

# บทที่ 7

## ความเที่ยงตรงและความเชื่อมั่น

### I จุดประสงค์การเรียนรู้

หลังจากจบบทเรียนนี้แล้วผู้เรียนสามารถปฏิบัติได้ดังนี้

1. อธิบายความหมายของความเที่ยงตรงและความเชื่อมั่น
2. อธิบายความเที่ยงตรงตามเนื้อหา ความเที่ยงตรงตามเกณฑ์และความเที่ยงตรงตามโครงสร้าง
3. วิเคราะห์องค์ประกอบที่มีอิทธิพลต่อค่าความเที่ยงตรง
4. แสดงการหาค่าความเชื่อมั่นแต่ละวิธี
5. บอกแหล่งความคลาดเคลื่อนในการหาค่าความเชื่อมั่น
6. วิเคราะห์องค์ประกอบที่มีอิทธิพลต่อค่าความเชื่อมั่น

### II เนื้อหาสาระสำคัญ

#### 1. ความเที่ยงตรง (Validity)

หมายถึงคุณลักษณะของเครื่องมือวัดที่จะสามารถวัดได้ในสิ่งที่ต้องการวัด ดังนั้น เครื่องมือวัดอย่างหนึ่งอาจมีความเที่ยงตรงสำหรับจุดมุ่งหมายอย่างหนึ่ง และอาจมีความเที่ยงตรง ปานกลางหรือน้อยมากสำหรับการวัดในอีกจุดมุ่งหมายหนึ่ง เช่น แบบทดสอบที่วัดความสามารถทางช่างกล อาจมีค่าความเที่ยงตรงสูงเมื่อใช้ในการพยากรณ์เกรดของนักเรียนช่างกล หรือใช้คัดเลือกผู้จะเข้ามาเรียนทางช่างกล ในขณะเดียวกันก็อาจมีความเที่ยงตรงต่ำถ้าจะนำมาใช้คัดเลือกผู้หญิงที่จะเข้าทำงานในโรงงานอุตสาหกรรม เป็นต้น

จะเห็นได้ว่าสิ่งที่เป็นตัวแปรสำคัญที่จะกำหนดว่าเครื่องมือวัดนั้น ๆ มีค่าความเที่ยงตรงอยู่ในระดับใดก็คือสิ่งที่จะวัด หรือสิ่งที่เป็นตัวแปรเกณฑ์ (Criteria variable) การกำหนด

ขอบเขต หรือให้คำนิยามเชิงปฏิบัติการ (Operational Definition) ของสิ่งที่จะนำมาใช้เป็นตัวเกณฑ์ เมื่อเครื่องมือวัดไม่เที่ยงตรงตามเกณฑ์ที่จะใช้วัดแล้ว ข้อมูลที่ได้รับมา ก็จะไม่สามารถนำมาใช้ตัดสินใจอย่างโดยย่างหนึ่งได้

ในการหาค่าความเที่ยงตรงนี้ ก็ยังไม่มีกฎเกณฑ์ใดจะมากำหนดให้ว่าเครื่องมือชนิดใด มีความเที่ยงตรงสูงหรือต่ำ การจะพิจารณาตัดสินใจเกี่ยวกับการหาค่าความเที่ยงตรงนั้น จำเป็นจะต้องมาจาก การฝึกฝน หรือจากประสบการณ์ของผู้ที่คุ้นเคยกับเครื่องมือวัดชนิดนั้น ๆ ดังนั้น การพิจารณาค่าความเที่ยงตรงนั้นนอกจากจะใช้วิธีการคำนวณทางสถิติโดยการหาค่า สหสมพันธ์ระหว่างข้อมูลที่ได้จากเครื่องมือวัดนั้น ๆ กับข้อมูลที่ได้จากการตัวแปรเกณฑ์แล้ว ยัง อาจพิจารณาจากการใช้ภารณฑ์ทางตรรกวิทยา ก็ได้ ซึ่งการใช้วิธีการทางตรรกวิทยา มีความจำเป็นอย่างยิ่งที่จะต้องอาศัยประสบการณ์ และความแม่นยำในสิ่งที่ต้องการจะวัดของ ผู้ประเมิน

ความเที่ยงตรงมีหลายประเภท สมาคมจิตวิทยาอเมริกัน (American Psychological Association) ร่วมกับสมาคมวิจัยการศึกษาอเมริกัน (American Educational Research Association) และสภากារวัดผลการศึกษาแห่งชาติ (National Council of Measurement in Education) ได้ ร่วมกันแบ่งความเที่ยงตรงออกเป็น 3 ประเภทใหญ่ ๆ คือ

1. ความเที่ยงตรงตามเนื้อหา (Content Validity)
2. ความเที่ยงตรงตามเกณฑ์ (Criterion Related Validity)
3. ความเที่ยงตรงตามโครงสร้าง (Construct Validity)

### 1.1 ความเที่ยงตรงตามเนื้อหา (Content Validity)

ความเที่ยงตรงชนิดนี้ หมายถึง ความสอดคล้องของเนื้อหาในแบบทดสอบกับเนื้อหาที่ ต้องการจะวัด มีบางแห่งอาจเรียกว่าความเที่ยงตรงชนิดนี้ว่า logical validity, course validity, curricular validity, text book validity เป็นต้น ความเที่ยงตรงตามเนื้อหานี้มีความสำคัญยิ่งต่อ เครื่องมือวัดผลสัมฤทธิ์ทางการเรียน เพราะสามารถจะแสดงถึงคุณลักษณะของเครื่องมือวัด นั้น ๆ ว่าวัดได้ตรงตามเนื้อเรื่อง ตรงตามจุดประสงค์เชิงพฤติกรรมที่ควรกำหนดไว้หรือไม่ ควร ได้ครอบคลุมเนื้อเรื่องทั้งหมดหรือไม่

การพิจารณาหาค่าความเที่ยงตรงตามเนื้อหานี้ไม่จำเป็นจะต้องใช้วิธีการทางสถิติมา ช่วย แต่มักนิยมใช้ผู้เชี่ยวชาญในสาขาวิชานั้น ๆ หลาย ๆ คนเป็นผู้ร่วมพิจารณา โดยทั่วไป มักพิจารณาดังนี้

1. ดูว่าเนื้อหาในแบบทดสอบกับเนื้อหาที่ต้องการจะวัดสอดคล้องกันหรือไม่

2. ดูว่าแบบทดสอบทั้งฉบับวัดได้ครอบคลุมเนื้อหาที่กำหนดขอบเขตว่าจะวัดได้ครบถ้วนหรือไม่
3. จำนวนข้อที่กำหนดไว้ในตารางวิเคราะห์หลักสูตรทดสอบคล้องและครบถ้วนหรือไม่
4. ลักษณะของเครื่องมือวัดตรงตามจุดมุ่งหมายเชิงพฤติกรรมที่ได้ระบุไว้หรือไม่
5. เรื่องราวต่าง ๆ ที่นำมาอ้างในแบบทดสอบหรือเครื่องมือวัดยังทันต่อเหตุการณ์ปัจจุบันหรือล้าสมัยไปแล้ว

## 1.2 ความเที่ยงตรงตามเกณฑ์ (Criterion Related Validity)

เป็นความเที่ยงตรงที่มีความสัมพันธ์กับเกณฑ์ เป็นกลวิธีในการศึกษาความสัมพันธ์ของข้อมูลในเครื่องมือวัดกับสภาพความเป็นจริงซึ่งได้จากการปฏิบัติ บางครั้งอาจเรียกว่า Empirical Validity ความเที่ยงตรงชนิดนี้นับว่าเป็นความเที่ยงตรงที่จะต้องอาศัยการคำนวณทางสถิติมาช่วยในการหาค่าความเที่ยงตรงมากที่สุด ซึ่งเป็นการหาค่าสัมพันธ์ระหว่างคะแนนที่ได้จากเครื่องมือวัดกับคะแนนที่ได้จากเกณฑ์ทั้งในสภาพปัจจุบันและใช้ทำนายผลในอนาคต จะเห็นได้ว่าค่าความเที่ยงตรงชนิดนี้จะเปลี่ยนแปลงไปตามเกณฑ์ที่จะนำมาหาค่าความสัมพันธ์กับความเที่ยงตรงชนิดนี้จึงแบ่งเป็น 2 ประเภทด้วยกัน คือ

1.2.1 ความเที่ยงตรงตามสภาพ (Concurrent Validity) เป็นค่าความเที่ยงตรงที่ได้จากการทดสอบคล้องของค่าที่วัดได้จากเครื่องมือวัดกับสภาพความเป็นจริงของเด็กในปัจจุบัน โดยนำเอาสภาพความเป็นจริงของเด็กในปัจจุบันเป็นตัวเกณฑ์ เป็นเครื่องแสดงให้เห็นว่าเมื่อครูกำหนดจุดมุ่งหมายการเรียนการสอนไปแล้ว การจะตรวจสอบว่าในสภาพที่เป็นอยู่ปัจจุบันเด็กบรรลุเป้าหมายเพียงใด เช่น ครูกำหนดจุดมุ่งหมายให้นักเรียนรู้จักแปรรูปน้อยกวันละ 2 ครั้ง ครูจะสร้างเครื่องมือชนิดใดจึงจะตรวจสอบได้ว่าเด็กปฏิบัติตามหรือไม่ ถ้าเครื่องมือนั้นสามารถตรวจสอบได้ว่านักเรียนคนใดปฏิบัติตาม คนใดไม่ปฏิบัติตาม ก็สามารถถกล่าวได้ว่าเครื่องมือนั้น ๆ มีความเที่ยงตรงตามสภาพ

1.2.2 ความเที่ยงตรงตามพยากรณ์ (Predictive Validity) เป็นค่าความเที่ยงตรงที่แสดงความสอดคล้องของผลการวัดจากเครื่องมือวัดนั้น ๆ กับสภาพเหตุการณ์ในอนาคตของเด็ก ถ้าเครื่องมือวัดชนิดใดสามารถพยากรณ์คุณลักษณะหรือความสามารถในอนาคตของเด็กได้ เช่น การนำผลการทดสอบความถนัดทางการเรียนไปหาค่าความสัมพันธ์กับเกรดเฉลี่ยของนักเรียนในสาขาวิชาหนึ่ง ๆ การหาค่าความเที่ยงตรงเชิงพยากรณ์จึงมักจะนำมาใช้ในการแนะนำเกี่ยวกับการศึกษาต่อและการเลือกอาชีพของนักเรียนได้

อย่างไรก็ตามจะเห็นได้ว่าความเที่ยงตรงตามสภาพและความเที่ยงตรงตามพยากรณ์นั้นต่างกับเป็นการหาค่าความสัมพันธ์ของคะแนนที่ได้จากเครื่องมือวัดกับเกณฑ์นั้นเอง ความแตกต่างของค่าความเที่ยงตรงของเครื่องมือวัดทั้ง 2 ชนิด ก็คือ เวลาที่ใช้ในการเก็บข้อมูลที่ใช้เป็นเกณฑ์ ถ้าเป็นข้อมูลที่ได้ในเวลาใกล้เคียงกับข้อมูลที่ได้จากเครื่องมือวัดก็เรียกว่าเป็นความเที่ยงตรงตามสภาพ แต่ถ้าเก็บข้อมูลจากเกณฑ์ภายหลังจากได้มีการสอบวัดไปแล้วระยะหนึ่ง ก็จะเรียกว่าเป็นความเที่ยงตรงตามพยากรณ์ นอกจากนั้นความแตกต่างของความเที่ยงตรงทั้ง 2 ชนิดนี้อีกประการหนึ่งก็คือ จุดมุ่งหมายที่ต้องการจะวัด ถ้ามีจุดมุ่งหมายที่จะตรวจสอบว่าคะแนนที่ได้จากเครื่องมือวัดนั้น ๆ วัดได้ตรงตามสภาพจริงหรือไม่ ก็เป็นความเที่ยงตรงตามสภาพ แต่ถ้ามีจุดมุ่งหมายที่จะนำมาใช้พยากรณ์เหตุการณ์ที่จะเกิดขึ้นในอนาคตก็เป็นความเที่ยงตรงตามพยากรณ์

การคำนวณหาค่าความเที่ยงตรงตามเกณฑ์นั้น คำนวณได้จากการหาค่าสหสัมพันธ์ระหว่างคะแนนจากแบบทดสอบและคะแนนจากเกณฑ์ เช่น หากความสัมพันธ์ระหว่างคะแนนข้อสอบวัดความถนัดทางการเรียนของนักเรียนชั้นม.1 กับคะแนนเฉลี่ยเมื่อนักเรียนเรียนจบชั้น ม.1 โดยวิธีของ Pearson Product Moment หลังจากนั้นก็นำค่ามาระบุทางสถิติเพื่อหาสมการพยากรณ์ต่อไป ค่าความเที่ยงตรงที่ได้ลักษณะนี้ก็เรียกว่าเป็นค่าความเที่ยงตรงเชิงพยากรณ์นั้นเอง

### 1.3 ความเที่ยงตรงตามโครงสร้าง (Construct Validity)

เป็นความเที่ยงตรงที่สำคัญที่สุดที่ใช้กันในทางจิตวิทยา โดยเฉพาะอย่างยิ่งใช้ในแบบทดสอบวัดทางจิตวิทยา เป็นคุณลักษณะของแบบทดสอบที่สามารถวัดโครงสร้างหรือคุณลักษณะ (traits) ที่ต้องการจะวัด

คำว่า โครงสร้างเป็นการอธิบายถึงเรื่องราวทางวิชาการ หรือสาระสำคัญเดียว ที่ปรากฏอยู่ตามทฤษฎีใดทฤษฎีหนึ่ง ซึ่งอาจจะเขียนออกมาในรูปของสมมุติฐานที่สามารถพิสูจน์ได้ เช่น คำว่าเชาว์ปัญญา จัดว่าเป็นโครงสร้างที่เกี่ยวกับความสามารถของมนุษย์ ซึ่งจะต้องจำแนกออกมามาว่ามีคุณลักษณะ (traits) แบบใดบ้าง ผู้ที่กำหนดโครงสร้างหรือคุณลักษณะออกมานั้น และต้องการพิสูจน์ว่าสิ่งที่เขากำหนดนั้นเป็นจริงหรือไม่ก็จะต้องทดสอบ ถ้าการทดสอบปรากฏว่าโครงสร้างนั้น ๆ ประกอบด้วยคุณลักษณะตามที่กำหนดไว้ ก็ถือว่าแบบทดสอบนั้นมีค่าความเที่ยงตรงตามโครงสร้าง แต่ถ้าปรากฏว่าโครงสร้างที่ทดสอบไม่สัมพันธ์กับคุณลักษณะที่กำหนด อาจมีสาเหตุมาจากการ

#### 1. แบบทดสอบมิได้วัดคุณลักษณะที่กำหนดได้จริง

2. ทฤษฎีที่กล่าวไว้ไม่ถูกต้อง และควรจะมีการปรับปรุงแก้ไข

3. การทดสอบมีช่องโหว่

การคำนวณหาค่าความเที่ยงตรงเชิงโครงสร้าง สามารถทำได้ดังนี้

1. วิธี Known group technique เป็นวิธีการหาความเที่ยงตรงตามโครงสร้างกับกลุ่มที่เราทราบคุณสมบัติอยู่แล้วกับกลุ่มทั่วไป เช่น สมมติเราทราบว่าผู้บริหารจะมีลักษณะของผู้นำ (leadership) มากกว่าคนทั่ว ๆ ไป ก็นำแบบทดสอบนั้นไปสอบกับกลุ่มผู้บริหารและกลุ่มคนทั่วไป แล้วนำคะแนนมาเปรียบเทียบกัน ซึ่งคะแนนความเป็นผู้นำของกลุ่มผู้บริหารควรจะสูงกว่าของกลุ่มบุคคลทั่วไป

2. วิธี Pretest and Post test technique คือวิธีการทดสอบก่อนและหลัง โดยยึดหลักการว่าระหว่างการสอบครั้งแรกและครั้งหลัง กลุ่มผู้ทดสอบควรจะได้รับการพัฒนาขึ้น คะแนนครั้งหลังควรจะสูงกว่าครั้งแรก

3. วิธีหาความคงที่ภายในของแบบทดสอบ (Internal Consistency) เป็นวิธีหาความสัมพันธ์ระหว่างคะแนนแต่ละข้อกับคะแนนรวม หรือหาความสัมพันธ์ระหว่างข้อคำถามด้วยกัน หรือระหว่างตอนของแบบทดสอบนั้น ๆ ถ้าพบว่ามีค่าสหสัมพันธ์ภายนอกสูงก็แสดงว่าแบบทดสอบนั้นได้วัดลักษณะที่ต้องการวัด

4. วิธี Multitrait-Multimethod เป็นวิธีการหาค่าความเที่ยงตรงภายใต้ใช้เมื่อคุณลักษณะที่ต้องการจะวัดมี 2 วิธีขึ้นไป เช่น มีคุณลักษณะที่ต้องการวัด 2 คุณลักษณะ คือ การเป็นคนเก็บตัว และการเป็นโรคประสาท และใช้วิธีวัด 2 วิธี คือ เห็นผู้ทดสอบทำข้อสอบถูก-ผิด กับทำข้อสอบแบบเลือกตอบ ดังนั้นกลุ่มตัวอย่างก็จะต้องทำแบบทดสอบ 4 ครั้ง ก็จะต้องนำเอาค่าสหสัมพันธ์นั้นมาเรียงเป็น matrix และดูค่าความสัมพันธ์ระหว่างคุณลักษณะที่จะวัด กับวิธีการที่ใช้วัด

5. ใช้วิเคราะห์ตัวประกอบ (Factor Analysis) เป็นวิธีการที่หาค่าสหสัมพันธ์ภายนอกของคะแนนที่ได้จากแบบทดสอบหลายชนิด และใช้วิธีการทางสถิติคือการวิเคราะห์ตัวประกอบหาคุณลักษณะร่วมกันอookma วิธีการนี้จำเป็นต้องใช้คอมพิวเตอร์มาช่วยคำนวณจึงจะสะดวก

องค์ประกอบที่มีอิทธิพลต่อค่าความเที่ยงตรงตามเกณฑ์

ในการแปลค่าสหสัมพันธ์นั้นจำเป็นจะต้องใช้เหตุผลในการพิจารณาอย่างรอบคอบ ทั้งนี้ก็เนื่องจากถ้าค่าสหสัมพันธ์สูงก็แปลว่าค่าความเที่ยงตรงจะต้องสูง แต่เนื่องจากค่าความเที่ยงตรงนั้นมีอิทธิพลบางประการมีผลทำให้ค่าสหสัมพันธ์เปลี่ยนแปลงไปได้ องค์ประกอบดังกล่าวได้แก่

## 1. ตัวแปรในแบบทดสอบ

แบบทดสอบบางฉบับอาจจะมีความเที่ยงตรงในตัวเองหรือในตัวเกณฑ์บางตัว แต่อาจจะไม่มีความสัมพันธ์กับตัวเกณฑ์อื่น ๆ ตัวอย่างเช่น เกรดของนักเรียนตอนปลายภาคอาจจะเหมาะสมที่จะใช้เป็นตัวเกณฑ์กับแบบทดสอบคัดเลือกเข้าเรียน (เช่นแบบทดสอบความถนัด) ซึ่งตัวอย่างนี้มีเหตุผลอยู่ในตัวของมันเอง แต่ถ้าเราจะหาเกณฑ์ในการวัดความวิตกกังวลอาจจะมีความยุ่งยากมากขึ้น จะเห็นได้ว่าถ้าเรายังไม่สามารถหาเกณฑ์ที่ดี ที่เหมาะสมได้แล้ว เรา ก็ไม่ควรจะหวังว่าจะหาค่าสัมประสิทธิ์สหสัมพันธ์กับตัวเกณฑ์ได้สูง ดังนั้นค่าความเที่ยงตรงจึงน่าจะอยู่ที่ว่าสามารถวัดคุณลักษณะมากกว่าจะต้องนำไปหาความสัมพันธ์กับตัวเกณฑ์

## 2. กลุ่มตัวอย่าง

ในการหาความเที่ยงตรงตามเกณฑ์นั้นบางครั้งเราอาจจะพบว่าแบบทดสอบฉบับหนึ่ง มีค่าความเที่ยงตรงสูงในเด็กกลุ่มหนึ่ง แต่พอนำไปใช้กับเด็กกลุ่มอื่นแล้วมักจะพบว่ามีค่าความเที่ยงตรงต่ำ ตัวอย่างเช่นแบบทดสอบบางฉบับสามารถจำแนกเด็กเก่งเด็กอ่อนได้ดีสำหรับเด็กประถมศึกษา แต่ถ้าเราเอาแบบทดสอบฉบับนี้ไปสอบกับเด็กมัธยมศึกษาค่าความเที่ยงตรงจะต่ำ เพราะเด็กมัธยมศึกษาสามารถทำแบบทดสอบได้ถูกหมด ดังนั้นถ้าบุคคลในกลุ่มที่ถูกทดสอบมีพิสัยของความสามารถต่ำ หรือคนในกลุ่มนี้มีความสามารถใกล้เคียงกัน ค่าความเที่ยงตรงของแบบทดสอบก็จะมีค่าต่ำ แต่ในทางตรงกันข้าม ถ้าคนในกลุ่มนี้ถูกทดสอบมีความสามารถแตกต่างกันมาก ๆ ค่าความเที่ยงตรงก็จะสูง

## 3. เกณฑ์

เนื่องจากความเที่ยงตรงตามเกณฑ์นี้เราจะต้องนำคะแนนจากแบบทดสอบไปสัมพันธ์กับเกณฑ์ ตัวเกณฑ์จึงนับว่ามีความสำคัญมาก ถ้าเราสามารถหาตัวเกณฑ์ที่ตรงกันกับแบบทดสอบ ค่าความเที่ยงตรงก็จะสูง แต่บางครั้งเกณฑ์อาจจะมีคุณสมบัติไม่เหมาะสมค่าความเที่ยงตรงก็จะต่ำ เราจะสรุปว่าแบบทดสอบนี้ไม่มีความเที่ยงตรงคงไม่ได้ ทั้งนี้เพราะควรจะพิจารณาองค์ประกอบของเกณฑ์ด้วยว่ามีความเหมาะสมเพียงใด ค่าความเที่ยงตรงจึงอาจมีมากหรือน้อยได้ตามเกณฑ์ที่นำไปสัมพันธ์ด้วย

## 4. ข้อมูลประกอบอื่น ๆ

การที่จะสรุปว่าแบบทดสอบได้เป็นแบบทดสอบที่ดีโดยดูเฉพาะที่ค่าความเที่ยงตรงนั้น ก็อาจจะไม่ชัดเจนนัก ทั้งนี้เพราะเราจะต้องพิจารณาด้วยว่าค่าความเที่ยงตรงนั้นให้รายละเอียดแก่เรามากน้อยเพียงใด แบบทดสอบฉบับหนึ่งอาจจะมีค่าความเที่ยงตรงสูงกับตัวแปรที่เป็น

ตัวเกณฑ์แต่อาจจะให้ประโยชน์อะไรไม่ได้มากนัก ลักษณะเช่นนี้มักจะปรากฏในแบบทดสอบที่นำไปสัมพันธ์กับเกณฑ์ที่เรารู้แล้วซึ่ดแล้วว่าจะต้องมีความสัมพันธ์กันแน่ ๆ ถ้าเป็นเช่นนี้เราก็คงไม่ได้อะไรใหม่จากการหาค่าความเที่ยงตรงของแบบทดสอบเลย

จากที่กล่าวมาข้างต้นเราได้รู้ว่า ค่าความเที่ยงตรงนั้นมีอิทธิพลจากองค์ประกอบอื่น ๆ ทำให้ค่าสูงหรือต่ำได้ ดังนั้นจึงไม่ต้องกังวลใจถ้าพบว่าแบบทดสอบที่เราสร้างมาแล้วอย่างรอบคอบแต่ปรากฏว่ามีค่าความเที่ยงตรงต่ำ แต่ถ้าองค์ประกอบต่าง ๆ ที่กล่าวมาแล้วไม่มีข้อใดบกพร่อง เราจึงจะสรุปว่าถ้าความเที่ยงตรงของแบบทดสอบมีค่าสูงมาก ๆ ทำให้แบบทดสอบมีประสิทธิภาพยิ่งขึ้น

## 2. ความเชื่อมั่น (Reliability)

หมายถึงความคงที่ของคะแนนที่ได้จากการสอบวัด ความเชื่อมั่น จึงนับว่าเป็นคุณลักษณะที่สำคัญของเครื่องมือวัดทุกชนิด เครื่องมือวัดที่วัดเด็กกลุ่มใดแล้วได้ผลคงที่ก็ถือได้ว่า เครื่องมือวัดนั้นมีความเชื่อมั่นหรือเชื่อถือได้ เช่น แบบทดสอบฉบับหนึ่งไปวัดกับเด็กกลุ่มหนึ่ง แล้วต่อมาอีก 3-5 วัน ก็นำแบบทดสอบฉบับเดียวกันนี้ไปสอบวัดซ้ำกับเด็กกลุ่มเดียวกันนี้อีกครั้งหนึ่ง ถ้าผลของการทดสอบทั้ง 2 ครั้ง ของเด็กแต่ละคนไม่แตกต่างไปมากนัก ก็อาจจะกล่าวได้ว่าแบบทดสอบนั้นมีค่าความเชื่อมั่นสูง อย่างไรก็ตามเหตุการณ์ที่สมมุติขึ้นนี้ มิอาจจะเกิดขึ้นได้มากนัก เพราะยังมีสาเหตุอื่น ๆ ที่ทำให้เด็กคะแนนเดียวกันลดลง หรือเท่าเดิม อีกมากมาย

### วิธีการหาค่าความเชื่อมั่น

การหาค่าความเชื่อมั่นของแบบทดสอบ ก็คือ การหาค่าความเชื่อมั่นได้ของคะแนนที่ได้จากการสอบวัด คะแนนที่ได้จากเครื่องมือแต่ละชนิดย่อมแตกต่างกัน เช่น แบบทดสอบบางชนิดให้คะแนนข้อถูกเป็น 1 คะแนน และให้คะแนนข้อผิดเป็น 0 คะแนน แบบทดสอบบางชนิดก็ให้น้ำหนักของคะแนนในแต่ละข้อแตกต่างกัน และเครื่องมือวัดบางอย่าง เช่น แบบสอบถาม ก็อาจให้คะแนนเป็นแบบมาตราส่วนประมาณค่า ดังนั้นวิธีการหาค่าความเชื่อมั่นของเครื่องมือวัดจึงแตกต่างกัน พอยังจำแนกได้ดังนี้

1. การวัดความคงที่ หรือการสอบซ้ำ (Measures of Stability or Test-Retest)
2. การวัดโดยใช้แบบทดสอบคู่ขนาน (Measures of Equivalence or Parallel Form)
3. การวัดความคงที่ภายใน (Internal Consistency)
4. การวัดโดยใช้คะแนนของผู้ตัดสิน (Reliability of Raters or Judges)

**1. การวัดความคงที่หรือการสอบซ้ำ (Measures of Stability or Test-Retest)** วิธีนี้ เป็นการอธิบายความคงที่ของคะแนนของผู้เข้าสอบซึ่งทำการทดสอบซ้ำ 2 ครั้ง โดยใช้แบบทดสอบฉบับเดียวกัน การกำหนดช่วงห่างของเวลาในการสอบทั้ง 2 ครั้งนั้น จะต้องพิจารณาตัดสินใจว่าแบบทดสอบชนิดนี้ ๆ จะมีความคลาดเคลื่อนจากช่วงห่างของเวลามากน้อยเพียงใด ปกติมักจะสมมุติว่า คุณลักษณะที่จะวัดแบบทดสอบนี้ ๆ จะไม่เปลี่ยนแปลง ความคลาดเคลื่อนที่เกิดขึ้นอาจเนื่องมาจากการแบบทดสอบไม่ถูกต้อง อย่างไรก็ตามความคลาดเคลื่อนของการวัดความเชื่อมั่นชนิดนี้ขึ้นอยู่กับช่วงห่างของเวลาในการสอบครั้งแรกครั้งหลัง นอกจากนี้ยังขึ้นอยู่กับคุณลักษณะของผู้สอบครั้งแรกและครั้งหลังอาจมีการเปลี่ยนแปลงไป

อย่างไรก็ตาม การหาค่าความเชื่อมั่นแบบนี้นิยมใช้ในแบบทดสอบ Speed test เพราะแบบทดสอบชนิดนี้ผู้เข้าสอบมักทำข้อสอบไม่ทันตามเวลาที่กำหนด แต่ละข้อที่ทำเสร็จผู้เข้าสอบก็มีความมั่นใจมาก คะแนนมักจะคงที่ทั้งครั้งแรกและครั้งหลัง ค่าที่ได้ก็มักจะมีค่าสูงในการเว้นช่วงห่างของระยะเวลาส่วนมากนิยม เว้นไว้ประมาณ 2-4 สัปดาห์ แล้วนำคะแนนในการสอบทั้ง 2 ครั้งมาหาค่าสัมประสิทธิ์สหสัมพันธ์ โดยใช้สูตรของ Pearson Product Moment Correlation

**2. การวัดโดยใช้แบบทดสอบคู่ขนาน (Measures of Equivalence or Parallel Form)** การหาค่าความเชื่อมั่นวิธีนี้ เป็นการหาความคงที่ของคะแนนของผู้เข้าทดสอบจาก การสอบ 2 ครั้ง โดยใช้แบบทดสอบที่เทียบเท่ากัน แบบทดสอบคู่ขนานกันหมายถึงแบบทดสอบที่วัดเนื้อเรื่องเดียวกัน มีค่าเฉลี่ย ค่าความแปรปรวน และค่าความยากง่ายเท่า ๆ กัน วิธีการคือใช้แบบทดสอบคู่ขนานทดสอบกับบุคคลกลุ่มเดียวกันในวันเดียวกัน ดังนั้นแต่ละคนก็จะมีคะแนน 2 ชุด จากแบบทดสอบ 2 ฉบับ และนำคะแนนจากแบบทดสอบทั้ง 2 ชุด มาหาค่าสัมประสิทธิ์สหสัมพันธ์โดยใช้สูตรของ Pearson Product Moment Correlation

**3. การวัดความคงที่ภายใน (Internal Consistency)** เป็นการหาค่าความคงที่ของคะแนนโดยการสอบด้วยข้อสอบฉบับเดียวและสอบครั้งเดียว การหาความเชื่อมั่นวิธีนี้ถือว่า เป็นดัชนีที่แท้จริงในการวัดความเป็นอันหนึ่งอันเดียวกันของเครื่องมือวัด นั่นก็คือในข้อสอบย่อย ๆ ในแต่ละคำถามมีความเกี่ยวพันกับคะแนนรวมของเครื่องมือวัดทั้งฉบับอย่างไร เป็นวิธีที่ช่วยลดความผูกพันในกรณีที่ไม่มีแบบทดสอบคู่ขนาน หรือไม่มีเวลาที่จะนำข้อสอบฉบับเดิมไปสอบซ้ำกับผู้ถูกทดสอบซ้ำเดิมอีก วิธีการหาความคงที่ภายในมีอยู่ 4 วิธีคือ

**3.1 ใช้วิธีแบ่งครึ่งข้อสอบ (Split-half, Odd-even)** วิธีการแบ่งครึ่งข้อสอบนี้ก็ยังหลักการเกี่ยวกับแบบทดสอบคู่ช้าน แต่ถือว่าเป็นวิธีการหาความคงที่ภายในก์เพราระข้อคำถามที่เป็นคู่ช้านนั้นรวมอยู่ในแบบทดสอบชุดเดียวกัน แทนที่จะสอบทีละชุดก์สอบหนเดียวทั้งสองชุด ในฉบับเดียวกัน วิธีทางจึงแบ่งแบบทดสอบเป็นข้อคู่ชุดหนึ่งและข้อคู่อีกชุดหนึ่ง แล้วนำคะแนนจากแบบทดสอบข้อคู่และแบบทดสอบข้อคู่มาหาค่าสัมประสิทธิ์สหสัมพันธ์โดยวิธีของ Pearson Product Moment Correlation ค่าสัมประสิทธิ์ที่ได้จะเป็นค่าสหสัมพันธ์ของแบบทดสอบครึ่งฉบับเท่านั้น ( $r_{1/2}, \frac{1}{2}$ ) เรายังต้องนำค่าสหสัมพันธ์นั้นมาคำนวณหาความเชื่อมั่นของแบบทดสอบอีกครึ่งหนึ่ง โดยใช้สูตรของ Spearman-Brown Prophecy Formula ดังนี้

$$r_{xy} = \frac{2r_{1/2}}{1 + r_{1/2}}$$

เมื่อ  $r_{xy}$  = ค่าความเชื่อมั่นของแบบทดสอบทั้งฉบับ  
 $r_{1/2}$ ,  $\frac{1}{2}$  = ค่าความเชื่อมั่นของแบบทดสอบครึ่งฉบับ

**3.2 ใช้สูตรของ Kuder-Richardson Formula 21** การใช้วิธีนี้นับว่าสะดวกและรวดเร็วในการคำนวณหาค่าความเชื่อมั่น แต่การใช้วิธีนี้ก็มีข้อตกลงเบื้องต้นว่า จะใช้ได้ก็ต่อเมื่อการตรวจให้คะแนนนั้นกำหนดให้ข้อถูกได้ 1 คะแนน และข้อผิดได้ 0 คะแนน การใช้สูตร KR<sub>21</sub> นี้มักเป็นที่นิยมใช้สำหรับครุภัติ ไป ทั้งนี้โดยมีข้อสมมุติว่าข้อสอบแต่ละข้อมีความยากง่ายเท่า ๆ กัน สูตรนี้คือ

$$r_{KR_{21}} = \frac{k}{k - 1} \left[ 1 - \frac{\sum (k - X_i) S_i^2}{k S^2} \right]$$

เมื่อ  $k$  = จำนวนข้อ  
 $X_i$  = ค่าเฉลี่ยของคะแนน  
 $S^2$  = ค่าความแปรปรวนของคะแนน

**3.3 ใช้วิธีหาค่าสัมประสิทธิ์แอลfa (Coefficient Alpha)** วิธีการนี้ Cronbach ได้คิดขึ้นในปี ค.ศ. 1951 โดยคำนวณจาก KR<sub>20</sub> แต่ข้อคำถามไม่จำเป็นจะต้องให้คะแนนถูกเป็น 1 ผิดเป็น 0 เมื่อตอนของ KR<sub>20</sub> ดังนั้นวิธีนี้จึงใช้ได้กับข้อสอบที่เป็นความเรียง หรือแบบสอบถามที่ให้คะแนนตามมาตรฐานประมาณค่า มีสูตรดังนี้

$$r_{xy} = \frac{n}{n - 1} \left[ 1 - \frac{\sum S_i^2}{S^2} \right]$$

เมื่อ	$r_{xx}$	= ค่าความเชื่อมั่นของแบบทดสอบ
	n	= จำนวนผู้เข้าทดสอบ
	$S_i^2$	= ค่าความแปรปรวนของข้อสอบแต่ละข้อ
	$S_t^2$	= ค่าความแปรปรวนของคะแนนรวมทั้งหมด

**3.4 ใช้วิเคราะห์ความแปรปรวนของ Hoyt (Hoyt)** Hoyt ได้คิดหาค่าความเชื่อมั่นขึ้นในปี ค.ศ. 1941 ซึ่งผลของการคำนวณก็จะคล้ายกับ KR<sub>20</sub> หรือสัมประสิทธิ์แอลฟ่า วิธีการของ Hoyt เป็นวิธีการวิเคราะห์ความแปรปรวนที่ถือว่า ความแปรปรวนของคะแนนรวมชุดหนึ่งจะมีค่าเป็นสัดส่วนเท่ากับความแปรปรวนของความแตกต่างระหว่างบุคคล ระหว่างข้อสอบ และระหว่างคนกับข้อสอบร่วมกัน ซึ่งสรุปเป็นตารางได้ดังนี้

แหล่งของความแปรปรวน	SS	df	MS
ระหว่างบุคคล		n - 1	
ระหว่างข้อ		k - 1	
ความคลาดเคลื่อน		(n - 1) (k - 1)	
รวม		kn	

#### 4. การวัดโดยใช้คะแนนของผู้ตัดสิน (Reliability of Raters or Judges)

การหาค่าความเชื่อมั่นแบบนี้หมาย味着สำหรับการหาค่าความเชื่อถือได้ของแบบวัดที่เป็นมาตราส่วนประมาณค่า แบบวัดบุคคลิกภาพที่เกิดจากผู้ให้คะแนน หรือผู้ประเมินค่า ในกรณีที่มีผู้ให้คะแนนเพียง 2 คน อาจใช้วิธีการของ Pearson Product Moment หรือแบ่งครึ่งแบบของ Spearman-Brown แต่ถ้ามีผู้ให้คะแนนมากกว่า 2 คนขึ้นไป อาจต้องใช้วิเคราะห์ความแปรปรวนของ Hoyt หรือหาค่าสัมประสิทธิ์แอลฟ่าของ Cronbach

#### แหล่งของความคลาดเคลื่อนในการหาค่าความเชื่อมั่น

การประมาณค่าความเชื่อมั่นแต่ละวิธีจะมีค่าสูงหรือไม่ขึ้นอยู่กับแหล่งของความคลาดเคลื่อนซึ่งจะทำให้ค่าความเชื่อมั่นเปลี่ยนแปลงไป ดังนั้นการหาค่าความเชื่อมั่นจึงต้องพิจารณา ความเหมาะสมของแต่ละวิธีด้วย แหล่งของความคลาดเคลื่อนในการหาความเชื่อมั่นมีดังนี้

##### 1. เกิดจากตัวผู้เข้าสอบ

ความคลาดเคลื่อนที่เกิดจากตัวผู้เข้าสอบนี้มักจะมีมากกว่าองค์ประกอบอื่น ๆ เช่น ผู้เข้าสอบได้รับการกระตุ้นให้ทำแบบทดสอบแตกต่างกันจะมีผลต่อค่าความเชื่อมั่นทันที นอกจากนั้นผู้เข้าสอบบางคนอาจเกิดความเปลี่ยนแปลงภายในตัวเอง เช่น สุขภาพทางร่างกาย สุขภาพ

ทางจิต ความอดทน ความมีน้ำใจนักกีฬา ความไม่เข้าใจในตัวข้อทดสอบบางข้อ ลักษณะนิสัยในการทำงาน ฯลฯ สิงต่าง ๆ เหล่านี้เกิดจากตัวผู้เข้าสอบเองที่จะมีผลทำให้มีอิทธิพลต่อค่าความเชื่อมั่นต่าง ๆ การแก้ไขเรื่องนี้ไม่มีวิธีการใดที่จะทำได้ แต่ก็ควรพยายามทำให้เกิดน้อยที่สุด เช่น การเตรียมตัวเด็กให้พร้อม ให้ด้วยการทำข้อสอบ กระตุนให้เกิดความกระตือรือร้นที่จะทำแบบทดสอบให้ได้ดีที่สุด ในเมื่อเราไม่สามารถจัดความคลาดเคลื่อนชนิดนี้ได้ จึงต้องพยายามทำให้เกิดน้อยที่สุดเท่าที่จะทำได้

## 2. อิทธิพลจากผู้ตรวจให้คะแนน

แหล่งของความคลาดเคลื่อนชนิดนี้อาจเกิดขึ้นได้จากผู้ตรวจgrade ตามมาตรฐาน ให้เด็กเขียนตอบ (paper pencil test) ตัวอย่างเช่น ผู้ดำเนินการสอบไม่ดำเนินการสอบตามคู่มือ การทดสอบ โดยให้เวลามากไปหรือน้อยไป หรือบางที่อาจช่วยเหลือผู้เข้าสอบมากกว่าที่ได้บอกไว้ในคู่มือการทดสอบ ผลกระทบจากการทำเช่นนี้อาจจะเกิดขึ้นทั้งแบบทดสอบวัดผลสัมฤทธิ์ หรือแบบทดสอบวัดสติปัญญา การที่ผู้ดำเนินการสอบไม่ปฏิบัติตามคำแนะนำ เช่น ช่วยเด็กนอกเหนือจากที่กำหนดไว้ ซึ่งให้ถูกต้องในแบบทดสอบ ให้เวลาเกิน ฯลฯ และการที่ผู้ดำเนินการสอบเข้มงวดหรือปล่อยปละละเลย รวมทั้งอารมณ์ของผู้ตรวจให้คะแนน (ถ้าเป็นข้อสอบแบบให้เขียนตอบ) ก็ย่อมจะมีผลต่อค่าความเชื่อมั่นของแบบทดสอบทั้งสิ้น อิทธิพลที่เกิดจากผู้ตรวจให้คะแนนมักจะมีปรากฏอยู่ที่สุดกับผู้ที่ยังใหม่และขาดประสบการณ์ และผู้ดำเนินการสอบจะช่วยให้ความผิดพลาดนิดน้อยลงได้พยายามทำตามคำแนะนำในการดำเนินการสอบให้มากที่สุด

ผู้ตรวจให้คะแนนมีโอกาสจะทำให้เกิดความคลาดเคลื่อนได้มากที่สุด ในกรณีที่ต้องดำเนินการสอบเป็นรายบุคคล จากรายงานผลการวิจัยมักแสดงให้เห็นว่าการดำเนินการสอบที่แยกต่างกันมีผลให้ผู้สอบได้คะแนนแตกต่างกัน นอกจากนั้นยังมีข้อผิดพลาดที่ระบบทดสอบที่มีการตรวจสอบให้คะแนนแบบปรนัยจะมีโอกาสเกิดความคลาดเคลื่อนในการตรวจให้คะแนนน้อยกว่าแบบทดสอบความเรียง เช่นการให้คะแนนวิชาคณิตศาสตร์ย่อจะเกิดความคลาดเคลื่อนน้อยกว่าการให้คะแนนบุคคลิกภาพหรือสติปัญญาของเด็กแต่ละคนที่มีการให้คะแนนแบบอัตโนมัติในกรณีในการทำคู่มือแบบทดสอบจะต้องแสดงความเชื่อมั่นของคะแนนไว้ด้วย

## 3. เนื้อหาของแบบทดสอบ

สิ่งที่เกี่ยวพันกับค่าความเชื่อมั่นมากที่สุดอีกประการหนึ่งคือเนื้อหาของแบบทดสอบแบบทดสอบ 1 ฉบับคือเป็นตัวแทนของความรู้ทั้งหมดในเนื้อหาวิชานั้น ซึ่งมักจะทำให้เกิด error

variance ได้มาก ในขณะเดียวกันคะแนนที่ได้จากการตอบแบบทดสอบของนักเรียนกลุ่มนี้ย่อมจะแตกต่างกันออกไปถ้าค่าตามที่ใช้ถามนั้นแตกต่างกัน ถ้ามีแบบทดสอบที่ใช้แทนกันได้ เช่นแบบทดสอบคุณภาพเราก็อาจจะประมาณค่าความเชื่อมั่นได้ โดยการให้เด็กกลุ่มเดียวกันตอบแบบทดสอบทั้ง 2 ฉบับ แล้วหาค่าสัมพันธ์ของแบบทดสอบทั้ง 2 ฉบับนั้น

วิธีการหาความเชื่อมั่นโดยให้นักเรียนของแบบทดสอบมาร่วมอยู่ในแบบทดสอบฉบับเดียว ซึ่งจะทำให้ค่าความคงที่ภายใน (internal consistency) สูงนั้น นิยมใช้วิธีการหาความเชื่อมั่นโดยใช้สูตรของคูเดอร์ ริ查ร์ดสัน (Kuder-Richardson) สูตรนี้ขึ้นอยู่กับพื้นฐานของข้อสอบแต่ละข้อกับคะแนนรวมทั้งหมด สามารถใช้ได้กับแบบทดสอบหลายชนิด ซึ่งพอจะนำมาใช้ประมาณค่าความเชื่อมั่นเนือหาของแบบทดสอบได้

วิธีการอีกวิธีหนึ่งที่นิยมใช้คือวิธี Split-half ซึ่งมีวิธีการให้คะแนนข้อคู่ครึ่งหนึ่ง และข้อคู่อีกครึ่งหนึ่ง แล้วนำคะแนนทั้ง 2 ชุดมาหาค่าความสัมพันธ์กัน การหาค่าความเชื่อมั่นวิธีนี้ มีข้อสังเกตว่าค่าความเชื่อมั่นของข้อสอบที่มีมาก ๆ ข้อมากจะสูงกว่าข้อสอบที่มีน้อยข้อ

การหาค่าความเชื่อมั่นทั้งแบบหาความคงที่ภายใน (internal consistency) และวิธี Split-half นั้น มีข้อตกลงเบื้องต้นหลายประการที่บางครั้งเราอาจทำไม่ได้ ข้อควรระวังประการหนึ่งก็คือ ถ้าเป็นแบบทดสอบประเภทความเร็ว (Speed test) ไม่สามารถจะใช้วิธีการหาค่าความเชื่อมั่นที่กล่าวนี้ได้ ตัวอย่างเช่น

ข้อสอบความเร็วนับหนึ่ง กำหนดให้ผู้สอบทำเครื่องหมาย S ถ้าข้อความเหมือนกัน แต่ถ้าต่างกันให้กาตัว D ให้เวลาทำเพียง 5 นาที (ข้อสอบ Speed test นี้จะมีมาก ๆ ข้อให้เวลาห้อย ๆ ซึ่งถ้าปล่อยให้ทุกคนโดยไม่จำกัดเวลาแล้วจะทำได้หมด) ดังนั้นพอเราแยกข้อสอบออก เป็นข้อคู่ ข้อคู่แล้วเราก็จะได้คะแนนทั้ง 2 ชุดเท่ากัน ดังนั้นการคำนวณค่าอุอกมาจึงสูงเกินกว่า ที่ควรจะเป็น นับว่าเป็นการใช้วิธีหาค่าความเชื่อมั่นที่ไม่ถูกต้อง แบบทดสอบประเภทนี้เข้าสอบจะมีความแตกต่างกันในด้านของอัตราความเร็วที่จะสามารถทำงานได้ไม่ใช่ด้วยความสามารถในการทำงาน

ถ้าสิ่งอื่น ๆ เท่ากันหมด แบบทดสอบที่มีมาก ๆ ข้อจะมีค่าความเชื่อมั่นสูงกว่าแบบทดสอบที่มีน้อยข้อทั้งนี้ เพราะเนื้อหาของแบบทดสอบจะเป็นตัวแทนของเนื้อหาทั้งหมดที่ต้องการทดสอบได้มากกว่า

#### 4. สถานการณ์ในการสอน

สถานการณ์ในขณะสอบก็เป็นส่วนหนึ่งที่ทำให้คะแนนจากการทดสอบแตกต่างกัน สถานการณ์ที่ว่ามีน้ำใจแก่ สภาพภัยในห้องสอบ เสียงรบกวน แสง จำนวนคนสอบ ที่รองเขียน

สอบ ยลน สิ่งเหล่านี้นับว่ามีอิทธิพลต่อค่าความเชื่อมั่นได้ นอกจากนี้การทุจริตในห้องสอบก็ทำให้เกิดความคลาดเคลื่อนได้ เพราะผู้เข้าสอบอาจจะได้คะแนนมากขึ้นหรือลดลงจากที่เข้าควรจะได้จริงถ้าไม่มีการทุจริต อิทธิพลของสิ่งต่าง ๆ เหล่านี้เป็นเรื่องยากที่จะประมาณค่าอภิมาได้ สิ่งแวดล้อมต่าง ๆ ภายในห้องสอบนี้ผู้ดำเนินการสอบควรควบคุมให้อยู่ในสภาพที่ปกติที่สุด เท่าที่จะทำได้ และผู้จะดำเนินการได้ตีกีดคู่ที่มีประสบการณ์

### 5. เวลา

เป็นความคลาดเคลื่อนอีกชนิดหนึ่งที่อาจเกิดขึ้นได้ เช่น แบบทดสอบฉบับเดียวกัน สอบกับคนกลุ่มเดียวกัน แต่ต่างเวลาออกไป ซึ่งการทำเช่นนี้จะต้องมีความคลาดเคลื่อนอันเนื่องจากระยะเวลาเกิดขึ้น เราจะหาความคลาดเคลื่อนนี้ได้โดยดูความสัมพันธ์ของคะแนนในการสอบครั้งแรกและครั้งหลัง

สมมติว่าในการสอบครั้งที่สองเราทำในเวลาใกล้ ๆ กับการสอบครั้งแรก เด็กบางคนอาจจำข้อสอบได้ ซึ่งย่อมจะมีผลต่อคะแนนในการสอบครั้งที่สอง นอกจากนั้นแบบทดสอบบางชนิดยังเกี่ยวข้องกับความยากง่ายของแบบทดสอบอีกด้วย แต่ถ้าหากว่าการสอบครั้งแรกและครั้งที่สองห่างกันมากก็อาจจะเกิดการเปลี่ยนแปลงอย่างอื่น ๆ ได้อีก เช่น ถ้าเราสอบดันปีครั้งหนึ่งแล้วไปสอบปลายปีเป็นครั้งที่สอง นักเรียนทุกคนได้เรียนเหมือนกันหมด แต่บางคนอาจได้ความรู้มากกว่าคนอื่น ๆ ซึ่งผลจากการเรียนรู้นั้นจะทำให้เขารอแบบทดสอบครั้งที่สองดีขึ้น มีผลทำให้คะแนนที่ได้แตกต่างกันไป

จากที่กล่าวมาแล้วจะเห็นว่าความเชื่อมั่นบางค่ามีความคลาดเคลื่อนหลาย ๆ อย่าง ซึ่งพิสูจน์ได้ในตาราง 7.1 ดังนี้

ตาราง 7.1 เปรียบเทียบแหล่งของความคลาดเคลื่อนในการประมาณค่าความเชื่อมั่นโดยวิธีต่าง ๆ

แหล่งของความคลาดเคลื่อน	วิธีการหาความเชื่อมั่น		
	สอบช้า	คุ่นนาน	ความคงที่ภายใน
ตัวผู้สอบ	X	X	
ตัวแบบทดสอบ	X	X	X
การดำเนินการสอบ	X	X	
ลักษณะกลุ่มผู้สอบ			X

## องค์ประกอบที่มีอิทธิพลต่อค่าความเชื่อมั่น

ค่าความเชื่อมั่นจะมีค่าสูงหรือต่ำย่อมขึ้นอยู่กับอิทธิพลขององค์ประกอบต่อไปนี้

### 1. ความยาวของข้อสอบ

จำนวนข้อของแบบทดสอบมีอิทธิพลต่อค่าความเชื่อมั่น แบบทดสอบใดมีน้อยข้อมากมีค่าความเชื่อมั่นต่ำ ถ้ามีจำนวนข้อมากจะมีค่าความเชื่อมั่นสูง จากสูตรที่แสดงให้เห็นถึงความสัมพันธ์ระหว่างความยาวของแบบทดสอบกับความเชื่อมั่นคือสูตรที่ว่าไปของ Spearman Brown ที่ว่า

$$r_{xx} = \frac{kr}{1 + (k - 1)r}$$

เมื่อ

$r_{xx}$  = ค่าพยากรณ์ความเชื่อมั่นของแบบทดสอบที่เพิ่มขึ้น  $k$  เท่า

$k$  = อัตราส่วนของจำนวนข้อในแบบทดสอบใหม่กับจำนวนข้อใน

แบบทดสอบเดิม

$r$  = ค่าความเชื่อมั่นของเดิมของแบบทดสอบ

ตัวอย่างเช่น ข้อสอบฉบับหนึ่งมีค่าความเชื่อมั่น .60 และถ้าต้องการขยายแบบทดสอบให้ยาวเป็น 3 เท่า ( เช่นเดิมมี 20 ข้อ จะขยายให้เป็น 60 ข้อ ) จะพยากรณ์ความเชื่อมั่นของแบบทดสอบที่เพิ่มความยาวได้เท่าไร

$$\begin{aligned} r_{xx} &= \frac{3(.60)}{1 + (3 - 1)(.60)} \\ &= .818 \end{aligned}$$

และถ้าจะลดข้อสอบลงครึ่งหนึ่ง เหลือ 10 ข้อจะมีค่าความเชื่อมั่นใหม่ดังนี้

$$\begin{aligned} r_{xx} &= \frac{\frac{1}{2} (.60)}{1 + (\frac{1}{2} - 1)(.60)} \\ &= \frac{(.5)(.60)}{1 + (.5)(.60)} \\ &= \frac{.30}{1.30} = 0.2307 \end{aligned}$$

จากตัวอย่างจะเข้าให้เห็นว่าข้อสอบยิ่งยาวความเชื่อมั่นก็ยิ่งสูงขึ้น ข้อสอบที่สั้นลงค่าความเชื่อมั่นก็ต่ำลง แต่ความสัมพันธ์ระหว่างจำนวนข้อกับความเชื่อมั่นดังแสดงในตาราง 7.2

ตาราง 7.2 แสดงความสัมพันธ์ระหว่างจำนวนข้อกับค่าพยากรณ์ความเชื่อมั่น

จำนวนข้อ	5	10	20	40	60	120	$\infty$
ความเชื่อมั่น	.10	.23	.60	.75	.82	.9	1.00

จากตาราง 7.2 แสดงให้เห็นว่า ถ้าข้อสอบเพิ่มขึ้น 5 ข้อ (จาก 5-10) ความเชื่อมั่นเพิ่มขึ้น .13 (.23-.10) แต่ข้อสอบเพิ่มขึ้น 100 ข้อ (120-20) ความเชื่อมั่นเพิ่มขึ้น 0.3 (.9-.6)

## 2. ระดับความสามารถที่แตกต่างกันของกลุ่มที่เข้าทดสอบ

ถ้าปรากฏว่ากลุ่มตัวอย่างที่นำมาใช้สอบแต่ละคนมีความสามารถแตกต่างกันมาก เรียกว่า เป็นกลุ่มวิวิชพันธ์ (heterogeneous group) ผลจากการวัดจะมีลำดับที่คงเดิมอยู่มาก เพราะช่วงความสามารถของแต่ละคนแตกต่างกันมาก แต่ถ้ากลุ่มตัวอย่างมีความสามารถใกล้เคียงกัน หรือเรียกว่าเป็นกลุ่มเอกพันธ์ (homogeneous group) โอกาสที่จะทำให้ตำแหน่งผิดพลาดจะมีมาก เราอาจใช้สูตรอธิบายเรื่องนี้ได้ดังนี้ (Mehren, 1984: 280)

$$r_{xx} = 1 - \frac{S_t^2}{S_x^2}$$

เมื่อ

$r_{xx}$  = ค่าความเชื่อมั่น

$S_t^2$  = ความแปรปรวนของคะแนนจริง

$S_x^2$  = ความแปรปรวนของคะแนนสอบได้

จากสูตรนี้จะเห็นว่าถ้าค่า  $S_t^2$  คงที่ แล้วค่า  $S_x^2$  มากขึ้น ในกรณีที่กลุ่มตัวอย่างแตกต่างกันมาก ค่าความเชื่อมั่นก็จะเพิ่มขึ้น แต่ถ้า  $S_x^2$  ลดลง ซึ่งก็หมายถึงเป็นกลุ่มที่มีความสามารถใกล้เคียงกัน ค่าความเชื่อมั่นก็จะลดลงด้วย

3. ข้อสอบแบบ Speed test เป็นข้อสอบง่าย ๆ แต่มีมากข้อ ผู้ตอบต้องอาศัยความรวดเร็วในการตอบคำถามส่วนมากนักเรียนจะตอบถูกทุกข้อที่ทำทันซึ่งก็จะได้คะแนนเท่านั้นเสมอ ดังนั้น สอบแต่ละครั้งจึงมีคะแนนคงเดิมเสมอจึงทำให้ข้อสอบประเภทนี้มีความเชื่อมั่นสูง

4. ความยากของข้อสอบ ซึ่งจะมีผลให้การกระจายของคะแนนมีมากน้อยต่างกัน การกระจายของคะแนนยิ่งมากก็จะทำให้ค่าความเชื่อมั่นสูงกว่าคะแนนที่มีการกระจายน้อยเนื่องจากข้อสอบง่ายมากนักเรียนส่วนใหญ่ทำได้ คะแนนจึงไม่แตกต่างกัน

5. ความเป็นปรนัย ข้อสอบใดที่ผู้ตรวจข้อสอบสามารถให้คะแนนได้เที่ยงตรง ข้อสอบนั้นก็จะมีความเชื่อมั่นสูง ดังนั้นจึงมักพบว่าข้อสอบอัตนัยที่ขาดความเป็นปรนัยในการให้คะแนน

## **มีความเชื่อมั่นต่ำกว่าข้อสอบปรนัยเสมอ**

เมื่อเปรียบเทียบค่าความเชื่อมั่นและความเที่ยงตรงแล้วจะเห็นว่าค่าความเที่ยงตรงเป็นการเปรียบเทียบค่าทางสถิติของแบบทดสอบกับตัวแปรภายนอกแบบทดสอบ ดังนั้นความคลาดเคลื่อนใด ๆ ที่เกิดภายใต้ตัวแบบทดสอบก็ย่อมมีผลโดยตรงต่อค่าความเที่ยงตรง ส่วนค่าความเชื่อมั่นนั้นไม่คำนึงว่าแบบทดสอบนั้นจะดีอย่างไรเป็นเพียงค่าความคงที่ของคะแนนในแบบทดสอบนั้น ๆ เท่านั้น ดังนั้นวิธีการดำเนินการสอบจึงมีผลโดยตรงต่อค่าความเชื่อมั่นและจะส่งผลกระทบอย่างมีนัยสำคัญต่อค่าความเที่ยงตรงด้วย

## สรุปย่อท้ายบท

1. ความเที่ยงตรงเป็นสิ่งที่สามารถจะวัดได้ในสิ่งที่ต้องการวัด ซึ่งจะนำผลไปอ้างอิงได้ ดังนั้นแบบทดสอบ 1 ฉบับ อาจจะวัดได้หลายจุดมุ่งหมาย จึงไม่จำเป็นว่าแบบทดสอบ 1 ฉบับ จะมีค่าความเที่ยงตรงเพียงค่าเดียว
2. ความเที่ยงตรงมี 3 ประเภท คือความเที่ยงตรงตามเนื้อหาความเที่ยงตรงตามเกณฑ์ และความเที่ยงตรงตามโครงสร้าง
3. ความเที่ยงตรงตามเนื้อหา คือความแม่นยำของแบบทดสอบที่จะตรงตามขอบเขตของเนื้อหาที่อ้างอิงว่าจะวัด
4. ความเที่ยงตรงตามเกณฑ์ เป็นเทคนิคการศึกษาความสัมพันธ์ระหว่างคะแนนการสอบกับตัวแปรอื่นภายนอกที่รัดได้ ดังนั้นการวัดตัวเกณฑ์ จะต้องตรงเป้าเชือถือได้ และไม่ล้าเอียง
5. ความเที่ยงตรงตามโครงสร้าง คืออัตราที่คะแนนการสอบจะสามารถอธิบายโครงสร้างของสิ่งที่ต้องการรู้
6. วิธีการหาความเที่ยงตรงบางอย่างจะต้องใช้ค่าสัมประสิทธิ์สหสัมพันธ์ ค่าสัมประสิทธิ์ในการตัดสินใจ ค่าความคลาดเคลื่อนในการคาดคะเน รวมทั้งการสร้างตารางพยากรณ์
7. องค์ประกอบที่มีอิทธิพลต่อค่าความเที่ยงตรงของแบบทดสอบคือ ตัวแบบทดสอบ กลุ่มตัวอย่าง เกณฑ์และข้อมูลอื่น ๆ
8. ความเชื่อมั่น เป็นอัตราความคงที่ระหว่างการวัดของสิ่งเดียวกันแต่เวัด 2 ครั้ง ซึ่งความไม่คงที่นั้นมีสาเหตุมาจากการความคลาดเคลื่อนของกลุ่มตัวอย่าง ผู้ตรวจให้คะแนนเนื้อหาในแบบทดสอบ สถานการณ์ในการสอบ และระยะเวลาในการสอบ
9. การคาดคะเนค่าความเชื่อมั่นได้จากการวัดความคงที่ของคะแนน การสอบด้วยแบบทดสอบคู่ขนาน และการหาความคงที่ภายในตัวแบบทดสอบ
10. การวัดความคงที่ของคะแนน คือการสอบช้ำแล้วเอาคะแนน 2 ครั้งมาหาความสัมพันธ์กัน ความคลาดเคลื่อนของคะแนน คือความเปลี่ยนแปลงในการสอบครั้งที่ 1 และครั้งที่ 2
11. การวัดจากแบบทดสอบคู่ขนาน คือ การให้ผู้สอบคนเดียวกันทำแบบทดสอบ 2 ฉบับ ในเวลาเดียวกัน และหากความสัมพันธ์ของคะแนนจากแบบทดสอบ 2 ฉบับนั้น
12. การวัดความคงที่ภายใน ใช้การสอบเพียงครั้งเดียว

13. การใช้ชีวิธีการคาดคะเนความเชื่อมั่นที่มีชีวิธีการต่างกันจะทำให้เกิดแหล่งความคลาดเคลื่อนต่างกัน

14. องค์ประกอบที่มีอิทธิพลต่อค่าความเชื่อมั่นได้แก่ ความยาวของข้อสอบ ความสามารถของกลุ่มที่แตกต่างกัน ข้อสอบแบบ Speed test ไม่ควรใช้ชีวิธีการหาความเชื่อมั่นแบบการหาความคงที่ภายในตัวแบบทดสอบ

15. ความยากของข้อสอบ และความเป็นปรนัย มีผลต่อค่าความเชื่อมั่นของแบบทดสอบ

16. ความเชื่อมั่นเป็นสิ่งจำเป็นแต่ก็ยังน้อยกว่าความเที่ยงตรง

## คำตามท้ายบท

1. จงเปรียบเทียบความหมายของความเที่ยงตรงและความเชื่อมั่น
2. วิธีการหาค่าความเที่ยงตรงมืออย่างไรบ้าง จงยกตัวอย่างการหาค่าความเที่ยงตรงแต่ละประเภท
3. จงกล่าวถึงอิทธิพลขององค์ประกอบที่มีต่อค่าความเที่ยงตรงแต่ละประเภท
4. จงนำตัวอย่างแสดงการคำนวณหาค่าความเชื่อมั่นของแบบทดสอบ 1 ฉบับ โดยกำหนดให้มีข้อสอบ 100 ข้อ สอบกับคน 30 คน (นักศึกษาอาจนำข้อมูลจากการสอบจริง ๆ มาทดลองหาค่าความเชื่อมั่น) และลองเปรียบเทียบค่าที่ได้แต่ละค่า
5. แหล่งของความคลาดเคลื่อนของความเชื่อมั่นมืออย่างไรบ้าง จงอธิบาย
6. จงแสดงเหตุผลที่องค์ประกอบแต่ละด้านมีอิทธิพลต่อการหาค่าความเชื่อมั่นแต่ละชนิด