



บทที่ 1

บทนำ

ลักษณะและความสำคัญของแบบสอบอิงเกณฑ์

แบบสอบที่ใช้วัดผลสัมฤทธิ์ทางการเรียนที่มีมาก่อน ได้แก่ แบบสอบที่วัดในระบบของการอิงร้อยละของคะแนนที่ผู้สอบแต่ละคนทำได้จากแบบสอบที่กำหนด การประเมินผลสัมฤทธิ์ทางการเรียนในระยะนี้จะเป็นการเปรียบเทียบระหว่างร้อยละของคะแนนที่ผู้สอบแต่ละคนทำได้กับเกณฑ์ คือร้อยละของคะแนนที่ถือว่าผู้สอบสอบผ่าน แบบสอบดังกล่าวนี้ถือได้ว่าเป็นแบบสอบอิงเกณฑ์ในระบบเก่า คืออิงร้อยละของคะแนนที่ได้จากการสอบ ระยะต่อมาเมื่อได้มีการพัฒนาทฤษฎีการวัด เชาวน์ปัญญาและทฤษฎีทางการวัดผลขึ้น การวัดผลสัมฤทธิ์ทางการเรียนนิยมใช้แบบสอบอิงกลุ่ม (Norm-Referenced Tests) แทนหลักการสำคัญของการใช้แบบสอบอิงกลุ่มจะมุ่งวัดผลสัมฤทธิ์ทางการเรียนเพื่อบ่งชี้ระดับความสามารถของผู้เรียน โดยการเปรียบเทียบกันเองภายในกลุ่มผู้เรียนว่า ใครมีความสามารถสูงหรือต่ำกว่ากัน แนวคิดการประเมินแบบอิงกลุ่มได้รับการนำมาใช้เพื่อวัดผลสัมฤทธิ์ทางการเรียนของผู้เรียนอย่างกว้างขวางในช่วงประมาณคริสต์ทศวรรษที่ 20

(Popham 1978 : 74) แต่หลังจากช่วงปี 1960 เป็นต้นมา ใ้มีนักวิชาการหลายท่าน เขียนบทความชี้ให้เห็นถึงข้อบกพร่องของการใช้แบบสอบอิงกลุ่มในการวัดผลสัมฤทธิ์ทางการเรียน ในขณะเดียวกันก็ได้มีการนำเสนอแนวคิดและพัฒนาแบบสอบอิงเกณฑ์

(Criterion-Referenced Tests) ในแนวใหม่ขึ้นมาทดแทน โดยมีหลักการสำคัญว่า การวัดผลสัมฤทธิ์ทางการเรียนควรจะมีจุดมุ่งที่จะชี้บ่งถึงระดับผลสัมฤทธิ์ในก้านความรู้หรือศักยภาพของผู้เรียนแต่ละคนในเรื่องใด เรื่องหนึ่งว่า นี้อยู่ในระดับใดออกมาได้อย่างชัดเจน สำหรับข้อบกพร่องของการใช้แบบสอบอิงกลุ่มเพื่อประเมินผลสัมฤทธิ์ทางการเรียนของนักเรียนที่นักวิชาการได้อภิปรายไว้นั้น มีสาระสำคัญพอประมวลได้คือ ประเด็นแรกได้แก่ปัญหาความเป็นตัวแทนของเนื้อหาที่ต่อการจะวัดและการแปลความหมายของคะแนนที่ได้จากการวัด ในปี ค.ศ. 1962 อีเบล (Ebel in Mehrens and Ebel 1976)

ได้เขียนบทความชี้ให้เห็นว่า คะแนนจากแบบสอบจะมีความหมายก็ต่อเมื่อมีการระบุขอบเขต (Domain) ของเนื้อหาที่ต้องการสอบให้เห็นชัด และการแปลความหมายของคะแนนที่ได้ ก็ควรจะทำในลักษณะของการพิจารณาจากจำนวนเปอร์เซ็นต์ที่ผู้สอบทำได้อีกด้วย ต่อมา เกลเซอร์ (Glaser 1963) ได้ชี้ให้เห็นว่า การแปลความหมายของคะแนนที่ได้จากแบบสอบผลสัมฤทธิ์ทางการเรียนนั้นควรจะมีเกณฑ์มาตรฐานที่กำหนดขึ้นเป็นหลักในการแปลความหมายของคะแนนที่ผู้สอบแต่ละคนทำได้ มากกว่าที่จะแปลความหมายในเชิงของการเปรียบเทียบกันเองภายในกลุ่มของผู้สอบ เกลเซอร์ได้เน้นไว้อย่างชัดแจ้งว่า เกณฑ์มาตรฐานนั้นจะต้องหมายถึงระดับความสามารถของผู้เรียนแต่ละคนภายใต้ขอบเขตของเนื้อหาหรืองานที่ได้รับกำหนดขึ้นไว้อย่างเด่นชัดก่อนที่จะมีการสอบ ออสเบอร์น (Osburn 1968) ก็ได้วิจารณ์แบบสอบอิงกลุ่มไว้ในแนวเดียวกันว่า เป็นแบบสอบที่สร้างขึ้นด้วยข้อสอบที่ไม่ได้รับการสุ่มขึ้นมาจากขอบเขตของเนื้อหาที่ได้รับกำหนดไว้แน่ชัดล่วงหน้า ในประเด็นนี้แบบสอบอิงเกณฑ์จึงได้รับการชี้ให้เห็นว่ามีความแตกต่างอย่างชัดแจ้งจากแบบสอบอิงกลุ่มคือ จะต้องมีการกำหนดขอบเขตของเนื้อหาที่ต้องการทดสอบขึ้นก่อนที่จะมีการสร้างแบบสอบ ข้อคำถามในแบบสอบอิงเกณฑ์จะต้องมีความเป็นตัวแทนของมวลความรู้ที่กำหนดขึ้น การสอบจะเป็นไปเพื่อการตรวจสอบว่า ผู้เรียนแต่ละคนมีระดับผลสัมฤทธิ์ทางการเรียนในเรื่องนั้นๆอยู่ในระดับใดเมื่อเทียบกับเกณฑ์มาตรฐานที่กำหนดขึ้นไว้ โดยไม่สนใจที่จะมุ่งเปรียบเทียบว่า ผู้เรียนคนใดมีความสามารถสูงหรือต่ำกว่าเมื่อเทียบกับบุคคลอื่น ๆ ในกลุ่ม (Popham and Husek 1969)

นอกจากประเด็นในค่านี้อีกสาระและการแปลความหมายของคะแนนแล้ว การใช้แบบสอบอิงกลุ่มเพื่อประกอบการดำเนินการเรียนการสอนยังได้รับการวิจารณ์ว่ามีข้อบกพร่อง คือ ข้อสอบที่ใช้ในแบบสอบอิงกลุ่มเป็นข้อสอบที่มุ่งเน้นในการจำแนกหรือกระจายความสามารถของคนออกให้เห็นเด่นชัด ข้อสอบที่ไต่ยากมากหรือง่ายมากเกินไป จะถูกคัดออก คงไว้แต่ข้อสอบที่มีความยากง่ายพอเหมาะและมีค่าอำนาจจำแนกสูง ประเด็นดังกล่าวนี้ เกลเซอร์ (Glaser 1963) ได้กล่าวไว้เช่นกันว่า ในการใช้การทดสอบเพื่อประกอบการเรียนการสอนนั้น ย่อมที่จะเป็นไปได้ที่ข้อสอบจะยากมากเมื่อนำไปทดสอบก่อนที่จะมีการจัดการเรียนการสอน และเป็นข้อสอบที่ง่ายมาก ทั้งไม่มีอำนาจจำแนกเมื่อมีการจัดการเรียนการสอนแล้ว แต่ข้อสอบนั้นๆยังคงเป็นข้อสอบที่มีความสำคัญในแง่ของ

การเรียนการสอน ไม่ควรจะถูกคัดออกไป ซึ่งโปแฟม (Popham 1978 : 83) ได้กล่าววิจารณ์ไว้ว่า โดยมาตรการของการคัดเลือกข้อสอบดังกล่าวในระบบของการสอนแบบอิงกลุ่ม แบบสอบที่ได้จะทำหน้าที่แทบจะไม่แตกต่างไปจากแบบสอบวัดความถนัด คือมุ่งจำแนกบุคคลมากกว่าที่จะมุ่งระบุระดับการรับรู้ของบุคคล

นอกจากนั้น โปแฟม (Popham 1978 : 81 - 82) ได้กล่าวไว้ด้วยว่า ในกรณีที่ผู้สอบได้คะแนนต่ำจากแบบสอบผลสัมฤทธิ์ทางการเรียนแบบอิงกลุ่มจะประสบความสำเร็จมากในการใช้ข้อมูลเพื่อการแก้ไขปรับปรุงการเรียนของตนเอง เพราะไม่สามารถระบุขอบเขตของเนื้อหาหรือสมรรถภาพที่ขาดไปได้ แต่ถ้าเป็นกรณีของแบบสอบอิงเกณฑ์แล้ว ผู้สอบจะสามารถทราบได้ชัดแจ้งว่าตนยังมีความรู้ หรือมีสมรรถภาพต่ำในก้านใบบ้าง ทำให้สามารถพัฒนาหรือปรับแก้ขอบกพร่องของตนได้ ในประเด็นนี้แบบสอบอิงเกณฑ์จึงเป็นประโยชน์ในการวินิจฉัยถึงขอบกพร่องของผู้เรียนได้ดีย

จากแนวคิดของการพัฒนาแบบสอบอิงเกณฑ์ เพื่อแก้ปัญหาต่างๆที่มีอยู่เมื่อใช้แบบสอบอิงกลุ่ม นักการศึกษาซึ่งมุ่งเน้นพัฒนาแบบสอบอิงเกณฑ์ได้มองเห็นว่า แนวทางของการพัฒนาแบบสอบจะต้องมีความแตกต่างไปจากแนวคิดของการพัฒนาแบบสอบอิงกลุ่มที่มีมาก่อน (Glaser and Nitko in Thorndike 1971 : 652) ดังนั้น นับจากสมัยของอีเบลและเกลเซอร์มาจนถึงปัจจุบัน แนวคิด ตลอดจนวิธีการต่างๆของการพัฒนาแบบสอบอิงเกณฑ์จึงได้รับการพัฒนาก้าวหน้าออกไปมาก เช่น ได้มีการให้ความหมายของคำว่า "การอิงเกณฑ์" ให้ความหมายของคำว่า "โอบเมน" อย่างชัดเจน มีการกำหนดขั้นตอนของการพัฒนาแบบสอบ มีวิธีการวิเคราะห์หาคุณภาพของข้อสอบโดยเฉพาะที่เป็นระบบของตนเอง มีวิธีการกำหนดเกณฑ์มาตรฐานขึ้นหลายวิธี มีมาตรการการหาค่าความเที่ยงที่เป็นเทคนิคเฉพาะของแบบสอบอิงเกณฑ์ เป็นต้น

ปัญหาการกำหนดความยาวของแบบสอบอิงเกณฑ์

ปัญหาที่สำคัญประการหนึ่งของแบบสอบอิงเกณฑ์ซึ่งยังไม่มีมาตรการการแก้ปัญหาที่เด่นชัด คือ ปัญหาของการกำหนดความยาวของแบบสอบอิงเกณฑ์ให้เหมาะสม โดยประเด็นของปัญหายุ่ตรงที่ว่า ในขอบเขตของเนื้อหาที่ได้รับการกำหนดขึ้น เพื่อสร้างข้อสอบวัดความรู้ของผู้เรียนนั้น ผู้วัดสามารถสร้างข้อสอบเพื่อใช้ในแบบสอบอิงเกณฑ์ได้เป็นจำนวนมาก

แต่ในทางปฏิบัตินั้น ผู้วัดไม่สามารถที่จะให้ผู้เรียนตอบคำถามต่างๆ ข้อใด เพราะจะเป็นการสิ้นเปลืองแรงงาน เวลา และค่าใช้จ่ายในการสอบ จึงได้มีการหาทางแก้ไขโดยวิธีการสุ่มเอาข้อสอบจำนวนหนึ่งขึ้นมาใช้สอบกับผู้เรียน โดยยึดหลักการว่าถ้าผู้สอบตอบคำถามที่สุ่มขึ้นมาได้ถูกต้องและถือว่าเป็นผู้มีความรอบรู้แล้ว จะมีความหมายว่าถ้าผู้สอบตอบคำถามต่างๆ ข้อในส่วนที่เหลือทั้งหมดก็จะยังคงเป็นผู้รอบรู้เช่นกัน (Glaser and Mitko in Thorndike 1971 : 660) จากแนวคิดดังกล่าวนี้ จึงเกิดคำถามต่อไปว่า ควรใช้ข้อสอบในการสอบแต่ละครั้ง เป็นจำนวนเท่าใด จึงจะมีความเหมาะสมที่สุด หรือโดยนัยของประสิทธิภาพแล้ว ควรจะใช้ข้อสอบเพียงจำนวนน้อยข้อ แต่ผลที่ได้จากการกักลินจะต้องมีความสอดคล้องกับการกักลินใจที่ เกิดจากการใช้แบบสอบที่ประกอบขึ้นด้วยข้อคำถามทั้งหมด ภายใต้ขอบเขตของเนื้อหาที่กำหนด

แนวคิดในความพยายามกำหนดความยาวของแบบสอบอิง เกณฑ์ที่เหมาะสม เริ่มมีปรากฏในงานเขียนของมิลล์แมน (Millman 1973) โดยเริ่มใช้การแจกแจงทวินาม (Binomial Distribution) เป็นหลักในการกำหนดความยาวของแบบสอบอิง เกณฑ์ แนววิธีการนี้ถูกต้องคือจะต้องทราบค่าระดับความสามารถที่แท้จริงของผู้สอบก่อนจึงจะสามารถคำนวณจำนวนข้อสอบที่ควรจะได้ ซึ่งปกติแล้วไม่มีทางที่จะทราบค่าความสามารถที่แท้จริงของบุคคล คอมมานอวิกและเลวิส (Novick and Lewis 1974) ได้เสนอแนวคิดของการใช้วิธีการเบย์เซียน (A Bayesian Model) กำหนดความยาวของแบบสอบอิง เกณฑ์ โดยมุ่งใช้ข้อสอบจำนวนน้อยข้อที่สุด ในการคำนวณจะมีการนำเอาข้อมูลเกี่ยวกับความสามารถของผู้เรียนที่มีมาก่อนค่า เนิ่นการสอบมาใช้ เป็นข้อมูลเบื้องต้น เพื่อใช้เป็นตัวช่วยในการคำนวณค่าหลักการของเบย์เซียน โดยอาศัยข้อตกลงเบื้องต้นว่า การกระจายของข้อมูลจะต้องเป็นแบบเบต้า ซึ่งข้อตกลงเบื้องต้นดังกล่าวจะไม่เป็นจริงถ้าผู้สอบมีความสามารถสูงมากหรือต่ำมาก (Millman in Popham 1974 : 355) ระยะเวลาได้มีการเสนอให้นำเอาทฤษฎีการตอบสนองข้อสอบ (Item Response Theory) มาใช้ในการกำหนดความยาวที่เหมาะสมของแบบสอบอิง เกณฑ์ โดยได้ข้อค้นพบที่สำคัญจากการศึกษาว่า ถ้าใช้ข้อสอบที่มีค่าความยาก (b) ตรงกับระดับความสามารถที่แท้จริงของผู้สอบแล้ว ความคลาดเคลื่อนในการประมาณความสามารถของผู้สอบ เมื่อเทียบกับระดับความสามารถที่แท้จริง จะต่ำกว่าความคลาดเคลื่อนในการประมาณความสามารถของผู้สอบ เมื่อใช้วิธีประมาณจาก

แบบสอบที่ประกอบขึ้นจากข้อสอบที่ได้รับการสุ่มขึ้นมา (Haladyna and Roid 1983) แต่การประยุกต์หลักการดังกล่าวนี้มีข้อจำกัดอยู่ตรงที่ว่า แบบสอบอิงเกณฑ์มุ่งเน้นการใช้ข้อสอบที่วัดได้ตรงตามขอบเขตของเนื้อหาที่กำหนดเป็นประเด็นหลัก ส่วนค่าประจำข้อสอบเป็นเพียงเครื่องช่วยเสริมเท่านั้น ในการประยุกต์ทฤษฎีการตอบสนองข้อสอบนั้น กลับยึดเอาค่าประจำข้อสอบเป็นคุณสมบัติสำคัญในการพิจารณา ซึ่งไม่ใช่หลักการที่สำคัญที่สุดของแบบสอบอิงเกณฑ์

นอกจากนี้โปแฟม (Popham 1978 : 100) ได้อภิปรายไว้ว่า ปัญหาการกำหนดความยาวของแบบสอบอิงเกณฑ์ที่เหมาะสมยังไม่มีความชัดเจน และเป็นปัญหาที่มีความสำคัญ เพราะถ้าใช้แบบสอบที่สั้นจนเกินไป ค่าความสามารถที่ประมาณได้ก็จะเป็นความคงที่ แต่ถ้าใช้แบบสอบที่มีความยาวมากเกินไปก็จะเสียเวลาและค่าใช้จ่ายในการสอบมากเกินไป ทั้งต้องสร้างข้อสอบเพิ่ม โปแฟมได้อภิปรายเชิงชี้แนะไว้ด้วยว่า ความจริงแล้วไม่ว่าจะสามารถกำหนดความยาวของแบบสอบอิงเกณฑ์ที่มีความเหมาะสมกับทุกสถานการณ์ได้ ปัจจัยต่างๆที่ควรจะต้องคำนึงถึงในการกำหนดความยาวของแบบสอบอิงเกณฑ์ อย่างน้อยควรมีอยู่ 3 ประการ คือ

1. การกำหนดเกณฑ์มาตรฐานที่ต่างกัน เช่น 50% หรือ 90% ของข้อสอบรวม จะต้องใช้แบบสอบที่มีความยาวต่างกัน
2. อัตราความคลาดเคลื่อนในการจัดประเภทบุคคลคือ การตัดสินให้ผู้รอบรู้สอบตก และผู้ไม่รอบรู้สอบผ่าน ที่จะยอมให้มีโอกาสเกิดขึ้นได้
3. ระดับความสามารถที่แท้จริงของผู้สอบ

นอกจากนี้ โปแฟมยังได้สรุปไว้ด้วยว่า ในสถานการณ์จริงนั้น ปัจจัยต่างๆทั้งสามประการดังกล่าวจะมีการผันแปรไปพร้อมๆกัน จึงทำให้การกำหนดความยาวของแบบสอบอิงเกณฑ์มีความยุ่งยากขึ้นไปอีก

ต่อมา เบรนนัน (Brennan 1981 : 74) ได้กล่าวถึงปัญหาเกี่ยวกับความยาวของแบบสอบอิงเกณฑ์ไว้ว่า ไม่มีวิธีการทางสถิติใดๆที่จะสร้างเป็นกฎเกณฑ์กำหนดความยาวของแบบสอบอิงเกณฑ์ที่ใช้ได้กับทุกกรณี ทั้งนี้ เพราะการกำหนดความยาวของแบบสอบอิงเกณฑ์ขึ้นอยู่กับปัจจัยที่สำคัญประการหนึ่ง คือการกำหนดอัตราความคลาดเคลื่อนที่จะยอมให้เกิดขึ้นได้ ซึ่งการกำหนดอัตราดังกล่าวมีภาวะเป็นอคติด้วย ทั้งการกำหนดอัตราความคลาด-

เคลื่อนที่ยัง จะต้อง ไร้ข้อมูล เกี่ยวกับภาวะแวดล้อมทางการสอบ และการทัศนคติในประเด็น
ต่างๆ เกี่ยวกับแบบสอบเข้ามาร่วมพิจารณา แต่อย่างไรก็ดี เบรนนัน ได้สรุปว่า มาตรการ
ทางสถิติก็ยังคงช่วยให้ข้อมูลประกอบการตัดสินใจ เกี่ยวกับการกำหนดความยาวของแบบสอบ
อิงเกณฑ์ภายใต้ภาวะอัติวิสัยได้

เมื่อวิเคราะห์แนวความคิดของ โปแฟมและเบรนนันแล้ว พบว่าบุคคลทั้งสองมีความ
คิดที่สอดคล้องกันอยู่ประการหนึ่งคือ ทั้ง โปแฟมและเบรนนัน เห็นว่าการกำหนดความยาวของ
แบบสอบอิง เกณฑ์นั้นคง ไม่สามารถกำหนดเป็นมาตรฐานที่ใช้ได้กับทุกกรณี แต่ควร จะกำหนด
ความยาวที่เหมาะสมของแบบสอบอิง เกณฑ์ เป็น เฉพาะกรณีไป โดยการนำเอาข้อมูลต่างๆที่
เกี่ยวข้อง กับภาวะการตัดสินใจจากการใช้แบบสอบมาร่วมพิจารณา ทั้งนี้ เพื่อพิจารณาจาก
หลักการพัฒนาแบบสอบอิง เกณฑ์ที่โปแฟมและฮูเสก (Popham and Husek 1969) ได้
กล่าวไว้ และปัญหาการกำหนดความยาวของแบบสอบอิง เกณฑ์ที่โปแฟมและเบรนนัน ได้ถีก-
ปราชญ์แล้วนั้น มาตรการที่จะนำมาใช้อย่างน้อยควรมีคุณสมบัติ 4 ประการ คือ

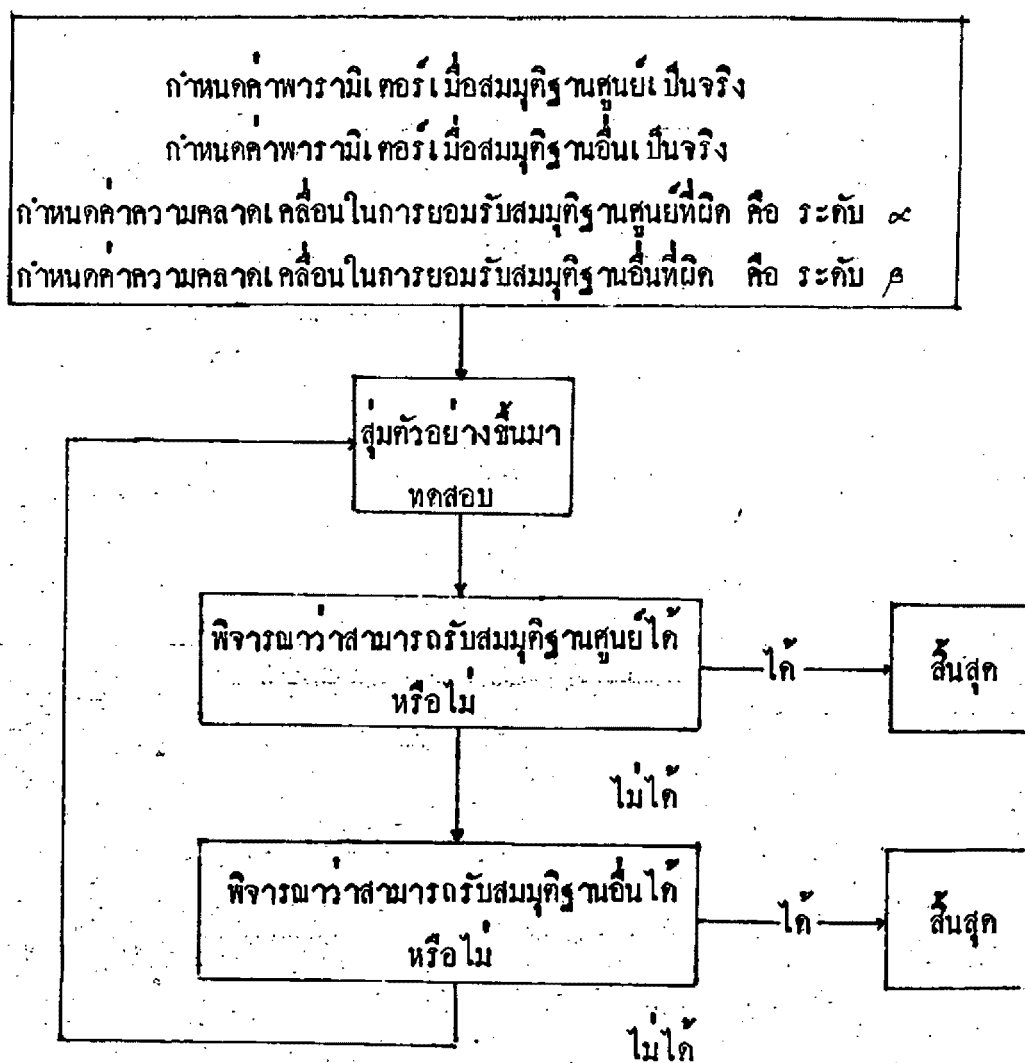
1. เป็นวิธีการที่ยึดเอาความตรงตามเนื้อหา เป็นหลักการสำคัญมากกว่าจะยึดเอา
ค่าประจำข้อสอบ เป็นหลัก
2. มีความยืดหยุ่นได้ในประเด็นที่ว่า ถ้ามีการเปลี่ยนแปลง เกณฑ์มาตรฐานใน
การตัดสินใจจัดประเภทบุคคลตามระดับการรอบรู้ อัตรากวามคลาดเคลื่อนในการจัดประ เภท
บุคคลนี้ที่จะยอมให้ เกิดขึ้นได้ หรือระดับความสามารถของผู้สอบที่แตกต่างกันแล้ว วิธีการที่
จะนำมาใช้จะต้องสามารถปรับเปลี่ยน ไป เพื่อให้ เหมาะสมกับสถานการณ์นั้นๆ
3. แนวคิดที่จะนำมาใช้นอกจากจะมีหลักการทางคณิตศาสตร์แล้ว ควรจะสามารถ
ตรวจสอบได้ด้วยข้อมูลเชิงประจักษ์
4. แนวคิดที่จะนำมาใช้ในการกำหนดความยาวของแบบสอบอิง เกณฑ์ นอกจาก
จะให้ข้อสรุปเกี่ยวกับความยาวที่เหมาะสมของแบบสอบอิง เกณฑ์เฉพาะกรณีแล้ว ควรจะต้อง
เป็นแนวคิดที่เสนอแนะขั้นตอน วิธีดำเนินการ เพื่อกำหนดความยาวของแบบสอบอิง เกณฑ์ใน
กรณีอื่นซึ่งอาจจะมีสภาพการณ์แตกต่าง ไปจากกรณีเฉพาะที่ได้ศึกษากัน

การประยุกต์เทคนิคการวิเคราะห์ตามลำดับขั้น (Sequential Analysis) เพื่อกำหนดความยาวที่เหมาะสมของแบบสอบถามเชิงคุณภาพ

เกลเซอร์และนิตโก (Glaser and Nitko in Popham 1971) และ มิลล์แมน (Millman 1973) ได้กล่าวพากพิงไปถึงการใช้ "เทคนิคการวิเคราะห์ตามลำดับขั้น" ว่าจะสามารถใช้เป็นวิธีการกำหนดความยาวของแบบสอบถามเชิงคุณภาพได้ ในช่วงปี 1950 - 1970 มีรายงานการวิจัยเกี่ยวกับการประยุกต์ใช้วิธีการดังกล่าวเพื่อลดความยาวของแบบสอบถามที่ใช้ในชั้นเรียน และพบว่าหลักการนี้สามารถลดความยาวของแบบสอบถามได้ โดยการจักประเภทนักเรียนตามระดับการรับรู้ยังคงเดิม คือ ไม่แตกต่างกันไปจากการจักประเภทจากแบบสอบถามเดิม ภายใต้อัตราความคลาดเคลื่อนที่กำหนดขึ้น ส.ศ. สไตน์และพิริยานุวัฒน์ (Stecklein and Pitiyanuwat 1975) กล่าวไว้เช่นกันว่าจะนำเอาแนวคิดนี้มาใช้ในการกำหนดความยาวของแบบสอบถามเชิงคุณภาพ

เทคนิคการวิเคราะห์ตามลำดับขั้นพัฒนาขึ้นโดย แอברהฮัม วอลด์ (Abraham Wald) มีปรากฏหลักฐานที่พิมพ์ในวารสาร "The Annals of Mathematical Statistics" และ "Journal of American Statistical Association" ในปี ค.ศ. 1945 แนวคิดดั้งเดิมมุ่งที่จะพัฒนาขึ้นใช้ในงานอุตสาหกรรม เทคนิคการวิเคราะห์ตามลำดับขั้นแตกต่างไปจากหลักการของสถิติสรุปพากพิง (Inferential Statistics) ที่ใช้กันอยู่ทั่วไปตรงที่ว่า สถิติสรุปพากพิงนั้นจะมีการกำหนดค่าพารามิเตอร์เมื่อสมมุติฐานศูนย์ เป็นจริงขึ้นไว้ แล้วสุ่มตัวอย่างจำนวนหนึ่งขึ้นมาทดสอบ เพื่อที่จะยอมรับหรือปฏิเสธสมมุติฐานศูนย์ แต่ในการใช้เทคนิคการวิเคราะห์ตามลำดับขั้น จะมีการกำหนดค่าพารามิเตอร์ทั้งเมื่อสมมุติฐานศูนย์เป็นจริง และค่าพารามิเตอร์เมื่อสมมุติฐานอื่นเป็นจริงขึ้นไว้ ในการทดสอบสมมุติฐานจะมีการสุ่มตัวอย่างจำนวนหนึ่งขึ้นมาทดสอบ โดยในการทดสอบถ้าไม่สามารถยอมรับสมมุติฐานตัวใดตัวหนึ่งได้ ภายใต้อัตราความคลาดเคลื่อนประเภทที่หนึ่ง และอัตราความคลาดเคลื่อนประเภทที่สองที่กำหนดขึ้นไว้ ก็จะมีการสุ่มตัวอย่างเข้ามาอีก และนำค่าที่ได้เข้าไปคำนวณร่วมกับค่าต่างๆที่มีอยู่เดิม แล้วพิจารณาต่อไปว่าจะยอมรับสมมุติฐานตัวใดตัวหนึ่งไว้หรือไม่ ถ้ายังไม่สามารถจะยอมรับสมมุติฐานตัวใดตัวหนึ่งได้ก็จะมีการสุ่มตัวอย่างเพิ่มเข้าไปอีก ทั้งนี้สามารถเขียนเป็นแผนภูมิอธิบายได้ดังนี้

แผนภูมิที่ 1 ลำดับขั้นตอนการทดสอบสมมุติฐานของเทคนิคการวิเคราะห์ตามลำดับชั้น



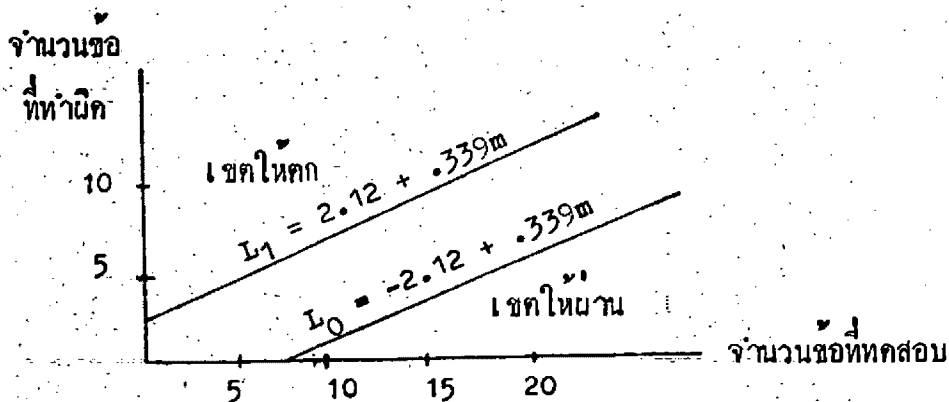
ในทางการศึกษา ได้มีการประยุกต์เทคนิคการวิเคราะห์ความล่าช้าชั้นในกรณีเฉพาะคือ การแจกแจงแบบทวินามมาใช้เพื่อลดจำนวนข้อสอบที่ใช้ในการสอบแต่ละครั้งลง โดยคาดว่าเมื่อลดจำนวนข้อสอบที่ใช้ในการสอบแต่ละครั้งลง โดยคาดว่าเมื่อลดจำนวนข้อสอบที่ใช้ในการสอบลงแล้ว การตัดสินใจประเภทบุคคลตามระดับการรอบรู้จะยังคงมีความถูกต้องไม่แตกต่างไปจากการตัดสินใจโดยวิธีแบบสอบที่มีความยาวเต็มฉบับ ทั้งนี้ในการประยุกต์จะต้องมีการกำหนดค่าพารามิเตอร์ชั้น 4 ค่า ซึ่งจะมีความหมายเฉพาะในทางการวัดผลคือ

- α คืออัตราความคลาดเคลื่อนที่เกิดจากการตัดสินใจให้ผู้รอบรู้สอบตก
 - β คืออัตราความคลาดเคลื่อนที่เกิดจากการตัดสินใจให้ผู้ไม่รอบรู้สอบได้
 - P_0 คือจำนวนสัดส่วนข้อสอบที่ยอมให้ผู้สอบทำผิดได้ และถือว่าเขาเป็นผู้รอบรู้
 - P_1 คือจำนวนสัดส่วนข้อสอบที่ให้ผู้สอบทำผิด จะถือว่าเขาเป็นผู้ไม่รอบรู้
- เช่น ถ้ากำหนด $\alpha = .05$ $\beta = .05$
 $P_0 = .20$ $P_1 = .50$

จะได้เส้นสมการ คือ $L_1 = 2.12 + .339m$
 $L_0 = -2.12 + .339m$

โดยเขียนแผนภูมิอธิบายได้ดังนี้

แผนภูมิที่ 2 การกำหนดเส้นสมการ เพื่อตัดสินใจประเภทบุคคลตามระดับการรอบรู้ โดยการประยุกต์เทคนิคการวิเคราะห์ความล่าช้าชั้น



จากแผนภูมิจะมีความหมายว่า ถ้าผู้สอบคนใดคนหนึ่งทำข้อสอบไปได้จำนวนหนึ่ง เช่น 15 ข้อ ถ้าหากว่าในจำนวน 15 ข้อแรกนั้นทำผิดไม่เกิน 3 ข้อ จะถือว่าผู้สอบเป็นผู้รอบรู้และยุติการทดสอบ ถ้าผู้สอบทำผิดตั้งแต่ 8 ข้อขึ้นไปจะถือว่า เป็นผู้ไม่รอบรู้และยุติการทดสอบเช่นกัน แต่ถ้าผู้สอบทำผิดอยู่ในช่วง 4 - 7 ข้อก็จะยังไม่มี การตัดสินผลว่า ผู้สอบเป็นผู้รอบรู้หรือไม่รอบรู้ แต่จะเปิดโอกาสให้ทำแบบทดสอบต่อไปอีกระยะหนึ่ง เพื่อตัดสินระดับการรอบรู้

การประยุกต์เทคนิคการวิเคราะห์ความล่าช้าขั้น เพื่อลดความยาวของแบบสอบปรากฏในปี 1946 โดยโคว์เดน (Cowden 1946) ได้ทดลองนำมาใช้กับผลการสอบของนักศึกษาที่เรียนวิชาสถิติในมหาวิทยาลัยแคลิฟอร์เนีย โดยใช้แบบสอบถูก-ผิดที่มีความยาว 200 ข้อ กำหนด $\alpha = .30$ $\beta = .40$ $P_0 = .20$ และ $P_1 = .10$ ในการวิเคราะห์เน้นตรงผลการตัดสินผู้สอบได้ เกรด ก และเกรด จ ในที่ทุกระยะความยาว 20 ข้อ เป็นหลัก พบว่าการตัดสินเมทคนได้ เกรด ก และ จ จะทำได้เร็วกว่าคนที่ผลการเรียนในระดับกลาง โคว์เดนได้ตั้งคำถามเพื่อการศึกษาต่อไปนี้ว่า

1. จำนวนข้อสอบที่ควรนำมาใช้ควรเป็นสักกี่ข้อจึงจะสามารถตัดสินใจได้
2. ผลที่ได้จากการตัดสินใจจัดประเภทบุคคลตามระดับการรอบรู้ดังกล่าว มีความถูกต้องเพียงใด

ในปี 1950 มูแนน (Moonan 1950) ได้นำหลักการเดียวกันนี้มาใช้กับนักศึกษาจำนวน 39 คน ที่เรียนวิชาสถิติพื้นฐานภาคบรรยาย โดยใช้แบบสอบที่มีความยาว 75 ข้อ เป็นแบบสอบ 5 ตัวเลือก พบว่าเทคนิคนี้สามารถจำแนกผู้สอบที่มีระดับความสามารถสูงหรือทำได้อย่างรวดเร็ว คำสั่งประสิทธิภาพสัมพันธ์ของแบบสอบย่อยที่ระดับ α , β , P_0 และ P_1 ต่างกัน 5 ลักษณะ กับคะแนนจากแบบสอบเต็มฉบับมีค่าโดยเฉลี่ยสูงถึง 0.90 แต่พบว่าจากการวิเคราะห์ความแปรปรวนค่ามัธยฐานและควินิลในส่วนของคะแนนจากแบบสอบต่างกัน 5 ลักษณะ เมื่อได้รับการปรับแล้วมีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 มูแนนได้ตั้งคำถามเพื่อการศึกษาต่อไปนี้

1. เมื่อมีการกำหนดช่วงเวลาที่จะใช้ทดสอบและอัตราการความคลาดเคลื่อนที่กำหนด จำนวนข้อสอบที่ควรจะใช้กับนักเรียนทุกคนในการสอบแต่ละครั้งที่จะทำให้เกิดการจัดประเภทบุคคลตามระดับการรอบรู้ถูกต้องมากที่สุดควรจะเป็นจำนวนเท่าใด

2. เป็นไปไ้หรือไม่ที่จะมีความสัมพันธ์ระหว่างเวลาในการทดสอบกับระดับความสามารถของผู้สอบ

3. จำนวนข้อสอบที่จำเป็นจะต้องใช้ เพื่อที่จะแยกความแตกต่างระหว่างกลุ่มสูงกับกลุ่มต่ำควรจะเป็นจำนวนเท่าใด

ทอม่า ฟิสก์และโจนส์ (Fiske and Jones 1958) ได้เขียนบทความเผยแพร่การนำเอาแนวคิดนี้มาใช้กับงานวิจัยทางการศึกษา โดยบุคคลทั้งสองได้เสนอว่า เทคนิคการวิเคราะห์ตามลำดับชั้น ควรจะนำมาใช้ในกรณีต่อไปนี้

1. เป็นการตัดสินใจเกี่ยวกับค่าพารามิเตอร์สองค่า
2. ในการเก็บรวบรวมข้อมูล เสียค่าใช้จ่ายสูงและต้องการประหยัดค่าใช้จ่ายถึงกลา

3. จำนวนข้อมูลที่จะนำมาใช้เพื่อการตัดสินใจมีอย่างไม่จำกัด

เคลียร์ ลินน์ และรอก (Cleary, Linn and Rock 1968) ได้กล่าวอ้างถึงงานวิจัยของแพทเทอร์สัน (Patterson) ซึ่งทำไว้ในปี 1962 จากการใช้ข้อมูลที่มีสมมุติฐาน เพื่อเปรียบเทียบผลการวิเคราะห์โดยใช้เทคนิคการวิเคราะห์ตามลำดับชั้น กับการทดสอบโดยวิธีปกติพบว่า แม้ว่าการกระจายของความสามารถในประชากรจะมีลักษณะต่างกันออกไป หลักการดังกล่าวนี้สามารถจัดประเภทบุคคลที่มีความสามารถสูงสุดและต่ำสุดได้ถูกต้องมากกว่าการทดสอบในภาวะปกติ และจากผลการวิจัยของบุคคลทั้งสามซึ่งมุ่งศึกษาเปรียบเทียบวิธีการจัดแบ่งระดับความสามารถของนักเรียนเกรด 11 จำนวน 4,885 คน ออกเป็น 4 กลุ่มตามระดับความสามารถ โดยใช้ข้อมูลการตอบแบบสอบถามจำนวน 190 ข้อ เปรียบเทียบผลการจัดประเภทบุคคลโดยใช้วิธีการจัดแบ่ง 4 วิธี คือ

1. การทดสอบที่เป็นลักษณะสองขั้นตอน วิธีการนี้ใช้การแบ่งผู้สอบออกเป็น 2 กลุ่ม โดยครั้งแรกใช้ข้อสอบที่มีความยากตรง 0.50 จำนวน 10 ข้อ เป็นหลักในการแบ่ง จากนั้นในแต่ละกลุ่มจะได้รับการแบ่งออกเป็น 2 กลุ่มเท่าๆกัน โดยใช้ข้อสอบอีกกลุ่มละ 10 ข้อ

2. การจัดสอบจากข้อสอบที่มีความยากแตกต่างกันมาก วิธีการนี้ใช้ข้อสอบจำนวน 20 ข้อ ที่มีค่าความยากจากช่วง 0.19 - 0.91 เพื่อจำแนกผู้เรียนออกเป็น 4 กลุ่ม

3. การใช้ผลความแตกต่างจากการตอบภายในกลุ่ม วิธีการนี้แบ่งผู้สอบออกเป็น 4 กลุ่มจากคะแนนเต็มฉบับ แล้วคัดเลือกเอาข้อสอบที่มีสัดส่วนการตอบถูกจากกลุ่มค่าไปหา กลุ่มสูงที่เพิ่มขึ้นตามลำดับที่ต่ำที่สุดจำนวน 20 ข้อ เป็นเกณฑ์

4. การใช้เทคนิคการวิเคราะห์ตามลำดับชั้น วิธีการนี้ใช้ข้อสอบที่มีค่าสัมประสิทธิ์สหสัมพันธ์กับคะแนนจากแบบสอบเต็มฉบับสูงสุดจำนวน 23 ข้อ เป็นเกณฑ์ในการจัดแบ่ง

ผลการศึกษาจากข้อมูลย้อนหลังที่เก็บแยกไว้เพื่อการตรวจสอบ (Cross Validation) จำนวน 2,407 คน พบว่าวิธีการต่างๆ ให้ผลการจัดประเภทนักเรียนได้ถูกต้อง ตรงกับวิธีการจัดประเภทจากคะแนนเต็มฉบับเป็น 59.12% 60.90% 71.20% และ 72.70% ตามลำดับ ทอมาปรากฏในรายงานของ ลินน์ รอค และเคลียร์ (Linn, Rock and Cleary 1972) ได้อ้างถึงงานวิจัยของกรีน (Green) ซึ่งทำในปี 1971 พบว่าการวิเคราะห์ตามลำดับชั้นสามารถลดเวลาในการสอบลงไปได้ครึ่งหนึ่งจากเวลาที่ใช้ปกติ และปรากฏผลจากการวิจัยของบุคคลทั้งสาม เมื่อใช้คะแนนผลการทดสอบกับนักศึกษา ระดับวิทยาลัยจำนวน 4,840 คนที่มีอยู่ในรายวิชาต่างๆ คือ ภาษาอังกฤษ วิทยาศาสตร์-ธรรมชาติ วิชาละ 100 ข้อ และคณิตศาสตร์จำนวน 75 ข้อ พบว่า ณ ที่ระดับความถูกต้อง ในการตัดสินใจจำแนกผู้สอบออก เป็นผู้รอบรู้และผู้ไม่รอบรู้ ได้ถูกต้องเท่ากัน การวิเคราะห์ตามลำดับชั้นสามารถให้ข้อสอบเพียงครึ่งหนึ่งของการสอบตามปกติ

การประยุกต์เทคนิคการวิเคราะห์ตามลำดับชั้น เพื่อลดความยาวของแบบสอบอิงเกณฑ์ การประยุกต์เทคนิคการวิเคราะห์ตามลำดับชั้น เพื่อลดความยาวของแบบสอบในช่วงปี 1946 - 1972 ที่กล่าวมานั้น เป็นการมุ่งเน้นการลดความยาวของแบบสอบอิงกลุ่ม เมื่อแนวคิดของการพัฒนาแบบสอบอิงเกณฑ์ได้ เริ่มพัฒนามากขึ้น สเตคไลน์และพิริยานุวัฒน์ (Stecklein and Pitayanuwat 1975) ได้อภิปรายชี้แนะการนำเอาเทคนิคการวิเคราะห์ตามลำดับชั้น มาใช้กับการทดสอบในระบบอิงเกณฑ์ไว้ มีสาระสำคัญ 2 ประการ คือ

1. แบบสอบอิงเกณฑ์ที่มีความเป็นเอกพันธ์ในเนื้อหาสูง ทั้งมีการระบุระดับความสามารถของผู้สอบที่แบบสอบต้องการวัดไว้ชัดเจน ดังนั้นปัญหาที่ว่าข้อสอบที่สุ่มขึ้นมาจะไม่มีความเป็นตัวแทนของเนื้อหาทั้งหมดไป

2. ในภาวะของการเรียนรอบรู้ สเตคไลน์และพิริยานุวัฒน์กล่าวไว้ว่า ข้อได้เปรียบของแบบสอบที่มีระดับความยาวพอเหมาะคือสามารถที่จะตัดสินใจผู้สอบ สอบได้หรือ

สอบตก โดยใช้ข้อสอบเพียงจำนวนน้อยข้อ เพื่อว่านักเรียนที่มีระดับความรู้ความที่ต้องการสามารถที่จะเลื่อนไปทำกิจกรรมต่อไป หรือเรียนบทเรียนต่อไปได้ สำหรับผู้ที่ยังไม่พร้อมที่จะรับการทดสอบก็จะได้ไม่ต้องเสียเวลาไปกับการสอบ แต่ควรมุ่งใช้เวลาไปกับการเรียนซ่อม เสริมหรือฝึกปฏิบัติงานแทน

นอกจากนั้น สเตคไลน์ยังได้อภิปรายชี้แนะจากผลการวิจัยที่พบอีกด้วยว่า แม้ในทางทฤษฎีค่าอำนาจจำแนกน่าจะเป็นดัชนีที่สำคัญที่สุดที่จะนำมาใช้ในการวิเคราะห์ตามลำดับชั้น แต่จากการศึกษาพบว่าปัจจัยที่น่าจะมีความสำคัญกว่า ได้แก่ ค่าความยากของข้อสอบ คือพหุมีข้อสอบที่ง่ายและมีค่าอำนาจจำแนกที่คนกลุ่มอ่อนตอบไม่ได้ และมีข้อสอบที่ยากและคนกลุ่มเก่งตอบไม่ได้ เงื่อนไขสำคัญจึงอยู่ที่ว่า ข้อสอบที่น่ามาใช้ไม่ควรจะมีค่าความยากแตกต่างกันมากจนเกินไป

งานวิจัยที่โคเคนความาทั้งหมด จะมีลักษณะร่วมกันประการหนึ่งคือเป็นการศึกษาความถูกต้องของการใช้เทคนิคการวิเคราะห์ตามลำดับชั้นโดยนำเอาข้อมูลที่มีอยู่มาศึกษาหรือเป็นการศึกษาย้อนรอย ไม่ได้มีการศึกษาในสถานการณ์จริง โดยเฉพาะในการศึกษาปฏิบัติการของผู้สอบซึ่งได้รับการทดสอบในสภาพดังกล่าว หรือศึกษาในภาวะที่ผู้สอบจะสามารถทราบผลการตอบของตนในทันทีเมื่อเลือกตัวเลือกของข้อคำถามหนึ่งๆ ในปี 1980 สเตคไลน์ พิธิยานุวัฒน์ และลิงค์ (Stecklein, Pitayanuwat and Ling 1980) ได้ศึกษาหาความยาวของแบบสอบที่เหมาะสมโดยใช้เทคนิคการวิเคราะห์ตามลำดับชั้น โดยกำหนด

$\alpha = .05$, $\beta = .10$ ค่า $P_1 = .45$ และค่า $P_0 = .25$ จากนั้นได้กำหนดจำนวนข้อสอบที่ถ้าผู้สอบตอบถูกต้องจะได้ เกรด k คือ ทำถูกต้อง 8 ใน 8 ข้อแรก ทำถูกต้อง 10 ใน 11 ข้อแรก ทำถูกต้อง 12 ใน 14 ข้อแรก ทำถูกต้อง 14 ใน 17 ข้อแรก และทำถูกต้อง 16 ใน 20 ข้อแรก อย่างใดอย่างหนึ่ง จากการศึกษาพบว่าผลการตัดสินใจในแบบสอบที่ช่วงความยาวซึ่งจำกัดลงมีความถูกต้องในการจัดประเภทบุคคลเมื่อเทียบกับผลการตัดสินใจจากแบบสอบเต็มฉบับสูง เช่น ในรายวิชาแรกสามารถตัดสินใจตัดสินผลการทดสอบจากการใช้เทคนิคการวิเคราะห์ตามลำดับชั้นได้ถึง 68 ใน 152 คน โดยไม่มีความคลาดเคลื่อนในการให้ผู้รอบรู้สอบไม่ผ่านเลย และมีความคลาดเคลื่อนในการให้ผู้ไม่รอบรู้สอบผ่าน 3 คน อีกรายวิชาหนึ่ง สามารถตัดสินใจได้ 28 คนในจำนวนผู้เข้าสอบ 53 คน โดยไม่มีความ-

คลาดเคลื่อนในการให้ผู้นอบบรรุสอบไม่ผ่านเลยเช่นกัน และมีความคลาดเคลื่อนในการให้ผู้นอบบรรุสอบผ่าน 3.6% และรายวิชาสุดท้ายสามารถตัดสินใจได้ 27 ใน 31 คน

จากปัญหาการกำหนดความยาวที่เหมาะสมของแบบสอบอิงเกณฑ์ และผลของการประยุกต์เทคนิคการวิเคราะห์ตามลำดับชั้นในทางการศึกษามีประเด็นที่น่าสนใจคือ

ความข้อเขียนของโปแฟมที่กล่าวไว้ว่า วิธีการกำหนดเกณฑ์มาตรฐานของแบบสอบอิงเกณฑ์นั้นอย่างน้อยควรจะคำนึงถึงปัจจัยต่างๆ คือ ระดับความคาดหวัง หรือเกณฑ์มาตรฐานในการตัดสินใจ สัดส่วนของความคลาดเคลื่อนในการให้ผู้นอบบรรุสอบไม่ผ่าน และผู้นอบบรรุสอบได้ ตลอดจนระดับความสามารถที่แท้จริงของผู้สอบแล้ว เทคนิคการวิเคราะห์ตามลำดับชั้นสามารถที่จะแก้ปัญหาดังกล่าวได้ คือ

1. การกำหนดเกณฑ์มาตรฐานและสัดส่วนของความคลาดเคลื่อนทั้งสองประเด็นดังกล่าวที่จะใช้ในการสร้างสมการเพื่อกำหนดจำนวนข้อสอบที่ผู้สอบจะต้องทำได้หรือไม่ได้ก่อนที่จะได้รับการตัดสินว่าเป็นผู้นอบบรรุสอบหรือไม่สามารถที่จะทำได้โดยอิสระในการทดสอบแต่ละครั้ง หรือในทางคณิตศาสตร์จะมีความหมายว่า การกำหนดค่า α , β , P_0 และ P_1 สามารถกระทำโดยอิสระตามความต้องการของผู้ใช้ในแต่ละสถานการณ์ โดยไม่ผูกพันกับเงื่อนไขการคำนวณ

2. ในประเด็นของความยาวของแบบสอบ โควด์น (Cowden) แพทเทอร์สัน (Patterson) เคลียร์, ลินน์ และ รอค (Cleary, Linn and Rock) ต่างก็รายงานผลการวิจัยที่ทำในสถานที่ต่างกัน ต่างกลุ่มตัวอย่าง และต่างเวลา แต่ให้ผลตรงกันว่าสามารถลดความยาวของแบบสอบลงได้ โดยเฉพาะในกลุ่มที่มีความสามารถสูงหรือต่ำกว่า เกณฑ์นั้นสามารถที่จะใช้แบบสอบที่สั้นกว่าผู้สอบที่มีความสามารถยังตัดสินใจไม่ได้อีกด้วย

3. เทคนิคการวิเคราะห์ตามลำดับชั้น ไม่ได้นั้นตรงค่าประจำข้อสอบดังเช่นวิธีการของทฤษฎีการตอบสนองข้อสอบ (Item Response Theory) หลักการสำคัญของการใช้ข้อสอบของแบบสอบยังคงยึดเอาความตรงตามเนื้อหา เป็นสำคัญ

4. แนวทางการประยุกต์วิธีการดังกล่าวนี้เพื่อกำหนดความยาวของแบบสอบอิงเกณฑ์ น่าจะกระทำในลักษณะเดียวกันกับที่ โควด์น (Cowden) มูนาน (Moonan) สเตคไลน์, พิธิยานุวัฒน์ และลิงก์ (Stecklein, Pitiyanuwat and Ling) เคยทำไว้คือตรวจสอบตัดสินใจจากช่วงความยาวของแบบสอบเป็นช่วงๆ เหมือนจุดตรวจสอบ มากกว่าที่จะใช้

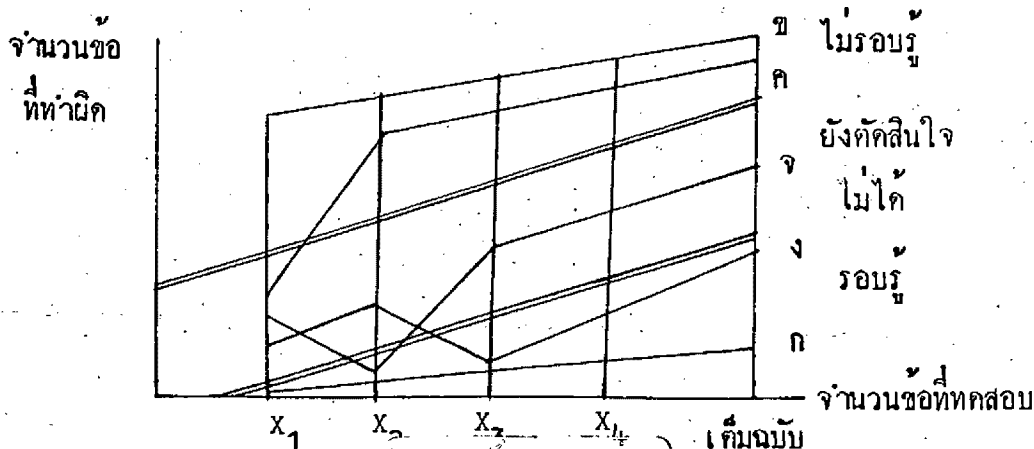
การตรวจสอบทุกๆครั้งที่ผู้สอบได้ตอบข้อคำถามหนึ่งๆ

5. ในรายงานการวิจัยเกี่ยวกับการประยุกต์เทคนิคการวิเคราะห์ตามลำดับชั้น เพื่อลดความยาวของแบบสอบถาม การตรวจสอบความถูกต้องในการตัดสินใจ ณ ที่ระดับความยาวต่างๆดำเนินการโดยใช้การตรวจสอบกับความสอดคล้องในการตัดสินใจให้ผู้สอบผ่านหรือไม่ผ่าน ณ ที่ระดับความยาวหนึ่งๆกับตัวเกณฑ์ คือผลการตัดสินใจให้ผู้สอบผ่านหรือไม่ผ่าน จากแบบสอบถามที่มีความยาวเต็มฉบับ

จากแนวคิดข้างต้น การที่จะใช้เทคนิคการวิเคราะห์ตามลำดับชั้นเพื่อกำหนดความยาวของแบบสอบถามซึ่งเกณฑ์ที่น่าจะสามารถดำเนินการได้ในลักษณะเดียวกัน คือถ้าสามารถสร้างแบบสอบถามที่ใหม่ที่มีความยาวมากพอที่จะถือได้ว่า สามารถตรวจสอบวัดความรู้ของผู้สอบในโดเมนนั้นๆได้อย่างถูกต้องแล้ว การที่จะลดความยาวของแบบสอบถามก็สามารถกระทำได้โดยการตรวจสอบความสอดคล้องของการตัดสินใจให้ผู้สอบผ่านหรือไม่ผ่าน หรือมีความสามารถในระดัวยังตัดสินใจไม่ได้ โดยใช้ผลการตัดสินใจจากแบบสอบถามที่มีความยาวเต็มฉบับซึ่งถือเป็นตัวเกณฑ์กับแบบสอบถามที่มีความยาวสั้นกว่าที่จะนำมาเป็นตัวอิงเกณฑ์

อย่างไรก็ดี หลักการดังกล่าวนี้ ขณะนี้ยังไม่มีการดำเนินการค้นคว้าหาหลักฐานความเป็นไปได้ โดยการใช้อยู่จริงประจักษ์มายืนยัน นอกจากนี้ก็ยังมีปัญหาที่สำคัญประการหนึ่งในการนำเอาเทคนิคการวิเคราะห์ตามลำดับชั้นมาใช้ในการกำหนดความยาวของแบบสอบถาม คือภาวะความไม่คงที่ในการตัดสินใจผลที่อาจจะเกิดขึ้นได้ในระยะความยาวซึ่งสั้นเกินไปของแบบสอบถามที่กำหนดขึ้น เป็นจุดตรวจสอบ โดยอธิบายได้จากแผนภูมิดังนี้

แผนภูมิที่ 3 ความคงที่ในการตัดสินใจจัดประเภทบุคคลตามระดับการรอบรู้ ณ ช่วงความยาวต่างๆของแบบสอบถามโดยการวิเคราะห์ตามลำดับชั้น



จากแผนภูมิกังกล่าว จะมีความหมายว่า

1. ผลการสอบของ ก มีความคงที่ตลอดทุกช่วงความยาวที่ตรวจสอบ คือ สอบผ่านตลอด แสดงชี้ว่า ก เป็นผู้รอบรู้ ส่วนผลการสอบของ ข มีความคงที่ในทุกๆช่วงความยาวที่ตรวจสอบเช่นกัน คือ สอบไม่ผ่านตลอด แสดงชี้ว่า ข เป็นผู้ไม่รอบรู้

2. ผลการสอบของ ค มีความไม่คงที่ในระยะที่ความยาวสั้นๆของการสอบ คือ ในตอนต้นผลการสอบจะยังอยู่ในภาวะยังไม่ตัดสินใจ ก่อนที่จะปรากฏชี้ว่าเป็นผู้ไม่รอบรู้ ส่วนผลการสอบของ ง มีความไม่คงที่ในช่วงความยาวสั้นๆของการสอบ คือ ตอนต้นจะอยู่ในภาวะยังไม่ตัดสินใจไม่ได้ ก่อนที่จะปรากฏชี้ในภายหลังว่า เป็นผู้รอบรู้

3. ผลการสอบของ จ ปรากฏในตอนต้นว่า เป็นผู้รอบรู้ ก่อนที่จะปรากฏชี้ในภายหลังว่า ผลการสอบของเขามีความคงที่ คือมีความสามารถอยู่ในระดับต่ำกว่าเกณฑ์มาตรฐานที่กำหนด แต่ในขณะที่เดียวกันเขาก็มีความสามารถสูงกว่าเกณฑ์มาตรฐานขั้นต่ำที่กำหนด ทั้งนี้แล้วก็ควรจะต้องถือว่า ผลการสอบของ จ อยู่ในภาวะที่ยังตัดสินใจไม่ได้

จากแผนภูมิกังกล่าว เมื่อแบ่งระดับการรอบรู้ของผู้สอบออกเป็น 3 ระดับ คือ รอบรู้ ยังตัดสินใจไม่ได้ และไม่รอบรู้ เมื่อใช้คะแนนจากแบบสอบเต็มฉบับเป็นตัวเกณฑ์ในการตัดสินใจแล้วพิจารณาต่อไปว่า ทรงความยาวของแบบสอบตรงจุด x_1 x_2 x_3 และ x_4 จุดใดจะเป็นจุดของช่วงความยาวแบบสอบที่สั้นและให้ผลในการตัดสินใจจัดประเภทบุคคลตามระดับการรอบรู้เป็นเช่นเดียวกันกับเมื่อตัดสินใจโดยใช้คะแนนจากแบบสอบเต็มฉบับ ดังนั้นจะพบว่า ถ้าตัดสินใจใช้ความยาวตรงจุด x_1 ก็จะเกิดความคลาดเคลื่อนในการจัดประเภทบุคคลตามระดับการรอบรู้อยู่ 2 คน คือ ค และ ง ซึ่งจะสามารถสรุปให้เห็นชัดแจ้งได้จากตารางโครงสร้างดังนี้

ตารางที่ 1 การจัดประเภทบุคคลตามระดับการรอบรู้ที่ช่วงความยาว x_1

x_1 / เกณฑ์	รอบรู้	ยังตัดสินใจไม่ได้	ไม่รอบรู้
รอบรู้	ก		
ยังตัดสินใจไม่ได้		จ	ค
ไม่รอบรู้			ข

จากตาราง ๗ ความยาว x_1 ผลการตัดสินใจจัดประเภทบุคคลตามระดับการรอบรู้ของ ก จ และ ข จะสอดคล้องกับผลการตัดสินใจเมื่อใช้คะแนนจากแบบสอบถามเต็มฉบับ คือ ก เป็นผู้รอบรู้ จ เป็นผู้ที่มีความสามารถในระดับยังตัดสินใจไม่ได้ และ ข เป็นผู้ไม่รอบรู้ เรียกปรากฏการณ์ที่เกิดขึ้นเช่นนี้ว่า "ความสอดคล้องที่ปรากฏ" ซึ่งตามตารางโครงสร้างจะเกิดขึ้นในแนวทะแยงทั้งหมด และที่ความยาว x_1 จะเกิดความคลาดเคลื่อนในการตัดสินใจจัดประเภทบุคคลตามระดับการรอบรู้เมื่อเทียบกับแบบสอบถามเต็มฉบับอยู่ 2 คน คือ ง และ ค คือที่ความยาว x_1 ตัดสิน ง เป็นผู้ที่มีความสามารถในระดับยังตัดสินใจไม่ได้ แต่แบบสอบถามเต็มฉบับตัดสินว่า ง เป็นผู้รอบรู้ และตัดสินว่า ค เป็นผู้ที่มีความสามารถในระดับยังตัดสินใจไม่ได้ แต่แบบสอบถามเต็มฉบับตัดสิน ค ว่าเป็นผู้ไม่รอบรู้ เรียกปรากฏการณ์เช่นนี้ว่า "ความคลาดเคลื่อนที่ปรากฏ" ซึ่งจะปรากฏอยู่นอกแนวทะแยงมุมในตารางโครงสร้าง

โดยการศึกษาในลักษณะเดียวกันนี้ ถ้าหากตัดสินใจใช้ความยาวตรงจุด x_2 ก็จะทำให้เกิดความคลาดเคลื่อนในการจัดประเภทบุคคลตามระดับการรอบรู้อยู่ 2 คน คือ ค และ จ แต่ถ้าใช้ความยาวตรงจุด x_3 และ x_4 การตัดสินใจจัดประเภทบุคคลตามระดับการรอบรู้จะมีความสอดคล้องกับผลที่ตัดสินใจเมื่อใช้คะแนนจากแบบสอบถาม หรือจัดสรุปได้ในตารางโครงสร้างต่อไปนี้

ตารางที่ 2 การจัดประเภทบุคคลตามระดับการรอบรู้ที่ช่วงความยาว x_3 และ x_4

$x_3 : x_4$	เต็มฉบับ รอบรู้	ยังตัดสินใจไม่ได้	ไม่รอบรู้
รอบรู้	ก จ		
ยังตัดสินใจไม่ได้		จ	
ไม่รอบรู้			ข ค

ซึ่งจากตารางโครงสร้างดังกล่าว จะไม่ปรากฏว่ามีใครได้รับการจัดประเภทอยู่นอกแนวทะแยง แสดงว่าไม่มีความคลาดเคลื่อนในการตัดสินใจ เกิดขึ้นเลย

จากปรากฏการณ์ที่สมมุติขึ้นดังกล่าว การกำหนดความยาวของแบบสอบถาม ๖ ข้อที่ควร
จะใช้ตรงช่วงความยาว x_3 เพราะไม่ก่อให้เกิดความคลาดเคลื่อนในการจัดประเภทบุคคล
ตามระดับการรอบรู้ และเป็นช่วงความยาวที่สั้นกว่าช่วง x_4

อย่างไรก็ดี ในทางปฏิบัติจริงย่อมจะเป็นไปได้ว่า ผู้ใช้แบบสอบถามจะยอมให้เกิด
ความคลาดเคลื่อนในการจัดประเภทบุคคลตามระดับการรอบรู้ขึ้นไต่บ้างเล็กน้อย เช่น อาจจะ
ยอมให้เกิดขึ้น 1 คน 5 คน หรือ 10 คน ในจำนวนผู้เข้าสอบ 100 คน โดยยังคงถือว่าการ
จัดประเภทบุคคลตามระดับการรอบรู้ที่ช่วงความยาวนั้นๆ ของแบบสอบถามยังคงมีประสิทธิภาพสูง
เรียกความคลาดเคลื่อนที่ยอมให้ เกิดขึ้นไต่ว่า "ความคลาดเคลื่อนที่กำหนด"

จากผลงานวิจัยและแนวคิดที่กล่าวมาแล้วทั้งหมด มีประเด็นที่พิจารณาคือ

การใช้เทคนิคการวิเคราะห์ตามลำดับขั้น เพื่อจัดประเภทบุคคลตามระดับการรอบรู้
พบว่าสามารถที่จะลดความยาวของแบบสอบถามที่ใช้ในชั้นเรียนลงได้ โดยวิธีการตรวจสอบที่ช่วง
ความยาวสั้นๆ เป็นช่วงๆ และใช้คะแนนจากแบบสอบถามเต็มฉบับเป็นตัวเกณฑ์ตัดสินความถูกต้อง
ในการตัดสินใจจัดประเภทบุคคลตามระดับการรอบรู้ ทั้งนี้ ลักษณะการประยุกต์แนวคิดนี้มาใช้
เป็นมาตรการในการกำหนดความยาวที่เหมาะสมของแบบสอบถาม ๖ ข้อ เพื่อจัดประเภทบุคคลตาม
ระดับการรอบรู้นั้นก็ยังคงมีประเด็นที่สำคัญ คือ ความคลาดเคลื่อนในการจัดประเภทบุคคลตาม
ระดับการรอบรู้ ณ ที่ระดับความยาวในช่วงต่างๆ ของแบบสอบถามว่าจะแตกต่างกันไปจากอัตราความ
คลาดเคลื่อนที่กำหนดขึ้นไว้หรือไม่ ปัญหาที่กล่าวมานี้แม้จะมีการกล่าวหาเพียงถึงอยู่ตลอดเวลา แต่
ก็ยังไม่ได้รับการหยิบยกขึ้นมาพิจารณาให้เด่นชัด และทั้งยังขาดมาตรการในการตรวจสอบทาง
สถิติให้ทราบถึงผลที่เกิดขึ้นอย่างจริงจัง ซึ่งถ้าสามารถศึกษาปัญหานี้ให้ถ่องแท้แล้ว ก็น่าที่จะสามารถ
กำหนดความยาวของแบบสอบถาม ๖ ข้อที่ตรงช่วงความยาวที่ให้เกิดความคลาดเคลื่อนในการจัดประเภท
บุคคลตามระดับการรอบรู้ของบุคคลที่น้อยที่สุดได้

วัตถุประสงค์ในการวิจัย

การวิจัยครั้งนี้มีวัตถุประสงค์ทั่วไปคือ ต้องการหาช่วงความยาวของแบบสอบถาม ๖ ข้อ
เพื่อใช้ตัดสินใจจัดประเภทบุคคลตามระดับการรอบรู้ โดยที่มีความคลาดเคลื่อนในการจัดประเภท
ต่ำสุด เมื่อใช้คะแนนจากแบบสอบถามเต็มฉบับเป็นเกณฑ์ในการตัดสินผล โดยประยุกต์เทคนิคการวิเคราะห์
ตามลำดับขั้นเป็นมาตรการทางสถิติ

การวิจัยครั้งนี้มีวัตถุประสงค์เฉพาะเพื่อตรวจสอบว่า ณ ช่วงความยาวต่างๆของแบบสอบ คือ 10 ข้อ 14 ข้อ 20 ข้อ 25 ข้อ 28 ข้อ และ 30 ข้อ ช่วงความยาวใดที่มีอัตราความคลาดเคลื่อนที่ปรากฏไม่แตกต่างกันไปจากอัตราความคลาดเคลื่อนที่กำหนดอย่างมีนัยสำคัญทางสถิติ

ข้อตกลงเบื้องต้น

ในการวิจัยครั้งนี้ มีข้อตกลงเบื้องต้นต่อไปนี้

1. โอกาสที่ผู้สอบจะตอบข้อสอบข้อหนึ่งๆถูกหรือผิด จะคงที่ไม่ว่าข้อสอบข้อนั้นๆได้รับบริการจัดเรียงให้อยู่เป็นลำดับที่เท่าใดในแบบสอบ
2. ผู้ที่มีประสิทธิภาพในการสอนสามารถที่จะระบุจำนวนข้อสอบที่จะใช้ในแต่ละรูปแบบเฉพาะประจำข้อสอบ เพื่อตัดสินผลระดับการรอบรู้ของผู้สอบในเบื้องต้นได้
3. แบบสอบที่มีความยาวเต็มฉบับสามารถที่จะประมาณคะแนนที่แท้จริงของผู้สอบภายใต้ข้อสมมุติของเนื้อหาที่กำหนดขึ้นได้
4. ข้อสอบทุกข้อในแบบสอบเต็มฉบับมีความเป็นเอกพันธ์ในเนื้อหาวิชา

สมมุติฐานในการวิจัย

จากการศึกษางานวิจัยต่างๆที่ประยุกต์เทคนิคการวิเคราะห์ทศมูลค่าดัชนีในทางการศึกษาคั้งที่ได้อ้างอิงถึงแล้ว แม้จะยังไม่มีการมุ่งใช้เพื่อกำหนดความยาวของแบบสอบอิงเกณฑ์ แต่เมื่อพิจารณาจากข้อค้นพบต่างๆเท่าที่มีการศึกษามาแล้วทั้งหมด มีข้อชี้แนะว่ากรณีที่มีผู้สอบมีความสามารถสูงหรือต่ำกว่ามาตรฐานที่กำหนดอย่างชัดเจน ผลการตอบข้อสอบในระยะสั้นๆเป็นจำนวนไม่มากนักก็จะสามารถตรวจวัดระดับความสามารถได้อย่างมีความแม่นยำสูง เมื่อเทียบกับคะแนนที่ได้จากแบบสอบเต็มฉบับ เช่น จากผลการวิจัยของมูแนน (Moonan 1950) หรือจากผลการวิจัยของเคลียร์, ลินน์ และรอก (Cleary, Linn and Rock 1968) และจากผลการวิจัยของลินน์, เคลียร์ และรอก (Linn, Rock and Cleary 1972 : 90 - 91) พบว่าเมื่อเพิ่มจำนวนข้อสอบในแบบสอบ สัดส่วนความถูกต้องในการตัดสินใจจัดประเภทบุคคลตามระดับการรอบรู้ก็จะสูงตามไปด้วย แต่ทั้งนี้ถ้าใช้จำนวนข้อสอบที่น้อยจนเกินไป คือประมาณไม่เกิน 10 ข้อ การตัดสินใจจัดประเภทบุคคลตามระดับการรอบรู้ก็ยังคงมีความคลาดเคลื่อนสูง จากข้อค้นพบที่สามารถสรุปได้ดังกล่าว เมื่อนำมาเป็นแนวคิดในการกำหนดความยาวที่เหมาะสมของแบบสอบอิงเกณฑ์ ผู้วิจัยจึงได้กำหนดสมมุติฐานไว้ดังนี้

"แบบสอบที่ช่วงความยาว 14 ข้อ 20 ข้อ หรือ 25 ข้อ ช่วงใดช่วงหนึ่งใน ความยาวทั้งสามช่วงดังกล่าว จะเป็นแบบสอบที่มีอัตราความคลาดเคลื่อนที่ปรากฏไม่แตกต่าง ไปจากอัตราความคลาดเคลื่อนที่กำหนดอย่างมีนัยสำคัญทางสถิติ"

ขอบเขตของการวิจัย

1. การวิจัยครั้งนี้มุ่งศึกษาเพื่อกำหนดความยาวของแบบสอบอิงเกณฑ์ตามเนื้อหา เรื่อง "การวิเคราะห์ข้อสอบ" ซึ่งใช้กับนักศึกษาที่ศึกษาตามหลักสูตรการฝึกหัดครู สภาการ- ฝึกหัดครู พุทธศักราช 2519
2. การศึกษาค้นคว้าดำเนินการในช่วงปีการศึกษา 2528 โดยใช้นักศึกษาวิทยาลัย- ครูสุรินทร์ ในระดับประกาศนียบัตรวิชาการศึกษาชั้นสูง ในรายวิชาเอกต่างๆ เป็นผลวิจัย
3. การศึกษาค้นคว้านี้มุ่งตรวจสอบการประยุกต์เทคนิคการวิเคราะห์ตามลำดับชั้นเป็น มาตรการทางสถิติในการกำหนดความยาวของแบบสอบ แต่จะไม่มุ่งเปรียบเทียบผลที่ได้จากการ กำหนดความยาวของแบบสอบโดยเทคนิคการวิเคราะห์ตามลำดับชั้นกับเทคนิคอื่น ๆ
4. ในการศึกษาค้นคว้านี้ ได้กำหนดอัตราความคลาดเคลื่อนในการตัดสินใจจัดประเภท บุคคลตามระดับการรอบรู้ใน 2 ลักษณะ คือ การให้ผู้นับรอบรู้สอบตก (α) และผู้นับรอบรู้สอบ ผ่าน (β) เพื่อใช้ในสมการตามเทคนิคการวิเคราะห์ตามลำดับชั้นทรงระดับ 0.05 ทั้งสอง กรณี ทั้งนี้เนื่องจากว่าในทางปฏิบัติจริง ถ้าห้องเรียนหนึ่งๆมีนักศึกษาประมาณ 40 คน ผลการ ทดสอบแต่ละครั้งที่เกิดความคลาดเคลื่อนในการจัดประเภทบุคคลผิด 2 คนนั้น น่าจะอยู่ในภาวะ วิสัยที่ยอมให้เกิดขึ้นได้ ทั้งนี้เมื่อพิจารณาจากรายงานการวิจัยที่น่าเอาเทคนิคการวิเคราะห์ตาม ลำดับชั้นมาใช้ในทางการศึกษา ส่วนใหญ่จะกำหนดอัตราความคลาดเคลื่อนทั้งสองประเภทสูงกว่า ระดับ 0.05 แทนทั้งสิ้น การกำหนดอัตราดังกล่าวจึงมีความเข้มงวดมาก
5. ในการตรวจสอบความคลาดเคลื่อนในการตัดสินใจที่ปรากฏ ได้ศึกษาโดยใช้แบบ สอบที่ช่วงความยาวต่างๆ คือ 10 ข้อ 14 ข้อ 20 ข้อ 25 ข้อ 28 ข้อ และ 30 ข้อ ตาม ลำดับ เหตุผลที่ใช้แบบสอบที่ช่วงความยาว 10 ข้อก็เพราะว่าเป็นแบบสอบที่สั้นที่สุด ถ้าแบบสอบที่ ช่วงความยาวนี้สามารถจัดประเภทบุคคลตามระดับการรอบรู้ โดยเกิดความคลาดเคลื่อนไม่เกิน อัตราที่กำหนดแล้ว แบบสอบที่ใดจะเป็นแบบสอบที่ครอบคลุมขึ้นท้ายข้อสอบที่น้อยที่สุด ซึ่งมีข้อดีในแง่ของการประหยัดค่าใช้จ่ายและเวลาในการสอบ ส่วนที่ช่วงความยาว 14 ข้อนั้นเป็นช่วงความ

ยาวที่ประกอบด้วยข้อสอบจากรูปแบบเฉพาะประจำข้อสอบรูปแบบละ 1 ข้อ และที่ช่วงความยาว 28 ข้อก็จะประกอบด้วยข้อสอบจากรูปแบบเฉพาะประจำข้อสอบรูปแบบละ 2 ข้อ ส่วนที่ช่วงความยาว 20 ข้อ 25 ข้อ และ 30 ข้อ ก็เป็นระยะความยาวที่คาดเดาได้ว่าจะได้นำมาศึกษา

ความไม่สมบูรณ์ในการวิจัย

การวิจัยครั้งนี้อาจเกิดความไม่สมบูรณ์ขึ้นได้ ทั้งนี้เนื่องจากสาเหตุ 2 ประการ คือ

1. การทดสอบประกอบการเรียนการสอนที่จัดให้กับนักศึกษาในรายวิชาเอกต่างๆ ทำในเวลาต่างกัน จึงอาจเป็นไปได้ว่านักศึกษาในรายวิชาเอกที่สอบก่อนอาจชี้แนะแนวทางของข้อสอบให้กับนักศึกษาที่สอบทีหลัง ทั้งนี้การควบคุมให้นักศึกษาทุกรายวิชาเอกสอบพร้อมกันไม่สามารถทำได้ เพราะในการศึกษาครั้งนี้แบ่งการศึกษาออกเป็นระยะของการพัฒนาแบบสอบและระยะศึกษาจริงซึ่งต้องการใช้ช่วงเวลาต่างกันในภาคการศึกษาที่ดำเนินการ และนอกจากนั้นแม้การศึกษาจะอยู่ในระยะเดียวกันการนัดหมายเพื่อทดสอบนักศึกษาพร้อมกันก็ทำไม่ได้ เพราะจะไปกระทบต่อแผนการเรียนการสอนในรายวิชาอื่นๆ ซึ่งนักศึกษาแต่ละรายวิชาเอกจะต้องเรียน

2. ในการสร้างตัวเลือกที่เป็นตัวลวงประจำข้อสอบต่างๆ ผู้วิจัยได้สร้างขึ้นโดยใช้หลักการว่า แม้ผู้สอบจะใช้วิธีการนึกคิดหาคำตอบก็จะมีตัวลวงที่เลือกตอบได้ (ดูภาคผนวก หน้า 168) เหตุนี้จึงอาจจะเป็นไปได้ว่า นักศึกษาส่วนน้อยส่วนหนึ่งที่ขาดเรียนในบางครั้งที่มีการจัดการเรียนการสอนเกิดความเข้าใจในหลักการวิเคราะห์ข้อสอบไม่ลึกซึ้งพอ หรือนักศึกษามบางคนที่ไม่รอบคอบอาจจะเลือกคำตอบผิดในการตอบข้อคำถามต่างๆนั้น เนื่องมาจากลักษณะการลวงของตัวเลือก ซึ่งจะมีผลทำให้คะแนนที่สอบได้จากแบบสอบเต็มฉบับต่ำกว่าคะแนนที่ควรจะได้ตามระดับความสามารถที่แท้จริง

คำนิยามศัพท์

แบบสอบ เต็มฉบับ	หมายถึง	แบบสอบเรื่อง "การวิเคราะห์ข้อสอบ" จำนวน 72 ข้อ ซึ่งใช้เป็นตัวเกณฑ์ในการจัดประเภทบุคคลตามระดับการรอบรู้ และเป็นแบบสอบที่มีข้อสอบในคอนตันจำนวนหนึ่ง เป็นแบบสอบที่ช่วงความยาวต่างๆ
-----------------	---------	---

แบบสอบที่สั้นที่สุด	หมายถึง	แบบสอบที่ประกอบด้วยจำนวนข้อสอบในตอนต้นของแบบสอบ เติมฉบับจำนวนน้อยข้อ ซึ่งถ้าผู้สอบทำถูกต้องจะได้รับรางวัลประเภท เป็นผู้รอบรู้ และยุติการทดสอบ ทั้งนี้เมื่อยึดเอาเทคนิคการวิเคราะห์ความล่าช้าขึ้น เป็นเกณฑ์ตัดสิน ซึ่งในที่นี้เป็นแบบสอบที่ประกอบด้วยข้อสอบ จำนวน 10 ข้อ
เกณฑ์มาตรฐานในการให้ผ่าน	หมายถึง	จำนวนข้อสอบที่ยอมให้ผู้สอบทำผิดได้จากจำนวนข้อสอบทั้งหมดในแบบสอบ เติมฉบับ โดยถือว่าผู้สอบเป็นผู้รอบรู้ซึ่งนิยมเสนอ เป็นสัดส่วน แทนควยสัญลักษณ์ $P_0 = .20$
เกณฑ์มาตรฐานในการให้ตก	หมายถึง	จำนวนข้อสอบที่ถ้าผู้สอบทำผิดจากจำนวนข้อสอบทั้งหมดในแบบสอบ เติมฉบับ จะถือว่าผู้สอบเป็นผู้ไม่รอบรู้ ซึ่งนิยมเสนอ เป็นสัดส่วน แทนควยสัญลักษณ์ $P_1 = .40$
ผู้รอบรู้	หมายถึง	ผู้สอบที่ตอบข้อสอบจากแบบสอบ เติมฉบับผิดไม่เกินอัตรา $P_0 \times N$ เมื่อ N คือจำนวนข้อสอบทั้งหมด
ผู้ไม่รอบรู้	หมายถึง	ผู้สอบที่ตอบข้อสอบจากแบบ เติมฉบับผิดตั้งแต่อัตรา $P_1 \times N$ ขึ้นไป
ผู้สอบที่ยังคัดสินใจไม่ได้	หมายถึง	ผู้สอบที่ตอบข้อสอบจากแบบสอบ เติมฉบับผิดอยู่ในช่วงอัตรา $P_0 \times N - P_1 \times N$
แบบสอบที่ช่วงความยาวต่างๆ	หมายถึง	แบบสอบย่อยที่ประกอบด้วยข้อสอบในตอนต้นของแบบสอบ เติมฉบับที่ระยะ 10 ข้อ 14 ข้อ 20 ข้อ 25 ข้อ 28 ข้อ และ 30 ข้อ ตามลำดับ ซึ่งใช้เป็นช่วงในการพิจารณาจักประเภทบุคคลความระกับการรอบรู้

ความสอดคล้องที่ปรากฏ หมายถึง ภาวะที่ผลการตัดสินใจจัดประเภทบุคคลตามระดับการ
 ครอบรู้ของผู้เข้าสอบคนหนึ่งจากแบบสอบเต็มฉบับ กับ
 แบบสอบที่ช่วงความยาวหนึ่งๆ ได้ผลตรงกันว่า ผู้เข้า
 สอบมีความรอบรู้ในระดับใด

ความคลาดเคลื่อนที่ปรากฏ หมายถึง ภาวะที่ผลการตัดสินใจจัดประเภทบุคคลตามระดับการ
 ครอบรู้ของผู้เข้าสอบคนหนึ่งๆ จากแบบสอบเต็มฉบับกับ
 แบบสอบที่ช่วงความยาวหนึ่งๆ ได้ผลไม่ตรงกัน ทั้งนี้ ความ
 คลาดเคลื่อนที่ปรากฏยังสามารถแยกออกได้เป็น 3 ลักษณะ
 ย่อย คือ

ความคลาดเคลื่อน ในการจัดประเภท บุคคลตามระดับการ ครอบรู้ที่มีความสามารถ คาบจุก เกณฑ์มาตรฐาน หมายถึง ความคลาดเคลื่อนซึ่ง เกิดจากการจัดประเภทบุคคลตาม
 ระดับการ ครอบรู้จากแบบสอบที่ช่วงความยาวหนึ่งๆ ไม่
 ตรงกับผลการจัดประเภทบุคคลจากแบบสอบเต็มฉบับ
 อันเนื่องมาจากคะแนนจากแบบสอบเต็มฉบับที่ผู้สอบทำได้
 มีช่วงคะแนนจริงคาบจุก เกณฑ์มาตรฐาน

ความคลาดเคลื่อน ณ จุดตัดสินใจที่กำหนด หมายถึง ความคลาดเคลื่อนที่เกิดจากเทคนิคการวิเคราะห์
 ความล่าช้าขึ้น ถ้าผู้สอบทำแบบสอบที่ช่วงความยาวสั้นๆ
 ไปไคระยะหนึ่ง และผลการทดสอบยังไม่สามารถตรวจ
 วักระดับความสามารถที่แท้จริงของผู้สอบได้ จะได้รับการ
 จัดประเภทเป็น "ยังไม่ตัดสินใจ" และจะต้องดำเนินการ
 ทดสอบต่อ เพื่อค้นหาระดับความสามารถที่แท้จริงของผู้สอบ
 แต่ในการออกแบบการวิจัยครั้งนี้ถือว่าการจัดประเภท
 บุคคลตามระดับการ ครอบรู้สิ้นสุดลงในแต่ละช่วงความยาว
 ของแบบสอบที่กำหนดขึ้น เป็นจุดตรวจสอบ

ความคลาดเคลื่อนที่แท้จริง	หมายถึง	ความคลาดเคลื่อนที่นอกเหนือไปจากความคลาดเคลื่อนในการจัดประเภทบุคคลตามระดับการรอบรู้ที่ความสามารถความถูกต้องมาตรฐาน และความคลาดเคลื่อน ณ จุดทัศนคติที่กำหนด
ความคลาดเคลื่อนที่กำหนด	หมายถึง	จำนวนผู้สอบที่ยอมให้เกิดความคลาดเคลื่อนในการจัดประเภทบุคคลตามระดับการรอบรู้เมื่อใช้แบบสอบที่ช่วงความยาวหนึ่งๆ เป็นตัวเทียบกับผลการทัศนคติเมื่อใช้คะแนนจากแบบเต็มฉบับ ทั้งนี้ กำหนดไว้ว่าจะเกิดขึ้นไม่เกินอัตราร้อยละ 5 ที่ระดับความเชื่อมั่น 0.95
ความยาวที่เหมาะสมของแบบสอบอิงเกณฑ์	หมายถึง	แบบสอบอิงเกณฑ์ที่ประกอบขึ้นด้วยข้อสอบจำนวนน้อยข้อ แต่เกิดความคลาดเคลื่อนในการจัดประเภทบุคคลตามระดับการรอบรู้ไม่แตกต่างไปจากอัตราความคลาดเคลื่อนที่กำหนดอย่างมีนัยสำคัญทางสถิติ
ผู้ที่มีประสบการณ์ในการสอน	หมายถึง	<p>ผู้ที่มีคุณสมบัติดังต่อไปนี้</p> <ol style="list-style-type: none"> 1. เป็นข้าราชการประจำซึ่งทำหน้าที่เป็นอาจารย์บรรยายในสถาบันการศึกษาระดับอุดมศึกษาที่มีเนื้อหาเรื่อง "การวิเคราะห์ข้อสอบ" เป็นส่วนหนึ่งของการบรรยาย 2. มีระดับการศึกษาไม่ต่ำกว่าปริญญาโท โดยเน้นทางด้านการวัดผล สถิติ หรือ วิจัยการศึกษา 3. มีประสบการณ์ในการสอนเนื้อหาวิชาดังกล่าวมาแล้ว ไม่น้อยกว่า 5 ปี

รูปแบบเฉพาะประจำข้อสอบ หมายถึง วิธีการหนึ่งในการที่จะสร้างข้อสอบให้ได้เป็นจำนวนมาก โดยที่ข้อสอบทั้งหมดมุ่งวัดในวัตถุประสงค์เดียวกัน คำศัพท์คำนี้มุ่งใช้ให้ตรงกับคำว่า "Item Form" ในภาษาอังกฤษ เช่น ในการวัดความรู้ในเรื่องบัญญัติไตรยางค์ อาจจะสามารถกำหนดรูปแบบเฉพาะประจำข้อสอบลักษณะหนึ่งขึ้นมาได้

ซื้อสินค้ามา	X	ชิ้น	ราคา	Y	บาท		
นำเข้ามา	Z	ชิ้น	ราคา	?	บาท		
กำหนด	X	เป็นจำนวนชิ้นของสินค้าโดยที่			1	X	50
	Y	เป็นราคาของสินค้า X ชิ้น โดยที่			50	Y	200
	Z	เป็นจำนวนสินค้าที่กองการซื้อ โดยที่			10	Z	60
		โดยที่			$X \neq Z$		

ประโยชน์ที่ได้รับ

ประโยชน์ที่ได้รับจากการศึกษาคำนี้ แยกออกได้เป็น 3 ประเด็น คือ

1. ทำให้ทราบจำนวนข้อสอบที่พอเหมาะในการนำไปใช้ทดสอบ เพื่อจัดประเภทบุคคลตามระดับการรอบรู้ โดยเกิดจากความคลาดเคลื่อนที่ปรากฏไม่แตกต่างกันไปจากอัตราความคลาดเคลื่อนที่กำหนดอย่างมีนัยสำคัญทางสถิติในเรื่อง "การวิเคราะห์ข้อสอบ" ที่ใช้กับนักศึกษาฝึกหัดครู ตามหลักสูตรการฝึกหัดครู สภาการฝึกหัดครู พุทธศักราช 2519

2. ได้แนวทางหรือลำดับขั้นคอนในการปฏิบัติ เพื่อกำหนดความยาวของแบบสอบอิง-เกณฑ์ให้เหมาะกับสภาพการณ์หนึ่ง โดยใช้เทคนิคการวิเคราะห์ความลำดับขั้นเป็นมาตรการทางสถิติ

3. ได้แบบสอบอิงเกณฑ์ เรื่อง "การวิเคราะห์ข้อสอบ"