

## ปัญหาและความเป็นมาของปัญหา

ปัจจุบันวงการศึกษาศึกษาศาสตร์และการทหาร ใช้การสอบเป็นเครื่องมือวัดความสามารถเพื่อคัดเลือกบุคคล ตรวจสอบความรู้ความสามารถของบุคคล สนับสนุนการเลื่อนตำแหน่ง และการออกใบรับรองหรือใบอนุญาตกันอย่างแพร่หลาย มีการนำผลการสอบของผู้เข้าสอบไปใช้ประกอบการพิจารณาตัดสินใจในเรื่องต่าง ๆ เหล่านี้เพิ่มมากขึ้นทำให้ปัญหาเรื่องความไม่ยุติธรรมหรือความลำเอียงของข้อสอบหรือแบบทดสอบที่ใช้ในการสอบแต่ละครั้งเป็นประเด็นสำคัญที่ใช้พิจารณาความตรงของแบบทดสอบ เนื่องจากข้อสอบหรือแบบทดสอบที่ลำเอียงเข้าข้างกลุ่มผู้สอบย่อยบางกลุ่มของผู้สอบทั้งหมด อาจทำให้ผู้สอบกลุ่มย่อยนั้นได้เปรียบผู้เข้าสอบกลุ่มย่อยกลุ่มอื่นๆ ทั้งๆ ที่สอบด้วยข้อสอบเดียวกันหรือแบบทดสอบฉบับเดียวกัน ผู้เข้าสอบในการสอบแต่ละครั้ง อาจจำแนกออกเป็นกลุ่มย่อยๆ ได้ ตามลักษณะที่แตกต่างกันในด้านเชื้อชาติ เผ่าพันธุ์ เพศ ศาสนา ภาษา อายุ ประสบการณ์ ทูพลาภาพ หรือภูมิหลังอื่นๆ ของผู้เข้าสอบ ที่ทำให้กลุ่มผู้สอบย่อยๆ บางกลุ่มเกิดการเสียเปรียบ (Scheuneman and Bleistein, 1989 : 255)

ตัวอย่างของความลำเอียงของเครื่องมือวัดในลักษณะที่อาจทำให้เกิดการได้เปรียบและเสียเปรียบในการสอบ ระหว่างกลุ่มผู้สอบย่อยที่มีความสามารถเท่ากันในเรื่องที่เครื่องมือวัดต้องการวัด ได้แก่ การใช้แบบวัดความถนัดกับผู้สอบในประเทศที่มีประชากรประกอบด้วยประชากรกลุ่มย่อยๆ หลายเชื้อชาติที่มีภูมิหลังด้านวัฒนธรรมแตกต่างกัน เช่น ประเทศสหรัฐอเมริกา ซึ่งประกอบด้วยชนชาวผิวขาว ชนพื้นเมืองชนผิวดำและผิวสีอื่น ๆ ซึ่งต่างก็มีสัญชาติอเมริกันเช่นเดียวกัน แบบวัดความถนัดที่ใช้ร่วมกันจะมีความตรงและไม่ลำเอียงก็ต่อเมื่อผลที่ได้จากการวัดเป็นเรื่องของความถนัดเพียงอย่างเดียว ความแตกต่างของผลการสอบระหว่างกลุ่มย่อยๆ ดังกล่าว เป็นความแตกต่างของระดับความถนัดเท่านั้น ไม่ใช่เพราะความแตกต่างด้านวัฒนธรรมพื้นฐานของผู้สอบที่มีภูมิหลังต่างกัน กล่าวอีกนัยหนึ่ง แบบทดสอบต้องปราศจากความลำเอียง ด้านเชื้อชาติ และวัฒนธรรม (Culture Free และ Culture Fair) มิฉะนั้นผลการสอบที่แตกต่างกัน จะไม่สามารถอธิบายได้ว่า เป็นความแตกต่างอันเนื่องมาจากความถนัดวัฒนธรรมหรืออื่น ๆ ลักษณะของกลุ่มประชากรย่อยๆ ที่แตกต่างกันในตัวแปรบางตัว ทำให้

คะแนนจากแบบทดสอบหรือข้อสอบขาดความตรงในการแปลความหมายเช่นนี้เรียกว่า แบบทดสอบหรือ ข้อสอบนั้นมีความลำเอียง ทำให้ผู้สอบต่างกลุ่มกันมีความได้เปรียบเสียเปรียบกัน เพราะแบบทดสอบ หรือข้อสอบนั้นไม่ได้วัดมิติหรือลักษณะแฝง (Latent Trait) ที่สนใจเท่านั้น แต่วัดมิติอื่น ๆ อยู่ด้วยทำให้การแปลความหมายของคะแนนไม่บริสุทธิ์ และทำให้ผลจากการวัดเบี่ยงเบนไปจากวัตถุประสงค์ที่กำหนดไว้ ผู้ที่มีความรู้ในมิติที่แฝงเข้ามามากกว่าจะมีโอกาสตอบแบบทดสอบ หรือข้อสอบได้ถูกต้องมากกว่า (กาญจนา วัฒนสุนทร, 2538:2)

การศึกษาเกี่ยวกับผลการสอบของกลุ่มผู้เข้าสอบย่อยของผู้เข้าสอบทั้งหมดมีมานานแล้ว แต่เรื่องความยุติธรรมในการสอบระหว่างผู้เข้าสอบกลุ่มย่อย ๆ เพิ่งมีการศึกษากันอย่างจริงจังในช่วงปลายทศวรรษ 1960 โดยมีการเสนอวิธีการต่าง ๆ เพื่อนำไปใช้ตรวจสอบความลำเอียงของแบบทดสอบ (Test-bias) หรือความลำเอียงในการคัดเลือกผู้เข้าสอบ (Selection-bias) ขึ้นหลายวิธี ขณะเดียวกันในช่วงเวลานั้นนักพัฒนาแบบทดสอบก็สนใจวิธีการจำแนกข้อสอบที่ไม่เหมาะสมกับผู้เข้าสอบบางกลุ่มออกจากแบบทดสอบก่อนจะพัฒนาเป็นแบบทดสอบฉบับสมบูรณ์ ทำให้จำเป็นต้องพัฒนาวิธีการตรวจสอบความลำเอียงของข้อสอบ (Item-bias) ขึ้นเพื่อใช้เป็นแนวทางในการจำแนกข้อสอบที่ลำเอียงกับผู้เข้าสอบบางกลุ่มออกจากแบบทดสอบหรือคลังข้อสอบ ปัจจุบันการตรวจสอบความลำเอียงของข้อสอบ เป็นส่วนหนึ่งของกระบวนการพัฒนาและการประเมินแบบทดสอบ เช่นเดียวกับการวิเคราะห์ข้อสอบและการตรวจสอบความเชื่อมั่นของแบบทดสอบ (Hambleton, et al. 1993:237)

วิธีการทางสถิติที่ใช้ตรวจค้นความลำเอียงของข้อสอบ แบ่งได้หลายวิธี ซึ่งแฮมเบิตตัน และคณะ (Hambleton, et al. 1993:237) ได้จำแนกออกเป็น 3 กลุ่มใหญ่ ๆ ดังนี้

#### 1. กลุ่มวิธีที่ใช้ทฤษฎีการสอบแบบคลาสสิกอล(Methods Using Classical Test Theory)

วิธีการตรวจค้นความลำเอียงของข้อสอบในกลุ่มนี้พัฒนามาจากหลักการของทฤษฎีการสอบแบบคลาสสิกอล โดยปกติแล้วใช้คะแนนที่สังเกตได้ของผู้เข้าสอบ (Observed Scores) แต่ละคนเป็นเกณฑ์การจับคู่กลุ่มผู้เข้าสอบย่อยและเปรียบเทียบค่าความยากของข้อสอบแต่ละข้อระหว่างกลุ่มผู้เข้าสอบย่อยเหล่านั้น วิธีการในกลุ่มนี้ ได้แก่ การวิเคราะห์ความแปรปรวน (Analysis of Variance) วิธีสหสัมพันธ์ (Correlation Methods) (Green & Draper, 1972 quoted in Scheuneman Bleitein, 1989) วิธีแปลงค่าความยากของข้อสอบ (Transformed Item Difficulty Method) หรือวิธีกำหนดจุดค่าเดต้า (Delta - plot Method) (Angoff, 1982) การวิเคราะห์ตัวลวง (Distractor Analysis) (Scheuneman, 1982) วิธีสหสัมพันธ์บางส่วน (Partial Correlation Methods) (Strecker, 1982) และวิธีการทำให้เป็นมาตรฐาน (Standardization Method) (Dorans & Kulick, 1983 : 1986)

ข้อได้เปรียบของวิธีในกลุ่มนี้คือ กระบวนการตรวจสอบความลำเอียงของข้อสอบไม่ยุ่งยาก เสียค่าใช้จ่ายไม่สูงนัก ใช้ตรวจสอบกับกลุ่มตัวอย่างขนาดเล็กได้และสามารถอธิบายให้คนทั่วไปเข้าใจได้ง่าย ส่วนข้อเสียเปรียบก็คือ ค่าสถิติของข้อสอบเปลี่ยนไปตามกลุ่มตัวอย่าง เมื่อกกลุ่มตัวอย่างเปลี่ยนไปผลการตรวจพบข้อสอบที่มีความลำเอียงก็เปลี่ยนไปทำให้การอ้างอิงผลการศึกษาไปยังกลุ่มประชากรอาจมีความเชื่อถือได้น้อยลง

## 2. กลุ่มวิธีที่ใช้ทฤษฎีการตอบสนองข้อสอบ (Methods Using Item Response Theory)

วิธีการในกลุ่มนี้ตรวจสอบความลำเอียงของข้อสอบ ตามกรอบแนวคิดของทฤษฎีการตอบสนองข้อสอบ โดยปกติแล้วใช้การเปรียบเทียบโค้งลักษณะข้อสอบ (Item Characteristic Curves : ICCs) ของกลุ่มผู้เข้าสอบย่อยตามระดับความสามารถของผู้เข้าสอบ ถ้าโค้งลักษณะข้อสอบของกลุ่มผู้เข้าสอบย่อยสองกลุ่มมีรูปร่างเหมือนกัน แสดงว่าข้อสอบข้อนั้นไม่มีความลำเอียง แต่ถ้าโค้งลักษณะข้อสอบของกลุ่มผู้สอบย่อยสองกลุ่มมีรูปร่างต่างกัน แสดงว่าข้อสอบข้อนั้นทำหน้าที่ต่างกัน ค่าพารามิเตอร์ของโค้งลักษณะข้อสอบ ได้แก่ ค่าความยากของข้อสอบ (Item Difficulty, b-parameter) ค่าอำนาจจำแนกของข้อสอบ (Item Discrimination, a-parameter) และค่าการเดาข้อสอบ (Pseudo Guessing Parameter, c-parameter) วิธีการในกลุ่มนี้ได้แก่ วิธี Analysis of Fit (Durovic, 1975, quoted in Hambleton & Others, 1993) ซึ่งใช้โมเดล IRT แบบหนึ่งพารามิเตอร์ วิธี IRT Area (Ironson and Subkoviak, 1979; Raju, 1988; 1990) และวิธี Two Stage (Lord, 1980) ซึ่งใช้โมเดล IRT แบบสองพารามิเตอร์ หรือสามพารามิเตอร์ และวิธี Plot (Hambleton and Rogers, 1991, quoted in Hambleton and Others, 1993)

ข้อได้เปรียบวิธีการในกลุ่มนี้คือ การแก้ไขข้อบกพร่องของทฤษฎีการสอบแบบคลาสสิกอล ทำให้ค่าสถิติของข้อสอบไม่เปลี่ยนไปตามกลุ่มตัวอย่างที่สุ่มมาจากประชากรเดียวกัน การประมาณค่าความสามารถของผู้สอบเป็นอิสระจากค่าความยากของแบบทดสอบ โมเดลทางคณิตศาสตร์ง่ายต่อการจับคู่โค้งลักษณะข้อสอบตามระดับความสามารถของผู้สอบทำให้สามารถศึกษาความแตกต่างของผลการตอบข้อสอบตามระดับความสามารถของผู้สอบย่อยได้ ไม่ต้องมีข้อตกลงเบื้องต้นเรื่องแบบทดสอบคู่ขนานในการหาค่าสัมประสิทธิ์ความเชื่อมั่นของแบบทดสอบ และถ้าผลการตอบข้อสอบของกลุ่มผู้สอบสอดคล้องกับข้อตกลงเบื้องต้นของโมเดล IRT แล้ว วิธีในกลุ่มนี้ก็มักจะเป็นวิธีการตรวจสอบความลำเอียงของข้อสอบที่ได้ผลดีวิธีหนึ่ง เนื่องจากเป็นวิธีที่มีทฤษฎีการตอบสนองข้อสอบสนับสนุนและใช้การประมาณค่าความสามารถที่แท้จริงของผู้สอบแทนคะแนนที่สังเกตได้ ดังเช่นที่ใช้ในกลุ่มวิธีที่ใช้ทฤษฎีการสอบแบบคลาสสิกอล ส่วนข้อเสียเปรียบของวิธีการในกลุ่มนี้ คือ กระบวนการตรวจสอบความลำเอียงของข้อสอบสลับซับซ้อน เสียค่าใช้จ่ายในการดำเนินการสูง และต้องการกลุ่มตัวอย่างขนาดใหญ่

### 3. กลุ่มวิธีที่ใช้วิธีไคสแควร์ (Methods Using ChiSquare Methods)

วิธีการตรวจสอบความลำเอียงของข้อสอบในกลุ่มนี้ บางครั้งก็เรียกว่ากลุ่มวิธีไคสแควร์ (Chisquare Methods) วิธีการในกลุ่มนี้ใช้ค่าสถิติไคสแควร์แสดงความลำเอียงของข้อสอบ และใช้คะแนนของแบบทดสอบหรือคะแนนของแบบทดสอบที่ทำให้บริสุทธิ์ (Purified Test Score) เป็นเกณฑ์การจับคู่กลุ่มผู้สอบย่อยสองกลุ่มที่ศึกษาก่อนการเปรียบเทียบผลการตอบข้อสอบ วิธีการในกลุ่มนี้ ได้แก่ วิธีตารางการณัจจร (Contingency Table Method) (Scheuneman, 1975 ; 1979) วิธีตารางการณัจจรปรับขยาย (Modified Contingency Table Method) (Veale, 1977, quoted in Hambleton and Others, 1993) วิธีล็อก - ลินีเยร์ (Log - Linear Models) (Mellenbrgh, 1982) วิธีแมนเทล-เฮนส์เซล (Mantel-Haenszel method : MH) (Holland and Thayer, 1986; 1988) และวิธีการถดถอยโลจิสติก (Logistic Regression Method) (Swaminathan and Rogers, 1990)

ข้อได้เปรียบของวิธีการในกลุ่มนี้ก็คือ กระบวนการตรวจสอบความลำเอียงของข้อสอบไม่ยุ่งยากซับซ้อน เสียค่าใช้จ่ายในการดำเนินการไม่สูง ใช้กับกลุ่มตัวอย่างขนาดไม่ใหญ่มาก บางวิธีการมีหลักการที่ดีในการจับคู่กลุ่มผู้สอบย่อยตามความสามารถของผู้สอบ และมีการทดสอบนัยสำคัญ ส่วนข้อเสียเปรียบของวิธีการในกลุ่มนี้ก็คล้าย ๆ กับวิธีที่ใช้ทฤษฎีการสอบแบบคลาสสิกอล

จากการศึกษาเปรียบเทียบการวิเคราะห์ความลำเอียงของข้อสอบที่ผ่านมาจนถึงปัจจุบัน มีวิธีวิเคราะห์หลายวิธีแต่วิธีที่นิยมใช้กันมากในการเปรียบเทียบ ได้แก่ วิธีแปลงค่าความยาก (Transformed Item Difficulty) การวิเคราะห์ความแปรปรวน (Analysis of Variance) การวิเคราะห์ด้วย ไคสแควร์ (Chi - Square) วิธีวิเคราะห์องค์ประกอบ (Factor Analysis) วิธีวิเคราะห์การถดถอย (Regression Analysis) วิธีล็อกลินีเยร์ (Log - Linear) วิธีการถดถอยโลจิสติก (Logistic Regression) วิธีค่าอำนาจจำแนกของข้อสอบ (Item Discrimination Indices) วิธีของแมนเทล-เฮนส์เซล (Mantel - Haenszel) ซึ่งเป็นวิธีที่ใช้หลักการของทฤษฎีการวัดแบบคลาสสิกอล (Classical Test Theory) นอกจากวิธีวิเคราะห์ความลำเอียงดังกล่าวแล้ว ยังมีวิธีวิเคราะห์ความลำเอียงโดยใช้หลักการของทฤษฎีการตอบข้อคำถาม (Item Response Theory) ได้แก่ วิธีโค้งลักษณะข้อสอบ 3 พารามิเตอร์ (Item Characteristic Curve-3 Parameter) วิธีโค้งลักษณะข้อสอบ 2 พารามิเตอร์ (Item Characteristic Curve-2 Parameter) วิธีโค้งลักษณะข้อสอบ 1 พารามิเตอร์ (Item Characteristic Curve-1 Parameter) ซึ่งวิธีการวิเคราะห์ความลำเอียงของข้อสอบแต่ละวิธีที่กล่าวมามีจุดเด่นและจุดด้อยแตกต่างกันดังที่กล่าวมาแล้ว มีนักวิจัยหลายท่านได้ทำการศึกษาเปรียบเทียบวิธีการวิเคราะห์ความลำเอียงของข้อสอบแต่ละวิธี เช่น รัตเนอร์และคนอื่น ๆ (Subkoviak, 1984 : 5, quoting Rudner and Others, 1978) ได้ศึกษาเปรียบเทียบวิธี

วิเคราะห์ความลำเอียงของข้อสอบ 4 วิธี คือ วิธีแปลงค่าความยาก วิธีโคสแควร์ วิธีโค้งลักษณะข้อสอบ และวิธีวิเคราะห์องค์ประกอบ ผลการวิจัยปรากฏว่า วิธีโคสแควร์ วิธีโค้งลักษณะข้อสอบ และวิธีแปลงค่าความยาก เป็นวิธีที่เหมาะสม สงบ ลักษณะ (Burrill, 1982 : 168, quoting Laksana, 1979) ศึกษาเปรียบเทียบวิธีวิเคราะห์ความลำเอียง 2 วิธี คือ วิธีโค้งลักษณะข้อสอบ 3 พารามิเตอร์ และวิธีวิเคราะห์ความแปรปรวน พบว่า วิธีโค้งลักษณะข้อสอบ 3 พารามิเตอร์เป็นวิธีที่ให้ผลครอบคลุม แต่วิธีวิเคราะห์ความแปรปรวนใช้ได้ดีในทางปฏิบัติ เชปาร์ดและคณะ (Berk, 1982 : 169, quoting Shepard, Camilli and Averrill, 1980) ได้ทำการศึกษาเปรียบเทียบวิธีวิเคราะห์ความลำเอียง 5 วิธี คือ วิธีโคสแควร์ วิธีแปลงค่าความยาก วิธีพอยท์ไบซีเรียล วิธีโค้งลักษณะข้อสอบ 3 พารามิเตอร์ และวิธีโค้งลักษณะข้อสอบ 1 พารามิเตอร์ พบว่า วิธีโค้งลักษณะข้อสอบ 3 พารามิเตอร์ เป็นวิธีเหมาะสมในทางทฤษฎี ถูกต้องและตรงที่สุด สำหรับวิธีโคสแควร์ให้ผลสอดคล้องกับวิธีโค้งลักษณะข้อสอบ 3 พารามิเตอร์ ส่วนสวามินาธานและโรเจอร์ส (Swaminathan and Rogers, 1990:361) ศึกษาความลำเอียงของข้อสอบด้วยวิธีการลดรอยแบบโลจิสและวิธีของแมนเทล-แฮนส์เซล พบว่า ทั้งสองวิธีให้ผลใกล้เคียงกันและสามารถใช้แทนกันได้ ทศนีย์ พิรมนตรี (2530 : 76-78) ได้ศึกษาเปรียบเทียบวิธีวิเคราะห์ความลำเอียงของข้อสอบ 3 วิธี คือ วิธีกำหนดจุดเดลตา วิธีโคสแควร์ และวิธีโค้งลักษณะข้อสอบ 3 พารามิเตอร์ พบว่า วิธีโค้งลักษณะข้อสอบ 3 พารามิเตอร์ พบข้อสอบมีความลำเอียงมากที่สุด สุรศักดิ์ อมรรัตนศักดิ์ (2531 : 200 – 207) ได้ทำการศึกษาเปรียบเทียบวิธีวิเคราะห์ความลำเอียงของข้อสอบ 4 วิธี คือ วิธีวิเคราะห์ความแปรปรวน วิธีแปลงค่าความยาก วิธีโค้งลักษณะข้อสอบ 1 พารามิเตอร์ และวิธีโค้งลักษณะข้อสอบ 3 พารามิเตอร์ พบว่า วิธีโค้งลักษณะข้อสอบ 3 พารามิเตอร์ พบข้อสอบที่มีความลำเอียงมากที่สุด รองลงมาเป็นวิธีวิเคราะห์ความแปรปรวนและวิธีโคสแควร์ ส่วนวิธีแปลงค่าความยากของข้อสอบเป็นวิธีที่พบข้อสอบที่มีความลำเอียงน้อยที่สุด

จากการศึกษารายงานการวิจัยแสดงให้เห็นว่า แบบทดสอบวัดความสามารถด้านต่าง ๆ มีข้อสอบบางข้อที่มีความลำเอียง เมื่อนำไปทดสอบกับกลุ่มที่มีความแตกต่างกันทางด้านเพศ เชื้อชาติ ฐานะทางเศรษฐกิจ สังคม วัฒนธรรม ศาสนา และสภาพภูมิศาสตร์ จะทำให้ผลการสอบที่แตกต่างกัน ทั้งนี้เพราะระหว่างกลุ่มผู้สอบย่อนั้นมีอิทธิพลมาจากสิ่งแวดล้อม วัฒนธรรม และความเคยชินที่แตกต่างกัน จึงเป็นเหตุให้การได้เปรียบเสียเปรียบเกี่ยวกับเนื้อหาของข้อสอบ ซึ่งในประเทศไทยของเราก็มีวัฒนธรรม ศาสนา และภาษาในแต่ละท้องถิ่นที่แตกต่างกัน แต่มีการศึกษาวิจัยที่เกี่ยวข้องกับวัฒนธรรม ศาสนาและภาษาว่าจะส่งผลกระทบต่อแบบทดสอบ ทำให้เกิดมีความลำเอียงขึ้นหรือไม่ ยังมีอยู่น้อยมาก

ด้วยเหตุผลดังกล่าวข้างต้นการศึกษาค้นคว้าครั้งนี้ผู้วิจัยจะศึกษาวิธีวิเคราะห์ความลำเอียงของข้อสอบ 3 วิธี คือ วิธีแปลงค่าความยากของข้อสอบ วิธีโค้งลักษณะข้อสอบแบบ 3 พารามิเตอร์ และวิธีของแมนเทิล – แอนส์เซต จะให้ผลการวิเคราะห์ความลำเอียงแตกต่างกันหรือสอดคล้องกันเพียงไร ในด้านจำนวนข้อสอบที่มีความลำเอียง ค่าความเชื่อมั่นและค่าฟังก์ชันสารสนเทศของแบบทดสอบหลังจากคัดข้อสอบที่มีความลำเอียงออกแล้ว โดยศึกษากับกลุ่มนักเรียนที่ใช้ภาษาในชีวิตประจำวันต่างกัน อันจะเป็นแนวทางในการนำวิธีการวิเคราะห์ความลำเอียงมาใช้ในทางปฏิบัติ ตลอดจนการปรับปรุงแก้ไขข้อสอบที่มีความลำเอียงให้สามารถนำไปใช้ได้อย่างยุติธรรมต่อไป

### วัตถุประสงค์ของการวิจัย

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อวิเคราะห์ความลำเอียงของข้อสอบจากแบบทดสอบ โจทย์ปัญหาทางคณิตศาสตร์ ด้วยวิธีวิเคราะห์ความลำเอียงแตกต่างกัน 3 วิธี คือ วิธีแปลงค่าความยากของข้อสอบ วิธีโค้งลักษณะข้อสอบ 3 พารามิเตอร์ และวิธีของแมนเทิล – แอนส์เซต สำหรับกลุ่มผู้สอบจำแนกตามภาษาที่ใช้ในชีวิตประจำวัน โดยมีวัตถุประสงค์ดังนี้

1. เปรียบเทียบจำนวนข้อสอบที่มีความลำเอียงระหว่างวิธีวิเคราะห์ความลำเอียงต่างกัน ของแบบทดสอบ โจทย์ปัญหาทางคณิตศาสตร์
2. เปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบหลังการคัดเลือกข้อสอบที่มีความลำเอียงออก ระหว่างวิธีวิเคราะห์ความลำเอียงต่างกัน ของแบบทดสอบ โจทย์ปัญหาทางคณิตศาสตร์
3. เปรียบเทียบค่าฟังก์ชันสารสนเทศของแบบทดสอบหลังการคัดเลือกข้อสอบที่มีความลำเอียงออก ระหว่างวิธีวิเคราะห์ความลำเอียงต่างกัน ของแบบทดสอบ โจทย์ปัญหาทางคณิตศาสตร์

### สมมติฐานของการวิจัย

1. เมื่อใช้วิธีวิเคราะห์ความลำเอียงต่างกัน แบบทดสอบ โจทย์ปัญหาทางคณิตศาสตร์จะมีจำนวนข้อสอบที่ลำเอียงแตกต่างกัน
2. เมื่อใช้วิธีวิเคราะห์ความลำเอียงต่างกัน แบบทดสอบ โจทย์ปัญหาทางคณิตศาสตร์ที่คัดเลือกข้อสอบที่มีความลำเอียงออกแล้ว จะมีค่าความเชื่อมั่นแตกต่างกัน
3. เมื่อใช้วิธีวิเคราะห์ความลำเอียงต่างกัน แบบทดสอบ โจทย์ปัญหาทางคณิตศาสตร์ที่คัดเลือกข้อสอบที่มีความลำเอียงออกแล้ว จะมีค่าฟังก์ชันสารสนเทศแตกต่างกัน

## ความสำคัญของการวิจัย

1. ผลจากการวิจัยจะช่วยให้ทราบว่าวิธีการวิเคราะห์ความลำเอียงทั้ง 3 วิธี ให้ผลสอดคล้องหรือแตกต่างกันเพียงไร หากผลการวิจัยพบว่าวิธีวิเคราะห์ดังกล่าวมีความสอดคล้องกันมาก ก็ทำให้สามารถเลือกใช้วิธีที่ง่าย สะดวก แต่หากผลการวิจัย พบว่าการวิเคราะห์ดังกล่าวมีความสอดคล้องกันน้อย ก็ทำให้สามารถเลือกใช้วิธีการวิเคราะห์ที่เหมาะสมและมีประสิทธิภาพสูงกว่า
2. เพื่อนำเสนอวิธีการพัฒนาคุณภาพแบบทดสอบที่มีความยุติธรรม สำหรับการที่จะใช้แบบทดสอบในการตัดสินเกี่ยวกับผู้สอบได้อย่างถูกต้องและไม่ลำเอียงต่อไป
3. เพื่อเป็นแนวทางในการศึกษาเกี่ยวกับวิธีการวิเคราะห์ความลำเอียงแบบต่าง ๆ ให้แพร่หลายมากยิ่งขึ้น

## ขอบเขตของการวิจัย

### 1. ประชากรที่ใช้ในการศึกษาวิจัย

ประชากรที่ใช้ในการศึกษาวิจัยครั้งนี้เป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ภาคเรียนที่ 2 ปีการศึกษา 2542 ของโรงเรียนสังกัดสำนักงานคณะกรรมการการประถมศึกษาจังหวัดยะลา จำนวน 210 โรงเรียน 271 ห้องเรียน มีจำนวนนักเรียนทั้งหมด 7,531 คน

### 2. ตัวแปรในการวิจัย

2.1 ตัวแปรอิสระ ได้แก่ วิธีวิเคราะห์ความลำเอียง 3 วิธี คือ

2.1.1 วิธีแปลงค่าความยากของข้อสอบ

2.1.2 วิธีโค้งลักษณะข้อสอบ 3 พารามิเตอร์

2.1.3 วิธีของแมนเทล - แชนต์เซล

2.2 ตัวแปรตาม ได้แก่

2.2.1 จำนวนข้อสอบที่มีความลำเอียง

2.2.2 ความเชื่อมั่นของแบบทดสอบหลังคัดเลือกข้อสอบที่มีความลำเอียงออกแล้ว

2.2.3 ฟังก์ชันสารสนเทศของแบบทดสอบหลังคัดเลือกข้อสอบที่มีความลำเอียงออกแล้ว

3. เครื่องมือที่ใช้ในการวิจัย เป็นแบบทดสอบโจทย์ปัญหาทางคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 ในเรื่อง จำนวน และการบวก การลบ การคูณ การหาร ตามหลักสูตรประถมศึกษา พุทธศักราช 2521 (ฉบับปรับปรุง พ.ศ. 2533) เป็นแบบเลือกตอบ ชนิด 4 ตัวเลือก จำนวน 40 ข้อ
4. ในการวิจัยครั้งนี้ผู้วิจัยคำนวณค่าความเชื่อมั่นของแบบทดสอบหลังจากคัดเลือกข้อที่มีความลำเอียงออกแล้ว โดยใช้ข้อมูลผลการสอบเดิม (ฉบับสมบูรณ์ที่มี 40 ข้อ) โดยการคำนวณค่าเฉพาะข้อมูลผลการสอบของข้อที่ผ่านการคัดเลือกไว้เท่านั้น และมีการปรับขยาย เพื่อหาค่าประมาณความเชื่อมั่นของแบบทดสอบที่มี 40 ข้อ อีกครั้งตามวิธีการของสเปียร์แมน-บราว : Spearman-Brown (Allen and Yen, 1979) เพื่อใช้ในการเปรียบเทียบค่าประมาณความเชื่อมั่นต่อไป

### นิยามศัพท์เฉพาะ

1. แบบทดสอบโจทย์ปัญหาทางคณิตศาสตร์ หมายถึง ข้อสอบที่วัดความสามารถทางด้านการแก้ปัญหา ซึ่งแบบทดสอบประเภทนี้ประกอบด้วยคำถามประเภทโจทย์ต่าง ๆ ที่มีภาษาเข้ามาเกี่ยวข้องทำงานองเดียวกันกับโจทย์คณิตศาสตร์ทั่วไป แต่ลักษณะของคำถามมุ่งที่จะวัดความสามารถในการแปลความหมาย และการหาความสัมพันธ์ของตัวเลข จำนวน และการใช้เทคนิคหรือวิธีการแก้ปัญหาต่าง ๆ
2. ความลำเอียงของข้อสอบ หมายถึง ข้อสอบที่ทำให้ผลการสอบของผู้สอบที่มีความสามารถเท่ากันในเรื่องที่ต้องการวัด มีโอกาสตอบข้อสอบนั้นได้ถูกต้องไม่เท่ากัน เนื่องจากอยู่ในกลุ่มต่างกัน ในการวิจัยครั้งนี้ศึกษาความลำเอียงของข้อสอบ เมื่อใช้กับกลุ่มผู้สอบที่ใช้ภาษาในชีวิตประจำวันต่างกัน
3. การวิเคราะห์ความลำเอียง หมายถึง การหาค่าดัชนีความลำเอียงของข้อสอบในแบบทดสอบ โจทย์ปัญหาทางคณิตศาสตร์ สำหรับการวิจัยครั้งนี้ใช้ตัวแปรที่วัดความลำเอียงคือ การใช้ภาษาในชีวิตประจำวัน ซึ่งจำแนกเป็นใช้ภาษาไทยในชีวิตประจำวันและไม่ใช้ภาษาไทยในชีวิตประจำวัน
4. วิธีวิเคราะห์ความลำเอียงของข้อสอบ หมายถึง วิธีการที่ใช้ในการวิเคราะห์หาค่าดัชนีความลำเอียงของข้อสอบ ซึ่งในที่นี้ใช้วิธีวิเคราะห์ 3 วิธี คือ วิธีแปลงค่าความยากของข้อสอบ วิธีโค้งลักษณะข้อสอบ 3 พารามิเตอร์ และวิธีของแมนเทล-แฮนส์เซล
5. วิธีแปลงค่าความยากของข้อสอบ (Transformed Item Difficulties Approach) หมายถึง การวิเคราะห์หาค่าดัชนีความลำเอียงของข้อสอบโดยพิจารณาค่าระยะห่างตั้งฉากจากคู่อันดับเคลตาของข้อสอบแต่ละข้อไปยังเส้นแกนหลัก ข้อสอบที่มีความลำเอียงจะต้องมีค่าระยะห่างตั้งฉากมากกว่า  $3S_d$  หรือน้อยกว่า  $-3S_d$



6. วิธีโค้งลักษณะข้อสอบ (Item Characteristic Curve Method) หมายถึง การวิเคราะห์หาดัชนีความลำเอียงของข้อสอบโดยหาค่าพารามิเตอร์ของข้อสอบ 3 ตัว ซึ่งคำนวณโดยใช้โปรแกรม BILOG (version 3.04) แล้วสร้างเป็นโค้งลักษณะข้อสอบ แล้วดูความแตกต่างของพื้นที่ใต้จากโค้งลักษณะข้อสอบของกลุ่มผู้สอบที่ใช้ภาษาไทยในชีวิตประจำวันกับกลุ่มผู้สอบที่ไม่ใช้ภาษาไทยในชีวิตประจำวัน ซึ่งคำนวณโดยใช้โปรแกรม IRTDIF ข้อสอบที่มีความลำเอียงจะต้องมีพื้นที่ใต้โค้งต่างกันตั้งแต่ .40 ขึ้นไป
7. โค้งลักษณะข้อสอบ หมายถึง เส้นกราฟของข้อสอบซึ่งแสดงความสัมพันธ์ระหว่างระดับความสามารถของผู้สอบกับค่าความน่าจะเป็นของการตอบข้อสอบข้อนั้นถูก โค้งลักษณะข้อสอบนี้จะถูกกำหนดด้วยพารามิเตอร์ของข้อสอบ
8. ค่าพารามิเตอร์ของข้อสอบ หมายถึง ค่าความยากของข้อสอบ ค่าอำนาจจำแนกของข้อสอบ และค่าสัมประสิทธิ์การเดา ซึ่งคำนวณโดยใช้โปรแกรม BILOG (version 3.04)
9. วิธีของแมนเทล-แฮนส์เซล หมายถึง การวิเคราะห์หาดัชนีความลำเอียงของข้อสอบ โดยพิจารณาเปรียบเทียบผลการตอบข้อสอบของกลุ่มผู้สอบ 2 กลุ่ม ซึ่งคำนวณโดยใช้โปรแกรม MHDIF ข้อสอบที่มีความลำเอียง จะต้องมียค่า  $\alpha_{MH}$  มากกว่า 1 หรือ  $\chi^2_{MH}$  ที่คำนวณได้มีนัยสำคัญทางสถิติที่ระดับ .05
10. กลุ่มอ้างอิง (Reference Group : R) หมายถึง กลุ่มผู้สอบที่คาดว่าจะได้ประโยชน์จากข้อสอบที่ลำเอียง คือ เป็นกลุ่มที่มีความน่าจะเป็นในการตอบข้อสอบที่ลำเอียงได้ถูกต้องมากกว่าผู้สอบอีกกลุ่มหนึ่งที่ได้คะแนนจากแบบทดสอบเท่ากัน สำหรับกลุ่มอ้างอิงในการวิจัยครั้งนี้ คือ กลุ่มผู้สอบที่ใช้ภาษาไทยในชีวิตประจำวัน
11. กลุ่มเปรียบเทียบ (Focal Group : F) หมายถึง กลุ่มผู้สอบที่คาดว่าจะได้ประโยชน์จากข้อสอบที่ลำเอียง คือ เป็นกลุ่มที่มีความน่าจะเป็นในการตอบข้อสอบที่ลำเอียงได้ถูกต้องต่ำกว่าผู้สอบกลุ่มอ้างอิงที่มีคะแนนจากแบบทดสอบเท่ากัน สำหรับกลุ่มเปรียบเทียบในการวิจัยครั้งนี้ คือ กลุ่มผู้สอบที่ไม่ใช้ภาษาไทยในชีวิตประจำวัน (หมายถึง กลุ่มผู้สอบที่ใช้ภาษามลายูท้องถิ่น)
12. ผลการวิเคราะห์ความลำเอียงของข้อสอบ หมายถึง ผลที่ได้จากการคำนวณด้วยวิธีวิเคราะห์ต่างกัน ซึ่งได้แก่ วิธีแปลงค่าความยากของข้อสอบ วิธีโค้งลักษณะข้อสอบ 3 พารามิเตอร์และวิธีของแมนเทล-แฮนส์เซล ในการวิจัยครั้งนี้ หมายถึงผลดังนี้

12.1 จำนวนข้อสอบที่มีความลำเอียง หมายถึง จำนวนข้อสอบในแบบทดสอบที่ถูกตัดสินว่ามีความลำเอียงระหว่างกลุ่มผู้สอบที่ใช้ภาษาไทยในชีวิตประจำวันและกลุ่มผู้สอบที่ไม่ใช้ภาษาไทยในชีวิตประจำวัน ซึ่งวัดได้จากแบบทดสอบที่ผู้วิจัยสร้างขึ้น และทำการวิเคราะห์ความลำเอียงด้วยวิธีต่างกัน

12.2 ค่าความเชื่อมั่น (Reliability) ของแบบทดสอบ หมายถึง คุณภาพของแบบทดสอบ ซึ่งพิจารณาได้จากความสอดคล้องภายใน ซึ่งในการวิจัยครั้งนี้จะประมาณค่าความเชื่อมั่นแบบสัมประสิทธิ์ความสอดคล้องภายใน (Coefficient of Internal Consistency) ของแบบทดสอบ หลังจากคัดเลือกข้อที่มีความลำเอียงออกแล้ว โดยใช้สูตรคูเดอริชาร์ดสัน 20 (Kuder-Richardson 20) และทำการปรับขยายแบบทดสอบให้เต็มฉบับ โดยใช้สูตรของสเปียร์แมน-บราว (Spearman-Brown)

12.3 ค่าฟังก์ชันสารสนเทศของแบบทดสอบ (Test Information Function : TIF) หมายถึง ผลรวมของฟังก์ชันสารสนเทศของข้อสอบทุก ๆ ข้อในแบบทดสอบ ณ ระดับความสามารถของผู้สอบต่าง ๆ ถ้ามีค่าสูงที่ระดับความสามารถใด แสดงว่าแบบทดสอบนั้นสามารถจำแนกความสามารถของผู้สอบได้ดี ณ ระดับความสามารถนั้น