

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

ในการวิจัยครั้งนี้ผู้วิจัยได้เสนอเอกสารและงานวิจัยที่เกี่ยวข้องไว้ตามลำดับดังนี้

2.1 ทฤษฎีการวัดทางการศึกษา

2.1.1 ทฤษฎีการวัดแบบดั้งเดิม (Classical Test Theory)

2.1.2 ทฤษฎีการตอบสนองข้อสอบ (Item Response Theory)

2.2 ความหมายของความลำเอียงของข้อสอบ

2.3 วิธีการวิเคราะห์ความลำเอียงของข้อสอบ

2.4 โครงสร้างหลักสูตรกลุ่มวิชาสร้างเสริมประสบการณ์ชีวิต

2.5 งานวิจัยที่เกี่ยวข้อง

2.1 ทฤษฎีการวัดทางการศึกษา

ทฤษฎีการวัดทางการศึกษาในปัจจุบันที่มีบทบาทในวงการศึกษากว้าง ๆ ไป ประกอบด้วย 2 ทฤษฎีที่สำคัญคือ ทฤษฎีการวัดแบบดั้งเดิม (Classical Test Theory) และ ทฤษฎีการตอบสนองข้อสอบ (Item Response Theory)

2.1.1 ทฤษฎีการวัดแบบดั้งเดิม (Classical Test Theory : CTT) เป็นทฤษฎีพื้นฐานทฤษฎีแรกและใช้กันมานานเป็นที่รู้จักและเข้าใจกันดี โดยมีลักษณะดังนี้ (สัมพันธ์ พันธุ์พฤษย์, 2536)

1. ข้อตกลงเบื้องต้น (Assumptions)

ข้อตกลงเบื้องต้นของทฤษฎีการวัดแบบดั้งเดิม (CTT) เป็นข้อตกลงเบื้องต้นที่อ่อน (Weak assumptions) ซึ่งมีดังนี้คือ

1.1 คะแนนที่สังเกตได้ (Observed score) หรือคะแนนที่ได้จากการวัดของผู้สอบแต่ละคน เป็นผลรวมของคะแนนจริง (True score) และคะแนนความคลาดเคลื่อน (Error score) ของผู้สอบนั้น

$$X = T + E$$

1.2 คะแนนจริงเป็นค่าคาดหวังหรือค่าเฉลี่ยของคะแนนที่สังเกตได้

$$E(X) = T$$

หรือคะแนนจริง T เป็นค่าเฉลี่ยของการแจกแจงตามทฤษฎีของคะแนนสังเกตได้ (X) ที่บุคคลนั้นทำการสอบแบบสอบฉบับหนึ่งหลาย ๆ ครั้ง โดยการสอบแต่ละครั้งเป็นอิสระต่อกัน

1.3 คะแนนความคลาดเคลื่อน และคะแนนจริงที่ประชากรผู้สอบทำแบบสอบฉบับหนึ่งไม่มีความสัมพันธ์กัน

$$\rho_{ET} = 0$$

ข้อตกลงนี้อาจถูกฝ่าฝืนได้ ถ้าผู้สอบมีคะแนนจริงต่ำ ลอกคำตอบจากผู้สอบที่มีคะแนนจริงสูง หรือในการเรียนการสอนในชั้นเรียน ถ้าครูลำเอียงให้เด็กอ่อนมานั่งสอบข้างหน้าและให้เด็กเก่งไปนั่งสอบข้างหลังห้องที่อาจได้ยินคำสั่งของครูไม่ชัดเจน ในสถานการณ์อย่างนี้ก็จะส่งผลให้คะแนนความคลาดเคลื่อน มีความสัมพันธ์ทางลบกับคะแนนจริง

1.4 คะแนนความคลาดเคลื่อนในการทำแบบสอบฉบับหนึ่ง (E_1) และคะแนนความคลาดเคลื่อนในการทำแบบสอบอีกฉบับหนึ่ง (E_2) ไม่มีความสัมพันธ์กัน

$$\rho_{E_1 E_2} = 0$$

คะแนนความคลาดเคลื่อนของแบบสอบสองฉบับที่แตกต่างกันไม่มีความสัมพันธ์กัน กล่าวคือ ผู้สอบคนหนึ่งมีคะแนนความคลาดเคลื่อนเป็นทางบวกในการทำแบบสอบฉบับหนึ่ง และอาจจะมีคะแนนความคลาดเคลื่อนในทางบวกหรือทางลบในการทำแบบสอบอีกฉบับหนึ่งก็ได้

1.5 คะแนนความคลาดเคลื่อนจากแบบสอบฉบับที่หนึ่ง (E_1) ไม่มีความสัมพันธ์กับคะแนนจริงในแบบสอบฉบับที่สอง (T_2)

$$\rho_{E_1 T_2} = 0$$

ข้อตกลงเบื้องต้นนี้อาจถูกฝ่าฝืนได้ ถ้าแบบสอบฉบับที่สองวัดคุณลักษณะด้านบุคลิกภาพหรือมิติความสามารถที่มีอิทธิพลต่อการเกิดความคลาดเคลื่อนในแบบสอบฉบับที่หนึ่ง

1.6 แบบสอบฉบับ A และ B เป็นแบบสอบคู่ขนาน (Parallel tests) ถ้าคะแนนที่สังเกตได้จากแบบสอบ A และ B เป็น X และ X' ที่เป็นไปตามข้อตกลงเบื้องต้นของ 1.1 - 1.5 เมื่อ $X_1 = T + E$ และ $X' = T' + E'$ แล้วสำหรับประชากรผู้สอบทุกคนจะได้ $T = T'$ และ $\sigma^2_E = \sigma^2_{E'}$

1.7 แบบสอบฉบับ A และ B เป็น Essentially $\mathcal{J}$ -equivalent test ถ้า

คะแนนที่สังเกตได้จากแบบสอบ A และ B เป็น X_1 X_2 ที่เป็นไปตามข้อตกลงเบื้องต้นข้อที่ 1.1 - 1.5 และเมื่อ $X_1 = T_1 + E_1$ และ $X_2 = T_2 + E_2$ แล้วสำหรับประชากรทุกคน จะได้ $T_1 = T_2 + C_{12}$ เมื่อ C_{12} เป็นค่าคงที่

โมเดลของทฤษฎีการวัดแบบดั้งเดิม

การประมาณค่าความเที่ยง ตามทฤษฎีการวัดแบบดั้งเดิม (CTT) ต้องใช้โมเดลของทฤษฎีการทดสอบดั้งเดิมซึ่งมีลักษณะดังนี้ (สัมพันธ์ พันธุ์พฤกษ์, 2536)

1. เป็น Additive Model

$$X = T + E$$

2. เป็น Simple Model ถือว่าความคลาดเคลื่อนที่เกิดโดยสุ่มมาจากแหล่งเดียว (Single random error) และอธิบายถึงความคลาดเคลื่อนของการวัดที่มีอิทธิพลต่อคะแนนสังเกต

3. ไม่สามารถทดสอบโมเดลของ ทฤษฎีการวัดแบบดั้งเดิมได้เนื่องจากไม่สามารถหาคะแนนจริงได้

4. โมเดลอยู่ภายใต้ข้อตกลงเบื้องต้นที่มีการกำหนดเงื่อนไขบางอย่างให้เป็นจริง ดังนั้นข้อสรุปที่ได้จากโมเดลจะสมเหตุสมผล เพียงใดขึ้นอยู่กับความสมเหตุสมผลของข้อตกลงเบื้องต้น

ข้อพึงระวังในการประเมินค่าความเที่ยงตามทฤษฎีการวัดแบบดั้งเดิม (CTT) ภายใต้ Weak Assumptions ของ CTT ส่งผลให้การประมาณค่าความเที่ยงมีความคลาดเคลื่อนดังนี้

1. จากการกำหนดให้แหล่งความคลาดเคลื่อนมีเพียงแหล่งเดียว โดยการรวมความคลาดเคลื่อนจากแหล่งต่าง ๆ ทั้งหมดรวมเป็นแหล่งเดียว (Single error Component) จึงทำให้การประมาณค่าความเที่ยงโดย Spearman-Brown มีค่าสูงกว่าที่ควรจะเป็น และถ้าใช้ Kuder Richardson 21 หรือ Cronbach Alpha's Coefficient จะได้ค่าต่ำกว่าที่ควรจะเป็น

2. ค่าสถิติที่อธิบายลักษณะของข้อกระทบ (Item statistics) อาทิ เช่น ค่าความยากจะแปรผันตามลักษณะของกลุ่มผู้สอบ ดังนั้น ถ้านำแบบสอบไปใช้กับกลุ่มผู้สอบมีระดับความสามารถไม่เหมือนกับกลุ่มตัวอย่างที่ใช้ในการวิเคราะห์ข้อสอบแล้ว จะต้องทำการประมาณค่าความเที่ยงใหม่ เนื่องจากค่าความเที่ยงมีค่าน้อยเพียงใดขึ้นอยู่กับลักษณะการกระจายของคะแนนผู้สอบ ความยาวแบบสอบ ค่าความยาก และค่าอำนาจจำแนก

3. การประมาณค่าความเที่ยงตามทฤษฎีการวัดแบบดั้งเดิม (CTT) ยึด Parallelism assumptions ซึ่งในทางปฏิบัติจะให้ $\sigma_{\epsilon}^2 = \sigma_{\epsilon}^2$ เป็นไปได้ยาก จึงทำให้เกิดความคลาดเคลื่อนในการประมาณ

4. จากการกำหนดความคลาดเคลื่อนจากการวัดของแต่ละคนมีค่าเท่ากัน กล่าวคือ ความคลาดเคลื่อนมาตรฐานของการวัด $= \sigma_x \sqrt{1 - \rho_{xx'}}$ ซึ่งในความเป็นจริงความคลาดเคลื่อนของการทำข้อสอบจะขึ้นอยู่กับระดับความสามารถของผู้สอบแต่ละคน

5. ในการประมาณความเที่ยงวิเคราะห์จากตัวประกอบตัวเดียว ซึ่งได้แก่ ข้อกระทง (Item) ไม่พิจารณาตัวประกอบอื่น จึงทำให้ค่าที่ประมาณได้มีความคลาดเคลื่อนและการนำผลที่ได้ไปสรุปอ้างอิงจึงไม่กว้างขวาง

2.1.2 ทฤษฎีการตอบสนองข้อสอบ (Item Response Theory : IRT)

