



หน้า

หน้า

ความเป็นมาและความสำคัญของปัญหา

ความลำเอียงหรือความไม่ยุติธรรมของแบบสอบเป็นเรื่องที่ได้รับการวิพากษ์วิจารณ์และมีผู้ให้ความสนใจกันมาก สาเหตุที่เป็นเช่นนี้คงเป็นเพราะแบบสอบเป็นเครื่องมือสำคัญในการวัดการนำเครื่องมือที่มีความลำเอียงในการวัดมาใช้จึงไม่อาจแน่ใจได้ในเรื่องความถูกต้องของผลการวัด ข้อวิพากษ์วิจารณ์ดังกล่าวมักเกี่ยวข้องกับเรื่องความได้เปรียบเสียเปรียบของกลุ่มผู้สอบที่มีลักษณะบางประการแตกต่างกันในเรื่องเพศ เชื้อชาติ ศาสนา หรือภาษาถิ่น สาเหตุที่ทำให้มีการได้เปรียบเสียเปรียบระหว่างผู้สอบต่างกลุ่มกันอาจเนื่องมาจากสาเหตุหลายประการ แต่สาเหตุประการหนึ่งน่าจะเนื่องมาจากความลำเอียงของแบบสอบที่ใช้

เกี่ยวกับเรื่องความลำเอียงนี้ได้มีผู้ให้ความหมายกันมากมาย ดังเช่น

Lord (1980:212) ได้ให้ความหมายของข้อสอบที่มีความลำเอียงว่าเป็นข้อสอบที่ผู้สอบที่มีความสามารถเท่ากันแต่อยู่ต่างกลุ่มกันมีโอกาสที่จะตอบถูกไม่เท่ากัน

Popham (1981:181) ได้กล่าวว่าความลำเอียงหมายถึงความชอบหรือความเอนเอียงที่ทำให้การพิจารณาตัดสินเป็นไปอย่างไม่ยุติธรรม

Allen and Yen (1979:129) ให้ความหมายของข้อสอบที่มีความลำเอียงว่าหมายถึงข้อสอบที่วัดความสามารถหรือคุณลักษณะทางจิตวิทยาของผู้สอบแต่ละกลุ่มได้ไม่ตรงกัน เช่นเมื่อนำ

ข้อสอบวัดความสามารถด้านจิตภาษาสรรที่ใช้ภาษาสื่อความหมายของโจทช์ปัญหา ไปทดสอบกับผู้สอบ 2 กลุ่ม ที่มีความสามารถทางการอ่านไม่เท่ากัน ข้อสอบที่จะมีความลำเอียงต่อผู้สอบกลุ่มที่มีความสามารถทางการอ่านสูง ข้อสอบเหล่านี้จะมีโค้งลักษณะข้อสอบของผู้สอบแต่ละกลุ่มแตกต่างกัน

Cleary and Hilton (1968: 61) ได้ให้คำจำกัดความของข้อสอบที่มีความลำเอียงว่าหมายถึงข้อสอบที่เมื่อนำไปทดสอบกับบุคคลกลุ่มใดกลุ่มหนึ่งแล้วปรากฏว่า มีคะแนนเฉลี่ยของข้อสอบเดียวกันจากแต่ละกลุ่มไม่เท่ากัน

การนำแบบสอบที่มีความลำเอียงทางการวัดไปใช้จึงเป็นการไม่ยุติธรรมแก่ผู้สอบและผลที่ได้จากการวัดก็ไม่สามารถสะท้อนถึงความสามารถที่แท้จริงของผู้สอบได้ ทั้งนี้เพราะแต่ละกลุ่มมีการได้เปรียบเสียเปรียบกันในการตอบข้อสอบข้อเดียวกัน ตัวอย่างเช่น กลุ่มเพศชายเคยผ่านหรือเรียนรู้เนื้อหาที่ถูกทดสอบมากกว่ากลุ่มเพศหญิงแล้วจะทำให้กลุ่มเพศชายมีโอกาสในการตอบข้อสอบข้อนั้นมากกว่าเพศหญิง ในกรณีดังกล่าวอาจเป็นไปได้ว่ากลุ่มเพศชายตอบคำถามโดยใช้ความสามารถทางสมองขั้นความรู้ความจำ ในขณะที่เพศหญิงจะตอบคำถามข้อนั้นได้ต้องใช้ความสามารถทางสมองสูงกว่าขั้นความรู้ความจำ หรืออีกตัวอย่างหนึ่งได้แก่ การนำเอาภาษาถิ่นหรือคำศัพท์ทางด้านวิชาการที่ทราบกันอยู่ในวงจำกัดมาใช้ ข้อสอบข้อนั้นจึงอาจเป็นข้อสอบที่ยากสำหรับกลุ่มหนึ่งและง่ายสำหรับอีกกลุ่มหนึ่ง ซึ่งทำให้ข้อสอบเกิดความลำเอียงในผลการวัดได้

จากตัวอย่างที่ยกมาจะเห็นได้ว่า ถ้าข้อสอบมีความลำเอียงกับกลุ่มผู้สอบบางกลุ่มที่มีลักษณะแตกต่างกันทางด้านเพศ เชื้อชาติ ศาสนา ภาษาถิ่นหรือสถานแวดล้อมทางภูมิศาสตร์แล้วจะทำให้มีการได้เปรียบเสียเปรียบเกี่ยวกับเนื้อหาของข้อสอบ

เมื่อไม่นานมานี้ได้มีผู้ให้ความสนใจเกี่ยวกับวิธีการหาความลำเอียงของข้อสอบกันอย่างกว้างขวาง และได้มีการเสนอแนะวิธีการหาความลำเอียงของข้อสอบเป็นจำนวนมาก ซึ่งวิธีการ

ในการวิเคราะห์ความลำเอียงของข้อสอบเท่าที่ใช้กันอยู่หากจะแบ่งเป็นประเภทใหญ่ ๆ อาจแบ่งได้เป็น 2 ประเภทคือ การวิเคราะห์ความลำเอียงโดยยึดทฤษฎีการวัดแบบคลาสสิก (Classical Test Theory) และการวิเคราะห์ความลำเอียงโดยยึดทฤษฎีการตอบสนองต่อข้อสอบ (Item Response Theory)

วิธีการวิเคราะห์ความลำเอียงของข้อสอบโดยอาศัยทฤษฎีการวัดแบบคลาสสิก เท่าที่มีผู้เคยทำการศึกษากันไว้ได้แก่ การวิเคราะห์ความแปรปรวน (Analysis of Variance) การแปลงค่าความยาก (Transformed Item Difficulty) การวิเคราะห์องค์ประกอบ (Factor Analysis) และการวิเคราะห์ด้วยไคสแควร์ (Chi - square) วิธีการวิเคราะห์ความลำเอียงของข้อสอบแต่ละวิธีที่ได้กล่าวมา ต่างนิยามความลำเอียงของข้อสอบแตกต่างกัน ดังจะเห็นได้ต่อไปนี้

การวิเคราะห์ความแปรปรวน นิยามความลำเอียงของข้อสอบว่าถ้า Interaction Effect ระหว่างข้อ (Items) กับกลุ่มผู้สอบ (Subgroups) มีนัยสำคัญทางสถิติ (Significant) แสดงว่าข้อสอบลำเอียง (Laksana 1979:12)

การแปลงค่าความยากของข้อสอบ นิยามความลำเอียงข้อสอบดังนี้ ข้อสอบที่เบี่ยงเบนไปจากเส้นแกนหลัก (Principal Axis) มาก แสดงว่าข้อสอบข้อนี้มีความลำเอียง (Angoff cited in Berk 1982:97)

การวิเคราะห์องค์ประกอบ นิยามข้อสอบที่ไม่ได้วัดสิ่งเดียวกันในผู้สอบต่างกลุ่มกันว่า เป็นข้อสอบที่มีความลำเอียง (Laksana and Coffman 1980:2)

การวิเคราะห์ด้วยไคสแควร์ นิยามข้อสอบที่ไม่ลำเอียงต่อกลุ่มผู้สอบว่า หมายถึงข้อสอบที่ผู้สอบที่มีความสามารถเท่ากันมีโอกาสตอบถูกเหมือนกันโดยไม่คำนึงว่า จะเป็นกลุ่มที่มีเชื้อชาติแตกต่างกันหรือไม่ (Scheuneman 1975 cited in Intasuwan 1979 : 26)

การวิเคราะห์ข้อสอบโดยอาศัยทฤษฎีการวัดแบบคลาสสิกมีจุดอ่อนที่ควรพิจารณาก็คือ ค่าสถิติต่าง ๆ ของข้อสอบจะแปรเปลี่ยนไปตามกลุ่มตัวอย่างที่มีระดับความสามารถแตกต่างกัน

จะเห็นได้ว่าทฤษฎีการวัดแบบคลาสสิกไม่สามารถแก้ปัญหาเกี่ยวกับการแปรเปลี่ยนของค่าสถิติได้ นักวัดจึงได้ช่วยกันคิดและสร้างทฤษฎีขึ้นมาใหม่เพื่อแก้ปัญหาดังกล่าว ทฤษฎีหนึ่งที่นักวัดช่วยกันพัฒนาขึ้นมาคือ ทฤษฎีการตอบสนองต่อข้อสอบ (Item Response Theory) ผู้ที่ทำการบุกเบิกทฤษฎีนี้กันอย่างจริงจังก็คือ Lord, Birnbaum และ Rasch ทฤษฎีการตอบสนองต่อข้อสอบนี้ บางทีก็เรียกว่าทฤษฎีความสามารถแฝง (Latent Trait Theory) หรือทฤษฎีโด่งลักษณะข้อสอบ (Item Characteristic Theory)

ทฤษฎีการตอบสนองต่อข้อสอบเป็นทฤษฎีที่เข้ามามีบทบาทในการวัดทางการศึกษา และจิตวิทยาในราวปี ค.ศ. 1950 - 1960 แต่ทฤษฎีนี้ไม่ได้รับความสนใจจากนักวัดทางการศึกษา และจิตวิทยาเท่าที่ควร สาเหตุที่เป็นเช่นนี้อาจเนื่องมาจากความยุ่งยากซับซ้อนในการคำนวณ แต่ในปัจจุบันนักวัดทั้งหลายต่างก็หันมาสนใจทฤษฎีนี้กันมากขึ้น ทั้งนี้อาจเป็นเพราะปัจจุบันมีโปรแกรมคอมพิวเตอร์สำหรับการวิเคราะห์ข้อสอบตามทฤษฎีนี้ จึงทำให้ลดความยุ่งยากลงได้มาก และทฤษฎีนี้ยังสามารถแก้จุดอ่อนของทฤษฎีการวัดแบบคลาสสิกได้ ตลอดจนสามารถนำไปประยุกต์เพื่อแก้ปัญหาทางการวัดซึ่งทฤษฎีการวัดแบบคลาสสิกไม่สามารถจะแก้ได้ (Lord 1977 : 117 - 138)

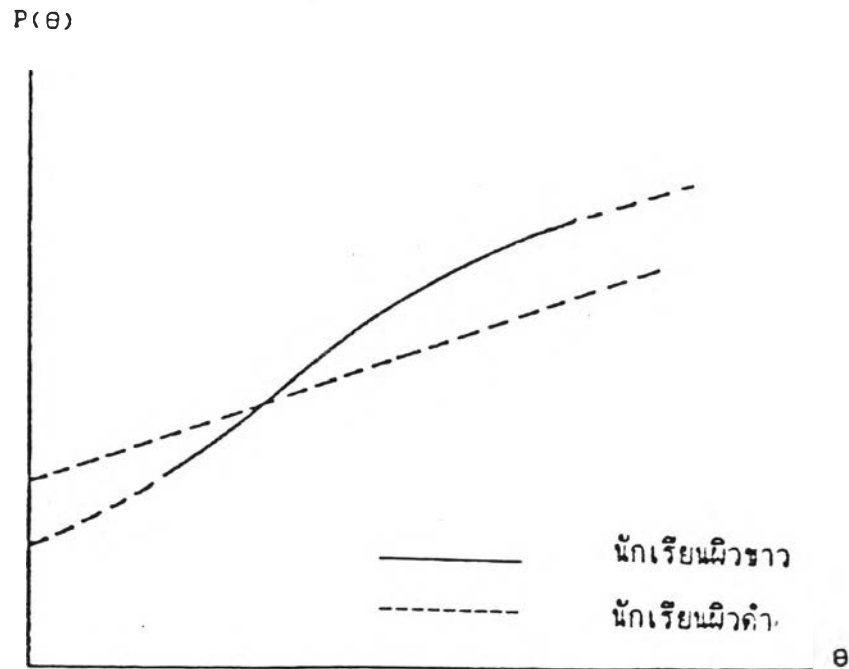
ทฤษฎีการตอบสนองต่อข้อสอบมีความเชื่อว่า พฤติกรรมที่บุคคลตอบสนองต่อข้อสอบนั้น ถูกกำหนดโดยลักษณะบางอย่างภายในตัวบุคคล (Latent Trait) ทฤษฎีนี้จึงพยายามที่จะอธิบายความสัมพันธ์ของผลการตอบสนองต่อข้อสอบหรือคะแนน (Performance or Score) กับความสามารถ (Ability or Trait) ดังนั้นทฤษฎีการตอบสนองต่อข้อสอบจึงเป็นทฤษฎีที่อธิบายความสัมพันธ์ดังกล่าวนี้โดยอธิบายออกมาในรูปแบบจำลองทางคณิตศาสตร์ดังนี้

$$P = f(\theta)$$

เมื่อ	P	แทนผลการสอบ (Performance)
	θ	แทนความสามารถ (Ability or Trait)
	f	แทนฟังก์ชัน (Function)

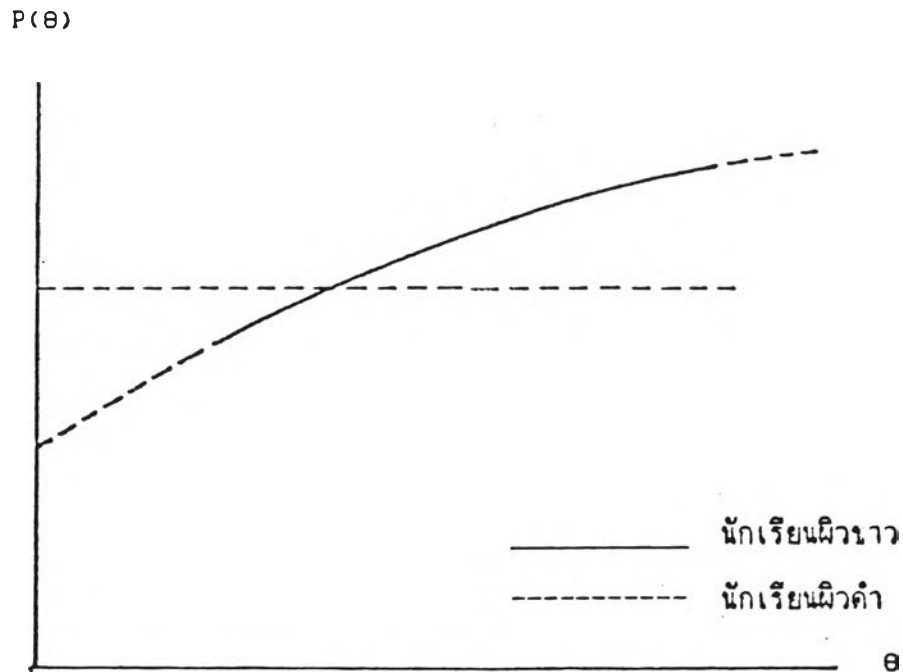
จะเห็นได้ว่าจุดประสงค์ของทฤษฎีการตอบสนองต่อข้อสอบก็คือ การมุ่งสรุปเกี่ยวกับลักษณะภายในตัวบุคคลโดยอาศัยเหตุการณ์การตอบสนองต่อข้อสอบของบุคคลนั้น โดยอาศัยแบบจำลองทางคณิตศาสตร์เป็นตัวเชื่อมที่สำคัญ

ในด้านของทฤษฎีการตอบสนองต่อข้อสอบ (IRT) ได้นิยามข้อสอบที่มีความลำเอียงทางการวัดไว้ว่า หมายถึงข้อสอบที่มีโครงสร้างข้อสอบจากผู้สอบต่างกลุ่มกันแตกต่างกัน เพื่อให้เข้าใจง่ายขึ้น ในที่นี้จะขอเสนอตัวอย่างผลการวิเคราะห์ความลำเอียงของข้อสอบ Verbal Scholastic Aptitude ที่ใช้ทดสอบกับนักเรียนผิวขาวและผิวดำในสหรัฐอเมริกา (Lord 1977 : 136 - 137) จำนวน 3 ข้อ ดังภาพที่ 1 - 3



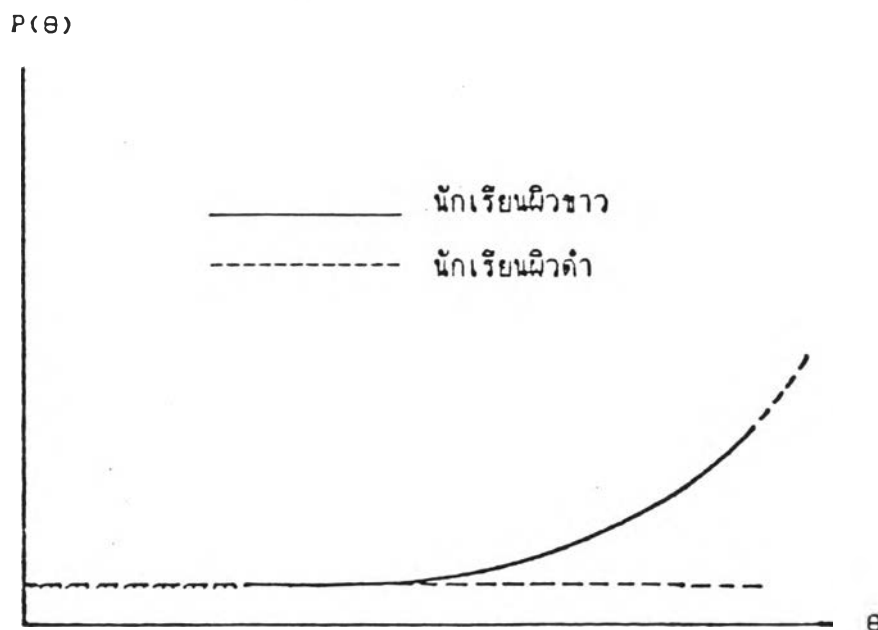
ภาพที่ 1 แสดงโค้งลักษณะข้อสอบข้อ 71 ระหว่างนักเรียนผิวขาวและผิวดำ

จากภาพ แสดงให้เห็นว่านักเรียนผิวขาวที่มีความสามารถสูงมีโอกาที่จะตอบข้อสอบข้อนี้ถูกมากกว่านักเรียนผิวดำที่มีความสามารถสูง แต่สำหรับผู้ที่มีความสามารถต่ำแล้วนักเรียนผิวดำมีโอกาที่จะตอบถูกมากกว่านักเรียนผิวขาว



ภาพที่ 2 แสดงโค้งลักษณะข้อสอบข้อ 2 ระหว่างนักเรียนผิวขาวและผิวดำ

จากภาพ แสดงให้เห็นว่าข้อสอบข้อนี้ไม่สามารถจำแนกนักเรียนผิวดำได้ เพราะไม่ว่านักเรียนผิวดำจะมีความสามารถในระดับใดก็ตามต่างก็มีโอกาสที่จะตอบข้อสอบข้อนี้ได้ถูกเท่า ๆ กัน เฉพาะกลุ่มผู้สอบที่มีความสามารถสูงนักเรียนผิวขาวมีโอกาสที่จะตอบข้อสอบข้อนี้ถูกมากกว่านักเรียนผิวดำ ซึ่งตรงกันข้ามกับในกลุ่มผู้สอบที่มีความสามารถต่ำ นักเรียนผิวดำมีโอกาสดตอบข้อนี้ถูกมากกว่านักเรียนผิวขาว



ภาพที่ 3 แสดงโค้งลักษณะข้อสอบข้อ 24 ระหว่างนักเรียนผิวขาวและผิวดำ

จากภาพ แสดงให้เห็นว่าข้อสอบข้อนี้ยาก (โอกาสในการตอบถูกต่ำมาก) สำหรับนักเรียนผิวดำในทุก ๆ ระดับความสามารถ และยากสำหรับนักเรียนผิวขาวที่มีความสามารถในระดับต่ำ ๆ ด้วย ส่วนนักเรียนผิวขาวที่มีความสามารถสูงมีแนวโน้มที่จะตอบถูกมากขึ้น

จากตัวอย่างข้อสอบ 3 ข้อที่ยกมานี้ แสดงให้เห็นถึงข้อสอบที่คิดว่ามีความลำเอียงปรากฏอยู่ อย่างไรก็ตามค่าดัชนีความลำเอียงของข้อสอบที่ยกมานี้อาจไม่สูงมากนัก (โค้งลักษณะข้อสอบของผู้สอบ 2 กลุ่ม แตกต่างกันไม่มาก)

ทฤษฎีการตอบสนองต่อข้อสอบก็ยังมีข้อจำกัดและจุดด้อยบางประการ เช่น การคำนวณก่อนข้างจะยุ่งยากซับซ้อน ต้องใช้กลุ่มตัวอย่างขนาดใหญ่และค่าใช้จ่ายในการคำนวณก่อนข้างแพง

จะเห็นได้ว่าวิธีการวิเคราะห์ความลำเอียงของแบบสอบ ไม่ว่าจะยึดทฤษฎีการวัดแบบคลาสสิก หรือยึดทฤษฎีการตอบสนองต่อข้อสอบ ต่างก็มีจุดเด่นและจุดด้อยช่วยกันทั้งสองวิธี ดังนั้นจึงน่าที่จะทำการเปรียบเทียบวิธีการวิเคราะห์ความลำเอียงของข้อสอบที่ยึดทฤษฎีการวัดแบบคลาสสิก และยึดทฤษฎีการตอบสนองต่อข้อสอบว่าให้ผลแตกต่างกันหรือไม่ หรือสอดคล้องกันเพียงไร พร้อมทั้งทำการเปรียบเทียบความแตกต่างของผลการคัดเลือก เมื่อตัดข้อสอบที่มีความลำเอียงออกจากแบบสอบ วิธีการวิเคราะห์ความลำเอียงของแบบสอบ ที่ผู้วิจัยเห็นว่าน่าจะนำมาเปรียบเทียบกันได้แก่วิธีวิเคราะห์ความแปรปรวน (Analysis of Variance : ANOVA) วิธีแปลงค่าความยากของข้อสอบ (Transformed Item Difficulty : TID) วิธีโค้งลักษณะข้อสอบที่มีพารามิเตอร์ 1 ตัว (One - parameter Item Characteristic Curve : ICC 1) และวิธีโค้งลักษณะข้อสอบที่มีพารามิเตอร์ 3 ตัว (Three-parameter Item Characteristic Curve : ICC 3) วิธีการวิเคราะห์ 2 วิธีแรกนี้เป็นการหาความลำเอียงของข้อสอบโดยยึดทฤษฎีการวัดแบบคลาสสิก ส่วน 2 วิธีหลังเป็นการหาความลำเอียงของข้อสอบโดยยึดทฤษฎีการตอบสนองต่อข้อสอบ ซึ่งรายละเอียดของวิธีการวิเคราะห์ความลำเอียงจะกล่าวถึงในบทที่ 2

ส่วนที่เลือกใช้วิธีวิเคราะห์ที่มีพารามิเตอร์ 3 ตัว และที่มีพารามิเตอร์ 1 ตัว ก็เพราะการวิเคราะห์โดยใช้พารามิเตอร์ 1 ตัว มีความสะดวก ง่ายและรวดเร็ว ซึ่งโดยทั่วไปการวิเคราะห์ก็ไม่จำเป็นต้องประมาณค่าพารามิเตอร์ของข้อสอบทุกตัว และหากต้องการให้ผลการวิเคราะห์มีความสมบูรณ์ก็ใช้ วิธีที่มีพารามิเตอร์ 3 ตัว เพราะสามารถวิเคราะห์และอธิบายถึงโอกาสในการตอบข้อสอบของผู้สอบในกลุ่มที่มีความสามารถต่ำมาก ๆ ได้

ส่วนแบบสอบที่จะนำมาวิเคราะห์ความลำเอียงก็คือ แบบสอบแข่งขันเพื่อบรรจุเข้ารับราชการเป็นข้าราชการครู สังกัดสำนักงานคณะกรรมการการประถมศึกษาแห่งชาติปี พ.ศ. 2529 ทั้งนี้เพราะผู้วิจัยเห็นว่าแบบสอบที่ใช้คัดเลือกดังกล่าวนอกจากจะต้องมีคุณภาพแล้ว ยังต้องสามารถใช้ทดสอบบุคคลที่มีความสามารถในระดับเดียวกันแต่อยู่ต่างกลุ่มกัน โดยไม่มีความลำเอียง แต่เท่าที่

ผ่านมาแบบสอบที่ใช้คัดเลือกทั่ว ๆ ไป หาได้มีคุณลักษณะดังที่กล่าวมาแล้วไม่ แม้กระทั่งแบบสอบที่
 ให้นิยามเป็นมาตรฐานแล้วก็ตาม ซึ่งไม่ว่าจะเป็นแบบสอบวัดผลสัมฤทธิ์ หรือแบบสอบวัดความถนัด
 ที่ยึดทฤษฎีการวัดแบบคลาสสิกเป็นหลักในการพัฒนาแบบสอบ ซึ่งมักจะปรากฏว่ามีข้อสอบอยู่จำนวน
 หนึ่งที่มีความลำเอียงกับกลุ่มผู้สอบที่มีลักษณะบางประการที่แตกต่างกัน เช่นแตกต่างกันในเรื่องเพศ
 สาเหตุที่เป็นเช่นนี้อาจเนื่องมาจากในการพัฒนาแบบสอบโดยทั่วๆ ไปนั้นนอกจากจะมุ่งปรับปรุงคุณภาพ
 ของแบบสอบแล้ว ยังมุ่งที่จะพิจารณาว่าข้อสอบแต่ละข้อสามารถวัดในสิ่งที่ต้องการวัดหรือไม่ นั่นคือ
 เป็นการตรวจสอบความตรงเชิงเนื้อหาของข้อสอบนั่นเอง แต่ไม่มีการตรวจสอบเกี่ยวกับโครงสร้าง
 ของพฤติกรรมที่ต้องการวัด กล่าวคือถ้าจำแนกประชากรออกเป็นกลุ่มย่อย ๆ ตามคุณลักษณะบาง
 ประการที่แตกต่างกัน เช่นเพศชายและเพศหญิงแล้วข้อสอบข้อนั้นสามารถวัดพฤติกรรมของผู้สอบทั้ง
 สองกลุ่มนี้ได้ตรงกันหรือไม่ หากวัดพฤติกรรมได้ไม่ตรงกันก็ย่อมแสดงว่าข้อสอบนั้นมีความลำเอียง
 ทางกรวัด ซึ่งถ้าหากนำข้อสอบข้อนี้มาทำการวิเคราะห์ด้วยทฤษฎีการตอบสนองต่อข้อสอบแล้วก็
 จะพบว่าได้งลักษณะข้อสอบระหว่างกลุ่มเพศชายและเพศหญิงแตกต่างกัน

เหตุผลอีกประการหนึ่งที่ผู้วิจัยต้องการวิเคราะห์ความลำเอียงของแบบสอบแข่งขัน
 เพื่อบรรจุเข้ารับราชการเป็นข้าราชการครู ก็เพราะผู้วิจัยเห็นว่าการวิเคราะห์ความลำเอียง
 ของข้อสอบจะช่วยให้ได้ข้อมูลเกี่ยวกับคุณภาพของข้อสอบในอีกแง่มุมหนึ่ง ซึ่งจะช่วยให้ผู้ที่มีหน้าที่
 เกี่ยวข้องกับการคัดเลือกได้นำไปใช้เป็นข้อมูลส่วนหนึ่งในการพิจารณาตัดสินผลการคัดเลือกให้มี
 ความถูกต้องและเป็นธรรมมากยิ่งขึ้น ทั้งนี้เพราะสิ่งที่คณะกรรมการหรือผู้ที่มีหน้าที่เกี่ยวข้องกับการ
 การคัดเลือกจะต้องระลึกถึงอยู่เสมอ ก็คือผลการสอบคัดเลือกเป็นเครื่องชี้และตัดสินอนาคต
 ของผู้สอบ ดังนั้นถ้าเครื่องมือที่ใช้ในการคัดเลือกมีความลำเอียงเกิดขึ้นแล้วก็จะเป็นการไม่
 ยุติธรรมแก่ผู้สอบ ซึ่งผู้ที่มีส่วนเกี่ยวข้องกับเรื่องนี้ น่าจะได้ตระหนักถึงปัญหาดังกล่าวด้วยความ
 รอบคอบ

สำหรับที่ผู้วิจัยเลือกเพศมาเป็นตัวแปรหนึ่งในการศึกษาเกี่ยวกับความลำเอียงของข้อสอบ
 ครั้งนี้ ก็เพราะมีผลการวิจัยจำนวนมากที่พบว่า แบบสอบที่ใช้มีความลำเอียงทางการวัดต่อกลุ่ม

ผู้สอบเพศชายและเพศหญิง ดังเช่นผลการวิจัยของ Coffman (1961 cited in Laksana 1979) ที่นำแบบสอบความถนัดด้านภาษามาศึกษาความลำเอียงของข้อสอบด้วยวิธีการแปลงค่าความยากของข้อสอบ ผลการวิจัยของ Coffman พบว่าความยากของข้อสอบข้อเดียวกันจากกลุ่มเพศชายและเพศหญิงมีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ถึง 9 ข้อ ในจำนวนข้อสอบทั้งหมด 60 ข้อ นอกจากนี้ยังมีผลการวิจัยของ Plake, Hoover และ Loyd (1978 cited in Laksana 1979) ที่สนับสนุนข้อความรู้งี้ดังกล่าว โดย Plake และคณะได้นำแบบสอบวัดความสามารถของผู้สอบชุด คำศัพท์ การอ่าน และมโนทัศน์ทางคณิตศาสตร์จำนวน 3 ฉบับ จากแบบสอบ ITBS มาวิเคราะห์ความลำเอียงโดยวิธีวิเคราะห์ความแปรปรวน ผลการวิจัยปรากฏว่าผลรวม (Interaction) ระหว่างข้อสอบกับกลุ่มผู้สอบชายและหญิงมีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ นั้นย่อมแสดงว่าแบบสอบที่นำมาวิเคราะห์มีความลำเอียงทางการวัดนั่นเอง และในระยะเวลาต่อมา Laksana (1979) ก็ได้นำแบบสอบ ITBS ฉบับดังกล่าวมาวิเคราะห์ความลำเอียง โดยใช้วิธีวิเคราะห์ความแปรปรวนและวิธีโค้งลักษณะข้อสอบ ผลการศึกษาของ Laksana พบว่าแบบสอบดังกล่าวมีความลำเอียงระหว่างเพศชายและเพศหญิง

นอกจากนี้ Title (1973 cited in Chase 1976) ยังได้แสดงให้เห็นว่าข้อสอบส่วนมากมักจะมีสถานการณ์เกี่ยวข้องกับเพศชายมากกว่าเพศหญิง ฉะนั้นเพศหญิงจึงมีการเสียเปรียบเพศชาย

ด้วยเหตุผลดังที่ได้กล่าวมา ผู้วิจัยเห็นว่าน่าจะนำวิธีการทางสถิติมาวิเคราะห์ความลำเอียงของแบบสอบแข่งขันเพื่อบรรจุเข้ารับราชการเป็นข้าราชการครู ระหว่างผู้สอบเพศชายและเพศหญิงด้วยวิธีการวิเคราะห์ความลำเอียง 4 วิธี พร้อมทั้งทำการเปรียบเทียบวิธีวิเคราะห์ความลำเอียงเหล่านั้นว่าให้ผลแตกต่างหรือสอดคล้องกันเพียงไร อันจะเป็นประโยชน์ในการพัฒนาปรับปรุงแก้ไขข้อสอบที่มีความลำเอียงให้สามารถใช้อย่างยุติธรรมต่อไป ตลอดจนเนื้อให้ผลการสอบแข่งขันในครั้งต่อ ๆ ไปได้รับการพิจารณาอย่างรอบคอบเป็นธรรมมากยิ่งขึ้น

วัตถุประสงค์ของการวิจัย

1. เพื่อศึกษาผลของวิธีการวิเคราะห์ความลำเอียงของข้อสอบด้วยวิธีวิเคราะห์ความแปรปรวน วิธีแปลงค่าความยากของข้อสอบ วิธีโค้งลักษณะข้อสอบที่มีพารามิเตอร์ 1 ตัว และวิธีโค้งลักษณะข้อสอบที่มีพารามิเตอร์ 3 ตัว ในด้านต่าง ๆ ดังต่อไปนี้

- 1.1 ดัชนีความลำเอียงของข้อสอบเป็นรายข้อ
- 1.2 สัมประสิทธิ์สหสัมพันธ์ระหว่างวิธีการวิเคราะห์ทั้ง 4 วิธี

2. เพื่อเปรียบเทียบความแตกต่างของผลการคัดเลือกก่อนและหลังการศึกษาความลำเอียงของข้อสอบตามวิธีการคิดคะแนนรวมที่แตกต่างกัน 6 วิธี ในด้านต่าง ๆ ดังต่อไปนี้

- 2.1 จำนวนผู้ที่ได้รับการคัดเลือก
- 2.2 สัดส่วนของชาย : หญิงที่ได้รับการคัดเลือก
- 2.3 ความเที่ยงของแบบสอบ

ความไม่สมบูรณ์ในการวิจัย

การวิจัยครั้งนี้ไม่อาจบอกแหล่งหรือสาเหตุที่ทำให้ข้อสอบมีความลำเอียงทางการวัดได้ ทั้งนี้เนื่องจากผู้วิจัยไม่สามารถดูเนื้อหาจากแบบสอบได้เพราะแบบสอบที่ใช้ในการคัดเลือกครั้งนี้ถือเป็นความลับของทางราชการ

คำนิยามศัพท์ที่ใช้ในการวิจัย

ข้อสอบที่มีความลำเอียง หมายถึง ข้อสอบที่ทำให้ผู้สอบซึ่งมีระดับความสามารถเท่ากัน แต่อยู่ในกลุ่มหญิงและกลุ่มชายมีโอกาสตอบข้อสอบข้อนั้นถูกไม่เท่ากัน

วิธีวิเคราะห์ความลำเอียงของข้อสอบ 4 วิธี หมายถึง วิธีการในการตรวจสอบความลำเอียงของข้อสอบด้วยวิธีการวิเคราะห์ความแปรปรวน การแปลงค่าความยากของข้อสอบ วิธีโค้งลักษณะข้อสอบที่มีพารามิเตอร์ 1 ตัว และวิธีโค้งลักษณะข้อสอบที่มีพารามิเตอร์ 3 ตัว

โค้งลักษณะข้อสอบ หมายถึง กราฟที่แสดงความสัมพันธ์ระหว่างระดับความสามารถของผู้สอบกับโอกาสที่จะตอบถูก ซึ่งโค้งลักษณะข้อสอบนี้จะถูกกำหนดด้วยค่าพารามิเตอร์ของข้อสอบ

พารามิเตอร์ของข้อสอบ หมายถึง ค่าความยาก ค่าอำนาจจำแนก และ ค่าการเดา ซึ่งคำนวณได้ด้วยทฤษฎีการตอบสนองต่อข้อสอบ

ดัชนีความลำเอียงของข้อสอบ หมายถึง ระดับความลำเอียงของข้อสอบซึ่งสามารถคำนวณได้หลายวิธี ทั้งนี้ขึ้นอยู่กับวิธีการวิเคราะห์ความลำเอียงดังนี้

1. วิธีวิเคราะห์ความแปรปรวน ดัชนีความลำเอียงของข้อสอบ หมายถึงค่าสถิติที่ใช้ทดสอบความแตกต่างของโอกาสในการตอบถูกเป็นรายข้อ
2. วิธีแปลงค่าความยากของข้อสอบ ดัชนีความลำเอียงของข้อสอบ หมายถึงระยะทางที่ข้อสอบแต่ละข้อห่างไปจากเส้นแกนหลัก
3. วิธีโค้งลักษณะข้อสอบที่มีพารามิเตอร์ 1 ตัว ดัชนีความลำเอียงของข้อสอบ หมายถึงความแตกต่างของค่าความยากระหว่างกลุ่มชายกับกลุ่มหญิง
4. วิธีโค้งลักษณะข้อสอบที่มีพารามิเตอร์ 3 ตัว ดัชนีความลำเอียงของข้อสอบ หมายถึงความแตกต่างของพื้นที่ที่คำนวณจากโค้งลักษณะข้อสอบจากผู้สอบกลุ่มชายกับกลุ่มหญิง

วิธีการคิดคะแนน 6 วิธี หมายถึง วิธีการรวมคะแนนที่ได้จากผลการสอบด้วยวิธีที่แตกต่างกัน 6 วิธีดังนี้

1. นำคะแนนดิบจากแบบสอบแต่ละวิชามารวมกัน แล้วตัดสินผลการคัดเลือกโดยเรียงตามลำดับผลรวมของคะแนนดิบ

2. นำคะแนนดิบแต่ละวิชามาแปลงให้เป็นคะแนนมาตรฐานที่ปกติ (Normalized T - score) หลังจากนั้นจึงนำเอาคะแนนมาตรฐานที่ปกติจากแต่ละวิชามารวมกัน แล้วตัดสินผลการคัดเลือกโดยเรียงตามลำดับผลรวมของคะแนนมาตรฐานที่ปกติ

3. ตัดข้อสอบที่วิเคราะห์ด้วยวิธีโค้งลักษณะข้อสอบที่มีพารามิเตอร์ 1 ตัว แล้วพบว่ามีความลำเอียงออกจากแบบสอบแต่ละฉบับ นำคะแนนดิบของแต่ละวิชามาแปลงให้เป็นคะแนนมาตรฐานที่ปกติ หลังจากนั้นจึงนำเอาคะแนนมาตรฐานที่ปกติจากแต่ละวิชารวมกัน แล้วตัดสินผลการคัดเลือกโดยเรียงตามลำดับผลรวมของคะแนนมาตรฐานที่ปกติ

4. ตัดข้อสอบที่วิเคราะห์ด้วยวิธีโค้งลักษณะข้อสอบที่มีพารามิเตอร์ 3 ตัว แล้วพบว่ามีความลำเอียงออกจากแบบสอบแต่ละฉบับ นำคะแนนดิบของแต่ละวิชามาแปลงให้เป็นคะแนนมาตรฐานที่ปกติ หลังจากนั้นจึงนำเอาคะแนนมาตรฐานที่ปกติจากแต่ละวิชารวมกัน แล้วตัดสินผลการคัดเลือกโดยเรียงตามลำดับผลรวมของคะแนนมาตรฐานที่ปกติ

5. ตัดข้อสอบที่วิเคราะห์ด้วยวิธีวิเคราะห์ความแปรปรวนแล้วพบว่ามีความลำเอียงออกจากแบบสอบแต่ละฉบับ นำคะแนนดิบของแต่ละวิชามาแปลงให้เป็นคะแนนมาตรฐานที่ปกติ หลังจากนั้นจึงนำเอาคะแนนมาตรฐานที่ปกติจากแต่ละวิชารวมกัน แล้วตัดสินผลการคัดเลือกโดยเรียงตามลำดับผลรวมของคะแนนมาตรฐานที่ปกติ

6. ตัดข้อสอบที่วิเคราะห์ด้วยวิธีแปลงค่าความยากของข้อสอบแล้วพบว่า มีความลำเอียงออกจากแบบสอบแต่ละฉบับ นำคะแนนดิบของแต่ละวิชามาแปลงให้เป็นคะแนนมาตรฐานที่ปกติ หลังจากนั้นจึงนำเอาคะแนนมาตรฐานที่ปกติจากแต่ละวิชารวมกัน แล้วตัดสินผลการคัดเลือกโดยเรียงตามลำดับผลรวมของคะแนนมาตรฐานที่ปกติ

แบบสอบแข่งขันเพื่อบรรจุเข้ารับราชการ หมายถึง แบบสอบที่ใช้แข่งขันเพื่อบรรจุเข้ารับราชการเป็นข้าราชการครู สังกัดสำนักงานคณะกรรมการการประถมศึกษาแห่งชาติ ปี พ.ศ. 2529 ซึ่งมี 4 ฉบับ คือ

- ฉบับที่ 1 แบบสอบวิชาเอก (ความรู้เกี่ยวกับการประถมศึกษา)
- ฉบับที่ 2 แบบสอบวิชาการศึกษาและกฎหมาย
- ฉบับที่ 3 แบบสอบภาษาไทย
- ฉบับที่ 4 แบบสอบความสามารถทั่วไป

ประโยชน์ที่คาดว่าจะได้รับจากการวิจัย

1. เพื่อเป็นแนวทางในการปรับปรุงแก้ไขข้อสอบที่มีความลำเอียงให้สามารถใช้ได้ อย่างยุติธรรมสำหรับนักเรียนต่างกลุ่ม อันจะทำให้ผลการคัดเลือกมีความถูกต้อง และเป็นธรรม มากขึ้น
2. ผลจากการวิจัยจะช่วยให้ทราบว่าวิธีการวิเคราะห์ความลำเอียงทั้ง 4 วิธี ให้ผล สอดคล้องหรือแตกต่างกันเพียงไร หากผลการวิจัยพบว่าวิธีวิเคราะห์ดังกล่าวมีความสอดคล้องกัน มาก ก็ทำให้สามารถเลือกใช้วิธีที่ง่าย สะดวก แต่หากผลการวิจัย พบว่าวิธีการวิเคราะห์ ดังกล่าวมีความสอดคล้องกันน้อย ก็ทำให้สามารถเลือกใช้วิธีการวิเคราะห์ที่ให้ผลการวิจัยที่ ชัดเจนกว่า
3. ผลจากการวิจัยครั้งนี้สามารถนำไปเป็นข้อเสนอแนะให้คณะกรรมการ หรือผู้ที่มี หน้าที่เกี่ยวข้องกับ การสอบคัดเลือกได้นำไปใช้เป็นข้อมูลส่วนหนึ่งในการพิจารณาตัดสินผลการ คัดเลือก เพื่อให้ผลการคัดเลือกมีความถูกต้องและเป็นธรรมมากยิ่งขึ้น ตลอดจนเพื่อให้คณะ กรรมการได้พิจารณาเรื่องนี้อย่างรอบคอบ
4. เพื่อเป็นแนวทางในการศึกษาเกี่ยวกับวิธีการวิเคราะห์ความลำเอียงแบบต่าง ๆ ให้แพร่หลายมากยิ่งขึ้น