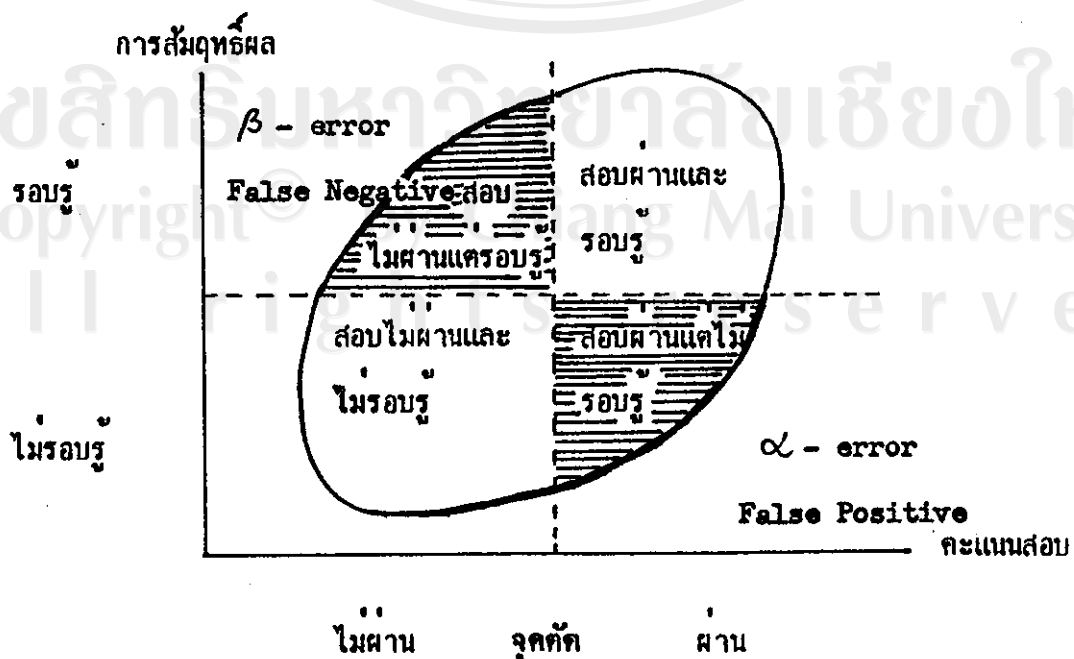


ในการศึกษาเอกสารและผลงานวิจัยที่เกี่ยวข้องกับการศึกษาค้นคว้านี้ ผู้วิจัยได้ศึกษาเนื้อหา และรวบรวมเป็นหมวดหมู่ ตามลำดับต่อไปนี้คือ

1. การกำหนดจุดตัดของแบบทดสอบอิงเกณฑ์
2. ทฤษฎีคุณลักษณะแฝง
3. งานวิจัยที่เกี่ยวข้อง

การกำหนดจุดตัดของแบบทดสอบอิงเกณฑ์

ในการประเมินผลการเรียนแบบอิงเกณฑ์ การกำหนดคะแนนเกณฑ์หรือคะแนนจุดตัด เป็นสิ่งสำคัญมาก ถ้ากำหนดคะแนนจุดตัดต่ำไปก็อาจจะเกิดปัญหากลุ่มที่สอบผ่านแต่ยังไม่รอบรู้ แต่ ถ้ากำหนดคะแนนจุดตัดสูงเกินไปก็อาจจะเกิดปัญหากลุ่มที่สอบไม่ผ่าน ซึ่งจะก่อให้เกิด ปัญหาในด้านการประเมินผลแบบอิงเกณฑ์ขึ้น ดังภาพข้างล่างนี้ (อังคณา สายยศ 2525 หน้า 70-71)



จากภาพส่วนที่แรเงาจะเป็นส่วนที่เกิดความคลาดเคลื่อนจากการประเมินผล เมื่อกำหนดเกณฑ์ผิดพลาด จะมีความคลาดเคลื่อนซึ่งอาจจำแนกได้เป็น 2 ประเภท

1. กลุ่มที่สอบไม่ผ่านแต่รอบรู้ เป็นการผิดพลาดที่เรียกว่า เบต้าเออร์เรอร์ (β - error) ซึ่งเกิดจากการที่กำหนดคะแนนจุดตัดสูงเกินไป ทำให้เด็กสอบตกมาก ทำให้เสียเวลาและกำลังใจ สิ้นเปลืองแรงงานในการที่จะต้องสอนซ้ำ คัดทอนกำลังคนที่มีประโยชน์ของชาติ
2. กลุ่มที่สอบผ่านแต่ไม่จริง เป็นการผิดพลาดที่เรียกว่า แอลฟา เออร์เรอร์ (α - error) ซึ่งเกิดจากการที่กำหนดคะแนนจุดตัดต่ำเกินไป ทำให้เด็กที่มีคุณภาพไม่ถึงมาตรฐานสอบผ่านก็อาจมีผลเสีย เช่น เด็กจะมีความสามารถไม่พอในการเรียนรู้งานหมายใหม่หรือเรียนในระดับสูงต่อไป หรืออาจเป็นพลเมืองที่ไม่มีคุณภาพคือ

จากความผิดพลาดดังกล่าวทำให้นักวิจัยหลายท่านได้พยายามศึกษาและหาวิธีการกำหนดจุดตัดที่ทำให้เกิดความคลาดเคลื่อนน้อยที่สุด ซึ่งเกลส (Glass, 1978 : 243-257) ได้สรุปและรวบรวมไว้ดังนี้

1. วิธีใช้ผลที่ได้จากการปฏิบัติของคนอื่น ๆ เป็นเกณฑ์ (Performance of Other Criterion)

ระดับเกณฑ์หรือคะแนนจุดตัดแบบนี้สร้างขึ้นโดยอ้างอิงค่าพารามิเตอร์ของประชากรผู้เข้าสอบ ดังนั้นคะแนนจุดตัดจึงอาจกำหนดด้วยคะแนนมัธยฐาน (Median) ของคะแนนจากแบบทดสอบ หรืออาจกำหนดเป็นค่าคะแนนเปอร์เซ็นต์ไทล์ เป็นต้น แต่วิธีนี้นักวิจัยหลายท่านแย้งว่าไม่เหมาะสม เพราะยังคงเป็นการอิงกลุ่มอยู่นั่นเอง

2. วิธีนับถอยหลังจาก 100% (Counting Backward from 100%)

เนื่องจากแบบทดสอบอิงเกณฑ์สร้างมาจากจุดประสงค์เชิงพฤติกรรม ดังนั้นการกำหนดคะแนนจุดตัดของแบบทดสอบจึงใช้การพิจารณาจากความสำคัญของจุดประสงค์เชิงพฤติกรรมเป็นหลัก ถ้าจุดประสงค์ใดมีความสำคัญมาก ค่าเกณฑ์ที่ต้องการต้องเป็น 100% ถ้าจุดประสงค์ใดมีความสำคัญรองลงมา ค่าเกณฑ์ที่ต้องการก็จะลดลงมาจาก 100% แต่ตามหลักเหตุผล และประสบการณ์ ส่วนมากแล้วการกำหนดคะแนนจุดตัดหรือระดับเกณฑ์ที่สมบูรณ์คือ 100% ย่อมเป็นไปได้ เพราะนักเรียนอาจมีความบกพร่องในคันต่าง ๆ เช่น มีความสะเพร่าในการคิดหรือเขียน

ตอบ ได้ข้อมูลที่ผิดพลาดหรือขาดความตั้งใจ สิ่งเหล่านี้ย่อมทำให้ระดับเกณฑ์คลาดเคลื่อนไปจาก
อุดมคติที่คาดหวัง ดังนั้นการกำหนดคะแนนจุดตัดจึงมักกำหนดให้ต่ำกว่า 100% ประมาณ 5 -
20% แผนบางครั้งอาจกำหนดให้ต่ำกว่าขึ้นอยู่กับความสำคัญของจุดประสงค์ที่กล่าวไว้

3. วิธีการเพิ่มคะแนนเกณฑ์อื่น ๆ (Boot Strapping Another Criterion Score)

วิธีนี้คะแนนจุดตัดของแบบทดสอบจะเป็นการกำหนดโดยอาศัยเกณฑ์ภายนอกมาระบุเพิ่ม
ขึ้นว่าสำเร็จ หรือรอบรู้อย่างชัดเจน เช่น อาจใช้เกณฑ์ว่าผู้ได้รับประกาศนียบัตรในอาชีพต่าง ๆ
เช่น ตั๊กแตน นายหน้า นักสถิติ เป็นต้น ถือเป็นผู้มีความสามารถ ส่วนผู้ที่ไม่ได้รับประกาศนียบัตร
ถือเป็นผู้ไม่มีความสามารถ พิจารณาการกระจายของคะแนนที่ได้จากการทำแบบทดสอบของคน
เหล่านี้แล้วกำหนดคะแนนจุดตัดเพื่อแยกผู้ที่มีความสามารถออกจากผู้ที่ไม่มีความสามารถ

เบอร์ก (Berk, 1976 อ้างถึงในสุรินทร์ แท่งจันทร์ 2528 หน้า 50 - 51) ได้นำ
วิธีนี้มาประยุกต์ใช้ ซึ่งมีลำดับขั้นตอนในการหาคะแนนจุดตัดดังนี้คือ

1. กำหนดเกณฑ์ภายนอกว่าใครได้รับการสอนหรือไม่ได้รับการสอนตามจุดประสงค์การ
เรียนรู้ที่ต้องการจะวัด
2. แบ่งนักเรียนออกเป็น 2 กลุ่ม ตามเกณฑ์ภายนอกที่กำหนดไว้
3. นำแบบทดสอบอิงเกณฑ์ไปสอบกับนักเรียนทั้ง 2 กลุ่ม
4. นำคะแนนที่ได้ไปเขียนกราฟร่วมกัน จุดที่กราฟหรือฟังก์ชันทั้ง 2 ตัดกัน จะกำหนด

ให้เป็นคะแนนพยากรณ์ (Predicted Score)

5. ใช้คะแนนพยากรณ์แบ่งนักเรียนเป็น 2 กลุ่มคือ กลุ่มที่ได้คะแนนสูงกว่าหรือเท่ากับ
คะแนนพยากรณ์เรียกว่า พวกรอบรู้ออกจากการพยากรณ์ (PM) และกลุ่มที่ได้คะแนนต่ำกว่าคะแนน
พยากรณ์เรียกว่าผู้ไม่รอบรู้ออกจากการพยากรณ์ (PN)

6. เขียนตาราง 2 x 2 แบ่งนักเรียนเป็น 4 กลุ่ม ดังนี้

เกณฑ์ภายนอก

ได้รับการศึกษา	ไม่ได้รับการศึกษา
รอบรู้อจริง (TM)	รอบรู้อไม่จริง (FM)
ไม่รอบรู้อไม่จริง (FN)	ไม่รอบรู้อจริง (TN)

รอบรู้อ
คะแนนพยากรณ์
ไม่รอบรู้อ

7. ใช้สูตร

$$P(TM) = TM / (M + N)$$

$$P(TN) = TN / (M + N)$$

$$P(FM) = FM / (M + N)$$

$$P(FN) = FN / (M + N)$$

เมื่อ M คือ จำนวนนักเรียนที่ไม่ได้รับการศึกษา

N คือ จำนวนนักเรียนที่ได้รับการศึกษา

8. หากคะแนนจุดตัดโดยการพิจารณาจากคะแนนพยากรณ์ทำให้ $P(TM) + P(TN)$ มีค่าสูงสุด หรือ $P(FM) + P(FN)$ มีค่าต่ำสุด คะแนนพยากรณ์นั้นจะถือเป็นคะแนนจุดตัดของแบบทดสอบ

อย่างไรก็ตามการกำหนดจุดตัดโดยวิธีนี้อาจมีข้อหา 2 ประการคือ

1. ถ้าแบบทดสอบที่เราต้องการกำหนดคะแนนจุดตัดมีค่าสหสัมพันธ์กับเกณฑ์ภายนอกต่ำ เราก็ไม่สามารถกำหนดคะแนนจุดตัดให้สอดคล้องกับเกณฑ์ภายนอกได้อย่างสมบูรณ์
2. ถ้าจะกำหนดคะแนนจุดตัดให้สอดคล้องกับเกณฑ์ภายนอก ก็จะมีปัญหาว่าเกณฑ์ภายนอกที่ระบุขึ้นมานั้นเลือกมาได้อย่างไร

4. การพิจารณาตัดสินจากความสามารถต่ำสุด (Judging Minimal Competency)

การกำหนดคะแนนจุดตัดด้วยวิธีนี้อาศัยแนวคิดของเกลเซอร์ (Glaser) ที่เสนอไว้ว่า ความสามารถจะต่อเนื่องกัน จากไม่มีเลยจนถึงปฏิบัติได้โดยสมบูรณ์ทั้งหมด และคะแนนจุดตัดจะแบ่งผู้ที่มีความสามารถและไม่มีความสามารถออกจากกัน ดังนั้นการหาคะแนนจุดตัดวิธีนี้ต้องให้ผู้เชี่ยวชาญเนื้อหาวิชาหรือครูประจำวิชาที่ศึกษาข้อสอบแล้วระบุว่าคะแนนใด จึงจะเป็นที่ยอมรับว่าผู้สอบผ่านเป็นผู้รอบรู้ สำหรับวิธีการนี้ผู้ศึกษาและได้เสนอเทคนิคในการหาคะแนนความสามารถต่ำสุดไว้หลายคนด้วยกัน เช่น อีเบล (Ebel) มีความเห็นว่าคะแนนจุดตัดที่กำหนดเป็นเปอร์เซ็นต์ของคะแนนสอบทั้งหมดอาจมีข้อบกพร่องเนื่องจากข้อสอบนั้นยากหรือง่ายเกินไป หรือค่าอ่านอาจจำแนกน้อยกว่าที่ผู้สร้างตั้งใจไว้ ในกรณีเช่นนี้การที่จะสอบผ่านหรือไม่อาจเนื่องมาจากตัวข้อสอบมากกว่าระดับความสามารถที่แท้จริงของผู้สอบเอง อีเบลจึงเสนอวิธีแก้ไขโดยการให้ผู้เชี่ยวชาญหรือครูประจำ

วิชาวิเคราะห์เนื้อหาของแบบทดสอบ ในแง่ลักษณะข้อสอบและระดับความยากของข้อสอบโดยเทียบกับเปอร์เซ็นต์การสอบผ่านที่คาดหวัง ซึ่งเขาได้ทำการจำแนกไว้แล้ว 4 กลุ่ม ดังตาราง

ตาราง 1 แสดงค่าเปอร์เซ็นต์การสอบผ่านที่คาดหวังจากลักษณะของข้อสอบและระดับความยากของข้อสอบ

ลักษณะข้อสอบ	ระดับความยาก (%)		
	ง่าย	ปานกลาง	ยาก
ตรงกับเนื้อหาและจำเป็นอย่างมาก	100	-	-
ตรงกับเนื้อหาและสำคัญ	90	70	-
ตรงกับเนื้อหา	80	60	40
ไม่แน่ใจว่าตรงกับเนื้อหา	70	50	30

วิธีการหาคะแนนจุดตัดตามแนวคิดของอ็อบมีลัวกับชั้นดังนี้ (บุญเชิด ภิญโญอนันตพงษ์ 2527 หน้า 125 - 127)

- นำข้อสอบที่ต้องการวิเคราะห์ไปให้กับครูประจำวิชาพิจารณาว่า ข้อสอบมีลักษณะตรงกับเนื้อหามากน้อยเพียงไร มีระดับความยากง่ายเพียงไร
- จำแนกข้อสอบตามลักษณะและความยาก โดยแจกแจงเป็นความดีรวม
- กำหนดเปอร์เซ็นต์การสอบผ่านตามลักษณะข้อสอบและความยากของข้อสอบ
- คูณความดีจำนวนข้อกับเปอร์เซ็นต์การสอบผ่านที่คาดหวังและรวมผลคูณทั้งหมด
- หาค่าเปอร์เซ็นต์การสอบผ่านที่คาดหวังโดยเฉลี่ย
- กำหนดคะแนนจุดตัดจากค่าเปอร์เซ็นต์การสอบผ่านที่คาดหวังโดยเฉลี่ยจากขั้นที่ 5

นีเทลสกี (Nedelisky) ได้เสนอเทคนิคในการหาคะแนนผ่านต่ำสุด (Minimum Passing Score) ซึ่งมีลำดับขั้นตอนในการหาดังนี้คือ

- นำแบบทดสอบไปให้ครูผู้สอนทุกคนพิจารณาว่าตัวเลือกใดในข้อสอบแต่ละข้อที่นักเรียน

ในระบัตที่ต่ำสุดที่จะผ่าน(พวกที่ไคเกรด D) นอกไคว่าผิด ซึ่งอาจมีมากกว่า 1 ตัวเลือก

2. หาส่วนกลับของจำนวนตัวเลือกที่เหลืออยู่ในแต่ละข้อ เช่น ข้อสอบ 5 ตัวเลือก ข้อหนึ่งมีตัวเลือก 2 ตัวที่นักเรียนไคเกรด D รู้ว่าผิดและจะไม่เลือก ดังนั้นส่วนกลับของจำนวนตัวเลือกในข้อนี้คือ $1/3$

3. หาผลรวมของส่วนกลับที่ไคจากข้อ 2 แล้วหาค่าเฉลี่ยของส่วนกลับของข้อสอบแต่ละข้อซึ่งจะถือเป็นความน่าจะเป็นในการเาคตอบถูกของข้อสอบแต่ละข้อ

4. หาผลรวมของความน่าจะเป็นในการเาคตอบถูก (M_{FD}) และส่วนเบี่ยงเบนมาตรฐานของความน่าจะเป็นในการเาคตอบถูก (S_{FD})

5. ให้ครูผู้สอนทุกคนพิจารณากำหนดค่า K โดยยึดหลักว่า ถ้าตัวเลือกผิดเกินซัคจะใช้ค่า 0.5 ถ้าตัวเลือกแตกต่างกันเล็กน้อยจะใช้ค่า 1.0 ถ้าตัวเลือกคล้ายคลึงกันมากจะใช้ค่า 1.5

6. หาคะแนนจุดคัค (D) จากสูตร

$$D = M_{FD} + K \cdot S_{FD}$$

แองกอฟ (Angoff) เสนอวิธีการกำหนดเกณฑ์ผ่านระดับต่ำสุดดังนี้คือ

1. ให้ผู้สอนกลุ่มหนึ่งพิจารณาแบบทดสอบโดยดูจากความยากและเนื้อหาของข้อสอบ
2. ผู้สอนกลุ่มเิมพิจารณาว่านักเรียนที่จะยอมรับไคว่าผ่านในระดับต่ำสุด (ไคเกรด D) มีโอกาสหรือความน่าจะเป็นที่จะทำข้อสอบแต่ละข้อถูกเป็นเท่าไร

3. หาผลรวมของความน่าจะเป็นที่จะทำข้อสอบถูก จากผู้สอนทุกคนและจากข้อสอบทุกข้อ

4. หาเปอร์เซ็นต์เฉลี่ยของความน่าจะเป็นแล้วนำมากำหนดเป็นคะแนนจุดคัคโดยเทียบกับจำนวนข้อสอบทั้งหมด

5. ใช้วิธีวิจัยเชิงปฏิบัติ (Operation Research Methods)

วิธีนี้เป็นการค้นหาคะแนนจุดคัคที่เหมาะสมจากผลการทดลองที่ตามมาโดยอาศัยไคถึงการแจกแจงทางคณิตศาสตร์หรือกราฟช่วยในการตัดสินใจความสัมพันธ์ของคะแนนจุดคัคจากแบบทดสอบกับเกณฑ์ภายนอกที่กำหนดขึ้น เช่น วิธีของบลอค (Block) ซึ่งมีลำดับขั้นในการหาคะแนนจุดคัคดังนี้คือ (บุญเชิด บุญไคอนันตพงษ์ 2527 หน้า 136)

1. สุ่มตัวอย่างนักเรียนที่ต้องการศึกษามาจำนวนหนึ่ง สุ่มแยกเป็นกลุ่มย่อย ๆ ประมาณ 5 - 10 กลุ่ม แต่ละกลุ่มต้องมีความสามารถหักเหเหมือนกัน

2. กำหนดคะแนนจุดตัดหรือระดับเกณฑ์ความรู้ให้มีระดับต่าง ๆ กัน เช่น 40%, 50%, 60% ..., 100%

3. ทำการสอนนักเรียนแต่ละกลุ่มจนกระทั่งนักเรียนบรรลุความรู้อย่างที่กำหนดโดยใช้แบบทดสอบอิงเกณฑ์ทำการทดสอบ

4. เลือกเกณฑ์ภายนอกโดยใช้คุณลักษณะที่เกิดจากการสอนอย่างใดอย่างหนึ่ง เช่น ความคงทนของการเรียนรู้ การถ่ายโอนการเรียนรู้ ความสำเร็จในการทำงาน เป็นต้น แล้วทำการวัดคุณลักษณะดังกล่าว

5. นำค่าวัดที่ได้จากแบบทดสอบอิงเกณฑ์ และค่าที่วัดได้จากเกณฑ์ภายนอกมาเขียนกราฟ

6. กำหนดคะแนนจุดตัดของแบบทดสอบอิงเกณฑ์ โดยพิจารณาจากคะแนนจุดตัดที่สัมพันธ์กับจุดโค้งกลับของกราฟที่สูงที่สุด

6. การกำหนดจุดตัดโดยใช้ทฤษฎีการตัดสินใจ (Decision-Theoretic Approaches)

เกลส ได้เสนอวิธีการกำหนดคะแนนจุดตัดโดยใช้ทฤษฎีการตัดสินใจ ซึ่งอาศัยเกณฑ์ภายนอกและคะแนนจากแบบทดสอบอิงเกณฑ์ที่จะทำให้ค่าฟังก์ชันของคะแนนจุดตัด $f(C_x)$ มีค่าน้อยที่สุด ซึ่งมีวิธีการหาดังนี้คือ

1. เลือกเกณฑ์ภายนอกที่สอดคล้องกับแบบทดสอบอิงเกณฑ์เพื่อจำแนกนักเรียนเป็น 2 กลุ่ม เช่น ได้รับการสอนและไม่ได้รับการสอน

2. นำแบบทดสอบอิงเกณฑ์ไปสอบกับกลุ่มตัวอย่างนักเรียนที่เลือกไว้ทั้งหมด

3. กำหนดคะแนนจุดตัด (C_x) ของแบบทดสอบอิงเกณฑ์หลาย ๆ ค่า

4. จำแนกนักเรียนเป็น 4 กลุ่ม ในแต่ละคะแนนจุดตัดซึ่งเกิดจากคะแนนจุดตัดกับเกณฑ์ภายนอก จำแนกรวมกัน

		เกณฑ์ภายนอก		
		ผ่าน	ไม่ผ่าน	
แบบทดสอบอิงเกณฑ์	ไม่ผ่าน	P_A	P_B	$1-P_C$
	ผ่าน	P_C	P_D	P_C
		P_E	$1-P_E$	1

5. หากค่าฟังก์ชันของคะแนนจุดตัด $f(C_X)$ แต่ละคะแนนจากสมการ

$$f(C_X) = (P_A + P_D)/(P_B + P_C) \dots\dots\dots (1)$$

หรือ

$$f(C_X) = (\alpha P_A + \beta P_D)/(P_B + P_C) \dots\dots\dots (2)$$

ขึ้นอยู่กับกำหนัดความสำคัญของการจำแนก ผลทั้ง 2 แบบ

6. กำหนดคะแนนจุดตัดที่เหมาะสม โดยดูจากผลการใช้คะแนนจุดตัดที่ทำให้ค่าฟังก์ชันของคะแนนจุดตัด $f(C_X)$ มีค่าน้อยที่สุดหรือค่าที่สุด

สวามินาธาน และคณะ (Swaminathan and others, 1975 อ้างถึงใน
ชวลิต โห้จันทร์ 2528 หน้า 19 - 20) ได้ศึกษาวิธีการกำหนดจุดตัดของเบส์ (Bayesian
Decision Procedure) ซึ่งเป็นวิธีการที่อาศัยคะแนนมวลความรู้ (Domain Score) เปรียบ
เทียบกับจุดตัดที่กำหนดขึ้นเพื่อบ่งบอกความรอบรู้ของผู้สอบ มีวิธีการโดยสรุปดังนี้

1. แปลงคะแนนสอบ (X_1) แต่ละคนให้เป็นอาร์คไซน์ (g_1)

$$g_1 = \sin^{-1} \sqrt{(x_1 + 3/8)/(n+3/4)}$$

เมื่อ n = จำนวนข้อสอบในแบบทดสอบ

2. กำหนดคะแนนเกณฑ์โดยประมาณ (π_0) แล้วเปลี่ยนให้เป็น v_0 จากสูตร

$$v_0 = \sin^{-1} \sqrt{\pi_0}$$

3. หากค่า v_1 จากคะแนนมวลความรู้ของผู้สอบแต่ละคน (π_1) จากสูตร

$$v_1 = \sin^{-1} \sqrt{\pi_1}$$

4. หากการแจกแจงครั้งแรกของ ρ โดยระบบค่าพารามิเตอร์ v และ λ เมื่อ v เป็นค่า degree of freedom เท่ากับ 8

$$\lambda = \bar{\rho}(v-2), \quad \bar{\rho} = (\Delta t+2)^{-1}$$

t แทน จำนวนข้อสอบที่เหมาะสมที่จะให้คะแนนวัดความรู้ของผู้สอบมากที่สุด

5. หาค่า p^* และ σ^2 จากตารางของแวง เมื่อทราบค่า v และ λ

6. หาค่าเฉลี่ยเลขคณิต (μ_i) และความแปรปรวน (σ_i^2) ของ v_i จากสูตร

$$\mu_i = \bar{g} + p^*(g_i - \bar{g})$$

$$\sigma_i^2 = \frac{1+(m-1)p^*}{(4n+2)m} + (g_i - \bar{g})^2 \sigma^{*2}$$

เมื่อ m แทน จำนวนนักเรียนที่เข้าสอบ

$$\bar{g} = \sum_{i=1}^m g_i / m$$

7. หาส่วนเบี่ยงเบนปกติมาตรฐาน (Z_{01}) ของคะแนน v_i โดย

$$Z_{01} = (v_i - \mu_i) \sigma_i^{-1}$$

8. หาค่า $\text{Prob}[Z < Z_{01} | \text{Data}]$ และ $\text{Prob}[Z \geq Z_{01} | \text{Data}]$ จากตารางพื้นที่ใต้โค้งปกติเมื่อ

$\text{Prob}[Z < Z_{01} | \text{Data}]$ แทน ความน่าจะเป็นในการสอบไม่ผ่าน

$\text{Prob}[Z \geq Z_{01} | \text{Data}]$ แทน ความน่าจะเป็นในการสอบผ่าน

9. หากความสูญเสียที่คาดหวังของการตัดสินใจให้ผ่านและไม่ผ่านของคะแนนจากผู้สอบดังนี้

$l_{12} \text{Prob}[Z < Z_{01} | \text{Data}]$ แทน ความสูญเสียที่คาดหวังของการตัดสินใจให้ผ่าน

$l_{21} \text{Prob}[Z \geq Z_{01} | \text{Data}]$ แทน ความสูญเสียที่คาดหวังของการตัดสินใจให้ไม่ผ่าน

เมื่อ l_{12} แทน ความสูญเสียที่เกิดจากความคลาดเคลื่อนในการยอมรับสิ่งที่ผิด

l_{21} แทน ความสูญเสียที่เกิดจากความคลาดเคลื่อนในการปฏิเสธสิ่งที่ผิด

10. เลือกคะแนนที่น้อยที่สุดที่ทำให้ค่า $l_{21} \text{Prob} [Z \geq Z_{0i} | \text{Data}]$ มากกว่า
ค่า $l_{12} \text{Prob} [Z < Z_{0i} | \text{Data}]$ เป็นจุดตัด

ฮวน (Huynh, อ้างถึงใน รังสรรค์ รมย์เล็ก 2527 หน้า 30) ได้เสนอแนะวิธีการ
กำหนดจุดตัดแบบมินิมแมกซ์ (Minimax) ซึ่งวิธีการกำหนดจุดตัดวิธีนี้มีกระบวนการคล้าย ๆ กับ
วิธีการกำหนดจุดตัดของเบส์ (Bayes) แต่วิธีมินิมแมกซ์นี้จะกำหนดจุดตัดเป็นรายบุคคล และไม่
ต้องการข้อสันนิษฐานแรก (Prior Information) เกี่ยวกับตัวผู้สอบในการหาคำนวณ
จุดตัด

นอกจากวิธีทั้งหมดที่กล่าวมาแล้ว ค่อมภายหลังได้มีนักวิจัยบางท่านนำเอาแนวคิดเกี่ยวกับ
กับทฤษฎีที่เกิดขึ้นใหม่ คือ ทฤษฎีคุณลักษณะแฝงเข้ามาประยุกต์ใช้ในการกำหนดจุดตัดของแบบทดสอบ
อิงเกณฑ์ ซึ่งจะกล่าวถึงต่อไป

แวน เดอ ลินเดน (Van Der Linden, อ้างจาก ชวลิต โพธิ์นคร 2528 หน้า 26-
28) ใช้ทฤษฎีคุณลักษณะแฝง (Latent Trait Theory) ช่วยในการกำหนดจุดตัดตามวิธี
ของนีเคลสกี และ แองกอฟ ทั้งมีวิธีการโดยสรุปดังนี้

1. เลือกโมเดลของทฤษฎีคุณลักษณะแฝงเพื่อคำนวณหาค่าพารามิเตอร์ของข้อสอบ และ
ทดสอบความเหมาะสม (Fit) ของข้อสอบกับโมเดล

2. กำหนดค่าความน่าจะเป็นของความสำเร็จของข้อสอบแต่ละข้อ ($P_i(S)$) ตามวิธี
ของนีเคลสกี และ แองกอฟ

3. คำนวณค่าคะแนนรวมความรู้ (π_0)

3.1 วิธีแองกอฟ

$$\sum_{i=1}^n P_i (+|\theta_0) \quad (1)$$

เมื่อ $P_i (+|\theta)$ แทน ความน่าจะเป็นที่นักเรียนตอบข้อสอบ i ถูก
ที่ระดับความรู้ θ

θ แทน ระดับความรู้ของนักเรียนในระดับความ

$$E(U_i | \theta) = 1 \cdot P_i(+|\theta) + 0 \cdot [1 - P_i(+|\theta)] \quad (2)$$

เมื่อ U_i แทน คะแนนการตอบข้อสอบแต่ละข้อซึ่งมีคะแนนแบบ (0,1)
 E แทน ค่าคะแนนที่คาดหวังของนักเรียนที่มีระดับความรู้ θ

$$\sum_{i=1}^n P_i(+|\theta_c) = \sum_{i=1}^n E(U_i | \theta_c) = E\left(\sum_{i=1}^n U_i | \theta_c\right) = E(X | \theta_c) = \pi_c \quad (3)$$

3.2 วิธีเคลสกี

$$P_i(+|\theta_c) = [g_i - k_i^{(c)}]^{-1} \quad (4)$$

$$\pi_c \equiv \sum_{i=1}^n [g_i - k_i^{(c)}]^{-1} \quad (5)$$

เมื่อ g_i คือ จำนวนในการเลือกข้อสอบ i
 $k_i^{(c)}$ คือ จำนวนในการเลือกที่ใช้ตัดสินว่านักเรียนที่มีความรู้ในระดับ
 ความรู้ θ_c ควรจะเลือกข้อสอบนั้นผิด

4. เปลี่ยนค่า π_c เป็น θ_c จากสมการ 3 หรือ 5

5. หาค่า $\hat{P}_i$ จากสมการ

$$P_i(+|\theta) = c_i + (1 - c_i) \{1 + \exp[-a_i(\theta - b_i)]\}^{-1} \quad (6)$$

เมื่อ c_i (การคาดเดา), a_i (การอ่านใจจำแนก) และ b_i (ค่าความยากของข้อสอบ) เป็นค่าพารามิเตอร์ของข้อสอบแต่ละข้อที่คำนวณจากข้อ 1

6. คำนวณค่าเฉลี่ยความคลาดเคลื่อนสัมบูรณ์ของส่วนเบี่ยงเบนเฉพาะจากค่าความเป็นไปได้อย่างสูงในการวัดบนช่วงสเกลแบบ 0,1 (C) เมื่อ

$$c_1 \equiv \frac{M-E}{M}$$

$$M \equiv \sum_{i=1}^n e_i^{(u)}/n$$

$$e_i \equiv p_i^{(s)} - p_i$$

$$e_i^{(u)} \equiv \max \left\{ p_i, 1 - p_i \right\}$$

$$E \equiv \sum_{i=1}^n |p_i^{(s)} - p_i| / n \equiv \sum_{i=1}^n |e_i| / n$$

เมื่อ $e_i^{(u)}$ เป็นค่าสัมบูรณ์สูงสุดของความคลาดเคลื่อนของค่าเฉพาะเมื่อ

$$p_i^{(s)} = 0 \text{ หรือ } 1$$

E เป็นค่าเฉลี่ยความคลาดเคลื่อนสัมบูรณ์ของค่าเฉพาะ
ในกรณีที่กำหนดโดยวิธีของนีเคลสกี ต้องหาค่า λ ซึ่งหาได้จาก

$$\lambda = c_2 - c_1$$

$$c_2 = \frac{M-E}{M-m}$$

$$m \equiv \sum_{i=1}^n e_i^{(1)} / n$$

$$e_i^{(1)} \equiv \left| [g_i - k^*]^{-1} - p_i \right|$$

เมื่อ $e_i^{(1)}$ เป็นค่าสัมบูรณ์ต่ำสุดของความคลาดเคลื่อนของค่าเฉพาะเมื่อ

$$|p_i^{(s)} - p_i| \text{ มากกว่า } 0$$

λ

เป็นค่าแสดงการลดลงของความถี่

7. หาค่าเฉลี่ยเลขคณิตของ E , c_1 , c_2 และ λ จากผู้ตัดสินทั้งหมด

วิลคอกซ์ (Wilcoxon) ได้กำหนดจุดตัดโดยอาศัยคุณสมบัติลักษณะแบ่ง ซึ่งมีขั้นตอนทั้ง

ต่อไปนี้ (Hambleton and Grujter, 1983 : 357-361)

1. เตรียมข้อสอบไว้มาก ๆ ข้อ ข้อสอบเหล่านี้ต้องวัดในโดเมนเดียวกัน

2. หาค่าพารามิเตอร์ของข้อสอบโดยใช้คุณสมบัติลักษณะแบ่งจากผู้เข้าสอบที่มีจำนวนมาก

3. ตรวจสอบดูว่าข้อมูลที่ไ้มีความสอดคล้องกับไ้งลักษณะของข้อสอบตามทฤษฎีหรือไม่
4. เลือกคะแนนจุดตัด (π_0) และช่วงที่มีคะแนนโกลเมที่มีค่าเกือบองที่ (π_e, π_u)
5. แปลงค่า π_e, π_0 และ π_u ให้เป็น θ_e, θ_0 และ θ_u ตามลำดับ โดยใช้สมการ

$$\pi = \frac{1}{m} \sum_{i=1}^m P_i(\theta)$$

เมื่อ m คือ จำนวนข้อของข้อสอบ

6. กำหนดค่าความน่าจะเป็นในการจำแนกคนผิดพลาด (P^*)
7. รวบรวมข้อสอบที่จะให้ค่าใกล้เคียงกับ θ_0 แล้วสุ่มออกมาคำนวณ โดยใช้สมการ

$$\theta_i^* = b_i + \frac{1}{Da_i} \log_1 .5 (1 + \sqrt{1 + 8 C_i})$$

8. เปลี่ยนค่า θ_e, θ_0 และ θ_u ไปเป็น π_e^*, π_0^* และ π_u^* ตามลำดับโดยใช้ไ้งลักษณะของข้อสอบ (Item Characteristic Curve) ที่มีข้อสอบ k ข้อ คำนวณค่า P_e, P_u และ P_m สำหรับ P_m คือค่าสูงสุดของ P_e หรือ P_u ก็ได้ และเปลี่ยนจุดตัดไปเรื่อย ๆ ให้ใกล้เคียงกับ $\sum_{i=1}^k P_i(\theta)$ ค่านี้จะเป็นจุดตัดที่ควรเป็น

ส่ง ลักษณะ (2524, อ้างถึงใน วุฒิชัย เสาวราต 2526 หน้า 37 - 39) เสนอวิธีการใช้ไ้งลักษณะของข้อสอบโดยใช้โมเดลโลจิสติกแบบพารามิเตอร์ 3 ตัว ของเบิร์นบอร์ม (Birnbaurm) กำหนดจุดตัดโดยการพิจารณาความสามารถของคน (θ) เปรียบเทียบกับความยากของข้อสอบ (b_g)

ชวลิต โพธิ์นคร (2528 หน้า 63 - 65) ใช้วิธีประยุกต์ราซคโมเดลในการกำหนดจุดตัดโดยใช้ค่าความยากเฉลี่ยของข้อสอบทั้งฉบับไปเปรียบเทียบกับค่าความสามารถของคน (θ) โดยถือว่าค่าความยากเฉลี่ยของข้อสอบ ($\bar{b}$) เท่ากับค่าความสามารถของคนตรงกับคะแนนรวมใด ให้ใช้คะแนนนั้นเป็นจุดตัด

จากที่กล่าวมาแล้วทั้งหมดในเรื่องของคะแนนจุดตัดนั้น มิลแมน (Millman) ได้ให้ข้อสังเกตว่าควรพิจารณาถึงสิ่งต่าง ๆ ประกอบการกำหนดจุดตัดดังต่อไปนี้ (อังกฤษ สายยศ 2525 หน้า 72 - 73)

1. คะแนนของผู้สอบคนอื่น ๆ (Performance of other) คือการกำหนดคะแนนจุดตัด ควรคำนึงถึงเปอร์เซ็นต์ของนักเรียนที่จะสอบผ่านก่อน แล้วเลือกจากผู้ที่มีความสามารถสูงที่สุดลงมา นอกจากนี้อาจคำนึงจากระดับคะแนนของผู้ที่ได้รับการรับรองแล้วว่ามีความรู้ในเนื้อหาจริงเป็นจุดตัด

2. เนื้อหาในแต่ละข้อคำถาม (Item Content) ควรให้ผู้เชี่ยวชาญพิจารณาเพื่อตัดสินว่ามีความสำคัญที่จะใช้ในการเรียนเนื้อหาใหม่หรือไม่ เช่น ในการกำหนดคะแนนจุดตัดจะต้องดูข้อคำถามแต่ละข้อที่ถาม เนื้อหาที่มีความสำคัญ และความยากง่ายเพียงใดถึงข้อใด (Ebel) พิจารณาจากความสำคัญและความยากง่ายของเนื้อหา แล้วเสนอสัดส่วนของข้อสอบในแต่ละด้านเป็นจำนวนที่คาดว่าผู้สอบมีความสามารถขั้นต่ำสุดที่ควรจะทำได้

3. ผลการศึกษาที่ผ่านมา (Education Consequence) ถ้าตั้งเกณฑ์ค่าเกินไป นักเรียนที่ผ่านไปเรียนบทเรียนต่อไป อาจจะเรียนได้อย่างไม่มีประสิทธิภาพ ถ้าตั้งเกณฑ์สูงเกินไปก็จะทำให้นักเรียนเสียเวลาในการเรียนซ่อมเสริมโดยไม่จำเป็น ทำให้เกิดผลเสียทางจิตวิทยาและทางกาย เศรษฐกิจ ดังนั้น สิ่งที่ต้องพิจารณาคือ ควรตั้งคะแนนจุดตัดเท่าใด จึงจะให้เกิดประโยชน์ทางการศึกษามากที่สุด โดยพิจารณาว่าถ้าความรู้และทักษะที่วัดเป็นพื้นฐานของสิ่งที่เรียนต่อไปควรตั้งเกณฑ์ให้สูง แต่ถาเนื้อหาไม่ใช่สิ่งที่จะต้องใช้ในการเรียนต่อไป คะแนนจุดตัดควรต่ำลง

4. ความคลาดเคลื่อนจากการเดาและการสุ่มข้อสอบ (Error Due to Guessing and Item Sampling) การกำหนดเกณฑ์สูงจะช่วยให้การเดาลดลง และการใช้สูตรปรับแก้การเดามาตรฐาน แล้วนำมาเปรียบเทียบกับคะแนนจุดตัดที่กำหนด จะช่วยควบคุมการเดาได้บางส่วน ส่วนความคลาดเคลื่อนจากการสุ่มข้อสอบที่ไม่แน่ใจว่าเนื้อหาในแบบทดสอบนั้นเป็นตัวแทนที่ศึกษาหรือไม่ ควรมีการปรับคะแนนจุดตัดให้สูงขึ้นเพื่อป้องกันความคลาดเคลื่อนที่เกิดจากการจัดประเภทกลุ่ม ผู้รอบรู้และไม่รอบรู้ผิดไป

แฮมเบิลตัน (Hambleton, 1978 : 288-289) ได้ให้ข้อเสนอแนะเกี่ยวกับการกำหนดคะแนนจุดตัดซึ่งได้มาจากประสบการณ์ในการทำงานในหลาย ๆ โรงเรียนดังนี้

1. การพิจารณาคะแนนจุดตัดควรทำโดยกลุ่มที่ทำงานร่วมกันหลาย ๆ กลุ่ม กลุ่มเหล่านี้ ได้แก่ ครู ผู้ปกครอง ผู้เชี่ยวชาญด้านหลักสูตร ผู้บริหารโรงเรียน และนักเรียน (ในกรณีที่เป็นักเรียน

เรียนในระดับมัธยมศึกษาตอนปลาย) จำนวนคนในแต่ละกลุ่มขึ้นอยู่กับความสำคัญของแบบทดสอบ ที่นำมาพิจารณา และจำนวนของตารางวิเคราะห์โคเมนของข้อสอบ โดยอย่างน้อยที่สุดควรมีจำนวน คนเพียงพอที่จะแบ่งผู้พิจารณาเป็น 2 กลุ่ม เพื่อเปรียบเทียบความคงที่ของการพิจารณาตัดสินจาก บุคคล 2 กลุ่มนี้

2. ผู้ทำการตัดสินควรนำตารางวิเคราะห์โคเมนของข้อสอบมาเสนอและอภิปรายรวม กันด้วย

3. การใช้วิธีของอ็อบล หรือ นีเคสส์ ควรทำการทดลองกับกลุ่มตัวอย่างหลายพวก เพื่อตัดสินและแก้ปัญหาเกี่ยวกับความแตกต่างระหว่างกลุ่ม

4. พยายามจัดเวลาให้เหมาะสมกับการวิเคราะห์โคเมนข้อสอบแต่ละตาราง หาก ตารางวิเคราะห์โคเมนใดมีความซับซ้อนหรือมีความสำคัญมาก ควรให้เวลามากขึ้น

5. พยายามทำให้เกิดความมั่นใจว่าผู้ตัดสินมีความรู้เกี่ยวกับการใช้แบบทดสอบ และรู้ ว่าจะใช้กับนักเรียนกลุ่มใด

6. หากมีความสัมพันธ์ระหว่างตารางวิเคราะห์โคเมน ควรนำมาพิจารณาด้วย เช่น ถ้าตารางวิเคราะห์โคเมนใดเป็นสิ่งที่ต้องการหรือจำเป็นมากกว่าควรกำหนดคะแนนจุดตัดที่สูงกว่า

7. เมื่อใดก็ตามที่เป็นไปก็พยายามใช้บุคคล 2 กลุ่ม หรือมากกว่าในการตัดสินคะแนน จุดตัด ควรศึกษาความคงที่ในการกำหนดคะแนนและความแตกต่างในการกำหนดน้ำหนักของผู้ตัดสิน และทำการอภิปรายเพื่อหาการตัดสินที่สอดคล้องกัน

8. ถ้าข้อมูลเกี่ยวกับการสอบครั้งก่อน ๆ มีประโยชน์ก็สามารถนำมาใช้เพื่อพิจารณา แกไขเกี่ยวกับการกำหนดคะแนนจุดตัดและเวลาที่ใช้สอนได้

9. หากข้อมูลที่ไต่จากการสอบเหมาะที่จะนำมาใช้พิจารณาเพื่อเป็นดัชนีของผู้รอบรู้และ ผู้ไม่รอบรู้ในแต่ละวัตถุประสงค์ ก็ควรนำมาศึกษา หากคะแนนที่ไต่จากการทำแบบทดสอบมีที่หาว่า จะเป็นไปในลักษณะผิดปกติ (Out of line) เราก็สามารถอธิบายสาเหตุได้โดยการพิจารณา ข้อสอบของแบบทดสอบ (ซึ่งอาจจะไม่มีความเที่ยงตรงก็ได้) ระดับของคะแนนจุดตัด ความแปรผัน ของคะแนน สอบระหว่างชั้นเรียนและปริมาณเวลาที่ใช้สอน

10. เปรียบเทียบสภาพการรอบรู้ของผู้สอบที่ยังไม่ได้เรียนและที่เรียนแล้ว โดยกลุ่มที่

เรียนแล้วควรมีจำนวนผู้รอบรู้มาก กลุ่มที่ยังไม่ไ้เรียนรู็ควรมีจำนวนผู้ไม่รอบรู้มาก หากมีนักเรียนจำนวนมากได้รับการจัดประเภทผิดต้องพิจารณาคะแนนจุดตัดที่มีความเที่ยงตรงมากกว่าคะแนนจุดตัดที่ใช้

11. หากหลักสูตรมีการเปลี่ยนแปลงซึ่งจะมีผลต่อคะแนนจุดตัด จำเป็นต้องมีการตรวจสอบคะแนนจุดตัดใหม่ให้เหมาะสม

จากข้อเสนอแนะเกี่ยวกับการกำหนดจุดตัดที่กล่าวมาแล้วนั้น พอจะสรุปได้ว่าการกำหนดคะแนนจุดตัดจะต้องคำนึงถึงสิ่งสำคัญ 2 ประการต่อไปนี้

1. การกำหนดคะแนนจุดตัดไว้ล่วงหน้าตายตัว เป็นวิธีที่ไม่เหมาะสม
2. การเลือกคะแนนจุดตัดสูง ๆ ไม่มีส่วนทำให้การเรียนสูงขึ้นเสมอไป

ทฤษฎีคุณลักษณะแฝง

การวัดผลทางการศึกษาและจิตวิทยานั้น สิ่งที่เราต้องการจะวัดตรวจสอบมักเป็นคุณลักษณะ (Traits) ต่าง ๆ ที่เป็นคุณสมบัติภายในที่แอบแฝง (Latent Traits) และไม่สามารถสังเกตได้โดยตรง งานของนักวัดผลที่พยายามสร้างทฤษฎีเป็นเวลานานก็คือการหาทางอธิบายหรือทำนายคุณลักษณะภายในของบุคคลโดยอาศัยการตอบข้อสอบในแบบทดสอบที่กำหนดให้ ทฤษฎีที่เราต้องการจะต้องสามารถอธิบายหรือทำนายความสามารถแท้ (True Ability หรือ True Score) ของบุคคลได้อย่างแม่นยำ เชื่อถือได้ ทฤษฎีการวัดผลที่เคยใช้กันระยะหนึ่งที่ผ่านมาแล้วจนถึงปัจจุบันถูกมองว่าเป็นทฤษฎีดั้งเดิม (Classical Theory) ที่ไม่เพียงพอต่อการประมาณค่าความสามารถแท้ และไม่สามารถแก้ปัญหาคือหลายประการของการวัดผลในปัจจุบัน ดังต่อไปนี้

1. การตัดสินความสามารถของผู้สอบจากคะแนนที่ได้ (Observe Score) เป็นการตัดสินบนพื้นฐานของความคลาดเคลื่อน (Error) ยังไม่สามารถประมาณค่าความสามารถแท้หรือคะแนนจริงที่ชัดเจน
2. ผู้สอบคนเดิม ข้อสอบชุดขนาน 2 ชุด ที่วัดความสามารถเดียวกัน ผลออกมาคือคะแนนที่ต่างกันทั้ง ๆ ที่สอบในเวลาใกล้เคียงกัน
3. ค่าสถิติของข้อสอบแต่ละข้อ เช่น ค่าความยากง่าย (P) ค่าอำนาจจำแนก (r)

เปลี่ยนไปตามกลุ่มของผู้เข้าสอบ ทำให้เกิดปัญหาในการวิเคราะห์ข้อสอบฉบับเดิม แต่เปลี่ยนกลุ่มผู้สอบ

4. มีปัญหาในการประมาณค่าคะแนนจริงตามสูตรของทฤษฎีทั้งเดิม เพราะค่าความคลาดเคลื่อนของการวัด (*Error of Measurement*) เป็นค่าที่ผันแปรไปตามฉบับของแบบทดสอบและแปรตามกลุ่มเฉพาะของผู้เข้าสอบ ผู้เข้าสอบ 1 คน ถ้าใช้ข้อสอบที่วัดสิ่งเดียวกันต่างฉบับหรือไปสอบกับกลุ่มเก่ง กลุ่มอ่อนต่างกันไป คะแนนจริงของผู้เข้าสอบคนนี้จะผันแปรไปด้วย

5. ในการทดลองข้อสอบมาตรฐาน (*Try out*) ตามทฤษฎีแบบดั้งเดิมจำเป็นต้องใช้ข้อสอบจำนวนมากและกลุ่มตัวอย่างขนาดใหญ่จึงทำให้เกิดความไม่สะดวกในการทำธนาคารข้อสอบ (*Item banking*)

6. การเทียบความสามารถแท้ (*Equating*) ระหว่างการสอบข้อสอบชุดต่าง ๆ ที่วัดสิ่งเดียวกัน แต่ต่างเวลากันยังไม่สามารถทำได้อย่างชัดเจน

7. การจัดข้อสอบที่สอดคล้องตรงกับระดับความสามารถของผู้สอบ (*Tailored Testing*) และหาคะแนนความสามารถแท้ของผู้สอบได้ตลอดเวลาที่กำลังรับการทดสอบตามโปรแกรมเรียน รายบุคคล (*Individualized Instruction*) นั้นยังไม่อาจทำได้ชัดเจนโดยทฤษฎีทั้งเดิม

8. การตัดสินสภาพการรู้รอบ (*Mastery*) ในหน่วยการเรียนรู้ใด ๆ นั้น ปัจจุบันมีข้อคลาดเคลื่อนยากต่อการตัดสิน

ปัญหาหลัก ๆ ทั้ง 8 ประการ เป็นการท้าทายนักวัดผลให้ค้นหาทฤษฎีใหม่ ๆ ซึ่งทฤษฎีคุณลักษณะแฝง (*Latent Trait Theory*) หรือบางที่เรียกว่าทฤษฎีการตอบสนองข้อสอบ (*Item Response Theory : IRT*) สามารถแก้ปัญหาดังกล่าวได้เป็นที่น่าพอใจ

ทฤษฎีคุณลักษณะแฝง เป็นทฤษฎีที่ใช้โมเดลทางคณิตศาสตร์แสดงความสัมพันธ์ระหว่างความสามารถแท้กับการตอบสนองข้อสอบมีความสำคัญดังนี้คือ ผลการสอบของผู้สอบจากแบบทดสอบใด ๆ สามารถพยากรณ์หรืออธิบายได้จากความสามารถของคน ๆ นั้น (สำเริง บุญเรืองรัตน์ 2529 หน้า 42 - 43)

ข้อตกลงเบื้องต้นของทฤษฎีคุณลักษณะแฝง

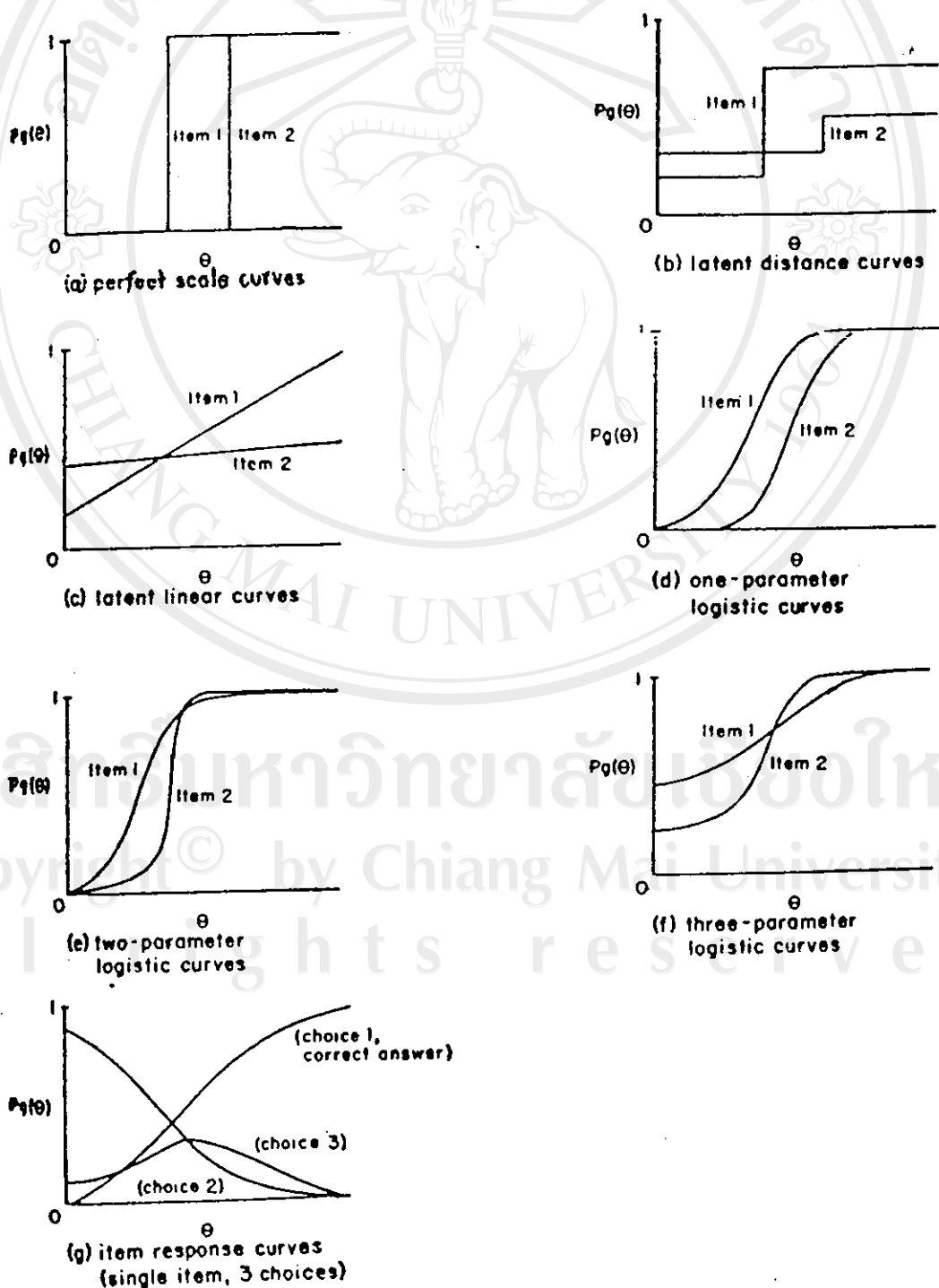
สำเร็จ บุญเรืองรัตน์ (2525 หน้า 3 - 7) ได้กล่าวถึงข้อตกลงเบื้องต้นของทฤษฎีไว้ดังนี้

1. ข้อตกลงเกี่ยวกับมิติที่จะวัด (Dimensional) กล่าวคือ โดยทั่วไปแล้วการสร้างแบบทดสอบมักมีข้อตกลงเกี่ยวกับมิติเดียว (Unidimensional) เพราะจะทำให้ผู้สร้างแบบทดสอบรวมคำถามที่มุ่งวัดความสามารถอย่างใดอย่างหนึ่งในฉบับเดียวกัน รวมคะแนนผลการตอบใ้กายและแปลความหมายของคะแนนได้สะดวกขึ้น แบบทดสอบที่มีลักษณะเป็นมิติเดียว (Unidimensional) จะมีลักษณะดังนี้ สมมติว่าแบบทดสอบฉบับหนึ่งมี n ข้อ และใช้ผู้สอบ x กลุ่ม และสมมติอีกว่าแบบทดสอบวัดกลุ่มบุคคลที่มีความสามารถเท่ากับ θ ด้านคะแนนจากการสอบทั้ง x กลุ่มนั้นมาแจกแจงความถี่แล้ว คะแนนมีลักษณะการแจกแจงเป็นอย่างเดียวกัน ก็ถือว่าแบบทดสอบนั้นมีลักษณะเป็นมิติเดียว แต่การกระจายของคะแนนมีลักษณะแตกต่างกันไป ก็แสดงว่าแบบทดสอบนั้นต่างวัดสิ่งอื่น ๆ อีกด้วย ไม่ได้วัดความสามารถเดียวเท่านั้น เทคนิคที่จะตรวจสอบข้อตกลงว่าแบบทดสอบนั้นมีลักษณะความเป็นมิติเดียวหรือไม่ ทำได้โดยการวิเคราะห์องค์ประกอบพบว่าเมื่อองค์ประกอบสำคัญอยู่มากกว่าหนึ่งองค์ประกอบ ก็ถือว่าแบบทดสอบนั้นไม่เป็นมิติเดียว ดังเช่น ในรายงานของ ทองหล่อ วิภาวิน (2522) ลอร์ด และ โนวิค (Lord and Novick, 1968) เสนอให้ใช้สหสัมพันธ์แบบเตตระคอรริก (Tetrachoric) ระหว่างข้อสอบข้อต่าง ๆ ในการทำการวิเคราะห์องค์ประกอบ

2. ข้อตกลงเกี่ยวกับความเป็นอิสระต่อกัน (Local Independence) ข้อตกลงนี้กล่าวว่าโอกาสผู้ตอบคำถามข้อหนึ่งนั้นไม่เกี่ยวกับการตอบคำถามข้ออื่น ๆ ในแบบทดสอบนั้น การทดสอบว่าข้อสอบแต่ละข้อเป็นไปตามข้อตกลงของความเป็นอิสระต่อกันหรือไม่นั้น ลอร์ด (Lord) ได้เสนอแนะว่าให้ตรวจสอบรูปแบบแผนการตอบคำถามของผู้สอบที่มีความสามารถปาน ๆ กัน ว่าข้อสอบแต่ละข้อ เช่น ข้อ 1 และข้อ 2 ข้อ 2 และข้อ 3 นั้น ลักษณะการตอบแตกต่างกันหรือไม่ ซึ่งสามารถตรวจสอบได้จากการคำนวณค่าไคสแควร์เพื่อทดสอบความเป็นอิสระของการตอบคำถามเหล่านั้นได้

3. ข้อตกลงเกี่ยวกับลักษณะของโค้งการตอบข้อสอบ (Item Characteristic Curves: ICC) ลักษณะของโค้งการตอบข้อสอบจะไม่ขึ้นอยู่กับการกระจายความสามารถของกลุ่มผู้เข้าสอบ และเป็นโค้งที่มีลักษณะเพิ่มขึ้นทางเดียว (Monotonic increasing) ตามระดับความสามารถที่เพิ่มขึ้น จึงเห็นได้ว่าลักษณะของข้อคำถามและคุณสมบัติของคำถามนี้ไม่ขึ้นอยู่กับลักษณะของกลุ่มผู้สอบ จึงเป็นข้อคำถามที่มีคุณสมบัติคงที่ตลอดเวลา

ภาพ 1 แสดงโค้งลักษณะของข้อสอบ (ICC) ของโมเดลต่าง ๆ ในทฤษฎีคุณลักษณะแฝง หรือทฤษฎีการตอบสนองของข้อสอบ (IRT) (Hambleton and Cook, อ้างถึงใน ปณิศา วิษัยณะ 2528 หน้า 16 - 17)



รูป (a) แสดงถึงโอกาสที่จะตอบข้อสอบแต่ละข้อได้ถูกเป็น 0 หรือ 1 มีลักษณะเป็นฟังก์ชันแบบสเตป (Step Function) ใช้กับมาตราของกัทต์แมน (Guttman's scale) จากรูปแสดงว่าข้อสอบข้อที่ 2 ยากกว่าข้อที่ 1

รูป (b) แสดงถึงโอกาสที่จะตอบข้อสอบแต่ละข้อได้ถูกไม่เป็น 0 หรือ 1 มีลักษณะเป็นฟังก์ชันแบบสเตป (Step Function) เช่นเดียวกับรูป (a) จากรูปแสดงว่าข้อสอบข้อที่ 2 ยากกว่าข้อที่ 1

รูป (c) แสดงถึงความสัมพันธ์ของโอกาสที่จะทำข้อสอบได้ถูกกับระดับความสามารถ มีลักษณะเป็นเส้นตรง จากรูปแสดงว่าข้อสอบข้อที่ 1 จำแนกคนได้ดีกว่าข้อที่ 2

รูป (d) แสดงถึงโอกาสที่จะทำข้อสอบได้ถูกขึ้นอยู่กับความยากของข้อสอบ (b) เพียงอย่างเดียว และทุกข้อมีอำนาจจำแนก (a) เท่ากับ 1 จากรูปแสดงว่าข้อสอบข้อที่ 2 ยากกว่าข้อที่ 1

รูป (e) แสดงถึงโอกาสที่จะทำข้อสอบได้ถูกขึ้นอยู่กับค่าความยาก (b) และค่าอำนาจจำแนก (a) ของข้อสอบ จากรูปแสดงว่าข้อสอบข้อที่ 2 จำแนกคนได้ดีกว่าข้อที่ 1

รูป (f) แสดงถึงโอกาสที่จะทำข้อสอบได้ถูกขึ้นอยู่กับค่าความยาก (b) ค่าอำนาจจำแนก (a) และการเกาของข้อสอบ (c) จากรูปแสดงว่าข้อสอบข้อที่ 2 มีค่าการเกาน้อยกว่าข้อที่ 1 และมีค่าอำนาจจำแนกสูงกว่าข้อที่ 1

รูป (g) แสดงถึงการเปรียบเทียบโครงสร้างลักษณะของข้อสอบข้อเดียวกันที่มี 3 ตัวเลือก จากรูปแสดงว่าตัวเลือกที่ 3 เป็นตัวเลือกที่ไม่ดี

สงวนลิขสิทธิ์ (2525 หน้า 91) กล่าวว่า ถ้าข้อตกลงเบื้องต้นของการสอบตรงตามทฤษฎีที่ต้องการ และเราสามารถคำนวณค่าพารามิเตอร์ได้เหมาะสมกับโมเดลแล้วผลที่ได้ก็คือ

1. ค่าพารามิเตอร์ของข้อสอบ คือ ค่าความยาก ค่าอำนาจจำแนก และค่าสัมประสิทธิ์การเกา จะเป็นค่าที่ไม่แปรเปลี่ยนไปตามกลุ่มผู้สอบไม่ว่าจะนำไปสอบกับผู้ที่ใดก็ตาม
2. เมื่อทราบลักษณะการตอบข้อสอบแต่ละข้อของผู้สอบคนใด เราจะสามารถคำนวณหาความสามารถที่แท้จริงของบุคคลนั้นได้

โกวิท ประวาลพุกษ์ (2525 หน้า 63) ได้กล่าวถึงค่าพารามิเตอร์ของข้อสอบที่ได้

จากการวิเคราะห์ข้อสอบตามทฤษฎีคุณลักษณะแฝงว่า มีคุณสมบัติ 3 ประการคือ

1. ค่าพารามิเตอร์ไม่ขึ้นอยู่กับกลุ่มตัวอย่าง
2. สามารถเปรียบเทียบความยากง่ายกันได้โดยอิสระ
3. เป็นค่าสถิติที่เพียงพอสำหรับการใช้อธิบายคุณภาพข้อทดสอบ ไม่ต้องอาศัยค่าอื่น ๆ

อีกเพิ่มเติม

ค่าพารามิเตอร์ในทฤษฎีคุณลักษณะแฝง แบ่งออกเป็นค่าพารามิเตอร์ของข้อสอบ (Item Parameter) ได้แก่ ความยาก (b_g) ค่าอำนาจจำแนก (a_g) ค่าสัมประสิทธิ์การเกาะ (c_g) และค่าพารามิเตอร์ของผู้สอบ (Examinee Parameter) ได้แก่ ระดับความสามารถของผู้สอบ (θ) พิสัยของค่าพารามิเตอร์ต่าง ๆ เหล่านี้มีดังนี้ (Hambleton and Cook, อ้างถึงในจิราภรณ์ ไกรสรศิริเวช 2528 หน้า 13)

b_g มีค่าตั้งแต่ $-\infty$ ถึง $+\infty$ แต่ในทางปฏิบัติส่วนใหญ่จะมีค่าอยู่ระหว่าง -2 ถึง $+2$ ค่า -2 แสดงว่าข้อสอบขง้นง่ายมาก และค่า $+2$ แสดงว่าข้อสอบนี้ยากมาก

a_g มีค่าตั้งแต่ $-\infty$ ถึง $+\infty$ แต่ในทางปฏิบัติส่วนใหญ่ a_g มีค่าตั้งแต่ 0 ถึง $+2$ เพราะค่า a_g ที่เป็นลบ แสดงว่าข้อสอบไม่ใช้ไม่ได้จึงตัดข้อสอบข้อนั้นทิ้งไป ค่า 0 แสดงว่าข้อสอบไม่มีอำนาจจำแนก หรือมีค่าอำนาจจำแนกต่ำ ค่า $+2$ แสดงว่าข้อสอบมีค่าอำนาจจำแนกสูง

θ มีค่าอยู่ระหว่าง $-\infty$ ถึง $+\infty$ แต่ในทางปฏิบัติส่วนใหญ่มีค่าอยู่ระหว่าง -3 ถึง $+3$ ค่า -3 แสดงว่ามีความสามารถต่ำ และค่า $+3$ แสดงว่ามีความสามารถสูง

ทฤษฎีคุณลักษณะแฝง ได้มีการพัฒนาไปในรูปแบบต่าง ๆ ทำให้เกิดเป็นโมเดลเฉพาะขึ้น โดยที่แต่ละโมเดลก็จะแตกต่างกันด้วยรูปแบบของฟังก์ชันทางคณิตศาสตร์และจำนวนพารามิเตอร์ที่ใช้ในการอธิบายฟังก์ชันโค้งลักษณะของข้อสอบ (Item Characteristic Function) สำหรับโมเดลที่พัฒนาแล้วก้าวหน้าไปมากก็คือ โลจิสติกโมเดล (ผจงจิต อินทสุวรรณ 2525 หน้า 57)

โลจิสติกโมเดล

ในปี ค.ศ. 1968 เบิร์นบอม (Birnbaum) เป็นผู้เสนอโลจิสติกโมเดลที่ใช้กับพารามิเตอร์ 2 ตัว คือ ค่าความยาก (b) และค่าอำนาจจำแนก (a) โดยที่โมเดลนี้สามารถดัดแปลงให้ใช้กับพารามิเตอร์ตัวเดียวคือค่าความยาก (b) และพารามิเตอร์สามตัวคือ ความยาก (b) ค่าอำนาจจำแนก (a) และค่าการเกาะ (c) ได้ (Warr, อ้างถึงใน ปันดิกา วิยวัฒน์ 2528 หน้า 15)

1. โมเดลที่ใช้พารามิเตอร์ 2 ตัว (Two-parameter logistic model) มีสมการดังนี้ (ผจงจิต อินทสุวรรณ 2525 หน้า 59)

$$P_i(\theta) = \frac{e^{D_{i1}(\theta - b_{i1})}}{1 + e^{D_{i1}(\theta - b_{i1})}} \quad (i = 1, 2, \dots, n)$$

θ คือ ระดับความสามารถที่แท้จริงที่คำนวณจากคะแนนจริงแต่ปรับหน่วยให้เป็นมาตรฐาน

$P_i(\theta)$ คือ ค่าความน่าจะเป็นที่ผู้สอบที่มีความสามารถ θ จะทำข้อสอบ i ได้ถูกต้อง

D คือ ค่าคงที่ซึ่งมีค่าเท่ากับ 1.7

a_{i1} คือ ค่าอำนาจจำแนกซึ่งมีค่าเป็นสัดส่วนโดยตรงกับค่าความชันของโค้ง ณ จุดเปลี่ยนโค้งหรือจุดที่ชันที่สุด

b_{i1} คือ ค่าความยากที่แสดงระดับความสามารถที่แท้จริง ที่จุดโค้งชันที่สุด หรือในกรณีที่ไม่มีการเคา ค่า b_{i1} คือ θ ณ จุดความน่าจะเป็น .50

2. โมเดลที่ใช้พารามิเตอร์ 3 ตัว (Three-parameter logistic model) เป็นโมเดลที่ดัดแปลงมาจากโมเดลที่ใช้พารามิเตอร์ 2 ตัว โดยเพิ่มพารามิเตอร์อีก 1 ตัว คือค่าสัมประสิทธิ์การเคา (C_i) มีสมการดังนี้คือ

$$P_i(\theta) = C_i + (1 - C_i) \frac{e^{D_{i1}(\theta - b_{i1})}}{1 + e^{D_{i1}(\theta - b_{i1})}} \quad (i = 1, 2, \dots, n)$$

3. โมเดลที่ใช้พารามิเตอร์ตัวเดียว (One-parameter logistic model) หรือ ราชโมเดล (Rasch Model) ราช (1960) ได้พัฒนาทฤษฎีคุณลักษณะแฝงโดยที่ฟังก์ชันของโมเดลนั้นสามารถอธิบายได้ด้วยพารามิเตอร์ของข้อสอบเพียงตัวเดียวคือค่าความยาก (b_i) ซึ่งตรงกับโมเดลที่ใช้พารามิเตอร์ตัวเดียวที่พัฒนาขึ้น โดยเบิร์นบอม (Birnbaum, 1968) โมเดลนี้ ถือว่าไม่มีค่าการเคาและค่าอำนาจจำแนกของข้อสอบคงที่ตลอดทั้งฉบับ จึงสามารถเขียนสมการได้ดังนี้ (ผจงจิต อินทสุวรรณ 2525 หน้า 61)

$$P_i(\theta) = \frac{e^{(\theta - b_i)}}{1 + e^{(\theta - b_i)}} \quad (i = 1, 2, \dots, n)$$

ความแกร่งของทฤษฎีคุณลักษณะแฝง (Robustness of Latent Trait)

จากการศึกษาข้อตกลงเบื้องต้นของทฤษฎีคุณลักษณะแฝงจะเห็นว่าข้อตกลงทางด้านการเป็นมิติเดียว (Unidimensionality) และความเป็นอิสระของข้อคำถาม (Local Independent) เป็นการยากที่จะสร้างแบบทดสอบใหม่คุณสมบัติดังกล่าว จากการศึกษามาก่อน (Hambleton) และนักวิจัยอื่น ๆ ในเรื่องนี้ ถึงแม้แต่ละคนใช้วิธีการแตกต่างกันแต่ผลที่สรุปได้ก็เหมือนกัน คือตามข้อตกลงในโมเดลของทฤษฎีคุณลักษณะแฝงนั้นไม่มีข้อมูลใด ๆ ที่จะมีคุณสมบัติครบถ้วนสมบูรณ์ตามโมเดลของทฤษฎีคุณลักษณะแฝงที่กำหนดไว้ อย่างไรก็ตาม โมเดลนี้ยังสามารถใช้ได้กับข้อมูลประเภทนี้เพราะผลการวิเคราะห์ข้อมูลลักษณะนี้ยังสอดคล้อง (Fit) กับโมเดลที่สรุปแล้วโมเดลของทฤษฎีคุณลักษณะแฝงมีความแกร่ง (บุญมี พันธุ์ไทย 2528 หน้า 21)

งานวิจัยที่เกี่ยวข้อง

งานวิจัยที่เกี่ยวข้องกับการกำหนดจุดตัดของแบบทดสอบอิงเกณฑ์

ในต่างประเทศได้มีการศึกษาค้นคว้าวิจัยเกี่ยวกับเรื่องจุดตัดของแบบทดสอบอิงเกณฑ์มานานแล้ว สำหรับในประเทศไทยนั้น แบบทดสอบอิงเกณฑ์เริ่มเข้ามามีบทบาทมากขึ้น ดังนั้น การศึกษาค้นคว้าวิจัยเกี่ยวกับเรื่องจุดตัดของแบบทดสอบอิงเกณฑ์ก็จะมีเพิ่มมากขึ้นตามลำดับ ผู้วิจัยเสนอผลงานที่เกี่ยวข้องกับจุดตัดของแบบทดสอบอิงเกณฑ์ดังนี้

บล็อก (Block อ้างถึงใน สำเร็จ บุญเรืองรัตน์ 2527 หน้า 66) ได้ทำการทดลองเพื่อค้นหาว่าเกณฑ์ระดับใดจึงจะเหมาะสมโดยแบ่งนักเรียนเป็นกลุ่ม ๆ และตั้งเกณฑ์ไว้ดังนี้ 65, 75, 85 และ 95 เปอร์เซนต์ เมื่อทำการสอนจนได้ครบตามเกณฑ์ที่กำหนดไว้ เขาพบว่ากลุ่มที่ตั้งเกณฑ์ไว้สูงสุดคือ 95% จะมีผลสัมฤทธิ์ทางการเรียนที่ดีที่สุด แต่เมื่อจบการศึกษาไปนาน ๆ ความสนใจและทัศนคติจะไม่ดีต่อการสอน ซับโคเวียค (Subkoviak , 1976 อ้างถึงใน ชวลิต โพธิ์นคร 2528 หน้า 32) พบว่าการกำหนดจุดตัดไว้ต่ำหรือสูงสุดจะทำให้ค่าความเชื่อมั่นของข้อสอบมีค่าสูงสุดและเมื่อกำหนดจุดตัดไว้ต่ำกลาง ๆ จะทำให้ค่าความเชื่อมั่นมีค่าต่ำสุด คลาร์ตี (Klarty , 1982 อ้างถึงใน ชวลิต โพธิ์นคร 2528 หน้า 33) ศึกษาเกี่ยวกับตัวแปรที่ส่งผล

ต่อการกำหนดจุดตัดพบว่าการกำหนดจุดตัดของครูขึ้นอยู่กับประสบการณ์ของครูผู้กำหนดจุดตัด เบอร์ก (Berk, 1976: 5 - 7) ได้ศึกษาวิธีการเลือกจุดตัดของคะแนนในแบบทดสอบอิงเกณฑ์พบว่าเกณฑ์ 50 เปอร์เซ็นต์คือเกณฑ์ที่ตัดสินได้ถูกต้องที่สุด มาเลช (Malech, 1978 อ้างถึงใน สุทธิวรรณ หิรัศกิจโสภณ 2529 หน้า 39) ได้ทำการศึกษาเปรียบเทียบวิธีการกำหนดจุดตัดที่แบบในแบบทดสอบวัดความสามารถในการอ่านของนักเรียนระดับ 12 การหาคะแนนจุดตัดที่แบบคือ แบบไบโนเมียล แบบเบส์ และแบบการใช้วิธีหาความน่าจะเป็นอีก 2 แบบ จากการศึกษาพบว่าวิธีใช้ข้อมูลแบบไม่เป็นอิสระ ให้จุดตัดสูงกว่าข้อมูลแบบอิสระ และพบว่าแบบทดสอบที่มีจำนวนข้อของข้อสอบน้อย จะให้จุดตัดที่ตรงกัน และทุกวิธีให้จุดตัดที่อยู่กลาง ๆ มากกว่าจะอยู่ค่อนข้างต่ำหรือค่อนข้างสูง

สำหรับในประเทศไทย สมหวัง พิธิยานุวัฒน์ (2520 หน้า 13) ได้ศึกษาพบว่าคะแนนจุดตัดหรือเกณฑ์ที่สอดคล้องกับการเรียนวิชาสถิติศาสตร์คือเกณฑ์ขั้นต่ำระดับสูงเป็น 53.57 เปอร์เซ็นต์ และเกณฑ์ขั้นสูงระดับต่ำเป็น 70 เปอร์เซ็นต์ ต่อมาในปี 2521 กาญจนา วัธนสุนทร (2521 หน้า 106 - 108) ได้สร้างแบบทดสอบอิงเกณฑ์วิชาคณิตศาสตร์ชั้นมัธยมศึกษาปีที่ 1 เรื่องสมการจำนวน 4 ฉบับและกำหนดจุดตัดเป็น 60, 70 และ 80 เปอร์เซ็นต์ พบว่าเมื่อใช้จุดตัดเป็น 60% จะให้ความเชื่อมั่นและความเที่ยงตรงสูงสุด ในปี 2523 ชมภู จันทรมรรพ (2523 หน้า 75 - 76) ได้ปรับปรุงแบบทดสอบวิชาคณิตศาสตร์ที่กาญจนา วัธนสุนทร สร้างขึ้น และพบว่าแบบทดสอบนี้มีเกณฑ์ดังนี้ ฉบับที่ 1 เป็น 60 เปอร์เซ็นต์ ฉบับที่ 2, 3 และ 4 มีเกณฑ์เป็น 50 เปอร์เซ็นต์ ต่อมาในปี 2524 ไพฑูรย์ เวทการ (2524 หน้า 67 - 72) ได้สร้างแบบทดสอบอิงเกณฑ์วิชาคณิตศาสตร์เรื่องอัตราส่วนของชั้นมัธยมศึกษาปีที่ 1 จำนวน 50 ข้อโดยแยกเป็น 5 ฉบับ ฉบับละ 10 ข้อ ศึกษาจุดตัดที่เหมาะสมโดยคำนวณจากทฤษฎีการตัดสินใจของแกลสพบว่าจุดตัดที่เหมาะสมของแบบทดสอบฉบับที่ 1, 2 และ 5 เป็น 60 เปอร์เซ็นต์ ฉบับที่ 3 และ 4 เป็น 50 เปอร์เซ็นต์ สมฉวี วิจิตรวรรณ (2524 หน้า 115 - 120) ได้สร้างแบบทดสอบอิงเกณฑ์วิชาคณิตศาสตร์ เรื่องการหาร ของชั้นประถมศึกษาปีที่ 2 จำนวน 4 ฉบับ พบว่าคะแนนจุดตัดที่คำนวณโดยใช้ค่าเฉลี่ยของครูผู้สอน และวิธีของเบอร์ก มีค่าตั้งแต่ 50 - 80 เปอร์เซ็นต์ ต่อมาในปี 2525 บุญเลิศ คำหอม (2525 หน้า 83 - 88) ได้สร้างแบบทดสอบอิงเกณฑ์วิชาคณิตศาสตร์ เรื่องสมการและอสมการของชั้นมัธยมศึกษาปีที่ 2 โดยแยกเป็น 4 ฉบับ หาคะแนนจุดตัดโดยวิธีของ

เบอร์คและจากการกำหนดของครูผู้สอนพบว่าคะแนนจุดตัดที่เหมาะสมของฉบับที่ 1 และ 4 คือ 50 เปอร์เซนต์ ส่วนฉบับที่ 2 และ 3 คือ 60 เปอร์เซนต์

ต่อมาในปี 2526 เกศริน บุญเกิด (2526 หน้า 84) ได้สร้างแบบทดสอบอิงเกณฑ์วิชาภาษาไทย เรื่องการสะกดคำของชั้นประถมศึกษาปีที่ 4 โดยแยกเป็น 3 ฉบับ หาคะแนนจุดตัดโดยใช้ทฤษฎีการตัดสินใจของแกลส พบว่าคะแนนจุดตัดของแบบทดสอบฉบับที่ 1 และ 2 คือ 60 เปอร์เซนต์ และฉบับที่ 3 คือ 67 เปอร์เซนต์ ในปีเดียวกัน มัทนี เผื่อนน้อย (2526 หน้า 96 - 101) ได้สร้างแบบทดสอบอิงเกณฑ์วิชาคณิตศาสตร์ เรื่องร้อยละของชั้นประถมศึกษาปีที่ 4 จำนวน 4 ฉบับ และศึกษาหาคะแนนจุดตัดที่เหมาะสมโดยใช้ทฤษฎีการตัดสินใจของแกลสพบว่า ฉบับที่ 1, 2 และ 3 มีคะแนนจุดตัดเป็น 50 เปอร์เซนต์ ส่วนฉบับที่ 4 มีคะแนนจุดตัดเป็น 60 เปอร์เซนต์ สุโชค สันตกิจวงศ์โชย (2526 หน้า 102 - 104) ได้สร้างแบบทดสอบอิงเกณฑ์วิชาคณิตศาสตร์ เรื่องสมการควอดรติก ชั้นมัธยมศึกษาปีที่ 3 โดยใช้แบบทดสอบย่อยจำนวน 6 ฉบับ และหาคะแนนจุดตัดตามวิธีของเบอร์คมีค่าอยู่ระหว่าง 30 - 50 เปอร์เซนต์ และคะแนนจุดตัดตามวิธีของแกลส วิธีของเบอร์ค และวิธีของเบส พบว่าวิธีที่ทำให้คะแนนจุดตัดมีค่าสูงสุดคือวิธีการของเบสและค่าต่ำสุดคือวิธีของเบอร์ค

ต่อมาในปี 2527 ประเทือง หาลีแสง (2527 หน้า 112) ได้สร้างแบบทดสอบอิงเกณฑ์ วิชาคณิตศาสตร์ เรื่องค่าตัวตั้งและกราฟของชั้นมัธยมศึกษาปีที่ 1 โดยสร้างเป็นแบบทดสอบคู่ขนาน 3 ชุด และหาคะแนนจุดตัดโดยวิธีให้ผู้เชี่ยวชาญพิจารณาเกณฑ์นับผลจาก 100% กับวิธีของชวณ พบว่าคะแนนจุดตัดของแบบทดสอบคู่ขนานชุดที่ 1, 2 และ 3 คือ 80, 87 และ 90 เปอร์เซนต์ ตามลำดับ เชาวลิต โอภาสวัฒนา (2527 บทคัดย่อ) ได้ศึกษาความสัมพันธ์ระหว่างระดับของคะแนนจุดตัดกับผลสัมฤทธิ์และเจตคติต่อวิชาคณิตศาสตร์ พบว่านักเรียนที่ใช้จุดตัด 80% และ 60% มีผลสัมฤทธิ์ทางการเรียนสูงกว่านักเรียนที่ใช้จุดตัด 40% แต่นักเรียนทั้ง 3 กลุ่มมีเจตคติไม่แตกต่างกัน รังสรรค์ มณีเล็ก (2527 หน้า 73 - 75) ได้เปรียบเทียบความเชื่อมั่นของแบบทดสอบอิงเกณฑ์โดยใช้วิธีการกำหนดจุดตัด 4 วิธีคือ วิธีนับถอยหลังจาก 100% วิธีของนิตลสกี วิธีของเบอร์ค และวิธีของเบส พบว่าทั้ง 4 วิธีมีค่าความเชื่อมั่นแตกต่างกันอย่างไม่มีนัยสำคัญและวิธีให้คะแนนจุดตัดในระดับกลางคือวิธีของเบอร์คและวิธีของเบส

ต่อมาในปี 2528 กำนัน เกตุสุวรรณ (2528 หน้า 126 - 132) ได้สร้างแบบทดสอบ
อิงโคเนวิชคณิตศาสตร์เรื่องลำดับของชั้นมัธยมศึกษาปีที่ 5 โดยสร้างเป็นแบบทดสอบภูเขานาจำ-
นวน 4 ชุด และหาคะแนนจุดตัดตามวิธีของชวณ พบว่าคะแนนจุดตัดของแบบทดสอบชุดที่ 1, 2, 3
และ 4 คือ 70, 45, 44 และ 50 เปอร์เซนต์ตามลำดับ ชวลิต โพธิ์นคร (2528 หน้า 77 - 90)
ได้เปรียบเทียบผลการกำหนดจุดตัดตามวิธี ของเบอร์ค วิธีประยุกต์ราชคโมเดล และวิธีกำหนดเกณฑ์
ผ่านระดับต่ำสุดพบว่าคะแนนจุดตัดที่คำนวณโดยวิธีกำหนดเกณฑ์ผ่านระดับต่ำสุดมีค่าสูงสุดส่วนคะแนนจุดตัด
ที่คำนวณโดยวิธีของเบอร์ค และวิธีประยุกต์ราชคโมเดล นั้น มีค่ากลาง ๆ คือประมาณ 50% ของคะแนน
เต็มและคะแนนจุดตัดที่ได้โดยวิธีประยุกต์ราชคโมเดลนั้นมีแนวโน้มที่จะไม่แปรเปลี่ยนไปตามระดับ
ความสามารถของกลุ่มตัวอย่าง ส่วนคะแนนจุดตัดที่ได้โดยวิธีของเบอร์คมีแนวโน้มที่จะแปรเปลี่ยนตาม
ความสามารถของกลุ่มตัวอย่าง

ต่อมาในปี 2529 สุทธิธรรม หิรัญศิริโสภณ (2529 หน้า 75 - 80) ได้ศึกษาคะแนนจุดตัด
ที่คำนวณโดยวิธีของชวณ วิธีของแกลสและวิธีของเบอร์ค พบว่าในแบบทดสอบฉบับเดียวกันจะคะแนน
จุดตัดที่เหมาะสมเพียงค่าเดียวคือมีค่าเป็นร้อยละ 50 ของคะแนนเต็ม

จากผลงานวิจัยดังกล่าว จะเห็นได้ว่าคะแนนจุดตัดของแบบทดสอบอิงเกณฑ์ จะมีค่ากลาง ๆ
หรือประมาณ 50% ของคะแนนเต็ม แต่อย่างไรก็ตามคะแนนจุดตัดที่ได้นี้ก็อาจจะแปรเปลี่ยนตามระดับ
ความสามารถของผู้เข้าสอบ ฉะนั้นจากงานวิจัยของชวลิต โพธิ์นคร ที่ได้นำเอาทฤษฎีคุณลักษณะแฝง
เข้ามาช่วยในการกำหนดจุดตัดพบว่า ทำให้เกิดผลดีในแง่ของคะแนน จุดตัดไม่แปรเปลี่ยนตามระดับ
ความสามารถของกลุ่มตัวอย่าง ดังนั้น ในการศึกษาครั้งนี้ก็จะได้นำเอาทฤษฎีคุณลักษณะแฝงในโมเดล
โลจิสติกส์อื่น ๆ เข้ามาเปรียบเทียบกับราชคโมเดลในแง่ของการประยุกต์ใช้โดยกำหนดจุดตัดแต่ละวิธี
และศึกษาเปรียบเทียบผลที่ตามมา

งานวิจัยที่เกี่ยวข้องกับทฤษฎีคุณลักษณะแฝง

ฮัทเทน (Hutten, 1982 อ้างถึงใน ปันดดา วิวัฒน์ 2528 หน้า 28 - 29) ได้
ทำการศึกษาดังความสัมพันธ์ระหว่างการวิเคราะห์ข้อสอบแบบราชคโมเดล กับการวิเคราะห์แบบพารา-
มิเตอร์ 3 ตัว ตามข้อตกลงเบื้องต้นของแต่ละโมเดล ผลปรากฏว่า จำนวนข้อของข้อสอบที่เหมาะสม
(Fit) กับโมเดลทั้งสองไม่แตกต่างกัน และพบว่าขนาดของกลุ่มตัวอย่างที่จะทำการวิเคราะห์แบบราชค
โมเดลใช้ประมาณ 250 คนขึ้นไป ส่วนการวิเคราะห์แบบพารามิเตอร์ 3 ตัวใช้กลุ่มตัวอย่างประมาณ

1,000 คนขึ้นไป เจมส์ รี (James Ree, 1981 อ้างถึงใน นภา กาญจนกิจโสภณ 2528 หน้า 11) พบว่า ความสัมพันธ์ระหว่างค่าความสามารถที่แท้จริงและค่าความสามารถจากการประมาณจะสูงมากถ้าในการสอบนั้นมีจำนวนข้อสอบที่มากเพียงพอ ค่าความคลาดเคลื่อนมาตรฐานของการทำนายค่าพารามิเตอร์ของข้อสอบและของผู้เข้าสอบสามารถทำให้น้อยลงได้เมื่อมีข้อสอบอย่างน้อย 200 ข้อ และกลุ่มตัวอย่าง 2,000 คน ชูศักดิ์ ชัมภลลิขิต (2525 หน้า 33 - 43) ทำการศึกษาเปรียบเทียบผลของการให้คะแนนแบบให้น้ำหนักคะแนนรายข้อต่างกันตามทฤษฎีการตอบสนองของข้อสอบ (IRT) โมเดลที่มีพารามิเตอร์ 3 ตัว กับการให้คะแนนแบบให้น้ำหนักคะแนนรายข้อเท่ากัน คะแนนมาตรฐาน (Z) และคะแนนความสามารถ (θ) โดยศึกษากับข้อมูลจากสภาพการสอบคัดเลือกจริง โดยที่ข้อสอบคัดเลือกส่วนใหญ่เป็นข้อสอบที่เขียนขึ้นมาใหม่และยังไม่มีผลการทดลองใช้ จากการศึกษาพบว่า ในการสอบคัดเลือกสามารถนำคะแนนจากการสอบไปใช้กับทฤษฎีการตอบสนองของข้อสอบ โมเดลที่มีพารามิเตอร์ 3 ตัวได้ คุณภาพของข้อสอบที่เขียนขึ้นมาใหม่ไม่เป็นปัญหาต่อการให้คะแนนและการให้คะแนนแบบให้น้ำหนักรายข้อต่างกันทำให้ค่าความเชื่อมั่นและค่าอำนาจจำแนกของแบบทดสอบสูงกว่าการให้คะแนนแบบอื่น ๆ อวยพร วิบูลย์กาญจน์ (2526 หน้า 47 - 48) ได้นำเอาวิธีการวิเคราะห์ข้อสอบแบบราชภัฏโมเดลมาใช้ในการศึกษา เรื่องการเปรียบเทียบผลการวิเคราะห์แบบอุปมาอุปไมย ด้วยทฤษฎีสถิติของโมเดลกับราชภัฏโมเดลพบว่า จำนวนข้อสอบที่คัดเลือกไว้ด้วยทฤษฎีสถิติของโมเดลมีมากกว่าการคัดเลือกด้วยราชภัฏโมเดลและค่าความเชื่อมั่นของแบบทดสอบที่ประกอบด้วยข้อสอบที่คัดเลือกไว้ด้วยทฤษฎีสถิติของโมเดล มีค่าสูงกว่าของราชภัฏโมเดล สมพร บุญคุ้ม (2528 หน้า 51 - 53) ได้ศึกษาความคงที่ของค่าพารามิเตอร์ความยากในการวิเคราะห์ข้อสอบด้วยราชภัฏโมเดลพบว่า ค่าความยากของข้อสอบจะมีค่าคงที่ในข้อสอบที่มีค่าความยากปานกลาง ประสิทธิ์ เรืองตระกูล (2528 หน้า 80 - 84) ได้สร้างแบบทดสอบวินิจัยวิชาคณิตศาสตร์สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 1 โดยใช้โมเดลโลจิสติก พบว่า แบบทดสอบที่สร้างขึ้นมีความยากตั้งแต่ -3.75 ถึง 1.90 ค่าอำนาจจำแนกตั้งแต่ 0.19 ถึง 2.00 และค่าการเค้นค่าตั้งแต่ 0.0 ถึง 0.5 ปณิศา วิยวัฒน์ (2528 หน้า 65 - 70) ได้เปรียบเทียบผลการวิเคราะห์ข้อสอบโดยวิธีโลจิสติกโมเดลกับวิธีทฤษฎีสถิติของโมเดลในวิชาคณิตศาสตร์ชั้นมัธยมศึกษาปีที่ 1 พบว่า จำนวนข้อของข้อสอบที่ได้รับการคัดเลือกไว้จากการวิเคราะห์ข้อสอบแบบพารามิเตอร์ตัวเดียว พารามิเตอร์ 2 ตัว และพารามิเตอร์ 3 ตัว ไม่แตกต่างกัน แต่มีจำนวนข้อแตกต่างกับการวิเคราะห์โดยวิธีทฤษฎีสถิติของโมเดล และการคัดเลือก

ข้อสอบจากการวิเคราะห์แบบพหุวิธีสองตัวกับพหุวิธีสามตัว มีความสอดคล้องกัน
 จากผลงานการวิจัยที่มีผู้นำเอาทฤษฎีคุณลักษณะแห่งมาทำการศึกษาดังกล่าว จะเห็นได้
 ว่าในปัจจุบันทฤษฎีนี้ได้เข้ามามีบทบาทต่อการพัฒนาการศึกษาในสถานการณต่าง ๆ และในการ
 ศึกษาครั้งนี้จะได้นำเอาโมเดลโลจิสติกมาประยุกต์ในการกำหนดจุดตัดของแบบทดสอบอิงเกณฑ์
 วิชาคณิตศาสตร์ที่ผู้วิจัยสร้างขึ้น และเปรียบเทียบผลซึ่งเกิดจากการกำหนดจุดตัดนั้น

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
 Copyright© by Chiang Mai University
 All rights reserved