

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

การเสนอเนื้อหาสาระ รายงาน เอกสารและงานวิจัยที่เกี่ยวข้องกับการวิจัยในครั้งนี้ผู้วิจัยแยกเสนอเป็น 7 ตอน ตอนแรกว่าด้วยเรื่องมโนทัศน์เกี่ยวกับความคลาดเคลื่อนในการประเมิน ตอนที่ 2 มโนทัศน์เกี่ยวกับการป้องกันความคลาดเคลื่อนในการประเมิน ตอนที่ 3 มโนทัศน์เกี่ยวกับการประมาณค่าความคลาดเคลื่อนในการประเมินผล ตอนที่ 4 มโนทัศน์เกี่ยวกับการปรับแก้ความคลาดเคลื่อนในการประเมินผล ตอนที่ 5 มโนทัศน์เกี่ยวกับการใช้ทฤษฎีการสรุปอ้างอิงในการตรวจสอบและประมาณค่าความคลาดเคลื่อนในการประเมินผล ตอนที่ 6 มโนทัศน์เกี่ยวกับการประเมินผลการปฏิบัติงานของครู และตอนที่ 7 คืองานวิจัยที่เกี่ยวข้อง และกรอบความคิดในการวิจัย

ตอนที่ 1 มโนทัศน์เกี่ยวกับความคลาดเคลื่อนในการประเมินผล

ความคลาดเคลื่อน (error) มีความหมายตามพจนานุกรมของเว็บสเตอร์ (Webster's Encyclopaedia Dictionary, 1994) คือความผิดพลาดในการตัดสินใจหรือภาวะที่เชื่ออย่างไม่จริงไม่ถูกต้อง

โดยทั่วไปในงานประเมินมักใช้มาตรฐานค่าเป็นเครื่องมือในการประเมินเพราะสร้างง่ายและใช้สะดวก ในการใช้มาตรฐานค่ามีข้อตกลงเบื้องต้นว่าคนเป็นเครื่องมือที่ดีที่สุดในการประเมินเพราะสามารถให้ข้อความที่ถูกต้องและเป็นปรนัยสามารถตอบแบบประเมินได้ถูกต้องตรงตามประเด็นที่นักประเมินต้องการถึงกระนั้นแม้ว่านักประเมินจะสามารถเชื่อมั่นในตัวผู้ประเมิน (rater) แต่ต้องตระหนักว่าอาจเกิดความคลาดเคลื่อนอันเนื่องมาจากความลำเอียงในการประเมิน อันเนื่องมาจากผู้ประเมิน (Guilford, 1954)

Guilford (1954) ได้แบ่งความคลาดเคลื่อนในการประเมิน ออกเป็น 6 ชนิดคือ

1. ความคลาดเคลื่อนแบบกต/ปล่อยคะแนน (leniency error) ความคลาดเคลื่อนชนิดนี้เป็นคำศัพท์เรียกรวม ๆ ความคลาดเคลื่อนอันเนื่องมาจากผู้ประเมินที่มีลักษณะเป็นความคลาดเคลื่อนคงที่ ที่ผู้ประเมินให้คุณค่าของผู้ถูกประเมินดีเกินความเป็นจริง แสดงว่าเกิดความคลาดเคลื่อนแบบกต/ปล่อยคะแนนเชิงบวก (positive leniency error) ในทางตรงกันข้าม

หากผู้ถูกประเมินได้รับการประเมินคุณค่าต่ำจนเกินความเป็นจริงแสดงว่าเกิดความคลาดเคลื่อนแบบกด/ปล่อยคะแนนเชิงลบ (negative leniency error)

2. ความคลาดเคลื่อนแบบแนวโน้มเข้าสู่ส่วนกลาง (central tendency error) ความคลาดเคลื่อนชนิดนี้เกิดขึ้นเมื่อผู้ประเมินพยายามหลีกเลี่ยงที่จะประเมินคุณค่าให้มากไปทางใดทางหนึ่งจนเกินไปประกอบกับว่าผู้ประเมินอาจจะไม่รู้จักผู้ถูกประเมินดีพอจึงมีผลให้ผู้ประเมินพยายามประเมินให้คุณค่าในระดับปานกลางอย่างสม่ำเสมอ

3. ความคลาดเคลื่อนแบบฮาโล (halo error) ความคลาดเคลื่อนชนิดนี้เกิดขึ้นเมื่อผู้ถูกประเมินมีอิทธิพลต่อผู้ประเมินโดยผู้ประเมินจะให้คุณค่ามากหรือน้อยขึ้นอยู่กับความพอใจของผู้ประเมินต่อผู้ถูกประเมิน ความคลาดเคลื่อนแบบฮาโลมักจะพบในการประเมินคุณลักษณะ (trait) ที่มีลักษณะดังนี้ (1) เป็นคุณลักษณะที่สังเกตได้ยาก (2) เป็นคุณลักษณะที่ไม่ค่อยเป็นที่รู้จัก (3) เป็นคุณลักษณะที่ไม่มีนิยามที่ชัดเจน (4) เป็นคุณลักษณะที่ต้องอาศัยปฏิบัติการโต้ตอบจากผู้อื่น (5) เป็นคุณลักษณะที่เกี่ยวข้องกับศีลธรรมจรรยา Guilford (1954) ได้เสนอแนะแนวทางในการแก้ไขความคลาดเคลื่อนแบบฮาโล โดยให้ผู้ประเมินทำการประเมินผู้ถูกประเมินทุกคนทีละคุณลักษณะ

4. ความคลาดเคลื่อนแบบตรรกวิทยา (logical error) ความคลาดเคลื่อนชนิดนี้เกิดขึ้นเมื่อผู้ประเมินนำตัวแปรอื่นที่ตนเองคิดว่ามีความสัมพันธ์กับตัวแปรที่ประเมินมามีส่วนในการประเมินด้วย ความคลาดเคลื่อนประเภทนี้มิได้เกิดจากความลำเอียงแต่เกิดจากการที่ผู้ประเมินได้รับรู้รับฟังเรื่องราวของผู้ถูกประเมินมาก่อน

ความคลาดเคลื่อนแบบฮาโล (halo error) และความคลาดเคลื่อนแบบตรรกวิทยา (logical error) มีลักษณะใกล้เคียงกัน ข้อแตกต่างอยู่ที่กรณีที่เป็นความคลาดเคลื่อนแบบฮาโล คุณลักษณะที่ประเมินมีความสัมพันธ์กันตามลักษณะบุคคล ส่วนในกรณีที่เป็นความคลาดเคลื่อนแบบตรรกวิทยา คุณลักษณะที่ประเมินมีความสัมพันธ์กันในเชิงตรรกโดยไม่จำเป็นต้องมีความสัมพันธ์กันตามลักษณะบุคคล

5. ความคลาดเคลื่อนจากความใกล้เคียง (proximity error) ความคลาดเคลื่อนชนิดนี้เกิดขึ้นเมื่อ มีการประเมินข้อความที่แสดงคุณลักษณะ 2 คุณลักษณะ (trait) อยู่ใกล้กันหรือนำมาใช้ในการประเมินในเวลาใกล้เคียงกัน ผลการประเมินจะมีค่าสูง แต่ถ้านำคุณลักษณะที่มีความคล้ายคลึงกันนี้มาวางให้ห่างกันหรือทิ้งช่วงเวลาในการประเมินผลการประเมินที่ได้จะมีค่าต่ำลง ผู้ประเมินสามารถแก้ไขความคลาดเคลื่อนจากความใกล้เคียงได้โดยสร้างเครื่องมือให้ข้อที่ประเมินคุณลักษณะคล้ายกันอยู่ห่างกัน

6. ความคลาดเคลื่อนแบบขัดแย้ง (contrast error) ความคลาดเคลื่อนชนิดนี้เกิดขึ้นเมื่อผู้ประเมินให้ระดับคุณค่าแก่ผู้ถูกประเมินตรงกันข้ามกับคุณลักษณะของตน กล่าวคือผู้ประเมินที่คิดว่าตนเองมีคุณลักษณะ (trait) ที่ประเมินสูงมีแนวโน้มที่จะประเมินผู้ถูกประเมินให้มีลักษณะที่ต่ำกว่าที่ตนเองมีอยู่ และในทางตรงกันข้ามผู้ประเมินที่คิดว่าตนเองมีคุณลักษณะที่ประเมินต่ำมีแนวโน้มที่จะประเมินผู้ถูกประเมินสูงกว่าที่ตนเองมีอยู่

นอกจากนี้ Guilford ยังได้เสนอแนะเพิ่มเติมว่าการให้ความรู้เรื่องความคลาดเคลื่อนแก่ผู้ประเมิน (training rater) จะสามารถช่วยลดความคลาดเคลื่อนอันเนื่องมาจากผู้ประเมิน

Selltiz (1959) ได้แบ่งแหล่งความคลาดเคลื่อนในการประเมินมี 3 แหล่งดังนี้

1. ความคลาดเคลื่อนที่เกิดจากกระบวนการประเมิน ความคลาดเคลื่อนชนิดนี้เป็นความคลาดเคลื่อนที่เกิดจากข้อบกพร่องในการประเมินได้แก่การกำหนดคุณลักษณะหรือสิ่งที่ต้องการประเมินนั้นมีจำนวนมากเกินไปแต่มีผู้ประเมินเพียงคนเดียวจึงส่งผลให้การประเมินเกิดความคลาดเคลื่อนได้ ข้อบกพร่องอีกประการหนึ่งคือ กรณีที่ใช้ผู้ประเมิน (rater) เพียงคนเดียวทำการประเมินผู้ถูกประเมิน (ratee) หลาย ๆ คน สามารถก่อให้เกิดความคลาดเคลื่อนได้เช่นกัน ควรแก้ไขโดยใช้ผู้ประเมินหลาย ๆ คน อันเป็นวิธีที่ช่วยลดความคลาดเคลื่อนที่เกิดจากผู้ประเมินแต่ละคนได้ ผลการประเมินที่ได้จะมีความแม่นยำมากขึ้น

2. ความคลาดเคลื่อนที่เกิดจากเครื่องมือที่ใช้ในการประเมิน ความคลาดเคลื่อนชนิดนี้เป็นความคลาดเคลื่อนที่เกิดจากการใช้เครื่องมือที่เป็นมาตรฐานค่าที่ขาดการนิยามหรือการบรรยายคุณลักษณะ (characteristic) ของสิ่งที่ประเมินไว้อย่างชัดเจน ทำให้เกิดความเข้าใจที่ไม่ตรงกันระหว่างผู้ประเมิน เป็นผลให้การประเมินนั้นขาดความเที่ยง หรืออาจเป็นเพราะภาษาที่ใช้ไม่เหมาะสมหรือไม่สามารถบอกสิ่งที่ต้องการวัดได้ เป็นต้น

3. ความคลาดเคลื่อนที่เกิดจากตัวผู้ประเมิน (rater) แบ่งเป็น 3 ประเภท ดังมีรายละเอียดดังนี้

3.1 ความคลาดเคลื่อนที่เกิดจากความลำเอียงหรือความมีอคติส่วนบุคคลของผู้ประเมิน (personal bias) เป็นความคลาดเคลื่อนที่เกิดขึ้นเฉพาะลักษณะนิสัยประจำตัวของผู้ประเมินที่มักมีรูปแบบการประเมินตายตัว (fixed pattern) ซึ่งมี 3 ชนิดคือ

3.1.1 ความคลาดเคลื่อนแบบกด/ปล่อยคะแนนในเชิงบวก (positive leniency error generosity error) เกิดขึ้นเมื่อผู้ประเมิน ประเมินค่าคุณลักษณะต่าง ๆ ของผู้ถูกประเมินทุกคนในระดับที่สูงกว่าความเป็นจริงในลักษณะที่ค่อนข้างบวกเพียงด้านเดียวอย่างสม่ำเสมอ

3.1.2 ความคลาดเคลื่อนแบบกด/ปล่อยคะแนนในเชิงลบ (negative leniency error) หรือความคลาดเคลื่อนอันเนื่องมาจากความเข้มงวด (severity error) เกิดขึ้นเมื่อผู้ประเมินประเมินค่าคุณลักษณะต่าง ๆ ของผู้ถูกประเมินทุกคนต่ำกว่าความเป็นจริงในลักษณะที่ค่อนข้างไปทางลบเพียงด้านเดียวอย่างสม่ำเสมอ

3.1.3 ความคลาดเคลื่อนแบบแนวโน้มเข้าสู่ส่วนกลาง (central tendency error) คือความคลาดเคลื่อนที่เกิดขึ้นเมื่อผู้ประเมินพยายามหลีกเลี่ยงการประเมินค่าทางระดับค่อนข้างไปทางบวกและลบ มักให้คะแนนระดับปานกลางอย่างสม่ำเสมอ ทำให้ค่าประเมินที่ได้มีค่าต่ำกว่าหรือสูงกว่าสภาพที่เป็นจริง สาเหตุของความคลาดเคลื่อนชนิดนี้อาจเนื่องมาจากข้อใดข้อหนึ่งต่อไปนี้

3.1.3.1 ผู้ประเมินไม่ทราบว่าสิ่งที่ประเมินนั้นมีพิสัยอยู่ในทางสูงหรือทางต่ำ

3.1.3.2 ผู้ประเมินไม่มีความมั่นใจอย่างยิ่งในการจำแนกความแตกต่างของสิ่งที่ประเมิน

3.1.3.3 บุคลิกของผู้ประเมินไม่มั่นคงและไม่กล้าที่จะประเมินค่าในทางสูงหรือทางต่ำทางใดทางหนึ่ง

3.1.3.4 ผู้ประเมินรู้สึกที่ไม่สามารถที่จะตัดสินใจได้อย่างเป็นธรรมและเกรงว่าจะทำร้ายจิตใจคน

3.1.3.5 ผู้ประเมินเกรงว่าการที่ประเมินค่าเพียงระดับใดระดับหนึ่งเพียงระดับเดียวนั้นจะก่อให้เกิดผลอันไม่พึงปรารถนา การที่ผู้ประเมิน ประเมินค่าเพียงระดับใดระดับหนึ่งเพียงระดับเดียวนั้น จะก่อให้เกิดผลอันไม่พึงปรารถนา 2 ประการ (Gronlund, 1981) คือ

ประการที่ 1 ทำให้ค่าที่ประเมินไม่ตรงตามสภาพเนื่องจากผู้ประเมินมักจะประเมินทางสูง หรือทางต่ำ หรือประเมินกลาง ๆ อย่างใดอย่างหนึ่งเพียงระดับเดียวผลการประเมินจึงเป็นการสะท้อนถึงลักษณะการมองบุคคลของผู้ประเมิน (personal outlook) มากกว่าที่จะสะท้อนคุณลักษณะของผู้ถูกประเมิน

ประการที่ 2 การที่ผู้ประเมิน มักจะประเมินค่าเพียงระดับใดระดับหนึ่งเพียงระดับเดียวเป็นการจำกัดขอบเขตในการประเมินทำให้ผลการประเมินมีอำนาจจำแนกต่ำลงและความเที่ยงต่ำลงด้วย

3.2 ความคลาดเคลื่อนแบบฮาโล (halo error) ความคลาดเคลื่อนชนิดนี้ เป็นความคลาดเคลื่อนที่เกิดขึ้นเมื่อความรู้สึกของผู้ประเมินที่มีต่อผู้ถูกประเมินมีอิทธิพลต่อการประเมิน

คุณค่าลักษณะของผู้ถูกประเมิน เช่น ถ้าผู้ประเมินมีความชอบพอบต่อผู้ถูกประเมินมาก่อนแล้ว ผู้ประเมินมักจะประเมินคุณลักษณะต่าง ๆ ของผู้ถูกประเมินในทางบวก หรือในทางตรงกันข้ามถ้า ผู้ประเมินมีความไม่พึงพอใจต่อผู้ถูกประเมินมาก่อนแล้ว ผู้ประเมินมักจะประเมินคุณลักษณะต่าง ๆ ของผู้ถูกประเมินในทางลบ

3.3 ความคลาดเคลื่อนแบบตรรกวิทยา (logical error) ความคลาดเคลื่อนชนิดนี้เป็นความคลาดเคลื่อนที่เกิดขึ้นเมื่อผู้ประเมินได้รับรู้รับฟังเรื่องราวของผู้ถูกประเมินมาก่อนและนำตัวแปรอื่นที่คิดว่ามีความสัมพันธ์กับตัวแปรที่กำลังประเมินมามีส่วนในการประเมินด้วย เช่น การประเมินความฉลาด ครูมักจะประเมินค่าความฉลาดของนักเรียนจากผลสัมฤทธิ์ทางการเรียนของนักเรียน ถ้านักเรียนมีผลสัมฤทธิ์ทางการเรียนในระดับสูงนักเรียนคนนั้นมักจะได้รับการประเมินความฉลาดในระดับสูง ทั้งนี้เพราะครูมีความเชื่อว่า คุณลักษณะทั้งสองคุณลักษณะคือ ความฉลาดกับผลสัมฤทธิ์ทางการเรียนมีความสัมพันธ์กัน เป็นต้น

ความคลาดเคลื่อนอันเนื่องมาจากผู้ประเมิน (rater) มีหลายชนิด แต่ความคลาดเคลื่อนที่มีระบบและสามารถบอกทิศทางได้ว่า มีคะแนนสูงหรือต่ำกว่าความเป็นจริงแบ่งเป็น 3 ชนิดคือ ความคลาดเคลื่อนแบบกต/ปล่อยคะแนน (leniency error) ความคลาดเคลื่อนฮาโล (halo error) และความคลาดเคลื่อนแบบขัดแย้ง (contrast error) (Guilford, 1954) Guilford ได้สร้างโมเดลสำหรับการประมาณค่าความคลาดเคลื่อนทั้งสามประเภทโดยมีข้อตกลงเบื้องต้นว่า โมเดลเป็นแบบบวก (additive model) โดยมีสมการแสดงความสัมพันธ์ระหว่างผลการประเมินคุณค่าที่วัดได้ ผลการประเมินคุณค่าที่แท้จริงและความคลาดเคลื่อนอื่น ๆ ดังสมการ

$$X_{ijk} = X_{ijt} + X_{kl} + X_{ki} + X_{kj} + X_{ijk_r} \quad \text{-----} (2)$$

$$\text{หรือ} \quad X_{ijk} = X_{ijt} + X_{ijk_{re}} \quad \text{-----} (3)$$

$$\text{เมื่อ} \quad X_{ijk_{re}} = X_{kl} + X_{ki} + X_{kj} + X_{ijk_r} \quad \text{-----} (4)$$

ในที่นี้ X_{kl} = ความคลาดเคลื่อนแบบกต/ปล่อยคะแนน (leniency error = l) ของผู้ประเมินคนที่ k

X_{ki} = ความคลาดเคลื่อนแบบฮาโล (halo error) ของผู้ถูกประเมินคนที่ i โดย ผู้ประเมินคนที่ k

- X_{kj} = ความคลาดเคลื่อนแบบขัดแย้ง (contrast error) ของการประเมินคุณลักษณะที่ j โดยผู้ประเมินคนที่ k
- X_{ijk} = ความคลาดเคลื่อนอื่น ๆ (residual = r) จากการประเมินผู้ถูกประเมินคนที่ i ในการประเมินคุณลักษณะที่ j โดยผู้ประเมินคนที่ k นอกเหนือจากความคลาดเคลื่อนแบบกด/ปล่อยคะแนน (leniency error) ความคลาดเคลื่อนแบบฮาโล (halo error) และความคลาดเคลื่อนแบบขัดแย้ง (contrast error)

เหตุที่สมการที่ (2) และสมการที่ (4) มีเพียงความคลาดเคลื่อน 3 ชนิดคือ ความคลาดเคลื่อนแบบกด/ปล่อยคะแนน (leniency error) ความคลาดเคลื่อนแบบฮาโล (halo error) และความคลาดเคลื่อนแบบขัดแย้ง (contrast error) เนื่องจากความคลาดเคลื่อนทั้ง 3 ชนิดนี้สามารถบอกทิศทางได้ชัดเจนว่าให้คะแนนสูงหรือต่ำกว่าความเป็นจริงในสมการนี้ไม่มีความคลาดเคลื่อน อีก 3 ชนิดคือความคลาดเคลื่อนแบบแนวโน้มเข้าสู่ส่วนกลาง (central tendency error) ความคลาดเคลื่อนแบบตรรกวิทยา (logical error) และความคลาดเคลื่อนจากความใกล้เคียง (proximity error) เพราะว่าความคลาดเคลื่อนแบบแนวโน้มเข้าสู่ส่วนกลาง (central tendency error) สามารถเป็นได้ทั้งความคลาดเคลื่อนเชิงบวกหรือเชิงลบก็ได้ไม่สามารถสังเกตได้ชัดเจนจึงจัดอยู่ในส่วนคะแนนความคลาดเคลื่อนที่เหลือ (residual error) ส่วนความคลาดเคลื่อนแบบตรรกวิทยา (logical error) และความคลาดเคลื่อนจากความใกล้เคียง (proximity error) ความคลาดเคลื่อนที่เกิดขึ้นไม่ได้อยู่ในรูปคะแนนที่เพิ่มขึ้นแต่จะไปมีผลต่อสหสัมพันธ์ระหว่างคุณลักษณะโดย ในการวิจัยครั้งนี้ผู้วิจัยสนใจศึกษาเฉพาะความคลาดเคลื่อนแบบกด/ปล่อยคะแนน (leniency error) ความคลาดเคลื่อนแบบฮาโล (halo error)

จากโมเดลของ Guilford ในสมการที่ (2) เมื่อพิจารณาความแปรปรวนของผลการประเมินคุณค่าที่สังเกตได้ (X_{ijk}) ผู้วิจัยพบว่าความแปรปรวนของผลการประเมินที่มีมาจากความแตกต่างของผู้ประเมิน (rater) จากความแตกต่างของผู้ถูกประเมิน (ratee) จากความแตกต่างของคุณลักษณะ (trait) ของผู้ประเมิน และจากความคลาดเคลื่อนแบบสุ่ม (random error) เมื่อนำมาเขียนเป็นสมการตามโมเดลการวิเคราะห์ความแปรปรวน (ANOVA model) เมื่อไม่มีอิทธิพลปฏิสัมพันธ์จะได้ดังสมการ

$$X_{ijk} = X_{it} + \alpha_k + \beta_i + \gamma_j + e_{ijk} \quad \text{-----} (5)$$

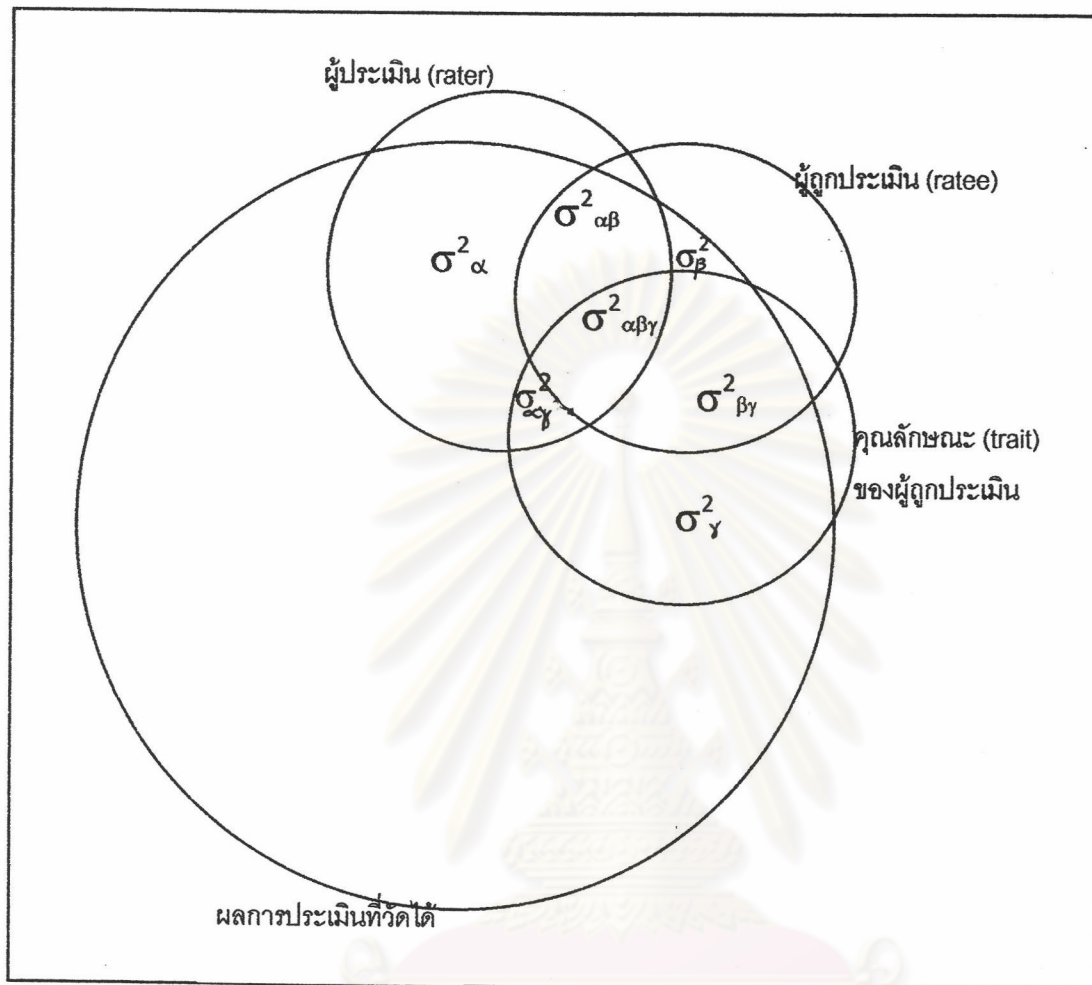
ในกรณีที่มีการวิเคราะห์อิทธิพลปฏิสัมพันธ์ด้วยโมเดลการวิเคราะห์ความแปรปรวนในสมการที่ (5) จะเขียนได้ดังสมการ

$$X_{ijk} = X_{ijt} + \alpha_k + \beta_i + \gamma_j + (\alpha_k\beta_i) + (\alpha_k\gamma_j) + (\beta_i\gamma_j) + (\alpha_k\beta_i\gamma_k) + e'_{ijk} \quad (6)$$

- ในที่นี้ α_k = อิทธิพลหลักของผู้ประเมิน (rater main effect) ที่ k
 β_i = อิทธิพลหลักของผู้ถูกประเมิน (ratee main effect) ที่ i
 γ_j = อิทธิพลหลักของคุณลักษณะ (trait) ที่ j ของผู้ถูกประเมิน
 $\alpha_k\beta_i$ = อิทธิพลปฏิสัมพันธ์ระหว่างผู้ประเมินที่ k กับผู้ถูกประเมินที่ i
 $\alpha_k\gamma_j$ = อิทธิพลปฏิสัมพันธ์ระหว่างผู้ประเมินที่ k กับคุณลักษณะที่ j ของผู้ถูกประเมิน
 $\beta_i\gamma_j$ = อิทธิพลปฏิสัมพันธ์ระหว่างผู้ถูกประเมินที่ i กับคุณลักษณะที่ j ของผู้ถูกประเมิน
 $\alpha_k\beta_i\gamma_j$ = อิทธิพลปฏิสัมพันธ์ระหว่างผู้ประเมินที่ k กับผู้ถูกประเมินที่ i กับคุณลักษณะที่ j ของผู้ถูกประเมิน
 e'_{ijk} = ความคลาดเคลื่อนสุ่ม

การวิเคราะห์ความแปรปรวนของผลการประเมินที่วัดได้ตามสมการที่ (6) แสดงได้ดังภาพและสมการต่อไปนี้

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย



ภาพที่ 1 ความแปรปรวนของผลการประเมินที่วัดได้ เมื่ออธิบายด้วยความแปรปรวนของผู้ประเมิน ผู้ถูกประเมิน และคุณลักษณะของผู้ประเมิน

ศูนย์วิทยาศาสตร์การ
จุฬาลงกรณ์มหาวิทยาลัย

$$\sigma^2 = \sigma^2_T + \sigma^2_\alpha + \sigma^2_\beta + \sigma^2_\gamma + \sigma^2_{\alpha\beta} + \sigma^2_{\alpha\gamma} + \sigma^2_{\beta\gamma} + \sigma^2_{\alpha\beta\gamma} + \sigma^2_e$$

----- (7)

ในที่นี้	σ^2_T	= ความแปรปรวนของผลการประเมินที่วัดได้
	σ^2_α	= ความแปรปรวนเนื่องจากอิทธิพลหลัก α_k
	σ^2_β	= ความแปรปรวนเนื่องจากอิทธิพลหลัก β_i
	σ^2_γ	= ความแปรปรวนเนื่องจากอิทธิพลหลัก γ_j
	$\sigma^2_{\alpha\beta}$	= ความแปรปรวนเนื่องจากอิทธิพลปฏิสัมพันธ์ $\alpha_k\beta_i$
	$\sigma^2_{\alpha\gamma}$	= ความแปรปรวนเนื่องจากอิทธิพลปฏิสัมพันธ์ $\alpha_k\gamma_j$
	$\sigma^2_{\beta\gamma}$	= ความแปรปรวนเนื่องจากอิทธิพลปฏิสัมพันธ์ $\beta_i\gamma_j$
	$\sigma^2_{\alpha\beta\gamma}$	= ความแปรปรวนเนื่องจากอิทธิพลปฏิสัมพันธ์ $\alpha_k\beta_i\gamma_j$
	σ^2_e	= ความแปรปรวนเนื่องจากความคลาดเคลื่อนสุ่ม

ตามโมเดลของ Guilford โดยวิธีวิเคราะห์ความแปรปรวนแบบสองทาง (two-way ANOVA) จากข้อมูลที่ประกอบด้วยผู้ประเมิน (rater) และผู้ถูกประเมิน (ratee) หลายๆ คน ทำการประเมินข้อมูลหลาย ๆ กลุ่มคุณลักษณะ (trait) ให้ค่าคะแนนความคลาดเคลื่อนได้ทั้งค่าบวกและค่าลบ โดยผลจากการวิเคราะห์ความแปรปรวนแบบสองทางแสดงถึงความคลาดเคลื่อนแต่ละชนิดดังนี้

ความคลาดเคลื่อนแบบกต/ปล่อยคะแนน (leniency error) พิจารณาได้จากการมีนัยสำคัญทางสถิติของอิทธิพลหลักผู้ประเมิน (rater main effect) หากอิทธิพลหลักผู้ประเมินมีนัยสำคัญทางสถิติ (significant) แสดงว่าคะแนนโดยรวมของผู้ประเมิน (rater) แต่ละคนแตกต่างกัน ผู้ประเมินบางคนให้คะแนนสูงกว่าคนอื่น ๆ ในทุก ๆ คุณลักษณะ (trait) ทุก ๆ ผู้ถูกประเมิน (ratee) นั่นคือมีความคลาดเคลื่อนแบบกต/ปล่อยคะแนน

ความคลาดเคลื่อนแบบฮาโล (halo error) พิจารณาได้จากการมีนัยสำคัญทางสถิติของปฏิสัมพันธ์ระหว่างผู้ประเมินกับผู้ถูกประเมิน (rater x ratee interaction) หากปฏิสัมพันธ์ระหว่างผู้ประเมินกับผู้ถูกประเมินมีนัยสำคัญทางสถิติ (significant) แสดงว่าผู้ประเมิน (rater) ประเมินผู้ถูกประเมิน (ratee) ไม่สอดคล้องกันผู้ประเมินจะประเมินให้มากกว่าหรือน้อยกว่าความเป็นจริงขึ้นอยู่กับความพอใจของผู้ประเมินที่มีต่อผู้ถูกประเมินนั่นคือมีความคลาดเคลื่อนแบบฮาโล

ความคลาดเคลื่อนแบบขัดแย้ง (contrast error) พิจารณาได้จากการมีนัยสำคัญทางสถิติของปฏิสัมพันธ์ระหว่างผู้ถูกประเมินกับคุณลักษณะที่ประเมิน (ratee x trait interaction) หากปฏิสัมพันธ์ระหว่างผู้ถูกประเมินกับคุณลักษณะที่ประเมินมีนัยสำคัญทางสถิติ (significant) แสดงว่าผู้ประเมิน (rater) แต่ละคนประเมินคุณลักษณะ (trait) เดียวกันของผู้ถูกประเมิน (ratee) ไม่สอดคล้องกันนั่นคือมีความคลาดเคลื่อนแบบขัดแย้ง

ตอนที่ 2 มโนทัศน์เกี่ยวกับการป้องกันความคลาดเคลื่อนในการประเมินผล

ความคลาดเคลื่อนที่ได้นำเสนอในหัวข้อที่ผ่านมาเป็นความคลาดเคลื่อนที่เกิดขึ้นอย่างเป็นระบบ (systematic error) ที่ทำให้ความตรงลดลง ความคลาดเคลื่อนมีระบบดังกล่าวสามารถปรับปรุงแก้ไขให้การประเมินที่ให้ผลถูกต้องและน่าเชื่อถือได้ Guilford (1954) ได้เสนอแนะให้มีการป้องกันความคลาดเคลื่อนก่อนการประเมินโดยการปรับปรุงมาตรฐานค่าให้เป็นมาตรฐานที่มีค่าเอนเอียงไปทางด้านใดด้านหนึ่ง เพื่อป้องกันไม่ให้เกิดความคลาดเคลื่อนแบบกด/ปล่อยคะแนน เนื่องจากการปรับแก้ความคลาดเคลื่อนมีขั้นตอนที่ค่อนข้างจะยุ่งยาก ดังนั้นหากสามารถป้องกันความคลาดเคลื่อนไม่ให้เกิดขึ้นหรือลดการเกิดความคลาดเคลื่อนได้จะเป็นการดีกว่า การป้องกันความคลาดเคลื่อนแบบกด/ปล่อยคะแนน (leniency error) ตามแนวคิดของ Guilford ใช้วิธีการสร้างมาตรฐานค่าแบบอสมมาตรให้มีระดับของมาตรฐานค่าเอนเอียงไปทางด้านบวก (positive or favorable) หรือทางด้านลบ (negative or unfavorable) อย่างใดอย่างหนึ่งดังตัวอย่าง

ต้องปรับปรุง	พอใช้	ดี	ดีมาก	ดีเยี่ยม
--------------	-------	----	-------	----------

ดังนั้นผู้วิจัยจึงได้ใช้มาตรฐานค่าเป็นวิธีหนึ่งในการป้องกันความคลาดเคลื่อนในการประเมินโดยสนใจที่จะตรวจสอบว่า มาตรฐานค่าใช้ป้องกันความคลาดเคลื่อนแบบกด/ปล่อยคะแนน (leniency error) ได้หรือไม่

Thomdike และ Hagen (1997) ได้สรุปหลักการในการป้องกันความคลาดเคลื่อนไว้เป็นแนวทางที่ควรปฏิบัติ ดังต่อไปนี้

1. การประเมินควรจะทำเฉพาะลักษณะที่ปรากฏให้เห็นชัดเจน
2. ลักษณะที่จะประเมินต้องแสดงออกมาให้เห็นในรูปของพฤติกรรมและการตัดสินใจกระทำจากพฤติกรรมที่เห็นนั้นสามารถบรรยายละเอียดของพฤติกรรมที่จะประเมินได้อย่างชัดเจน

3. รูปแบบของมาตราประมาณค่า (rating form) ต้องปรับปรุงเพื่อบังคับให้ผู้ประเมินบอกความแตกต่างหรือเพื่อควบคุมความแตกต่างของมาตรฐานในการตัดสินระหว่างผู้ประเมินแต่ละคน
4. ผู้ประเมินต้องเป็นผู้ที่มีโอกาสมากที่สุดในการสังเกตพฤติกรรมในสถานการณ์ตามที่ต้องการ
5. ผู้ประเมินต้องได้รับการบอกเล่าถึงคุณค่าของการประเมินค่าและได้รับการฝึกฝนการใช้เครื่องมือ
6. ถ้าสามารถมีผู้ประเมินหลายคนก็ควรรวบรวมผลการประเมินจากผู้ประเมินหลาย ๆ คน

นอกจากนี้ Thorndike และ Hagen (1997) ได้กล่าวโดยสรุปได้ว่ามาตราประมาณค่าแบบกราฟฟิก (graphic rating scale) มีลักษณะเป็นมาตราที่ใช้คำคุณศัพท์บ่งบอกถึงความแตกต่างของแต่ละระดับของมาตราทำให้ผู้ประเมินอาจกำหนดเกณฑ์แตกต่างกันไปตามลักษณะการให้คะแนนซึ่งเชื่อต่อการ ผู้ประเมินจะลำเอียงให้คุณค่าผู้ถูกประเมินมากหรือน้อยขึ้นอยู่กับความพอใจของผู้ประเมินต่อผู้ถูกประเมิน ซึ่งต่างจากการใช้มาตราประมาณค่าเชิงพฤติกรรม (BARS) ที่มีลักษณะเป็นข้อความอธิบายถึงความแตกต่างของแต่ละระดับของมาตราทำให้ผู้ประเมินมีเกณฑ์ร่วมกันในการประเมินและตรงกับเกณฑ์ของผู้ทำการวิจัย ดังนั้นการใช้มาตราประมาณค่าเชิงพฤติกรรม (BARS) จึงควรจะป้องกันการให้คะแนนที่มีความมีความคลาดเคลื่อนซึ่งเกิดจากการมีอคติ ลำเอียง หรือความคลาดเคลื่อนอันเนื่องมาจากความสนิทสนมได้ จากข้อเสนอแนะของ Thorndike และ Hagen ที่กล่าวข้างต้นแสดงให้เห็นว่าถ้าผู้ประเมินมีหลายคนและผู้ประเมินได้ทำการประเมินลักษณะพฤติกรรม โดยมีรายละเอียดของพฤติกรรมที่ชัดเจนและมีเกณฑ์ประมาณค่าเชิงพฤติกรรม (Behavior Anchor Rating Scale; BARS) ที่ชัดเจนน่าจะทำให้มีความคลาดเคลื่อนในการประเมินน้อยลง ผู้วิจัยเห็นว่าความคลาดเคลื่อนในลักษณะที่ Thorndike และ Hagen กล่าวข้างต้นเป็นลักษณะของความคลาดเคลื่อนแบบฮาโล ซึ่งการใช้มาตราประมาณค่าเชิงพฤติกรรม (BARS) ควรที่จะป้องกันความคลาดเคลื่อนแบบฮาโลได้

ดังนั้นผู้วิจัยจึงสนใจศึกษาการใช้มาตราประมาณค่าเชิงพฤติกรรม (BARS) เป็นวิธีหนึ่งในการป้องกันความคลาดเคลื่อนในการประเมินโดยสนใจที่จะตรวจสอบว่า มาตราประมาณค่าเชิงพฤติกรรม (BARS) ใช้ป้องกันความคลาดเคลื่อนแบบฮาโล (halo error) ได้หรือไม่

โดยสรุปในงานวิจัยนี้สนใจศึกษาเปรียบเทียบคุณภาพของวิธีป้องกันความคลาดเคลื่อน 2 วิธี คือ (1) วิธีป้องกันความคลาดเคลื่อนแบบกด/ปล่อยคะแนนโดยการใช้มาตราแบบบอสมมาตร และ(2) วิธีป้องกันความคลาดเคลื่อนแบบฮาโลโดยการใช้มาตราประเมินค่าเชิงพฤติกรรม (BARS)

เนื่องจากการป้องกันความคลาดเคลื่อนตามเสนอแนะของ Thomdike และ Hagen ได้แก่ การกำหนดให้มีผู้ประเมินหลายคน แต่เนื่องจากยังไม่มีรายงานการวิจัยเป็นหลักฐานสนับสนุนว่า การใช้จำนวนผู้ประเมินหลายจะเป็นวิธีการป้องกันความคลาดเคลื่อนที่ได้ผลจริงหรือไม่ ประกอบกับผู้วิจัยคาดว่าจำนวนผู้ประเมินไม่สามารถป้องกันความคลาดเคลื่อนแต่อย่างได้เหมือนวิธีของ Guilford ดังนั้นในการวิจัยครั้งนี้ผู้วิจัยจึงได้ใช้ตัวแปร จำนวนผู้ประเมินเป็นตัวแปรในการควบคุม

ตอนที่ 3 มโนทัศน์เกี่ยวกับการประมาณค่าความคลาดเคลื่อนในการประเมินผล

การเสนอเนื้อหาสาระในตอนนี้แบ่งเป็นการนำเสนอวิธีการประมาณค่าความคลาดเคลื่อนในการประเมินผล โดยการวิเคราะห์ความแปรปรวน ตามแนวคิดของ Guilford (1954) การประมาณค่าความคลาดเคลื่อนที่มีผลต่อการประเมิน (estimation of error contributions to ratings) สามารถทำได้เฉพาะ ความคลาดเคลื่อนแบบกด/ปล่อยคะแนน (leniency error) ความคลาดเคลื่อนแบบฮาโล (halo error) และความคลาดเคลื่อนแบบขัดแย้ง (contrast error) ดังสมการ (2) สามารถเขียนให้อยู่ในรูปของสมการประมาณค่าสมการที่ (8) ซึ่งใช้สัญลักษณ์และโมเดล ดังต่อไปนี้

$$\begin{aligned}
 X_{ijk} &= \text{ผลการประเมินคุณค่า (observed score) ของผู้ถูกประเมินคนที่ } i \text{ ในการ} \\
 &\quad \text{ประเมินคุณลักษณะที่ } j \text{ โดยผู้ประเมินคนที่ } k \\
 X'_{ijt} &= \text{ค่าประมาณผลการประเมินคุณค่าที่แท้จริง (true = t) ของผู้ถูกประเมินคนที่ } i \\
 &\quad \text{ในการประเมินคุณลักษณะที่ } j \\
 X'_{ki} &= \text{ค่าประมาณความคลาดเคลื่อนแบบกด/ปล่อยคะแนน (leniency error = l)} \\
 &\quad \text{ของผู้ประเมินคนที่ } k \\
 X'_{ki} &= \text{ค่าประมาณความคลาดเคลื่อนแบบฮาโล (halo error) ของผู้ถูกประเมินคนที่ } i \\
 &\quad \text{โดยผู้ประเมินคนที่ } k \\
 X'_{kj} &= \text{ค่าประมาณความคลาดเคลื่อนแบบขัดแย้ง (contrast error) ของคุณลักษณะที่ } j \\
 &\quad \text{โดยผู้ประเมินคนที่ } k
 \end{aligned}$$

X'_{ijk} = ค่าประมาณความคลาดเคลื่อนอื่น ๆ (residual = r) จากการประเมินผู้ถูกประเมินคนที่ i ในการประเมินคุณลักษณะที่ j โดยผู้ประเมินคนที่ k นอกเหนือจากความคลาดเคลื่อนแบบกด/ปล่อยคะแนน (leniency error) ความคลาดเคลื่อนแบบฮาโล (halo error) และความคลาดเคลื่อนแบบขัดแย้ง (contrast error)

$$X_{ijk} = X'_{ijt} + X'_{kj} + X'_{ki} + X'_{kj} + X'_{ijk} \text{ -----(8)}$$

สมการที่ (2) และ (8) มีลักษณะต่างกันตรงที่ สัญลักษณ์ที่ใช้ในสมการที่ (8) เป็นค่าประมาณ ส่วนเทอมต่าง ๆ ในสมการทั้งสองมีลักษณะตรงกัน

Guilford (1954) ได้นำเสนอการประมาณค่าความคลาดเคลื่อน 3 แบบในสมการที่ (8) ข้างต้นโดยการวิเคราะห์ความแปรปรวนสองทาง (two-way ANOVA) โดยมีข้อตกลงเบื้องต้นว่า คะแนนที่ได้จากการประเมินต้องประกอบด้วยผู้ประเมิน (rater) และผู้ถูกประเมินหลาย ๆ คน (ratee) ทำการประเมินข้อมูลหลาย ๆ กลุ่มคุณลักษณะ (trait) วิธีประมาณค่าความคลาดเคลื่อนทั้ง 3 แบบควรใช้การวิเคราะห์ความแปรปรวนแบบสองทาง โดยวิเคราะห์ 3 ครั้ง ทั้งนี้อาจเป็นเพราะเทคโนโลยีการดำเนินการในยุคนั้นยังไม่สะดวกพอที่จะวิเคราะห์ความแปรปรวนแบบสามทาง

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 1 คะแนนเฉลี่ยคะแนนประเมินของผู้ประเมินที่ k ผู้ถูกประเมินที่ i และค่าประมาณความคลาดเคลื่อนกดปล่อยคะแนน (X'_{kl})

rater \ ratee	1	2	3	k	n	all raters	X'_{kl}
1	$X_{.11}$	$X_{.12}$	$X_{.13}$	$X_{.1k}$	$X_{.1n}$		
2	$X_{.21}$	$X_{.22}$	$X_{.23}$	$X_{.2k}$	$X_{.2n}$		
3	$X_{.31}$	$X_{.32}$	$X_{.33}$	$X_{.3k}$	$X_{.3n}$		
⋮	⋮	⋮	⋮		⋮		⋮		
i	$X_{.i1}$	$X_{.i2}$	$X_{.i3}$	$X_{.ik}$	$X_{.in}$		
⋮	⋮	⋮	⋮		⋮		⋮		
m	$X_{.m1}$	$X_{.m2}$	$X_{.m3}$	$X_{.mk}$	$X_{.mn}$		

จากตารางข้างต้นนี้ จะประมาณค่าความคลาดเคลื่อนทั้ง 3 แบบได้ดังนี้

- 1) ความคลาดเคลื่อนแบบกดปล่อยคะแนน (leniency error) คืออิทธิพลหลักของผู้ประเมิน (rater main effect) ซึ่งประมาณค่าได้ตามสูตร ซึ่งประมาณค่าได้ตามสูตร
 - 2) ความคลาดเคลื่อนแบบฮาโล (halo error) คืออิทธิพลปฏิสัมพันธ์ระหว่างผู้ประเมินกับผู้ถูกประเมิน (rater x ratee interaction) ซึ่งประมาณค่าได้ตามสูตร
 - 3) ความคลาดเคลื่อนแบบขัดแย้ง (contrast error) คืออิทธิพลปฏิสัมพันธ์ระหว่างผู้ประเมินกับคุณลักษณะของผู้ถูกประเมิน (rater x trait interaction) ซึ่งประมาณค่าได้ตามสูตร
- ต่อไปนี้จะวิจันำเสนอสู่ข้อมูลและวิธีการประมาณค่าความคลาดเคลื่อนทั้ง 3 แบบของ Guilford เพื่อประมาณค่าความคลาดเคลื่อน ในกรณีที่มีผู้ประเมิน (rater) 3 คน (α, β, γ) และมีผู้ถูกประเมิน (ratee) 7 คนโดยประเมิน 5 คุณลักษณะ (A,B,C,D,E) ดังตารางต่อไปนี้

ตารางที่ 2 ผลการประเมินคุณลักษณะ 5 ด้าน ของผู้ถูกประเมิน 7 คน โดยผู้ประเมิน 3 คน

rater ratee	Trait A			Trait B			Trait C			Trait D			Trait E		
	α	β	γ	α	β	γ	α	β	γ	α	β	γ	α	β	γ
1	5	6	5	5	5	5	3	4	5	5	6	7	3	3	3
2	9	8	7	7	7	7	5	5	5	8	7	7	5	2	3
3	3	4	3	3	5	5	3	3	5	7	6	5	1	6	5
4	7	5	5	3	6	3	1	4	3	3	5	3	3	5	1
5	9	2	9	7	4	7	7	3	7	8	2	7	5	3	1
6	3	4	3	5	4	3	3	6	3	5	4	5	1	2	3
7	7	3	7	7	3	7	5	5	7	5	5	5	5	4	7

ขั้นตอนที่ 1 การประมาณค่าความคลาดเคลื่อนแบบกด/ปล่อยคะแนน (leniency error = X'_{ki}) มีขั้นตอนดังนี้

1.1 หาค่าเฉลี่ยคะแนนการประเมิน ในแต่ละเซลล์ ให้ได้ ค่าเฉลี่ยผลการประเมินทุกคุณลักษณะของผู้ถูกประเมิน ตัวอย่างเช่น คะแนนเฉลี่ยของผู้ประเมิน α ผู้ถูกประเมิน 1 มีค่าเท่ากับ $\frac{1}{5}(5+5+3+5+3)$ ซึ่งมีค่าเท่ากับ 4.20 ดังตารางที่ 3

1.2 หาค่าเฉลี่ยคะแนนประเมินของผู้ประเมิน (rater) แต่ละคน (mean all rates) ตามตารางที่ 2 ตัวอย่างเช่น คะแนนเฉลี่ยของผู้ประเมิน α มีค่าเท่ากับ $\frac{1}{7}(4.2+6.8+3.4+3.4+7.2+3.4+5.8)$ ซึ่งมีค่าเท่ากับ 4.89 ดังตารางที่ 3

1.3 หาค่าเฉลี่ยคะแนนประเมินของผู้ถูกประเมิน (ratee) แต่ละคน (mean all raters) ตามตารางที่ 2 ตัวอย่างเช่นคะแนนเฉลี่ยของผู้ถูกประเมิน 1 มีค่าเท่ากับ $\frac{1}{3}(4.2+4.8+5.0)$ ซึ่งมีค่าเท่ากับ 4.67 ดังตารางที่ 3

1.4 หาค่าเฉลี่ยรวม (grand mean) ซึ่งในที่นี้ได้เท่ากับ 4.84

1.5 ค่าประมาณความคลาดเคลื่อนแบบกด/ปล่อยคะแนน (leniency error = X'_{ki}) (mean all rates - grand mean) ตัวอย่างเช่นค่าประมาณความคลาดเคลื่อนแบบกด/ปล่อยคะแนนของผู้ประเมิน α มีค่าเท่ากับ $4.89-4.84$ ซึ่งมีค่าเท่ากับ .05 ดังตารางที่ 3

การแปลความหมายค่าความคลาดเคลื่อนแบบกด/ปล่อยคะแนนมีดังนี้

หาก $X'_{ki} = 0$ หมายถึง ผลการประเมินของผู้ประเมินคนที่ k ไม่มีความคลาดเคลื่อนแบบกต/ปล่อยคะแนน (leniency error)

หาก $X'_{ki} > 0$ หมายถึง ผลการประเมินของผู้ประเมินคนที่ k มีความคลาดเคลื่อนแบบกต/ปล่อยคะแนน โดยให้คะแนนสูงกว่าความเป็นจริง (positive leniency error)

หาก $X'_{ki} < 0$ หมายถึง ผลการประเมินของผู้ประเมินคนที่ k มีความคลาดเคลื่อนแบบกต/ปล่อยคะแนน โดยให้คะแนนต่ำกว่าความเป็นจริง (negative leniency error)

1.6 หากค่า d_i คือ ผลต่างของค่าเฉลี่ยผลการประเมินของผู้ถูกประเมินแต่ละคน (mean all raters) กับค่าเฉลี่ยรวม (grand mean) โดย $d_i = \text{mean all raters} - \text{grand mean}$

เมื่อดำเนินการตามขั้นตอนทั้ง 5 ขั้นตอนจะได้ผลดังตารางที่ 3 ซึ่งอธิบายได้ว่า ผู้ประเมิน γ มีผลการประเมินที่มีความคลาดเคลื่อนแบบกต/ปล่อยคะแนน โดยให้คะแนนสูงกว่าความเป็นจริง .33 รองลงมาคือผู้ประเมิน α มีผลการประเมินที่มีความคลาดเคลื่อนแบบกต/ปล่อยคะแนน โดยให้คะแนนสูงกว่าความเป็นจริง .05 และผู้ประเมิน β มีผลการประเมินที่มีความคลาดเคลื่อนแบบกต/ปล่อยคะแนน โดยให้คะแนนต่ำกว่าความเป็นจริง .38 ดังแสดงในตารางที่ 3

ตารางที่ 3 ผลการประมาณค่าความคลาดเคลื่อนแบบกต/ปล่อยคะแนน

ratee rater	1	2	3	4	5	6	7	All rates	X'_{ki}
α	4.20	6.80	3.40	3.40	7.20	3.40	5.80	4.89	+0.05
β	4.80	5.80	4.80	5.00	2.80	4.00	4.00	4.46	-0.38
γ	5.00	6.20	4.60	3.00	7.40	3.40	6.60	5.17	+0.33
All raters	4.67	6.27	4.27	3.80	5.80	3.60	5.47	4.84	.00
d_i	-0.17	+1.43	-0.57	-1.04	-0.96	-1.24	+0.63		

ขั้นตอนที่ 2 การประมาณค่าความคลาดเคลื่อนแบบฮาโล (halo error = X'_{ki}) มีขั้นตอนดังนี้

2.1 จากข้อมูลในตารางที่ 3 นำมาหาค่าเฉลี่ยคะแนนการประเมินในแต่ละเซลล์ ที่หักออกด้วยค่าประมาณความคลาดเคลื่อนแบบกต/ปล่อยคะแนน (X'_{ki}) และค่า d_i ตัวอย่างเช่น

จากตารางที่ 4 เมื่อนำค่าเฉลี่ยคะแนนผลการประเมินในแต่ละเซลล์เป็นตัวตั้งลบด้วยค่าเฉลี่ยรวม (grand mean) จะได้ค่าประมาณความคลาดเคลื่อนแบบฮาโล ตัวอย่างเช่นค่าประมาณความคลาดเคลื่อนแบบฮาโลของผู้ประเมิน α ผู้ถูกประเมินที่ 1 มีค่าเท่ากับ (4.32-4.84) ซึ่งมีค่าเท่ากับ -.52 ดังตารางที่ 5

ตารางที่ 5 ค่าความคลาดเคลื่อนแบบฮาโล (halo error : X'_{ki})

ratee raters	ความคลาดเคลื่อนแบบฮาโล							
	1	2	3	4	5	6	7	Σ
α	-.52	+.48	-.92	-.45	+1.35	-.25	+.28	-.03
β	+.51	-.09	+.91	+1.58	-2.62	+.78	-1.09	-.02
γ	.00	-.40	.00	-1.13	+1.27	-.53	+.80	+.01
Σ	-.01	-.01	-.01	.00	.00	.00	-.01	-.04

ขั้นตอนที่ 3 การประมาณค่าความคลาดเคลื่อนแบบขัดแย้ง (X'_{ki})

นอกจากความคลาดเคลื่อนแบบกด/ปล่อยคะแนนและความคลาดเคลื่อนแบบฮาโลแล้ว Guilford (1954) ยังได้เสนอวิธีการประมาณค่าความคลาดเคลื่อนแบบขัดแย้ง ซึ่งใช้การวิเคราะห์ความแปรปรวนแบบสองทาง โดยมีแนวความคิดว่าหากอิทธิพลปฏิสัมพันธ์ระหว่างผู้ประเมินและคุณลักษณะที่ประเมินมีนัยสำคัญทางสถิติแสดงว่ามีความคลาดเคลื่อนแบบขัดแย้ง (Guilford 1954) วิธีการประมาณค่าความคลาดเคลื่อนแบบขัดแย้งตามวิธี Guilford เริ่มต้นจากการนำข้อมูล (จากตารางที่ 2) มานำเสนอในรูปตารางการวิเคราะห์ความแปรปรวนแบบสองทางจำแนกตามผู้ประเมินและคุณลักษณะที่ประเมิน ตามตารางที่ 6 โดยดำเนินการดังนี้

3.1 หาค่าเฉลี่ยคะแนนการประเมิน ในแต่ละเซลล์ ให้ได้ ค่าเฉลี่ยผลการประเมินทุกคุณลักษณะของผู้ถูกประเมิน ตัวอย่างเช่น คะแนนเฉลี่ยของผู้ประเมิน α คุณลักษณะ A มีค่าเท่ากับ $\frac{1}{5}(5+9+3+7+9+3+7)$ ซึ่งมีค่าเท่ากับ 6.14 ดังตารางที่ 6

3.2 หาค่าเฉลี่ยคะแนนประเมินของผู้ประเมิน (rater) ได้ค่าเฉลี่ยในคอลัมน์รวมทุกคุณลักษณะ (all traits) ตามตารางที่ 5 ตัวอย่างเช่น คะแนนเฉลี่ยของผู้ประเมิน α มีค่าเท่ากับ $\frac{1}{5}(6.14+5.29+3.86+5.86+3.29)$ ซึ่งมีค่าเท่ากับ 4.89 ดังตารางที่ 6

3.3 หาค่าเฉลี่ยคะแนนประเมินของคุณลักษณะ (trait) ได้ค่าเฉลี่ยในแถว รวมทุกผู้ประเมิน (all raters) ตามตารางที่ 6 ตัวอย่างเช่นคะแนนเฉลี่ยของคุณลักษณะ A มีค่าเท่ากับ $\frac{1}{3}(6.14+4.57+5.57)$ ซึ่งมีค่าเท่ากับ 5.43

3.4 หาค่าเฉลี่ยรวม (grand mean) ซึ่งในที่นี้ได้เท่ากับ 4.84

3.5 ค่าประมาณความคลาดเคลื่อนแบบกต/ปล่อยคะแนน (leniency error = X'_{ki}) (mean all rates - grand mean) ตัวอย่างเช่นค่าประมาณความคลาดเคลื่อนแบบกต/ปล่อยคะแนนของผู้ประเมิน α มีค่าเท่ากับ 4.89-4.84 ซึ่งมีค่าเท่ากับ .05 ดังตารางที่ 6

ตารางที่ 6 ผลการประมาณค่าความคลาดเคลื่อนแบบกต/ปล่อยคะแนน

traits rater	A	B	C	D	E	All traits	X'_{ki}
α	6.14	5.29	3.86	5.86	3.29	4.89	+0.05
β	4.57	4.86	4.29	5.00	3.57	4.46	-.38
γ	5.57	5.29	5.00	5.57	4.43	5.17	+.33
All traits	5.43	5.14	4.38	5.48	3.76	4.84	.00
d_i	+.59	+.30	-.46	+.64	-1.08		

จากตารางที่ 6 แสดงผลการประมาณค่าความคลาดเคลื่อนแบบกต/ปล่อยคะแนนซึ่งอธิบายได้ว่าผลการประเมินของผู้ประเมิน γ มีความคลาดเคลื่อนแบบกต/ปล่อยคะแนน โดยให้คะแนนสูงกว่าความเป็นจริง .33 รองลงมาคือผู้ประเมิน α มีผลการประเมินที่มีความคลาดเคลื่อนแบบกต/ปล่อยคะแนน โดยให้คะแนนสูงกว่าความเป็นจริง .05 และผู้ประเมิน β มีผลการประเมินที่มีความคลาดเคลื่อนแบบกต/ปล่อยคะแนน โดยให้คะแนนต่ำกว่าความเป็นจริง .38 เหมือนผลการวิเคราะห์ในตารางที่ 3

3.6 หาค่า d_i จาก mean all raters - grand mean

3.7 หาค่าเฉลี่ยคะแนนการประเมินในแต่ละเซลล์ ที่หักออกด้วยค่าประมาณความคลาดเคลื่อนแบบกต/ปล่อยคะแนน (X'_{ki}) และค่า d_i ตัวอย่างเช่น คะแนนผลการประเมินของผู้ประเมิน α คุณลักษณะ A มีค่าเท่ากับ $6.14 - (0.05) - (-0.17)$ ซึ่งมีค่าเท่ากับ 5.50 ดังตารางที่ 7

ตารางที่ 7 ค่าเฉลี่ยผลการประเมินที่จำแนกตามคุณลักษณะที่ประเมินปรับแก้ด้วยค่า X'_{kj} และค่า d_j

traits rater	A	B	C	D	E	All traits
α	5.50	4.94	4.27	5.17	4.32	4.84
β	4.36	4.94	5.13	4.74	5.03	4.84
γ	4.65	4.66	5.13	4.60	5.18	4.84
All traits	4.84	4.84	4.84	4.84	4.84	4.84

เมื่อนำผลการประเมินในแต่ละเซลล์จากตารางที่ 5 ลบด้วยค่าประมาณความคลาดเคลื่อนแบบกต/ปล่อยคะแนน (X'_{kj}) และค่า d_j จะได้ค่าประมาณผลการประเมินที่ปรับแก้ด้วยความคลาดเคลื่อนแบบกต/ปล่อยคะแนน

3.8 นำผลจากขั้นตอน 3.7 ลบด้วยค่าเฉลี่ยรวม (grand mean) ตัวอย่างเช่น คะแนนผลการประเมินของผู้ประเมิน α คุณลักษณะ A มีค่าเท่ากับ $5.50 - 4.84$ ซึ่งมีค่าเท่ากับ 0.66 ดังตารางที่ 8 ค่าที่คำนวณได้แต่ละเซลล์ คือค่าประมาณความคลาดเคลื่อนแบบขัดแย้ง (contrast error = X'_{kj}) ซึ่งมีการแปลผลดังนี้

- หาก $X'_{kj} = 0$ หมายถึง ผลการประเมินคุณลักษณะที่ j ของผู้ประเมินคนที่ k ไม่มี ความคลาดเคลื่อนแบบขัดแย้ง (contrast error)
- หาก $X'_{kj} > 0$ หมายถึง ผลการประเมินคุณลักษณะที่ j ของผู้ประเมินคนที่ k ไม่มี ความคลาดเคลื่อนแบบขัดแย้ง (contrast error) โดยให้ คะแนนสูงกว่าความเป็นจริง (overvalue)
- หาก $X'_{kj} < 0$ หมายถึง ผลการประเมินคุณลักษณะที่ j ของผู้ประเมินคนที่ k ไม่มี ความคลาดเคลื่อนแบบขัดแย้ง (contrast error) โดยให้ คะแนนต่ำกว่าความเป็นจริง (undervalue)

ตารางที่ 8 ค่าประมาณความคลาดเคลื่อนแบบขัดแย้ง (contrast error : X'_{kj})

traits rater	A	B	C	D	E	Σ
α	+0.66	+0.10	-0.57	+0.33	-0.52	.00
β	-0.48	+0.10	+0.29	-0.10	+0.19	.00
γ	-0.19	-0.22	+0.29	-0.24	+0.34	-0.02
Σ	-0.01	-0.02	+0.01	-0.01	+0.01	-0.02

จากตารางที่ 7 เมื่อนำค่าในแต่ละเซลล์ลบด้วยค่าเฉลี่ยรวม (grand mean) จะได้ค่าประมาณความคลาดเคลื่อนแบบขัดแย้ง (contrast error = X'_{kj}) ตัวอย่างเช่น ผู้ประเมิน β คุณลักษณะ A มีค่าเท่ากับ 4.36 - 4.84 ซึ่งมีค่าเท่ากับ -0.44 แสดงว่าผลการประเมินของผู้ประเมิน β ที่ประเมินคุณลักษณะ A มีความคลาดเคลื่อนแบบขัดแย้งโดยประเมินต่ำกว่าความเป็นจริง 0.48

จากการประมาณค่าความคลาดเคลื่อนตามวิธีของ Guilford (1954) ผู้วิจัยนำมาประยุกต์ให้เข้ากับบริบทในการประเมินผลการปฏิบัติงานของครูในการวิจัยครั้งนี้โดยมีรายละเอียดดังนี้

ผู้ประเมิน (rater) คือบุคลากร 4 ประเภทได้แก่ ครูในโรงเรียนอัสสัมชัญ ครูต่างโรงเรียน ผู้ปกครองนักเรียนในโรงเรียนอัสสัมชัญ ผู้ปกครองนักเรียนต่างโรงเรียน แบ่งเป็น 2 กรณีตามจำนวนผู้ประเมินต่างกันเป็น แบบมีผู้ประเมิน 4 คน และ 8 คน

ผู้ถูกประเมิน (ratee) ได้แก่ ครูในโรงเรียนอัสสัมชัญ จำนวน 6 คนที่สุ่มเลือกโดยวิธีการสุ่มแบบแบ่งชั้นแยกตามวุฒิทางการศึกษาและอายุงานเป็นเกณฑ์ในการแบ่งชั้น

คุณลักษณะที่ประเมินผลการปฏิบัติงานของครูในห้องเรียน (trait) ประกอบด้วย 5 ด้าน คือ (1) การสื่อสารการใช้ภาษาพูดและภาษาเขียนของครู (2) การสร้างบรรยากาศการเรียนรู้ในห้องเรียน (3) การกำหนดกิจกรรมและการมอบหมายงาน (4) คุณภาพของข้อมูลป้อนกลับให้แก่ นักเรียน (5) การแสดงความยืดหยุ่นในการปรับบทเรียน

ในงานวิจัยนี้ ผู้วิจัยประยุกต์วิธีประมาณค่าความคลาดเคลื่อนโดยการวิเคราะห์ความแปรปรวน ตามแนวคิดของ Guilford โดยใช้ทั้งวิธีการวิเคราะห์ความแปรปรวน ทั้งแบบสองทางและแบบสามทางดังมีรายละเอียดดังนี้

การวิเคราะห์ความแปรปรวน (ANalysis Of VAriance: ANOVA) เป็นสถิติวิเคราะห์สำหรับข้อมูลที่ตัวแปรตามมีระดับการวัดเป็นระดับอันดับหรือระดับอัตราส่วนและมีตัวแปรอิสระเป็นแบบตัวแปรจัดประเภทที่ใช้ดังนี้

1. ใช้ในการทดสอบความแตกต่างระหว่างค่าเฉลี่ยจากกลุ่มประชากรมากกว่า 2 กลุ่ม
2. ใช้กับการวิจัยที่ออกแบบซับซ้อนมีตัวแปรอิสระมากกว่า 1 ตัวได้ในกรณีที่มีตัวแปรอิสระตั้งแต่สองตัวขึ้นไปสามารถทดสอบอิทธิพลปฏิสัมพันธ์ (interaction) ระหว่างตัวแปรอิสระได้ด้วย
3. ใช้ระบุขนาดความสัมพันธ์เชิงเส้นโค้งระหว่างตัวแปรอิสระและตัวแปรตามการวิเคราะห์ความแปรปรวนมีข้อตกลงเบื้องต้น (ศิริชัย กาญจนวาสี, ทวีวัฒน์ ปิตยานนท์ และ ดิเรก ศรีสุโข, 2540) ดังนี้

ข้อตกลงเบื้องต้น (basic assumptions)

1. กลุ่มตัวอย่างแต่ละกลุ่มเลือกมาอย่างสุ่มจากประชากรที่มีการแจกแจงเป็นโค้งแบบปกติ (normality)
2. กลุ่มตัวอย่างแต่ละกลุ่มเป็นอิสระต่อกัน (independent)
3. ค่าความแปรปรวนของประชากร (σ^2) ของกลุ่มตัวอย่างแต่ละกลุ่ม (หรือแต่ละ cell) มีค่าเท่ากัน (homogeneity of variance)

การวิเคราะห์ความแปรปรวนทางสามารจจำแนกได้ตามจำนวนตัวแปรอิสระ (ตัวแปรต้น) ดังนี้

- 1) การวิเคราะห์ความแปรปรวนทางเดียว (one-way ANOVA) คือการวิเคราะห์ความแปรปรวนในกรณีที่มีตัวแปรอิสระ (ตัวแปรต้น) เพียงตัวแปรเดียวและมีตัวแปรตามเพียงตัวแปรเดียว โดยตัวแปรอิสระอาจจะเป็นตัวแปรเชิงคุณลักษณะ (qualitative) หรือตัวแปรเชิงปริมาณ (quantitative) (Linderman R.H.and others 1980, ศิริชัย กาญจนวาสี, ทวีวัฒน์ ปิตยานนท์ และ ดิเรก ศรีสุโข 2540 ,บุญชม ศรีสะอาด 2541)

แนวความคิดพื้นฐานการวิเคราะห์ความแปรปรวนคือการวิเคราะห์การแปรผัน (variation) คือผลรวมของกำลังสองของความเบี่ยงเบนหรือที่เรียกว่า ผลรวมของกำลังสอง (sum of squares) แยกออกเป็นสองส่วนแต่ละส่วนใช้สัญลักษณ์และมีสมการแสดงความสัมพันธ์ระหว่างการแปรผันดังต่อไปนี้

- SS_T = การแปรผันทั้งหมด (total sum of squares)
 SS_b = การแปรผันระหว่างกลุ่ม (between groups sum of squares) ซึ่งเป็นการแปรผันที่เป็นระบบ
 SS_w = การแปรผันภายในกลุ่ม (within groups sum of squares) ซึ่งเป็นการแปรผันที่เป็นความคลาดเคลื่อน

$$SS_T = SS_b + SS_w \quad \text{-----} \quad (9)$$

สมมติฐานหลัก (H_0) ในการทดสอบมีสมมติฐานเดียวคือค่าเฉลี่ยของประชากรกลุ่มต่าง ๆ ไม่แตกต่างกัน

$$H_0 = \mu_1 = \mu_2 = \dots = \mu_k$$

เมื่อ $\mu_1, \mu_2, \dots, \mu_k$ แทน ค่าเฉลี่ยของประชากรกลุ่มที่ 1 กลุ่มที่ 2 ไปจนถึงกลุ่มที่ K

ตามลำดับ

k แทน จำนวนกลุ่ม

ตารางที่ 9 การวิเคราะห์ความแปรปรวนทางเดียว

แหล่งของความแปรผัน	ผลรวมของกำลังสอง (SS)	df	ค่าประมาณของความแปรปรวน (MS)	F	P
ระหว่างกลุ่ม (between)	SS_b	k-1	MS_b	$F = \frac{MS_b}{MS_w}$	
ภายในกลุ่ม (within)	SS_w	N-k	MS_w		
ทั้งหมด	SS_T	N-1			

2) การวิเคราะห์ความแปรปรวนสองทาง (two-way ANOVA)

การวิเคราะห์ความแปรปรวนสองทางคือการวิเคราะห์ความแปรปรวนในกรณีที่มีตัวแปรอิสระ (ตัวแปรต้น) สองตัวและมีตัวแปรตามเพียงตัวแปรเดียว ข้อดีของการวิเคราะห์แบบนี้คือนอกจากจะสามารถทดสอบอิทธิพลหลักของตัวแปรอิสระสองตัวไปพร้อม ๆ กันแล้วยังสามารถทดสอบอิทธิพลร่วมหรืออิทธิพลปฏิสัมพันธ์ (interaction) ระหว่างตัวแปรทั้งสองได้อีกด้วย (Linderman R.H. and others 1980, ศิริชัย กาญจนวาสี, ทวีวัฒน์ ปิตยานนท์ และดิเรก ศรีสุข, 2540, บุญชม ศรีสะอาด, 2541)

หลักการวิเคราะห์ความแปรปรวนสองทาง คือการวิเคราะห์การแปรผัน (variation) หรือผลรวมของกำลังสองของความเบี่ยงเบนทั้งหมด (SS_T) แยกออกเป็นการแปรผันเนื่องมาจากความแตกต่างแต่ละตัวแปร (SS_A และ SS_B) และการแปรผันของปฏิสัมพันธ์ระหว่างตัวแปรทั้งสองร่วมกัน (SS_{AB}) และการแปรผันภายในกลุ่ม (SS_W) เขียนได้ดังสมการที่ (10)

$$SS_T = SS_A + SS_B + SS_{AB} + SS_W \quad \text{----- (10)}$$

สมมติฐานการทดสอบในการวิเคราะห์ความแปรปรวนสองทางมีสมมติฐานในการทดสอบรวมทั้งหมด 3 สมมติฐานคือ

1. ค่าเฉลี่ยของประชากรกลุ่มต่าง ๆ ที่จำแนกกลุ่มตามตัวแปรแรก (A) ไม่แตกต่างกัน

$$H_0 = \mu_{A_1} = \mu_{A_2} = \dots = \mu_{A_k}$$

2. ค่าเฉลี่ยของประชากรต่าง ๆ ที่จำแนกกลุ่มตามตัวแปรที่สอง (B) ไม่แตกต่างกัน

$$H_0 = \mu_{B_1} = \mu_{B_2} = \dots = \mu_{B_k}$$

3. ค่าเฉลี่ยของประชากรในแต่ละเซลล์ (AB) ไม่แตกต่างกัน

$$H_0 = \mu_{A_1B_1} = \mu_{A_2B_2} = \dots = \mu_{A_kB_k}$$

ตารางสรุปผลการวิเคราะห์ความแปรปรวนแสดงได้ดังตารางที่ 10

ตารางที่ 10 การวิเคราะห์ความแปรปรวนสองทาง

แหล่งของ ความแปรผัน	ผลรวมของ กำลังสอง (SS)	df	ค่าประมาณ ของความแปรปรวน (MS)	F	P
A	SS_A	a-1	MS_A	$F_A = \frac{MS_A}{MS_W}$	
B	SS_B	b-1	MS_B	$F_B = \frac{MS_B}{MS_W}$	
ปฏิสัมพันธ์ (AB)	SS_{AB}	(a-1)(b-1)	MS_{AB}	$F_{AB} = \frac{MS_{AB}}{MS_W}$	
ภายในกลุ่ม (W)	SS_W	N-ab	MS_W		
ทั้งหมด	SS_T	N-1			

3) การวิเคราะห์ความแปรปรวนสามทาง (three-way ANOVA)

การวิเคราะห์ความแปรปรวนสามทางคือการวิเคราะห์ความแปรปรวนในกรณีที่มีตัวแปรอิสระ (ตัวแปรต้น) สามตัวและมีตัวแปรตามเพียงตัวแปรเดียว ข้อดีของการวิเคราะห์แบบนี้คือนอกจากจะสามารถทดสอบอิทธิพลหลักของตัวแปรอิสระสามตัวไปพร้อม ๆ กันแล้วยังสามารถทดสอบอิทธิพลร่วมหรือที่เรียกว่าอิทธิพลปฏิสัมพันธ์ (interaction) ระหว่างตัวแปรทั้งสามอีกด้วย

หลักการวิเคราะห์ความแปรปรวนสามทาง คือการวิเคราะห์การแปรผัน (variation) หรือผลรวมของกำลังสองของความเบี่ยงเบนทั้งหมด (SS_T) แยกออกเป็น การแปรผันเนื่องมาจากความแตกต่างของแต่ละตัวแปร (SS_A , SS_B และ SS_C) และการแปรผันของปฏิสัมพันธ์ระหว่างตัวแปรทั้งสอง (SS_{AB} , SS_{AC} , SS_{BC} , SS_{ABC}) และการแปรผันภายในกลุ่ม (SS_W) เขียนได้ดังสมการที่ (11)

$$SS_T = SS_A + SS_B + SS_C + SS_{AB} + SS_{AC} + SS_{BC} + SS_{ABC} + SS_W \quad \text{----- (11)}$$

สมมติฐานการทดสอบในการวิเคราะห์ความแปรปรวนสามทาง มีสมมติฐานในการทดสอบทั้งหมด 7 สมมติฐานคือ

1. ค่าเฉลี่ยของประชากรของกลุ่มต่าง ๆ ที่จำแนกกลุ่มตามตัวแปรแรก (A) ไม่แตกต่างกัน

$$H_0 = \mu_{A_1} = \mu_{A_2} = \dots = \mu_{A_k}$$

2. ค่าเฉลี่ยของประชากรของกลุ่มต่าง ๆ ที่จำแนกกลุ่มตามตัวแปรที่สอง (B) ไม่แตกต่างกัน

$$H_0 = \mu_{B_1} = \mu_{B_2} = \dots = \mu_{B_k}$$

3. ค่าเฉลี่ยของประชากรของกลุ่มต่าง ๆ ที่จำแนกกลุ่มตามตัวแปรที่สาม (C) ไม่แตกต่างกัน

$$H_0 = \mu_{C_1} = \mu_{C_2} = \dots = \mu_{C_k}$$

4. ค่าเฉลี่ยของประชากรในแต่ละเซลล์ AB ไม่แตกต่างกัน

$$H_0 = \mu_{A_1B_1} = \mu_{A_2B_2} = \dots = \mu_{A_kB_k}$$

5. ค่าเฉลี่ยของประชากรในแต่ละเซลล์ AC ไม่แตกต่างกัน

$$H_0 = \mu_{A_1C_1} = \mu_{A_2C_2} = \dots = \mu_{A_kC_k}$$

6. ค่าเฉลี่ยของประชากรในแต่ละเซลล์ BC ไม่แตกต่างกัน

$$H_0 = \mu_{B_1C_1} = \mu_{B_2C_2} = \dots = \mu_{B_kC_k}$$

7. ค่าเฉลี่ยของประชากรในแต่ละเซลล์ ABC ไม่แตกต่างกัน

$$H_0 = \mu_{A_1B_1C_1} = \mu_{A_2B_2C_2} = \dots = \mu_{A_kB_kC_k}$$

นำค่าที่คำนวณได้เขียนลงในตารางสรุปเพื่อรายงานผลการวิเคราะห์ความแปรปรวน
ดังตารางที่ 11

ตารางที่ 11 การวิเคราะห์ความแปรปรวนสามทาง

แหล่งของ ความแปรปรวน	ผลรวมของ กำลังสอง (SS)	df	ค่าประมาณของ ความแปรปรวน (MS)	F	P
A	SS_A	a-1	MS_A	$F_A = \frac{MS_A}{MS_W}$	
B	SS_B	b-1	MS_B	$F_B = \frac{MS_B}{MS_W}$	
C	SS_C	c-1	MS_C	$F_C = \frac{MS_C}{MS_W}$	
ปฏิสัมพันธ์ (AB)	SS_{AB}	(a-1)(b-1)	MS_{AB}	$F_{AB} = \frac{MS_{AB}}{MS_W}$	
ปฏิสัมพันธ์ (AC)	SS_{AC}	(a-1)(c-1)	MS_{AC}	$F_{AC} = \frac{MS_{AC}}{MS_W}$	
ปฏิสัมพันธ์ (BC)	SS_{BC}	(b-1)(c-1)	MS_{BC}	$F_{BC} = \frac{MS_{BC}}{MS_W}$	
ปฏิสัมพันธ์ (ABC)	SS_{ABC}	(a-1)(b-1)(c-1)	MS_{ABC}	$F_{ABC} = \frac{MS_{ABC}}{MS_W}$	
ภายในกลุ่ม (W)	SS_W	N-abc	MS_W		
ทั้งหมด	SS_T	N-1			

ในการวิจัยครั้งนี้ผู้วิจัยใช้การวิเคราะห์ความแปรปรวนแบบสามทางโดยประยุกต์ให้เข้ากับบริบทในการประเมินผลการปฏิบัติงานของครูในครั้งนี้ดังนี้

- A แทน แหล่งของความแปรปรวนเนื่องมาจากผู้ประเมิน (rater)
- B แทน แหล่งของความแปรปรวนเนื่องมาจากผู้ถูกประเมิน (ratee)
- C แทน แหล่งของความแปรปรวนเนื่องมาจากคุณลักษณะที่ประเมินผลการปฏิบัติงานของครู (trait)

ตอนที่ 4 มโนทัศน์เกี่ยวกับการปรับแก้ความคลาดเคลื่อนในการประเมิน

จากแนวคิดของ Guilford (1954) นอกจากจะนำเสนอการประมาณค่าความคลาดเคลื่อนแล้ว Guilford ยังได้นำเสนอการปรับแก้ความคลาดเคลื่อนในการประเมิน (Guilford, 1954:285-290) โดยเขียนในรูปของสมการ ดังสมการที่ (12) ซึ่งใช้สัญลักษณ์และโมเดล ดังต่อไปนี้

- X'_{ijk} = ค่าประมาณผลการประเมินคุณค่า (observed score) ของผู้ถูกประเมินคนที่ i ในการประเมินคุณลักษณะที่ j โดยผู้ประเมินคนที่ k
- X_{ijk} = ผลการประเมินคุณค่า (observed score) ของผู้ถูกประเมินคนที่ i ในการประเมินคุณลักษณะที่ j โดยผู้ประเมินคนที่ k
- X'_{ki} = ค่าประมาณความคลาดเคลื่อนแบบกต/ปล่อยคะแนน (leniency error = l) ของผู้ประเมินคนที่ k
- X'_{ki} = ค่าประมาณความคลาดเคลื่อนแบบฮาโล (halo error) ของผู้ถูกประเมินคนที่ i โดยผู้ประเมินคนที่ k
- X'_{kj} = ค่าประมาณความคลาดเคลื่อนแบบขัดแย้ง (contrast error) ของคุณลักษณะที่ j โดยผู้ประเมินคนที่ k

$$X'_{ijk} = X_{ijk} - X'_{ki} - X'_{kj} \text{ ----- (12)}$$

จากตารางที่ 3, ตารางที่ 5 และตารางที่ 9 ซึ่งแสดงค่าประมาณความคลาดเคลื่อนแบบกต/ปล่อยคะแนน, แบบฮาโลและแบบขัดแย้งตามลำดับ ผู้วิจัยนำมาเสนอเป็นตารางเดียว ดังตารางที่ 12

จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 12 ผลการประมาณค่าความคลาดเคลื่อนแบบกด/ปล่อยคะแนน,แบบฮาโลและ

แบบขัดแย้ง

rater	X'_{ki} (leniency error)	X'_{ki} (halo error)							X'_{kj} (contrast error)				
		Rate	Rate	Rate	Rate	Rate	Rate	Rate	Trai	Trai	Trai	Trai	Trai
		e	e	e	e	e	e	e	t	t	t	t	t
		1	2	3	4	5	6	7	1	2	3	4	5
α	+0.05	-0.52	+0.48	-0.92	-0.45	+1.35	-0.25	+0.28	+0.66	+0.10	-0.57	+0.33	-0.52
β	-0.38	+0.51	-0.09	+0.91	+1.58	-2.62	+0.78	-1.09	-0.48	+0.10	+0.29	-0.10	+0.19
γ	+0.33	.00	-0.40	.00	-1.13	+1.27	-0.53	+0.80	-0.19	-0.22	+0.29	-0.24	+0.34

เมื่อปรับแก้ความคลาดเคลื่อนทั้ง 3 แบบตามสมการที่ (12) โดยนำความคลาดเคลื่อนไปหักออกจากผลการประเมินเดิม (ตารางที่ 2) ดังตัวอย่างเช่น ข้อมูลของผู้ถูกประเมินที่ 1 ในคุณลักษณะ A โดยผู้ประเมิน α ($X'_{1A\alpha}$) มีค่า $5 - (.05) - (-.52) - (.66)$ ซึ่งมีค่าเท่ากับ 4.81 ดังตารางที่ 13 หรือข้อมูลของผู้ถูกประเมินคนที่ 7 ในคุณลักษณะ A โดยผู้ประเมิน β ($X'_{7A\beta}$) มีค่า $3 - (-.38) - (-1.09) - (-.48)$ ซึ่งมีค่าเท่ากับ 4.95 ดังตารางที่ 13

เพื่อความสะดวกในการพิจารณาผู้วิจัยจึงนำข้อมูลก่อนการปรับแก้และข้อมูลที่ปรับแก้แล้วมานำเสนอเป็นตารางเดียวดังตารางที่ 13

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 13 คะแนนผลการประเมินก่อนและหลังการปรับแก้ความคลาดเคลื่อนทั้ง 3 แบบ

ratee	ก่อนปรับแก้															หลังปรับแก้														
	Trait A			Trait B			Trait C			Trait D			Trait E			Trait A			Trait B			Trait C			Trait D			Trait E		
	α	β	γ	α	β	γ	α	β	γ	α	β	γ	α	β	γ	α	β	γ	α	β	γ	α	β	γ	α	β	γ	α	β	γ
1	5	6	5	5	5	5	3	4	5	5	6	7	3	3	3	4.8	6.4	4.9	5.4	4.8	4.9	4.0	3.6	4.4	5.1	6.0	6.9	4.0	2.7	2.3
2	9	8	7	7	7	7	5	5	5	8	7	7	5	2	3	7.8	9.0	7.3	6.4	7.4	7.3	5.0	5.0	4.8	7.1	7.6	7.3	5.0	2.3	4.7
3	3	4	3	3	5	5	3	3	5	7	6	5	1	6	5	3.2	4.0	2.9	3.8	4.4	4.9	4.4	2.2	4.4	7.5	5.6	4.9	2.4	5.3	4.3
4	7	5	5	3	6	3	1	4	3	3	5	3	3	5	1	6.7	4.3	6.0	3.3	4.7	4.0	2.0	2.5	3.5	3.1	3.9	4.0	3.9	3.6	1.5
5	9	2	9	7	4	7	7	3	7	8	2	7	5	3	1	6.9	5.5	7.6	5.5	6.9	5.6	6.2	5.7	5.1	6.3	5.1	5.6	4.1	5.8	5.1
6	3	4	3	5	4	3	3	6	3	5	4	5	1	2	3	2.5	4.1	3.4	5.1	3.5	3.4	3.8	5.3	2.9	4.9	3.7	5.4	1.7	1.4	2.9
7	7	3	7	7	3	7	5	5	7	5	5	5	5	4	7	6.0	5.0	6.1	6.6	4.4	6.1	5.2	6.2	5.6	4.3	6.6	4.1	5.2	5.3	5.5

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตอนที่ 5 การใช้ทฤษฎีการสรุปอ้างอิง (Generalizability Theory) ในการตรวจสอบและประมาณค่าความคลาดเคลื่อน

เนื่องจากการประมาณค่าความคลาดเคลื่อนในการประเมินตามวิธีของ Guilford (1954) เป็นวิธีการที่ใช้หลักการเรื่องการวิเคราะห์ความแปรปรวน (ANOVA) และหลักการทฤษฎีการสรุปอ้างอิงเป็นการวิเคราะห์โดยใช้หลักการวิเคราะห์ความแปรปรวนผู้วิจัยจึงได้นำหลักการทฤษฎีการสรุปอ้างอิง (Generalizability Theory) มาใช้ในการประมาณค่าความคลาดเคลื่อนด้วยการวิเคราะห์โดยใช้ทฤษฎีการสรุปอ้างอิงสามารถประมาณค่าความคลาดเคลื่อนจากหลายแหล่งที่เป็นไปได้ทั้งหมด ทั้งจากผลหลัก (main effect) และผลของปฏิสัมพันธ์ (interaction effect) ซึ่งจะทำให้ได้สารสนเทศที่มีความถูกต้อง แม่นยำมากยิ่งขึ้น นอกจากนี้ทฤษฎีนี้ยังสามารถพิจารณาความแปรปรวนของความคลาดเคลื่อนสัมพัทธ์ และความแปรปรวนของความคลาดเคลื่อนสัมบูรณ์ได้อีกด้วย ในตอนนี้ผู้วิจัยนำเสนอสาระเกี่ยวกับทฤษฎีการสรุปอ้างอิง ดังมีรายละเอียดต่อไปนี้

ทฤษฎีการสรุปอ้างอิง (Generalizability Theory)

คำศัพท์ที่เกี่ยวข้องกับทฤษฎีทางการสรุปอ้างอิง

ทฤษฎีการสรุปอ้างอิงมีคำศัพท์เฉพาะที่สำคัญดังนี้ (ศิริชัย กาญจนวาสี, 2541)

ประชากร (population) ประชากรตามทฤษฎีนี้หมายถึงสิ่งที่ต้องการวัด (object of measurement) หรือสิ่งที่มุ่งวัดทั้งหมด สิ่งที่มีวัดมักได้แก่คุณลักษณะของบุคคล หรือผู้ถูกประเมิน

facet ที่ต้องการศึกษา facet ที่ต้องการศึกษาเป็นชุดของเงื่อนไขของการวัดและประเมินที่มีลักษณะคล้ายกัน หรือกลุ่มเงื่อนไขของการวัดและประเมิน ซึ่งเป็นองค์ประกอบที่คาดว่าจะมีผลต่อการวัดและประเมินความคลาดเคลื่อน เช่น จำนวนครั้งของการประเมิน จำนวนผู้ประเมิน โดย facet ที่ต้องการศึกษาอาจเป็นองค์ประกอบสุ่มหรือองค์ประกอบที่เจาะจง ถ้าเงื่อนไขการวัดและประเมินถูกเลือกมาอย่างเจาะจงจากองค์ประกอบที่ศึกษาแสดงว่าผู้วิจัยสามารถทำการสรุปอ้างอิงไปยังองค์ประกอบเฉพาะในระดับของเงื่อนไขที่เลือกมาศึกษาเท่านั้น แต่ถ้าเงื่อนไขการวัดและประเมินได้รับการสุ่มเพื่อเป็นตัวแทนองค์ประกอบที่ศึกษา แสดงว่าผู้วิจัยสามารถทำการสรุปอ้างอิงไปยังระดับต่าง ๆ ขององค์ประกอบที่ศึกษาได้ facet ของการวัดและประเมิน เช่น ผู้ประเมิน (rater facet)

เงื่อนไขของการวัด (condition of measurement) เงื่อนไขของการวัดและประเมิน เป็นส่วนต่าง ๆ ของ facet หรือระดับของ facet ของการวัดและประเมิน เช่น facet ของจำนวนผู้ประเมิน อาจกำหนดจำนวนระดับเป็น 1, 2 และ 3 คน facet ผู้ประเมินแต่ละคนเป็นเงื่อนไขหนึ่งของ facet ผู้ประเมิน

เอกภพ (universe) เอกภพหมายถึงเงื่อนไขของการวัดและประเมินทั้งหมดของแต่ละ facet หรือกล่าวได้ว่าเป็นเงื่อนไขของการวัดและประเมิน ที่สนใจทั้งหมด เช่นจำนวนผู้ประเมินทั้งหมด

เอกภพของการสังเกตที่ยอมรับได้หรือเอกภพของค่าที่ได้จากการสังเกตทั้งหมด (universe of admissible observation) เป็นกลุ่มเงื่อนไขของการวัดและประเมิน ที่เป็นไปได้ซึ่งสามารถสังเกตได้ในแต่ละ facet เช่น facet ของผู้ประเมิน เอกภพของการสรุปอ้างอิง (universe of generalization) เป็นเงื่อนไขการวัดและประเมินทั้งหมดที่เป็นเป้าหมายของการสรุปอ้างอิง กล่าวได้ว่า เป็นการวัดและประเมินที่ครอบคลุมเงื่อนไขที่สนใจทั้งหมด หรือเป็นเงื่อนไขในเอกภพของการสังเกตที่ยอมรับได้ทั้งหมด ซึ่งอาจประกอบด้วยเซตย่อยของเงื่อนไขในเอกภพของการสังเกตที่ยอมรับได้

การศึกษา G (G-study) และการศึกษา D (D-study)

ทฤษฎีการสรุปอ้างอิง ประกอบด้วยขั้นตอนของการศึกษาที่สำคัญ 2 ขั้นตอน คือ

(1) การศึกษาเพื่อการสรุปอ้างอิงหรือการศึกษา G (generalizability study or G-study) เป็นการสรุปอ้างอิงผลที่ได้จากการศึกษาตัวอย่างการวัดและประเมิน ตามเงื่อนไขที่สนใจบรรยายความแปรปรวนของความคลาดเคลื่อนจากแหล่งความคลาดเคลื่อนต่าง ๆ เพื่อสรุปอ้างอิงไปยังเอกภพ

(2) การศึกษาเพื่อการตัดสินใจหรือการศึกษา D (decision study or D-study) เป็นการใช้ข้อมูลจากการศึกษา G ที่สอดคล้องกับจุดประสงค์เฉพาะของการวัด ตัดสินใจเลือกใช้รูปแบบหรือวิธีใด ในสถานการณ์ต่าง ๆ ที่ศึกษา

จุดประสงค์ของการศึกษา G คือความต้องการประมาณค่าความแปรปรวนของคะแนนจริงและความแปรปรวนของคะแนนความคลาดเคลื่อนจากแหล่งความคลาดเคลื่อนต่าง ๆ ที่สนใจ และใช้เป็นข้อมูลสำหรับวางแผนเพื่อตัดสินใจ ในการศึกษา D จะเกี่ยวกับความเที่ยงในสถานการณ์ของการวัดและประเมินต่าง ๆ ดังนั้น การออกแบบ G-study จึงควรครอบคลุมเงื่อนไขที่ต้องการตัดสินใจนำไปใช้ใน D-study

การวิเคราะห์ความแปรปรวนโดยใช้ทฤษฎีการสรุปอ้างอิง

ทฤษฎีการสรุปอ้างอิง สามารถประมาณค่าความแปรปรวนจากแหล่งต่าง ๆ ด้วยการใช้กระบวนการวิเคราะห์ความแปรปรวน (analysis of variance procedure) ดังที่ Shavelson (1991: 21,83-90) ได้อธิบายความหมายของความแปรปรวนและค่าสัมประสิทธิ์การสรุปอ้างอิง อันเป็นหลักการสำคัญไว้ดังนี้

ความแปรปรวนของคะแนน ความแปรปรวนที่สำคัญของทฤษฎีการสรุปอ้างอิง ได้แก่ ความแปรปรวน 2 ประเภท คือ (1) ความแปรปรวนของคะแนนความคลาดเคลื่อนสัมบูรณ์ (absolute error variance) และ (2) ความแปรปรวนของคะแนนความคลาดเคลื่อนสัมพัทธ์ (relative error variance)

สำหรับทฤษฎีการสรุปอ้างอิงนั้น คะแนนเอกภพ (universe score : μ_p) คือค่าเฉลี่ยของคะแนนการวัดหรือประเมิน ซ้ำหลาย ๆ ครั้งตามเงื่อนไขการวัดในการสรุปอ้างอิง สำหรับความคลาดเคลื่อนของการวัดและประเมิน (E_{pi}) แบ่งออกเป็นความคลาดเคลื่อนจาก facet หรือกลุ่มเงื่อนไขของการวัดและประเมิน (E_i) และความคลาดเคลื่อนจากแหล่งที่เหลืออื่น ๆ (e_{pi}) การวัดแต่ละครั้งไม่จำเป็นต้องใช้แบบสอบคู่ขนานเหมือนทฤษฎีการวัดแบบดั้งเดิม (classical test theory) ส่วนความแปรปรวนของค่าคาดหวังของคะแนนที่สังเกตได้เป็นผลรวมของความแปรปรวนของคะแนนเอกภพ ($\sigma_{\mu_p}^2$ หรือเขียนย่อ ๆ ว่า σ_p^2) กับความแปรปรวนของคะแนนความคลาดเคลื่อนจาก facet หรือองค์ประกอบ (i) ต่าง ๆ ของการวัด ($\sigma_{E_i}^2$) และความแปรปรวนของความคลาดเคลื่อนจากแหล่งอื่น ๆ ($\sigma_{e_p}^2$) โดยมีความสัมพันธ์เป็นสมการเชิงเส้นตรงดังนี้

$$X_{pi} = T_{pi} + E_{pi} + e_{pi} \quad \text{-----} \quad (13)$$

$$\sigma_{x_p}^2 = \sigma_{\mu_p}^2 + \sigma_{E_i}^2 + \sigma_{e_p}^2 \quad \text{-----} \quad (14)$$

ความแปรปรวนของคะแนนความคลาดเคลื่อนจากองค์ประกอบต่าง ๆ ของการวัดและประเมินแบ่งออกเป็น 2 ประเภท ได้แก่

1) ความแปรปรวนของคะแนนความคลาดเคลื่อนแบบสัมบูรณ์ (absolute error variance: σ_{ABS}^2 หรือ σ_{Δ}^2) หรือ ความแปรปรวนของ $(\mu_p - X_p)$ ซึ่งคำนวณได้จากผลรวมของความแปรปรวนของคะแนนจากแหล่งต่าง ๆ ยกเว้น $\sigma_{\mu_p}^2$ หรือ σ_p^2

2) ความแปรปรวนของคะแนนความคลาดเคลื่อนแบบสัมพัทธ์ (relative error variance: σ_{REL}^2 หรือ σ_s^2) หรือ ความแปรปรวนของ $(\mu_p - X_p)$ ซึ่งคำนวณได้จากผลรวมของความแปรปรวนของคะแนนจากแหล่งต่าง ๆ ที่มีปฏิสัมพันธ์ร่วมกับผู้ถูกประเมิน (p)

ค่าสัมประสิทธิ์การสรุปอ้างอิง (generalizability coefficient : E_p^2) ค่าสัมประสิทธิ์การสรุปอ้างอิง เป็นอัตราส่วนระหว่างความแปรปรวนของคะแนนเอกภพ และความแปรปรวนของค่าคาดหวังของคะแนนที่สังเกตได้ดังนี้

$$G - coefficient = \frac{\sigma_p^2}{\sigma_p^2 + error\ variance} \quad (15)$$

เนื่องจากความแปรปรวนของคะแนนความคลาดเคลื่อนจากองค์ประกอบต่าง ๆ ของการวัดมี 2 ประเภทจึงทำให้สัมประสิทธิ์การสรุปอ้างอิงมี 2 ประเภท ดังนี้

(1) สัมประสิทธิ์การสรุปอ้างอิงสำหรับการตัดสินใจสัมบูรณ์ (ρ^2_s)

เมื่อคะแนนความคลาดเคลื่อน เป็นความแปรปรวนของคะแนนความคลาดเคลื่อนสัมบูรณ์ สัมประสิทธิ์นี้บ่งบอกความเที่ยงของแบบสอบ ในสถานการณ์ของการตัดสินใจที่ขึ้นกับคะแนนของผู้สอบตามลำพัง ไม่มีการเปรียบเทียบภายในกลุ่มหรือระหว่างกลุ่ม เช่น แบบสอบอิงเกณฑ์ซึ่งใช้ตรวจสอบความสามารถของผู้สอบเทียบกับเกณฑ์ที่กำหนดไว้ล่วงหน้า

(2) สัมประสิทธิ์การสรุปอ้างอิงสำหรับการตัดสินใจสัมพัทธ์ (ρ^2_Δ)

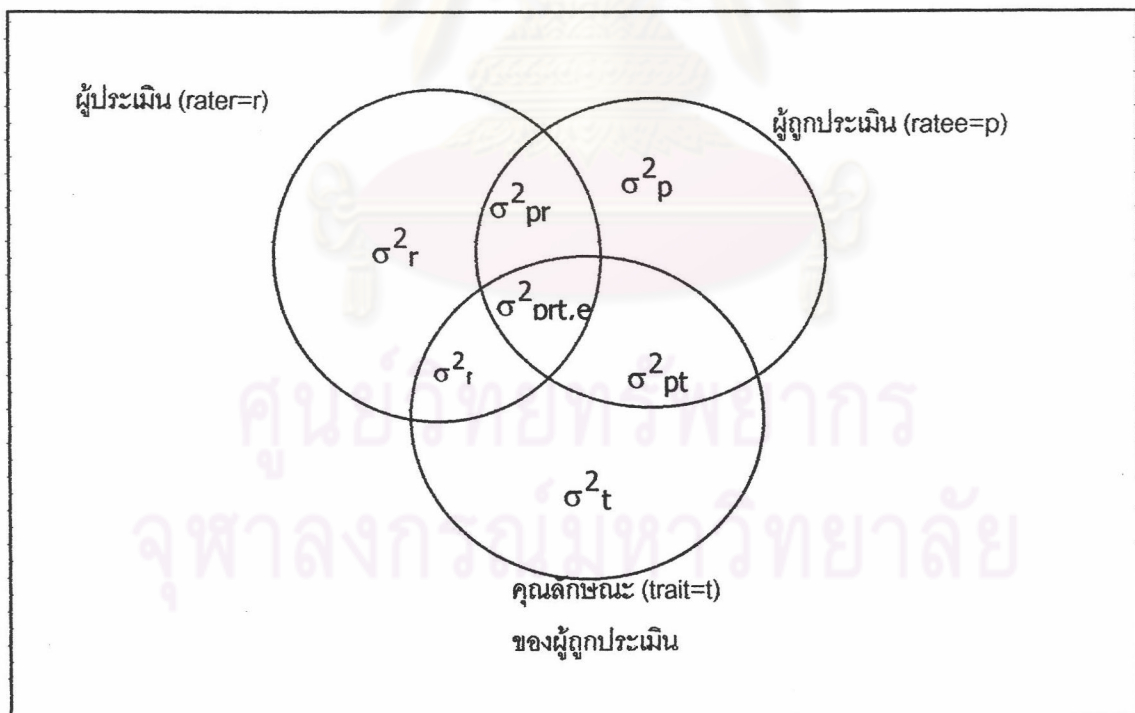
เมื่อคะแนนความคลาดเคลื่อน เป็นความแปรปรวนของคะแนนความคลาดเคลื่อนสัมพัทธ์ สัมประสิทธิ์นี้บ่งบอกความเที่ยงของแบบสอบในสถานการณ์ของการตัดสินใจที่มีการเปรียบเทียบคะแนนระหว่างผู้สอบ เช่น แบบสอบอิงกลุ่มซึ่งให้เปรียบเทียบคะแนนระหว่างผู้สอบด้วยการเข้าโค้งการแจกแจงคะแนนและตัดเกรดหรือให้ระดับผลการเรียนหรือการสอบแข่งขันเพื่อหาผู้ที่ได้คะแนนสูงเข้าศึกษา

ถึงแม้ว่าค่าสัมประสิทธิ์การสรุปอ้างอิง จะมีความหมายที่ใช้เช่นเดียวกับความเที่ยงของทฤษฎีการวัดแบบดั้งเดิม (classical test theory) แต่ก็มีประเด็นที่แตกต่างกันคือดังนี้

1. มีค่าสัมประสิทธิ์การสรุปอ้างอิงในการวิเคราะห์ได้มากกว่า 1 ค่า
2. การอ้างอิงไปยังเอกภพใดจะต้องระบุและอธิบายเอกภพนั้นให้ชัดเจนและต้องสุ่มเงื่อนไขนั้นมาศึกษาด้วย

3. คำสัมประสิทธิ์การสรูปอ้างอิงสามารถบอกถึงความเป็นเอกภาพได้ถ้าสิ่งที่ศึกษาเป็นตัวอย่างสุ่มจากเอกภาพ ข้อสอบที่มีความเป็นเอกพันธ์ จะสามารถใช้คะแนนสังเกตแทนคะแนนเอกภาพได้อย่างมั่นใจ

ในการวิจัยครั้งนี้ผู้วิจัยประยุกต์หลักทฤษฎีการสรูปอ้างอิงในการวิเคราะห์ข้อมูลโดยการสุ่มตัวอย่างมาศึกษา ข้อมูลสำหรับการวิเคราะห์ในที่นี้ประกอบด้วยกลุ่มตัวอย่างผู้ประเมิน (rater) n_r คน ประเมินคุณลักษณะ (trait) n_t ข้อของผู้ถูกประเมิน (ratee) n_p คน การศึกษาแบบ G-study จะมี $p \times t \times r$ design ในกรณีนี้จะประกอบด้วย แหล่งความแปรปรวน 7 แหล่งเป็น facet หลัก (main effect) 3 แหล่ง คือ (1) ผลจากผู้ถูกประเมิน (ratee= p) (2) ผลจากคุณลักษณะที่ประเมิน (trait= t) และ (3) ผลจากผู้ประเมิน (rater= r) และเป็นผลของปฏิสัมพันธ์อีก 4 แหล่งได้แก่ ปฏิสัมพันธ์ระหว่าง (1) ผู้ถูกประเมินกับคุณลักษณะที่ประเมิน ($p \times t$) (2) ผู้ถูกประเมินกับผู้ประเมิน ($p \times r$) (3) คุณลักษณะที่ประเมินกับผู้ประเมิน ($t \times r$) และ (4) ผู้ถูกประเมินกับคุณลักษณะที่ประเมินและผู้ประเมิน ($p \times t \times r$) ในกรณีนี้จะมีผลจากแหล่งต่าง ๆ ดังภาพที่ 2 และตารางที่ 14 (Shavelson 1991: 24,33)



ภาพที่ 2 ความแปรปรวนของผลการประเมินที่วัดได้ตามแนวคิดทฤษฎีการสรูปอ้างอิง เมื่ออธิบายด้วยความแปรปรวนของผู้ประเมิน ผู้ถูกประเมินและคุณลักษณะของผู้ประเมิน

ตารางที่ 14 การวิเคราะห์ความแปรปรวนรูปแบบ $p \times t \times r$ design ของชั้นการศึกษา G (G- Study)

Source of variance	df	MS	Expected mean square
ratee (p)	$n_p - 1$	$\frac{SS_p}{(n_p - 1)}$	$\sigma_{ptr,e}^2 + n_t \sigma_{pr}^2 + n_r \sigma_{pt}^2 + n_r n_t \sigma_p^2$
trait (t)	$n_t - 1$	$\frac{SS_t}{(n_t - 1)}$	$\sigma_{ptr,e}^2 + n_p \sigma_{tr}^2 + n_r \sigma_{pt}^2 + n_p n_r \sigma_t^2$
rater (r)	$n_r - 1$	$\frac{SS_r}{(n_r - 1)}$	$\sigma_{ptr,e}^2 + n_p \sigma_{tr}^2 + n_t \sigma_{pr}^2 + n_p n_t \sigma_r^2$
Interaction rateextrait (pt)	$(n_p - 1)(n_t - 1)$	$\frac{SS_{pt}}{(n_p - 1)(n_t - 1)}$	$\sigma_{ptr,e}^2 + n_r \sigma_{pt}^2$
Interaction rateextrater (pr)	$(n_p - 1)(n_r - 1)$	$\frac{SS_{pr}}{(n_p - 1)(n_r - 1)}$	$\sigma_{ptr,e}^2 + n_t \sigma_{pr}^2$
Interaction traitxrater (tr)	$(n_t - 1)(n_r - 1)$	$\frac{SS_{tr}}{(n_t - 1)(n_r - 1)}$	$\sigma_{ptr,e}^2 + n_p \sigma_{tr}^2$
Residual (ptr,e)	$(n_p - 1)(n_t - 1)(n_r - 1)$	$\frac{SS_{ptr}}{(n_p - 1)(n_t - 1)(n_r - 1)}$	$\sigma_{ptr,e}^2$

เมื่อนำหลักของการวิเคราะห์ความแปรปรวน (ANOVA) และหลักของทฤษฎีการสรุปอ้างอิง (Generalizability Theory) มาประยุกต์ใช้กับข้อมูลจะได้ผลว่าคะแนนผลการประเมินผลการปฏิบัติงานของครูที่สังเกตได้ (observe score) จะประกอบด้วยส่วนประกอบ 5 ส่วนคือ (1) ผลการประเมินผลการปฏิบัติงานของครูที่แท้จริง (true score) (2) อิทธิพลจากผู้ประเมิน (rater effect) (3) อิทธิพลจากผู้ถูกประเมิน (ratee effect) (4) อิทธิพลจากคุณลักษณะที่ประเมิน (trait

effect) (5) ความคลาดเคลื่อนในการประเมินผลการปฏิบัติงานของครู (residual) ซึ่งมีสัญลักษณ์และโมเดลดังสมการที่ (16)

X_{ijk} = ผลการประเมินผลการปฏิบัติงานของครูที่สังเกตได้ (observed score) ของครูคนที่ i ในการประเมินคุณลักษณะที่ j โดยผู้ประเมินคนที่ k

α_i = อิทธิพลจากผู้ประเมินคนที่ i (rater effect)

β_j = อิทธิพลจากผู้ถูกประเมินคนที่ j (ratee effect)

γ_k = อิทธิพลจากคุณลักษณะที่ k (trait effect)

$\alpha\beta$ = ปฏิสัมพันธ์ระหว่างผู้ประเมินกับผู้ถูกประเมิน (rater x ratee interaction)

$\alpha\gamma$ = ปฏิสัมพันธ์ระหว่างผู้ประเมินกับคุณลักษณะที่ประเมิน (rater x trait interaction)

$\beta\gamma$ = ปฏิสัมพันธ์ระหว่างผู้ถูกประเมินกับคุณลักษณะที่ประเมิน (ratee x trait interaction)

$\alpha\beta\gamma$ = ปฏิสัมพันธ์ระหว่างผู้ประเมินกับผู้ถูกประเมิน และคุณลักษณะที่ประเมิน (rater x ratee x trait interaction)

e = ความคลาดเคลื่อนในการประเมินผลการปฏิบัติงานของครู (residual)

$$X_{ijk} = X_{ijt} + \alpha_i + \beta_j + \gamma_k + \alpha\beta + \alpha\gamma + \beta\gamma + \alpha\beta\gamma + e \text{ -----(16)}$$

สมการข้างต้นนี้เหมือนกับสมการที่ (6) ที่ได้นำเสนอไว้ตอนต้น แต่เนื่องจากข้อมูลชุดนี้เมื่อวิเคราะห์ความแปรปรวนแบบสามทาง (three way ANOVA) จะไม่สามารถประมาณค่าความแปรปรวนปฏิสัมพันธ์ $\alpha\beta\gamma$ ได้ เพราะความแปรปรวนดังกล่าวจะไม่รวมอยู่ในทอมของความคลาดเคลื่อน แต่การวิเคราะห์ความแปรปรวนแบบสามทางจะให้ผลการวิเคราะห์ถูกต้องมากกว่าวิธีการวิเคราะห์ความแปรปรวนแบบสองทาง (two way ANOVA) ของ Guilford ดังนั้นการวิจัยในครั้งนี้ ผู้วิจัยจะพิจารณาประยุกต์วิธีการของ Guilford โดยใช้การวิเคราะห์ความแปรปรวนแบบสามทางในการวิเคราะห์ด้วยทฤษฎีการสรุปอ้างอิง

ตอนที่ 6 มโนทัศน์เกี่ยวกับการประเมินผลการปฏิบัติงานของครู

การเสนอเนื้อหาสาระในตอนนี้แบ่งเป็น 2 ตอนคือ (1) ความหมาย วัตถุประสงค์ ประโยชน์ และเทคนิคในการประเมินผลการปฏิบัติงานของครู และ (2) เกณฑ์การประเมินผลการปฏิบัติงาน มีรายละเอียดดังนี้

6.1 ความหมาย วัตถุประสงค์ ประโยชน์และเทคนิคในการประเมินผลการปฏิบัติงานของครู

การประเมินผลการปฏิบัติงานเป็นสิ่งจำเป็นในการบำรุงรักษาและการพัฒนาบุคลากรรวมทั้งมีผลกระทบต่อประสิทธิภาพของการทำงานของบุคลากร ในยุคปัจจุบันองค์การต่างๆ มีการพัฒนาและการแข่งขันกันสูงในทุกด้านการบริหารงานบุคคลอย่างมีประสิทธิภาพ จึงเป็นสิ่งจำเป็นอย่างยิ่งและเป็นที่ยอมรับกันในวงการบริหารงานสมัยใหม่ว่า การประเมินผลการปฏิบัติงานกำลังมีความสำคัญเพิ่มมากขึ้นอย่างต่อเนื่อง องค์การต่าง ๆ ทั้งภาครัฐบาลและเอกชนตระหนักดีว่าการประเมินผลการปฏิบัติงานของบุคลากรทุกชั้นตอนอย่างต่อเนื่องเป็นหน้าที่สำคัญประการหนึ่งในการบริหารงานบุคคล (Mondy และ Noe 1990)

6.1.1 ความหมายของการประเมินผลการปฏิบัติงาน

คำว่า "การประเมินผลการปฏิบัติงาน" ในภาษาอังกฤษใช้แตกต่างกันตามความประสงค์ในการใช้งานของนักวิชาการ เช่น Performance Appraisal, Staff Appraisal, Performance Evaluation และ Performance Assessment คำเหล่านี้มีความหมายคล้ายคลึงกันแต่คำที่ใช้กันมากคือ "Performance Appraisal" และ "Performance Evaluation" ซึ่งหมายถึง "การประเมินผลการปฏิบัติงาน" (Mondy และ Noe 1990)

นักวิชาการหลายท่านให้ความหมายของการประเมินผลการปฏิบัติงานไว้ต่างกัน ดังต่อไปนี้

Werther และ Devis (1989) ได้กล่าวโดยสรุปได้ว่าการปฏิบัติงานเป็นกระบวนการทำงานภายในองค์กรให้ประสบผลสำเร็จตามเป้าหมายขององค์กร

Schermehorn และ Osborn (1991) ให้ความหมายของผลการปฏิบัติงานไว้ว่าผลการปฏิบัติงาน หมายถึง ปริมาณและคุณภาพของการทำงานให้สำเร็จของบุคคล กลุ่ม หรือองค์กร

Steers (1991) เสนอว่า การปฏิบัติงาน หมายถึง ขอบเขตที่บุคคลทำงานสำเร็จบรรลุเป้าหมายงานที่ได้กำหนดไว้ซึ่งรวมถึงผลผลิตของงานและปัจจัยด้านอื่น ๆ ผลการปฏิบัติงานของ

บุคลากรในองค์กร ได้รับอิทธิพลมาจาก ความสามารถของบุคคล ความสนใจ แรงจูงใจในการทำงาน บรรยากาศที่ทำงาน รูปแบบการบริหารงาน

Cumming และ Schwab (1973) ได้เสนอว่าปัจจัยที่มีผลต่อผลการปฏิบัติงานของบุคลากรในองค์กร คือปัจจัยสภาพแวดล้อมเช่น การบริหารงาน ความสัมพันธ์กับเพื่อนร่วมงาน สภาพการทำงาน ตลอดจนปัจจัยด้านแรงจูงใจที่ทำให้เกิดความพึงพอใจในการทำงาน และปัจจัยบุคคลที่มีผลต่อการปฏิบัติงานของบุคลากรในองค์กร

Beach (1975) นิยามว่า “การประเมินผลการปฏิบัติงานคือ การสร้างระบบการประเมินค่าบุคลากรในด้านการปฏิบัติงาน”

Mondy และ Noe (1990) ได้ให้ความหมายว่าไว้คล้ายคลึงกันว่า “การประเมินผลการปฏิบัติงาน หมายถึง ระบบที่จัดขึ้นอย่างเป็นทางการเพื่อประเมินค่าการปฏิบัติงานของบุคคล”

จะเห็นได้ว่า ผลการปฏิบัติงาน หมายถึง ปริมาณ คุณภาพ ประสิทธิภาพหรือความพยายามในการทำงานและผลผลิตของงานและปัจจัยด้านอื่น ๆ ของบุคคลที่ทำงานให้สำเร็จบรรลุเป้าหมายของบุคคล กลุ่ม องค์กรหรือหน่วยงานการศึกษาเกี่ยวกับการปฏิบัติงาน สามารถประเมินผลการปฏิบัติงาน ซึ่งเป็นสิ่งสำคัญของการตัดสินใจประสิทธิภาพขององค์กร และยังเป็น การประเมินประสิทธิภาพการทำงานของบุคคลในองค์กรอีกด้วย การประเมินผลการปฏิบัติงาน เป็นกระบวนการอย่างเป็นทางการในการประเมินผลการปฏิบัติงาน เพื่อได้ทราบประสิทธิภาพการทำงาน และนำไปปรับปรุงการปฏิบัติงานให้ดียิ่งขึ้น และเป็นพื้นฐานในการพัฒนาองค์กรให้ก้าวหน้าต่อไป (Szilagyi & Wallace, 1990)

การประเมินผลการปฏิบัติงานของบุคคลสามารถกระทำได้สองแนวทางคือ การประเมินกิจกรรมของบุคคลหรือพฤติกรรมของบุคคล (activity or behavioral measures) และการประเมินผลผลิตหรือผลลัพธ์ของกิจกรรม (output measures) ทั้งนี้ในการประเมินจะเน้นแนวทางใดแนวทางหนึ่ง หรือทั้งสองแนวทางก็ได้แต่ ระบบการประเมินผลการปฏิบัติงานที่จะสามารถบรรลุวัตถุประสงค์ของทั้งบุคคลและองค์กร จะต้องมีการประเมินทั้งด้านกิจกรรมหรือพฤติกรรมที่ต้องปฏิบัติ (Porter & Hackman, 1987)

ในงานวิจัยนี้ผู้วิจัยได้นิยามผลการปฏิบัติงานของครูโดยสรุปจากนิยามที่นำเสนอว่าหมายถึง ผลการประเมินคุณภาพในการทำหน้าที่ในความรับผิดชอบของครูผู้สอนโดยประเมินจากกิจกรรมและพฤติกรรมของครูในห้องเรียน

6.1.2 วัตถุประสงค์ของการประเมินผลการปฏิบัติงาน

ในการประกอบกิจการใดก็ตามไม่ว่าจะเป็นองค์การของทางราชการหรือองค์การทางเอกชนหรือสถานศึกษาต่าง ๆ จะต้องกำหนดเป้าหมายหรือวัตถุประสงค์ในการปฏิบัติงานเสนอวัตถุประสงค์เหล่านี้อาจมีรายละเอียดแตกต่างกันไปตามลักษณะของหน่วยงาน แต่จะมีวัตถุประสงค์สำคัญเหมือนกันคือ มุ่งนำผลการประเมินไปใช้ประโยชน์ให้เกิดความสำเร็จในการปฏิบัติงาน อย่างมีประสิทธิภาพ การที่จะทราบว่าหน่วยงานหรือสถานศึกษานั้นมีประสิทธิภาพและประสิทธิภาพเพียงใดนั้นสามารถพิจารณาข้อมูลได้จากการประเมินผลการปฏิบัติงานนั่นเอง วัตถุประสงค์ของการประเมินผลการปฏิบัติงานมีนักวิชาการกล่าวไว้หลายท่าน คือ

Henderson (1984) ได้กล่าวถึงวัตถุประสงค์ของการนำผลการประเมินไปใช้ประโยชน์ได้ในประเด็นต่าง ๆ รวม 6 ประเด็นต่อไปนี้ คือ

1. การตัดสินใจที่เกี่ยวกับการให้ค่าตอบแทน
2. การกำหนดวัตถุประสงค์ของการฝึกอบรม
3. การพัฒนาผู้ปฏิบัติงาน
4. การตั้งเกณฑ์การคัดเลือกบุคลากร
5. การตัดสินใจที่เกี่ยวกับการเปลี่ยนแปลงสถานภาพของผู้ปฏิบัติงานเช่น การเลื่อนตำแหน่ง การไม่เลื่อนตำแหน่ง การโยกย้าย การลดตำแหน่ง และการให้พ้นจากงาน
6. การวางแผนกำลังคนขององค์การ

Crane (1986) กล่าวไว้ว่า การประเมินผลการปฏิบัติงานมีวัตถุประสงค์ดังนี้

1. เพื่อที่จะเป็นข้อมูลป้อนกลับให้ผู้ปฏิบัติงานได้ทราบถึงผลการการปฏิบัติงานของตน
2. เพื่อใช้เป็นหลักฐานในการเปลี่ยนแปลงแก้ไขพฤติกรรมอันจะทำให้ผลการปฏิบัติงานมีประสิทธิภาพมากขึ้น
3. เพื่อใช้เป็นข้อมูลในการมอบหมายงานการเลื่อนตำแหน่งและการให้ค่าตอบแทน
4. เพื่อจัดหรือแก้ไขปรับปรุงผลการปฏิบัติงานของผู้ที่มีผลการปฏิบัติงานต่ำหรือน้อยเกินไป
5. เพื่อเป็นข้อมูลในการจัดทำโครงการฝึกอบรมได้อย่างมีประสิทธิภาพ
6. เพื่อปรับปรุงการให้ค่าตอบแทนกับผู้ปฏิบัติงาน

7. เพื่อเป็นข้อมูลในการคาดคะเนเกี่ยวกับการคัดเลือกและโยกย้ายบุคคล
ประยูร บุญประเสริฐ (2523) กล่าวว่า วัตถุประสงค์ของการประเมินผลการปฏิบัติงาน
 ไว้ดังนี้

1. เพื่อความเหมาะสมในการพิจารณาเลื่อนเงินเดือนค่าจ้างหรือค่าตอบแทน
 อื่น ๆ
2. เพื่อให้ผู้บริหารมีข้อมูลที่จะแสดงต่อผู้ปฏิบัติงานเมื่อเกิดข้อขัดแย้งหรือความ
 ไม่พอใจขึ้นว่าใครเหมาะสมที่จะได้รับความดีความชอบระดับนั้น ๆ
3. เพื่อการเลื่อนตำแหน่ง เลื่อนขั้น สับเปลี่ยนงาน หรือให้ออกจากงาน
4. เพื่อเป็นเครื่องมือในการตัดสินใจว่าผู้ปฏิบัติงานแต่ละคนมีความแตกต่างกัน
 อย่างไร ใครเป็นบุคคลที่เหมาะสม

สมพงษ์ เกษมสิน (2526) กล่าวว่า วัตถุประสงค์ของการประเมินผลการปฏิบัติงานมี
 6 ประการ คือ

1. เพื่อการพิจารณาแต่งตั้งบุคคลเข้าปฏิบัติงาน
2. เพื่อการพิจารณาเลื่อนขั้น และเลื่อนเงินเดือน
3. เพื่อพิจารณาปรับปรุงสมรรถภาพของผู้ปฏิบัติงานให้ดีขึ้น
4. เพื่อพิจารณาเกี่ยวกับการจัดสวัสดิการผู้ปฏิบัติงาน
5. เพื่อพิจารณาความดีความชอบประจำปี
6. เพื่อพิจารณาให้ผู้ปฏิบัติงานที่หย่อนสมรรถภาพและไม่สามารถปรับปรุง
 แก้ไขตนเองให้ดีขึ้น ให้พ้นจากหน้าที่การงานไป

เสนาะ ดีเยาว์ (2532) กล่าวว่า การประเมินผลการปฏิบัติงานอาจกระทำด้วย
 วัตถุประสงค์ที่แตกต่างกัน ขึ้นอยู่กับความต้องการของผู้บังคับบัญชา แต่วัตถุประสงค์ที่สำคัญ
 ก็คือ

1. เพื่อพิจารณาความเหมาะสมในด้านเงินเดือน ค่าจ้าง ค่าตอบแทนที่จ่ายให้
 กับการทำงาน
2. เพื่อการพิจารณาเปลี่ยนแปลงสถานะของผู้ปฏิบัติงานด้วยการ
 เลื่อนตำแหน่ง เลื่อนขั้นหรือย้ายงาน
3. เพื่อวัดค่าของการปฏิบัติงานของพนักงานอันเป็นสิ่งชี้ว่าพนักงานคนนั้นควร
 อยู่ปฏิบัติงานต่อไปในองค์การหรือไม่

4. เพื่อหาจุดเด่นและจุดด้อยของพนักงานองค์การจะได้ส่งเสริมให้ได้ใช้ประโยชน์จากจุดเด่นต่อไป ในแง่ของจุดด้อยก็จะได้แก้ไขปรับปรุง

กล่าวโดยสรุปวัตถุประสงค์ของการประเมินผลการปฏิบัติงานมี 4 ประการคือ (1) เพื่อใช้เป็นข้อมูลในการคัดเลือกบุคลากร (2) เพื่อใช้เป็นข้อมูลในการวางแผนกำลังคนขององค์กร (3) เพื่อใช้เป็นข้อมูลในการตัดสินใจเกี่ยวกับการให้ค่าตอบแทน การพิจารณาเลื่อนขั้น เลื่อนตำแหน่ง การโยกย้าย การสับเปลี่ยนหน้าที่ การให้พ้นจากงาน (4) เพื่อใช้เป็นข้อมูลย้อนกลับต่อผู้ปฏิบัติงานและใช้เป็นข้อมูลสำหรับการวางแผนพัฒนาผู้ปฏิบัติงานในองค์การได้อย่างเหมาะสมและมีประสิทธิภาพ

6.1.3 ประโยชน์ของการประเมินผลการปฏิบัติงาน

วัตถุประสงค์ของระบบการประเมินผลการปฏิบัติงานทั้งหมดก็คือ การประเมินค่าและการให้ข้อมูลย้อนกลับต่อผู้ปฏิบัติงาน ซึ่งจะช่วยปรับปรุงบุคลากรและประสิทธิผลขององค์กรหรือสถานศึกษาประโยชน์ของการประเมินผลการปฏิบัติงานได้มีนักวิชาการกล่าวไว้หลายท่าน ได้แก่

ภิญโญ สาร (2516) กล่าวถึง ประโยชน์ที่เห็นได้ชัดเจนในการประเมินผลการปฏิบัติงานอย่างมีหลักเกณฑ์ ไว้ดังนี้

1. รักษาความเป็นธรรมของผู้บริหาร ไว้ในการพิจารณาความดีความชอบของบุคคล ทั้งในด้านการขึ้นเงินเดือน เลื่อนตำแหน่ง และการให้ลาศึกษาต่อ
2. ช่วยการตัดสินใจของผู้บริหารในการปูนบำเหน็จบุคคลได้ถูกต้อง
3. เป็นการบำรุงขวัญและกระตุ้นบุคคลให้ขยันขันแข็งในการปฏิบัติงาน
4. ช่วยให้มีการควบคุมค่าจ้างได้อย่างเหมาะสม กับการทำงานของแต่ละบุคคลและลักษณะงานแต่ละอย่างไม่เป็นเหตุให้จ้างบุคคลแพงเกินไปหรือถูกเกินไป
5. เมื่อมีการทักท้วงจากบุคคลว่าไม่ยุติธรรมผู้บริหารมีหลักฐานยืนยันว่า ได้ใช้หลักเกณฑ์ตามหลักวิชาแล้วและมีความบริสุทธิ์ใจในการพิจารณาความดีความชอบ
6. เป็นการจูงใจให้บุคคลที่ดีและมีความสามารถเข้ามาสมัครงาน

เยาเวดี ราชชัยกุล และคณะ (2522) กล่าวถึงประโยชน์ที่ได้จากการประเมินผลการปฏิบัติงานบุคคลว่า ถ้าการดำเนินการเป็นไปอย่างสุจริตรอบคอบและด้วยความซื่อสัตย์แล้วผลที่จะบังเกิดขึ้นคือ

1. การเลื่อนตำแหน่งพนักงานเจ้าหน้าที่ที่มีผลงานเป็นที่น่าพึงพอใจทำให้เกิดขวัญและกำลังใจว่าผู้บังคับบัญชาเอาใจใส่ ติดตามผลความก้าวหน้าของตนอย่างใกล้ชิด
2. เป็นแนวทางอย่างเป็นระบบแบบแผนที่ผู้บังคับบัญชาจะวางแผนฝึกอบรมให้แก่พนักงานเจ้าหน้าที่ของตนต่อไป
3. ให้ความมั่นใจต่อการปฏิบัติงานของพนักงานเจ้าหน้าที่มากกว่าเป็นการตัดสินใจการกระทำเฉพาะหน้า
4. ช่วยในการโอนย้ายบรรจุบุคลากรตามความสามารถและได้ประโยชน์มากที่สุด
5. ช่วยในการบันทึกความสามารถของพนักงานเจ้าหน้าที่ว่ามีความสามารถด้านใด
6. ช่วยในการให้พนักงานเจ้าหน้าที่ได้มีโอกาสปรึกษาผู้บังคับบัญชาอย่างใกล้ชิดในปัญหาทางาน การสนใจในงานและอนาคตเป็นต้น

สมพงษ์ เกษมสิน (2526) กล่าวว่าการประเมินผลการปฏิบัติงานมีประโยชน์ดังนี้

1. เพื่อเป็นการป้อนำเห็นแก่ผู้ปฏิบัติงานได้ดำเนินไปด้วยความยุติธรรม มีเหตุผล มีระบบเป็นแบบแผน ซึ่งใช้ต่อบุคคลในองค์การ เดียวกันโดยเสมอหน้าและถี่ถ้วนเป็นการป้องกันการเล่นพรรคเล่นพวก
2. เพื่อให้การเลื่อนขั้นเลื่อนตำแหน่งเป็นไปด้วยความยุติธรรมมีข้อเปรียบเทียบเป็นหลักฐานในการพิจารณาป้องกันการกินแหนงแคลงใจแก่ผู้ที่มีได้รับการเลื่อนขั้น เลื่อนตำแหน่ง ให้ได้ทราบข้อบกพร่องด้วยความเป็นเป็นธรรม
3. เพื่อใช้เป็นเครื่องมือในการพิจารณาว่าบุคคลใดมีสมรรถภาพในการปฏิบัติงาน หรือผู้ใดหย่อนสมรรถภาพ ไม่สมควรจะให้ปฏิบัติงานในองค์การต่อไปซึ่งเท่ากับเป็นการลดค่าใช้จ่ายอันไม่จำเป็นขององค์การออกไปทั้งก่อให้เกิดความเป็นธรรมในระหว่างผู้ปฏิบัติงานด้วยกัน
4. เพื่อประโยชน์ในการกำหนดวิธีการฝึกอบรมพนักงานเจ้าหน้าที่ให้ถูกต้องตามความจำเป็น และความต้องการในการฝึกอบรมของเจ้าหน้าที่แต่ละประเภท แต่ละหน้าที่และแต่ละสาขาวิชาชีพที่ปฏิบัติงานอยู่ในองค์การ

5. เพื่อประโยชน์ในการปรับเปลี่ยนโยกย้ายพนักงานเจ้าหน้าที่ได้ปฏิบัติหน้าที่ให้เหมาะสมกับความรู้ ความสามารถของตน เพื่อประสิทธิผลแก่งาน
6. เพื่อประกอบการพิจารณาปรับปรุงวิธีการสรรหาบุคคลที่องค์การใช้เลือกสรรบุคคลเข้ามาปฏิบัติงานในองค์การให้มีประสิทธิภาพยิ่งขึ้น กล่าวคือจะได้ทราบว่าวิธีการทดสอบที่ใช้อยู่สามารถวัดและเลือกสรรบุคคลได้ตามความประสงค์ขององค์การหรือไม่หากมีข้อบกพร่องจะได้ปรับปรุงแก้ไขได้ทันที
7. เพื่อเป็นเครื่องจูงใจให้พนักงานเจ้าหน้าที่ในองค์การได้ปฏิบัติหน้าที่ด้วยความเต็มใจเต็มกำลังความรู้ความสามารถไม่ต้องคอยกังวลที่จะประจบผู้บังคับบัญชาด้วยวิธีที่ไม่ถูกต้องแต่จะประจบผู้บังคับบัญชาด้วยการปฏิบัติงานโดยปราศจากความกังวลใจ

สรุปได้ว่าการประเมินผลการปฏิบัติงานมีประโยชน์ทั้งกับผู้บริหารในการวางแผนงานบุคคลและการวางแผนการอบรมคือช่วยการพัฒนาคุณภาพและประสิทธิภาพของบุคลากร สร้างขวัญและกำลังใจต่อผู้ปฏิบัติงาน เป็นข้อมูลในการพิจารณาค่าตอบแทนการพิจารณาความดีความชอบเป็นกรณีพิเศษ การเลื่อนตำแหน่ง การโอนย้าย การให้พ้นจากงาน ให้เป็นไปอย่างยุติธรรมป้องกันการเล่นพรรคเล่นพวก นอกจากนี้ยังเป็นประโยชน์ต่อผู้ปฏิบัติงานกล่าวคือ ทำให้ผู้ปฏิบัติงานรู้ข้อบกพร่องในการปฏิบัติงานของตนเองและสามารถที่จะแก้ไขพัฒนาต่อไปได้ และผู้ปฏิบัติงานได้รับความเสมอภาคและยุติธรรมในการพิจารณาค่าตอบแทน การพิจารณาความดีความชอบหรือ การโยกย้าย เลื่อนตำแหน่ง

6.1.4 เทคนิคการประเมินผลการปฏิบัติงาน

Henderson (1984) ได้แบ่งเทคนิคการประเมินผลการปฏิบัติงานแยกตามประเภทของวิธีได้ทั้งหมด 6 วิธีคือ (1) การวิเคราะห์วิธีการทำงาน (work methods analysis) (2) เทคนิคการทบทวนเชิงบรรยาย (narrative description review technique) (3) เทคนิคการจัดลำดับ (ranking technique) (4) เทคนิคการตรวจรายการ (check list technique) (5) เทคนิคมาตรฐานค่า (rating scale technique) (6) เทคนิคเชิงเป้าหมาย (goal setting technique)

เริงศักดิ์ ปานเจริญ (2535) ได้แบ่งเทคนิคการประเมินผลการปฏิบัติงานออกเป็น 5 ประเภทตามลักษณะเครื่องมือคือ (1) การประเมินผลตามสเกล (graphic rating scale) (2) วิธีการประเมินแบบจัดลำดับที่ (ranking) (3) วิธีการประเมินเชิงเปรียบเทียบ (paired-comparison) (4) วิธีการประเมินจากทางเลือกที่กำหนดไว้ให้ (forced choice) (5) วิธีการประเมินโดยการให้น้ำหนัก (weight checklist)

ในบรรดาเทคนิคทั้งหลายเหล่านี้เทคนิคมาตรฐานประมาณค่าเป็นวิธีที่ใช้แพร่หลายและง่ายมากที่สุด ผู้ประเมินสามารถเข้าใจได้ง่ายและสามารถสร้างให้ครอบคลุมเนื้อหาที่ต้องการประเมินได้ (Henderson, 1984, เริงศักดิ์ ปานเจริญ, 2535) ในงานวิจัยนี้ผู้วิจัยจึงเลือกใช้ เทคนิคมาตรฐานประมาณค่าในการประเมินผลการปฏิบัติงาน

6.2 เกณฑ์การประเมินผลการปฏิบัติงาน

การประเมินผลการปฏิบัติงานมีนักวิชาการเสนอเกณฑ์ในการประเมินผลการปฏิบัติงานที่หลากหลายดังนี้

Abdel- Halim (1980) ได้เสนอแนะเกณฑ์การประเมินผลการปฏิบัติงานโดยประเมินจากคุณภาพการทำงาน ปริมาณความพยายามในการทำงาน ผลผลิตของงาน ความเร็วในการทำงาน การปฏิบัติงานโดยรวม

Kesselman และคณะ (1974) กล่าวโดยสรุปได้ว่าเกณฑ์ที่ใช้ในการประเมินผลการปฏิบัติควรครอบคลุม 3 ด้านที่สำคัญ คือ คุณภาพการปฏิบัติงาน ผลผลิตของงาน และปริมาณความพยายามในการทำงาน

สำหรับในสถานศึกษาการประเมินผลการปฏิบัติงานมุ่งเน้นการประเมินผลการปฏิบัติงานของครูโดยมีเกณฑ์ที่ใช้ในการประเมินผลการปฏิบัติงานของครู เรียกว่า เกณฑ์มาตรฐานวิชาชีพครู สำนักงานคุรุสภา (2537) ให้นิยามว่า เกณฑ์มาตรฐานวิชาชีพครู เป็นเกณฑ์ของการปฏิบัติตนและปฏิบัติหน้าที่ของครู เพื่อประกันคุณภาพการทำงานและการเป็นครู เกณฑ์มาตรฐานวิชาชีพครูครอบคลุมความสามารถในการปฏิบัติงานและพฤติกรรมที่พึงปรารถนาของครู (สุวิทย์ มูลคำ 2542) ซึ่งกำหนดไว้ 11 มาตรฐานดังนี้

- มาตรฐานที่ 1 ปฏิบัติกิจกรรมทางวิชาการเกี่ยวกับการพัฒนาวิชาชีพครูอยู่เสมอ
- มาตรฐานที่ 2 ตัดสินใจปฏิบัติกิจกรรมต่าง ๆ โดยคำนึงถึงผลที่จะเกิดกับผู้เรียน
- มาตรฐานที่ 3 มุ่งมั่นพัฒนาผู้เรียนให้เต็มตามศักยภาพ

- มาตรฐานที่ 4 พัฒนาแผนการสอนให้สามารถปฏิบัติได้เกิดผลจริง
- มาตรฐานที่ 5 พัฒนาสื่อการเรียนการสอนให้มีประสิทธิภาพอยู่เสมอ
- มาตรฐานที่ 6 จัดกิจกรรมการเรียนการสอนโดยเน้นผลถาวรที่เกิดแก่ผู้เรียน
- มาตรฐานที่ 7 รายงานผลการพัฒนาคุณภาพของผู้เรียนได้อย่างมีระบบ
- มาตรฐานที่ 8 ปฏิบัติตนเป็นแบบอย่างที่ดีแก่ผู้เรียน
- มาตรฐานที่ 9 ร่วมมือกับผู้อื่นในสถานศึกษาอย่างสร้างสรรค์
- มาตรฐานที่ 10 ร่วมมือกับผู้อื่นอย่างสร้างสรรค์ในชุมชน
- มาตรฐานที่ 11 แสวงหาและใช้ข้อมูลข่าวสารในการพัฒนา

เกณฑ์การประเมินผลการปฏิบัติงานของครูตามแนวคิดของ Beerens

จากงานวิจัยของ Beerens, D.R. (2000) ได้สรุปเกณฑ์ในการประเมินการปฏิบัติการสอนแบบมืออาชีพของครูออกเป็น 4 ด้าน คือ

1. ด้านการวางแผนและการเตรียมการสอน คือมีความรู้ในสิ่งต่อไปนี้
 - 1.1 มีความรู้ในเนื้อหาที่สอน ความรู้ในความสัมพันธ์พื้นฐานที่ต้องมี มีความรู้ในเนื้อหาเกี่ยวกับวิชาครู
 - 1.2 มีความรู้ในเรื่องเกี่ยวกับนักเรียน มีความรู้เกี่ยวกับพัฒนาการ ทักษะการเรียนรู้ และความสนใจและมรดกทางวัฒนธรรมของผู้เรียน
 - 1.3 มีความรู้เกี่ยวกับการกำหนดเป้าหมายในการสอน สามารถกำหนดเป้าหมายในการสอนอย่างมีคุณค่า/ประโยชน์ มีความชัดเจนเหมาะสมกับนักเรียนที่มีความแตกต่างกันหรือความเท่าเทียมสมดุลงัน
 - 1.4 การแสดงความรู้จักแหล่งความรู้ต่าง ๆ แหล่งสำหรับการสอนและแหล่งสำหรับนักเรียน
 - 1.5 สามารถออกแบบการสอนที่ต่อเนื่องราบรื่น คือสามารถออกแบบ/เลือกใช้วัสดุ อุปกรณ์และแหล่งการสอน ใช้กิจกรรมการเรียนรู้ สนับสนุนการเรียนการสอนได้อย่างเหมาะสมมีความสามารถในการจัดโครงสร้างของหน่วยการเรียนรู้และบทเรียน หรือจัดกลุ่มทางการสอนได้อย่างเหมาะสม
 - 1.6 มีความสามารถในการประเมินการเรียนรู้ของนักเรียนเลือกใช้วิธีการประเมินที่สอดคล้องกับเป้าหมายในการสอน มีเกณฑ์และมาตรฐานการประเมินที่ชัดเจน และใช้ผลการประเมินเพื่อการวางแผนการสอนต่อไป

2. ด้านการจัดสิ่งแวดล้อมในห้องเรียน
 - 2.1 สร้างบรรยากาศแห่งความรักใคร่ปรองดอง การให้ความเคารพและสร้างสัมพันธภาพที่ดีทั้งระหว่างผู้เรียนด้วยกันและครูผู้สอนกับผู้เรียน
 - 2.2 สร้างวัฒนธรรมในการเรียนรู้ ให้ความสำคัญของเนื้อหา ส่งเสริมให้นักเรียนมีความภาคภูมิใจในผลงานของตนและมีแรงจูงใจใฝ่สัมฤทธิ์
 - 2.3 จัดระเบียบการปฏิบัติในห้องเรียน จัดแบ่งตำแหน่งหน้าที่การแบ่งกลุ่มในการสอน และการปฏิบัติหน้าที่มิใช่การสอน การจัดสื่อการใช้สื่อ การตรวจตราในการอาสาสมัครและเทียบเคียงมืออาชีพ
 - 2.4 การจัดการพฤติกรรมของนักเรียนติดตามพฤติกรรมของนักเรียนแสดงความคิดเห็น ตักเตือนแก้ไข พฤติกรรมที่ไม่เหมาะสมของนักเรียน
 - 2.5 ดูแลความปลอดภัย และจัดหาอุปกรณ์ในห้องเรียน จัดเตรียมเนื้อที่ว่างให้นักเรียนสามารถเข้าไปเรียนรู้และใช้แหล่งทางกายภาพได้
3. ด้านการจัดการเรียนการสอน
 - 3.1 มีการสื่อสารที่ชัดเจนถูกต้อง แม่นยำ กระชับ ทั้งในการพูดและการเขียนให้คำแนะนำ และ ข้อปฏิบัติที่สามารถสื่อได้ตรงกัน
 - 3.2 ใช้การอภิปรายและใช้คำถามที่ดีมีคุณภาพ ผู้เรียนมีส่วนร่วมในการเรียนการสอน
 - 3.3 ทำให้ผู้เรียนสนใจเรียนโดยอาจจะสอดแทรกกิจกรรมที่หลากหลาย การนำเสนอตัวอย่างเนื้อหา กิจกรรมและการมอบหมายงาน การแบ่งกลุ่ม การนิยามโครงสร้างเนื้อหาและการทำแบบฝึกหัด ใช้สื่อ/วัสดุและแหล่งการเรียนรู้
 - 3.4 ให้ผลย้อนกลับ ที่ถูกต้อง เหมาะสม ตามเวลา สร้างสรรค์ (ในเชิงก่อ) แก่ผู้เรียนเป็นรายบุคคล
 - 3.5 สามารถยืดหยุ่น ปรับปรุงบทเรียน และสามารถตอบสนองต่อนักเรียนที่มีความมุ่งมั่นในการพัฒนานักเรียน
4. ด้านความรับผิดชอบในวิชาชีพ
 - 4.1 มีการประเมินการสอนที่ถูกต้อง แม่นยำผลการประเมินนำไปใช้ในการปรับปรุงการสอนต่อไป
 - 4.2 มีการเก็บรักษา บันทึกข้อมูล ผลงานและความก้าวหน้าของนักเรียนอย่างถูกต้องตรงตามความเป็นจริง และบันทึกข้อมูลที่ไม่ใช่การสอน

- 4.3 ติดต่อกับผู้ปกครองนักเรียนแลกเปลี่ยนข้อมูลเพื่อความเข้าใจและร่วมกันพัฒนา
นักเรียน
- 4.4 สนับสนุน/มีส่วนช่วยเหลือโรงเรียน/ท้องถิ่น เสียสละ มีมนุษยสัมพันธ์ที่ดีกับเพื่อน
ร่วมงาน
- 4.5 เพิ่มพูนความรู้ในเนื้อหาและทักษะในวิชาครู เผยแพร่และพัฒนาในด้านวิชาชีพ
ให้บริการชุมชนในฐานะผู้เชี่ยวชาญ
- 4.6 แสดงออกถึงความเป็นมืออาชีพเช่นการตัดสินใจการเป็นผู้สนับสนุน การให้
บริการแก่นักเรียน

Seyfart (1991) ได้เสนอเกณฑ์ในการพิจารณาการประเมินผลการปฏิบัติงานของครูดังนี้

1. ความรู้ในเนื้อหาวิชาการ
2. การเตรียมตัวก่อนสอนและการวางแผนการสอน
3. วิธีการถ่ายทอดและวิธีการใช้สื่อ วัสดุ อุปกรณ์ ประกอบการสอน
4. การประเมินผลนักเรียน
5. ความสามารถในการสร้างบรรยากาศในห้องเรียน

Gibson and Hunt (1965) ได้ให้แนวคิดในการสร้างเกณฑ์การประเมินผลการปฏิบัติงานของครู 3 ด้านต่อไปนี้

1. ด้านความรู้ คือรู้ในเรื่องการสอน เชี่ยวชาญเป็นพิเศษในสายงานของตน รู้ใน
เนื้อหาสาระที่จะต้องสอน รู้เกี่ยวกับพัฒนาการของเด็ก และพัฒนาการเรียนรู้
รู้ในการปฏิบัติงาน เทคนิควิธีการสอนต่าง ๆ
2. ด้านทักษะ คือมีทักษะในการใช้ภาษาพูดและเขียน สามารถถ่ายทอดความรู้
นักเรียนเข้าใจง่าย สามารถเข้าร่วมสัมพันธ์กับผู้อื่น มีมนุษยสัมพันธ์ดี มีทักษะ
ในการค้นคว้าวิจัยศึกษาหาความรู้เพิ่มเติม การแก้ไขปัญหาต่าง ๆ อย่าง
ถูกต้องวิชาการ ตรงกับความเป็นจริง มีนิสัยรักอาชีพครู
3. ด้านความสนใจ ได้แก่ ความสนใจในตัวนักเรียน ดูแลเอาใจใส่อย่างดี สนใจใน
วิชาที่สอน สนใจที่จะถ่ายทอดความรู้แก่นักเรียน สนใจที่จะร่วมมือให้ความ
ช่วยเหลือแก่ผู้อื่น โดยเฉพาะเพื่อนร่วมงาน สนใจที่จะพัฒนาอาชีพครูให้
ก้าวหน้า

Shreeve และคณะ (1993) สรุปเกณฑ์ที่ใช้ในการประเมินผลการปฏิบัติงานของครูไว้ 22 ข้อ ดังนี้

1. มีความรู้ ความเข้าใจ ในเนื้อหาที่สอนอย่างชัดเจน
2. สามารถใช้สื่ออุปกรณ์ต่าง ๆ เพื่อช่วยให้การสอนมีประสิทธิภาพมากยิ่งขึ้น
3. มีเทคนิคการสอนที่หลากหลาย ช่วยให้นักเรียนเกิดความรู้ ความเข้าใจอย่างถูกต้องชัดเจน
4. สามารถใช้กิจกรรมต่าง ๆ สอดแทรกในเนื้อหาเพื่อให้เกิดการเรียนรู้ที่ดีได้อย่างเหมาะสม
5. ให้งานหรือกิจกรรมต่าง ๆ เพื่อให้เกิดการเรียนรู้มากยิ่งขึ้น
6. ดูแลเอาใจใส่ห้องเรียนให้เป็นระเบียบเรียบร้อย
7. สร้างบรรยากาศในห้องเรียนเพื่อก่อให้เกิดการเรียนรู้ที่ดี
8. สนใจความรู้ ความสามารถของนักเรียนแต่ละชั้น
9. มีการประเมินผลในเนื้อหาและทักษะด้านต่าง ๆ ของนักเรียนเป็นประจำและสม่ำเสมอ
10. สนใจในอุปสรรคปัญหาต่าง ๆ ที่เกิดขึ้นกับนักเรียนเป็นรายบุคคล
11. ให้คำแนะนำตักเตือนแก่นักเรียนที่ต้องการหรือมีปัญหา
12. ให้กำลังใจนักเรียนที่จะพัฒนาตนเองในด้านความรู้ ความประพฤติ ลักษณะนิสัยต่าง ๆ
13. มีความเป็นกันเองและยืดหยุ่นได้เสมอ
14. เป็นผู้มองการณ์ไกล ศึกษาหาความรู้เพิ่มเติมเสมอ
15. เสริมสร้างปลูกฝังทัศนคติ ค่านิยมที่ถูกต้องให้นักเรียน
16. มีความรับผิดชอบในหน้าที่
17. มีทัศนคติที่ดีต่อวิชาชีพครู
18. รู้จักประเมินตนเองในทุก ๆ ด้าน และยอมรับในข้อบกพร่องที่เกิดขึ้นเพื่อปรับปรุงแก้ไขตนเองอยู่เสมอ
19. มีมนุษยสัมพันธ์ที่ดีต่อบุคคลทุกระดับ
20. เป็นผู้ที่คอยเอาใจใส่ให้ความช่วยเหลือในกรณีที่เกิดปัญหาต่าง ๆ ในโรงเรียน
21. ติดต่อกับผู้ปกครองนักเรียนเพื่อให้เกิดความเข้าใจอันดีต่อกัน

22. มีความซื่อสัตย์สุจริต มั่นคงในอารมณ์และในการตัดสินใจ

จากแนวคิดการประเมินผลการปฏิบัติงานของครูดังได้กล่าวมานี้ผู้วิจัยได้ปรับและนำมาประยุกต์ใช้กับการวิจัยในครั้งนี้ โดยพิจารณาผลการปฏิบัติงานของครูในห้องเรียน 5 ด้าน ดังนี้

- ด้านที่ 1 การสื่อสารการใช้ภาษาพูดภาษาเขียนของครู
- ด้านที่ 2 การสร้างบรรยากาศการเรียนรู้ในห้องเรียน
- ด้านที่ 3 การกำหนดกิจกรรมและการมอบหมายงาน
- ด้านที่ 4 คุณภาพของข้อมูลป้อนกลับให้แก่ักเรียน(คุณภาพด้าน:ความถูกต้อง, หลักฐาน, การวิจารณ์เชิงสร้างสรรค์, ชี้ประเด็นให้แก้ไข)
- ด้านที่ 5 การแสดงความยืดหยุ่นในการปรับบทเรียน

ตอนที่ 7 งานวิจัยที่เกี่ยวข้อง และกรอบความคิดในการวิจัย

7.1 งานวิจัยที่เกี่ยวข้อง

Beerens (2000) ได้ทำการวิจัยประเมินคุณภาพการปฏิบัติงานของครูเปรียบเทียบกับระหว่างครูใหม่และครูผู้มีประสบการณ์ในการสอน โดยมีการทดสอบเป็นระยะ ๆ และการให้ครูรายงานตนเอง ในช่วงสองปีแรก ในปีที่ 3 เป็นการประเมินเพื่อพัฒนาและปรับปรุง โดยมีหลักการ ดังนี้คือ 1) เป็นการประเมินเพื่อสรุปความเข้าใจ 2) ไม่ใช่การตรวจสอบพฤติกรรมการสอน 3) ไม่ใช่การจับผิดเพื่อบันทึกความผิดในการให้ใบอนุญาตของรูปแบบการสอนที่เฉพาะ และ 4) สามารถใช้กับบริบทต่าง ๆ กันในแนวทางที่ต่างจากผลการวิจัยเชิงประจักษ์ ในการวิจัยครั้งนี้มีโครงสร้างของเกณฑ์ในการปฏิบัติการสอนที่เป็นมืออาชีพในการประเมินผลการปฏิบัติงานของครูออกเป็น 4 ด้าน คือ 1) ด้านการวางแผนและการเตรียมการสอน 2) ด้านการจัดสิ่งแวดล้อมในห้องเรียน 3) ด้านการจัดการเรียน การสอน 4) ด้านความรับผิดชอบในวิชาชีพ

Morrison (1996) ได้ศึกษาผลการปฏิบัติงานของพนักงานธุรกิจที่ประสบความสำเร็จและได้พัฒนาความพึงพอใจในงาน เครื่องมือในงานวิจัยคือแบบสอบถามแบบมาตราส่วนค่า เกือบรวบรวมข้อมูลโดยการสำรวจจากพนักงานโรงงานอุตสาหกรรม ในประเทศสหรัฐอเมริกาจำนวน 307 คน โดยสรุปได้ผลการวิจัยว่าปัจจัยที่มีผลต่อผลการการปฏิบัติงานของพนักงานคือ ลักษณะงาน และลักษณะบุคคล และความพึงพอใจในงานเป็นตัวทำนายผลการปฏิบัติงานได้ด้วย

Naizer (1992) ได้เสนอแนวคิดพื้นฐานของทฤษฎีการสรุปอ้างอิง โดยเป็นทฤษฎีที่มีวิถีทางที่มีพลังในการเฝ้าความเที่ยง ทฤษฎีนี้มีความสัมพันธ์อย่างใกล้ชิดกับทฤษฎีการวัดแบบดั้งเดิม (Classical Test Theory) เขาเสนอว่าทฤษฎีการสรุปอ้างอิง เป็นทฤษฎีที่เกี่ยวกับความเชื่อถือได้ของการวัดพฤติกรรมซึ่งยอมให้มีการประมาณค่าความแปรปรวนของความคลาดเคลื่อนจากหลาย ๆ แหล่งได้พร้อม ๆ กัน ทฤษฎีนี้ได้แสดงความแตกต่างระหว่างการตัดสินใจแบบสัมบูรณ์และแบบสัมพัทธ์และเป็นทฤษฎีที่มีกลไกในการประมาณความแปรปรวนของความคลาดเคลื่อนในชั้นการศึกษาเพื่อการตัดสินใจ (D-Studies) อันเป็นสิ่งที่ช่วยนักวิจัยในการพัฒนาการวัดที่จะลดความคลาดเคลื่อน นอกจากนี้ เขาได้เสนอตัวอย่างของการใช้ทฤษฎีการสรุปอ้างอิง จากการศึกษาดัวอย่าง 20 คนโดยให้ทำงานด้านปฏิบัติ จำนวน 3 ครั้งและประเมินโดยผู้ประเมินจำนวน 2 คน

Crawford (1990) ได้ศึกษาความเที่ยงและความตรงของการประเมินครูในรัฐเท็กซัส ข้อมูลในการประเมิน ได้รวบรวมจากโรงเรียนชานเมืองแห่งหนึ่งในรัฐเท็กซัสซึ่งมีจำนวนนักเรียนประมาณ 6,000 คน ครู 386 คน และใช้ผู้ประเมิน 21 คนโดยใช้ทฤษฎีการสรุปอ้างอิง ในการประเมินองค์ประกอบของความแปรปรวนจากตัวแปรต่อไปนี้ เวลา (time) ผู้ประเมิน (reter) และข้อกระทง (item) ผลการศึกษาโดยใช้ทฤษฎีการสรุปอ้างอิง พบว่า ความแปรปรวนในคะแนนของครูมีมากกว่าแหล่งความแปรปรวนอื่น ๆ สำหรับผลการศึกษาคำตรงต่าง ๆ ได้แก่ ความตรงเชิงเนื้อหา ซึ่งได้จากการตรวจสอบ วรรณกรรมของการประเมินครูและงานวิจัยของการสอนที่มีประสิทธิผลได้แสดงให้เห็นว่า การสอนที่มีประสิทธิผลเป็นสิ่งที่มีความสำคัญต่อการปรับปรุงคะแนนการสอบของนักเรียน พฤติกรรมของครูที่แตกต่างกันก็จะมีส่วนสำคัญในการแก้ปัญหาให้นักเรียนสำหรับความตรงเชิงโครงสร้าง ได้ตรวจสอบจากการใช้การวิเคราะห์ความแปรปรวน โดยมีตัวแปรที่ใช้ในการเปรียบเทียบในการประเมินครู คือ วิทยาเขต ระดับ สถานภาพด้านคุณธรรม เชื้อชาติ เพศ และจำนวนปีที่มีประสบการณ์ ผลการศึกษาพบว่า วิทยาเขต เพศ และจำนวนปีที่มีประสบการณ์ มีนัยสำคัญในความแปรเปลี่ยนของคะแนนครู นอกจากนี้ได้ศึกษาความตรงเชิงเกณฑ์สัมพันธ์โดยทำการเปรียบเทียบคะแนนของครูในกลุ่ม 5 เปอร์เซนต์ คือกลุ่มที่ได้รับการบ่งชี้จากครูใหญ่ว่าเป็นผู้ที่มีคุณธรรมสูงกับคะแนนของครูในกลุ่มที่เหลือ ผลการศึกษาพบว่าความตรงเชิงเกณฑ์สัมพันธ์ได้รับการสนับสนุนจากกลุ่มครูที่มีคุณธรรม มากกว่ากลุ่มครูที่มีคุณธรรมน้อยกว่า อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05

Choi (1989) ได้ใช้ทฤษฎีการสรุปอ้างอิง ในการวิเคราะห์แหล่งของความแปรปรวนที่ได้จากการประเมินพฤติกรรมของครูในระบบชั้นเรียน และการประมาณค่าองค์ประกอบความแปรปรวนของความคลาดเคลื่อน ที่เนื่องมาจากผู้ประเมิน (rater) โอกาส (occasion) ครู (teacher) และปฏิสัมพันธ์ ผลจากการศึกษาพบว่า องค์ประกอบความแปรปรวนสัมพัทธ์ขนาดใหญ่เกิดเนื่องจากฟุ้งโอกาส ในขณะที่องค์ประกอบความแปรปรวนของผู้ประเมินมีค่าน้อยมาก ในส่วนของปฏิสัมพันธ์นั้น องค์ประกอบความแปรปรวนของปฏิสัมพันธ์ระหว่างโอกาสและครู มีค่ามากกว่าปฏิสัมพันธ์ของครูและผู้ประเมินนอกจากนี้ความแปรปรวนที่เนื่องจากโอกาส มีขนาดใหญ่กว่าความแปรปรวนที่มาจากผู้ประเมิน ดังนั้น การเพิ่มจำนวนโอกาสจะทำให้เกิด ค่าสัมประสิทธิ์การสรุปอ้างอิง มีค่าสูงกว่าสัมประสิทธิ์การอ้างอิงสรุปที่มาจากเพิ่มจำนวน ผู้ประเมิน

Cook (1988) ได้ศึกษาผลของการใช้มาตราประมาณค่าที่มีรูปแบบแตกต่างกัน 2 รูปแบบ คือ มาตราประมาณค่าเชิงพฤติกรรม (Behavior Anchor Rating Scale; BARS) กับมาตราประมาณค่าแบบ Likert ซึ่งมีจำนวนช่วงมาตร 6 ช่วง และผลของการฝึกทักษะแก่ ผู้ประเมินที่มีต่อค่าความคลาดเคลื่อนแบบกต/ปล่อยคะแนน (leniency error) และความคลาดเคลื่อนแบบฮาโล (halo error) โดยผู้วิจัยได้ใช้มาตราประมาณค่า 2 รูปแบบดังกล่าวให้นักศึกษา 2 กลุ่ม คือ กลุ่มที่ได้รับการฝึกทักษะกับกลุ่มที่ไม่ได้รับการฝึกทักษะทำการประเมินพฤติกรรมการสอนของศาสตราจารย์ 1 ท่านแล้วทำการวิเคราะห์หาค่าเฉลี่ยเลขคณิต และส่วนเบี่ยงเบนมาตรฐานของคะแนนผลการประเมินแต่ละมิติเปรียบเทียบระหว่างกลุ่มที่ได้รับการฝึกทักษะกับกลุ่มที่ไม่ได้รับการฝึกทักษะพร้อมทั้งเปรียบเทียบคะแนนผลการประเมินระหว่างมาตราประมาณค่าต่างรูปแบบกัน เฉพาะกลุ่มที่ได้รับการฝึกทักษะโดยใช้การตรวจสอบความคลาดเคลื่อนแบบฮาโล (halo error) ผู้วิจัยได้ใช้ส่วนเบี่ยงเบนมาตรฐานของคะแนนผลการประเมิน (standard deviation) เป็นคะแนนบ่งชี้คะแนนความคลาดเคลื่อนแบบฮาโล ภายใต้เงื่อนไขว่า ถ้าคะแนนความคลาดเคลื่อนแบบฮาโล มีค่ามากแสดงถึงการเกิดความคลาดเคลื่อนแบบฮาโล น้อยกว่ากรณีที่คะแนนความคลาดเคลื่อนแบบฮาโล มีค่าน้อย ผลการวิจัยพบว่าการฝึกทักษะมีผลต่อการลดความคลาดเคลื่อนแบบฮาโล การประเมินโดยมาตราประมาณค่าเชิงพฤติกรรม ความคลาดเคลื่อนแบบกต/ปล่อยคะแนน (leniency error) น้อยกว่า มาตราประมาณค่าแบบ Likert ในขณะที่มาตราประมาณค่าแบบ Likert เกิดความคลาดเคลื่อนแบบ ฮาโล น้อยกว่ามาตราประเมินค่าเชิงพฤติกรรม

Ibrahim (1984) ได้ประยุกต์ทฤษฎีการสรุปอ้างอิง เพื่อประมาณค่าความแปรปรวน ที่มีผลต่อการประเมินวัตถุประสงค์ทางการศึกษา (The Rating of Evaluational Goals) โดยสุ่มตัวอย่างครู 80 คน และนักศึกษา 80 คน ในประเทศชูดาน ประเมินวัตถุประสงค์ทางการศึกษา 2 ชนิด คือ วัตถุประสงค์ที่สำคัญจริง ๆ และวัตถุประสงค์ที่คาดหวังองค์ประกอบในการศึกษาคือ ผู้ประเมิน กลุ่มผู้ประเมิน จำนวนครั้งของการประเมิน ถิ่นที่อยู่ของผู้ประเมิน ชนิดของวัตถุประสงค์ สถานที่ทำงานของผู้ประเมิน และเพศของผู้ประเมิน ผลการวิจัยพบว่าองค์ประกอบที่มีผลต่อการประเมินมากที่สุดได้แก่ผู้ประเมินและกลุ่มผู้ประเมินส่วนสถานที่ทำงานของผู้ประเมิน มีผลเล็กน้อย ส่วนองค์ประกอบอื่นที่เหลือไม่มีผลต่อการประเมินเลย

เขาวลักษณะ แสงสร้อย (2542) ศึกษาปัจจัยที่ส่งผลกระทบต่อประสิทธิภาพการปฏิบัติงานวิชาการของผู้บริหารโรงเรียนประถมศึกษาโดยใช้การวิเคราะห์ด้วยโมเดลเชิงเส้นตรงระดับลดหลั่น (hierarchical linear model= HLM) เก็บรวบรวมข้อมูลจากผู้บริหารโรงเรียนประถมศึกษาและครูหัวหน้าฝ่ายวิชาการของโรงเรียนประถมศึกษาในสังกัดสำนักงานคณะกรรมการการประถมศึกษาแห่งชาติ เขตการศึกษา 7 จังหวัดนครสวรรค์ จำนวน 513 โรงเรียน ผลการวิจัยพบว่า ตัวแปรระดับ โรงเรียนที่มีอิทธิพลทางบวกอย่างมีนัยสำคัญทางสถิติต่อประสิทธิภาพการปฏิบัติงานวิชาการของผู้บริหารโรงเรียนได้แก่ ความเป็นผู้นำทางวิชาการของผู้บริหารโรงเรียนและจำนวนครั้งที่ได้รับการอบรมงานวิชาการในช่วง 3 ปีของผู้บริหารโรงเรียน ตัวแปรระดับโรงเรียนที่มีอิทธิพลทางลบอย่างมีนัยสำคัญทางสถิติต่อประสิทธิภาพการปฏิบัติงานวิชาการของผู้บริหารโรงเรียนได้แก่ อัตราส่วนครูต่อนักเรียนและระยะทางระหว่างโรงเรียนกับอำเภอ ตัวแปรระดับชุมชนที่มีอิทธิพลทางบวกอย่างมีนัยสำคัญทางสถิติต่อประสิทธิภาพการปฏิบัติงานวิชาการของผู้บริหารโรงเรียน ได้แก่ ค่าเฉลี่ยของจำนวนเงินที่ได้รับบริจาคจากบุคคลหรือหน่วยงานหรือองค์กรในชุมชน ค่าเฉลี่ยของปริมาณความร่วมมือของชาวบ้านในชุมชน ค่าเฉลี่ยของปริมาณความร่วมมือของผู้บริหารภายในกลุ่มโรงเรียน และสัดส่วนของผู้บริหารโรงเรียนที่พักอาศัยอยู่ในชุมชน

บวร คุ่มศิริพิทักษ์ (2541) ศึกษาเปรียบเทียบความคลาดเคลื่อนแบบฮาโล (halo error) ความคลาดเคลื่อนแบบขัดแย้ง (contrast error) ความคลาดเคลื่อนแบบกต/ปล่อยคะแนน (leniency error) โดยใช้แบบประเมินคุณภาพการสอนของครูซึ่งมีลักษณะเป็นมาตราประมาณค่าที่มีจำนวนช่วงมาตรฐานเป็น 4, 5, 6 และ 7 ช่วง กลุ่มตัวอย่างเป็นครุฑมวดวิทยาศาสตร์ 13 คน และนักเรียนสาย คณิต-วิทย์ ระดับชั้นมัธยมศึกษาปีที่ 5 จำนวน 27 คน โดยให้กลุ่มตัวอย่าง แบบ

ประเมินคุณภาพการสอนของครู วิทยาศาสตร์ 2 คน นำข้อมูลมาวิเคราะห์ความแปรปรวนสองทาง (two-way ANOVA) เพื่อวิเคราะห์เปรียบเทียบความแปรปรวนแต่ละชนิดผลการวิจัยพบว่า นักเรียนประเมินคุณภาพการสอนของครู โดยใช้มาตร 4 ช่วง 6 ช่วง 7 ช่วง มีความคลาดเคลื่อนแบบกด/ปล่อยคะแนน (leniency error) ในขณะที่ มาตร 5 ช่วง ไม่พบความคลาดเคลื่อนทั้ง 3 ชนิด ส่วนครูเพื่อนร่วมงานประเมินคุณภาพการสอนของครูมาตร 4 ช่วง 6 ช่วง มีความคลาดเคลื่อนแบบกด/ปล่อยคะแนน (leniency error) ในขณะที่ มาตร 5 ช่วง 7 ช่วง ไม่พบความคลาดเคลื่อนทั้ง 3 ชนิดนอกจากนี้ยังไม่พบ ปฏิสัมพันธ์ระหว่างจำนวนช่วงมาตรกับ สถานภาพ ของผู้ประเมินในการเกิดความคลาดเคลื่อนทั้ง 3 ชนิด

สังวรณ์ รัตกระโทก (2541) ใช้โมเดลสมการโครงสร้างพหุระดับ ตรวจสอบความตรงของ โมเดลแสดงความสัมพันธ์ระหว่างปัจจัยครู ปัจจัยโรงเรียนกับความพึงพอใจในการปฏิบัติงานของ ครู โดยการวิเคราะห์ด้วยโปรแกรมลิสเรล (Linear Structural RELationship model ; LISREL) ที่พัฒนาให้มีความสามารถในการวิเคราะห์ข้อมูลที่มีลักษณะลัดหลั่นได้ และเปรียบเทียบ ประสิทธิภาพของผลของการวิเคราะห์ระหว่างโปรแกรมลิสเรล กับโปรแกรมเอชแอลเอ็ม (Hierarchical Linear Model : HLM) ข้อมูลที่ใช้ในการวิจัยเป็นข้อมูลทุติยภูมิจากโครงการวิจัย ประสิทธิภาพการใช้ครู: การวิเคราะห์เชิงปริมาณระดับมหภาค ของสำนักงานคณะกรรมการการ ศึกษาแห่งชาติ กลุ่มตัวอย่างที่ใช้ประกอบด้วยครูจำนวน 9,599 คน และผู้บริหารโรงเรียนจำนวน 1,290 คน จากโรงเรียน 1,290 โรงเรียนในสังกัดสำนักงานการศึกษากรุงเทพมหานคร สำนักงาน การศึกษาท้องถิ่น สำนักงานคณะกรรมการการประถมศึกษาแห่งชาติ กรมสามัญศึกษาและสำนักงาน คณะกรรมการการศึกษาเอกชน ตัวแปรที่ใช้ในการวิจัยมีจำนวน 16 ตัวแปร ประกอบด้วยตัว แปรระดับครู จำนวน 7 ตัวแปร และตัวแปรระดับโรงเรียน 9 ตัว การวิเคราะห์ข้อมูลใช้สถิติบรรยาย การวิเคราะห์สหสัมพันธ์ การวิเคราะห์ LISREL และการวิเคราะห์ HLM ผลการวิจัยสรุปได้ว่าตัว แปรระดับครูที่มีอิทธิพลอย่างมีนัยสำคัญต่อความพึงพอใจในการปฏิบัติงานของครูมากที่สุดมี 5 ตัว ได้แก่ ความก้าวหน้ารองลงมาคือ ความผูกพันต่ออาชีพครู คุณภาพของงาน เพศชาย และ จำนวน ปีที่ปฏิบัติงานในโรงเรียน ตามลำดับ ส่วนอิทธิพลของตัวแปรระดับโรงเรียนทั้ง 9 ตัวแปร มีอิทธิพล ต่อค่าเฉลี่ยความพึงพอใจในการปฏิบัติงานของครูอย่างไม่มีนัยสำคัญทางสถิติ แต่มีตัวแปรตัวที่มี โรงเรียนสังกัดกรุงเทพมหานครส่งอิทธิพลต่อค่าเฉลี่ยความผูกพันต่ออาชีพของครู โรงเรียนสังกัด สำนักงานการศึกษาท้องถิ่นส่งอิทธิพลต่อค่าเฉลี่ยความก้าวหน้าและค่าเฉลี่ยความผูกพันต่ออาชีพ โรงเรียนสังกัดสำนักงานการประถมศึกษาแห่งชาติและโรงเรียนสังกัดสำนักงานคณะกรรมการ

การศึกษาเอกชน ส่งอิทธิพลต่อค่าเฉลี่ยความก้าวหน้า ค่าเฉลี่ยคุณภาพของงาน และค่าเฉลี่ยความผูกพันต่ออาชีพ ประสิทธิภาพของผลการวิเคราะห์ด้วย LISREL และ HLM แสดงว่าแต่ละโปรแกรมมีลักษณะต่างกัน โปรแกรม LISREL ให้ผลการวิเคราะห์ในภาพรวมแต่ยังไม่สามารถวิเคราะห์ความชันในฐานะตัวแปรตามสุ่ม ในขณะที่โปรแกรม HLM วิเคราะห์ความชันในฐานะตัวแปรตามสุ่มได้ แต่ยังไม่สามารถให้ผลการวิเคราะห์แยกเป็นอิทธิพลทางตรงและทางอ้อมได้ รวมทั้งยังไม่สามารถวิเคราะห์โดยยอมให้เทอมความคลาดเคลื่อนสัมพันธ์กันได้ สัมประสิทธิ์ไชว์ระดับจากโปรแกรม LISREL และ HLM ส่วนใหญ่มีความแตกต่างกันทั้งขนาดและทิศทาง แต่เมื่อนำสัมประสิทธิ์ที่ได้จากโปรแกรม HLM ไปกำหนดในการวิเคราะห์ด้วยโปรแกรม LISREL พบว่าสัมประสิทธิ์จากโปรแกรม HLM ทำให้โมเดลค่อนข้างสอดคล้องกับข้อมูลเชิงประจักษ์

โดยสรุปจากการศึกษารายงานวิจัยที่เกี่ยวข้องกับการวิจัยในครั้งนี้ 10 เรื่อง มีรายงานวิจัยเกี่ยวกับการประเมินผลการปฏิบัติงาน 3 เรื่อง โดยใช้แบบสอบถามมาตรฐานค่ามีรายงานการวิจัยที่ใช้ทฤษฎีการสรุปอ้างอิงในการประมาณค่าองค์ประกอบแหล่งความแปรปรวนของความคลาดเคลื่อนในการประเมิน 4 เรื่องทั้ง 4 เรื่องไม่ได้นำเสนอการปรับแก้ความคลาดเคลื่อนให้ได้ผลการประเมินที่ถูกต้องเพียงแต่ประมาณค่าความแปรปรวนของความคลาดเคลื่อนและประมาณค่าความเที่ยงและความตรงของการประเมินมีรายงานวิจัย 1 เรื่องใช้มาตรฐานค่าเชิงพฤติกรรม (BARS) และมาตรฐานค่าแบบ Likert ในการตรวจสอบความคลาดเคลื่อนแบบฮาโล มีรายงานวิจัย 1 เรื่องที่เปรียบเทียบปริมาณความคลาดเคลื่อนแบบฮาโล แบบกด/ปล่อยคະแนน และแบบขัดแย้ง โดยใช้การวิเคราะห์ความแปรปรวนสถานภาพต่าง ๆ กัน โดยไม่มีการประมาณค่าและปรับแก้ความคลาดเคลื่อน ด้วยเหตุนี้ผู้วิจัยจึงสนใจทำการวิจัย ประมาณค่าความคลาดเคลื่อนและปรับแก้ ความคลาดเคลื่อนรวมทั้งป้องกันความคลาดเคลื่อนจากการประเมินผลการปฏิบัติงาน

7.2 กรอบความคิดในการวิจัย

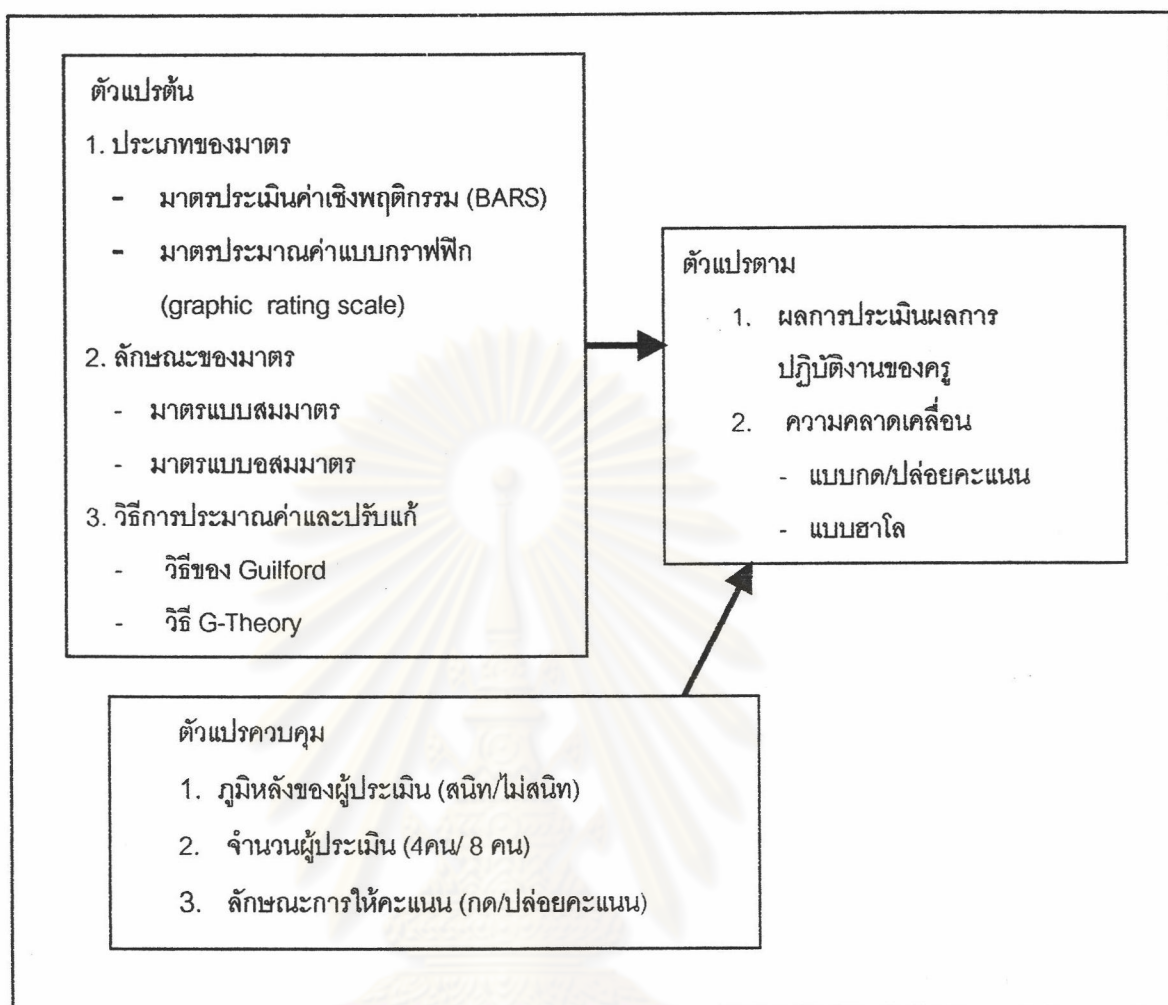
การวิจัยครั้งนี้ผู้วิจัยสนใจศึกษา (1) เปรียบเทียบความสอดคล้องของผลการประเมินผลการปฏิบัติงานของครูและความคลาดเคลื่อนแบบกด/ปล่อยคະแนน (leniency error) ระหว่างวิธีการ 2 วิธีคือใช้วิธีป้องกันความคลาดเคลื่อนแบบกด/ปล่อยคະแนนโดยการใช้มาตรฐานแบบสมมาตร และ วิธีการใช้มาตรฐานแบบสมมาตรร่วมกับการประมาณค่าและปรับแก้ความคลาดเคลื่อนแบบกด/ปล่อยคະแนน (leniency error) (2) เปรียบเทียบความสอดคล้องของผลการประเมินผลการปฏิบัติงานของครูและความคลาดเคลื่อนแบบฮาโล (halo error) ระหว่างวิธีการ 2 วิธีคือ วิธี

ป้องกันความคลาดเคลื่อนแบบฮาโลโดยการใช้อนุกรมประเมินค่าเชิงพฤติกรรม (BARS) และวิธีการ
 ใช้อนุกรมประเมินค่าแบบกราฟิก (graphic rating scale) ร่วมกับการประเมินค่าและปรับแก้
 ความคลาดเคลื่อนแบบฮาโล (halo error) และ (3) เปรียบเทียบความสอดคล้องของผลการ
 ประเมินผลการปฏิบัติงานของครูและความคลาดเคลื่อนแบบกต/ปล่อยคะแนน (leniency error)
 และความคลาดเคลื่อนแบบฮาโล (halo error) ระหว่างวิธีการ 3 วิธีคือ วิธีป้องกันความ
 คลาดเคลื่อนแบบกต/ปล่อยคะแนนและแบบฮาโล โดยใช้อนุกรมประเมินค่าเชิงพฤติกรรม (BARS)
 แบบอสมมาตร, วิธีการใช้อนุกรมประเมินค่าแบบกราฟิก (graphic rating scale) แบบสมมาตร
 ร่วมกับการประเมินค่าและปรับแก้ความคลาดเคลื่อนของ Guilford และวิธีการวิเคราะห์ความ
 แปรปรวนโดยใช้ทฤษฎีการสรุปอ้างอิง (Generalizability Theory)

ซึ่งสรุปได้ดังภาพที่ 3



คุรุวิทยาลัย
 จุฬาลงกรณ์มหาวิทยาลัย



ภาพที่ 3 กรอบความคิดในการวิจัย

การวิจัยครั้งนี้คือต้องการศึกษาเปรียบเทียบคะแนนการประเมินผลการปฏิบัติงานและค่าประมาณความความคลาดเคลื่อนสองชนิดคือ ความคลาดเคลื่อนแบบกด/ปล่อยคะแนน และความคลาดเคลื่อนแบบฮาโล โดยมีตัวแปรต้น 3 ตัวแปรคือ (1) ประเภทของมาตรแบ่งเป็นการใช้มาตรประเมินค่าเชิงพฤติกรรม (BARS) กับการใช้มาตรประเมินค่าแบบกราฟฟิก (graphic rating scale) (2) ลักษณะของมาตรที่ใช้ในการประเมินแบ่งเป็นมาตรแบบสมมาตรและมาตรแบบอสมมาตร (3) วิธีการประมาณค่าและปรับแก้ความคลาดเคลื่อนมีสองวิธีคือ วิธีการวิเคราะห์ของ Guilford และ วิธีการวิเคราะห์ความแปรปรวนโดยใช้ทฤษฎีการสรุปอ้างอิง (Generalizability Theory) โดยมีตัวแปรควบคุมในการศึกษาครั้งนี้ 3 ตัวแปรคือ (1) ภูมิหลังของผู้ประเมินแบ่งเป็นสนิทและไม่สนิทกับผู้ถูกประเมิน (2) จำนวนผู้ประเมินแบ่งเป็นกรณีเมื่อประเมินโดยผู้ประเมิน 4 คนและกรณีเมื่อประเมินโดยผู้ประเมิน 8 คน (3) ลักษณะการให้คะแนนของผู้ประเมิน แบ่งเป็นผู้ประเมินที่มีลักษณะการให้คะแนนแบบกดคะแนนและผู้ประเมินที่มีลักษณะการให้คะแนนแบบปล่อยคะแนน