

วรรณคดีที่เกี่ยวข้อง

ในการนำเสนอวรรณคดีที่เกี่ยวข้อง ผู้วิจัยได้นำเสนอเนื้อหาออกเป็น 6 หัวข้อ คือความหมายของมาตราประมาณค่า ลักษณะของมาตราประมาณค่า รูปแบบของมาตราประมาณค่า ความคลาดเคลื่อนที่เกิดจากการใช้มาตราประมาณค่า วิธีการประเมินและงานวิจัยที่เกี่ยวข้องกับมาตราประมาณค่า

ความหมายของมาตราประมาณค่า

มาตราประมาณค่า ได้ถูกนำมาใช้เป็นเครื่องมือในการตัดสินคุณค่าของสิ่งต่าง ๆ ซึ่งอาจจะเป็นพฤติกรรมของบุคคล สถานการณ์ใด ๆ ก็ตาม ซึ่งผลของการประเมินจะอยู่ในรูปเชิงปริมาณมากกว่าเชิงคุณภาพ โดยให้ผู้ประเมินตัดสินคุณค่าของสิ่งที่ต้องการลงบนสเกล หรืออาจกล่าวได้ว่า มาตราประมาณค่าเป็นการแปลงค่าเชิงคุณภาพให้เป็นเชิงปริมาณ ด้วยการประเมินคุณค่าสิ่งนั้นในเชิงเปรียบเทียบกันเป็นกลุ่มเรียงตามลำดับมากขึ้น (บุญธรรม กิจปรีดาบริสุทธิ์ 2527 : 180)

สำหรับความหมายของมาตราประมาณค่า ได้มีผู้ให้ความหมายไว้ดังต่อไปนี้

วิเชียร เกตุสิงห์ (2523 : 180) กล่าวว่า มาตราประมาณค่า หมายถึง แบบสอบถามที่มีข้อคำถาม หรือข้อความเกี่ยวกับเรื่องที่จะให้ผู้ตอบพิจารณาพร้อมกับมีคำตอบที่แสดงความเข้มของความเห็นเกี่ยวกับเรื่องนั้น

Maslow (1982 : 411) กล่าวว่า มาตราประมาณค่า เป็นมาตราจัดอันดับ (Ordinal Scale) ที่มีถ้อยคำ หรือตัวเลขแสดงไว้อย่างเป็นลำดับขั้น

Gronlund (1981 : 441) กล่าวว่า มาตราประมาณค่าเป็นกระบวนการที่เป็นระบบ เพื่อที่จะให้มาซึ่งผลการตัดสินให้คุณค่าของผู้สังเกต ซึ่งมาตราประมาณค่าประกอบด้วยคุณลักษณะหรือคุณภาพของสิ่งที่จะประเมิน และระดับคุณค่าของสิ่งที่ปรากฏ

Guilford (1954 : 263) กล่าวว่า มาตรฐานค่าเป็นวิธีการวัดทางจิตวิทยา (Psychological - measurement) ที่ขึ้นอยู่กับ การตัดสินใจของผู้ประเมิน ในการประเมินพฤติกรรม บุคลิกลักษณะ หรือผลผลิต

สรุปได้ว่า มาตรฐานค่า คือ เครื่องมือที่ใช้สำหรับประเมินค่าสิ่งต่าง ๆ ที่เป็นคุณภาพให้อยู่ในเชิงปริมาณ โดยได้กำหนดสเกลของลักษณะที่ต้องการประเมินมาให้ ผู้ตอบเลือกอย่างเป็นลำดับขั้น

ลักษณะของมาตรฐานค่า

มาตรฐานค่าโดยทั่วไปจะประกอบด้วยลักษณะทั่ว ๆ ไป 2 องค์ประกอบด้วยกัน (R.L. Thorndike and E.P.Hagen 1977 : 462) คือ

1. คุณลักษณะของสิ่งที่ประเมิน (a set of stimulus variable)
2. รูปแบบคำตอบที่บ่งบอกถึงระดับคุณภาพของสิ่งที่ประเมิน (a pattern of response options)

ซึ่งมาตรฐานค่าที่นิยมใช้กันโดยทั่วไปจะใช้เพื่อประเมินพฤติกรรม (Behavior Ratings) หรือประเมินผลผลิต (Product Ratings) (Popham, W. James. 1981 : 315-317)

ในทางปฏิบัตินี้เราสามารถประยุกต์มาตรฐานค่าเพื่อใช้ในกิจกรรมทางการศึกษาและ เพื่อประเมินผลการเรียนรู้และพัฒนากิจการต่าง ๆ ซึ่งสามารถจำแนกได้เป็น 3 กลุ่มใหญ่ ๆ ดังนี้

1. ใช้ประเมินกระบวนการ (Procedures Evaluation) เช่น ใช้ในการประเมินทักษะการพูด ทักษะเชิงดนตรี ทักษะการเล่นกีฬา เป็นต้น
2. ใช้ในการประเมินผลผลิต (Products Evaluation) เช่น ภาพวาดสิ่งประดิษฐ์ต่าง ๆ เป็นต้น
3. ใช้ในการประเมินพัฒนาการทางบุคลิกภาพและทางสังคม (Evaluating Personal-Social Development) เช่น ความสนใจ ทศนคติ ภาวะผู้นำ ความรับผิดชอบการปรับตัว เป็นต้น

จากที่กล่าวมาจะเห็นได้ว่า มาตรฐานค่าเป็นเครื่องมือที่ให้ประโยชน์อย่างยิ่ง สำหรับการแปลความหมายข้อมูลเชิงคุณภาพให้เป็นเชิงปริมาณ โดยอาศัยการกำหนด มาตรฐานค่าของรายการที่ต้องการประเมินให้อยู่ในรูปของทางเลือกต่าง ๆ ที่เรียง เอาไว้เป็นลำดับขั้นและสามารถใช้ในกิจกรรมประเมินค่าต่าง ๆ ได้อย่างหลากหลาย

รูปแบบของมาตรฐานค่า

ด้วยเหตุผลที่เป็นเครื่องมือที่สร้างได้ไม่ยากนัก ใช้งบประมาณในการสร้างน้อย สะดวกต่อการนำไปใช้ และสามารถประยุกต์ใช้ในสถานการณ์การประเมินต่าง ๆ ได้อย่าง กว้างขวาง ส่งผลให้มาตรฐานค่ากลายเป็นเครื่องมือที่นิยมใช้กันอย่างแพร่หลายใน การวิจัย (Turney and Robb 1971 : 136) และสามารถกล่าวได้ว่า ในบรรดา วิธีการวัด ทางจิตวิทยาทั้งหลายที่อาศัยการตัดสินของคน ซึ่งมาตรฐานค่าจัดได้ว่าเป็น เครื่องมือที่ได้รับการนิยมใช้มากที่สุด (Guilford : 263) ซึ่งภายใต้สถานการณ์ที่นิยม ใช้กันอย่างแพร่หลายนี้เอง จึงได้มีการพัฒนารูปแบบของมาตรฐานค่าที่ปรากฏราย ละเอียดแตกต่างกันหลายรูปแบบ

Guilford ได้แบ่งประเภทของมาตรฐานค่าตามที่นิยมใช้กันอยู่ออกเป็น 5 แบบ คือ

1. แบบตัวเลข (Numerical Scales) กำหนดตัวเลขพร้อมกับนิยามหรือคำ บรรยายไว้ให้ผู้ประเมิน ดังตัวอย่างมาตรวัดที่ใช้ในการประเมินค่าของสีและกลิ่น ต่อไปนี้

- 10 Most pleasant imaginable
- 9 Most pleasant
- 8 Extremely pleasant
- 7 Moderately pleasant
- 6 Mildly pleasant
- 5 Indifferent

- 4 Midly impleasant
 3 Moderately impleasant
 2 Most impleasant
 1 Most impleasant imaginable

ในบางครั้งอาจไม่มีตัวเลขให้ มีแต่เพียงข้อความ (Descriptive Cues)

ดังนี้

Very Heavy

Heavy

Medium

Light

Very Light

2. แบบกราฟิก (Graphic Scales) เป็นแบบที่นิยมใช้กันอย่างกว้างขวางมากที่สุด มีได้หลายรูปแบบ เน้นการจัดข้อความเข้ากับเส้นตรง ดังตัวอย่าง

ในวงสนทนาคุณเป็นอย่างไร _____
 ช่างพูด พูดน้อย พูดเมื่อ ชอบฟัง ไม่เคย
 จำเป็น มากกว่า พูดเลย

เขาคิดได้ช้าหรือเร็ว _____

ช้ามาก ค่อยๆคิด ธรรมดา ว่องไว รวดเร็วมาก

3. แบบสเกลมาตรฐาน (Standard Scales) ในแบบนี้เป็นแบบที่ได้กำหนดค่าสเกล (Scales Value) เป็นค่าประจำของชุดมาตรฐานที่จะใช้ในการเปรียบเทียบไว้ล่วงหน้าผู้ประเมินจะต้องทำการเปรียบเทียบรายการที่ให้ประเมินกับชุดมาตรฐาน

4. แบบแต้มสะสม (Rating by Cumulative Points) แตกต่างจากแบบอื่นที่วิธีการให้คะแนน วิธีการตรวจรายการ (Check-List) เป็นตัวอย่างหนึ่งของรูปแบบนี้ได้มีผู้ใช้วิธีการตรวจรายการในการประเมินคุณลักษณะของเด็ก โดยกำหนดรายการที่มุ่ง

คุณลักษณะที่ดีและไม่ดีมาให้เลือก เช่น ให้ความร่วมมือ ช่างคิด ดูร้าย เห็นแก่ตัว เป็นต้น ผู้ประเมินแต่ละคนจะเลือกทุก ๆ คุณลักษณะที่ตรงกับเด็กคนนั้น คะแนนของเด็กแต่ละคนจะเป็นผลรวมของคะแนนทั้งหมด ซึ่งได้กำหนดให้ 1 สำหรับทุกคุณลักษณะที่ดี และให้คะแนน -1 สำหรับทุกคุณลักษณะที่ไม่ดี

5. แบบบังคับให้เลือก (Forced-Choice Rating) เป็นแบบที่พัฒนาขึ้นเพื่อประเมินบุคคล ผู้ประเมินไม่ต้องบอกว่าผู้ถูกประเมินมีลักษณะใด มากน้อยเพียงใด แต่ให้เปรียบเทียบเป็นคู่ ๆ แล้วบอกว่ามีลักษณะใดมากกว่ากัน

โดยปกติมาตรประมาณค่ามีอยู่ด้วยกันหลายรูปแบบ Turney และ Robb (1971 : 136-137) ได้จำแนกชนิดของมาตรประมาณค่าไว้ 3 แบบ คือ

1. แบบตัวเลข (Numerical Scale) ใช้ตัวเลขบอกระดับของลักษณะหรือพฤติกรรมที่ต้องการประเมิน เช่น ให้ประเมินพฤติกรรมของนักเรียนตามรายการต่อไปนี้ โดยให้วงกลมล้อมรอบหมายเลข 1 เมื่ออยู่ในระดับต่ำ, 2 เมื่อต่ำกว่าปานกลาง, 3 เมื่ออยู่ระดับปานกลาง, 4 เมื่อสูงกว่าปานกลาง และ 5 เมื่ออยู่ในระดับสูง

(1) ความร่วมมือกับเพื่อน	1	2	3	4	5
(2) ความร่วมมือกับครู	1	2	3	4	5

2. แบบกราฟิก (Graphic Rating Scale) ให้ทำเครื่องหมายลงบนเส้นตรง เพื่อแสดงระดับลักษณะที่ประเมิน เช่น ให้แสดงความเห็นเกี่ยวกับลักษณะของนักเรียน โดยทำ เครื่องหมายกากบาท (X) ลงบนเส้นตรงในตำแหน่งที่ตรงกับลักษณะของนักเรียนคนนั้นมากที่สุด

(1) ความร่วมมือ						
	ชดชวาง					เต็มใจเสมอ
(2) ความคงที่ของอารมณ์						
	ไม่คงที่					ควบคุมได้ดี

3. แบบบรรยาย (Descriptive Scale) มีคำบรรยายหรือข้อความแสดงระดับของลักษณะที่ต้องการประเมิน เช่น ให้ทำเครื่องหมายกากบาท (X) ลงในช่องที่ตรงกับพฤติกรรมของนักเรียนในแต่ละรายการต่อไปนี้

	ไม่เคย	นาน ๆ ครั้ง	บางครั้ง	บ่อยครั้ง	เสมอ
(1) ฟังความคิดเห็นของคนอื่น					
(2) ยอมรับฟังคำติเชิงก่อก					

รูปแบบของมาตรประมาณค่าตามที่ใช้ในวิทยานิพนธ์ส่วนมากเป็นมาตรประมาณค่าแบบบรรยาย แบบตัวเลข และแบบที่มีทั้งตัวเลขและคำบรรยาย นอกจากนี้ยังมีแบบที่ใช้สัญลักษณ์อื่น ๆ ที่ไม่ใช่ตัวเลขอีกด้วย

Gronlund, Norman E. (1981 : 441-449) ได้แบ่งรูปแบบของมาตรประมาณค่าออกเป็น 3 รูปแบบ คือ

1. มาตรประมาณค่าแบบตัวเลข (Numerical Rating Scale) เป็นรูปแบบมาตรประมาณค่าที่ให้ผู้ประเมินทำเครื่องหมายตรงตัวเลขที่เป็นดัชนีชี้ถึงระดับของคุณลักษณะที่ปรากฏ
2. มาตรประมาณค่าแบบกราฟิก (Graphic Rating Scale) เป็นมาตรประมาณค่าที่คุณลักษณะหรือสิ่งที่จะประเมินปรากฏอยู่บนเส้นในแนวนอน (Horizontal line) โดยผู้ประเมินจะทำเครื่องหมายบนเส้นตรงนั้น
3. มาตรประมาณค่าบรรยายแบบกราฟิก (Descriptive Graphic Rating Scale) เป็นมาตรประมาณค่าที่มีวลีบรรยายแต่ละจุดบนสเกล ซึ่งวลีที่เพิ่มเข้ามานั้นจะเป็นค่าที่แสดงถึงพฤติกรรมที่แตกต่างกันบนสเกล โดยทั่วไปแล้วมาตรประมาณค่าชนิดนี้จะมี 2 ลักษณะด้วยกันคือ



3.1 ลักษณะที่มีคำบรรยาย 3 จุด คือ มีคำบรรยายจุดปลายสเกลทั้งสอง
ด้านและตรงกลาง เช่น

1. To what extent does the pupil participate in discussion ?

--	--	--	--

Never participates Participates as Participates more
quiet, passive much as other than any other
group members group member

Comment

3.2 ลักษณะที่มีคำบรรยายทุกจุดบนสเกล

ไม่ว่ามาตรฐานคำบรรยายแบบกราฟิกจะมีลักษณะใดก็ตาม แต่ละข้อกระทง
จะมีที่ว่างไว้เพื่อให้ผู้ประเมินใส่เหตุผลประกอบการประเมิน หรือบันทึกพฤติกรรมที่ปรากฏ
ในขณะที่ทำการประเมิน

จากรูปแบบมาตรฐานค่าทั้ง 3 รูปแบบที่กล่าวมาแล้ว Gronlund ได้ให้
ความเห็นอีกว่า มาตรฐานคำบรรยายแบบกราฟิกเป็นรูปแบบที่ใช้กันอย่างแพร่หลายใน
โรงเรียน เนื่องจากเป็นมาตรฐานค่าที่ทั้งครูและนักเรียนสามารถที่จะทำความเข้าใจ
พฤติกรรมที่จะประเมินได้ชัดเจน

Popham (1981 : 317-322) ได้แบ่งประเภทของมาตรฐานค่าตามที่นิยม
ใช้กันโดยทั่วไปออกเป็น 3 รูปแบบ คือ

1. แบบตัวเลข (Numericals Rating Scales) เป็นมาตรฐานค่าที่มี
ค่าตัวเลขที่ต่อเนื่องสำหรับให้ผู้ประเมินเป็นผู้ตัดสิน โดยกำหนดความหมายให้กับค่าของตัวเลข

ตัวอย่างเช่น การประเมินพฤติกรรมการสอนของครู

กำหนดให้	5	หมายถึง	ดีเลิศ	(Excellent)
	4	หมายถึง	ดี	(Above Average)
	3	หมายถึง	ปานกลาง	(Average)

2 หมายถึง ควรปรับปรุง (Below Average)

1 หมายถึง ไม่ดี (Poor)

[] การสร้างปฏิสัมพันธ์ที่ดีกับนักเรียน

1 2 3 4 5

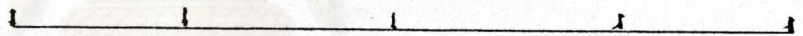
[] การนำเสนอเนื้อหาได้อย่างชัดเจน

1 2 3 4 5

2. แบบกราฟิก (Graphic Rating Scales) เป็นมาตราประมาณค่าที่กำหนดคุณลักษณะที่จะประเมินควบคู่ไปกับเส้นตรงในแนวนอน (horizontal line) สำหรับให้ผู้ประเมินทำเครื่องหมาย

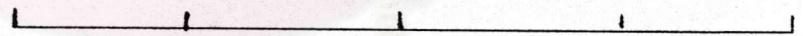
ตัวอย่างเช่น การประเมินการมีปฏิสัมพันธ์ในการร่วมกิจกรรมนอกหลักสูตรของนักเรียน

1. การเข้าร่วมกิจกรรมในชมรม



เสมอ ๆ บ่อยครั้ง บางครั้งบางครั้ง นาน ๆ ครั้ง ไม่เคย

2. การให้ความร่วมมือในที่มียมนาสดิก



เสมอ ๆ บ่อยครั้ง บางครั้งบางครั้ง นาน ๆ ครั้ง ไม่เคย

3. แบบจัดอันดับ (Ranking) การประเมินโดยใช้มาตราประมาณค่าแบบจัดอันดับเป็นวิธีการที่แตกต่างไปจากรูปแบบอื่น ๆ ซึ่งมาตราประมาณค่าแบบจัดอันดับเป็นที่รู้จักโดยทั่วไปในรูปของ การเปรียบเทียบรายคู่ (paired-comparison method)

