บทที่ 4
ความน่าเชื่อถือของแบบทดสอบ

เค้าโครงเรื่อง

1. การหาค่าสหสัมพันธ์และการตีความหมาย
2. การหาค่าความน่าเชื่อถือ
3. การตีความหมายความน่าเชื่อถือ
4. ค่าความน่าเชื่อถือจะสูงหรือต่ำขึ้นอยู่กับอะไรบ้าง

สาระสำคัญ

1. สหสัมพันธ์ หมายถึง ความเกี่ยวข้องหรือความสัมพันธ์ระหว่างคะแนน 2 ชุด โดยความสัมพันธ์จะเป็นไปได้ 2 แบบ ได้แก่ positive และ negative

2. การหาค่าความน่าเชื่อถือหาได้ 4 วิธีคือ แบบทดสอบซ้ำ แบบคู่ขนาน แบบแบ่งครึ่งและวิธีการสถิติ

3. ค่าความน่าเชื่อถือจะสูงหรือต่ำขึ้นอยู่กับพิสัยของกลุ่ม ระดับความสามารถภายในกลุ่ม ความยาวของแบบทดสอบและวิธีการที่ใช้ในการหาค่าความน่าเชื่อถือ

วัตถุประสงค์ของการเรียน

1. หาค่าสหสัมพันธ์และค่าความน่าเชื่อถือได้
2. อธิบายความแตกต่างของค่าสหสัมพันธ์และค่าความน่าเชื่อถือได้
3. อธิบายได้ว่าค่าความน่าเชื่อถือที่สูงหรือต่ำขึ้นอยู่กับอะไรและมีผลต่อการเลือกใช้แบบทดสอบอย่างไร

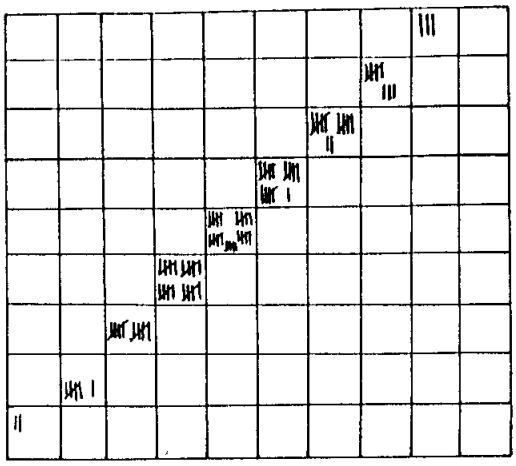
ในบทนี้จะกล่าวถึงเรื่องของความน่าเชื่อถือซึ่งจะเกี่ยวข้องกับวิธีการในการหาค่าสหสัมพันธ์ด้วย นอกจากนี้แล้วยังจะกล่าวถึงการตีความหมายความน่าเชื่อถือตลอดจนเรื่องของค่าของความน่าเชื่อถือว่าจะสูงหรือต่ำเกี่ยวข้องกับอะไรบ้าง และนำไปใช้เพื่อประโยชน์ในการเลือกใช้แบบทดสอบอย่างไร

ความน่าเชื่อถือของแบบทดสอบ หมายถึง ความคงที่ของคะแนนที่แต่ละบุคคลทำได้เมื่อทำการทดสอบใหม่ โดยใช้แบบทดสอบเดิมแต่ต่างเวลากัน หรือโดยแบบทดสอบที่ต่างกันแต่ใกล้เคียงกัน หรือในการทดสอบที่มีตัวแปรอื่น ๆ การทดสอบความน่าเชื่อถือของแบบทดสอบ คือ การคำนวณหาความผิดพลาดในการวัด (error of measurement) คือ การคำนวณหาอัตราของการที่ผลที่จะเกิดขึ้นเป็นผลมาจากการเสี่ยง (chance)

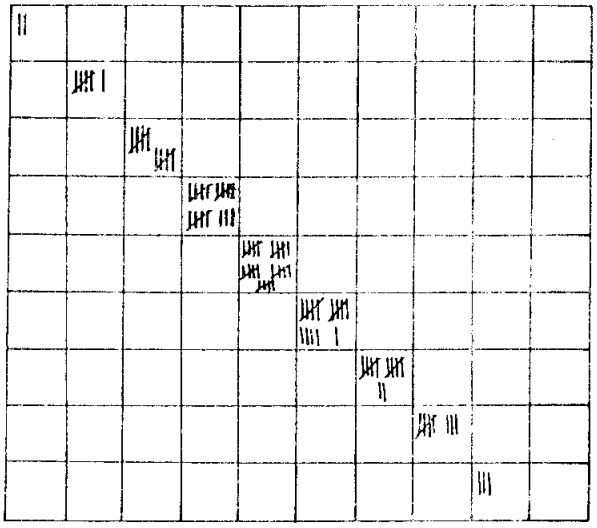
1. ค่าสหสัมพันธ์ (Correlation)

ความหมายของค่าสหสัมพันธ์ (r) อธิบายถึง อัตราของความเกี่ยวข้องหรือความสัมพันธ์ระหว่างคะแนน 2 ชุด ถ้าคนที่ได้คะแนนสูงสุดในคะแนนชุดที่ 1 ก็ได้คะแนนสูงสุดในคะแนนชุดที่ 2 คนที่ได้คะแนนเป็นที่ 2 ในคะแนนชุดที่ 1 ก็ได้คะแนนสูงเป็นที่ 2 ในคะแนนชุดที่ 2 หรือในทางตรงข้าม คนที่ได้คะแนนต่ำสุดในคะแนนชุดที่ 1 ก็ได้คะแนนต่ำสุดในคะแนนชุดที่ 2 แสดงว่าเป็นความสัมพันธ์ที่สมบูรณ์แบบ Positive Perfect Correlation ซึ่งจะมีค่า = +1.00 ซึ่งเราสามารถมองเห็นได้ ในตารางภาพที่ 4.1 ในตารางภาพที่ 4.2 เป็นความสัมพันธ์ที่สมบูรณ์แบบ Negative Perfect Correlation ซึ่งมีค่า = −1.00 ซึ่งในที่นี้คือความสัมพันธ์ระหว่างคะแนน 2 ชุด เป็นไปโดยที่คนที่ได้คะแนนสูงในคะแนนชุดที่ 1 จะได้คะแนนต่ำสุดในคะแนนชุดที่ 2 และเป็นเช่นนี้ไปเรื่อย ๆ

ส่วนการที่ค่าสหสัมพันธ์ = 0 หมายถึง การที่ข้อมูลนั้นไม่มีความสัมพันธ์กันเลย ซึ่งมันจะเกิดขึ้นได้โดยการเสี่ยง เมื่อค่า r = 0 นั่นหมายถึง การที่เราไม่สามารถที่จะทำนายผลของข้อมูลชุดที่ 2 จากชุดที่ 1 ได้