ความเป็นมาของทฤษฎีการตอบสนองข้อสอบ

Ferguson and Lawley (1942, 1943) เป็นผู้ริเริ่มทฤษฎีการตอบสนองข้อสอบ โดยมีหลักการว่า ผลการสอบของผู้สอบจากข้อสอบใด ๆ ขึ้นอยู่กับความสามารถ (Ability or Skill) ของผู้สอบ ต่อมาในปี ค.ศ. 1952 Lord ได้เสนอทฤษฎีใหม่ในรูปโค้งลักษณะข้อสอบ (Item Characteristic Curve : ICC) โดย Lord ได้เสนอว่า โค้งลักษณะข้อสอบมีลักษณะเป็นโค้งปกติสะสม ซึ่งเรียกว่า Normal Ogive Model ซึ่งโมเดลนี้จะกล่าวถึงพารามิเตอร์ 2 ตัว คือ ค่าความยากและค่าอำนาจจำแนก แต่ทฤษฎีนี้ในระยะแรกพัฒนาไปช้ามาก เนื่องจากความยุ่งยากในการคำนวณและขาดแคลนเครื่องคอมพิวเตอร์ที่ใช้ในการวิเคราะห์ข้อมูล จึงทำให้ทฤษฎีนี้ต้องใช้เวลาในการศึกษาและพัฒนาอยู่ระยะหนึ่ง ในปี ค.ศ. 1960 Rasch ได้ศึกษาและเสนอแนวคิดในรูปของพารามิเตอร์ตัวเดียวคือ ค่าความยาก (b) หรือที่เรียกกันว่า Rasch Model และในปี ค.ศ. 1965 Lord ได้กลับมาสนใจและพัฒนาทฤษฎีการตอบสนองข้อสอบใหม่ ในปี ค.ศ. 1968 Birnbaum ได้เสนอ Logistic Model ที่ใช้พารามิเตอร์ 2 ตัว คือ ค่าความยาก และ ค่าอำนาจจำแนก ซึ่งเป็นโมเดลที่ง่ายกว่าของ Lord จึงทำให้เป็นที่นิยมแพร่หลายและมีการพัฒนาขึ้นเรื่อย ๆ จนกระทั่งสามารถใช้ได้กับ พารามิเตอร์ 1 ตัว และพารามิเตอร์ 3 ตัว (Warm, 1979 อ้างถึงใน สุรศักดิ์ อมรรัตนศักดิ์, 2530) และต่อมาได้มีการพัฒนา Logistic Model ใช้กับพารามิเตอร์ 4 ตัว คือ ค่าความยาก (b) ค่าอำนาจจำแนก (a) ค่าการเดา (c) และค่าความรอบคอบ (r) (Hambleton and

Swaminathan, 1985 อ้างถึงใน เจน จันทรแสง, 2537)

ข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองข้อสอบ

ทฤษฎีการตอบสนองข้อสอบ ((Item Response Theory) มีข้อตกลงเบื้องต้นดังนี้ (สัมพันธ์ พันธุ์พฤษ, 2536)

1. ความเป็นเอกมิติ (Unidimensional) โมเดลการตอบสนองรายข้อได้กำหนดให้ต้องวัดความสามารถที่เป็นคุณลักษณะภายในที่สำคัญเพียงลักษณะเดียว (Single latent ability)
2. ความเป็นอิสระ (Local independent) การตอบข้อกระทงแต่ละข้อในแบบสอบฉบับหนึ่งของผู้ตอบแต่ละคนมีความเป็นอิสระต่อกัน กล่าวคือ ข้อกระทงแต่ละข้อเป็นอิสระต่อกัน และไม่ว่าจะให้ข้อกระทงแต่ละข้ออยู่ตำแหน่งใด ก็ไม่ส่งผลต่อการตอบของผู้สอบ และรวมทั้งผู้สอบแต่ละคนจะตอบข้อกระทงแต่ละข้ออย่างเป็นอิสระต่อกัน
3. โคงลักษณะข้อกระทง (Item Characteristic Curve : ICC) แต่ละโคงเป็นฟังก์ชันทางคณิตศาสตร์ ที่แสดงความสัมพันธ์ระหว่างระดับความสามารถของผู้ตอบ (θ) และโอกาสที่จะตอบข้อนั้นถูก $P_i(\theta)$
4. การบริหารการสอบ ไม่อยู่ในเงื่อนไขของการจำกัดเวลา (Speededness) กล่าวคือ ผู้สอบทุกคนมีโอกาสทำข้อกระทงทุกข้อไม่ใช่เป็นการสอบแบบจำกัดเวลา

ลักษณะเด่นของทฤษฎีการตอบสนองข้อสอบ

ทฤษฎีการตอบสนองข้อสอบมีความเหนือกว่า (Superiority) ทฤษฎีการวัดแบบดั้งเดิม ดังนี้ (สุรศักดิ์ อมรรตสันศักดิ์, 2530 เจน จันทรแสง, 2537)

1. ความไม่แปรเปลี่ยนของค่าพารามิเตอร์ (Invariant of Item Parameter) กล่าวคือ ไม่ว่าจะเป็น การประมาณค่าคุณสมบัติของข้อสอบจากกลุ่มตัวอย่างที่มีความสามารถระดับใดก็ตาม ค่าพารามิเตอร์จะไม่แปรเปลี่ยนไปตามกลุ่มของผู้สอบ
2. ความเป็นอิสระจากกลุ่ม (Person-Free) นั่นคือไม่ว่าจะนำข้อสอบไปใช้สอบกับบุคคลกลุ่มใด โคงลักษณะข้อสอบก็จะคงเดิม
3. ความเป็นอิสระของข้อ (Item-Free) ในการประมาณค่าความสามารถแท้ (θ) ของผู้สอบ จะใช้ข้อสอบชุดใด จำนวนข้อเท่าใดก็ได้ บางครั้งอาจใช้ข้อสอบเพียง 3-5 ข้อ ก็สามารถประมาณค่าความสามารถแท้ของผู้สอบได้

ข้อควรคำนึงในการใช้ทฤษฎีการตอบสนองข้อสอบ (IRT)

เนื่องจากข้อตกลงเบื้องต้นของ IRT เป็น Strong assumption ดังนั้นก่อนใช้ทฤษฎีการตอบสนองข้อสอบ (IRT) ต้องพิจารณาในเรื่องต่อไปนี้ (สัมพันธ์ พันธุ์พฤษ, 2536)

1. คำนึงถึงวัตถุประสงค์ของการศึกษา ซึ่งอาจเป็นวัตถุประสงค์ของการพัฒนาข้อสอบ หรือการวิเคราะห์ข้อมูลที่ได้จากข้อสอบ
2. ทดสอบว่าข้อสอบนั้นเป็น Unidimensional ตามข้อตกลงเบื้องต้นหรือไม่
3. ตรวจสอบข้อมูลที่จะศึกษาว่าเหมาะสมกับโมเดลใดได้ดีที่สุด
4. ต้องศึกษาจากกลุ่มตัวอย่างที่มีขนาดใหญ่พอต่อการหาพารามิเตอร์ของข้อกระทง และพารามิเตอร์ความสามารถ
5. เนื่องจากการใช้ IRT ต้องศึกษาจากกลุ่มตัวอย่างขนาดใหญ่ และใช้คอมพิวเตอร์ จึงควรพิจารณาว่า งานนั้นเหมาะสมกับการใช้เพียงใด
6. แม้ว่าทฤษฎีการตอบสนองข้อสอบ (IRT) จะศึกษาทั้งลักษณะข้อกระทง (Items) และลักษณะของผู้สอบหรือระดับความสามารถ (θ) ก็ตามแต่ถ้าในกรณีต้องการศึกษาในลักษณะ Multifaceted Universes การใช้ IRT จะไม่ได้ประสิทธิภาพมากนักต้องใช้ทฤษฎีการสรุปอ้างอิงจึงจะเหมาะสม

2.2 ความหมายของความลำเอียงของข้อสอบ

เกี่ยวกับเรื่องความหมายของความลำเอียงของข้อสอบนี้มีผู้ให้ความหมายไว้มากมาย ซึ่งผู้วิจัยรวบรวมมาได้ดังนี้

Allen and Yen (1979) ได้ให้ความหมายของข้อสอบที่มีความลำเอียงว่า หมายถึงข้อสอบที่วัดความสามารถหรือคุณลักษณะทางจิตวิทยาของผู้สอบแต่ละกลุ่มได้ไม่ตรงกัน เช่น เมื่อนำข้อสอบวัดความสามารถด้านคณิตศาสตร์ที่ใช้ภาษาสื่อความหมายของโจทย์ปัญหาไปทดสอบกับผู้สอบ 2 กลุ่มที่มีความสามารถทางการอ่านไม่เท่ากัน ข้อสอบก็จะมี ความลำเอียงต่อผู้สอบกลุ่มที่มีความสามารถทางการอ่านสูง ข้อสอบเหล่านี้จะมีโครงสร้างข้อสอบของผู้สอบแต่ละกลุ่มแตกต่างกัน

ซัชชัย เผ่าพงศ์ (2527) ได้ให้ความหมายของข้อสอบที่มีความลำเอียงว่า หมายถึงข้อสอบที่มีการนำไปทดสอบกับกลุ่มที่มีความแตกต่างกันทางด้านเพศ เชื้อชาติ ศาสนา สังคม วัฒนธรรม สภาพภูมิศาสตร์ และเศรษฐกิจ โดยสภาพการณ์ต่าง ๆ ของกลุ่มเหล่านี้มีอิทธิพลต่อการพัฒนา สร้างเสริมประสบการณ์และการเรียนรู้ของบุคคลในแต่ละกลุ่มให้มีความแตกต่างกัน

และทำให้บุคคลในแต่ละกลุ่มมีการได้เปรียบเสียเปรียบจากการสอบข้อสอบข้อเดียวกันไม่เท่ากัน

Cleary and Hilton (1968, อ้างถึงใน สุรศักดิ์ อมรรตน์ศักดิ์, 2530) ได้ให้ความหมายของข้อสอบที่มีความลำเอียงว่าหมายถึงข้อสอบที่เมื่อนำไปทดสอบกับบุคคลกลุ่มใดกลุ่มหนึ่ง แล้วปรากฏว่า มีคะแนนเฉลี่ยของข้อสอบข้อเดียวกันแตกต่างจากคะแนนเฉลี่ยที่ได้จากบุคคลในกลุ่มอื่น ๆ ซึ่งความแตกต่างของคะแนนเฉลี่ยนี้อาจจะมากกว่าหรือน้อยกว่าที่คาดไว้จากข้อสอบข้ออื่น ๆ ในแบบสอบฉบับเดียวกันก็ได้

Lord (1980, อ้างถึงใน พชร บิณฑิต, 2531) ได้ให้ความหมายของข้อสอบที่มีความลำเอียงว่า เป็นข้อสอบที่ผู้สอบที่มีความสามารถเท่ากันแต่อยู่ต่างกลุ่มกันมีโอกาสที่จะตอบข้อสอบถูกไม่เท่ากัน

Merhrens and Lehmann (1978, อ้างถึงใน เจน จันทรแสง, 2537) ได้ให้ความหมายของข้อสอบที่มีความลำเอียงทางวัฒนธรรม (Cultural Bias) ว่าหมายถึงข้อสอบที่ทำให้ผู้สอบแต่ละกลุ่มมีคะแนนเฉลี่ยแตกต่างกัน หรือเป็นข้อสอบที่วัดคุณลักษณะทางจิตวิทยาหรือผลสัมฤทธิ์ในเนื้อหาเดียวกันไม่ตรงกัน และให้ความหมายของข้อสอบที่มีความลำเอียงระหว่างเพศ ว่าหมายถึง ข้อสอบที่มีเนื้อหาลำเอียงต่อเพศใดเพศหนึ่งโดยเฉพาะ และเพศ มีผลต่อการทำข้อสอบแต่ละข้อ

บุญชม ศรีสะอาด (2535) ได้ให้ความหมายของความลำเอียงของข้อสอบว่า หมายถึง ข้อสอบที่มองค้ประกอบบางอย่างที่ทำให้ผู้สอบคนหนึ่ง หรือกลุ่มหนึ่งทำข้อสอบได้ดีกว่าคนอื่นหรือกลุ่มอื่นที่มีความสามารถไม่แตกต่างกัน ตัวแปรที่ทำให้เกิดความลำเอียงได้แก่ เพศ เชื้อชาติ ศาสนา สภาพภูมิศาสตร์ สภาพเศรษฐกิจและสังคม ตลอดจนประสบการณ์ส่วนตัว เป็นต้น

นิรมล ชัยชลิต (2537) ได้ให้ความหมายของความลำเอียงของข้อสอบว่า หมายถึง ข้อสอบที่เมื่อนำไปสอบกับกลุ่มผู้สอบใด ๆ ที่มีความสามารถเท่ากัน แต่มีลักษณะของกลุ่มต่างกันจะปรากฏว่ามีคะแนนเฉลี่ยของข้อสอบข้อนั้นไม่เท่ากัน หรือ เป็นข้อสอบที่วัดความสามารถ หรือ คุณลักษณะทางจิตวิทยาของผู้สอบแต่ละกลุ่มได้ไม่ตรงกัน

เจน จันทรแสง (2537) ได้ให้ความหมายของข้อสอบที่มีความลำเอียงว่าเป็นข้อสอบที่นำไปสอบกับกลุ่มผู้สอบต่างกลุ่มกัน แล้วทำให้บางกลุ่มมีการได้เปรียบและเสียเปรียบในการทำข้อสอบข้อเดียวกัน

จากความหมายข้างต้น พอสรุปได้ว่า ข้อสอบที่มีความลำเอียง หมายถึง ข้อสอบที่นำไปสอบกับกลุ่มนักเรียนหรือกลุ่มผู้เข้าสอบที่มีความสามารถเท่ากัน แต่มีความแตกต่างกันทางด้าน เพศ เชื้อชาติ ศาสนา วัฒนธรรม สถานที่ตั้งของสถานศึกษา สภาพทางภูมิศาสตร์และเศรษฐกิจ

แล้วทำให้นักเรียนหรือกลุ่มผู้เข้าสอบในแต่ละกลุ่มมีการได้เปรียบเสียเปรียบกัน จากการตอบข้อสอบข้อเดียวกัน

2.3 วิธีการวิเคราะห์ความลำเอียงของข้อสอบ

การวิเคราะห์ความลำเอียงของข้อสอบแบ่งเป็น 2 วิธี คือ การวิเคราะห์ความลำเอียงของข้อสอบโดยใช้เกณฑ์ภายนอก และใช้เกณฑ์ภายใน โดยมีหลักและวิธีการดังนี้ (Runder, Getson, and Knight 1980, Plake 1981 อ้างถึงใน ชัชชัย เผ่าพงศ์, 2527, สุรศักดิ์ อมรรัตนศักดิ์, 2530)

2.3.1 การวิเคราะห์ความลำเอียงของข้อสอบโดยใช้เกณฑ์ภายนอก

(External Methods Item Bias)

การวิเคราะห์ความลำเอียงของข้อสอบโดยวิธีนี้ สามารถวิเคราะห์ได้ทั้งรายฉบับและรายข้อ โดยการเขียนกราฟแสดงความสัมพันธ์ระหว่างคะแนนที่ใช้เป็นเกณฑ์ภายนอกกับตัวแปรทำนาย เพื่อวิเคราะห์ความถดถอยของตัวแปรทั้งสองแล้วทำการเปรียบเทียบค่าความชัน (Slope) และค่าตัดแกน (Intercept) ของเส้นกราฟ ในการวิเคราะห์ความลำเอียงของข้อสอบทั้งฉบับจะใช้คะแนนรวมเป็นตัวแปรทำนาย แต่ถ้าเป็นการวิเคราะห์รายข้อจะใช้ค่าความยาก (P) ของข้อสอบแต่ละข้อเป็นตัวแปรทำนาย ส่วนตัวแปรเกณฑ์ภายนอกจะใช้คะแนนรวมหรือเกรด ซึ่งเป็นผลสัมฤทธิ์ของผู้สอบเกี่ยวกับงานบางอย่างที่ทำภายในโรงเรียน ฉะนั้นการวิเคราะห์ความลำเอียงของข้อสอบด้วยวิธีนี้จึงเป็นการเปรียบเทียบเส้นกราฟระหว่างกลุ่มผู้สอบ โดยถ้าเส้นกราฟดังกล่าวมีค่าความชันและค่าตัดแกนแตกต่างกันในแต่ละกลุ่ม ข้อสอบนั้นก็จะมีความลำเอียงต่อกลุ่มผู้สอบที่มีค่าตัดแกน และค่าความชันมากกว่า จะเห็นได้ว่าการวิเคราะห์ความลำเอียงของข้อสอบตามวิธีการนี้จึงมีจุดอ่อนตรงที่ในทางปฏิบัติแล้วเป็นการยากมากที่จะหาตัวแปรเกณฑ์ภายนอกที่มีความตรงเชิงพยากรณ์ และมีความยุติธรรม ซึ่งถ้าหากตัวแปรเกณฑ์ภายนอกขาดคุณสมบัติดังกล่าว จะทำให้ผลการวิเคราะห์ความลำเอียงของข้อสอบขาดความสมบูรณ์ และไม่มีความถูกต้องเท่าที่ควร การวิเคราะห์ความลำเอียงของข้อสอบวิธีนี้เหมาะสำหรับจะนำมาใช้กับการสอบคัดเลือกบุคคลเข้าทำงานหรือศึกษาต่อในระดับสูง จากความยุ่งยากในการสร้างเกณฑ์ภายนอก นักวัดผลจึงหันมาตรวจสอบโครงสร้างภายในของข้อสอบ วิธีการนี้จะใช้ข้อมูลผลการตอบคำถามในข้อสอบเท่านั้น จึงเรียกวิธีการนี้ว่า การวิเคราะห์ความลำเอียงของข้อสอบโดยใช้เกณฑ์ภายใน

2.3.2 การวิเคราะห์ความลำเอียงโดยใช้เกณฑ์ภายใน

(Internal Methods Item Bias)

การวิเคราะห์ความลำเอียงของข้อสอบวิธีนี้ยึดหลักการที่ว่าข้อสอบจะมีความยุติธรรมโดยไม่ลำเอียงต่อกลุ่มผู้สอบก็ต่อเมื่อผู้สอบที่อยู่ต่างกลุ่มกัน มีความสามารถเท่ากันแล้วจะมีคะแนนจริงของผลการสอบเท่ากัน การวิเคราะห์ความลำเอียงโดยใช้เกณฑ์ภายในมีหลายวิธีได้แก่

2.3.2.1 วิธีแปลงค่าความยากของข้อสอบ (Transformed Item Difficulty:TID)

วิธีนี้จะคำนวณหาค่าความยาก (P) ของข้อสอบแต่ละข้อจากผู้สอบแต่ละกลุ่ม แล้วแปลงค่าความยาก (P) ไปเป็นคะแนนมาตรฐานซี (Z) และแปลงค่า Z ให้เป็นค่าความยากมาตรฐานเคลต้า (Δ) โดยใช้สมการ $\Delta = 4z + 13$ จากนั้นคำนวณค่าระยะห่างตั้งฉากจากจุดศูนย์กลางของค่าเคลต้าของข้อสอบแต่ละข้อไปยังเส้นแกนหลัก จากสูตรดังนี้ (Angoff, 1982 : 96-112)

$$d_i = \frac{ax_i - y_i + b}{\sqrt{a^2 + 1}}$$

$$a = \frac{(s_y^2 - s_x^2) \pm \sqrt{(s_y^2 - s_x^2)^2 + 4r_{xy}^2 s_x^2 s_y^2}}{2r_{xy} s_x s_y}$$

$$b = M_y - aM_x$$

เมื่อ d_i = ค่าดัชนีความลำเอียงซึ่งก็คือ ค่าระยะห่างตั้งฉากจากจุดศูนย์กลางของเคลต้าของข้อสอบข้อที่ i ไปยังเส้นแกนหลัก

a = ค่าความชันเส้นกราฟ

b = ค่าที่ตัดแกน y

x_i, y_i = ค่าเคลต้าของข้อสอบข้อที่ i จากกลุ่มที่ 1 และ 2

M_x, M_y = ค่าเฉลี่ยของค่าเคลต้า กลุ่มที่ 1 และ 2

s_x, s_y = ค่าเบี่ยงเบนมาตรฐานของค่าเคลต้า กลุ่มที่ 1 และ 2

r_{xy} = ค่าสหสัมพันธ์ระหว่างค่าเคลต้า ของกลุ่มที่ 1 และ 2

แล้วคำนวณหาค่าเบี่ยงเบนมาตรฐานของระยะห่างตั้งฉาก จากสูตร $s_d = \sqrt{1 - r_{xy}}$ ข้อสอบข้อใดที่มีระยะห่างจากเส้นแกนหลักมากกว่าหรือเท่ากับ $\pm 1.96s_d$ ถือว่าเป็นข้อสอบที่มีความลำเอียง

2.3.2.2 วิธีวิเคราะห์ความลำเอียงโดยใช้ไค-สแควร์ (Chi-Square : CHI) มีขั้นตอนดังนี้

1. เรียงคะแนนจากผู้ที่ได้คะแนนรวมน้อยที่สุด ไปหาผู้ที่ได้คะแนนรวมมากที่สุด
2. แบ่งคะแนนออกเป็นช่วงโดยมีหลักการคือ ความถี่ของการตอบข้อสอบถูกแต่ละกลุ่มมีค่าน้อย 10 และความถี่คาดหวังมีค่าน้อยที่สุด 5 และในแต่ละช่วงกว้างของคะแนนในแต่ละชั้นมีกลุ่มตัวอย่างใกล้เคียงกัน ซึ่งในการวิจัยครั้งนี้ผู้วิจัยแบ่งเป็น 5 ชั้นคะแนน ซึ่งในแต่ละชั้นคะแนนนับความถี่ของการตอบถูก (O_j) ในแต่ละข้อ

3. คำนวณหาค่าสัดส่วนการตอบถูก (P_j) จากสูตรดังนี้

$$P_j = \frac{O_j}{n_j}$$

เมื่อ O_j แทน ความถี่ที่สังเกตได้ในแต่ละชั้นคะแนน

n_j แทน จำนวนนักเรียนรวมในแต่ละชั้น

4. คำนวณความถี่คาดหวัง (E_{ij}) จากสูตร

$$E_{ij} = P_j n_{ij}$$

เมื่อ P_j แทนสัดส่วนการตอบถูกของกลุ่มตัวอย่างทั้งหมดในแต่ละชั้นคะแนน

n_{ij} แทนจำนวนนักเรียนรวมในแต่ละกลุ่มและแต่ละชั้นคะแนน

5. และคำนวณหาดัชนีความลำเอียงของข้อสอบ จากสูตรดังนี้ (Ironson, 1982:126)

$$\chi^2 = \sum_{j=1}^J \left[\frac{(E_{1j} - O_{1j})^2}{E_{1j}} \right] + \sum_{j=1}^J \left[\frac{(E_{2j} - O_{2j})^2}{E_{2j}} \right]; df = (I-1)(J-1)$$

เมื่อ χ^2 = ค่าดัชนีความลำเอียงของข้อสอบ

O_{1j}, O_{2j} = ความถี่ที่สังเกตได้ของการตอบข้อสอบถูกของกลุ่มที่ 1 และ 2 ในแต่ละชั้นคะแนน ตามลำดับ

E_{1j}, E_{2j} = ความถี่ที่คาดหวังของการตอบข้อสอบถูกของกลุ่มที่ 1 และ 2 ในแต่ละชั้นคะแนน ตามลำดับ

