

การเปรียบเทียบคุณภาพของแบบสอบแบบเลือกตอบ ที่มีจำนวนตัวเลือกไม่เท่ากัน โดยการใช้ทฤษฎี การตอบสนองข้อสอบ

ธีระศักดิ์ อูร์จันนันท

ความเป็นมาและความสำคัญของปัญหา

การวัดผลการศึกษาโดยเฉพาะการทดสอบ เป็นการนำเครื่องมือที่ใช้วัดคุณลักษณะต่าง ๆ ของบุคคลไปเร้าเพื่อให้บุคคลนั้นตอบสนองออกมาให้สังเกตได้ ดังนั้นจึงต้องตระหนักถึงจุดเด่นและจุดด้อยของเครื่องมือแต่ละประเภท เพื่อจะได้เลือกใช้เครื่องมือได้ถูกต้อง อันจะนำมาซึ่งความถูกต้องและแม่นยำของข้อมูลที่ได้วัดได้ (สมหวัง พิธิยานุวัฒน์ 2524 : 21) เครื่องมือที่นิยมใช้กันอย่างแพร่หลายและมีบทบาทมากทางการศึกษา คือ แบบสอบ ซึ่งรูปแบบของแบบสอบที่ใช้กันอยู่ในปัจจุบันมีหลายรูปแบบด้วยกัน

ถ้าจะทำการสอบวัดความสามารถในการเขียนเสนอความคิดสร้างสรรค์ มักนิยมใช้แบบสอบแบบความเรียง ถ้าจะวัดความสามารถในการปฏิบัติหรือกระทำจริง ก็ใช้แบบสอบภาคปฏิบัติ นอกเหนือจากด้านการเขียนเสนอความคิดสร้างสรรค์และการปฏิบัติ มักนิยมใช้แบบสอบแบบจำกัดคำตอบกันอย่างแพร่หลาย

จนเกิดปัญหาในการเลือกรูปแบบของแบบสอบแบบจำกัดคำตอบ ซึ่งมีแบบถูกผิด แบบเติมคำตอบแบบจับคู่และแบบเลือกตอบ การจะเลือกรูปแบบใดนั้นแล้วแต่โอกาสและความเหมาะสม เพราะทุกรูปแบบต่างมีข้อดีและข้อเสียแตกต่างกัน แต่เมื่อเปรียบเทียบข้อดีและข้อเสียกันแล้วแบบสอบแบบเลือกตอบมีความเหมาะสมที่จะใช้มากกว่าแบบอื่น ๆ (ไพศาล หวังพานิช

2526 : 72) โดยเฉพาะอย่างยิ่งในต่างประเทศ แบบสอบที่ถือว่าเป็นแบบสอบมาตรฐาน จะเป็นแบบสอบเลือกตอบเป็นส่วนมาก

แบบสอบแบบเลือกตอบนี้ แต่ละข้อได้ กำหนดคำตอบไว้แล้ว ผู้ตอบเพียงแค่เลือกคำตอบที่ถูกต้องที่สุดหรือดีที่สุดเท่านั้น ข้อสอบแบบนี้ประกอบด้วยคำถามกับตัวเลือกที่เป็นตัวถูกและตัวลวง แบบสอบแบบเลือกตอบถ้าแบ่งออกตามลักษณะของแบบฟอร์มแบ่งได้เป็น 4 ลักษณะ คือ 1. มี คำตอบถูกเพียงคำตอบเดียว 2. มีคำตอบที่ดีที่สุดเพียงคำตอบเดียว 3. ประเภทการเปรียบเทียบ และ 4. ประเภทกลับตรงกันข้าม (Mehrens and Lehmann 1975 : 275) ถึงแม้ว่าแบบสอบแบบเลือกตอบ จะมีความเหมาะสมแต่ก็มีข้อบกพร่องและปัญหาคล้าย เหมือนกัน ปัญหาอย่างหนึ่งที่ทำให้แบบสอบแบบเลือกตอบมีคุณภาพไม่ดี อยู่ที่การสร้างคำถามและตัวเลือกไม่ดี นอกจากนี้ยังมีปัญหาเกี่ยวกับจำนวนตัวเลือกที่เหมาะสม นักวัดและประเมินผลการศึกษาหลายท่านได้สนใจ และ ทำการศึกษาวิจัยเกี่ยวกับจำนวนตัวเลือกของแบบสอบแบบเลือกตอบในตำแหน่งต่าง ๆ เพื่อให้แบบสอบแบบเลือกตอบมีคุณภาพมากที่สุด

จากการศึกษาและวิจัยคุณภาพ ของแบบสอบแบบเลือกตอบเกี่ยวกับจำนวนตัวเลือกที่เหมาะสมนั้น ได้อาศัยแนวทฤษฎีการทดสอบ

ตามแนวคลาสสิกอล (Classical Test Theory) ใช้อธิบายพบว่า มีผลการวิจัยขัดแย้งกันเกี่ยวกับค่าความเที่ยง ค่าอำนาจจำแนก และค่าความยาก ซึ่งเป็นคุณลักษณะสำคัญในการตรวจสอบคุณภาพของแบบสอบ เพราะเป็นการเปรียบเทียบคุณลักษณะของแบบสอบที่ละประเด็น ซึ่งแต่ละคุณลักษณะนี้อาจมีบางคุณลักษณะที่ถึงมาตรฐาน แต่บางคุณลักษณะไม่ถึงมาตรฐาน ดังนั้นจึงไม่สามารถที่จะสรุปคุณภาพของแบบสอบที่ชัดเจนและสมบูรณ์ ได้โดยรอบทฤษฎีนี้

ทฤษฎีการตอบสนองข้อสอบ (Item Response Theory) เป็นทฤษฎีการวัดทฤษฎีหนึ่งที่เข้ามามีบทบาทในวงการวัดและประเมินผลการศึกษาตั้งแต่ปี ค.ศ. 1936 และมีบทบาทมากขึ้นเรื่อย ๆ ในปัจจุบันเนื่องจากสามารถแก้จุดอ่อนของทฤษฎีคลาสสิกอลได้ โดยให้แนวคิดอย่างชัดเจนเกี่ยวกับการเปรียบเทียบคุณภาพของแบบสอบ (Hambleton and Swaminathan 1985:101) ซึ่งมาจากแนวคิดที่ว่า เมื่อใช้แบบสอบวัดความสามารถของบุคคล โดยให้บุคคลนั้น ตอบสนองต่อข้อสอบในแบบสอบ แล้วใช้ผลการตอบสนองต่อข้อสอบหรือผลการวัดความสามารถนั้น สรุปอ้างอิงเกี่ยวกับความสามารถที่มีอยู่ภายในบุคคล ดังนั้นแบบสอบที่มีคุณภาพจึงควรเป็นแบบสอบ

ที่ให้ผลการวัดความสามารถที่สามารถนำไปใช้ประมาณค่าความสามารถที่ต้องการวัดได้ถูกต้องแม่นยำ (Precision) มากที่สุด คชนี้ที่ใช้ประมาณค่าความสามารถที่ต้องการวัดได้ถูกต้องแม่นยำ คือ ค่าฟังก์ชันสารสนเทศ (Information Function) ซึ่งค่าฟังก์ชันสารสนเทศนี้เป็นฟังก์ชันของค่าพารามิเตอร์ของข้อสอบที่เป็นคชนี้ผสมโดยสร้างจากคชนี้ที่กำหนดคุณภาพของข้อสอบหลายๆ คชนี้ตัวด้วยกันจึงไม่ต้องเปรียบเทียบคุณลักษณะที่ละประเด็นอีกต่อไป ค่าฟังก์ชันสารสนเทศที่ใช้พิจารณาคุณภาพของแบบสอบ ได้ใช้วิธีการพิจารณาจากค่าฟังก์ชันสารสนเทศแบบสอบ (Test Information Function) ซึ่งหาได้จากผลรวมของค่าฟังก์ชันสารสนเทศข้อสอบ (Item Information Function) และเป็นค่าที่อธิบายคุณภาพของแบบสอบที่เพียงพอแล้ว โดยไม่ต้องอาศัยค่าอื่น ๆ เพิ่มเติม (สำเร็จ บุญเรือง รัตน 2527 : 130)

ค่าฟังก์ชันสารสนเทศแบบสอบจะมีค่าที่แตกต่างกัน ถ้าเราใช้ผลการวัดความสามารถด้วยแบบสอบเดียวกันประมาณค่าความสามารถในระดับที่แตกต่างกัน ณ ระดับความสามารถใดที่ค่าฟังก์ชันสารสนเทศแบบสอบสูง แสดงว่าช่วงการประมาณค่าความสามารถจะแคบ

ที่ระดับนั้นมีความถูกต้องแม่นยำมาก แต่ถ้ามีค่าน้อยหมายความว่า การประมาณค่าความสามารถแสดงว่าช่วงการประมาณค่าความสามารถจะกว้าง นั่นคือ การประมาณค่าความสามารถที่ระดับนั้นมีความถูกต้องแม่นยำน้อย ดังนั้น ถ้าเรามีแบบสอบหลายฉบับที่วัดความสามารถเดียวกันแล้ว เราก็สามารถเปรียบเทียบคุณภาพของแบบสอบเหล่านั้นได้ โดยการนำเอาค่าฟังก์ชันสารสนเทศแบบสอบของแต่ละแบบสอบที่ต้องการเปรียบเทียบ 2 ฉบับ ณ ระดับความสามารถเดียวกัน คือ ค่าประสิทธิภาพสัมพัทธ์ (Relative Efficiency) ของแบบสอบ จะเห็นว่าค่าประสิทธิภาพสัมพัทธ์ของแบบสอบนี้ใช้ค่าฟังก์ชันสารสนเทศแบบสอบซึ่งเป็นคชนี้ที่กำหนดคุณภาพของแบบสอบมาเปรียบเทียบคุณภาพของแบบสอบ ตามกรอบทฤษฎีการตอบสนองข้อสอบนี้ก็มีจุดอ่อนอยู่ที่การเปรียบเทียบคุณภาพของแบบสอบในแต่ละระดับความสามารถ แต่ก็สามารถแก้จุดอ่อนของทฤษฎีคลาสสิกได้ โดยไม่ต้องเปรียบเทียบคุณลักษณะสำคัญในการตรวจสอบคุณภาพของแบบสอบที่ละประเด็นในแต่ละแบบสอบ

วัตถุประสงค์ของการวิจัย

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อศึกษา ค่าประสิทธิภาพสัมพัทธ์ของแบบสอบ ของ

แบบ สอบแบบเลือกตอบฉบับเดียวกันที่มีการลดจำนวนตัวเลือกจาก 5 ตัวเลือก เป็น 4 ตัวเลือกและ 3 ตัวเลือก

สมมติฐานของการวิจัย

การวิจัยครั้งนี้ ผู้วิจัยได้ตั้งสมมติฐานดังนี้

สมมติฐานที่ 1 แบบสอบแบบเลือกตอบที่มี 5 ตัวเลือก เมื่อเปรียบเทียบกับแบบสอบแบบเลือกตอบที่มี 3 ตัวเลือก น่าจะมีค่าประสิทธิภาพสัมพัทธ์สูงกว่า 1 ทุกกลุ่มพลวิจัยระดับความสามารถ ซึ่งหมายความว่า แบบสอบแบบเลือกตอบที่มี 5 ตัวเลือก น่าจะมีคุณภาพสูงกว่าแบบสอบแบบเลือกตามที่มี 3 ตัวเลือกทุกกลุ่มพลวิจัยระดับความสามารถ

สมมติฐานที่ 2 แบบสอบแบบเลือกตอบที่มี 5 ตัวเลือก เมื่อเปรียบเทียบกับแบบสอบแบบเลือกตามที่มี 4 ตัวเลือก น่าจะมีค่าประสิทธิภาพสัมพัทธ์สูงกว่า 1 ทุกกลุ่มพลวิจัยระดับความสามารถ ซึ่งหมายความว่า แบบสอบแบบเลือกตอบที่มี 5 ตัวเลือก น่าจะมีคุณภาพสูงกว่าแบบสอบแบบเลือกตอบที่มี 4 ตัวเลือกทุกกลุ่มพลวิจัยระดับความสามารถ

สมมติฐานที่ 3 แบบสอบแบบเลือกตอบที่มี 4 ตัวเลือก เมื่อเปรียบเทียบกับแบบ

สอบแบบเลือกตอบที่มี 3 ตัวเลือก น่าจะมีค่าประสิทธิภาพสัมพัทธ์สูงกว่า 1 ทุกกลุ่มพลวิจัยระดับความสามารถ ซึ่งหมายความว่าแบบสอบแบบเลือกตอบที่มี 4 ตัวเลือกน่าจะมีคุณภาพสูงกว่าแบบสอบแบบเลือกตอบที่มี 3 ตัวเลือกทุกกลุ่มพลวิจัยระดับความสามารถ

ขอบเขตของการวิจัย

1. แบบสอบที่ใช้ในการวิจัยครั้งนี้ คือแบบสอบความถนัดประเภทความสามารถในการใช้เหตุผลเชิงภาษาแบบอุปมาอุปไมยสำหรับนักเรียนชั้นมัธยมศึกษาของ วิทยาลัยบุญสุวรรณ

2. ในการศึกษาค่าประสิทธิภาพสัมพัทธ์ของแบบสอบ ตามกรอบของทฤษฎีการตอบสนองข้อสอบ โมเดลที่ใช้คือ โมเดลโลจิสติกที่มีพารามิเตอร์ 3 ตัว

3. ในการวิจัยครั้งนี้ได้ศึกษาตัวแปรดังนี้

3.1 ตัวแปรต้น คือ แบบสอบแบบเลือกตอบที่มีตัวเลือกต่างกัน คือ แบบสอบแบบเลือกตอบที่มี 5 ตัวเลือก 4 ตัวเลือก และ 3 ตัวเลือก

3.2 ตัวแปรตาม คือ ค่าประสิทธิภาพสัมพัทธ์ของแบบสอบ

ข้อตกลงเบื้องต้น

1. คะแนนที่ได้จากการสอบเป็นคะแนนที่ได้จากนักเรียนที่เข้าสอบทุกคนตั้งใจทำแบบสอบอย่างเต็มความสามารถ

2. ระดับความสามารถแบ่งออกเป็น 3 กลุ่ม ดังนี้

2.1 $\theta > 1$ S.D. คือ กลุ่มพลวิจัยระดับความสามารถสูง

2.2 $\theta \pm 1$ S.D. คือ กลุ่มพลวิจัยคือ กลุ่มพลวิจัยระดับความสามารถปานกลาง

2.3 $\theta < 1$ S.D. คือ กลุ่มพลวิจัยระดับความสามารถต่ำ

การดำเนินการวิจัย

ประชากรพลวิจัย

ประชากรพลวิจัยในการวิจัยครั้งนี้ คือนักเรียนชั้นมัธยมศึกษาปีที่ 1 ภาคเรียนที่ 2 มีการศึกษา 2530 ในจังหวัดเพชรบุรี จำนวน 3,320 คน

กลุ่มตัวอย่างพลวิจัย

กลุ่มตัวอย่างพลวิจัยในการวิจัยครั้งนี้ คือนักเรียนชั้นมัธยมศึกษาปีที่ 1 ภาคเรียนที่ 2 ปีการศึกษา 2530 ในจังหวัดเพชรบุรี จำนวน 2,960 คน โดยแบ่งตอบแบบสอบแบบเลือกตอบที่มี 5 ตัวเลือก 4 ตัวเลือก และ 3 ตัวเลือก ตามลำดับดังนี้คือ 989 คน และ 985 คน และ

986 คน เนื่องจากให้เหมาะสมกับรูปแบบการวิเคราะห์ข้อมูลที่ใช้โปรแกรมโลจิสต์ 5 (Logist 5) และให้การประมาณค่าพารามิเตอร์ของข้อสอบมีความคงที่มากขึ้น ซึ่งควรใช้จำนวนผู้สอบประมาณ 1,000 คน ในการตอบแบบสอบแต่ละฉบับ (Warm 1978: 109)

เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัยครั้งนี้ เป็นแบบสอบความถนัดประเภทความสามารถในการใช้เหตุผลเชิงภาษาแบบอุปมาอุปไมยแบบเลือกตอบที่มี 5 ตัวเลือก จำนวน 32 ข้อ สำหรับนักเรียนชั้นมัธยมศึกษาของ วิบูลย์ บุญสุวรรณ และในการวิจัยได้แยกแบบสอบนี้ออกเป็น 3 ฉบับ โดยในแต่ละฉบับจะมีตอนนำหรือข้อคำถามเหมือนกันและมีจำนวนข้อเท่ากัน แต่จำนวนตัวเลือกของแต่ละข้อในแบบสอบแต่ละฉบับมีไม่เท่ากัน ซึ่งเป็นแบบสอบแบบเลือกตอบที่มี 5 ตัวเลือก 4 ตัวเลือกและ 3 ตัวเลือก ในการลดหรือตัดตัวเลือกทั้ง ได้ใช้วิธีสุ่มตัดตัวเลือกออกทีละข้อโดยการสุ่มอย่างง่าย

การเก็บรวบรวมข้อมูล

ผู้วิจัยนำแบบสอบแบบเลือกตอบทั้ง 3 ฉบับไปทดลองสอบก่อน เพื่อซักซ้อมการดำเนินการสอบรวมทั้งศึกษานโยบายต่าง ๆ ที่จะแก้ไขในการดำเนินการสอบจริง แล้วนำไป

สอบกับกลุ่มตัวอย่างพลวิจัย 2,960 คน โดยในการสอบแต่ละห้องเรียน ได้แจกแบบสอบถามเลือกตอบทั้ง 3 ฉบับ เป็นจำนวนเท่ากันหรือใกล้เคียงกัน โดยวิธีการสุ่มอย่างง่าย

การวิเคราะห์ข้อมูล

1. ตรวจสอบคุณสมบัติการวัดเพียงมิติเดียว และความเป็นอิสระในการตอบข้อสอบของแบบสอบถามเลือกตอบทั้ง 3 ฉบับ โดยใช้วิธีวิเคราะห์ตัวประกอบ

2. หาค่าพารามิเตอร์ของข้อสอบคือค่าอำนาจจำแนกของข้อสอบ ค่าความยากของข้อสอบ และค่าการเดาของข้อสอบของแบบสอบถามเลือกตอบทั้ง 3 ฉบับ โดยโปรแกรมโลจิสต์ 5

3. นำค่าพารามิเตอร์ของข้อสอบมาคำนวณหาค่าฟังก์ชันसारสนเทศแบบสอบของแบบสอบถามเลือกตอบทั้ง 3 ฉบับ ณ ระดับความสามารถต่าง ๆ

4. นำค่าฟังก์ชันसारสนเทศแบบสอบของแบบสอบถามเลือกตอบแต่ละฉบับมาคำนวณค่าประสิทธิภาพสัมพัทธ์ของแบบสอบ

ผลการวิจัย

1. แบบสอบถามเลือกตอบที่มี 5 ตัวเลือกมีคุณภาพสูงกว่าแบบสอบถามเลือกตอบที่มี 3 ตัวเลือก เฉพาะในกลุ่มพลวิจัยระดับ

ความสามารถสูงและต่ำ ส่วนแบบสอบถามเลือกตอบที่มี 3 ตัวเลือกกลับมีคุณภาพสูงกว่าแบบสอบถามเลือกตอบที่มี 5 ตัวเลือกเฉพาะในกลุ่มพลวิจัยระดับความสามารถปานกลาง

2. แบบสอบถามเลือกตอบที่มี 5 ตัวเลือกมีคุณภาพสูงกว่าแบบสอบถามเลือกตอบที่มี 4 ตัวเลือก เฉพาะในกลุ่มพลวิจัยระดับความสามารถสูงและปานกลาง ส่วนแบบสอบถามเลือกตอบที่มี 4 ตัวเลือก กลับมีคุณภาพสูงกว่าแบบสอบถามเลือกตอบที่มี 5 ตัวเลือกเฉพาะในกลุ่มพลวิจัยระดับความสามารถต่ำ

3. แบบสอบถามเลือกตอบที่มี 4 ตัวเลือกมีคุณภาพสูงกว่าแบบสอบถามเลือกตอบที่มี 3 ตัวเลือก เฉพาะในกลุ่มพลวิจัยระดับความสามารถสูงและต่ำ ส่วนแบบสอบถามเลือกตอบที่มี 3 ตัวเลือก กลับมีคุณภาพสูงกว่าแบบสอบถามเลือกตอบที่มี 4 ตัวเลือก เฉพาะในกลุ่มพลวิจัยระดับความสามารถปานกลาง

4. แบบสอบถามเลือกตอบที่มี 5 ตัวเลือกมีคุณภาพสูงสุดในกลุ่มพลวิจัยระดับความสามารถสูง รองลงไปที่คือ แบบสอบถามเลือกตอบที่มี 4 ตัวเลือกและ 3 ตัวเลือกตามลำดับ แบบสอบถามเลือกตอบที่มี 3 ตัวเลือกมีคุณภาพสูงสุดในกลุ่มพลวิจัยระดับความสามารถปานกลาง รองลงไปที่คือแบบสอบถามเลือกตอบที่มี 5 ตัวเลือกและ 4 ตัวเลือกตามลำดับ

และแบบสอบแบบเลือกตอบที่มี 4 ตัวเลือกมีคุณภาพสูงสุดในกลุ่มผลวิจัยระดับความสามารถต่ำ รองลงไปคือแบบสอบแบบเลือกตอบที่มี 5 ตัวเลือกและ 3 ตัวเลือก

อภิปรายผลการวิจัย

จากผลการวิจัยนี้พบว่า เมื่อนำแบบสอบแบบเลือกตอบทั้ง 3 ฉบับมาเปรียบเทียบคุณภาพของแบบสอบที่ละคู่ในกลุ่มผลวิจัยระดับความสามารถสูงทำให้มีผลการวิจัยที่สอดคล้องกับสมมติฐาน กล่าวคือแบบสอบแบบเลือกตอบที่มี 5 ตัวเลือกและ 4 ตัวเลือกมีคุณภาพสูงกว่าแบบสอบแบบเลือกตอบที่มี 3 ตัวเลือก และแบบสอบแบบเลือกตอบที่มี 5 ตัวเลือกมีคุณภาพสูงกว่าแบบสอบแบบเลือกตอบที่มี 4 ตัวเลือก แต่เมื่อนำมาเปรียบเทียบคุณภาพของแบบสอบในกลุ่มผลวิจัยระดับความสามารถปานกลางและต่ำนั้น ทำให้มีผลการวิจัยที่ขัดแย้งกับสมมติฐาน น่าจะเป็นเพราะว่าเมื่อพิจารณาค่าฟังก์ชันสารสนเทศแบบสอบสูงสุดของแบบสอบแบบเลือกตอบที่มี 5 ตัวเลือก 4 ตัวเลือก และ 3 ตัวเลือกอยู่ที่ระดับความสามารถ 1.6, 1.6 และ 1.4 ตามลำดับ แสดงว่าแบบสอบแบบเลือกตอบทั้ง 3 ฉบับมีความเหมาะสมในการสอบวัดกับกลุ่มผลวิจัยระดับความสามารถ

สูง เมื่อมาเปรียบเทียบคุณภาพของแบบสอบในกลุ่มผลวิจัยระดับความสามารถปานกลางและต่ำ อาจทำให้มีการแกว่ง (Fluctuation) ของคุณลักษณะภายในหรือความสามารถที่วัดนั้นไม่แจ่มแจ้งแบบปกติ จึงน่าจะทำให้มีผลการวิจัยที่ขัดแย้งกับสมมติฐาน จึงไม่สามารถสรุปได้อย่างชัดเจนในกลุ่มผลวิจัยระดับความสามารถปานกลางและต่ำ

ข้อเสนอแนะที่ได้จากการวิจัย

ข้อเสนอแนะในการนำไปใช้

ในการสร้างจำนวนตัวเลือกที่เหมาะสมของแบบสอบแบบเลือกตอบตามกรอบของทฤษฎีการตอบสนองข้อสอบเพื่อให้แบบสอบมีคุณภาพสูงสุด โดยในกลุ่มผลวิจัยระดับความสามารถสูง ควรสร้างแบบสอบแบบเลือกตอบที่มี 5 ตัวเลือก รองลงไปคือ แบบสอบแบบเลือกตอบที่มี 4 ตัวเลือกและ 3 ตัวเลือกตามลำดับ

ข้อเสนอแนะในการวิจัยเพิ่มเติม

ควรจะศึกษาวิจัยต่อไปว่า ถ้าใช้แบบสอบที่มีความเหมาะสมกับการสอบวัดความสามารถกับกลุ่มผลวิจัยระดับความสามารถปานกลางและต่ำ ผลที่ได้จะสอดคล้องกันหรือไม่

บรรณานุกรม

- ธีระศักดิ์ อัจฉนนนท์. “การเปรียบเทียบคุณภาพของแบบสอบแบบเลือกตอบที่มีจำนวนตัวเลือกไม่เท่ากัน โดยการใช้ทฤษฎีการตอบสนองข้อสอบ.”
วิทยานิพนธ์ปริญญาโทบริหารศึกษาศาสตร์ ภาควิชาวิจัยการศึกษา บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย, 2530.
- ไพศาล หวังพานิช. *การวัดผลการศึกษา*. กรุงเทพมหานคร : สำนักพิมพ์ไทยวัฒนาพานิช, 2523.
- สมหวัง พิธิยานุวัฒน์. *การประเมินโครงการประชุม : หลักการและการประยุกต์ใช้*. กรุงเทพมหานคร : สำนักพิมพ์โอเคียนส์โตร, 2524.
- สำเริง บุญเรืองรัตน์. *ทฤษฎีการวัดและการประเมินผลการศึกษา*. กรุงเทพมหานคร : สำนักทดสอบทางการศึกษาและจิตวิทยา มหาวิทยาลัยศรีนครินทรวิโรฒประสานมิตร, 2527.
- Allen, Merry J. and Yen, Wendy M. *Introduction to Measurement Theory*. California : Wadsworth, Inc., 1979.
- Hambleton, R.K., and Swaminathan Hariharan. *Item Response Theory Principles and Application*. Boston, Dordrecht, Lancaster : Kluwer-Nijhoff Publishing, 1985.
- Hulin, C.L., Drawgow, F. and Parsons, C.K. *Item Response Theory : Application to Psychological Measurement*. Homewood, Illinois : Dow Jones-Irwin, 1983.
- Lord, F.M. *Application of Item Response Theory to Practical Testing Problems*. Hillsdale N.J. : Lawrence Erlbaum Associates, 1980.
- Lord, F.M. and Novick, M.R. *Statistical Theories of Mental Test Scores*. Massachusetts : Addison-Wesley Publishing Company, Inc., 1968.
- Mehrens, William A., and Lehmann, Irvin J. *Measurement and Evaluation in Education and Psychology*. New York : Holt, Rinehart and Winston, 1975.
- Warm, Thomas A. *A Primer of Item Response Theory*. Oklahoma U.S. Coast Guard Institute, 1978.