

เอกสารและงานวิจัยที่เกี่ยวข้อง

การนำเสนอผลการศึกษาค้นคว้าเอกสาร งานวิจัยที่เกี่ยวข้องกับงานวิจัยในครั้งนี้ ผู้วิจัยได้นำเสนอผลการศึกษา โดยแบ่งออกเป็น 5 ตอน ดังต่อไปนี้

ตอนที่ 1 แนวความคิดเกี่ยวกับการประเมินการสอน

ตอนที่ 2 แนวความคิดเกี่ยวกับมาตรฐานค่า

- รูปแบบของมาตรฐานค่า
- ความคลาดเคลื่อนที่เกิดจากการใช้มาตรฐานค่า
- แนวทางการปรับปรุงคุณภาพของมาตรฐานค่า

ตอนที่ 3 แนวความคิดเกี่ยวกับมาตรฐานค่าแบบลิเคอร์ท์

- การสร้างมาตรฐานค่าแบบลิเคอร์ท์
- ข้อดีและข้อเสียของมาตรฐานค่าแบบลิเคอร์ท์

ตอนที่ 4 แนวความคิดเกี่ยวกับมาตรฐานค่าแบบยึดพฤติกรรม

- ลักษณะของมาตรฐานค่าแบบยึดพฤติกรรม
- ขั้นตอนการพัฒนามาตรฐานค่าแบบยึดพฤติกรรม
- สภาพการประยุกต์ใช้ มาตรฐานค่าแบบยึดพฤติกรรมในการวิจัย

ตอนที่ 5 งานวิจัยที่เกี่ยวข้องกับการศึกษาเปรียบเทียบคุณภาพของมาตรฐานค่าในด้านความเที่ยงและความคลาดเคลื่อนในการประมาณค่า

ตอนที่ 1 แนวความคิดเกี่ยวกับการประเมินการสอน

ในการจัดการศึกษาให้บรรลุตามจุดหมายของหลักสูตร จำเป็นต้องอาศัยองค์ประกอบที่เกี่ยวข้องหลายด้านด้วยกัน การสอนนับเป็นองค์ประกอบหนึ่งที่มีความสำคัญมากที่สุด ทั้งนี้เพราะเป็นกระบวนการในการเสริมสร้างให้ผู้เรียนได้เกิดความรู้สึกของอกงามทั้งร่างกาย อารมณ์ สังคม และสติปัญญา โดยผู้ที่มีหน้าที่เกี่ยวข้องโดยตรง ก็คือ ครูผู้สอน ซึ่งเป็นบุคคลที่จะนำพาผู้เรียนให้ประสบความสำเร็จในการเรียนรู้ ดังนั้นจึงจำเป็นอย่างยิ่งที่ครูผู้สอนจะต้องตระหนักถึงหน้าที่และความรับผิดชอบของตนในการปรับปรุงและพัฒนาการสอนของตนให้ดีและมีประสิทธิภาพอยู่เสมอ

การที่ครูผู้สอนจะรู้ว่าการสอนของตนประสบความสำเร็จหรือมีประสิทธิภาพมากน้อยเพียงใด นั้น จำเป็นที่ครูจะต้องมีการประเมินผลการสอนของตนเอง เพื่อให้รู้การสอนของตนที่เป็นอยู่มีข้อดี ข้อบกพร่องอะไรบ้างที่ควรนำมาปรับปรุงแก้ไขให้ดีขึ้น ซึ่งเกี่ยวกับเรื่องของการประเมินผลการสอนของครูนั้น ไพทวีย์ สีนลาร์ตน (2624) ได้กล่าวสรุปว่า โดยทั่วไปแล้วการประเมินผลการสอนของครูจะมีจุดมุ่งหมายดังนี้

1. เพื่อปรับปรุงการสอนโดยที่การประเมินการสอนจะเป็นเสมือนกระจกเงาสท้อนพฤติกรรมการสอนให้ครูผู้สอนเห็นการสอนของตนเองว่าดี หรือไม่ดีอย่างไรจะได้นำส่วนไม่ดีมาปรับปรุงแก้ไขให้ดีขึ้น
2. เพื่อการตัดสินใจทางการบริหาร โดยนำผลการสอนมาใช้ประกอบการพิจารณาตัดสินใจทางการบริหาร อันได้แก่ เพื่อเลื่อนตำแหน่งทางวิชาการ เพื่อเลื่อนขั้นเงินเดือน เพื่อบริหารการเรียนการสอน
3. เพื่อประกอบการเลือกเรียนของนักศึกษา ผลการประเมินประสิทธิภาพการสอนของอาจารย์จะเป็นประโยชน์สำหรับนักศึกษาที่จะใช้ประกอบการเลือกที่จะเรียนวิชาใด และเป็นแนวทางให้นักศึกษาได้ปรับตัวเข้ากับวิชาที่เลือกเรียนนั้นด้วย
4. เพื่อการวิจัย ซึ่งส่วนใหญ่ทำเฉพาะในหมู่นักวิชาการ ทั้งนี้มุ่งนำผลไปพัฒนาการเรียนการสอนโดยตรง เช่น การศึกษาลักษณะการสอนที่มีประสิทธิภาพ รูปแบบการสอนที่ดี เป็นต้น

ส่วนผู้ที่มีส่วนเกี่ยวข้องในการประเมินการสอนนั้น อุทุมพร ทองอุไทย (2623) ได้กล่าวว่า ผู้ประเมินการสอนนั้นเป็นอีกสิ่งหนึ่งที่ต้องพิจารณาเพราะอาจมีความคลาดเคลื่อนจากผู้ให้ข้อมูลได้ ดังนั้นในการประเมินจึงควรให้ผู้ประเมินจากบุคคลหลายกลุ่มประกอบกัน อันได้แก่

1. ตัวอาจารย์เอง

เนื่องจากจะเป็นผู้หนึ่งที่อยู่ตัวว่าการสอนของตนเป็นเช่นไรอย่างไรก็ดี การให้ครูผู้สอนประเมินตนเองนั้นมักจะประสบปัญหาในเรื่องของการนำความลำเอียงเข้ามาเกี่ยวข้องฉะนั้นถ้าให้ครูผู้สอนประเมินการสอนของตนเองอย่างมีระบบ โดยปราศจากอคติใด ๆ แล้ว ข้อมูลที่ได้น่าจะเที่ยงตรงและเชื่อถือได้มาก

2. เพื่อนอาจารย์

ซึ่งควรใช้เพื่อนอาจารย์ที่ทำงานอยู่ในภาควิชาหรือสาขาเดียวกัน เนื่องจากจะรู้จักจุดมุ่งหมายของวิชาและใกล้ชิดกับอาจารย์ที่ถูกประเมิน สามารถให้ข้อมูลบางอย่างที่ถูกต้องและชัดเจนได้เป็นอย่างดี โดยจะให้ผลดีอย่างแท้จริง ทั้งอาจารย์ที่ถูกประเมินและเพื่อนอาจารย์ ต้องใจกว้าง เข้าอกเข้าใจกัน มีความจริงใจต่อกัน และพร้อมที่จะรับฟังซึ่งกันและกัน

3. ผู้บริหาร
ซึ่งเป็นผู้หนึ่งที่มีส่วนในการประเมินการสอนของอาจารย์ทั้งนี้เพราะมีความใกล้ชิดและเป็นผู้พิจารณาความคิดเห็นความชอบของอาจารย์
4. นักศึกษา
เป็นบุคคลที่รับรู้ได้เห็นและติดตามกิจกรรมการสอนของอาจารย์ตลอดเวลาและเป็นผู้ได้รับผลจากการสอนของอาจารย์โดยตรง
5. คณะกรรมการเฉพาะกิจ หรือผู้เชี่ยวชาญทางการประเมินประสิทธิภาพการสอน
ซึ่งมักจะให้ประเมินประสิทธิภาพการสอนรวมๆของอาจารย์ทั้งภาควิชาคณะหรือมหาวิทยาลัย เพื่อประโยชน์ในการบริหารและการจัดการทางการศึกษา

ไพฑูรย์ สีนลาร์ตัน (2524) ได้ให้แนวคิดเพิ่มเติมว่า แหล่งที่มาของข้อมูลในการประเมินการสอนของครูนั้นควรมาจาก

1. อาจารย์ผู้สอนเอง เนื่องจากเป็นบุคคลที่สำคัญที่สุดในการสอนและปรับปรุงการสอน ซึ่งการให้ผู้สอนประเมินผลการสอนของตนเอง จะทำให้ได้ข้อมูลที่ชัดเจนที่สุด แต่ในขณะเดียวกันบางครั้งก็ทำให้มีการลำเอียงเข้าข้างตนเอง
2. จากเพื่อนร่วมงาน โดยหลักการแล้วการได้ข้อมูลจากเพื่อนร่วมงานเป็นสิ่งที่ดี และควรกระทำเพราะเพื่อนร่วมงานเป็นผู้ที่เข้าใจงานของเราเองอย่างดีที่สุด
3. จากผู้บังคับบัญชาหรือผู้ว่าจ้าง เป็นการประเมินผลโดยมุ่งไปที่การพิจารณาความคิดเห็นความชอบเป็นหลักสำคัญ ผู้บังคับบัญชาที่จะเป็นแหล่งข้อมูลที่ดีนั้นจะต้องสนใจและติดตามผลการสอนโดยตลอด
4. จากผลการสอบ ผลการสอบของนักศึกษาจะเป็นข้อมูลหรือที่มาของการประเมินผลได้ถ้าการสอนได้ดำเนินไปอย่างมีประสิทธิภาพ ตั้งจุดมุ่งหมายชัดเจน ดำเนินการสอนอย่างเหมาะสมและมีระบบการประเมินผลที่ดี
5. จากตัวนักศึกษาเอง ข้อมูลที่ได้จากตัวนักศึกษาเป็นข้อมูลที่น่าสนใจที่สุดและอาจเป็นประโยชน์มากที่สุดเพราะในการเรียนการสอน ผู้เรียนจะเป็นผู้รู้และรับผลโดยตรงจากอาจารย์ผู้สอน ดังนั้นการให้ผู้เรียนประเมินจึงเป็นเครื่องมือวัดที่ดีว่าผู้สอนสอนแล้วประสบความสำเร็จมากน้อยเพียงใด

ทั้งนี้ ในกรณีของการใช้ข้อมูลจากนักศึกษานั้น ถึงแม้ว่าข้อมูลจะดีและน่าสนใจเพียงใดก็ตาม แต่ก็มีความวิจารณ์อยู่หลายประการ เช่น ผู้เรียนยังไม่มีวุฒิภาวะพอ ผู้เรียนจะพอใจแต่อาจารย์ใจดีให้คะแนนดี ถ้าผลการประเมินออกมาต่ำจะทำให้เกิดการปกป้องตนเองของผู้สอน หรือทำให้ผู้เรียนไม่มีสัมมาคารวะต่อผู้สอน เป็นต้น

นอกจากนี้บอลลิงตันและคณะ (Bollington และคณะ, 1990) ได้กล่าวไว้ว่า ผู้ทำหน้าที่ประเมิน ซึ่งเป็นบุคคลต่าง ๆ ในโรงเรียน คือ ครูเป็นผู้ประเมินตนเอง ผู้บริหารเป็นผู้ประเมิน ผู้ร่วมงานซึ่งเป็นเพื่อนครูด้วยกันเป็นผู้ประเมิน และนักเรียนเป็นผู้ประเมิน

อย่างไรก็ตาม สิ่งที่จะพิจารณาต่อไปเกี่ยวกับการประเมินการสอนนั้น ก็คือ ประเด็นที่ใช้วัดประสิทธิภาพการสอน ซึ่งได้มีรายงานการวิจัยที่เกี่ยวข้องทั้งในและต่างประเทศที่สรุปได้สอดคล้องกันว่า การประเมินในประเด็น ดังต่อไปนี้

1. บุคลิกลักษณะของอาจารย์
2. ความสามารถทางวิชาการ
3. เจตคติของอาจารย์ต่อวิชาที่สอน
4. เจตคติของอาจารย์ต่อนักศึกษา
5. ความสัมพันธ์ของอาจารย์กับนักศึกษา
6. เทคนิคการสอน
7. การใช้วัสดุอุปกรณ์การเรียนการสอน
8. การวัดและประเมินผลการสอน

สำหรับเทคนิค หรือวิธีการเก็บรวบรวมข้อมูลเพื่อการประเมินการสอนนั้น ก็เป็นสิ่งหนึ่งที่สำคัญ และจำเป็นอย่างยิ่งที่จะต้องทำอย่างมีระบบเพื่อให้ได้ข้อมูลที่มีความตรงและความเที่ยงมากที่สุด ฉะนั้น การที่จะให้ได้ข้อมูลที่มีความตรงและความเที่ยงนั้นจึงขึ้นอยู่กับวิธีการรวบรวมข้อมูลและคุณภาพของเครื่องมือที่ใช้เป็นสำคัญ โดยวิธีการที่นำมาใช้ในการรวบรวมข้อมูลอาจแบ่งได้ 3 วิธี ดังนี้ (บุญธรรม กิจปรีดาภิรุตย์, 2535)

1. การสังเกต ซึ่งเป็นการรวบรวมข้อมูลด้วยการใช้ประสาทสัมผัส โดยเฉพาะตา โดยตรง โดยผู้สังเกตจะเข้าไปร่วมอยู่ในห้องเรียน คอยติดตามการสอนและพฤติกรรมของอาจารย์ที่ถูกประเมิน โดยอาจจะใช้อุปกรณ์ช่วยสังเกต เช่น วีดีโอเทป เทป บันทึกภาพและเสียงไว้ โดยเฉพาะการใช้วีดีโอเทปนั้น แม้จะมีราคาแพงแต่ก็นำไปใช้ในลักษณะประเมินตนเองเพื่อปรับปรุงการสอนโดยเฉพาะ จะได้ผลดีมาก

2. การสัมภาษณ์ โดยการพูดคุยซักถามกับผู้ที่มีข้อมูล เช่น นักศึกษา เพื่อนอาจารย์ ตัวอาจารย์เอง เป็นต้น วิธีการสัมภาษณ์อาจจะใช้ในลักษณะที่มีระบบแน่นอน มีการกำหนดคำถามไว้ล่วงหน้า หรือใช้วิธีที่ไม่มีระบบที่แน่นอนก็ได้

3. การใช้แบบสอบถาม การรวบรวมข้อมูลโดยใช้แบบสอบถามจัดเป็นวิธีที่นิยมกันทั่วไป เนื่องจากง่าย และสะดวกในการใช้ ซึ่งลักษณะแบบสอบถามที่ใช้ส่วนมากจะเป็นแบบประเมินค่า แบบตรวจสอบรายการ (Check list) และแบบให้เติมคำหรือเลือกตอบ ทั้งนี้ การใช้แบบสอบถามรวบรวมข้อมูลในการประเมินการสอนนั้น สิ่งสำคัญที่ต้องพิจารณา คือ ประเด็นที่ใช้ถาม ความครอบคลุมของเนื้อหา และความเป็นปรนัยได้ใจความชัดเจน

อุทุมพร ทองอุไทย (2523) ได้เสนอแนวทางการใช้เครื่องมือสำหรับประเมินการสอนไว้ในลักษณะต่าง ๆ ดังนี้ คือ

1. เครื่องมือที่ประเมินโดยใช้เทคโนโลยีช่วย เช่น วิดีโอเทป เครื่องบันทึกเสียง ภาพยนตร์ ซึ่งเป็นวิธีที่ดีมากเพราะผู้ถูกประเมินจะสังเกตพฤติกรรมของตนได้โดยตรง แต่เสียค่าใช้จ่ายสูงมากจึงมักจะไม่ใช่ที่แพร่หลาย

2. เครื่องมือประเมินที่เป็นมาตรฐาน ได้แก่ เครื่องมือที่ได้รับการสร้างและทดลองใช้จนเป็นที่เชื่อถือแล้ว เช่น แบบวัดต่าง ๆ

3. เครื่องมือประเมินที่สร้างเอง โดยผู้ใช้ต้องมีความรู้ในเรื่องประเภทต่าง ๆ ของเครื่องมือ วิธีสร้างและวิธีใช้ สามารถอธิบายได้ดังนี้

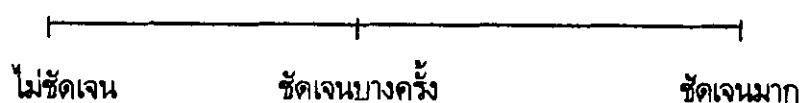
- แบบสอบถาม (Questionnaire) เป็นเครื่องมือที่รวบรวมความจริงที่ยังไม่ทราบเพื่อนำไปเปรียบเทียบกับความจริงที่ทราบแล้วซึ่งแบ่งได้ 2 ประเภท คือ แบบสอบถามปลายเปิดและแบบสอบถามปลายปิด

- แบบตรวจสอบรายการ (Checklists) เป็นแบบที่ใช้กับพฤติกรรมที่ต้องการประเมินแต่ไม่สามารถประมาณค่า (rate) หรือจัดเรียงลำดับได้ตรวจสอบเพียงแต่ต้องการจะรู้ว่าพฤติกรรมนั้นเกิดขึ้นหรือไม่ เป็นต้น

- แบบประมาณค่า (Rating scale) เป็นเครื่องมือที่ช่วยสังเกตเฉพาะพฤติกรรมที่ปรากฏหรือไม่ปรากฏก็ได้ แต่ผู้สังเกตต้องเปรียบเทียบความมากน้อยหรือความแตกต่างของพฤติกรรมให้ได้แบบประมาณค่าที่ใช้มี 3 ประเภท คือ

1. แบบตัวเลข ประกอบด้วยรายการพฤติกรรมและมีตัวเลขกำกับ เช่น 1, 2, 3, 4, 5 เพื่อให้ผู้ประมาณค่าทำเครื่องหมาย

2. แบบกราฟิก มีลักษณะคล้ายแบบแรกแต่จะใช้ภาษาแทนตัวเลขในสเกล และใช้เส้นแทนช่องโดยมีข้อความกำกับข้างท้าย เช่น



3. แบบกราฟิกเชิงบรรยาย เป็นแบบที่รวมเอาแบบที่ 1 และ 2 เข้าด้วยกัน คือมีทั้งตัวเลขและข้อความกำกับข้างท้าย

- การสัมภาษณ์ (Interview) เป็นวิธีการโต้ตอบระหว่างคน 2 ฝ่าย ในที่นี้คือ นักเรียนกับครู ผู้สอน ครูผู้สอนสามารถทำได้ง่ายและสะดวกที่สุดในรูปแบบของการพูดคุยซักถามจากตัวผู้เรียนเองโดยสามารถทำเป็นรายบุคคลหรือเป็นกลุ่ม การสัมภาษณ์มีหลายประเภท ได้แก่ การสัมภาษณ์อย่างไม่เป็นทางการและการสัมภาษณ์อย่างเป็นทางการ

- การให้ลำดับคุณค่า (Ranking) เป็นการจัดลำดับเพื่อเปรียบเทียบลักษณะบางอย่างตามเกณฑ์ที่กำหนด เช่น ให้เปรียบเทียบการสอนของอาจารย์ 5 คน ที่ทำการสอนอยู่โดยจัดเรียงลำดับจากคนที่หนึ่ง คือ คนที่ตนคิดว่าสอนดีที่สุดจนถึงสุดท้ายคือคนที่ตนคิดว่าสอนไม่ดีที่สุดในจำนวนทั้งหมด 5 คน

- การใช้แบบสอบกับผู้เรียน (Test) เป็นเครื่องมืออย่างหนึ่งที่สะท้อนให้เห็นประสิทธิผลในการเรียนของนิสิตและการสอนของครู ผลจากการสอบจะเป็นเครื่องชี้ที่ชัดเจนหนึ่งว่าการเรียนของนิสิตอยู่ระดับไหน ซึ่งแบบสอบแบ่งออกเป็น 2 ประเภท คือ แบบสอบอิงเกณฑ์และแบบสอบอิงกลุ่ม

- การใช้สเกลซีแมนติกดิฟเฟอเรนเชียล (Semantic differential scale) เป็นเครื่องมือที่ใช้ประเมินอาจารย์ทั้งทางปริมาณและคุณภาพได้อีกแบบหนึ่ง ลักษณะของเครื่องมือ คือ ให้ผู้ประเมินทำการตัดสินข้อความ เช่น ลักษณะอาจารย์ พฤติกรรมอาจารย์ ลงในสเกล 7 ช่วง ที่ปลายข้างหนึ่งแสดงลักษณะทางบวกและอีกปลายหนึ่งทางลบ เช่น

บทเรียนวันนี้

เหมาะสม.....ไม่เหมาะสม

ชัดเจน.....ไม่ชัดเจน

ให้ผู้ประเมินทำเครื่องหมายลงบนเส้นใดเส้นหนึ่งที่กำหนดให้ ซึ่งสะท้อนให้เห็นทิศทางและความมากน้อยของการตัดสินของผู้ประเมิน ทำการวิเคราะห์โดยให้ค่า 1, 2, 3,...,7 ลงในตำแหน่งในแต่ละเรื่อง

- การใช้สเกลของกัตต์แมน (Guttman scale) สเกลของกัทต์แมนมีหลักการ คือ ใช้ประเมินพฤติกรรมที่จัดเรียงลำดับกันได้ ถ้าสามารถจัดเรียงลำดับจากต่ำขึ้นไปหาสูงได้ก็ใช้สเกลของกัตต์แมนได้ดังตัวอย่าง

ง)ครูสามารถทำให้นักเรียนบรรลุผลด้วยวิธีสอน ก.

ค)ครูสามารถใช้วิธีสอน ก.

ข)ครูสามารถประยุกต์วิธีสอน ก.

ก)ครูรู้วิธีสอน ก.

ถ้าพฤติกรรมของครูสามารถจัดเรียงลำดับได้ การตอบก็น่าจะเรียงจาก ก→ข→ค→ง ซึ่งสเกลของกัตต์แมนสามารถใช้ประเมินพฤติกรรมของครูคนใดคนหนึ่งได้โดยไม่ต้องนำไปเปรียบเทียบกับคนอื่น

- การตัดสินใจของผู้เชี่ยวชาญ เป็นเครื่องมือที่เคยใช้ในสมัยก่อนที่จะมีเครื่องมือแบบใหม่ ๆ เกิดขึ้นมักจะใช้กับแบบประมาณค่าสิ่งที่เป็นนามธรรมเชิงคุณภาพโดยอาศัยผู้เชี่ยวชาญเป็นผู้ประเมิน
- การสังเกตอย่างมีระบบ มักจะประกอบด้วยแบบฟอร์มที่มีรายการที่ต้องสังเกตปรากฏอยู่ ซึ่งมักจะได้แก่ พฤติกรรมของครูในห้องเรียนและทำการสังเกตโดยผู้ที่ได้รับการฝึกหัดอย่างดีให้ทำการสังเกตพฤติกรรมที่ปรากฏในช่วงเวลาที่กำหนดแล้วทำการบันทึกความถี่ของพฤติกรรมที่ปรากฏเหล่านั้น

Lewis (1973) ได้ให้แนวคิดเกี่ยวกับเครื่องมือที่ใช้ในการประเมินผลการปฏิบัติงานของครูไว้ดังนี้คือ

1. การตรวจสอบรายการ (Checklists) ซึ่งผู้ประเมินจะเป็นผู้ตรวจสอบข้อมูลในลักษณะที่เป็นรูปธรรม ในรายการต่าง ๆ ที่เกี่ยวกับการปฏิบัติงานของครู
2. แบบสอบถามปลายเปิด (Open-ended questionnaire) ซึ่งผู้ประเมินจะเป็นผู้กรอกข้อมูลในลักษณะเชิงบรรยายที่เกี่ยวกับการปฏิบัติงานของครู
3. แบบมาตราประเมินค่า (Rating scales) โดยผู้ประเมินเป็นผู้ให้น้ำหนักคะแนนตามมาตรวัดที่มีอยู่ซึ่งเป็นการประเมินผลการปฏิบัติงานในด้านต่าง ๆ
4. การสังเกต (Observation) เป็นการสังเกตดูพฤติกรรมต่าง ๆ ที่ครูปฏิบัติไม่ว่าจะอยู่ในชั้นเรียนหรือภายนอกชั้นเรียน

Cyril และPoster (1993) ได้ให้แนวคิดเกี่ยวกับเครื่องมือที่ใช้ในการประเมินผลการปฏิบัติงานของครู ไว้ดังนี้คือ

1. การสังเกตการณ์ในห้องเรียน (Classroom observation) ซึ่งถือว่าเป็นส่วนสำคัญมาก เพราะครูใช้เวลาในด้านการเรียนการสอนในห้องเรียนเป็นส่วนใหญ่
2. การใช้แบบสอบถาม (Questionnaire) ในลักษณะต่าง ๆ ที่ถามเกี่ยวกับการปฏิบัติหน้าที่ภารกิจต่าง ๆ ที่ได้รับมอบหมาย รวมทั้งทำที่บุคลิกลักษณะและทัศนคติต่าง ๆ

Bollington และคณะ (1993) ได้ให้แนวคิดเกี่ยวกับเครื่องมือที่ใช้ในการประเมินผลการปฏิบัติงานของครู ไว้ดังนี้คือ

1. แบบสอบถาม ซึ่งเป็นมาตราประมาณค่า (Rating scales) เพื่อสอบถามถึงพฤติกรรมของครูที่แสดงออกในด้านต่าง ๆ ที่เกี่ยวข้องกับภารกิจหน้าที่ที่ได้รับมอบหมาย
2. การสังเกตการสอน (Observation) เป็นการสังเกตถึงพฤติกรรมในการสอนของครูในห้องเรียน
3. การสัมภาษณ์ เป็นการสัมภาษณ์ผู้รับการประเมินผลในด้านภารกิจหน้าที่ รวมทั้งทำที่สนคดีต่อวิชาชีพครู
4. การบันทึกภาพและเสียงจากวิดีโอ เพื่อทราบถึงพฤติกรรมต่าง ๆ ของครูในห้องเรียนว่า ได้แสดงพฤติกรรมอะไรบ้างตลอดเวลาการเรียนการสอนในแต่ละชั่วโมง

จากแนวคิดและหลักการในการประเมินการสอน ทั้งหมดที่กล่าวมานี้พอสรุปได้ว่า การประเมินการสอนมีวัตถุประสงค์หลักเพื่อปรับปรุงการสอน โดยผู้ประเมินอาจเป็นนักศึกษา เพื่อนอาจารย์ และตัวอาจารย์เองด้วยการให้ประเมิน อาทิ นุคลิกลักษณะ ความสามารถทางวิชาการ เจตคติต่อนักศึกษา เทคนิคการสอน การวัดและประเมินผลการเรียน เป็นต้น สำหรับวิธีการและเครื่องมือที่ใช้ในการประเมินการสอน สามารถใช้การสังเกต การสัมภาษณ์ การใช้แบบสอบถาม การใช้มาตราประมาณค่า ตลอดจนใช้การบันทึกภาพหรืออาจใช้หลาย ๆ วิธีประกอบกัน

ตอนที่ 2 แนวความคิดเกี่ยวกับมาตราประมาณค่า

มาตราประมาณค่า (Rating scale) เป็นเครื่องมือวัดผลทางการศึกษาที่สำคัญอย่างหนึ่ง ใช้วัดคุณลักษณะภายในของบุคคล มาตราประมาณค่าประกอบด้วยข้อคำถามหรือสิ่งเร้ากับตัวเลือกที่ให้ผู้ตอบตอบสนองเพื่อที่จะให้ประมาณค่าคุณลักษณะสิ่งใดสิ่งหนึ่งของบุคคลหรือสิ่งของ (Wiersma and Jurs, 1990)

ในอดีตที่ผ่านมา การที่จะตัดสินคุณค่าของสิ่งใด ๆ ผู้ตัดสินสามารถลงข้อสรุปหรือตัดสินด้วยการบรรยายคุณค่าของสิ่งนั้นๆออกมาในรูปของความเรียงซึ่งให้อิสระกับผู้ตัดสินที่จะเขียนบรรยายออกมาในรูปแบบใดก็ได้ มีผลทำให้ข้อสรุปขาดแบบแผนที่ชัดเจนแน่นอน ด้วยเหตุผลดังกล่าว นักวัดผลได้พยายามค้นหาวิธีการในการตัดสินคุณค่าของสิ่งต่าง ๆ ให้เป็นรูปแบบที่มีความเป็นปรนัยมากยิ่งขึ้น

การใช้มาตราประมาณค่า (Rating scale) เป็นแนวคิดหนึ่งที่ได้รับการพัฒนาขึ้นเพื่อแก้ไขปัญหาดังกล่าว โดยเน้นให้ผู้ประเมินตัดสินคุณค่าของสิ่งที่ต้องการลงบนสเกล หรือระดับคุณภาพที่ได้กำหนดไว้ล่วงหน้า (สุพัทธ์ร์ พินุลย์, 2534 : 14)

มาตราประมาณค่าโดยทั่วไปจะประกอบด้วยลักษณะทั่ว ๆ ไป 2 องค์ประกอบด้วยกัน (Thorndike and Hagen, 1977: 462) คือ

1. คุณลักษณะของสิ่งที่ประเมิน
2. รูปแบบคำตอบที่บ่งบอกถึงระดับคุณภาพของสิ่งที่ประเมิน

ซึ่งมาตราประมาณค่าโดยทั่วไปจะนิยมใช้เพื่อประเมินพฤติกรรม (Behavior rating) หรือ ประเมินผลผลิต (Product rating) (Popham, 1981: 315-317)

สำหรับในทางปฏิบัตินั้น เราสามารถประยุกต์มาตราประมาณค่าเพื่อใช้ในกิจกรรมทางการศึกษา และเพื่อประเมินผลการเรียนรู้และพัฒนาการต่าง ๆ ซึ่งสามารถจำแนกได้เป็น 3 กลุ่มใหญ่ ๆ ดังนี้

1. ใช้ประเมินกระบวนการ (Process evaluation)
เช่น ประเมินทักษะการพูด ทักษะการบรรยาย ทักษะการทำงานเป็นกลุ่ม ทักษะการเล่น กีฬา เป็นต้น
2. ใช้ประเมินผลผลิต (Product evaluation)
เช่น ประเมินผลงานต่าง ๆ ของนักเรียน เป็นต้น
3. ใช้ประเมินพัฒนาการทางบุคลิกภาพและสังคม (Evaluating personal social development)
เช่น ประเมินความสนใจ ทักษะคิด ความรับผิดชอบ เป็นต้น

จากที่กล่าวมาจะเห็นได้ว่า มาตราประมาณค่าเป็นเครื่องมือที่ได้รับการพัฒนาขึ้นเพื่อมุ่งหวังให้เกิดความเป็นปรนัยในการแปลความหมายข้อมูลโดยอาศัยการกำหนดสเกลของลักษณะที่ต้องการ ประเมินมาให้ผู้ตอบตอบในรูปของทางเลือกต่าง ๆ ที่เรียงเอาไว้เป็นลำดับขั้น ซึ่งสามารถนำมาประยุกต์ใช้ในกิจกรรมทางการประเมินต่าง ๆ ได้อย่างหลากหลาย

รูปแบบของมาตรประมาณค่า

เนื่องจากเหตุผลที่มาตรประมาณค่าเป็นเครื่องมือที่สร้างได้ไม่ยากนัก ใช้งบประมาณในการสร้างน้อย สะดวกต่อการนำไปใช้ ตลอดจนสามารถนำไปประยุกต์ใช้ในสถานการณ์ของการประเมินต่างๆ ได้อย่างกว้างขวาง จึงได้มีการพัฒนารูปแบบของมาตรประมาณค่าขึ้นอย่างหลากหลายรูปแบบ ดังรายชื่อที่นักจิตวิทยาและนักวัดผลที่ได้แสดงไว้ในตารางที่ 1 ต่อไปนี้

ตารางที่ 1 รูปแบบและลักษณะของมาตรประมาณค่าตามรายชื่อผู้จำแนก

ชื่อผู้จำแนก	รูปแบบของมาตรประมาณค่า	ลักษณะมาตรประมาณค่า
Guilford (1954)	1. แบบตัวเลข (Numerical scales)	<ul style="list-style-type: none"> กำหนดตัวเลขพร้อมกัมนิยามหรือคำบรรยายไว้ให้ผู้ประเมิน
	2. แบบกราฟิก (Graphic scales)	<ul style="list-style-type: none"> มีข้อความและมีเส้นให้ประเมิน
	3. แบบสเกลมาตรฐาน (Standard scales)	<ul style="list-style-type: none"> มีชุดของคุณลักษณะมาตรฐานให้เปรียบเทียบ
	4. แบบแต้มสะสม (Cumulated points scales)	<ul style="list-style-type: none"> กำหนดตัวเลือกหลาย ๆ ตัว ให้ผู้ประเมินเลือกตัวเลือกที่มีความเหมาะสมกับคุณลักษณะที่ต้องการวัด
	5. แบบตัวเลือกบังคับตอบ (Forced -choice rating)	<ul style="list-style-type: none"> กำหนดตัวเลือกและค่าคะแนนให้กับตัวเลือก
Turney และ Robb (1971) Rummel (1964)	1. แบบตัวเลข (Numerical scales)	<ul style="list-style-type: none"> ใช้ตัวเลขซึ่งบอกระดับของลักษณะหรือพฤติกรรมที่ต้องการประเมิน
	2. แบบกราฟิก (Graphic scales)	<ul style="list-style-type: none"> ให้ผู้ตอบทำเครื่องหมายลงบนเส้นตรงเพื่อแสดงระดับลักษณะที่ประเมิน
	3. แบบบรรยาย (Descriptive scales)	<ul style="list-style-type: none"> มีคำบรรยายหรือข้อความแสดงระดับของลักษณะที่ต้องการให้ประเมิน

ตารางที่ 1 (ต่อ)

ชื่อผู้จำแนก	รูปแบบของมาตรประมาณค่า	ลักษณะมาตรประมาณค่า
Gronlund (1981)	<ol style="list-style-type: none"> แบบตัวเลข (Numerical rating scales) แบบกราฟิก (Graphic rating scales) แบบบรรยายแบบกราฟิก (Descriptive Graphic rating - scales) 	<ul style="list-style-type: none"> ● ให้ผู้ประเมินทำเครื่องหมายตรงตัวเลขที่เป็นดัชนีชี้ถึงระดับของคุณลักษณะที่ปรากฏ ● มีคุณลักษณะหรือสิ่งที่จะประเมินปรากฏอยู่บนเส้นในแนวนอนโดยผู้ประเมินทำเครื่องหมายบนเส้นตรงนั้น ● มีวลีบรรยายแต่ละจุดบนสเกล อาทิ <ul style="list-style-type: none"> - มีคำบรรยายที่จุดปลายสเกลทั้งสองด้านและตรงกลาง - มีคำบรรยายทุกจุดบนสเกล
Popham (1981)	<ol style="list-style-type: none"> แบบตัวเลข (Numerical rating scales) แบบกราฟิก (Graphic rating scales) แบบจัดอันดับ (Ranking) 	<ul style="list-style-type: none"> ● มีค่าตัวเลขที่ต่อเนื่องสำหรับผู้ประเมินเป็นผู้ตัดสิน โดยกำหนดความหมายให้กับค่าของตัวเลข ● มีการกำหนดคุณลักษณะที่จะประเมินควบคู่ไปกับเส้นตรงในแนวนอน สำหรับให้ผู้ประเมินทำเครื่องหมาย ● มีการให้เปรียบเทียบคุณลักษณะที่จะประเมินเป็นรายคู่
Richard (1989)	<ol style="list-style-type: none"> Category ratings แบบตัวเลข (Numerical rating scales) แบบกราฟิก (Graphic rating scales) แบบบังคับตอบ (Force-choice rating) 	<ul style="list-style-type: none"> ● ให้ผู้ตอบเลือกคำตอบที่ดีที่สุดเพื่อตอบคำถามหรือบรรยายคำถามที่ถาม ● มีลักษณะกำหนดค่าตัวเลขที่ต่อเนื่องให้กับคำที่อธิบายคำตอบให้ผู้ประเมินเลือกตัวเลขเป็นคำตอบ ● ใช้เส้นตรงเป็นตัวแทนค่าที่ต่อเนื่องโดยมีช่วงที่เท่าๆ กัน ● ให้ผู้ตอบเลือกคำตอบจากชุดของคำตอบ

ความคลาดเคลื่อนที่เกิดจากการใช้มาตรฐานค่า

ความคลาดเคลื่อนที่เกิดขึ้นจากการใช้มาตรฐานค่า เป็นผลมาจากข้อบกพร่องในการสร้างเครื่องมือ เป็นต้นว่า ลักษณะที่ต้องการให้ประมาณค่าเป็นลักษณะที่อยู่ภายใน สังเกตได้ยากและอาจแสดงออกไม่เหมือนกัน เช่น ความตั้งใจ เครียด บางครั้งความหมายและทิศทางของลักษณะไม่เด่นชัด ไม่มีขอบเขตที่แน่นอน ทำให้ผู้ประเมินตีความหมายไปในแนวทางที่แตกต่างกันและที่สำคัญคือ ผู้ประเมินขาดเกณฑ์ในการประมาณค่าที่เป็นอันหนึ่งอันเดียวกัน เช่น " ดี" ของคนหนึ่งอาจเป็นแค่ระดับ "พอใช้" ของอีกคนหนึ่งก็ได้ การที่ผู้ประเมินมีขอบเขตและเกณฑ์ในการประมาณค่าที่แตกต่างกันนี้มีผลทำให้ค่าความเที่ยงในระหว่างผู้ประเมินลดลง (Sellitz และคณะ, 1959: 352)

จากการศึกษาของนักจิตวิทยาโดยเฉพาะกิลฟอร์ด (Guilford, 1954) ได้พบว่า การใช้มาตรฐานค่าเป็นเครื่องมือในการเก็บรวบรวมข้อมูลมักทำให้ข้อมูลที่ได้มีความคลาดเคลื่อนสูงโดยมีแหล่งความคลาดเคลื่อนที่สำคัญอยู่ 3 แหล่ง คือ

1. ความบกพร่องของเครื่องมือ ข้อบกพร่องที่สำคัญของเครื่องมือที่เป็นมาตรฐานประมาณค่าคือ ขาดการนิยามหรือบรรยายคุณลักษณะของสิ่งที่จะประเมินไว้อย่างชัดเจน ทำให้เกิดความเข้าใจที่ไม่ตรงกัน จึงเป็นผลทำให้การประมาณค่านั้นขาดความเที่ยง หรืออาจเป็นเพราะภาษาที่ใช้ไม่เหมาะสม หรือไม่สามารวัดสิ่งที่ต้องการวัดได้ เป็นต้น

2. ความบกพร่องที่เกิดจากกระบวนการในการประเมิน ข้อบกพร่องที่มักจะพบคือ คุณลักษณะหรือสิ่งที่ต้องการให้ประเมินนั้นมีจำนวนมาก จึงทำให้ผลการประเมินเกิดความคลาดเคลื่อนได้และข้อบกพร่องที่สำคัญอีกประการหนึ่งคือกรณีที่มีผู้ประเมิน(Rater)เพียงคนเดียวทำการประเมินผู้ถูกประเมิน(Ratee) หลาย ๆ คน อาจทำให้เกิดความคลาดเคลื่อนขึ้นได้ แต่ถ้ามีผู้ประเมินหลาย ๆ คน จะเป็นวิธีที่ช่วยลดความคลาดเคลื่อนที่เกิดจากแต่ละบุคคลได้ ทำให้ผลการประเมินมีความแม่นยำยิ่งขึ้น

3. ความคลาดเคลื่อนที่เกิดจากตัวผู้ประเมินซึ่งมีความคลาดเคลื่อนหลายประการที่พึงควรระมัดระวัง ได้แก่ ความลำเอียงหรือความมีอคติของแต่ละบุคคล (Personal bias) ซึ่งเป็นผลมาจากการมีปฏิสัมพันธ์มาก่อน (Halo effect) ความคลาดเคลื่อนแบบใช้ตรรกผิด (Logical error) และความคลาดเคลื่อนเชิงปล่อยหรืออคติคณน (Error of leniency) ดังรายละเอียดต่อไปนี้

- 3.1 ความลำเอียงหรือความมีอคติส่วนบุคคลของผู้ประเมิน (Personal bias) เป็นความคลาดเคลื่อนที่เกิดขึ้นเฉพาะลักษณะนิสัยประจำตัวของผู้ประเมิน(Rater)ที่มักมีรูปแบบการประเมินที่ตายตัว (Fixed pattern) ซึ่งมี 3 ชนิด คือ

3.1.1 ความคลาดเคลื่อนแบบประเมินค่าสูงกว่าความเป็นจริง (Generosity error) เป็นความคลาดเคลื่อนที่เกิดขึ้นเมื่อผู้ประเมินประเมินค่าคุณลักษณะต่าง ๆ ของทุกคนในระดับที่ค่อนข้างสูงเพียงด้านเดียว กล่าวคือ เมื่อคุณลักษณะของผู้ถูกประเมินมีค่าอยู่ในระดับปานกลางหรือต่ำ แต่เมื่อมีความคลาดเคลื่อนชนิดนี้เกิดขึ้นผู้นั้นจะได้รับการประเมินให้อยู่ในระดับสูงหรือค่อนข้างสูง เป็นผลทำให้ค่าที่ประเมินสูงกว่าความเป็นจริง

3.1.2 ความคลาดเคลื่อนแบบประเมินค่าต่ำกว่าความเป็นจริง (Severity error) เป็นความคลาดเคลื่อนที่เกิดขึ้นเมื่อผู้ประเมินประเมินค่าคุณลักษณะต่าง ๆ ของทุก ๆ คน ในระดับต่ำหรือค่อนข้างต่ำเพียงด้านเดียว เป็นผลทำให้ค่าที่ประเมินได้ต่ำกว่าที่ควรจะเป็น กล่าวคือ เมื่อคุณลักษณะของผู้ถูกประเมินมีค่าอยู่ในระดับดีหรือปานกลาง เมื่อมีความคลาดเคลื่อนชนิดนี้เกิดขึ้น ผู้ถูกประเมินผู้นั้นจะได้รับการประเมินให้อยู่ในระดับต่ำหรือค่อนข้างต่ำ

3.1.3 ความคลาดเคลื่อนแบบโน้มเข้าสู่ตรงกลาง (Central tendency error) ความคลาดเคลื่อนชนิดนี้เกิดขึ้นเมื่อผู้ประเมินพยายามหลีกเลี่ยงการประเมินค่าทางระดับค่อนข้างสูงและระดับค่อนข้างต่ำมาเป็นระดับปานกลาง ทำให้ค่าที่ประเมินได้ต่ำกว่าหรือสูงกว่าสภาพที่เป็นจริงซึ่งสาเหตุของความคลาดเคลื่อนชนิดนี้อาจเป็นเพราะ

1. ผู้ประเมินอาจจะไม่ทราบว่สิ่งที่เขาประเมินนั้นมีพิสัยอยู่ในทางสูงหรือทางต่ำ
2. ผู้ประเมินอาจจะไม่มีความเคยชินในการจำแนกความแตกต่างของสิ่งที่ประเมิน
3. บุคลิกภาพของผู้ประเมินเป็นคนไม่มั่นคง และไม่กล้าที่จะประเมินค่าทางสูงหรือทางต่ำทางใดทางหนึ่ง
4. ผู้ประเมินอาจรู้สึกว่าจะตัดสินได้อย่างเป็นธรรมและอาจเกรงว่าจะเป็นการทำร้ายจิตใจคน

ดังจะเห็นได้ว่าการที่ผู้ประเมิน ประเมินค่าเพียงระดับใดระดับหนึ่งเพียงระดับเดียว จะทำให้เกิดผลที่ไม่น่าพึงปรารถนา 2 ประการ คือ

ประการแรก ทำให้ค่าที่ประเมินได้เป็นค่าที่ตรงสภาพเพียงใด ทั้งนี้เพราะการที่ผู้ประเมินมักประเมินเฉพาะทางสูงหรือทางต่ำทางเดียวน่าจะสะท้อนให้เห็นถึงการมองของแต่ละบุคคล (Personal outlook) มากกว่าที่จะเป็นคุณลักษณะของผู้ถูกประเมิน

ประการที่สอง การที่ผู้ประเมินมักจะประเมินเฉพาะระดับใดระดับหนึ่งเป็นการจำกัดขอบเขตในการประเมิน ทำให้อ่านาจจำแนกต่ำลง ความเที่ยงย่อมต่ำลงด้วย

3.2 ผลจากการมีปฏิสัมพันธ์มาก่อน (Halo effect) เป็นความคลาดเคลื่อนที่เกิดขึ้นเมื่อความรู้สึกของผู้ประเมินที่มีต่อผู้ถูกประเมินมีอิทธิพลต่อการประเมินค่าคุณลักษณะของผู้ถูกประเมิน เช่น ถ้าผู้ประเมินมีทัศนคติในทางที่ดีต่อผู้ถูกประเมินแล้ว เขาก็มักจะประเมินคุณลักษณะต่าง ๆ ของผู้ถูกประเมินในทางสูง แต่ในทางตรงกันข้ามถ้าผู้ประเมินมีทัศนคติในทางลบต่อผู้ถูกประเมินแล้ว เขาก็มักจะประเมินคุณลักษณะต่าง ๆ ของผู้ถูกประเมินในทางต่ำ ความคลาดเคลื่อนแบบนี้จะเกิดขึ้นกับเฉพาะบางคนที่ถูกประเมิน ซึ่งแตกต่างจากความคลาดเคลื่อนแบบ Generosity error และ Severity error ที่ผู้ประเมินมักจะประเมินทุกคนในทางสูงเหมือน ๆ กัน และในทางต่ำเหมือน ๆ กัน

3.3 ความคลาดเคลื่อนแบบใช้ตรรกผิด (Logical error) ความคลาดเคลื่อนชนิดนี้เกิดขึ้นเมื่อผู้ประเมินนำตัวแปรอื่นที่คิดว่ามีความสัมพันธ์กับตัวแปรที่กำลังประเมินมามีส่วนในการประเมินด้วย เช่น การประเมินความฉลาด ความมักจะประเมินค่าความฉลาดของนักเรียนจากผลสัมฤทธิ์ทางการเรียนของนักเรียน กล่าวคือ ถ้านักเรียนมีผลสัมฤทธิ์ทางการเรียนในระดับสูง เขาก็มักจะประเมินความฉลาดของนักเรียนผู้นั้นในระดับสูง ทั้งนี้เพราะเขามีความคาดหวังทางตรรกว่า คุณลักษณะ 2 สิ่งนั้นเป็นสิ่งที่เกิดควบคู่กันหรือสัมพันธ์กัน หรือบางคนก็เชื่อว่า นักเรียนที่มีพรสวรรค์จะมีการปรับตัวทางสังคมต่ำจึงมักประเมินคุณลักษณะทางสังคมของนักเรียนเหล่านั้นในระดับต่ำ ซึ่งเป็นความเชื่อที่ผิด ความคลาดเคลื่อนประเภทนี้มิได้เป็นผลมาจากความลำเอียงหรือความมีอคติต่อนักเรียนคนใดคนหนึ่ง หรือสถานการณ์ใดสถานการณ์หนึ่ง หากแต่เป็นเพราะผู้ประเมินมีการรับรู้มาก่อนเกี่ยวกับธรรมชาติของคน จึงคิดว่าคุณลักษณะนั้น ๆ มีความสัมพันธ์โดยตรง ซึ่งต่างจากที่เป็นจริง ดังนั้นจึงทำให้เกิดความคลาดเคลื่อนขึ้น

3.4 ความคลาดเคลื่อนเชิงปล่อยหรือกดคะแนน (Error of leniency) ความคลาดเคลื่อนแบบนี้คล้าย Generosity error และ Severity error ต่างกันที่ผู้ประเมินพยายามประเมินผลงานในด้านดีเกินไป (Easy rate) หรือในทางตรงกันข้าม ผู้ประเมินจะประเมินผลงานให้ต่ำกว่าความเป็นจริง (Hard rate)

ความคลาดเคลื่อนดังที่ได้กล่าวมานี้เป็นความคลาดเคลื่อนอย่างเป็นระบบ (Systematic error) ที่ทำให้ความตรงลดลง (Selltiz และคณะ, 1959: 352)

นอกจากนี้ Thorndike และ Hagen (1955) และ Wolf (1988) ได้กล่าวเพิ่มเติมว่า ในการใช้มาตรฐานประเมินค่าเป็นเครื่องมือในการประเมินตนเอง ประเมินผู้อื่น หรือประเมินกระบวนการก็ตาม ข้อมูลที่ได้มักขาดความน่าเชื่อถือ ทั้งนี้มีสาเหตุมาจาก

1. ความไม่เต็มใจและไม่พอใจที่จะตัดสินใจอย่างละเอียดรอบคอบของผู้ให้ข้อมูล เนื่องจากเหตุผลสำคัญคือ ไม่ต้องการให้ตนเองเกิดความลำบากใจในการตัดสินใจ และไม่ต้องการทำร้ายความรู้สึกของผู้ถูกประเมิน

2.ความสามารถในการตอบให้ตรงกับความเป็นจริงของผู้ตอบหรือผู้ให้ข้อมูล เป็นสิ่งที่ทำได้ยาก เนื่องจากเหตุผลคือ โอกาสในการสังเกตหรือรู้จักสิ่งที่จะประเมินมีน้อย สิ่งที่จะประเมินหรือตอบเป็นเรื่องที่ไม่อาจตอบตามความเป็นจริงได้ เช่น เรื่องที่ไวต่อความรู้สึกต่าง ๆ ไม่เข้าใจคุณลักษณะที่จะประเมินว่าคืออะไร ไม่สามารถจำแนกระดับความเข้มของพฤติกรรมหรือคำตอบได้ และเกิดจากบุคลิกภาพส่วนตัวของผู้ตอบหรือผู้ประเมินเอง

ดังนั้นจากเหตุผลที่กล่าวมา นักจิตวิทยาและนักการศึกษาจึงได้เสนอแนะแนวทางในการปรับปรุงแก้ไขการใช้มาตราประมาณค่าให้ได้ผลถูกต้องแม่นยำมากยิ่งขึ้น ดังนี้

แนวทางการปรับปรุงคุณภาพของมาตราประมาณค่า

Guilford (1954) กล่าวว่า การให้ความรู้และฝึกฝนผู้ประเมินอย่างดีเป็นวิธีการที่มีประสิทธิภาพมากที่สุดในการปรับปรุงแก้ไขการประมาณค่า โดยเฉพาะในการลดความคลาดเคลื่อนอย่างมีระบบ เพราะผู้ประเมินที่ได้รับการฝึกฝนอย่างดีจะสามารถหลีกเลี่ยงไม่ให้เกิดความคลาดเคลื่อนได้

Gronbach (1960) กล่าวว่า ในกรณีที่สามารถรวบรวมความคิดเห็นจากผู้ประเมินหลาย ๆ คน ถ้าผู้ประเมินเหล่านั้นรู้จักผู้ถูกประเมินดีพอก็จะช่วยให้ได้ค่าความเที่ยงในการประมาณค่าสูงขึ้นเนื่องจากอคติที่เกิดขึ้นเป็นค่าเฉลี่ยนอกจากนั้นยังเป็นการเพิ่มเติมข้อมูลจากผู้ประเมินบางคนไม่มีโอกาสในการสังเกตเห็นอีกด้วย

Oppenhiem (1966) ได้เสนอแนะว่า ในการสร้างหรือพัฒนามาตราประมาณค่าจะต้องกำหนดขอบเขต และทิศทางของลักษณะที่จะประเมินให้เป็นที่เข้าใจตรงกันเสียก่อน จากนั้นจึงค่อยกำหนดระดับในการประมาณค่าและกำหนดความหมายในแต่ละระดับและเพื่อให้ลักษณะที่จะประเมินมีความชัดเจน ลักษณะนั้น ๆ จะต้องอธิบายไว้อย่างเฉพาะเจาะจง แต่ละลักษณะต้องหมายถึง กิจกรรมเพียงอย่างเดียว และเป็นลักษณะที่คงที่ การตัดสินแต่ละลักษณะต้องมีพื้นฐานมาจากอดีต หรือปัจจุบันมากกว่า เป็นการมองอนาคต ถ้าเป็นการประเมินตนเอง ไม่ควรมีลักษณะที่คนส่วนมากจะต้องตัดสินสูงหรือต่ำกว่าความเป็นจริงเสมอ

Tuckman (1975) ได้เสนอแนะวิธีการในการสร้างมาตรประมาณค่าให้มีคุณภาพ ดังนี้

1. เลือกคุณลักษณะหรือพฤติกรรมที่จะประเมินค่า ซึ่งควรมีลักษณะดังต่อไปนี้
 - ต้องมีความสำคัญและตรงกับสิ่งที่ต้องการวัดมากที่สุด
 - ต้องเป็นอิสระไม่เกี่ยวข้องกับคุณลักษณะอื่น ๆ
 - ต้องเป็นสิ่งที่สามารถสังเกตได้โดยตรง
 - แต่ละองค์ประกอบของพฤติกรรมมีความเป็นเอกภาพ
 - คุณลักษณะหรือพฤติกรรมต่าง ๆ ไม่ควรมีมากเกินไป ปกติควรมี 5-6 รายการ
2. นิยามหรือบรรยายคุณลักษณะหรือพฤติกรรมที่จะประเมินค่าให้ชัดเจนเพื่อให้เกิดความเข้าใจที่ตรงกัน และสามารถตัดสินได้อย่างเป็นปรนัยและเที่ยง
3. ออกแบบช่วงของมาตรประมาณค่าที่จะใช้ประเมิน ดังนี้
 - ปกติมักจะแบ่งช่วงออกเป็น 3-7 ช่วง การแบ่งช่วงยิ่งมากยิ่งแสดงความแตกต่างได้มาก การแบ่งช่วงเป็นจำนวนคี่จะมีข้อดีที่ทำให้มีค่ากลาง ซึ่งในบางครั้งอาจใช้ในการตีค่าการประเมินค่า นั้นตัดสินใจไม่ได้ มาตรประมาณค่าแบบ 5 ช่วง จึงเป็นที่นิยมมาก
 - มีการนิยามหรืออธิบายความต่อเนื่องในแต่ละช่วงให้ชัดเจน
 - ควรมีการสลับช่วงจากน้อยไปหามาก และจากมากไปหาน้อยทั้งนี้เพื่อป้องกันมิให้ผู้ประเมินประเมินค่าโดยอัตโนมัติ

Thorndike และ Hagen (1977) ได้เสนอแนะแนวทางการปรับปรุงแก้ไขการประมาณค่าให้ได้ผลถูกต้องแน่นอน ไว้ดังนี้

1. การประเมินควรกระทำเฉพาะลักษณะที่ปรากฏให้เห็นชัดเจน
2. ลักษณะที่จะประเมินต้องแสดงออกมาให้เห็นในรูปพฤติกรรมและการตัดสินจะกระทำจากพฤติกรรมที่เห็นนั้น
3. รูปแบบของการประมาณค่า (Rating form) ต้องปรับปรุงเพื่อบังคับให้ผู้ประเมินบอกความแตกต่าง หรือเพื่อควบคุมการที่ผู้ประเมินมีมาตรฐานในการตัดสินที่ต่างกัน
4. ผู้ประเมินต้องเป็นผู้ที่มีโอกาสมากที่สุดในการสังเกตพฤติกรรมในสถานการณ์ตามที่ต้องการ
5. ผู้ประเมินต้องได้รับการบอกเล่าถึงคุณค่าของการประมาณค่า และได้รับการฝึกฝนในการใช้เครื่องมือ
6. ถ้าสามารถมีผู้ประเมินหลายคนก็ควรรวบรวมผลการประเมินจากผู้ประเมินหลายๆ คน

Mathis (1979) ได้เสนอแนะว่า ในการที่จะลดอคติของผู้ประเมิน (Rater bias) หรือลดความคลาดเคลื่อนอย่างเป็นระบบในการประเมินนั้น นอกเหนือไปจากการให้ความรู้ความเข้าใจในเรื่องความคลาดเคลื่อนลักษณะต่าง ๆ แก่ผู้ประเมินแล้ว สิ่งที่สำคัญยิ่งอีกประการหนึ่งก็คือ การนิยามหรือกำหนดความหมายของระดับการประเมินแต่ละระดับให้ชัดเจน อีกทั้งกำหนดขอบเขตในการพิจารณาในแต่ละรายการหรือแต่ละข้อให้ผู้ประเมินแต่ละคนสามารถเข้าใจได้ตรงกันก็จะสามารถแก้ปัญหาในเรื่องการเกิดความคลาดเคลื่อนในการประเมินได้ซึ่งเขาได้ชี้แนะว่ามาตรประมาณค่าแบบยึดพฤติกรรม (Behaviorally anchored rating scales) เป็นอีกรูปแบบหนึ่งของเครื่องมือที่สามารถช่วยลดความคลาดเคลื่อนอย่างเป็นระบบในการประเมินได้

ตอนที่ 3 แนวความคิดเกี่ยวกับมาตรประมาณค่าแบบลิเคอร์ท

มาตรประมาณค่าแบบลิเคอร์ท (Likert rating scale) ได้พัฒนาขึ้นในปี ค.ศ. 1932 โดย Rensis Likert ซึ่งเขาได้ให้ชื่อมาตรวัดตัวแปรทางจิตวิทยาของเขาว่า Likert's summated rating scale มีจุดมุ่งหมายเพื่อใช้วัดทัศนคติซึ่งเป็นการวัดโดยใช้ข้อความเกี่ยวกับเรื่องใดเรื่องหนึ่งสอบถามความคิดเห็นของบุคคลที่มีต่อเรื่องนั้น แล้วให้บุคคลนั้นแสดงความรู้สึกต่อข้อความดังกล่าว ซึ่งการตอบสนองต่อข้อความนั้นอาจจะเห็นด้วย ไม่เห็นด้วย หรือไม่แน่ใจกับข้อความนั้นก็

วิธีการวัดของลิเคอร์ทเกิดจากข้อจำกัดของวิธีของ Thurstone ที่ให้กลุ่มคนตัดสินข้อความ ลิเคอร์ทใช้การตอบของแต่ละคนทำให้ยกเลิกข้อตกลงเบื้องต้นของ Thurstone ที่ว่ากลุ่มผู้ตัดสินจะตัดสินข้อความเหมือนกับกลุ่มผู้ตอบจริง ข้อแตกต่างอีกประการหนึ่ง คือ แทนที่ผู้ตัดสินจะตัดสินข้อความบางข้อ ผู้ตัดสินต้องตอบทุกข้อลงในมาตรา 5 ช่อง ซึ่งเรียงลำดับจาก เห็นด้วยอย่างยิ่ง เห็นด้วย ไม่แน่ใจ ไม่เห็นด้วย และไม่เห็นด้วยอย่างยิ่ง วิธีของลิเคอร์ทเป็นการจัดลำดับวัตถุตามความรู้สึก เช่น มาตรประมาณค่า 5 ช่อง จัดแบ่งความรู้สึกเป็น 1-เห็นด้วยอย่างยิ่ง 2-เห็นด้วย 3-ยังไม่ตัดสินใจ 4-เห็นด้วย 5-ไม่เห็นด้วยอย่างยิ่ง โดยใช้หลักการของการจัดเรียงลำดับแบบสะสม ลิเคอร์ทใช้วิธีจัดช่องให้รู้สึกเท่ากัน วัดความรู้สึกพึงพอใจ (Preference) และการตัดสินใจของคนต่อวัตถุตามความรู้สึกพึงพอใจ ค่าบนมาตรได้จากการวิเคราะห์รายข้อ (อุทุมพร จามรมาน, 2537)

มาตรประมาณค่าของลิเคอร์ทนี้โดยพื้นฐานของมาตรวัดนั้นจัดอยู่ในมาตราเรียงอันดับ (Ordinal scale) แต่ในทางสถิติลิเคอร์ทได้แสดงให้เห็นประจักษ์ว่าสามารถกำหนดค่าคะแนนของข้อความโดยใช้วิธีการให้ค่าคะแนนคำตอบ (1,2,3,4,5) ตามลักษณะของข้อความได้ (ไม่เห็นด้วยอย่างยิ่ง ไม่เห็นด้วย ไม่แน่ใจ เห็นด้วย เห็นด้วยอย่างยิ่ง) ดังนั้นในการวิเคราะห์ทางสถิติจึงมีการกำหนดข้อตกลงเบื้องต้น

ให้ค่าคะแนนของข้อความในมาตราวัดของลิเคอร์ท์อยู่ในระดับอันตรภาค (Interval scale) (Edward, 1957; Berk, 1979: 653)

การสร้างมาตรประมาณค่าแบบลิเคอร์ท์

การสร้างมาตรประมาณค่าแบบลิเคอร์ท์ สามารถดำเนินการได้โดยมีขั้นตอนดังต่อไปนี้
(อุทุมพร จามรمان, 2537)

1. กำหนดวัตถุประสงค์ในการนำมาตราวัดไปใช้งานว่าต้องการวัดทัศนคติของใครที่มีต่ออะไร
2. กำหนดเนื้อหา (Content) หรือประเด็นที่ต้องการจะใช้ในมาตราวัด
3. กำหนดลักษณะของมาตราวัดว่าจะเป็นแบบใด เช่น แบบมีตัวเลขกำกับ แบบมีคำกำกับ หรือแบบอื่น ๆ และจำนวนช่องของคำตอบว่าจะให้มีจำนวนกี่ช่อง
4. เขียนข้อความที่ใช้ในมาตราวัด โดยให้ครอบคลุมคุณลักษณะที่ต้องการจะวัด
5. ตรวจสอบข้อความที่สร้างขึ้น ซึ่งอาจทำได้โดยผู้สร้างข้อความเองหรือนำไปให้ผู้เชี่ยวชาญตรวจสอบโดยพิจารณาว่าข้อความที่สร้างขึ้นมีความครอบคลุมวัตถุประสงค์ของการสร้างเพื่อวัดคุณลักษณะนั้น ๆ หรือไม่ รวมทั้งตรวจสอบความเหมาะสมของภาษาที่ใช้
6. นำไปทดลองใช้ (Try out) โดยนำข้อความที่ตรวจสอบแล้วไปทดลองใช้กับกลุ่มตัวอย่างจำนวนหนึ่ง เพื่อดูว่ายังมีข้อความที่ยังไม่ชัดเจนหรือไม่ แล้วปรับปรุงแก้ไขให้เป็นมาตราวัดที่มีความเชื่อถือได้

ข้อดีของมาตรประมาณค่าแบบลิเคอร์ท์

มาตรประมาณค่าตามวิธีการของลิเคอร์ท์เป็นมาตรประมาณค่าที่ได้รับความนิยมนำมาใช้มากที่สุดเพราะเป็นมาตรประมาณค่าที่มีข้อดีดังนี้ คือ (วิเชียร เกตุสิงห์, 2530; บุญธรรม กิจบริดาภิรสุทธิ, 2531)

1. มีวิธีการสร้างที่ง่ายประหยัดรวมทั้งสะดวกในการนำไปใช้การให้คะแนนและการวิเคราะห์ผล
2. ไม่ต้องหากกลุ่มผู้เชี่ยวชาญที่จะมาเป็นผู้ตัดสินเพื่อกำหนดค่าประจำข้อ
3. สามารถนำไปปรับใช้ในการวัดเจตคติที่มีต่อสิ่งต่าง ๆ หรือสถานการณ์ต่าง ๆ ได้อย่างกว้างขวาง และยังสามารถวัดได้ทั้งทิศทางและความเข้มของเจตคติได้ด้วย

4. เมื่อสร้างข้อคำถามได้ดีและมีคุณภาพ จะสามารถวัดเจตคติได้ตรงและเชื่อถือได้มากกว่า มาตรการวัดเจตคติแบบเทอร์สโตนเพราะผู้ตอบจะต้องตอบคำถามทุกข้อซึ่งการตอบตามวิธีของเทอร์สโตนผู้ตอบสามารถเลือกตอบเพียงบางข้อทำให้มีโอกาสที่ผู้ตอบจะบิดเบือนความจริงได้มากกว่าวิธีของลิเคอร์ท

ข้อเสียของมาตรการประมาณค่าแบบลิเคอร์ท

แม้ว่ามาตรการประมาณค่าแบบลิเคอร์ทจะมีจุดเด่นอยู่หลายประการ แต่ก็มีข้อเสียหลัก ๆ ดังนี้คือ

1. ผู้ตอบมาตรการที่มีลักษณะการเลือกคำตอบที่แตกต่างกันแต่ได้คะแนนรวมเท่ากัน นั่นคือ ผู้ตอบที่ได้คะแนนเฉลี่ย 3 อาจได้จากการเลือกคำตอบ " เป็นกลาง " ซึ่งมีค่านำหนักคะแนนเท่ากับ 3 ทุกข้อหรืออาจได้จากการเลือกคำตอบ " เห็นด้วยอย่างยิ่ง " และ " ไม่เห็นด้วยอย่างยิ่ง " อย่างละครึ่ง การแปลผลข้อมูลก็แตกต่างกัน จากปัญหาดังกล่าวนักวิชาการได้เสนอแนะว่า ควรวิเคราะห์ข้อมูลด้วยการจัดกระทำข้อมูลเป็นแบบมาตรอันดับมากกว่าจัดเป็นแบบมาตรอันตรภาค (Shaw and Wright, 1967) .

2. ปัญหาในการตอบ เพราะแต่ละบุคคลมีแนวโน้มในการตอบคำถามแตกต่างกัน บางคนชอบตอบแบบสุดขั้ว แต่บางคนชอบตอบเป็นกลาง ๆ ส่วนมากผู้ตอบนิยมตอบแต่คำตอบที่เป็นกลาง จึงทำให้ไม่ทราบเจตคติที่แท้จริงของผู้ตอบได้ ซึ่งเกิดมาจากสาเหตุ 6 ประการดังต่อไปนี้ คือ (Guilford, 1964; Cronbach, 1970)

1). ความรู้สึกในเรื่องของการเสี่ยง โดยการพยายามตอบเป็นกลาง ๆ ไว้เพื่อป้องกันความเสียหายที่จะเกิดขึ้นกับตัวเอง

2). ความเข้าใจความหมายของภาษาไม่ตรงกัน ซึ่งอาจจะเนื่องมาจากการใช้ข้อความที่หลากหลายหรือเนื่องมาจากการจัดเรียงลำดับข้อคำถามทางบวกและทางลบ จึงทำให้แต่ละบุคคลเข้าใจความหมายของการเห็นด้วยหรือไม่เห็นด้วยแตกต่างกัน

3). ขาดแรงจูงใจในการตอบ ซึ่งการให้เลือกคำตอบอาจทำให้บุคคลเลือกตอบโดยไม่คิดอะไร

4). การยอมรับเรื่องที่ถามโดยบุคคลที่เห็นด้วยกับเรื่องที่ถามจะสามารถตอบคำถามได้ตรงกับความเป็นจริงมากกว่าบุคคลที่ไม่ยอมรับเรื่องที่ถาม

5). ปัญหาในเรื่องเวลาที่ใช้ตอบซึ่งต้องให้เวลาที่เพียงพอในการพิจารณาคำตอบผู้ตอบก็จะตอบข้อคำถามอย่างรอบคอบ

6). ความรู้สึกซ่อนเร้นของผู้ตอบ โดยผู้ตอบพยายามปิดบังส่วนเสียดังกล่าวและจะแสดงออกมาเฉพาะลักษณะที่ดี สภาพการณ์เช่นนี้ทำให้ผู้ตอบเลือกคำตอบที่ไม่แสดงลักษณะที่แท้จริงของตนเอง

ตอนที่ 4 แนวความคิดเกี่ยวกับมาตรประมาณค่าแบบยึดพฤติกรรม (Behaviorally anchored rating scales)

ลักษณะของมาตรประมาณค่าแบบยึดพฤติกรรม

มาตรประมาณค่าแบบยึดพฤติกรรม (Behaviorally anchored rating scales - BARS) ได้รับการพัฒนาขึ้นโดย Smith และ Kendall (1963) ซึ่งได้นำมาใช้เป็นครั้งแรก ในการประเมินพฤติกรรมกรรมการปฏิบัติงานด้านการพยาบาล ซึ่งต่อมามีวิธีการนี้ได้ถูกนำมาประยุกต์ใช้ในการประเมินการปฏิบัติงานของบุคคลในหลากหลายอาชีพภายใต้ชื่อเรียกเป็นภาษาอังกฤษที่ต่างกันไปหลายชื่อ อาทิ Behaviorally Anchored Rating Scales (Witkin, 1984: 79) Behaviorally Based Rating Scales (Campbell และคณะ, 1973: 15) Behavioral Expectation Scales (Keaveny และ McGann, 1975: 695) เป็นต้น ส่วนชื่อที่ใช้ในภาษาไทยมีการใช้คำว่า มาตรประมาณค่าเชิงพฤติกรรม มาตรประมาณค่าแบบยึดพฤติกรรม ซึ่งในการวิจัยครั้งนี้ผู้วิจัยจะใช้คำว่ามาตรประมาณค่าแบบยึดพฤติกรรม

มาตรประมาณค่าแบบยึดพฤติกรรม ถือเป็นวิธีการใหม่ที่ออกแบบขึ้นเพื่อแก้ปัญหาในเรื่องของการจัดระดับชั้นคุณภาพของรายการที่ให้ประเมินด้วยวิธีการกำหนดหรือนิยามพฤติกรรมกรรมการปฏิบัติงานให้เห็นชัดเจน (Mathis, 1979: 305) โดยมุ่งหวังว่าสามารถช่วยลดโอกาสการเกิดความคลาดเคลื่อนอันเนื่องมาจากการตีความหมายข้อความหรือการรับรู้ข้อความที่ต่างกันไประหว่างผู้ประเมินต่างกลุ่มต่างคนกัน

มาตรประมาณค่าแบบยึดพฤติกรรมถือเป็นรูปแบบหรือลักษณะหนึ่งของมาตรประมาณค่า (Rating scales) ลักษณะของมาตรวัดในแต่ละมิติที่จะให้ประเมิน จะประกอบด้วยข้อความที่กระชับชัดเจน อธิบายพฤติกรรมกรรมการปฏิบัติงานต่างๆ ที่เกี่ยวข้องกับมิตินั้นๆ ข้อความดังกล่าวจะถูกวางในตำแหน่งอย่างเป็นลำดับและมีความต่อเนื่องบนมาตรวัด (Witkin, 1984)

สำหรับแนวคิดในการพัฒนามาตรประมาณค่าแบบยึดพฤติกรรมนั้น Crane (1986: 413) ได้กล่าวว่า วิธีการนี้เป็นวิธีการที่ส่งเสริมหรือเพิ่มประสิทธิภาพของการบริหารงานตามแนวทางของการบริหารโดยยึดวัตถุประสงค์ (Management by objectives approach - MBO) ทั้งนี้เพราะข้อความในมาตรประเมินจะแสดงให้เห็นว่า "อย่างไรจึงจะเรียกว่า เป็นพฤติกรรมกรรมการปฏิบัติงานที่มีประสิทธิภาพ"

ในการปฏิบัติงาน” ซึ่งลักษณะสำคัญที่ทำให้มาตรฐานค่าแบบยึดพฤติกรรมเป็นวิธีการที่แตกต่างไปจากวิธีการเดิม ๆ ทั่วไป คือ

1. เป็นวิธีการที่เน้นเป้าหมายเชิงพัฒนา หรือชี้แนะแนวทางการพัฒนา
2. เน้นในเรื่องพฤติกรรมกรรมการปฏิบัติงานของบุคคล
3. เป็นวิธีการที่เน้นการนิยามพฤติกรรมอย่างชัดเจน สังเกตได้ และวัดได้
4. เป็นวิธีการที่แสดงออกซึ่งความแตกต่างระหว่างคำว่าพฤติกรรมโดยทั่วไป (Behaviors) พฤติกรรมการปฏิบัติงาน (Performance) และประสิทธิภาพของการปฏิบัติงาน (Effectiveness)

ลักษณะของมาตรฐานค่าแบบยึดพฤติกรรม แสดงได้ดังตัวอย่างต่อไปนี้

คำถาม การตรวจสอบความเข้าใจของนักเรียนขณะสอน

- | | |
|--------|--|
|5 | ใช้คำถาม หรือโจทย์ตัวอย่าง ให้ผู้เรียนได้ใช้ความสามารถในชั้นประยุกต์ใช้กฎหรือทฤษฎี ด้วยการซักถามนักเรียนในห้องเป็นระยะ ๆ |
|4 | ใช้คำถามหรือโจทย์ประเภทวัดความรู้-ความจำ หรือความเข้าใจสู่สมถมนักเรียนเป็นระยะ ๆ |
|3 | ใช้คำถาม หรือโจทย์ตัวอย่างโดยไม่คำนึงว่าเป็นการวัดความสามารถระดับใด สอบถามนักเรียนให้ตอบพร้อมกัน |
|2 | ยกตัวอย่างคำถาม หรือโจทย์ แล้วบอกให้นักเรียนนำไปทดลองทำที่บ้านเป็นการตรวจสอบความสามารถของตนเอง |
|1 | ซักถามนักเรียนเป็นระยะ ๆ ว่ามีปัญหา หรือไม่เข้าใจประเด็นใดหรือไม่ |

ขั้นตอนการพัฒนามาตรฐานค่าแบบยึดพฤติกรรม

การพัฒนามาตรฐานค่าแบบยึดพฤติกรรมมีวิธีดำเนินการตามลำดับขั้นตอนดังต่อไปนี้

1. รวบรวมรายการเหตุการณ์ หรือรายการพฤติกรรมที่แสดงออกเกี่ยวกับการปฏิบัติงานในหน้าที่อย่างเฉพาะเจาะจง ทั้งพฤติกรรมที่แสดงถึงความมีประสิทธิภาพ และพฤติกรรมที่ด้อยประสิทธิภาพ จากกลุ่มบุคคลผู้มีความรู้ ความเข้าใจเกี่ยวกับงานที่จะประเมินเป็นอย่างดี
2. จัดกลุ่มเหตุการณ์หรือพฤติกรรมที่แสดงออกต่าง ๆ ออกเป็นกลุ่ม ๆ กำหนดชื่อ หรือนิยามความหมายของแต่ละกลุ่ม ถือเป็นมิติหรือด้าน (Dimensions) ที่ต้องการประเมิน ขณะเดียวกันก็กำหนดระดับของคุณลักษณะนั้น ๆ อย่างคร่าว ๆ ว่า มาก ปานกลาง หรือน้อยคืออย่างไร

3. ตรวจสอบยืนยันการจัดกลุ่ม การกำหนดนิยาม หรือการให้ชื่อกลุ่ม ในข้อ 2 โดยใช้กลุ่มผู้ตัดสินอีกกลุ่มหนึ่ง ทั้งนี้เพื่อขจัดความกำกวม ความสับสนของข้อความ และขจัดรายการข้อความที่เห็นว่าไม่มีสาระออกไป พร้อมทั้งให้กลุ่มตัวอย่างกลุ่มนี้จัดกลุ่มรายการข้อความใหม่อีกครั้งหนึ่ง โดยตัดสินรายการข้อความแต่ละรายการว่าควรอยู่ในกลุ่มใด ในขั้นนี้อาจใช้เกณฑ์ตามความสอดคล้องของความคิดเห็น ระหว่างร้อยละ 50 - 80 ของจำนวนผู้ตัดสินเป็นเกณฑ์ในการตัดสินว่ารายการข้อความใดควรจะอยู่ในกลุ่มใด

4. แปลงรายการพฤติกรรมให้อยู่ในรูปแบบของมาตรวัด โดยให้กลุ่มผู้ตัดสินกลุ่มที่ 3 ประเมินค่ารายการข้อความแต่ละรายการว่าได้แสดงออกถึงคุณลักษณะ (Trait) ที่มุ่งประเมินในระดับมากน้อยเพียงใด แล้วทำการวิเคราะห์ค่าเฉลี่ยเลขคณิต และส่วนเบี่ยงเบนมาตรฐานของคะแนนผลการประเมินทำการเปรียบเทียบค่าเฉลี่ยเลขคณิต และส่วนเบี่ยงเบนมาตรฐานระหว่างรายการข้อความต่าง ๆ แล้วเรียงลำดับรายการข้อความอย่างต่อเนื่องตามลำดับจากมากไปหาน้อย

5. สร้างแบบประเมิน ซึ่งจะประกอบด้วยชุดของมาตรวัดตามแนวตั้งจำนวนหนึ่งตามจำนวนมิติที่ได้สรุปในข้อ 3 ในแต่ละมาตรวัดไม่จำเป็นจะต้องเรียงลำดับรายการข้อความในลักษณะที่เป็นช่วงเท่ากัน (Equal interval) ทดสอบประสิทธิภาพของมาตรวัดอาจทำได้โดยใช้ผู้ประเมิน 3 คน ทำการประเมินพฤติกรรมของบุคคลใดบุคคลหนึ่ง แล้วนำผลการประเมินมาวิเคราะห์ค่าสหสัมพันธ์ และถือเป็นค่าความเที่ยงระหว่างผู้ประเมิน (Interrater reliability)

สภาพการประยุกต์ใช้มาตรประมาณค่าแบบยึดพฤติกรรมในการวิจัย

นักวิจัยและนักประเมินได้ประยุกต์ใช้มาตรประมาณค่าแบบยึดพฤติกรรมในการประเมินพฤติกรรม การปฏิบัติงานของบุคคลหลายลักษณะ อาทิ Campbell และคณะ (1973: 15-22) ได้พัฒนามาตรประมาณค่าแบบยึดพฤติกรรมสำหรับการประเมินพฤติกรรม การปฏิบัติงานของผู้จัดการฝ่ายในห้างขายส่งสินค้า Harari และคณะ (1973: 261-265) ได้พัฒนามาตรประมาณค่าแบบยึดพฤติกรรมเพื่อประเมินการสอนของคณะ Keaveny และ McGann (1975: 695-703) ได้ประเมินพฤติกรรม การปฏิบัติงานของลูกจ้างโดยเปรียบเทียบผลการประเมินระหว่างการใช้มาตรประมาณค่าแบบยึดพฤติกรรมกับมาตรประมาณค่าแบบกราฟิกและ Witkin (1984: 83-85) ได้ใช้มาตรประมาณค่าแบบยึดพฤติกรรมในกิจกรรมการประเมินความต้องการจำเป็นในเรื่องต่าง ๆ

ตอนที่ 5 งานวิจัยที่เกี่ยวข้องกับการศึกษาเปรียบเทียบคุณภาพของมาตรฐานค่าในด้าน ความเที่ยงและความคลาดเคลื่อนในการประมาณค่า

นักประเมินและนักวิจัยหลายท่านได้ศึกษาคุณภาพของมาตรฐานค่าในด้านความเที่ยงและความคลาดเคลื่อนในการประมาณค่าในสถานการณ์หรือประเด็นต่าง ๆ ดังต่อไปนี้

Campbell และคณะ (1973: 15-22) ต้องการที่จะพัฒนาแบบประเมินการปฏิบัติงานของผู้จัดการฝ่ายในห้างขายส่งสินค้า จึงได้ทำการศึกษาเปรียบเทียบระหว่างมาตรฐานค่าสองแบบ คือ แบบที่ใช้เทคนิคการประมาณค่าแบบรวมสรุป (Summated rating technique) ของ Likert และมาตรฐานค่าแบบของ Smith และ Kendall ซึ่งกำหนดสเกลด้วยลักษณะพฤติกรรม (Scaled behavioral anchors) เครื่องมือทั้งสองชุด มีพฤติกรรมที่ให้ประเมินรวม 9 ข้อ เหมือนกัน แต่แตกต่างกันที่ส่วนสเกล ผู้ที่ประเมินคือ ผู้จัดการห้างและผู้ช่วยผู้จัดการ ผลที่ได้ปรากฏว่า มาตรฐานค่าแบบยึดพฤติกรรมเกิด Leniency error และ Halo error น้อยกว่ามาตรฐานค่าแบบรวมสรุป ส่วนในด้านความตรงตามภาวะสันนิษฐานซึ่งดำเนินการโดยวิเคราะห์เมตริกซ์สหสัมพันธ์แบบ Multitrait - Multimethod เพื่อตรวจสอบ Convergent และ Discriminant validity นั้น พบว่า มาตรฐานค่าทั้งสองรูปแบบให้ผลการประเมินที่คมชัดเป็นไปตามเกณฑ์ของ Campbell และ Fiske โดยที่มาตรฐานค่าแบบยึดพฤติกรรมปรากฏ Discriminant validity สูงกว่ามาตรฐานค่าแบบรวมสรุป

Showers (1974: 5731-A) พบว่า ในการประเมินการสอนของครูนั้น นักเรียนมักเกิด Leniency error กล่าวคือ จะเลือกเฉพาะระดับสเกลที่มีค่าสูงเพียง 2-3 ค่าเท่านั้น ทำให้การประมาณค่าไม่ได้ผลเนื่องจากไม่สามารถบอกความแตกต่างของครูแต่ละคนได้ เขามีความคิดเห็นว่าอคติที่เกิดขึ้นอาจจะทำให้ลดลงได้ด้วยการปรับปรุงด้วยค่าที่ใช้บนสเกล ดังนั้นเขาจึงทำการวิจัยเพื่อปรับปรุงแบบประเมินการสอนของครูโดยทดลองใช้รูปแบบคำตอบเป็น 3 แบบ คือ 1. แบบลิเคอร์ท (Likert) 2. แบบประเมินค่า (Evaluative cues) และ 3. แบบเลือกตอบจากคำบรรยายสั้น ๆ (Short descriptive cues) ผลปรากฏว่า รูปแบบที่ 2 คือ แบบประเมินค่าทำให้เกิดอคติน้อยที่สุด ส่วนแบบลิเคอร์ทซึ่งเป็นที่ยอมรับมากนั้นมีแนวโน้มที่จะทำให้เกิดอคติมากที่สุด อย่างไรก็ตามค่าความเที่ยงของผู้ประเมินทั้งสามรูปแบบไม่แตกต่างกัน

Pierson (1975: 2718-A) ต้องการทราบว่า ในการวัดทัศนคติของนักเรียนเกรดหกที่มีผลสัมฤทธิ์ในการอ่านแตกต่างกันนั้น ระหว่างเทคนิคจำแนก (Semantic differential technique) กับมาตรประมาณค่าแบบกราฟิก รูปแบบใดจะให้ค่าความเที่ยงสูงกว่ากัน จากการทดลองกับนักเรียนเกรดหกจำนวน 585 คน ซึ่งถูกแบ่งตามระดับของทักษะในการอ่านเป็น 3 กลุ่ม คือ สูง กลาง ต่ำ ผลการวิจัยพบว่า ค่าความเที่ยงของแบบวัดทั้งสองรูปแบบนี้ขึ้นอยู่กับระดับของทักษะในการอ่านและแบบวัดที่ใช้เทคนิคจำแนกให้ค่าความเที่ยงสูงกว่า

Keaveny และ McGann (1975: 695-703) ได้พัฒนาและศึกษาเปรียบเทียบประสิทธิภาพของมาตรคาดหวังเชิงพฤติกรรม (Behavioral expectation scales) กับมาตรประมาณค่าแบบกราฟิกในการประเมินพฤติกรรมการสอนของศาสตราจารย์ 4 คน โดยนักศึกษาที่ลงทะเบียนเรียนกับศาสตราจารย์เหล่านั้นเป็นผู้ประเมิน โดยมาตรวัดทั้งสองรูปแบบประกอบด้วยรายการที่ให้ประเมินที่เหมือนกันจำนวน 13 รายการพฤติกรรม จะแตกต่างกันเฉพาะในส่วนของสเกลคำตอบเท่านั้น แล้วนำมาตรวัดทั้งสองรูปแบบไปให้นักศึกษา 4 กลุ่ม ทำการประเมินพฤติกรรมการสอนของผู้สอน แล้วทำการวิเคราะห์ค่าเฉลี่ยเลขคณิตและส่วนเบี่ยงเบนมาตรฐานของคะแนนผลการประเมิน วิเคราะห์เมตริกซ์สหสัมพันธ์แบบ Multitrait - Multimethod เพื่อตรวจสอบ Convergent และ Discriminant validity ของมาตรวัดทั้ง 2 รูปแบบ โดยถือว่ารายการพฤติกรรมที่แตกต่างกัน 13 รายการพฤติกรรม เป็นเสมือนคุณลักษณะที่แตกต่างกัน (Heterotrait) ผลการวิจัยพบว่า มาตรคาดหวังเชิงพฤติกรรมเกิด Halo error น้อยกว่ามาตรประมาณค่าแบบกราฟิก ทั้งนี้โดยพิจารณาจากค่าสหสัมพันธ์ระหว่างคู่รายการพฤติกรรมใด ๆ เปรียบเทียบระหว่างกรณีใช้มาตรคาดหวังเชิงพฤติกรรมเป็นเครื่องมือกับกรณีใช้มาตรประมาณค่าแบบกราฟิกเป็นเครื่องมือประเมิน ซึ่งปรากฏว่าค่าสหสัมพันธ์ภายใต้การใช้มาตรคาดหวังเชิงพฤติกรรมมีค่าน้อยกว่า อันแสดงถึงความเป็นอิสระจากกันของผลการประเมินแต่ละรายการพฤติกรรม ในการตรวจสอบ Leniency error จากการเปรียบเทียบค่าเฉลี่ยเลขคณิตของผลการประเมินรายการพฤติกรรมต่าง ๆ ระหว่างกรณีใช้มาตรวัดต่างรูปแบบกัน ปรากฏข้อสรุปว่ามาตรคาดหวังเชิงพฤติกรรมไม่สามารถลด Leniency error ได้ และจากการวิเคราะห์เมตริกซ์สหสัมพันธ์แบบ Multitrait-Multimethod พบว่ามาตรคาดหวังเชิงพฤติกรรมให้ความตรงเชิงจำแนก (Discriminant validity) สูงกว่ามาตรประมาณค่าแบบกราฟิก

Slater (1977: 6203-A) ต้องการจะพัฒนาแบบประเมินประสิทธิภาพการสอนของครูเพื่อที่จะลดอคติที่เกิดขึ้นเช่นเดียวกับ Shower เขาจึงได้ศึกษาโดยเปรียบเทียบระหว่างมาตรประมาณค่าแบบกราฟิก 2 แบบ คือ แบบที่มีสเกลเป็นคำคุณศัพท์ (Adjective anchored scales) กับแบบที่มีสเกลเป็นตัวอย่างพฤติกรรม (Behaviorally anchored scales) โดยที่แบบแรกจะมีข้อความ ตีมาก-ปาน

กลาง-เลวมาก กำหนดบนสเกล ส่วนแบบที่สองจะมีตัวอย่างพฤติกรรมของผู้สอนแทนค่าคุณศัพท์บนสเกล การวิจัยนี้มีสมมติฐานว่า การประเมินอาจารย์ผู้สอนด้วยแบบประเมินที่มีสเกลเป็นแบบพฤติกรรมจะเกิดอคติประเภท Leniency น้อยกว่าแบบประเมินที่มีสเกลเป็นแบบค่าคุณศัพท์ โดยได้นำแบบประเมินทั้งสองที่ให้ประเมินพฤติกรรม 5 พฤติกรรมไปให้นักศึกษาประเมินอาจารย์ 15 คน ซึ่งนักศึกษาที่เรียนกับอาจารย์แต่ละคนจะถูกสุ่มมาครั้งหนึ่งเพื่อให้แบบประเมินแบบแรกและที่เหลือจะใช้รูปแบบที่ 2 ผลจากการวิเคราะห์ความแปรปรวน พบว่า ไม่มีรูปแบบใดที่สามารถลดอคติได้สำเร็จ สำหรับค่าความเที่ยงของคะแนนเฉลี่ยในการประมาณค่าของแต่ละรูปแบบ พบว่า แบบที่มีสเกลเป็นค่าคุณศัพท์ให้ค่าความเที่ยงสูงกว่าเล็กน้อย 3 ใน 5 ของพฤติกรรมที่ให้ประเมิน

Cook (1988: 31-45) ได้ศึกษาผลของการใช้มาตรฐานประมาณค่าที่มีรูปแบบแตกต่างกัน 2 รูปแบบ คือ มาตรฐานประมาณค่าแบบยึดพฤติกรรม (Behaviorally anchored rating scales) กับ มาตรฐานประมาณค่าแบบลิเคิร์ต (Likert scale) และผลของการฝึกทักษะทางด้านการประเมินแก่ผู้ประเมินต่อค่าความคลาดเคลื่อนเชิงปล่อยคะแนน (Leniency error) และความคลาดเคลื่อนอันเนื่องมาจากการเคยมีปฏิสัมพันธ์มาก่อน (Halo error) โดยผู้วิจัยได้ใช้มาตรฐานค่าทั้ง 2 รูปแบบดังกล่าวให้นักศึกษา 2 กลุ่ม คือ กลุ่มที่ได้รับการฝึกทักษะกับกลุ่มที่ไม่ได้รับการฝึกทักษะ ทำการประเมินพฤติกรรมการสอนของศาสตราจารย์ 1 ท่าน แล้วทำการวิเคราะห์หาค่าเฉลี่ยเลขคณิตและส่วนเบี่ยงเบนมาตรฐานของคะแนนผลการประเมินแต่ละมิติ เปรียบเทียบระหว่างกลุ่มที่ได้รับการฝึกทักษะกับกลุ่มที่ไม่ได้รับการฝึกทักษะ พร้อมทั้งเปรียบเทียบคะแนนผลการประเมินระหว่างมาตรฐานวัดต่างรูปแบบกัน เฉพาะในกลุ่มที่ได้รับการฝึกทักษะ โดยในการตรวจสอบ Halo error ผู้วิจัยได้ใช้ส่วนเบี่ยงเบนมาตรฐานของคะแนนผลการประเมิน (Standard deviation) เป็นคะแนนบ่งชี้ความคลาดเคลื่อน (Halo score) ภายใต้เงื่อนไขว่า ถ้าคะแนนความคลาดเคลื่อนมีค่ามาก แสดงถึงการเกิด Halo error น้อยกว่ากรณีที่คะแนนความคลาดเคลื่อนมีค่าน้อย ผลการวิจัยพบว่า การฝึกทักษะมีผลต่อการลด Halo error และการประเมินโดยมาตรฐานประมาณค่าแบบยึดพฤติกรรมเกิด Leniency error น้อยกว่า มาตรฐานประมาณค่าแบบลิเคิร์ต ในขณะที่มาตรฐานประมาณค่าแบบลิเคิร์ตเกิด Halo error น้อยกว่า

สุพักตร์ พิบูลย์ (2534) ได้เปรียบเทียบคุณภาพของมาตรฐานประมาณค่าเชิงพฤติกรรมกับมาตรฐานประมาณค่าแบบกราฟิก ในด้านความตรง ความเที่ยง และลักษณะความคลาดเคลื่อนในการประมาณค่า ในสถานการณ์การประเมินพฤติกรรมด้านการประเมินผลของครู ผลการวิจัย พบว่า มาตรฐานประมาณค่าเชิงพฤติกรรมให้ผลการประเมินที่มีคุณภาพมากกว่ามาตรฐานประมาณค่าแบบกราฟิก คือ (1) มีความตรงตามภาวะสัมพันธ์ (Construct validity) สูงกว่าเมื่อตรวจสอบโดยวิธีวิเคราะห์เมตริกซ์สหสัมพันธ์แบบ Multitrait - Multimethod และวิธีตรวจสอบความแตกต่างระหว่างกลุ่มครูที่ปฏิบัติงาน

มีประสิทธิภาพต่างกัน (2) ให้ค่าความเที่ยงระหว่างผู้ประเมิน (Rater reliability) สูงกว่าและ(3) ปรากฏลักษณะความคลาดเคลื่อนในเชิงปล่อยคะแนน (Leniency error) และความคลาดเคลื่อนอันเนื่องมาจากการเคยมีปฏิสัมพันธ์มาก่อนระหว่างผู้ประเมินกับเป้าหมายของการประเมิน (Halo error) น้อยกว่า

จากที่กล่าวมาข้างต้น สามารถสรุปงานวิจัยที่เกี่ยวข้องกับการศึกษาเปรียบเทียบคุณภาพของมาตรฐานประเมินค่าในด้านความเที่ยงและความคลาดเคลื่อนในการประเมินค่าจำแนกตามสถานการณ์กลุ่มตัวอย่าง และประเด็นที่ทำการศึกษาดังตารางที่ 2 ดังนี้



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 2 สรุปงานวิจัยที่เกี่ยวข้องกับการศึกษาเปรียบเทียบคุณภาพของมาตรฐานค่าในด้านความเที่ยงและความคลาดเคลื่อนในการประมาณค่าจำแนกตามสถานการณ์ กลุ่มตัวอย่าง และประเด็นที่ทำการศึกษา

ผู้วิจัย/ปีที่ทำ	สถานการณ์ที่ศึกษา	กลุ่มตัวอย่าง	ประเด็นที่ทำการศึกษา				ผลการศึกษา
			Leniency error	Halo error	ความตรง	ความเที่ยง	
Campbell และ คณะ (1973)	เปรียบเทียบระหว่างมาตรฐานค่าแบบรวมสรุปและมาตรฐานค่าแบบยึดพฤติกรรม	<ul style="list-style-type: none"> ผู้จัดการห้าง ผู้ช่วยผู้จัดการ 	/	/	/	-	<ul style="list-style-type: none"> BARS เกิด Leniency error และ Halo error น้อยกว่า BARS ปรากฏ Discriminant validity สูงกว่า
Shower (1974)	เปรียบเทียบรูปแบบของแบบประเมินที่ใช้ในการประเมินการสอนของครู 3 รูปแบบ คือ แบบลิเคอร์ท แบบประเมินค่า และแบบเลือกตอบจากคำบรรยายสั้น ๆ	<ul style="list-style-type: none"> นักเรียน 	/	/	-	/	<ul style="list-style-type: none"> แบบประเมินค่าเกิดอคติน้อยที่สุดและแบบลิเคอร์ทเกิดอคติมากที่สุดโดยที่ทั้ง 3 รูปแบบมีค่าความเที่ยงไม่แตกต่างกัน
Pierson (1975)	เปรียบเทียบความเที่ยงจากการวัดทัศนคติของนักเรียนเกรดหกที่มีผลสัมฤทธิ์ในการอ่านต่างกันระหว่างเทคนิคน้อยจำแนก (Semantic differential technique) กับมาตรฐานค่าแบบกราฟิก	<ul style="list-style-type: none"> นักเรียน 	-	-	-	/	<ul style="list-style-type: none"> แบบเทคนิคน้อยจำแนกให้ค่าความเที่ยงสูงกว่ามาตรฐานค่าแบบกราฟิก

ตารางที่ 2 (ต่อ)

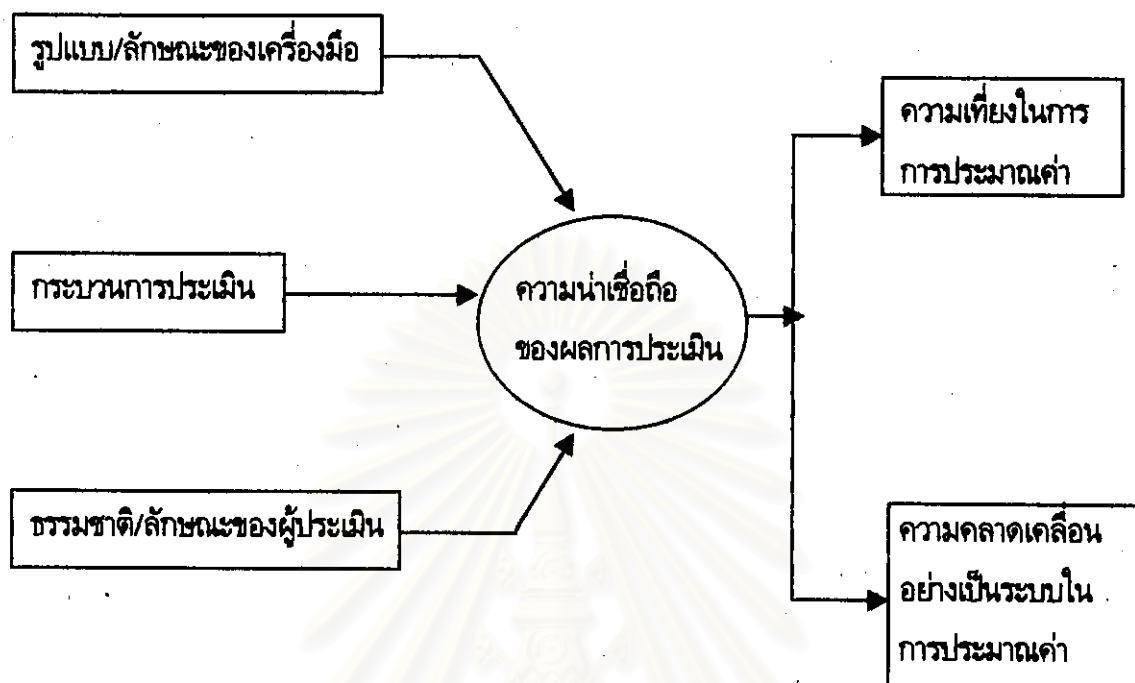
ผู้วิจัย/ปีที่ทำ	สถานการณ์ที่ศึกษา	กลุ่มตัวอย่าง	ประเด็นที่ทำการศึกษา				ผลการศึกษา
			Leniency error	Halo error	ความตรง	ความเที่ยง	
Keaveny และ McGann (1975)	เปรียบเทียบมาตรฐานคาดหวังพฤติกรรม (Behavioral expectation scales) กับมาตรฐานค่าแบบกราฟิกในการประเมินการสอนของอาจารย์	● นักเรียน	/	/	/	-	<ul style="list-style-type: none"> ● BARS เกิด Halo error น้อยกว่ามาตรฐานประมาณค่าแบบกราฟิก ● BARS ไม่สามารถลด Leniency error ได้ ● BARS ให้ความตรงเชิงจำแนกสูงกว่า
Slater (1977)	เปรียบเทียบรูปแบบของมาตรฐานค่าที่มีสเกลเป็นคำคุณศัพท์กับมีสเกลเป็นตัวอย่างพฤติกรรมในการประเมินประสิทธิภาพการสอนของครู	● นักเรียน	/	-	-	/	<ul style="list-style-type: none"> ● ไม่มีรูปแบบใดที่ลดอคติได้สำเร็จ ● แบบที่มีสเกลเป็นคำคุณศัพท์ให้ค่าความเที่ยงสูงกว่าเล็กน้อย 3 ใน 5 ของพฤติกรรมที่ให้ประเมิน
Cook (1988)	ผลของการใช้รูปแบบมาตรฐานค่าและการฝึกทักษะแก่ผู้ประเมิน ที่มีต่อความคลาดเคลื่อนในการประเมิน	● นักเรียน	/	/	-	-	● BARS เกิด Leniency error และ Halo error น้อยกว่า
สุพัทธ์ พิบูลย์ (2534)	เปรียบเทียบคุณภาพของมาตรฐานค่าเชิงพฤติกรรมกับมาตรฐานค่าแบบกราฟิกในการประเมินพฤติกรรมด้านการประเมินผลของครู	<ul style="list-style-type: none"> ● ครูผู้สอน ● หัวหน้าหมวดวิชา 	/	/	/	/	● BARS ให้ Construct validity และค่าความเที่ยงระหว่างผู้ประเมิน (Rater reliability) สูงกว่า รวมทั้ง เกิด Leniency error และ Halo error น้อยกว่า

จากที่กล่าวมาทั้งหมด จะเห็นได้ว่า ได้มีการศึกษาเปรียบเทียบคุณภาพของมาตรฐานค่ารูปแบบต่าง ๆ ในด้านความเที่ยงและความคลาดเคลื่อนในการประมาณค่าในสถานการณ์ที่แตกต่างกันซึ่งส่วนใหญ่จะเป็นการศึกษาเปรียบเทียบในลักษณะของการให้บุคคลหนึ่งประเมินพฤติกรรมของอีกบุคคลหนึ่งโดยสถานการณ์ที่ได้รับความสนใจศึกษากันอย่างแพร่หลายที่สุด คือ การให้นักเรียนประเมินพฤติกรรมของครูผู้สอน จากประเด็นของการศึกษาดังกล่าวผู้วิจัยมีความเห็นว่า มาตรฐานค่าแบบยึดพฤติกรรมเป็นอีกรูปแบบหนึ่งของมาตรฐานค่าที่สามารถช่วยลดความคลาดเคลื่อนอย่างเป็นระบบในการประมาณค่าได้ นอกจากนี้ยังมีความตรงเชิงจำแนกและความเที่ยงระหว่างผู้ประเมินสูงกว่ามาตรฐานค่ารูปแบบอื่น ซึ่งจากการศึกษางานวิจัยต่างประเทศที่ทำการศึกษารูปแบบของมาตรฐานค่าแบบยึดพฤติกรรมกับมาตรฐานค่าแบบลิเคอร์ท ผู้วิจัยมีความเห็นว่า งานวิจัยส่วนใหญ่จะศึกษาเฉพาะในสถานการณ์ของการประเมินผู้อื่นเท่านั้น ส่วนสถานการณ์ของการประเมินตนเองนั้นยังไม่มีผู้ใดศึกษาและสรุปไว้ ประกอบกับสภาพการใช้เครื่องมือในการประเมินทั้งประเมินตนเองและประเมินผู้อื่นในบริบทของการประเมินการสอนของครูในประเทศไทยนั้นส่วนมากจะนิยมใช้มาตรฐานค่าแบบลิเคอร์ทเป็นเครื่องมือในการประเมิน

ดังนั้นด้วยเหตุผลดังกล่าวผู้วิจัยจึงสนใจที่จะศึกษาเปรียบเทียบผลของการใช้มาตรฐานค่าแบบลิเคอร์ทและมาตรฐานค่าแบบยึดพฤติกรรมด้านความเที่ยงและความคลาดเคลื่อนอย่างเป็นระบบในการประมาณค่าในการประเมินการสอนของครู ซึ่งจะศึกษาในสถานการณ์ที่ให้ครูผู้สอนประเมินตนเองและให้นักเรียนประเมินครูผู้สอน ทั้งนี้เพื่อให้ได้ข้อสรุปที่ชัดเจนและเป็นประโยชน์เกี่ยวกับคุณภาพของมาตรฐานค่าทั้งสองรูปแบบในฐานะที่เป็นเครื่องมือที่ใช้ในการประเมินการสอนของครูต่อไป

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

กรอบแนวคิดตามทฤษฎี



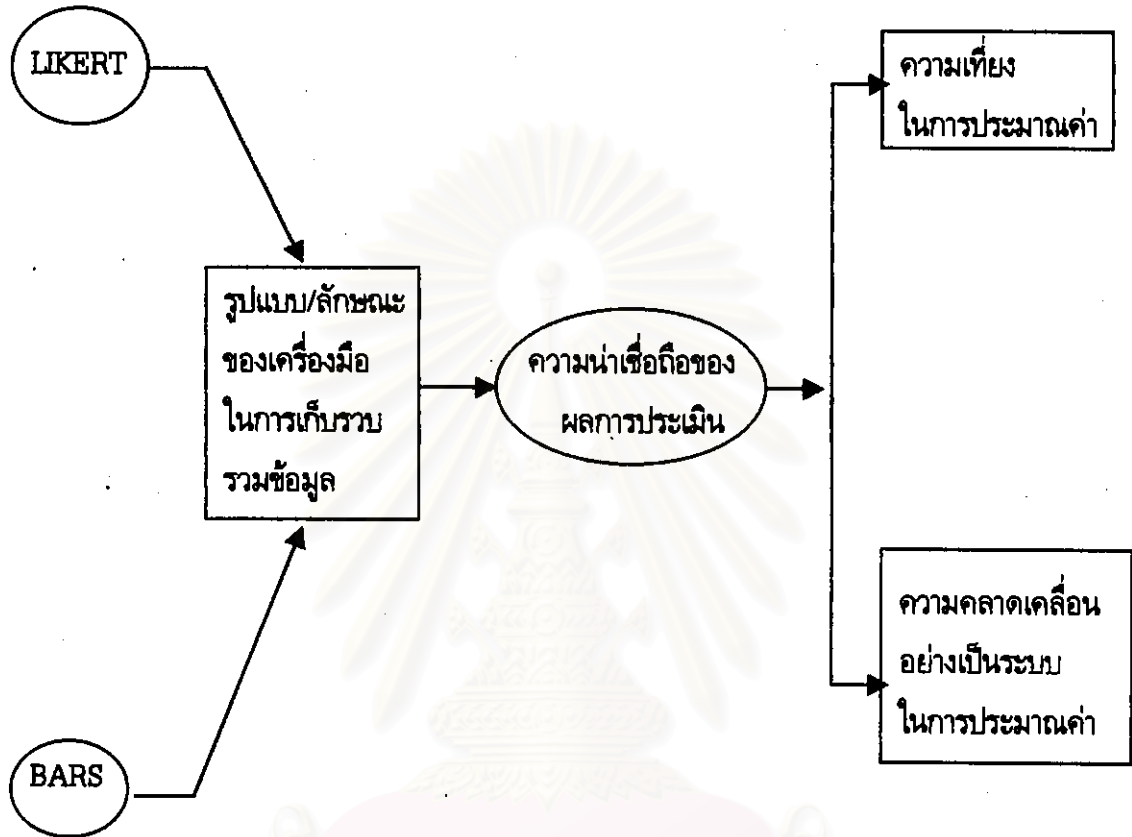
แผนภาพที่ 1 ตัวแปรที่ส่งผลต่อความน่าเชื่อถือของผลการประเมิน

จากแผนภาพที่ 1 แสดงให้เห็นตัวแปรที่ส่งผลต่อความน่าเชื่อถือของผลการประเมิน ซึ่งได้แก่ รูปแบบหรือลักษณะของเครื่องมือในการเก็บรวบรวมข้อมูล กระบวนการประเมิน และธรรมชาติหรือลักษณะของผู้ประเมิน

สำหรับการวิจัยครั้งนี้ ผู้วิจัยมีความสนใจที่จะศึกษาเฉพาะตัวแปรด้านรูปแบบหรือลักษณะของเครื่องมือเก็บรวบรวมข้อมูลที่นำมาใช้ในการทวนสอบการประเมินว่าจะส่งผลต่อความน่าเชื่อถือของผลการประเมินด้านความเที่ยงและความคลาดเคลื่อนอย่างเป็นระบบในการประมาณค่าอย่างไรบ้าง

ดังนั้นกรอบแนวคิดในการวิจัยครั้งนี้ จึงสามารถแสดงได้ดังแผนภาพที่ 2

กรอบแนวคิดที่ใช้ในการวิจัย



แผนภาพที่ 2 ตัวแปรที่สนใจศึกษาในการเปรียบเทียบความเที่ยงและความคลาดเคลื่อนอย่างเป็นระบบในการประมาณค่าของมาตรฐานค่าแบบลิเคอร์ท์และมาตรฐานค่าแบบยึดพฤติกรรม

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย