

สารบัญ

	หน้า
กำหนดการ	1
รายชื่อผู้ร่วมอบรม (แบบแบ่งกลุ่ม)	5
รายชื่อวิทยากร (ช่วงแพทยศาสตรศึกษา)	7
เอกสารประกอบการอบรม	
(Week 1) : 6 – 8 November 2019	9
6 November 2019	11
หัวข้อ : What is curriculum?	13
หัวข้อ : Education objectives and lesson plan	41
หัวข้อ : Learning psychology	79
หัวข้อ : Teaching a large class	95
หัวข้อ : Preparing slides for a class	105
หัวข้อ : Questioning	149
7 November 2019	155
หัวข้อ : Feedback	157
หัวข้อ : Teaching in laboratory	161
หัวข้อ : Clinical supervision	172
หัวข้อ : Small group teaching	209
หัวข้อ : Ambulatory teaching	233
หัวข้อ : Bedside teaching	269
8 November 2019	275
หัวข้อ : How to teach attitudes and ethics	277
หัวข้อ : Simulation	281

สารบัญ

	หน้า
<i>เอกสารประกอบการอบรม</i>	
(Week 2) : 13 – 15 November 2019	289
13 November 2019	291
หัวข้อ : Basic principles of assessment	293
หัวข้อ : MCQ	308
หัวข้อ : MCQ item analysis	327
หัวข้อ : Portfolio	356
14 November 2019	427
หัวข้อ : Essential skills for thesis advisors	429
หัวข้อ : Long case exam	441
หัวข้อ : EPA and WPBA	463
หัวข้อ : Performance assessment	473
หัวข้อ : Standard setting	477
15 November 2019	487
หัวข้อ : Educational project proposal	489
หัวข้อ : Summary	495
ช่องทางการติดต่อสื่อสาร	497

6-8 November 2019

(สัปดาห์ที่ 1) โรงแรมสามพรานริเวอร์ไซด์ จังหวัดนครปฐม

วัน เวลา	จันทร์ 4 Nov	อังคาร 5 Nov	พุธ 6 Nov	พฤหัสบดี 7 Nov		ศุกร์ 8 Nov		วัน เวลา										
				Room 1	Room 2	Room 1	Room 2											
07:00 - 07:15			อาหารเช้า					07:00 - 07:15										
07:15 - 07:30								ลงทะเบียน					07:15 - 07:30					
07:30 - 07:45													ลงทะเบียน					07:30 - 07:45
07:45 - 08:00																		ลงทะเบียน
08:00 - 08:15			What is curriculum? อ.รุ่งนรินทร์, อ.สุพจน์	Reflection อ.กษณา	Professionalism of teachers อ.สุพจน์		08:00 - 08:15											
08:15 - 08:30			Educational objectives and lesson plan อ.สุประพัฒน์	Feedback อ.อนิรุต		How to teach attitudes and ethics อ.สุพจน์		08:15 - 08:30										
08:30 - 08:45				Feedback อ.อนิรุต		How to teach attitudes and ethics อ.สุพจน์		08:30 - 08:45										
08:45 - 09:00								Feedback อ.อนิรุต		How to teach attitudes and ethics อ.สุพจน์		08:45 - 09:00						
09:00 - 09:15				Feedback อ.อนิรุต		How to teach attitudes and ethics อ.สุพจน์						09:00 - 09:15						
09:15 - 09:30			Feedback อ.อนิรุต					How to teach attitudes and ethics อ.สุพจน์		09:15 - 09:30								
09:30 - 09:45					Feedback อ.อนิรุต		How to teach attitudes and ethics อ.สุพจน์			09:30 - 09:45								
09:45 - 10:00			Feedback อ.อนิรุต						How to teach attitudes and ethics อ.สุพจน์		09:45 - 10:00							
10:00 - 10:15					Feedback อ.อนิรุต		How to teach attitudes and ethics อ.สุพจน์				10:00 - 10:15							
10:15 - 10:30			Feedback อ.อนิรุต						How to teach attitudes and ethics อ.สุพจน์		10:15 - 10:30							
10:30 - 10:45					Feedback อ.อนิรุต		How to teach attitudes and ethics อ.สุพจน์				10:30 - 10:45							
10:45 - 11:00			Feedback อ.อนิรุต						How to teach attitudes and ethics อ.สุพจน์		10:45 - 11:00							
11:00 - 11:15					Feedback อ.อนิรุต		How to teach attitudes and ethics อ.สุพจน์				11:00 - 11:15							
11:15 - 11:30			Feedback อ.อนิรุต						How to teach attitudes and ethics อ.สุพจน์		11:15 - 11:30							
11:30 - 11:45					Feedback อ.อนิรุต		How to teach attitudes and ethics อ.สุพจน์				11:30 - 11:45							
11:45 - 12:00			Feedback อ.อนิรุต						How to teach attitudes and ethics อ.สุพจน์		11:45 - 12:00							
12:00 - 12:15					อาหารกลางวัน						12:00 - 12:15							
12:15 - 12:30			อาหารกลางวัน							12:15 - 12:30								
12:30 - 12:45										อาหารกลางวัน					12:30 - 12:45			
12:45 - 13:00															อาหารกลางวัน			
13:00 - 13:15								Teaching a large class อ.เชิดศักดิ์	Teaching on the run อ.ภูมิ									
13:15 - 13:30			Teaching a large class อ.เชิดศักดิ์	Teaching on the run อ.ภูมิ		เดินทางกลับกรุงเทพฯ		13:15 - 13:30										
13:30 - 13:45				Teaching on the run อ.ภูมิ				เดินทางกลับกรุงเทพฯ		13:30 - 13:45								
13:45 - 14:00				Teaching on the run อ.ภูมิ						เดินทางกลับกรุงเทพฯ		13:45 - 14:00						
14:00 - 14:15				Teaching on the run อ.ภูมิ								เดินทางกลับกรุงเทพฯ		14:00 - 14:15				
14:15 - 14:30			Teaching on the run อ.ภูมิ		เดินทางกลับกรุงเทพฯ		14:15 - 14:30											
14:30 - 14:45			Teaching on the run อ.ภูมิ				เดินทางกลับกรุงเทพฯ		14:30 - 14:45									
14:45 - 15:00			Teaching on the run อ.ภูมิ						เดินทางกลับกรุงเทพฯ		14:45 - 15:00							
15:00 - 15:15			Teaching on the run อ.ภูมิ								เดินทางกลับกรุงเทพฯ		15:00 - 15:15					
15:15 - 15:30			Teaching on the run อ.ภูมิ		เดินทางกลับกรุงเทพฯ								15:15 - 15:30					
15:30 - 15:45			Teaching on the run อ.ภูมิ				เดินทางกลับกรุงเทพฯ						15:30 - 15:45					
15:45 - 16:00			Teaching on the run อ.ภูมิ						เดินทางกลับกรุงเทพฯ				15:45 - 16:00					
16:00 - 16:15			พักผ่อนตามอัธยาศัย								16:00 - 16:15							
16:15 - 16:30											พักผ่อนตามอัธยาศัย					16:15 - 16:30		
16:30 - 16:45								พักผ่อนตามอัธยาศัย								16:30 - 16:45		
16:45 - 17:00																พักผ่อนตามอัธยาศัย		
17:00 - 17:15			พักผ่อนตามอัธยาศัย															
17:15 - 17:30													พักผ่อนตามอัธยาศัย					
17:30 - 17:45								พักผ่อนตามอัธยาศัย										
17:45 - 18:00																		พักผ่อนตามอัธยาศัย
18:00 - 18:15			พักผ่อนตามอัธยาศัย															
18:15 - 18:30													พักผ่อนตามอัธยาศัย					
18:30 - 18:45								พักผ่อนตามอัธยาศัย										
18:45 - 19:00																		พักผ่อนตามอัธยาศัย
19:00 - 19:15			พักผ่อนตามอัธยาศัย															
19:15 - 19:30													พักผ่อนตามอัธยาศัย					
19:30 - 19:45								พักผ่อนตามอัธยาศัย										
19:45 - 20:00																		พักผ่อนตามอัธยาศัย

หมายถึง ช่วงเวลาพัก

13-15 November 2019
(สัปดาห์ที่ 2) โรงแรมสามพรานริเวอร์ไซด์ จังหวัดนครปฐม

วัน เวลา	จันทร์ 11 Nov	อังคาร 12 Nov	พุธ 13 Nov	พฤหัสบดี 14 Nov		ศุกร์ 15 Nov	วัน เวลา
				Room 1	Room 2		
07:00 - 07:15			อาหารเช้า				07:00 - 07:15
07:15 - 07:30			อาหารเช้า				07:15 - 07:30
07:30 - 07:45			อาหารเช้า				07:30 - 07:45
07:45 - 08:00			อาหารเช้า				07:45 - 08:00
08:00 - 08:15			Basic principles of assessment อ.เชิดศักดิ์	Essential skills for thesis advisors	Long case exam	QA in education	08:00 - 08:15
08:15 - 08:30				อ.เพทาย, อ.ถาวรชัย, อ.บดินทร์	อ.พรพรรณ	อ.รุ่งนิรันดร์	08:15 - 08:30
08:30 - 08:45							08:30 - 08:45
08:45 - 09:00							08:45 - 09:00
09:00 - 09:15			MCQ		OSCE item development	Educational project proposal	09:00 - 09:15
09:15 - 09:30			อ.พรพรรณ		อ.สุประพัฒน์, อ.อนิรุต	อ.เชิดศักดิ์, อ.พรพรรณ	09:15 - 09:30
09:30 - 09:45						อ.ธวัชวรรณ, อ.ภูมิ	09:30 - 09:45
09:45 - 10:00						อ.สุประพัฒน์, อ.อนิรุต	09:45 - 10:00
10:00 - 10:15			พักรับประทานอาหารว่าง	พักรับประทานอาหารว่าง	พักรับประทานอาหารว่าง	พักรับประทานอาหารว่าง	10:00 - 10:15
10:15 - 10:30			MCQ item review	Essential skills for thesis advisors	OSCE item review		10:15 - 10:30
10:30 - 10:45			อ.สุพจน์, อ.พรพรรณ, อ.เชิดศักดิ์,	(continued)	อ.พรพรรณ, อ.เชิดศักดิ์		10:30 - 10:45
10:45 - 11:00			อ.อนิรุต, อ.สุประพัฒน์, อ.กษณา	อ.เพทาย, อ.ถาวรชัย, อ.บดินทร์	อ.ตรีภพ, อ.สุประพัฒน์, อ.อนิรุต		10:45 - 11:00
11:00 - 11:15			MCQ item analysis				11:00 - 11:15
11:15 - 11:30			อ.เชิดศักดิ์				11:15 - 11:30
11:30 - 11:45						Summary	11:30 - 11:45
11:45 - 12:00						อ.เชิดศักดิ์	11:45 - 12:00
12:00 - 12:15			อาหารกลางวัน				12:00 - 12:15
12:15 - 12:30			อาหารกลางวัน				12:15 - 12:30
12:30 - 12:45			อาหารกลางวัน				12:30 - 12:45
12:45 - 13:00			อาหารกลางวัน				12:45 - 13:00
13:00 - 13:15			Constructed response item exam	Essential skills for thesis advisors	EPA and WPBA		13:00 - 13:15
13:15 - 13:30			อ.สุประพัฒน์	(continued)	อ.เชิดศักดิ์, อ.อนิรุต		13:15 - 13:30
13:30 - 13:45				อ.เพทาย, อ.ถาวรชัย, อ.บดินทร์			13:30 - 13:45
13:45 - 14:00			Constructed response item review				13:45 - 14:00
14:00 - 14:15			อ.สุพจน์, อ.พรพรรณ, อ.เชิดศักดิ์				14:00 - 14:15
14:15 - 14:30			อ.อนิรุต, อ.สุประพัฒน์, อ.กษณา	Performance assessment			14:15 - 14:30
14:30 - 14:45				อ.เชิดศักดิ์		พักรับประทานอาหารว่าง	14:30 - 14:45
14:45 - 15:00			พักรับประทานอาหารว่าง	พักรับประทานอาหารว่าง			14:45 - 15:00
15:00 - 15:15			Portfolio	Standard setting		เดินทางกลับกรุงเทพฯ	15:00 - 15:15
15:15 - 15:30			อ.ตรีภพ	อ.เชิดศักดิ์			15:15 - 15:30
15:30 - 15:45							15:30 - 15:45
15:45 - 16:00							15:45 - 16:00
16:00 - 16:15			พักผ่อนตามอัธยาศัย				16:00 - 16:15
16:15 - 16:30			พักผ่อนตามอัธยาศัย				16:15 - 16:30
16:30 - 16:45			พักผ่อนตามอัธยาศัย				16:30 - 16:45
16:45 - 17:00			พักผ่อนตามอัธยาศัย				16:45 - 17:00
17:00 - 17:15			พักผ่อนตามอัธยาศัย				17:00 - 17:15
17:15 - 17:30			พักผ่อนตามอัธยาศัย				17:15 - 17:30
17:30 - 17:45			พักผ่อนตามอัธยาศัย			17:30 - 17:45	
17:45 - 18:00			พักผ่อนตามอัธยาศัย			17:45 - 18:00	
18:00 - 18:15			อาหารเย็น			18:00 - 18:15	
18:15 - 18:30			อาหารเย็น			18:15 - 18:30	
18:30 - 18:45			อาหารเย็น			18:30 - 18:45	
18:45 - 19:00			อาหารเย็น			18:45 - 19:00	
19:00 - 19:15			พักผ่อนตามอัธยาศัย			19:00 - 19:15	
19:15 - 19:30			พักผ่อนตามอัธยาศัย			19:15 - 19:30	
19:30 - 19:45			พักผ่อนตามอัธยาศัย			19:30 - 19:45	
19:45 - 20:00			พักผ่อนตามอัธยาศัย			19:45 - 20:00	

หมายถึง ช่วงเวลาพัก

การอบรมหลักสูตรสู่ความเป็นเลิศของอาจารย์แพทย์ศิริราช รุ่นที่ 12 (ประจำปีงบประมาณ 2562)

วันที่ 4-22 พฤศจิกายน 2562 ณ โรงแรมสามพานริเวอร์ไซด์ จังหวัดนครปฐม และในคณะแพทยศาสตร์ศิริราชพยาบาล

ลำดับที่	SAP ID	ชื่อ-สกุล	ภาควิชา/หน่วยงาน	กลุ่ม	กลุ่ม
1	10020312	นพ.อดิสร รัตนโยธา	กายวิภาคศาสตร์	1	1
2	10030609	นพ.ภาคภูมิ ปิยะมาน	กายวิภาคศาสตร์	1	
3	10032479	อ. พญ.อนุช ดุรงค์พันธุ์	กายวิภาคศาสตร์	1	
4	10014312	อ.ดร. พญ.วิชชุดา กมลวิทย์	จุลชีววิทยา	1	
5	10028905	พญ.กวิสรา รัตนธรรมวัฒน์	พยาธิวิทยาคลินิก	1	
6	10010894	อ. นพ.เมธา ใหญ่กว่าวงศ์	เภสัชวิทยา	1	
7	10033867	นพ. ฆนัท จันทรทองดี	สรีรวิทยา	1	
8	10018599	พญ.ทักษอร กิตติภัสสร	สรีรวิทยา	1	
9	10025254	อ. นพ.วรธีร์ เดชารักษ์	โรงเรียนกายอุปกรณ์สิรินธร	2	2
10	10022175	อ.ศิริรัตน์ เส็งเอียด	โรงเรียนกายอุปกรณ์สิรินธร	2	
11	10020579	น.ส.สาวิตรี ศรีทำบุญ	โรงเรียนกายอุปกรณ์สิรินธร	2	
12	10034278	พญ.ไพรินทร์ เลหาสินณรงค์	เวชศาสตร์ฟื้นฟู	2	
13	10034279	พญ.ปณิติกา ปราชญ์โกสินทร์	เวชศาสตร์ฟื้นฟู	2	
14	10023214	พญ.จิตภาภา ว่องเจริญวัฒนา	ศัลยศาสตร์ออร์โธปิดิกส์และกายภาพบำบัด	2	
15	10033810	นพ.ภาคภูมิ เรืองสมบูรณ์	ศัลยศาสตร์ออร์โธปิดิกส์และกายภาพบำบัด	2	
16	10036639	ดร.นิษฐา พุฒิमानรติกุล	ฝ่ายการศึกษา	3	3
17	10022250	นายวิศิษฐ์ สมบัติถาวรกุล	สถานเทคโนโลยีการศึกษาแพทยศาสตร์	3	
18	10034386	นายสถาปนา นิยะโมสถ	สถานเทคโนโลยีการศึกษาแพทยศาสตร์	3	
19	10036943	อ. นพ.สมบูรณ์ หทัยอยู่สุข	จิตเวชศาสตร์	3	
20	10034400	พญ.กัรติ พัฒนเสรี	จิตเวชศาสตร์	3	
21	10037058	พญ.กัณฑ์กมล จัยสิน	จิตเวชศาสตร์	3	
22	10037057	พญ.จุฑาวดี หล่อตระกูล	จิตเวชศาสตร์	3	

ลำดับที่	SAP ID	ชื่อ-สกุล	ภาควิชา/หน่วยงาน	กลุ่ม	กลุ่ม
23	10035042	พญ.นิดา วงศ์ชัยสุวัฒน์	จักษุวิทยา	4	4
24	10028979	อ. นพ.พัศยุ อยู่เย็น	วิสัญญีวิทยา	4	
25	10028953	อ. นพ.ภานพ หลิมรัตน์	วิสัญญีวิทยา	4	
26	10030571	พญ.ภัทราพรรณ วงศ์ศรีภูมิเทศ	วิสัญญีวิทยา	4	
27	10023561	ดร. พญ.วรรณนันทน์ มะกรสาร	วิสัญญีวิทยา	4	
28	10026839	ดร. นพ.ภาวิต สมนึก	วิสัญญีวิทยา	4	
29	10036836	พญ.ณิชา ศรีสุวรรณนท์	ศัลยศาสตร์	4	
30	10032666	พญ.กัณทรिता ศรีพานิชกุลชัย	เวชศาสตร์ป้องกันและสังคม	5	5
31	10028943	นพ.สิทธิ ลิ้อฤทธิพงษ์	เวชศาสตร์ป้องกันและสังคม	5	
32	10023512	นพ.หฤษฎ์ ปิ่นณะรัส	เวชศาสตร์ป้องกันและสังคม	5	
33	10016212	อ. ดร.ณัชร ล้ำเลิศกิจ	สถานการแพทย์แผนไทยประยุกต์	5	
34	10025215	นพ.ภัทรพันธุ์ เลิศฤทธิวิมานแมน	อายุรศาสตร์	5	
35	10032755	อ. พญ.ชิตาภา กาวีต๊ะ	อายุรศาสตร์	5	
36	10031116	พญ.ยุวดี พิทักษ์ปฐพี	อายุรศาสตร์	5	
37	10020630	พญ.ชญามน ทักษะประดิษฐ์	กุมารเวชศาสตร์	6	6
38	10032598	อ. นพ.ธีร กิจมาตรสสุวรรณ	กุมารเวชศาสตร์	6	
39	10035205	พญ.ชญาดา ชัยบุตร	ตจวิทยา	6	
40	10037012	พญ.ซุดา รุจิธารณวงศ์	ตจวิทยา	6	
41	10034271	นพ.กฤตชาติ บุตรเนียร	รังสีวิทยา	6	
42	10034679	พญ.ศุลยพร วงศ์แสวง	รังสีวิทยา	6	
43	10028909	นพ.ธีรดนัย ตรีชัยรัมย์	รังสีวิทยา	6	
44	10036772	นพ.วศิน เอียบสกุล	รังสีวิทยา	6	

ใบรายชื่อวิทยากรโครงการสู่ความเป็นเลิศของอาจารย์แพทย์ศิริราช (รุ่นที่ 12)

ช่วงแพทยศาสตรศึกษา (Medical Education)

ลำดับ	ชื่อ - สกุล	ภาควิชา / หน่วยงาน
1	รศ. นพ.รุ่งนรินทร์ ประดิษฐ์สุวรรณ	รองคณบดีฝ่ายการศึกษา ก่อนปริญญา
2	ศ.ดร. เพทาย เย็นจิตโสมนัส	สถานส่งเสริมการวิจัย
3	รศ.ดร. นพ.ถาวรชัย ลิ้มจินดาพร	รองคณบดีฝ่ายการศึกษา หลังปริญญา
4	รศ. นพ.สุพจน์ พงศ์ประสพชัย	อายุรศาสตร์
5	ผศ.ดร. พญ.วัฒนา วัฒนาภา	สรีรวิทยา
6	ผศ. นพ.ยิ่งยง ชินธรรมมิตร	อายุรศาสตร์
7	รศ. พญ.พรพรรณ กุ้มานะชัย	อายุรศาสตร์
8	อ.ดร. นพ.บดีนทร์ ทรัพย์สมบูรณ์	สรีรวิทยา
9	รศ.ดร. นพ.เชิดศักดิ์ ไอรมนรัตน์	ศัลยศาสตร์
10	ผศ. นพ.สุประพัฒน์ สนใจพาณิชย์	กุมารเวชศาสตร์
11	ผศ. นพ.ตรีภพ เลิศบรรณพงษ์	สูติศาสตร์-นรีเวชวิทยา
12	ผศ. พญ.กษณา รักษาภรณ์	วิสัญญีวิทยา
13	ผศ. พญ.ธัชวรรณ จิระติวานนท์	วิสัญญีวิทยา
14	อ. นพ.ภูมิ ตรีตระการ	วิสัญญีวิทยา
15	อ.นพ.อนิรุต วรวาท	นิติเวชศาสตร์
16	อ.ดร. นพ.พรนพ นัยเนตร	วิทยาภูมิคุ้มกัน

โครงการสู่ความเป็นเลิศของอาจารย์แพทย์ศิริราช (รุ่นที่ 12)
(สัปดาห์ที่ 1) วันที่ 6 - 8 พฤศจิกายน พ.ศ. 2562
ณ โรงแรมสามพรานริเวอร์ไซด์ จังหวัดนครปฐม

เอกสารประกอบการอบรม

The logo for the TEACH program features the word "TEACH" in large, bold, dark blue capital letters with a white outline. The text is centered against a light blue background that includes a white ECG (heart rate) line. The background also has a subtle pattern of light blue and white squares, and some faint, glowing white circles at the bottom.

TEACH

6 November 2019

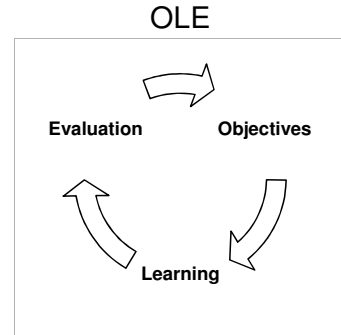
New staff workshop, Faculty of Medicine
Siriraj Hospital

Nov 2019

โครงการสู่ความเป็นเลิศของอาจารย์แพทย์ศิริราชรุ่น 12

- แพทยศาสตรศึกษา
- 6 - 8 พฤศจิกายน 2561 และ 13 - 15 พฤศจิกายน 2562
- วัตถุประสงค์
 - ภาพรวมของหลักสูตร
 - วัตถุประสงค์การเรียนรู้
 - วิธีการจัดประสบการณ์การเรียนรู้
 - การประเมินผล

1



2

หลักสูตรการอบรม

- Objectives
 - Knowledge
 - Skills
 - Attitudes
- Learning experience
 - Interactive workshop
 - Group activities
- Evaluation
 - Attendance
 - Education projects

3

Group Scoring

- สมาชิกกลุ่มมาตรฐานเป็นกลุ่มแรก รอบเช้า / บ่าย 20
- สมาชิกกลุ่มมาตรฐานเป็นกลุ่มที่สอง รอบเช้า / บ่าย 10
- ขวัญใจอาจารย์ผู้สอน แต่ละคาบ 10
- รวมคะแนน รายวัน ก่อนเริ่มเรียนตอนเช้าของแต่ละวัน
- รางวัลใหญ่ รวมคะแนนจนถึงคาบสุดท้ายวันที่ 15 พย 62
- กลุ่มรวมคะแนน เรียงลำดับ กลุ่ม 1,2,3,4,5,6 ตามลำดับ

4

Questions ?

เชิดศักดิ์
090-9905757
Line ID: chersak_si
5

chersakiramanerat@gmail.com

1

The Curriculum

รศ. นพ.รุ่งนรินทร์ ประดิษฐ์สุวรรณ
รองคณบดีฝ่ายการศึกษาก่อนปริญญา

“Curriculum”

All the planned **learning experiences** of a school or educational institution

BMJ 2003;326:268

หลักสูตร

รายวิชา

การสอนรายคาบ/หัวข้อ

OLE

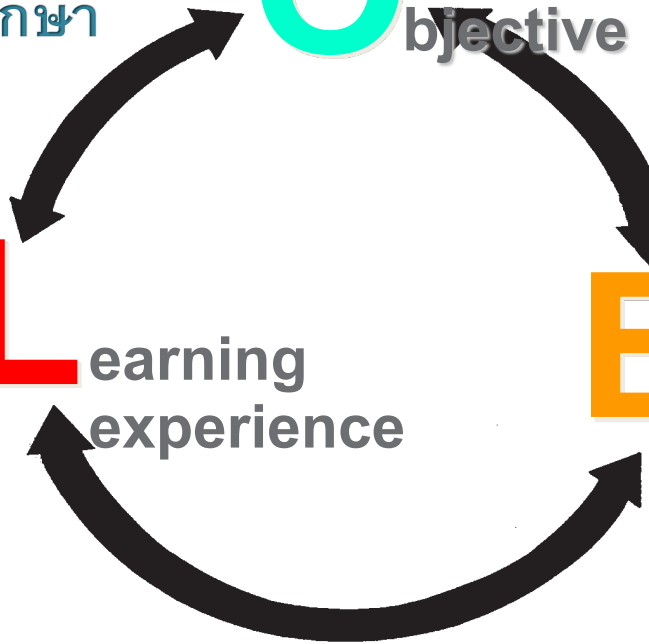
องค์ประกอบหลักใน
การจัดการศึกษา

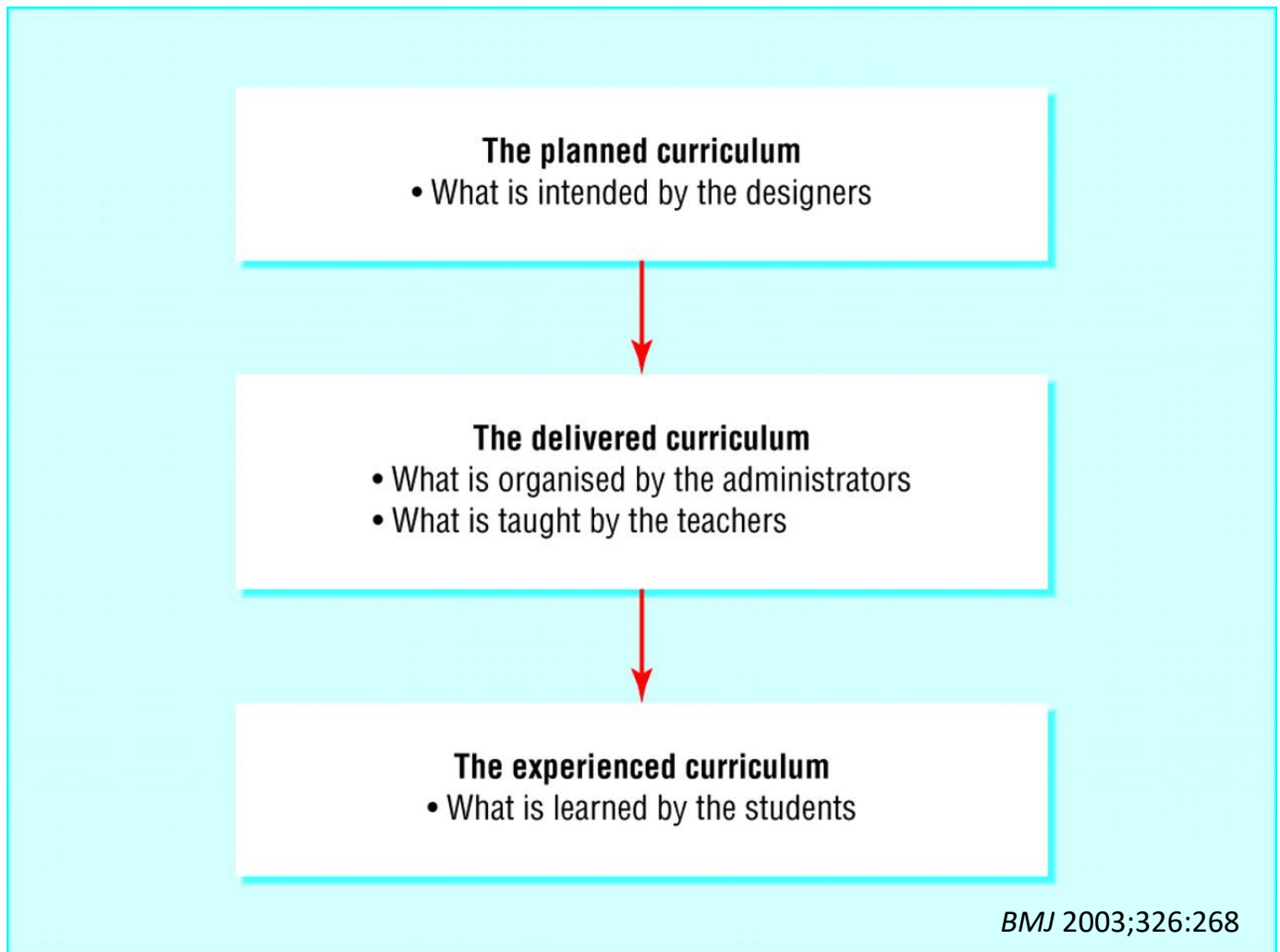
O สอน/เรียนรู้อะไร
Objective

สอน/เรียนรู้
อย่างไร

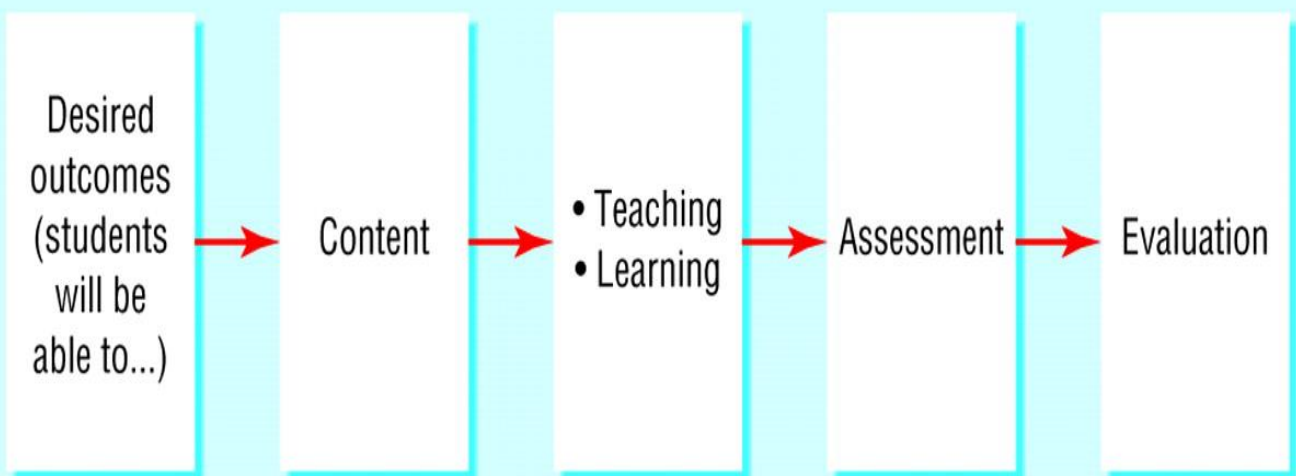
Learning
experience

สอน/เรียน
ได้ผลหรือไม่
Evaluation

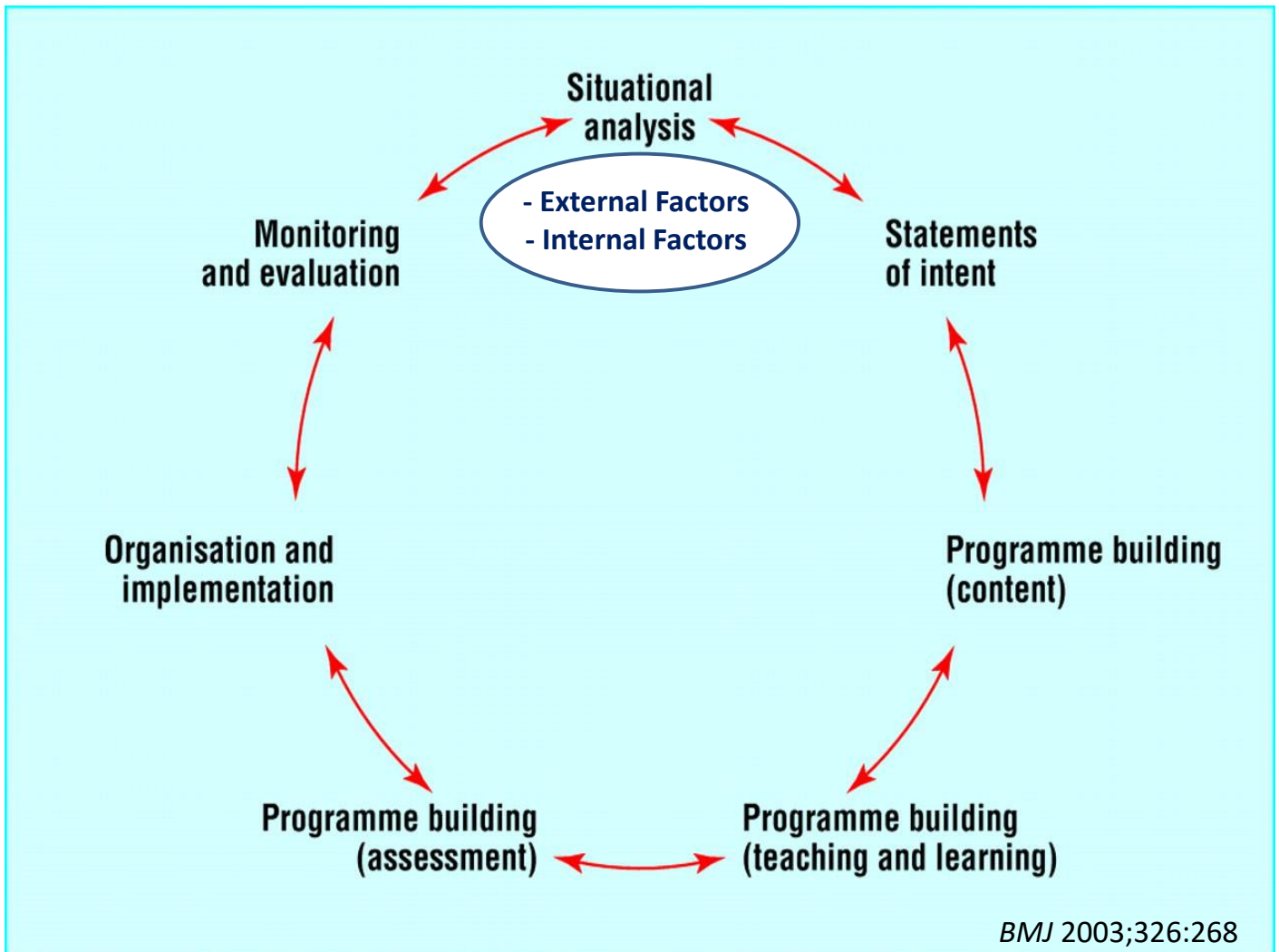




Outcome-based Curriculum



BMJ 2003;326:268



หลักสูตรแพทยศาสตรบัณฑิต คณะแพทยศาสตร์ศิริราชพยาบาล ปรับปรุง พ.ศ. 2557

21st Century Learners



- *Information, media and technology savvy*
- *Insist that they are real multi-tasker*

21st Century Learners

SHORT attention span



- ชอบการเรียนรู้
 - ผ่านสื่อที่เคลื่อนไหว จับต้องได้ หรือมีปฏิสัมพันธ์กับผู้เรียน
 - โดยลองปฏิบัติ
 - ชอบพิสูจน์สมมุติฐาน
 - สิ่งที่เชื่อว่ามีประโยชน์กับตน
 - คู่กับความสนุก
 - แบบทำงานเป็นทีม
- กล้าแสดงออก อยากให้คนอื่นชื่นชม

21st Century Skills

1. **LEARNING and INNOVATION SKILLS**

- *Creativity and innovation*
- *Critical thinking and problem solving*
- *Communication and Collaboration*

2. **INFORMATION, MEDIA and TECHNOLOGY SKILLS**

3. **LIFE and CAREER SKILLS**

- *Flexibility and adaptability*
- *Initiative and self-direction*
- *Social and cross-cultural skills*
- *Productivity and accountability*
- *Leadership and responsibility*

Medical Education Strategies

SPICES Model

S	Student-centred	←	→	Teacher-centred
P	Problem-based	←	→	Information gathering
I	Integrated	←	→	Discipline-based
C	Community-based	←	→	Hospital-based
E	Electives	←	→	Standard programme
S	Systematic	←	→	Apprenticeship-based/ Opportunistic

Image from www.kjme.kr/m/journal/view.php?number=300



WFME

**WORLD FEDERATION
FOR
MEDICAL EDUCATION**

หลักสูตรแพทยศาสตรบัณฑิต ปรับปรุง พ.ศ. 2557

Outcome-based Curriculum

คุณลักษณะที่พึงประสงค์ของบัณฑิตแพทย์ศิริราช

SKILLS

- **Soul** (*professional habits, attitudes, moral, and ethics*)
- **Knowledge** (*professional knowledge*)
- **Information** (*information, media, technology literacy*)
- **Learning** (*learning and innovation skills: Creativity, Critical thinking, Communication, Collaboration*)
- **Leadership** (*leadership and life skills*)
- **Skills** (*professional skills*)

Siriraj's curriculum



แผนการศึกษาชั้นปีที่ 1 (41 หน่วยกิต)

ภาคการศึกษาที่ 1

	หน่วยกิต
1. แคลคูลัสและระบบสมการเชิงอนุพันธ์สามัญ	3
2. เคมีทั่วไป	3
3. ปฏิบัติการชีววิทยา 1	1
4. ชีววิทยาสาระสำคัญ	2
5. ปฏิบัติการฟิสิกส์ทั่วไป	1
6. ฟิสิกส์พื้นฐานสำหรับวิทยาศาสตร์การแพทย์	2
7. ภาษาอังกฤษ ระดับ 3	3
รวม	15

ภาคการศึกษาที่ 2

	หน่วยกิต
1. สถิติศาสตร์สำหรับวิทยาศาสตร์การแพทย์	2
2. เคมีอินทรีย์	3
3. ฟิสิกส์สำหรับวิทยาศาสตร์การแพทย์	3
4. ภาษาอังกฤษ ระดับ 4	3
5. การศึกษาวิชาแพทย์	1
6. วิชาชีพแพทย์	1
7. วิชาเลือกเสรี	2
รวม	15

รายวิชาที่จัดต่อเนื่องทั้งสองภาคการศึกษาหรือการศึกษาไปสิ้นสุดในภาคปลาย

	หน่วยกิต
1. ปฏิบัติการเคมี	1
2. ศิลปะการใช้ภาษาไทยเพื่อการสื่อสาร	3
3. การศึกษาทั่วไปเพื่อพัฒนามนุษย์	2
4. สังคมศึกษาเพื่อพัฒนามนุษย์	3
5. ศิลปะวิทยาการเพื่อพัฒนามนุษย์	2
รวม	11 หน่วยกิต

Siriraj's curriculum



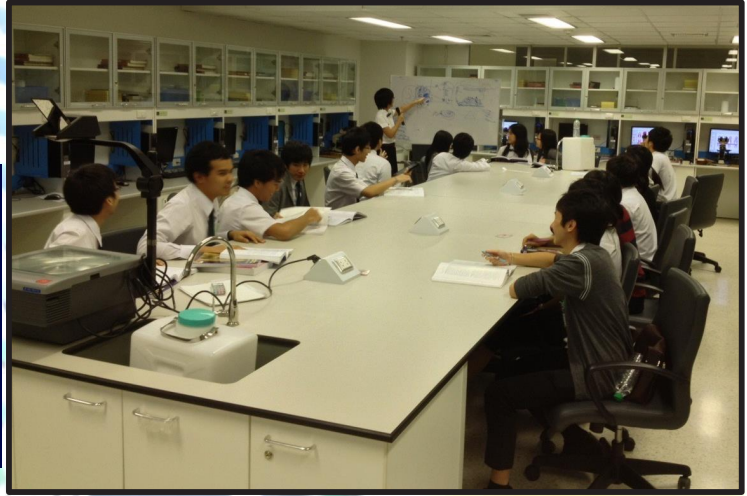
2-3

Foundational Sciences

ขอบคุณ อ. ดร. นพ. ยอดยิ่ง แดงประไพ สำหรับภาพประกอบ

Active Learning

เน้นการมีส่วนร่วมของนักเรียน
เน้นกระบวนการคิด
การค้นหา การจัดการข้อมูล
เพื่อหาคำตอบที่เหมาะสม



**Integration
&
Early Clinical Exposure**



แผนการศึกษาชั้นปีที่ 2 (40 หน่วยกิต)

ภาคการศึกษาที่ 1

	หน่วยกิต
1. Foundation of the Human Body: From Molecules to Body	3
2. Foundation of the Human Body: Neural and Hormonal Regulation	2
3. Foundation of the Human Body: Energy and Metabolism	2
4. Integumentary and Musculoskeletal System I	2
5. Blood and Lymphoid System I	1
6. Gross Anatomy Laboratory I	3
7. Electives	2
รวม	15

ภาคการศึกษาที่ 2

	หน่วยกิต
1. Circulatory System I	2
2. Respiratory System I	2
3. Gastrointestinal and Hepatobiliary System I	2
4. Urinary System I	2
5. Reproductive System I	2
6. Nervous System I	4
7. Gross Anatomy Laboratory II	2
รวม	16

รายวิชาที่จัดต่อเนื่องทั้งสองภาคการศึกษาหรือการศึกษาไปสิ้นสุดในภาคปลาย

	หน่วยกิต	
1. The Human Life	2	}
2. Applied Preclinical Knowledge	1	
3. Humanistic Medicine	1	
4. Epidemiology and Biostatistics in Public Health Practice	1	
5. Life and Social Skills	2	
6. Critical Review of Contemporary Thai Society	2	
รวม		9 หน่วยกิต

แผนการศึกษาชั้นปีที่ 3 (40 หน่วยกิต)

ภาคการศึกษาที่ 1

	หน่วยกิต
1. Immune Responses and Inflammation	3
2. Principles of Microbiology and Parasitology	4
3. Principles of Genetics and Neoplasia	1
4. Integumentary and Musculoskeletal System II	2
5. Circulatory System II	2
6. Blood and Lymphoid System II	3
รวม	15

ภาคการศึกษาที่ 2

	หน่วยกิต
1. Respiratory System II	2
2. Gastrointestinal and Hepatobiliary System II	2
3. Urinary System II	2
4. Reproductive System II	2
5. Nervous System II	3
6. Disorders of Endocrine and Multi-organ Systems	2
7. Symptomatology	2
8. Electives	2
รวม	17

รายวิชาที่จัดต่อเนื่องทั้งสองภาคการศึกษาหรือการศึกษาไปสิ้นสุดในภาคปลาย

	หน่วยกิต
1. Principles of Therapeutic Medicine	2
2. Preventive Medicine and Health Promotion	2
3. Evidence-based Medicine	1
4. Medical Ethics and Laws	1
5. Society and Health	2
รวม	8 หน่วยกิต

Siriraj's curriculum



4-6 Clinical Experiences

ขอบคุณ อ. ดร.นพ.ยอดยิ่ง แดงประไพ สำหรับภาพประกอบ

แผนการศึกษา ปี 4 (43 หน่วยกิต)

พร้อมกันทั้งชั้นปี (5 สัปดาห์)	สัปดาห์	หน่วยกิต
1. ทักษะพื้นฐานทางคลินิก	2	2
2. การแก้ปัญหาทางคลินิก	3	2
หมุนเวียนเป็นกลุ่ม (8 สัปดาห์)		
1. อายุรศาสตร์ 1	8	8
2. ศัลยศาสตร์ 1	8	8
3. สูติศาสตร์และนรีเวชวิทยา 1	4	4
4. กุมารเวชศาสตร์ 1	4	4
5. เวชศาสตร์ป้องกันและสังคม-เวชศาสตร์ครอบครัว	4	4
6. เวชศาสตร์ชุมชน	4	3
7. จิตเวชศาสตร์	4	4
8. วิชาเลือกเสรี	2	2
9. รังสีวิทยา	2	2
รวม	45	43

แผนการศึกษา ปี 5 (46 หน่วยกิต)

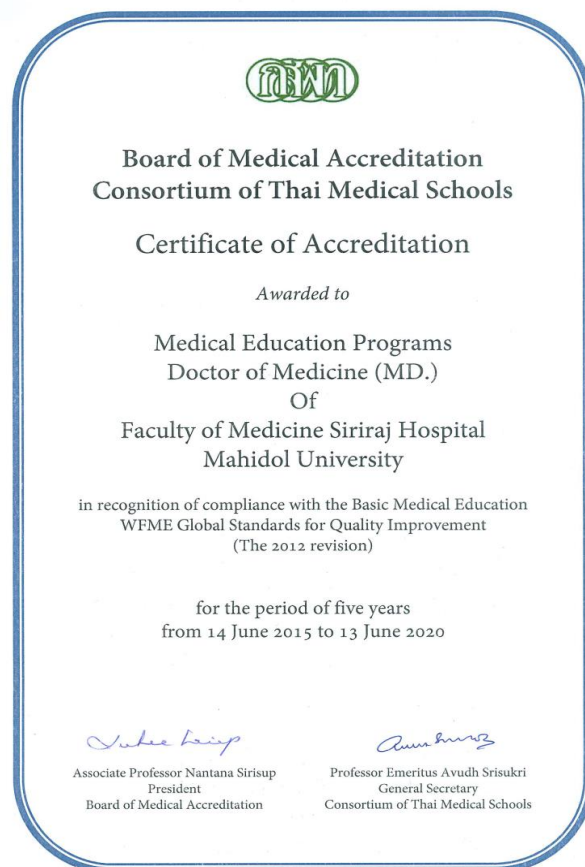
หมวดเรียนเป็นกลุ่ม (6 สัปดาห์)	สัปดาห์	หน่วยกิต
1. อายุรศาสตร์ 2	6	6
2. ศัลยศาสตร์ 2	6	6
3. สูติศาสตร์และนรีเวชวิทยา 2	6	6
4. กุมารเวชศาสตร์ 2	6	6
5. ศัลยศาสตร์ออร์โธปิดิกส์ 1	3	3
6. เวชศาสตร์ฉุกเฉินและอุบัติเหตุ 1	3	3
7. วิสัญญีวิทยา	2	2
8. จักษุวิทยา	2	2
9. โสต นาสิก ลาริงซ์วิทยา	2	2
10. เวชศาสตร์ฟื้นฟู	2	2
11. เวชศาสตร์ผู้ป่วยนอกและตจวิทยา 1	2	3
12. นิติเวชศาสตร์	2	2
13. เวชศาสตร์ชุมชนและการสร้างเสริมสุขภาพ	4	3
(พร้อมกันทั้งชั้นปี)	รวม	46

ภาคฤดูร้อน

แผนการศึกษา ปี 6 (42 หน่วยกิต)

กลุ่มที่ 1 (24 สัปดาห์)	สัปดาห์	หน่วยกิต
1. อายุรศาสตร์ 3	4	4
2. ศัลยศาสตร์ 3	4	4
3. สูติศาสตร์และนรีเวชวิทยา 3	4	4
4. กุมารเวชศาสตร์ 3	4	4
5. เวชศาสตร์ฉุกเฉินและอุบัติเหตุ 2	4	4
6. เวชบำบัดวิกฤต	2	2
7. เวชศาสตร์ผู้ป่วยนอกและตจวิทยา 2	2	2
กลุ่มที่ 2 (24 สัปดาห์)		
8. โรงพยาบาลร่วมสอน (5 ภาคฯ)	20	15
9. วิชาเลือกเสรี	4	3
รวม	48	42

หลักสูตร พ.บ.
ได้รับการรับรองมาตรฐาน
WFME 2012
เป็นเวลา 5 ปี
(14 มิ.ย. 2558 – 13 มิ.ย. 2563)



2562 - 2564

ปีแห่งการ**ทบทวน**หลักสูตร พ.บ.

ปีแห่งการ**รับรอง**หลักสูตร พ.บ.
ตามเกณฑ์ **WFME 2015**

ไม่ว่าจะปรับหลักสูตรกี่ครั้ง
อาจารย์
สำคัญที่สุด



Mahidol University
Faculty of Medicine
Siriraj Hospital



Siriraj Health science Education Excellence center

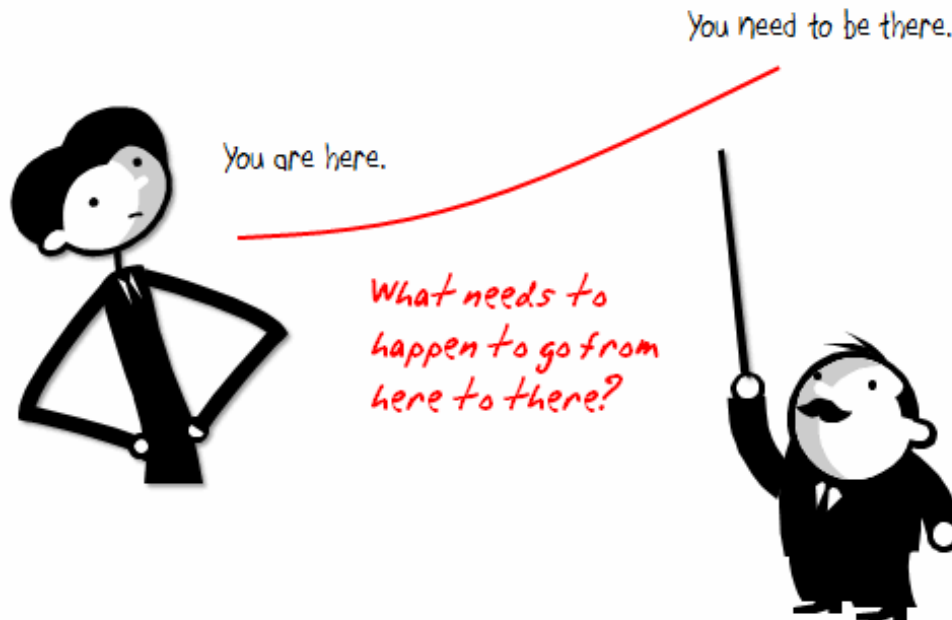
Educational Objectives and Lesson Plan

Suprapath Sonjaipanich MD.

Department of Pediatrics

Faculty of Medicine Siriraj Hospital

Educational Objectives



<http://blogs.articulate.com/images/blogs/rel/uploads/2008/11/here2there.gif>

Outline

- Definition, Level and Domain of educational objectives
- Purpose of educational objectives
- Writing learning objectives
 - characteristics of good learning objectives: SMART
 - components of learning objectives: ABCD



Educational objectives



a description of a performance you want learners to be able to exhibit before you consider them competent

It describes an intended **result** of instruction, rather than the **process** of instruction itself.

Robert F. Mager (1984)





Levels of Educational Objectives

I. Primary level

- Program / Curriculum objectives or Course objectives
- **Broad statement** of what the students will be able to do when they have completed the course / program



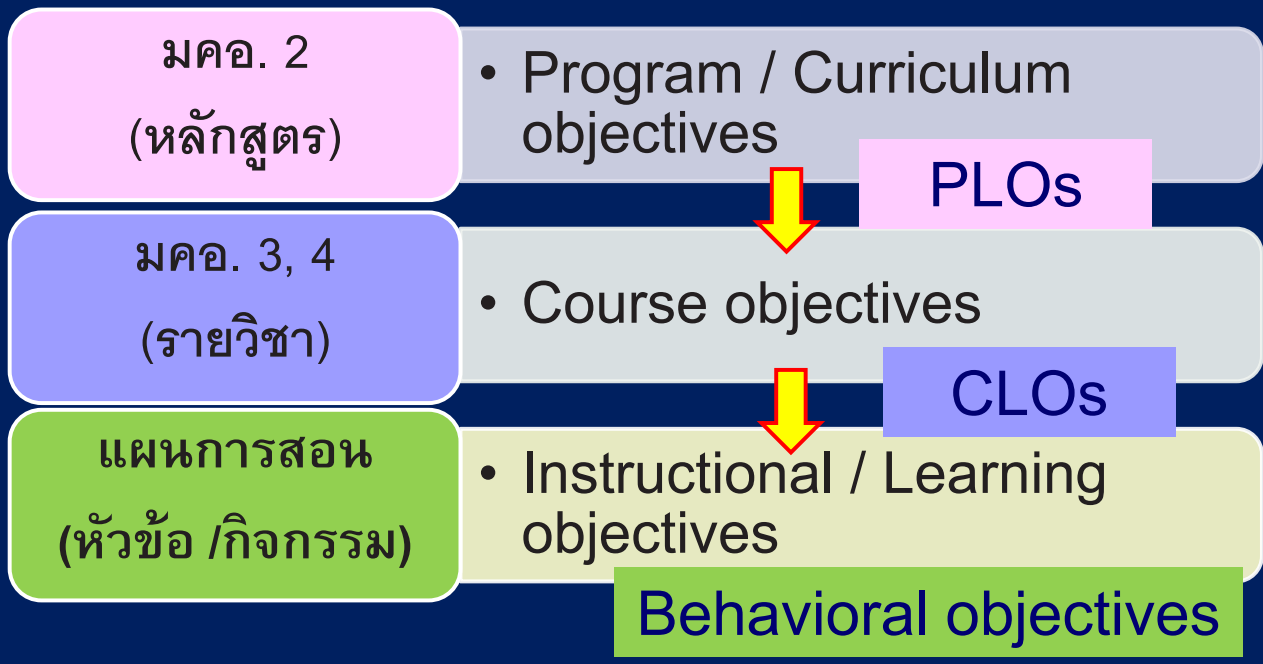
Levels of Educational Objectives

II. Second level

- Learning or Instructional objectives
- **Concise and specific** statement of **observable** student **behaviors** that can be assessed at the end of learning activities



Level of educational objectives





SIMD Curriculum objectives

- เพื่อสร้างบัณฑิตแพทย์ที่มีความรู้ ความสามารถ ทักษะ และเจตคติ ตามเกณฑ์มาตรฐานผู้ประกอบการ วิชาชีพเวชกรรมของแพทยสภา พ.ศ. 2555 และกรอบมาตรฐานคุณวุฒิระดับอุดมศึกษา สาขาแพทยศาสตร์ โดยเน้นคุณลักษณะที่พึงประสงค์ของบัณฑิตแพทย์ศิริราช ดังต่อไปนี้
 1. Soul หมายถึง มีจิตวิญญาณของความเป็นแพทย์ ประกอบด้วย คุณธรรมแห่งวิชาชีพ การคำนึงถึงผู้ป่วย.....
 2. Knowledge หมายถึง มีความรู้ทางวิชาชีพ (professional knowledge)
 3. Information หมายถึง



Course objectives in Pediatrics

- เมื่อสิ้นสุดการฝึกเวชปฏิบัติในโรงพยาบาลร่วมสอน เป็นเวลา 4 สัปดาห์ นักศึกษาแพทย์ชั้นปีที่ 6 จะมีความรู้ความสามารถด้านกุมารเวชศาสตร์ ดังนี้.....

เมื่อเสร็จสิ้นการบรรยายแบบมีส่วนร่วม และการทำกิจกรรม
อาจารย์ผู้เข้าอบรมสามารถ

1. อธิบายความสำคัญของวัตถุประสงค์การเรียนรู้ได้ครบถ้วน
2. บอกองค์ประกอบและลักษณะที่สำคัญของการกำหนด
วัตถุประสงค์การเรียนรู้ได้ครบถ้วน
3. เขียนวัตถุประสงค์การเรียนรู้ในแผนการศึกษา (Lesson
plan) ในรูปแบบ behavioral objectives สำหรับกิจกรรมการ
สอนที่อาจารย์เกี่ยวข้องได้ถูกต้องตามหลักการ อย่างน้อย 3
วัตถุประสงค์



Why are educational objectives important?

- To inform learners of what (knowledge, skills or attitudes) they expect to achieve from the course / learning activities
- To guide teachers in planning the teaching (e.g.. content, teaching strategies, materials)
- To help teachers in designing appropriate assessment tools
- To ensure that the desired course or learning objectives are achieved



Classification of educational objectives

Three domains of learning

I. **C**ognitive or **K**nowledge

II. **P**sychomotor or **S**kills

III. **A**ffective or **A**ttitudes

Bloom's taxonomy 1956



Knowledge objectives

- เมื่อสิ้นสุดการเรียนคาบนี้ นักศึกษาแพทย์สามารถ
 - ระบุสาเหตุที่พบบ่อยของภาวะไข้เฉียบพลันในเด็กอายุ 3 เดือน – 3 ปี ได้อย่างน้อย 3 ข้อ
 - เลือกการสืบค้นที่จำเป็นพร้อมแปลผล เพื่อวินิจฉัยและวางแผนการรักษาผู้ป่วยเด็กที่มีอาการไข้เฉียบพลัน



Skills objectives

- ภายหลังการเรียนและฝึกปฏิบัติในศูนย์ฝึกทักษะทางคลินิก นักศึกษาแพทย์ชั้นปี 4 จะสามารถเย็บแผลใน model โดยวิธี simple suture ได้ในเวลา 5 นาที
- เมื่อสิ้นสุดการปฏิบัติงานที่หอผู้ป่วยทารกแรกเกิด 2 สัปดาห์ นักศึกษาแพทย์ชั้นปี 5 จะสามารถประเมินอายุครรภ์ของทารกเกิดก่อนกำหนดโดยใช้ Ballard score ได้ถูกต้อง



21st century skills: Information

- เมื่อสิ้นสุดการเรียน การทำกิจกรรมกลุ่มย่อยและ การศึกษาเพิ่มเติมด้วยตนเอง นักศึกษาสามารถ
 - แสดงการค้นหาข้อมูลทางการแพทย์จากฐานข้อมูลแบบ peer-review ได้ในเวลาที่กำหนด
 - วิเคราะห์ข้อมูลโดยระบุ bias/confounder ที่เป็นไปได้จาก งานวิจัยที่ค้นหา อย่างน้อย 1 อย่าง พร้อมบอกวิธี จัดการที่เหมาะสม



Attitudinal objective

- เมื่อสิ้นสุดการเรียนแบบกลุ่มย่อย นักศึกษาแพทย์สามารถ ตระหนักถึงความสำคัญของปัญหาการใช้สารเสพติดในวัยรุ่นไทยที่ส่งผลต่อสุขภาพและคุณภาพชีวิตของผู้เสพและครอบครัว



Characteristics of good learning objectives

SMART

- **Specific:** *precisely states what the learner will be able to do*
- **Measurable:** *can be observed during or after the training session*
- **Action-oriented:** *uses an active verb that represents a behavior change*
- **Realistic:** *is appropriate to the time and scope*
- **Time-bound:** *is expected to achieved within the time-span of the session*



Interactive lecture: การสั่งยาในผู้ป่วยเด็ก

เมื่อสิ้นสุดการบรรยายและการทำกิจกรรมในชั้นเรียน นศพ.ชั้นปี 5
สามารถ
(reasonable, time-bound)

(action-oriented)

1. อธิบายประเด็นสำคัญของขั้นตอนการสั่งและจ่ายยาที่อาจเกิด

ความผิดพลาดได้ครบทุกข้อ (specific, measurable)

(action-oriented)

2. ระบุปัจจัยทั้ง 4 ข้อ ที่แพทย์ควรพิจารณาเมื่อสั่งยาให้ผู้ป่วย

เด็ก (specific, measurable)

(action-oriented)

3. เขียนใบสั่งยาผู้ป่วยนอกกรณีศึกษา 3 ราย ที่มีข้อมูลถูกต้อง

ครบถ้วน (specific, measurable)



Writing an objective: components

A: Audience (ผู้เรียน)

B: Behaviors (พฤติกรรมที่ครูจะเปลี่ยน)

C: Conditions (เงื่อนไขปฏิบัติ)

D: Degree (ระดับความสามารถที่ทำได้)

ABCD model

Heinich R, Molenda M, RussellJ, and Smaldino S. Institutional Media and Technologies for Learning 2002. (7th ed.)



A: Audience / Learner (ผู้เรียน)

A: Who is/are the learner(s)?

e.g.

- Fourth-year medical students
- Second-year pediatric residents



B: Behaviors (พฤติกรรมที่ครูจะเปลี่ยน)

B: What should the learner be able to do?

What is the measurable behavior?

- an action verb describes an observable performance eg. “identify, differentiate, solve, list, compare, show, create”
- avoid vague, poorly defined, non-specific verbs *such as “know, understand, appreciate, enjoy”*

Verb List for Writing Behavioral Objectives



Cognitive Domain Action Verbs

	Description	Useful action verbs
Knowledge	Recall of information	Define, list, name, record
Comprehension	Grasping the meaning	Describe, explain, discuss, recognize
Application	Using the rules and principles	Apply, use, demonstrate, illustrate, practice
Analysis	Breaking down components to clarify	Distinguish, analyze, calculate, test, inspect
Synthesis	Arranging and assembling elements into a whole	Design, organize, formulate, propose
Evaluation	Ability to judge X for a purpose	Judge, appraise, evaluate, compare, assess



C: Conditions (เงื่อนไขปฏิบัติ)

C: the important conditions under which the performance is to occur

- under supervision
- in emergency setting
- in simulated environment
- at the end of clinical practice in a nursery ward
- during hands on workshop



D: Degree (ระดับความสามารถที่ทำได้)

D: how well must the behavior be done?
criterion of acceptable performance, *if applicable*

- a degree of accuracy e.g.. score 80% or higher
- a stated proportion, e.g. 3 out of 5 or a minimum of 3
- within a given period of time e.g. perform within 5 minutes (in OSCE station)



ตัวอย่าง (1)

- เมื่อสิ้นสุดการอบรมในหัวข้อนี้ อาจารย์ผู้เข้ารับการอบรมทุกท่าน สามารถเขียนวัตถุประสงค์การเรียนรู้ ในหัวข้อที่ตนเองสอน ได้ถูกต้องอย่างน้อย 2 วัตถุประสงค์



ตัวอย่าง (2)

- เมื่อสิ้นสุดการเรียนในภาคเรียนที่ 1 นักศึกษา
แพทย์ชั้นปี 2 สามารถอธิบายระบบการไหลเวียน
เลือดแดงได้อย่างถูกต้อง ได้คะแนนไม่ต่ำกว่าร้อยละ
60 ในการสอบอัตนัย



ตัวอย่าง (3)

- เมื่อสิ้นสุดการฝึกปฏิบัติหัตถการพื้นฐานในผู้ป่วย
เด็กที่ศูนย์ฝึกทักษะทางคลินิก นักศึกษาแพทย์ชั้นปี 5
สามารถแสดงการเจาะน้ำไขสันหลังในหุ่นจำลองได้
ถูกต้องทุกขั้นตอน



Conclusion

Good learning objectives

- Learning center and precise description of what learners are expected to achieve at the end of the teaching session
- Using ABCD formula and should be SMART
- Linking to the instructional methods and assessment instruments



Questions and Comments

SUPRAPATH.SON@MAHIDOL.AC.TH



29



“ An instructor will function in a fog of their own making until they know just what they want their students to be able to do at the end of instruction.”

Mager R. F.

Writing Instructional Objectives

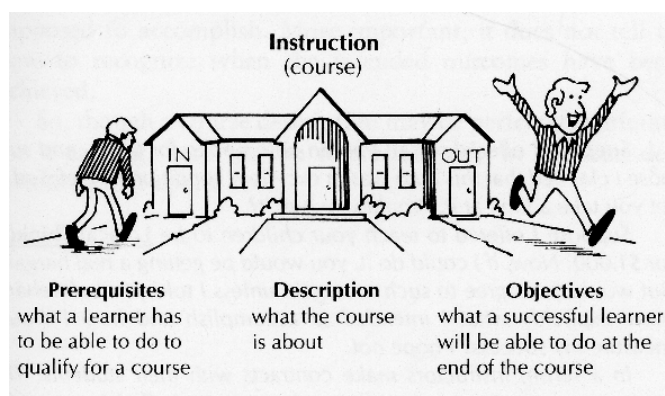
Babbi J. Winegarden, Ph.D.
Assistant Dean for Educational Development and Evaluation
UCSD School of Medicine
Department of Medical Education

Begin with the end in mind...

An *objective* is a description of a performance you want learners to be able to exhibit before you consider them competent. An objective describes an intended *result* of instruction, rather than the *process* of instruction itself.
(Mager, p. 5)

Course Description, Goals and Objectives

It is important to be able to differentiate the course description from the course objectives. A course description simply tells what the course is about. You might consider the GOALS of the course to be linked to the course description; they are broad educational statements fitting the mission and description of the course. Specific measurable objectives, however, tell what the learner will be able to do upon successful completion of the course. *Begin with the end in mind...*



(Mager, p. 11)

Why are well-written objectives important?

First, they provide some basis and guidance for the selection of instructional content and procedures. Second, they help in evaluating the success of the instruction. Third, they help the student organize his/her efforts to accomplish the intent of the instruction.

How Does One Write a Good Objective?

Most of us are fairly proficient at writing a course description and course goals. We have a good sense of what we want the course to be about. Writing behavioral objectives, however, can take some practice. So let's begin with just two simple rules.

1. Rule #1: A good objective communicates your intent well and leaves little room for interpretation. There are words that we often use that are open to many interpretations, and there are words that we can use that leave less to the

imagination. So, when writing behavioral objectives, stick to the words that leave less room for interpretation. Consider the following:

WORDS OPEN TO MANY INTERPRETATIONS	WORDS OPEN TO FEWER INTERPRETATIONS
to know	to write
to understand	to recite
to <u>really</u> understand	to identify
to appreciate	to sort
to <u>fully</u> appreciate	to solve
to grasp the significance of	to construct
to enjoy	to build
to believe	to compare
to have faith in	to contrast

The verbs below may prove useful as you write your instructional objectives and are based on Blooms' Taxonomy of cognitive behavior: For more information on Bloom's taxonomy, visit:

<http://faculty.washington.edu/krumme/guides/bloom.html>

Knowledge		Comprehension	
Cite	Recite	Associate	Express
Count	Recognize	Classify	Extrapolate
Define	Record	Compare	Interpret
Draw	Relate	Compute	Interpolate
Identify	Repeat	Contrast	Locate
Indicate	Select	Describe	Predict
List	State	Differentiate	Report
Name	Tabulate	Discuss	Restate
Point	Tell	Distinguish	Review
Quote	Trace	Explain	Translate
Read	Write	Estimate	
Application		Analysis	
Apply	Predict	Analyze	Distinguish
Calculate	Practice	Appraise	Experiment
Complete	Relate	Contract	Infer
Demonstrate	Report	Criticize	Inspect
Dramatize	Restate	Debate	Inventory
Employ	Review	Detect	Question
Examine	Schedule	Diagram	Separate
Illustrate	Sketch	Differentiate	Summarize
Interpret	Solve		
Interpolate	Translate		
Locate	Use		
Operate	Utilize		
Order			
Synthesis		Evaluation	
Arrange	Integrate	Appraise	Measure
Assemble	Manage	Assess	Rank
Collect	Organize	Choose	Rate
Compose	Plan	Critique	Recommend
Construct	Prepare	Determine	Review
Create	Prescribe	Estimate	Score
Design	Produce	Evaluate	Select
Detect	Propose	Grade	Test
Formulate	Specify	Judge	
Generalize			

2. **RULE #2:** There are three characteristics that help communicate intent when writing an objective: Performance, Conditions, and Criterion.

Performance: An objective always states what a learner is expected to be able to *DO*.

Conditions: An objective often describes the conditions under which a student is able to *DO* or perform the task.

Criterion: If possible, an objective clarifies how well the student must perform the task, in order for the performance to be acceptable.

...Other ways to think about writing instructional objectives include the following:

1. Who
2. Will do
3. How much (how well)
4. Of what
5. By when

or

ABCDs of Writing Objectives

- **A-Audience:** The who. "The student will be able to..."
- **B-Behavior:** What a learner is expected to be able to do or the product or result of the doing. The behavior or product should be observable.
- **C-Condition:** The important conditions under which the performance is to occur.
- **D-Degree:** The criterion of acceptable performance. How well the learner must perform in order for the performance to be considered acceptable.

Learner objectives can relate to Cognitive, Affective, and Psychomotor domains.

Here are some examples of fairly well-written objectives: There is always room for improvement...

Cognitive:

By the end of the Neurology curriculum, the learner will describe in writing a cost-effective approach to the initial evaluation and management of patients with dementia (an approach that includes at least six of the eight elements listed on their handout).

By the end of their internal medicine ambulatory clerkship, each third-year medical student will have achieved cognitive proficiency in the diagnosis and management of

hypertension, diabetes, angina, chronic obstructive pulmonary disease, hyperlipidemia, alcohol and drug abuse, smoking, and asymptomatic HIV infection, as measured by acceptable scores on interim tests and the final examination.

The student will be able to name the five stages of mitosis. (Not all objectives need all components specified above. This is a nice, clearly measurable objective).

The student will be able to describe protein folding and its relation to protein function.

The student will be able to define stress; discuss typical neurobiological and behavioral responses to stress and their implications for physical and mental functioning.

Affective:

By the end of the HIV curriculum, all residents will have identified their attitudes and beliefs regarding HIV patients who are drug abusers and will have discussed with their colleagues and attending physicians how these might influence their management of such patients.

Interact with others with honesty, civility, and respect.

Maintain composure when dealing with hostile or difficult patients. These skills can be assessed by direct or videotaped observation by an instructor.

From UCSD's Curricular Objectives

Behaviors

To demonstrate:

- Compassion and sensitivity in the care of patients and respect for their privacy and dignity.
- Respect for patient autonomy.
- The ability to build therapeutic relationships with patients.
- Honesty and integrity in all interactions.
- Responsibility and trustworthiness in the execution of all duties.
- The ability to accept criticism and to understand the limitations of one's own knowledge and skills.
- Adherence to all of the principles of UCSD's Professionalism Code.
- Demonstrate a commitment to excellence and ongoing professional development.

In addition, graduates of the UCSD School of Medicine should possess the following:

- Dedication to the well-being and needs of patients.
- Dedication to lifelong learning and an appreciation for the role of science in medical advances.
- Dedication to continual enhancement of clinical skills.
- An understanding of the threats posed by conflicts of interest in the practice of medicine and the performance of research.
- A commitment to promote the health and well-being of the communities they serve.
- The willingness to lead when leadership is required.

Psychomotor Skill:

By the end of the curriculum, all medical students will have demonstrated proficiency in assessing alcohol abuse by utilizing all four of the CAGE questions with one simulated and one real patient. These skills can be assessed by direct or videotaped observation by an instructor.

Demonstrate effective use of rapport-building techniques, including empathic listening, facilitation, clarification, validation, reflection, and appropriate nonverbal cues. These skills can be assessed by direct or videotaped observation by an instructor.

Assess patient’s expectations, values and goals. These skills can be assessed by direct or videotaped observation by an instructor.

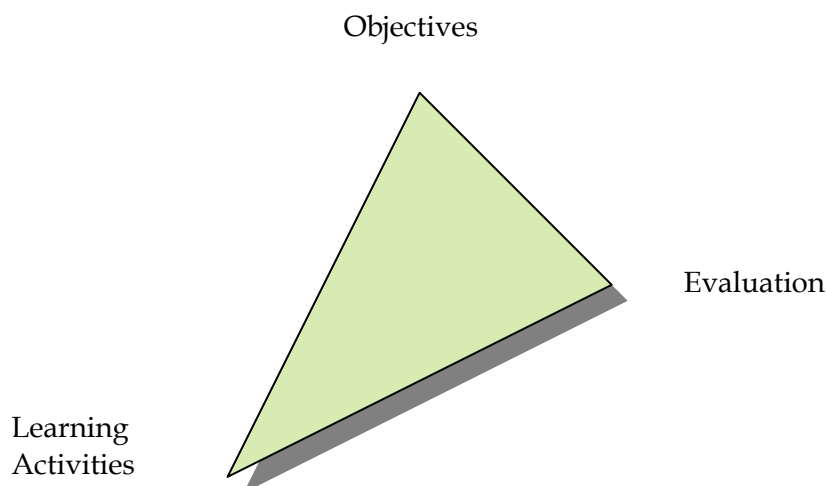
Use proper hand-washing technique prior to starting the physical exam. These skills can be assessed by direct or videotaped observation by an instructor.

Perform a basic psychiatric interview, including risk of suicide and potential of harm to self or others. These skills can be assessed by direct or videotaped observation by an instructor.

Some objectives involve more than one domain:

The learner will be able to: orally present (skill) a new patient’s case in a logical manner, chronologically developing the present illness, summarizing the pertinent positive and negative findings as well as the differential diagnosis and plans for further testing and treatment (cognitive).

We end with the Magic Triangle:



If we begin with the end in mind, the objectives inform the learning activities and evaluation confirms that the expected learning did indeed occur.

References

<http://www.uams.edu/oed/resources/objectives.asp#abcde>

<http://med.fsu.edu/education/FacultyDevelopment/objectives.asp>

<http://faculty.washington.edu/krumme/guides/bloom.html>

http://meded.ucsd.edu/Catalog/overview/curricular_obj.html

registrar.mc.duke.edu/forms/GoalsObjectives.pdf

Kern, D.E., Thomas, P.A., Howard, D.M., & Bass, E.B. (1998). Curriculum Development for Medical Education: A Six-Step Approach. Johns Hopkins University Press: Baltimore, Maryland.

Mager, R. F. (1984). Preparing Instructional Objectives (2nd edition). Lake Publishing Company: Belmont, California.

Verb List for Writing Behavioral Objectives

Goals and objectives are critical to planning educational experiences. A goal is a general description of what the learner will gain from instruction; an objective is a statement in specific and measurable terms that describes what the learner will know or be able to do as a result of the learning activity. Objectives serve as guides in planning, delivering, and evaluating instruction. They help learners know what is expected of them and help them prioritize content. They provide a template for tests and other evaluation tools. Learning objectives typically are composed of four parts, also known as the ABCD's of objective writing:

- A = audience who
- B = behavior will do what (in measurable terms)**
- C = condition under what conditions
- D = degree how much, or how well, or both

The most important component of a learning objective is the action verb that specifies the performance required. When creating objectives, the more precise the action verb the better. When choosing verbs for objectives, the emphasis is on sing verbs that are specific and unambiguous. The three domains of learning objectives are: cognitive (knowledge), psychomotor (skill), and affective (attitude)

Cognitive (Knowledge) Domain

The New Version of Bloom's Taxonomy for Objectives in the Cognitive Domain

Level	Cognitive process	Verbs
Lower Level Objectives		
Remembering	Remembering learned material	define, list, state, name, recall, recite, repeat, state, point out, tell, write, cite
Understanding	Explaining material that has been learned	identify, explain, recognize, discuss, describe, restate, express, translate, review, paraphrase, summarize
Higher Level Objectives		
Applying	Using knowledge to find or develop new solutions	apply, demonstrate, illustrate, interpret, dramatize, illustrate, operate, practice, perform, use, order, classify, utilize, calculate, employ, interpolate, adapt, consolidate, develop, extrapolate, modify, predict
Analyzing	The ability to break down material into parts so that is organizational structure can be understood	analyze, categorize, compare, differentiate, examine, contrast, diagram, distinguish, experiment, inspect, question, relate, test, infer, inventory, audit, correlate, deduce, discriminate, investigate, solve, verify
Evaluating	Making decisions and supporting views	choose, critique, rate, appraise, assess, estimate, evaluate, judge, measure, revise, score, select, value, determine, grade, recommend, conclude, confirm, justify, prioritize, prove, research, validate, support
Creating	Combining information, concepts and theories	design, formulate, plan, manage, construct, compose, arrange, create, organize, plan, prepare, propose, assemble, diagnose, generalize, integrate, prescribe, produce, specify, build, devise, generate, integrate, synthesize

After Waller K.V. *Writing Instructional Objectives* (see references)

Created by Teri L Turner, MD, MPH, MEd and Sandra M Sanguino, MD September 2013

Psychomotor Objectives:

Psychomotor (skill) objectives are easier to measure than affective or cognitive objectives because they are readily observable. It is important to delineate the degree or standard of performance expected. The degrees for performance of psychomotor objectives tend to be very explicit to allow the learners (and evaluators) to know if the objectives were achieved relative to standards. An example would be, “The learner will intubate a child older than one year of age who does not have underlying airway problems utilizing a curved blade. The outcome will be considered successful if the learner performs the task within two attempts, does not break or dislodge any teeth, and positions the tip of the endotracheal tube 5-15 mm above the carina.”

align	integrate	measure	percuss	disinfect
arrange	collate	palpate	utilize	assemble
hold	close	prepare	perform	dismantle
clean	straighten	insert	repair	tie
tighten	manipulate	operate	extract	standardize
compose	react	troubleshoot	arrange	prepare

Affective Objectives:

The affective domain involves attitudes, feelings, values, and beliefs. While this domain is very important, it is also the most difficult to teach. Because objectives for the affective domain involve attitudes, beliefs, and values, they are difficult to write and extremely difficult to measure. Despite these limitations, every effort should be made to include affective objectives when possible.

Level	Judgment	Verbs	Example
Receiving (attending)	Learners are willing to receive the subject matter	Accept, acknowledge, attend (to) follow, listen, meet, observe, receive	The physician will listen attentively while the patient expresses his beliefs about the cause of his illness
Responding	Learners prefer the subject matter	Agree, answer, ask, attempt, choose, comply, conform, cooperate, exhibit, help, notify, offer, participate (in), read, respond, try	The resident will answer a call for volunteers to work with patients displaced by Hurricane Katrina
Valuing	Learners are committed to the subject matter	Adopt, care (for), compliment, contribute, encourage, endorse, foster, initiate, join, praise, propose, respect, seek, subscribe, suggest support, thank, uphold, express appreciation	The physician will express appreciation for the contributions of all team members in the care of his patients
Organization	Learners are forming a life philosophy	Collaborate, confer coordinate, direct, establish, facilitate, follow through, lead, manage, organize, oversee, recommend, choose (to)	The medical student will choose to eat nutritious food over food obtained from the snack achiness while in the hospital
Characterization by value or value complex	The learners’ values consistently guide their behaviors without conscious forethought	Act, advocate, challenge, promote, promulgate, reject, serve, strive, commit (to), defend, empathize, enhance, endeavor, forgive, influence, motivate, negotiate, persevere, persist, promulgate, profess, reject, serve, strive, volunteer (for)	The physician will habitually abide by the standards outlined in the Hippocratic Oath

Created by Teri L Turner, MD, MPH, MEd and Sandra M Sanguino, MD September 2013

Basic Learning Psychology

เชิดศักดิ์ ไอรณรัตน์

ภาควิชาศัลยศาสตร์ คณะแพทยศาสตร์ศิริราชพยาบาล

"If we teach today as we were taught yesterday, we rob our children of tomorrow."

John Dewey

Cognitive Psychology

- What is learning? What is cognition?
- How does learning occur?
- How can we influence learning in our students?

Objectives

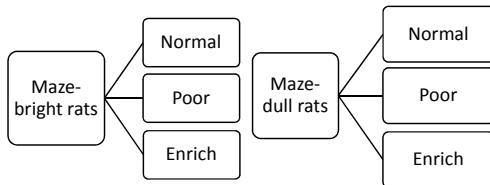
- เมื่อสิ้นสุดการเรียนรู้ในคาบนี้แล้ว อาจารย์ผู้เข้าอบรมสามารถ
 - อธิบายบทบาทของอาจารย์ในการส่งเสริมการเรียนรู้ของนักศึกษา
 - เสนอแนวทางในการกระตุ้นให้นักศึกษามีความสนใจในบทเรียน
 - วางแผนการสอนที่ทำให้ผู้เรียนเกิดการเรียนรู้ได้อย่างมีประสิทธิภาพ

Outline

- Teacher's roles in student's learning
- Motivation
- Cognitive information processing

Teachers' roles in Students' Learning

A Research: Genetics and Learning in Rats



Cooper RM, Zubeck JP. Effects of enriched and restricted early environments on the learning ability of bright and dull rats. Canadian Journal of Psychology 1958, 12: 159-64.

- Teacher’s roles in providing appropriate learning environment is critical to allow students to prosper to their full potential.

Motivation

Extrinsic Motivation

- A person performs a task because of a stimulus outside of the task or activity.
- Examples
 - Money
 - Exam scores
 - Threat of punishment

LEGO Bionicle

- Participants build up some Lego Bionicles.
- Condition A: Each Bionicle get \$2. The next Bionicle get \$ 0.11 less.
- Condition B: The same incentive, with disassembly of Bionicles right away.

Ariely D, et al. Man’s search for meaning: The case of Legos. J Econ Behav & Organization 2008.

IKEA

- Assembling a piece of IKEA furniture demands a significant amount of time and effort. People tend to gain satisfaction from completing the task and love the furniture more.

Norton M, et al. The IKEA effect: When labor leads to love. J Consumer psychology 2012.

New staff workshop

Nov 2019

Intrinsic Motivation

- A person performs a task because of rewards inherent to a task or activity itself
- Examples:
 - Playing jigsaw puzzle for fun
 - Drawing pictures for relaxation of mind

Curiosity

- People are motivated to learn when they see or perceive of new things.
- Novel, complex, or unique patterns in the environment are good learning motivators.
 - Examples:
 - Teach M4 students about surgical hemostasis by showing interesting tools for hemostasis and then lead to the lesson

Goals and Goal Orientation

- When learning something new, people set up their goals. Most of the time, people keep trying until they determine that they have achieved the goal.
 - Examples:
 - Persistent practice of knot tying by students during a surgical rotation

Goals and Goal Orientation

- Key determinants:
 - The specificity of the goal (specific vs general)
 - Time to achieve the goal (proximal vs distal)
 - Determination of achievement (learning vs performance)

Goal Orientation

1. The specificity of the goal
 - Specific goal: being able to tie a knot
 - General goal: knowing principles of hemostasis
2. Time to achieve the goal
3. Determination of achievement

Goal Orientation

1. The specificity of the goal
2. Time to achieve the goal
 - Proximal goal: knowing how to draw venous blood sample
 - Distal goal: Becoming a good anesthesiologist
3. Determination of achievement

cherdsakiramaneeerat@gmail.com

3

New staff workshop

Nov 2019

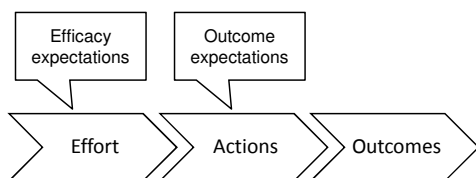
Goal Orientation

1. The specificity of the goal
2. Time to achieve the goal
3. Determination of achievement
 - Performance goal: gaining favorable judgment
 - Learning goal: increase their competence, skills, knowledge

Self-Efficacy

- Self-efficacy is the belief that one is capable of performing in a certain manner to attain certain goals.
- Albert Bandura (1977, 1982, 1997) proposed self-efficacy as a belief system that is causally related to behavior and outcomes

Bandura's Theory of Self-Efficacy



People with higher level of self-efficacy generally put more effort into their learning, persevere in the face of obstacles, and achieve better learning outcomes.

Self-Efficacy Beliefs

1. Enactive mastery experiences
2. Vicarious experiences
3. Verbal persuasion
4. Physiological reactions

Enactive Mastery Experiences

- A learner's own previous success at a task
- The most influential source of self-efficacy

Vicarious Experiences

- A learner's observation of a role model attaining success at a task

cherdsakiramaneeerat@gmail.com

4

Verbal Persuasion

- Teachers or others persuade a learner that he or she is capable of succeeding at a particular task

Physiological Reactions

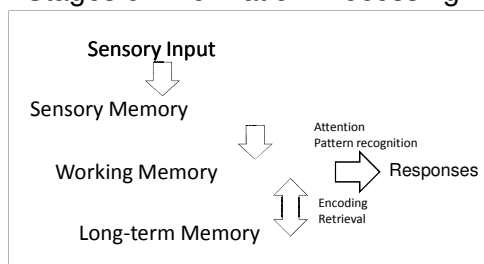
- Physiological reactions (heart rate, sweating, breathing pattern, shaking of hands, etc.) that result from encountering the task

Attribution Theory

- Weiner (1985, 1986)
 - Three dimensions that people categorize the causes of their success or failure
 1. Internal or external
 2. Unstable or stable
 3. Controllable or uncontrollable

Cognitive Information Processing

Stages of Information Processing



Atkinson RC, Shiffrin RM. Human memory: A proposed system and its control processes. In: Spence K, Spence J, editors. The psychology of learning and motivation. New York: Academic Press, 1968.

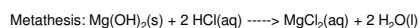
Attention

- The process by which people select some of the environmental input for further cognitive processing
- Catching attention
 - Meaning
 - Competing tasks
 - Task complexity or difficulty
 - Individual differences: age, IQ, disabilities
 - Presentation: Size, Novelty, Color, etc.

Pattern Recognition

- Comparing the incoming information with the prototype (past experience, prior learning) => perception of meaning/significance

Can you read this sentence? Of course, you can.



Working Memory

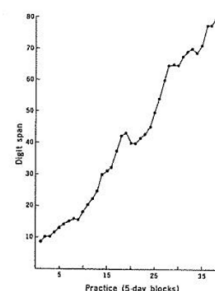
- Typical memory span: 7 ± 2 units of information

Acquisition of a Memory Skill

- An undergraduate (S.F.) with average memory abilities and average intelligence engage in the memory span task for 1 hr/day x 3 – 5 days/wk, for one and a half year

Ericsson KA, Chase WG, Falon S. Acquisition of a memory skill. Science 208, 1980: 1181 – 2.

A Digit Memorization Task



Long-term Memory

- Rehearsal: Repeating the information to yourself over and over again
- Encoding: Relating incoming information to concepts or ideas already in memory in such a way that the information is more memorable
 - Categorization (grouping)
 - Hierarchies or diagrams
 - Mnemonics, stories
 - Songs
 - Images (Illustrations)

A Research Study

- 124 university students age 18 – 24 years
- Subject: English reading comprehension
- 2 x 3 groups
- Two learning approaches
 - Group A: Study, Study
 - Group B: Study, Test
- Three testing times: 5 min, 2 days, 1 week

Roediger HL, Karpicke JD. Test-enhanced learning: Taking memory tests improves long-term retention. Psychological Science 2006, 17(3): 249-55.

New staff workshop

Nov 2019

A Research Study

- 180 university students age 18 – 24 years
- Subject: English reading comprehension
- 3 x 2 groups
- Three learning approaches
 - Group A: Study, Study, Study, Study
 - Group B: Study, Study, Study, Test
 - Group C: Study, Test, Test, Test
- Two testing times: 5 min, 1 week

Roediger HL, Karpicke JD. Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science* 2006, 17(3): 249-55.

The Benefit of Testing

- Repeated testing is an effective learning strategy to promote long term memory.
- Self-test should be done early.

Testing Effect or Test-enhanced learning

Karpicke JD, Butler AC, Roediger HL. Metacognitive strategies in student learning: Do students practise retrieval when they study on their own? *Memory* 2009, 17(4): 471-9.
Roediger HL, Karpicke JD. Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science* 2006, 17(3): 249-55

Summary

- Teacher's roles in student's learning
- Motivation
- Cognitive information processing

“Education is the most powerful weapon which you can use to change the world.”

Nelson Mandela

cherdsakiramaneeerat@gmail.com

7

Iramaneerat C. Motivation: Part I [Thai]. Medical Education Pamphlet 2008; 4(1): 1-2.

การสร้างแรงจูงใจให้นักเรียนสนใจเรียน (Motivation)

(ตอนที่ 1)

เชิดศักดิ์ ไชรมณีรัตน์

แรงจูงใจ (Motivation) เป็นสิ่งที่มีความสำคัญมากในการเรียนรู้ หากนักเรียนไม่มีแรงจูงใจในการเรียน นักเรียนก็จะขาดความเอาใจใส่ในการอ่านหนังสือ หรือฟังบรรยาย นักเรียนจะเอาเวลาไปทำกิจกรรมอื่นที่ไม่เกี่ยวกับการเรียนรู้แต่มีความน่าสนใจ แทนที่จะใช้เวลาทำความเข้าใจในบทเรียน การขาดแรงจูงใจในการเรียนเป็นปัญหาที่สำคัญของการศึกษาในประเทศไทยในทุกๆระดับ ดังจะเห็นได้จากสิ่งที่เกิดขึ้นในห้องเรียนทั่วไป เมื่อสิ้นสุดการบรรยาย หรือการสนทนากลุ่มย่อยในรูปแบบต่างๆ จะพบว่านักเรียนไทยในแทบทุกระดับจะไม่มีปัญหาหรือข้อสงสัยใดๆที่จะถามอาจารย์ ซึ่งต่างจากนักเรียนในประเทศทางตะวันตกซึ่งมีความกระตือรือร้นที่จะอภิปรายแสดงความคิดเห็น หรือซักถามข้อสงสัย ในบทความนี้ผมจะขอกล่าวถึงการสร้างแรงจูงใจในการเรียนซึ่งน่าจะเป็นประโยชน์มากหากอาจารย์แพทย์สามารถนำไปประยุกต์ใช้ในการสอนเพื่อทำให้นักเรียน และแพทย์ประจำบ้านมีความสนใจในบทเรียนมากขึ้น

โดยทฤษฎีจิตวิทยาการศึกษา แรงจูงใจให้คนอยากเรียนรู้เกิดขึ้นได้จาก 3 ปัจจัย ได้แก่ (1) จุดมุ่งหมายในการเรียน (2) ความเชื่อเกี่ยวกับความสามารถของตนเอง และ (3) ความคาดหวังในผลการเรียน ผมจะขอแบ่งเนื้อหาของบทความนี้เป็น 4 ตอน โดยเป็นการกล่าวถึงปัจจัยสร้างแรงจูงใจในการเรียนรู้แต่ละปัจจัยใน 3 ตอนแรก และตอนสุดท้ายเป็นการอภิปรายถึงกลยุทธ์การจัดการสอนที่ใช้หลักการสร้างแรงจูงใจเป็นพื้นฐาน

ปัจจัยประการแรกของแรงจูงใจในการเรียน คือ จุดมุ่งหมายในการเรียน (learning goals) การมีจุดมุ่งหมายที่แน่ชัดในการเรียนเป็นสิ่งสำคัญ แต่จุดมุ่งหมายแต่ละอย่างสามารถสร้างแรงจูงใจให้นักเรียนสนใจเรียนได้ไม่เท่ากัน ลักษณะสำคัญ 3 ประการในจุดมุ่งหมายในการเรียนที่ส่งผลต่อแรงจูงใจในการเรียน คือ

(1) ความเฉพาะเจาะจงของจุดมุ่งหมาย

จุดมุ่งหมายที่มีความเฉพาะเจาะจงมากจะสร้างแรงจูงใจในการเรียนได้ดีกว่า นักเรียนที่ตั้งจุดมุ่งหมายในการเรียนว่าเมื่อเรียนจบชั่วโมงนี้แล้วเขาจะผูกด้ายผ่าตัดด้วยมือข้างเดียวเป็น จะมีความกระตือรือร้นที่จะเรียนมากกว่านักเรียนที่ตั้งจุดมุ่งหมายว่าเมื่อเรียนจบชั่วโมงนี้แล้วเขาจะเข้าใจหลักการพื้นฐานของวิธีการผูกด้ายผ่าตัด

(2) ระยะเวลาที่ใช้ในการไปถึงจุดมุ่งหมาย

จุดมุ่งหมายที่สามารถบรรลุผลได้ในระยะเวลาอันสั้นมักจะสร้างแรงจูงใจในการเรียนได้ดีกว่าจุดมุ่งหมายที่ต้องใช้เวลาอีกนานกว่าจะเห็นผล นักเรียนที่ตั้งจุดมุ่งหมายว่าเมื่อเรียนจบชั่วโมงนี้แล้วเขาจะสามารถให้การวินิจฉัยผู้ป่วยที่มีอาการปวดท้องเฉียบพลันได้ จะมีความสนใจในการเรียนมากกว่านักเรียนที่ตั้งจุดมุ่งหมายว่าเมื่อเรียนจบจากโรงเรียนแพทย์แล้วเขาจะเป็นศัลยแพทย์

(3) เกณฑ์การประเมินความสำเร็จ

นักเรียนสามารถประเมินการบรรลุจุดมุ่งหมายได้ 2 ทาง ทางแรกคือการมองว่าความสำเร็จคือการแสดงให้เห็นว่าตนมีความสามารถ (performance goal) ส่วนทางที่สองคือการมองว่าความสำเร็จคือการที่ตนได้เกิดการเรียนรู้ (learning goal) การประเมินความสำเร็จแบบแรกนั้นเป็นการพุ่งความสนใจไปที่ คะแนนสอบ หรือ เกณฑ์ที่ได้รับจากการเรียน นักเรียนที่ประเมินความสำเร็จของตนในลักษณะนี้จะตั้งใจเรียนเฉพาะเมื่อตนมองเห็นว่าจะสามารถทำคะแนนสอบได้ดี นักเรียนที่ไม่มี

ความสามารถมากมักจะรู้สึกว่าคุณไม่มีทางจะได้คะแนนดีได้ต่อให้พยายามตั้งอกตั้งใจเรียนเท่าไรก็ตาม จึงไม่ใช้ความพยายามมากนักในการเรียน ในทางกลับกัน นักเรียนที่ใช้การประเมินความสำเร็จแบบที่สองนั้นจะไม่ใส่ใจมากนักกับคะแนนสอบ หรือเกรดเฉลี่ย แต่จะพยายามขวนขวายทำให้ตนเองได้รับความรู้ใหม่ อาจารย์แพทย์ควรจะพยายามส่งเสริมให้นักเรียนมองว่าการมีความรู้หรือความเข้าใจในวิชาการแพทย์นั้นถือว่าเป็นความสำเร็จในการเรียน โดยที่ไม่จำเป็นต้องได้คะแนน หรือเกรดเฉลี่ยที่ดีเลิศ นอกจากนี้การเปลี่ยนแปลงรูปแบบการตัดเกรดด้วยวิธีการอิงกลุ่ม (norm-reference) ซึ่งนักเรียนต้องทำคะแนนได้ดีกว่านักเรียนคนอื่นๆในชั้นเรียนจึงจะได้เกรดดี เป็นการตัดเกรดด้วยวิธีอิงเกณฑ์ (criterion-reference) ซึ่งนักเรียนที่ทำคะแนนถึงเกณฑ์ที่กำหนดไว้จะได้เกรดดี โดยไม่จำเป็นต้องแข่งขันกับนักเรียนอื่นในชั้น ก็จะทำให้นักเรียนที่ยึดติดกับคะแนน (performance goal) มีแรงจูงใจในการเรียนได้ดีขึ้นเนื่องจากนักเรียนสามารถได้เกรดดีเมื่อคะแนนผ่านเกณฑ์ที่กำหนดโดยไม่จำเป็นต้องเป็นคะแนนที่สูงกว่านักเรียนส่วนใหญ่ในชั้น

Iramaneerat C. Motivation: Part II [Thai]. Medical Education Pamphlet 2008; 4(2): 1.

การสร้างแรงจูงใจให้นักเรียนสนใจเรียน (Motivation)

(ตอนที่ 2)

เชิดศักดิ์ ไชยมณีรัตน์

ในบทความนี้ผมจะขอลงถึงปัจจัยที่สองในการสร้างแรงจูงใจในการเรียน คือ ความเชื่อเกี่ยวกับความสามารถของตัวนักเรียนเอง (Self-efficacy beliefs) ความเชื่อนี้ส่งผลต่อพฤติกรรม การเรียน และผลการเรียน งานวิจัยแสดงให้เห็นว่านักเรียนที่มีความเชื่อมั่นในตนเองว่าเขาสามารถที่จะทำกิจกรรมอย่างใดอย่างหนึ่งได้สำเร็จจะมีความสนใจในกิจกรรมนั้นร่วมมือในกิจกรรมนั้น มีความพยายามที่จะทำกิจกรรมนั้น และใช้เวลาทำกิจกรรมนั้น มากกว่านักเรียนที่ไม่มี ความเชื่อมั่นในตนเอง ความเชื่อมั่นในตนเองนี้ไม่ใช่คุณลักษณะที่คงที่ในนักเรียนแต่ละคน หากแต่เป็นคุณลักษณะที่มีความแปรปรวนตามปัจจัยต่างๆ เช่น นักเรียนที่มีความเชื่อมั่นว่าเขาจะเรียนวิชาชีวเคมีได้ดีอาจไม่มีความเชื่อมั่นในการเรียนวิชากายวิภาค หรือนักเรียนที่มั่นใจว่าเขาจะทำคะแนนสอบวิชาสรีรวิทยาได้ดีเมื่อสอบด้วยข้อสอบปรนัยอาจไม่มั่นใจในการสอบวิชาเดียวกันด้วยข้อสอบอัตนัย เป็นต้น งานวิจัยจำนวนมากแสดงให้เห็นว่ายิ่งนักเรียนมีความเชื่อมั่นในการเรียนวิชาหนึ่งๆ หรือการทำกิจกรรมอย่างใดอย่างหนึ่งมากเท่าไร นักเรียนคนนั้นจะตั้งเป้าหมายการเรียนรู้ หรือความสำเร็จในกิจกรรมนั้นๆ สูงขึ้นเท่านั้น และส่งผลให้เขามีความพยายามที่จะไปให้ถึงเป้าหมายนั้นมากขึ้น ทำให้เขาเกิดการเรียนรู้ในเรื่องนั้นๆ มากกว่านักเรียนที่ขาดความเชื่อมั่นในตนเอง

จะเห็นได้ว่าการสร้างความเชื่อมั่นในตนเองให้เกิดขึ้นในตัวนักเรียนแต่ละคนนั้นมีความสำคัญที่จะทำให้นักเรียนมีความสนใจเรียน ตั้งอกตั้งใจเรียน และประสบผลสำเร็จในการเรียน ในที่นี้ผมจะขอลงถึงปัจจัย 4 ประการที่ส่งผลต่อความเชื่อมั่นในตนเอง

1. ประสบการณ์ความสำเร็จที่เคยเกิดขึ้นกับตนเอง (enactive mastery experiences) ปัจจัยนี้เป็นปัจจัยที่สำคัญที่สุด หากนักเรียนเคยประสบความสำเร็จในการสอบครั้งแรกของรายวิชาหนึ่งแล้ว นักเรียนคนนั้นจะมีความเชื่อมั่นในตนเองว่าเขาจะประสบความสำเร็จในการสอบครั้งที่ 2 ในรายวิชาเดียวกัน
2. การสังเกตเห็นความสำเร็จของเพื่อน (vicarious experiences) หากนักเรียนเห็นว่าเพื่อนที่มีความสามารถใกล้เคียงกันกับเขาสามารถประสบความสำเร็จในการเรียนหรือการทำกิจกรรมหนึ่ง นักเรียนคนนั้นก็จะมีความเชื่อมั่นว่าเขาเองก็น่าที่จะประสบความสำเร็จได้เช่นกัน
3. การจูงใจด้วยคำพูด (verbal persuasion) การให้กำลังใจ หรือการชักจูงให้นักเรียนเชื่อว่าเขาสามารถทำสิ่งใดสิ่งหนึ่งได้ สามารถเพิ่มความเชื่อมั่นในตนเองของนักเรียนได้ อย่างไรก็ตามการชักจูงจะได้ผลดีต่อความเชื่อมั่นในตนเองเฉพาะเมื่อทำอย่างพอเหมาะ ไม่ชักจูงมากเกินไปจนความเป็นจริง การให้กำลังใจแก่นักเรียนที่เคยทำคลอดสำเร็จมาก่อนแล้วและเคยฝึกปฏิบัติการใช้ forceps ช่วยคลอดกับแบบจำลองมาแล้วว่าเขาสามารถใช้ forceps ช่วยคลอดภายใต้การดูแลของอาจารย์ได้เป็นการช่วยเพิ่มความเชื่อมั่นให้นักเรียนได้ แต่การบอกนักเรียนที่ไม่เคยทำคลอดเลย หรือไม่เคยฝึกปฏิบัติการใช้ forceps ช่วยคลอดกับแบบจำลองมาก่อน ว่าเขาสามารถใช้ forceps ช่วยคลอดในมารดาที่มีปัญหาคลอดยากได้ คงเป็นการชักจูงที่มากเกินไปจนความจริงและไม่ช่วยสร้างความเชื่อมั่นให้เกิดขึ้น
4. ปฏิกิริยาทางสรีรวิทยา (physiological reactions) ความรู้สึกเครียด อ่อนล้า ตื่นเต้น หรือกลัวที่เกิดขึ้นก่อนทำกิจกรรมบางอย่างเป็นกลไกการตอบสนองทางสรีรวิทยาที่สามารถส่งผลต่อความเชื่อมั่นในตนเองของนักเรียนได้ นักเรียนบาง

คนมีอาการมือสั่น ใจสั่น เหงื่อออกมาก เมื่อทำผ่าตัดครั้งแรก ปฏิบัติงานเหล่านี้มีผลลดความมั่นใจในตนเองของนักเรียนลง อาจารย์สามารถช่วยลดปฏิบัติการเหล่านี้ได้บ้างโดยการช่วยแนะนำให้นักเรียนผ่อนคลาย และไม่เพิ่มความเครียดให้นักเรียน

Iramaneerat C. Motivation: Part III [Thai]. Medical Education Pamphlet 2008; 4(3): 2.

การสร้างแรงจูงใจให้นักเรียนสนใจเรียน (Motivation)

(ตอนที่ 3)

เชิดศักดิ์ ไอรรมณีรัตน์

ในบทความนี้ผมจะขอกล่าวถึงปัจจัยที่สามในการสร้างแรงจูงใจในการเรียน คือ ความคาดหวังในผลการเรียน โดยทั่วไปผลที่เกิดขึ้นจากการเรียน หรือการทำกิจกรรมใดๆที่คนทั่วไปคาดหวังมี 3 ประการ ได้แก่

1. ผลที่เกิดกับร่างกายโดยตรง เช่น ความพึงพอใจ ความสุข การลดความเจ็บปวด เป็นต้น
2. ผลทางสังคม เช่น การยอมรับจากเพื่อนฝูง การได้รับรางวัลหรือผลตอบแทนทางการเงิน การทำโทษ เป็นต้น
3. การประเมินตนเอง ว่า ตนเองมีความสามารถมากน้อยเพียงใด มีข้อบกพร่องด้านใด ควรต้องพัฒนาด้านใดบ้าง

ความคาดหวังในผลลัพธ์ทั้ง 3 ประการนี้ส่งผลต่อแรงจูงใจของนักเรียน ความคาดหวังผลทางด้านบวก (รางวัล การยอมรับทางสังคม ฯลฯ) ช่วยเพิ่มแรงจูงใจให้นักเรียน ในทางกลับกัน ความคาดหวังผลในทางลบ (สอบไม่ผ่าน ถูกทำโทษ ถูกเพื่อนเยาะเย้ย ฯลฯ) ลดแรงจูงใจในการเรียนหรือการทำกิจกรรมของนักเรียน

นอกจากความคาดหวังในผลทั้ง 3 ประการนี้จะมีผลต่อแรงจูงใจให้นักเรียนสนใจเรียน ก่อนที่จะเริ่มเรียนหรือทำกิจกรรมใดๆ แล้ว ความคาดหวังเหล่านี้ยังมีความสำคัญในการรักษาความสนใจในการเรียนหรือการทำกิจกรรมของนักเรียนอีกด้วย เมื่อการเรียนหรือการทำกิจกรรมสิ้นสุดลงและนักเรียนได้รับผลของการเรียนหรือการทำกิจกรรมนั้นแล้ว นักเรียนจะยังสนใจที่จะเรียนเรื่องนั้นๆต่อไปในระดับที่สูงขึ้น หรือทำกิจกรรมนั้นๆ ต่อไปในระดับที่ยากขึ้นหรือไม่ขึ้นอยู่กับปัจจัยสำคัญ 2 ประการคือ

1. ความพึงพอใจในผลการเรียน หรือผลของกิจกรรมที่ได้ทำ หากการเรียนหรือการทำกิจกรรมนั้นทำให้นักเรียนได้รับผลดีตามที่คาดหวังไว้ หรือ ดีเกินความคาดหวัง นักเรียนก็จะคงความสนใจในการเรียนเรื่องนั้นๆ หรือการทำกิจกรรมนั้นๆ ต่อไปในระดับที่สูงขึ้นหรือยากขึ้น การได้รับผลดีเกินความคาดหวังจากการเรียนที่สำคัญมากประการหนึ่งคือการทำให้นักเรียนสามารถนำความรู้หรือทักษะที่ได้รับจากการเรียนนั้นไปใช้ประโยชน์ได้จริงในทางปฏิบัติ หากภายหลังจากที่นักเรียนได้เรียนวิธีทำหัตถการอย่างหนึ่ง แล้วนักเรียนได้มีโอกาสได้ใช้ทักษะดังกล่าวในผู้ป่วยจริงแล้วได้เห็นว่าการที่พวกเขาได้เรียนรู้มานั้นช่วยทำให้ผู้ป่วยดีขึ้น เขาจะเกิดความรู้สึกพึงพอใจในการเรียนมาก และจะสนใจเรียนรู้หัตถการต่างๆ เพิ่มขึ้น อย่างไรก็ตามความรู้หรือทักษะทางการแพทย์ที่นักเรียนได้รับการสอนจำนวนมากนั้นนักเรียนจะไม่มีโอกาสได้ใช้จริงกับผู้ป่วยที่ตนดูแลในทันที การจะเพิ่มความพึงพอใจในผลการเรียนอีกวิธีที่อาจารย์สามารถทำได้คือการให้รางวัล หรือคำชมเชยเมื่อนักเรียนทำได้ดี

2. การอธิบายถึงปัจจัยที่นำไปสู่ผลการเรียน หรือผลการทำกิจกรรมของนักเรียน หากนักเรียนไม่ได้รับผลดีจากการเรียนหรือทำกิจกรรมตามที่คาดหวังไว้ นักเรียนจะพยายามหาคำอธิบายว่าเหตุใดเขาจึงไม่ประสบความสำเร็จตามที่คาดไว้ หากสาเหตุที่นักเรียนใช้อธิบายความผิดหวังนั้นเป็นสิ่งที่อยู่ในตัวนักเรียนเอง (internal) สามารถเปลี่ยนแปลงได้ (unstable) และเขาสามารถควบคุมได้ (controllable) (เช่น อ่านหนังสือไม่จบก่อนสอบ) นักเรียนคนนั้นจะมีแรงจูงใจที่จะทำให้ดีขึ้นในอนาคต ในทางตรงข้ามหากนักเรียนคิดว่าสาเหตุของความผิดหวังนั้นเป็นสิ่งที่อยู่นอกตัวของนักเรียน (external) ไม่สามารถเปลี่ยนแปลงได้ (stable) และอยู่นอกเหนือความควบคุมของเขา (uncontrollable) (เช่น อาจารย์ออกข้อสอบในเรื่องที่ไม่เกี่ยวข้องที่เรียน) นักเรียนคนนั้นก็จะไม่มีแรงจูงใจที่จะพยายามทำให้ดีขึ้น หากอาจารย์สามารถช่วยชี้ให้นักเรียนเห็นว่าเขาไม่ประสบ

ความสำเร็จตามที่คาดหวังไว้เนื่องจากปัจจัยที่อยู่ในตัวนักเรียนเอง ที่เขาควบคุมได้ และ เป็นสิ่งที่สามารถเปลี่ยนแปลงได้หากมีความตั้งใจ นักเรียนก็จะสามารถคงความตั้งใจที่จะเรียนเรื่องนั้นต่อไปได้

Iramaneerat C. Motivation: Part IV [Thai]. Medical Education Pamphlet 2008; 4(4): 1.

การสร้างแรงจูงใจให้นักเรียนสนใจเรียน (Motivation)

(ตอนที่ 4)

เชิดศักดิ์ ไอรมณีรัตน์

ดังที่เห็นแล้วว่าแรงจูงใจเป็นสิ่งที่มีความสำคัญมากต่อการเรียนรู้ของนักเรียน ในบทความนี้ผมจะขอแนะนำเทคนิคในการสอนที่เน้นการกระตุ้นความสนใจของผู้เรียนซึ่งได้รับการพัฒนาขึ้นโดย John Keller การสอนในรูปแบบนี้ประกอบด้วยกระบวนการที่ต่อเนื่องกัน 4 ขั้นตอนคือ (1) ดึงดูดความสนใจ (attention), (2) สร้างความเกี่ยวข้องของเนื้อหากับความต้องการของนักเรียน (relevance), (3) สร้างความเชื่อมั่นในตนเองของนักเรียน (confidence), และ (4) ทำให้นักเรียนได้รับความพึงพอใจในการเรียน (satisfaction)

(1) ดึงดูดความสนใจ (attention) เป็นขั้นตอนแรกของการสอน อาจารย์สามารถสร้างความสนใจในบทเรียนให้เกิดขึ้นได้หลายวิธี เช่น การเริ่มต้นด้วยเรื่องตลกที่เกี่ยวข้องกับเรื่องที่จะสอน หรือ แสดงรูปที่น่าสนใจ หรือ วิดิทัศน์สั้นๆ ที่แสดงให้เห็นปัญหาที่จะต้องใช้ความรู้เรื่องที่จะเรียนในการแก้ไข นอกจากนี้การสร้างความสนใจในบทเรียนให้เกิดขึ้นเมื่อเริ่มต้นสอนแล้ว อาจารย์ยังต้องมีเทคนิคในการรักษาความสนใจของนักเรียนตลอดการสอนด้วย อาจารย์ควรเปลี่ยนรูปแบบการสอน หรือ สอดแทรกกิจกรรมต่างๆ เพื่อกระตุ้นความตื่นตัวของนักเรียนทุกๆ 20 นาที

(2) สร้างความเกี่ยวข้องของเนื้อหาด้วยความต้องการของนักเรียน (relevance) เมื่อนักเรียนให้ความสนใจในสิ่งที่อาจารย์จะสอนแล้ว อาจารย์สามารถจูงใจให้นักเรียนตั้งใจเรียนมากขึ้นโดยการทำให้นักเรียนเห็นว่าเรื่องที่อาจารย์กำลังสอนนี้ นักเรียนสามารถนำไปใช้ประโยชน์ได้จริง ปัจจัยสำคัญประการหนึ่งที่จะช่วยสร้างความเกี่ยวข้องคือความคุ้นเคย นักเรียนจะรู้สึกว่าการเรียนหนึ่งๆ มีประโยชน์ต่อเขามากหากนักเรียนมีความคุ้นเคยกับเรื่องดังกล่าวเช่น เคยอ่านเรื่องดังกล่าวมาก่อน เคยเห็นผู้ป่วยที่มีลักษณะดังกล่าวมาก่อน เคยช่วยทำหัตถการนั้นๆ มาก่อน เป็นต้น หากอาจารย์สามารถโยงบทเรียนเข้ากับประสบการณ์ที่นักเรียนมีมาก่อนได้ จะทำให้นักเรียนสนใจในบทเรียนมาก

(3) สร้างความเชื่อมั่นในตนเองของนักเรียน (confidence) ดังได้กล่าวมาแล้วในบทความตอนก่อนหน้าเห็นว่า นักเรียนที่มีความเชื่อมั่นว่าเขาจะเรียนเรื่องหนึ่งๆ ได้ดี จะมีความสนใจในบทเรียน และพยายามทำความเข้าใจในบทเรียนมากกว่านักเรียนที่ขาดความมั่นใจในตนเอง อาจารย์สามารถสร้างความเชื่อมั่นให้เกิดขึ้นได้ด้วย 3 เทคนิคคือ (1) ระบุเป้าหมายของการเรียนให้ชัดเจนว่าอาจารย์หวังว่านักเรียนควรรู้หรือทำอะไรได้บ้างเมื่อเรียนจบ และตั้งไต่ถามนอกเหนือจากความคาดหวัง (2) เปิดโอกาสให้นักเรียนได้แสดงความรู้หรือความสามารถในเรื่องที่ได้เรียนไป และ (3) ให้ความช่วยเหลือแก่นักเรียนในการทำกิจกรรมหรือแสดงความสามารถในเรื่องที่เรียนตามความเหมาะสม ไม่มากไปจนนักเรียนรู้สึกว่าเขาไม่สามารถทำสิ่งนั้นได้หากอาจารย์ไม่ช่วยเขา แต่ก็ไม่น้อยไปจนนักเรียนรู้สึกว่าเขาถูกทอดทิ้งให้ทำการอันยากเกินความสามารถของเขาโดยลำพัง

(4) ทำให้นักเรียนได้รับความพึงพอใจในการเรียน (satisfaction) อาจารย์สามารถทำให้นักเรียนเกิดความพึงพอใจในความรู้หรือทักษะที่เขาได้เรียนไปได้หลายวิธีเช่น เปิดโอกาสให้นักเรียนได้ใช้ความรู้หรือทักษะนั้นในการช่วยดูแลผู้ป่วยจริง การให้คำชมเชย ให้รางวัล เป็นต้น

Iramaneerat C. Cognitive information processing [Thai]. Medical Education Pamphlet 2005; 1(3): 2.

การประมวลข้อมูลในการเรียน (Cognitive Information Processing)

เชิดศักดิ์ ไอรมนิรัตน์

ในบทความนี้ผมขอแนะนำทฤษฎีที่ใช้อธิบายกระบวนการที่มนุษย์ใช้ในการประมวลข้อมูลที่ได้รับไปเก็บในความทรงจำ (cognitive information processing) ตามหลักการของทฤษฎีนี้ การเรียนรู้เกิดขึ้นเมื่อข้อมูลที่ได้รับจากสิ่งแวดล้อมได้รับการประมวลผลและถูกบันทึกไว้ในความทรงจำ แล้วผู้เรียนสามารถดึงเอาข้อมูลที่เก็บไว้ดังกล่าวมาใช้ได้เมื่อต้องการ ทฤษฎีนี้มุ่งเน้นความสนใจไปที่กระบวนการที่เกิดขึ้นภายในระบบความคิดของผู้เรียนระหว่างการรับรู้ (sensory input) กับการแสดงออก (response) ของผู้เรียน

การประมวลข้อมูลในระบบความคิดประกอบด้วย 3 ขั้นตอน คือ ข้อมูลจากการรับรู้ (sensory memory) ข้อมูลในระบบความจำชั่วคราว (short-term memory) และ ความทรงจำถาวร (long-term memory) ในขั้นตอนแรก ผู้เรียนรับข้อมูลจากระบบประสาท (ภาพ เสียง กลิ่น รส สัมผัส) ข้อมูลที่ได้รับในขั้นตอนแรกนี้มีปริมาณมาก แต่ประกอบด้วยข้อมูลทั้งที่มีและไม่มีประโยชน์ ข้อมูลนี้จะอยู่ในระบบความคิดเพียงเสี้ยววินาทีก่อนที่ข้อมูลจากสิ่งแวดล้อมชุดใหม่เข้ามาแทนที่ ผู้เรียนจะเลือกเอาเฉพาะข้อมูลที่มีประโยชน์เท่านั้นเข้าสู่ระบบความจำชั่วคราว ระบบความจำชั่วคราวนี้มีความจุจำกัด ข้อมูลในระบบความจำชั่วคราวนี้จะคงอยู่เพียง 20 – 30 วินาที เมื่อมีเรื่องใหม่เข้ามาให้คิด ข้อมูลในระบบความจำชั่วคราวนี้จะถูกแทนที่ด้วยข้อมูลใหม่ การที่ข้อมูลจะถูกบันทึกในระบบความจำถาวรได้นั้นจะต้องได้รับการเชื่อมโยงเข้ากับพื้นฐานความรู้ของผู้เรียนที่มีอยู่เดิมโดยการใส่รหัส (encoding) (ใช้บทกลอน คำย่อ หรือ ภาพช่วยสร้างความหมายให้กับข้อมูลที่ต้องการจดจำ) หรือ โดยการท่องจำหลายๆครั้ง (rehearsal)

ความเข้าใจเกี่ยวกับระบบการประมวลข้อมูลนี้มีประโยชน์มากในการจัดการเรียนการสอนในทุกๆระดับ ผมจะขอ ยกตัวอย่างการประยุกต์ทฤษฎีนี้มาใช้เป็นหลักการในการสอนนักเรียนแพทย์สักเล็กน้อยนะครับ

1. จัดการสอนที่เป็นระเบียบ (organized instruction) เนื้อหาที่นักเรียนแพทย์ต้องเรียนรู้นั้นมีปริมาณมาก หากอาจารย์ทำการสอนโดยบรรยายเนื้อหาตามที่ปรากฏในตำราให้นักเรียนฟัง ข้อมูลส่วนใหญ่จะได้รับเข้าสู่ระบบข้อมูลการรับรู้ หรือ ความจำชั่วคราวเท่านั้น ข้อมูลส่วนน้อยเท่านั้นที่ได้รับการถ่ายทอดไปสู่ระบบความทรงจำถาวร หากอาจารย์สามารถจัดเนื้อหาการสอนให้นักเรียนในรูปแบบที่มีระเบียบ ง่ายต่อการจดจำ และเชื่อมโยงกับพื้นฐานความรู้ของนักเรียนได้ เนื้อหาดังกล่าวจะถูกบันทึกเข้าในระบบความจำถาวร และไม่ถูกลืมโดยง่าย การจัดเนื้อหาให้เป็นระเบียบสามารถทำได้หลายวิธีเช่น การใช้แผนภาพ (diagram, flow chart, clinical pathway) หรือ รูปภาพ (conceptual drawing, real photo)

2. จัดให้มีการฝึกปฏิบัติ หากนักเรียนได้รับแบบฝึกหัดสำหรับเนื้อหาที่ได้รับการสอนด้วยจะทำให้นักเรียนจดจำบทเรียนได้ดีขึ้น การฝึกฝนทำให้นักเรียนได้ทบทวนเนื้อหาวิชา (rehearsal) ซึ่งช่วยในการถ่ายทอดความรู้เข้าสู่ระบบความทรงจำถาวร นอกจากนี้ การฝึกฝนยังทำให้นักเรียนสร้างความเชื่อมโยงความรู้เข้ากับพื้นฐานความรู้เดิม และ สถานการณ์ต่างๆ ได้ดีขึ้น ซึ่งช่วยในการใส่รหัส (encoding) ให้กับเนื้อหาดังกล่าวด้วย

3. ช่วยเหลือให้นักเรียนสร้างทักษะในการช่วยจำ นักเรียนจำนวนมากขาดทักษะในการจดจำเนื้อหาวิชา หากได้รับการแนะนำวิธีการที่ช่วยในการถ่ายข้อมูลเข้าสู่ระบบความจำถาวร นักเรียนเหล่านี้จะเรียนได้อย่างมีประสิทธิภาพมากขึ้น อาจารย์สามารถสอดแทรกเกร็ดในการช่วยจำเนื้อหาให้แก่ให้นักเรียนได้หลายรูปแบบ เช่น การใช้บทกลอน การเชื่อมโยงเนื้อหาเข้ากับตัวอย่างผู้ป่วย การสร้างความเชื่อมโยงเนื้อหาเข้ากับพื้นฐานความรู้เดิมในหลายด้าน การแบ่งเนื้อหาออกเป็นส่วนย่อยที่จดจำได้ง่าย การจดบันทึกด้วยคำพูดของตนเอง เป็นต้น

ผมหวังว่าเกร็ดความรู้เกี่ยวกับกระบวนการประมวลข้อมูลดังกล่าวมานี้คงเป็นประโยชน์ในการพัฒนาการสอนนักเรียนไม่มากนักน้อยน่ะครับ

New staff workshop

Nov 2019

Teaching a Large Class

รศ.นพ.เชิดศักดิ์ ไอร่มณีรัตน์
ภาควิชาศัลยศาสตร์
คณะแพทยศาสตร์ศิริราชพยาบาล

1

Objectives

- เมื่อสิ้นสุดการอบรมแล้ว อาจารย์ผู้เข้าอบรมสามารถ
 - บอกแนวปฏิบัติที่เหมาะสมในการสอนบรรยายแบบมีปฏิสัมพันธ์กับผู้เรียน

2

Lecture Tip #1

- Start by gaining attention
 - Story
 - Picture
 - Cases
 - Videoclips
 - Newspaper headline
 - Humor

3

Lecture Tip #2

- Provide clear learning objectives
 - An objective
 - indicates the purpose of instruction.
 - informs students what to expect and what they should be able to do after the instruction
 - reminds the teacher of the lecture focus.

4

Lecture Tip #3

- Focus on concepts, not on facts
 - Provide only essential facts
 - Frame facts within concepts
 - Explain concepts using different examples

5

Lecture Tip #4

- Prepare your lectures in three parts
 1. Introduction
 2. Body
 3. Conclusion

6

cherdsakirameerat@gmail.com

1

Lecture Tip #5

- Gather important information prior to your lecture
 - Content: scope, depth
 - Audience: experience level, specialty, special needs, number
 - Goal of the lecture: limit to only a single goal
 - Environment: room, audiovisual equipments
 - Time: time of the day, amount of time provided

7

Lecture Tip #6

- Use your voice and gestures effectively to engage your audience
 - Voice: Tone, pitch, pace
 - Gesture: Body position, hand movement

8

Lecture Tip #7

- Prepare not only what you will present, but also how you will present it.
 - Suggestions:
 - Keep eye contact
 - Speak clearly
 - Vary your pace
 - Pause from time to time
 - Limit filler words (umm..., ahh....)
 - Un-root yourself (move around)

9

Lecture Tip #8

- Involve your audience in your lecture
 - Questions
 - Cases
 - Examples
 - Presentations

10

Lecture Tip #9

- Prepare a good handout
 - A handout with
 - Background information
 - Main topics
 - Spaces for students to write:
 - Key facts
 - Concepts

11

Lecture Tip #10

- Don't forget to summarize your lesson
 - Intermittent summary of your main points
 - Provide concrete examples
 - Ask questions
 - Conclusion at the end of your lecture

12

New staff workshop

Nov 2019

Summary

- Tips for effective interactive lecturing
 - Attention
 - Objectives
 - Concepts
 - Three parts
 - Information
 - Voice and gesture
 - Rehearsal
 - Audience
 - Handout
 - Summary

“The success of your presentation will be judged not by the knowledge you send, but by what the listeners receives.”

Lily Walters

13

14

cherdsakiramaneeerat@gmail.com

3

Iramaneerat C. Preparing for a lecture [Thai]. Medical Education Pamphlet 2007; 3(4): 4.

การเตรียมการสอนบรรยาย (Preparing for a lecture)

เชิดศักดิ์ ไอรมนรัตน์

การสอนบรรยาย (lecture) เป็นการสอนที่ใช้อย่างมากในโรงเรียนแพทย์ ถึงแม้จะมีการใช้รูปแบบการสอนอื่นมากขึ้นเรื่อยๆ ในปัจจุบัน เช่น การสอนกลุ่มย่อย, problem-based learning, computer-assisted learning แต่การสอนบรรยายก็ยังเป็นการสอนที่มีความสำคัญมาก เนื่องจากเป็นรูปแบบการสอนที่สามารถทำได้ง่าย ใช้บุคลากรน้อย ให้ความรู้แก่นักเรียนได้จำนวนมากในเวลาพร้อมกัน การสอนบรรยายจัดเป็นรูปแบบการสอนที่เหมาะสมสำหรับให้ความรู้พื้นฐาน (factual knowledge and conceptual understanding) แต่ไม่เหมาะสำหรับการสอนการประยุกต์ความรู้ (application) หรือการแก้ปัญหา (problem solving) ถึงแม้ว่าคนส่วนใหญ่เข้าใจว่าการสอนบรรยายนั้นทำได้ง่าย ไม่ต้องเตรียมการมาก ในความเป็นจริง การสอนบรรยายให้ดีนั้นทำได้ยาก และต้องมีการเตรียมตัวที่ดีจึงจะสอนได้มีประสิทธิภาพ ในบทความนี้ผมจะขอเสนอเกร็ดความรู้เกี่ยวกับแนวทางการเตรียมตัวเพื่อสอนบรรยายให้มีประสิทธิภาพ

ก่อนอื่นอาจารย์ต้องหาข้อมูลสำคัญ 5 ประการ เกี่ยวกับการบรรยาย ได้แก่

1. Content (เรื่องที่ต้องสอน) จะต้องสอนเรื่องอะไร มีความครอบคลุมมากน้อยแค่ไหน
2. Audience (ผู้เรียน) ต้องสอนใคร นักเรียนที่จะมาเรียนมีความรู้พื้นฐานของเรื่องที่จะสอนมากน้อยเพียงใด นักเรียนที่จะเข้าฟังบรรยายมีจำนวนเท่าไร
3. Goals (วัตถุประสงค์ของการสอน) เมื่อสิ้นสุดการสอนแล้ว อาจารย์คาดหวังว่านักเรียนจะได้อะไรจากการบรรยายนี้บ้าง
4. Environment (ลักษณะของห้องบรรยาย) ห้องบรรยายอยู่ที่ใด มีขนาดใหญ่หรือเล็กเกินไปสำหรับจำนวนนักเรียนหรือไม่ ห้องบรรยายมีอุปกรณ์อะไรบ้าง
5. Time (เวลา) มีเวลาในการสอนมากน้อยเพียงใด

เมื่อรู้ข้อมูลพื้นฐานเหล่านี้แล้ว อาจารย์สามารถเริ่มทำการเตรียมการสอนได้ โดยมีสิ่งที่ต้องคำนึงถึงในการเตรียมตัวตามลำดับต่อไปนี้

- ขั้นแรกให้ใช้เวลาไม่เกินครึ่งชั่วโมงทบทวนตำราพื้นฐาน (Standard introductory textbook) ของเรื่องที่จะทำการสอน เพื่อเป็นการฟื้นความจำว่าหัวข้อสำคัญในเรื่องนั้นมีอะไรบ้าง
- หลังจากนั้นให้ร่างแผนการสอนอย่างคร่าวๆว่าจะสอนเรื่องอะไรบ้าง และเรียงลำดับอย่างไร คำนึงโดยทั่วไปคือให้อาจารย์แบ่งเนื้อหาออกเป็น 3 กลุ่ม เนื้อหาที่นักเรียนต้องรู้ เนื้อหาที่นักเรียนควรรู้ และเนื้อหาที่นักเรียนอาจไม่จำเป็นต้องรู้ก็ได้ ให้อาจารย์วางแผนการสอนให้เน้นการให้ความรู้ในกลุ่มแรก และกลุ่มที่สอง ควรสอนเนื้อหาในกลุ่มที่ 3 เฉพาะเมื่อมีเวลามากเพียงพอ
- เขียนวัตถุประสงค์การสอน (Educational objectives) ที่มีความชัดเจน
- วางโครงสร้างของการบรรยายว่าจะใช้เวลาสอนในแต่ละหัวข้อนานเท่าไร มีกิจกรรม หรือ โจทย์ปัญหาอะไรบ้างที่สามารถใช้เรียกความสนใจของนักเรียนในแต่ละช่วงของการบรรยายแต่ละกิจกรรมใช้เวลาเท่าไร
- เขียนรายการของเรื่องที่ต้องทำการค้นคว้าเพิ่มเติมก่อนการบรรยาย และหาหนังสืออ่านเพิ่มเติมเพื่อทำให้ตนเองมั่นใจว่ามีความรู้ในเรื่องนั้นๆ แน่นพอที่จะสอนนักเรียนได้อย่างราบรื่นและสามารถตอบข้อสงสัยของนักเรียนได้

- หลังจากนั้นให้ลงรายละเอียดของแผนการสอนว่าจะพูดถึงอะไรบ้าง จะถามคำถามอะไรบ้าง จะใช้เวลาอย่างไรในการบรรยาย และจัดทำ slide ช่วยสอนตามความเหมาะสม การวางแผนการสอนในขั้นตอนนี้ต้องระมัดระวังอย่าสอนรายละเอียดมากเกินไปจนเกินกว่าที่นักเรียนจะรับไหว การสอนมากเกินไปได้หมายความว่านักเรียนจะเรียนรู้มาก การสอนที่ดีคือการสอนเนื้อหาในปริมาณที่พอเหมาะที่นักเรียนสามารถรับได้ โดยทั่วไปนักเรียนมักมีความสนใจในเนื้อหาบรรยายไม่เกิน 20 นาที ดังนั้นอาจารย์จึงไม่ควรเตรียมการบรรยายให้นักเรียนนั่งฟังอย่างเดียวนานเกิน 20 นาที ทุก 15 – 20 นาทีของการบรรยายควรมีช่วงพักให้นักเรียนได้ทำกิจกรรมเพื่อกระตุ้นความสนใจของนักเรียนให้กลับมาตั้งใจฟังการบรรยายต่อ เช่นถามคำถาม ให้นักเรียนอภิปราย เป็นต้น

- และสิ่งที่สำคัญประการสุดท้ายคือ อาจารย์ควรคิดว่า จะประเมินผลการสอนของตนอย่างไร จะรู้ได้อย่างไรว่าตนสอนบรรยายได้อย่างมีประสิทธิภาพ มีข้อบกพร่องในเรื่องใดบ้าง เพื่อเป็นแนวทางในการพัฒนาการสอนของตนในอนาคต

Iramaneerat C. Structured lectures [Thai]. Medical Education Pamphlet 2007; 3(5):1.

การสอนบรรยายอย่างมีโครงสร้าง (Structured Lectures)

เชิดศักดิ์ ไชยมณีรัตน์

การสอนบรรยายนั้นเปรียบเหมือนการเล่าเรื่อง การเล่าเรื่องที่จะทำให้ผู้ฟังสนใจติดตามตั้งแต่ต้นจนจบและมีความเข้าใจในเรื่องนั้นๆมักต้องอาศัยการวางแผนที่ดี และมีการจัดโครงสร้างของการดำเนินเรื่องอย่างเป็นระเบียบ แบ่งออกเป็น 3 ส่วนหลักด้วยกัน คือ การเปิดฉาก การดำเนินเรื่อง และตอนจบ ในทำนองเดียวกัน การสอนบรรยายก็ควรมีการจัดโครงสร้างให้ดีทำให้นักเรียนติดตามเรื่องราวได้อย่างไม่สับสน โดยแบ่งการบรรยายออกเป็น 3 ส่วน คือ

1. Context บอกว่าจะสอนเรื่องอะไร มีความสำคัญอย่างไร เกี่ยวเนื่องกับความรู้ที่นักเรียนมีอยู่เดิมอย่างไร
2. Content บอกถึงเนื้อหาหลักที่ตั้งใจจะให้นักเรียนเรียนรู้
3. Closure สรุปให้นักเรียนฟังอีกครั้งว่าประเด็นสำคัญที่ได้สอนไปนั้นมีอะไรบ้าง

ในบทความนี้จะขอขยายความองค์ประกอบของการบรรยายทั้ง 3 ส่วนนี้เพื่อเป็นแนวทางให้อาจารย์สามารถเตรียมการสอนบรรยายได้อย่างมีประสิทธิภาพ

Context การเปิดฉากการบรรยายประกอบด้วยสิ่งสำคัญหลายประการด้วยกัน คือ

- การเรียกความสนใจของนักเรียน
- แนะนำหัวข้อที่จะทำการสอน
- ทบทวนความรู้พื้นฐานที่สำคัญต่อการทำความเข้าใจในเรื่องที่จะบรรยาย
- อธิบายความสำคัญของเนื้อหาว่าเกี่ยวข้องกับเวชปฏิบัติอย่างไร เพื่อกระตุ้นให้นักเรียนสนใจที่จะเรียนรู้
- แจ้งวัตถุประสงค์การเรียนรู้ (educational objectives) ให้นักเรียนรู้ว่าอาจารย์คาดหวังให้นักเรียนมีความรู้หรือความสามารถในเรื่องใดภายหลังจากการบรรยายแล้ว

Content เมื่อนักเรียนมีความพร้อมที่จะรับความรู้ใหม่แล้ว อาจารย์ก็สามารถเริ่มบรรยายเนื้อหาที่ต้องการได้ ในขั้นตอนนี้มีสิ่งที่จะต้องคำนึงถึงหลายประการด้วยกัน ได้แก่

- อาจารย์ต้องควบคุมเนื้อหาที่จะบรรยายให้มีความเหมาะสม ปัญหาที่พบบ่อยมากในการสอนบรรยายคือ อาจารย์ต้องการบรรยายเนื้อหามากเกินไป เกินกว่าที่นักเรียนจะรับได้ การสอนบรรยายนั้นไม่ใช่การบอกทุกอย่างที่มีในตำราให้นักเรียนท่อง หากแต่เป็นการชี้ประเด็นสำคัญให้นักเรียนทำความเข้าใจ
- อาจารย์ควรจัดให้การบรรยายมีการเปลี่ยนแปลงรูปแบบเป็นระยะ การบรรยายที่ประกอบด้วยคำพูดเพียงอย่างเดียวตลอด 1 ชั่วโมงเป็นการบรรยายที่น่าเบื่อ และไม่มีนักเรียนคนไหนจะตั้งใจฟังได้ตลอด การบรรยายที่ดีต้องมีการปรับเปลี่ยนกิจกรรมเป็นระยะๆ เพื่อกระตุ้นให้นักเรียนไม่ง่วงนอน เช่น มีการอภิปรายปัญหาผู้ป่วย มีการแสดงวิดิทัศน์การทำหัตถการ การสลับให้นักเรียนออกมาแสดงความคิดเห็นหน้าชั้น เป็นต้น
- อาจารย์ต้องจัดเนื้อหาให้เป็นระเบียบ การสอนเนื้อหาที่เป็นระเบียบทำให้นักเรียนติดตามเนื้อหาได้ง่าย ไม่สับสน การจัดเนื้อหาให้เป็นระเบียบทำได้หลายรูปแบบ เช่น การบรรยายตามลำดับการเกิดเรื่องก่อนหลัง การเปรียบเทียบข้อดี ข้อเสีย การเรียงลำดับโรคตามระบบอวัยวะ หรือตามลำดับความพบบ่อย เป็นต้น
- การบรรยายที่ดีต้องกระตุ้นให้นักเรียนใช้ความคิด การบอกเนื้อหาให้นักเรียนฟังอย่างเดียวโดยไม่ต้องคิดมีแต่จะทำให้ให้นักเรียนฟังเข้าหูซ้ายทะลุหูขวา ไม่ได้ก่อให้เกิดความเข้าใจ

Closure เมื่ออาจารย์ได้บรรยายเนื้อหาจนจบแล้ว อาจารย์ต้องทำการสรุปประเด็นสำคัญให้นักเรียนฟังอีกครั้ง เทคนิคที่ใช้ได้ง่ายคือการทบทวนวัตถุประสงค์การสอนที่ได้แจ้งไปเมื่อต้นชั่วโมง แล้วอภิปรายกับนักเรียนว่านักเรียนเกิดการเรียนรู้ตามวัตถุประสงค์ที่ตั้งไว้หรือไม่ อย่างไร

Iramaneerat C. Lecture handout [Thai]. Medical Education Pamphlet 2008; 4(5): 1.

การเตรียมเอกสารประกอบการสอนบรรยาย (Lecture Handout)

เชิดศักดิ์ ไอรมนรัตน์

เอกสารประกอบการบรรยายที่ดีสามารถทำให้การสอนเป็นไปได้อย่างราบรื่นและนักเรียนเข้าใจบทเรียนได้ดีขึ้น เอกสารประกอบการบรรยายสามารถใช้เป็นได้ทั้งเอกสารให้ข้อมูลเกี่ยวกับเนื้อหาที่ทำการบรรยาย และ เป็นสื่อการสอนที่ให้นักเรียนมีส่วนร่วมในการเรียนมากขึ้น

1. เอกสารการสอนเพื่อให้อ่าน อาจารย์สามารถเตรียมเอกสารการสอนชนิดนี้ได้ 2 ลักษณะ คือ เอกสารที่มีเนื้อหาครบถ้วนตามการบรรยาย (full lecture note) และ เอกสารที่มีเนื้อหาไม่ครบถ้วนตามการบรรยาย (partial lecture note) เอกสารการสอนแบบแรกมีข้อดีคือมีความครอบคลุมเนื้อหาที่อาจารย์ต้องการบรรยายทั้งหมด นักเรียนไม่ต้องพะวงกับการจดบันทึกเลย แต่การแจกเอกสารการสอนที่มีเนื้อหาครบถ้วนนี้เป็นการส่งเสริมให้นักเรียนไม่สนใจการบรรยาย นักเรียนจะสนใจเพียงจะเอาเอกสารกลับไปอ่าน หากสามารถฝากเพื่อนเอาเอกสารให้ได้ก็ไม่จำเป็นต้องเข้าฟังบรรยายเลย เพื่อหลีกเลี่ยงปัญหาการไม่ตั้งใจฟังบรรยายเนื่องจากมีเอกสารการสอนที่ครบถ้วน อาจารย์สามารถใช้เอกสารการสอนที่มีเนื้อหาไม่ครบถ้วนได้ใน 2 ลักษณะ คือ

(1) เอกสารที่มีเฉพาะหัวข้อหรือคำศัพท์ที่สำคัญในเรื่องที่บรรยาย หรือมีเนื้อหาของการบรรยายเฉพาะส่วนที่สำคัญ หรือมีแผนภาพที่สำคัญที่นักเรียนไม่สามารถจดได้ทันในระหว่างการบรรยาย หรือ

(2) เอกสารที่ไม่มีเนื้อหาของการบรรยาย แต่มีเนื้อหาที่ช่วยขยายความเข้าใจจากการบรรยายโดยมีตัวอย่างผู้ป่วย หรือ คำแนะนำเอกสารอ้างอิงให้นักเรียนไปศึกษาเพิ่มเติมหากสนใจ

2. เอกสารการสอนที่สร้างกิจกรรมให้นักเรียนมีส่วนร่วมในการเรียน อาจารย์สามารถพัฒนาเอกสารการสอนในรูปแบบต่างๆ ที่กระตุ้นให้นักเรียนได้คิดตามในการบรรยาย เอกสารในลักษณะนี้สามารถช่วยให้นักเรียนตั้งใจฟังบรรยายมากขึ้น ในขณะที่เดียวกับที่มีการใช้ความคิดวิเคราะห์ และสังเคราะห์เนื้อหาเพื่อทำกิจกรรมที่มีในเอกสารการสอน ตัวอย่างเช่น

- เอกสารการสอนที่มีช่องว่างให้นักเรียนเติมคำ หรือข้อความ
- เอกสารการสอนที่มี แผนภาพที่ไม่สมบูรณ์ ซึ่งนักเรียนต้องเติม ลูกศร หรือ องค์ประกอบบางอย่างลงไป
- เอกสารการสอนที่มีคำถามแทรกเป็นระยะๆ ซึ่งเว้นที่ว่างไว้ให้นักเรียนเติมคำตอบซึ่งได้จากการอภิปรายกันใน

ห้องเรียน

อีกประเด็นหนึ่งที่มีความสำคัญในการเตรียมเอกสารการสอนคือการเลือกเวลาที่จะแจกจ่ายเอกสารการสอนที่เหมาะสม เอกสารการสอนที่ต้องการลดภาระของนักเรียนในการจดบันทึก หรือ เอกสารการสอนที่มีกิจกรรมให้นักเรียนทำในระหว่างบรรยาย ควรแจกเมื่อเริ่มบรรยาย, เอกสารการสอนที่เน้นการขยายความเข้าใจของนักเรียนโดยยกตัวอย่างผู้ป่วย หรือชี้แนะเอกสารอ้างอิงให้นักเรียนไปศึกษาเพิ่มเติมควรแจกเมื่อสิ้นสุดการบรรยาย, สำหรับเอกสารการสอนที่เป็นการทบทวนเนื้อหาซึ่งนักเรียนต้องใช้ในการทำความเข้าใจเรื่องที่จะบรรยาย หรือจะให้นักเรียนอภิปรายในชั้นเรียน ควรแจกให้นักเรียนศึกษาหลายวันก่อนที่จะบรรยายเพื่อให้นักเรียนมีเวลาอ่านเตรียมมาก่อนเข้าฟังบรรยาย

การจัดเตรียมสไลด์นำเสนอทางวิชาการ

Preparing academic slide presentation

ตริภพ เลิศบรรณพงษ์

*“A good presentation should be like a mini skirt: as short as possible to catch everyone’s attention,
and just long enough to cover what it needs to cover”*

Jacques Gruwez

บทนำ

การกำเนิดขึ้นของโปรแกรมไมโครซอฟต์ เพาเวอร์พอยท์ (Microsoft powerpoint) เมื่อ พ.ศ. 2530 โดยบริษัท ไมโครซอฟท์ ทำให้รูปแบบการใช้สไลด์เพื่อประกอบการนำเสนอเปลี่ยนแปลงไปอย่างมาก การนำเสนอแบบเดิม ๆ ด้วยการใช้เครื่องฉายแผ่นใสหรือแผ่นสไลด์ชนิด 35 มิลลิเมตร ไม่ปรากฏให้เห็นแล้วในปัจจุบัน⁽¹⁾ ทั้งยังมีการพัฒนาโปรแกรม นำเสนออีกมากมายให้สามารถเลือกใช้ได้อย่างหลากหลาย เช่น โปรแกรม keynote ของเครื่องคอมพิวเตอร์ Mac (Apple Inc.) หรือโปรแกรม Prezi หรือโปรแกรม google slide ซึ่งให้บริการสร้างสไลด์แบบออนไลน์ เป็นต้น อย่างไรก็ตามสิ่งที่ผู้สอนต้องระลึกไว้เสมอคือ ในการนำเสนอที่นำจดจำนั้น ผู้สอนคือคนสำคัญที่ต้องโดดเด่นที่สุด และสไลด์ประกอบการนำเสนอเหล่านี้เป็นเพียงตัวช่วยหรือสนับสนุนการนำเสนอเท่านั้น มิใช่การปล่อยให้สไลด์เป็นเสมือนผู้สอนและลดบทบาทของตนเองลงไปเป็นเพียงตัวประกอบการนำเสนอ ราวกับว่าแม้ไม่มีผู้สอน ผู้เรียนก็สามารถเรียนรู้ข้อมูลจากการอ่านสไลด์ที่อยู่บนจอภาพได้อย่างครบถ้วน⁽¹⁻³⁾

ย้อนกลับไปเมื่อเดือนกุมภาพันธ์ พ.ศ.2546 ภาวะสวอวกาศโคลัมเบียเกิดอุบัติเหตุ ระเบิดระหว่างทางที่กลับสู่พื้นโลกเนื่องจากโฟมที่ปกคลุมข้างหลุดแล้วไปกระทบทำให้เกิดความเสียหายกับปีกด้านซ้าย เป็นเหตุให้นักบินอวกาศทั้ง 7 คนบนกระสวยอวกาศเสียชีวิตทั้งหมด แม้สาเหตุหลักของการระเบิดจะเกิดจากความผิดพลาดทางเทคนิค แต่ก็มีข้อสังเกตว่าอีกปัจจัยหนึ่งที่สำคัญซึ่งทำให้การรับรู้ข้อผิดพลาดของทีมนักบินอวกาศเกิดความล่าช้า ก็คือ การนำเสนอข้อมูลของบริษัทผู้สร้างผ่านภาพสไลด์ที่ซับซ้อน ไม่ชัดเจน และอัดแน่นไปด้วยข้อมูลที่มากเกินไปจนผู้รับสารไม่สามารถรับรู้หรือเข้าใจเป้าหมายที่สื่อผ่านชุดสไลด์ดังกล่าวได้^(3, 4)

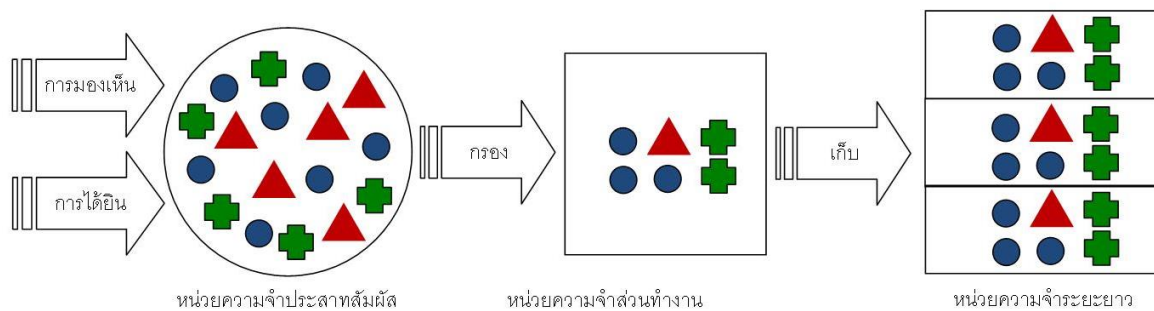
การเตรียมสไลด์เป็นทักษะที่จำเป็นอย่างยิ่งของครูแพทย์ในศตวรรษที่ 21⁽⁵⁾ จำเป็นต้องมีการเรียนรู้ ผูกพันและเข้าใจจิตวิทยาการเรียนรู้ของมนุษย์ บทความนี้ได้ทำการสรุปพื้นฐานสำคัญของการจัดเตรียมสไลด์ประกอบการนำเสนอ เพื่อให้ผู้อ่านสามารถสร้างสรรค์สไลด์ที่เรียบง่าย แต่มีคุณภาพ และสามารถโน้มน้าวใจผู้เรียนให้เข้าใจเป้าหมายของผู้สอนได้อย่างมีประสิทธิภาพ

ข้อแนะนำก่อนการเตรียมสไลด์

1. ผู้เรียนรับรู้ได้อย่างไร

ผู้เรียนคือหัวใจสำคัญที่สุดของการสอนบรรยาย⁽²⁾ การที่อาจารย์ผู้สอนเข้าใจพื้นฐานการรับรู้ของผู้เรียนจะทำให้สามารถเตรียมสไลด์ได้อย่างมีประสิทธิภาพ ทำให้ผู้เรียนสามารถเรียนรู้ได้อย่างเต็มที่ ทฤษฎี “หน่วยความจำส่วนทำงาน (working memory)” ของ George A. Miller แสดงให้เห็นถึงกลไกการรับรู้ของมนุษย์ขณะทำการสื่อสาร สมองสามารถรับรู้ได้ผ่านการทำงานของหน่วยความจำ 3 ระบบ^(6,7) (รูปที่1) ได้แก่

1. หน่วยความจำประสาทสัมผัส (sensory memory)
2. หน่วยความจำส่วนทำงาน (working memory)
3. หน่วยความจำระยะยาว (long term memory)



รูปที่1 การรับรู้ของผู้เรียนผ่านหน่วยความจำหลัก 3 ชนิด (ภาพประกอบโดย ผศ. นพ.ตรีภพ เลิศบรรณพงษ์)

หน่วยความจำประสาทสัมผัส (sensory memory)

สมองของผู้เรียนเริ่มต้นรับรู้จากการรับสิ่งเร้าที่เป็นภาพหรือเสียงขณะฟังการสอนบรรยาย เข้าสู่หน่วยความจำประสาทสัมผัส และพยายามรับรู้สิ่งเร้าที่เข้ามาทั้งหมดโดยใช้เวลานับวินาที เพียง 1-2 วินาทีต่อการรับรู้แต่ละสิ่งเร้า การรับรู้ในขั้นตอนนี้จะไม่เกิดความเข้าใจในเนื้อหา หากแต่จะพยายามรับรู้สิ่งเร้าให้ได้มากที่สุดก่อนและส่งต่อข้อมูลเหล่านั้นไปยังหน่วยความจำส่วนทำงาน

หน่วยความจำส่วนทำงาน (working memory)

เนื่องจากสมองมีพื้นที่ในการทำงานจำกัดมาก เมื่อสมองรับรู้สิ่งเร้าจำนวนมากจากหน่วยความจำประสาทสัมผัสแล้ว ข้อมูลดังกล่าวจะถูกส่งต่อไปยังหน่วยความจำส่วนทำงานเพื่อทำการพิจารณาและคัดกรองว่าสิ่งเร้าใดจำเป็นหรือไม่จำเป็น ควรจะให้ความสนใจหรือไม่ และทำการแปลผลสิ่งเร้าที่กรองแล้วนั้น ให้กลายเป็นข้อมูลหรือความรู้ใหม่ เพื่อส่งไปเก็บในหน่วยความจำระยะยาวต่อไป หน่วยความจำส่วนงานนี้ถือว่าเป็นแก่นสำคัญที่สุดในการรับรู้ของสมอง

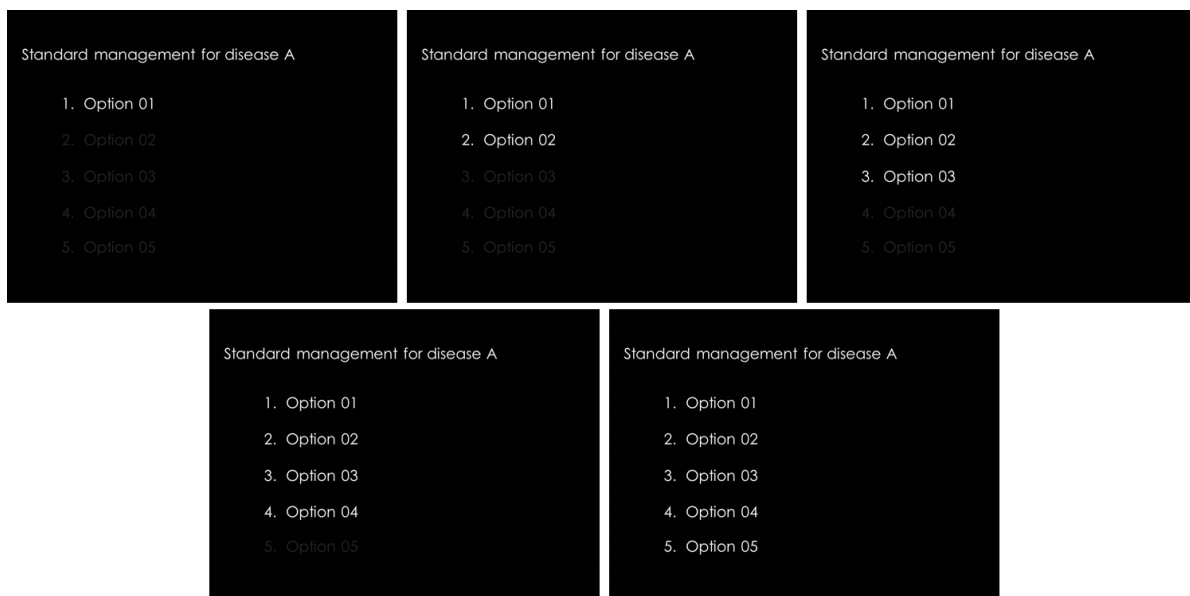
หน่วยความจำระยะยาว (long term memory)

หน่วยความจำส่วนนี้คล้ายกับเป็นห้องหรือฮาร์ดดิสสำหรับเก็บสะสมข้อมูลหรือความรู้ของมนุษย์ในระยะยาว ทำให้มนุษย์สามารถจดจำข้อมูล ความรู้ ประสบการณ์ในอดีต เพื่อนำกลับมาใช้ได้ตลอดเวลา หรือนำมาผสมผสานข้อมูลเดิมกับข้อมูลใหม่ที่ได้จากหน่วยความจำส่วนทำงาน แล้วปรับเปลี่ยนไปเป็นองค์ความรู้ใหม่เก็บไว้ต่อไป

เมื่อผู้สอนทราบกลไกการรับรู้ของมนุษย์เช่นนี้แล้ว จะเห็นได้ว่าผู้เรียนมีข้อจำกัดในการรับรู้ และหากจัดเตรียมสไลด์ได้ตามทฤษฎีดังกล่าวข้างต้น ผู้เรียนก็จะสามารถเรียนรู้ได้ดีที่สุด ดังนั้นการเตรียมสไลด์ที่ดีจึงต้องยึดหลักดังต่อไปนี้

1. ข้อมูลในสไลด์ต้องสอดคล้องกับเนื้อหาการสอน (รูปที่2)

การรับรู้ของผู้เรียนเป็นการรับรู้ที่ซับซ้อน เนื่องจากต้องใช้ไฮตัสสัมผัส 2 อย่างในเวลาเดียวกัน พร้อม ๆ กัน ได้แก่ การมองเห็นและการได้ยิน ด้วยเหตุนี้การทำสไลด์ที่ดีจึงไม่ควมใส่ข้อมูลทุกอย่างไว้ทั้งหมดในสไลด์เดียว แต่ควรใส่ข้อมูลให้สอดคล้องกับเนื้อหาการสอนที่กำลังดำเนินการบรรยายอยู่ โดยยึดหลักง่าย ๆ ในการทำสไลด์ คือ **“หนึ่งสไลด์ ต่อ หนึ่งประเด็นที่ต้องการสอน”**^(1, 8) ตัวอย่างเช่น หากผู้สอนต้องการแสดงหลักการรักษาโรค 5 ข้อ ก็ควรให้หลักการรักษาปรากฏให้เห็นทีละข้อ มากกว่าให้ผู้เรียนเห็นหลักการรักษาทั้ง 5 ข้อในครั้งเดียว เพื่อให้ผู้เรียนตั้งใจจดจำ และรับรู้ตามลำดับ ทีละข้อ เป็นต้น



รูปที่2 การนำเสนอโดยใช้หลัก “หนึ่งสไลด์ ต่อ หนึ่งประเด็นที่ต้องการสอน” โดยให้ตัวเลือกการรักษาค่อย ๆ ปรากฏขึ้นทีละข้อตามลำดับ มากกว่าให้ผู้เรียนเห็นในครั้งเดียว และใช้สไลด์ทั้งหมดรวม 5 สไลด์ในการสอนบรรยาย ไม่ใช่สไลด์เดียว

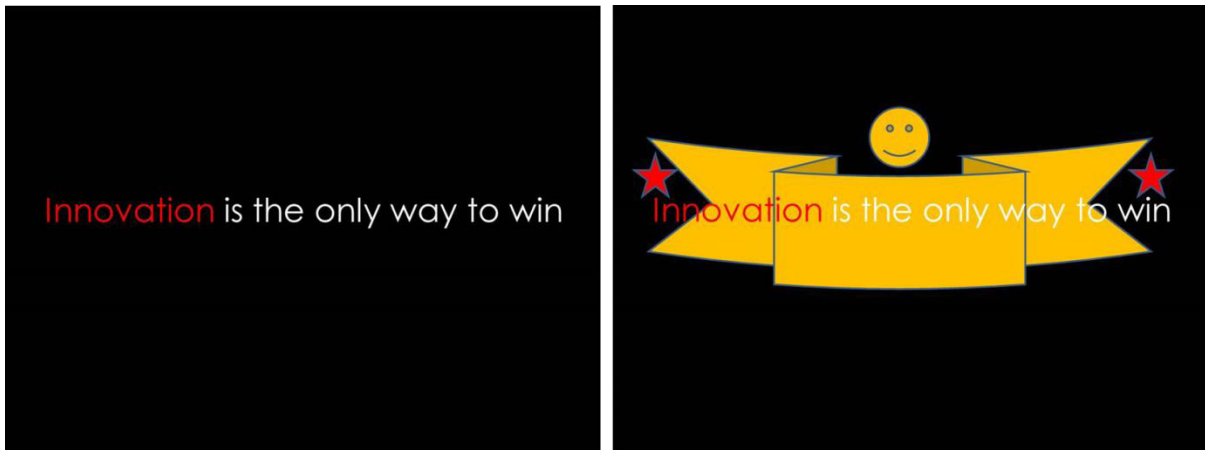
(ภาพประกอบโดย ผศ. นพ.ตรีภพ เลิศบรรณพงษ์)

2. ข้อมูลในสไลด์ต้องเป็นแก่นของเรื่องที่จะสอนและมีองค์ประกอบเท่าที่จำเป็น

ปัญหาหนึ่งที่พบบ่อยเมื่อสอนบรรยาย คือ ผู้สอนใส่ตัวอักษรเข้าไปในสไลด์มากเกินไป ผู้สอนส่วนใหญ่มักไม่ทราบว่า การอ่านตัวอักษรบนสไลด์นั้นต้องใช้ไฮตัสประสาทรับรู้ทั้งการมองเห็นและการได้ยินพร้อม ๆ กัน

ดังนั้นเมื่อผู้เรียนต้องอ่านตัวอักษรบนไลต์ จะรับรู้การสื่อสารที่ผู้สอนบรรยายได้ลดลงหรือไม่ได้เลย โดยเฉพาะอย่างยิ่งเมื่อมีการเขียนเป็นประโยคยาว ๆ หลายประโยค ดังนั้นสไลด์ที่ดีต้องมีตัวอักษรน้อยที่สุดเท่าที่จำเป็น เพื่อให้ผู้เรียนได้ใช้สายตาประสาทการได้ยิน รับฟังคำอธิบายจากผู้สอนให้ได้มากที่สุด และใช้สายตาประสาทการมองเห็นในการมองสไลด์เพื่อประกอบความเข้าใจเท่านั้น⁽⁶⁾

นอกจากนี้ผู้สอนต้องทราบข้อเท็จจริงที่ว่า **สมองของมนุษย์มีความสามารถในการรับรู้สิ่งเราได้จำกัด เฉลี่ยเพียง 4-7 สิ่งเราต่อการรับรู้หนึ่งครั้ง**^(6,7) ด้วยเหตุนี้การใส่องค์ประกอบที่ไม่จำเป็นเข้าไปในสไลด์มากเกินไป โดยไม่จำเป็น เช่น ตัวการ์ตูนตกแต่ง ภาพหรือฉากหลังที่ไม่สอดคล้องกัน เป็นต้น ก็เป็นอีกประเด็นหนึ่งที่รบกวนการรับรู้ของผู้เรียนเช่นเดียวกัน ด้วยเหตุนี้ผู้สอนต้องไม่ใส่องค์ประกอบในสไลด์มากเกินไป และต้องเน้นเฉพาะส่วนที่จำเป็นเท่านั้น (รูปที่3)



รูปที่3 ผู้เรียนต้องการความเรียบง่ายของสไลด์การนำเสนอ ภาพซ้ำๆ มีองค์ประกอบเท่าที่จำเป็น ภาพขวา มีองค์ประกอบที่เกินจำเป็นและรบกวนการรับรู้ของผู้เรียน (ภาพประกอบโดย ผศ. นพ.ตรีภพ เลิศบรรณพงษ์)

2. สร้างเป้าหมายให้ชัดเจน

ภายหลังเสร็จสิ้นการนำเสนอหลายครั้ง มักจะได้รับเสียงตอบกลับจากผู้เรียนว่า “ไม่รู้เรื่อง ไม่เข้าใจ และไม่ทราบว่าผู้สอนจะสื่อประเด็นสำคัญใดให้รับรู้ เราอาจเรียนการนำเสนอเหล่านั้นว่า “การนำเสนอที่สูญเปล่า” เพราะเป็นการทิ้งทรัพยากรต่าง ๆ อย่างสูญเปล่าทั้งเวลาของผู้สอนและผู้เรียน เวลาของเจ้าหน้าที่ผู้ควบคุมห้องเรียน สาธารณูปโภคที่อยู่ภายในห้องเรียน เวลาและค่าใช้จ่าย รวมทั้งเสียความรู้สึกของผู้เรียน โดยไม่มีประโยชน์ใด ๆ เกิดขึ้นจากการนำเสนอครั้งนั้นเลยแม้แต่ชนิดเดียว⁽⁶⁾

สิ่งที่ผู้สอนต้องจำไว้เสมอคือ การนำเสนอที่ขาดเป้าหมายเท่ากับไม่ได้นำเสนอ เป้าหมายของการนำเสนอคือการทำให้ผู้เรียนรับรู้แก่นสำคัญของเรื่องราวที่นำเสนอให้ได้ ด้วยเหตุนี้ผู้สอนต้องคิด พิจารณาและกลั่นกรองเนื้อหาที่จะนำเสนออย่างดีก่อน และตั้งคำถามกับตนเองเสมอว่า “แก่นแท้ของเรื่องราวที่จะนำเสนอและต้องการให้ผู้เรียนจดจำได้ไปตลอด” คืออะไร^(1, 3, 9)

3. เริ่มต้นด้วยการสร้างเรื่องราวในกระดาก่อนเสมอ

ด้วยความชำนาญ เชี่ยวชาญหรือประสบการณ์อันมากมายของผู้สอน สิ่งที่คุณสอนหลายคนมักจะทำเสมอเมื่อต้องสอนบรรยาย คือ การนั่งลง เปิดคอมพิวเตอร์ แล้วเปิดโปรแกรมสร้างสไลด์นำเสนอ จากนั้นก็เริ่มลงมือสร้างสไลด์ประกอบการนำเสนอในทันทีโดยไม่รีรอ การทำเช่นนี้มักจะทำให้เกิดสไลด์ที่ด้อยคุณภาพและเสียเวลาในการทำสไลด์โดยใช่เหตุ เนื่องจากสไลด์เหล่านี้มักจะมีเนื้อหามากมายเกินความจำเป็นและมีแนวโน้มจะเป็นเอกสารประกอบการบรรยายเสียมากกว่า ทั้งยังเป็นการเสียเวลาผู้สอนในการค้นหาตัวเลือกร่างต่าง ๆ ในโปรแกรมหรือการแก้ไขที่อาจเกิดขึ้นอย่างมากโดยไม่จำเป็น

สิ่งที่คุณสอนควรกระทำก่อนการสร้างสไลด์ประกอบการสอน คือ การคิดถึงแก่นแท้ของการนำเสนอเรื่องนั้น ๆ ให้เรียบร้อยก่อน แล้วสร้างเรื่องราวลงในกระดาษคล้าย ๆ กับการสร้างบทภาพ (storyboard)^(1, 5, 6, 10) เพื่อให้ผู้สอนได้มีโอกาสกลั่นกรอง แก้ไข และเรียบเรียงเรื่องราว โดยมีเป้าหมายเพื่อให้ผู้เรียนได้รับรู้แก่นแท้ของการบรรยายเป็นลำดับ ลดสิ่งรบกวนที่อาจเกิดขึ้นในสไลด์และสร้างสรรคให้เกิดเรื่องราวที่น่าจดจำหลังจบการบรรยาย เมื่อทำเสร็จแล้วจึงไปสร้างสไลด์ประกอบการนำเสนอด้วยโปรแกรมคอมพิวเตอร์เป็นขั้นตอนสุดท้าย กระบวนการเช่นนี้จะทำให้ได้สไลด์ประกอบการนำเสนอที่ดีและดูง่ายที่สุด และเสียเวลาน้อยที่สุดในการทำสไลด์ด้วยเช่นเดียวกัน

4. เขียนเอกสารประกอบการบรรยาย หากต้องการสื่อรายละเอียดของเนื้อหา

การนำเสนอหลายครั้งมักเต็มไปด้วยตัวอักษรจำนวนมากบนสไลด์ ซึ่งมากมายจนทำให้ผู้เรียนไม่สามารถจับแก่นแท้ของการบรรยายได้ ไม่ว่าเนื้อหาอันมากมายนั้นจะมีวัตถุประสงค์เพื่อไม่ให้ผู้สอนลืมเนื้อหาสำคัญ หรือเพราะผู้สอนต้องการเน้นรายละเอียดของเนื้อหาก็ตาม สิ่งนี้มักทำให้การนำเสนอครั้งนั้นล้มเหลวและไม่สร้างความประทับใจให้ผู้เรียนแต่อย่างใด หากผู้สอนต้องการให้ผู้เรียนมีความเข้าใจเนื้อหาที่ซับซ้อนหรือรายละเอียดของเนื้อหาที่มากขึ้น สิ่งที่คุณควรทำไม่ใช่การอัดข้อมูลหรือเนื้อหาเหล่านั้นเข้าไปในสไลด์ประกอบการนำเสนอเพียงแผ่นเดียวหรือทุกแผ่น แต่แนะนำให้สร้างเอกสารประกอบการบรรยายเพิ่มขึ้นอีกฉบับหนึ่ง เพื่อให้ผู้เรียนสามารถนำไปอ่านเพิ่มเติมและสร้างความเข้าใจด้วยตนเองได้ในภายหลัง⁽¹⁾

ข้อเสนอแนะสำหรับการเตรียมสไลด์

เมื่อผู้สอนได้เตรียมตัวก่อนการสร้างสไลด์เรียบร้อยแล้ว เนื้อหาต่อไปนี้จะเสนอแนะเบื้องต้นในการสร้างสไลด์อย่างมีประสิทธิภาพ และสร้างความประทับใจแก่ผู้เรียนเมื่อสิ้นสุดการบรรยาย โดยยึดหลักง่าย ๆ คือ "การสร้างสไลด์ให้ดูเรียบง่ายที่สุด"^(5, 8)

1. พื้นหลัง

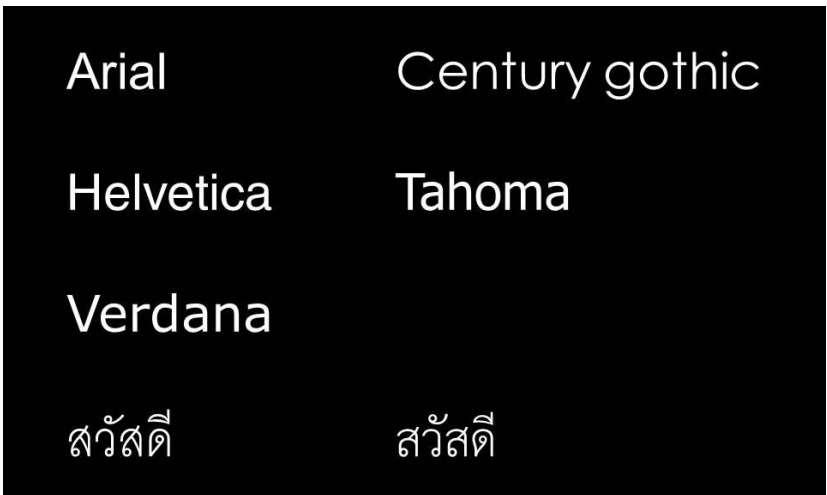
ในการสร้างสไลด์ พื้นหลังเปรียบเสมือนพื้นที่ว่างสำหรับการใส่องค์ประกอบต่าง ๆ ลงไปในสไลด์ ด้วยเหตุนี้ผู้สอนต้องไม่ทำให้พื้นหลังดึงดูดความสนใจของผู้ฟัง แข่งกับเนื้อหาที่คุณสอนนำเสนอ เช่น พื้นหลังที่ถูกตกแต่งด้วยภาพสัญลักษณ์ หรือสีต่าง ๆ จนทำให้โดดเด่นเกินกว่าที่ผู้เรียนจะจดจ่ออยู่กับเนื้อหาบนสไลด์ได้ หลักการพื้นฐานง่าย ๆ ของการเลือกพื้นหลังคือ หลีกเลี่ยงการใช้พื้นหลังสำเร็จรูปที่โปรแกรมเตรียมไว้ให้ เนื่องจากพื้นหลังสำเร็จรูปดังกล่าวไม่ได้สร้างความสนใจให้กับผู้ฟังเพราะใคร ๆ ก็เคยเห็นพื้นหลังแบบนี้มาแล้วทั้งนั้น แนะนำให้เลือกใช้พื้นหลังสีเดียวที่กำหนดขึ้นเอง เช่น สีดำ หรือ สีขาว เป็นต้น จากข้อมูลข้างต้น ผู้เรียนมีข้อจำกัดในการรับรู้สิ่งเร้าในแต่ละครั้ง ด้วยเหตุนี้พื้นหลังจึงไม่ควรเป็นปัจจัยรบกวนการเรียนรู้ของผู้เรียน การใช้สัญลักษณ์ของสถาบันหรือการแสดงชื่อพร้อมตำแหน่งของ

ผู้สอนควรจำกัดไว้เพียงแค่สไลด์แรกเท่านั้น ไม่มีความจำเป็นต้องใส่สัญลักษณ์ของสถาบันหรือตัวอักษรแสดงสิ่งใด ๆ ซึ่งไม่ได้เกี่ยวข้องกับเป้าหมายการนำเสนอในสไลด์ถัดไป เพราะผู้เรียนได้ทราบตั้งแต่สไลด์แรกแล้วว่า ผู้สอนเป็นใครและมาจากที่ไหน เช่นเดียวกับการฟังภาพบนพื้นหลังในทุกสไลด์ก็เป็นสิ่งรบกวนการรับรู้ของผู้เรียนเช่นเดียวกัน^(1, 6)

การใช้พื้นหลังสีเดียว นอกจากจะไม่รบกวนการรับรู้ของผู้เรียนแล้ว สีของพื้นหลังยังสื่ออารมณ์การนำเสนอได้อีกด้วย สีพื้นหลังที่แนะนำสำหรับการทำสไลด์ ได้แก่ สีดำ หรือ สีขาว ซึ่งจากเป็นสีที่ผู้สอนสามารถสร้างความแตกต่างได้มากที่สุดบนสไลด์ เพราะสีทั้งสองนี้เปรียบเสมือนการไม่มีสีนั่นเอง การใช้สีดำเป็นพื้นหลังจะให้ความรู้สึกที่เป็นทางการและเหมาะกับการสอนบรรยายในห้องเรียนขนาดใหญ่หรือการประชุมวิชาการซึ่งห้องมักจะมืด เพราะสีดำจะช่วยให้ตัวอักษรเด่นชัดขึ้น ขณะที่การใช้พื้นหลังสีขาวจะให้ความรู้สึกเป็นกันเอง ไม่เป็นทางการ และเหมาะกับห้องเรียนขนาดเล็กหรือห้องเรียนสว่างหรือใช้แสงธรรมชาติ การใช้พื้นหลังสีขาวยังทำให้ง่ายต่อการวางรูปภาพที่มีพื้นหลังสีขาว เนื่องจากไม่ต้องเสียเวลาไปลบพื้นหลังออก^(1, 2, 5)

2. ตัวอักษร

การสอนบรรยายทางการแพทย์ส่วนใหญ่มักใช้ตัวอักษรภาษาอังกฤษพิมพ์ลงในสไลด์ ดังนั้นการเลือกใช้ตัวอักษรที่เหมาะสมจะสามารถทำให้ผู้สอนสามารถสื่อสารได้อย่างมีประสิทธิภาพ ทั้งนี้ตัวอักษรที่แนะนำให้ใช้เพื่อประกอบการนำเสนอบนสไลด์และเป็นที่ยอมรับ ได้แก่ **ตัวอักษรชนิด san serif** (รูปที่4) เช่น Arial, Century Gothic, Helvetica, Tahoma, Verdana เป็นต้น^(1, 5, 10) เนื่องจากอักษรกลุ่มนี้จะมีลักษณะใหญ่ หนา และอ่านง่าย เหมาะกับการสื่อสารด้วยข้อความหรือวลีสั้น ๆ ทั้งนี้แนะนำว่าในการทำสไลด์ประกอบการนำเสนอแต่ละครั้งควรใช้ตัวอักษรไม่เกิน 2 ชนิด (สำหรับหัวเรื่อง 1 ชนิดและสำหรับข้อความอีก 1 ชนิด) และใช้ตัวพิมพ์เล็กจะอ่านง่ายกว่าตัวพิมพ์ใหญ่ ทั้งนี้เพื่อลดการเพิ่มการรับรู้ของผู้เรียน^(1, 2, 10) สำหรับผู้นิพนธ์นั้น เมื่อต้องทำสไลด์ประกอบการบรรยาย นิยมใช้ตัวอักษรชนิด Century Gothic หรือ Helvetica โดยจะเลือกใช้เพียงชนิดเดียวเท่านั้น และในกรณีตัวอักษรภาษาไทยนั้น ยังไม่มีข้อเสนอแนะที่ชัดเจน ผู้นิพนธ์นิยมใช้ชุดอักษร cordia หรือ TH sarabun เนื่องจากมีลักษณะโค้งมนและอ่านง่าย



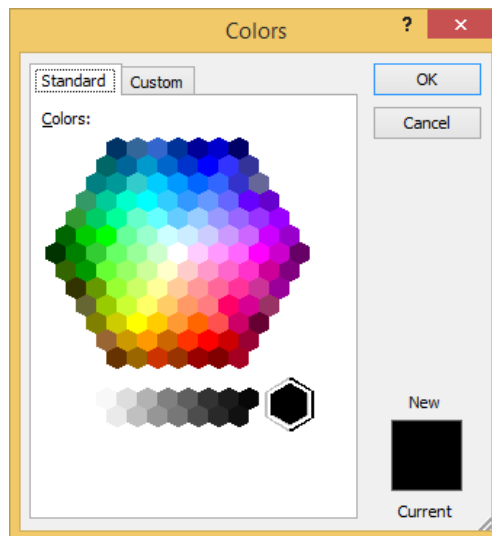
รูปที่4 แสดงตัวอักษรภาษาอังกฤษชนิด san serif ที่แนะนำให้ใช้ในการเตรียมสไลด์ และตัวอักษรภาษาไทย (ซ้าย:cordia new และขวา: TH sarabun) ที่ผู้นิพนธ์นิยมใช้ในการเตรียมสไลด์ (ภาพประกอบโดย ผศ. นพ.ตรีภพ เลิศบรรณพงษ์)

สไลด์ประกอบการนำเสนอไม่ได้ถูกออกแบบมาให้เขียนข้อความประโยค หากแต่ถูกออกแบบให้นำเสนอเฉพาะคำหรือวลีที่สำคัญ⁽⁸⁾ เพื่อช่วยสนับสนุนการรับรู้ของผู้เรียนขณะสอนบรรยาย และตัวอักษรที่เขียนขึ้นไม่ควรยาวมากจนต้องขึ้นบรรทัดใหม่ ขณะที่จำนวนบรรทัดที่เหมาะสมสอดคล้องกับการรับรู้ของผู้เรียนควรจำกัดไว้ที่ไม่เกิน 6 บรรทัดหรือ 6 กลุ่มข้อมูลต่อ 1 สไลด์^(8, 10) ในกรณีต้องใส่แหล่งที่มาเพื่ออ้างอิง แนะนำให้ย่อจนสั้นมากพอจนเป็นวลีที่เข้าใจได้ โดยไม่ต้องคัดลอกต้นฉบับออกมาทั้งหมด

ผู้สอนหลายคนตั้งคำถามถึงขนาดของตัวอักษรว่า ขนาดเท่าไรจึงจะเหมาะสมสำหรับการทำสไลด์ประกอบการนำเสนอ และพบว่ามีย่อแนะนำมากมายเกี่ยวกับข้อสงสัยนี้ บ้างก็แนะนำตัวอักษรว่าอย่างน้อยต้องมีขนาดไม่ต่ำกว่า 28⁽¹¹⁾ บางคนก็แนะนำว่าต้องมีขนาด 30 เป็นอย่างต่ำ⁽¹¹⁾ อย่างไรก็ตามข้อแนะนำที่เหมาะสมมากที่สุดของขนาดตัวอักษรที่เหมาะสมที่สุดในการสอนบรรยาย คือ “ขนาดตัวอักษรที่มีขนาดใหญ่พอให้ผู้เรียนที่นั่งอยู่ข้างหลังสุดของห้องเรียนมองเห็นได้ชัดเจน”^(5, 6, 9) เพราะผู้สอนไม่ได้สอนบรรยายในห้องเรียนขนาดเดียวเสมอไป บางครั้งนำเสนอให้ผู้เรียนเพียง 30 คน บางครั้งต้องนำเสนอในหอประชุมขนาดความจุ 300 คน การใช้ตัวอักษรสำหรับการนำเสนอจึงไม่ควรถูกกำหนดให้จำกัดเพียงขนาดเดียวเท่านั้น หากแต่ผู้สอนต้องพึงระลึกอยู่เสมอว่าขนาดตัวอักษรที่เหมาะสมในสไลด์นั้นต้องทำให้ผู้เรียนที่นั่งด้านหลังสุดมองเห็นเสมอ

3. สี

หลายครั้งเมื่อเสร็จสิ้นการนำเสนอ ผู้สอนมักได้ยินเสียงบ่นจากผู้เรียนว่า สไลด์ไม่ชัดเจน มองเห็นยาก หรือต้องเพ่งสายตาวงมากเพื่อมองสไลด์ ปัญหาเหล่านี้เป็นผลจากการที่ผู้สอนเลือกใช้โทนสีไม่เหมาะสมบนสไลด์ประกอบการนำเสนอ โปรแกรมการทำสไลด์ประกอบการนำเสนอส่วนใหญ่จะมีวงล้อสี (รูปที่ 2) มาให้เลือกอย่างมากมายและเพียงพอที่จะจัดเรียงโทนสีที่เข้มที่สุดไว้ด้านนอกสุด โดยโทนสีจะค่อย ๆ จางลงในทิศเข้าสู่ศูนย์กลาง และยังสามารถปรับแต่งความเข้มจาง ความอึมครึมได้อีกด้วย⁽⁵⁾ การเลือกใช้โทนสีจึงมีความสำคัญอย่างมากสำหรับการทำสไลด์ประกอบการนำเสนอเนื่องจากโทนสีที่แตกต่างกันใช้สื่ออารมณ์ที่แตกต่างกัน (ตารางที่ 1) การเข้าคู่สีที่เหมาะสมจะทำให้สไลด์ดูง่ายและสบายตาผู้เรียนมากขึ้น รวมทั้งการรู้จักชุดสีที่เหมาะสมก็จะทำให้สไลด์ดูน่าสนใจมากขึ้นเช่นกัน



รูปที่ 2 วงล้อสีในชุดโปรแกรมไมโครซอฟท์ (ภาพประกอบโดย ผศ. นพ.ตรีภาพ เลิศบรรณพงษ์)

ตารางที่ 1 ตัวอย่างความสัมพันธ์ของสีกับการสื่ออารมณ์⁽⁵⁾

สี	การสื่ออารมณ์
แดง	กล้าแสดงออก มีพลัง เร่งด่วน เน้น เกรี้ยวกราด รัก ความหลงใหล อันตราย สัญญาณหยุด ปีกาจร้าย ความตาย
เหลือง	ความสุข ความร่าเริง ความคาดหวังที่ดี ความสดชื่น ความสนุกสนาน แรงบันดาลใจ แสงแดด ฤดูร้อน
เขียว	ธรรมชาติ ความสมดุล ความกลมกลืน รักรักรัก สุขภาพดี ความสงบ โชคดี การเกิดใหม่ สัญญาณให้ไป
น้ำเงิน	ความเป็นมืออาชีพ ความสำเร็จ ความน่าเชื่อถือ เกียรติยศ ความจงรักภักดี เสรีภาพ ความสงบ ความรู้สึกเชิงบวก ความรู้สึกเศร้า
ม่วง	การไตร่ตรอง ความหรูหรา ความจงรักภักดี ความเจ็บแสบ ความคิดสร้างสรรค์ แรงบันดาลใจ ความมีจิตวิญญาณ
ส้ม	ความอบอุ่น ความเห็นใจ ความตื่นตัว ความกระตือรือร้น ความมีจิตวิญญาณ ความสนุกสนาน ฤดูใบไม้ร่วง
ชมพู	โรแมนติก ความเป็นผู้หญิง ความรัก สุขภาพ สุขสบาย
ดำ	อำนาจ ความยอดเยี่ยม ความรู้สึกเป็นทางการ ความเรียบง่าย มีพลัง ความตาย การสูญเสีย ความยากลำบาก ความเศร้าใจ
ขาว	ความบริสุทธิ์ ไร้เดียงสา ความสะอาด ความใหม่ ความง่าย ความกว้างใหญ่ ความหนาวเย็น
เทา	ความเป็นกลาง ความนับถือ ความอ่อนน้อมถ่อมตน ความคงที่ ฉลาด และความเรียบง่าย ความขุ่นมัวครั้ง ๆ กลาง ๆ ความเศร้าหมอง ความไม่ชัดเจน

การเลือกใช้โทนสีสำหรับการเตรียมสไลด์สามารถเลือกใช้ได้หลายรูปแบบ^(1, 5) ได้แก่

- **สีโทนเดียว (monochromatic)** ได้แก่ การเลือกใช้สีเพียงสีเดียว แต่หลายเฉดสี โดยเติมสีขาวเพื่อให้จางลง หรือ เติมสีดำเพื่อให้เข้มขึ้น ทำให้เกิดชุดสีที่น่าสนใจและดูเป็นมืออาชีพมากขึ้น
- **สีใกล้เคียง (analogous)** ได้แก่ การใช้ชุดสีในโทนใกล้เคียงกันของวงล้อ ทำให้เกิดความรู้สึกผสมผสานอย่างกลมกลืน เช่น ชุดสีม่วง แดง และส้ม แสดงถึงความมีพลัง ชุดสีฟ้า น้ำทะเล สีเขียวแก่ และสีใบตอง แสดงถึงความรู้สึกสงบ เป็นต้น
- **สีตรงข้าม (complementary)** ได้แก่ การใช้สีชุดที่อยู่ตรงข้ามกันของวงล้อสี ทำให้เกิดความรู้สึกแตกต่างอย่างกลมกลืน เช่น ชุดสีส้ม น้ำเงิน แสดงถึงความแข็งแกร่ง ชุดสีชมพู เขียว แสดงถึงความอ่อนโยนของสตรี เป็นต้น
- **สีสามเส้า (Triadic)** ได้แก่ ชุดสีที่มีระยะห่างในวงล้อสีในองศาที่เท่ากัน การใช้สีลักษณะนี้ช่วยเพิ่มความน่าสนใจของสไลด์มากขึ้น เนื่องจากสีทั้งสามมีอารมณ์ที่แตกต่างกันมาก เช่น ชุดสีแดง น้ำเงิน เหลือง สื่อถึงอารมณ์สนุกสนาน ชุดสีเขียวแก่ ม่วง และทอง สื่อถึงการเชิญชวน เป็นต้น

ปัจจัยหนึ่งที่ผู้สอนต้องคำนึงเกี่ยวกับการใช้สีอยู่เสมอ คือ **ผู้เรียน 1 ใน 12 คนมักมีปัญหาตาบอดสี⁽¹⁾** ดังนั้นเมื่อทำสไลด์เสร็จสิ้นแล้ว แนะนำให้พิมพ์สไลด์ออกมาพิจารณาแบบสเกลสีเทา (greyscale) หากมีส่วนใดของสไลด์ที่ไม่สามารถอ่านหรือมองเห็นได้ชัดเจน แนะนำให้เปลี่ยนโทนสีในตำแหน่งนั้น เพื่อป้องกันความบอพร่องในการรับรู้ของผู้เรียน

ที่ตาบอดสี ด้วยเหตุนี้จึงมีข้อแนะนำในการเตรียมสไลด์ด้วยชุดสีขาว เทา ดำ หรือแบบไม่มีสี (achromatic) เพื่อเอื้อประโยชน์ต่อผู้เรียนกลุ่มนี้ และอาจมีการเน้นเพียงบางส่วนของสไลด์ด้วยสีเดียวอื่น ๆ เสริมให้เด่นขึ้นมา เช่น สีแดง หรือ สีเหลือง เป็นต้น

นอกจากนี้ผู้สอนต้องมีความเข้าใจหลักการใช้สีโทนร้อนหรือเย็นด้วย กล่าวคือ สีโทนร้อน ได้แก่ สีแดง ส้ม เหลือง และน้ำตาล สีกลุ่มนี้มีความโดดเด่นสูงและมักให้ความรู้สึกลอยออกจากสไลด์ สีโทนเย็น ได้แก่ สีฟ้า น้ำเงิน เขียว ม่วง สีกลุ่มนี้มีความโดดเด่นต่ำ และมีแนวโน้มกลมกลืนไปกับพื้นหลัง ด้วยเหตุนี้ผู้สอนควรเลือกสีโทนร้อนเป็นองค์ประกอบของข้อความหรือใช้เน้นส่วนสำคัญของสไลด์ และเลือกสีโทนเย็นเป็นองค์ประกอบของพื้นหลัง^(5, 6) ตัวอย่างเช่น การใช้สีน้ำเงินเป็นพื้นหลัง และใช้ตัวอักษรสีเหลือง เป็นต้น

4. รูปภาพและวิดีโอ

ในขณะที่ฟังสอนบรรยาย สมอของผู้เรียนสามารถจดจำรูปภาพได้รวดเร็วและง่ายกว่าการอ่านตัวหนังสือ การใช้รูปภาพที่สอดคล้องกับเนื้อหาการบรรยายจึงดึงดูดใจและสร้างความน่าจดจำของเนื้อหาบรรยายได้ดีกว่าการใช้ตัวหนังสือปริมาณมากในสไลด์^(3, 5) ด้วยเหตุนี้การใช้รูปภาพที่เหมาะสมจึงเป็นข้อแนะนำพื้นฐานสำหรับการเตรียมสไลด์ที่มีประสิทธิภาพ อย่างไรก็ตามการใช้รูปภาพคุณภาพต่ำที่ไม่ผ่านการปรับแต่ง และไม่สอดคล้องกับเรื่องราวที่บรรยาย กลับนำไปสู่การสอนที่ล้มเหลว หลายครั้งที่มักจะพบว่าผู้สอนไม่ใส่ใจกับการเตรียมรูปภาพในสไลด์ ทำให้สไลด์ที่สร้างขึ้นขาดความเป็นเอกภาพ ไม่ดึงดูดใจหรือบางครั้งอาจทำให้ผู้เรียนเกิดความสับสนในการรับรู้ได้

ข้อแนะนำเบื้องต้นในการเลือกใช้รูปภาพสำหรับการเตรียมสไลด์⁽⁵⁾ ได้แก่

1. อย่าใช้รูปภาพเป็นเพียงเครื่องตกแต่งสไลด์ เพราะมันจะเบี่ยงเบนความสนใจของผู้เรียน แนะนำให้เลือกใช้รูปภาพที่สอดคล้องกับเนื้อหาการบรรยายเพื่อเพิ่มความสนใจและทำให้ผู้เรียนเข้าใจเนื้อหามากขึ้น
2. เลือกใช้รูปภาพขนาดใหญ่เพียงรูปเดียวหรือไม่เกินสองรูปเพื่อสื่อแก่นของการบรรยายแก่ผู้เรียน อย่างจงใจใส่รูปภาพจำนวนมากเกินไปจนดูรกและเพิ่มภาระการรับรู้ของผู้เรียนโดยใช้เหตุ หากเป็นไปได้ แนะนำให้ใช้รูปภาพที่เหมาะสมเพียงภาพเดียววางให้เต็มสไลด์เพื่อสื่อแก่นสำคัญของสไลด์นั้น ๆ
3. เลือกรูปภาพที่มีความละเอียดสูงและสัดส่วนเหมาะสม ไม่น้อยกว่า 800x600 พิกเซลหรือ 1024x768 พิกเซล สำหรับหน้าจออัตราส่วน 4:3 และความละเอียดไม่น้อยกว่า 1280x720 พิกเซลสำหรับหน้าจออัตราส่วน 16:9
4. จัดวางทิศทางของรูปภาพอย่างเหมาะสม แนะนำให้วางรูปภาพในทิศทางเดียวกับเนื้อหาที่อยู่ในสไลด์ เช่น สายตาของรูปคนในสไลด์ต้องมองเข้าหาตัวอักษรบนสไลด์เสมอ
5. ชนิดของไฟล์รูปภาพที่เหมาะสมสำหรับการเตรียมสไลด์ ได้แก่ JPEG, PNG และ TIFF
6. ระวังการใช้รูปภาพที่มีลายน้ำในสไลด์ เพราะนอกจากจะผิดลิขสิทธิ์แล้ว ยังสะท้อนถึงความมั่งคั่งและไม่เป็นมืออาชีพของผู้สอนด้วย
7. หลีกเลี่ยงการใช้ภาพกฤตศิลป์ (clip art) หรือตัวการ์ตูน มาประกอบในสไลด์เพราะภาพเหล่านี้ล้าสมัยและไม่ดึงดูดความสนใจของผู้ฟัง ทั้งยังสะท้อนถึงความไม่เป็นมืออาชีพของผู้สอนอีกด้วย
8. ด้วยเทคโนโลยีการถ่ายภาพที่ง่ายขึ้นในปัจจุบัน ผู้สอนสามารถถ่ายภาพจากโทรศัพท์มือถือหรือกล้องดิจิทัล และเก็บเป็นคลังภาพของตนเองได้โดยง่ายและไม่ผิดลิขสิทธิ์

9. รูปดี ๆ สวยงาม และมีความละเอียดสูง สามารถหาได้ฟรีจากเว็บไซต์ต่อไปนี้โดยไม่ผิดลิขสิทธิ์

- <https://picjumbo.com>
- <https://unsplash.com>
- <https://stocksnap.io>
- www.pexels.com
- www.pixabay.com

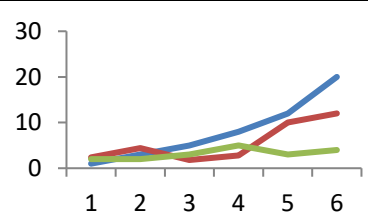
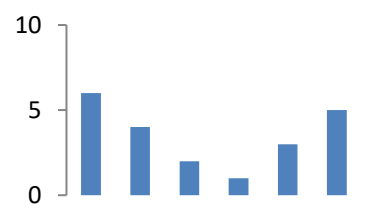
เช่นเดียวกับการใช้รูปภาพในการนำเสนอ วิดีโอก็มีประโยชน์ในการดึงดูดความสนใจของผู้เรียนเช่นเดียวกัน เพียงแต่ว่าผู้สอนต้องทราบถึงหลักการพื้นฐานในการนำวิดีโอมาใช้เพื่อประกอบการนำเสนอ⁽⁵⁾ ดังนี้

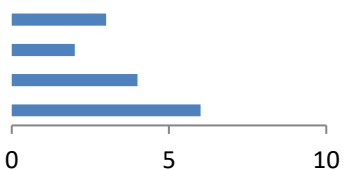
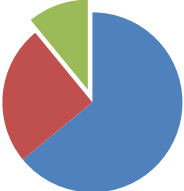
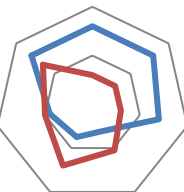
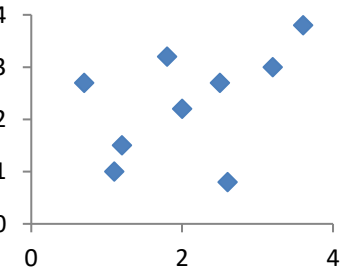
1. วิดีโอที่ใช้ต้องมีความละเอียดสูง และชัดเจนเมื่อนำเสนอ ในปัจจุบันความละเอียดของวิดีโอสำหรับการเตรียมสไลด์ที่แนะนำ คือระดับความละเอียดสูง 1280x720 พิกเซล (high definition; HD 720p) ขึ้นไป
2. เลือกวิดีโอที่มีเนื้อหาสอดคล้องกับเนื้อหาการบรรยาย
3. ความยาวของวิดีโอที่เหมาะสม คือ ไม่ควรยาวเกิน 60 วินาที
4. หลีกเลี่ยงการนำเสนอวิดีโอผ่านระบบอินเทอร์เน็ต เพราะมีโอกาสเสี่ยงต่อการเกิดความผิดพลาดของระบบเครือข่ายได้บ่อย

5. กราฟและแผนภูมิ

การนำเสนอด้วยกราฟและแผนภูมิเป็นรูปแบบที่พบได้บ่อยในการบรรยายทางการแพทย์ ผู้สอนมีความจำเป็นอย่างยิ่งที่จะต้องเข้าใจเป้าหมายของการใช้กราฟและแผนภูมิแต่ละชนิด (ตารางที่2) เพื่อให้สามารถสื่อความหมายไปยังผู้เรียนได้อย่างมีประสิทธิภาพ นอกจากนี้เป็นหน้าที่ของผู้สอนที่จะต้องพิจารณาและกรองเฉพาะแก่นสำคัญในกราฟและแผนภูมิล่านั้นว่า แก่นที่แท้จริงนั้นคืออะไร เพื่อจะได้ลดทอนหรือตัดเนื้อหาที่ไม่จำเป็นออกไปและทำให้กราฟและแผนภูมิล่านั้นดูง่ายที่สุด^(5, 6, 8)

ตารางที่2 เป้าหมายของกราฟและแผนภูมิชนิดต่าง ๆ ที่นิยมใช้ในปัจจุบัน^(2, 6, 8)

ชนิด	รูปภาพ	เป้าหมาย
กราฟเส้น		ใช้แสดงแนวโน้มการเปลี่ยนแปลงของข้อมูลที่น่าสนใจ ข้อมูลที่แสดงไม่ควรมากกว่า 4 เส้น
แผนภูมิแท่งแนวนิ่ง		ใช้แสดงข้อมูลการเปลี่ยนแปลงของปริมาณ ณ ช่วงเวลาหนึ่ง ข้อมูลที่แสดงควรอยู่ระหว่าง 4-8 ข้อมูล ใช้ภาพ 2 มิติเหมาะสมที่สุด

<p>แผนภูมิแท่ง แนวนอน</p>		<p>ใช้แสดงข้อมูลเพื่อเปรียบเทียบปริมาณของสิ่งที่สนใจ ข้อมูลที่แสดงควรอยู่ระหว่าง 4-8 ข้อมูล ใช้ภาพ 2 มิติเหมาะสมที่สุด</p>
<p>แผนภูมิรูปวงกลม</p>		<p>ใช้แสดงข้อมูลที่เป็นร้อยละ ข้อมูลที่แสดงควรอยู่ระหว่าง 4-6 ข้อมูล ข้อมูลที่ต้องการเน้นสามารถแสดงโดยใช้สีที่แตกต่าง หรือทำให้ลอยออกมาจากขอบเขตของวงกลมได้</p>
<p>แผนภูมิรูปใยแมงมุม</p>		<p>ใช้แสดงข้อมูลเพื่อเปรียบเทียบกับค่าปกติ ข้อมูลที่แสดงควรมีแกนเปรียบเทียบอยู่ระหว่าง 3-9 แกน</p>
<p>แผนภาพการกระจายของตัวแปร X และ Y</p>		<p>ใช้แสดงความสัมพันธ์ของข้อมูลให้เห็นเด่นชัด</p>

6. จำนวนสไลด์

คำถามที่พบบ่อยเมื่อต้องเตรียมสไลด์ประกอบการนำเสนอ คือ มีเวลาจำกัดเท่านี้ จะเตรียมสไลด์กี่แผ่นดี ภายคาวาซากิ ผู้เชี่ยวชาญด้านการนำเสนอแนะนำว่าจำนวนสไลด์ที่เหมาะสมควรอยู่ในอัตราส่วน 10 สไลด์ต่อการนำเสนอ 20 นาที⁽¹¹⁾ แต่ในความจริงนั้นจำนวนสไลด์ที่เหมาะสมแปรผันกับรูปแบบการนำเสนอของผู้สอนเป็นสำคัญ ผู้สอนหลายคนใช้สไลด์เพียง 10 สไลด์หรือน้อยกว่านั้นในการนำเสนอ 1 ชั่วโมง ขณะที่ผู้สอนบางคนนำเสนอหลายร้อยสไลด์ภายใน 1 ชั่วโมง ประเด็นสำคัญคือ ในฐานะผู้สอน คุณมีรูปแบบการนำเสนอแบบใด ที่จะทำให้ผู้ฟังได้รับแก่นแท้ของการนำเสนอมากที่สุด โดยมีองค์ประกอบที่ไม่จำเป็นน้อยที่สุด ผ่านจำนวนสไลด์ที่เหมาะสมที่สุด⁽¹⁾ เท่านั้นเอง จากประสบการณ์ของผู้นิพนธ์ จะยึดหลักการ 1 สไลด์ต่อการนำเสนอเฉลี่ย 2 นาที ด้วยเหตุนี้หากต้องสอนบรรยายในระยะเวลา 1 ชั่วโมง จะเตรียมสไลด์ 30-35 สไลด์เท่านั้น

อย่างไรก็ตาม สิ่งที่สำคัญกว่าจำนวนสไลด์คือ ผู้เรียนมักจะมีสมาธิหรือมีความสนใจลดลงเมื่อการสนทนาบรรยายดำเนินไปแล้วประมาณ 20 นาที ดังนั้นผู้สอนจำเป็นต้องดึงดูดความสนใจของผู้เรียนด้วยกิจกรรมที่แตกต่างอยู่เสมอในช่วงเวลาดังกล่าว เช่น การถามคำถาม การอภิปราย การฉายวิดีโอ เล่าเรื่องราวหรือประสบการณ์ เป็นต้น ทั้งนี้เพื่อกระตุ้นให้ผู้เรียนตื่นตัวอยู่เสมอ⁽⁹⁾

7. ภาพเคลื่อนไหวและการเปลี่ยนสไลด์

โปรแกรมสร้างสไลด์ประกอบการนำเสนอส่วนใหญ่ได้เตรียมรูปแบบภาพเคลื่อนไหวและการเชื่อมต่อสไลด์ที่หลากหลายไว้ให้ผู้สอน ทั้งแบบที่น่าตื่นเต้น หมุนเกลียว ควงสว่าง หรือแม้แต่เสียงอันน่าตกใจจำนวนมาก แต่ไม่ได้หมายความว่าผู้สอนต้องเลือกใช้องค์ประกอบนี้ทุกครั้งไป ผู้สอนต้องพึงระลึกเสมอว่าผู้เรียนต้องการสไลด์ที่เรียบง่ายและความตื่นเต้นของภาพเคลื่อนไหวเหล่านี้เป็นสิ่งไม่จำเป็น ทั้งยังเป็นการเพิ่มภาระของสมองผู้เรียนในการรับรู้สิ่งเร้าโดยใช้เหตุ ด้วยเหตุนี้แนะนำว่าการใช้ภาพเคลื่อนไหวในสไลด์ ต้องมีวัตถุประสงค์ที่ชัดเจน และสอดคล้องกับหลักการข้อใดข้อหนึ่งต่อไปนี้^(1, 5)

- 1.1. เพื่อเน้นส่วนสำคัญของบนสไลด์ให้ผู้เรียนสนใจและเข้าใจมากขึ้น
- 1.2. เพื่อชักจูงของผู้เรียนไปหาแก่นของเนื้อหาบรรยายตามลำดับ
- 1.3. เพื่อส่งสัญญาณนำเข้าสู่ประเด็นสำคัญของเนื้อหาบรรยาย
- 1.4. เพื่อสร้างการขับเคลื่อนไปสู่ประเด็นสำคัญถัดไป

การเปลี่ยนสไลด์ก็เช่นเดียวกัน โปรแกรมสร้างสไลด์นำเสนอได้มีรูปแบบการเปลี่ยนสไลด์มาให้เลือกใช้มากมาย แต่แนะนำให้ผู้สอนยึดหลักสำคัญ คือ เลือกใช้รูปแบบการเปลี่ยนสไลด์ให้น้อยที่สุด ไม่เกิน 2-3 รูปแบบ และไม่จำเป็นต้องใช้คุณสมบัตินี้กับทุก ๆ สไลด์ที่ใช้ในการสอนบรรยาย⁽²⁾

8. การจัดวางองค์ประกอบบนสไลด์ (รูปที่3)

เพื่อตอบสนองต่อการรับรู้ที่จำกัดของผู้เรียน การจัดองค์ประกอบภายในสไลด์จึงมีความสำคัญอย่างมาก ทั้งนี้ก็เพื่อให้สมองของผู้เรียนไม่ถูกใช้งานในการรับรู้มากเกินไป หลักการง่าย ๆ ของการจัดวางองค์ประกอบบนสไลด์คือการใช้ **ทฤษฎีการจัดวางองค์ประกอบภาพด้วยกฎ 3X3⁽⁵⁾** เช่นเดียวกับการจัดองค์ประกอบของภาพถ่าย เพื่อให้ได้องค์ประกอบบนสไลด์ที่สมดุล ดูสบายตา มีพื้นที่ว่างเหมาะสม และไม่เพิ่มภาระการรับรู้ของผู้เรียน โดยการวางองค์ประกอบต่าง ๆ บนสไลด์ให้อยู่ในระนาบหรือแนวของเส้นแบ่งที่กำหนดไว้ และจัดวางองค์ประกอบต่าง ๆ ให้เป็นกลุ่ม ๆ ไม่เกิน 4-7 กลุ่ม⁽⁶⁾ เพียงเท่านี้ก็จะทำให้สไลด์มีความน่าสนใจและดึงดูดผู้เรียนได้มากขึ้น ตัวอย่างที่เห็นได้ชัดของการจัดวางองค์ประกอบด้วยทฤษฎีนี้ได้แก่ การจัดวางสำหรับในกล่องข้าวของชาวญี่ปุ่นหรือ Bento นั่นเอง



รูปที่3 ตัวอย่างการวางภาพด้วยการใช้ทฤษฎีจัดวางองค์ประกอบภาพด้วยกฎ 3X3 ภาพซ้าย รูปการจัดวางอาหารในข้าวกล่องญี่ปุ่น (bento) ให้ความรู้สึกน่ารับประทาน ภาพขวา การจัดวางองค์ประกอบภาพทิวทัศน์ ให้ความรู้สึกแปลกตา และสวยงาม (ภาพประกอบโดย ผศ. นพ.ตรีภพ เลิศบรรณพงษ์)

บทสรุป

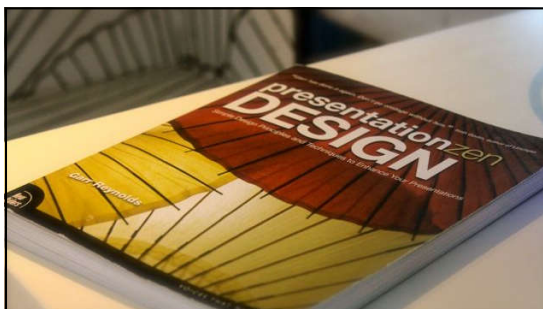
การเตรียมสไลด์สำหรับการนำเสนอเป็นทักษะพื้นฐานสำคัญของครูแพทย์ที่มักถูกละเลย และไม่ได้รับความสนใจเท่าที่ควร ทักษะนี้ไม่ใช่เรื่องยาก ขณะเดียวกันก็ไม่ได้ง่ายจนไม่ต้องเรียนรู้เลย ความเข้าใจการรับรู้ของผู้เรียนและการตั้งเป้าหมายการนำเสนอที่ชัดเจน เน้นเฉพาะแก่นแท้ของเรื่องราวที่ต้องการนำเสนอ เป็นองค์ประกอบสำคัญลำดับแรกที่ต้องใส่ใจ การให้เวลากับการสร้างเรื่องราวในกระดาน และการยับยั้งชั่งใจในการใส่เนื้อหาที่มากเกินไปในสไลด์ด้วยการแยกสไลด์ประกอบการนำเสนอกับเอกสารประกอบการบรรยายออกจากกัน เป็นสิ่งที่พึงกระทำในลำดับถัดมา เมื่อพร้อมแล้วจึงค่อยลงมือสร้างสรรค์สไลด์ด้วยการจัดวางองค์ประกอบ การใช้พื้นหลัง ตัวอักษร สี รูปภาพ วิดีโอ รวมทั้งการเชื่อมต่ออย่างเหมาะสม จะช่วยส่งเสริมให้ผู้สอนโดดเด่นและเป็นที่น่าสนใจ และผู้ที่ได้รับประโยชน์มากที่สุดภายหลังเสร็จสิ้นการสอนบรรยายก็คือ “ผู้เรียน” นั่นเอง

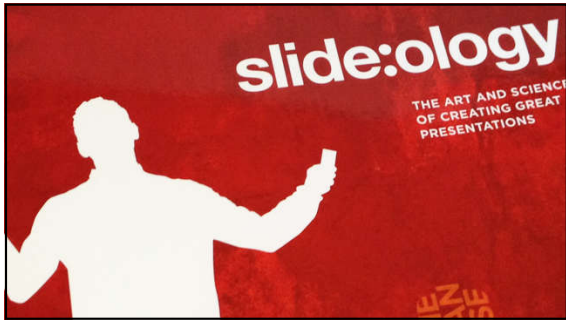
เอกสารอ้างอิง

1. Duarte N. Slide:ology : the art and science of creating great presentations. 1 ed. Sebastopol: O'Reilly Media; 2008. 296 p.
2. Reynolds G. Top ten slide tips 2 0 1 6 [cited 2 0 1 6 May 2]. Available from: <http://www.garreynolds.com/preso-tips/design/>.
3. Harden RM. Death by PowerPoint - the need for a 'fidget index'. Med Teach. 2008;30:833-35.
4. Tufte ER. The cognitive style of powerpoint: pitching out corrupts within. 2 ed. Connecticut: Graphics Press; 2006. 32 p.
5. Reynolds G. Presentation Zen Design : A simple visual approach to presenting in today's world. 2 ed. San Francisco: New Riders; 2014. 277 p.
6. Gruwez E. Presentation thinking and design : create better presentations, quicker. 1 ed. London: FT Publishing International; 2014. 225 p.
7. Young JQ, Van Merriënboer J, Durning S, Ten Cate O. Cognitive Load Theory: implications for medical education: AMEE Guide No. 86. Med Teach. 2014;36:371-84.
8. Vadrnais D. Presentation Dos and Don'ts: tips for preparing great slides. Connecting People to Useful Information - Guidelines for Effective Data Presentations Washington, DC: Population Reference Bureau, MEASURE Communication (Project); 2004. p. 9-16.
9. Brown G, Manogue M. AMEE Medical Education Guide No. 22 : Refreshing lecturing: a guide for lecturers. Med Teach. 2001;23:231-44.
10. Holzl J. Twelve tips for effective PowerPoint presentations for the technologically challenged. Med Teach. 1997;19:175-9.
11. Kawasaki G. The 10/20/30 Rule of PowerPoint. 2005 [cited 2016 May 2]. Available from: http://guykawasaki.com/the_102030_rule/.









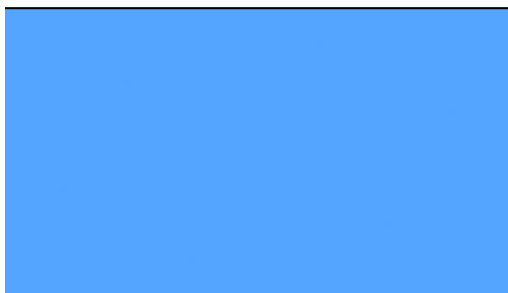


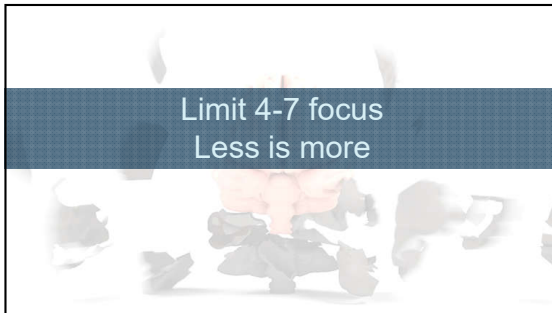


5/12/2019



















5/12/2019





Item	Self-admission rate (%)		Odds ratio (RTS/control)	Two-tailed p (likelihood ratio test)	Defensibility rating (across groups)
	Control group	RTS group			
1. In a paper failing to report all of a study's dependent measures	63.4	46.5	1.14	.23	1.94 (0.39)
2. Deciding whether to collect more data after looking to see whether the results were significant	55.9	58.0	1.08	.46	1.79 (0.44)
3. In a paper failing to report all of a study's conditions	27.7	27.4	0.98	.90	1.77 (0.49)
4. Stopping collecting data earlier than planned because one found the result that one had been looking for	15.6	22.5	1.57	.00	1.76 (0.48)
5. In a paper, "rounding off" a p value (e.g., reporting that a p value of .056 is less than .05)	22.0	23.3	1.07	.58	1.68 (0.57)
6. In a paper, selectively reporting studies that "worked"	45.0	50.0	1.18	.13	1.66 (0.53)
7. Deciding whether to exclude data after looking at the impact of doing so on the results	38.2	43.4	1.23	.06	1.61 (0.55)
8. In a paper, reporting an unexpected finding as having been predicted from the start	27.0	35.0	1.45	.00	1.50 (0.60)
9. In a paper, claiming that results are unaffected by demographic variables (e.g., gender) when one is actually unsure (or knows that)	3.0	4.5	1.52	.16	1.32 (0.60)
0. Falsifying data	0.6	1.7	2.75	.07	0.16 (0.38)

5/12/2019

“คำ” หรือ “วลี”... ดีที่สุด

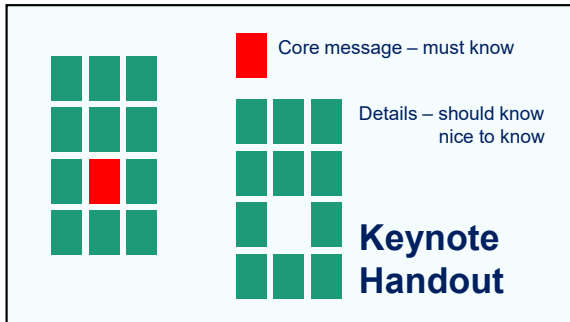
7%

ใส่ตั้งอีกเสบ





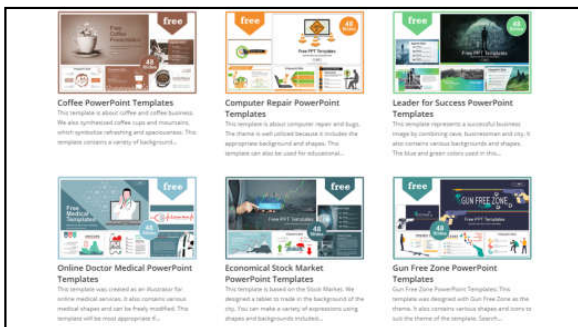


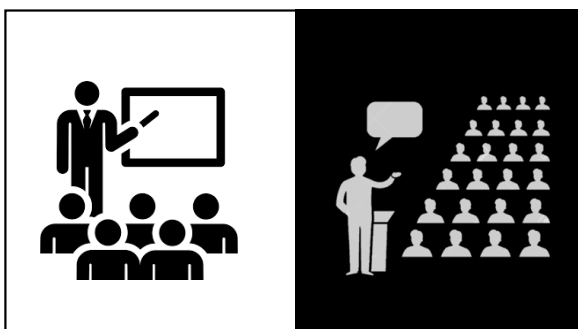


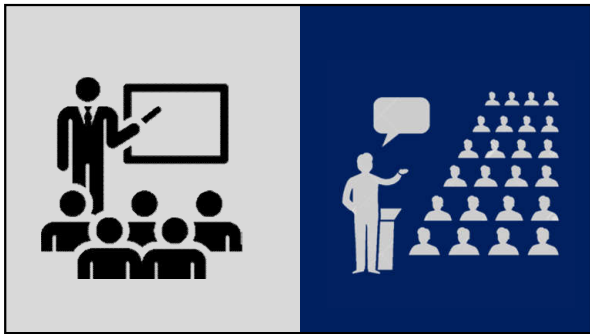


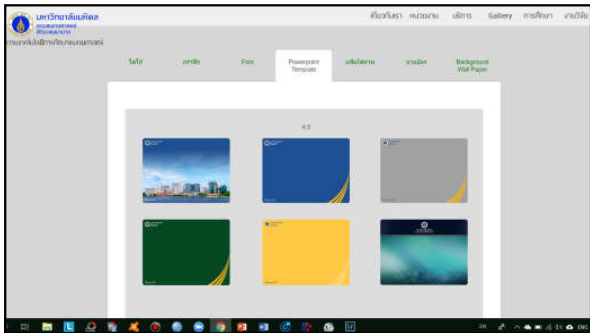


01Background











02Alphabet



use **Sans Serif**

<p>Bookman old style Garamond Times New Roman Rockwell</p>	<p>Arial Century Gothic Helvetica Tahoma</p>
--	--

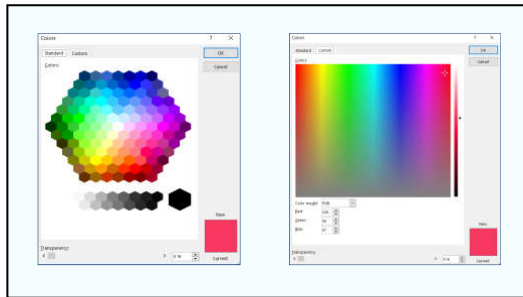
No more than **2**

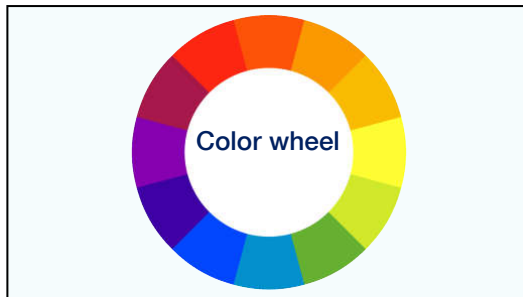
<p>Heading</p> <p>Background Objective Methods Conclusion</p>	<p>Heading</p> <p>Background Objective Methods Conclusion</p>
--	--

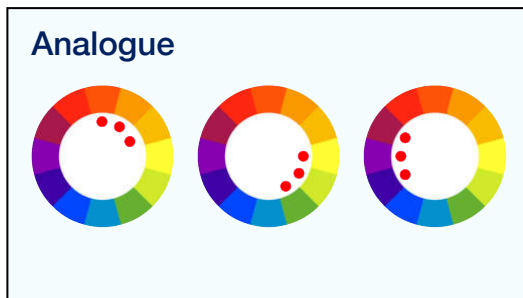
avoid CAPITAL LETTER

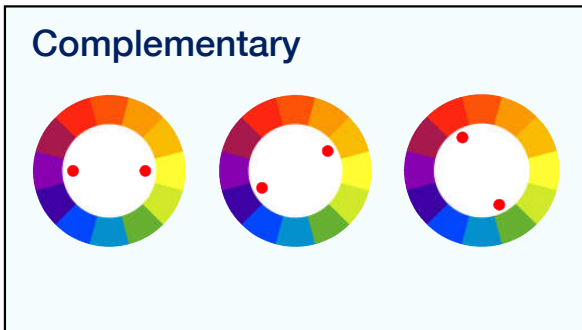
HEADING	Heading
BACKGROUND	Background
OBJECTIVE	Objective
METHODS	Methods
CONCLUSION	Conclusion

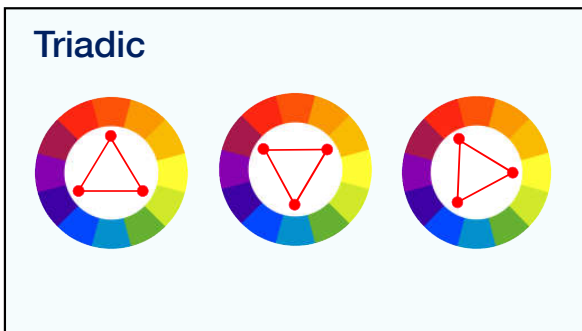
03Color

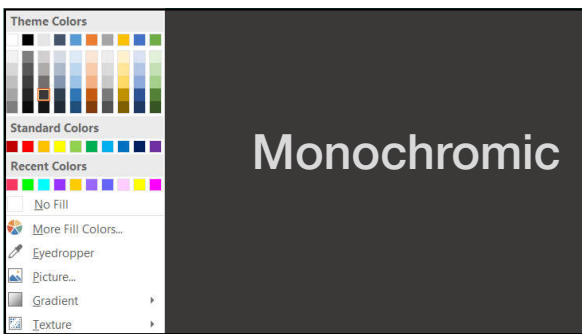












No more than **3**

1/12
Color blindness

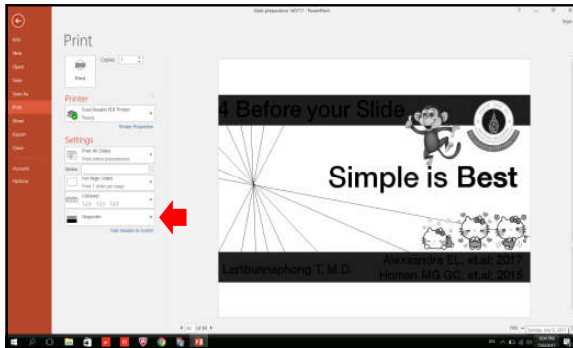
4 Before your Slide

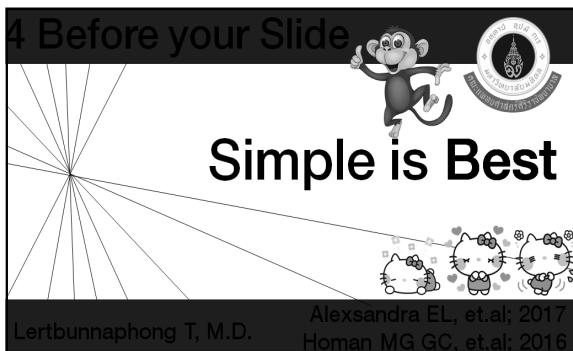


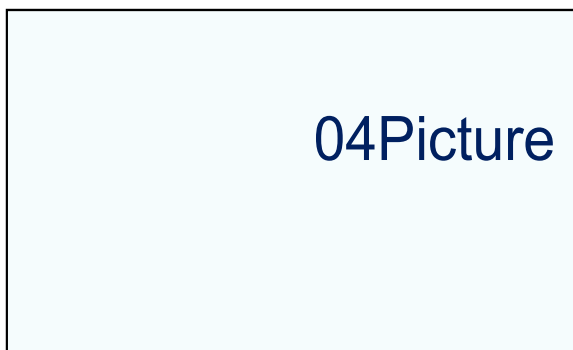
Simple is Best

Lertbunnaphong T, M.D. Alexandra EL, et.al; 2017
Homan MG GC, et.al; 2016

5/12/2019







5/12/2019

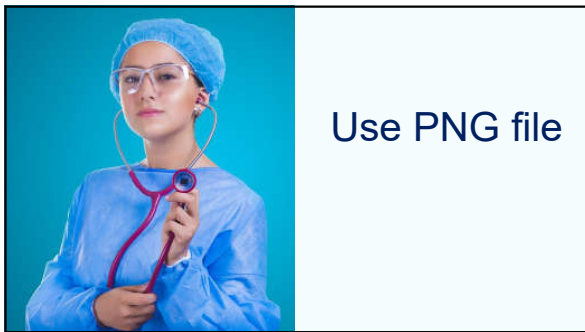






5/12/2019







Emotion

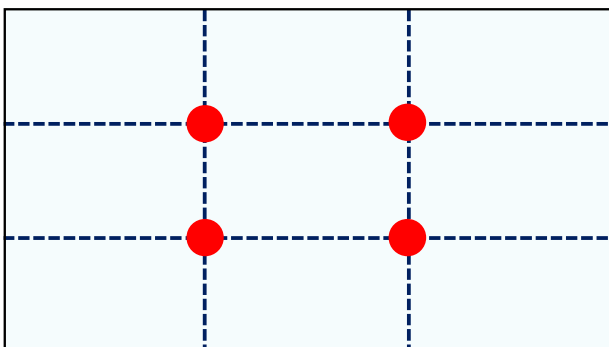




5/12/2019



3x3 rule

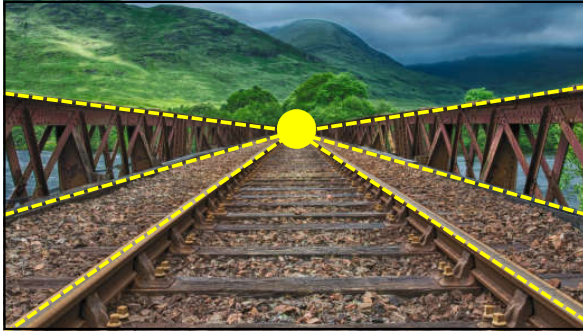


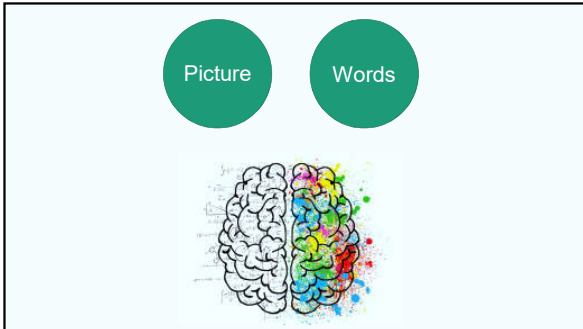
5/12/2019





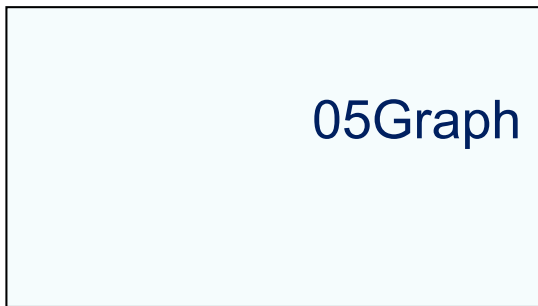
Leading line

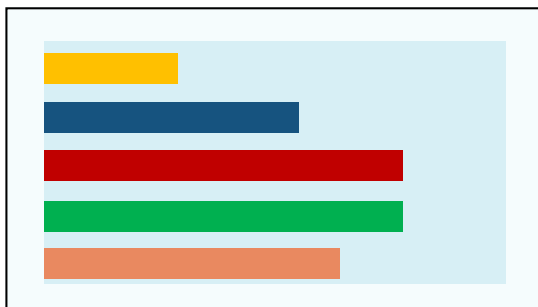




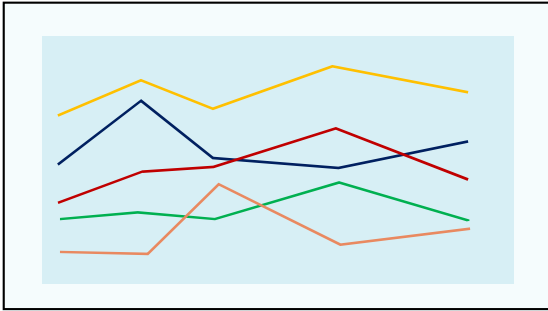


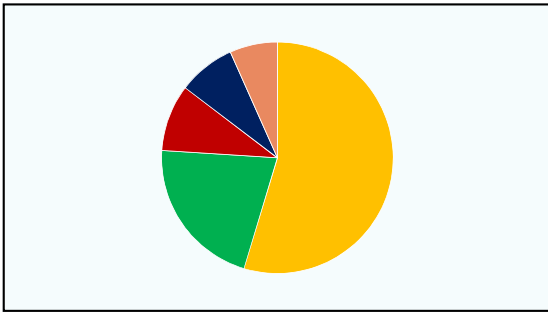




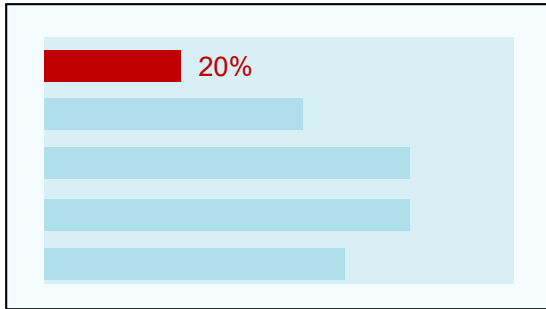


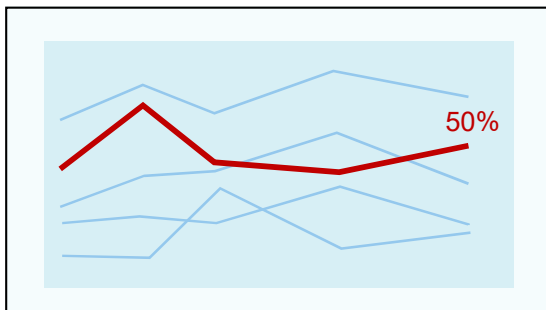
5/12/2019

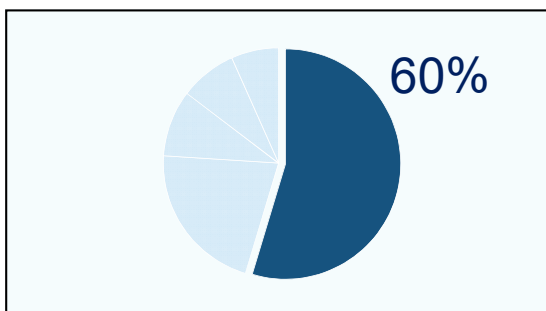


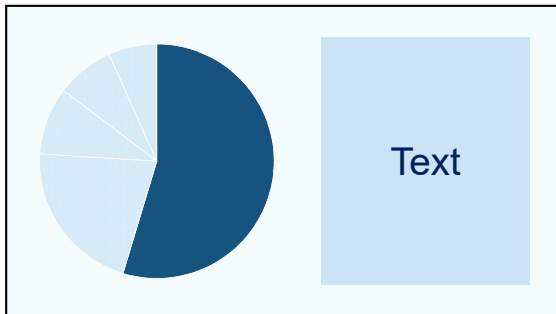


God VS Devil





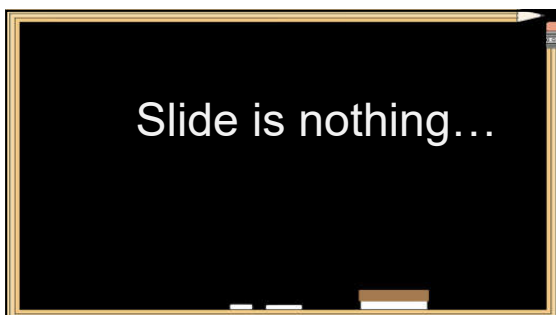




Item	Self-admission rate (%)		Odds ratio (BTS/control)	Two-tailed p (likelihood ratio test)	Defensibility rating (across groups)
	Control group	BTS group			
1. In a paper, failing to report all of a study's dependent measures	63.4	66.5		.23	1.84 (0.39)
2. Deciding whether to collect more data after looking to see whether the results were significant	55.9			.46	1.79 (0.44)
3. In a paper, failing to report all of a study's conditions				.90	1.77 (0.49)
4. Stopping collecting data more than planned because one got the results that one had been looking for				.00	1.76 (0.48)
5. In a paper, "rounding off" a p value (e.g., reporting that a p value of .004 is less than .01)				.58	1.68 (0.57)
6. In a paper, selectively reporting studies that "worked"				.13	1.66 (0.53)
7. Deciding whether to exclude studies after looking at the impact of a finding on the results				.06	1.61 (0.59)
8. In a paper, reporting an unexpected finding as having been predicted from the start				.00	1.50 (0.60)
9. In a paper, claiming that results are unaffected by demographic variables (e.g., gender) when one is actually unsure (or knows that they do)			1.52	.16	1.32 (0.60)
10. Falsifying data	0.6	1.7	2.75	.07	0.16 (0.38)

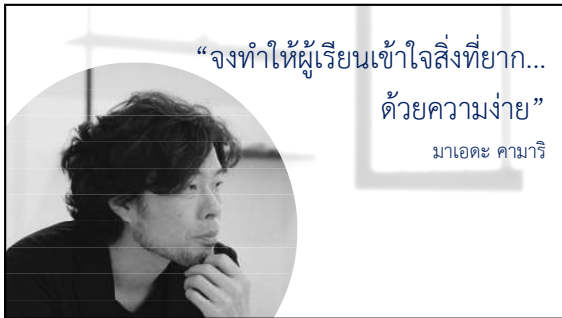
CAUTION

Notes: Items are listed in decreasing order of rated defensibility. Respondents who admitted to having engaged in a given behavior were asked to rate whether they thought it was defensible to have done so (0 = no, 1 = possibly, and 2 = yes). Standard deviations are given in parentheses. BTS = Bayesian Trust System. Applying the Bonferroni correction for multiple comparisons, we adjusted the critical alpha level downward to .005 (i.e., .05/10 admissions).



5/12/2019





การใช้คำถามในการสอน

สุพจน์ พงศ์ประสพชัย

To question well is to teach well
In the skillful use of questions, more than anything else,
lines the fine art of teaching

Earnst Sachs

บทนำ

การถาม (questioning) เป็นเครื่องมือในการสอนที่มีพลังเป็นอย่างยิ่งที่ครูควรมีทักษะเป็นอย่างดี เนื่องจาก การใช้คำถามในการสอนที่ดีจะก่อให้เกิดประโยชน์ทั้งต่อครูผู้สอนเองและผู้เรียน ในขณะเดียวกันเนื่องจากธรรมชาติ ของครู (แพทย์) ส่วนใหญ่มีความถนัดในการหาคำตอบมากกว่าการตั้งคำถาม จึงประสบปัญหาในการใช้คำถามในการ สอนได้บ่อย รวมทั้งอาจก่อปัญหาจากการใช้คำถามของตนเสียเองก็มี ในบทนี้ผู้เขียนจะกล่าวถึงประโยชน์ของการใช้ คำถามในการสอน ชนิดของคำถาม คำถามแบบโซเครติก และปัจจัยที่จะทำให้การใช้คำถามมีประสิทธิภาพ

ประโยชน์ของการใช้คำถาม

ประโยชน์ต่อครู

1. ได้ทบทวนเนื้อหาที่จะสอน ว่าส่วนใดเป็นแก่น (core) ซึ่งควรสอนเน้นโดยเฉพาะอย่างยิ่งการใช้คำถาม และ เนื้อหาใดเป็นรายละเอียด (detail) ซึ่งไม่จำเป็นต้องเน้น อาจสอนด้วยการบอกไปเลย หรือไม่สอนเลยก็ได้
2. ใช้ประเมินผู้เรียน
3. ใช้กระตุ้นผู้เรียน ให้ตื่นตัว คิด วิเคราะห์ วิจาร์ณ สร้างสรรค์ และอภิปรายอย่างเต็มศักยภาพของตน ตาม การเรียนรู้แบบผู้ใหญ่
4. ใช้กระตุ้นการมีส่วนร่วมของผู้เรียนทุกคน
5. ใช้สร้างบรรยากาศการสอนให้มีสีสัน สนุก และอบอุ่นไปด้วยความสุข
6. ใช้สอนทักษะการแก้ปัญหา (problem solving) และการใช้เหตุผล (clinical reasoning) การสอนทักษะทั้ง สองนี้ไม่มีวิธีใดดีไปกว่าการใช้คำถาม

ประโยชน์ต่อผู้เรียน

1. กระตุ้นให้ผู้เรียนต้องคิด (thinking) เพราะการถูกถามทำให้ผู้เรียนต้องคิดอย่างเลียงไม่ได้ ซึ่งแตกต่างจาก การฟังเลคเชอร์ซึ่งผู้เรียนอาจฟังแต่ไม่ได้คิดก็ได้
2. ทำให้ผู้เรียนได้นำความรู้พื้นฐานเดิมมาประยุกต์ (application) เพื่อตอบคำถาม โดยเฉพาะถ้าครูเลือกใช้ คำถามขั้นสูงที่มีใช้ถามแต่ความจำ
3. ทำให้ผู้เรียนได้สังเคราะห์ (synthesis) ความรู้ใหม่ จากองค์ความรู้เดิมที่มีรวมกับการชี้แนะโดยครู
4. ทำให้ผู้เรียนได้รับการแก้ไขความรู้ความเข้าใจที่ผิด (correction) เพราะคำตอบที่ผู้เรียนตอบออกมาเป็น การเปิดเผยตัวตน ความรู้ความเข้าใจของผู้เรียนที่อาจจะถูกหรือผิดมาโดยตลอด (ไม่เคยมีผู้เรียนที่จะแก้งตอบผิด)
5. ทำให้ผู้เรียนเกิดแรงจูงใจในการเรียน (motivation) เพราะทำให้ผู้เรียนทราบความสำคัญของเรื่องนั้นและ เกิดความอยากเรียนรู้

6. ทำให้ผู้เรียนจำได้นาน (retention)
7. ฝึกให้ผู้เรียนกล้าคิด กล้าพูด กล้าแสดงออก (expression)
8. ฝึกให้ผู้เรียนเรียนรู้ที่จะรับฟังความเห็นของผู้อื่น (listening) รู้จักการให้เหตุผลเพื่อสนับสนุนแนวคิดของตนเอง (defending) และรู้จักการอภิปรายโต้แย้งความเห็นของผู้อื่นอย่างสุภาพและสร้างสรรค์ (discussion)

ชนิดของคำถาม

1. Cognitive memory question เป็นคำถามที่แคบ ได้แก่คำถามให้ตอบใช่-ไม่ใช่ บอกชื่อ บอกความหมาย บอกคำจำกัดความ เป็นต้น คำถามชนิดนี้ผู้ตอบไม่ต้องใช้ความคิดหนักนอกจากความจำ จึงไม่สนับสนุนให้เกิดการคิดและการเรียนรู้มากนัก แต่เป็นการเรียนแบบนกแก้วนกขุนทอง (rote learning) แต่มีประโยชน์ในแง่ถ้าต้องการระดมความคิดเห็นจากกลุ่ม ใช้ดึงความสนใจของนักเรียนที่ไม่สนใจหรือนักเรียนที่อ่อนให้มีส่วนร่วมได้ง่าย

2. Convergent question คำถามที่กว้างกว่า ได้แก่ คำถามให้อธิบายความหมาย บอกความสัมพันธ์ บอกความเหมือนหรือความแตกต่าง เป็นต้น คำถามชนิดนี้ผู้เรียนต้องมีความรู้ระดับหนึ่งจึงจะสามารถคิดต่อได้ แต่อย่างไรก็ตามก็ยังถือว่าเป็นคำถามที่แคบเนื่องจากมักมีคำตอบที่ถูกที่สุดเพียงคำตอบเดียว

3. Divergent question เป็นคำถามที่กว้างเนื่องจากสามารถมีคำตอบที่ถูกต้องได้หลากหลายคำตอบ ได้แก่ คำถามให้คาดการณ์ ตั้งสมมุติฐาน สร้างความคิดหรือไอเดียใหม่ๆ divergent question เป็นคำถามที่ดีในการกระตุ้นให้เกิดความคิดในผู้เรียน

4. Evaluative question เป็นคำถามที่กว้างที่ต้องใช้ความคิดขั้นสูงที่สุด ได้แก่ คำถามให้ประเมิน ตัดสิน เปรียบเทียบคุณค่า เลือทางที่ดีที่สุด เป็นต้น คำถามนี้เป็นคำถามที่ผู้เรียนต้องใช้ความคิดมากที่สุด

เมื่อพิจารณาตาม Bloom's Taxonomy จะพบว่าคำถามที่แคบจะกระตุ้นให้ผู้เรียนใช้ความคิดเพียงในระดับ lower order of thinking ในขณะที่คำถามที่กว้างจะกระตุ้นให้ผู้เรียนต้องใช้ higher order of thinking ซึ่งเป็นสิ่งที่เราต้องการให้ผู้เรียนได้

การสอนโดยการถามแบบโซเครติก (Socratic questioning)

การถามแบบโซเครติก คือ การถามเพื่อสอน แทนที่จะใช้การบอกไปเลย การถามเพื่อสอนมีข้อดีกว่าการบอกหลายประการดังที่ได้กล่าวไปแล้วข้างต้น การถามแบบโซเครติกจึงมักใช้ divergent หรือ evaluative question เป็นหลัก ที่ใช้บ่อยๆ มี 6 แบบได้แก่

1. คำถามให้กระจ่าง (clarification หรือ “tell me more”) ได้แก่

- “คุณลองอธิบายเพิ่มเติมอีกหน่อยได้มั๊ย?”
- “เรื่องที่คุณพูดมันเกี่ยวกับผู้ช่วยรายนี้อย่างไร?”
- “ที่คุณพูดหมายความว่าอย่างไร?”
- “ไหนบอกตัวอย่างผมสักหน่อยได้ไหม?”

2. คำถามถึงข้อสรุป (probe assumption) ได้แก่

- “ตกลงคุณสรุปว่าไง?”
- “คุณเห็นด้วยกับข้อสรุปนี้หรือไม่ เพราะอะไร?”
- “อธิบายชี้ว่าทำไมเราจึงสรุปเช่นนี้?”

3. คำถามถึงเหตุผล (reason) หรือหลักฐาน (evidence) ได้แก่

- “คุณมีเหตุผลอะไรสำหรับคำตอบนี้?”
- “คุณคิดว่าอะไรทำให้เกิดเหตุนี้?”
- “มีเหตุผลอะไรที่ไม่สนับสนุนโรคนี้?”

4. คำถามเกี่ยวกับมุมมอง (viewpoint) และแนวความคิด (perspective) ได้แก่

- “มีทางเลือกอะไรอีกไหม?”
- “มีทางมองปัญหานี้แนวอื่นไหม?”
- “อะไรคือข้อดี ข้อเสีย ของวิธีนี้?”
- “วิธีนี้เทียบกับวิธีอื่นเป็นอย่างไร?”

5. คำถามเกี่ยวกับการนำไปใช้ (implication) และผลที่จะตามมา (consequence) ได้แก่

- “ผลที่ตามมาถ้าเรารักษาด้วยวิธีนี้คืออะไร?”
- “เราจะนำหลักการนี้ไปใช้กับผู้ป่วยรายนี้ได้อย่างไร?”

6. คำถามที่เกี่ยวกับคำถาม ได้แก่

- “คำถามนี้สำคัญอย่างไร?”
- “รู้มั๊ยผมถามคำถามนี้คุณทำไม?”

การใช้คำถามในการสอนให้ประสบความสำเร็จ

การใช้คำถามในการสอนที่จะประสบความสำเร็จประกอบด้วยองค์สาม ได้แก่ การสร้างบรรยากาศที่ดี (good atmosphere) การเลือกใช้คำถามที่ดี (good question) และการใช้เทคนิคที่ดี (good technique)

การสร้างบรรยากาศที่ดี

เป็นปัจจัยขั้นแรกสุดที่จะนำไปสู่การใช้คำถามที่ประสบความสำเร็จ ครูควรชี้แจง (orientate) ให้ผู้เรียนทราบตั้งแต่ก่อนเริ่มสอนว่าครูจะสอนโดยใช้คำถาม เพราะมีข้อดีมากกว่าการบอกไปเลยอย่างมากมาย ผู้เรียนจะได้ไม่ตกใจ หวาดระแวง หรือเครียดเมื่อถูกถาม ในการถามควรให้เกียรติโดยเรียกผู้เรียนด้วยชื่อจริง¹ ไม่ควรเรียกด้วยชื่อเล่น ไม่ควรเรียกว่า “น้อง” “ลูก” หรืออื่นๆ ซึ่งเป็นการให้เกียรติน้อยกว่าการเรียกด้วยชื่อจริง (ครูบางคนรู้สึกว่าการเรียกผู้เรียนว่า “น้องๆ” ทำให้สนิทสนม ซึ่งไม่เป็นความจริง ผู้เรียนมักไม่รู้สึกเท่ากับถูกเรียกด้วยชื่อ) การทำให้บรรยากาศเป็นมิตร ปลอดภัย และทำให้ผู้เรียนอยากตอบทำได้โดยใช้หลัก SARA (ตารางที่ 1)

ตารางที่ 1 การสร้างบรรยากาศในการใช้คำถามที่ดีโดยใช้หลัก SARA

S	Smile	ยิ้มแย้ม ไบหน้าผ่อนคลาย
A	Attentive	สบตา ตั้งใจฟังคำตอบ ไม่ขัดจังหวะ ไม่พูดแทรก
R	Respect	ให้เกียรติผู้เรียน ให้เกียรติในคำตอบของผู้เรียน ไม่เยาะเย้ย ไม่เสียดสี
A	Acknowledge	ไม่เฉย ไม่บั่นปึ้งเมื่อผู้เรียนตอบคำถาม อย่างน้อยควรพยักหน้า ผงกศีรษะ กระตุ้นให้ผู้เรียนพูดต่อ ชมเชยถ้าตอบได้ดี แก่ไขถ้าตอบไม่ถูกต้อง

การเลือกใช้คำถามที่ดี

มีหลักการดังต่อไปนี้

1. ควรถามคำถามอย่างมีจุดประสงค์เพื่อให้เกิดแนวทางที่สนับสนุนบทเรียนในชั่วโมงนั้นๆ มิใช่ถามไปเรื่อยเปื่อย
2. ถามให้ชัดเจน ไม่กำกวม ถามครั้งละ 1 คำถาม ไม่ถามครั้งเดียวหลายคำถาม เช่น “ผู้ป่วยรายนี้วินิจฉัยเป็นอะไร จะสืบค้นอะไร และจะรักษาอย่างไร?”
3. พยายามใช้คำถามแบบโซเครติก คำถามปลายเปิด ที่กระตุ้นกระบวนการคิดขั้นสูง (higher order of thinking) มักเป็นคำถามว่า “ทำไม?” และ “อย่างไร?”
4. คำถามที่ควรหลีกเลี่ยง ได้แก่

- คำถามใช่-ไม่ใช่ เห็นด้วย-ไม่เห็นด้วย เนื่องจากเป็นคำถามที่แคบ เกิดประโยชน์น้อย
- คำถามกำกวม (vague) เช่น “แล้วยังไงต่อ?” “เคสนี้เอาอย่างไรดี?”
- คำถามชักคะเย่อ (tugging) เช่น “มีอีกมั้ยๆ ...เอาอีกโรคหนึ่งๆ ...”
- คำถามให้เดา (guessing) โดยไม่ใช้ความรู้ เช่น “โรคอะไรเอ่ย? ขึ้นต้นด้วยตัว D ลงท้ายด้วยตัว E มี 2 พยางค์” “ลองเดาสิว่าคุณคิดถึงโรคอะไร?”
- คำถามชี้นำ (leading) เช่น “น่าจะอีกเสบตรงอวัยวะที่อยู่ท้องน้อยด้านขวาล่าง ห้อยอยู่ตรง caecum คือโรคอะไร?” “ลงท้ายด้วย _itis แปลว่าอักเสบ ดังนั้น salpingitis แปลว่าอะไร?”

การใช้เทคนิคที่ดี¹⁻³

1. การรอ (Wait Time หรือ Thinking Time)

หลังจากถาม ควรรอสักอย่างน้อย 5-10 วินาทีสำหรับคำตอบเสมอ (หรือนับการหายใจของตน 3 ครั้ง) เพราะผู้เรียนต้องใช้เวลาทำความเข้าใจกับคำถาม คาดเดาคำตอบของตัวเอง คิดหาคำตอบ เลือกคำตอบที่คิดว่าดีที่สุด แล้วจึงรวบรวมกำลังขวัญตอบออกมา ซึ่งกระบวนการนี้ต้องใช้เวลาอย่างน้อย 5-10 วินาที หรือนานกว่าถ้าคำถามยากมาก มิงานวิจัยที่พบว่าถ้ามีการรอที่นานพอเหมาะดังกล่าว การตอบของผู้เรียนจะเพิ่มขึ้น ตอบยาวขึ้น และตอบว่าไม่ทราบน้อยลง แต่ถ้าครูไม่รอ คำตอบที่ได้จากผู้เรียนจะน้อย เนื่องจากในขณะที่ผู้เรียนกำลังคิดอยู่ก็โดนตัดบท เฉลยหรือเปลี่ยนคำถามไปเสียก่อน เพราะครูเองก็มักขวยเขิน อึดอัด ขุ่นเคือง ที่ถามไปแล้วเงียบ (จากงานวิจัยพบว่าครูส่วนใหญ่คิดว่าตัวเองรอถึง 10 วินาทีแล้ว แต่จริงๆ รอเพียง 2.5 วินาทีเท่านั้น⁴) ในที่สุดผู้เรียนก็เรียนรู้และเกิดพฤติกรรมที่จะตอบว่า ไม่ทราบ หรือเงียบ เพราะรู้ว่าครูจะทนไม่ไหวและบอกเองในที่สุด

ในทางตรงกันข้าม ครูก็ไม่ควรรอนานเกินไป เช่น รอนานกว่า 30 วินาที เพราะจะเกิดบรรยากาศการกดดัน ผู้เรียนให้เครียด³ รู้สึกว่าครูเริ่มโกรธ หรือเริ่มลงโทษ เป็นต้น

2. เทคนิค Pose-Pause-Pounce-Bounce¹

เป็นเทคนิคที่ใช้ในกรณีที่สอนเป็นวงที่มีผู้เรียนหลายๆคน กล่าวคือ

Pose หมายถึง การตั้งคำถามให้แก่กลุ่มทั้งกลุ่ม

Pause หมายถึง หยุดรอฟังคำตอบ ซึ่งก็คือ 5-10 วินาที

Pounce หมายถึง กรณีที่ไม่มีผู้ตอบ ให้ถามระบุไปยังคนใดคนหนึ่ง

Bounce หมายถึง การสะท้อนไปยังผู้อื่น เมื่อผู้เรียนคนหนึ่งตอบ อาจถามผู้เรียนคนอื่นต่อในประเด็นนี้ หรือให้แสดงความเห็นต่อคำตอบนี้ เป็นต้น

3. การทำให้ทุกคนมีส่วนร่วม

เทคนิค ได้แก่

1. พยายามกระจายคำถามไปสู่ทุกๆ คนในกลุ่มแบบสุ่ม ไม่ควรถามแบบระบุไล่เรียงตามคิว ตามอาวุโส หรืออื่นๆ เพราะจะทำให้เกิดความกดดัน ถ้าจะถามแบบระบุคน ควรถามไปกลางกลุ่มก่อน เพื่อให้ทุกคนได้คิด ก่อนที่จะระบุไปยังคนใดคนหนึ่ง ไม่ควรเรียกระบุคนก่อนแล้วถาม เพราะคนอื่นจะไม่คิดเพราะคิดว่าตนรอดแล้ว

2. ผู้เรียนที่อ่อนหรือดูไม่ค่อยสนใจ ครูอาจรู้สึกขุ่นเคืองได้ง่าย ครูพึงระลึกไว้ว่า “ผู้เรียนที่ดูไม่ค่อยน่าสอน คือผู้เรียนที่ต้องการครูมากที่สุด” ผู้เรียนที่อ่อน หรือดูไม่ค่อยสนใจ ครูควรถามด้วยคำถามที่แคบ ปลายปิด มากขึ้น เล็กน้อย เพื่อให้ผู้เรียนเกิดกำลังใจ

3. ผู้เรียนที่เก่ง โดดเด่น ช่างตอบ ครูอาจเพลินสนุกไปด้วย แต่ “ผู้เรียนที่ดูน่าสอน คือผู้เรียนที่จริงๆ ไม่ต้องการครูเท่าใดนัก” ควรถามโดยระบุตัวไปยังผู้เรียนคนอื่นๆ ก่อน แล้วค่อยกลับมาให้ผู้เรียนคนนี้สรุป หรือเพิ่มเติม ส่วนที่เพื่อนยังไม่ได้อธิบายหรือไม่สมบูรณ์จะดีกว่า

4. ในกรณีที่ผู้เรียนหลายระดับ คำถามที่รุ่นพี่ตอบไม่ได้ ไม่ควรไปถามรุ่นน้องอีก เพราะทำให้เกิดความอับอาย ไม่เกิดประโยชน์ใดๆ

5. อย่าขัดขวางหรือตัดบทผู้เรียนที่กำลังพยายามจะตอบ

6. กระตุ้นให้ผู้เรียนตอบคำถามยาวๆ ยิ่งยาวยิ่งดี ยิ่งมีหลายคำตอบยิ่งดี ยิ่งมีการถกเถียงกันในกลุ่มยิ่งดี

7. กระตุ้นให้ผู้เรียนแสดงความเห็นกับคำตอบของเพื่อนในกลุ่ม

8. ถ้าผู้เรียนถามคำถามครู พยายามอย่าตอบคำถามในทันที แต่จงโยนคำถามนี้ไปให้กลุ่มตอบ ก่อนที่ครูจะตอบ

4. การถามแบบต่อห้อง

เมื่อผู้เรียนตอบคำถามแรกได้ ถ้าครูเห็นว่าเหมาะสม ควรถามต่อลงลึกมากขึ้น (deep exploring) หรือออกกว้างมากขึ้น (expanding) เพื่อขยายกรอบความคิดของผู้เรียนให้เข้าใจดีขึ้น แดกฉานขึ้น หรือกว้างขวางขึ้นในเรื่องนั้นๆ ผู้เรียนจะรู้สึกสนุกกับการได้เรียนรู้อะไรที่มากขึ้นเรื่อยๆ แต่ก็ควรให้เหมาะสมกับเวลาที่มี

5. การตอบสนองต่อผู้เรียน

1. ถ้าผู้เรียนตอบถูกต้อง ตอบได้ดี ครูควรชื่นชมอย่างจริงใจ

2. ถ้าผู้เรียนตอบผิด ไม่ควรรีบบอกว่าเป็นผิด ควรถามผู้เรียนต่อว่าทำไมจึงคิดเช่นนั้น (tell me more) เพื่อให้ทราบสาเหตุของความเข้าใจผิดและแก้ไขได้ตรงจุด อาจถามความเห็นในกลุ่มให้แสดงความเห็นสนับสนุนหรือโต้แย้งก็ได้ หรือครูอาจถามซ้ำโดยลดระดับคำถามให้พื้นฐานมากขึ้น หรือชี้แนะบางอย่างที่จะทำให้ผู้เรียนแก้ไขคำตอบใหม่ได้เอง หรือท้ายที่สุดครูขอขอบคุณผู้เรียนที่ช่วยตอบอย่างจริงใจ แล้วแก้ไขความเข้าใจที่ไม่ถูกต้องนั้นให้

3. ถ้าผู้เรียนถามคำถามครูกลับที่ครูเองก็ตอบไม่ได้ อย่าวิตก และจงภูมิใจในตัวผู้เรียน ครูอาจโยนคำถามนี้ให้กลุ่มก่อนซึ่งครูอาจประหลาดใจที่พบว่ามีคนตอบได้ ครูเองควรตอบผู้เรียนอย่างจริงใจว่า “ครูก็ไม่รู้”⁵ คำตอบนี้ไม่ทำให้ความน่าเชื่อถือของครูลดลง มีแต่ทวีให้ความน่าเคารพของครูมากยิ่งขึ้นเสียอีก ไม่ควรพยายามรักษาหน้า บ่ายเบี่ยง เฉลไฉนเปลี่ยนเรื่อง หรือให้การบ้านผู้เรียนเหมือนเป็นการลงโทษผู้เรียนไปเสียอีก

6. ถามแล้วผู้เรียนไม่ตอบ

มีเหตุผลมากมายที่ถามแล้วผู้เรียนไม่ตอบ ไม่ว่าจะเป็น ไม่เข้าใจคำถาม กำลังคิดอยู่ กำลังหากระแวง กลัวไม่กล้าตอบ (เพราะเคยตอบแล้วเจ็บปวดมาแล้ว) หรือกำลังวิตกแข่งกับครูว่าฝ่ายใดจะทนได้มากกว่า ดังนั้นสิ่งที่ครูควรทำคือ

1. ควรถามว่า “เข้าใจคำถามของครูมั๊ย” และปรับคำถามใหม่ถ้าผู้เรียนไม่เข้าใจ

2. รอคำตอบให้นานพอ (5-10 วินาที)

3. ประเมินว่าบรรยากาศเป็นอย่างไร ตั้งเครียดหรือไม่

4. กระตุ้น เชื้อเชิญให้ตอบเล็กน้อย ด้วยความจริงใจ ในที่สุดจะมีผู้เรียนตอบเสมอ (มักเป็นผู้เรียนที่เก่ง หรือมีจิตใจดี ไม่อยากให้ครูอึดอัดลำบากใจ)

5. ถ้าไม่มีคนตอบอีก ให้ถามระบุคน (pounce)

ปัญหาเหล่านี้มักเกิดขึ้นกับคำถามแรกๆ หรือเมื่อครูมิได้ชี้แจงผู้เรียนก่อนว่าจะใช้คำถาม แต่หลังจากมีผู้ตอบคำถามแรกแล้วพบว่าครูมีการตอบสนองอย่างดี ปัญหานี้จะหมดไป

7. ไม่ถามจนมากเกินไป

ควรมีความสมดุลระหว่างการถามคำถามชั้นสูง กับคำถามปลายปิดหรือแคบ การถามคำถามชั้นสูงอย่างต่อเนื่องนานเกินไป ทำให้ผู้เรียนต้องใช้ความคิดมากๆ อาจเหนื่อยล้าหรือเครียดได้ ในทางกลับกันการถามคำถามปลายปิดแคบๆ ติดต่อกัน ผู้เรียนจะรู้สึกว่าคุณคุกคามเช่นกัน จึงควรสลับลักษณะคำถามบ้างเป็นระยะๆ เพื่อคงความสดชื่นของกลุ่มไว้บ้าง ควรเลือกใช้การถามกับเรื่องที่เป็นประเด็นสำคัญๆ ก็เพียงพอ

สรุป

การถามเป็นวิธีสอนที่มีประโยชน์และประสิทธิภาพเป็นอย่างยิ่ง การใช้คำถามให้ประสบความสำเร็จประกอบด้วย การสร้างบรรยากาศ (SARA) การเลือกคำถามที่ดี (คำถามโซเครติก ไม่ทำกวม ที่ละ 1 คำถาม) และเทคนิคที่ดี (การรอ, Pose-Pause-Pounce-Bounce และการตอบสนองต่อคำตอบของผู้เรียนอย่างให้เกียรติ)

เอกสารอ้างอิง

1. Lake FR, Vickery AW, Ryan G. Teaching on the run tips 7: Effective use of questions. Med J Aust 2005;182:126-7.
2. Al-Umran K. Teaching tips - questioning. J Family Community Med 2004;11:73.
3. Nicholl HM, Tracey CA. Questioning: a tool in the nurse educator's kit. Nurse Educ Pract 2007;7:285-92.
4. Cho YH, Lee SY, Jeong DW, Im SJ, Choi EJ, Lee SH, et al. Analysis of questioning technique during classes in medical education. BMC Med Educ 2012;12:39.
5. Smith R. Thoughts for new medical students at a new medical school. BMJ 2003;327:1430-3.

เอกสารประกอบการอบรม



TEACH

7 November 2019

การให้ข้อมูลย้อนกลับ (Feedback)

วันชัย เดชสมฤทธิ์ฤทัย

ภาควิชาอายุรศาสตร์ คณะแพทยศาสตร์ศิริราชพยาบาล

การให้ข้อมูลย้อนกลับ (Feedback) เป็นองค์ประกอบหนึ่งที่สำคัญของการเรียนการสอนและฝึกอบรมทุกระดับ โดยเฉพาะอย่างยิ่ง ในปัจจุบันที่หลักสูตรการศึกษามีทิศทางไปสู่ outcome-based/competency-based education มากขึ้น การให้ข้อมูลย้อนกลับแก่ผู้เรียนอย่างถูกวิธี จะช่วยให้ผู้เรียน มีผลสัมฤทธิ์ทางการศึกษา และ ฝึกอบรม ตามผลลัพธ์ที่กำหนดไว้ได้ดีขึ้น

การให้ข้อมูลย้อนกลับ หมายถึง การให้ข้อมูลที่เกี่ยวข้องกับสถานะ สมรรถนะ พฤติกรรม ของผู้เรียนในกิจกรรมที่เกิดขึ้น โดยมีจุดมุ่งหมายให้เป็นแนวทางสำหรับผู้เรียนในการพัฒนาให้บรรลุถึงผลลัพธ์ และศักยภาพที่สูงสุด ของผู้เรียนแต่ละคน

แม้ว่าทั้งครูและผู้เรียนจะทราบถึงความสำคัญของการให้ข้อมูลย้อนกลับที่มีต่อการศึกษาและฝึกอบรม แต่จากการสำรวจพบว่า มีอุปสรรคหลายประการที่ทำให้ การให้ข้อมูลย้อนกลับไม่มีประสิทธิผลเท่าที่ควร อุปสรรคเหล่านี้ ได้แก่ การที่เป้าประสงค์ของการให้ข้อมูลย้อนกลับไม่ชัดเจน ผู้สอนไม่มีเวลา ไม่มั่นใจในวิธีการให้ข้อมูลย้อนกลับ รวมไปถึงประสบการณ์ในอดีตที่ไม่ดีที่เคยได้รับ ตลอดจน วัฒนธรรมองค์กรที่สั่งสมมา ปัจจัยเหล่านี้ล้วนทำให้ผู้สอนมีความลังเลที่จะให้ข้อมูลย้อนกลับแก่ผู้เรียน

ดังนั้น หากผู้สอนได้เรียนรู้และฝึกฝนทักษะของการให้ข้อมูลย้อนกลับที่ดี จะทำให้เกิดความมั่นใจมากขึ้น ในการให้ข้อมูลย้อนกลับ หลักการและแนวทางปฏิบัติต่อไปนี้ ผู้สอนสามารถนำไปประยุกต์ใช้ ให้เหมาะสมกับบริบทของตนเองต่อไป ได้แก่

การสร้างบรรยากาศของความไว้วางใจและกำหนดเป้าหมายร่วมกัน

ผู้เรียนจะยอมรับข้อมูลย้อนกลับหากเชื่อถือและไว้วางใจผู้สอน การสร้างบรรยากาศของความเชื่อถือไว้วางใจ ทำได้โดยที่ผู้สอนมีท่าทีรับฟัง การทำความตกลงกันก่อนว่า การให้ข้อมูลย้อนกลับ เป็นส่วนหนึ่งของกระบวนการเรียนการสอน และต้องอาศัยความร่วมมือของทั้งผู้เรียนและผู้สอน เป็นการสื่อสารสองทาง และผู้เรียนมีบทบาทสำคัญ ในการประเมินตนเอง การให้ข้อมูลย้อนกลับเป็นประจำสม่ำเสมอ รวมทั้ง การมีตารางเวลาชัดเจน จะช่วยให้ผู้เรียนเห็นว่าการให้ข้อมูลย้อนกลับเป็นเรื่องปกติในกระบวนการเรียนรู้ นอกจากนี้ การสื่อสารถึงความคาดหวังและเป้าประสงค์ ตั้งแต่แรก เพื่อความเข้าใจที่ตรงกันจะทำให้ การให้ข้อมูลย้อนกลับ ในเวลาต่อมาทำได้ง่าย

ให้ข้อมูลย้อนกลับโดยใช้ข้อมูลจากการสังเกตโดยตรง

ผู้เรียนจะยอมรับข้อมูลย้อนกลับที่ได้จากการสังเกตโดยตรงของผู้สอน ดังนั้นการสังเกตโดยตรง จึงเป็นพื้นฐานที่สำคัญมากต่อการให้ข้อมูลย้อนกลับ ทักษะทางวิชาชีพทั้งหลาย ได้แก่ การซักประวัติ การตรวจร่างกาย การสื่อสารกับผู้ป่วย และญาติ ฯลฯ ล้วนต้องการข้อมูลจากการสังเกตโดยตรง เพื่อให้ข้อมูลย้อนกลับ ด้วยเหตุนี้ ผู้ที่เหมาะสมในการให้ข้อมูลย้อนกลับ จึงมักจะเป็นผู้ที่ใกล้ชิดและมีโอกาสที่จะสังเกตผู้เรียนได้โดยตรงนั่นเอง อย่างไรก็ตาม บางครั้ง ผู้สอนอาจได้ข้อมูลจากแหล่งอื่น หากจำเป็น ควรที่จะตรวจสอบความถูกต้องของข้อมูลก่อนให้ข้อมูลย้อนกลับ และควรเปิดโอกาสให้ผู้เรียนได้ให้ข้อมูลก่อนเสมอ

ให้ข้อมูลย้อนกลับในจังหวะเวลาที่เหมาะสมและสม่ำเสมอ

ควรให้ข้อมูลย้อนกลับเร็วที่สุดเท่าที่โอกาสเอื้ออำนวย การปล่อยเวลาเนิ่นนานเกินไป จะทำให้ทั้งผู้เรียนและผู้สอนนึก รายละเอียดของเหตุการณ์ได้ไม่ครบถ้วน ทำให้อาจขาดประเด็นสำคัญไป ทั้งนี้ก่อนให้ข้อมูลย้อนกลับ ควรคำนึงถึง ปัจจัยต่อไปนี้ด้วย ได้แก่ อารมณ์ของผู้สอนในขณะนั้นควรเป็นกลาง อยู่ในสถานที่ปลอดบุคคลอื่น โดยเฉพาะอย่างยิ่ง หากเป็นการให้ข้อมูลย้อนกลับด้านลบ นอกจากนี้ อาจให้ข้อมูลย้อนกลับแบบไม่เป็นทางการ เช่น การขอพูดคุยกับผู้เรียนเป็นการส่วนตัวภายหลังเสร็จสิ้นกิจกรรม เป็นต้น การให้ข้อมูลย้อนกลับทั้งด้านบวกและลบสม่ำเสมอ จะทำให้ผู้เรียนรู้สึกว่าเป็นเรื่องปกติและยอมรับได้

เริ่มด้วยการให้ผู้เรียนประเมินตนเอง

กระบวนการให้ข้อมูลย้อนกลับ ควรเริ่มต้นโดยการเปิดโอกาสให้ผู้เรียนย้อนความคิด หรือประเมินตนเองก่อน โดยการ ใช้คำถามปลายเปิด ผู้เรียนอาจเปิดประเด็นที่ต้องการความเห็นหรือคำแนะนำจากผู้สอน ซึ่งก็จะเป็นจุดเริ่มต้นของการสนทนา บ่อยครั้งประเด็นนี้จะตรงกับที่ผู้สอนต้องการให้ข้อมูลย้อนกลับอยู่แล้ว ซึ่งจะทำให้การให้ข้อมูลย้อนกลับง่ายขึ้น ที่สำคัญคือผู้เรียนจะยอมรับข้อมูลย้อนกลับได้ดีขึ้น

ให้ข้อมูลย้อนกลับด้านบวก (positive feedback) ก่อน

ผู้สอนควรให้ข้อมูลย้อนกลับด้านบวกก่อน โดยอาจเป็นคำชม หรือเป็นการเน้นย้ำความถูกต้องในประเด็นนั้นๆ ข้อมูลเหล่านี้ จะช่วยเสริมความมั่นใจของผู้เรียนและกระตุ้นให้ทำต่อไป การให้ข้อมูลด้านบวกก่อน ยังช่วยให้เกิดบรรยากาศที่ดีต่อการสนทนา นอกจากนี้ การให้ข้อมูลด้านบวกเป็นประจำสม่ำเสมอก่อนหน้านี้ จะมีส่วนช่วยให้ผู้เรียนเกิดการยอมรับข้อมูลย้อนกลับด้านลบได้ดีขึ้น

ให้ข้อมูลย้อนกลับด้านลบ (negative feedback) ที่จำเพาะ และบรรยายเชิงพฤติกรรม

การให้ข้อมูลย้อนกลับด้านลบที่เป็นเชิงสร้างสรรค์ (constructive) ต่างจากการตำหนิติเตียน (criticism) ตรงที่การให้ข้อมูลย้อนกลับอย่างสร้างสรรค์นั้นมีความจำเพาะ ชี้ให้เห็นจุดที่บกพร่อง หรือไม่ตรงกับความคิดที่กำหนดไว้ ทำให้ผู้เรียนทราบได้ว่า จะพัฒนาปรับปรุงอย่างไร นอกจากนี้ การใช้คำพูดแนะนำให้ใช้วิธีการบรรยายเชิงพฤติกรรม หลีกเลี่ยงคำที่บ่งถึงตัวตน หรือบุคลิกภาพของผู้เรียน และหลีกเลี่ยงคำที่บ่งถึงการตัดสิน ตัวอย่างประโยคที่ควรหลีกเลี่ยง เช่น "คุณนี่แย่มาก ไม่รับผิดชอบ" ควรเปลี่ยนเป็น "การที่คุณขึ้นวอร์ดสายเป็นประจำ ถือว่าผิดระเบียบที่ตกลงกันได้" เป็นต้น นอกจากนี้ อวัจนภาษาที่เหมาะสม ก็มีผลสำคัญ เช่น การหลีกเลี่ยงการนั่งเผชิญหน้า การใช้โทนเสียงที่เหมาะสม เป็นต้น

ให้ข้อมูลย้อนกลับในปริมาณที่เหมาะสมและเลือกประเด็นที่สำคัญและแก้ไขได้ก่อน

บางกรณี อาจมีประเด็นที่ผู้สอนต้องการให้ข้อมูลย้อนกลับหลายประเด็น ผู้สอนควรประเมินสถานการณ์ว่า ควรให้ข้อมูลย้อนกลับปริมาณใดจึงเหมาะสม โดยอาจวางแผนการให้ข้อมูลย้อนกลับ แบ่งเป็นหลายครั้งได้ตามความเหมาะสม ควรเลือกประเด็นที่สำคัญและประเด็นที่สามารถแก้ไขได้มาให้ข้อมูลย้อนกลับก่อน

การให้คำแนะนำและการวางแผนแก้ไขปัญหาร่วมกัน

การให้ข้อมูลย้อนกลับควรเป็นการสื่อสารสองทาง ผู้สอนควรเปิดโอกาสให้ผู้เรียนซักถาม ด้วยท่าทีที่รับฟัง ในขณะเดียวกันผู้สอนควรประเมินว่าผู้เรียนมีความเข้าใจในประเด็นที่ให้ข้อมูลย้อนกลับเพียงใด รวมไปถึงการยอมรับของผู้เรียนต่อข้อมูลย้อนกลับที่ได้ ผู้สอนควรเปิดโอกาสให้ผู้เรียนเสนอวิธีการพัฒนาตนเองก่อน จากนั้นจึงให้คำแนะนำเพิ่มเติมหรือเสนอทางเลือกอื่น และกำหนดระยะเวลาและนัดหมายเพื่อติดตามผล สำหรับวิธีการให้

คำแนะนำแก่ผู้เรียนนั้น ควรใช้รูปประโยคที่ขึ้นต้นด้วยสรรพนามของผู้พูด ที่เรียกว่า "I message" ยกตัวอย่างเช่น "ผมอยากให้คุณขึ้นวอร์ดตรงเวลา" แทนที่จะพูดว่า "คุณควรขึ้นวอร์ดตรงเวลา" เป็นต้น อย่างไรก็ตาม ผู้สอนบางรายกลัวการให้ข้อมูลด้านลบ จึงพยายามเลี่ยงการให้ข้อมูลย้อนกลับด้านลบมาใช้เป็นการให้คำแนะนำแทน การทำเช่นนี้จะทำให้ผู้เรียนไม่ทราบถึงสถานะของตนเอง จึงทำให้ขาดพลังในการเปลี่ยนพฤติกรรมของผู้เรียน ดังนั้น จึงไม่ควรเลี่ยงประโยค "ข้อมูลย้อนกลับ" ซึ่งถือว่าเป็นส่วนสำคัญที่สุดของกระบวนการไป

สรุป

ทักษะการให้ข้อมูลย้อนกลับที่ดีนั้นมียุทธศาสตร์ประกอบหลักด้วยกัน 2 ส่วน คือ ความปรารถนาดี และ วิธีการที่ถูกต้อง หากขาดองค์ประกอบใดไปก็ทำให้การให้ข้อมูลย้อนกลับนั้นไม่เกิดประสิทธิผลที่ดี นอกจากนี้ ยังต้องอาศัยการฝึกฝนและการประยุกต์ให้เข้ากับบริบทและสถานการณ์ที่ต่างกันด้วย

เอกสารอ่านเพิ่มเติม

1. Ende J. Feedback in clinical medical education. JAMA 1983; 250:777-781.
2. Ramani S, Krackov SK. Twelve tips for giving feedback effectively in the clinical environment. Med Teach 2012; 34:787-791.

Learning activity ระดับชั้นปรีคลินิก

Lecture: Traditional vs. Interactive

Activity: Laboratory

Flipped classroom

Team-based learning

Group activity

ในการเรียนชั้นปริคlinik ยังต้องจัดให้เรียน **lab** อยู่หรือไม่?

ในการเรียนชั้นปริคlinik เราจัดการเรียนการสอน **lab** เพื่ออะไร?

Aims of the **lab** classes:

1. To underpin essential knowledge
2. To develop essential skills



Aims of the **lab** classes:

1. To underpin essential knowledge

ลำดับของการสอนทฤษฎีและปฏิบัติการ?

สอนทฤษฎีก่อน

สอนปฏิบัติการก่อน

สอนคู่กันไป



Aims of the **lab** classes:

2. To develop essential skills

ทักษะที่สามารถเรียนรู้ในห้องปฏิบัติการ?



Teaching strategies for the lab classes:

1. Controlled sessions
2. Experimental investigations



Organizing **lab** classes:

Outcomes

Evaluation

When *timing*

How *steps & duration, safety, errors*

Tasks **What** *tools, materials, staffs, time required*



**Facilitating Laboratory
and Practical Classes**

Teaching Development | Whānau Whakapakari Ako

Designing **lab** manuals:

Clear instructions

Proper illustration

Interesting questions

Elizabeth H. Hegarty & Adrian Lee (1979) How to...: Organize Effective Laboratory Teaching in Medicine. Part 2, Design, Medical Teacher, 1:5, 227-234.

Teaching skills in **lab** classes:

Performing (the tasks)

Describing & Explaining

Questioning

Directing

Giving feedbacks



Assessment:



Aims of the lab classes:

To underpin essential knowledge

To develop essential skills

เราจะพัฒนาการจัดการเรียน **lab** ในระดับชั้นปรีคลินิกได้อย่างไรบ้าง?



Mahidol University
Faculty of Medicine
Siriraj Hospital

Clinical Supervision

Suprapath Sonjaipanich MD.
Department of Pediatrics



A Competent Physician



- Clinical **skills** (Hx taking, PE)
- Technical and Procedural **skills**
- Communication and counseling **skills**



Definition: Supervision

“The provision of *guidance and feedback* on matters of *personal, professional & educational development* in the context of a trainee’s experience of *providing safe and appropriate patient care*”

AMEE Guide No. 27: Effective educational and clinical supervision. *Med Teach* 2007; 29: 2–19



Clinical Supervision

Increased
skill set of
learners

Improved
outcome
of patient
care





Clinical supervision

การสอนนักศึกษาฝึกฝนทักษะที่ต้องปฏิบัติต่อ
ผู้ป่วย

(ทักษะทางคลินิก ทักษะหัตถการ)

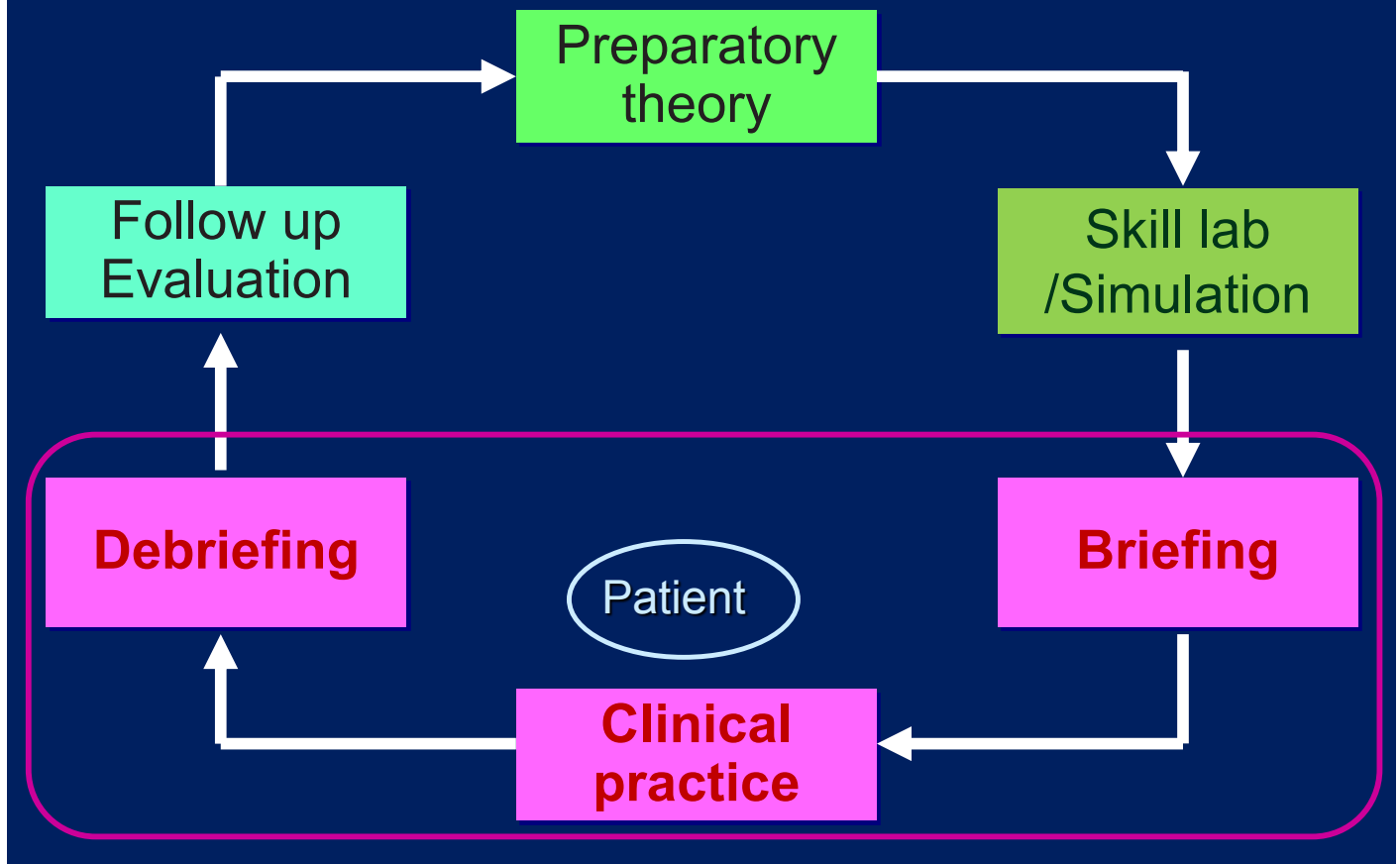


Learning objectives

- อธิบายหลักการและขั้นตอนการสอนในรูปแบบ supervision
- ประยุกต์ใช้หลักการและแนวทางของวิธีดังกล่าวในการสอนนักศึกษาฝึกทักษะหัตถการและทักษะทางการแพทย์ที่ต้องปฏิบัติกับผู้ป่วยได้ถูกต้องและมีประสิทธิภาพ



Clinical Learning Cycle





Supervision process





Supervision process

Briefing

- สอบถามประสบการณ์ที่ผ่านมาของนักศึกษา
- ทบทวนขั้นตอนของหัตถการ โดยเน้นเทคนิค ขั้นตอนที่สำคัญ เพื่อลดโอกาสเกิดข้อผิดพลาด
- กำหนดข้อตกลง หากนักศึกษาปฏิบัติไม่ถูกต้อง เกิดความไม่ปลอดภัยกับผู้ป่วย



Supervision process

Clinical
practice

- แนะนำตนเองและนักศึกษากับผู้ป่วย
- หลีกเลี่ยงการพูด สนทนา สอนระหว่าง
การปฏิบัติกับผู้ป่วย
- ใช้ภาษากาย/ท่าทาง เมื่อนักศึกษาปฏิบัติ
ไม่ถูกต้อง
- คำนึงถึงความปลอดภัยของผู้ป่วย

10



Supervision process

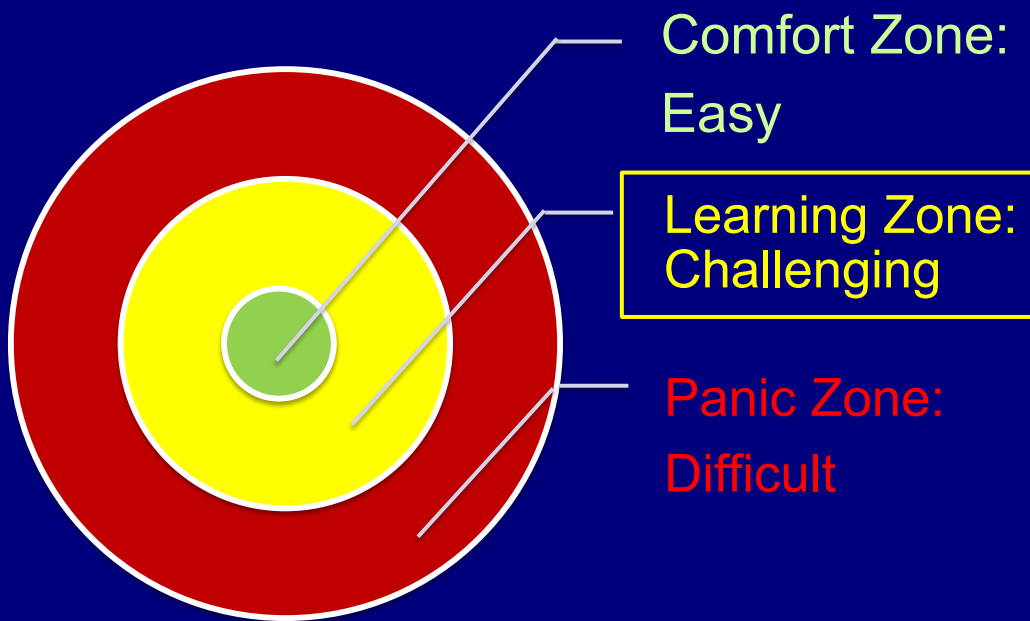
Debriefing

- นักศึกษา – reflection
- อาจารย์ – feedback (positive and negative feedback including suggestion)

11



Zone of learning





Supervision activities

- **Ensuring patients safety/care**
- **Educating the trainee**
- **Promoting high standard**
- **Identifying trainee problems**
- **Supporting the trainee**
- **Monitoring trainee progress**

AMEE Guide No. 27: Effective educational and clinical supervision. *Med Teach* 2007; 29: 2 -19



Effective Clinical Supervision

- Show a significant improvement in the process of patient care (quality and safety)
- Help trainees gain skills more rapidly

Snowdon et al. *BMC Health Serv Res* 2017

Tomlinson J. *BMC Med Educ* 2015

Kilminster & Jolly. *Med Educ* 2000



Questions & Comments

suprapath.son@mahidol.ac.th



15

การสอนนักศึกษาฝึกฝนทักษะที่ต้องปฏิบัติต่อผู้ป่วย (Clinical Supervision)

ผู้ช่วยศาสตราจารย์ นายแพทย์สุพัฒน์ สนใจพาณิชย์
ภาควิชากุมารเวชศาสตร์ คณะแพทยศาสตร์ศิริราชพยาบาล

“Supervision is an opportunity to bring someone back to their own mind, to shown them how good they can be.”

Nancy Kline

ในการเรียนระดับชั้นคลินิก ผู้เรียนจำเป็นต้องฝึกฝนทักษะต่างๆ เพื่อใช้ในการดูแลรักษาผู้ป่วย เช่น ทักษะการซักประวัติ การตรวจร่างกาย รวมถึงการทำหัตถการต่างๆ ซึ่งครูมีบทบาทสำคัญมากที่จะเป็นผู้นำหน้าที่กำกับดูแลให้คำแนะนำแก่ผู้เรียนในการฝึกปฏิบัติทักษะเหล่านั้น จนเกิดความชำนาญ

บทความนี้จะกล่าวถึงรูปแบบการสอนหนึ่งที่จะช่วยให้ครูสอนทักษะที่ต้องปฏิบัติต่อผู้ป่วยได้อย่างมีประสิทธิภาพ คือ การกำกับดูแล (supervision) โดยมุ่งเน้นที่ทักษะหัตถการ (procedural skill) การสอนโดยใช้เทคนิคนี้ได้อย่างถูกต้อง ย่อมส่งผลให้ผู้เรียนพัฒนาทักษะต่างๆ ได้อย่างรวดเร็ว

คำจำกัดความ

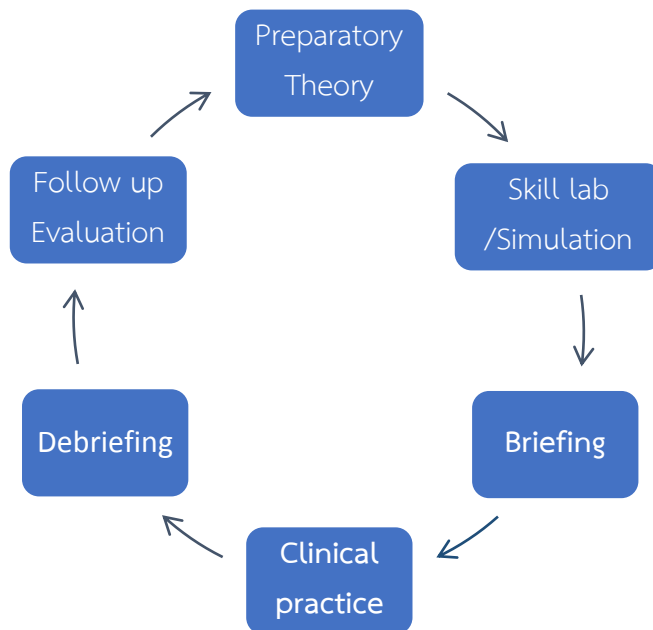
Kilminster และคณะ⁽¹⁾ กล่าวว่า การกำกับดูแล (supervision) เป็นกระบวนการสอนที่มุ่งเน้นการพัฒนาทักษะของผู้เรียนเพื่อให้มีความสามารถในการปฏิบัติต่อผู้ป่วยอย่างเหมาะสมและมีความปลอดภัย

Butterworth และคณะ (1965), Burton และ Launer (2003) และ Clark และคณะ (2006) ให้ความเห็นว่า⁽²⁾ การกำกับดูแลทางคลินิก (clinical supervision) ถือเป็นกระบวนการหลักของการพัฒนาทักษะทางวิชาชีพ และพัฒนาการเรียนรู้ตลอดชีวิต ซึ่งจะมีส่วนช่วยในการฝึกฝนทักษะที่ซับซ้อนในบริบทของสถานการณ์จริงที่มีความหลากหลาย

การฝึกทักษะทางคลินิกของผู้เรียน

การฝึกทักษะทางคลินิก ผู้เรียนควรได้รับการเตรียมความพร้อมด้านความรู้ และฝึกปฏิบัติ ในสถานการณ์จำลอง เช่น ในห้องฝึกทักษะ (skills lab) เพื่อให้เกิดความคุ้นเคยและมั่นใจก่อนที่จะปฏิบัติกับผู้ป่วยจริง ดังแผนภาพที่ 1

1 |Clinical Supervision



แผนภาพที่ 1 Clinical learning cycle

โดยเฉพาะทักษะหัตถการที่อาจมีความเสี่ยงในแง่ความปลอดภัยต่อผู้ป่วย ด้วยเหตุนี้จึงมีความจำเป็นอย่างยิ่งที่ผู้เรียนควรฝึกปฏิบัติกับหุ่นจำลองตามลำดับขั้นตอนต่างๆ ที่เสมือนจริงจนคล่อง โดยการฝึกทักษะหัตถการ ในสถานการณ์จำลองนี้ ครูสามารถสอนผู้เรียนโดยปฏิบัติ 4 ขั้นตอน⁽³⁾ ตามลำดับตามต่อไปนี้

1. Demonstration ครูแสดงการทำหัตถการโดยไม่ต้องพูดหรือบรรยาย
2. Deconstruction ครูสาธิตร่วมกับการอธิบายแต่ละลำดับขั้นตอนของหัตถการ
3. Comprehension ครูเป็นผู้ทำหัตถการโดยให้ผู้เรียนพูดขั้นตอนต่างๆ
4. Performance ผู้เรียนแสดงการทำหัตถการร่วมกับอธิบายแต่ละขั้นตอนของหัตถการด้วยตนเอง

การกำกับดูแลเมื่อฝึกปฏิบัติกับผู้ป่วย

ภายหลังจากผู้เรียนได้ฝึกทักษะในสถานการณ์จำลองจนเกิดความชำนาญและมั่นใจในระดับหนึ่ง ก็ต้องฝึกทักษะในสถานการณ์ที่มีผู้ป่วยจริงภายใต้การกำกับดูแลของครู ในกรณีที่ครูต้องดูแลควบคุมผู้เรียนทำหัตถการกับผู้ป่วย สามารถแบ่งการสอนเป็น 3 ขั้นตอน ดังแผนภาพที่ 1 ดังนี้

1. การซักซ้อม (Briefing) ก่อนผู้เรียนจะทำหัตถการกับผู้ป่วย ครูควรสอบถามถึงประสบการณ์การทำหัตถการที่ผ่านมา และให้ผู้เรียนทบทวนขั้นตอนต่างๆ ของหัตถการนั้น เพื่อเป็นการเตรียมความพร้อมและลดโอกาสเกิดข้อผิดพลาด ทั้งนี้ครูควรเน้นเทคนิค ขั้นตอนที่สำคัญที่มีโอกาสเกิดข้อผิดพลาดได้ง่าย เช่น หัตถการเย็บแผล สิ่งที่ต้องทบทวนคือ ขั้นตอนการทำให้ปลอดเชื้อ การฉีดยาเฉพาะที่ และเทคนิคการเย็บแผล เป็นต้น นอกจากนี้ ครูควรทำข้อตกลงกับผู้เรียนหากผู้เรียนปฏิบัติไม่ถูกต้อง เกิดข้อผิดพลาดหรืออาจเกิดความปลอดภัย

2 |Clinical Supervision

กับผู้ป่วยระหว่างทำหัตถการ เช่น ครูจะให้ผู้เรียนหยุดทำในกรณีใด หรือให้ผู้เรียนทำบางขั้นตอน โดยเฉพาะหัตถการที่ยุ่งยาก ซับซ้อนมีหลายขั้นตอน

2. การปฏิบัติกับผู้ป่วย (Clinical practice) ก่อนที่ผู้เรียนจะเริ่มทำหัตถการ ครูควรแนะนำตนเองและผู้เรียนกับผู้ป่วยรวมถึงขออนุญาตผู้ป่วยเพื่อทำหัตถการ โดยแจ้งว่าครูและผู้เรียนจะช่วยกันดูแลในการทำหัตถการนั้นๆ เป็นทีม เพื่อให้ผู้ป่วยเกิดความมั่นใจว่าจะได้รับการดูแลรักษาอย่างดี จากนั้นเปิดโอกาสให้ผู้เรียนอธิบายขั้นตอน วิธีการทำหัตถการแก่ผู้ป่วยพอสังเขป นอกจากนี้ควรจัดทำผู้ป่วยให้เกิดความสบาย เช่น นอนบนเตียง รวมถึงแนะนำให้ปิดตาผู้ป่วยเพื่อลดความกังวล สำหรับการทำหัตถการกับผู้ป่วยเด็ก หากเป็นไปได้แนะนำให้ผู้ปกครองรออยู่นอกห้อง

ขณะทำหัตถการครูควรเตรียมพร้อมที่จะเข้าช่วยเหลือเมื่อผู้เรียนพบปัญหาหรืออุปสรรคที่อาจส่งผลกระทบต่อความปลอดภัยของผู้ป่วย เช่น ในหัตถการเย็บแผล ครูควรใส่ถุงมือปลอดเชื้อตั้งแต่เริ่มต้น นอกจากนี้ควรหลีกเลี่ยงการสอนด้วยการพูดคุย สนทนา หากจำเป็นควรใช้ภาษากายท่าทางแทนการสื่อสารด้วยวาจา เช่น การพยักหน้าเพื่อแสดงว่าทำได้ถูกต้อง การแตะที่มือผู้เรียนเพื่อให้หยุดเมื่อปฏิบัติไม่ถูกต้อง เป็นต้น ไม่ควรกล่าวคำตำหนิ หรือแม้แต่คำชมขณะนักเรียนกำลังทำหัตถการ เพราะล้วนแต่ทำให้ผู้ป่วยรู้สึกเชิงลบทั้งสิ้น ในขั้นตอนนี้หากครูประเมินแล้วผู้เรียนปฏิบัติขั้นตอนใดของหัตถการไม่ถูกต้อง ครูอาจแสดงตัวอย่างให้ผู้เรียนได้เห็น เช่น ในการเย็บแผล ครูฉีดยาชาข้างหนึ่งของแผล และให้ผู้เรียนฉีดยาอีกข้างหนึ่ง ครูเย็บแผลเข็มแรกเพื่อให้ผู้เรียนสังเกตการปักเข็มที่ถูกต้อง เป็นต้น ทั้งนี้ต้องคำนึงถึงความปลอดภัยของผู้ป่วยเป็นสำคัญ

3. การทบทวน (Debriefing) ภายหลังเสร็จสิ้นการทำหัตถการ ครูควรเปิดโอกาสให้ผู้เรียนประเมินตนเองโดยสะท้อนความรู้สึกกับประสบการณ์ที่ได้รับ (reflection) หลังจากนั้นจึงเป็นช่วงเวลาที่คุณให้ข้อมูลป้อนกลับ (feedback) แก่ผู้เรียนทั้งสิ่งที่ปฏิบัติได้ดีและสิ่งที่ยังปฏิบัติไม่ถูกต้อง รวมถึงสิ่งที่ควรไปฝึกฝนเพิ่มเติม ซึ่งขั้นตอนนี้จะช่วยให้ผู้เรียนได้พัฒนาทักษะหัตถการนั้นได้ดียิ่งขึ้น

การกำกับดูแลผู้เรียนทำให้ผู้ป่วยเสี่ยงหรือไม่?

จากงานวิจัยพบว่า การสอนโดยวิธีการกำกับดูแลไม่ได้ส่งผลให้เกิดภาวะแทรกซ้อนของหัตถการเพิ่มขึ้น แต่เวลาที่ใช้ในการหัตถการอาจจะนานขึ้นเมื่อเปรียบเทียบกับ การลงมือปฏิบัติเองโดยครู

บทบาทของครูในการกำกับดูแล

การสอนโดยการกำกับดูแลครูเป็นผู้ที่มีบทบาทต่อทั้งผู้เรียนและผู้ป่วย^(1,4) สรุปได้ดังนี้

1. สร้างความมั่นใจให้กับผู้ป่วยทั้งในแง่ความปลอดภัยและมาตรฐานการรักษา
2. สอนทักษะการปฏิบัติต่อผู้ป่วยโดยส่งเสริมให้ผู้เรียนปฏิบัติได้อย่างมีมาตรฐานของทักษะนั้นๆ
3. ชี้ให้ผู้เรียนเห็นข้อบกพร่องหรือสิ่งที่ควรพัฒนาและฝึกฝนเพิ่มเติม
4. ให้ความช่วยเหลือและติดตามความก้าวหน้าของผู้เรียน

คุณลักษณะที่ดีของผู้ที่จะสอนโดยการกำกับดูแล

คุณลักษณะที่ดีของครูที่จะสอนโดยการกำกับดูแล^(1,5) ได้แก่

1. มีความรู้ความสามารถในทักษะต่างๆ ทางคลินิก
2. มีทักษะในการสื่อสารและสร้างสัมพันธภาพที่ดี
3. มีทักษะในการสอน
4. มีทักษะในการให้ข้อมูลย้อนกลับ และการให้คำแนะนำ
5. สามารถเชื่อมโยงความรู้ทางทฤษฎีสู่การปฏิบัติจริง
6. แสดงแบบอย่างที่ดีในการปฏิบัติต่อผู้ป่วย

สรุป

การกำกับดูแลผู้เรียนเพื่อฝึกทักษะทางคลินิกเป็นบทบาทหน้าที่ที่สำคัญของครู หากครูใช้เทคนิคที่ถูกต้อง จะส่งผลดีต่อทั้งผู้ป่วยและผู้เรียน กล่าวคือ ผู้ป่วยมีผลการรักษาที่ดี ขณะเดียวกันผู้เรียนได้ฝึกฝนทักษะต่างๆจนเกิดความชำนาญและมีความมั่นใจ

เอกสารการอ่านเพิ่มเติม

1. Kilmimster S, Cottrell D, Grant J & Jolly B. AMEE Guide No.27: Effective educational and clinical supervision. *Med Teach* 2007; 29: 2-19
2. Davis M, Dent J. Getting started with educational supervision: Making the most of your role as an educational supervisor. Centre for medical Education, University of Dundee; 2008.
3. George JH, Doto FX. A simple five-step method for teaching clinical skills. *Fam Med* 2001; 33: 577-8
4. Halpern H, McKimm J. Clinical Teaching Made Easy: Supervision. *Brit J Hosp Med* 2009; 70: 226-9
5. Hesketh E, Laidlaw J. Developing the teaching instinct: Supervision. *Med Teach* 2002; 24: 364-7

AMEE Guide No. 27: Effective educational and clinical supervision

SUE KILMINSTER¹, DAVID COTTRELL¹, JANET GRANT² & BRIAN JOLLY³

¹University of Leeds, UK, ²Open University Centre for Education in Medicine, UK, ³University of Monash, Australia

Abstract

Background: This guide reviews what is known about educational and clinical supervision practice through a literature review and a questionnaire survey. It identifies the need for a definition and for explicit guidelines on supervision. There is strong evidence that, whilst supervision is considered to be both important and effective, practice is highly variable. In some cases, there is inadequate coverage and frequency of supervision activities. There is particular concern about lack of supervision for emergency and 'out of hours work', failure to formally address under-performance, lack of commitment to supervision and finding sufficient time for supervision. There is a need for an effective system to address both poor performance and inadequate supervision.

Supervision is defined, in this guide as: 'The provision of guidance and feedback on matters of personal, professional and educational development in the context of a trainee's experience of providing safe and appropriate patient care.' A framework for effective supervision is provided:

(1) Effective supervision should be offered in context; supervisors must be aware of local postgraduate training bodies' and institutions' requirements; (2) Direct supervision with trainee and supervisor working together and observing each other positively affects patient outcome and trainee development; (3) Constructive feedback is essential and should be frequent; (4) Supervision should be structured and there should be regular timetabled meetings. The content of supervision meetings should be agreed and learning objectives determined at the beginning of the supervisory relationship. Supervision contracts can be useful tools and should include detail regarding frequency, duration and content of supervision; appraisal and assessment; learning objectives and any specific requirements; (5) Supervision should include clinical management; teaching and research; management and administration; pastoral care; interpersonal skills; personal development; reflection; (6) The quality of the supervisory relationship strongly affects the effectiveness of supervision. Specific aspects include continuity over time in the supervisory relationship, that the supervisees control the product of supervision (there is some suggestion that supervision is only effective when this is the case) and that there is some reflection by both participants. The relationship is partly influenced by the supervisor's commitment to teaching as well as both the attitudes and commitment of supervisor and trainee; (7) Training for supervisors needs to include some of the following: understanding teaching; assessment; counselling skills; appraisal; feedback; careers advice; interpersonal skills. Supervisors (and trainees) need to understand that: (1) *helpful supervisory behaviours* include giving direct guidance on clinical work, linking theory and practice, engaging in joint problem-solving and offering feedback, reassurance and providing role models; (2) *ineffective supervisory behaviours* include rigidity; low empathy; failure to offer support; failure to follow supervisees' concerns; not teaching; being indirect and intolerant and emphasizing evaluation and negative aspects; (3) in addition to supervisory skills, effective supervisors need to have good interpersonal skills, good teaching skills and be clinically competent and knowledgeable.

Introduction

Why the Guide?

What is good educational supervision and who are the good supervisors? Documentation from the UK Department of Health (DoH 1996) and General Medical Council (GMC 2005 (New Doctor), 2006 (Good Medical Practice), 1999) has highlighted the need for good educational supervision, appraisal and assessment in postgraduate education. However, it is not always clear what supervision entails, who should or could supervise, what the effects of supervision are

and moreover, what its benefits to patients and the service in general are. It is clear that some doctors receive excellent supervision. It is also clear that others receive inadequate supervision (Grant et al. 2003).

Effective supervision of trainees involves skills that are different from other more general competences expected of a teacher or trainer (Harden & Crosby 2000; Hesketh et al. 2001). Supervision includes ensuring the safety of the trainee and patient in the course of clinical care; giving feedback on performance, both informally and through appraisal; initial training and continuing education planning; monitoring progress; ensuring provision of careers advice; ensuring an

Correspondence: David Cottrell, Department of Child and Adolescent Psychiatry, University of Leeds, UK; email: d.j.cottrell@leeds.ac.uk

2

ISSN 0142-159X print/ISSN 1466-187X online/07/010002-18 © 2007 Informa UK Ltd.
DOI: 10.1080/01421590701210907

RIGHTS LINK
Copyright Clearance Center

Practice points

- Although supervision is recognized as important and effective, actual practice is very variable and there is a need for a definition and explicit guidelines.
- This guide provides a framework for effective supervision—direct supervision, constructive feedback, structure and the quality of the supervisory relationship are particularly important.
- Helpful supervisory behaviours include giving direct guidance on clinical work, linking theory and practice, engaging in joint problem-solving, offering feedback and reassurance and providing role models.
- There is a need for an effective system to address both poor performance and inadequate supervision.

appropriate level and amount of clinical duties. Supervision features more observation, continual feedback and sharing of clinical judgement. Supervision has been the least researched and supported aspect of medical education and yet is central to the effective training of physicians. This guide is the culmination of a research project designed to investigate the role of educational and clinical supervision and the skills required. It draws on relevant empirical and theoretical work to offer a practical, informative guide to good supervision.

Who is it for?

The Guide is for anyone who supervises others in medical clinical practice settings. It is based on work conducted in the UK but can be applied elsewhere. Although it has been targeted at education in the postgraduate setting it may also have some relevance in undergraduate medical education. It will also be useful for trainees. We have attempted to describe the roles and terms of reference of all the key players, with descriptions of the supervisory tasks necessary at each level of training from newly qualified doctor to Consultant and including the roles of the regulatory and statutory bodies. The Guide should also be of help to those managing, monitoring or delivering training.

What is supervision?

There are various understandings and definitions of supervision; based on the findings of our research project *Good supervision: Guiding the clinical educator of the 21st century* (Kilminster et al. 2000) we define supervision as:

The provision of guidance and feedback on matters of personal, professional and educational development in the context of a trainee's experience of providing safe and appropriate patient care.

We would hope that the trainee's care would be safe and appropriate at all times. However, the definition recognizes that some benefit can be derived from analysis of errors, their management and resultant lessons. The anticipatory element of supervision is necessary to isolate and deal with threats to patients' safety. The 'personal' issue in the definition is an

attempt to acknowledge that many problems with competence can arise from personality-related variables and that these are often the most difficult aspect to deal with for the supervisor and trainee.

Educational and clinical supervision in context

Ensuring patient and trainee safety

Postgraduate medical training is the process whereby newly qualified doctors—in the UK these are pre-registration House Officers (PRHOs)—progress through a series of training stages until they become trained and certified specialists or general practitioners. From an educational point of view, different processes are involved in this progression. Some of the body of knowledge and many of the skills and attitudes that the trainee doctor acquires whilst progressing along this road will come from his/her own self-directed private study; some will also come from the formal educational activities of the training programme on which he/she is enrolled. However the most important element of training for nearly all doctors is the opportunity to undertake medical practice in their chosen speciality under appropriate supervision.

Clinical supervision must have patient safety and the quality of patient care as its primary purposes but must also fit in with the trainee's educational objectives. Clinical responsibility for patient well-being lies with the supervising consultant who is in turn responsible to the chief executive of the clinical service, hospital or Trust.

The nature of clinical supervision will vary from speciality to speciality and from unit to unit. The nature of the speciality (surgical or non-surgical for example), location (primary care or hospital) and the structure of the clinical team providing the service will be the primary determinants of the sort of supervision required, but in all cases the object of supervision will be the same: to provide the patient with the best possible quality service under the prevailing circumstances and to provide the community from which that patient comes with the quality of service which meets its needs. The processes that ensure patient safety are essentially educational and form the backbone of the trainee's clinical learning.

Patient safety. It may be perfectly safe for a highly competent practitioner to see and examine a patient in the home, where conditions are often less than ideal. Put an inexperienced trainee in that same position and it becomes less safe. The role of the supervisor may also be considered at different levels. A senior manager or consultant supervisor may well have responsibility for the working environment, whereas a Specialist Registrar (SpR or senior resident) supervising a Senior House Officer (SHO or junior resident) carrying out an emergency appendectomy will have limited responsibility for the work environment, but does have a great responsibility for ensuring that the procedure itself is carried out safely. It is important therefore that supervisors understand their responsibilities with regard to patient safety.

The practice of medicine has evolved in a way that has left many trainees working with minimal supervision. Whilst this

3

S. Kilminster et al.

may be perfectly reasonable, it does place great responsibility on the supervisor to ensure that the trainee is competent and performing only at an appropriate level. Accomplishing this without compromising patient safety may be very difficult. In the absence of a clear, explicit indication of the level of competence of a trainee it will be necessary for the supervisor to ascertain this either by direct observation or enquiry from other staff. Only then can the correct levels for practice and supervision be determined and applied without compromising patient safety. It is important to ensure that appropriately qualified supervisors are available, for example, when trainees are providing out-of-hours care.

Trainee safety. Ensuring the safety of trainees in the course of their clinical duties is an essential element of supervision. Trainees are less likely to acquire new competences in environments where they are in fear of being exposed to risk—the major factor in determining levels of psychological distress in trainees is their confidence in carrying out the clinical tasks expected of them (Williams et al. 1997).

Supervisors should ensure that trainees work within their competences and that they are adequately supervised when acquiring new skills. In addition:

- Trainees should receive adequate induction to training placements.
- Opportunities must be made available to reduce stress by ensuring availability of stress counselling and training in communication skills. Trainees must be made aware of these resources.
- Trainees who are required to undertake procedures that may expose them to risk (e.g. handling of surgical equipment or making up toxic drugs) should receive adequate instruction and protection.
- Trainee safety should not be compromised by onerous duty rotas or excessive service commitments.
- Adequate procedures must be in place for prevention and control of transmissible infectious diseases.
- Personal safety from attack must be ensured.
- Procedures in the event of fire and other emergencies must be in place and trainees must be made aware of them.

Overseas trainees have special needs; it cannot be assumed they have the same level of understanding of local healthcare systems as doctors who have trained in the country concerned and therefore they may need more carefully planned induction.

Supervisors themselves need to be competent in the skills to be acquired and in dealing with the complications that may arise from using these skills. Trainees need to have confidence in their supervisors: this is particularly important when responsibility for teaching has been delegated to staff other than the supervisor. Supervisors need to monitor the quality and effectiveness of education and supervision carried out in their name.

4

Ensuring trainee competence and level of supervision

The content of what needs to be supervised at different levels will change but the level of supervision will vary according to the grade and relevant experience of the trainee. Supervisors need to make judgements as to whether they should be:

- present in the same room as the person being supervised, providing direct supervision (direct supervision);
- nearby and immediately available to come to the aid of the person being supervised (immediately available supervision);
- in the hospital or primary care premises and available at short notice, able to offer immediate help by telephone and able to come to the aid of the person within a short time (local supervision);
- on call and available for advice, able to come to the trainee's assistance in an appropriate time (distant supervision).

Training log books can be useful tools in helping to determine the level of supervision required.

Supervision in clinical teams

Clinical teams are hierarchically structured and the responsibility for clinical supervision does not lie solely with the consultant or general practitioner principal who is at the head of the team. For example the main responsibility for the clinical supervision of a medical team on emergency take usually lies with a specialist registrar or senior resident who will directly supervise the activities of the more junior staff who are delivering the care. This produces a system of great complexity for all the team members. Responsibility is distributed in an uneven fashion throughout the team. The consultant has responsibility for the overall functioning of the medical team and for the individual clinical performance of all the team members. The consultant has clinical responsibility for the decisions that lead to individual team members working without direct supervision. It is clearly not possible, nor would it be appropriate, for the consultant to allocate work on a case-by-case basis. However, it is essential for the consultant to understand and orchestrate the process by which individual team members are working with more or less direct supervision in different clinical areas. The processes whereby this may happen and how they need to be negotiated also needs to be understood within the team. It is important for the development of even the most junior team member that he/she has areas of clinical activity for which he/she takes direct responsibility and only reports back to a more senior member when he/she judges the need to do so. It is self-evident that the extent of the less directly supervised domain will be large for experienced senior trainees and much smaller for junior trainees. Senior trainees require instruction in, and experience of, supervising more junior staff. Although a junior trainee may refer to them as their first line of advice and assistance, both the junior and senior trainee will be subject to supervision from a designated consultant. There will be some occasions during highly specialized training when it will be inappropriate

for senior trainees to act as supervisors—they may themselves require direct supervision.

The situation becomes further complicated in teams with more than one consultant supervisor, now the norm rather than the exception. A single named consultant may take on the role of 'educational supervisor' for the team. This role concerns arranging appropriate educational support for trainees and ensuring that they are exposed to appropriate clinical experience and responsibility. The task of clinical supervision and the process of taking vicarious responsibility for patient care delivered by trainees will fall to each of the consultants whose patients are looked after by the clinical team in question.

Employers' responsibility for supervision

Employers need to ensure that the arrangements for the delivery and monitoring of supervision are practical, robust and transparent although, ultimately, designated supervisors are responsible for ensuring that clinical supervision takes place in accordance with local clinical guidelines and external advice (in the UK from regional training committees, Royal Colleges and the General Medical Council).

Individual trainers have to manage the conflict between the need to provide a clinical service and the need to provide adequate supervision for the trainees for whom they are responsible. Within healthcare organizations, the lines of responsibility are through clinical directors and medical directors to the chief executive. In all cases, the trainee-trainer interface is the local level of accountability. Trainer and trainee have at their disposal advice and support from the local course organizers, speciality training bodies and external regulatory bodies. (In the UK this includes, Directors of Postgraduate Medical Education, College tutors and programme directors, Royal Colleges and the Specialist Training Agency.)

In addition to ensuring that all doctors in training receive adequate supervision in an appropriate environment, employers (Trusts in the UK) will need to ensure that they have in place systems that can deal with:

New doctors who have not worked in the hospital or practice before:

- How is an assessment of competence made?
- How much direct supervision is needed before allowing the person 'clinical freedom'?
- How much trust can be placed on the appointment process to select doctors who can be relied upon to perform at an appropriate level of competence?

Locum doctors:

- Who has responsibility for the clinical performance of locums?
- How much direct supervision should occur before allowing the locum 'clinical freedom'?
- How much reliance can be placed upon the agencies that provide the locums?

Non-consultant career-grade doctors:

- Who is responsible for supervising the clinical performance of staff-grade doctors?
- Who is responsible for ensuring that staff-grade doctors avail themselves of educational opportunities and keep up to date with developments and current practice in their speciality?

Some lines of accountability within individual organizations are relatively straightforward as clinical activity is delivered either by clinical teams or by individual departments or practices. The main line of accountability will involve doctors within the same speciality or practice, some of whom will have a designated supervisory role. However, there will also be circumstances in which accountability involves doctors from another speciality. An example would be a consultant anaesthetist supervising aspects of a trainee surgeon's work. The consultant anaesthetist might be expected to carry some responsibility for ensuring that the trainee surgeon performs at an appropriate level. There are also inter-professional lines of accountability involving other healthcare personnel, for example, nurses, technicians, operating department assistants.

The quality of clinical supervision of trainees is therefore a central problem for clinical governance organizations within the medical management structure, and these organizations need to assure themselves that appropriate supervision is being undertaken. The complexity of lines of responsibility for trusts and individual consultants is shown by the example in Box 1.

Box 1: The outpatient treatment of varicose veins by injection

What if the patient has a cardiac arrest during the injection of varicose veins? Resuscitation equipment must be readily to hand, in good working order and the staff trained to use it. There is a management role here in 'supervising' facilities and in ensuring the training of nursing staff. The consultant or competent specialist registrar supervising a junior doctor in training has a duty to ensure that the training of the junior doctor encompasses the possibility of a cardiac arrest. Does the junior doctor know where the equipment is kept? Is she/he competent in resuscitation techniques? If the answer is 'no' to either of these questions, it is surely the duty of the trainer as a supervisor to see that these deficits are rectified. It should be appreciated that the supervisor does not necessarily need to train the junior doctor in resuscitation, but *does* need to ensure that proper induction has been organized and that the necessary training takes place.

Within the UK, the responsibility for good clinical governance in Trusts lies with chief executives and through them medical directors, clinical directors and individual consultants. The General Medical Council has emphasized that Trusts must ensure that the time and resources necessary for encouraging and sustaining a culture of education are available, and that the environment is adequate.

5

S. Kilminster et al.

Poor performance

Every employer will have to create a system for identifying and dealing with doctors who exhibit a persistent pattern of poor performance. The concept of Clinical Governance has encouraged the creation of clinical standards committees and/or clinical governance committees, which carry out their functions with clearly defined lines of accountability to the Trust Board and Chief Executive. The educational system has a role when poor performance relates to educational issues and a failure to progress, with increasing competence, in a placement. The Education Supervisor would then discuss this with either the Clinical Tutor or College Tutor, and might then refer on to the Deanery, particularly if the trainee is in a managed scheme. The Trust would be involved when issues of personal conduct such as lateness, rudeness or neglect of clinical responsibilities were the reason for poor performance.

What is known about current supervisory practice?

The literature

This section summarizes the literature on supervision in practice settings in order to identify what is known about effective supervision. Relatively, there is a limited amount of published medical literature addressing supervision; particularly, there are few empirical studies (Kilminster & Jolly 2000). Supervision is a complex activity, occurring in a variety of settings, and has various definitions, functions and modes of delivery. It usually includes an interpersonal exchange. This complexity means that research into supervisory practice presents methodological problems and adequate research methodologies have yet to be established.

What are the understandings and definitions of supervision and its purposes?

There appears to be general agreement that the essential aspects of supervision are that it should ensure patient/client safety and promote professional development. Clearly, there may be some occasions when these two aspects are in opposition.

There is also agreement in the general literature that supervision has three functions—educative, supportive and managerial or administrative. In medicine, this would include guiding patient management.

What are the theoretical models of supervision?

Various models are presented in the psychotherapy, social work and nursing literature. Common features of these models include the idea that supervisory behaviours can be categorized and that supervision needs vary according to the recipient's level of experience. Most models stress the need to use supervision approaches that are appropriate to the trainee's level of experience and training.

There is some limited empirical support for the proposition that supervision needs vary according to the trainee's experience and level of training.

There are no adequate theoretical accounts of supervision in medicine; such an account of supervision in medicine might draw on ideas developed in adult learning theories, experiential and work-based learning as well as understandings about apprenticeship and development of expertise (for example, Kolb 1984; Patel & Groen 1990; Lave & Wenger 1991; Boud et al. 1993; Eraut 1994; Tenant 1999) but would also need to connect with educational strategies used throughout medical education, including the problem-based learning approach, skill development and apprenticeship.

How is supervision delivered—what is its structure and content?

The evidence indicates that there are wide variations in the frequency and amount of supervision that trainees in the UK receive (Kilminster et al. 2000). In particular, there are marked variations across and between specialties. Where guidelines exist they are not always met. The variation is so great that it cannot simply be explained by variations in individual learning. Problems with the extent and availability of supervision have been identified across the professions. The quality of supervisory interactions remains to be investigated in depth. Supervision can occur 'on the job', usually whilst a practical task is being carried out; informally; in a one-to-one meeting; in peer supervision; in group supervision; and in networking. There is empirical evidence (including some of our own work) indicating that finding sufficient time for supervision can be a problem; some strategies have been suggested to address this but more are needed.

Is supervision effective and how can this be determined?

There is some convincing quantitative evidence, across health and social care professions, that supervision has a positive effect on patient outcome and that lack of supervision is harmful to patients. In particular, empirical evidence shows that direct supervision is very important and can positively affect patient outcome and trainee development, especially when combined with focused feedback.

Review evidence suggests that increased deaths are associated with less supervision of junior doctors in surgery, anaesthesia, trauma and emergencies, obstetrics and paediatrics (McKee & Black 1992). These authors argue that the balance of evidence shows that patient care suffers when trainees are unsupervised even though some trainees claim to benefit from the experience that lack of supervision gives them. However, they also argue that unsupervised experience can lead to the acceptance of lower standards of care because the trainee may not learn correct practice without appropriate supervision.

In the USA, strong evidence for the importance of direct supervision was obtained by comparing attendings' (senior doctors equivalent to UK consultants) own findings regarding patients with their ratings of residents' (equivalent to specialist

registrars) reports and history taking, assessment of severity of the patients' illness, diagnoses, treatment and follow-up plan (Genniss & Genniss 1993). The researchers found that the attendings' assessments of the residents were more critical after seeing the patients and that they considered seeing the patients themselves to be important for both teaching and management. The patients were seen as more seriously ill, and there were frequent changes in diagnosis and management. The authors indicate that there were some weaknesses in the study design (it was not a randomized trial so the results could be due to the order of evaluation and changes in treatment were often minor and therefore could be due to differences in opinion). They do conclude, however, that, when supervisors see the patient themselves rather than relying on trainees' reports there is a significant difference in their assessments of residents' skills and patient management.

The effects of supervision on quality of care were examined in five Harvard teaching hospitals (Sox et al. 1998). A range of measures was used—residents' compliance with process-of-care guidelines (assessed by record review), patients' satisfaction and patients' reported problems with care. Over a seven-month period all 3667 patients presenting with abdominal pain, asthma/COPD, chest pain, hand laceration, head trauma and vaginal bleeding were included; residents were unaware of the purpose of the study. All patients were given a questionnaire to complete on site and some were randomly selected for a 10-day follow up interview. Analyses were adjusted for case mix, degree of urgency and chief complaints. Using these measures the researchers found that the quality of care was higher when the resident was directly supervised, i.e. when the attending also saw the patient. The benefits of direct supervision of residents applied regardless of the level of training and urgency of the cases. The authors point out that there are limits to the generalizability of the study because the five hospitals did not have emergency medicine training programmes, there may be between-hospital variations in quality and frequency of supervision, patients were not randomized to different groups and there was no control for the speciality of the attending physician.

Faculty involvement was investigated, over a 12-month period, for each surgical procedure and all resuscitation and operations in the trauma service in one hospital (Fallon et al. 1993). Faculty involvement was ranked on a five-point scale and these data were matched to outcomes of death or complications that were reported in the weekly departmental complications conference. The results suggested that supervision had a greater impact where the trainee was less experienced. The authors acknowledge a number of limitations to their study but conclude that close supervision of general surgical residents during their rotations to subspecialties is important and that the effect of supervision can be evaluated by using probability of survival data in trauma. They also argue that there is a need to establish measurable standards of supervision.

Griffiths et al. (1996) compared tests (X-rays, arterial blood gases (ABG) and electrolytes) ordered in the neonatal intensive care unit by staff with different levels of experience. They found that as workload increases newly qualified doctors order more ABG, especially when they are less supervised.

To summarize, empirical evidence from the literature review shows that:

- Direct supervision seems to help trainees gain skills more rapidly.
- The quality of the supervisory relationship strongly affects the effectiveness of supervision. Particularly important aspects are continuity over time in the supervisory relationship, the trainees having some control over the supervision (there is some suggestion that supervision is only effective when this is the case) and that there is some reflection by both participants.
- Behavioural changes can occur relatively quickly as a result of supervision whilst changes in thinking and attitude take longer. This is particularly important because there may be relatively frequent changes of supervisor due to rotations.
- Self-supervision is not effective; input from a supervisor is required.

The supervision environment is extremely important because medical students have strategies to appear as competent as possible, which can conflict with opportunities to learn (Jolly & MacDonald 1986). In addition, trainees can perceive 'one to one consultations as problematic and risky situations in which they struggle[d] for a balance between the opportunity to learn and the need to perform in and manage the consultation process' (Somers et al. 1994, p. 587). There is compelling evidence that postgraduate trainees engage in similar behaviours (Arluke 1980). Clearly, such defensive behaviours are likely to have an effect on the supervision process and, ultimately, that may not be beneficial to patients.

What skills and qualities do effective supervisors need?

Empirical and review evidence indicates that, to be effective, in addition to supervisory skills supervisors need to have good interpersonal skills, good teaching skills and be clinically competent and knowledgeable. The distinction between supervision and teaching is not easily made. However, empirical and review evidence indicates that:

- (1) *Helpful supervisory behaviours* include giving direct guidance on clinical work; linking theory and practice; engaging in joint problem-solving and offering feedback, reassurance and providing role models.
- (2) *Ineffective supervisory behaviours* include rigidity; low empathy; failure to offer support; failure to follow supervisees' concerns; not teaching; being indirect and intolerant and emphasizing evaluation and negative aspects.
- (3) *Good interpersonal skills* include involving trainees in patient care; negotiation and assertiveness skills; counselling skills; appraisal skills; self-awareness; warmth; empathy; respect for others; listening skills; expressing one's own emotions appropriately; offering support; being positive; having enthusiasm.
- (4) *Clinical competence* includes being seen as a good clinician and having up-to-date theoretical and clinical knowledge.

7

S. Kilminster et al.

- (5) *Teaching skills* include offering opportunities to carry out procedures; giving direction; giving feedback; having knowledge of teaching resources; knowledge of certification requirements; individualizing the teaching approach; being available and having evaluation skills.

Studies reporting on characteristics of effective clinical teachers have some relevance for supervisors. The characteristics include having clinical credibility; having knowledge of context, learners and general principles of teaching including the importance of feedback and evaluation; being a positive role model and appearing to enjoy teaching.

In addition trainees need clear feedback on their errors; corrections must be conveyed unambiguously so that trainees are aware of mistakes and any weaknesses they may have.

How these skills should be assessed and how supervisors should be selected is not discussed in the literature. Some empirical and review evidence suggests that race and gender dynamics are areas of potential difficulty in supervisory relationships.

What training do supervisors need and how can its effectiveness be determined?

The need for training is widely accepted and there is some evidence that it can be effective. There is agreement that training probably needs to include at least some of the following: understanding teaching; assessment; counselling skills; appraisal; feedback; careers advice; interpersonal skills. Course content should emphasize the importance of understanding the concept and purposes of supervision; understanding the content and type of training undertaken by the supervisee; understanding the structure and types of supervision including the importance of a supervision contract, giving and receiving criticism, counselling skills and interpersonal dynamics.

Some commentators consider there should be some criteria regulating entry into supervisor training courses or for acceptance as a supervisor.

Supervision research project: empirical findings

Supervision, both educational and clinical, is an essential part of Specialist Registrar (SpR) training (DoH 1996) although there is relatively little guidance as to how and where this should take place. There are no large-scale studies describing supervision practices in medical education (Kilminster & Jolly 2000) and so relatively little is known about how supervision takes place in different specialities. Therefore, as part of a Department of Health funded project investigating supervision (Kilminster et al. 2000), we undertook a national questionnaire survey to identify the range and effectiveness of supervisory methods for SpRs in current usage. The purpose of the survey was to establish what supervisory methods were being used and to determine how effective, particularly in relation to effects on patient care, education supervisors (ESs), specialist registrars (SpRs) and medical directors (MDs) perceived these

8

methods to be. We were interested in the general situation rather than a detailed examination of one particular area (geographic and/or speciality) and intended to evaluate the findings in the context of clinical governance. We also undertook an exploratory critical incident study to identify key features of effective supervision from the perspectives of SpRs and ESs (Cottrell et al. 2002).

Our findings suggest that whilst supervision is considered to be both important and effective, practice is highly variable (Grant et al. 2003). This would not necessarily give cause for concern except that there are clear indications that there is inadequate coverage and frequency of supervision activities (although supervision is considered to be effective), together with significant differences in the perceptions of SpRs and ESs particularly in relation to monitoring performance, feedback, planning learning and support of the trainee. At the least this indicates there is a need for more explicit guidance for ESs and SpRs.

Purposes of supervision

Respondents were asked about educative, managerial and supportive functions of supervision because these three functions are frequently identified as the purpose of supervision in much health, social care and education literature. Activities reflecting each of these functions were considered to be of significant importance to the purpose of supervision in medical education (see Box 2).

Box 2: Supervision activities rated as of significant importance

- (1) Ensuring patient safety/care.
- (2) Educating the trainee.
- (3) Promoting high standards.
- (4) Identifying trainee problems.
- (5) Supporting the trainee.
- (6) Monitoring trainee progress.

Respondents were asked to rate each suggested purpose on a five-point scale (where 1 was not important); each purpose was rated as at least important (3 or more on the scale) by all respondents but where there were significant differences in the ratings SpRs placed more emphasis on educative functions of supervision whilst the educational supervisors prioritized managerial and supportive functions. This difference probably reflects different priorities and concerns of trainers and trainees in an environment where there can be a tension between service and education.

Organization of supervision

In the UK, approximately 90% of SpRs reported having a named supervisor, a similar number to those reported in other studies (for example, Bools and Cottrell 1994; Davies et al. 2000).

Traditionally, there has been an expectation that all consultants should be supervisors. However, in our study,

the majority of ESs and MDs considered that there was a difference between an educational supervisor and a consultant to whom the trainee is answerable although only slightly less than half (47%) of SpRs recognized this difference. However, apart from general practice and psychiatry, it is clear that this practice is not systematic. It varies between departments, hospitals and specialities. Furthermore, almost all the respondents indicated that only those consultants with an interest and commitment to supervision should be supervisors, rather than all consultants. Most SpRs would like to be able to choose supervisors, although they only rated this issue as of 'some importance' in factors that support good supervision. These views probably reflect a change in perceptions regarding roles of supervising consultants, which may be a result of the relatively recent changes in UK specialist training.

Although four out of five SpRs report that they have regular supervision meetings there is a wide range in the length (10–240 minutes) and the frequency (daily–six monthly) of these meetings. GP trainees and psychiatry SpRs have a mandatory requirement for weekly meetings/supervision meetings. In those specialities (anaesthesia, laboratory science, medicine, paediatrics and surgery) where there is no such requirement meetings are shorter and less frequent. Again, practice is highly variable, as has been reported elsewhere (Davies & Campbell 1995; Panayiatou & Fotherby 1996).

There were also consistent differences between ESs and SpRs in ratings in relation to the frequency of supervision, those activities that are supervised and the effectiveness of this supervision. SpRs reported lower frequency and effectiveness of supervision. It is not simply that SpRs consistently under-report all meetings—they reported receiving more frequent tutorials than the ESs reported giving. Also both groups rated supervisor and trainee availability as good (although there was a significant difference on ES and SpR ratings of ES availability). Both considered ESs to have good approachability. Therefore, the difference in SpR and ES perceptions are probably not due to availability or approachability of supervisors although availability was reported as a problem in the critical incident study. The reasons for this disparity are not clear; it may be that activities ESs recognize as supervision are not recognized as such by SpRs. Other studies have reported trainee dissatisfaction with supervision but most concentrate on trainee perceptions and/or experiences rather than comparing trainee and supervisor perceptions.

Supervision practices

The questionnaire data relating to supervision activities (see Box 3) give some cause for concern. None of the activities, including ensuring patient safety, was rated as receiving significant or full coverage either by SpRs or ESs. In other words, none of the activities was rated as occurring to a sufficient extent or with sufficient frequency. Almost all the activities showed a significant difference between SpRs' and ESs' ratings. ESs thought there was more coverage than did SpRs. Some of the largest differences occurred on items dealing with monitoring performance, feedback, planning

learning and support of the trainee. These activities might be seen as particularly important with regard to trainee development. Although this difference between ES and SpR perceptions is not explained in our findings, the most important aspect is that neither group rated any supervision activity as receiving significant or full coverage.

Box 3: Supervision activities (shown in decreasing order of extent and frequency of occurrence)

- (1) Discuss individual patients.
- (2) Ensure patient safety.
- (3) Provide informal feedback.
- (4) Monitor the trainee's performance.
- (5) Discuss (away from the bedside) the management of specific disorders.
- (6) Ensure that the trainee has an appropriate level and amount of clinical duties.
- (7) Provide feedback through appraisal.
- (8) Give advice relevant to personal and professional development.
- (9) Give support relevant to personal and professional development.
- (10) Address successes/problems in trainee performance.
- (11) Give career development advice.
- (12) Develop teamwork skills.
- (13) Ensure the safety of the trainee.
- (14) Discuss/review the process of supervision.
- (15) Teach specific techniques and procedures.
- (16) Plan the trainee's learning.
- (17) Develop interpersonal skills.
- (18) Develop communication skills.
- (19) Develop presentation skills.
- (20) Bedside teaching.
- (21) Use videotaped consultations.

SpRs, ESs and MDs all considered that supervision activities were at least moderately effective. Again, where there was a significant difference in perceptions of effectiveness, then SpRs rated the activity as less effective. ES reported giving significantly more feedback than SpRs reported receiving. ESs also considered this feedback to be more effective than did SpRs. There is considerable scope here for training courses aimed at creating more congruence concerning feedback.

Although both SpRs and ESs considered supervision during specific procedures/tasks (for example, outpatient clinics, ward rounds, tutorials and informal supervision) to be important, it occurred infrequently. Evidence from our literature review demonstrated the importance of supervision in relation to patient care and that direct supervision is effective but is often insufficient. The critical incident study had similar findings. Whilst quantity does not necessarily equate directly with quality, these data do suggest that existing supervision is insufficient. It is clear that SpRs think they need more feedback and direct supervision than they report receiving.

S. Kilminster et al.

Good supervision

Generally, we found consensus regarding good supervision between SpRs, ESs and MDs. The attitudes and commitment of supervisor and trainee, the relationship between them, protected time, importance of positive feedback and regular meetings were rated as of significant importance in supporting good supervision and this is supported by the literature. Although there appears to be general agreement on what good supervision is, other findings indicate that it is not always practised. Finding time for supervision is clearly important but there would appear to be other factors involved.

SpRs rated the relationship between supervisor and trainee as of significant importance. SpRs also rated the need for guidelines, a definition of supervision and choice of supervisor higher than did ESs. These items all relate to control of the supervisory process and relationship and suggest that SpRs want more control over this. In the remainder of the survey SpRs consistently made lower ratings than ESs. There were only three speciality-specific significant differences in views concerning factors supporting good supervision. General practice gave highest ratings to the importance of the supervisor's teaching skills and the need for training, assessment and monitoring whilst psychiatry gave highest rating to 'trainee having regular meetings with the supervisor'. These ratings are noteworthy because training is mandatory for GP supervisors and supervision meetings for psychiatry trainees. The lowest rankings for all aspects of supervision were in medicine, where there was also least supervision.

Difficulties in supervision

Respondents were concerned about lack of supervision for emergency and 'out of hours work', failure to formally address underperformance, lack of commitment to supervision and finding sufficient time for supervision. These problems have serious implications in the context of clinical governance and audit. There is a need for an effective system to address both poor performance and inadequate supervision. Where there were significant differences, across specialities, in ratings of difficulties it is interesting to note that time, supervisor availability and lack of training of the supervisor caused the greatest difficulty in anaesthesia and medicine, and least difficulty in general practice where supervisors have to be trained and weekly meetings are mandatory. The large numbers of trainees in anaesthesia are perceived to be causing problems although it is not clear why this should be so. Where respondents gave figures there did not appear to be a severe imbalance between numbers of trainees and numbers of consultants. It might be expected there would be a similar problem in surgery but this was not apparent.

A framework for effective supervision

In this section we suggest a framework for effective supervision, which is based on our research findings and the literature. This framework must be understood as located in

the external framework for training and the guidance provided on necessary training experiences by bodies responsible for postgraduate training. Training is a partnership between supervisor and supervisee and requires the active involvement of both—it is not something that trainers 'do' to trainees. Within this partnership trainers and trainees both have obligations and responsibilities.

Early planning meetings, agreement about learning objectives, written contracts and review of trainee placements and progress by the programme director are an essential component of well-run training programmes and will prevent many problems arising. The differences between specialities in their ratings regarding difficulties in supervision and factors supporting good supervision suggest that having minimum requirements for supervision and training of supervisors reduces problems and promotes good supervision.

Our work has demonstrated that there is a need for a clear definition of supervision (which we have provided) and guidelines concerning supervision. In the following paragraphs we identify the features and mechanics of effective supervision.

Features of effective supervision

- (1) Direct supervision—trainee and supervisor working together and observing each other—positively affects patient outcome and trainee development.
- (2) Constructive feedback is essential and should be frequent.
- (3) Supervision should be structured and there should be regular timetabled meetings. The content of supervision meetings should be agreed and learning objectives determined at the beginning of the supervisory relationship. Supervision contracts can be useful tools and should include details of frequency, duration and content of supervision; appraisal and assessment; learning objectives; and any specific requirements.
- (4) Supervision should include clinical management; teaching and research; management and administration; pastoral care; interpersonal skills; personal development; reflection.
- (5) The supervision process should be informed by a '360 degree perspective'. This includes patient feedback, inter-professional supervision and training as well as reviewing written work and records. This will be supplemented by formal processes such as appraisal meetings and the results of examinations and formal assessments.

The quality of the supervisory relationship strongly affects the effectiveness of supervision. Specific aspects include continuity over time in the supervisory relationship, that the trainees control the content of supervision (there is some suggestion that supervision is only effective when this is the case) and that there is some reflection by both participants. The relationship is partly influenced by the supervisor's commitment to teaching as well as both the attitudes and the commitment of

supervisor and trainee. Supervisors (and trainees) need to understand that:

- (1) *Helpful supervisory behaviours* include giving direct guidance on clinical work, linking theory and practice, engaging in joint problem-solving and offering feedback, reassurance and providing role models.
- (2) *Ineffective supervisory behaviours* include rigidity; low empathy; failure to offer support; failure to follow supervisees' concerns; not teaching; being indirect and intolerant and emphasizing evaluation and negative aspects.
- (3) In addition to supervisory skills, effective supervisors need to have good interpersonal skills, good teaching skills and be clinically competent and knowledgeable.
- (4) *Training*. There is agreement that training for supervisors probably needs to include at least some of the following: understanding teaching; assessment; counselling skills; appraisal; feedback; careers advice; interpersonal skills.

In addition, our research evidence indicates that there are two areas of particular concern:

- It might be expected that there would be particular difficulties in supervision with regard to time, availability and approachability of supervisors. However, these issues appear to present fewer problems in specialities where there is a formal requirement for weekly supervision meetings than those where there is no such requirement. There is therefore a strong argument that all specialities should have a formal requirement specifying the frequency of supervision meetings.
- There are problems with 'out of hours' supervision and formally addressing underperformance and inadequate supervision. There is an urgent need for effective systems to resolve these issues.

Continuity of supervision

Continuity is a vital element in effective supervision of training and the delivery of a safe and effective service. It is essential for trainees who rotate through different placements. Establishing the level of competence of the trainee (for example, by direct observation or enquiry of others) is an essential first step in supervision otherwise training cannot commence and judgements cannot be made concerning the closeness of supervision needed to ensure patient safety.

Continuity of supervision needs to start early. Ideally, foundation trainees should have a portfolio documenting their strengths, weaknesses and achievements as an undergraduate. Training schemes need to organize themselves in such a way as to ensure accurate information about trainees is communicated effectively to supervisors as trainees rotate from post to post. Regardless of whether undergraduate or previous training information is available or not, an early meeting with the trainee (within the first two weeks of starting the post) needs to take place. At this meeting the structure and ground rules need to be agreed. These should indicate agreement on time and place of future meetings, issues of

confidentiality and accessibility of supervisor outside normal meetings.

A suggested format for this early meeting is shown below:

- Review progress to date (and any hand-over information).
- Review together speciality training guidelines.
- Formulate/review educational/training contract with timescales.
- Identify methods of achieving objectives or goals (the subsequent meeting should be used to review progress).

At the final meeting at the end of any training placement, an overall review should be undertaken to ensure that the trainee is able to progress to the next level and to identify in which area training should now take place. This information needs to be communicated to trainers in the next placement.

The issue of who should provide continuity of supervision is difficult and different specialities may adopt different solutions. For training to occur in a planned and coherent way, supervision of a trainee is best overseen by a single individual who will be involved with the trainee for a significant amount of time. Additionally, if problems are identified, they are more likely to be addressed by a supervisor who has responsibility for the trainee over, say, two years, than by someone who only sees the trainee for six months.

In some disciplines a programme director or the post-graduate tutor may be best placed to provide this overseeing role. In others where there are large numbers of trainees, this may be logistically impossible and here an individual consultant may take responsibility for a trainee throughout his/her time in the training programme. Irrespective of who takes on this role, it is essential that trainees and trainers are aware of the roles and responsibilities of the various people involved in providing training and who has ultimate responsibility for the trainee's progress.

Trainees

Trainees should be familiar with the overall training objectives for their chosen speciality and the agreed objectives for any particular placement. They should keep a record of their training experiences and achievements in relation to agreed objectives that can be used to inform discussions on future training. Different specialities require different recording procedures but, increasingly, trainees are being encouraged to keep detailed learning portfolios.

Trainees should attend supervision meetings punctually and should have prepared for any agreed tasks.

Supervision sessions should be trainee led, with trainees taking responsibility for their learning by suggesting topics for discussion. This does not preclude consultant supervisors from also suggesting topics.

Trainees must be prepared to develop a capacity for self-awareness and reflection on their practice that will enable them to identify, and bring forward for discussion in supervision, any areas where they feel their performance needs improving. They also need to be able to constructively criticize

S. Kilminster et al.

local services where service organization issues interfere with delivery of training.

Trainee needs

Trainees also have certain specific needs in relation to supervision:

- graduated responsibility over time with direct supervision of assessment and management of patients, prescribing, practical procedures, administrative duties etc. shifting to less direct supervision over time;
- regular constructive feedback;
- establishment of learning objectives at the outset of each placement and identification of strategies for achieving them;
- periodic assessment and appraisal;
- time to attend specialist courses and specific instruction for examinations;
- supervision of their teaching/supervision of others;
- development of management, audit skills and involvement in the processes required by clinical governance;
- pastoral care and the provision of appropriate role models;
- provision of a safe environment for training and clinical work;
- career planning and advice based on the best workforce data available.

Supervisors

Supervisors must contribute to the provision of a well-organized and comprehensive training programme and ensure that trainees placed with them have the necessary opportunities to achieve agreed objectives.

Supervisors must be accessible and should arrange regular uninterrupted meetings with trainees for supervision as well as being clear about how and under what circumstances they can be contacted between meetings.

Supervisors must observe their trainees in practice and make arrangements to gather information from others who have observed the trainee.

Supervisors must provide a safe environment in which trainees feel able to discuss their own perceived deficiencies and empowered to make any relevant constructive criticism of their training, including the supervision process. Trainees should see that action has been taken on problems they have identified. Supervisors need to cultivate an atmosphere of openness throughout the departments for which they are responsible.

Supervisors must be able to provide honest, fair and constructive feedback on trainee performance at regular intervals (see Box 4).

Box 4: How to give constructive feedback

Constructive feedback aims to improve performance. It should identify and reinforce the strengths of a person's performance and identify the weaknesses *whilst suggesting ways to improve them*. Feedback is most effective when it is *timely*—close to the event.

Giving feedback.

Ask the trainee to comment first and to identify which aspects of his/her performance went well. Then let him/her identify areas of difficulty and possibilities for change or development.

Respond to his/her comments before offering your own comments.

Again, begin with the positive. Be specific and descriptive, for example, 'The way you analysed the patient's problems and arranged appropriate investigations was excellent' rather than simply saying 'Very good'.

Prioritize and do not give a lot of negative feedback in one big bundle. Refer to behaviour that can be changed; for example, 'I know you are nervous but you will make the patient more comfortable if you make eye contact while you are talking to him'. Offer alternatives—try not simply to criticize but offer an alternative way of doing it. 'I think the patient was uncooperative partly because you did not explain what you were going to do. Try explaining the procedure now and then go back and tell her in simpler terms.'

Agree the next steps.

Feedback should be regular but can be brief and still very effective.

Feedback should be given as close to the event as possible.

Skills required of a good supervisor

Supervision for junior staff must be offered in a supportive environment whilst ensuring patient safety. The skills required to deliver this supervision are many and varied (Box 5).

All training placements should start with a detailed 'educational needs assessment' and identification of clear learning objectives for the placement. This requires *appraisal skills* and the ability to establish the level of competence of trainees through observation of performance. Of particular importance is the ability to recognize unsatisfactory performance and progress and the willingness to act appropriately in the interests of the trainee and the patient. Supervisors need the ability to *observe and reflect* on practice and to provide trainees with clear and constructive *feedback* on their performance (see Box 4).

Effective supervisors need *formal skills in teaching* and facilitating learning. They need to be able to plan and organize teaching sessions, formulate relevant and achievable learning objectives, and facilitate trainee involvement in the learning process. In supervision sessions, helping the trainee to develop his/her own solutions requires the supervisor to have skills in *identifying alternatives* and *problem-solving*. Supervisors will also, at times, need the ability to *motivate* trainees.

Managing the tension between facilitating self-directed learning and directing the learning of the trainee is not easy. It may feel safest to monitor the trainee closely and this may be very appropriate in the early stages of training but supervisors need to be able to progressively advance the ability of the trainee to work independently without compromising patient safety and thus need skills in *fostering autonomy*.

The supervisor must be a skilled *information provider* able to understand and transmit the training and legal requirements of relevant statutory bodies (for example, medical Royal Colleges and the General Medical Council). The supervisor must also communicate the policy and procedures of the local department/unit and Trust and in turn ensure that other team members are aware of the training requirements and responsibilities of the trainee.

Supervision takes place in a context and the supervisor usually has a key part to play in creating the best possible environment for training. This requires good *service management skills* to ensure that department/unit affairs are well organized and run smoothly and that all staff are clear about their roles and responsibilities. Role modelling good clinical practice, leadership, teamwork and open communication, and critical self-evaluation of performance within the clinical service are essential components of good supervision.

Creating a context for delivering effective clinical services includes ensuring an appropriate balance between service and educational activities, constructing timetables and rotas that are coherent with training requirements, and seeking funds to provide the necessary physical facilities and materials for training. It also requires the *ability to foster a supportive culture* that promotes the personal and professional development of staff.

The supervisor may have to be an *advocate* for the trainee, to ensure he/she has adequate resources for training and that his/her training needs are being met. This will on occasion require *negotiation* skills. Ensuring that there is time for supervision whilst meeting clinical service needs requires *time management and organizational skills*.

Finally, the supervisor needs *self-appraisal skills* and the willingness to reflect on his or her own personal supervisory style and initiate change if it is not shown to be supportive and enabling.

Box 5: Effective supervisors are able to

- Observe and reflect on practice
- Give constructive feedback (see Box 4)
- Teach
- Identify alternatives
- Problem-solve
- Motivate
- Foster autonomy
- Provide information
- Appraise self and others
- Manage a service
- Create a supportive climate
- Advocate
- Negotiate
- Manage time
- Organize

Dealing with problems in supervision

There are many reasons why supervision may not be effective; these include:

- poorly organized training programmes;
- trainers who have poor supervisory skills;
- tension between service delivery and supervision/training needs;
- whether the trainee is able to learn from experience and to manage errors;
- whether trainees feel confident enough to acknowledge/address difficulties.

Many problems can be resolved through effective organization of training and appropriate mechanisms for appraisal and feedback. However, at the heart of supervision is the relationship between trainer and trainee and considerable difficulties can ensue if there are problems in this relationship.

Hierarchy and power

The innate hierarchy and power in a supervisor trainee relationship may be used as a positive or destructive force on either side, although the potential for abuse is probably greater on the supervisor's side.

Working closely together over a period of time can produce a feeling of mutual trust between the supervisor and trainee and a much greater understanding of the problems encountered by both parties. Obstacles to training that are identified can form the basis of supervisory sessions where the supervisor can help the trainee to arrive at his/her own solutions. However, if it is not possible the supervisor can step in when required. This might happen if problems of service work override educational needs—a trainee may be able to address this by making minor adjustments in timetabling but, for example, a consultant intervention may be required to prevent trainees being asked to do extra clinics for other consultants. The trainee's difficulties with other health professionals can be highlighted and might be dealt with by consultant intervention.

13

S. Kilminster et al.

However, the relationship is open to abuse, particularly as the trainee may feel in a subservient position, often dependent on the supervisor to progress to the next level of training and for job references. In primary care, the trainee is an employee of the practice and of the trainer. The supervisor also has the power to manipulate the trainee's timetable to ensure that service—not training—needs are met. Consequently the trainee may not feel able to reveal clinical weakness or emotional/psychological problems. If these problems are revealed, the supervisor may constantly focus on these problems and not the solutions, gradually undermining the trainee's confidence.

Dealing with personality issues in supervision

Personality issues may arise in supervision in a number of ways. A 'personality clash' between trainer and trainee may impede effective supervision; some personality issues are almost inevitable within supervision at some stage, even in the best run training schemes. More seriously, supervisors may become concerned about trainee attitudes to patients and to the other staff in the healthcare team. In such cases there may not necessarily be any particular problems in the trainer/trainee relationship, highlighting the importance of canvassing the opinions of other members of the team.

Possible solutions

The process of supervision is a finely balanced one and abuse of the system on either side may well tip the relationship into a potentially destructive one. All training programmes should have clear guidance regarding the conduct of supervision and well-publicized systems in place to address difficulties. Guidance on appropriate conduct should exclude teaching by 'humiliation', bullying, sexual harassment, and relationships between trainers and trainees. Trainees should know whom to contact if problems arise that cannot be resolved within the placement. Trainees should be discouraged from receiving medical treatment from trainers, for example GP trainees registering with their training practice.

Ideally problems should be discussed with the supervisor, as part of the regular process of reflecting on supervision within supervision sessions. Trainees need clear feedback and constructive suggestions on action. These can be related to the speciality learning objectives and also to other relevant publications such as the GMC's *Good medical practice* (1998). If problems cannot be resolved within supervision, there should be clear mechanisms for trainer and trainee to involve a third party as a mediator to help resolve issues. Programme Directors, Postgraduate Tutors or Postgraduate Deans and their nominees are most likely to be involved in this in the UK. A well-defined process of appeal should be identified if all else fails.

When there is concern about 'personality issues' it is important to ensure that trainers and/or trainees are not suffering from treatable physical or psychiatric disorders, or experiencing adverse life events. Careful assessment of

the situation and information regarding past progress and problems will be helpful here.

Many trainers are reluctant to raise concerns about attitudinal problems with trainees, as they can be difficult to resolve. However, the advent of clinical governance and recent advice from the GMC place an obligation on trainers to report such issues if they cannot resolve them. If problems cannot be reconciled, then clearly defined sanctions need to be in place to either prevent the progress of the trainee to the next level or allow for the removal of the trainee from a particular supervisor or trainer. As a last resort, local 'three wise people' procedures can be involved or the national professional regulatory body may need to be contacted if there are serious, unresolved concerns about a trainee's attitude to patients.

Supervision at different levels and in different specialities

Supervision at different levels

It is clear from our definition that all clinical staff should receive supervision, irrespective of grade. This should apply to consultants, principals in general practice and non-training grade doctors as much as to doctors in training. It is illogical that the process of reflection on and coordination of learning, which now takes place for all junior staff, should cease on leaving the training grades. All staff should participate in a programme of continuing professional development and ensure that they are up to date with new procedures, practices and knowledge.

Staff at all levels are likely to be receiving supervision and at the same time supervising others. Even foundation trainees will be 'supervising' medical students.

Trainees need to acquire responsibility in a graded fashion as they achieve competences, with the aim of becoming independent practitioners. The amount of direct clinical supervision required will be maximal at the foundation trainee level and at the beginning of each grade as new and unfamiliar problems are encountered and will decrease with time and experience. Paradoxically, much of the 'supervised' work of more senior trainees such as SpRs will take place without direct supervision. The process of development into an independent specialist requires that as experience is gained so the trainees are able to take more and more responsibility themselves. Clinical decisions are therefore reported to supervisors after the event or may not be reported at all if they form part of the daily currency of the work of a senior trainee. Middle- and senior-grade trainees will also be supervising others as well as receiving supervision themselves although, ultimately, responsibility will lie with consultant supervisors. Thus the capacity to supervise is also an essential part of the training process.

The content of what needs to be supervised at different levels will change but the 'closeness' of supervision will vary according to the grade and amount of progress within the grade. Trainers need to make judgements regarding levels of supervision (See section on 'Ensuring trainee

competence and levels of supervision') as to whether they should:

- be present in same room as the person being supervised;
- be nearby and immediately available to come to the aid of the person being supervised;
- be in the hospital or primary care premises and available at short notice, able to offer immediate help by telephone and able to come to the aid of the person within a short time;
- on call and available for advice, able to come to the trainee's assistance in an appropriate length of time.

Setting out the supervision needs of trainees at each of the different training grades is counterproductive as so much varies according to speciality. Although the content of supervision will vary according to grade, the basic structure of supervision needs is broadly similar at each level. Similarly, the generic skills required of the supervisor remain the same at each level (see section on 'Skills required of a good supervisor'). The personal contribution of the consultant will vary with the amount of supervision also available from intermediate grades; for example, the consultant will be the only person supervising an SpR, whereas a foundation trainee will receive supervision from SHOs and SpRs and other members of the healthcare team as well as the consultant. Where supervision is less direct, as in the situation where a SpR may be providing direct supervision of a SHO or foundation trainee, consultants must set up systems requiring the SpR to report to the consultant on trainee progress of an SHO, staff grade or foundation trainee. This in turn provides valuable supervision opportunities for discussion of and reflection on the SpR's role as a trainer and supervisor.

The supervision of trainees in general practice needs to acknowledge the change from hospital to primary care. The transfer from the confines of hospital work to the open-ended environment of primary care is a culture shock not to be underestimated. The new trainee will need time and space to adjust to the new environment. The registrar must be able to work within his/her competence. After the initial orientation, she/he will be learning new skills, not least in the realm of clinical assessment, consultation skills and living with uncertainty. It is the trainer's job to monitor closely and teach the new skills and attitudes required slowly over the first weeks and months as there will be a gradual increase in responsibility and clinical load. The 'sink or swim' approach is to be strongly deprecated. The paramount aim of supervision is patient safety, now and in the future. Formative assessment, regular tutorials and an educational culture that allows sharing of both knowledge and ignorance is essential.

'Supervision' for consultants, principals and staff grades

The principle of 'partnership' is of paramount importance for consultants and principals in general practice where

individuals may enter into arrangements for peer consultation/supervision of work with colleagues as part of a programme of continuing professional development.

It is important that staff grades should not be exploited in the name of supporting the training grades. The needs of these staff with regard to supervision are similar if not identical to the needs of those in the training grades, albeit that consultants and principals are likely to be receiving 'supervision' from peers. However, the lack of a formal structure to monitor training and supervision has led to many difficulties in ensuring that consultants, principals and staff grades continue to benefit from education and supervision. In the future, the advent of appraisal, revalidation, personal learning portfolios and clinical governance should ensure that this state of affairs does not continue.

Supervision across the specialities

Although the mechanics of supervision vary across the specialities there is a generic structure and skills in all supervision. Here we give examples of supervisory practice from disparate specialities and it will be evident that they have general applicability to supervision issues in other specialities. The examples are taken from case studies, written by experienced supervisors, regarding their personal experiences of supervision. The speciality from which the vignette is taken is indicated in each box (Boxes 6-13).

S. Kilminster et al.

Box 6: Assessing trainee competence

Although surgical experience is carefully documented in log books, trainers are *not* at present required to sign off the competence of individuals in particular procedures. This gives rise to several problems:

- There is a delay when changing training paths while competence and training requirements are assessed by the new trainer.
- It makes it difficult for the new trainer to formulate and agree a training plan with the trainee.
- It makes it difficult for other supervisors such as the programme director at appraisal (RITA) to monitor progress of the trainee through the training scheme and remedy any deficits.
- It makes it difficult to defend assessment decisions, particularly if the trainee is deemed to be not competent.

Consequently the Vascular Surgical Society of Great Britain and Ireland instituted a simple chart that itemizes the key or index vascular procedures essential for subspecialty training. Included in this chart is open space to record the training received in these specialist procedures and *the level of competence* acquired. This latter is 'signed off' by the trainers and so creates a permanent progressive record of achievement. We adapted this form for the year 1–3 trainees on a local higher surgical training scheme. It now forms part of the RITA for these trainees and is used as part of the training agreement between SpRs and their individual trainers. The information is also gathered and analysed at regional level, forming a valuable source of data on the efficiency of operative training.

*Surgery***Box 7: Supervision practices**

I have managed to divide my outpatient work into new and review clinics. This means those review clinics can be run in a more meaningful way. As far as possible, doctors follow up their own patients. This gives continuity for patient and doctor.

In the afternoon after the morning clinic, all the doctors meet and each presents (consultant included) the patients they have seen and discusses the difficulties and their management plan. This is the time for any doctor to ask for advice about a particular patient. It works extremely well, junior doctors feel supported, patients can be confident that the consultant is still overseeing their case and patients are not subjected to endless, non-productive follow up. Areas of lack of knowledge can be highlighted and addressed. At the end of the post, both SpRs and SHOs have spontaneously expressed enthusiasm for this—regardless of their seniority.

*Medicine***Box 8: Continuity of supervision**

There is a five-year training programme for the subspecialty of geriatric medicine involving a series of clinical placements and experience. In one local area supervisors decided that trainees should spend at least two years in one hospital site. We think that the advantage is that trainees have increased experience in one unit and a greater chance of longitudinal follow up of patients thereby enhancing their experience of disease progression. In addition, the trainees are more secure in their geographical placement with less disruption to their personal life. The advantage to the hospital is fewer changes of personnel. The trainee is associated with one supervisor for a longer period of time and thereby they get to know each other better and develop a deeper professional association.

*Geriatric Medicine***Box 9: Useful supervision techniques**

Although various relatively objective and recordable systems of supervision for procedures (e.g. observe, assist at etc.) have been developed, it is more difficult to make an objective assessment of the development of trainee doctors' diagnostic, consulting and medical management skills. A number of techniques are used in general practice to identify whether the trainee's work is developing satisfactorily and that the trainee's management of patients is of an appropriately high standard:

- random case note analysis;
- analysis of consultation on video;
- critical event analysis (events such as deaths or perceived clinical errors are analysed to see if anything might have been done better);
- analysis of prescription rates;
- analysis of investigation rates;
- analysis of hospital referral patterns, referral letters and replies;
- analysis of complaints.

General practice

Box 10: Levels of supervision

The consequences of poor supervision in anaesthesia can be very serious indeed and there is a stringent requirement for all trainees to receive appropriate clinical supervision at all stages in their training in anaesthesia. It is recommended that full-time, direct supervision should be provided at all times during the first 12 weeks of training. If the trainee does take part in the on-call rota then the supervision will probably be provided by other trainees who are further advanced in their training.

The level of clinical supervision is determined by the previous experience/training of the anaesthetist being supervised and the specific clinical situation.

Anaesthetists also have major involvement in intensive care units and pain management clinics. Clinical supervision for anaesthetists in intensive care should follow a similar pattern to that described for anaesthesia. There is normally much less urgency about clinical situations in a chronic pain clinic but high levels of supervision are usually warranted during both consultations and treatment sessions.

At varying times during their training anaesthetists require enhanced supervision. This may be whilst a single procedure is being performed or during introduction to a new sub-speciality.

Anaesthetics

Box 11: Continuity in supervision

Although each trainee had a supervisor there were some problems with continuity as well as personality clashes between trainee and supervisor so a mentor system was instituted in one region. The mentor is a consultant in A&E in another department in the region. She/he meets regularly with the trainee and reviews their progress in the light of their own assessment and feedback from the operational level. Any problems identified are then addressed appropriately. Participants have found that the process enables the strategic education plan to develop appropriately over time, even when the trainee moves hospital. In addition a more balanced assessment can be made during the bi-annual strategic meeting with the trainee.

Accident and Emergency

Box 12: Complexities of supervision in practice

The operating theatre can be a hostile environment for trainees. The trainee has to contend not only with the supervisor/teacher and the process of learning but also with the stresses of administering anaesthesia, the demands of the surgeon, time pressures, cost pressures, the presence of other staff such as nurses and, last but not least, patient expectations. Supervision of a trainee during an operating list may be subject to many interruptions and frequent inability to complete episodes of teaching.

Anaesthetics

Box 13: Problems in supervision

A trainee was enthusiastic about a career in front-line acute paediatrics. Early reports from both nurses and junior members of the department caused concern about the trainee's competence because of panic decision-making, indecisive leadership, failure in delegation of tasks and signs of stress. This led the educational supervisor to sit down with the trainee to list the skills necessary for acute intensive clinical work. But there was no evidence that these skills were improving at repeated reviews. The supervisor helped identify the trainee's areas of strength and identified a career pathway in which the trainee was more likely to succeed. This approach, emphasizing strengths not weaknesses, was successful. The trainee took the career advice enthusiastically, and with relief as she/he did have insight into his/her problems.

A second trainee lacked insight into his/her own difficulties with interpersonal relationships. She/he was brilliant in some areas of basic science and clinical medicine but was not a 'team player'. The educational supervisor arranged regular meetings and offered opportunities for skills development. However, these opportunities were poorly attended and relevant questioning at the trainee's appraisal meeting indicated that she/he had a lack of awareness and understanding of the difficulties. Progress to the next part of the training programme was deferred and the trainee protested. This situation was very difficult to manage and was referred to the Postgraduate Dean who supported the decision of the appraisal panel—that it was very unlikely the trainee would achieve a successful appraisal in the future. The trainee left the training programme to work in research.

Paediatrics

These examples have been chosen to reflect some of issues that cause difficulties in supervision and to show how they have been addressed in different specialities. They illustrate the importance of structure, continuity, supervision techniques, direct supervision, complexities of supervision in practice and dealing with problems in supervision.

Conclusion

The content of this guide is informed by both empirical work and practitioners' experiences. We have identified the need for a definition of and for explicit guidelines on supervision. There is strong evidence that, whilst supervision is considered to be both important and effective, practice is highly variable. In some cases, there is inadequate coverage and frequency of supervision activities. There is particular concern about lack of supervision for emergency and 'out of hours work', failure to formally address under-performance, lack of commitment to supervision and finding sufficient time for supervision. There is a need for an effective system to address both poor performance and inadequate supervision. We have offered both a definition and a framework for effective supervision that is intended to be of practical use to practitioners.

17

Acknowledgements

The authors wish to thank the following people who contributed to this guide through their participation in a two-day development workshop and/or by writing about their experiences of supervision: Sushma Acquilla; Janet Anderson; Dee Dawkins; Peter Driscoll; Vin Diwaker; Gordon Jackson; Doug Justins; Richard Morgan; David Newble; Geoff Norris; Trudie Roberts; Denis Wilkins; Simon Bennett; Robin Cairncross; Claire Waring; John Wilkinson.

Notes on contributors

SUE KILMINSTER is Senior Research Officer in the Medical Education Unit, University of Leeds, UK.

DAVID COTTRELL is Professor of Child & Adolescent Psychiatry at the University of Leeds, UK. He remains an active clinician and is Associate Medical Director in his local NHS Trust.

JANET GRANT is Professor of Education in Medicine at the Open University Institute of Educational Technology and Director of the Centre for Education in Medicine.

BRIAN JOLLY is Professor of Medical Education and Director, Centre for Medical and Health Sciences Research, University of Monash, Australia.

References

- Arluke A. 1980. Roundsmanship: inherent control on a medical teaching ward. *Soc Sci Med* 14A:297–302.
- Boud D, Cohen R, Walker D. 1993. *Using Experience for Learning* (Buckingham, SRHE & Open University Press).
- Bools C, Cottrell D. 1994. Future child and adolescent psychiatrists: a further survey of senior registrar training. *Psychiatr Bull* 14:611–615.
- Cottrell D, Kilminster SM, Jolly B, Grant J. 2002. What is effective supervision and how does it happen? *Med Educ* 36:1042–1049.
- Davies BW, Campbell WB. 1995. Inguinal hernia repair: see one, do one, teach one? *Ann Roy Coll Surg Engl* 77:299–301.
- Davies JH, Tan K, Jenkins HR. 2000. The current status of senior house officer postgraduate education in a single region. *Med Educ* 34:367–370.
- Department of Health. 1996. *A Guide to Specialist Registrar Training* (London, Department of Health).
- Ende J, Pomerantz A, Erickson F. 1995. Preceptors' strategies for correcting residents in an ambulatory care medicine setting: a qualitative analysis. *Acad Med* 70:224–229.
- Eraut M. 1994. *Developing Professional Knowledge and Competence* (London, Falmer Press).
- Fallon Jr WF, Wears RL, Tepas III JJ. 1993. Resident supervision in the operating room: does this impact on outcome? *J Trauma* 35:556060 (discussion 560–561).
- General Medical Council. 1997. *The New Doctor* (London, GMC).
- General Medical Council. 1998. *Good Medical Practice* (London, GMC).
- General Medical Council. 1999. *The Early Years* (London, GMC).
- Genniss VM, Genniss MA. 1993. Supervision in the outpatient clinic: effects on teaching and patient care. *J Gen Intern Med* 8:378–380.
- Grant J, Kilminster SM, Jolly B, Cottrell D. 2003. Clinical supervision of SpR's. Where does it happen, when does it happen and is it effective? *Med Educ* 37:140–149.
- Griffiths CH, Desai NS, Wilson EA, Powell KJ, Rich EC. 1996. Housestaff experience, workload and test ordering in a neonatal intensive care unit. *Acad Med* 71:106–1108.
- Harden RM, Crosby JR. 2000. AMEE Education Guide no 20: the good teacher is more than a lecturer—the twelve roles of the teacher. *Med Teach* 22:334–347.
- ▶ Hesketh EA, Bagnall G, Buckley EG, Friedman M, Goodall E, Harden RM, Laidlaw JM, Leighton-Beck L, McKinlay P, Newton R, Oughton R. 2001. A framework for developing excellence as a clinical educator. *Med Educ* 35:555–564.
- Jolly BC, Macdonald MM. 1986. Practical experience in the pre-registration year in relation to undergraduate preparation. *Proc Annu Conf Res Med Educ* 25:171–176.
- Kilminster SM, Jolly BC. 2000. Effective supervision in clinical practice settings: a literature review. *Med Educ* 34:827–840.
- Kilminster SM, Jolly B, Grant J, Cottrell D. 2000. *Good Supervision: Guiding the Clinical Educator of the 21st Century* (Sheffield, University of Sheffield).
- ▶ Kilminster SM, Jolly B, van der Vleuten C. 2002. A framework for training effective supervisors. *Med Teach* 24:385–389.
- Kolb DA. 1984. *Experiential Learning Experience as the Source of Learning and Development* (Englewood Cliffs, NJ, Prentice Hall).
- Lave J, Wenger E. 1991. *Situated Learning: Legitimate Peripheral Participation* (Cambridge, Cambridge University Press).
- McKee M, Black N. 1992. Does the current use of junior doctors in the United Kingdom affect the quality of medical care. *Soc Sci Med* 34:549–558.
- Panayiatou BN, Fotherby MD. 1996. Junior hospital doctors' views on their training in the UK. *Postgrad Med J* 72:547–550.
- Patel V, Groen P. 1990. The general and specific nature of medical expertise: a critical look, in: KA. Ericsson & J. Smith (Eds), *Towards a General Theory of Expertise: Prospects and Limits* (Cambridge, Cambridge University Press).
- Schon DA. 1995. *The Reflective Practitioner: How Professionals Think in Action* (Aldershot, Arena).
- Somers PS, Muller JH, Saba GW, Draisin JA, Shore WA. 1994. Reflections-on-action: medical students' accounts of their implicit beliefs and strategies in the context of one-to-one clinical teaching. *Acad Med* 69:584–587.
- ▶ Sox CM, Burstin HR, Orav EJ, Conn A, Setnick G, Rucker DW, Dasse P, Brennan TA. 1998. The effect of supervision of residents on quality of care in five university-affiliated emergency departments. *Acad Med* 73:776–782.
- Tenant M. 1999. Is learning transferable?, in: D. Boud & J. Garrick (Eds), *Understanding Learning at Work* (London, Routledge).
- Williams S, Dale J, Glucksman E, Wellesley A. 1997. Senior house officers' work related stressors, psychological distress and confidence in performing clinical tasks in accident and emergency: a questionnaire study. *BMJ* 314:713–718.

Appendix: the UK regulatory framework

Supervision and clinical governance

Clinical governance is defined in the 1998 White Paper, *A First Class Service*, as:

...a framework through which NHS organisations are accountable for continuously improving the quality of their services and safe guarding high standards of care by creating an environment in which excellence in clinical care will flourish.

The object of training is to provide the patients of the future with high-quality specialists who have had a wide range of useful and informative experience during their training years. Both the interests of the patients of today

and the quality of the training experience depend on good clinical and educational supervision of trainees during their training years.

The quality of clinical supervision of trainees is therefore a central problem for the clinical governance organizations within Trusts, and these organizations will need to assure themselves that appropriate supervision is being undertaken. Although the arrangements for the management of educational supervision have improved out of recognition throughout the UK over the last decade, it is still relatively unusual for Trusts to have identifiable management systems which are capable of assuring the clinical governance organization within the Trust that the level of clinical supervision of trainees is adequate to ensure the delivery of services of appropriate quality. However, appropriate supervision is central to the process of clinical governance and such management systems will need to be developed.

การสอนกลุ่มย่อย (Small group teaching)

รองศาสตราจารย์แพทย์หญิง พรพรรณ กุ้มานะชัย
สาขาวิชาโรคติดเชื้อและอายุรศาสตร์เขตร้อน
ภาควิชาอายุรศาสตร์
คณะแพทยศาสตร์ศิริราชพยาบาล

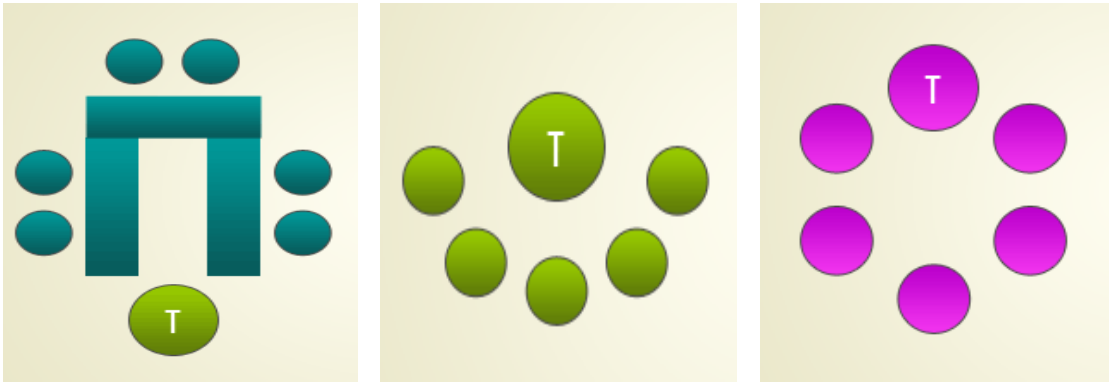
การสอนกลุ่มย่อยเป็นรูปแบบที่มีความสำคัญในแพทยศาสตรศึกษาเพราะการสอนกลุ่มย่อยมีข้อดีหรือข้อได้เปรียบ คือ สามารถกระตุ้นให้ผู้เรียนใช้กระบวนการคิดจากพื้นฐานความรู้ที่มีและนำมาอภิปรายร่วมกับผู้เรียนอื่นๆ โดยมีผู้สอนเป็นผู้คอยชี้แนะและให้ข้อมูลสำคัญต่างๆ เท่าที่จำเป็น การสอนแบบกลุ่มย่อยที่ได้ประสิทธิผลสูงสุดนั้นต้องเน้นให้ผู้เรียนเป็นศูนย์กลาง (learner-centered) และให้ความสำคัญแก่ผู้เรียนเป็นหลัก โดยขนาดของสถานที่หรือห้องเรียน หรือจำนวนของผู้เรียนมีความสำคัญรองลงมา

รูปแบบต่างๆของการสอนกลุ่มย่อย ได้แก่

- การอภิปรายสัมมนา
- การติว
- การเรียนโดยตั้งต้นจากปัญหา
- การอภิปรายจากตัวอย่างผู้ป่วย
- การเรียนทางห้องปฏิบัติการ การฝึกหัดทางห้องปฏิบัติการ
- การเรียนรู้จากการดูแลผู้ป่วยที่หอผู้ป่วยหรือผู้ป่วยนอก
- การประชุมกลุ่มปฏิบัติการย่อย

ลักษณะของการสอนกลุ่มย่อย

1. ขนาดของกลุ่มโดยทั่วไปประกอบไปด้วยผู้เรียนประมาณ 6-8 คน อย่างไรก็ตามจำนวนผู้เรียนอาจมีมากหรือน้อยกว่านี้ได้แตกต่างกันไปตามความเหมาะสมของกิจกรรม
2. ต้องมีการกำหนดจุดประสงค์หรือสิ่งที่จะเรียนร่วมกันให้จำเพาะและชัดเจน การสอนกลุ่มย่อยที่ดีต้องมีลักษณะของการเรียนร่วมกัน มีการอภิปราย และมีการร่วมกันแสดงความคิดเห็น ซึ่งในส่วนนี้มีความสำคัญมากที่สุดและมากกว่าการกำหนดขนาดของกลุ่ม หลังจากจบการอภิปรายควรมีการสรุปสิ่งที่ได้จากการเรียน
3. การเตรียมรูปแบบห้องเรียน หรือ สถานที่เรียน เช่น
 - บรรยากาศสบาย ไม่ร้อนไม่หนาว ไม่มีเสียงรบกวน ขนาดห้องเหมาะสมกับจำนวนผู้เรียน
 - อุปกรณ์ช่วยสอนอื่นๆ เช่น คอมพิวเตอร์ เครื่องฉายภาพ กระดาน หรืออื่นๆ
 - รูปแบบการจัดกลุ่มการเรียน ดังแสดงในรูป โดยเน้นให้สะดวกต่อการมีปฏิสัมพันธ์ของผู้เรียนทุกคนและผู้สอน



ข้อดีของการสอนกลุ่มย่อย

1. กระตุ้นให้ผู้เรียนฝึกฝนความสามารถด้านต่างๆ หรือแสดงออกถึงความสามารถ ดังนี้
 - 1.1 ความสามารถในการอภิปราย
 - 1.2 ความสามารถในการแสดงออกและการมีส่วนร่วม
 - 1.3 แสดงให้เห็นถึงความเข้าใจของผู้เรียน
 - 1.4 ความสามารถในการแสดงความคิดเห็น กระบวนการคิด
 - 1.5 สามารถสะท้อนและวิจารณ์สิ่งที่ได้เรียนรู้
2. พัฒนาทักษะที่สำคัญ ดังนี้
 - 2.1 ทักษะการแก้ปัญหา
 - 2.2 ทักษะการสื่อสาร การแสดงออก และ ปฏิสัมพันธ์กับผู้อื่น
 - 2.3 ทักษะการทำงานร่วมกับผู้อื่นและการเป็นผู้นำกลุ่ม
 - 2.4 ทักษะการนำเสนอ

บทบาทและหน้าที่ของผู้สอน

1. ประธานการอภิปราย (Chair person) เพื่อช่วยให้มีการเริ่มการอภิปรายประเด็นปัญหาเริ่มด้วยการใช้คำถามและมักเป็นคำถามปลายเปิดกว้างเพื่อให้ผู้เรียนช่วยกันคิดต่อ ขณะเดียวกันผู้สอนก็จะช่วยนำทางกระบวนการอภิปรายไปในแนวทางตามจุดประสงค์หรือหัวข้อของการเรียนที่ได้ตั้งไว้
2. ที่ปรึกษา (Consultant) เนื่องจากระหว่างการอภิปรายอาจมีคำถามหรือข้อสงสัยที่ผู้สอนจำเป็นต้องเป็นผู้ให้ข้อมูล หรือความรู้ เพื่อให้การอภิปรายดำเนินต่อไปได้
3. สังเกตการณ์และแสดงความคิดเห็นในจังหวะเวลาที่เหมาะสม (Observer)
4. ทำหน้าที่ผู้ร้าย หรือฝ่ายตรงข้าม (Devil's advocate) ในกรณีนี้เป็นบทบาทอีกชั้นหนึ่งที่ผู้สอนจะสามารถทดสอบผู้เรียนทั้งหมดว่าสิ่งที่สรุปได้หรือเห็นตรงกันซึ่งเป็นคำตอบที่ถูกต้องนั้นเป็นเพราะผู้เรียนรู้อย่างถ่องแท้หรือเป็นเพราะเคยเห็นหรือจำได้ โดยผู้สอนอาจใช้คำถามตรงกันข้ามหรือถามเหตุผลว่าทำไมไม่ตอบคำตอบอื่นๆ ที่ผู้สอนรู้ว่าเป็นคำตอบที่ผิด

5. ผู้แนะนำ หรือวางตัวเป็นอุปทูต (Counselor) โดยทำหน้าที่สร้างบรรยากาศการอภิปรายไปในทางสร้างสรรค์ ลดความเครียดระหว่างการอภิปราย เพิ่มโอกาสให้ผู้เรียนที่มีบทบาทน้อย และลดบทบาทของผู้เรียนที่มีความโดดเด่นจนมากเกินไป

ดังที่ได้กล่าวถึงบทบาทของผู้สอน ดังนั้นผู้สอนแบบกลุ่มย่อยจึงควรมีลักษณะที่สำคัญ คือ

- มีความกระตือรือร้นในการสอน
- มีความเข้าใจในลักษณะต่างๆ ของผู้เรียน
- เข้าใจว่าผู้เรียนจะเกิดการเรียนรู้ได้จากหลายๆ วิธี
- ต้องสามารถบริหารจัดการชั่วโมงการเรียนการสอนนั้นๆ ได้เป็นอย่างดี
- เข้าใจจุดประสงค์และหลักการขอการสอนกลุ่มย่อย
- ใช้ทักษะต่างๆ ในการสอนได้แก่ การตั้งคำถาม (questioning), การฟัง (listening), การโต้ตอบ (reacting) การสรุปใจความสำคัญ (summarising) และการเป็นผู้นำ (leadership)

การใช้คำถามเพื่อให้เกิดกระบวนการคิดและการอภิปรายควรใช้คำถามที่เป็น “อย่างไร” “ทำไม” “ให้ผู้เรียนยกตัวอย่าง” ผู้สอนควรกระตุ้นให้ผู้เรียนเชื่อมโยงคำตอบต่างๆ เข้ากับคำถามว่าเป็นเหตุเป็นผล และช่วยให้เกิดการอภิปรายต่อเนื่อง เมื่อกลุ่มสามารถอภิปรายไปได้ตามความเหมาะสม ผู้สอนสามารถกระตุ้นให้ผู้เรียนคิดเชื่อมโยงต่อไปได้แล้วยังเหลือคำถามหรือข้อสงสัยใดๆ ที่ยังไม่ได้คำตอบ และนำไปสู่การวางแผนว่าจะทำอย่างไรในการเรียนหรือการอภิปรายครั้งต่อไป เช่น ควรมีการเตรียมตัวมาก่อน มีการหาข้อมูลต่างๆ มาก่อน เป็นต้น

ปัญหาที่พบบ่อยในการสอนกลุ่มย่อยให้มีประสิทธิภาพ

ปัญหาในส่วนของเนื้อหา

- มีคำถามมากเกินไปเนื่องจากไม่กำหนดวัตถุประสงค์ของการอภิปรายให้ชัดเจน
- ระดับของคำถามเป็นการถามความจำมากกว่ากระตุ้นกระบวนการคิด

ปัญหาในส่วนของผู้สอน

- ผู้สอนให้การบรรยาย (lecture) มากกว่าเป็นผู้ดำเนินการอภิปราย
- ผู้สอนมีบทบาทมากเกินไป
- ผู้สอนไม่ตอบสนองต่อการอภิปราย ไม่กระตือรือร้นที่จะช่วยตอบคำถามที่เกิดจากการอภิปราย และไม่พยายามส่งเสริมให้เกิดการอภิปราย

ปัญหาในส่วนของผู้เรียน

- ผู้เรียนไม่มีการเตรียมตัวก่อนการเรียน หรือ การอภิปราย
- ผู้เรียนชอบที่จะได้คำตอบจากผู้สอนมากกว่าที่จะคิดหรืออภิปราย
- ผู้เรียนที่มีบทบาทโดดเด่นเกินไปจะทำให้ผู้เรียนอื่นๆ ไม่ต้องการอภิปรายหรือไม่ต้องการมีส่วนร่วม

การสรุปหลังจบการอภิปราย

ควรให้มีการประเมินสิ่งที่ได้รับจากการเรียนการสอน ข้อดี และข้อควรพัฒนาได้ ทั้งจากผู้เรียนและผู้สอน การประเมินทำได้หลายวิธี ได้แก่ การทดสอบโดยข้อสอบต่างๆ หรือ ให้ผู้เรียนได้สะท้อนความรู้สึก และสิ่งที่ได้จากการเรียนการสอนกลุ่มย่อย หรืออาจใช้วิธีการทบทวนจากเทปบันทึกการเรียนการสอน (ถ้ามี)

เอกสารแนะนำอ่านเพิ่มเติม

1. Steinert Y. Student perceptions of effective small group teaching. Med Educ. 2004;38(3):286-93
2. McKimm J, Morris C. Small group teaching. Br J Hosp Med (Lond). 2009;70(11):654-7
3. Edmunds S, Brown G. Effective small group learning: AMEE Guide No. 48. Med Teach. 2010;32(9):715-26
4. Learning in small group. In: R. M., Laidlaw J.M., Kanter S.L., 2012. Essential Skills for a Medical Teacher: An Introduction to Teaching and Learning in Medicine. Elsevier Ltd., 137-143

Small Group Teaching – key theories and methods

1. What the literature tells us about using small groups for learning
2. Factors influencing the success of small group learning
3. Small group teaching methods
4. Managing difficult groups/students; tips and pitfalls

Acknowledgement: Some of this material has been adapted from the course notes for PHCM9302 Learning in Small Groups developed by Sue Toohey and convened by Sophie di Corpo.

1. What the literature tells us about using small groups

It is useful to firstly remind ourselves what constitutes a small group.

Fisher and Ellis (1990) emphasise that most of the definitions of a group indicate the **sharing** element among members as the key factor which defines the existence of a group. The sharing can be around perceptions, motivation or goals, as well as around tasks, such as in a scenario group session. This sharing element can be greatly influenced by the **group dynamic** or climate of the group.

The **structure** of the group is another defining element - the roles, norms, values and power relationships that influence the behaviour of group members and tie them to the group, providing the 'glue' of group structure. The structure of a group can influence the level and success of interaction in a group.

Small group work (also known as cooperative or collaborative learning or peer learning) involves a high degree of **interaction**. The effectiveness of learning groups is determined by the extent to which the interaction enables members to clarify their own understanding, build upon each other's contributions, sift out meanings, ask and answer questions.

Secondly, what does the literature tell us? Studies have shown that when looking at long term retention, the **ability to apply knowledge and solve problems, critical thinking and development of positive attitudes**, results consistently favour small discussion classes (McKeachie & Kulik, 1975, McKeachie, 1994). Jaques (2004) argues that the purposes and benefits of group learning coincide closely with the goals of higher education in general.

Benefits of learning in small groups include:

- allowing students to discover and engage with a range of perspectives, ideas, and backgrounds
- providing students the opportunity for more active involvement

Small Group Teaching – OME 2009

- assisting students to clarify their attitudes to and ideas about the subject matter, as they test their own ideas and attitudes against those of others
- helping students develop a sense of academic rigour and a willingness to share ideas
- providing opportunities for students to receive more immediate feedback on their learning
- encouraging students towards self-directed and independent learning
- providing more opportunities for peer learning and sharing responsibility for learning
- providing opportunities for students to more easily gain awareness of their emotional reactions
- providing opportunities for students to learn and develop cooperative behaviour including critical thinking and the process of group problem solving
- more easily establishing rapport between teacher and student
- providing more opportunity to develop skills in communication (listening, responding, interacting) and interpersonal relations

References

FISHER BA & ELLIS DG (1990) *Small Group Decision Making*, (3rd ed) McGraw Hill, Publishing Co, New York.

JAQUES, D (2004) *Small Group Teaching*, Oxford Centre for Staff and Learning Development, UK.

McKEACHIE, W.J. & KULIK, J.A. (1975) *Effective College Teaching*. In F.N. Kerlinger (ed.) *Review of Research in Education*. Itaska, Ill. Peacock.

McKEACHIE, W.J. (1994) *Teaching Tips: Strategies, Research and Theory for College and University Teachers*. (9th ed.) Lexington, Mass. Heath and Co.

TIBERIUS, R.G. (1990) *Small Group Teaching: A Trouble-Shooting Guide*, Toronto, OISE Press and the Ontario Institute for Studies in Education.

UNSW L&T website:

http://learningandteaching.unsw.edu.au/content/LT/teaching_support/smallgroup.cfm?ss=2

Small Group Teaching – MESO 2009 (*Sophie di Corpo*)

2. Factors influencing the success of small group learning

Group Climate

Group climate is the general psychological or emotional state of the group. You may be able to identify from your own experience, groups in which the climate was suspicious, competitive or antagonistic. In such situations it is unlikely that much learning will happen. Some groups have a set of formal rules governing such things as attendance (eg. Scenario groups sessions require 80% attendance) and preparation (eg. students must do the required reading before each class). Even more powerful though is the set of informal rules, which are established over time as the group members learn to work with each other.

David Jaques (1991) points out that many of the difficulties that students have with expressing themselves in groups stem from uncertainty about what the rules of the game are. They may believe that they would be stepping on the group leader's toes if they were to propose topics for discussion, ask questions, or propose a change in direction or procedure. Because they don't want to embarrass a fellow student they may be reluctant to ask questions about another student's opinion or presentation. Students may be afraid to speak up for fear that they will be ridiculed or embarrassed if they make a mistake. For all of these reasons it is productive to devote some time to establishing 'ground rules' for how the group should operate.

You can do this by suggesting some rules to the group and having them discuss and agree on which they would like to use. Rules which are often proposed, include –

- Students can initiate or redirect discussion
- Speakers will be allowed to finish what they have to say
- Speak whenever you wish but after you have spoken try waiting until two or three others have contributed before speaking again, to avoid having a few people dominate
- Treat other people and their contributions with respect
- Everyone takes responsibility for the working of the group process

Group Structure

Group members need to understand what is expected of them. In learning groups, this means that the instructions for any activities are clearly spelled out. Roles such as leader, recorder or reporter are allocated or negotiated and clearly agreed. Everyone understands exactly what has to be achieved and what the time frame is.

Small Group Teaching – OME 2009

Role clarity is particularly important. This includes the role, skills and behaviours of the leader/facilitator of the group. Once a group is well established group leaders/facilitators and members should share many of the group building and the task roles. In the initial stages, it is often up to the group leader/facilitator to diagnose what is needed and ensure that essential social functions occur (such as ensuring that everyone is introduced, that people know something about other members in the group so that they will feel comfortable expressing opinions in front of them) and that essential task functions also occur (everyone is clear about the purpose of the group, the kinds of activities that will be undertaken and the way the group will operate.)

Important roles when facilitating small group learning most commonly relate to supporting the group by building and maintaining good relationships among group members and getting the tasks done. The behaviours associated with these roles are detailed below.

a) Maintaining and supporting the group

These behaviours or roles are aimed at developing the social side of the group. They contribute to building good relationships among members. The examples are based on Jaques (1991).

Encouraging - being friendly, warm and responsive to others, acknowledging others and their ideas, agreeing with and accepting the contributions of others.

For example, *"Hafeez, why don't you tell all of us what you told me about the observations you made in clinical?"*

Mediating - harmonizing, conciliating differences in point of view, making compromises

by saying, *"Other people here might be as worried as you are Aeysha, about the delay in treatment for Mrs Bruce, but we need to take a second look at some of the possible reasons. You had something to say about that, Sue...?"*

Diagnosing - determining and pointing out blocks to group progress -

"we seem to be going over the same ground all over again, is that because we've run out of new ideas?"

Consensus Taking - testing group opinions and decisions by stating them and asking whether or not members agree

Small Group Teaching – MESO 2009 (*Sophie di Corpo*)

"What you're saying Liz is that we should limit the time each member speaks on this issue, is that what other members are thinking?"

Gate Keeping - trying to make it possible for another member to make a contribution, or suggesting limited taking-time for everyone so that all will have a chance to be heard.

"We haven't heard from everyone yet, let's see, Jim, then Sunil."

Standard Setting - expressing standards for the group to use in choosing its subject matter or procedures, rules of conduct, ethical values.

"It is important that everyone feels free to express an opinion in the group, that's the first ground rule, the next..."

Following - going along with the group, accepting the ideas of others, serving as an audience during group discussion, being a good listener.

"That's a good point", "I see", "Let me check with you to see whether I've understood the point you were making"

Relieving Tension - draining off negative feeling by joking or diverting attention from unpleasant to pleasant matters".

"Am I right, we all seem a bit tense at present - let's take a few moments and get some fresh air..." or "I can assure you, this subject is not nearly as difficult as I seem to be making it"

Many of us, who are concerned to get the task done efficiently, tend to underestimate the importance of these behaviours aimed at making the group a pleasant and rewarding place to be. But if they are lacking it is unlikely that the group will perform effectively. Some people will drop out, others will withdraw and make little contribution, a few will dominate and just about everyone will feel that their time and work is under-appreciated.

b) Getting the task done

This can include:

Initiating - suggesting new ideas or a changed way of looking at the group problem or goal, proposing new activities.

Information seeking - asking for relevant facts or information.

Information giving - providing relevant facts or authoritative information or relating personal experience pertinent to the group task.

Small Group Teaching – OME 2009

Opinion Seeking - asking for opinions, judgements or feelings of other group members, seeking clarification of values.

Opinion Giving - stating a pertinent belief or opinion about something the group is considering.

Elaborating - building on a previous comment, enlarging on it, giving examples.

Coordinating - showing or clarifying the relationships among various ideas, trying to pull ideas and suggestions together.

Orienting - defining the progress of the discussion in terms of the group's goals, raising questions about the direction the discussion is taking.

Testing - checking with the group to see if it is ready to make a decision or to take some action.

Energising - stimulating the group, encouraging activity and movement toward group goals

Summarising - revising the content of past discussion

Recording - writing down ideas, suggestions or decisions made by the group

Timekeeping – keeping the group on schedule or to plan.

The roles of group support and the task roles listed above reflect the positive aspects of group work. In reality, some of the behaviour that occurs in groups is not productive and prevents the group from making progress. This is likely to be self-centred behaviour that does not contribute to the group goals but satisfies personal needs. 'Nonfunctional' roles are listed below and can apply equally to students or teachers/facilitators of small groups.

c) Nonfunctional Behaviour

Monopolizing - talking so often or so long that others do not get a chance to speak.

Blocking - interfering with the progress of the group by going off on a tangent, citing personal experiences unrelated to the group's problem, arguing too much on a point the rest of the group has resolved, rejecting ideas without consideration, preventing a vote.

Aggression - criticising or blaming others, showing hostility toward the group or some individual without relation to what has happened in the

Small Group Teaching – MESO 2009 (*Sophie di Corpo*)

group, attacking the motives of others, deflating the ego or status of others.

Seeking Recognition - attempting to call attention to one's self by excessive talking, extreme ideas, boasting, boisterousness.

Special Pleading - introducing or supporting ideas related to one's own pet concerns or philosophies beyond reason, attempting to speak 'for the grass roots', the 'patients', 'the common man', and so on.

Withdrawing - day-dreaming, sleepiness, becoming indifferent or passive, resorting to excessive formality, doodling, whispering to others.

We will cover some strategies for dealing with problem behaviour in **Managing difficult groups/students; tips and pitfalls**

References

JAUQUES, D (1991) *Learning In Groups*, (2nd ed) Kogan Page, London.

Small Group Teaching – OME 2009

3. Small group teaching methods

Good small group work rarely just happens. It relies to a great extent on preparation by the teacher and involves being clear about what the session is designed to achieve, identifying useful problems, cases or other material which might form the basis of learning and identifying crucial questions which will get people thinking.

Planning your small group session

Before meeting with your group you need to plan your session. At the very least make sure you review the following:

Content:

What are the most important points for the session? What are the likely errors that students may make in trying to understand the topic/issue/skill? Are there important principles, or key concepts that you want students to understand?

Aims:

It is important to be clear about what you want to achieve in your small group teaching session. Brookfield (1990) suggests that the following aims are well suited to discussion-based teaching. Consider whether any of these could describe the aims of your small group teaching:

1. To engage students in exploring a range of perspectives and discovering new perspectives
2. To emphasize the complexity and ambiguity of issues, topics or themes.
3. To help students recognize the assumptions underlying their habitual ideas and behaviours.
4. To increase intellectual agility
5. To encourage active listening
6. To increase students' interest and involvement with a topic;
7. To show students that their opinions and experiences are valued;
8. To help develop a sense of group identity;
9. To encourage democratic habits such as valuing participation, respect for others' opinions and tolerance of diversity.

Activities and Questions:

What kinds of activities might you use or what kinds of questions might you ask that help to raise students' interest in this topic, establish what is and is not an example of this concept or condition, expose students' misunderstandings and/or help students understand the complexities of this issue?

Small Group Teaching – MESO 2009 (*Sophie di Corpo*)

Conditions:

How might you need to modify what you can do according to the number of students, the time available, and the space you are working in? For example, should you break the students up into smaller groups for all or part of the session? There are many different reasons why you might want to split your group up into smaller subgroups; to give everyone an opportunity to discuss an issue or to get some 'hands-on' experience in a new skill or technique; to encourage quieter students; to discourage the more dominant students. Although the physical arrangement of chairs and tables is one of the most basic tasks in providing comfort, it is also highly influential in the flow of discussion in the group.

Techniques for Small Group Work

In small group work the teacher or facilitator sets up activities, asks questions, listens and responds to students' comments and questions, occasionally explaining, often asking further questions. S/he is often the one to pull the session to a close, by summarising the understanding that has been reached and the questions that remain. Students too, ask and answer questions, explain their ideas and summarise each other's arguments.

It follows that the skills needed in small group teaching are predominantly questioning, listening, responding, explaining, and summarising and that students need to develop them just as much as teachers. In addition, teachers need to be able to prepare materials and activities for students to work on and they need to be able to prepare students so that they understand what small group teaching is about and develop the skills described above.

Many different techniques have been developed to accomplish the different purposes of small group learning. We will look briefly at two, discussion groups and briefing and debriefing practical, clinical or experiential learning.

Discussion groups (open)

A good discussion class allows students the opportunity to expose their individual conceptions and misconceptions and to compare their ideas with those of others. Cognitively, the act of putting material into one's own words is a very important step in developing understanding, as well as providing an opportunity to begin using the language of the discipline.

One important point to be made about open discussion is that although it appears to be quite spontaneous and even chaotic, its success depends to a considerable extent on students doing some preparatory work and on the questions that the teacher/facilitator poses in order to start the discussion.

Educational reasons for questioning may differ from clinical purposes. They include to

Small Group Teaching – OME 2009

- stimulate learning and thinking
- assist the learner in organizing and clarifying concepts
- correct misunderstandings or faulty reasoning
- assist in showing special or obscure relationships
- strengthen the learner's ability to synthesize and analyse
- correct attitudes or behaviour.

Teachers/facilitators have an interesting dilemma here. The choice is basic to the whole philosophy of teaching and learning in the clinical professions. Clinical teachers must ensure that the student '*gets the right answer*' in the interests of patient safety. But is learning the right answer on this particular patient the best path to being right on the next patient when the teacher is not there? What is the focus of your small group session? *Knowing that...or Knowing how to work it out?*

The issue for each clinical teacher is how much to concentrate on 'the facts' and how much on 'reasoning from the facts'. What may seem to be a quibble about balance, actually profoundly affects the way a tutorial is conducted. To polarize the extremes, some tutorials are a 'lecture to a small group' or a question and answer session - teacher questions, student answers. Or, alternatively, the session may function around student questions, hypotheses and guesses to be explored and justified; and the teacher is used by students as a resource, not as the source.

The teaching skills needed are quite different between the two approaches. The position of the teacher in relation the students is quite different. Some teachers are very uncomfortable at being questioned by students. Some students resent not being told the facts which the teacher obviously knows.

Some question types that are useful to use in discussions include

Asking for more evidence: How do you know that? What data is that claim based on? Do you have any evidence for that?

Asking for clarification: Can you put that another way? Can you give us an example of what you are talking about? What do you mean by...?

Open questions: How ... do you think that may work? Why ...

Linking questions: Is there any connection between what you've just said and what X said...? How does your idea support what has been said so far? Hypothetical questions: How would this change if the xxx was xxx?

Cause and effect questions: What would be the effect of?

Summary and synthesis: What are two of the most important ideas that have emerged from this discussion? What do you understand better as a result of this discussion? What remains unresolved or contentious?

Small Group Teaching – MESO 2009 (*Sophie di Corpo*)

Brookfield and Preskill (1999) suggest that discussion topics are always more interesting when framed as a question rather than a statement. It's also important to pick a topic that is not too factual or uncontroversial. You may well want students to acquire quite a lot of factual information but they can equally well learn it by considering questions like – What are two feasible explanations for what's going on here? When is it *not* desirable to intervene with the standard treatment?

Apart from using a provocative question to start discussion, Brookfield and Preskill suggest the following strategies:

Frame the discussion around student questions

Split students into pairs or small groups (3-5) and ask them to identify what they think are the most important questions that need to be answered about a particular case, situation or problem. Questions can then be put on the board and agreement reached about which are the most important or interesting ones which will be addressed first. Students can also be given this task as homework so that they come to the class with questions prepared.

Start with a sentence completion exercise

Students are asked to choose one of the following statements and complete it, then to share their statement with a subgroup if the class is large or the whole group if it is not. In the groups students can choose the statement that they find most interesting and want to explore further.

The statements:

What most struck me about the reading (or lecture / case / data / other stimulus material) was.....

The point I most take issue with in the reading etc was.....

The point I found most confusing was.....

The question I would most like to ask the author / patient / consultant is...

The part of that experience I found most confusing was....

Generate truth statements

In small groups ask students to generate statements which they believe to be true about the topic. This technique may be particularly useful when dealing with a topic about which there are many popular misconceptions. The complexity and ambiguity of knowledge is revealed

Small Group Teaching – OME 2009

as each group presents their truth statements and other groups raise questions about them or refute them. The idea is not to generate statements which are factually true but to identify issues for further research and exploration.

Start with a personal experience

Ask students to volunteer their personal experiences with a topic or to give their personal reaction to a case history, video etc. Students who are new to discussion may be reluctant to speak when they don't feel knowledgeable enough. However most people feel they are experts on their own lives. As students progress it may be useful to ask them what they thought and felt about the experience at the time and whether their underlying assumptions have now changed or they would interpret the situation differently now. A possible educational benefit is to help people see their stories from different perspectives and understand their experiences in new ways.

In the early life of a learning group, members may benefit from practising how to *listen* effectively, how to process the contribution of other members and in sharpening their own responses.

Briefing and debriefing practical, clinical or experiential learning

When **briefing** students before a practical, clinical or experiential learning session, you can ask students to think for a few moments and to make some notes for themselves as to what they hope to learn from the forthcoming experience and how they expect to learn. Then ask students to discuss and compare their expectations. Prompt students to think about how they might make the most of the experience – will they be doing all of their learning on the ward or should they be doing follow up reading? What kinds of contacts can they initiate for themselves? What is the protocol?

You might also explore their emotional readiness for the work ahead – how confident or anxious do they feel? Do they feel adequately prepared? Do they have the level of knowledge necessary for to-day's learning?

Remind students that in the debriefing later, what and how they learned will be discussed. If the experience is to extend over some time, it may be useful to ask students to keep a record of any 'critical incidents' – incidents that were significant learning experiences for them. These can provide rich material for discussion in the debriefing session.

If **debriefing** were simply a matter of checking what each student had learned there would be no point in meeting as a group. Each student has observed from an individual perspective. Getting the 'whole picture' is important and is not often possible without group discussion following clinical exposure.

Small Group Teaching – MESO 2009 (*Sophie di Corpo*)

Debriefing may be necessary during as well as after. An individual student, the group as a whole, or the tutor could initiate a debriefing session as a particular issue or patient problem emerges.

In a debriefing session you can:

Provide support by showing interest in students' problems, recognising concerns and acknowledging commendable performance, recognising difficulties, praising where due.

Provide opportunities for students to review their progress by assisting students to determine further learning activities in relation to their outcomes and encouraging students' to review their own progress.

Acknowledge partnership in learning by offering reviews of your own clinical or teaching performance.

Give feedback on students' performance by providing feedback requested by students in briefing sessions and using information from direct observation and by providing concrete examples and checking that the feedback is congruent with students' perception of performance.

Invite reflection on the events of the attachment by prompting students to go over what happened, what was surprising, different, frightening, satisfying, disappointing and so on ; encouraging expression of feelings about what happened during the attachment; inviting examples of new discoveries, new knowledge and/or insights as they applied their knowledge to patient's problems; prompting students to draw meanings from their personal experiences ; encouraging students to determine their own outcomes for further learning.

References

BROOKFIELD, S.D. (1990) *The Skillful Teacher*. San Francisco, Jossey Bass.

BROOKFIELD, S.D. and PRESKILL, S. (1999) *Discussion as a Way of Teaching*. Buckingham, S.R.H.E. & Open University Press

Small Group Teaching – OME 2009

4. Managing difficult groups/students; tips and pitfalls

Let's think first about the pitfalls as the tips cover some strategies to avoid these. Common pitfalls or likely problems of small groups have been grouped by Tiberius (1990) under the following headings:

Group goals; they are unclear, unattainable or unacceptable

Group interaction; it is lacking, teacher dominates, students participate unequally

Group motivation and emotion; students are tuned out, teacher is tuned out or students don't cooperate.

What is most useful in Tiberius' approach is that he identifies possible causes for each of these. Rather than apportion blame to the students, he lists other underlying factors that may be contributing. Even though his book is now 10 years old, his trouble-shooting guide is applicable in the small groups we have today and I strongly recommend you get a copy.

A very brief summary of some of his most useful points follow around the three areas referred to above:

1. Pitfall - Group goals; they are unclear, unattainable or unacceptable

Possible causes:

failure to establish goals for the group; digression from goals; process not matched to goals; poor time management; teacher and students perceive goals differently

Suggestions:

set clear goals; establish goals for each session/meeting; make goals relevant to those of the students; agree on timeframe and remind students

2. Pitfall - Group interaction; it is lacking, teacher dominates, students participate unequally

Possible causes:

lack of, or bad, experience with small group learning; students not rewarded for participation; low level of trust; teacher's authority is overwhelming; dominant speakers monopolise the discussion

Suggestions:

make clear what skills are needed for small group learning and explain the benefits; agree on a set on group or ground rules; reflect on both content and process of the group; encourage students to reward one another; reward students contributions by using them; remember who said what; talk less; provide opportunities for students to cooperate and trust one another; talk to the dominant student privately and/or assign a task to this student.

Small Group Teaching – MESO 2009 (*Sophie di Corpo*)

- 3. Pitfall - Group motivation and emotion;** students are tuned out, teacher is tuned out or students don't cooperate.

Possible causes:

little interest in topic; relevance of group process unclear; students are preoccupied with the exam; lack of institutional support for teaching; lack of feedback from students; students do not accept the assumption of the course; excessive competition among students; disruptive behaviour

Suggestions:

explore students' personal motivations to study topic; provide evidence topic is relevant or interesting; tap into students' experiences; review group rules if needed; address the exam agenda and define activity as useful to exam; support teaching and document teaching activities; reward excellence in teaching; ask for formative feedback from your students; don't let discussion of presuppositions replace the content; expose competitive behaviour; emphasise cooperative learning; break the group up into pairs or smaller groups; ask students to take different roles; if needed speak to any students privately.

Some other pitfalls you may encounter more specifically include

Pitfall 1: Students don't prepare

Ask the students why. Consider beginning the class by giving students short extracts to read or data to review so that all the class is familiar with the material they are to work on. If you think it is reasonable to get them to prepare ahead of time:

- emphasise the importance of preparation,
- consider calling off a class if you find that most students are not ready, in order to make your point, and
- *make sure to use what the students have prepared in the class.*

Students will quickly realise that whatever you say about the importance of preparation there is not much point in doing it if no one will notice.

Pitfall 2: Students don't participate or seem disengaged

This relates to Tiberius' motivation. Is it all the students or just some? If no one wants to participate consider whether past experiences (in this group or others) have made participation a risky business because of the likelihood of being criticised, embarrassed or humiliated for making a mistake. It may also be that the teacher or facilitator who had the group before dominated the discussions and they expect you to do the same.

Reiterate why you think small group work is important and (re)establish ground rules for discussion. (Try including 'No put downs'.)

Small Group Teaching – OME 2009

Begin the session by breaking students into pairs or subgroups to work on a task. Have reporters report on the group opinion.

Start with a question that's easy to answer such as 'What has been your personal experience with.....?'

Pose a question and give students a few minutes to think about it and make some notes before calling on someone.

Refer to students' points in the discussion and when summarising (eg 'as Joe said.....') so that they know that you have been listening and that their contribution was valuable.

Pitfall 3: One person dominates

This relates to Tiberius' interaction. Thank the talkative person for their contribution and then invite others to speak. Interrupt them and invite other comments ('Before you go on, I would like to see if anyone else has an opinion on that')

Use structured participation, such as going around the group (each person speaks or passes.)

Break into subgroups. Ask the talkative person to be the scribe.

Rearrange the seating so that you are sitting beside the talkative person.

Refer to ground rules (if you made a rule about valuing wide participation.)

Speak to them privately. Explain that while you understand that they like to participate actively and that you appreciate their enthusiasm, you are concerned that their confidence and articulateness may inhibit others from participating. Ask them to hold back a bit so that you can encourage others to come in.

Pitfall 4: Students complain about how you run the group

Check what their goals are for the group. Is the problem that students do not see how the group work is contributing to their goals - which usually involve passing exams or other forms of assessment?

Explain why you do things and how what you do contributes to their goals (short term and long term professional goals). If it doesn't contribute, consider how it might and negotiate.

Ask for suggestions about how the group might be better run. Discuss with the group and negotiate alternative strategies.

Some useful tips when teaching a small group which can be used as a checklist include:

Tip 1: Be prepared

As we have already stated, good small group work rarely just happens. It relies to a great extent on preparation by the teacher and involves being clear about what the session is designed to achieve.

Tip 2: Make introductions and set ground rules

Introductions are important because it is difficult to have an open discussion with someone when you don't know anything about them, even their name. It is also a great opportunity for the group leader/facilitator to find out a little more about students' backgrounds with the topic. You may also wish to hear the students' previous experiences of tutorials, especially what worked well for them, and what didn't.

We have already covered the importance of negotiating and clarifying ground rules for discussion so that students know what the 'rules of the game' are. Of course if you are working with students who have already had extensive experience in small group work this need not be a lengthy process but it is probably still worth doing for the value it has in heading off later conflicts.

Tip 3: Use questioning effectively

Asking questions, considering the answer, knowing when to respond with a comment or explanation and when to use a follow up question or re-direct the question to someone else are key skills in keeping a discussion going and keeping students interested and involved.

Tip 4: Explain at the appropriate time

George Brown who wrote extensively about the art of lecturing and explaining, pointed out that when it comes to small group sessions, knowing *when* to explain was probably more important than knowing how to explain. Because the objective of small group teaching is to encourage students to think and discuss, it is unwise for the teacher to offer too much in the way of explanation too early. Students will quickly realise that all they have to do is remain silent and the teacher will provide all of the answers.

It is usually much better to provide any explanation needed after the students have made a good attempt at the task for the session. It may be a good idea to hold the explanation until you are summarising at the end of a task or the end of a session. At this point you can draw together the responses of the group, correct any misconceptions that may have arisen and

make use of the students' own contributions in any explanation that is given. Having their contributions recognised in this way rewards students for participating and builds up their confidence. It is likely to encourage them to contribute to future sessions (Brown and Atkins, 1988).

Tip 5: Summarise and Close

Summarising may be used at the end of an activity as well as at the end of a session to bring together the key points that have been made, the key understandings that have been arrived at and the unresolved questions that remain for further discussion or research.

If you are aiming for students to improve their teamwork and communication skills it may also be useful to summarise the processes that have been used, the stages that the group has gone through, for example, in problem solving, and the progress that has been made.

Summaries help to show students what is important in a topic and how that links to related topics. They are useful in helping students develop the 'well-structured knowledge base' that is the foundation for expertise. Good judgement is required in deciding what to highlight and what to omit.

You can also, of course, delegate the process of summarising to students, asking a couple of students to highlight the key points that have been made and someone else to identify important questions that remain unresolved.

Closing a session also involves other courtesies. Thanking the group for their contributions and pointing out what has been achieved is good for morale and helps to develop the cohesiveness of the group.

Tip 6: Evaluate your teaching

Ask your students how you are going. Ask a peer to come in and sit in on a session, using the Peer Review Checklist, see attached. Make use of the institutional evaluation processes, such as CATEI.

References

BROWN, G. & ATKINS, M. (1989) *Effective Teaching In Higher Education*. London, Routledge

TIBERIUS, R.G. (1990) *Small Group Teaching: A Trouble-Shooting Guide*, Toronto, OISE Press and the Ontario Institute for Studies in Education.

Teaching at Ambulatory Care

ยิ่งยง ชินธรรมมิตร

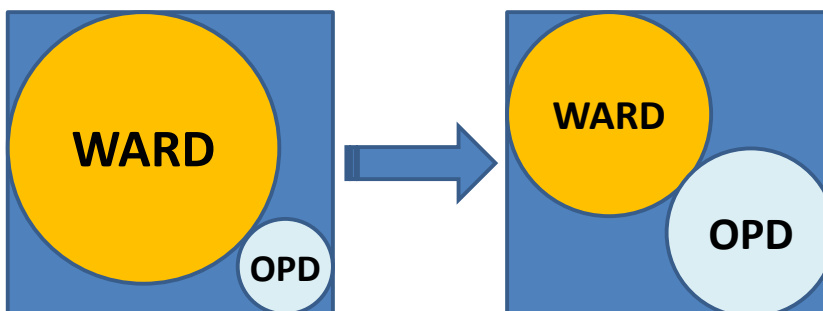
คณะแพทยศาสตร์ศิริราชพยาบาล

Outline

- Importance and challenge
- Cycle of learning
- Teaching method
- Model

Why Ambulatory Setting

- Changes in hospital practice
- Changes in student numbers
- Changes in educational requirements
- OPD has advantage



Med Teacher 2005;27:302

Challenge

- Time : the pace is rapid
- Reduced opportunity for direct observation
- Variability of patients' problem
- Lack of continuity
- Decrease Patient satisfaction?

MJA 2006;185:166

Cycle of Learning in the OPD Setting

- Planning
- Learning
- Appraisal and assessment
- Reflection

MJA 2006;185:166

Planning

Learning

**Appraisal/
Assessment**

Reflection

- Define course **outcome** (eg, managing common presentations) & methods of assessment
- **Orientation** to the practice, patient care (time spent with a patient, focused interaction), learning and resources
- Consider **organization of the clinic** (eg, having a second room, “wave scheduling”)

MJA 2006;185:166



- **Pre-select patients** for review (based on the experience level of the learner) and tell the trainee what complaint to focus on
- Orient students to patients
- Ensure authentic **patient contact** (watch-interact-manage patient) (**independently if possible**)
- Students present their cases ... **Discuss** case
- Approach patients
- **Help student to do part of the chart preparation**
- Set aside time for a tutorial
- Use other members of the team for teaching (eg, nurse, patient educator)

MJA 2006;185:166; JAMA 2000;283:2362

Planning

Learning

**Appraisal/
Assessment**

Reflection

- Allow direct observation and give **feedback** (may gather data from others, the receptionist often knows whether patients want to see your trainee again)
- Use **questions** to ascertain understanding

MJA 2006;185:166



- Help **students** do effective reflection
- Self-reflection by the **teacher** will improve subsequent teaching
- Consider whether the student/trainee experience was optimal
- **Plan** the next session

MJA 2006;185:166

Structured Framework for Teaching

- One-minute teacher
- SNAPPS – learner-driven approach of case presentation

1-Minute Teaching

5-Step Microskills

METRC

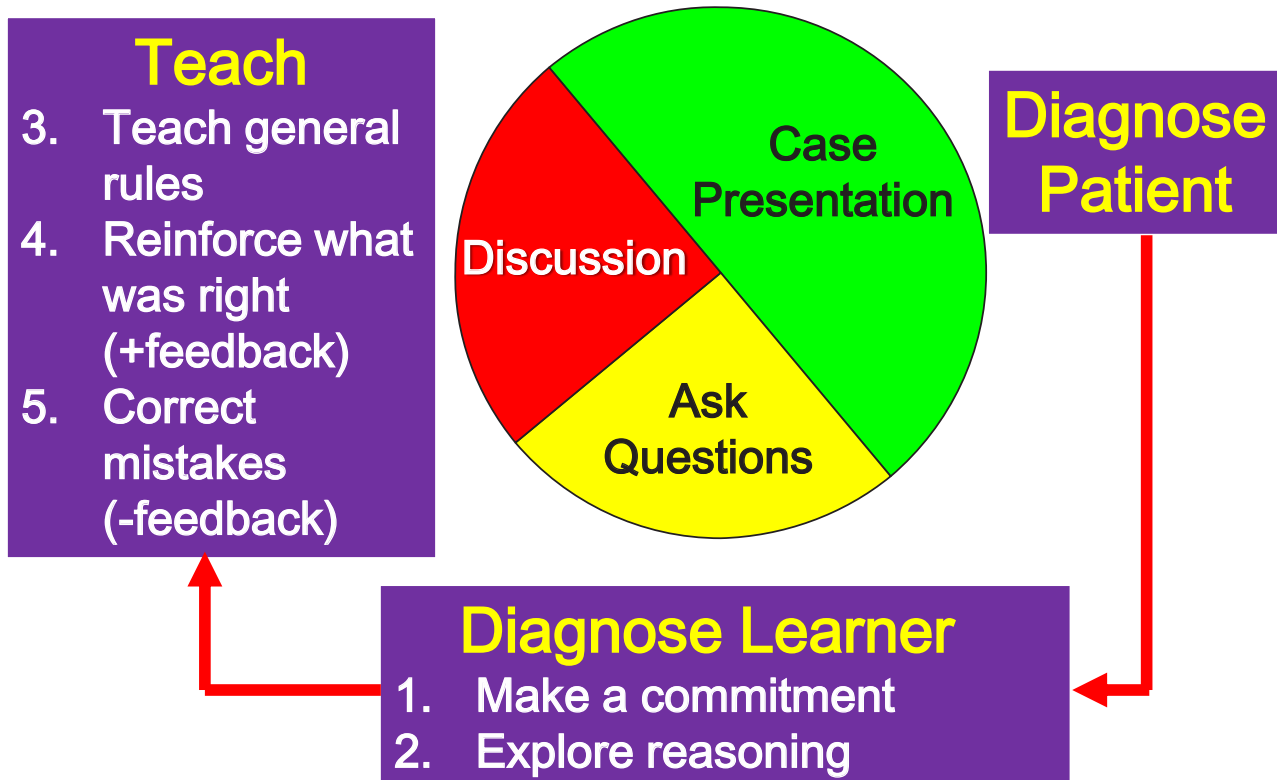
1-Minute Teaching

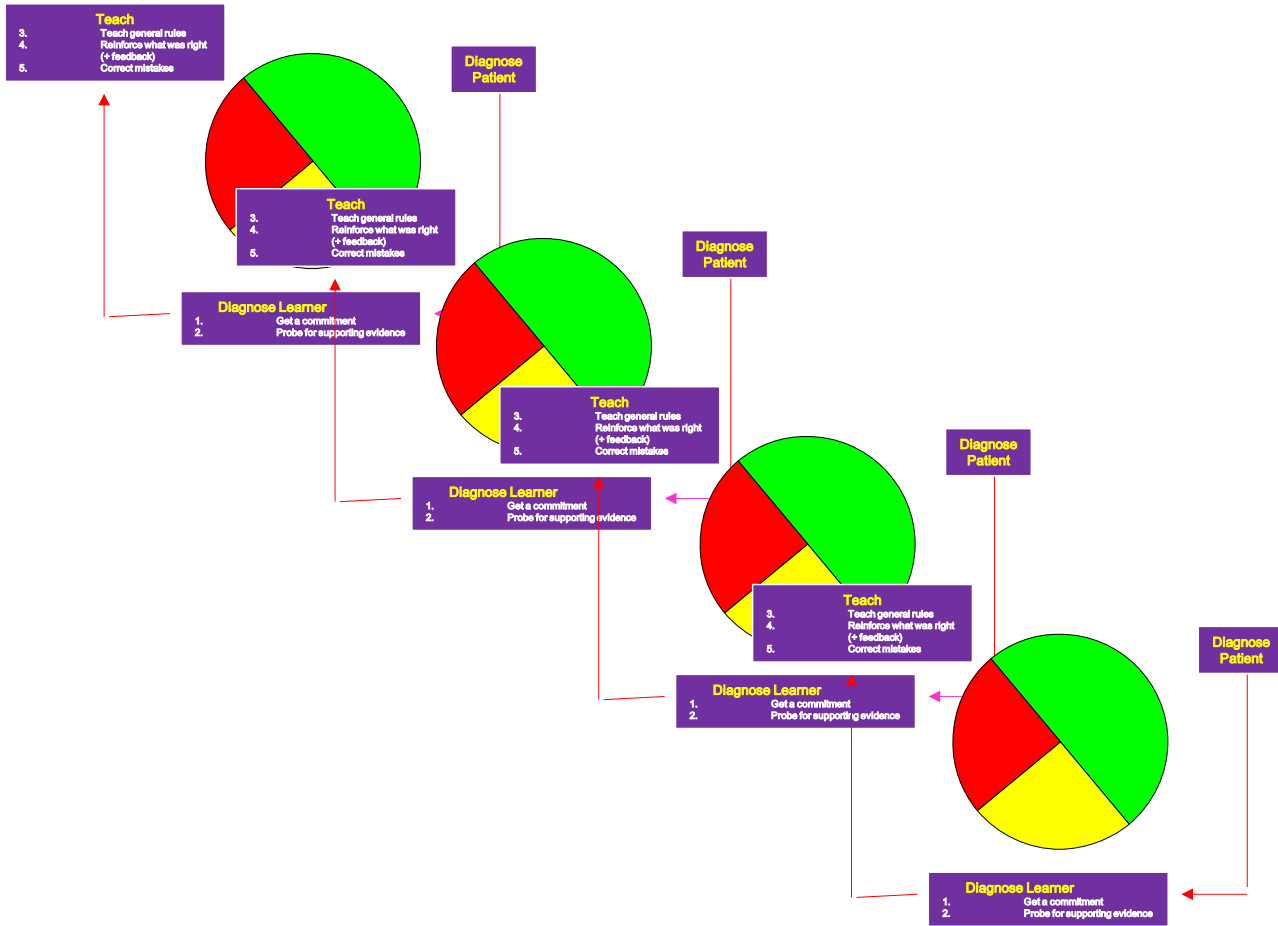
5-Step Microskills

- **M**ake a commitment Questioning
- **E**xplore reasons Questioning
- **T**each general rules
- **R**einforce what was right + feedback
- **C**orrect mistakes - feedback

1-Minute Teaching

5-Step Microskills





SNAPPS used by Student

- **S**ummarize the history and physical findings
- **N**arrow down the differential diagnosis (DDx) to 2-3 possibilities
- **A**nalyze the DDx by comparing and contrasting the possibilities
- **P**robe the teacher with questions about uncertainties, difficulties, or alternative approaches
- **P**lan management
- **S**elect a case-related issue for self-directed learning

Acad Med 2003;78:893

Time-efficient Strategies

1. Assess the student's clinical abilities and tailoring the learning experience to their level
2. Using at least 2 examination rooms
3. Allowing the student to perform the initial interview and physical exam independently
4. Help student to do part of the chart preparation

JAMA 2000;283:2362

Strategies for Joint Consultation with Clinician and Student Present

- Before the observed consultation, give the learner a framework for thinking, and discuss his or her reflections later
- Watch the learner take the history or perform physical exam, and provide feedback
- Use structured frameworks for teaching
- Ask the learner to look up medications or side effects during consultation
- Get the learner to provide the information on lifestyle changes (eg, smoking cessation)
- Get the learner to record observations in the patient notes

ประสบการณ์
จำนวน



ห้องรวม
ห้องแยก

ความยากง่าย
จำนวน/เวลา
ความเป็นส่วนตัว

Models for Organizing Student/Patient Contacts in OPD Clinics

- One student/one clinician
- Several students/one clinician
- Several students/several clinicians

Med Teach 2005;27:302

Several students/
One clinician

Grandstand model

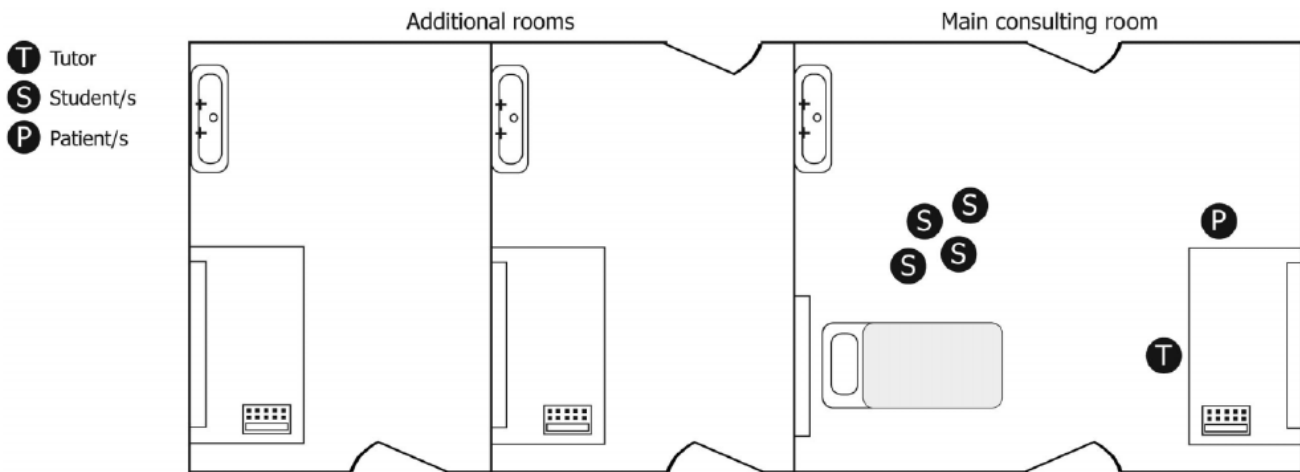


Figure 1. Grandstand model.

Several students/
One clinician

Breakout model

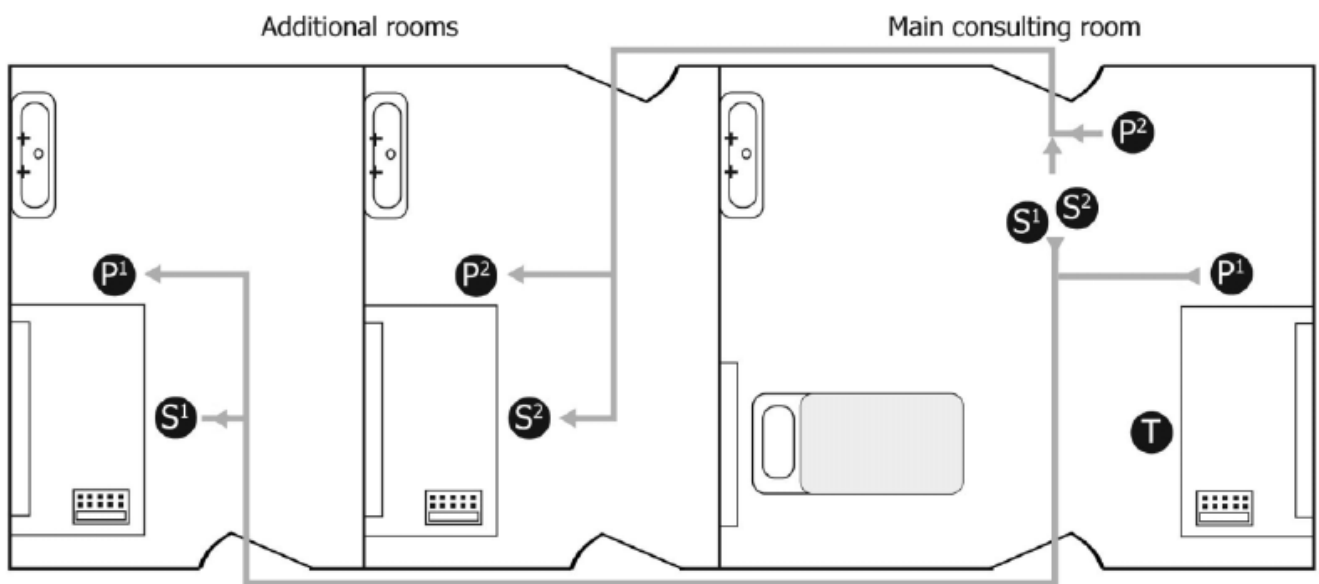


Figure 4. Breakout model.

Several students/
One clinician

Supervising model

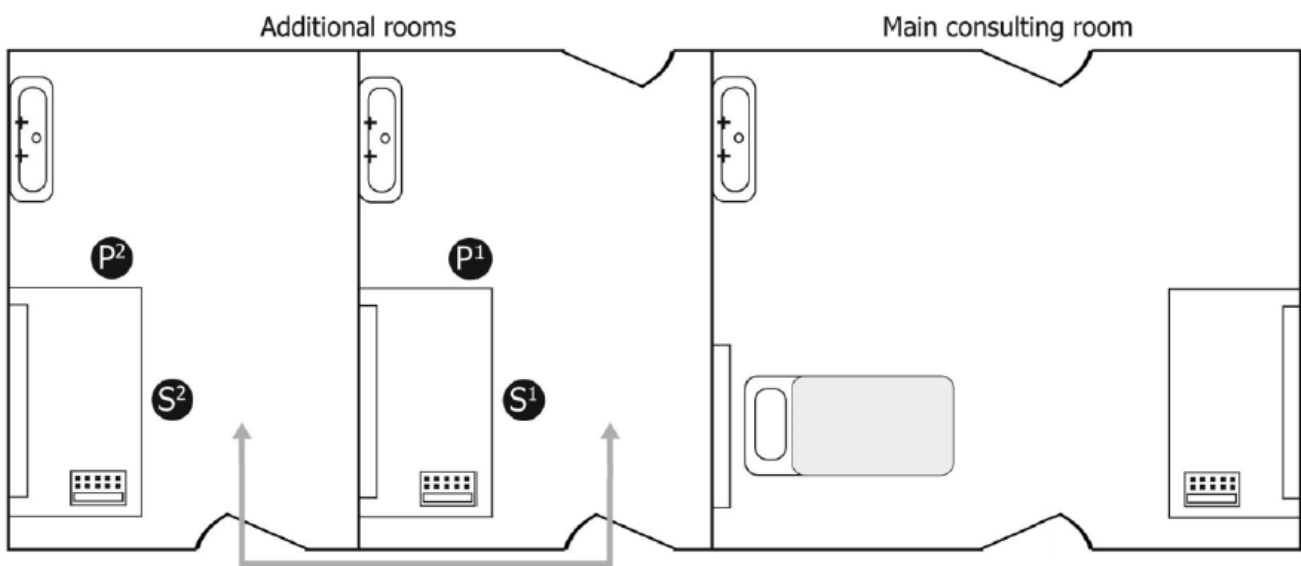


Figure 2. Supervising model.

Several students/
One clinician

Report-back model

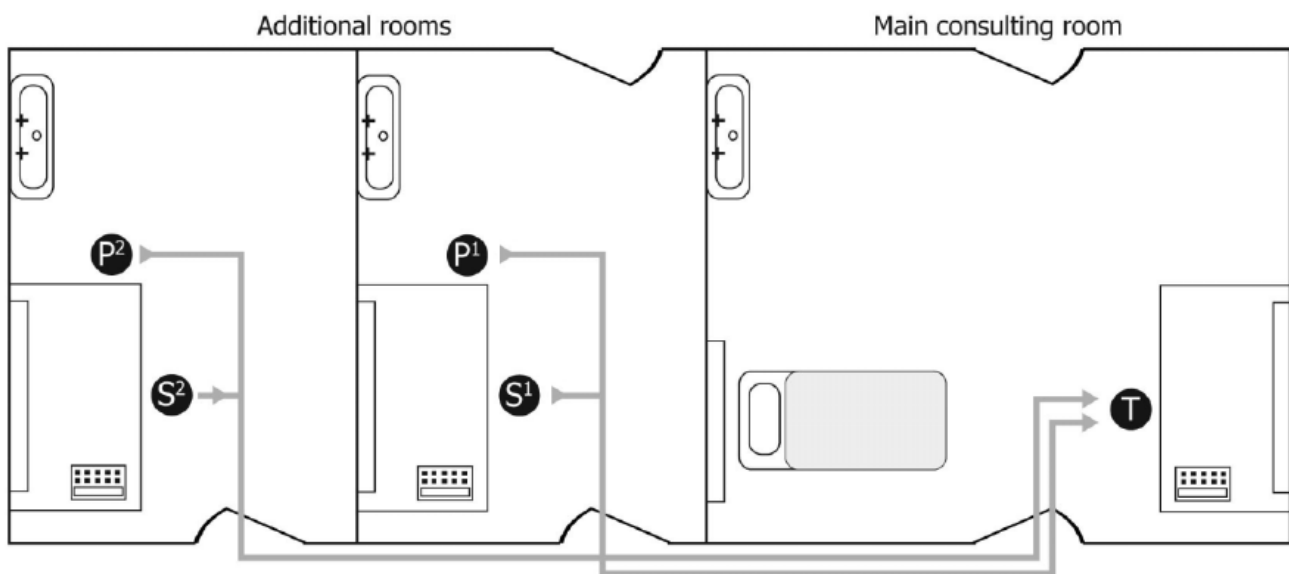


Figure 3. Report-back model.

	Room 1	Room 2
9.00-10.00		
10.00-10.30		Empty
10.30-11.30		
11.30-12.00	Empty	

Acad Med 2002;77:593

One student / One clinician

- Sitting-in model – student as observer
- Apprenticeship model – tutor as observer
- Team member model – separate room for (senior) student

Med Teach 2005;27:302

Several students / Several clinicians

- **Shuttle** model – students move to a clinician with interested case
- **Division** model – student group is divided
- **Flip-flop** model – switch group half of time
- **Tutor** model – one teacher with a limited number of selected patients for students (other patients seen by other clinicians)

Med Teach 2005;27:302

Ambulatory Care Teaching Centre

U. of Dundee, Scotland UK



Ambulatory Care Teaching Centre

U. of Dundee, Scotland UK

- **Problem**

- **IPD case:**

- more acutely ill and may be less willing or able to participate in student teaching
 - too serious condition, too complicated (referral)
 - pre-admit full work-up

- **Routine OPD clinics:**

- too busy to provide the constructive supervision

Ambulatory Care Teaching Centre

U. of Dundee, Scotland UK

- **Problem**
- **Solution**
 - If dedicated **space** can be made available, a suitable environment can be created where students may spend time with selected patients under appropriate supervision.
 - That offers students a **bridge** between their experiences in a **clinical skills centre** and subsequent visits to wards or OPD departments.

Ambulatory Care Teaching Centre

Advantage

- Optimal learning environment
- Adequate supervision
- Not jeopardize or hamper patient care
- All students received the same experiences
- Routine OPD service is not jeopardized

Disadvantage

- Find a definite space
- Require a **budget** – maintain facilities, reimbursement of patients or clinical volunteers
- Need a full-time tutor or **coordinator**

Med Teach 2005;27:302

Ambulatory Care Teaching Centre

U. of Dundee, Scotland UK

12 tips to develop ACTC (design-**implement**-**evaluation**)

1. Allow development time
2. Integrate curriculum needs and identify organizational constraints
3. Identify interested parties and their strategic role as a steering group
4. Find suitable accommodation

Med Teach 2005;27:302

Ambulatory Care Teaching Centre

U. of Dundee, Scotland UK

12 tips to develop ACTC (design-**implement**-**evaluation**)

5. Secure a budget
6. Acquire suitable resources and equipment
- 7. Recruit and train enthusiastic staff**
- 8. Evolve an implementation function for the steering group**

Med Teach 2005;27:302

Ambulatory Care Teaching Centre

U. of Dundee, Scotland UK

12 tips to develop ACTC (design-**implement**-**evaluation**)

9. Build up a bank of referred patients or clinical volunteers

10. Implement a teaching plan

11. Develop a multifaceted evaluation process

12. Develop a research and development function for the steering group

Med Teach 2005;27:302

Ambulatory Care Teaching Centre

U. of Dundee, Scotland UK

- The teaching method and style are more important than the environment for attaining learning outcomes.
- Students and patients prefer the learning environment of a dedicated ACTC.
- The ACTC is advantageous to patient welfare and quality of service.

Med Teach 2005;27:358

การสอนข้างเตียง (Bedside Teaching)

สุมาลี นิมนานิตย์

*There should be "no teaching without a patient for a text, and
the best teaching is that taught by the patient himself"*

Sir William Osler 1903

การสอนข้างเตียง เป็นวิธีการสอนที่มีมาช้านาน ครูแพทย์ในสมัยก่อนเน้นย้ำความสำคัญของวิธีการสอนนี้มาก แต่ปัจจุบันวิวัฒนาการด้านวิชาการและเทคโนโลยีรุดหน้าไปอย่างรวดเร็ว ทำให้วิธีการตรวจ เครื่องมือ/อุปกรณ์ และยาที่มีประสิทธิภาพ ซึ่งช่วยให้การวินิจฉัยโรคแม่นยำขึ้น และการรักษาที่ได้ผลดีขึ้น ขณะเดียวกันก็มีพัฒนาการด้านการเรียนการสอน ทำให้วิธีการสอนอื่นๆ เพิ่มขึ้น เช่น การสอนกลุ่มย่อยในห้องเรียน ซึ่งเปิดโอกาสให้นักศึกษาคิดวิเคราะห์และแสดงความคิดเห็น เป็นต้น นอกจากนี้ยังมีการนำเทคโนโลยีอิเล็กทรอนิกส์มาสร้างสื่อการสอนต่างๆ อีกมากมาย ดังนั้นจึงน่าจะได้พิจารณาบทบาทของการสอนข้างเตียงนั้น สอนอะไร ยังจำเป็นต้องมีหรือไม่ ทดแทนด้วยวิธีการสอนและสื่อการสอนสมัยใหม่ เครื่องมือ/อุปกรณ์และยาที่ทันสมัย ได้หรือไม่

ในความเห็นของผู้เขียน การสอนข้างเตียงเป็นการสอนให้นักศึกษา "เป็นหมอ" และมี "ความเป็นหมอ" การ "เป็นหมอ" นั้น คือ การมีความสามารถในการวินิจฉัยและรักษาโรคได้ถูกต้องแม่นยำ ส่วน "ความเป็นหมอ" นั้น หมายถึงว่าจะต้องมีจิตใจที่ต้องการช่วยให้ผู้อื่นพ้นจากความทุกข์ทรมานทั้งทางกายและทางใจ และถึงพร้อมด้วยจริยธรรมและคุณธรรม ที่แสดงออกได้ด้วยภาษากาย (body language) และวาจา ดังนั้นการสอนข้างเตียงจึงไม่สามารถทดแทนด้วยตำรา วิธีการสอนอื่น สื่อการสอนสมัยใหม่ เครื่องมือ/อุปกรณ์ และเทคโนโลยีที่ทันสมัยได้ และยังผู้วิชาการและเทคโนโลยีด้านการแพทย์ก้าวรุดหน้าไปรวดเร็วเท่าใด การสอนข้างเตียงก็ยิ่งมีความสำคัญเพิ่มมากขึ้นเท่านั้น

การสอนข้างเตียงเป็นวิธีการสอนที่มีความพิเศษหลายประการ ประการที่สำคัญที่สุด คือ มีผู้ป่วยร่วมอยู่ในการเรียนการสอนด้วย ครูและนักศึกษา (ในที่นี้หมายรวมถึงนักศึกษาแพทย์และแพทย์ประจำบ้าน) จึงมีโอกาสที่จะร่วมกันทำความเข้าใจกับผู้ป่วยและปัญหาของผู้ป่วย และร่วมกันแก้ปัญหา ครูมีโอกาสที่จะรู้จักนักศึกษา รู้ว่านักศึกษามีความรู้ ทักษะ และเจตคติอะไรและอย่างไรจากการเฝ้าสังเกต นอกจากนั้นยังสามารถสอดแทรกเจตคติ จริยธรรม และให้ "feedback" ในขณะนั้นได้ทันที ดังนั้นนักศึกษาจึงมีโอกาสได้รับรู้ ทักษะ และเจตคติต่างๆ ดังนี้

1. ได้เรียนรู้วิธีการปฏิบัติต่อผู้ป่วย โดยมีครูผู้สอนเป็นแบบอย่าง ในการวางท่าทาง ท่าที วิธีการพูด/การฟัง การสัมภาษณ์ การตั้งคำถาม วิธีสร้างความสัมพันธ์และความไว้วางใจ การตอบสนองต่ออารมณ์ของผู้ป่วย และวิธีการปลอบใจ ประคองใจ และให้กำลังใจผู้ป่วยด้วยความเมตตากรุณา ตลอดจนวิธีการให้คำอธิบาย คำแนะนำและคำปรึกษาแก่ผู้ป่วย และที่สำคัญยิ่ง คือ เห็นผู้ป่วยเป็น "คน" ไม่ใช่เป็นเพียง "ไข้" และเมื่อเห็นผู้ป่วยเป็น "คน" ก็จะเห็นภาพสิ่งอื่น ๆ ที่เกี่ยวข้องเชื่อมโยงกับผู้ป่วยด้วย เช่น อารมณ์และความรู้สึก ครอบครัว หน้าที่การงาน ความเป็นอยู่ รวมถึงสังคมและวัฒนธรรม ซึ่งมีส่วนเกี่ยวข้องกับความเจ็บป่วยและผลการรักษาของผู้ป่วยทั้งสิ้น
2. ได้เรียนรู้วิธีการตรวจร่างกายที่ถูกต้องและแม่นยำ โดยการ "ดู คลำ เคาะ ฟัง"
3. ได้ฝึกทักษะดังกล่าวข้างต้น (การสัมภาษณ์ interpersonal skills การตรวจร่างกาย) โดยการฝึกปฏิบัติกับผู้ป่วย โดยมีครูผู้สอนเป็นผู้ชี้แนะและแก้ไขให้ถูกต้อง ทำให้เกิดความรู้ความสามารถจากการปฏิบัติและประสบการณ์

4. รู้จักรวบรวมข้อมูลจากประวัติ การตรวจร่างกาย การเลือกใช้ในการตรวจทางห้องปฏิบัติการ
5. ได้ฝึกทักษะในการวิเคราะห์ และสังเคราะห์ข้อมูลโดยการเชื่อมโยงและแยกแยะ รวมทั้งการแก้ปัญหาทางคลินิก
6. ได้ฝึกการนำเสนอผู้ป่วย
7. ได้เสริมสร้างเจตคติ จริยธรรมและคุณธรรม

จากที่กล่าวมานี้จะเห็นได้ว่าการสอนข้างเตียงเป็นการสอนทั้งการ “เป็นหมอ” และ “ความเป็นหมอ” แก่ นักศึกษา เพราะเป็นการสอนทักษะพื้นฐานทางคลินิก (basic clinical skills) รวมทั้งเจตคติ จริยธรรมและคุณธรรมที่สำคัญมากสำหรับนักศึกษา ซึ่งต้องการพัฒนาต่อไปตลอดชีวิตของการเป็นแพทย์ การสอนข้างเตียงเป็นโอกาสเดียวที่นักศึกษาจะได้เห็น “แบบอย่าง” การปฏิบัติทางคลินิก (clinical performance) ของครูแพทย์ ซึ่งจะมีอิทธิพลต่ออนาคตในการ “เป็นหมอ” และ “ความเป็นหมอ” ของนักศึกษา จึงนำที่ครูแพทย์จะให้ความสำคัญต่อการสอนข้างเตียง และพยายามเพิ่มพูน พัฒนาทักษะของตนในการสอนวิธีนี้อย่างต่อเนื่องตลอดชีวิตการเป็นครูแพทย์เช่นกัน

จะสอนข้างเตียงอย่างไร

เมื่อการสอนข้างเตียงเป็นหัวใจของการเรียนการสอนแพทย์ทางคลินิกดังกล่าวข้างต้นแล้ว ครูแพทย์จะสอนข้างเตียงอย่างไรให้มีประสิทธิภาพและประสิทธิผล ผู้เขียนใคร่ขอเสนอข้อคิดเห็นและวิธีการที่ได้จากการศึกษาและประสบการณ์ ดังนี้

หลักการ

- การสอนข้างเตียงควรต้องถือนักศึกษาเป็นศูนย์กลางในการสอน กล่าวคือ สอนสิ่งที่นักศึกษาควรรู้ และยังไม่รู้ มิใช่สอนสิ่งที่ครูชอบหรืออยากสอนเพียงเพื่อแสดงความเก่งของครู ควรสอนในสิ่งที่ไม่สามารถเรียนได้จากตำรา ครูจึงจำเป็นต้องทำความเข้าใจกับนักศึกษารู้และไม่รู้อะไร นักศึกษาคิดอะไรและอย่างไร ซึ่งทำได้โดยการถามและการพูดคุย ครูควรให้โอกาสนักศึกษาได้แสดงความคิดและศักยภาพของตนเองอย่างอิสระ รวมทั้งกระตุ้นให้คิดโดยการตั้งคำถามนำให้คิด

- ครูควรต้องเป็นแบบอย่างของแพทย์ที่ดี ทั้งในด้านความรู้ความสามารถทางวิชาการ เจตคติ และจริยธรรม
- การสอนข้างเตียง เป็นการสอนและการฝึกทักษะในการปฏิบัติต่อผู้ป่วย และการมีปฏิสัมพันธ์

(interaction) กับผู้ป่วย เช่น การสัมภาษณ์ การตรวจร่างกาย การเลือกการตรวจทางห้องปฏิบัติการ นอกจากนี้ นักศึกษายังสามารถฝึกวิธีการรวบรวมข้อมูล วิธีการคิดเชื่อมโยงและแยกแยะข้อมูล แล้วนำองค์ความรู้ต่างๆ มาประยุกต์ใช้ในการแก้ปัญหาของผู้ป่วย เพื่อนำไปสู่การวินิจฉัยที่ถูกต้องแม่นยำและการรักษาที่ได้ผล การสอนข้างเตียงมิใช่การบรรยายเนื้อหาวิชาที่ข้างเตียงผู้ป่วยโดยครูผู้สอน หรือการอภิปรายทฤษฎีต่างๆ ทางวิชาการที่ไม่เกี่ยวข้องกับผู้ป่วย จะเป็นที่น่าเสียดายโอกาสและเวลาอันมีค่าในการสอนข้างเตียงมาก หากครูผู้สอนดำเนินการสอนโดยการบรรยายถึงแต่ความรู้ตามตำรา ในขณะที่มีผู้ป่วยอยู่ด้วย และนักศึกษาไม่มีโอกาสแสดงทักษะทางคลินิกให้ครูตรวจสอบ เพื่อครูจะได้ให้คำแนะนำและชี้แนะวิธีการที่ถูกต้อง และนี่เป็นสาเหตุสำคัญอย่างหนึ่งที่ทำให้นักศึกษาขาดทักษะและความรู้ความสามารถทางคลินิก ซึ่งนำไปสู่การวินิจฉัยและการรักษาที่ไม่ถูกต้องและไม่ได้มาตรฐานดังที่ปรากฏอยู่ทั่วโลกในปัจจุบัน ทั้งนี้เพราะทักษะทางคลินิกนั้นไม่สามารถเรียนได้จากตำรา หรือ ผลการตรวจทางห้องปฏิบัติการต่างๆ

วิธีสอน

วิธีการสอนข้างเตียงทำได้หลายรูปแบบ อาจเป็นการสอนโดยมีผู้ป่วยเพียงหนึ่งคนหรือมีหลายคนก็ได้ในการสอนแต่ละครั้งครูควรกำหนดวัตถุประสงค์ให้ชัดเจน ว่าต้องการสอนหรือเน้นด้านใดบ้าง เพื่อจะได้เตรียมผู้ป่วยให้เหมาะสม วิธีที่ควรปฏิบัติ คือ ครูสอนแสดงให้นักศึกษาดูเป็นตัวอย่าง ตั้งแต่การทักทายผู้ป่วย การสัมภาษณ์ การตรวจร่างกาย และเปิดโอกาสให้นักศึกษาได้ทดลองปฏิบัติ โดยตลอดเวลาที่ทำการเรียนการสอนทั้งครูและนักศึกษาต้องคำนึงถึงความรู้สึกและความสบายของผู้ป่วย ตลอดจนให้เกียรติผู้ป่วยด้วย การสัมภาษณ์และการตรวจร่างกาย ควรกระทำอย่างสุภาพ นุ่มนวล ละมุนละม่อม ไม่ควรทำพร้อมกันที่ละหลายคน จนเป็นสภาพ “มະรุ่มมะตุ่ม” ซึ่งนอกจากเป็นการไม่ให้เกียรติแล้วยังทำให้ผู้ป่วยสับสนและอึดอัดด้วย ในการอธิบายข้างเตียง ควรให้ผู้ป่วยมีส่วนร่วม ควรใช้ถ้อยคำที่สุภาพ ไม่พูดในสิ่งที่จะทำให้ผู้ป่วยตื่นตระหนกหรือวิตกกังวล ควรอธิบายให้ผู้ป่วยเข้าใจถึงโรค การตรวจและการรักษา ครูควรสอบถามและกระตุ้นให้นักศึกษาแสดงความคิดเห็น และเปิดโอกาสให้ถาม พร้อมทั้งชี้แนะและแก้ไขทักษะและความเข้าใจที่ยังไม่ถูกต้อง โดยกระทำอย่างระมัดระวังไม่ให้ผู้ป่วยเกิดความรู้สึกที่ไม่ดีต่อนักศึกษา

ศาสตราจารย์ Ken Cox ที่เป็นปรมาจารย์ด้านแพทยศาสตรศึกษาคนหนึ่งของออสเตรเลีย ได้เสนอวิธีการสอนข้างเตียง ซึ่งตีพิมพ์ในวารสาร The Medical Journal of Australia ฉบับที่ 158 วันที่ 15 กุมภาพันธ์ ค.ศ. 1993 โดยมีขั้นตอนต่างๆ ซึ่งแบ่งได้เป็น 2 กลุ่ม ดังนี้

กลุ่มที่ 1 ประสบการณ์ทางคลินิก (clinical experience)

ขั้นตอนที่ 1 การเตรียมความพร้อม (preparation)

ครูต้องรู้ว่านักศึกษามีความพร้อมในการเรียนหรือไม่ และจะเตรียมความพร้อมให้นักศึกษาอย่างไร รวมทั้งนักศึกษาจะอะไรและทำอะไรได้แล้วบ้าง นักศึกษาควรมีพื้นฐานความรู้และทักษะทั่วไปก่อนที่จะดูแลผู้ป่วยอะไรบ้าง ครูจะต้องรู้จักและศึกษาหลักสูตรการเรียนการสอนอย่างดี และต้องรู้ระดับความรู้ความสามารถของนักศึกษาของตน

ขั้นตอนที่ 2 การให้คำแนะนำแก่นักศึกษาก่อนพบผู้ป่วย (briefing before seeing the patient)

ก่อนพบกับผู้ป่วย ครูควรแนะนำนักศึกษาก่อนว่า ผู้ป่วยและโรคของผู้ป่วยที่จะพบมีลักษณะใดที่น่าสนใจบ้าง รวมทั้งการปฏิบัติต่อผู้ป่วยอย่าง “คน” ที่เจ็บป่วย มิใช่เป็นเพียงตัวอย่างโรคหรือวัสดุประกอบการเรียนทางคลินิก การให้คำแนะนำแก่นักศึกษาก่อนพบผู้ป่วยนี้เป็นการวางแผนการสอนล่วงหน้า เพื่อจัดการเรียนการสอนให้สอดคล้องกับวัตถุประสงค์ที่ได้วางไว้ โดยครูสามารถบอกนักศึกษาได้ว่าควรจะเรียนอย่างไร มีโอกาสเรียนรู้อะไร และคาดหวังอะไรจากกิจกรรมในแต่ละครั้งได้บ้าง

ขั้นตอนที่ 3 การเรียนจากการฝึกปฏิบัติกับผู้ป่วย (clinical experience with the patient)

เมื่อเตรียมความพร้อมให้นักศึกษาแล้ว นักศึกษาจะได้พบกับผู้ป่วย ซึ่งเป็นส่วนที่มีความสำคัญที่สุดต่อการเรียนรู้ทางคลินิก กิจกรรมในขั้นนี้ประกอบด้วย การสัมภาษณ์และการพูดคุยเกี่ยวกับความเจ็บป่วยของผู้ป่วยกับผู้ป่วย การตรวจร่างกาย การแปลผลข้อมูลที่มีอยู่ และการวางแผนการตรวจรักษา ซึ่งผู้ป่วยต้องมีส่วนร่วมในการแสดงความคิดเห็นเช่นกัน ความรู้ที่เกิดขึ้นเป็นความรู้จากประสบการณ์จริง ซึ่งจะมีผลต่อเนื่องตลอดชีวิตการเป็นแพทย์ของนักศึกษา

ขั้นตอนที่ 4 การให้คำแนะนำแก่นักศึกษาหลังได้พบผู้ป่วย (debriefing after leaving the patient)

หลังจากตรวจผู้ป่วยเสร็จแล้ว ครูควรทบทวนสิ่งที่ได้ประสบมากับนักศึกษาอีกครั้งว่า นักศึกษาได้เห็น ฟัง และสัมผัสอะไรบ้าง จะแปลผลข้อมูลต่างๆ ที่รวบรวมได้อย่างไร และได้เรียนรู้อะไรจากผู้ป่วยบ้าง

ครูพึงตระหนักว่านักศึกษาแต่ละคนมีความสามารถในการเรียนรู้ การสังเกต และการจับประเด็นสำคัญได้แตกต่างกัน ขั้นตอนนี้ครูจะมีโอกาสรู้ว่านักศึกษารู้และไม่รู้อะไร เข้าใจหรือไม่เข้าใจในสิ่งที่พบเห็นเกี่ยวกับผู้ป่วย และครูสามารถเสริมสร้างความรู้หรืออธิบายให้นักศึกษาเข้าใจมากขึ้น นอกจากนั้นในการอธิบายร่วมกันในหมู่นักศึกษากับครู

นักศึกษาจะได้เรียนจากเพื่อนด้วย หลังจากนั้นครูและนักศึกษาควรร่วมกันเชื่อมโยง/แยกแยะข้อมูลที่รวบรวมได้ แล้วสรุปให้เห็นภาพและปัญหาของโรคอย่างชัดเจน

จากขั้นตอนนี้อาจเตรียมนักศึกษาสำหรับการสอนข้างเตียงในผู้ป่วยคนต่อไป หรือสอนต่อในขั้นตอนที่ 5

กลุ่มที่ 2 การอธิบายเหตุการณ์ทางคลินิกที่ได้ประสบ (explanation of the clinical events)

ขั้นตอนที่ 5 การทบทวนประสบการณ์ที่ได้รับ (reflection)

ขั้นตอนนี้เป็นกาให้นักศึกษาทบทวนสิ่งที่ได้พบได้เห็นและได้ฟังมา เพื่อเชื่อมโยงกับประสบการณ์เดิมที่มีอยู่ และเพื่อค้นหาว่าตนเองรู้และไม่รู้อะไร เป็นการสำรวจตนเอง ทำให้เกิดการเรียนรู้เพิ่มขึ้นอีก ทั้งนี้เพราะแม้จะได้เรียนจากประสบการณ์ตรง ซึ่งเป็นวิธีเรียนรู้ที่ดีที่สุด แต่ก็ยังไม่เพียงพอ ยังควรต้องนำความรู้ที่ได้รับใหม่มาเชื่อมโยงกับความรู้เดิมที่มีอยู่ เพื่อให้เกิดความรู้ความเข้าใจเพิ่มขึ้นและลึกซึ้งขึ้น ขั้นตอนนี้แตกต่างจากขั้นตอนที่ 4 ที่ในขั้นตอนที่ 4 นั้น เป็นการเชื่อมโยงสิ่งที่พบและรู้ว่าอะไรเกิดขึ้นบ้าง ส่วนขั้นตอนที่ 5 นี้ เป็นการทำให้รู้และเข้าใจความหมายของสิ่งที่ได้พบ

ขั้นตอนที่ 6 การทำความเข้าใจกับประสบการณ์ที่ได้รับ (explication of the experience)

ขั้นตอนนี้เป็นการรวม (incorporate) ประสบการณ์ที่ตนเองได้รับเข้ากับของผู้อื่น ซึ่งอาจเป็นสิ่งที่ได้จากการสังเกต การได้เรียนรู้ หรืออาจเป็นการศึกษาจากตำรา หรือจากผู้ป่วยรายอื่น หรือจากทฤษฎีและงานวิจัยที่มีอยู่และเกี่ยวกับข้อกับผู้ป่วยที่เพิ่งศึกษามา ซึ่งอาจเป็นความรู้ด้าน ชีวเคมี สรีรวิทยา พยาธิวิทยา จิตเวชศาสตร์หรือสังคมศาสตร์ก็ได้ ขั้นตอนนี้จะทำให้นักศึกษารู้กว้างและรู้สึก นักศึกษาจะได้ใช้ภูมิปัญญาอย่างเต็มที่ในการรวบรวมเหตุปัจจัยต่างๆ ที่ได้จากการปฏิสัมพันธ์กับผู้ป่วย ครู และเพื่อนนักศึกษาในการสอนข้างเตียง ทำให้เข้าใจการเจ็บป่วยและโรคของผู้ป่วยอย่างชัดเจน ซึ่งจะนำไปสู่การตรวจและการรักษาที่ถูกต้องและเหมาะสม

ขั้นตอนที่ 7 การวางแผนงานและแผนการเรียนจากประสบการณ์ที่ได้รับ

เมื่อเข้าใจเหตุการณ์และปัญหาของผู้ป่วยและโรคอย่างถ่องแท้แล้ว สิ่งที่นักศึกษาควรทำ คือ ถามตนเองว่า จะต้องทำอะไรและอย่างไรต่อไป สิ่งที่ได้คิดและได้ทำไปแล้วดีหรือยัง สามารถทำอย่างอื่นที่ดีกว่าได้หรือไม่จากประสบการณ์ที่ได้รับ นักศึกษาควรนำมาวางแผนงานและแผนการเรียนของตนเองต่อไป สิ่งปฏิบัติเหล่านี้จะเป็นรากฐานในการปฏิบัติทางคลินิกและในการเป็นผู้ประกอบวิชาชีพเวชกรรมที่มีความสามารถต่อไป

ครูมีหน้าที่สำคัญในขั้นตอนนี้ โดยช่วยกระตุ้นหรือชี้แนะให้นักศึกษาคิดว่า ถ้าพบปัญหาเช่นนี้อีกจะทำอย่างไรอะไรเป็นข้อมูลที่สำคัญและมีความหมายมากที่สุดที่จะต้องมองหา และควรมองหาอะไรก่อน โดยนักศึกษาอาจทำแผนกลยุทธ์ในการปฏิบัติหรือแผนผังการปฏิบัติ (flow chart) หรือเลือกว่าควรจะตรวจหาอะไรก่อนหลัง หรืออาจทำแนวทางการตัดสินใจ สำหรับใช้ในครั้งต่อไป เมื่อพบผู้ป่วยหรือปัญหาเช่นเดียวกันนี้อีก เป็นต้น

ขั้นตอนที่ 8 การเตรียมตัวสำหรับการดูแลผู้ป่วยคนต่อไป (preparation for the next patient)

ในขั้นตอนนี้ นักศึกษาต้องขอความช่วยเหลือหาความรู้ที่ยังขาดอยู่เพิ่มเติม เพื่อเตรียมพร้อมในการดูแลผู้ป่วยในครั้งต่อไป

จะเห็นว่าวิธีการของศาสตราจารย์ Ken Cox นั้น มีการวางแผนก่อนเริ่มต้นทำ และเมื่อทำแล้วให้หยุดและคิดวิเคราะห์ในสิ่งที่ได้ทำไป เพื่อนำมาเป็นแนวทางในการวางแผนในการทำครั้งต่อไป หากครูแพทย์จะได้พิจารณาวิธีการดังกล่าวและนำมาประยุกต์ใช้ก็จะเป็นประโยชน์ต่อการสอนข้างเตียงอย่างยิ่ง โดยอาจปรับเวลาให้เหมาะสม ในความเห็นของผู้เขียน วิธีการสอนข้างเตียงดังกล่าวไม่ยากเกินศักยภาพของครูแพทย์ไทยและสามารถทำให้เกิดขึ้นได้จริง หากครูแพทย์ให้ความสนใจและบริหารจัดการเวลาเป็น และหน้าที่ครูแพทย์ทั้งหลายจะได้ช่วยกันพยายามทำให้เกิดขึ้น เพราะ “A teacher..... can never tell where his influence stops” (โดย Henry Adams) และนี่เป็นโอกาสที่ครูแพทย์จะได้หล่อ

หลอมนักศึกษาแพทย์ให้เป็นแพทย์ที่ถึงพร้อมซึ่งวิชาความรู้ด้านวิชาชีพและจริยธรรมคุณธรรม เพื่อจรรโลงวิชาชีพแพทย์และสังคมไทยต่อไป

เอกสารอ่านเพิ่มเติม

1. Cox KR, Ewan CE. The medical teacher. Edinburgh: Churchill Livingstone; 1982.
2. Cox K. Planning bedside teaching – 2. Preparation before entering the wards. Med J Australia 1993; 158: 355 – 357.
3. Cox K. Planning bedside teaching – 4. Teaching around the patient. Med J Australia 1993; 158: 493 – 495.

เอกสารประกอบการอบรม

The word "TEACH" is written in large, bold, dark blue capital letters with a white outline. It is centered on a light blue background that features a white ECG (heart rate) line. The background also has a subtle pattern of light blue squares and some glowing white circles.

TEACH

8 November 2019

การสอนเจตคติและจรรยาวิชาชีพ (Teaching Attitude and Professionalism)

สุพจน์ พงศ์ประสพชัย

บทนำ

เจตคติ (attitude) และจรรยาวิชาชีพ (professionalism) เป็นองค์ประกอบ 1 ใน 3 ของการเป็นวิชาชีพใดๆ ที่ดีงามและพึงปรารถนาอันได้แก่ มีความรู้ดี มีทักษะดี และมีเจตคติดี ดูเหมือนว่าเจตคติเป็นด้านที่สำคัญที่สุดใน 3 ด้าน เนื่องจากเจตคติเป็นเข็มทิศที่จะชี้นำบุคคลให้หาความรู้และทักษะไปในทางที่ถูกต้อง ดีงาม เพื่อประโยชน์ส่วนรวม บุคคลที่มีความรู้และทักษะดีแต่เจตคติไม่ดีก็ไม่ต่างจากมหาโจร ที่จะก่อให้เกิดความเลวร้ายได้อย่างมากมาย เนื่องจากการมีความรู้และทักษะที่ดีมาก

ในอดีตเรื่องเจตคติและจรรยาวิชาชีพมักไม่มีการสอนอย่างเป็นทางการเป็นเรื่องราว มักอาศัยการเรียนรู้จากประสบการณ์การทำงาน การได้ใกล้ชิดได้สัมผัสกับตัวอย่างดีๆ (role model) จากครูผู้ศิษย์โดยตรง วิธีนี้ยืนยันมาได้หลายพันปี แต่ผลที่เห็นในปัจจุบันคือ วงการแพทย์เริ่มพบแพทย์ที่มีปัญหาทางเจตคติและจรรยาวิชาชีพมากขึ้นและรุนแรงขึ้นเรื่อยๆ จนเกิดความเคลือบแคลงและมีเสียงสะท้อนจากสังคม ซึ่งปฏิภานี่สำคัญยิ่ง เพราะความเป็นแพทย์นั้นเป็นอภิสิทธิ์ (privilege) ที่สังคมมอบให้แพทย์ด้วยความเต็มใจ แต่ก็ไม่ใช่อภิสิทธิ์ (right) ของแพทย์ ดังนั้นสังคมก็มีสิทธิ์ทวงคืนกลับไปได้ตลอดเวลาหากแพทย์ประพฤติตนไม่เหมาะสม ดังนั้นวงการแพทย์จึงควรเหลียวกลับมามองว่า เราคงต้องมีการสอนเจตคติและจรรยาวิชาชีพอย่างจริงจังมากกว่านี้ เพราะการสอนโดยรูปแบบ role model เพียงอย่างเดียวอาจไม่เพียงพออีกต่อไป

อุปสรรคของการสอนเจตคติและจรรยาวิชาชีพ

แม้เจตคติและจรรยาวิชาชีพจะมีความสำคัญที่สุด แต่กลับเป็นด้านที่ครูให้ความสำคัญในการสอนผู้เรียนน้อยที่สุด ครูมักใช้เวลาส่วนใหญ่ในการสอน

ความรู้และทักษะเป็นหลัก จึงเหมือนว่าครูได้สอน “อาวุธ” อย่างครบครันแก่ผู้เรียน แต่หลงลืมให้คุณธรรมในการนำอาวุธนั้นไปใช้ อุปสรรคที่ทำให้ครูสอนเจตคติและจรรยาวิชาชีพแก่ผู้เรียนน้อยมีหลายประการ

(ตารางที่ 1)

ตารางที่ 1 อุปสรรคของการสอนเจตคติและจรรยาวิชาชีพ

- ไม่มั่นใจในนิยามหรือเกณฑ์
- ประเมินและวัดได้ยาก
- กลัวผู้เรียนเบื่อ ไม่สนใจ
- การสอนด้วยวิธีธรรมดาไม่ได้ผล
- ไม่ทราบวิธีการสอน
- รู้สึกว่าตนเองยังไม่ดีพอ
- รู้สึกว่าไม่ใช่ธุระของตน

ไม่มั่นใจในนิยามหรือเกณฑ์

เกณฑ์การบอกว่าเจตคติดีหรือไม่ดีบางครั้งเป็นเรื่องยาก เพราะบางครั้งเกณฑ์อาจคลุมเครือ หรือแตกต่างกันในแต่ละสังคม เช่น แพทย์อาจถือว่าการใส่กระโปรงสั้นมาก รัตรูป ไม่ถูกต้อง แต่สิ่งเดียวกันเป็นที่ยอมรับในวิชาชีพอื่น เป็นต้น เกณฑ์บางอย่างอาจแปรเปลี่ยนไปตามกาลเวลา ตามการยอมรับ แต่ต้องคำนึงอยู่เสมอว่า สิ่งที่สังคมยอมรับ ไม่จำเป็นต้องเป็นสิ่งที่จะต้องเสมอไป เช่น การคอร์รัปชัน การโกหก ค่านิยมผิดๆ ในวัยรุ่น เป็นต้น นอกจากนี้เกณฑ์บางอย่างก็อาจเริ่มหย่อนยานหรือเลือนเมื่อเวลาผ่านไปหรือสังคมเปลี่ยนไป

การแก้ปัญหาเหล่านี้คือ กลุ่มวิชาชีพต่างๆ ควรกำหนดเกณฑ์ทางเจตคติหรือจรรยาวิชาชีพของตนให้ชัดเจน เพื่อเป็นเกณฑ์บรรทัดฐาน ตัวอย่างเกณฑ์จรรยาวิชาชีพที่สมาคมแพทย์นานาชาติขณะนี้อ้างอิงถึงเป็นส่วนใหญ่ ได้แก่ เกณฑ์ที่เป็นการร่วมกำหนดกันโดย American Board of Internal Medicine, American College of Physicians และ European Federation of Internal Medicine ค.ศ. 2002¹ แต่ละสถาบันก็ควร

กำหนดนิยามของเจตคติและจรรยาวิชาชีพของตนให้ชัดเจนและแจ้งให้ทราบโดยทั่วกัน

ประเมินและวัดได้ยาก

การประเมินวัดผลทางเจตคติและจรรยาวิชาชีพไม่สามารถใช้การสอบข้อเขียน หรือสอบปฏิบัติได้ เพราะผู้เรียนที่เจตคติไม่ดี สามารถเสแสร้งให้ดูมีเจตคติดีได้ และมักตอบข้อสอบได้ดี การวัดผลจึงต้องทำโดยผู้เรียนไม่รู้ตัว เช่น ใช้การสังเกตจากครู เพื่อนร่วมงาน หรือผู้ป่วย (การประเมิน 360 องศา) เป็นสำคัญ ในต่างประเทศอาจมีการประเมินโดยใช้ผู้ป่วยจำลองที่ได้รับการฝึกมาเป็นอย่างดีมีปฏิสัมพันธ์กับแพทย์โดยที่แพทย์ไม่รู้ตัว แล้วจึงทำการประเมินแพทย์ เป็นต้น

กลัวผู้เรียนเบื่อ ไม่สนใจ

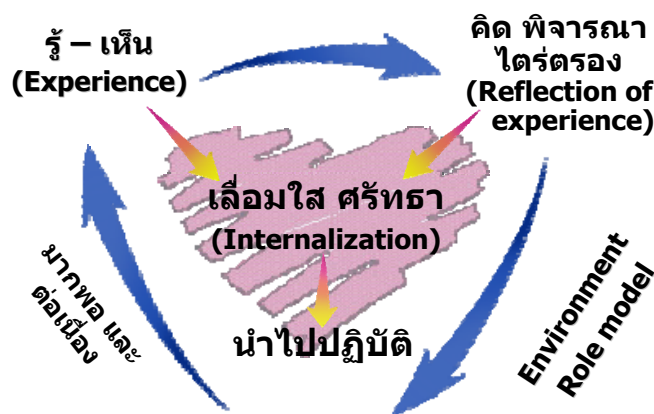
ผู้เรียนอาจไม่ค่อยสนใจจริงๆ อาจเป็นเพราะไม่เห็นความสำคัญ ไม่เคยได้คิดทบทวนให้เห็นความสำคัญของเรื่องนี้อย่างจริงจัง ไม่มีการสอบหรือประเมินผล ซึ่งถ้าให้มีการประเมินและสอบเรื่องเจตคติและจรรยาวิชาชีพอย่างจริงจัง ผู้เรียนก็จะให้ความสนใจมากขึ้นเอง แต่ที่สำคัญคือ ครูเองก็ไม่ค่อยคิดจะสอนด้วย เพราะครูอาจรู้สึกว่าสอนเรื่องนี้แล้วจะน่าเบื่อ ไม่น่าสนใจ ซึ่งไม่เป็นความจริงนัก การสอนเจตคติและจรรยาวิชาชีพเป็นเรื่องดีงาม อาจมีความจริงจัง อาจมีความซาบซึ้ง สะเทือนใจ แต่ไม่ใช่ต้องเครียด การสอนเรื่องนี้ไม่จำเป็นต้องน่าเบื่อได้

เป็นเรื่องที่เป็นนามธรรม การสอนด้วยวิธีธรรมดาไม่ได้ผล

เจตคติและจรรยาวิชาชีพเป็นนามธรรมจริง และเป็นการสอนที่ “ใจ” ไม่ใช่ที่ “สมอง” แบบการสอนความรู้หรือทักษะ แต่เป็นเรื่องที่สอนได้² การสอนแบบบรรยายหรือให้ความรู้ความเข้าใจในเรื่องเจตคติ (cognitive-based professionalism) อย่างเดียวไม่เพียงพอ แต่ก็ยังมีความจำเป็นต้องมี^{1, 3, 4} เพราะเป็นการทำให้ผู้เรียนทุกคนรับทราบการมีอยู่ของมัน ความสำคัญ ที่มา เหตุผล และข้อบังคับต่างๆ ที่มีอยู่ และควรพูดเป็นครั้งแรกในการสอนเจตคติและจรรยาวิชาชีพ^{5, 6}

การสอนเจตคติและจรรยาวิชาชีพที่มีประสิทธิภาพไม่ใช่การสอนในห้องเรียน แต่ต้องใช้การสอนแบบ situated หรือ experiential learning⁷ ซึ่งประกอบด้วย

1. ต้องให้ผู้เรียนได้ประสบการณ์ (experience) ไม่ว่าจะเป็นการได้เห็น ได้ฟัง หรือประสบด้วยตัวเอง
2. ผู้เรียนต้องได้คิดวิเคราะห์ ไตร่ตรอง และสะท้อนความรู้สึกในเรื่องนั้นๆ (reflection of experience) ขั้นตอนนี้เป็นหัวใจสำคัญในการเปลี่ยนประสบการณ์ที่ได้รับให้กลายเป็นจิตสำนึกที่ยั่งยืน
3. ต้องประสบซ้ำๆ อย่างมากพอและต่อเนื่อง (repetitive-continuous) โดยมีสภาพแวดล้อมที่เอื้อ และมีตัวอย่าง (role model) ที่มากพอและต่อเนื่อง
4. ผู้เรียนเชื่อ ศรัทธา และนำสิ่งนั้นเข้าไปในใจ (internalization) และนำไปเปลี่ยนพฤติกรรมของตน^{5, 6} (ภาพที่ 1)



ภาพที่ 1 หลักการสอนเจตคติโดยวิธี experiential learning

รู้สึกว่าคุณเองยังไม่ดีพอ

ครูจำนวนมากไม่กล้าสอนเจตคติและจรรยาวิชาชีพเพราะคิดว่าตนเองไม่ดีพอ ไม่มีใครดีพร้อมทุกด้าน แต่ทุกคนสามารถสอนเจตคติที่ดีๆ และเป็น role model แก่ผู้เรียนได้เสมอ (role model หมายถึง บุคคลคนหนึ่ง ที่สิ่งที่ตนแสดงออก กลายเป็นแบบอย่างให้กับอีกคนหนึ่ง) โดยเฉพาะด้านที่ตนทำได้และเป็นตัวอย่างที่ดีได้ ไม่มีใครรู้ตัวว่าตนกำลังเป็น role model ให้ใครอยู่ เพราะผู้ที่เป็นคนเลือก role model คือผู้เรียน ไม่ใช่ครู ครูจึงควรทำตนให้ดีที่สุดเท่าที่จะทำได้ ถ้าครูทำตัวเป็นตัวอย่างที่ไม่ดี ครูก็อาจกำลังเป็น role model ปลุกฝังสิ่งที่ไม่ดีนั้นแก่ผู้เรียนหลายๆ คนได้เช่นกัน⁸

รู้สึกว่าไม่ใช่ฐานะของตน

การสอนเจตคติและจรรยาวิชาชีพไม่ใช่ฐานะของใครคนหนึ่ง ไม่ใช่หน้าที่ของแพทย์อาวุโส ไม่ใช่หน้าที่ของครูวิชาจริยธรรม ยิ่งสอนหลายคน ยิ่งแสดงตัวอย่างให้ดูหลายคน ยิ่งได้ผล ยิ่งมีประสิทธิภาพ เพราะจะทำให้บรรยากาศอบอุ่นไปด้วยตัวอย่างที่ดีๆ ครูอายุน้อยอาจสอนเจตคติได้ดีกว่าครูอาวุโส ด้วยวัยที่ไม่ต่างกันมากทำให้มีความใกล้ชิดสนิทสนมกับผู้เรียนได้มากกว่า ผู้เรียนอาจเชื่อมากกว่า ควรสอนเจตคติให้รู้สึกเหมือนพี่สอนน้อง พี่รักน้องพี่ก็จะสอนน้อง ถ้าเห็นน้องทำตัวไม่ดี ผู้เป็นพี่ก็จะไม่ลังเลที่จะวักกล่าวตักเตือนน้อง

การจัดการสอนเจตคติและจรรยาวิชาชีพ

มีการจัดได้ 2 แบบคือ

1. Formal curriculum คือ จัดเป็นส่วนหนึ่งของหลักสูตรอย่างชัดเจน ได้แก่

1.1. สถาบันต้องให้ความสำคัญกับเรื่องนี้ อย่างเต็มที่ และประสานเสียงกันในทุกส่วน มิใช่ต่างคนต่างสอน

1.2. กำหนดนิยามหรือกฎเกณฑ์ทางจริยธรรมวิชาชีพให้ชัดเจน และมีการสอนหรือแจ้งให้ทราบอย่างชัดเจน^{1, 5, 6}

1.3. มีการวัดผลและประเมินผล

1.4. มีบทบาทปฏิบัติ ตักเตือน หรือลงโทษผู้ที่มีปัญหาทางเจตคติและจรรยาวิชาชีพ ไม่ว่าจะเป็นผู้เรียนหรือแม้แต่ครูเอง

1.5. การเรียนอภิปรายกลุ่มย่อย เช่น อภิปรายปัญหาทางจริยธรรมจากกรณีศึกษา⁹ จากข่าวในหนังสือพิมพ์ การแสดง (role play), interactive virtual patient¹⁰ และจากศิลปะหรือภาพยนตร์¹¹ เป็นต้น โดยหัวใจสำคัญคือ ต้องให้ผู้เรียนได้สะท้อนความรู้สึก (reflection) ออกมาด้วยตนเอง โดยมีครูคอยชี้แนะ แต่มิใช่บอกกล่าวหรือชี้หน้า

1.6. การจัดกิจกรรมพิเศษ เช่น กิจกรรมให้ผู้เรียนได้พูดคุยสอบถามกับผู้ป่วยถึงประเด็นทางจริยธรรมของแพทย์ ให้ผู้เรียนพาผู้ป่วยไปตรวจผู้ป่วยนอกแบบชีวิตจริง เพื่อจะได้เข้าใจความยากลำบากของผู้ป่วย เป็นต้น

1.7. หนังสืออ่านนอกเวลา ที่ให้แรงบันดาลใจด้านเจตคติและจรรยาวิชาชีพ¹¹

1.8. หนังสือประวัติศาสตร์ทางการแพทย์ การให้ผู้เรียนได้ศึกษาประวัติศาสตร์ของวิชาชีพแพทย์² จะช่วยให้ผู้เรียนได้ระลึก ตระหนักถึงเกียรติภูมิ และจรรยาแพทย์ที่บูรพคณาจารย์ได้พยายามผดุงไว้ยิ่งกว่าชีวิตจนมาถึงแพทย์รุ่นเรา

2. Hidden curriculum เป็นการสอนที่ไม่ได้อยู่ในหลักสูตรอย่างเป็นทางการ แต่แทรกซึมอยู่ในหลายๆ ที่ หลายๆ เวลา แต่เป็นวิธีที่ได้ผลยิ่งกว่า formal curriculum ตัวอย่างเช่น

2.1. การให้ผู้เรียนได้ประสบ ได้เห็น ได้ฟัง (experience) ประเด็นทางเจตคติและจรรยาวิชาชีพเมื่อมีโอกาสทอง (teachable moment) ช่วงใดก็ตามในระหว่างปฏิบัติงาน โดยเปิดโอกาสให้ผู้เรียนได้สะท้อนความรู้สึก (reflection) โดยมีครูหรือรุ่นพี่เป็นผู้ชี้แนะ

2.2. การจัดสิ่งแวดล้อม (environment) ให้เอื้อให้คนเกิดกำลังใจที่จะทำสิ่งดีงามและเกรงกลัวต่อการทำไม่ดี เช่น มีบรรยากาศ มีกิจกรรม มีพิธีการ (เช่น การกล่าวสัตย์ปฏิญาณตน, white coat ceremony เป็นต้น) มีการยกย่องคนดี (เช่น การกล่าวยกย่อง การมอบรางวัลแพทย์ดีเด่น เป็นต้น)

2.3. เรื่องเล่าสู่กันฟัง (parable) ได้แก่ เรื่องดีๆ เรื่องที่สร้างสรรค์ ข้อคิด อุทาหรณ์ต่างๆ ให้ฟังกัน จากครูสู่ศิษย์ จากที่สู่น้องระหว่างการเรียนหรือการทำงาน

2.4. การทำให้ดู (role modeling) เป็นวิธีที่สำคัญที่สุดและทรงประสิทธิภาพที่สุดในการสอนเจตคติและจรรยาวิชาชีพและสามารถเปลี่ยนชีวิตของแพทย์คนหนึ่งๆได้เลย อย่างไรก็ตามการสอนโดยเป็น role model ให้ดูอย่างเดียวยังไม่เพียงพอ แต่ควรร่วมกับให้ผู้เรียนได้สะท้อนความรู้สึกต่อบทบาทที่ผู้เรียนเห็นด้วย¹² ครูทุกคนสามารถเป็น role model ได้ จึงควรให้กำลังใจให้แรงบันดาลใจแก่กันและกันในหมู่ครูที่จะช่วยกันประพฤติดีๆ ทำสิ่งดีๆ เพื่อเป็น role model ที่ดีแก่ลูกศิษย์อย่างพร้อมหน้ากัน ไม่ต้องวิตกว่า เราไม่ดีพอ ไม่ต้องวิตกว่าเราไม่มีเวลา เพราะปัจจัยที่ทำให้ครูคนหนึ่งกลายเป็น role model ของผู้เรียนคนหนึ่งนั้นไม่ขึ้นกับระยะเวลา (duration) ที่ทั้งสองได้พบหรือมีปฏิสัมพันธ์กัน แต่อยู่ที่การมีช่วงเวลาร่วมกันที่มีคุณค่าและมีความประทับใจเป็นสำคัญ

สรุป

เจตคติและจรรยาวิชาชีพเป็นเรื่องที่ครูต้องสอนศิษย์ เป็นหน้าที่ของครูทุกคนและทุกคนสามารถสอนได้ วิธีการสอนเจตคติและจรรยาวิชาชีพ ได้แก่ การกำหนดนิยามหรือเกณฑ์ที่ชัดเจน มีการสอนหรือแจ้งให้ทราบอย่างชัดเจน แต่การสอนที่สำคัญที่สุดคือ การให้ผู้เรียนได้ประสบการณ์ไม่ว่าด้วยการเห็น ฟัง หรือประสบโดยตรง ได้สะท้อนความรู้สึก ต่อเนื่องซ้ำแล้วซ้ำอีกจนเข้าใจและนำไปปฏิบัติ วิธีการสอนที่มีพลังที่สุดคือการเป็น role model

The mediocre teacher tells.

The good teacher explains.

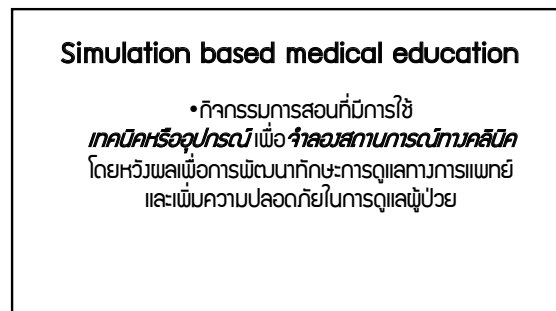
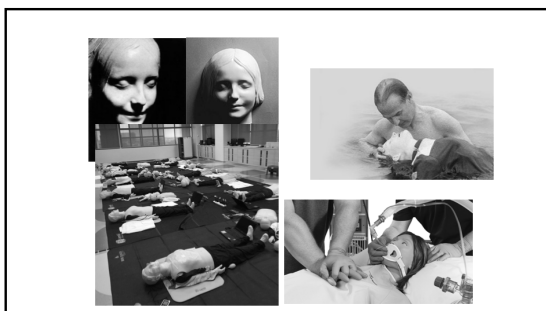
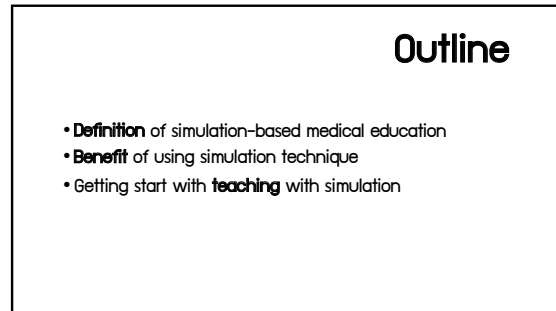
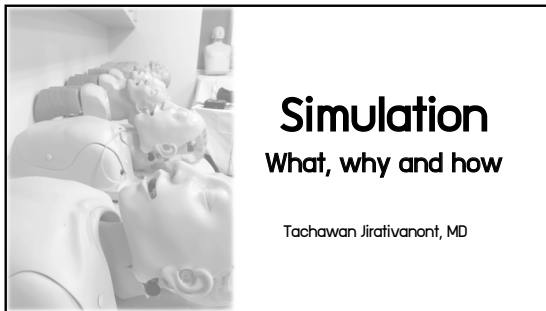
The superior teacher demonstrates.

The great teacher inspires.

William Arthur Ward


เอกสารอ้างอิง

1. Medical professionalism in the new millennium: a physician charter. *Ann Intern Med* 2002;136:243-6.
2. Collier R. Professionalism: can it be taught? *CMAJ* 2012;184:1234-6.
3. Cruess SR, Cruess RL. Professionalism must be taught. *BMJ* 1997;315:1674-7.
4. Swick HM. Toward a normative definition of medical professionalism. *Acad Med* 2000;75:612-6.
5. Cruess RL. Teaching professionalism: theory, principles, and practices. *Clin Orthop Relat Res* 2006;449:177-85.
6. Cruess RL, Cruess SR. Teaching professionalism: general principles. *Med Teach* 2006;28:205-8.
7. Maudsley G, Strivens J. Promoting professional knowledge, experiential learning and critical thinking for medical students. *Med Educ* 2000;34:535-44.
8. Feudtner C, Christakis DA, Christakis NA. Do clinical clerks suffer ethical erosion? Students' perceptions of their ethical environment and personal development. *Acad Med* 1994;69:670-9.
9. Gunderman RB, Brown BP. Teaching professionalism through case studies. *Acad Radiol* 2013;20:1183-5.
10. McEvoy M, Butler B, MacCarrick G. Teaching professionalism through virtual means. *Clin Teach* 2012;9:32-6.
11. Charon R. The patient-physician relationship. Narrative medicine: a model for empathy, reflection, profession, and trust. *JAMA* 2001;286:1897-902.
12. Stern DT, Papadakis M. The developing physician--becoming a professional. *N Engl J Med* 2006;355:1794-9.



Healthcare Simulation

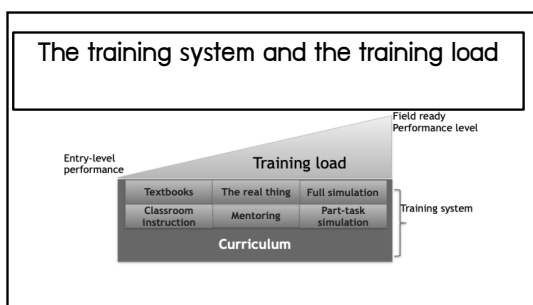
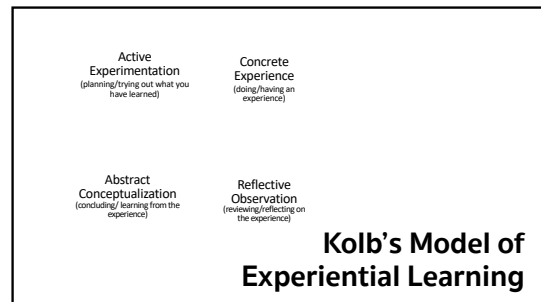
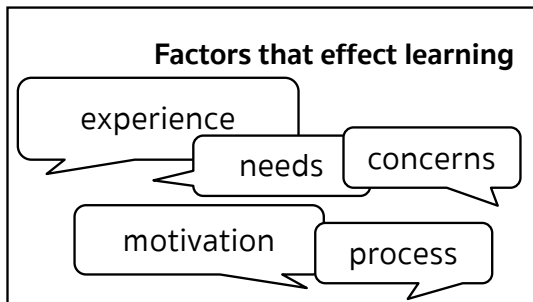
Simulation is a *technique* –not a technology– to replace or amplify real experiences with *guided experiences*.



David Gaba

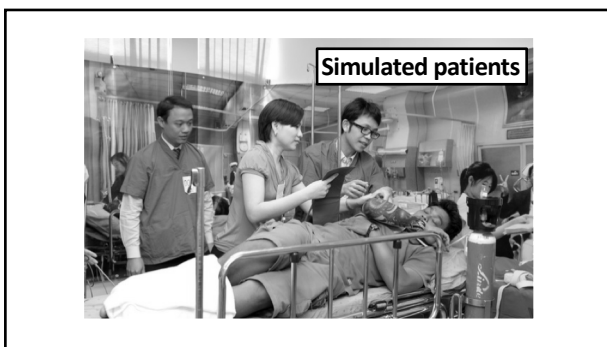
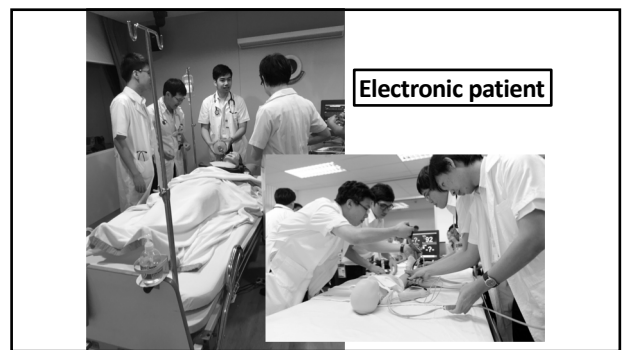
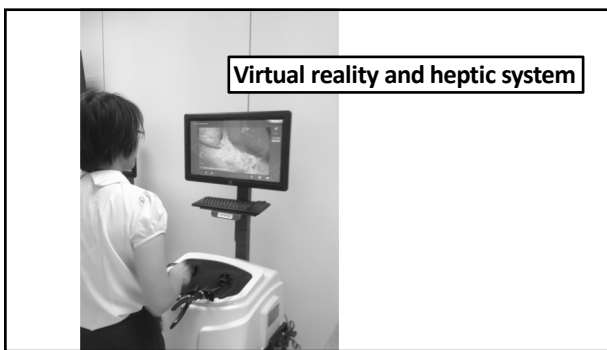
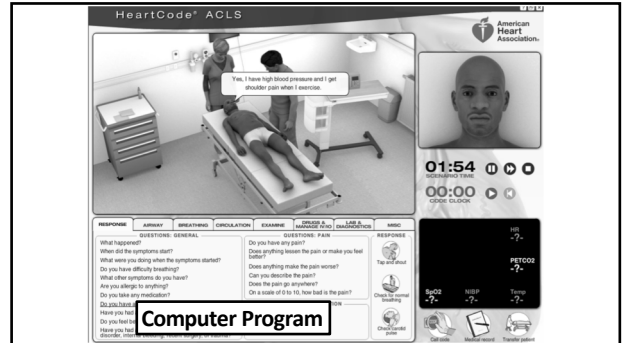
ทำไมต้อง simulation

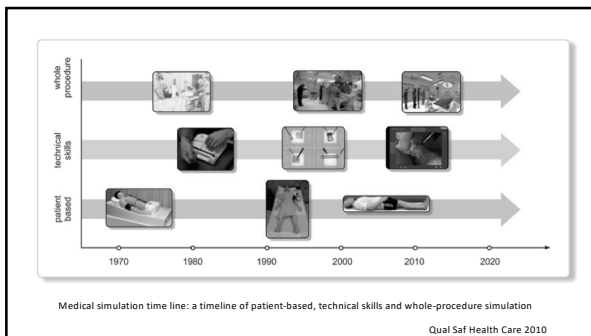
- พวกเราเรียนรู้กันแบบผู้ใหญ่ ... *adult learning*
- เราต้องการสถานการณ์ที่ปลอดภัยและควบคุมได้เพื่อการเรียนรู้



Type of Simulation Technology

- Part-task trainer
- Computer program
- Virtual reality and heptic system
- Standardized patient
- Electronic patient





Big Question!!!

When and how to use simulation most effectively and cost-efficiently.

When should we consider simulation?

- The decision to use simulation should be based on
 - Acuity:** the potential severity of an event and their subsequent impact on the patient.
 - Opportunity:** the frequency in which persons are actively involved in the management of the event

Christina G. Med Teach. 2012. e1-e15

Zone of simulation matrix

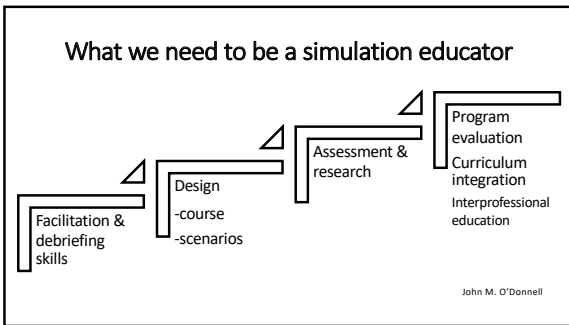
The zone of simulation : situations where healthcare simulation may be **advantageous** over other instructional media

Component of An Effective Simulation Session

- An introduction or briefing
- The simulation session
- The debriefing
- An evaluation

ขอให้ท่านสังเกต VDO ต่อไปนี้ แล้วพิจารณาว่า หากท่านต้องสอนในลักษณะเดียวกัน

- บทบาทของท่านคืออะไร
- ท่านต้องการอะไรในการสอนลักษณะนี้



How to design a simulation event

- Purpose :**
 - Objective of the course
- Participant :**
 - Level of learners' experience
- Process :**
 - Simulator chosen
 - Scenario development
 - Simulation class

Adapted from: Swanwick T, Understanding Medical Education, 2011

Scenario objectives: hypotensive baby

- เมื่อสิ้นสุดการทำ simulation ผู้เรียนต้องสามารถ

- วินิจฉัยภาวะ shock ในเด็กทารกได้
- วิเคราะห์สาเหตุของภาวะ shock ในเด็กได้
- สามารถดูแลเบื้องต้นในผู้ป่วยเด็กที่มีภาวะ hypovolemic shock ได้
- ทราบข้อบ่งชี้และข้อจำกัดในการทำ intraosseous infusion
- แสดงทักษะการให้สารน้ำด้วยวิธี intraosseous infusion ได้
- สามารถให้การดูแล intraosseous infusion และระมัดระวัง complication ได้

How to design a simulation event

- Purpose :**
 - Objective of the course
- Participant :**
 - Level of learners' experience
- Process :**
 - Simulator chosen
 - Scenario development
 - Simulation class

นักศึกษาแพทย์ชั้นปีที่ 6
 ทารก (newborn) (สามารถแสดง vital sign ได้),
 ทนสำหรับรับการทำ intraosseous

Adapted from: Swanwick T, Understanding Medical Education, 2011

การเตรียมผู้ป่วยจำลอง

พยาบาล
 เป็นพยาบาลประจำห้องฉุกเฉิน มีหน้าที่ให้ความช่วยเหลือ เตรียมอุปกรณ์ ให้ข้อมูลที่จำเป็นเพิ่มเติม สามารถแนะนำการรักษาและบอกข้อมูลเพื่อให้สถานการณ์ดำเนินไป

ญาติผู้ป่วย (สามารถให้อ.ผู้คุมเครื่องเป็นผู้พากย์เสียง)
 เป็นมารดาของผู้ป่วย ให้ประวัติว่าบุตรถ่ายเหลวเป็นน้ำ 10 ครั้ง
 ผู้ป่วยไม่มีโรคประจำตัว ไม่มีไข้ มีประวัติคล้าย URI นำมาก่อน

รายละเอียดผู้ป่วย

- ผู้ป่วยเด็กอายุ 1 ปี หนัก 10กก.
- ถ่ายเหลวเป็นน้ำ 10 ครั้งก่อนมาโรงพยาบาล

การดำเนินสถานการณ์จำลอง

	สมการผู้ป่วย	สิ่งขับเคลื่อนสถานการณ์ต่อไป	ประเด็นการ Debriefing
Scene 1 (initial status)	BP 65/50 mHg, P 160/min, RR 50/min, O ₂ sat 95%	-หากผู้เรียนยังไม่ได้ให้ O ₂ supplement เวลา 1 นาที ให้ไปที่ scene 2 -หากผู้เรียนขอให้เปิดเส้นให้น้ำเกลือ ให้แจ้งว่าเปิดเส้นไม่ได้ หลังจากเวลาผ่านไปประมาณ 1 นาที -หากไม่ขอให้เปิดเส้น ให้นอกถามผู้เรียนว่าต้องเปิดเส้นหรือไม่	การวินิจฉัยและประเมินผู้ป่วยที่มีภาวะ Shock

Scene 2 (O ₂ supplement)	BP 65/52mmHg, P 157/min, RR 50/min, O ₂ sat 88%	-หากผู้เรียนขอให้เปิดเส้นให้น้ำเกลือ ให้แจ้งว่าเปิดเส้นไม่ได้ หลังจากเวลาผ่านไปประมาณ 1 นาที -หากไม่ขอให้เปิดเส้น ให้นอกถามผู้เรียนว่าต้องเปิดเส้นหรือไม่	-ถ้าผู้เรียนขอทำ IO cannulation ไปที่ scene 3 -ถ้าพยายามเปิดน้ำเกลือเอง ให้นอกว่าเปิดไม่ได้ หากผู้เรียนไม่พิจารณาการทำ IO ให้ พยายามเสนอความเห็นเพื่อทำ IO	การดูแลผู้ป่วย shock เบื้องต้น สรรพภัณฑ์ใช้: ชนิด ปริมาณ และวิธีการให้
Scene 3 (IO cannulation)	BP 55/50 mHg, P 160/min, RR 50/min, O ₂ sat 95%	-ตามผู้เรียนว่าจะเจาะเลือดอะไรบ้าง -ตามผู้เรียนว่าจะย้ายผู้ป่วยอย่างไร	-ทำ IO สำเร็จไป scene 4	-indication และ limitation ของการทำ IO -ทักษะการทำ IO

จบสถานการณ์อย่างไร??

Scene 4 (improved)	BP 79/51 mHg, P 130/min, RR 40/min, O ₂ sat 99%		-การดูแล IO
-----------------------	--	--	-------------

The Debriefing

- **Debriefing:**
 - Postexperience analysis
 - The period of structured **reflection, discussion and feedback** after a simulation exercise

Flagman B, Inc. Manual simulation in healthcare, debriefing theory and technique.

The Debriefing

Allow trainees to **explain, analyze,** and **synthesize** information and emotional states

Being a Facilitator

- Be an **active listener** and a **vigilant observer**
- Guides process
 - ≠ presents information
- Provides the right questions
 - ≠ provides the right answers

Model of debriefing: *GAS model*

Phase	Goal
Gather	Understanding the students' perspective
Analyze	Facilitate student reflection and analysis of their actions
Summarize	Review of lessons learned

Phrampus R. Donnell J. Symposium of nursing simulation, WISER

Model of debriefing: *Plus & delta*

Two by two table to gather the participant feedback

Plus : positive aspect of the course

Delta : wanted to improve aspect of the course

The Features of the Effective Learning



- Educational feedback
- Repetitive practice
- Curriculum integration

Issenberg SB, medical teacher 2005

Learning is the process
whereby knowledge is created through
the **transformation of experience**

David A. Kolb

Q&A



tachypang@gmail.com

โครงการสู่ความเป็นเลิศของอาจารย์แพทย์ศิริราช (รุ่นที่ 12)
(สัปดาห์ที่ 2) วันที่ 13 - 15 พฤศจิกายน พ.ศ. 2562
ณ โรงแรมสามพรานริเวอร์ไซด์ จังหวัดนครปฐม

เอกสารประกอบการอบรม

The word "TEACH" is written in large, bold, dark blue capital letters with a white outline. It is centered on a light blue background that features a white ECG (heart rate) line. The background also has a subtle pattern of light blue and white squares, and some faint, glowing white circles at the bottom.

TEACH

13 November 2019

New staff workshop

Nov 2019

Basic Principles of Assessment

รศ.นพ.เชิดศักดิ์ ไอรรมณรัตน์
ภาควิชาศัลยศาสตร์
คณะแพทยศาสตร์ศิริราชพยาบาล

“Purposeful assessment drives instruction and affects learning.”

Wisconsin's guiding principles for teaching and learning

Objectives

- เมื่อสิ้นสุดการเรียนรู้ในหัวข้อนี้แล้ว อาจารย์แพทย์สามารถ
 - บอกความแตกต่างของการสอบ formative กับ summative ได้
 - บอกถึงเกณฑ์การประเมินคุณภาพการประเมินผลได้ทั้ง 7 ข้อ
 - เลือกใช้วิธีการประเมินผลที่เหมาะสมในการประเมินนักศึกษา

Outline

- Assessment and instruction
- Basic considerations in planning an assessment
- Guidelines for effective assessment
- Assessment levels: Miller's pyramid

Assessment

- The process of documenting, usually in measurable terms, knowledge, skills, attitudes and beliefs.

Assessment drives instruction.

Assessment and Instructional Process

- Placement
 - Aims at determining the readiness of students for the planned instruction
- Formative
 - Aims at providing feedback to students and teachers concerning learning successes and failures
- Summative
 - Aims at determining the extent to which instructional goals have been achieved; used primarily for assigning grades

cherdsakiramaneeerat@gmail.com

1

New staff workshop

Nov 2019

Criteria for Good Assessment

- Validity
- Reliability (Reproducibility)
- Equivalence
- Feasibility
- Educational Effect
- Catalytic Effect
- Acceptability

Norcini J, et al. Criteria for good assessment: Consensus statement and recommendations from the Ottawa 2010 conference. Medical Teacher 2011; 33: 206 – 214.

1. Validity

- The extent to which an assessment instrument measures what it intends to measure
- The degree to which evidence and theory support the interpretations of test scores entailed by the proposed uses of tests

Validity Threats

- **Construct Underrepresentation**

(วัดไม่ครอบคลุม)

The degree to which a test fails to capture important aspects of the construct. The test does not adequately sample some parts of the content

- **Construct-Irrelevant Variance**

(วัดปัจจัยอื่น)

The degree to which test scores are affected by processes that are extraneous to its intended construct

2. Reliability

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

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

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

11

12

cherdsakiramaneerat@gmail.com

2

New staff workshop

Nov 2019

3. Equivalence

- การทดสอบหัวข้อเดียวกันกับนักศึกษาระดับชั้นเรียนเดียวกันที่จัดสอบกันต่างเวลา ได้คะแนนที่เทียบเคียงกันได้

4. Feasibility

ความเป็นไปได้ของการจัดสอบ

The assessment is practical, realistic, and sensible, given appropriate contexts:

- Time
- Money
- Expertise
- Administration

5. Educational Effect

- การประเมินผลนั้นกระตุ้นให้ผู้เรียนมีการเรียนรู้ในเรื่องที่ควรเรียนรู้ ... educational benefit

6. Catalytic Effect

- การประเมินผลก่อให้เกิดการนำผลของการสอบไปใช้ให้ feedback เพื่อสร้าง หรือส่งเสริม หรือสนับสนุนการเรียนรู้ของนักศึกษา

7. Acceptability

- ผู้เกี่ยวข้อง (stakeholders) ทั้งหมดเชื่อถือผลประเมิน

Guidelines for Effective Assessment (1)

1. Effective assessment requires a clear conception of all intended learning outcomes.
2. Effective assessment requires that a variety of assessment procedures be used.
3. Effective assessment requires that the instructional relevance of the procedures be considered.

cherdsakiramaneeerat@gmail.com

3

New staff workshop

Nov 2019

Guidelines for Effective Assessment (2)

4. Effective assessment requires an adequate sample of student performance.
5. Effective assessment requires that the procedures be fair to everyone.
6. Effective assessment requires the specifications of criteria for judging successful performance.

Guidelines for Effective Assessment (3)

7. Effective assessment requires feedback to students.
8. Effective assessment must be supported by a comprehensive grading and reporting system.

What is grading?

- Grading is an exercise in professional judgment. It involves the collection and evaluation of evidence on students' achievement or performance over a specified period of time. Through this process, various types of descriptive information and measures of students' performance are converted into grades that summarize students' accomplishments.

Absolute Grading

- Strengths
 - Grades relate directly to student performance
 - All students can obtain high grades
 - Students have clear vision of how to get good grades
- Limitations
 - Standards can be arbitrary.
 - Performance standards tend to vary due to variations in test difficulty, student ability, and instructional effectiveness.

Relative Grading

- Strengths
 - Guarantee a constant proportion of grades in every group of students.
- Limitations
 - The percent of students receiving each grade is arbitrary.
 - The meaning of grades varies with the students' ability.
 - Prevent students from helping each other.
 - Cannot link students' grades to the accomplishment of medical competencies

Guidelines for Fair Grading

1. Inform students at the beginning of the course what grading procedures is used.
2. Base grades on student achievement, and achievement only.
3. Base grades on a wide variety of valid assessment data.
4. Use a proper technique to combine scores.
5. If there is no quota limitation, use absolute grading.
6. Review all borderline cases by reexamining all test scores.

cherdsakiraneerat@gmail.com

4

New staff workshop

Nov 2019

Assessment Approaches

Does

Shows how

Knows how

Knows

Miller's Pyramid

25

Assessment Methods in Health Professions Education

- Knows: Multiple-choice questions (MCQ)
- Knows how: Essay, Modified Essay questions (MEQ), Oral exam, chart-simulated recall (CSR)
- Shows how: Objective structured clinical examination (OSCE), Long case exam
- Does: clinical performance ratings, workplace-based assessment, portfolio

Summary

- Assessment and instruction
- Basic considerations in planning an assessment
- Guidelines for effective assessment
- Assessment levels: Miller's pyramid

**“Students can
escape bad teaching,
but they can’t escape
bad assessment”**

David Boud

cherdsakiramaneeerat@gmail.com

5

The metric of medical education

Validity: on the meaningful interpretation of assessment data

Steven M Downing

Context All assessments in medical education require evidence of validity to be interpreted meaningfully. In contemporary usage, all validity is construct validity, which requires multiple sources of evidence; construct validity is the whole of validity, but has multiple facets. Five sources – content, response process, internal structure, relationship to other variables and consequences – are noted by the *Standards for Educational and Psychological Testing* as fruitful areas to seek validity evidence.

Purpose The purpose of this article is to discuss construct validity in the context of medical education and to summarize, through example, some typical sources of validity evidence for a written and a performance examination.

Summary Assessments are not valid or invalid; rather, the scores or outcomes of assessments have more or less evidence to support (or refute) a specific interpretation (such as passing or failing a course). Validity is approached as hypothesis and uses theory, logic and the scientific method to collect and assemble data to

support or fail to support the proposed score interpretations, at a given point in time. Data and logic are assembled into arguments – pro and con – for some specific interpretation of assessment data. Examples of types of validity evidence, data and information from each source are discussed in the context of a high-stakes written and performance examination in medical education.

Conclusion All assessments require evidence of the reasonableness of the proposed interpretation, as test data in education have little or no intrinsic meaning. The constructs purported to be measured by our assessments are important to students, faculty, administrators, patients and society and require solid scientific evidence of their meaning.

Keywords Education, Medical, Undergraduate/ *standards, Educational measurement, Reproducibility of results.

Medical Education 2003;37:830–837

Introduction

The purpose of this paper is to discuss validity in the context of assessment in medical education and to present examples of the five types of validity evidence typically sought to support or refute the valid interpretations of assessment data.¹ This essay builds on and expands the older and more traditional view of test validity expressed in the first article in this series² and extends the validity discussion into state-of-the-art 21st century educational measurement.

Validity refers to the evidence presented to support or refute the meaning or interpretation assigned to assessment results. All assessments require validity

evidence and nearly all topics in assessment involve validity in some way. Validity is the *sine qua non* of assessment, as without evidence of validity, assessments in medical education have little or no intrinsic meaning.

Validity is always approached as hypothesis, such that the desired interpretative meaning associated with assessment data is first hypothesized and then data are collected and assembled to support or refute the validity hypothesis. In this conceptualization, assessment data are more or less valid for some very specific purpose, meaning or interpretation, at a given point in time and only for some well-defined population. The assessment itself is never said to be ‘valid’ or ‘invalid’ rather one speaks of the scientifically sound evidence presented to either support or refute the proposed interpretation of assessment scores, at a particular time period in which the validity evidence was collected.

In its contemporary conceptualization,^{1,3–14} validity is a unitary concept, which looks to multiple sources of

Department of Medical Education (MC 591), College of Medicine, University of Illinois at Chicago, Chicago, Illinois, USA

Correspondence: S M Downing, University of Illinois at Chicago, College of Medicine, Department of Medical Education (MC 591), 808 South Wood Street, Chicago, Illinois 60612-7309, USA. Tel.: +1 312 996 6428; Fax: +1 312 413 2048, E-mail: sdowning@uic.edu

Key learning points

Validity is a unitary concept, with construct validity as the whole of validity.

Assessments are not valid or invalid, rather assessment scores have more (or less) validity evidence to support the proposed interpretations.

Validity requires multiple sources of evidence to support or refute meaningful score interpretation.

Validity is always approached as hypothesis.

Validation research uses theory, data and logic to argue for or against specific score interpretations.

evidence. These evidentiary sources are typically logically suggested by the desired types of interpretation or meaning associated with measures. All validity is construct validity in this current framework, described most eloquently by Messick⁸ and embodied in the current *Standards of Educational and Psychological Measurement*.¹ In the past, validity was defined as three separate types: content, criterion and construct, with criterion-related validity usually subdivided into concurrent and predictive depending on the timing of the collection of the criterion data.^{2,15}

Why is construct validity now considered the sole type of validity? The complex answer is found in the philosophy of science⁸ from which, it is posited, there are many complex webs of inter-related inference associated with sampling content in order to make meaningful and reasonable inferences to a domain or larger population of interest. The more straightforward answer is: Nearly all assessments in the social sciences, including medical education, deal with *constructs* – intangible collections of abstract concepts and principles which are inferred from behavior and explained by educational or psychological theory. *Educational achievement* is a construct, usually inferred from performance on assessments such as written tests over some well-defined domain of knowledge, oral examinations over specific problems or cases in medicine, or highly structured standardized patient examinations of history-taking or communication skills.

Educational *ability* or *aptitude* is another example of a familiar construct – a construct that may be even more intangible and abstract than *achievement* because there is less agreement about its meaning among educators and psychologists.¹⁶ Tests that purport to measure educational ability, such as the Medical College Admissions Test (MCAT), which is relied on heavily

in North America for selecting prospective students for medical school admission, must present scientifically sound evidence, from multiple sources, to support the reasonableness of using MCAT test scores as one important selection criterion for admitting students to medical school. An important source of validity evidence for an examination such as the MCAT is likely to be the predictive relationship between test scores and medical school achievement.

Validity requires an evidentiary chain which clearly links the interpretation of the assessment scores or data to a network of theory, hypotheses and logic which are presented to support or refute the reasonableness of the desired interpretations. Validity is never assumed and is an ongoing process of hypothesis generation, data collection and testing, critical evaluation and logical inference. The validity argument^{11,12} relates theory, predicted relationships and empirical evidence in ways to suggest which particular interpretative meanings are reasonable and which are not reasonable for a specific assessment use or application.

In order to meaningfully interpret scores, some assessments, such as achievement tests of cognitive knowledge, may require fairly straightforward content-related evidence of the adequacy of the content tested (in relationship to instructional objectives), statistical evidence of score reproducibility and item statistical quality and evidence to support the defensibility of passing scores or grades. Other types of assessments, such as complex performance examinations, may require both evidence related to content and considerable empirical data demonstrating the statistical relationship between the performance examination and other measures of medical ability, the generalizability of the sampled cases to the population of skills, the reproducibility of the score scales, the adequacy of the standardized patient training and so on.

Some typical sources of validity evidence, depending on the purpose of the assessment and the desired interpretation are: evidence of the content representativeness of the test materials, the reproducibility and generalizability of the scores, the statistical characteristics of the assessment questions or performance prompts, the statistical relationship between and among other measures of the same (or different but related) constructs or traits, evidence of the impact of assessment scores on students and the consistency of pass-fail decisions made from the assessment scores.

The higher the stakes associated with assessments, the greater the requirement for validity evidence from multiple sources, collected on an ongoing basis and continually re-evaluated.¹⁷ The ongoing documentation of validity evidence for a very high-stakes testing

programme, such as a licensure or medical specialty certification examination, may require the allocation of many resources and the contributions of many different professionals with a variety of skills – content specialists, psychometricians and statisticians, test editors and administrators.

In the next section, five major sources of validity evidence are discussed in the contexts of example assessments in medical education.

Sources of evidence for construct validity

According to the *Standards*: ‘Validity refers to the degree to which evidence and theory support the interpretations of test scores entailed by proposed uses of tests’¹ (p. 9). The current *Standards*¹ fully embrace this unitary view of validity, following closely on Messick’s work^{8,9} that considers all validity as construct validity, which is defined as an investigative process through which constructs are carefully defined, data and evidence are gathered and assembled to form an argument either supporting or refuting some very specific interpretation of assessment scores.^{11,12} Historically, the methods of validation and the types of evidence associated with construct validity have their foundations on much earlier work by Cronbach,³⁻⁵ Cronbach and Meehl⁶ and Messick.⁷ The earliest unitary conceptualization of validity as construct validity dates to 1957 in a paper by Loevinger.¹⁸ Kane¹¹⁻¹³ places validity into the context of an interpretive argument, which must be established for each assessment; Kane’s work has provided a useful framework for validity and validation research.

The Standards

The *Standards*¹ discuss five distinct sources of validity evidence (Table 1): content, responses, internal structure, relationship to other variables and consequences. Each source of validity evidence (Table 1) is associated with some examples of the types of data that might be collected to support or refute specific assessment interpretations (validity). Some types of assessment demand a stronger emphasis on one or more sources of evidence as opposed to other sources and not all sources of data or evidence are required for all assessments. For example, a written, objectively scored test covering several weeks of instruction in microbiology, might emphasize content-related evidence, together with some evidence of response quality, internal structure and consequences, but very likely would not seek much or any evidence concerning relationship to other variables. On the other hand, a high-stakes

Table 1 Some sources of validity evidence for proposed score interpretations and examples of some types of evidence

Content	Response process	Internal structure	Relationship to other variables	Consequences
<ul style="list-style-type: none"> • Examination blueprint • Representativeness of test blueprint to achievement domain • Test specifications • Match of item content to test specifications • Representativeness of items to domain • Logical/empirical relationship of content tested to achievement domain • Quality of test questions • Item writer qualifications • Sensitivity review 	<ul style="list-style-type: none"> • Student format familiarity • Quality control of electronic scanning/scoring • Key validation of preliminary scores • Accuracy in combining different formats scores • Quality control/accuracy of final scores/marks/grades • Subscore/subscale analyses: Accuracy of applying pass-fail decision rules to scores • Quality control of score reporting to students/faculty • Understandable/accurate descriptions/interpretations of scores for students 	<ul style="list-style-type: none"> • Item analysis data: <ol style="list-style-type: none"> 1. Item difficulty/discrimination 2. Item/test characteristic curves (ICCs/TCCs) 3. Inter-item correlations 4. Item-total correlations • Score scale reliability • Standard errors of measurement (SEM) • Generalizability • Dimensionality • Item factor analysis • Differential Item Functioning (DIF) • Psychometric model 	<ul style="list-style-type: none"> • Correlation with other relevant variables • Convergent correlations - internal/external: <ol style="list-style-type: none"> 1. Similar tests 2. Divergent correlations-internal/external <ol style="list-style-type: none"> 1. Dissimilar measures 2. Test-criterion correlations • Generalizability of evidence 	<ul style="list-style-type: none"> • Impact of test scores/results on students/society • Consequences on learners/future learning <ul style="list-style-type: none"> • Positive consequences outweigh unintended negative consequences? • Reasonableness of method of establishing pass-fail (cut) score • Pass-fail consequences: <ol style="list-style-type: none"> 1. P/F Decision reliability- Classification accuracy 2. Conditional standard error of measurement at pass score (CSEM) <ul style="list-style-type: none"> • False positives/negatives • Instructional/learner consequences

summative Objective Structured Clinical Examination (OSCE), using standardized patients to portray and rate student performance on an examination that must be passed in order to proceed in the curriculum, might require all of these sources of evidence and many of the data examples noted in Table 1, to support or refute the proposed interpretation of the scores.

Sources of validity evidence for example assessments

Each of the five sources of validity evidence will now be considered, in the context of a written assessment of cognitive knowledge or achievement and a performance examination in medical education. Both example assessments are high-stakes, in that the consequences of passing or failing are very important to students, faculty and, ultimately, patients. The written assessment is a summative comprehensive examination in the basic sciences – a test consisting of 250 multiple-choice questions (MCQs) covering all the pre-clinical instruction in the basic sciences – and a test that must be passed in order to proceed into clinical training. The performance examination is a standardized patient (SP) examination, administered to medical students toward the end of their clinical training, after having completed all of their required clerkship rotations. The purpose of the SP examination is to comprehensively assess graduating medical students' ability to take a history and do a focused physical examination in an ambulatory primary care setting. The SP examination consists of 10 20-minute SP cases, presented by a lay, trained standardized patient who simulates the patient's presenting problem and rates the student's performance at the conclusion of the examination. The SP examination must be passed in order to graduate medical school.

Documentation of these five sources of validity evidence consists of the systematic collection and presentation of information and data to present a convincing argument that it is reasonable and defensible to interpret the assessment scores in accordance with the purpose of the measurement. The scores have little or no intrinsic meaning; thus the evidence presented must convince the skeptic that the assessment scores can reasonably be interpreted in the proposed manner.

Content evidence

For the written assessment, documentation of validity evidence related to the content tested is the most essential. The outline and plan for the test, described by a detailed test blueprint or test specifications, clearly

relates the content tested by the 250 MCQs to the domain of the basic sciences as described by the course learning objectives. The test blueprint is sufficiently detailed to describe subcategories and subclassifications of content and specifies precisely the proportion of test questions in each category and the cognitive level of those questions. The blueprint documentation shows a direct linkage of the questions on the test to the instructional objectives. Independent content experts can evaluate the reasonableness of the test blueprint with respect to the course objectives and the cognitive levels tested. The logical relationship between the content tested by the 250 MCQs and the major instructional objectives and teaching/learning activities of the course should be obvious and demonstrable, especially with respect to the proportionate weighting of test content to the actual emphasis of the basic science courses taught. Further, if most learning objectives were at the application or problem-solving level, most test questions should also be directed to these cognitive levels.

The quality of the test questions is a source of content-related validity evidence. Do the MCQs adhere to the best evidence-based principles of effective item-writing.¹⁹ Are the item-writers qualified as content experts in the disciplines? Are there sufficient numbers of questions to adequately sample the large content domain? Have the test questions been edited for clarity, removing all ambiguities and other common item flaws? Have the test questions been reviewed for cultural sensitivity?

For the SP performance examination, some of the same content issues must be documented and presented as validity evidence. For example, each of the 10 SP cases fits into a detailed content blueprint of ambulatory primary care history and physical examination skills. There is evidence of faculty content-expert agreement that these specific 10 cases are representative of primary care ambulatory cases. Ideally, the content of the 10 clinical cases is related to population demographic data and population data on disease incidence in primary care ambulatory settings. Evidence is documented that expert clinical faculty have created, reviewed and revised the SP cases together with the checklists and ratings scales used by the SPs, while other expert clinicians have reviewed and critically critiqued the SP cases. Exacting specifications detail all the essential clinical information to be portrayed by the SP. Evidence that SP cases have been competently edited and that detailed SP training guidelines and criteria have been prepared, reviewed by faculty experts and implemented by experienced SP trainers are all important sources of content-related validity evidence.

There is documentation that during the time of SP administration, the SP portrayals are monitored closely to ensure that all students experience nearly the same case. Data are presented to show that a different SP, trained on the same case, rates student case performance about the same. Many basic quality-control issues concerning performance examinations contribute to the content-related validity evidence for the assessment.²⁰

Response process

As a source of validity evidence, response process may seem a bit strange or inappropriate. *Response process* is defined here as evidence of data integrity such that all sources of error associated with the test administration are controlled or eliminated to the maximum extent possible. Response process has to do with aspects of assessment such as ensuring the accuracy of all responses to assessment prompts, the quality control of all data flowing from assessments, the appropriateness of the methods used to combine various types of assessment scores into one composite score and the usefulness and the accuracy of the score reports provided to examinees. (Assessment data quality-control issues could also be discussed as content evidence.)

For evidence of response process for the written comprehensive examination, documentation of all practice materials and written information about the test and instructions to students is important. Documentation of all quality-control procedures used to ensure the absolute accuracy of test scores is also an important source of evidence: the final key validation after a preliminary scoring – to ensure the accuracy of the scoring key and eliminate from final scoring any poorly performing test items; a rationale for any combining rules, such as the combining into one final composite score of MCQ, multiple true-false and short-essay question scores.

Other sources of evidence may include documentation and the rationale for the type of scores reported, the method chosen to report scores and the explanations and interpretive materials provided to explain fully the score report and its meaning, together with any materials discussing the proper use and any common misuses of the assessment score data.

For the SP performance examination, many of the same response process sources may be presented as validity evidence. For a performance examination, documentation demonstrating the accuracy of the SP rating is needed and the results of an SP accuracy study is a particularly important source of response process evidence. Basic quality control of the large amounts of data from an SP performance examination is important

to document, together with information on score calculation and reporting methods, their rationale and, particularly, the explanatory materials discussing an appropriate interpretation of the performance-assessment scores (and their limitations).

Documentation of the rationale for using global versus checklist rating scores, for example, may be an important source of response evidence for the SP examination. Or, the empirical evidence and logical rationale for combining a global rating-scale score with checklist item scores to form a composite score may be one very important source of response evidence.

Internal structure

Internal structure, as a source of validity evidence, relates to the statistical or psychometric characteristics of the examination questions or performance prompts, the scale properties – such as reproducibility and generalizability, and the psychometric model used to score and scale the assessment. For instance, scores on test items or sets of items intended to measure the same variable, construct, or content area should be more highly correlated than scores on items intended to measure a different variable, construct, or content area.

Many of the statistical analyses needed to support or refute evidence of the test's internal structure are often carried out as routine quality-control procedures. Analyses such as item analyses – which computes the difficulty (or easiness) of each test question (or performance prompt), the discrimination of each question (a statistical index indicating how well the question separates the high scoring from the low scoring examinees) and a detailed count of the number or proportion of examinees who responded to each option of the test question, are completed. Summary statistics are usually computed, showing the overall difficulty (or easiness) of the total test scale, the average discrimination and the internal consistency reliability of the test.

Reliability is an important aspect of an assessment's validity evidence. Reliability refers to the reproducibility of the scores on the assessment; high score reliability indicates that if the test were to be repeated over time, examinees would receive about the same scores on retesting as they received the first time. Unless assessment scores are reliable and reproducible (as in an experiment) it is nearly impossible to interpret the meaning of those scores – thus, validity evidence is lacking.

There are many different types of reliability, appropriate to various uses of assessment scores. In both example assessments described above, in which the

stakes are high and a passing score has been established, the reproducibility of the pass-fail decision is a very important source of validity evidence. That is, analogous to score reliability, if the ultimate outcome of the assessment (passing or failing) can not be reproduced at some high level of certainty, the meaningful interpretation of the test scores is questionable and validity evidence is compromised.

For performance examinations, such as the SP example, a very specialized type of reliability, derived from generalizability theory (GT)^{21,22} is an essential component of the internal structure aspect of validity evidence. GT is concerned with how well the specific samples of behaviour (SP cases) can be generalized to the population or universe of behaviours. GT is also a useful tool for estimating the various sources of contributed error in the SP exam, such as error due to the SP raters, error due to the cases (case specificity), and error associated with examinees. As rater error and case specificity are major threats to meaningful interpretation of SP scores, GT analyses are important sources of validity evidence for most performance assessments such as OSCEs, SP exams and clinical performance examinations.

For some assessment applications, in which sophisticated statistical measurement models like Item Response Theory (IRT) models^{23,24} the measurement model itself is evidence of the internal structure aspect of construct validity. In IRT applications, which might be used for tests such as the comprehensive written examination example, the factor structure, item-inter-correlation structure and other internal structural characteristics all contribute to validity evidence.

Issues of bias and fairness also pertain to internal test structure and are important sources of validity evidence. All assessments, presented to heterogeneous groups of examinees, have the potential of validity threats from statistical bias. Bias analyses, such as differential item functioning (DIF)^{25,26} analyses and the sensitivity review of item and performance prompts are sources of internal structure validity evidence. Documentation of the absence of statistical test bias permits the desired score interpretation and therefore adds to the validity evidence of the assessment.

Relationship to other variables

This familiar source of validity evidence is statistical and correlational. The correlation or relationship of assessment scores to a criterion measure's scores is a typical design for a 'validity study', in which some newer (or simpler or shorter) measure is 'validated'

against an existing, older measure with well known characteristics.

This source of validity evidence embodies all the richness and complexity of the contemporary theory of validity in that the relationship to other variables aspect seeks both confirmatory and counter-confirmatory evidence. For example, it may be important to collect correlational validity evidence which shows a strong positive correlation with some other measure of the same achievement or ability and evidence indicating no correlation (or a strong negative correlation) with some other assessment that is hypothesized to be a measure of some completely different achievement or ability.

The concept of convergence and divergence of validity evidence is best exemplified in the classic research design first described by Campbell and Fiske.²⁷ In this 'multitrait multimethod' design, different measures of the same trait (achievement, ability, performance) are correlated with different measures of the same trait. The resulting pattern of correlation coefficients may show the convergence and divergence of the different assessment methods on measures of the same and different abilities or proficiencies.

In the written comprehensive examination example, it may be important to document the correlation of total and subscale scores with achievement examinations administered during the basic science courses. One could hypothesize that a subscale score for biochemistry on the comprehensive examination would correlate more highly with biochemistry course test scores than with behavioural science course scores. Additionally, the correlation of the written examination scores with the SP final examination may show a low (or no) correlation, indicating that these assessment methods measure some unique achievement, while the correlation of the SP scores with other performance examination scores during the students' clinical training may be high and positive.

As with all research, issues of the generalizability of the results of these studies and the limitations of data interpretation pertain. Interpretation of correlation coefficients, as validity coefficients, may be limited due to the design of the study, systematic bias introduced by missing data from either the test or the criterion or both and statistical issues such as restriction of the range of scores (lack of variance).

Consequences

This aspect of validity evidence may be the most controversial, although it is solidly embodied in the current *Standards*.¹ The consequential aspect of validity refers to the impact on examinees from the assessment

scores, decisions and outcomes, and the impact of assessments on teaching and learning. The consequences of assessments on examinees, faculty, patients and society can be great and these consequences can be positive or negative, intended or unintended.

High-stakes examinations abound in North America, especially in medicine and medical education. Extremely high-stakes assessments are often mandated as the final, summative hurdle in professional education. For example, the United States Medical Licensure Examination (USMLE) sequence, sponsored by the National Board of Medical Examiners (NBME), consists of three separate examinations (Steps 1, 2 and 3) which must be passed in order to be licensed as a physician. The consequences of failing any of these examinations is enormous, in that medical education is interrupted in a costly manner or the examinee is not permitted to enter graduate medical education or practice medicine. Likewise, most medical specialty boards in the USA mandate passing a high-stakes certification examination in the specialty or subspecialty, after meeting all eligibility requirements of postgraduate training. The consequences of passing or failing these types of examinations are great, as false positives (passing candidates who should fail) may do harm to patients through the lack of a physician's specialized knowledge or skill and false negatives (failing candidates who should pass) may unjustly harm individual candidates who have invested a great deal of time and resources in graduate medical education.

Thus, consequential validity is one very important aspect of the construct validity argument. Evidence related to consequences of testing and its outcomes is presented to suggest that no harm comes directly from the assessment or, at the very least, more good than harm arises from the assessment. Much of this evidence is more subjective than other sources.

In both example assessments, sources of consequential validity may relate to issues such as passing rates (the proportion who pass), the subjectively judged appropriateness of these passing rates, data comparing the passing rates of each of these examinations to other comprehensive examinations such as the USMLE Step 1 and so on. Evaluations of false positive and false negative outcomes relate to the consequences of these two high-stakes examinations.

The passing score (or grade levels) and the process used to determine the cut scores, the statistical properties of the passing scores, and so on all relate to the consequential aspects of validity.²⁸ Documentation of the method used to establish a pass-fail score is key consequential evidence, as is the rationale for the

selection of a particular passing score method. The psychometric characteristics of the passing score judgments and the qualification and number of expert judges – all may be important to document and present as evidence of consequential validity.

Other psychometric quality indicators concerning the passing score and its consequences (for both example assessments) include a formal, statistical estimation of the pass-fail decision reliability or classification accuracy²⁹ and some estimation of the standard error of measurement at the cut score.³⁰

Equally important consequences of assessment methods on instruction and learning have been discussed by Newble and Jaeger.³¹ The methods and strategies selected to evaluate students can have a profound impact on what is taught, how and exactly what students learn, how this learning is used and retained (or not) and how students view and value the educational process.

Threats to validity

The next essay in this series will discuss the many threats to the meaningful interpretation of assessment scores and suggest methods to control these validity threats.

Conclusion

This paper has reviewed the contemporary meaning of validity, a unitary concept with multiple facets, which considers construct validity as the whole of validity. Validity evidence refers to the data and information collected in order to assign meaningful interpretation to assessment scores or outcomes, which were designed for a specific purpose and at one specific point in time. Validity always refers to score interpretations and never to the assessment itself. The process of validation is closely aligned with the scientific method of theory development, hypothesis generation, data collection for the purpose of hypothesis testing and forming conclusions concerning the accuracy of the desired score interpretations. Validity refers to the impartial, scientific collection of data, from multiple sources, to provide more or less support for the validity hypothesis and relates to logical arguments, based on theory and data, which are formed to assign meaningful interpretations to assessment data.

This paper discussed five typical sources of validity evidence – content, response process, internal structure, relationship to other variables and consequences – in the context of two example assessments in medical education.

Acknowledgements

The author wishes to thank Michael T Kane, PhD, for his critical review of this manuscript.

Funding

There was no external funding for this project.

References

- American Educational Research Association, American Psychological Association, National Council on Measurement in Education. *Standards for Educational and Psychological Testing*. Washington, DC: American Educational Research Association 1999.
- Crossley J, Humphris G, Jolly B. Assessing health professionals. *Med Educ* 2002;**36**:800-4.
- Cronbach LJ. Test validation. In: *Educational Measurement*, 2nd edn. Ed: Thorndike RL. Washington, DC: American Council on Education 1971:443-507.
- Cronbach LJ. Five perspectives on validity argument. In: *Test Validity*. Eds: Wainer H, Braun H. Hillsdale, NJ: Lawrence Erlbaum 1988:3-17.
- Cronbach LJ. Construct validation after 30 years. In: *Intelligence: Measurement, Theory, and Public Policy*. Ed: Linn RE. Urbana, IL: University of Illinois Press 1989:147-71.
- Cronbach LJ, Meehl PE. Construct validity in psychological tests. *Psychol Bull* 1955;**52**:281-302.
- Messick S. The psychology of educational measurement. *J Educ Measure* 1984;**21**:215-37.
- Messick S. Validity. In: *Educational Measurement*, 3rd edn. Ed: Linn RL. New York: American Council on Education and Macmillan 1989:13-104.
- Messick S. Validity of psychological assessment: validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *Am Psychologist* 1995;**50**:741-9.
- Messick S. Standards of validity and the validity of standards in performance assessment. *Educ Measure Issues Prac* 1995;**14**:5-8.
- Kane MT. An argument-based approach to validation. *Psychol Bull* 1992;**112**:527-35.
- Kane MT. Validating interpretive arguments for licensure and certification examinations. *Evaluation Health Professions* 1994;**17**:133-59.
- Kane MT. Current concerns in validity theory. *J Educ Measure* 2001;**38**:319-42.
- Kane MT, Crooks TJ, Cohen AS. Validating measures of performance. *Educ Measure Issues Prac* 1999;**18**:5-17.
- Cureton EE. Validity. In: *Educational Measurement*. Ed: Lingquist EF. Washington, DC: American Council on Education 1951:621-94.
- Lohman DF. Teaching and testing to develop fluid abilities. *Educational Reser* 1993;**22**:12-23.
- Linn RL. Validation of the uses and interpretations of results of state assessment and accountability systems. In: *Large-Scale Assessment Programs for All Students: Development, Implementation, and Analysis*. Eds: Tindal G, Haladyna T. Mahwah, NJ: Lawrence Erlbaum 2002.
- Loevinger J. Objective tests as instruments of psychological theory. *Psychol Reports, Monograph* 1957;**3** (Suppl.) 635-94.
- Haladyna TM, Downing SM, Rodriguez MC. A review of multiple-choice item-writing guidelines for classroom assessment. *Appl Measure Educ* 2002;**15**:309-34.
- Boulet JR, McKinley DW, Whelan GP, Hambleton RK. Quality assurance methods for performance-based assessments. *Adv Health Sci Educ* 2003;**8**:27-47.
- Brennan RL. *Generalizability Theory*. New York: Springer-Verlag 2001.
- Crossley J, Davies H, Humphris G, Jolly B. Generalisability; a key to unlock professional assessment. *Med Educ* 2002;**36**:972-8.
- Van der Linden WJ, Hambleton RK. Item response theory. Brief history, common models, and extensions. In: *Handbook of Modern Item Response Theory*. Eds: van der Linden WJ, Hambleton RK. New York: Springer-Verlag 1997:1-28.
- Downing SM. Item response theory: Applications of modern test theory in medical education. *Med Educ* 2003;**37**:1-7.
- Holland PW, Wainer H, eds. *Differential Item Functioning*. Mahwah, NJ: Lawrence Erlbaum 1993.
- Penfield RD, Lam RCM. Assessing differential item functioning in performance assessment: review and recommendations. *Educ Measure Issues Prac* 2000;**19**:5-15.
- Campbell DT, Fiske DW. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psych Bull* 1959;**56**:81-105.
- Norcini JJ. Setting standards on educational tests. *Med Educ* 2003;**37**:464-9.
- Subkoviak MJ. A practitioner's guide to computation and interpretation of reliability indices for mastery tests. *J Educ Measure* 1988;**25**:47-55.
- Angoff WH. Scales, norms, and equivalent scores. In: *Educational Measurement*, 2nd edn. Ed: Thorndike RL. Washington, DC: American Council on Education 1971:508-600.
- Newble DI, Jaeger K. The effects of assessment and examinations on the learning of medical students. *Med Educ* 1983;**17**:165-71.

Received 29 May 2003; accepted for publication 3 June 2003

Iramaneerat C. Validity threats [Thai]. Medical Education Pamphlet 2006; 2(9): 1.

สิ่งไม่พึงประสงค์ในการสอบ
เชิดศักดิ์ ไอรมนิรัตน์

ในบทความนี้ผมจะขอกล่าวถึงสิ่งอันไม่พึงประสงค์ในการสอบ (Validity threats) ที่เราต้องคำนึงถึงในการจัดสอบ ดังที่ได้กล่าวในบทความก่อนหน้านี้แล้วว่า Validity นั้นคือการประเมินคุณค่าของการแปลผลและการนำผลสอบไปใช้ ดังนั้น สิ่งอันไม่พึงประสงค์ในการสอบ หรือ validity threats ก็คือสิ่งใดก็ตามที่เข้ามาบงกการแปลผลสอบ สิ่งบงกเหล่านี้แยกได้เป็น 2 ปัจจัยหลัก คือ construct underrepresentation และ construct-irrelevant variance

Construct underrepresentation หมายถึงการประเมินผลที่ไม่ครอบคลุมสิ่งที่ต้องการวัดอย่างเพียงพอ ทำให้ผลการสอบไม่สามารถบ่งบอกถึงความสามารถของนักเรียนผู้สอบในเรื่องที่ต้องการวัดผลอย่างครบถ้วน ตัวอย่างเช่นในการสอบ OSCE เพื่อวัดความสามารถของแพทย์ประจำบ้านในการให้คำแนะนำปรึกษาแก่ผู้ป่วย หากเกณฑ์การให้คะแนนมีเพียงหัวข้อที่เกี่ยวกับการพูดกับผู้ป่วย แต่ไม่มีหัวข้อที่เกี่ยวกับการใช้ อวัจนภาษา เช่น การใช้ท่าทาง น้ำเสียง การรับฟังปัญหา เป็นต้น ก็จัดว่า ทำการประเมินไม่ครอบคลุมเนื้อหา ผลการประเมินก็นำไปใช้บอกได้เพียงว่าแพทย์ประจำบ้านให้ข้อมูลผู้ป่วยครบถ้วน แต่ไม่สามารถบอกได้ว่าแพทย์ประจำบ้านทำการสื่อสารกับผู้ป่วยได้ดีในทุกด้าน ในการสอบข้อเขียนสำหรับวัดความรู้ของนักเรียน หากใช้ข้อสอบที่สั้นเกินไป มีจำนวนข้อสอบไม่กี่ข้อ ก็จะมีปัญหาที่ไม่สามารถวัดความรู้ของนักเรียนได้ครอบคลุมเนื้อหาที่ต้องการวัดผล

Construct-irrelevant variance หมายถึง ปัจจัยอื่นที่นอกเหนือไปจากความรู้ความสามารถของนักเรียนที่สามารถส่งผลต่อคะแนนสอบของนักเรียนได้ ปัจจัยนี้อาจบงกการคะแนนสอบ multiple-choice examination ได้แก่

- ข้อสอบที่ไม่มีคุณภาพ โจทย์คำถามกำกวม มีตัวเลือกที่ถูกต้องมากกว่า 1 ตัวเลือก ทำให้นักเรียนที่มีความรู้ตอบผิด หรือโจทย์คำถามบอกใบ้ให้นักเรียนตอบถูกโดยไม่ต้องใช้ความรู้ ข้อสอบเก่าที่รั่วไหลออกจากคลังข้อสอบทำให้นักเรียนที่รู้ข้อสอบมาก่อนสามารถตอบได้โดยไม่ต้องคิด
 - นักเรียนที่ทุจริตในการสอบ ลอกข้อสอบของเพื่อน หรือใช้วิธีการอื่นในการได้มาซึ่งคำตอบโดยที่ไม่ได้ใช้ความรู้ในเรื่องที่ทำการสอบ
 - อาจารย์ที่บอกข้อสอบให้นักเรียนในการสอน ทำให้นักเรียนที่ท่องคำตอบเข้าไปสอบ ทำข้อสอบได้โดยไม่ต้องคิด สำหรับการสอบในรูปแบบอื่นที่ต้องใช้กรรมการให้คะแนน เช่น OSCE การสอบข้อสอบบรรยาย หรือการสอบปากเปล่า นั้นจะมีปัจจัยที่เกี่ยวข้องเกี่ยวกับกรรมการผู้ให้คะแนนเข้ามาบงกการแปลผลคะแนนสอบได้ด้วย เช่น
 - ความไม่เสมอภาคของอาจารย์ในเกณฑ์การให้คะแนน นักเรียนที่สอบกับอาจารย์ที่กดคะแนน เสียเปรียบนักเรียนที่สอบกับอาจารย์ที่ใจดี และปล่อยคะแนน
 - ความไม่สม่ำเสมอของอาจารย์ในการให้คะแนน อาจารย์บางท่านมีแนวโน้มจะให้คะแนนต่ำลงในกลุ่มนักเรียนที่สอบตอนท้าย เนื่องด้วยความเหนื่อยล้า ในขณะที่อาจารย์บางท่านมีแนวโน้มจะให้คะแนนสูงขึ้นในตอนท้ายของการสอบ เนื่องจากได้เห็นความสามารถของนักเรียนจำนวนหนึ่งแล้วพบว่าเกณฑ์ที่ตั้งเป้าไว้นั้นสูงเกินความสามารถของนักเรียนส่วนใหญ่จึงปรับเกณฑ์การให้คะแนนให้ง่ายลง ทำให้นักเรียนในกลุ่มหลังได้คะแนนง่ายขึ้น
 - การจำกัดช่วงของคะแนน ที่พบบ่อยคืออาจารย์บางท่านนิยมเดินสายกลาง ไม่ว่าจะนักเรียนจะทำดีมากหรือน้อยเพียงใด ก็มักจะให้คะแนนอยู่ในเกณฑ์ปานกลาง ไม่กล้าให้คะแนน 0 ในรายที่ทำไม่ดี แต่ก็ไม่กล้าให้คะแนนเต็มในนักเรียนที่ทำได้ดี
- ปัจจัยต่างๆ เหล่านี้ เป็นสิ่งที่ผู้จัดสอบต้องคำนึงถึงเสมอในการจัดสอบและตั้งมาตรฐานเพื่อควบคุมและกำจัดปัจจัยบงกเหล่านี้จากการสอบ เพื่อให้ได้ผลการสอบที่มีความเที่ยงตรง เป็นธรรม และสามารถใช้อธิบายความรู้ ความสามารถของนักเรียนได้ตามที่ต้องการ

หลักการพื้นฐานของข้อสอบปรนัย

รศ.นพ. เชิดศักดิ์ ไอร่มณีรัตน์
ภาควิชาศัลยศาสตร์ คณะแพทยศาสตร์ศิริราชพยาบาล
มหาวิทยาลัยมหิดล

Multiple-Choice Questions

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

3

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

- 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

4

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

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

การวางแผนและ
จัดทำตารางกำหนดเนื้อหาข้อสอบ

รศ.นพ. เชิดศักดิ์ ไอร่มเจริญรัตน์
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

Content Definition

- What content is to be tested?
- Human judgment

Categorization of the Test Items

1. Nature of the content
2. Nature of learning

Cognitive Hierarchy

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

A Simplified Cognitive Hierarchy

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

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

การสร้างโจทย์และตัวเลือกข้อสอบ

รศ. นพ. เชิดศักดิ์ ไอร่มณีรัตน์
ภาควิชาศัลยศาสตร์ คณะแพทยศาสตร์ศิริราชพยาบาล
มหาวิทยาลัยมหิดล

Twelve Steps for Test Development

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

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

A Good MCQ Item

1. Content
2. Structure

Guidelines for MCQ items

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

4

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

5

Format Guidelines

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

6

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

7

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

8

Guidelines for MCQ items

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

9

Questions & Comments

การสร้างข้อสอบปรนัย

ผู้ช่วยศาสตราจารย์ นายแพทย์เชิดศักดิ์ โสมณรัตน์

ภาควิชาศัลยศาสตร์, คณะแพทยศาสตร์ศิริราชพยาบาล, มหาวิทยาลัยมหิดล, กรุงเทพมหานคร ๑๐๓๐๐.

ข้อสอบปรนัย (multiple-choice question) เป็นรูปแบบการประเมินผลที่นิยมใช้กันอย่างแพร่หลาย ในวงการแพทยศาสตรศึกษาเนื่องด้วยคุณสมบัติที่ดีหลายประการด้วยกัน ได้แก่ ประสิทธิภาพในการประเมินความรู้ปริมาณมากในเวลาอันสั้น ผลการประเมินที่ไม่มีผลกระทบจากความรู้สึกส่วนตัวของผู้ตรวจให้คะแนน คะแนนที่มีความเที่ยงสูง รวมถึงผลการวิจัยจำนวนมากที่สนับสนุนความถูกต้องของผลการประเมินด้วยข้อสอบปรนัย^{๑-๖} ข้อสอบปรนัยที่พัฒนาขึ้นอย่างดีนั้นสามารถวัดความรู้ได้ทั้งระดับการจดจำ การทำความเข้าใจ และการประยุกต์ความรู้ไปใช้ในการดูแลคนไข้^{๗-๙} อย่างไรก็ตาม การศึกษาค้นคว้าวิจัยเกี่ยวกับคุณภาพของข้อสอบปรนัยที่พัฒนาขึ้นใช้ในโรงเรียนแพทย์หลายแห่งพบว่าข้อสอบจำนวนไม่น้อยมีลักษณะที่ไม่เหมาะสม^{๑๐-๑๒} ข้อสอบปรนัยที่ถูกพัฒนาขึ้นอย่างไม่ถูกหลักการนั้นส่งผลเสียหลายอย่าง เช่น ทำให้ข้อสอบยากขึ้นโดยไม่จำเป็น ทำให้ผู้สอบเกิดความสับสน ทำให้ผู้สอบบางกลุ่มเสียเปรียบผู้สอบคนอื่น ทำให้การตัดสินใจผิดพลาด เป็นต้น^{๑๓-๑๕} ดังนั้นการออกข้อสอบปรนัยที่ดี วางอยู่บนหลักการที่ต้องจึงมีความสำคัญมากในการควบคุมคุณภาพการศึกษาในโรงเรียนแพทย์ บทความนี้จะจึงถูกเขียนขึ้นเพื่อเป็นการรวบรวมหลักการพื้นฐานในการออกข้อสอบปรนัยที่ได้รับการยอมรับกันทั่วไปในวงการวัดและประเมินผล ผู้ที่หวังว่าข้อแนะนำต่าง ๆ ที่ได้นำเสนอในบทความนี้จะ เป็นแนวทางที่เป็นประโยชน์ในการพัฒนาข้อสอบปรนัยที่มีคุณภาพให้ผู้อ่านไม่มากนักน้อย

รูปแบบพื้นฐานของข้อสอบปรนัย

ข้อสอบปรนัยคือข้อสอบชนิดที่มีคำถามแล้วมีตัวเลือกให้ผู้สอบเลือกตัวเลือกที่เหมาะสมเพื่อตอบคำถามดังกล่าว ข้อสอบปรนัยสามารถแบ่งออกได้เป็น ๒ รูปแบบ^{๑๖} ได้แก่

๑. ข้อสอบถูกผิด (True/false item)

ในข้อสอบประเภทนี้จะมีข้อความให้ผู้สอบพิจารณาว่าถูกหรือผิด ในยุคแรกข้อสอบเหล่านี้แต่ละข้อจะแยกเป็นอิสระจากกัน ผู้สอบตัดสินใจว่าข้อความแต่ละข้อถูกหรือผิดโดยไม่เกี่ยวข้องข้อข้อความในข้ออื่น ต่อมาเมื่อพัฒนาข้อสอบเป็นชุดของข้อความ (multiple true/false หรือ K-type item) โดยในแต่ละข้อจะมีข้อความ ผู้สอบต้องพิจารณาว่าแต่ละข้อความถูกหรือผิด แล้วทำการเลือกตัวเลือกที่บรรยายจำนวนข้อความที่ต้องได้อย่างเหมาะสม (เช่น ตอบ ก. เมื่อข้อความที่ ๑, ๒, และ ๓ ถูกต้อง, ตอบ ข. เมื่อข้อความที่ ๑ และ ๓ ถูกต้อง ฯลฯ)

ข้อสอบชนิดถูกผิดนี้เคยเป็นที่นิยมมากในวงการแพทยศาสตรศึกษาอยู่ระยะหนึ่งเนื่องจากสามารถทดสอบความรู้ได้ปริมาณมาก แต่ข้อสอบชนิดนี้มีข้อจำกัดที่สำคัญคือสามารถใช้ได้เฉพาะกับเนื้อหาที่มีความถูกต้องชัดเจนเท่านั้น ซึ่งการตัดสินใจทางการแพทย์ส่วนมากไม่เป็นเช่นนั้น การตัดสินใจในการวินิจฉัย การตรวจค้นเพิ่มเติม หรือการรักษามือป่วยส่วนใหญ่นั้นแพทย์ตัดสินใจเลือกกระหว่างทางเลือกที่แตกต่างกันสามสี่อย่างซึ่งทุกทางเลือกมีความเป็นไปได้ มีส่วนถูก หรือมีความเหมาะสมในบางด้าน

เวชบันทึกศิริราช

บทความทั่วไป

แต่ก็มีความไม่เหมาะสมในด้านอื่นด้วย เช่นการเลือกใช้ยาในผู้ป่วยที่มีการติดเชื้อ นักศึกษาแพทย์มักรู้ว่าควรใช้ยาปฏิชีวนะ ซึ่งยาปฏิชีวนะหลายชนิดก็รักษาการติดเชื้อชนิดนั้นๆ ได้ แต่นักศึกษาต้องเลือกระหว่างยาที่ล้วนใช้ได้ในการรักษานั้นว่ายาใดที่มีประสิทธิภาพสูงสุด เหมาะสมที่สุดกับชนิดของเชื้อก่อโรคที่พบบ่อยในการติดเชื้อนั้น มีผลข้างเคียงน้อยที่สุด และราคาเหมาะสมด้วย ซึ่งในสถานการณ์ข้อสอบชนิดถูกผิดจะนำมาใช้ได้ยาก ด้วยเหตุนี้ทำให้ข้อสอบชนิดถูกผิดไม่เป็นที่นิยมกันมากนักในปัจจุบัน

๒. ข้อสอบเลือกคำตอบที่ถูกที่สุด (one best response item)

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

เนื่องจากข้อสอบชนิดที่มีใช้กันแพร่หลายในวงการแพทยศาสตรศึกษาในประเทศไทยในปัจจุบันคือข้อสอบประเภทที่มีตัวเลือก ๔-๕ ตัวเลือก (A-type) ผู้นิพนธ์จะขอเน้นหลักการสำหรับการออกข้อสอบประเภทนี้เป็นสำคัญ

องค์ประกอบของข้อสอบปรนัยชนิดเลือกคำตอบที่ถูกที่สุด

ข้อสอบปรนัยแต่ละข้อมีส่วนประกอบสำคัญ ๒ ส่วนด้วยกันคือ

๑. โจทย์ (stem) เป็นข้อมูลของโรค หรือภาวะ หรือผู้ป่วยตามด้วยคำถาม หรือเว้นช่องว่างสำหรับเติมคำ หรือข้อความที่เหมาะสมลงไป

๒. ตัวเลือก (options) คือคำ หรือข้อความที่

ผู้ออกข้อสอบนำเสนอตามหลังจากโจทย์เพื่อให้ผู้สอบเลือกไปใช้ตอบคำถาม หรือเติมลงในช่องว่างในโจทย์

๒.๑ ตัวเลือกที่ถูกต้อง (correct option) เป็นคำตอบที่ถูกต้องมีเพียงตัวเลือกเดียวต่อข้อสอบข้อหนึ่ง

๒.๒ ตัวลวง (distractors) เป็นคำตอบที่ผิดหรือไม่เหมาะสม มีไว้ลวงให้ผู้สอบที่ไม่มีความรู้ หรือมีความเข้าใจไม่ถูกต้องในเนื้อหาที่นำมาออกข้อสอบเลือกตอบ ตัวลวงไม่จำเป็นต้องเป็นคำตอบที่ผิดชัดเจนเสมอไป ตัวลวงที่ดีมักมีส่วนถูกบ้าง แต่มีระดับของความถูกต้องเหมาะสมน้อยกว่าคำตอบที่ถูก

ข้อแนะนำพื้นฐานของการเขียนข้อสอบปรนัย

มีผู้เชี่ยวชาญทางการประเมินผลให้ข้อแนะนำจำนวนมากในการเขียนข้อสอบปรนัย เคยมีผู้รวบรวมไว้ถึง ๔๓ ข้อ^{๑๖} ในที่นี้ผู้นิพนธ์ขอนำเสนอเฉพาะข้อแนะนำที่ได้รับการยอมรับอย่างกว้างขวางและสามารถประยุกต์ใช้ได้ชัดเจนในการพัฒนาข้อสอบทางการแพทย์ โดยจะทำการจัดหมวดหมู่ของข้อแนะนำเหล่านี้ออกเป็น ๔ กลุ่มด้วยกัน ได้แก่ (๑) เนื้อหาข้อสอบ, (๒) การจัดรูปแบบข้อสอบ, (๓) การเขียนโจทย์, และ (๔) การเขียนตัวเลือก

๑. เนื้อหาข้อสอบ

๑.๑ ข้อสอบหนึ่งข้อควรมุ่งเน้นประเมินความรู้เพียงเรื่องเดียว

ก่อนเริ่มเขียนข้อสอบอาจารย์ผู้ออกข้อสอบควรตั้งวัตถุประสงค์ให้ชัดเจนว่าต้องการประเมินความรู้ของผู้สอบในเรื่องใด และเขียนโจทย์เพื่อทดสอบองค์ความรู้ดังกล่าวเท่านั้น เนื่องจากเนื้อหาวิชาทางการแพทย์มีมาก อาจารย์แต่ละท่านเมื่อทำการสอนไปแล้วจึงอยากจะทดสอบความรู้ในหลายเรื่องที่ได้สอนไป แต่กลับมีโควตาจำกัดในการออกข้อสอบ ทำให้อาจารย์จำนวนไม่น้อยเขียนข้อสอบหนึ่งข้อถามทั้งเรื่องการวินิจฉัยโรค การตรวจค้นเพิ่มเติม การรักษาโรค และภาวะแทรกซ้อนของโรคไปพร้อมกัน ลักษณะข้อสอบเช่นนี้ไม่ควรใช้ เพราะมักซับซ้อนเกินไป เมื่อผู้สอบตอบข้อสอบผิด ก็ไม่สามารถวินิจฉัยได้ว่าผู้สอบขาดความรู้ ความเข้าใจในเรื่องใด

๑.๒ หลีกเลี่ยงการถามความรู้ในรายละเอียดปลีกย่อยที่ไม่มีที่ใช้ทางคลินิก (trivial content)

๓๓๐

มกราคม-มิถุนายน ๒๕๕๕, ปีที่ ๕, ฉบับที่ ๑

เวบบ์ทีกีรราช

บทความทั่วไป

องค์ความรู้ทางการแพทย์นั้นมีปริมาณมาก ไม่มีผู้ใดที่จดจำเนื้อหาที่มีในตำรา หรือวารสารทางการแพทย์ได้ทั้งหมด แม้ว่าองค์ความรู้หลายเรื่องมีความน่าสนใจ แต่มีประโยชน์ในการประยุกต์ใช้ทางคลินิกค่อนข้างน้อย องค์ความรู้ดังกล่าวจัดเป็นรายละเอียดปลีกย่อย (trivial content) ซึ่งไม่แนะนำให้ทำการทดสอบ สิ่งที่ดีกว่าการประเมินคือความสามารถในการประยุกต์ใช้ความรู้ในทางคลินิก (application of knowledge) ไม่แนะนำการทดสอบวัดความสามารถในการจดจำเป็นหลัก อย่างไรก็ตามก็ตามการที่แนะนำให้ออกข้อสอบที่เน้นการประยุกต์ใช้ความรู้ ไม่ได้หมายความว่า การแก้ปัญหาผู้ป่วยนั้นไม่ต้องใช้ความจำเลย ตรงกันข้ามการจดจำเนื้อหาเป็นพื้นฐานที่สำคัญในการแก้ปัญหาทางคลินิก ผู้สอบย่อมต้องจำเนื้อหาได้บ้าง จึงจะประยุกต์องค์ความรู้ดังกล่าวไปแก้ไขภัยปัญหาที่นำเสนอได้

๑.๓ หลีกเลียงการถามความรู้ในเรื่องที่ยังมีความขัดแย้งกันในแนวทางปฏิบัติ (controversy)

ความรู้ทางการแพทย์ในหลายหัวข้อยังเป็นเรื่องและผู้เชี่ยวชาญยังมีความเห็นแตกต่างกัน ผู้ป่วยรายเดียวกันไปพบแพทย์สองคนอาจได้รับการรักษาที่แตกต่างกันซึ่งวิธีการรักษาทั้งสองวิธีก็มีงานวิจัยสนับสนุนด้วยทั้งคู่ อย่างไรก็ตามยังคงมีความขัดแย้ง (controversy) ในเรื่องดังกล่าวอยู่ เนื้อหาในลักษณะนี้ไม่ควรนำมาออกสอบด้วยข้อสอบปรนัย เนื่องจากในขณะที่ทำข้อสอบอยู่นั้น ผู้สอบไม่มีทางรู้ได้เลยว่าอาจารย์ผู้ออกข้อสอบอ้างอิงจากตำราหรือบทความวิชาการใด เนื้อหาที่ยังมีความขัดแย้ง ที่ผู้เชี่ยวชาญจากต่างสถาบันมีแนวทางในการปฏิบัติที่ต่างกันในแนะนำให้ใช้ข้อสอบในรูปแบบอื่นในการทดสอบเช่นข้อสอบอัตนัย เป็นต้น

๑.๔ หลีกเลียงการลอกประโยคหรือข้อความจากตำราโดยตรง

ดังได้กล่าวแล้วว่าข้อสอบที่ดีควรมุ่งเน้นการประเมินความเข้าใจ หรือ การประยุกต์ใช้ความรู้ ไม่ควรออกข้อสอบที่ประเมินความสามารถในการจำรายละเอียดปลีกย่อย การออกข้อสอบโดยวิธีการเปิดตำราแล้วคัดลอกประโยคจากตำราโดยตรงมักจะลงเอยด้วยข้อสอบที่ทดสอบความจำว่าผู้สอบท่องเนื้อหาในตำราตรงส่วนนั้นได้หรือไม่

ข้อสอบที่ดีควรได้จากการดูผู้ป่วย ใจหายที่ดีควรเป็นปัญหาของผู้ป่วยที่พบในการทำงานนั่นเอง ตัวเลือกก็ได้จากข้อผิดพลาดที่นักศึกษาหรือแพทย์ประจำบ้านมักปฏิบัติกับผู้ป่วยแล้วทำให้ผลการรักษาไม่ดีขึ้นเอง

๑.๕ หลีกเลียงการนำเสนอข้อสอบที่ประเมินความรู้ในเรื่องเดียวกันสองข้อในข้อสอบชุดเดียวกัน

เนื่องจากเนื้อหาวิชาที่ต้องทำการประเมินในการสอบแต่ละครั้งนั้นมีมาก ดังนั้นองค์ความรู้ในแต่ละเรื่องแต่ละโรคจึงมักมีสัดส่วนของข้อสอบที่จะออกได้เพียงหนึ่งหรือสองข้อเท่านั้น การที่อาจารย์ออกข้อสอบในเรื่องหรือโรคเดียวกันซ้ำสองข้อในชุดข้อสอบเดียวกันจึงมักเป็นการลดโอกาสในการประเมินความรู้เรื่องอื่นซึ่งก็มีความสำคัญเช่นกัน การออกข้อสอบที่ตื้นนั้นควรต้องครอบคลุมวัตถุประสงค์การเรียนรู้ตามที่กำหนดในหลักสูตร หรือในเกณฑ์มาตรฐานผู้ประกอบวิชาชีพเวชกรรมอย่างสมดุล การที่จะบรรจุเป้าหมายดังกล่าวได้นั้นต้องเริ่มต้นจากการกำหนดสัดส่วนข้อสอบสร้างเป็นตารางกำหนดจำนวนข้อสอบ (table of specification) เมื่ออาจารย์ได้รับมอบหมายให้ออกข้อสอบควรต้องตรวจสอบให้ชัดเจนว่าเนื้อหาที่ต้องออกข้อสอบนั้นอยู่ในส่วนใดของตารางดังกล่าว การออกข้อสอบซ้ำซ้อนในเนื้อหาเรื่องเดียวกันเป็นสัญญาณบอกว่าอาจไม่ได้สร้างข้อสอบตามข้อกำหนดในตาราง นอกจากนี้การมีใจหายสองข้อประเมินความรู้เรื่องเดียวกันมีความเป็นไปได้สูงที่เนื้อหาในข้อสอบข้อหนึ่งอาจบอกคำตอบในข้อสอบอีกข้อหนึ่งได้

๒. การจัดรูปแบบข้อสอบ

๒.๑ เลือกใช้คำศัพท์หรือรูปประโยคที่ง่ายต่อการทำความเข้าใจ

อาจารย์ผู้ออกข้อสอบต้องระลึกไว้เสมอว่าข้อสอบที่อาจารย์ออกเพื่อใช้ในการประเมินผลนักศึกษาแพทย์หรือแพทย์ประจำบ้านนั้นมีวัตถุประสงค์เพื่อทดสอบความรู้ทางการแพทย์เป็นสำคัญ มิใช่การประเมินความรู้ทางภาษาศาสตร์ ดังนั้นการเขียนข้อสอบของอาจารย์ควรเลือกใช้รูปแบบประโยคที่ง่ายต่อการทำความเข้าใจ อย่าเขียนประโยคซับซ้อนที่มีความยาวประโยคหลายบรรทัด มุ่งเน้นให้ภาษาเป็นสื่อในการนำเสนอความคิดของอาจารย์ผู้ออกข้อสอบไปยังผู้สอบ อย่าให้

เวบบันทึกศิริราช

บทความทั่วไป

ภาษาเป็นอุปสรรคในการสื่อสาร การจะเลือกใช้ภาษาใดในการเขียนข้อสอบนั้นให้พิจารณาตามข้อกำหนดขององค์กรหรือหน่วยงานที่ควบคุมการสอบที่อาจารย์ส่งข้อสอบไปให้ใช้ ข้อสอบที่ใช้ในระดับการศึกษาหลักสูตรแพทยศาสตรบัณฑิตทั้งในระดับคณะ หรือข้อสอบที่ใช้ในการสอบระดับประเทศในปัจจุบันยังนิยมใช้ข้อสอบที่เขียนด้วยภาษาไทยโดยมีการใช้ศัพท์เทคนิคเป็นภาษาอังกฤษเหมือนดังภาษาที่แพทย์ใช้สื่อสารกันในการทำงานปกติ ส่วนข้อสอบในระดับหลังปริญญาตรีมีหลายการสอบที่ภาควิชา หรือราชวิทยาลัยที่เกี่ยวข้องกำหนดให้ใช้ภาษาอังกฤษทั้งหมด ก่อนที่อาจารย์จะสร้างข้อสอบต้องมีการศึกษาข้อกำหนดของแต่ละการสอบให้ดี

๒.๒ หลีกเลี่ยงการนำเสนอข้อมูลที่ไม่เกี่ยวข้องกับการแก้ปัญหาของโจทย์ข้อนั้น

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

๒.๓ จัดให้มีการตรวจสอบเนื้อหา คำศัพท์ และรูปประโยคที่ใช้ในข้อสอบแต่ละข้อก่อนนำไปใช้

ถึงแม้ว่าอาจารย์ผู้เขียนข้อสอบจะได้มีการอ่านทวนสิ่งที่ตนเองเขียนแล้วเข้าใจเนื้อหาได้ดีและคิดว่าข้อสอบอยู่ในรูปแบบที่สามารถนำไปใช้ได้แล้ว ก็ไม่ควร

นำข้อสอบข้อนั้นไปใช้สอบเลย ควรให้มีคณะกรรมการข้อสอบซึ่งประกอบไปด้วยอาจารย์หลายท่านช่วยกันตรวจสอบและพิจารณาปรับแก้ข้อสอบทุกข้อก่อนนำไปใช้จริงเสมอ เนื่องจากผู้เขียนข้อสอบย่อมเข้าใจสิ่งที่ตนเขียนเสมอ แต่เมื่อผู้อื่นอ่านแล้วอาจพบว่ามีเนื้อหาที่กำกวมหรือเข้าใจโจทย์ต่างออกไปได้ การปรับแก้เนื้อหาที่มีความกำกวม หรือเฉลยซึ่งอาจารย์บางท่านอาจไม่เห็นด้วยให้ได้ข้อสอบที่มีความชัดเจน และอาจารย์ทุกท่านยอมรับในค่าเฉลยได้ก่อนจะนำข้อสอบไปทำการสอบจริงย่อมเป็นสิ่งที่ดีกว่าการตรวจพบปัญหาหลังจากสอบเสร็จแล้วซึ่งต้องมาตัดสินใจกันอีกว่าจะทำอย่างไรกับการคิดคะแนนของข้อสอบข้อดังกล่าว

๓. การเขียนโจทย์

๓.๑ เขียนโจทย์ให้มีความชัดเจน ผู้สอบทุกคนอ่านแล้วมีความเข้าใจตรงกัน

ข้อแนะนำนี้อาจดูเหมือนตรงไปตรงมา แต่กลับเป็นปัญหาที่พบบ่อยมากในการพัฒนาข้อสอบปรนัยประเด็นสำคัญคือโจทย์ที่ดีนั้นต้องมีความสมบูรณ์ในตัวเองโดยไม่ต้องอาศัยตัวเลือก โจทย์ข้อสอบที่ดีนั้นเมื่ออ่านโจทย์เสร็จแล้ว หากผู้สอบมีความรู้ในเรื่องที่ทำการประเมินนั้น เขาจะบอกคำตอบได้โดยไม่จำเป็นต้องอ่านตัวเลือกเลย ดังนั้นเมื่ออาจารย์เขียนข้อสอบเสร็จแล้วแนะนำให้ลองปิดตัวเลือกแล้วอ่านเฉพาะโจทย์ดู หากอาจารย์อ่านแล้วบอกได้ว่าโจทย์ถามอะไรและบอกได้ว่าควรตอบอะไรโดยไม่ต้องอ่านตัวเลือกจัดว่าข้อสอบข้อดังกล่าวมีโจทย์ที่มีความชัดเจน

๓.๒ เรียบเรียงเนื้อหาให้ใจความสำคัญของข้อสอบอยู่ในโจทย์

เนื่องจากข้อสอบปรนัยมีตัวเลือกที่อาจารย์ต้องสร้างขึ้นหลายตัวเลือก บางครั้งอาจารย์ผู้พัฒนาข้อสอบอาจเผลอเรอณาเอาใจความสำคัญไปใส่ไว้ในตัวเลือกซึ่งทำให้เนื้อหาในโจทย์ขาดสาระสำคัญ อ่านโจทย์แล้วไม่เข้าใจว่าผู้ออกข้อสอบต้องการถามความรู้เรื่องอะไร ตัวอย่างข้อสอบที่ไม่เป็นไปตามข้อแนะนำนี้คือข้อสอบที่ถามว่า ข้อใดต่อไปนี้เป็นข้อที่ต้อง หรือข้อใดต่อไปนี้เป็นข้อที่ไม่ถูกต้องแล้วเขียนรายละเอียดเกี่ยวกับโรค หรือการรักษาบางอย่างในตัวเลือกแต่ละข้อ ข้อสอบในลักษณะนี้มักทำให้

เวบบ์ทีกศิริราช

บทความทั่วไป

ผู้สอบต้องอ่านข้อสอบย้อนไปมาหลายรอบกว่าจะเข้าใจจุดประสงค์ของข้อสอบ แล้วจึงตัดสินใจเลือกคำตอบโดยทั่วไปแนะนำให้อาจารย์นำเสนอรายละเอียดต่าง ๆ ไว้ในตัวอย่างให้มากที่สุด ส่วนตัวเลือกเขียนเป็นคำหรือข้อความสั้น ๆ

๓.๓ หลักเขียนการเขียนโจทย์ที่มีรูปประโยคเป็นเชิงปฏิเสธ

โจทย์ที่ดีไม่ควรอยู่ในประโยคเชิงปฏิเสธ เช่นถามถึงสิ่งที่เป็นข้อยกเว้น สิ่งที่ไม่ควรปฏิบัติ สิ่งที่มีน้อยที่สุด หรือสิ่งที่ไม่แน่นอนถึงเป็นต้น งานวิจัยส่วนใหญ่พบว่าข้อสอบที่มีโจทย์ในรูปแบบปฏิเสธเหล่านี้มีระดับความยากง่ายไม่ต่างจากข้อสอบอื่น ๆ แต่งานวิจัยบางชิ้นพบว่าข้อสอบที่มีโจทย์ในรูปแบบปฏิเสธมีความยากมากกว่าข้อสอบอื่นชัดเจนโดยเฉพาะในข้อสอบวัดความรู้ระดับสูง^{๑๐-๑๒} แต่ผู้เชี่ยวชาญในการประเมินผลส่วนใหญ่มีความเห็นพ้องกันว่าข้อสอบประเภทนี้สามารถสร้างความสับสนให้กับผู้สอบได้ จึงไม่แนะนำให้ใช้ แต่หากอาจารย์ผู้ออกข้อสอบมีความจำเป็นต้องใช้ข้อสอบที่มีการใช้คำปฏิเสธในโจทย์ แนะนำให้พิมพ์คำปฏิเสธให้เด่นชัด โดยใช้ตัวหนาและขีดเส้นใต้เพื่อให้ผู้สอบเห็นชัด^{๑๑}

๔. การเขียนตัวเลือก

๔.๑ เขียนตัวเลือกที่มีประสิทธิภาพให้มีจำนวนมากที่สุดเท่าที่เหมาะสมกับบริบท

เรื่องจำนวนตัวเลือกที่เหมาะสมนี้เป็นเรื่องที่คุณเชี่ยวชาญด้านการประเมินผลจำนวนมากสนใจ มีงานวิจัยเกี่ยวกับเรื่องจำนวนตัวเลือกที่เหมาะสมในข้อสอบปรนัยอยู่มากมาย^{๑๓} อาจารย์ผู้ออกข้อสอบส่วนมากจะคุ้นเคยกับข้อสอบปรนัยชนิดที่มีห้าตัวเลือก บ่อยครั้งที่อาจารย์ออกข้อสอบแล้วนึกตัวเลือกได้เพียงสามหรือสี่ตัว จึงเกิดคำถามว่าจำเป็นต้องมีตัวเลือกครบห้าตัวเลือกหรือไม่ งานวิจัยบางชิ้นพบว่าการลดจำนวนตัวเลือกลงทำให้ข้อสอบง่ายขึ้น^{๑๓-๑๔} แต่งานวิจัยบางชิ้นพบว่าการลดจำนวนตัวเลือกลงทำให้ได้ข้อสอบยากขึ้น^{๑๕-๑๖} ผู้เชี่ยวชาญในการประเมินผลเสนอว่าข้อสอบปรนัยที่มีตัวเลือกเพียงสามตัวเลือกก็สามารถทดสอบความรู้ได้อย่างมีประสิทธิภาพ^{๑๗-๑๙} แต่มีอาจารย์จำนวนไม่น้อยที่ไม่สบายใจที่มีตัวเลือกในข้อสอบแต่ละข้อน้อยกว่าห้าตัว

เลือกด้วยกังวลว่าจะทำให้มีโอกาสสูงที่ผู้สอบที่ไม่มีความรู้จะเดาสุ่มได้คำตอบที่ถูกต้อง แต่จากข้อมูลที่ปรากฏในปัจจุบันพบว่าผู้สอบในการสอบในระดับสูงนั้นพฤติกรรมเดาสุ่มโดยที่ผู้สอบปราศจากความรู้นั้นน่าจะมีบทบาทน้อยมาก ผู้สอบส่วนใหญ่มักพอมีความรู้บ้างและสามารถตัดตัวเลือกที่ไม่สมเหตุสมผลอย่างชัดเจนได้^{๑๐} ในการศึกษาข้อสอบปรนัยส่วนใหญ่พบตัวเลือกที่ไม่ทำงานเป็นจำนวนไม่น้อย^{๑๐} ข้อมูลที่ได้จากการวิเคราะห์ข้อสอบปรนัยที่ใช้ในทางแพทยศาสตรศึกษาในประเทศไทยหลายครั้งก็สอดคล้องกับงานวิจัยในต่างประเทศที่พบว่าข้อสอบส่วนใหญ่มักมีตัวเลือกที่ทำงานจริงราวสามหรือสี่ตัวเลือก มีข้อสอบน้อยข้อมากที่ตัวเลือกทั้งห้าตัวเลือกทำงานอย่างมีประสิทธิภาพ

ด้วยข้อมูลจากการศึกษาต่างๆ 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

เวบบันทึทศิริราช

บทความทั่วไป

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 ไม่ใช่กริยา 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 ซึ่งข้อสังเกตนี้ผู้สอบจำนวนมากไม่น้อยก็ทราบดี และผู้สอบส่วนมากเมื่อไม่ทราบคำตอบก็มักเลือกตัวเลือกที่มีความยาวมากที่สุด ดังนั้นอาจารย์ผู้ออกข้อสอบควรระมัดระวังไม่ให้ตัวเลือกตัวใดตัวหนึ่งมีความยาวแตกต่างไปจากตัวเลือกอื่นชัดเจน เพราะจะทำให้ผู้สอบเดาคำตอบที่ถูกได้ง่าย

๔.๖ หลีกเลี่ยงการใช้ตัวเลือก “ถูกทุกข้อ” หรือ “ไม่มีข้อใดถูก”

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

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

การใช้ตัวเลือก “ไม่มีข้อใดถูก” จะยังเป็นปัญหามากขึ้นในข้อสอบที่ถามถึงสิ่งที่ไม่ควรทำ เช่นยาใดไม่ควรใช้ในผู้ป่วย ซึ่งนอกจากยาที่น่าเสนอในตัวเลือกแล้วย่อมมียาชนิดอื่นอีกมากมายในบัญชียาที่ไม่เหมาะสม ซึ่งไม่มี

เวบบ์ทีกศิริราช

บทความทั่วไป

ทางที่ใครจะรู้ได้ว่าการที่ผู้สอบเลือกตอบ “ไม่มีข้อใดถูก” นั้นเขาคิดถึงยาใด และยานั้นไม่เหมาะสมมากไปกว่ายาที่มีอยู่ในตัวเลือกหรือไม่ งานวิจัยทั้งหมดที่ศึกษาถึงตัวเลือกชนิดนี้ได้ข้อสรุปที่ตรงกันว่าข้อสอบที่ใช้ตัวเลือกประเภทนี้เพิ่มระดับความยากให้ข้อสอบ^{๑๖} โดยทั่วไปแล้วจึงไม่แนะนำให้ใช้ตัวเลือกประเภทนี้ในการสอบทางแพทยศาสตรศึกษาซึ่งทางเลือกสำหรับสถานการณ์ที่น่าเสนาหามีได้มากและการตัดสินใจเลือกคำตอบต้องอาศัยการเปรียบเทียบข้อดีข้อเสียของแต่ละตัวเลือก

สรุป

ในบทความนี้ผู้พิมพ์ได้กล่าวถึงข้อแนะนำขั้นพื้นฐานในการพัฒนาข้อสอบปรนัยชนิดเลือกคำตอบที่ถูกที่สุดโดยสรุปข้อแนะนำเหล่านี้ออกเป็นสี่กลุ่มด้วยกัน ได้แก่ (๑) เนื้อหาข้อสอบ, (๒) การจัดรูปแบบข้อสอบ, (๓) การเขียนโจทย์, และ (๔) การเขียนตัวเลือก ผู้พิมพ์หวังว่าข้อแนะนำเหล่านี้คงพอเป็นแนวทางสำหรับอาจารย์แพทย์ในการพัฒนาข้อสอบปรนัยที่มีคุณภาพเพื่อใช้ในการประเมินนักศึกษาแพทย์และแพทย์ประจำบ้านได้บ้าง อย่างไรก็ตามบทความนี้เป็นการกล่าวถึงข้อแนะนำเบื้องต้นเท่านั้น ยังมีข้อแนะนำอื่น ๆ ที่ผู้พิมพ์ไม่ได้นำมารวบรวมไว้ในบทความนี้เพื่อต้องการทำให้เนื้อหากระชับโดยข้อแนะนำอื่น ๆ ที่ผู้พิมพ์ไม่ได้กล่าวถึงนี้พบว่าเป็นปัญหาน้อยในการออกข้อสอบทางการแพทย์ หรือเป็นข้อแนะนำที่ไม่ได้รับการสนับสนุนอย่างกว้างขวางจากผู้เชี่ยวชาญทางการวัดและประเมินผล หากผู้อ่านสนใจรายละเอียดของข้อแนะนำอื่น ๆ ที่มีผู้กล่าวไว้สามารถศึกษาเพิ่มเติมได้จากเอกสารอ้างอิงที่แสดงไว้ท้ายบทความ

มีข้อควรพิจารณาในการประยุกต์ใช้ข้อแนะนำเหล่านี้ในการพัฒนาข้อสอบที่ผู้พิมพ์ขอกกล่าวถึงประการหนึ่งคือ แม้ว่าข้อแนะนำที่กล่าวถึงเหล่านี้หลายข้อมีการศึกษาวิจัยสนับสนุนที่ชัดเจน แต่สิ่งเหล่านี้ก็เป็นเพียงข้อแนะนำว่าผู้ออกข้อสอบควรปฏิบัติ ไม่ใช่กฎเกณฑ์ตายตัว การเขียนข้อสอบปรนัยนั้นเป็นงานที่ต้องอาศัยทั้งศาสตร์และศิลปะผสมผสานกันอย่างเหมาะสม

หาใช้สูตรคณิตศาสตร์ที่ไม่มีข้อยกเว้น ผู้พิมพ์ไม่คาดหวังให้อาจารย์ผู้พัฒนาข้อสอบยึดข้อแนะนำเหล่านี้เสมือนกฎเกณฑ์ตายตัวที่ต้องทำตามในทุกกรณี หากแต่ต้องการให้อาจารย์ใช้เป็นแนวทางในการสร้างข้อสอบ ในบางบริบทผู้ออกข้อสอบอาจเลือกที่จะไม่ปฏิบัติตามข้อแนะนำบางประการได้บ้าง แต่การที่จะไม่ปฏิบัติตามข้อแนะนำเหล่านี้มันต้องมีเหตุผลที่เหมาะสม และควรทำไม่บ่อยนัก ยกตัวอย่างเช่นข้อแนะนำว่า โจทย์ไม่ควรเขียนถามข้อยกเว้น จะพบได้ว่ามีบางบริบทที่การรู้ข้อยกเว้น หรือข้อห้ามปฏิบัติก็เป็นองค์ความรู้ที่สำคัญในการดูแลรักษาผู้ป่วย ดังนั้นในบริบทที่เหมาะสมผู้พิมพ์เองก็เห็นด้วยว่าอาจเขียนโจทย์ที่ถามข้อยกเว้นได้ แต่อย่างไรก็ตามการจะไม่ปฏิบัติตามข้อแนะนำนี้ต้องไม่ทำบ่อยจนเกินจำเป็น หากออกข้อสอบ ๑๐๐ ข้อ จะมีข้อสอบที่ถามข้อยกเว้น ประมาณบ้าง ๒-๓ ข้อ ย่อมเป็นสิ่งที่ยอมรับได้ แต่หากในชุดข้อสอบมีข้อสอบถึงร้อยละ ๒๐ – ๓๐ ที่โจทย์เขียนในรูปประโยคปฏิเสธ ถามสิ่งที่ไม่ควรปฏิบัติ หรือสิ่งที่ไม่ถูกต้อง อย่างนี้ย่อมจัดว่าละเอียดแนวทางในการพัฒนาข้อสอบอย่างไม่เหมาะสม ซึ่งย่อมส่งผลให้คุณภาพของข้อสอบด้อยลงอย่างชัดเจน

เอกสารอ้างอิง

- Downing SM. Assessment of knowledge with written test forms. In: Norman GR, van der Vleuten C, Newble DI, editors. International handbook of research in medical education. Dordrecht: Kluwer Academic Publishers, 2002:647 - 72.
- Haladyna TM, Downing SM. A taxonomy of multiple-choice item-writing rules. Appl Meas Educ 1989;2:37-50.
- Haladyna TM. Developing and validating multiple-choice test items, 3rd ed. Mahwah, NJ: Lawrence Erlbaum Associates, 2004.
- 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.
- Jozefowicz RF, Koepfen BM, Case S, Galbraith R, Swanson D, Glew RH. The quality of in-house medical school examinations. Acad Med 2002;77(2):156-61.
- 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.

เวบบ์นทักศึรราช

บทความทัวไป

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.

ข้อผิดพลาดที่ควรระวังในการสร้าง ข้อสอบปรนัย

ผู้ช่วยศาสตราจารย์ นายแพทย์เชดศักดิ์ โอบนรัตน

ภาควิชาศัลยศาสตร์, คณะแพทยศาสตร์ศิริราชพยาบาล, มหาวิทยาลัยมหิดล, กรุงเทพมหานคร ๑๐๓๐๐.

ข้อผิดพลาดที่ควรระวังในการสร้างข้อสอบปรนัย

ข้อสอบปรนัย (multiple-choice question) เป็นรูปแบบการประเมินผลที่นิยมใช้กันอย่างแพร่หลายในวงการแพทยศาสตรศึกษา ข้อสอบชนิดนี้เป็นที่ชื่นชอบของนักศึกษาผู้เข้าสอบจำนวนมากเนื่องจากมีคำตอบให้เลือก หากไม่มีความรู้ก็สามารถเดาได้ ซึ่งต่างไปจากข้อสอบประเภทอัตนัยซึ่งผู้สอบต้องเขียนคำตอบจากความคิดของตนเอง^๑ ดังนั้นข้อสอบปรนัยจึงเป็นข้อสอบที่ผู้สอบทำได้ง่าย แต่ในทางตรงข้ามข้อสอบปรนัยเป็นข้อสอบที่สร้างปัญหาให้กับอาจารย์ผู้สร้างข้อสอบไม่น้อย เนื่องจากในกระบวนการเขียนข้อสอบปรนัยแต่ละข้อนั้นต้องใช้ทักษะอย่างมาก ต้องใช้ทั้งศาสตร์และศิลป์ และบ่อยครั้งอาจารย์ผู้สร้างข้อสอบก็ถูกขอให้ทำการปรับแก้ข้อสอบเนื่องจากคณะกรรมการพิจารณาข้อสอบมีความเห็นว่ารายละเอียดในข้อสอบไม่เหมาะสม มีการศึกษาวิจัยพบว่าคุณภาพของข้อสอบปรนัยที่พัฒนาขึ้นในโรงเรียนแพทย์หลายแห่งนั้นไม่สู้ดีนัก มีข้อสอบที่มีลักษณะไม่เหมาะสมอยู่จำนวนมาก^{๒-๓} ข้อสอบปรนัยที่มีลักษณะไม่เหมาะสมเหล่านี้ส่งผลเสียต่อการสอบได้หลายประการ เช่น ทำให้ข้อสอบยากขึ้น สร้างความสับสนให้ผู้สอบ ทำให้ผู้สอบบางกลุ่มเสียเปรียบและทำให้การตัดสินผลสอบผิดพลาด เป็นต้น^{๓-๕} ดังนั้นการออกข้อสอบปรนัยที่มีคุณภาพดีจึงเป็นงานที่มีความสำคัญและท้าทายความสามารถ

การสร้างข้อสอบปรนัยที่มีคุณภาพดีนั้นควรเริ่มต้นจากการมีองค์ความรู้พื้นฐานในการสร้างข้อสอบแล้ว เกิดการฝึกฝนทักษะ สังเกตประสบการณ์ในการออกข้อสอบ จนเกิดความชำนาญ ปัญหาที่พบบ่อยในโรงเรียนแพทย์หลายแห่งคือมีอาจารย์จำนวนไม่น้อยที่ได้รับมอบหมายให้ออกข้อสอบปรนัย โดยไม่ได้มีการพัฒนาองค์ความรู้พื้นฐานที่เหมาะสมก่อน ซึ่งเป็นเหตุให้มีข้อสอบปรนัยที่มีลักษณะไม่เหมาะสมตามหลักการออกข้อสอบปะปนมาในข้อสอบที่ให้นักศึกษาแพทย์และแพทย์ประจำบ้านทำอยู่บ้าง ผู้นิพนธ์จึงเห็นความสำคัญของการเผยแพร่องค์ความรู้พื้นฐานของการออกข้อสอบปรนัย องค์ความรู้พื้นฐานในการสร้างข้อสอบปรนัยนั้นมีสองส่วน ส่วนแรกเป็นหลัก การของการสร้างข้อสอบทั่วไปซึ่งได้มีผู้รวบรวมเป็นข้อแนะนำตีพิมพ์ในตำราและวารสารทางวิชาการอยู่บ้าง^{๑, ๕-๗} ส่วนที่สองเป็นข้อผิดพลาดในการสร้างข้อสอบที่อาจารย์ผู้ออกข้อสอบพึงหลีกเลี่ยง ในบทความนี้ผู้นิพนธ์จะมุ่งเน้นในส่วนที่สองนี้ โดยจะรวบรวมข้อผิดพลาดในการสร้างข้อสอบปรนัย ที่อาจเป็นตัวบอกรับให้ผู้สอบที่ไม่มีความรู้ในเรื่องที่ทำการทดสอบสามารถเลือกคำตอบที่ถูกต้องได้ ดังนั้นการที่อาจารย์ผู้ออกข้อสอบทราบถึงสิ่งเหล่านี้และหลีกเลี่ยงเสียจะส่งผลให้ข้อสอบปรนัยที่สร้างขึ้นสามารถใช้วัดองค์ความรู้ทางการแพทย์ได้จริง โดยปราศจากปัจจัยรบกวนจากการสังเกตพบสิ่งบอกรับคำตอบ

๓/๓๗

กรกฎาคม-ธันวาคม ๒๕๕๕, ปีที่ ๕, ฉบับที่ ๒

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

ในบทความนี้ผู้พิมพ์ขอนำเสนอข้อผิดพลาดในการออกข้อสอบ ๗ กลุ่มด้วยกัน ได้แก่ (๑) ข้อผิดพลาดในไวยากรณ์, (๒) การไปคำตอบด้วยหลักตรรกะ, (๓) การใช้คำคุณศัพท์บอกระดับของความแน่ชัด, (๔) ความยาวของตัวเลือก, (๕) การใช้คำซ้ำในโจทย์และตัวเลือก, (๖) การเข้าพวกของคำ หรือข้อความที่ปรากฏในตัวเลือก, และ (๗) การบอกรับคำตอบโดยโจทย์ข้ออื่น

๑. ข้อผิดพลาดในไวยากรณ์

ตัวเลือกทุกตัวต้องสามารถตอบโจทย์ได้อย่างถูกต้องตามหลักไวยากรณ์ บ่อยครั้งอาจารย์ผู้ออกข้อสอบมุ่งความสนใจไปที่คำตอบที่ถูก และให้ความสนอกสนใจกับตัวลวงน้อยไปจนทำให้ตัวลวงผิดหลักไวยากรณ์ 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 gonorrhoeae*
- B. *Chlamydia trachomatis* and *Giardia lamblia*
- C. *Treponema pallidum*
- D. *Ureaplasma urealyticum* and *Mycoplasma genitalium*

ในตัวอย่างที่ ๒ นี้โจทย์ถามหาเชื้อก่อโรค โดยให้รูปประโยคถามหาคำตอบที่เป็นเอกพจน์ ดังนั้นคำตอบที่ถูกต้องย่อมมีเชื้อก่อโรคตัวเดียว ผู้สอบที่มีทักษะการทำข้อสอบดีจะตัดตัวเลือก B และ D ได้โดยไม่ต้องใช้ความรู้ทางการแพทย์

๒. การไปคำตอบด้วยหลักตรรกะ

ในการเขียนตัวเลือก อาจารย์ผู้ออกข้อสอบต้องระมัดระวังไม่ให้ผู้สอบสามารถตัดตัวเลือกได้ด้วยหลักตรรกศาสตร์ เนื่องจากผู้สอบที่มีทักษะการทำข้อสอบดีจะสามารถพิจารณาความเป็นไปได้ของตัวเลือกต่าง ๆ และตัดตัวลวงที่ไม่มีทางเป็นไปได้ตามหลักของเหตุและผลออกไปได้โดยไม่ต้องอาศัยความรู้เรื่องที่ว่าอาจารย์ตั้งเป้าหมายว่าจะทดสอบ

ตัวอย่างที่ ๓. ภาวะไส้เลื่อนบริเวณขาหนีบ (inguinal hernia)

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

ในตัวอย่างที่ ๓ นี้อาจารย์ผู้ออกข้อสอบต้องการวัดความรู้เรื่องอุบัติการณ์ของไส้เลื่อนขาหนีบ แต่หาก

เวชบัณฑิตศิริราช

บทความทั่วไป

พิจารณาตามหลักตรรกศาสตร์แล้ว ตัวเลือก A, B, และ C เพียงสามตัวเลือกก็ครอบคลุมสิ่งที่เป็นไปได้ทั้งหมดแล้ว (เนื่องจากมนุษย์มีสองเพศ ภาวะใส่เลื้อนนี้หากไม่มีอัตราการเกิดเท่ากันในสองเพศแล้วก็ต้องมีเพศใดเป็นมากกว่าอีกเพศหนึ่ง) ดังนั้นผู้สอบที่มีทักษะการทำข้อสอบดีสามารถตัดตัวเลือก D และ E ได้โดยไม่ต้องมีความรู้เรื่องใส่เลื้อนเลย

๓. การใช้คำคุณศัพท์บอกระดับของความแน่ชัด

อาจารย์ผู้ออกข้อสอบพึงระมัดระวังการใช้คำคุณศัพท์ที่บ่งบอกถึงความแน่ชัดของข้อความ ซึ่งจะมีหลายระดับ โดยทั่วไปแล้วคำคุณศัพท์ที่แสดงความแน่ชัดมาก แสดงความมั่นใจมาก (เช่น always, never) มักไม่ถูกต้อง เนื่องจากในทางการแพทย์นั้นมีความไม่แน่นอนเกิดขึ้นเป็นประจำ ข้อความที่บอกเล่าถึงสิ่งที่เป็นไปได้โดยไม่ชี้ชัดลงไปว่าต้องเกิดขึ้นแน่นอน (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 นี้ผู้สอบต้องเลือกข้อความเกี่ยวกับใส่เลื้อนขาหนีบที่ถูกต้องหนึ่งข้อความ หากสังเกตดูทั้งสี่ข้อความมีการใช้คำคุณศัพท์บอกระดับของความแน่ชัดได้แก่ may (ตัวเลือก A), always (ตัวเลือก B), the only (ตัวเลือก C), inevitably (ตัวเลือก D) ซึ่งจะเห็นว่าตัวเลือก B, C, และ D เป็นข้อความที่แสดงความแน่ชัดว่าต้องเป็นแน่ ต้องใช่แน่นอน ไม่มีทางเลี่ยงได้ ข้อความทำนองนี้มีโอกาสสูงที่จะผิด ในทางตรงข้ามตัวเลือก A เป็นข้อความบอกว่ามีโอกาสเป็นไปได้โดยไม่ชี้ชัดว่าต้องเกิด

ผู้สอบที่มีทักษะการทำข้อสอบดีจะตัดตัวเลือก B, C, และ D ได้โดยไม่ต้องอาศัยความรู้ทางการแพทย์เลย

๔. ความยาวของตัวเลือก

มีการตั้งข้อสังเกตว่าอาจารย์แพทย์มักชอบสอบและอธิบายแม้กระทั่งในการสอบอาจารย์แพทย์หลายท่านก็ติดนิสัยรักการสอบนั้นมาด้วย ทำให้อาจารย์มักเขียนตัวเลือกที่ถูกต้องที่มีคำอธิบายประกอบอย่างครบถ้วนทำให้ตัวเลือกที่ถูกมักมีความยาวมากกว่าตัวลวง นักศึกษาผู้เข้าสอบจำนวนไม่น้อยรู้ถึงความจริงข้อนี้และมักเลือกตัวเลือกที่มีความยาวมากที่สุด หากเขาไม่สามารถหาคำตอบได้ด้วยความรู้ทางการแพทย์ที่เขา

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

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

ในตัวอย่างที่ ๕ นี้จะสังเกตเห็นว่าตัวเลือก C มีการอธิบายเหตุผลประกอบส่งผลให้มีความยาวมากกว่าตัวเลือกอื่นชัดเจน ลักษณะเช่นนี้จะเป็นการบอกใบ้ให้นักศึกษาเลือกตัวเลือกนี้

๕. การใช้คำซ้ำในโจทย์และตัวเลือก

การใช้คำเดียวกัน หรือคำที่มีความหมายเหมือนกันในโจทย์และตัวเลือก มักเป็นการบอกใบ้ว่าตัวเลือกลงกล่าวเป็นตัวเลือกที่ถูกต้อง

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

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

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

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

ในตัวอย่างที่ ๖ นี้โจทย์ถามถึง sacular theory ซึ่งหากแปลความหมายก็น่าจะเป็นเรื่องที่เกี่ยวข้องกับถุง (sac) ผู้สอบที่มีทักษะการทำข้อสอบดีจะหาตัวเลือกที่มีคำที่มีความหมายเกี่ยวกับถุง แล้วเลือกตัวเลือกดังกล่าวทันที ซึ่งในที่นี้จะพบคำว่า diverticulum ซึ่งมีความหมายว่าถุงในข้อ B การที่มีคำที่มีความหมายซ้ำกันเช่นนี้เป็นตัวบอกใบ้คำตอบที่อาจารย์ผู้ออกข้อสอบต้องตรวจตราให้ดีก่อนนำข้อสอบไปใช้

๖. การเข้าพวของคำ หรือข้อความที่ปรากฏในตัวเลือก

ข้อสอบจำนวนไม่น้อยนำเสนอรายการของหลายอย่างในตัวเลือก (เช่น ชื่อการตรวจค้นเพิ่มเติม ชื่อโรค ชื่อยา ฯลฯ) มีผู้เชี่ยวชาญในการประเมินผลตั้งข้อสังเกตว่าในข้อสอบเหล่านี้ตัวเลือกที่ถูกต้องมักมีลักษณะเข้าพวกับตัวเลือกอื่นมากที่สุด หากเป็นรายการของตัวเลือกที่ถูกก็คือข้อที่มีจำนวนรายการซ้ำกับตัวเลือกอื่นมากที่สุด^๔ ดังนั้นในการนำเสนอตัวเลือกอาจารย์ผู้ออกข้อสอบพึงระมัดระวังอย่าให้ตัวเลือกที่ถูกต้องมีลักษณะที่เข้าพวกันได้อย่างชัดเจน พยายามทำตัวลวงอื่นให้มีลักษณะเข้าพวกับตัวเลือกเดียวกับตัวเลือกที่ถูกต้อง

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) ส่วนโรคที่เหลือกกล่าวถึงโรคละครั้งเดียว ดังนั้นตัวเลือกที่มีพวมากที่สุดคือตัวเลือก A ซึ่งเป็นคำตอบที่ถูกต้อง

การเข้าพวของตัวเลือกที่ถูกต้องนั้นไม่จำเป็นต้องเป็นลักษณะของการมีจำนวน หรือความถี่ของคำมากที่สุดเพียงเท่านั้น อาจหมายรวมถึงการมีรูปร่างลักษณะ หรือความหมายคล้ายคลึงกันได้ด้วย

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

A. Amoxicillin PO

B. Ceftazidime IV + Amikacin IV

C. Amphotericin B IV + Ceftazidime IV

D. Cloxacillin IV + Metronidazole IV

ในตัวอย่างที่ ๘ นี้โจทย์ถามถึงยาที่ควรให้กับผู้ป่วย ในตัวเลือกสี่ตัวเลือกนี้มียาเกินข้อเดียว (A) ที่เหลือเป็นยาฉีดสองขนานควบกัน ดังนั้นตัวเลือกข้อ A ไม่เข้าพว จะถูกตัดทิ้งได้โดยง่าย ในบรรดา ยาฉีดจะเห็นว่ามียาต้านเชื้อราที่ไม่เข้าพว (ตัวเลือก C) ดังนั้นจะเหลือตัวเลือกที่นักศึกษาต้องคิดเลือกจริง ๆ เพียงตัวเลือก B กับ D ซึ่งหากดูกลุ่มยา ก็จะพบว่ายากลุ่ม Cephalosporin เข้าพวมากที่สุด ทำให้ผู้สอบที่มีทักษะการทำข้อสอบดีสามารถเลือกคำตอบที่ถูกต้อง (ตัวเลือก B) ได้โดยไม่ต้องมีความรู้เรื่องการรักษาผู้ป่วย febrile neutropenia

๗. การบอกใบ้คำตอบโดยโจทย์ข้ออื่น

ข้อผิดพลาดนี้เป็นข้อผิดพลาดที่ตัวผู้เขียนข้อสอบไม่ค่อยรู้ แต่ผู้ที่จะตรวจพบข้อผิดพลาดนี้คืออาจารย์ผู้เลือกข้อสอบไปใช้ เนื่องจากในการสอบแต่ละครั้งใช้ข้อสอบจำนวนมาก หากเลือกข้อสอบโดยไม่ระมัดระวังอาจมีข้อสอบสองข้อที่ถามเกี่ยวกับโรคหรือกลุ่มอาการเดียวกัน ซึ่งข้อมูลจากโจทย์ในข้อหนึ่งอาจเป็นตัวบอกใบ้คำตอบของข้อสอบอีกข้อได้ ดังนั้นเมื่อทำการเลือกข้อสอบเสร็จแล้วจัดหน้ากระดาษเข้ารูปเล่มข้อสอบแล้วอาจารย์ควรอ่านข้อสอบฉบับสมบูรณ์อีกครั้งหนึ่งหรือสองรอบก่อนส่ง

เวบบันทึกศิริราช

บทความทั่วไป

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 ซึ่งผู้สอบที่มีทักษะการทำข้อสอบดีสามารถย้อนกลับไปอ่านโจทย์ในข้อก่อนหน้านั้นแล้วได้ข้อมูลว่าผู้ป่วยที่น่าเสนอว่าเป็น febrile neutropenia มีไข้ขึ้น ๑๔ วันหลังได้ยาเคมีบำบัด ก็สามารถตอบข้อสอบข้อนี้ได้ง่าย

สรุป

ผู้นิพนธ์ได้รวบรวมข้อผิดพลาดในการสร้างข้อสอบปรนัยที่ผู้สอบอาจใช้เป็นแนวทางในการเลือกคำตอบที่ถูกต้องโดยไม่ต้องอาศัยความรู้ทางการแพทย์ที่อาจารย์ต้องการประเมินผล โดยเรียบเรียงเป็นเจ็ดกลุ่มข้อผิดพลาดด้วยกัน ผู้อ่านทุกท่านพึงตระหนักว่าสิ่งเหล่านี้ไม่ใช่หลักการทางวิทยาศาสตร์ที่ชัดเจนดังกฎทางคณิตศาสตร์หรือฟิสิกส์ หากแต่เป็นการรวบรวมข้อสังเกต

และคำแนะนำของผู้เชี่ยวชาญทางการวัดและประเมินผล จึงเป็นเพียงแนวทางเบื้องต้นในการพิจารณาตรวจสอบเนื้อหาของข้อสอบเท่านั้น การประยุกต์ใช้องค์ความรู้นี้คงต้องอาศัยศิลปะพอสมควรเพื่อที่จะได้ข้อสอบที่ดีสามารถวัดองค์ความรู้ทางการแพทย์ของนักศึกษาหรือแพทย์ประจำบ้านที่เข้าสอบได้ตามวัตถุประสงค์ของการสอบ

เอกสารอ้างอิง

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

June 2018

ปัญหาในการสร้างข้อสอบปรนัย

รศ.นพ. เชิดศักดิ์ ไอร่มณีรัตน์
ภาควิชาศัลยศาสตร์ คณะแพทยศาสตร์ศิริราชพยาบาล
มหาวิทยาลัยมหิดล

Common Problems

1. Objectives
2. Stem
3. Options

Objectives

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

Stem

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

Options

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

Questions & Comments

Cherdsak.ira@mahidol.ac.th

MCQ Item Analysis

รศ.นพ.เชิดศักดิ์ ไอรรมณีรัตน์

ภาควิชาศัลยศาสตร์

คณะแพทยศาสตร์ศิริราชพยาบาล

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.

Objectives

- เมื่อสิ้นสุดการอบรมแล้ว อาจารย์ผู้เข้าอบรมสามารถ
 - อธิบายผลการวิเคราะห์ข้อสอบ MCQ ที่ใช้บ่อยทางแพทยศาสตร์ ศึกษาได้อย่างถูกต้อง
 - นำผลการวิเคราะห์ข้อสอบไปเป็นแนวทางในการพัฒนาคุณภาพของข้อสอบ MCQ ในภาควิชาของตนได้
 - บอกถึงข้อควรระวัง และข้อจำกัดในการวิเคราะห์ผลการสอบ MCQ

Outline

- Item statistics
- Test statistics
- Applications
- Limitations

MCQ item analysis

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

MCQ item analysis

Item Statistics

Looking at individual test items

MCQ item analysis

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

MCO item analysis

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

MCO item analysis

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

MCO item analysis

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

MCO item analysis

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	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

MCO item analysis

12

Example 2

Number 145	Correct answer = 3				
P-VALUE = 0.79	PT BISERIAL =0.34				Total number of examinees
DISTRACTOR	1	2	3	4	5
N OF PEOPLE	7	27	190	9	242
MEAN SCORE	77	78.11	85.81	78.22	75.89
P-VALUE	0.03	0.11	0.79	0.04	0.04
PT BISERIAL	-0.12	-0.21	0.34	-0.11	-0.16

MCO item analysis

13

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

Test Statistics

Looking at the whole test

MCQ item analysis

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

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

$$n = \text{number of testlets}$$

$$= \text{score variance of total scores}$$

$$= \text{score variance of the } i^{\text{th}} \text{ testlet}$$

MCQ item analysis

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

MCQ item analysis

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

MCQ item analysis

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.

MCQ item analysis

Applications

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

MCQ item analysis

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.

Summary

- Item statistics
- Test statistics
- Applications
- Limitations

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

Bill Gates

การวิเคราะห์ข้อสอบปรนัย

อาจารย์ นายแพทย์เชตศักดิ์ โอรมนิรัตน์

ภาควิชาศัลยศาสตร์, คณะแพทยศาสตร์ศิริราชพยาบาล, มหาวิทยาลัยมหิดล, กรุงเทพมหานคร ๑๐๓๑๐.

การวิเคราะห์ข้อสอบปรนัย (Item analysis) เป็นการใช้วิธีการทางสถิติเพื่อวิเคราะห์คำตอบที่ผู้สอบตอบข้อสอบปรนัยในการสอบครั้งหนึ่ง เพื่อประเมินว่าข้อสอบที่นำมาใช้ในการสอบครั้งนั้นมีคุณสมบัติอย่างไร ทำงานได้ตามที่ต้องการหรือไม่ มีระดับความยากง่ายของข้อสอบเหมาะสมหรือไม่ มีข้อบกพร่องหรือไม่ และควรได้รับการปรับปรุงแก้ไขอย่างไร การวิเคราะห์ข้อสอบเป็นศาสตร์ที่ได้รับการพัฒนาอย่างต่อเนื่องมาเป็นเวลานาน มีเทคนิคและวิธีการต่าง ๆ มากมายที่ผู้วิเคราะห์สามารถใช้เพื่อบอกคุณสมบัติของข้อสอบแต่ละข้อ ตั้งแต่วิธีการง่าย ๆ ไปจนถึงวิธีการที่มีความซับซ้อนมาก โดยแต่ละเทคนิคการวิเคราะห์ก็มีจุดประสงค์แตกต่างกันไป ตั้งแต่การบอกระดับความยากง่าย การบอกถึงความสามารถในการแยกผู้สอบที่เก่งออกจากผู้สอบที่ไม่เก่ง ไปจนถึงเทคนิคขั้นสูงที่สามารถบอกได้ว่าข้อสอบมีความลำเอียงต่อผู้สอบเพศใดเพศหนึ่ง หรือผู้สอบจากสถาบันใดสถาบันหนึ่งเป็นพิเศษหรือไม่ มีการเดาข้อสอบมากน้อยเพียงใด ผู้สอบรู้ข้อสอบมาก่อนเข้าสอบหรือไม่ หรือมีความน่าจะเป็นมากน้อยเพียงใดที่ผู้สอบลอกคำตอบ ในบทความนี้ผู้ประพันธ์ไม่ได้ตั้งเป้าประสงค์ที่จะรวบรวมและอภิปรายเทคนิคการวิเคราะห์ข้อสอบทุกวิธีที่มีใช้อยู่ในปัจจุบัน แต่ต้องการเพียงนำเสนอความรู้พื้นฐานที่เกี่ยวกับการวิเคราะห์ข้อสอบและอธิบายถึงวิธีการวิเคราะห์ข้อสอบที่นิยมใช้กันในทางแพทยศาสตรศึกษา โดยเฉพาะในประเทศไทย โดยประสงค์ให้อาจารย์ผู้อ่านสามารถนำเอาความรู้ที่ได้จากบทความนี้ไปใช้แปลผลการวิเคราะห์ข้อสอบที่ตน

เกี่ยวข้อง และดำเนินการปรับปรุงคุณภาพของข้อสอบได้อย่างเหมาะสม

ความรู้พื้นฐานเกี่ยวกับข้อสอบปรนัย

ก่อนที่จะกล่าวถึงรายละเอียดในการวิเคราะห์ข้อสอบ ผู้ประพันธ์ก็จะขอทบทวนความรู้พื้นฐานเกี่ยวกับข้อสอบปรนัยก่อน 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 ทฤษฎีการสอบแบบดั้งเดิมนี้อาศัยฐานอยู่บนสมมติฐานว่าคะแนนสอบที่ได้มานั้นประกอบไปด้วยคะแนนที่แท้จริง (true score) กับความผิดพลาดจากการวัด (error) ซึ่งสมมติฐานดังกล่าวต่อมาพบว่ามีข้อจำกัดหลายประการด้วยกัน ในราว ค.ศ. ๑๙๗๐ จึงได้มีความพยายามพัฒนาทฤษฎีที่ใช้ในการวิเคราะห์ข้อสอบแบบใหม่ขึ้นซึ่งให้หลักการของความน่าจะเป็นมาวิเคราะห์ข้อสอบ ทำให้สามารถแยกผลการวิเคราะห์ข้อสอบแต่ละข้อเป็นอิสระจากข้อสอบข้ออื่นในการสอบเดียวกัน ทฤษฎีใหม่นี้เรียกว่าทฤษฎีการตอบสนองต่อข้อสอบ (item response theory) ทฤษฎีใหม่นี้มีข้อได้เปรียบกว่าทฤษฎีเดิมหลายประการด้วยกัน ได้แก่ ความสามารถในการปรับตัวเข้ากับสถานการณ์ต่าง ๆ (flexibility) ความมีประสิทธิภาพในการใช้ข้อมูล (efficiency) และความสามารถในการวิเคราะห์ถึงคุณภาพของข้อสอบ และผู้สอบโดยละเอียด (in-depth analysis)^๓ จึงเป็นเหตุให้ทฤษฎีการตอบสนองต่อข้อสอบนี้ได้รับความนิยมอย่างกว้างขวางตั้งแต่ในค.ศ. ๑๙๙๐ ในปัจจุบันการสอบต่าง ๆ ได้ถูกวิเคราะห์ด้วยทฤษฎีการตอบสนองต่อข้อสอบนี้มากขึ้นเรื่อย ๆ

เนื่องจากการวิเคราะห์ข้อสอบในวงการแพทยศาสตรศึกษาในประเทศไทยทั้งหมดในปัจจุบันยังใช้เทคนิคต่าง ๆ ตามทฤษฎีการสอบแบบดั้งเดิมอยู่ ดังนั้นผู้นิพนธ์จะขอกล่าวถึงเทคนิคการวิเคราะห์ข้อสอบตามทฤษฎีการสอบแบบดั้งเดิมเท่านั้น เพราะจะเป็นสิ่งที่อาจารย์แพทย์ทุกท่านจะได้พบและใช้งานเป็นประจำ

การวิเคราะห์ข้อสอบตามทฤษฎีการสอบแบบดั้งเดิม

การวิเคราะห์ข้อสอบตามทฤษฎีการสอบแบบดั้งเดิมนี้อาศัยการประกอบไปด้วย ๒ ส่วนใหญ่ ๆ คือ (๑) การ

วิเคราะห์ข้อสอบรายข้อ (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 หากข้อสอบข้อนั้นเป็นข้อสอบที่ง่ายผู้สอบทุกคนตอบถูกค่า p ก็จะเป็น ๑ หากไม่มีผู้สอบคนใดตอบถูกเลยข้อสอบข้อนั้นก็จะมีค่า p เป็น ๐ หากมีคนตอบถูก ๗๐% ข้อสอบข้อนั้นก็จะมีค่า 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

๑.๒ ความสามารถในการจำแนกผู้สอบตามระดับความสามารถ (item discrimination, r)

ความสามารถในการจำแนกผู้สอบ หมายถึงความสามารถของข้อสอบข้อหนึ่ง ๆ ในการแยกผู้สอบที่ทำคะแนนได้ดี ออกจากผู้สอบที่ทำคะแนนได้ไม่ดี ข้อสอบที่มีความสามารถในการแยกแยะได้ดีนั้นผู้สอบที่ตอบข้อสอบข้อนั้นถูกมักจะได้คะแนนสูง และผู้สอบที่ตอบข้อสอบข้อนั้นผิดมักจะได้คะแนนต่ำ ดัชนีที่ใช้วัดความสามารถในการจำแนกผู้สอบที่ใช้กันมากที่สุดในปัจจุบันคือค่า point-biserial correlation ซึ่งนิยมใช้อักษรย่อเป็น $r^{๑,๕}$ ซึ่งสามารถคำนวณได้จากสูตรต่อไปนี้^๕

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

๓๓๓

นภกรณ-นภษน ๒๕๕๓, ปีที่ ๓๒, ฉบับที่ ๑

เวบบันทึทศิริราช

บทความทั่วไป

- 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 เป็นข้อสอบที่ไม่ผู้ติ นัก โดยเฉพาะอย่างยิ่งข้อสอบที่มีค่า point-biserial ต่ำกว่า ๐ ไม่ควรนำมาคิดคะแนน^{๑๖} (โดยทั่วไปแล้วข้อสอบที่มีค่า point-biserial ติดลบ ให้สงสัยว่าจะเฉลยผิด)

๑.๓ ประสิทธิภาพของตัวลวง (distractor functionality)

ตัวลวงที่มีประสิทธิภาพนั้นมีคุณสมบัติ ๒ ประการคือ^{๑๗}

(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 หรือลวงให้ผู้ที่มี

ความรู้ดีมาเลือกจัดเป็นตัวลวงที่ไม่ดี สมควรพิจารณาตัดทิ้งหรือปรับเปลี่ยน

๒. การวิเคราะห์ข้อสอบโดยรวม (test analysis)

การวิเคราะห์ข้อสอบโดยรวมเป็นการพิจารณาว่าเมื่อข้อสอบทั้งหมดทำงานร่วมกันแล้วผลสอบที่ได้ออกมาเป็นอย่างไร มีระดับความยากง่ายเป็นอย่างไร มีการกระจายตัวของคะแนนเป็นอย่างไร มีความน่าเชื่อถือของคะแนนสอบมากน้อยเพียงใด ดัชนีต่าง ๆ ที่ต้องพิจารณาได้แก่

๒.๑ ความเที่ยงตรงของคะแนนสอบ (internal consistency reliability)

การประเมินความเที่ยงตรงของคะแนนสอบเป็นการตรวจสอบว่าคะแนนที่ได้ออกมานั้นมีความน่าเชื่อถือเพียงใด เป็นการตอบคำถามว่าหากนำผู้สอบมาสอบใหม่ในสภาวะการณ์เดิม ด้วยข้อสอบที่มีระดับความยากง่ายเท่าเดิม และผู้สอบมีความรู้เท่าเดิมไม่ได้ไปศึกษาหาความรู้เพิ่มเติม จะได้คะแนนสอบเท่าเดิมหรือไม่^{๑๘}

ดัชนีชี้วัดความเที่ยงตรงของคะแนนสอบที่นิยมใช้ในการรายงานผลสอบด้วยข้อสอบปรนัยคือค่าสัมประสิทธิ์ อัลฟา (Coefficient Alpha) ซึ่งสามารถคำนวณได้จากสูตร^{๑๙}

$$\alpha = \frac{n}{n-1} \left(1 - \frac{\sum \sigma_{x_i}^2}{\sigma_x^2} \right)$$

เมื่อ α = สัมประสิทธิ์ อัลฟา (Coefficient Alpha)

n = จำนวนชุดย่อยของข้อสอบที่ทำการแบ่งออกเพื่อหาความเที่ยง

σ_x^2 = การกระจายตัว (variance) ของคะแนนรวม

$\sigma_{x_i}^2$ = การกระจายตัว (variance) ของคะแนนข้อสอบย่อยชุดที่ i

ค่าสัมประสิทธิ์อัลฟานี้มีค่าอยู่ในช่วง ๐ - ๑ ค่าต่ำแสดงว่าคะแนนที่ได้มีความเชื่อถือได้น้อย ไม่แตกต่างไปจากการเดาสุ่ม ค่าสูงแสดงว่าคะแนนที่ได้ี้มีความน่าเชื่อถือมาก หากทำการทดสอบซ้ำคะแนนที่ได้ี้ก็จะใกล้เคียงเดิม โดยทั่วไประดับของความเที่ยงตรง

เวบบันทึทศิริราช

บทความทั่วไป

ของคะแนนสอบที่ยอมรับได้นั้นขึ้นกับว่าต้องการนำเอาคะแนนสอบไปใช้ทำอะไร หากการตัดสินผลสอบนั้นมีความสำคัญมาก (high-stakes examination) เช่น การตัดสินผลสอบขอรับใบประกอบวิชาชีพเวชกรรม หรือประกาศนียบัตรแพทย์ผู้เชี่ยวชาญเฉพาะสาขา มักต้องการคะแนนสอบที่มีค่าสัมประสิทธิ์ อัลฟา ไม่ต่ำกว่า ๐.๙ หากการตัดสินผลสอบนั้นมีความสำคัญปานกลาง (medium-stakes examination) เช่นการสอบลงกอง การสอบเลื่อนชั้นเรียน มักต้องการคะแนนสอบที่มีค่าสัมประสิทธิ์ อัลฟา อยู่ในช่วง ๐.๘ - ๐.๘๙ หากการตัดสินผลสอบนั้นมีความสำคัญน้อย (low-stakes examination) เช่นการสอบย่อยในชั้นเรียน การสอบแบบ formative assessment มักต้องการคะแนนสอบที่มีค่าสัมประสิทธิ์ อัลฟา อยู่ในช่วง ๐.๗ - ๐.๗๙^{๑๒}

ประเด็นสำคัญที่ต้องพิจารณาคือเมื่อได้คะแนนสอบที่มีค่าสัมประสิทธิ์ อัลฟาต่ำ จะต้องดำเนินการอย่างไรเพื่อพัฒนาให้การสอบครั้งต่อไปไม่ประสบปัญหาเรื่องความไม่น่าเชื่อถือของคะแนนสอบอีก ปัจจัยหลักที่จะช่วยเพิ่มความเที่ยงตรงของคะแนนสอบปรนัยมี ๓ ปัจจัยด้วยกัน^{๑๓} คือ

(๑) เพิ่มจำนวนข้อสอบให้มากขึ้น ยังมีข้อสอบมากข้อคะแนนที่ได้ก็จะมีค่าสัมประสิทธิ์เพิ่มขึ้น

(๒) ปรับให้ข้อสอบมีการคละกันของข้อสอบที่ยากและง่ายอย่างเหมาะสม เพื่อปรับให้คะแนนมีการกระจายตัวมากขึ้น หากข้อสอบทั้งหมดประกอบด้วยข้อสอบที่ง่ายหมด ผู้สอบเกือบทั้งหมดได้คะแนนสูงมาก จะทำให้มีความแตกต่างของคะแนนน้อย โอกาสที่จะแยกแยะผู้สอบที่มีความรู้ดีออกจากผู้ที่มีความรู้ปานกลาง หรือไม่สู้ดีได้อย่างมั่นใจก็เป็นไปได้น้อย ดังนั้นหากอาจารย์ปรับให้มีการคละกันของข้อสอบยากและง่ายอย่างเหมาะสม ก็จะทำให้ผู้สอบมีระดับคะแนนแตกต่างกันมาก ค่าสัมประสิทธิ์อัลฟาก็จะสูงขึ้นด้วย

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

๒.๒ การกระจายตัวของคะแนน และคะแนน

เฉลี่ย (standard deviation and mean score)

การตรวจดูลักษณะพื้นฐานของคะแนนสอบนี้จะช่วยบอกได้คร่าว ๆ ว่าการเรียนการสอนมีประสิทธิภาพเพียงใด หากอาจารย์สอนได้ดี นักเรียนทั้งชั้นเรียนเข้าใจเนื้อหาดี คะแนนสอบที่ได้ออกมาก็ควรจะกระจายตัวมากนัก (คะแนนเกาะกลุ่มกัน) และคะแนนเฉลี่ยก็ควรจะค่อนข้างสูงเมื่อเทียบกับนักเรียนรุ่นอื่น ๆ หากคะแนนสอบของนักเรียนมีการกระจายตัวมากเกินไป แสดงว่าอาจมีปัญหาบางประการในการเรียนการสอนทำให้นักเรียนบางคนมีความรู้ความเข้าใจดี แต่มีนักเรียนบางกลุ่มที่ไม่ค่อยรู้เรื่อง^{๑๔}

๒.๓ ค่าความยากง่ายเฉลี่ยของข้อสอบ (average difficulty)

จากการวิเคราะห์ข้อสอบรายข้อ เราได้ค่าความยากง่ายของข้อสอบแต่ละข้อ (p) เมื่อนำค่า p ของข้อสอบทุกข้อมาหาค่าเฉลี่ย เราก็จะได้ค่าความยากง่ายของข้อสอบทั้งหมด ค่าที่ได้มานี้ใช้เป็นตัวชี้วัดว่าข้อสอบทั้งหมดโดยรวมแล้วมีระดับความยากง่ายเป็นอย่างไร หากผู้สอบเป็นนักศึกษาในกลุ่มใหญ่พอที่เราจะตั้งสมมติฐานว่าระดับความสามารถมีการกระจายตัวอย่างเหมาะสมและไม่ต่างจากระดับความสามารถเฉลี่ยของกลุ่มผู้สอบปีก่อน ๆ เราก็สามารถนำค่าความยากง่ายของข้อสอบทั้งหมดนี้มาเทียบได้ว่าข้อสอบที่นำมาใช้ในปีนี้อาจง่ายกว่าข้อสอบปีก่อน ๆ ซึ่งอาจารย์อาจนำข้อมูลนี้มาใช้พิจารณาปรับเกณฑ์การตัดเกรดด้วยว่าต้องมีการปรับระดับคะแนนที่ได้เกรดต่าง ๆ หรือไม่ อย่างไร

๒.๔ ค่าความสามารถในการแยกแยะผู้สอบเฉลี่ย (average discrimination)

การนำค่า point-biserial correlation ของข้อสอบทั้งหมดมาหาค่าเฉลี่ย เป็นการบอกคร่าว ๆ ว่าโดยรวมแล้วข้อสอบชุดนี้มีความสามารถในการแยกแยะผู้สอบตามระดับความสามารถเพียงใด ยิ่งได้ค่าสูงก็ยิ่งดี แต่มีข้อควรระวังในการแปลผลในกรณีที่มีการเรียนการสอนเป็นไปได้ดี และผู้สอบทั้งหมด หรือเกือบทั้งหมดทำคะแนนได้สูง ค่า point-biserial correlation เฉลี่ยของข้อสอบทั้งหมดจะไม่สูงแต่ไม่ได้แปลว่าข้อสอบที่ใช้มีคุณภาพไม่ดี^{๑๕}

เวบบิ้นทักสิริราช

บทความทั่วไป

การนำผลการวิเคราะห์ข้อสอบไปใช้

ผลการวิเคราะห์ข้อสอบด้วยดัชนีชี้วัดต่าง ๆ ดังกล่าวข้างต้นสามารถนำไปใช้ประโยชน์ได้หลายประการ เช่น

๑. ใช้เป็นประโยชน์ในการปรับแก้คะแนนสอบ

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

๒. ใช้เป็นประโยชน์ในการปรับปรุงคุณภาพข้อสอบ

ภายหลังจากการรายงานคะแนนสอบเป็นที่เรียบร้อยแล้ว คณะกรรมการสอบสามารถนำผลการวิเคราะห์ข้อสอบแต่ละข้อมาพิจารณาโดยละเอียดเพื่อดูว่าข้อสอบข้อใดสมควรได้รับการปรับปรุงแก้ไข ข้อสอบที่พบว่ายากเกินไปอาจเกิดจากโจทย์คำถามมีความคลุมเครือ ต้องทำการปรับแก้ให้โจทย์ชัดเจนขึ้น หรือเพิ่มเติมข้อมูลบางประการเข้าไปเพื่อให้การวินิจฉัย

ชัดเจนขึ้น ข้อสอบที่พบว่ายากเกินไปอาจพิจารณาปรับให้ยากขึ้นโดยการแก้ไขโจทย์หรือตัวเลือก ข้อสอบที่มีค่า point-biserial ต่ำมักเกิดจากโจทย์ที่คลุมเครือ สร้างความสับสนให้ผู้สอบ สมควรได้รับการปรับแก้โจทย์คำถามใหม่

นอกจากนี้อาจารย์ยังต้องพิจารณาถึงการ ทำงานของตัวเลือกด้วย ปัญหาที่พบบ่อยมากในการวิเคราะห์ข้อสอบปรนัยคือมีตัวลวงจำนวนมากที่ไม่ทำงาน (มีผู้สอบเลือกน้อยมาก หรือลวงเฉพาะผู้ที่มีความรู้ดีให้มาเลือก) จากการศึกษาวิจัยข้อสอบปรนัยจำนวนมากพบว่าข้อสอบส่วนใหญ่มักมีตัวเลือกที่ทำงานจริงเพียง ๓ ตัวเลือกเท่านั้น^๕ ตัวเลือกที่เหลือเป็นตัวลวงที่ไม่มีประโยชน์ พิมพ์ลงมาในข้อสอบก็เป็นการเปลืองเนื้อที่หน้ากระดาษ และเสียเวลาอ่านโดยใช่เหตุ อาจารย์ควรพิจารณาตัดตัวลวงที่ไม่ทำงานออกเสีย หรือเปลี่ยนเป็นตัวลวงอื่นที่น่าจะมีประสิทธิภาพมากขึ้น

๓. ใช้เป็นประโยชน์ในการบริหารคลังข้อสอบ

ข้อสอบแต่ละข้อนั้นได้มาด้วยความยากลำบาก อาจารย์แต่ละท่านต้องใช้เวลาและความคิดอย่างมากเพื่อพัฒนาข้อสอบที่ดีขึ้นมาใช้ ดังนั้นเมื่อนำข้อสอบมาใช้แล้วผลการวิเคราะห์ข้อสอบแสดงว่าข้อสอบข้อใดเป็นข้อสอบที่ดี มีระดับความยากง่ายเหมาะสม มีความสามารถในการจำแนกผู้สอบที่ดีก็ควร จะพิจารณาเลือกเก็บข้อสอบดังกล่าวไว้ในคลังข้อสอบ เพื่อที่จะได้นำกลับมาใช้ใหม่ในอนาคต ในการเก็บข้อสอบเข้าในคลังข้อสอบก็ต้องมีการแนบข้อมูลเกี่ยวกับประวัติการใช้งานและผลการวิเคราะห์ข้อสอบในแต่ละครั้งไว้คู่กันด้วย เพื่อที่จะได้เป็นประโยชน์ในการเลือกข้อสอบมาใช้งาน หากอาจารย์ต้องการข้อสอบที่มีระดับความยากง่าย หรือความสามารถในการจำแนกผู้สอบมากนักเพียงใดจะได้ดึงเอาข้อสอบที่มีคุณลักษณะตามต้องการออกมาใช้ได้ตามต้องการ

๔. ใช้เป็นประโยชน์ในการพัฒนาคุณภาพการสอน

การพิจารณาผลการวิเคราะห์ข้อสอบโดยละเอียดในหัวข้อที่อาจารย์ท่านใดท่านหนึ่งรับผิดชอบ

เวบบันทึทศิริราช

บทความทั่วไป

ในการสอนนักเรียนหรือแพทย์ประจำบ้านอยู่นั้นจะทำให้ได้ข้อมูลที่เป็นประโยชน์ในการพัฒนาการเรียนการสอนได้ กล่าวคืออาจารย์สามารถตรวจสอบดูได้ว่านักเรียนหรือแพทย์ประจำบ้านมีความเข้าใจที่ถูกต้องในเรื่องดังกล่าวหรือไม่ ประเด็นใดที่มีผู้เข้าใจผิดอยู่มากก็สมควรที่อาจารย์จะทำการเน้นย้ำในบรรดานักเรียนหรือแพทย์ประจำบ้านในการสอนครั้งต่อ ๆ ไปเพื่อแก้ไขความเข้าใจผิดดังกล่าว ประเด็นใดที่นักเรียนหรือแพทย์ประจำบ้านมีความเข้าใจดีมากอยู่แล้ว อาจารย์อาจไม่ต้องใช้เวลามากนักในการสอนเรื่องดังกล่าว แต่เอาเวลาไปใช้สอนในเรื่องที่นักเรียนหรือแพทย์ประจำบ้านยังไม่ค่อยเข้าใจให้มากขึ้นได้

ข้อจำกัดของการวิเคราะห์ข้อสอบ

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

พื้นฐานสำคัญที่เป็นข้อจำกัดของผลการวิเคราะห์ข้อสอบด้วยทฤษฎีการสอบแบบดั้งเดิมคือค่าต่าง ๆ ที่ได้มาจากการวิเคราะห์นั้นขึ้นอยู่กับกลุ่มตัวอย่างที่ใช้ในการเก็บข้อมูล^{๓๓,๓๔} หากได้ข้อมูลมาจากกลุ่มตัวอย่างที่มีขนาดใหญ่พอและมีกระจายตัวของระดับความสามารถของผู้สอบที่เหมาะสม ค่าต่าง ๆ ที่ได้ (p , r , coefficient alpha) จะค่อนข้างเที่ยงตรง ปัญหาที่สำคัญในการวิเคราะห์ข้อสอบในโรงเรียนแพทย์คือการสอบจำนวนมากจัดในนักศึกษาในกลุ่มเล็ก และ

นักศึกษาแต่ละกลุ่มก็มีการกระจายตัวของระดับความสามารถแตกต่างกัน นักศึกษาบางกลุ่มมีความสามารถสูงกว่านักศึกษากลุ่มอื่น ดังนั้นผลการวิเคราะห์ข้อสอบไม่ว่าจะเป็นค่า p , r , coefficient alpha, mean, หรือ standard deviation อาจเปลี่ยนแปลงไปในแต่ละกลุ่มของนักศึกษา ดังนั้นการนำผลการวิเคราะห์ข้อสอบไปใช้ในทางปฏิบัติจึงมีข้อควรระวังดังต่อไปนี้

การพิจารณาว่าข้อสอบยากหรือง่ายโดยใช้ค่า p นั้นเป็นค่าที่ไม่คงที่ ขึ้นอยู่กับกลุ่มผู้สอบ หากนำข้อสอบข้อหนึ่งไปไปใช้กับนักเรียนกลุ่มที่มีความรู้ดีนักเรียนส่วนใหญ่จะทำข้อสอบได้ถูกต้องทำให้ค่า p สูง แต่เมื่อนำข้อสอบข้อเดิมไปใช้กับนักเรียนกลุ่มที่ความรู้ไม่ดีสัก สักส่วนของนักเรียนที่ทำข้อสอบข้อเดียวกันได้ถูกต้องจะลดลงทำให้ค่า p ลดลง นอกจากนี้ในข้อสอบที่เน้นการท่องจำที่เคยใช้แล้ว เมื่อนำกลับมาใช้ใหม่ในนักเรียนกลุ่มใหม่ อาจมีนักเรียนจำนวนหนึ่งที่สามารถตอบข้อสอบถูกต้องเนื่องจากรู้ข้อสอบมาก่อนก็จะทำให้ค่า p สูงขึ้นกว่าเดิมได้

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

ค่าสัมประสิทธิ์อัลฟา เป็นค่าที่มีความเฉพาะเจาะจงกับการสอบของนักเรียนกลุ่มใดกลุ่มหนึ่งเท่านั้น หากใช่เป็นคุณสมบัติติดตัวข้อสอบแต่ละข้อไม่ หากข้อสอบชุดหนึ่งทำการสอบกับนักเรียนกลุ่มหนึ่งแล้วพบว่าคะแนนสอบที่ได้มานั้นมีค่าสัมประสิทธิ์อัลฟาสูงในระดับที่ต้องการก็ไม่ได้เป็นตัวรับประกันว่าหากนำข้อสอบชุดเดิมนั้นไปทำการสอบกับนักเรียนกลุ่มอื่นจะได้ค่าสัมประสิทธิ์อัลฟาที่สูงเช่นเดียวกัน นอกจากนี้ค่าสัมประสิทธิ์อัลฟาที่สูงไม่ได้เป็นตัวบอกถึงคุณภาพของข้อสอบรายข้อแต่อย่างใด

ค่าสัมประสิทธิ์อัลฟาที่สูงช่วยบอกแค่เพียงว่า

๓๖

เมษายน-เมษายน ๒๕๕๓, ปีที่ ๓, ฉบับที่ ๑

เว็บบัณฑิตศิริราช

บทความทั่วไป

คะแนนสอบในข้อสอบข้อหนึ่งมีความผันแปรไปในทิศทางเดียวกันกับคะแนนสอบในข้อสอบข้ออื่นในการสอบชุดเดียวกัน นั่นคือในข้อสอบชุดที่มีค่าสัมประสิทธิ์อัลฟาสูงก็อาจประกอบไปด้วยข้อสอบที่ดี และข้อสอบที่ไม่ดีรวมกันอยู่ ต้องไปตรวจสอบดัชนีชี้วัดคุณภาพของข้อสอบตัวอื่น ๆ ในแต่ละข้ออีกครั้ง

ข้อควรจำในการวิเคราะห์ข้อสอบที่ผู้นิพนธ์ขอ ย้ำในตอนท้ายของบทความนี้ก็คือค่าดัชนีชี้วัดคุณภาพต่าง ๆ ของข้อสอบที่กล่าวมาทั้งหมดนี้เป็นเพียงตัวช่วยให้อาจารย์เข้าใจข้อสอบดีขึ้นและช่วยแนะแนวทางในการพัฒนาปรับปรุงข้อสอบให้ดีขึ้น ดัชนีเหล่านี้ไม่ใช่ค่าตัดสินหรือตัวชี้ชะตาของข้อสอบ ไม่มีดัชนีใดที่ได้จากการวิเคราะห์ข้อสอบจะมาทดแทนดุลยพินิจของอาจารย์ไปได้ ดัชนีคุณภาพของข้อสอบไม่ว่าจะคำนวณมาด้วยวิธีการที่ถูกต้องแล้วก็ตามก็เป็นเพียงตัวเลขที่สามารถเกิดความผิดพลาดในการแปลผลได้ดังเช่นการแปลผลการวิเคราะห์ทางสถิติต่าง ๆ บทบาทของอาจารย์ในการวิเคราะห์ข้อสอบคงไม่ใช่การยึดถือตัวเลขดัชนีต่าง ๆ เป็นกฎตายตัว หากแต่ใช้ดัชนีเหล่านี้ช่วยเป็นแนวทางในการพิจารณาข้อสอบ หากดัชนีตัวใดระบุว่าข้อสอบอาจมีปัญหา อาจารย์ก็นำข้อสอบนั้นมาพิจารณากันโดยคณะกรรมการข้อสอบ หากหลังจากการพิจารณาโดยถี่ถ้วนแล้วอาจารย์คิดว่าข้อสอบข้อนั้นเหมาะสมแล้ว ไม่ควรทำการปรับแก้เนื้อหา อาจารย์ก็ยืนยันไปว่าไม่แก้ไข อาจารย์คงไม่ตัดสินการรักษาผู้ป่วยโดยใช้ผลเลือดตัวใดตัวหนึ่งเป็นเกณฑ์โดยไม่พิจารณาอาการและอาการแสดงของผู้ป่วยร่วมด้วย ฉะนั้นดิฉันนั้น อาจารย์

ไม่ควรตัดสินชะตากรรมของข้อสอบโดยใช้เพียงค่า 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

๓๓๗

นภกรณ-เมฆงน ๒๕๕๓, ปีที่ ๑๒, ฉบับที่ ๑



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)

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

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

=====

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 $\sqrt{1-rtt}$)

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

การสร้างข้อสอบอัตนัยประยุกต์

รองศาสตราจารย์ นายแพทย์เชิดศักดิ์ โอรมนิรัตน์ พ.บ., ป.ชั้นสูง (ศัลยศาสตร์), ว.ว. ศัลยศาสตร์, MHPE, Ph.D.
ภาควิชาศัลยศาสตร์, คณะแพทยศาสตร์ศิริราชพยาบาล, มหาวิทยาลัยมหิดล, กรุงเทพมหานคร 10700.

ข้อสอบอัตนัยประยุกต์ (modified essay question, MEQ) เป็นรูปแบบการประเมินผลที่นิยมใช้กับนักศึกษาแพทย์ระดับคลินิกเพื่อประเมินความสามารถในการแก้ปัญหา และตัดสินใจเลือกการตรวจรักษาที่เหมาะสมสำหรับผู้ป่วย ในปัจจุบันมีการใช้ข้อสอบอัตนัยประยุกต์ในการสอบของนักศึกษาแพทย์ในหลายภาควิชา รวมทั้งใช้ในการสอบขั้นตอนที่สามของการประเมินความรู้ความสามารถในการประกอบวิชาชีพเวชกรรม ของแพทยสภาด้วย อย่างไรก็ตาม จากการติดตามเนื้อหาของโจทย์ข้อสอบอัตนัยประยุกต์ ร่วมกับการพิจารณาเกณฑ์การให้คะแนนของข้อสอบเหล่านี้ที่ใช้กับการสอบของนักศึกษาแพทย์ในหลายการสอบ ผู้นิพนธ์ยังคงพบเห็นปัญหาในการสร้างข้อสอบชนิดนี้อยู่พอสมควร บทความนี้จึงได้รับการเขียนขึ้นเพื่อสร้างความเข้าใจในหลักการพื้นฐาน และแนวปฏิบัติที่เหมาะสมในการสร้างข้อสอบอัตนัยประยุกต์สำหรับการประเมินความรู้ทางการแพทย์

ลักษณะพื้นฐานของข้อสอบอัตนัยประยุกต์

ข้อสอบอัตนัยประยุกต์เป็นรูปแบบหนึ่งของข้อสอบอัตนัย (Essay question) ซึ่งในรูปแบบดั้งเดิม (traditional essay) นั้นผู้ออกข้อสอบจะเขียนโจทย์คำถามแล้วให้ผู้สอบเขียนคำตอบด้วยตนเองในขั้นตอนเดียว โดยไม่มีตัวเลือกให้ ในการเขียนคำตอบอาจเขียนตอบเป็นคำ หรือวลีสั้น ๆ (Short essay) หรือ ตอบเป็นบทความที่มีความยาวเป็นย่อหน้า หรือ หลายย่อหน้า (Long essay) ซึ่งผู้ออกข้อสอบคาดหวังการสอบในลักษณะที่ผู้สอบไม่มี

ตัวเลือก แต่ต้องคิดคำตอบด้วยตนเองนี้จะสามารถวัดความรู้ขั้นสูงในระดับการวิเคราะห์ สังเคราะห์ หรือประเมินคุณค่าได้^{1,2}

อย่างไรก็ตามข้อสอบในรูปแบบอัตนัยแบบดั้งเดิมนั้นประสบปัญหาในการใช้ประเมินความรู้ทางการแพทย์อยู่หลายประการ ทั้งความยากในการตรวจให้คะแนน ความจำกัดในปริมาณเนื้อหาที่สามารถสอบได้ในเวลาที่มี ความเห็นที่แตกต่างกันของผู้ตรวจให้คะแนน ความไม่เที่ยงของคะแนนสอบ เป็นต้น^{1,2} ปัญหาที่สำคัญยิ่งที่ทำให้การสอบอัตนัยแบบดั้งเดิมไม่ได้รับความนิยมในการประเมินความรู้ในระดับคลินิกคือ การที่ข้อสอบอัตนัยแบบดั้งเดิมนั้นมักวัดความรู้ในระดับการท่องจำ หรือความเข้าใจพื้นฐานเท่านั้น และรูปแบบการคิดวิเคราะห์เพื่อตอบโจทย์ข้อสอบอัตนัยแบบดั้งเดิมนั้นมีลักษณะแตกต่างไปจากกระบวนการแก้ปัญหาในระดับคลินิกที่แพทย์ปฏิบัติจริง

ข้อสอบอัตนัยแบบดั้งเดิมที่ตีนั้นผู้ออกข้อสอบสามารถประเมินทักษะการคิดวิเคราะห์ขั้นสูงได้ แต่อุปสรรคสำคัญที่ทำให้ไม่สามารถบรรลุวัตถุประสงค์ดังกล่าวได้คือการสร้างข้อสอบที่ผู้สอบตั้งใจให้ตรวจให้คะแนนได้ง่ายเป็นสำคัญ ทำให้ข้อสอบอัตนัยแบบดั้งเดิมส่วนใหญ่ทำการประเมินเพียงความรู้ระดับความจำหรือความเข้าใจพื้นฐานเท่านั้น

สมมติฐานพื้นฐานในการตอบข้อสอบอัตนัยแบบดั้งเดิมคือการวิเคราะห์และหาแนวทางแก้ปัญหาเป็นกระบวนการที่ทำในขั้นตอนเดียว ดังนั้นข้อสอบจึง

เวบบิ้นทักสิริราช

บทความทั่วไป

นำเสนอข้อมูลทั้งหมดในขั้นตอนเดียวแล้วให้ผู้เข้าสอบ แสดงการวิเคราะห์และแก้ปัญหา ซึ่งเป็นกระบวนการแก้ปัญหาทางคลินิกที่แพทย์ใช้ในกรณีเจอผู้ป่วยที่ไม่ซับซ้อนที่ไม่ต้องการกระบวนการคิดวิเคราะห์ขั้นสูงมากนัก อย่างไรก็ตามปัญหาผู้ป่วยที่มีความซับซ้อนและต้องการวิเคราะห์มากมักต้องการกระบวนการแก้ปัญหาหลายขั้นตอน แพทย์จะต้องทำการประเมินข้อมูลพื้นฐานที่ได้จากผู้ป่วย แล้วซักประวัติ หรือตรวจร่างกายเพื่อเก็บข้อมูลเพิ่มเติมอย่างเหมาะสม เมื่อได้ข้อมูลพื้นฐานมาแล้ว แพทย์ต้องทำการตั้งสมมติฐานถึงโรคที่ผู้ป่วยน่าจะเป็น แล้วทำการสืบค้นเพิ่มเติมด้วยการตรวจทางห้องปฏิบัติการ หรือใช้ภาพถ่ายรังสี ในบางกรณีแพทย์จำเป็นต้องให้การรักษารักษาเบื้องต้นก่อน พร้อมกับทำการสืบค้นเพิ่มเติม ซึ่งเมื่อเวลาผ่านไปแพทย์จะได้รับข้อมูลของผู้ป่วยมากขึ้นเรื่อย ๆ จากผลตรวจทางห้องปฏิบัติการ หรือการตอบสนองต่อการรักษาที่ให้ เมื่อได้ข้อมูลมากขึ้นแพทย์จะต้องทำการประเมินสถานการณ์ใหม่ ข้อมูลที่เพิ่มขึ้นอาจทำให้แพทย์สามารถให้การวินิจฉัยที่แน่ชัด และวางแผนการรักษาที่เหมาะสมได้ จะเห็นได้ว่ากระบวนการแก้ปัญหาของแพทย์มักทำเป็นหลายขั้นหลายตอน แต่ละขั้นตอนจะได้ข้อมูลเพิ่มเติมขึ้นเรื่อย ๆ การตัดสินใจในแต่ละขั้นเมื่อได้เลือกที่จะตรวจหรือให้การรักษาใดแก่ผู้ป่วยแล้ว ไม่สามารถย้อนเวลากลับไปแก้ไขการตัดสินใจที่ทำผิดพลาดไปก่อนหน้านี้ได้

จากข้อจำกัดของข้อสอบอัตนัยแบบดั้งเดิมที่กล่าวมาข้างต้น ทำให้มีการพัฒนารูปแบบการสอบเป็นข้อสอบอัตนัยประยุกต์ (modified essay question, MEQ) ซึ่งเป็นข้อสอบที่เริ่มจากการให้สถานการณ์ของผู้ป่วย แล้วมีโจทย์ถามให้ผู้สอบตอบคำถามที่เกี่ยวกับการแก้ปัญหาผู้ป่วยในสถานการณ์นั้นโดยไม่มีตัวเลือกให้ เมื่อผู้สอบตอบคำถามแล้วจะมีการเปิดเผยข้อมูลเพิ่มเติมเกี่ยวกับผู้ป่วยมากขึ้นทีละน้อย และมีโจทย์ถามคำถามเพิ่มเติมเป็นลำดับ โดยที่ผู้สอบไม่มีโอกาสย้อนกลับไปแก้ไขคำตอบของตนเองที่ได้ตอบไปในขั้นตอนก่อนหน้านี้^{1,3} รูปแบบของข้อสอบอัตนัยประยุกต์ที่นิยมใช้กันมากในยุคแรก ๆ มีลักษณะเป็นการสอบถามกระบวนการดูแลผู้ป่วยตั้งแต่ต้นจนจบในรูปแบบที่เรียกว่าการจัดการ

ปัญหาของผู้ป่วย (Patient management problem, PMP)^{1,4,5}

เนื่องจากข้อสอบอัตนัยประยุกต์ที่ใช้ในทางการแพทย์มักมุ่งเน้นการประเมินทักษะการวินิจฉัยโรค ผู้นิพนธ์จึงขอทบทวนทฤษฎีเกี่ยวกับกระบวนการวินิจฉัยโรคเล็กน้อยก่อนนำเข้าสู่หลักการสร้างข้อสอบ โดยทั่วไปแล้ววิธีการที่แพทย์ใช้ในการวินิจฉัยโรคมีสามวิธีหลักได้แก่ (1) วิธีจำได้จากแบบแผนของความผิดปกติที่พบ (pattern recognition), (2) วิธีปฏิบัติตามขั้นตอนวิธีที่มีแบบแผน (algorithm), และ (3) วิธีทดสอบสมมติฐาน (hypothesis testing)⁶ ซึ่งในวิธีทดสอบสมมติฐานนี้สามารถแบ่งออกเป็นวิธีการย่อยได้สองวิธีคือ (3.1) การแก้ปัญหาด้วยวิธีอุปนัย (inductive reasoning) ซึ่งแพทย์จะรวบรวมข้อมูลอย่างครบถ้วนตามแบบแผนก่อนจึงตั้งสมมติฐาน และ (3.2) การแก้ปัญหาด้วยวิธีนรนัย (deductive reasoning) ซึ่งแพทย์จะเริ่มตั้งสมมติฐานตั้งแต่เมื่อเริ่มเก็บข้อมูลจากผู้ป่วยเพียงเล็กน้อย แล้วใช้สมมติฐานที่ได้มานั้นเป็นแนวทางในการซักประวัติ และตรวจร่างกายอย่างมีจุดหมายเพื่อทดสอบสมมติฐานที่ตั้งขึ้นจนค่อย ๆ ตัดโรคที่ไม่สอดคล้องกับข้อมูลที่ได้รับออกไปเรื่อย ๆ โดยทั่วไปแล้ววิธีอุปนัยเป็นวิธีที่มีประสิทธิภาพน้อยกว่าวิธีนรนัย เนื่องจากการเก็บข้อมูลเป็นไปอย่างขาดจุดหมายทำให้เสียเวลาและอาจพลาดการเก็บข้อมูลที่สำคัญไป⁶

การสร้างข้อสอบอัตนัยประยุกต์ที่มีคุณภาพดีควรเริ่มจากความเข้าใจในปรัชญาพื้นฐานของการประเมินผลว่าข้อสอบอัตนัยประยุกต์นั้นได้รับการพัฒนาขึ้นเพื่อประเมินทักษะการแก้ปัญหาด้วยวิธีนรนัยเป็นสำคัญ ข้อผิดพลาดที่พบบ่อยของการสร้างข้อสอบอัตนัยประยุกต์ประการหนึ่งคือการสร้างข้อสอบที่ให้ข้อมูลผู้ป่วยสั้นมาก (จนไม่มีทางตั้งสมมติฐานที่ชัดเจนได้) แล้วตั้งโจทย์ให้ผู้เข้าสอบเขียนรายการประวัติที่จะสอบถามหรือการตรวจร่างกายที่จะดำเนินการในผู้ป่วยดังกล่าว เช่น ให้สถานการณ์เป็นหญิงอายุ 45 ปี ปวดท้อง 1 วัน แล้วตั้งโจทย์ว่า จงทำการซักประวัติที่เหมาะสม ซึ่งการให้สถานการณ์ในลักษณะนี้มีโรคที่สามารถเป็นไปได้มากมาย ในหลายระบบ สิ่งที่จะประเมินได้จากการตอบ

เวบบ์ทีกศิริราช

บทความทั่วไป

คำถามลักษณะนี้คือความจำขึ้นพื้นฐาน (simple recall) ว่าแบบแผนการชักประวัติผู้ป่วยปวดท้องเฉียบพลันมีอะไรบ้าง ซึ่งผู้เข้าสอบเขียนอะไรมาก็หน้าจะถูกต้อง ไม่มีการชักประวัติที่ไม่เข้าประเด็น เนื่องจากข้อมูลจากโจทย์ไม่มีรายละเอียดมากพอที่จะจำกัดโรคที่ควรนึกถึง ข้อสอบอัตนัยประยุกต์ที่ดีควรเริ่มจากข้อมูลที่สามารถสร้างสมมติฐานที่ชัดเจนพอได้ เช่น หญิงอายุ 50 ปี จุกแน่นลิ้นปี่และได้ชายโครงขวาเป็น ๆ หาย ๆ 4 เดือน มีอาการปวดท้องได้ชายโครงขวามาก ร่วมกับมีไข้ต่ำ ๆ 7 ชั่วโมง การให้ข้อมูลที่มีรายละเอียดพอสมควรนี้ผู้สอบที่มีความรู้จะตั้งสมมติฐานได้ว่าผู้ป่วยน่าจะเป็นโรคใด หากโจทย์กำหนดให้ชักประวัติเพิ่มเติม ผู้สอบที่มีความรู้จะสามารถสอบถามอาการที่สอดคล้องกับการวินิจฉัยที่เหมาะสมได้ ในกรณีนี้คำตอบที่ไม่สอดคล้อง (เช่นสมมติฐานที่เหมาะสมคือภาวะถุงน้ำดีอักเสบเฉียบพลัน แต่ผู้สอบชักประวัติประจำเดือน ประวัติเพศสัมพันธ์) ไม่ควรได้คะแนน

พัฒนาการของข้อสอบอัตนัยประยุกต์

หลังจากที่มีรายงานการใช้ข้อสอบอัตนัยประยุกต์ในการประเมินผลทางแพทยศาสตรศึกษาตั้งแต่ปี พ.ศ. 2514 โดยราชวิทยาลัยแพทย์เวชปฏิบัติทั่วไปเพื่อประเมินทักษะการแก้ปัญหาทางคลินิกแล้ว^{3,7,8} ข้อสอบอัตนัยประยุกต์ก็ได้ถูกใช้ในการประเมินทางการแพทย์และสาธารณสุขในหลากหลายบริบท⁹⁻¹² โดยรูปแบบที่เป็นที่นิยมกันมากเป็นการสอบถามกระบวนการดูแลผู้ป่วยตั้งแต่ต้นจนจบในรูปแบบ การจัดการปัญหาของผู้ป่วย (Patient management problem, PMP) ซึ่งการแก้ปัญหาผู้ป่วยแต่ละรายมักใช้เวลานานมาก ทำให้การสอบแต่ละครั้งมักมีจำนวนสถานการณ์ผู้ป่วยที่นำมาสอบไม่มากนัก¹³

จากการใช้ข้อสอบอัตนัยประยุกต์ในรูปแบบการจัดการปัญหาของผู้ป่วย พบว่ามีข้อจำกัดบางประการ กล่าวคือ ข้อสอบส่วนใหญ่มุ่งเน้นวัดความครบถ้วนสมบูรณ์ของคำตอบมากกว่าการตัดสินใจแก้ปัญหา จำนวนสถานการณ์ผู้ป่วยที่มีจำนวนน้อยทำให้ไม่สามารถครอบคลุมองค์ความรู้ที่ต้องการประเมินได้ครบ และความ

เที่ยงของคะแนนสอบที่ต่ำ^{4,13,14} ปัญหาที่สำคัญยิ่งในการสอบด้วยสถานการณ์ผู้ป่วยจำนวนน้อยคือ ทักษะในการแก้ปัญหาทางคลินิกมีความจำเพาะต่อบริบทของผู้ป่วยแต่ละราย (case specificity)¹⁵⁻¹⁸ การที่ผู้เข้าสอบสามารถแก้ปัญหาผู้ป่วยที่มีอาการเจ็บหน้าอกได้ดีนั้นไม่สามารถจะบอกได้ว่าผู้เข้าสอบคนดังกล่าวจะสามารถแก้ปัญหาผู้ป่วยที่มีอาการปวดศีรษะได้ดีด้วยหรือไม่ ดังนั้นหลักการที่สำคัญประการหนึ่งในการสร้างข้อสอบอัตนัยประยุกต์ก็คือการจัดทำข้อสอบให้มีหลากหลายสถานการณ์ เพื่อให้สามารถประเมินการแก้ปัญหาของผู้เข้าสอบได้ในหลากหลายบริบท ในหลายระบบอวัยวะ จากปัญหาในการใช้ข้อสอบอัตนัยประยุกต์ต่าง ๆ เหล่านี้ ทำให้นักการศึกษาได้มีการพัฒนารูปแบบข้อสอบอัตนัยประยุกต์ให้ต่างไปจากรูปแบบดั้งเดิม รูปแบบข้อสอบที่ผู้เชี่ยวชาญในการประเมินผลแนะนำในปัจจุบันคือ การแก้ปัญหาสำคัญ (key features problems, KFP)

ข้อสอบอัตนัยประยุกต์แบบแก้ปัญหาสำคัญนี้ได้รับการพัฒนามาบนหลักการสำคัญคือในการแก้ปัญหาผู้ป่วยแต่ละรายมีประเด็นปัญหาที่เป็นหัวใจสำคัญเพียงไม่กี่ประเด็นเท่านั้น ซึ่งประเด็นปัญหาเหล่านี้เรียกว่า ปัญหาสำคัญ (key features)¹⁹ ซึ่งในผู้ป่วยแต่ละรายจะมีปัญหาสำคัญที่แพทย์ต้องให้ความสนใจต่างกันไป บางรายเป็นเรื่องการชักประวัติ บางรายเป็นการเลือกการส่งตรวจทางห้องปฏิบัติการ ในขณะที่บางรายเป็นการตัดสินใจเลือกวิธีการรักษาที่เหมาะสม เป็นต้น ในข้อสอบอัตนัยประยุกต์รูปแบบการแก้ปัญหาสำคัญจะมุ่งเน้นตั้งใจถามเฉพาะประเด็นปัญหาสำคัญเหล่านี้เท่านั้น ไม่จำเป็นต้องถามกระบวนการดูแลผู้ป่วยตั้งแต่ต้นจนจบในผู้ป่วยทุกราย การสร้างข้อสอบอัตนัยประยุกต์ในลักษณะนี้ทำให้ผู้สอบใช้เวลาในการแก้ปัญหาผู้ป่วยแต่ละรายไม่มากนัก และสามารถประเมินทักษะการแก้ปัญหาได้ในหลากหลายสถานการณ์ คะแนนสอบที่ได้จึงมีความเที่ยงสูง มีรายงานค่าความเที่ยงของคะแนนสอบถึง 0.8 ในการสอบอัตนัยประยุกต์แบบแก้ปัญหาสำคัญเป็นเวลาสี่ชั่วโมง¹⁴

เวบบันทึกศิริราช

บทความทั่วไป

ตัวอย่างข้อสอบอัตนัยประยุกต์แบบแก้ปัญหาสำคัญ
ตอนที่ 1 ชาย 36 ปี น้ำหนักตัว 55 กิโลกรัม ท้องร่วงถ่ายเป็นน้ำ 20 ครั้งในเวลา 1 วัน ตรวจร่างกายพบ อุณหภูมิ 36.9 องศาเซลเซียส ชีพจร 112 ครั้งต่อนาที ตรวจความดันโลหิตท่านอน 104/56 มิลลิเมตรปรอท ความดันโลหิตท่านั่ง 90/50 มิลลิเมตรปรอท

คำถามที่ 1.1 ให้ผู้สอบเขียนปัญหาสำคัญที่สุดของผู้ป่วยรายนี้ 1 อย่าง

ตอนที่ 2 ผู้ป่วยได้รับการประเมินว่ามีภาวะขาดสารน้ำปานกลางถึงรุนแรง ท่านต้องการให้สารน้ำทางหลอดเลือดดำแก่ผู้ป่วย

คำถามที่ 2.1 จงเขียนคำสั่งการรักษาเพื่อให้สารน้ำที่เหมาะสมแก่ผู้ป่วย

คำถามที่ 2.2 จงส่งตรวจเพิ่มเติมทางห้องปฏิบัติการเพื่อช่วยวินิจฉัยผู้ป่วยรายนี้ 2 การตรวจ

จากตัวอย่างข้างต้นจะเห็นว่าผู้ออกข้อสอบไม่ได้เริ่มจากการถามว่าจะซักประวัติ หรือตรวจร่างกายอะไรในผู้ป่วยที่มีภาวะท้องร่วงรุนแรง เนื่องจากผู้ออกข้อสอบเห็นว่าปัญหาสำคัญในการดูแลผู้ป่วยในภาวะนี้เป็นเรื่องการประเมินความรุนแรงของการขาดสารน้ำและการให้น้ำเกลือทดแทนในปริมาณที่เหมาะสมร่วมกับการสืบค้นหาสาเหตุของท้องร่วง ดังนั้นโจทย์ข้อนี้จึงมีเพียงสองตอนและใช้เวลาสอบไม่เกินสิบนาที

ขั้นตอนการสร้างข้อสอบอัตนัยประยุกต์

การสร้างข้อสอบอัตนัยประยุกต์ที่มีคุณภาพดีควรมีการดำเนินการเป็นขั้นตอน ดังนี้^{4,20}

1. ตั้งกลุ่มพัฒนาข้อสอบ

ข้อสอบอัตนัยประยุกต์ที่ดีควรเป็นการแก้ปัญหาที่อาศัยความรู้จากหลากหลายวิชา การมีทีมคณาจารย์ที่มีประสบการณ์และความชำนาญแตกต่างกันมาช่วยกันสร้างข้อสอบจะได้สถานการณ์ผู้ป่วยที่เหมือนจริงในเวชปฏิบัติและสามารถประเมินความรู้ของผู้เข้าสอบได้ครอบคลุมสหสาขาวิชา และมั่นใจได้ว่าการเฉลยคำตอบทำได้อย่างรอบคอบ

2. เลือกปัญหาทางคลินิกที่จะทำการประเมินผู้สอบ

ขั้นตอนนี้เป็นขั้นตอนที่สำคัญมาก เนื่องจากโดยลักษณะข้อสอบอัตนัยประยุกต์จะทำให้ทำการสอบได้จำนวนข้อไม่มากนัก จึงเป็นไปได้ที่จะทำให้สถานการณ์ที่เป็นปัญหาทางคลินิกทุกอย่างจะมาปรากฏอยู่ในชุดข้อสอบ ดังนั้นการเลือกปัญหาทางคลินิกที่จะทำการสอบจึงต้องทำอย่างเป็นระบบ ควรมีการจัดทำตารางกำหนดลักษณะข้อสอบที่ชัดเจนว่าในการสอบครั้งหนึ่ง ๆ จะมีข้อสอบกี่ข้อ จะประเมินความรู้ในระบบอวัยวะใด และจัดสรรให้ข้อสอบไม่ซ้ำซ้อนกัน (ไม่ควรมีข้อสอบสองข้อถามความรู้ในระบบอวัยวะเดียวกัน ในขณะที่บางระบบอวัยวะไม่มีข้อสอบเลย)

ลักษณะปัญหาทางคลินิกที่ควรเลือกมาสอบด้วยข้อสอบอัตนัยประยุกต์ได้แก่

- ปัญหาที่พบได้บ่อยในเวชปฏิบัติ
- ปัญหาที่แพทย์เกิดความผิดพลาดในการดูแลผู้ป่วยค่อนข้างบ่อย
- ปัญหาที่ยังไม่สามารถวินิจฉัยสาเหตุได้ชัดเจน
- ปัญหาที่มีความเกี่ยวข้องกับหลายระบบ

เมื่อทีมคณาจารย์กำหนดปัญหาทางคลินิกที่จะทำการประเมินได้ชัดเจนแล้ว (เช่น ปัญหาตัวเหลือง, น้ำหนักลด เป็นต้น) สิ่งที่ต้องดำเนินการต่อคือการสร้างสถานการณ์ผู้ป่วยที่แสดงถึงปัญหาดังกล่าวขึ้น โดยกำหนดรายละเอียดต่าง ๆ ให้ผู้เข้าสอบอ่านแล้วนึกภาพผู้ป่วยได้ ในสถานการณ์ควรมีรายละเอียดเกี่ยวกับอายุ เพศ อาการสำคัญ บริบทของการดูแลผู้ป่วย (เช่น ห้องฉุกเฉินของโรงพยาบาลชุมชน หรือ หอผู้ป่วยในโรงพยาบาลมหาวิทยาลัย เป็นต้น)

3. กำหนดปัญหาสำคัญ

เมื่อทีมคณาจารย์เลือกปัญหาทางคลินิกที่จะทำการสอบแล้ว คณาจารย์ต้องตั้งคำถามว่าขั้นตอนใดในการดูแลผู้ป่วยที่มีปัญหาดังกล่าวจัดเป็นขั้นตอนสำคัญที่สุดในการจัดการปัญหานั้น ซึ่งขั้นตอนดังกล่าวจะได้รับการกำหนดให้เป็น ปัญหาสำคัญของสถานการณ์ผู้ป่วยที่จะใช้สอบ ในบางกรณีที่มีทีมคณาจารย์ไม่สามารถเลือกขั้นตอนสำคัญในปัญหาทางคลินิกนั้น ๆ จากวิธีดังกล่าวได้

เวบบันเทิงศิริราช

บทความทั่วไป

อาจใช้คำถามว่าขั้นตอนใดในการดูแลผู้ป่วยที่มีปัญหา ดังกล่าวเป็นขั้นตอนที่นักศึกษาแพทย์หรือแพทย์ประจำบ้านทำผิดพลาดมากที่สุด⁴

มีข้อเสนอแนะสองประการสำหรับการกำหนด ปัญหาสำคัญในแต่ละสถานการณ์ได้แก่

- สิ่งที่ต้องตัดสินใจในผู้ป่วยแม้เป็นสิ่งที่ถูกต้อง และควรปฏิบัติอาจไม่ได้เป็นขั้นตอนสำคัญที่จะต้องนำมาสอบเสมอไป การปฏิบัติต่อผู้ป่วยหลายอย่างที่ทำกัน เป็นปกติ โดยไม่ต้องคิดวิเคราะห์ เป็นขั้นตอนที่ไม่ค่อยทำผิดพลาด มักไม่ใช่ปัญหาสำคัญในสถานการณ์นั้น

- ปัญหาสำคัญไม่จำกัดอยู่เฉพาะประเด็นปัญหาทาง ชีววิทยาการแพทย์ (biomedical) เท่านั้น ในบางสถานการณ์ปัญหาสำคัญอาจเป็นประเด็นทางจริยธรรม กฎหมาย หรือ การส่งเสริมสุขภาพและป้องกันโรคก็ได้

4. เขียนใจหทัยคำถาม

เมื่อมีสถานการณ์ผู้ป่วยและขั้นตอนที่เป็นปัญหาสำคัญในสถานการณ์นั้นแล้ว ทีมคณาจารย์ต้องเขียน ใจหทัยคำถามที่มีความชัดเจน เพื่อประเมินว่าผู้เข้าสอบมีความสามารถในการตัดสินใจในการแก้ปัญหาสำคัญใน สถานการณ์ดังกล่าวหรือไม่ โดยทั่วไปแล้วลักษณะใจหทัย คำถามที่ช่วยในข้อสอบอัตนัยประยุกต์ได้แก่

- จงสอบถามประวัติที่สำคัญเพิ่มเติม
- จงบอกการตรวจร่างกายที่สำคัญที่ต้องมองหา (หรือตรวจเพิ่มเติม) ในผู้ป่วย
- จงให้การวินิจฉัย (หรือ การวินิจฉัยแยกโรค)
- จงสั่งการตรวจค้นเพิ่มเติมเพื่อให้การวินิจฉัยโรค
- จงสั่งการรักษาที่เหมาะสมให้ผู้ป่วย

โดยทั่วไปแล้วสถานการณ์ผู้ป่วยหนึ่ง ๆ ควรมี คำถามราว 2 – 3 ข้อ แต่ละข้อประเมินความสามารถในการจัดการกับปัญหาสำคัญ 1 ประเด็น^{4,21} ในการเขียน ใจหทัยคำถามแต่ละข้อนั้นแนะนำให้มีการกำหนดจำนวน คำตอบที่สามารถตอบได้ไว้ด้วย เช่น

- จงบอกชื่อโรคที่ผู้ป่วยรายนี้น่าจะเป็นมากที่สุด 1 โรค
- จงบอกผลการตรวจร่างกายที่สำคัญที่จะช่วย ยืนยันการวินิจฉัยโรคมา 3 ประการ

- จงระบุการตรวจเพิ่มเติมทางห้องปฏิบัติการที่ จะช่วยในการวินิจฉัยโรค 1 การตรวจ

การกำหนดจำนวนคำตอบนี้จะทำให้ผู้เข้าสอบ ต้องเลือกสิ่งที่ถูกต้องเหมาะสมที่สุดเท่านั้นมาเขียนตอบ หากผู้เข้าสอบเขียนคำตอบเกินจำนวนที่กำหนด อาจารย์ ผู้ตรวจข้อสอบจะไม่อ่านคำตอบที่เกินมา การปฏิบัติเช่นนี้ จะช่วยกำจัดปัญหาการตรวจกระดาษคำตอบที่ผู้เข้าสอบ เขียนคำตอบแบบหว่านแห ให้ครอบคลุมทุกอย่างโดยที่ ผู้เข้าสอบเองไม่มีความรู้ ความเข้าใจว่าสิ่งใดเป็นประเด็น สำคัญในการดูแลผู้ป่วยในขั้นตอนนั้น ๆ

เมื่อทำการเขียนใจหทัยคำถามและจำนวนคำตอบ ที่ต้องการแล้ว ให้อาจารย์ระบุเวลาที่ใช้ในการตอบคำถาม ตอนนั้นด้วย เนื่องจากข้อสอบอัตนัยประยุกต์มีการดำเนิน ของสถานการณ์ผู้ป่วยที่กำหนดให้โดยมีการให้ข้อมูลที่ละ ส่วน ผู้เข้าสอบจำเป็นต้องรู้เวลาที่มิในการทำข้อสอบ แต่ละตอนก่อนที่จะต้องส่งคำตอบและสถานการณ์ผู้ป่วย ดำเนินต่อไป ในการกำหนดเวลาในการทำข้อสอบแต่ละ ตอนให้อาจารย์ผู้ออกข้อสอบพิจารณาจากทั้งเวลาที่ ต้องใช้ในการอ่าน และเวลาที่ต้องใช้ในการเขียนคำตอบ ในข้อสอบตอนที่ต้องอ่านเนื้อหาใจหทัยมาก หรือต้องเขียน คำตอบหลายบรรทัด ควรต้องมีการให้เวลาในการทำ ข้อสอบมากพอ หากเป็นไปได้ควรมีการลองทำการ อ่านใจหทัยและเขียนคำตอบโดยตัวอาจารย์ผู้ออกข้อสอบ เองหรือเพื่อนอาจารย์แล้วลองจับเวลาที่อาจารย์ใช้ในการ ทำข้อสอบตอนนั้น ๆ เวลาที่ได้จะเป็นเวลาที่ผู้เชี่ยวชาญใช้ แก้ปัญหาผู้ป่วยในสถานการณ์ดังกล่าว หากให้นักศึกษา ทำ ควรเพิ่มเวลาให้ร้อยละ 30 – 50 ของเวลาที่อาจารย์ใช้

5. กำหนดเกณฑ์การให้คะแนน

ขั้นตอนสุดท้ายในการสร้างข้อสอบอัตนัย ประยุกต์คือการกำหนดเกณฑ์การให้คะแนน ซึ่งเป็นขั้นตอนที่มีความท้าทาย และสร้างความลำบากใจให้แก่ อาจารย์ผู้ออกข้อสอบหลายท่าน เนื่องด้วยเกรงว่าจะเฉลย คำตอบไม่ครอบคลุมสิ่งที่ผู้เข้าสอบจะเขียนตอบมา หรือ เกิดความไม่เป็นธรรมขึ้น ในที่นี้ผู้เขียนขอเสนอแนะแนว ทางในการกำหนดเกณฑ์ให้คะแนนดังนี้

- แนะนำให้กำหนดคะแนนเต็มในการแก้ปัญหา

เวบบันทึกศิริราช

บทความทั่วไป

สถานการณ์หนึ่ง ๆ เป็น 100 คะแนน เท่ากันในทุกสถานการณ์ เพื่อให้ไม่ต้องทำการปรับคะแนนสอบหลังการตรวจข้อสอบ

- กรณีที่มีคำตอบที่ถูกต้องยอมรับได้เพียงคำตอบเดียว เช่นข้อมูลจากโจทย์มีความชัดเจนว่าผู้ป่วยเป็นโรคอะไร แล้วโจทย์ให้ผู้เข้าสอบตอบชื่อโรค หากผู้เข้าสอบตอบตรงตามเฉลยที่ตั้งไว้ให้ได้คะแนนเต็ม หากตอบคำตอบอื่นนอกจากนั้นไม่ได้คะแนน

- ในกรณีที่มีคำตอบที่เป็นไปได้หลายคำตอบ เช่นถามการวินิจฉัยแยกโรค 3 โรค ในกรณีนี้ผู้ออกข้อสอบควรเตรียมเฉลยไว้หลายคำตอบ (มากกว่าที่กำหนดให้ตอบ) โดยแต่ละคำตอบสามารถมีน้ำหนักคะแนนไม่เท่ากันได้ โดยคำตอบที่ถูกต้องมาก สอดคล้องกับสิ่งที่ควรคิดถึงหรือปฏิบัติในขั้นตอนดังกล่าว จะได้คะแนนสูง ในขณะที่สิ่งที่สามารถเป็นไปได้หรือควรปฏิบัติน้อยกว่าจะได้คะแนนลดลงไป แต่เมื่อรวมคะแนนจากทุกคำตอบที่ผู้เข้าสอบตอบมาแล้วคะแนนสูงสุดที่ผู้เข้าสอบจะได้ต้องไม่สูงเกินคะแนนที่กำหนดไว้เป็นคะแนนเต็มของข้อสอบตอนนั้น

- คำตอบบางลักษณะมีการเขียนเนื้อหาที่มีความครบถ้วนสมบูรณ์แตกต่างกันได้ การกำหนดเกณฑ์สามารถกำหนดให้คำตอบที่มีความสมบูรณ์ได้คะแนนเต็ม ส่วนคำตอบที่ไม่สมบูรณ์จะได้คะแนนลดลงไปตามความเหมาะสม (เช่น โจทย์ถามเรื่องการให้สารน้ำทางหลอดเลือดดำ คำตอบ Normal saline solution 1000 ml IV drip 200 ml/hr จะได้คะแนนเต็ม 4 คะแนน แต่หากเขียนตอบ Normal saline solution โดยไม่บอกอัตราเร็วของการให้ ได้เพียง 2 คะแนน หากบอกอัตราเร็วให้ถูกต้องให้ 2 คะแนน)

- คำตอบที่ไม่ถูกต้อง ไม่สมควรปฏิบัติแก่ผู้ป่วยโดยทั่วไปแล้วพิจารณาไม่ให้คะแนน ซึ่งก็จัดเป็นการทำโทษในระดับหนึ่งแล้ว เพราะผู้สอบมีสิทธิเขียนคำตอบได้จำนวนจำกัด การที่ไม่ให้คะแนนในคำตอบที่ไม่เหมาะสม ก็จะทำให้คะแนนสูงสุดที่ผู้สอบจะทำได้ลดลงไปแล้ว การปฏิบัติที่ไม่ถูกต้องที่มีผลเสียรุนแรงต่อผู้ป่วยเท่านั้นที่ควรจะพิจารณาให้คะแนนติดลบ และแม้มีการให้คะแนนติดลบก็ไม่ควรมีการติดลบข้ามไปถึงข้อสอบข้ออื่นในชุดข้อสอบนั้น

- การกำหนดเกณฑ์การให้คะแนน ไม่ควรใช้อาจารย์ท่านเดียวในการกำหนด เพราะมักได้คำตอบที่ไม่ครอบคลุม ควรใช้ทีมคณาจารย์หลายท่านช่วยกันคิดว่าคำตอบที่ผู้เข้าสอบอาจจะตอบได้ในสถานการณ์ดังกล่าว ซึ่งจะได้เกณฑ์การให้คะแนนที่สมบูรณีกว่า อย่างไรก็ตาม ถึงแม้ว่าจะใช้คณาจารย์หลายท่านช่วยกันคิดคำตอบแล้วก็ตาม จะพบว่าในการตรวจข้อสอบอัตโนมัติประยุกต์หลายครั้ง จะพบคำตอบที่ผู้เข้าสอบตอบมาที่ น่าจะได้คะแนนแต่อาจารย์ผู้ออกข้อสอบไม่ได้กำหนดเกณฑ์คะแนนไว้ล่วงหน้าอยู่ประปราย ดังนั้นในการนำข้อสอบอัตโนมัติประยุกต์ที่สร้างขึ้นใหม่มาใช้ในการสอบ 2-3 รอบแรกแนะนำให้อาจารย์ผู้ออกข้อสอบและมีความเชี่ยวชาญชำนาญในการดูแลผู้ป่วยในสถานการณ์นั้น ๆ เป็นผู้ทำการตรวจข้อสอบ เพื่อให้สามารถพิจารณาได้ว่าคำตอบใดที่น่าจะเพิ่มเข้าไปในเกณฑ์การให้คะแนนด้วย ซึ่งเมื่อทำไป 2-3 รอบการสอบแล้วมักจะได้เกณฑ์การให้คะแนนที่มีความครอบคลุมคำตอบที่ผู้สอบจะตอบมาได้ทั้งหมด แล้วจึงมอบหมายให้อาจารย์ท่านอื่นช่วยตรวจให้คะแนนข้อสอบต่อไป

เมื่อทำการกำหนดเกณฑ์การให้คะแนนในข้อสอบเสร็จทุกข้อย่อยแล้วกระบวนการขั้นตอนสุดท้ายในการสร้างข้อสอบอัตโนมัติประยุกต์คือการกำหนดเกณฑ์ผ่านของโจทย์สถานการณ์นั้น กล่าวคือจากคะแนนเต็ม 100 คะแนน ผู้สอบต้องทำคะแนนได้อย่างน้อยที่สุดกี่คะแนนจึงจะจัดว่าสอบผ่านในการแก้ปัญหาสถานการณ์นั้น ๆ วิธีการตั้งเกณฑ์ผ่านทำได้หลายวิธี แต่วิธีที่เป็นที่นิยมมากที่สุดสำหรับข้อสอบอัตโนมัติประยุกต์ และเป็นวิธีที่คณะแพทยศาสตร์ศิริราชพยาบาลใช้เป็นประจำในการตัดสินผลสอบอัตโนมัติประยุกต์คือวิธี Modified Angoff ซึ่งมีขั้นตอนที่สำคัญสามขั้นตอนคือ

(1) กำหนดลักษณะของผู้ที่มีความรู้ ความสามารถคาบเส้น (borderline examinee) ว่าในความเห็นของคณาจารย์แล้วผู้ที่มีความรู้เทียบเท่าระดับต่ำสุดของเกณฑ์มาตรฐานการทำงานในการแก้ปัญหาเรื่องนั้น ๆ น่าจะทำอะไรได้ ทำอะไรไม่ได้

(2) ไล่ดูโจทย์คำถามทีละข้อพร้อมเฉลย แล้วทำสัญลักษณ์ * ไว้ในคำตอบที่คาดว่าผู้ที่มีความรู้ความสามารถคาบเส้นจะตอบในข้อสอบแต่ละตอน

เวบบันทึกศิริราช

บทความทั่วไป

(3) ทำการรวมค่าคะแนนที่ได้รับการทำ
สัญลักษณ์ * ไว้ตั้งแต่ข้อแรกจนถึงข้อสุดท้าย จะได้
คะแนนเกณฑ์ผ่านในการแก้ปัญหาสถานการณ์นั้น ๆ²²

แนวทางการพัฒนาข้อสอบอัตนัยประยุกต์ในคณะ แพทยศาสตร์ศิริราชพยาบาล

คณะแพทยศาสตร์ศิริราชพยาบาลมีการใช้
ข้อสอบอัตนัยประยุกต์ในการประเมินความรู้ของนักศึกษา
แพทย์ชั้นคลินิกมานานแล้ว โดยเริ่มต้นจากการสอบของ
แต่ละภาควิชา และต่อมาเมื่อศูนย์ประเมินและรับรอง
ความรู้ความสามารถในการประกอบวิชาชีพเวชกรรม
กำหนดให้การสอบอัตนัยประยุกต์เป็นส่วนหนึ่งของ
การประเมินขั้นตอนที่ 3 ในการขอใบประกอบวิชาชีพ
เวชกรรมตั้งแต่ปีการศึกษา 2550 ทางคณะแพทยศาสตร์
ศิริราชพยาบาลก็ได้มีการจัดสอบประมวลความรู้
ทางการแพทย์สหสาขาวิชา ด้วยข้อสอบอัตนัยประยุกต์
(comprehensive MEQ examination) ในนักศึกษา
แพทย์ปีที่ 6 อย่างต่อเนื่อง ตลอดช่วงเวลาที่มีการใช้
ข้อสอบอัตนัยประยุกต์ในคณะฯ ได้มีการพัฒนาข้อสอบ
ประเภทนี้อย่างต่อเนื่อง จากเดิมเคยจัดสอบข้อสอบอัตนัย
ประยุกต์ในรูปแบบข้อสอบกระดาษ จนพัฒนาให้จัดสอบ
อัตนัยประยุกต์ด้วยการนำเสนอข้อมูลผู้ป่วยบนจอภาพ
คอมพิวเตอร์ ร่วมกับการเขียนคำตอบในกระดาษคำตอบ
ตั้งแต่ปีการศึกษา 2552 จนถึงปัจจุบัน แต่ถึงแม้ว่าฝ่าย
การศึกษาจะมีการพัฒนาระบบจัดสอบข้อสอบอัตนัย
ประยุกต์ให้มีประสิทธิภาพมากขึ้น อำนาจความสะดวกให้
ผู้เข้าสอบมากขึ้น และเพิ่มความพึงพอใจในประสบการณ์
การสอบขึ้นอย่างต่อเนื่อง จากการเก็บรวบรวมข้อมูลการ
วิเคราะห์ข้อสอบ วิเคราะห์คะแนน และแบบสำรวจความ
พึงพอใจของผู้สอบที่ผ่านมาผู้นิพนธ์มีความเห็นว่าการ
จัดสอบประมวลความรู้ทางการแพทย์ด้วยข้อสอบ
อัตนัยประยุกต์ของนักศึกษาแพทย์ยังสามารถพัฒนาให้
มีคุณภาพดีขึ้นได้อีกในหลายด้าน ดังนี้

(1) เนื้อหาข้อสอบ

ข้อสอบอัตนัยประยุกต์ที่ใช้ในการสอบประมวล
ความรู้ทางการแพทย์ของคณะแพทยศาสตร์ศิริราช
พยาบาลที่ผ่านมาหลายข้อเป็นเนื้อหาวิชาที่ยากและเป็น
ความรู้ลึกในระดับผู้เชี่ยวชาญเฉพาะทาง แนวทางการ

พัฒนาการสอบอัตนัยประยุกต์อันดับแรกคือการพัฒนา
เนื้อหาให้เหมาะสมกับการประเมินความรู้ของแพทย์เวช
ปฏิบัติทั่วไป

เนื้อหาข้อสอบอัตนัยประยุกต์สำหรับการสอบ
ประมวลความรู้ไม่ควรมุ่งเน้นเนื้อหาที่เป็นสหสาขา
วิชา กล่าวคือต้องอาศัยองค์ความรู้ที่นักศึกษาได้ศึกษา
มาจากหลายภาควิชามาช่วยกันแก้ปัญหาผู้ป่วย ข้อสอบ
อัตนัยประยุกต์ที่นำมาสอบนักศึกษาแพทย์ทุกข้อใน
ปัจจุบันล้วนมีความเป็นสหสาขาวิชาทั้งสิ้น มีอาจารย์จาก
หลากหลายภาควิชามาร่วมกันออกข้อสอบ แต่อย่างไร
ก็ตามข้อสอบบางข้ออาจมีลักษณะการใช้ความรู้สหสาขา
วิชาแบบแยกเป็นส่วน ๆ กล่าวคืออาจารย์ต่างภาควิชากัน
ใช้การแบ่งงานออกเป็นส่วน ๆ อาจารย์ภาควิชาที่หนึ่งออก
ข้อสอบในตอนหนึ่งกับสอง อาจารย์ภาควิชาที่สองออก
ข้อสอบในตอนที่สามกับสี่ และอาจารย์ภาควิชาที่สามออก
ข้อสอบในตอนสุดท้ายกับหก ข้อสอบลักษณะนี้มักจะยาก
มาก เนื่องจากเป็นการใช้ความรู้เชิงลึกของแต่ละภาควิชา
ทีละเรื่อง เช่นซักประวัติ ตรวจร่างกายแล้วก็ไม่สามารถ
วินิจฉัยโรคได้ ต้องส่งต่อไปทำการตรวจเพิ่มเติมในอีก
ภาควิชาหนึ่ง ซึ่งผลการตรวจเพิ่มเติมก็แปลผลได้ยาก เมื่อ
ได้ข้อสรุปแล้วก็ต้องส่งต่อไปให้แพทย์อีกสาขาวิชาหนึ่ง
ทำการรักษา เมื่อรักษาแล้วก็มีการตรวจย้อนต้องส่งต่อ
ให้แพทย์อีกสาขาวิชาหนึ่งทำการแก้ไขภาวะแทรกซ้อนให้
เป็นต้น โดยทั่วไปแล้วข้อสอบอัตนัยประยุกต์ที่ใช้ความรู้
สหสาขาวิชาที่เป็นที่ต้องการในการสอบประมวลความรู้
รอบรู้นั้นไม่ควรเป็นการประเมินความรู้ในเชิงลึกทีละวิชา
ในข้อสอบแต่ละตอน แต่ควรเป็นการผสมผสานความรู้
จากหลากหลายสาขาวิชาในทุกขั้นตอน เช่น หญิงอายุ
30 ปี ปวดท้องน้อยคือ ๆ ตลอดเวลา 6 ชั่วโมง มีไข้ต่ำ ๆ
คลื่นไส้เล็กน้อย โจทย์ให้ผู้สอบซักประวัติเพื่อการวินิจฉัย
โรคซึ่งผู้สอบที่จะตอบคำถามได้ดีต้องอาศัยความรู้ทั้งโรค
ในระบบทางเดินอาหาร ทางเดินปัสสาวะ ฮอร์โมนสืบพันธุ์
สตรี กระดูกและกล้ามเนื้อ เป็นต้น

ข้อแนะนำในเรื่องเนื้อหาที่สำคัญคืออาจารย์
ผู้ออกข้อสอบต้องตระหนักว่าการสอบนี้เป็นการประเมิน
ความรู้เวชปฏิบัติทั่วไป มิใช่การประเมินความรู้เชิงลึก
ในศาสตร์ของแต่ละสาขาวิชา โรคหรือภาวะที่นำมาออก

เวบบันทึกศิริราช

บทความทั่วไป

ข้อสอบส่วนใหญ่ควรอยู่ในเกณฑ์มาตรฐานผู้ประกอบวิชาชีพเวชกรรมในกลุ่มที่ 1 หรือ 2 (โรคหรือภาวะที่แพทย์เวชปฏิบัติทั่วไปสามารถให้การดูแลด้วยตนเองได้ และพิจารณาส่งต่อในกรณีที่มีโรครุนแรงหรือซับซ้อน) โรคหรือภาวะที่อยู่ในเกณฑ์มาตรฐานฯ กลุ่มที่ 3 (โรคหรือภาวะที่แพทย์เวชปฏิบัติทำการดูแลเบื้องต้นแล้วให้ส่งต่อไปยังผู้เชี่ยวชาญ) ควรนำมาออกข้อสอบไม่มากนัก หากจะนำโรคหรือภาวะในเกณฑ์มาตรฐานฯ กลุ่มที่ 3 มาออกสอบ ต้องมุ่งเน้นการดูแลรักษาเบื้องต้นที่แพทย์เวชปฏิบัติทั่วไปพึงทำได้ ไม่ควรมุ่งประเด็นไปที่การรักษาโดยผู้เชี่ยวชาญ เฉพาะสาขามากจนเกินไป

(2) รูปแบบคำถาม

หลักการสำคัญของการวัดและประเมินผลคือ การเลือกใช้เครื่องมือที่เหมาะสมในการวัดผลการเรียนรู้ ข้อสอบอัตนัยประยุกต์ได้รับการพัฒนาขึ้นเพื่อประเมินทักษะในการตัดสินใจทางคลินิกเป็นสำคัญ สิ่งที่ยังเป็นปัญหาในข้อสอบอัตนัยประยุกต์บางข้อคือการเลือกถามคำถามในรูปแบบที่ไม่ตรงตามเป้าประสงค์ของการสอบอัตนัยประยุกต์ เช่นถามความจำขั้นพื้นฐาน โดยไม่ต้องคิดวิเคราะห์และตัดสินใจว่าจะทำหรือไม่ทำสิ่งใดกับผู้ป่วย รูปแบบคำถามที่ไม่เหมาะสมเหล่านี้เช่น ผู้ชายอายุ 40 ปี มีไข้สองเดือน จงถามประวัติ การใช้รูปแบบคำถามลักษณะนี้จะวัดเพียงว่าผู้เข้าสอบจดจำหัวข้อทั้งหมดของการซักประวัติในผู้ป่วยที่มีไข้เรื้อรังได้หรือไม่ และผู้สอบคนใดเขียนได้เร็วและครบถ้วนกว่ากัน ซึ่งอาจารย์สามารถใช้เครื่องมือประเมินผลชนิดอื่นในการวัดความจำขั้นพื้นฐานได้ดีกว่าการใช้ข้อสอบอัตนัยประยุกต์ การใช้ข้อสอบอัตนัยประยุกต์ควรมุ่งเน้นคำถามประเมินความสามารถในการวิเคราะห์ปัญหาผู้ป่วย และตัดสินใจสั่งการตรวจ หรือรักษาผู้ป่วยอย่างเหมาะสม

(3) จำนวนสถานการณ์ผู้ป่วยที่ใช้สอบ

ในการสอบประมวลความรู้ด้วยข้อสอบอัตนัยประยุกต์ของคณะแพทยศาสตร์ศิริราชพยาบาลที่ผ่านมา มีการใช้สถานการณ์ผู้ป่วยในข้อสอบตั้งแต่ 5 ถึง 8 ราย ถึงแม้ว่าจำนวนสถานการณ์ในการสอบระยะหลังมี

แนวโน้มเพิ่มขึ้น แต่หากพิจารณาในแง่ของความจำเพาะต่อบริบทของผู้ป่วย (case specificity) ที่ได้อภิปรายไปก่อนหน้านี้แล้วจะเห็นได้ว่ากรณีที่ผู้สอบแก้ปัญหาผู้ป่วยได้ 5 ถึง 8 รายนั้นน่าจะยังครอบคลุมประเด็นปัญหาทางคลินิกได้ไม่มากเพียงพอ และคะแนนสอบที่ได้มาน่าจะพัฒนาให้มีความเที่ยงสูงขึ้นได้อีกหากในการสอบมีจำนวนสถานการณ์มากขึ้น เนื่องด้วยรูปแบบข้อสอบอัตนัยประยุกต์ที่ใช้ในการสอบของคณะฯ ยังเน้นการสอบถามการจัดการปัญหาของผู้ป่วยตลอดตั้งแต่ต้นจนจบ (Patient management problem, PMP) จึงทำให้เวลาที่ใช้ในการสอบในแต่ละสถานการณ์ค่อนข้างนาน (แต่ละสถานการณ์มีคำถามย่อย 4 – 8 ข้อ ใช้เวลา 15 ถึง 30 นาทีต่อสถานการณ์) จึงทำให้ไม่สามารถสอบได้หลายสถานการณ์

หากพิจารณาจากข้อเสนอแนะของผู้เชี่ยวชาญในการประเมินผลที่ได้อภิปรายไปก่อนหน้านี้ที่แนะนำให้ใช้ข้อสอบอัตนัยประยุกต์แบบแก้ปัญหาสำคัญ แนวทางการพัฒนาข้อสอบอัตนัยประยุกต์ของคณะฯ ให้มีความครอบคลุมสถานการณ์ผู้ป่วยที่มากขึ้น และมีความเที่ยงของคะแนนสอบมากขึ้นคือการใช้ข้อสอบแบบแก้ปัญหาสำคัญมาแทนการจัดการปัญหาของผู้ป่วยตั้งแต่ต้นจนจบ กล่าวคือในแต่ละสถานการณ์ผู้ป่วย ข้อสอบควรมุ่งถามคำถามสำคัญเพียงสองหรือสามข้อ และเพิ่มจำนวนสถานการณ์ผู้ป่วยให้มากขึ้นนั่นเอง

(4) การนำเสนอข้อสอบ

การทำข้อสอบอัตนัยประยุกต์ ผู้สอบต้องทำงานภายใต้ข้อจำกัดด้านเวลา เวลาที่ใช้ในการตอบข้อสอบอัตนัยประยุกต์เป็นผลรวมของเวลาที่ใช้อ่านโจทย์ คิดวิเคราะห์ และเขียนคำตอบ ปัญหาสำคัญประการหนึ่งที่สร้างความลำบากให้กับผู้สอบคือปริมาณข้อมูลที่น่าเสนอให้ผู้สอบอ่านในสถานการณ์ผู้ป่วยแต่ละรายนั้นมีมาก ทำให้ผู้สอบต้องใช้เวลาในการอ่านมากและเหลือเวลาสำหรับเขียนคำตอบน้อย ถึงแม้ว่าในการนำเสนอข้อมูลของข้อสอบอัตนัยประยุกต์จะได้มีการแยกข้อมูลเดิมที่เคยนำเสนอไปก่อนหน้านี้ ออกจากข้อมูลใหม่ที่เพิ่มเติมขึ้นมาในการนำเสนอข้อสอบแต่ละตอนแล้วก็ตาม ด้วย

เวชบันทึกศิริราช

บทความทั่วไป

รายละเอียดที่น่าเสนอมีมาก ผู้สอบก็ยังคงมีความจำเป็น ต้องประมวลผลข้อมูลปริมาณมากอยู่ดี จากการทบทวน เนื้อหาของข้อสอบอัตนัยประยุกต์ที่ได้จัดสอบไปหลาย ครั้งพบว่าข้อสอบหลายข้อใช้ข้อมูลเพียงส่วนน้อยของที่ นำเสนอเท่านั้นก็สามารถนำไปสู่การแก้ปัญหาและการ ตัดสินใจเลือกการส่งตรวจหรือให้การรักษาผู้ป่วยได้อย่าง ถูกต้อง ดังนั้นแนวทางในการพัฒนาคุณภาพของข้อสอบ อัตนัยประยุกต์อีกทางหนึ่งคือการที่อาจารย์ผู้ออกข้อสอบ พึงตระหนักถึงข้อจำกัดเรื่องเวลาในการทำข้อสอบของ นักศึกษาและเขียนสถานการณ์ผู้ป่วยให้มีความกระชับ นำ เสนอเฉพาะข้อมูลที่มีความจำเป็นในการตัดสินใจให้การ ดูแลรักษาผู้ป่วยเท่านั้น ในการนำเสนอข้อมูลแต่ละตอน ควรต้องทบทวนว่าข้อมูลเก่าที่เคยให้ในขั้นตอนก่อนหน้า นั้นมีความจำเป็นต่อนำเสนอซ้ำทั้งหมดหรือไม่ หากทำได้ ควรทำการสรุปข้อมูลให้ผู้เข้าสอบ และตัดทอนข้อมูลที่ ไม่จำเป็นในการแก้ปัญหาขั้นตอนนั้น ๆ ออกไป ตัวอย่าง เช่น ในข้อสอบตอนที่หนึ่งมีการนำเสนอประวัติผู้ป่วยสั้น ๆ แล้วมีโจทย์ถามถึงประวัติที่จะซักเพิ่มเติม และการตรวจ ร่างกายที่จะทำเพื่อนำไปสู่การวินิจฉัยโรค ในข้อสอบตอนที่ สองอาจารย์นำเสนอประวัติและผลการตรวจร่างกาย เพิ่มเติมให้ แล้วมีโจทย์ถามถึงการวินิจฉัยโรค และการ ส่งตรวจทางห้องปฏิบัติการที่เหมาะสม ในข้อสอบตอนที่ สามอาจารย์นำเสนอข้อมูลการวินิจฉัยโรคของผู้ป่วย พร้อมผลการตรวจทางห้องปฏิบัติการ แล้วถามแนวทางการรักษา การนำเสนอข้อสอบในลักษณะนี้ในข้อสอบ หลายข้อมีการนำเสนอข้อมูลของโจทย์ซ้ำเดิมและค่อย ๆ เพิ่มข้อมูลขึ้นในทุกขั้นตอน ในข้อสอบตอนที่สองก็นำเสนอ ข้อมูลที่เสนอในตอนหนึ่งกับสอง ในข้อสอบตอนที่สามก็นำเสนอข้อมูลที่เสนอในตอนหนึ่ง สอง และ สาม ซึ่งเมื่อ ผ่านการสอบไปหลายตอนจะมีข้อมูลสะสมจำนวนมาก ที่ผู้สอบต้องอ่าน การนำเสนอข้อสอบที่มีประสิทธิภาพ มากกว่าควรมีการสรุปข้อมูลอย่างเหมาะสม ในข้อสอบ ตอนที่สาม หากได้ข้อสรุปการวินิจฉัยโรคแล้ว จะถาม แนวทางการรักษาโรค อาจารย์ควรพิจารณาตัดข้อมูล ประวัติและการตรวจร่างกายออก หากการสั่งการรักษา จำเป็นต้องทราบข้อมูลจากประวัติ หรือการตรวจร่างกาย บางอย่าง เช่น น้ำหนักตัว หรือ โรคร่วมที่ส่งผลต่อการ

วางแผนการรักษา ก็ให้นำเสนอเฉพาะข้อมูลที่ส่งผลต่อ การตัดสินใจในขั้นตอนนั้นเท่านั้น

การนำเสนอข้อสอบอัตนัยประยุกต์ด้วยระบบ คอมพิวเตอร์ก็เป็นอีกแนวทางหนึ่งที่คณะแพทยศาสตร์ ศิริราชพยาบาลเห็นความสำคัญ และได้ดำเนินการพัฒนา อย่างต่อเนื่อง คณะแพทยศาสตร์ศิริราชพยาบาลมีความ พร้อมในการพัฒนาด้านนี้มากพอสมควร เนื่องด้วยมีห้อง คอมพิวเตอร์ที่มีจำนวนคอมพิวเตอร์มากพอที่จะจัดให้ ผู้เข้าสอบทุกคนมีจอคอมพิวเตอร์ส่วนตัว มีการวางระบบ เครือข่ายให้มีการส่งผ่านข้อมูลระหว่างเครื่องคอมพิวเตอร์ ได้ดี และมีความเสถียรของระบบพอสมควร มีการวาง มาตรการรักษาความปลอดภัยของข้อมูลในระบบที่ดี สามารถควบคุมการเข้าออกของข้อมูลจากระบบเครือข่ายคอมพิวเตอร์ได้ จึงส่งผลให้คณะได้ปรับรูปแบบ การจัดสอบอัตนัยประยุกต์จากระบบสอบด้วยข้อสอบ กระดาษมาเป็นการนำเสนอข้อสอบบนจอคอมพิวเตอร์ ตั้งแต่ปีการศึกษา 2552 ซึ่งจากการสำรวจความเห็นของ นักศึกษาผู้เข้าสอบได้รับการตอบรับดีมาก นักศึกษาพึง พอใจกับการสอบในระบบนี้ในระดับมากถึงมากที่สุด อย่างไรก็ตามระบบการสอบนี้ยังมีโอกาสที่จะพัฒนา ให้ดีขึ้นได้อีก ในระบบการจัดสอบปัจจุบันของคณะฯ ยังคงเป็นรูปแบบที่ไม่ได้ใช้คอมพิวเตอร์อย่างเต็มรูปแบบ ยังคงให้ผู้สอบเขียนคำตอบลงในกระดาษคำตอบและเก็บ กระดาษในตอนท้ายของการสอบในแต่ละสถานการณ์ ผู้ป่วย การใช้ประโยชน์ของคอมพิวเตอร์ในการสอบ ปัจจุบันเน้นไปในการนำเสนอข้อมูลที่ให้ผู้สอบสามารถ เห็นภาพถ่ายรังสี ภาพการตรวจทางห้องปฏิบัติการ แผนภาพ ตาราง รวมถึงรูปของผู้ป่วยได้ โดยผู้สอบทุกคน เห็นภาพที่มีความละเอียดสูงเท่าเทียมกัน และทำให้การบริหารการสอบทำได้มีประสิทธิภาพมากขึ้น ดัดปัญหา ผู้สอบลักลอบเปิดดูข้อสอบในตอนต่อไปล่วงหน้า หรือทำ ข้อสอบในบางตอนเกินเวลา การแสดงเวลาที่เหลือในการ ทำข้อสอบแต่ละตอนบนหน้าจอทำให้ผู้สอบบริหารเวลา ในการทำข้อสอบได้ดีขึ้น

ระบบจัดสอบอัตนัยประยุกต์ด้วยคอมพิวเตอร์ อย่างเต็มรูปแบบที่ไม่ต้องมีการเขียนตอบในกระดาษเลย นั้นมีการจัดทำในต่างประเทศ^{12,23} แต่ต้องยอมรับว่าการ

เวชบันทึกศิริราช

บทความทั่วไป

สร้างระบบการจัดสอบอัตโนมัติประยุกต์ด้วยคอมพิวเตอร์อย่างเต็มรูปแบบนั้นเป็นงานที่ซับซ้อนและมีความท้าทายหลายอย่าง ทั้งในด้านผู้จัดสอบ ระบบเครือข่ายคอมพิวเตอร์ และผู้เข้าสอบ ในอนาคตอันใกล้นี้ทางการศึกษาฯ ยังไม่มีแนวทางที่จะพัฒนาการสอบอัตโนมัติประยุกต์เป็นระบบคอมพิวเตอร์อย่างเต็มรูปแบบ ด้วยข้อจำกัดสำคัญสามประการคือ ความพร้อมของผู้เข้าสอบ ความพร้อมของผู้ตรวจข้อสอบ และความพร้อมของระบบการสื่อสารระหว่างผู้ใช้กับคอมพิวเตอร์ กล่าวคือ ผู้เข้าสอบจำนวนไม่น้อยยังไม่คุ้นเคยกับการพิมพ์คำตอบที่มีทั้งภาษาไทยและภาษาอังกฤษผสมกันภายในเวลาที่จำกัด อาจารย์ผู้ตรวจข้อสอบจำนวนไม่น้อยยังไม่สะดวกที่จะทำการตรวจข้อสอบและกรอกคะแนนบนหน้าจอคอมพิวเตอร์ในสถานที่และเวลาที่กำหนด และการสร้างระบบการสื่อสารระหว่างคอมพิวเตอร์กับผู้ใช้ให้ทั้งนำเสนอข้อมูลผู้ป่วยที่มีรายละเอียดมาก พร้อมกับตอบรับคำตอบที่มีทั้งอักษร ตัวเลข และสัญลักษณ์พิเศษ ที่ผู้เข้าสอบจะพิมพ์เข้าเครื่องพร้อม ๆ กันหลายร้อยคนโดยมีการควบคุมเวลาอย่างรัดกุมด้วย ยังเป็นสิ่งที่ยากในในระบบเครือข่ายคอมพิวเตอร์ในปัจจุบัน ดังนั้นในอนาคตอันใกล้นี้ทิศทางการพัฒนาระบบการจัดสอบข้อสอบอัตโนมัติยังคงมุ่งเน้นไปในรูปแบบการนำเสนอข้อสอบผ่านจอภาพคอมพิวเตอร์ ร่วมกับการเขียนตอบในกระดาษคำตอบอยู่

แต่ถึงแม้ว่าจะคงการจัดสอบอัตโนมัติในรูปแบบผสมผสานเช่นนี้ ผู้นิพนธ์ก็ยังเห็นว่าสิ่งที่จะระบบการนำเสนอข้อมูลผ่านจอคอมพิวเตอร์สามารถทำให้ดีขึ้นได้ เช่นการทำให้ภาพมีรายละเอียดสูงขึ้น การเปิดโอกาสให้ผู้เข้าสอบสามารถขยายภาพเพื่อดูรายละเอียดในบางส่วน การปรับรูปแบบการนำเสนออักษร และพื้นหลังของจอภาพให้ผู้เข้าสอบอ่านข้อมูลได้ง่ายขึ้น เป็นต้น ซึ่งสิ่งเหล่านี้จะได้มีการศึกษาหาแนวทางในการพัฒนาในการสอบอัตโนมัติประยุกต์ครั้งต่อไป แต่อย่างไรก็ตามด้วยศักยภาพของระบบการจัดสอบในปัจจุบัน ผู้นิพนธ์ยังมีความเห็นว่าอาจารย์ผู้ออกข้อสอบก็ยังไม่ได้อาศัยศักยภาพของระบบอย่างเต็มที่ ยังมีข้อสอบหลายข้อที่ใช้การบรรยายสิ่งตรงพบที่สามารถมองเห็นเป็นภาพได้

แต่นำมาเขียนเป็นอักษรบรรยายสิ่งตรงพบดังกล่าวซึ่งทำให้ผู้เข้าสอบไม่ได้คิดวิเคราะห์และแปลผลการตรวจด้วยตนเอง แนวทางการพัฒนาข้อสอบอัตโนมัติประยุกต์ที่สมควรได้รับการส่งเสริมในระบบการจัดสอบปัจจุบันคือการใช้สื่อที่เป็นรูปภาพในข้อสอบให้มากขึ้น ไม่ว่าจะเป็นการตรวจร่างกายจากการดู การดูภาพรังสี การดูคลื่นไฟฟ้าหัวใจ การดูสิ่งตรงพบด้วยกล้องจุลทรรศน์ ล้วนแล้วแต่ควรนำเสนอเป็นรูปภาพทั้งสิ้น

บทสรุป

ในบทความนี้ผู้นิพนธ์ได้กล่าวถึงความรู้พื้นฐานในการสร้างข้อสอบอัตโนมัติประยุกต์โดยได้สรุปลักษณะพื้นฐานของข้อสอบอัตโนมัติประยุกต์ พัฒนาการของข้อสอบประเภทนี้จากรูปแบบการจัดการปัญหาผู้ป่วยเป็นการแก้ปัญหาสำคัญ มีการสรุปขั้นตอนสำคัญในการสร้างข้อสอบอัตโนมัติประยุกต์ห้าขั้นตอน ได้แก่ (1) ตั้งกลุ่มพัฒนาข้อสอบ, (2) เลือกปัญหาทางคลินิก, (3) กำหนดปัญหาสำคัญ, (4) เขียนโจทย์คำถาม, และ (5) กำหนดเกณฑ์การให้คะแนน และในตอนท้ายได้มีการนำหลักการพัฒนาข้อสอบต่าง ๆ ที่กล่าวมาแล้วมาวิเคราะห์สถานการณ์การจัดสอบอัตโนมัติประยุกต์สำหรับนักศึกษาแพทย์คณะแพทยศาสตร์ศิริราชพยาบาลและเสนอแนะแนวทางในการพัฒนาคุณภาพการสอบอัตโนมัติประยุกต์สี่แนวทาง ได้แก่ (1) เนื้อหาข้อสอบ, (2) รูปแบบคำถาม, (3) จำนวนสถานการณ์ผู้ป่วย, และ (4) การนำเสนอข้อสอบ ผู้นิพนธ์เชื่อมั่นว่าหากการจัดสอบอัตโนมัติประยุกต์ได้รับการพัฒนาอย่างเหมาะสมจะนำไปสู่การประเมินความรู้ และทักษะการตัดสินใจดูแลผู้ป่วยในระดับคลินิกที่มีประสิทธิภาพ

เอกสารอ้างอิง

1. Downing SM. Assessment of knowledge with written test forms. In: Norman GR, van der Vleuten C, Newble DI, editors. International handbook of research in medical education. Dordrecht: Kluwer Academic Publishers, 2002:647 - 72.
2. Epstein RM. Assessment in medical education. New Engl J Med 2007;356:387-96.
3. The Board of Censors of the Royal College of General Practitioners. The modified essay question. J Roy Coll Gen Practit 1971;21:373-6.
4. Farmer EA, Page G. A practical guide to assessing clinical decision-making skills using the key features approach. Med Educ 2005;39:1188 -94.

เวบบันทึกศิริราช

บทความทั่วไป

5. McGuire CH, Babbott D. Simulation technique in the measurement of problem solving skills. *J Educ Meas* 1967;4:1-10.
6. จินตนา สิรินาวิน, สาริต วรรมแสง. ทักษะทางคลินิก, พิมพ์ครั้งที่ 2. กรุงเทพฯ: หมอชาวบ้าน, 2549.
7. Hodgkin K, Knox JDE. Problem centered learning. London, United Kingdom: Churchill Livingstone, 1975.
8. Stratford P, Pierce-Fenn H. Modified essay question. *Phys Ther* 1985; 65(1075-9).
9. Feletti GI, Smith EK. Modified essay questions: Are they worth the effort? *Med Educ* 1986;20:126 - 32.
10. Rabinowitz HK. The modified essay question: An evaluation of its use in a family medicine clerkship. *Med Educ* 1987;21:114-8.
11. Wallerstedt S, Erickson G, Wallerstedt SM. Short answer questions or modified essay questions - More than a technical issue. *Int J Clin Med* 2012;3:28-30.
12. Lim EC, Seet RC, Oh VMS, Chia B, Aw M, S Q, et al. Computer-based testing of the modified essay question: The Singapore experience. *Med Teach* 2007;29:e261-8.
13. Norman G, Bordage G, Curry L, et al. Review of recent innovations in assessment. In: Wakeford R, editor. Directions in clinical assessment: Report of the Cambridge conference on the Assessment of Clinical competence. Cambridge: Office of the Regius Professor of Physic, Cambridge University School of clinical Medicine, 1985:8-27.
14. Page G, Bordage G. The Medical Council of Canada's key features project: A more valid written examination of clinical decision-making skills. *Acad Med* 1995;70:104-10.
15. Neufeld VR, Norman GR, Barrows HS, Feightner JW. Clinical problem solving by medical students: A longitudinal and cross-sectional analysis. *Med Educ* 1981;15:315-22.
16. Perkins DN, Salomon G. Are cognitive skills context-bound? *Educ Researcher* 1989;18:16-25.
17. van der Vleuten CPM, Swanson DB. Assessing clinical skills with standardized patients: The state of the art. *Teach Learn Med* 1990;2 (58-76).
18. Eva KW. On the generality of specificity. *Med Educ* 2003;37(7): 587-88.
19. Bordage G, Page G. An alternate approach to PMPs, the key feature concept. In: Hart I, Harden R, editors. Further developments in assessing clinical competence. Montreal: Can-Heal Publications, 1987:57-75.
20. Page G, Bordage G, Allen T. Developing key features problems and examinations to assess clinical decision-making skills. *Acad Med* 1995;70:194-201.
21. Norman G, Bordage G, Page G, Keane D. How specific is case specificity? *Med Educ* 2006;40:618-23.
22. Hambleton RK, Pitoniak MJ. Setting performance standards. In: Brennan RL, editor. Educational measurement, 4th ed. Westport, CT: Praeger publishers, 2006:433-70.
23. Federation of State Medical Boards of the United States, National Board of Medical Examiners. USMLE Step 3: Content description and general information, Available from http://www.usmle.org/pdfs/step-3/2014content_Step3.pdf. June 2014.

ตามปกหน้าเวบบันทึกศิริราช ปีที่ 7 ฉบับที่ 2 กรกฎาคม-ธันวาคม 2557 หน้า 74-83 เรื่อง
“หน้ากากครอบบกล่องเสียง Laryngeal Mask Airway (LMA)” โดย อรุณทัย ศิริอัศวกุล

ขอแก้ไขเป็น

เวบบันทึกศิริราช

ปีที่ 7 ฉบับที่ 2 กรกฎาคม-ธันวาคม 2557 หน้า 74-83 เรื่อง

“หน้ากากครอบบกล่องเสียง Laryngeal Mask Airway (LMA)” โดย อังศุมาศ หวังดี

และได้ทำการแก้ไข pdf เรียบร้อยแล้ว

ผลลัพธ์การปฏิบัติงานของ



นายแพทย์ X

อาจารย์ที่ปรึกษา อาจารย์ A

ตามการประเมินด้วยแฟ้มสะสมผลงานการ (Portfolio)

ปีการศึกษา 2554-2556

Competency based portfolio assessment

Academic year 2011-2013

สาส์นจากหัวหน้าภาควิชา

ภาควิชาสูติศาสตร์-นรีเวชวิทยา คณะแพทยศาสตร์ศิริราชพยาบาล มหาวิทยาลัยมหิดล ขอแสดงความยินดีกับ **นายแพทย์ A** ที่สำเร็จการฝึกอบรมแพทย์ประจำบ้าน สาขาสูติศาสตร์-นรีเวชวิทยา ระหว่างปีการศึกษา 2553-2555

ตลอดระยะเวลาสามปีที่ผ่านมา ภาควิชาฯ ได้ดำเนินการประเมินคุณสมบัติด้านต่างๆ ของท่าน ได้แก่ ความรู้ ทักษะหัตถการ การวิจัย และพฤติกรรมกรปฏิบัติงาน ในรูปแบบ Portfolio ดังผลสรุปในเอกสารฉบับนี้

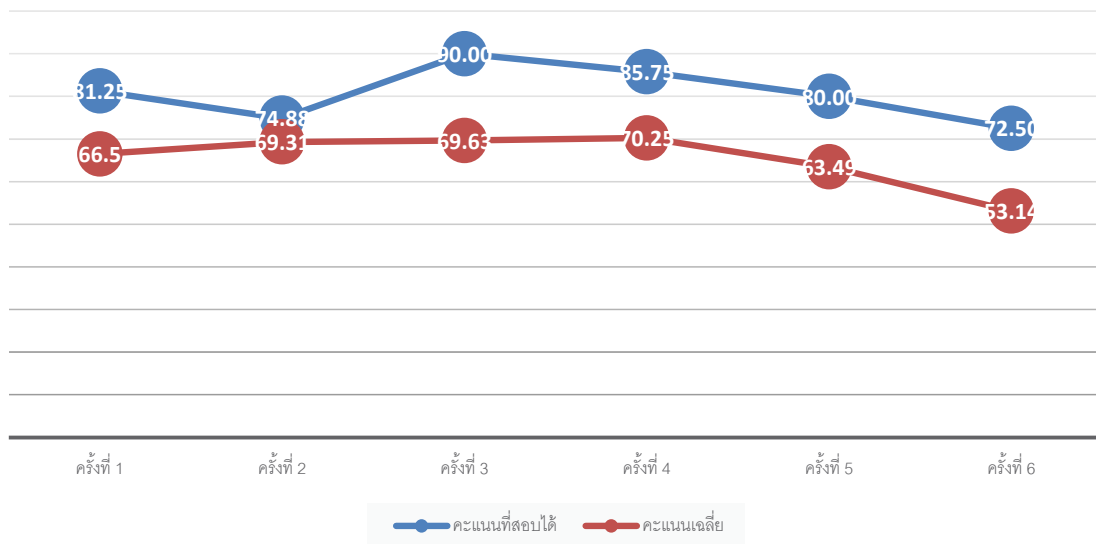
ภาควิชาฯ ขออำนวยการให้ท่านประสบความสำเร็จในการดำเนินชีวิตครอบครัว และหน้าที่การงาน ตลอดไป

ศาสตราจารย์คลินิก นายแพทย์ชาญชัย วันทนาศิริ
หัวหน้าภาควิชาสูติศาสตร์-นรีเวชวิทยา
คณะแพทยศาสตร์ศิริราชพยาบาล มหาวิทยาลัยมหิดล

การประเมินความรู้ทางสูติศาสตร์-นรีเวชวิทยา

(Knowledge assessment)

ผลการสอบวัดระดับความรู้ทางวิชาการ



ชั้นปีที่ 1

การสอบครั้งที่	คะแนนรวม	คะแนนเฉลี่ย	คะแนนที่สอบได้	ลำดับที่
1	100	66.50	81.25	1
2	100	69.31	74.88	3
3	100	69.63	90.00	1
4	100	70.25	85.75	1
5	100	63.49	80.00	1
6	100	53.14	72.50	2

ผลการสอบตามหลักสูตรประกาศนียบัตรบัณฑิตชั้นสูงสาขาวิทยาศาสตร์การแพทย์คลินิก:

The Higher Graduate Diploma (Clinical Medical Sciences) คณะแพทยศาสตร์ศิริราชพยาบาล

ผ่าน ได้รับประกาศนียบัตรเมื่อ 25 พฤษภาคม 2555

ไม่ผ่าน

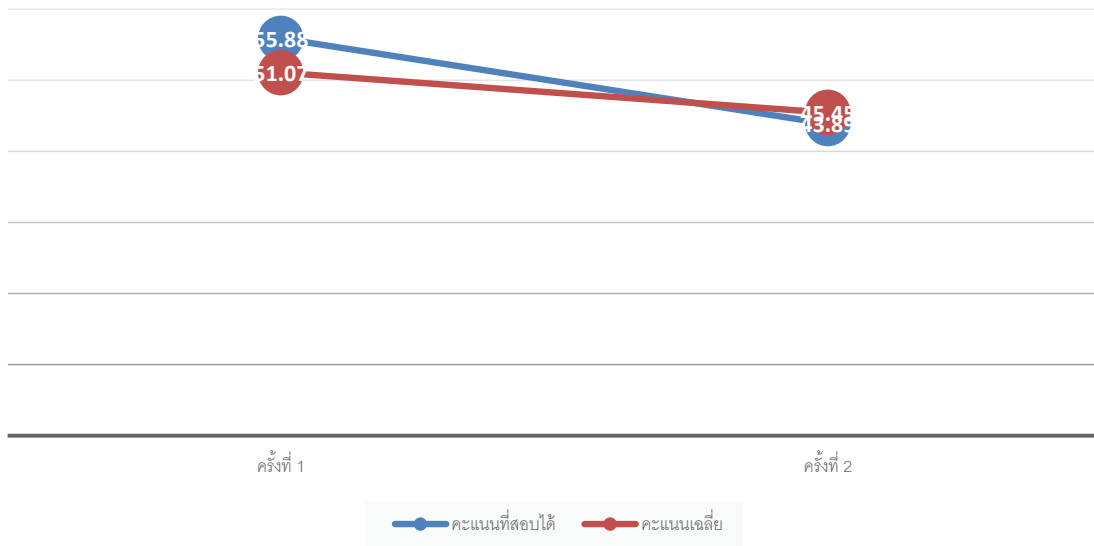
การสอบวิชาภาษาอังกฤษและกฎหมายทางการแพทย์ครั้งที่ 1

ผ่าน ไม่ผ่าน

การสอบวิชาภาษาอังกฤษและกฎหมายทางการแพทย์ครั้งที่ 2 (กรณีสอบไม่ผ่านครั้งแรก)

ผ่าน ไม่ผ่าน

ผลการสอบวัดระดับความรู้ทางวิชาการ



ชั้นปีที่ 2

การสอบครั้งที่	คะแนนรวม	คะแนนเฉลี่ย	คะแนนที่สอบได้	ลำดับที่
1	100	51.07	55.88	5
2	100	45.45	43.89	10

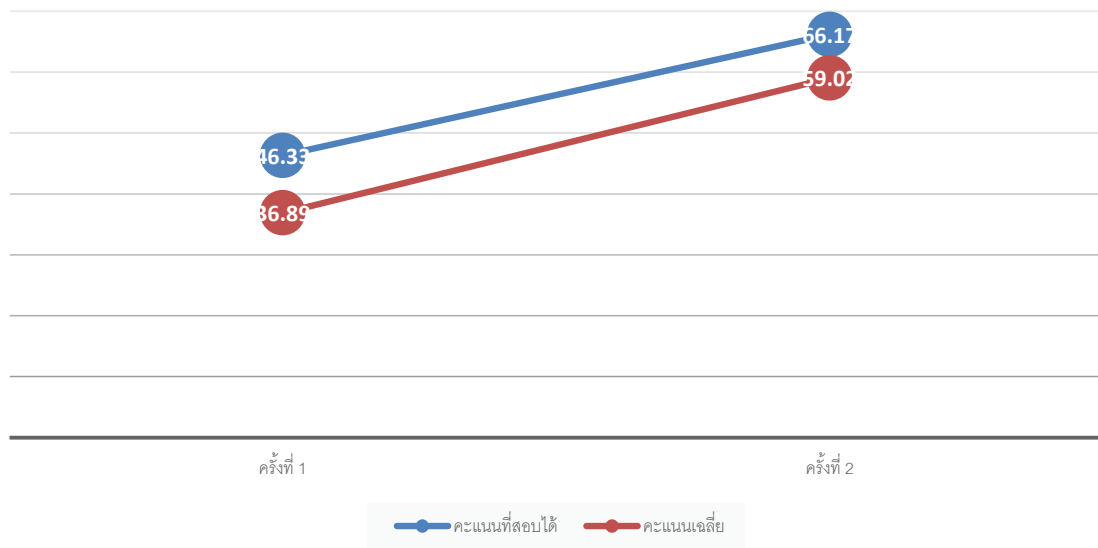
การสอบ OSLER ในสถาบัน ครั้งที่ 1

ผ่าน ไม่ผ่าน

การสอบ Basic science ราชวิทยาลัยสูตินรีแพทย์แห่งประเทศไทย

ผ่าน ไม่ผ่าน

ผลการสอบวัดระดับความรู้ทางวิชาการ



ชั้นปีที่ 3

การสอบครั้งที่	คะแนนรวม	คะแนนเฉลี่ย	คะแนนที่สอบได้	ลำดับที่
1	100	36.89	46.33	2
2	100	59.02	66.17	1

การสอบ OSLER ในสถาบัน ครั้งที่ 2

ผ่าน ไม่ผ่าน

การสอบ OSLER ราชวิทยาลัยสูตินรีแพทย์แห่งประเทศไทย ครั้งที่ 1

ผ่าน ไม่ผ่าน

การสอบ OSLER ราชวิทยาลัยสูตินรีแพทย์แห่งประเทศไทย ครั้งที่ 2 (กรณีสอบครั้งแรกไม่ผ่าน)

ผ่าน ไม่ผ่าน

การสอบงานวิจัย ราชวิทยาลัยสูตินรีแพทย์แห่งประเทศไทย

ผ่าน ไม่ผ่าน

หัตถการสำคัญทางสูติศาสตร์-นรีเวชวิทยาที่ปฏิบัติ
ขณะเป็นแพทย์ประจำบ้านชั้นปีที่ 3
 (Clinical skills assessment when being the 3rd year resident)

การผ่าตัดทางนรีเวช

การผ่าตัด	จำนวน
Total abdominal hysterectomy +/- bilateral salpingoophorectomy	19
Vaginal hysterectomy +/- AP repair	4
Adnexal surgery: Salpingectomy/Salpingotomy/Salpingostomy	21
Cervical conization	11

การผ่าตัดทางสูติศาสตร์

การผ่าตัด	จำนวน
Cesarean delivery	55
Tubal sterilization	3
Dilatation and curettage	16
Vacuum extraction/Forceps extraction	4
Breech assisting	
Manual removal of placenta	2

หมายเหตุ

จำนวนหัตถการเป็นจำนวนโดยประมาณ เนื่องจากอยู่ระหว่างกระบวนการพัฒนาและปรับปรุงระบบเก็บข้อมูลหัตถการแพทย์ประจำบ้าน ภาควิชาสูติศาสตร์-นรีเวชวิทยา

การทำงานวิจัยระดับแพทย์ประจำบ้าน
(Research competency)

เรื่อง **Prevalence and Associating Factors of Sexual Dysfunction in Women Who Use Intrapartum Device (IUD)**

อาจารย์ผู้ควบคุมผู้ช่วยศาสตราจารย์นายแพทย์ธันยรัตน์ วงศ์วนานุรักษ์

ข้อมูลสำคัญสำหรับงานวิจัย

1. ผ่าน SIRB เมื่อ 21 กุมภาพันธ์ 2555
เลขที่ 813/2554 (EC3)
2. ประกวดการนำเสนองานวิจัยในการประชุมราชวิทยาลัยสูตินรีแพทย์แห่งประเทศไทย
วันที่ 26 พฤศจิกายน 2556
 เข้าร่วมนำเสนอ ไม่ได้รับรางวัล
 เข้าร่วมนำเสนอ ได้รับรางวัล ชมเชย
3. การตีพิมพ์ในวารสารวิชาการ
 ไม่ได้ตีพิมพ์
 ได้รับการตีพิมพ์ (ระบุรายละเอียดวารสาร) J Med Assoc Thai 2014
Full text. E-Journal: <http://Jmatonline.com>

ผลการประเมินเจตคติและพฤติกรรมการปฏิบัติงานของแพทย์ประจำบ้าน
(Multisources feedback)

แพทย์ประจำบ้านจะได้รับการประเมินในประเด็นต่อไปนี้

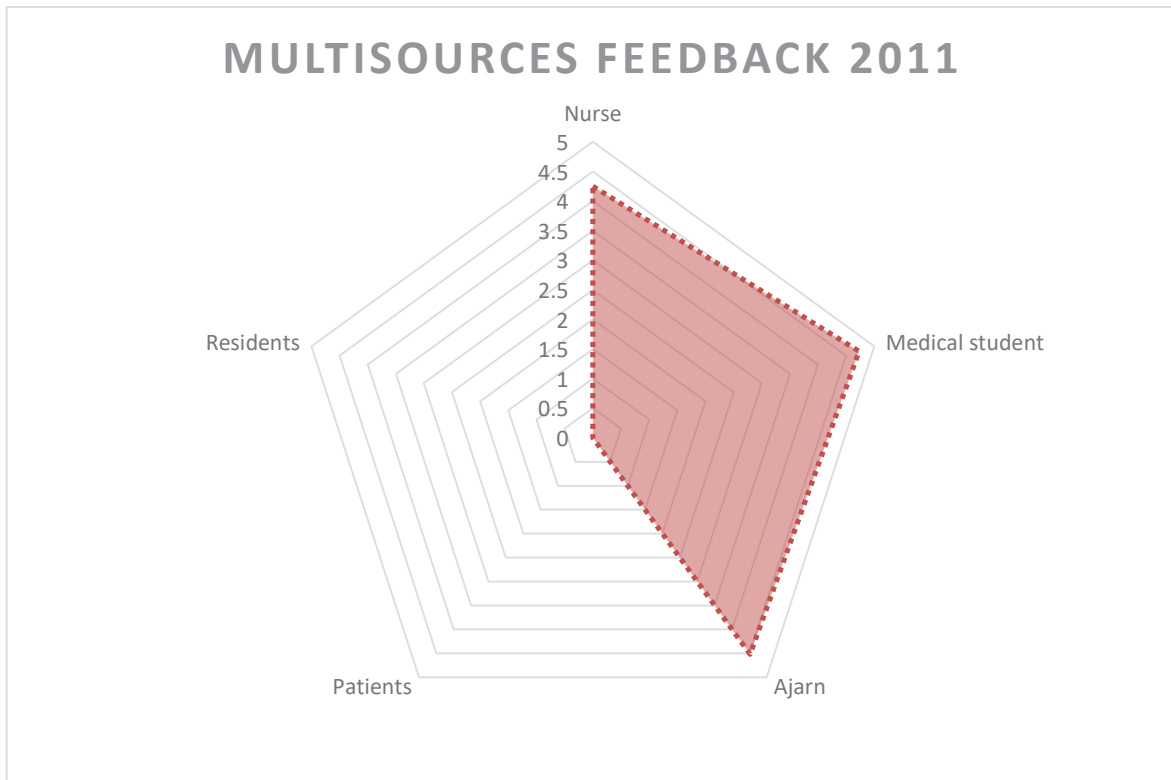
1. ความรู้ความสามารถด้านวิชาการ

2. ทักษะพื้นฐานในการปฏิบัติงาน

ได้แก่ ทักษะการสื่อสารกับเพื่อนร่วมงานและผู้ป่วย/ญาติ การบันทึกรายงานผู้ป่วย การทำงานร่วมกับผู้อื่น และบุคลิกภาพขณะปฏิบัติงาน

3. คุณธรรมและจริยธรรม

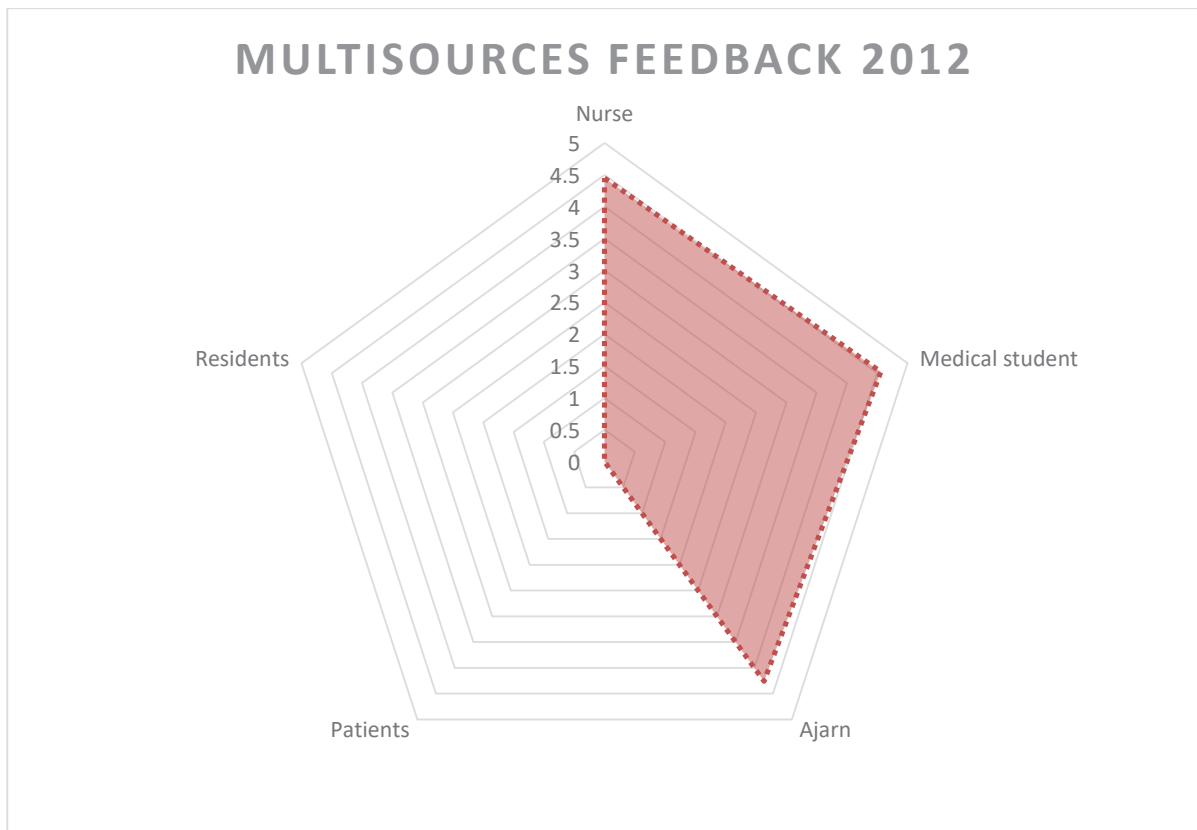
ได้แก่ ความรับผิดชอบ ความเสียสละ ความตรงต่อเวลา ความซื่อสัตย์ การปฏิบัติตามระเบียบข้อบังคับ และอภัย/น้ำใจ/ความเอื้อเฟื้อต่อผู้อื่น



ชั้นปีที่ 1 ปีการศึกษา 2554

Rotation	อาจารย์ (5 คะแนน)	นักศึกษาแพทย์ (5 คะแนน)	พยาบาล (5 คะแนน)	แพทย์ ประจำบ้าน	ผู้รับบริการ
พระศรีฯ 9/2			4.61		
LR เข้า		5.00	3.76		
LR พิเศษเข้า			4.61		
นรีเวช 1	4.90	5.00	4.00		
นรีเวช 1 (2)	4.50	4.90	4.00		
พระศรีฯ 10/2			4.46		
พระศรีฯ 9/1+ANC			5.00		
LR ดึก			4.00		
LR พิเศษป่วย			4.30		
นรีเวช 2	4.20	4.50	4.56		
Onco	4.50	4.30	3.84		
พระศรีฯ 10/3		5.00	4.30		
พระศรีฯ 10/1		4.46	3.91		
คะแนนเฉลี่ย	4.52	4.73	4.25		

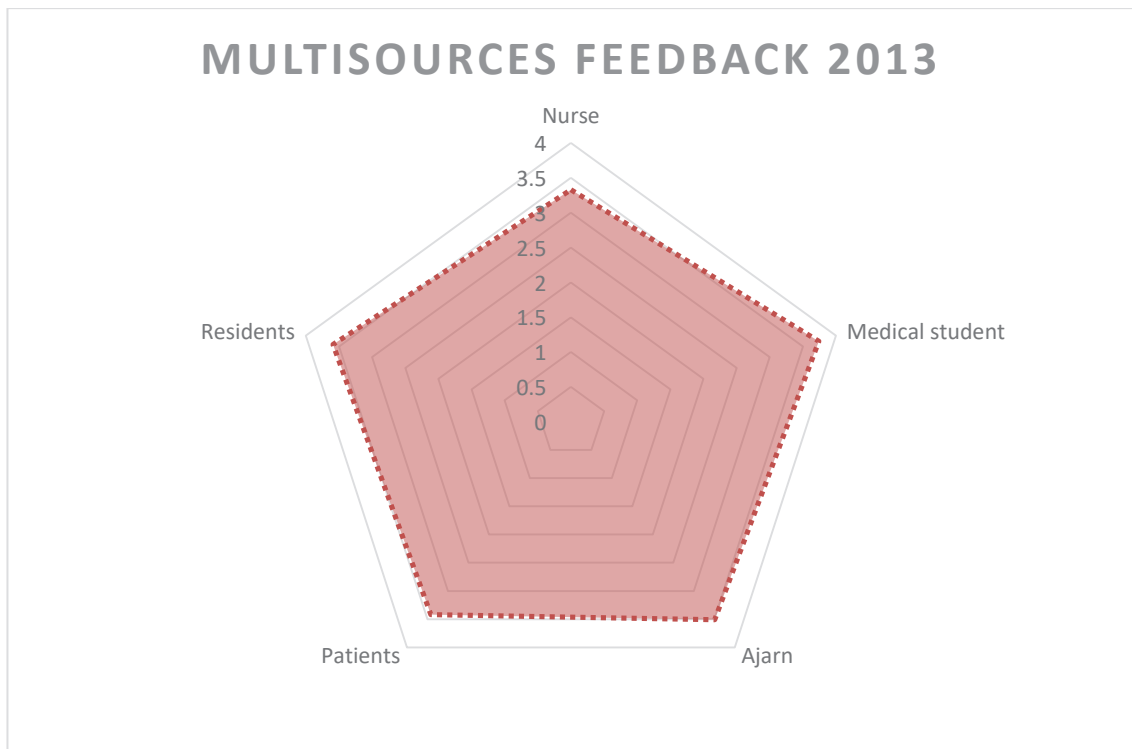
*ยังไม่มีกรประเมินจากแพทย์ประจำบ้านและผู้รับบริการในปีการศึกษา 2554



ชั้นปีที่ 2 ปีการศึกษา 2555

Rotation	อาจารย์ (5 คะแนน)	นักศึกษาแพทย์ (5 คะแนน)	พยาบาล (5 คะแนน)	แพทย์ ประจำบ้าน	ผู้รับบริการ
นรีเวช 1	3.93	4.00	4.53		
เลิดสิน	4.67				
พระศรีฯ 13/1	4.35		4.61		
LR ดึก			4.00		
Onco	4.17	4.20	3.23		
พระศรีฯ 14/2			5.00		
นรีเวช 2	4.11	4.50	5.00		
สระบุรี	4.47				
พระศรีฯ 13/2	4.40		4.70		
พระศรีฯ 10/1		4.70	4.50		
พระศรีฯ 14/1	4.00		4.23		
LR เช้า		5.00	4.69		
พระศรีฯ 10/3		5.00	4.56		
คะแนนเฉลี่ย	4.26	4.56	4.45		

*ยังไม่มีการประเมินจากแพทย์ประจำบ้านและผู้รับบริการในปีการศึกษา 2555



ชั้นปีที่ 3 ปีการศึกษา 2556

Rotation	อาจารย์ (4 คะแนน)	แพทย์ประจำบ้าน (4 คะแนน)	พยาบาล (4 คะแนน)	นักศึกษาแพทย์ (4 คะแนน)	ผู้รับบริการ (4 คะแนน)
นรีเวช 1	3.50	3.80	3.40	3.90	3.03
STD	3.70		3.20		
พระศรีฯ 10/1		2.62	4.00	3.70	3.36
LR พิเศษ		3.90	3.08		
OPD GYN			2.90		3.40
Septic		3.75	3.10	4.00	3.26
วิสัญญี	3.75				
นรีเวช 2	3.90	4.00	3.85	3.87	3.74
Infertile	3.20				
นครปฐม	3.00				
OPD ANC			3.75		3.73
ONCO	3.60	3.81	3.02		
LR เข้า		3.25	3.08	3.50	
Surgery	3.47				
คะแนนเฉลี่ย	3.51	3.59	3.33	3.74	3.42

*เริ่มการประเมินจากนักศึกษาแพทย์และผู้รับบริการ ในปีการศึกษา 2556

ผลลัพธ์การปฏิบัติงานของ



แพทย์หญิง Y

อาจารย์ที่ปรึกษา อาจารย์ B

ตามการประเมินด้วยแฟ้มสะสมพัฒนาการ (Portfolio)

ปีการศึกษา 2554-2556

Competency based portfolio assessment

Academic year 2011-2013

สาส์นจากหัวหน้าภาควิชา

ภาควิชาสูติศาสตร์-นรีเวชวิทยา คณะแพทยศาสตร์ศิริราชพยาบาล มหาวิทยาลัยมหิดล ขอแสดงความยินดีกับ **แพทย์หญิง B** ที่สำเร็จการฝึกอบรมแพทย์ประจำบ้าน สาขาสูติศาสตร์-นรีเวชวิทยา ระหว่างปีการศึกษา 2553-2555

ตลอดระยะเวลาสามปีที่ผ่านมา ภาควิชาฯ ได้ดำเนินการประเมินคุณสมบัติด้านต่างๆ ของท่าน ได้แก่ ความรู้ ทักษะหัตถการ การวิจัย และพฤติกรรมกรปฏิบัติงาน ในรูปแบบ Portfolio ดังผลสรุปในเอกสารฉบับนี้

ภาควิชาฯ ขออำนวยการให้ท่านประสบความสำเร็จในการดำเนินชีวิตครอบครัว และหน้าที่การงาน ตลอดไป

ศาสตราจารย์คลินิก นายแพทย์ชาญชัย วันทนาศิริ

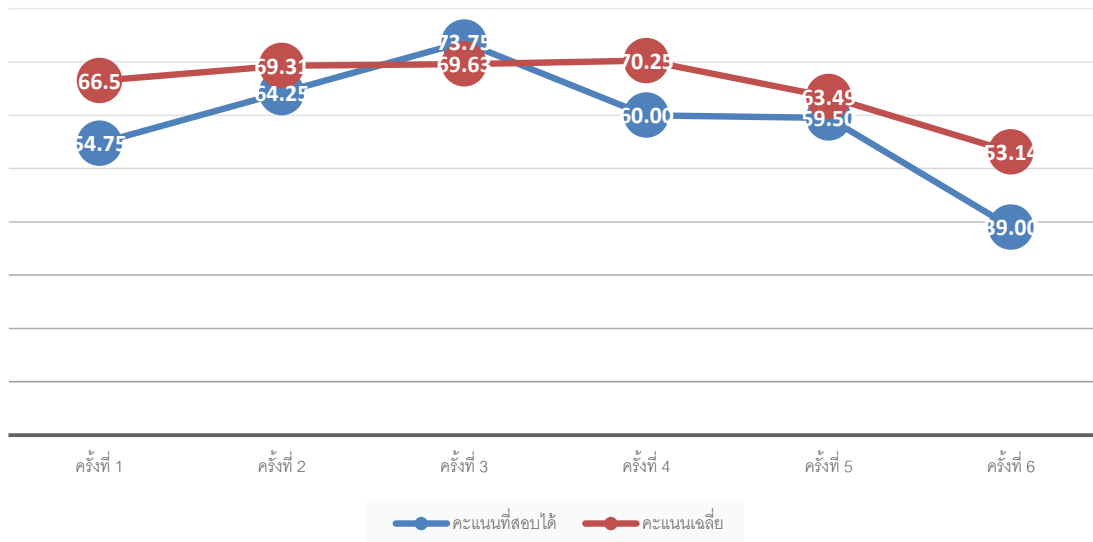
หัวหน้าภาควิชาสูติศาสตร์-นรีเวชวิทยา

คณะแพทยศาสตร์ศิริราชพยาบาล มหาวิทยาลัยมหิดล

การประเมินความรู้ทางสูติศาสตร์-นรีเวชวิทยา

(Knowledge assessment)

ผลการสอบวัดระดับความรู้ทางวิชาการ



ชั้นปีที่ 1

การสอบครั้งที่	คะแนนรวม	คะแนนเฉลี่ย	คะแนนที่สอบได้	ลำดับที่
1	100	66.50	54.75	13
2	100	69.31	64.25	11
3	100	69.63	73.75	4
4	100	70.25	60.00	13
5	100	63.49	59.50	11
6	100	53.14	39.00	13

ผลการสอบตามหลักสูตรประกาศนียบัตรบัณฑิตชั้นสูงสาขาวิทยาศาสตร์การแพทย์คลินิก:

The Higher Graduate Diploma (Clinical Medical Sciences) คณะแพทยศาสตร์ศิริราชพยาบาล

ผ่าน ได้รับประกาศนียบัตรเมื่อ 25 พฤษภาคม 2555

ไม่ผ่าน

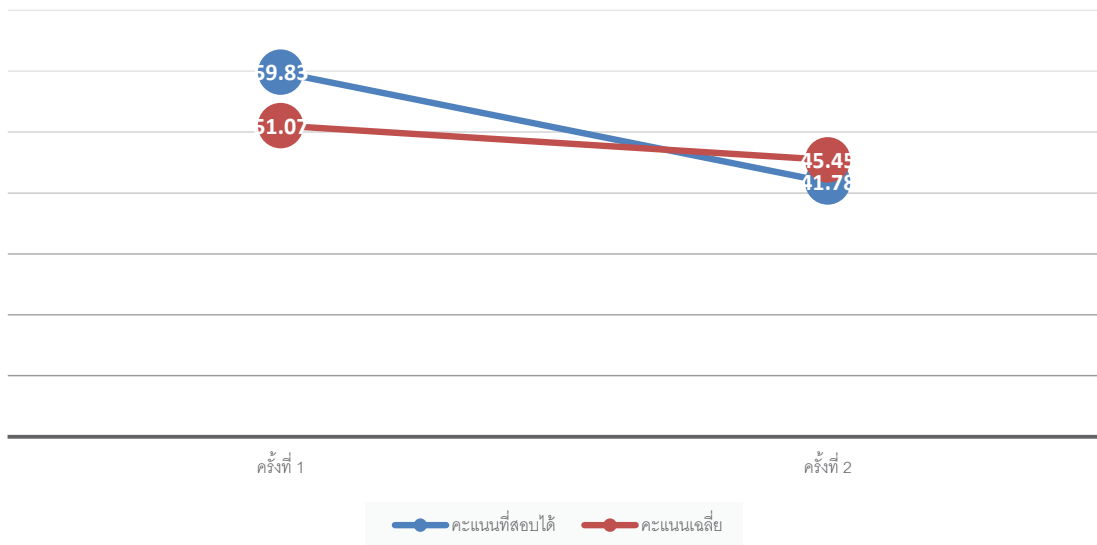
การสอบวิชาภาษาอังกฤษและกฎหมายทางการแพทย์ครั้งที่ 1

ผ่าน ไม่ผ่าน

การสอบวิชาภาษาอังกฤษและกฎหมายทางการแพทย์ครั้งที่ 2 (กรณีการสอบครั้งที่ 1 ไม่ผ่าน)

ผ่าน ไม่ผ่าน

ผลการสอบวัดระดับความรู้ทางวิชาการ



ชั้นปีที่ 2

การสอบครั้งที่	คะแนนรวม	คะแนนเฉลี่ย	คะแนนที่สอบได้	ลำดับที่
1	100	51.07	59.83	4
2	100	45.45	41.78	12

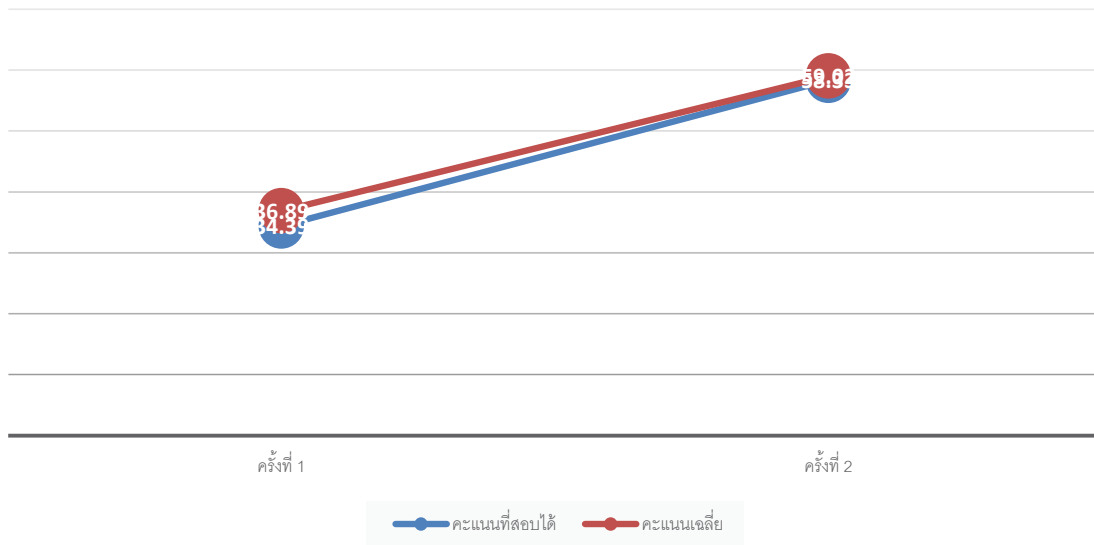
การสอบ OSLER ในสถาบัน ครั้งที่ 1

ผ่าน ไม่ผ่าน

การสอบ Basic science ราชวิทยาลัยสูตินรีแพทย์แห่งประเทศไทย

ผ่าน ไม่ผ่าน

ผลการสอบวัดระดับความรู้ทางวิชาการ



ชั้นปีที่ 3

การสอบครั้งที่	คะแนนรวม	คะแนนเฉลี่ย	คะแนนที่สอบได้	ลำดับที่
1	100	36.89	34.39	10
2	100	59.02	58.33	10

การสอบ OSLER ในสถาบัน ครั้งที่ 2

ผ่าน ไม่ผ่าน

การสอบ OSLER ราชวิทยาลัยสูตินรีแพทย์แห่งประเทศไทย ครั้งที่ 1

ผ่าน ไม่ผ่าน

การสอบ OSLER ราชวิทยาลัยสูตินรีแพทย์แห่งประเทศไทย ครั้งที่ 2 (กรณีสอบครั้งแรกไม่ผ่าน)

ผ่าน ไม่ผ่าน

การสอบงานวิจัย ราชวิทยาลัยสูตินรีแพทย์แห่งประเทศไทย

ผ่าน ไม่ผ่าน

หัตถการสำคัญทางสูติศาสตร์-นรีเวชวิทยาที่ปฏิบัติ

ขณะเป็นแพทย์ประจำบ้านชั้นปีที่ 3

(Clinical skills assessment when being the 3rd year resident)

การผ่าตัดทางนรีเวช

การผ่าตัด	จำนวน
Total abdominal hysterectomy +/- bilateral salpingoophorectomy	14
Vaginal hysterectomy +/- AP repair	7
Adnexal surgery: Salpingectomy/Salpingotomy/Salpingostomy	4
Cervical conization	2

การผ่าตัดทางสูติศาสตร์

การผ่าตัด	จำนวน
Cesarean delivery	43
Tubal sterilization	1
Dilatation and curettage	5
Vacuum extraction/Forceps extraction	5
Breech assisting	
Manual removal of placenta	6

หมายเหตุ

จำนวนหัตถการเป็นจำนวนโดยประมาณ เนื่องจากอยู่ระหว่างกระบวนการพัฒนาและปรับปรุงระบบเก็บข้อมูลหัตถการ
แพทย์ประจำบ้าน ภาควิชาสูติศาสตร์-นรีเวชวิทยา

การทำงานวิจัยระดับแพทย์ประจำบ้าน
(Research competency)

เรื่อง Prevalence of Abnormal Menstrual Patterns among Copper Intrauterine Devices
(IUDs)Users in Women Attending Family Planning Clinic, Siriraj Hospital

อาจารย์ผู้ควบคุม ผู้ช่วยศาสตราจารย์นายแพทย์สุรศักดิ์ อังสุวัฒนา

ข้อมูลสำคัญสำหรับงานวิจัย

- ผ่าน SIRB เมื่อ 28 สิงหาคม 2555
เลขที่ 415/2555(EC3)
- ประกวดการนำเสนองานวิจัยในการประชุมราชวิทยาลัยสูตินรีแพทย์แห่งประเทศไทย
วันที่ 26 พฤศจิกายน 2556
 เข้าร่วมนำเสนอ ไม่ได้รับรางวัล
 เข้าร่วมนำเสนอ ได้รับรางวัล
- การตีพิมพ์ในวารสารวิชาการ
 ไม่ได้ตีพิมพ์
 ได้รับการตีพิมพ์ (ระบุรายละเอียดวารสาร)

ผลการประเมินเจตคติและพฤติกรรมการทำงานของแพทย์ประจำบ้าน
(Multisources feedback)

แพทย์ประจำบ้านจะได้รับการประเมินในประเด็นต่อไปนี้

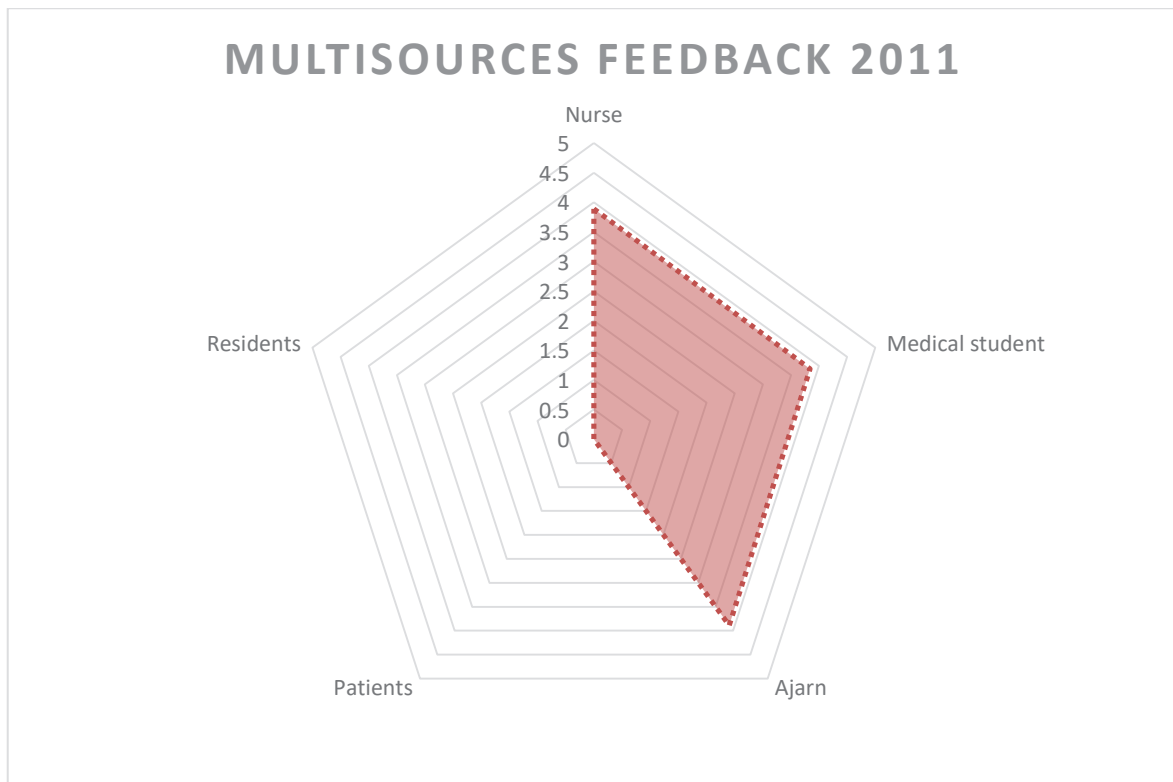
1. ความรู้ความสามารถด้านวิชาการ

2. ทักษะพื้นฐานในการปฏิบัติงาน

ได้แก่ ทักษะการสื่อสารกับเพื่อนร่วมงานและผู้ป่วย/ญาติ การบันทึกรายงานผู้ป่วย การทำงานร่วมกับผู้อื่น และบุคลิกภาพขณะปฏิบัติงาน

3. คุณธรรมและจริยธรรม

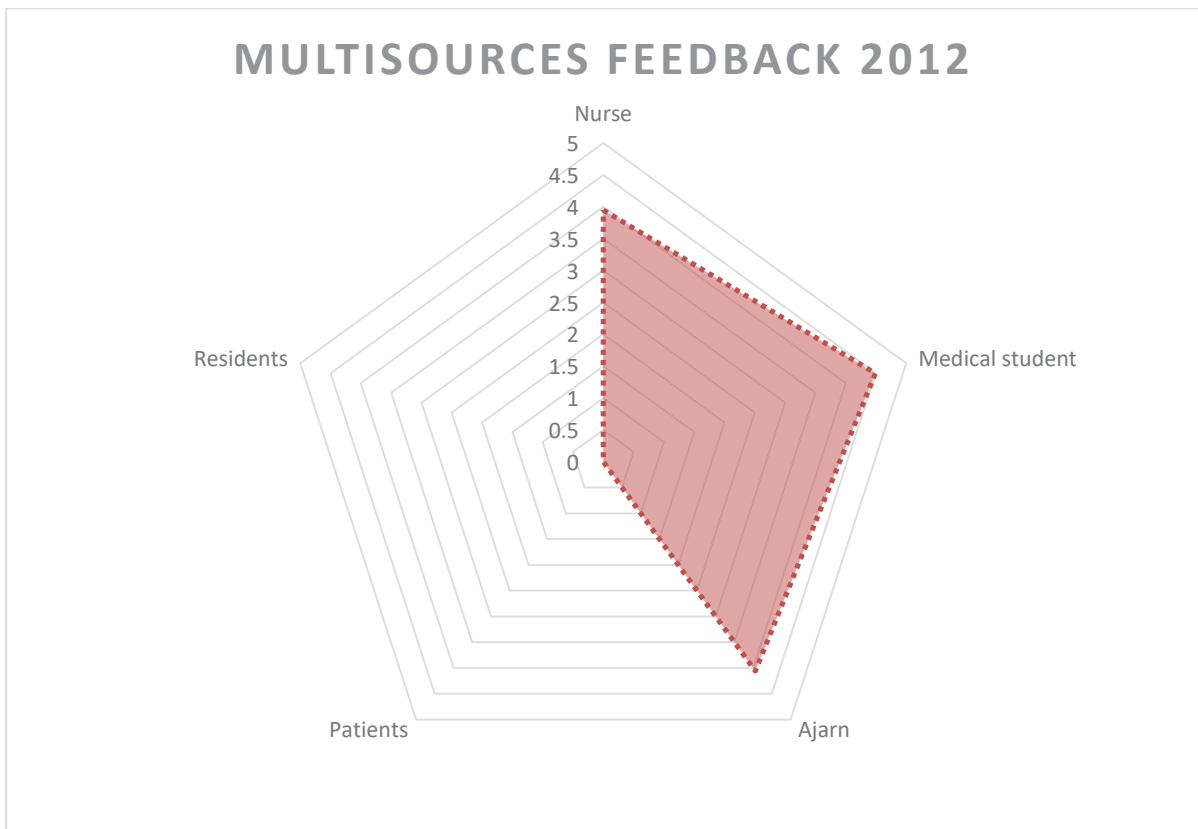
ได้แก่ ความรับผิดชอบ ความเสียสละ ความตรงต่อเวลา ความซื่อสัตย์ การปฏิบัติตามระเบียบข้อบังคับ และอภัย/น้ำใจ/ความเอื้อเฟื้อต่อผู้อื่น



ชั้นปีที่ 1 ปีการศึกษา 2554

Rotation	อาจารย์ (5 คะแนน)	นักศึกษาแพทย์ (5 คะแนน)	พยาบาล (5 คะแนน)	แพทย์ ประจำบ้าน	ผู้รับบริการ
พระศรีฯ 9/2			4.00		
LR เข้า		3.58	4.00		
LR พิเศษเข้า			3.90		
นริเวช 1	3.50	3.40	4.10		
นริเวช 1 (2)	4.00	3.41	3.92		
พระศรีฯ 10/2			3.92		
พระศรีฯ 9/1+ANC			4.03		
LR ดึก			3.76		
LR พิเศษป่วย			3.23		
นริเวช 2	4.20	3.17	5.00		
Onco	3.88	5.00	4.07		
พระศรีฯ 10/3		3.83	2.92		
พระศรีฯ 10/1		4.50	3.84		
คะแนนเฉลี่ย	3.89	3.84	3.89		

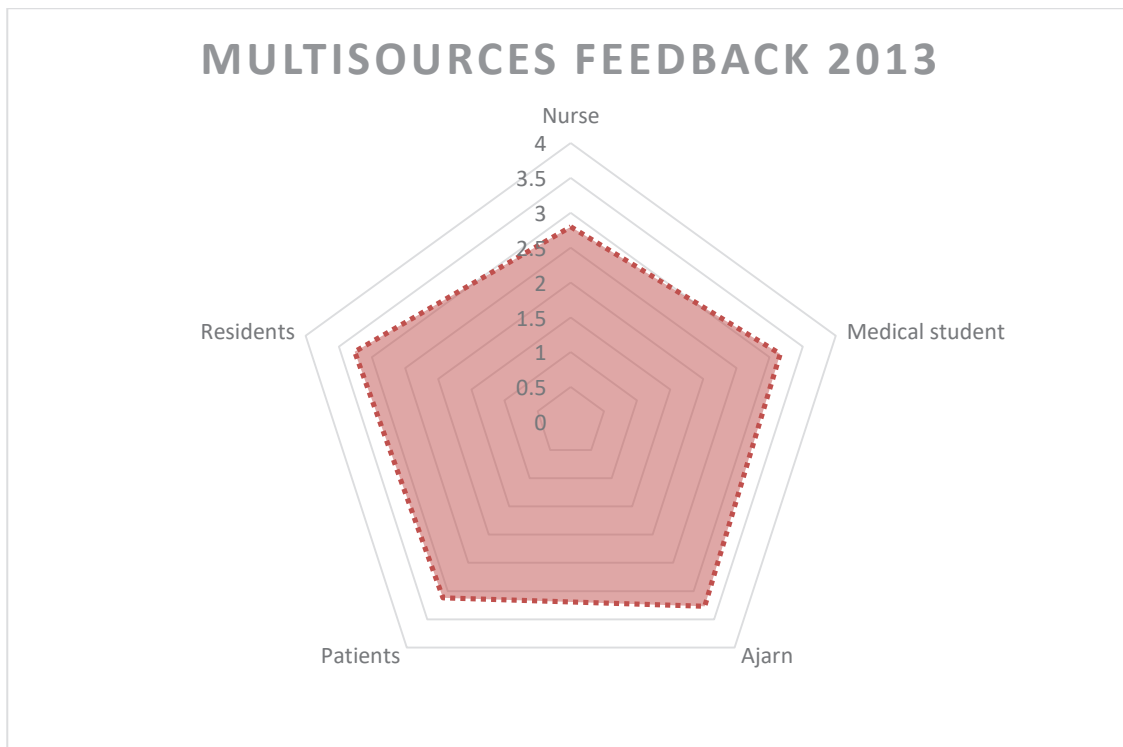
*ยังไม่มีผลการประเมินจากแพทย์ประจำบ้านและผู้รับบริการในปีการศึกษา 2554



ชั้นปีที่ 2 ปีการศึกษา 2555

Rotation	อาจารย์ (5 คะแนน)	นักศึกษาแพทย์ (5 คะแนน)	พยาบาล (5 คะแนน)	แพทย์ ประจำบ้าน	ผู้รับบริการ
นรีเวช 1	3.95	4.67	4.84		
เลือดสิน	4.00				
พระศรีฯ 13/1	4.00		3.69		
LR ดึก			4.07		
Onco	4.05	3.77	3.76		
พระศรีฯ 14/2			4.42		
นรีเวช 2	3.82	4.50	4.23		
พระศรีฯ 13/2	4.35		3.08		
พระศรีฯ 10/1		5.00	3.25		
พระศรีฯ 14/1	4.30		4.00		
LR เช้า		4.58	4.20		
พระศรีฯ 10/3		4.42	4.00		
คะแนนเฉลี่ย	4.06	4.49	3.95		

*ยังไม่มีผลการประเมินจากแพทย์ประจำบ้านและผู้รับบริการในปีการศึกษา 2555



ชั้นปีที่ 3 ปีการศึกษา 2556

Rotation	อาจารย์ (4 คะแนน)	แพทย์ประจำบ้าน (4 คะแนน)	พยาบาล (4 คะแนน)	นักศึกษาแพทย์ (4 คะแนน)	ผู้รับบริการ (4 คะแนน)
นรีเวช 1	3.30	3.40	2.29	3.12	3.28
STD	3.50		3.00		
พระศรีฯ 10/1		3.30	2.20	2.66	2.64
LR พิเศษ			3.06		
OPD GYN			3.40		3.07
Septic		3.50	3.00	3.40	3.33
วิสัญญี	2.70				
นรีเวช 2	3.40	3.48	3.13	3.44	3.40
Infertile	3.30				
นครปฐม	3.30				
OPD ANC			2.85		3.04
ONCO	3.05	2.80	2.21	3.75	
LR เช้า		3.10	2.95	2.63	
Surgery	3.62				
คะแนนเฉลี่ย	3.27	3.26	2.80	3.16	3.12

*เริ่มการประเมินจากนักศึกษาแพทย์และผู้รับบริการ ในปีการศึกษา 2556



Portfolios for Assessment and Learning

Jan van Tartwijk
Erik W Driessen

AMEE GUIDE
Assessment

45






AMEE Guides in Medical Education

www.amee.org

Welcome to AMEE Guides Series 2

The AMEE Guides cover important topics in medical and healthcare professions education and provide information, practical advice and support. We hope that they will also stimulate your thinking and reflection on the topic. The Guides have been logically structured for ease of reading and contain useful take-home messages. Text boxes highlight key points and examples in practice. Each page in the guide provides a column for your own personal annotations, stimulated either by the text itself or the quotations. Sources of further information on the topic are provided in the reference list and bibliography.

Guides are divided into series according to subject:

-  **Teaching and Learning**
-  **Research Methods**
-  **Education Management**
-  **Curriculum Planning**
-  **Assessment**

The Guides are designed for use by individual teachers to inform their practice and can be used to support staff development programmes.

'Living Guides'

An important feature of this new Guide series is the concept of supplements, which will provide a continuing source of information on the topic. Published supplements will be available to all who have purchased the Guide.

If you would like to contribute a supplement based on your own experience, please contact the Guides Series Editor, Professor Trevor Gibbs (tjg.gibbs@gmail.com).

Supplements may comprise either a 'Viewpoint', when you communicate your views and comments on the Guide or the topic more generally, or a 'Practical Application', where you report on implementation of some aspect of the subject of the Guide in your own situation. Submissions for consideration for inclusion as a Guide supplement should be maximum 1,000 words.

Other Guides in the new series

A list of topics in this exciting new series is listed on the back inside cover.

Institution/Corresponding address:

Dr Jan van Tartwijk, ICLON – Leiden University Graduate School of Teaching, Leiden University,
PO Box 905, 2300 AX Leiden, The Netherlands

Tel: +31 71 527 3845

Fax: +31 71 527 5342

Email: jtartwijk@iclon.leidenuniv.nl

The authors:

Dr Jan van Tartwijk works at the ICLON – Leiden University Graduate School of Teaching. In his research and teaching he focuses on teacher-student communication processes in the classroom and the use of portfolios in medical education and teacher education.

Dr Erik Driessen works at the Department of Educational Development and Research at Faculty of Medicine of the University of Maastricht. He specializes in assessment and the use of portfolios in medical education.

Both have a long history with working with portfolios. Jan van Tartwijk started experimenting with portfolios in teacher education and faculty development in 1994. In 1999, he joined Erik Driessen and Cees van der Vleuten at Maastricht University, where they implemented portfolios in the undergraduate program of the Faculty of Medicine of the University of Maastricht. Since then, they have published a series of articles and books about using portfolios in higher education and have advised numerous faculties and originations in medical education and elsewhere about the use of portfolio for learning and assessment. Their corporation is not limited to the topic of portfolios; they also work together on research on how to stimulate and assess self-critical thinking and reflection.

Part of this AMEE Guide was first published in *Medical Teacher*:

Van Tartwijk J & Driessen EW (2009). Portfolios for assessment and learning. *AMEE Guide No.45. Medical Teacher*, 31(9): 790-801.

Guide Series Editor: Trevor Gibbs (tjg.gibbs@gmail.com)

Published by: Association for Medical Education in Europe (AMEE), Dundee, UK

Designed by: Lynn Thomson

© AMEE 2010

ISBN: 978-1-903934-57-9

Contents

Abstract	1
Introduction	2
Portfolio goals, content, and organization	4
Portfolios as a multipurpose instrument	4
Electronic portfolios	7
Portfolios and learning from experience	9
Theoretical background	9
Reflection and professional development	10
Using portfolios as tools for assessment	14
Factors influencing the success of the introduction of a portfolio	21
People	21
Academic leadership	23
Infrastructure	23
Concluding remarks	24
References	25

Abstract

In 1990, Miller wrote that no tools were available for assessment of what a learner does when functioning independently at the clinical workplace (Miller 1990). Since then portfolios have filled this gap and found their way into medical education, not only as tools for assessment of performance in the workplace, but also as tools to stimulate learning from experience.

We give an overview of the content and structure of various types of portfolios, describe the potential of electronic portfolios, present techniques and strategies for using portfolios as tools for stimulating learning and for assessment, and discuss factors that influence the success of the introduction. We conclude that portfolios have a lot of potential but that their introduction also often leads to disappointment, because they require a new perspective on education from mentors and learners and a significant investment of time and energy.

TAKE HOME MESSAGES

- The goals of working with a portfolio need to be clear.
- It is not problematic to use portfolios concurrently to formatively promote learning as well as for summative assessment. Summative assessment is important to ensure that portfolio learning maintains its status alongside other assessed subjects.
- The effectiveness of learning is enhanced when a mentor supports the portfolio process. Mentorship requires a substantial time investment but is crucial for the successful use of portfolios. The effectiveness of assessment can be enhanced by combining the portfolio with an interview.
- Use a flexible learner-centred portfolio format. A rigid structure in which every detail of portfolio content is prescribed will elicit negative reactions from portfolio users.
- Too much structure is a greater risk than too little structure, but learners do need clear directions and guidance to support the development and assessment of broad competencies.
- Working with a portfolio is time consuming both for learners and mentors. This is more of a problem in postgraduate training and continuous medical education than in undergraduate education.

Introduction

Today's doctors find themselves confronted not only with patients who are increasingly knowledgeable and assertive, but also with pressure to apply new findings and evidence in day-to-day practice, and with the necessity to collaborate with other health professionals in ever larger teams and communities. To deal with these complexities, doctors need generic competencies to enhance effective communication, organization, teamwork and professionalism. These generic competencies are sometimes labelled as doctors' "soft skills" in contrast to "hard clinical skills". In recent years, learning, teaching and assessment of these generic competencies has gained unexpected urgency among politicians and the general public. Headlines decrying incidents involving dysfunctional doctors and hospital departments with dramatic impact on morbidity and mortality figures catapulted generic competencies to the forefront of attention as indispensable qualities for doctors. As a result, professional associations and governments began to voice increasingly urgent demands to include these generic competencies in education and assessment (General Medical Council, 2000). At the same time, consistent with the general trend towards outcome-based education, the focus in medical education shifted from the educational process itself towards the competencies of doctors at the end of training and at important junctures during the training process (Norcini et al., 2008). The competencies described by professional organizations such as the Royal College of Physicians and Surgeons of Canada (1996) became the framework for assessment and, as a consequence, for the content and organization of programmes for medical education in many countries.

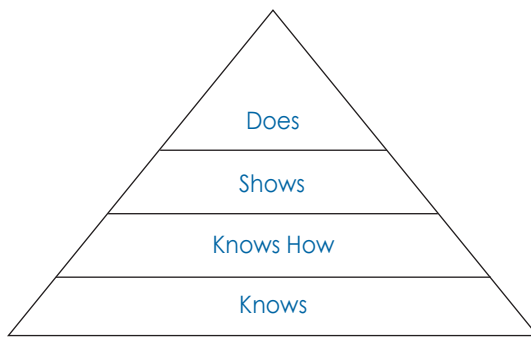
However, stimulating the development of competencies (Box 1) and the assessment of its result is complicated. Already in 1990, Miller described the challenges involved in assessing clinical competence. He presented a framework for clinical assessment, shaped like a pyramid (Figure 1), whose layers from bottom to top represent increasingly complex levels of mastery, with the lower levels providing the foundation for the higher levels (Miller, 1990).

BOX 1 Competence

The concept of competence is much used and much debated (Stoof et al., 2002; Dreyfus, 2004). Here, we define it as an integrated body of knowledge, skills, and (professional) attitudes enabling proficient performance in certain real life settings, i.e. the "Does" level in Miller's framework.

...doctors need generic competencies to enhance effective communication, organization, teamwork and professionalism.

FIGURE 1
Framework for clinical assessment: Miller's Pyramid (cf. Miller, 1990)



The bottom level is concerned with *knowledge*. This is the knowledge relating to the skills that learners must master for their future professional practice. This knowledge is best assessed by written tests. The next level represents application of the knowledge from level 1. Learners should know *how* to apply their knowledge when performing skills. For instance, at this level, learners are expected to know how to diagnose a patient and which aspects of a patient's presentation to attend to. The *knows how* level can also be assessed by written tests. One level up, at level 3, the issue of interest is that learners demonstrate their ability to use their knowledge to *take appropriate action in a simulated environment*. This level combines knowledge and action (cognition and behaviour). Not only should learners know how to diagnose a patient, they should also be able to actually perform the appropriate actions, for example a physical examination in a simulated patient (*shows how*). The top of the pyramid is concerned with *independent performance within the complex environment of day-to-day practice*. This requires integration of knowledge, skills, attitudes, and personal characteristics. Performance at the top of the pyramid is manifested when learners are working independently in professional practice. Typically, adequate performance at this level requires integrated performance of different roles; not only the role of medical expert but also that of counsellor, participant in the doctor- patient relationship, a leadership role in relation to nursing staff, etc. Good performance at the Does level (of Miller's Pyramid) implies competence.

In 1990, Miller observed that there were no instruments to evaluate performance consistent with the top of the pyramid (Miller, 1990). At the same time, scholars in the field of teacher education and teacher assessment were struggling with the same problem (Bird, 1990). Here too, the key challenge was how to assess performance in real life settings. Shulman (1998) describes the Teacher Assessment Project that was set up with the purpose of exploring and developing new approaches to the evaluation of teaching in primary and secondary education. He recounts that it was considered undesirable to assess teacher competence solely on the basis of ratings in assessment centres, because experiments showed that the information provided by assessment centres alone was not enough to identify competent and excellent teachers. Information about whether teachers succeeded in making the most of their pupils' learning opportunities *within* their own complex working environment was needed as well. It was also

Good performance at the Does level (of Miller's Pyramid) implies competence.

recognised that there can be striking variations among teaching settings. For instance, it makes quite a difference whether one teaches at an urban school in a deprived area with its myriad of social problems or at a high school in a middle class suburban environment. As part of efforts to achieve fair judgement of teacher performance in a broad array of settings and situations, the *portfolio* concept was borrowed from the arts and architecture (Box 2).

BOX 2 Portfolio

Portfolios that are used in education contain evidence of how learners fulfil tasks and their competence is progressing. They may be digital or paper based and content may be prescribed or left to the learners' discretion. Despite variations in content and format, portfolios basically report on work done, feedback received, progress made, and plans for improving competence (Driessen et al., 2007b).

Since portfolios were introduced in medical education in the early 1990s (Royal College of General Practitioners, 1993), their use as an instrument for both assessment and encouraging professional growth has increased enormously (Snadden et al., 1999; Friedman Ben David et al., 2001). However, the evidence to date suggests that the introduction of portfolios for these purposes has met with mixed success (Driessen et al., 2007b; Tochel, et al., 2009, Buckley et al., 2009). Although potentially powerful instruments in education, the use of portfolios has proved to be vulnerable.

The aim of this AMEE Guide is to help medical teachers and educators to make full use of the possibilities that portfolios offer and prevent difficulties occurring. Based on an analysis of what portfolios help achieve, it is our purpose to provide practical clues about the design, implementation and use of portfolios in medical education.

Firstly, we will describe how portfolio content and structure relate to the various goals that they are designed to achieve. Next, we will focus on the use of portfolios as instruments that can encourage professional growth by stimulating learning from experience and subsequently, we will elaborate on the use of portfolios as instruments for assessment. Each of these goals requires specific content and organization of portfolios. Finally, we will focus on the factors that are important for the successful introduction of portfolios in (medical) education.

Portfolio goals, content, and organization

Portfolios as a multipurpose instrument

- **Portfolios for assessment:** When portfolios were originally introduced in education as instruments for authentic assessment, they closely resembled the portfolios of architects and artists that Lyons (1998) describes as a portable case for keeping, usually without folding, loose sheets of papers, drawings or photographs. Building on the principle of triangulation (Denzin, 1978; Denzin & Lincoln, 2000) all kinds of evidence can be brought

together in those portfolios that, in combination, give the possibility to draw valid conclusions about competence (Box 3).

BOX 3

Combining evidence to improve the quality of conclusions

In the literature, combining data from various sources with the aim to improve the quality of conclusions is often referred to as triangulation. The aim of triangulation is to avoid biases and problems, such as those related to the reliability and trustworthiness of data that are derived from one single source.

Procedures for multisource feedback or 360-degree feedback use a similar strategy by stimulating learners to gather feedback from different sources. Lockyer & Clyman (2008) describe a procedure involving a questionnaire survey among medical colleagues, nurses, and patients and their families to collect data about learners' specific competencies. The same questionnaire is completed by the learners themselves. By aggregating these data, reliability is improved.

However, in one of the first explorations of portfolios for teacher assessment, Bird (1990) wrote that the portfolio procedures for assessment might easily degenerate into exercises in amassing paper. He suggested that the evidence in a portfolio should be organised according to the competencies that the person compiling the portfolio wants to show. Both for the learner compiling the portfolio and for an assessor this would be helpful. Instructions starting with "Show how you..." might clarify for portfolio owners that they are asked to provide specific evidence about their performance. A portfolio organised by tasks or competencies might be helpful for assessors, because it indicates what the material in the portfolio is supposed to show. Based on initial experiments with portfolios, Collins (1991) suggested that captions should be attached to the evidence in the portfolio:

One essential component of the portfolio was the document caption. The caption is a little sheet attached to each document stating what the document is (...) and why it is valuable evidence. (...) Captions proved to be essential to the portfolio development process. Documents without captions were meaningless to the raters. (p. 153)

- **Portfolios for learning:** Soon after the introduction of portfolios in medical education, Snadden & Thomas introduced the term "portfolio learning" (Snadden & Thomas, 1998b):

Portfolio learning is a method of encouraging adult and reflective learning for professionals. Derived from the graphic arts it is based on developing a collection of evidence that learning has taken place (p. 192)

They emphasise the importance the importance of supervision and critical reflection for portfolio learning:

The system works well when it operates through the interaction of a learner and mentor using the material as a catalyst to guide further learning. It is essential that the portfolio does not become a mere collection of events seen or experienced, but contains critical reflections on these and the learning that has been made from them (p.192).

...portfolio procedures for assessment might easily degenerate into exercises in amassing paper.

Portfolio learning is a method of encouraging adult and reflective learning for professionals. Derived from the graphic arts it is based on developing a collection of evidence that learning has taken place.

A portfolio can also stimulate reflection, because collecting and selecting work samples, evaluations and other types of materials that are illustrative of the work done, compels learners to look back on what they have done and analyse what they have and have not yet accomplished.

In many cases, portfolios are assembled over a longer period of time. That is why they can also be used to support planning and monitoring in professional development. One way to do so is to include learning objectives in the portfolio as well as a document trail of related learning activities and accomplishments (Mathers et al. 1999; Oermann, 2002).

As a consequence, reflections and overviews of personal development have secured a prominent place in many portfolios. Portfolios that are primarily geared to assessment will remain organised around all kinds of materials that provide 'evidence' of competencies. In portfolios that are primarily used to monitor and plan learners' development, overviews will take centre stage. Portfolios whose primary objective is to foster learning by stimulating learners to reflect on and discuss their development will be organised around learners' reflections.

- **A multipurpose instrument¹:** Inevitably, these developments have widened the applicability of the label *portfolio* to a broad range of instruments. Some portfolios might equally and aptly be labelled *Personal Development Plan* or *Reflective Essay*. Because of the tremendous variety in portfolios, careful and critical appraisal of the strengths and weaknesses of different portfolios is advisable before deciding which one to implement in a particular setting.

The question to be answered is whether a certain portfolio is fit for its intended purpose. And just as someone else's shoes are unlikely to fit comfortably, portfolios tailored to one particular educational setting may not fit into the educational configuration(s) of other settings (Spandel, 1997). An ill-fitting portfolio will inevitably be discarded sooner or later. To assist in determining whether a portfolio is appropriate for its intended purpose the triangle in Figure 2 helps to define the nature of a portfolio. It does so by inviting positioning of a portfolio in the area of the triangle where it is most likely to achieve its intended principal objectives.

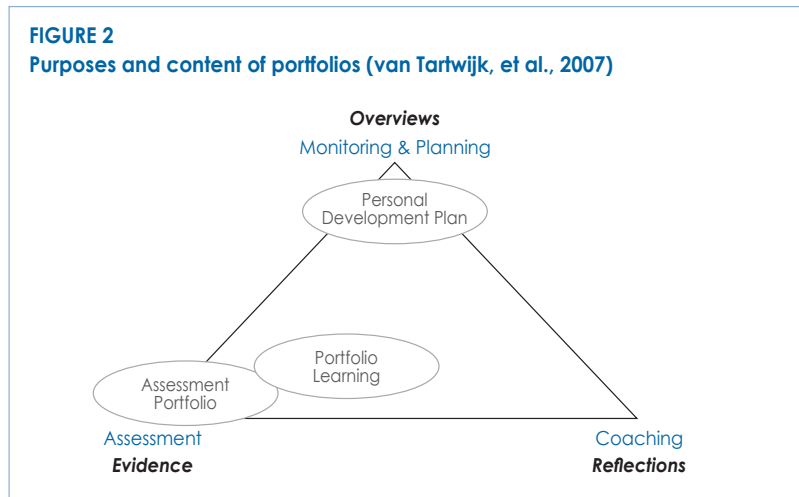
Obviously, a portfolio can be used to achieve more than one goal. When a portfolio is to serve a combination of goals, its position in the triangle will shift towards the centre because its strengths have to be distributed more evenly over evidence, overviews and reflections. In practice, the majority of portfolios are not situated in one of the corners of the triangle (Buckley et al., in press). A controversial issue in the literature on educational portfolios is whether it is acceptable to have one portfolio for both assessment and reflection (Snyder et al. 1998). An argument against this dual function is that assessment may jeopardise the quality of reflection thereby detracting from the portfolio's effectiveness for mentoring purposes. Learners may be reluctant to expose their less successful efforts at specific tasks and to reflect on strategies for addressing weaknesses if

A portfolio can also stimulate reflection...

¹ Parts of this section were published in the journal *Quality in Higher Education* (van Tartwijk, et al., 2007)

they believe they are at risk of having 'failures' turned against them in an assessment situation. Portfolios that are not assessed, on the other hand, do not "reward" learners for the time and energy they invest in them. As a result, learners are likely to take the portfolio and any associated learning activities less seriously. A recent BEME review showed that most portfolios were also assessed for summative purposes (Buckley et al., 2009).

FIGURE 2
Purposes and content of portfolios (van Tartwijk, et al., 2007)



An effective portfolio has a clear but flexible structure, giving individual learners opportunities to describe their own unique development (Pearson & Heywood, 2004; Driessen et al. 2005b; Grant et al. 2007). Clear instructions are important, but when the content of a portfolio is prescribed in detail, portfolios are often experienced as highly bureaucratic instruments (Davis et al., 2001; O'Sullivan et al. 2004; Pearson & Heywood, 2004; Kjaer et al. 2006). Portfolios meet with stronger appreciation when learners have a certain amount of freedom to determine the content of their own portfolios (Snadden & Thomas, 1998a; Driessen et al., 2005b).

An effective portfolio has a clear but flexible structure, giving individual learners opportunities to describe their own unique development.

Electronic portfolios

A growing number of medical schools use electronic portfolios (e-portfolios) instead of paper-based portfolios (Fung Kee Fung et al., 2000; Lawson et al., 2004; Woodward & Nanlohy, 2004; van Tartwijk et al., 2007; Driessen et al. 2007a). This preference is based on a number of considerations:

- In e-portfolios, hyperlinks can be inserted to make connections between evidence, overviews, and reflections. This can be useful, for instance, when learners want to illustrate reflections with evidence that is stored somewhere else in the portfolio, or want to illustrate a schematic overview of their development by making hyperlinks to materials and reflections. Hyperlinks can also be useful to make a table of contents of the portfolio. For instance by including a list of captions in the portfolio and making hyperlinks to related materials. Mentors or assessors can browse through this list of captions, obtain a quick overview of all the evidence in the portfolio, and just click on the evidence that is relevant to their specific purpose.

- A paper-based portfolio can be cumbersome because of its bulk. Imagine an assessor who needs to take 15 paper portfolios home! Furthermore, there is generally only one copy of a paper portfolio. Whenever learners hand their paper portfolios to their mentor or assessor, the portfolio is literally out of their hands. Not only do they run the risk of the portfolio getting lost, it is also more difficult for them to prepare to discuss the portfolio with their mentor or assessor. Another advantage of e-portfolios is that they are easier to keep up to date.

Of course there are disadvantages as well:

- Mentors who do not like to read a portfolio on screen will still have to print it. In most systems it is not possible to make notes on the portfolio itself (although making notes on the learner's paper portfolio might not be desirable as well).
- E-portfolios can only be used by learners and teachers who are sufficiently skilled in using the relevant software and hardware.
- An e-portfolio requires a stable and high quality information technology infrastructure that is not always available.

Nowadays, many dedicated portfolio systems are available, which are usually user-friendly (Dornan et al., 2002; www.eportfolioservice.nl). These systems can provide specific functionalities for specific portfolio goals: options to include work-based assessment instruments, such as multisource feedback or mini clinical evaluation exercises (mini-CEX) in portfolios for clinical training; to invite specific individuals to inspect the portfolio, either wholly or in part, while denying access to everyone else.

Apart from dedicated systems, learners can produce an e-portfolio using standard word-processors or HTML editors, preferably ones that they and their teachers are familiar with (Gibson & Barrett, 2003). The cost of dedicated portfolio software is not the only reason to support this choice: for many purposes the hyperlink functionality of generic software is all that learners need. Furthermore, generic software allows a learner to impart his or her own flavour to the portfolio. This can enhance the learners' motivation to work with the instrument. Another reason is that many portfolio systems are limited because they are built to accommodate no more than one or two portfolio types. Finally, portfolios built with dedicated software need to be accessible with generic software for later maintenance and presentation. This may well be the case after a learner has left the setting in which the portfolio was produced, or in the event that the vendor in question ceases to do business. In summary, standard software tools have disadvantages from the perspective of managing access to the portfolio using the internet or to include work-based assessment instruments, but they usually provide all the options learners need to produce a portfolio that works well and looks great.

In a study comparing web-based and paper-based portfolios (Driessen et al., 2007a), not only did learners add more personal touches to content and form and invested more time in their portfolios, but mentors were unanimous in their appreciation of the greater ease of use of web-based portfolios compared to the more familiar paper-based ones. Information was

...standard software tools have disadvantages from the perspective of managing access to the portfolio using the internet or to include work-based assessment instruments, but they usually provide all the options learners need to produce a portfolio that works well and looks great.

easy to locate without having to turn pages to find certain content and the portfolios could be accessed from different locations were two reasons cited for preferring web-based portfolios. Other authors have also reported on the user friendliness of electronic portfolios (Fung Kee Fung et al., 2000; Lawson et al., 2004). In these studies, tutors appreciated the easy electronic access and reduction in the amount of paper used. However, the same authors also reported certain situations that make web-based portfolios less user-friendly than paper-based portfolios. For instance, limited computer access in the clinical workplace cancels out the advantages of user-friendliness and may even have an opposite effect.

Portfolios and learning from experience

Research shows that the role of the mentor is crucial to the successful use of portfolios aimed at learning from experience (Finlay et al. 1998; Snadden & Thomas, 1998a Mathers et al., 1999; Pearson & Heywood, 2004; Driessen et al., 2005b; Grant et al., 2007). In this section, we focus on the strategies mentors can use to promote learning from experience with a portfolio.

Theoretical background

The contemporary view of learning, based on constructivism, is that people "construct" new knowledge and understanding based on what they already know and believe (Bransford et al. 2000). What people know and believe can be represented as cognitive structures that guide their perception of reality. Evidently, a perception of reality based on individual cognitive structures does not afford an objective view of reality, but, by definition, an individual, idiosyncratic view. It is this personal perception of reality that guides a person's actions.

Reflection is an important concept in this framework, which relates to changing cognitive structures. Research has shown that meta-cognitive skills, such as reflection, increase the degree to which learners transfer what they have learned to new settings and events (Bransford et al., 2000). Despite considerable confusion about the precise definition of the term reflection (Hatton & Smith, 1995; Mann et al. 2007) all authors writing about reflection share the constructivist view that human behaviour is guided by mental structures that are not static but flexible, evolving, and changing in response to experiences. Based on this consensus view, reflection can be defined as the mental process of organising or reorganising cognitive structures that represent existing knowledge and beliefs and guide perceptions of experiences, situations, and problems (Korthagen et al. 2001). To put it in simpler terms: reflection means exploring and elaborating one's *understanding* of an experience (Eva & Regehr, 2008). Building on Van Manen's work (1977), Hatton & Smith (1995) distinguish three types or levels of reflection. The first type is concerned with the *means* to achieve certain ends. The second type is not only about means, but also about *goals*, the *assumptions* upon which they are based, and the actual *outcomes*. The third type of reflection is referred to as *critical reflection*. Here, moral and ethical criteria are also taken into consideration. Judgements are made about whether professional activity is equitable, just, and respectful to persons or not. Hatton and Smith emphasise that these three types of reflection should

Research shows that the role of the mentor is crucial to the successful use of portfolios aimed at learning from experience.

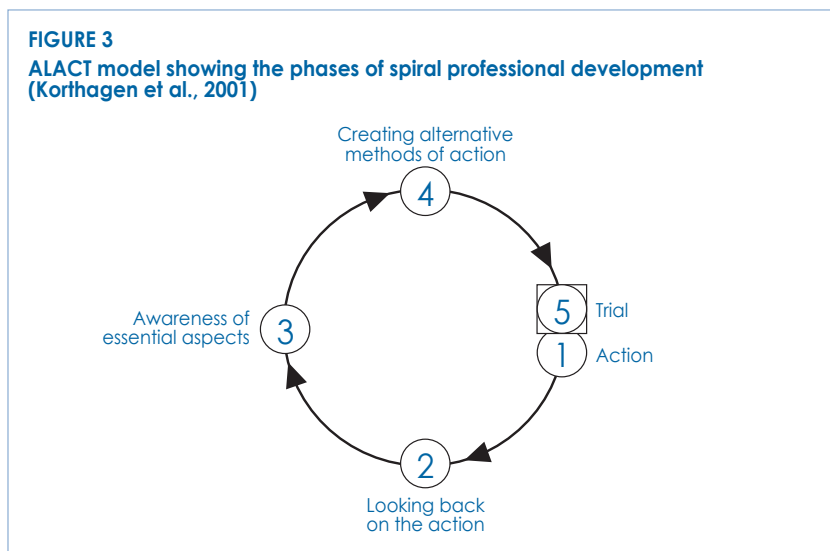
...meta-cognitive skills, such as reflection, increase the degree to which learners transfer what they have learned to new settings and events.

not be viewed as hierarchical. Different (educational) contexts and situations may lend themselves more to one kind of reflection than to another.

Reflection and professional development

For medical teachers who want to help learners learn from practice, the key question to answer is: "How can I stimulate my learners to reflect on their experiences and learn from them?" For this AMEE Guide the additional question is: "... and how can a portfolio help to improve the quality of reflection?".

Korthagen designed the **ALACT** model (**A**ction, **L**ooking back, **A**wareness, **C**reating alternative methods, **T**rial) (Figure 3) to describe the spiralling process that effective learners go through when faced with a situation for which no routine solution is available (Korthagen et al., 2001). This model resembles the three step model described by Snadden & Thomas (1998b) which focused on evaluation, reflection, and formulating a learning plan. We will describe the ALACT model, explain the potential contribution of working with a portfolio in each of the stages, and give suggestions for coaching strategies (Driessen et al., 2008).



ALACT

A Action: The cycle starts with action undertaken for a specific purpose (e.g. for developing a specific competence). Learners can be helped to improve their existing routines and concurrently acquire new ones by pre-selecting experiences from which they can learn, for example a mixture of patients who are more or less easy to diagnose. Ericsson's research predicts that expertise will grow not just from the weight of experience but also from engaging in activities specifically designed or selected to improve performance (Ericsson, 2006).

Learners can be helped to improve their existing routines and concurrently acquire new ones by pre-selecting experiences from which they can learn.

L Looking back on action: self-directed assessment seeking: The ALACT cycle then moves to the stage where learners look back on a previous action, usually when that action was not successful or something unexpected happened. This looking back on action is assumed to be accompanied by an evaluation of whether the goals were realised and the learner's part in this. In many cases this can be regarded as a form of *self assessment*. Eva & Regehr (2008) write that most of the time self-assessment is conceptualised according to a "guess your grade" model of which the quality is generally poor (Davis et al., 2006). As an alternative they propose *self-directed assessment seeking*, which they describe as a process by which a learner takes personal responsibility for looking outward, explicitly seeking feedback and information from external sources of assessment data, to direct performance improvements that can help them to validate their self-assessment.

The role of the portfolio: Seeking and selecting evidence (documents, feedback, work-based assessments, etc.) for inclusion in a portfolio can be regarded as self-directed assessment seeking. To improve the quality of this process, it is important to use a variety of evidence from various sources. The validity of the results of self-directed assessment seeking will be maximised if the learner's self-reflections are consistent with all the information that is brought together in a portfolio.

Teaching strategies: Research has shown that a mentor can play a decisive role in determining whether the use of portfolios in education is successful or not (e.g. Driessen et al., 2007b). At the very least, learners may expect their mentors to pay serious attention to their portfolios, for after all they did spend a lot of time and energy to put their portfolio together. But even more importantly, careful scrutiny of their own performance may be confronting for learners. Effective mentors have an important role in this respect. In Box 4, we give suggestions for a number of strategies to be used by medical teachers in this phase, derived from the work by Korthagen and colleagues (Korthagen et al. 2002).

A Awareness of essential aspects: reflection: After conclusions have been drawn about the quality of performance and the characteristics of the situation, the next step in the ALACT model is to foster awareness of essential aspects. In this phase, learners try to develop a new and better understanding of what has happened, i.e. they reflect on their performance.

They can focus on the *means* they used to achieve a goal and try to understand why their strategy was successful or not. They can also consider whether they had selected a suitable *goal* for this particular situation. And finally they may consider what they want to achieve from a *moral* or *ethical* perspective.

Seeking and selecting evidence (documents, feedback, work-based assessments, etc.) for inclusion in a portfolio can be regarded as self-directed assessment seeking.

BOX 4**Strategies to stimulate self-directed assessment seeking**

- Provide a safe environment by distinguishing between learners as individuals and their performance.
- Focus on description.
- Stimulate learners to be concrete in their reports. When learners give general evaluations about a situation and their performance, ask questions:
 - What went well?
 - What went wrong?
 - How did you solve this?
 - What effect did this have?
- Stimulate learners to carefully scrutinise all the information in their portfolio. Learners could be asked to go through all the available evidence and answer questions:
 - Which information in your portfolio supports your answers/evaluation?
 - Which information in your portfolio contradicts your answers/evaluation?
- Stimulate learners to take the perspective of other stakeholders. Ask questions:
 - What did you want? What do you think the patient/your colleague/the nurse wanted?
 - What did you think? What did the others think?
 - What did you do? What did the others do?
 - What emotions did you experience? What emotions did the other people involved experience?

The role of the portfolio: Language is important in supporting thinking. Writing things down can help to stimulate reflection (Korthagen et al., 2001). Written reflections were not a part of the original portfolios, like the ones in which artists presented a selection from their work, but almost immediately after the introduction of portfolios in education, written reflections became a fixture of portfolios (Paulson et al. 1991). Embedding a written reflection in a portfolio has the advantage that it can be built on the self-assessment that was validated by the evidence in the portfolio. This is a form of facilitated reflection (Conlon, 2003). The learner can also use the evidence to illustrate a reflection with a concrete example.

Teaching strategies: To stimulate learners to reflect and learn from their experiences, mentors do not need to have all the right answers. The most important thing for them is to ask the right questions. In Box 5 we give a number of examples of questions that mentors can ask.

Language is important in supporting thinking. Writing things down can help to stimulate reflection.

To stimulate learners to reflect and learn from their experiences, mentors do not need to have all the right answers. The most important thing for them is to ask the right questions.

BOX 5

Questions to stimulate reflection

Means

- Which strategies did you consider? Why did you select this strategy? Which are the advantages and disadvantages of the strategy you used?
- Which part of your strategy was effective and which part was not effective? Why was it effective / not effective?
- Would this strategy have been more /less effective in a different situation?

Goals, assumptions, outcomes

- What did you want to achieve? Were you successful? What do you consider successful?
- Why is this particular goal important?/Why did you pursue this goal?

Critical reflection

- Do you think patients / patients' families / medical colleagues / nurses / administrators are satisfied with these outcomes? What are their primary interests?

Confront with discrepancies

- I read in your portfolio that you are happy with the result, but when we talk about it, your face tells a different story.
- You write here that this is what you want to achieve, but you are pleased with your results even though they do not match your goals.
- You do not actually do what you say you want to do.

Generalize across experiences

- Which differences and similarities do you recognise between what is happening now and what happened in situations that you described in your portfolio?
- When do these things happen?
- Do you recognise a pattern?

C Creating or identifying alternative methods of action: change: Analysing previous actions may trigger a search for alternative strategies, or abandonment of original goals. It is important to explicate (new) goals and alternative strategies. A recent review showed that goal setting stimulates learning and that a mentor has an important role to play in this respect (Shute, 2008). Learners who work with a mentor set more specific goals and improve more than those who do not work with a mentor (Smither et al. 2003). Very often, agreement about what should be done differently and which goals should be achieved are written down in a document that is referred to as a Personal Development Plan (PDP).

The role of the portfolio: In many portfolios, the central goal is to keep track of the learner's development. In these portfolios, PDPs can have an important place. Snadden & Thomas for instance, (Snadden & Thomas 1998b) propose that when a portfolio is used for professional development and to track progress, it is important to attach to the portfolio some kind of learning plan.

Teaching Strategies: Both mentors and learners should commit to the agreements in the PDP and it should be on the agenda of their next progress meeting. The plans in the PDP are often too vague. It is important that mentors stimulate learners to be very concrete. It can be helpful to keep in mind that the learning goals in the plan should be formulated in a SMART way (Box 6).

Learners who work with a mentor set more specific goals and improve more than those who do not work with a mentor.

**BOX 6
SMART**

Specific	(Straightforward, not ambiguous)
Measurable	(It is clear under which conditions the goals are achieved)
Acceptable	(The goals should be acceptable to all stakeholders)
Realistic	(The learner should be able to achieve the goals)
Time-bound	(It should be clear when the goal is to be achieved)

T Trial: The last step in the ALACT cycle is trial. This is also the start of a new cycle in the spiral of professional development in this model.

Using portfolios as tools for assessment

In the introduction, we quoted Shulman (1998), who wrote that the reason for introducing portfolios in education as tools for assessment is that in a portfolio information can be brought together about how a person performs and how his or her competencies develop in his or her own complex working environment. From the perspective of assessment, the strength of the portfolio is also its weakness. The evidence held by a portfolio is often not standardised and its meaning often depends on the context from which it originates.

Assessing non-standardised portfolios requires a different perspective on assessment than the traditional quantitative perspective that is best suited for analysing quantitative test scores or results from standardised observations. Authors like Snadden (1999) and Webb (2003) all come to the conclusion that we should not try to fit non-standardised portfolios to standardised psychometric assessment criteria. They point out that portfolio assessment is primarily concerned with interpreting various forms of qualitative information and suggest that assessment procedures should be developed that are based on methods used in qualitative research.

In the next section, we will translate the insights of this literature into recommendations for portfolio assessment. We will structure this section according to five questions that, according to Harden (1979), should always be asked and answered by medical teachers in relation to assessment:

- What is assessed?
- Why is this assessed?
- How is this assessed?
- Who assesses?
- When is this assessed?

What? Although portfolios are also used in undergraduate medical education to assess reflective ability or communication skills (Driessen et al. 2003), portfolios are particularly suited to work-based assessment. In other words, they have added value at the does level of Miller's pyramid (Miller 1990).

The evidence held by a portfolio is often not standardised and its meaning often depends on the context from which it originates.

Many medical curricula are based on competency criteria developed by organisations such as the General Medical Council (GMC), the American Council of Graduate Medical Education (ACGME), and the Royal College of Physicians and Surgeons of Canada (RCPSC). More often than not, additional detail is required to fit the competency criteria to assessment procedures. In aligning competency descriptions with assessment procedures it is of the essence to strike the right balance between very concrete but also very detailed and long lists of “is able to” statements, on the one hand, and very global descriptions providing an overview but too little to support assessment, on the other hand. The extremes of this continuum have been referred to as an analytical versus a global approach. Both approaches have their pros and cons (Box 7).

BOX 7

Analytical versus global assessment

In an analytical assessment, various aspects of a competency are assessed separately. A formula is used to combine the partial assessments into one final score.

Because the criteria are explicitly defined and each partial competence is explicitly assessed, the result is very transparent and usually more reliable and more informative for the learner. Criteria are usually defined in terms of: “The candidate is able to...”.

Problems that may occur are:

- Learners may adapt their learning activities to ‘ticking’ specified criteria. This may result in unnatural activities in the workplace where competencies are acquired.
- Analytical assessment is very labour intensive. It may be experienced as bureaucratic.
- It can be difficult for assessors to give a truly distinct assessment of each individual criterion (‘halo effect’).
- Assessors have limited freedom to take account of specific competencies or extremely good (or poor) performance: if it is not in the criteria, it is not assessed. The assessor may feel curtailed in his/her freedom by the criteria.

In a global assessment, the assessors study the entire portfolio and give an assessment based on their overall impression. A global assessment is far less labour intensive than an analytical assessment. It also enables assessors to take account of learners’ special qualities.

Disadvantages are:

- It is less clear to learners on which criteria the assessment is based. The assessment may also be less reliable. As a result the assessment will be less acceptable to learners.
- Some assessors will feel less certain about their judgement. As a result they will study the material over and over again, which will take even more time than an analytical assessment.
- This type of assessment is relatively vulnerable to assessor preferences and sequence effects (the contrast with the previous candidate may influence the assessment).

A way to combine the best of both approaches is to use scoring rubrics. A *scoring rubric* is a global performance descriptor that lists the criteria for a competency and articulates a limited number of gradations of quality for each criterion. Gradations can be unsatisfactory, sufficient, good, and excellent. Scoring rubrics can be presented as tables, with the criteria in the rows and the grades in the columns. In each cell of this table, performance at that particular level of competence is described. Box 8 provides an example.

BOX 8
Rubrics used for the assessment for final year medical students (source Maastricht University)

	BELOW EXPECTATION	AS EXPECTED	ABOVE EXPECTATION
Clinical performance	Slow in taking a history and performing a physical examination. Considers irrelevant aspects. Slow in making a diagnosis. Misses important conclusions. Frequently unable to formulate management plan and needs considerable guidance.	Adequate speed in taking a history and performing a physical examination. Relevant aspects are considered. Adequate speed in making a diagnosis. Diagnosis contains important conclusions. Formulates an adequate management plan for simple clinical presentations. Needs some guidance. Achieves these goals in the second half of the internship.	Conducts an adequate and efficient history and physical examination. Arrives at an accurate diagnosis within adequate time. Formulates an adequate management plan for simple clinical presentations. Needs little guidance. Has achieved these goals at the start of the internship.
Professionalism (for instance as judged by 360 degree feedback)	Does not keep commitments. Occasionally fails to ask for supervision when this is necessary. Reacts defensively to feedback. Is unable to cope with stress Does not pay attention to his/her personal appearance. Frequently shows awkward behaviour or behaves disrespectfully.	Keeps commitments. Asks for supervision when this is necessary. Needs help in reflecting and considering alternatives and responds adequately to feedback. Occasionally needs help in coping with stress. Appropriate personal appearance; behaves respectfully.	Keeps commitments. Asks for supervision when this is necessary. Is able to reflect critically; responds adequately to feedback and is prepared to acknowledge errors. Is able to cope with stress adequately. Looks well cared for and behaves respectfully.
Has critically assessed his/her performance and formulated appropriate learning goals. This is evidenced by an adequate analysis of strengths and weaknesses and the development plan.	Incomplete, limited or one-sided description of strengths and weaknesses in performance (e.g. only strengths or only weaknesses, limited to one competency). No explanations only lists of facts or situations. No learning goals, learning goals do not match the analysis or are not specific.	A fair number of strengths and weaknesses are not explained or explanations are limited to external attributions (for instance mini-CEX at the wrong moment) Some of the learning goals are not specified.	Above expectation (authentic, recognizable, and well explained). A good analysis of strengths and weaknesses. Also internal attributions and references to evidence in the portfolio. Logical, detailed (based on the analysis) and attainable learning goals.

For learners and their mentors, scoring rubrics can be a roadmap for competence development. It can help them diagnose a learner's current level of competence and point the way to further development. Assessors should not use scoring rubrics as a checklist,

but as a list of arguments to underpin their assessment when they explain it to learners. Learners can also use scoring rubrics to organise their portfolio. They can organise the evidence in their portfolio in chapters corresponding to the different competencies to be assessed and use captions to explain what the evidence shows about a specific competency.

Why? Assessing competencies can be done for three reasons: selection, diagnosis, and certification.

Selection: Determining whether a person is suitable for a certain position. Assessments for selection purposes can take place before entering an educational programme, but also, for instance, before starting a new job.

Diagnosis: In the course of an education programme, the development of learners' competencies is assessed. The purpose of this type of assessment is to give feedback to learners and help them identify new learning goals. Sometimes, this assessment is also used to determine whether or not a learner is allowed to continue with a programme.

Certification: The goal of assessment at the end of an educational or training programme is to establish whether learners have attained the competencies required for graduation or certification. Obviously, the quality of any assessment is important. Poor quality of assessment for selection purposes, for instance, can harm the interests of prospective learners and waste talent. Similarly, poor quality of diagnostic assessment can cause frustration and delay in learners' development. Nevertheless, with graduation and certification decisions the quality of assessment is crucial. Learners who pass but should have failed will become (or continue to be) certified doctors and may become a risk to the community!

How? The quality of the assessment of competencies is crucially determined by the procedure that is used. In the introduction to this section about portfolio assessment, we wrote that the standard psychometric procedures that are used to determine the quality of tests and standardised observations are not very well suited to portfolios with their non-standardised content. In medical education, Webb and colleagues (2003) pointed out that portfolio assessment is primarily concerned with qualitative information and they introduced the idea to use routines developed for qualitative research. Guba & Lincoln's (1989) strategies to achieve *credibility* and *dependability* of assessment can be translated to portfolio assessment (Webb et al., 2003; Tigelaar et al. 2005). In Box 9, we discuss how these strategies can be used.

...the standard psychometric procedures that are used to determine the quality of tests and standardised observations are not very well suited to portfolios with their non-standardised content.

BOX 9**Strategies for portfolio assessment derived from the methodology of qualitative research**

- Incorporate feedback cycles into the mentoring process that accompanies the portfolio to ensure that the mentor's final recommendation does not come as a(n) - unpleasant - surprise to the learner; this approach relates to the credibility strategies of prolonged engagement and member checking.
- Maintain a careful balance between the roles of the mentor as coach and assessor. The aim is to ensure that the person who knows the learner best provides the most relevant information while minimizing any damaging effect on the mentor-learner relationship by using an assessment committee to assess the portfolio; this approach relates to the credibility strategy of prolonged engagement.
- Involve the learner in the decision process to ensure commitment on the part of the learner and allow the learner to communicate a different point of view to that of the mentor; this approach relates to the credibility strategy of member checking.
- Use a sequential judgement procedure in which conflicting information necessitates more information gathering. This ensures the efficient use of resources by limiting the use of additional resources to cases where this is necessary to achieve reliable judgement. This approach relates to the credibility strategy of triangulation.
- Document the different steps of the assessment process. For example a formal assessment plan approved by the Examination Board; portfolio and assessment guidelines; overviews of the results per phase, and written assessment forms per learner. This approach relates to the dependability strategy of audit trail.

The major problem with qualitative research methods as well as with portfolio assessment is the required substantial time investment. At Maastricht University, we developed a portfolio assessment procedure that uses many of these strategies while at the same time aiming for optimal efficiency (Driessen et al., 2005a). This procedure is described in Box 10.

Who? A problem that is much debated in the portfolio literature is the feasibility and acceptability of combining the roles of mentor and assessor into one person. Tigelaar et al. interviewed nine portfolio experts about their views on the use of portfolios in education (Tigelaar et al. 2004). While some of the experts agreed that the mentor is the most appropriate person to advise an assessment committee about a candidate, others argued that it is unethical for mentors to undertake the assessor role. The latter group argued that candidates must feel free to reflect on their professional development together with their mentors, knowing that the mentor will not pass any information on to others. For this reason, the majority of the experts were of the opinion that mentors should not be involved in summative assessment nor make recommendations to an assessment committee. However, there was a minority who agreed with Snyder and colleagues, who wrote that: *"The tension between assessment for support and assessment for high stakes decision making will never disappear. Still, that tension is constructively dealt with daily by teacher educators throughout the nation"* (Snyder, et al., 1998, p. 59).

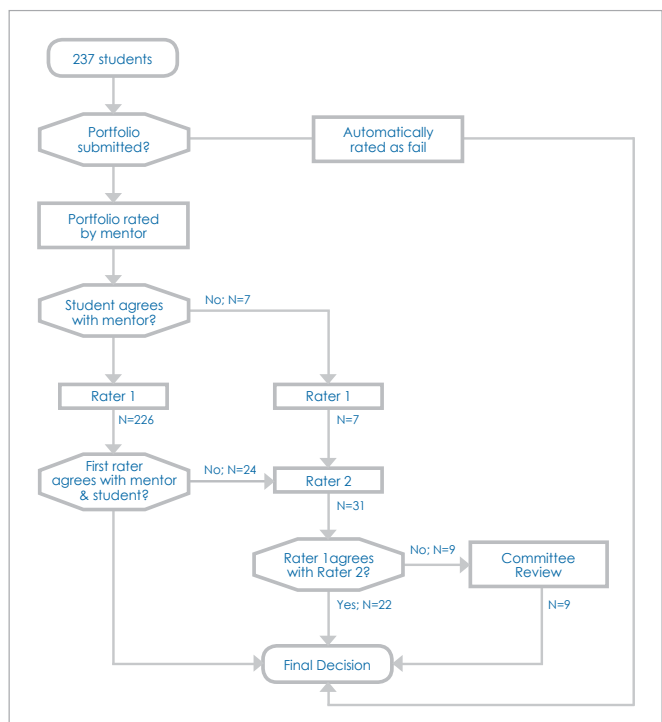
The tension between assessment for support and assessment for high stakes decision making will never disappear.

BOX 10
A procedure for portfolio assessment (Driessen et al., 2005a)

The student submits the portfolio to the mentor, who examines the portfolio and writes a recommendation regarding the grading of the portfolio to be submitted to the assessment committee.

In their final meeting of the academic year the student and the mentor discuss this recommendation. When student and mentor agree on the grade, the student signs the recommendation. If the student disagrees, he or she does not sign.

Subsequently, the portfolio is submitted to the assessment committee. This committee consists of all the mentors. The committee members do not grade the portfolios of the students they mentored themselves. Portfolios on which student and mentor agree are rated by one committee member, who does not study the portfolio in any great detail, but typically scans the work of the student and mentor and checks whether all procedures have been followed correctly. When rater and mentor agree on the grading, the recommendation becomes the final decision.



Striking the right balance between support and judgement is the challenge facing assessors/mentors with whom learners talk about their portfolios. A number of scenarios can be chosen in a procedure (Box 11). Which one is the most appropriate depends, amongst other things, on the educational context and the level of experience of the learners in question.

When? The answer to the question “when is this assessed?” depends on the answers to the other questions in this section.

Decisions about *selection* are made before the actual start of a programme or training period or after a first “trial” period, in which learners are observed and can prove themselves. The important question is whether a prospective learner matches the criteria for admission and whether this learner has the potential to finish an education or training programme.

Diagnostic assessment can be a frequent occurrence during an education or training programme. In fact, every time a mentor and a learner meet to discuss the learner’s progress using information from the learner’s portfolio, it can be qualified as diagnostic/ formative assessment. This implies that having easy access to a portfolio, for instance on-line, can be very helpful for mentors.

Decisions about *certification* are made when a learner's competencies match all the criteria or when the time available for a programme has run out. In an outcome based programme, this means that when the learner and his or her mentor conclude that the learner's competence meets all the criteria an assessment for certification purposes can take place. The logical consequence would be that if a person meets the competency criteria on entering an educational or training programme, he or she is exempt from training and awarded a certificate right away.

BOX 11

Portfolio assessors: scenarios

Combining the role of the mentor and assessor is often considered problematic. On the hand, most people will agree that the mentor is probably the person who is best informed about the learner's competencies. As a consequence, ignoring the mentor's opinion in assessing the portfolio can be considered as missing the chance to improve the validity of the assessment. On the other hand, combining the roles of assessor and mentor can put a strain on the relationship between mentor and learner, because learners may be reluctant to discuss any difficulties they are facing for fear of repercussions in the assessment. Below we use the metaphors of the mentor as teacher, PhD supervisor, driving instructor, and coach to distinguish between four (non exclusive) scenarios. When mentors are in the role of a teacher, their role of assessor is most prominent. When they are in the role of a coach, they do not assess at all.

The teacher: This is the most common assessment scenario in education. Just like most teachers in primary, secondary, and higher education, mentors discuss their learners' performance and progress and assess their level of competence at the end of a course.

PhD supervisor: In some scenarios the role of the mentors in the assessment procedure of portfolios can be compared with the role of supervisors of PhD students. In many countries, the formal assessment of theses/portfolios is the responsibility of a committee. Supervisors invite their peers to sit on the committee but they themselves are not a member of the committee. A negative assessment of the thesis/portfolio would harm their reputation among their peers. For this reason they are highly unlikely to invite their peers to sit on the committee unless they are convinced the portfolio meets the criteria. As a consequence, mentors and students have the same interest: to produce a thesis or portfolio that merits a positive judgment.

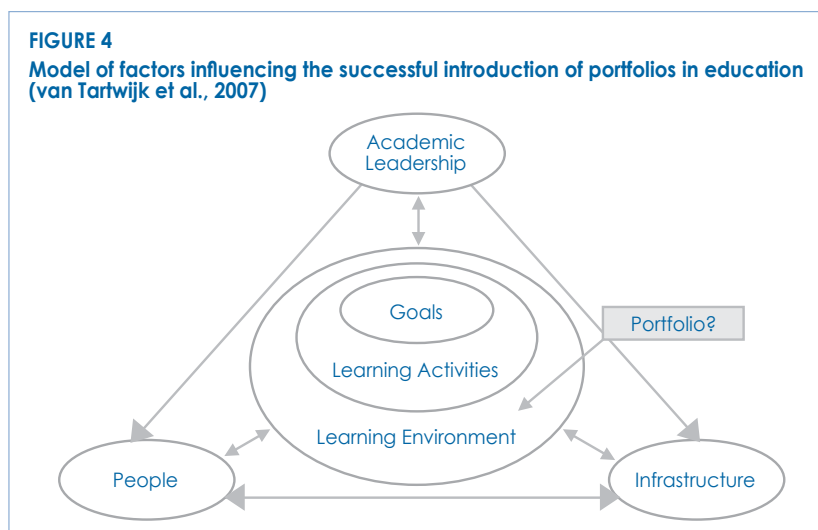
Driving license instructor: In this model the roles of the mentor and the assessor are strictly separated. The mentor/driving instructor mentors the learner in acquiring the required competencies, which are shown in the portfolio. If the mentor thinks the learner is competent, he invites an assessor from a professional body (i.e. the examiner from the Driver and Vehicle Licensing Agency) to assess the competence of the learner result. The learners can also approach the licensing agency themselves.

Coach: In this model, the learners themselves have the initiative. They can ask, for instance, a senior colleague to coach them until they have achieved the required level of competence. This scenario is likely, for instance, when a professional wants to acquire an additional qualification. The assessor would be someone from an external body.

Factors influencing the success of the introduction of a portfolio²

In the previous sections, we have argued that it is important to tailor portfolios to the intended purposes and to introduce portfolios only in situations in which they can serve a useful purpose. However, these conditions do not suffice to guarantee a successful introduction. In the literature on educational change, winning the hearts and minds of the people involved, both teachers and learners, as well as the quality of leadership are identified as key factors for lasting educational improvement (Martin et al. 2003; Hargreaves & Fink, 2004;).

Figure 4 presents a model in which portfolios are presented as part of the learning environment and in which three conditional factors are presented that influence whether an educational portfolio is introduced successfully or not: people (the teachers and learners), leadership, and infrastructure. The importance of these three conditional factors is discussed below.



People

Educational innovations involving the use of portfolios usually imply a transfer from teacher-directed education with a strong focus on conveying knowledge, to education in which the development of students' competencies in the workplace is emphasised. In most cases, teachers are expected to invest more time and effort in coaching and assessment than they were used to. Almost inevitably, this change in roles and routines will cause uncertainty and evoke resistance (Hammerness et al., 2005). Not only does it imply that teachers need to rethink key ideas, practices, and values, but for many teachers it also means that they need to invest in developing new competencies for coaching and assessment.

Educational innovations involving the use of portfolios usually imply a transfer from teacher-directed education with a strong focus on conveying knowledge, to education in which the development of students' competencies in the workplace is emphasised.

² Parts of this chapter were published before in Quality in Higher Education (van Tartwijk, et al., 2007)

In discussions about these innovations, the important questions are which educational problems need to be resolved and what is the most effective and efficient way to do that. Very often however, discussions concentrate on the portfolio, which becomes the visible "symbol" of the innovation. As a consequence, resistance to the innovation is likely to be projected onto the portfolio, while the important questions are not discussed.

Teachers are more likely to support and invest in educational changes if they acknowledge and subscribe to the educational value of the new learning approach, internalise and support the innovation, and are empowered to assume ownership of it. They are more likely to do so when it is clear to them how the innovation helps solve concrete problems that they have to cope with in their everyday teaching practice (Hargreaves et al. 1998). The risk that the important questions are not discussed can be reduced if teachers are involved in educational innovations at an early stage of decision-making. They are more likely to support and invest in working with a portfolio if the decision to work with this instrument was their own decision, based on their personal understanding and endorsement of the educational innovation and the role of the portfolio in it. From this perspective, the option should be kept of not using a portfolio when a better alternative is found. Teachers who have had a say in the decision to use a portfolio will feel a stronger commitment to it and will be more inclined to look for solutions and less likely to lay the instrument aside when faced with problems and inevitable design faults in the curriculum and the portfolio.

In the literature on educational change the importance of teachers as change agents is emphasised (Darling-Hammond et al., 2005) but the input of learners is crucial too. The successful introduction of a portfolio in education also depends on how much time and energy learners are willing to invest in their portfolios. In general, learners will only put effort into portfolios if this effort is rewarded in some way. The most obvious reward is that the portfolio is graded. In education, a very strong relationship exists between summative assessment and learning: assessment drives learning (Frederiksen, 1984; Driessen & van der Vleuten, 2000; van der Vleuten et al., 2000). Although assessment influences whether learners accept and put effort into a portfolio, assessment in itself is not enough. For learners, developing a portfolio implies putting a lot of effort into making their development visible. Thus, it is very frustrating for them if they discover that nobody takes a good look at the result of all their hard work. Mentors who take an interest in learners and their portfolios have been found to be a key factor in learners' appreciation of working with portfolios (Pearson & Heywood, 2004; Tigelaar et al. 2006).

A last condition for a successful introduction of portfolios related to learners and their mentors is their *understanding* of the portfolio and of what working with portfolios entails. Experience has shown that, although in theory portfolios can have a clear function in education, in practice the introduction of portfolios often leads to confusion and, consequently, frustration (Anderson & DeMeulle, 1998; Pearson & Heywood, 2004; Kjaer, et al., 2006; Davis et al. 2009). Most students who enrol in a medical curriculum are accustomed to teacher directed education. Self-assessment, asking for feedback, reflection and identifying personal learning needs, which are fundamental to portfolio learning (Snadden & Thomas, 1998b; Driessen et al. 2008), are perceived as

Although assessment influences whether learners accept and put effort into a portfolio, assessment in itself is not enough. For learners, developing a portfolio implies putting a lot of effort into making their development visible.

strange and sometimes even threatening by learners for whom education is synonymous with lectures and exams. Instructions are necessary that not only explain how to work with a portfolio, but also help learners and their mentors understand what a portfolio is and why it used in education. A study by Duque and colleagues (Duque et al., 2006) demonstrated that hands-on introduction with a proper briefing of learners by staff on the portfolio's purpose and procedures had a positive effect on portfolio scores and learner satisfaction with the portfolio. We have experimented with the use of the analogy between a portfolio and a CV to help learners better understand what a portfolio is and what working with a portfolio entails (van Tartwijk et al. 2008).

Academic leadership

Commitment by educational leaders is another vital condition for the successful introduction of portfolios. In a study on perceptions of leadership in academic contexts, Martin and her colleagues (2003) found that the quality of student learning is affected by the way leadership is constituted and experienced in academic contexts. A group of educational leaders was identified who were successful in stimulating teachers to adopt a student-focused approach to teaching. A characteristic of these educational leaders is that they discuss and negotiate these changes with the teachers. Similar findings are reported by Bland and her colleagues (2000), who reviewed the available literature with the aim to identify a set of characteristics that are associated with successful curricular change in medical education. They write that leadership comes up again and again as critical to the success of curricular change. The literature shows that successful and less successful leaders in medical education use organizational authority at about the same rate, but also that successful leaders more often seek input from others. When educational innovations ask teachers to change their roles and routines, these teachers must know that they can rely on educational leaders who support and value their commitment in every respect (Malden, 1994; van Veen et al. 2005). And finally, of course, commitment of the academic leaders is also reflected in the allocation of sufficient financial resources to ensure that the intended changes can actually be implemented.

Infrastructure

An increasing number of Faculties of Medicine are choosing to work with electronic rather than paper portfolios. In the section on e-portfolios, we described the reasons for this choice. We also wrote that research shows that adverse conditions like limited computer access in the workplace may cancel out the advantages of an e-portfolio. In general we conclude that e-portfolios are vulnerable to adverse conditions, because the demands of the technical infrastructure are large. If the electronic part of the portfolio system malfunctions, that is usually all the excuse that the adversaries of the use of portfolios need to drop the idea of a portfolio altogether, including the curriculum innovation for which the portfolio very often is a symbol.

Concluding remarks

In curricula with a strong focus on the development and assessment of competencies a portfolio can be a valuable instrument. They have the potential to make learning visible on the *Does* level of Miller's pyramid (Miller 1990), which describes independent performance in the workplace. However, portfolios are also vulnerable. Portfolio learning requires reflection by learners and investment in coaching by teachers. The quality of portfolio assessment depends on investing in the interpretation of and discussion about qualitative data. Not only does it require a new perspective on education from mentors and learners, many of whom are used to teacher-directed learning with a strong emphasis on the acquisition of knowledge, it also asks teachers and learners for a significant investment of time and energy. The literature shows that many conditions need to be fulfilled to enable successful introduction of a portfolio (Driessen et al., 2007b), and even then a portfolio is not a cure for all pains.

We conclude this Guide for using portfolios for assessment and learning by referring to Spandel once more (Spandel, 1997), who wrote:

"..... introducing portfolios is just like buying shoes: the best choice depends on purpose and comfort comes with wearing".

We would like to add that portfolios are like expensive shoes and even during the process of getting used to them, there will inevitably be times when one's toes are really hurting. However, for those owners who persist, the portfolio has the potential to become one of their best purchases.

Portfolio learning requires reflection by learners and investment in coaching by teachers. The quality of portfolio assessment depends on investing in the interpretation of and discussion about qualitative data.

"..... introducing portfolios is just like buying shoes: the best choice depends on purpose and comfort comes with wearing".

References

- ANDERSON RS & DEMEULLE L (1998). Portfolio use in twenty-four teacher education programs. *Teacher Education Quarterly*, 25: 23-32.
- BIRD T (1990). The schoolteacher's portfolio: an essay on possibilities. In: J Millman & L Darling-Hammond (Eds), *The new handbook of teacher evaluation: Assessing elementary and secondary school teachers*, pp. 241-256 (Newbury Park, CA, Corwin Press, inc).
- BLAND CJ, STARNAMAN S, WERSAL L, MOORHEAD-ROSENBERG L, ZONIA S & HENRY R (2000). Curricula change in medical schools: How to succeed. *Academic Medicine*, 75: 575-594.
- BRANSFORD J, BROWN AL & COCKING RR (Eds) (2000). *How people learn: Brain, mind, experience, and school*. (Washington D.C., National Academy Press).
- BUCKLEY S, ASHCROFT T, DAVIS J, KHAN KS, MORLEY D, POLLARD D, POPOVIC C, SAYERS J, SUSARLA R, THOMAS H & ZAMORA J (in press). The educational effects of portfolios on undergraduate student learning: A Best Evidence Medical Education systematic review. *Medical Teacher*.
- COLLINS A (1991). Portfolios for biology teacher assessment. *Journal of Personnel Evaluation in Education*, 5: 147-167.
- CONLON M (2003). Appraisal: The catalyst of personal development. *British Medical Journal*, 327: 389-391.
- DARLING-HAMMOND L, PACHECO A, MICHELLI N, LEPAGE P, HAMMERNESSE K & YOUNG P (2005). Implementing curriculum renewal in teacher education: managing organizational and policy change. In: L Darling-Hammond, J Bransford, P LePage, K Hammerness & H Duffy (Eds), *Preparing teachers for a changing world: What teachers should learn and be able to do*, pp. 442-479 (San Francisco, Jossey-Bass).
- DAVIS DA, MAZMANIAN PE, FORDIS M, VAN HARRISON R, THORPE KE & PERRIER L (2006). Accuracy of physician self-assessment compared with observed measures of competence: a systematic review. *JAMA*, 296: 1094-1102.
- DAVIS MH, FRIEDMAN BEN DAVID M, HARDEN RM, HOWIE P, KER J, MCGHEE C, et al. (2001). Portfolio assessment in medical students' final examinations. *Medical Teacher*, 23: 357-366.
- DAVIS MH, PONNAMPERUMA GG, & KER JS (2009). Student perceptions of a portfolio assessment process. *Medical Education*, 43: 89-98.
- DENZIN NK (1978). *Sociological Methods: A Sourcebook* (2nd ed.). New York: McGraw Hill.
- DENZIN NK & LINCOLN YS (2000). *Handbook of qualitative research*. Thousand Oaks, CA: Sage.
- DORNAN T, CARROLL C & PARBOOSHING J (2002). An electronic learning portfolio for reflective continuing professional development. *Medical Education*, 36: 767-769.
- DREYFUS SE (2004). The five-stage model of adult skill acquisition. *Bulletin of Science Technology and Society*, 24: 117-181.
- DRIESSEN EW, MUIJTJENS AMM, VAN TARTWIJK J & VAN DER VLEUTEN CPM (2007a). Web- or paper-based portfolios: is there a difference? *Medical education*, 41: 1067-1073.
- DRIESSEN EW & VAN DER VLEUTEN CPM (2000). Matching student assessment to problem based learning: lessons from experience in a law faculty. *Studies in Continuing Education*, 22: 235-248.
- DRIESSEN EW, VAN DER VLEUTEN CPM, SCHUWIRTH L, VAN TARTWIJK J & VERMUNT JD (2005a). Credibility of portfolio assessment as an alternative for reliability evaluation: a case study. *Medical Education*, 39: 214-220.
- DRIESSEN EW, VAN TARTWIJK J & DORNAN T (2008). The self-critical doctor: Helping students become more reflective. *BMJ*, 336: 827-830.
- DRIESSEN EW, VAN TARTWIJK J, OVEREEM K, VERMUNT JD & VAN DER VLEUTEN CPM (2005b). Conditions for successful reflective use of portfolios in undergraduate medical education. *Medical Education*, 39: 1230-1235.
- DRIESSEN EW, VAN TARTWIJK J, VAN DER VLEUTEN CPM, & WASS V (2007b). Portfolios in medical education: Why do they meet with mixed success? A systematic review. *Medical Education*, 41: 1224-1233.

- DRIESSEN EW, VAN TARTWIJK J, VERMUNT JD & VAN DER VLEUTEN CPM (2003). Use of portfolio in early undergraduate medical training. *Medical Teacher*, 25: 18-23.
- DUQUE G, FINKELSTEIN A, ROBERT A, TABATABAIA D, GOLD SL & WINER LR (2006). Learning while velauating: the use of an electronic evaluation portfolio in a geriatric medicine clerckship. *BMC Medical Education*, 6: 1-7.
- ERICSSON KA (2006). The influence of experience and deliberate practice on the development of expert performance. In: KA Ericsson, N Charness, PJ Feltovich & RR Hoffman (Eds), *The Cambridge handbook of expertise and expert performance* (pp. 683-704). New York: Cambridge University Press.
- EVA KW & REGEHR G (2008). "I'll never play professional football" and other fallacies of self-assessment. *Journal of Continuing Education in the Health Professions*, 28: 14-19.
- FINLAY IG, MAUGHAN TS & WEBSTER DJ (1998). A randomized controlled study of portfolio learning in undergraduate cancer education. *Medical Education*, 32: 172-176.
- FREDERIKSEN N (1984). The real test bias: Influences of testing on teaching and learning. *American Psychologist*, 39: 193-202.
- FRIEDMAN BEN DAVID M, DAVIS MH, HARDEN RM, HOWIE PW, KER J & PIPPARD MJ (2001). *AMEE Medical Education Guide No. 24: Portfolios as a method of student assessment* (Dundee, Association for Medical Education in Europe).
- FUNG KEE FUNG M, WALKER M, FUNG KEE FUNG K, TEMPLE L, LAJOIE F, BELLEMARE G, et al. (2000). An Internet-based learning portfolio in resident education: The KOALA-super (TM) multicentre programme. *Medical Education*, 34: 474-479.
- GENERAL MEDICAL COUNCIL (2000). *Revalidating doctors: Ensuring standards, securing the future*. London: GMC.
- GIBSON D & BARRETT H (2003). Directions in Electronic Portfolio Development. *Contemporary Issues in Technology and Teacher Education*, 2: 559-576.
- GRANT AJ, VERMUNT JD, KINNERSLEY P & HOUSTON H (2007). Exploring students' perceptions of the use of a significant event analysis as part of a portfolio assessment process in general practice, as a tool for learning how to use ferlection in learning. *BMC Medical Education*: 7:5.
- GUBA EG & LINCOLN YS (1989). Judging the quality of fourth generation evaluation. In: EG Guba & YS Lincoln (Eds), *Fourth Generation Evaluation* (London, Sage).
- HAMMERNES K, DARLING-HAMMOND L, BRANSFORD J, BERLINER DC, COCHRAN-SMITH M, MCDONALD M, et al. (2005). How teachers learn and develop. In: L Darling-Hammond, J Bransford, P LePage, K Hammerness & H Duffy (Eds), *Preparing teachers for a changing world: What teachers should learn and be able to do*, pp. 358-389 (San Francisco, Jossey-Bass).
- HARDEN RM (1979). How to assess students: An overview. *Medical Teacher*, 1: 65-70.
- HARGREAVES A & FINK D (2004). The seven principles of sustainable leadership. *Educational Leadership*, April 2004: 8-13.
- HARGREAVES A, LIEBERMAN A, FULLAN M & HOPKINS D (Eds) (1998). *International handbook of educational change* (Dordrecht: Kluwer Academic Publishers).
- HATTON N & SMITH D (1995). Reflection in teacher education: Towards definition and implementation. *Teaching and Teacher Education*, 11: 33-49.
- KJAER NK, MAAGARD R & WIES S (2006). Using an online portfolio in postgraduate training. *Medical Teacher*, 28: 708-712.
- KORTHAGEN FAJ, KESSELS J, KOSTER B, LAGERWERF B & WUBBELS T (2001). *Linking theory and practice: The pedagogy of realistic teacher education* (Mahwah, NY, Lawrence Erlbaum Associates).
- KORTHAGEN FAJ, KOSTER B, MELIEF K & TIGCHELAAR A (2002). *Teach teachers to reflect: Systematic reflection in the training and coaching of teachers* [In Dutch: Docenten leren reflecteren: Systematische reflectie in de opleiding en begeleiding van leraren] (Soest, Uitgeverij Nelissen).
- LAWSON M, NESTEL D & JOLLY B (2004). An e-portfolio in heath professional education. *Medical Education*, 38: 569-570.
- LOCKYER JM & CLYMAN SG (2008). Multisource feedback (360-degree feedback). In: ES Holmboe & RE Hawkins (Eds), *Practical guide to the evaluation of clinical competence*, pp. 75-85 (Philadelphia, Pa, Mosby Elsevier).

- LYONS N (1998). Reflection in teaching: Can it be developmental? A portfolio perspective. *Teacher Education Quarterly*, Winter 1998: 115-127.
- MALDEN B (1994). The micropolitics of education: mapping the multiple dimensions of power relations in school policies. *Journal of Educational Policy*, 9: 147-167.
- MANN K, GORDON J & MACLEOD A (2007). Reflections and reflective practice in health profession education: A systematic review. *Advanced Health Science Education*, (First published online November 2007): 1-27.
- MARTIN E, TRIGWELL K, PROSSER M & RAMSDEN P (2003). Variations in the experience of leadership of teaching in higher education. *Studies in Higher Education*, 28: 247-259.
- MATHERS NJ, CHALLIS MC, HOWE AC & FIELD NJ (1999). Portfolios in continuing medical education – effective and efficient? *Medical Education*, 33: 521-530.
- MILLER GE (1990). The assessment of clinical skills/competence/performance. *Academic Medicine*, 65: S63-67.
- NORCINI JJ & BURCH VC (Eds) (2007). *Workplace-based assessment as an educational tool*, AMEE Guide 31 (Dundee, UK, AMEE).
- NORCINI JJ, HOLMBOE ES & HAWKINS RE (2008). Evaluation challenges in the era of outcome based education. In: ES Holmboe & RE Hawkins (Eds), *Practical guide to the evaluation of clinical competence*, pp. 1-9 (Philadelphia, PA, Mosby Elsevier).
- O'SULLIVAN PS, RECKASE MD, MCCLAIN T, SAVIDGE MA & CLARDY JA (2004). Demonstration of portfolios to assess competency of residents. *Advances in Health Sciences Education*, 9: 1-15.
- OERMANN MH (2002). Developing a professional portfolio in Nursing. *Orthopaedic Nursing*, 21: 73-78.
- PAULSON FL, PAULSON PR & MEYER CA (1991). What makes a portfolio a portfolio? Eight thoughtful guidelines will help educators encourage self directed learning. *Educational Leadership*, February 1991: 60-63.
- PEARSON DJ & HEYWOOD P (2004). Portfolio use in general practice vocational training: A survey of GP registrars. *Medical Education*, 38: 87-95.
- ROYAL COLLEGE OF GENERAL PRACTITIONERS (1993). *Portfolio-based learning in general practice: Report of a working group on higher professional education*, Occasional paper 63 (London, Royal College of General Practitioners).
- ROYAL COLLEGE OF PHYSICIANS AND SURGEONS OF CANADA (1996). *Canmeds 2000 Project: Skills for the New Millennium. Report on the societal needs working group* (Ottawa, The Royal College of Physicians and Surgeons of Canada).
- SHULMAN LS (1998). Teacher portfolios: a theoretical activity. In: N Lyons (Ed), *With portfolio in hand: validating the new teacher professionalism*, pp. 23-38 (New York, Teachers College Press).
- SHUTE VJ (2008). Focus on formative feedback. *Review of Educational Research*, 78: 153-189.
- SMITHER JW, LONDON M, FLAUTT R, VARGAS Y & KUCINE I (2003). Can working with an executive coach improve multisource feedback ratings over time? A quasi-experimental field study. *Personal Psychology*, 56: 23-44.
- SNADDEN D (1999). Portfolios – attempting to measure the unmeasurable? [Commentary]. *Medical Education*, 33(7): 478-479.
- SNADDEN D, CHALLIS M, & THOMAS ML (1999). *AMEE Medical Education Guide No. 11: Portfolio-based learning and assessment* (Dundee, Association for Medical Education in Europe).
- SNADDEN D & THOMAS ML (1998a). Portfolio learning in general practice vocational training - does it work? *Medical Education*, 32: 401-406.
- SNADDEN D & THOMAS ML (1998b). The use of portfolio learning in medical education. *Medical Teacher*, 20: 192-199.
- SNYDER J, LIPPINCOTT A & BOWER D (1998). The inherent tensions in the multiple uses of portfolios in teacher education. *Teacher Education Quarterly*, 25: 45-60.
- SPANDEL V (1997). Reflections on portfolios. In: GD Phye (Ed), *Handbook of academic learning: Construction of knowledge* (pp. 573-591). San Diego: Academic Press.

- STOOF A, MARTENS RL, VAN MERRIËNBOER J & BASTIAENS TJ (2002). The boundary approach of competence: a constructivist aid for understanding and using the concept of competence. *Human resource development review*, 1, pp. 345-365.
- TIGELAAR DEH, DOLMANS DHJM, DE GRAVE WS, WOLFHAGEN HAP & VAN DER VLEUTEN CPM (2006). Participants opinions about the usefulness of a teaching portfolio. *Medical Education*, 40(4): 371-378.
- TIGELAAR DEH, DOLMANS DHJM, WOLFHAGEN HAP & VAN DER VLEUTEN CPM (2004). Using a conceptual framework and the opinion of portfolio experts to develop a teaching portfolio prototype. *Studies in Educational Evaluation*, 30: 305-321.
- TIGELAAR DEH, DOLMANS DHJM, WOLFHAGEN HAP & VAN DER VLEUTEN CPM (2005). Quality issues in judging portfolio: Implications for organizing teaching portfolio assessment procedures. *Studies in Higher Education*, 30: 595-610.
- TOCHEL C, HAIG A, HESKETH A, CADZOW A, BEGGS K, COLTHART L, et al. The effectiveness of portfolios for post-graduate assessment and education: a Best Evidence Medical Education systematic review. *Medical Teacher* (in press).
- VAN DER VLEUTEN CPM, DOLMANS DHJM & SCHERPBIER AJJA (2000). The need for evidence in education. *Medical Teacher*, 22: 246-250.
- VAN MANEN M (1977). Linking ways of knowing with ways of being practical. *Curriculum Inquiry*, 6: 205-228.
- VAN TARTWIJK J, DRIESSEN EW, STOKKING K & VAN DER VLEUTEN CPM (2007). Factors influencing the successful introduction of portfolios. *Quality in Higher Education*, 13: 69-79.
- VAN TARTWIJK J, VAN RIJSWIJK M, TUIHOF H & DRIESSEN EW (2008). Using an analogy in the introduction of a portfolio. *Teaching and Teacher Education*, 24: 927-938.
- VAN VEEN K, SLEEGERS P, & VAN DE VEN P (2005). One teacher's identity, emotions, and commitment to change: A case study into the cognitive-affective processes of a secondary school teacher in the context of reforms. *Teaching and Teacher Education*, 21: 917-934.
- WEBB C, ENDACOTT R, GRAY MA, JASPER MA, MCCULLAN M & SCHOLES J (2003). Evaluating portfolio assessment systems: What are the appropriate criteria? *Nurse Education Today*, 23: 600-609.
- WOODWARD H & NANLOHY P (2004). Digital portfolios: Fact or fashion. *Assessment & Evaluation in Higher Education*, 29: 227-238.

Series 2

- 30 Peer Assisted Learning: a planning and implementation framework**
Michael Ross & Helen Cameron (2007)
ISBN: 978-1-903934-38-8
Primarily designed to assist curriculum developers, course organisers and educational researchers develop and implement their own PAL initiatives.
- 31 Workplace-based Assessment as an Educational Tool**
John Norcini & Vanessa Burch (2008)
ISBN: 978-1-903934-39-5
Several methods for assessing work-based activities are described, with preliminary evidence of their application, practicability, reliability and validity.
- 32 e-Learning in Medical Education**
Rachel Ellaway & Ken Masters (2008)
ISBN: 978-1-903934-41-8
An increasingly important topic in medical education – a 'must read' introduction for the novice and a useful resource and update for the more experienced practitioner.
- 33 Faculty Development: Yesterday, Today and Tomorrow**
Michelle McLean, Francois Cilliers & Jacqueline M van Wyk (2010)
ISBN: 978-1-903934-42-5
Useful frameworks for designing, implementing and evaluating faculty development programmes.
- 34 Teaching in the clinical environment**
Subha Ramani & Sam Leinster (2008)
ISBN: 978-1-903934-43-2
An examination of the many challenges for teachers in the clinical environment, application of relevant educational theories to the clinical context and practical teaching tips for clinical teachers.
- 35 Continuing Medical Education**
Nancy Davis, David Davis & Ralph Bloch (2010)
ISBN: 978-1-903934-44-9
Designed to provide a foundation for developing effective continuing medical education (CME) for practicing physicians.
- 36 Problem-Based Learning: where are we now?**
David Taylor & Barbara Mifflin (2010)
ISBN: 978-1-903934-45-6
A look at the various interpretations and practices that claim the label PBL, and a critique of these against the original concept and practice.
- 37 Setting and maintaining standards in multiple choice examinations**
Raja C Bandaranayake (2010)
ISBN: 978-1-903934-51-7
An examination of the more commonly used methods of standard setting together with their advantages and disadvantages and illustrations of the procedures used in each, with the help of an example.
- 38 Learning in Interprofessional Terms**
Marilyn Hammick, Lorna Olckers & Charles Campion-Smith (2010)
ISBN: 978-1-903934-52-4
Clarification of what is meant by Interprofessional learning and an exploration of the concept of teams and team working.
- 39 Online eAssessment**
Reg Dennick, Simon Wilkinson & Nigel Purcell (2010)
ISBN: 978-1-903934-53-1
An outline of the advantages of on-line eAssessment and an examination of the intellectual, technical, learning and cost issues that arise from its use.
- 40 Creating effective poster presentations**
George Hess, Kathryn Tosney & Leon Liegel (2009)
ISBN: 978-1-903934-48-7
Practical tips on preparing a poster – an important, but often badly executed communication tool.
- 41 The Place of Anatomy in Medical Education**
Graham Louw, Norman Eizenberg & Stephen W Carmichael (2010)
ISBN: 978-1-903934-54-8
The teaching of anatomy in a traditional and in a problem-based curriculum from a practical and a theoretical perspective.
- 42 The use of simulated patients in medical education**
Jennifer A Cleland, Keiko Abe & Jan-Joost Rethans (2010)
ISBN: 978-1-903934-55-5
A detailed overview on how to recruit, train and use Standardized Patients from a teaching and assessment perspective.
- 43 Scholarship, Publication and Career Advancement in Health Professions Education**
William C McGaghie (2010)
ISBN: 978-1-903934-50-0
Advice for the teacher on the preparation and publication of manuscripts and twenty-one practical suggestions about how to advance a successful and satisfying career in the academic health professions.
- 44 The Use of Reflection in Medical Education**
John Sandars (2010)
ISBN: 978-1-903934-56-2
A variety of educational approaches in undergraduate, postgraduate and continuing medical education that can be used for reflection, from text based reflective journals and critical incident reports to the creative use of digital media and storytelling.
- 45 Portfolios for Assessment and Learning**
Jan van Tartwijk & Erik W Driessen (2010)
ISBN: 978-1-903934-57-9
An overview of the content and structure of various types of portfolios, including eportfolios, and the factors that influence their success.
- 46 Student Selected Components**
Simon C Riley (2010)
ISBN: 978-1-903934-58-6
An insight into the structure of an SSC programme and its various important component parts.
- 47 Using Rural and Remote Settings in the Undergraduate Medical Curriculum**
Maira Maley, Paul Worley & John Dent (2010)
ISBN: 978-1-903934-59-3
A description of an RRME programme in action with a discussion of the potential benefits and issues relating to implementation.
- 48 Effective Small Group Learning**
Sarah Edmunds & George Brown (2010)
ISBN: 978-1-903934-60-9
An overview of the use of small group methods in medicine and what makes them effective.

To see the full list of guides available, and to order, see the website www.amee.org.

About AMEE

What is AMEE?

AMEE is an association for all with an interest in medical and healthcare professions education, with members throughout the world. AMEE's interests span the continuum of education from undergraduate/basic training, through postgraduate/specialist training, to continuing professional development/continuing medical education.

- **Conferences:** Since 1973 AMEE has been organising an annual conference, held in a European city. The conference now attracts over 2300 participants from 80 countries.
- **Courses:** AMEE offers a series of courses at AMEE and other major medical education conferences relating to teaching, assessment, research and technology in medical education.
- **MedEdWorld:** AMEE's exciting new initiative has been established to help all concerned with medical education to keep up to date with developments in the field, to promote networking and sharing of ideas and resources between members and to promote collaborative learning between students and teachers internationally.
- **Medical Teacher:** AMEE produces a leading international journal, Medical Teacher, published 12 times a year, included in the membership fee for individual and student members.
- **Education Guides:** AMEE also produces a series of education guides on a range of topics, including Best Evidence Medical Education Guides reporting results of BEME Systematic Reviews in medical education.
- **Best Evidence Medical Education (BEME):** AMEE is a leading player in the BEME initiative which aims to create a culture of the use of best evidence in making decisions about teaching in medical and healthcare professions education.

Membership categories

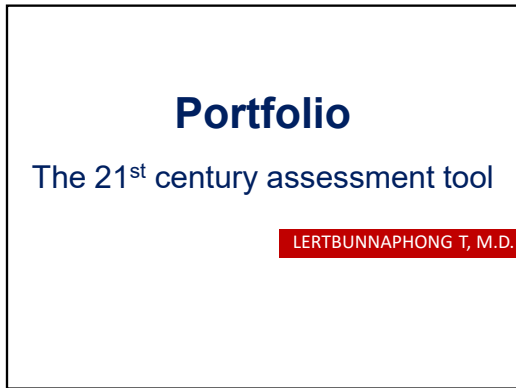
- **Individual and student members (£85/£39 a year):** Receive Medical Teacher (12 issues a year, hard copy and online access), free membership of MedEdWorld, discount on conference attendance and discount on publications.
- **Institutional membership (£200 a year):** Receive free membership of MedEdWorld for the institution, discount on conference attendance for members of the institution and discount on publications.

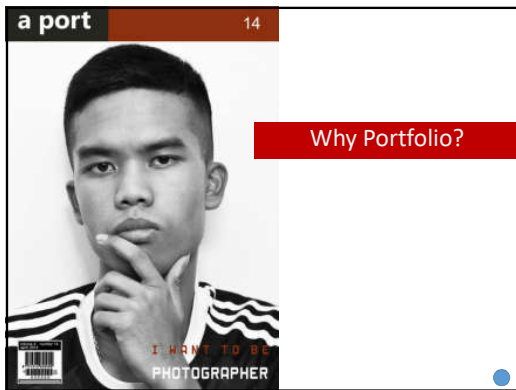
See the website (www.amee.org) for more information.

If you would like more information about AMEE and its activities, please contact the AMEE Office:
Association for Medical Education in Europe (AMEE), Tay Park House, 484 Perth Road, Dundee DD2 1LR, UK
Tel: +44 (0)1382 381953; Fax: +44 (0)1382 381987; Email: amee@dundee.ac.uk

www.amee.org

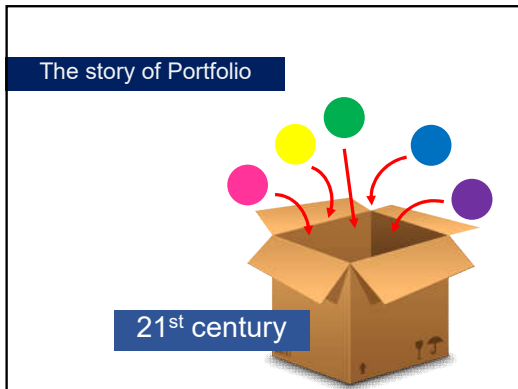
Scottish Charity No. SC 031618







3/8/2019







3/8/2019

It is your **DESIGN** as long as...
It reflects your **LEARNING OUTCOMES**

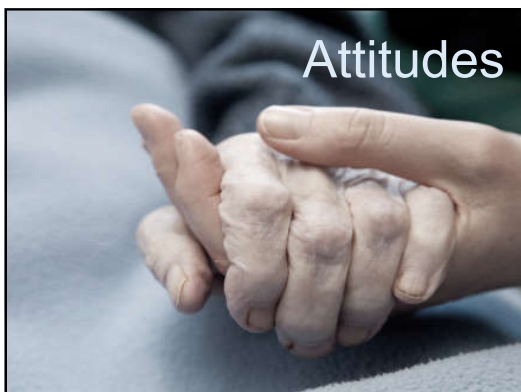
แต่ละกลุ่มกำหนดเป้าหมายและหลักฐานเพื่อสะสมใน Portfolio (เวลา 10 นาที)	
บัณฑิตศึกษา นักศึกษาแพทย์ปี6 แพทย์ประจำบ้าน แพทย์ประจำบ้านดอยอด	
Learning outcomes and why?	Evidences and why?

1. กำหนดเป้าหมาย
VISION
MISSION
POLICY
DESIRED DOCTOR

3/8/2019







3/8/2019

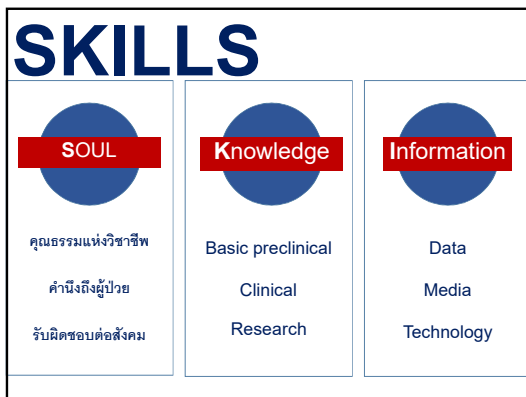


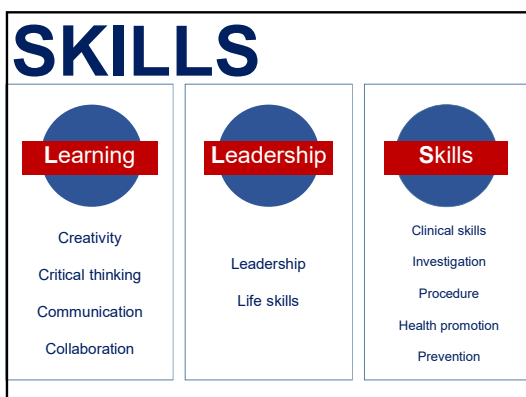


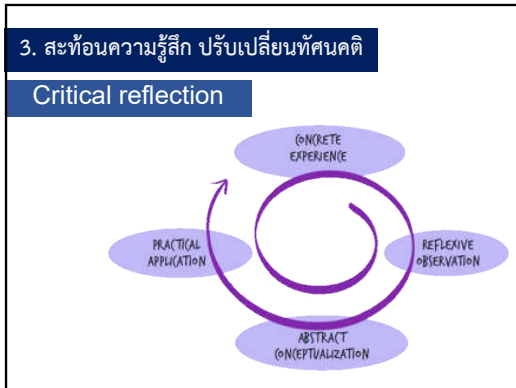


3/8/2019







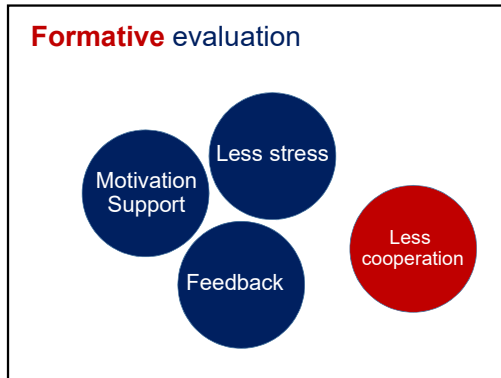


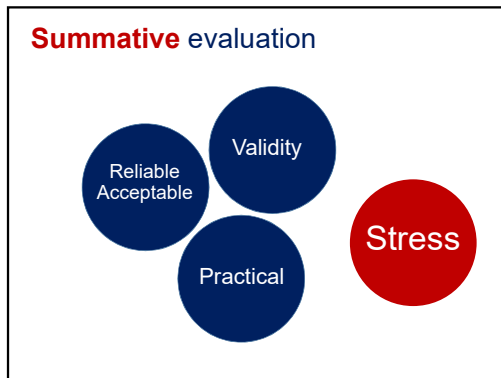
Learning without reflection is **waste**
Confucius

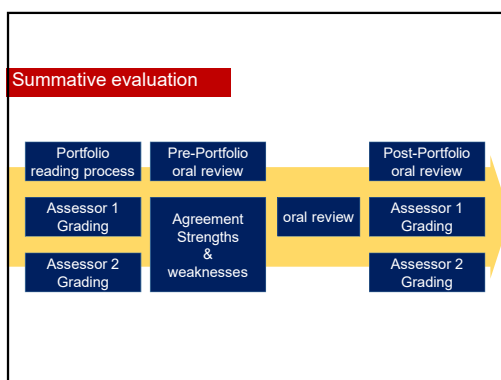
4. ประเมินผล

Summative
or
Formative

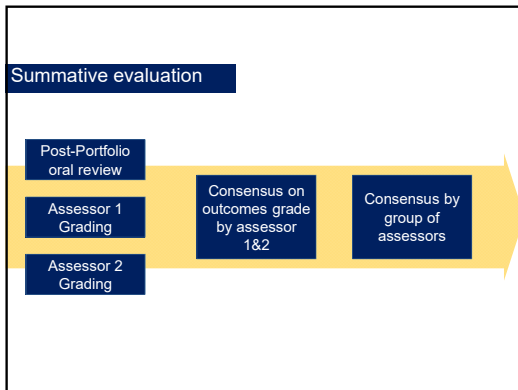
3/8/2019

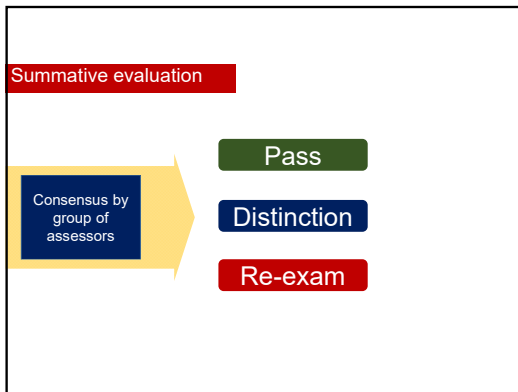






3/8/2019





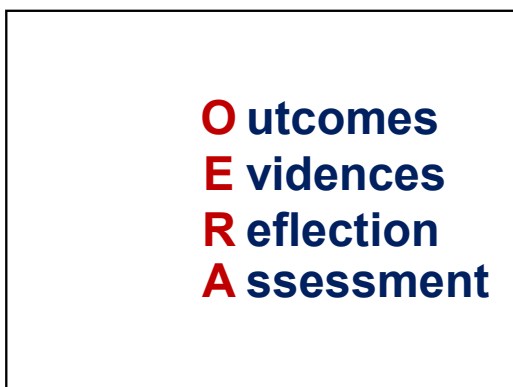
Formative or Summative?

It depends

3/8/2019







เอกสารประกอบการอบรม



TEACH

14 November 2019

Essential Skills for Thesis Advisor 2019

Thawornchai Limjindaporn, MD, PhD

Pa-thai Yenchitsomanus, PhD

Bordin Sapsomboon, MD, PhD

As an advisor

- 1. How do you educate and evaluate the students?***
- 2. How do you communicate to the students?***
- 3. How do you build a team and leadership skill for the students?***

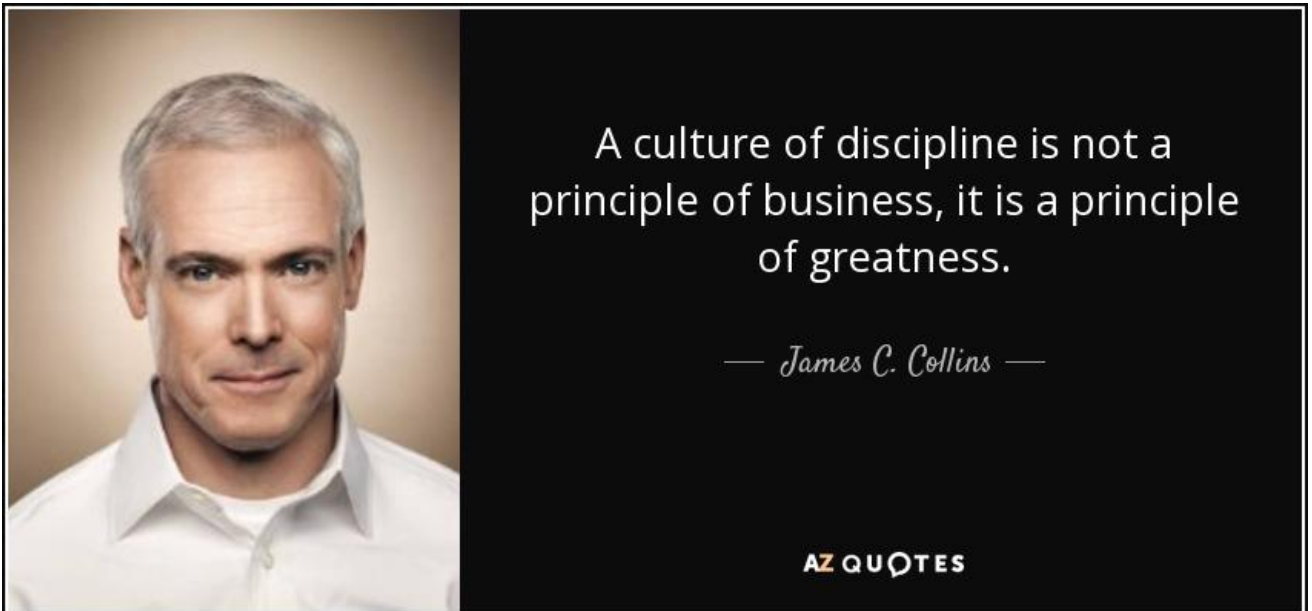
Educate and Evaluate the student

- ***Assigning work and follow up***
 - ***Give assignment from simplicity to difficulty***
 - ***Understand the caliber of your student***
 - ***Check the timing***
- ***Observing performance***
 - ***Watch over students not only to assess work quality but also to evaluate conduct, appearance, vitality, attitude and eagerness to learn***

Educate and Evaluate the student

- ***Documenting performance***
 - ***Treat documentation as a communication tool to get facts thereby removing ambiguities***
- ***Evaluating performance***
 - ***Give student ongoing feedback on performance both good and bad so that they know what they are doing right and what they need to improve as it has a lasting impact on the student***

Treat the students with discipline and cultural manner



Communicate to the students

- ***Inspire and appraise the student***
 - ***Taking every opportunity to appreciate their contribution and to urge them to excellence***
- ***Availability for the student***
 - ***maintain a proper timing in office and lab***
- ***Mastering constructive confrontation***
 - ***Speak with clarity and purpose***
 - ***Investigate complaints properly***
 - ***Take an unbiased fact-based approach when investigation of complaints***

Communication has profound effects in terms of how teachers and students can learn from each other



I'm a great believer that any tool that enhances communication has profound effects in terms of how people can learn from each other, and how they can achieve the kind of freedoms that they're interested in.

— *Bill Gates* —

AZ QUOTES

Building a team and leadership skill for the students

- **Building your team**
 - ***Select students and get them into a shared goal***
 - ***Year plan for a shared goal of group and for a student***
- **Resolving the team conflicts**
 - ***Focus on a shared goal to referee disputes***
- **Managing unfit acting and violence**
 - ***Know the conditions that breed violence and protect your workplace from the conflicts***

Building a team and leadership skill for the students

- **Be a discerning listener**
 - *An important trait of a good leader is someone who listens to suggestions, ideas, and feedback from other people, and build on them*
- **Learn to follow**
 - *A true leader has no problem yielding control to another person when appropriate.*
- **Keep learning**
 - *The best path to becoming a good leader is to always keep learning new things. It keeps your mind sharp, and your skills fresh.*

Get students into a shared goal



TEAMWORK...means never having to
take all the blame yourself.

— *Stephen Hawking* —

AZ QUOTES

References

- <https://www.wrike.com/blog/9-ways-develop-leadership-skills/>
- <https://blog.azendoo.com/8-ways-to-improve-team-communication/>

AMEE Guide No. 25: The assessment of learning outcomes for the competent and reflective physician

J.M. SHUMWAY¹ & R.M. HARDEN²

¹Office of Medical Education, West Virginia University, USA;

²International Virtual Medical School, 484 Perth Road, Tay Park House, Dundee, UK

SUMMARY Two important features of contemporary medical education are recognized. The first is an emphasis on assessment as a tool to ensure quality in training programmes, to motivate students and to direct what they learn. The second is a move to outcome-based education where the learning outcomes are defined and decisions about the curriculum are based on these. These two trends are closely related. If teachers are to do a better job of assessing their students, they need an understanding of the assessment process, an appreciation of the learning outcomes to be assessed and a recognition of the most appropriate tools to assess each outcome. Assessment tools selected should be valid, reliable, practical and have an appropriate impact on student learning. The preferred assessment tool will vary with the outcome to be assessed. It is likely to be some form of written test, a performance test such as an OSCE in which the student's competence can be tested in a simulated situation, and a test of the student's behaviour over time in clinical practice, based on tutors' reports and students' portfolios. An assessment profile can be produced for each student which highlights the learning outcomes the student has achieved at the required standard and other outcomes where this is not the case. For educational as well as economic reasons, there should be collaboration across the continuum of education in test development as it relates to the assessment of learning outcomes and in the implementation of a competence-based approach to assessment.

The importance of assessment

Assessment plays a major role in the process of medical education, in the lives of medical students, and in society by certifying competent physicians who can take care of the public. The very foundation of medical curricula is built around assessment milestones for students. For example, in the United States medical students must pass a series of steps towards licensure before graduating from medical school. It is assessment and evaluation that often drives the curricula of medical schools and students measure their progress through the curriculum by the examinations they have passed. Assessment becomes a motivating force for them to learn. Society has the right to know that physicians who graduate from medical school and subsequent residency training programmes are competent and can practise their profession in a compassionate and skilful manner. It is the responsibility of the medical school to demonstrate that such competence has been achieved, and the responsibility of accreditation agencies to certify that the educational programmes in medical schools can do what they promise. Assessment is of fundamental importance because it is central to public accountability.

The General Medical Council (GMC) has the responsibility to ensure that graduates of a UK medical school have met the requirements for their next posts as house officers. In 1993 they issued their recommendations on undergraduate medical education (GMC, 1993). More recent recommendations (GMC, 2002) place greater emphasis on learning outcomes and on the assessment of the outcomes. "In line with current educational theory and research, we have adopted an outcomes-based model. This sets out what is to be achieved and assessed at the end of the medical course in terms of knowledge, skills and behaviour" (Rubin & Franchi-Christopher, 2002). Table 1 contains a summary of those recommendations as they relate to assessment.

In the United States, the Liaison Committee for Medical Education (LCME) is the accreditation agency for North American medical schools (USA and Canada). Medical schools in North America have traditionally been accredited on the quality of the elements that make up the student educational programme (e.g. faculty, research, facilities, courses and clerkships). There are essentially four questions asked during accreditation: (1) What are the goals?; (2) What did students actually learn?; (3) What is the evidence?; and (4) What needs to be changed? The LCME has instituted standards focusing on the assessment of outcomes (LCME, 2003). In outcome-based assessment the educational programme goals or learning outcomes are defined and their accomplishment is assessed. North American medical education institutions are now required to document educational outcomes in light of their institutional purposes and missions. The LCME standards pertaining to the assessment of these outcomes are included in Table 2.

Outcome-based assessment for a competent and reflective physician

Assessment is an intrinsic component of outcome-based education. Outcome-based education and performance assessment are closely related paradigms (Friedman Ben-David, 1999). Outcome-based education involves an educational approach in which the decisions about the curriculum and evaluation are driven by the learning outcomes that students should achieve (Harden *et al.*, 1999 a). In this approach, the product (student learning outcomes) defines the process (instructional methods and learning opportunities). This is distinctively different from earlier educational

Correspondence: Professor James M. Shumway, Office of Medical Education, West Virginia University, School of Medicine, PO Box 9007, Morgantown, WV 26506, USA. Email: jshumway@hsc.wvu.edu

J.M. Shumway & R.M. Harden

Table 1. Recommendations of the General Medical Council in the UK relating to assessing student performance and competence (GMC, 2002).

The Principles of Assessment

Schemes of assessment must support the curriculum and allow students to prove that they have achieved the curricular outcomes. This means assessment must allow students to demonstrate the breadth and depth of their knowledge, and to show what they can do. Professional attitudes and behaviour must also be assessed.

Medical schools should use a range of assessment techniques that are appropriate for testing the curricular outcomes. Medical schools should determine the most appropriate scheme of assessment for their curriculum. However, schemes must meet best practice in assessment, and medical schools must be able to provide evidence that the schemes are valid and reliable, and that they have processes for setting standards and making decisions about student performance.

When students get close to graduating, their knowledge, skills, attitudes and behaviour must be thoroughly assessed to determine their fitness to practice as pre-registration house officers.

Assessment Procedures

Schemes of assessment must be open, fair and meet appropriate standards. Medical schools must make sure that:

- a. there is a clear indication of how the scheme of assessment deals with all the curricular outcomes;
- b. there is a clear indication of how individual assessments and examinations contribute to the overall assessment of the curricular outcomes;
- c. when they design individual assessments, there is a clear indication of how the targeted curricular outcomes have been met;
- d. students have clear guidance about what is expected of them in any examination or assessment;
- e. examiners are trained to carry out the role and to apply the medical school's assessment criteria consistently;
- f. examiners have clear guidelines for marking assessments, which indicate how performance against targeted curricular outcomes should be rewarded;
- g. systems are in place to determine the pass mark; and
- h. external examiners are employed to make sure that standards are met.

Note: The methods are listed for each of the assessment outcomes in order of importance.

Table 2. LCME standards related to the assessment of learning outcomes (LCME, 2003).

“Medical schools must evaluate educational program effectiveness by documenting the achievements of their students and graduates in verifiable and internally consistent ways that show the extent to which institutional and program purposes are met.”

“Medical schools should use a variety of measures to evaluate program quality, such as data on student performance/achievement, acceptance into residency programs, postgraduate performance and licensing, and emerging measures that may prove to be valid.”

“The results of such evaluations should be used to determine how well schools are fulfilling their objectives and to assess the need for program improvement.”

“Schools also should evaluate the performance of their students and graduates in the framework of national norms of accomplishment.”

approaches that relied on inputs. These approaches, having defined the educational programme, accepted whatever outcomes resulted from the process. The assumption was that a ‘better’ process would result in ‘better’ outcomes. In outcome-based education, agreement on predetermined student learning outcomes defines the processes used to achieve them. Put simply, outcome-based education has two requirements:

- (1) Expected learning outcomes are initially agreed and then communicated to all involved in the educational process (students, instructors, employers, the public, etc.).
- (2) Learning outcomes determine curriculum content, teaching methods and assessment.

All decisions concerning the curriculum are based on achieving the desired learning outcomes.

Outcome-based education programmes, argued Friedman Ben-David (1999), are faced with the need to develop non-traditional teaching and assessment techniques which capture both the learning and performance of broad abilities. The clearer the definition of learning outcomes, the more effectively can the assessment process be planned and implemented.

Over the past 10 years, educators and individuals in a number of countries have been developing approaches to outcome-based education (Harden, 2002). In the USA, Kassebaum *et al.* (1997) reviewed accreditation reports of medical schools and found that only a small percentage of schools had ‘robust’ institutional objectives that guided their educational programmes. They noted that schools where there was a lack of institutional objectives were more likely to have accreditation citations for shortcomings in curricular management. The Association of American Medical Colleges

(AAMC) formed an advisory group to recommend guidelines to US medical schools on the objectives of medical education. The Medical Schools Objectives Project (MSOP) identified attributes based on society's expectations of a good physician. They were grouped into four categories: (1) physicians must be altruistic; (2) physicians must be knowledgeable; (3) physicians must be skilful; and (4) physicians must be dutiful (AAMC, 1998). Some individual medical schools in the USA have developed their own competences. Brown University in Providence, Rhode Island, described a list of nine abilities (Smith & Dollase, 1999; Smith, 1999). Likewise, in Canada, physician groups developed the essential competences and roles of their profession. The CanMEDS 2000 Project Societal Needs Working Group reported seven roles of specialist physicians: (1) medical expert, (2) communicator, (3) collaborator, (4) manager, (5) health advocate, (6) scholar, and (7) professional (CanMEDS, 2000). Within each of these categories, a number of specific attributes or objectives were identified. The Accreditation Council for Graduate Medical Education (ACGME) in the USA also specified learning outcomes for the training of the doctor (ACGME, 2003). The International Institute of Medical Education has produced an international consensus on the learning outcomes or minimum essential requirements expected of a student on graduation from medical school (IIME, 2002).

In the UK, all five Scottish medical schools have adopted the same framework for learning outcomes (Simpson *et al.*, 2002). This is based on the three-circle model, devised to classify the learning outcomes at the University of Dundee School of Medicine as illustrated in Figure 1 (Harden *et al.*, 1999a, 1999b). The model is based on the three essential aspects of competence of a generalist physician. The inner sphere describes those things the physician is able to do. These include the clinical, procedural, investigation, management, health promotion, communication, and information-handling skills. The middle layer represents how the physician approaches the skills with knowledge and understanding, ethical/legal principles, and clinical reasoning and decision-making abilities. The outer layer represents the development of the personal characteristics of the physician. Such characteristics include understanding physicians' roles in society and their personal development as lifelong learners and professionals. This model has also been used in other

countries such as Spain and Sweden and in the training of junior doctors and specialists in the UK.

The outcomes specified by these different bodies have similarities and embrace a similar set of competences or abilities. For the purpose of this guide we have used the 12 learning outcomes as identified in the Dundee three-circle model.

Aim of this guide

Many scholars in the field of assessment have contributed to descriptions of approaches to assessment and the assessment instruments that are referred to in this guide. Less has been written, however, on the application of these assessment approaches to the specific learning outcomes of medical education (ACGME and ABMS, 2000; Scottish Deans' Medical Curriculum Group, 2002).

A critical examination of practice relating to student assessment should be undertaken in the context of the changing expectations with regard to students' learning outcomes. There has been a change in emphasis from knowledge acquisition and factual recall to more widely embracing learning outcomes such as problem solving, clinical judgement, communication skills, attitudes and professionalism. To date, however, these changes in the curriculum have not been fully reflected in the assessment process.

In the United States, multiple-choice questions continue to play a major role as an assessment tool, especially in the early years of the curriculum. While clinical and technical skills are increasingly assessed by approaches such as the objective structured clinical examination (OSCE) and the use of standardized patients (SPs), little has been done to assess attitudes and other professional behaviours and characteristics of physicians in training. Such mismatches between the choice of assessment tools and the expected learning outcomes cannot be attributed to the lack of appropriate assessment instruments. An extensive number of assessment instruments is now available to the examiner.

The range of instruments from which examiners can make their choice includes traditional approaches such as written questions, approaches introduced over the past few decades such as the OSCE and relative newcomers to the assessment scene such as portfolios.

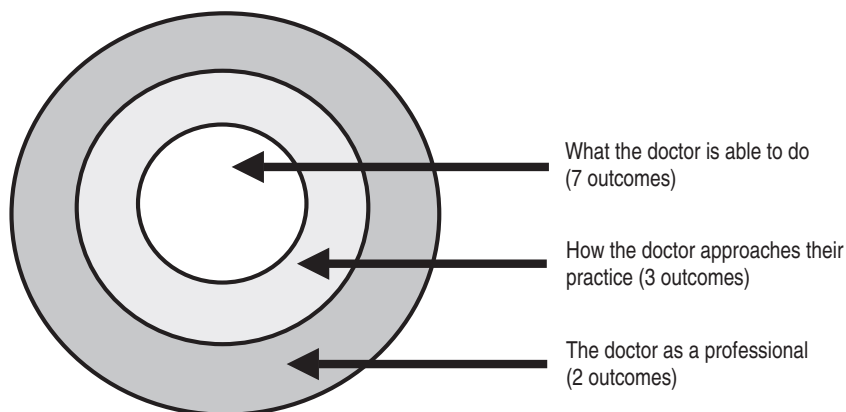


Figure 1. The three-circle framework for learning outcomes.

J.M. Shumway & R.M. Harden

In designing and implementing a programme of student assessment, examiners must be knowledgeable about the expected learning outcomes and the methods available which might be applied to their measurement. They must also have an understanding of the assessment process in order to allow them to make an appropriate choice of methods and implement them in practice. It is important to understand how to use new assessment tools, and how to evaluate their strengths, weaknesses and applicability under various instructional conditions.

This guide describes assessment approaches for each of the specific competences expected of students or trainees (learners). All schools and institutions engaged in the move to competence-based approaches to medical education, with the associated need to assess expected learning outcomes, should find the suggestions helpful.

The guide takes a multi-dimensional view of assessment (Figure 2). It describes how as educators we can do a better job of assessing students by assessing the full range of learning outcomes, by choosing appropriate instruments for the purpose and by having an understanding of the assessment process.

Improved understanding of the assessment process

Assessment instruments can be described according to certain prescribed criteria that are evidence based and recognized by professionals in the field. The criteria most commonly referred to are: (1) validity, (2) reliability, (3) impact on the learner and educational programme, and (4) practicality including cost.

Validity

The validity of an instrument is the degree to which an instrument measures what it is supposed to measure (Goetz *et al.*, 1992; Atkinson *et al.*, 1993; Pressley & McCormick, 1995). It is concerned with whether there is anything about a test that affects an examinee’s score so that the test fails to measure the learning outcomes intended. Unfortunately, a gold standard to measure the validity of an assessment instrument against is often unavailable. For assessment instruments, validity concerns itself with a specific measurement in a specific situation with a specific group of individuals. In other words, it is not the instrument or approach that is valid but the scores that are valid (ACGME and ABMS, 2000). What is being measured depends as much on the content of the assessment as on any characteristic of the method (Van der Vleuten, 1996).

Validity can be broken down into three major categories: content validity, criterion-related validity and construct validity. Content validity is determined by a review of the assessment instrument and the extent to which it measures what it is intended to measure. Criterion-related validity refers to a comparison of the test score against a known criterion of the expected performance. Construct validity refers to a collection of indirect information that the assessment instrument measures what it purports to measure. For example, construct validity is corroborated if the instrument is able to distinguish between different levels of trainees (Winckel *et al.*, 1994). A simple way to remember the concept of validity is to think of a ‘bulls-eye target’, such as

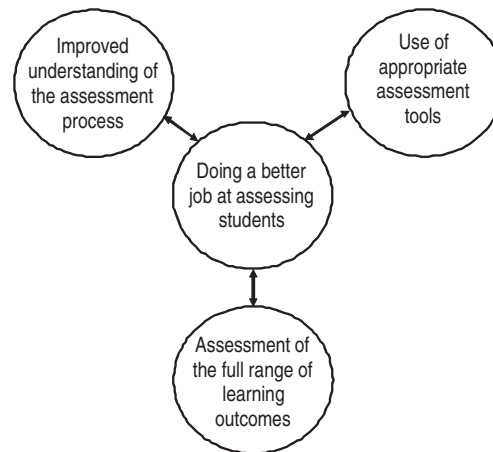


Figure 2. A multi-dimensional model of assessment.

used in darts and archery. The degree of accuracy to which items end up in the desired central targeted area indicates the degree of validity of measuring what was intended.

Reliability

The reliability of a test is the consistency, generalizability or reproducibility of a test. It is the extent to which examinees’ scores would be similar if they were retested. Reliability is the degree to which an instrument consistently measures the phenomena over time (Goetz *et al.*, 1992; Atkinson *et al.*, 1993; Pressley & McCormick, 1995). If the test is repeated over time, then the new test results should be similar to the earlier scores for the same assessment instrument on the same group of individuals (ACGME and ABMS, 2000). This is called the stability of test scores or test-retest reliability. Another approach to estimating reliability is to prepare two forms of a test and compare the results. This is called equivalent or alternative-forms reliability. Internal consistency is the reliability approach most typically practised. It involves the extent to which a student’s scores on two or more portions of a test agree with each other (Goetz *et al.*, 1992). Split-half reliability is the simplest approach for measuring internal consistency.

Reliability is usually measured as a correlation coefficient with 1.0 being a perfect correlation and zero being no correlation. Values above 0.70 are considered to indicate a reliable instrument although some aim for a figure above 0.80. Inter-rater reliability refers to the level of agreement among different independent evaluators of the same examinee (Winckel *et al.*, 1994). An instrument is unlikely to be reliable if it has only a small number of items sampling the intended behaviour. So reliability is a matter of adequately sampling the content. Reliability is also a function of the time spent in the test setting (Van der Vleuten, 2000). Again, thinking of a bulls-eye target is also helpful with regard to reliability. If the results of the assessment cluster around each other like the arrows on the target, then it can be said that the evaluations are reliable, even if the cluster is off centre of the target.

Impact on learning

An important but often ignored characteristic of an evaluation instrument is its impact on the learner. Nowhere is this

more true than in medical education. An old adage in medical education is that ‘examinations drive the curriculum’. The type of examination can determine how students study, thus how they learn. It is also true that the importance and frequency of examinations also impact on the student’s learning environment. Student responses to examinations may include memorization, studying at the last moment, concentrating on their perception of ‘what is the right answer’, and cheating. Van der Vleuten (2000) has emphasized that medical educators need to pay more attention to these characteristics than is currently done. Van der Vleuten (1996) has described at least four ways that assessment can drive the learning process. Assessment drives learning through: (1) its content, (2) its structure or format, (3) what is asked (e.g. the information given), and (4) its frequency, timing and the number of repeat examinations.

Practicality including cost

The practicality of an assessment method is also important. Whether an assessment tool is practical or not may depend on the resources and expertise available and its cost. Cost is an important characteristic of an assessment instrument if it is to be adopted for widespread use in a medical school. The cost of assessment instruments is difficult to determine and wide variations have been reported in the literature (Fowell *et al.*, 2000). Calculating the cost is even more complex than estimating the expenses related to the administration of a

particular approach. Cost needs to include the start-up and continuing resources needed for development and implementation. The cost of an assessment should be considered in relation to the benefit to teaching and learning. Van der Vleuten (1996) has argued that an investment in a good assessment is also an investment in teaching and learning. What may be perceived initially as resource-intensive assessment approaches often turn out to be feasible in practice and, when the overall benefits are taken into consideration, they are deemed worth the expense.

These four factors—validity, reliability, impact on learning and practicality including cost—should be taken into account when instruments are chosen to assess the different learning outcomes.

Use of appropriate assessment tools

In the literature on assessment, numerous instruments are described which should find a place in the examiner’s tool kit of assessment methods. This guide provides only a short overview of the different methods. It is beyond the scope of the guide to review or critique each of these assessment instruments in detail. The instruments can be considered in five categories—written assessments, clinical or practical assessments, observations, portfolios and other records of performance, and peer and self-assessment. Examples in each category are given in Table 3.

Table 3. Categories of assessment instruments with examples in each category.

Assessment category	Representative instruments
Written Assessments	Essay Short Answer Questions Completion Questions Multiple Choice Questions (MCQs) Extended Matching Items (EMIs) Modified Essay Questions (MEQs) Patient Management Problems (PMPs) Progress Test Dissertation Report
Clinical/Practical Assessments	Long Cases Practical Examination Spot Examination Objective Structured Clinical Examination (OSCE) Objective Structured Practical Examination (OSPE) Objective Structured Long Examination Record (OSLER) Group Objective Structured Clinical Examination (GOSCE)
Observation	Tutor’s report Checklists Rating scales Patient report
Portfolio and Other Records of Performance	Logbooks Portfolios Procedural Logs
Peer and Self-Assessment	Peer report Self-report

J.M. Shumway & R.M. Harden

A short description of each of the approaches is provided below.

Written assessments

Description. It is highly probable that readers are familiar with written examinations and assessments. They have been the most commonly used form of assessment in medical education. Essays and short-answer questions were formerly a prominent feature of examinations. There has been a general move in the past three decades, however, to more 'objective' or multiple-choice type of questions (MCQ). MCQs may be of the 'one from five' type where a statement is followed by 'distractors' that are included to discriminate the learner's ability to recall or apply factual information. In the UK, multiple true-false items where a common stem is followed by five statements, each of which may be true or false, were used. In the extended matching type a series of four or so items have to be matched with one from an extended list of options (e.g. appropriate investigations or treatment). Despite the increased emphasis on performance testing, the multiple-choice question format continues to play a part in low- and high-stake assessment. There is a trend to set MCQs in the context of patient scenarios. Modified essay questions and patient management problems (PMPs) offered potential advantages but are less popular now because of difficulties with question setting, with marking and with standardization.

The progress test was initially developed at Maastricht University School of Medicine about 20 years ago. It is a comprehensive test in MCQ format which samples a student's knowledge across the entire undergraduate medical education. It was designed so that, because of its depth, breadth and comprehensiveness, students could not prepare for the examination by cramming and memorizing of content. It is given four times per year to students in all years of the course and consists of approximately 250 items. In each administration a newly constructed test is prepared from the large item bank. As we would expect, students in the early years perform considerably less well than more advanced students nearing graduation. In other centres, it has been argued that short constructed-response answers are a more appropriate test of core knowledge and can be marked objectively. A progress test using this form of question has been described (Friedman Ben-David *et al.*, 2001a).

'Written' assessments are not necessarily limited to paper-and-pencil administration. They can be administered using computers. This provides a number of advantages and conveniences of administration, scoring, and sampling of students' mastery of content, including adapting the test to the needs of the individual student—so-called 'adaptive testing' (Hambleton, 1996). If medical students take their licensing exams on computers, as in the USA, they should be given opportunities for practice in their undergraduate programme.

Strengths and weaknesses. Written examinations, in particular the MCQ, have traditionally enjoyed high reliability. They can be conveniently administered but are time consuming to construct properly. They have been referred to as 'objective' as examiners do not make individual judgements about the quality of a student answer. MCQs, when constructed using a

'test blueprint', can sample a large domain of knowledge in an effective and efficient manner. Parallel forms of written examinations can be developed with equivalent content, quality and psychometric characteristics (Case, 2000). An over-reliance on the multiple-choice format to measure the recall of knowledge instead of higher level learning has resulted in a measure of disenchantment with multiple-choice testing. Critics of multiple-choice-type items have suggested that because they test discrete pieces of knowledge, and because they are designed to discriminate what a student knows or does not know, the questions tend to test 'trivial' facts or superficial knowledge of a subject rather than a deep understanding. A MCQ's cuing effects has been another criticism but the research does not appear to identify this as a major problem.

Impact on learning. The use of written assessment, particularly MCQs, has a significant impact on how students study and what they learn. To put it simply, students learn to pass the test. This can have unintended negative outcomes on their training to care for patients and to be competent and caring physicians. Written examinations are at the 'heart of the hidden curriculum' (Van der Vleuten, 1996). Students will study for the test rather than learn the information as an integrated whole. Who can blame their choice of increasing their chances to do well on the test! The effect of written exams on students is unpredictable and may have unanticipated consequences not predicted (Van der Vleuten, 1996). Faculty are also influenced by the use of written tests. Instruction is designed around what needs to be learned to pass the examinations and packaged according to when the examinations will be administered.

Practicality including cost. The design and construction of MCQ examinations is costly if it is to be done properly. However, MCQs can be shared between institutions and their administration and marking on a large scale is less expensive than other assessment approaches.

Clinical or practical assessments

Description. The multiple-station examination or the objective structured clinical exam (OSCE) is an assessment approach primarily used to measure clinical competence (Harden & Gleeson, 1979). Students are assessed at a number of stations on discrete focused activities that simulate aspects of clinical competence. Each student is exposed to the same stations and assessment. Examinees move from station to station in sequence on the same time schedule. The OSCE stations are designed to measure a number of predetermined clinical competences. Standardized patients (SPs), real patients or simulators may be used in the OSCE (Collins & Harden, 1998). OSCE stations may also incorporate the assessment of interpretation, non-patient skills and technical skills.

The OSCE is more than just a 'multi-station' exam. It is in essence a clinical or practical examination in which aspects of clinical competences are sampled to determine students' clinical skills and abilities related to their competence to practise medicine. The OSCE assesses performance and is concerned with what students 'can do' rather than what students 'know' as is the case with more traditional

assessment approaches. While the term OSCE refers to the class of examinations that use stations to measure performance, other names appear in the literature such as OSPE (objective structured practical examination). The OSPRE (objective structured performance-related exam) is used as a summative assessment in the police force in the UK. All of the examinations have similar characteristics and are essentially some modified special-purpose version of the original OSCE.

Standardized patients, simulations and models were introduced in medical education as the importance of an examination of the learner's clinical abilities was recognized. Written and observation approaches were inadequate as they did not accurately assess the required skills in a clinical encounter. Viewed simply, the clinical encounter has three variables—the learner, the patients, and the examiner. To assess a learner accurately, the patient and the examiner should be held constant across all the different learners to be assessed. In the literature there is a difference between what is meant by the terms 'simulated' and 'standardized' patients. A simulated patient is an individual who simulates the role of a real patient. A standardized patient is a patient, an actor or another individual who has undergone training to provide a realistic and consistent representation of a clinical encounter and who has been trained according to specific criteria to play the role of a patient with certain conditions. The terms standardized patients and simulated patients, however, are often used interchangeably and are often referred to as SPs. Simulated or standardized patients are particularly useful to help teach medical students history and physical examination skills. In the USA and Canada approximately 80% of medical schools use SPs in their medical educational programmes (Cushing, 2000) and we now have a great deal of information on how to use SPs effectively in OSCEs (Adamo, 2003). SPs have been used in assessment primarily in two ways: in the context of formal examinations such as an OSCE and in the day-to-day practice setting to assess the learner's or physician's performance in routine practice. Gorter *et al.* (2001) described, for example, how to introduce incognito standardized patients into outpatient clinics of specialists in rheumatology.

Simulations are approximations of reality and attempt to 'simulate' as near as possible a real clinical scenario or encounter. The use of simulations in medical education with paper-and-pencil encounters, computers or actors is widespread. One area where simulation is particularly prominent is in educating and assessing learners in history and physical diagnostic skills in the early as well as the later years of the undergraduate medical curriculum. Simulation is also used frequently in the surgical and anaesthesiology areas. Simulations have been used extensively to assess competence in clinical reasoning, patient examination, patient management and the performance of procedures. Simulations such as Harvey, the cardiac simulator, are now widely used in assessment (Issenberg *et al.*, 2001). Virtual reality offers the latest opportunity to simulate patient encounters and is an approach to watch out for in the future.

Strengths and weaknesses. The OSCE offers many advantages as a practical, reliable and valid tool to assess clinical competence. Norman (2002) has suggested that "The objective structured clinical examination, with its multiple

samples of performance has come to dominate performance assessment". Generally, the greater the number of stations in the OSCE, the greater is its reliability and content validity. It has been stipulated that an OSCE with about 20 stations obtains the minimum reliability needed (ACGME and ABMS, 2000; Van der Vleuten & Swanson, 1990). Reliability is high when checklists are used with the standardized patient stations (Cushing, 2000). In a study of the use of the OSCE, the Accreditation Council for Graduate Medical Education found in paediatric certifying examinations a correlation between the OSCE and pre-certification examinations ranging between 0.59 and 0.71 (Carraccio & Englander, 2000). Given an adequate number of stations, the OSCE can be a reliable test with modest validity.

Probert *et al.* (2003) found that student performance in a traditional clinical final examination was not a good indicator of his/her rating as a junior doctor by a consultant. In contrast there was a positive relationship between student performance in a final year of OSCE with his/her rating as a junior doctor.

The advantage of using simulations for assessment purposes is that procedures which are normally difficult to assess can be assessed under conditions with high 'fidelity' without putting the learner or the patient at risk. In simulations, feedback to the learner can be given instantly so that the learner may be able to respond to that information and correct the behaviour. In Van der Vleuten and Swanson's review of the literature (1990) they found that in the use of multiple SPs very little measurement error was introduced when SPs were trained to play the same patient role. Vu *et al.* (1992) found that SPs were very good and consistent over time when recording checklist items. Reliability of SPs has been shown to be acceptable when there is adequate training and standardization (Tamblyn *et al.*, 1991).

In the assessment of students' clinical competence, each approach has its own advantages. One should not be limited to one assessment approach. Real patients, standardized or simulated patients, patient simulators or models may be the most appropriate approach depending on the particular skill being assessed and the context in which it is being assessed.

Impact on learning. Student participation in an OSCE has a positive impact on learning. One of the clear benefits is that the students' attention is focused on the acquisition of clinical skills. Students are examined in settings that better approximate the skills and competences that will be required when they practise medicine. The nature of the OSCE is to sample and simulate the examination and management of a patient. The OSCE provides formative evaluation as the student is participating in it. Participation in an OSCE was found to improve competence at subsequent stations and improve the quality of the learning experiences for examinees (Carraccio & Englander, 2000).

The OSCE also has a potentially negative impact on learning. An OSCE contains multiple stations that sample aspects of clinical medicine. The student may prepare for the examination by compartmentalizing the skills and not completely understanding the connection and flow of the skills.

The use of SPs, simulations and manikins for assessment allow practical and clinical competences to be assessed in a formal examination and focus the student's attention on these competences.

J.M. Shumway & R.M. Harden

Practicality including cost. The cost of implementing an OSCE varies greatly from centre to centre. The time and resources required to administer an OSCE are the two major costs. Estimated costs have been hard to pin down, and have been reported as ranging from US\$7 per student per station to between US\$54 and US\$496 per student to participate in an OSCE (Carraccio & Englander, 2000). Contributing to the costs are payment for standardized patients, required facilities and administrative and examiners' time. While these are charges in some centres, particularly in the USA, in others only marginal costs are incurred, standardized patients provide their services free and no charge is being made for the use of the facilities where the examination is held. Where an OSCE is mounted within an institution or medical school, costs relating to staff and facilities may be provided at no additional cost. Where the OSCE is administered in a special facility designed for the purpose, the costs may be significant. The costs of the preparation and planning of the examination may be considerable. Maintaining exam security is another potential cost.

The recruitment, training and scheduling of standardized patients may be for some centres one of the most expensive and most variable aspects contributing to the overall cost of standardized patients. The start-up and continuing costs of a standardized patient programme can seem overwhelming and often discourage institutions from investing in such an activity. Approaches vary from programme to programme. There is also a considerable investment in staff time and resources, which may include special facilities for the SPs to be examined by learners. Recruitment of SPs is another time-consuming task. It has been estimated that the time required to train SPs ranges from one to 15 hours depending on the clinical scenario being simulated. Costs also vary according to whether laypersons, actors or volunteers are used.

Observation

Description. A time-tested and well-used method of assessing students in medicine is the observation of the student by clinical faculty members during a clinical attachment. In the United States it is the predominant method used to assess students' performance during the third- and fourth-year clinical clerkships. Recently the use of observation has been gaining popularity during the first two years.

Faculty usually record their assessment of students on a checklist or rating scale. The student can be marked or rated against each of the expected learning outcomes, e.g. communication skills and attitudes. A global rating of the student's overall competence may also be recorded. For practical procedures, the faculty may simply 'check' whether the procedure was performed or not. It is especially helpful if the examiner includes narrative comments with the ratings.

Strengths and weaknesses. One of the major issues in the assessment of students by faculty is the lack of reliability of faculty ratings. This may be due to the fact that staff have had insufficient contact with the student to make an informed judgement of the student's abilities. The consistency of ratings may be improved with training of the staff. Scores may be biased because of different standards applied by untrained raters who do not want to include the ends of the rating scales in their evaluations and who do not have the ability to

discriminate between the different categories on the rating scale. Ratings for the acquisition of knowledge appear to be more reliable than ratings for interpersonal or communication skills (ACGME and ABMS, 2000).

We have described the rating of students during clinical attachments. Students' performance and abilities may also be rated by staff in other situations, such as tutorials or PBL sessions.

Impact on learning. If a student's performance on learning outcomes, not assessed with other instruments, is rated over a period of time, the student's attention is directed to these outcomes that might otherwise have been ignored. Such assessments can also be a powerful source of feedback to students. If the ratings are used for summative purposes there is a danger that the students will be encouraged to hide their weaknesses and deficiencies rather than reveal them to the teacher with a view to rectifying them.

Practicality including cost. The time and commitment needed for the evaluation of students over time by faculty should not be underestimated. The bottom line is that good assessment practice requires staff time, training and motivation.

Portfolios and other reports of performance

Description. A portfolio is a collection of material made by a professional or a learner that documents his or her achievements and includes a reflection on those achievements. Portfolio-based learning can be defined as "the collection of evidence that learning has taken place, usually set within some agreed objectives or a negotiated set of learning activities" (Snadden, 1999). They contain material collected by the learner over time and include a critical review and analysis by the student of the material. Portfolios can contain just about anything. The following are typical elements of a portfolio: critical incidents or events, a reflective journal or diary, tutorials and learning plans, clinical experiences, exam preparation materials, recorded consultations, audits and project work, critical reviews of articles, and management material (Snadden, 1999). Portfolios used for assessment purposes (both formative and summative) need to be written primarily so they can be reviewed. Issues of public and private viewing need to be made explicit and clear from the beginning. The learner should control who has access and can see her/his portfolio because of the potential personal nature of reflection that may occur. Assessment of portfolios concentrates on whether the learner has demonstrated in the portfolio that the learning outcomes have been achieved. Portfolios are very useful to document that learners have achieved the desired learning outcomes (Davis *et al.*, 2001, Friedman Ben-David *et al.*, 2001b).

Logbooks, like portfolios, document the student's experiences. However, they are usually more limited in scope than portfolios and are focused on data collected in a specific area or activity. Reflection is not normally part of logs. At least three kinds of logs have been documented—procedural, operative and case logs (ACGME and ABMS, 2000). Procedural logs usually document how many and when procedures were performed by the learner. Operative logs are similar but document what was done and when. Case logs

record which patients were seen and with what diseases, within a given time period.

Strengths and weaknesses. Portfolios are a valuable instrument for inclusion in the examiner's toolkit, if for no other reason than that they assess learning outcomes such as critical thinking and self-assessment, which are not easily assessed using other instruments. They also provide a record of the student's performance over a period of time and are not just a snapshot at one specific point in time. The reliability of portfolios is at least in part due to the ability of raters to agree on standards and criteria for the content and assessment of portfolios. Their reliability is enhanced by the triangulation of evidence from a number of sources (Friedman Ben-David *et al.*, 2001a). The validity of portfolios will be determined by the extent to which they document accurately those experiences that are indicators of the mastery of the desired learning outcomes.

Logbooks are limited in their assessment powers. Logbooks involve simply recording whether something has been performed or not. They do not document competence or quality. The literature is weak in determining whether logbooks are accurately completed by the learners (ACGME and ABMS, 2000). The reporting of validity and reliability data for logbooks is unavailable.

Impact on learning. Portfolio assessments have a positive impact on the learning environment because they document what the learner has done and ask the learner to reflect on what she/he has accomplished. They ask the learner to put together large amounts of disparate information into a whole that tells a story about the learner's accomplishments. The portfolio is more than an accumulation of what the learner has done. It is a story of the development of the learner and, seen as such, can greatly contribute to the positive nature of how a student learns and studies.

Logbooks are often considered by learners to be boring or repetitive. Yet, viewed from the preceptor's point of view, they can paint a picture of what the learner has or has not done. The learner also has an opportunity to see what he or she has accomplished by reviewing the documented list of tasks contained in the logbook.

Practicality including cost. It goes without saying that these assessment approaches require staff time and resources. The portfolio assessment has its own associated costs, especially those related to the external review and marking of the portfolios. There is cost associated with developing electronic portfolios or logbooks. These resource investments can be justified, however, in the light of the valuable insight gained into the learner's abilities and competence to practise medicine.

Peer and self-assessment

Description. Peer and self-assessment focuses on an important consideration in assessment: who should do the evaluating? Most evaluation methods are designed by faculty who 'naturally' take on the responsibility for assessing learners. However, important information can be gained if

we also ask the learner's peers and the learners themselves what they are able to do and how comfortably they are performing a task.

Peer and self-assessment may carry less weight than ratings by trained faculty examiners. It is likely, therefore, that peer and self-assessment will be used in conjunction with ratings by faculty and other trained health professionals. Checklists and rating scales can be used. Peers have the advantage of observing each other performing the tasks and procedures that are being learned under real clinical conditions (e.g. residents are more likely than the attending physician to observe their peer's performance). Peers can be more discriminating than faculty in their evaluations because of their increased exposure to the observed performance. Peer ratings have particularly been used in the area of the assessment of attitudes and communication skills among medical students. One of the concerns in the education of medical students is: 'Will the student have the ability to know when he/she doesn't know, and will he/she seek help?' In other words, the learner needs to be able to self-evaluate and take the appropriate action from that assessment. The use of self-assessment shows promise in understanding how to assess this important attribute of lifelong learning and self-discriminating abilities.

Strengths and weaknesses. Studies have shown that peer evaluations correlate highly with faculty evaluations of the rating of the same behaviours and skills (Gray, 1996). Self-assessment correlates moderately with the rating of a trained examiner. It is reported that mean ratings by self-raters tend to be lower (more critical) than those of trained examiners (Cushing, 2000). One of the difficulties in the use of peer and self-assessment approaches is the training of the raters. It has been found that the provision of benchmarks helps to improve the reliability of such ratings. A benchmark is a standard or expected level of performance with an example that helps the rater standardize his/her rating compared with those of others. Reports of the use of peer and self-assessment suggest that this approach to assessment is little used in the majority of medical education experiences. Where it has been used is mainly to assess communication skills in small groups and other similar learning settings, for example PBL (Fowell *et al.*, 2000). The use of peer and self-assessment are areas that need further study and development.

Impact on learning. The use of peer and self-assessment can have a profound impact on the students' educational programme. On the positive side, their use has the power to transform the perceived nature of an examination, especially in the clinical years, and to develop in the student skills of self-appraisal. On the negative side, if done poorly, their use could cause mistrust, suspicion and peer rivalries that would be devastating to a medical education programme. It is difficult to predict the results of a widespread use of such approaches. Which will it be: negative or positive? It is up to the institution and the climate of trust and acceptance of change to determine the final outcome.

Practicality including cost. The main cost and time commitment is the training of peers and of the learners themselves to be reliable and accurate raters.

J.M. Shumway & R.M. Harden

Assessment of the full range of learning outcomes

An overview

Miller (1990) proposed a pyramid of learning with increasing professional authenticity, starting with the learner’s cognition and moving towards a focus on the learner’s behaviour. Van der Vleuten (2002) linked a hierarchy of assessment approaches of increasing authenticity with the Miller Pyramid. This model is useful with regard to assessment in so far as different assessment instruments are appropriate at each level of the pyramid.

Figure 3 shows the Miller pyramid with each of the four levels matched with one or more of the assessment approaches described above. Written assessment is the predominant instrument at the ‘know’ and ‘know how’ levels, clinical assessment and the OSCE at the ‘shows how’ level and observation, portfolios and logbooks at the ‘does’ level.

Assessment and the learning outcomes

This section of the guide looks at assessment instruments with regard to their appropriateness as tools to measure the attainment by the students of the different learning outcomes. It describes 12 learning outcomes and provides a commentary on their assessment.

Table 4 lists for each of the 12 learning outcomes the most appropriate assessment approaches. All of the methods can be used to a greater or lesser extent to assess each of the outcomes. The methods are arranged according to the authors’ estimation of their relative importance for each outcome, with the most important or useful instrument for each learning outcome listed first.

Knowledge is embedded in each of the learning outcomes, for example knowledge about clinical theory, knowledge about management, knowledge about ethical principles and knowledge about the healthcare team and teamwork. Thus instruments used to assess knowledge may be appropriate in relation to all 12 outcomes. With regard to the technical skills of the student, as exemplified in outcomes one to seven, the assessment should include the demonstration of the relevant skill. Clinical and practical assessment methods are therefore particularly important. Attitudes, decision making and professionalism can be demonstrated most effectively through instruments designed to assess the student’s or doctor’s performance in practice in the clinical context. The

four levels of the Miller pyramid are relevant to all 12 learning outcomes. Figure 4 lists for each of the learning outcomes what is reported as the most important level of the Miller pyramid for the purposes of assessment.

The assessment of technical competences
(‘What the doctor is able to do’)

Learning Outcome 1: Competence in clinical skills. The doctor must be able to take a complete and focused history, perform an appropriate physical exam, interpret findings, and formulate an action plan to characterize the problem and reach a diagnosis.

To allow a judgement to be made with regard to a student’s competence in clinical skills such as taking a history or examining a patient’s abdomen, the student needs to demonstrate that he or she can perform the necessary clinical skill. The student must ‘show how’ he/she takes a history, examines a patient as appropriate, interprets the findings and formulates an action plan in order to reach a final diagnosis. Clinical examinations, particularly the OSCE, are appropriate for this purpose and may provide sufficient information on which the student’s abilities with regard to these outcomes can be assessed.

It may be possible, however, to obtain information about the achievement of these outcomes from reports of observers such as clinical attachment supervisors or from student entries in a diary or logbook presented as part of a portfolio. Some teachers choose to ignore such additional sources of information about the student’s clinical skills because of the potential unreliability of the data or the cost of collecting and analysing it.

Learning Outcome 2: Competence in practical procedures. The doctor should be able to undertake a range of procedures on a patient for diagnostic or therapeutic purposes. This usually involves using an instrument or some device, e.g. suturing a wound or catheterization.

Competence in some procedures can be assessed in the same way as clinical skills, with a practical examination such as an OSCE. This is not possible or desirable for all procedures. A record of the student’s performance of the required procedure in practice, certified by a member of staff, may be an important part of a student portfolio or logbook which is assessed by the examiners. Information about the student’s competence in practical procedures may also be

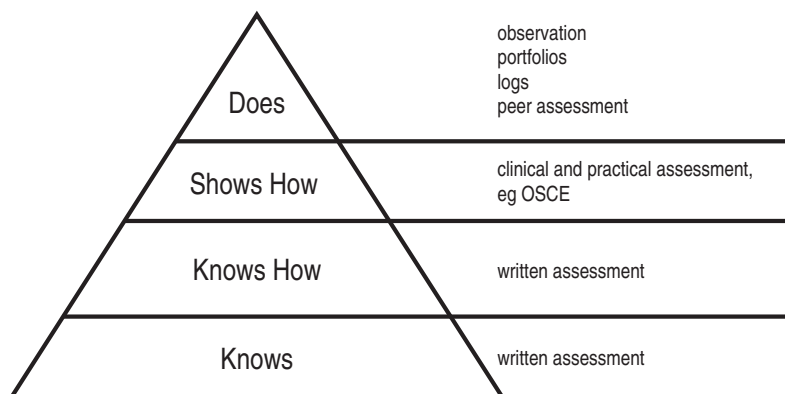


Figure 3. The learning assessment pyramid.

Table 4. Recommended assessment methods for the 12 learning outcomes of a competent and reflective physician.

Learning outcome	Assessment methods
<i>What the doctor is able to do</i>	
1 Clinical Skills	OSCE Observation Logbooks Written examination
2 Practical Procedures	OSCE Portfolios and logbooks Observation Written examination
3 Patient Investigation	Written examination OSCE Observation Portfolio
4 Patient Management	Written examination OSCE Observation Portfolios
5 Health Promotion and Disease Prevention	OSCE Portfolios Observation Written Assessment
6 Communication	OSCE Observation Peer/self assessment Portfolio
7 Information Management Skills	Portfolio OSCE Observation Written examination
<i>How doctors approach their practice</i>	
8 Principles of Social, Basic and Clinical Sciences	Written examination Portfolios Observation OSCE
9 Attitudes, Ethics and Legal Responsibilities	Observation Portfolio Peer/self assessment OSCE Written examination
10 Decision Making, Clinical Reasoning and Judgement	Portfolio Observation Written examination OSCE Peer/self assessment
<i>Doctors as professionals</i>	
11 Role as a Professional	Observation Peer/self assessment Portfolio OSCE Written examination
12 Personal Development	Portfolio Observation Peer/self assessment OSCE Written examination

obtained from staff responsible for supervising a student during an attachment.

Learning Outcome 3: Competence to investigate a patient. The doctor should be competent to arrange appropriate investigations for a patient and, where appropriate, interpret these. The investigations are carried out on the patient or on samples of fluid or tissue taken from the patient. The investigations are usually carried out by personnel trained for the purpose, e.g. a clinical biochemist or radiographer, but may in some instances be carried out by the doctor.

Written assessment instruments can be used to provide the student with the opportunity to demonstrate that he/she knows how to investigate a patient and that he/she can, where appropriate, interpret the results of patient investigation. Students may be observed doing this in an OSCE. Their competence over time can be assessed by clinical supervisors during clinical attachments and in a portfolio.

Learning Outcome 4: Competence in patient management. The doctor is competent to identify appropriate treatment for the patient and to deliver this personally or to refer the patient to the appropriate colleague for treatment. It includes interventions such as surgery and drug therapy and in contexts for care such as acute care and rehabilitation.

The examiner, using a combined approach to assessment with written instruments, OSCEs, observation and portfolios, is able to assess whether the student can successfully integrate knowledge and apply it competently to the management of a patient. Both indirect (written) and direct (OSCE) assessment approaches work together to understand the abilities of the learner.

Learning Outcome 5: Competence in health promotion and disease prevention. The doctor recognizes threats to the health of individuals or communities at risk. The doctor is able to implement, where appropriate, the basic principles of disease prevention and health promotion.

The assessment of a learner's ability to practise medicine with a population medicine perspective aimed at disease prevention and health promotion requires the application of assessment instruments in all three domains: cognitive, skill and affective. The assessment should occur in either a simulated or preferably a practice setting because of the need to assess a more complex process encompassing the broader perspective of health and how to promote it. OSCEs and portfolios are particularly useful. The knowledge underpinning health promotion and disease prevention and its application can be assessed with written examinations.

Learning Outcome 6: Competence in communication. The doctor is proficient in a range of communication skills including written and oral, both face-to-face and by telephone. He or she communicates effectively with patients, relatives of patients, the public and colleagues.

The assessment of communication skills has not in the past attracted as much attention as it deserves. It is, however, one of the competences that are critical to the care of the patient. Good care is only as good as the ability to communicate clearly and accurately with fellow health workers and other professionals, and with patients and their families. Communication skills are best assessed under 'real'

J.M. Shumway & R.M. Harden

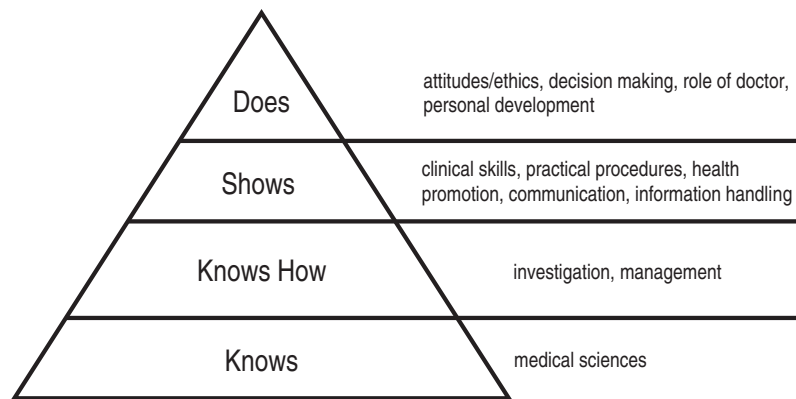


Figure 4. The 12 learning outcomes matched for assessment purposes against the most appropriate level of the Miller Pyramid.

or simulated conditions in the care of patients. The assessment of communication skills can occur continuously in the curriculum from the early years of medical school and in the student’s progress in the area recorded. The assessment of communication skills is a sensitive area because of the personal nature of learners’ communication abilities, styles, background, etc. This assessment has a profound impact on the learning environment that should not be underestimated. Careful planning has to occur in the assessment of communication skills so as not to interrupt or intrusively interfere with the patient–physician relationship. The assessment of communication skills should be high on the list of assessed competences of any medical school. A good measure of this competence can tell one much about learners’ potential success and their ability to become competent and caring physicians. OSCE’s, observation, peer assessment and portfolios have been used effectively to assess different aspects of communication skills.

Learning Outcome 7: Competence in handling and retrieval of information. The doctor is competent in retrieving, recording and analysing information using a range of methods including computers.

Students should be assessed as to their competence to retrieve, record and analyse information obtained from a range of sources including information technology resources. This learning outcome can be assessed by the student, documenting her/his experience in a portfolio or demonstrating these skills in an OSCE. The students’ competence can also be assessed over time by observation.

Assessment of intellectual, emotional, analytical and creative competences (‘How the doctor approaches his/her practice’)

Learning Outcome 8: Approach practice with an understanding of basic and clinical sciences. Doctors should understand the basic, clinical and social sciences that underpin the practice of medicine. This ensures that they not only have the technical competence outlined in outcomes 1 to 7 but that they also have an understanding of what they are doing, and why they are doing it. This includes an awareness of the psychosocial dimensions of medicine.

The assessment of an understanding of basic and clinical sciences is heavily concentrated in the cognitive domain. The classic approach to the assessment of the ‘sciences’ in medicine is through written tests. Portfolios and logbooks are also helpful in that they require students to reflect on the relationship of what they know to the application of what they do in a care setting.

Learning Outcome 9: Approach practice with appropriate attitudes, ethical stance and legal responsibilities. Doctors adopt appropriate attitudes, ethical behaviour and legal approaches to the practice of medicine. This includes issues relating to informed consent, confidentiality and the practice of medicine in a multicultural society.

The assessment of attitudes, ethical understanding and legal responsibilities is a relatively new area of assessment in medical education. It is now recognized, however, to be a key aspect of competence for the future physician. “Medical schools need to be satisfied that each of their students has reached a minimum standard of ethical competence, just as they need to be satisfied with the clinical competence of their graduates” (Wong & Cheung, 2003). Attitudes, ethical understanding and legal responsibilities can be assessed in a variety of ways. The assessment of attitudes, ethics and legal responsibilities is high on the learning pyramid and occurs best in the ‘real’ setting during care giving. It follows that these outcomes are best assessed directly over time by staff observing the student’s behaviour. Peer assessment and portfolios can be powerful assessment tools in this area. Written assessment approaches, computer simulations and OSCEs can be used as a secondary measure to assess the application of students’ attitudes, ethics and understanding of their legal responsibilities in the simulated or practice setting. The use of multiple instruments and observations based on multiple samples of behaviour is particularly important in this difficult area.

Caution needs to be exercised when students’ behaviours are assessed in practice to ensure that the assessment itself does not interfere with the student’s behaviour. Students may unknowingly alter their caring behaviour with patients because of a fear of being criticized. Caution also needs to be exercised in the use of feedback to learners in this area. Learners may be particularly sensitive to criticism.

Learning Outcome 10: Approach practice with appropriate decision making, clinical reasoning and judgement. Doctors apply clinical judgement and evidence-based medicine to their practice. They understand research and statistical methods. They can cope with uncertainty and ambiguity. Medicine requires, in some cases, instant recognition, response and unreflective action, and at other times deliberate analysis and decisions and actions following a period of reflection and deliberation. This outcome also recognizes the creative element in problem solving that can be important in medical practice.

Problem solving is a higher-order cognitive activity. Its assessment needs to occur in the setting where problem solving is applied and justified. Decision-making skills and clinical reasoning are difficult to assess because there are no generic or effective universal approaches to their assessment. This is due to several factors, the most notable of which is that clinical reasoning approaches are case dependent. In other words, different decision-making approaches are applied related to the nature of the problem. It appears that the difference between novice and expert physicians in their clinical reasoning skills is attributed to 'problem scripts' that are developed over time based on accumulated experience. The assessment of clinical reasoning on the part of the medical student can best be thought of as a reflective process with demonstrated accumulated evidence. Patient management problems (PMPs) used in the early 1970s were difficult to score and it was difficult to determine the exact 'decision path' that a learner was following. Extended matching items (EMIs) get at the knowledge aspect of clinical decision making and how students use the knowledge in this context. Portfolios can be used conveniently to assess the student's reflective practice. Students demonstrate that they can reflect on what they have accomplished, and examine and explain their problem-solving approaches on a case-by-case basis. Problem solving and decision making can also be assessed in the OSCE. An assessment approach, the Diagnostic Thinking Inventory (Bordage *et al.*, 1999), was designed specifically to assess a student's problem-solving abilities.

Assessment of personal competences ('The doctor as a professional')

Learning Outcome 11: Physicians should have an appreciation (understanding) of the doctor's role in the health service. Doctors should understand the healthcare system within which they are practising and the roles of other professionals within the system. They should appreciate the role of the doctor as physician, teacher, manager and researcher. There should be a willingness by the doctor to contribute to research even in a modest way and to build up the evidence base for medical practice. This outcome also recognizes that most doctors have some management and teaching responsibility.

Many of these competences are best assessed by observing directly or indirectly the student's or the doctor's behaviour using observations, ratings, reports or portfolios. The OSCE can assess the learning outcomes in a simulated situation. For example, a healthcare team including doctors and nurses can be assessed as a team in a group OSCE. Knowledge components can be assessed in written tests.

Learning Outcome 12: Physicians should have an aptitude for personal development. The doctor should be a self-learner and should be able to assess his or her own performance. The doctor has to take responsibility for his or her own personal and professional development, including personal health and career development. These are important learning outcomes.

Their assessment is, however, not easy. One of the most important qualities a physician can possess is the ability to judge his or her personal limits or abilities. A specific assessment instrument to determine whether a student is a self-directed and lifelong learner has not yet been devised. These skills and behaviours are best directly observed and assessed in the clinical phase of a medical student's education by a clinical supervisor. The assessment of professionalism may ultimately best be done by oneself and one's peers. Another approach which provides a glimpse into professionalism and self-directedness is the use of portfolios for assessment purposes. The portfolio provides the learner with the opportunity to demonstrate evidence of self-reflection on the accumulated evidence of her/his experiences and accomplishments during her/his medical education.

Discussion and conclusions

The development in medical education institutions of appropriate approaches to assessment underpinned by a related philosophy has lagged behind developments that have occurred in other aspects of the curriculum. The implications of this may be serious. Implementation of a new curriculum without changes to the approach to assessment may result in little or no change at all. More attention must be paid to assessment and the underlying principles (Brown *et al.*, 1996; Fowell *et al.*, 2000). Faculty and staff at each institution should spend time developing a cohesive assessment philosophy and engaging in a staff development programme. Engagement with the assessment process and ensuring that students achieve the required learning outcomes is an important role of a medical teacher (Harden & Crosby, 2000).

This guide outlines what we are looking for in an assessment instrument and describes the wide range of tools that are available to assess the required range of learning outcomes. It emphasizes the move from assessing knowledge and technical competence to more complex learning outcomes such as attitudes, teamwork and professionalism. Written assessments, OSCEs, standardized patients, simulations and models, observation, checklists and rating scales, portfolios and logs, and peer and self-assessment all deserve attention. How they fit and complement one another in the assessment of the required competences is as important as understanding the details of each.

It is unlikely that one assessment instrument can address all of the learning outcomes. In general, what is required is a focus on the construction of test blueprints to adequately sample the learning outcomes to be assessed. It is necessary to choose assessment approaches that will do the job. New assessment instruments have been developed and introduced, and can be used to assess the range of learning outcomes. More traditional instruments such as MCQs that have been used for decades have dominated the assessment process. We are likely to see in the future a move away from selected-

J.M. Shumway & R.M. Harden

response questions as in MCQs to constructed-response questions. There will also be a move to the assessment of learning outcomes which are not currently tested. Institutions need a basic package of assessment methods that will get the job done. The basic assessment package should consist of some sort of written assessment, for example, constructed-response questions (CRQs), and/or extended-matching items. The toolkit should also include a performance assessment such as the OSCE. Finally some measure of the students over time to assess other outcomes such as critical thinking and self-assessment is necessary. Portfolios can be used for this purpose. This basic package is not meant to be exhaustive, indeed institutions are encouraged to develop additional assessment strategies that meet the unique and particular needs of their own settings.

What is needed is a move to assess what learners do in practice and how they apply their knowledge of basic and clinical sciences to the care of patients. While many schools have moved to the assessment of competence in stimulated situations, only a few have placed a major emphasis on the assessment of performance in clinical practice through direct observational approaches or through the use of indirect measures such as portfolios.

Qualitative assessment approaches have also been under-used in medical education. Qualitative assessment has been associated wrongly with subjective assessment. Evidence-based literature exists that outlines the principles of good qualitative assessment (MacQueen *et al.*, 1998; Murphy *et al.*, 1998). Qualitative assessment methods provide the assessor with a sense of the environment and conditions in which the learning and its practice take place. It allows the assessor to better understand not just what students know or do not know, but what they do with that knowledge in the real settings that require the application of the learner's knowledge, attitudes and skills. Three principles apply in the use of qualitative assessment: triangulation, frequent and continuous assessment, and training of the evaluators. Triangulation means measuring the same phenomenon from different angles or vantage points using a variety of measurement strategies. In other words, assessors 'triangulate' on the phenomena being measured. Cushing

(2000) described triangulation as a means whereby data are acquired from different sources and found to be consistent so that they can corroborate validity and reliability. Multiple measures need to be employed because of the complexity of the assessment of higher-order application of knowledge, attitudes and skills in the care of patients. The second principle is that assessment should be continuous and frequent. In order to sample the learner's abilities and attributes, one cannot rely on a single measure at a single point in time. Such samples are inherently unreliable, and their validity can be questioned. As acquired knowledge is applied in the clinical setting to the actual practice of medicine, the application of that knowledge becomes more case or situation specific. In order to adequately sample the students' competence, more frequent assessment needs to occur. Assessments that occur beforehand are helpful to learners so that they can assess where they are and what they need to do to improve. Learners are developing and changing with each new learning experience. It seems only logical that a true assessment of their competence would be the latest assessment. The third principle is that assessors need to be trained. Nowhere else is this more true than in the clinical setting. In order for faculty to assess learners' abilities they need to be trained how to do it. A specific example is the need for inter-rater agreement on standards and on marking scale scores. The practice of evidence-based assessment means that the instruments used and the application of those instruments will be valid and reliable. The need for staff and faculty development as assessors underpins the ability to achieve the desired standards.

We have referred to the assessment of each learning outcome and to the choice of the most appropriate tool to be used for the purpose. While an assessment of each competence or learning outcome is necessary, an indication of a student's strength and weakness across all competences is the ultimate goal. This can be presented as a profile for a student which shows his/her levels of attainment for each of the competences. Figure 5 illustrates a hypothetical profile for a student. In the example provided, the student reaches the required standard in all of the outcomes with the exception of those relating to attitudes and professionalism. Such a profile

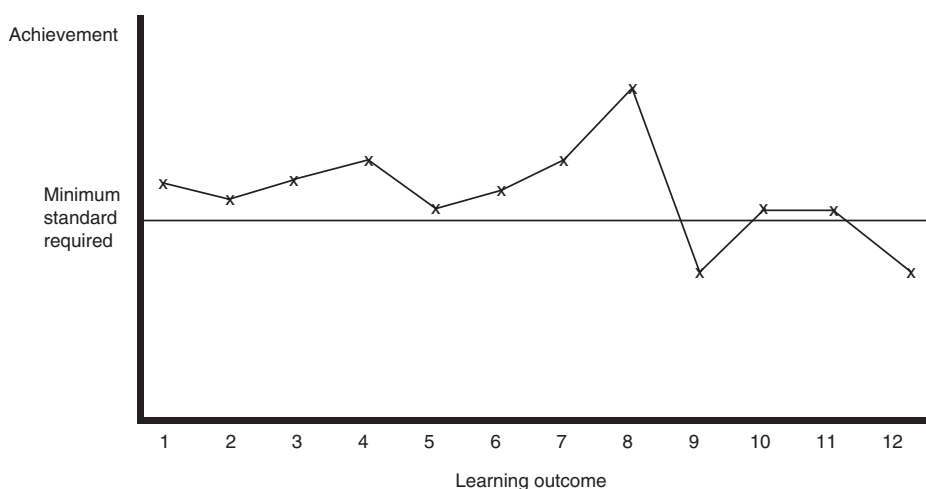


Figure 5. Profile for a student.

can prove to be very useful to the educational director, to the student and to curriculum developers.

In response to increasing public demands for a greater measure of accountability for the physicians we educate, rapid progress needs to be made in designing competence-based curricula and assessing students in increasingly realistic ways to show they can practise medicine. In this guide we demonstrate how to do a better job of assessing our students: we need an understanding of the assessment process and knowledge of the tools available and how they can be used to assess the range of learning outcomes. This model of assessment of learning outcomes needs to be applied across the different phases of medical education from undergraduate medical education to postgraduate education and continuing professional development. We need more studies of how professional competence is measured and how those measurement data are applied so that desired changes in behaviour are achieved. We have an opportunity to work across the medical education continuum to significantly improve what we do educationally, to have an impact on the future of the practice of medicine, and to guarantee to the public the competence of our physician workforce. In the end we will look back and be able to confidently educate the 'competent and reflective' physician.

Notes on contributors

JAMES M. SHUMWAY is Associate Dean for Medical Education and Professor of Internal Medicine and Pediatrics at West Virginia University School of Medicine

RONALD M. HARDEN is Education Director of the International Virtual Medical School (IVIMEDS) and former Director of the Centre for Medical Education, University of Dundee.

References

- AAMC (1998) *Report I: Learning Objectives for Medical Student Education*, Medical Schools Objectives Project, January (Washington, AAMC).
- ACGME (2003) Accreditation Council for Graduate Medical Education [website: <http://www.acgme.org>].
- ACGME and ABMS (2000) *Toolbox of Assessment Methods*, A Product of the Joint Initiative, Accreditation Council for Graduate Medical Education (ACGME) Outcomes Project, American Board of Medical Specialties (ABMS), Version 1.1, September 2000 [Copyright 2000 Accreditation Council for Graduate Medical Education and American Board of Medical Specialties].
- ADAMO, G. (2003) Simulated and standardized patients in OSCEs achievements and challenges 1992–2003 *Medical Teacher*, 5(3), pp. 262–270.
- ATKINSON, R.L., ATKINSON, R.C., SMITH, E.E. & BEM, D.J. (1993) *Introduction to Psychology*, 11th edn (New York, Harcourt Brace Jovanovich College Publishers).
- BORDAGE, G., GRANT, J. & MARSDEN, P. (1999) Quantitative assessment of diagnostic ability, *Medical Education*, 24, pp. 413–425.
- BROWN, S., RACE, P. & SMITH, B. (1996) *500 Tips on Assessment* (London, Kogan Page.)
- CanMEDS (2000) Extract from the CanMEDS 2000 Project Societal Needs Working Group Report, *Medical Teacher*, 22(6), pp. 549–554.
- CARRACCIO, C. & ENGLANDER, R. (2000) The objective structured clinical examination: a step in the direction of competency-based evaluation, *Archives of Pediatric Adolescent Medicine*, 154, pp. 736–741.
- CASE, S. (2000) Developments in written assessment, oral presentation given at the 9th International Ottawa Conference on Medical Education, Cape Town, South Africa, 28 February–3 March [<http://www.educ.unimaas.nl/ottawa/>].
- COLLINS, J.P. & HARDEN, R.M. (1998) AMEE Education Guide No. 13: The use of real patients, simulated patients and simulators in clinical examinations, *Medical Teacher* 20(6), pp. 508–521.
- CUSHING, A. (2000) Developments in attitude and professional behaviour assessment, oral presentation given at the 9th International Ottawa Conference on Medical Education, Cape Town, South Africa, 28 February–3 March [<http://www.educ.unimaas.nl/ottawa/>].
- CUSHING, A. (2002) Assessment of Non-Cognitive Factors, in: G.R. Norman (Ed.) *International Handbook of Research in Medical Education*, Ch. 22 (Dordrecht: Kluwer).
- DAVIS, M.H., FRIEDMAN, BEN-DAVID, M., HARDEN, R.M., HOWIE, P., KER, J., MCGHEE, C., PIPPARD, M.J. & SNADDEN, D. (2001) Portfolio assessment in medical students' final examinations, *Medical Teacher*, 23, pp. 357–366.
- FOWELL, S.L., MAUDSLEY, G., MAGUIRE, P., LEINSTER, S.J. & BLYTH, J. (2000) Report of findings: student assessment in undergraduate medical education in the United Kingdom 1998, *Medical Education*, 34(Suppl. 1, September), pp. 1–78.
- FRIEDMAN BEN-DAVID, M. (1999) AMEE Guide No 14: Outcome-based education, part 3: Assessment in outcome-based education, *Medical Teacher*, 21(1), pp. 23–25.
- FRIEDMAN BEN-DAVID, M., HUNTER, I. & HARDEN, R.M. (2001a) Introduction of a progress test as a tool for defining core undergraduate curriculum, *Educacion Medica XV Congreso de la Sociedad Nacional de Educacion Medica*, 4(3), pp. 97–99.
- FRIEDMAN BEN-DAVID, M., DAVIS, M.H., HARDEN, R.M., HOWIE, W., KER, J., MCGHEE, C., PIPPARD, M.J. & SNADDEN, D. (2001b) AMEE Education Guide No. 24: Portfolios as a method of student assessment, *Medical Teacher*, 23(6), pp. 535–552.
- General Medical Council (1993) *Tomorrow's Doctors: Recommendations on Undergraduate Medical Education* (London, GMC).
- General Medical Council (2002) *Recommendations on Undergraduate Medical Education* (London, GMC).
- GOETZ, E.T., ALEXANDER, P.A. & ASH, M.J. (1992) *Educational Psychology A Classroom Perspective* (New York, Maxwell Macmillan International).
- GORTER, S.L., RETHANS, J.J., SCHERPBIER, A.J.J.A., Van Der LINDEN, S. et al. (2001) How to introduce incognito standardized patients into outpatient clinics of specialists in rheumatology, *Medical Teacher*, 23(2), pp. 138–144.
- GRAY, J. (1996) Global rating scales in residency education, *Academic Medicine* 71, pp. S55–63.
- HAMBLETON, R.K. (1996). Advances in assessment models, methods, and practices, in: D.C. Berliner & R.C. Calfee (Eds) *Handbook of Educational Psychology*, pp. 899–925 (New York, Simon & Schuster Macmillan).
- HARDEN, R.M. (2002) Developments in outcome-based education, *Medical Teacher*, 24(2), pp. 117–120.
- HARDEN, R.M. & CROSBY, J.R. (2000) AMEE Education Guide No. 20: The good teacher is more than a lecturer—the twelve roles of the teacher, *Medical Teacher*, 22(4), pp. 334–347.
- HARDEN, R.M., CROSBY, J.R. & DAVIS, M.H. (1999a) An introduction to outcome-based education, *Medical Teacher*, 21(2), pp. 7–14.
- HARDEN, R.M., CROSBY, J.R., DAVIS, M.H. & FRIEDMAN, M. (1999b). From competency to meta-competency: a model for the specification of learning outcomes, *Medical Teacher*, 21(6), pp. 546–552.
- HARDEN, R.M. & GLEESON, F.A. (1979) ASME Medical Education Booklet No. 8: Assessment of clinical competence using an objective structured clinical examination (OSCE), *Medical Education*, 13, pp. 41–54.
- IIME Core Committee (2002) Global minimum essential requirements in medical education, *Medical Teacher*, 24(2), pp. 130–135.
- ISSENBERG, B., GORDON, M.S., GORDON, D.L., SAFFORD, R.E. & HART, I.R. (2001) Simulation and new learning technologies, *Medical Teacher*, 23(1), pp. 16–23.
- KASSEBAUM, D.G., EAGLEN, R.H. & CUTLER, E.R. (1997). The objectives of medical education: reflections in the accreditation looking glass, *Academic Medicine*, 72(7), pp. 648–656.
- LCME (2003) Liaison Committee on Medical Education [website: <http://www.lcme.org>].

J.M. Shumway & R.M. Harden

- MACQUEEN, K.M., McLELLAN, E., KAY, K. & MILSTEIN, B. (1998) *Codebook Development for Team-based Qualitative Analysis* (Atlanta, GA, Centers for Disease Control).
- MILLER, G.E. (1990) The assessment of clinical skills/competence/performance, *Academic Medicine*, 65(Suppl.), pp. S63–S67.
- MURPHY, E., DINGWALL, R., GREATBATCH, D., PARKER, S. & WATSON, P. (1998) Qualitative research methods in health technology assessment: a review of the literature, *Health Technology Assessment*, 2(16).
- NORMAN, G. (2002) Research in medical education: three decades of progress, *British Medical Journal*, 324, pp. 1560–1562.
- PRESSLEY, M. & McCORMICK, C.B. (1995) *Advanced Educational Psychology for Educators, Researchers, and Policymakers* (New York, HarperCollins College Publishers).
- PROBERT, C.S., CAHILL, D.J., McCANN, G.L. & Ben-SHLOMO, Y. (2003) Traditional finals and OSCEs in predicting consultant and self-reported clinical skills of PRHOs: a pilot study, *Medical Education*, 37, pp. 597–602.
- RUBIN, P. & Franchi-CHRISTOPHER, D. (2002) New Edition of Tomorrow's Doctors. *Medical Teacher*, 24(4), pp. 368–369.
- Scottish Deans Medical Curriculum Group (2002) The Scottish Doctor—Undergraduate Learning outcomes and their assessment: a foundation for competent and reflective practitioners [<http://www.scottishdoctor.org>].
- SIMPSON, J.G., FURNACE, J., CROSBY, J., CUMMING, A.D. *et al.* (2002) The Scottish doctor—learning outcomes for the medical undergraduate in Scotland: a foundation for competent and reflective practitioners, *Medical Teacher*, 24(2), pp. 136–143.
- SMITH, S.R. & DOLLASE, R. (1999) Planning, implementing and evaluating a competency-based curriculum, *Medical Teacher*, 21(1), pp. 15–22.
- SMITH, S.R. (1999) An Educational Blueprint for the Brown University School of Medicine, institutional publication, Brown University.
- SNADDEN, D. (1999) AMEE Education Guide No. 11 (revised): Portfolio-based learning and assessment in medical education, *Medical Teacher*, 4, pp. 370–386.
- TAMBLYN, R.M., KLASS, D.J., SCHNABL, G.K. & KOPELOW, M.L. (1991) Sources of unreliability and bias in standardized-patient rating, *Teaching and Learning in Medicine*, 3(2), pp. 74–85.
- VAN DER VLEUTEN, C.P.M. (1996) The assessment of professional competence: developments, research and practical implications, *Advances in Health Sciences Education*, 1(1), pp. 41–67.
- VAN DER VLEUTEN, C.P.M. (2000) A paradigm shift in education: how to proceed with assessment, oral presentation given at the 9th International Ottawa Conference on Medical Education, Cape Town, South Africa, 28 February–3 March [<http://www.educ.unimaas.nl/ottawa/>].
- VAN DER VLEUTEN, C.P.M. & SWANSON, D. (1990) Assessment of clinical skills with standardized patients: state of the art, *Teaching and Learning in Medicine*, 2, pp. 58–76.
- VU, N.V., MARCY, M.M., COLLIVER, J.A., VERHULST, S.J., TRAVIS, T.A. & BARROWS, H.S. (1992) Standardized (simulated) patients' accuracy in recording clinical performance checklist items, *Medical Education*, 16, pp. 99–104.
- WINCKEL, C.P., REZNICK, R., COHEN, R. & TAYLOR, B. (1994) Reliability and construct validity of a structured technical skills assessment form, *American Journal of Surgery*, 167, 423–427.
- WONG, J.G.W.S. & CHEUNG, E.P.T. (2003) Ethics assessment in medical students, *Medical Teacher*, 25(1), pp. 5–8.



A standardized, structured long-case examination of clinical competence of senior medical students

LUIZ E.A. TRONCON, ROBERTO O. DANTAS, JOSÉ FERNANDO C. FIGUEIREDO, EDUARDO FERRIOLLI, JÚLIO C. MORIGUTI, ANA L.C. MARTINELLI & JÚLIO C. VOLTARELLI

Departamento de Clínica Médica, Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Brazil

SUMMARY *Aiming at improving the assessment of senior medical students a standardized and structured modification to the traditional long-case examination is proposed. Students are presented with a sequence of two long cases, with each case being observed by a different examiner. After approaching the patient, the student is interviewed by the examiner who uses a set of four standardized questions. A 10-item checklist covering five groups of major clinical skills was delineated. The time frames for both patient approach (25 minutes) and student interview (10 minutes) were standardized. After the examination, immediate feedback is given to the student. Preliminary results obtained with the first 27 students taking the exam showed that none has failed and 20 of them had combined marks higher than 8.0 (range: 6.6–10.0). Failing performances on individual checklist items were more common in physical examination and less frequent regarding communication with patients. Agreement between different examiners on the assessment of the various skills of the same student was 89%. It is concluded that modifying the format to the long-case examination increased its value in the assessment of student clinical competence.*

Introduction

Assessment of clinical competence of undergraduate medical students plays a key role in their education (Lowry, 1993). However, the application of valid, reliable and feasible examination methods represents a continuing challenge for medical educators (Newble, 1992).

Clerkship in Medicine is an important step in the acquisition of knowledge and clinical skills (McLeod & Harden, 1985). In Brazilian medical schools, assessment of clerks is usually based on the perceptions of supervisors, as there is no tradition of performing objective examination of clinical competence (Troncon *et al.*, 1994).

At the Faculty of Medicine of Ribeirão Preto (University of São Paulo, Brazil), recent changes in the undergraduate curriculum leading to modifications in clerkship configuration provided an opportunity to revise the assessment of senior students. Faculty members in charge of organizing the clerkship in Internal Medicine agreed that a more objective, patient-centered examination was needed for a more valid and reliable summative assessment of the clerks.

The Department of Clinical Medicine has a small amount of experience in running an OSCE (Objective Structured Clinical Examination) (Harden & Gleeson, 1979) for the assessment of basic clinical skills of junior medical students (Troncon *et al.*, 1996). Nevertheless, it was felt that an

examination with a strict OSCE format would not provide comprehensive information on the clinical competence that interns should exhibit in the patient encounter.

We therefore proposed and implemented an examination method containing some elements of both the 'observed long case' (Newble, 1991) and Gleeson's OSLE (Objective Structured Long Examination Record) (Gleeson, 1994, 1977), and running in an OSCE format (Harden & Gleeson, 1979). A main feature of the proposed examination is presenting the student with a sequence of two long cases, each observed by a separate examiner.

In this paper, we describe this method and present the preliminary results of the assessment of a small group of clerks as well as the estimated rate of agreement between different examiners.

Methods

Settings

At the Faculty of Medicine of Ribeirão Preto, students aged 17–19 years enter the medical school shortly after finishing the second cycle of basic education. The current medical school curriculum comprises 2 years of integrated basic sciences, one semester (3rd year) of pre-clinical disciplines and 3 semesters (3rd and 4th years) of clinical disciplines. The two final years (5th and 6th) are spent mainly in five clerkships of Internal Medicine, Pediatrics, Gynecology & Obstetrics, Surgery & Orthopedics and Social & Community Medicine. Each clerkship period lasts 8 weeks per year. During the Internal Medicine periods, students typically engage in intensive ward work in the mornings and in outpatient visits in the afternoons. While in the wards, clerks have considerable responsibility for the management of inpatients. Overnight calls in the Emergency room (1–2 weekly) complete the program.

Examination design

The proposed examination consisted of a sequence of two long cases entirely observed and organized in a multiple-station format. The time frames for both student–patient encounter and examiner–student interview are standardized. Students have 25 minutes with the first patient (focused

Correspondence: Dr Luiz E.A. Troncon, Departamento de Clínica Médica, Hospital das Clínicas - Campus da USP, 14048-900, Ribeirão Preto, SP, Brazil. Tel: +55 16 602 2457; fax: +55 16 633 6695; email: ledatron@fmrp.usp.br

history taking and physical examination), after which 2 minutes are allowed for data organization and preparation for the interview with the examiner. This takes the next 8 minutes and is structured over four standardized questions aimed at assessing clinical acumen, clinical reasoning and patient management. The interview is followed by the provision of immediate feedback to the student on his/her performance. The student then moves to the next station, where the same sequence is repeated with another patient and examiner.

Performance assessment

Student performance on both patient approach and interview is assessed and graded using a 10-item checklist covering five groups of clinical skills: (a) history-taking; (b) physical examination; (c) communication and interaction with the patient; (d) case presentation and clinical reasoning and (e) patient management (Figure 1). The last four checklist items are covered during the interview, which is carried out over a set of four standard questions (Figure 2). The checklist also allows for the evaluation of case difficulty during actual student assessment. Since sources of difficulty may relate to different domains, the checklist contains grades (low, average and high) for three aspects: patient ability to communicate, physical examination findings and clinical problem as a whole.

Grading and marking

For each of the 10 checklist items, a four-grade marking system is applied: (a) fully adequate (mark 1.0); (b) barely adequate (mark 0.75); (c) inadequate (mark 0.25) and (d) poor (mark 0). An intermediate grade corresponding to a mark 0.5 was deliberately omitted in order to avoid borderline evaluations. A detailed description of student behaviors to be associated with each grade for all items was also constructed. For example a 'fully adequate' grade on 'characterization of the main complaint' should be attributed to students who 'obtained information on type, intensity, chronology and modifying circumstances of the chief complaint presented by the patient, as well as on the influence of the present disease on patient's ordinary activities and general well being'.

Preparation for the examination

The preparation for the examination involved a number of meetings between faculty members and student representatives aimed at reaching consensus on the main points of the examination. Case characteristics and degree of difficulty, as well as the range of specific contents and clinical problems to be included in the examination, were carefully discussed. Those faculty who would probably act as examiners received detailed written instructions for the entire procedure and were put in charge of selecting and recruiting real patients to be included in the exam. The criteria for selection included stable clinical conditions, absence of communication problems, presence of at least one well-defined diagnostic abnormality detectable by a standard physical examination maneuver, well-defined diagnosis fitting into the previously elaborated list of contents and problems and an average degree of overall difficulty. All recruited patients had also to give their informed consent.

Running the examination

The actual run-up to the examination took place in a special ward unit containing six rooms located around a living room. Each room was arranged as an office and accommodated one patient. For each run of the examination, a pool of 9–12 patients was selected, which was sufficient for the patients to alternate between working with students (25 min) and resting (15–50 min). Each selected patient was approached by students from two to four times during the examination day. Each station was staffed by one member of faculty, who was fully aware of the main findings and clinical problems of the patients alternating in that particular station. Immediately before the examination, both examiners and students were briefed on the main points of the exam. Clerks were examined in groups of six. Since each run took 40 minutes and every student was examined in two consecutive stations, the time needed for the assessment of up to 18 interns was 4 hours.

Data analysis

In the first two runs of the proposed examination a group of 27 students was examined. For every student, the marks given by each examiner to the various checklist items were added to form a 'total mark', which could range from 0 to 10. The average of the total marks attributed by the two examiners comprised the 'combined mark'. In order to comply with local regulations, students had to obtain a combined mark of at least 5.0 to 'pass' the examination.

For the assessment of overall student performance on the various checklist items, any of the two 'adequate' grades were regarded as a 'pass' event and any 'inadequate' or 'poor' grade corresponded to a 'fail' event. The frequency of 'fail' events was calculated as a proportion of 540, the total number of grades attributed (27 students times 10 items times 2 examiners).

Inter-examiner variability was estimated by determining the rates of 'full agreement', 'partial agreement' and 'disagreement' concerning the 10 checklist items, as a proportion of the total number of decisions taken. A 'full agreement' was recorded when both examiners attributed an identical grade for a given item. A 'partial agreement' was recorded when both grades, despite being different, were in the same 'pass' or 'fail' ranges. A 'disagreement' was recorded when grades fell in different ranges ('pass' or 'fail'). Considering that the limited numbers of students examined would preclude an appropriate statistical analysis, no attempt was made to calculate reliability coefficients more accurately.

Results

All 27 students examined carried out the expected tasks and all the 54 checklist forms were thoroughly fulfilled by examiners. Also, all the selected patients complied with the proposed schedule and no incidents or drop-outs were recorded in any of the two runs of the examination.

Student performance

Among the first 27 interns examined, none failed. Combined marks ranged from 6.6 to 10.0 (median: 8.5) and were equal or higher than 8.0 for 20 students.

Practice-based assessment of clerks in internal medicine				
Student name:.....				
Examiner name:.....				
Date:...../...../.....				
Checklist				
History taking	Grading(*)			
1. Characterization of main complaint.....	(FA)	(BA)	(I)	(P)
2. Characterization of associated manifestations and other personal and familial data.....	(FA)	(BA)	(I)	(P)
Physical examination				
3. Approach to the most likely affected organ or system (technique and thoroughness).....	(FA)	(BA)	(I)	(P)
4. Routine general examination.....	(FA)	(BA)	(I)	(P)
Communication and interaction with patient				
5. Technical aspects of medical interview.....	(FA)	(BA)	(I)	(P)
6. Personal respect and consideration.....	(FA)	(BA)	(I)	(P)
Case presentation and clinical reasoning				
7. Summary of main points of the case.....	(FA)	(BA)	(I)	(P)
8. Diagnostic hypothesis.....	(FA)	(BA)	(I)	(P)
Patient management				
9. Indication of diagnostic measures.....	(FA)	(BA)	(I)	(P)
10. Indication of therapeutic measures.....	(FA)	(BA)	(I)	(P)
(*) FA = fully adequate (mark 1.0); BA = barely adequate (0.75); I = inadequate (0.25); P = poor (0.0).				
Observations				
.....				
.....				
Case difficulty	Grading (#)			
A. Patient communication	(H)	(A)	(L)	
B. Physical findings	(H)	(A)	(L)	
C. Clinical problem	(H)	(A)	(L)	
(#) H = high; A = average; L = low.				

Figure 1. Assessment form for the proposed examination of clinical competence.

'Fail' performances on any of the checklist items were recorded in only 36 (6.6%) out of the 540 grades attributed. Table 1 shows that 26 of these events were recorded for the groups of skills directly related to patient approach (physical examination and history taking) and only 10 'fail' performances were detected for the items involving cogni-

tive performance (clinical reasoning and patient management).

Case characteristics

The two runs of the proposed examination involved 18 patients. Data regarding case difficulty are presented in

Practice-based assessment of clerks in internal medicine	
Standardized questions for clerk interview	
1.	Would you please make a summary of the main points regarding this case that you gather from the clinical history and the physical examination ?
2.	Would you please describe briefly what are your main diagnostic hypotheses, ranked from the most to the least probable?
3.	Taking into account your findings, would you please mention briefly what tests you would need to confirm your diagnostic hypothesis?
4.	Would you please describe briefly what kind of general and specific therapeutic measures you would take in order to both solve the patient's problem and improve his/her well-being?

Figure 2. Standardized questions for the assessment of clerks' competences on case presentation, clinical reasoning and patient management.

Table 1. Number of student 'fail' performances on the standardized, structured long-case examination.

Clinical skills group	'Fail' performances (%)
History taking	9
Physical examination	15
Communication with patient	2
Case presentation and clinical reasoning	4
Patient management	6

Table 2, which shows that only a small proportion of cases (1.8%) were regarded as having difficulties in communicating with students. Physical examination was considered as having low or average difficulty in 92.6% of cases. As expected from the requirements for inclusion in the examination, most of the cases included (87%) had average or low difficulty concerning the nature of the clinical problem presented.

Agreement between examiners

Data for the rates of agreement between examiners concerning student performance throughout the 10 checklist items are presented in Table 3. In 89% of the decisions taken, there was full or partial agreement between the two examiners and the rate of disagreement was quite low (11.0%).

Discussion

In many medical schools throughout the world, assessment of clinical competence of senior students is carried out by

Table 2. Degree of difficulty of clinical cases included in the observed long-case examination.

Category	Degree of difficulty (%)		
	High	Average	Low
Patient communication	1.8*	9.2	89.2
Physical findings	7.4	18.5	74.1
Clinical problem	13.0	50.0	37.0

Note: *Percentage of total number of decisions taken.

Table 3. Examiners' agreement on student clinical competence.

Category	% of grades attributed
Full agreement*	43.0
Partial agreement [#]	46.0
Disagreement ^φ	11.0

Notes: *both examiners attribute identical grades; [#] examiners attribute different grades but in the same range ('pass' or 'fail'); ^φ examiners attribute grades in different ranges.

the traditional long-case examination (Stokes, 1974; Weatherhall, 1991; Newble, 1992; Lowry, 1993), which has been criticized for its unreliability. Nevertheless, the place of the long-case examination in the assessment of students seems to be justifiable not only for its tradition and practicality, but also because this method has an adequate degree of face validity (Newble, 1991; 1992; Gleeson, 1997).

Modifications of the traditional long case (Newble, 1992; Price & Byrne, 1994; Abouna & Hamdy, 1999), including the OSLER (Gleeson, 1994, 1997) have improved the method, but it still has a number of shortcomings. For example, in the 'observed long case' (Newble, 1991) the checklist seems to be too generic to guarantee that individual components of competence are properly assessed. On the other hand, in Gleeson's OSLER (Gleeson, 1997), students are observed only partially and the examination seems to put much more emphasis on the assessment of case presentation and the ability of the student to communicate with the examiner, rather than on basic clinical skills.

In the long-case examination herein described, the student performance is entirely observed and the focus of the assessment is directed to genuine clinical skills. In fact, six out of the 10 items comprising the checklist are devoted to patient-centered skills. Moreover, we have used real patients presenting with clinical features fitting into a pre-defined set of contents and clinical problems. Therefore, the actual observation of the student performing relevant clinical tasks, complemented by the use of standard questions for the assessment of the cognitive aspects involved, implies coverage of the most important components of medical competence (Newble, 1992). Thus, our proposed examination is in line with recent work aimed at improving student

Med Teach Downloaded from informahealthcare.com by FACULDADE DE ODONTOLOGIA - USP For personal use only.

L.E.A. Troncon et al.

evaluation using a more standardized performance-based exam (Abouna & Hamdy, 1999).

Like others (Newble, 1991), we also proposed a sequence of two long cases and a more rigid time control, which were implemented in a two-station OSCE format (Harden & Gleeson, 1979), so as to ensure that all clerks were assessed in similar conditions. While exposing the candidate to two different cases increased the sample of contents (Newble, 1991; Van der Vleuten, 1996), the use of more than one observer is likely to have improved exam accuracy (Van der Vleuten, 1996; Gleeson, 1997; Abouna & Hamdy, 1999).

Unlike others (Newble, 1991, Price & Byrne, 1994, Gleeson, 1994, 1997; Abouna & Handy, 1999), we have not used two simultaneous examiners per case, which was done for a number of practical reasons. The number of examiners available in our department is quite small and students unaccustomed to being observed pointed out that the presence of two examiners would probably make the exam more stressful. Also, faculty expressed concerns about the efficacy of the 'confer and agree' process between concomitant examiners (Gleeson, 1997) on 10 checklist items within a controlled time frame. Also taken into consideration was the opinion of international experts, who think that presenting the candidate with two cases, each one with a different single examiner, is likely to improve reliability of the long-case examination (Van der Vleuten, 1996). The low rate of disagreement between examiners (11%) that we found in our long-case examination is consistent with this view.

The use of one examiner per station also seemed to increase practicality, by means of a more convenient usage of the human resources available. Indeed, the organizational difficulties of running this examination were felt to be lower than those of a conventional OSCE for assessing basic clinical skills of junior medical students (Troncon *et al.*, 1996). The examination of an entire group of interns demanded six examiners, for one morning's work, which was perceived as being reasonable in terms of our manpower utilization.

Our preliminary results showed that not a single intern failed and the combined marks were rather higher than one could predict. A possible explanation for this could be that examiners were reluctant to attribute grades in the 'fail' range, since nearly all examiners were clerkship supervisors and involve intensively with students in their daily work on the wards. A possible way to overcome this limitation would be inviting external examiners (Walters *et al.*, 1995), which would probably increase exam accuracy. Nevertheless, this would probably work against the feasibility of the proposed examination.

As far as the overall student performance on individual items is concerned, we found a relatively high frequency of deficiencies regarding physical examination, which cannot be explained by difficulties expressed by the selected cases (Table 2). This is consistent with data reported elsewhere (Gleeson, 1994), which has been interpreted as a possible consequence of a gradual decline in the time available for teaching clinical skills in the undergraduate curriculum (Gleeson, 1994).

Provision of feedback to students immediately after the examination has been recognized as a substantial improvement in the educational value of current methods of assess-

ment of clinical competence (Newble, 1992; Lowry, 1993; Gleeson, 1994, 1997; Van der Vleuten, 1996). Although we have not attempted to evaluate the response of students to the proposed examination, many clerks expressed informally their satisfaction concerning this particular point of the exam. We also believe that the involvement of students in the delineation of the assessment has contributed to both increase student compliance with the implementation of the exam and improve its educational meaning.

Acknowledgement

This work received financial support from the following Brazilian agencies: FAEPA/HCFMRP, CAPES and CNPq.

Notes on contributors

LUIZ E.A. TRONCON, MD, is an associate professor in the Department of Clinical Medicine and former Chairman of the Committee of Undergraduate Studies, Ribeirão Preto Faculty of Medicine, University of São Paulo, Brazil.

ROBERTO O. DANTAS, MD, is an associate professor in the Department of Clinical Medicine and former Coordinator of the Internship in Internal Medicine, Ribeirão Preto Faculty of Medicine, University of São Paulo, Brazil.

JOSE FERNANDO C. FIGUEIREDO, MD, is an assistant professor in the Department of Clinical Medicine, Ribeirão Preto Faculty of Medicine, University of São Paulo, Brazil.

EDUARDO FERRIOLLI, MD, is an assistant physician and clerk supervisor in the Department of Clinical Medicine, Ribeirão Preto Faculty of Medicine, University of São Paulo, Brazil.

JULIO C. MORIGUTI, MD, is an assistant physician and clerk supervisor in the Department of Clinical Medicine, Ribeirão Preto Faculty of Medicine, University of São Paulo, Brazil.

ANA L. C. MARTINELLI, MD, is an assistant professor in the Department of Clinical Medicine, Ribeirão Preto Faculty of Medicine, University of São Paulo, Brazil.

JULIO C. VOLTARELLI, MD, is an associate professor in the Department of Clinical Medicine, Ribeirão Preto Faculty of Medicine, University of São Paulo, Brazil.

References

- ABOUNA, G.M. & HAMDY, H. (1999) The integrated direct observation clinical encounter examination (IDOCCE)—an objective assessment of students' clinical competence in a problem-based learning curriculum. *Medical Teacher*, 21, pp. 67–72.
- GLEESON, F. (1994) The effect of immediate feedback on clinical skills using the OSLE, in: A.I. ROTHMAN & R. COHEN (Eds) *Proceedings of the Sixth Ottawa Conference on Medical Education*, pp. 412–415 (Toronto, University of Toronto Bookstore Custom Publishing).
- GLEESON, F. (1997) *AMEE Education Guide No. 9: Assessment of Clinical Competence using the Objective Structured Long Examination Record (OSLER)* (Dundee, Association for Medical Education in Europe).
- HARDEN, R.M. & GLEESON, F.A. (1979) Assessment of clinical competence using an objective structured clinical examination (OSCE), *Medical Education*, 13, pp. 39–45.
- LOWRY, S. (1993) Assessment of students, *British Medical Journal*, 306, pp. 51–54.
- MCLEOD, P.J. & HARDEN, R.M. (1985) Clinical teaching strategies for clinicians, *Medical Teacher*, 7, pp. 173–189.
- NEWBLE, D.I. (1991) The observed long case in clinical assessment, *Medical Education*, 25, pp. 369–373.
- NEWBLE, D.I. (1992) Assessing clinical competence at the undergraduate level, *Medical Education*, 26, pp. 504–511.

Examination of clinical competence

- PRICE, J. & BYRNE, G.J.A. (1994) The direct clinical examination: an alternative method for the assessment of clinical psychiatry skills in undergraduate medical students, *Medical Education*, 28, pp. 120–125.
- STOKES, J. (1974). *ASME Medical Education Booklet No.2: The Clinical Examination: Assessment of Clinical Skills* (Dundee, Association for the Study of Medical Education).
- TRONCON, L.E.A., FOSS, N.T., VOLTARELLI, J.C. & DANTAS, R.O. (1996) Avaliação de habilidades clinicas por exame objetivo estruturado por estações, com emprego de paciente padronizados: uma aplicação no Brasil (Parte II), *Revista Brasileira de Educação Médica*, 20, pp. 53–60.
- TRONCON, L.E.A., RODRIGUES, M.L.V., PICCINATO, C.E. & CIANFLONE, A.R.L. (1994) Responses of students and faculty members to changes in assessment in a Brazilian medical school, in: A.I. ROTHMAN & R. COHEN (Eds) *Proceedings of the Sixth Ottawa Conference on Medical Education*, pp. 126–131 (Toronto, University of Toronto Bookstore Custom Publishing).
- VAN DER VLEUTEN, C. (1996) Making the best of the long case, *Lancet*, 347, pp. 704–705.
- WALTERS, W.A., SIVANESARATNAM, V. & HAMILTON, J.D. (1995) External examiners, *Lancet*, 345, pp. 1093–1095.
- WEATHERHALL, D.J. (1991) Examining undergraduate examinations, *Lancet*, 338, pp. 37–39.

New staff workshop

Nov 2019

EPA and WPBA

เชิดศักดิ์ ไอรรมณีรัตน์
ศูนย์ความเป็นเลิศด้านการศึกษาวิทยาศาสตร์สุขภาพ
คณะแพทยศาสตร์ศิริราชพยาบาล

Entrustable Professional Activities (EPA)

รศ.นพ.เชิดศักดิ์ ไอรรมณีรัตน์
ภาควิชาศัลยศาสตร์
คณะแพทยศาสตร์ศิริราชพยาบาล

Outline

- EPA
 - Definitions: EPA, competencies, milestones
 - Key concepts
 - Assessment in EPA framework

EPA

- Entrustable Professional Activity
 - A unit of professional practice, defined as tasks or responsibilities that trainees are entrusted to perform unsupervised once they have attained sufficient specific competence

AAMC. Core entrustable professional activities for entering residency: Faculty and learners' guide, Washington DC, 2014.

Competency

- Competency: An observable ability of a professional, integrating multiple components such as knowledge, skills, values, and attitudes

AAMC. Core entrustable professional activities for entering residency: Faculty and learners' guide, Washington DC, 2014.

Key Concepts

- EPAs are not an alternative for competencies, but a means to translate competencies into clinical practice.
- Competencies are descriptors of physicians.
- EPAs are descriptors of work.
- An EPA usually requires multiple competencies in an integrative, holistic nature.

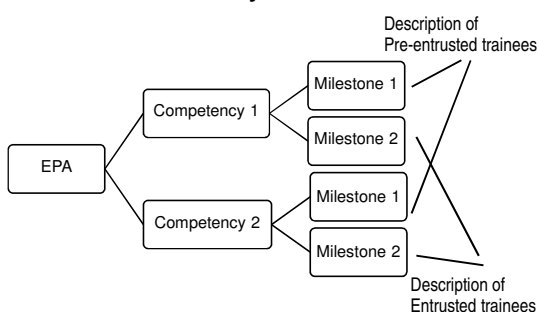
EPA and Competencies

EPA	Med knowledge	Patient care	Interpersonal skills	Professionalism	Practice-based learning	Systems-based practice
Performing appendectomy	x	x				
Executing a patient handover	x	x	x			x
Designing therapy protocol	x				x	
Chairing multidisciplinary meeting		x	x	x		x
Request organ donation			x	x		
Manage CRF		x	x	x		x

Milestone

- Stages in the development of specific competencies
- Milestones may link to a supervisor’s EPA decisions

How They Related?



Assessing Trainees

1. Observation but no execution
2. Execution with direct, proactive supervision
3. Execution with reactive supervision (i.e., on request)
4. Supervision at a distance
5. Supervision provided by the trainee to more junior colleagues

Cate OT. Nuts and bolts of entrustable professional activities. JGME 2013.

Summary

- EPA
 - Definitions: EPA, competencies, milestones
 - Key concepts
 - Assessment in EPA framework

Workplace-Based Assessment

เชิดศักดิ์ ไชยมณีรัตน์
 ภาควิชาศัลยศาสตร์ คณะแพทยศาสตร์ศิริราชพยาบาล
 มหาวิทยาลัย มหิดล

New staff workshop

Nov 2019

Outline

- WPBA
 - Characteristics
 - Examples: Mini-CEX, DOPS, CbD, PBA
 - Implementation

Workplace-based Assessment

- A number of assessment methods, suitable for providing feedback based on observation of trainee performance in the workplace.
 - Mini-clinical Evaluation Exercise (mini-CEX)
 - Clinical Encounter Card (CEC)
 - Blinded Patient Encounter (BPE)
 - Direct Observation of Procedural Skills (DOPS)
 - Procedure-based Assessment (PBA)
 - Case-based Discussion (CbD)
 - Multisource Feedback (MSF)

WPBA: Characteristics

- เป็นการประเมินที่ให้ผู้เรียนเป็นผู้เริ่มต้น
- ผู้เรียนสามารถขอให้อาจารย์ประเมินได้ตลอดช่วงเวลาปฏิบัติงาน
- เป็นการประเมินในสถานที่ปฏิบัติงานจริง
- ประเมินซ้ำได้ หากคิดคะแนนจะใช้คะแนนครั้งที่ดีที่สุด
- จุดมุ่งหมายสำคัญคือการเปิดโอกาสให้อาจารย์ได้ **feedback**

WPBA: Strengths

- Validity: assessment of “does” level
- Identify students in needs of support early
- Provide feedback
- Create a nurturing culture
- Samples widely in many workplaces
- Utilize a number of assessors

General Medical Council. Workplace based assessment: A guide for implementation, April 2010.

WPBA: Limitations

- Low reliability
- Can be opportunistic
- Trainees may delay or avoid assessment
- Learner dependent and vulnerable
- Require time and training
- Bias due to the interaction between trainers and trainees

General Medical Council. Workplace based assessment: A guide for implementation, April 2010.

Mini-Clinical Evaluation Exercise

- นักศึกษาแสดงการ **approach** ผู้ป่วยจริงในคลินิกหรือหอผู้ป่วย ขณะที่ได้รับการสังเกตการณ์โดยอาจารย์
 - Focused history taking
 - Focused physical examination
 - Making clinical diagnosis
 - Develop a management plan
- ใช้เวลาในการ **approach** ผู้ป่วย 15 นาทีต่อราย ตามด้วยการ **ให้ feedback** จากอาจารย์อีก 5 นาที
- อาจารย์ให้คะแนนแต่ละทักษะด้วย **rating scale 1-9**

Cherdsakramaneerat@gmail.com

3

New staff workshop

Nov 2019

Direct Observation of Procedural Skills (DOPS)

- ประเมินทักษะการทำหัตถการในขณะทำงานกับผู้ป่วยจริง
- อาจารย์สังเกตขั้นตอนการทำหัตถการแล้วให้คะแนนด้วย **rating scale 1-6** คะแนนในแต่ละมิติของการประเมิน
- แต่ละหัตถการทำการประเมินโดยอาจารย์หลายท่าน ในหลายบริบท
- แต่ละหัตถการใช้เวลาสังเกต **15 นาที** และ **feedback 5 นาที**
- ตัวอย่างหัตถการ: endotracheal intubation, nasogastric tube insertion, IV injection, arterial blood sampling, etc.

Case-based Discussion (CbD)

- นักศึกษาเลือกผู้ป่วย **2** รายที่ตนเคยดูแลนำเสนอให้อาจารย์
- อาจารย์ผู้ประเมินเลือก **1** ใน **2** ผู้ป่วยนั้นเพื่อทำการอภิปรายรายละเอียดของผู้ป่วย
 - Clinical assessment
 - Investigations
 - Treatment
 - Follow-up and future plan
- วัดดูประสงค์เพื่อประเมิน **clinical reasoning skills**
- การอภิปรายผู้ป่วยแต่ละรายใช้เวลาไม่เกิน **20 นาที** และมี **feedback 5 นาที**

Procedure-Based Assessment (PBA)

- A form of workplace-based assessment
- An assessor completes the form based on observation of a trainee performs a surgical procedure
- Six domains: consent, pre-operative planning, exposure and closure, intraoperative technique, postoperative management
- Two groups of items: general items, task-specific items

Marriott J et al. Evaluation of procedure-based assessment for assessing trainees' skills in the operating theatre. BJS 2011; 98: 450-7.

Our Experience with WPBA

- Department of Surgery
 - Start with undergraduate learners: DOPS
 - Postgraduate learners: PBA

A Journey

- Choose wisely using a consensus
- A working group
- Multiple rounds of presentation
- Plenty of time for revision
- Allow for subjective judgment
- Instruction for teachers and learners
- Availability of assistance
- Evolving the system

23

Summary

- Workplace-based assessment (WPBA)
 - Characteristics
 - Formative: Increase feedback
 - Strengths and weaknesses
 - Examples: Mini-CEX, DOPS, CbD, PBA
 - Implementation: Siriraj Hospital's experience
 - Undergrad: DOPS
 - Postgrad: PBA

RIP OUT

Nuts and Bolts of Entrustable Professional Activities

OLLE TEN CATE, PHD

The Challenge

The entrustable professional activity (EPA) concept allows faculty to make competency-based decisions on the level of supervision required by trainees. Competency-based education targets standardized levels of proficiency to guarantee that all learners have a sufficient level of proficiency at the completion of training.¹⁻⁶ Collectively, the competencies (ACGME or CanMEDS) constitute a framework that describes the qualities of professionals. Such a framework provides generalized descriptions to guide learners, their supervisors, and institutions in teaching and assessment. However, these frameworks must translate to the world of medical practice. EPAs were conceived to facilitate this translation, addressing the concern that competency frameworks would otherwise be too theoretical to be useful for training and assessment in daily practice.

What Is Known

Trust is a central concept for safe and effective health care. Patients must trust their physicians, and health care providers must trust each other in a highly interdependent health care system. In teaching settings, supervisors decide when and for what tasks they entrust trainees to assume clinical responsibilities. Building on this concept, EPAs are units of professional practice, defined as tasks or responsibilities to be entrusted to the unsupervised execution by a trainee once he or she has attained sufficient specific competence. EPAs are independently executable, observable, and measurable in their process and outcome, and therefore, suitable for entrustment decisions. Sequencing EPAs of increasing difficulty, risk, or sophistication can serve as a backbone for graduate medical education.⁶

How Do EPAs Differ From Competencies?

- EPAs are not an alternative for competencies, but a means to translate competencies into clinical practice.
- Competencies are descriptors of physicians, EPAs are descriptors of work.
- EPAs usually require multiple competencies in an integrative, holistic nature. TABLE 1 shows how different EPAs require proficiency in several competency domains.

Olle ten Cate, PhD, is Professor of Medical Education and Director of the Center for Research & Development of Education at the University Medical Center Utrecht, the Netherlands.

Corresponding author: Th J (Olle) ten Cate, PhD, PO Box 85500, 3508 GA Utrecht, the Netherlands, t.j.tencate@umcutrecht.nl

DOI: <http://dx.doi.org/10.4300/JGME-D-12-00380.1>

What Is Included in a Full EPA Description?

An EPA must be described at a sufficient level of detail to set trainee expectations and guide supervisor's assessment and entrustment decisions (see TABLE 2 for guidelines).

How Do EPAs Relate to Milestones?

Milestones, as defined by the ACGME, are stages in the development of specific competencies. Milestones may link to a supervisor's EPA decisions (eg, direct proactive supervision versus distant supervision). The Pediatrics Milestone Project provides examples of how milestones can be linked to entrustment decisions.^{7,8}

What Do Entrustment Decisions Require?

Entrustment decisions involve clinical skills and abilities as well as more general facets of competence, such as understanding one's own limitations and knowing when to ask for help. Making entrustment decisions for unsupervised practice requires observed proficiency, usually on multiple occasions.

In practice, entrustment decisions are affected by 4 groups of variables: (1) attributes of the trainee (tired, confident, level of training); (2) attributes of the supervisors (eg, lenient or strict); (3) context (eg, time of the day, facilities available); and (4) the nature of the EPA (rare, complex versus common, easy). Entrustment decisions can be further distinguished as ad hoc (eg, happening during a night shift) or structural (establishing the recognition that a trainee may do this activity at a specific level of supervision from now on). In the clinical context, many ad hoc entrustment decisions happen every day. Structural entrustment decisions formally acknowledge that a trainee has passed a threshold that allows for decreased supervision. The certificate awarded at such occasions has been called a *statement of awarded responsibility* (STAR) and should be carefully documented.²

Linking an EPA with a competency framework emphasizes essential competency domains when observing a trainee executing the EPA.

How You Can Start TODAY

Decide how many EPAs are useful for training.

While there can be many EPAs that serve to make ad hoc entrustment decisions, EPAs that lead to structural entrustment decisions (ie, certification or STARs) should involve broad-based responsibilities and be limited in number. For a graduate medical education program, no more than 20 to 30 EPAs are recommended.

RIP OUT

TABLE 1 EXAMPLES OF EPAs RELATED TO THEIR MOST IMPORTANT ACGME COMPETENCY DOMAINS

Illustrative EPAs	ACGME Competencies					
	MK	PC	ISC	P	PBLI	SBP
Performing an appendectomy	•	•				
Executing a patient handover	•	•	•			•
Designing a therapy protocol	•				•	
Chairing a multidisciplinary meeting		•	•	•		•
Requesting organ donation			•	•		
Chronic disease management		•	•	•		•

Abbreviation: EPAs, entrustable professional activities; ACGME, Accreditation Council for Graduate Medical Education; MK, Medical Knowledge; PC, Patient Care; ISC, Interpersonal Skills and Communication; P, Professionalism; PBLI, Practice-Based Learning and Improvement and SBP, Systems-Based Practice.

Use of EPAs in Assessing Trainees

EPAs can be the focus of assessment. The key question is: Can we trust this trainee to execute this EPA? The answer may be translated to 5 levels of supervision for the EPA:

1. Observation but no execution, even with direct supervision
2. Execution with direct, proactive supervision
3. Execution with reactive supervision, ie, on request and quickly available

4. Supervision at a distance and/or post hoc
5. Supervision provided by the trainee to more junior colleagues

What You Can Do LONG TERM

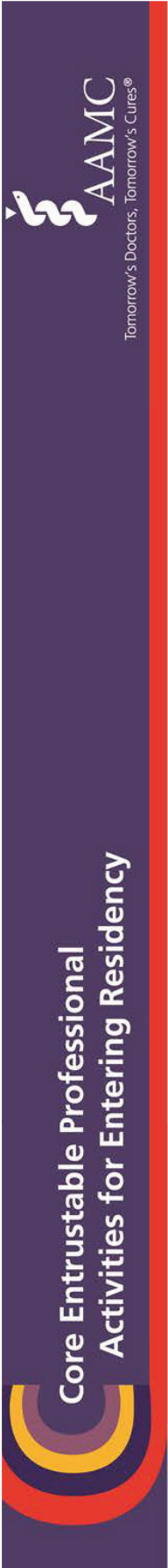
- Review the specialty requirements and milestones, and work with your professional organization and local colleagues to identify EPAs.
- Detail the EPAs, following TABLE 2.
- Prepare faculty to provide EPA-based assessments.
- Use structural entrustment decisions as a “license” for trainees to execute EPAs with distant supervision.

Resources

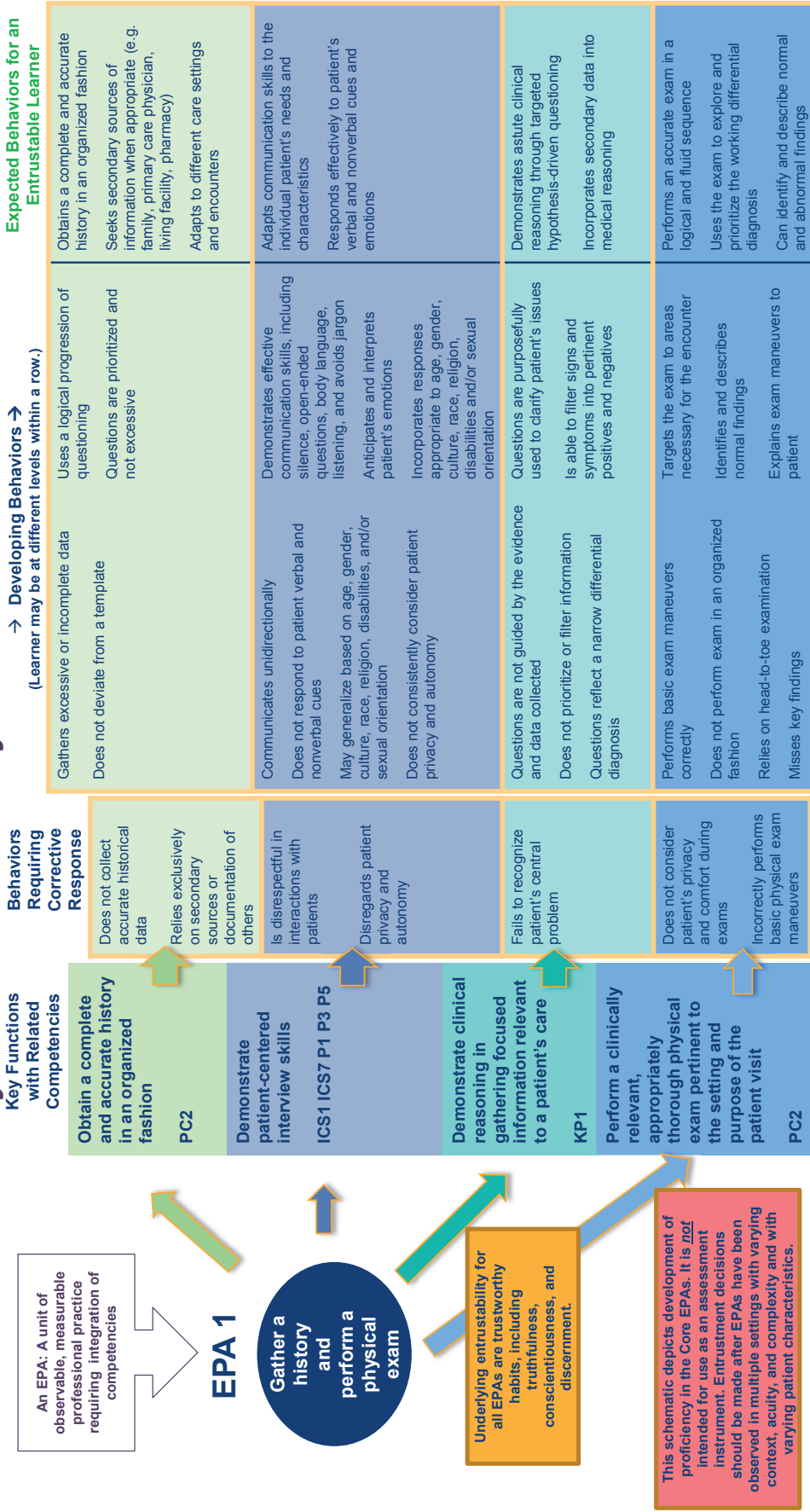
- 1 ten Cate O. Entrustability of professional activities and competency-based training. *Med Educ.* 2005;39(12):1176–1177.
- 2 ten Cate O, Scheele F. Competency-based postgraduate training: can we bridge the gap between theory and clinical practice? *Acad Med.* 2007;82(6):542–547.
- 3 Mulder H, ten Cate O, Daalder R, Berkvens J. Building a competency-based workplace curriculum around entrustable professional activities: the case of physician assistant training. *Med Teach.* 2010;32(10):e453–e459.
- 4 ten Cate O, Young JQ. The patient handover as an entrustable professional activity: adding meaning in teaching and practice. *BMJ Qual Saf.* 2012. 2012 21: i9–i12. doi: 10.1136/bmjqs-2012-001213.
- 5 Chang A, Bowen JL, Buranosky RA, Frankel RM, Ghosh N, Rosenblum MJ, et al. Transforming primary care training-patient-centered medical home entrustable professional activities for internal medicine residents [published online ahead of print September 21, 2012]. *J Gen Int Med.* DOI: 10.1007/s11606-012-2193-3
- 6 Nasca TJ. The Next Accreditation System, June 2012. <http://www.acgme-nas.org/assets/pdf/Nasca%20NAS%20June%202012%20Presentation%20Slide%20Show.pdf>. Accessed October 21, 2012.
- 7 Hicks PJ, Schumacher DJ, Benson BJ, Burke AE, Englander R, Guralnick S, et al. The pediatrics milestones: conceptual framework, guiding principles, and approach to development. *J Grad Med Educ.* 2010;2(3):410–418.
- 8 Pediatrics Milestone Project. http://www.acgme.org/acgmeweb/Portals/o/PFAAssets/ProgramResources/320_PedsMilestonesProject.pdf. Accessed October 14, 2012.

TABLE 2 GUIDELINES FOR FULL ENTRUSTABLE PROFESSIONAL ACTIVITIES DESCRIPTIONS

1. Title	Make it short; avoid words related to proficiency or skill. Ask yourself: Can a trainee be scheduled to do this? Can an entrustment decision for unsupervised practice for this EPA be made and documented?
2. Description	To enhance universal clarity, include everything necessary to specify the following: What is included? What limitations apply? Limit the description to the actual activity. Avoid justifications of why the EPA is important, or references to knowledge and skills.
3. Required Knowledge, Skills, and Attitudes (KSAs)	Which competency domains apply? Which subcompetencies apply? Include only the most relevant ones. These links may serve to build observation and assessment methods.
4. Required KSAs	Which KSAs are necessary to execute the EPA? Formulate this in a way to set expectations. Refer to resources that reflect necessary or helpful standards (books, a skills course, etc).
5. Information to assess progress	Consider observations, products, monitoring of knowledge and skill, multisource feedback.
6. When is unsupervised practice expected?	Estimate when full entrustment for unsupervised practice is expected, acknowledging the flexible nature of this. Expectations of entrustment moments can shape an individual workplace curriculum.
7. Basis for formal entrustment decisions	How many times must the EPA be executed proficiently for unsupervised practice? Who will judge this? What does formal entrustment look like (documented, publicly announced)?



EPA 1: Gather a History and Perform a Physical Examination



Barron, B., Orlander, P., Schwartz, M.L., Obeso V., Brown D., Phillipi C., eds.; for Core EPAs for Entering Residency Pilot Program. Adapted from the Association of American Medical Colleges (AAMC). Core entrustable professional activities for entering residency, 2014.

การประเมินผลในบริบทของการทำงาน (Workplace-based assessment)

รศ.นพ.เชิดศักดิ์ ไอรณณรัตน์

ภาควิชาศัลยศาสตร์

การประเมินผลมีบทบาทในการส่งเสริมการเรียนรู้ของนักศึกษาได้จากหลายกลไก ซึ่งกลไกหนึ่งที่มีความสำคัญมากและอาจารย์แพทย์ควรมีการใช้มากขึ้นคือการใช้การประเมินผลในระหว่างเรียน (formative assessment) เพื่อให้ได้ข้อมูลว่านักศึกษามีระดับความรู้ ความสามารถมากน้อยเพียงใด ยังต้องพัฒนาในด้านใดบ้าง แล้วนำข้อมูลดังกล่าวให้แก่นักศึกษา (feedback) เพื่อให้ นักศึกษาได้พัฒนาตนเองให้ดีขึ้นก่อนที่จะถูกประเมินผลในตอนสิ้นสุดการเรียน (summative assessment) งานวิจัยหลายชิ้นแสดงให้เห็นว่านักศึกษาแพทย์และแพทย์ประจำบ้านได้รับการสังเกตและประเมินผลการทำงานในลักษณะของ formative assessment นี้ไม่เพียงพอ ซึ่งส่งผลให้นักศึกษาและแพทย์ประจำบ้านเหล่านี้ขาดโอกาสที่จะพัฒนาความรู้ และทักษะพื้นฐานในการดูแลผู้ป่วย อย่างมีประสิทธิภาพ

การประเมินผลในบริบทของการทำงาน (Workplace-based assessment) เป็นกลุ่มของวิธีการประเมินผลที่ถูกพัฒนาขึ้น เพื่อให้อาจารย์แพทย์ได้มีโอกาสประเมินความรู้ และทักษะต่างๆทางคลินิกของนักศึกษาในขณะที่ทำงานกับผู้ป่วยจริง และได้นำผลการประเมินนั้นมาชี้แนะแนวทางในการพัฒนาความรู้ และทักษะของนักศึกษา วิธีการประเมินผลในกลุ่มนี้มีลักษณะสำคัญต่างๆ ดังนี้

1. เป็นการประเมินผลที่ให้นักศึกษาเป็นผู้เริ่มต้น กล่าวคือ นักศึกษาไม่ต้องรอถึงวันที่กำหนดโดยอาจารย์ว่าจะทำการสอบในวันใด เมื่อไรที่นักศึกษาพบโอกาสเหมาะได้พบผู้ป่วยที่ตนสามารถแสดงระดับความรู้ และทักษะทางคลินิกของตนให้ อาจารย์แพทย์ประเมินได้ก็ทำการขอให้อาจารย์ช่วยประเมินผลได้ทันที
 2. นักศึกษาสามารถขอให้อาจารย์ประเมินได้ตลอดช่วงเวลาที่ทำการปฏิบัติงานในภาควิชา ไม่จำเป็นต้องรอถึงสิ้นสุดการปฏิบัติงาน การเปิดโอกาสให้ประเมินได้ตั้งแต่เริ่มปฏิบัติงานทำให้อาจารย์สามารถเห็นจุดบกพร่องของนักศึกษาแต่แรกใน ขณะที่ยังมีเวลาให้นักศึกษาได้ฝึกฝนพัฒนาตนเอง
 3. นักศึกษาสามารถขอรับการประเมินทักษะเดิมซ้ำได้จนกว่านักศึกษาจะมีทักษะดังกล่าวดีเป็นที่น่าพอใจ โดยคะแนนที่จะนำไปตัดสินผลการศึกษาคือคะแนนครั้งที่นักศึกษาทำได้ดีที่สุด
 4. จุดมุ่งหมายสำคัญของการประเมินผลนี้คือการเปิดโอกาสให้อาจารย์ได้สังเกตนักศึกษาปฏิบัติงานกับผู้ป่วยจริงแล้วให้ ข้อมูลย้อนกลับ (feedback) แก่นักศึกษา
 5. นักศึกษาเป็นผู้รับผิดชอบในการดำเนินการรวบรวมคะแนนของตนเอง และตรวจสอบว่าตนยังต้องทำการประเมินทักษะใด อีกบ้าง
- วิธีการประเมินผลกลุ่มนี้มีด้วยกันหลายวิธี เช่น
1. Mini-clinical evaluation exercise (Mini-CEX) เป็นการประเมินทักษะการตรวจรักษาผู้ป่วยที่แผนกผู้ป่วยนอก หรือในหอผู้ป่วยโดยอาจารย์แพทย์ให้เวลานักศึกษาซักประวัติ ตรวจร่างกายผู้ป่วยราว 15 นาทีแล้วจึงทำการอภิปรายและทำการให้ คะแนนร่วมกับให้ feedback แก่นักศึกษา

2. Direct Observation of Procedural Skills (DOPS) เป็นการประเมินทักษะการทำหัตถการพื้นฐาน โดยอาจารย์สังเกตนักศึกษาทำหัตถการดังกล่าวกับผู้ป่วย ซึ่งมักเป็นหัตถการที่ใช้เวลาทำไม่นานนัก เสร็จแล้วอาจารย์ให้คะแนนและให้ feedback แก่นักศึกษา
3. Case-based discussion (CbD) เป็นการประเมินทักษะการคิดวิเคราะห์แก้ปัญหาทางคลินิกของนักศึกษา โดยนักศึกษาเลือกผู้ป่วยสองรายที่ตนเคยดูแลเสนอให้อาจารย์ผู้ประเมิน อาจารย์ทำการเลือกหนึ่งในสองรายของผู้ป่วยให้นักศึกษานำเสนอประวัติและแนวทางในการตรวจวินิจฉัย และการรักษา ร่วมกับการที่อาจารย์ซักถามเพื่อประเมินความเข้าใจของผู้ป่วยของนักศึกษา เมื่ออภิปรายผู้ป่วยดังกล่าวเสร็จแล้ว อาจารย์ให้คะแนนและให้ feedback แก่นักศึกษา
4. Multisource feedback (MSF) เป็นการเก็บรวบรวมใบประเมินการปฏิบัติงานที่กรอกโดยบุคลากรที่ทำงานร่วมกับนักศึกษาที่หลากหลาย ได้แก่ อาจารย์ พยาบาล แพทย์ประจำบ้าน เพื่อนักศึกษา แล้วนำผลการประเมินที่รวบรวมได้มาสรุปแล้วนำไปให้ feedback แก่นักศึกษาให้เห็นข้อมูลในภาพรวมว่าการทำงานของเขาในสายตาของเพื่อนร่วมงานนั้นมีประสิทธิภาพเพียงใด มีสิ่งใดที่นักศึกษาควรปรับปรุงบ้าง

หากอาจารย์มีความต้องการจะนำการประเมินผลในบริบทของการทำงานมาใช้มีข้อเสนอแนะแนวปฏิบัติบางประการเพื่อช่วยให้การประเมินผลนี้บรรลุตามวัตถุประสงค์

1. ต้องทำการชี้แจงให้ทั้งนักศึกษาและอาจารย์ที่เกี่ยวข้องในการประเมินผลทุกท่านเข้าใจถึงวัตถุประสงค์ของการประเมินนี้ที่ชัดเจนว่ามุ่งเน้นที่ formative assessment
2. ต้องทำการสร้างบรรยากาศที่เอื้อให้เกิดการประเมินผลในรูปแบบนี้ โดยทำให้ทุกคนเข้าใจว่าเป็นการประเมินเพื่อให้เกิดการพัฒนาในตัวนักศึกษา ดังนั้นไม่จำเป็นต้องคาดหวังว่าจะต้องได้คะแนนสูงในการประเมินทุกครั้ง และต้องสร้างกระบวนการช่วยกระตุ้นให้นักศึกษาล้ำที่จะเชิญชวนให้อาจารย์ทำการประเมินนักศึกษาเมื่อมีโอกาส และให้นักศึกษาสามารถเข้าหาอาจารย์ได้หลากหลายท่าน
3. ต้องมีการกำหนดบทบาทที่ชัดเจนของอาจารย์ว่าในการประเมินรูปแบบนี้หน้าที่ของอาจารย์ไม่เพียงแต่ให้คะแนนตามใบประเมินแล้ว อาจารย์ยังต้องมีหน้าที่สอน และให้ข้อมูลย้อนกลับ (feedback) แก่นักศึกษาด้วย
4. ต้องสร้างความเข้าใจในกลุ่มนักศึกษาว่าตัวนักศึกษาเองมีความรับผิดชอบในการหาโอกาสที่จะรับการประเมินด้วยตนเอง สนใจรับฟัง feedback จากอาจารย์เพื่อนำไปปรับปรุงตนเองให้ดีขึ้น และตรวจสอบว่าตนเองจะต้องประเมินทักษะใดบ้าง
5. การดำเนินการให้ประสบความสำเร็จต้องอาศัยความช่วยเหลือจากเจ้าหน้าที่สายสนับสนุนทางการศึกษาในการรวบรวมข้อมูลการประเมินผล และตรวจสอบว่ามีปัญหาใดเกิดขึ้นในกระบวนการประเมินผลหรือไม่ เช่น จำนวนผู้ป่วยที่เหมาะสมสำหรับการประเมินนักศึกษามีเพียงพอหรือไม่ มีอาจารย์ท่านใดไม่มีส่วนร่วมในการประเมินนักศึกษาหรือไม่ มีอุปสรรคใดทำให้นักศึกษาไม่สามารถรับการประเมินได้หรือไม่ เป็นต้น

หากอาจารย์สามารถนำรูปแบบการประเมินผลในบริบทของการทำงานไปประยุกต์ใช้ในการพัฒนาการเรียนการสอนในภาควิชาต่างๆได้ น่าจะทำให้เพิ่ม constructive feedback ให้แก่นักศึกษาและก่อให้เกิดการพัฒนาความรู้ และทักษะพื้นฐานทางการแพทย์ในนักศึกษาแพทย์ได้อย่างมีประสิทธิภาพมากขึ้น

staff workshop

Nov 2019

Performance Assessment

เชิดศักดิ์ ไอรมนรัตน์
ภาควิชาศัลยศาสตร์ คณะแพทยศาสตร์ศิริราชพยาบาล
มหาวิทยาลัย มหิดล

Competence and Performance

- Competence = The capacity of a person to perform a defined task (Maximal ability)
- Performance = The actual act in carrying out or execute the duty (Typical ability)

**An ounce of performance is
Worth pounds of promises.**

Mae West

Clinical Performance Ratings

Ratings of clinical performance based on observing real-life clinical practice by attending faculty members

Objectives

- เมื่อสิ้นสุดการอบรมแล้ว อาจารย์ผู้เข้าอบรมสามารถ
 - ร่วมในการประเมินนักศึกษาแพทย์ด้วยแบบประเมิน clinical performance ratings ได้อย่างเหมาะสม
 - พัฒนาแบบประเมิน clinical performance ratings ที่มีคุณภาพดี ซึ่งนำไปสู่การประเมินที่ถูกต้อง และเที่ยงตรง

Rater Errors

- Construct-irrelevance variance in performance ratings that is associated with raters' behavior, not with the actual performance of rates
- Valid use of clinical performance assessment requires monitoring and controlling of rater errors.

Myford, C. M., & Wolfe, E. W. (2003). Detecting and measuring rater effects using many-facet Rasch measurement: Part I. Journal of Applied Measurement, 4, 386-422.

dsakiraneerat@gmail.com

1

New staff workshop

Nov 2019

Reducing Rater Errors

- Improving raters
- Improving a rating instrument

Improving Raters

1. Rater training
2. Rater monitoring
3. Rater feedback

Writing Effective Items

- Remember your purpose
- Keep it simple
- Focused: include only one topic per item
- Start with easy-to-respond items
- Group items into sections, position these sections in a logical order

Characteristics of A Good Scale

1. Well-defined category
2. Appropriate number of categories
3. Proper handling of middle category
4. Ordered
5. Research-based

Key Points: Performance Ratings

- Remember what to observe
- Rate when you still remember the students
- Multiple ratings: multiple raters, time points
- Rate when you are in a stable emotional state
- Be consistent in your rating standards (within and across groups)
- Rate each item independently: avoid halo effect
- Use the full range of scores: avoid restriction of range

If my future were determined just by my performance on a standardized test, I would not be here. I guarantee you that.

Michelle Obama

แบบประเมินการปฏิบัติงาน นักศึกษาแพทย์ชั้นปีที่ 6

นศพ. ชื่นบาน แซ่มชื่นใจ

รหัสประจำตัวนักศึกษา 6001999

ให้อาจารย์ทำเครื่องหมายกากบาท (X) ในระดับคะแนนที่เหมาะสม

ข้อ	10 (ดีมาก)	9	8	7	6 (พอใช้)	5	4	3	2	1 (ไม่ดี อย่างยิ่ง)
1. ความรู้พื้นฐานทางการแพทย์										
2. ทักษะการซักประวัติ และตรวจร่างกาย										
3. ทักษะการวินิจฉัยโรค										
4. ความสามารถในการเลือกการตรวจค้น เพิ่มเติมที่เหมาะสม										
5. การเลือกการรักษาที่เหมาะสม										
6. การคิดอย่างมีเหตุผล										
7. การสื่อสาร										
8. มนุษยสัมพันธ์										
9. บุคลิกภาพ และการแต่งกาย										
10. ความเป็นวิชาชีพแพทย์										

**แบบประเมินการปฏิบัติงานของนักศึกษาแพทย์ปี 6
คณะแพทยศาสตร์ศิริราชพยาบาล**

นศ. พ.
ฝึกปฏิบัติงานที่
หอผู้ป่วย

รหัส
ภาควิชา/แผนก
ช่วงเวลาปฏิบัติงาน

กลุ่ม
ถึง

หัวข้อการประเมิน	%	ดีเกิน (10)	ดี (8-9)	ผ่าน (6-7)	ไม่ผ่าน (<6)	หมายเหตุ	
1. ความรู้		มีความรู้พื้นฐานที่สำคัญอย่างดีและสามารถนำมาประยุกต์ใช้ในการดูแลผู้ป่วยเป็นอย่างดี	มีความรู้พื้นฐานที่สำคัญอย่างดีแต่นำมาประยุกต์ใช้ในการดูแลผู้ป่วยได้นิดนึ้	มีความรู้พื้นฐานที่สำคัญแต่ไม่สามารถนำมาประยุกต์ใช้ในการดูแลผู้ป่วย	ขาดความรู้พื้นฐานที่สำคัญ		
2. ทักษะ		รวบรวมข้อมูลปัญหาได้สมบูรณ์ คิดวิเคราะห์แก้ปัญหาได้ด้วยตนเองดี	รวบรวมข้อมูลปัญหาได้สมบูรณ์ คิดวิเคราะห์แก้ปัญหาได้ด้วยตนเอง	รวบรวมข้อมูลปัญหาได้สมบูรณ์ แต่ต้องชี้แนะการคิดวิเคราะห์แก้ปัญหา	การรวบรวมข้อมูลปัญหาและการคิดวิเคราะห์แก้ปัญหาบกพร่อง		
2.1 การแก้ปัญหาทางคลินิก		เลือกการสืบค้นและการรักษาได้ถูกต้อง สามารถบอกเหตุผล และคำวินิจฉัยพร้อมด้วย	เลือกการสืบค้นและการรักษาได้ถูกต้อง สามารถบอกเหตุผล แต่ยังไม่สามารถอธิบายถึงปัจจัยทางด้าน	เลือกการสืบค้นและการรักษาได้ถูกต้อง แต่ไม่สามารถบอกเหตุผลได้ชัดเจน	ไม่สามารถเลือกการสืบค้นและการรักษาได้ถูกต้อง		
2.2 ความสามารถในการดูแลผู้ป่วยและการตัดสินใจ		มีข้อมูลสำคัญครบถ้วน เป็นระเบียบ อ่านง่าย ลงลายมือชื่อและรหัส	มีข้อมูลสำคัญครบถ้วน แต่ไม่เป็นระเบียบ อ่านยาก หรือ ไม่ลงลายมือชื่อ/รหัส	ขาดข้อมูลสำคัญบางอย่าง เช่น ประวัติส่วนตัวและสังคม ประวัติยา progress note, procedure/surgical note, etc.	ขาดข้อมูลที่สำคัญหลายอย่าง ไม่เขียน progress note		
2.3 การบันทึกเวชระเบียน		ทำหัตถการที่สำคัญได้อย่างแต่ละครั้ง ชัดเจนการที่ถูกต้องมีความชำนาญในการใช้เครื่องมือ และติดตามดูแลผู้ป่วยหลังทำหัตถการอย่างเหมาะสม	สามารถทำหัตถการที่สำคัญได้แต่ไม่แต่ละครั้ง อาจต้องการความช่วยเหลือในบางขั้นตอน มีการติดตามดูแลผู้ป่วยหลังทำหัตถการอย่างเหมาะสม	สามารถทำหัตถการที่สำคัญได้แต่ต้องการความช่วยเหลือค่อนข้างมาก หรือขาดการติดตามดูแลผู้ป่วยหลังผ่าตัด	ไม่สามารถทำหัตถการที่สำคัญได้ แม้จะได้รับการชี้แนะแล้ว ไม่รู้ขั้นตอนการทำหัตถการ และ/หรือขาดทักษะพื้นฐานในการทำหัตถการ		
2.4 การทำหัตถการ		เป็นขั้นตอนดีมาก เข้าใจง่าย	เป็นขั้นตอนดี เข้าใจ โดยอาจต้องการเพิ่มเติมเล็กน้อย	ไม่เป็นขั้นตอน ต้องการเพิ่มเติมค่อนข้างมากจึงจะเข้าใจ	ล้มเหลวมาก ก็เรียนรู้ไม่มีความเข้าใจในเรื่องที่นำเสนอ		
2.5 ทักษะการนำเสนอ		ดีมาก ผู้ป่วยและญาติพึงพอใจมาก	ดี ผู้ป่วยและญาติเข้าใจโรคที่เป็น	ผู้ป่วยและญาติบางคนไม่เข้าใจโรค	สร้างความสับสนให้แก่ผู้ป่วยและญาติ		
2.6 การสื่อสารกับผู้ป่วย/ญาติ							
3. ความเป็นวิชาชีพแพทย์							
3.1 ความสามารถในการเรียนรู้ด้วยตนเอง		แสดงถึงความใฝ่รู้ ค้นคว้าเพิ่มเติมได้ด้วยตนเอง	แสดงถึงความใฝ่รู้ ค้นคว้าเพิ่มเติมได้โดยต้องชี้แนะวิธีการ	ต้องกระตุ้นและชี้แนะวิธีการจึงจะค้นคว้าเพิ่มเติม	ขาดความใฝ่รู้ แม้จะได้รับการกระตุ้นและชี้แนะ		
3.2 การวางตัวที่เหมาะสม		ตรงต่อเวลา บุคลิกภาพ ความประพฤติ การแต่งกายถูกระเบียบทุกกาลเทศะ	ตรงต่อเวลา บุคลิกภาพ ความประพฤติ การแต่งกายถูกระเบียบเป็นส่วนหนึ่ง	ไม่ตรงต่อเวลา บุคลิกภาพ การแต่งกายถูกระเบียบเป็นส่วนใหญ่	มีพฤติกรรมที่ไม่เหมาะสมและไม่ปฏิบัติตามหลังจากได้รับคำตักเตือน		
3.3 ความรับผิดชอบ		รับผิดชอบดีมาก หรือ ได้รับความไว้วางใจในการดูแลผู้ป่วยและการอยู่เวรอย่างมาก	รับผิดชอบดีในการดูแลผู้ป่วยและการอยู่เวร	ไม่มีความรับผิดชอบเรื่องความรับผิดชอบในการดูแลผู้ป่วยและการอยู่เวร	ไม่รับผิดชอบ หรือ ปล่อยให้ผู้ป่วยหรือเรียนในการดูแลผู้ป่วยและการอยู่เวร		
3.4 เจตคติและจริยธรรม		ดูแลผู้ป่วยทั้งร่างกายและจิตใจอย่างดี เคารพสิทธิของผู้ป่วย	ดูแลผู้ป่วยทั้งร่างกายและจิตใจ เคารพสิทธิของผู้ป่วย	การดูแลผู้ป่วยขาดมีจิตสำนึกที่ดี และเคารพสิทธิของผู้ป่วย	การดูแลผู้ป่วยขาดมีจิตสำนึกที่ดี และไม่เคารพสิทธิของผู้ป่วย		
3.5 มนุษยสัมพันธ์กับผู้ร่วมงาน		มีมนุษยสัมพันธ์ที่ดีมาก การทำงานเป็นทีมดีมาก	มีมนุษยสัมพันธ์ที่ดี ทำงานร่วมกับผู้อื่นได้	ขาดมนุษยสัมพันธ์ หรือมีปัญหาในการทำงานร่วมกับผู้อื่น	มนุษยสัมพันธ์ไม่ดี และ ไม่สามารถทำงานร่วมกับผู้อื่นได้		
เวลาปฏิบัติงาน		ครบ	ป่วย.....วัน	ลา.....วัน	ขาด.....วัน		
ความคิดเห็นเพิ่มเติม					ผู้ประเมิน
					(.....)
					วันที่

หมายเหตุ กรุณาให้คะแนนในช่องสี่เหลี่ยมหน้าของที่ต้องการ (ไม่มีจุดทศนิยม) , NA = ไม่สามารถประเมินได้

New staff workshop

Nov 2019

การตั้งเกณฑ์ผ่านการสอบ

นพ.เชิดศักดิ์ ไอรมนิรัตน์
ภาควิชาศัลยศาสตร์
คณะแพทยศาสตร์ศิริราชพยาบาล

1

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

2

Objectives

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

3

Outline

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

4

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

5

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

6

cherdsakiramaneeerat@gmail.com

1

New staff workshop

Nov 2019

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

7

1. Types of Standards

- Absolute standard
- Relative standard

8

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.

9

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 some way is easy for curriculum management

10

2. Methods for Setting Standards

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

11

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

12

cherdsakiraneerat@gmail.com

2

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

13

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

14

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 = $1/(\text{the number of options not eliminated})$
 - The passing score of the test = the sum of all the probability of correct answers of all items on the test

15

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

16

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

17

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

18

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

19

2.Methods for Setting Standards

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

20

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

21

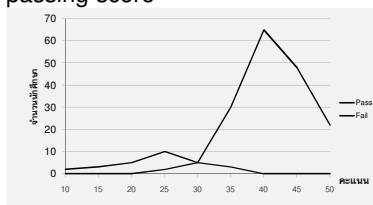
Borderline-Group Method

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

22

Contrasting-Groups Method

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



23

2.Methods for Setting Standards

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

24

New staff workshop

Nov 2019

Compromised Method

- Combining relative and absolute standard setting methods
 - Hofstee method

25

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

26

3. Selecting Judges

- The number of judges
- The qualification of judges

27

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.

28

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

29

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

30

cherdsakiramaneerat@gmail.com

5

New staff workshop

Nov 2019

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

31

5. Calculating Standard

- Outliers

32

6. Checking Standard

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

33

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

34

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

Confucius

35

35

cherdsakiramaneeerat@gmail.com

6

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

วิธีการตั้งเกณฑ์สอบผ่าน (passing standard) (ตอนที่ 1)

เชิดศักดิ์ ไอรมนรัตน์

เกณฑ์สอบผ่าน (passing standard) คือคะแนนสอบที่น้อยที่สุดที่คณาจารย์ยินยอมให้นักเรียนสามารถสอบผ่าน นักเรียนที่สอบได้คะแนนน้อยกว่าเกณฑ์สอบผ่านจะถูกตัดสินว่าสอบตก การตั้งเกณฑ์สอบผ่านจัดเป็นขั้นตอนที่มีความสำคัญมาก ในการจัดสอบ แต่กลับไม่ได้รับความสนใจเท่าที่ควรในการวัดผลทางแพทยศาสตรศึกษาจำนวนมาก ในบทความนี้ผมขอเสนอ เกร็ดความรู้เกี่ยวกับวิธีการตั้งเกณฑ์สอบผ่าน ผมหวังว่าอาจารย์ผู้อ่านจะสามารถนำเกร็ดความรู้ไปใช้พัฒนาคุณภาพของการตั้ง เกณฑ์สอบผ่านได้ไม่มากนักยอครับ

เกณฑ์สอบผ่านในทางแพทยศาสตรศึกษาจัดว่ามีความสำคัญมากเนื่องจากเกณฑ์สอบผ่านเป็นการแสดงออกถึง มาตรฐานของวิชาชีพที่อาจารย์ยอมรับ เกณฑ์สอบผ่านที่ดีต้องได้รับการตั้งขึ้นโดยใช้ดุลยพินิจของคณาจารย์ผู้เชี่ยวชาญใน สาขาวิชานั้นๆ เพื่อรักษามาตรฐานการประกอบวิชาชีพเพื่อให้สังคมได้รับบริการทางการแพทย์ที่มีคุณภาพ ในขณะเดียวกันกับให้ ความเป็นธรรมกับนักเรียนผู้สอบ เนื่องจากเกณฑ์สอบผ่านเป็นการแสดงออกถึง "ความยอมรับได้" ในดุลยพินิจของคณาจารย์ ผู้เชี่ยวชาญ จึงไม่มีวิธีการทางวิทยาศาสตร์ใดที่จะตัดสินว่าเกณฑ์ที่ตั้งขึ้นนั้นถูกหรือผิด สิ่งที่สำคัญที่สุดในการตั้งเกณฑ์สอบผ่าน หาใช่ "ตัวเลข" คะแนนที่จะใช้ตัดสินได้ตก หากแต่เป็น "กระบวนการ" ให้ได้มาซึ่งเกณฑ์ดังกล่าว เกณฑ์สอบผ่านที่ตั้งขึ้นโดยใช้ อาจารย์ 1 ท่านเลือกตัวเลข 1 ตัวเลขขึ้นมาโดยไม่ได้พิจารณาถึงข้อสอบหรือนักเรียนผู้สอบ เป็นวิธีการตั้งเกณฑ์ที่ล่อแหลมต่อการ ถูกวิจารณ์ (และประท้วง) โดยผู้ที่ไม่พอใจในผลสอบ วิธีการตั้งเกณฑ์สอบผ่านที่ดีนั้นต้องมีหลักการและเหตุผลประกอบ และผ่าน ดุลยพินิจของคณาจารย์ จำนวนของอาจารย์ผู้เชี่ยวชาญที่ต้องใช้ในการตั้งเกณฑ์นั้นขึ้นกับความสำคัญของการสอบนั้นๆ ในการ สอบที่มีความสำคัญสูงเช่นการสอบวุฒิบัตรแพทย์ผู้เชี่ยวชาญ แนะนำให้ใช้คณาจารย์อย่างน้อย 6 – 8 ท่าน ในการตั้งเกณฑ์ แต่ หากเป็นการสอบเล็กๆ เช่น การทดสอบหลังการสอบกลุ่มย่อย อาจใช้อาจารย์เพียง 1 ท่านก็ได้

การตั้งเกณฑ์สอบผ่านมี 2 ชนิดคือ การตัดสินแบบอิงเกณฑ์ (criterion-referenced standard, absolute standard) และการตัดสินแบบอิงกลุ่ม (norm-referenced standard, relative standard) การตัดสินแบบอิงเกณฑ์ เป็นการตั้งว่า คะแนน เท่าไร จึงจัดว่าผ่านการสอบ ในทางตรงข้าม การตัดสินแบบอิงกลุ่ม เป็นการตั้งว่า จะให้ นักเรียนจำนวนเท่าไร ผ่านการสอบ การ ตัดสินแบบอิงเกณฑ์นั้นเหมาะกับการสอบเพื่อวัดว่าผู้สอบมีความรู้ความสามารถในด้านใดด้านหนึ่งเพียงพอหรือไม่ ส่วนการสอบ แบบอิงกลุ่มนั้นเหมาะสำหรับการสอบแข่งขันเพื่อเข้าศึกษาต่อ หรือ ทำงาน ในสถาบันที่มีตำแหน่งที่จะรับได้จำกัด เช่น การสอบ เข้าโรงเรียนแพทย์ หรือ การสอบคัดเลือกแพทย์ประจำบ้าน การสอบส่วนใหญ่ในทางแพทยศาสตรศึกษานั้นเหมาะกับการตัดสิน แบบอิงเกณฑ์ หากผู้สอบทุกคนมีความสามารถเพียงพอก็ไม่จำเป็นต้องมีผู้สอบตก การใช้การตัดสินแบบอิงกลุ่มเพื่อวัดความรู้ ความสามารถในสถานการณ์อื่นนอกจากการสอบคัดเลือกนั้นเป็นการส่งเสริมให้นักเรียนเกิดความแข่งขันกัน (แทนที่จะช่วยกัน เรียน) โดยไม่จำเป็น

เนื่องจากการสอบทางแพทยศาสตรศึกษาแทบทั้งหมดเหมาะกับการตั้งเกณฑ์สอบผ่านแบบอิงเกณฑ์ ผมจะขอขยาย ความวิธีการตั้งเกณฑ์สอบผ่านแบบอิงเกณฑ์ที่สำคัญและใช้บ่อย 2 วิธีใหญ่ๆ คือ 1. การตั้งเกณฑ์โดยพิจารณาข้อสอบ และ 2. การ ตั้งเกณฑ์โดยพิจารณาจากผู้สอบ ในบทความตอนต่อไปครับ

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

วิธีการตั้งเกณฑ์สอบผ่าน (passing standard) (ตอนที่ 2)

เชิดศักดิ์ ไชยมณีรัตน์

ในบทความนี้ผมจะขอแนะนำวิธีการตั้งเกณฑ์สอบผ่านโดยพิจารณาตัวข้อสอบที่ใช้สอบ วิธีการตั้งเกณฑ์ผ่านแบบนี้เหมาะสำหรับการสอบ multiple-choice questions ซึ่งอาจารย์ผู้ตั้งเกณฑ์ผ่านสามารถประเมินความน่าจะเป็นของการตอบข้อสอบแต่ละข้อถูกต้อง การตั้งเกณฑ์ผ่านแบบนี้ประกอบด้วย 3 ขั้นตอนหลักคือ

1. ระบุลักษณะของนักเรียน"คาบเส้น" (borderline examinees): นักเรียนในกลุ่มคาบเส้นนี้คือนักเรียนที่มีความรู้ความสามารถอยู่ระหว่าง "ยอมรับได้" กับ "ยอมรับไม่ได้" นักเรียนกลุ่มนี้มีความรู้ไม่มากพอที่อาจารย์จะตัดสินใจให้สอบผ่านได้อย่างสบายใจ แต่ก็มีความรู้ไม่น้อยจนอาจารย์จะตัดสินใจให้สอบตกได้โดยไม่มีข้อสงสัย คณะกรรมการตั้งเกณฑ์สอบผ่านต้องระบุลักษณะของนักเรียนในกลุ่มคาบเส้นนี้อย่างชัดเจนว่า ในเนื้อหาวิชาที่ทำการสอบ นักเรียนกลุ่มนี้ควรมีความรู้ในเรื่องใด และไม่มีความรู้ในเรื่องใด ขั้นตอนนี้อาจทำได้ง่ายขึ้นหากอาจารย์แต่ละท่านนึกภาพของนักเรียนจริงที่อาจารย์เคยรู้จักที่สมควรถูกจัดให้อยู่ในกลุ่มนักเรียนคาบเส้น แล้วบรรยายลักษณะของนักเรียนคนนั้นๆ ว่าทำอะไรได้ และทำอะไรไม่ได้ รู้เรื่องอะไรบ้าง ไม่รู้เรื่องอะไรบ้าง
2. ให้กรรมการแต่ละท่านพิจารณาข้อสอบแต่ละข้อ และตัดสินใจว่านักเรียนคาบเส้นน่าจะมีโอกาสตอบข้อสอบถูกมากน้อยเพียงใด ขั้นตอนนี้สามารถทำได้หลายวิธีด้วยกัน ผมขอยกตัวอย่างวิธีที่เป็นที่แพร่หลายมาก 2 วิธีด้วยกัน คือ
 - 2.1. Angoff's method: ให้อาจารย์ระบุว่าหากนักเรียนคาบเส้น 100 คนทำข้อสอบข้อนั้น จะมีนักเรียนกี่คนที่ตอบข้อสอบข้อนั้นถูก (หรือความน่าจะเป็นที่นักเรียนคาบเส้นตอบข้อสอบข้อนั้นถูก)
 - 2.2. Ebel's method: ให้อาจารย์สร้างตารางแยกประเภทข้อสอบตามความสำคัญของเนื้อหาและตามความยากง่ายของข้อสอบและระบุว่าในข้อสอบแต่ละกลุ่ม 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 ความน่าจะเป็นของการตอบข้อสอบถูก สำหรับข้อสอบในกลุ่มนั้น) 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		

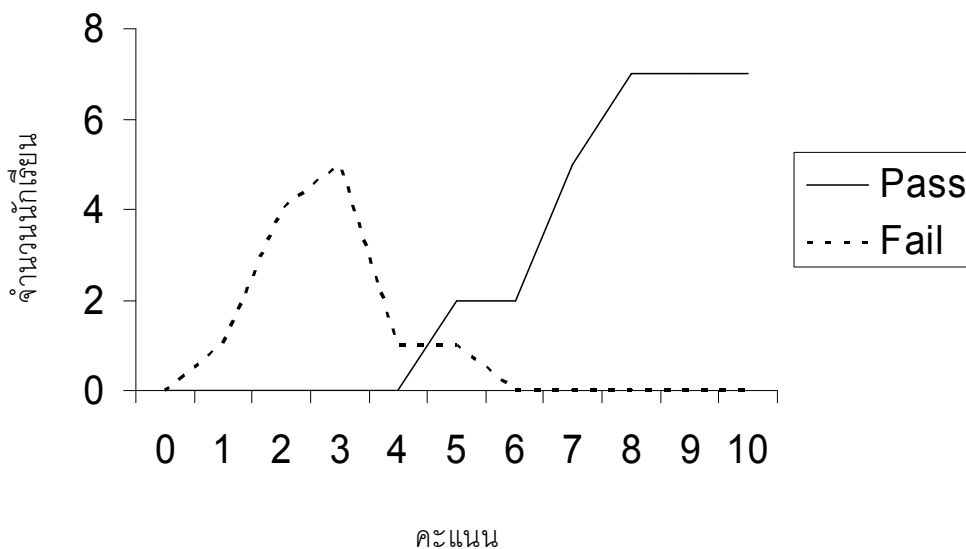
Iramaneerat C. Passing standard: Part III [Thai]. Medical Education Pamphlet 2006; 2(3): 1.

วิธีการตั้งเกณฑ์สอบผ่าน (passing standard) (ตอนที่ 3)

เชิดศักดิ์ ไอรมณีรัตน์

ในบทความนี้ผมจะขอแนะนำวิธีการตั้งเกณฑ์สอบผ่านโดยพิจารณานักเรียนผู้สอบ วิธีการตั้งเกณฑ์ผ่านแบบนี้เหมาะสำหรับการสอบวัดทักษะ การสอบสัมภาษณ์ หรือการประเมินการปฏิบัติงาน ซึ่งมักตัดสินการสอบผ่านโดยดูจากความสามารถของผู้สอบโดยรวมได้ง่ายกว่าดูจากคะแนนที่ได้ในหัวข้อประเมินแต่ละข้อ วิธีการตั้งเกณฑ์ผ่านลักษณะนี้ที่ใช้อยู่มีด้วยกัน 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)

เอกสารประกอบการอบรม



TEACH

15 November 2019

Nov 2019

Educational Project Proposal

รศ.นพ.เชดศักดิ์ ไอร่มณีรัตน์
ภาควิชาศัลยศาสตร์
คณะแพทยศาสตร์ศิริราชพยาบาล

**Education is the most powerful
Weapon which you can use
to change the world.**

Nelson Mandela

2

Educational Project

• Objectives

—อาจารย์ผู้เข้าอบรมสามารถ

- ประยุกต์ความรู้ทางการศึกษาที่ได้เรียนเพื่อวางแผนพัฒนางานการศึกษาที่ตนเกี่ยวข้อง
- เขียนโครงการพัฒนางานทางการศึกษาได้อย่างเหมาะสม

3

Assignment 1

- Educational project proposal
 - Background
 - Problem statement
 - Plan of action
 - Timeframe
 - Evaluation

4

Group Discussion of Proposal

- นำเสนอที่ละโครงการ
- หลังนำเสนอ รับฟังข้อเสนอแนะจากเพื่อน และ อาจารย์ ประจำกลุ่ม
- เวลา โครงการละ 5 - 8 นาที
- อาจารย์ให้ข้อเสนอแนะหลังนำเสนอแต่ละโครงการ

5

Scoring Criteria

1. ความชัดเจน
2. ความสำคัญ
3. ความเป็นไปได้
4. จริยธรรม
5. เทคนิคนำเสนอ

6

cherdsakiramaneerat@gmail.com

1

Project Proposal

แบบประเมินการนำเสนอโครงการพัฒนากายศึกษา (Educational project proposal)

	ดีมาก (10)	ดี (6)	ควรปรับปรุง (4)
1. ความชัดเจน	6-7 ประเด็น	4-5 ประเด็น	1-3 ประเด็น
2. ความสำคัญ	ระดับคณะ ทั่วประเทศ นานาชาติ	ระดับเขตภาค ระดับชาติ	ระดับกลุ่มเล็ก ผู้เริ่มเฉพาะ ชุมชนเฉพาะกลุ่ม
3. ความเป็นไปได้	เหมาะสมถึง 2 ด้าน	เหมาะสม 1 ด้าน	ไม่เหมาะสมถึง 2 ด้าน
4. จีเยอร์ม	ถูกต้อง 3 ประเด็นในขุมมุ่ง	มีจุดถึงประเด็นถึง 1 ประเด็น	มีจุดถึงประเด็นถึง 2 ประเด็นขึ้นไป ไร้ทิศทางหลัก ประเด็นใหญ่ 1 ประเด็น
5. เทคนิคนำเสนอ	ถูกต้อง 90 - 100%	ถูกต้อง 70 - 89%	ถูกต้องน้อยกว่า 70%

7

Educational Portfolio

- การทบทวนประสบการณ์การพัฒนากายศึกษา
สอนของตนเอง
 - บรรยายสิ่งที่วางแผน
 - เล่าถึงประสบการณ์ที่เกิดขึ้นจริง
 - สรุปผลลัพธ์ที่ได้จากการทำกิจกรรมนั้น
 - สะท้อนคิด (reflection): การเรียนรู้ที่เกิดขึ้น

Educational Portfolio (Project Summary)

- ส่งผลงาน
 - ทาง email: tipaporn.yod@mahidol.ac.th
 - หรือ ส่งด้วยตนเอง ที่คุณทิพาพร งานแพทยศาสตรศึกษา
ตึกอคูเลยเดซีวีกรมชั้น 6
 - กำหนดส่ง: ก่อนวันนำเสนอ อย่างน้อย 1 สัปดาห์

Evaluation Criteria

แบบประเมินการนำเสนอสรุปโครงการพัฒนากายศึกษา (Educational project summary)

	ดีมาก (10)	ดี (6)	ควรปรับปรุง (4)
1. ประเด็นเด่นชัด	4 ประเด็น	3 ประเด็น	1-2 ประเด็น
2. ความตั้งใจ	โครงการมีประโยชน์ สู่ประเทศชาติได้	ผ่านสู่ประเทศไม่ยากนักได้ โครงการมีเจตนาดี	ไม่เกิดประโยชน์หรือ อุปสรรคขัดแย้ง
3. เนื้อหา	มีสาระ มีคุณค่า ประโยชน์ต่อประเทศ นานาชาติ	มีสาระการปรับปรุง การศึกษาในระดับภาค คณะ ราชวิทยาลัย	มีสาระแต่ยังไม่ชัดเจน หรือมีผลเพียงระยะสั้นๆ
4. การสะท้อนคิด	Critical reflection	Practical reflection	Superficial reflection
5. เทคนิคนำเสนอ	ถูกต้อง 90 - 100%	ถูกต้อง 70 - 89%	ถูกต้องน้อยกว่า 70%

10

กิจกรรมการนำเสนอโครงการ

Time	Activities
1200 - 1230	Lunch
1230 - 1300	Guest lecture
1300 - 1600	Oral presentation - Presentation 10 min - Comment 10 min

แบบประเมินการนำเสนอโครงการพัฒนาการศึกษา (Educational project proposal)

ชื่อผู้นำเสนอ
ภาควิชา
โครงการ

	ดีมาก (10)	ดี (6)	ควรปรับปรุง (4)
1. ความชัดเจน <ul style="list-style-type: none"> - ทำอะไร - ผลลัพธ์ของโครงการคืออะไร - ทำเมื่อไร - ทำที่ไหน - ทำไมถึงต้องทำ - ทำอย่างไร - ใครเป็นคนทำ ใครได้รับผล 	6-7 ประเด็น	4-5 ประเด็น	1-3 ประเด็น
2. ความสำคัญ <p>ส่งผลกระทบต่อ การเรียนการสอน การวัดประเมินผล การทำหลักสูตร สร้างความรู้ใหม่ทางการศึกษา</p>	ระดับคณะ ประเทศ นานาชาติ	ระดับภาควิชา รายวิชา	คนกลุ่มเล็ก ๆ ผู้เรียนเฉพาะรุ่น เฉพาะกลุ่ม
3. ความเป็นไปได้ <ul style="list-style-type: none"> - ทรัพยากรที่ใช้ : เงิน บุคลากร ระยะเวลา - ผลลัพธ์ที่คาดหวัง 	เหมาะสมทั้ง 2 ด้าน	เหมาะสม 1 ด้าน	ไม่เหมาะสมทั้ง 2 ด้าน
4. จริยธรรม <ul style="list-style-type: none"> - respect for autonomy - beneficence, non-maleficence - justice 	ถูกหลัก 3 ประการ ในทุกแง่	ผิดหลักประเด็นเล็ก 1 ประเด็น	ผิดหลักประเด็นเล็ก 2 ประเด็นขึ้นไป หรือผิดหลักประเด็นใหญ่ 1 ประเด็น
5. เทคนิคนำเสนอ <ul style="list-style-type: none"> - การใช้สื่อการสอน - การสื่อสาร นำเสนอข้อมูล ด้วยภาษาพูด (วจนภาษา) - การสื่อสาร นำเสนอข้อมูล ด้วยภาษากาย (อวจนภาษา): น้ำเสียง สีหน้า ท่าทาง - การรับฟังความเห็นของผู้อื่น และ ตอบสนองอย่างเหมาะสม 	ถูกต้อง 90 - 100%	ถูกต้อง 70 - 89%	ถูกต้องน้อยกว่า 70%

รวมคะแนน

ข้อเสนอแนะ

อาจารย์ผู้ประเมิน

เกณฑ์พิจารณา

1. ความชัดเจน

- ทำอะไร
- ผลลัพธ์ของโครงการคืออะไร
- ทำเมื่อไร
- ทำที่ไหน
- ทำไมถึงต้องทำ
- ทำอย่างไร
- ใครเป็นคนทำ ใครได้รับผล

2. ความสำคัญ: ผลกระทบที่เกิดขึ้นจากการทำโครงการด้านการศึกษา ส่งผลต่อ การเรียนการสอน การวัดประเมินผล การจัดทำหลักสูตร การสร้างองค์ความรู้ใหม่ทางวิชาการศึกษา ในระดับใด

3. ความเป็นไปได้: ความเป็นไปได้ที่จะดำเนินโครงการสำเร็จ

- ทรัพยากรที่ใช้ : เงิน บุคลากร ระยะเวลา
- ผลลัพธ์ที่คาดหวัง

4. จริยธรรม

- Respect for autonomy เคารพให้เกียรติผู้ร่วมโครงการ มีการให้ข้อมูลแก่ผู้เกี่ยวข้องอย่างเหมาะสม ขอความยินยอมจากผู้เกี่ยวข้องอย่างถูกต้อง เปิดโอกาสให้ผู้ที่ไม่อยากร่วมโครงการปฏิเสธได้
- Beneficence and non-maleficence สร้างประโยชน์แก่ผู้ร่วมโครงการ ป้องกันอันตรายและผลเสียที่จะเกิดแก่ผู้ร่วมโครงการ
- Justice การดำเนินโครงการอย่างเป็นธรรม ไม่เลือกผู้เข้าร่วมโครงการอย่างลำเอียง ไม่ตัดโอกาสผู้ที่ควรได้รับประโยชน์อย่างไม่เหมาะสม

5. เทคนิคนำเสนอ: อาจารย์ประยุกต์ใช้เทคนิคในการนำเสนอที่ดีที่วิทยากรได้แนะนำในช่วงอบรมสู่ความเป็นเลิศ ฯ

- การใช้สื่อการสอน
- การสื่อสาร นำเสนอข้อมูล ด้วยภาษาพูด (วจนภาษา)
- การสื่อสาร นำเสนอข้อมูล ด้วยภาษากาย (อวจนภาษา): น้ำเสียง สีหน้า ท่าทาง
- การรับฟังความเห็นของผู้อื่น และตอบสนองอย่างเหมาะสม

แบบประเมินการนำเสนอสรุปโครงการพัฒนาการศึกษา (Educational project summary)

ชื่อผู้นำเสนอ
ภาควิชา
โครงการ

	ดีมาก (10)	ดี (6)	ควรปรับปรุง (4)
1. ประเด็นครบถ้วน - สิ่งที่วางแผน - สิ่งที่ทำจริง - ผลลัพธ์ - reflection	4 ประเด็น	3 ประเด็น	1-2 ประเด็น
2. ความตั้งใจ - แสดงถึงความตั้งใจในการพัฒนาการศึกษามาจะพยายาม ไม่ย่อท้อต่ออุปสรรค	โครงการซับซ้อน ฝ่าฟันอุปสรรคที่เกิดขึ้นได้ดี	ผ่านอุปสรรคที่ไม่ยากนักได้ โครงการสำเร็จด้วยดี	ล้มเลิกโครงการเมื่อเจออุปสรรคโดยไม่พยายาม
3. ผลลัพธ์ - ผลการดำเนินงาน	ตีพิมพ์ หรือนำเสนอในที่ประชุมวิชาการระดับประเทศหรือนานาชาติ	ส่งผลต่อการปรับปรุงการศึกษาในระดับภาควิชา คณะฯ ราชวิทยาลัย	เกิดผลกระทบต่อคนกลุ่มเล็กๆ หรือมีผลเพียงระยะสั้นๆ
4. การสะท้อนคิด	Critical reflection	Practical reflection	Superficial reflection
5. เทคนิคนำเสนอ - การใช้สื่อการสอน - การนำเสนอด้วยภาษาพูด (วงจนาษา) - อวจนาษา: น้ำเสียง สีหน้า ท่าทาง - การรับฟังความเห็นของผู้อื่น	ถูกต้อง 90 - 100%	ถูกต้อง 70 - 89%	ถูกต้องน้อยกว่า 70%

รวมคะแนน _____

ข้อเสนอแนะ

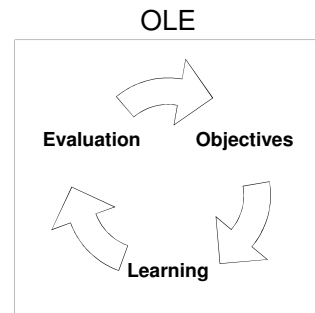
อาจารย์ผู้ประเมิน

เกณฑ์พิจารณา

1. ประเด็นครบถ้วน: บรรยายการดำเนินโครงการครบถ้วนทั้งที่ประเด็นตามที่ได้รับมอบหมาย
 - บรรยายสิ่งที่วางแผนไว้: สรุป **proposal** ที่เคยคิดว่าจะทำ
 - บรรยายประสบการณ์จริง: เล่าการดำเนินโครงการจริง ประสบปัญหาอะไรบ้าง แก้ปัญหาอย่างไร ต้องปรับเปลี่ยนโครงการอย่างไรบ้าง
 - ผลลัพธ์ของโครงการ: เล่าผลที่เกิดขึ้นจากการดำเนินโครงการ วัดผลอย่างไร มีตัวชี้วัดใดบ้าง ตัวชี้วัดใดดีขึ้น ตัวชี้วัดใดไม่ดีขึ้น อภิปรายผลที่เกิดขึ้น
 - ทบทวนประสบการณ์: **reflection** ว่าได้เรียนรู้อะไรบ้าง ปรับแนวคิดอย่างไรบ้าง
2. ความตั้งใจ: อาจารย์แสดงถึงความตั้งใจในการพัฒนาการศึกษาอย่างเต็มที่
 - มีความพยายามที่จะทำโครงการที่มีความท้าทาย โครงการที่ซับซ้อน ยากลำบาก
 - มานะพยายาม ไม่ย่อท้อต่ออุปสรรคที่เกิดขึ้น หาแนวทางแก้ปัญหา พินิจอุปสรรคอย่างเหมาะสม
3. ผลลัพธ์: ผลกระทบที่เกิดขึ้นจากการทำโครงการด้านการศึกษา ส่งผลต่อ การเรียนการสอน การวัดประเมินผล การจัดทำหลักสูตร การสร้างองค์ความรู้ใหม่ทางวิชาการศึกษา ในระดับใด
4. การสะท้อนคิด: มีการแสดงการสะท้อนคิด (**reflection**) อย่างเหมาะสม เพื่อนำไปสู่การพัฒนาตนเองอย่างต่อเนื่อง มีการสะท้อนคิดในระดับสูงได้
 - **Superficial reflection** การสะท้อนคิดอย่างผิวเผิน เพียงบรรยายสิ่งที่ได้พบ ได้เห็น
 - **Practical reflection** การสะท้อนคิดที่นำเอาประสบการณ์ที่ได้รับมาคิดวิเคราะห์ว่าเกิดอะไรขึ้น เพราะเหตุใด
 - **Critical reflection** การสะท้อนคิดอย่างลึกซึ้ง เอาบทเรียนที่ได้รับมาคิดและเสนอแนวทางปรับเปลี่ยนตนเองหรือระบบให้ดีขึ้นในอนาคต
5. เทคนิคนำเสนอ: อาจารย์ประยุกต์ใช้เทคนิคในการนำเสนอที่ดีที่วิทยากรได้แนะนำในช่วงอบรมสู่ความเป็นเลิศ ฯ
 - การใช้สื่อการสอน
 - การสื่อสาร นำเสนอข้อมูล ด้วยภาษาพูด (วจนภาษา)
 - การสื่อสาร นำเสนอข้อมูล ด้วยภาษากาย (อวจนภาษา): น้ำเสียง สีหน้า ท่าทาง
 - การรับฟังความเห็นของผู้อื่น และตอบสนองอย่างเหมาะสม

Nov 2019

Summary

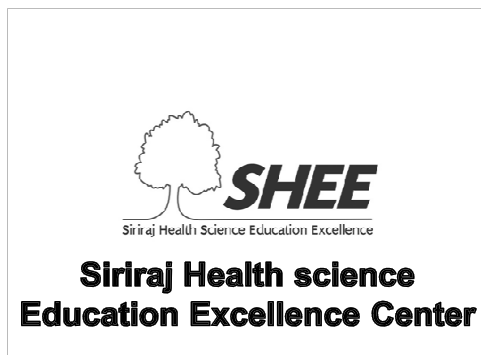
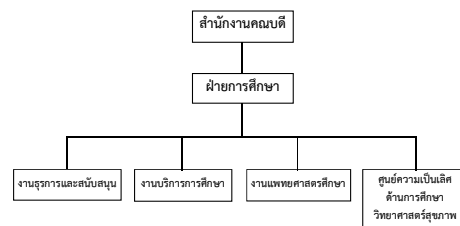


Assignment 2

- ขอให้อาจารย์แต่ละท่านเขียน

1. สิ่งที่ได้เรียนรู้จากการอบรมความรู้พื้นฐานทางการศึกษาที่เป็นประโยชน์ที่สุดคือ ...
2. ปณิธานในการเป็นครูแพทย์ของข้าพเจ้า

โครงสร้างการบริหารงาน



Shee.si.mahidol.ac.th

cherdsakiramaneeerat@gmail.com

1

Nov 2019

พันธกิจ

1. บริการวิชาการด้านการศึกษาสำหรับแพทย์ และบุคลากรทางวิทยาศาสตร์สุขภาพ
2. ส่งเสริมการวิจัยทางการศึกษาเพื่อพัฒนาด้านการศึกษาสำหรับแพทย์ และบุคลากรทางวิทยาศาสตร์สุขภาพ
3. ผลิตนักการศึกษา ซึ่งมีพื้นฐานทางวิทยาศาสตร์สุขภาพ โดยมีคุณธรรม และความรู้ ความสามารถ ทักษะ และเจตคติตามมาตรฐานสากล

Assignment 2

- ขอให้อาจารย์แต่ละท่านเขียน
1. สิ่งที่ได้เรียนรู้จากการอบรมความรู้พื้นฐานทางการศึกษาที่เป็นประโยชน์ที่สุดคือ ...
 2. ปณิธานในการเป็นครูแพทย์ของข้าพเจ้า

Objectives

Evaluation Learning

CherdsakIramaneerat@gmail.com
Tel: (062)823-4418

▶ Question & Comments

ศูนย์ความเป็นเลิศด้านการศึกษาวิทยาศาสตร์สุขภาพ (ศศว)
Siriraj Health science Education Excellence center (SHEE)

ฝ่ายการศึกษาก่อนปริญญา คณะแพทยศาสตร์ศิริราชพยาบาล

สำนักงาน: ตึกอดุลยเดชวิกรม ชั้น 6 (ห้อง 656)

Tel. 02 419 9978, 02 419 96637 Fax. 02 412 3901



shee.si.mahidol.ac.th



shee.mahidol@gmail.com



mahidol.shee



SHEE FC



Siriraj Health science Education Excellence center

