

วรรณคดีที่เกี่ยวข้อง

การวิเคราะห์ข้อสอบ คือการพิจารณาคุณสมบัติของข้อสอบเป็นรายข้อว่ามีลักษณะตรงตามวัตถุประสงค์ที่เราต้องการหรือไม่ มีมากน้อยเพียงใด และเป็น การหาค่าสถิติของข้อสอบแต่ละข้อมีลักษณะที่ไม่ดีในทางใดจำนวนเท่าไร การ วิเคราะห์ข้อสอบจะทำให้ได้แก่นำข้อสอบไปใช้สอบจริง ๆ กับเด็กแล้ว และนำ ผลที่เด็กตอบข้อสอบนั้นมาวิเคราะห์ ข้อสอบที่นำมาวิเคราะห์จะต้องเป็นข้อสอบปรนัย ชนิดเลือกตอบที่มี 4 หรือ 5 ตัวเลือกจึงจะดีที่สุด

ข้อสอบต่าง ๆ ที่นำมาวิเคราะห์หมักทำการวิเคราะห์ 2 ด้าน¹ คือ

1. การวิเคราะห์ด้านคุณภาพ (Qualitatively) หมายถึงการ วิเคราะห์ทางด้านเนื้อหาพฤติกรรมและแบบฟอร์มการเขียนข้อสอบว่าครบถ้วนตาม เนื้อหาพฤติกรรมหรือถูกแบบหรือไม่ ซึ่งเป็นการพิจารณาความตรงตามเนื้อหา (Content Validity) และใช้ประเมินผลว่า การเขียนข้อสอบนั้นใช้วิธีที่ถูก ต้องหรือไม่

2. การวิเคราะห์ด้านปริมาณ (Quantitatively) หมายถึงการวิเคราะห์ ในรูปของค่าสถิติ ซึ่งเป็นการวิเคราะห์ข้อสอบเป็นรายข้อโดยมีจุดมุ่งหมาย 2 ประการคือ

- 2.1 ทหาระดับความยาก (Item Difficulty) ของข้อสอบโดย จำนวนออกมาในรูปของสัดส่วน (p) หรือเปอร์เซ็นต์ (P) ของผู้ที่ตอบข้อสอบขอ นั้นถูกต้อง

¹ Anne Anastasi, Psychological Testing, (3d.ed., New York: The MacMillan Company, 1968), pp. 158-175.

2.2 ทหาอำนาจจำแนก (Power of Discrimination or Item Validity) ของข้อสอบ อำนาจจำแนก หมายถึง ข้อสอบข้อนั้นสามารถแยกผู้ตอบได้เป็นสองกลุ่ม คือกลุ่มเก่ง และไม่เก่ง ในการหาอำนาจจำแนกนั้นมีอยู่หลายวิธีด้วยกัน เช่น วิธีของจอห์นสัน (Johnson) ทหาอำนาจจำแนกโดยใช้ผลต่างของจำนวนผู้ตอบข้อสอบแต่ละข้อถูก หรือผิดจำนวนผู้ตอบข้อสอบทั้งหมดในกลุ่มสูงหรือกลุ่มต่ำ หรือจะหาอำนาจจำแนกโดยใช้ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนทั้งฉบับกับคะแนนในข้อสอบข้อนั้น ๆ โดยใช้วิธีของเพียร์สัน (Pearson Product-moment correlation) หรือไบซีเรียล หรือพอยท์ไบซีเรียล (Biserial or Point Biserial correlation) ก็ได้ เป็นต้น ค่าอำนาจจำแนก (r) จะมีค่าตั้งแต่ -1.00 ถึง $+1.00$

ในการวิเคราะห์ข้อสอบค่านปริมาณ นอกจากจะหาค่าสถิติดังกล่าวแล้วยังนิยมหาค่าความเที่ยงและความตรงของแบบสอบอีกด้วย เพราะเป็นลักษณะที่สำคัญของแบบสอบและยังขึ้นอยู่กับลักษณะของข้อสอบแต่ละข้อด้วย

ลำดับขั้นในการวิเคราะห์ข้อสอบส่วนใหญ่จะทำเหมือน ๆ กัน อีเบล (Ebel)¹ ได้เขียนลำดับขั้นของการวิเคราะห์ข้อสอบอย่างง่าย ๆ สมควรที่ครูจะนำไปใช้ ดังนี้

1. จัดเรียงกระดาษคำตอบจากคะแนนสูงสุดไปหาต่ำสุด
2. แยกกระดาษคำตอบออกเป็นสองกลุ่มโดยอาศัยเทคนิคกลุ่มสูง กลุ่มต่ำมีจำนวน 27% ของผู้เข้าสอบ การนับกลุ่มสูงนับจากกระดาษคำตอบ คะแนนสูงสุดลงไปจนครบจำนวนตามต้องการ จะได้เป็นกลุ่มสูง ส่วนกลุ่มต่ำนับกระดาษคำตอบจากคะแนนต่ำสุดขึ้นไปจนครบจำนวนตามต้องการเป็นกลุ่มต่ำ จำนวนคนในกลุ่มสูงและกลุ่มต่ำจะต้องเท่ากัน
3. นับจำนวนผู้ตอบถูกในแต่ละข้อกระทงในกลุ่มสูง บันทึกลงในกระดาษในกลุ่มต่ำก็ทำเช่นเดียวกัน

¹ Robert L. Ebel. Measuring Educational Achievement, (New Jersey: Prentice-Hall Inc., 1965), p. 347.

4. รวมจำนวนผู้ตอบถูกในกลุ่มสูงและกลุ่มต่ำแล้วหารด้วยจำนวนผู้ตอบทั้งหมดในกลุ่มสูงและกลุ่มต่ำ เสร็จแล้วคูณด้วย 100 เพื่อให้ค่าที่ได้เป็นจำนวนเปอร์เซ็นต์และค่าที่ได้นี้เรียกว่า ดัชนีระดับความยาก (The index of item difficulty)

5. เอาจำนวนผลต่างของผู้ตอบถูกในกลุ่มสูงกับกลุ่มต่ำ หารด้วยจำนวนผู้ตอบในกลุ่มสูงหรือกลุ่มต่ำ กลุ่มใดกลุ่มหนึ่ง ค่าที่ได้ เรียกว่า ดัชนีอำนาจจำแนก (The index of item discrimination)

การวิเคราะห์ข้อสอบ เทคนิคที่ใช้ในการแบ่งกลุ่มสูงกลุ่มต่ำมีหลายวิธี เช่น เทคนิค 25%, 27%, 33% และ 50% แต่ในทางปฏิบัติจริง ๆ เราไม่มีเครื่องมือในการคำนวณขมูมมาก ๆ ใช้ในโรงเรียน ดังนั้นการวิเคราะห์ข้อสอบจึงต้องอาศัยเทคนิค 25%, 27% หรือ 33% ซึ่งเป็นเพียงส่วนปลายสุดของคะแนน เทคนิคที่นิยมใช้และพบมากที่สุดในการนำมาใช้วิเคราะห์ข้อสอบ คือเทคนิค 27% ทั้งนี้ตัวอย่างประชากรจะตองมาก และมีคะแนนสอบแจกแจงเป็นแบบการแจกแจงปกติ (Normal Distribution) เคลเลย์¹(Kelley, 1939) ได้แสดงให้เห็นว่าหากใช้กลุ่มสูงและกลุ่มต่ำ กลุ่มละ 27% มาทำการคำนวณจะได้ค่าประมาณใกล้เคียงกับการที่จะนำขอมูลทั้งหมดมาวิเคราะห์ แต่ถาการแจกแจงของคะแนนไม่เป็นแบบการแจกแจงปกติ Cureton² (1957) ได้แสดงให้เห็นว่าตองใช้ถึง 33% จึงจะได้ค่าใกล้เคียง

การพิจารณาลักษณะของขอกระหวาขอใดสมควรเก็บไวตอไป คัดทิ้งหรือปรับปรุงแก้ไข ภายหลังกการวิเคราะห์ข้อสอบจะตองดูที่ระดับความยากและอำนาจจำแนกของแต่ละขอวาอยู่ในเกณฑ์ขอกระหวาที่คี่หรือไม่ เกณฑ์ระดับความยากที่จ้ควาพอเหมาะจะอยู่ระหว่าง 20% ถึง 80%³ ส่วนอีเบล (Ebel)⁴ ได้ประเมินผลขอกระหวาตามลักษณะของอำนาจจำแนกดังนี้

¹Anastasi, op.cit., p. 169.

²Loc. cit.

³ชวาล แพร์ตกุล, เรื่องเดิม, หน้า 309.

⁴Ebel, Measuring Educational Achievement, p. 364.

| ดัชนีอำนาจจำแนก | การประเมินผลของข้อกระทง |
|--------------------|---|
| ตั้งแต่ .40 ขึ้นไป | ข้อกระทงที่มากสมควรเก็บไว้ใช้ |
| .30 ถึง .39 | ข้อกระทงที่ แต่อาจมีการปรับปรุงบ้าง |
| .20 ถึง .29 | ข้อกระทงควรปรับปรุงใหม่ |
| ต่ำกว่า .19 | ข้อกระทงไม่สมควรตัดทิ้ง หรือปรับปรุงแก้ไขใหม่ |

การวิเคราะห์ข้อสอบนอกจากจะหาค่าความยากและอำนาจจำแนกของข้อสอบแต่ละข้อแล้วยังนิยมหาความเที่ยงของข้อสอบ โดยวิธีทดสอบซ้ำ (Test-Retest) บ้าง ใช้สูตรของคูเคอร์ ริชาร์ดสัน (Kuder Richardson Formula) บ้าง และหาความตรงของข้อสอบโดยใช้คะแนนเฉลี่ยประจำปี หรือประจำภาคการศึกษาเป็นเกณฑ์ (criterion)

รายงานการวิจัยในประเทศไทย

รายงานการวิจัยต่าง ๆ ที่เกี่ยวกับการวิเคราะห์ข้อสอบในประเทศไทยมีมากพอสมควรซึ่งส่วนใหญ่จะเป็นการวิเคราะห์ข้อสอบวัดความถนัด ข้อสอบคัดเลือกเข้าศึกษาต่อ เพื่อปรับปรุงแก้ไขให้เป็นข้อสอบที่มีคุณภาพดีขึ้น หรือบางรายงานเป็นการสร้างข้อสอบมาตรฐานเพื่อไว้ใช้ต่อไป เช่นแบบสอบวัดสมรรถภาพทางสมอง แบบสอบสัมฤทธิ์ผลในวิชาต่าง ๆ เป็นต้น งานวิจัยมีผู้จัดทำมาแล้วในประเทศไทยดังนี้

ในปีพ.ศ. 2506 วุฒิ เขาวภา¹ ได้วิเคราะห์ข้อสอบสัมฤทธิ์ผลวิชา

¹ วุฒิ เขาวภา, "ระดับความยากง่าย ระดับอำนาจจำแนกเป็นรายข้อและความเที่ยงของข้อสอบสัมฤทธิ์ผลชั้นประถมศึกษาปีที่ 5, 6 และ 7" (วิทยานิพนธ์ คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย, 2506).

