

บทที่ 1

บทนำ



ความเป็นมาและความสำคัญของปัญหา

การศึกษาทุกระดับชั้นมักประกอบด้วยองค์ประกอบ 3 ประการ คือ จุดมุ่งหมายทางการศึกษา การจัดกระบวนการเรียนการสอน การวัดและประเมินผล องค์ประกอบดังกล่าวมีความสัมพันธ์กันจะขาดสิ่งหนึ่งสิ่งใดมิได้ การวัดและประเมินผลเป็นกิจกรรมสำคัญที่จะช่วยตรวจสอบความก้าวหน้าของผู้เรียน และประสิทธิภาพของการจัดกระบวนการเรียนการสอนว่าบรรลุผลตามจุดมุ่งหมายของการจัดการศึกษาหรือไม่ จึงเป็นองค์ประกอบที่สำคัญยิ่งประการหนึ่งในการจัดการศึกษา และการพัฒนาการศึกษาให้มีคุณภาพยิ่งขึ้น

การวัดผลสัมฤทธิ์ทางการเรียนนั้น การทดสอบ นับว่าเป็นวิธีที่ได้รับความนิยมมากที่สุด ในระบบการจัดการศึกษาทุกระดับชั้น เครื่องมือที่ใช้ในการทดสอบส่วนใหญ่ ได้แก่ แบบสอบซึ่งมีอยู่ 2 ประเภทใหญ่ ๆ ด้วยกัน คือ แบบสอบแบบอัตนัย (subjective type) และแบบสอบแบบปรนัย (objective type) (อุทุมพร จามรمان, 2532 : 48) แต่เนื่องจากแบบสอบแบบอัตนัยนี้มีข้อจำกัดหลายประการในการนำไปใช้ เช่น ผู้ตอบต้องใช้เวลามากในการตอบข้อสอบแต่ละข้อ ทำให้การสอบทำได้ไม่ครอบคลุมสิ่งที่ต้องการวัด การตอบคำถามแต่ละข้อ ผู้ตอบต้องใช้ความสามารถหลายประการ จนไม่อาจจะระบุได้อย่างชัดเจนว่าข้อสอบข้อหนึ่ง ๆ นั้นวัดความสามารถใด การตรวจให้คะแนนมีความเป็นปรนัยน้อย มีผลให้การสอบขาดความเที่ยงและความตรง นอกจากนั้นการตรวจให้คะแนนต้องใช้เวลาในการตรวจมาก จึงทำให้เกิดปัญหาในเรื่องการรายงานผลโดยเฉพาะในกรณีที่มีผู้สอบเป็นจำนวนมาก ในปัจจุบันสถานศึกษาส่วนใหญ่จึงหันมานิยมใช้แบบสอบแบบปรนัยกันอย่างแพร่หลาย เนื่องจากมีข้อดีหลายประการ เช่น คำถามมีความเฉพาะเจาะจงมากขึ้น การตัดสินความถูกต้องของคำตอบทำได้ง่าย สามารถระบุได้ค่อนข้างแน่ชัดว่า การตอบข้อสอบแต่ละข้อนั้นผู้ตอบจะต้องใช้ความรู้ความสามารถด้านใด การสอบวัดทำได้ครอบคลุมเนื้อหา หรือพฤติกรรมที่ต้องการวัดมากขึ้น การตรวจให้คะแนนมีความเป็นปรนัยสูงทำได้รวดเร็ว การรายงานผลการสอบจึงทันกับความต้องการ แม้จะมีผู้เข้าสอบเป็นจำนวนมากก็ตาม

แบบสอบเลือกตอบ (multiple choice) เป็นแบบสอบปรนัยที่ได้รับความนิยมมากกว่าแบบสอบปรนัยรูปแบบอื่น ๆ เพราะสามารถใช้วัดจุดมุ่งหมายในการสอนได้ทุก ๆ ระดับของปริศนาคำตอบ (Cognitive Domain) ตั้งแต่ความรู้ ความเข้าใจ การนำไปใช้ การวิเคราะห์ การสังเคราะห์ และการประเมินค่า (Marshall; Jon C. and Hales; Loyde W., 1972: 46) นอกจากนี้ทั้ง สวีส์ ประทุมราช (2531:64) สำเร้ง บุญเรืองรัตน์ (2527:35) และ อีเบล (Ebel, 1965 : 60) ยังมีความเห็นสอดคล้องกันว่า แบบสอบเลือกตอบเป็นแบบสอบแบบปรนัยที่ดีที่สุด ดังนั้น ความสนใจในเรื่องการพัฒนาแบบสอบเลือกตอบเพื่อใช้เป็นเครื่องมือวัดและประเมินผลการศึกษาจึงเกิดขึ้นอย่างกว้างขวาง แต่การจะสร้างแบบสอบเลือกตอบให้มีคุณภาพดีนั้น ขึ้นอยู่กับการใช้คำถาม และตัวเลือกเป็นสำคัญ

แบบสอบเลือกตอบที่ครูใช้ส่วนใหญ่ นั้น นิยมให้นักเรียนพิจารณาคำตอบที่ถูกต้องเพียงคำตอบเดียว ทั้งที่บางเนื้อหาหรือพฤติกรรมที่ต้องการวัดนั้นอาจมีคำตอบได้หลายคำตอบ และเป็นคำตอบที่ถูกต้องพอ ๆ กัน การให้นักเรียนเลือกคำตอบที่ถูกต้องเพียงคำตอบเดียวนั้น หากเขาเลือกได้ถูกต้องก็ไม่ได้หมายความว่า เขาจะทราบคำตอบอื่น ๆ ซึ่งเราไม่ได้นำมาเป็นตัวเลือก หากต้องการทราบว่าเขารู้คำตอบทั้งหมดหรือไม่ ก็คงจะต้องตั้งคำถามเดิมซ้ำอีก และใช้คำตอบอื่น ๆ ที่เป็นไปได้มาเป็นผู้เลือกที่ถูกต้องใหม่ ซึ่งเป็นการถามคำถามซ้ำ ทำให้ข้อสอบมีจำนวนข้อมากขึ้นโดยไม่จำเป็น ปัญหาการมีคำตอบถูกได้หลายคำตอบ และเราต้องการทราบว่าผู้ตอบรู้คำตอบที่ถูกต้องทั้งหมดของปัญหานั้นหรือไม่ มีอยู่เสมอนในหลาย ๆ วิชา ในกรณีเช่นนี้ข้อสอบเลือกตอบตัวเลือกธรรมดาที่ใช้กันอยู่อาจทำหน้าที่ได้ไม่ดีเท่าที่ควร นอกจากนี้การเขียนตัวเลือกให้มีคำตอบถูกเพียงคำตอบเดียว โดยเขียนตัวลวงให้เหมาะสมมีแรงดึงดูดให้นักเรียนที่ไม่มีความรู้จริงเลือกตอบ และตัวลวงทุกตัวมีทำที่ว่าจะเป็นตัวถูกแต่แท้จริงผิดได้ทุกตัวนั้นเป็นสิ่งที่ทำได้ยาก จากสภาพปัญหาดังกล่าวทำให้นักวัดผลพยายามที่จะหาวิธีการเขียนข้อสอบเลือกตอบที่แตกต่างกันออกไป เช่น การเขียนตัวเลือกให้มีคำตอบถูกหลายตัว เพราะการสร้างตัวเลือกถูกหลาย ๆ ตัวนั้น บางครั้งทำได้ง่ายกว่าการสร้างตัวเลือกผิดที่มีลักษณะใกล้เคียงกับตัวเลือกถูกหลาย ๆ ตัว (ลีซล สังก์ศิริ, 2527 : 40) แต่เนื่องจากแบบสอบที่มีคำตอบถูกหลายตัวนั้นมักจะประสบปัญหาเกี่ยวกับการตรวจให้คะแนน คือ ให้คะแนนได้ไม่คงที่ในแต่ละข้อ การตรวจให้คะแนนจึงยุ่งยากกว่าการตรวจแบบสอบเลือกตอบที่มีคำตอบถูกตัวเดียว ดังนั้นจึงได้มีการพัฒนารูปแบบของแบบสอบเลือกตอบที่มีคำตอบถูกไม่จำกัดให้มีลักษณะต่างออกไป เพื่อแก้ปัญหาเรื่อง การตรวจให้คะแนน แบบสอบที่คิดแปลงนี้ เป็นลักษณะของแบบสอบเลือกตอบชนิดหนึ่ง ที่ครูนิยม

นำมาใช้ในหลาย ๆ วิชา ก็คือ แบบสอบเลือกตอบแบบตัวเลือกซ้อน (double or complex multiple choice) ซึ่งคำถามประเภทนี้ประกอบด้วย ตัวตั้ง หรือตอนนำ ที่มีตัวเลือกหลายตัว ในตัวเลือกเหล่านี้อาจมีคำตอบถูกหลายตัว นำตัวเลือกที่ให้ไว้มาสร้างเป็นตัวเลือกอีกต่อหนึ่ง โดยผสมตัวเลือกเดิมออกเป็นชุด ๆ เป็นตัวเลือกใหม่ (เอนก พ.อนุกุลบุตร, 2522 : 228) การรวมคำตอบที่ถูกต้องมาอยู่ในข้อเดียวกัน จะทำให้สะดวกในการตอบและการตรวจมากกว่า ให้นักเรียนตอบตัวเลือกถูกทุก ๆ ตัวลงในกระดาษคำตอบ และยังทำให้เหมือนกับที่เราได้วัดเนื้อหา หรือพฤติกรรมเดียวกันซ้ำหลาย ๆ ครั้ง โดยไม่ต้องเพิ่มจำนวนข้อสอบ จึงไม่ทำให้ผู้ตอบ เกิดความรู้สึกเบื่อหน่ายที่จะต้องทำข้อสอบคล้าย ๆ กันซ้ำหลายครั้งอีกด้วย การตอบข้อสอบ เลือกตอบตัวเลือกซ้อนนั้น ผู้ตอบจะต้องรู้ว่าสถานการณ์ใดถูกหรือผิดทุกสถานการณ์ จึงจะสามารถ เลือกคำตอบที่ถูกต้องได้อย่างถูกต้อง ลักษณะเช่นนี้จึงน่าจะทำให้มั่นใจได้มากขึ้นว่า เราวัดความรู้ ความสามารถของนักเรียนได้ถูกต้องแม่นยำมากขึ้น จึงน่าจะช่วยให้ปัญหาเรื่องการเดาเด้น้อยลง ในแง่การเขียนข้อสอบ ก็น่าจะทำให้ตัวเลือกที่มีลักษณะผิดหรือถูกชัดเจนลดลงอีกด้วย นอกจากนี้ การตรวจให้คะแนน ก็สามารถทำได้สะดวกเหมือนกับการตรวจข้อสอบเลือกตอบตัวเลือกธรรมดา ที่ใช้กันอยู่

เมื่อพิจารณาถึงการรวมตัวเลือกแล้ว การเขียนตัวเลือก "ถูกทั้งข้อ ก. และข้อ ข." หรือ "ถูกทั้งข้อ ก. และข้อ ค." ในแบบสอบเลือกตอบตัวเลือกธรรมดาน่าจะเป็นตัวเลือกแบบหนึ่ง ที่จะช่วยลดตัวเลือกที่มีลักษณะถูกหรือผิดอย่างชัดเจนลงเช่นกัน ด้วยเหตุผลดังกล่าวจึงน่าจะได้มีการศึกษาถึงคุณภาพของแบบสอบเลือกตอบตัวเลือกซ้อน และแบบสอบเลือกตอบตัวเลือกผสม ให้ชัดเจนยิ่งขึ้นว่ามีคุณภาพเป็นอย่างไร

จากการศึกษารายงานการวิจัยเกี่ยวกับคุณภาพของแบบสอบเลือกตอบชนิดตัวเลือกซ้อน เมื่อเปรียบเทียบกับแบบสอบเลือกตอบชนิดตัวเลือกธรรมดา นั้น Weiten (1982 : 46-50) ได้ศึกษาคุณภาพของแบบสอบเลือกตอบตัวเลือกธรรมดา กับตัวเลือกซ้อนในวิชาจิตวิทยากับนักศึกษาปริญญาตรี พบว่า แบบสอบเลือกตอบตัวเลือกธรรมดามีค่าอำนาจจำแนก และค่าความเที่ยง สูงกว่าแบบสอบเลือกตอบชนิดตัวเลือกซ้อน แบบสอบเลือกตอบตัวเลือกธรรมดาง่ายกว่าแบบสอบเลือกตอบชนิดตัวเลือกซ้อน ส่วนด้านความตรงของแบบทั้งสองชนิดไม่แตกต่างกัน

Hughes and Trimble (1965: 119-125) ได้ศึกษาแบบสอบที่มีตัวเลือก "ถูกทั้งข้อ 1 และข้อ 2" หรือ "ผิดทั้งข้อ 1 และข้อ 2" ซึ่งเรียกว่าตัวเลือกแบบผสม ในวิชาจิตวิทยา

กับนักศึกษาปริญญาตรีปีที่ 1 พบว่า แบบสอบเลือกตอบตัวเลือกผสมยากกว่าแบบสอบเลือกตอบตัวเลือกธรรมดา ส่วนค่าอำนาจจำแนกนั้นไม่แตกต่างกัน

สำหรับการวิจัยในประเทศไทยนั้น วรรณญา ปราณี (2529) ได้ศึกษาถึงคุณภาพของแบบสอบเลือกตอบตัวเลือกธรรมดา กับตัวเลือกซ้อนที่เงื่อนไขมีโอกาสได้รับเลือกเป็นตัวเลือกเท่ากัน และแบบตัวเลือกซ้อนที่เงื่อนไขมีโอกาสได้รับเลือกเป็นตัวเลือกไม่เท่ากัน ในวิชาวิทยาศาสตร์กับนักเรียนชั้นมัธยมศึกษาปีที่ 2 พบว่า แบบสอบเลือกตอบตัวเลือกซ้อนที่เงื่อนไขมีโอกาสได้รับเลือกเป็นตัวเลือกไม่เท่ากันมีค่าความยากมากที่สุด และมีความเที่ยงสูงที่สุด แบบสอบเลือกตอบตัวเลือกธรรมดาง่ายที่สุด และมีค่าความเที่ยงต่ำที่สุด แต่ค่าอำนาจจำแนกและเปอร์เซ็นต์การเดาไม่แตกต่างกัน ต่อมาในปี 2530 กมลทิพย์ อธิการยานนท์ ได้ศึกษาถึงคุณภาพของแบบสอบเลือกตอบชนิดคำตอบถูกตัวเดียว ค่าตอบถูกไม่จำกัด และคำตอบรวมหรือแบบตัวเลือกซ้อน พบว่า ค่าความเที่ยงและความตรงของแบบสอบคำตอบถูกตัวเดียวกับแบบคำตอบรวมไม่แตกต่างกัน แต่นักเรียนตอบแบบสอบแบบคำตอบถูกตัวเดียวมีปริมาณการเดาสูงกว่าแบบสอบแบบคำตอบรวม นอกจากนี้ ปราณี ร่วมทอง (2528) ได้ศึกษาถึงคุณภาพของแบบสอบเลือกตอบชนิดตัวเลือกธรรมดา และชนิดตัวเลือกซ้อนในการวัดระดับความรู้ที่แตกต่างกัน ในวิชาวิทยาศาสตร์กับนักเรียนชั้นมัธยมศึกษาปีที่ 3 พบว่า ในการวัดความรู้ขั้นต่ำ แบบสอบเลือกตอบตัวเลือกธรรมดามีค่าความเที่ยง และค่าอำนาจจำแนกสูงกว่าแบบสอบเลือกตอบชนิดตัวเลือกซ้อน ส่วนด้านความตรงไม่แตกต่างกัน สำหรับการวัดความรู้ขั้นสูง แบบสอบทั้งสองชนิดมีค่าความเที่ยง ความตรง ความยาก และอำนาจจำแนกไม่แตกต่างกัน นักเรียนที่มีผลสัมฤทธิ์ทางการเรียนสูงและต่ำตอบแบบสอบเลือกตอบทั้งสองชนิดได้ไม่แตกต่างกัน แต่ในกลุ่มนักเรียนที่มีผลสัมฤทธิ์ทางการเรียนปานกลาง ตอบแบบสอบเลือกตอบชนิดตัวเลือกธรรมดาได้คะแนนเฉลี่ยสูงกว่าแบบสอบเลือกตอบชนิดตัวเลือกซ้อน นอกจากนี้ นวลน้อย แต่บรรพกุล (2520) ได้ศึกษาแบบสอบเลือกตอบ 4 รูปแบบ คือ ตัวเลือกแบบธรรมดา ตัวเลือกแบบปลายเปิด "ผิดทุกข้อ" ตัวเลือกแบบปลายปิด "ถูกทุกข้อ" และตัวเลือกแบบผสม "ถูกทั้งข้อ ข และข้อ ค" หรือ "ถูกทั้งข้อ ก และ ข้อ ค" ในวิชาภาษาไทยกับนักเรียนชั้นประถมศึกษาปีที่ 7 พบว่า ค่าความเที่ยงและค่าอำนาจจำแนกของแบบสอบทั้ง 4 ฉบับไม่แตกต่างกัน แบบสอบตัวเลือกแบบผสมมีค่าความยากมาตรฐานสูงสุด

จากผลการวิจัยที่กล่าวมา แสดงให้เห็นว่าแบบสอบเลือกตอบที่มีรูปแบบต่างกันมีผลทำให้คุณภาพของแบบสอบแตกต่างกัน แต่ผลการวิจัยยังมีข้อสรุปที่ขัดแย้งกันอยู่ เช่น คุณภาพด้านความเที่ยง เมื่อพิจารณาถึงทฤษฎีการวัดที่ใช้เป็นแนวทางในการศึกษาเกี่ยวกับประสิทธิภาพของแบบสอบ

ต่าง ๆ นั้น ได้อาศัยทฤษฎีการวัดแบบคลาสสิก (Classical Test Theory) ซึ่งปัจจุบันทฤษฎีนี้ถูกมองว่ามีข้อบกพร่องหลายประการ ประการแรกคือ การศึกษาคุณภาพของแบบสอบโดยทฤษฎีนี้ ค่าสถิติที่ได้ของข้อสอบและแบบสอบ มักแปรเปลี่ยนไปตามกลุ่มตัวอย่างที่ใช้ในการศึกษา ประการที่สองการเปรียบเทียบความสามารถของผู้สอบต้องใช้แบบสอบฉบับเดียวกัน หรือแบบสอบคู่ขนาน ประการที่สามนิยามความเที่ยงของแบบสอบในรูปของการวัดคู่ขนานเป็นไปได้ยากในทางปฏิบัติ ประการที่สี่ไม่มีพื้นฐานในการที่จะบอกว่าผู้สอบคนหนึ่งจะมีโอกาสตอบข้อสอบข้อหนึ่งถูกเพียงใด และประการที่ห้าข้อตกลงเบื้องต้นของทฤษฎีที่กล่าวว่าความแปรปรวนของความคลาดเคลื่อนในการวัดของผู้สอบทุกคนเท่ากัน (Hambleton and Swaminathan, 1985:1-3) จุดอ่อนดังกล่าวทำให้การสรุปข้อความรู้ในเรื่องคุณภาพของแบบสอบทำได้ไม่ชัดเจนพอ หากได้รับการตรวจสอบด้วยวิธีวิเคราะห์โดยอาศัยกรอบทฤษฎีใหม่ ที่สามารถอธิบายปรากฏการณ์เกี่ยวกับการวัดได้ดีกว่า ก็น่าจะทำให้การตรวจสอบคุณภาพของแบบสอบมีความชัดเจนยิ่งขึ้น

ทฤษฎีการตอบสนองข้อสอบ (Item Response Theory) เป็นทฤษฎีการวัดทฤษฎีหนึ่ง ที่เข้ามามีบทบาทต่อวงการวัดผลทางการศึกษาและจิตวิทยาเพิ่มมากขึ้นในปัจจุบัน เนื่องจากทฤษฎีนี้สามารถแก้จุดอ่อนของทฤษฎีการวัดแบบคลาสสิกคอลได้ และสามารถนำไปประยุกต์เพื่อแก้ปัญหาการวัดทางการศึกษาและจิตวิทยาได้หลายเรื่อง โดยทฤษฎีการตอบสนองข้อสอบนี้เชื่อว่า พฤติกรรมที่บุคคลตอบสนองต่อข้อสอบนั้น ถูกกำหนดโดยลักษณะบางอย่างภายในตัวบุคคล (latent trait) ซึ่งไม่สามารถจะสังเกตได้ จึงพยายามที่จะอธิบายความสัมพันธ์ระหว่างลักษณะภายในตัวบุคคล (trait) กับพฤติกรรมการตอบสนองต่อข้อสอบ (performance) โดยอาศัยแบบจำลองทางคณิตศาสตร์เป็นตัวเชื่อมที่สำคัญ จุดเด่นของทฤษฎีนี้ก็คือ ค่าพารามิเตอร์ของข้อสอบไม่ขึ้นอยู่กับลักษณะของกลุ่มตัวอย่างผู้ตอบข้อสอบ และการประมาณค่าความสามารถของบุคคลไม่ขึ้นอยู่กับกลุ่มตัวอย่างของข้อสอบ นั่นคือ แม้ผู้สอบจะตอบข้อสอบที่ไม่ใช่ชุดเดียวกันก็สามารถประมาณค่าความสามารถเปรียบเทียบกันได้

ทฤษฎีการตอบสนองข้อสอบดังกล่าว ได้แสดงความสัมพันธ์ระหว่างผลการตอบข้อสอบแต่ละข้อ กับระดับความสามารถของผู้สอบด้วยฟังก์ชันการตอบสนองข้อสอบ (Item Response Function) หรือโค้งลักษณะของข้อสอบ (Item Characteristic Curve: ICC) ฟังก์ชันนี้คือความน่าจะเป็นของการตอบข้อสอบถูก โดยมีข้อสมมุติ (assume) ว่า ความน่าจะเป็นของการตอบข้อสอบถูก $P(\theta)$ จะเพิ่มขึ้นเมื่อความสามารถ (θ) เพิ่มขึ้น

นอกจากนี้ทฤษฎีการตอบสนองข้อสอบ ยังให้แนวคิดที่ชัดเจนเกี่ยวกับการประมาณค่าความสามารถของผู้ตอบข้อสอบ โดยใช้ผลการตอบสนองข้อสอบนั้นสรุปอ้างอิงเกี่ยวกับความสามารถที่มีอยู่ในตัวบุคคล ดัชนีที่ใช้วัดระดับความถูกต้องแม่นยำ (precision) ในการประมาณค่าความสามารถที่ต้องการวัด ได้แก่ ค่าสารสนเทศ (information function) (Birnbbaum, 1968: 419) ค่าสารสนเทศนี้จะมีค่ามากหรือน้อยขึ้นอยู่กับค่าพารามิเตอร์ของข้อสอบซึ่งเป็นตัวกำหนดลักษณะของข้อสอบ ค่าสารสนเทศจึงมีลักษณะเป็นดัชนีผสม (composite Index) ที่หาได้จากค่าพารามิเตอร์ของข้อสอบ คือ ค่าอำนาจจำแนก (a) ค่าความยาก (b) และค่าการเดา (c) การศึกษาคุณภาพของแบบสอบจึงมีความชัดเจนกว่า

จากแนวคิดตามกรอบทฤษฎีการตอบสนองข้อสอบ ผู้วิจัยจึงสนใจที่จะศึกษาว่ารูปแบบของตัวเลือกในแบบสอบเลือกตอบที่มีรูปแบบต่างกัน คือ แบบสอบเลือกตอบตัวเลือกธรรมดา แบบสอบเลือกตอบตัวเลือกผสม และแบบสอบเลือกตอบตัวเลือกซ้อน จะมีค่าฟังก์ชันสารสนเทศของข้อสอบ (IIF) และค่าฟังก์ชันสารสนเทศของแบบสอบ (TIF) แตกต่างกันอย่างไรรู้ข้อความรู้ที่ได้จะเป็นประโยชน์แก่ผู้เขียนข้อสอบ ในการเลือกใช้รูปแบบของแบบสอบเลือกตอบ ในลักษณะที่จะให้ค่าสารสนเทศของข้อสอบ และค่าสารสนเทศของแบบสอบที่สูงกว่า ซึ่งจะให้ความแม่นยำในการวัดมากกว่า

วัตถุประสงค์ของการวิจัย

การวิจัยครั้งนี้ มีวัตถุประสงค์ที่จะศึกษาเปรียบเทียบ ค่าฟังก์ชันสารสนเทศของข้อสอบ (IIF) และค่าฟังก์ชันสารสนเทศของแบบสอบ (TIF) ของแบบสอบเลือกตอบ 3 ฉบับ ที่มีตัวเลือกต่างกัน คือ แบบสอบเลือกตอบตัวเลือกธรรมดา แบบสอบเลือกตอบตัวเลือกผสม และแบบสอบเลือกตอบตัวเลือกซ้อน โดยใช้โมเดลโลจิสติกที่มีพารามิเตอร์ 3 ตัว (Three Parameter Logistic Model) ดังนี้

1. เปรียบเทียบค่าฟังก์ชันสารสนเทศของข้อสอบรายข้อ (Item Information Function: IIF)
2. เปรียบเทียบค่าฟังก์ชันสารสนเทศของแบบสอบ (Test Information Function: TIF)

สมมติฐานของการวิจัย

เนื่องจากการตอบข้อสอบเลือกตอบตัวเลือกซ้อนนั้น ผู้ตอบจะต้องรู้ว่าสถานการณ์ใดถูกหรือผิดทุกสถานการณ์ จึงจะสามารถเลือกคำตอบที่ถูกได้อย่างถูกต้อง จากการวิจัยของ Weiten (1982 : 46-50) และวรัญญา ปราบปัญจะ (2529) พบว่า แบบสอบเลือกตอบตัวเลือกซ้อนมีความยากมากกว่าแบบสอบเลือกตอบตัวเลือกธรรมดา ลักษณะดังกล่าวทำให้ผู้สอบต้องใช้ความสามารถมากขึ้น ผู้ที่มีความสามารถระดับสูงจะมีโอกาสตอบถูกได้มาก ในขณะที่ผู้ที่มีความสามารถปานกลางและต่ำจะมีโอกาสตอบถูกน้อย และเดาคำตอบได้ยากขึ้น ในทำนองเดียวกันการตอบข้อสอบเลือกตอบตัวเลือกผสมนั้น จากการศึกษาของ Hughes and Trimble (1965 : 119-125) และนวนน้อย แต่บรรพกุล (2520) พบว่า แบบสอบเลือกตอบตัวเลือกผสมมีความยากมากกว่าแบบสอบเลือกตอบตัวเลือกธรรมดา ทั้งนี้เนื่องจากผู้ตอบมักจะเลือกตัวเลือกผสมเป็นคำตอบมากกว่าตัวเลือกอื่น ดังนั้นเมื่อตัวเลือกผสมทำหน้าที่เป็นตัวลวงจะทำให้ข้อสอบยากขึ้นในด้านค่าอำนาจจำแนก พบว่า แบบสอบเลือกตอบตัวเลือกผสมมีแนวโน้มจะมีค่าอำนาจจำแนกที่ต่ำกว่าซึ่งจะมีผลต่อค่าพารามิเตอร์ของข้อสอบ และจะส่งผลถึงค่าสารสนเทศของข้อสอบ เพราะค่าฟังก์ชันสารสนเทศจะเพิ่มขึ้นเมื่อค่าอำนาจจำแนกเพิ่มขึ้น และค่าฟังก์ชันสารสนเทศจะลดลงเมื่อค่าการเดาเพิ่มขึ้น (Warm, 1978 : 69-70) ดังนั้นผู้วิจัยจึงตั้งสมมติฐานในการวิจัย ดังนี้

1. ค่าฟังก์ชันสารสนเทศของข้อสอบ (IIF) และค่าฟังก์ชันสารสนเทศของแบบสอบ (TIF) ของแบบสอบเลือกตอบที่ใช้ตัวเลือกธรรมดา น่าจะมีค่าสูงกว่าค่าฟังก์ชันสารสนเทศของข้อสอบ (IIF) และค่าฟังก์ชันสารสนเทศของแบบสอบ (TIF) ของแบบสอบเลือกตอบตัวเลือกผสม และแบบสอบเลือกตอบตัวเลือกซ้อน เมื่อเปรียบเทียบกันในช่วงตำแหน่ง $\theta < 1$

2. ค่าฟังก์ชันสารสนเทศของข้อสอบ (IIF) และค่าฟังก์ชันสารสนเทศของแบบสอบ (TIF) ของแบบสอบเลือกตอบตัวเลือกซ้อน น่าจะมีค่าสูงกว่าค่าฟังก์ชันสารสนเทศของข้อสอบ (IIF) และค่าฟังก์ชันสารสนเทศของแบบสอบ (TIF) ของแบบสอบเลือกตอบตัวเลือกธรรมดา และแบบสอบเลือกตอบตัวเลือกผสม เมื่อเปรียบเทียบกันในช่วงตำแหน่ง $\theta > 1$

3. ค่าฟังก์ชันสารสนเทศของข้อสอบ (IIF) และค่าฟังก์ชันสารสนเทศของแบบสอบ (TIF) ของแบบสอบเลือกตอบตัวเลือกผสม น่าจะมีค่าต่ำกว่าค่าฟังก์ชันสารสนเทศของข้อสอบ (IIF) และค่าฟังก์ชันสารสนเทศของแบบสอบ (TIF) ของแบบสอบเลือกตอบตัวเลือกธรรมดา และแบบสอบเลือกตอบตัวเลือกซ้อน ทุกตำแหน่งที่เปรียบเทียบ

ขอบเขตของการวิจัย

1. แบบสอบที่ใช้ในการวิจัยครั้งนี้คือ แบบสอบเลือกตอบกลุ่มสร้างเสริมประสบการณ์ชีวิต เรื่องสิ่งที่มีชีวิต หน่วยย่อยที่ 1 เรื่องตัวเรา ชนิด 4 ตัวเลือก สำหรับนักเรียนชั้นประถมศึกษาปีที่ 6

2. การเปรียบเทียบค่าฟังก์ชันสารสนเทศของข้อสอบ(IIF)และค่าฟังก์ชันสารสนเทศของแบบสอบ(TIF)นั้น ศึกษาตามกรอบทฤษฎีการตอบสนองข้อสอบ โมเดลโลจิสติกที่มีพารามิเตอร์ 3 ตัว

3. ตัวแปรที่ศึกษา ได้แก่

3.1 แบบสอบเลือกตอบที่มีรูปแบบตัวเลือก 3 รูปแบบ คือ แบบสอบเลือกตอบตัวเลือกธรรมดา แบบสอบเลือกตอบที่ใช้ตัวเลือกผสม และแบบสอบเลือกตอบตัวเลือกซ้อน ที่วัดเนื้อหาและพฤติกรรมเดียวกัน

3.2 ค่าอติเนต (ordinate) บนโค้งสารสนเทศของข้อสอบ และแบบสอบที่ครอบคลุมตำแหน่ง e ต่าง ๆ ดังนี้

1) ค่าอติเนตบนโค้งสารสนเทศของข้อสอบ(IIF) แต่ละข้อที่ตำแหน่ง e เท่ากับ $-3.0, -2.5, -2.0, -1.5, -1.0, -0.5, 0.0, 0.5, 1.0, 1.5, 2.0, 2.5$ และ 3.0 ตามลำดับ

2) ค่าอติเนตบนโค้งสารสนเทศของแบบสอบ(IIF) ทั้ง 3 ฉบับที่ตำแหน่ง e เท่ากับ $-3.0, -2.5, -2.0, -1.5, -1.0, -0.5, 0.0, 0.5, 1.0, 1.5, 2.0, 2.5$ และ 3.0 ตามลำดับ

ข้อตกลงเบื้องต้น

การดำเนินการสอบในเวลา สถานที่ที่แตกต่างกัน ไม่มีผลต่อการตอบสนองต่อข้อสอบของนักเรียนที่เป็นกลุ่มตัวอย่าง และนักเรียนที่เข้าสอบทุกคนตั้งใจทำแบบสอบอย่างเต็มความสามารถ

ความจำกัดในการวิจัย

1. แบบสอบตัวเลือกธรรมดา แบบสอบตัวเลือกผสม และแบบสอบตัวเลือกซ้อนที่ศึกษาขึ้น เป็นแบบสอบกลุ่มสร้างเสริมประสบการณ์ชีวิต ชั้นประถมศึกษาปีที่ 6 เรื่องสิ่งที่มีชีวิต หน่วยย่อย เรื่องตัวเราเท่านั้น ผลการวิจัยจึงไม่ครอบคลุมถึงเนื้อหาวิชาอื่น ๆ ด้วย

2. วัตถุประสงค์เชิงพฤติกรรมที่ใช้ในการออกข้อสอบ เป็นวัตถุประสงค์ที่อิงกับ แผนการสอนกลุ่มสร้างเสริมประสบการณ์ชีวิต ชั้นประถมศึกษาปีที่ 6 และสมุดประจำชั้น(ป.02) ชั้นประถมศึกษาปีที่ 6 เท่านั้น

3. จำนวนข้อสอบที่ใช้ในชั้นศึกษาจริง มีจำนวน 19 ข้อ โดยคัดเลือกมาจากข้อสอบ ที่สร้างขึ้น ตามตารางโครงสร้างน้ำหนักเนื้อหาและวัตถุประสงค์เชิงพฤติกรรมเดิมที่กำหนดไว้ จำนวน 25 ข้อ เนื่องจากข้อจำกัดในเรื่องระยะเวลาในการวิจัย จึงไม่สามารถพัฒนาข้อสอบ ให้มีคุณภาพตามเกณฑ์ที่จะนำไปใช้ศึกษาจริงได้ครบทุกข้อ และถือว่าข้อสอบ 19 ข้อดังกล่าว เป็น ตัวแทนของข้อสอบ 25 ข้อที่กำหนดไว้

คำจำกัดความที่ใช้ในการวิจัย

1. แบบสอบเลือกตอบ หมายถึง แบบสอบวัดผลสัมฤทธิ์กลุ่มสร้างเสริมประสบการณ์ชีวิต ชั้นประถมศึกษาปีที่ 6 เรื่องสิ่งที่มีชีวิต หน่วยย่อยที่ 1 เรื่องตัวเรา ในที่นี้มี 3 แบบ คือ

1.1 แบบสอบเลือกตอบตัวเลือกธรรมดา หมายถึง แบบสอบที่ข้อสอบแต่ละข้อ แบ่งออกเป็น 2 ส่วน คือ ส่วนที่เป็นตัวปัญหา (stem) และส่วนที่เป็นตัวเลือก (option) ซึ่งมีจำนวน 4 ตัวเลือกผู้ตอบต้องพิจารณาว่าคำตอบใดเป็นคำตอบที่ดีที่สุดเพียงคำตอบเดียว

1.2 แบบสอบเลือกตอบตัวเลือกผสม หมายถึง แบบสอบเลือกตอบที่มีลักษณะ เหมือนแบบสอบเลือกตอบตัวเลือกธรรมดา แต่มีตัวเลือกตัวสุดท้ายเขียนโดยรวมตัวเลือกไว้ด้วยกัน เช่น "ถูกต้องทั้งข้อ ก. และข้อ ข." หรือ "ถูกต้องทั้งข้อ ก. และข้อ ค." ตามความเหมาะสม

1.3 แบบสอบเลือกตอบตัวเลือกซ้อน หมายถึง แบบสอบเลือกตอบที่ข้อสอบ แบ่งออกเป็น 3 ส่วน คือ ส่วนที่เป็นตัวปัญหา (stem) ส่วนที่เป็นข้อความหรือสถานการณ์ซึ่งทำหน้าที่เป็นตัวเลือกในแบบสอบเลือกตอบตัวเลือกธรรมดา ได้กำหนดไว้เป็นข้อ ๆ และส่วนที่เป็นตัวเลือก สร้างมาจากข้อความสถานการณ์ที่กำหนดมาให้ นั้น ตัวเลือกแต่ละตัวอาจมี 1, 2, 3 หรือ 4 ข้อความ/สถานการณ์แล้วแต่ความเหมาะสม ซึ่งมีจำนวน 4 ตัวเลือก ผู้ตอบต้องพิจารณาว่าคำตอบใดเป็นคำตอบที่ดีที่สุดเพียงคำตอบเดียว

2. ค่าพารามิเตอร์ของข้อสอบ หมายถึง ค่าอำนาจจำแนกของข้อสอบ (a) ค่าความยากของข้อสอบ (b) และค่าการเดา (c) โดยที่

2.1 ค่าอำนาจจำแนกของข้อสอบ (a) หมายถึง ค่าที่เป็นสัดส่วนโดยตรงกับความชันของโค้งลักษณะข้อสอบ ณ จุดเปลี่ยนโค้ง เป็นค่าที่แสดงถึงความสามารถของข้อสอบที่จะจำแนกผู้สอบระหว่างผู้ที่มีระดับความสามารถต่ำกว่าค่าความยากของข้อสอบ และผู้ที่มีระดับความสามารถสูงกว่าค่าความยากของข้อสอบ มีค่าที่เป็นไปได้ในช่วง 0 ถึง + ∞ ในทางปฏิบัติมักจะคัดเลือกข้อสอบที่มีค่าอำนาจจำแนกอยู่ในช่วง 0.5 ถึง 2.0

2.2 ค่าความยากของข้อสอบ (b) หมายถึง ค่าที่แสดงถึงค่าความสามารถที่จุดโค้งลักษณะข้อสอบมีความชันมากที่สุด มีค่าเท่ากับระดับความสามารถของผู้สอบ θ ที่มีโอกาสตอบข้อสอบข้อนั้นได้ถูก $(1+C) / 2$ มีค่าที่เป็นไปได้ ตั้งแต่ $-\infty$ ถึง $+\infty$ ในทางปฏิบัติจะคัดเลือกข้อสอบที่มีค่าความยากอยู่ในช่วง -2.5 ถึง $+2.5$ ถ้า $b = -2.5$ แสดงว่าข้อสอบง่ายมาก $b = 0$ แสดงว่าข้อสอบยากปานกลาง และถ้า $b = +2.5$ แสดงว่าข้อสอบยากมาก

2.3 ค่าการเดาของข้อสอบ (c) หมายถึง ค่าความน่าจะเป็นที่ผู้สอบที่มีความสามารถต่ำมากจะมีโอกาสตอบข้อสอบข้อนั้นถูก มีค่าเป็นไปได้อยู่ในช่วง 0 ถึง 1 ในทางปฏิบัติมักจะคัดเลือกข้อสอบที่มีค่า $c < .3$

3. ค่าฟังก์ชันสารสนเทศข้อสอบ (IIF) หมายถึง ดัชนีที่แสดงถึงความแม่นยำในการประมาณค่าความสามารถจริงด้วยผลการตอบข้อสอบข้อนั้น ๆ มีค่าเท่ากับ อัตราส่วนของกำลังสองของความชันของโค้งลักษณะข้อสอบ ต่อค่าความแปรปรวนของข้อสอบข้อนั้น ณ ระดับความสามารถ (θ) ในที่นี้มี 13 ค่า คือค่าออกติเนตบนโค้งสารสนเทศของข้อสอบ ที่ตำแหน่ง θ เท่ากับ $-3.0, -2.5, -2.0, -1.5, -1.0, -0.5, 0.0, 0.5, 1.0, 1.5, 2.0, 2.5$ และ 3.0 ตามลำดับ

คำนวณค่าฟังก์ชันสารสนเทศของข้อสอบ โดยสูตร

$$I(\theta_j, u_i) = \frac{[P_1(\theta_j)]^2}{P_1(\theta_j) Q_1(\theta_j)}$$

เมื่อ $I(\theta_j, u_i)$ คือ ค่าฟังก์ชันสารสนเทศของข้อสอบ

$P_1(\theta_j)$ คือ ความชันของโค้งลักษณะข้อสอบที่ระดับความสามารถ θ_j

$P_1(\theta_j)$ คือ ค่าความน่าจะเป็นที่ผู้ตอบที่มีความสามารถ θ_j

จะตอบข้อสอบข้อ i ได้ถูก

$Q_1(\theta_j)$ คือ $1 - P_1$

4. ประสิทธิภาพสัมพัทธ์ของข้อสอบ (RE) หมายถึง อัตราส่วนระหว่างค่า
สารสนเทศของข้อสอบต่างแบบที่นำมาเปรียบเทียบกัน

5. ค่าฟังก์ชันสารสนเทศแบบสอบ (TIF) หมายถึง ผลรวมของค่าสารสนเทศของ
ข้อสอบทุก ๆ ข้อในแบบสอบ สำหรับความสามารถของผู้สอบต่าง ๆ กัน ถ้ามีค่าสูงที่ระดับ
ความสามารถของผู้สอบใด แสดงว่า แบบสอบนั้นสามารถจำแนกความสามารถของผู้สอบได้ดี
ณ ระดับความสามารถนั้น ในที่นี้หมายถึงค่าฟังก์ชันสารสนเทศของแบบสอบเลือกตอบตัวเลือก
ธรรมดา แบบสอบเลือกตอบตัวเลือกผสม และแบบสอบเลือกตอบตัวเลือกซ้อน ที่มีค่าออกเินเต
นโดลิงสารสนเทศของแบบสอบที่ตำแหน่ง θ เท่ากับ $-3.0, -2.5, -2.0, -1.5, -1.0, -0.5,$
 $0.0, 0.5, 1.0, 1.5, 2.0, 2.5$ และ 3.0 ตามลำดับ รวม 13 ค่า

คำนวณค่าฟังก์ชันสารสนเทศของแบบสอบ โดยสูตร

$$I(\theta_i) = \frac{n}{\sum_{i=1}^n [P_i(\theta_i)]^2 / P_i(\theta_i) Q_i(\theta_i)}$$

6. ระดับความสามารถ (θ) หมายถึง ระดับความรู้ความเข้าใจในเนื้อหา
การเรียนรู้กลุ่มสร้างเสริมประสบการณ์ชีวิต เรื่องสิ่งที่มีชีวิต หน่วยย่อยที่ 1 เรื่องตัวเรา ของ
นักเรียนชั้น ป.6 วัดในหน่วย θ ที่มีค่าเฉลี่ยเท่ากับ θ ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1
ในที่นี้แบ่งระดับความสามารถออกเป็น 3 ระดับ ดังนี้

$\theta > 1$ คือ กลุ่มที่มีความสามารถสูง

$1 \geq \theta \geq -1$ คือ กลุ่มที่มีความสามารถปานกลาง

$\theta < -1$ คือ กลุ่มที่มีความสามารถต่ำ

7. นักเรียน หมายถึง นักเรียนชั้นประถมศึกษาปีที่ 6 ภาคเรียนที่ 1 ปีการศึกษา
2535 ในสังกัดสำนักงานการประถมศึกษาจังหวัดชัยนาท จำนวน 4722 คน

ประโยชน์ที่ได้รับจากการวิจัย

1. จะทำให้ทราบถึงคุณลักษณะต่าง ๆ ของแบบสอบเลือกตอบ ที่มีตัวเลือกแตกต่างกัน คือ แบบสอบเลือกตอบตัวเลือกธรรมดา แบบสอบเลือกตอบตัวเลือกผสม และแบบสอบเลือกตอบตัวเลือกซ้อน ตามกรอบทฤษฎีการตอบสนองข้อสอบ ซึ่งข้อความรู้ที่ได้นี้จะช่วยเสริมข้อความรู้เดิมเกี่ยวกับแบบสอบทั้ง 3 แบบดังกล่าวให้สมบูรณ์ชัดเจนยิ่งขึ้น
2. ข้อความรู้ที่ได้จากการศึกษาครั้งนี้ จะเป็นประโยชน์สำหรับการเขียนข้อสอบแบบเลือกตอบ โดยพิจารณาเลือกรูปแบบตัวเลือกที่มีประสิทธิภาพสูงในการวัดผลสัมฤทธิ์ทางการเรียน นอกจากนี้ หากต้องการใช้แบบสอบเพื่อการประมาณค่าความสามารถของผู้สอบ ก็สามารถเลือกรูปแบบของตัวเลือกที่ให้ค่าสารสนเทศสูงที่สุดในระดับความสามารถที่ต้องการประมาณค่านั้นได้
3. ข้อความรู้ที่ได้จากการวิจัย จะเป็นประโยชน์สำหรับผู้สนใจจะศึกษาวิจัยในทำนองเดียวกันนี้ ซึ่งจะเป็นประโยชน์ต่อวงการวัดผลการศึกษาต่อไป

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย