

## บทที่ 4

### ผลการวิเคราะห์ข้อมูล

การวิจัยนี้มีจุดมุ่งหมายหลักเพื่อ (1) เปรียบเทียบความสอดคล้องของผลการประเมินผล การปฏิบัติงานของครูและความคลาดเคลื่อนแบบกต/ปล่อยคะแนน (leniency error) ระหว่างวิธี การ 2 วิธีคือใช้วิธีป้องกันความคลาดเคลื่อนแบบกต/ปล่อยคะแนนโดยการใช้มาตราแบบอสมมาตร และวิธีการใช้มาตราแบบสมมาตรร่วมกับการประมาณค่าและปรับแก้ความคลาดเคลื่อนแบบกต/ ปล่อยคะแนน (2) เปรียบเทียบความสอดคล้องของผลการประเมินผล การปฏิบัติงานของครูและ ความคลาดเคลื่อนแบบฮาโล (halo error) ระหว่างวิธีการ 2 วิธีคือ วิธีป้องกันความคลาดเคลื่อน แบบฮาโลโดยการใช้มาตราประเมินค่าเชิงพฤติกรรม (BARS) และวิธีการใช้มาตราประเมินค่าแบบ กราฟฟิก (graphic rating scale) ร่วมกับการประมาณค่าและปรับแก้ความคลาดเคลื่อนแบบ ฮาโล และ (3) เปรียบเทียบความสอดคล้องของผลการประเมินผล การปฏิบัติงานของครูและ ความคลาดเคลื่อนแบบกต/ปล่อยคะแนน และความคลาดเคลื่อนแบบฮาโล ระหว่างวิธีการ 3 วิธี คือ วิธีป้องกันความคลาดเคลื่อนแบบกต/ปล่อยคะแนนและแบบฮาโล โดยใช้มาตราประเมินค่าเชิง พฤติกรรม (BARS) แบบอสมมาตร, วิธีการใช้มาตรากราฟฟิก แบบสมมาตร ร่วมกับการ ประมาณค่าและปรับแก้ความคลาดเคลื่อนของ Guilford และวิธีการวิเคราะห์ความแปรปรวนโดย ใช้ทฤษฎีการสรุปอ้างอิง (Generalizability theory)

ผลการวิเคราะห์ข้อมูลแบ่งเป็น 6 ตอนดังนี้ สองตอนแรกเป็นการวิเคราะห์ความแปรปรวน เพื่อตรวจสอบนัยสำคัญของความคลาดเคลื่อนแบบกต/ปล่อยคะแนน และความคลาดเคลื่อนแบบ ฮาโล ตามแนวคิดของ Guilford และ โดยใช้ทฤษฎีการสรุปอ้างอิง ตอนที่ 3 การศึกษาความ สอดคล้องของผลการประเมินเมื่อเทียบเกณฑ์การประเมินของผู้ทรงคุณวุฒิ ตอนที่ 4 การ เปรียบเทียบผลการใช้มาตราอสมมาตรกับการใช้มาตราแบบสมมาตรร่วมกับการประมาณค่าและ ปรับแก้ความคลาดเคลื่อนแบบกต/ปล่อยคะแนน ตอนที่ 5 การเปรียบเทียบผลการใช้มาตราประเมิน ค่าเชิงพฤติกรรม (BARS) และการใช้มาตรากราฟฟิกร่วมกับการประมาณค่าและปรับแก้ ความคลาดเคลื่อนแบบฮาโล และตอนที่ 6 เปรียบเทียบผลการใช้มาตราประเมินค่าเชิงพฤติกรรม แบบอสมมาตรกับการใช้มาตรากราฟฟิกแบบสมมาตรร่วมกับการประมาณค่าและปรับแก้ ความคลาดเคลื่อน

**ตอนที่ 1 การวิเคราะห์ความแปรปรวนเพื่อตรวจสอบความคลาดเคลื่อนแบบกต/ปล่อย  
คะแนนและความคลาดเคลื่อนแบบฮาโล ตามแนวคิดของ Guilford**

จากแนวคิดของ Guilford (1954) ได้นำเสนอวิธีการตรวจสอบความคลาดเคลื่อนแบบกต/ปล่อยคะแนน และความคลาดเคลื่อนแบบฮาโล ด้วยการวิเคราะห์ความแปรปรวนสองทาง (two-way anova) โดยความคลาดเคลื่อนแบบกต/ปล่อยคะแนนพิจารณาได้จากการมีนัยสำคัญทางสถิติระหว่างผู้ประเมิน (rater main effect) และความคลาดเคลื่อนแบบฮาโลพิจารณาได้จากการมีนัยสำคัญทางสถิติของปฏิสัมพันธ์ระหว่างผู้ประเมินกับผู้ถูกประเมิน (raterxratee interaction)

ในงานวิจัยนี้วิเคราะห์โดยจำแนกตามประเภทของมาตร กลุ่มผู้ประเมินและจำนวนผู้ประเมิน ผู้วิจัยขอยกตัวอย่างกรณี ผลการวิเคราะห์เมื่อใช้มาตรประเมินค่าแบบกราฟฟิก (graphic rating scale) แบบสมมาตร กับผู้ประเมินที่มีลักษณะการให้คะแนนแบบปล่อยคะแนน จำนวน 4 คน ชุดที่ 1 (AL1, BL5, CL9, และ DL13) ดังมีรายละเอียดดังนี้

**ตารางที่ 21** การวิเคราะห์ความแปรปรวน กรณีผู้ประเมินปล่อยคะแนน จำนวน 4 คน ชุดที่ 1 โดยมาตรประเมินค่าแบบกราฟฟิก (graphic rating scale) แบบสมมาตร

Source	SS	df	MS	F
rater	9.87	3	3.29	9.07***
ratee	3.6	5	0.72	1.99
raterxratee	10.53	15	0.70	1.94*
within sets	34.80	96	0.36	
total	58.80	119	0.49	

\*  $p < .05$ , \*\*\*  $p < .001$

จากตารางที่ 19 ผลการวิเคราะห์ความแปรปรวนแบบสองทางเมื่อใช้มาตรประเมินค่าแบบกราฟฟิก (graphic rating scale) แบบสมมาตรกับผู้ประเมินที่มีลักษณะการให้คะแนนแบบปล่อยคะแนน จำนวน 4 คน ชุดที่ 1 ความแปรปรวนระหว่างผู้ประเมิน (rater main effect) และความแปรปรวนของปฏิสัมพันธ์ระหว่างผู้ประเมินกับผู้ถูกประเมิน (raterxratee interaction) มีนัยสำคัญ

ทางสถิติที่ระดับ .001 และ ระดับ .05 ตามลำดับ แสดงว่ามีความคลาดเคลื่อนแบบกด/ปล่อย  
คะแนน และมีความคลาดเคลื่อนแบบฮาโล

ส่วนการตรวจสอบความคลาดเคลื่อนแบบกด/ปล่อยคะแนนและความคลาดเคลื่อนแบบ  
ฮาโล ด้วยการวิเคราะห์ความแปรปรวนแบบสองทาง กรณีอื่น ๆ สรุปได้ดังตารางที่ 22



ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 22 สรุปการตรวจสอบความคลาดเคลื่อนแบบกต/ปล่อยคะแนนและความคลาดเคลื่อนแบบฮาโล ด้วยการวิเคราะห์ความแปรปรวนแบบสองทาง

ประเภท มาตร	กลุ่ม ผู้ประเมิน	จำนวน ผู้ประเมิน	กรณี ที่	สมมาตร		อสมมาตร	
				กต/ปล่อย คะแนน	ฮาโล	กต/ปล่อย คะแนน	ฮาโล
BARS	ปล่อย คะแนน	4 คนชุด 1	1	-	-	-	-
		4 คนชุด 2	2	-	-	-	✓
		8 คน	3	-	-	-	-
	กต คะแนน	4 คนชุด 1	4	✓	-	-	-
		4 คนชุด 2	5	-	-	-	-
		8 คน	6	-	-	-	-
graphic	ปล่อย คะแนน	4 คนชุด 1	7	✓	✓	✓	✓
		4 คนชุด 2	8	✓	-	✓	-
		8 คน	9	✓	✓	✓	✓
	กต คะแนน	4 คนชุด 1	10	-	-	-	-
		4 คนชุด 2	11	✓	-	-	-
		8 คน	12	✓	-	-	-
BARS	สนิท	4 คนชุด 1	13	✓	-	✓	-
		4 คนชุด 2	14	✓	-	✓	-
		8 คน	15	✓	-	-	-
	ไม่สนิท	4 คนชุด 1	16	✓	-	✓	-
		4 คนชุด 2	17	✓	-	✓	-
		8 คน	18	✓	-	✓	-
graphic	สนิท	4 คนชุด 1	19	✓	✓	✓	✓
		4 คนชุด 2	20	✓	✓	✓	-
		8 คน	21	✓	✓	✓	✓
	ไม่สนิท	4 คนชุด 1	22	✓	✓	✓	-
		4 คนชุด 2	23	✓	✓	✓	-
		8 คน	24	✓	-	✓	-

หมายเหตุ ✓ คือ พบว่ามีความคลาดเคลื่อน

-- คือ ไม่พบว่ามีความคลาดเคลื่อน

จากตารางที่ 22 จะเห็นได้ว่า เมื่อใช้มาตรประเมินค่าแบบกราฟฟิก (graphic rating scale) แบบสมมาตรจะพบความคลาดเคลื่อนแบบกด/ปล่อยคะแนน ในทุกกลุ่มของผู้ประเมิน เมื่อใช้มาตรประเมินค่าแบบกราฟฟิก (graphic rating scale) แบบอสมมาตรจะไม่พบความคลาดเคลื่อนแบบกด/ปล่อยคะแนนเมื่อผู้ประเมินมีลักษณะการให้คะแนนแบบกดคะแนน หากใช้มาตรประเมินค่าเชิงพฤติกรรม (BARS) ทั้งแบบสมมาตรและแบบอสมมาตรจะไม่พบความคลาดเคลื่อนแบบกด/ปล่อยคะแนน ทั้งในกลุ่มผู้ประเมินที่มีลักษณะการให้คะแนนแบบปล่อยคะแนน และกดคะแนน

เมื่อพิจารณาตามความสนิทของผู้ประเมินกับโรงเรียนอัสสัมชัญจะพบความคลาดเคลื่อนแบบฮาโลเมื่อใช้มาตรประเมินค่าแบบกราฟฟิก (graphic rating scale) กับผู้ประเมินที่สนิทกับโรงเรียนอัสสัมชัญ เท่านั้น

ตารางที่ 22 ที่นำเสนอเป็นผลของการตรวจสอบความคลาดเคลื่อนแบบกด/ปล่อยคะแนน และความคลาดเคลื่อนแบบฮาโลโดยใช้การวิเคราะห์ความแปรปรวนแบบสองทาง โดยวิเคราะห์ 3 ครั้ง ตามข้อเสนอแนะของ Guilford (1954) ทั้งนี้อาจเป็นเพราะเทคโนโลยีการดำเนินการในยุคนั้นยังไม่สะดวกพอที่จะวิเคราะห์ความแปรปรวนแบบสามทาง ผู้วิจัยจึงเห็นว่าการจะวิเคราะห์ความแปรปรวนแบบสามทางจะให้สารสนเทศที่เที่ยงตรงมากกว่าและสามารถวิเคราะห์ได้ไม่ยากในปัจจุบันดังนั้นผู้วิจัยจึงประยุกต์ใช้วิธีตรวจสอบความคลาดเคลื่อนความคลาดเคลื่อนของ Guilford โดยใช้การวิเคราะห์ความแปรปรวนแบบสามทาง

ผู้วิจัยขอยกตัวอย่างกรณีเดียวกันกับที่นำเสนอการวิเคราะห์ความแปรปรวนแบบสองทางเพื่อเปรียบเทียบกัน โดยใช้ผลการวิเคราะห์เมื่อใช้มาตรประเมินค่าแบบกราฟฟิก (graphic rating scale) แบบสมมาตร กับผู้ประเมินที่มีลักษณะการให้คะแนนแบบปล่อยคะแนน จำนวน 4 คน ชุดที่ 1 (AL1, BL5, CL9, และ DL13) ได้ผลการวิเคราะห์ความแปรปรวนแบบสามทางดังนี้

ศูนย์วิจัยที่ปรึกษา  
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 23 การวิเคราะห์ความแปรปรวนแบบสามทาง กรณีผู้ประเมินปล่อยคะแนน จำนวน 4 คน ชุดที่ 1 โดยมาตรประเมินค่าแบบกราฟฟิก (graphic rating scale) แบบสมมาตร

Source	SS	df	MS	F
rater	9.87	3	3.29	16.38***
ratee	3.60	5	0.72	3.59
trait	10.72	4	2.68	13.34
raterxratee	10.53	15	0.70	3.50***
raterxtrait	7.55	12	0.63	3.13
rateextrait	4.48	20	0.22	1.12
raterxrateextrait	12.05	60	0.20	
total	58.80	119	0.49	

\*  $p < .05$ , \*\*\*  $p < .001$

จากตารางที่ 23 ผลการวิเคราะห์ความแปรปรวนแบบสามทางเมื่อใช้ มาตรประเมินค่าแบบกราฟฟิก (graphic rating scale) แบบสมมาตรกับผู้ประเมินที่มีลักษณะการให้คะแนนแบบปล่อยคะแนน จำนวน 4 คน ชุดที่ 1 ความแปรปรวนระหว่างผู้ประเมิน (rater main effect) และความแปรปรวนของปฏิสัมพันธ์ระหว่างผู้ประเมินกับผู้ถูกประเมิน (raterxratee interaction) มีนัยสำคัญทางสถิติที่ระดับ .001 ทั้งคู่ แสดงว่ามีความคลาดเคลื่อนแบบกด/ปล่อยคะแนน และมีความคลาดเคลื่อนแบบฮาโล

ส่วนการตรวจสอบความคลาดเคลื่อนแบบกด/ปล่อยคะแนนและความคลาดเคลื่อนแบบฮาโล ด้วยการวิเคราะห์ความแปรปรวนแบบสามทาง กรณีอื่น ๆ สรุปได้ดังตารางที่ 24

จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 24 สรุปการตรวจสอบความคลาดเคลื่อนแบบกด/ปล่อยคะแนนและความคลาดเคลื่อนแบบฮาโลโดยใช้การวิเคราะห์ความแปรปรวนแบบสามทาง

ประเภท มาตร	กลุ่ม ผู้ประเมิน	จำนวน ผู้ประเมิน	กรณี ที่	สมมาตร		อสมมาตร	
				กด/ปล่อย คะแนน	ฮาโล	กด/ปล่อย คะแนน	ฮาโล
BARS	ปล่อย คะแนน	4 คนชุด 1	1	✓	✓	---	--
		4 คนชุด 2	2	✓	✓	---	✓
		8 คน	3	✓	✓	---	---
	กด คะแนน	4 คนชุด 1	1	✓	✓	--	---
		4 คนชุด 2	2	--	---	---	✓
		8 คน	3	--	---	---	---
graphic	ปล่อย คะแนน	4 คนชุด 1	10	✓	✓	✓	✓
		4 คนชุด 2	11	✓	✓	✓	✓
		8 คน	12	✓	✓	✓	✓
	กด คะแนน	4 คนชุด 1	10	--	✓	--	✓
		4 คนชุด 2	11	✓	--	--	✓
		8 คน	12	✓	✓	✓	--
BARS	สนิท	4 คนชุด 1	13	✓	--	✓	--
		4 คนชุด 2	14	✓	✓	--	--
		8 คน	15	✓	--	✓	---
	ไม่สนิท	4 คนชุด 1	16	✓	--	✓	✓
		4 คนชุด 2	17	✓	--	✓	--
		8 คน	18	✓	---	✓	✓
graphic	สนิท	4 คนชุด 1	19	✓	✓	✓	✓
		4 คนชุด 2	20	✓	✓	✓	✓
		8 คน	21	✓	✓	✓	✓
	ไม่สนิท	4 คนชุด 1	22	✓	✓	✓	---
		4 คนชุด 2	23	✓	✓	✓	--
		8 คน	24	✓	---	✓	---

หมายเหตุ ✓ คือ พบว่ามีความคลาดเคลื่อน

-- คือ ไม่พบว่ามีความคลาดเคลื่อน

เมื่อใช้การวิเคราะห์ความแปรปรวนแบบสามทางจะให้สารสนเทศมากกว่าและตรวจพบความคลาดเคลื่อนได้มากกว่าการวิเคราะห์ความแปรปรวนแบบสองทาง จากตารางที่ 24 พบว่าเมื่อใช้มาตรฐานค่าเชิงพฤติกรรม (BARS) จะพบความคลาดเคลื่อนทั้งสองแบบนี้ต่ำกว่าใช้มาตรฐานประเมินค่าแบบกราฟฟิก (graphic rating scale) โดยมาตร BARS แบบอสมมาตรจะพบความคลาดเคลื่อนน้อยที่สุด

## ตอนที่ 2 การวิเคราะห์ความแปรปรวนโดยทฤษฎีการสรุปอ้างอิง

การวิเคราะห์ความแปรปรวนตามแนวทฤษฎีการสรุปอ้างอิง (Generalizability Theory) ในครั้งนี้ ผู้วิจัยวิเคราะห์โดยใช้โปรแกรมสำเร็จรูป GENOVA (Crick & Brennan, 1983) ซึ่งเป็นโปรแกรมที่พัฒนาขึ้น เพื่อใช้ในการวิเคราะห์ข้อมูลที่เกี่ยวข้องกับทฤษฎีการสรุปอ้างอิง โดยมี 2 ขั้นตอนคือ

2.1 ขั้นการศึกษาเชิงสรุปอ้างอิงหรือขั้นการศึกษา G (Generalizability study or G-study) มีจุดประสงค์เพื่อประมาณค่าความแปรปรวนของคะแนนจริงและความแปรปรวนของคะแนนความคลาดเคลื่อนจากแหล่งความคลาดเคลื่อนต่าง ๆ ที่สนใจ

2.2 ขั้นการศึกษาเพื่อการตัดสินใจหรือขั้นการศึกษา D (decision study or D-study) เป็นการศึกษาความเที่ยงในสถานการณ์ของการวัดและประเมินต่าง ๆ เพื่อใช้เป็นข้อมูลสำหรับวางแผนเพื่อตัดสินใจ

ผู้วิจัยได้ยกตัวอย่างผลการวิเคราะห์ข้อมูล เมื่อมีผู้ประเมิน 16 คนทำการประเมินผลการปฏิบัติงานของครูด้วยแบบประเมินมาตรฐานประเมินค่าแบบกราฟฟิก (graphic rating scale) แบบสมมาตร ซึ่งได้ผลดังนี้

ศูนย์วิทยุทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย



ตารางที่ 25 การวิเคราะห์และการประมาณค่าความแปรปรวนของคะแนนการประเมิน การปฏิบัติงานของครู รูปแบบ  $p \times i \times t$  design ของชั้นการศึกษา G

analysis of variance	df	SS	MS	Estimate Variance Component	Percentage of Total Variance
ratee (P)	5	133.07	26.61	0.31	18.06
rater (I)	15	280.39	18.69	0.53	30.83
trait (T)	4	27.72	6.93	0.05	2.69
rateextrater (PI)	75	94.33	1.26	0.18	10.59
rateextrait (PT)	20	18.56	0.93	0.04	2.11
raterxtrait (IT)	60	115.48	1.92	0.26	15.30
Residual (pit,e)	300	105.04	0.35	0.35	20.42
TOTAL	479	774.59	1.62	1.71	100.00

จากตารางที่ 25 การวิเคราะห์ความแปรปรวนของคะแนนการประเมินการปฏิบัติงานของครู มีแหล่งของความแปรปรวน 7 แหล่ง คือ ผู้ถูกประเมิน (ratee=P) ผู้ประเมิน (rater=I) คุณลักษณะที่ประเมิน (trait=T) ผลร่วมระหว่างผู้ถูกประเมินและผู้ประเมิน (rateeXrater interaction=PI) ผลร่วมระหว่างผู้ถูกประเมินและคุณลักษณะที่ประเมิน (rateeXtrait interaction=PT) ผลร่วมระหว่างผู้ประเมินและคุณลักษณะที่ประเมิน (raterXtrait interaction=IT) พบว่าค่าผลรวมความเบี่ยงเบนกำลังสองของผู้ประเมินมีค่าสูงสุดคือมีค่า 280.39 รองลงมาคือค่าผลรวมความเบี่ยงเบนกำลังสองของผู้ถูกประเมินมีค่า 133.07 และค่าผลรวมความเบี่ยงเบนกำลังสองของผลร่วมระหว่างผู้ประเมินและคุณลักษณะที่ประเมินมีค่าน้อยที่สุดคือ 18.55

สำหรับกำลังสองเฉลี่ยนั้น ค่ากำลังสองเฉลี่ยของผู้ถูกประเมินมีค่าสูงสุดคือมีค่า 26.61 รองลงมาคือค่ากำลังสองเฉลี่ยของผู้ประเมินมีค่า 18.69 และค่ากำลังสองเฉลี่ยของผลร่วมระหว่างผู้ถูกประเมินและผู้ประเมินและคุณลักษณะที่ประเมินมีค่าน้อยที่สุดคือ 0.35

เมื่อเอกภพของการสังเกตประกอบด้วยผู้ถูกประเมิน ผู้ประเมินและคุณลักษณะที่ประเมิน ซึ่งมีขนาดไม่จำกัดจะเป็นการวิเคราะห์ความแปรปรวน 7 แหล่งโดยเป็นการ

ประมาณค่าความแปรปรวนของแหล่งต่าง ๆ เมื่อเทียบกับความแปรปรวนรวมมีค่าเรียงลำดับจากมากไปหาน้อยดังนี้

แหล่งความแปรปรวนที่มีค่ามากที่สุดคือความแปรปรวนของผู้ประเมิน ( $\sigma^2_i$ ) ซึ่งแสดงถึงความแตกต่างของการให้คะแนนประเมินของผู้ประเมินแต่ละคน มีค่า 0.53 หรือประมาณร้อยละ 30.83 ของความแปรปรวนรวม

ความแปรปรวนส่วนที่เหลือ ( $\sigma^2_{pit,e}$ ) ซึ่งแสดงถึงความแปรปรวนของผลร่วมระหว่างผู้ถูกประเมิน ผู้ประเมินและคุณลักษณะที่ประเมิน ความแปรปรวนที่เป็นระบบที่ไม่ได้อยู่ในรูปแบบและความแปรปรวนที่ไม่เป็นระบบ มีค่า 0.35 หรือประมาณร้อยละ 20.42 ของความแปรปรวนรวม

ความแปรปรวนของผู้ถูกประเมิน ( $\sigma^2_p$ ) ซึ่งแสดงถึงความแตกต่างของคะแนนผลการประเมินที่ผู้ถูกประเมินแต่ละคนได้รับมีค่า 0.31 หรือประมาณร้อยละ 18.06 ของความแปรปรวนรวม

ความแปรปรวนของผลร่วมระหว่างผู้ถูกประเมินและคุณลักษณะที่ประเมิน ( $\sigma^2_{pi}$ ) ซึ่งหมายถึงความแตกต่างของคะแนนการประเมินผลการปฏิบัติงานของผู้ถูกประเมินแต่ละคนในแต่ละคุณลักษณะมีค่า 0.26 หรือประมาณร้อยละ 15.30 ของความแปรปรวนรวม

ความแปรปรวนของผลร่วมของผู้ถูกประเมินและผู้ประเมิน ( $\sigma^2_{pi}$ ) ซึ่งแสดงถึงความไม่คงเส้นคงวาของผู้ประเมินแต่ละคนในการประเมินผลการปฏิบัติงานของผู้ถูกประเมินแต่ละคนมีค่า 0.18 หรือประมาณร้อยละ 10.59 ของความแปรปรวนรวม

ความแปรปรวนของคุณลักษณะที่ประเมิน ( $\sigma^2_e$ ) ซึ่งหมายถึงความแตกต่างของคะแนนการประเมินผลการปฏิบัติงานในแต่ละคุณลักษณะมีค่า 0.05 หรือประมาณร้อยละ 2.69 ของความแปรปรวนรวม

ความแปรปรวนที่มีค่าน้อยที่สุดคือผลร่วมระหว่างผู้ประเมินและคุณลักษณะที่ประเมิน ( $\sigma^2_{ie}$ ) ซึ่งหมายถึงความแตกต่างของคะแนนที่ผู้ถูกประเมินแต่ละคนให้ในการประเมินในแต่ละคุณลักษณะมีค่า 0.07 หรือประมาณร้อยละ 4.78 ของความแปรปรวนรวม

เมื่อนำข้อมูลนี้มาวิเคราะห์และการประมาณค่าสัมประสิทธิ์การอ้างอิงสรุปในชั้นการศึกษา D ได้ผลดังตารางที่ 26

ตารางที่ 26 การวิเคราะห์และการประมาณค่า ชั้นการศึกษา D ค่าสัมประสิทธิ์การสุรปร่างเชิงการ  
ประเมินการปฏิบัติงานของครู รูปแบบ  $p \times i \times t$  design

analysis of variance	Estimate Variance Component	$n'_i=4$	8	16
		$n'_i=5$	5	5
ratee (p)	0.31	0.31	0.31	0.31
rater (i)	0.53	0.13	0.07	0.033
trait (t)	0.05	0.009	0.0092	0.0092
ratexrater (pi)	0.18	0.045	0.023	0.011
rateextrait (pt)	0.04	0.007	0.007	0.007
raterextrait (it)	0.26	0.013	0.007	0.003
Residual (pit,e)	0.35	0.018	0.009	0.004
TOTAL	1.71			
$\sigma^2_{\epsilon}$	0.568	0.0701	0.0387	0.0229
$\sigma^2_{\Delta}$	1.405	0.2246	0.1205	0.0685
$\hat{\rho}^2_{\epsilon}$	0.353	0.8154	0.8890	0.9310
$\hat{\rho}^2_{\Delta}$	0.181	0.5796	0.7199	0.8189

จากตารางที่ 26 ผลของการศึกษา D ประกอบด้วยการประมาณค่าความแปรปรวนของจำนวนผู้ประเมินจำนวนหลากหลาย ตั้งแต่ 4 คน, 8 คน และ 16 คน โดยความแปรปรวนจากแหล่งต่าง ๆ มีค่าลดลงเมื่อจำนวนผู้ประเมินเพิ่มขึ้น ความแปรปรวนของคะแนนแบบสัมพัทธ์ ( $\sigma^2_{\epsilon}$ ) มีค่าลดลงเมื่อจำนวนผู้ประเมินเพิ่มขึ้น โดยมีค่า 0.0701, 0.0387 และ 0.0229 เมื่อมีผู้ประเมินจำนวน 4 คน 8 คน และ 16 คน ตามลำดับ เช่นเดียวกับ ความแปรปรวนของคะแนนแบบสัมบูรณ์ ( $\sigma^2_{\Delta}$ ) มีค่าลดลงเมื่อจำนวนผู้ประเมินเพิ่มขึ้น โดยมีค่า 0.2246, 0.1205 และ 0.0685 เมื่อมีผู้ประเมินจำนวน 4 คน 8 คน และ 16 คน ตามลำดับ

ค่าสัมประสิทธิ์การอ้างอิงสุรมีค่าเพิ่มขึ้น เมื่อจำนวนผู้ประเมินเพิ่มขึ้น โดยค่าสัมประสิทธิ์การอ้างอิงสุรสำหรับการตัดสินใจแบบสัมพัทธ์ ( $\hat{\rho}^2_{\epsilon}$ ) มีค่า 0.8154, 0.8890 และ

0.9310 ค่าสัมประสิทธิ์การอ้างอิงสรุปสำหรับการตัดสินใจแบบสัมบูรณ์ ( $\rho^2_{\Delta}$ ) มีค่า 0.5796, 0.7199 และ 0.8189 เมื่อมีผู้ประเมินจำนวน 4 คน 8 คน และ 16 คน ตามลำดับ จะเห็นได้ว่าจากข้อมูลนี้ ค่าความเที่ยงหรือค่าสัมประสิทธิ์การอ้างอิงสรุปมีค่าเพิ่มขึ้นเมื่อจำนวนผู้ประเมินเพิ่มขึ้น แต่ค่าที่เพิ่มขึ้นไม่แตกต่างกันนัก เมื่อเป็นค่าสัมประสิทธิ์การอ้างอิงสรุปสำหรับการตัดสินใจแบบสัมพัทธ์ ( $\rho^2_{\delta}$ ) ในกรณีของค่าสัมประสิทธิ์การอ้างอิงสรุปสำหรับการตัดสินใจแบบสัมบูรณ์ ( $\rho^2_{\Delta}$ ) เมื่อเพิ่มผู้ประเมินจากจำนวน 4 คน เป็นจำนวน 8 คน ค่า  $\rho^2_{\Delta}$  เพิ่มขึ้นมากคือจาก 0.5796 เป็น 0.7199 แต่เมื่อเทียบกรณีผู้ประเมินจำนวน 8 คน และ 16 คน ค่าไม่ต่างกันนัก ดังนั้นหากการประเมินโดยให้ผู้ประเมินจำนวน 8 คนมีความสะดวกในทางปฏิบัติมากกว่าก็ควรจะให้ผู้ประเมินจำนวน 8 คนก็เพียงพอแล้ว

ผู้วิจัยได้ทำการเปรียบเทียบกรณีที่ประเมินผลการปฏิบัติงานของครูโดยให้ผู้ประเมิน 16 คนเมื่อใช้แบบประเมินแบบมาตราประเมินค่าแบบกราฟฟิก (graphic rating scale) แบบสมมาตร และแบบประเมินค่าเชิงพฤติกรรม (BARS) แบบสมมาตร ซึ่งสรุปได้ดังตารางที่ 27 ตารางที่ 27 ผู้ประเมิน 16 คน ประเมินผลการปฏิบัติงานของครูด้วยแบบประเมินมาตรกราฟฟิก ( graphic scale ) แบบสมมาตร

analysis of variance	df	ss	ms	$n'_i$	1	4	8	16	4	8	16
				$n'_T$	1	5	5	5	10	10	10
ratee (p)	5	134.12	26.82	$\sigma^2_p$	0.30	0.30	0.30	0.30	0.30	0.30	0.30
rater(I)	15	271.39	18.09	$\sigma^2_i$	0.56	0.14	0.07	0.035	0.140	0.070	0.035
trait(T)	4	46.55	11.64	$\sigma^2_T$	0.11	0.021	0.0211	0.0211	0.0106	0.0106	0.0106
rateexrater(PI)	75	98.48	1.31	$\sigma^2_{pi}$	0.20	0.051	0.025	0.013	0.051	0.025	0.013
rateextrait(PT)	20	29.68	1.48	$\sigma^2_{pT}$	0.07	0.015	0.015	0.015	0.007	0.007	0.007
raterextrait(IT)	60	19.32	0.3	$\sigma^2_{iT}$	0.00	0.000	0.000	0.000	0.000	0.000	0.000
P X I X T	300	90.06	0.30	$\sigma^2_{piT}$	0.30	0.015	0.008	0.004	0.008	0.004	0.002
TOTAL	479	689.59	1.44								
Relative error variance				$\sigma^2_{\delta}$	g-re	0.08	0.048	0.0311	0.066	0.04	0.02
Absolute error variance				$\sigma^2_{\Delta}$	g-ab	0.24	0.139	0.087	0.216	0.117	0.07
G-coefficient for relative decision				$\rho^2_{\delta}$	co-re	0.79	0.865	0.907	0.823	0.893	0.93
G-coefficient for absolute decision				$\rho^2_{\delta}$	co-ab	0.56	0.69	0.78	0.59	0.729	0.82

จากตารางที่ 27 การวิเคราะห์ความแปรปรวนด้วย G-theory โดยใช้ข้อมูลผู้ประเมินทั้ง 16 คนใช้มาตรแบบสมมาตร ในการศึกษา D-study เมื่อเพิ่มจำนวนผู้ประเมิน (rater) เป็น 4 คน, 8 คน และ 16 คน จะทำให้ความคลาดเคลื่อนของผลการประเมิน (error variance) ลดน้อยลง โดยความคลาดเคลื่อนแบบสัมพัทธ์ ( $\sigma^2_e$ ) จะมีค่าเป็น .0701, .0387 และ .0229 ตามลำดับ สำหรับมาตรประเมินค่าแบบกราฟฟิก (graphic rating scale) แบบสมมาตร และมีค่า .028, .018 และ .012 ตามลำดับ สำหรับมาตรประมาณค่าเชิงพฤติกรรม (BARS) แบบสมมาตร

ส่วนความคลาดเคลื่อนแบบสัมบูรณ์ ( $\sigma^2_\Delta$ ) ของมาตรประเมินค่าแบบกราฟฟิก (graphic rating scale) แบบสมมาตร มีค่าเท่ากับ .0701, .0387 และ .0229 ส่วนมาตรประมาณค่าเชิงพฤติกรรม (BARS) แบบสมมาตร มีค่าความคลาดเคลื่อนแบบสัมบูรณ์ เท่ากับ .111, .066 และ .044 เมื่อจำนวน ผู้ประเมิน เป็น 4 คน, 8 คน และ 16 คน ตามลำดับ จะเห็นได้ว่ามาตรประมาณค่าเชิงพฤติกรรม (BARS) แบบสมมาตรมีความคลาดเคลื่อนของผลการประเมินน้อยกว่า มาตรประเมินค่าแบบกราฟฟิก (graphic rating scale) แบบสมมาตร

ความเที่ยงหรือค่าสัมประสิทธิ์การสรุปอ้างอิงในการประเมิน (generalizability coefficient) เมื่อเพิ่มจำนวนผู้ประเมินมากขึ้นเป็น 4 คน, 8 คน และ 16 คน ค่าสัมประสิทธิ์การสรุปอ้างอิงในการประเมินจะเพิ่มขึ้นตามลำดับดังนี้ มาตรประเมินค่าแบบกราฟฟิก (graphic rating scale) แบบสมมาตร ค่าสัมประสิทธิ์การสรุปอ้างอิงแบบสัมพัทธ์ ( $\rho^2_e$ ) มีค่าเท่ากับ .8154, .8890 และ .9310 ค่าสัมประสิทธิ์การสรุปอ้างอิงแบบสัมบูรณ์ ( $\rho^2_\Delta$ ) มีค่าเท่ากับ .5796, .7199 และ .8189 มาตรประมาณค่าเชิงพฤติกรรม (BARS) แบบสมมาตร ค่าสัมประสิทธิ์การสรุปอ้างอิงแบบสัมพัทธ์ ( $\rho^2_e$ ) มีค่าเท่ากับ .624, .727 และ .792 ค่าสัมประสิทธิ์การสรุปอ้างอิงแบบสัมบูรณ์ ( $\rho^2_\Delta$ ) มีค่าเท่ากับ .299, .415 และ .516 ตามลำดับ จะเห็นได้ว่ามาตรประเมินค่าแบบกราฟฟิก (graphic rating scale) แบบสมมาตร มีความเที่ยง (ค่าสัมประสิทธิ์การสรุปอ้างอิงในการประเมิน) สูงกว่ามาตรประเมินค่า เชิงพฤติกรรม (BARS) แบบสมมาตร

ในกรณีมาตรอสมมาตร ผู้วิจัยแยกวิเคราะห์ความแปรปรวนตามลักษณะการให้คะแนน (กต/ปล่อยคะแนน) ของผู้ประเมินโดยใช้ข้อมูลในการวิเคราะห์ดังนี้ (1) เมื่อผู้ประเมินที่ปล่อยคะแนน 8 คนประเมินด้วยมาตรประเมินค่าแบบกราฟฟิก (graphic rating scale) แบบอสมมาตร (2) เมื่อผู้ประเมินที่กตคะแนน 8 คนประเมินด้วยมาตรประเมินค่าแบบกราฟฟิก (graphic rating scale) แบบอสมมาตร (3) เมื่อผู้ประเมินที่ปล่อยคะแนน 8 คนประเมินด้วยมาตรประมาณค่าเชิงพฤติกรรม (BARS) แบบอสมมาตร (4) เมื่อผู้ประเมินที่กตคะแนน 8 คนประเมินด้วยมาตรประมาณค่าเชิงพฤติกรรม (BARS) แบบอสมมาตร ผลดังตารางที่ 28 และแยกตามความสนิท/ไม่

สนิทธระหว่างผู้ประเมินกับโรงเรียนอัสสัมชัญแบ่งเป็น (1) ผู้ประเมินที่สนิทกับโรงเรียนอัสสัมชัญ 8 คนประเมินด้วยมาตราประเมินค่าแบบกราฟฟิก (graphic rating scale) แบบอสมมาตร (2) เมื่อผู้ประเมินที่ไม่สนิทกับโรงเรียนอัสสัมชัญ 8 คนประเมินด้วยมาตราประเมินค่าแบบกราฟฟิก (graphic rating scale) แบบอสมมาตร (3) ผู้ประเมินที่สนิทกับโรงเรียนอัสสัมชัญ 8 คนประเมินด้วยมาตราประมาณค่าเชิงพฤติกรรม (BARS) แบบอสมมาตร (4) ผู้ประเมินที่ไม่สนิทกับโรงเรียน อัสสัมชัญ 8 คนประเมินด้วยมาตราประมาณค่าเชิงพฤติกรรม (BARS) แบบอสมมาตร ผลดังตาราง ที่ 29

ตารางที่ 28 ผู้ประเมิน 16 คน ประเมินผลการปฏิบัติงานของครูด้วยแบบประเมินแบบมาตรากราฟฟิก ( graphic scale ) แบบอสมมาตร

analysis of variance	df	ss	ms	$n'_i$	1	4	8	16	4	8	16
				$n'_T$	1	5	5	5	10	10	10
ratee (p)	5	29.54	5.91	$\sigma^2_p$	0.06	0.06	0.06	0.06	0.06	0.06	0.06
rater(I)	15	386.93	25.80	$\sigma^2_i$	0.82	0.21	0.10	0.051	0.205	0.103	0.051
trait(T)	4	11.97	2.99	$\sigma^2_T$	0.01	0.003	0.0029	0.0029	0.0014	0.0014	0.0014
rateextrater(PI)	75	22.99	0.31	$\sigma^2_{pi}$	0.03	0.007	0.003	0.002	0.007	0.003	0.002
rateextrait(PT)	20	15.35	0.77	$\sigma^2_{pT}$	0.04	0.007	0.007	0.007	0.004	0.004	0.004
raterextrait(IT)	60	60.78	1.01	$\sigma^2_{iT}$	0.14	0.007	0.003	0.002	0.003	0.002	0.001
P X I X T	300	52.78	0.18	$\sigma^2_{piT}$	0.18	0.009	0.004	0.002	0.004	0.002	0.001
TOTAL	479	580.35	1.21								
Relative error variance				$\sigma^2_\delta$	0.239	0.023	0.015	0.011	0.015	0.009	0.006
Absolute error variance				$\sigma^2_\Delta$	1.215	0.238	0.124	0.067	0.225	0.115	0.060
G-coefficient for relative decision				$\rho^2_\delta$	0.208	0.734	0.806	0.848	0.811	0.872	0.907
G-coefficient for absolute decision				$\rho^2_\delta$	0.049	0.208	0.335	0.482	0.218	0.352	0.510

ตารางที่ 29 ผู้ประเมิน 16 คน ประเมินผลการปฏิบัติงานของครูด้วยแบบมาตราประเมินค่าเชิง  
พฤติกรรม (BARS) แบบสมมาตร

analysis of variance	df	ss	ms	$n'_i$	1	4	8	16	4	8	16
				$n'_T$	1	5	5	5	10	10	10
ratee (p)	5	23.36	4.67	$\sigma^2_p$	0.046	0.046	0.046	0.046	0.046	0.046	0.046
rater(I)	15	127.88	8.53	$\sigma^2_i$	0.26	0.06	0.03	0.016	0.065	0.032	0.016
trait(T)	4	33.65	8.41	$\sigma^2_T$	0.08	0.015	0.0150	0.0150	0.0075	0.0075	0.0075
rateextrater(PI)	75	32.61	0.43	$\sigma^2_{pi}$	0.03	0.008	0.004	0.002	0.008	0.004	0.002
rateextrait(PT)	20	17.05	0.85	$\sigma^2_{pT}$	0.04	0.007	0.007	0.007	0.004	0.004	0.004
raterextrait(IT)	60	37.21	0.62	$\sigma^2_{iT}$	0.06	0.003	0.001	0.001	0.001	0.001	0.000
P X I X T	300	84.49	0.28	$\sigma^2_{piT}$	0.28	0.014	0.007	0.004	0.007	0.004	0.002
TOTAL	479	356.25	0.74								
Relative error variance				$\sigma^2_s$	0.348	0.029	0.018	0.013	0.018	0.011	0.007
Absolute error variance				$\sigma^2_\Delta$	0.738	0.111	0.067	0.044	0.092	0.051	0.031
G-coefficient for relative decision				$\hat{p}^2_s$	0.116	0.614	0.718	0.785	0.715	0.808	0.864
G-coefficient for absolute decision				$\hat{p}^2_s$	0.058	0.292	0.407	0.508	0.333	0.471	0.594

จากตารางที่ 28 และ 29 จะเห็นได้ว่าผลการวิเคราะห์สอดคล้องกับตารางที่ 23 คือในการวิเคราะห์สรุปอ้างอิงเมื่อเพิ่มจำนวนผู้ประเมินมากขึ้นเป็น 4 คน, 8 คน และ 16 คน ความคลาดเคลื่อนของผลการประเมิน (error variance) ทั้งความคลาดเคลื่อนแบบสัมพัทธ์ ( $\sigma^2_s$ ) และ ความคลาดเคลื่อนแบบสัมบูรณ์ ( $\sigma^2_\Delta$ ) จะมีค่าลดน้อยลง และมีความเที่ยงหรือค่าสัมประสิทธิ์การสรุปอ้างอิงในการประเมิน (generalizability coefficient  $\hat{p}^2_s$  และ  $\hat{p}^2_\Delta$ ) ทั้งค่าสัมประสิทธิ์การสรุปอ้างอิงแบบสัมพัทธ์ ( $\hat{p}^2_s$ ) และค่าสัมประสิทธิ์การสรุปอ้างอิงแบบสัมบูรณ์ ( $\hat{p}^2_\Delta$ ) เพิ่มมากขึ้น นอกจากนี้ยังพบว่าเมื่อใช้มาตราประเมินค่าแบบกราฟฟิก (graphic rating scale) ความคลาดเคลื่อนแบบกต/ปล่อยคะแนน ( $\sigma^2_p$ ) และความคลาดเคลื่อนแบบฮาโล ( $\sigma^2_{pi}$ ) มากกว่าเมื่อประเมิน ด้วยมาตราประเมินค่าเชิงพฤติกรรม (BARS) และเมื่อใช้มาตราแบบสมมาตรจะมีความคลาดเคลื่อนทั้งสองแบบมากกว่าเมื่อใช้มาตราแบบอสมมาตร หากจำแนกผู้ประเมินตามภูมิภาค (สนิท/ไม่สนิทกับโรงเรียนอัสสัมชัญ) พบว่ากลุ่มผู้ประเมินที่สนิทกับโรงเรียนอัสสัมชัญ เมื่อใช้มาตราประเมินค่าเชิงพฤติกรรม (BARS) แบบอสมมาตรจะมีความคลาดเคลื่อนทั้งแบบกต/ปล่อย

คะแนนและ ความคลาดเคลื่อนแบบฮาโลน้อยที่สุด

ค่าความเที่ยงหรือค่าสัมประสิทธิ์การสรุปอ้างอิง เมื่อใช้ มาตรฐานประเมินค่าแบบกราฟฟิก (graphic rating scale) จะมีค่าสูงกว่าเมื่อใช้มาตรฐานประเมินค่าเชิงพฤติกรรม (BARS) ทั้งมาตรฐานแบบสมมาตรและมาตรฐานแบบอสมมาตร เมื่อจำแนกผู้ประเมินตามลักษณะการให้คะแนน แบบกด/ปล่อยคะแนน หากจำแนกผู้ประเมินตามภูมิหลัง (สนิท/ไม่สนิทกับโรงเรียนอัสสัมชัญ) ค่าสัมประสิทธิ์การสรุปอ้างอิงแบบสัมพัทธ์ ( $\rho^2_r$ ) จะมีค่ามากที่สุด หากใช้มาตรฐานประเมินค่าแบบกราฟฟิก (graphic rating scale) ในกลุ่มผู้ประเมินที่ไม่สนิทกับโรงเรียนอัสสัมชัญ โดยมีค่าเท่ากับ .779, .861 และ .909 เมื่อมีจำนวนผู้ประเมิน เป็น 4 คน, 8 คน และ 16 คน ตามลำดับ ค่าสัมประสิทธิ์การสรุปอ้างอิงแบบสัมบูรณ์ ( $\rho^2_a$ ) จะมีค่ามากที่สุดเมื่อใช้มาตรฐานประเมินค่าเชิงพฤติกรรม (BARS) ในกลุ่มผู้ประเมินที่สนิทกับโรงเรียนอัสสัมชัญ ได้ .468, .511 และ .536 เมื่อมีจำนวนผู้ประเมิน เป็น 4 คน, 8 คน และ 16 คน ตามลำดับ

เนื่องจากหลักการของ G-theory และ วิธี Guilford ใช้หลักการเดียวกันและผลจากการวิเคราะห์ ความแปรปรวนของทั้งสองวิธีมีค่าเท่ากันต่างกันในระดับทศนิยม (รายละเอียดภาคผนวก) ดังนั้นเมื่อนำมาประมาณค่าและปรับแก้ความคลาดเคลื่อนจึงไม่แตกต่างกัน ดังนั้นผู้วิจัยจึงไม่นำเสนอผลประมาณค่าและปรับแก้ความคลาดเคลื่อนด้วยวิธี G-theory

### ตอนที่ 3 การศึกษาความสอดคล้องของผลการประเมินเมื่อเทียบเกณฑ์การประเมินของผู้ทรงคุณวุฒิ

#### 3.1 เปรียบเทียบความสอดคล้องของผลการประเมินตามสูตรของ Downie และ Heath

การวิจัยในครั้งนี้ ผู้วิจัยได้เปรียบเทียบความสอดคล้องของผลการประเมินโดยใช้การหาขนาดความสอดคล้องของคะแนนที่ได้จาก ผู้ประเมินมากกว่า 2 คนตามสูตรของ Downie และ Heath (Downie and Heath 1974 :215)

$$r_{xx} = \frac{MS_r - MS_E}{MS_r + (k - 1)MS_E} \quad \text{เมื่อ } k = \text{จำนวนผู้ประเมิน}$$

ผู้วิจัยขอยกตัวอย่างการคำนวณด้วยข้อมูลการประเมินที่ใช้มาตรฐานประเมินค่าแบบกราฟฟิกแบบสมมาตร ดังมีรายละเอียดดังนี้



ตัวอย่าง ผลการประเมินการปฏิบัติงานของครู ของผู้ทรงคุณวุฒิ 3 คน เมื่อใช้มาตราประเมิน  
ค่าแบบกราฟฟิก แบบสมมาตร

rater	ผู้ประเมิน			รวม	ยกกำลังสอง			รวม
	ผู้ทรง คุณวุฒิ 1	ผู้ทรง คุณวุฒิ 2	ผู้ทรง คุณวุฒิ 3		ผู้ทรง คุณ วุฒิ 1	ผู้ทรง คุณวุฒิ 2	ผู้ทรง คุณวุฒิ 3	
ratee	1	2	3		1	2	3	
ratee1	2	2.5	1.4	5.9	4	6.3	1.9	12.21
ratee2	3.2	3.8	4.3	11.3	10.3	14.4	18.1	42.74
ratee3	3.5	4	3.5	11	12.3	16.0	12.3	40.50
ratee4	4.2	3.8	4	12	17.6	14.4	16.0	48.08
ratee5	3.5	4.4	4.2	12.1	12.3	19.4	17.6	49.25
ratee6	4.2	4.4	4.5	13.1	17.6	19.4	20.3	57.25
รวม	20.6	22.9	21.9	$\sum X=65.35$	74.0	89.9	86.2	$\sum X^2 =250.0$

จากข้อมูลข้างต้น นำมาคำนวณดังนี้

$$SS_T = \sum X^2 - \frac{(\sum X)^2}{N}$$

$$= 250.03 - \frac{(65.35)^2}{18}$$

$$= 12.78$$

$$SS_r = \frac{(5.9)^2 + (11.3)^2 + (11)^2 + (12)^2 + (12.1)^2 + (13.1)^2}{3} - \frac{(65.35)^2}{18}$$

$$= 10.87$$

$$SS_e = \frac{(20.6)^2 + (22.9)^2 + (21.9)^2}{6} - \frac{(65.35)^2}{18}$$

$$= .44$$

$$SS_E = SS_T - (SS_r + SS_e)$$

$$= 12.78 - (10.87 + 0.44)$$

$$= 1.46$$

เขียนในรูปการวิเคราะห์ความแปรปรวนได้ผลดังนี้

Source of Variation	df	SS	MS	F
ratee	5	10.87	2.17	14.86***
expert	2	.44	.22	1.51
Error	10	1.46	.15	
total	17	36.28		

\*\*\*  $p < .001$ , \*  $p < .05$

จากการวิเคราะห์ ความแปรปรวนข้างต้น พบว่าผู้ถูกประเมิน (ratee) แต่ละคนมีผลการปฏิบัติงานแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .001 และผู้ทรงคุณวุฒิทั้ง 3 คนให้ผลการประเมินสอดคล้องกัน สามารถคำนวณหาขนาดของความเที่ยงในรูปความสอดคล้องระหว่างผู้ทรงคุณวุฒิทั้ง 3 คน และสำหรับผู้ทรงคุณวุฒิคนเดียว ได้ดังนี้

ขนาดของความสอดคล้องของผู้ทรงคุณวุฒิ  
ทั้ง 3 คน

$$r_{tt} = \frac{MS_r - MS_E}{MS_r}$$

$$r_{tt} = \frac{2.17 - 0.146}{2.17}$$

$$= .93$$

ขนาดของความสอดคล้องเฉลี่ยของ  
ผู้ทรงคุณวุฒิแต่ละคน

$$r_{xx} = \frac{MS_r - MS_E}{MS_r + (k - 1)MS_E}$$

$$r_{xx} = \frac{2.17 - 0.146}{2.17 + (3 - 1)0.146}$$

$$= .82$$

ผลการคำนวณข้างต้นพบว่า ขนาดของความสอดคล้องของผู้ทรงคุณวุฒิ ทั้ง 3 คน (reliability of all rater taken together) เท่ากับ .93 และขนาดของความสอดคล้องเฉลี่ยของผู้ทรงคุณวุฒิแต่ละคน (reliability of a singer rater's ratings) เท่ากับ .82

ขนาดของความสอดคล้องเฉลี่ยของผู้ทรงคุณวุฒิทั้ง 3 คน กรณีที่ใช้เครื่องมือการวิจัยฉบับอื่น ๆ สรุปได้ดังต่อไปนี้

ขนาดความสอดคล้องในการประเมินของผู้ทรงคุณวุฒิทั้ง 3 คน

ประเภทมาตร	ลักษณะมาตร	$r_{xx}$
Graphic	สมมาตร	.82
	อสมมาตรทางบวก (ปล่อยคะแนน)	.80
	อสมมาตรทางลบ (กดยคะแนน)	.84
BARS	สมมาตร	.87
	อสมมาตรทางบวก (ปล่อยคะแนน)	.85
	อสมมาตรทางลบ (กดยคะแนน)	.84

นอกจากนี้ผู้วิจัยยังได้ใช้สูตรของ Downie และ Heath (1974) ในการพิจารณาความสอดคล้องของผลการประเมินของผู้ประเมินแต่ละท่านเมื่อเทียบกับผลการประเมินของผู้ทรงคุณวุฒิทั้ง 3 คน โดยผู้วิจัยขอยกตัวอย่างการพิจารณาความสอดคล้องของผลการประเมินของผู้ประเมิน AL1 กับผู้ทรงคุณวุฒิทั้ง 3 คน เมื่อใช้มาตรประเมินค่าแบบกราฟฟิก แบบสมมาตรในการประเมิน ดังมีรายละเอียดดังนี้

ตัวอย่าง ผลการประเมินการปฏิบัติงานของครูของผู้ทรงคุณวุฒิ 3 คนกับผู้ประเมิน AL1 เมื่อใช้มาตรประเมินค่าแบบกราฟฟิก แบบสมมาตร

	ผู้ประเมิน				รวม	ยกกำลังสอง				รวม
	ผู้ทรงคุณวุฒิ	ผู้ทรงคุณวุฒิ	ผู้ทรงคุณวุฒิ	ผู้ประเมิน		ผู้ทรงคุณวุฒิ	ผู้ทรงคุณวุฒิ	ผู้ทรงคุณวุฒิ	ผู้ประเมิน	
	1	2	3	AL1		1	2	3	AL1	
ratee1	2	2.5	1.4	4	9.9	4	6.3	1.9	16.0	28.2
ratee2	3.2	3.8	4.3	4.8	16.1	10.3	14.4	18.1	23.0	65.8
ratee3	3.5	4	3.5	5	16	12.3	16.0	12.3	25.0	65.5
ratee4	4.2	3.8	4	4.8	16.8	17.6	14.4	16.0	23.0	71.1
ratee5	3.5	4.4	4.2	5	17.1	12.3	19.4	17.6	25.0	74.3
ratee6	4.2	4.4	4.5	5	18.1	17.6	19.4	20.3	25.0	82.25
รวม	20.6	22.9	21.9	28.6	$\sum X = 93.95$	74.0	89.9	86.2	137.1	$\sum X^2 = 387.1$

$$\begin{aligned}
 SS_T &= \sum X^2 - \frac{(\sum X)^2}{N} \\
 &= 387.1 - \frac{(93.95)^2}{24} \\
 &= 19.34
 \end{aligned}$$

$$SS_r = \frac{(11)^2 + (16.1)^2 + (16)^2 + (16.8)^2 + (17.1)^2 + (18.1)^2}{4} - \frac{(93.95)^2}{24}$$

$$= 10.69$$

$$SS_e = \frac{(20.6)^2 + (22.9)^2 + (21.9)^2 + (28.6)^2}{6} - \frac{(93.95)^2}{24}$$

$$= 6.25$$

$$SS_E = SS_r - (SS_r + SS_e)$$

$$= 19.34 - (10.69 + 6.25)$$

$$= 2.39$$

Source of Variation	df	SS	MS	F
ratee	5	10.69	2.14	13.38***
rater	3	6.25	2.08	13***
Error	15	2.39	0.16	
total	23	36.28		

\*\*\* p<.001

จากการวิเคราะห์ ความแปรปรวนข้างต้น พบว่าผู้ถูกประเมิน (ratee) แต่ละคนมีผลการปฏิบัติงานแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .001 และผู้ประเมิน AL1 กับผู้ทรงคุณวุฒิทั้ง 3 คนให้ผลการประเมินไม่สอดคล้องกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .001 เมื่อพิจารณาขนาดของความสอดคล้องได้ผลดังนี้

ขนาดของความสอดคล้องของผู้ทรงคุณวุฒิ  
และผู้ประเมิน AL1

$$r_{tt} = \frac{MS_r - MS_E}{MS_r}$$

$$r_{tt} = \frac{2.14 - 0.16}{2.14}$$

$$= .92$$

ขนาดของความสอดคล้องเฉลี่ยของ  
ผู้ทรงคุณวุฒิ และผู้ประเมิน AL1

$$r_{xx} = \frac{MS_r - MS_E}{MS_r + (k - 1)MS_E}$$

$$r_{xx} = \frac{2.14 - 0.16}{2.14 + (4 - 1)0.16}$$

$$= .75$$

จะเห็นได้ว่าเมื่อใช้มาตรกราฟฟิกแบบสมมาตร พิจารณารายการของความสอดคล้องของผู้ทรงคุณวุฒิ และผู้ประเมิน AL1 มีค่าเท่ากับ .92 ขนาดของความสอดคล้องเฉลี่ยของผู้ทรงคุณวุฒิ และผู้ประเมิน AL1 มีค่าเท่ากับ .75 ซึ่งน้อยกว่าขนาดของความสอดคล้องเฉลี่ยของผู้ทรงคุณวุฒิ 3 คน ซึ่งมีค่า .82 แสดงว่าค่าที่ลดลงเนื่องมาจากผลการประเมินของผู้ประเมิน AL1 กรณีผู้ประเมินคนอื่น ๆ และกรณีการใช้เครื่องมือฉบับอื่น ๆ ในการประเมิน สรุปได้ดังนี้



ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 30 ขนาดความสอดคล้องของผลการประเมินของผู้ประเมินแต่ละท่าน ( $r_{xx}$ ) เมื่อเทียบกับ

เกณฑ์การประเมินของผู้ทรงคุณวุฒิ

ผู้ประเมิน	มาตรฐานกราฟฟิก			มาตรฐาน BAR		
	สมมาตร	อสมมาตร ทางบวก	อสมมาตร ทางลบ	สมมาตร	อสมมาตร ทางบวก	อสมมาตร ทางลบ
ผู้ทรงคุณวุฒิทั้ง 3 ท่าน	.82	.80	.84	.87	.85	.84
ผู้ทรงคุณวุฒิทั้ง 3 ท่าน และผู้ประเมิน AL1	*.75	*.79		*.81	.82	
ผู้ทรงคุณวุฒิทั้ง 3 ท่าน และผู้ประเมิน AL2	*.77	.80		*.82	.86	
ผู้ทรงคุณวุฒิทั้ง 3 ท่าน และผู้ประเมิน AS3	*.76		*.77	*.78		.80
ผู้ทรงคุณวุฒิทั้ง 3 ท่าน และผู้ประเมิน AS4	*.78		*.79	*.80		.82
ผู้ทรงคุณวุฒิทั้ง 3 ท่าน และผู้ประเมิน BL5	*.70	*.77		*.82	.85	
ผู้ทรงคุณวุฒิทั้ง 3 ท่าน และผู้ประเมิน BL6	*.69	*.75		*.79	.87	
ผู้ทรงคุณวุฒิทั้ง 3 ท่าน และผู้ประเมิน BS7	*.73		*.79	*.82		.83
ผู้ทรงคุณวุฒิทั้ง 3 ท่าน และผู้ประเมิน BS8	*.76		*.76	.83		.83
ผู้ทรงคุณวุฒิทั้ง 3 ท่าน และผู้ประเมิน CL9	*.78	*.79		.85	.86	
ผู้ทรงคุณวุฒิทั้ง 3 ท่าน และผู้ประเมิน CL10	*.79	.80		.84	.84	
ผู้ทรงคุณวุฒิทั้ง 3 ท่าน และผู้ประเมิน CS11	*.80		.84	.86		.85
ผู้ทรงคุณวุฒิทั้ง 3 ท่าน และผู้ประเมิน CS12	*.74		.81	*.80		.80
ผู้ทรงคุณวุฒิทั้ง 3 ท่าน และผู้ประเมิน DL13	*.77	.81		.82	.82	
ผู้ทรงคุณวุฒิทั้ง 3 ท่าน และผู้ประเมิน DL14	*.76	.80		.84	.84	
ผู้ทรงคุณวุฒิทั้ง 3 ท่าน และผู้ประเมิน DS15	*.80		.85	.87		.82
ผู้ทรงคุณวุฒิทั้ง 3 ท่าน และผู้ประเมิน DS16	*.75		*.78	.87		.83

หมายเหตุ \* หมายถึง แตกต่างอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

จากการพิจารณา ขนาดความสอดคล้องของผู้ประเมินแต่ละคนเมื่อเทียบกับผู้ทรงคุณวุฒิ จากตารางข้างต้น พบว่า เมื่อใช้มาตรฐานแบบสมมาตร ขนาดความสอดคล้องของผู้ประเมินแต่ละคนกับผู้ทรงคุณวุฒิทั้ง 3 คน มีค่าน้อยกว่า ขนาดความสอดคล้องของผู้ทรงคุณวุฒิ 3 คน และมีผลการประเมินไม่สอดคล้องกัน (มีความแตกต่าง) อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 โดยมาตรฐานกราฟฟิกแบบสมมาตร ผลการประเมินของผู้ประเมินทั้ง 16 คน ไม่สอดคล้องกับผู้ทรงคุณวุฒิ ส่วนมาตรฐานประเมินค่าเชิงพฤติกรรม (BARS) แบบสมมาตร ไม่สอดคล้อง 8 คนจาก 16 กรณีใช้มาตรฐานแบบสมมาตร พบว่า มีผู้ประเมินที่ผลการประเมินสอดคล้องกับผู้ทรงคุณวุฒิ มากกว่า เมื่อใช้มาตรฐานแบบสมมาตร และ เมื่อใช้ มาตรฐานประเมินค่าเชิงพฤติกรรม (BARS) แบบสมมาตร ทั้งสองมาตรฐาน ทางบวกและทางลบ พบว่า ผู้ประเมินทุกคนมีผลการประเมินสอดคล้องกับผู้ทรงคุณวุฒิ ดังนั้นจึงสรุปได้ว่า ผลการประเมินจากมาตรฐานสมมาตรมีขนาดความสอดคล้องกับผู้ทรงคุณวุฒิ มากกว่าผลการประเมินจากมาตรฐานแบบสมมาตร และการใช้มาตรฐานประเมินค่าเชิงพฤติกรรม (BARS) ให้ผลการประเมินที่สอดคล้องกับผลการประเมินของผู้ทรงคุณวุฒิมากกว่า มาตรฐานกราฟฟิก

### 3.2 เปรียบเทียบ ค่าเฉลี่ยรวม (grand mean) ด้วยการทดสอบที (t-test )

นอกจากนี้ผู้ประเมินยังได้ วิเคราะห์ด้วยการทดสอบที (t-test) เพื่อเปรียบเทียบ ค่าเฉลี่ยรวม (grand mean) ของผลการประเมินของผู้ประเมินและผู้ทรงคุณวุฒิปรากฏผลดังนี้

ศูนย์วิทยพัทยากร  
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 31 เปรียบเทียบ ค่า grand mean ด้วยการทดสอบที (t-test)

ประเภทมาตร	กลุ่มผู้ประเมิน	จำนวนผู้ประเมิน	กรณี ที่	ค่า t-test ระหว่าง grand mean ของผลการประเมินของผู้ทรงคุณวุฒิกับผู้ประเมิน	
				มาตร Graphic	มาตร BARS
สมมาตร	กตคะแนน	4 คนชุด 1	4	15.237*	9.235*
		4 คนชุด 2	5	11.126*	8.622
		8 คน	6	9.455	9.500
	ปล้่อยคะแนน	4 คนชุด 1	1	18.485*	10.365*
		4 คนชุด 2	2	21.188*	10.378*
		8 คน	3	13.333*	8.322
อสมมาตร	กตคะแนน	4 คนชุด 1	10	9.235	8.235
		4 คนชุด 2	11	8.362	5.622
		8 คน	12	8.510	5.500
	ปล้่อยคะแนน	4 คนชุด 1	7	14.235*	8.985
		4 คนชุด 2	8	17.622*	7.691
		8 คน	9	10.123*	6.300
สมมาตร	สนิท	4 คนชุด 1	13	17.270*	4.271
		4 คนชุด 2	14	21.691*	6.143
		8 คน	15	11.319*	7.519
	ไม่สนิท	4 คนชุด 1	16	11.945*	5.375
		4 คนชุด 2	17	15.331*	6.833
		8 คน	18	9.571	6.571
อสมมาตร	สนิท	4 คนชุด 1	19	12.365*	4.652
		4 คนชุด 2	20	15.391*	7.569
		8 คน	21	13.519*	6.958
	ไม่สนิท	4 คนชุด 1	22	10.165*	7.659
		4 คนชุด 2	23	9.365*	7.396
		8 คน	24	8.833	5.254

\* p&lt;.05



เพื่อเปรียบเทียบ ค่าเฉลี่ยรวม (grand mean) ของผลการประเมินของผู้ประเมินและผู้ทรงคุณวุฒิพบว่า เมื่อใช้มาตราประเมินค่าแบบกราฟฟิก (graphic rating scale) มีกรณีที่ค่าเฉลี่ยรวม (grand mean) ต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 มี 18 กรณี จาก 24 กรณี โดยเป็นมาตรฐานมาตรฐาน 10 กรณีและมาตรฐานมาตรฐาน 8 กรณี แต่เมื่อใช้มาตราประเมินค่าเชิงพฤติกรรม (BARS) พบว่ามี เพียง 3 กรณี ที่ค่าเฉลี่ยรวม (grand mean) ต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 จึงสรุปได้ว่า ใช้มาตราประเมินค่าแบบกราฟฟิก ค่าเฉลี่ยรวม (grand mean) ของผลการประเมินของผู้ประเมินและผู้ทรงคุณวุฒิไม่สอดคล้องกัน แต่เมื่อใช้มาตราประเมินค่าเชิงพฤติกรรม (BARS) พบว่า ค่าเฉลี่ยรวม (grand mean) ของผลการประเมินของผู้ประเมินและผู้ทรงคุณวุฒิสอดคล้องกัน

เนื่องจาก ผู้ประเมิน ในงานวิจัยนี้ คัดเลือกมาโดยจำแนกตามลักษณะการให้คะแนนแบบกต/ปล่อยคะแนน และความสนิท/ไม่สนิทกับผู้ถูกประเมิน ผลการประเมินที่ได้จึงไม่สอดคล้องกับผลของผู้ทรงคุณวุฒิ ดังนั้นในงานวิจัยนี้ จึงได้ ใช้การประมาณค่าและการปรับแก้ความคลาดเคลื่อน ตามวิธีการของ Guilford (1954) ซึ่งใช้ ค่าเฉลี่ยรวม (grand mean) ไปหักออกจากค่าเฉลี่ย (cell mean หรือ marginal mean) ก็จะทำให้ความคลาดเคลื่อนลดน้อยลง และเนื่องจากค่าเฉลี่ยรวม (grand mean) ของผู้ประเมินคลาดเคลื่อนไปจาก ค่าเฉลี่ยรวม (grand mean) ของผู้ทรงคุณวุฒิ ผู้วิจัยจึงพิจารณาปรับวิธีการดังกล่าว โดยปรับแก้ ค่าเฉลี่ยรวม (grand mean) ของผู้ประเมิน โดยใช้ ค่าเฉลี่ยรวม (grand mean) ของผู้ทรงคุณวุฒิ ดังรายละเอียดที่จะแสดงในหัวข้อที่ 3.3

### 3.3 การประมาณค่าความคลาดเคลื่อนโดยการปรับค่า grand mean ให้เท่ากับค่า grand mean ของผลการประเมินของผู้ทรงคุณวุฒิ

ผู้วิจัยได้คำนวณการประมาณค่าความคลาดเคลื่อนแบบกต/ปล่อยคะแนน และค่าความคลาดเคลื่อนแบบฮาโล โดยการปรับค่าเฉลี่ยรวม (grand mean) ให้เท่ากับ ค่าเฉลี่ยรวม (grand mean) ของผลการประเมินของผู้ทรงคุณวุฒิ โดยได้ยกตัวอย่างการคำนวณโดยใช้ข้อมูลผู้ประเมินปล่อยคะแนน จำนวน 4 คน เมื่อประเมินด้วยมาตราประเมินค่าแบบกราฟฟิก แบบสมมาตร มีค่าประมาณความคลาดเคลื่อนแบบกต/ปล่อยคะแนนและแบบฮาโล ดังนี้

ตารางที่ 32 ค่าประมาณความคลาดเคลื่อนแบบกด/ปล่อยคะแนนของผู้ประเมินคนที่  $k$  ( $X'_{kl}$ )

Rating (symmetry 4)	ratee 1	ratee 2	ratee 3	ratee 4	ratee 5	ratee 6	all ratees	$X'_{kl}$
rater AL1	4.8	4.8	5	4.8	5	4.8	4.9	0.27
rater BL5	4.2	5.0	4.2	4.6	5.0	4.6	4.6	0.00
rater CL9	4.0	3.6	5.0	3.6	4.8	3.8	4.1	-0.47
rater DL13	4.8	5.0	4.4	4.8	5.0	4.8	4.8	0.20
all raters	4.45	4.60	4.65	4.45	4.95	4.50	4.60	

ตารางที่ 33 ค่าประมาณความคลาดเคลื่อนแบบฮาโล ( $X'_{kl}$ )

rating(sy4)	ratee 1	ratee 2	ratee 3	ratee 4	ratee 5	ratee 6	sum
rater AL1	0.083	-0.067	0.083	0.083	-0.217	0.033	0.000
rater BL5	-0.250	0.400	-0.450	0.150	0.050	0.100	0.000
rater CL9	0.017	-0.533	0.817	-0.383	0.317	-0.233	0.000
rater DL13	0.150	0.200	-0.450	0.150	-0.150	-0.450	-0.550
sum	0.000	0.000	0.000	0.000	0.000	-0.550	

จากข้อมูลข้างต้นค่าเฉลี่ยรวม (grand mean) ของผู้ประเมินปล่อยคะแนน จำนวน 4 คน เมื่อประเมินด้วยมาตราประเมินค่าแบบกราฟฟิก แบบสมมาตร มีค่าเท่ากับ 4.6 ซึ่งเมื่อเทียบกับค่าเฉลี่ยรวม (grand mean) ของผู้ทรงคุณวุฒิ ทั้ง 3 คน มีค่าเท่ากับ 3.6 หากปรับข้อมูลให้มีค่าเฉลี่ยรวม (grand mean) เท่ากับค่าของ ผู้ทรงคุณวุฒิ คือ 3.63 ต้องลบทุกเซลล์ ด้วยค่าคงที่ 1.0 จะได้ผลดังนี้

ตารางที่ 34 ค่าประมาณความคลาดเคลื่อนแบบกต/ปล่อยคะแนนของผู้ประเมินคนที่  $k$  ( $X'_{ki}$ ) เมื่อปรับค่าเฉลี่ยรวม (grand mean) ให้เท่ากับ ค่าเฉลี่ยรวม (grand mean) ของผู้ทรงคุณวุฒิ

rating(sy4)	ratee 1	ratee 2	ratee 3	ratee 4	ratee 5	ratee 6	all ratees	$X'_{ki}$
rater AL1	3.8	3.8	4.0	3.8	4.0	3.8	3.9	0.27
rater BL5	3.2	4.0	3.2	3.6	4.0	3.6	3.6	0.00
rater CL9	3.0	2.6	4.0	2.6	3.8	2.8	3.1	-0.47
rater DL13	3.8	4.0	3.4	3.8	4.0	3.8	3.8	0.20
all raters	3.45	3.60	3.65	3.45	3.95	3.50	3.60	

ตารางที่ 35 ค่าประมาณความคลาดเคลื่อนแบบฮาโล ( $X'_{ki}$ ) เมื่อปรับค่าเฉลี่ยรวม (grand mean) ให้เท่ากับ ค่าเฉลี่ยรวม (grand mean) ของผู้ทรงคุณวุฒิ

rating(sy4)	ratee 1	ratee 2	ratee 3	ratee 4	ratee 5	ratee 6	sum
rater AL1	0.083	-0.067	0.083	0.083	-0.217	0.033	0.000
rater BL5	-0.250	0.400	-0.450	0.150	0.050	0.100	0.000
rater CL9	0.017	-0.533	0.817	-0.383	0.317	-0.233	0.000
rater DL13	0.150	0.200	-0.450	0.150	-0.150	-0.450	-0.550
sum	0.000	0.000	0.000	0.000	0.000	-0.550	

จากตารางข้างต้น จะเห็นได้ว่าเมื่อปรับค่าเฉลี่ยรวม (grand mean) ให้เท่ากับ ค่าเฉลี่ยรวม (grand mean) ของผู้ทรงคุณวุฒิแล้ว ค่าประมาณความคลาดเคลื่อนแบบกต/ปล่อยคะแนน และค่าประมาณความคลาดเคลื่อนแบบฮาโล ก่อนและหลังการปรับแก้ไข ค่าเฉลี่ยรวม (grand mean) เท่ากับ ค่าเฉลี่ยรวม (grand mean) ของผู้ทรงคุณวุฒิ จะได้ค่าประมาณความคลาดเคลื่อนเท่าเดิม ดังตารางที่ 32 กับ 34 และตารางที่ 33 กับ 35 ตามลำดับ ดังนั้นผู้วิจัยจึงไม่ได้แสดงการคำนวณการประมาณค่าความคลาดเคลื่อนโดยวิธีนี้ ในกรณีอื่น ๆ

#### ตอนที่ 4 การเปรียบเทียบผลของการใช้มาตรฐานสมมาตรกับการใช้มาตรฐานค่าแบบสมมาตรร่วมกับการประมาณค่าและปรับแก้ความคลาดเคลื่อนแบบกต/ปล่อยคะแนน

ขั้นตอนนี้เป็นขั้นตอนการวิเคราะห์ข้อมูล ที่มีจุดมุ่งหมายเพื่อเปรียบเทียบความสอดคล้องของผลการประเมินผลการปฏิบัติงานของครูและความคลาดเคลื่อนแบบกต/ปล่อยคะแนน (leniency error) ระหว่างวิธีการ 2 วิธีคือ (1) การใช้วิธีป้องกันความคลาดเคลื่อนแบบกต/ปล่อยคะแนนโดยการใช้มาตรฐานแบบสมมาตร และ (2) การใช้ มาตรฐานแบบสมมาตรร่วมกับการประมาณค่าและปรับแก้ความคลาดเคลื่อนแบบกต/ปล่อยคะแนน (leniency error) ตามวัตถุประสงค์การวิจัยข้อที่ 1 ผู้วิจัยวิเคราะห์ตามลำดับขั้นตอนนี้

##### 4.1 การประมาณค่าความคลาดเคลื่อนแบบกต/ปล่อยคะแนนตามแนวคิด

###### Guilford

ผู้วิจัยขอยกตัวอย่างข้อมูลและการประมาณค่าความคลาดเคลื่อนแบบ กต/ปล่อยคะแนนตามแนวคิด Guilford เมื่อใช้มาตรฐานประเมินค่าแบบกราฟฟิก (graphic rating scale) แบบสมมาตรกับผู้ประเมินที่มีลักษณะการให้คะแนนแบบปล่อยคะแนน จำนวน 4 คน ชุดที่ 1 (AL1, BL5, CL9 และ DL13)

ตารางที่ 36 ข้อมูลผู้ประเมินปล่อยคะแนน จำนวน 4 คน โดยมาตรฐานประเมินค่าแบบกราฟฟิก แบบสมมาตร

	rater AL1					rater BL5					rater CL9					rater DL13				
trait	T1	T2	T3	T4	T5	T1	T2	T3	T4	T5	T1	T2	T3	T4	T5	T1	T2	T3	T4	T5
ratee 1	5	5	5	5	4	5	4	4	4	4	4	5	4	4	3	5	5	5	5	4
ratee 2	5	5	5	5	4	5	5	5	5	5	5	5	3	3	2	5	5	5	5	5
ratee 3	5	5	5	5	5	4	5	4	5	3	5	5	5	5	5	4	5	5	5	3
ratee 4	5	5	5	5	4	5	4	5	5	4	4	5	4	2	3	5	5	5	5	4
ratee 5	5	5	5	5	5	5	5	5	5	5	5	5	5	4	5	5	5	5	5	5
ratee 6	5	5	5	5	4	5	4	5	5	4	5	5	3	3	3	5	5	5	5	4

จากข้อมูลผู้ประเมินปล่อยคะแนน จำนวน 4 คน โดยมาตรฐานประเมินค่าแบบกราฟฟิก แบบสมมาตร นำมาประมาณค่าความคลาดเคลื่อนแบบกต/ปล่อยคะแนนตามแนวคิด Guilford ดังนี้

1. หาค่าเฉลี่ยของคะแนนการประเมินที่ผู้ประเมิน (rater) แต่ละคนทำการประเมินผู้ถูกประเมิน (ratee) แต่ละคน

2. หาค่าเฉลี่ยของคะแนนประเมินผู้ถูกประเมิน (ratee) ที่ได้รับการประเมินจากผู้ประเมิน (rater) คนเดียวกัน (mean all rates)
  3. หาค่าเฉลี่ยของคะแนนของผู้ประเมิน (rater) ทั้งหมดที่ประเมินผู้ถูกประเมิน (ratee) แต่ละคน (mean all raters)
  4. หาค่าเฉลี่ยรวม (grand mean)
  5. หาค่า  $X'_{ki}$  (ค่าประมาณความคลาดเคลื่อนแบบกต/ปล่อยคะแนน) จาก mean all rates - grand mean
- เมื่อดำเนินการตามขั้นตอนทั้ง 5 ขั้นตอนจะได้ผลดังนี้

ตารางที่ 37 ค่าประมาณความคลาดเคลื่อนแบบกต/ปล่อยคะแนนของผู้ประเมินคนที่ k ( $X'_{ki}$ )

Rating (symmetry 4)	ratee 1	ratee 2	ratee 3	ratee 4	ratee 5	ratee 6	all ratees	( $X'_{ki}$ )
rater AL1	4.8	4.8	5	4.8	5	4.8	4.9	0.27
rater BL5	4.2	5.0	4.2	4.6	5.0	4.6	4.6	0.00
rater CL9	4.0	3.6	5.0	3.6	4.8	3.8	4.1	-0.47
rater DL13	4.8	5.0	4.4	4.8	5.0	4.8	4.8	0.20
all raters	4.45	4.60	4.65	4.45	4.95	4.50	4.60	

6. นำค่าประมาณความคลาดเคลื่อนแบบกต/ปล่อยคะแนนมาปรับแก้จากผลการประเมินก่อนการปรับแก้ความคลาดเคลื่อนจะได้ผลดังตารางต่อไปนี้

ตารางที่ 38 ผลการประเมินที่ปรับแก้ความคลาดเคลื่อนแบบกต/ปล่อยคะแนน

trait	rater AL1					rater BL5					rater CL9					rater DL13				
	T1	T2	T3	T4	T5	T1	T2	T3	T4	T5	T1	T2	T3	T4	T5	T1	T2	T3	T4	T5
ratee 1	4.73	4.73	4.73	4.73	3.73	5.00	4.00	4.00	4.00	4.00	4.47	5.47	4.47	4.47	3.47	4.80	4.80	4.80	4.80	3.80
ratee 2	4.73	4.73	4.73	4.73	3.73	5.00	5.00	5.00	5.00	5.00	5.47	5.47	3.47	3.47	2.47	4.80	4.80	4.80	4.80	4.80
ratee 3	4.73	4.73	4.73	4.73	4.73	4.00	5.00	4.00	5.00	3.00	5.47	5.47	5.47	5.47	5.47	3.80	4.80	4.80	4.80	2.80
ratee 4	4.73	4.73	4.73	4.73	3.73	5.00	4.00	5.00	5.00	4.00	4.47	5.47	4.47	2.47	3.47	4.80	4.80	4.80	4.80	3.80
ratee 5	4.73	4.73	4.73	4.73	4.73	5.00	5.00	5.00	5.00	5.00	5.47	5.47	5.47	4.47	5.47	4.80	4.80	4.80	4.80	4.80
ratee 6	4.73	4.73	4.73	4.73	3.73	5.00	4.00	5.00	5.00	4.00	5.47	5.47	3.47	3.47	3.47	4.80	4.80	4.80	4.80	3.80

7. นำผลการประเมินที่ปรับแก้ความคลาดเคลื่อนแบบกต/ปล่อยคะแนนมาประมาณค่าความคลาดเคลื่อนแบบกต/ปล่อยคะแนน อีกครั้งได้ผลดังนี้

ตารางที่ 39 ค่าประมาณความคลาดเคลื่อนแบบกต/ปล่อยคะแนนที่ปรับแก้ความคลาดเคลื่อนแบบกต/ปล่อยคะแนนแล้ว

rating(sy4)	ratee 1	ratee 2	ratee 3	ratee 4	ratee 5	ratee 6	all ratees	$X'_{ki}$
rater AL1	4.53	4.53	4.73	4.53	4.73	4.53	4.6	0.00
rater BL5	4.2	5.0	4.2	4.6	5.0	4.6	4.6	0.00
rater CL9	4.5	4.1	5.5	4.1	5.3	4.3	4.6	0.00
rater DL13	4.6	4.8	4.2	4.6	4.8	4.6	4.6	0.00
all raters	4.45	4.60	4.65	4.45	4.95	4.50	4.60	

จะเห็นได้ว่าเมื่อคะแนนที่ปรับแก้ความคลาดเคลื่อนแบบกต/ปล่อยคะแนนแล้ว ค่าประมาณความคลาดเคลื่อนแบบกต/ปล่อยจะลดลงในที่นี้มีค่าเป็นศูนย์ทุกผู้ประเมิน แสดงว่า ในกรณีนี้ เมื่อปรับแก้ความคลาดเคลื่อนแล้ว จะได้ผลการประเมินที่ไม่มีความคลาดเคลื่อนแบบกต/ปล่อยคะแนน

#### 4.2 เปรียบเทียบความสอดคล้องค่าประมาณความคลาดเคลื่อนแบบกต/ปล่อยคะแนน

เมื่อนำประมาณค่าความคลาดเคลื่อนแบบกต/ปล่อยคะแนนของข้อมูลจากมาตรฐานมาตรฐานปรับแก้ความคลาดเคลื่อนแบบกต/ปล่อยคะแนนตามวิธี Guilford เทียบกับค่าความคลาดเคลื่อนแบบกต/ปล่อยคะแนนของข้อมูลจากมาตรฐานมาตรฐานที่ไม่ปรับแก้ความคลาดเคลื่อน จะได้ผลดังนี้

ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 40 เปรียบเทียบค่าประมาณความคลาดเคลื่อนแบบกต/ปล่อยคะแนน ของมาตรอสมมาตรและมาตรสมมาตรที่ปรับแก้ความคลาดเคลื่อนแล้ว

กลุ่มผู้ประเมิน			มาตรกราฟฟิก (graphic scale )												มาตรประเมินค่าเชิงพฤติกรรม (behavior anchor rating = BAR)											
			อสมมาตร (asymmetry)						สมมาตร ปรับแก้แล้ว						อสมมาตร (asymmetry)						สมมาตร ปรับแก้แล้ว					
			ปล่อยคะแนน			กตคะแนน			ปล่อยคะแนน			กตคะแนน			ปล่อยคะแนน			กตคะแนน			ปล่อยคะแนน			กตคะแนน		
			8 คน	4คน ชุด1	4คน ชุด2	8 คน	4คน ชุด1	4คน ชุด2	8 คน	4คน ชุด1	4คน ชุด2	8 คน	4คน ชุด1	4คน ชุด2	8 คน	4คน ชุด1	4คน ชุด2	8 คน	4คน ชุด1	4คน ชุด2	8 คน	4คน ชุด1	4คน ชุด2	8 คน	4คน ชุด1	4คน ชุด2
ครูอัสสัมชัญ	ปล่อย คะแนน	AL1	.50	.35																						
		AL2	.50		.65																					
	กต คะแนน	AS3				-.14	-.02						.00	.00										.00	.00	
		AS4				-.08		-.10					.00		.01									.00		.00
ครูต่างโรงเรียน	ปล่อย คะแนน	BL5	.00	-.15											.02	.07								.00	.01	
		BL6	-1.6		-1.5											-.04		-.08						.00		.10
	กต คะแนน	BS7				-.04	-.02						.00	.00											.00	.00
		BS8				.02		.00					.00		.00										.00	.00
ผู้ปกครอง อัสสัมชัญ	ปล่อย คะแนน	CL9	-.30	-.45											-.44	-.40								.00	.00	
		CL10	.23		.38											.16		.12							.00	.10
	กต คะแนน	CS11				.09	.12						.00	.00											.00	.00
		CS12				.16		.13					.00		.00										.00	.00
ผู้บริหาร อัสสัมชัญ	ปล่อย คะแนน	DL13	0.4	.25											-.11	-.07										
		DL14	.27		.42											-.24		-.28						.00	.00	
	กต คะแนน	DS15				-.01	.12						.00	.00										.11	.00	.00
		DS16				-.01		-.03					.00		.00									.00		.00

ตารางที่ 40 เปรียบเทียบค่าประมาณความคลาดเคลื่อนแบบกต/ปล่อยคะแนน ของมาตรอสมมาตรและมาตรสมมาตรที่ปรับแก้ความคลาดเคลื่อนแล้ว

กลุ่มผู้ประเมิน			มาตรกราฟฟิก (graphic scale )											มาตรประเมินค่าเชิงพฤติกรรม (behavior anchor rating = BAR)												
			อสมมาตร (asymmetry)						สมมาตร ปรับแก้แล้ว					อสมมาตร (asymmetry)						สมมาตร ปรับแก้แล้ว						
			สนิทกับอัสสัมชัญ			ไม่สนิทกับอัสสัมฯ			สนิทกับอัสสัมชัญ			ไม่สนิทกับอัสสัมฯ		สนิทกับอัสสัมชัญ			ไม่สนิทกับอัสสัมฯ			สนิทกับอัสสัมชัญ			ไม่สนิทกับอัสสัมฯ			
			8 คน	4 คน	4 คน	8 คน	4 คน	4 คน	8 คน	4 คน	4 คน	8 คน	4 คน	4 คน	8 คน	4 คน	4 คน	8 คน	4 คน	4 คน	8 คน	4 คน	4 คน	8 คน	4 คน	4 คน
ครูอัสสัมชัญ	ปล่อย คะแนน	AL1	1.2	1.2				.00	.00					.95	1.0					.00	.00					
		AL2	1.2		1.1				.00		.00				.89		.83				.00		.20			
	กต คะแนน	AS3	-1.0	-.96				.00	.00						-.81	-.75					.00	.00				
		AS4	-.98		-1.1				.00		.00				-.78		-.84				.00		-.61			
ครูต่างโรงเรียน	ปล่อย คะแนน	BL5				.64	.64				.00	.00					.72	.68					.00	.30		
		BL6				.71		.71				.00		.00				.65		.69				.00		.68
	กต คะแนน	BS7				-.89	-.89					.00	.00					-.58	-.63					.00	.30	
		BS8				-.83		-.83				.00		.00				-.68		-.64				.00		.68
ผู้ปกครอง อัสสัมชัญ	ปล่อย คะแนน	CL9	.36	.44				.00	.00					.15	.22					.00	.00					
		CL10	.89		.81				.00		.00				.75		.69				.00		.20			
	กต คะแนน	CS11	-.81	-.73					.00	.00					-.55	-.48					.00	.00				
		CS12	-.74		-.83				.00		.00				-.61		-.68				.00		.21			
นอกรั้วโรงเรียน	ปล่อย คะแนน	DL13				1.1	-.89				.00	.00					.58	.54					.00	.29		
		DL14				.98		.98				.00		.00				.45		.49				.00		-2.03
	กต คะแนน	DS15				-.86	1.1				.00	.00						-.55	-.59					.00	-.89	
		DS16				-.86		-.86				.00		.00				-.58		-.54				.00		.68



จากตารางข้างต้น จะเห็นได้ว่าผลการประเมินจากมาตรแบบอสมมาตรมีค่าประมาณ ความคลาดเคลื่อนแบบกต/ปล่อยคะแนนมากกว่าผลการประเมินจากมาตรแบบสมมาตรที่ปรับแก้ ความคลาดเคลื่อนแบบกต/ปล่อยคะแนนแล้ว โดยผลการประเมินจากมาตรประมาณค่า เริงพฤติกรรม (BARS) มีค่าประมาณความคลาดเคลื่อนแบบกต/ปล่อยคะแนนสูงกว่าค่าประมาณ ความคลาดเคลื่อนจากมาตรประเมินค่าแบบกราฟฟิก (graphic rating scale)

เพื่อให้เห็นถึงความสอดคล้องของค่าประมาณความคลาดเคลื่อนแบบกต/ปล่อยคะแนน ของมาตรแบบอสมมาตรที่ไม่ปรับแก้ความคลาดเคลื่อนและค่าประมาณความคลาดเคลื่อนแบบ กต/ปล่อยคะแนนของมาตรแบบสมมาตรที่ปรับแก้ความคลาดเคลื่อนแบบกต/ปล่อยคะแนนผู้วิจัย จึงเปรียบเทียบความแตกต่างของค่าประมาณความคลาดเคลื่อนแบบกต/ปล่อยคะแนน โดยใช้ เกณฑ์ว่าผลการประมาณค่าความคลาดเคลื่อนจะสอดคล้องกันก็ต่อเมื่อ ขนาดของความแตกต่าง ต้องมีขนาดไม่เกิน .05 หากเกินต้องมีจำนวนน้อยกว่าร้อยละ 80 ของทั้งหมด (Joreskog, 1970) ผลการเปรียบเทียบขนาดของความแตกต่างของค่าประมาณความคลาดเคลื่อนให้ผลดังตาราง ต่อไปนี้



ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 41 ขนาดของความแตกต่างของค่าประมาณความคลาดเคลื่อนแบบกต/ปล่อยคะแนน  
ของมาตรสสมมาตรก่อนปรับแก้และมาตรสสมมาตรหลังปรับแก้แล้ว

กลุ่มผู้ประเมิน			ขนาดของความแตกต่าง											
			มาตรประเมินค่าแบบกราฟฟิก (graphic rating scale)						มาตรประเมินค่าเชิงพฤติกรรม BARS					
			ปล่อยคะแนน			กตคะแนน			ปล่อยคะแนน			กตคะแนน		
			8 คน	4 คน ชุด1	4 คน ชุด2	8 คน	4 คน ชุด1	4 คน ชุด2	8 คน	4 คน ชุด1	4 คน ชุด2	8 คน	4 คน ชุด1	4 คน ชุด2
ครูอัสสัมชัญ	ปล่อย คะแนน	AL1	.51	.35					.36	.41				
		AL2	.5		.83				.29		.57			
	กต คะแนน	AS3				-.14	-.02					-.12	-.14	
		AS4				-.08		-.11				-.09	-.07	
ครูต่างโรงเรียน	ปล่อย คะแนน	BL5	.0	-.15					.02	.06				
		BL6	-1.6		-1.6				-.04		-.18			
	กต คะแนน	BS7				-.04	-.02					.01	-.01	
		BS8				.02		.0				-.09	-.07	
ผู้ปกครองอัสสัมชัญ	ปล่อย คะแนน	CL9	-0.3	-.45					-.44	-0.4				
		CL10	.23		.32				.16		.02			
	กต คะแนน	CS11				.09	.12					.15	.13	
		CS12				.16		.13				.08	.1	
ผู้ปกครอง ต่างโรงเรียน	ปล่อย คะแนน	DL13	.4	.25					-.11	-.07				
		DL14	.25		.36				-.24		-.28			
	กต คะแนน	DS15				-.01	.12					.05	.03	
		DS16				-.01		-.03				.01		.03

ตารางที่ 41 ขนาดของความแตกต่างของค่าประมาณความคลาดเคลื่อนแบบกต/ปล่อยคะแนน  
ของมาตรอสมมาตรก่อนปรับแก้และมาตรสมมาตรหลังปรับแก้แล้ว (ต่อ)

กลุ่มผู้ประเมิน			ขนาดของความแตกต่าง										
			มาตรประเมินค่าแบบกราฟฟิก (graphic rating scale)						มาตรประเมินค่าเชิงพฤติกรรม (BARS)				
			สนิทกับอัสสัมชัญ			ไม่สนิทกับอัสสัมฯ			สนิทกับอัสสัมชัญ		ไม่สนิทกับอัสสัมฯ		
			8 คน	4 คน ชุด1	4 คน ชุด2	8 คน	4 คน ชุด1	4 คน ชุด2	8 คน	4 คน ชุด1	4 คน ชุด2	8 คน	4 คน ชุด1
ครูอัสสัมชัญ	ปล่อย	AL1	1.2	1.2					.95	1			
	คะแนน	AL2	1.2		1.1				.89		.63		
	กต	AS3	-1	-.96					-.81	-.75			
	คะแนน	AS4	-.98		-1.1				-.78		-.23		
ครูต่างโรงเรียน	ปล่อย	BL5				.64	.64					.72	.38
	คะแนน	BL6				.71		.71				.65	.01
	กต	BS7				-.89	-.89					-.58	-.93
	คะแนน	BS8				-.83		-.83				-.68	-1.32
ผู้ปกครอง อัสสัมชัญ	ปล่อย	CL9	.36	.44					.15	.22			
	คะแนน	CL10	.89		.81				.75		.49		
	กต	CS11	-.81	-.73					-.55	-.48			
	คะแนน	CS12	-.74		-.83				-.61		-.89		
ผู้ปกครอง ต่างโรงเรียน	ปล่อย	DL13				1.1	-.89					.58	.25
	คะแนน	DL14				.98		.98				.45	2.52
	กต	DS15				-.86	1.1					-.55	.3
	คะแนน	DS16				-.86		-.86				-.58	-1.22

จากตารางข้างต้น ผลการเปรียบเทียบขนาดของความแตกต่างของค่าประมาณความคลาดเคลื่อนแบบกต/ปล่อยคะแนนของมาตรอสมมาตรที่ไม่ปรับแก้ความคลาดเคลื่อนและมาตรสมมาตรที่ปรับแก้ความคลาดเคลื่อน พบว่าการใช้มาตรสมมาตรปรับแก้ความคลาดเคลื่อนแบบกต/ปล่อยคะแนนทำให้ค่าประมาณความคลาดเคลื่อนมีขนาดเล็กลง เมื่อพิจารณาความสอดคล้องของค่าประมาณความคลาดเคลื่อนแบบกต/ปล่อยคะแนนของทั้งสองวิธี โดยใช้เกณฑ์ขนาดของความแตกต่างต้องมีขนาดไม่เกิน .05 มากกว่าร้อยละ 80 ของค่าประมาณทั้งหมด (Joreskog, 1970) ผลการวิเคราะห์ความสอดคล้องสรุปได้ดังตาราง ต่อไปนี้

ตารางที่ 42 สรุปกรณีการวิเคราะห์ความสอดคล้องของค่าประมาณความคลาดเคลื่อนแบบ  
กต/ปล่อยคะแนน ตามวัตถุประสงค์ข้อที่ 1

ประเภทมาตร	กลุ่มผู้ประเมิน	จำนวนผู้ประเมิน	กรณี ที่	สมมาตร ปรับแก้ leniency error	อสมมาตร ไม่ปรับแก้ ความคลาดเคลื่อน
BARS	กตคะแนน	4 คนชุด 1	1	ไม่สอดคล้อง	
		4 คนชุด 2	2	ไม่สอดคล้อง	
		8 คน	3	ไม่สอดคล้อง	
	ปล่อยคะแนน	4 คนชุด 1	4	ไม่สอดคล้อง	
		4 คนชุด 2	5	ไม่สอดคล้อง	
		8 คน	6	ไม่สอดคล้อง	
GRAPHIC	กตคะแนน	4 คนชุด 1	7	ไม่สอดคล้อง	
		4 คนชุด 2	8	ไม่สอดคล้อง	
		8 คน	9	ไม่สอดคล้อง	
	ปล่อยคะแนน	4 คนชุด 1	10	ไม่สอดคล้อง	
		4 คนชุด 2	11	ไม่สอดคล้อง	
		8 คน	12	ไม่สอดคล้อง	
BARS	สนิท	4 คนชุด 1	13	ไม่สอดคล้อง	
		4 คนชุด 2	14	ไม่สอดคล้อง	
		8 คน	15	ไม่สอดคล้อง	
	ไม่สนิท	4 คนชุด 1	16	ไม่สอดคล้อง	
		4 คนชุด 2	17	ไม่สอดคล้อง	
		8 คน	18	ไม่สอดคล้อง	
GRAPHIC	สนิท	4 คนชุด 1	19	ไม่สอดคล้อง	
		4 คนชุด 2	20	ไม่สอดคล้อง	
		8 คน	21	ไม่สอดคล้อง	
	ไม่สนิท	4 คนชุด 1	22	ไม่สอดคล้อง	
		4 คนชุด 2	23	ไม่สอดคล้อง	
		8 คน	24	ไม่สอดคล้อง	

จากตารางสรุปการวิเคราะห์ความสอดคล้องของค่าประมาณความคลาดเคลื่อนแบบกต/ปล่อยคะแนนจะเห็นได้ว่า ไม่สอดคล้องทุกกรณี เนื่องจากขนาดของความแตกต่างของค่าประมาณความคลาดเคลื่อนแบบกต/ปล่อยคะแนนจากทั้งสองวิธีที่มีขนาดเล็กกว่า .05 มีไม่ถึงร้อยละ 80 ของค่าประมาณทั้งหมด อาจกล่าวได้ว่ามาตรฐานไม่สามารถป้องกันความคลาดเคลื่อนแบบกต/ปล่อยคะแนนได้

#### 4.2 เปรียบเทียบความสอดคล้องของผลการประเมินผลการปฏิบัติงานของครู

นอกจากความคลาดเคลื่อนแล้วผู้วิจัยยังได้พิจารณาความสอดคล้องของผลการประเมินการปฏิบัติงานของครูของมาตรฐานแบบอสมมาตรที่ไม่ปรับแก้ความคลาดเคลื่อน และมาตรฐานแบบสมมาตรที่ปรับแก้ความคลาดเคลื่อนแบบกต/ปล่อยคะแนน โดยเปรียบเทียบใน 3 กรณีดังนี้ คือ

กรณีที่ 1 เปรียบเทียบความสอดคล้องของผลการประเมินการปฏิบัติงานของครูของมาตรฐานแบบสมมาตรที่ปรับแก้ความคลาดเคลื่อนแบบกต/ปล่อยคะแนน (leniency error) และผลการประเมินการปฏิบัติงานของครูของมาตรฐานแบบอสมมาตรที่ไม่ได้ปรับแก้ความคลาดเคลื่อน เทียบกับเกณฑ์การประเมินจากผู้ทรงคุณวุฒิ โดยหาขนาดความสอดคล้องของคะแนนตามสูตรของ Downie และ Heath (Downie and Heath 1974 :215)

$$r_{xx} = \frac{MS_r - MS_E}{MS_r + (k - 1)MS_E} \quad \text{เมื่อ } k = \text{จำนวนผู้ประเมิน}$$

กรณีที่ 2 เปรียบเทียบความสอดคล้องของผลการประเมินการปฏิบัติงานของครูของมาตรฐานแบบสมมาตรที่ปรับแก้ความคลาดเคลื่อนแบบกต/ปล่อยคะแนน (leniency error) เทียบกับผลการประเมินการปฏิบัติงานของครูของมาตรฐานแบบอสมมาตรที่ไม่ได้ปรับแก้ความคลาดเคลื่อน

โดยพิจารณาจากค่าสัมประสิทธิ์สหสัมพันธ์ และค่าเฉลี่ยผลต่างกำลังสอง (mean square difference) ดังสมการที่ (18)

$$\text{mean square difference} = \sqrt{\frac{\sum (x_i - y_i)^2}{n}} \quad \text{-----(18)}$$

เมื่อ	$X_i$	แทน	คะแนนผลการประเมินก่อนการปรับแก้
	$Y_i$	แทน	คะแนนผลการประเมินหลังการปรับแก้
	$n$	แทน	จำนวนข้อมูล

ได้ผลการวิเคราะห์ดังนี้

##### 4.2.1 เปรียบเทียบกับเกณฑ์การประเมินจากผู้ทรงคุณวุฒิ

ตารางที่ 43 ขนาดความสอดคล้องของผลการประเมินของผู้ประเมินแต่ละคน ( $r_{xx}$ ) เมื่อเทียบกับผู้ทรงคุณวุฒิ ตามวัตถุประสงค์การวิจัยข้อที่ 1

ผู้ประเมิน	สมมาตรปรับแก้ leniency error		สมมาตรไม่ปรับแก้ความคลาดเคลื่อน			
	graphic สมมาตร	BAR สมมาตร	graphic		BAR	
			สมมาตร ทางบวก	สมมาตร ทางลบ	สมมาตร ทางบวก	สมมาตร ทางลบ
ผู้ทรงคุณวุฒิทั้ง 3 ท่าน	.82	.87	.80	.84	.85	.84
ผู้ทรงคุณวุฒิทั้ง 3 ท่าน และผู้ประเมิน AL1	*.79	.82	*.79		.82	
ผู้ทรงคุณวุฒิทั้ง 3 ท่าน และผู้ประเมิน AL2	*.79	.85	.80		.86	
ผู้ทรงคุณวุฒิทั้ง 3 ท่าน และผู้ประเมิน AS3	.79	.81		*.77		.80
ผู้ทรงคุณวุฒิทั้ง 3 ท่าน และผู้ประเมิน AS4	.82	.81		*.79		.82
ผู้ทรงคุณวุฒิทั้ง 3 ท่าน และผู้ประเมิน BL5	.79	.83	*.77		.85	
ผู้ทรงคุณวุฒิทั้ง 3 ท่าน และผู้ประเมิน BL6	.79	.88	*.75		.87	
ผู้ทรงคุณวุฒิทั้ง 3 ท่าน และผู้ประเมิน BS7	.80	.85		*.79		.83
ผู้ทรงคุณวุฒิทั้ง 3 ท่าน และผู้ประเมิน BS8	.81	.84		*.76		.83
ผู้ทรงคุณวุฒิทั้ง 3 ท่าน และผู้ประเมิน CL9	.81	.87	*.79		.86	
ผู้ทรงคุณวุฒิทั้ง 3 ท่าน และผู้ประเมิน CL10	.81	.84	.80		.84	
ผู้ทรงคุณวุฒิทั้ง 3 ท่าน และผู้ประเมิน CS11	.83	.86		.84		.85
ผู้ทรงคุณวุฒิทั้ง 3 ท่าน และผู้ประเมิน CS12	.88	.85		.81		.80
ผู้ทรงคุณวุฒิทั้ง 3 ท่าน และผู้ประเมิน DL13	.82	.82	.81		.82	
ผู้ทรงคุณวุฒิทั้ง 3 ท่าน และผู้ประเมิน DL14	.81	.86	.80		.84	
ผู้ทรงคุณวุฒิทั้ง 3 ท่าน และผู้ประเมิน DS15	.85	.84		.85		.82
ผู้ทรงคุณวุฒิทั้ง 3 ท่าน และผู้ประเมิน DS16	.85	.87		*.78		.83

หมายเหตุ \* หมายถึง แตกต่างอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ผลการประเมินจากมาตรแบบสมมาตร เมื่อปรับแก้ความคลาดเคลื่อนแบบกด/ปล่อย คะแนนแล้ว จะมีขนาดความสอดคล้องกับเกณฑ์การประเมินของผู้ทรงคุณวุฒิมากขึ้น โดยมาตรกราฟฟิกแบบสมมาตร ก่อนปรับแก้ความคลาดเคลื่อนไม่มีผู้ประเมิน 2 ที่มีผลการประเมินสอดคล้องกับผู้ทรงคุณวุฒิเลย แต่เมื่อปรับแก้ความคลาดเคลื่อนแบบกด/ปล่อย คะแนนแล้ว พบว่ามีผู้ประเมินเพียง 2 คนที่มีผลการประเมินไม่สอดคล้องกับผู้ทรงคุณวุฒิ นอกจากนี้ยังพบว่า ผลของมาตรกราฟฟิกแบบสมมาตร ที่ปรับแก้ความคลาดเคลื่อนแบบกด/ปล่อยคะแนน ส่วนใหญ่มีค่าขนาดความสอดคล้อง ( $r_{xx} > .8$ ) มากกว่ามาตรกราฟฟิกแบบสมมาตรที่ไม่ได้ปรับแก้ความคลาดเคลื่อน อย่างไรก็ตามจะเห็นได้ว่า เมื่อปรับแก้ความคลาดเคลื่อนแบบกด/ปล่อยคะแนนแล้ว จะมีผลการประเมินของผู้ประเมินเพียง 4 ท่านที่มีขนาดความสอดคล้องมากกว่า เกณฑ์ความสอดคล้องของผู้ทรงคุณวุฒิทั้ง 3 ท่าน ( $r_{xx} > .82$ )

เมื่อพิจารณามาตรประเมินค่าเชิงพฤติกรรม (BARS) ทั้งมาตรแบบสมมาตร ที่ปรับแก้ความคลาดเคลื่อนแบบกด/ปล่อยคะแนนแล้ว และมาตรแบบสมมาตรที่ไม่ได้ปรับแก้ความคลาดเคลื่อน พบว่า ผลการประเมินจากผู้ประเมินทุกท่าน สอดคล้องกับผู้ทรงคุณวุฒิ ( $r_{xx} > .8$ ) โดยมาตรประเมินค่าเชิงพฤติกรรม (BARS) ทั้งมาตรแบบสมมาตร ที่ปรับแก้ความคลาดเคลื่อนแบบกด/ปล่อยคะแนนแล้ว มีขนาดของความสอดคล้องมากกว่ามาตรประเมินค่าเชิงพฤติกรรม (BARS) แบบสมมาตรที่ไม่ได้ปรับแก้ความคลาดเคลื่อน

สรุปได้ว่าการปรับแก้ความคลาดเคลื่อนแบบกด/ปล่อย คะแนน ทำให้ได้ผลการประเมินที่มีคุณภาพสอดคล้องกับผลการประเมินของผู้ทรงคุณวุฒิมากยิ่งขึ้น กล่าวคือ มาตรแบบสมมาตร ที่ปรับแก้ความคลาดเคลื่อนแบบกด/ปล่อยคะแนน ให้ผลการประเมินที่สอดคล้อง กับผลการประเมินของผู้ทรงคุณวุฒิ มากกว่ามาตรแบบสมมาตรที่ไม่ได้ปรับแก้ความคลาดเคลื่อน

4.2.2 เปรียบเทียบความสอดคล้องของผลการประเมินการปฏิบัติงานของครูของมาตรแบบสมมาตรที่ปรับแก้ความคลาดเคลื่อนแบบกด/ปล่อยคะแนน (leniency error) เทียบกับผลการประเมินการปฏิบัติงานของครูของมาตรแบบสมมาตรที่ไม่ได้ปรับแก้ความคลาดเคลื่อน

สรุปได้ดังตารางต่อไปนี้

ตารางที่ 44 กรณีการวิเคราะห์ความสอดคล้องผลการปฏิบัติงานของครูตามวัตถุประสงค์ข้อที่ 1

ประเภท มาตร	กลุ่ม ผู้ประเมิน	จำนวน ผู้ประเมิน	กรณีที่	สมมาตร ปรับแก้ leniency error เทียบสมมาตร ไม่ ปรับแก้ความคลาดเคลื่อน	
				Mean square	Corr
BARS	กวดคะแนน	4 คนชุด 1	1	.763	.720
		4 คนชุด 2	2	.724	.692
		8 คน	3	.745	.512
	ปล่อยคะแนน	4 คนชุด 1	4	.660	.754
		4 คนชุด 2	5	.721	.526
		8 คน	6	.690	.772
GRAPHIC	กวดคะแนน	4 คนชุด 1	7	.771	.426
		4 คนชุด 2	8	1.147	.314
		8 คน	9	.874	.429
	ปล่อยคะแนน	4 คนชุด 1	10	.973	.324
		4 คนชุด 2	11	.505	.217*
		8 คน	12	1.015	.209
BARS	สนิท	4 คนชุด 1	13	1.002	.314
		4 คนชุด 2	14	.932	.414
		8 คน	15	1.026	.287
	ไม่สนิท	4 คนชุด 1	16	.877	.310
		4 คนชุด 2	17	1.721	.312
		8 คน	18	.875	.358
GRAPHIC	สนิท	4 คนชุด 1	19	1.193	.247
		4 คนชุด 2	20	1.326	.219
		8 คน	21	1.222	.306
	ไม่สนิท	4 คนชุด 1	22	1.234	.214
		4 คนชุด 2	23	1.159	.303
		8 คน	24	1.196	.294

\*  $p < .05$



ผลการเปรียบเทียบความสอดคล้องพบว่า ผลการประเมินจากมาตรแบบสมมาตรที่ปรับแก้ความคลาดเคลื่อนแบบกด/ปล่อยคะแนน (leniency error) และผลการประเมินจากมาตรแบบอสมมาตรที่ไม่ปรับแก้ความคลาดเคลื่อน มีกรณีที่สอดคล้องกันเพียงกรณีเดียว จาก 24 กรณี จึงสรุปได้ว่า ผลการประเมินการปฏิบัติงานของครู จากวิธีป้องกันความคลาดเคลื่อนแบบกด/ปล่อยคะแนน (leniency error) ด้วยมาตรอสมมาตร และวิธีการใช้มาตรแบบสมมาตรร่วมกับการประมาณค่าและปรับแก้ความคลาดเคลื่อนแบบกด/ปล่อยคะแนน ให้ผลไม่สอดคล้องกัน

#### ตอนที่ 5 เปรียบเทียบผลการใช้มาตรประเมินค่าเชิงพฤติกรรม (BARS) และการใช้มาตรกราฟฟิกร่วมกับการประมาณค่าและปรับแก้ความคลาดเคลื่อนแบบฮาโล

ขั้นตอนนี้เป็นขั้นตอนการวิเคราะห์ข้อมูล ที่มีจุดมุ่งหมาย เพื่อเปรียบเทียบความสอดคล้องของผลการประเมินผลการปฏิบัติงานของครูและความคลาดเคลื่อนแบบฮาโล (halo error) ระหว่างวิธีการ 2 วิธีคือ (1) การใช้วิธีป้องกันความคลาดเคลื่อนแบบฮาโลโดยการใช้มาตรประเมินค่าเชิงพฤติกรรม (BARS) และ (2) การใช้มาตรประเมินค่าแบบกราฟฟิก (graphic rating scale) ร่วมกับการประมาณค่าและปรับแก้ความคลาดเคลื่อนแบบฮาโล (halo error) ตามวัตถุประสงค์การวิจัยข้อที่ 2 มีลำดับขั้นตอนการวิเคราะห์ดังนี้

##### 5.1 การประมาณค่าความคลาดเคลื่อนแบบฮาโล ตามแนวคิด Guilford

ผู้วิจัยขอยกตัวอย่างข้อมูลและวิธีการประมาณค่าความคลาดเคลื่อนแบบฮาโล เมื่อใช้มาตรประเมินค่าแบบกราฟฟิก (graphic rating scale) แบบสมมาตร กับผู้ประเมินที่มีความสนิทกับโรงเรียนอัสสัมชัญ จำนวน 4 คน ชุดที่ 1 (AL1, AS3, CL9 และ CS11) เป็นตัวอย่างการวิเคราะห์ดังนี้

ตารางที่ 45 คะแนนผลการประเมินเมื่อใช้มาตรประเมินค่าแบบกราฟฟิก (graphic rating scale) แบบสมมาตร กับผู้ประเมินที่มีความสนิทกับโรงเรียนอัสสัมชัญ จำนวน 4 คน ชุดที่ 1

trait	rater AL1					rater AS3					rater CL9					rater CS11				
	T1	T2	T3	T4	T5	T1	T2	T3	T4	T5	T1	T2	T3	T4	T5	T1	T2	T3	T4	T5
ratee 1	5	5	5	5	4	1	2	3	4	2	4	5	4	4	3	2	3	3	3	3
ratee 2	5	5	5	5	4	2	1	3	2	1	5	5	3	3	2	3	2	3	2	1
ratee 3	5	5	5	5	5	5	5	4	5	4	5	5	5	5	5	5	4	4	3	4
ratee 4	5	5	5	5	4	2	3	3	2	2	4	5	4	2	3	4	3	3	2	2
ratee 5	5	5	5	5	5	5	5	5	3	3	5	5	5	4	5	4	5	5	4	4
ratee 6	5	5	5	5	4	2	2	2	1	2	5	5	3	3	3	4	3	3	3	2

จากข้อมูลประมาณค่าความคลาดเคลื่อนแบบกต/ปล่อยคะแนนตามขั้นตอนของ Guilford จะได้ผลดังนี้

ตารางที่ 46 ค่าประมาณความคลาดเคลื่อนแบบกต/ปล่อยคะแนนเมื่อใช้มาตรประเมินค่าแบบกราฟฟิก (graphic rating scale) แบบสมมาตร กับผู้ประเมินที่มีความสนิทกับโรงเรียนอัสสัมชัญ จำนวน 4 คน ชุดที่ 1

rating(sy4)	ratee 1	ratee 2	ratee 3	ratee 4	ratee 5	ratee 6	all ratees	KL (LENIEN)
rater AL1	4.8	4.8	5	4.8	5	4.8	4.9	1.10
rater AS3	2.4	1.8	4.6	2.4	4.2	1.8	2.9	-0.90
rater BL5	4.0	3.6	5.0	3.6	4.8	3.8	4.1	0.37
rater BS7	2.8	2.2	4.0	2.8	4.4	3.0	3.2	-0.57
all raters	3.50	3.10	4.65	3.40	4.60	3.35	3.77	
di	-0.27	-0.67	0.88	-0.37	0.83	-0.42		

จากข้อมูลในตารางข้างต้น นำมาหาค่าเฉลี่ยคะแนนการประเมินในแต่ละเซลล์ ที่หักออกด้วยค่าประมาณความคลาดเคลื่อนแบบกต/ปล่อยคะแนน ( $X'_{ki}$ ) และค่า  $d_i$  และลบด้วยค่า grand mean ค่าที่คำนวณได้แต่ละเซลล์ คือ ค่าประมาณความคลาดเคลื่อนแบบฮาโล (halo error =  $X'_{ki}$ ) ดังต่อไปนี้

ตารางที่ 47 ค่าประมาณความคลาดเคลื่อนแบบฮาโล เมื่อใช้มาตรประเมินค่าแบบกราฟฟิก (graphic rating scale) แบบสมมาตร กับผู้ประเมินที่มีความสนิทกับโรงเรียนอัสสัมชัญ จำนวน 4 คน ชุดที่ 1

rating(sy4)	ratee 1	ratee 2	ratee 3	ratee 4	ratee 5	ratee 6	sum
rater AL1	0.200	0.600	-0.750	0.300	-0.700	0.350	0.000
rater AS3	-0.200	-0.400	0.850	-0.100	0.500	-0.650	0.000
rater CL9	0.133	0.133	-0.017	-0.167	-0.167	0.083	0.000
rater cS11	-0.133	-0.333	-0.083	-0.033	0.367	-0.083	-0.300
sum	0.000	0.000	0.000	0.000	0.000	-0.300	

นำค่าประมาณความคลาดเคลื่อนแบบฮาโล จากตารางข้างต้น ปรับแก้ผลการประเมินจะได้ผลดังตารางต่อไปนี้

ตารางที่ 48 ผลการปรับแก้ความคลาดเคลื่อนแบบฮาโล

	rater AL1					rater AS3					rater CL9					rater CS11				
trait	T1	T2	T3	T4	T5	T1	T2	T3	T4	T5	T1	T2	T3	T4	T5	T1	T2	T3	T4	T5
ratee 1	4.8	4.8	4.8	4.8	3.8	1.2	2.2	3.2	4.2	2.2	3.9	4.9	3.9	3.9	2.9	2.3	3.1	3.1	3.1	3.1
ratee 2	4.4	4.4	4.4	4.4	3.4	2.4	1.4	3.4	2.4	1.4	4.9	4.9	2.9	2.9	1.9	3.3	2.3	3.3	2.3	1.3
ratee 3	5.7	5.7	5.7	5.7	5.7	4.1	4.1	3.2	4.2	3.2	5	5	5	5	5	5.1	4.1	4.1	3.1	4.1
ratee 4	4.7	4.7	4.7	4.7	3.7	2.1	3.1	3.1	2.1	2.1	4.2	5.2	4.2	2.2	3.2	4	3	3	2	2
ratee 5	5.7	5.7	5.7	5.7	5.7	4.5	4.5	4.5	2.5	2.5	5.2	5.2	5.2	4.2	5.2	3.6	4.6	4.6	3.6	3.6
ratee 6	4.6	4.6	4.6	4.6	3.6	2.6	2.6	2.7	1.7	2.7	4.9	4.9	2.9	2.9	2.9	4.1	3.1	3.1	3.1	2.1

นำข้อมูลที่ปรับแก้ความคลาดเคลื่อนแบบฮาโล มาประมาณค่าความคลาดเคลื่อนแบบฮาโล อีกครั้งได้ผลดังนี้

ตารางที่ 49 เปรียบเทียบค่าประมาณความคลาดเคลื่อนแบบฮาโลของมาตรประเมินค่าแบบกราฟฟิค แบบสมมาตร ก่อนและหลังการปรับแก้ความคลาดเคลื่อนแบบฮาโล

rating(sy4)	ก่อนปรับแก้ halo error							หลังปรับแก้ halo error						
	ratee 1	ratee 2	ratee 3	ratee 4	ratee 5	ratee 6	sum	ratee 1	ratee 2	ratee 3	ratee 4	ratee 5	ratee 6	sum
rater AL1	.200	.600	-.750	.300	-.700	.350	.000	.004	.014	.014	.014	.014	-.061	.000
rater AS3	-.200	-.400	.850	-.100	.500	-.650	.000	.004	.014	.014	.014	.014	-.061	.000
rater CL9	.133	.133	-.017	-.167	-.167	.083	.000	.004	.014	.014	.014	.014	-.061	.000
rater CS11	-.133	-.333	-.083	-.033	.367	-.083	-.300	-.013	-.043	-.043	-.043	-.043	-.043	-.225
sum	.000	.000	.000	.000	.000	-.300		.000	.000	.000	.000	.000	-.225	

จากตารางข้างต้น จะเห็นได้ว่าค่าประมาณความคลาดเคลื่อนแบบฮาโลของหลังการปรับแก้ความคลาดเคลื่อนแบบฮาโลแล้วมีขนาดเล็กลง จะมีก็เพียงผลของผู้ประเมิน CS11 (ผู้ปกครองนักเรียนโรงเรียนอัสสัมชัญที่มีลักษณะการให้คะแนนแบบกดคะแนน คนที่ 1) ที่มีขนาดเพิ่มขึ้นคือ จาก -.033 เป็น -.043 จะเห็นได้ว่าก่อนการปรับแก้ผู้ประเมินที่มีภูมิหลังสนิทกับโรงเรียนอัสสัมชัญ จะมีขนาดค่าประมาณความคลาดเคลื่อนแบบฮาโลสูงโดยครูโรงเรียนอัสสัมชัญที่มีลักษณะกดคะแนนจะมีความคลาดเคลื่อนแบบฮาโลมากที่สุดคือ .650 โดยมีลักษณะให้คะแนนต่ำกว่าความเป็นจริง รองลงมาคือครูโรงเรียนอัสสัมชัญที่มีลักษณะปล่อยคะแนนจะมีความคลาดเคลื่อนแบบฮาโล .35 โดยมีลักษณะให้คะแนนสูงกว่าความเป็นจริง

## 5.2 เปรียบเทียบความสอดคล้องค่าประมาณความคลาดเคลื่อนแบบฮาโล

เมื่อนำผลการประเมินจากมาตรประเมินค่าแบบกราฟฟิก (graphic rating scale) ปรับแก้ความคลาดเคลื่อนแบบฮาโลเทียบกับผลการประเมินของมาตรประเมินค่าเชิงพฤติกรรม (BARS) ที่ไม่ปรับแก้ความคลาดเคลื่อน ตามวัตถุประสงค์การวิจัยข้อที่ 2 ได้ผลดังนี้

ตารางที่ 50 เปรียบเทียบค่าประมาณความคลาดเคลื่อนแบบฮาโล ของข้อมูลผู้ประเมินที่มีความสนิทกับโรงเรียนอัสสัมชัญ จำนวน 4 คน ชุดที่ 1

rating(sy4)	BARS ไม่ปรับแก้							Graphic Rating Scale หลังปรับแก้ halo error						
	ratee 1	ratee 2	ratee 3	ratee 4	ratee 5	ratee 6	sum	ratee 1	ratee 2	ratee 3	ratee 4	ratee 5	ratee 6	sum
rater AL1	.200	.600	-.750	.300	-.700	.350	.000	.004	.014	.014	.014	.014	-.061	.000
rater AS3	-.200	-.400	.850	-.100	.500	-.650	.000	.004	.014	.014	.014	.014	-.061	.000
rater CL9	.133	.133	-.017	-.167	-.167	.083	.000	.004	.014	.014	.014	.014	-.061	.000
rater CS11	-.133	-.333	-.083	-.033	.367	-.083	-.300	-.013	-.043	-.043	-.043	-.043	-.043	-.225
sum	.000	.000	.000	.000	.000	-.300		.000	.000	.000	.000	.000	-.225	

ตารางที่ 51 ขนาดความแตกต่างของค่าประมาณความคลาดเคลื่อนแบบฮาโลระหว่างมาตรกราฟฟิกที่ปรับแก้ความคลาดเคลื่อนแบบฮาโลและมาตรประเมินค่าเชิงพฤติกรรม (BARS) ที่ไม่ปรับแก้ความคลาดเคลื่อน

rating(sy4)	ขนาดความแตกต่างของค่าประมาณความคลาดเคลื่อนแบบฮาโล						
	ratee 1	ratee 2	ratee 3	ratee 4	ratee 5	ratee 6	sum
rater AL1	0.196	0.586	-0.764	0.286	-0.714	0.411	0
rater AS3	-0.204	-0.414	0.836	-0.114	0.486	-0.589	0
rater CL9	0.129	0.119	-0.031	-0.181	-0.181	0.144	0
rater CS11	-0.12	-0.29	-0.04	0.01	0.41	-0.04	-0.075
sum	0	0	0	0	0	-0.075	

จากขนาดของความแตกต่างของค่าประมาณความคลาดเคลื่อนแบบฮาโลระหว่างมาตรกราฟฟิกที่ปรับแก้ความคลาดเคลื่อนแบบฮาโลและมาตรประเมินค่าเชิงพฤติกรรม (BARS) ที่ไม่ปรับแก้ความคลาดเคลื่อน จะเห็นได้ว่า ขนาดของความแตกต่างที่มีค่าต่ำกว่า .05 มีจำนวนเพียง 4 ค่าจาก 24 ค่าน้อยกว่าร้อยละ 80 ของทั้งหมดแสดงว่าค่าประมาณความคลาดเคลื่อนแบบฮาโลจากมาตรทั้งสองไม่สอดคล้องกัน (Joreskog, 1970) ส่วนการเปรียบเทียบขนาดของ

ความแตกต่างของค่าประมาณความคลาดเคลื่อนแบบฮาโล กรณีอื่น ๆ สรุปได้ดังนี้  
 ตารางที่ 52 สรุปความสอดคล้องของค่าประมาณความคลาดเคลื่อนแบบฮาโลระหว่างมาตร  
 กราฟฟิก (graphic rating scale) ที่ปรับแก้ความคลาดเคลื่อนแบบฮาโลและมาตรประเมินค่าเชิง  
 พฤติกรรม (BARS) ที่ไม่ปรับแก้ความคลาดเคลื่อน

ประเภทมาตร	กลุ่มผู้ประเมิน	จำนวนผู้ประเมิน	กรณี ที่	Graphic rating scale ปรับแก้ halo error	BARS ไม่ปรับแก้ ความคลาดเคลื่อน
สมมาตร	ปล่อยคะแนน	4 คนชุด 1	1		ไม่สอดคล้อง
		4 คนชุด 2	2		ไม่สอดคล้อง
		8 คน	3		ไม่สอดคล้อง
	กวดคะแนน	4 คนชุด 1	4		ไม่สอดคล้อง
		4 คนชุด 2	5		ไม่สอดคล้อง
		8 คน	6		สอดคล้อง
อสมมาตร	ปล่อยคะแนน	4 คนชุด 1	7		ไม่สอดคล้อง
		4 คนชุด 2	8		ไม่สอดคล้อง
		8 คน	9		ไม่สอดคล้อง
	กวดคะแนน	4 คนชุด 1	10		ไม่สอดคล้อง
		4 คนชุด 2	11		ไม่สอดคล้อง
		8 คน	12		สอดคล้อง
สมมาตร	สนิท	4 คนชุด 1	13		ไม่สอดคล้อง
		4 คนชุด 2	14		ไม่สอดคล้อง
		8 คน	15		ไม่สอดคล้อง
	ไม่สนิท	4 คนชุด 1	16		ไม่สอดคล้อง
		4 คนชุด 2	17		ไม่สอดคล้อง
		8 คน	18		สอดคล้อง
อสมมาตร	สนิท	4 คนชุด 1	19		ไม่สอดคล้อง
		4 คนชุด 2	20		ไม่สอดคล้อง
		8 คน	21		ไม่สอดคล้อง
	ไม่สนิท	4 คนชุด 1	22		สอดคล้อง
		4 คนชุด 2	23		สอดคล้อง
		8 คน	24		สอดคล้อง

จากตารางสรุปความสอดคล้องของค่าประมาณความคลาดเคลื่อนแบบฮาโลระหว่างมาตรประเมินค่าแบบกราฟฟิก (graphic rating scale) ที่ปรับแก้ความคลาดเคลื่อนแบบฮาโลและมาตรประเมินค่าเชิงพฤติกรรม (BARS) ที่ไม่ปรับแก้ความคลาดเคลื่อน จะเห็นได้ว่าเมื่อใช้มาตรแบบผสมมาตรกรณี ผู้ประเมินไม่สนิทกับผู้ถูกประเมินผลการเปรียบเทียบจะสอดคล้องกัน อย่างไรก็ตามพบว่ามีการที่สอดคล้องกันเพียง 6 กรณีจาก 24 กรณี จึงสรุปว่าค่าประมาณความคลาดเคลื่อนจากวิธีทั้งสองไม่สอดคล้องกัน กล่าวคือมาตรประเมินค่าเชิงพฤติกรรม (BARS) ป้องกันความคลาดเคลื่อนแบบฮาโลไม่ได้

### 5.3 เปรียบเทียบความสอดคล้องของผลการประเมินผลการปฏิบัติงานของครู

ผู้วิจัยพิจารณาความสอดคล้องของผลการประเมินการปฏิบัติงานของครูของมาตรประเมินค่าเชิงพฤติกรรม (BARS) ที่ไม่ได้ปรับแก้ความคลาดเคลื่อน และผลการประเมินการปฏิบัติงานของครูของมาตรแบบกราฟฟิกที่ปรับแก้ความคลาดเคลื่อนแบบฮาโล โดยเปรียบเทียบใน 2 กรณีดังนี้ คือ

กรณีที่ 1 เปรียบเทียบความสอดคล้องของผลการประเมินการปฏิบัติงานของครูของมาตรแบบกราฟฟิกที่ปรับแก้ความคลาดเคลื่อนแบบฮาโล และ มาตรประเมินค่าเชิงพฤติกรรม (BARS) ที่ไม่ได้ปรับแก้ความคลาดเคลื่อน เทียบกับเกณฑ์การประเมินจากผู้ทรงคุณวุฒิ โดย ขนาดความสอดคล้องของคะแนนตามสูตรของ Downie และ Heath (1974)

กรณีที่ 2 เปรียบเทียบความสอดคล้องของผลการประเมินการปฏิบัติงานของครูของมาตรแบบกราฟฟิกที่ปรับแก้ความคลาดเคลื่อนแบบฮาโล (halo error) เทียบกับผลการประเมินการปฏิบัติงานของครูของมาตรประเมินค่าเชิงพฤติกรรม (BARS) ที่ไม่ได้ปรับแก้ความคลาดเคลื่อน

โดยพิจารณาจากค่าสัมประสิทธิ์สหสัมพันธ์และค่าเฉลี่ยผลต่างกำลังสอง (mean square difference) ได้ผลดังนี้

#### 5.3.1 เทียบกับเกณฑ์การประเมินจากผู้ทรงคุณวุฒิ

ตารางที่ 53 ขนาดความสอดคล้องของผลการประเมินของผู้ประเมินแต่ละคน ( $r_{xx}$ ) เมื่อเทียบกับผู้ทรงคุณวุฒิ ตามวัตถุประสงค์การวิจัยข้อที่ 2

ผู้ประเมิน	มาตรฐานค่าแบบกราฟฟิก ปรับแก้ halo error			มาตรฐาน BARS ไม่ปรับแก้ความคลาดเคลื่อน		
	สมมาตร	อสมมาตร ทางบวก	อสมมาตร ทางลบ	สมมาตร	อสมมาตร ทางบวก	อสมมาตร ทางลบ
ผู้ทรงคุณวุฒิทั้ง 3 ท่าน	.82	.80	.84	.87	.85	.84
ผู้ทรงคุณวุฒิทั้ง 3 ท่าน และผู้ประเมิน AL1	.83	.82		*,.81	.82	
ผู้ทรงคุณวุฒิทั้ง 3 ท่าน และผู้ประเมิน AL2	.85	.87		*,.82	.86	
ผู้ทรงคุณวุฒิทั้ง 3 ท่าน และผู้ประเมิน AS3	.83		.80	*,.78		.80
ผู้ทรงคุณวุฒิทั้ง 3 ท่าน และผู้ประเมิน AS4	.82		.82	*,.80		.82
ผู้ทรงคุณวุฒิทั้ง 3 ท่าน และผู้ประเมิน BL5	.82	.87		*,.82	.85	
ผู้ทรงคุณวุฒิทั้ง 3 ท่าน และผู้ประเมิน BL6	*,.79	.87		*,.79	.87	
ผู้ทรงคุณวุฒิทั้ง 3 ท่าน และผู้ประเมิน BS7	.83		.81	*,.82		.83
ผู้ทรงคุณวุฒิทั้ง 3 ท่าน และผู้ประเมิน BS8	.84		.82	.83		.83
ผู้ทรงคุณวุฒิทั้ง 3 ท่าน และผู้ประเมิน CL9	.87	.87		.85	.86	
ผู้ทรงคุณวุฒิทั้ง 3 ท่าน และผู้ประเมิน CL10	.86	.87		.84	.84	
ผู้ทรงคุณวุฒิทั้ง 3 ท่าน และผู้ประเมิน CS11	.87		.85	.86		.85
ผู้ทรงคุณวุฒิทั้ง 3 ท่าน และผู้ประเมิน CS12	.87		.83	*,.80		.80
ผู้ทรงคุณวุฒิทั้ง 3 ท่าน และผู้ประเมิน DL13	.83	.84		.82	.82	
ผู้ทรงคุณวุฒิทั้ง 3 ท่าน และผู้ประเมิน DL14	.81	.85		.84	.84	
ผู้ทรงคุณวุฒิทั้ง 3 ท่าน และผู้ประเมิน DS15	.84		.85	.87		.82
ผู้ทรงคุณวุฒิทั้ง 3 ท่าน และผู้ประเมิน DS16	.85		.84	.87		.83

ผลการประเมินของผู้ประเมินจากมาตรประเมินค่าแบบกราฟฟิก เมื่อปรับแก้ความคลาดเคลื่อนแบบฮาโลแล้วสอดคล้องกับผลของผู้ทรงคุณวุฒิ 15 จาก 16 คน และ ส่วนใหญ่มีค่าขนาดความสอดคล้อง ( $r_{xx} > .8$ ) มากกว่าผลของมาตรประเมินค่าเชิงพฤติกรรม (BARS) ที่ไม่ได้ปรับแก้ความคลาดเคลื่อน โดยเฉพาะอย่างยิ่งผู้ประเมินที่มีความสนิทคุ้นเคยกับผู้ประเมินกับผู้ถูกประเมิน (AL1,AL2,AS3, AS4,CL9,CL10, CS11 และ CS12) เมื่อปรับแก้ความคลาดเคลื่อนแบบฮาโล จะได้ผลการประเมินที่มีขนาดความสอดคล้องมากขึ้น และมากกว่า เกณฑ์ความสอดคล้องของผู้ทรงคุณวุฒิ

ดังนั้นจึงสรุปได้ว่าการปรับแก้ความคลาดเคลื่อนแบบฮาโล ทำให้ได้ผลการประเมินที่มีคุณภาพสอดคล้องกับเกณฑ์การประเมินของผู้ทรงคุณวุฒิ มากกว่าการใช้มาตรประเมินค่าเชิงพฤติกรรม (BARS) ที่ไม่ได้ปรับแก้ความคลาดเคลื่อน

5.3.2 เปรียบเทียบความสอดคล้องของผลการประเมินการปฏิบัติงานของครูของมาตรแบบกราฟฟิกที่ปรับแก้ความคลาดเคลื่อนแบบฮาโล (halo error) เทียบกับผลการประเมินการปฏิบัติงานของครูของมาตรประเมินค่าเชิงพฤติกรรม (BARS) ที่ไม่ได้ปรับแก้ความคลาดเคลื่อน

ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย



ตารางที่ 54 กรณีกาการวิเคราะห์ความสอดคล้องผลการปฏิบัติงานของครูตามวัตถุประสงค์ข้อที่ 2

ลักษณะมาตร	กลุ่มผู้ประเมิน	จำนวนผู้ประเมิน	กรณีที	กราฟฟิค ปรับแก้ halo error เทียบมาตร BARS ไม่ปรับแก้ความคลาดเคลื่อน	
				Mean square	Corr
สมมาตร	กตคะแนน	4 คนชุด 1	1	.763	.720
		4 คนชุด 2	2	.724	.692
		8 คน	3	.745	.512
	ปล่อยคะแนน	4 คนชุด 1	4	.660	.754
		4 คนชุด 2	5	.721	.526
		8 คน	6	.690	.772
อสมมาตร	กตคะแนน	4 คนชุด 1	7	.771	.426
		4 คนชุด 2	8	1.147	.314
		8 คน	9	.874	.429
	ปล่อยคะแนน	4 คนชุด 1	10	.973	.324
		4 คนชุด 2	11	.505	.217*
		8 คน	12	1.015	.209
สมมาตร	สนิท	4 คนชุด 1	13	1.002	.314
		4 คนชุด 2	14	.932	.414
		8 คน	15	1.026	.287
	ไม่สนิท	4 คนชุด 1	16	.877	.310
		4 คนชุด 2	17	1.721	.312
		8 คน	18	.875	.358
อสมมาตร	สนิท	4 คนชุด 1	19	1.193	.247
		4 คนชุด 2	20	1.326	.219
		8 คน	21	1.222	.306
	ไม่สนิท	4 คนชุด 1	22	1.234	.214
		4 คนชุด 2	23	1.159	.303
		8 คน	24	1.196	.294

\* p &lt; .05

เมื่อเปรียบเทียบผลของมาตรประเมินค่าแบบกราฟฟิค ที่ปรับแก้ความคลาดเคลื่อนแบบฮาโล และผลของมาตรประเมินค่าเชิงพฤติกรรมที่ไม่ได้ปรับแก้ความคลาดเคลื่อนพบว่า มีกรณีที่สอดคล้องกันเพียงกรณีเดียว จาก 24 กรณี จึงสรุปได้ว่า ผลการประเมินการปฏิบัติงานของครู จากวิธีป้องกันความคลาดเคลื่อนแบบฮาโล ด้วยมาตรประเมินค่าเชิงพฤติกรรม (BARS) และวิธีการใช้มาตรประเมินค่าแบบกราฟฟิค ร่วมกับการประมาณค่าและปรับแก้ความคลาดเคลื่อนแบบฮาโล ให้ผลไม่สอดคล้องกัน

## ตอนที่ 6 เปรียบเทียบผลการใช้มาตรประเมินค่าเชิงพฤติกรรมแบบอสมมาตรและการใช้มาตรกราฟฟิคแบบสมมาตรร่วมกับการประมาณค่าและปรับแก้ความคลาดเคลื่อน

ขั้นตอนนี้เป็นขั้นตอนการวิเคราะห์ข้อมูล ที่มีจุดมุ่งหมายเพื่อเปรียบเทียบความสอดคล้องของผลการประเมินผลการปฏิบัติงานของครูและความคลาดเคลื่อนแบบกด/ปล่อยคะแนน (leniency error) และความคลาดเคลื่อนแบบฮาโล (halo error) ระหว่างวิธีการ 3 วิธีคือ (1) วิธีป้องกันความคลาดเคลื่อนแบบ กด/ปล่อยคะแนนและแบบฮาโล โดยใช้มาตรประเมินค่าเชิงพฤติกรรม ( BARS) แบบอสมมาตร (2) วิธีการใช้มาตรประเมินค่าแบบกราฟฟิค (graphic rating scale) แบบสมมาตร ร่วมกับการประมาณค่าและปรับแก้ความคลาดเคลื่อนของ Guilford (3) วิธีการวิเคราะห์ความแปรปรวนโดยใช้ทฤษฎีการสรุปอ้างอิง (generalizability theory) ตามวัตถุประสงค์การวิจัยข้อที่ 3

แต่เนื่องจากหลักการของ G-theory และ วิธี Guilford ใช้หลักการแนวเดียวกัน ผลจากการวิเคราะห์ความแปรปรวนของทั้งสองวิธีจึงมีค่าเท่ากัน (ภาคผนวก ค) ค่าประมาณความคลาดเคลื่อนและค่าปรับแก้ความคลาดเคลื่อนจึงเท่ากันด้วย ดังนั้นผู้วิจัยจึงไม่นำเสนอผลประมาณค่าและปรับแก้ความคลาดเคลื่อนด้วยวิธี G-theory

### 6.1 การประมาณค่าความคลาดเคลื่อนทั้งสามแบบตามแนวคิด Guilford

ผู้วิจัยขอยกตัวอย่างข้อมูลและการประมาณค่าความคลาดเคลื่อนทั้ง 2 แบบ (แบบกด/ปล่อยคะแนน และแบบฮาโล) ตามแนวคิด Guilford เมื่อใช้มาตรประเมินค่าแบบกราฟฟิค (graphic rating scale) แบบสมมาตร กับผู้ประเมินที่มีความสนิทกับโรงเรียนอัสสัมชัญ จำนวน 4 คน ชุดที่ 1 (AL1, AS3, CL9 และ CS11)

จากคะแนนผลการประเมินเมื่อใช้มาตรประเมินค่าแบบกราฟฟิค (graphic rating scale) แบบสมมาตร กับผู้ประเมินที่มีความสนิทกับโรงเรียนอัสสัมชัญ จำนวน 4 คน ชุดที่ 1 (ตารางที่ 46) ประมาณค่าความคลาดเคลื่อนแบบกด/ปล่อยคะแนน และแบบฮาโล ได้ค่าดังตารางที่ 47 และตารางที่ 48 ลำดับ

ประมาณค่าความคลาดเคลื่อนแบบขัดแย้ง ( $X'_{kj}$ ) โดยดำเนินการดังนี้

1. หาค่าเฉลี่ยคะแนนการประเมินในแต่ละเซลล์ ให้ได้ค่าเฉลี่ยผลการประเมินทุกคุณลักษณะของผู้ถูกประเมิน
2. หาค่าเฉลี่ยคะแนนประเมินของผู้ประเมิน (rater) ได้ค่าเฉลี่ยในคอลัมน์ all traits
3. หาค่าเฉลี่ยคะแนนประเมินของคุณลักษณะ (trait) ได้ค่าเฉลี่ยในแถว all raters
4. หาค่าเฉลี่ยรวม (grand mean)
5. หาค่าประมาณความคลาดเคลื่อนแบบกด/ปล่อยคะแนน (leniency error =  $X'_{kl}$ )  
(mean all rates - grand mean)
6. หาค่า  $d_i$  จาก mean all raters - grand mean
7. เฉลี่ยคะแนนการประเมินในแต่ละเซลล์ ที่หักออกด้วยค่าประมาณความคลาดเคลื่อนแบบกด/ปล่อยคะแนน ( $X'_{kl}$ ) และค่า  $d_i$  และ ลบด้วยค่าเฉลี่ยรวม (grand mean) ค่าที่คำนวณได้แต่ละเซลล์ คือค่าประมาณความคลาดเคลื่อนแบบขัดแย้ง (contrast error =  $X'_{kj}$ ) ซึ่งได้ผลดังตารางต่อไปนี้

ตารางที่ 55 ค่าประมาณความคลาดเคลื่อนแบบขัดแย้ง

rating(sy4)	trait 1	trait 2	trait 3	trait 4	trait 5	sum
rater AL1	-0.142	-0.183	-0.058	0.358	0.025	0.000
rater AS3	-0.308	-0.183	0.275	0.192	0.025	0.000
rater CL9	0.258	0.550	-0.325	-0.408	-0.075	0.000
rater CS11	0.192	-0.183	0.108	-0.142	0.025	0.000
sum	0.000	0.000	0.000	0.000	0.000	

ปรับแก้ความคลาดเคลื่อนในการประเมินตามสูตรของ Guilford ดังสมการที่ 17

$$X_{ijk} = X_{ijk} - X_{kl} - X_{ki} - X_{kj} \quad \text{-----} \quad (17)$$

ผลการประเมินที่ปรับแก้ตามสูตรในสมการที่ (17) ได้ผลดัง ตารางต่อไปนี้

ตารางที่ 56 ผลการประเมินที่ปรับแก้ความคลาดเคลื่อนแบบกต/ปล่อยคะแนน ความคลาดเคลื่อนแบบฮาโล และความคลาดเคลื่อนแบบขัดแย้ง

trait	rater AL1					rater AL2					rater AS3					rater AS4				
	T1	T2	T3	T4	T5	T1	T2	T3	T4	T5	T1	T2	T3	T4	T5	T1	T2	T3	T4	T5
ratee 1	3.84	3.88	3.76	3.34	2.68	2.41	3.28	3.83	4.91	3.08	3.24	3.95	3.82	3.91	2.57	2.71	3.89	3.60	3.85	3.68
ratee 2	3.44	3.48	3.36	2.94	2.28	3.61	2.48	4.03	3.11	2.28	4.24	3.95	2.82	2.91	1.57	3.71	3.09	3.80	3.05	1.88
ratee 3	4.79	4.83	4.71	4.29	4.63	5.36	5.23	3.78	4.86	4.03	4.39	4.10	4.97	5.06	4.72	5.46	4.84	4.55	3.80	4.63
ratee 4	3.74	3.78	3.66	3.24	2.58	3.31	4.18	3.73	2.81	2.98	3.54	4.25	4.12	2.21	2.87	4.41	3.79	3.50	2.75	2.58
ratee 5	4.74	4.78	4.66	4.24	4.58	5.71	5.58	5.13	3.21	3.38	4.54	4.25	5.12	4.21	4.87	4.01	5.39	5.10	4.35	4.18
ratee 6	3.69	3.73	3.61	3.19	2.53	3.86	3.73	3.28	2.36	3.53	4.29	4.00	2.87	2.96	2.62	4.46	3.84	3.55	3.80	2.63

นำผลการประเมินที่ปรับแก้แล้วมาคำนวณหาค่าประมาณความคลาดเคลื่อนแบบกต/ปล่อยคะแนนและค่าประมาณความคลาดเคลื่อนแบบฮาโล อีกครั้งหนึ่ง ได้ผลดังตารางที่ 57 และ ตารางที่ 58 ตามลำดับ

ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 57 ค่าประมาณความคลาดเคลื่อนแบบกต/ปล่อยคะแนนของผลการประเมินที่ปรับแก้แล้ว

BARS (asy4)	ratee 1	ratee 2	ratee 3	ratee 4	ratee 5	ratee 6	all ratees	KL (LENIEN)
rater AL1	3.386	3.286	3.586	3.236	3.736	3.286	3.4	0.11
rater AS3	1.7	1.6	1.9	1.5	2.0	1.6	1.7	-1.62
rater CL9	4.4	4.3	4.6	4.2	4.7	4.3	4.4	1.11
rater CS11	3.6	3.5	3.8	3.4	3.9	4.0	3.7	0.40
all raters	3.26	3.15	3.45	3.10	3.60	3.28	3.31	
di	-0.05	-0.15	0.15	-0.21	0.30	-0.03		

ตารางที่ 58 ค่าประมาณความคลาดเคลื่อนแบบฮาโลของผลการประเมินที่ปรับแก้แล้ว

ratee 1	ratee 2	ratee 3	ratee 4	ratee 5	ratee 6	sum
0.017	0.022	0.022	0.022	0.022	-0.103	0.000
0.017	0.022	0.022	0.022	0.022	-0.103	0.000
0.017	0.022	0.022	0.022	0.022	-0.103	0.000
-0.050	-0.065	-0.065	-0.065	-0.065	-0.065	-0.375
0.000	0.000	0.000	0.000	0.000	-0.375	

## 5.2 เปรียบเทียบความสอดคล้องค่าประมาณความคลาดเคลื่อน

วัตถุประสงค์การวิจัยข้อที่ 3 ศึกษาเปรียบเทียบค่าประมาณความคลาดเคลื่อน 2 แบบคือ

### 5.2.1 เปรียบเทียบความสอดคล้องค่าประมาณความคลาดเคลื่อนแบบ กต/ปล่อยคะแนน

การเปรียบเทียบค่าประมาณความคลาดเคลื่อนแบบกต/ปล่อยคะแนน ของวิธีการทั้งสองวิธี ศึกษาโดยการเปรียบเทียบจากขนาดของความแตกต่างของค่าประมาณความคลาดเคลื่อนได้ผลดังนี้

ตารางที่ 59 ค่าประมาณความคลาดเคลื่อนแบบกต/ปล่อย

ผู้ประเมิน			BAR อสมมาตร (asymmetry) ไม่ปรับแก้						มาตรกราฟฟิกแบบสมมาตร ปรับแก้ 3 อย่าง						ขนาดความแตกต่าง					
			ปล่อยคะแนน			กตคะแนน			ปล่อยคะแนน			กตคะแนน			ปล่อยคะแนน			กตคะแนน		
			8 คน	4คน ชุด1	4คน ชุด2	8 คน	4คน ชุด1	4คน ชุด2	8 คน	4คน ชุด1	4คน ชุด2	8 คน	4คน ชุด1	4คน ชุด2	8 คน	4คน ชุด1	4คน ชุด2	8 คน	4คน ชุด1	4คน ชุด2
ครูอัตราจ้าง	ปล่อย	AL1	.36	.40					0	-.05				.36	.45					
	คะแนน	AL2	.29		.25				0		-.4			.29		.65				
	กต	AS3				-.12	-.14						-.01	-.03		0	0			
	คะแนน	AS4				-.09		-.07					.05		.01	0		0		
ครูต่างโรงเรียน	ปล่อย	BL5	.02	.07					0	-.05							0	0		
	คะแนน	BL6	-.04		-.08				-.04		-.07						0		0	
	กต	BS7				.01	-.01						.06	-.02				-.05	.01	
	คะแนน	BS8				-.09		-.07					-.08		.01			-.01	-.08	
ผู้ปกครองชั้นต้น	ปล่อย	CL9	-.44	-.40					0	.04				-.44	-.44					
	คะแนน	CL10	.16		.12				0		.58			.16		-.46				
	กต	CS11				.15	.13						.06	.07		0	0			
	คะแนน	CS12				.08		.10					-.03		.02	0		0		
นายจ้าง	ปล่อย	DL13	-.11	-.07					0	.06							0	0		
	คะแนน	DL14	-.24		-.28				.02		-.11						0		0	
	กต	DS15				.05	.03						-.02	-.02				.07	.05	
	คะแนน	DS16				.01		.03					-.03	-.04				.04		.07

จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 59 ค่าประมาณความคลาดเคลื่อนแบบกต/ปล่อย (ต่อ)

ผู้ประเมิน			BAR อสมมาตร (asymmetry) ไม่ปรับแก้						มาตรกราฟฟิกแบบสมมาตร ปรับแก้ 3 อย่าง						ขนาดความแตกต่าง					
			สนิทกับอัสสมิตรี			ไม่สนิทกับอัสสมิตรี			สนิทกับอัสสมิตรี			ไม่สนิทกับอัสสมิตรี			สนิทกับอัสสมิตรี			ไม่สนิทกับอัสสมิตรี		
			8 คน	4คน ชุด1	4คน ชุด2	8 คน	4คน ชุด1	4คน ชุด2	8 คน	4คน ชุด1	4คน ชุด2	8 คน	4คน ชุด1	4คน ชุด2	8 คน	4คน ชุด1	4คน ชุด2	8 คน	4คน ชุด1	4คน ชุด2
ครูอัสสมิตรี	ปล่อย	AL1	.01	-.03				.95	1.0					-.94	-1.03					
	คะแนน	AL2	.02		.02			.89		.83				-.87		-.81				
	กต	AS3	.01	-.03				-.81	-.75					.82	.72					
	คะแนน	AS4	-.1		.03			-.78		-.84				.68		.87				
ครูสำรวจโรงเรียน	ปล่อย	BL5				.03	.04					.72	.68				-.69	-.64		
	คะแนน	BL6				.03		-.46				.65		.69			-.62		-1.15	
	กต	BS7				.03	.04					-.58	-.63				.61	.67		
	คะแนน	BS8				-.24		-1.69				-.68		-.64			.44		-1.05	
ผู้ปกครอง	ปล่อย	CL9	.02	.03				.15	.22					-.13	-.19					
	คะแนน	CL10	.01		-.22			.75		.69				-.74		-.91				
	กต	CS11	.01	.03				-.55	-.48					.56	.51					
	คะแนน	CS12	.01		.17			-.61		-.68				.62		.85				
ผู้จบปริญญา	ปล่อย	DL13				.03	.04					.58	.54				-.55	-.5		
	คะแนน	DL14				.04		1.79				.45	.49				-.41		1.3	
	กต	DS15				.03	-.13					-.55	-.59				.58	.46		
	คะแนน	DS16				.03		.36				-.58		-.54			.61		.9	

ผลการเปรียบเทียบขนาดของความแตกต่างของค่าประมาณความคลาดเคลื่อนแบบกต/ปล้อยคะแนนของมาตรประเมินค่าแบบกราฟฟิก (graphic rating scale) แบบสมมาตรเมื่อปรับแก้ความคลาดเคลื่อนทั้ง 3 แบบ (ความคลาดเคลื่อนแบบกต/ปล้อยคะแนน, ความคลาดเคลื่อนแบบฮาโล และความคลาดเคลื่อนแบบขัดแย้ง) และมาตรประเมินค่าเชิงพฤติกรรม (BARS) แบบอสมมาตรที่ไม่ปรับแก้ความคลาดเคลื่อน จะเห็นได้ว่าผลต่างมีขนาดเล็กกว่า .05 จำนวน 19 ค่าจาก 64 ค่าซึ่งน้อยกว่าร้อยละ 80 ของค่าประมาณทั้งหมด จึงสรุปได้ว่ามาตรประเมินค่าเชิงพฤติกรรม (BARS) แบบอสมมาตรที่ไม่ปรับแก้ความคลาดเคลื่อน และมาตรประเมินค่าแบบกราฟฟิก (graphic rating scale) แบบสมมาตรเมื่อปรับแก้ความคลาดเคลื่อนทั้ง 3 แบบ ไม่สอดคล้องกัน แสดงว่า การใช้มาตรประเมินค่าเชิง พฤติกรรม (BARS) แบบอสมมาตรป้องกันความคลาดเคลื่อนแบบกต/ปล้อยคะแนนไม่ได้

### 6.2.2 เปรียบเทียบความสอดคล้องค่าประมาณความคลาดเคลื่อนแบบฮาโล

พิจารณาความสอดคล้องของค่าประมาณความคลาดเคลื่อนแบบฮาโลของมาตรประเมินค่าแบบกราฟฟิก (graphic rating scale) แบบสมมาตรเมื่อปรับแก้ความคลาดเคลื่อนทั้ง 3 แบบ (ความคลาดเคลื่อนแบบกต/ปล้อยคะแนน, ความคลาดเคลื่อนแบบฮาโล และความคลาดเคลื่อนแบบขัดแย้ง) และมาตรประเมินค่าเชิงพฤติกรรม (BARS) แบบอสมมาตร ที่ไม่ปรับแก้ความคลาดเคลื่อน ผู้วิจัยขอยกตัวอย่างการเปรียบเทียบค่าประมาณความคลาดเคลื่อนแบบฮาโล ทั้ง 2 วิธี กรณี ผู้ประเมินสนิทกับโรงเรียนอัสสัมชัญ 4 คน ชุดที่ 1

ตารางที่ 60 ขนาดความแตกต่างของค่าประมาณความคลาดเคลื่อนแบบฮาโลระหว่างมาตรกราฟฟิกแบบสมมาตรที่ปรับแก้ความคลาดเคลื่อนทั้ง 3 แบบและมาตรประเมินค่าเชิงพฤติกรรม (BARS) แบบอสมมาตรที่ไม่ปรับแก้ความคลาดเคลื่อน กรณีผู้ประเมินสนิท 4 คน ชุด 1

	ขนาดความแตกต่างของค่าประมาณความคลาดเคลื่อนแบบฮาโล						
	ratee 1	ratee 2	ratee 3	ratee 4	ratee 5	ratee 6	sum
rater AL1	0.129	0.019	0.119	-0.131	-0.031	-0.106	0
rater AS3	0.096	-0.014	0.086	0.036	-0.064	-0.139	0
rater CL9	-0.271	0.019	-0.081	0.069	0.169	0.094	0
rater CS11	0.045	-0.024	-0.124	0.026	-0.074	-0.124	-0.525
sum	0	0	0	0	0	-0.275	

จากขนาดของความแตกต่างของค่าประมาณความคลาดเคลื่อนแบบฮาโลระหว่างมาตร



ประเมินค่าแบบกราฟฟิค (graphic rating scale) แบบสมมาตรที่ปรับแก้ความคลาดเคลื่อนทั้ง 3 แบบ และมาตรประเมินค่าเชิงพฤติกรรม (BARS) แบบอสมมาตร ที่ไม่ปรับแก้ความคลาดเคลื่อน จะเห็นได้ว่า ขนาดของความแตกต่างที่มีค่าต่ำกว่า .05 มีจำนวนเพียง 8 ค่าจาก 24 ค่า ซึ่งน้อยกว่าร้อยละ 80 ของทั้งหมดแสดงว่าค่าประมาณความคลาดเคลื่อนแบบฮาโลจากมาตรทั้งสองไม่สอดคล้องกัน (Joreskog, 1970) ส่วนการเปรียบเทียบขนาดของความแตกต่างของค่าประมาณความคลาดเคลื่อนแบบฮาโล กรณีอื่น ๆ สรุปได้ดังนี้

ตารางที่ 61 สรุปความสอดคล้องของค่าประมาณความคลาดเคลื่อนแบบกต/ปล่อยคะแนนและแบบฮาโลระหว่างมาตรกราฟฟิคแบบสมมาตรที่ปรับแก้ความคลาดเคลื่อนกับมาตรประเมินค่าเชิงพฤติกรรม (BARS) แบบอสมมาตรที่ไม่แก้ความคลาดเคลื่อน

กลุ่มผู้ประเมิน	จำนวนผู้ประเมิน	กรณีที่	graphic rating scale แบบสมมาตร ปรับแก้ความคลาดเคลื่อนทั้ง 3 แบบ เทียบกับ BARS แบบอสมมาตร ไม่ปรับแก้ ความคลาดเคลื่อน	
			ความสอดคล้องของค่าประมาณ leniency error	ความสอดคล้องของค่าประมาณ halo error
สนิท	4 คน ชุด 1	1	ไม่สอดคล้อง	ไม่สอดคล้อง
	4 คน ชุด 2	2	ไม่สอดคล้อง	ไม่สอดคล้อง
	8 คน	3	ไม่สอดคล้อง	ไม่สอดคล้อง
ไม่สนิท	4 คน ชุด 1	4	ไม่สอดคล้อง	ไม่สอดคล้อง
	4 คน ชุด 2	5	ไม่สอดคล้อง	ไม่สอดคล้อง
	8 คน	6	ไม่สอดคล้อง	ไม่สอดคล้อง
กตคะแนน	4 คน ชุด 1	7	ไม่สอดคล้อง	ไม่สอดคล้อง
	4 คน ชุด 2	8	ไม่สอดคล้อง	ไม่สอดคล้อง
	8 คน	9	สอดคล้อง	ไม่สอดคล้อง
ปล่อยคะแนน	4 คน ชุด 1	10	ไม่สอดคล้อง	ไม่สอดคล้อง
	4 คน ชุด 2	11	ไม่สอดคล้อง	สอดคล้อง
	8 คน	12	ไม่สอดคล้อง	ไม่สอดคล้อง

จากตารางสรุปความสอดคล้องของค่าประมาณความคลาดเคลื่อนแบบกต/ปล่อยคะแนนระหว่างมาตรกราฟฟิคแบบสมมาตรที่ปรับแก้ความคลาดเคลื่อนทั้ง 3 แบบและมาตรประเมินค่าเชิงพฤติกรรม (BARS) แบบอสมมาตร ที่ไม่ปรับแก้ความคลาดเคลื่อน มีกรณีที่สอดคล้องกันเพียงกรณีเดียว คือกรณีที่ 9 เมื่อผู้ประเมินมีลักษณะการให้คะแนนแบบกตคะแนน 8 คน

ความสอดคล้องของค่าประมาณความคลาดเคลื่อนแบบฮาโลระหว่าง มาตรประเมินค่าแบบกราฟฟิค แบบสมมาตรที่ปรับแก้ความคลาดเคลื่อนทั้ง 3 แบบ และมาตรประเมินค่าเชิง

พฤติกรรม (BARS) แบบอสมมาตรที่ไม่ปรับแก้ความคลาดเคลื่อน จะเห็นมีกรณีที่สอดคล้องกันเพียงกรณีเดียว คือ กรณีที่ 11 เมื่อผู้ประเมินมีลักษณะการให้คะแนนแบบปล่อยคะแนน 4 คน ชุดที่ 1 จึงสรุปว่าค่าประมาณความคลาดเคลื่อนจากวิธีทั้งสองไม่สอดคล้องกัน กล่าวคือมาตรประเมินค่าเชิงพฤติกรรม (BARS) แบบอสมมาตรป้องกันความคลาดเคลื่อนแบบฮาโลไม่ได้

### 6.3 เปรียบเทียบความสอดคล้องของผลการประเมินผลการปฏิบัติงานของครู

ในการวิเคราะห์ความสอดคล้องของผลการประเมินการปฏิบัติงานของครูของมาตรประเมินค่าเชิงพฤติกรรม (BARS) แบบอสมมาตร ที่ไม่ได้ปรับแก้ความคลาดเคลื่อนและผลการประเมินการปฏิบัติงานของครูของมาตรแบบกราฟฟิกแบบสมมาตรที่ปรับแก้ความคลาดเคลื่อนแบบกด/ปล่อยคะแนน,แบบฮาโล และแบบขัดแย้ง โดยเปรียบเทียบใน 2 กรณีดังนี้ คือ กรณีที่ 1 เปรียบเทียบความสอดคล้องของผลการประเมินการปฏิบัติงานของครูของมาตรแบบกราฟฟิก แบบสมมาตรที่ปรับแก้ความคลาดเคลื่อนทั้ง 3 แบบและประเมินค่าเชิงพฤติกรรม (BARS) แบบอสมมาตร ที่ไม่ได้ปรับแก้ความคลาดเคลื่อน เทียบกับเกณฑ์การประเมินจากผู้ทรงคุณวุฒิ โดย หาขนาดความสอดคล้องของคะแนนตามสูตรของ Downie และ Heath (1974)

กรณีที่ 2 เปรียบเทียบความสอดคล้องของผลการประเมินการปฏิบัติงานของครูของมาตรแบบกราฟฟิกแบบสมมาตรที่ปรับแก้ความคลาดเคลื่อน ทั้ง 3 แบบ เทียบกับผลการประเมินการปฏิบัติงานของครูของมาตรประเมินค่าเชิงพฤติกรรม (BARS) แบบอสมมาตร ที่ไม่ได้ปรับแก้ความคลาดเคลื่อน โดยพิจารณาจากค่าสัมประสิทธิ์สหสัมพันธ์และค่าเฉลี่ยผลต่างกำลังสอง (mean square difference) ได้ผลดังนี้

#### 6.3.1 เปรียบเทียบกับเกณฑ์การประเมินจากผู้ทรงคุณวุฒิ

ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 62 ขนาดความสอดคล้องของผลการประเมินของผู้ประเมินแต่ละคน ( $r_{xx}$ ) เมื่อเทียบกับ ผู้ทรงคุณวุฒิ ตามวัตถุประสงค์การวิจัยข้อที่ 3

ผู้ประเมิน	มาตร Graphic แบบสมมาตร ปรับแก้ leniency error และ halo error	BARS อสมมาตร ไม่ปรับแก้ความคลาดเคลื่อน	
		อสมมาตรทางบวก	อสมมาตรทางลบ
ผู้ทรงคุณวุฒิทั้ง 3 ท่าน	.82	.85	.84
ผู้ทรงคุณวุฒิทั้ง 3 ท่านและผู้ประเมิน AL1	.83	.82	
ผู้ทรงคุณวุฒิทั้ง 3 ท่านและผู้ประเมิน AL2	.86	.86	
ผู้ทรงคุณวุฒิทั้ง 3 ท่านและผู้ประเมิน AS3	.83		.80
ผู้ทรงคุณวุฒิทั้ง 3 ท่านและผู้ประเมิน AS4	.83		.82
ผู้ทรงคุณวุฒิทั้ง 3 ท่านและผู้ประเมิน BL5	.84	.85	
ผู้ทรงคุณวุฒิทั้ง 3 ท่านและผู้ประเมิน BL6	.86	.87	
ผู้ทรงคุณวุฒิทั้ง 3 ท่านและผู้ประเมิน BS7	.84		.83
ผู้ทรงคุณวุฒิทั้ง 3 ท่านและผู้ประเมิน BS8	.82		.83
ผู้ทรงคุณวุฒิทั้ง 3 ท่านและผู้ประเมิน CL9	.86	.86	
ผู้ทรงคุณวุฒิทั้ง 3 ท่านและผู้ประเมิน CL10	.85	.84	
ผู้ทรงคุณวุฒิทั้ง 3 ท่านและผู้ประเมิน CS11	.85		.85
ผู้ทรงคุณวุฒิทั้ง 3 ท่านและผู้ประเมิน CS12	.81		.80
ผู้ทรงคุณวุฒิทั้ง 3 ท่านและผู้ประเมิน DL13	.83	.82	
ผู้ทรงคุณวุฒิทั้ง 3 ท่านและผู้ประเมิน DL14	.83	.84	
ผู้ทรงคุณวุฒิทั้ง 3 ท่านและผู้ประเมิน DS15	.83		.82
ผู้ทรงคุณวุฒิทั้ง 3 ท่านและผู้ประเมิน DS16	.84		.83

หมายเหตุ \* หมายถึง แตกต่างอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

จากตารางข้างต้น ผลจากมาตรกราฟฟิกแบบสมมาตร ที่ปรับแก้ความคลาดเคลื่อนทั้ง 3 แบบ (ความคลาดเคลื่อนแบบกด/ปล่อยคะแนน, ความคลาดเคลื่อนแบบฮาโล และ ความคลาดเคลื่อนแบบขัดแย้ง) พบว่ามีผู้ประเมิน 15 จาก 16 คน ที่ผลการประเมินมีขนาดความสอดคล้องมากกว่าเกณฑ์การประเมินของผู้ทรงคุณวุฒิ ( $r_{xx} > .82$ ) และพบว่าขนาดความสอดคล้องของผลการประเมินจากมาตรกราฟฟิกแบบสมมาตร ที่ปรับแก้ความคลาดเคลื่อนทั้ง 3 แบบ และผลการประเมินจากมาตรประเมินค่าเชิงพฤติกรรม (BARS) แบบอสมมาตรที่ไม่ได้ปรับแก้ความคลาดเคลื่อนมีค่าใกล้เคียงกัน

ดังนั้น การใช้มาตรประเมินค่าเชิงพฤติกรรม (BARS) แบบอสมมาตร ให้ผลการประเมินที่มีคุณภาพสอดคล้องกับเกณฑ์การประเมินของผู้ทรงคุณวุฒิใกล้เคียงกับการใช้ มาตรกราฟฟิกแบบสมมาตร ที่ปรับแก้ความคลาดเคลื่อนทั้ง 3 แบบ

ตารางที่ 63 กรณีการวิเคราะห์ความสอดคล้องผลการปฏิบัติงานของครูตามวัตถุประสงค์ข้อที่ 3

กลุ่มผู้ประเมิน	จำนวนผู้ประเมิน	กรณี	กราฟฟิค แบบสมมาตร ปรับแก้ ความคลาดเคลื่อน เทียบ BARS อสมมาตร ไม่ปรับแก้ความคลาดเคลื่อน	
			Mean square	Corr
กตคะแนน	4 คนชุด 1	1	0.467	.651*
	4 คนชุด 2	2	0.695	.523*
	8 คน	3	0.557	.491
ปล่อยคะแนน	4 คนชุด 1	4	0.622	.562*
	4 คนชุด 2	5	0.398	.753*
	8 คน	6	0.552	.498*
สนิท	4 คนชุด 1	7	0.898	.331
	4 คนชุด 2	8	0.825	.234*
	8 คน	9	0.866	.332*
ไม่สนิท	4 คนชุด 1	10	0.798	.512*
	4 คนชุด 2	11	0.625	.223*
	8 คน	12	0.719	.415*

ผลการเปรียบเทียบความสอดคล้องพบว่ามาตรกราฟฟิคแบบสมมาตรที่ปรับแก้ความคลาดเคลื่อนทั้ง 3 แบบ และผลของมาตรประเมินค่าเชิงพฤติกรรมแบบอสมมาตรที่ไม่ได้ปรับแก้ความคลาดเคลื่อนพบว่ามีกรณีที่สอดคล้องกัน 10 กรณีเดียว จาก 12 กรณี จึงสรุปได้ว่ามาตรกราฟฟิคแบบสมมาตรที่ปรับแก้ความคลาดเคลื่อนทั้ง 3 แบบ และมาตรประเมินค่าเชิงพฤติกรรมแบบอสมมาตรที่ไม่ได้ปรับแก้ความคลาดเคลื่อนให้ผลการประเมินการปฏิบัติงานของครูสอดคล้องกัน