คณิตศาสตร์ชั้นประถมศึกษาปีที่ 5, 6 และ 7 ซึ่งแผนกวิชาวิจัยการศึกษา คณะครุศาสตร์ เป็นผู้สร้างไคเดลครั้งนี้ ขอสอบแต่ละขอของแบบสอบในแต่ละชั้นมีทั้งข้อที่มีอำนาจจำแนกดีและไม่ดี มีข้อที่ยากมากและน้อยแตกต่างกัน แบบสอบชุดนี้มีความเที่ยงสำหรับเด็กนักเรียนชั้นประถมศึกษาปีที่ 5, 6 และ 7 เท่ากับ .85, .86 และ .84 ตามลำดับ ส่วนพิชชา เอกานนท์¹ ได้วิเคราะห์ขอสอบความพร้อมที่ดัดแปลงมาจากเมโทรโพลิแตน

(Metropolitan Readiness Test - Verbal) และขอสอบเขาวนที่ดัดแปลงมาจากแบบสอบโอทิส (Otis test) ซึ่งเป็นขอสอบคัดเลือกเข้าศึกษาต่อชั้นประถมศึกษาปีที่ 1 ของโรงเรียนสาธิตจุฬาลงกรณ์มหาวิทยาลัย ปีการศึกษา 2506 ผลการวิเคราะห์ขอสอบความพร้อมในเรื่องความเข้าใจคำ 19 ข้อ มีระดับความยากอยู่ระหว่าง 18% ถึง 95% และอำนาจจำแนกระหว่าง .10 ถึง .53 ในเรื่องความเข้าใจประโยค 14 ข้อ มีระดับความยากอยู่ระหว่าง 40% ถึง 97% อำนาจจำแนกระหว่าง .07 ถึง .58 ในเรื่องความเข้าใจเนื้อเรื่อง 14 ข้อ มีระดับความยากอยู่ระหว่าง 56% ถึง 92% อำนาจจำแนกระหว่าง .13 ถึง .50 ในเรื่องความเข้าใจการเปรียบเทียบ 19 ข้อ มีระดับความยากอยู่ระหว่าง 27% ถึง 82% อำนาจจำแนกระหว่าง .33 ถึง .59 ในเรื่องการนับจำนวน 24 ข้อ มีระดับความยากอยู่ระหว่าง 37% ถึง 90% อำนาจจำแนกระหว่าง .17 ถึง .73 และในเรื่องความสามารถในการเลียนแบบ 10 ข้อ มีระดับความยากอยู่ระหว่าง 43% ถึง 83% อำนาจจำแนกอยู่ระหว่าง .32 ถึง .59

พิชชา เอกานนท์, "ระดับความยากง่าย และระดับการเฟ้นจำแนก เป็นรายขอของขอสอบแบบความพร้อม ซึ่งดัดแปลงมาจาก Metropolitan Readiness Test (Verbal) และแบบสอบเขาวนของ Otis ซึ่งดัดแปลงมาจาก Otis Test (Alpha Test Form As) . และตำแหน่งเปอร์เซ็นต์ไทล์ของคะแนนนักเรียนที่สอบเข้าศึกษาชั้นประถมศึกษาปีที่ 1 ของโรงเรียนสาธิต จุฬาลงกรณ์มหาวิทยาลัย ปีการศึกษา 2506" (วิทยานิพนธ์ คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย, 2506).

ในปี พ.ศ. 2507 สุรีย์ เกิดผล¹ ได้สร้างข้อสอบสัมฤทธิ์ผลวิชา
คณิตศาสตร์ เพื่อสอบความรู้อะดับชั้นมัธยมปีที่ 1, 2 วิชาเลขคณิต และพีชคณิต ได้
อำนาจจำแนกเฉลี่ย 39.53 และระดับความยากเฉลี่ย 57.43% ส่วน บุพิน
ไชยวงศ์เกียรติ และเพื่อน² ได้ช่วยกันวิเคราะห์ข้อสอบคัดเลือกเข้าศึกษาต่อ
ระดับประกาศนียบัตรวิชาการศึกษา ปีการศึกษา 2506 ของกรมฝึกหัดครู ซึ่งแบบ
สอบแยกเป็น 4 วิชา คือภาษาไทย ภาษาอังกฤษ คณิตศาสตร์ และวิทยาศาสตร์
พบว่า หมวดคณิตศาสตร์ ภาษาไทย ภาษาอังกฤษ และวิทยาศาสตร์ มีข้อกระทง
ที่ดี คือมีระดับความยากระหว่าง 20% ถึง 80% และอำนาจจำแนก .20 ขึ้นไป
จำนวน 44, 70, 72 และ 55 ข้อตามลำดับ ข้อที่ไม่ดีคือมีอำนาจจำแนกต่ำกว่า
.20 จำนวน 13, 34, 17 และ 34 ข้อตามลำดับ ของง่ายคือข้อที่มีระดับความยาก
มากกว่า 80% ขึ้นไป แต่มีอำนาจจำแนก .20 ขึ้นไป จำนวน 1, 13, 9 และ 1
ข้อตามลำดับ และข้อกระทงที่ยาก คือข้อที่มีระดับความยากต่ำกว่า 20% ลงมา แต่
มีอำนาจจำแนก .20 ขึ้นไป จำนวน 2, 3, 2 และ 10 ข้อตามลำดับ ส่วนผลการ
วิเคราะห์ข้อเลือกพบว่าส่วนใหญ่ยังไม่ดีต้องแก้ไข และปรับปรุง

ในปีเดียวกันนี้เอง สมประสงค์ สถาปิตานนท์³ ได้สร้างข้อสอบการสะกดคำ

¹สุรีย์ เกิดผล, "การวัดผลการเรียนการสอนวิชาคณิตศาสตร์ตามหลักสูตร
ชั้นมัธยมปีที่ 1 พ.ศ. 2507" (วิทยานิพนธ์ คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย, 2507).

²บุพิน ไชยวงศ์เกียรติ, สุมาลี สะอาดบัว, วันเพ็ญ พิศาลพงศ์ และมาลินี
นกรพนม, "การวิเคราะห์ข้อสอบคัดเลือกเข้าเรียนประกาศนียบัตรวิชาการศึกษาใน
สถาบันฝึกหัดครู จังหวัดพระนครและธนบุรี ปีการศึกษา 2506" (วิทยานิพนธ์ คณะครุศาสตร์
จุฬาลงกรณ์มหาวิทยาลัย, 2507).

³สมประสงค์ สถาปิตานนท์, "การสร้างข้อสอบเขียนสะกดคำชั้นประถมปีที่ 1
และศึกษาดุลสัมฤทธิ์ในการเรียนสะกดคำของนักเรียน 11 โรงเรียน ในภาคศึกษา 1"
(วิทยานิพนธ์การศึกษามหาบัณฑิต วิทยาลัยวิชาการศึกษาระสาณมิตร, 2507).

ชั้นประถมศึกษาปีที่ 1 และศึกษาสัมฤทธิ์ผลในการเรียนสะกดคำของนักเรียนภาคการศึกษา 1 โดยสอบกับนักเรียนชั้นประถมศึกษาปีที่ 2 และคัดเลือกข้อสอบที่มีคุณภาพดีไว้ใช้

พ.ศ. 2508 ขวาด แพร์ทกุล, วิลาส สิงหวิสัย และ วัน สังข์สะอาด¹ ได้รายงานเกี่ยวกับโครงสร้างแบบสอบมาตรฐานสำหรับวัดความถนัดทางการเรียน (CESAT หรือ College of Education Scholastic Aptitude Test) ของวิทยาลัยวิชาการศึกษาประสานมิตร แบบสอบประกอบควย แบบสอบย่อย 5 ชุด คือ

1. แบบสอบคำถามจัดเข้าพวก (Classification)
2. แบบสอบสรุปความ (Inference Exercise)
3. แบบสอบคณิตศาสตร์พื้นฐาน (Arithmetic Fundamental)
4. แบบสอบมิติสัมพันธ์ (Spatial Relationship) 005417
5. แบบสอบตีความหมาย (Interpretation)

แบบสอบย่อยแต่ละชุดมีข้อกระทง 50 ข้อ แต่ละชุดมี 5 ตัวเลือก ใช้เวลา 2 ชั่วโมง 20 นาที นำไปทดสอบกับผู้สมัครเข้าศึกษาในวิทยาลัยวิชาการศึกษาประสานมิตร ในเดือน กรกฎาคม 2508 ในการวิเคราะห์ใช้กลุ่มตัวอย่าง 1,624 คน ผลการวิเคราะห์พบว่า มีระดับความยากเรียงจากยากไปหาง่ายดังนี้ มิติสัมพันธ์ ตีความหมาย จัดเข้าพวก คณิตศาสตร์ และสรุปความ ส่วนอำนาจจำแนกใกล้เคียงกันทุกชุด การแจกแจงของคะแนนในแต่ละแบบสอบย่อย (Subtest) เป็นโค้งปกติ และสามารถจำแนกผู้สอบตามความถนัดได้อีกด้วย

¹ ขวาด แพร์ทกุล, วิลาส สิงหวิสัย และ วัน สังข์สะอาด, รายงานความก้าวหน้าของโครงการแบบสอบมาตรฐานความถนัดทางการเรียนของวิทยาลัยวิชาการศึกษา (ฉบับที่ 1 สำนักงานทดสอบทางการศึกษาและจิตวิทยา, 2508)

พ.ศ. 2509 สหิส สุชะเคชะ¹ สร้างแบบสอบความถนัดด้านการเรียน
 ของนักเรียนชั้นประถมศึกษาปีที่ 4 และ 5 ซึ่งมีความถนัด 4 ด้าน คือด้านตัวเลข
 ความเข้าใจในการอ่าน การเขียนสะกดคำ และวิทยาศาสตร์ แล้วใช้เป็นเครื่อง
 มือเปรียบเทียบความถนัด 4 ด้าน ระหว่างเด็กชายกับเด็กหญิงที่มีอายุเท่ากัน และ
 ระหว่างเด็กที่มีอายุต่างกัน ปรากฏว่าเด็กชายมีความถนัดในทั้ง 4 ด้านมากกว่าเด็ก
 หญิง และเด็กที่มีอายุต่างกันมีความถนัดไม่แตกต่างกัน ในปีเดียวกันนี้ พจน สะเพียรชัย²
 ได้ทำการวิจัยเรื่อง การวิจัยองค์ประกอบของแบบสอบความถนัดทางการเรียนสำหรับ
 ชั้นประถมศึกษาปีที่ 7 โดยใช้แบบสอบความถนัดทางการเรียนประกอบด้วยแบบสอบย่อย
 6 ฉบับ คือ

1. แบบสอบอุปมาอุปมัย
2. แบบสอบความสามารถในการอ่านตาราง
3. แบบสอบความสามารถในการเห็นรูปซ้อนและรูปสี่เหลี่ยมทรงตัน
4. แบบสอบความสามารถในการคำนวณ
5. แบบสอบความสามารถเชิงกล
6. แบบสอบความสามารถทางภาษา (ศัพท์)

¹สหิส สุชะเคชะ, "การสร้างแบบสอบความถนัดในการเรียนสำหรับชั้น
 ประถมศึกษาปีที่ 4 และความถนัดในการเรียนชั้นประถมศึกษาปีที่ 5," (วิทยานิพนธ์
 การศึกษามหาบัณฑิต แขนงจิตวิทยาพัฒนาการ วิทยาลัยวิชาการศึกษาประสานมิตร,
 2509).

²พจน สะเพียรชัย, การวิจัยองค์ประกอบของแบบสอบความถนัดทาง
 การเรียนสำหรับชั้นประถมศึกษาปีที่ 7, (คณะวิชาวิจัยการศึกษา วิทยาลัยวิชาการศึกษา
 ศึกษาประสานมิตร, 2509).

กลุ่มตัวอย่างที่ใช้ในการทดสอบนี้เป็นนักเรียนชั้นประถมศึกษาปีที่ 7 จาก จังหวัดพระนคร เชียงใหม่ อ่างทอง ปัตตานี ศรีสะเกษ จังหวัดละ 2 โรงเรียน โรงเรียนละประมาณ 60 ถึง 70 คน รวมทั้งหมด 669 คน ผลการวิจัยปรากฏว่า ทักษะวิชาภาษา สังคม และคณิตศาสตร์มีความตรงร่วมสมัย (Concurrent Validity) ระหว่าง .54 ถึง .78 เมื่อใช้คะแนนรวมผลสัมฤทธิ์ทุกวิชา ความตรงร่วมสมัย มีค่าระหว่าง .60 ถึง .75 ส่วนแบบสอบความสามารถเชิงกลและความสามารถในการเห็นรูปซ่อน มีค่าความตรงตามทฤษฎี (Construct Validity) ระหว่าง .54 ถึง .84 และมีความเที่ยง (Reliability) ระหว่าง .70 ถึง .90

ในปีพ.ศ. 2511 บวรศรี ยายาคัน¹ ได้พัฒนาข้อสอบสมรรถภาพทางสมองเพื่อใช้กับนักเรียนไทยในชั้นมัธยมศึกษาตอนต้น โดยอาศัยแนวจากข้อสอบมาตรฐานของต่างประเทศ ข้อสอบทั้งหมดมี 6 ชุด คือ อุปมาอุปมัย ความสัมพันธ์เชิงที่ว่าง อนุกรมตัวเลข แพนคาตัวเลข โจทย์เลขคณิตและสรุปความ ใช้สอบกับนักเรียนชั้นมัธยมศึกษาตอนต้นของโรงเรียนราษฎร์ และโรงเรียนรัฐบาลจำนวน 2412 คน ผลการวิจัยปรากฏว่า แบบสอบอุปมาอุปมัยมีความยากอยู่ระหว่าง 23% ถึง 83% อำนาจจำแนกอยู่ระหว่าง .25 ถึง .66 ความเที่ยงมีค่าเท่ากับ .43 และความตรงเท่ากับ .09 มีประสิทธิภาพในการทำนายผลสำเร็จเป็น 4% แบบสอบความสัมพันธ์เชิงที่ว่าง มีระดับความยากอยู่ระหว่าง 32% ถึง 79% อำนาจจำแนกอยู่ระหว่าง .18 ถึง .51 ความเที่ยงเท่ากับ .65 ความตรงเท่ากับ .40 มีประสิทธิภาพในการทำนายผลสำเร็จ 8% แบบสอบอนุกรมตัวเลข มีระดับความยากอยู่ระหว่าง 26% ถึง 80% อำนาจจำแนกอยู่ระหว่าง .32 ถึง .63 ความเที่ยงเท่ากับ .64 ความตรงเท่ากับ .26

¹บวรศรี ยายาคัน, "การพัฒนาข้อสอบสมรรถภาพทางสมองเพื่อใช้กับนักเรียนไทยในชั้นมัธยมศึกษาตอนต้น" (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต แผนกวิชา ศึกษาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย, 2511).

มีประสิทธิภาพในการทำนายผลสำเร็จ 3% แบบสอบการแทนค่าตัวเลข มีระดับความยากอยู่ระหว่าง 31% ถึง 82% อำนาจจำแนก $-.02$ ถึง $.71$ ความเที่ยงเท่ากับ $.71$ ความตรง $.31$ มีประสิทธิภาพในการทำนายผลสำเร็จ 14% แบบสอบแบบโจทย์เลขคณิต มีระดับความยากอยู่ระหว่าง 30% ถึง 85% อำนาจจำแนก $.20$ ถึง $.69$ ความเที่ยงเท่ากับ $.60$ ความตรงเท่ากับ $.14$ มีประสิทธิภาพในการทำนายผลสำเร็จ 2% แบบสอบแบบสรุปความ มีระดับความยากอยู่ระหว่าง 33% ถึง 78% อำนาจจำแนก $.26$ ถึง $.61$ ความเที่ยงเท่ากับ $.43$ ความตรงเท่ากับ $-.05$ มีประสิทธิภาพในการทำนายผลสำเร็จในทางตรงข้าม 2% ข้อสอบรวมทั้งหมด มีความเที่ยงเท่ากับ $.71$ และความตรงเท่ากับ $.26$ และมีประสิทธิภาพในการทำนายผลสำเร็จ 3%

ในปีเดียวกันนี้ จีรพันธ์ จันทรศรีวงศ์¹ ได้พัฒนาข้อสอบสมรรถภาพทางสมองเพื่อใช้กับนักเรียนไทยชั้นประถมศึกษาตอนปลาย โดยอาศัยแนวข้อสอบมาตรฐานของพินเนอร์ (Pintner General Ability Test of Mental Maturity (Long Form) Level 2 and 3) ผลการวิจัยพบว่าหลังจากที่ได้ปรับปรุงแก้ไขแล้ว ข้อสอบฉบับสุดท้ายมีดัชนีความยากอยู่ระหว่าง 21% ถึง 91% อำนาจจำแนกระหว่าง $.28$ ถึง $.83$ ความเที่ยงของข้อสอบชั้นประถมศึกษาปีที่ 5, 6 และ 7 เป็น $.648$, $.722$ และ $.731$ ตามลำดับ ความตรงของข้อสอบชั้นประถมศึกษาปีที่ 5, 6 และ 7 เป็น $.661$, $.603$ และ $.753$ ตามลำดับ

¹จีรพันธ์ จันทรศรีวงศ์, "การพัฒนาข้อสอบสมรรถภาพทางสมองเพื่อใช้กับนักเรียนไทยในชั้นประถมศึกษาตอนปลาย" (วิทยานิพนธ์ปริญญาโทบัณฑิตศึกษาด้านจิตวิทยาการศึกษาระดับปริญญาโท คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย, 2511).

ในปี 2514 ศักดา บุญไวโรจน์¹ ได้ทำการวิเคราะห์หขอสอบความถนัดเชิงกลขององค์การโทรศัพท์แห่งประเทศไทย โดยใช้คะแนนจากผู้สมัครสอบคัดเลือกเข้าทำงานในองค์การโทรศัพท์แห่งประเทศไทยในปี 2513 จำนวน 629 คน ที่สอบแบบสอบซึ่งประกอบด้วยแบบสอบย่อย 3 ชุด คือ คำนวณตัวเลข คำนวณความสัมพันธ์ด้านรูปร่าง และคำนวณแรงและเครื่องกล ผลการวิจัยปรากฏว่า แบบสอบความถนัดคำนวณแรงและเครื่องกลมีค่าสัมประสิทธิ์ความเที่ยง .8803 สัมประสิทธิ์ความตรงเท่ากับ .7072 แบบสอบความถนัดคำนวณความสัมพันธ์ด้านรูปร่าง มีสัมประสิทธิ์ความเที่ยงเท่ากับ .9227 และสัมประสิทธิ์ความตรงเท่ากับ .8869 ส่วนแบบสอบความถนัดคำนวณตัวเลขมีสัมประสิทธิ์ความเที่ยงเท่ากับ .8574 และสัมประสิทธิ์ความตรง .5884

ในปี พ.ศ. 2515 จรรยา มีวาสนา² ได้ทำการวิเคราะห์หขอสอบคัดเลือกเข้าศึกษาต่อระดับประกาศนียบัตรวิชาการศึกษา วิทยาลัยครูสวนสุนันทา ปีพ.ศ. 2513 แบบสอบมี 5 ชุด คือ แบบสอบวิทยาศาสตร์ คณิตศาสตร์ ภาษาไทย ภาษาอังกฤษ และความรู้ทั่วไป โดยคณะกรรมการของวิทยาลัยเป็นผู้สร้างข้อสอบเอง ผลการวิจัยปรากฏว่า แบบสอบแต่ละวิชามีข้อสอบที่ตีตามเกณฑ์เรียงตามลำดับดังนี้ ภาษาไทย 42 ข้อ วิทยาศาสตร์ 41 ข้อ ภาษาอังกฤษ 40 ข้อ คณิตศาสตร์ 36 ข้อ และความรู้ทั่วไป 28 ข้อ ค่าสัมประสิทธิ์ความเที่ยงชนิดความคงที่ภายในของแบบสอบคณิตศาสตร์ วิทยาศาสตร์ ภาษาไทย ภาษาอังกฤษ และความรู้ทั่วไปเท่ากับ .6995, .5504, .5861, .7167 และ .4095 ตามลำดับ และความเที่ยงชนิดการทดสอบซ้ำ (Test-Retest)

¹ ศักดา บุญไวโรจน์, "การวิเคราะห์หขอสอบความถนัดเชิงกลขององค์การโทรศัพท์แห่งประเทศไทย" (วิทยานิพนธ์ปริญญาโทบริหารศึกษาศาสตร์ ภาควิชาการศึกษาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย, 2514).

² จรรยา มีวาสนา, "การวิเคราะห์หขอสอบคัดเลือกนักศึกษาวิชาครูประกาศนียบัตรวิชาการศึกษาของวิทยาลัยครูสวนสุนันทา," (วิทยานิพนธ์ปริญญาโทบริหารศึกษาศาสตร์ ภาควิชาการศึกษาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย, 2515).

ในวิชาเหล่านี้มีค่าเท่ากับ .7682, .5616, .5047, .6819 และ .3510 ตามลำดับ ส่วนสัมประสิทธิ์ความตรงตามทำนายของวิชาดังกล่าวมีอันที่เฉลี่ยของผลการสอบปลายภาคเรียนของแต่ละวิชาเป็นเกณฑ์ใดค่าความตรงเท่ากับ .5955, .5481, .4209, .6586 และ .1760 ตามลำดับ สัมประสิทธิ์ความตรงตามทำนายของวิชาความรู้ทั่วไปไม่สามารถใช้ทำนายได้อย่างมีนัยสำคัญ สัมประสิทธิ์ความตรงตามทำนายของคะแนนรวมทั้ง 5 วิชา กับเกณฑ์เท่ากับ .5476

ในปี พ.ศ. 2517 ประนิศรี สาริกัลยะ¹ ได้วิเคราะห์แบบสอบวิชาภาษาอังกฤษฉบับที่ 16 ที่อาจารย์คณะครุศาสตร์สามท่านเป็นผู้สร้างขึ้นเพื่อใช้สอบคัดเลือกนิสิตที่จบปริญญาตรี ทดลองเข้าศึกษาต่อปริญญาโทของคณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ผลการวิเคราะห์ปรากฏว่า ค่ามัธยเลขคณิตของคะแนนสอบในแบบสอบดังกล่าวเท่ากับ 59.97 ส่วนค่าเบี่ยงเบนมาตรฐานเท่ากับ 14.29 การแจกแจงของคะแนนกลุ่มตัวอย่างมีลักษณะใกล้เคียงโค้งปกติ ค่าสัมประสิทธิ์สหสัมพันธ์ของแบบสอบตอนที่ 1 และที่ 2 เท่ากับ .88 ค่ามัธยเลขคณิตของระดับความยากเท่ากับ 65.5 ค่ามัธยเลขคณิตของอำนาจจำแนกเท่ากับ .42 ข้อกระทงที่อยู่ในเกณฑ์คือมีระดับความยากระหว่าง 20% ถึง 80% และอำนาจจำแนก .20 ขึ้นไป มีจำนวนรอยละ .72 ความเที่ยงชนิดความสอดคล้องภายในของแบบสอบตอนที่ 1, 2 และทั้งฉบับเท่ากับ .86, .94 และ .72 ตามลำดับ จัดว่าอยู่ในระดับสูง ความตรงตามทำนายของแบบสอบเท่ากับ .11 ซึ่งค่อนข้างต่ำ ความตรงตามทฤษฎีของแบบสอบพบว่า คะแนนเฉลี่ยของกลุ่มตัวอย่างที่เลือกการสอบภาษาอังกฤษเป็นวิชาเอกสูงกว่ากลุ่มที่เลือกวิชาอื่น ๆ เป็นวิชาเอกอย่างมีนัยสำคัญ

¹ ประนิศรี สาริกัลยะ, "การวิเคราะห์แบบสอบวิชาภาษาอังกฤษฉบับ 16" (วิทยานิพนธ์ปริญญาโทฉบับที่ 10 แผนกวิชาวิจัยการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย, 2517).

ในปีเดียวกันนี้เอง สมคิด แก้วอรสาณ¹ ได้ทำการวิเคราะห์แบบสอบความถนัดเชิงวิศวกรรม φόρμ 16 ซึ่งคณะครุศาสตร์ และคณะวิศวกรรมศาสตร์ร่วมกันสร้างขึ้น แบบสอบประกอบด้วยแบบสอบย่อย 3 ชุด คือชุดความถนัดเชิงคณิตศาสตร์ ความถนัดเชิงความสัมพันธ์ทางรูปร่าง และความถนัดเชิงเหตุผลทางวิศวกรรมศาสตร์ จำนวนข้อทั้งหมด 60 ข้อ ใช้เวลาสอบ 90 นาที ผลปรากฏว่า แบบสอบความถนัดเชิงวิศวกรรม φόρμ 16 มีความยากอยู่ระหว่าง .175 ถึง .986 อำนาจจำแนกระหว่าง .002 ถึง .988 สัมประสิทธิ์แห่งความเที่ยงเป็น .512 สัมประสิทธิ์แห่งความตรงตามทำนายเท่ากับ .266 และมีความตรงเชิงเปรียบเทียบกลุ่ม คือความถนัดเชิงวิศวกรรมของบุคคลที่ผ่านการสอบคัดเลือกเข้าศึกษาต่อในคณะวิศวกรรมศาสตร์ สูงกว่าความถนัดเชิงวิศวกรรมของกลุ่มอื่นอย่างมีนัยสำคัญที่ .01

รายงานการวิจัยในต่างประเทศ

การวิจัยในต่างประเทศได้มีการวิเคราะห์แบบสอบกันมานานแล้ว จนมีบริษัทสร้างแบบสอบมาตรฐานเกิดขึ้นหลายบริษัท งานวิจัยที่ยกมาเป็นตัวอย่างมีดังนี้

ในปี ค.ศ. 1961 บลูมและปีเตอร์ (Bloom and Peter)² ได้รวบรวม

¹สมคิด แก้วอรสาณ, "การวิเคราะห์แบบสอบความถนัดเชิงวิศวกรรม φόρμ 16" (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต แผนกวิชาวิจัยการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย, 2517).

²Benjamin S. Bloom, and Frank R. Peter, The Use of Academic Prediction Scale for Counselling and Selection College Entrants (New York: The Free Press of Glencoe, Inc., 1961), pp. 8-9.

การวิเคราะห์แบบสอบเกี่ยวกับการพยากรณ์ความสำเร็จในการเรียนชั้นมัธยมศึกษา พบว่า การใช้คะแนนสะสมเฉลี่ยเป็นตัวพยากรณ์ไม่ใคร่โดยลึกลับ แต่การใช้แบบสอบความถนัดวิชานั้นเป็นเกณฑ์ และคะแนนเฉลี่ยสะสมในวิชาเดียวกันนั้นเป็นตัวพยากรณ์โดยลึกลับกว่า วิชาอะไรควรจะใช้คะแนนของวิชาเดียวกันเป็นตัวพยากรณ์และเป็นเกณฑ์ เช่น ใ้คะแนนวิชา คณิตศาสตร์ชั้นมัธยมศึกษาปีที่ 1 ในการพยากรณ์ความสำเร็จในการเรียนคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 2 และ 3 ไม่ใช่วิชาหนึ่งไปพยากรณ์ความสำเร็จอีกวิชาหนึ่ง โดยที่ทั้ง 2 วิชานั้นไม่เกี่ยวข้องกันหรือมีค่างเนื้อหากันซึ่งทำให้ความเชื่อมั่นในการทำงานต่ำ ควยเหตุนี้จึงนิยมใช้แบบสอบความถนัดทางการเรียนไปพยากรณ์ผลสัมฤทธิ์ทางการเรียนมากยิ่งขึ้น

ในปีเดียวกันนี้ แอนนาสเทซี (Anastasi)¹ ได้ศึกษาลักษณะของแบบสอบและรายงานว่ แบบสอบความเร็ว (Speed Test) ซึ่งเป็นข้อสอบค่อนข้างง่ายให้ใช้เวลาดอบน้อยที่สุด ใช้สูตรคูณเคอร์ ริชาร์ดสัน หรือวิธีแบ่งครึ่ง (Split-Half) หากความเที่ยงของแบบสอบไม่ได้เพราะผู้เขาสอบทุกคนไม่มีโอกาสตอบข้อสอบใดครบทุกข้อ ึ่ง ๆ ที่ตอบได้ เนื่องจากเวลาจำกัดควรใช้วิธีทดสอบซ้ำ (Test-Retest) เพื่อหาความเที่ยงของแบบสอบ ส่วนแบบสอบไม่จำกัดเวลา (Power Test) ข้อสอบชนิดนี้ค่อนข้างยากมากกว่าแบบสอบชนิดอาศัยความเร็ว และใช้เวลานานพอที่ทุกคนสามารถตอบได้ทุกข้อ ค่าความเที่ยงของแบบสอบหาได้ทั้งใช้สูตรคูณเคอร์ ริชาร์ดสัน หรือวิธีแบ่งครึ่ง (Split-Half) หรือวิธีทดสอบซ้ำ ดังนั้นการหาความเที่ยงของแบบสอบทั้งสองชนิดจึงค่างกัน การใช้เทคนิคหรือวิธีหาความเที่ยงที่ไม่เหมาะสมจะทำให้ค่าที่ได้ค่างกัน

¹ Anastasi, op.cit., pp. 89-91.

ในปีค.ศ. 1964 บริจแมน (Brigdman)¹ ได้ทำการวิจัยเรื่องความสัมพันธ์ระหว่างการหาค่าอำนาจจำแนก 2 วิธี คือวิธีหาโดยเอาผลต่างของผู้ตอบถูกในกลุ่มสูงกับกลุ่มค่าหารควยจำนวนคนในกลุ่มสูงหรือกลุ่มต่ำ ค่าที่ได้ใช้สัญลักษณ์แทนควย D. อีกวิธีหนึ่งเป็นการหาโดยใช้สัมประสิทธิ์สหสัมพันธ์ระหว่างสองตัวแปร ซึ่งมีการกระจายปกติควยสูตรไบซีเรียล (Biserial Correlation Coefficient) ค่าที่ได้ใช้สัญลักษณ์แทนควย r ผลการวิจัยพบว่าความสัมพันธ์ของการหาอำนาจจำแนกทั้งสองวิธีขึ้นอยู่กับขนาดของกลุ่มตัวอย่าง และระดับความยากของข้อสอบ และมีความสัมพันธ์กันเกือบเป็นเส้นตรง และถ้าใช้เทคนิค 27% แบ่งกลุ่มสูงกลุ่มต่ำและระดับความยากที่ 50% ค่า D. จะมีค่าใกล้เคียงกับค่า r

ในปีค.ศ. 1966 เคอราแกน (Curran) และเพื่อน² ได้สร้างแบบสอบวิชาปรัชญาการศึกษาเพื่อทดสอบกับนักศึกษาระดับปริญญาตรีที่กำลังศึกษาวิชาปรัชญาการศึกษา และนักศึกษาที่เรียนวิชานี้จบแล้วในมหาวิทยาลัยฟลอริดา แบบสอบมีจำนวน 100 ข้อ ผลการวิจัยพบว่ามี 25 ข้อที่ดี และได้รับการคัดเลือกไว้ใช้ต่อไป ค่าสัมประสิทธิ์แห่งความเที่ยงของแบบสอบโดยวิธีทดสอบซ้ำ (Test-Retest Reliability) มีค่าเท่ากับ .82 และค่าสัมประสิทธิ์แห่งความตรงของแบบสอบโดยวิธีหาสหสัมพันธ์แบบพอยท์ ไบซีเรียล (Point Biserial Correlation) มีค่าเท่ากับ .63 โดยการใช้การฝึกทดลองของนักศึกษาเป็นเกณฑ์ คือผู้ที่สำเร็จแล้วถือว่ามี การฝึกมาก ส่วนนักศึกษาที่กำลังเรียนถือว่ามี การฝึกน้อยที่สุด

¹C.S. Brigdman, "The Relation of the Upper-Lower Item Discrimination Index D, to The Bivariate Normal Correlation Coefficient," Educational and Psychological Measurement, (Vol.24, No. 1, 1964), pp. 85-90.

²R.L. Curran, I.J. Gordon and J.F. Doyle, "A Short Test of One's Educational Philosophy," Educational and Psychological Measurement (Vol.26, No.2, 1966), pp. 383-393.

ในปี ค.ศ. 1972 เฮลส์ (Hales)¹ ได้ศึกษาวิชาหาค่าดัชนีอำนาจจำแนก โดยเปรียบเทียบการหา 3 วิธี คือ วิธีที่หนึ่งหาอำนาจจำแนกด้วยการใช้ผลต่างของ ผู้ตอบถูกในแต่ละข้อของกลุ่มสูงกับกลุ่มต่ำ หาค่ายจำนวนคนในกลุ่มสูงหรือกลุ่มต่ำ ค่าที่ได้ใช้สัญลักษณ์แทนด้วย D วิธีที่สองเปิดตารางค่าอำนาจจำแนก (r) ของ ฟลานานาแกน (Flanagan) ซึ่งค่า r นี้ Flanagan ได้ทำขึ้นในปีพ.ศ. 1939 โดยใช้สูตรสหสัมพันธ์ขรรคมคา (Product-Moment Correlation Coefficient) ในการคำนวณค่าที่ได้เรียกว่า Flanagan r วิธีที่สามจาก ค่า Flanagan r ในปีค.ศ. 1949 เดวิส (Davis) ได้นำมาแก้ไขให้ค่า Flanagan r มีความอคติน้อยที่สุด โดยหาความสัมพันธ์กับค่าของฟิชเชอร์ ซี (Fisher z) เรียกค่านี้อีกว่า Flanagan r_c ในการศึกษาใช้กลุ่มตัวอย่าง 3 กลุ่มเป็นนักเรียนเกรด 10, 11 และเกรด 10 และ 11 รวมกัน ผลปรากฏดังนี้ ในนักเรียนแต่ละกลุ่มค่าสัมประสิทธิ์สหสัมพันธ์ของค่า r_c ในแบบสอบสูงกว่าค่า r อย่างมีนัยสำคัญนัยสำคัญ ส่วนค่าสัมประสิทธิ์สหสัมพันธ์ของค่า D อยู่ระหว่างค่า r กับ r_c ความเที่ยงของแบบสอบในนักเรียนแต่ละกลุ่มหาโดยใช้สูตรคูเคอร์ริชาร์ดสันที่ 20 (Kuder - Richardson 20) และแบบแบ่งครึ่ง (Split - Half) มีค่าไม่แตกต่างกัน และอยู่ระหว่าง .88 ถึง .95

¹Loyde W. Hales, "Method of Obtaining the Index of Discrimination For Item Selection and Selected Test Characteristics: A Comparative Study" Educational and Psychological Measurement (Vol.32, No.4, 1972), pp. 929-937.