Richard L. Luftig (1989 : 165-166) ได้แบ่งมาตราประมาณค่าออกเป็น 4 ประเภท คือ

1. Category Ratings เป็นรูปแบบมาตราประมาณค่าที่ให้ผู้ตอบเลือกคำตอบที่ดีที่สุดเพื่อตอบคำถามหรือบรรยายคำถามนั้น ซึ่งรวมถึง Multiple-choice question ด้วย

2. Numerical Ratings เป็นรูปแบบมาตรประมาณค่าที่มีลักษณะกำหนดค่าตัวเลขที่ต่อเนื่องให้กับคำที่อธิบายคำตอบ ซึ่งผู้ประเมินจะเลือกตัวเลขเป็นคำตอบ
3. Graphic Ratings เป็นรูปแบบมาตรประมาณค่าที่ใช้เส้นตรงเป็นตัวแทนค่าที่ต่อเนื่องโดยมีช่วงที่เท่า ๆ กัน
4. Force-Choice Ratings เป็นรูปแบบมาตรประมาณค่าที่ให้ผู้ตอบเลือกคำตอบจากชุดของคำตอบ ซึ่งแต่ละคำตอบจะมีลักษณะความดึงดูด หรือคำตอบที่มีน้ำหนักเท่าเทียมกัน

อุทุมพร จามรมาน (2530 : 105-106) ได้กล่าวถึงมาตรประมาณค่าว่าเป็นเครื่องมือช่วยสังเกตพฤติกรรมที่ปรากฏ หรือไม่ปรากฏก็ได้ แต่ผู้สังเกตต้องเปรียบเทียบความมากน้อยหรือความแตกต่างของพฤติกรรมให้ได้ และได้จำแนกแบบประเมินค่าเป็น 3 ประเภท พร้อมทั้งยกตัวอย่าง ดังนี้

1. Numerical Rating Scales ประกอบด้วย รายการพฤติกรรมและเลขกำกับ เช่น 1, 2, 3, 4, 5 ให้ผู้ประเมินค่าทำวงกลมรอบเลขที่เลือก ตัวอย่างเช่น

บุคลิกภาพของอาจารย์ X	เลขประมาณค่า				
ก. แต่งกายเรียบร้อย	1	2	3	4	5
ข. ตรงเวลา	1	2	3	4	5

2. Graphic Rating Scales คล้ายแบบที่ 1 แต่ใช้ภาษาแทนตัวเลขในสเกลและใช้เส้นแทนช่อง โดยมีข้อความกำกับหัวท้าย เช่น

ก. การใช้อุปกรณ์การสอนของอาจารย์	ไม่เหมาะสม	_____	เหมาะสม
ข. การอธิบาย		_____	
		ไม่ชัดเจนเลย	ชัดเจนบางครั้ง
		ชัดเจน	

สเกลแบบนี้ให้ผู้ประเมินค่าใส่เครื่องหมายลงที่ใดที่หนึ่งบนเส้น

3. Descriptive Graphic Rating Scales เป็นแบบประเมินค่าของ 1 และ 2 รวมกัน คือมีทั้งเลขและภาษากำกับหัวท้าย เช่น

น้ำเสียงของอาจารย์	เวลาสอน	_____			
		1	2	3	4
		เบามาก		กำลังดี	ดังเกินไป

หรืออาจจัดสเกลเรียงลำดับจากมากไปน้อย หรือน้อยไปมาก เช่น

น้ำเสียงของอาจารย์ X					
เวลาสอน	1	2	3	4	5
	ดั่งมาก	ดั่งมาก	ดั่ง	เบา	ไม่ได้ยิน
	เกินไป		พอสมควร		เลย

วิธีการประเมิน

ในการจำแนกมาตรฐานประมาณค่านี้ ถ้าจำแนกตามวิธีการใช้แล้ว จำแนกได้ 3 ชนิด (Wallage 1957 : 166) คือ

1. มาตรฐานประมาณค่าที่ใช้ประเมินตนเอง
2. มาตรฐานประมาณค่าที่ใช้ประเมินผู้อื่น
3. มาตรฐานประมาณค่าที่ใช้สำหรับประเมินตนเองและผู้อื่น

Mathis (1979 : 298-301) ได้เสนอวิธีการประเมินพฤติกรรมการทำงานที่นิยมใช้มี 4 รูปแบบ คือ

1. วิธีการประเมินโดยกลุ่มผู้ใต้บังคับบัญชา (Subordinate Rating Approach) เป็นการให้ผู้ใต้บังคับบัญชาหรือกลุ่มสมาชิกเป็นผู้ประเมินผู้บังคับบัญชา
2. วิธีการประเมินโดยกลุ่ม (Group Rating Approach) เป็นการประเมินบุคคลใดบุคคลหนึ่งโดยใช้มติที่ประชุม หรือสมาชิกแต่ละคนประเมินอย่างเป็นอิสระแยกจากกัน แล้วจึงสรุปผลการประเมิน
3. วิธีการให้ผู้บังคับบัญชาประเมินผู้ใต้บังคับบัญชา (Superior Rating of Subordinates)
4. วิธีการประเมินตนเองโดยยึดเป้าหมายงานเป็นเกณฑ์ (Self-Appraisal through the use of Management by Objectives) เป็นวิธีการที่เปิดโอกาสให้บุคลากรสามารถประเมินผลงานของตนเองโดยยึดผลงานเป็นเกณฑ์

สำหรับการประเมินการเรียนการสอน Dressel (1977 : 1486-1487) Marris (1986 : 102) และนักการศึกษาไทย อุตุมพร จามรมาน (2530 : 101) บุญธรรม กิจปรีดาวิสุทธ์ (2526 : 12) มีความเห็นสอดคล้องกันว่า ผู้ที่ทำการประเมิน การสอนของครูควรประกอบด้วย

1. นักเรียน
2. เพื่อนครู หรือเพื่อนร่วมงาน
3. ครูผู้สอนประเมินตนเอง
4. ผู้บริหารหรือผู้สังเกต

ความคลาดเคลื่อนที่เกิดจากการใช้มาตรฐานค่า

ถึงแม้ว่ามาตรฐานค่าเป็นเครื่องมือที่สร้างง่าย ใช้สะดวก จึงเป็นที่นิยมใช้กันอย่างกว้างขวาง โดยเฉพาะในการสังเกตและประเมินพฤติกรรมของมนุษย์ แต่การใช้มาตรฐานค่าก็ต้องใช้ด้วยความระมัดระวัง เนื่องจากมีข้อจำกัดอันเกิดจากความคลาดเคลื่อนในการวัด

ความคลาดเคลื่อนที่เกิดขึ้นในการใช้มาตรฐานค่า เป็นผลมาจากข้อบกพร่องในการสร้างเครื่องมือ เป็นต้นว่า ลักษณะที่ต้องการให้ประมาณค่าเป็นลักษณะที่อยู่ภายในสังเกตได้ยาก และอาจแสดงออกไม่เหมือนกัน เช่น ความตั้งใจ เครียด บางครั้งความหมายและทิศทางของลักษณะไม่เด่นชัด ไม่มีขอบเขตที่แน่นอน ทำให้ผู้ประเมินตีความไปในแนวทางที่แตกต่างกัน และที่สำคัญคือ ผู้ประเมินขาดเกณฑ์ในการประมาณค่าที่เป็นอันหนึ่งอันเดียวกัน เช่น "ดี" ของคนหนึ่งอาจเป็นแค่ระดับ "พอใช้" ของอีกคนหนึ่งก็ได้ การที่ผู้ประเมินมีขอบเขต และเกณฑ์ในการประมาณค่าที่แตกต่างกันนี้มีผลทำให้ค่าความเที่ยงในระหว่างผู้ประเมินลดลง (Selltiz, et al. 1959 : 352)

การใช้เครื่องมือเก็บรวบรวมข้อมูลที่เป็นมาตรฐานค่ามีความคลาดเคลื่อนที่อาจเกิดขึ้นได้จาก 3 แหล่ง ดังนี้คือ

1. ความบกพร่องของเครื่องมือ ข้อบกพร่องที่สำคัญของเครื่องมือที่เป็นมาตรฐานค่าคือ ขาดการนิยามหรือบรรยายคุณลักษณะ (characteristic) ของสิ่งที่จะ

ประเมินไว้อย่างชัดเจน ทำให้เกิดความเข้าใจที่ไม่ตรงกัน จึงเป็นผลทำให้การประเมินค่านั้นขาดความเที่ยง หรืออาจเป็นเพราะภาษาที่ใช้ไม่เหมาะสม หรือไม่สามารวัดสิ่งที่ต้องการวัดได้ เป็นต้น

2. ความบกพร่องที่เกิดจากกระบวนการในการประเมิน ข้อบกพร่องที่มักจะพบคือคุณลักษณะหรือสิ่งที่ต้องการให้ประเมินนั้นมีจำนวนมาก จึงทำให้ผลการประเมินเกิดความคลาดเคลื่อนได้ และข้อบกพร่องที่สำคัญอีกประการหนึ่ง คือ กรณีที่มีผู้ประเมิน (Rater) เพียงคนเดียว ทำการประเมิน ผู้ถูกประเมิน (Ratee) หลาย ๆ คน อาจทำให้เกิดความคลาดเคลื่อนขึ้นได้ แต่ถ้ามีผู้ประเมินหลาย ๆ คน จะเป็นวิธีที่ช่วยลดความคลาดเคลื่อนที่เกิดจากแต่ละบุคคลได้ ทำให้ผลการประเมินมีความแม่นยำยิ่งขึ้น

3. ความคลาดเคลื่อนที่เกิดจากตัวผู้ประเมิน ซึ่งมีความคลาดเคลื่อนหลายประการที่พึงควรระมัดระวัง ได้แก่ ความลำเอียงหรือความมีอคติของแต่ละบุคคล (Personal Bias) ซึ่งเป็นผลมาจากการมีปฏิสัมพันธ์มาก่อน (Halo effect) ความคลาดเคลื่อนแบบใช้ตรรกผิด (Logical error) และความคลาดเคลื่อนเชิงปล่อยหรือกตคะแนน (Error of Leniency)

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 แต่เมื่อมีความคลาดเคลื่อนชนิดนี้เกิดขึ้น ผู้นั้นจะได้รับการประเมินให้อยู่ในระดับสูงหรือค่อนข้างสูง เป็นผลทำให้ค่าที่ประเมินสูงกว่าความเป็นจริง

3.1.2 ความคลาดเคลื่อนแบบประเมินค่าต่ำกว่าความเป็นจริง (Severity Error) เป็นความคลาดเคลื่อนที่เกิดขึ้นเมื่อผู้ประเมิน ประเมินค่าคุณลักษณะต่าง ๆ ของทุก ๆ คน ในระดับต่ำหรือค่อนข้างต่ำเพียงด้านเดียว เป็นผลทำให้ค่าที่

ประเมินได้ต่ำกว่าที่ควรจะเป็น กล่าวคือ เมื่อคุณลักษณะของผู้ถูกประเมินมีค่าอยู่ในระดับดีหรือปานกลาง เมื่อมีความคลาดเคลื่อนชนิดนี้เกิดขึ้น ผู้ถูกประเมินผู้นั้นจะได้รับการประเมินให้อยู่ในระดับต่ำหรือค่อนข้างต่ำ

3.1.3 ความคลาดเคลื่อนแบบโน้มเข้าสู่ตรงกลาง (Central Tendency Error) ความคลาดเคลื่อนชนิดนี้เกิดขึ้นเมื่อผู้ประเมินพยายามหลีกเลี่ยงการประเมินค่าทางระดับค่อนข้างสูงและระดับค่อนข้างต่ำมาเป็นระดับปานกลาง ทำให้ค่าที่ประเมินได้ต่ำกว่าหรือสูงกว่าสภาพที่เป็นจริง ซึ่งสาเหตุของความคลาดเคลื่อนชนิดนี้อาจเป็นเพราะ (Wallage 1957 : 167)

3.1.3.1 ผู้ประเมินอาจจะไม่ทราบว่สิ่งที่เขาประเมินนั้นมีพิสัยอยู่ในทางสูงหรือทางต่ำ

3.1.3.2 ผู้ประเมินอาจจะไม่มีความละเอียดในการจำแนกความแตกต่างของสิ่งทีประเมิน

3.1.3.3 อาจเป็นเพราะบุคลิกภาพของผู้ประเมินเป็นคนไม่มั่นคง และไม่กล้าที่จะประเมินค่าทางสูงหรือทางต่ำทางใดทางหนึ่ง

3.4.3.4 ผู้ประเมินอาจจะรู้สึกที่ไม่สามารถที่จะตัดสินใจได้อย่างเป็นธรรมและอาจเกรงว่าจะเป็นการทำร้ายจิตใจคน

จะเห็นว่า การที่ผู้ประเมิน ประเมินค่าเพียงระดับใดระดับหนึ่งเพียงระดับเดียว จะทำให้เกิดผลที่ไม่น่าพึงปรารถนา 2 ประการ (Gronlund 1976 : 442-443) คือ

ประการแรก ทำให้ค่าที่ประเมินได้เป็นค่าที่ตรงสภาพเพียงใด ทั้งนี้เพราะการที่ผู้ประเมินมักประเมินเฉพาะทางสูงหรือทางต่ำทางเดียว น่าจะสะท้อนให้เห็นถึงการมองของแต่ละบุคคล (Personal Outlook) มากกว่าที่จะเป็นคุณลักษณะของผู้ถูกประเมิน

ประการที่สอง การที่ผู้ประเมินมักจะประเมินเฉพาะระดับใดระดับหนึ่งเป็นการจำกัดขอบเขตในการประเมิน ทำให้อำนาจจำแนกต่ำลง ความเที่ยงย่อมต่ำลงด้วย

3.2 ผลจากการมีปฏิสัมพันธ์มาก่อน (Halo Effect) เป็นความคลาดเคลื่อนที่เกิดขึ้น เมื่อความรู้สึกของผู้ประเมินที่มีต่อผู้ถูกประเมินเมื่ออิทธิพลต่อการประเมินค่าคุณลักษณะของผู้ถูกประเมิน เช่น ถ้าผู้ประเมินมีทัศนคติในทางที่ดีต่อผู้ถูกประเมินแล้ว เขาก็มักจะประเมินคุณลักษณะต่าง ๆ ของผู้ถูกประเมินในทางสูง แต่ในทางตรงกันข้าม ถ้าผู้ประเมินมีทัศนคติในทางลบต่อผู้ถูกประเมินแล้ว เขาก็มักจะประเมินคุณลักษณะต่าง ๆ ของผู้ถูกประเมินในทางต่ำ ความคลาดเคลื่อนแบบนี้จะเกิดขึ้นกับเฉพาะบางคนที่ถูกประเมิน ซึ่งแตกต่างจากความคลาดเคลื่อนแบบ Generosity Error และ Severity Error ที่ผู้ประเมินมักจะประเมินทุกคนในทางสูงเหมือนกัน และในทางต่ำเหมือนกัน

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

3.4 ความคลาดเคลื่อนเชิงปล่อยหรือกตคะแนน (Error of Leniency) ความคลาดเคลื่อนแบบนี้คล้าย Generosity Error และ Severity Error ต่างกันที่ผู้ประเมินพยายามประเมินผลงานในด้านดีเกินไป (Easy Rate) หรือในทางตรงกันข้าม ผู้ประเมินจะประเมินผลงานให้ต่ำกว่าความเป็นจริง (Hard Rate)

ความคลาดเคลื่อนดังที่ได้กล่าวมานี้เป็นความคลาดเคลื่อนอย่างเป็นระบบ (Systematic error) ที่ทำให้ความตรงลดลง (Selltiz, et. al. 1959 : 352)

การปรับปรุงแก้ไขการประมาณค่าให้ได้ผลที่ถูกต้องแน่นอน นั้น Thorndike และ Hagen (1977 : 487) ได้สรุปไว้ว่าควรปฏิบัติตามเกณฑ์ต่อไปนี้

1. การประเมินควรจะทำเฉพาะลักษณะที่ปรากฏให้เห็นชัดเจน
2. ลักษณะที่จะประเมินต้องแสดงออกมาให้เห็นในรูปพฤติกรรม และการตัดสินใจกระทำจากพฤติกรรมที่เห็นนั้น
3. รูปแบบของมาตราประมาณค่า (Rating Form) ต้องปรับปรุงเพื่อบังคับให้ผู้ประเมินบอกความแตกต่าง หรือเพื่อความคุมการที่ผู้ประเมินมีมาตรฐานในการตัดสินที่ต่างกัน
4. ผู้ประเมินต้องเป็นผู้ที่มีโอกาสมากที่สุดในการสังเกตพฤติกรรมในสถานการณ์ตามที่ต้องการ
5. ผู้ประเมินต้องได้รับการบอกเล่าถึงคุณค่าของการประเมินค่าและได้รับการฝึกฝนในการใช้เครื่องมือ
6. ถ้าสามารถมีผู้ประเมินหลายคนก็ควรรวบรวมผลการประเมินจากผู้ประเมินหลาย ๆ คน

ในการประเมินค่านี้ แม้จะมีความคลาดเคลื่อนเกิดขึ้นได้ แต่ก็มีวิธีที่จะช่วยลดความคลาดเคลื่อนได้โดยออกแบบมาตราส่วนประมาณค่าให้ถูกต้องและเหมาะสม (Gronlund 1976 : 443) Guilford (1954, อ้างถึงใน Chase 1978 : 165-166), Gronlund (1976 : 443-444), Tuckman (1975 : 196-199), Adam (1964 : 278), Wallage (1957 : 168) และ Travers (1955 : 218-219) ได้เสนอแนะวิธีการในการสร้างมาตราประมาณค่าให้มีคุณภาพ ซึ่งพอจะสรุปสาระสำคัญในการสร้างเป็น 3 ขั้นตอน ดังนี้

1. เลือกคุณลักษณะหรือพฤติกรรมที่จะประเมินค่า ซึ่งควรมีลักษณะดังต่อไปนี้
 - 1.1 ต้องมีความสำคัญและตรงกับสิ่งที่ต้องการวัดมากที่สุด
 - 1.2 ต้องเป็นอิสระไม่เกี่ยวข้องกับคุณลักษณะอื่น ๆ
 - 1.3 ต้องเป็นสิ่งที่สามารถสังเกตได้โดยตรง
 - 1.4 แต่ละองค์ประกอบของพฤติกรรมมีความเป็นเอกภาพ
 - 1.5 คุณลักษณะหรือพฤติกรรมต่าง ๆ ไม่ควรมีมากเกินไป ปกติควรมี 5-6

2. นิยามหรือบรรยายคุณลักษณะหรือพฤติกรรมที่จะประเมินค่าให้ชัดเจนเพื่อให้เกิดความเข้าใจที่ตรงกัน และสามารถตัดสินได้อย่างเป็นปรนัยและเที่ยง
3. ออกแบบช่วงของมาตราส่วนที่จะใช้ประเมิน ดังนี้
 - 3.1 ปกติมักจะแบ่งช่วงออกเป็น 3-7 ช่วง การแบ่งช่วงยิ่งมากยิ่งแสดงความแตกต่างได้มาก การแบ่งช่วงเป็นจำนวนคี่จะมีข้อดีที่ทำให้มีค่ากลาง และในบางครั้งอาจใช้ในกรณีที่มีการประเมินค่าที่ตัดสินใจไม่ได้ มาตรฐานประมาณค่าแบบ 5 ช่วง เป็นที่นิยมใช้กันมาก
 - 3.2 มีการนิยามหรืออธิบายความต่อเนืองในแต่ละช่วงให้ชัดเจน
 - 3.3 ควรมีการสลับช่วงจากน้อยไปหามาก และจากมากไปหาน้อย ทั้งนี้เพื่อป้องกันมิให้ผู้ประเมินประเมินค่าโดยอัตโนมัติ

งานวิจัยที่เกี่ยวข้อง

ด้วยเหตุที่มาตรฐานค่าเป็นเครื่องมือที่นิยมใช้กันอย่างแพร่หลายในการเก็บรวบรวมข้อมูล ในขณะที่มาตรฐานค่ามีความคลาดเคลื่อนเกิดขึ้นจากตัวผู้ประเมิน เครื่องมือ และกระบวนการประเมิน ดังนั้น การปรับปรุงเครื่องมือให้มีประสิทธิภาพจึงเป็นเรื่องที่สำคัญมาก จึงได้มีการศึกษาและวิจัยเพื่อหาแนวทางในการพัฒนามาตรฐานค่าในรูปแบบต่าง ๆ ดังต่อไปนี้

Campbell และคณะ (1973 : 15-22) ได้ทำการศึกษาเปรียบเทียบมาตรฐานค่า 2 แบบคือ มาตรฐานค่าแบบรวม (Summated Rating Technique) ของ Likert และมาตรฐานค่าเชิงพฤติกรรม (Behavioral Anchore Scale) ของ Smith และ Kendall ผลการวิจัยพบว่า มาตรฐานค่าเชิงพฤติกรรมเกิดความคลาดเคลื่อนเชิงปล่อยหรือกดคะแนน (Leniency error) และความคลาดเคลื่อนอันเนื่องมาจากการมีปฏิสัมพันธ์มาก่อน (Halo Error) น้อยกว่ามาตรฐานค่าแบบรวม สำหรับด้านความตรงพบว่า มาตรฐานค่าเชิงพฤติกรรม มีความคมชัดในแง่ของ Convergent และ Discriminant Validity สูงกว่ามาตรฐานค่าแบบรวม

Shower (1974 : 5731-A) พบว่า ในการประเมินการสอนของครูนั้น นักเรียนมักจะเกิด Leniency bias โดยจะเลือกเฉพาะระดับสเกลที่มีค่าสูงเพียง 2-3 ค่าเท่านั้น ทำให้การประมาณค่าไม่ได้ผลเนื่องมาจากไม่สามารถจะบอกความแตกต่างของครูแต่ละคนได้ ดังนั้น จึงได้ทำการวิจัยเพื่อปรับปรุงแบบประเมินการสอนของครูโดยทดลองใช้รูปแบบคำตอบเป็น 3 แบบคือ (1) แบบลิเคิร์ต (Fix Alternative Likert Cues) (2) แบบตัวเลือกคงที่ที่ใช้คำคุณศัพท์เชิงประเมิน (Fix Alternative Evaluation Cues-Superior - Inferior) และ (3) แบบทางเลือกเชิงบรรยาย (Descriptive Cues) ผลการวิจัยพบว่า แบบที่ 2 คือ แบบตัวเลือกคงที่ที่ใช้คำคุณศัพท์เชิงประเมิน เกิด Leniency bias น้อยที่สุด ส่วนแบบลิเคิร์ตซึ่งเป็นแบบที่นิยมใช้กันมากที่สุดที่แนวโน้มที่จะเกิดอคติในการประเมินสูงสุด อย่างไรก็ตามเมื่อเปรียบเทียบค่าความเที่ยงของผู้ประเมิน (Rater Reliability) ทั้ง 3 รูปแบบ ไม่แตกต่างกัน

Keaveny และ McGann (1975 : 695-703) ได้ศึกษาและเปรียบเทียบมาตรฐานประมาณค่าแบบพฤติกรรมที่คาดหวัง (Behavioral Expectation Scales) ตามแบบของ Smith และ Kendall กับมาตรฐานประมาณค่าแบบกราฟิก ในการประเมินพฤติกรรมการสอนของศาสตราจารย์ 4 คน โดยนักศึกษาเป็นผู้ประเมินโดยใช้เครื่องมือประเมินทั้ง 2 รูปแบบ ผลการวิจัยพบว่า มาตรฐานประมาณค่าแบบพฤติกรรมที่คาดหวังเกิด Halo Error น้อยกว่ามาตรฐานประมาณค่าแบบกราฟิกและมาตรฐานประมาณค่าเชิงพฤติกรรมที่คาดหวังไม่สามารถลด Leniency Error ได้ ส่วนในแง่ของความตรงเชิงจำแนก พบว่ามาตรฐานประมาณค่าแบบพฤติกรรมที่คาดหวัง ให้ค่าความตรงเชิงจำแนก (Discriminant Validity) สูงกว่ามาตรฐานประมาณค่าแบบกราฟิก

Slater (1977 : 6203-A) ได้พัฒนามาตรประมาณค่า 2 รูปแบบ คือ แบบที่ใช้คำคุณศัพท์ผูกติดกับระดับการประเมิน (Adjective-Anchored Scales) และแบบที่สอง เป็นแบบที่ใช้ตัวอย่างพฤติกรรมผูกติดกับระดับการประเมินในแต่ละสเกล (Behaviorally-Anchored Scales) ซึ่งผู้วิจัยได้นำแบบประเมินทั้งสองรูปแบบไปใช้ในการประเมินพฤติกรรมการสอนของศาสตราจารย์จำนวน 15 คน โดยนักศึกษาเป็นผู้

ประเมิน ผลการวิจัยพบว่า ไม่มีรูปแบบใดที่สามารถลดอคติได้สำเร็จ สำหรับค่าความเที่ยงของคะแนนเฉลี่ยในการประมาณค่าของแต่ละรูปแบบ พบว่า แบบที่มีสเกลเป็นคำคุณศัพท์ให้ค่าความเที่ยงสูงกว่าเล็กน้อย

Cook (1988: 31-45) ได้ศึกษาผลของการใช้มาตรประมาณค่าที่มีรูปแบบแตกต่างกัน 2 รูปแบบ คือ มาตรประมาณค่าเชิงพฤติกรรม (Behaviorally-Anchored Rating Scales) กับ มาตรประมาณค่าแบบลิเคิร์ต (Likert Scales) และผลของการฝึกทักษะแก่ผู้ประเมินต่อค่าความคลาดเคลื่อนเชิงปล่อยคะแนน (Likert Error) และความคลาดเคลื่อนอื่นเนื่องมาจากการเคยมีปฏิสัมพันธ์มาก่อน (Halo Error) โดยผู้วิจัยได้ใช้มาตรวัด 2 รูปแบบดังกล่าว ให้นักศึกษา 2 กลุ่ม คือ กลุ่มที่ได้รับการฝึกทักษะกับกลุ่มที่ไม่ได้รับการฝึกทักษะ ทำการประเมินพฤติกรรมการสอนของศาสตราจารย์ 1 ท่าน แล้วทำการวิเคราะห์หาค่าเฉลี่ยเลขคณิตและส่วนเบี่ยงเบนมาตรฐานของคะแนนผลการประเมินแต่ละมิติ เปรียบเทียบระหว่างกลุ่มที่ได้รับการฝึกทักษะกับกลุ่มที่ไม่ได้รับการฝึกทักษะ พร้อมทั้งเปรียบเทียบคะแนนผลการประเมินระหว่างมาตรวัดต่างรูปแบบกันเฉพาะในกลุ่มที่ได้รับการฝึกทักษะ โดยในการตรวจสอบ Halo Error ผู้วิจัยได้ใช้ส่วนเบี่ยงเบนมาตรฐานของคะแนนผลการประเมิน (Standard Deviation) เป็นคะแนนบ่งชี้ความคลาดเคลื่อน (Halo Score) ภายใต้เงื่อนไขว่า ถ้าคะแนนความคลาดเคลื่อน (Halo Score) มีค่ามาก แสดงถึงการเกิด Halo Error น้อยกว่ากรณีที่คะแนนความคลาดเคลื่อนมีค่าน้อย ผลการวิจัยพบว่า การฝึกทักษะมีผลต่อการลด Halo Error การประเมินโดยมาตรประมาณค่าเชิงพฤติกรรมเกิด Leniency Error น้อยกว่า มาตรประมาณค่าแบบลิเคิร์ต ในขณะที่มาตรประมาณค่าแบบลิเคิร์ตเกิด Halo Error น้อยกว่า

สุภัทตร์ พิบูลย์ (2534:114) ได้เปรียบเทียบคุณภาพของมาตรประมาณค่าเชิงพฤติกรรมกับมาตรประมาณค่าแบบกราฟิก ในการประเมินพฤติกรรมด้านการประเมินผลของครู ผลการวิจัยพบว่า มาตรประมาณค่าเชิงพฤติกรรมปรากฏลักษณะความคลาดเคลื่อนเชิงปล่อยหรือกดคะแนนและความคลาดเคลื่อนอื่นเนื่องจากปฏิสัมพันธ์ น้อยกว่าการใช้มาตรประมาณค่าแบบกราฟิกเป็นเครื่องมือประเมิน

จากที่กล่าวมา จะเห็นได้ว่า มาตรฐานค่ามีความคลาดเคลื่อนเกิดขึ้นได้ทั้ง จากตัวผู้ประเมิน (Rater) ผู้ได้รับการประเมิน (Ratee) วิธีดำเนินการประเมิน (Procedure) รวมทั้งการใช้เครื่องมือ ซึ่งในการเลือกใช้รูปแบบมาตรฐานค่า (Rating Format) ที่มีคุณภาพ และวิธีการประเมินที่สอดคล้องกับมาตรฐานค่า จะทำให้ผลการวิจัยมีความตรง (Validity) ดังนั้นผู้วิจัยจึงสนใจที่จะศึกษาว่า รูปแบบมาตรฐานค่า (Rating Format) และผู้ประเมินที่แตกต่างกันจะมีความคลาดเคลื่อนของการประเมินมากน้อยเพียงใด เพื่อที่จะได้ข้อค้นพบที่เป็นประโยชน์ต่องานที่วิจัยที่ใช้มาตรฐานค่า (Rating Scale) เป็นเครื่องมือในการเก็บรวบรวมข้อมูล

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย