



บทที่ 1

บทนำ

ความเป็นมาและความสำคัญของปัญหา

วิธีการเชิงระบบ (System Approach) เป็นเครื่องมือและกรอบความคิดที่มีคุณค่าอย่างยิ่งต่อการบริหารและการจัดการศึกษาในปัจจุบัน เป็นกระบวนการที่เน้นการวิเคราะห์ปัญหาและกำหนดแนวทางในการแก้ไขปัญหาอย่างมียุทธวิธี (Kaufman, 1972: 2) การดำเนินงานตามขั้นตอนของวิธีการเชิงระบบจะเริ่มด้วยการวิเคราะห์และกำหนดปัญหาของระบบ ต่อจากนั้นก็วิเคราะห์ทางเลือกในการแก้ไขปัญหา คัดเลือกวิธีการหรือทางเลือกที่เห็นว่าเหมาะสมที่สุด กำหนดแนวทางในการดำเนินการ ดำเนินการตามแนวทางที่กำหนดไว้ และประเมินผลการดำเนินงาน โดยในการดำเนินการแต่ละขั้นตอนดังกล่าวเหล่านี้ จะต้องมีการเก็บรวบรวมข้อมูล วิเคราะห์ข้อมูลและใช้ข้อมูลประกอบการตัดสินใจปรับปรุง-เปลี่ยนแปลงการทำงานอยู่ตลอดเวลา (Kaufman, 1968-1970; Corrigan, 1969 อ้างถึงใน Kaufman, 1972: 7) จนกระทั่ง Huse (1979: 112) ได้เรียกชื่อกระบวนการนี้ในอีกชื่อหนึ่งว่า กระบวนการตัดสินใจ หรือวิธีการเชิงเหตุผล (The Decision-Making Process or The Rationale Approach)

ในการดำเนินงานโดยทั่วไปที่ยึดถือวิธีการเชิงระบบเป็นกรอบปฏิบัติ จะพบว่ากิจกรรมการประเมิน (Evaluation Activities) ซึ่งได้รับการนิยามว่า "คือการเก็บรวบรวมข้อมูลและใช้ข้อมูลเพื่อประกอบการตัดสินใจ" (Cronbach, 1983: 101) เป็นกิจกรรมที่มีบทบาทสำคัญยิ่งและมีส่วนเกี่ยวข้องกับการดำเนินงานในเกือบทุกขั้นตอน นับตั้งแต่ขั้นของการวิเคราะห์และกำหนดปัญหาของระบบ หรือที่เรียกว่า การประเมินความต้องการจำเป็น (Needs Assessment) จนถึงขั้นของการประเมินสรุปผลการดำเนินงาน

ในทางการศึกษา ซึ่งได้พยายามประยุกต์แนวคิดเชิงระบบเพื่อการบริหาร กิจกรรมการประเมินก็มีบทบาทต่อการตัดสินใจเป็นอย่างมากเช่นกัน Gephart (n.d. : 2) ได้สรุปว่าจุดมุ่งหมายของการประเมินในทางการศึกษานั้นมีมากมายหลายประการ เช่น เพื่อตัดสินคุณค่าของ

โปรแกรมการเรียนการสอนที่โรงเรียนได้ดำเนินการไปแล้ว เพื่อช่วยในการวางแผนการสอน เพื่อชี้แจงหรือวิเคราะห์เชิงเปรียบเทียบคุณภาพของวิธีการใหม่ ๆ ทางการศึกษา เพื่อพัฒนาสื่อการสอนหรือวิธีสอนรูปแบบต่าง ๆ เป็นต้น

หากพิจารณาการบริหารในอีกบทบาทหนึ่ง คือมองในแง่ของการบริหารงานบุคคลซึ่งถือได้ว่าเป็นเงื่อนไขที่สำคัญยิ่งสำหรับการอยู่รอดหรือความเจริญก้าวหน้าขององค์การในปัจจุบัน (Crane, 1985: 4) กิจกรรมการประเมินก็ยังคงได้รับการยอมรับว่าเป็นสิ่งจำเป็นอย่างยิ่งที่จะต้องดำเนินการและเป็นกิจกรรมที่เป็นประโยชน์อย่างมากต่อการดำเนินงานที่เกี่ยวข้องกับการบริหารงานบุคคลในหลายลักษณะ อาทิ เพื่อตอบแทนในผลการดำเนินงาน (Compensation) เพื่อจัดวางตัวบุคคล-เลื่อนตำแหน่ง-ปรับเปลี่ยนหน้าที่การงาน (Placement) หรือ เพื่อการฝึกอบรม และพัฒนาความสามารถของบุคลากร (Training and Development) เป็นต้น (Mathis, 1979: 290-291)

ภายใต้ความสำคัญของ "การประเมิน" ดังที่ได้กล่าวมาแล้วข้างต้น ประเด็นปัญหาประการหนึ่งที่นักวัดผล นักประเมิน หรือบุคคลที่ทำหน้าที่ในการตัดสินใจต้องกังวลใจหรือให้ความสนใจเป็นอย่างยิ่งก็คือ ในเรื่องของความถูกต้อง ความน่าเชื่อถือ หรือความเชื่อถือได้ของผลการประเมิน ซึ่งจะมีความถูกต้อง หรือเชื่อถือได้มากน้อยเพียงใดนั้น ส่วนหนึ่งก็ขึ้นอยู่กับเทคนิควิธีหรือเครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูล

จากอดีตจนถึงปัจจุบัน นักวัดผล นักประเมิน หรือนักวิจัย ได้พยายามพัฒนาเทคนิควิธีในการเก็บรวบรวมข้อมูล เสนอแนะวิธีการ ตลอดจนรูปแบบในการประเมินไว้อย่างหลากหลาย อย่างไรก็ตาม ในทางปฏิบัติ อาจเป็นเพราะข้อจำกัดในเรื่องของทรัพยากรสำหรับดำเนินการ ประกอบกับภาวะการที่นักบริหารหรือผู้ทำหน้าที่ตัดสินใจต้องการใช้ข้อมูลเพื่อการตัดสินใจอย่างรวดเร็ว ทำให้การใช้เทคนิควิธีที่มีความซับซ้อน หรือต้องลงทุนมากไม่สามารถนำมาใช้ได้ จึงเปิดโอกาสให้วิธีการสำรวจ โดยเฉพาะอย่างยิ่งการสำรวจด้วยแบบสอบถาม (Questionnaire Survey) ได้เข้ามามีบทบาทต่อการตัดสินใจมากยิ่งขึ้น ทั้งนี้ด้วยเงื่อนไขที่ว่า สะดวกต่อการเก็บรวบรวมข้อมูลง่ายต่อการประมวลผล และให้ข้อสรุปอย่างรวดเร็วทันต่อการตัดสินใจ

การสำรวจด้วยแบบสอบถามในกิจกรรมการประเมินได้ปรากฏให้เห็นกันโดยทั่วไปและมีแนวโน้มว่าจะมีการใช้ต่อไปเรื่อย ๆ และในบรรดาแบบสอบถามที่ใช้กันอยู่ทั่วไปนั้นพบว่าแบบสอบถามประเภทมาตรประมาณค่า (Rating Scales) เป็นแบบที่ได้รับความนิยมและใช้กันอย่างแพร่หลายมากที่สุด (Guilford, 1954: 263, Turney and Robb, 1971: 36) เฉพาะในประเทศสหรัฐอเมริกา Monahan (1975: 421) กล่าวว่าได้มีการใช้กันอย่างแพร่หลายเสมือนการระบาดของวิธีการวัด (Measurement Myopia) หรือในกรณีของการบริหารงานบุคคล Mathis (1979: 301-302) ได้สรุปไว้ว่าเครื่องมือประมาณค่าประเภทกำหนดรายการไว้อย่างเป็นลำดับชั้น โดยเฉพาะอย่างยิ่งมาตรประมาณค่าแบบกราฟิก (Graphic Rating Scales) เป็นแบบที่ได้รับความนิยมกันอย่างแพร่หลายมากที่สุด สำหรับในประเทศไทย มัลลิกา ยุวณะเดมิย์ (2527: 2) ได้ทำการสำรวจเครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูลในการทำวิทยานิพนธ์ของนิสิตบัณฑิตวิทยาลัย คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ที่ได้ดำเนินการระหว่างปีการศึกษา 2521-2525 โดยวิธีการสุ่มอย่างง่ายจากบัตรรายการจำนวน 50 บัตร พบว่าได้ใช้แบบสอบถามชนิดมาตรประมาณค่าจำนวน 34 ฉบับ คิดเป็นร้อยละ 68 ของจำนวนที่สุ่มมาศึกษา ซึ่งถือว่าเป็นสัดส่วนที่สูงมากพอสมควร อีกทั้งจากการที่ผู้วิจัยได้ทำการศึกษางานวิจัยกลุ่มหนึ่งที่เน้นการศึกษาเชิงประยุกต์แนวความคิดในการประเมินความต้องการจำเป็นทางการศึกษา (Educational Needs Assessment) ที่ได้ดำเนินการระหว่างปี พ.ศ. 2525 ถึง 2531 ประมาณ 20 เรื่อง พบว่า ทุกเรื่องได้ใช้แบบสอบถามประเภทมาตรประมาณค่าเป็นเครื่องมือในการเก็บรวบรวมข้อมูล และลักษณะของมาตรประมาณค่าที่ใช้ส่วนใหญ่เป็นแบบตัวเลข (Numerical Rating Scales) หรือแบบกราฟิก (Graphic Rating Scales)

จากการศึกษางานวิจัยเชิงประเมินความต้องการจำเป็นที่นิยมใช้มาตรประมาณค่าเป็นเครื่องมือในการเก็บรวบรวมข้อมูล ผู้วิจัยได้พบข้อสังเกตที่สำคัญประการหนึ่งคือ ข้อสรุปจากงานวิจัยเหล่านั้น ไม่สามารถจำแนกรายการ หรือชี้บ่งความแตกต่างระหว่างรายการทั้งหลายที่ให้ประเมิน กล่าวคือ ในกรณีที่การศึกษาครั้งนั้นเป็นการศึกษาปัญหาในเรื่องใดเรื่องหนึ่ง มักจะปรากฏผลการประเมินโดยครู หรือผู้บริหารว่า เป็นปัญหาในระดับน้อย หรือน้อยมาก เกือบทุกรายการทุกด้านที่ให้ประเมิน ในกรณีที่มุ่งประเมินความต้องการเพื่อวางแผนพัฒนาในเรื่องใดเรื่องหนึ่ง มักจะปรากฏผลว่ามีความต้องการในระดับมาก หรือมากที่สุด เกือบทุกรายการ ทุกด้านที่ให้ประเมิน

หรือ ในกรณีที่ให้ประเมินสภาพการปฏิบัติงานที่เป็นจริง และสภาพการปฏิบัติงานที่ควรจะเป็น ในเรื่องใดเรื่องหนึ่ง ก็มักจะปรากฏผลการประเมินในลักษณะที่ไม่สามารถจำแนกความแตกต่างของ รายการที่ให้ประเมินเช่นกัน กล่าวคือ ในกรณีที่ให้ประเมินสภาพการปฏิบัติงานจริง มักจะพบว่าผลการประเมินในเกือบทุกรายการ ทุกด้าน ปรากฏคะแนนในระดับปานกลาง หรือค่อนข้างมากอย่างสม่ำเสมอ ส่วนการประเมินพฤติกรรมการปฏิบัติงานที่ควรจะเป็น ผลการประเมินจะปรากฏคะแนนในระดับมาก หรือมากที่สุดอย่างสม่ำเสมอเช่นกัน ซึ่งการที่ผลการวิจัยเชิงประเมินในสถานการณ์ที่แตกต่างกัน แต่ปรากฏภาพรวมผลการประเมินในลักษณะที่คล้ายคลึงกันเช่นนี้ น่าจะสรุปได้ว่าส่วนหนึ่งเป็นผลสืบเนื่องมาจากการเกิดความคลาดเคลื่อนในขั้นของการวัด หรือเครื่องมือวัดไม่สามารถจำแนกและตรวจสอบความแตกต่างระหว่างรายการที่ให้ประเมินได้ ทั้งนี้เพราะว่า โดยลักษณะของความน่าจะเป็นแล้ว ในการประเมินเป้าหมายใดเป้าหมายหนึ่งในแง่มุมที่ต่างกัน หรือประเมินในหลาย ๆ ด้าน หลายมิติ ไม่น่าจะเป็นไปได้ที่ผลการประเมินจะปรากฏคะแนนเฉลี่ยในลักษณะใกล้เคียงกัน อย่างสม่ำเสมอ ทุกครั้ง ทุกมิติที่ให้ประเมิน

การเกิดความคลาดเคลื่อนในขั้นของการวัดอันเนื่องมาจากการใช้เทคนิคการประมาณค่าแบบตัวเลข หรือแบบกราฟิก อาจเกิดจากแหล่งความคลาดเคลื่อนหลายแหล่ง แหล่งของความคลาดเคลื่อนที่อาจเป็นไปได้ เช่น เกิดจากลักษณะธรรมชาติของมาตรวัด เกิดจากธรรมชาติของการตอบของกลุ่มตัวอย่างประชากร เกิดจากความร่วมมือในการตอบ หรือเกิดจากลักษณะเนื้อหาที่ให้ประเมิน เป็นต้น ในส่วนของความคลาดเคลื่อนอันเนื่องมาจากลักษณะของมาตรวัดนั้น อาจเป็นไปได้ว่าในกรณีของการใช้เทคนิคการประมาณค่าแบบตัวเลข หรือแบบกราฟิก ซึ่งเป็นแบบที่นิยมใช้กันอย่างแพร่หลายในการประเมินสภาพปัญหา ความต้องการ สภาพการปฏิบัติงานจริง หรือสภาพการปฏิบัติงานที่ควรจะเป็น ในเรื่องใดเรื่องหนึ่งนั้น ผู้ประเมินอาจเกิดความสับสนในแง่ของการเปรียบเทียบในเชิงคุณภาพของสิ่งที่ให้ประเมิน ทำให้ยากแก่การประเมินได้อย่างคมชัด (Tobin, H.M. and Others, 1979: 171-172) หรืออาจเกิดจากจุดอ่อนบางประการของเทคนิคนี้ เช่น ในกรณีที่ขาดเกณฑ์ในการตัดสินใจที่ชัดเจน ผู้ประเมินมีแนวโน้มที่จะหลีกเลี่ยงการประเมินในสเกลสูงสุด หรือต่ำสุด และมักประเมินในระดับกลาง ๆ (Guilford, 1954: 265) หรือในแง่ของการใช้คุณศัพท์เชิงปริมาณเพียงอย่างเดียวในกรณีของการประมาณค่าแบบตัวเลข หรือแบบกราฟิกนั้น คำว่า "ดี" "ดีเยี่ยม" "น่าพอใจ" ฯลฯ เป็นคำที่ก่อให้เกิดความแตกต่างใน

การตีความหมายของผู้ประเมินแต่ละคน คำว่า "ดี" ของคนหนึ่งอาจจะเป็นเพียงระดับพอใช้ของอีกคนหนึ่ง (Mathis, 1979: 295) ในส่วนความคลาดเคลื่อนที่เกิดจากธรรมชาติ หรือลักษณะของผู้ตอบนั้น ภายใต้สถานการณ์ที่มาตรประเมินขาดเกณฑ์ในการตัดสินใจที่ชัดเจนดังที่ปรากฏในมาตรประมาณค่าแบบตัวเลข หรือมาตรแบบกราฟิกนั้น ส่งผลให้ผู้ประเมินแต่ละคนไม่สามารถตัดสินใจได้อย่างชัดเจน ผู้ประเมินบางคนมีแนวโน้มปล่อยคะแนน (Easy Rater) ก็จะประเมินในสเกลสูง ๆ ในขณะที่ผู้ประเมินบางคนมีแนวโน้มจะกดคะแนน (Hard Rater) ก็จะประเมินในระดับคะแนนต่ำ ๆ (Guilford, 1954: 278) อีกทั้งในกรณีที่ขาดเกณฑ์ในการตัดสินใจเช่นนี้ การประเมินเป้าหมายใด ๆ ในเรื่องใดเรื่องหนึ่ง ข้อสรุปจะขึ้นอยู่กับความรู้ของผู้ให้ข้อมูล ค่านิยม ตลอดจนวัฒนธรรมของชุมชนด้วย (Siegel, Attkisson and Carson 1978: 20) รวมทั้งเพศ ศาสนา เชื้อชาติ ก็อาจส่งผลต่อคุณภาพผลการประเมินเช่นกัน (Mathis 1979: 295) สำหรับในกรณีของเนื้อหาสาระที่ให้ประเมินนั้น คุณลักษณะบางอย่างเป็นลักษณะที่แฝงอยู่ภายในตัวบุคคลที่มีโอกาสเกิดความขัดแย้งในการตัดสิน เช่น ความเมตตา ความกล้าหาญ ในขณะที่ลักษณะแฝงบางประการสามารถตัดสินได้อย่างสอดคล้อง เช่น ความสามารถ ความคิดริเริ่ม เป็นต้น (Guilford 1954: 296)

ในภาวะที่กิจกรรมการประเมินได้กลายเป็นเงื่อนไขที่จำเป็นสำหรับการบริการงานที่มีประสิทธิภาพ ประกอบกับจากการวิเคราะห์สถานการณ์ปัจจุบันของกิจกรรมการประเมิน ซึ่งพบว่าวิธีการสำรวจด้วยแบบสอบถาม โดยเฉพาะอย่างยิ่งแบบสอบถามประเภทมาตรประมาณค่าแบบตัวเลข หรือแบบกราฟิก ยังเป็นที่นิยมใช้กันอย่างแพร่หลายทั่วไปในกิจกรรมการประเมิน และยังคงมีแนวโน้มที่จะใช้กันอย่างหลากหลายต่อไปเรื่อย ๆ ทั้ง ๆ ที่โดยลักษณะธรรมชาติของมาตรวัดรูปแบบนี้และข้อสรุปในงานวิจัยที่ใช้เทคนิคนี้ในการเก็บรวบรวมข้อมูลได้ปรากฏลักษณะที่เป็นจุดอ่อนหลายประการดังที่ได้นำเสนอไปแล้วนั้น ในภาวะการณ์เช่นนี้ผู้วิจัยเห็นว่า การศึกษาวิจัยเพื่อพัฒนาเทคนิควิธีการประมาณค่า หรือการศึกษาในเชิงตรวจสอบคุณภาพของเครื่องมือที่นิยมใช้กันโดยทั่วไป น่าจะเป็นสิ่งจำเป็นและมีคุณค่ายิ่ง โดยในการศึกษาวิจัยดังกล่าว จะต้องเน้นการเปรียบเทียบคุณภาพของเครื่องมือหรือเทคนิควิธีที่พัฒนาขึ้นกับเทคนิควิธีอันเป็นรูปแบบที่นิยมใช้ในเชิงประจักษ์อย่างชัดเจน เพื่อผู้ปฏิบัติงานด้านการประเมินสามารถมองเห็นจุดเด่น จุดด้อยของเครื่องมือหรือเทคนิควิธี ทำให้อาจสามารถตัดสินใจเลือกใช้เทคนิควิธีในการวัดได้อย่างเหมาะสม

จากประเด็นความคิดที่เห็นว่า ควรจะมีการพัฒนาเทคนิควิธีในการประมาณค่าให้มีคุณภาพมากยิ่งขึ้น ผู้วิจัยจึงได้ศึกษาวิเคราะห์ข้อเสนอนั้นของนักวัดผลหลายท่านที่ได้พยายามให้ข้อเสนอแนะเพื่อการปรับปรุงคุณภาพของการประมาณค่า อาทิ Thorndike และ Hagen (1977: 487) ได้เสนอแนะว่า ในการประเมินนั้นควรทำการประเมินเฉพาะลักษณะที่ปรากฏให้เห็นชัดเจน ลักษณะที่จะประเมินจะต้องแสดงออกมาให้เห็นในรูปของพฤติกรรม ฯลฯ หรือ Oppenheim (1977: 86-87) เสนอแนะว่าในการสร้างมาตราประมาณค่า นั้น จะต้องกำหนดขอบเขตและทิศทางของลักษณะที่จะประเมินให้เป็นที่เข้าใจตรงกันเสียก่อน จากนั้นจึงค่อยกำหนดระดับในการประมาณค่าและกำหนดความหมายในแต่ละระดับ และเพื่อให้ลักษณะที่จะประเมินมีความชัดเจน จะต้องอธิบายลักษณะนั้น ๆ ไว้อย่างเฉพาะเจาะจง เป็นต้น ในขณะที่เดียวกันได้ศึกษาแนวความคิดในการพัฒนามาตราประมาณค่าเชิงพฤติกรรม (Behaviorally - Anchored Rating Scales) ซึ่งพบว่า ลักษณะของมาตรวัดตามแนวคิดนี้สอดคล้องกับเงื่อนไขหรือข้อเสนอสำหรับการปรับปรุงการประมาณค่าให้มีความคมชัดหลายประการ เป็นวิธีการใหม่ในการประเมินพฤติกรรมการทำงาน (Mathis, 1979: 305) เป็นวิธีการเชิงพัฒนา (Developmental Approach to Appraisal) (Crane 1985: 408-409) ที่แฝงไว้ด้วยช่องทางในการพัฒนาตนเองของบุคลากรที่ได้รับการประเมิน และได้ปรากฏข้อมูลยืนยันเกี่ยวกับคุณภาพของมาตรวัดตามแนวคิดนี้ในงานวิจัยในต่างประเทศหลายเรื่อง หลายลักษณะ ผู้วิจัยจึงเห็นว่ามาตราประมาณค่าเชิงพฤติกรรม น่าจะเป็นเทคนิคทางเลือกที่น่าสนใจในรูปแบบหนึ่งสมควรได้รับการพัฒนาและตรวจสอบยืนยันคุณภาพในเชิงประจักษ์ในภาวะแวดล้อมการประเมินของไทย

ด้วยแนวคิดที่จะพัฒนาและศึกษาคุณภาพของมาตราประมาณค่าเชิงพฤติกรรมในภาวะแวดล้อมการประเมินของคนไทย ดังกล่าวข้างต้น ผู้วิจัยจึงได้กำหนดกรอบความคิดในการวิจัยครั้งนี้ขึ้น โดยมุ่งเน้นการพัฒนาและตรวจสอบคุณภาพของมาตราประมาณค่ารูปแบบนี้ เปรียบเทียบกับมาตราประมาณค่าแบบกราฟิก อันเป็นรูปแบบหนึ่งของมาตราประมาณค่าที่ได้รับความนิยมใช้กันอย่างแพร่หลายและเคยได้รับการตรวจสอบยืนยันคุณภาพในด้านความเที่ยง (Reliability) ในภาวะแวดล้อมการประเมินของไทยมาแล้วว่าให้ค่าความเที่ยงสูงที่สุด เมื่อเทียบกับรูปแบบอื่น ๆ (มัลลิกา ยวนะเทมิย์, 2527) โดยในการเปรียบเทียบคุณภาพของมาตราประมาณค่าครั้งนี้ได้เน้นการเปรียบเทียบด้วยดัชนีชี้คุณภาพหลายรายการทั้งในด้าน ความตรง ความเที่ยง และลักษณะความคลาดเคลื่อนในการประมาณค่า

และได้เลือกสถานการณ์การประเมินพฤติกรรมการทำงานของคุณ (Performance Assessment) เป็นสถานการณ์เงื่อนไขในการตรวจสอบคุณภาพของมาตรฐานค่าครั้งนี้ ทั้งนี้ด้วยเห็นว่าการประเมินพฤติกรรมการทำงานเป็นกิจกรรมที่สำคัญยิ่งประการหนึ่งในกระบวนการบริหารงานบุคคลที่หน่วยงานต่าง ๆ ได้ให้ความสนใจและดำเนินการเพื่อตอบสนองผลการดำเนินงาน เลื่อนระดับ ปรับปรุง เปลี่ยนแปลงตำแหน่ง หรือเพื่อวิเคราะห์ความต้องการจำเป็นในการพัฒนาบุคลากร (Training Needs / Staff Development Needs) รวมทั้งการศึกษาเชิงพัฒนาหรือตรวจสอบคุณภาพของเครื่องมือในการประเมินพฤติกรรมการทำงานของคุณในลักษณะเช่นนี้ยังจะเป็นประโยชน์สำหรับกิจกรรมการประเมินในลักษณะของการติดตามผลการฝึกอบรม หรือติดตามผลผู้ที่ผ่านการศึกษาในหลักสูตรต่าง ๆ ได้อีกด้วย ภายใต้กรอบความคิดในการวิจัยที่ถูกกำหนดขึ้นในการวิจัยนี้ จะทำให้สามารถชี้บ่งเทคนิควิธีในการวัดที่มีคุณภาพที่สามารถประเมินพฤติกรรมการทำงานได้อย่างแม่นยำ สามารถอธิบายลักษณะความคลาดเคลื่อนในการวัดลักษณะต่าง ๆ ได้ อันจะเป็นประโยชน์อย่างยิ่งต่อการเลือกใช้เทคนิควิธี ตลอดจนการตีความหมายหรือแปลความหมาย ข้อมูลผลการประเมินได้อย่างถูกต้องมากยิ่งขึ้นในโอกาสต่อไป

วัตถุประสงค์ของการวิจัย

เพื่อเปรียบเทียบคุณภาพของมาตรฐานค่าเชิงพฤติกรรม กับ มาตรฐานค่าแบบกราฟิก ในด้านความตรง ความเที่ยง และลักษณะความคลาดเคลื่อนในการประเมินค่า ภายใต้สถานการณ์การประเมินพฤติกรรมด้านการประเมินผลของครู

ขอบเขตของการวิจัย

1. เป้าหมายของการประเมิน

1.1 พฤติกรรมเป้าหมายในการศึกษาครั้งนี้ คือ พฤติกรรมการทำงานด้านการประเมินผลของครู ซึ่งครอบคลุมลักษณะการทำงาน 4 ด้านย่อย คือ การวางแผนการวัดผล-ประเมินผล การวัด-ประเมินผลในชั้นสอน การวัด-ประเมินผลในชั้นสรุปผลการเรียน และการตรวจสอบและพัฒนาระบบการวัด-ประเมินผล

1.2 กลุ่มบุคคลที่เป็นเป้าหมายของการประเมินเพื่อการตรวจสอบคุณภาพของ มาตรฐาน (Target Group for Scale Evaluation) ในการศึกษาครั้งนี้ คือ ครูคณิตศาสตร์ ที่สอนในระดับชั้นมัธยมศึกษาตอนต้น ในภาคต้น ปีการศึกษา 2533 ในโรงเรียนมัธยมศึกษา สังกัดกรมสามัญศึกษา ในเขตกรุงเทพมหานคร

2. เงื่อนไขหรือสถานการณ์การประเมิน

การศึกษาครั้งนี้ จะเน้นการวิเคราะห์คุณภาพมาตรฐานค่าในสถานการณ์การ ประเมิน 2 สถานการณ์ คือ (1) สถานการณ์ของการประเมินตนเองของครูผู้สอน และ (2) สถานการณ์การประเมินพฤติกรรมของครูโดยหัวหน้าหมวดวิชา

3. ตัวแปรในการวิจัย

3.1 ตัวแปรอิสระ

ตัวแปรอิสระในการศึกษาครั้งนี้ คือ รูปแบบของมาตรฐานค่า มี 2 รูปแบบ คือ (1) มาตรฐานค่าเชิงพฤติกรรม (Behaviorally-Anchored Rating Scales) และ (2) มาตรฐานค่าแบบกราฟิก (Graphic Rating Scales)

3.2 ตัวแปรตาม

ตัวแปรตามในการศึกษาครั้งนี้ คือ คุณภาพของมาตรฐานค่า ในด้านต่อไปนี้

3.2.1 ความตรงในการประมาณค่า เป็นการพิจารณาความตรงตามภาวะ สัมพันธภาพภายใต้เงื่อนไขการตรวจสอบ 2 รายการ คือ (1) ตรวจสอบความตรงชนิด Convergent and Discriminant Validity โดยการวิเคราะห์เมตริกซ์สหสัมพันธ์แบบ Multitrait-Multimethod ตามหลักการของ Campbell และ Fiske (1959) และ (2) ตรวจสอบความตรง ของมาตรฐานค่าในการจำแนกความแตกต่างระหว่างกลุ่มครูที่ปฏิบัติงานมีประสิทธิภาพต่างกัน

3.2.2 ความเที่ยงในการประมาณค่า เป็นการพิจารณาจากความสอดคล้อง ระหว่างคะแนนของผู้ประเมิน 2 กลุ่ม คือ ครูผู้สอน กับหัวหน้าหมวดวิชา ถือเป็นความเที่ยงระหว่าง ผู้ประเมิน (Rater Reliability)

3.2.3 ลักษณะความคลาดเคลื่อนในการประมาณค่า เป็นการพิจารณาลักษณะ ความคลาดเคลื่อนอย่างเป็นระบบ 2 ลักษณะ คือ (1) ความคลาดเคลื่อนในเชิงปล่อยคะแนนของ ผู้ประเมิน (Leniency Error) และ (2) ความคลาดเคลื่อนอันเนื่องมาจากการเคยมีปฏิสัมพันธ์ มาก่อนระหว่างผู้ประเมินกับเป้าหมายของการประเมิน (Halo Error)

สมมติฐานการวิจัย

จากข้อเสนอแนะของ Thorndike และ Hagen (1977) ที่เสนอแนะว่าในการปรับปรุงแก้ไขการประมาณค่าให้ได้ผลถูกต้องนั้น "จะต้องจัดทำรายการที่จะประเมินให้เห็นเป็นรูปธรรมที่ชัดเจน ทำให้ผู้ประเมินสามารถตัดสินใจได้โดยพิจารณาจากรายการพฤติกรรมที่กำหนดให้" หรือตามข้อเสนอแนะของ Oppenheim (1966) ที่กล่าวว่า "ในการสร้างมาตรประมาณค่า นั้น จะต้องกำหนดขอบเขตและทิศทางของลักษณะที่จะประเมินให้เป็นที่เข้าใจตรงกันเสียก่อน จากนั้นจึงค่อยกำหนดระดับในการประมาณค่า" รวมทั้งจากการศึกษางานวิจัยที่มุ่งพัฒนาและศึกษาคุณภาพของมาตรประมาณค่าเชิงพฤติกรรมในสถานการณ์การประเมินต่าง ๆ ในต่างประเทศบางลักษณะ คือ Campbell (1973) ได้พบว่ามาตรประเมินค่าเชิงพฤติกรรมเกิดความคลาดเคลื่อนในการวัดน้อยกว่ามาตรประมาณค่าชนิดที่ไม่มีตัวอย่างพฤติกรรมผูกติดบนมาตรวัด และมีความตรงเชิงจำแนกสูงกว่าด้วย และ Keaveny และ McGinn (1975) ก็ได้พบข้อสรุปในเชิงยืนยันคุณภาพของมาตรประมาณค่าเชิงพฤติกรรมในการใช้ประเมินพฤติกรรมการสอนของครู 4 คน ว่า เกิดความคลาดเคลื่อนในการวัดน้อยกว่ามาตรประมาณค่าแบบกราฟิก และมีความตรงเชิงจำแนกสูงกว่าด้วย เช่นกัน ดังนั้น ในการศึกษาครั้งนี้ ผู้วิจัยจึงได้กำหนดสมมติฐานของการวิจัยดังนี้

1. มาตรประมาณค่าเชิงพฤติกรรม มีความตรงตามภาวะสัมพันธฐาน สูงกว่ามาตรประมาณค่าแบบกราฟิก
2. มาตรประมาณค่าเชิงพฤติกรรมให้ค่าความเที่ยงระหว่างผู้ประเมิน สูงกว่ามาตรประมาณค่าแบบกราฟิก
3. ผลการประเมินโดยมาตรประมาณค่าเชิงพฤติกรรม ปรากฏลักษณะความคลาดเคลื่อนในการประมาณค่า น้อยกว่ามาตรประมาณค่าแบบกราฟิก

ข้อตกลงเบื้องต้น

ในการปฏิบัติงานด้านการเรียนการสอนในโรงเรียน หัวหน้าหมวดวิชาเป็นผู้ทำหน้าที่กำกับ ดูแล และให้การนิเทศ แนะนำแก่ครูผู้สอน ดังนั้นถือว่าเป็นผู้ที่รู้เห็นพฤติกรรมการทำงาน ของครูผู้สอนเป็นอย่างดี และสามารถจำแนกได้ว่าครูคนใดเป็นผู้ปฏิบัติงานเป็นระบบ-มีประสิทธิภาพ หรือครูคนใดปฏิบัติงานยังไม่เป็นระบบ-ขาดประสิทธิภาพ

นิยามศัพท์ที่เกี่ยวข้อง

มาตรฐานค่าเชิงพฤติกรรม
(Behaviorally-Anchored
Rating Scales)

เป็นรูปแบบหนึ่งของมาตรฐานค่า มีลักษณะเด่น
ที่ใช้ตัวอย่างพฤติกรรมที่ชัดเจน ผู้กติดไว้บนทางเลือกแต่ละ
ระดับบนมาตรวัด เพื่อช่วยให้ผู้ประเมินสามารถตัดสินใจได้
โดยการเปรียบเทียบพฤติกรรมการทำงานจริงกับพฤติกรรม
ตัวอย่างในแต่ละทางเลือก

มาตรฐานค่าแบบกราฟิก
(Graphic Rating Scales)

หมายถึง รูปแบบหนึ่งของมาตรฐานค่าที่ใช้
คำคุณศัพท์เชิงคุณภาพผูกติดไว้ที่หัวท้ายของเส้นกราฟ เช่น
เหมาะสม---ไม่เหมาะสม เป็นระบบ---ไม่เป็นระบบ
เป็นต้น

คุณภาพของการประมาณค่า

หมายถึงคุณสมบัติของมาตรวัด ซึ่งในที่นี้คือ มาตรฐาน
ประมาณค่าเชิงพฤติกรรม และมาตรฐานค่าแบบกราฟิก
ทั้งในด้านความตรง (Validity) ความเที่ยง
(Reliability) และความคลาดเคลื่อนในการประมาณค่า
(Error)

ความตรงในการประมาณค่า

ในที่นี้หมายถึงความตรงตามภาวะสันนิษฐานของ
มาตรวัด 2 รูปแบบ ภายใต้เงื่อนไขการตรวจสอบ 2 รายการ
คือ (1) Convergent & Discriminant Validity
ตรวจสอบด้วยการวิเคราะห์เมตริกซ์สหสัมพันธ์ตามแนวทาง
ของวิธีการ Multitrait-Multimethod ซึ่งเสนอโดย
Campbell และ Fiske (1959) และ (2) ความตรง
ในการจำแนกความแตกต่างระหว่างกลุ่มครูที่ปฏิบัติงานมี
ประสิทธิภาพต่างกัน ตรวจสอบโดยการเปรียบเทียบผลการ
ประเมินพฤติกรรมตนเองของครูผู้สอนกลุ่มปฏิบัติงานเป็น
ระบบ กับกลุ่มที่ปฏิบัติงานยังไม่เป็นระบบ

ความเที่ยงในการประมาณค่า

หมายถึง ความสอดคล้องระหว่างคะแนนของผู้ประเมิน 2 กลุ่ม (Rater Reliability) คำนวณโดยการวิเคราะห์สัมประสิทธิ์สหสัมพันธ์ที่เรียกว่า Intraclass Correlation Coefficient (ICC) ตามวิธีการของ Shrout (1979: 420-428)

ลักษณะความคลาดเคลื่อน

ในการประมาณค่า

หมายถึงลักษณะความคลาดเคลื่อนที่แสดงออกซึ่งการปล่อยคะแนนของผู้ประเมิน (Leniency Error) และความคลาดเคลื่อนอันเนื่องมาจากการเคยมีปฏิสัมพันธ์มาก่อนระหว่างผู้ประเมินกับเป้าหมายของการประเมิน (Halo Error)

ความคลาดเคลื่อนเชิงปล่อย
คะแนน (Leniency Error)

ในที่นี้คือแนวโน้มของการปล่อยคะแนนในเชิงเปรียบเทียบระหว่างผลการประเมินจากมาตรฐานประมาณค่า 2 รูปแบบ

ความคลาดเคลื่อนอันเนื่อง
มาจากการเคยมีปฏิสัมพันธ์มาก่อน
ระหว่างผู้ประเมินกับเป้าหมายของ
การประเมิน (Halo Error)

เป็นสถานการณ์ที่ผู้ประเมินมีอคติในการประเมิน ถ้ามีแนวโน้มพิจารณารายการพฤติกรรมด้านใดด้านหนึ่งของเป้าหมายที่กำลังประเมินในทางบวกก็มักจะประเมินรายการอื่น ๆ ในทางบวกคล้ายตามกันหรือถ้าพิจารณาเป้าหมายที่ให้ประเมินในทางลบ ผลการประเมินพฤติกรรมการปฏิบัติงานก็มีแนวโน้มทางลบ หรือทางคะแนนต่ำอย่างคล้องตามกัน เฉพาะในกรณีของการประเมินตนเองในการวิจัยนี้ ถือว่าอคติเช่นนี้ก็เกิดขึ้นได้ ซึ่งอาจเนื่องมาจากความรู้สึกภาคภูมิใจในตนเอง หรือมองตนเองว่ายังไม่ประสบความสำเร็จเท่าที่ควร ภาพสะท้อนของผลการประเมินที่แสดงออกซึ่งการเกิดความคลาดเคลื่อน ในลักษณะนี้คือ การกระจายของคะแนนระหว่างรายการพฤติกรรมที่ผู้ประเมินมีค่าต่ำ

กลุ่มครูที่ปฏิบัติงาน
มีประสิทธิภาพต่างกัน

หมายถึง กลุ่มครู ที่หัวหน้าหมวดวิชา และ
ศึกษานิเทศก์ระบุได้อย่างสอดคล้องว่าเป็นผู้ปฏิบัติงาน
เป็นระบบ-มีประสิทธิภาพและกลุ่มครูที่หัวหน้าหมวดระบุว่า
เป็นกลุ่มที่ปฏิบัติงานยังไม่เป็นระบบควรได้รับการพัฒนา
กลุ่มครูทั้งสองกลุ่มนี้ เป็นกลุ่มเป้าหมายสำหรับการตรวจสอบ
ความตรงของมาตรวัดในการจำแนกความแตกต่างระหว่าง
กลุ่ม

ประโยชน์ที่คาดว่าจะได้รับ

การศึกษาครั้งนี้ เป็นการขยายขอบเขตความรู้ในเรื่องของเทคนิคในการวัด โดยเฉพาะ
ในสถานการณ์ของการประเมินเชิงสำรวจด้วยแบบสอบถาม ประโยชน์ที่คาดว่าจะได้รับจากการ
ศึกษาในลักษณะนี้ คือ

1. การพัฒนามาตรประมาณค่าเชิงพฤติกรรมอย่างมีลำดับขั้นตอนในการวิจัยนี้ จะทำให้
ผู้สนใจมองเห็นแนวทางการพัฒนาเครื่องมือวัดรูปแบบนี้ว่าเป็นรูปธรรมชัดเจน เป็นการเพิ่มทางเลือก
ในการพัฒนาเครื่องมือวัดประเภทมาตรประมาณค่า เพื่อการศึกษาริวิจัยในโอกาสต่อไปได้เป็น
อย่างดี
2. การตรวจสอบคุณภาพของมาตรประมาณค่าเชิงพฤติกรรม เปรียบเทียบกับมาตร
ประมาณค่าแบบกราฟิกอันเป็นรูปแบบที่ได้รับความนิยมใช้กันอย่างแพร่หลายทั่วไป ดังที่ได้ดำเนินการ
ในการวิจัยนี้ จะทำให้นักวิจัยและนักวัดผลมองเห็นจุดอ่อนหรือข้อจำกัดของมาตรประมาณค่าแบบ
กราฟิกได้อย่างชัดเจนยิ่งขึ้น ซึ่งจะเป็ประโยชน์อย่างยิ่งสำหรับผู้วิจัยหรือผู้บริโภคนผลการวิจัย ใน
การแปลความหมายข้อมูลผลการวิจัยในโอกาสต่อไป
3. การศึกษาครั้งนี้ ทำการศึกษาภายใต้สถานการณ์ของการประเมิน 2 ลักษณะ คือ
การประเมินตนเอง และการประเมินผู้อื่น โดยได้ทำการวิเคราะห์สภาพผลการประเมินในแต่ละ
สถานการณ์ด้วยดัชนีบ่งชี้คุณภาพหลายรายการ ดังนั้น ข้อสรุปจากงานวิจัยนี้จะเป็นประโยชน์
สำหรับการแปลความหมายข้อมูลในสถานการณ์การประเมินต่าง ๆ ทำให้นักวิจัยทราบลักษณะ
ธรรมชาติของผลการประเมินโดยกลุ่มบุคคลต่าง ๆ

4. การประเมินพฤติกรรมการปฏิบัติงาน ถือเป็นบทบาทที่สำคัญยิ่งประการหนึ่งในการบริหารงานบุคคล ซึ่งพบว่าในปัจจุบันหน่วยงานต่าง ๆ ได้ดำเนินการอย่างจริงจังมากขึ้น อาทิ สำนักงานคณะกรรมการข้าราชการครู (ก.ค.) ได้ใช้วิธีการนี้สำหรับการพิจารณาเลื่อนระดับของข้าราชการครู เป็นต้น รวมทั้งหน่วยงานอื่น ๆ ที่ทำหน้าที่ในการบริหารงานบุคคลก็ต้องดำเนินการในลักษณะที่คล้ายคลึงกัน ดังนั้นการศึกษาคั้งนี้ จะเป็นประโยชน์ต่อหน่วยงานต่าง ๆ จำนวนมาก

5. ในด้านการประเมินผลโครงการ โดยเฉพาะโครงการฝึกอบรม หรือ กรณีการติดตามผลผู้สำเร็จการศึกษาหลักสูตรต่าง ๆ ก็จะได้รับประโยชน์จากการศึกษาคั้งนี้ สามารถตัดสินใจเลือกใช้เทคนิควิธี หรือ เครื่องมือ ตลอดจนแนวทางในการเก็บรวบรวมข้อมูลที่เหมาะสม-มีประสิทธิภาพ ในโอกาสต่อไป