

บทที่ 2
วรรณคดีที่เกี่ยวข้อง

ในบทนี้จะกล่าวถึงที่มาของแนวคิดในการวิเคราะห์แบบแผนการตอบข้อสอบ ดัชนีที่ใช้วัดความสอดคล้องของแบบแผนการตอบข้อสอบรายบุคคล และงานวิจัยที่เกี่ยวข้องกับดัชนีดังกล่าว โดยแบ่งรายละเอียดเป็น 3 ตอน คือ

1. ที่มาของแนวคิด : ความสอดคล้องของคะแนนรายชื่อของผู้สอบทั้งกลุ่ม กับค่าความยากของข้อสอบ
2. ดัชนีที่ใช้วัดความสอดคล้องของแบบแผนการตอบข้อสอบรายบุคคล
3. งานวิจัยที่เกี่ยวข้องกับดัชนีที่ใช้วัดความสอดคล้องของแบบแผนการตอบข้อสอบรายบุคคล

ตอนที่ 1 ที่มาของแนวคิด : ความสอดคล้องของคะแนนรายชื่อของผู้สอบทั้งกลุ่ม กับค่าความยากของข้อสอบ

ก่อนที่จะมีการวิเคราะห์แบบแผนการตอบข้อสอบรายบุคคล กัตแมน (Guttman 1940) นับว่าเป็นผู้บุกเบิกแนวคิดนี้ โดยการเสนอดัชนีที่บ่งชี้ความสอดคล้องของคะแนนรายชื่อของผู้สอบทั้งกลุ่มกับค่าความยากของข้อสอบ และเรียกดัชนีนี้ว่าดัชนีผลผลิตใหม่ (Reproducibility Index: R) ต่อมา แจคสัน (Jackson 1949) ได้พัฒนาดัชนีอัตราส่วนเปอร์เซ็นต์เชิงบวก (Plus Percentage Ratio: PPR) ขึ้นเพื่อแก้ไขจุดอ่อนบางประการของดัชนีผลผลิตใหม่ นอกจากนี้ยังมีดัชนีความเป็นเอกพันธ์ (Index of Homogeneity: H) ของ โลวีนเจอร์ (Loevinger 1947) และดัชนีความสอดคล้อง (Index of Consistency: I) ของ กรีน (Green 1956) ซึ่งดัชนีที่กล่าวมานี้จะบอกถึงความสอดคล้องดังกล่าวของข้อสอบเป็นรายชื่อ และของแบบสอบทั้งฉบับ

1. ดัชนีผลผลิตใหม่ (Reproducibility Index: R)

กัตแมน พัฒนาดัชนีผลผลิตใหม่ขึ้นเพื่อบ่งชี้ความสอดคล้องของคะแนนรายชื่อของผู้สอบทั้งกลุ่มกับค่าความยากของแบบสอบทั้งฉบับ และของข้อสอบรายข้อ โดยพิจารณาจากคะแนนรายชื่อของผู้สอบ

1.1 ข้อตกลงเบื้องต้นของกัตแมน (Allen and Yen 1979: 186)

1.1.1 ข้อสอบข้อใดก็ตามที่ผู้สอบที่มีความสามารถระดับหนึ่งตอบได้ถูกต้องแล้ว ผู้สอบทุกคนที่มีความสามารถสูงกว่าย่อมจะตอบได้ถูกต้อง ในทางตรงกันข้ามข้อสอบข้อใดก็ตามที่ผู้สอบที่มีความสามารถระดับหนึ่งตอบผิดแล้ว ผู้สอบทุกคนที่มีความสามารถต่ำกว่าย่อมจะตอบผิดด้วย

1.1.2 ผู้สอบคนใดก็ตามที่ตอบข้อสอบที่มีความยากระดับหนึ่งได้ถูกต้องแล้ว เขาจะตอบข้อสอบทุกข้อที่ง่ายกว่าข้อดังกล่าวได้ถูกต้องด้วย ในทางตรงกันข้ามเมื่อผู้สอบตอบข้อสอบที่มีความยากระดับหนึ่งผิด เขาย่อมจะตอบข้อสอบทุกข้อที่ยากกว่าข้อดังกล่าวผิดด้วย

กัตแมน เสนอให้จัดเรียงลำดับข้อสอบจากข้อที่ง่ายที่สุด (ข้อที่มีจำนวนผู้สอบภายในกลุ่มตอบได้ถูกต้องมากที่สุด) ไปจนถึงข้อที่ยากที่สุด (ข้อที่มีจำนวนผู้สอบภายในกลุ่มตอบได้ถูกต้องน้อยที่สุด) คะแนนรายข้อของผู้สอบแต่ละคนจะถูกนำมาเรียงลำดับจากข้อที่ง่ายที่สุดไปจนถึงข้อที่ยากที่สุดตามแนวสดมภ์ และเรียกคะแนนรายข้อของผู้สอบแต่ละคนที่ถูกนำมาจัดเรียงลำดับนี้ว่า แบบแผนการตอบข้อสอบ (item response pattern) จากนั้นนำแบบแผนการตอบข้อสอบของผู้สอบทั้งกลุ่มมาเรียงลำดับตามแนวแถวจากบนลงล่างตามคะแนนรวมมากไปน้อย เป็นเมตริกของคะแนนรายข้อของผู้สอบ N คน ที่ตอบข้อสอบ K ข้อ ลักษณะเมตริกของกัตแมนนี้จะมีความสอดคล้องสมบูรณ์เป็นไปตามข้อตกลงเบื้องต้นเมื่อคะแนนรายข้อที่อยู่มุมบนด้านซ้ายมือของเมตริกมีค่าเป็น 1 หมด และคะแนนรายข้อที่อยู่มุมล่างด้านขวามือมีค่าเป็น 0 หมด และเรียกเมตริกที่มีลักษณะเช่นนี้ว่า กัตแมนสมบูรณ์ (Perfect Guttman) ดังตัวอย่างในตารางที่ 1

ตารางที่ 1 แบบแผนการตอบข้อสอบ 5 ข้อ ของผู้สอบ 4 คน ที่มีลักษณะเป็นเมตริก กัตแมนสมบูรณ์

ผู้สอบ	ข้อกระถง				
	2	4	3	1	5
ค	1	1	1	1	0
ก	1	1	1	0	0
ข	1	1	0	0	0
ง	1	0	0	0	0

เนื่องจากแบบแผนการตอบข้อสอบตามแบบกัตแมนสมบูรณ์ถือว่าเป็นแบบแผนในอุดมคติ ตามสภาพความเป็นจริงแบบแผนการตอบข้อสอบมักจะมี ความคลาดเคลื่อนเกิดขึ้น โดยที่กัตแมนได้ นิยามความคลาดเคลื่อนเป็น 2 ประเภท คือ ประเภทที่ 1 ความคลาดเคลื่อนที่เกิดจากการตอบข้อ สอบที่จัดว่าง่ายกว่าความสามารถของเขาผิด และประเภทที่ 2 ความคลาดเคลื่อนที่เกิดจากการ ตอบข้อสอบที่จัดว่ายากกว่าความสามารถของเขาได้ถูกต้อง (Cudeck 1980: 119)

1.2 สูตรคำนวณดัชนีผลผลิตใหม่ (R) (White and Saltz 1957: 85)

1.2.1 ดัชนีผลผลิตใหม่ของข้อสอบ (R_i)

$$R_i = 1 - \frac{E}{N}$$

โดยที่

R_i คือ ค่าดัชนีผลผลิตใหม่ของข้อสอบข้อที่ i

E คือ จำนวนของคะแนนรายข้อที่มีความคลาดเคลื่อนของข้อสอบ

ข้อที่ i

N คือ จำนวนผู้สอบทั้งหมดที่ตอบข้อสอบข้อที่ i

1.2.2 ดัชนีผลผลิตใหม่ของแบบสอบ (R_c)

$$R_c = 1 - \frac{\sum_{i=1}^K E_i}{NK}$$

โดยที่

R_c คือ ค่าดัชนีผลผลิตใหม่ของแบบสอบ

i คือ ลำดับที่ของข้อสอบ ; $i = 1, 2, \dots, K$ โดยที่ข้อสอบ

ลำดับที่ 1 เป็นข้อที่ง่ายที่สุด และข้อสอบลำดับที่ K เป็นข้อที่ยากที่สุด

$\sum_{i=1}^K E_i$ คือ ผลรวมของจำนวนของคะแนนรายข้อที่มีความคลาดเคลื่อนของ

ข้อสอบทั้งหมด K ข้อ

N คือ จำนวนผู้สอบทั้งหมด

K คือ จำนวนข้อสอบทั้งหมด

ตารางที่ 2 ซึ่งจะกล่าวต่อไปนี้เป็นตัวอย่างการคำนวณค่าดัชนีผลผลิตใหม่ของข้อสอบเป็นรายข้อ และของแบบสอบทั้งฉบับที่ได้จากการตอบข้อสอบ 6 ข้อ ของผู้สอบ 10 คน

ตารางที่ 2 ตัวอย่างการคำนวณค่าดัชนีผลผลิตใหม่ของข้อสอบเป็นรายข้อและของแบบสอบทั้งฉบับ ที่ได้จากแบบแผนการตอบข้อสอบ 6 ข้อ ของผู้สอบ 10 คน

ผู้สอบ	ข้อสอบ						คะแนนรวม
	2	6	4	1	5	3	
ง	1	1	1	(0)	1	<u>1</u>	5
จ	1	1	1	1	<u>1</u>	0	5
ข	1	1	1	<u>1</u>	0	0	4
ฉ	1	1	1	0	(1)	0	4
ช	1	1	1	0	0	0	3
ญ	(0)	1	<u>1</u>	(1)	0	0	3
ซ	1	(0)	0	0	0	(1)	2
ฉ	1	<u>1</u>	0	0	0	0	2
ค	<u>1</u>	0	0	0	0	0	1
ก	0	0	0	(1)	0	0	1
จำนวนผู้ที่ตอบถูก	8	7	6	4	3	2	
E_i	1	1	0	3	1	1	
R_i	.90	.90	1.00	.70	.90	.90	

หมายเหตุ (0) หรือ (1) คือ คะแนนรายข้อที่นับว่าเป็นคะแนนที่มีความคลาดเคลื่อน

— คือ เส้นแบ่งเขตของคะแนนรายข้อตามลักษณะกัณฑ์แมนสมบูรณ์

E_i คือ จำนวนของคะแนนรายข้อที่มีความคลาดเคลื่อน

R_i คือ ค่าของดัชนีผลผลิตใหม่ของข้อสอบข้อที่ i

จากตัวอย่างที่เสนอในตารางที่ 2 จะเห็นได้ว่าคะแนนรายชื่อของผู้สอบทั้ง 10 คน ที่ตอบข้อสอบจำนวน 6 ข้อ มีลักษณะที่แตกต่างไปจากกัตแมนสมบรูณ์ ดังนั้นการลากเส้นแบ่งเขตของคะแนนรายชื่อตามลักษณะกัตแมนสมบรูณ์ของข้อสอบแต่ละข้อ จะต้องพิจารณาให้จำนวนของคะแนนรายชื่อที่มีความคลาดเคลื่อนมีจำนวนน้อยที่สุด ยกตัวอย่างเช่น ข้อสอบข้อที่ 1 ซึ่งมีจำนวนของคะแนนรายชื่อที่มีความคลาดเคลื่อนเท่ากับ 3 แต่ถ้าเลื่อนเส้นแบ่งเขตของคะแนนรายชื่อตามกัตแมนสมบรูณ์ไปอยู่ที่ผู้สอบคนที่ ๖ หรือ คนที่ ๗ จะมีผลให้จำนวนของคะแนนรายชื่อที่มีความคลาดเคลื่อนเพิ่มขึ้นเป็น 4 สรุปได้ว่าหลักการลากเส้นแบ่งเขตของคะแนนรายชื่อตามกัตแมนสมบรูณ์ หรือที่เรียกว่า จุดตัด (cutting point) นี้ คือ การพิจารณาให้จำนวนของคะแนนรายชื่อที่มีความคลาดเคลื่อนมีจำนวนน้อยที่สุด (minimum error) และในกรณีที่มีจุดตัดมากกว่า 1 แห่งที่ทำให้จำนวนของคะแนนรายชื่อที่มีความคลาดเคลื่อนมีจำนวนเท่ากันและเป็นจำนวนที่น้อยที่สุด ให้เลือกจุดตัดที่อยู่ในตำแหน่งที่ไกลเคียงแต่ลดหลั่นจากจุดตัดของข้อที่ง่ายกว่า ยกตัวอย่างเช่นข้อสอบข้อที่ 6 ซึ่งมีจำนวนผู้ตอบถูก 7 คน ถ้าลากเส้นแบ่งเขตของคะแนนรายชื่อตามกัตแมนที่ผู้สอบคนที่ ๗ หรือ คนที่ ๘ จะทำให้จำนวนของคะแนนรายชื่อที่มีความคลาดเคลื่อนมีจำนวนเท่ากันคือเท่ากับ 1 แต่โดยหลักการแล้วควรเลือกจุดตัดตรงคนที่ ๘ เพราะเป็นจุดตัดที่อยู่ไกลเคียงแต่ลดหลั่นไปจากจุดตัดของข้อสอบข้อที่ 2 ซึ่งเป็นข้อที่ง่ายกว่า

ดังตัวอย่างที่เสนอในตารางที่ 2 เป็นการนำคะแนนรายชื่อมาจัดเรียงให้เป็นเมตริกตามกัตแมน และแสดงวิธีการกำหนดเส้นแบ่งเขตคะแนนรายชื่อตามกัตแมนสมบรูณ์ การคำนวณค่าของดัชนีผลผลิตใหม่รายชื่อ และสามารถจะคำนวณค่าดัชนีผลผลิตใหม่ของแบบสอบทั้งฉบับ (R_c) ได้โดยการรวมจำนวนของคะแนนรายชื่อที่มีความคลาดเคลื่อนของข้อสอบทั้ง 6 ข้อ ซึ่งในที่นี้มีค่าเท่ากับ 7 เมื่อแทนค่าต่างๆในสูตรคำนวณแล้ว ดัชนีผลผลิตใหม่ของแบบสอบทั้งฉบับจะมีค่าเท่ากับ 0.88 [0.88 ได้มาจาก $1 - (7 / (6 \times 10))$]

1.3 นิสัย และค่าของจุดตัดของดัชนีผลผลิตใหม่

ดัชนีผลผลิตใหม่มีค่าต่ำสุดไม่แน่นอน ขึ้นอยู่กับค่าความยากของข้อสอบในแบบสอบฉบับนั้นๆ (Jackson and others อ้างถึงใน White and Saltz 1957: 86) ส่วนค่าสูงสุดของดัชนีมีค่าเท่ากับ 1.0 เหตุที่ดัชนีผลผลิตใหม่มีค่าต่ำสุดไม่แน่นอนนี้เกิดขึ้นเนื่องจากค่าความยากของข้อสอบ มีผลให้เกิดความจำกัดของจำนวนของคะแนนรายชื่อที่มีความคลาดเคลื่อน และข้อสอบที่มีค่าความยากเท่ากับ 0.5 จะมีจำนวนคะแนนรายชื่อที่มีความคลาดเคลื่อนสูงสุด คือมีค่าเท่ากับจำนวนผู้สอบทั้งหมด โดยที่ครึ่งหนึ่งของจำนวนผู้สอบทั้งหมดนับจากคนที่มีความรวมสูงสุด

ตอบข้อสอบข้อดังกล่าวซึ่งจัดว่าเป็นข้อที่ง่ายกว่าความสามารถของเขาผิดทุกคน ในขณะที่จำนวนผู้สอบอีกครั้งหนึ่งตอบข้อสอบข้อดังกล่าว ซึ่งจัดว่าเป็นข้อที่ยากกว่าความสามารถของเขาได้ถูกต้องทุกคน จากตารางที่ 2 จะเห็นได้ว่า ค่าดัชนีผลผลิตใหม่รายข้อ (R_i) มีค่าต่ำสุดเท่ากับ 0.70 การที่ดัชนีผลผลิตใหม่มีค่าต่ำสุดไม่แน่นอนขึ้นอยู่กับค่าความยากของข้อสอบนี้ ทำให้มีปัญหาในการกำหนดค่าจุดตัดของดัชนีที่จะบอกความสอดคล้องของคะแนนรายข้อของผู้สอบทั้งกลุ่มกับค่าความยากของข้อสอบเป็นรายข้อหรือของแบบสอบทั้งฉบับ แต่อย่างไรก็ตาม กัทแมนได้เสนอแนะว่าข้อสอบหรือแบบสอบฉบับใดๆจะถือว่ามิระดับของความสอดคล้องที่ยอมรับได้จะต้องมีค่าของดัชนีผลผลิตใหม่ตั้งแต่ 0.9 ขึ้นไป (White and Saltz 1957: 86)

2. ดัชนีอัตราส่วนเปอร์เซ็นต์เชิงบวกของแจคสัน (Jackson's Plus Percentage Ratio: PPR)

เพื่อแก้ไขจุดอ่อนในเรื่องค่าต่ำสุดที่ไม่แน่นอนของดัชนีผลผลิตใหม่ ในปีค.ศ. 1949 แจคสัน จึงได้พัฒนาดัชนีของเขาขึ้นเพื่อแก้ไขจุดอ่อนดังกล่าว หลักการประมาณค่าดัชนีอัตราส่วนเปอร์เซ็นต์เชิงบวกของแจคสัน คือการหาค่าดัชนีผลผลิตใหม่ต่ำสุด (Minimum Reproducibility Index: MR) ซึ่งก็คือค่าต่ำสุดของดัชนีผลผลิตใหม่ แล้วนำค่าที่ได้ไปลบออกจากค่าดัชนีผลผลิตใหม่ที่คำนวณได้ ค่าความแตกต่างที่ได้นี้เรียกว่าเปอร์เซ็นต์เชิงบวก (Plus %) จากนั้นนำค่าเปอร์เซ็นต์เชิงบวกนี้ มาหารด้วยค่าผลต่างของ 1.0 กับค่าดัชนีผลผลิตใหม่ต่ำสุด (1.0 คือค่าสูงสุดของดัชนีผลผลิตใหม่)

2.1 สูตรคำนวณค่าดัชนีผลผลิตใหม่ต่ำสุดและค่าดัชนีอัตราส่วนเปอร์เซ็นต์เชิงบวก

2.1.1 ค่าดัชนีผลผลิตใหม่ต่ำสุดคำนวณได้จากสูตรต่อไปนี้ (White and

Saltz 1957: 86)

2.1.1.1 ค่าดัชนีผลผลิตใหม่ต่ำสุดรายข้อ (MR_i) มีสูตรคำนวณคือ

$$MR_i = \frac{\# \text{ rights or wrongs whichever is larger}}{N}$$

โดยที่

MR_i คือ ค่าของดัชนีผลผลิตใหม่ต่ำสุดของ

ข้อสอบข้อที่ i

ได้ถูกต้อง

rights คือ จำนวนผู้สอบที่ตอบข้อสอบข้อที่ i

ผิด

wrongs คือ จำนวนผู้สอบที่ตอบข้อสอบข้อที่ i

rights or # wrongs whichever is larger คือ จำนวนผู้สอบที่ตอบข้อสอบข้อที่ i ได้ถูกต้อง หรือจำนวนผู้สอบที่ตอบข้อสอบข้อที่ i ผิด โดยเลือกใช้จำนวนที่มากกว่า

N คือ จำนวนผู้สอบทั้งหมด

2.1.1.2 ค่าดัชนีผลผลิตใหม่ต่ำสุดของแบบสอบทั้งฉบับ (MR_c) มี

สูตรคำนวณคือ

$$MR_c = \frac{\sum_{i=1}^k \# \text{ rights or wrongs whichever is larger}}{NK}$$

โดยที่

MR_c คือ ค่าดัชนีผลผลิตใหม่ต่ำสุดของแบบ-

สอบทั้งฉบับ

rights คือ จำนวนผู้สอบที่ตอบข้อสอบข้อที่ i

ได้ถูกต้อง

wrongs คือ จำนวนผู้สอบที่ตอบข้อสอบข้อที่ i

ผิด

rights or # wrongs whichever is larger คือ จำนวนผู้สอบที่ตอบข้อสอบข้อที่ i ได้ถูกต้อง หรือ จำนวนผู้สอบที่ตอบข้อสอบข้อที่ i ผิด โดยเลือกใช้จำนวนที่มากกว่า

N คือ จำนวนผู้สอบทั้งหมด

K คือ จำนวนข้อสอบทั้งหมด

2.1.2 ค่าดัชนีอัตราส่วนเปอร์เซ็นต์เชิงบวก (PPR)

2.1.2.1 ค่าดัชนีอัตราส่วนเปอร์เซ็นต์เชิงบวกรายข้อ (PPR_c) มี

สูตรคำนวณคือ (White and Saltz 1957: 86)

$$PPR_i = \frac{R_i - MR_i}{1 - MR_i}$$

โดยที่

PPR_i คือ ค่าอัตราส่วนเปอร์เซ็นต์เชิงบวกของข้อสอบข้อที่ i

R_i คือ ค่าดัชนีผลผลิตใหม่ของข้อสอบข้อที่ i

MR_i คือ ค่าดัชนีผลผลิตใหม่ต่ำสุดของข้อสอบข้อที่ i

2.1.2.2 ค่าดัชนีอัตราส่วนเปอร์เซ็นต์เชิงบวกของแบบสอบทั้งฉบับ

(PPR_c) มีสูตรคำนวณคือ (White and Saltz 1957: 87)

$$PPR_c = \frac{R_c - MR_c}{1 - MR_c}$$

โดยที่

PPR_c คือ ค่าดัชนีอัตราส่วนเปอร์เซ็นต์เชิงบวกของแบบสอบ

ทั้งฉบับ

R_c คือ ค่าดัชนีผลผลิตใหม่ของแบบสอบทั้งฉบับ

MR_c คือ ค่าดัชนีผลผลิตใหม่ต่ำสุดของแบบสอบทั้งฉบับ

ตารางที่ 3 ที่จะนำเสนอต่อไปนี้ เป็นตัวอย่างของขั้นตอนและวิธีการหาค่าดัชนีอัตราส่วนเปอร์เซ็นต์เชิงบวกของแฉกสั้น ที่ได้จากการตอบข้อสอบ 6 ข้อ ของผู้สอบ 10 คน ซึ่งเป็นข้อมูลเดียวกับตารางที่ 2

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 3 ค่าของดัชนีอัตราส่วนเปอร์เซ็นต์เชิงบวก ที่ได้จากการตอบข้อสอบ 6 ข้อ
ของผู้สอบ 10 คน

ผู้สอบ	ข้อสอบ												คะแนนรวม
	2		6		4		1		5		3		
	+	-	+	-	+	-	+	-	+	-	+	-	
ง	1		1		1		(0)		1		1		5
จ	1		1		1		1		1			0	5
ข	1		1		1		1			0		0	4
ฉ	1		1		1		0		(1)			0	4
ช	1		1		1		0		0			0	3
ญ		(0)	1		1		(1)		0			0	3
ซ	1			(0)		0		0	0		(1)		2
ฉ	1		1			0		0	0			0	2
ค	1			0		0		0	0			0	1
ก		(0)		0		0		(1)		0		0	1
# right	8		7		6		4		3		2		
# wrong	2		3		4		6		7		8		
E_i	1		1		0		3		1			1	
R_i	.90		.90		1.00		.70		.90			.90	
MR_i	.80		.70		.60		.60		.70			.80	
PP_i	.10		.20		.40		.10		.20			.10	
PPR_i	.50		.67		1.00		.25		.67			.50	

หมายเหตุ (0) หรือ (1) คือคะแนนรายข้อที่มีความคลาดเคลื่อน

-- คือเส้นแบ่งเขตของคะแนนรายข้อตามกัณฑ์สมบูรณ; MR_i คือค่าของดัชนีผลผลิตใหม่ต่ำสุดของข้อ i

E_i คือจำนวนคะแนนรายข้อที่มีความคลาดเคลื่อนของข้อ i ; $PP_i = R_i - MR_i$

R_i คือค่าดัชนีผลผลิตใหม่ของข้อ i ; PPR_i คือค่าดัชนีอัตราส่วนเปอร์เซ็นต์เชิงบวกของข้อสอบข้อ i

จากข้อมูลตัวอย่างที่เสนอในตารางที่ 3 จะเห็นได้ว่าเป็นการประมาณค่าอัตราส่วนเปอร์เซ็นต์เชิงบวกของแฉกล้านนั้น นอกจากจะมีการเรียงลำดับข้อสอบและผู้สอบแล้ว ยังต้องแยกสดมภ์คะแนนรายข้อที่ถูกต้อง (1) กับคะแนนรายข้อที่ผิด (0) ของข้อสอบแต่ละข้อด้วย เพื่อจะแยกจำนวนผู้สอบที่ตอบข้อสอบได้ถูกต้องกับจำนวนผู้สอบที่ตอบผิดของข้อสอบแต่ละข้อให้ชัดเจนสำหรับนำมาคำนวณค่าดัชนีผลผลิตใหม่ต่ำสุด ส่วนใหญ่ค่าของดัชนีผลผลิตใหม่รายข้อจะมีค่ามากกว่าค่าของดัชนีอัตราส่วนเปอร์เซ็นต์เชิงบวกเสมอ ยกเว้นเมื่อข้อสอบข้อนั้นมีความสอดคล้องสมบูรณ์ตามกัตแมน ค่าของดัชนีทั้งสองจะมีค่าเท่ากันและเท่ากับ 1.0 ดังตัวอย่างข้อสอบข้อที่ 4 สำหรับค่าดัชนีผลผลิตใหม่ต่ำสุดของแบบสอบฉบับนี้มีค่าเท่ากับ .70 (.70 ได้มาจาก $1 - \frac{(8+7+6+6+7+8)}{(10 \times 6)}$) และค่าของดัชนีอัตราส่วนเปอร์เซ็นต์เชิงบวกของแบบสอบทั้งฉบับ มีค่าเท่ากับ .61 (.61 ได้มาจาก $(.88 - .70)/(1 - .70)$)

2.2 นิสัยและค่าของจุดตัดของดัชนี

จุดเด่นของดัชนีอัตราส่วนเปอร์เซ็นต์เชิงบวกของแฉกล้านคือ มีค่าต่ำสุด และค่าสูงสุดแน่นอน ไม่ว่าแบบสอบฉบับนั้นจะประกอบด้วยข้อสอบที่มีค่าความยากเท่าใดก็ตาม นิสัยของดัชนีมีค่าอยู่ระหว่าง 0 ถึง 1.0 โดยที่ดัชนีจะมีค่าเท่ากับ 0 เมื่อค่าของดัชนีผลผลิตใหม่เท่ากับค่าของดัชนีผลผลิตใหม่ต่ำสุด ซึ่งหมายความว่าค่าความยากของข้อสอบได้มาจากการที่ผู้สอบที่มีคะแนนรวมน้อยตอบข้อดังกล่าวได้ถูกต้อง (คะแนนรายข้อของผู้สอบทั้งกลุ่มไม่มีความสอดคล้องกับค่าความยากของข้อสอบข้อดังกล่าว) ในทางตรงกันข้ามดัชนีจะมีค่าเท่ากับ 1.0 เมื่อค่าของดัชนีผลผลิตใหม่มีค่าสูงสุดคือ มีค่าเท่ากับ 1.0 ซึ่งหมายความว่าค่าความยากของข้อสอบได้มาจากการที่ผู้สอบที่มีคะแนนรวมสูงตอบข้อดังกล่าวได้ถูกต้อง (คะแนนรายข้อของผู้สอบทั้งกลุ่มมีความสอดคล้องกับค่าความยากของข้อสอบข้อดังกล่าว)

แฉกล้านเสนอแนะว่า ข้อสอบหรือแบบสอบใดๆก็ตามจะถือว่ามิระดับของความสอดคล้องที่ยอมรับได้ จะต้องมีการอัตราส่วนเปอร์เซ็นต์เชิงบวกตั้งแต่ 0.7 ขึ้นไป (White and Saltz 1957: 87)

3. ดัชนีความเป็นเอกพันธ์ของโลวิงเจอร์ (Loevinger's Index of Homogeneity:

H)

การพัฒนาดัชนีเพื่อวัดความสอดคล้องของคะแนนรายข้อของผู้สอบทั้งกลุ่มกับค่าความยากของโลวิงเจอร์นี้แตกต่างไปจากแนวของกัตแมน โดยการนิยามความเป็นเอกพันธ์ของแบบสอบ

ในเทอมของความน่าจะเป็น (probability) คือแบบสอบฉบับใดๆก็ตามที่มีการเรียงลำดับที่ของข้อสอบจากข้อที่ง่ายที่สุดไปจนถึงข้อที่ยากที่สุด จะเป็นแบบสอบที่มีความเป็นเอกพันธ์สมบูรณ์ (perfectly homogeneous test) ก็ต่อเมื่อข้อสอบทุกข้อที่มีลำดับที่ก่อนข้อสอบข้อที่ผู้สอบตอบได้ถูกต้องนั้น จะต้องมีความน่าจะเป็นในการตอบข้อดังกล่าวได้ถูกต้องเท่ากับ 1.0 ในทางตรงกันข้ามแบบสอบฉบับใดๆก็ตามจะเป็นแบบสอบที่มีความเป็นวิวิธพันธ์สมบูรณ์ (perfectly heterogeneous test) ก็ต่อเมื่อความน่าจะเป็นของการตอบข้อสอบที่มีลำดับที่ก่อนข้อที่ผู้สอบตอบได้ถูกต้องมีค่าเท่ากับ หรือไม่เท่ากับความน่าจะเป็นของข้อที่เขตอบได้ถูกต้อง

ดัชนีความเป็นเอกพันธ์ของโลวินเจอร์นี้พัฒนาขึ้น เพื่อบ่งชี้ความเป็นเอกพันธ์ของแบบสอบโดยมีหลักการคือ เมื่อเรียงข้อสอบจากข้อที่ง่ายที่สุดไปจนถึงข้อที่ยากที่สุดแล้ว โลวินเจอร์เสนอให้คำนวณหาสัดส่วนของผู้สอบที่ตอบข้อสอบข้อที่ i และข้อที่ j ได้ถูกต้องทั้งคู่ (P_{ij}) แล้วนำค่าสัดส่วนที่ได้นี้ ลบด้วยผลคูณของสัดส่วนของผู้ที่ตอบข้อสอบข้อที่ i ได้ถูกต้อง กับสัดส่วนของผู้ที่ตอบข้อสอบข้อที่ j ได้ถูกต้อง ($P_i P_j$) ซึ่งผลคูณนี้ก็คือค่าสัดส่วนเชิงทฤษฎีที่ผู้สอบจะตอบข้อสอบข้อที่ i และข้อที่ j ได้ถูกต้องทั้งคู่ เมื่อข้อสอบทั้ง 2 ข้อมีความเป็นอิสระจากกัน และทำเช่นนี้กับทุกๆ คู่ของข้อสอบ ก็จะได้ผลรวมของค่าความแตกต่างระหว่างสัดส่วนของผู้สอบที่ตอบข้อสอบทั้ง 2 ข้อได้ถูกต้องตามข้อมูลจริงกับทางทฤษฎี เมื่อข้อสอบทั้ง 2 มีความเป็นอิสระจากกัน และเรียกผลรวมดังกล่าวนี้ว่าค่าเอส (S) เพราะฉะนั้นค่าเอสจึงคำนวณได้จากสูตรต่อไปนี้ (White and Saltz 1957: 87)

$$S = \sum_{i=1}^{K-1} \sum_{j=i+1}^K P_{ij} - P_i P_j$$

โดยที่

i และ j คือ ลำดับที่ของข้อสอบที่เรียงจากข้อที่ง่ายที่สุดไปจนถึงข้อที่ยากที่สุด ; $i = 1, 2, \dots, K-1$ และ $j = 2, 3, \dots, K$

P_{ij} คือ ค่าสัดส่วนของผู้สอบที่ตอบข้อสอบข้อที่ i และข้อที่ j ได้ถูกต้องทั้งคู่

P_i คือ ค่าสัดส่วนของผู้สอบที่ตอบข้อสอบข้อที่ i ได้ถูกต้อง

P_j คือ ค่าสัดส่วนของผู้สอบที่ตอบข้อสอบข้อที่ j ได้ถูกต้อง

$P_i P_j$ คือ ผลคูณของค่าสัดส่วนของผู้สอบที่ตอบข้อสอบข้อที่ i ได้ถูกต้อง

กับค่าสัดส่วนของผู้สอบที่ตอบข้อสอบข้อที่ j ได้ถูกต้อง เมื่อข้อสอบข้อที่ i และข้อที่ j มีความเป็นอิสระจากกัน

ผลลัพธ์ของค่าเอส ค่าเอสจะมีค่าต่ำสุดเท่ากับ 0 เมื่อค่า $P_{ij} = P_i P_j$ ซึ่งหมายความว่าข้อสอบแต่ละข้อของแบบสอบเป็นอิสระจากกัน นั่นคือแบบสอบมีความเป็นวิวิธพันธุ์สมบูรณ์ สำหรับค่าสูงสุดของค่าเอส (S_{max}) จะมีค่าเท่ากับค่าสัดส่วนของผู้ที่ตอบข้อสอบข้อที่ j ได้ถูกต้อง (P_j) ลบด้วยผลคูณของค่าสัดส่วนของผู้ที่ตอบข้อสอบข้อที่ i ได้ถูกต้องกับค่าสัดส่วนของผู้ที่ตอบข้อสอบข้อที่ j ได้ถูกต้อง ($P_i P_j$) ของข้อสอบแต่ละคู่ เพราะฉะนั้นค่าสูงสุดของค่าเอสจึงคำนวณได้จากสูตรต่อไปนี้ (White and Saltz 1957: 87)

$$S_{max} = \sum_{i=1}^{K-1} \sum_{j=i+1}^K P_j - P_i P_j$$

โดยที่

i และ j คือ ลำดับที่ของข้อสอบที่เรียงจากข้อที่ง่ายที่สุดไปจนถึงข้อที่ยากที่สุด ; $i = 1, 2, \dots, K-1$ และ $j = 2, 3, \dots, K$

P_i คือ ค่าสัดส่วนของผู้สอบที่ตอบข้อสอบข้อที่ i ได้ถูกต้อง

P_j คือ ค่าสัดส่วนของผู้สอบที่ตอบข้อสอบข้อที่ j ได้ถูกต้อง

$P_i P_j$ คือ ผลคูณของค่าสัดส่วนของผู้สอบที่ตอบข้อสอบข้อที่ i ได้ถูกต้องกับค่าสัดส่วนของผู้สอบที่ตอบข้อสอบข้อที่ j ได้ถูกต้อง เมื่อข้อสอบข้อที่ i และข้อที่ j มีความเป็นอิสระจากกัน

จากค่าสูงสุดของเอสจะเห็นว่า เมื่อค่า $P_{ij} = P_i P_j$ แสดงว่าแบบสอบมีความเป็นเอกพันธุ์สมบูรณ์ตามนิยามที่โลวินเจอร์กำหนดขึ้น ทั้งนี้เพราะการที่ผู้สอบตอบข้อสอบข้อที่ j ได้ถูกต้อง แสดงว่าความน่าจะเป็นในการตอบข้อสอบข้อที่ i ได้ถูกต้องเท่ากับ 1.0

สำหรับดัชนีความเป็นเอกพันธุ์ของโลวินเจอร์นั้น คำนวณได้จากค่าเอสหารด้วยค่าเอสสูงสุด ดังสูตรคำนวณต่อไปนี้ (White and Saltz 1957: 87)

$$H_c = \frac{S}{S_{max}}$$

ผลลัพธ์ของค่าดัชนีความเป็นเอกพันธุ์ของโลวินเจอร์ มีค่าอยู่ระหว่าง 0 ถึง 1.0 โดย

มีค่าเท่ากับ 0 เมื่อค่าเอสเท่ากับ 0 ซึ่งหมายความว่าแบบสอบถามประกอบด้วยข้อสอบที่มีความเป็นอิสระจากกัน ในทางตรงกันข้ามดัชนีจะมีค่าเท่ากับ 1.0 เมื่อค่าเอสมีค่าเท่ากับค่าสูงสุดของมัน ซึ่งหมายความว่าแบบสอบถามมีความเป็นเอกพันธ์สมบูรณ์ ถ้าแบบสอบถามมีความเป็นเอกพันธ์สมบูรณ์แล้ว จะเห็นได้ว่าแบบแผนการตอบข้อสอบ K ข้อ ของผู้สอบ N คน ย่อมมีลักษณะเดียวกับกัตแมนสมบรูณ์นั่นเอง ดังนั้นค่าสูงสุดของดัชนีความเป็นเอกพันธ์จึงมีค่าเท่ากับค่าสูงสุดของดัชนีผลผลิตใหม่ของแบบสอบถาม และก็มีค่าเท่ากับค่าสูงสุดของดัชนีอัตราส่วนเปอร์เซ็นต์เชิงบวกของแจคสันด้วย

4. ดัชนีความสอดคล้องของแบบสอบ (Index of Consistency: I)

กรีน (Green) ได้พัฒนาดัชนีความสอดคล้องของแบบสอบขึ้น โดยนำหลักการและจุดเด่นของแจคสันและโลวินเจอร์มาใช้โดยปรับให้การคำนวณง่ายขึ้น การคำนวณหาค่าดัชนีความสอดคล้องของแบบสอบจะเหมือนกับดัชนีอัตราส่วนเปอร์เซ็นต์เชิงบวก

$$I = \frac{Rep - Rep_{ind}}{1 - Rep_{ind}}$$

โดยที่

I คือ ดัชนีความสอดคล้องของแบบสอบของกรีน

Rep คือ ค่าดัชนีผลผลิตใหม่ของกรีน

Rep_{ind} คือ ค่าดัชนีผลผลิตใหม่ที่คำนวณจากชุดข้อสอบที่มีความยากเหมือนกัน และข้อสอบดังกล่าวเป็นอิสระจากกัน

การประมาณค่า Rep และ Rep_{ind} จะได้จากข้อสอบที่มีลำดับที่ติดกัน (adjacent) คล้ายกับของโลวินเจอร์ แต่ไม่ต้องจับคู่ข้อสอบทุกคู่ ค่า Rep และ Rep_{ind} คำนวณได้จากสูตรต่อไปนี้

$$Rep = 1 - \frac{1}{NK} \sum_{i=1}^{K-1} n_{i, i+1} - \frac{1}{NK} \sum_{i=2}^K n_{i-1, i, i+1, i+2}$$

และ

$$Rep_{ind} = 1 - \frac{1}{N^2 K} \sum_{i=1}^{K-1} n_i n_{i+1} - \frac{1}{N^4 K} \sum_{i=2}^K n_{i-1} n_{i+1} n_{i+2} n_{i-1}$$

โดยที่

i คือ ลำดับที่ของข้อสอบเรียงจากข้อที่ง่ายที่สุดไปจนถึงข้อที่ยากที่สุด: $i = 1, 2, 3, \dots, K-2, K-1, K$

N คือ จำนวนผู้สอบทั้งหมด

K คือ จำนวนข้อสอบทั้งหมด

$n_{i, i+1}$ คือ จำนวนผู้สอบที่ตอบข้อสอบข้อที่ i ผิด แต่ตอบข้อสอบข้อที่ $i+1$ ได้

ถูกต้อง

$n_{i-1, i, i+1, i+2}$ คือ จำนวนผู้สอบที่ตอบข้อสอบข้อที่ $i-1$ และข้อที่ i ผิด แต่ตอบข้อสอบข้อที่ $i+1$ และข้อที่ $i+2$ ได้ถูกต้อง

n_{i-1} คือ จำนวนผู้สอบที่ตอบข้อสอบข้อที่ $i-1$ ผิด

n_{i-1} คือ จำนวนผู้สอบที่ตอบข้อสอบข้อที่ $i-1$ ผิด

n_{i+1} คือ จำนวนผู้สอบที่ตอบข้อสอบข้อที่ $i+1$ ได้ถูกต้อง

n_{i+2} คือ จำนวนผู้สอบที่ตอบข้อสอบข้อที่ $i+2$ ได้ถูกต้อง

ตารางที่ 4 ซึ่งจะนำเสนอต่อไปนี้เป็นตัวอย่างของการคำนวณค่าของดัชนีความสอดคล้องของแบบสอบ จากการตอบข้อสอบ 6 ข้อ ของผู้สอบ 10 คน ซึ่งเป็นข้อมูลเดียวกับตารางที่ 2

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 4 ตัวอย่างการคำนวณค่าดัชนีความสอดคล้องของแบบสอบของกรีน จากการตอบข้อสอบ 6 ข้อ ของผู้สอบ 10 คน

ผู้สอบ	ข้อสอบ						คะแนนรวม
	2	6	4	1	5	3	
ง	1	1	1	0	1	1	5
จ	1	1	1	1	1	0	5
ช	1	1	1	1	0	0	4
ฉ	1	1	1	0	1	0	4
ซ	1	1	1	0	0	0	3
ญ	0	1	1	1	0	0	3
ช	1	0	0	0	0	1	2
ฉ	1	1	0	0	0	0	2
ค	1	0	0	0	0	0	1
ก	0	0	0	1	0	0	1
จำนวนผู้ตอบถูก (n_i)	8	7	6	4	3	2	
จำนวนผู้ตอบผิด ($n_{i'}$)	2	3	4	6	7	8	
ลำดับที่ของข้อสอบ	1	2	3	4	5	6	
$n_{i, i+1}$		1	0	1	2	1	
$n_{i-1, i, i+1, i+2}$			-	0	0	0	-

$$Rep = 1 - \frac{1}{10 \times 6} (1 + 0 + 1 + 2 + 1) - \frac{1}{10 \times 6} (0 + 0 + 0) = 0.917$$

$$Rep_{ind} = 1 - \frac{1}{10^2 \times 6} \{ (7)(2) + (6)(3) + (4)(4) + (3)(6) + (2)(7) \} \\ - \frac{1}{10^4 \times 6} \{ (4)(6)(3)(2) + (3)(4)(4)(3) + (2)(3)(6)(4) \} = .860$$

$$I = (.917 - .860) / (1 - .860) = .407$$

จากข้อมูลที่เสนอในตารางที่ 4 แสดงให้เห็นถึงวิธีการหาความคลาดเคลื่อนของข้อมูล เมื่อพิจารณาเปรียบเทียบระหว่างคู่หรือชุดของข้อสอบที่มีลำดับที่ติดกัน วิธีนี้นับว่าสะดวกกว่าของโลวินเจอร์ตรงที่ไม่ต้องจับคู่ข้อสอบทุกคู่ สำหรับนิสัยของดัชนีความเป็นเอกพันธ์จะมีค่าอยู่ระหว่าง 0 ถึง 1.0 โดยมีค่าเป็น 0 เมื่อค่า Rep เท่ากับ Rep_{ind} ซึ่งหมายความว่าแบบสอบประกอบด้วยข้อสอบที่เป็นอิสระจากกัน นั่นคือแบบสอบมีความเป็นวิวิธพันธุ์สมบูรณ์ เมื่อดัชนีมีค่าเป็น 1.0 หมายความว่าคะแนนรายชื่อในเมตริกการตอบ (response matrix) มีลักษณะเป็นกักตมสมบูรณ์ แบบสอบจึงมีความเป็นเอกพันธุ์สมบูรณ์หรือที่กรีน เรียกว่ามีความสอดคล้องสมบูรณ์ กรีนเสนอแนะว่าแบบสอบที่มีค่าดัชนีความสอดคล้องตั้งแต่ 0.5 ขึ้นไปจึงจะถือว่ามีความสอดคล้องที่ยอมรับได้

การประมาณค่าความสอดคล้องของคะแนนรายชื่อกับค่าความยากของข้อสอบ ทั้ง 4 แนวคิดที่ได้กล่าวมาแล้วนั้น จะเห็นได้ว่าความหมายของ ค่าดัชนีผลผลิตใหม่ ค่าดัชนีอัตราส่วนเปอร์เซ็นต์เชิงบวก ค่าดัชนีความเป็นเอกพันธ์ และค่าดัชนีความสอดคล้องนั้น เมื่อดัชนีมีค่าสูงสุดแล้วคะแนนรายชื่อที่ได้จากการตอบข้อสอบ K ข้อ ของผู้สอบ N คน จะมีลักษณะเหมือนกัน คือ เป็นกักตมสมบูรณ์ และจะพบว่าข้อมูลลักษณะดังกล่าวจะให้ค่าความแปรปรวนร่วมระหว่างข้อสอบ (interitem covariances) และค่าความแปรปรวนของคะแนนรวม (total scores variance) สูงสุด ซึ่งหมายความว่าค่าความเที่ยงของแบบสอบที่ประมาณด้วยวิธีความสอดคล้องภายใน (internal consistency) จะมีค่าสูงสุดเช่นกัน

จากข้อมูลการตอบข้อสอบ 6 ข้อ ของผู้สอบ 10 คน ตามตารางที่ 4 จะให้ค่าดัชนีความสอดคล้องของคะแนนรายชื่อกับค่าความยากของข้อสอบเปรียบเทียบได้ดังนี้ คือ ค่าดัชนีผลผลิตใหม่ของแบบสอบเท่ากับ .88 ค่าดัชนีอัตราส่วนเปอร์เซ็นต์เชิงบวกของแบบสอบเท่ากับ .61 ค่าดัชนีความเป็นเอกพันธ์เท่ากับ .23 และค่าดัชนีความสอดคล้องของแบบสอบของกรีนเท่ากับ .41 ในขณะที่ค่าความสอดคล้องภายในของแบบสอบที่ประมาณด้วยวิธีความสอดคล้องภายใน มีค่าเท่ากับ 0.47 สำหรับข้อมูลชุดนี้เมื่อกำหนดให้ค่าความยากของข้อสอบแต่ละข้อเท่าเดิม แต่มีแบบแผนการตอบข้อสอบรายบุคคลเป็นกักตมสมบูรณ์ จะทำให้ค่าดัชนีความสอดคล้องภายในมีค่าเท่ากับ 0.83 ที่เป็นเช่นนั้นเพราะค่าความเที่ยงของแบบสอบขึ้นอยู่กับค่าความยากของข้อสอบนั่นเอง

จะเห็นได้ว่าพัฒนาการของดัชนีความสอดคล้องของคะแนนรายชื่อกับค่าความยากของข้อสอบที่ได้กล่าวมาแล้วทั้งหมด เป็นการพิจารณาความสอดคล้องของข้อสอบเป็นรายชื่อ และทั้งของแบบสอบทั้งฉบับ ซึ่งสามารถสรุปข้อดีของดัชนีแต่ละตัวได้ดังนี้คือ

1. มีสูตรที่คำนวณได้ทั้งความสอดคล้องของข้อสอบรายข้อ และของแบบสอบทั้งฉบับ

ได้แก่ ดัชนีผลผลิตใหม่ และดัชนีอัตราส่วนเปอร์เซ็นต์เชิงบวก

2. มีค่าพิสัยคงที่ไม่แปรผันตามค่าความยากของข้อสอบ ได้แก่ ดัชนีอัตราส่วนเปอร์เซ็นต์เชิงบวก ดัชนีความเป็นเอกพันธ์ และดัชนีความสอดคล้อง ซึ่งดัชนีทั้ง 3 มีพิสัยอยู่ในช่วง 0 กับ 1.00

3. คำนิยามถึงความคลาดเคลื่อนในการตอบข้อสอบข้อยากได้ถูกต้อง และการตอบข้อสอบข้อง่ายผิดได้แก่ ดัชนีผลผลิตใหม่ ดัชนีอัตราส่วนเปอร์เซ็นต์เชิงบวก และดัชนีความสอดคล้อง

4. สูตรคำนวณไม่ยุ่งยาก ได้แก่ ดัชนีผลผลิตใหม่ และดัชนีอัตราส่วนเปอร์เซ็นต์เชิงบวก

5. ไม่ต้องกำหนดเส้นแบ่งเขตคะแนนรายข้อตามกัตแมน ได้แก่ ดัชนีความเป็นเอกพันธ์ และดัชนีความสอดคล้อง

ซึ่งหลักการและข้อดีของดัชนีดังกล่าวมาแล้วทั้ง 5 ข้อ เป็นประเด็นที่ผู้วิจัยนำมาพิจารณาประกอบการพัฒนาดัชนีของผู้วิจัย สำหรับนิยามของกัตแมนสมบูรณ์ และลักษณะการเรียงลำดับที่ของข้อสอบและผู้สอบตามเมตริกกัตแมนนั้น ได้มีผู้นำมาประยุกต์ใช้วิเคราะห์แบบแผนการตอบข้อสอบรายบุคคลเพื่อใช้เป็นข้อมูลประกอบการตัดสินใจเกี่ยวกับคะแนนรวม (total score) ของผู้สอบ

ตอนที่ 2 ดัชนีความสอดคล้องแบบแผนการตอบข้อสอบรายบุคคล (Indices of an Individual's Item Response Pattern)

การวิเคราะห์แบบแผนการตอบข้อสอบรายบุคคลเป็นประเด็นที่ได้รับความนิยมในช่วง 10 ปีที่ผ่านมา ผู้เชี่ยวชาญหลายท่านที่ให้ความสนใจกับการตอบถูกหรือการตอบผิดของข้อสอบแต่ละข้อจากแบบสอบที่วัดคุณลักษณะเดียวกัน และสามารถจัดเรียงข้อสอบดังกล่าวจากข้อที่ง่ายที่สุด (ข้อที่มีจำนวนผู้สอบภายในกลุ่มตอบได้ถูกต้องมากที่สุด) ไปจนถึงข้อที่ยากที่สุด (ข้อที่มีจำนวนผู้สอบภายในกลุ่มตอบได้ถูกต้องน้อยที่สุด) ส่วนคะแนนรายข้อของข้อสอบทั้งหมดที่นำมาจัดเรียงลำดับกันตามความยากของข้อสอบนี้เรียกว่า แบบแผนการตอบข้อสอบ (item response pattern) และมีความเชื่อว่าถ้าคะแนนรวมที่ผู้สอบได้มาจากการตอบแบบสอบเลือกตอบนั้น เป็นคะแนนที่เขาทำได้อย่างมีความรู้จริงแล้ว จะต้องได้มาจากการตอบข้อสอบข้อที่ง่ายที่สุดไปจนถึงข้อสอบข้อที่มีลำดับที่ต่ำกว่ากับคะแนนรวมได้ถูกต้องทุกข้อและผู้สอบที่มีคะแนนรวมเท่ากันควรจะมีแบบแผนการตอบข้อสอบเหมือนกัน คือเป็นแบบสอบที่มีลักษณะเป็นกัตแมนสมบูรณ์เหมือนกัน แต่ในสถานที่เป็นจริงมักจะพบว่าผู้สอบที่มี

คะแนนรวมเท่ากันกลับมีแบบแผนการตอบข้อสอบที่แตกต่างกัน ยิ่งเมื่อนำหลักทางคณิตศาสตร์มาวิเคราะห์แบบแผนการตอบข้อสอบที่เป็นไปได้จะพบว่าข้อสอบ K ข้อ จะเกิดแบบแผนการตอบข้อสอบถึง 2^K แบบ (Birenbaum 1985: 523) สำหรับดัชนีความสอดคล้องของแบบแผนการตอบข้อสอบจะบอกถึงระดับความสอดคล้องของแบบแผนการตอบข้อสอบที่ผู้สอบทำได้กับแบบแผนการตอบข้อสอบที่มีลักษณะเป็นกัตแมนสมบูรณ์ โดยมีรายละเอียดของดัชนีต่างๆดังต่อไปนี้คือ

1. ดัชนีพอยท์ไบซีเรียลของแต่ละบุคคล (Personal Point Biserial Index: r_i^*)

ในปี ค.ศ. 1980 เบรนแนน (Brennan 1980) ได้พัฒนาดัชนีพอยท์ไบซีเรียลของแต่ละบุคคลขึ้น เพื่อบอกระดับความสอดคล้อง (agreement) ระหว่างแบบแผนการตอบข้อสอบกับค่าความยากของข้อสอบ ค่าของดัชนีคำนวณได้จากค่าสหสัมพันธ์พอยท์ไบซีเรียลระหว่างคะแนนรายข้อกับค่าความยากของข้อสอบข้อนั้น โดยมีสูตรคำนวณดังนี้ (Harnisch and Linn 1981: 136)

$$r_i^* = \frac{\bar{P}_i - \bar{P}_R}{S_P} \sqrt{\frac{P_R}{1 - P_R}}$$

โดยที่

r_i^* คือ ค่าดัชนีพอยท์ไบซีเรียลของผู้สอบคนที่ i

\bar{P}_i คือ ค่าเฉลี่ยของค่าความยากของข้อสอบทุกข้อที่ผู้สอบคนที่ i ตอบได้ถูกต้อง

\bar{P}_R คือ ค่าเฉลี่ยของค่าความยากของแบบสอบทั้งฉบับ

S_P คือ ค่าส่วนเบี่ยงเบนมาตรฐานของค่าความยากของแบบสอบทั้งฉบับ

P_R คือ ค่าสัดส่วนของจำนวนข้อสอบที่ผู้สอบคนที่ i ตอบได้ถูกต้อง

พิสัยของค่าดัชนีพอยท์ไบซีเรียลของแต่ละบุคคล มีค่าอยู่ระหว่าง -1.0 ถึง 1.0

โดยที่ดัชนีจะมีค่าเข้าใกล้ 1.0 เมื่อผู้สอบมีแบบแผนการตอบข้อสอบที่มีลักษณะใกล้เคียงกับกัตแมนสมบูรณ์ และดัชนีจะมีค่าเท่ากับ 1.0 เมื่อแบบแผนการตอบข้อสอบของผู้สอบมีลักษณะเป็นกัตแมนสมบูรณ์ ในทางตรงกันข้ามดัชนีจะมีค่าเข้าใกล้ -1.0 เมื่อผู้สอบมีแบบแผนการตอบที่มีลักษณะแตกต่างจากกัตแมนสมบูรณ์ และดัชนีมีค่าเท่ากับ -1.0 เมื่อแบบแผนการตอบข้อสอบของผู้สอบมีลักษณะตรงกันข้ามกับกัตแมนสมบูรณ์

2. ดัชนีไบบีเรียลของแต่ละบุคคล (Personal Biserial Index: r_i)

ในปี ค.ศ. 1968 ดอนลอนและฟิสเชอร์ (Donlon and Fischer 1968) ได้พัฒนาดัชนีไบบีเรียลของแต่ละบุคคลขึ้น เพื่อบอกความสัมพันธ์ระหว่างคะแนนรายข้อกับค่าความยากของข้อสอบข้อนั้นจากค่าสหสัมพันธ์ไบบีเรียลโดยมีสูตรคำนวณดังนี้ (Donlon and Fischer 1968: 107)

$$r_i = \frac{\bar{P}_i - \bar{P}_R}{S_P} \sqrt{\frac{P_R}{u}}$$

โดยที่

r_i คือ ค่าดัชนีไบบีเรียลของแต่ละบุคคลของผู้สอบคนที่ i

\bar{P}_i คือ ค่าเฉลี่ยของค่าความยากของข้อสอบทุกข้อที่ผู้สอบคนที่ i ตอบได้ถูกต้อง

\bar{P}_R คือ ค่าเฉลี่ยของค่าความยากของแบบสอบทั้งฉบับ

S_P คือ ค่าส่วนเบี่ยงเบนมาตรฐานของค่าความยากของแบบสอบทั้งฉบับ

P_R คือ ค่าสัดส่วนระหว่างจำนวนข้อสอบที่ผู้สอบคนที่ i ตอบได้ถูกต้อง

u คือ ค่าอติเนต (ordinate) ของการแจกแจงแบบโค้งปกติมาตรฐาน ซึ่ง
เป็นค่าที่แบ่งพื้นที่ของโค้งปกติมาตรฐานเป็น P_R และ $1 - P_R$

นิสัยของค่าดัชนีไบบีเรียลของแต่ละบุคคล มีค่าโดยประมาณอยู่ระหว่าง -1.0 ถึง 1.0 โดยที่ดัชนีจะมีค่าเข้าสู่ 1.0 เมื่อผู้สอบมีแบบแผนการตอบข้อสอบที่มีลักษณะใกล้เคียงกับกัตแมนสมบูรณ์ ในทางตรงกันข้ามดัชนีจะมีค่าเข้าสู่ -1.0 เมื่อผู้สอบมีแบบแผนการตอบที่มีลักษณะตรงกันข้ามกับกัตแมนสมบูรณ์

3. ดัชนีความสอดคล้อง (The Agreement Index: A_i)

ในปี ค.ศ. 1980 เคน และ เบนแนน (Kane and Brennan 1980) ได้พัฒนาดัชนีขึ้น เพื่อบอกความสอดคล้องของแบบแผนการตอบข้อสอบที่ผู้สอบทำได้กับแบบแผนการตอบข้อสอบที่มีลักษณะเป็นกัตแมนสมบูรณ์ โดยที่ดัชนีมีสูตรคำนวณดังนี้ (Harnisch and Linn 1981: 137)

$$A_i = \frac{J}{\sum_{j=1}^J U_{ij} P_j}$$

โดยที่

A_i คือ ค่าดัชนีความสอดคล้องของผู้สอบคนที่ i

j คือ ลำดับที่ของข้อสอบที่เรียงจากข้อที่ง่ายที่สุดไปจนถึงข้อที่ยากที่สุด; $j =$

1, 2, 3, ..., J

U_{ij} คือ คะแนนรายข้อของข้อสอบข้อที่ j ของผู้สอบคนที่ i โดยที่ U_{ij} มีค่าเท่ากับ 1 เมื่อผู้สอบตอบได้ถูกต้อง และ U_{ij} มีค่าเท่ากับ 0 เมื่อผู้สอบตอบผิด

P_j คือ ค่าความยากของข้อสอบข้อที่ j

นิลย์ของดัชนีความสอดคล้อง สำหรับผู้สอบที่มีคะแนนรวม n_i คะแนนนั้น ดัชนีจะมีค่าสูงสุด ($\max A_i$) เท่ากับ $\sum_{j=1}^J P_j$ เมื่อผู้สอบมีแบบแผนการตอบข้อสอบที่มีลักษณะเป็นกักตแมนสมบูรณ์ ในทางตรงกันข้ามดัชนีจะมีค่าต่ำสุด ($\min A_i$) เท่ากับ $\sum_{j=n_i+1}^J P_j$ เมื่อผู้สอบมีแบบแผนการตอบข้อสอบที่มีลักษณะที่ตรงกันข้ามกักตแมนสมบูรณ์

เนื่องจากค่าสูงสุดของดัชนีความสอดคล้องมีค่ามากน้อยตามคะแนนรวม จึงทำให้เกิดปัญหาต่อการแปลความหมายของค่าดัชนีเมื่อผู้สอบมีคะแนนรวมแตกต่างกัน โดยเฉพาะเมื่อผู้สอบที่มีคะแนนรวมน้อยแต่มีแบบแผนการตอบข้อสอบที่มีลักษณะเป็นกักตแมนสมบูรณ์ กลับมีค่าของดัชนีน้อยกว่าของผู้สอบที่มีคะแนนรวมสูงกว่าแต่มีแบบแผนการตอบข้อสอบที่แตกต่างจากลักษณะกักตแมนสมบูรณ์ ด้วยเหตุนี้ เคน และ เบรินแนน จึงได้พัฒนาดัชนีความไม่สอดคล้อง ดัชนีความเชื่อถือได้ และดัชนีความเชื่อถือได้ที่ปรับแก้การเดาขึ้น ดังที่จะเสนอต่อไป

4. ดัชนีความไม่สอดคล้อง (The Disagreement Index: D_i)

เคน และ เบรินแนน ได้นำค่าของดัชนีความสอดคล้องที่คำนวณได้จากแบบแผนการตอบข้อสอบของผู้สอบ ไปลบออกจากค่าสูงสุดของดัชนี ในกรณีที่ผู้สอบคนดังกล่าวมีแบบแผนการตอบข้อสอบที่มีลักษณะเป็นกักตแมนสมบูรณ์ ค่าของความแตกต่างนี้จึงบอกถึงความแตกต่างระหว่างแบบแผนการตอบข้อสอบที่ผู้สอบทำได้กับที่มีลักษณะเป็นกักตแมนสมบูรณ์ และเรียกค่าความแตกต่างนี้ว่าค่าความไม่สอดคล้อง ซึ่งมีสูตรคำนวณ คือ (Harnisch and Linn 1981: 138)

$$D_i = \max A_i - A_i$$

โดยที่

0_i คือ ค่าดัชนีความไม่สอดคล้องของผู้สอบคนที่ i

$\max A_i$ คือ ค่าสูงสุดของดัชนีความสอดคล้องของผู้สอบคนที่ i ในกรณีที่เขามีแบบแผนการตอบข้อสอบที่มีลักษณะเป็นกัตแมนสมบูรณ์ ดังนั้นผู้สอบคนที่ i จะมีค่า $\max A_i = \sum_{j=1}^{n_i} P_j$

A_i คือ ค่าของดัชนีความสอดคล้องของผู้สอบคนที่ i

พิสัยของค่าดัชนีความไม่สอดคล้อง มีค่าอยู่ระหว่าง 0 ถึง $\max A_i - \min A_i$ โดยที่ดัชนี จะมีค่าเข้าใกล้ 0 เมื่อผู้สอบมีแบบแผนการตอบข้อสอบที่มีลักษณะใกล้เคียงกับกัตแมนสมบูรณ์ และดัชนีจะมีค่าเท่ากับ 0 เมื่อแบบแผนการตอบข้อสอบของผู้สอบมีลักษณะเป็นกัตแมนสมบูรณ์ ในทางตรงกันข้ามดัชนีจะมีค่าเข้าใกล้ $\max A_i - \min A_i$ เมื่อผู้สอบมีแบบแผนการตอบที่มีลักษณะแตกต่างจากกัตแมนสมบูรณ์ และดัชนีมีค่าเท่ากับ $\max A_i - \min A_i$ เมื่อแบบแผนการตอบข้อสอบของผู้สอบมีลักษณะตรงกันข้ามกับกัตแมนสมบูรณ์

5. ดัชนีความเชื่อถือได้ (The Dependability Index: θ_i)

นอกจากนี้ เคนและเบรนนัน ยังได้นำค่าดัชนีความสอดคล้องที่คำนวณได้จากแบบแผนการตอบข้อสอบของผู้สอบ มาหารด้วยค่าสูงสุดของดัชนี ในกรณีที่ผู้สอบมีแบบแผนการตอบข้อสอบที่มีลักษณะเป็นกัตแมนสมบูรณ์ และเรียกดัชนีตัวใหม่นี้ว่าดัชนีความเชื่อถือได้ ซึ่งมีสูตรคำนวณคือ (Harnisch and Linn 1981: 138)

$$\theta_i = \frac{A_i}{\max A_i}$$

โดยที่

θ_i คือ ค่าดัชนีความเชื่อถือได้ของผู้สอบคนที่ i

A_i คือ ค่าดัชนีความสอดคล้องของผู้สอบคนที่ i

$\max A_i$ คือ ค่าสูงสุดของดัชนีความสอดคล้องของผู้สอบคนที่ i

พิสัยของค่าดัชนีความเชื่อถือได้ มีค่าอยู่ระหว่าง $\min A_i / \max A_i$ ถึง 1.0 โดยที่ดัชนีจะมีค่าเข้าใกล้ 1.0 เมื่อผู้สอบมีแบบแผนการตอบข้อสอบที่มีลักษณะใกล้เคียงกับกัตแมนสมบูรณ์ และดัชนีจะมีค่าเท่ากับ 1.0 เมื่อแบบแผนการตอบข้อสอบของผู้สอบมีลักษณะเป็นกัตแมนสมบูรณ์ ในทางตรงกันข้ามดัชนีจะมีค่าเข้าใกล้ $\min A_i / \max A_i$ เมื่อผู้สอบมีแบบแผนการตอบที่มีลักษณะแตกต่างจากกัตแมนสมบูรณ์ และดัชนีมีค่าเท่ากับ $\min A_i / \max A_i$ เมื่อแบบแผนการตอบข้อสอบของผู้สอบมี

ลักษณะตรงกันข้ามกับกัตแมนสมบรูณ์

6. ดัชนีความเชื่อถือได้ที่ปรับแก้การเดา (The Dependability Corrected for Chance Index: θ_{ci})

เคน และ เบนแนน ได้พัฒนาดัชนีความเชื่อถือได้ที่ปรับแก้การเดาขึ้น โดยนิยามค่าคาดหวังของค่าดัชนีความสอดคล้อง ($E(A_i)$) ของผู้สอบที่มีคะแนนรวม n_i คะแนน ว่ามีโอกาสที่จะตอบข้อสอบแต่ละข้อได้ถูกต้อง เท่ากับคะแนนรวมของผู้สอบหารด้วยจำนวนข้อสอบทั้งหมด (n_i/J) เพราะฉะนั้นโอกาสที่ผู้สอบคนที่ i ซึ่งมีคะแนนรวม n_i คะแนน จะได้คะแนนจากการตอบแบบสอบฉบับหนึ่งๆ โดยการเดา จึงมีค่าเท่ากับคะแนนรวมของผู้สอบหารด้วยจำนวนข้อสอบทั้งหมด แล้วนำค่าสัดส่วนที่ได้นี้มาคูณด้วยผลรวมของค่าความยากของข้อสอบทุกข้อ (ดังนั้น ค่าโอกาสที่ผู้สอบคนที่ i จะได้คะแนนจากการตอบแบบสอบฉบับหนึ่ง โดยการเดา จึงมีค่าเท่ากับ $(n_i/J)(\sum_{j=1}^J P_j)$) และเรียกค่าโอกาสดังกล่าวว่า ค่าการเดา (chance: A_{ci}) ของผู้สอบคนที่ i ซึ่งมีคะแนนรวม n_i คะแนน จากนั้นนำค่าดังกล่าวไปใช้ในการหาค่าดัชนีความเชื่อถือได้ที่ปรับแก้การเดา ดังสูตรคำนวณต่อไปนี้ (Harnisch and Linn 1981: 138)

$$\theta_{ci} = \frac{A_i - A_{ci}}{\max A_i - A_{ci}}$$

โดยที่

θ_{ci} คือ ค่าดัชนีความเชื่อถือได้ที่ของผู้สอบคนที่ i

A_i คือ ค่าดัชนีความสอดคล้องของผู้สอบคนที่ i

$\max A_i$ คือ ค่าสูงสุดของดัชนีความสอดคล้องของผู้สอบคนที่ i

A_{ci} คือ ค่าการเดาของผู้สอบคนที่ i ซึ่งมีค่าเท่ากับคะแนนรวมหารด้วย

จำนวนข้อสอบทั้งหมด แล้วคูณด้วยผลรวมของค่าความยากของข้อสอบทุกข้อ ($A_{ci} = n_i/J(\sum_{j=1}^J P_j)$)

นิสัยของค่าดัชนีความเชื่อถือได้ที่ปรับแก้การเดา มีค่าอยู่ระหว่าง $-(A_{ci} - \min A_i) / (\max A_i - A_{ci})$ ถึง 1.0 โดยที่ดัชนีจะมีค่าเข้าใกล้ 1.0 เมื่อผู้สอบมีแบบแผนการตอบข้อสอบที่มีลักษณะใกล้เคียงกับกัตแมนสมบรูณ์ และดัชนีจะมีค่าเท่ากับ 1.0 เมื่อแบบแผนการตอบข้อสอบของผู้สอบมีลักษณะเป็นกัตแมนสมบรูณ์ ในทางตรงกันข้ามดัชนีจะมีค่าเข้าใกล้ $-(A_{ci} - \min A_i) / (\max A_i - A_{ci})$ เมื่อผู้สอบมีแบบแผนการตอบที่มีลักษณะแตกต่างจากกัตแมนสมบรูณ์ และดัชนีมีค่า

เท่ากับ $-(A_{ci} - \min A_i) / (\max A_i - A_{ci})$ เมื่อแบบแผนการตอบข้อสอบของผู้สอบมีลักษณะตรงกันข้ามกับกัตแมนสมบรูณ์

7. ดัชนีชี้เตือนของซาโต (Sato's Caution Index: C_i)

ในปี ค.ศ. 1975 ซาโต (Sato 1975) ได้พัฒนาดัชนีชี้เตือนของซาโตขึ้น โดยการคำนวณหาค่าความแปรปรวนร่วมระหว่างคะแนนรายชื่อของผู้สอบกับค่าความยากของข้อสอบหารด้วยค่าความแปรปรวนร่วมระหว่างคะแนนรายชื่อที่มีลักษณะตามกัตแมนสมบรูณ์กับค่าความยากของข้อสอบ แล้วนำค่าสัดส่วนที่ได้ขึ้นไปลบออกจาก 1.0 ดังสูตรคำนวณต่อไปนี้ (Tatsuoka 1984: 97)

$$C_i = 1 - \frac{\text{Cov}(X_{ij}, Y_j)}{\text{Cov}(U_{ij}, Y_j)}$$

โดยที่

C_i คือ ค่าดัชนีชี้เตือนของซาโตของผู้สอบคนที่ i

j คือ ลำดับที่ของข้อสอบที่เรียงจากข้อที่ง่ายที่สุดไปจนถึงข้อที่ยากที่สุด: $j = 1, 2, 3, \dots, J$

X_{ij} คือ คะแนนรายชื่อของข้อสอบข้อที่ j ของผู้สอบคนที่ i โดยที่ X_{ij} มีค่าเท่ากับ 1 เมื่อผู้สอบตอบได้ถูกต้อง และ X_{ij} มีค่าเท่ากับ 0 เมื่อผู้สอบตอบผิด

Y_j คือ ค่าความยากของข้อสอบข้อที่ j

U_{ij} คือ คะแนนรายชื่อที่มีลักษณะตามกัตแมนสมบรูณ์ของข้อสอบข้อที่ j ของผู้สอบคนที่ i โดยที่ U_{ij} มีค่าเท่ากับ 1 เมื่อข้อสอบข้อที่ j มีลำดับที่น้อยกว่าหรือเท่ากับคะแนนรวม ($j \leq n_i$) และ U_{ij} มีค่าเท่ากับ 0 เมื่อข้อสอบข้อที่ j มีลำดับที่มากกว่าคะแนนรวม ($j > n_i$)

จากสูตรคำนวณข้างต้น สามารถเขียนเป็นสูตรคำนวณที่สะดวกได้ดังนี้ (Sato 1980:

19)

$$C_i = \frac{\sum_{j=1}^{n_i} (1 - U_{ij}) n_{.j} - \sum_{j=n_i+1}^J U_{ij} n_{.j}}{\sum_{j=1}^{n_i} n_{.j} + \frac{(n_{.i}) (\sum_{j=1}^J n_{.j})}{J}}$$

โดยที่

C_i คือ ค่าดัชนีชี้เตือนของชาโตของผู้สอบคนที่ i

j คือ ลำดับที่ของข้อสอบที่เรียงจากข้อที่ง่ายที่สุดไปจนถึงข้อที่ยากที่สุด; $j = 1, 2, 3, \dots, J$

U_{ij} คือ คะแนนรายข้อของข้อสอบข้อที่ j ของผู้สอบคนที่ i โดยที่ U_{ij} มีค่าเท่ากับ 1 เมื่อผู้สอบตอบได้ถูกต้อง และ U_{ij} มีค่าเท่ากับ 0 เมื่อผู้สอบตอบผิด

n_i คือ คะแนนรวมของผู้สอบคนที่ i

n_j คือ จำนวนผู้สอบที่ตอบข้อสอบข้อที่ j ได้ถูกต้อง

J คือ จำนวนข้อสอบทั้งหมด

นิสยของดัชนีชี้เตือนของชาโตมีค่าอยู่ระหว่าง 0 ถึง $(\sum_{j=1}^{n_i} n_j - \sum_{j=J-n_i+1}^J n_j) / (\sum_{j=1}^{n_i} n_j - (n_i) (\sum_{j=1}^J n_j / J))$ โดยที่ดัชนีจะมีค่าเข้าใกล้ 0 เมื่อผู้สอบมีแบบแผนการตอบข้อสอบที่มีลักษณะใกล้เคียงกับกัตแมนสมบรูณ์ และดัชนีจะมีค่าเท่ากับ 0 เมื่อแบบแผนการตอบข้อสอบของผู้สอบมีลักษณะเป็นกัตแมนสมบรูณ์ ในทางตรงกันข้ามดัชนีจะมีค่าเข้าใกล้ $(\sum_{j=1}^{n_i} n_j - \sum_{j=J-n_i+1}^J n_j) / (\sum_{j=1}^{n_i} n_j - (n_i) (\sum_{j=1}^J n_j / J))$ เมื่อผู้สอบมีแบบแผนการตอบข้อสอบที่มีลักษณะแตกต่างจากกัตแมนสมบรูณ์ และดัชนีมีค่าเท่ากับ $(\sum_{j=1}^{n_i} n_j - \sum_{j=J-n_i+1}^J n_j) / (\sum_{j=1}^{n_i} n_j - (n_i) (\sum_{j=1}^J n_j / J))$ เมื่อแบบแผนการตอบข้อสอบของผู้สอบมีลักษณะตรงกันข้ามกับกัตแมนสมบรูณ์

8. ดัชนีชี้เตือนที่ปรับแก้แล้ว (The Modified Caution Index: C_i^*)

ในปี ค.ศ. 1981 ฮาร์นิชและลินน์ (Harnisch and Linn 1981) ได้พัฒนาดัชนีขึ้นโดยการปรับแก้สูตรคำนวณของดัชนีชี้เตือนของชาโต เพื่อให้มีค่านิสยอยู่ระหว่าง 0 ถึง 1.0 หลักการปรับแก้ของฮาร์นิชและลินน์อยู่ที่การปรับค่าของเทอม $(n_i) (\sum_{j=1}^J n_j / J)$ ในสูตรคำนวณของชาโตซึ่งก็คือค่าการเตา (A_{ci}) ของ เคน และ เบนแนน ให้มีค่าเท่ากับค่าต่ำสุดของค่าดัชนีความสอดคล้อง ($\min A_i$) คือ $\sum_{j=J-n_i+1}^J n_j$ ดังนั้นดัชนีชี้เตือนที่ปรับแก้แล้ว จึงมีสูตรคำนวณคือ (Harnisch and Linn 1981: 135)

$$C_i^* = \frac{\sum_{j=1}^{n_i} (1 - U_{ij}) n_j - \sum_{j=J-n_i+1}^J U_{ij} n_j}{\sum_{j=1}^{n_i} n_j + \sum_{j=J-n_i+1}^J n_j}$$

โดยที่

C_i^* คือ ค่าดัชนีชี้เตือนที่ปรับแก้แล้วของผู้สอบคนที่ i

j คือ ลำดับที่ของข้อสอบที่เรียงจากข้อที่ง่ายที่สุดไปจนถึงข้อที่ยากที่สุด; $j = 1, 2, 3, \dots, J$

U_{ij} คือ คะแนนรายข้อของข้อสอบข้อที่ j ของผู้สอบคนที่ i โดยที่ U_{ij} มีค่าเท่ากับ 1 เมื่อผู้สอบตอบได้ถูกต้อง และ U_{ij} มีค่าเท่ากับ 0 เมื่อผู้สอบตอบผิด

n_i คือ คะแนนรวมของผู้สอบคนที่ i

n_j คือ จำนวนผู้สอบที่ตอบข้อสอบข้อที่ j ได้ถูกต้อง

J คือ จำนวนข้อสอบทั้งหมด

พิสัยของค่าดัชนีชี้เตือนที่ปรับแก้แล้ว มีค่าอยู่ระหว่าง 0 ถึง 1.0 โดยที่ดัชนีจะมีค่าเข้าใกล้ 0 เมื่อผู้สอบมีแบบแผนการตอบข้อสอบที่มีลักษณะใกล้เคียงกับกัตแมนสมบูรณ์ และดัชนีจะมีค่าเท่ากับ 0 เมื่อแบบแผนการตอบข้อสอบของผู้สอบมีลักษณะเป็นกัตแมนสมบูรณ์ ในทางตรงกันข้ามดัชนีจะมีค่าเข้าใกล้ 1.0 เมื่อผู้สอบมีแบบแผนการตอบข้อสอบที่มีลักษณะแตกต่างจากกัตแมนสมบูรณ์ และดัชนีจะมีค่าเท่ากับ 1.0 เมื่อแบบแผนการตอบข้อสอบของผู้สอบมีลักษณะตรงกันข้ามกับกัตแมนสมบูรณ์ ฮาร์นิชและลินน์กำหนดว่าดัชนีชี้เตือนที่ปรับแก้แล้วที่มีค่าตั้งแต่ .3 ขึ้นไป จะบ่งชี้ว่าแบบแผนการตอบนั้นมีลักษณะที่แตกต่างไปจากกัตแมนสมบูรณ์ ดังนั้นการแปลความหมายของคะแนนรวมของผู้สอบที่มีค่าดัชนีดังกล่าวจะต้องเป็นไปอย่างระมัดระวัง

9. ดัชนียูฟโรม ของ แวน เดอร์ ฟลิเออร์ (The van der Flier Index: U'_1)

ในปี ค.ศ. 1977 แวน เดอร์ ฟลิเออร์ (van der Flier 1977) ได้พัฒนาดัชนีขึ้น โดยการนำคะแนนรายข้อของข้อสอบทุกข้อมาจับคู่กัน แล้วนับจำนวนคู่ของคะแนนรายข้อที่มีลักษณะตามกัตแมน คือ (0, 1) แล้วนำค่าที่ได้มาหารด้วยค่าสูงสุดของดัชนี (U_{max}) ซึ่งค่าสูงสุดนี้ได้มาจากผลคูณระหว่างจำนวนข้อสอบที่ผู้สอบตอบได้ถูกต้องกับจำนวนข้อสอบที่ผู้สอบตอบผิด

แวน เดอร์ ฟลิเออร์ เรียกผลรวมของจำนวนคู่ของคะแนนรายข้อที่มีลักษณะตามกัตแมนว่า ค่ายู (U) และเสนอให้คำนวณค่ายูจากผลรวมของจำนวนเลข 1 ที่อยู่ขวามือของเลข 0 ทุกตัว ดัชนียูฟโรมจึงมีสูตรคำนวณ คือ (van der Flier 1977: 32-33)

$$U'_1 = \frac{U}{U_{max}}$$

โดยที่

U_i คือ ค่าของดัชนียูนิฟิรไมของผู้สอบคนที่ i

U คือ ผลรวมของจำนวนตัวเลข 1 ที่อยู่ด้านขวามือของทุกๆตัวเลข 0 ดังนั้น

$$U = \sum_{j=1}^{K-1} \sum_{k=j+1}^K (1-U_k) U_j$$

U_{max} คือ ค่าสูงสุดของ U มีค่าเท่ากับผลคูณของจำนวนข้อสอบที่ผู้สอบตอบได้ถูกต้องกับจำนวนข้อสอบที่ผู้สอบตอบผิด

U_k คือ คะแนนรายชื่อของข้อสอบข้อที่ k โดย U_k มีค่าเท่ากับ 1 เมื่อผู้สอบตอบได้ถูกต้อง และ U_k มีค่าเท่ากับ 0 เมื่อผู้สอบตอบผิด

U_j คือ คะแนนรายชื่อของข้อสอบข้อที่ j โดย U_j มีค่าเท่ากับ 1 เมื่อผู้สอบตอบได้ถูกต้อง และ U_j มีค่าเท่ากับ 0 เมื่อผู้สอบตอบผิด

นิสัยของค่าดัชนียูนิฟิรไม มีค่าอยู่ระหว่าง 0 ถึง 1.0 โดยที่ดัชนีจะมีค่าเข้าใกล้ 0 เมื่อผู้สอบมีแบบแผนการตอบข้อสอบที่มีลักษณะใกล้เคียงกับกัตแมนสมบรูณ์ และดัชนีจะมีค่าเท่ากับ 0 เมื่อแบบแผนการตอบข้อสอบของผู้สอบมีลักษณะเป็นกัตแมนสมบรูณ์ ในทางตรงกันข้ามดัชนีจะมีค่าเข้าใกล้ 1.0 เมื่อผู้สอบมีแบบแผนการตอบข้อสอบที่มีลักษณะแตกต่างจากกัตแมนสมบรูณ์ และดัชนีจะมีค่าเท่ากับ 1.0 เมื่อแบบแผนการตอบข้อสอบของผู้สอบมีลักษณะตรงกันข้ามกับกัตแมนสมบรูณ์

10. ดัชนีการคล้อยตามกลุ่ม (Norm Conformity Index: NCI_i)

ในปี ค.ศ. 1980 ทาทซูกะและทาทซูกะ (Tatsuoka and Tatsuoka 1980) ได้พัฒนาดัชนีขึ้น โดยการจับคู่คะแนนรายชื่อของข้อสอบทุกๆข้อแล้วนับจำนวนคู่ของคะแนนรายชื่อที่มีลักษณะตามกัตแมน (0,1) ซึ่งกำหนดเป็นสัญลักษณ์คือ U_{01} กับจำนวนคู่ที่มีลักษณะตรงกันข้ามกับกัตแมน (1,0) ซึ่งกำหนดเป็นสัญลักษณ์คือ U_{10} จากนั้นนำค่าดังกล่าวมาคำนวณค่าดัชนีการคล้อยตามกลุ่ม ซึ่งมีสูตรคำนวณคือ (Tatsuoka and Tatsuoka 1983: 222-223)

$$NCI_i = \frac{2U_{01}}{U} - 1$$

โดยที่

NCI_i คือ ค่าดัชนีการคล้อยตามกลุ่มของผู้สอบคนที่ i

U_{01} คือ ผลรวมของจำนวนคู่ของคะแนนรายชื่อที่มีลักษณะตามกัตแมน

U_{10} คือ ผลรวมของจำนวนคู่ของคะแนนรายชื่อที่มีลักษณะตรงกันข้ามกัตแมน

U คือ ผลคูณระหว่างจำนวนข้อสอบที่ผู้สอบตอบได้ถูกต้องกับจำนวนข้อสอบที่ผู้สอบตอบผิด (U มีค่าเท่ากับ $U_u + U_d$ เสมอ)

นอกจากนี้ค่าของ U_u และ U_d ที่ได้กล่าวมาแล้ว ยังสามารถคำนวณได้จากวิธีการต่อไปนี้ คือ (Tatsuoka and Tatsuoka 1982: 219-220)

กำหนดให้

S คือ เวกเตอร์ของแบบแผนการตอบข้อสอบของผู้สอบ ตัวอย่างเช่น

[1 1 0 1 0]

\tilde{S} คือ ทรานโพสิทของคอมพลิเมนต์ของ S ตามตัวอย่าง \tilde{S} คือ

$$\tilde{S} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

$$N = \tilde{S} S \quad \text{ตามตัวอย่างคือ} \quad \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} [1 \ 1 \ 0 \ 1 \ 0] = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \end{bmatrix}$$

ตามแนวคิดนี้ U_u มีค่าเท่ากับผลรวมของจำนวนที่อยู่ใต้เส้นทแยงมุมหลัก ในที่นี้ U_u เท่ากับ 5

U_d มีค่าเท่ากับผลรวมของจำนวนที่อยู่เหนือเส้นทแยงมุมหลัก ในที่นี้ U_d เท่ากับ 1

และ U มีค่าเท่ากับ $U_u + U_d$ ในที่นี้ U เท่ากับ 5 + 1 เท่ากับ 6

เพราะฉะนั้น ผู้สอบคนนี้มีค่าดัชนีการคล้อยตามกลุ่มเท่ากับ $\{(2)(5)/6\} - 1 = .67$

ค่าของดัชนีการคล้อยตามกลุ่มนี้ มีความสัมพันธ์เชิงลบอย่างสมบูรณ์ (perfectly negative correlation) กับ ดัชนีสุ่มไร้มองวน เฮอร์ ไฟล์เออร์ คือ $NCI = 1 - 2U$ (Harnisch and Linn 1981: 139)

พิสัยของค่าของดัชนีการคล้อยตามกลุ่ม มีค่าอยู่ระหว่าง -1.0 ถึง 1.0 โดยที่ดัชนีจะมีค่าเข้าใกล้ 1.0 เมื่อแบบแผนการตอบข้อสอบของผู้สอบมีลักษณะใกล้เคียงกับกัตแมนสมบูรณ์ และดัชนีจะมีค่าเท่ากับ 1.0 เมื่อแบบแผนการตอบข้อสอบของผู้สอบมีลักษณะเป็นกัตแมนสมบูรณ์ ใน

ทางตรงกันข้ามดัชนีการคล้ายตามกลุ่มจะมีค่าเข้าใกล้ -1.0 เมื่อแบบแผนการตอบข้อสอบของผู้สอบมีลักษณะแตกต่างจากกัตแมนสมบูรณ์ และดัชนีจะมีค่าเท่ากับ -1.0 เมื่อผู้สอบมีแบบแผนการตอบข้อสอบที่มีลักษณะตรงกันข้ามกับกัตแมนสมบูรณ์

11. ดัชนีดีเอฟซี (DFC Index: DFC)

ในปี พ.ศ. 2534 ส้าราญ มีแจ้งได้พัฒนาดัชนีดีเอฟซีขึ้น เพื่อแก้ไขจุดอ่อนในประเด็นที่ค่าของดัชนีมีค่าเปลี่ยนแปลงไปตามแบบแผนการตอบข้อสอบของผู้สอบภายในกลุ่ม โดยที่ดัชนีดีเอฟซีที่เสนอนั้นมี 2 สูตร คือ ดัชนีดีเอฟซี 1 และดัชนีดีเอฟซี 2 ซึ่งแนวทางในการแก้จุดอ่อนดังกล่าวของ ดัชนีดีเอฟซี 1 คือการใช้ค่าโอกาสการตอบข้อสอบข้อที่ j ได้ถูกต้องของผู้สอบคนที่ i มาแทนค่าความยากของข้อสอบข้อที่ j และดัชนีดีเอฟซี 2 ใช้ค่าเฉลี่ยของค่าโอกาสการตอบข้อสอบข้อที่ j ได้ถูกต้องของผู้สอบทั้งกลุ่ม มาแทนค่าความยากของข้อ j ในสูตรคำนวณของดัชนีดีเอฟซีของซาโต สำหรับ ดัชนีดีเอฟซี 1 และ ดัชนีดีเอฟซี 2 ดังสูตรต่อไปนี้ (ส้าราญ มีแจ้ง 2534: 75-80)

$$DFC 1_i = \frac{\sum_{j=1}^{n_i} (1 - U_{ij}) P_{ij} - \sum_{j=n_i+1}^J U_{ij} P_{ij}}{\sum_{j=1}^{n_i} P_{ij} + \frac{(n_i \cdot) (\sum_{j=1}^J P_{ij})}{J}}$$

$$DFC 2_i = \frac{\sum_{j=1}^{n_i} (1 - U_{ij}) G_j - \sum_{j=n_i+1}^J U_{ij} G_j}{\sum_{j=1}^{n_i} G_j + \frac{(n_i \cdot) (\sum_{j=1}^J G_j)}{J}}$$

โดยที่

DFC 1_i คือ ค่าของดัชนีดีเอฟซี 1 ของผู้สอบคนที่ i

DFC 2_i คือ ค่าของดัชนีดีเอฟซี 2 ของผู้สอบคนที่ i

j คือ ลำดับที่ของข้อสอบที่เรียงจากข้อที่ง่ายที่สุดไปยังข้อที่ยากที่สุด;

$j = 1, 2, 3, \dots, J$

U_{ij} คือ คะแนนรายชื่อของข้อสอบข้อที่ j ของผู้สอบคนที่ i โดยที่ U_{ij} มีค่าเท่ากับ 1 เมื่อผู้สอบตอบได้ถูกต้อง และ U_{ij} มีค่าเท่ากับ 0 เมื่อผู้สอบตอบผิด

P_{ij} คือ ค่าโอกาสการตอบข้อสอบข้อที่ j ได้ถูกต้องของผู้สอบคนที่ i

G_j คือ ค่าเฉลี่ยของค่าโอกาสการตอบข้อสอบข้อที่ j ได้ถูกต้องของผู้

สอบทั้งกลุ่ม

นิสัยของดัชนีดีเอฟซี มีค่าอยู่ระหว่าง - ถึง โดยที่ดัชนีดีเอฟซีจะมีค่า
 เข้าสู่ 0 เมื่อผู้สอบมีแบบแผนการตอบข้อสอบที่มีลักษณะใกล้เคียงกับกัตแมนสมบูรณ์ และดัชนีจะมีค่า
 เท่ากับ 0 เมื่อผู้สอบมีแบบแผนการตอบข้อสอบที่มีลักษณะเป็นกัตแมนสมบูรณ์ (สำราญ มีแจ้ง 2534:
 80)

ตารางที่ 5 ที่จะนำเสนอต่อไปนี้เป็น การเปรียบเทียบค่าของดัชนีทั้ง 12 ดัชนี
 ที่กล่าวมาแล้ว จากข้อมูลเชิงทฤษฎีที่แสดงแบบแผนการตอบข้อสอบที่เป็นไปได้ทั้งหมดของการตอบข้อ
 สอบ 5 ข้อ ซึ่งข้อสอบทั้ง 5 ข้อนี้ มีค่าความยากตามแนวทฤษฎีดั้งเดิม (Classical Test
 Theory) เท่ากับ 0.67 0.56 0.39 0.33 และ 0.17 ตามลำดับ และมีค่าความยาก
 ตามแนวทฤษฎีการตอบสนองรายข้อ (Item Response Theory) เท่ากับ -2 -1 0 1
 และ 2 ตามลำดับ สำหรับข้อสอบทั้ง 5 ข้อนี้ได้จัดเรียงลำดับจากข้อที่ง่ายที่สุดไปจนถึงข้อที่ยากที่สุด

ดัชนีทั้ง 12 ดัชนี ที่นำมาเปรียบเทียบกัน คือ

- c คือ ดัชนีชี้เตือนของซาโต
- c* คือ ดัชนีชี้เตือนที่ปรับแก้แล้ว
- r* คือ ดัชนีของไบบีเรียลของแต่ละบุคคล
- r คือ ดัชนีไบบีเรียลของแต่ละบุคคล
- A คือ ดัชนีความสอดคล้อง
- D คือ ดัชนีความไม่สอดคล้อง
- O คือ ดัชนีความเชื่อถือ
- Oc คือ ดัชนีความเชื่อถือที่ปรับแก้การเตา
- U คือ ดัชนียูโรม
- NCI คือ ดัชนีการคล้อยตามกลุ่ม
- DFC1 คือ ดัชนีดีเอฟซี1
- DFC2 คือ ดัชนีดีเอฟซี2

ตารางที่ 5 ค่าของดัชนีความสอดคล้องทั้ง 12 ดัชนี ที่คำนวณจากข้อมูลเชิงทฤษฎีที่

แสดงแบบแผนการตอบข้อสอบที่เป็นไปได้ทั้งหมดจากการตอบข้อสอบ 5 ข้อ

แบบแผน การตอบ	ค่าความ สามารถ	ข้อสอบ					คะแนน รวม	ดัชนี											
		1	2	3	4	5		c	c'	r'	r	A	D	O	Oc	U	NCI	DFC1	DFC2
1	1.8	1	1	1	1	1	5	0	0	0	0	2.10	0	1	1	0	1	1	1
2	1.8	1	1	1	1	0	4	0	0	.73	1.05	1.94	0	1	1	0	1	0	0
3	1.6	1	1	1	0	1	4	.65	.33	.26	.36	1.78	.17	.91	.35	.25	.50	.75	.43
4	1.6	1	1	0	1	1	4	.89	.44	.12	.15	1.70	.24	.88	.85	.50	0	1.15	1
5	1.6	1	0	1	1	1	4	1.52	.78	-.34	-.43	1.56	.38	.80	.76	.75	-.50	1.65	1.54
6	1.6	0	1	1	1	1	4	1.96	1	-.65	-.81	1.44	.50	.74	.69	1	-1	1.85	1.94
7	1.4	1	1	1	0	0	3	0	0	.81	1.02	1.61	0	1	1	0	1	0	0
8	1.2	1	1	0	1	0	3	.16	.08	.68	.86	1.56	.06	.97	.84	.17	.67	.38	.34
9	1.0	1	1	0	0	1	3	.65	.31	.29	.36	1.39	.22	.86	.35	.33	.33	.79	.63
10	0.8	1	0	1	1	0	3	1.02	.49	-.02	-.03	1.19	.36	.72	-.08	.33	.33	.94	.86
11	0.8	1	0	1	0	1	3	1.13	.54	-.10	-.13	1.22	.39	.76	-.13	.50	0	1.05	.98
12	0.8	1	0	0	1	1	3	1.29	.62	-.21	-.26	1.17	.44	.73	.68	.67	-.33	3.00	2.58
13	0.8	0	1	0	1	1	3	1.47	.77	-.46	-.58	1.06	.55	.66	.60	.83	-.67	3.14	2.85
14	0.8	0	1	1	0	1	3	1.45	.67	-.34	-.43	1.11	.50	.69	.63	.67	-.33	1.20	1.28
15	0.8	0	1	1	1	0	3	.97	.46	.04	.05	1.28	.33	.80	.76	.50	0	.83	.98
16	0.8	0	0	1	1	1	3	2.10	1	-.83	-1.04	.89	.72	.55	.47	1	-1	3.42	3.21
17	0.6	1	1	0	0	0	2	0	0	.88	1.12	1.22	0	1	1	0	1	0	0
18	0.2	1	0	1	0	0	2	.44	.23	.49	.63	1.06	.17	.86	.56	.17	.67	.35	.35
19	0	1	0	0	1	0	2	.59	.31	.36	.46	1	.22	.82	.41	.33	.33	.75	.70
20	-0.2	0	1	1	0	0	2	.74	.38	.23	.30	.94	.28	.77	.26	.33	.33	.64	.61
21	-0.4	1	0	0	0	1	2	1.32	.69	-.02	-.03	.83	.39	.68	.63	.50	0	.91	1.02
22	-0.4	0	1	0	1	0	2	.88	.46	.10	.10	.89	.33	.73	.12	.50	0	1.01	.96
23	-0.4	0	1	0	0	1	2	1.32	.69	-.27	-.34	.72	.50	.59	.52	.50	0	1.21	1.28
24	-0.4	0	0	1	1	0	2	1.32	.69	-.27	-.34	.72	.50	.59	.52	.67	-.33	2.77	3.26
25	-0.4	0	0	1	0	1	2	1.76	.92	-.64	-.80	.56	.66	.46	.37	.83	-.67	2.90	3.57
26	-0.4	0	0	0	1	1	2	1.91	1	-.76	-.95	.50	.72	.41	.31	1	-1	2.58	3.21
27	-0.6	1	0	0	0	0	1	0	0	.70	1	.67	0	1	1	0	1	0	0
28	-1.0	0	1	0	0	0	1	.45	.22	.38	.55	.56	.11	.83	.55	.25	.50	.57	.41
29	-1.2	0	0	1	0	0	1	1.14	.56	-.10	-.14	.39	.28	.58	-.14	.50	0	1.15	.97
30	-1.4	0	0	0	1	0	1	1.36	.67	-.26	-.36	.33	.33	.50	-.36	.75	-.50	1.47	1.52
31	-1.6	0	0	0	0	1	1	2.04	1	-.70	-.88	.17	.50	.25	.15	1	-1	1.58	2.06
32	-1.8	0	0	0	0	0	1	0	0	0	0	0	0	1	1	0	1	0	0

จากตัวอย่างข้อมูลที่ได้อธิบายในตารางที่ 5 เป็นรายละเอียดเกี่ยวกับดัชนีความสอดคล้องของแบบแผนการตอบข้อสอบรายบุคคลที่ได้กล่าวมาแล้วทั้งหมด ซึ่งสามารถให้ข้อสรุปเกี่ยวกับค่าของดัชนี คุณลักษณะที่ร่วมกัน และจุดอ่อนของดัชนีดังกล่าวได้ดังนี้

ข้อสรุปเกี่ยวกับดัชนีความสอดคล้องของแบบแผนการตอบข้อสอบรายบุคคล

1. ข้อตกลงเบื้องต้น เมื่อแบบสอบฉบับหนึ่งๆวัดคุณลักษณะเดียวกัน และประกอบด้วยข้อสอบที่สามารถนำมาจัดเรียงลำดับตามจำนวนผู้สอบภายในกลุ่มที่ตอบข้อดังกล่าวได้ถูกต้องแล้ว ผู้สอบคนใดก็ตามที่ตอบข้อสอบที่มีความยากระดับหนึ่งได้ถูกต้องแล้ว ย่อมจะตอบข้อสอบทุกข้อที่ง่ายกว่าข้อดังกล่าวได้ถูกต้องด้วย ในทางตรงกันข้ามเมื่อผู้สอบตอบข้อสอบที่มีความยากระดับหนึ่งผิด ย่อมจะตอบข้อสอบทุกข้อที่ยากกว่าข้อดังกล่าวผิดด้วย

2. แบบแผนการตอบข้อสอบของผู้สอบ หมายถึง ชุดของคะแนนรายชื่อของผู้สอบที่ถูกนำมาจัดเรียงลำดับจากข้อสอบข้อที่ง่ายที่สุด (ข้อที่มีจำนวนผู้สอบภายในกลุ่มตอบได้ถูกต้องมากที่สุด) ไปจนถึงข้อสอบข้อที่ยากที่สุด (ข้อที่มีจำนวนผู้สอบภายในกลุ่มตอบได้ถูกต้องน้อยที่สุด)

3. เส้นแบ่งเขตคะแนนรายชื่อตามลักษณะกัณฑ์สมมาตร ดัชนีดังกล่าวกำหนดให้ข้อสอบที่มีลำดับที่เท่ากับคะแนนรวม (n_x) ที่ผู้สอบทำได้ เป็นเส้นแบ่งเขตระหว่างข้อสอบที่จัดว่าง่ายกว่าหรือข้อที่มีความยากเท่ากับความสามารถของผู้สอบ กับข้อสอบที่จัดว่ายากกว่าความสามารถของผู้สอบ

4. ความสอดคล้อง (agreement or consistency) ของแบบแผนการตอบข้อสอบ หมายถึง ความสอดคล้องของแบบแผนการตอบข้อสอบที่ผู้สอบทำได้ กับ แบบแผนการตอบข้อสอบที่มีลักษณะเป็นกัณฑ์สมมาตร โดยที่

4.1 ค่าความสอดคล้องจะมีค่าสูงสุด เมื่อผู้สอบตอบข้อสอบทุกข้อที่จัดว่าเป็นข้อที่ง่ายกว่าหรือข้อที่มีความยากเท่ากับความสามารถของเขาได้ถูกต้องทุกข้อ และตอบข้อสอบทุกข้อที่ยากกว่าความสามารถของเขาผิด

สรุปได้ว่าค่าความสอดคล้องจะมีค่าสูงสุด เมื่อผู้สอบมีแบบแผนการตอบข้อสอบที่มีลักษณะเป็นกัณฑ์สมมาตร

4.2 ค่าความสอดคล้องจะมีค่าต่ำสุด เมื่อผู้สอบตอบข้อสอบทุกข้อที่จัดว่าเป็นข้อที่ง่ายกว่าหรือข้อที่มีความยากเท่ากับความสามารถของเขาผิด และตอบข้อสอบทุกข้อที่จัดว่าเป็นข้อที่ยากกว่าความสามารถของเขาได้ถูกต้อง

สรุปได้ว่าค่าความสอดคล้องจะมีค่าต่ำสุด เมื่อผู้สอบมีแบบแผนการตอบข้อสอบที่มีลักษณะตรงกันข้ามกับกัณฑ์สมมาตร

5. ผู้สอบทุกคนมีโอกาสที่จะมีแบบแผนการตอบข้อสอบที่มีลักษณะเป็นกัตแมนสมบรูณ์ หรือมีแบบแผนการตอบข้อสอบที่มีลักษณะที่ตรงกันข้ามกัตแมนสมบรูณ์

6. ความหมายของค่าความสอดคล้องของดัชนีทั้ง 12 ดัชนี แบ่งได้เป็น 2 ลักษณะ คือ

6.1 ดัชนีที่มีค่าสูงสุด ซึ่งหมายความว่าแบบแผนการตอบข้อสอบของผู้สอบมีลักษณะเป็นกัตแมนสมบรูณ์ ดัชนีในกลุ่มนี้ ได้แก่ ดัชนีพอยท์ไบซีเรียลของแต่ละบุคคล ดัชนีไบซีเรียลของแต่ละบุคคล ดัชนีความสอดคล้อง ดัชนีความเชื่อถือได้ ดัชนีความเชื่อถือได้ที่ปรับแก้การเดา และดัชนีการคล้อยตามกลุ่ม โดยที่มีความหมายของค่าของดัชนีว่ายิ่งดัชนีมีค่าสูง ยิ่งแสดงว่าแบบแผนการตอบข้อสอบของผู้สอบมีลักษณะใกล้เคียงกับกัตแมนสมบรูณ์

6.2 ดัชนีที่มีค่าต่ำสุด ซึ่งหมายความว่าแบบแผนการตอบข้อสอบของผู้สอบมีลักษณะเป็นกัตแมนสมบรูณ์ ดัชนีในกลุ่มนี้ ได้แก่ ดัชนีชี้เตือนของซาโต ดัชนีชี้เตือนที่ปรับแก้แล้ว ดัชนีความไม่สอดคล้อง ดัชนีฮูไรม และดัชนีดีเอฟซี โดยที่มีความหมายของค่าของดัชนีว่ายิ่งดัชนีมีค่าต่ำ ยิ่งแสดงว่าแบบแผนการตอบข้อสอบของผู้สอบมีลักษณะใกล้เคียงกัตแมนสมบรูณ์

7. จุดอ่อนของดัชนีที่มีอยู่ ได้แก่

7.1 ในกรณีที่ผู้สอบมีแบบแผนการตอบข้อสอบเป็นกัตแมนสมบรูณ์เหมือนกัน ดัชนีกลับมีค่าแตกต่างกัน ถ้าผู้สอบมีคะแนนรวมต่างกัน

ยกตัวอย่างเช่นแบบแผนการตอบที่ 2 ที่ 7 ที่ 17 และที่ 27 ที่อยู่ในตารางที่ 5 ซึ่งมีคะแนนรวมต่างกันเป็น 4 คะแนน 3 คะแนน 2 คะแนน และ 1 คะแนน ตามลำดับ และผู้สอบทั้ง 4 คนนี้ มีแบบแผนการตอบข้อสอบที่มีลักษณะเป็นกัตแมนสมบรูณ์เหมือนกัน แต่มีค่าของดัชนีพอยท์ไบซีเรียลของแต่ละบุคคลเท่ากับ 0.73 0.81 0.88 และ 0.70 ตามลำดับ

นอกจากดัชนีพอยท์ไบซีเรียลของแต่ละบุคคลแล้ว ยังมีดัชนีไบซีเรียลของแต่ละบุคคล และดัชนีความสอดคล้อง ที่มีจุดอ่อนในข้อนี้

จุดอ่อนของความไม่คงที่ของดัชนี เมื่อผู้สอบมีแบบแผนการตอบข้อสอบที่มีลักษณะเป็นกัตแมนสมบรูณ์นี้ ทำให้ไม่สามารถกำหนดค่าสูงสุดที่บ่งชี้ว่าแบบแผนการตอบข้อสอบมีลักษณะเป็นกัตแมนสมบรูณ์ได้ เพราะฉะนั้นค่าของดัชนีจึงไม่มีความหมายแต่อย่างใด

7.2 ค่าของดัชนีได้รับผลกระทบ จากการเปลี่ยนแปลงของแบบแผนการตอบข้อสอบของผู้สอบคนอื่นๆภายในกลุ่ม ที่เป็นเช่นนี้เพราะเมื่อแบบแผนการตอบข้อสอบของผู้สอบคนอื่นเปลี่ยนแปลงจะมีผลทำให้จำนวนผู้สอบภายในกลุ่มที่ตอบข้อดังกล่าวได้ถูกต้อง เปลี่ยนไป ดังนั้นค่าของดัชนีทุกตัว ที่นำจำนวนผู้สอบภายในกลุ่มที่ตอบข้อดังกล่าวได้ถูกต้อง หรือนำค่าความยากของข้อสอบมา

เป็นตัวประกอบในการกำหนดความสอดคล้องดังกล่าวจะมีค่าเปลี่ยนไปทั้งๆที่แบบแผนการตอบข้อสอบของผู้สอบยังคงเดิม ดัชนีที่มีจุดอ่อนในข้อนี้ ได้แก่ ดัชนีใบชี้เรีลของแต่ละบุคคล ดัชนีรอยที่ใบชี้เรีลแต่ละของบุคคล ดัชนีความสอดคล้อง ดัชนีความไม่สอดคล้อง ดัชนีความเชื่อถือได้ ดัชนีความเชื่อถือได้ที่ปรับแก้การเดา ดัชนีชี้เตือนของซาโต และดัชนีชี้เตือนที่ปรับแก้แล้ว

7.3 ค่าของดัชนีแสดงความคลาดเคลื่อนสูงสุด ทั้งๆที่แบบแผนการตอบข้อสอบของผู้สอบไม่ได้มีความหมายเช่นนั้น จุดอ่อนข้อนี้จะเกิดขึ้นในกรณีที่ผู้สอบคะแนนรวมมากกว่า $1/2$ ของจำนวนข้อสอบทั้งหมด และได้คะแนนรวม \times คะแนน จากการตอบข้อสอบจำนวน \times ข้อสุดท้ายได้ถูกต้องทุกข้อ ยกตัวอย่างเช่นแบบแผนการตอบที่ 6 และที่ 16 ในตารางที่ 5 ซึ่งมีคะแนนรวม 4 คะแนน จากการตอบข้อสอบจำนวน 4 ข้อสุดท้ายได้ถูกต้อง และที่มีคะแนนรวม 3 คะแนน จากการตอบข้อสอบจำนวน 3 ข้อสุดท้ายได้ถูกต้อง จะมีค่าดัชนีการคล้อยตามกลุ่มเท่ากับ -1.0 (ซึ่งค่าดังกล่าวแสดงว่าแบบแผนการตอบข้อสอบของผู้สอบมีความคลาดเคลื่อนสูงสุด) ทั้งๆที่การตอบข้อสอบลำดับที่ 2 ถึงลำดับที่ 4 ได้ถูกต้องของแบบแผนการตอบที่ 6 และการตอบข้อสอบลำดับที่ 3 ได้ถูกต้องของแบบแผนการตอบที่ 16 เป็นสิ่งที่มีความสมเหตุสมผลตามความหมายของกัตแมน เพราะข้อดังกล่าวเป็นข้อที่จัดว่าเป็นข้อที่ง่ายกว่า หรือเป็นข้อที่มีความยากเท่ากับความสามารถของเขา เพราะฉะนั้นจึงไม่ควรจะมีค่าของดัชนีที่แสดงความคลาดเคลื่อนสูงสุด ซึ่งเมื่อนำมาเปรียบเทียบกับแบบแผนการตอบที่ 26 ซึ่งมีคะแนนรวม 2 คะแนน จากการตอบข้อสอบลำดับที่ 4 และลำดับที่ 5 ได้ถูกต้อง หรือแบบแผนการตอบที่ 32 ซึ่งมีคะแนนรวม 1 คะแนน จากการตอบข้อสอบลำดับที่ 5 ได้ถูกต้องนั้นการตอบข้อดังกล่าวได้ถูกต้องนับว่าเป็นคะแนนรายข้อที่มีความคลาดเคลื่อน เพราะข้อดังกล่าวจัดว่าเป็นข้อที่ยากกว่าความสามารถของเขา (แบบแผนการตอบที่ 26 และที่ 32 เป็นแบบแผนการตอบข้อสอบที่มีความคลาดเคลื่อนสูงสุด) ด้วยเหตุนี้แบบแผนการตอบที่ 6 และที่ 16 จึงไม่น่าจะมีค่าของดัชนีการคล้อยตามกลุ่มเท่ากับของผู้สอบคนที่ 26 และที่ 32

สำหรับดัชนีที่มีจุดอ่อนข้อนี้ ได้แก่ ดัชนีการคล้อยตามกลุ่ม และดัชนียูนิโรม

7.4 นิสัยของดัชนีมีค่าไม่แน่นอน เมื่อพิจารณาจากนิสัยของดัชนีต่างๆที่ได้กล่าวมาแล้ว พบว่า

7.4.1 ดัชนีที่มีค่าสูงสุด ที่แสดงว่าแบบแผนการตอบข้อสอบของผู้สอบมีลักษณะเป็นกัตแมนสมบูรณ์นั้น พบว่าค่าต่ำสุดของดัชนีขึ้นอยู่กับค่าความยากของข้อสอบ ดัชนีที่มีจุดอ่อนในข้อนี้ ได้แก่ ดัชนีความเชื่อถือได้ และดัชนีความเชื่อถือได้ที่ปรับแก้การเดา

7.4.2 ดัชนีที่มีค่าต่ำสุด ที่แสดงว่าแบบแผนการตอบข้อสอบของผู้สอบมีลักษณะ

เป็นกัตแมนสมบูรณ์นั้น พบว่าค่าสูงสุดของดัชนีบางดัชนีขึ้นอยู่กับค่าความยากของข้อสอบ ดัชนีที่มีจุดอ่อนในข้อนี้ ได้แก่ ดัชนีความไม่สอดคล้อง ดัชนีชี้เตือนของซาโต และดัชนีดีเฟนซี

7.4.3 ดัชนีที่มีค่าสูงสุดและค่าต่ำสุดไม่แน่นอน ขึ้นกับจำนวนข้อและค่าความยากของข้อสอบ ดัชนีที่มีจุดอ่อนในข้อนี้ ได้แก่ ดัชนีความสอดคล้อง ดัชนีพอที่ไบซีเรียลของแต่ละบุคคล และดัชนีไบซีเรียลของแต่ละบุคคล

ตอนที่ 3 งานวิจัยที่เกี่ยวกับดัชนีความสอดคล้องของแบบแผนการตอบข้อสอบรายบุคคล

งานวิจัยที่จะกล่าวถึงต่อไปนี้จะเกี่ยวข้องกับดัชนีความสอดคล้องของแบบแผนการตอบข้อสอบรายบุคคล ที่ยึดตามแนวคิดของกัตแมน

ทาซุโอกะและเบอเรนบวม (Tatsuoka and Birenbaum 1979 อ้างถึงใน Tatsuoka and Tatsuoka 1982: 221) นำดัชนีการคล้อยตามกลุ่มมาเป็นดัชนีบ่งชี้ความแตกต่างของแบบแผนการตอบข้อสอบของกลุ่มผู้สอบที่ถูกลบยั้งการเรียนรู้ โดยการจัดโครงสร้างของความคิดรวบยอดให้แตกต่างไปจากการเรียนครั้งก่อน ซึ่งจะมีผลให้การเรียนรู้ครั้งหลังของผู้เรียนช้าลง แบบแผนการตอบของกลุ่มตัวอย่างทั้งหมดจะถูกนำมาวิเคราะห์กลุ่ม (cluster analysis) ซึ่งผลการวิเคราะห์สามารถจัดแบ่งกลุ่มตัวอย่างได้เป็น 4 กลุ่ม ที่มีแบบแผนการตอบที่แตกต่างกันอย่างชัดเจน ค่าเฉลี่ยและค่าส่วนเบี่ยงเบนมาตรฐานของดัชนีการคล้อยตามกลุ่มของผู้สอบทั้ง 4 กลุ่ม คือ กลุ่มที่ 1 (จำนวน 34 คน) มีค่าเฉลี่ยเท่ากับ .18 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ .26 กลุ่มที่ 2 (จำนวน 27 คน) มีค่าเฉลี่ยเท่ากับ .41 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ .28 กลุ่มที่ 3 (จำนวน 20 คน) มีค่าเฉลี่ยเท่ากับ .35 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ .31 และกลุ่มสุดท้าย (จำนวน 10 คน) มีค่าเฉลี่ยเท่ากับ .18 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ .43

ฮาร์นิชและลินน์ (Harnisch and Linn 1981: บทคัดย่อ) เสนอดัชนีที่ใช้วัดความสอดคล้องของแบบแผนการตอบข้อสอบขึ้นเรียกว่า ดัชนีชี้เตือนที่ปรับแก้แล้ว และได้นำดัชนีดังกล่าวมาหาค่าความสัมพันธ์กับดัชนีชี้เตือนของซาโตและผลการเรียนของผู้เรียน สำหรับข้อมูลที่นำมาวิเคราะห์นี้ได้จากการตรวจสอบความสามารถทางคณิตศาสตร์ของนักเรียนอายุ 13 ปี ที่จัดทำโดย The National Assessment of Educational Progress (NAEP) ที่ตรวจสอบเมื่อปี ค.ศ. 1978 ผลการศึกษานพบว่า ดัชนีที่พัฒนาขึ้น และดัชนีชี้เตือนของซาโต มีค่าสัมประสิทธิ์สหสัมพันธ์เข้าใกล้ 1.0 ค่าของดัชนีชี้เตือนที่ปรับแก้แล้วได้รับการอภิปรายผลในด้านความเป็นประโยชน์ของ

การแปลผลความสามารถจากคะแนนรวมที่ทำได้ หรือในการค้นหาจุดแข็ง หรือความแตกต่างด้าน
ความรู้ของผู้สอบ

ฮาร์นิชและลินน์ (Harnisch and Linn 1981: 133-146) ศึกษาความสัมพันธ์ระหว่างคะแนนรวมกับดัชนีที่วัดความสอดคล้องของแบบแผนการตอบข้อสอบ 10 ดัชนี กลุ่มตัวอย่าง
เป็นนักเรียนเกรด 4 เกรด 8 และเกรด 11 จำนวน 6,300 คน จากโรงเรียนทั้งหมด 110
โรงเรียน แต่ละระดับของแต่ละโรงเรียนจะถูกสุ่มมา จำนวน 18-22 คน เพื่อทดสอบด้วยแบบสอบ IIEP
(Illinois Inventory of Educational Progress) ซึ่งประกอบด้วยส่วนที่เป็นทักษะทาง
ด้านคณิตศาสตร์และการอ่าน ผลการศึกษาค่าความสัมพันธ์ดังกล่าวของนักเรียนเกรด 4 ที่สอบ
ทักษะคณิตศาสตร์จำนวน 40 ข้อ และด้านการอ่าน จำนวน 28 ข้อ ปรากฏตามตารางที่ 6



ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 6 ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างค่าของดัชนีความสอดคล้องของแบบแผนการตอบข้อสอบกับคะแนนรวม โดยค่าสัมประสิทธิ์สหสัมพันธ์ที่อยู่เหนือเส้นทแยงมุมได้จากการแบบสอบทักษะทางคณิตศาสตร์ ส่วนทักษะด้านการอ่านจะอยู่ที่เส้นทแยงมุม

ดัชนี	C_t	C_t^*	r_t^*	r_t	A_t	D_t	θ_t	U_t'	คะแนนรวม
$C_t (-\theta_{C_t})$	--	.99	-.97	-.99	-.28	.84	-.76	.96	-.17
C_t^*	.96	--	-.95	-.96	-.13	.82	-.65	.93	-.02
r_t^*	-.96	-.90	--	.97	.30	-.77	.74	-.93	.18
r_t	-.95	-.89	.96	--	.39	-.84	.80	-.97	.28
A_t	-.54	-.34	.63	.73	--	-.33	.77	-.40	.99
D_t	.92	.94	-.88	-.91	-.47	--	-.72	.85	-.25
θ_t	-.93	-.79	.90	.93	.74	.80	--	-.80	.70
$U_t' (NCI_t)$.94	.89	-.93	-.95	-.64	.89	-.90	--	-.30
คะแนนรวม	-.42	-.21	.50	.63	.99	-.36	.65	-.54	--

หมายเหตุ ดัชนีต่อไปนี้ คือ

C_t คือ ดัชนีชี้เดือนของซาโต

D_t คือ ดัชนีความไม่สอดคล้อง

C_t^* คือ ดัชนีชี้เดือนที่ปรับแก้แล้ว

θ_t คือ ดัชนีความเชื่อถือได้

r_t^* คือ ดัชนีพอยท์ไบซีเรียลของแต่ละบุคคล

θ_{C_t} คือ ดัชนีความเชื่อถือได้ที่ปรับแก้การเตา

r_t คือ ดัชนีไบซีเรียลของแต่ละบุคคล

U_t คือ ดัชนียูไนต์

A_t คือ ดัชนีความสอดคล้อง

NCI_t คือ ดัชนีการคล้อยตามกลุ่ม

จากตารางที่ 6 จะเห็นได้ว่า ดัชนีดังกล่าวมีความสัมพันธ์กันทั้งในเชิงบวกและลบในระดัที่ต่างกัน เมื่อพิจารณาจากแบบแผนการตอบที่ได้จากแบบสอบที่วัดทักษะด้านคณิตศาสตร์พบว่าค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างดัชนีดังกล่าวมีค่าอยู่ระหว่าง -0.99 ถึง 0.99 โดยที่ดัชนีชี้เดือนที่ปรับแก้แล้ว มีความสัมพันธ์ต่ำสุดกับดัชนีความสอดคล้อง และเป็นความสัมพันธ์ในเชิงผกผันสำหรับทักษะด้านการอ่านให้ผล เช่นเดียวกับด้านทักษะทางคณิตศาสตร์เป็นส่วนใหญ่

เมื่อนิยามความสัมพันธ์กับคะแนนรวม พบว่าแบบแผนการตอบทั้งจากด้านคณิตศาสตร์ และด้านการอ่านให้ผลที่สอดคล้องกัน คือ ดัชนีชี้เตือนที่ปรับแก้แล้วมีความสัมพันธ์กับคะแนนรวมต่ำสุด และเป็นในทางลบ คือ ค่าสัมประสิทธิ์สหสัมพันธ์เป็น -0.02 และ -0.21 สำหรับด้านคณิตศาสตร์ และด้านการอ่าน ตามลำดับ ในขณะที่ดัชนีความสอดคล้องมีค่าสัมประสิทธิ์สหสัมพันธ์กับคะแนนรวม สูงสุด คือ เท่ากับ 0.99 ทั้งด้านคณิตศาสตร์ และการอ่าน

จากผลที่ได้ตามตารางที่ 6 ฮาร์นิชและลินน์ จึงได้คัดเลือกเอาดัชนีชี้เตือนที่ปรับแก้แล้ว มาใช้เพื่อวินิจฉัยความแตกต่างในด้านการจัดการเรียนการสอนของโรงเรียนในแต่ละเขตการศึกษา เพราะเป็นดัชนีที่ค่อนข้างเป็นอิสระจากคะแนนรวมของผู้สอบ การศึกษาครั้งนี้ครอบคลุมเขตการศึกษาที่ 1 ถึงเขตการศึกษาที่ 5 ของมลรัฐอิลลินอย ซึ่งมีโรงเรียนที่อยู่ในแต่ละเขตเป็นจำนวนดังนี้ คือ เขตการศึกษาที่ 1 มี 50 โรงเรียน เขตการศึกษาที่ 2 มี 14 โรงเรียน เขตการศึกษาที่ 3 มี 18 โรงเรียน เขตการศึกษาที่ 4 มี 15 โรงเรียน และเขตการศึกษาที่ 5 มี 13 โรงเรียน ผลการวิเคราะห์ ความแตกต่างด้วยการวิเคราะห์ความแปรปรวน (ANOVA) ปรากฏตามตารางที่ 7

ตารางที่ 7 ผลการวิเคราะห์ความแตกต่างของโรงเรียนในแต่ละเขตการศึกษาด้วยค่าสถิติทดสอบเอฟ (F-test)

ผล	คณิตศาสตร์	การอ่าน
เขตการศึกษา	0.57	2.12
โรงเรียนในแต่ละเขตการศึกษา	1.64**	1.74**
โรงเรียนในเขตการศึกษาที่ 1	1.69*	1.97**
โรงเรียนในเขตการศึกษาที่ 2	2.78**	1.73
โรงเรียนในเขตการศึกษาที่ 3	0.46	0.66
โรงเรียนในเขตการศึกษาที่ 4	1.37	1.35
โรงเรียนในเขตการศึกษาที่ 5	2.17*	2.83**

* $p < 0.01$

** $p < 0.001$

จากตารางที่ 7 สำหรับด้านคณิตศาสตร์ พบว่าโรงเรียนภายในเขตการศึกษาที่ 1 เขตการศึกษาที่ 2 และเขตการศึกษาที่ 5 ให้ค่าของดัชนีชี้วัดที่ปรับแก้แล้วที่แตกต่างกันอย่างมีนัยสำคัญ ส่วนด้านการอ่านพบว่าโรงเรียนภายในเขตการศึกษาที่ 1 และเขตการศึกษาที่ 5 มีค่าของดัชนีที่แตกต่างกันอย่างมีนัยสำคัญ จากการวิเคราะห์เนื้อหาแบบสอยย่อยพบว่าสาเหตุของความแตกต่างด้านการจัดการเรียนการสอนนั้นเนื่องมาจากลักษณะเฉพาะของเนื้อหา และความครอบคลุมของเนื้อหา

ฮาร์นิช (Harnisch 1981: 1-27) วิเคราะห์แบบแผนการตอบข้อสอบรายบุคคลที่ได้จากแบบสอบอิงเกณฑ์ โดยมีเกณฑ์การเรียงทักษะย่อยจากยากไปง่าย 2 เกณฑ์ คือ เรียงตามจำนวนผู้สอบทั้งกลุ่มที่ตอบทักษะดังกล่าวถูก และเรียงตามจำนวนข้อสอบที่ผู้สอบเป็นรายคนนั้นตอบถูก ซึ่งงานวิจัยครั้งนี้ ใช้ดัชนีการคล้อยตามกลุ่มในกรณีที่เรียงทักษะย่อยตามเกณฑ์จำนวนผู้สอบทั้งกลุ่มที่ตอบทักษะย่อยดังกล่าวได้ถูกต้อง และใช้ดัชนีความสอดคล้องรายบุคคล (Individual Consistency Index) ในกรณีที่เรียงทักษะย่อยตามจำนวนข้อสอบที่ผู้สอบเป็นรายคนนั้นตอบได้ถูกต้อง กลุ่มตัวอย่างที่ศึกษามี 3 กลุ่ม คือ กลุ่มนักเรียนเกรด 3 กลุ่มนักเรียนเกรด 6 และกลุ่มนักเรียนเกรด 7-8 จำนวน 357 คน 396 คน และ 643 คน ตามลำดับ ซึ่งได้รับการทดสอบด้วยแบบสอบอิงเกณฑ์ที่ประเมินทักษะทางด้านคณิตศาสตร์พื้นฐาน (Basic Arithmetic Skills Evaluation: BASE) โดยที่แบบสอบดังกล่าวประกอบด้วยทักษะย่อย จำนวน 23 ทักษะ 21 ทักษะ และ 40 ทักษะ สำหรับเกรด 3 เกรด 6 และเกรด 7-8 ตามลำดับ และแต่ละทักษะย่อยจะประกอบด้วยข้อสอบจำนวน 3 ข้อ ผลการวิเคราะห์พบว่า

1. ค่าคะแนนควอไทล์ของดัชนีการคล้อยตามกลุ่ม และดัชนีความสอดคล้องรายบุคคลของกลุ่มนักเรียนเกรด 3 และเกรด 7-8 มีค่าใกล้เคียงกัน ในขณะที่ค่าของดัชนีดังกล่าวของนักเรียนเกรด 6 มีค่าต่ำมาก การที่ดัชนีการคล้อยตามกลุ่มมีค่าต่ำแสดงว่าการเรียงลำดับของทักษะย่อยตามจำนวนผู้สอบทั้งกลุ่มที่ตอบทักษะดังกล่าวถูกนั้น ไม่สอดคล้องกับแบบแผนการตอบข้อสอบของผู้สอบ และยิ่งดัชนีความสอดคล้องรายบุคคลมีค่าต่ำยิ่งแสดงว่ามีความแตกต่างในการเรียนรู้ทักษะย่อยๆ ของผู้สอบในกลุ่ม

2. ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างดัชนีการคล้อยตามกลุ่ม และดัชนีความสอดคล้องรายบุคคล ของนักเรียนเกรด 3 อยู่ในช่วง .714 ถึง .788 ส่วนค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างดัชนีการคล้อยตามกลุ่มกับคะแนนรวมมีค่าสูงสุดในนักเรียนเกรด 6 คือ มีค่าเท่ากับ .412 ในขณะที่

นักเรียนเกรด 3 และเกรด 7-8 มีค่าเท่ากับ .156 และ .147 ตามลำดับ ส่วนค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างดัชนีความสอดคล้องรายบุคคลกับคะแนนรวมมีค่าสูงสุดในกลุ่มนักเรียนเกรด 6 คือ มีค่าเท่ากับ .326 และมีค่าเท่ากับ .163 และ .125 ในกลุ่มนักเรียนเกรด 3 และเกรด 7-8 ตามลำดับ

3. ความแตกต่างของค่าดัชนี เมื่อผู้สอบมีเชื้อชาติ และมีผู้สอนต่างกัน พบว่า

3.1 กลุ่มนักเรียนระดับเกรด 6 และเกรด 7-8 พบว่า ดัชนีการคล้อยตามกลุ่มจะมีค่าแตกต่างกันเมื่อผู้สอนต่างกัน

3.2 เฉพาะกลุ่มนักเรียนเกรด 7-8 เท่านั้น ที่พบว่า ดัชนีความสอดคล้องรายบุคคลจะมีค่าแตกต่างกันเมื่อผู้สอนต่างกัน

3.3 เมื่อนักเรียนมีเชื้อชาติต่างกัน กลุ่มนักเรียนทั้ง 3 กลุ่ม มีค่าดัชนีการคล้อยตามกลุ่มไม่แตกต่างกัน เช่นเดียวกับดัชนีความสอดคล้องรายบุคคล

4. ความแตกต่างของค่าดัชนี เมื่อนำคะแนนรวมเป็นตัวแปรร่วมในการวิเคราะห์ความแปรปรวนร่วม (ANCOVA) พบว่าดัชนีการคล้อยตามกลุ่มจะมีค่าแตกต่างกันเมื่อผู้สอนต่างกันในเฉพาะในกลุ่มนักเรียนเกรด 7-8 เท่านั้น ส่วนตัวแปรด้านเชื้อชาติไม่มีผลให้ดัชนีการคล้อยตามกลุ่มมีค่าแตกต่างกัน ซึ่งผลที่ได้นี้สอดคล้องกับที่ได้จากดัชนีความสอดคล้องรายบุคคล

เบอร์นฮอมและทาทซูกะ (Birenbaum and Tatsuoka in press อ้างถึงใน Tatsuoka and Tatsuoka 1982: 221) นำคะแนนรวมและดัชนีการคล้อยตามกลุ่ม มาวินิจฉัยความแตกต่างของวิธีการสอนคณิตศาสตร์เรื่องการบวก-ลบ 2 วิธี กลุ่มตัวอย่างเป็นนักเรียนเกรด 7 ของ โรงเรียนมัธยมศึกษาตอนปลาย ซึ่งถูกแบ่งแบบสุ่มเป็น 2 กลุ่ม กลุ่มที่ 1 มีจำนวน 67 คน และกลุ่มที่ 2 มีจำนวน 62 คน กลุ่มที่ 1 เรียนจากโปรแกรมสำเร็จรูปที่สร้างขึ้นไว้ในคอมพิวเตอร์ ส่วนกลุ่มที่ 2 ครูสอนเอง หลังจากนั้นนำผู้สอบทั้งสองกลุ่มมาทดสอบด้วยแบบสอบที่วัดเนื้อหาเรื่องการบวก-ลบ จำนวน 52 ข้อ ผลการศึกษาพบว่าค่าเฉลี่ยของคะแนนรวมของผู้สอบทั้งสองกลุ่มซึ่งทดสอบด้วยสถิติทดสอบที (t-test) มีค่าไม่แตกต่างกัน (ค่าเฉลี่ยของคะแนนรวมของกลุ่มที่ 1 และกลุ่มที่ 2 มีค่าเท่ากับ 20.06 และ 18.36 ตามลำดับ) ส่วนค่าเฉลี่ยของดัชนีการคล้อยตามกลุ่มของผู้สอบทั้งสองกลุ่มซึ่งทดสอบด้วยสถิติทดสอบที มีค่าแตกต่างกันอย่างมีนัยสำคัญ (ค่าเฉลี่ยของดัชนีการคล้อยตามกลุ่มของกลุ่มที่ 1 และกลุ่มที่ 2 มีค่าเท่ากับ .375 และ 0.216 ตามลำดับ) สรุปได้ว่าการใช้ดัชนีการคล้อยตามกลุ่มเป็นดัชนีวิเคราะห์ความแตกต่างของ

วิธีการสอนนั้นมีประสิทธิภาพกว่าการใช้คะแนนรวม

ฮาร์นิชและทอร์เรส (Harnisch and Torres 1983: บทคัดย่อ) นำดัชนีชี้เดือนที่ปรับแก้แล้ว มาใช้วิเคราะห์แบบแผนการตอบข้อสอบที่ได้จากการตอบแบบสอบวัดความเข้าใจการอ่าน กลุ่มตัวอย่างเป็นเด็กอายุ 9 ปี 13 ปี และ 17 ปี ซึ่งทำการประเมินโดย The National Assessment of Educational Progress (NAEP) ในปี ค.ศ. 1979 การศึกษาครั้งนี้เพื่อจะวิเคราะห์ความแตกต่างของแบบแผนการตอบข้อสอบที่เนื่องมาจากตัวแปรด้านโรงเรียนและลักษณะของผู้เรียน ซึ่งมีผลให้ค่าของดัชนีดังกล่าวมีค่าสูงคือทำให้แบบแผนการตอบข้อสอบมีลักษณะที่แตกต่างไปจากกัตแมนสมบูรณ์ ผลการวิจัยนี้ช่วยอธิบายถึงความแตกต่างของสภาพแวดล้อมของโรงเรียนในท้องถิ่นซึ่งนักเรียนเหล่านี้ขาดการเตรียมพร้อมสำหรับการทดสอบ เนื้อหาที่จัดสอนไม่สอดคล้องกับจุดประสงค์ทางการศึกษา

ฮาร์นิช (Harnisch 1983: 191-205) ศึกษาแบบแผนการตอบข้อสอบของผู้สอบเป็นรายบุคคล กลุ่มตัวอย่างเป็นผู้สอบที่มีความสามารถอยู่ในระดับเกรด 4 จำนวน 24 คน ที่ตอบข้อสอบ จำนวน 44 ข้อ ข้อมูลที่ศึกษานี้จำลองขึ้นด้วยโปรแกรม ชื่อ SPP (Student-Problem Package) เพื่อสร้างแบบแผนการตอบข้อสอบของผู้สอบจำนวนดังกล่าว จากนั้นนำแบบแผนการตอบข้อสอบที่ได้มาคำนวณค่าดัชนีชี้เดือนที่ปรับแก้แล้วและคะแนนรวมเพื่อนำมาจำแนกกลุ่มผู้สอบ

ฮาร์นิชได้จำแนกประเภทของผู้สอบตามคะแนนรวม และค่าของดัชนีออกเป็น 4 กลุ่ม คือ กลุ่ม A คือผู้ที่มีคะแนนรวมตั้งแต่ร้อยละ 50 ของคะแนนเต็มขึ้นไป และมีค่าของดัชนีน้อยกว่าหรือเท่ากับ 0.30 พบว่ามีผู้สอบจำนวน 7 คนที่อยู่ในกลุ่มนี้

กลุ่ม B คือผู้ที่มีคะแนนรวมตั้งแต่ร้อยละ 50 ของคะแนนเต็มขึ้นไป และมีค่าของดัชนีมากกว่า 0.3 ขึ้นไป กลุ่มนี้เป็นกลุ่มที่ฮาร์นิชถือว่าข้อสอบที่ผู้สอบตอบผิดส่วนใหญ่มาจากความเลินเล่อ สะเพร่า จากข้อมูลไม่พบว่าผู้สอบคนใดอยู่ในกลุ่มนี้เลย

กลุ่ม C คือผู้ที่มีคะแนนรวมต่ำกว่าร้อยละ 50 ของคะแนนเต็ม และมีค่าของดัชนีน้อยกว่าหรือเท่ากับ 0.3 พบว่ามีผู้สอบจำนวน 11 คนที่ตกอยู่ในกลุ่มนี้ ซึ่งฮาร์นิชอธิบายว่าเป็นผู้ที่ต้องใช้ความพยายามในการเตรียมสอบเรื่องนั้นให้มากขึ้นจึงจะผ่านเกณฑ์

กลุ่ม D คือผู้ที่มีคะแนนรวมต่ำกว่าร้อยละ 50 ของคะแนนเต็มและมีค่าของดัชนีมากกว่า 0.3 ขึ้นไป พบว่ามีผู้สอบจำนวน 6 คนที่ตกอยู่ในกลุ่มนี้ ผู้ที่ตกอยู่ในกลุ่มนี้แสดงว่าขาดความพร้อมในการสอบครั้งนี้ เพราะแบบแผนการตอบข้อสอบขาดความคงเส้นคงวาคือมีลักษณะเป็นแบบสุ่ม

จากข้อมูลนี้จะเห็นได้ว่า ผู้ที่มีความสามารถสูงมักจะทำแบบสอบด้วยความรู้ความสามารถ ดังนั้นจึงมีแบบแผนการตอบข้อสอบที่มีลักษณะใกล้เคียงกับกัตแมนสมบูรณ์ ซึ่งตรงกันข้ามกับผู้ที่มีความสามารถต่ำจะทำข้อสอบโดยการเดาหรือตอบอย่างไม่รู้จริงแบบแผนการตอบข้อสอบจึงมักจะมีลักษณะที่แตกต่างจากกัตแมนสมบูรณ์

ทาซูกะโอกะและทาซูกะโอกะ (Tatsuoka and Tatsuoka 1988: 221-230) นำดัชนีความสอดคล้องของแบบแผนการตอบข้อสอบรายบุคคล จำนวน 3 ดัชนี คือ ดัชนีการคล้อยตามกลุ่ม ดัชนีความสอดคล้องเป็นรายบุคคล และดัชนีชี้เตือนที่ปรับแก้แล้ว มาใช้วินิจฉัยความคลาดเคลื่อนของแบบแผนการตอบข้อสอบเรื่องการบวก-ลบ ของนักเรียนเกรด 8 จำนวน 180 คน ที่เรียนเนื้อหาดังกล่าวจากโปรแกรมคอมพิวเตอร์ที่สร้างขึ้น และของนักเรียนเกรด 7 จำนวน 127 คน ที่ได้รับการสอนในชั้นเรียนเป็นเวลา 3 สัปดาห์ แบบสอบที่ใช้วัดประกอบด้วย 12 ทักษะย่อย แต่ละทักษะมีข้อสอบ 4 ข้อที่ค้ำขนานกัน รวมทั้งสิ้น 48 ข้อ โดยที่แต่ละทักษะจะแทนด้วยสัญลักษณ์ L และ S และทั้ง 12 ทักษะนี้ ครอบคลุมเรื่อง การบวก-ลบ เลขจำนวนบวกและจำนวนลบรวมทั้งจำนวนในวงเล็บ รายละเอียดของทักษะย่อยทั้ง 12 ทักษะ และตัวอย่างของแบบแผนการตอบข้อสอบของนักเรียนเกรด 8 (อยู่ในข้อมูลชุด 1) และของนักเรียนเกรด 7 (อยู่ในข้อมูลชุด 2) จำนวนชั้นเรียนละ 3 คน รวมทั้งค่าดัชนีทั้ง 3 ปรากฏตามตารางที่ 8

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 8 ตัวอย่างของแบบแผนการตอบข้อสอบของนักเรียนเกรด 8 (ข้อมูลชุดที่ 1) และเกรด 7 (ข้อมูลชุดที่ 2) ระดับละ 3 คน และค่าดัชนีการคล้อยตามกลุ่ม ดัชนีความสอดคล้อง เป็นรายบุคคล และดัชนีชี้เตือนที่ปรับแก้แล้ว

ข้อมูลชุดที่ 1	#S1	#S2	#S3	ข้อมูลชุดที่ 2	#S1	#S2	#S3
-L - (-S)	0000	1110	1111	L + -S	1111	1010	1111
S - L	0100	1110	0000	-L + -S	1111	0010	0000
L + -S	1111	1010	1111	-S + L	1011	1111	1111
L - S	0100	1111	1111	S + -L	1111	1111	0111
-L + S	1111	1101	1111	-L + S	1111	1101	1111
-S + L	1011	1111	1111	L - S	0100	1111	1111
S + -L	1111	1111	0111	-L - (-S)	0000	1110	1111
-L + -S	1111	0010	0000	-S - (-L)	0000	1010	1000
-S - (-L)	0000	1010	1000	-S - +L	0111	0100	0000
-S - L	1111	0000	0000	S - L	0100	1110	0000
-S - +L	0111	0100	0000	S - (-L)	0000	0000	0000
S - (-L)	0000	0000	0000	-S - L	1111	0000	0000
คะแนนรวม	28	27	24		28	27	24
C_t	.93	.43	.16		.30	.60	.44
NCI_t	-.10	.62	.69		.49	.44	.67
ICI_t	.96	.88	.99		.96	.88	.99

หมายเหตุ C_t คือ ดัชนีชี้เตือนที่ปรับแก้แล้ว

NCI_t คือ ดัชนีการคล้อยตามกลุ่ม

ICI_t คือ ดัชนีความสอดคล้องรายบุคคล

จากตารางที่ 8 สำหรับข้อมูลชุดที่ 1 จะเห็นว่าผู้สอบคนที่ 1 ซึ่งนับว่าทำข้อสอบได้ดีและมีปัญหาเฉพาะเรื่องการบวกจำนวนติดลบที่อยู่ในวงเล็บเท่านั้น ซึ่งเป็นทักษะส่วนน้อยและเกิดความสับสนได้ง่าย ดังนั้นความคลาดเคลื่อนของเขาจึงเกิดจากคำถามที่ยุ่งยากเกินความจำเป็น (sophistication) ซึ่งเป็นแบบแผนการตอบข้อสอบที่พบได้น้อยเมื่อเทียบกับคนอื่นๆในกลุ่ม ผู้สอบคนนี้มีค่าดัชนีการคล้อยตามกลุ่มต่ำ (-.10) และมีค่าดัชนีชี้เตือนที่ปรับแก้แล้วสูง (.93) ซึ่งถ้าแปลผลตามค่าของดัชนีทั้งสองผู้สอบคนนี้ต้องได้รับการสอนซ่อมเสริม และการนำคะแนนรวมของเขามาใช้จะต้องแปลผลอย่างระมัดระวัง แต่ถ้าพิจารณาที่ค่าของดัชนีความสอดคล้องรายบุคคลแล้วผู้สอบคนนี้มีค่าเท่ากับ .96 ซึ่งเป็นค่าที่สูงมาก และเป็นการบ่งชี้ว่าแบบแผนการตอบของเขามีความสอดคล้องกันดี ไม่ใช่แบบแผนการตอบข้อสอบที่มีลักษณะแตกต่างจากกัตแมนสมบูรณ์ ส่วนผู้สอบคนที่ 2 และคนที่ 3 มีลักษณะของการตอบทักษะย่อยที่เหมือนกับผู้สอบในกลุ่ม ค่าดัชนีการคล้อยตามกลุ่มจึงมีค่าค่อนข้างสูง คือมีค่าเท่ากับ .62 และ .69 ตามลำดับ และมีค่าดัชนีชี้เตือนที่ปรับแก้แล้วต่ำ คือมีค่าเท่ากับ .43 และ .16 ตามลำดับ และมีค่าดัชนีความสอดคล้องรายบุคคลเท่ากับ 0.88 และ 0.99 ตามลำดับ สำหรับผู้สอบสองคนนี้ค่าของดัชนีทั้งสามจะให้ความหมายของแบบแผนการตอบข้อสอบที่สอดคล้องกัน คือผู้สอบคนที่ 3 มีแบบแผนการตอบข้อสอบที่มีระบบแบบแผนมากกว่าผู้สอบคนที่ 2

สำหรับข้อมูลชุดที่ 2 เมื่อพิจารณาจากค่าดัชนีชี้เตือนที่ปรับแก้แล้ว จะเห็นได้ว่าผู้สอบคนที่ 1 มีแบบแผนการตอบข้อสอบที่มีระบบแบบแผนมากที่สุด รองลงมาได้แก่ผู้สอบคนที่ 3 และผู้สอบคนที่ 2 มีแบบแผนการตอบข้อสอบที่มีระบบแบบแผนน้อยที่สุด ค่าดัชนีชี้เตือนที่ปรับแก้แล้วของผู้สอบคนที่ 1 คนที่ 2 และคนที่ 3 มีค่าเท่ากับ .30 .60 และ .44 ตามลำดับ แต่เมื่อพิจารณาจากค่าดัชนีการคล้อยตามกลุ่ม และดัชนีความสอดคล้องรายบุคคล ซึ่งให้ผลสอดคล้องกันว่าผู้สอบคนที่ 3 มีแบบแผนการตอบข้อสอบที่มีระบบแบบแผนมากที่สุด รองลงมาได้แก่ผู้สอบคนที่ 1 และผู้สอบคนที่ 2 มีแบบแผนการตอบที่มีระบบแบบแผนน้อยที่สุด ค่าดัชนีการคล้อยตามกลุ่มของผู้สอบคนที่ 1 คนที่ 2 และคนที่ 3 มีค่าเท่ากับ .49 .44 และ .67 ตามลำดับ และค่าดัชนีความสอดคล้องรายบุคคลของผู้สอบคนที่ 1 คนที่ 2 และคนที่ 3 มีค่าเท่ากับ .96 .88 และ .99 ตามลำดับ

การที่ดัชนีทั้งสามให้ความหมายของแบบแผนการตอบข้อสอบของผู้สอบแตกต่างกันไป เนื่องจากดัชนีการคล้อยตามกลุ่ม และดัชนีชี้เตือนที่ปรับแก้แล้ว มีการเรียงทักษะย่อยจากง่ายไปยากโดยใช้จำนวนผู้สอบในกลุ่มที่ตอบทักษะดังกล่าวถูก ดังนั้นแบบแผนการตอบของผู้สอบจึงขึ้นกับความรู้ของผู้สอบในกลุ่ม ในขณะที่ดัชนีความสอดคล้องรายบุคคลนั้นจะเรียงทักษะย่อยจากง่ายไปยากโดยใช้จำนวนข้อสอบที่ผู้สอบคนนั้นๆตอบทักษะดังกล่าวได้ถูกต้อง การคำนวณค่าดัชนีของผู้สอบแต่ละ

คนต้องมีการจัดเรียงทักษะใหม่ตามความสามารถของเขา แต่ก็มีควมจำกัดว่าจะต้องมีข้อสอบที่เป็น
 คำนานกันตั้งแต่ 3 ข้อขึ้นไป

รัดเนอร์ (Rudner 1983: 207-219) ศึกษาความตรงของดัชนีที่ใช้ประเมินความ
 ถูกต้องของคะแนนสอบที่ยึดตามแนวกัตแมน จำนวน 4 ดัชนี คือ ดัชนีของไบบีเรียลของแต่ละบุคคล
 ดัชนีไบบีเรียลของแต่ละบุคคล ดัชนีการคล้อยตามกลุ่ม และดัชนีชี้เตือนที่ปรับแก้แล้ว กับดัชนีใน
 กลุ่มที่ยึดตามแนวทฤษฎีการตอบสนองรายข้อ จำนวน 5 ดัชนี คือ ดัชนีดับเบิลยู 1 (W1) ดัชนีดับ
 เบิลยู 3 (W3) ดัชนียู 1 (U1) ดัชนียู 3 (U3) และดัชนีแอล 3 (L3) ข้อมูลที่ใช้จำลอง
 ขึ้นด้วยโปรแกรมคอมพิวเตอร์ โดยกำหนดลักษณะของค่าพารามิเตอร์ของข้อสอบ แล้วจึงจำลองค่า
 ความสามารถของผู้สอบ และแบบแผนการตอบข้อสอบของผู้สอบ แบบสอบฉบับที่ 1 กำหนดให้มีค่า
 พารามิเตอร์ตามแบบสอบวัดความถนัดทางการเรียนด้านภาษาที่ใช้สำหรับการค้าของลอร์ด จำนวน
 80 ข้อ ซึ่งมีค่าเฉลี่ยของค่าความยาก อำนาจจำแนก และการเดา เท่ากับ 1.07 0.58 และ
 0.16 ตามลำดับ แบบสอบฉบับที่ 2 กำหนดให้มีพารามิเตอร์ตามแบบสอบวิชาชีววิทยาที่ใช้ในชั้น
 เรียนของบีจาร์ ไวสส์และคิงส์บิวรี่ (Bejar Weiss and Kingsbury) จำนวน 45 ข้อ ซึ่งมี
 ค่าเฉลี่ยของค่าความยาก อำนาจจำแนก และการเดา เท่ากับ 1.09 .08 และ .25 ตามลำดับ
 ความสามารถของกลุ่มผู้สอบที่จำลองขึ้นมีลักษณะการแจกแจงแบบโค้งปกติที่มีค่ามัธยฐานเลขคณิตเท่ากับ
 0 และค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1.00 หลังจากที่ได้จำลองแบบแผนการตอบของผู้สอบได้แล้ว
 นำมาปรับแก้ให้เป็นกลุ่มที่มีคะแนนสอบสูงกว่าความสามารถจริง (spuriously high total
 scores) โดยการลุ่มแก้คำตอบที่ถูกบางข้อให้เป็นคำตอบที่ผิด และให้เป็นกลุ่มที่มีคะแนนสอบต่ำ
 กว่าความสามารถจริง (spuriously low total scores) โดยการลุ่มแก้คำตอบที่ผิดบางข้อ
 ให้เป็นคำตอบที่ถูก ความตรงของดัชนีพิจารณาจากสัดส่วนของจำนวนผู้สอบที่มีคะแนนสอบสูงกว่า
 ความสามารถจริง และจำนวนผู้สอบที่มีคะแนนสอบต่ำกว่าความสามารถจริง ต่อจำนวนผู้สอบทั้งหมด
 ซึ่งวินิจฉัยได้จากค่าของดัชนี ผลการวิเคราะห์พบว่า

1. เมื่อจำนวนข้อสอบเพิ่มขึ้น (จาก 5 ข้อ เป็น 10 ข้อ และ 15 ข้อ) ดัชนีส่วนใหญ่
 จะยังมีความถูกต้อง ยกเว้นดัชนีไบบีเรียลของแต่ละบุคคล และดัชนียู 1 ในกรณีของแบบสอบฉบับ
 ที่ 1 และกลุ่มผู้สอบที่มีคะแนนสอบสูงกว่าความสามารถจริง กับดัชนียู 3 ในกรณีของแบบสอบฉบับที่
 2 ในกลุ่มผู้สอบที่มีคะแนนสอบสูงกว่าความสามารถจริง

2. แบบสอบฉบับที่ 1 พบว่า

2.1 ในกลุ่มผู้สอบที่มีคะแนนสอบสูงกว่าความสามารถจริงนั้น ดัชนีดับเบิลยู 1 และ ดัชนีแอล 3 มีความสามารถในการวินิจฉัยความถูกต้องของคะแนนสอบสูง

2.2 ในกลุ่มผู้สอบที่มีคะแนนสอบต่ำกว่าความสามารถจริงนั้น ดัชนีแอล 3 มีความสามารถในการวินิจฉัยความถูกต้องของคะแนนสอบสูงสุด

3. แบบสอบฉบับที่ 2 พบว่า

3.1 ในกลุ่มผู้สอบที่มีคะแนนสอบสูงกว่าความสามารถจริงนั้น ดัชนีดับเบิลยู 3 มีความสามารถในการวินิจฉัยความถูกต้องของคะแนนสอบสูงสุด

3.2 ในกลุ่มผู้สอบที่มีคะแนนสอบต่ำกว่าความสามารถจริงนั้น ดัชนีดับเบิลยู 1 และ ดัชนีแอล 3 มีความสามารถในการวินิจฉัยความถูกต้องของคะแนนสอบสูง

4. ดัชนีการคล้อยตามกลุ่ม และดัชนีชี้เตือนที่ปรับแก้แล้ว พบว่ามีความสามารถในการวินิจฉัยความถูกต้องของคะแนนสอบได้นอกกันทั้งจากแบบสอบฉบับที่ 1 และฉบับที่ 2 และจากกลุ่มผู้สอบที่มีคะแนนสอบต่ำกว่าความสามารถจริงและกลุ่มผู้สอบที่มีคะแนนสอบสูงกว่าความสามารถจริง

บลีซท์ และ ไดเนอโร (Blixt and Dinero 1985: 293-299) ศึกษาคุณภาพด้านความตรงของดัชนีชี้เตือนของซาโต แต่ได้เลือกดัชนีชี้เตือนที่ปรับแก้แล้วมาใช้แทนเพราะมีจุดเด่นในด้านพิสัยของดัชนี และการกำหนดจุดตัดที่จะบอกว่าแบบแผนนั้นมีความผิดปกติหรือไม่ กลุ่มตัวอย่างเป็นนักศึกษาระดับบัณฑิตศึกษา จำนวน 60 คน ที่เรียนวิชาสถิติเบื้องต้นซึ่งมีการสอบย่อย (quiz) ทุก 2 สัปดาห์ โดยมีการเฉลยคำตอบและแจกแบบสอบถามสั้นๆ มีข้อความถามอยู่ 13 ข้อ ให้ผู้สอบเลือกว่าทำไมเขาถึงตอบผิดในข้อสอบข้อนั้นๆ บลีซท์และไดเนอโรได้นำผลการสอบย่อยที่เขาคิดว่าเชื่อถือได้มาศึกษาถึงจำนวนความถี่ของการเลือกเหตุผลที่ตอบผิดของผู้สอบที่เขาจัดแบ่งเป็น 4 ประเภท โดยบลีซท์และไดเนอโรได้กำหนดประเภทของเหตุผลที่ได้จากแบบสอบถาม คะแนนรวม และค่าของดัชนีชี้เตือนที่ปรับแก้แล้ว จากการตอบข้อสอบจำนวน 14 ข้อ ออกเป็น 4 กลุ่มดังนี้ คือ

กลุ่ม A หมายถึงกลุ่มที่ผู้สอบทำคะแนนรวมได้ตั้งแต่ 11 คะแนนขึ้นไป และมีค่าดัชนีน้อยกว่า 0.20

เหตุผลที่คาดหวังสำหรับผู้สอบกลุ่ม A คือ "เหตุผลที่ให้ทั้งหมดไม่มีข้อใดเหมาะสม"

กลุ่ม B หมายถึงกลุ่มที่ผู้สอบทำคะแนนรวมได้ตั้งแต่ 11 คะแนนขึ้นไป และมีค่าของดัชนีตั้งแต่ 0.20 ขึ้นไป

เหตุผลที่คาดหวังสำหรับผู้สอบกลุ่ม B คือ

1. ไม่ได้อ่านข้อสอบข้อนี้อย่างระมัดระวังพอ
2. ทำเครื่องหมายคำตอบในกระดาษคำตอบผิดพลาด
3. ไม่ได้อ่านตัวเลือกทุกตัว จึงทำให้พลาดข้อที่ถูกต้องที่สุด

กลุ่ม C หมายถึงกลุ่มที่ผู้สอบทำคะแนนรวมได้น้อยกว่า 10 คะแนนและมีค่าของดัชนีน้อยกว่า

0.20

เหตุผลที่คาดหวังสำหรับผู้สอบกลุ่ม C คือ

1. ไม่ได้อ่านเอกสารในส่วนที่ข้อสอบข้อนี้ถาม
2. อ่านเอกสารเพียงคร่าว ๆ เลขจํารายละเอียดตรงนี้ไม่ได้
3. ข้อความในข้อสอบข้อนี้ไม่ชัดเจน

กลุ่ม D หมายถึงกลุ่มที่ผู้สอบทำคะแนนรวมได้น้อยกว่า 10 คะแนน และมีค่าของดัชนี

ตั้งแต่ 0.20 ขึ้นไป

เหตุผลที่คาดหวังสำหรับผู้สอบกลุ่ม D คือ

1. ไม่มีความรู้พื้นฐานทางคณิตศาสตร์ที่จำเป็นต่อการตอบคำถามข้อนี้
2. เข้าใจเอกสารมากมายแต่ไม่มีส่วนที่ข้อสอบข้อนี้ถาม
3. เนื้อหาวิชาที่เกี่ยวข้องกับข้อคำถามข้อนี้เป็นเนื้อหาที่ยาก
4. เวลาที่ใช้สอนเกี่ยวกับเนื้อหาของข้อคำถามข้อนี้น้อยเกินไป

ผลจากการตอบแบบสอบถามเกี่ยวกับเหตุผลที่ตอบผิดของผู้สอบทั้ง 60 คนซึ่งมีคะแนนรวม และค่าดัชนีชี้เตือนที่ปรับแก้แล้ว ที่เป็นไปตามประเภทที่จัดแบ่งไว้ ปรากฏดังภาพที่ 1

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ภาพที่ 1 ความถี่ของการให้เหตุผลในแต่ละประเภท สำหรับผู้เรียนทั้ง 4 กลุ่ม

ค่าของดัชนีชี้เตือนที่ปรับแก้แล้ว

		ต่ำ (0-0.19)		สูง (0.2-0.71)		
		คำตอบที่ได้จากกลุ่ม		คำตอบที่ได้จากกลุ่ม		
		ความถี่	%	ความถี่	%	
สูง (11 - 14) คะแนนรวม	ประเภท B	4	22.22	ประเภท B	11	68.75
	ของ C	13	72.22	ของ C	5	31.25
	เหตุผล D	1	5.55	เหตุผล D	0	-
		n = 19 (13) กลุ่ม A		n = 10 (10) กลุ่ม B		
ต่ำ (0 - 10)	ประเภท B	16	25.40	ประเภท B	14	19.72
	ของ C	36	57.14	ของ C	29	40.85
	เหตุผล D	11	17.46	เหตุผล D	28	39.44
		n = 17 (16) กลุ่ม C		n = 14 (14) กลุ่ม D		

หมายเหตุ ตัวเลขในวงเล็บ คือ จำนวนผู้สอบที่ตอบมากกว่า 1 ประเภท

จากภาพที่ 1 สรุปได้ว่าผู้สอบทั้ง 60 คน ทำคะแนนรวมได้ระหว่าง 2 ถึง 14 คะแนน โดยที่การกระจายของคะแนนรวมมีลักษณะเป็นแบบเบ้ซ้าย ในขณะที่ค่าของดัชนีชี้เตือนมีค่าอยู่ระหว่าง 0

ถึง 0.71 โดยมีลักษณะการกระจายแบบเบ้ขวา เมื่อพิจารณาความถี่ของเหตุผลในแต่ละประเภทกับกลุ่มที่จัดให้ จะเห็นว่ามีความเฉพาะกลุ่ม B เท่านั้น ที่มีความสอดคล้องของเหตุผลที่ผู้สอบในกลุ่มตอบกับที่คาดหวัง คือ มีผู้ที่ให้เหตุผลประเภท B ถึง 68.75 % ในขณะที่กลุ่ม A และกลุ่ม C มีผู้สอบที่ให้เหตุผลประเภท C สูงสุดถึง 72.22 % และ 57.14 % ตามลำดับ ส่วนกลุ่ม D ให้เหตุผลประเภท C และ D ใกล้เคียงกัน คือ 40.85 % และ 39.44 % ผลที่ปรากฏเช่นนี้ บลิซท์และไดเนอโร อธิบายว่าน่าจะเป็นเพราะปัจจัยแทรกซ้อนเกี่ยวกับการยอมรับของสังคม (social desirability) จึงทำให้ผู้สอบส่วนใหญ่เลือกคำตอบประเภท C ซึ่งเป็นกลุ่มของการให้เหตุผลว่าตนเองยังเตรียมตัวไม่พอสำหรับการสอบและต้องการการกระตุ้นว่าเพิ่มเติม

บูซี (Buxie 1986: บทคัดย่อ) ได้พัฒนาดัชนีความสอดคล้องของแบบแผนการตอบข้อสอบของผู้สอบเป็นรายบุคคลเรียกว่าดัชนีความสอดคล้องเป็นรายบุคคล (Individual Consistency Index: ICI) และนำมาเปรียบเทียบกับดัชนีที่เตือนที่ปรับแก้แล้ว (The Modified Caution Index: C₁) เพื่อศึกษาการเปลี่ยนแปลงเมื่อค่าความยากของข้อสอบ จำนวนข้อสอบ และลักษณะการเดาแตกต่างกันไป ผลการศึกษาพบว่าเมื่อแบบสอบมีจำนวนข้อต่ำสุด ทำให้ค่าของดัชนีทั้งสองขาดความไวในการบ่งชี้แบบแผนการตอบข้อสอบที่มีลักษณะแตกต่างไปจากกัตแมนสมบูรณ์ สำหรับดัชนีของบูซี จะมีค่าสูงสุดเมื่อแบบสอบมีความยากปานกลาง และมีค่าต่ำสุดเมื่อแบบสอบยาก และจะบอกถึงการเดาเมื่อมีการเดาถึง 70 % ของข้อสอบทั้งหมด สำหรับดัชนีที่เตือนที่ปรับแก้แล้วพบว่า มีปฏิสัมพันธ์ระหว่างค่าความยากของข้อสอบกับความยาวของแบบสอบ ส่วนใหญ่ค่าของดัชนีจะมีค่าต่ำสุดเมื่อแบบสอบยากที่สุด ยกเว้นเมื่อแบบสอบมีจำนวนข้อต่ำสุด ค่าของดัชนีที่ได้จากแบบสอบที่มีค่าความยากสูงสุดจะบอกถึงระดับของการเดามากที่สุด โดยสรุปคือค่าของดัชนีที่เตือนที่ปรับแก้แล้วจะมีค่าเฉลี่ยต่ำสุด ค่าความแปรปรวนน้อยสุด และมีความไวต่อการเปลี่ยนแปลงของความยาวของแบบสอบต่ำสุดเมื่อแบบสอบนั้นยากที่สุด

สุนันท์ ศลโกสม (2530: 112-132) ได้เปรียบเทียบผลการวิเคราะห์และศึกษาความสัมพันธ์ระหว่างวิธีวิเคราะห์ผลการสอบวัดผลสัมฤทธิ์วิชาภาษาไทยด้วยทฤษฎีการทดสอบดั้งเดิม ดัชนีที่เตือนของซาโต และทฤษฎีการตอบสนองข้อกระทงซึ่งใช้ค่าความสามารถและค่าดัชนีชี้นำของทาทชูโอกะ ข้อมูลที่ใช้วิเคราะห์ได้จากการตรวจสอบคุณภาพการศึกษาวិชาภาษาไทย ชั้นมัธยมศึกษาปีที่ 6 ทั่วประเทศ ในปีการศึกษา 2528 ของสำนักงานทดสอบทางการศึกษา กรมวิชาการ กระทรวง

ศึกษาธิการ แบบสอบที่ใช้เป็นแบบสอบเลือกตอบ 4 ตัวเลือก จำนวน 60 ข้อ ครอบคลุมเนื้อหา และพฤติกรรมสมรรถภาพของวิชาภาษาไทย ด้านสื่อความเข้าใจข้อความภาษาไทย ด้านความสามารถใช้ภาษาอย่างมีประสิทธิภาพ ด้านความสามารถคิดอย่างมีวิจารณญาณ และด้านความสามารถใช้ภาษาเชิงสร้างสรรค์ กลุ่มตัวอย่างได้มาจากการสุ่มแบบหลายขั้นตอน จากประชากร จำนวน 50,201 คน จากเขตการศึกษาที่ 9, 10 และ 11 จำนวน 5,774 คน นำผลการสอบมาวิเคราะห์ความแตกต่างของค่าเฉลี่ยของผลสัมฤทธิ์วิชาภาษาไทยจากค่าคะแนนรวม ค่าดัชนีชี้เตือนของซาโต ค่าความสามารถและค่าดัชนีชี้นำของทาทซึโอกะ โดยใช้แบบแผนการวิเคราะห์ด้วยไฮราจิคัล $p \times q \times r$ ($p \times q \times r$ Hierarchical Design) จากระดับเขตการศึกษา (p) ระดับจังหวัดในเขตการศึกษา (q) และระดับโรงเรียนในจังหวัด (r) และเปรียบเทียบความเข้มของความสัมพันธ์ระหว่างสถานของสถานศึกษากับผลสัมฤทธิ์วิชาภาษาไทย ที่วิเคราะห์ด้วยวิธีทั้ง 4 รวมทั้งหาความสัมพันธ์ระหว่างดัชนีที่ได้จากการวิเคราะห์ทั้ง 4 วิธีนั้น ผลการวิจัยพบว่า

1. การวิเคราะห์ความแตกต่างของผลสัมฤทธิ์วิชาภาษาไทย

1.1 ความแตกต่างระหว่างเขตการศึกษา ดัชนีชี้นำของทาทซึโอกะเท่านั้นที่บอกความแตกต่างดังกล่าวได้อย่างมีนัยสำคัญ

1.2 ความแตกต่างระหว่างจังหวัดในเขตการศึกษา ดัชนีชี้นำของซาโต และค่าประมาณความสามารถเป็นดัชนีที่บอกความแตกต่างดังกล่าวได้อย่างมีนัยสำคัญ

1.3 ความแตกต่างระหว่างโรงเรียนในจังหวัด ทุกดัชนีจะบอกถึงความแตกต่างได้อย่างมีนัยสำคัญ

2. ความเข้มของความสัมพันธ์ระหว่างสถานของสถานศึกษากับผลสัมฤทธิ์วิชาภาษาไทย

2.1 อิทธิพลของเขตการศึกษาที่ทำให้เกิดความแตกต่างของผลสัมฤทธิ์ ระหว่างเขตการศึกษา พบว่า ดัชนีชี้นำของทาทซึโอกะเท่านั้นที่แสดงถึงอิทธิพลดังกล่าว โดยมีความเข้มของความสัมพันธ์ระหว่างเขตศึกษากับคะแนนผลสัมฤทธิ์เท่ากับร้อยละ 1

2.2 อิทธิพลของจังหวัดที่ทำให้เกิดความแตกต่างของผลสัมฤทธิ์ระหว่างจังหวัดในเขตการศึกษา พบว่าดัชนีชี้เตือนของซาโตและค่าประมาณความสามารถเป็นดัชนีที่แสดงถึงอิทธิพลดังกล่าว โดยมีค่าความเข้มเท่ากับร้อยละ 3.3 และร้อยละ 3.4 ตามลำดับ

2.3 อิทธิพลของโรงเรียนที่ทำให้เกิดความแตกต่างของผลสัมฤทธิ์ ระหว่างโรงเรียนในจังหวัด พบว่า ดัชนีทุกตัวแสดงถึงอิทธิพลดังกล่าว โดยที่คะแนนรวม ค่าประมาณความสามารถ ค่าดัชนีชี้เตือนของซาโต และค่าดัชนีชี้นำของทาทซึโอกะ ให้ค่าความเข้มเท่ากับร้อยละ 7.3 ร้อย

ละ 5.0 ร้อยละ 3.8 และร้อยละ 1.4 ตามลำดับ

3. ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างดัชนีทั้ง 4 พบว่า คะแนนรวมมีความสัมพันธ์อย่างมีนัยสำคัญกับค่าประมาณความสามารถ ดัชนีชี้เดือนของซาโต และดัชนีชี้หน้าของทาทซึโอกะ โดยมีค่าสัมประสิทธิ์สหสัมพันธ์ เท่ากับ 0.8297 เท่ากับ -0.5698 และเท่ากับ 0.3445 ตามลำดับ สำหรับดัชนีชี้เดือนของซาโต มีความสัมพันธ์อย่างมีนัยสำคัญกับค่าความสามารถ และดัชนีชี้หน้าของทาทซึโอกะ โดยมีค่าสัมประสิทธิ์สหสัมพันธ์เท่ากับ -0.8210 และ 0.2889 ตามลำดับ แต่ค่าประมาณความสามารถไม่มีความสัมพันธ์กับดัชนีชี้หน้าของทาทซึโอกะ

สำราญ มีแจ้ง (2534: 150-157) ได้พัฒนาดัชนีชี้ความบกพร่องของผู้ตอบแบบทดสอบ โดยการวิเคราะห์แบบแผนการตอบข้อสอบ และเปรียบเทียบคุณภาพในการวินิจฉัยความบกพร่องของผู้ตอบแบบทดสอบในแบบแผนการตอบระหว่างดัชนีชี้ความบกพร่องที่พัฒนาขึ้น กับดัชนีชี้เดือนของซาโต การศึกษาครั้งนี้ใช้เทคนิคมอนติคาร์โลซิมูเลชัน จำลองการทดลองด้วยเครื่องคอมพิวเตอร์ โดยกำหนดลักษณะประชากรให้มีการแจกแจงความสามารถแตกต่างกันเป็น 3 ลักษณะ คือ กลุ่มผู้สอบที่มีความสามารถสูง ปานกลาง และต่ำ แต่ละกลุ่มมีการสุ่มตัวอย่างออกมา 3 ขนาด คือ 35 คน 50 คน และ 200 คน แบบสอบที่ใช้ในการวิจัยกำหนดขึ้นเป็นแบบสอบเลือกตอบ ประกอบด้วยแบบสอบ 4 ฉบับ ซึ่งมีจำนวนข้อสอบ 4 ขนาด คือ 30 ข้อ 60 ข้อ 90 ข้อ และ 120 ข้อ แต่ละฉบับ มีค่าความยากตั้งแต่ -3 ถึง +3 ผลการวิจัยพบว่าดัชนีชี้ความบกพร่องที่พัฒนาขึ้น

1. ดัชนีดีเอฟซี1 และดัชนีดีเอฟซี2 มีความตรงในการจำแนกผู้ตอบแบบทดสอบที่มีความบกพร่อง โดยปรากฏว่าค่าความคลาดเคลื่อนมาตรฐานของการวัดที่คำนวณจากกลุ่มที่คัดเลือกผู้ตอบแบบสอบที่มีความบกพร่องในการตอบออกด้วยการใช้ดัชนีดีเอฟซี1 และดีเอฟซี2 มีค่าต่ำกว่าค่าความคลาดเคลื่อนมาตรฐานของการวัดที่คำนวณจากกลุ่มผู้สอบทั้งหมดอย่างมีนัยสำคัญทางสถิติ ในทุกขนาดของกลุ่มตัวอย่างที่แบบสอบมีจำนวนข้อสอบเท่ากับ 30 ข้อ และ 60 ข้อ

2. ดัชนีดีเอฟซี1 และดัชนีดีเอฟซี2 มีความสัมพันธ์ในทางบวกกับดัชนีชี้เดือนของซาโตอย่างมีนัยสำคัญ โดยที่ดัชนีดีเอฟซี2 มีความสัมพันธ์กับดัชนีชี้เดือนของซาโตสูงกว่าดัชนีดีเอฟซี1 (มีค่าสัมประสิทธิ์สหสัมพันธ์อยู่ระหว่าง .921 ถึง .990)

3. ดัชนีดีเอฟซี2 มีความสัมพันธ์กับคะแนนรวมน้อยกว่าดัชนีชี้เดือนของซาโตอย่างมีนัยสำคัญ ในกลุ่มประชากรที่มีความสามารถสูงและต่ำ ที่ตอบแบบสอบที่มีจำนวนข้อสอบเท่ากับ 30 ข้อ 90 ข้อ และ 120 ข้อ ส่วนดัชนีดีเอฟซี1 มีความสัมพันธ์กับคะแนนรวมน้อยกว่าดัชนีชี้เดือนของ

ชาติอย่างมีนัยสำคัญ ในกลุ่มประชากรที่มีความสามารถปานกลางและมีขนาดของกลุ่มตัวอย่างเป็น 25 คน 50 คน และ 100 คน ซึ่งตอบแบบสอบถามที่มีจำนวนข้อสอบเท่ากับ 30 ข้อ และ 120 ข้อ

4. ดัชนีดีเอ็นพี 1 มีประสิทธิภาพในการวินิจฉัยความบกพร่องของผู้ตอบแบบทดสอบในแบบแผนการตอบสูงกว่าดัชนีดีเอ็นพี 2 และดัชนีเดือนของชาติ เมื่อประชากรมีความสามารถปานกลาง ส่วนดัชนีดีเอ็นพี 2 มีประสิทธิภาพในการวินิจฉัยความบกพร่องของผู้ตอบแบบทดสอบในแบบแผนการตอบสูงกว่าดัชนีดีเอ็นพี 1 และดัชนีนี้เดือนของชาติ เมื่อกลุ่มประชากรมีความสามารถสูงและต่ำ



ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย