บทที่ 5

สรุปผล อภิปราย และข้อเสนอแนะ

การใช้คะแนนผลการสอบเพื่อจุดประสงค์บางอย่างเช่นการคัดเลือกนั้นเป็นสิ่งที่ผู้เกี่ยวข้องกับการสอบจะต้องพิถีพิถันและให้ความสำคัญเป็นพิเศษ โดยเฉพาะอย่างยิ่งในเรื่องคุณภาพของข้อสอบ ทั้งนี้เพราะคุณภาพของข้อสอบนับได้ว่าเป็นปัจจัยสำคัญยิ่งประการหนึ่งที่จะทำให้ผลการตัดสินมีความถูกต้อง ยุติธรรม การพิจารณาคุณภาพของข้อสอบโดยพิจารณาเฉพาะความเที่ยงและความตรงของแบบสอบจึงเป็นสิ่งที่ยังไม่เพียงพอ การวิเคราะห์หาความลำเอียงของข้อสอบเป็นสิ่งหนึ่งที่ควรจะต้องนำมาพิจารณาควบคู่ไปกับคุณภาพของข้อสอบด้านอื่น ๆ ด้วย ทั้งนี้เพราะจะช่วยลดปัญหาเกี่ยวกับความได้เปรียบเสียเปรียบระหว่างผู้สอบต่างกลุ่มกัน การวิจัยครั้งนี้เป็นการวิจัยเชิงประจักษ์เพื่อเปรียบเทียบผลของวิธีวิเคราะห์ความลำเอียงของข้อสอบที่แตกต่างกัน 4 วิธี

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อศึกษาความลำเอียงของข้อสอบจากวิธีวิเคราะห์ที่แตกต่างกัน 4 วิธี คือวิธีวิเคราะห์ความแปรปรวน วิธีแปลงค่าความยากของข้อสอบ วิธีโค้งลักษณะข้อสอบที่มีพารามิเตอร์ 1 ตัว และวิธีโค้งลักษณะข้อสอบที่มีพารามิเตอร์ 3 ตัว โดยหาสัมประสิทธิ์สหสัมพันธ์ของดัชนีความลำเอียงระหว่างวิธีวิเคราะห์ความลำเอียง 4 วิธี และเปรียบเทียบความแตกต่างของผลการคัดเลือกก่อนและหลังการศึกษาความลำเอียงของข้อสอบตามวิธีการคิดคะแนนรวมที่แตกต่างกัน 6 วิธี ในด้านจำนวนผู้ได้รับการคัดเลือก สัดส่วนของชาย : หญิง ที่ได้รับการคัดเลือก ความเที่ยงของแบบสอบ

ผลวิจัยที่ใช้ในการวิจัยครั้งนี้คือผู้เข้าสอบแข่งขันเพื่อบรรจุเข้ารับราชการเป็นข้าราชการครูในปี พ.ศ. 2529 จำนวน 2,340 คน เป็นชาย 1,170 คน เป็นหญิง 1,170 คน เครื่องมือที่ใช้ในการรวบรวมข้อมูลได้แก่แบบสอบที่ใช้สอบแข่งขัน เพื่อบรรจุเข้ารับราชการเป็นข้าราชการครูในปี พ.ศ. 2529 ซึ่งมี 4 ฉบับ

ในการดำเนินการวิจัยนั้นได้ดำเนินการดังนี้คือ วิเคราะห์หาค่าดัชนีความลำเอียงของข้อสอบแต่ละข้อด้วยวิธีวิเคราะห์ความแปรปรวน วิธีแปลงค่าความยากของข้อสอบ วิธีโค้งลักษณะข้อสอบที่มีพารามิเตอร์ 1 ตัว และวิธีโค้งลักษณะข้อสอบที่มีพารามิเตอร์ 3 ตัว แล้วนำค่าดัชนีความลำเอียงนี้ไปคำนวณค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างวิธีวิเคราะห์ทั้ง 4 วิธี จากนั้นก็ตัดข้อสอบข้อที่มีความลำเอียงในแต่ละวิธีวิเคราะห์ออกไป แล้วรวมคะแนนเพื่อจัดอันดับผู้สอบจากวิธีรวมคะแนนที่แตกต่างกัน 6 วิธี โดยกำหนดอัตราส่วนการคัดเลือกเป็น 10% 15% และ 20% ของผู้สอบทั้งหมด ในการนี้ผู้วิจัยได้เปรียบเทียบความแตกต่างของผลการคัดเลือกก่อนและหลังการศึกษาความลำเอียงของข้อสอบในด้านจำนวนผู้ได้รับการคัดเลือก สัดส่วนของชาย : หญิง ที่ได้รับการคัดเลือก และความเที่ยงของแบบสอบ

ในการวิเคราะห์ข้อสอบตามทฤษฎีการตอบสนองต่อข้อสอบโดยใช้คอมพิวเตอร์โปรแกรม LOGIST นั้นปรากฏว่าจะต้องตัดข้อสอบออกไปจำนวนหนึ่งเนื่องจากไม่สามารถประมาณค่าพารามิเตอร์ได้ โดยแบบสอบแต่ละฉบับจะตัดข้อสอบออกดังนี้คือ ฉบับที่ 1 ตัดออก 38 ข้อ ฉบับที่ 2 ตัดออก 32 ข้อ ฉบับที่ 3 ตัดออก 38 ข้อ และฉบับที่ 4 ตัดออก 29 ข้อ (ข้อสอบเดิมมีฉบับละ 100 ข้อ) ดังนั้นจึงต้องตัดข้อสอบข้อดังกล่าวออกจากแต่ละวิธีวิเคราะห์ ทั้งนี้เพื่อให้มีจำนวนข้อสอบเท่ากันในทุกวิธีวิเคราะห์ โดยแบบสอบฉบับที่ 1,2,3, และ 4 มีจำนวนข้อสอบเป็น 62,68,62 และ 71 ข้อ ตามลำดับ

<u>สรุปผลการวิจัย</u>

การศึกษาเปรียบเทียบผลของวิธีวิเคราะห์ความลำเอียงของข้อสอบที่แตกต่างกัน 4 วิธี ปรากฏดังนี้

1. จากการทดสอบข้อตกลงเบื้องต้นเกี่ยวกับความเป็นมิติเดียว ปรากฏว่าแบบสอบทุกฉบับไม่ว่าจะแยกวิเคราะห์ตามกลุ่มรวมชายหญิง กลุ่มผู้สอบชาย หรือกลุ่มผู้สอบหญิงมีข้อมูลสนับสนุนว่ามีความเป็นมิติเดียว ดังนั้นจึงมีความสมเหตุสมผลที่จะใช้ทฤษฎีการตอบสนองต่อข้อสอบ มาอธิบายพฤติกรรมการตอบข้อสอบจากแบบสอบทั้ง 4 ฉบับได้

2. ผลการวิเคราะห์ความลำเอียงของข้อสอบโดยใช้วิธีวิเคราะห์ความแปรปรวนพบว่าแบบสอบฉบับที่ 3 มีจำนวนข้อสอบที่ลำเอียงคิดเป็นเปอร์เซ็นต์สูงสุดคือ 38.71% หรือมีข้อสอบที่ลำเอียงจำนวน 24 ข้อ จากข้อสอบทั้งหมด 62 ข้อ โดยเป็นข้อสอบที่ลำเอียงต่อต้านกลุ่มชาย 17 ข้อ ข้อสอบที่ลำเอียงต่อต้านกลุ่มหญิง 7 ข้อ รองลงมาได้แก่ แบบสอบฉบับที่ 4 ซึ่งมีข้อสอบที่ลำเอียง 38.03% หรือมีข้อสอบที่ลำเอียงทั้งสิ้น 27 ข้อ จากข้อสอบทั้งหมด 71 ข้อ เป็นข้อสอบที่ลำเอียงต่อต้านกลุ่มชายจำนวน 5 ข้อ เป็นข้อสอบที่ลำเอียงต่อต้านกลุ่มหญิง 22 ข้อ

ส่วนแบบสอบฉบับที่ 2 มีข้อสอบที่ลำเอียง 36.76% หรือมีข้อสอบที่ลำเอียง 25 ข้อจากข้อสอบทั้งหมด 68 ข้อ เป็นข้อสอบที่ลำเอียงต่อต้านกลุ่มชาย 5 ข้อ เป็นข้อสอบที่ลำเอียงต่อต้านกลุ่มหญิง 20 ข้อ และแบบสอบฉบับที่ 1 เป็นแบบสอบที่มีจำนวนข้อสอบที่ลำเอียงคิดเป็นเปอร์เซ็นต์ต่ำสุดคือ 35.48% หรือมีข้อสอบที่ลำเอียง 22 ข้อ จากข้อสอบทั้งหมด 62 ข้อ เป็นข้อสอบที่ลำเอียงต่อต้านกลุ่มชาย 11 ข้อ เป็นข้อสอบที่ลำเอียงต่อต้านกลุ่มหญิง 11 ข้อ

เมื่อพิจารณารวมทั้ง 4 ฉบับ พบว่ามีข้อสอบที่ลำเอียงทั้งสิ้นจำนวน 98 ข้อ จากข้อสอบทั้งหมด 263 ข้อ หรือประมาณ 37.26% โดยเป็นข้อสอบที่ลำเอียงต่อต้านกลุ่มชายจำนวน 38 ข้อ และเป็นข้อสอบที่ลำเอียงต่อต้านกลุ่มหญิงจำนวน 60 ข้อ

3. ผลการวิเคราะห์ความลำเอียงของข้อสอบโดยใช้วิธีแปลงค่าความยากของข้อสอบ พบว่าแบบสอบฉบับที่ 2 มีจำนานข้อสอบที่ลำเอียงคิดเป็นเปอร์เซ็นต์สูงสุดคือ17.65% หรือมีข้อสอบที่ลำเอียงจำนวน 12 ข้อ จากข้อสอบทั้งหมด 68 ข้อ โดยเป็นข้อสอบที่ลำเอียงต่อต้านกลุ่มชาย 7 ข้อ เป็นข้อสอบที่ลำเอียงต่อต้านกลุ่มหญิง 5 ข้อ รองลงมาได้แก่ แบบสอบฉบับที่ 4 ซึ่งมีข้อสอบที่ลำเอียง 15.49% หรือมีข้อสอบที่ลำเอียง 11 ข้อ จากข้อสอบทั้งหมด 71 ข้อ เป็นข้อสอบที่ลำเอียงต่อต้านกลุ่มชายจำนวน 4 ข้อ เป็นข้อสอบที่ลำเอียงต่อต้านกลุ่มหญิงจำนวน 7 ข้อ

ส่วนแบบสอบฉบับที่ 3 มีข้อสอบที่ลำเอียง 11.29% หรือมีข้อสอบที่ลำเอียง 7 ข้อจากข้อสอบทั้งหมด 62 ข้อ เป็นข้อสอบที่ลำเอียงต่อต้านกลุ่มชายจำนวน 2 ข้อ เป็นข้อสอบที่ลำเอียงต่อต้านกลุ่มหญิงจำนวน 5 ข้อ และแบบสอบฉบับที่ 1 เป็นแบบสอบที่มีความลำเอียงคิดเป็นเปอร์เซ็นต์ต่ำสุดคือ 9.68% หรือมีข้อสอบที่ลำเอียง 6 ข้อจากข้อสอบทั้งหมด 62 ข้อ เป็นข้อสอบที่ลำเอียงต่อต้านกลุ่มชาย 3 ข้อ เป็นข้อสอบที่ลำเอียงต่อต้านกลุ่มหญิง 3 ข้อ

เมื่อพิจารณารวมทั้ง 4 ฉบับ พบว่ามีข้อสอบที่ลำเอียงทั้งสิ้นจำนวน 36 ข้อ จากข้อสอบทั้งหมด 263 ข้อหรือประมาณ 13.68% โดยเป็นข้อสอบที่ลำเอียงต่อต้านกลุ่มชายจำนวน 16 ข้อ และเป็นข้อสอบที่ลำเอียงต่อต้านกลุ่มหญิงจำนวน 20 ข้อ

4. ผลการวิเคราะห์ความลำเอียงของข้อสอบ โดยใช้วิธีโค้งลักษณะข้อสอบที่มีพารามิเตอร์ 1 ตัว พบว่าแบบสอบฉบับที่ 1 มีจำนวนข้อสอบที่ลำเอียงคิดเป็นเปอร์เซ็นต์สูงสุดคือ 30.64% หรือมีข้อสอบที่ลำเอียงจำนวน 19 ข้อจากข้อสอบทั้งหมด 62 ข้อ โดยเป็นข้อสอบที่ลำเอียงต่อต้านกลุ่มชายจำนวน 8 ข้อ เป็นข้อสอบที่ลำเอียงต่อต้านต่อกลุ่มหญิงจำนวน 11 ข้อ รองลงมาได้แก่แบบสอบฉบับที่ 4 ซึ่งมีข้อสอบที่ลำเอียง 25.35% หรือมีข้อสอบที่ลำเอียง 18 ข้อ จากข้อสอบทั้งหมด 71 ข้อ เป็นมีข้อสอบที่ลำเอียงต่อต้านกลุ่มชายจำนวน 7 ข้อ เป็นข้อสอบที่ลำเอียงต่อต้านกลุ่มหญิง จำนวน 11 ข้อ

ส่วนแบบสอบฉบับที่ 3 มีข้อสอบที่ลำเอียง 24.19% หรือมีข้อสอบที่ลำเอียง 15 ข้อจากข้อสอบทั้งหมด 62 ข้อ เป็นข้อสอบที่ลำเอียงต่อต้านกลุ่มชายจำนวน 8 ข้อ เป็นข้อสอบที่ลำเอียงต่อต้านกลุ่มหญิงจำนวน 7 ข้อ และแบบสอบฉบับที่ 2 เป็นแบบสอบที่มีความลำเอียงคิดเป็นเปอร์เซ็นต์ต่ำสุดคือ 17.64% หรือมีข้อสอบที่ลำเอียง 12 ข้อจากข้อสอบทั้งหมด 68 ข้อ เป็นข้อสอบที่ลำเอียงต่อต้านกลุ่มชาย 6 ข้อ เป็นข้อสอบที่ลำเอียงต่อต้านกลุ่มหญิง 6 ข้อ

เมื่อพิจารณารวมทั้ง 4 ฉบับ พบว่ามีข้อสอบที่ลำเอียงทั้งสิ้นจำนวน 64 ข้อ จากข้อสอบทั้งหมด 263 ข้อ หรือประมาณ 24.33% โดยเป็นข้อสอบที่ลำเอียงต่อต้านกลุ่มชายจำนวน 29 ข้อ และเป็นข้อสอบที่ลำเอียงต่อต้านกลุ่มหญิงจำนวน 35 ข้อ

5. ผลการวิเคราะห์ความลำเอียงของข้อสอบ โดยใช้วิธีโค้งลักษณะข้อสอบที่มีพารามิเตอร์ 3 ตัว พบว่าแบบสอบฉบับที่ 4 มีจำนวนข้อสอบที่ลำเอียงคิดเป็นเปอร์เซ็นต์สูงสุดคือ 45.07% หรือมีข้อสอบที่ลำเอียงจำนวน 32 ข้อ จากข้อสอบทั้งหมด 71 ข้อ โดยเป็นข้อสอบที่ลำเอียงต่อต้านกลุ่มชายจำนวน 7 ข้อ เป็นข้อสอบที่ลำเอียงต่อต้านกลุ่มหญิง จำนวน 25 ข้อ รองลงมาได้แก่แบบสอบฉบับที่ 1 ซึ่งมีข้อสอบที่ลำเอียง 37.09% หรือมีข้อสอบที่ลำเอียง 23 ข้อจากข้อสอบทั้งหมด 62 ข้อ เป็นข้อสอบที่ลำเอียงต่อต้านกลุ่มชายจำนวน 6 ข้อ เป็นข้อสอบที่ลำเอียงต่อต้านกลุ่มหญิงจำนวน 17 ข้อ

ส่วนแบบสอบฉบับที่ 3 มีข้อสอบที่ลำเอียง 33.87% หรือมีข้อสอบที่ลำเอียง 21 ข้อจากข้อสอบทั้งหมด 62 ข้อ เป็นข้อสอบที่ลำเอียงต่อต้านกลุ่มชาย จำนวน 14 ข้อ เป็นข้อสอบที่ลำเอียงต่อต้านกลุ่มหญิง 7 ข้อ และแบบสอบฉบับที่ 2 เป็นแบบสอบที่มีความลำเอียงคิดเป็นเปอร์เซ็นต์ต่ำสุดคือ 33.82% หรือมีข้อสอบที่ลำเอียง 23 ข้อ จากข้อสอบทั้งหมด 68 ข้อ เป็นข้อสอบที่ลำเอียงต่อต้านกลุ่มชายจำนวน 5 ข้อ เป็นข้อสอบที่ลำเอียงต่อต้านกลุ่มหญิงจำนวน 18 ข้อ

เมื่อพิจารณารวมทั้ง 4 ฉบับ พบว่ามีข้อสอบที่ลำเอียงทั้งสิ้นจำนวน 99 ข้อ จากข้อสอบทั้งหมด 263 ข้อ หรือประมาณ 37.64% โดยเป็นข้อสอบที่ลำเอียงต่อต้านกลุ่มชายจำนวน 32 ข้อ และเป็นข้อสอบที่ลำเอียงต่อต้านกลุ่มหญิงจำนวน 67 ข้อ

6. จากผลการวิเคราะห์ความลำเอียงของข้อสอบ 4 วิธี ทำให้พบว่ามีข้อสอบอยู่จำนวนหนึ่งที่มีความลำเอียงในทุกวิธีวิเคราะห์ รายละเอียดของข้อสอบที่ลำเอียงในทุกวิธีวิเคราะห์จากแบบสอบทั้ง 4 ฉบับ ปรากฏในตาราง 50

ตาราง 50 แสดงจำนวนข้อสอบที่ลำเอียงในทุกวิธีวิเคราะห์จากแบบสอบทั้ง 4 ฉบับ

แบบสอบ	ข้อสอบข้อที่ลำเอียง	
	ต่อต้านกลุ่มชาย	ต่อต้านกลุ่มหญิง
ฉบับที่ 1	27,30,89	13,88,90
ฉบับที่ 2	33,66,78,79	2,14,28,31,88
ฉบับที่ 3	25,76	27,28,62,70
ฉบับที่ 4	29,53,59,82	4,8,77,91,94

7. ในการหาสัมประสิทธิ์สหสัมพันธ์ของดัชนีความลำเอียงระหว่างวิธีวิเคราะห์ความลำเอียง 4 วิธี พบว่าวิธีวิเคราะห์ทั้ง 4 วิธีมีความสัมพันธ์ในทางบวกซึ่งกันและกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .001 ค่าสัมประสิทธิ์สหสัมพันธ์ของดัชนีความลำเอียงระหว่างวิธีวิเคราะห์ทั้ง 4 วิธีมีค่าสูงมากโดยมีค่า r อยู่ระหว่าง .7535 ถึง .9921 และค่าสัมประสิทธิ์สหสัมพันธ์ที่หาได้นี้ให้ผลค่อนข้างจะสอดคล้องกันมาก กล่าวคือในแบบสอบฉบับที่ 2, 3 และ 4 ต่างก็พบว่าวิธีวิเคราะห์ความแปรปรวนกับวิธีโค้งลักษณะข้อสอบที่มีพารามิเตอร์ 1 ตัวเป็นวิธีที่มีความสัมพันธ์กันสูงสุด จะมีแตกต่างไปก็เฉพาะแบบสอบฉบับที่ 1 เท่านั้น ที่พบว่าวิธีแปลงค่าความยากของข้อสอบกับวิธีโค้งลักษณะข้อสอบที่มีพารามิเตอร์ 1 ตัว เป็นวิธีที่มีความสัมพันธ์กันสูงที่สุด

8. ในการเปรียบเทียบความแตกต่างของผลการคัดเลือกก่อนและหลังการศึกษาความลำเอียงของข้อสอบตามวิธีการคิดคะแนนที่แตกต่างกัน 6 วิธี ปรากฏผลดังนี้

8.1 ผลการเปรียบเทียบความแตกต่างของจำนวนผู้ได้รับการคัดเลือกระหว่างวิธีคะแนนดิบกับวิธีรวมคะแนนแบบอื่น ๆ อีก 5 วิธี พบว่าถ้ากำหนดอัตราส่วนในการคัดเลือกเป็น 10% ของผู้สอบทั้งหมด วิธีตัดข้อที่ลำเอียงจากวิธีโค้งลักษณะข้อสอบที่มีพารามิเตอร์ 3 ตัว เป็นวิธีที่มีจำนวนผู้ได้รับการคัดเลือกแตกต่างจากวิธีคะแนนดิบมากที่สุด กล่าวคือมีจำนวนผู้ได้รับการคัดเลือกแตกต่างกัน 23.93% ส่วนวิธีแปลงคะแนนดิบเป็นคะแนนมาตรฐานทีปกติจะมีจำนวนผู้ได้รับการคัดเลือกแตกต่างจากวิธีคะแนนดิบน้อยที่สุดคือแตกต่างกัน 4.70% และหากอัตราส่วนในการคัดเลือกเป็น 15% หรือ 20% ของผู้สอบทั้งหมดจะพบว่าวิธีตัดข้อสอบที่ลำเอียงจากวิธีวิเคราะห์ความแปรปรวนกับวิธีคะแนนดิบเป็นวิธีที่มีจำนวนผู้ได้รับการคัดเลือกแตกต่างกันมากที่สุด ส่วนวิธีแปลงคะแนนดิบเป็นคะแนนมาตรฐานทีปกติกับวิธีคะแนนดิบเป็นวิธีที่มีจำนวนผู้ได้รับการคัดเลือกแตกต่างกันน้อยที่สุด โดยเปอร์เซ็นต์ความแตกต่างของจำนวนคนที่ได้รับการคัดเลือกระหว่างวิธีคะแนนดิบกับวิธีรวมคะแนนแบบอื่น ๆ อีก 5 วิธี มีค่าอยู่ระหว่าง 3.85% ถึง 23.93%

8.2 ผลการเปรียบเทียบความแตกต่างของจำนวนผู้ได้รับการคัดเลือกระหว่างวิธีแปลงคะแนนดิบเป็นคะแนนมาตรฐานทีปกติกับวิธีรามคะแนนแบบอื่น ๆ อีก 4 วิธี พบว่าถ้าอัตราส่วนในการคัดเลือกเป็น 10% ของผู้สอบทั้งหมด วิธีตัดข้อที่ลำเอียงจากวิธีโค้งลักษณะข้อสอบที่มีพารามิเตอร์ 3 ตัวเป็นวิธีที่มีจำนวนผู้ได้รับการคัดเลือกแตกต่างไปจากวิธีแปลงคะแนนดิบเป็นคะแนนมาตรฐานทีปกติถึง 22.65% ในขณะที่วิธีตัดข้อที่ลำเอียงจากวิธีแปลงค่าความยากของข้อสอบมีจำนวนผู้ได้รับการคัดเลือกต่างไปจากวิธีแปลงคะแนนดิบเป็นคะแนนมาตรฐานทีปกติเพียง 9.40% และถ้าอัตราส่วนในการคัดเลือกเป็น 15% หรือ 20% ของผู้สอบทั้งหมด วิธีตัดข้อที่ลำเอียงจากวิธีวิเคราะห์ความแปรปรวน จะมีจำนวนผู้ได้รับการคัดเลือกแตกต่างไปจากวิธีแปลงคะแนนดิบเป็นคะแนนมาตรฐานทีปกติมากที่สุด (17.66% สำหรับอัตราส่วนการคัดเลือก 15% และ 16.67% สำหรับอัตราส่วนการคัดเลือก 20%) ในขณะที่วิธีตัดข้อที่ลำเอียงจากวิธีแปลงค่าความยากของข้อสอบมีจำนวนผู้ได้รับการคัดเลือกแตกต่างไปจากวิธีแปลงคะแนนดิบเป็นคะแนนมาตรฐานทีปกติน้อยที่สุด (5.98% สำหรับอัตราส่วนการคัดเลือก 15% และ 7.05% สำหรับอัตราส่วนการคัดเลือก 20%) โดยเปอร์เซ็นต์ความแตกต่างของจำนวนผู้ได้รับการคัดเลือกระหว่างวิธีแปลงคะแนนดิบเป็นคะแนนมาตรฐานทีปกติกับวิธีรวมคะแนนแบบอื่น ๆ อีก 4 วิธีมีค่าอยู่ระหว่าง 5.98% ถึง 22.65%

8.3 สัดส่วนของชาย : หญิง ที่ได้รับการคัดเลือกก่อนและหลังการศึกษาความลำเอียงของข้อสอบ พบว่าก่อนการศึกษาความลำเอียงของข้อสอบนั้นสัดส่วนของชายที่ได้รับการคัดเลือกจะสูงกว่าหญิง กล่าวคือผู้สอบชายมีโอกาสได้รับการคัดเลือกประมาณ 54% ถึง 57% ในขณะที่ผู้สอบหญิงมีโอกาสได้รับคัดเลือกประมาณ 43% ถึง 46% แต่หลังจากการศึกษาความลำเอียงและตัดข้อที่ลำเอียงออกไปจะพบว่าสัดส่วนของผู้สอบหญิงที่ได้รับการคัดเลือกใกล้เคียงกับสัดส่วนของผู้สอบชาย

8.4 ค่าความเที่ยงของแบบสอบ พบว่าหลังจากที่ตัดข้อสอบที่ลำเอียงออกไป ค่าความเที่ยงของแบบสอบจะลดลง แต่เมื่อทำการปรับขยายให้แบบสอบแต่ละฉบับจากวิธีวิเคราะห์ทั้ง 4 มีจำนวนข้อสอบเท่ากับแบบสอบเดิมแล้ว ปรากฏว่าค่าความเที่ยงของแบบสอบส่วนใหญ่ จะมีค่าสูงกว่าแบบสอบเดิม ยกเว้นแบบสอบฉบับที่ 4 ที่พบว่าค่าความเที่ยงต่ำกว่าแบบสอบเดิม อย่างไรก็ตามค่าความเที่ยงของแบบสอบฉบับที่ 4 นี้ ต่ำกว่าค่าความเที่ยงของแบบสอบเดิมเพียงเล็กน้อย

เท่านั้น นอกจากนี้ยังพบว่าวิธีโค้งลักษณะข้อสอบที่มีพารามิเตอร์ 3 ตัว เป็นวิธีที่ให้ค่าความเที่ยงสูงสุดในแบบสอบฉบับที่ 2 และ 3 ส่วนวิธีวิเคราะห์ความแปรปรวน และวิธีโค้งลักษณะข้อสอบที่มีพารามิเตอร์ 1 ตัว เป็นวิธีที่ให้ค่าความเที่ยงสูงสุดในแบบสอบฉบับที่ 1 และ 4 ตามลำดับ

อภิปรายผล

1. จากการวิเคราะห์ความลำเอียงของข้อสอบด้วยเทคนิควิธีวิเคราะห์ที่แตกต่างกัน 4 วิธี พบว่าแต่ละวิธีวิเคราะห์จะพบจำนวนข้อที่ลำเอียงแตกต่างกัน โดยวิธีโค้งลักษณะข้อสอบที่มีพารามิเตอร์ 3 ตัว พบจำนวนข้อที่ลำเอียงมากที่สุด รองลงมาได้แก่ วิธีวิเคราะห์ความแปรปรวน ซึ่งพบว่าในแบบสอบ 4 ฉบับ (263 ข้อ) มีข้อสอบที่ลำเอียงจำนวน 99 และ 98 ข้อ ตามลำดับ ส่วนวิธีที่พบจำนวนข้อที่ลำเอียงน้อยที่สุดคือ วิธีแปลงค่าความยากของข้อสอบที่พบว่ามีข้อที่ลำเอียงจำนวน 36 ข้อ อย่างไรก็ตามดัชนีความลำเอียงของข้อสอบที่วิเคราะห์ด้วยวิธีโค้งลักษณะข้อสอบที่มีพารามิเตอร์ 3 ตัว ยังคงอยู่ในระดับต่ำเป็นจำนวนมาก ($\emptyset < .40$) (Rudner 1977) สำหรับข้อสอบข้อที่ควรให้ความสนใจมากเป็นพิเศษคือข้อที่พบว่ามีความลำเอียงในทุกวิธีวิเคราะห์ ซึ่งปรากฏผลดังนี้ แบบสอบฉบับที่ 1 ข้อที่ลำเอียงได้แก่ข้อ 13,27,30,88, 89, และ 90 แบบสอบฉบับที่ 2 ข้อที่ลำเอียงได้แก่ข้อ 2,14,28,31,33,66,78,79 และ 88 แบบสอบฉบับที่ 3 ข้อที่ลำเอียงได้แก่ข้อ 25,27,28,62,70 และ 76 แบบสอบฉบับที่ 4 ข้อที่ลำเอียงได้แก่ข้อ 4,8,29,53,59,77,82,91 และ 94 ในการศึกษาวิธีวิเคราะห์ความลำเอียงครั้งนี้ไม่พบข้อสอบที่ลำเอียงสลับกลุ่มเมื่อใช้วิธีวิเคราะห์แตกต่างกัน กล่าวคือจะไม่มีข้อสอบใดที่ลำเอียงต่อต้านกลุ่มหนึ่งในวิธีวิเคราะห์หนึ่ง และลำเอียงต่อต้านอีกกลุ่มหนึ่งในอีกวิธีวิเคราะห์หนึ่ง จากผลดังนี้ทำให้เห็นว่าวิธีการวิเคราะห์ความลำเอียงที่นำมาใช้ในการศึกษาครั้งนี้ให้ผลค่อนข้างจะสอดคล้องกัน

จากผลการศึกษาที่พบยังไม่อาจกล่าวได้ว่าวิธีวิเคราะห์ใดเป็นวิธีที่ดีที่สุดถึงแม้วิธีโค้งลักษณะข้อสอบที่มีพารามิเตอร์ 3 ตัว เป็นวิธีที่ค้นพบจำนวนข้อที่ลำเอียงได้มากที่สุด ทั้งนี้เพราะเกณฑ์ที่ใช้ในการตัดสินความลำเอียงของข้อสอบยังเป็นเกณฑ์ที่กำหนดขึ้นโดยผู้วิจัย ไม่มีเกณฑ์ที่แน่นอน

ชัดเจน ดังนั้นถ้าเกณฑ์เปลี่ยน จำนวนข้อที่พบว่ามีความลำเอียงก็จะเปลี่ยนตามไปด้วย ข้อสอบที่พบว่ามีความลำเอียงนั้นไม่ได้หมายความว่าเป็นข้อสอบที่ไม่ดี ขาดคุณภาพ แต่เป็นเพียงการให้ข้อมูลเกี่ยวกับข้อสอบในอีกด้านหนึ่ง การจะตัดสินว่าข้อสอบใดขาดคุณภาพหรือวัดได้ไม่ตรงนั้น จำเป็นต้องใช้ข้อมูลอื่น ๆ ประกอบอีกหลายด้าน เทคนิควิธีวิเคราะห์ความลำเอียงของข้อสอบไม่ว่าจะเป็นวิธีใดก็ตามไม่สามารถจะช่วยให้แบบสอบมีคุณภาพดีขึ้น การที่แบบสอบจะดีหรือไม่นั้นขึ้นอยู่กับความสามารถของผู้เขียนแบบสอบเป็นสำคัญ เทคนิควิธีวิเคราะห์ เป็นเพียงดัชนีตัวหนึ่งที่สนับสนุนว่าแบบสอบที่เขียนขึ้นนั้นมีข้อบกพร่องในแง่ใดหรือไม่ แม้ว่าวิธีโค้งลักษณะข้อสอบที่มีพารามิเตอร์ 3 ตัว ดูว่าจะเป็นวิธีที่สอดคล้อง ถูกต้องตามทฤษฎีการวัดผลมากที่สุด แต่วิธีนี้ก็เป็นวิธีที่มีความจำกัด คือ ต้องใช้กลุ่มตัวอย่างขนาดใหญ่ (กลุ่มตัวอย่างไม่ควรน้อยกว่า 1000 คน) และข้อสอบจำนวนมาก (ข้อสอบไม่ควรน้อยกว่า 40 ข้อ) (Wood, Wingersky and Lord, 1976) จึงจะทำให้การประมาณค่าพารามิเตอร์ต่าง ๆ มีความถูกต้องและคงที่ นอกจากนี้ Heebara (1979) และ Lord (1975) ยังได้ให้ความเห็นว่าวิธีโค้งลักษณะข้อสอบที่มีพารามิเตอร์ 3 ตัว เป็นวิธีที่ค่อนข้างซับซ้อน ค่าใช้จ่ายในการคำนวณสูง และความถูกต้องในการประมาณค่าการเดา (c) ค่อนข้างต่ำ ด้วยเหตุดังที่ได้กล่าวมานักวิจัยส่วนใหญ่จึงสนับสนุนให้ใช้วิธีวิเคราะห์แบบอื่นที่ให้ผลพอใช้ได้ถึงแม้จะไม่ลึกซึ้งเท่าวิธีโค้งลักษณะข้อสอบที่มีพารามิเตอร์ 3 ตัวก็ตาม วิธีการที่ได้รับการเสนอแนะบ่อย ๆ ก็คือ วิธีวิเคราะห์ความแปรปรวน

ไม่ว่าจะเป็นวิธีวิเคราะห์แบบใดก็ไม่อาจบอกแหล่ง หรือสาเหตุของความลำเอียงในตัวคำถามได้ วิธีการหนึ่งที่จะตอบคำถามนี้ได้ ก็คือการวิเคราะห์ภายหลัง เพื่อหาแหล่งของความลำเอียง (Sources of Bias) (Scheuneman 1982 : 180 - 195) โดยการนำเอาข้อสอบที่พบว่ามีความลำเอียงมาทบทวน (Review) โดยผู้สร้างแบบสอบอีกครั้งหนึ่ง ขั้นตอนในการทบทวนสิ่งที่ทำให้เกิดความลำเอียงมีอยู่ในแบบฟอร์มการทบทวนในภาคผนวก A ของหนังสือ Handbook of Method for Detecting Test Bias (Berk 1982 : 306 - 308)

2. วิธีวิเคราะห์ความลำเอียงทั้ง 4 วิธี มีความสัมพันธ์กันสูงมาก กล่าวคือมีค่า r อยู่ระหว่าง .7535 ถึง .9921 ซึ่งสอดคล้องกับผลการวิจัยของ Ironson and Craig (1982) ที่พบว่าสหสัมพันธ์ระหว่างวิธีวิเคราะห์ความลำเอียงทั้ง 5 วิธีมีค่าสูง และเป็นไปตามที่ผู้วิจัยคาดหมายไว้กล่าวคือ วิธีการทางสถิติสำหรับวิเคราะห์ความลำเอียงของข้อสอบไม่ว่าจะเป็นวิธีการใดก็ตามจะต้องมีความสัมพันธ์กันสูง จากข้อค้นพบที่ว่าวิธีการวิเคราะห์ทั้ง 4 วิธี มีความสัมพันธ์กันสูงมาก จึงทำให้ สามารถสรุปได้ว่าในการวิเคราะห์ความลำเอียงของข้อสอบนั้นจะเลือกใช้วิธีการใดก็ได้ เพราะจะให้ผลไม่แตกต่างกันมากนัก ซึ่งในที่นี้ผู้วิจัยเห็นว่าน่าจะเลือกใช้วิธีที่ง่าย สะดวก และเสียค่าใช้จ่ายน้อย เช่น วิธีวิเคราะห์ความแปรปรวน วิธีแปลงค่าความยากของข้อสอบ หรือวิธีโค้งลักษณะข้อสอบที่มีพารามิเตอร์ 1 ตัว วิธีการที่เสนอทั้ง 3 วิธีนี้ต่างก็มีความสัมพันธ์กับวิธีโค้งลักษณะข้อสอบที่มีพารามิเตอร์ 3 ตัว ค่อนข้างสูง ดังรายละเอียดในตอนที่ 5 ของบทที่ 4

3. จากการเปรียบเทียบความแตกต่างของผลการคัดเลือกก่อนและหลังการศึกษาความลำเอียงของข้อสอบตามวิธีการคิดคะแนนรวมที่แตกต่างกัน 6 วิธีนั้น ปรากฏผลดังนี้คือถ้าตัดข้อสอบที่ลำเอียงจากแต่ละวิธีวิเคราะห์ออกไปจะทำให้จำนวนผู้ได้รับการคัดเลือกแตกต่างไปจากวิธีคะแนนดิบระหว่าง 7% ถึง 24% โดยเปอร์เซ็นต์ของความแตกต่างนี้จะขึ้นอยู่กับจำนวนข้อที่ตัดออกไป กล่าวคือถ้าตัดข้อที่ลำเอียงออกไปมากเท่าใด จำนวนผู้ได้รับการคัดเลือกก็จะแตกต่างจากวิธีคะแนนดิบมากเท่านั้น ดังเห็นได้จากวิธีตัดข้อที่ลำเอียงจากวิธีโค้งลักษณะข้อสอบที่มีพารามิเตอร์ 3 ตัว และวิธีวิเคราะห์ความแปรปรวนที่ตัดข้อสอบที่ลำเอียงออกประมาณ 37% ของข้อสอบทั้งหมด ซึ่งทำให้จำนวนผู้ได้รับการคัดเลือกแตกต่างไปจากวิธีคะแนนดิบประมาณ 16% ถึง 24% อย่างไรก็ตามความแตกต่างของจำนวนผู้ได้รับการคัดเลือกนี้ยังขึ้นอยู่กับสัดส่วนในการคัดเลือกอีกด้วย กล่าวคือถ้าสัดส่วนในการคัดเลือกสูงความแตกต่างของจำนวนผู้ได้รับการคัดเลือกจะน้อย เมื่อเทียบกับสัดส่วนในการคัดเลือกที่ต่ำ

การตัดข้อสอบข้อที่ลำเอียงในการวิจัยครั้งนี้นั้นพิจารณาจากค่าสถิติที่คำนวณได้เป็นหลัก กล่าวคือถ้าปรากฏว่าข้อสอบข้อใดมีดัชนีความลำเอียงอยู่ในเกณฑ์ที่ถือว่ามีความลำเอียงในแต่ละวิธีวิเคราะห์ ก็จะตัดข้อสอบนั้น ๆ ออก ซึ่งการตัดข้อสอบออกในลักษณะเช่นนี้อาจไม่ถูกต้องนักเพราะเกณฑ์สำคัญประการหนึ่งในการพิจารณาคุณภาพของข้อสอบก็คือความตรง (Validity) นั่นคือถึงแม้ว่าข้อสอบข้อนั้นจะมีความลำเอียงแต่หากมีความตรงดีก็ไม่ควรจะตัดข้อสอบนั้นออก ควรหาทางแก้ไขปรับปรุงข้อสอบเหล่านี้ให้ไม่ลำเอียงจะดีกว่า

4. เมื่อพิจารณาในรายละเอียดเกี่ยวกับสัดส่วนของชาย : หญิงที่ได้รับการคัดเลือกปรากฏว่าหากไม่มีการศึกษาความลำเอียงของข้อสอบ และใช้วิธีคะแนนดิบเป็นวิธีการในการคัดเลือกแล้วผู้สอบหญิงมีโอกาสจะได้รับการคัดเลือกน้อยกว่าผู้สอบชาย โดยผู้สอบชายมีโอกาสจะได้รับการคัดเลือกประมาณ 54% ถึง 57% ในขณะที่ผู้สอบหญิงมีโอกาสได้รับการคัดเลือกประมาณ 43% ถึง 46% ซึ่งเป็นการไม่ยุติธรรมแก่ผู้สอบหญิง แต่หลังจากที่ได้ตัดข้อสอบที่ลำเอียงออกไปแล้ว จะเห็นว่าสัดส่วนของผู้สอบหญิงที่ได้รับการคัดเลือกเพิ่มขึ้น โดยสัดส่วนของผู้สอบหญิงที่ได้รับการคัดเลือกจะเพิ่มขึ้นตามจำนวนข้อที่ลำเอียงต่อต้านกลุ่มหญิงที่ถูกตัดออก ดังจะเห็นได้จากวิธีวิเคราะห์ความแปรปรวนที่ตัดข้อสอบที่ลำเอียงต่อต้านกลุ่มชายจำนวน 38 ข้อ และตัดข้อสอบที่ลำเอียงต่อต้านกลุ่มหญิงจำนวน 60 ข้อ ที่ปรากฏว่าเมื่อรวมคะแนนและจัดอันดับผู้สอบหลังจากที่ตัดข้อสอบที่ลำเอียงออกแล้ว สัดส่วนของผู้สอบหญิงจะได้รับการคัดเลือกเพิ่มจาก 43% ถึง 46% มาเป็น 47% ถึง 48% ซึ่งทำให้ใกล้เคียงกับสัดส่วนชายที่ได้รับการคัดเลือก ส่วนวิธีโค้งลักษณะข้อสอบที่มีพารามิเตอร์ 3 ตัวซึ่งตัดข้อสอบที่ลำเอียงต่อต้านกลุ่มชายจำนวน 32 ข้อ และตัดข้อสอบที่ลำเอียงต่อต้านกลุ่มหญิงจำนวน 67 ข้อ ก็มีผลเช่นเดียวกัน คือทำให้สัดส่วนของผู้สอบหญิงได้รับการคัดเลือกมากขึ้น คือมีโอกาสได้รับการคัดเลือกอยู่ระหว่าง 47% ถึง 49% อย่างไรก็ตามถึงแม้ว่าสัดส่วนของผู้สอบหญิงที่ได้รับการคัดเลือกจะเพิ่มขึ้นกว่าวิธีคะแนนดิบไม่มากนัก แต่ผู้วิจัยเห็นว่าการตัดข้อสอบที่มีความลำเอียงออก แล้วจึงทำการรวมคะแนนเพื่อจัดอันดับเป็นวิธีการที่ถูกต้องและให้ความยุติธรรมแก่ผู้สอบ โดยไม่คำนึงถึงความแตกต่างในเรื่องเพศทั้งนี้เพราะผลการสอบเป็นการตัดสินอนาคตของผู้สอบ ดังนั้นถึงแม้ว่าสัดส่วนของผู้สอบจะเพิ่มขึ้นเพียงเล็กน้อยก็เป็นสิ่งที่ควรจะต้องได้รับการ

พิจารณา นอกจากนี้ผู้วิจัยยังเห็นว่าน่าจะต้องคำนึงถึงความยุติธรรมในด้านอื่น ๆ อีกเช่น เชื้อชาติ ศาสนา วัฒนธรรม สภาพภูมิศาสตร์ และอื่น ๆ

ข้อค้นพบสำคัญประการหนึ่งจากผลการวิจัยนี้ก็คือ หลังจากตัดข้อสอบข้อที่ลำเอียงในแต่ละวิธีวิเคราะห์ออกไป สัดส่วนของผู้สอบหญิงที่ได้รับการคัดเลือกจะสูงขึ้นกว่าวิธีคะแนนดิบ และมีแนวโน้มว่าสัดส่วนของหญิงและชายที่ได้รับการคัดเลือกจะใกล้เคียงกันมากขึ้น เหตุที่เป็นเช่นนี้เนื่องมาจากแบบสอบที่ใช้ในการคัดเลือกครั้งนี้ มีจำนวนข้อสอบที่ลำเอียงต่อต้านผู้สอบหญิงมากกว่าผู้สอบชาย กล่าวคือในแบบสอบทั้ง 4 ฉบับนั้น วิธีวิเคราะห์ความแปรปรวนพบว่ามีข้อที่ลำเอียงต่อต้านกลุ่มหญิง 60 ข้อ และลำเอียงต่อต้านกลุ่มชาย 38 ข้อ วิธีแปลงค่าความยากของข้อสอบ พบจำนวนข้อที่ลำเอียงต่อต้านกลุ่มหญิง 20 ข้อ และลำเอียงต่อต้านกลุ่มชายจำนวน 16 ข้อ วิธีโค้งลักษณะข้อสอบที่มีพารามิเตอร์ 1 ตัว พบจำนวนข้อที่ลำเอียงต่อต้านกลุ่มหญิง 35 ข้อ ลำเอียงต่อต้านกลุ่มชาย 29 ข้อ ในขณะที่วิธีโค้งลักษณะข้อสอบที่มีพารามิเตอร์ 3 ตัว พบว่ามีข้อสอบที่ลำเอียงต่อต้านกลุ่มหญิงจำนวน 67 ข้อ และลำเอียงต่อต้านกลุ่มชาย 32 ข้อ จากข้อค้นพบดังกล่าวแสดงให้เห็นว่าในการสอบคัดเลือกนั้น ปัญหาเกี่ยวกับความลำเอียงของข้อสอบเป็นปัญหาที่ควรจะต้องนำมาพิจารณา และพยายามหาวิธีการที่จะทำให้ข้อสอบปราศจากความลำเอียงก่อนการนำข้อสอบไปใช้

5. ในเรื่องเกี่ยวกับความเที่ยงของแบบสอบนั้น พบว่าเมื่อตัดข้อสอบข้อที่ลำเอียงออกไปยิ่งมากเท่าไร ก็จะทำให้ค่าความเที่ยงของแบบสอบลดลงมากขึ้นเท่านั้น ข้อค้นพบนี้สอดคล้องกับทฤษฎีการวัดผลที่เกี่ยวกับองค์ประกอบที่มีอิทธิพลต่อค่าความเที่ยงของแบบสอบที่ว่า ข้อสอบยิ่งยาวความเที่ยงยิ่งสูง ข้อสอบยิ่งสั้นความเที่ยงยิ่งต่ำ อย่างไรก็ตามเมื่อทำการปรับขยายให้แบบสอบแต่ละฉบับมีจำนวนข้อสอบเท่ากับแบบสอบเดิมในทุกวิธีวิเคราะห์ แล้วคำนวณค่าความเที่ยงตามสูตรของ Spearman - Brown จะเห็นว่าค่าความเที่ยงของแบบสอบส่วนใหญ่หลังจากที่ตัดข้อสอบที่ลำเอียงออกไป ในแต่ละวิธีวิเคราะห์จะมีค่าความเที่ยงสูงกว่าค่าความเที่ยงของแบบสอบเดิม ยกเว้นแบบสอบฉบับที่ 4 ที่ค่าความเที่ยงต่ำกว่าแบบสอบเดิมเล็กน้อยจากข้อค้นพบนี้ทำให้อาจสรุปได้ว่าถึงแม้จะตัดข้อสอบข้อที่ลำเอียงออกไป ค่าความเที่ยงของแบบสอบที่เหลืออยู่ก็ยังเป็นค่าความเที่ยงที่สูงเมื่อเทียบกับแบบสอบเดิม

เหตุที่ค่าความเที่ยงของแบบสอบที่ตัดข้อที่ลำเอียงตามวิธีแปลงค่าความยากของข้อสอบให้ค่าความเที่ยงสูงกว่าค่าความเที่ยงจากการตัดข้อสอบตามวิธีการวิเคราะห์แบบอื่น ๆ นั้นน่าจะเนื่องมาจากวิธีแปลงค่าความยากของข้อสอบได้ตัดข้อสอบออกไปเป็นจำนวนน้อยที่สุด เมื่อเทียบกับวิธีวิเคราะห์แบบอื่น ๆ การที่วิธีแปลงค่าความยากของข้อสอบค้นพบจำนวนข้อสอบที่ลำเอียงได้น้อยกว่าวิธีวิเคราะห์อื่น ๆ เป็นเพราะเกณฑ์ที่กำหนดไว้ในการตัดสินความลำเอียงของข้อสอบค่อนข้างสูง กล่าวคือข้อสอบจะลำเอียงก็ต่อเมื่อข้อสอบข้อนั้นอยู่ห่างจากเส้นแกนหลัก $\geq \pm 3$ Sd

6. การศึกษาครั้งนี้ไม่ได้พิจารณาดัชนีความลำเอียงในลักษณะที่หาความแตกต่างของพื้นที่ตามช่วงที่โค้งลักษณะข้อสอบตัดกัน แล้วนำความแตกต่างในแต่ละช่วงมารวมกัน หากหาดัชนีความลำเอียงของข้อสอบตามวิธีนี้แล้ว ผลที่ได้จะแตกต่างไปจากวิธีที่ผู้วิจัยใช้ในการศึกษาในกรณีที่โค้งลักษณะข้อสอบจากผู้สอบ 2 กลุ่มตัดกัน แต่ในกรณีที่โค้งลักษณะของข้อสอบจากผู้สอบ 2 กลุ่มไม่ตัดกัน ดัชนีความลำเอียงตามวิธีการที่กล่าวมา กับวิธีการที่ผู้วิจัยศึกษาก็จะไม่แตกต่างกัน จำนวนข้อที่พบว่ามีความลำเอียงหากหาโดยวิธีที่ผู้วิจัยกล่าวถึงกับวิธีที่ผู้วิจัยใช้ในการศึกษาครั้งนี้ ก็จะแตกต่างกันด้วย ซึ่งสามารถเห็นได้จากตาราง 55 ถึง 58 ในภาคผนวก ก. เหตุที่เสนอแนะให้พิจารณาดัชนีความลำเอียงในลักษณะดังกล่าว ก็เพราะเห็นว่าน่าจะสอดคล้องกับหลักการในทฤษฎีการตอบสนองต่อข้อสอบที่จะพิจารณาโอกาสในการตอบข้อสอบแต่ละข้อถูกตามระดับความสามารถที่กำหนด อย่างไรก็ตามการหาดัชนีความลำเอียงตามวิธีการดังกล่าว จะไม่สามารถบอกได้ว่าโดยสรุปแล้วข้อสอบข้อนั้นลำเอียงต่อต้านกลุ่มใด เพราะข้อสอบเดียวกันอาจลำเอียงต่อต้านกลุ่มหนึ่งในระดับความสามารถหนึ่ง แต่จะลำเอียงต่อต้านอีกกลุ่มหนึ่งในอีกระดับความสามารถหนึ่ง การนำความแตกต่างของพื้นที่ใน 2 ช่วงมารวมกัน จึงบอกไม่ได้ว่าข้อสอบข้อนั้นลำเอียงต่อต้านกลุ่มใด การแปลผลดัชนีความลำเอียงในลักษณะนี้ จึงไม่ควรแปลผลรวม แต่ควรแปลผลเป็นช่วง ๆ ไป ว่าข้อสอบข้อนั้นลำเอียงต่อต้านผู้สอบกลุ่มใด ในช่วงระดับความสามารถใด

ข้อเสนอแนะ

ข้อเสนอแนะจากการวิจัยครั้งนี้ สามารถแยกได้ 2 ประการคือ

1. ข้อเสนอแนะในการนำไปใช้ ข้อค้นพบเกี่ยวกับการเปรียบเทียบผลของวิธีวิเคราะห์ความลำเอียงของข้อสอบที่แตกต่างกัน 4 วิธีนั้น สามารถนำไปประยุกต์ใช้ได้ดังนี้

1.1 ในการสร้างแบบสอบนั้นควรมีการกำหนดให้การวิเคราะห์เนื้อหาเพื่อหาความลำเอียงของข้อสอบเป็นขั้นตอนหนึ่งของการสร้างแบบสอบ ทั้งนี้เพื่อจะช่วยลดความลำเอียงที่อาจเกิดขึ้น ซึ่งวิธีการดังกล่าวนี้จะเป็นการตรวจสอบด้วยการวิเคราะห์เชิงเหตุผลก่อนที่จะนำแบบสอบไปใช้จริง

1.2 ในการจะตัดสินว่าควรใช้ข้อสอบที่มีความลำเอียงต่อไปหรือไม่นั้นขึ้นอยู่กับปัจจัยหลายอย่าง เช่น จุดประสงค์ของการสอบ ข้อสอบที่ลำเอียงนั้นจะไม่ได้ลำเอียงตลอดไปในทุก ๆ สถานการณ์ และข้อที่ลำเอียงก็ไม่ได้หมายความว่าเป็นข้อที่ไม่ดีหรือขาดคุณภาพในการวัด ผลจากการศึกษาครั้งนี้พบว่ามีข้อสอบจำนวนมากที่มีค่าดัชนีความลำเอียงต่ำ อยู่ในเกณฑ์ที่สามารถนำไปใช้ได้ หรือแม้กระทั่งข้อที่มีความลำเอียงสูงก็อาจนำไปใช้กับกลุ่มอื่น ที่ไม่ได้คำนึงถึงความแตกต่างในเรื่องเพศได้

2. ข้อเสนอแนะเพื่อการศึกษาเพิ่มเติม

2.1 ควรมีการศึกษาแหล่งหรือสาเหตุที่ทำให้เกิดความลำเอียงของข้อสอบ โดยอาจให้ครู หรือผู้สร้างแบบสอบเป็นผู้วิเคราะห์ข้อสอบที่พบว่ามีความลำเอียงตามวิธีการทางสถิติว่าความลำเอียงนั้นน่าจะมาจากสาเหตุใด อันจะเป็นประโยชน์ในการปรับปรุงแก้ไขข้อสอบที่ลำเอียงให้สามารถใช้ได้อย่างยุติธรรมสำหรับนักเรียนต่างกลุ่มกัน ซึ่งจะทำให้ผลการคัดเลือกมีความถูกต้องและยุติธรรมมากขึ้น

2.2 ควรจะมีการศึกษาเปรียบเทียบวิธีวิเคราะห์ความลำเอียงของข้อสอบโดยใช้ข้อมูลสมมุติ (Simulate Data) ทั้งนี้เพื่อยืนยันข้อความรู้เกี่ยวกับเรื่องความลำเอียงของข้อสอบในอีกแง่มุมหนึ่ง

2.3 ควรทำการศึกษาเกี่ยวกับวิธีการวิเคราะห์ความลำเอียงของข้อสอบโดยใช้เกณฑ์ภายนอก ซึ่งจะต้องพยายามหาตัวแปรเกณฑ์ภายนอกที่มีความตรงเชิงพยากรณ์และมีความยุติธรรม

2.4 ควรจะทำการศึกษาถึงความสัมพันธ์ระหว่างวิธีวิเคราะห์ความลำเอียงโดยใช้ค่าดัชนีความลำเอียงตามวิธีการทางสถิติกับดัชนีความลำเอียงที่ได้จากการให้ผู้เชี่ยวชาญ หรือนักวัดผลเป็นผู้ประเมิน

2.5 น่าจะมีการศึกษาถึงวิธีการในการหาประสิทธิภาพของวิธีวิเคราะห์ความลำเอียงของข้อสอบ วิธีการหนึ่งซึ่งอาจเป็นไปได้ก็คือ การจงใจสร้างข้อสอบให้มีความลำเอียง เช่นสร้างคำถามให้เข้าข้างผู้สอบเพศชาย หรือเข้าข้างคนในกรุงเทพฯ แล้วดูว่าวิธีวิเคราะห์ใดสามารถค้นพบจำนวนข้อสอบที่ลำเอียงได้มากที่สุด วิธีการนั้นก็อาจกล่าวได้ว่าเป็นวิธีที่มีประสิทธิภาพในการวิเคราะห์ความลำเอียงของข้อสอบ