I = จำนวนกลุ่มตัวอย่างที่ทำการวิเคราะห์

J = จำนวนชั้นคะแนน ซึ่งเท่ากับ 5 ชั้นคะแนน

2.3.2.2 วิเคราะห์ความลำเอียงโดยใช้ไค-สแควร์ (Chi-Square : CHI) มีขั้นตอนดังนี้

1. เรียงคะแนนจากผู้ที่ได้คะแนนรวมน้อยที่สุด ไปหาผู้ที่ได้คะแนนรวมมากที่สุด
2. แบ่งคะแนนออกเป็นช่วงโดยมีหลักการคือ ความถี่ของการตอบข้อสอบถูกแต่ละกลุ่มมีค่าน้อย 10 และความถี่คาดหวังมีค่าน้อยที่สุด 5 และในแต่ละช่วงกว้างของคะแนนในแต่ละชั้นมีกลุ่มตัวอย่างใกล้เคียงกัน ซึ่งในการวิจัยครั้งนี้ผู้วิจัยแบ่งเป็น 5 ชั้นคะแนน ซึ่งในแต่ละชั้นคะแนนนับความถี่ของการตอบถูก (O_j) ในแต่ละข้อ

3. คำนวณหาค่าสัดส่วนการตอบถูก (P_j) จากสูตรดังนี้

$$P_j = \frac{O_j}{n_j}$$

เมื่อ O_j แทน ความถี่ที่สังเกตได้ในแต่ละชั้นคะแนน

n_j แทน จำนวนนักเรียนรวมในแต่ละชั้น

4. คำนวณความถี่ที่คาดหวัง (E_{ij}) จากสูตร

$$E_{ij} = P_j n_{ij}$$

เมื่อ P_j แทนสัดส่วนการตอบถูกของกลุ่มตัวอย่างทั้งหมดในแต่ละชั้นคะแนน

n_{ij} แทนจำนวนนักเรียนรวมในแต่ละกลุ่มและแต่ละชั้นคะแนน

5. และคำนวณหาดัชนีความลำเอียงของข้อสอบ จากสูตรดังนี้ (Ironson, 1982:126)

$$\chi^2 = \sum_{j=1}^J \left[\frac{(E_{1j} - O_{1j})^2}{E_{1j}} \right] + \sum_{j=1}^J \left[\frac{(E_{2j} - O_{2j})^2}{E_{2j}} \right]; df = (I-1)(J-1)$$

เมื่อ χ^2 = ค่าดัชนีความลำเอียงของข้อสอบ

O_{1j}, O_{2j} = ความถี่ที่สังเกตได้ของการตอบข้อสอบถูกของกลุ่มที่ 1 และ 2 ในแต่ละชั้นคะแนน ตามลำดับ

E_{1j}, E_{2j} = ความถี่ที่คาดหวังของการตอบข้อสอบถูกของกลุ่มที่ 1 และ 2 ในแต่ละชั้นคะแนน ตามลำดับ

I = จำนวนกลุ่มตัวอย่างที่ทำการวิเคราะห์

J = จำนวนชั้นคะแนน ซึ่งเท่ากับ 5 ชั้นคะแนน

6. ถ้าค่าไค-สแควร์ (χ^2) ที่คำนวณได้มากกว่าค่าวิกฤติ ของไค-สแควร์ (χ^2) ในตารางแสดงว่าข้อสอบข้อนั้นมีความลำเอียง

2.3.2.3 วิธีวิเคราะห์ความแปรปรวน (Analysis of Variance : ANOVA)

การวิเคราะห์ความแปรปรวนสามารถนำมาประยุกต์ใช้ในการหาความลำเอียงของข้อสอบได้ โดยถือว่าข้อสอบที่ไม่วัดความสามารถเดียวกัน จากผู้สอบต่างกลุ่มกันเป็นข้อสอบที่มีความลำเอียง ซึ่งหมายความว่าถ้ามีข้อสอบที่ง่ายสำหรับคนกลุ่มหนึ่ง แต่ยากสำหรับคนอีกกลุ่มหนึ่งก็อาจตั้งข้อสงสัยได้ว่าข้อสอบนั้นเป็นข้อสอบที่มีความลำเอียง ซึ่งมีขั้นตอนการวิเคราะห์ดังนี้ (สุรศักดิ์ อมรรัตนศักดิ์, 2530 และ เจน จันทรแสง, 2537)

1. จัดผลการตอบแบบสอบ (Items) ลงในแนวสดมภ์ (column) และกลุ่มผู้สอบลงในแนวแถว (Row)
2. คำนวณค่าสัดส่วนในการตอบแบบสอบถูกแต่ละข้อระหว่างกลุ่ม 2 กลุ่ม
3. ทดสอบปฏิกริยาร่วม (Interaction Effect) ระหว่างข้อ (Items) กับกลุ่มผู้สอบ (Subgroups) หากพบว่ามีนัยสำคัญทางสถิติ แสดงว่าแบบสอบฉบับนั้นมีความลำเอียง
4. ถ้า Interaction Effect ระหว่างข้อกับกลุ่มผู้สอบมีนัยสำคัญทางสถิติให้ทดสอบความแตกต่างของโอกาสในการตอบถูกเป็นรายข้อ โดยใช้การทดสอบ t-test ดังนี้ (Walpole , 1978 อ้างถึงใน สัมพันธ์ พันธุ์พฤกษ์ , 2539)

4.1 ใช้การทดสอบที่ แบบ Pooled variance estimate

$$\text{กรณี } \sigma_1^2 = \sigma_2^2 \text{ หรือ } n_1 = n_2$$

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - d_0}{S_p \sqrt{(1/n_1 + 1/n_2)}} ; df = n_1 + n_2 - 2$$

$$S_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 2)s_2^2}{n_1 + n_2 - 2}$$

4.2 ใช้การทดสอบที แบบ Separate variance estimate

กรณี $\sigma_1^2 \neq \sigma_2^2$ หรือ $n_1 \neq n_2$

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - d_0}{\sqrt{(s_1^2/n_1 + s_2^2/n_2)}}$$

$$df = \frac{(s_1^2/n_1 + s_2^2/n_2)^2}{\frac{(s_1^2/n_1)^2}{n_1 - 1} + \frac{(s_2^2/n_2)^2}{n_2 - 1}}$$

5. ค่าที่คำนวณได้จากข้อ 4 คือค่าดัชนีความลำเอียงของข้อสอบ โดยใช้เกณฑ์ตัดสินคือ ถ้าหากค่า t ที่คำนวณได้มีค่ามากกว่า ค่าวิกฤติของ t ในตารางถือว่าข้อสอบลำเอียง

2.3.2.4 วิเคราะห์องค์ประกอบ (Factor Analysis) เป็นเทคนิคที่ช่วยให้ทราบ ว่าแต่ละข้อวัดองค์ประกอบเดียวกันมากน้อยเพียงใด โดยพิจารณาน้ำหนักตัวประกอบ (Factor Loading) ขององค์ประกอบนั้น ความแตกต่างกันของ Factor Loading จากผู้สอบต่างกลุ่มกันหรือความแตกต่างกันอย่างมีนัยสำคัญระหว่างค่าเฉลี่ยของ Factor score จากแต่ละกลุ่มที่นำมาเปรียบเทียบกัน เมื่อทดสอบความมีนัยสำคัญพบว่ามีความแตกต่างกันอย่างมีนัยสำคัญถือว่า ข้อสอบข้อนั้นมีค่าความลำเอียง

2.3.2.5 วิธีการให้ผู้เชี่ยวชาญพิจารณา (Judgmental methods) เป็นการนำแบบสอบทั้งฉบับไปให้ผู้เชี่ยวชาญด้านเนื้อหาและผู้เชี่ยวชาญด้านการสร้างแบบสอบเป็นผู้พิจารณา ข้อคำถามข้อความ สถานการณ์หรือองค์ประกอบบางอย่างที่ปรากฏอยู่ในข้อสอบที่จะเอื้อให้ผู้คล ที่มีระดับความสามารถเท่ากันแต่อยู่ต่างกลุ่มกัน มีโอกาสทำข้อสอบข้อนั้นถูกต้องด้วยความน่าจะเป็น ต่างกันหรือไม่อย่างไร ถ้าต่างกันกลุ่มใดจะได้เปรียบหรือเสียเปรียบ เมื่อผู้เชี่ยวชาญพิจารณา เสร็จแล้วก็นำคะแนนที่ได้จากความเห็นของผู้เชี่ยวชาญแต่ละคนมาหาค่าเฉลี่ย ข้อสอบที่มีความ ลำเอียงคือข้อสอบที่มีค่าเฉลี่ยมากกว่า .50 (ไชยยา อะการะวัง, 2536)

2.3.2.6 วิเคราะห์ด้วยโค้งลักษณะข้อสอบที่มีพารามิเตอร์ 1 ตัว วิธีนี้เป็นการ เปรียบเทียบโค้งลักษณะข้อสอบ (Item Characteristic Curve) โดยโค้งลักษณะข้อสอบ เป็นกราฟของข้อสอบซึ่งแสดงถึง ความสัมพันธ์ระหว่างระดับความสามารถของผู้สอบ (θ) กับ ความน่าจะเป็นของการตอบข้อสอบถูก $P(\theta)$ วิธีนี้จะประมาณค่าความสามารถของผู้สอบ และ

พารามิเตอร์ของข้อสอบเป็นรายข้อ แล้วนำค่าพารามิเตอร์มาคำนวณหาโอกาสที่จะตอบข้อสอบถูกของผู้สอบแต่ละคนในแต่ละกลุ่ม นำค่าโอกาสที่จะตอบข้อสอบถูกมาเขียนโค้งลักษณะข้อสอบของแต่ละกลุ่มไว้บนกราฟเดียวกัน ข้อใดมีโค้งลักษณะข้อสอบของ 2 กลุ่มต่างกันถือว่าเป็นข้อสอบที่มีความลำเอียง

2.3.2.7 วิธีวิเคราะห์ด้วยโค้งลักษณะข้อสอบที่มีพารามิเตอร์ 3 ตัว การวิเคราะห์ความลำเอียงของข้อสอบวิธีนี้ โดยการประมาณค่าพารามิเตอร์ 3 ตัว นำค่าพารามิเตอร์ที่ได้มาคำนวณหาโอกาสที่จะตอบข้อสอบถูกของผู้สอบแต่ละกลุ่ม แล้วนำมาเขียนเป็นโค้งลักษณะข้อสอบโดยเขียนโค้งลักษณะข้อสอบของผู้สอบ 2 กลุ่มไว้ในภาพเดียวกัน ซึ่งจะพบว่าในภาพเดียวกันของโค้งลักษณะข้อสอบข้อหนึ่งอาจจะมีจุดที่โค้งลักษณะข้อสอบตัดกันมากกว่า 1 จุดก็ได้ ค่าดัชนีความลำเอียงของข้อสอบก็จะคำนวณหาเป็นช่วง ๆ โดยคำนวณหาพื้นที่ระหว่างโค้งลักษณะข้อสอบที่มีความแตกต่างกัน ดังนั้นข้อสอบข้อหนึ่ง ๆ จึงอาจจะมีค่าดัชนีความลำเอียงมากกว่าหนึ่งค่าซึ่งขึ้นอยู่กับจำนวนจุดที่โค้งลักษณะข้อสอบตัดกัน ซึ่งคำนวณโดยใช้สูตรดังนี้

(Hambleton.Ronald K.,1991)

$$P_i(\theta) = c_i + (1-c_i) \frac{e^{Da_i(\theta-b_i)}}{1+e^{Da_i(\theta-b_i)}} \dots\dots\dots (1)$$

เมื่อ $P_i(\theta)$ = ความน่าจะเป็นที่ผู้เข้าสอบคนที่มีระดับความสามารถ θ จะตอบข้อสอบข้อที่ i ได้ถูกต้อง

a_i = ค่าอำนาจจำแนกของข้อสอบข้อที่ i

b_i = ค่าความยากของข้อสอบข้อที่ i

c_i = ค่าโอกาสที่ผู้สอบจะตอบข้อสอบถูกโดยการเดาในข้อที่ i

D = ค่าคงที่มีค่าเท่ากับ 1.7

e = ค่าคงที่มีค่าเท่ากับ 2.718

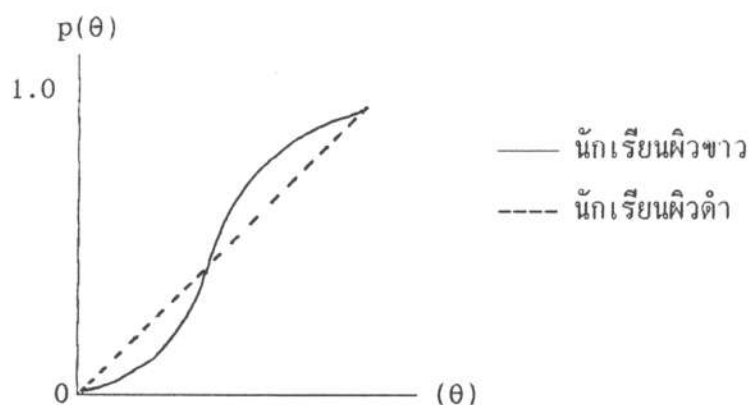
θ = ระดับความสามารถของผู้สอบ

จากสมการที่ 1 เมื่อได้โค้งลักษณะข้อสอบของผู้สอบ 2 กลุ่มในภาพเดียวกันแล้วก็คำนวณหาพื้นที่ในแต่ละกลุ่มตามช่วงของระดับความสามารถที่มีเส้นโค้งลักษณะข้อสอบแตกต่างกัน ซึ่งสามารถหาได้ดังนี้ (ชัชชัย เผ่าพงศ์, 2527 สุรศักดิ์ อมรรัตนศักดิ์, 2530)

$$A_j = \int_s^m p_i(\theta) d\theta \quad \dots\dots\dots (2)$$

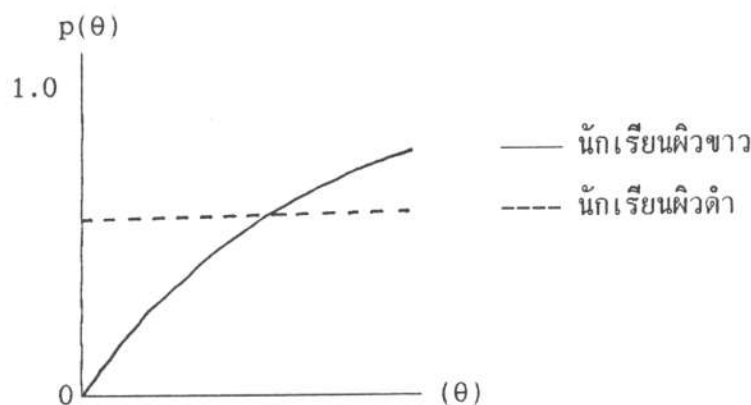
เมื่อ A_j = พื้นที่ของโค้งลักษณะข้อสอบของผู้สอบกลุ่ม j
 s และ m = ช่วงระดับความสามารถที่ต้องการ

ดังนั้นดัชนีความลำเอียงของข้อสอบในช่วงของความสามารถ s ถึง m จึงมีค่าเท่ากับผลต่างของพื้นที่ A_j สองลักษณะและคอฟแมน (Sa-ngob and Coffman, 1980) เสนอให้ใช้ค่าดัชนีความลำเอียงมากกว่าหรือเท่ากับ .20 เป็นค่าที่กำหนดว่าข้อสอบข้อนั้นมี ความลำเอียง ซึ่งภาพแสดงโค้งลักษณะข้อสอบที่มีความลำเอียงในที่นี้จะยกตัวอย่างข้อสอบวัดความถนัดทางด้านภาษา (Verbal Scholastic Aptitude) ซึ่งแสดงให้เห็นถึงข้อสอบที่มีความลำเอียงต่อกลุ่มนักเรียนผิวขาวและนักเรียนผิวดำในสหรัฐอเมริกา จำนวน 3 ข้อ ดังภาพ (Lord, 1977)



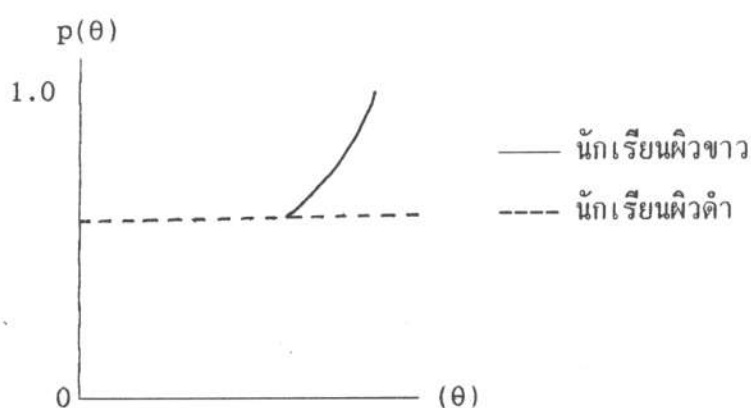
ภาพที่ 1 โค้งลักษณะข้อสอบข้อ 71 ระหว่างนักเรียนผิวขาวและผิวดำ

จากภาพที่ 1 แสดงให้เห็นว่านักเรียนผิวขาวที่มีความสามารถสูงมีโอกาที่จะตอบข้อสอบข้อนี้ถูกมากกว่านักเรียนผิวดำที่มีความสามารถสูง แต่สำหรับผู้ที่มีความสามารถต่ำแล้วนักเรียนผิวดำมีโอกาที่จะตอบข้อสอบถูกมากกว่านักเรียนผิวขาว



ภาพที่ 2 โค้งลักษณะข้อสอบข้อ 2 ระหว่างนักเรียนผิวขาวและผิวดำ

จากภาพที่ 2 แสดงให้เห็นว่าข้อสอบข้อที่ 2 ไม่สามารถจำแนกนักเรียนผิวดำได้ไม่ว่า นักเรียนผิวดำจะมีความสามารถในระดับใดก็ตามต่างก็มีโอกาสที่จะตอบข้อสอบข้อนี้ได้เท่า ๆ กัน สำหรับผู้ที่มีความสามารถสูงของนักเรียนผิวขาวมีโอกาสตอบข้อสอบถูกมากกว่านักเรียนผิวดำ แต่กลุ่มผู้มีความสามารถต่ำ พบว่านักเรียนผิวดำมีโอกาสตอบข้อสอบถูกมากกว่านักเรียนผิวขาว



ภาพที่ 3 โค้งลักษณะข้อสอบข้อ 24 ระหว่างนักเรียนผิวขาวและผิวดำ

จากภาพที่ 3 แสดงให้เห็นว่าข้อสอบข้อนี้ยากสำหรับนักเรียนผิวดำทุก ๆ ระดับความสามารถ และยากสำหรับนักเรียนผิวขาวที่มีความสามารถต่ำด้วย โดยพิจารณาจากค่าความน่าจะเป็นที่นักเรียนผิวดำและนักเรียนผิวขาวที่จะตอบข้อสอบข้อที่ 24 ได้ถูกต้องและนักเรียนผิวขาวที่มีความสามารถสูงมีโอกาสที่จะตอบข้อสอบถูกมากขึ้นกว่านักเรียนผิวดำ



2.4 โครงสร้างหลักสูตรกลุ่มวิชาสร้างเสริมประสบการณ์ชีวิต

หลักสูตรประถมศึกษาพุทธศักราช 2521 (ฉบับปรับปรุง พ.ศ. 2533) มีความคาดหวังให้ผู้เรียน คิดเป็น ทำเป็น แก้ปัญหาเป็น และทำงานอย่างเป็นระบบ เพื่อให้ผู้เรียนสามารถพัฒนาคุณภาพชีวิตให้พร้อมที่จะทำประโยชน์กับสังคมตามบทบาทหน้าที่ของตนโดยจะต้องปลูกฝังให้ผู้เรียนเกิดประโยชน์ในด้านต่าง ๆ ต่อไปนี้

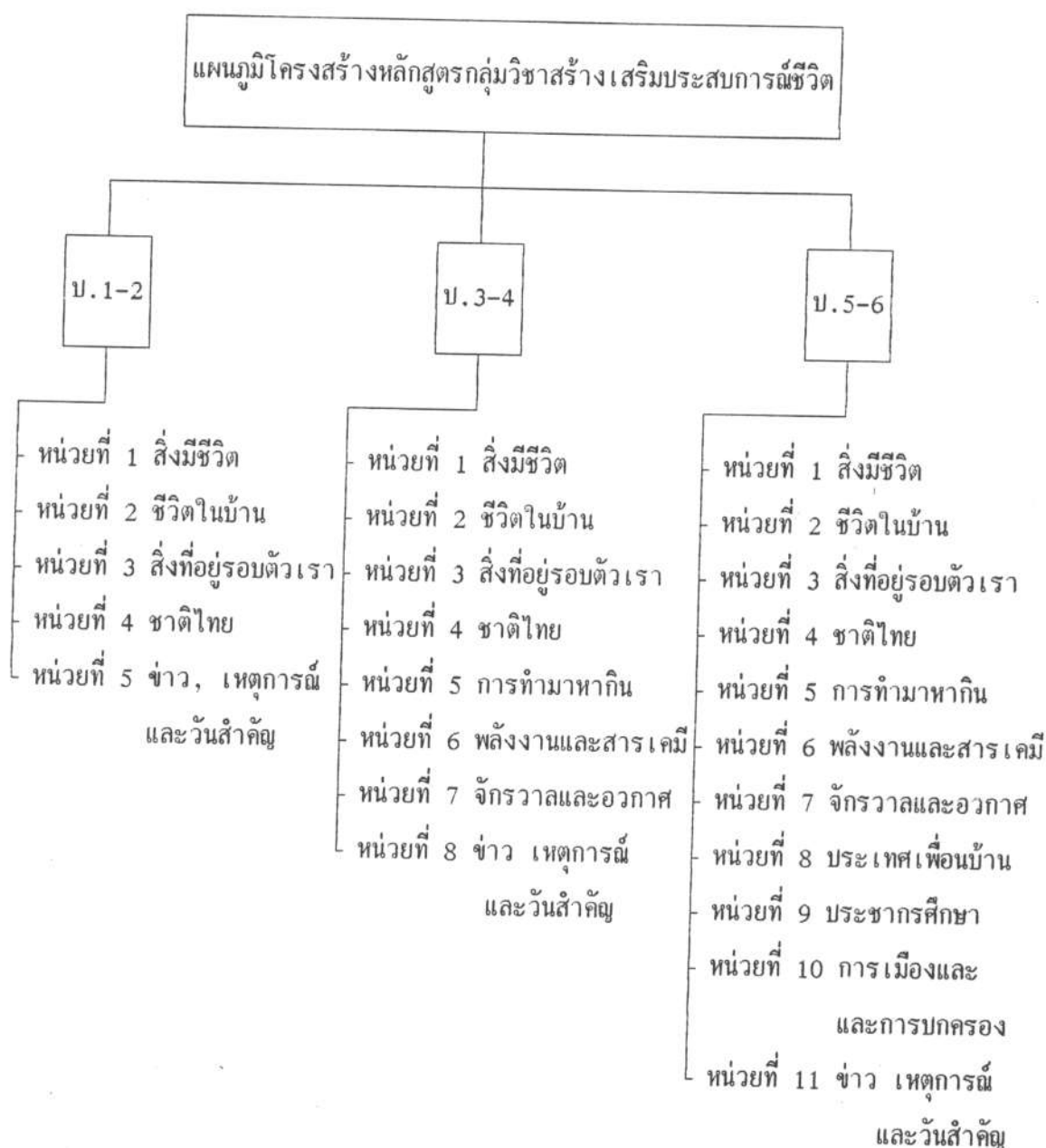
1.1 การพัฒนาคนได้แก่ มีความรู้พื้นฐาน มีสุขภาพกายใจสมบูรณ์ แก้ปัญหาเป็น เสียสละ และมุ่งพัฒนา