ตารางภาพที่ 4.1 *แสดงลักษณะความสัมพันธ์ที่มีค่าสหสัมพันธ์ = +1.00*



ตารางภาพที่ 4.2 *แสดงลักษณะความสัมพันธ์ที่มีค่าสหสัมพันธ์ = -1.00*

วิธีการในการคำนวณหาค่าสหสัมพันธ์นั้นสามารถที่จะคำนวณหาได้หลายวิธี แต่วิธีที่ง่ายที่สุด และเป็นที่รู้จักที่สุดคือ Pearson Product-Moment Correlation Coefficient ซึ่งวิธีคิดนี้ไม่ใช่แต่จะบอกตำแหน่งของแต่ละบุคคลในกลุ่มเท่านั้น แต่ยังบอกให้ทราบถึงตำแหน่งของแต่ละบุคคลที่อยู่เหนือหรือต่ำกว่ามัชฌิมเลขคณิตของกลุ่มไปเท่าไรด้วย

เราคงจะจำได้ว่า ถ้าบุคคลใดอยู่เหนือกว่าค่าเฉลี่ยโดยการอธิบายในรูปของคะแนนมาตรฐานนั้น ถ้าคนนั้นอยู่เหนือค่าเฉลี่ยของกลุ่ม นั่นคือ ค่าของ S.D. เป็น + และถ้าอยู่ต่ำกว่าค่าเฉลี่ย ค่าของ S.D. จะเป็น – ถ้าคนนั้นได้คะแนนสูงในคะแนน 2 กลุ่ม คือ ได้คะแนนเป็นทั้ง + ทั้ง 2 ชุด และในทำนองเดียวกันถ้าได้คะแนนต่ำก็จะได้ค่าเป็น – ทั้ง 2 ชุด

นักเรียน	X	Y	$x - \bar{x}$ x	$y - \bar{y}$ y	x^2	y^2	xy
	คณิตศาสตร์	การอ่าน					
1	41	17	+1	-4	1	16	-4
2	38	28	-2	+7	4	49	-14
3	48	22	+8	+1	64	1	8
4	32	16	-8	-5	64	25	40
5	34	18	-6	-3	36	9	18

นักเรียน	X	Y	$x-\bar{x}$ x	$y-\bar{y}$ y	x^2	y^2	xy
	คณิตศาสตร์	การอ่าน					
6	36	15	−4	−6	16	36	24
7	41	24	+1	+3	1	9	3
8	43	20	+3	−1	9	1	−3
9	47	23	+7	+2	49	4	14
10	40	27	0	+6	0	36	0
	400	210	0		244	186	86
M($\bar{X}$)	40	21					

ตารางที่ 4.1 *แสดงการคำนวณโดยการใช้ Pearson-Product Moment*

สูตร $$r_{xy} = \frac{\Sigma xy}{N\sigma_x\sigma_y}$$

$$\sigma_x = \sqrt{\frac{\Sigma x^2}{N}} = \sqrt{\frac{244}{10}} = \sqrt{24.40} = 4.94$$

$$\sigma_y = \sqrt{\frac{\Sigma y^2}{N}} = \sqrt{\frac{186}{10}} = \sqrt{18.6} = 4.31$$

$$r_{xy} = \frac{86}{10(4.94)(4.31)} = \frac{86}{212.91}$$

$$= .40$$

นัยสำคัญทางสถิติ ค่าสหสัมพันธ์ = .40 จากตารางที่ 4.1 แสดงถึงอัตราปานกลางของความสัมพันธ์ของคะแนน 2 ชุด ระหว่างคณิตศาสตร์และการอ่าน นั่นคือ มีโอกาสเป็นไปได้ในการที่เด็กทำคณิตศาสตร์ได้ดี ก็จะได้คะแนนดีในการอ่าน และในทางกลับกันถึงแม้ว่าความสัมพันธ์นี้จะไม่ใกล้ชิดกันนัก แต่อย่างไรก็ตาม ถ้าเราสนใจในเด็กทั้ง 10 คนนี้ก็อาจกล่าวได้ว่าความสัมพันธ์ที่เกิดขึ้นยังไม่ถูกต้อง ในงานการวิจัยทางจิตวิทยานั้น เราจะสนใจในการที่จะนำเอาตัวอย่างของแต่ละกลุ่มมาเป็นตัวแทนของประชากรทั้งหมด เช่น ต้องการจะดูความสัมพันธ์ของความสามารถทางคณิตศาสตร์กับการอ่านของเด็กนักเรียนทั่วประเทศในระดับชั้นเรียนหนึ่ง เพราะฉะนั้น การใช้ตัวอย่างเพียง 10 คน จากตัวอย่างจึงไม่เพียงพอ

กิจกรรมการเรียนที่ 1

ให้ค่าสหสัมพันธ์จากข้อมูลต่อไปนี้โดยใช้สูตร Pearson-Product-Moment

คนที่	x	y
1.	5	7
2.	7	8
3.	8	9
4.	10	10
5.	10	11

2. การหาค่าความน่าเชื่อถือ

ในเรื่องของความน่าเชื่อถือ นั้นเป็นการวัดความน่าเชื่อถือแค่ไหน ในการวัดนั้นไม่ได้ถามแค่ว่า วัดอะไรเท่านั้น แต่เราวัดโดยสนใจต่อไปอีกว่า วัดได้แม่นยำแค่ไหน นั่นคือคะแนนที่ได้มาคงเส้นคงวาแค่ไหน ถ้าวัดคน ๆ เดียวกันอีกครั้งจะได้ผลเหมือนเก่าหรือไม่ ถ้าจะวัดน้ำหนักของนักศึกษาแต่ละคนในห้อง เป็นการวัดครั้งที่ 1 ในวันรุ่งขึ้นมาวัดอีก โดยอาจจะใช้เครื่องมือที่ดีกว่าเดิม ผลการวัดทั้ง 2 ครั้งนี้จะไม่เหมือนกัน บางคนอาจจะเหมือนเดิมแต่บางคนอาจจะไม่เหมือน ทั้งนี้อาจเป็นเพราะว่าวันแรกคน ๆ นั้นใส่เสื้อหนากว่าหรือกินข้าวมามากกว่าหรือผู้วัดเองเมื่อเข็มก้ำกึ่งกัน คนหนึ่งอาจจะอ่านว่า 42 อีกคนอาจจะอ่าน 43 ซึ่งทำให้การวัด 2 ครั้งนี้ต่างกัน ข้อผิดพลาดนี้เรียกว่า Chance error ซึ่งหมายถึงว่าความผิดพลาดที่เกิดขึ้นนี้ คราวต่อไปอาจจะไม่เกิดขึ้นก็ได้ เพราะฉะนั้นจึงมีการแสดงให้เห็นถึงความน่าเชื่อถือ หรือพูดง่าย ๆ ว่าความน่าเชื่อถือ ความแม่นยำของการวัดชุดหนึ่งมีวิธีแสดงได้ 2 วิธี

1. วัดคน ๆ เดียวซ้ำ ๆ กันหลาย ๆ หน และดูว่าเปลี่ยนแปลงแค่ไหน แทนที่จะวัดน้ำหนักของคนเดียว 2 ครั้ง เราอาจจะวัด 200 ครั้ง ซึ่งจากการวัดมากครั้งทำให้ได้ frequency Distribution of Scores ที่เป็นน้ำหนักตัวของคน ๆ นั้น จากคนนั้นเราสามารถหาส่วนเบี่ยงเบนมาตรฐาน ซึ่ง S.D. นี้จะบอกให้ทราบถึงการกระจัดกระจายของคะแนนซึ่งในที่นี้จะเรียกว่าความผิดพลาดในการวัด (Standard Error of Measurement) เพราะเป็น S.D. ของความผิดพลาดในการวัด

2. การจะวัดว่าน่าเชื่อถือแค่ไหนก็ต่อเมื่อแต่ละคนอยู่ในตำแหน่งเดิมของการวัดแต่ละครั้ง คนที่หนักที่สุดในการวัดครั้งแรก ควรจะเป็นคนที่หนักที่สุดในการวัดครั้งต่อไป ค่าสหสัมพันธ์จะเป็นตัวเลขทางสถิติที่บอกให้เราทราบถึง ขอบเขตที่การวัด 2 อย่างของแต่ละ-

กลุ่มจะไปด้วยกัน ถ้าค่าที่หามาได้แทนที่จะเป็นการวัด 2 อย่าง แต่เป็นการวัดอย่างเดียวหรือเรื่องเดียวกัน แต่วัด 2 ครั้ง ความสัมพันธ์ที่ได้นี้จะบอกให้ทราบถึงความน่าเชื่อถือ ซึ่งเรียกชื่อใหม่ว่า Reliability Coefficient ถ้าการวัด 2 ครั้งได้ผลใกล้เคียงกันเท่าไร ค่าสหสัมพันธ์ก็ยิ่งสูงขึ้น ข้อสอบก็จะยิ่งน่าเชื่อถือมากขึ้นเท่านั้น การวัดของเราจะน่าเชื่อถือขึ้นกับขอบเขตที่บุคคลแต่ละคนจะมีค่าเดิมในการวัดซ้ำ ๆ ซึ่งแสดงได้ 2 อย่างคือ

1. ค่าของความผิดพลาดในการวัด (Standard Error of Measurement) ต่ำ
2. ค่าของความน่าเชื่อถือ (Reliability Coefficient) สูง

ปัญหาคือ เราจะได้ข้อ 1, 2 ได้อย่างไร ซึ่งจะมีวิธีหาอยู่ 4 วิธี คือ

1. การทดสอบซ้ำ (Retest with the same Test)
2. การทดสอบแบบคู่ขนาน (Parallel-Form)
3. การทดสอบแบบแบ่งครึ่ง (Split-Half Method หรือ Subdivided Test)
4. ใช้วิธีการทางสถิติ (Reliability via Interval Consistency หรือ Items Statistics)

2.1 การทดสอบซ้ำ เราต้องการจะทราบว่าเราจะประมาณค่าน้ำหนักตัวของแต่ละบุคคลได้น่าไว้วางใจแค่ไหน เราสามารถที่จะชั่งน้ำหนักได้สองครั้ง โดยต่างคนต่างวัดเพราะไม่ต้องการให้ผลจากการวัดครั้งแรกไปมีผลต่อการวัดครั้งหลัง การวัดอาจจะเป็นการวัดแบบต่อเนื่องกันไป หรือต่างวันกันไป ขึ้นอยู่กับว่าเราสนใจในเรื่องใด ถ้าเราต้องการทราบว่าเราวัดน้ำหนักคนได้แม่นแค่ไหน การวัด 2 ครั้งก็ควรจะกระทำติด ๆ กันถ้าวัดติด ๆ กัน ค่าที่ต่างกัน หรือความผิดพลาดที่เกิดขึ้นเป็นผลของการวัด แต่ในบางครั้งเราต้องการจะวัดความแตกต่างที่เกิดขึ้นในตัวบุคคลนั้น การที่จะศึกษาถึงคุณสมบัติของความน่าเชื่อถือทางกายภาพของคน เช่น การวัดความเร็วของการตอบสนอง ความแข็งแรงทางกล้ามเนื้อถ้าจะทำซ้ำ ๆ ก็ทำได้ แต่ถ้าจะดูความน่าเชื่อถือของข้อสอบความเข้าใจในการอ่านถ้าข้อสอบนั้นมีข้อความ 6 เรื่อง แต่ละเรื่องมีคำถาม 10 ข้อ ทำสอบแล้วหยุดไป แล้วมาทำข้อสอบนั้นซ้ำอีก จะไม่เหมือนกับการวัดน้ำหนักซ้ำ ๆ เพราะผู้ถูกทดสอบจะไม่อ่านหมดจะอ่านแต่เฉพาะที่ไม่รู้เรื่อง เพราะฉะนั้น คำตอบในครั้งที่สองจะมีผลมาจากครั้งแรก และครั้งแรกจะทำไม่เสร็จ แต่ครั้งหลังจะทำได้เสร็จ เพราะฉะนั้น ในการทำการทดสอบซ้ำนั้นเราจะเพ่งเล็งในงานที่จะต้องทำหรือ ในข้อสอบข้อความที่ผู้ถูกทดสอบชอบเป็นพิเศษ จะทำได้คะแนนดีเป็นพิเศษ เช่น ข้อสอบเป็นเรื่องใดเรื่องหนึ่งในวิชาจิตวิทยานักศึกษาจิตวิทยากับนักศึกษาภาษา นักศึกษาจิตวิทยาจะทำได้คะแนนดีกว่า และในขณะเดียวกัน ข้อสอบการอ่าน นักศึกษาภาษาจะทำได้ดีกว่า เพราะฉะนั้น ข้อสอบนี้ก็จะทำได้มากกว่าที่ควรจะเป็น (over estimate) ถ้าทำครั้งที่ 2

ก็จะทำได้มากกว่าที่ควรจะเป็นอีก ความผิดพลาดจากครั้งแรกก็จะติดไปเป็นข้อผิดพลาดหนที่ 2 ด้วย เพราะฉะนั้น ความผิดพลาดก็จะเป็นความผิดพลาดที่คงที่ (Constance error) ซึ่งจะทำให้ดูเหมือนกับว่าข้อสอบนั้นน่าเชื่อถือ เพราะฉะนั้นในลักษณะงานแต่ละชนิดที่จะให้ทำจะต้องคำนึงถึงเรื่องประสบการณ์เฉพาะเรื่อง หรือพื้นฐาน ของแต่ละคน เพราะจะมีผลทำให้ผู้ถูกทดสอบทำได้ดีในเรื่องหนึ่ง และทำไม่ได้ดีในอีกเรื่องหนึ่ง

ในเรื่องของการทดสอบเพื่อจะดูความน่าเชื่อถือของแบบทดสอบนั้น จะทำให้มีสาเหตุใหญ่ ๆ ที่ทำให้เกิดความแตกต่าง (Variation) ขึ้น 3 อย่างคือ

1. ความแตกต่างในการตอบสนองที่มีต่อแบบทดสอบในขณะนั้นของการทดสอบแต่ละครั้ง

2. ความแตกต่างในตัวบุคคลจากเวลาหนึ่งไปยังอีกเวลาหนึ่ง

3. ความแตกต่างเนื่องมาจากงานที่เราสุ่มมาในแต่ละครั้ง

ในการทดสอบซ้ำจะมีความแตกต่างในข้อ 1, 2 แต่ในข้อ 3 จะไม่เกิดขึ้นแต่มีผลของความจำและการฝึกฝนเข้ามาเกี่ยวข้อง

2.2 การทดสอบแบบคู่ขนานในการทดสอบที่จะต้องเลือกงานมา ซึ่งการเลือกงานจะเป็นต้นเหตุของความผิดพลาดจากความแตกต่างของข้อที่ 3 ซึ่งเราอาจที่จะขจัดโดยการหาสหสัมพันธ์ของแบบทดสอบคือ แบบทดสอบหลายแบบที่สร้างขึ้นเพื่อวัดรายละเอียดอันเดียวกัน แต่เป็นตัวอย่างของพฤติกรรมหลายอัน เช่น แบบทดสอบเกี่ยวกับการอ่านความสมดุลของแบบทดสอบนี้ ควรจะเป็นแบบทดสอบที่มีข้อความที่จะอ่าน (Reading passage) และมีคำถามในระดับความยากง่ายเท่ากัน และคำถามแบบเดียวกัน ในเนื้อหาทั่ว ๆ ไปคล้ายคลึงกัน การหาความน่าเชื่อถือ คือ การให้ผู้ถูกทดสอบแต่ละคนทำแบบทดสอบชุดหนึ่งก่อน แล้วทำอีกชุดทีหลัง ซึ่งเราอาจทำโดยติดต่อกันไปเลย ถ้าเราไม่สนใจว่าแต่ละคนจะมีการเปลี่ยนแปลงอะไรบ้าง ในช่วงระยะเวลาที่เราทิ้งไว้ ถ้าเราจะสนใจว่าในช่วงเวลาที่ทิ้งไว้จะมีการเปลี่ยนแปลง ก็ทิ้งไว้ก่อน เพราะฉะนั้น แต่ละคนจะมีคะแนนสองชุด และเอาคะแนนมาหาสหสัมพันธ์ก็จะบอกให้ทราบถึงค่าของความน่าเชื่อถือ (Reliability Coefficient) แต่ถ้าเว้นช่วงเอาไว้ก็จะมีความแตกต่างของทั้งสามแบบเข้ามาเกี่ยวข้องถ้าทำแบบทดสอบแบบนี้ แล้วยังได้ค่าที่คงเส้นคงวาแสดงว่าเป็นแบบที่ดีที่สุด เพราะฉะนั้นจะเห็นว่าการใช้แบบทดสอบคู่ขนาน จะเป็นมาตรฐานที่ดีที่สุดที่จะใช้ประเมินผลแบบทดสอบ แต่ในบางแบบทดสอบ ไม่สามารถจะทำคู่ขนานได้เพราะไม่ใช่ว่าตัวแบบทดสอบต้องมีหลายชุดเท่านั้น เราจะต้องมีเวลา กำลังเงินที่จะใช้แบบทดสอบทั้งสองแบบในช่วงเวลาที่ต่างกัน ซึ่งการทำเช่นนี้ ยังเป็นการน่าเบื่อและ

สิ้นเปลือง เพราะฉะนั้นคนทำแบบทดสอบมักจะใช้วิธีประเมินแบบทดสอบ จากการทำแบบทดสอบเพียงครั้งเดียว แต่วิธีนี้ในช่วงเวลาที่ห่างกันไม่มากอาจเป็นวันหรือสัปดาห์ (ไม่ห่างจนเกิดการเรียนรู้) เป็นวิธีที่ดีที่สุดในการที่จะบอกให้ทราบถึงค่าของความน่าเชื่อถือ

2.3 การแบ่งครึ่งแบบทดสอบใช้แพร่หลายในการประเมินค่าความน่าเชื่อถือ ในการทดสอบครั้งเดียว เราจะแบ่งแบบทดสอบออกเป็น 2 ส่วนที่เทียบกันได้ว่าเท่ากัน และการแบ่งนี้จะพิจารณาโดยละเอียดถึงเนื้อหา ความยากง่ายของแต่ละข้อ ให้ 2 ส่วนนี้มีความสมดุลกัน วิธีที่ใช้คือ แบ่งโดยสลับเป็นข้อคู่ อย่างละครึ่ง การที่ทำเช่นนี้เพราะแต่ละข้อส่วนมากจะมีเนื้อหา หรือความยากง่ายอยู่รวมกัน โดยเฉพาะแบบทดสอบที่ยาวประมาณ 60 ข้อ ถ้าแบ่งแบบนี้มักจะให้ความสมดุลดี ถ้าแบ่งแบบทดสอบแบบนี้ก็จะมีความสมดุลกัน การจะแบ่งแบบทดสอบเป็น 2 ส่วนนี้ แบ่งเพียงเพื่อจะคิดคะแนนเท่านั้น ตัวผู้ทำแบบทดสอบจะไม่ทราบเลยว่า แบบทดสอบถูกแบ่ง เพราะผู้ทดสอบจะทดสอบเหมือนไม่ได้แบ่งผู้ถูกทดสอบก็ทำเพียงครั้งเดียวและเวลาเดียวกัน แต่เอาคะแนนมาแปรเป็น 2 พวก พวกหนึ่งคะแนนจากข้อเลขคู่ อีกพวกหนึ่งคะแนนจากข้อเลขคี่สหสัมพันธ์ระหว่างคะแนน 2 พวกนี้คือความแม่นที่แบบทดสอบนี้วัดแต่ละบุคคล แต่ค่าสหสัมพันธ์ที่ได้มาเป็นแต่เพียงค่าสหสัมพันธ์ที่ได้จาก $\frac{1}{2}$ ของแบบทดสอบเท่านั้น ค่านี้จะเอามาใช้เป็นค่าของแบบทดสอบทั้งหมดไม่ได้ เพราะฉะนั้น เราจะต้องทำให้ค่าสหสัมพันธ์ระหว่าง $\frac{1}{2}$ ของแบบทดสอบให้เป็นแบบทดสอบที่มีค่าของความน่าเชื่อถือของแบบทดสอบทั้งหมด โดยใช้สูตร

$$r_{11} = \frac{2\, r_{\frac{1}{2}\,\frac{1}{2}}}{1 + r_{\frac{1}{2}\,\frac{1}{2}}}$$

r_{11} = ค่าความน่าเชื่อถือของแบบทดสอบทั้งหมด

$r_{\frac{11}{22}}$ = ค่าสหสัมพันธ์ระหว่างแบบทดสอบที่แบ่งครึ่ง

โดยทั่วไปแล้ว ตัวอย่างของพฤติกรรมใหญ่เท่าไร การวัดก็จะน่าเชื่อถือมากเท่านั้น ถ้าตัวอย่างของพฤติกรรมมากเท่าไร ก็จะหนีพ้นจากผลจากการเสี่ยง (Chance) ได้มากขึ้นเท่านั้น สูตรนี้เรียกว่า Spearman Brown Prophecy Formular ซึ่ง Spearman และ Brown เป็นผู้คิดขึ้น ซึ่งจากสูตรนี้จะเห็นว่า เราประเมินค่าของความน่าเชื่อถือ จากการทำแบบทดสอบเพียงครั้งเดียว ซึ่งเป็นวิธีที่ง่ายและสะดวกมาก แต่ข้อเสีย คือ

1. ค่าที่ได้ไม่เปิดโอกาสให้มีความแตกต่างในการเปลี่ยนแปลงในคนนั้น ในช่วงเวลาหนึ่ง ๆ

2. ค่าที่ได้จะหมดความหมาย ถ้าความแตกต่างที่ได้ เป็นความแตกต่างในเรื่องความเร็ว (speed) การทำผิดหรือทำไม่ได้ อาจไม่เกิดขึ้น เพราะฉะนั้น ยิ่งมีเรื่องของความเร็ว เข้ามาเกี่ยวข้องมากเท่าใด ค่าของความน่าเชื่อถือยิ่งสูงขึ้นเท่านั้น

2.4 วิธีการทางสถิติ-การทำแบบทดสอบเพียงครั้งเดียว โดยไม่แบ่งแบบทดสอบเป็น 2 ส่วนโดยการใช้สูตร Kuder-Richardson Reliability Coefficient ซึ่ง Kuder-Richardson จะให้ค่าความน่าเชื่อถือที่ตีความหมายโดย

1. เราไม่เอาความแตกต่างในตัวบุคคลจากช่วงระยะเวลาหนึ่ง ไปยังอีกช่วงระยะเวลาหนึ่งมาคิด

2. ไม่เหมาะสำหรับแบบทดสอบความเร็ว

โดยที่ค่าความน่าเชื่อถือที่ได้นั้น ได้จากข้อสมมติที่ว่า ข้อสอบแต่ละข้อในข้อสอบมีอะไรที่คล้าย ๆ กัน และในทุก ๆ ข้อ (item) วัดเรื่องราวเรื่องเดียวกัน

Kuder มี 2 สูตร คือ

≠20

$$r_{11} = \left(\frac{n}{n+1}\right)\left(\frac{St^2 - \Sigma pq}{St^2}\right)$$

r_{11} = ค่าความน่าเชื่อถือที่เราจะประมาณ

n = จำนวนของข้อทดสอบในแบบทดสอบ

St = S.D. ของแบบทดสอบ (ในการทดสอบคนกลุ่มหนึ่งทำ มีคนได้คะแนนต่างกันแล้วหาค่าเฉลี่ย)

Σ = ผลบวกของ pq

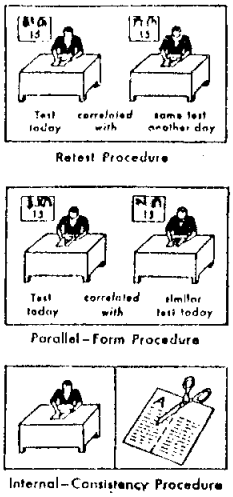
p = สัดส่วน (poportion) ของคนที่ทำข้อหนึ่งได้

q = สัดส่วน (poportion) ของคนที่ทำข้อนั้นไม่ได้

#21

$$r_{11} = \left(\frac{n}{n-1}\right)\left[1 - \frac{Mt(1 - Mt/N)}{St^2}\right]$$

Mt = mean Scores ของกลุ่ม (คะแนนเฉลี่ยของกลุ่ม)



ตารางภาพที่ 4.3 *แสดงการหาค่าความน่าเชื่อถือโดยวิธีต่างกัน*

กิจกรรมการเรียนที่ 2

จับคู่เอาตัวเลขทางขวามือมาใส่ไว้ทางซ้ายมือ

.....	1. การวัดคน ๆ เดียวกันซ้ำแล้วดูว่าคะแนนกระจายไปแค่ไหน	1. Retest
.....	2. การวัดคน ๆ เดียวกันซ้ำแล้วผู้ถูกวัดยังคงอยู่ในตำแหน่งเดิม	2. Split-Half
.....	3. ทำการทดสอบซ้ำเรื่องเดิมใหม่	3. Reliability
.....	4. ทำการทดสอบโดยใช้แบบทดสอบ 2 ชุดที่ใกล้เคียงกัน	4. Error of Measurement
.....	5. ทำแบบทดสอบครั้งเดียวแล้วแบ่งครึ่งแบบทดสอบ	5. Parallel-Form

3. การตีความหมายค่าของความน่าเชื่อถือ

สมมติว่า ผลที่ได้จากการทดสอบเด็กชั้นประถม ในเรื่องความถนัด ได้ค่าความน่าเชื่อถือ .85 เราจะแปลความหมายค่าความน่าเชื่อถือนี้ได้อย่างไร มันเกี่ยวข้องกับคะแนนของแต่ละคนอย่างไร เราจะพอใจหรือไม่กับค่าความน่าเชื่อถือนี้ เราพบว่าการตีความหมายของค่าความน่าเชื่อถือจะเกี่ยวข้องกับเรื่องของความผิดพลาดจากการวัด (Standard Error of Measurement)

คงจะจำได้ว่า ค่าความผิดพลาดจากการวัด คือค่าโดยประมาณของส่วนเบี่ยงเบนมาตรฐาน ซึ่งได้จากการทดสอบคน ๆ เดียวกันซ้ำหลาย ๆ ครั้ง ซึ่งค่าความผิดพลาดในการวัดสามารถที่จะคำนวณหาได้จากสูตร

$$Sm = St\sqrt{1 - r_{11}}$$

เมื่อ Sm = ค่าความผิดพลาดในการวัด

St = ส่วนเบี่ยงเบนมาตรฐานของแบบทดสอบ

r_{11} = ค่าความน่าเชื่อถือของแบบทดสอบ

สมมติว่าค่า r_{11} = .85 และส่วนเบี่ยงเบนมาตรฐาน = 15

ค่า $Sm = 15\sqrt{1 - .85}$

$= 15\sqrt{.15}$

$= 15(.387)$

$= 5.8$

ในตัวอย่างนี้แสดงว่า ในการวัดใครคนใดคนหนึ่ง ค่าของส่วนเบี่ยงเบนมาตรฐานจะ = 5.8 คะแนน ซึ่งถ้าดูความสัมพันธ์กับเรื่องของเส้นโค้งปกติ

โดยนำมาดูกับในกรณีที่ค่าความผิดพลาดในการวัด = 5.8 นั่นคือ การที่ผู้ถูกทดสอบจำนวน 1 ใน 3 หรือ 31.8% ของผู้ถูกทดสอบ ที่มีโอกาสที่จะได้คะแนนผิดไปจากที่เป็นจริงไป 5.8 คะแนน และเพียง 4.6% หรือ 1 ใน 20 ที่จะได้คะแนนผิดไปจากที่เป็นจริงถึง 2 Sm หรือ 11.6 คะแนน (ดูจากตารางที่ 3.2)

ค่า r_{11} ของแบบทดสอบชุดนี้ = .85 และส่วนเบี่ยงเบนมาตรฐานคือ 15 จัดได้ว่า เป็นค่าที่ค่อนข้างสูงสำหรับค่าความน่าเชื่อถือ ถ้าค่าส่วนเบี่ยงเบนมาตรฐานเป็น 15 หรือ 16 คะแนน การที่คะแนนของผู้ถูกทดสอบต่างกัน 5 หรือ 10 คะแนนในแบบทดสอบที่ต่างกัน เราสามารถที่จะคิดได้เลยว่า เป็นความแตกต่างอันเนื่องมาจากความผิดพลาดในการวัด เป็นการไม่ฉลาดเลยถ้าผู้ทดสอบพยายามชี้ให้เห็นความสามารถที่ต่างกันของผู้ถูกทดสอบ 2 คน ที่ทดสอบแบบทดสอบเดียวกัน หรือผู้ถูกทดสอบคนเดียวกันทดสอบแบบทดสอบเดียวกันแล้วได้คะแนนต่างกัน 5 คะแนน ทั้งนี้ เพราะเราจะคำนึงถึงข้อผิดพลาดจากการวัดด้วยนั่นเอง

โดยความจริงแล้ว ความผิดพลาดในการวัด จะสัมพันธ์กับความน่าเชื่อถือ ซึ่งจะแสดงให้เห็นในตารางที่ 4.2 ซึ่งจะเห็นได้ว่า ความผิดพลาดจะยิ่งลดลง เมื่อค่าความน่าเชื่อถือสูงขึ้น แต่ความผิดพลาดในการวัดก็ยังคงมีอยู่ ถึงแม้ว่าค่าความน่าเชื่อถือจะสูงถึง .90 หรือ .95 ก็ตาม เพราะฉะนั้น ในการแปลผลแบบทดสอบ ค่าความผิดพลาดในการวัด เป็นสิ่งที่เราต้องคำนึงอยู่ด้วยเสมอ ในกรณีของแบบทดสอบความถนัดที่กล่าวมานี้ถ้าได้คะแนนจากแบบทด-

สอบ 90 เราค่อนข้างจะแน่ใจได้ว่า คะแนนจะอยู่ระหว่าง 80 และ 100

ตารางที่ 4.2 *แสดงค่าความผิดพลาดในการวัด ในระดับที่ค่าความน่าเชื่อถือต่างกัน*

ค่าความน่าเชื่อถือ	เมื่อ SI = 10
.50	7.1
.60	6.3
.70	5.5
.80	4.5
.85	3.9
.90	3.2
.95	2.2
.98	1.4

นอกจากค่าของความน่าเชื่อถือจะบอกให้ทราบถึงความผิดพลาดในการวัด ซึ่งจะช่วยตีความหมายของคะแนนได้แล้ว ค่าความน่าเชื่อถือ ยังมีไว้สำหรับเปรียบเทียบแบบทดสอบต่าง ๆ ถ้าแบบทดสอบที่มีอะไรเหมือน ๆ กัน หรือเท่ากัน เราก็เลือกแบบทดสอบที่มีความน่าเชื่อถือสูง โดยที่เราคำนึงถึงความเที่ยงตรงและความเหมาะสมด้วย ความน่าเชื่อถือสำคัญในฐานะที่เป็นภาวะที่จำเป็นในการวัด เพื่อให้มีความเที่ยงตรง ถ้าสิ่งที่เราจะวัดคงที่ แต่ถ้าวัดได้แล้วไม่คงที่ เราก็ไม่ได้วัดในสิ่งที่เราต้องการจะวัด เพราะฉะนั้นถ้าแบบทดสอบนั้นไม่มีความน่าเชื่อถือ ก็ไม่มีความเที่ยงตรง ความน่าเชื่อถือเป็นคุณสมบัติขั้นต้นที่จะต้องมี แต่แบบทดสอบที่มีความเชื่อถือสูงอาจเป็นแบบทดสอบที่ไม่เที่ยงตรงเลยก็ได้

กิจกรรมการเรียนที่ 3

ถูก-ผิด ถ้าข้อใดถูกให้เขียนเครื่องหมาย ✓ และเขียน × เมื่อข้อความนั้นผิด

..... 1. ความน่าเชื่อถือไม่เกี่ยวข้องกับส่วนเบี่ยงเบนมาตรฐานเลย

..... 2. ผลการทดสอบมีผลมาจากความผิดพลาดในการวัดด้วย

..... 3. ความน่าเชื่อถือจะเกี่ยวข้องกับความผิดพลาดในการวัด

..... 4. ถ้าค่าความน่าเชื่อถือสูงมาก ๆ จะไม่มีความผิดพลาดในการวัดเลย

..... 5. เมื่อทำการทดสอบอย่างดีและรอบคอบจะไม่พบความผิดพลาดในการวัดเลย

4. ค่าของความน่าเชื่อถือ จะสูงหรือต่ำขึ้นอยู่กับ

4.1 พิสัยของกลุ่ม (Range of the Group) ค่าความน่าเชื่อถือจะบอกให้ทราบว่าแต่ละคนอยู่ลำดับที่ต่างกันได้คงสม่ำเสมอขนาดไหน ถ้ามีการขยับที่กันเล็กน้อย เมื่อทำแบบทดสอบใหม่ ถ้าไม่ขยับเลยแบบทดสอบนั้นก็มีความน่าเชื่อถืออย่างสมบูรณ์ แต่ถ้าขยับนิดหน่อยก็ถือความน่าเชื่อถือสูง ถ้าขยับมากก็มีความน่าเชื่อถือต่ำ การขยับมากหรือน้อย แค่ไหนขึ้นอยู่กับเดิมที่ของแต่ละคนนั้นอยู่ใกล้แค่ไหน

ถ้าค่าความน่าเชื่อถือเป็นของเด็ก ป.2 - ป.7 จะเห็นได้ว่าเราไม่ต้องใช้การวัดละเอียดมากนักก็จะเห็นความแตกต่าง แต่ถ้าจะวัดความสามารถในการอ่านของเด็ก ป.2 เท่านั้น ก็จำเป็นต้องใช้แบบทดสอบที่วัดได้ละเอียดกว่าแบบทดสอบแรก คนที่มีความสามารถใกล้กันเราก็จัดลำดับที่ได้ยากหรือวัดได้แม่นยำน้อยลง นั่นคือ ความสามารถของเด็กในหมู่ต่างกันมาก ถ้าไปหาความสัมพันธ์จะได้ความน่าเชื่อถือที่แลดูสูง แต่แบบทดสอบอันเดียวกัน เอาไปวัดในกลุ่มที่มีความสามารถใกล้เคียงกันเห็นว่าค่าความน่าเชื่อถือจะต่ำกว่าในตอนแรก เพราะฉะนั้นค่าของความน่าเชื่อจะถือต้องดูความสามารถของกลุ่มด้วย

4.2 ระดับความสามารถภายในกลุ่ม (Level of Ability in the Group) ระดับความสามารถของบุคคลที่ถูกวัดมีส่วนที่เกี่ยวข้องกับยุทธศาสตร์ที่ถูกวัดด้วย ถ้าแบบทดสอบยากไปหรือง่ายไป ทำให้ความแม่นผิดไปด้วย แต่บอกไม่ได้ว่า ยากแค่ไหน ถ้าคนทำแบบทดสอบไม่ได้ คือยากเกินไป ส่วนใหญ่ของคำตอบจะเกิดจากการเดา เมื่อมีการเดาก็จะมีโอกาสผิดบ้างถูกบ้าง เพราะฉะนั้น คะแนนที่แต่ละคนได้ แทนที่จะเป็นความสามารถของแต่ละคนจะกลายเป็นความสามารถบวกโอกาส (Chance) จะเห็นได้ว่ายิ่งแบบทดสอบยาก ความแม่นยำจะยิ่งต่ำลง ถ้าแบบทดสอบง่ายไปทุกคนก็ทำได้เกือบหมดทุกข้อ ถ้าแบบสอบมี 100 ข้อ คนที่ทำได้น้อยที่สุด 80 ข้อ จะแยกความสามารถของกลุ่มได้ยากขึ้น เพราะมีคนได้คะแนนเท่ากันมากเมื่อทุกคนทำข้อง่าย ๆ ได้ ข้อง่ายก็ไม่แบ่งแยกความสามารถของแต่ละคนได้ เพราะฉะนั้นนักสร้างแบบทดสอบจะต้องรายงานคือ Standard error ในระดับของคะแนนต่างกันด้วยเพราะคะแนนที่ได้ ไม่ใช่เป็นคะแนนที่ได้จริง ๆ อาจเป็นเพียง 95% เท่านั้นที่คะแนนเราจะต่างกัน เช่น ในแบบทดสอบของบินเน่ ได้ I.Q. 130 จะได้ Standard Error เป็น 5.2 จะเห็นได้ว่าบินเน่ จะวัดพวกที่มี I.Q. สูงได้แม่นยำน้อยกว่าพวกมี I.Q. ต่ำ ถ้า Standard Error ต่ำแบบทดสอบนั้นก็วัดได้แม่นยำ ส่วนของเวกสเล่อร์ดูอายุเป็นปี เด็ก 7 1/2 ปี Standard Error = 4.2 10 1/2 ปี Standard Error = 3.4 13 1/2 ปี Standard Error = 3.7

4.3 ความยาวของแบบทดสอบ ตอนที่แบ่งออกเป็นข้อคู่, คี่ จะเห็นว่าค่าความน่าเชื่อถือจะขึ้นอยู่กับความยาวของแบบทดสอบ นักสถิติจิตวิทยาสร้างสูตรว่า ถ้าคุณสมบัติของข้อทดสอบและธรรมชาติของผู้ถูกทดสอบอันเดิม ความยาวของแบบทดสอบจะเป็นตามสูตร

$$r_{nn} = \frac{Nr11}{1+(n-1)r11}$$

เมื่อ r_{nn} เป็นความน่าเชื่อถือของแบบทดสอบเมื่อยาวเป็น n เท่าของแบบทดสอบเดิม

r_{11} เป็นความน่าเชื่อถือของแบบทดสอบเดิม

n เป็นจำนวนเท่าของความยาวที่แบบทดสอบอันใหม่มีต่อแบบทดสอบเก่า ถ้าแบบทดสอบเดิม ค่าความน่าเชื่อถือ = 0.5 แบบทดสอบใหม่ยาวขึ้นเป็น 5 เท่า

$$r = \frac{5\times.5}{1+(5-1).5} = \frac{2.5}{1+2} = .83$$

แต่สูตรนี้จะใช้ได้เมื่อคุณสมบัติของแต่ละข้อของแบบทดสอบคงเดิมและธรรมชาติของผู้ถูกทดสอบคงเดิม

เราจะพบว่า ยิ่งแบบทดสอบยาวเท่าใด ความน่าเชื่อถือยิ่งสูงขึ้น เพราะความผิดพลาดที่เกิดจากโชคหรือความบังเอิญก็ยิ่งยากขึ้น คะแนนที่ได้จะไว้ใจได้มากขึ้น แต่อย่างไรก็ตาม ในทางปฏิบัติ เราต้องคำนึงถึง

1. เวลาที่จะพอมีให้ได้ในการที่จะทำแบบทดสอบ
2. ความเบื่อหน่ายของผู้ทำ
3. การสร้างข้อสอบดี ๆ ไม่ง่าย

4.4 วิธีการที่ใช้ในการหาค่าความน่าเชื่อถือ วิธีที่ได้ค่าความน่าเชื่อถือ เพราะวิธีต่างกัน ค่าความน่าเชื่อถือที่ได้มาก็ไม่เท่ากัน เพราะโอกาสของความแตกต่าง (variation) จะเกิดขึ้น แต่ละวิธีจะต่างกัน

ตารางที่ 4.3 *แสดงที่มาของความแตกต่างที่จะเกิดขึ้นเมื่อใช้วิธีการในการหาค่าความน่าเชื่อถือต่างกัน*

ที่มาของความแตกต่าง	วิธีการที่ใช้ในการหาค่าความน่าเชื่อถือ				
	ทดสอบซ้ำในทันทีทันใด	ทดสอบซ้ำในเวลาต่อมา	ทดสอบแบบคู่ขนานต่อกัน	แบ่งครึ่งแบบทดสอบ	ใช้วิธีการทางสถิติ
ค่าที่ได้จะต่างกันอันเป็นผลมาจาก					
1. ความแตกต่างที่เกิดขึ้นจากวิธีการที่ใช้ในการวัด	x	x	x	x	x
2. การเปลี่ยนแปลงในตัวของผู้ถูกทดสอบ		x		x	
3. การเปลี่ยนแปลงที่เกิดจากลักษณะงานที่ต่างกัน		x	x	x	x
4. การเปลี่ยนแปลงที่เกิดขึ้นจากความเชื่อในการทดสอบของผู้ถูกทดสอบ	x	x	x	x	

กิจกรรมการเรียนที่ 4

ถูก-ผิด เขียนเครื่องหมาย ✓ ลงหน้าข้อความที่ถูกและเขียน × หน้าข้อความที่ผิด

..... 1. ยิ่งแบบทดสอบง่ายค่าความน่าเชื่อถือจะสูง

..... 2. ความสามารถของคนในกลุ่มต่างกันมากค่าความน่าเชื่อถือจะต่ำ

..... 3. แบบทดสอบยิ่งยาวก็ยิ่งทำให้ค่าความน่าเชื่อถือต่ำ

..... 4. วิธีการที่ใช้ในการหาค่าความน่าเชื่อถือทำให้ค่าที่ได้ต่างกันด้วย

..... 5. แบบทดสอบสั้น ๆ ทำให้สร้างแบบทดสอบได้ง่ายและไม่น่าเบื่อสำหรับผู้ทำ

สรุป

ความน่าเชื่อถือของแบบทดสอบคือความคงที่ของคะแนนที่แต่ละบุคคลทำได้เมื่อทำการทดสอบใหม่โดยใช้แบบทดสอบเดิมโดยต่างเวลากันหรือโดยแบบทดสอบที่ต่างกันแต่ใกล้เคียงกัน การแสดงความน่าเชื่อถือของแบบทดสอบได้แก่ การวัดคน ๆ เดียวกันซ้ำแล้วได้ค่าความผิดพลาดในการวัดต่ำ และค่าความน่าเชื่อถือสูง

การหาความน่าเชื่อถือทำได้ 4 วิธีคือ

1. การทดสอบซ้ำ (Retest with the Same Test)
2. การทดสอบแบบคู่ขนาน (Parallel Form)
3. การทดสอบแบบแบ่งครึ่ง (Split-Half Method)
4. วิธีการทางสถิติ (Item Statistics)

ในการทดสอบเพื่อหาค่าความน่าเชื่อถือนั้น สิ่งที่ควรระวังคือการเกิดความผิดพลาดคงที่ หรือ Constance Error

ค่าความน่าเชื่อถือจะสูงหรือต่ำขึ้นอยู่กับ

1. พิสัยของกลุ่ม
2. ระดับความสามารถภายในกลุ่ม
3. ความยาวของแบบทดสอบ
4. วิธีการที่ใช้ในการหาค่าความน่าเชื่อถือ

แบบฝึกหัด

ตอนที่ 1 *ถูก-ผิด ให้นักศึกษาเขียนเครื่องหมาย ✓ ลงหน้าข้อความที่ถูก และเขียน × หน้าข้อความที่ผิด*

..... 1. ค่าสหสัมพันธ์จะต้องไม่เกิน ± 1

..... 2. ค่าความน่าเชื่อถือเป็นได้ทั้ง + และ –

..... 3. เด็กสามารถพูดได้เก่งขึ้นตามการเพิ่มขึ้นของอายุเป็นความสัมพันธ์ทางบวก

..... 4. ราคาที่ดินเพิ่มขึ้นตามการเพิ่มขึ้นของจำนวนปี เป็นความสัมพันธ์ทางลบ

..... 5. ก. สอบได้เป็นที่ 1 ในวิชา A และได้เป็นที่สุดท้ายในวิชา B แสดงว่าวิชา A และ B ไม่มีความสัมพันธ์กันเลย

..... 6. ค่าความน่าเชื่อถือมากขึ้น หรือลดลงขึ้นอยู่กับค่าความผิดพลาดในการวัดด้วย

..... 7. ถ้าคนในกลุ่มมีความสามารถต่างกันมาก จะยังทำให้ค่าความน่าเชื่อถือสูงยิ่งขึ้น

..... 8. เราสามารถจะวิเคราะห์ค่าความน่าเชื่อถือด้วยวิธีการทางสถิติเท่านั้น

..... 9. ลักษณะของ "งาน" ที่ให้ทำในแบบทดสอบ ไม่มีผลต่อค่าความน่าเชื่อถือ

..... 10. ความน่าเชื่อถือ จะเกี่ยวข้องกับตำแหน่งที่คงที่ในการทดสอบด้วย

..... 11. การวัดเรื่องเดียวกัน คนเดียวกัน เป็นเรื่องของความเที่ยงตรง

..... 12. ความสัมพันธ์ของการวัดเรื่องเดียวกัน คนเดียวกัน เรียกว่า สหสัมพันธ์

..... 13. ความผิดพลาดคงที่ (Constance Error) มีผลทำให้ค่าความน่าเชื่อถือสูง

..... 14. Chance Error จะไม่เกิดขึ้นในการทดสอบทุกประเภท

..... 15. ประสบการณ์ และพื้นฐานของผู้ถูกทดสอบ มีผลต่อค่าความน่าเชื่อถือ

..... 16. ความแตกต่างที่เกิดขึ้นในตัวบุคคลจากเวลาหนึ่ง ไปยังอีกเวลาหนึ่ง อาจจะเกิดขึ้นได้ในการทดสอบแบบทดสอบซ้ำ (Retest with the same Test)

..... 17. ความแตกต่างในการตอบสนองที่มีต่อแบบทดสอบในขณะนั้น จะเกิดขึ้นในการทดสอบทุกชนิด

..... 18. ความจำ และการฝึกฝน จะมีผลต่อการทดสอบแบบคู่ขนาน (Parallel Form)

..... 19. การแบ่งครึ่งแบบทดสอบ (Split-Half Method) ไม่เป็นที่นิยมเพราะต้องทำแบบทดสอบยาวขึ้น

..... 20. ยิ่งข้อทดสอบมากเท่าไหร่ ยิ่งทำให้ได้ผลถูกต้องมากขึ้นเท่านั้น

..... 21. การทำผิดหรือทำไม่ได้ของผู้ถูกทดสอบ ไม่มีผลเกี่ยวกับการทดสอบแบบใช้ความเร็ว (Speed-Test)

..... 22. เราสามารถหาค่าความน่าเชื่อถือของแบบทดสอบ โดยใช้สูตร Kuder Richardson ได้

..... 23. ความสามารถที่ใกล้เคียงกันของผู้ถูกทดสอบมีผลทำให้ค่าความน่าเชื่อถือสูงได้

..... 24. ความยาก ง่าย ของแบบทดสอบ ไม่มีผลทำให้ค่าความน่าเชื่อถือ สูง ต่ำ ได้

ตอนที่ 2 *จับคู่ ให้นักศึกษานำเลขหน้าข้อความทางขวามือมาใส่ไว้หน้าข้อที่เกี่ยวข้องสัมพันธ์กัน*

ชุดที่ 1

.....	1. การทดสอบคน ๆ เดียวกันซ้ำแล้วดูว่าคะแนนคงที่แค่ไหน	1. Positive Correlation
.....	2. ความผิดพลาดที่เกิดขึ้นแล้วคราวหน้าจะไม่เกิดขึ้นอีก	2. Negative Correlation
.....	3. ความผิดพลาดจากการทดสอบครั้งแรกตัดไปเป็นความผิดพลาดครั้งที่ 2	3. Reliability
.....	4. ราคารถลดลงตามจำนวนปีที่เพิ่มขึ้น	4. Constance Error
.....	5. ยิ่งอายุมากขึ้นยิ่งสวยขึ้น	5. Chance Error

ชุดที่ 2 *ข้อ 6–10*

.....	6. ความแตกต่างที่อาจเกิดขึ้นเนื่องมาจากงานที่เราสุ่มมาให้แต่ละครั้ง	1. Split-Half
.....	7. ไม่นำมาใช้กับแบบทดสอบที่เกี่ยวข้องกับเรื่องความเร็ว	2. Item-Statistics
.....	8. ไม่มีปัญหาเรื่องความแตกต่างในช่วงเวลาหนึ่งไปยังอีกเวลาหนึ่ง	3. Parallel-Form
.....	9. มีผลทำให้ค่าความน่าเชื่อถือของแบบทดสอบสูงหรือต่ำได้	4. Retest
.....	10. ลักษณะงานจะไม่มีผลทำให้เกิดความแตกต่างได้	5. Range of the Group

ตอนที่ 3 *ให้ตอบคำถามต่อไปนี้*

1. จงกล่าวถึงวิธีการในการหาค่าความน่าเชื่อถือมาตามที่ท่านทราบ

2. ในการทดสอบใหม่แต่ละครั้งจะมีความแตกต่าง (Variation) อะไรเกิดขึ้นได้บ้าง

3. พิสัยของกลุ่มหมายถึงอะไรและเกี่ยวข้องกับความน่าเชื่อถืออย่างไร

4. ความผิดพลาดในการวัดเกี่ยวข้องกับส่วนเบี่ยงเบนมาตรฐานหรือไม่อย่างไร