

วรรณคดีที่เกี่ยวข้อง



การวิเคราะห์ข้อกระทง

แนวความคิดและเทคนิคของการวิเคราะห์ข้อกระทงก็เช่นเดียวกับความรู้ในการสร้างแบบสอบอื่น ๆ ซึ่งช่วยผู้ใช้แบบสอบในการประเมินผลแบบสอบต่าง ๆ แบบสอบจะใช้ประโยชน์ได้อย่างมีค่าขึ้นอยู่กับเทคนิคการเขียนข้อกระทงและระเบียบวิธีสถิติที่ใช้ การวิเคราะห์ข้อกระทงจะวิเคราะห์ทั้งด้านคุณภาพและปริมาณ ด้านคุณภาพวิเคราะห์หาและแบบฟอร์มที่ใช้ ด้านปริมาณวิเคราะห์ในรูปของคุณสมบัติทางสถิติ กล่าวคือ การวิเคราะห์ด้านคุณภาพรวมถึงความตรงเชิงเนื้อหาและการประเมินผลวิธีการเขียนข้อกระทง (Item Writing) การวิเคราะห์ด้านปริมาณมุ่งวัดระดับความยาก และอำนาจจำแนกของข้อกระทง ความตรงและความเที่ยงของแบบสอบขึ้นอยู่กับคุณสมบัติของข้อกระทงแต่ละข้อ เราสามารถจะสร้างแบบสอบให้มีความตรงและความเที่ยงสูงได้ เพราะหลังจากการวิเคราะห์ข้อ แบบสอบจะได้รับการปรับปรุงโดยการเลือก การตัดแทน การแก้ไขปรับปรุงข้อกระทง โดยทั่วไปแบบสอบที่ยาวกว่าจะมีความตรงและความเที่ยงมากกว่าแบบสอบที่สั้นกว่า แต่การวิเคราะห์ข้อกระทงจะช่วยทำให้แบบสอบสั้นลงในขณะเดียวกันจะเพิ่มความตรงและความเที่ยงได้โดยใช้สูตรของสเปียร์แมน บราวน์ (Spearman Brown) คาคคะเนผล ดังนั้นเมื่อใดทำให้แบบสอบสั้นลงโดยกำจัดข้อกระทงที่ไม่เป็นที่น่าพอใจทิ้งไป แบบสอบที่สั้นกว่าก็อาจมีความตรงและความเที่ยงมากกว่าแบบสอบยาวฉบับเดิม¹

การวิเคราะห์ข้อกระทง เป็นเทคนิคสำหรับตรวจคุณภาพข้อกระทงเป็นรายข้อว่าแต่ละ

¹Anne Anastasi, Psychological Testing (New York : The Mcmillan Company, 1968), p. 158.

ข้อนี้คุณลักษณะตรงตามวัตถุประสงค์ของการเพียงใด คำถามแต่ละข้อมีคุณค่าทางใด ความหมายชั้นแรกของการวิเคราะห์แบบสอบก็เพื่อทราบคุณภาพของข้อกระทง 2 ประการคือ ระดับความยาก (Item Difficulty) และอำนาจจำแนก (Item Discrimination) วิธีวิเคราะห์หลายแบบตั้งแต่วิธีง่าย ๆ จนถึงวิธีที่ค่อนข้างการคำนวณอย่างพิสดาร หลักใหญ่ของการวิเคราะห์ข้อกระทงคือเพื่อสร้างข้อกระทงและนำข้อกระทงไปใช้กับเด็กแล้วนำเอาผลการสอบมาตรวจให้คะแนนดูว่า ใครทำถูกกี่ข้อ ข้อใดบ้าง ได้คะแนนเท่าใด แล้วจึงนับเป็นรายข้อว่าข้อใดมีเด็กทำถูกกี่คน คิดเป็นเปอร์เซ็นต์ จากนั้นก็ใช้วิธีสถิติหาค่าต่าง ๆ ที่ต้องการ ข้อกระทงที่อยู่ในเกณฑ์ดี หมายความว่าข้อนั้นมีคุณภาพง่ายพอเหมาะและมีอำนาจจำแนกดี เหมาะสมที่จะรวมเป็นข้อสอบ แต่การวิเคราะห์ให้ได้รับรองว่าข้อกระทงนั้นมีคุณภาพดีในด้านการวัด ตรงตามพฤติกรรมที่เราปรารถนา อาจเป็นข้อสอบที่ดีในการวัดด้านความจำอย่างเดียวกันได้¹

ระดับความยาก (Item Difficulty) หมายถึง เปอร์เซ็นต์ของคนที่เลือกคำตอบที่ถูกในแต่ละข้อกระทง ผู้ร่างแบบสอบต้องการหาระดับความยากและการแจกแจงของระดับความยากด้วย เพื่อจะได้เลือกใช้ข้อกระทงให้ตรงตามความประสงค์ เช่น ใช้แบบสอบเพื่อเลือกนักเรียนเข้าเรียนหลักสูตรในระดับสูงจะเลือกข้อกระทงที่มีระดับความยากมากกว่าในระดับต้น แนวความคิดเกี่ยวกับระดับความยากนี้ใช้ได้อย่างมีความหมายกับแบบสอบสัมฤทธิ์ผล และแบบสอบวัดความสามารถ (Ability Test) ซึ่งเป็นแบบสอบที่มีคำตอบที่ถูก และนักเรียนจะพยายามทำให้ได้คะแนนสูง แต่ระดับความยากถ้านำไปใช้กับแบบสอบวัดบุคลิกภาพแล้ว ผลที่ได้ไม่ชัดเจนเพราะแบบสอบชนิดนี้ไม่มีคำตอบที่ถูก²

¹ชวาล แพร์ทกุล, เทคนิคการวัดผล (ฉบับปรับปรุง) พิมพ์ครั้งที่ 4. (พระนคร : โรงพิมพ์อักษรเจริญทัศน์, 2507), หน้า 275 - 325.

²Frederick G. Brown, Principal of Educational and Psychological Testing (Illinois : The Dryden Press Inc., 1970), p.36.

อำนาจจำแนก (Item Discrimination) อำนาจจำแนกที่ใช้เป็นตัวชี้บอกความตรง
ของข้อระทง คือข้อระทงสามารถวัดสิ่งที่ต้องการวัดได้ เกณฑ์ที่ใช้วัดสำหรับแบบสอบสัมฤทธิ์
ผลคือคะแนนสะสมปลายปี เปรียบเทียบกับคะแนนข้อระทง อาจกล่าวได้ว่า คนที่ตอบข้อระทง
ถูกเป็นคนเดียวกับคนที่ได้คะแนนสะสมปลายปีสูง ถ้าข้อระทงสามารถวัดได้ดังนี้ ข้อระทงนั้น
จะมีความตรง ข้อระทงใดก็ไม่ได้ ข้อระทงนั้นจะต้องปรับปรุงใหม่ ในกานปฏิบัติเครื่องมือ
ที่ใช้วัดยังไม่สมบูรณ์พอจึงต้องใช้วิธีทดแทนกัน คือหาสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนข้อ
ระทงกับคะแนนรวมของแบบสอบ (Item-Total Score Correlation) คนที่ได้คะแนน
รวมของแบบสอบสูงจะตอบข้อระทงถูกเป็นจำนวนมากกว่าคนที่ได้คะแนนรวมต่ำกว่า วิธีการนี้
จะได้ค่าอำนาจจำแนกของข้อระทงของแบบสอบ แบบสอบวัดบุคลิกภาพและแบบสอบอื่น ๆ ที่
ไม่มีค่าตอบถูกตงก็ใช้วิธีการวิเคราะห์อย่างอื่น¹

การพิจารณาตัวเลือกของข้อระทง (Alternative Response) การพิจารณา
จำนวนคนที่เลือกตอบค่าตอบต่าง ๆ ของแต่ละข้อระทงนั้น ถ้าตัวเลือกใดที่คนเลือกน้อย
หรือตัวเลือกที่ไม่ถูกเลือกเลย ตัวเลือกเหล่านี้จะต้องแก้ไข กล่าวคือ ตัวเลือกใดที่ไม่มีคน
เลือกตอบตัวเลือกนั้นไม่มีอำนาจจำแนกคนในระดับความรู้ต่าง ๆ กัน เช่นเดียวกันตัวเลือก
ของข้อระทงใดที่มีคนเลือกตอบทั้งหมดหรือเลือกเกือบหมดทุกคน ตัวเลือกตัวนั้นก็ขาดอำนาจ
จำแนก²

ลำดับขั้นในการวิเคราะห์ข้อระทง การวิเคราะห์ข้อระทงควรทำตามลำดับขั้นดังนี้

1. ตรวจสอบคะแนนแบบสอบแต่ละข้อระทง แต่ละตอนและทั้งฉบับ
2. คำนวณค่าสัมประสิทธิ์สหสัมพันธ์ ส่วนเบี่ยงเบนมาตรฐาน ความแปรปรวนของคะแนน
ของแบบสอบ
3. หาลักษณะรูปร่างของการแจกแจงของคะแนน

¹Brown, loc.cit.

²Brown. loc.cit.

4. คำวนัการะดับความยาก เช่น กำหนดการะดับความยากเท่ากับ .40 ถึง .60
5. ขัจขอกระทงทมิการะดับความยากไมอยูใน เกมที่กำหนดไว้ทังไป
6. หากคำสัมประสิทธีสหสัมพันธ์ระหว่างคะแนนขอกระทงกับคะแนนรวมแบบสอบ หมายถึง อำนาจจำแนกของขอกระทง เราจะกำหนดเป็น เกมที่ขึ้นมา เช่น อำนาจจำแนก .50 ขึ้นไป
7. คัดขอกระทงที่มีอำนาจจำแนกไมอยูใน เกมที่ทังไป
8. เขียนแบบสอบใหม่ทดแทนขอกระทงที่คัดทังไป
9. นำแบบสอบไปใช้กับคนจำนวนมาก
10. ทำซ้ำตั้งแต่ขอ 1 ถึง ขอ 9
11. ถ้าแบบสอบมี 2 ถึง 3 ตอน หาสหสัมพันธ์ระหว่างขอและวิเคราะห้ตัวประกอบ
12. หาปกติวิสัยของแบบสอบ

การวิเคราะห์ขอกระทงช่วยให้เราเห็นสภาพความจริงของแบบสอบ ช่วยครูในการสร้างแบบสอบและปรับปรุงแบบสอบ ตลอดจนคัดเลือกให้มีคุณค่า ผลการวิเคราะห์นำไปใช้ในการวินิจฉัยผลการเรียนของเด็กเป็นรายบุคคลได้¹

แบบสอบ

แบบสอบประกอบด้วยขอกระทงจำนวนหนึ่ง ผลการสอบขึ้นอยู่กับขอกระทงแต่ละข้อด้วย² แบบสอบเป็นเครื่องมือที่ใช้วัดตัวอย่างพฤติกรรมของบุคคล การวัดต้องมีระบบเพื่อลดอิทธิพลเกี่ยวกับตัวแปรต่าง ๆ ทั้งที่เกี่ยวกับบุคคลและสิ่งแวดล้อมให้มีผลต่อคะแนนน้อยที่สุด การวัดจะ

¹ Anastasi, Ibid., p. 159.

² Robert L. Thorndik, Educational Measurement (Washington D.C. : American Council on Education, 1972), p. 81.

วัดเนื้อหา วิธีดำเนินการสอน (Administration) และการให้คะแนน วิธีดำเนินการสอน
การสอบต้องเป็นมาตรฐาน การให้คะแนนต้องเป็นปรนัย¹

วัตถุประสงค์ของการใช้แบบสอบ



1. วัดความแตกต่าง ระหว่างบุคคล
2. เป็นเครื่องมือในการตัดสินใจของบุคคล
3. เป็นเครื่องมือในการเลือก จัดตำแหน่ง วินิจฉัย ประเมินผลของบุคคล

แบบสอบความถนัด (Aptitude Test) เป็นแบบสอบทางจิตวิทยาที่นิยมใช้วัด

ความแตกต่างระหว่างบุคคล ใช้เป็นเครื่องมือวัดความสามารถด้านต่าง ๆ ของผู้เรียน²
การเลือกแบบสอบความถนัดสำหรับสมัครคัดเลือกเพื่อวัดความสามารถที่จะเรียนสำเร็จของ
ผู้เรียน ผู้ใช้แบบสอบต้องเลือกแบบสอบที่มีประสิทธิภาพและได้มาตรฐาน ต้องมีการวิจัยเพื่อ
จะหาว่าแบบสอบมีความตรงและความเที่ยงเพียงใดในการที่จะใช้เป็นแบบสอบคัดเลือกบุคคล
เข้าศึกษาในสถาบันการศึกษา การมีแบบสอบความถนัดไว้สอบนักเรียนก่อนเข้าเรียนหรือก่อน
เลือกวิชาเรียนเป็นสิ่งสำคัญยิ่ง เป็นการเพิ่มประสิทธิภาพในการเรียนการสอนวิธีหนึ่งด้วย
และถ้าหากผลการวิจัยปรากฏว่าแบบสอบไม่เหมาะสมจะใช้วัดความถนัดของผู้เรียนก็เป็นแนว
ทางในการปรับปรุงแก้ไขข้อบกพร่องต่อไป³

¹Brown, Ibid., pp. 2 - 3.

²Anastasi, Ibid., pp. 1-3.

³W.M. Rothney, What Research Says to the Teacher (Washington D.C. : The National Education Association, 1955), p. 5.

แบบสอบผลสัมฤทธิ์ ใช้วัดความรู้ความเข้าใจตลอดจนความคิดในเนื้อหาวิชานั้น ๆ หรือจะกล่าวว่าเป็นการวัดคุณภาพการสอนของครูว่านักเรียนเรียนรู้ได้แค่ไหนจากประสบการณ์ที่ได้จากการเรียนในหลักสูตรใดหลักสูตรหนึ่ง แบบสอบชนิดนี้มีประโยชน์ต่อวงการศึกษและการวิจัยทางการศึกษามาก เพราะในการวิจัยทางการศึกษามักจะเกี่ยวข้องกับวิธีสอน และผลสัมฤทธิ์ของ เด็กอยู่เสมอ แบบสอบชนิดนี้เป็นแบบสอบที่ครูสร้างขึ้นเองและแบบสอบมาตรฐาน

แบบสอบมาตรฐาน (Standardized Test) เป็นแบบสอบที่วัดเนื้อหาวิชากว้างๆ และได้รับการสร้างอย่างดีโดยผู้เชี่ยวชาญทางจิตวิทยา มีความเข้าใจในเทคนิคการสร้างแบบสอบ การวิเคราะห์และระเบียบวิธีสถิติเป็นอย่างดี แบบสอบมาตรฐานจะรวมความเป็นมาตรฐานในเรื่องของเนื้อหาวิชา (Standard Content) วิธีดำเนินการสอบ และ

การให้คะแนน นอกจากนี้แบบสอบนี้ยังมีปกติวิสัยในการแปลความหมายของ คะแนน มีความเที่ยงและความตรงด้วย แบบสอบมาตรฐานแบ่งออกเป็นแบบทั่วไป และแบบพิเศษ แบบทั่วไปมักจะทำเป็นชุด (Batteries) คือในแบบสอบฉบับเดียวอาจจะมีแบ่งเป็นตอน ๆ หรือชุด โดยวัดความรู้หลายอย่าง เช่น การใช้ภาษา คำศัพท์ การอ่าน เลขคณิต และสังคม แต่แบบสอบมาตรฐานชนิดพิเศษมักจะทำเป็นรายวิชา

แบบสอบที่ผู้สอนต้องนำเอาประสบการณ์การเรียนรู้นำมาใช้ในการทำแบบสอบ เรียกว่าแบบสอบความสามารถ (Proficiency Test) แบบสอบความสามารถทางภาษาเป็นแบบสอบที่ใช้วัดว่าคนหนึ่งรู้ภาษาใดภาษาหนึ่งแค่ไหน (ไม่จำเป็นต้องเป็นนักเรียน) บางทีแบบสอบผลสัมฤทธิ์หรือแบบสอบความสามารถใช้จัดแบ่งนักเรียนให้อยู่ในชั้นปีที่ 1, 2, 3 ในหลักสูตรภาษาอังกฤษ แบบสอบเช่นนี้เรียกว่า แบบสอบจัดตำแหน่ง (Placement Test) ซึ่งความจริงก็คือแบบสอบผลสัมฤทธิ์ หรือแบบสอบความสามารถนั่นเอง¹

¹Robert Lado, Language Testing (New York : Longmans Green and Co. Ltd., 1961), pp. 342 - 369.

เมื่อมีการวัดและกำหนดสิ่งต่าง ๆ ออกมาเป็นตัวเลข ข้อมูลที่เราได้จากการวัดนั้น ถ้าเป็นข้อมูลซึ่งได้มาจากพฤติกรรมของมนุษย์ ข้อมูลเหล่านั้นอาจมีความคลาดเคลื่อนหลายอย่าง เช่น เครื่องมือวัดนั้น ๆ ไม่แน่นอน ผู้วัดเองอาจจะพลั้งเผลอวัดได้ไม่แม่นยำ หรือ สิ่งที่วัดนั้นมีอะไรแปรปรวนอยู่ภายใน สิ่งเหล่านี้ทำให้การวัดขาดความเที่ยงและความตรง

ความเที่ยงของแบบสอบ¹

หมายถึง ความคงที่ ความสอดคล้อง ความถูกต้องของแบบสอบ หรืออาจกล่าวได้ว่า ความเที่ยง หมายถึง ผลลัพธ์คงที่ไม่เปลี่ยนแปลงของเครื่องมือวัด ถ้าหากนำเอาเครื่องมือนั้น ไปทดลองอีกครั้งหนึ่ง

ข้อตกลงเบื้องต้น (Assumption) ของความเที่ยง คือการวัดความคลาดเคลื่อนที่เกิดขึ้นโดยวิธีการสุ่มตัวอย่าง สัมประสิทธิ์แห่งความเที่ยงแสดงสัดส่วนของความแปรปรวนที่เกิดขึ้นจริงของคะแนนสอบ

เทคนิคการหาความเที่ยงของแบบสอบอาจหาได้หลายวิธี ซึ่งรวบรวมแล้วได้ 4 แบบใหญ่คือ

1. การสอบซ้ำ (Test - Retest) หาสัมประสิทธิ์แห่งความคงที่ (Coefficient of Stability) คือหาค่าสหสัมพันธ์ของการสอบ 2 ครั้ง โดยใช้แบบสอบชุดเดิม และผู้สอบกลุ่มเดิม จำนวนควยสูตรของเพียร์สัน

ข้อตกลงเบื้องต้น (Assumption) มีดังนี้

1. คุณลักษณะที่จะวัดควยแบบสอบนั้นคงที่ตลอดเวลา
2. ผลจากการปฏิบัติไม่แตกต่างกัน
3. การเรียนรู้ที่เกิดขึ้นระหว่างวิธีดำเนินการสอบ 2 ครั้งไม่แตกต่างกัน

2. ใช้ฟอร์มคู่ขนาน (Parallel Forms) แบบสอบคู่ขนานมีลักษณะที่แต่ละ

¹Brown, Ibid., pp. 60 - 69.

ฉบับวัดเนื้อหาวีชาเดียวกัน ความยากง่ายเหมือนกัน และชนิดของข้อกระทงเหมือนกัน (Type of Item) การใช้ฟอร์มคูณหานี้ให้สัมประสิทธิ์แห่งความสมนัย (Coefficient of Equivalence) คือให้แบบสอบคูณหานไปสอบกลุ่มตัวอย่างกลุ่มเดียว แล้วหาคาสหสัมพันธ์ระหว่างคะแนนจากทั้งสองฟอร์มควยสูตรของ เพียร์สัน ปัญหาคือการใช้ข้อสอบวิธีแก้ปัญห ถ้าให้ปัญหา เช่น เกือบผู้สอบก็จะจำได้ ปัญหา เช่นนั้นไม่ควรจะให้ทำซ้ำอีกเพราะเกี่ยวกับความจำไม่ได้ใช้ความสามารถในการแก้ปัญหา กรณีปัญหา เช่นนี้ไม่ควรใช้แบบสอบคูณหาน

ข้อตกลงเบื้องต้น ในการคำนวณหาสัมประสิทธิ์แห่งความเที่ยงมีดังนี้

แบบสอบทั้งคู่จะต้องสมนัยกันในเนื้อหาวีชา แบบฟอร์ม ความยาว ความยากของแบบสอบ และมีความซับซ้อน เลขคณิตและค่าส่วนเบี่ยงเบนมาตรฐานเท่ากัน

3. การแบ่งครึ่งแบบสอบ (Split-Half Reliability) คือการใช้แบบสอบชุดเดียวกันไปสอบแล้วแบ่งครึ่งข้อสอบเป็น 2 ชุด หาคาสหสัมพันธ์และแก้ควยสูตรของสเปียร์แมน บราวน์ (Spearman-Brown) การแบ่งครึ่งแบบสอบมักจะแบ่งโดยใช้ข้อคู่และข้อคี่เป็นหลัก กล่าวคือ เมื่อนำแบบสอบไปสอบแล้ว ก็ทำการตรวจให้คะแนนโดยนับคะแนนเฉพาะข้อที่เป็นเลขที่รวมครึ่งหนึ่ง แล้วก็นับคะแนนเฉพาะข้อที่เป็นเลขคี่อีกครึ่งหนึ่ง ฉะนั้นผู้เข้าสอบแต่ละคนก็จะมีคะแนน 2 พวก คือ คะแนนข้อคู่และคะแนนข้อคี่ แล้วหาความสัมพันธ์ของคะแนนข้อคู่และข้อคี่ ก็จะได้ค่าความเที่ยงของแบบสอบซึ่งมีความยาวเป็นครึ่งหนึ่งของแบบสอบเดิม แล้วจึงนำค่าความเที่ยงนั้นมาแก้โดยใช้สูตรของสเปียร์แมน บราวน์

ข้อตกลงเบื้องต้น ความเที่ยงขึ้นอยู่กับความยาวของแบบสอบ

4. ความสอดคล้องภายใน (Internal Consistency) เทคนิคนี้ใช้ค่าประมาณความเที่ยงจากลักษณะค่าสถิติของแบบสอบนั้น ๆ ค่าความสอดคล้องภายในเป็นดัชนีที่จะบอกว่าข้อกระทงทุกข้อในแบบสอบวัดลักษณะที่เหมือนกัน ลักษณะเดียวกัน เทคนิคนี้มีข้อดีตรงที่ว่าสอบครั้งเดียวแล้วหาความเที่ยงได้ เทคนิคนี้มีข้อตกลงพื้นฐานว่าข้อสอบแต่ละข้อจะต้องวัดลักษณะเดียวกัน สูตรเบื้องต้นก็คือสูตรคูเคอร์ ริชาร์ดสัน สูตรที่ 20 แต่ค่าความยากง่ายของข้อสอบไม่แตกต่างกันมากอาจจะใช้สูตรคูเคอร์ ริชาร์ดสันสูตรที่ 21 การใช้สูตร

ของคูเคอร์ วิชา รัศัน สูตรที่ 21 นี้ มักจะประมาณค่าต่ำสุดของความเที่ยง ทั้งนี้เพราะลักษณะของข้อสอบแต่ละข้อไม่ เป็นไปตามข้อตกลงเบื้องต้นของสูตร เทคนิคการหาความเที่ยงที่เป็นความสอดคล้องภายในของแบบสอบนิยมใช้กันมากเพราะอาศัยค่าสถิติเบื้องต้นที่ได้จากข้อสอบนั้น ๆ

ตัวประกอบที่มีอิทธิพลต่อการหาสัมประสิทธิ์แห่งความเที่ยง

1. พิสัยของความแตกต่างระหว่างบุคคล การกระจายของคะแนนของกลุ่มตัวอย่างจะมีอิทธิพลต่อสัมประสิทธิ์สหพันธ์แห่งความเที่ยงของแบบสอบ

2. ความยากของแบบสอบ (Test Difficulty) ถ้าข้อสอบง่าย ยากไป ช่วงคะแนนจะแคบ อันมีผลต่อความเที่ยง แบบสอบจะมีความเที่ยงมากที่สุดที่ระดับความยากซึ่งทำให้คะแนนมีการกระจายมาก

3. ความยาวของแบบสอบ (Length of the Test) ถ้าเพิ่มความยาวมาก จะเพิ่มช่วงกว้างของคะแนน ถ้าคะแนนกระจายมากทำให้ความเที่ยงสูง

4. ความเร็ว (Speedness) การจำกัดเวลามีอิทธิพลต่อความเที่ยงของแบบสอบ นักเรียนที่ทำงานช้าจะทำไม่เสร็จหมดทุกข้อ

ถ้าทราบความเที่ยงของแบบสอบแล้ว ผู้สร้างแบบสอบต้อง การจะเพิ่มค่าความเที่ยงให้สูงขึ้นก็อาจทำได้โดยเพิ่มความยาวของแบบสอบ

ค่าความเที่ยง ควรจะเบ้มเท่าใดจึงจะเป็นที่พอใจนั้น Kelly¹ บอกว่าข้อสอบควรมีค่าความเที่ยงสูงแค่ไหนนั้นแล้วแต่ว่าจะไปใช้ประโยชน์ในแง่ไหน เขาได้กำหนดค่าความเที่ยงของแบบสอบแต่ละประเภทไว้ดังนี้

1. ถ้าต้องการประเมินผลสัมฤทธิ์ของคนกลุ่มใดกลุ่มหนึ่ง ค่าความเที่ยงควรมีตั้งแต่ .50 ขึ้นไป

001505

¹ E.F. Linquist, Educational Measurement (Washington D.C.: American Council on Education, 1951), p. 604.

2. ถ้าต้องการเปรียบเทียบผลสัมฤทธิ์ของคน 2 กลุ่มขึ้นไป จะต้องมีความเที่ยงตั้งแต่ .90 ขึ้นไป
 3. ถ้าต้องการประเมินผลสัมฤทธิ์ของคนเป็นรายบุคคล ค่าความเที่ยงตั้งแต่ .94 ขึ้นไป อยู่ในเกณฑ์
 4. ถ้าจะประเมินผลเปรียบเทียบความแตกต่างตั้งแต่ 2 คนขึ้นไป จะต้องมีความเที่ยงตั้งแต่ .98 ขึ้นไป
- เมื่อต้องการเทียบคุณภาพของแบบสอบ มักจะนำค่าสัมประสิทธิ์สหสัมพันธ์ของแบบสอบมาเปรียบเทียบกัน

ความตรงของแบบสอบ

ความตรงของแบบสอบ หมายถึง การที่แบบสอบวัดสิ่งที่ต้องการจะวัดได้และวัดได้อย่างแม่นยำ¹ ถ้าแบบสอบสามารถวัดได้แสดงว่าแบบสอบนั้นมีความตรง การที่จะให้แบบสอบมีความเที่ยงสูงก็ต้องพยายามทำแบบสอบนั้นให้มีความตรง (Validity) แนวความคิดเกี่ยวกับความตรงของแบบสอบคือถ้าไม่มีเกณฑ์กำหนดไว้แล้วสัมประสิทธิ์แห่งความตรงของแบบสอบจะหาออกมาไม่ได้ ผู้สร้างแบบสอบได้ใช้วิธีการต่าง ๆ หาเหตุผลที่จะประกันว่าแบบสอบที่สร้างขึ้นนั้นใช้ได้ตรงตามวัตถุประสงค์ เช่น หาความตรงตามเนื้อหาว่าแบบสอบนั้นประกอบด้วยเนื้อหาที่เป็นตัวแทนหรือตัวอย่างที่เพียงพอของเนื้อหาต่าง ๆ หรือไม่ แบบสอบจะมีความตรงเชิงเนื้อหาได้โดยไม่ต้องประกอบด้วยเนื้อหาทั้งหมดตามความมุ่งหมายของหลักสูตร เราจะพบบ่อยๆว่า ตัวอย่างที่ได้ออกมาโดยวิธีสุ่มมีผลน่าพอใจเท่ากัน และมีประสิทธิภาพมากพอที่จะพิจารณาได้ บางครั้งจึงใช้การตัดสินอย่างมีเหตุผล²

¹Fred N. Kerlinger, Foundations of Behavioral Research (New York : Holt, Rinehart and Winston, Inc., 1965), p. 429.

²Adkins Wood Dorothy, Test Construction (Ohio : Charles E. Merrill Books Inc., 1961), p. 21.



ชนิดของความตรงมี 5 ชนิดคือ

1. ความตรงตามเกณฑ์เทียบ (Criterion-Related Validity) ความตรงตามเกณฑ์เทียบแบ่งเป็นความตรงตามทำนาย (Predictive Validity) และความตรงร่วมสมัย (Concurrent Validity) ความตรงตามเกณฑ์เทียบใช้ทำนายพฤติกรรมที่เป็นเกณฑ์บางอย่าง ความตรงที่เกี่ยวข้องกับเกณฑ์ของแบบสอบจะแสดงผลของแบบสอบในการทำนายพฤติกรรมของบุคคลในสถานการณ์หนึ่งโดยเฉพาะ

ความตรงตามทำนายจะตรงจุดมุ่งหมายมากที่สุด ถ้าใช้แบบสอบเป็นเครื่องมือในการเลือกและจัดบุคคล เช่น การคัดเลือกนักเรียนเข้าเรียนในโรงเรียนหรือมหาวิทยาลัย แล้วหาความสัมพันธ์ระหว่างคะแนนสอบคัดเลือกกับเกณฑ์ซึ่งอาจจะเป็นคะแนนเฉลี่ยสะสม คะแนนการสอบอ่านเอาเรื่อง เปรียบเทียบกับคะแนนวิชาภาษาอังกฤษ นอกจากนี้หาความสัมพันธ์ระหว่างแบบสอบใหม่และแบบสอบอันก่อนก็มีประโยชน์ในการอ้างถึงการหาความตรง¹

ความตรงร่วมสมัยมีลักษณะคล้ายกับความตรงตามทำนาย ต่างกันในแง่เวลาเท่านั้น ความตรงตามทำนายหมายถึงทำนายเกณฑ์ในอนาคต แต่ความตรงร่วมสมัย หมายถึงทำนายเกณฑ์ในปัจจุบัน

อย่างไรก็ดี สำหรับการหาความตรงของแบบสอบ ความตรงตามเกณฑ์เทียบบางครั้งนิยมใช้ความตรงตามทำนาย การหาความตรงของแบบสอบนั้น วิธีที่ใช้อยู่ ๗ ได้แก่การคำนวณค่าสัมประสิทธิ์แห่งความตรงคือหาความสัมพันธ์ระหว่างคะแนนสอบหรือตัวทำนาย (x) กับคะแนนที่ใช้เป็นเกณฑ์ (y) โดยใช้สูตรของเพียร์สัน² (ดูในภาคผนวก)

การวิเคราะห์ข้อกระทง เพื่อเลือกข้อกระทงที่มีความตรงนั้นอาจจะหาความตรงตามทำนาย (Predictive Validity) ของข้อกระทงโดยหาความสัมพันธ์ชนิดไปซีเรียล

¹Anastasi, op.cit., pp. 105 - 121.

²Brown, Ibid., p. 105.



(Biserial Correlation)¹ ระหว่างคะแนนของแต่ละข้อกับคะแนนรวมของแบบ
สอบ

แนวคิดที่สำคัญในการวัดความตรง คือ แบบสอบนั้นวัดอะไร และวัดได้แค่ไหน
อันนี้จะดูได้จากค่าความแปรปรวนที่เกิดขึ้นจริง ความตรงของแบบสอบหาจากความสัมพันธ์
ระหว่างคะแนนรวมกับ เกณฑ์ที่จะวัด แบบสอบทางจิตวิทยาใช้ทำนายการปฏิบัติงานของคนใน
อนาคตบ่อยครั้ง เช่น ทำนายคะแนนเฉลี่ยของนักศึกษามหาวิทยาลัยหรือการปฏิบัติงานของ
นิสิต ความตรงในกรณีนี้คือคะแนนจากแบบสอบได้ทำนายฐานะหรือการปฏิบัติงานของเขาไว้ได้
แค่ไหน ความแน่นอนในการทำนายจะเป็นเครื่องชี้ความตรงของแบบสอบ คะแนนรวมของ
แบบสอบจะเป็นตัวสำคัญที่จะทำนายเกณฑ์เท่านั้น ดังนั้นในการปรับปรุงแบบสอบ เขาจะเลือก
ข้อกระทงที่จะทำนายเกณฑ์เอาไว้ เพราะฉะนั้นแบบสอบจะทำหน้าที่เป็นตัวทำนายมากกว่าที่
จะเป็นตัวแทนของพฤติกรรมที่ต้องการ²

ความตรงตามเกณฑ์ที่เทียบอาจแสดง โดยการเปรียบเทียบคะแนนสอบกับตัวแปรตัวอื่นๆ
หรืออาจจะใช้ค่าสหสัมพันธ์ของคะแนนจากแบบสอบกับเกณฑ์วัด³

2. ความตรงตามเนื้อหา (Content Validity) มีความสำคัญมากสำหรับ
วัดแบบสอบผลสัมฤทธิ์และแบบสอบความสามารถ ความตรงตามเนื้อหาเราหาว่าเนื้อหาของ
แบบสอบนั้นเป็นกลุ่มตัวอย่างที่ดีของ เนื้อหาทั้งหมดได้แค่ไหน⁴ เนื้อหาของแบบสอบสามารถ

¹J.P. Guilford, Fundamental Statistics in Psychology and Education (Tokyo : Kagakusha Company Ltd., 1965), p. 502.

²Frederick G. Brown, Principles of Education and Psychological Testing (Hinsdale, Illinois : The Dryden Press Inc., 1970),
p. 97.

³Brown, Ibid., pp. 98 - 99.

⁴Anne Anastasi, op.cit., pp. 100 - 102.

ครอบคลุมเนื้อหาต่าง ๆ หรือเป็นตัวแทนของพฤติกรรมที่จะวัดหรือไม่ หมายถึงการเป็นตัวแทน
ทั้งในรายละเอียดและหัวข้อต่าง ๆ ใด ในขั้นตอนนี้ต้องอาศัยการตัดสินใจของครูผู้สอนหรือผู้ออก
ข้อสอบ

ความตรงตามเนื้อหาที่ใช้ประเมินค่าแบบสอบผลสัมฤทธิ์นั้น ใช้การสุ่มตัวอย่างเนื้อหา
ที่มีประมาณภาพพอ และในขณะเดียวกันใช้การตัดสินใจของครูผู้สอนด้วย และต้องเป็นการตัดสินใจ
ใจอย่างมีเหตุผลที่สุดโดยไม่มีอคติใดๆ คำนึงถึงการสร้างแบบสอบต้องพยายามให้เนื้อหา
เฉพาะเจาะจง มีความเป็นปรนัย สะดวกในการให้คะแนนและเขียนคำสั่งที่ชัดเจน¹

3. ความตรงตามทฤษฎี (Construct Validity) แบบสอบทางจิตวิทยาใช้
ศึกษาลักษณะเฉพาะและพฤติกรรมที่แสดงออกของบุคคล คำถามที่ใช้คือแบบสอบวัดอะไร ความ
ตรงตามทฤษฎีของแบบสอบทำให้เห็นลักษณะของสิ่งที่วัดได้อย่างชัดเจน เช่น คะแนนรวมของ
แบบสอบแสดงความแตกต่างระหว่างกลุ่ม (Group Differentiation) เมื่อเราทราบลักษณะ
ใดลักษณะหนึ่งหรือหลายลักษณะที่แบบสอบวัดได้ เราใช้แบบสอบศึกษาความแตกต่างของบุคคล
วิธีหาความตรงเราหาจากหลักฐานที่รวบรวมได้ หลักฐานรวบรวมได้จากหลายวิธีและหลายแหล่ง
ข้อมูล ความสัมพันธ์ระหว่างลักษณะเฉพาะกับแบบสอบต้องชัดเจน และการกำหนดความหมาย
ของลักษณะเฉพาะจะต้องชัดเจน เช่น เกี่ยวกับความตรงตามเนื้อหา ไม่มีดัชนีจำนวนที่แสดงจำนวน
ที่แสดงความตรงตามทฤษฎีของแบบสอบ ดังนั้นการประเมินผลจึงใช้การตัดสินใจ

การทำให้แบบสอบมีความตรงตามทฤษฎีนั้น มีใจจะดูแลแบบสอบแต่จะต้องดูไปถึงทฤษฎี
ที่อยู่เบื้องหลังอีกชั้นหนึ่ง **ครอนบาค** ได้อธิบายว่า การทำให้แบบสอบมีความตรงตามทฤษฎี
มี 3 ตอน คือ ต้องคว้ามองถึงประกอบอะไร อยู่ในแบบสอบนั้น แล้ววางสมมุติฐานจากทฤษฎีที่
เกี่ยวข้องและสอบสมมุติฐานนั้นโดยการทดลองอย่างถูกวิธี

การทดสอบสมมุติฐานเกี่ยวกับทัศนคติทางการศึกษาของคนสามกลุ่มคือกลุ่มที่จบมหาวิทยาลัย
แล้ว กลุ่มที่กำลังศึกษาและกลุ่มที่อยู่นอกมหาวิทยาลัยนั้น ผู้วิจัยได้ใช้แบบสอบเป็นเครื่องมือ

¹Frederick G. Brown, op.cit., pp. 99 - 100.

ที่ชี้ให้เห็นความแตกต่างของทัศนคติทางการศึกษาของบุคคลทั้งสามกลุ่ม วิธีการเช่นนี้เรียกว่าวิธีการที่ใช้กลุ่มวิถึกันณะ บางครั้งใช้วิธีนี้หาความแตกต่างของค่าเฉลี่ยของกลุ่ม ตัวอย่างประชากรสองกลุ่ม คือกลุ่มเอ และ บี เราจะยอมรับสมมุติฐานถ้าแบบสอบถามได้ว่าประชากรสองกลุ่มนี้ต่างกันจริง คือคนที่ได้คะแนนสอบสูงจะได้คะแนนรวมจากเกณฑ์สูงด้วย ถ้าค่าเฉลี่ยของคะแนนสองกลุ่มแตกต่างกันอย่างมีนัยสำคัญ แสดงว่าแบบสอบถามนั้นมีความตรง นอกจากนี้อาจจะใช้วิธีอื่น เช่น คะแนนรวมของแบบสอบทั้งชุดมีความสัมพันธ์กับคะแนนรวมของแบบสอบย่อยซึ่งขอกระหนงแต่ละข้อว่าสิ่งเดียวกัน ขอกระหนง เช่นนี้มีความตรงตามทฤษฎี¹

ปกติวิสัย (Norms)

ปกติวิสัยใช้แปลความหมายของคะแนนซึ่งขึ้นอยู่กับคนที่สอบและวัตถุประสงค์ของแบบสอบ คะแนนดิบบอกอะไรไม่ได้ ปกติวิสัยใช้แปลความหมายคะแนนสอบของบุคคลโดยเปรียบเทียบคะแนนของบุคคลกับคะแนนของกลุ่มซึ่งเรียกว่าปกติวิสัย² การให้คะแนนเป็นเปอร์เซ็นต์หรือร้อยละขึ้นอยู่กับความยากง่ายของแบบสอบ ไม่ใ้บอกอะไรแน่นอน ดังนั้นในการแปลความหมายของคะแนนจึงต้องใช้ปกติวิสัย (Norms) เป็นเครื่องมือในการแปลผลการศึกษาของแต่ละคน เมื่อเทียบกับปกติวิสัยแล้วอยู่ในชั้นเฉลี่ยต่ำกว่าหรือสูงกว่าของปกติวิสัย

คานจิตวิทยาได้แบ่งค่าปกติวิสัยที่ใช้ทั่วไปเป็นหน่วยวัดผลดังนี้

1. คะแนนอายุ (Age Score) เริ่มใช้ตั้งแต่ปี ค.ศ. 1908 เมื่อสร้างแบบสอบวัดอายุสมองของ บิเน็ต (Binet) คะแนนเฉลี่ยแต่ละอายุไว้วัดอายุสมอง (Mental Age) โดยเอาผลเฉลี่ยของเด็กแต่ละอายุ เช่น เด็กอายุ 10 ขวบ โดยเฉลี่ยทำแบบสอบฉบับหนึ่งได้ 40 ข้อ เด็กคนไหนทำแบบสอบเดียวกันนั้นได้ 40 ข้อ เด็กคนนั้นมีอายุสมองเท่ากับเด็ก 10 ขวบ

¹Kerlinger, op.cit., pp. 453 - 454.

²Brown, Ibid., pp. 46 - 48.

2. ลำดับที่เปอร์เซ็นต์ไทล์ (Percentile Rank) หมายถึง ณ คะแนนที่กำหนดให้ มีคนร้อยละเท่าไรของกลุ่มตัวอย่างใดคะแนนต่ำกว่าคะแนนนั้น ลำดับที่เปอร์เซ็นต์ไทล์ใช้วัดความสามารถของคนเมื่อเทียบกับกลุ่ม ถ้าลำดับที่เปอร์เซ็นต์ไทล์ที่คำนวณได้มีค่าสูง แสดงว่านักเรียนคนนั้นมีความสามารถสูง เมื่อเทียบกับคนในกลุ่มนั้น

3. คะแนนมาตรฐาน (Standard Scores) ที่นิยมใช้ได้แก่

3.1 คะแนนทีปกติ (T - Scores) เป็นวิธีที่เปลี่ยนคะแนนดิบให้เป็นคะแนนมาตรฐาน ค่าทีปกติมีค่าเฉลี่ย (Mean) เท่ากับ 50 และส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation) เท่ากับ 10

3.2 สแตแนย (Stanine) มาจากคำว่า Standard Line ระบบนี้แบ่งหน่วยที่ใช้วัดออกเป็น 9 หน่วย มีค่าเฉลี่ย (Mean) เท่ากับ 5 และส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation) เท่ากับ 2 วิธีนี้นิยมใช้มาก เพราะใช้เลขหลักเดียวตลอด ทำให้คิดง่าย

อย่างไรก็ดี ในการพิจารณาเลือกใช้หน่วยวัดผลนั้นแต่ละวิธีมีข้อจำกัดใช้ได้ในลักษณะบางอย่าง เช่น หากลำดับที่เปอร์เซ็นต์ไทล์ คะแนนที่จะวัดนั้นจะต้องจัดเรียงกันตามลำดับ (Ordinal Scale) และหาปกติวิสัยจากคะแนนมาตรฐาน (Standard Scores) คะแนนดิบจะต้องจัดแบ่งเป็นชั้นคะแนนต่อเนื่องกัน¹ (Interval Scale) เป็นต้น

งานค้นคว้าและวิจัยที่เกี่ยวข้อง

พ.ศ. 2508 ชาวาล แพร์ตกุล, วิลาศ สิงหวิสัย และ วัน สังข์สะอาด² ได้รายงาน

¹Brown, *Ibid.*, p. 163.

²ชาวาล แพร์ตกุล, วิลาศ สิงหวิสัย และ วัน สังข์สะอาด, รายงานความก้าวหน้าของโครงการสร้างแบบสอบมาตรฐาน ความถนัดทางการเรียนของวิทยาลัยวิชาการศึกษา,

1 (พระนคร : สำนักงานทดสอบทางการศึกษาและจิตวิทยา, 2508).

เกี่ยวกับโครงสร้างแบบสอบมาตรฐานสำหรับวัดความถนัดทางการเรียน (CESAT หรือ College of Education Scholastic Aptitude Test) ของวิทยาลัยวิชาการศึกษา ประสานมิตร แบบสอบนี้ประกอบด้วยแบบสอบย่อย 5 ชุด

1. แบบสอบคำถามจัดเข้าพวก (Classification)
2. แบบสอบสรุปความ (Inference Exercise)
3. แบบสอบคณิตศาสตร์พื้นฐาน (Arithmetic Fundamental)
4. แบบสอบมิติสัมพันธ์ (Spatial Relationship)
5. แบบสอบตีความหมาย (Interpretation)

ผู้วิจัยได้นำแบบสอบไปทดสอบผู้สมัครเข้าศึกษาในวิทยาลัยวิชาการศึกษาบางแสนและประสานมิตร ในเดือนกรกฎาคม 2508 การวิเคราะห์ขั้นต้นใช้ผลสอบที่วิทยาลัยวิชาการศึกษา ประสานมิตรเพียงแห่งเดียว โดยใช้กลุ่มตัวอย่างจำนวน 1,624 คน ผลการวิเคราะห์แบบสอบชุดนี้ ปรากฏว่ามีระดับความยากเรียงจากยากไปทาง่ายดังนี้ มิติสัมพันธ์ ตีความหมาย จัดเข้าพวก คณิตศาสตร์และสรุปความ แบบสอบทุกชุดมีอำนาจจำแนกใกล้เคียงกัน การแจกแจงของคะแนนในแต่ละแบบสอบย่อย (Sub-Test) เป็นโค้งปกติและสามารถจำแนกผู้สอบตามความถนัดได้อีกด้วย แบบสอบชุดนี้มีจำนวน 50 ข้อกระทง แต่ละข้อกระทงมีตัวเลือก 5 ตัว ใช้เวลา 2 ชั่วโมง 20 นาที

พ.ศ. 2512 พจน์ สะเพียรชัย¹ ได้ทำการวิจัยเรื่อง การวิจัยองค์ประกอบของแบบสอบความถนัดทางการเรียนสำหรับชั้นประถมศึกษาปีที่ 7 โดยใช้แบบสอบความถนัดทางการเรียน ประกอบด้วยแบบสอบย่อย 6 ฉบับคือ

1. แบบสอบอนุภาคปฐมย
2. แบบสอบความสามารถในการอ่านตาราง

¹พจน์ สะเพียรชัย (กร.), "การวิจัยองค์ประกอบของแบบสอบความถนัดทางการเรียนสำหรับชั้นประถมศึกษาปีที่ 7," (คณะวิชาวิจัยการศึกษา วิทยาลัยวิชาการศึกษาประสานมิตร, 2509).



3. แบบสอบความสามารถในการคำนวณ
4. แบบสอบความสามารถในการเห็นรูปช้อนและรูปสี่เหลี่ยมทรงตัน
5. แบบสอบความสามารถเชิงกล
6. แบบสอบความสามารถทางภาษา (ศัพท์)

กลุ่มตัวอย่างที่ใช้ในการทดสอบครั้งนี้ เป็นนักเรียนชั้นประถมศึกษาปีที่ 7 จากจังหวัดพระนครศรีอยุธยา อ่างทอง ปัตตานี ศรีสะเกษ จังหวัดละ 2 โรงเรียน โรงเรียนละประมาณ 60 ถึง 70 คน รวมทั้งหมด 669 คน ผลการวิจัยปรากฏว่า ทักษะวิชาภาษา สังคม และ คณิตศาสตร์ มีความตรงรวมสมัย (Concurrent Validity) ระหว่าง .54 ถึง .78 เมื่อใช้คะแนนรวมผลสัมฤทธิ์ทุกวิชา มีค่าความตรงรวมสมัยระหว่าง .60 ถึง .75 ส่วนแบบสอบความสามารถเชิงกลและความสามารถในการเห็นรูปช้อน มีค่าความตรงตามทฤษฎี (Construct Validity) ระหว่าง .54 ถึง .84 และความเที่ยง (Reliability) มีค่าระหว่าง .70 ถึง .90

พ.ศ. 2512 สงบ ลักษณะ¹ ได้ทำการศึกษาความสัมพันธ์ระหว่างคะแนนสอบคัดเลือกคะแนนจากแบบสอบติดตามผลและผลการเรียนของนักเรียนฝึกหัดครูประกาศนียบัตรวิชาการศึกษาปีการศึกษา 2509 โดยศึกษาข้อสอบคัดเลือกนักเรียนฝึกหัดครูประกาศนียบัตรวิชาการศึกษา พ.ศ. 2508 ถึง พ.ศ. 2509 ใช้คะแนนเฉลี่ยสะสม 5 ภาคการศึกษาของนักเรียน เป็นเกณฑ์พบว่า แบบสอบคัดเลือกประเภทแบบสอบผลสัมฤทธิ์นั้น วิชาภาษาอังกฤษมีความสัมพันธ์กับเกณฑ์มากกว่าวิชาอื่น

พ.ศ. 2513 เสนาะ หัตถิ² ได้ทำการวิจัยศึกษาแบบสอบทางภาษาที่ส่งผลสัมฤทธิ์ใน

¹ สงบ ลักษณะ, "การศึกษาความสัมพันธ์ระหว่างคะแนนสอบคัดเลือก คะแนนจากแบบสอบติดตามผล และผลการสอบข้อเขียนของนักเรียนฝึกหัดครูประกาศนียบัตรวิชาการศึกษาปีการศึกษา 2509," (ปริทัศน์นิพนธ์การศึกษามหาบัณฑิต วิทยาลัยวิชาการศึกษาประสานมิตร, 2512).

² เสนาะ หัตถิ, "การศึกษาแบบสอบทางด้านภาษาที่ส่งผลสัมฤทธิ์ในวิชาเรียงความของนักเรียนชั้นมัธยมศึกษาปีที่ 1 ปีการศึกษา 2512," (ปริทัศน์นิพนธ์การศึกษามหาบัณฑิต วิทยาลัยวิชาการศึกษาประสานมิตร, 2513).

วิชา เรียงความของนักเรียนชั้นมัธยมศึกษาปีที่ 1 ปีการศึกษา 2512 โดยศึกษาความสัมพันธ์ระหว่างแบบสอภาษา กับผลสัมฤทธิ์ในการเขียนเรียงความ นำแบบสอภาษา 10 ชุด ไปทดสอบนักเรียนชั้นมัธยมศึกษาปีที่ 1 ของโรงเรียนในจังหวัดลพบุรี ปีการศึกษา 2512 โดยใช้แบบสอ เป็นตัวพยากรณ์ และคะแนนที่ได้จากการเขียนเรียงความของนักเรียนเป็นเกณฑ์ ได้ค่าสหสัมพันธ์ทุกคู่ของตัวพยากรณ์ทั้ง 10 กับเกณฑ์ เป็น 0.2625

พ.ศ. 2514 ศักดา บุญยไวโรจน์¹ ได้ทำการวิจัยเรื่องการวิเคราะห์ขอข้อความถนัดเชิงกลขององค์การโทรศัพท์แห่งประเทศไทย โดยคำนวณหาอำนาจจำแนกและระดับความยาก สัมประสิทธิ์แห่งความเที่ยง และปกติวิสัยของแบบสอ โดยใช้คะแนนจากผู้ที่มีครสอบคัดเลือกเข้าทำงานในองค์การโทรศัพท์แห่งประเทศไทยประจำปี 2513 จำนวน 629 คน ผลการวิจัยปรากฏว่า แบบสอย่อย 3 ชุด ได้แก่ แบบสอความถนัดด้านแรง และเครื่องมือกล มีค่าสัมประสิทธิ์แห่งความเที่ยง .8803 สัมประสิทธิ์แห่งความตรงมีค่า .7072 สัมประสิทธิ์แห่งความเที่ยงของแบบสอความถนัดความสัมพันธ์ด้านรูปร่าง มีค่า .9227 และสัมประสิทธิ์แห่งความตรงมีค่า .8869 สัมประสิทธิ์แห่งความเที่ยงของแบบสอความถนัดความสามารถด้านตัวเลข มีค่า .8574 และสัมประสิทธิ์แห่งความตรงมีค่า .5884 หากค่าสัมประสิทธิ์แห่งความเที่ยงของแบบสอทุกชุดโดยใช้วิธีหาความสอดคล้องภายในแบบสอความถนัดด้านแรงและเครื่องมือกล และแบบสอความถนัดความสามารถด้านตัวเลขนั้นคำนวณหาค่าความเที่ยงโดยใช้สูตรของคูเคอร์ ริชาร์ดสัน สูตรที่ 21 ส่วนแบบสอความสัมพันธ์ด้านรูปร่างใช้วิธีแบ่งครึ่งแบบสอเป็นข้อคู่และข้อคี่ หากค่าสหสัมพันธ์แล้วนำมาคำนวณแยกค่า โดยใช้สูตรของสเปียร์แมน บรรวณค่าสัมประสิทธิ์แห่งความตรงของแบบสอย่อยแต่ละชุดใช้คะแนนรวมเป็นเกณฑ์คำนวณโดยใช้สูตรการหาสัมประสิทธิ์สหสัมพันธ์ระหว่างส่วนย่อยกับส่วนรวม

¹ ศักดา บุญยไวโรจน์, "การวิเคราะห์แบบสอความถนัดเชิงกลขององค์การโทรศัพท์แห่งประเทศไทย," (วิทยานิพนธ์ครุศาสตรมหาบัณฑิต แผนกวิชาวิจัยการศึกษา บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย, 2514).

ค.ศ. 1961 ซีเกล¹ (Segel) ได้รวบรวมการวิเคราะห์แบบสอบเกี่ยวกับการพยากรณ์ความสำเร็จในการเรียนชั้นมัธยมศึกษา พบว่า การใช้คะแนนสะสมเฉลี่ยเป็นตัวพยากรณ์ไม่ใคร่ได้ผลดี แต่การใช้แบบสอบความถนัดวิชานั้นเป็นเกณฑ์ และคะแนนเฉลี่ยสะสมของวิชาเดียวกันนั้นเป็นตัวพยากรณ์ได้ผลดีกว่า วิชาอะไรควรจะใช้คะแนนของวิชาเดียวกันเป็นตัวพยากรณ์และเป็นเกณฑ์ เช่น ใช้คะแนนวิชาภาษาไทย ชั้นมัธยมศึกษาปีที่ 1 ไปพยากรณ์ความสำเร็จการเรียนภาษาไทยชั้นมัธยมศึกษาปีที่ 2 และ 3 ไม่ใช่ใช้วิชาหนึ่งไปพยากรณ์ความสำเร็จของอีกวิชาหนึ่ง โดยที่ทั้งสองวิชานั้นไม่เกี่ยวข้องหรือต่างเนื้อหากัน ทำให้ความเชื่อมั่นในการทำงานต่ำ ภัยเหตุนี้จึงนิยมใช้แบบสอบความถนัดการเรียนไปพยากรณ์ผลสัมฤทธิ์ทางการเรียนมากยิ่งขึ้น

ค.ศ. 1969 แอนาสตาซี² (Anastasi) ได้ศึกษาลักษณะของแบบสอบ รายงานว่า แบบสอบใช้ความเร็ว (Speed Test) ซึ่งเป็นข้อสอบค่อนข้างง่าย ใช้สูตร คูเคอร์ ริชาร์ดสัน (Kuder Richardson) หาความเที่ยงของแบบสอบไม่ได้ เพราะผู้เข้ารับการสอบทุกคนไม่มีโอกาสตอบข้อกระทงทุกข้อ ทั้ง ๆ ที่ตอบได้แต่เนื่องจากเวลาจำกัดมาก ทำให้ความเที่ยงคลาดเคลื่อนจากความเป็จริง ควรใช้วิธีทดสอบซ้ำ (Test - Retest) ส่วนแบบสอบใช้ความสามารถ (Power Test) นั้นเป็นแบบสอบค่อนข้างยากและยากกว่าแบบสอบความเร็ว และให้เวลานานพอที่ทุกคนสามารถตอบทันทุกข้อกระทง แบบสอบความสามารถวิชาภาษาอังกฤษของมิชิแกน³ (Michigan Test of English Language Proficiency) แบบสอบนี้ออกแบบขึ้นมาเพื่อจะประมาณความสามารถที่จะศึกษาวิชาภาษาอังกฤษในมหาวิทยาลัยหรือวิทยาลัยของนักศึกษาซึ่งไม่ใช่ภาษาอังกฤษ

¹Bloom Benjamin Segel, The Use of Academic Prediction Scale for Counselling and Selection College Entrance (New York : The Feu Press of Glencoe Inc., 1961), pp. 8 - 9.

²Anastasi, Ibic., pp. 89 - 91.

³Oscar K. Buro (ed.) The Sixth Mental Measurement Year Book. (Highland Park, New Jersey : The Gryphon Press, 1965), p. 385.

เป็นภาษาประจำชาติว่ามีความสามารถแค่ไหน

ลักษณะของแบบสอบมีดังนี้

ตอนแรก ให้ฟังข้อความภาษาอังกฤษแล้วให้แต่งประโยคเขียนเรียงความตอนสั้น ๆ ใช้เวลา 30 นาที

ตอนที่สอง เป็นแบบปรนัย มี 100 ข้อ เกี่ยวกับไวยากรณ์ ให้เติมคำในช่องว่าง เลือกจากตัวเลือกที่กำหนดให้ 4 ตัว เลือก จำนวน 40 ข้อ เกี่ยวกับคำศัพท์ ตามเกี่ยวกับความหมายของคำ มีจำนวน 40 ข้อ เกี่ยวกับการอ่านเอาเรื่อง มี 4 คอลัมน์ เป็นคอลัมน์สั้น ๆ ความยาวประมาณ 139 ถึง 217 คำ แต่ละคอลัมน์มี 5 คำตาม รวมเป็น 20 ข้อ ใช้เวลา 75 นาที สำหรับนักเรียนที่ทำซำก็จะทำไม่ถึงอ่านเอาเรื่อง ข้อ तरह ได้รับการคัดเลือกมาอย่างดี และสร้างเป็นแบบสอบที่ยากสำหรับคนที่ไม่ใช่ภาษาอังกฤษเป็นภาษาพูด

นักเรียนต่างประเทศจะทำได้ดีกว่ามากถ้าแบบสอบได้ใช้ภาษาของเขา หากว่าเขาอ่านบทความสั้น ๆ ตอนหนึ่งของแบบสอบแล้วเขาไม่สามารถที่จะตอบคำถามที่ง่าย ๆ หรือใช้เหตุผลธรรมดาแสดงว่านักเรียนทำไม่ได้เอง ภาษาไม่เป็นอุปสรรค คะแนนเฉลี่ยจากแบบสอบเท่ากับ 75.35 ความเที่ยง มีค่าเท่ากับ .965 (Split - Half) ความตรงตามทำนายหาได้จากความสัมพันธ์ระหว่าง คะแนนรวมของ แบบสอบและคะแนนเฉลี่ยของผลการสอบวิชาภาษาอังกฤษในมหาวิทยาลัยในโปรแกรมต่าง ๆ 6 โปรแกรม ค่า r จะอยู่ระหว่าง .26 ถึง .77 ถ้าความตรงตามทำนายนี้จะบอกเป็นการประมาณว่านักศึกษาซึ่งไม่ใช่ภาษาอังกฤษเป็นภาษาประจำชาติของเขาสามารถเรียนในมหาวิทยาลัยหรือวิทยาลัยที่สอนภาษาอังกฤษได้หรือไม่เท่านั้น แต่จะไม่ทำนายความสำเร็จในการศึกษาของเขา

ในปีพุทธศักราช 2503 ได้มีการปรับปรุงแบบสอบภาษาต่างประเทศซึ่งไร้ทดสอบความสามารถในการอ่าน เขียน ฟัง และพูดภาษาอังกฤษซึ่งมีประโยชน์มากในการสอบคัดเลือกเข้าศึกษาในมหาวิทยาลัยในประเทศสหรัฐอเมริกา จึงมีแบบสอบภาษา โดยเฉพาะเรียกว่า ซีอีบี (CEEB ย่อมาจาก College Entrance Examination Board) และได้เริ่มใช้เมื่อปีพุทธศักราช 2510¹

¹ Buros, *op.cit.*, pp. 395 - 360.

ปี พ.ศ. 2502 เบนเน็ตต์ ซีฮอร์ และเวสแมน¹ (Benett, Seashore and Wesman) ได้ปรับปรุงคู่มือแบบทดสอบความถนัดปี พ.ศ. 2495 (Manual for the Differential Aptitude Tests) โดยเพิ่มผลการวิจัยอื่นที่สำคัญซึ่งเกี่ยวข้องกับแบบสว
นี้ ทาปกติวิสัยรวมของแบบสวชุดเหตุผลคำพูด (Verbal Reasoning) กับชุดความ
สามารถคำนวณตัวเลข (Numerical Ability) แบบสวความถนัดนี้ประกอบด้วยแบบสว
ย่อย 8 ชุดด้วยกันคือ

1. แบบสวเหตุผลเกี่ยวกับคำพูด (Verbal Reasoning)
2. แบบสวความสามารถคำนวณตัวเลข (Numerical Ability)
3. แบบสวเหตุผลทางนามธรรม (Abstract Reasoning)
4. แบบสวความสัมพันธ์เชิงอากาศ (Space Relative)
5. แบบสวเหตุผลทางจักรกล (Mechanical Reasoning)
6. แบบสวความเร็วและความแม่นยำเกี่ยวกับเสมียนพนักงาน (Clerical Speed and Accuracy)
7. แบบสวการใช้ภาษา (Language Usage)
 - 7.1 แบบสวการสะกดคำ (Spelling)
 - 7.2 แบบสวประโยค (Sentence)

แบบสวทุกชุดมี 2 ฟอร์ม คือ ฟอร์ม เอ (Form A) และฟอร์ม บี (Form B) แบบ
สวที่เป็นแบบสวมาตรฐานใช้สวกับนักเรียนเกรด 8 ถึง เกรด 12 จำนวน 47,000 คน
จากรัฐต่าง ๆ ของอเมริกา เพื่อหาปกติวิสัยโดยถือชั้นเรียนเป็นเกณฑ์ (Grade Norms)
แบบเปอร์เซนต์ ค่าความค่าสัมประสิทธิ์แห่งความเที่ยงโดยวิธีแบ่งครึ่งแบบสว (Split-
Half Method) และค่าสัมประสิทธิ์แห่งความตรงเชิงพยากรณ์ (Predictive Validity)

¹George R. Benett, Harold G. Seashore and Alexander G. Wesman, Manual for the Differential Aptitude Tests. (New York : The Psychological Cooperation, 1957).

เพื่อทำนายผลการเรียน ผลการสอบองศาและสอบสัมฤทธิ์ผลทางการเรียนและอาชีพตลอดจนความสำเร็จในการเรียน

แบบสอบความถนัด ที เอ ที (The Differential Aptitude Tests : DAT) เป็นแบบสอบความถนัดมาตรฐานที่นิยมใช้กันมากชุดหนึ่ง เกรย์¹ (Gray) ได้ทำการวิจัยโดยนำแบบสอบเฉพาะแบบสอบย่อยเกี่ยวกับเหตุผลเกี่ยวกับคำพูด (Verbal Reasoning) เหตุผลทางนามธรรม (Abstract Reasoning) ความสามารถสะกดคำ (Spelling) และประโยค (Sentence) มาใช้กับวงการศึกษาในปี พ.ศ.2505 เกรย์ ได้นำไปสอบทหารที่ฟอร์ท บลิส (Fort Bliss, Texas) เมื่อคำนวณหาสัมประสิทธิ์แห่งสหสัมพันธ์ระหว่างคะแนนสอบกับคะแนนสัมฤทธิ์ผลทางการเรียน ปรากฏว่าสหสัมพันธ์ระหว่างเหตุผลเกี่ยวกับคำพูด (Verbal Reasoning) กับสัมฤทธิ์ผลทางการเรียนมีค่าสูง แสดงว่าแบบสอบย่อยชุดเหตุผลเกี่ยวกับคำพูดมีประสิทธิภาพสูง เหมาะที่สุดในการทำนายผลการเรียนและคัดเลือกบุคคลเข้าเรียนในวิชาการกานนี้



¹Beanard Gray, "The Differentiate Aptitude Test Military Academic Setting," The Journal of Educational Research, 68 : 352-354, April, 1965.