



ความเป็นมาและความสำคัญของปัญหา

ในการจัดกิจกรรมการเรียนการสอนให้สามารถบรรลุผลตามจุดมุ่งหมายของหลักสูตรนั้นครูนับว่าเป็นผู้ที่มีความสำคัญเป็นอย่างยิ่ง ทั้งนี้เนื่องจากเป็นผู้ที่มีบทบาทในการกระตุ้น แนะนำตลอดจนช่วยเหลือให้นักเรียนให้เกิดการเรียนรู้และพัฒนาผู้เรียนให้เกิดความเจริญงอกงามในทุก ๆ ด้าน ซึ่งการที่ครูจะสามารถจัดกิจกรรมการเรียนการสอนให้ดำเนินไปถึงจุดมุ่งหมายดังกล่าวนั้นต้องอาศัยความสามารถของครูในการจัดการเรียนการสอนที่เหมาะสมกับผู้เรียน นั่นก็คือการจัดการเรียนการสอนของครูจะบรรลุวัตถุประสงค์มากน้อยเพียงใดย่อมขึ้นอยู่กับประสิทธิภาพในการสอนของครูเป็นสำคัญ

การที่จะรู้ว่าการจัดการเรียนการสอนของครูจะมีประสิทธิภาพหรือประสบความสำเร็จมากน้อยเพียงใดนั้น จำเป็นอย่างยิ่งที่จะต้องมีการประเมินผลการปฏิบัติงานหรือประเมินผลการสอนของครู ทั้งนี้เพื่อที่ครูจะได้ทราบข้อบกพร่องของตนเองและนำข้อมูลที่ได้มาปรับปรุงแก้ไขการสอนของตนเองให้มีประสิทธิภาพมากยิ่งขึ้น รวมทั้งอาจนำผลการประเมินการสอนไปพิจารณาถึงความก้าวหน้าของการปฏิบัติงานด้วย (บุญธรรม กิจปรีดานวิสุทธิ, 2536) ซึ่งในการประเมินการสอนของครูนั้นสิ่งที่ควรคำนึงถึง ได้แก่ แหล่งข้อมูลของการประเมิน วิธีการประเมิน ตลอดจนเครื่องมือที่ใช้ในการประเมิน โดยสิ่งต่างๆดังกล่าวนี้นับว่าเป็นปัจจัยสำคัญที่มีผลต่อคุณภาพของผลการประเมินโดยเฉพาะเครื่องมือที่ใช้ในการประเมินจะต้องมีคุณภาพทั้งทางด้านความตรงและความเที่ยงอันเป็นคุณลักษณะที่สำคัญของเครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูลที่ดีและมีคุณภาพ (ศิริชัย กาญจนวาสี และคณะ, 2539)

สำหรับวิธีการหรือเครื่องมือที่สามารถนำมาใช้ในการประเมินประสิทธิภาพการสอนหรือผลการสอนของครูนั้นมีอยู่ด้วยกันหลายชนิด ได้แก่ การใช้วิดีโอเทป เครื่องบันทึกเสียง แบบสอบถาม แบบตรวจสอบรายการ แบบประมาณค่า การสัมภาษณ์ การสังเกตและแบบสอบ เป็นต้น (อุทุมพร จามรมาน, 2530) ซึ่งในบรรดาวิธีการหรือเครื่องมือที่นำมาใช้ในการประเมินประสิทธิภาพการสอนดังกล่าว แบบประมาณค่าหรือมาตราประมาณค่า (Rating scale) นับเป็นรูปแบบหนึ่งของเครื่องมือที่นิยมนำมาใช้ในการประเมินผลการสอนของครู จากการสำรวจวิทยานิพนธ์ที่ศึกษาเกี่ยวกับเรื่องการประเมินประสิทธิภาพการสอนของครูผู้สอนทั้งที่ประเมินโดยตนเอง ประเมินจากเพื่อนร่วมงานและประเมินจากนักเรียน ผู้วิจัยพบว่าเครื่องมือที่มีลักษณะเป็นมาตราประมาณค่าถูกนำมาใช้ในการประเมินมากที่สุด โดย

เฉพาะอย่างยิ่งมาตรฐานค่าแบบลิเคอร์ท (Likert rating scale) ซึ่งสอดคล้องกับ Berk (1979) ที่กล่าวว่า มาตรฐานค่าที่ใช้เป็นเครื่องมือในการประเมินครูผู้สอนนั้น มาตรฐานค่าแบบลิเคอร์ท นิยมนำมาใช้มากที่สุดเนื่องจากสามารถสร้างและแปลผลได้ง่าย กล่าวคือ ครูผู้สอนสามารถสร้างเองได้ เมื่อนำไปใช้กับนักศึกษา ก็ง่ายต่อการเข้าใจและผู้บริหารสามารถแปลผลได้ง่าย โดยเฉพาะอย่างยิ่งเป็น เครื่องมือที่สามารถสร้างให้มีความเที่ยงและความตรงได้แม้ว่าจะประกอบด้วยจำนวนข้อไม่มาก

สำหรับการใช้มาตรฐานค่าเป็นเครื่องมือในการเก็บรวบรวมข้อมูลนั้น กิลฟอร์ด (Guilford, 1954) ได้กล่าวว่า มักทำให้ข้อมูลที่ได้มีความคลาดเคลื่อนสูง โดยมีแหล่งความคลาดเคลื่อนที่สำคัญอยู่ 3 แหล่ง คือ แหล่งแรกความคลาดเคลื่อนเนื่องจากเครื่องมือบกพร่อง เช่น ภาษาที่ใช้ไม่เหมาะสม กากวม ขาดการนิยามความหมายของคุณลักษณะที่จะให้ประเมิน แหล่งที่สองความคลาดเคลื่อนจาก กระบวนการประเมิน เช่น คุณลักษณะที่จะให้ประเมินมีมากเกินไป และแหล่งที่สามความคลาดเคลื่อนที่ เกิดจากตัวผู้ตอบหรือผู้ประเมิน ได้แก่ ความลำเอียงที่เกิดจากบุคลิกภาพของผู้ตอบ เช่น ประเมินสูงกว่า ที่เป็นจริง (Leniency error) ประเมินผิดจากความเป็นจริงอันเนื่องมาจากการมีความรู้สึกบางอย่างต่อ ผู้ถูกประเมิน (Halo error) ประเมินทุกคนในระดับปานกลางหรือสูงกว่าปานกลางในทุกลักษณะ (Generosity error) หรือประเมินเฉพาะค่ากลาง ๆ เพื่อป้องกันความผิดพลาดเนื่องจากมีข้อมูลเกี่ยวกับ ผู้ถูกประเมินไม่เพียงพอ (Error of central tendency) ซึ่งความคลาดเคลื่อนทั้งหมดที่กล่าวมานี้ส่งผล ให้ข้อมูลที่ได้มีความตรงและความเที่ยงต่ำ ซึ่งสอดคล้องกับ Wolf (1988) และ Thorndike & Hagen (1955) ที่กล่าวไว้ว่า ในการใช้มาตรฐานค่าเป็นเครื่องมือในการประเมินตนเอง ประเมินผู้อื่นหรือ ประเมินกระบวนการนั้น ข้อมูลที่ได้มักขาดความน่าเชื่อถือทั้งนี้โดยมีสาเหตุมาจาก

1. ความไม่เต็มใจและไม่พอใจที่จะตัดสินใจอย่างละเอียดรอบคอบของผู้ให้ข้อมูลเนื่องจากไม่ ต้องการให้ตนเองเกิดความลำบากในการตัดสินใจและไม่ต้องการที่จะทำร้ายความรู้สึกของผู้ถูกประเมิน
2. ความสามารถในการตอบให้ตรงกับความเป็นจริงของผู้ตอบหรือผู้ให้ข้อมูลเป็นสิ่งที่ทำได้ยาก เนื่องจากโอกาสในการสังเกตหรือรู้จักสิ่งที่จะประเมินมีน้อย ไม่เข้าใจคุณลักษณะที่จะให้ประเมินว่าคือ อะไร ไม่สามารถจำแนกระดับความเข้มของพฤติกรรมหรือคำตอบได้และเกิดจากบุคลิกภาพส่วนตัวของผู้ประเมินเอง

ด้วยเหตุผลดังกล่าว นักจิตวิทยาและนักการศึกษาจึงได้เสนอแนะแนวทางแก้ไขเพื่อที่จะลด ความคลาดเคลื่อนและเพิ่มความน่าเชื่อถือในการใช้มาตรฐานค่าให้ได้ผลถูกต้องแม่นยำมากยิ่งขึ้น อาทิ

Guilford (1954: 280) ได้เสนอแนะว่า การให้ความรู้และฝึกฝนผู้ประเมินอย่างดีเป็นวิธีการที่มีประสิทธิภาพมากที่สุดในการประมาณค่าโดยเฉพาะการลดความคลาดเคลื่อนอย่างเป็นระบบ

Thomdike และHagen (1977: 487) ได้เสนอแนะว่า การประเมินควรจะทำเฉพาะลักษณะที่ปรากฏให้เห็นชัดเจน ลักษณะที่จะให้ประเมินต้องแสดงออกมาให้เห็นในรูปพฤติกรรมและการตัดสินใจกระทำจากพฤติกรรมที่เห็นนั้น รูปแบบของการประมาณค่าต้องปรับปรุงเพื่อบังคับให้ผู้ประเมินบอกความแตกต่างหรือเพื่อความคุมการที่ผู้ประเมินมีมาตรฐานในการตัดสินใจที่ต่างกัน

Oppenhiem (1966: 86-87) ได้เสนอแนะว่า ในการสร้างมาตรฐานประมาณค่านั้นจะต้องกำหนดขอบเขตและทิศทางของลักษณะที่จะประเมินให้เป็นที่เข้าใจตรงกันเสียก่อน จากนั้นจึงกำหนดระดับในการประมาณค่าและกำหนดความหมายให้แก่แต่ละระดับ และเพื่อให้ลักษณะที่จะประเมินมีความชัดเจนควรมีการอธิบายลักษณะนั้น ๆ ไว้อย่างเฉพาะเจาะจง

Mathis (1979: 297) ได้เสนอแนะว่า ในการที่จะลดอคติของผู้ประเมินนั้น นอกจากการให้ความรู้ความเข้าใจในเรื่องความคลาดเคลื่อนลักษณะต่าง ๆ แก่ผู้ประเมินแล้ว สิ่งสำคัญยิ่งอีกประการหนึ่งก็คือ การนิยามหรือกำหนดความหมายของระดับการประเมินแต่ละระดับให้ชัดเจน อีกทั้งกำหนดขอบเขตในการพิจารณาในแต่ละรายการหรือแต่ละข้อให้ผู้ประเมินแต่ละคนสามารถเข้าใจตรงกันก็จะสามารถแก้ปัญหาหรือลดปัญหาของการเกิดความคลาดเคลื่อนในการประเมินได้ พร้อมกันนี้ได้เสนอแนะว่า มาตรฐานประมาณค่าแบบยึดพฤติกรรม (Behaviorally anchored rating scales) เป็นรูปแบบหนึ่งของความพยายามที่จะลดความคลาดเคลื่อนอย่างเป็นระบบในการประเมิน ซึ่งลักษณะของมาตรฐานแบบนี้ถือเป็นวิธีการใหม่ที่ออกแบบขึ้นเพื่อแก้ปัญหาในเรื่องของการจัดระดับชั้นคุณภาพของรายการที่ให้ประเมินด้วยวิธีการกำหนดหรือนิยามพฤติกรรมการปฏิบัติงานให้เห็นชัดเจน สังเกตได้และวัดได้

จากการพิจารณาข้อเสนอแนะของนักวัดผลดังกล่าว ผู้วิจัยตั้งข้อสังเกตว่ามาตรฐานประมาณค่าแบบยึดพฤติกรรมเป็นแนวทางหนึ่งที่มีลักษณะสอดคล้องกับเงื่อนไขหรือข้อเสนอแนะสำหรับการปรับปรุงการประมาณค่าให้มีความน่าเชื่อถือหลายประการประกอบกับการมีข้อมูลยืนยันเกี่ยวกับคุณภาพของมาตรฐานประมาณค่าชนิดนี้ในงานวิจัยต่างประเทศหลายเรื่อง ที่แสดงให้เห็นว่ามาตรฐานประมาณค่าชนิดนี้ให้ผลการประเมินที่ช่วยแก้ปัญหาในเรื่องของการเกิดความคลาดเคลื่อนในการประเมิน ได้อาทิ

Campbell และคณะ (1973: 15-22) ต้องการที่จะพัฒนาแบบประเมินการปฏิบัติงานของผู้จัดการฝ่าย ในทั้งชายส่งสินค้า จึงได้ทำการศึกษารียบเทียบระหว่างมาตรประมาณค่าสองรูปแบบ คือแบบที่ใช้เทคนิคการประมาณค่าแบบรวมสรุป (Summated rating technique) ของ Likert และมาตรประมาณค่าแบบของ Smith และ Kendall ซึ่งกำหนดสเกลด้วยลักษณะพฤติกรรม (Scaled behavioral anchors) โดยเครื่องมือทั้งสองชุดมีพฤติกรรมที่ให้ประเมินรวม 9 ข้อเหมือนกัน แต่แตกต่างกันที่ส่วนของสเกล ผู้ที่ประเมินคือ ผู้จัดการห้างและผู้ช่วยผู้จัดการ ผลการวิจัยพบว่ามาตรประมาณค่าแบบยึดพฤติกรรมเกิด Leniency error และ Halo error น้อยกว่ามาตรประมาณค่ารวมสรุป ส่วนในด้านความตรงตามภาวะสันนิษฐานซึ่งใช้การวิเคราะห์เมตริกซ์สหสัมพันธ์แบบ Multitrait-Multimethod นั้นพบว่ามาตรประมาณค่าแบบยึดพฤติกรรมปรากฏ Discriminant validity สูงกว่ามาตรประมาณค่าแบบรวมสรุป

Keaveny และ McGann (1976: 695-703) ได้พัฒนาและศึกษารียบเทียบประสิทธิภาพของมาตรคาดหวังพฤติกรรม (Behavioral expectation scales) กับมาตรประมาณค่าแบบกราฟิกในการประเมินพฤติกรรมการสอนของศาสตราจารย์ 4 คนโดยนักศึกษาที่ลงทะเบียนเรียนกับศาสตราจารย์เหล่านั้นเป็นผู้ประเมิน โดยมาตรวัดทั้งสองรูปแบบประกอบด้วยรายการที่ให้ประเมินที่เหมือนกันจำนวน 13 รายการพฤติกรรม จะแตกต่างกันเฉพาะในส่วนของสเกลคำตอบเท่านั้น แล้วนำมาตราวัดทั้งสองรูปแบบไปให้นักศึกษา 4 กลุ่ม ทำการประเมินพฤติกรรมการสอนของผู้สอน แล้วทำการวิเคราะห์ค่าเฉลี่ยเลขคณิตและส่วนเบี่ยงเบนมาตรฐานของคะแนนผลการประเมิน วิเคราะห์เมตริกซ์สหสัมพันธ์แบบ Multitrait - Multimethod เพื่อตรวจสอบ Convergent และ Discriminant validity ของมาตรวัดทั้ง 2 รูปแบบ โดยถือว่า รายการพฤติกรรมที่แตกต่างกัน 13 รายการพฤติกรรมเป็นเสมือนคุณลักษณะที่แตกต่างกัน (Heterotrait) ผลการวิจัยพบว่า มาตรคาดหวังพฤติกรรมเกิด Halo error น้อยกว่ามาตรประมาณค่าแบบกราฟิก ทั้งนี้โดยพิจารณาจากค่าสหสัมพันธ์ระหว่างรายการพฤติกรรมใด ๆ เปรียบเทียบระหว่างกรณีใช้มาตรคาดหวังพฤติกรรมเป็นเครื่องมือกับกรณีใช้มาตรประมาณค่าแบบกราฟิกเป็นเครื่องมือประเมินซึ่งปรากฏว่าค่าสหสัมพันธ์ภายใต้การใช้มาตรคาดหวังพฤติกรรมมีค่าน้อยกว่า อันแสดงถึงความเป็นอิสระจากกันของผลการประเมินแต่ละรายการพฤติกรรม ในการตรวจสอบ Leniency error จากการเปรียบเทียบค่าเฉลี่ยเลขคณิตของผลการประเมินรายการพฤติกรรมต่าง ๆ ระหว่างกรณีใช้มาตรวัดต่างรูปแบบกัน ปรากฏข้อสรุปว่ามาตรคาดหวังพฤติกรรมไม่สามารถลด Leniency error ได้ และจากการวิเคราะห์เมตริกซ์สหสัมพันธ์แบบ Multitrait - Multimethod พบว่า มาตรคาดหวังพฤติกรรมให้ความตรงเชิงจำแนก (Discriminant validity) สูงกว่ามาตรประมาณค่าแบบกราฟิก

Cook (1988: 31-45) ได้ศึกษาผลของการใช้มาตรฐานค่าที่มีรูปแบบแตกต่างกัน 2 รูปแบบ คือ มาตรฐานค่าแบบยึดพฤติกรรม (Behaviorally anchored rating scales) กับ มาตรฐานค่าแบบลิเคอร์ท (Likert scale) และผลของการฝึกทักษะแก่ผู้ประเมินต่อค่าความคลาดเคลื่อนเชิงปล่อยคะแนน (Leniency error) และความคลาดเคลื่อนอันเนื่องมาจากการเคยมีปฏิสัมพันธ์มาก่อน (Halo error) โดยผู้วิจัยได้ใช้มาตรฐานค่าทั้ง 2 รูปแบบดังกล่าวให้นักศึกษา 2 กลุ่ม คือ กลุ่มที่ได้รับการฝึกทักษะกับกลุ่มที่ไม่ได้รับการฝึกทักษะทำการประเมินพฤติกรรมการสอนของศาสตราจารย์ 1 ท่าน แล้วทำการวิเคราะห์หาค่าเฉลี่ยเลขคณิตและส่วนเบี่ยงเบนมาตรฐานของคะแนนผลการประเมินแต่ละมิติ เปรียบเทียบระหว่างกลุ่มที่ได้รับการฝึกทักษะกับกลุ่มที่ไม่ได้รับการฝึกทักษะ พร้อมทั้งเปรียบเทียบคะแนนผลการประเมินระหว่างมาตรวัดต่างรูปแบบกันเฉพาะในกลุ่มที่ได้รับการฝึกทักษะ โดยในการตรวจสอบ Halo error ผู้วิจัยได้ใช้ส่วนเบี่ยงเบนมาตรฐานของคะแนนผลการประเมิน (Standard Deviation) เป็นคะแนนงซึ่งความคลาดเคลื่อน (Halo score) ภายใต้เงื่อนไขว่า ถ้าคะแนนความคลาดเคลื่อนมีค่ามากแสดงถึงการเกิด Halo error น้อยกว่ากรณีที่คะแนนความคลาดเคลื่อนมีค่าน้อย ผลการวิจัยพบว่า การฝึกทักษะมีผลต่อการลด Halo error และการประเมินโดยมาตรฐานค่าแบบยึดพฤติกรรมเกิด Leniency error น้อยกว่า มาตรฐานค่าแบบลิเคอร์ท ในขณะที่มาตรฐานค่าแบบลิเคอร์ทเกิด Halo error น้อยกว่า

สำหรับในประเทศไทยมีงานวิจัยที่ได้ศึกษาคุณภาพของมาตรฐานค่าแบบยึดพฤติกรรมในลักษณะดังนี้

สุพักตร์ พิบูลย์ (2534) ได้เปรียบเทียบคุณภาพของมาตรฐานค่าเชิงพฤติกรรมกับมาตรฐานค่าแบบกราฟิก ในด้านความตรง ความเที่ยง และลักษณะความคลาดเคลื่อนในการประเมินค่า ในสถานการณ์การประเมินพฤติกรรมด้านการประเมินผลของครู ผลการวิจัย พบว่า มาตรฐานค่าเชิงพฤติกรรมให้ผลการประเมินที่มีคุณภาพมากกว่ามาตรฐานค่าแบบกราฟิก คือ (1) มีความตรงตามภาวะสันนิษฐาน (Construct validity) สูงกว่าเมื่อตรวจสอบโดยวิธีวิเคราะห์เมตริกซ์สหสัมพันธ์แบบ Multitrait - Multimethod และวิธีตรวจสอบความแตกต่างระหว่างกลุ่มครูที่ปฏิบัติงานมีประสิทธิภาพต่างกัน (2) ให้ค่าความเที่ยงระหว่างผู้ประเมิน (Rater reliability) สูงกว่า และ (3) ปรากฏลักษณะความคลาดเคลื่อนในเชิงปล่อยคะแนน (Leniency error) และความคลาดเคลื่อนอันเนื่องมาจากการเคยมีปฏิสัมพันธ์มาก่อนระหว่างผู้ประเมินกับเป้าหมายของการประเมิน (Halo error) น้อยกว่า

จากที่กล่าวมาจะเห็นได้ว่า มาตรฐานประมาณค่าแบบยึดพฤติกรรมเป็นเครื่องมือที่สามารถช่วยลดความคลาดเคลื่อนในการประมาณค่าได้ ซึ่งจากการศึกษางานวิจัยต่างประเทศจะเห็นว่ามีการนำมาตรฐานประมาณค่าแบบยึดพฤติกรรมมาศึกษาเปรียบเทียบคุณภาพกับมาตรฐานประมาณค่าแบบกราฟิกและมาตรฐานประมาณค่าแบบลิเคอร์ท์ แต่สำหรับงานวิจัยในประเทศไทยนั้นพบว่ามีการศึกษาเปรียบเทียบเฉพาะกับมาตรฐานประมาณค่าแบบกราฟิกเท่านั้น ส่วนมาตรฐานประมาณค่าแบบลิเคอร์ท์นั้นยังไม่ได้มีการนำมาศึกษาเปรียบเทียบและสรุปไว้ ดังนั้นผู้วิจัยจึงสนใจที่จะศึกษาเปรียบเทียบการใช้มาตรฐานประมาณค่าแบบยึดพฤติกรรมและมาตรฐานประมาณค่าแบบลิเคอร์ท์ในบริบทของการประเมิน โดยจะศึกษาเปรียบเทียบด้านความเที่ยงและความคลาดเคลื่อนอย่างเป็นระบบในการประมาณค่าในด้านการประเมินการสอนของครู ทั้งนี้เพื่อให้ได้ข้อสรุปที่ชัดเจนเกี่ยวกับคุณภาพของมาตรฐานประมาณค่าทั้งสองรูปแบบเพื่อเป็นประโยชน์ต่อองค์ความรู้เกี่ยวกับการใช้มาตรฐานประมาณค่าในการประเมินต่อไป

วัตถุประสงค์ของการวิจัย

เพื่อเปรียบเทียบความเที่ยงและความคลาดเคลื่อนอย่างเป็นระบบในการประมาณค่าของมาตรฐานประมาณค่าแบบลิเคอร์ท์และมาตรฐานประมาณค่าแบบยึดพฤติกรรมในการประเมินการสอนของครู

ขอบเขตของการวิจัย

1. การวิจัยครั้งนี้สนใจศึกษาเฉพาะค่าความเที่ยงและความคลาดเคลื่อนอย่างเป็นระบบในการประมาณค่าของการใช้เครื่องมือทั้ง 2 รูปแบบเท่านั้น มิได้สนใจที่จะศึกษาเนื้อหาตามที่ปรากฏในเครื่องมือที่ใช้

2. ประชากรที่ใช้ในการวิจัยครั้งนี้ เป็นครูผู้สอนวิชาคณิตศาสตร์ในระดับชั้นมัธยมศึกษาปีที่ 5 ภาคเรียนที่ 2 ปีการศึกษา 2541 และนักเรียนที่กำลังศึกษาอยู่ระดับชั้นมัธยมศึกษาปีที่ 5 ในโรงเรียนมัธยมศึกษา สังกัดกรมสามัญศึกษา ในเขตกรุงเทพมหานคร

3. สถานการณ์ในการศึกษา ได้แก่

- 1) สถานการณ์การประเมินตนเองของครูผู้สอน
- 2) สถานการณ์การประเมินพฤติกรรมครูผู้สอนโดยนักเรียน

4. ตัวแปรที่ศึกษาในครั้งนี้ ได้แก่

1. ตัวแปรอิสระ คือ รูปแบบของมาตรฐานประมาณค่า ซึ่งแบ่งเป็น 2 รูปแบบ ได้แก่ มาตรฐานประมาณค่าแบบลิเคอร์ท์ และมาตรฐานประมาณค่าแบบยึดพฤติกรรม
2. ตัวแปรตาม คือ ค่าความเที่ยงและความคลาดเคลื่อนอย่างเป็นระบบในการประมาณค่า

5. พฤติกรรมในการศึกษาค้างนี้ คือ พฤติกรรมในการสอนของครู ซึ่งประกอบด้วย พฤติกรรมดังต่อไปนี้ คือ

- 1). พฤติกรรมในขั้นก่อนสอน
- 2). พฤติกรรมในขั้นขณะสอนหรือระหว่างดำเนินการสอน
- 3). พฤติกรรมในขั้นหลังการสอน

สมมติฐานการวิจัย

จากการศึกษาของ Campbell และคณะ (1973) ที่ได้ศึกษาเปรียบเทียบระหว่างมาตรประมาณค่าสองแบบคือ แบบที่ใช้เทคนิคการประมาณค่าแบบรวมสรุป (Summated rating technique) ของ Likert และมาตรประมาณค่าแบบของ Smith และ Kendall ซึ่งกำหนดสเกลด้วยลักษณะพฤติกรรม ผลการวิจัยพบว่า มาตรประมาณค่าแบบยึดพฤติกรรมเกิด Leniency error และ Halo error น้อยกว่ามาตรประมาณค่าแบบรวมสรุปและยังมีความตรงเชิงจำแนก (Discriminant validity) สูงกว่ามาตรประมาณค่าแบบรวมสรุป

และงานวิจัยของ Cook (1988) ที่ทำการศึกษาผลของการใช้มาตรประมาณค่าแบบยึดพฤติกรรม กับมาตรประมาณค่าแบบลิเคอร์ทและผลของการฝึกทักษะแก่ผู้ประเมินต่อค่าความคลาดเคลื่อนเชิงปล่อยคะแนน Leniency error และความคลาดเคลื่อนจากการเคยมีปฏิสัมพันธ์มาก่อน Halo error โดยได้ใช้มาตรประมาณค่าทั้ง 2 รูปแบบ ให้นักศึกษา 2 กลุ่ม คือ กลุ่มที่ได้รับการฝึกทักษะทางด้าน การประเมินกับกลุ่มที่ไม่ได้รับการฝึกทักษะทางด้าน การประเมิน ทำการประเมินพฤติกรรม การสอนของ ศาสตราจารย์ 1 ท่าน ผลการวิจัยพบว่า การฝึกทักษะมีผลต่อการลด Halo error และการประเมินโดย มาตรประมาณค่าแบบยึดพฤติกรรมเกิด Leniency error น้อยกว่ามาตรประมาณค่าแบบลิเคอร์ท ในขณะที่มาตรประมาณค่าแบบลิเคอร์ทเกิด Halo error น้อยกว่ามาตรประมาณค่าแบบยึดพฤติกรรม ดังนั้นในการวิจัยครั้งนี้ ผู้วิจัยจึงตั้งสมมติฐานการวิจัยไว้ดังนี้

- 1.มาตรประมาณค่าแบบยึดพฤติกรรมให้ค่าความเที่ยงระหว่างผู้ประเมินสูงกว่ามาตรประมาณค่าแบบลิเคอร์ท
- 2.มาตรประมาณค่าแบบยึดพฤติกรรมเกิดความคลาดเคลื่อนเชิงปล่อยคะแนน(Leniency error) น้อยกว่ามาตรประมาณค่าแบบลิเคอร์ท
- 3.มาตรประมาณค่าแบบยึดพฤติกรรมเกิดความคลาดเคลื่อนอันเนื่องมาจากการเคยมีปฏิสัมพันธ์มาก่อน(Halo error) น้อยกว่ามาตรประมาณค่าแบบลิเคอร์ท

ข้อตกลงเบื้องต้น

การจัดการเรียนการสอนภายในชั้นเรียนนักเรียนจะเป็นผู้ที่ได้รับผลจากการสอนของครูโดยตรง และเห็นพฤติกรรมการสอนของครูโดยตลอด ดังนั้นจึงถือได้ว่านักเรียนมีความสามารถประเมินครูผู้สอนได้

คำจำกัดความที่ใช้ในการวิจัย

มาตรฐานค่า หมายถึง เครื่องมือที่ช่วยในการตีค่าคุณลักษณะทางจิตวิทยาออกมาเป็นค่าเชิงปริมาณอย่างเป็นระบบ ซึ่งใช้ได้ทั้งการวัดและประเมินพฤติกรรม ผลผลิตและบุคลิกภาพทางสังคมของมนุษย์ โดยมีข้อตกลงว่า มีสเกลเป็นแบบช่วงเท่า ๆ กัน ซึ่งมาตรฐานค่าจะประกอบด้วย 2 ส่วน คือ ส่วนที่เป็นข้อความแสดงพฤติกรรม และส่วนที่เป็นคำตอบแสดงถึงระดับความเข้มของพฤติกรรม

มาตรฐานค่าแบบยึดพฤติกรรม (Behaviorally anchored rating scales) หมายถึง มาตรฐานค่าที่มีข้อความแสดงตัวอย่างพฤติกรรมที่ชัดเจนผูกติดไว้บนทางเลือกแต่ละระดับบนมาตรวัด เพื่อช่วยให้ผู้ประเมินสามารถตัดสินใจได้ โดยการเปรียบเทียบพฤติกรรมการทำงานจริงกับพฤติกรรมตัวอย่างในแต่ละทางเลือก

มาตรฐานค่าแบบลิเคิร์ต (Likert rating scale) หมายถึง มาตรฐานค่าที่มีชุดของคำตอบแสดงระดับความเข้มของพฤติกรรมเป็นแบบคำคุณศัพท์ ได้แก่ มากที่สุด มาก ปานกลาง น้อย น้อยที่สุด

ความเที่ยงระหว่างผู้ประเมิน หมายถึง ความสอดคล้องระหว่างคะแนนของผู้ประเมิน 2 กลุ่ม (Rater reliability) ซึ่งคำนวณโดยการวิเคราะห์สัมประสิทธิ์สหสัมพันธ์ที่เรียกว่า Intraclass Correlation Coefficient (ICC) ตามวิธีการของ Shrout (1979: 420-428)

ความคลาดเคลื่อนอย่างเป็นระบบในการประมาณค่า หมายถึง ลักษณะความคลาดเคลื่อนที่แสดงออกถึงการประมาณค่าเชิงปล่อยหรือกดคะแนน (Leniency error) และความคลาดเคลื่อนอันเนื่องมาจากการเคยมีปฏิสัมพันธ์มาก่อนระหว่างผู้ประเมินกับเป้าหมายของการประเมิน (Halo error)

ความคลาดเคลื่อนเชิงปล่อยหรืออคติคะแนน (Leniency error) หมายถึง สภาพของการประมาณค่าที่ผู้ประเมินมีความเอนเอียงที่จะประเมินสูงกว่าหรือต่ำกว่าที่เป็นจริง ซึ่งตรวจสอบโดยใช้วิธีเปรียบเทียบค่าเฉลี่ยเลขคณิตของผลการประเมินพฤติกรรมเป็นรายด้านระหว่างการใช้มาตรฐานประมาณค่าทั้ง 2 รูปแบบแล้วทำการทดสอบความแตกต่างระหว่างกรณีใช้มาตรฐานประมาณค่าต่างรูปแบบกันด้วย The Wilcoxon Signed Ranks Test

ความคลาดเคลื่อนอันเนื่องมาจากการเคยมีปฏิสัมพันธ์มาก่อน (Halo error) หมายถึง สภาพของการประเมินที่ผู้ประเมินมีความรู้สึกบางอย่างต่อผู้ถูกประเมินทำให้ผลการประเมินผิดจากความเป็นจริง อันเนื่องมาจากความรู้สึกนั้นอาจเป็นได้ทั้งทางบวก และทางลบ ซึ่งตรวจสอบโดยใช้วิธีการเปรียบเทียบค่าเฉลี่ยเลขคณิตของดัชนีบ่งชี้การกระจายของคะแนนหรือคะแนนความคลาดเคลื่อน (Halo score) ระหว่างกรณีใช้มาตรฐานประมาณค่าต่างรูปแบบกัน และทดสอบความแตกต่างในประชากรด้วย t-test กรณีกลุ่มตัวอย่างไม่เป็นอิสระจากกัน (Dependent samples) ซึ่งคำนวณด้วยสูตร Guilford (1978: 158)

ประโยชน์ที่คาดว่าจะได้รับจากการวิจัย

1. ทำให้ทราบว่ามาตรฐานประมาณค่ารูปแบบใดที่มีค่าความเที่ยงมากที่สุดและมีลักษณะความคลาดเคลื่อนน้อยที่สุด ในการประเมินการสอนของครู
2. เพื่อเป็นทางเลือกในการตัดสินใจสำหรับผู้วิจัยหรือผู้ที่จะนำมาตรฐานประมาณค่าไปใช้เป็นเครื่องมือในการประเมินต่อไป

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย