่ากัน (ในกรณีตัวอย่างนี้คือ 5 คะแนน) (คณะกรรมการตั้งเกณฑ์ผ่านอาจปรับเกณฑ์ผ่านได้เพื่อปรับอัตรา false positive และ false negative passing ได้ตามวัตถุประสงค์ของการสอบ)

23 June 2017

หัวข้อ : การจัดทำคลังข้อสอบ

การจัดทำคลังข้อสอบ
Item banking

ผศ. พญ. อนัญญา พงษ์ไพบูลย์
คุณทศพร มาสวัสดิ์

Items – Item bank

Items

- The fundamental block that contains one set of question and responses
- It is the smallest exchangeable object in a test

Item bank

- A collection of test items that may be easily accessed for use in preparing examinations

Functions of item banking

1. Item entry and storage (ใส่ เก็บ)
2. Item retrieval (ดึง)
3. Maintenance of item history (บันทึกประวัติ)

Options for Setting Up an Item Bank

1. **Purchase** of a bank of items along with the software to operate it
2. **Purchase** item banking software with local **development** of items
3. Local **development** of software and items

Key Issues in Item Banking

- Security
- Flexibility
- Retrievability
- **Storage of pertinent information**

Flexibility

- Storage - Sort (เก็บ - เรียง)
- Search based on multiple variables (สืบค้นได้หลากหลายรูปแบบ)
- Adaptable to the needs of test developers (ปรับปรุงพัฒนา)

Security

- Three security threats
  - Theft (ขโมย)
  - Disappearance (ข้อสอบสูญหาย)
  - Exposure (ข้อสอบถูกใช้งาน)
- Two security measures
  - Access control
  - Encryption

Vale CD. Computerized item banking. In: Downing SM, Haladyna TM. Handbook of test development, Mahwah, NJ: Lawrence Earlbaum Associates 2006.

Access Control

- Physical control
  - Secure areas with locked doors (+/- guards)
- Logical control
  - Software systems restricting access to those with no authorization
  - Three approaches
    - Secrets: password, PIN
    - Tokens: card or electronic key
    - Biometrics: voiceprint, fingerprint, retinal scan

M-Tech. Password management best practices. Calgary, Alberta: M-Tech information technology 2003.

Encryption

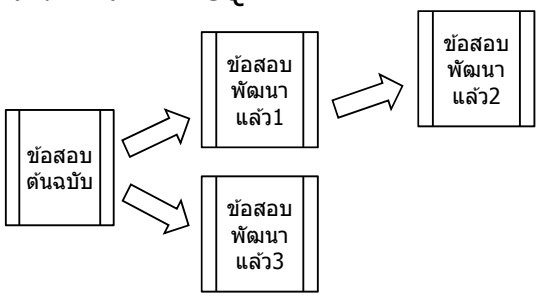
- An item bank with plaintext item files is like a locked gate without a fence.
- Item database must be encrypted using one of the modern encryption standards.
- Encrypt the bank so that only privileged test developers can decrypt it.

Vale CD. Computerized item banking. In: Downing SM, Haladyna TM. Handbook of test development, Mahwah, NJ: Lawrence Earlbaum Associates 2006.

โปรแกรมคลังข้อสอบ
คณะแพทยศาสตร์ศิริราชพยาบาล

โปรแกรมคลังข้อสอบ
(Item Banking)

พัฒนาข้อสอบ MCQ



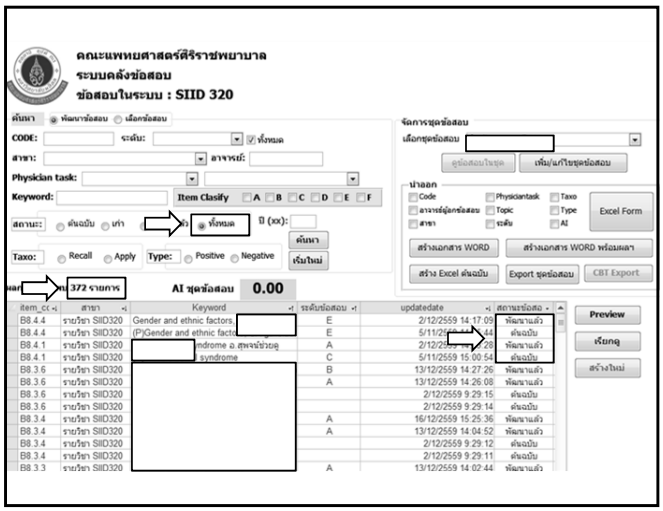
ฝ่ายการศึกษา คณะแพทยศาสตร์ศิริราชพยาบาล

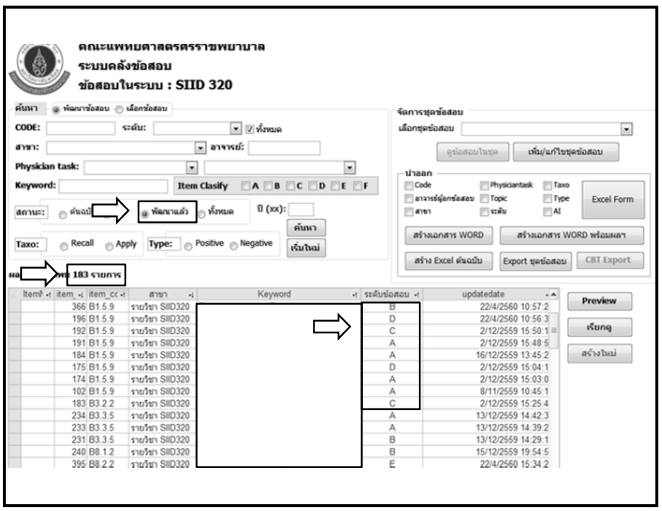
โปรแกรมคลังข้อสอบ
(Item Banking)

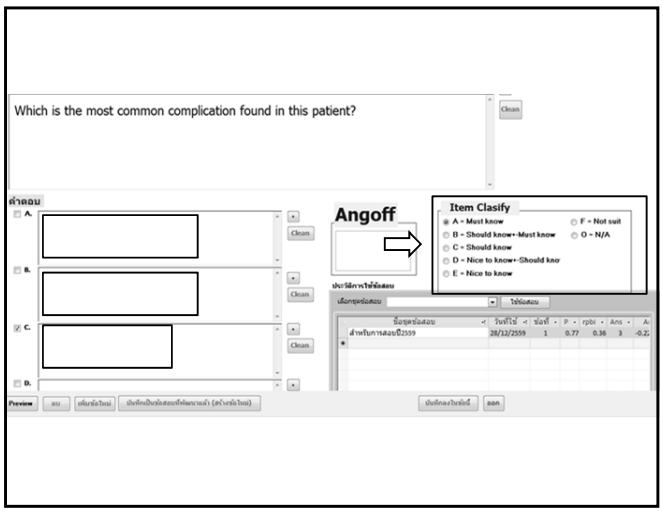
เครื่องคอมพิวเตอร์คลังข้อสอบ

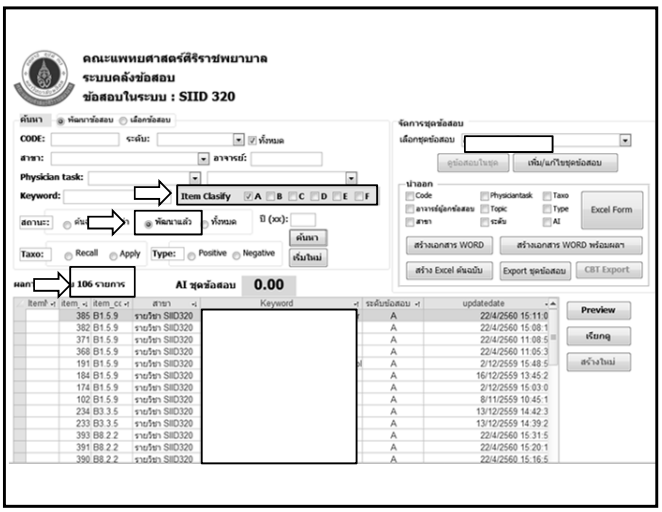
1. ไม่เชื่อมต่อเครือข่าย
2. มีห้องทำข้อสอบโดยเฉพาะ
3. จำกัดจำนวนผู้รู้รหัสเข้าเครื่องและโปรแกรม
4. มีเครื่องสำรองไฟ (UPS)
5. Update Virus สม่ำเสมอ
6. Backup ข้อมูล

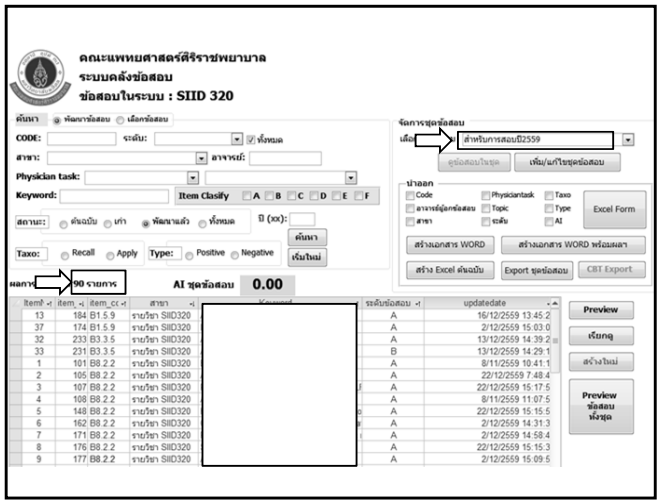
ฝ่ายการศึกษา คณะแพทยศาสตร์ศิริราชพยาบาล

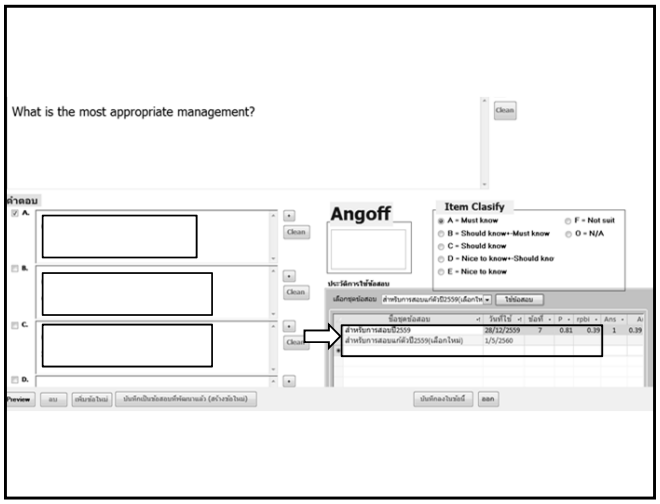

คณะแพทยศาสตร์ศิริราชพยาบาล
ระบบคลังข้อสอบ
ข้อสอบในระบบ : SIID 320
372 รายการ
AI ชุดข้อสอบ 0.00


คณะแพทยศาสตรราชพยาบาล
ระบบคลังข้อสอบ
ข้อสอบในระบบ : SIID 320
183 รายการ


Which is the most common complication found in this patient?
Angoff
Item Clasify


คณะแพทยศาสตร์ศิริราชพยาบาล
ระบบคลังข้อสอบ
ข้อสอบในระบบ : SIID 320
106 รายการ
AI ชุดข้อสอบ 0.00





	พัฒนาแล้ว (152)	เลือกใช้ปี59 (90)
B1.5.9	7	2
B3.3.5	3	2
B8.1.2	1	0
B8.2.2	96	58
B8.2.3	7	5
B8.2.4	16	8
B8.2.5	4	3
B8.2.6	4	4
B8.3.3	11	10
B8.4.1	1	0
B8.4.4	1	0

Total	90
B1.5.9	2
B8.2.2	58
B8.2.3	5
B8.2.4	8
B8.2.5	3
B8.2.6	4
Pharmaco	12

B8.2.2	58
Upper GI	8
Lower GI	19
pancreas	11
Liver	19

Total	90
A-must know	58
B	18
C-should know	13
D	1

รายการข้อสอบที่เลือกเข้าชุด

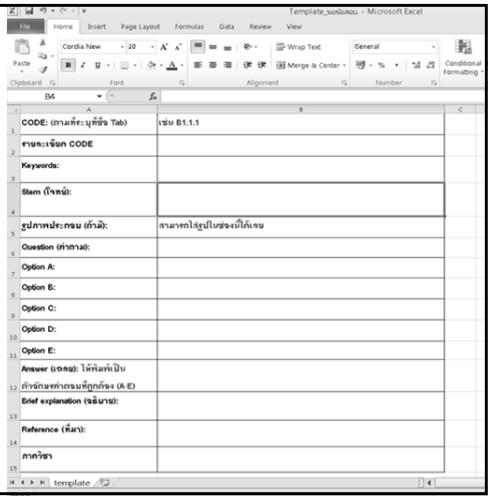
A	B	C
item_code	Keyword	ระดับข้อสอบ
B1.5.9	Aeromonas hydrophilla	A
B1.5.9	Infectious diarrhea	A
	Lab	2
B3.3.5	Antiemetics, food poisoning	A
B3.3.5	Antiemetics, extrapyramidal symptoms	B
B8.3.1	Functional dyspepsia, treatment	A
B8.3.2	colicky abdominal pain, treatment	A
B8.3.2	Diabetic gastroparesis, treatment	B
B8.3.2	GERD, treatment	B
B8.3.3	Variceal bleeding, prevention	A
B8.3.3	NSAIDs-induced PU, management	A
B8.3.3	Stress-related mucosal disease, prevention, gastric secretion	B
B8.3.4	Cholera, treatment	A
B8.3.4	Acute diarrhea, anti-muscarinic drug, treatment	A
B8.3.6	Laxatives, hepatic encephalopathy	A
	Pharmaco	12
B8.2.3	dysphagia, etiology	B
B8.2.3	Intestinal obstruciton, adhesion	A
B8.2.3	Achalasia, pathophysiology	A
B8.2.3	Intussusception, etiology	C
B8.2.3	Hirschsprung disease, pathology	C
	B8.2.3	5
B8.2.4	SAAG, peritoneal metastasis	A
B8.2.4	Colorectal carcinoma, stage	C
B8.2.4	HNPCC, carcinogenesis	B
B8.2.4	FAP, APC gene	A
B8.2.4	Esophageal adenocarcioma	B
B8.2.4	Liver tumor, metastasis	A
B8.2.4	Insulinoma	B
B8.2.4	signet-ring cell carcinoma, diagnosis	D
	B8.2.4	8
B8.2.5	NAFLD	A
B8.2.5	IBS, clinical presentation	A
B8.2.5	Pigment gallstones	B
	B8.2.5	3
B8.2.6	Esophageal/gastric varices, pathophysiology	A
B8.2.6	Mallory-Weiss Tears	C
B8.2.6	Hemorrhoids	A
B8.2.6	Angiodysplasia	C
	B8.2.6	4

ประโยชน์ของการใช้โปรแกรมคลังข้อสอบ

1. เก็บข้อมูลเป็นระบบ → สืบค้น และบริหารได้ง่าย → วิเคราะห์ชุดข้อสอบ (Table of specification) ได้
2. เมื่อจัดเป็นชุดข้อสอบ → สามารถนำไปจัดสอบด้วยระบบ fully computerized ได้

จะเริ่มต้นอย่างไร

ออกข้อสอบ
ใน excel file



กระดาษบันทึก

กระดาษบันทึก

กระดาษบันทึก

กระดาษบันทึก

กระดาษบันทึก

กระดาษบันทึก

Question & Comments

หน่วยพัฒนาแพทยศาสตรศึกษาและวิจัยการศึกษา
ฝ่ายการศึกษา คณะแพทยศาสตร์ศิริราชพยาบาล
สำนักงาน: ตึกอดุลยเดชวิกรม ชั้น 6 (ห้อง 656)
Tel. 02 419 9978 Fax. 02 412 3901

: www.si-merd.com
: merd.project@gmail.com
: MERD
: MERD FC