

สรุปผลการวิจัยและข้อเสนอแนะ

การวิจัยครั้งนี้มีจุดมุ่งหมายเพื่อที่จะนำรูปแบบของราสซ์มาประยุกต์ในการออกแบบโครงสร้างสารสนเทศของแบบสอบด้วยการจัดข้อสอบที่มีระดับความยากง่ายให้เหมาะสมกับระดับความสามารถของนักเรียน โดยการศึกษาความแม่นยำ (precision) ในการวัดของแบบสอบที่จัดข้อสอบตามระดับความสามารถของนักเรียนด้วยวิธีการหาค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของนักเรียนและค่าความเที่ยงของแบบสอบ กลุ่มตัวอย่างประชากรที่ใช้ในการวิจัยได้แก่ นักเรียนชั้นประถมศึกษาปีที่ 6 ที่เรียนในภาคเรียนที่ 1 ปีการศึกษา 2529 สังกัดสำนักงานการประถมศึกษาจังหวัดลพบุรีจำนวน 1,429 คน ซึ่งสุ่มมาโดยใช้วิธีการสุ่มแบบสอบชั้นตอนในแต่ละชั้น ใช้การสุ่มอย่างง่ายจากประชากรนักเรียน 10,527 คน

เครื่องมือที่ใช้ในการวิจัยได้แก่แบบสอบผลสัมฤทธิ์วิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 ชนิดเลือกตอบ 4 ตัวเลือกที่วัดจุดประสงค์เดียวกัน 2 ฉบับ ระยะเวลา 60 ข้อ การเก็บรวบรวมข้อมูลและการวิเคราะห์นั้น ผู้วิจัยได้ใช้แบบสอบผลสัมฤทธิ์ที่ผู้วิจัยสร้างขึ้นไปทดลองสอบกับกลุ่มตัวอย่างที่สุ่มไว้ทั้งนี้หลังจากการสอบแต่ละครั้งได้มีการปรับปรุงข้อสอบด้วย ต่อจากนั้นจึงได้นำแบบสอบที่ได้ปรับปรุงแล้วไปทดสอบกับกลุ่มตัวอย่าง ผู้วิจัยนำข้อมูลที่ได้จากการทดสอบกับกลุ่มตัวอย่างมาวิเคราะห์ด้วยโปรแกรมวิเคราะห์ข้อสอบและโปรแกรมไบคาล (BICAL) เพื่อหาค่าพารามิเตอร์ของข้อสอบและพารามิเตอร์ความสามารถ โดยทำการเทียบมาตรา (Equating) แบบสอบสองฉบับที่ทดสอบกับกลุ่มตัวอย่าง 1,429 คน เมื่อได้ประชากรความสามารถชุดเดียวกันจึงทำการจัดกลุ่มความสามารถของนักเรียนซึ่งใช้เกณฑ์การพิจารณาจากช่วงค่าส่วนเบี่ยงเบนมาตรฐาน เป็นเกณฑ์ในการแบ่งกลุ่มความสามารถออกเป็น 3 ระดับ คือ กลุ่มความสามารถสูง ความสามารถปานกลาง และความสามารถต่ำ นำข้อสอบที่เหมาะสมกับรูปแบบของราสซ์ มาคัดเลือกเพื่อออกแบบจัดข้อสอบให้เหมาะสมกับระดับความสามารถ โดยนำค่าความยากของข้อสอบ (b) มาคำนวณค่าฟังก์ชันแสดงสารสนเทศของข้อสอบและแบบสอบ (Item and Test information function) และเขียนโครงสร้างสารสนเทศของแบบสอบ

(Test information curve) ปรับปรุงโค้งให้เหมาะสมกับกลุ่มความสามารถ โดยพิจารณาค่าความยากของข้อสอบที่นำมาใช้ (แบบสอบแต่ละฉบับจะมีข้อสอบที่ครอบคลุมขอบเขตเนื้อหาทั้ง 7 ขอบเขตเนื้อหาที่นำมาสอบ) เมื่อได้โค้งแสดงสารสนเทศของแบบสอบที่ผู้วิจัยต้องการจึงนำข้อสอบที่บรรจุอยู่ในโค้งไปทดสอบกับกลุ่มตัวอย่างที่สุ่มไว้ทั้ง 3 ระดับกลุ่มละ 200 คน เพื่อหาค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถและค่าความเที่ยงของแบบสอบ เมื่อได้ค่าทั้งสองแล้วนำมาเปรียบเทียบทดสอบนัยสำคัญของความแตกต่างระหว่างกลุ่มผู้สอบทั้ง 3 ระดับ โดยค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของนักเรียนใช้ Kruskal Wallis Test ทดสอบ ส่วนค่าความเที่ยงของแบบสอบจะแปลงเป็นคะแนนมาตรฐานฟิชเชอร์ซี (Fisher Z :  $Z_F$ ) ทำการทดสอบรวมด้วยค่าไคสแควร์ (Chi square:  $\chi^2$ ) ก่อน เมื่อพบความแตกต่างจึงทดสอบเป็นรายคู่ด้วยสถิติทดสอบซี (z-test) ต่อไป

#### ผลการวิจัย

1. พิจารณาค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถ
  - 1.1 เมื่อไม่จำแนกความสามารถของผู้สอบ (พิจารณาความสามารถรวม) แล้วทำการทดสอบแบบสอบที่เหมาะสมกับความสามารถต่าง ๆ กันทั้ง 3 ฉบับ คือ แบบสอบที่เหมาะสมกับความสามารถสูง ปานกลาง และต่ำ ปรากฏว่า ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถไม่แตกต่างกันอย่างมีนัยสำคัญที่ระดับ .05 แสดงว่าลักษณะการกระจายของค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของแบบสอบที่เหมาะสมกับกลุ่มความสามารถทั้ง 3 ฉบับ ไม่แตกต่างกัน
    - 1.2 เมื่อกลุ่มผู้สอบมีความสามารถต่ำและสูง ทดสอบแบบสอบที่วัดความสามารถต่าง ๆ กัน 3 ฉบับให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถแตกต่างกันที่ระดับนัยสำคัญ .01 และไม่มี ความแตกต่างกันอย่างมีนัยสำคัญที่ระดับ .05 เมื่อกลุ่มผู้สอบที่มีความสามารถปานกลางทดสอบแบบสอบทั้ง 3 ฉบับทั้งกล่าว แสดงว่าลักษณะการกระจายของค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของแบบสอบที่เหมาะสมกับความสามารถต่าง ๆ กันทั้ง 3 ฉบับแตกต่างกันเมื่อทดสอบกับกลุ่มความสามารถต่ำและสูง ส่วนลักษณะการกระจายของค่า

ความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของแบบสอบที่เหมาะสมกับ  
กลุ่มความสามารถต่าง ๆ กันทั้ง 3 ฉบับไม่แตกต่างกันเมื่อทดสอบกับกลุ่มความสามารถ  
ปานกลาง

## 2. พิจารณาความเที่ยงของแบบสอบ

2.1 กลุ่มความสามารถต่ำเมื่อทดสอบแบบสอบที่วัดความสามารถต่าง ๆ  
กันทั้ง 3 ฉบับ มีค่าความเที่ยงอยู่ระหว่าง 0.0475 ถึง 0.5344 ค่าความเที่ยงดังกล่าว  
แตกต่างกันอย่างมีนัยสำคัญที่ระดับ .01 ระหว่างแบบสอบที่เหมาะสมกับความสามารถ  
ต่ำและปานกลาง และแบบสอบที่เหมาะสมกับความสามารถต่ำและสูง ส่วนค่าความ  
เที่ยงระหว่างแบบสอบที่เหมาะสมกับความสามารถปานกลางและสูงไม่มีนัยสำคัญแต่อย่างใด

2.2 กลุ่มความสามารถปานกลางเมื่อทดสอบแบบสอบที่วัดความสามารถ  
ต่าง ๆ กันทั้ง 3 ฉบับ มีค่าความเที่ยงอยู่ระหว่าง 0.3126 ถึง 0.5006 ค่าความเที่ยง  
ดังกล่าวแตกต่างกันอย่างมีนัยสำคัญที่ระดับ .01 ระหว่างแบบสอบที่เหมาะสมกับความสามารถ  
ปานกลางและสูง ส่วนค่าความเที่ยงระหว่างแบบสอบที่เหมาะสมกับความสามารถ  
ต่ำและปานกลาง และแบบสอบที่เหมาะสมกับความสามารถสูงและต่ำแตกต่างกันอย่าง  
มีนัยสำคัญที่ระดับ .05

2.3 กลุ่มความสามารถสูงเมื่อทดสอบแบบสอบที่วัดความสามารถต่าง ๆ  
กันทั้ง 3 ฉบับ มีค่าความเที่ยงอยู่ระหว่าง 0.6450 ถึง 0.7633 ค่าความเที่ยงดังกล่าว  
แตกต่างกันอย่างมีนัยสำคัญที่ระดับ .05 ระหว่างแบบสอบที่เหมาะสมกับความสามารถ  
ต่ำและปานกลาง และแบบสอบที่เหมาะสมกับความสามารถต่ำและสูง ส่วนค่าความ  
เที่ยงระหว่างแบบสอบที่เหมาะสมกับความสามารถปานกลางและสูงไม่มีความแตกต่าง  
กันอย่างมีนัยสำคัญแต่อย่างใด

### อภิปรายผลการวิจัย

1. ผลการวิจัยเกี่ยวกับค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่า  
ความสามารถจากการวิจัยครั้งนี้พบว่า ผลการวิจัยมีความสอดคล้องกับสมมติฐานที่ 1  
กล่าวคือ ในการวิจัยครั้งนี้การคำนวณค่าความคลาดเคลื่อนมาตรฐานในการประมาณ  
ค่าความสามารถ (Standard Error of Estimate :S.E.E) ในทฤษฎีการตอบ  
สนองข้อสอบนั้นสามารถคำนวณหาได้จากสูตรที่ได้เสนอไว้หน้า 8 ซึ่งค่า S.E.E.

เป็นสัดส่วนกลับของรากที่สองของค่าฟังก์ชันแสดงสารสนเทศของแบบสอบ

(Test information function) ดังนั้นเมื่อพิจารณาจากสูตรจะเห็นได้ว่า ถ้าค่าฟังก์ชันแสดงสารสนเทศของแบบสอบมีค่าสูงที่ระดับความสามารถใดก็จะทำให้ค่า S.E.E. น้อยลงที่ระดับความสามารถนั้น ๆ และในทำนองเดียวกันถ้าค่าฟังก์ชันแสดงสารสนเทศของแบบสอบมีค่าต่ำที่ระดับความสามารถใดก็จะทำให้ค่า S.E.E. สูงขึ้นที่ระดับความสามารถนั้น ๆ เช่นกัน เมื่อพิจารณาค่า S.E.E. ในช่วงความสามารถที่ผู้วิจัยแบ่งไว้จากการวิเคราะห์แยกออกเป็น 3 กลุ่มคือ กลุ่มความสามารถสูง ความสามารถปานกลาง และความสามารถต่ำ จะเห็นได้ว่า เมื่อกลุ่มความสามารถต่ำทดสอบแบบสอบที่เหมาะสมกับความสามารถต่าง ๆ กันทั้ง 3 ฉบับ ได้ค่า S.E.E. น้อยที่สุดเมื่อกลุ่มความสามารถต่ำทดสอบแบบสอบที่เหมาะสมกับความสามารถต่ำ ซึ่งในกลุ่มความสามารถสูงก็มีลักษณะเช่นเดียวกัน แต่ในกลุ่มความสามารถปานกลางเมื่อทดสอบแบบสอบที่เหมาะสมกับความสามารถต่าง ๆ กันทั้ง 3 ฉบับ ค่า S.E.E. ไม่แตกต่างกัน ผลดังกล่าวผู้วิจัยมีความเห็นว่า การที่กลุ่มความสามารถปานกลางทดสอบแบบสอบที่เหมาะสมกับความสามารถต่าง ๆ กันทั้ง 3 แบบแล้วค่า S.E.E. ไม่แตกต่างกันอาจมีสาเหตุมาจากข้อสอบที่ได้จากการวิเคราะห์ในการวิจัยครั้งนี้ส่วนใหญ่เป็นข้อสอบที่เหมาะสมกับความสามารถปานกลาง รองลงมาได้แก่ข้อสอบที่เหมาะสมกับความสามารถต่ำและสูง ตามลำดับ ดังนั้นการออกแบบจัดข้อสอบให้เหมาะสมกับความสามารถต่าง ๆ กัน จึงมีข้อสอบที่มาจากกลุ่มความสามารถปานกลางรวมอยู่ด้วยจำนวนหนึ่งในทุกกลุ่มความสามารถ จึงทำให้โอกาสในการตอบข้อสอบได้ถูกต้องในแต่ละระดับความสามารถได้ใกล้เคียงกันมากกว่ากลุ่มความสามารถต่ำและสูง จึงอาจเป็นผลที่ทำให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถไม่แตกต่างกัน ผลจากการวิจัยครั้งนี้แสดงให้เห็นว่าการนำแบบสอบที่เหมาะสมกับความสามารถนั้น ๆ มาทดสอบกับผู้สอบที่มีความสามารถตรงกับแบบสอบที่จัดขึ้นจะลดค่าความคลาดเคลื่อน ฯ ลงได้และยังสอดคล้องกับงานวิจัยของ Hambleton และ Gruijter (1983) ที่พบว่า การใช้ข้อสอบที่มีระดับความยากเหมาะสมกับระดับความสามารถของผู้สอบจะลดค่าความคลาดเคลื่อน ฯ ลงได้และยังสอดคล้องกับงานวิจัยของ Haladyna และ Roid (1983) ที่พบว่า การจัดข้อสอบให้มีค่าความยากตรงกับระดับความสามารถของผู้สอบจะทำให้ค่าความคลาดเคลื่อน ฯ ลดลง ซึ่งงานวิจัยของ Tinsley และ Dawis (1975) และ สมพร บุญอ้อม (2529)

ก็ให้ผลสรุปเช่นเดียวกันคือ การใช้แบบสอบกับกลุ่มผู้สอบกลุ่มไหนก็ตามต้องพิจารณาระดับความยากให้เหมาะสมกับความสามารถของผู้สอบด้วย

ผลของการอภิปรายที่ได้กล่าวมาแล้วในข้อ 1 ทั้งหมดแสดงให้เห็นว่าตามแนวความคิดของทฤษฎีการตอบสนองข้อสอบ (Item Response Theory) นี้ ผู้วิจัยหรือผู้สร้างแบบสอบสามารถที่จะควบคุมให้แบบสอบที่สร้างขึ้นแต่ละฉบับมีค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถมากหรือน้อยเพียงใดก็ได้ โดยการพิจารณาค่าฟังก์ชันแสดงสารสนเทศของข้อสอบและแบบสอบที่ต้องการสร้าง อาทิ ถ้าต้องการให้มี ความคลาดเคลื่อนน้อยที่ระดับความสามารถ  $(\theta) = 2.00$  ก็นำข้อสอบแต่ละข้อที่ให้ค่าฟังก์ชันแสดงสารสนเทศของข้อสอบสูงที่ระดับความสามารถ  $(\theta) = 2.00$  หรือใกล้เคียงหลาย ๆ ข้อ หรือทั้งฉบับ นำมาคำนวณหาค่าฟังก์ชันแสดงสารสนเทศของแบบสอบก็จะได้ค่าความคลาดเคลื่อน น้อยด้วยเนื่องจากค่า S.E.E. คำนวณได้จากสูตร  $S.E.E. = \frac{1}{\sqrt{I(\theta)}}$  ซึ่งค่า S.E.E. เป็นอัตราส่วนผกผันของค่าฟังก์ชันแสดงสารสนเทศของแบบสอบ จึงทำให้ค่า S.E.E. น้อยที่ระดับ  $(\theta) = 2.00$  ตามต้องการ ผลที่ได้นี้แสดงให้เห็นอย่างชัดเจนถึงความแตกต่างและความได้เปรียบของการนำค่าพารามิเตอร์ของข้อสอบที่มีความคงที่โดยการประมาณจากทฤษฎีการตอบสนองข้อสอบมาใช้ว่ามีความถูกต้องแม่นยำกว่าการใช้ค่าพารามิเตอร์ของข้อสอบในทฤษฎีมาตรฐานดั้งเดิม (Classical Test Theory) ซึ่งมีการเปลี่ยนแปลงไปตามกลุ่มผู้สอบ

ต่อไปจะขออภิปรายผลการสอบของนักเรียนที่ได้สุ่มไว้กลุ่มความสามารถละ 200 คน ที่ได้สอบแบบสอบที่มีความเหมาะสมกับความสามารถต่าง ๆ กันทั้ง 3 ฉบับดังกล่าวแล้วมาพิจารณาหาความเที่ยงโดยใช้แนวคิดของทฤษฎีมาตรฐานดั้งเดิมเพื่อพิจารณาความสอดคล้องว่าเมื่อผู้สอบมีความสามารถตรงกับแบบสอบและทำให้ค่า S.E.E. น้อยที่สุดจากแนวคิดของทฤษฎีการตอบสนองข้อสอบแล้วจะให้ค่าความเที่ยงสูงสุดหรือไม่ซึ่งผลที่ได้ ผู้วิจัยขออภิปรายในข้อที่ 2 ดังต่อไปนี้

2. ผลการวิจัยเกี่ยวกับค่าความเที่ยงของแบบสอบ จากผลการวิจัยครั้งนี้ ผลที่ได้มีทั้งสอดคล้องและขัดแย้งกับสมมติฐานที่ 2 ที่ผู้วิจัยตั้งไว้ ซึ่งผู้วิจัยขอพิจารณาผลที่เกิดขึ้นในประเด็นที่น่าสนใจต่อไปนี้

2.1 พิจารณาค่าความเที่ยงของแบบสอบเมื่อกลุ่มความสามารถต่าง ๆ กัน สอบแบบสอบที่วัดความสามารถต่าง ๆ กันทั้ง 3 ฉบับ ซึ่งในการวิจัยครั้งนี้จะได้อภิปรายโดยใช้สูตรในการคำนวณค่าความเที่ยงของแบบสอบจากสูตรของ Kuder และ Richardson

สูตรที่ 20 ที่มีข้อตกลงเบื้องต้นคล้ายคลึงกับลักษณะของข้อมูลที่ใช้ในการวิจัยครั้งนี้ ซึ่งสูตรดังกล่าวผู้วิจัยได้เสนอไว้ในหน้าที่ 8 แล้ว จากสูตรการหาค่าความเที่ยงดังกล่าวจะสังเกตได้ว่าค่าความเที่ยงขึ้นอยู่กับค่าความแปรปรวนของแบบสอบทั้งฉบับและความแปรปรวนของแบบสอบดังกล่าวนี้ก็ขึ้นอยู่กับความแปรปรวนของข้อสอบรายข้อ คะแนนของผู้สอบยิ่งแตกต่างกันมากก็ยิ่งทำให้ความแปรปรวนรายข้อและความแปรปรวนของคะแนนที่สอบได้มากขึ้น จึงทำให้ค่าความเที่ยงมากขึ้น จากผลการทดสอบกับกลุ่มตัวอย่างที่ผู้วิจัยนำแบบสอบที่เหมาะสมกับความสามารถต่าง ๆ กันทั้ง 3 ฉบับไปทดสอบกับกลุ่มความสามารถต่าง ๆ กัน ได้ผลดังนี้

2.1.1 กลุ่มความสามารถต่ำเมื่อทดสอบแบบสอบที่วัดความสามารถต่ำได้ค่าความเที่ยงสูงสุดโดยพิจารณาจากสูตรหาค่าความเที่ยง KR-20 พบว่า ความแปรปรวนที่กลุ่มความสามารถต่ำทดสอบแบบสอบที่เหมาะสมกับความสามารถต่ำมีค่าสูงสุด การกระจายของคะแนนก็กระจายมากกว่า แสดงว่าแบบสอบที่เหมาะสมกับความสามารถต่ำสามารถวัดความสามารถที่แท้จริงของผู้สอบที่มีความสามารถต่ำได้ความคงที่ของคะแนนก็มีมากกว่าแบบสอบที่วัดความสามารถปานกลางและสูง ดังนั้นจึงอาจกล่าวได้ว่า โอกาสในการทำข้อสอบถูกของกลุ่มความสามารถต่ำทำข้อสอบที่ง่าย ๆ ที่เหมาะสมกับความสามารถของตนเองได้มากกว่าข้อสอบที่ยากกว่าความสามารถของเขาซึ่งอาจจะมีการเดาคำตอบถูก

2.1.2 กลุ่มความสามารถปานกลางเมื่อทดสอบแบบสอบที่มีความเหมาะสมกับความสามารถปานกลางได้ค่าความเที่ยงสูงสุด เมื่อพิจารณาจากผลการใช้สูตรหาค่าความเที่ยง KR-20 พบว่า ความแปรปรวนที่กลุ่มความสามารถปานกลางทำแบบสอบที่ตรงกับความสามารถปานกลางมีค่าสูงสุด การกระจายของคะแนนก็มากกว่า แสดงว่า แบบสอบที่เหมาะสมกับความสามารถปานกลางวัดความสามารถที่แท้จริงของผู้สอบที่มีความสามารถปานกลางได้ ความคงที่ของคะแนนก็มีมากกว่าเมื่อสอบแบบสอบที่วัดความสามารถต่ำและสูง ดังนั้นจึงอาจกล่าวได้ว่า โอกาสในการทำข้อสอบได้ถูกต้องของกลุ่มความสามารถปานกลางจะทำข้อสอบที่ตรงกับความสามารถของตนเองได้มากกว่าและทำข้อสอบที่ง่าย ๆ ได้มากกว่าข้อสอบที่ยากกว่าความสามารถของเขา

2.1.3 กลุ่มความสามารถสูงเมื่อทดสอบแบบสอบที่เหมาะสมกับความสามารถสูงได้ค่าความเที่ยงสูงสุด โดยพิจารณาผลการใช้สูตรหาค่าความเที่ยง KR-20 พบว่า ความแปรปรวนที่กลุ่มความสามารถสูงทำแบบสอบที่ตรงกับความสามารถสูงมีค่าสูงสุด การกระจายของคะแนนก็กระจายมากกว่า แสดงว่าแบบสอบที่

เหมาะสมกับความสามารถสูงสามารถวัดความสามารถที่แท้จริงของผู้ที่มีความสามารถสูงได้ ความคงที่ของคะแนนก็มีมากกว่าเมื่อสอบแบบสอบที่เหมาะสมกับความสามารถต่ำและปานกลาง ดังนั้นจึงอาจกล่าวได้ว่า โอกาสในการทำข้อสอบได้ถูกต้องของกลุ่มความสามารถสูงจะทำข้อสอบที่ตรงกับความสามารถของตนเองได้มากกว่าแบบสอบฉบับอื่น ๆ และยังมีโอกาสทำข้อสอบที่ง่ายกว่าความสามารถของตนเองได้ด้วย

ผลที่ได้กล่าวไว้ในข้อ 2.1.1 - 2.1.3 เมื่อพิจารณาให้ดีแล้วจะเห็นว่าในกลุ่มความสามารถแต่ละกลุ่มซึ่งมีลักษณะเป็นกลุ่มเอกพันธ์ (Homogeneous) แต่เมื่อพิจารณาภายในกลุ่มของตนเองแล้วมีลักษณะเป็นกลุ่มวิวิธพันธ์ (Heterogeneous) ตัวอย่างเช่น กลุ่มความสามารถสูง ภายในกลุ่มของตนเองก็ยังแบ่งเป็นความสามารถสูงที่มีระดับความสามารถแตกต่างกันและความสามารถในการตอบข้อสอบฉบับเดียวกันก็ยังแตกต่างกันอีกด้วย ดังนั้นเมื่อกลุ่มความสามารถสูงทดสอบแบบสอบที่เหมาะสมกับความสามารถต่าง ๆ กันทั้ง 3 ฉบับ ความคงที่ของคะแนนที่วัดได้จะมีมากขึ้น การกระจายของคะแนนภายในกลุ่มของตนเองมีมาก จึงเป็นสาเหตุที่ทำให้ค่าความเที่ยงของแบบสอบมีค่าสูงขึ้นด้วย ซึ่งผลดังกล่าวนี้ในกลุ่มความสามารถต่ำและปานกลางก็มีลักษณะเช่นเดียวกัน ผลการวิจัยครั้งนี้สอดคล้องกับงานวิจัยของ Hambleton Mills และ Simon (1983) ที่ได้ศึกษาข้อมูลที่สร้างจากเครื่องคอมพิวเตอร์แล้วพบว่า แบบสอบที่มีความยากใกล้เคียงกันและตรงกับความสามารถของผู้สอบจะมีค่าความเที่ยงสูง และค่าความเที่ยงจะเพิ่มขึ้น ณ ระดับที่ความยาวของข้อสอบเท่ากัน และยังสอดคล้องกับผลงานวิจัยในเรื่องการทดสอบแบบเทเลอร์ (Tailored Testing) ที่พบว่าในกรณีที่ผู้สอบได้ตอบข้อสอบที่มีระดับความยากสูงกว่าระดับความสามารถของตนเองแล้ว ผู้สอบมีแนวโน้มจะตอบข้อสอบด้วยพฤติกรรมอื่น ๆ นอกเหนือไปจากความรู้ความสามารถที่มีอยู่ เช่น การอาศัยพฤติกรรม การเดา และในกรณีที่ผู้สอบคนใดได้ตอบข้อสอบที่มีระดับความยากต่ำกว่าระดับความสามารถของตนเองแล้ว ข้อสอบนั้นจะไม่ช่วยให้ผู้สอบได้ใช้ความสามารถในระดับสูงสุด ซึ่งมีผลทำให้ค่าความเที่ยงของคะแนนผลการสอบจากผู้สอบทั้งสองกลุ่มต่ำกว่าที่ควรจะเป็น (Weiss 1974 : 1)

จากข้ออภิปรายข้อ 1 - 2 อาจกล่าวได้ว่า แบบสอบผลสัมฤทธิ์ที่ผู้วิจัยสร้างขึ้นสามารถใช้ได้กับรูปแบบของราศีเนื่องจากมีความเหมาะสมกับรูปแบบของราศีและจากการพิจารณาค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถและค่า

ความเที่ยงของแบบสอบไม่ว่าจะพิจารณาที่แบบสอบแต่ละฉบับที่วัดความสามารถต่าง ๆ กัน หรือพิจารณาที่กลุ่มผู้สอบที่มีความสามารถแตกต่างกันก็ตาม จะเห็นได้ว่าถ้าจัดข้อสอบให้เหมาะสมและตรงกับระดับความสามารถของผู้สอบจะให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถที่ต่ำกว่าและมีค่าความเที่ยงสูงสุด ซึ่งเป็นคุณสมบัติที่ดีของเครื่องมือวัด นอกจากนั้นค่านำไปใช้ก็จะได้แบบสอบที่มีประสิทธิภาพในการวัดได้ตรงกับความสามารถระดับต่าง ๆ ที่ผู้ดำเนินการสอบต้องการ อีกทั้งยังเป็นประโยชน์ในการจัดเก็บข้อสอบที่มีคุณภาพไว้ใช้ในธนาคารข้อสอบเพื่อใช้ในการสอบครั้งต่อไปด้วย

### ข้อเสนอแนะ

จากผลการวิจัยครั้งนี้ทำให้ได้แนวคิดบางประการซึ่งอาจมีประโยชน์ต่อไปโดยผู้วิจัยขอแยกเป็น 2 ประเด็น คือ

#### ข้อเสนอแนะในการนำไปใช้

1. ในการออกแบบแบบสอบด้วยการจัดข้อสอบให้มีระดับความยากง่ายที่เหมาะสมกับระดับความสามารถของผู้สอบโดยใช้ความสัมพันธ์ของฟังก์ชันแสดงสารสนเทศของข้อสอบและแบบสอบ (Item and Test information function)

สามารถใช้ได้กับการจำแนกคนในกลุ่มความสามารถที่ผู้สร้างแบบสอบต้องการ ทั้งนี้ข้อสอบที่นำมาใช้น่าจะวิเคราะห์ด้วยรูปแบบต่าง ๆ ในทฤษฎีการตอบสนองข้อสอบเสียก่อน เช่น รูปแบบของราสส์ เพื่อให้ได้ข้อสอบที่เหมาะสมกับรูปแบบและข้อสอบที่นำมาใช้ควรมีอยู่ในคลังข้อสอบที่มากพอ เพื่อให้การออกแบบจัดข้อสอบเหมาะสมกับความสามารถที่ต้องการได้อย่างสมบูรณ์

2. การใช้แบบสอบกับกลุ่มผู้สอบกลุ่มไหนต้องพิจารณาค่าความยากให้เหมาะสมกับความสามารถของผู้สอบด้วย เพราะในข้อสอบที่ง่ายหรือยากเกินไปอาจไม่มีผลต่อการใช้สอบวัด อีกทั้งการคัดเลือกข้อสอบต้องให้ตรงกับวัตถุประสงค์ของการสอบวัดในแต่ละครั้ง เช่น การสอบคัดเลือกบุคคล การวัดผลสัมฤทธิ์ เป็นต้น

3. การประมาณค่าพารามิเตอร์ของข้อสอบโดยใช้รูปแบบของราสส์ ซึ่งเป็นรูปแบบหนึ่งในทฤษฎีการตอบสนองข้อสอบที่มีข้อตกลงเบื้องต้นว่าค่าพารามิเตอร์ของข้อสอบจะไม่เปลี่ยนแปลงเมื่อนำไปทดสอบกับกลุ่มผู้สอบที่มีความสามารถต่างกัน ดังนั้นจึงสามารถที่จะนำข้อสอบที่ได้วิเคราะห์โดยรูปแบบของราสส์เก็บรวบรวมไว้ใช้ต่อไปได้ ซึ่งเป็นแนวทางในการสร้างธนาคารข้อสอบ (Item Bank) เพื่อเก็บข้อสอบที่มีคุณภาพไว้ใช้ต่อไป



### ข้อเสนอแนะในการวิจัย

1. ศึกษางานวิจัยในทำนองเดียวกันนี้ แต่ใช้ข้อสอบที่อยู่ในคลังข้อสอบที่มีข้อสอบจำนวนมาก ๆ เพื่อให้ได้ข้อสอบที่เหมาะสมกับกลุ่มความสามารถที่แตกต่างกันมากๆ และสามารถนำค่าความยากของข้อสอบเหล่านั้นมาออกแบบโค้งแสดงสารสนเทศของแบบสอบ (Test information curve) ที่มีประสิทธิภาพในการวัดคนกลุ่มต่าง ๆ ได้อย่างชัดเจน
2. ศึกษางานวิจัยในทำนองเดียวกันนี้ แต่ใช้แบบสอบผลสัมฤทธิ์วิชาอื่น ๆ เช่น ภาษาไทย ภาษาอังกฤษ วิทยาศาสตร์ เป็นต้น ว่าจะได้ผลสอดคล้องกันหรือไม่
3. ศึกษางานวิจัยในทำนองเดียวกันนี้ โดยใช้รูปแบบ 2 พารามิเตอร์หรือ 3 พารามิเตอร์ แทนรูปแบบของราสส์ เพื่อพิจารณาผลความแตกต่างของค่าฟังก์ชันแสดงสารสนเทศของข้อสอบและแบบสอบเมื่อมีพารามิเตอร์อำนาจจำแนก (a) และค่าการเคา (c) เข้ามาเกี่ยวข้องด้วยว่าจะได้ผลสอดคล้องหรือแตกต่างกันหรือไม่

ศูนย์วิทยพัชร์พยากร  
จุฬาลงกรณ์มหาวิทยาลัย