



บทที่ 1

บทนำ

ความสำคัญและความเป็นมาของปัญหา

การวัดผลสัมฤทธิ์ของผู้เรียนหรือการประเมินโครงการเรียนแต่เดิมนั้นส่วนใหญ่ใช้การสอบที่เราเรียกว่า การสอบอิงกลุ่ม (Norm-Referenced) ซึ่งมีหลักการสำคัญเพื่อจะมุ่งวัดผลสัมฤทธิ์ทางการเรียนและบ่งชี้ระดับความสามารถของผู้เรียนจากการเปรียบเทียบกับคนภายในกลุ่มเดียวกันว่าใครมีความสามารถสูงหรือต่ำกว่ากัน ต่อมาได้มีนักวัดผลได้เสนอแนวคิดและหลักการสอบแนวใหม่ขึ้นมาเรียกว่า การสอบอิงเกณฑ์ (Criterion-Referenced) โดยมีหลักอยู่ว่าการวัดผลสัมฤทธิ์ทางการเรียนควรมุ่งบ่งชี้ถึงระดับผลสัมฤทธิ์ในด้านความรู้หรือศักยภาพของผู้เรียนแต่ละคนในเรื่องใดเรื่องหนึ่งว่ามีอยู่ในระดับใดออกมาอย่างชัดเจน (พิศิษฐ์ คัมภวณิช 2529: 1) หลักการของทั้งสองแนวคิดในการสร้างแบบสอบนี้ก็เพื่อจะวัดผลสัมฤทธิ์ พฤติกรรม หรือความสามารถของผู้เรียนซึ่งเป็นคุณลักษณะภายใน (Latent Traits) ซึ่งไม่สามารถจะสังเกตได้โดยตรงแต่ผลที่ได้ก็นักวัดผลได้กำหนดอยู่ในรูปคะแนนและใช้คะแนนที่ได้เป็นตัวแทนของความสามารถของผู้สอบซึ่งในทฤษฎีการวัดความสามารถต่าง ๆ ก็พยายามที่จะให้ได้ผลการวัดแต่ละครั้งตรงกับความสามารถที่แท้จริงของผู้สอบที่เรียกว่า คะแนนจริง (True Score) มากที่สุด

ในปัจจุบันทฤษฎีการวัดผลการศึกษาได้มีการพัฒนาไปอย่างกว้างขวางและที่ได้รับความนิยมใช้กันมากเรียกว่า ทฤษฎีมาตรฐานดั้งเดิมหรือทฤษฎีคลาสสิกอล (Classical Test Theory) ซึ่งเน้นการประมาณค่าคะแนนจริงหรือความสามารถที่แท้จริงโดยใช้คะแนนของกลุ่มผู้สอบเป็นสำคัญ ในการวิเคราะห์คุณสมบัติของข้อคำถามพิจารณาที่ค่าความยากและค่าอำนาจจำแนกที่มักเปลี่ยนแปลงไปตามกลุ่มผู้สอบในแต่ละครั้ง จึงถูกวิจารณ์ว่าเป็นทฤษฎีที่ยังไม่มีเหตุผลเพียงพอในการประมาณค่าความสามารถที่แท้จริงเป็นรูปแบบหรือโมเดลที่อ่อน ซึ่งปัญหาเหล่านี้นักวัดผลหลายท่านเช่น Alfred Binet หรือ L.L. Thurstone ได้กล่าวถึงมานานแล้ว นักทดสอบทางจิตวิทยาและนักวัดผลทางการศึกษาจำนวนมากได้สำรวจและพัฒนาทฤษฎีที่เหมาะสมกว่า ในการวัดความสามารถและได้มุ่งมาที่ทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT) หรือทฤษฎีคุณลักษณะแฝง

ซึ่งเริ่มด้วยรูปแบบทางคณิตศาสตร์ที่กล่าวถึงความสัมพันธ์ระหว่างการตอบข้อสอบของผู้สอบและระดับความสามารถของผู้สอบด้วยค่าพารามิเตอร์ และค่าพารามิเตอร์เหล่านี้สามารถนำไปอธิบายหรือทำนายโอกาสของผู้สอบในการทำข้อสอบ ถึงแม้ว่าผู้สอบจะมีลักษณะไม่คล้ายคลึงกับกลุ่มตัวอย่างที่ได้ค่าพารามิเตอร์นั้น ๆ มากก็ตาม รูปแบบของทฤษฎีการตอบสนองข้อสอบอยู่ในรูปฟังก์ชันทางคณิตศาสตร์ที่เรียกว่า โค้งลักษณะข้อสอบ (Item Characteristic Curve: ICC) ซึ่ง Frederic M. Lord ได้พัฒนาโดยใช้ Normal Ogive Model ต่อมา Birnbaum ได้พัฒนาโดยใช้ Logistic Model แทน Normal Ogive Model และ Georg Rasch ได้พัฒนารูปแบบที่ใช้พารามิเตอร์เพียงตัวเดียวและเป็นที่ยอมรับอย่างกว้างขวาง

Hambleton (Hambleton 1979: 14-15) ได้กล่าวถึงข้อบกพร่องที่ทฤษฎีมาตรฐานดั้งเดิม ยังไม่สามารถแก้ปัญหาได้ คือ

1. ค่าพารามิเตอร์ของข้อสอบของทฤษฎีมาตรฐานดั้งเดิม คือค่าความยากและค่าอำนาจจำแนกมีค่าไม่คงที่เปลี่ยนแปลงไปตามกลุ่มของผู้สอบที่มีความสามารถแตกต่างกัน ค่าพารามิเตอร์ดังกล่าวของข้อสอบจะใช้ประโยชน์ในการสร้างแบบสอบเฉพาะกับกลุ่มประชากรผู้สอบที่มีลักษณะคล้ายคลึงกับกลุ่มตัวอย่างผู้สอบที่ใช้ในการหาค่าพารามิเตอร์เหล่านั้น ซึ่งในหลาย ๆ สถานการณ์ของการสอบผู้เข้ารับการสอบต้องการค่าพารามิเตอร์ของข้อสอบที่ไม่เปลี่ยนแปลง ซึ่งทฤษฎีการตอบสนองข้อสอบมีคุณสมบัติของค่าพารามิเตอร์ที่จะไม่เปลี่ยนแปลงไปตามลักษณะความสามารถของกลุ่มผู้สอบ
2. การเปรียบเทียบผู้สอบในเรื่องของความสามารถที่ระดับหนึ่งซึ่งวัดจากแบบสอบจำกัดอยู่กับสถานการณ์ที่ผู้สอบถูกทดสอบด้วย นั่นคือ ใช้ข้อสอบชุดเดียวกันหรือเป็นคู่ขนานกัน โดยส่วนใหญ่จะเป็นแบบสอบมาตรฐานที่จะวัดผลสัมฤทธิ์เฉพาะกับนักเรียนที่มีความสามารถปานกลางซึ่งข้อสอบเหล่านี้ไม่ให้การประมาณค่าความสามารถที่แน่นอนสำหรับนักเรียนที่มีความสามารถระดับสูงหรือต่ำ ซึ่งในทฤษฎีการตอบสนองข้อสอบการเปรียบเทียบความสามารถไม่จำเป็นต้องใช้แบบสอบชุดเดียวกันเพียงแต่ใช้แบบสอบที่วัดคุณลักษณะเดียวกันก็ใช้ได้
3. เทคนิคการสอบที่ใช้อยู่ในปัจจุบันไม่ได้เป็นพื้นฐานที่จะบอกได้ว่าผู้สอบคนหนึ่งจะสามารถทำข้อสอบได้เพียงใดเมื่อได้เผชิญกับข้อสอบหรือคำถามชุดหนึ่ง ยกเว้นเมื่อได้มีการใช้ข้อสอบนั้นแล้วกับกลุ่มผู้สอบที่คล้ายคลึงกับบุคคลนั้น ซึ่งในทางปฏิบัติเราต้องการค่าพารามิเตอร์ของข้อสอบ (Item parameter) และพารามิเตอร์ของผู้สอบ (Examinee parameters) มาใช้เป็นตัวกำหนดลักษณะของข้อสอบเพื่อที่จะสามารถทำนายความน่าจะเป็น

ที่ผู้สอบคนหนึ่งจะตอบข้อสอบข้อหนึ่งได้ถูก แม้ว่าข้อสอบที่คล้ายคลึงกันนั้นจะยังไม่เคยใช้กับผู้สอบคนนี้เลย

ในทฤษฎีการตอบสนองข้อสอบมีหลายรูปแบบที่สามารถนำมาใช้ได้กับการสร้างแบบสอบเพื่อวัดผลสัมฤทธิ์และความสามารถที่แท้จริงของผู้สอบ รูปแบบของราสช์ (Rasch Model) เป็นรูปแบบที่มีความซับซ้อนน้อยที่สุดของทฤษฎีนี้โดยใช้พารามิเตอร์เพียง 1 ตัว มาใช้อธิบายลักษณะของข้อสอบคือ พารามิเตอร์ความยาก (b) โดยมีข้อตกลงว่าไม่มีการเดา (c) และค่าอ่านาจจำแนกของทุกข้อเท่ากัน (a) ซึ่งจะนำมาพิจารณากับพารามิเตอร์ที่อธิบายความสามารถของผู้สอบคือ ความสามารถ (θ) โดยที่โอกาสของการทำข้อสอบถูกจะแสดงโดยโค้งลักษณะของข้อสอบ (Item Characteristic Curve : ICC) ในปัจจุบันนี้รูปแบบของราสช์ได้เป็นที่สนใจของนักวิจัยทางการศึกษา ในการนำมาประยุกต์ใช้เกี่ยวกับการทดสอบทางการศึกษา

ในปี ค.ศ. 1975 Graham A. Douglass ได้เสนอรูปแบบของราสช์ในการใช้เป็นกลยุทธ์ (strategies) เกี่ยวกับการออกแบบแบบสอบเพื่อให้ผู้สร้างแบบสอบและผู้ใช้ได้พิจารณาถึงลักษณะของแบบสอบที่ดีในการนำไปใช้กับกลุ่มตัวอย่างต่าง ๆ โดยได้พัฒนาประสิทธิภาพในการออกแบบแบบสอบบนพื้นฐานของแบบสอบที่สัมพันธ์กับลักษณะเฉพาะเป้าหมายที่กำหนดขึ้น (Target Characteristics)

ในปี ค.ศ. 1977 Frederic M. Lord ได้เสนอการคัดเลือกข้อสอบและสร้างแบบสอบโดยพิจารณาค่าพารามิเตอร์ของข้อสอบแล้วคำนวณค่าฟังก์ชันแสดงสารสนเทศของข้อสอบ (Item information function) แต่ละข้อรวมกันเป็นฟังก์ชันแสดงสารสนเทศของแบบสอบ (Test information function) ต่อจากนั้นพิจารณารูปแบบของโค้งแสดงสารสนเทศเป้าหมายที่คงไว้กับความสัมพันธ์ของแต่ละข้อสอบที่บรรจุในแบบสอบว่ามีความสัมพันธ์กับโค้งแสดงสารสนเทศเป้าหมายเพียงใด ซึ่งวิธีนี้เป็นการพัฒนาการออกแบบแบบสอบเพื่อทดสอบตามกลุ่มตัวอย่างที่ต้องการ

ในปี ค.ศ. 1977-1978 Charles L. Hulin และคณะได้เสนอรูปแบบการพิจารณาข้อสอบจากโค้งแสดงสารสนเทศเป้าหมายที่กำหนดไว้คล้ายกับวิธีที่ Lord ได้เสนอไว้แต่ Hulin ได้ใช้ข้อมูลจาก Job Descriptive Index (JDI) ที่มีอยู่เป็นข้อมูลในการศึกษาโดยใช้รูปแบบ 2 พารามิเตอร์พิจารณา Attitude scale และรูปแบบ 3 พารามิเตอร์พิจารณาแบบสอบความสามารถทางภาษา (Test of Verbal Ability) ทำให้การออกแบบสร้างแบบสอบได้มีการพัฒนาขึ้นโดยอาศัยข้อมูลที่เก็บจากกลุ่มตัวอย่างจริง

จากข้อมูลที่ได้มีผู้วิจัยและลงข้อสรุปไว้ว่า หลักการวัดผลโดยใช้แบบสอบที่มีความยากพอเหมาะกับความสามารถของผู้สอบจะสามารถลดค่าความคลาดเคลื่อนลงได้ ทำให้ผู้วิจัยมีความสนใจที่จะสร้างข้อสอบและออกแบบลักษณะแบบสอบที่ใช้วัดได้เหมาะสมตรงกับระดับความสามารถของผู้สอบโดยใช้รูปแบบของราสช์มาประยุกต์ในการวิเคราะห์ค่าพารามิเตอร์ของข้อสอบที่สร้างขึ้น เนื่องจากรูปแบบของราสช์เป็นรูปแบบที่พิจารณาความสัมพันธ์ระหว่างค่าความยาก (b) ของข้อสอบกับค่าความสามารถ (θ) ของผู้สอบแต่ละคน ซึ่งเป็นรูปแบบที่ซับซ้อนน้อยที่สุดและสะดวกต่อการใช้มากที่สุดการประยุกต์รูปแบบของราสช์มาใช้ในการวิจัยครั้งนี้เนื่องจากมีข้อค้นพบปรากฏว่า รูปแบบของราสช์มีข้อได้เปรียบกว่ารูปแบบอื่น ๆ และเป็นที่น่าสนใจมาก อาทิ กลุ่มตัวอย่างประชากรมีขนาดเล็กประมาณ 100 คนก็ใช้ได้ แต่ถ้าจะให้ดีกลุ่มตัวอย่างควรใช้ประมาณ 400 คนขึ้นไปและข้อสอบควรมีความยาว 40 ข้อ (Wright 1977: 219) ค่าอำนาจจำแนกของข้อสอบนั้นไม่จำเป็นต้องเท่ากันและไม่จำเป็นต้องคำนึงถึงค่าการเดา (Wright and

Panchapakesan 1967: 25) การกระจายของความสามารถของกลุ่มตัวอย่างหรือความยาวของแบบสอบไม่จำเป็นต้องเป็นโค้งปกติเพียงแต่มีแนวโน้มเป็นโค้งปกติเท่านั้น และกลุ่มตัวอย่างไม่จำเป็นต้องมาจากการสุ่มก็ได้

วัตถุประสงค์ของการวิจัย

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อที่จะนำรูปแบบของราสช์มาประยุกต์ในการออกแบบโค้งแสดงสารสนเทศของแบบสอบด้วยการจัดข้อสอบที่มีระดับความยากง่ายให้เหมาะสมกับระดับความสามารถของนักเรียน โดยมีวัตถุประสงค์เฉพาะของการวิจัยดังนี้

1. ออกแบบโค้งแสดงสารสนเทศของแบบสอบด้วยการจัดข้อสอบตามค่าพารามิเตอร์ที่วิเคราะห์ได้ด้วยรูปแบบของราสช์ให้ผู้สอบมีระดับความสามารถต่างกัน 3 ระดับ คือ

- 1.1 กลุ่มผู้สอบที่มีความสามารถระดับสูง
- 1.2 กลุ่มผู้สอบที่มีความสามารถระดับปานกลาง
- 1.3 กลุ่มผู้สอบที่มีความสามารถระดับต่ำ

2. ตรวจสอบความเหมาะสมของแบบสอบที่จัดตามระดับความสามารถในข้อที่ 1 ด้วยวิธีการต่อไปนี้

2.1 ศึกษาความแม่นยำ (precision) ในด้านความเที่ยงและความคลาดเคลื่อนของแบบสอบที่จัดขึ้นตามระดับความสามารถ

สมมติฐานของการวิจัย

Hambleton และ Gruijter (1983) ได้ศึกษาการกำหนดความยาวของแบบสอบอิงเกณฑ์โดยใช้ทฤษฎีการตอบสนองข้อสอบแล้วพบว่าการใช้ข้อสอบที่มีระดับความยากพอเหมาะ กับระดับความสามารถของผู้สอบจะลดความคลาดเคลื่อนลงได้มากกว่าการใช้แบบสอบที่สร้างขึ้นจากการสุ่มข้อสอบขึ้นใช้ทดสอบ ซึ่งสอดคล้องกับ Haladyna และ Roid (1983) ที่ได้เปรียบเทียบผลของความคลาดเคลื่อนในการประมาณค่าความสามารถของบุคคลจากการใช้แบบสอบต่างกันโดยพบว่า การสอบที่จัดให้ข้อสอบมีค่าความยากตรงกับระดับความสามารถของผู้สอบจะทำให้ค่าความคลาดเคลื่อนลดลง และ Hambleton Mills และ Simon (1983) ได้ศึกษาข้อมูลที่สร้างขึ้นจากเครื่องคอมพิวเตอร์แล้วพบว่าแบบสอบที่มีค่าอำนาจจำแนกสูงขึ้นและตรงกับความสามารถของผู้สอบจะมีค่าความเที่ยงสูง จากข้อมูลดังกล่าวผู้วิจัยได้พิจารณาเปรียบเทียบค่าความแม่นยำ (precision) ในการวัด โดยพิจารณาตัวแปร 2 ตัวคือค่าความเที่ยง (reliability) และค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่า (Standard Error of Estimate : S.E.E.) ซึ่งผู้วิจัยขอตั้งสมมติฐานดังนี้

สมมติฐานที่ 1 แบบสอบที่ออกแบบสำหรับแต่ละกลุ่มความสามารถ เมื่อผู้สอบมีความสามารถต่างกันค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถน่าจะแตกต่างกัน

สมมติฐานที่ 2

1. แบบสอบที่ออกแบบสำหรับวัดกลุ่มความสามารถต่ำ เมื่อผู้สอบมีความสามารถต่ำ สอบจะให้ค่าความเที่ยงสูงกว่าเมื่อสอบแบบสอบที่เหมาะสมสำหรับกลุ่มความสามารถปานกลาง และสูงอย่างมีนัยสำคัญ
2. แบบสอบที่ออกแบบสำหรับวัดกลุ่มความสามารถปานกลาง เมื่อผู้สอบมีความสามารถปานกลางสอบจะให้ค่าความเที่ยงสูงกว่าเมื่อสอบแบบสอบที่เหมาะสมสำหรับกลุ่มความสามารถต่ำและสูงอย่างมีนัยสำคัญ
3. แบบสอบที่ออกแบบสำหรับวัดกลุ่มความสามารถสูง เมื่อผู้สอบที่มีความสามารถสูงสอบจะให้ค่าความเที่ยงสูงกว่าเมื่อสอบแบบสอบที่เหมาะสมสำหรับกลุ่มความสามารถต่ำและปานกลางอย่างมีนัยสำคัญ

ขอบเขตของการวิจัย

1. แบบสอบที่ใช้ในการวิจัยครั้งนี้เป็นแบบสอบผลสัมฤทธิ์วิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 ที่ผู้วิจัยสร้างขึ้นประกอบด้วยข้อสอบ 120 ข้อ เป็นแบบสอบชนิดเลือกตอบ 4 ตัวเลือก ตามหลักสูตรประถมศึกษาปพุทธศักราช 2521 ของกระทรวงศึกษาธิการ สร้างข้อสอบจากรูปแบบของข้อสอบ (Item Form)

การออกแบบโครงสร้างสารสนเทศของแบบสอบเพื่อจัดข้อสอบให้มีระดับความยากง่ายเหมาะสมกับกลุ่มความสามารถต่างกัน 3 ระดับ คือ กลุ่มความสามารถสูง ความสามารถปานกลาง และความสามารถต่ำ ในการวิจัยครั้งนี้เนื่องจากยังไม่มีคลังข้อสอบที่มีข้อสอบจำนวนมากพอ ผู้วิจัยใช้ข้อสอบ 120 ข้อที่สร้างขึ้นในการจัดแบบสอบจึงมีความจำกัดในเรื่องจำนวนข้อสอบ แบบสอบที่เหมาะสมกับกลุ่มความสามารถมีข้อสอบฉบับละ 40 ข้อ โดยข้อสอบ 40 ข้อนี้จะประกอบด้วยข้อสอบที่เหมาะสมในแต่ละกลุ่มความสามารถที่วิเคราะห์ได้ทั้งหมดเป็นหลัก นอกจากนั้นเลือกข้อสอบจากกลุ่มความสามารถอื่น ๆ โดยมีเงื่อนไขการเลือกให้ครอบคลุมเนื้อหาที่กำหนดไว้ ข้อสอบทั้ง 40 ข้อนำมาหาค่าฟังก์ชันแสดงสารสนเทศของข้อสอบและแบบสอบเพื่อให้ได้โครงสร้างสารสนเทศของแบบสอบที่ใกล้เคียงกับโครงสร้างเป้าหมายที่กำหนดไว้ให้มากที่สุด การพิจารณาโครงสร้างสารสนเทศของแบบสอบนี้พิจารณาค่าความสามารถให้มีช่วงใกล้เคียงช่วงความสามารถ -3.00 ถึง 3.00

2. เมื่อผู้วิจัยจัดข้อสอบตามระดับความสามารถแล้วทำการศึกษาความแม่นยำ (precision) ในการวัดของแบบสอบพิจารณาจากค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถและค่าความเที่ยงของแบบสอบในแต่ละกลุ่มความสามารถของนักเรียน โดยใช้รูปแบบของราสซ์ทำการวิเคราะห์

3. ในการวิจัยครั้งนี้ได้ศึกษาตัวแปรดังต่อไปนี้

3.1 ตัวแปรต้น คือ รูปแบบการจัดแบบสอบที่เหมาะสมกับระดับความสามารถของนักเรียนที่แตกต่างกัน 3 ระดับ คือ ความสามารถสูง ความสามารถปานกลาง และความสามารถต่ำ

3.2 ตัวแปรตาม ได้แก่

3.2.1 ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถ

3.2.2 ค่าความเที่ยงของแบบสอบ

ข้อตกลงเบื้องต้น

1. การแบ่งกลุ่มความสามารถของนักเรียนออกเป็น 3 กลุ่ม จะแบ่งตามมาตรฐาน (scale) ที่ได้จากการวิเคราะห์ด้วยรูปแบบของราสช์ ดังนี้

$\theta > 1$ SD. คือ กลุ่มความสามารถสูง

$\theta \pm 1$ SD. คือ กลุ่มความสามารถปานกลาง

$\theta < 1$ SD. คือ กลุ่มความสามารถต่ำ



2. การวิจัยครั้งนี้ไม่คำนึงถึงความแตกต่างด้านเพศและอายุของกลุ่มตัวอย่าง

คำจำกัดความที่ใช้ในการวิจัย

1. นักเรียน หมายถึง นักเรียนชั้นประถมศึกษาปีที่ 6 ตามหลักสูตรประถมศึกษา ปุทธศักราช 2521 ของกระทรวงศึกษาธิการในภาคเรียนที่ 1 ปีการศึกษา 2529 สังกัดสำนักงานการประถมศึกษาจังหวัดลพบุรี

2. ความสามารถที่แท้จริง หมายถึง ค่าความสามารถของผู้สอบที่ประมาณค่าได้จากคะแนนรวมของการทำข้อสอบจากรูปแบบของราสช์

3. ระดับความสามารถของผู้สอบ หมายถึง ความสามารถระดับสูง ความสามารถปานกลาง และความสามารถต่ำ ซึ่งพิจารณาจากสเกลความสามารถ (θ) ของผู้สอบที่ได้เก็บรวบรวมผลการสอบมาจำนวน 1,429 คน โดยพิจารณาช่วงค่าความเบี่ยงเบนมาตรฐาน เป็นเกณฑ์ในการจำแนกความสามารถ

4. การออกแบบโค้งแสดงสารสนเทศของแบบสอบ หมายถึง การจัดข้อสอบที่ได้จากการวิเคราะห์ด้วยรูปแบบของราสช์ โดยนำค่าความยากของข้อสอบที่ตรงกับระดับความสามารถมาคำนวณหาค่าฟังก์ชันแสดงสารสนเทศของข้อสอบและแบบสอบ (Item and Test information function) เพื่อกำหนดโค้งแสดงสารสนเทศของแบบสอบ (Test - information curve) ตามระดับความสามารถที่ต้องการ

5. รูปแบบของราสช์ (Rasch Model) หมายถึง วิธีการวิเคราะห์ข้อสอบรายข้อและวิธีการวัดคุณลักษณะแฝงของการสอบวิธีหนึ่งที่มีสมมติฐานเบื้องต้นว่า โอกาสที่บุคคลจะทำข้อสอบข้อใดได้ถูกต้องหรือไม่ขึ้นอยู่กับความสามารถของตนเอง (θ) และระดับความยากของข้อสอบ (b) เท่านั้น (Hambleton and Cook 1977: 82) และโอกาสของการตอบถูกจะแสดงด้วยโค้งลักษณะข้อสอบ (Item Characteristic Curve : ICC) ซึ่งเขียนเป็นฟังก์ชันได้ดังนี้

$$P_g(\theta) = \frac{\exp(\theta - b_g)}{1 + \exp(\theta - b_g)} \quad : g = 1, 2, \dots, n$$

$$\exp = 2.71828\dots$$

6. โค้งแสดงสารสนเทศเป้าหมายของแบบสอบ (Target information curve) หมายถึง โค้งของแบบสอบที่ได้จากการนำค่าพารามิเตอร์ของข้อสอบมาคำนวณหาฟังก์ชันแสดงสารสนเทศของข้อสอบและแบบสอบ ตามที่โครงการจะนำไปใช้ในการสอบที่แบ่งกลุ่มความสามารถของผู้สอบเป็น 3 กลุ่ม คือ กลุ่มความสามารถสูง ปานกลาง และต่ำ

7. ความคลาดเคลื่อนมาตรฐานในการประมาณค่า หมายถึง อัตราส่วนผกผันของค่าฟังก์ชันแสดงสารสนเทศของแบบสอบ ณ ระดับความสามารถ(θ) หนึ่ง ๆ ซึ่งสามารถเขียนเป็นสูตรคณิตศาสตร์ได้ดังนี้

$$S.E.E. = \frac{1}{\sqrt{I(\theta)}}$$

8. ความเที่ยง หมายถึง ความคงที่ของคะแนนที่ได้จากแบบสอบในการวัดแต่ละครั้งหรือความสม่ำเสมอของการตอบของผู้สอบต่อข้อสอบทุกข้อ โดยการสอบเพียงครั้งเดียว และสามารถประมาณค่าด้วยสูตรของคูเคอร์-ริชาร์คสัน สูตรที่ 20 ดังนี้


$$r_{tt} = \frac{K}{K-1} \left[\frac{s_t^2 - \Sigma pq}{s_t^2} \right]$$

9. ภาวะสารรูปสนิที่สุด (Goodness of fit) หมายถึง การพิจารณาว่าข้อมูลที่นำมาใช้ในการวิเคราะห์นั้นใช้ได้กับรูปแบบของราสส์หรือไม่เพียงใด ซึ่งการวิจัยครั้งนี้พิจารณาจากสูตร

$$V = \frac{N}{v} \sum_i \frac{L}{i} \sum_{vi} \frac{2}{vi} \frac{1}{(N-1)(L-1)}$$

ประโยชน์ที่คาดว่าจะได้รับจากการวิจัย

1. เป็นแนวทางในการจัดข้อสอบให้มีประสิทธิภาพเหมาะสมกับการสอบวัดกลุ่มนักเรียนที่มีความสามารถแตกต่างกัน
2. เป็นแนวทางในการสร้างธนาคารข้อสอบ (Item Bank) เพื่อเก็บข้อสอบที่มีคุณภาพไว้ใช้
3. เป็นแนวทางในการสร้างแบบสอบเพื่อคัดเลือกข้อสอบให้ตรงกับวัตถุประสงค์ในการสอบวัดแต่ละครั้ง เช่น การสอบคัดเลือกบุคคล การวัดผลสัมฤทธิ์ การจัดเก็บนักเรียนเข้าชั้นเรียน เป็นต้น



ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย