วิธีการประเมินความลำเอียงของข้อสอบ
(Method of Assessing Bias in Test Items : Janice Scheuneman)

ดร. มนตรี อนันตรักษ์ *

ในหลายปีที่ผ่านมา เรื่องความลำเอียงของข้อสอบเป็นเรื่องที่โต้เถียงกันมาก เพราะมีความสำคัญในการคัดเลือกบุคคลเข้าทำงานและเข้าเรียนต่อในวิทยาลัย ในการคัดเลือกต้องอาศัยข้อสอบและใช้ผลจากการสอบเปรียบเทียบกับเกณฑ์ ผู้สร้างข้อสอบได้หันมาสนใจในเรื่องความลำเอียงของข้อสอบก่อนที่จะนำไปใช้ในการสอบ เพื่อให้ได้ข้อสอบที่มีความยุติธรรมสำหรับผู้สอบ และผู้สร้างพยายามจะหาวิธีการที่จะขจัดความลำเอียงของข้อสอบ

วิธีหาความลำเอียงของข้อสอบ

1. Delta Plot Procedure (Angoff and Sharon, 1974)
2. Comparing responses to incorrect options of multiple choice items (Veale,1976)
3. Factor Analytic Procedure (Merz, 1976) (Green and Draper, 1972)
4. Latent Trait Theory and Item Characteristic Curve (Lord, 1977 and Rudner, 1977b)

Scheuneman กล่าวว่าเป็นวิธีการที่น่าสนใจมากกว่าวิธีการอื่น ๆ วิธีการนี้ต้องมีข้อตกลงว่า

1. test is fairly homogeneous
2. Measuring one underlying or latent ability

เขาได้วิจารณ์วิธีการของ Latent Trait Theory ว่า

1. เป็นวิธีการที่เสียค่าใช้จ่ายมาก (Vary Expensive) เพราะ
   1.1. ต้องเก็บข้อมูลกับคนจำนวนมาก เช่นของ Lord เก็บข้อมูลมากกว่า 2,000 ต่อกลุ่มและของ Green และ Draper ใช้ 1,000 คน
   1.2. ข้อสอบจะต้องมีจำนวนมาก อาจจะต้องถึง 60 ข้อ (Rudner.1977b)
2. ต้องใช้เครื่อง Computer ในการคำนวณ

---

* *อาจารย์ภาควิชาการวัดผลและวิจัยการศึกษา คณะศึกษาศาสตร์มหาวิทยาลัยศรีนครินทรวิโรฒ มหาสารคาม* *สรุปจาก Journal of Educational Measurement. Volume 16. No.3. Fall 1979*

วิธีการหาความลำเอียงของข้อสอบของ Scheuneman

หลักการ

ข้อสอบที่ไม่ลำเอียง (UNBIASED ITEM) หมายถึงข้อสอบที่มีความน่าจะเป็นของการตอบถูกในกลุ่มคนที่มีความสามารถเหมือนกันจะเท่ากัน โดยไม่คำนึงถึงความแตกต่างของวัฒนธรรม (ethnic qroup) ความสามารถวัดได้โดยคะแนนรวมของข้อสอบที่มีความเป็นเอกพันธ์ หรือตัวแทนของข้อสอบที่เราสนใจ ความแตกต่างของคะแนนรวมจะถูกแบ่งออกเป็นช่วง ๆ สัดส่วนของแต่ละคนภายในช่วงคะแนนซึ่งตอบข้อสอบถูกแต่ละข้อถือเป็นค่าประมาณความน่าจะเป็นในการตอบถูกสำหรับคะแนนของกลุ่มในช่วงนั้นและใช้ Chi Square (χ^2) ตรวจว่าข้อสอบแต่ละข้อมีความลำเอียงหรือไม่ตามความหมายนี้

วิธีการคำนวณความลำเอียงของข้อสอบ

SCHEUNEMAN

Table 1

Example of the Computation of Chi Square for One Item

Total Score on Subtest	Number with Scores in each Range			Obtained Frequencies (No. with item correct)			Proportion correct	Expected Frequencies	
								Black	White
	Black	White	Total	Black	White	Total	(T_0/T)	(p.B)	(p.W)
	B	W	T	B_0	W_0	T_0	p	B_e	W_e
13–14	25	315	340	22	300	322	.9471	23.68	298.32
12	24	110	134	18	99	117	.8731	20.96	96.04
10–11	48	118	166	23	93	116	.6988	33.54	82.46
1–9	65	92	157	14	33	47	.2994	19.46	27.54

$$\chi^2 = \Sigma \frac{(B_e - B_0)^2}{B_e} + \Sigma \frac{(W_e - W_0)^2}{W_e} = 7.912, \text{d.f.} = 3$$

ปัญหาของวิธีการนี้

1. การแบ่งช่วงคะแนนรวมเท่าไรจึงจะเหมาะสม

2. ความน่าจะเป็นในการที่จะ Reject Ho ว่าควรจะเท่าใดโดยปกติใช้ $\propto$ = .05 แต่ก็ขึ้นอยู่กับจุดมุ่งหมายของการวิเคราะห์ข้อสอบ อาจจะใช้ .10 หรือเท่าใดก็ได้

ข้อเสนอแนะ ความน่าจะเป็นในการตอบถูกในความสามารถแต่ละช่วงจะต้องน้อยกว่า 1 และควรจะมีจำนวนในการตอบถูกแต่ละช่วง เพราะ χ^2 จะต้องมีค่าความถี่ที่คาดหวัง (Expected Frequencies) อย่างน้อย 5 (จำนวนในแต่ละ Cell จะขึ้นอยู่กับความยากและความยาวของข้อสอบและจำนวนคนที่ไปสอบ)

ตัวอย่างการพิจารณาข้อสอบ

ASSESSING BIAS IN TEST ITEMS

Table 2

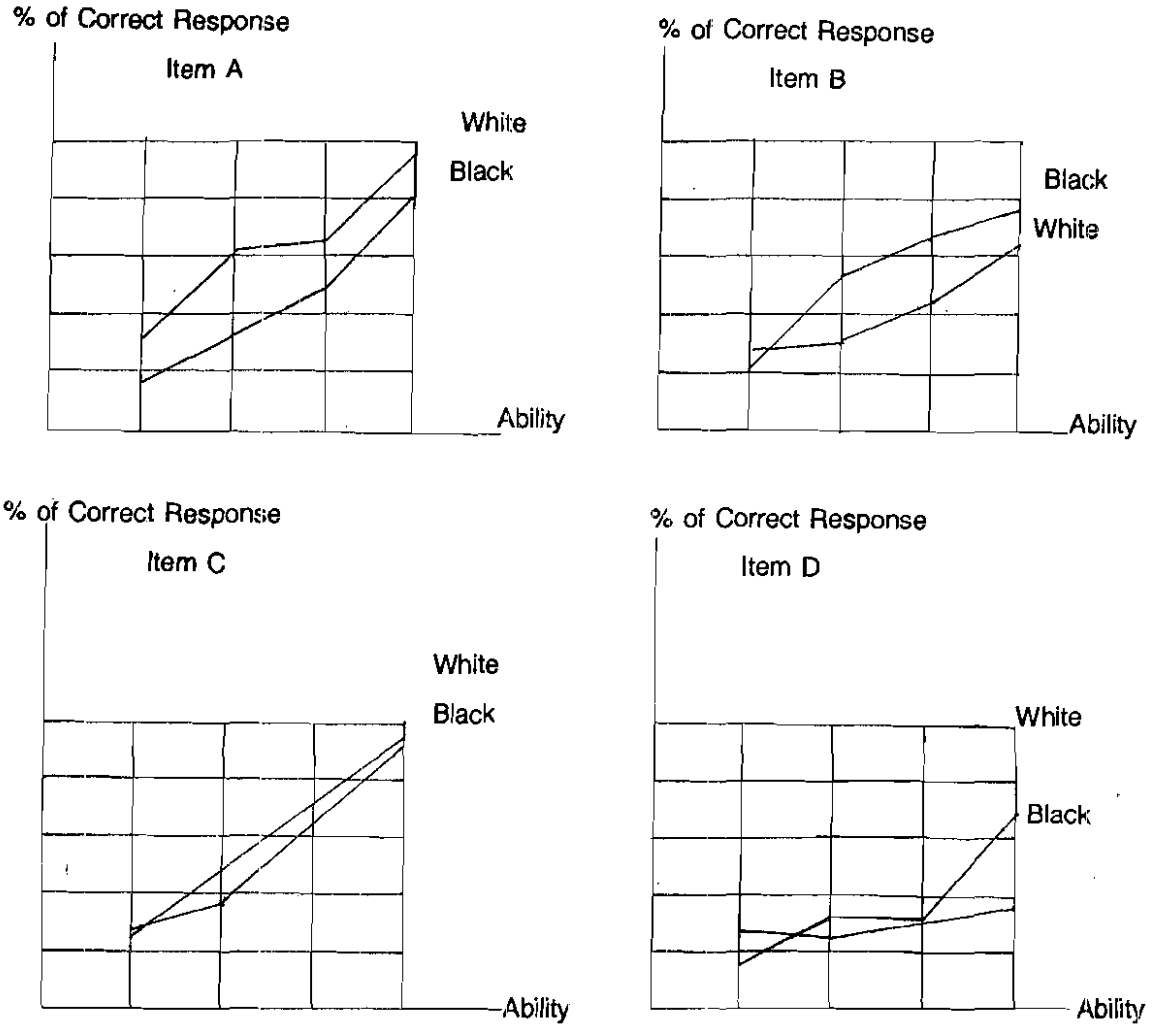
Sample Results for Selected Items from a Quantitative Contepts Subtest

Item	Population Group	Percent Correct	Point Biserial	% of Correct Responses by Score Interval				Chi Square[b]	Prob.[c]
				13-22[a]	10-12	8-9	3-7		
A	Black	40	.42	83	50	33	18	6.00	<.20
	White	78	.42	94	68	61	32		
B	Black	50	.34	75	65	54	21	6.34	>.10
	White	53	.33	65	43	29	28		
C	Black	50	.43	92	68	37	27	0.70	<.80
	White	78	.48	93	72	50	25		
D	Black	29	.09	36	29	25	27	4.66	<.20
	White	56	.48	70	30	32	17		

[a] For item C, the highest ability group was defined by a score range of 12-22; the next lower interval was 10-11.

[b] The chi square has 3 degrees of freedom for all examples given here.

[c] The probability of obtaining a chi square of that value or larger assuming that the



วิธีการนี้ถูกวิจารณ์ในแง่การใช้คะแนนรวมเป็นความสามารถของการวัด เพราะคะแนนรวมอาจจะมีความลำเอียงโดยตัวมันเอง แต่อย่างไรก็ตาม Rudner (1977b) ได้เปรียบเทียบการหาความลำเอียงของข้อสอบ โดยวิธีการต่าง ๆ พบว่าวิธีการของ Chi Square กับวิธี Item Characteristic Curve (3 parameter) กับกลุ่ม 2 กลุ่มมีความสัมพันธ์ = .67 และสัมพันธ์กับวิธี Delta Plot = .59 โดยที่ Delta Plot กับ ICC =.31 และ Ironson (1978) พบว่าวิธีการของ Chi Square กับวิธี ICC มีความสัมพันธ์ = .49 และสัมพันธ์กับวิธี Delta plot = .37 โดยที่ Delta Plot กับ ICC = .24 ผลของการศึกษานี้เขาก็แนะนำว่า การหาความลำเอียงของข้อสอบนิยมใช้กัน 3 วิธีคือ Chi Square, ICC และ Delta Plot

สรุปวิธีการของ Chi Square เป็นวิธีการที่คำนวณง่ายเหมาะสำหรับการที่นำไปใช้ เพราะเป็นวิธีที่ไม่ยุ่งยากซับซ้อนและไม่ต้องอาศัยบุคคลที่มีความรู้ทางสถิติมากนักและง่ายในการที่จะอธิบายให้คนอื่นเข้าใจ