1.2 การพัฒนาอาชีพ ได้แก่ การทำงานร่วมกับผู้อื่นได้ รักการทำงานและทำงานเป็น

1.3 การพัฒนาสังคม ได้แก่ การปฏิบัติหน้าที่ของตนต่อบ้าน ชุมชน ประเทศ และโลก

หลักสูตรประถมศึกษา พุทธศักราช 2521 (ฉบับปรับปรุง พ.ศ. 2533) ได้จัดมวลประสบการณ์ที่มุ่งให้เป็นความรู้พื้นฐานสำหรับผู้เรียนระดับประถมศึกษาออกเป็น 5 กลุ่ม คือ กลุ่มวิชาทักษะที่เป็นเครื่องมือการเรียนรู้ (ภาษาไทย+คณิตศาสตร์) กลุ่มวิชาสร้างเสริมประสบการณ์ชีวิต (สพข.) กลุ่มวิชาสร้างเสริมลักษณะนิสัย (สสน.) กลุ่มวิชาการงานและพื้นฐานอาชีพ (กพอ.) และกลุ่มวิชาประสบการณ์พิเศษ

กลุ่มวิชาสร้างเสริมประสบการณ์ชีวิต (สพข.) เป็นกลุ่มที่วัดด้วยความรู้ด้านต่าง ๆ ในการดำรงชีวิตของสังคมมนุษย์ ทางด้านสุขภาพอนามัย ความสัมพันธ์ระหว่างมนุษย์กับสิ่งแวดล้อม การเมือง การปกครอง พระพุทธศาสนา หลักธรรมคำสอน วัฒนธรรม ขนบธรรมเนียมประเพณี ความสัมพันธ์ของประเทศในระดับภูมิภาคและโลก วิทยาศาสตร์ และเทคโนโลยี ซึ่งความรู้ต่างๆ เหล่านี้จะเป็นสิ่งที่ช่วยแก้ปัญหาของชีวิตและสังคม เพื่อให้คนมีความเป็นอยู่และการดำรงชีวิตที่ดีขึ้น ดังนั้นจากความสำคัญที่กล่าวมาข้างต้น หลักสูตรประถมศึกษา พุทธศักราช 2521 (ฉบับปรับปรุง 2533) จึงได้พัฒนาขึ้นตามแนวคิดดังกล่าว โดยจัดทำเป็นหน่วย เริ่มจากตนเอง และขยายขอบเขตกว้างขึ้นตามวัยและระดับชั้น ดังแผนภูมิโครงสร้างหลักสูตรกลุ่มวิชาสร้างเสริมประสบการณ์ชีวิต



ภาพที่ 4 โครงสร้างหลักสูตรกลุ่มสร้างเสริมประสบการณ์ชีวิต

จากแผนภูมิโครงสร้างของหลักสูตรกลุ่มวิชาสร้างเสริมประสบการณ์ชีวิต (สพช.) ข้างต้น จะพบว่าหลักสูตรประถมศึกษา พุทธศักราช 2521 (ฉบับปรับปรุง พ.ศ. 2533) ได้ให้ความสำคัญในระดับประถมศึกษาปีที่ 5-6 มากที่สุด ถ้าพิจารณาในรายละเอียดเนื้อหาวิชาที่เรียนทั้งหมด 11 หน่วยการเรียนรู้ และถ้าพิจารณาเวลาเรียนของหลักสูตรกลุ่มสร้างเสริมประสบการณ์ชีวิตก็พบว่าในระดับประถมศึกษาปีที่ 5-6 มีเวลาเรียนมากที่สุด คือ ชั้นละ 750 คาบ หรือ 250 ชั่วโมง รองลงมาคือ ชั้นประถมศึกษาปีที่ 3-4 มีเวลาเรียนชั้นละ 600 คาบ หรือ 200 ชั่วโมง และ

ชั้นประถมศึกษาปีที่ 1-2 มีเวลาเรียนชั้นละ 450 คาบ หรือ 150 ชั่วโมง และถ้าพิจารณาอัตราเวลาเรียนรวมของกลุ่มประสบการณ์ทั้ง 5 กลุ่มในระดับประถมศึกษาปีที่ 5-6 ที่หลักสูตรได้จัดให้ผู้เรียนก็พบว่ากลุ่มวิชาสร้างเสริมประสบการณ์ชีวิต (สพช.) มีอัตราเวลาเรียนในหลักสูตรมากเป็นอันดับที่ 2 คิดเป็น 25% หรือ 750 คาบ/ปี รองจากกลุ่มวิชาการงานพื้นฐานอาชีพ (กพอ.) ที่มีอัตราเวลาเรียนคิดเป็น 30% หรือ 900 คาบ/ปี ดังนั้นจะพบว่า กลุ่มวิชาสร้างเสริมประสบการณ์ชีวิต (สพช.) เป็นกลุ่มวิชาที่หลักสูตรประถมศึกษา พุทธศักราช 2521 (ฉบับปรับปรุง พ.ศ. 2533) ให้ความสำคัญมาก เพราะกลุ่มวิชาสร้างเสริมประสบการณ์ชีวิต (สพช.) เป็นกลุ่มวิชาหนึ่งที่ทำให้เด็กเรียนที่เรียนจบไปแล้วสามารถปรับตัวให้เข้ากับสภาพแวดล้อมทั้งทางธรรมชาติและทางสังคม และสามารถแก้ปัญหาในชีวิตประจำวันได้เป็นอย่างดี (หลักสูตรประถมศึกษา 2521, 2534)

2.5 งานวิจัยที่เกี่ยวข้อง

งานวิจัยภายในประเทศ

ชัชชัย เผ่าพงศ์ (2527) ได้ทำการวิเคราะห์ความลำเอียงของข้อสอบมาตรฐานวัดความถนัดทางการเรียนด้านคณิตศาสตร์และภาษาในระดับมัธยมศึกษาตอนต้น ซึ่งพัฒนาโดยสำนักทดสอบทางการศึกษาและจิตวิทยา มหาวิทยาลัยศรีนครินทรวิโรฒประสานมิตรระหว่างกลุ่มผู้สอบเพศชายและเพศหญิง โดยการเปรียบเทียบโค้งลักษณะข้อสอบด้วยวิธีแบบจำลองโลจิสติกที่คำนวณค่าพารามิเตอร์ 3 ตัว ผลการวิเคราะห์พบว่าจากแบบทดสอบวัดความถนัดทางการเรียนด้านคณิตศาสตร์จำนวน 30 ข้อ มีข้อสอบที่ลำเอียงต่อกลุ่มใดกลุ่มหนึ่งโดยเฉพาะ 9 ข้อ คือ ลำเอียงต่อกลุ่มนักเรียนชาย 7 ข้อ และลำเอียงต่อกลุ่มนักเรียนหญิง 2 ข้อ ข้อสอบที่ลำเอียงต่อกลุ่มนักเรียนเพศชายและกลุ่มนักเรียนเพศหญิง ในระดับปานกลางขึ้นไปมีจำนวน 5 ข้อ ซึ่งวัดในเรื่องร้อยละ การหาปริมาตรและการหาความยาวของเส้นในด้านของรูปสามเหลี่ยมจำนวนเรื่องละ 1 ข้อ และเรื่องเกี่ยวกับโจทย์ปัญหาจำนวน 2 ข้อ ส่วนแบบทดสอบวัดความถนัดทางการเรียนด้านภาษาจำนวน 40 ข้อ มีข้อสอบที่ลำเอียงต่อนักเรียนชายโดยเฉพาะ 3 ข้อ และลำเอียงต่อนักเรียนหญิงโดยเฉพาะ 8 ข้อ

ทัศนีย์ พิรมนตรี (2529) ได้ทำการวิเคราะห์ความลำเอียงของแบบทดสอบวิชาคณิตศาสตร์โครงการตรวจสอบคุณภาพการศึกษาของสำนักทดสอบทางการศึกษา กรมวิชาการ กระทรวงศึกษาธิการ กลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 6 ปีการศึกษา 2526 โดยการเปรียบเทียบจำนวนข้อสอบที่มีความลำเอียงระหว่างกลุ่มนักเรียนในกรุงเทพมหานครกับกลุ่ม

นักเรียนในภาคภูมิศาสตร์ทั้ง 5 ภาค วิเคราะห์ความลำเอียง 3 วิธี คือ วิธีแปลงค่าความยากของข้อสอบ วิธีไค-สแควร์ และวิธีไค์ลักษณะข้อสอบที่ใช้พารามิเตอร์ 3 ตัว พบว่าข้อสอบมีความลำเอียงจำนวนข้อมากที่สุด เมื่อใช้พารามิเตอร์ 3 ตัว และในแต่ละวิธีพบว่ามีข้อสอบที่ลำเอียงซ้ำกันระหว่างกลุ่มนักเรียนในกรุงเทพมหานครกับกลุ่มนักเรียนในทุกภาคภูมิศาสตร์

สุรศักดิ์ อมรรัตนศักดิ์ (2530) ได้ทำการศึกษาเปรียบเทียบผลของวิธีวิเคราะห์ความลำเอียงของข้อสอบที่แตกต่างกัน 4 วิธี คือ วิธีวิเคราะห์ความแปรปรวน วิธีแปลงค่าความยากของข้อสอบ วิธีไค์ลักษณะข้อสอบที่มีพารามิเตอร์ 1 และ 3 ตัว โดยการหาสัมประสิทธิ์สหสัมพันธ์ของดัชนีความลำเอียงระหว่างวิธีวิเคราะห์ความลำเอียง 4 วิธี ผลการวิจัยพบว่า วิธีวิเคราะห์ความลำเอียงแต่ละวิธีพบจำนวนข้อที่ลำเอียงต่างกัน โดยวิธีไค์ลักษณะข้อสอบที่มีพารามิเตอร์ 3 ตัว พบจำนวนข้อสอบที่ลำเอียงมากที่สุด รองลงมาได้แก่วิธีวิเคราะห์ความแปรปรวน วิธีแปลงค่าความยากของข้อสอบ พบจำนวนข้อที่ลำเอียงน้อยที่สุด ส่วนค่าสัมประสิทธิ์สหสัมพันธ์ของดัชนีความลำเอียงระหว่างวิธีวิเคราะห์ความลำเอียง 4 วิธี พบว่ามีค่าอยู่ระหว่าง .7535 ถึง .9921

พัชรี ปิยภัณฑ (2531) ได้ทำการศึกษาการวิเคราะห์ความลำเอียงของข้อสอบจากแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนชั้นประถมศึกษาปีที่ 6 โดยทำการวิเคราะห์ความลำเอียงของข้อสอบที่มีต่อกลุ่มนักเรียนเพศหญิง กลุ่มนักเรียนเพศชาย และกลุ่มนักเรียนในกรุงเทพมหานครกับนักเรียนในจังหวัดสมุทรสาคร ด้วยวิธีวิเคราะห์ 3 วิธี คือวิธีแปลงค่าความยากของข้อสอบ วิธีไค-สแควร์ และวิธีไค์ลักษณะข้อสอบที่มีพารามิเตอร์ 1 ตัว ผลการวิจัยพบว่าในแต่ละวิธีพบจำนวนข้อสอบลำเอียงแตกต่างกัน วิธีไค์ลักษณะข้อสอบที่มีพารามิเตอร์ 1 ตัว มีจำนวนข้อที่ลำเอียงมากที่สุด 45 ข้อ ลำเอียงต่อกลุ่มนักเรียนเพศหญิง กลุ่มนักเรียนเพศชาย จำนวน 7 ข้อ ลำเอียงต่อนักเรียนในกรุงเทพมหานครกับสมุทรสาคร จำนวน 18 ข้อ วิธีแปลงค่าความยากข้อสอบตรวจพบข้อสอบลำเอียง 2 ข้อ ระหว่างกลุ่มนักเรียนเพศหญิงและกลุ่มนักเรียนเพศชาย ลำเอียงต่อนักเรียนในกรุงเทพมหานครกับสมุทรสาคร 1 ข้อ วิธีไค-สแควร์พบข้อสอบลำเอียงระหว่างนักเรียนชายและหญิง 3 ข้อ และพบข้อสอบลำเอียงระหว่างนักเรียนในกรุงเทพมหานครกับสมุทรสาคร 3 ข้อ

ไชยยา อะการะวัง (2536) ได้ทำการศึกษาความลำเอียงของแบบทดสอบวิชาสังคมศึกษาที่ใช้สอบคัดเลือกบุคคลเข้าศึกษาต่อระดับปริญญาตรีชั้นปีที่ 1 ของมหาวิทยาลัยศรีนครินทรวิโรฒ มหาสารคาม ประจำปีการศึกษา 2535 โดยวิธีรับตรง และเปรียบเทียบจำนวนข้อสอบที่มีความลำเอียงระหว่างเพศหญิงกับเพศชาย และระหว่างนักเรียนที่มาจากสถานศึกษาที่ตั้งอยู่ในเมือง

กับนอกเมือง ด้วยวิธีการวิเคราะห์ความลำเอียง 4 วิธี คือ วิธีแปลงค่าความยากของข้อสอบ วิธีไค-สแควร์ วิธีวิเคราะห์ความแปรปรวน และวิธีให้ผู้เชี่ยวชาญพิจารณาข้อสอบ ผลการวิจัยพบว่า

1. ในแบบทดสอบวิชาสังคมศึกษา ก จำนวน 100 ข้อ พบว่า วิธีวิเคราะห์ความแปรปรวนพบข้อสอบลำเอียงมากที่สุด 33 ข้อ วิธีไค-สแควร์ 17 ข้อ วิธีแปลงค่าความยากข้อสอบพบ 4 ข้อ วิธีให้ผู้เชี่ยวชาญตรวจสอบไม่พบข้อสอบที่มีความลำเอียง ส่วนการวิเคราะห์ความลำเอียงระหว่างกลุ่มในเมืองกับกลุ่มนอกเมือง พบว่าวิธีวิเคราะห์ความแปรปรวนพบจำนวนข้อสอบลำเอียงมากที่สุด 23 ข้อ วิธีไค-สแควร์ 4 ข้อ วิธีให้ผู้เชี่ยวชาญตรวจสอบพบ 2 ข้อ วิธีแปลงค่าความยากข้อสอบพบ 1 ข้อ

2. ในแบบทดสอบวิชาสังคมศึกษา กข จำนวน 100 ข้อ พบว่า วิธีวิเคราะห์ความแปรปรวนพบข้อสอบลำเอียงมากที่สุด 23 ข้อ วิธีแปลงค่าความยากข้อสอบ 14 ข้อ วิธีไค-สแควร์ พบ 1 ข้อ วิธีให้ผู้เชี่ยวชาญตรวจสอบไม่พบข้อสอบที่มีความลำเอียง ส่วนการวิเคราะห์ความลำเอียงระหว่างกลุ่มในเมืองกับกลุ่มนอกเมือง พบว่าวิธีการวิเคราะห์ความแปรปรวนพบจำนวนมากที่สุด 22 ข้อ วิธีแปลงค่าความยากพบ 14 ข้อ วิธีไค-สแควร์พบ 3 ข้อ และวิธีให้ผู้เชี่ยวชาญพิจารณาพบ 2 ข้อ

เจน จันทรแสง (2537) ได้ทำการศึกษาการวิเคราะห์ความลำเอียงของข้อสอบวัดผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 3 ที่เกิดจากตัวแปร เพศ และที่ตั้งของโรงเรียน ด้วยวิธีวิเคราะห์ความลำเอียง 4 วิธี คือ วิธีแปลงค่าความยากข้อสอบ วิธีวิเคราะห์ความแปรปรวน วิธีไค-สแควร์ และวิธีใช้ลักษณะข้อสอบที่มีพารามิเตอร์ 1 ตัว กลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 ปีการศึกษา 2536 ของโรงเรียนสังกัดกรมสามัญศึกษา จังหวัดนครสวรรค์จำนวน 1,600 คน ผลการวิจัยพบว่า

1. วิธีวิเคราะห์ความลำเอียงแต่ละวิธีตรวจพบจำนวนข้อสอบที่ลำเอียงแตกต่างกัน วิธีใช้ลักษณะข้อสอบที่มีพารามิเตอร์ 1 ตัว พบจำนวนข้อสอบลำเอียงมากที่สุด โดยลำเอียงตามตัวแปรเพศ 10 ข้อ ลำเอียงตามตัวแปรที่ตั้งโรงเรียน 15 ข้อ วิธีไค-สแควร์, ANOVA, ลำเอียงตามตัวแปรเพศ 7 ข้อ ลำเอียงตามตัวแปรที่ตั้งโรงเรียน 14 ข้อ และวิธีแปลงค่าความยากข้อสอบพบว่าลำเอียงตามตัวแปรเพศ 1 ข้อ ลำเอียงตามตัวแปรที่ตั้งโรงเรียน 5 ข้อ

2. สัมประสิทธิ์ สหสัมพันธ์ เมื่อวิเคราะห์ตามตัวแปรเพศพบว่ามีค่าระหว่าง .5284 ถึง .9970 และเมื่อวิเคราะห์ตามตัวแปรที่ตั้งโรงเรียนพบว่ามีค่าระหว่าง .6008 ถึง .9475

กาญจนา วัธนสุนทร (2538) ได้ทำการศึกษาพัฒนาเกณฑ์ตัดสินข้อสอบลำเอียงทางเพศ ด้วยข้อมูลเชิงประจักษ์สำหรับดัชนี 4 ตัว คือพื้นที่ระหว่างโค้งการตอบข้อสอบชนิดคิดเครื่องหมาย (SA) และไม่คิดเครื่องหมาย (UA) จากวิธีทฤษฎีการตอบข้อสอบโมเดล 2 พารามิเตอร์ ดัชนีแอลฟา (α_{MH}) จากวิธีแมนเทิล-แฮนส์เชล และเบต้า (β_{SIB}) จากวิธี SIBTEST โดยใช้ข้อมูลการตอบข้อสอบคัดเลือกบุคคลเข้าศึกษาต่อในสถาบันอุดมศึกษาของทบวงมหาวิทยาลัย ปีการศึกษา 2535 ในความยาวแบบสอบ 20, 30 และ 40 ข้อ สำหรับวิชาคณิตศาสตร์ และ 50, 60, 70 และ 80 สำหรับวิชาภาษาอังกฤษใช้ผู้สอบ 6 ขนาด คือ 100, 200, 400, 600 800 และ 1,000 คน ผลจากการวิจัยพบว่า

1. เกณฑ์พัฒนาจากข้อมูลเชิงประจักษ์เพื่อใช้ในการตัดสินความลำเอียงของข้อสอบ ระหว่างผู้สอบหญิงและชายคือ
 - (1) $SA > 80$ และ $UA > .50$ กรณีความยาวแบบสอบต่ำกว่า 50 ข้อ
 - (2) $SA > 40$ และ $UA > 1.20$ กรณีความยาวแบบสอบ 50 ข้อขึ้นไป
 - (3) $.60 > \alpha_{MH} > 1.40$ และ $\beta_{SIB} > .06$ ทุกความยาวแบบสอบและทุกขนาดผู้สอบ
 2. การตรวจค้นข้อสอบลำเอียงทางเพศมีความไม่คงที่ข้ามขนาดผู้สอบและความยาวแบบสอบ
 3. ความสอดคล้องในการตรวจค้นข้อสอบลำเอียงภายในวิธีเดียวกันข้ามขนาดผู้สอบค่อนข้างต่ำ แต่จะสูงขึ้นที่ขนาดผู้สอบตั้งแต่ 600 คนขึ้นไป
 4. ข้อสอบลำเอียงวิชาคณิตศาสตร์ส่วนใหญ่ลำเอียงเข้าข้างผู้สอบชายและวิชาภาษาอังกฤษลำเอียงเข้าหาผู้สอบหญิงเมื่อใช้ดัชนี SA และ α_{MH} แต่ดัชนี β_{SIB} ให้ผลตรงข้าม
- งานวิจัยต่างประเทศ**

ผจงจิต อินทสุวรรณ (Intasuwat 1979 อ้างถึงใน พัชรีย์ บิยภัณฑ, 2531) ได้ใช้วิธีการวิเคราะห์ความลำเอียง 3 วิธี วิธีของราสช์ (Rasch) วิธีโค้งลักษณะข้อสอบที่มี 3 พารามิเตอร์ และวิธีไค-สแควร์ กับกลุ่มนักเรียนอายุ 14 ปี ชาวอังกฤษ อเมริกัน และนิวซีแลนด์ จำนวน 8,601 คน เครื่องมือที่ใช้เป็นแบบทดสอบ International for the Evaluation of Educational Achievement ผลการวิจัยพบว่าสหสัมพันธ์ระหว่างทั้ง 3 วิธี มีค่าอยู่ระหว่าง .51-.98 โดยวิธีไค-สแควร์ และวิธีของ Rasch มีค่าสหสัมพันธ์สูงถึง .98

สงบ ลักษณะ (1979 อ้างถึงใน เจน จันทรแสง, 2537) ได้ประยุกต์ใช้วิธีโค้งลักษณะข้อสอบที่มีพารามิเตอร์ 3 ตัว และวิธีวิเคราะห์ความแปรปรวนในการวิเคราะห์ข้อสอบ

Iowa Tests of Basic Skills (ITBS) ในข้อสอบคำศัพท์และข้อสอบคณิตศาสตร์เพื่อหาความลำเอียงทางเพศและผิว ผลการวิจัยพบว่าวิธีวิเคราะห์ทั้ง 2 วิธี มีความสอดคล้องกัน วิธีวิเคราะห์ความแปรปรวนเป็นวิธีที่ใช้ได้ในทางปฏิบัติแต่เป็นวิธีไม่ครอบคลุม วิธีไค์ลักษณะข้อสอบเป็นวิธีที่ครอบคลุมกว่า แต่ไม่เหมาะสมในทางปฏิบัติเนื่องจากต้องใช้กลุ่มตัวอย่างขนาดใหญ่

Green and Draper (1972 อ้างถึงใน ชัชชัย เผ่าพงศ์, 2527) ได้ศึกษาความลำเอียงของข้อสอบจากค่าอำนาจจำแนก ซึ่งคำนวณจากค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนของข้อสอบรายข้อกับคะแนนรวมของข้อสอบทั้งฉบับ จากข้อสอบ California Achievement Test (CAT) ระหว่างกลุ่มนักเรียนผิวดำ และผิวขาว พบว่าจากจำนวนข้อสอบทั้งหมดร้อยละ 30 มีความลำเอียงกับกลุ่มนักเรียน ซึ่งมาจากสภาพภูมิศาสตร์ต่างกัน คือ นักเรียนผิวดำมาจากตอนใต้ นักเรียนผิวขาวมาจากตะวันตกเฉียงใต้ของอเมริกา

Ironson and Subkoviak (Subkoviak, 1984) ได้ศึกษาถึงวิธีการในการวิเคราะห์ความลำเอียงของข้อสอบโดยใช้วิธีแปลงค่าความยากข้อสอบ วิธีไค์-สแควร์ วิธีไค์ลักษณะข้อสอบ วิธีพอยท์ไบซีเรียล โดยนำข้อสอบมาตรฐานไปทดสอบกับนักเรียนผิวดำและผิวขาว ผลการวิจัยพบว่าวิธีไค์ลักษณะข้อสอบเป็นวิธีที่ดีที่สุด รองลงมาคือวิธีไค์-สแควร์ และวิธีการแปลงค่าความยากข้อสอบ

Rudner and Convey (Subkoviak, 1984 Citing Rudner and Convey 1978) ได้ทำการวิเคราะห์ความลำเอียงของข้อสอบโดยวิธีแปลงค่าความยากข้อสอบ วิธีไค์-สแควร์ วิธีไค์ลักษณะข้อสอบ และวิธีวิเคราะห์ห้วงค์ประกอบ ผลการวิจัยพบว่า วิธีไค์ลักษณะข้อสอบ วิธีไค์-สแควร์ และวิธีแปลงค่าความยากข้อสอบเป็นวิธีที่เหมาะสม

Baeza (1989) ได้ทำการศึกษาพฤติกรรมการตอบข้อสอบภาคปฏิบัติของนักเรียนชาวอเมริกันอินเดียและอเมริกันคอเคเซียน เพื่อพิจารณาระดับความลำเอียงทั้งภายในและภายนอกของข้อสอบโดยใช้ตัวแปรด้านเศรษฐกิจ สังคม และระดับคะแนนเฉลี่ยในชั้นมัธยมปลายเป็นตัวแปรจับคู่กลุ่มตัวอย่างนอกเหนือไปจากตัวแปรเพศ โรงเรียนที่เรียนอยู่ และปีการศึกษาที่เรียน วิธีที่ใช้ในการศึกษาความลำเอียงคือ แมนเทล - ฮันเซล (MH) และวิธีการจับคู่กลุ่มตัวอย่างของ McNemar ตัวแปรจับคู่สำหรับวิธี MH คือ คะแนนรวมจากแบบสอบส่วนตัวแปรจับคู่สำหรับวิธี McNemar คือ ตัวแปรด้านเศรษฐกิจ สังคม และระดับคะแนนเฉลี่ย ผลการวิจัยพบว่าวิธี MH ซึ่งใช้เกณฑ์ภายในตรวจสอบว่ามีข้อสอบที่มีความลำเอียงเล็กน้อย ส่วนวิธี McNemar ซึ่งใช้เกณฑ์ภายนอกนั้นพบว่าเมื่อใช้ตัวแปรจับคู่เป็นสถานภาพทางเศรษฐกิจ สังคม จะพบข้อสอบมีความลำเอียงต่ำกว่าการจับคู่ด้วยระดับคะแนนเฉลี่ย

Oshima (1989, อ้างถึงใน กาญจนนา วันสุนทร, 2538) ได้ทำการศึกษาผลของข้อมูลหลายมิติที่มีต่อการตรวจสอบความลำเอียงของข้อสอบ โดยการจำลองข้อมูล 42 ชุดจากข้อสอบ 40 ข้อ มีโครงสร้างความเป็นพหุมิติ 2 ชนิด คือชนิดที่มีลักษณะแฝงที่เด่นอยู่ลักษณะเดียวและมีลักษณะแฝงเด่น 2 ลักษณะ เกมที่ใช้ในการตัดสินใจได้จากการคำนวณค่าดัชนี SA UA SSOS (ผลบวกกำลังสองที่คิดเครื่องหมาย) USOS (ผลบวกกำลังสองที่ไม่คิดเครื่องหมาย) เกมที่การตัดสินใจความลำเอียงคือ ค่าที่มากกว่า 2 ส่วนเบี่ยงเบนมาตรฐานของ SA UA SSOS และ USOS ที่คำนวณได้จากกลุ่มอ้างอิง 2 กลุ่ม ผลการศึกษาพบว่าเมื่อผู้สอบ 2 กลุ่ม มีค่าเฉลี่ยของความสามารถแทรกซ้อน (Nuisance ability) เท่ากันพบว่าการระบุว่าข้อสอบลำเอียงทั้ง ๆ ที่จำลองขึ้นให้ไม่ลำเอียงซึ่งพบทั้งในข้อสอบที่มีลักษณะเอกมิติ และพหุมิติ และเมื่อค่าเฉลี่ยของความสามารถแทรกซ้อนต่างกัน ก็จะสามารถระบุข้อสอบที่ลำเอียงได้ถูกต้องในอัตราสูง โดยสรุปแล้วดัชนีจาก IRT สามารถตรวจสอบข้อสอบพหุมิติที่สร้างให้ลำเอียง และไม่ลำเอียงได้อย่างถูกต้องแต่อาจใช้ในการตรวจสอบลดลง ถ้าจำนวนข้อสอบลำเอียงในแบบสอบมีมากขึ้น

Merz and Grossen (1979 Cited in Subkoviak, 1984 : 51 อ้างถึงใน สุรศักดิ์ อมรรตน์ศักดิ์, 2530) ได้ทำการศึกษาวิธีวิเคราะห์ความลำเอียงของข้อสอบโดยใช้ข้อมูลสมมติ (Simulate Data) วิเคราะห์ความลำเอียง 4 วิธี คือ วิธีแปลงค่าความยากของข้อสอบ วิธีไค - สแควร์ วิธีวิเคราะห์หอคัพประกอบ และวิธีโค้งลักษณะข้อสอบที่มีพารามิเตอร์ 3 ตัว ผลการวิจัยพบว่า วิธีแปลงค่าความยากของข้อสอบเป็นวิธีที่ดีที่สุด วิธีโค้งลักษณะข้อสอบที่มีพารามิเตอร์ 3 ตัว และวิธีไค - สแควร์ต่างก็เป็นวิธีที่ใช้ได้

จากการศึกษารายงานการวิจัยดังกล่าวข้างต้น จะเห็นได้ว่าผลการวิจัยเกี่ยวกับการวิเคราะห์ความลำเอียงที่มีต่อตัวแปรเพศ และสถานที่ตั้งของโรงเรียน มีความสอดคล้องกัน กล่าวคือ จากผลการวิจัยของ เจน จันท์แสง (2537) ไชยยา อะการะวัง (2536) พชรปิยภักดิ์ (2531) และ ทศนีย์ พิรมนตรี (2530) พบว่าข้อสอบมีความลำเอียงต่อกลุ่มนักเรียนที่มีเพศต่างกัน และกลุ่มนักเรียนที่เรียนอยู่ในโรงเรียนที่ตั้งอยู่ในเมืองกับนอกเมือง และกลุ่มนักเรียนที่เรียนอยู่ในกรุงเทพมหานครกับกลุ่มนักเรียนที่เรียนอยู่ในต่างจังหวัด และจากผลการวิจัยของ สุรศักดิ์ อมรรตน์ศักดิ์ (2530) ชัชชัย เผ่าพงศ์ (2527) Green and Draper (1972) และ Ironson and Subkoviak (1984) พบว่าข้อสอบมีความลำเอียงต่อกลุ่มนักเรียนที่มีเพศต่างกัน จากผลการวิจัยที่กล่าวมาผู้วิจัยจึงนำมาเป็นกรอบแนวคิดในการวิจัยดังนี้

กรอบแนวคิดที่ใช้ในการวิจัย
Conceptual Framework

ตัวแปรต้น	ตัวแปรตาม
1. เพศ 1.1 เพศหญิง 1.2 เพศชาย 2. สถานที่ตั้งของโรงเรียน 2.1 โรงเรียนที่ตั้งอยู่ในเขตสุขภาพอำเภอ 2.2 โรงเรียนที่ตั้งอยู่นอกเขตสุขภาพอำเภอ 3. วิธีการวิเคราะห์ความลำเอียงของข้อสอบ 3.1 วิธีแปลงค่าความยากของข้อสอบ 3.2 วิธีไค - สแควร์ 3.3 วิธีวิเคราะห์ความแปรปรวน 3.4 วิธีไคถึงลักษณะข้อสอบที่มีพารามิเตอร์ 3 ตัว	1. ข้อสอบที่มีความลำเอียง (Item bias) ในแบบสอบวัดผลสัมฤทธิ์ทางการเรียน กลุ่มวิชาสร้างเสริมประสบการณ์ชีวิต (สบข.) ระดับชั้นประถมศึกษาปีที่ 6 จำนวน 2 ฉบับ ได้แก่ 1.1 แบบสอบวิชา สบข.ฉบับที่ 1 จำนวน 30 ข้อ เป็นแบบสอบที่ใช้วัดความรู้ความเข้าใจเกี่ยวกับทักษะกระบวนการทางวิทยาศาสตร์ และวัดความรู้ความเข้าใจเกี่ยวกับสุขภาพและสุขนิสัย 1.2 แบบสอบวิชา สบข.ฉบับที่ 2 จำนวน 30 ข้อ เป็นแบบสอบที่ใช้วัดความรู้ ความเข้าใจเกี่ยวกับชาติ ศาสนา พระมหากษัตริย์ ใช้วัดความรู้ ความเข้าใจเกี่ยวกับระบอบประชาธิปไตย และใช้วัดความรู้ ความเข้าใจเกี่ยวกับเศรษฐกิจการทำมาหากินและความสัมพันธ์กับประเทศเพื่อนบ้าน 2. ค่าสัมประสิทธิ์ความเที่ยงของแบบสอบหลังการตัดข้อสอบที่พบว่าลำเอียงออกไปแล้ว

ภาพที่ 5 กรอบแนวคิดที่ใช้ในการวิจัย