



บทที่ 1

บทนำ

ความเป็นมาและความสำคัญของปัญหา

ในอดีตการวัดและการประเมินผลการทดสอบเป็นการนำคะแนนที่ได้จากการสอบในแต่ละวิชามารวมกัน แล้วแปลงคะแนนทั้งหมดเป็นเปอร์เซ็นต์ ตัดสินการได้ตกจากการประเมินผลเป็นเปอร์เซ็นต์ ผู้ที่ได้เปอร์เซ็นต์สูงสุดคือผู้สอบได้อันดับที่ 1 และผู้ที่ได้เปอร์เซ็นต์รองลงมาคือผู้สอบได้อันดับที่ 2 และที่ 3 ตามลำดับ ซึ่งการประเมินผลในแนวคิดนี้ ไม่อาจจะบอกถึงความสามารถของผู้สอบว่ามีความรอบรู้ในเนื้อหาแต่ละวิชาเพียงใด นอกจากจะแปลผลในลักษณะของภาพรวมว่าผู้สอบเก่งหรืออ่อนกว่าคนในกลุ่มเท่านั้น แม้ว่าในระยะต่อมาจะได้มีการพิจารณาเป็นรายวิชา โดยนำคะแนนที่ได้จากการสอบที่เป็นคะแนนดิบมานับจำนวนข้อกระทงที่ตอบถูก แล้วนำคะแนนมาแปลงเป็นเปอร์เซ็นต์ไทล์หรือเป็นคะแนนมาตรฐาน (เยาวดี วิบูลย์ศรี 2526 : 23) ก็ยังเป็นการจัดอันดับความสามารถของบุคคลว่าสูงกว่า หรือต่ำกว่าในกลุ่ม ไม่อาจจะบอกได้ว่าผู้เรียนมีความรอบรู้ในเนื้อหาเพียงใด หรือถึงขั้นใดที่จะให้ผ่านไปเรียนในเนื้อหาต่อไป อันเป็นจุดอ่อนของการวัดและประเมินผลที่ต้องการรู้ถึงความสามารถของผู้สอบในแต่ละวิชาว่า ผู้สอบมีความสามารถในระดับใดของเนื้อหาที่ต้องการวัด นักการศึกษาจึงได้ศึกษาแนวทางที่จะปรับปรุงพัฒนาการวัดและประเมินผลให้มีประสิทธิภาพในการเรียนการสอน สามารถระบุได้ว่าผู้เรียนมีความสามารถถึงระดับที่ต้องการวัด ซึ่งจะทำให้การวัดและการประเมินผลมีความหมายขึ้น โดยนำคะแนนสอบที่ได้มาเปรียบเทียบกับเกณฑ์ที่กำหนดไว้ และแนวคิดนี้ได้นำมาใช้ในระเบียบการวัดและการประเมินผลการเรียนตามหลักสูตรมัธยมศึกษาตอนปลายพุทธศักราช 2524 จากลักษณะดังกล่าว การวัดและการประเมินผลการศึกษาก็อาจแบ่งเป็น 2 แบบอย่างกว้าง คือ การวัดแบบอิงกลุ่ม และการวัดแบบอิงเกณฑ์

การวัดแบบอิงกลุ่ม (Norm - Referenced Testing) เป็นการนำผลที่ได้จากการวัดไปเปรียบเทียบกับคนอื่น ๆ ในกลุ่มที่ถูกวัดด้วยแบบทดสอบฉบับเดียวกัน การวัดและการประเมินผลแบบนี้จึงเป็นการวัดที่ขึ้นอยู่กับพฤติกรรมของกลุ่มเป็นสำคัญ โดยมีความเชื่อในเรื่องความแตกต่างระหว่างบุคคลที่ว่า แต่ละคนมีความสามารถไม่เท่ากัน ดังนั้นการวัดและการประเมินผลแบบนี้จึงเป็นการหาความแตกต่างระหว่างบุคคลว่า แต่ละคนอยู่ตรงไหนของกลุ่ม และคะแนนที่ได้จะมีความหมายเมื่อนำไปเปรียบเทียบกับคะแนนในกลุ่มที่สอบด้วยแบบสอบฉบับเดียวกัน โดยนำคะแนนที่ได้มาตัดสินในรูปตำแหน่งเปอร์เซ็นต์ไทล์หรือคะแนนมาตรฐานอื่น ๆ จะเห็นได้ว่าการวัดและการประเมินผลแบบนี้เป็นการจัดอันดับความสามารถของผู้เรียนว่า ผู้เรียนมีความรู้ในเรื่องที่เรียนมากหรือน้อยกว่าคนอื่น ๆ ในกลุ่ม ซึ่งผลการสอบมิได้บอกให้เราทราบว่า นักเรียนมีความรอบรู้หรือมีความสามารถที่จะทำงานในด้านต่าง ๆ ตามจุดมุ่งหมายของการฝึกฝนทางการศึกษามากน้อยเพียงใด (Gronlund 1973) ในลักษณะของความเป็นจริงผลการสอบควรจะสามารถบอกได้ว่า ผู้สอบควรจะสามารถแสดงพฤติกรรมและสามารถทำอะไรได้บ้าง ดังนั้นการสอบวัดที่ดีในการเรียนการสอนควรมีข้อมูลที่จะใช้เพื่อการพัฒนาคนไปพร้อม ๆ กันด้วย กล่าวคือเขาได้เรียนรู้มีความสามารถด้านใด อยู่ในระดับใด และเขาจะต้องใช้กลวิธีแบบใดจึงจะเรียนรู้ในสิ่งที่กำหนดเป็นจุดมุ่งหมายได้ ระบบการเรียนการสอนที่ดีจึงควรเป็นระบบสะสมข้อมูลเกี่ยวกับตัวผู้เรียนในด้านต่าง ๆ ไว้สำหรับช่วยให้เขาเรียนรู้ได้ดียิ่งขึ้น เร็วยิ่งขึ้น การสอบจึงต้องคำนึงถึงแนวพัฒนาการของเด็กด้วยว่า เขาใช้แนวใดและเขาอยู่ในแนวระดับใดในแนวทางนั้น ๆ ดังนั้นแบบทดสอบที่ดีจึงควรคำนึงแนวพัฒนาการเหล่านี้ด้วย โดยเมื่อสอบมาแล้วเราก็สามารถบอกได้ทันทีว่าเขาอยู่ในระดับพัฒนาการในช่วงใด และเขาต้องพัฒนาไปอีกอย่างไร ด้วยลักษณะดังกล่าวนี้เท่านั้นที่จะช่วยให้การวัดมีคุณประโยชน์มากที่สุด สมตามเจตนารมณ์ของการพัฒนาในทุก ๆ ด้าน (โกวิท ประวาลพฤษษ์ 2529 : 11) ดังนั้นนักวัดผลจึงได้พยายามพัฒนาการวัดและการประเมินผลเพื่อให้สามารถวัดให้ได้ในลักษณะดังกล่าว ซึ่งการวัดและการประเมินผลในแนวนี้คือ การวัดแบบอิงเกณฑ์

การวัดแบบอิงเกณฑ์ (Criterion - Referenced Testing) การวัดในลักษณะนี้เป็นการวัดที่ไม่ได้นำผลของแต่ละบุคคลไปเปรียบเทียบกับคะแนนของคนอื่นในกลุ่ม แต่เป็นการวัดที่มุ่งไปในลักษณะที่ว่าผู้เรียนสามารถทำอะไรได้บ้าง ไม่สามารถทำอะไรได้บ้าง หลังจากที่มีการ

เรียนการสอนไปแล้ว โดยอาศัยเกณฑ์ ซึ่งเกณฑ์ในที่นี้ได้แก่ จุดมุ่งหมายของการเรียนการสอน ระดับความสามารถที่พึงปรารถนาของผู้เรียน พฤติกรรมหลังการสอน เป็นต้น จะเห็นว่าการวัด และการประเมินผลแบบนี้สามารถนำมาใช้ เพื่อแยกผู้เรียนที่รู้แล้วและที่ยังไม่รู้ได้โดยมีวัตถุประสงค์ เพื่อวินิจฉัยว่า ผู้เรียนบรรลุวัตถุประสงค์ที่วางไว้หรือไม่ ผู้เรียนคนใดที่ยังไม่บรรลุวัตถุประสงค์ ที่วางไว้ก็จะได้หาทางช่วยเหลือเช่น ถ้าผู้เรียนส่วนใหญ่ไม่บรรลุวัตถุประสงค์อาจจะต้องมีการเรียนการสอนวัตถุประสงค์นั้นใหม่ หรือจัดให้มีการสอนซ่อมเสริม เป็นต้น จากที่กล่าวมา การวัดแบบอิงเกณฑ์นั้น จะต้องอาศัยเกณฑ์หรือคะแนนจุดตัด ซึ่งเกลสเซอร์ (Glasser 1963) เป็นผู้หนึ่งที่ทำให้แนวคิดไว้ว่า แบบสอบอิงเกณฑ์จะต้องมีจุดตัด (Cut - off - Score) เพื่อวินิจฉัยว่านักเรียนคนใดเรียนรู้ครบถ้วนตามจุดประสงค์ที่กำหนดไว้ ดังนั้นแบบสอบที่ใช้เป็นเครื่องมือ วัดจะต้องสร้างให้สอดคล้องกับวัตถุประสงค์ และครอบคลุมโดเมนที่ต้องการให้แสดงพฤติกรรม ออกมา ซึ่งปัญหาในการสร้างนั้นสามารถแก้ไขได้โดยใช้เทคนิคการสร้างแบบสอบ ซึ่งมีนักวัดผล ให้แนวคิดในการสร้างไว้หลายคน อาทิเช่น โปพแฮม (Popham 1975) รอยด์และฮาไลดนา (Roid and Haladyna 1980) ออสเบอร์น (Osburn 1968) ไฮวัลลี (Hively 1974) เป็นต้น อันเป็นแนวทางที่ช่วยลดปัญหาการสร้างแบบสอบได้เป็นอย่างดี แต่ปัญหาที่สำคัญ ของการวัดผลแบบอิงเกณฑ์คือ การวัดผลในแนวนี้ขึ้นอยู่กับคะแนนจุดตัดเป็นสำคัญที่จะทำให้แยกผู้ ที่เรียนหรือผู้รู้แล้วออกจากผู้ที่ยังไม่ได้เรียนหรือยังไม่รู้ ซึ่งแฮมเบิลตันและโนวิก (Hambleton and Novick 1973 : 161) ก็ได้กล่าวว่าปัญหาเบื้องต้นของการวัดผลแบบอิงเกณฑ์คือ การ กำหนดคะแนนจุดตัดหรือคะแนนระดับความรอบรู้ (Master Level Score) ขึ้น เพื่อจัดนักเรียนทำข้อสอบในแบบทดสอบอิงเกณฑ์ ให้เป็นพวกรอบรู้ (Master) หรือไม่รอบรู้ (Nonmaster) แต่เกณฑ์ดังกล่าวนี้ควรจะกำหนดเป็นเท่าไรและกำหนดอย่างไร เป็นปัญหาที่ยังไม่อาจหาข้อยุติ ได้แน่นอน ทั้งนี้ส่วนใหญ่มักจะใช้เกณฑ์ของบลูม (Bloom 1976) คือ 80% หรือ 90% ของข้อสอบเป็นหลักผ่าน (ส.วาสนา ประवालนฤกษ์ 2517 : 80 - 85) เพื่อว่าผู้ที่สามารถ ทำคะแนนได้ 80% หรือ 90% ขึ้นไปเป็นผู้ผ่านเกณฑ์ แต่ถ้าต่ำกว่านี้ก็เป็นผู้ไม่ผ่านเกณฑ์ ดังนั้น หากว่าแบบสอบอิงเกณฑ์นี้ยากเกินไปก็อาจจะมีผู้สอบผ่านได้น้อยหรือ ถ้าหากแบบสอบอิงเกณฑ์ ง่ายเกินไปก็อาจมีผู้สอบผ่านหมด และที่สำคัญเกณฑ์นี้ก็อาจทำให้เกิดการจัดประเภทผิด โดย ปฏิเสธผู้ที่รอบรู้เป็นผู้ไม่ผ่านเกณฑ์ และยอมรับผู้ที่ไม่รอบรู้เป็นผู้ผ่านเกณฑ์ หรือที่เรารู้จักกันใน ลักษณะความคลาดเคลื่อนประเภทที่ 1 และประเภทที่ 2 (Type I error and Type II error) ตามลำดับ

อย่างไรก็ตามวิธีดังกล่าวก็ยังมีปัญหา มีข้อบกพร่องในแต่ละวิธี และยังคงใช้สถิติตามแนวทฤษฎีคลาสสิก กล่าวคือ แบบสอบอิงเกณฑ์ที่กลุ่มโดเมนตามลักษณะของทฤษฎีคลาสสิก มีปัญหาเรื่องค่าความยากของแบบสอบในการที่จะทำให้ค่าความยากคงที่ ซึ่งทำให้กำหนดคะแนนจุดตัดได้ไม่แน่นอน เมื่อกลุ่มเปลี่ยนไปหรือสถานการณ์เปลี่ยนไป อีกทั้งค่าความยากและคะแนนโดเมน (Domain Score) อยู่ในมาตราของการวัด (Scale) ต่างกัน ไม่อาจนำมาเทียบกันได้

จากข้อจำกัดดังกล่าวทำให้นักวัดผลบางกลุ่มเห็นว่า สถิติคลาสสิกไม่สามารถนำมาใช้แก้ปัญหาคำตอบได้ในบางแง่มุมปัญหา จึงมีผู้พัฒนาทฤษฎีคุณลักษณะแฝง หรือทฤษฎีการตอบสนองข้อกระทง (Latent Trait Theory หรือ Item Response Theory หรือ Item Characteristic Curve Theory หรือ Modern Test Theory) ขึ้นและสามารถนำมาใช้ในการวิเคราะห์ข้อสอบ เพื่อตรวจหาคุณภาพเป็นรายข้อและทั้งฉบับได้ข้อดีของทฤษฎีการตอบสนองข้อกระทงช่วยแก้ปัญหาที่ทฤษฎีตามแนวคลาสสิกไม่สามารถทำได้โดยที่ถ้าข้อมูลเป็นไปตามข้อตกลงแล้ว โมเดลของทฤษฎีนี้จะมีประโยชน์ในการแก้ปัญหาที่โมเดลแบบทฤษฎีคลาสสิกทำไม่ได้คือ พารามิเตอร์รายข้อกระทงจะไม่แปรเปลี่ยนไปตามกลุ่มตัวอย่างนักเรียน เขตข้อสอบที่อยู่ต่างเขตกันแต่วัดความสามารถหรือคุณลักษณะ (Trait) เดียวกันสามารถนำมาเปรียบเทียบกันได้ หรืออาจจะกล่าวได้ว่าสามารถใช้ข้อสอบคู่ขนานกันและสามารถประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดของผู้สอบแต่ละคนได้ทุกระดับความสามารถของผู้สอบ ซึ่งค่าความคลาดเคลื่อนนี้จะไม่ขึ้นอยู่กับความคลาดเคลื่อนมาตรฐานอื่น ๆ จึงทำให้ผลการวัดความสามารถของผู้สอบแต่ละคนมีความถูกต้องแม่นยำมาก จากการศึกษาของซิลวา (Silva 1984) ได้เปรียบเทียบวิธีการเลือกข้อกระทงของแบบสอบอิงเกณฑ์ โดยใช้วิธีทฤษฎีคลาสสิกกับวิธีของทฤษฎีการตอบสนองข้อกระทงได้ข้อสรุปว่า โมเดลของทฤษฎีการตอบสนองข้อกระทงที่ใช้ 3 พารามิเตอร์สามารถจำแนกผู้สอบเป็นผู้รอบรู้ได้ดีกว่าการเลือกข้อกระทงตามทฤษฎีคลาสสิก ดังนั้นผู้วิจัยจึงได้นำเอาทฤษฎีการตอบสนองข้อกระทงมาใช้ในการคัดเลือกข้อกระทง โดยที่โมเดลของทฤษฎีการตอบสนองข้อกระทงเป็นโมเดลที่แสดงความสัมพันธ์ของผลการสอบกับ ความสามารถในรูปของฟังก์ชัน ซึ่งความสัมพันธ์ในรูปของฟังก์ชันของทฤษฎีการตอบสนองข้อกระทงมีหลายโมเดลได้มีการพัฒนาขึ้นมาหลายแบบ แต่ละโมเดลแตกต่างกันด้วยรูปแบบของฟังก์ชันทางคณิตศาสตร์ที่ใช้และด้วยจำนวนของพารามิเตอร์ที่ใช้อธิบายฟังก์ชันแสดงโค้งลักษณะข้อกระทง (Item Characteristic Function) (ผจงจิต อินทสุวรรณ, 2525 : 57)

ฟังก์ชันที่นิยมใช้กัน คือ ฟังก์ชันโค้งปกติสะสม (Normal Ogive Function) กับ ฟังก์ชันโลจิสติก (Logistic Function) เนื่องจากฟังก์ชันโลจิสติกมีลักษณะการแจกแจงของ ฟังก์ชันใกล้เคียงกับฟังก์ชันโค้งปกติสะสม ความยุ่งยากในการคิดคำนวณน้อยกว่าฟังก์ชันโค้งปกติ สะสม ขณะเดียวกันได้มีการพัฒนาโปรแกรมคอมพิวเตอร์ทำให้สะดวกต่อการประมาณค่าของ พารามิเตอร์ งานวิจัยครั้งนี้ผู้วิจัยได้ใช้ฟังก์ชันโลจิสติกโมเดล 3 พารามิเตอร์ เพราะมีข้อสรุป ว่าโมเดลที่มีจำนวนพารามิเตอร์มากกว่าให้ความเหมาะสมดีกว่า (ผจงจิต อินทสุวรรณ 2525: 63) และเออร์รี่ (Urry 1970; cited by Warm 1978: 20) ได้ให้การสนับสนุนว่า แบบสอบอิงเกณฑ์ชนิดเลือกตอบโมเดลที่เหมาะสม คือ โมเดลที่ใช้พารามิเตอร์ 3 ตัว ซึ่ง สอดคล้องกับเครื่องมือที่ใช้ในการวิจัยครั้งนี้

นอกจากนี้แฮมเบิลตัน สวามินาทาน อัลจิงนา และ คูลสัน (Hambleton, Swaminathan, Algina and Coulson 1978) ใช้แบบสอบอิงเกณฑ์ในเรื่องของการ ประมาณคะแนนโดเมนและการจำแนกผู้สอบ ผู้วิจัยจึงได้ใช้คะแนนโดเมนมาทำการศึกษาเพราะ สามารถสรุปผลความรอบรู้หรือไม่รอบรู้ได้ชัดเจน คะแนนโดเมน (π) จะเป็นค่าที่ใช้คาดคะเน ระดับผลสัมฤทธิ์ทางการเรียนที่ปราศจากอคติ (Unbiased Estimate) และมีความพอเพียง (Sufficient Statistic) ก็ต่อเมื่อมีการระบุชุดของข้อสอบที่เป็นไปได้ในแต่ละโดเมนไว้ ล่วงหน้า และข้อสอบในแบบสอบอิงโดเมนต้องเป็นกลุ่มตัวอย่างแบบสุ่มของข้อสอบในโดเมนนั้น เป็นที่น่าสังเกตว่า คะแนนรวมใช้เป็นค่าคาดคะเนที่ดีได้ ก็ต่อเมื่อข้อสอบแต่ละข้อล้วนวัดในสิ่ง เดียวกัน หรือแบบสอบนั้นมีลักษณะเป็นเอกพันธ์ (Homogeneity) แบบสอบอิงโดเมน (ปัจจุบัน การทดสอบอิงโดเมนมักจะรู้จักกันในชื่อของการทดสอบอิงเกณฑ์) เป็นแบบสอบที่มีคุณสมบัติในเชิง เอกพันธ์สูง ซึ่งสามารถใช้คะแนนรวมเป็นตัวประมาณค่าผลสัมฤทธิ์ทางการเรียน หรือลักษณะ (Trait) ที่มุ่งวัดได้ (สมหวัง นิธิยานูวัฒน์ 2529: 33) ซึ่งค่าพารามิเตอร์ของข้อสอบ หลายข้อที่อยู่ในโดเมนเดียวกัน หรือค่าพารามิเตอร์ของข้อสอบชุดหนึ่ง ที่คิดว่าเป็นตัวแทนที่ดี ของโดเมนจะทำให้สามารถหาความสัมพันธ์ระหว่างคะแนนโดเมน และคะแนนความสามารถ ที่แท้จริง (θ) เนื่องจากคะแนนโดเมน คือ ค่าคาดหวัง (Expected Value) ของคะแนน จากข้อสอบทุกข้อในโดเมน (Hambleton and De Gruijter 1983: 358)

นอกจากนี้ในการให้เกณฑ์ตัดสินผลการเรียนเป็นระดับการเรียน A, B, C, D และ F หรือ 4, 3, 2, 1 และ 0 โดยทั่วไปจะใช้เกณฑ์ที่กรมวิชาการ กระทรวงศึกษาธิการได้เสนอแนวทางหนึ่งในการให้ระดับคะแนน ดังนี้

80% ขึ้นไป	ผลการเรียน คือ A หรือ 4
70% - 79%	ผลการเรียน คือ B หรือ 3
60% - 69%	ผลการเรียน คือ C หรือ 2
50% - 59%	ผลการเรียน คือ D หรือ 1
ต่ำกว่า 50%	ผลการเรียน คือ F หรือ 0

(คู่มือการประเมินผลตามหลักสูตรมัธยมศึกษาตอนปลายพุทธศักราช 2524 : 108) และสถานศึกษาต่าง ๆ จะใช้เกณฑ์ดังกล่าวในการตัดสินผลการเรียน ซึ่งเป็นเกณฑ์ที่ยังไม่เหมาะสมนัก เพราะแบบสอบอาจยากเกินไป ทำให้มีนักเรียนต่ำกว่าเกณฑ์ที่กำหนดมาก ทั้ง ๆ ที่นักเรียนอาจมีความรู้เพียงพอในเนื้อหา แต่เพราะแบบสอบยากเกินไปหรือแบบสอบง่ายเกินไป นักเรียนผ่านเกณฑ์ที่กำหนดไว้มาก ได้ผลการเรียนระดับสูงมาก ทั้ง ๆ ที่นักเรียนยังมีความรู้ไม่เพียงพอถึงระดับเนื้อหานั้นจากคะแนนโดเมนที่กล่าวมาข้างต้น ผู้วิจัยจึงเห็นว่าคะแนนโดเมนจะเป็นสิ่งที่กำหนดได้ชัดเจนและระบุว่านักเรียนมีความรอบรู้ของเนื้อหาในระดับใด ผู้วิจัยจึงกำหนดคะแนนโดเมนที่ต้องการวัดคือ .5, .6, .7 และ .8 เพื่อให้สัมพันธ์กับเกณฑ์ที่ระเบียบประเมินผลในปัจจุบันนี้ใช้อยู่ แต่บอกได้ชัดเจนขึ้นถึงความรอบรู้ในเนื้อหา

ส่วนการกำหนดความยาวแบบสอบนั้นจากผลงานวิจัยของนิศิษฐ์ ตัณฑวนิช (2528: 114) ได้ข้อค้นพบว่า แบบสอบที่นิยามโดเมนเนื้อหาไว้ชัดเจนความยาวแบบสอบ 20 ข้อ มีความยาวที่วัดได้ดีพอ ๆ กับความยาวแบบสอบ 30 ข้อ ดังนั้นผู้วิจัยจึงใช้ความยาว 20 และ 30 ข้อตามลำดับ ในการวิจัยครั้งนี้

จากเหตุผลดังกล่าวผู้วิจัยจึงสนใจศึกษาคะแนนจุดตัดของแบบสอบอิงเกณฑ์ถึงผลของคะแนนจุดตัดที่คำนวณได้จากทฤษฎีการตอบสนองข้อกระทงและตามทฤษฎีคลาสสิกเมื่อกำหนดคะแนนโดเมนและความยาวแบบสอบที่มีต่อค่าความเที่ยงในการตัดสินใจและความน่าจะเป็นในการจำแนกผิดพลาดเพื่อที่จะได้ใช้จำแนกกลุ่มผู้รอบรู้และไม่รอบรู้ในระดับของคะแนนโดเมนและความยาวแบบสอบที่ต้องการวัดได้ชัดเจน นอกจากนี้ยังใช้เป็นแนวทางในการเสนอแนะวิธีการหาคะแนน

จุดตัดเมื่อกำหนดคะแนนโดเมนและความยาวแบบสอบต่างกันตามทฤษฎีการตอบสนองข้อกระทง และทฤษฎีคลาสสิก

วัตถุประสงค์ของการวิจัย

1. เพื่อศึกษาค่าความน่าจะเป็นในการจำแนกผิดพลาดของคะแนนจุดตัดที่คำนวณโดยทฤษฎีการตอบสนองข้อกระทง และทฤษฎีตามแนวคลาสสิกในการจำแนกกลุ่มผู้รอบรู้และไม่รอบรู้ในระดับคะแนนโดเมน .5, .6, .7 และ .8 สำหรับความยาวแบบสอบ 20 ข้อ และคะแนนโดเมน .5, .6, .7 และ .8 สำหรับความยาวแบบสอบ 30 ข้อ ว่าค่าความน่าจะเป็นในการจำแนกผิดพลาดที่คะแนนจุดตัดแนวคิดทฤษฎีใดจะต่ำกว่ากัน

2. เพื่อศึกษาค่าความเที่ยงของคะแนนจุดตัดที่คำนวณโดยทฤษฎีการตอบสนองข้อกระทง และทฤษฎีตามแนวคลาสสิกว่าในระดับคะแนนโดเมน .5, .6, .7 และ .8 สำหรับความยาวแบบสอบ 20 ข้อ และคะแนนโดเมน .5, .6, .7 และ .8 สำหรับความยาวแบบสอบ 30 ข้อว่าค่าความเที่ยงของแนวทฤษฎีใดจะให้ค่าความเที่ยงที่มีแนวโน้มสูงกว่ากัน

สมมุติฐานของการวิจัย

ทฤษฎีตามแนวคลาสสิกมีจุดอ่อนในหลายประเด็น เช่น การตัดสินความสามารถของผู้สอบจากคะแนนที่ได้ (Observed Score) เป็นการตัดสินบนพื้นฐานของความคลาดเคลื่อน (Error) ยังไม่สามารถประมาณค่าความสามารถแท้จริงหรือคะแนนจริงที่ชัดเจนการตัดสินสภาพการรอบรู้ (Master) ในหน่วยการเรียนใด ๆ นั้น ปัจจุบันมีข้อคลาดเคลื่อนต่อการตัดสินและที่สำคัญค่าสถิติของข้อสอบแต่ละข้อ เช่น ค่าความยากง่าย (p) ค่าอำนาจจำแนก (r) เปลี่ยนไปตามกลุ่มของผู้เข้าสอบ ทำให้เกิดปัญหาในการวิเคราะห์ข้อสอบฉบับเดิมแต่เปลี่ยนกลุ่มผู้สอบ (สงบ ลักษณะ 2525: 49-49) อันเป็นจุดอ่อนที่สำคัญของทฤษฎีตามแนวคลาสสิก ซึ่งย่อมต้องมีผลต่อการศึกษาคะแนนจุดตัด ค่าความน่าจะเป็นในการจำแนกผิดพลาด ค่าความเที่ยงในการตัดสินใจ นักวัดผลจึงได้คิดค้นวิธีการแก้ปัญหาเหล่านี้ขึ้น และพบว่า ทฤษฎีการตอบสนองข้อกระทงให้ความกระจ่างในปัญหานี้ได้ ในการวิจัยเพื่อศึกษาคะแนนจุดตัดตามทฤษฎีการตอบสนองข้อกระทง ผู้วิจัยได้เลือกใช้ฟังก์ชันโลจิสติกที่มี 3 พารามิเตอร์ซึ่งเป็นโมเดลที่ให้ความ

เหมาะสมดีกว่า เนื่องจากมีข้อสรุปว่า โมเดลที่มีจำนวนพารามิเตอร์มากกว่าให้ความเหมาะสมดีกว่า (ผจงจิต อินทสุวรรณ 2525: 63) โดยเฉพาะอย่างยิ่ง กับแบบสอบเลือกตอบ นอกจากนี้พารามิเตอร์ของข้อสอบชุดหนึ่งที่เป็นตัวแทนของโดเมน สามารถที่จะหาความสัมพันธ์ระหว่างคะแนนโดเมนกับคะแนนความสามารถที่แท้จริง (θ) ได้ ซึ่งเกี่ยวข้องกับการหาคะแนนจุดตัด ความน่าจะเป็นในการจำแนกผิดพลาด และค่าความเที่ยงในการตัดสินใจ ผู้วิจัยจึงได้ตั้งสมมุติฐานในการวิจัยครั้งนี้ ดังนี้

1. การหาค่าความน่าจะเป็นในการจำแนกผิดพลาดสำหรับความยาวแบบสอบ 20 ข้อ
 - 1.1 เมื่อกำหนดคะแนนโดเมน .5 ความน่าจะเป็นในการจำแนกผิดพลาดของแบบสอบ ๗ คะแนนจุดตัดเพื่อให้ได้กลุ่มผู้รอบรู้และไม่รอบรู้ตามแนวคิดทฤษฎีการตอบสนองข้อกระทงน่าจะต่ำกว่าความน่าจะเป็นในการจำแนกผิดพลาดตามแนวคิดทฤษฎีคลาสสิก
 - 1.2 เมื่อกำหนดคะแนนโดเมน .6 ความน่าจะเป็นในการจำแนกผิดพลาดของแบบสอบ ๗ คะแนนจุดตัดเพื่อให้ได้กลุ่มผู้รอบรู้และไม่รอบรู้ตามแนวคิดทฤษฎีการตอบสนองข้อกระทงน่าจะต่ำกว่าความน่าจะเป็นในการจำแนกผิดพลาดตามแนวคิดทฤษฎีคลาสสิก
 - 1.3 เมื่อกำหนดคะแนนโดเมน .7 ความน่าจะเป็นในการจำแนกผิดพลาดของแบบสอบ ๗ คะแนนจุดตัดเพื่อให้ได้กลุ่มผู้รอบรู้และไม่รอบรู้ตามแนวคิดทฤษฎีการตอบสนองข้อกระทงน่าจะต่ำกว่าความน่าจะเป็นในการจำแนกผิดพลาดตามแนวคิดทฤษฎีคลาสสิก
 - 1.4 เมื่อกำหนดคะแนนโดเมน .8 ความน่าจะเป็นในการจำแนกผิดพลาดของแบบสอบ ๗ คะแนนจุดตัดเพื่อให้ได้กลุ่มผู้รอบรู้และไม่รอบรู้ตามแนวคิดทฤษฎีการตอบสนองข้อกระทงน่าจะต่ำกว่าความน่าจะเป็นในการจำแนกผิดพลาดตามแนวคิดทฤษฎีคลาสสิก
2. การหาค่าความน่าจะเป็นในการจำแนกผิดพลาดสำหรับความยาวแบบสอบ 30 ข้อ
 - 2.1 เมื่อกำหนดคะแนนโดเมน .5 ความน่าจะเป็นในการจำแนกผิดพลาดของแบบสอบ ๗ คะแนนจุดตัดเพื่อให้ได้กลุ่มผู้รอบรู้และไม่รอบรู้ตามแนวคิดทฤษฎีการตอบสนองข้อกระทงน่าจะต่ำกว่าความน่าจะเป็นในการจำแนกผิดพลาดตามแนวคิดทฤษฎีคลาสสิก
 - 2.2 เมื่อกำหนดคะแนนโดเมน .6 ความน่าจะเป็นในการจำแนกผิดพลาดของแบบสอบ ๗ คะแนนจุดตัดเพื่อให้ได้กลุ่มผู้รอบรู้และไม่รอบรู้ตามแนวคิดทฤษฎีการตอบสนอง

ข้อกระทงน่าจะต่ำกว่าความน่าจะเป็นในการจำแนกผิดพลาดตามแนวคิดทฤษฎีคลาสสิก

2.3 เมื่อกำหนดคะแนนโดเมน .7 ความน่าจะเป็นในการจำแนกผิดพลาดของแบบสอบ ๗ คะแนนจุดตัดเพื่อให้ได้กลุ่มผู้รอบรู้และไม่รอบรู้ตามแนวคิดทฤษฎีการตอบสนองข้อกระทงน่าจะต่ำกว่าความน่าจะเป็นในการจำแนกผิดพลาดตามแนวคิดทฤษฎีคลาสสิก

2.4 เมื่อกำหนดคะแนนโดเมน .8 ความน่าจะเป็นในการจำแนกผิดพลาดของแบบสอบ ๗ คะแนนจุดตัดเพื่อให้ได้กลุ่มผู้รอบรู้และไม่รอบรู้ตามแนวคิดทฤษฎีการตอบสนองข้อกระทงน่าจะต่ำกว่าความน่าจะเป็นในการจำแนกผิดพลาดตามแนวคิดทฤษฎีคลาสสิก

สำหรับค่าความเที่ยงตามแนวคิดทฤษฎีการตอบสนองข้อกระทงปกติแบบสอบที่จะจำแนกความรอบรู้โดยใช้แนวคิดทฤษฎีดังกล่าวนี้คือ แบบสอบข้อที่ให้โค้งแสดงสารสนเทศของข้อกระทง (Item Characteristic Curve) สูง ๆ ใกล้เคียง ๆ หรือเท่ากับค่าความสามารถที่ต้องการวัด เพื่อให้สัมพันธ์กับค่าความเที่ยงตามแนวคิดทฤษฎีคลาสสิก (CCT) วอร์ม (Warm 1978: 77) จึงได้ทำการประยุกต์จากโค้งแสดงสารสนเทศของข้อกระทงมาหาค่าความเที่ยง ซึ่งผู้วิจัยได้แนวทางการประยุกต์ของวอร์มมาใช้หาค่าความเที่ยงในการวิจัยตามแนวคิดทฤษฎีการตอบสนองข้อกระทงในครั้งนี้ ส่วนค่าความเที่ยงตามแนวคิดทฤษฎีคลาสสิกนั้นเนื่องจากการวิจัยครั้งนี้ใช้ การทดสอบครั้งเดียว การหาค่าความเที่ยงตามแนวคิดทฤษฎีคลาสสิกที่ทำการสอบครั้งเดียวมีหลายวิธี แต่วิธีที่เหมาะสมและสามารถหาค่าความเที่ยงได้ดี คือ วิธีของฮวีน (Subkoviak 1976) ผู้วิจัยจึงได้ใช้วิธีของฮวีนทำการวิจัยและตั้งสมมุติฐานในการหาค่าความเที่ยงในการวิจัย ดังนี้

3. การหาค่าความเที่ยงสำหรับความยาวแบบสอบ 20 ข้อ

3.1 เมื่อกำหนดคะแนนโดเมน .5 ค่าความเที่ยงในการตัดสินใจที่คะแนนจุดตัดของแนวคิดทฤษฎีการตอบสนองข้อกระทงน่าจะสูงกว่าความเที่ยงที่คะแนนจุดตัดตามทฤษฎีคลาสสิก

3.2 เมื่อกำหนดคะแนนโดเมน .6 ค่าความเที่ยงในการตัดสินใจที่คะแนนจุดตัดของแนวคิดทฤษฎีการตอบสนองข้อกระทงน่าจะสูงกว่าความเที่ยงที่คะแนนจุดตัดตามทฤษฎีคลาสสิก

3.3 เมื่อกำหนดคะแนนโดเมน .7 ค่าความเที่ยงในการตัดสินใจที่คะแนนจุดตัดของแนวคิดทฤษฎีการตอบสนองข้อกระทงน่าจะสูงกว่าความเที่ยงที่คะแนนจุดตัดตามทฤษฎีคลาสสิก

3.4 เมื่อกำหนดคะแนนโดเมน .8 ค่าความเที่ยงในการตัดสินใจที่คะแนนจุดตัดของแนวคิดทฤษฎีการตอบสนองข้อกระทงน่าจะสูงกว่าความเที่ยงที่คะแนนจุดตัดตามทฤษฎีคลาสสิก

4. การหาค่าความเที่ยงสำหรับความยาวแบบสอบ 30 ข้อ

- 4.1 เมื่อกำหนดคะแนนโดเมน .5 ค่าความเที่ยงในการตัดสินใจที่คะแนนจุดตัดของแนวคิดทฤษฎีการตอบสนองข้อกระทงน่าจะสูงกว่าความเที่ยงที่คะแนนจุดตัดตามทฤษฎีคลาสสิก
- 4.2 เมื่อกำหนดคะแนนโดเมน .6 ค่าความเที่ยงในการตัดสินใจที่คะแนนจุดตัดของแนวคิดทฤษฎีการตอบสนองข้อกระทงน่าจะสูงกว่าความเที่ยงที่คะแนนจุดตัดตามทฤษฎีคลาสสิก
- 4.3 เมื่อกำหนดคะแนนโดเมน .7 ค่าความเที่ยงในการตัดสินใจที่คะแนนจุดตัดของแนวคิดทฤษฎีการตอบสนองข้อกระทงน่าจะสูงกว่าความเที่ยงที่คะแนนจุดตัดตามทฤษฎีคลาสสิก
- 4.4 เมื่อกำหนดคะแนนโดเมน .8 ค่าความเที่ยงในการตัดสินใจที่คะแนนจุดตัดของแนวคิดทฤษฎีการตอบสนองข้อกระทงน่าจะสูงกว่าความเที่ยงที่คะแนนจุดตัดตามทฤษฎีคลาสสิก

ขอบเขตของการวิจัย

1. ประชากรที่ใช้ในการวิจัยครั้งนี้เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 ภาคเรียนที่ 1 ปีการศึกษา 2530 ในเขตกรุงเทพมหานคร สังกัดกรมสามัญศึกษากระทรวงศึกษาธิการ
2. ตัวแปรที่ศึกษา มีดังนี้
 - 2.1 ตัวแปรอิสระ ได้แก่
 - 2.1.1 วิธีหาคะแนนจุดตัด ซึ่งมีวิธีการหา 2 วิธี คือ
 - 2.1.1.1 วิธีการหาคะแนนจุดตัดโดยใช้ทฤษฎีการตอบสนองข้อกระทง ตามวิธีประยุกต์ของวิลคอกซ์ (Hambleton and De Gruijter 1983 : 361)
 - 2.1.1.2 วิธีการหาคะแนนจุดตัดโดยใช้ทฤษฎีตามแนวคลาสสิกใช้วิธีหาลัดส่วนความถูกต้อง (Hambleton and other 1978: 5)
 - 2.1.2 คะแนนโดเมน .5, .6, .7 และ .8
 - 2.1.3 ความยาวของแบบสอบ 20 และ 30 ข้อ

2.2 ตัวแปรตาม ได้แก่

2.2.1 ความน่าจะเป็นในการจำแนกความรอบรู้ของคะแนนจุดตัดที่คำนวณจากทฤษฎีการตอบสนองข้อกระทงและทฤษฎีคลาสสิก

2.2.2 ความเที่ยงของคะแนนจุดตัดที่คำนวณจากทฤษฎีการตอบสนองข้อกระทงและทฤษฎีคลาสสิก

3. เนื้อหาที่ใช้ในการศึกษา เป็นเนื้อหาวิชาคณิตศาสตร์ ค 311 เรื่องพหุนามเนื้อหาการคูณหารพหุนาม ซึ่งเป็นเนื้อหาตามหลักสูตรวิชาคณิตศาสตร์ชั้นมัธยมศึกษาปีที่ 3 ของกรมวิชาการ กระทรวงศึกษาธิการ

4. การศึกษาความแม่นยำ (Precision) ของคะแนนจุดตัด พิจารณาจากความน่าจะเป็นในการจำแนกผิดพลาดในการจำแนกผู้สอบออกเป็น 2 กลุ่ม คือ กลุ่มผู้รอบรู้และไม่รอบรู้ ใช้วิธีการหาตามรูปแบบของไบโนเมียลโมเดล (Binomial Model)

ข้อตกลงเบื้องต้น

1. นักเรียนทุกคนทำแบบสอบด้วยความสามารถและจริงใจ เนื่องจากคะแนนที่ได้จากการสอบบางโรงเรียนนำมาเป็นคะแนนระหว่างภาคเรียนของนักเรียน บางโรงเรียนใช้คะแนนที่ได้จากการสอบเพื่อวินิจฉัยนักเรียน นอกจากนี้ผู้วิจัยได้ชี้แจงให้นักเรียนทราบถึงวัตถุประสงค์ของการสอบ ประโยชน์ที่นักเรียนจะได้ทั้งทางตรงและทางอ้อม ทั้งยังเป็นการทดสอบความสามารถของนักเรียนกับนักเรียนโรงเรียนอื่นที่เป็นกลุ่มตัวอย่างในการสอบด้วย

2. วันเวลาที่ต่างกันในการสอบของแต่ละโรงเรียน และห้องเรียนไม่มีผลต่องานวิจัยนี้ ด้วยเหตุผลดังข้อ 1 และระหว่างการสอบทุกห้องปราศจากสิ่งรบกวนที่อาจจะดึงดูดความสนใจของนักเรียน

คำจำกัดความที่ใช้ในการวิจัย

แบบสอบอิงเกณฑ์ หมายถึง แบบสอบอิงโดเมนที่ใช้วัดผลสัมฤทธิ์วิชาคณิตศาสตร์ ค 311 ชั้นมัธยมศึกษาปีที่ 3 ซึ่งเป็นแบบสอบชนิดเลือกตอบ 5 ตัวเลือก สร้างขึ้นตามลักษณะเฉพาะของข้อสอบ ซึ่งมีลักษณะดังนี้

1. วัดผลสัมฤทธิ์ของเนื้อหาและพฤติกรรมที่กำหนดขึ้นในขอบเขต (Domain) ที่ชัดเจน โดยคำนึงถึงความเป็นเอกพันธ์ของเนื้อหา ลักษณะพฤติกรรมและความยาวที่เหมาะสม
2. เป็นตัวแทนของประชากรข้อสอบทั้งหมดที่วัดตามเนื้อหาและพฤติกรรมที่กำหนด
3. การสร้างแบบสอบสร้างขึ้นตามลักษณะเฉพาะของข้อสอบ (Item Specification) การกำหนดลักษณะเฉพาะของข้อสอบ เปรียบเสมือนการกำหนดขอบเขตประชากรข้อสอบ และข้อสอบที่สร้างขึ้นตามลักษณะเฉพาะของข้อสอบ เสมือนกับเป็นตัวแทนของข้อสอบที่สุ่มมาจากประชากรข้อสอบ (Hively 1974: 5)

ลักษณะเฉพาะของข้อสอบ หมายถึง แบบฟอร์มของการกำหนดระเบียบหรือกฎเกณฑ์สำหรับเขียนข้อสอบให้รัดกุมรอบคอบชัดเจน สมบูรณ์ด้วยคุณภาพต่าง ๆ เช่น มีเหตุผลเป็นปรนัย วัดในสิ่งที่ต้องการวัด

นักเรียน หมายถึง นักเรียนชั้นมัธยมศึกษาปีที่ 3 สังกัดกรมสามัญศึกษา ในเขตกรุงเทพมหานคร ที่เรียนวิชาคณิตศาสตร์ ค 311 จำนวน 2 หน่วยการเรียนรู้ ภาคเรียนที่ 1 ปีการศึกษา 2530

คะแนนโดเมน หมายถึง สัดส่วนของจำนวนข้อที่ตอบถูกต้องจำนวนข้อสอบทั้งหมด เช่น 1.00 หมายถึง ผู้สอบตอบแบบสอบถูก 100% จะอ้างอิงไปสู่ประมาณ 100% ของปริเซตของการสอบวัดนั้น การคำนวณค่าสถิติคะแนนโดเมนในแนวของทฤษฎีคลาสสิกคำนวณได้จากเปอร์เซ็นต์การตอบถูกหรือสัดส่วนการตอบถูก ส่วนการคำนวณค่าคะแนนโดเมนในเชิงทฤษฎีการตอบสนองข้อกระทง คำนวณจากเปอร์เซ็นต์ของผลรวมของโอกาสตอบถูกรายข้อของข้อสอบในแบบสอบในแต่ละค่าความสามารถ θ

ค่าความยากของข้อกระทง (b-parameter or Difficulty Power) หมายถึงระดับความสามารถของข้อทดสอบที่อยู่ตรงข้ามกับโค้งที่ชันที่สุดของโค้งแสดงลักษณะของข้อกระทง (Item Characteristic Curve : ICC) และเป็นจุดที่ผู้ทำข้อสอบที่มีความสามารถ (+) เท่ากับความยากง่าย (b) ของข้อทดสอบนั้นมีโอกาสที่จะทำข้อสอบข้อนั้น ๆ ได้ถูกต้อง ปกติแล้วข้อทดสอบจะมีค่า b ประมาณ 2.5 ข้อทดสอบที่ $b = -2.5$ แสดงว่าง่ายมากและข้อสอบที่มีค่า $b = 2.5$ แสดงว่าข้อสอบยากมาก แต่จากการวิเคราะห์ข้อสอบที่ง่าย ค่า b นี้มักจะมีค่า -3.0 ข้อทดสอบที่ $b = 0.00$ แสดงว่ายากง่ายปานกลาง แต่การเลือกข้อทดสอบมาใช้ควรเลือกข้อที่ค่า $|b| < 2.0$

ค่าอำนาจจำแนก (a-parameter or Discrimination Power) หมายถึงจุดบนเส้น ICC ณ ที่โค้งมีความชันมากที่สุด ซึ่งอยู่ตรงกันข้ามกับค่าความยากง่าย จุดนี้เป็นจุดแสดงความสามารถในการจำแนกผู้สอบที่มีความสามารถต่ำและสูงออกจากกันได้ อย่างเด่นชัด ค่าอำนาจจำแนกที่น้อยกว่า ในทางทฤษฎี ค่า a นี้มีค่าระหว่าง $\pm \alpha$ แต่ในทางปฏิบัติค่า a นี้นิยมใช้ระหว่าง 0.5 - 2.5 (Ree 1979: 372)

ค่าการเดา (c-parameter or pseudo - guessing parameter) หมายถึงค่าโอกาสที่ผู้สอบข้อทดสอบที่มีความสามารถต่ำ จะสามารถเดาข้อทดสอบนั้น ๆ ได้ถูกต้องมากน้อยเพียงใด โดยมากค่า c จะมีค่าระหว่าง 0.0 ถึง 0.40 ข้อทดสอบมีค่า c มากกว่า .30 ขึ้นไปจะเป็นข้อสอบที่ไม่ดี ดังนั้นข้อทดสอบควรมีค่า c น้อยกว่า .30 (Urry 1977)

ค่าความสามารถของผู้สอบ (θ) หมายถึง ความสามารถ ทักษะหรือศักยภาพของผู้สอบข้อทดสอบแต่ละข้อว่ามีโอกาสที่จะทำข้อทดสอบข้อนั้นถูกต้องมากหรือน้อยเท่าใด ในทางปฏิบัติจะมีค่าอยู่ระหว่าง -3 ถึง +3 ค่า θ ค่อนข้างมากทาง -3 แสดงว่าความสามารถต่ำ แต่ถ้าค่า $\theta = +3$ หรือค่าค่อนข้างมากทาง +3 แสดงว่าความสามารถสูง (สำ เร็จ บุญเรืองรัตน์ 2527: 104)

ฟังก์ชันการตอบสนองข้อกระทง หมายถึง ฟังก์ชันที่แสดงความน่าจะเป็นหรือ $P(\theta)$ ของการตอบข้อนั้น ๆ ถูก เมื่อกำหนด θ เขียนแทนด้วยฟังก์ชันโลจิสติกที่มี 3 พารามิเตอร์

ฟังก์ชันแสดงสารสนเทศข้อกระทง (Item Information Function) หมายถึง ฟังก์ชันที่ใช้จำแนกความสามารถเป็นรายข้อที่ใกล้จุดตัดได้ถูกต้อง เขียนแทนด้วยสัญลักษณ์ $I(\theta, u_i)$

ฟังก์ชันแสดงสารสนเทศของแบบสอบ (Test Information Function) หมายถึง ฟังก์ชันที่ใช้ตรวจสอบความถูกต้องของแบบสอบที่ θ_0 ได้เหมาะสม เขียนแทนด้วยสัญลักษณ์ $I(\theta)$

จุดประสงค์ หมายถึง จุดประสงค์ที่กำหนดขึ้นในการสอนซึ่งเขียนในลักษณะที่บอกการกระทำหรือพฤติกรรมของนักเรียน เมื่อสิ้นสุดการเรียนรู้เนื้อหาการคูณหารพหุนาม แล้วนักเรียนสามารถหาผลลัพท์การคูณหารพหุนามได้

ผลสัมฤทธิ์ทางการเรียน หมายถึง ผลที่ได้จากการเรียนเมื่อสิ้นสุดการเรียนรู้ ซึ่งเป็นคะแนนที่ได้จากการสอบแบบสอบอิงเกณฑ์วิชาคณิตศาสตร์ ค 311 ชั้นมัธยมศึกษาปีที่ 3 ภาคเรียนที่ 1 ปีการศึกษา 2530

คะแนนจุดตัด หรือเกณฑ์ หมายถึง การกำหนดคะแนนขั้นต่ำสุด (Minimum Requirement) ที่จะตัดสินว่าเป็นผู้รอบรู้และไม่รอบรู้ในระดับคะแนนโดเมนและความยาวของแบบสอบที่ต้องการ

วิธีการหาคะแนนจุดตัด หมายถึง ขบวนการที่คำนวณหาคะแนนที่แบ่งนักเรียนออกเป็นกลุ่มผู้รอบรู้และไม่รอบรู้ ในการวิจัยครั้งนี้ได้ใช้วิธีการหาคะแนนจุดตัดประยุกต์ใช้ทฤษฎีการตอบสนองข้อกระทงตามวิธีของวิลคอกซ์ (Wilcox 1976) และวิธีตามแนวคลาสสิก โดยใช้วิธีหาสัดส่วนความถูกต้อง (Hambleton et al. 1978: 5) เทียบกับคะแนนสอบ

ความเที่ยง หมายถึง ความคงที่ของคะแนนที่ได้จากแบบสอบ ในการวัดที่คะแนนจุดตัดที่คำนวณหาได้ ในทางปฏิบัติของทฤษฎีการตอบสนองข้อกระทง จะพิจารณาจากค่าของฟังก์ชันการแจกแจงของแบบสอบ สำหรับในการวิจัยครั้งนี้ใช้วิธีการประยุกต์ที่วอร์มเสนอไว้ (Warm 1978: 77) ส่วนการหาค่าความเที่ยงตามแนวทฤษฎีคลาสสิก ผู้วิจัยใช้วิธีของฮวินจากสัมประสิทธิ์แคปป่า (Huynh 1976: 258-259)

ความน่าจะเป็นในการจำแนกผิดพลาดสำหรับการแบ่งกลุ่มเป็นผู้รอบรู้และไม่รอบรู้ เมื่อคะแนนโดเมนและความยาวแบบสอบต่างกัน หมายถึง ค่าที่ใช้ตัดสินว่าวิธีคำนวณหาคะแนนจุดตัดโดยใช้ทฤษฎีการตอบสนองข้อกระทง และทฤษฎีคลาสสิก ทฤษฎีใดจะจำแนกได้ดีกว่ากัน โดยใช้รูปแบบของไบโนเมียลโมเดล เขียนแทนด้วยสัญลักษณ์ p

ประโยชน์ที่คาดว่าจะได้รับจากการวิจัย

1. เป็นแนวทางที่จะใช้คะแนนจุดตัดจำแนกนักเรียนเป็นผู้ที่รอบรู้หรือไม่รอบรู้ โดยมีความน่าจะเป็นในการจำแนกผิดพลาดน้อยที่สุด
2. เป็นแนวทางที่จะใช้ข้อสอบในธนาคารข้อสอบ (Item bank) ว่าควรจะใช้คะแนนจุดตัดเท่าใด เมื่อกำหนดความยาวของแบบสอบและกำหนดคะแนนโดเมนต่างกัน
3. เป็นแนวทางให้หาคะแนนจุดตัดด้วยวิธีทฤษฎีการตอบสนองข้อกระทง กับแบบสอบอิงเกณฑ์ที่มีลักษณะแตกต่างจากลักษณะที่ศึกษาในเรื่องอื่น ๆ และระดับอื่น ๆ อีกต่อไป
4. เป็นแนวทางให้หาคะแนนจุดตัดด้วยวิธีทฤษฎีการตอบสนองข้อกระทง โดยเปรียบเทียบกับวิธีอื่น ๆ
5. เป็นแนวทางสำหรับการคัดเลือกข้อกระทงที่เหมาะสม กับความสามารถของนักเรียน