

บทที่ 5

สรุป อภิปราย และข้อเสนอแนะ

การวิจัยครั้งนี้มีวัตถุประสงค์หลัก เพื่อวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ (DIF) ในแบบสอบวิชาภาษาไทยและวิชาภาษาอังกฤษ ด้วยวิธี MH สำหรับกลุ่มผู้สอบจำแนกตามเพศ ภูมิภาค ประสิทธิภาพในการสอบ และสังกัดของสถานศึกษา โดยมีวัตถุประสงค์เฉพาะ 3 ข้อ คือ

1. เพื่อศึกษาการระบุลักษณะของข้อสอบที่ทำหน้าที่ต่างกันแบบเอกกรุป (uniform DIF) และแบบอนเอกกรุป (nonuniform DIF) สำหรับกลุ่มผู้สอบจำแนกตามเพศ ภูมิภาค ประสิทธิภาพในการสอบ และสังกัดของสถานศึกษา

2. เพื่อเปรียบเทียบค่าความยาก ค่าอำนาจจำแนก และลักษณะเนื้อหาของข้อสอบที่ทำหน้าที่ต่างกัน สำหรับกลุ่มผู้สอบจำแนกตามเพศ ภูมิภาค ประสิทธิภาพในการสอบ และสังกัดของสถานศึกษา

3. เพื่อศึกษาค่าความเที่ยง ความตรง ของแบบสอบก่อน และหลังจากตัดข้อสอบที่ทำหน้าที่ต่างกัน ในแต่ละกลุ่มผู้สอบออกจากแบบสอบ

ผลวิจัยผู้ให้ข้อมูลในการวิจัยครั้งนี้คือผู้สมัครสอบคัดเลือกวิชาสอบร่วมวิชาภาษาไทย และวิชาภาษาอังกฤษในระดับบัณฑิตศึกษา ของศูนย์ทดสอบทางการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย รหัสการสอบที่ 383 ซึ่งประกอบด้วยผู้สอบวิชาภาษาไทยจำนวน 506 คน และผู้สอบวิชาภาษาอังกฤษจำนวน 501 คน

เครื่องมือวิจัยประกอบด้วย แบบสอบวิชาภาษาไทย (ศทท 383) สร้างโดยคณะกรรมการพัฒนาแบบสอบภาษาไทย ของศูนย์ทดสอบทางการศึกษา ซึ่งในการวิจัยนี้ใช้เฉพาะส่วนที่เป็นแบบสอบปรนัยชนิด 4 ตัวเลือก จำนวน 95 ข้อ และแบบสอบวิชาภาษาอังกฤษ (ศทอ 383) สร้างโดยคณะกรรมการพัฒนาแบบสอบภาษาอังกฤษของศูนย์ทดสอบทางการศึกษา ซึ่งในการวิจัยนี้ใช้เฉพาะส่วนที่เป็นแบบสอบปรนัยชนิด 5 ตัวเลือก จำนวน 80 ข้อ จากนั้นผู้วิจัยทำการสุ่มข้อสอบของแต่ละฉบับออกให้เหลือเท่ากับ 75 ข้อ เพื่อใช้ในการวิเคราะห์ข้อมูลสำหรับการวิจัยในครั้งนี้

การวิเคราะห์ข้อมูล ทำการวิเคราะห์ข้อมูลเบื้องต้น โดยการหาค่าสถิติพื้นฐานของคะแนนสอบวิชาภาษาไทยภาษาอังกฤษวิเคราะห์คุณภาพของแบบสอบด้านความเที่ยงแบบความสอดคล้องภายใน (Internal Consistency) ค่าความยาก และค่าอำนาจจำแนกของข้อสอบ โดยใช้โปรแกรมการวิเคราะห์ข้อสอบแนวทฤษฎีดั้งเดิม:CTIA (สุพัฒน์ สุขมลสันต์, 2535)

วิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ (DIF) โดยใช้โปรแกรมคอมพิวเตอร์สำเร็จรูป MH_{DIF} (Fidalgo, 1995) คำนวณค่าเฉลี่ย (\bar{X}) ส่วนเบี่ยงเบนมาตรฐาน (S.D.) และค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สันของข้อสอบเป็นรายด้าน โดยใช้โปรแกรมคอมพิวเตอร์ SPSS/PC⁺ วิเคราะห์ความตรงเชิงโครงสร้างหรือความตรงตามทฤษฎีของแบบสอบวิชาภาษาไทยและภาษาอังกฤษด้วยการวิเคราะห์องค์ประกอบเชิงยืนยัน (CFA) โดยใช้โปรแกรมคอมพิวเตอร์ LISREL 8.10

สรุปผลการวิจัย

1. ผลการระบุลักษณะของข้อสอบที่ทำหน้าที่ต่างกัน (DIF) จากการวิเคราะห์แบบเอกรูปและแบบอเนกรูป ในวิชาภาษาไทยและภาษาอังกฤษ พบว่า ส่วนใหญ่เป็นข้อสอบที่ DIF จากการวิเคราะห์แบบอเนกรูป และผลจากการวิเคราะห์ทั้งสองแบบมีความเป็นอิสระต่อกัน กล่าวคือ ข้อสอบที่ถูกระบุว่าพบ DIF จากการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบแบบเอกรูป อาจจะถูกระบุว่าเป็นข้อสอบที่พบ DIF จากการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบแบบอเนกรูป หรือไม่ก็ได้ และข้อสอบที่ระบุว่าจะไม่พบว่าทำหน้าที่ต่างกันระหว่างกลุ่ม (no DIF) จากการวิเคราะห์แบบเอกรูปอาจพบว่าเป็นข้อสอบที่พบ DIF เมื่อวิเคราะห์แบบอเนกรูป ก็ได้

สำหรับผลการเปรียบเทียบจำนวนข้อสอบที่พบ DIF เมื่อจำแนกกลุ่มผู้สอบตามเพศ ภูมิสำเนา ประสบการณ์ในการสอบ และสังกัดของสถานศึกษา จากการวิเคราะห์แบบเอกรูปและแบบอเนกรูป พบว่า ในวิชาภาษาไทย เมื่อจำแนกผู้สอบตามตัวแปรเพศ มีข้อสอบที่พบ DIF มากที่สุด จำนวน 14 ข้อ (ร้อยละ 18.67) รองลงมาคือเมื่อจำแนกกลุ่มผู้สอบตามภูมิสำเนา จำนวน 12 ข้อ (ร้อยละ 16.00) และสังกัดของสถานศึกษา จำนวน 10 ข้อ (ร้อยละ 13.33) ส่วนประสบการณ์ในการสอบเป็นตัวแปรที่มีข้อสอบที่พบ DIF น้อยที่สุดเพียง 1 ข้อ (ร้อยละ 1.33) นอกจากนี้ในแบบสอบวิชาภาษาอังกฤษ พบว่าตัวแปรในการจำแนกกลุ่มผู้สอบที่พบ DIF มากที่สุดคือ เมื่อจำแนกผู้สอบตามตัวแปรเพศ ซึ่งเท่ากับเมื่อจำแนกผู้สอบตามตัวแปรภูมิสำเนา จำนวน 8 ข้อ (ร้อยละ 10.67) รองลงมาคือเมื่อจำแนกผู้สอบตามสังกัดของสถานศึกษา จำนวน 7 ข้อ (ร้อยละ 9.33) และตัวแปรที่มีข้อสอบที่พบ DIF น้อยที่สุด คือ เมื่อจำแนกผู้สอบตามประสบการณ์ในการสอบ จำนวน 5 ข้อ (ร้อยละ 6.67)

2. ลักษณะของข้อสอบที่พบ DIF จากการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบแบบเอกรูป และแบบอเนกรูป กรณีจำแนกผู้สอบตามเพศ ภูมิสำเนา ประสบการณ์ในการสอบ และสังกัดของสถานศึกษา เมื่อพิจารณาข้อสอบที่พบ DIF ในด้านลักษณะของเนื้อหา ค่าความยาก และค่าอำนาจจำแนก ของแบบสอบวิชาภาษาไทย และภาษาอังกฤษ สรุปได้ดังนี้

2.1 วิชาภาษาไทย

2.1.1 เมื่อจำแนกผู้สอบตามเพศ พบว่า มีข้อสอบที่พบ DIF จากการวิเคราะห์แบบอนุกรมและแบบเอกรูป จำนวน 13 ข้อ และ 5 ข้อ คิดเป็นร้อยละ 17.33 และร้อยละ 6.67 ตามลำดับ เมื่อพิจารณาโดยรวมจากการวิเคราะห์ทั้งสองแบบพบว่า มีข้อสอบที่พบ DIF ทั้งหมด 14 ข้อ โดยข้อสอบที่พบ DIF เมื่อจำแนกผู้สอบตามเพศ เป็นข้อสอบที่เข้าข้างกลุ่มผู้สอบเพศหญิงจำนวน 11 ข้อ และเข้าข้างกลุ่มผู้สอบเพศชาย 3 ข้อ คิดเป็นร้อยละ 78.57 และร้อยละ 21.43 ตามลำดับ

สำหรับลักษณะเนื้อหาของข้อสอบที่พบ DIF เมื่อจำแนกผู้สอบตามเพศนั้น พบว่าเป็นข้อสอบที่มีเนื้อหาเกี่ยวกับหลักภาษาไทยมากที่สุด จำนวน 5 ข้อ คิดเป็นร้อยละ 35.71 และจากข้อสอบที่พบ DIF ทั้งหมด (14ข้อ) พบว่า ข้อสอบที่พบ DIF ส่วนใหญ่ร้อยละ 42.86 เป็นข้อสอบที่ง่ายมาก (มีค่าความยากอยู่ในช่วง 0.80 - 1.00) รองลงมาเป็นข้อสอบที่ยากปานกลาง (มีค่าความยากอยู่ในช่วง 0.40 - 0.59) และเป็นข้อสอบที่ค่อนข้างง่าย (มีค่าความยากอยู่ในช่วง 0.80 - 1.00) คิดเป็นร้อยละ 5.71 และ 14.29 ตามลำดับ และข้อสอบที่พบ DIF ส่วนใหญ่ร้อยละ 64.28 มีค่าอำนาจจำแนกค่อนข้างต่ำ (มีค่าอยู่ในช่วง 0.20 - 0.39)

เมื่อพิจารณาการวิเคราะห์ DIF แบบอนุกรม ซึ่งมีข้อสอบที่พบ DIF ทั้งหมดจำนวน 13 ข้อนั้น พบว่า เป็นข้อสอบที่เข้าข้างกลุ่มผู้สอบเพศหญิงที่ระดับความสามารถสูงมากที่สุด จำนวน 6 ข้อ คิดเป็นร้อยละ 46.15 รองลงมาเป็นข้อสอบที่เข้าข้างกลุ่มผู้สอบเพศหญิงที่ระดับความสามารถต่ำจำนวน 4 ข้อ คิดเป็นร้อยละ 30.77 ส่วนข้อสอบที่เข้าข้างกลุ่มผู้สอบเพศชายที่ระดับความสามารถสูงจำนวน 2 ข้อ คิดเป็นร้อยละ 15.38 และเป็นข้อสอบที่เข้าข้างกลุ่มผู้สอบเพศชายที่ระดับความสามารถต่ำน้อยที่สุดเพียง 1 ข้อ ส่วนในการวิเคราะห์ DIF แบบเอกรูปมีข้อสอบที่พบ DIF ทั้งหมดจำนวน 5 ข้อ เป็นข้อสอบที่เข้าข้างกลุ่มผู้สอบเพศหญิงจำนวน 3 ข้อ และเข้าข้างกลุ่มผู้สอบเพศชาย 2 ข้อ คิดเป็นร้อยละ 60.00 และร้อยละ 40.00 ตามลำดับ

2.1.2 เมื่อจำแนกผู้สอบตามภูมิภาค พบว่ามีข้อสอบที่พบ DIF จากการวิเคราะห์แบบอนุกรมและแบบเอกรูป จำนวน 10 ข้อ และ 4 ข้อ คิดเป็นร้อยละ 13.33 และร้อยละ 6.67 ตามลำดับ เมื่อพิจารณาโดยรวมจากการวิเคราะห์ทั้งสองแบบพบว่า มีข้อสอบที่พบ DIF ทั้งหมด 12 ข้อ โดยข้อสอบที่พบ DIF เมื่อจำแนกผู้สอบตามภูมิภาคเป็นข้อสอบที่เข้าข้างกลุ่มผู้สอบในเขตกรุงเทพมหานครและปริมณฑล จำนวน 9 ข้อ และเป็นข้อสอบที่เข้าข้างกลุ่มผู้สอบนอกเขตกรุงเทพมหานครและปริมณฑล จำนวน 3 ข้อ คิดเป็นร้อยละ 75.00 และ 25.00 ตามลำดับ

สำหรับลักษณะเนื้อหาของข้อสอบที่พบ DIF เมื่อจำแนกผู้สอบตามภูมิภาค พบว่า เป็นข้อสอบที่มีเนื้อหาเกี่ยวกับด้านตรรกวิทยา ด้านการใช้ภาษา 1 และด้านวรรณคดี

มากที่สุด จำนวนเท่า ๆ กันด้านละ 3 ข้อ คิดเป็นร้อยละ 25.00 โดยข้อสอบที่ DIF ส่วนใหญ่ ร้อยละ 33.33 เป็นข้อสอบที่ง่ายมาก (มีค่าความยากอยู่ในช่วง 0.80 - 1.00) รองลงมา เป็นข้อสอบที่ค่อนข้างง่าย (มีค่าความยากอยู่ในช่วง 0.60 - 0.79) และเป็นข้อสอบที่ยาก ปานกลาง (มีค่าความยากอยู่ในช่วง 0.40 - 0.59) คิดเป็นร้อยละ 25.00 เท่ากัน นอกจากนี้ ยังพบว่าข้อสอบที่พบ DIF ส่วนใหญ่ร้อยละ 50.00 เป็นข้อสอบที่มีค่าอำนาจจำแนกต่ำ (มีค่า อยู่ในช่วง 0.10 - 0.19)

เมื่อพิจารณาการวิเคราะห์ DIF แบบอนุกรม ซึ่งมีข้อสอบที่พบ DIF ทั้งหมด จำนวน 10 ข้อ นั้น เป็นข้อสอบที่เข้าข้างกลุ่มผู้สอบในเขตกรุงเทพมหานครและปริมณฑล ที่ ระดับความสามารถต่ำมากที่สุดจำนวน 4 ข้อ คิดเป็นร้อยละ 40.00 รองลงมาเป็นข้อสอบที่ เข้าข้างกลุ่มผู้สอบในเขตกรุงเทพมหานครที่ระดับความสามารถสูงจำนวน 3 ข้อ เป็นข้อสอบ ที่เข้าข้างกลุ่มผู้สอบนอกเขตกรุงเทพมหานครและปริมณฑลที่ระดับความสามารถต่ำจำนวน 3 ข้อ คิดเป็นร้อยละ 30.00 เท่ากัน ส่วนในกลุ่มผู้สอบนอกเขตกรุงเทพมหานครและปริมณฑล ที่ระดับความสามารถสูงไม่มีข้อสอบที่พบ DIF ส่วนในการวิเคราะห์ DIF แบบเอกรูป มีข้อสอบ ที่พบ DIF ทั้งหมดจำนวน 4 ข้อ เป็นข้อสอบที่เข้าข้างกลุ่มผู้สอบในเขตกรุงเทพมหานครและ ปริมณฑลหมดทั้ง 4 ข้อ

2.1.3 เมื่อจำแนกผู้สอบตามสังกัดของสถานศึกษา พบว่า มีข้อสอบที่พบ DIF จากการวิเคราะห์แบบเอกรูปและแบบอนุกรม จำนวน 7 ข้อ และ 5 ข้อ คิดเป็นร้อยละ 9.33 และ 6.67 ตามลำดับ เมื่อพิจารณาโดยรวมจากการวิเคราะห์ทั้งสองแบบพบว่ามีข้อสอบที่พบ DIF ทั้งหมด 10 ข้อ เป็นข้อสอบที่เข้าข้างกลุ่มผู้สอบที่สำเร็จการศึกษาในสังกัดทบวง มหาวิทยาลัยและนอกสังกัดทบวงมหาวิทยาลัย จำนวน 7 ข้อ และ 3 ข้อ คิดเป็นร้อยละ 70.00 และ 30.00 ตามลำดับ

สำหรับลักษณะเนื้อหาของข้อสอบที่พบ DIF เมื่อจำแนกผู้สอบตามสังกัดของ สถานศึกษานั้น พบว่าเป็นข้อสอบที่มีเนื้อหาเกี่ยวกับด้านการใช้ภาษา 1 จำนวน 3 ข้อ คิดเป็น ร้อยละ 30.00 ซึ่งเท่ากับด้านวรรณคดี และด้านครุวิทยา รองลงมาเป็นข้อสอบที่มีเนื้อหา ด้านหลักภาษา ซึ่งมี 1 ข้อ (ร้อยละ 10.00) โดยที่ข้อสอบที่พบ DIF ส่วนใหญ่ร้อยละ 40.00 เป็นข้อสอบที่ง่ายมาก (มีค่าความยากอยู่ในช่วง 0.80 - 1.00) รองลงมาเป็นข้อสอบที่ ยากปานกลาง (มีค่าความยากอยู่ในช่วง 0.40 - 0.59) ค่อนข้างง่าย (มีค่าความยากอยู่ใน ช่วง 0.60 - 0.79) และค่อนข้างยาก (มีค่าความยากอยู่ในช่วง 0.20 - 0.30) คิดเป็นร้อยละ 30.00, 20.00 และ 10.00 ตามลำดับ นอกจากนี้ยังพบว่าข้อสอบที่ DIF ส่วนใหญ่ร้อยละ 40.00 มีค่าอำนาจจำแนกค่อนข้างต่ำ (มีค่าอำนาจจำแนกอยู่ในช่วง 0.20 - 0.39)

เมื่อพิจารณาการวิเคราะห์ DIF แบบเอกรูป ซึ่งมีข้อสอบที่พบ DIF ทั้งหมด จำนวน 7 ข้อ เป็นข้อสอบที่เข้าข้างกลุ่มผู้สอบที่สำเร็จการศึกษาสังกัดทบวงมหาวิทยาลัย

จำนวน 5 ข้อ และเข้าข้างกลุ่มผู้สอบที่สำเร็จการศึกษานอกสังกัดทบวงมหาวิทยาลัยจำนวน 2 ข้อ คิดเป็นร้อยละ 71.43 และ 28.57 ตามลำดับ ส่วนในการวิเคราะห์ DIF แบบบอเนกรูป พบว่า มีข้อสอบที่พบ DIF จำนวน 5 ข้อ เป็นข้อสอบที่เข้าข้างกลุ่มผู้สอบที่สำเร็จการศึกษานอกสังกัดทบวงมหาวิทยาลัยที่ระดับความสามารถต่ำมากที่สุด 3 ข้อ คิดเป็นร้อยละ 60.00 รองลงมาเป็นข้อสอบที่เข้าข้างกลุ่มผู้สอบที่สำเร็จการศึกษานอกสังกัดทบวงมหาวิทยาลัยที่ระดับความสามารถสูง 1 ข้อ และเป็นข้อสอบที่เข้าข้างกลุ่มผู้สอบที่สำเร็จการศึกษานอกสังกัดทบวงมหาวิทยาลัยที่มีความสามารถต่ำอีก 1 ข้อ คิดเป็นร้อยละ 20.00 เท่ากัน

2.1.4 เมื่อจำแนกผู้สอบตามประเภทการดำเนินการสอบ พบว่า ในการวิเคราะห์ แบบบอเนกรูปไม่มีข้อสอบที่พบ DIF ส่วนในการวิเคราะห์แบบเอกรูป มีข้อสอบที่พบ DIF เพียง 1 ข้อ คิดเป็นร้อยละ 1.33 ซึ่งเป็นข้อสอบที่เข้าข้างกลุ่มผู้สอบที่ไม่มีประสบการณ์ในการสอบ โดยเป็นข้อสอบที่มีเนื้อหาด้านการใช้ภาษา 1 และพบว่าข้อสอบที่พบ DIF เป็นข้อสอบที่มีค่าความยากและค่าอำนาจจำแนกเท่ากับ 0.45 และ 0.32 ตามลำดับ

2.1.5 เมื่อพิจารณาลักษณะเนื้อหาของข้อสอบที่พบ DIF ในทุกเงื่อนไขของการวิเคราะห์ พบว่า แบบสอบวิชาภาษาไทยมีข้อสอบที่พบ DIF ในข้อสอบที่มีเนื้อหาด้านการใช้ภาษา 1 มากที่สุด รองลงมาเป็นข้อสอบที่มีเนื้อหาด้านหลักภาษาและด้านวรรณคดี ส่วนเนื้อหาที่มีข้อสอบที่พบ DIF น้อยที่สุด คือข้อสอบที่มีเนื้อหาด้านการใช้ภาษา 2 นอกจากนี้พบว่าข้อสอบที่พบ DIF ส่วนใหญ่เป็นข้อสอบที่มีง่ายมาก (0.80 - 1.00) และค่าอำนาจจำแนกค่อนข้างต่ำ (0.20 - 0.39)

2.2 วิชาภาษาอังกฤษ

2.2.1 เมื่อจำแนกผู้สอบตามเพศ พบว่ามีข้อสอบที่พบ DIF จากการวิเคราะห์ แบบบอเนกรูปและแบบเอกรูป จำนวน 6 ข้อ และ 2 ข้อ คิดเป็นร้อยละ 8.00 และ 2.67 ตามลำดับ เมื่อพิจารณาโดยรวมจากการวิเคราะห์ทั้งสองแบบ พบว่า มีข้อสอบที่พบ DIF ทั้งหมดจำนวน 8 ข้อ โดยข้อสอบที่พบ DIF เมื่อจำแนกผู้สอบตามเพศ เป็นข้อสอบที่เข้าข้างกลุ่มผู้สอบเพศหญิงและเพศชายกลุ่มละ 4 ข้อ คิดเป็นร้อยละ 50.00 เท่ากัน

สำหรับลักษณะเนื้อหาของข้อสอบที่พบ DIF เมื่อจำแนกผู้สอบตามเพศ พบว่า เป็นข้อสอบที่มีเนื้อหา ด้านการบอกความหมายคำศัพท์มากที่สุด จำนวน 3 ข้อ คิดเป็นร้อยละ 37.50 รองลงมาเป็นข้อสอบที่มีเนื้อหา ด้านการให้รายละเอียด จำนวน 2 ข้อ คิดเป็นร้อยละ 25.00 สำหรับเนื้อหา ด้านการจับใจความสำคัญและด้านการสรุปใจความไม่มีข้อสอบที่พบ DIF นอกจากนี้พบว่าข้อสอบที่พบ DIF ส่วนใหญ่ร้อยละ 50.00 เป็นข้อสอบที่ยากมาก (มีค่าความยากอยู่ในช่วง 0.00 - 0.19) รองลงมาเป็นข้อสอบที่ค่อนข้างยาก (มีค่าความยากอยู่ในช่วง

0.20 - 0.39) และยากปานกลาง (มีค่าความยากอยู่ในช่วง 0.40 - 0.59) คิดเป็นร้อยละ 25.00 เท่ากัน และพบว่าข้อสอบที่ DIF ส่วนใหญ่ร้อยละ 50.00 มีค่าอำนาจจำแนกค่อนข้างต่ำ (มีค่าอำนาจจำแนกอยู่ในช่วง 0.20 - 0.39)

เมื่อพิจารณาการวิเคราะห์ DIF แบบอนุกรม ซึ่งมีข้อสอบที่พบ DIF จำนวน 6 ข้อ นั้น พบว่า เป็นข้อสอบที่เข้าข้างเพศหญิงที่ระดับความสามารถต่ำ จำนวน 2 ข้อ เท่ากับ ข้อสอบที่เข้าข้างกลุ่มผู้สอบเพศชายที่ระดับความสามารถต่ำ 2 ข้อ เป็นข้อสอบที่เข้าข้างกลุ่มผู้สอบเพศหญิงที่ระดับความสามารถสูง 1 ข้อ เท่ากับข้อสอบที่เข้าข้างกลุ่มผู้สอบเพศชายที่ระดับความสามารถสูง 1 ข้อ ส่วนในการวิเคราะห์ DIF แบบเอกรูป ซึ่งมีข้อสอบที่พบ DIF จำนวน 2 ข้อ นั้นเป็นข้อสอบที่เข้าข้างกลุ่มผู้สอบเพศหญิงและกลุ่มผู้สอบเพศชายกลุ่มละ 1 ข้อ เท่ากัน

2.2.2 เมื่อจำแนกผู้สอบตามภูมิสำเนา พบว่า มีข้อสอบที่พบ DIF จากการวิเคราะห์แบบอนุกรมและแบบเอกรูป จำนวน 5 ข้อ และ 3 ข้อ คิดเป็นร้อยละ 6.67 และ 4.00 ตามลำดับ เมื่อพิจารณาโดยรวมจากการวิเคราะห์ทั้งสองแบบ พบว่า มีข้อสอบที่พบ DIF ทั้งหมด จำนวน 8 ข้อ โดยข้อสอบที่พบ DIF เมื่อจำแนกผู้สอบตามภูมิสำเนาเป็นข้อสอบที่เข้าข้างกลุ่มผู้สอบในเขตกรุงเทพมหานครและปริมณฑล จำนวน 5 ข้อ และเป็นข้อสอบที่เข้าข้างกลุ่มผู้สอบนอกเขตกรุงเทพมหานครและปริมณฑล จำนวน 3 ข้อ คิดเป็นร้อยละ 62.50 และ 37.50 ตามลำดับ

สำหรับลักษณะเนื้อหาของข้อสอบที่พบ DIF เมื่อจำแนกผู้สอบตามภูมิสำเนา พบว่า เป็นข้อสอบที่มีเนื้อหาด้านการให้รายละเอียด และด้านการบอกความหมายคำศัพท์มากที่สุด จำนวนด้านละ 2 ข้อ คิดเป็นร้อยละ 25.00 รองลงมาเป็นข้อสอบที่มีเนื้อหาด้านการถอดความ การอ้างอิง การสรุปอิงความและด้านการใช้ภาษาอีกด้านละ 1 ข้อ สำหรับเนื้อหาด้านการจับใจความสำคัญไม่มีข้อสอบที่พบ DIF นอกจากนี้พบว่าข้อสอบที่พบ DIF ส่วนใหญ่ร้อยละ 50.00 เป็นข้อสอบที่ยากมาก (มีค่าความยากอยู่ในช่วง 0.00 - 0.19) รองลงมาเป็นข้อสอบที่ค่อนข้างยาก (มีค่าความยากอยู่ในช่วง 0.20 - 0.39) และยากปานกลาง (มีค่าความยากอยู่ในช่วง 0.40 - 0.59) คิดเป็นร้อยละ 37.50 และ 12.50 ตามลำดับ และพบว่าข้อสอบที่ DIF ส่วนใหญ่ร้อยละ 50.00 มีค่าอำนาจจำแนกต่ำมาก (มีค่าอำนาจจำแนกอยู่ในช่วง 0.00 - 0.09)

เมื่อพิจารณาการวิเคราะห์ DIF แบบอนุกรม ซึ่งมีข้อสอบที่พบ DIF จำนวน 5 ข้อ นั้นเป็นข้อสอบที่เข้าข้างกลุ่มผู้สอบในเขตกรุงเทพมหานครและปริมณฑลที่ระดับความสามารถสูงและระดับความสามารถต่ำ จำนวน 2 ข้อ และ 1 ข้อ ตามลำดับ และเป็นข้อสอบที่เข้าข้างกลุ่มผู้สอบนอกเขตกรุงเทพมหานครและปริมณฑลที่ระดับความสามารถสูงและระดับความสามารถต่ำจำนวนกลุ่มละ 1 ข้อ ส่วนในการวิเคราะห์ DIF แบบเอกรูป มีข้อสอบที่พบ

DIF จำนวน 3 ข้อ เป็นข้อสอบที่เข้าข้างกลุ่มผู้สอบในเขตกรุงเทพมหานครและปริมณฑล 2 ข้อ และเป็นข้อสอบที่เข้าข้างกลุ่มผู้สอบนอกเขตกรุงเทพมหานครและปริมณฑลอีก 1 ข้อ

2.2.3 เมื่อจำแนกผู้สอบตามสังกัดของสถานศึกษา พบว่า ในการวิเคราะห์แบบเอกรูปนั้นไม่มีข้อสอบที่พบ DIF ส่วนในการวิเคราะห์แบบบอเนกรูป พบว่า มีข้อสอบที่พบ DIF ทั้งหมดจำนวน 7 ข้อ เป็นข้อสอบที่เข้าข้างกลุ่มผู้สอบที่สำเร็จการศึกษานอกสังกัดทบวงมหาวิทยาลัยมากที่สุด จำนวน 6 ข้อ คิดเป็นร้อยละ 85.71 เป็นข้อสอบที่เข้าข้างกลุ่มผู้สอบที่สำเร็จการศึกษาในสังกัดทบวงมหาวิทยาลัยอีก 1 ข้อ คิดเป็นร้อยละ 14.29 โดยข้อสอบที่พบ DIF ส่วนใหญ่มีเนื้อหาด้านการอ้างอิง และด้านการใช้ภาษามากที่สุดด้านละ 2 ข้อ คิดเป็นร้อยละ 28.57 เท่ากัน เมื่อพิจารณาค่าความยากและค่าอำนาจจำแนกโดยรวม พบว่า ข้อสอบที่พบ DIF ส่วนใหญ่ร้อยละ 57.14 เป็นข้อสอบที่ค่อนข้างยาก (มีค่าความยากอยู่ในช่วง 0.20 - 0.39) รองลงมาเป็นข้อสอบที่ยากปานกลาง (มีค่าความยากอยู่ในช่วง 0.40 - 0.59) คิดเป็นร้อยละ 42.86 และพบว่าข้อสอบที่พบ DIF ส่วนใหญ่ร้อยละ 57.14 เป็นข้อสอบที่มีค่าอำนาจจำแนกพอใช้ได้ (มีค่าความยากอยู่ในช่วง 0.40 - 0.59)

2.2.4 เมื่อจำแนกผู้สอบตามประสบการณ์ในการสอบ พบว่า มีข้อสอบที่พบ DIF จากการวิเคราะห์แบบบอเนกรูปและแบบเอกรูป จำนวน 3 ข้อ เท่ากัน คิดเป็นร้อยละ 4.00 เมื่อพิจารณาโดยรวมจากกการวิเคราะห์ทั้งสองแบบ พบว่า มีข้อสอบที่พบ DIF ทั้งหมด จำนวน 5 ข้อ โดยข้อสอบที่พบ DIF เมื่อจำแนกผู้สอบตามประสบการณ์ในการสอบ เป็นข้อสอบที่เข้าข้างกลุ่มผู้สอบที่ไม่มีประสบการณ์ในการสอบ จำนวน 3 ข้อ และเป็นข้อสอบที่เข้าข้างกลุ่มผู้สอบที่มีประสบการณ์ในการสอบ จำนวน 2 ข้อ คิดเป็นร้อยละ 60.00 และ 40.00 ตามลำดับ

สำหรับลักษณะเนื้อหาของข้อสอบที่พบ DIF เมื่อจำแนกผู้สอบตามประสบการณ์ในการสอบ พบว่า เป็นข้อสอบที่มีเนื้อหาด้านการให้รายละเอียดมากที่สุด จำนวน 3 ข้อ คิดเป็นร้อยละ 60.00 รองลงมาเป็นข้อสอบที่มีเนื้อหาด้านการถอดความและด้านการบอกความหมายคำศัพท์จำนวนด้านละ 1 ข้อ คิดเป็นร้อยละ 20.00 สำหรับเนื้อหาอื่น ๆ ไม่มีข้อสอบที่พบ DIF นอกจากนี้พบว่าข้อสอบที่พบ DIF ส่วนใหญ่ร้อยละ 60.00 เป็นข้อสอบที่ค่อนข้างยาก (มีค่าความยากอยู่ในช่วง 0.20 - 0.39) รองลงมาเป็นข้อสอบที่ยากมาก (มีค่าความยากอยู่ในช่วง 0.00-0.19) คิดเป็นร้อยละ 40.00 และพบว่าข้อสอบที่พบ DIF ส่วนใหญ่ร้อยละ 50.00 มีค่าอำนาจจำแนกต่ำมาก (มีค่าอำนาจจำแนกอยู่ในช่วง 0.00- 0.09)

เมื่อพิจารณาการวิเคราะห์ DIF แบบเอกรูป ซึ่งมีข้อสอบที่พบ DIF จำนวน 3 ข้อ นั้น เป็นข้อสอบที่เข้าข้างกลุ่มผู้สอบที่มีประสบการณ์ในการสอบ 2 ข้อ และเข้าข้างกลุ่มผู้สอบที่ไม่มีประสบการณ์สอบอีก 1 ข้อ ส่วนในการวิเคราะห์ DIF แบบบอเนกรูป พบว่า มีข้อสอบที่

DIF จำนวน 3 ข้อ เป็นข้อสอบที่เข้าข้างกลุ่มผู้ไม่มีประสบการณ์ในการสอบที่ระดับความสามารถสูง และความสามารถต่ำอีกกลุ่มละ 1 ข้อ และข้อสอบที่พบ DIF เป็นข้อสอบที่เข้าข้างกลุ่มผู้ไม่มีประสบการณ์ในการสอบที่ระดับความสามารถต่ำ 1 ข้อ

2.2.5 เมื่อพิจารณาลักษณะเนื้อหาของข้อสอบที่พบ DIF ในทุกเงื่อนไขของการวิเคราะห์ พบว่า แบบสอบวิชาอังกฤษมีข้อสอบที่พบ DIF ในข้อสอบที่มีเนื้อหาด้านการให้รายละเอียดมากที่สุด รองลงมาเป็นข้อสอบที่มีเนื้อหาด้านการอ้างอิงและด้านการบอกความหมายคำศัพท์ ส่วนเนื้อหาด้านการจับใจความสำคัญไม่มีข้อสอบที่พบ DIF นอกจากนี้พบว่าข้อสอบที่พบ DIF ส่วนใหญ่เป็นข้อสอบที่ยากมาก (0.00-0.19) และมีค่าอำนาจจำแนกค่อนข้างต่ำ (.20 - 0.39)

3. คุณภาพของแบบสอบวิชาภาษาไทย และภาษาอังกฤษของแบบสอบกรณีก่อนและหลังการตัดข้อสอบที่พบ DIF ในแต่ละกลุ่มผู้สอบออกจากแบบสอบ พบว่า ส่วนใหญ่ค่าความเที่ยงแบบความสอดคล้องภายในของแบบสอบในแต่ละกรณีไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ โดยมีแนวโน้มของการเพิ่มขึ้นเล็กน้อยทั้งในวิชาภาษาไทยและภาษาอังกฤษ สำหรับค่าความตรงตามโครงสร้างหรือความตรงตามทฤษฎี พบว่า กรณีก่อนและหลังการตัดข้อสอบที่พบ DIF ในแต่ละกลุ่มผู้สอบออกจากแบบสอบทั้งวิชาภาษาไทยและภาษาอังกฤษมีค่าความตรงตามทฤษฎีไม่แตกต่างกัน

อภิปรายผลการวิจัย

จากผลการวิจัยครั้งนี้ มีประเด็นสำคัญที่ควรนำมาอภิปรายดังต่อไปนี้

1. การวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบในวิชาภาษาไทยและภาษาอังกฤษ พบข้อสอบที่เกิด DIF ทั้งการวิเคราะห์ DIF แบบเอกรูปและแบบอนเอกรูป แต่ข้อสอบที่พบ DIF ส่วนใหญ่เป็นแบบอนเอกรูปมากกว่าแบบเอกรูป ทั้งนี้อาจเป็นเพราะ การวิเคราะห์แบบอนเอกรูปพิจารณาทั้งระดับความสามารถของผู้สอบ (กลุ่มอ้างอิง ; R และกลุ่มเปรียบเทียบ ; F) และโอกาสที่แต่ละกลุ่มจะตอบข้อสอบได้ถูก ซึ่งพบว่ามีปฏิสัมพันธ์ระหว่างกลุ่มย่อย (Subgroup) กับระดับความสามารถของผู้สอบ ส่วนการวิเคราะห์แบบเอกรูป พบข้อสอบที่เข้าข้างกลุ่มใดกลุ่มหนึ่งโดยเฉพาะทุกระดับความสามารถ ซึ่งในสภาพปกติ ข้อสอบที่พบ DIF โดยเข้าข้างกลุ่มใดกลุ่มหนึ่งทุกระดับความสามารถของผู้สอบจะมีความเป็นไปได้ยากกว่า จึงทำให้การวิเคราะห์ DIF แบบอนเอกรูปพบข้อสอบที่ DIF มากกว่าการวิเคราะห์แบบเอกรูป

2. ผลการระบุลักษณะของข้อสอบที่ทำหน้าที่ต่างกันของข้อสอบระหว่างกลุ่มผู้สอบ เมื่อจำแนกตามตัวแปรเพศ ภูมิภาค ประสิทธิภาพในการสอบ และสังกัดของสถานศึกษา พบว่ามีความแตกต่างกัน โดยเมื่อจำแนกผู้สอบตามเพศพบว่ามีข้อสอบที่มีการระบุว่าจะเกิด DIF มากที่สุด รองลงมาคือ ภูมิภาคของสถานศึกษา ตามลำดับ ส่วนตัวแปรประสิทธิภาพในการสอบมีข้อสอบที่พบ DIF น้อยที่สุด ซึ่งผลการวิจัยนี้สอดคล้องกันทั้งวิชาภาษาไทย และภาษาอังกฤษ ทั้งนี้อาจเป็นเพราะเพศ เป็นตัวแปรที่มีความสัมพันธ์ค่อนข้างสูงกับผลสัมฤทธิ์ทางการเรียนภาษา และลักษณะภาษาที่ใช้ในข้อสอบ มีปฏิสัมพันธ์กับเพศ ทำให้ผู้สอบที่มีระดับความสามารถเท่ากันแต่มีเพศต่างกัน มีโอกาสในการตอบข้อสอบได้ถูกต้องต่างกัน

ส่วนตัวแปรภูมิภาค มีความสำคัญในการระบุการทำหน้าที่ต่างกันของข้อสอบรองลงมาจากเมื่อจำแนกกลุ่มผู้สอบตามเพศ อาจเนื่องจากสภาพสังคม (เขตกรุงเทพมหานครและปริมณฑล กับนอกเขตกรุงเทพมหานครและปริมณฑล) เป็นสภาพแวดล้อมที่มีความสำคัญต่อการเรียนรู้ทางภาษาทั้งภาษาไทยและภาษาอังกฤษ

สำหรับสังกัดของสถานศึกษา และประสิทธิภาพในการสอบเป็นลักษณะทางกายภาพและคุณลักษณะแฝงซึ่งเป็นสภาพแวดล้อมของผู้สอบแต่ละคนที่อาจเปลี่ยนแปลง หรือพัฒนาได้อย่างอิสระ ทำให้มีความสัมพันธ์กับผลสัมฤทธิ์ทางการเรียนภาษาค่อนข้างต่ำ ตัวแปรสังกัดของสถานศึกษา และประสิทธิภาพในการสอบจึงเป็นตัวแปรที่มีผลต่อความสามารถทางภาษาทั้งภาษาไทยและภาษาอังกฤษ น้อยกว่าตัวแปรเพศและภูมิภาค ดังนั้นเมื่อวิเคราะห์ DIF จึงทำให้มีข้อสอบที่พบ DIF น้อยกว่า นอกจากนี้ยังพบว่าผลการวิเคราะห์ DIF แบบเอกรูป และแบบอนุกรุปในแต่ละตัวแปรทั้งแบบสอบวิชาภาษาไทยและวิชาภาษาอังกฤษ พบการระบุลักษณะของข้อสอบที่เกิด DIF ไม่ตรงกัน ซึ่งผลการวิจัยนี้สอดคล้องกับการศึกษาของ Raju และคณะ (1993) พบว่าการจำแนกกลุ่มผู้สอบระหว่างกลุ่มเพศชายและเพศหญิง กับกลุ่มผิวดำและผิวขาว พบข้อสอบที่ถูกระบุว่าจะเกิด DIF ไม่ซ้ำกัน

3. ลักษณะของข้อสอบที่พบ DIF เมื่อจำแนกผู้สอบตามเพศ ในแบบสอบวิชาภาษาไทยและวิชาภาษาอังกฤษ พบว่า ส่วนใหญ่เข้าข้างเพศหญิงมากกว่าเพศชาย ซึ่งสอดคล้องกับผลการศึกษาของ ชัยชัย เผ่าพงษ์ (2527) สุรศักดิ์ อมรรัตนศักดิ์ (2531) พัชรี ปิยภักดิ์ (2531) สุพัฒน์ สุขมลสันต์ (2534) และกาญจนา วัชรสุนทร (2537) ที่พบว่า ข้อสอบวัดความสามารถทางภาษาที่สำเียงส่วนมากเข้าข้างผู้สอบเพศหญิง และสอดคล้องกับคำกล่าวของไมเออร์ (Maier, N.R.F. and Casselman, G.G. ; 1970 อ้างถึงใน Peter, H. and Darwin, P.H. ; 1994) ที่ว่า เพศชายจะมีความสามารถในการแก้ปัญหาทางคณิตศาสตร์ ส่วนเพศหญิงจะมีความสามารถในทางภาษา และจากการศึกษาของ เฮอร์ล็อก (Hurlock, 1978 อ้างถึงใน ประภาศรี สุขนธชาติ, 2534) พบว่า เพศหญิงจะมีการพัฒนาทางภาษา และเรียนรู้เกี่ยวกับภาษาได้ดีกว่าเพศชายในทุกระดับอายุ

4. ลักษณะของข้อสอบที่พบ DIF เมื่อจำแนกผู้สอบตามภูมิสำเนา ในแบบสอบวิชาภาษาไทยและวิชาภาษาอังกฤษพบว่าส่วนใหญ่เข้าข้างผู้สอบกลุ่มผู้สอบในเขตกรุงเทพมหานครและปริมณฑลมากกว่ากลุ่มผู้สอบนอกเขตกรุงเทพมหานครและปริมณฑล ทั้งนี้อาจเป็นเพราะในเขตกรุงเทพมหานครและปริมณฑลมีแหล่งในการศึกษาค้นคว้าเป็นจำนวนมากและทันสมัย ตลอดจนเป็นลักษณะของสภาพสังคมที่มีการแข่งขันกันทั้งในการดำรงชีพ และการติดต่อสื่อสาร เพื่อรับข่าวสารและความรู้ต่างๆ ทำให้ผู้สอบในเขตกรุงเทพมหานครและปริมณฑล มีโอกาสพบสื่อวัสดุต่างๆ ทางภาษามากกว่า ดังนั้นอาจทำให้ผู้สอบกลุ่มนี้มีความคุ้นเคยกับคำศัพท์หรือมีทักษะการอ่านมากกว่า จึงอาจเป็นสาเหตุให้ผู้สอบที่อยู่ในเขตกรุงเทพมหานครและปริมณฑลมีโอกาสในการตอบข้อสอบได้ถูกต้องมากกว่าผู้สอบที่อยู่ในกลุ่มนอกเขตกรุงเทพมหานครและปริมณฑล ซึ่งผลการวิจัยนี้สอดคล้องกับการศึกษาของ ทศนีย์ พิรมนตรี (2530) ที่พบว่าข้อสอบที่ DIF ส่วนใหญ่เข้าข้างกลุ่มผู้สอบในกรุงเทพมหานคร

5. ลักษณะของข้อสอบที่พบ DIF เมื่อจำแนกผู้สอบตามสังกัดของสถานศึกษาพบว่า ในแบบสอบวิชาภาษาไทยข้อสอบที่พบ DIF ส่วนใหญ่เข้าข้างกลุ่มผู้สอบที่สำเร็จการศึกษาในสังกัดทบวงมหาวิทยาลัย ทั้งนี้อาจเป็นเพราะผู้สำเร็จการศึกษาในสังกัดทบวงมหาวิทยาลัยส่วนใหญ่ เป็นนักศึกษาที่มาจากต่างจังหวัด หรือชุมชนต่างๆ ทั่วประเทศ ทำให้มีโอกาสแลกเปลี่ยนประสบการณ์ทางภาษามากกว่าผู้ที่สำเร็จการศึกษานอกสังกัดทบวงมหาวิทยาลัย ซึ่งส่วนใหญ่เป็นนักศึกษาจากท้องถิ่น หรือชุมชนนั้นๆ จึงอาจเป็นผลให้ผู้สำเร็จการศึกษาในสังกัดทบวงมหาวิทยาลัย มีโอกาสในการตอบข้อสอบได้ถูกต้องมากกว่า ทั้งๆ ที่มีความสามารถเท่ากับผู้สำเร็จการศึกษาจากนอกสังกัดทบวงมหาวิทยาลัย แต่ในแบบสอบวิชาภาษาอังกฤษมีผลการวิจัยขัดแย้งกัน กล่าวคือ ข้อสอบที่พบ DIF ส่วนใหญ่เข้าข้างกลุ่มผู้สอบที่สำเร็จการศึกษานอกสังกัดทบวงมหาวิทยาลัย ทั้งนี้อาจเป็นเพราะลักษณะเนื้อหาวิชา มีปฏิสัมพันธ์กับผู้สอบและตัวแปรอื่นๆ จึงทำให้ผู้สอบมีโอกาสในการตอบข้อสอบได้ถูกต้องไม่เท่ากัน ทั้งๆ ที่พิจารณาเปรียบเทียบระหว่างกลุ่มที่มีความสามารถเท่ากัน ดังนั้นเมื่อพิจารณาโดยรวมจากแบบสอบวัดความสามารถทางภาษา พบว่ายังไม่สามารถสรุปได้แน่ชัดว่าแบบสอบวัดความสามารถทางภาษา (ภาษาไทยและภาษาอังกฤษ) ที่มีข้อสอบที่พบ DIF เมื่อจำแนกผู้สอบตามสังกัดของสถานศึกษาจะเป็นข้อสอบที่เข้าข้างกลุ่มใดกลุ่มหนึ่งโดยเฉพาะ

6. ลักษณะของข้อสอบที่พบ DIF เมื่อจำแนกผู้สอบตามประสบการณ์ในการสอบพบว่าทั้งแบบสอบวิชาภาษาไทยและภาษาอังกฤษมีข้อสอบที่พบ DIF เป็นจำนวนน้อย แสดงว่าประสบการณ์ในการสอบไม่ค่อยมีส่วนต่อการทำข้อสอบในคราวต่อไป อาจเป็นเพราะมาตรฐานของข้อสอบ การดำเนินการสอบ รวมทั้งตัวแปรประสบการณ์ในการสอบเป็นคุณลักษณะที่

ผู้สอบมีความคุ้นเคยกับการสอบทั่วไปอยู่แล้วจึงสามารถปรับตัวได้ ผลการสอบจึงอาจขึ้นอยู่กับ การฝึกฝนและการเตรียมตัวสอบของผู้สอบแต่ละคนเป็นสำคัญมากกว่าประสิทธิภาพในการสอบที่ผ่านมา ตลอดจนผู้สอบอาจผ่านการสอบจากสนามสอบอื่นๆ มาหลายครั้งแล้วผู้สอบทุกคนคงมีการเตรียมตัวสอบอย่างดี จึงทำให้ประสิทธิภาพในการสอบครั้งนี้ไม่ก่อให้เกิดความแตกต่างระหว่างกลุ่มผู้สอบมากนัก นอกจากนี้อาจเป็นเพราะในการสอบครั้งนี้มีผู้สอบที่มีประสิทธิภาพในการสอบจำนวนน้อยคือมีผู้เคยสอบในวิชาภาษาไทยร้อยละ 15.42 และวิชาภาษาอังกฤษเพียงร้อยละ 14.77 ของจำนวนผู้สอบทั้งหมด ซึ่งไม่ถึงกลุ่มละ 100 คน เมื่อวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบในครั้งนี้นี้จึงยังไม่อาจสรุปได้แน่ชัดนัก ลักษณะเช่นนี้ สอดคล้องกับที่ Hills (1989) ได้เสนอแนะไว้ว่าในการวิเคราะห์ DIF ด้วยวิธี MH นั้นจำนวนผู้สอบในกลุ่มอ้างอิง (R) หรือกลุ่มเปรียบเทียบ (F) กลุ่มใดกลุ่มหนึ่งที่เป็นกลุ่มขนาดเล็กควรมีผู้สอบอย่างน้อย 100 คน ดังนั้นตัวแปรประสิทธิภาพในการสอบ จึงน่าจะได้มีการศึกษาให้แน่ชัดต่อไป

7. ผลการวิเคราะห์ค่าความยาก และค่าอำนาจจำแนกของข้อสอบที่พบ DIF ในวิชาภาษาไทย พบว่าข้อสอบที่พบ DIF ส่วนใหญ่เป็นข้อสอบที่ค่อนข้างง่าย ซึ่งมีค่าความยากอยู่ในช่วง 0.80-1.00 และส่วนใหญ่เป็นข้อสอบที่มีค่าอำนาจจำแนกค่อนข้างต่ำ ซึ่งมีค่าอำนาจจำแนกอยู่ในช่วง 0.20-0.39 แสดงให้เห็นว่าค่าความยาก และค่าอำนาจจำแนกของข้อสอบที่พบ DIF มีความสัมพันธ์กัน โดยข้อสอบที่ง่ายเกินไปผู้สอบส่วนใหญ่ตอบถูก จึงทำให้ข้อสอบมีค่าอำนาจจำแนกต่ำ (ศิริชัย กาญจนวาสี, 2535 : สุพัฒน์ สุกมลสันต์, 2538)

ส่วนในวิชาภาษาอังกฤษพบว่าข้อสอบที่พบ DIF ส่วนใหญ่เป็นข้อสอบที่ยากโดยมีค่าความยากอยู่ในช่วง 0.00-0.19 และส่วนใหญ่มีค่าอำนาจจำแนกค่อนข้างต่ำ โดยมีค่าอำนาจจำแนกอยู่ในช่วง 0.20-0.39 ดังนั้นจึงพบว่าข้อสอบที่เกิด DIF มีความสัมพันธ์กับค่าความยากและค่าอำนาจจำแนก โดยข้อสอบที่พบ DIF ส่วนใหญ่เป็นข้อสอบที่ยาก และเป็นข้อสอบที่มีค่าอำนาจจำแนกต่ำหรือค่อนข้างต่ำซึ่งอาจเป็นเพราะข้อสอบที่มีค่าความยากสูงเกินไปนั้นไม่สามารถจำแนกผู้สอบที่อยู่ในกลุ่มเก่งและกลุ่มอ่อนได้ เนื่องจากข้อสอบที่ยากเกินไปผู้สอบส่วนใหญ่ตอบผิดหรืออาจตอบถูกเนื่องจากเกิดการเดา จึงทำให้ข้อสอบมีค่าอำนาจจำแนกต่ำ และทำให้เป็นข้อสอบที่ทำหน้าที่ต่างกันระหว่างกลุ่มผู้สอบอีกด้วย ซึ่งสอดคล้องกับที่ศิริชัย กาญจนวาสี (2535) กล่าวว่าข้อสอบที่ดีควรมีค่าความยากปานกลาง (0.41-0.60) และควรมีค่าอำนาจจำแนกสูง และสอดคล้องกับผลการวิจัยของ Sudweeks and Tolman (1990) ที่พบว่า ข้อสอบที่ DIF ส่วนใหญ่จะเป็นข้อสอบที่ยากหรือมีค่าความยากต่ำ นอกจากนี้ผลการวิจัยยังสอดคล้องกับการศึกษาของ Clauser และคณะ (1991) และ Zwick, R. (1994) ที่พบว่าระดับความยากและค่าอำนาจจำแนกมีผลต่อการตรวจค้นการทำหน้าที่ต่างกันของข้อสอบ

8. ผลการวิเคราะห์ค่าความเที่ยง และค่าความตรงเชิงโครงสร้างหรือทฤษฎีของแบบสอบวิชาภาษาไทยและภาษาอังกฤษ ฉบับก่อนและหลังการตัดข้อสอบที่พบ DIF ในแต่ละตัวแปรออกจากแบบสอบ พบว่า ส่วนใหญ่ค่าความเที่ยงของแบบสอบไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ แต่มีแนวโน้มของการเพิ่มขึ้นเล็กน้อยทั้งในวิชาภาษาไทยและภาษาอังกฤษ เช่นเดียวกับการศึกษาของ สุรศักดิ์ อมรรัตนศักดิ์ (2531) และ พัชรีย์ ปิยภัณฑ (2531) ที่พบว่าค่าความเที่ยงของแบบสอบฉบับก่อนและหลังการตัดข้อสอบที่สำคัญออกจากแบบสอบมีค่าไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ

ส่วนค่าความตรงเชิงโครงสร้างของแบบสอบทั้งฉบับก่อนและหลังการตัดข้อสอบที่พบ DIF ในแต่ละตัวแปรออกจากแบบสอบ พบว่า ดัชนีค่าความตรงเชิงโครงสร้างของแบบสอบจากการวิเคราะห์องค์ประกอบเชิงยืนยันไม่แตกต่างกัน ทั้งนี้อาจเป็นเพราะจำนวนข้อสอบที่พบ DIF ที่ตัดออกนั้นเป็นข้อสอบที่ไม่ดี และมีข้อสอบที่พบ DIF จำนวนน้อย จึงไม่ทำให้คะแนนรวมในแต่ละเนื้อหาของแบบสอบแตกต่างกัน แต่ถ้าในแบบสอบมีข้อสอบที่พบ DIF เป็นจำนวนมาก การตัดข้อสอบที่พบ DIF ออกจากแบบสอบอาจไม่ใช่แนวทางในการแก้ปัญหา ที่ควรเลือกใช้ เช่นเดียวกับที่สุรศักดิ์ อมรรัตนศักดิ์ (2531) และ O' Neill, A.K. and McPeck, W.M. (1993) ให้ข้อเสนอแนะไว้ว่าเมื่อเกิดกรณีที่พบ DIF ในแบบสอบ สิ่งที่สำคัญยิ่งคือการพิจารณาปรับปรุงแก้ไขข้อกระทงนั้น หรือพยายามให้มีจำนวนข้อสอบที่เข้าข้างผู้สอบกลุ่มใดกลุ่มหนึ่งโดยเฉพาะในจำนวนที่เท่าๆ กัน มากกว่าการที่จะมุ่งตัดข้อสอบที่พบ DIF ออกจากแบบสอบ

ข้อเสนอแนะ

ข้อเสนอแนะในการนำผลการวิจัยไปใช้

1. จากผลการวิจัยพบว่า การวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบในแบบสอบวิชาภาษาไทยและภาษาอังกฤษทั้งการวิเคราะห์แบบเอกรูปและแบบอเนกรูป ส่วนใหญ่มีข้อสอบที่ถูกระบุว่า DIF เมื่อจำแนกกลุ่มผู้สอบตามเพศมากที่สุด รองลงมาได้แก่สภาพสังคม ดังนั้นศูนย์ทดสอบทางการศึกษาและผู้สร้างแบบสอบทั่วไป น่าจะได้มีการศึกษาวิเคราะห์ถึงสาเหตุ เช่น ลักษณะของภาษา คำที่ใช้ เพื่อจะได้ปรับปรุงหรือหลีกเลี่ยงสิ่งที่จะนำไปสู่การเกิด DIF เพื่อให้ได้แบบสอบที่มีความยุติธรรมต่อผู้สอบ

2. ผลการวิเคราะห์ลักษณะของข้อสอบที่พบ DIF พบว่าข้อสอบที่พบ DIF ส่วนใหญ่เป็นข้อสอบที่ง่ายหรือยากมาก และเป็นข้อสอบที่มีค่าอำนาจจำแนกค่อนข้างต่ำ นอกจากนี้ควรพิจารณาถึงลักษณะเนื้อหาของรายวิชาอีกด้วย เนื่องจากในวิชาภาษาไทยข้อสอบที่พบ DIF ส่วนใหญ่เป็นข้อสอบที่ง่ายมีค่าอำนาจจำแนกค่อนข้างต่ำ และการสร้างแบบสอบควรระวังเนื้อหาเกี่ยวกับด้านการใช้ภาษาด้วย เพื่อให้มีความยุติธรรมต่อผู้สอบมากที่สุด ส่วนในวิชา

ภาษาอังกฤษข้อสอบที่พบ DIF ส่วนใหญ่เป็นข้อสอบที่ยาก มีค่าอำนาจจำแนกค่อนข้างต่ำ และส่วนใหญ่มีเนื้อหาเกี่ยวกับด้านการให้รายละเอียด ดังนั้นผู้พัฒนาแบบสอบจึงไม่ควรเลือกใช้ข้อสอบที่ง่ายหรือยากมาก แต่ควรใช้ข้อสอบที่มีค่าอำนาจจำแนกสูงเพื่อให้ข้อสอบหรือแบบสอบมีคุณภาพเชื่อถือได้ มีความยุติธรรม และมีความคลาดเคลื่อนในการวัดน้อยที่สุด

3. จากผลการวิจัยพบว่า ค่าความเที่ยงและค่าความตรงตามทฤษฎีของแบบสอบกรณีก่อนและหลังการตัดข้อสอบที่พบ DIF ออกจากแบบสอบไม่แตกต่างกัน ดังนั้นการตัดข้อสอบที่พบ DIF ออก จะทำให้ลดขนาดความยาวของแบบสอบซึ่งจะช่วยให้ประหยัดรายจ่ายได้จำนวนหนึ่งในขณะที่แบบสอบยังคงมีประสิทธิภาพเช่นเดิม นอกจากนี้หากได้มีการปรับปรุงข้อสอบภายหลังจากที่พบ DIF เป็นรายข้อ เช่นปรับปรุงภาษาหรือการใช้คำที่เหมาะสม น่าจะทำให้ได้แบบสอบที่มีประสิทธิภาพมากขึ้น

ข้อเสนอแนะในการวิจัยครั้งต่อไป

1. ควรได้มีการศึกษาวิเคราะห์ถึงสาเหตุของการเกิด DIF เช่น ลักษณะของภาษาคำที่ใช้ เพื่อจะได้ปรับปรุงหรือหลีกเลี่ยงสิ่งที่จะนำไปสู่การเกิด DIF เพื่อให้ได้แบบสอบที่สามารถตัดสินผู้สอบได้อย่างถูกต้องและมีความยุติธรรมยิ่งขึ้น

2. เนื่องจากการวิจัยนี้ ผู้วิจัยวิเคราะห์ค่าความตรงเชิงโครงสร้างของแบบสอบตามลักษณะของเนื้อหาเป็นรายด้าน จึงมิได้จำแนกความแตกต่างเกี่ยวกับคุณสมบัติเป็นรายข้อมากนัก สำหรับการวิจัยครั้งต่อไปถ้าสามารถคัดเลือกแบบสอบที่มีข้อสอบในจำนวนที่ใช้กับโปรแกรม LISREL เป็นรายข้อได้ ควรมีการวิเคราะห์องค์ประกอบเชิงยืนยันโดยพิจารณาเป็นรายข้อ ซึ่งจะทำให้ได้สารสนเทศที่เป็นประโยชน์และได้ข้อสรุปเกี่ยวกับการตัดสินใจปรับปรุงข้อสอบหรือความเหมาะสมของการตัดข้อสอบที่พบ DIF ออกจากแบบสอบอันเป็นประโยชน์ต่อการพัฒนาแบบสอบมากขึ้น

3. ควรได้มีการศึกษาปฏิสัมพันธ์ระหว่างตัวแปรต่างๆ เช่น เพศ กับ ภูมิฐานะหรือตัวแปรอื่นๆ เพื่อศึกษาว่าจะมีผลต่อการระบุลักษณะของข้อสอบที่พบ DIF หรือไม่

4. ควรได้มีการศึกษาการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบในข้อสอบอัตนัยหรือข้อสอบที่มีการให้คะแนนแบบหลายค่า (Polytomously scored Item) นอกเหนือไปจากการวิเคราะห์ข้อสอบที่ให้คะแนนแบบสองค่า (Dichotomously scored Item) ที่ใช้ในการวิจัยนี้

5. ในปัจจุบันวิธีการวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ ได้รับการพัฒนาขึ้นมาก น่าจะได้มีการศึกษาเปรียบเทียบผลการวิเคราะห์การระบุลักษณะข้อสอบที่พบ DIF หรือหาความสัมพันธ์ของวิธีการต่างๆ เช่น วิธี SIBTEST วิธี IRT วิธี Logistic Regression หรือวิธีอื่นๆ กับ วิธี MH เพื่อเปรียบเทียบกับผลที่ได้จากการวิจัยครั้งนี้

6. ควรได้มีการศึกษาวิเคราะห์การทำหน้าที่ต่างกันของแบบสอบ (Differential Test Functioning : DTF) ต่อไป