



เอกสารและรายงานวิจัยที่เกี่ยวข้อง

แบบสอบผลสัมฤทธิ์ เป็นแบบสอบที่ถูกออกแบบเพื่อใช้วัดความรู้และทักษะของผู้เรียนว่า มีการพัฒนาได้สัมพันธ์กับสถานการณ์ของการเรียนการสอนที่ผู้สอนจัดให้ แต่สถานการณ์การเรียน การสอนของแต่ละเนื้อหาย่อมมีขอบเขตของเนื้อหาที่ต้องการวัดในขนาดที่แตกต่างกัน เนื่องจาก มีระยะเวลาการสอนที่แตกต่างกัน เช่นการสอนในแต่ละวัน หรือการสอนในหนึ่งภาคการศึกษา หรือการจัดอบรมในระยะสั้นๆ เป็นต้น ถึงอย่างไรก็ตามผู้สร้างแบบสอบมักจะสร้างขึ้นมาเพื่อที่จะ ใช้วัดว่าอะไรคือสิ่งที่เขาได้เรียนรู้ไปบ้างแล้ว โดยอ้างอิงได้ถึงเวลาเฉพาะจุดใดจุดหนึ่ง หรือ อาจใช้อ้างอิงถึงเวลาในอดีตได้ เนื่องจากความรู้ที่เขามีอยู่นั้นเป็นผลที่เกิดจากประสบการณ์การ เรียนการสอนในอดีตก็ได้ เมื่อต้องการสร้างแบบสอบวัดผลสัมฤทธิ์ ผู้สร้างแบบสอบจะต้องทราบ หลักการและแนวทางการสร้างแบบสอบให้มีคุณภาพ ดังมีรายละเอียดดังต่อไปนี้

หลักการการสร้างแบบสอบเมื่อต้องการวัดผลสัมฤทธิ์ของผู้เรียน

นักวัดผลหลายท่าน ได้เสนอหลักการสร้างแบบสอบเพื่อใช้วัดผลสัมฤทธิ์ของผู้เรียน และ บราวน์ (Brown, F.G. : 1983) ได้เสนอหลักการสร้างแบบสอบไว้ 3 ข้อ ดังต่อไปนี้

1. แบบสอบที่สร้างขึ้นวัดได้อย่างครอบคลุมในขอบเขตเนื้อหา หรือขอบเขตของทักษะ ที่ต้องการวัดหรือไม่ แต่ก็มีข้อจำกัดของเวลาที่ทำเป็นอุปสรรคทำให้ไม่สามารถวัดได้ครอบคลุม ขอบเขตของเนื้อหาที่ต้องการวัดได้

2. แบบสอบต้องวัดในพฤติกรรมการเรียนรู้ที่สำคัญเท่านั้น

3. ผู้สอบต้องมีโอกาสที่จะ ได้ เรียนรู้ในสถานการณ์ต่างๆที่เกี่ยวข้องกับเนื้อหาที่ต้องการ วัดได้อย่างเท่าเทียมกัน เช่น สอน โดยครูคนเดียวกันหรือครูที่มีความสามารถทัดเทียมกัน เนื้อหาใน วิชาต้องเป็นมาตรฐานเดียวกัน มีการใช้สื่อการสอนที่มีคุณภาพทัดเทียมกัน

นอกจากนั้นบราวน์ (Brown, F.G. : 1983) ยังได้นำเสนอหลักในการทดสอบผลสัมฤทธิ์ ของผู้เรียนไว้หลายแนวทางดังนี้

1. แบบสอบที่สร้างขึ้น จะต้องวัดในเนื้อหาที่สอดคล้องกับวัตถุประสงค์ของการเรียน การสอน โดยที่ผู้สอนจะต้องตั้งวัตถุประสงค์ไว้ก่อน จะได้ใช้เป็นแนวทางในการประเมินผู้เรียน ได้ว่ามีความรู้ที่สอดคล้องตามวัตถุประสงค์ที่ตั้งไว้หรือไม่

2. ต้องมีการวิเคราะห์เนื้อหาที่ต้องการให้ผู้เรียนได้เรียนรู้ เช่นมีการทำตาราง กำหนดงานตารางวิเคราะห์หลักสูตร เพื่อจะได้กำหนดว่าเนื้อหาใดที่ผู้เรียนจำเป็นต้องมีความรู้

3. ต้องมีการเลือกใช้แบบสอบที่เหมาะสมกับเนื้อหาที่ต้องการวัด

4. แบบสอบที่สร้างขึ้นมาต้องเหมาะสมกับเป้าหมายของการสอบ เช่น ประเมินความ ก้าวหน้าของผู้เรียน การจัดกลุ่มผู้เรียน

5. แบบสอบจะต้องมีคุณภาพดี ที่สำคัญต้องมีทั้งความเที่ยงและความตรง ต้องมีความระมัดระวังเป็นพิเศษนำไปใช้ตัดสินการสอบได้/ตก ของผู้เรียน

6. ผลของการทดสอบควรมุ่งที่จะนำไปใช้ในการปรับปรุงการเรียนการสอน เพื่อจะได้พัฒนาการเรียนการสอนให้มีคุณภาพดียิ่งขึ้นไป

การพัฒนาแบบสอบผลสัมฤทธิ์ให้มีคุณภาพดี ผู้พัฒนาแบบสอบจะต้องประเมินคุณภาพของแบบสอบโดยให้ความสนใจในลักษณะ 2 ประการ คือ

1. แบบสอบฉบับนั้นวัดได้สอดคล้องกับคุณลักษณะที่ต้องการวัดหรือไม่
2. แบบสอบฉบับนั้นใช้อธิบายคุณลักษณะที่ต้องการวัดได้อย่างชัดเจนถูกต้องหรือไม่

ความสอดคล้อง (Consistency) หมายถึง ความสอดคล้องของคะแนนเมื่อมีการทดสอบซ้ำโดยใช้แบบสอบชุดเดียวกัน คะแนนจากแบบสอบที่มีความเที่ยงจะเป็นคะแนนที่บอกได้ว่ามีความสอดคล้องกับ เวลาและสถานการณ์ของการสอบ

ความตรง (Validity) หมายถึง ความสอดคล้องของคุณลักษณะที่วัดโดยใช้แบบสอบฉบับนั้นกับคุณลักษณะที่แฝงอยู่ในตัวผู้สอบ ได้ดีเพียงใด การประเมินความตรงของแบบสอบจะต้องตอบคำถามที่ว่า.

1. มีคุณลักษณะอะไรบ้างที่ถูกวัดด้วยแบบสอบฉบับนี้
2. แบบสอบที่สร้างขึ้นใช้วัดได้ดีเพียงใด
3. เนื้อหาของแบบสอบเป็นตัวแทนที่ดีของขอบเขตของเนื้อหาที่อธิบายไว้อย่างชัดเจนได้หรือไม่
4. มีพฤติกรรมความรู้ อะไรบ้างที่จะแปลความหมายได้จากคะแนนของแบบสอบฉบับนี้
5. วิธีการทำให้เกิดความตรงของแบบสอบทำได้อย่างถูกต้องเพียงใด

โดยสรุปการหาความตรงของแบบสอบ มีความเกี่ยวข้องกันวิธีการตรวจสอบว่าแบบสอบฉบับนั้นสามารถวัดสิ่งที่ต้องการวัดได้ดีเพียงใด กระบวนการหาความตรง เป็นกระบวนการที่ค้นหาคำตอบของคำถามที่ว่า

1. แบบสอบวัดอะไรได้บ้าง
2. แบบสอบนั้นสามารถใช้วัดได้ตามที่ออกแบบไว้ได้ดีเพียงไร

(Brown, F. G., 1983: 207-210)

คำจำกัดความของความตรง (Validity Defined)

ความตรง (Validity) หมายถึง สัดส่วนของความแปรปรวนของคะแนนที่ได้จากการสอบกับความแปรปรวนของคะแนนที่วัดในสิ่งที่ต้องการวัด จะคำนวณหาค่าสัมประสิทธิ์ความตรงของแบบสอบได้จากสมการดังต่อไปนี้.

$$r^2_{xy} = S^2_v / S^2_x \quad (1)$$

โดยที่ $S^2_v = S^2_T - S^2_i \quad (2)$

เมื่อ S^2_v = ความแปรปรวนของคะแนนที่วัดในสิ่งที่ต้องการวัด
 S^2_i = ความแปรปรวนของคะแนนที่ไม่มี ความตรง

และ

$$S^2_x = S^2_T + S^2_E \quad (3)$$

เมื่อ S^2_x = ความแปรปรวนของคะแนนที่วัดได้จากการสอบ
 S^2_T = ความแปรปรวนของคะแนนจริง
 S^2_E = ความแปรปรวนของความคลาดเคลื่อน

(Brown, F.G., 1983:55-71)

Messick, S. (1989) ได้ศึกษาถึงประวัติและความเป็นมาของความตรงของแบบสอบ และได้ข้อสรุปดังต่อไปนี้.

1946 Guildford กล่าวถึงความตรงของแบบสอบไว้ในลักษณะทั่วไปว่า "แบบสอบจะวัดได้ตรงกับสิ่งที่แบบสอบฉบับนั้นมีความสัมพันธ์ด้วย"

1954-1966 Standard APA ได้มีการแบ่งประเภทของความตรงของแบบสอบตามแนวคิดของ Guildford และตามเป้าหมายของการทดสอบออกเป็น 4 ประเภท คือ Content validity Predictive validity Concurrent validity และ Construct validity

1974 Standard APA แบ่งประเภทของความตรงของแบบสอบ ตามเป้าหมายการทดสอบและการแปลความหมายของคะแนน ได้กล่าวเน้นว่าแบบสอบใดจะมี "Content validity" ได้ จะต้องแสดงให้เห็นว่าแบบสอบฉบับนั้นมีความเป็นตัวแทนของ โดเมน เนื้อหาที่ต้องการวัดได้หรือไม่

1985 Standard APA ให้แนวคิดเกี่ยวกับการแปลความหมายของความตรงของแบบสอบถาม "เป็นการนำเอาคะแนนไปใช้อ้างอิงได้อย่างเฉพาะเจาะจง โดยมีความสอดคล้องกันระหว่าง ความหมายและประโยชน์ ได้อย่างเหมาะสม " มีการแบ่งความตรงตามพยานหลักฐานที่ใช้อ้างอิงออกเป็น 2 ชนิด คือ Content-related validity กับ Criterion-related validity

1954 Anastasi ได้แบ่งประเภทของความตรงออกเป็น Face validity, Content validity, Factorial validity และ Empirical validity

1961 Anastasi แบ่งประเภทความตรงออกเป็น Content validity, Predictive validity, Concurrent validity และ Construct validity ส่วนการตรวจสอบความตรงตามโครงสร้างของแบบสอบถามจะนำตัวแปรอายุมาเกี่ยวข้องด้วย สถิติที่นำมาใช้ในการวิเคราะห์ความตรงตามโครงสร้างคือ Factor analysis และหาค่าสหสัมพันธ์กับตัวแปรอื่น ๆ

1982 Anastasi แบ่งประเภทของความตรงออกเป็น Content validity, Criterion-related validity และ Construct validity เท่านั้น Anastasi มักจะใช้ศัพท์คำว่า "Validation" มากกว่า Validity ความหมายของ Validation คือ "กระบวนการทำให้เกิดความตรงของแบบสอบถาม"

1949 Cronbach แบ่งความตรงออกเป็น Logical validity, Empirical validity และ Content validity ได้เห็นว่าแบบสอบถามผลสัมฤทธิ์ทุกฉบับจะต้องมี Content validity

1984 Cronbach ได้กล่าวสรุปเกี่ยวกับความตรงว่า "เป้าหมายสุดท้ายของกระบวนการทำให้เกิดความตรงคือ Construct validation"

(cited in Linn, R.L., 1989: 18-20)

ผู้สร้างข้อสอบทุกท่านเมื่อต้องการพัฒนาแบบสอบถามให้มีความตรง ท่านจะต้องคำนึงถึงปัจจัยต่างๆที่ส่งผลกระทบต่อความตรงของแบบสอบถาม ทำให้ข้อสอบที่ท่านสร้างขึ้นนั้นมีความสามารถวัดความรู้ของผู้สอบได้อย่างถูกต้อง

Mehrens, W.A. and Lehmann, I.J. (1984) ได้กล่าวถึงปัจจัยที่มีผลต่อความตรงของแบบสอบถาม (Factors Affecting Validity) มีดังต่อไปนี้

- (1) แบบสอบถามที่สร้างขึ้นใช้วัดตัวแปรที่ต้องการวัดหรือไม่
- (2) แบบสอบถามที่สร้างขึ้นนั้นมีความยากง่ายเหมาะสมสำหรับผู้สอบหรือไม่
- (3) ค่าความเที่ยงของแบบสอบถามที่เป็นตัวทำนาย และแบบสอบถามที่เป็นตัวเกณฑ์ ก็ย่อมส่ง

ผลกระทบต่อความตรงของแบบสอบ ดังปรากฏในสมการ

$$r_{xy} < \sqrt{r_{xx} \times r_{yy}}$$

เมื่อ r_{xy} = ค่าสัมประสิทธิ์ความตรงแบบสอบ
 r_{xx} = ค่าสัมประสิทธิ์ความเที่ยงของแบบสอบที่เป็นตัวทำนาย
 r_{yy} = ค่าสัมประสิทธิ์ความเที่ยงของแบบสอบที่เป็นตัวเกณฑ์

(4) ความเป็นวิวิธพันธ์ของกลุ่มผู้สอบ (Heterogeneous) จะส่งผลทำให้ค่าสัมประสิทธิ์ความตรงของแบบสอบสูงไปด้วย ถ้ากลุ่มผู้สอบมีความเป็นเอกพันธ์ (Homogeneous) จะทำให้คะแนนของผู้สอบเกาะกลุ่มกันมากจะส่งผลให้ค่าสัมประสิทธิ์ความตรงของแบบสอบต่ำไปด้วย

Gronlund, N.E. (1971:98-102) ได้กล่าวถึงปัจจัยที่มีผลต่อความตรงของแบบสอบ ได้ดังมีรายละเอียดต่อไปนี้

(1) ปัจจัยภายในแบบสอบ ได้แก่

- 1.1 คำชี้แจงของแบบสอบ ถ้าไม่ชัดเจนอาจทำให้ผู้สอบทำข้อสอบผิดพลาดได้
- 1.2 การใช้ประโยคที่ยากเกินไป จะทำให้ผู้สอบเกิดความสับสน จะทำให้กลายเป็นการวัดความสามารถในการอ่านของผู้สอบ มากกว่าการวัดความรู้ของผู้สอบ
- 1.3 ระดับความยากของข้อสอบไม่เหมาะสม ข้อสอบที่ยากหรือง่ายเกินไปทำให้ไม่สามารถจำแนกผู้สอบได้
- 1.4 การสร้างข้อสอบที่ไม่รัดกุม ทำให้ข้อสอบบางข้อชี้แนะคำตอบที่ถูกต้องของข้ออื่นๆได้
- 1.5 การใช้ภาษาที่กำกวมของข้อสอบทำให้ผู้สอบตีความผิดได้
- 1.6 แบบสอบที่ใช้ไม่เหมาะสมกับพฤติกรรมที่ต้องการวัด
- 1.7 แบบสอบสั้นเกินไป ทำให้ตั้งคำถามได้ไม่ครอบคลุมวัตถุประสงค์ที่ต้องการวัด
- 1.8 การจัดเรียงข้อสอบไม่เหมาะสม ข้อสอบควรเรียงจากง่ายไปหายาก เพราะถ้าเรียงจากข้อยากก่อนผู้สอบจะใช้เวลาคิดนาน จนไม่มีเวลาทำข้อสอบที่ง่ายซึ่งอยู่ตอนท้ายของแบบสอบ ทำให้ไม่สามารถวัดความสามารถที่แท้จริงของผู้สอบได้
- 1.9 คำตอบที่เรียงเป็นรูปแบบเดียวกันทำให้ผู้สอบเดาคำตอบที่ถูกต้อง

(2) การบริหารแบบสอบและการให้คะแนน จะส่งผลให้ความตรงของแบบสอบเปลี่ยนแปลงได้ เช่น การไม่ดำเนินการสอบให้เป็นมาตรฐานเดียวกันตามคำชี้แจง การตรวจให้คะแนนไม่มีความเป็นปรนัยการแนะนำคำตอบให้กับผู้สอบ ฯลฯ

(3) ปัจจัยภายในตัวผู้สอบ ได้แก่ ความวิตกกังวล ความเครียด ทำให้ผู้สอบทำข้อสอบได้ไม่เต็มความสามารถ การเดาคำตอบของผู้สอบก็ทำให้คะแนนที่ได้ไม่ตรงกับความสามารถที่แท้จริงของผู้สอบย่อมส่งผลถึงความตรงของแบบสอบ

(4) ลักษณะของกลุ่มอ้างอิง และเกณฑ์ที่ใช้ในการตัดสิน โดยทั่วไปแล้วความตรงของแบบสอบควรคำนึงถึงลักษณะเฉพาะของกลุ่มผู้สอบด้วย เพราะสิ่งที่วัดได้จากผู้สอบมักจะได้รับอิทธิพลจากตัวแปรต่างๆ เช่น เพศ อายุ ระดับความสามารถ ภูมิหลังทางการศึกษา และวัฒนธรรมของผู้สอบ

Messick, S. (1989) ได้เน้นว่าแหล่งข้อมูลเกี่ยวกับความตรงของแบบสอบนั้นได้จากคะแนนมากกว่าแบบสอบ ดังนั้นความตรงจึงเป็นคุณสมบัติของคะแนนมากกว่าแบบสอบ เนื่องจากคะแนนได้จากคำตอบของผู้สอบและทำให้ทราบว่าผู้สอบรู้อะไรบ้าง หรือมีความรู้อะไรบ้าง การประเมินความตรง (Validity) ของแบบสอบประเภทต่าง ๆ นั้น จะต้องพิจารณาถึงความสอดคล้องกันระหว่างพยานหลักฐานเชิงประจักษ์ (Evidences) คือคะแนนจากแบบสอบ กับเกณฑ์ภายนอก ว่าวัดได้สอดคล้องกับความเป็นจริง เช่น สอดคล้องกับคุณลักษณะตามทฤษฎี หรือสอดคล้องกับพฤติกรรมอื่นๆ หรือสอดคล้องกับเนื้อหาที่ต้องการวัดได้หรือไม่ เพื่อให้มีการแปลความหมายของคะแนนที่ได้จากแบบสอบได้อย่างถูกต้องและน่าเชื่อถือ ส่งผลให้เกิดประโยชน์ เมื่อนำคะแนนไปใช้ประกอบในการตัดสินใจในการประเมินผู้สอบตามเป้าหมายที่ต้องการ เช่น การคัดเลือกบุคคล การวัดความสามารถของผู้เรียน การมอบหมายตำแหน่ง (Messick, S.; cited in Linn, R.L., 1989:13)

การแบ่งประเภทของความตรง

การแบ่งประเภทของความตรงของสำนักงานมาตรฐานการทดสอบและจิตวิทยา (APA, 1966; cited in Linn, R.L., 1989:16-18) ดังต่อไปนี้

1. Criterion-related Validity จะถูกประเมินโดยการหาความสัมพันธ์ระหว่างคะแนนจากแบบสอบกับตัวแปรภายนอกที่เป็นเกณฑ์ แบ่งย่อยออกเป็น 2 ประเภทคือ

1.1 Predictive Validity เป็นประเภทหนึ่งของความตรงตามเกณฑ์เป็นการแปลความหมายของคะแนนเพื่อใช้ในการทำนายความสามารถในอนาคตของผู้สอบ

1.2 Concurrent Validity เป็นประเภทหนึ่งของความตรงตามเกณฑ์เช่นกัน เป็นการแปลความหมายของคะแนน ว่าผู้สอบมีความสามารถอยู่ในระดับใดเมื่อนำไปเปรียบเทียบกับเกณฑ์

2. Construct Validity เป็นการประเมินว่าแบบสอบฉบับนั้นมีคุณภาพหรือไม่สามารถใช้วัดคุณลักษณะของผู้สอบตามโครงสร้างของทฤษฎีหรือไม่ และมีการรายงานผลการประเมินโครงสร้างของแบบสอบฉบับนั้น

3. **Content Validity** จะประเมินเนื้อหาของแบบสอบถามว่ามีระบบว่าเนื้อหาของแบบสอบถามนั้นถูกต้องมาจากสถานการณ์ใดของวิชา หรือถูกต้องมาจากเนื้อหาสาระใดของวิชา

ความตรงตามเกณฑ์ (Criterion-related Validity) เป็นความสามารถของแบบสอบถามที่จะใช้ทำนายพฤติกรรมของผู้สอบตามเกณฑ์ที่ผู้สร้างแบบสอบถามต้องการ จะคำนวณหาค่าความตรงประเภทนี้จากความสัมพันธ์ระหว่างคะแนนของแบบสอบ กับคะแนนที่วัดได้ตามเกณฑ์ภายนอก ความตรงตามเกณฑ์ของแบบสอบถามแบ่งได้เป็น 2 ชนิดคือ

(1) **ความตรงเชิงทำนาย (Predictive Validity)** หมายถึงความสามารถของแบบสอบถามที่จะใช้บ่งบอกถึงผลของสิ่งที่วัดในขณะนั้น ได้อย่างถูกต้องตามสภาพที่แท้จริงในอนาคต โดยอาศัยการแปลผลจากความสัมพันธ์ระหว่างคะแนนที่ได้จากแบบสอบถามนั้น กับคะแนนที่วัดตามเกณฑ์ที่เกิดขึ้นในอนาคต เช่น ความสัมพันธ์ของคะแนนที่ได้จากแบบสอบความถนัดทางการพยาบาล กับคะแนนเฉลี่ยสะสมปลายปีของนักศึกษาพยาบาล การคำนวณหาความตรงประเภทนี้ต้องใช้เวลาในการรอคอย จะทำให้บอกได้ว่าแบบสอบถามนั้นมีความตรงเชิงทำนายเป็นที่ยอมรับได้หรือไม่

(2) **ความตรงตามสภาพ (Concurrent Validity)** หมายถึงความสามารถของแบบสอบถามที่จะใช้บ่งบอกถึงสิ่งที่วัดได้อย่างถูกต้องตามสภาพที่แท้จริงในปัจจุบัน โดยอาศัยการแปลผลจากความสัมพันธ์ระหว่างคะแนนที่ได้จากแบบสอบถามนั้น กับคะแนนที่วัดจากเกณฑ์ที่เกิดขึ้นในเวลาใกล้เคียงกัน เช่น ความสัมพันธ์ของคะแนนที่ได้จากแบบสอบวัดทักษะการพยาบาลปีที่1 กับคะแนนที่ได้จากการประเมินผลการปฏิบัติพยาบาล จะทำให้บอกได้ว่าแบบสอบถามนั้นมีความตรงตามสภาพหรือไม่

ความตรงเชิงทำนายกับความตรงตามสภาพต่างกัน 2 ประเด็นคือ

1. ช่วงระยะเวลาของการเก็บข้อมูลที่ใช้เป็นเกณฑ์ เกณฑ์ของความตรงเชิงทำนายจะเกิดขึ้นภายหลังจากการเก็บคะแนนจากแบบสอบ ส่วนเกณฑ์ของความตรงตามสภาพจะเก็บข้อมูลพร้อมกันกับคะแนนจากแบบสอบ

2. ความมุ่งหมายของการทดสอบ ความตรงเชิงทำนายเป็นการนำเอาคะแนนที่ได้จากการทดสอบมาใช้ทำนายความสามารถหรือพฤติกรรมในอนาคตของผู้สอบ ส่วนความตรงตามสภาพเป็นการนำเอาคะแนนที่ได้จากการทดสอบมาเปรียบเทียบกับความสามารถในขณะนั้นของผู้สอบ

ในการศึกษาความตรงตามเกณฑ์ สิ่งที่สำคัญที่สุดคือคะแนนเกณฑ์ที่วัดมานั้นจะต้องมีคุณภาพที่สำคัญที่จะต้องพิจารณาจากสิ่งต่อไปนี้ (Thorndike, R. M., Cunningham, G. K., Thorndike, R. L. and Hagen, E. P., 1991:129-130)

1. มีความเกี่ยวข้องกันหรือไม่ (Relevance) การพิจารณาความเกี่ยวข้องของเกณฑ์โดยทั่วไปใช้วิธีการตัดสินคุณค่า มักจะไม่ได้วัดออกมาเป็นตัวเลข

2. มีความเที่ยงหรือไม่ (Reliability) เครื่องมือที่ใช้วัดคะแนนเกณฑ์จะต้องมี

ความคงที่ในการวัด ไม่ว่าจะใช้แบบวัดอื่นใดก็ตาม โดยที่แบบวัดฉบับเหล่านั้นสร้างขึ้นมาจากเพื่อใช้วัด
เกณฑ์เดียวกัน และตัวแปรเกณฑ์ที่วัดได้จะไม่เปลี่ยนแปลง

3.ความเป็นอิสระจากอคติใดทั้งสิ้น (Free from bias) จะต้องไม่ได้รับผล
กระทบจากอิทธิพลใดๆ เช่น คะแนนที่ได้จากการทำแบบสอบวัดทักษะทางการพยาบาล ทำให้ครูมี
ใจลำเอียงประเมินผลการปฏิบัติงานของนักศึกษาพยาบาลไปในทิศทางที่สูงหรือต่ำกว่าความ
สามารถที่แท้จริงของนักศึกษา

4.มีความเป็นไปได้ (Availability) ในการวัดตัวแปรเกณฑ์ต้องนึกถึงความ
เป็นไปได้ และมีความสะดวกแค่ไหน เช่น ต้องใช้ระยะเวลาในการรอคอยเพื่อวัดตัวแปรเกณฑ์ที่
ยาวนานแค่ไหน เพื่อที่จะไม่เกิดความเสียหายแก่งานวิจัย ต้องเสียค่าใช้จ่ายมากเกินไปหรือไม่
ในการวัดตัวแปรเกณฑ์ และจะมีความสะดวกในการเก็บรวบรวมข้อมูลของตัวแปรเกณฑ์หรือไม่
การประมาณค่าความตรงตามเกณฑ์ สามารถวิเคราะห์ได้จากการหาค่าสัมประสิทธิ์
สหสัมพันธ์ของคะแนนที่วัดได้จากแบบสอบ กับ คะแนนเกณฑ์ สูตรที่ใช้ในการคำนวณเป็นสูตร
คำนวณหาค่าสัมประสิทธิ์สหสัมพันธ์แบบต่าง ๆ นั้นเอง แต่ที่นิยมใช้กันมากที่สุดคือสูตรสัมประสิทธิ์
สหสัมพันธ์แบบเพียร์สัน

ความตรงตามโครงสร้าง (Construct Validity) เป็นความสามารถของแบบสอบใช้วัด
คุณลักษณะที่แฝงเร้นอยู่ในตัวผู้สอบนั้นว่า มีคุณลักษณะตามที่ทฤษฎีที่อธิบายไว้ หรือตามภาวะ
สันนิษฐานที่ตั้งไว้ สามารถหาความตรงตามโครงสร้าง โดยอาศัยข้อมูลพื้นฐานที่เป็นข้อมูล
เชิงประจักษ์ และต้องทำการพัฒนาเป็น 3 ระยะ คือ

(1) เป็นการนำเอาทฤษฎีมาอธิบายสนับสนุนโครงสร้างของแบบสอบ (Describing
of Theory) เป็นวิธีการที่มีความเป็นรูปธรรมน้อย จากทฤษฎีจะบอกให้ทราบว่า มีโดเมนใดบ้าง
เกี่ยวข้องกับโครงสร้างที่ต้องการวัด มีโดเมนใดบ้างที่ไม่เกี่ยวข้องควรตัดทิ้งไป

(2) ต้องทำการตรวจสอบว่าแบบสอบนั้นวัดได้มิติเดียวหรือไม่ (Unidimensionality)
จะทำให้ทราบว่าแบบสอบฉบับนั้นถูกพัฒนาขึ้นมาเพื่อวัดเพียงโครงสร้างเดียว หรือวัดเพียงคุณลักษณะ
เดียว (One Trait) เท่านั้น มักจะใช้การวิเคราะห์ตัวประกอบช่วยตรวจสอบ

(3) การวิเคราะห์ข้อมูลเชิงประจักษ์ (Empirical Analysis) เป็นขั้นของการ
ตรวจสอบกับหลักฐานเชิงประจักษ์ว่าแบบสอบใช้วัดโครงสร้างนั้นได้จริง แต่เนื่องจากการศึกษา
จากประชากรที่เป็นจริงต้องใช้เวลามาก และเสียค่าใช้จ่ายสูง นักวัดผลจึงศึกษาจากกลุ่มตัวอย่าง
เท่านั้น

(Brown, F.G., 1983:55-71)

วิธีการคำนวณหาความตรงตามโครงสร้างมีหลายวิธีดังต่อไปนี้

ก. วิธีการคุณลักษณะหลากหลาย-วิธีหลาย (Multitrait-multimethod) เป็นวิธีการที่ใช้ตรวจสอบความตรงเมื่อมีคุณลักษณะที่วัดตั้งแต่สองคุณลักษณะขึ้นไป และถูกวัดโดยวิธีการตั้งแต่สองวิธีขึ้นไป จะใช้การคำนวณในรูปแบบของค่าสัมประสิทธิ์สหสัมพันธ์มีอยู่ 2 ชนิด คือ

1. ความตรงเชิงลู่เข้า (Convergent Validity) เป็นความตรงที่ทราบได้จากค่าสัมประสิทธิ์สหสัมพันธ์ที่มีค่าสูง เมื่อใช้เครื่องมือวัดคุณลักษณะเดียวกันแต่มีวิธีการที่แตกต่างกัน

2. ความตรงเชิงจำแนก (Discriminant Validity) เป็นความตรงที่ทราบได้จากค่าสัมประสิทธิ์สหสัมพันธ์ที่มีค่าต่ำ เมื่อใช้เครื่องมือวัดคุณลักษณะที่แตกต่างกัน แต่ใช้วิธีการเดียวกัน

ข. วิธีการวิเคราะห์ตัวประกอบ (Factor Analysis) ต้องพิจารณาจากผลการวิเคราะห์ตัวประกอบ เป็นการใช้เทคนิคทางสถิติที่มีจุดมุ่งหมายคือต้องการแสดงความสัมพันธ์ของคุณลักษณะที่ต้องการวัดให้ปรากฏออกมาโดยใช้วิธีวิเคราะห์ทางสถิติ มีวิธีดำเนินการดังต่อไปนี้

(1) ศึกษาทฤษฎีที่เกี่ยวข้องกับคุณลักษณะที่ต้องการวัดให้ชัดเจน จนได้แนวคิดว่าคุณลักษณะนั้นจะประกอบด้วยตัวประกอบอะไรบ้าง ตัวอย่างเช่น เซอวปัญญาจะประกอบด้วยความสามารถทางด้านภาษา ความสามารถทางด้านตัวเลข ความสามารถทางด้านนามธรรม

(2) สร้างข้อสอบเพื่อใช้วัดตัวประกอบแต่ละตัวที่ศึกษาได้ตามทฤษฎีจนครบทุกตัวประกอบ

(3) นำแบบสอบที่สร้างขึ้นเพื่อใช้วัดตัวประกอบเหล่านั้นกับกลุ่มทดลองที่มีขนาดเหมาะสม

(4) นำข้อมูลที่ได้จากกลุ่มทดลอง มาวิเคราะห์ตัวประกอบ โดยใช้เทคนิคทางสถิติ แล้วพบว่าข้อสอบที่สร้างขึ้นมานั้น เกาะกลุ่มกันตามตัวประกอบที่ต้องการวัดหรือไม่ และเมื่อวิเคราะห์รายข้อเพื่อตรวจดูว่า ใช้วัดตัวประกอบเดียวกันหรือไม่ จะทำให้บอกได้ว่าแบบสอบฉบับนั้นมีความตรงตามโครงสร้าง

ความตรงเชิงเนื้อหา (Content Validity) เป็นความสามารถของแบบสอบที่ใช้วัดในเนื้อหาที่ต้องการวัดหรือไม่ โดยมีผู้เชี่ยวชาญเกี่ยวกับเนื้อหาเป็นผู้พิจารณาตัดสินแบบสอบฉบับนั้น ความหมายของความตรงเชิงเนื้อหานั้นมีการเสนอไว้หลายมุมมอง ฟิทซ์แพทริก (Fitzpatrick :1983) จึงได้เสนอแนวคิดเกี่ยวกับความหมายของความตรงเชิงเนื้อหาของแบบสอบ ไว้ 4 แนวคิดดังนี้ คือ

(1) กรอบแนวคิดได้จากการสุ่มโดเมนเนื้อหาที่ต้องการวัด

(The Concept of Domain Sampling)

(2) กรอบแนวคิดที่ว่ามีความเกี่ยวข้องกับโดเมนเนื้อหาที่ต้องการวัด

(The Concept of Domain of Relevance)

(3) กรอบแนวคิดที่ว่ามีการอธิบายโดเมนเนื้อหาที่ต้องการวัดได้อย่างแจ่มชัด

(The Concept of Domain Clarity)

(4) กรอบแนวคิดของการสร้างข้อสอบที่ใช้เทคนิคที่มีคุณภาพ
(The Concept of Technical Quality Items)

กรอบแนวคิดได้จากการสุ่มโดเมนเนื้อหาที่ต้องการวัด (The Concept of Domain Sampling)

มีการอธิบายความหมายของความตรงเชิงเนื้อหาของแบบสอบว่า แบบสอบที่มีความตรงเชิงเนื้อหาจะเป็นแบบสอบที่เป็นตัวแทนของโดเมนเนื้อหาที่ต้องการวัดได้หรือไม่ เพื่อให้เกิดความเข้าใจชัดเจนขึ้น ได้เสนอแนวคิดเกี่ยวกับการสุ่มโดเมนเนื้อหาที่ต้องการวัดไว้ 2 ลักษณะดังนี้คือ

ก. แบบสอบเป็นดังตัวอย่างของเนื้อหา (The Test as a Content Sample)

โดยจะพิจารณาจาก

(1) ข้อสอบแต่ละข้อในแบบสอบต้องมีความสอดคล้องกับเนื้อหาสาระที่มีการอธิบายไว้ในโดเมนเนื้อหา หรือกระทบต่อโดเมนเนื้อหาได้ดีเพียงใด

(2) แบบสอบจะต้องมีความเป็นตัวแทนของโดเมนเนื้อหาที่ต้องการวัดได้หรือไม่ โดยมีผู้เชี่ยวชาญเป็นผู้พิจารณาตัดสินว่ามีความเป็นตัวแทนที่ดีเพียงใด

ข. แบบสอบเป็นตัวอย่างของพฤติกรรมความรู้ (The Test as a Behavioral Sample) ในมุมมองนี้จะมองความตรงเชิงเนื้อหาของแบบสอบว่า คำตอบในแบบสอบเป็นดังตัวอย่างของพฤติกรรมความรู้ที่ต้องการวัด ในแง่นี้จะให้แนวคิดเกี่ยวกับการแปลความหมายของคำตอบในแบบสอบ โดยจะบอกให้ทราบว่าผู้สอบมีพฤติกรรมความรู้อะไรบ้าง ได้มีการอ้างถึงความหมายของความตรงเชิงเนื้อหาของสำนักงานมาตรฐานการทดสอบในสหรัฐอเมริกา (Standard, APA, 1974) ว่าหมายถึง "การสาธิตความตรงเชิงเนื้อหาของชุดคะแนนที่ได้จากแบบสอบ สิ่งหนึ่งที่ต้องสาธิตคือ ตัวอย่างของพฤติกรรมความรู้ในขอบเขตของเนื้อหาที่ต้องการวัด โดยการจัดเป็นสถานการณ์ทดสอบ" การตัดสินความเป็นตัวอย่างของพฤติกรรมความรู้มักจะทำการพิจารณาโดยผู้เชี่ยวชาญ

กรอบแนวคิดที่ว่ามี ความเกี่ยวข้องกับ โดเมนเนื้อหาที่ต้องการวัด

(The Concept of Domain of Relevance)

มีประเด็นที่เป็นสาระที่ควรกล่าวถึงดังต่อไปนี้

ก. ความเกี่ยวข้องของเนื้อหาของแบบสอบ (The Relevance of Test Content)

เนื้อหาในแบบสอบจะต้องครอบคลุมสาระสำคัญของจักรวาลเนื้อหา ตามเป้าหมายที่ผู้สร้างแบบสอบได้กำหนดไว้วิธีการพิจารณาตัดสินความเกี่ยวข้องของเนื้อหา มักจะกระทำโดยผู้เชี่ยวชาญเช่นกัน ผู้เชี่ยวชาญเนื้อหาอาจเป็นอาจารย์ผู้สอนที่มีประสบการณ์ อาจารย์ผู้จัดหลักสูตร ผู้แต่งตำรา

ข. ความเกี่ยวข้องของคำตอบในแบบสอบ (The Relevance of Test Responses) มองในแง่ที่ว่าคำตอบที่ได้จากแบบสอบ เป็นพฤติกรรมความรู้ที่ต้องถูกพิจารณาว่ามีความเกี่ยวข้องกับเนื้อหาที่ต้องการวัดมากน้อยเพียงใด

กรอบแนวคิดที่ว่าการอธิบายโดเมนเนื้อหาที่ต้องการวัดได้อย่างแจ่มชัด

(The Concept of Domain Clarity)

ต้องมีการจัดกระทำดังต่อไปนี้

ก. ต้องทำการพิจารณาเกี่ยวกับการจำกัดความของเทอม หรือตัวแปรที่ต้องการวัดว่าชัดเจนแค่ไหน มีการบอกวิธีการเก็บรวบรวมข้อมูล / คำพิจารณาตัดสินเรื่องความแจ่มชัดของเนื้อหาอย่างเป็นระบบ มีการจัดสร้างตารางวิเคราะห์ข้อสอบ (Items Specified) สร้างรูปแบบของข้อสอบ (Item Format)

ข. ต้องแสดงให้เห็นถึงความตรงเชิงเนื้อหาของแบบสอบอย่างชัดเจนที่สุด คือมีการอธิบายสาระสำคัญของเนื้อหาที่ต้องการวัด ที่นำไปสู่การปฏิบัติที่เป็นขั้นตอนได้

ค. คำจำกัดความนั้นต้องมีการอธิบายถึงรายละเอียดของลักษณะเนื้อหา และต้องบอกแนวทางของการตรวจให้คะแนนข้อสอบที่จะสร้างขึ้นด้วย

กรอบแนวคิดของการสร้างข้อสอบที่ใช้เทคนิคที่มีคุณภาพ

(The Concept of Technical Quality Items)

ผู้สร้างแบบสอบต้องทำการศึกษาเทคนิคในการสร้างข้อสอบที่มีคุณภาพ โดยเน้นการใช้กระบวนการเชิงประจักษ์ และมีความสมเหตุสมผล ระหว่างการสร้างข้อสอบควรมีผู้เชี่ยวชาญในการสร้างข้อสอบให้คำแนะนำ และต้องคำนึงถึงเสมอว่าความตรงเชิงเนื้อหาของแบบสอบเป็นสิ่งจำเป็น เทคนิคการสร้างข้อสอบที่ผู้เขียนได้เสนอแนะว่า จะต้องมีการวิเคราะห์เนื้อหาของแบบสอบและเนื้อหาของคำตอบจากแบบสอบฉบับนั้นด้วย

ลอว์ซี (Lawshe) ก็ให้คำจำกัดความของความตรงตามเนื้อหาเชิงปฏิบัติการ (An Operative Definition) ไว้เช่นกันว่า หมายถึง ความสอดคล้องกันระหว่างเนื้อหาของแบบสอบกับขอบเขตเนื้อหาที่ต้องการวัด ซึ่งได้มาจากการสรุปผลการพิจารณาตัดสินความสอดคล้อง โดยคณะผู้เชี่ยวชาญ

Allen, M. J. and Yen, W. M. (1979) ได้แบ่งประเภทของความตรงเชิงเนื้อหาของแบบสอบออกเป็น 2 ประเภทดังต่อไปนี้

(1) Face Validity บางครั้งอาจเรียกชื่อว่า "Armchair Validity" ความตรงเชิงเนื้อหาประเภทนี้มักจะเป็นวิธีการตรวจสอบที่รวบรัด แบบสอบจะถูกตรวจสอบโดยผู้เชี่ยวชาญแล้วให้ข้อสรุปว่าแบบสอบฉบับนั้นเมื่อมองดูโดยรูปภายนอกแล้วใช้วัดในเนื้อหาที่ต้องการวัดได้ โดยไม่ได้บอกเกณฑ์ที่ใช้ประกอบการพิจารณาตัดสิน เช่นนั้นของผู้เชี่ยวชาญ ดังนั้น Face Validity จึงไม่ได้รับความเชื่อถือเท่าใดนัก

(2) Logical Validity หรือ Sampling Validity วิธีการตรวจสอบความตรงเชิงเนื้อหาประเภทนี้ก็จะได้รับการตรวจสอบโดยผู้เชี่ยวชาญเช่นกัน ผู้เชี่ยวชาญจะต้องพิจารณาตัดสินถึงความสมเหตุสมผลของเนื้อหาของแบบสอบกับตารางวิเคราะห์เนื้อหา หรือพิจารณาถึงความสอดคล้องระหว่างเนื้อหาของแบบสอบ กับ เนื้อหาที่กำหนดไว้ในตารางวิเคราะห์เนื้อหาที่ต้องการวัด ดังนั้นจึงมีความซับซ้อนมากกว่า Face Validity ความตรงประเภทนี้จึงมีความสำคัญมากต่อการสร้างแบบสอบวัดผลสัมฤทธิ์ของผู้เรียน

ผู้สร้างข้อสอบทุกท่านเมื่อได้ทราบความหมายของความตรงเชิงเนื้อหาของแบบสอบแล้ว ควรจะได้ทราบเงื่อนไขของวิธีการตรวจสอบความตรงเชิงเนื้อหา และขั้นตอนของการปฏิบัติเมื่อต้องการตรวจสอบความตรงเชิงเนื้อหาของแบบสอบ ดังมีรายละเอียดต่อไปนี้.

วิธีการตรวจสอบความตรงเชิงเนื้อหาของแบบสอบ Guion (1977) ได้เสนอเงื่อนไข 5 ประการของการตรวจสอบความตรงเชิงเนื้อหา ต้องจัดกระทำดังต่อไปนี้

1. ขอบเขตของเนื้อหาที่ต้องการวัดจะต้องมีการอธิบายไว้อย่างชัดเจน
2. ขอบเขตของเนื้อหาที่ต้องการวัดต้องมีความเกี่ยวข้องกับเป้าหมายของการทดสอบ
3. การอธิบายความหมายของขอบเขตเนื้อหาที่ต้องการวัดต้องเป็นที่ยอมรับของบุคคลที่เกี่ยวข้อง
4. ต้องเป็นการตัดสินว่าเนื้อหาของแบบสอบนั้นมีความเป็นตัวแทนที่ดีของขอบเขตเนื้อหาที่ต้องการวัดได้หรือไม่
5. คำพิจารณาตัดสินของผู้เชี่ยวชาญเป็นสิ่งที่มีความน่าเชื่อถือมากน้อยเพียงใด

(cited in Brown, F.G., 1983:55-71)

Murphy, K.R. และ Davidshofor, C.O. (1988) กล่าวว่า การตัดสินความตรงเชิงเนื้อหาของแบบสอบนั้น จะมีความน่าเชื่อถือหรือไม่ ขึ้นกับความชัดเจนของการบรรยายขอบเขตเนื้อหา(Content Domain) นอกจากนั้นเขายังได้เสนอกระบวนการขั้นพื้นฐานของการประเมินความตรงเชิงเนื้อหาของแบบสอบ ว่าประกอบด้วย 3 ขั้นตอนดังนี้

1. การบรรยายโดเมนเนื้อหา
 2. กำหนดความสำคัญของ โดเมนเนื้อหาที่ต้องการวัด
 3. การเปรียบเทียบโครงสร้างของแบบสอบ (Structure) กับโดเมนเนื้อหา
- ตัวอย่างที่แสดงไว้ในตารางที่ 1 จะบอกรายละเอียดของการบรรยายโดเมนเนื้อหาบอกส่วนต่าง ๆ ที่ต้องมีอยู่ในโดเมน และการกำหนดน้ำหนักในแต่ละส่วนของ โดเมนที่ต้องการวัดเป็นเปอร์เซ็นต์

ตารางที่ 1 ตัวอย่างของการบรรยายขอบเขตเนื้อหา

1. โดเมนที่ถูกต้อง : ความรู้เกี่ยวกับสังคมศาสตร์โลกที่ครอบคลุมในหลักสูตรมาตรฐานของนักเรียนชั้นประถมศึกษาตอนปลาย

2. ส่วนต่าง ๆ ที่มีอยู่ในโดเมน :

(ก) ปัญหาเกี่ยวกับ	(ข) ในเขตพื้นที่ของ	(ค) ช่วงเวลา
1. สังคม	1. ยุโรป	1. ศตวรรษที่ 18
2. การเมือง	2. อเมริกา	2. ศตวรรษที่ 19
3. วัฒนธรรม	3. แอฟริกาและเอเชีย	

3. น้ำหนักของแต่ละส่วนของเนื้อหาที่ต้องการให้แบบสอบถามครอบคลุม : (คิดเป็นเปอร์เซ็นต์)

		<u>สังคม</u>	<u>การเมือง</u>	<u>วัฒนธรรม</u>
ยุโรป	ศตวรรษที่ 18	5	10	3
	ศตวรรษที่ 19	5	8	2
อเมริกา	ศตวรรษที่ 18	6	17	2
	ศตวรรษที่ 19	9	13	5
แอฟริกาและเอเชีย	ศตวรรษที่ 18	2	0	0
	ศตวรรษที่ 19	6	5	2

ค่าเปอร์เซ็นต์ที่แสดงไว้ในตารางที่ 1 จะเป็นตัวกำหนดชั่วโมงการสอน ชั่วโมงการอ่านหนังสือและจำนวนข้อสอบ ของแต่ละเนื้อหาที่ต้องการวัด

อุทฺตมพร จามรฺมาน (2535:78-80) ได้เสนอวิธีการตรวจสอบความตรงเชิงเนื้อหาของแบบสอบซึ่งเป็นวิธีการวิเคราะห์สัดส่วนความสอดคล้องของผู้เชี่ยวชาญ (Accordance Proportion) มีแนวทางปฏิบัติโดยสรุปดังต่อไปนี้

1. สร้างตารางโครงสร้างเพื่อแสดงให้เห็นว่าแบบสอบชุดนี้ สร้างครอบคลุมเนื้อหาอย่างครบถ้วนหรือไม่ ถ้าไม่แน่ใจในคุณภาพของตารางโครงสร้าง ให้เชิญผู้ทรงคุณวุฒิด้านเนื้อหา มาพิจารณาตารางโครงสร้าง โดยขอให้ผู้ทรงคุณวุฒิตอบคำถามต่อไปนี้

ก. การจำแนกเนื้อหาออกเป็นส่วน ๆ ได้อย่างถูกต้อง สมบูรณ์ครบถ้วน ตามหลักหรือไม่

ข. การระบุน้ำหนักของเนื้อหาที่จำเป็นต่อวัดมีความถูกต้องเหมาะสมหรือไม่

ค. ข้อความของข้อสอบแต่ละข้อสอดคล้องกับเนื้อหาและวัตถุประสงค์ที่จะวัดหรือไม่

2. กำหนดเกณฑ์ภายนอก เพื่อให้ผู้เชี่ยวชาญใช้ในการพิจารณาตัดสินเนื้อหาของข้อสอบที่ละข้อ คือ

ก. ความชัดเจน ถูกต้อง และครอบคลุมในเนื้อหาสาระ

ข. ความชัดเจน ถูกต้อง และครอบคลุมในจุดมุ่งหมายการทดสอบ

3. กลุ่มผู้เชี่ยวชาญต้องตรวจสอบข้อสอบที่ละข้อ ให้มีความสอดคล้องกับ ตารางเนื้อหา-จุดมุ่งหมายการทดสอบ

4. สรุปผลการพิจารณาตัดสินเนื้อหาของข้อสอบที่ละข้อ ว่าข้อสอบข้อนั้นวัดในเนื้อหาใด และด้วยจุดมุ่งหมายข้อใด มีข้อใดบ้างที่ไม่ตรง

5. คำนวณหาความสอดคล้องของคำพิจารณาตัดสินเนื้อหาของข้อสอบของกลุ่มผู้เชี่ยวชาญจะต้องมีค่าสูง เช่น เกินร้อยละ 80 จะเป็นค่าดัชนีของข้อสอบรายข้อ แล้วจึงสรุปผลเป็นค่าดัชนีของแบบสอบทั้งฉบับ

เฮวตี วิบูลย์ศรี (2528:71-73) ได้เสนอวิธีการตรวจสอบความตรงเชิงเนื้อหาไว้ดังนี้

1. กำหนดวัตถุประสงค์ของหลักสูตรให้ชัดเจน เพื่อจะได้นำไปสู่การตรวจสอบความสอดคล้องของแบบสอบ กับสถานการณ์ตามวัตถุประสงค์นั้น

2. แบบสอบที่ใช้วัดผลสัมฤทธิ์ทางการเรียน มักจะมีการสร้างตารางวิเคราะห์โครงสร้าง 2 ทาง เพื่อดูความสอดคล้องระหว่างขอบเขตเนื้อหา กับวัตถุประสงค์ของการวัด (เช่น วัดความรู้ความเข้าใจ วัดความระลึกได้ ฯลฯ) จะทำให้ตัดสินได้ว่าแบบสอบใช้วัดในเนื้อหาที่จำเป็นได้อย่างเหมาะสมเพียงใด

นอกจากนั้นเขาวดี วิบูลย์ศรี ยังได้เสนอกระบวนการที่จะทำให้มีเกณฑ์ เพื่อใช้ในการตัดสินความตรงเชิงเนื้อหาของแบบสอบ ว่าควรเป็นดังนี้

- (1) ให้มีการนิยาม หรือมีการอธิบายถึงขอบเขตของเนื้อหาอย่างชัดเจน
- (2) เกณฑ์ที่ใช้ในการตัดสินความเป็นตัวแทนของเนื้อหาที่ต้องการวัดของแบบสอบ ที่ผู้เชี่ยวชาญเป็นผู้พิจารณาควรมีความเป็นปรนัยและถูกต้องตามหลักวิชา

Hambleton, R.K. และ Zoal, J.N. (1991:9-17) ได้เสนอกระบวนการตรวจสอบความตรงตามเนื้อหาของแบบสอบ ว่าควรปฏิบัติดังต่อไปนี้ :

1. กำหนดคุณสมบัติของผู้เชี่ยวชาญ และคัดเลือกผู้เชี่ยวชาญ
2. ให้ผู้เชี่ยวชาญตรวจสอบข้อสอบว่ามีความสอดคล้องกับวัตถุประสงค์ มีความเป็นตัวแทนของเนื้อหาที่ต้องการวัด ไม่ซ้ำซ้อนกับเนื้อหาอื่น ๆ ไม่มีความลำเอียงของเนื้อหา
3. พิจารณาข้อสอบเพื่อเลือกใช้เทคนิคที่เหมาะสม

ลอร์ซี (Lawshe, C.H.) (cited in Pitiyanuwat, Somwang, 1981:173-189) ได้รับแรงกระตุ้นจากปัญหาของการออกกฎหมาย จึงได้ศึกษาถึงการตรวจสอบความตรงเชิงเนื้อหาของแบบสอบผลสัมฤทธิ์ ใช้คัดเลือกคนเข้าทำงาน เนื่องจากการคัดเลือกคนเข้าทำงานต้องได้บุคคลที่มีความรู้ดีและสามารถนำไปใช้ในการปฏิบัติงานได้เป็นอย่างดี จนเป็นที่ยอมรับจากสถาบันที่รับเข้าทำงาน วิชาชีพบางอย่างเช่น นักกฎหมาย แพทย์ พยาบาล ฯลฯ มีปัจจัยเสี่ยงสูงในการรับคนเข้าทำงาน จึงมีความเข้มงวดที่จะต้องรับคนที่มีความรู้และมีทักษะในการทำงานดีพอเข้าทำงาน จึงจะไม่ก่อให้เกิดอันตรายต่อผู้เข้ารับบริการ ลอร์ซีได้ตั้งเป้าหมายไว้ 2 ประการในการปรับปรุงวิธีการตรวจสอบความตรงเชิงเนื้อหาของแบบสอบไว้ดังนี้

1. ก่อให้เกิดความชัดเจนของกรอบแนวคิดเกี่ยวกับความตรงเชิงเนื้อหาของแบบสอบ
2. เสนอวิธีการตรวจสอบความตรงเชิงเนื้อหาของแบบสอบที่เป็นวิธีการเชิงปริมาณ

กรอบแนวคิดที่เกี่ยวข้องกับความตรงเชิงเนื้อหา

ลอร์ซีได้เสนอกรอบแนวคิดเกี่ยวกับความตรงเชิงเนื้อหาของแบบสอบไว้ดังต่อไปนี้

- (1) วิเคราะห์เนื้อหาที่จำเป็นต้องวัด โดยศึกษาจากหลักสูตรว่ามีการอธิบายขอบเขตของเนื้อหาไว้อย่างชัดเจนมากน้อยเพียงใด สามารถนำไปสร้างเป็นแบบสอบได้หรือไม่ ถ้าเนื้อหาที่ต้องการวัดนั้นมีความเป็นนามธรรมสูง ต้องวิเคราะห์ถึงโครงสร้างของเนื้อหา โดยการศึกษาจากงานวิจัยที่เกี่ยวข้องก็ได้ ทั้งนี้เพื่อจะได้ข้อสรุปว่าอะไรคือ เนื้อหาที่ต้องการวัด

(2) ต้องศึกษาธรรมชาติของหน้าที่ของผู้ที่มีความเกี่ยวข้องกับเนื้อหาที่ต้องการวัดว่าใครจะมีความรู้เกี่ยวกับเนื้อหาได้ดีกว่ากัน ลอว์ซีให้ความสำคัญกับหน้าที่ความรับผิดชอบที่มีต่อเนื้อหาที่ต้องการวัดหรือภาระงานที่มีศักยภาพในการทำงานต่างกัน ตัวอย่างเช่น เสมียน กับ หัวหน้างาน มีความรับผิดชอบต่อภาระงานในปริมาณที่แตกต่างกัน ย่อมมีศักยภาพและความรอบรู้ในเนื้อหาที่แตกต่างกันไปด้วย จึงส่งผลไปถึงความเชี่ยวชาญในเนื้อหาที่ต้องการวัดที่แตกต่างกันด้วยเช่นกัน จากแนวคิดนี้จึงนำไปสู่แนวทางในการคัดเลือกและแบ่งระดับความเชี่ยวชาญของผู้เชี่ยวชาญที่จะพิจารณาตัดสินเนื้อหาของแบบสอบด้วย

(3) การพิจารณาตัดสินเนื้อหาของแบบสอบ คณะผู้เชี่ยวชาญต้องประชุมพิจารณาเนื้อหาในหลักสูตรว่าเนื้อหาใดคือเนื้อหาที่จำเป็นต้องวัด เนื่องจากแบบสอบเป็นตัวแทนของจักรวาลเนื้อหาในหลักสูตร

วิธีการตรวจสอบความตรงเชิงเนื้อหาของแบบสอบ

ลอว์ซีได้เสนอวิธีการตรวจสอบความตรงเชิงเนื้อหาของแบบสอบไว้ดังนี้

(1) จัดทำบัญชีรายชื่อของคณะผู้เชี่ยวชาญที่จะเป็นผู้ประเมินเนื้อหาของแบบสอบ โดยมีการกำหนดคุณสมบัติของผู้เชี่ยวชาญว่าจะต้องเป็นผู้ที่มีภาระงานที่เกี่ยวข้องกับเนื้อหา แล้วจัดส่งแบบสอบที่ต้องการประเมิน พร้อมแบบประเมินเนื้อหาของแบบสอบไปให้ผู้เชี่ยวชาญคนละ 1 ชุด ผู้เชี่ยวชาญทุกคนจะต้องมีอิสระในการประเมินเนื้อหาของแบบสอบ มีเกณฑ์ในการประเมินเนื้อหาของแบบสอบ ดังตัวอย่าง คือ

เสนอข้อสอบ 1 ข้อ พร้อมคำถาม คือ

เนื้อหาที่ถูกวัดโดยข้อสอบข้อนี้ มี :

- จำเป็นที่สุด
- มีประโยชน์แต่ไม่จำเป็น
- ไม่จำเป็น

สำหรับการปฏิบัติงาน

จำนวนคำตอบของผู้เชี่ยวชาญที่บอกว่าข้อสอบข้อนี้วัดในเนื้อหาที่จำเป็น จะเป็นตัวกำหนดความตรงเชิงเนื้อหาของข้อสอบ

(2) เพื่อให้เกิดความตรงของคำพิจารณาตัดสินของผู้เชี่ยวชาญ คณะผู้เชี่ยวชาญต้องตกลงกันเพื่อกำหนดขอบเขตของเนื้อหาที่ต้องการวัด

(3) รวบรวมคำพิจารณาตัดสินของคณะผู้เชี่ยวชาญ

(4) คำนวณหาค่าดัชนีความตรงเชิงเนื้อหาของแบบสอบโดยหาค่าเฉลี่ยจากข้อสอบ

ทุกข้อในแบบสอบ

ลอร์วี ได้ให้สูตรที่ใช้ในการคำนวณหาสัดส่วนของความตรงตามเนื้อหา (Content Validity Ratio = CVR₁): ของข้อสอบเป็นดังนี้

$$CVR_1 = \frac{n_o - 1/2 N}{1/2 N}$$

เมื่อ n_o = จำนวนผู้เชี่ยวชาญที่ตัดสินข้อสอบว่า "จำเป็นที่สุด"

N = จำนวนผู้เชี่ยวชาญทั้งหมด

CVR_1 = ค่าที่ตัดแปลงจากจำนวนเปอร์เซ็นต์ของผู้ที่ตอบว่า "จำเป็นที่สุด"

การแปลความหมายของค่า CVR_1 ตามแนวคิดของลอร์วี

4.1 เมื่อได้ค่า CVR_1 เป็นลบ แสดงว่าผู้เชี่ยวชาญจำนวนน้อยกว่าครึ่งของจำนวนผู้เชี่ยวชาญทั้งหมดประเมินเนื้อหาของข้อสอบว่า "วัดในเนื้อหาที่จำเป็นต่อวัด"

4.2 เมื่อได้ค่า CVR_1 เป็น 0 แสดงว่าผู้เชี่ยวชาญจำนวนครึ่งหนึ่งประเมินเนื้อหาของข้อสอบว่า "ใช้วัดในเนื้อหาที่จำเป็นต่อวัด" และผู้เชี่ยวชาญอีกจำนวนครึ่งหนึ่งประเมินเนื้อหาของข้อสอบว่า "ใช้วัดในเนื้อหาที่ไม่จำเป็นต่อวัด"

4.3 เมื่อได้ค่า CVR_1 มีค่าอยู่ระหว่าง 0-0.99 แสดงว่าผู้เชี่ยวชาญจำนวนมากกว่าครึ่งของจำนวนผู้เชี่ยวชาญทั้งหมดแต่ไม่ใช่ทั้งหมด ประเมินเนื้อหาของข้อสอบว่า "ใช้วัดในเนื้อหาที่จำเป็นต่อวัด"

4.4 เมื่อได้ค่า CVR_1 เท่ากับ 1.00 แสดงว่าผู้เชี่ยวชาญจำนวนทั้งหมดประเมินเนื้อหาของข้อสอบว่า "ใช้วัดในเนื้อหาที่จำเป็นต่อวัด"

ดัชนีความตรงตามเนื้อหา (The Content Validity Index) = CVI

การคำนวณหาค่าดัชนีความตรงตามเนื้อหาของแบบสอบทั้งฉบับ (CVI) จะใช้วิธีการง่าย ๆ คือ หาค่าเฉลี่ยของค่า CVR_1 ของข้อสอบที่คงมีอยู่ในแบบสอบทั้งหมดนั่นเอง สูตรที่ใช้ในการคำนวณหาค่าดัชนีความตรงเชิงเนื้อหาเป็นดังนี้

$$CVI = \frac{\sum CVR_1}{n}$$

เมื่อ CVI = ค่าดัชนีความตรงเชิงเนื้อหาของแบบสอบ

CVR_1 = ค่าสัดส่วนความตรงเชิงเนื้อหาของข้อสอบข้อที่ i

n = จำนวนข้อสอบทั้งหมดที่มีอยู่ในแบบสอบ

(5) การคัดเลือกข้อสอบที่มีความตรงเชิงเนื้อหาไว้ ลอร์ซีได้เสนอว่าถ้าข้อสอบข้อใดถ้าไม่ผ่านเกณฑ์ขั้นต่ำของ Schipper ดังแสดงไว้ในตารางที่ 2 ให้ตัดทิ้งไป แต่ก่อให้เกิดปัญหาจำนวนข้อสอบที่น้อยเกินไป จนไม่สามารถนำไปใช้ในการตรวจสอบคุณภาพแบบสอบในลักษณะอื่นๆ เช่น ค่าอำนาจจำแนกของข้อสอบ

ตารางที่ 2 ค่าต่ำสุดของ CVR_1 ที่ใช้เป็นเกณฑ์ในการคัดเลือกข้อสอบไว้ใช้ในการสอบ ($p < .05$)

จำนวนผู้เชี่ยวชาญ	ค่า CVR_1 ต่ำสุด
2	0.99
3	0.99
4	0.99
5	0.99
6	0.99
7	0.99
8	0.78
9	0.78
10	0.62
11	0.59
12	0.56

รายงานวิจัยที่มีการตรวจสอบความตรงเชิงเนื้อหาของแบบสอบทุกฉบับ บุคคลที่มีบทบาทสำคัญในการตรวจสอบความตรงเชิงเนื้อหาของแบบสอบ คือ ผู้เชี่ยวชาญ ซึ่งจะต้องเป็นบุคคลที่มีความรอบรู้ในเนื้อหาที่ต้องการวัดเป็นอย่างดี จึงจะสามารถประเมินเนื้อหาของแบบสอบได้อย่างสอดคล้องกับความเป็นจริง ดังนั้น Livingston, S.A. and Zieky, M.J. (1989) ได้เสนอแนวทางที่จะนำมาใช้ประกอบการพิจารณาในการคัดเลือกผู้เชี่ยวชาญ ทั้งนี้ผู้เชี่ยวชาญจะเป็นผู้พิจารณาตัดสินเนื้อหาแบบสอบ และต้องตอบคำถามที่ว่า "แบบสอบฉบับนี้มีความตรงเชิงเนื้อหาหรือไม่" ผู้เชี่ยวชาญควรมีความสามารถที่จำเป็น 2 ประการ คือ

- (1) ทราบขอบเขตของเนื้อหาที่จำเป็นต้องการวัด
- (2) มีความสามารถในการแยกแยะวิเคราะห์เนื้อหาของแบบสอบถามว่าวัดได้ สอดคล้องกับขอบเขตเนื้อหาที่ต้องการวัดหรือไม่

นอกจากนั้น Livingston, S.A. and Zieky, M.J. ยังได้ให้ข้อเสนอแนะเกี่ยวกับ จำนวนของผู้เชี่ยวชาญว่าควรมีจำนวนมากจะดีกว่า แต่ก็จะทำให้มีความยุ่งยากในการเก็บรวบรวม ข้อมูล ดังนั้นจึงยังไม่มียืนยันว่าจำนวนของผู้เชี่ยวชาญควรจะเป็นจำนวนอย่างน้อยแค่ไหน ที่จะทำให้การประเมินเนื้อหาของแบบสอบถามได้อย่างเหมาะสม

การประเมินความตรงเชิงเนื้อหาให้เป็นเชิงปริมาณ (A Quantitative Approach to Content Validity)

การตรวจสอบความตรงเชิงเนื้อหาของแบบสอบถาม ขึ้นกับผู้สร้างข้อสอบถามมีความตระหนักถึงความสำคัญของความตรงเชิงเนื้อหาของแบบสอบถาม ถ้าผู้สร้างข้อสอบถามตระหนักถึงความสำคัญและมีความรู้เกี่ยวกับวิธีการตรวจสอบความตรงเชิงเนื้อหาของแบบสอบถามเป็นอย่างดี ก็จะทำ การตรวจสอบความตรงเชิงเนื้อหาอย่างเป็นระบบและใช้วิธีการตรวจสอบที่มีความชัดเจน ในกรณีที่ผู้สร้างข้อสอบถามไม่ตระหนักถึงความสำคัญก็จะใช้วิธีการตรวจสอบแบบรวบรัด เช่น มีการนำเสนอเพียงแบบสอบถามและบอกวัตถุประสงค์ของการวัดให้ผู้เชี่ยวชาญเพื่อประเมินเนื้อหาของแบบสอบถาม ว่ามีความสอดคล้องกับวัตถุประสงค์หรือไม่ เมื่อใช้วิธีการตรวจสอบความตรงเชิงเนื้อหาของแบบสอบถามอย่างชัดเจน ก็จะบอกลักษณะของผู้เชี่ยวชาญว่าควรประกอบด้วยลักษณะอย่างไร และบอกปริมาณของข้อมูลเกี่ยวกับเนื้อหาที่ต้องการวัดที่เสนอต่อผู้เชี่ยวชาญว่ามีรายละเอียดมากน้อยเพียงใด จึงจะก่อให้เกิดประสิทธิภาพในการประเมินความตรงเชิงเนื้อหาของแบบสอบถามได้ การตรวจสอบความตรงเชิงเนื้อหา มักจะให้ผู้เชี่ยวชาญพิจารณาตรวจสอบเนื้อหาของข้อสอบถามเป็นรายข้อ พร้อมกับมีการนำเสนอเกณฑ์ภายนอกให้ผู้เชี่ยวชาญนำมาใช้เป็นแนวทางในการพิจารณาตัดสินเนื้อหาของข้อสอบถาม ต่อไปนี้เป็นวิธีการตรวจสอบความตรงเชิงเนื้อหาของแบบสอบถามที่เป็นวิธีการเชิงปริมาณที่จะนำเสนอในรูปแบบของวิธีการที่แตกต่างกัน ดังมีรายละเอียดต่อไปนี้

(1) ตามแนวทางของ Rovinalli และ Hambleton (1977) มีการนำเสนอ วัตถุประสงค์ของการวัดพร้อมกับข้อสอบ ต่อผู้เชี่ยวชาญจำนวนหลายคน ผู้เชี่ยวชาญแต่ละคนต้อง ตรวจให้คะแนนข้อสอบแต่ละข้อ โดยมีเกณฑ์ภายนอกที่ใช้ในการพิจารณาตัดสินเนื้อหาของข้อสอบคือ

- 1 = แน่ใจว่าข้อสอบข้อนั้นวัดวัตถุประสงค์นั้นได้จริง
- 0 = ไม่แน่ใจว่าข้อสอบข้อนั้นวัดวัตถุประสงค์ได้จริง
- 1 = แน่ใจว่าข้อสอบข้อนั้นไม่วัดวัตถุประสงค์นั้น

ผู้วิเคราะห์จะรวมคะแนนของผู้เชี่ยวชาญในแต่ละข้อ หาค่าเฉลี่ยและค่าเบี่ยงเบนมาตรฐาน ตัวเลขนี้จะเรียกว่า "ดัชนีของความสอดคล้องกันระหว่างข้อสอบกับวัตถุประสงค์" (Item-Objective Congruence = IOC)

(2) ตามแนวทางของ Rovinalli และ Hambleton วิธีการนี้ใช้มาตราส่วน ประมาณค่าเป็นเครื่องมือ ผู้เชี่ยวชาญจะได้รับรายการวัตถุประสงค์ของการวัดกับข้อสอบรายข้อ เหมือนกับวิธีการแรก ผู้เชี่ยวชาญจะให้คะแนนของความเหมาะสมที่พิจารณาแล้วว่าข้อสอบแต่ละข้อ สามารถวัดวัตถุประสงค์ข้อนั้นได้ เช่น

- 5 คือ เหมาะสมมากที่สุด
- 1 คือ เหมาะสมน้อยที่สุด

แล้วนำเอาตัวเลขที่ผู้เชี่ยวชาญแต่ละคนตัดสินประมาณค่า มาหาค่าเฉลี่ยของแต่ละข้อ แล้วทำการ พิจารณาร่วมกับค่าเบี่ยงเบนมาตรฐาน ก็จะได้ค่าที่เป็นดัชนีบอกความตรงตามเนื้อหา = IOC

(3) ตามแนวทางของ Rovinalli และ Hambleton ให้ผู้เชี่ยวชาญหลายคนพิจารณา รายการของข้อสอบ และรายการของวัตถุประสงค์ต่าง ๆ แล้วให้ทำการจับคู่กันว่าข้อสอบข้อใด ใช้วัตถุประสงค์ข้อใด อาจใช้ค่าสถิติ ไค-สแควร์ ทดสอบความเป็นอิสระของข้อมูล หรืออาจใช้วิธี พิจารณาด้วยสายตา ก็พอจะเห็นความสอดคล้อง หรือไม่สอดคล้อง เป็นรายข้อ

(Cunningham ; 1986 : 120 - 132)

(4) วิธีการของ Cronbach แบบที่หนึ่ง เขาเสนอให้ใช้วิธีการที่มีผู้เขียนข้อสอบ 2 กลุ่ม ที่มีความสามารถทัดเทียมกัน เขียนข้อสอบกลุ่มละ 1 ชุด ใช้วัดวัตถุประสงค์เดียวกัน แล้วนำไปสอบกับนักเรียนที่เรียนตามวัตถุประสงค์นั้นแล้ว ถ้ามีการกำหนดวัตถุประสงค์อย่างชัดเจน และการสุ่มข้อสอบจากโตะเมททำได้อย่างเหมาะสมแล้ว ผลจากการทำข้อสอบทั้งสองชุดของนักเรียน กลุ่มนี้ต้องมีความสอดคล้องกัน ข้อสอบที่ไม่มีมีความสอดคล้องกันจะต้องถูกตัดออกไป เนื่องจากไม่มี ความตรงตามเนื้อหา

(5) วิธีการของ Cronbach แบบที่สอง (1990) เขาได้นำเสนอแบบตรวจสอบใช้สำหรับตรวจสอบเนื้อหาของข้อสอบแบบหลายตัวเลือก เพื่อใช้พิจารณาเนื้อหาของแบบสอบวัดผลสัมฤทธิ์ทางการเรียน โดยทำการนิยามที่ละข้อพร้อมกับ 1 วัตถุประสงค์ ในลักษณะตัวอย่างที่เสนอในตารางที่ 3

ตารางที่ 3 แบบตรวจสอบ ใช้สำหรับตรวจสอบเนื้อหาของแบบสอบถามผลสัมฤทธิ์ชนิดเลือกตอบตาม
แนวทางของ Cronbach (1990)

ให้กา ✓ สำหรับ "ใช่" : กา X สำหรับ "ไม่ใช่" : กา ? สำหรับ "ไม่แน่ใจ"

1. เนื้อหาของข้อสอบมีความสอดคล้องกับวัตถุประสงค์ใช่หรือไม่
2. ตัวเลือกเป็นอิสระจากเนื้อหาที่ไม่ต้องการวัดใช่หรือไม่
3. ภาษาที่ใช้ในข้อสอบ ทั้งตัวคำถาม และตัวเลือกมีความเหมาะสมใช่หรือไม่
4. คำถามของข้อสอบมีการนำเสนอเพียงปัญหาเดียวใช่หรือไม่
5. คำถามของข้อสอบมีความชัดเจนไม่คลุมเครือใช่หรือไม่
6. ข้อสอบมีการเน้นทางด้านลบใช่หรือไม่ **
7. ภาษาที่ใช้ในข้อสอบถูกหลักไวยากรณ์ใช่หรือไม่
8. มีคำตอบที่ถูกต้องชัดเจนที่สุดเพียงข้อเดียวใช่หรือไม่
9. มีการชี้แนะคำตอบที่ถูกต้องให้ข้อสอบข้ออื่นหรือไม่ **
10. เมื่อมีการใช้ภาพแสดง/ตาราง/ตารางภาพ สิ่งเหล่านี้มีความถูกต้องและนิมน์อย่างชัดเจนใช่หรือไม่
11. ข้อสอบข้ออื่นสามารถตอบโดยใช้เหตุผลง่าย ๆ หรือโดยใช้สามัญสำนึกใช่หรือไม่ **
12. มีการใช้ศัพท์ที่ชี้แนะถึงคำตอบที่ถูกต้อง เช่น เสมอ ไม่เคย ทุกอย่างๆ ใช่หรือไม่ **
13. มีคำซ้ำ ๆ หรือการแสดงซ้ำ ๆ ที่ควรถูกตัดออกไปใช่หรือไม่ **
14. ตัวลวงดูน่าเชื่อถือและดึงดูดความสนใจ ของผู้สอบที่ไม่รู้คำตอบที่ถูกต้อง ได้หรือไม่
15. ตัวเลือกทุกตัวมีความยาวเท่ากันโดยประมาณ ใช่หรือไม่

** ข้อใดที่มีเครื่องหมายต้องตอบ "ไม่ใช่" เท่านั้น

(6) วิธีการของ Carver (1970) เขาได้นำเสนอวิธีการตรวจสอบความตรงตามเนื้อหาของแบบสอบถามที่โดยใช้ค่าสถิติแบบง่าย ๆ ข้อมูลที่ใช้ในการวิเคราะห์ที่ได้จากผลการสอบของนักเรียนที่เรียนแล้ว กับนักเรียนที่ยังไม่เรียน ทำการคำนวณโดยใช้ข้อมูลจากตารางที่ 4

ตารางที่ 4 ตัวอย่างคะแนนที่ได้จากนักเรียน 2 กลุ่ม เมื่อใช้วิธีการคำนวณของ Carver

	กลุ่มที่ยังไม่เรียน	กลุ่มที่เรียนแล้ว
ผ่าน	b	a
ไม่ผ่าน	c	d

$$\begin{aligned} \text{ความตรงตามเนื้อหา} &= (a + c) / N \\ \text{โดยที่ } N &= a + b + c + d \end{aligned}$$

(อ้างใน ไทย ทินย์สุวรรณกุล, 2529:40-41)

(7) วิธีการของ Lawshe (1975)

วิธีการตรวจสอบความตรงเชิงเนื้อหาของ Lawshe เป็นวิธีการที่ต้องการประเมินความสอดคล้องของเนื้อหาของแบบสอบถามกับขอบเขตของเนื้อหาที่ต้องการวัด ได้มาจากการหาค่าเฉลี่ยของสัดส่วนความสอดคล้องของเนื้อหาของข้อสอบทุกข้อที่มีอยู่ในแบบสอบ

การวัดขนาดของความสอดคล้อง (Measuring The Extent of Overlap) ของเนื้อหาของข้อสอบกับเนื้อหาที่จำเป็นต้องวัดนั้น จะมีคณะผู้เชี่ยวชาญที่จะทำการประเมินเนื้อหาของแบบสอบ โดยจะทำการประเมินรายข้อ นั่นคือในข้อสอบ 1 ข้อผู้เชี่ยวชาญจะต้องตอบคำถามต่อไปนี้

(ooo) ทักษะนี้ หรือ ความรู้นี้ ... ที่ถูกวัดโดยข้อสอบข้อนี้มี :

... ความจำเป็นที่สุด (Essential)

... ประโยชน์แต่ไม่จำเป็นที่สุด

... ไม่จำเป็น

สำหรับการปฏิบัติงาน

คำตอบที่ได้จากคณะผู้เชี่ยวชาญจะถูกรวบรวม และจำนวนผู้เชี่ยวชาญที่ตอบว่าข้อสอบข้อนั้น ใช้วัด
ในเนื้อหาที่จำเป็นที่สุด จะเป็นตัวกำหนดความตรงตามเนื้อหาของข้อสอบ

ลอร์ซีได้ให้สูตรที่ใช้ในการคำนวณหาสัดส่วนของความตรงตามเนื้อหา (Content
Validity Ratio = CVR_1): ของข้อสอบเป็นดังนี้

$$CVR_1 = \frac{n_u - 1/2 N}{1/2 N}$$

เมื่อ n_u = จำนวนผู้เชี่ยวชาญที่ตัดสินข้อสอบว่า "จำเป็นที่สุด"
 N = จำนวนผู้เชี่ยวชาญทั้งหมด
 CVR_1 = ค่าที่ตัดแปลงจากจำนวนเปอร์เซ็นต์ของผู้ที่ตอบว่า "จำเป็นที่สุด"

ดัชนีความตรงตามเนื้อหา (The Content Validity Index) = CVI

การคำนวณหาค่าดัชนีความตรงตามเนื้อหาของแบบสอบทั้งฉบับ (CVI) จะใช้วิธีการง่าย ๆ คือ
หาค่าเฉลี่ยของค่า CVR ของข้อสอบที่คงมีอยู่ในแบบสอบทั้งหมดนั่นเอง สูตรที่ใช้ในการคำนวณหา
ค่าดัชนีความตรงเชิงเนื้อหาเป็นดังนี้

$$CVI = \frac{\sum CVR_1}{n}$$

เมื่อ CVI = ค่าดัชนีความตรงเชิงเนื้อหาของแบบสอบ
 CVR_1 = ค่าสัดส่วนความตรงเชิงเนื้อหาของข้อสอบข้อที่ i
 n = จำนวนข้อสอบทั้งหมดที่มีอยู่ในแบบสอบ

แนวทางของการพิจารณาตัดสินค่าดัชนีความตรงเชิงเนื้อหาของแบบสอบ

เกณฑ์ขั้นต่ำของการตัดสินค่าดัชนีความตรงเชิงเนื้อหาของแบบสอบ จะใช้เกณฑ์ต่ำสุด
ของค่า CVR_1 ของ Schipper, L. (cited in Pitayanuwat, S., 1981:179) มาเป็น
แนวทางในการเสนอเกณฑ์ตัดสินค่าดัชนีความตรงเชิงเนื้อหาของแบบสอบ ดังที่แสดงไว้ใน
ตารางที่ 5

ตารางที่ 5 เกณฑ์การตัดสินค่าดัชนีความตรงเชิงเนื้อหาของแบบสอบ

จำนวนผู้เชี่ยวชาญ	ค่า CVI ต่ำสุด
2	0.99
3	0.99
4	0.99
5	0.99
6	0.99
7	0.99
8	0.78
9	0.78
10	0.62
11	0.59
12	0.56

การแปลความหมายของค่า CVI ตามแนวคิดของลอร์ซี

1. เมื่อได้ค่า CVI เป็นลบ แสดงว่าโดยเฉลี่ยแล้วจำนวนน้อยกว่าครึ่งของจำนวนผู้เชี่ยวชาญทั้งหมดประเมินเนื้อหาของข้อสอบว่า "วัดในเนื้อหาที่จำเป็นต่อวัด"
2. เมื่อได้ค่า CVI เป็น 0 แสดงว่าโดยเฉลี่ยแล้วจำนวนครึ่งหนึ่งประเมินเนื้อหาของข้อสอบว่า "ใช้วัดในเนื้อหาที่จำเป็นต่อวัด" และผู้เชี่ยวชาญอีกจำนวนครึ่งหนึ่งประเมินเนื้อหาของข้อสอบว่า "ใช้วัดในเนื้อหาที่ไม่จำเป็นต่อวัด"
3. เมื่อได้ค่า CVI มีค่าอยู่ระหว่าง 0-0.99 แสดงว่าโดยเฉลี่ยแล้วจำนวนมากกว่าครึ่งของจำนวนผู้เชี่ยวชาญทั้งหมดแต่ไม่ใช่ทั้งหมด ประเมินเนื้อหาของข้อสอบว่า "ใช้วัดในเนื้อหาที่จำเป็นต่อวัด"
4. เมื่อได้ค่า CVI เท่ากับ 1.00 แสดงว่าโดยเฉลี่ยแล้วจำนวนทั้งหมดประเมินเนื้อหาของข้อสอบว่า "ใช้วัดในเนื้อหาที่จำเป็นต่อวัด"

ขอให้ดูแนวทางการพัฒนาแบบสอบภายหลังทราบค่า CVI ในหน้าถัดไป

แนวทางการพัฒนาแบบสอบภายหลังทราบค่า CVI แล้ว

เมื่อต้องการพัฒนาแบบสอบให้มีความตรงเชิงเนื้อหาสูงมากขึ้น จะต้องดำเนินการดังต่อไปนี้

ขั้นที่ 1 ให้พิจารณาความหมายของค่า CVR_n ใน 3 ลักษณะ ดังนี้

1. ข้อสอบข้อใดที่มีค่า CVR_n เป็นบวก และสูงกว่าเกณฑ์ขั้นต่ำของ Schipper, L. ดังที่แสดงไว้ในตารางที่ 2 จะเป็นข้อสอบที่มีความตรงเชิงเนื้อหาอยู่ในเกณฑ์ที่สูง ควรคัดเลือกไว้ใช้ในแบบสอบ

2. ข้อสอบข้อใดที่มีค่า CVR_n เป็นบวก แต่ต่ำกว่าเกณฑ์ เป็นข้อสอบที่ผู้เชี่ยวชาญยังมีความลังเลในความตรงเชิงเนื้อหาของข้อสอบข้อนั้นอยู่

3. ข้อสอบข้อใดที่มีค่า CVR_n เป็นลบ เป็นข้อสอบที่ควรตัดทิ้ง ไม่ควรนำมาใช้ในการสอบ

ขั้นที่ 2 ให้พิจารณาคัดเลือกข้อสอบรายข้อตามแนวทางดังต่อไปนี้

1. ให้เลือกใช้ข้อสอบเฉพาะที่มีค่า CVR_n เป็นบวกและสูงกว่าเกณฑ์ หากจำนวนข้อสอบมีมากเพียงพอและสอดคล้องกับตารางวิเคราะห์เนื้อหาแล้ว ก็ให้นำไปคำนวณหาค่า CVI ได้ จะได้แบบสอบที่มีความตรงเชิงเนื้อหาสูง

2. ถ้าได้ข้อสอบที่มีค่า CVR_n เป็นบวกและสูงกว่าเกณฑ์ แต่มีจำนวนข้อสอบไม่เพียงพอให้เลือกข้อสอบที่มีค่า CVR_n เฉพาะที่มีค่าบวกค่อนข้างสูงแต่ต่ำกว่าเกณฑ์มาเพิ่มเติม จนได้จำนวนข้อสอบครบตามต้องการแล้ว ก็ให้คำนวณค่า CVI ได้ จะทำให้ได้แบบสอบที่มีความตรงเชิงเนื้อหาเช่นกัน

หมายเหตุ ข้อสอบข้อใดที่มีค่า CVR_n เป็นลบ จะเป็นข้อสอบที่ผู้เชี่ยวชาญจำนวนมากกว่าครึ่งพิจารณาแล้วเห็นว่าไม่ได้วัดในเนื้อหาที่ต้องการวัดจึงไม่ควรนำมาใช้ในแบบสอบ ต้องนำมาแก้ไขปรับปรุงให้มีคุณภาพดีก่อน แล้วจึงนำไปให้ผู้เชี่ยวชาญตรวจสอบเนื้อหาซ้ำ

(8) วิธีการวิเคราะห์สัดส่วนความสอดคล้อง (Accordance Proportion Method)

เป็นวิธีตรวจสอบความตรงเชิงเนื้อหาของแบบสอบที่เป็นที่มักจะใช้กันในปัจจุบัน และคำนวณได้จากสัดส่วนของผู้เชี่ยวชาญที่พิจารณาตัดสินว่า ข้อสอบข้อนั้นวัดในเนื้อหาที่จำเป็นต่อการวัด จะทำการคำนวณได้จากสูตรนี้

$$AP_1 = \frac{n_1}{N}$$

- เมื่อ
- AP_1 = สัดส่วนของผู้เชี่ยวชาญที่ได้พิจารณาตัดสินว่าข้อสอบข้อนั้นวัด
ในเนื้อหาที่ต้องการวัด
- n_1 = จำนวนของผู้เชี่ยวชาญที่ได้พิจารณาตัดสินว่าข้อสอบข้อนั้นวัดใน
เนื้อหาที่ต้องการวัด
- N = จำนวนผู้เชี่ยวชาญที่ใช้ในการศึกษาทั้งหมด

ค่าดัชนีสัดส่วนความสอดคล้องของแบบสอบ (Accordance Proportion Index) ทั้งฉบับคำนวณ
ได้จากสูตร

$$API = \frac{\sum AP_1}{n}$$

- เมื่อ
- API = ค่าดัชนีการวิเคราะห์สัดส่วนความสอดคล้องของแบบสอบ
- AP_1 = สัดส่วนของเนื้อหาของข้อสอบที่วัดในเนื้อหาที่ต้องการวัด
โดยมีผู้เชี่ยวชาญเป็นผู้พิจารณาตัดสิน
- n = จำนวนข้อสอบที่มีอยู่ในแบบสอบทั้งฉบับ

รายงานวิจัยที่เกี่ยวข้อง

รายงานวิจัยที่นำเสนอครั้งนี้ จะนำเสนอเฉพาะการพัฒนาคุณภาพของแบบสอบถามด้านความตรงเชิงเนื้อหาของแบบสอบถาม จะทำให้ทราบว่ามีความหลากหลายของวิธีการตรวจสอบความตรงเชิงเนื้อหาของแบบสอบถาม ดังมีรายละเอียดต่อไปนี้

ลอร์ซีได้ศึกษาวิธีการตรวจสอบความตรงเชิงเนื้อหาของแบบสอบถาม โดยนำแบบสอบถามที่ได้จัดสร้างไว้สำหรับคัดเลือกบุคคลเข้ารับการฝึกทางช่าง เป็นแบบสอบถามที่ต้องประเมินความตรงเชิงเนื้อหา โดยมีจุดมุ่งหมายที่จะคัดเลือกคนที่มีความรู้และมีความสามารถทางช่างเป็นอย่างดีเข้าทำงาน คณะบุคคลที่จะเป็นผู้ประเมินเนื้อหาของแบบสอบถาม เป็นบุคคลต่างๆที่เกี่ยวข้องกับเนื้อหาจำนวน 175 คน มาจากโรงงาน 17 แห่ง ประกอบด้วย

- (1) หัวหน้าช่าง
- (2) ช่าง
- (3) บุคคลที่ฝึกเป็นช่าง
- (4) ครูผู้สอน
- (5) ผู้สมัครเข้าร่วมในการโครงการฝึก

จะเห็นได้ว่าไม่ได้ระบุลักษณะของผู้เชี่ยวชาญว่าควรเป็นเช่นไร แต่ระบุเพียงเป็นบุคคลที่เกี่ยวข้องกับเนื้อหาเท่านั้น จึงต้องใช้คณะบุคคลเป็นจำนวนมากถึง 175 คน คณะบุคคลที่จะประเมินเนื้อหาของแบบสอบถามจะได้รับเอกสารคนละ 1 ชุด คือ

- (1) แบบสอบถามที่ต้องการให้ประเมินความตรงเชิงเนื้อหา
- (2) แบบประเมินเนื้อหาของแบบสอบถาม

คณะผู้ประเมินเนื้อหาของแบบสอบถามจะต้องมีอิสระในการประเมินเนื้อหาของแบบสอบถาม ผลจากการพิจารณาตัดสินเนื้อหาของแบบสอบถามจะถูกนำไปคำนวณค่าความตรงเชิงเนื้อหาของแบบสอบถามแล้วไปเปรียบเทียบกับเกณฑ์ขั้นต่ำของ Schipper, L. ดังรายละเอียดที่แสดงไว้ในตารางที่ 6 แล้วนำมาหาค่าเฉลี่ยของค่า CVR₁ ของข้อสอบทุกข้อ จะได้ค่าดัชนีความตรงเชิงเนื้อหาของแบบสอบถามเท่ากับ .67

(cited in Pitayanuwat, S., 1981:181-182)

ตารางที่ 6 การแจกแจงของค่า CVR_1 ของแบบสอบคัดเลือกช่าง จำนวน 54 ชื่อ

CVR_1	f
90-94	13
85-89	12
80-84	3
75-79	0
70-74	1
65-69	2
60-64	7
55-59	0
50-54	1
45-49	2
40-44	2
35-39	2
30-34	3
25-29	3
20-24	2
15-19	0
10-14	0
5-9	0
0-4	1*
N = 54 (* = ไม่มีนัยสำคัญ)	

รายงานวิจัยฉบับของลอว์ซีเป็นการสังเคราะห์การตรวจสอบความตรงเชิงเนื้อหาของแบบสอบโดยทำงานร่วมกับสเตอร์นเบิร์ก (Lawshe and Sternberg, 1955; cited in Pitayanuwat, S., 1981:182-188) เป็นการตรวจสอบความตรงเชิงเนื้อหาที่เริ่มต้นจากการวิเคราะห์ขอบเขตของเนื้อหาที่ต้องการวัด โดยระบุขอบเขตของการศึกษาไว้ดังต่อไปนี้

- (1) ขอบเขตของเนื้อหาจะต้องได้รับการอธิบายให้กระจ่างชัด
- (2) แบบสอบที่สร้างขึ้นจะต้องสอดคล้องกับขอบเขตเนื้อหาที่ต้องการวัด

การวิเคราะห์ขอบเขตเนื้อหา จะถูกวิเคราะห์โดยทีมงานซึ่งจะประกอบด้วยนักวิเคราะห์ 8 คน เสมียน 4 คน ผู้ตรวจการ 4 คน คณะผู้เชี่ยวชาญจะประกอบด้วยบุคคลที่เกี่ยวข้องกับงานเชิงเสมียน เช่น หัวหน้าเสมียน เสมียน ผู้ตรวจการงานเสมียน ทุกคนจะมีประสบการณ์การทำงานอย่างน้อย 3 ปี คณะบุคคลชุดนี้จะทำหน้าที่ตรวจสอบขอบเขตเนื้อหาว่าเนื้อหาใดบ้างเป็นเนื้อหาที่จำเป็นต้องวัด

เนื้อหาที่นำมาใช้ในการสร้างแบบสอบจะเป็นแบบสอบวัดความสามารถเชิงเสมียน ข้อมูลเกี่ยวกับเนื้อหาได้มาจากมาตรฐานความสามารถเชิงเสมียน และนำมาใช้เป็นแนวทางในการประเมินเนื้อหาของแบบสอบวัดความสามารถเชิงเสมียนของมหาวิทยาลัยเพอร์ดู ผู้เชี่ยวชาญจะได้รับแบบสอบ 1 ชุด รายการมาตรฐานของความสามารถเชิงเสมียน และแบบประเมินเนื้อหาของแบบสอบ เมื่อจะกำหนดดัชนีความตรงเชิงเนื้อหาของแบบสอบแต่ละฉบับ ต้องกระทำดังนี้คือ

- (1) แยกแยะเนื้อหาย่อยใดบ้างที่มีค่า CVR ที่มียุทธศาสตร์สำหรับแบบสอบ
- (2) คำนวณหาค่าเฉลี่ยของค่า CVR เพื่อจะได้ทราบ CVI

ดังตัวอย่างที่แสดงไว้ในตารางที่ 7

ตารางที่ 7 ตัวอย่างเนื้อหาของย่อในภาระงาน A และค่า CVI ของแบบสอบความสามารถเชิง
เสมียน

เนื้อหาที่	เนื้อหาของย่อเชิงเสมียน	CVR
1	การคำนวณง่ายๆ เช่น การบวก หรือ การลบ	.99
3	การคำนวณที่เป็นขั้นตอนมากกว่า 1 ชั้น เช่น การคูณ การหาร	.99
4	การจดบันทึกข้อกระทง หรือ จำนวนช่วงเวลาทำการ โดยไม่ใช้ เครื่องมือช่วย	.99
45	กำหนดอัตรา ราคา จำนวน และการจัดประเภทข้อมูล ฯลฯ	.99
9	การคำนวณที่ซับซ้อน เช่น การถอดราก เลขยกกำลัง ฯลฯ	.83
38	การคำนวณเข้าเพื่อแก้ไขหรือตรวจสอบข้อมูลที่เป็นตัวเลข	.83
39	บอกความคลาดเคลื่อนที่พบ การบันทึกข้อมูล ฯลฯ	.83
ดัชนีความตรงเชิงเนื้อหาของแบบสอบ CVI (ค่าเฉลี่ยของ CVR)		.92

เพื่อกำหนดขอบเขตเนื้อหาที่จำเป็นต้องวัด โดยเลือกจากข้อสอบที่มีค่า CVR ที่มี
นัยสำคัญ แต่เนื้อหาในข้อสอบบางข้อไม่เกี่ยวข้อง ดังนั้นจึงต้องตรวจสอบกับความเป็นจริงตาม
พื้นฐานธรรมชาติของเนื้อหาแต่ละเนื้อหาด้วย งานวิจัยฉบับนี้จึงเป็นการทดลองนำมาใช้เพื่อกำหนด
ขอบเขตของเนื้อหาที่ต้องการวัดนั่นเอง

ดริงใจ หนูผลอำนวย (2534) ทำการวิจัยเรื่อง "การพัฒนาแบบสอบเฉพาะบุคคล
ในวิชาวิทยาศาสตร์สำหรับนักเรียน ชั้นมัธยมศึกษาปีที่ 3" เป็นงานวิจัยที่เน้นความสำคัญของความ
ตรงตามเนื้อหาอย่างชัดเจน โดยเฉพาะในระยะที่ 1 ของการวิจัยได้มีความพยายามสร้างข้อสอบ
ให้ครอบคลุมเนื้อหาทั้งหมด ระบุลักษณะของผู้เชี่ยวชาญว่าเป็นนักวัดผลจำนวน 3 คน เป็นอาจารย์
ผู้สอนวิชาวิทยาศาสตร์จำนวน 5 คน โดยจะต้องตรวจสอบความสอดคล้องระหว่างเนื้อหาของ
แบบสอบ กับ วัตถุประสงค์ของการวัด ผลของการวิจัยพบว่าแบบสอบเฉพาะบุคคลวิชาวิทยา
ศาสตร์ มีค่าดัชนีความสอดคล้องของข้อสอบกับวัตถุประสงค์อยู่ระหว่าง 0.75-1.00

ทนาย สิงห์พันธ์ (2534) ทำการวิจัยเรื่อง "การพัฒนาแบบวัดภาคปฏิบัติวิชาฟิสิกส์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 4" ได้พัฒนาคุณภาพของแบบสอบด้านความตรงเชิงเนื้อหาของแบบสอบ โดยมีผู้เชี่ยวชาญทางด้านฟิสิกส์จำนวน 11 คน ประเมินความสอดคล้องของเนื้อหาของข้อสอบ กับวัตถุประสงค์ของการทดลอง ใช้เกณฑ์ 0.5 ตัดสินว่าข้อสอบข้อนี้มีความตรงเชิงเนื้อหา

ทิพย์วรรณ มุลทองสุน (2534) ทำการวิจัยเรื่อง "การพัฒนาแบบสอบแบบ MBQ. เพื่อวัดความสามารถในการแก้ปัญหา สำหรับนักเรียนชั้นประถมศึกษาปีที่ 6" ได้พัฒนาคุณภาพของแบบสอบด้านความตรงเชิงเนื้อหาของแบบสอบ มีการระบุลักษณะของผู้เชี่ยวชาญ คือ เป็นนักวัดผลจำนวน 12 คน ช่วยตรวจดูความชัดเจนของภาษา และตรวจดูคำถามแต่ข้อว่าใช้วัดความสามารถในการแก้ปัญหาได้หรือไม่ เป็นอาจารย์สอนชั้นประถมศึกษาปีที่ 6 จำนวน 5 คน ช่วยตรวจดูว่าข้อสอบเหมาะสมกับระดับความสามารถของเด็กชั้นประถมศึกษาปีที่ 6 ได้หรือไม่ ไม่ัดชั้นบอกความตรงเชิงเนื้อหา ของแบบสอบ

เทียนพร ริงษ์อรุณกร (2532) ทำการวิจัยเรื่อง "การพัฒนาแบบสอบภาคปฏิบัติ ทมวตวิชาคหกรรมศาสตร์" ได้พัฒนาคุณภาพของแบบสอบด้านความตรงเชิงเนื้อหาของแบบสอบ โดยมีการระบุลักษณะของผู้เชี่ยวชาญว่าเป็นอาจารย์สอนวิชาคหกรรมศาสตร์ไม่น้อยกว่า 3 ปี และ เคยเรียนวิชาคหกรรมในระดับปริญญาตรีมาก่อน จำนวน 5 คน เป็นผู้รับผิดชอบหลักสูตรจำนวน 1 คน เป็นนักโภชนาการที่มีประสบการณ์ไม่น้อยกว่า 5 ปี จำนวน 1 คน ตรวจดูเนื้อหาของข้อสอบ โดยใช้เกณฑ์ คือ

- (1) ความเป็นตัวแทนของเนื้อหาที่ต้องการวัดหรือไม่
- (2) สร้างได้ครอบคลุมเนื้อหาที่สำคัญหรือไม่
- (3) เนื้อหาของข้อสอบสอดคล้องกับวัตถุประสงค์ของการวัดหรือไม่

เกณฑ์ที่ใช้ในการตัดสิน คือ ผู้เชี่ยวชาญจำนวน 6 ใน 7 คน มีความเห็นสอดคล้องกัน จะถือว่าข้อสอบข้อนี้มีความตรงเชิงเนื้อหา

ภาพร ลินสวัสดิ์ (2534) ทำการวิจัยเรื่อง "การพัฒนาแบบประเมินสมรรถภาพ การเรียนวิทยาศาสตร์สำหรับนักเรียนมัธยมศึกษาปีที่ 3" ได้พัฒนาคุณภาพของแบบสอบด้านความตรงเชิงเนื้อหาของแบบสอบ โดยมีผู้เชี่ยวชาญทางด้านหลักสูตร จำนวน 10 คน ช่วยตรวจดูการ เขียนคำนิยามของตัวแปรหลัก และรายการข้อคำถามที่ให้ผู้เชี่ยวชาญช่วยประเมินว่าใช้ได้ หรือ ใช้ไม่ได้ เกณฑ์เกินร้อยละ 80 ให้คงคำถามข้อนั้นไว้

บุญเลิศ คำหอม (2525) ทำการวิจัยเรื่อง "การสร้างแบบทดสอบอิงเกณฑ์ เรื่อง สมการและอสมการ สำหรับชั้นมัธยมศึกษาปีที่ 2" ได้พัฒนาคุณภาพของแบบสอบด้านความตรงเชิงเนื้อหา ของแบบสอบ โดยมีการวิเคราะห์เนื้อหาและจุดมุ่งหมายเชิงพฤติกรรม กำหนดให้ผู้เชี่ยวชาญ ประเมินเนื้อหาของข้อสอบกับวัตถุประสงค์เชิงพฤติกรรม โดยไม่ระบุลักษณะของผู้เชี่ยวชาญ จึงนำไปทดลองใช้กับกลุ่มตัวอย่าง 2 ครั้ง คำนวณหาค่าความตรงเชิงเนื้อหาโดยใช้สูตรของ Carver

ปิยะชาติ โชคนิพนธ์ (2527) ทำการศึกษาเรื่อง "ระดับความตรงของเนื้อหาของแบบสอบผลสัมฤทธิ์วิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 4 ของโรงเรียนสังกัดกรุงเทพมหานคร" นำแบบสอบวิชาคณิตศาสตร์จำนวน 24 ชุด มาวิเคราะห์เนื้อหาของแบบสอบ และสัมภาษณ์อาจารย์ผู้ออกข้อสอบจำนวน 43 คน แล้วนำข้อมูลที่ได้บันทึกลงในตาราง 2 มิติ จำแนกตามวัตถุประสงค์และพฤติกรรม ให้ผู้เชี่ยวชาญคืออาจารย์ผู้สอนคณิตศาสตร์ จำนวน 18 คน กำหนดเนื้อหาที่สำคัญที่จำเป็นต่อวัดของหลักสูตรลงในตาราง 2 มิติ ทำการคำนวณหาดัชนีของความลำเอียงของแบบสอบ โดยเปรียบเทียบความแตกต่างของสัดส่วนข้อสอบจากแบบสอบ กับข้อสอบที่ผู้เชี่ยวชาญระบุว่าจำเป็นต่อมีในหลักสูตร พบว่าแบบสอบวัดผลสัมฤทธิ์วิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 4 ของโรงเรียนในสังกัดกรุงเทพมหานคร มีความตรงตามเนื้อหาทั้งฉบับอยู่ในระดับปานกลาง

ดุสิต อัครชัยสุวิกรม (2536) ทำการวิจัยเรื่อง "การสร้างแบบทดสอบผลสัมฤทธิ์วิชางานผ้าและการตัดเย็บ 1 ทมวตวิชาคหกรรมศาสตร์ ชั้นมัธยมศึกษาปีที่ 2 ในจ.สระบุรี" ได้พัฒนาคุณภาพของแบบสอบด้านความตรงเชิงเนื้อหา มีการวิเคราะห์งานโดยผู้เชี่ยวชาญจำนวน 5 คน แล้วตรวจสอบหา Logical Validity ระหว่างเนื้อหาที่เป็นภาระงานกับวัตถุประสงค์เชิงพฤติกรรม คำนวณหาค่าดัชนีความตรงเชิงเนื้อหาที่บอกความสอดคล้องระหว่างเนื้อหาของข้อสอบกับวัตถุประสงค์เชิงพฤติกรรม (IOC) ใช้เกณฑ์ 0.5 ตัดสินเนื้อหาของข้อสอบว่ามีความตรงเชิงเนื้อหา

ระพีพรรณ ศรีวิเชียร (2534) ทำการวิจัยเรื่อง "การพัฒนาแบบสอบตามระดับความสามารถในวิชาคณิตศาสตร์ สำหรับนักเรียนมัธยมศึกษาชั้นปีที่ 3" ได้พัฒนาคุณภาพแบบสอบด้านความตรงเชิงเนื้อหาของแบบสอบ ได้ระบุคุณสมบัติของผู้เชี่ยวชาญว่ามีความรู้ในระดับปริญญาตรี มีประสบการณ์การสอนนาน 5 ปี จำนวน 5 คน ตรวจสอบความสอดคล้องของเนื้อหาของข้อสอบกับวัตถุประสงค์ของการวัด เกณฑ์ที่ใช้ตัดสิน คือ 0.5 จะถือว่าข้อสอบข้อนั้นมี ความตรงเชิงเนื้อหา

วลีรินทร์ เน้มกว่าเก่า (2530) ทำการวิจัยเรื่อง "การสร้างแบบทดสอบวัดคุณธรรม อริยมรรคมีองค์แปดด้านปัญญา (สัมมาทิฐิ และ สัมมาสังกัปปะ) สำหรับนักเรียนมัธยมศึกษาชั้นปีที่ 1" ได้พัฒนาคุณภาพของแบบสอบด้านความตรงเชิงเนื้อหาของแบบสอบ คือการหา Face Validity มีการนำแบบสอบที่ผู้วิจัยสร้างขึ้นไปให้ผู้เชี่ยวชาญตรวจสอบ มีการระบุผู้เชี่ยวชาญว่าเป็นผู้เชี่ยวชาญทางศาสนาและทางวัดผล จำนวน 5 คน ใช้เกณฑ์ความสอดคล้องของผู้เชี่ยวชาญ 60% จะถือว่าข้อสอบข้อนั้นมี ความตรงเชิงเนื้อหา

สุจินต์ สารงาม (2524) ทำการวิจัยเรื่อง "การสร้างแบบทดสอบวัดความถนัดทางการเรียนคหกรรมศาสตร์" ได้พัฒนาแบบทดสอบทางด้านความตรงเชิงเนื้อหาของแบบสอบ โดยมีการกำหนดลักษณะของผู้เชี่ยวชาญ คือ อาจารย์ประจำภาควิชาคหกรรมศาสตร์ จำนวน 40 คน เป็นผู้ประเมินเนื้อหาของแบบสอบว่าสร้างได้ครอบคลุมเนื้อหาที่ต้องการวัดหรือไม่ ได้ใช้แบบสำรวจว่าข้อสอบข้อใดวัดในเนื้อหาที่จำเป็น หรือไม่จำเป็น ใช้เกณฑ์ร้อยละ 50 ตัดสินว่าข้อสอบข้อใดมีความตรงเชิงเนื้อหาหรือไม่

สุภาวดี ตั้งบุปผา (2533) ทำการวิจัยเรื่อง "การสร้างแบบทดสอบวัดความสามารถในการคิดสร้างสรรค์ทางคณิตศาสตร์ ของนักเรียนชั้นมัธยมศึกษาปีที่ 3 กรุงเทพมหานคร" ได้พัฒนาคุณภาพของแบบสอบด้านความตรงเชิงเนื้อหา โดยผู้สร้างข้อสอบได้เขียนคำนิยามศัพท์ แล้วสร้างข้อสอบให้สอดคล้องกับคำนิยามศัพท์ นำแบบสอบไปให้ผู้เชี่ยวชาญตรวจสอบความสอดคล้องนั้น โดยไม่ระบุลักษณะของผู้เชี่ยวชาญและจำนวนของผู้เชี่ยวชาญ

อาภรณ์ ชุตวง (2534) ทำการวิจัยเรื่อง "การสร้างแบบสอบ MEQ เพื่อวัดความสามารถทางการพยาบาล" ได้พัฒนาคุณภาพแบบสอบด้านความตรงเชิงเนื้อหาของแบบสอบ โดยมีผู้เชี่ยวชาญเป็นพยาบาลประจำการ จำนวน 10 คน ช่วยตรวจสอบความสอดคล้องของปัญหาเกี่ยวกับวัตถุประสงค์การวัด ใช้เกณฑ์ความสอดคล้องเกิน 80% จะถือว่าข้อสอบข้อนั้นมีความตรงเชิงเนื้อหา

อารมณี สนั่นภู (2518) ทำการวิจัยเรื่อง "การพัฒนาแบบสอบมาตรฐานวิชาวิทยาศาสตร์ทั่วไป และ 2 เพื่อใช้กับนักศึกษาวิทยาลัยครูสกลนคร" ได้พัฒนาคุณภาพของแบบสอบด้านความตรงเชิงเนื้อหาของแบบสอบ โดยมีการกำหนดลักษณะของผู้เชี่ยวชาญเป็นอาจารย์ผู้สอนวิชาวิทยาศาสตร์นาน 2 เทอม จำนวน 17 คน ตรวจสอบดูความสอดคล้องของน้ำหนักของเนื้อหาที่ต้องการวัด กับ วัตถุประสงค์ของการสอน ไม่มีดัชนีที่ใช้บอกความตรงเชิงเนื้อหาของแบบสอบ

อาร์วิดสัน (Arvidson, 1990 บทคัดย่อ) ผู้วิจัยต้องการพัฒนาเครื่องมือที่ใช้ในการวิจัยใหม่ทั้งความเที่ยงและความตรง วิธีการทำให้เกิดความตรงเชิงเนื้อหาของเครื่องมือคือ การศึกษาทบทวนเอกสารและวรรณคดีที่เกี่ยวข้อง แล้วจึงสร้างเครื่องมือขึ้นมา และให้ผู้เชี่ยวชาญจำนวน 6 คน ทำการประเมินเนื้อหาของเครื่องมือ ไม่มีตัวเลขแสดงความตรงเชิงเนื้อหาของเครื่องมือ

ลิปส์คอมบ์ (Lipscomb, 1990 บทคัดย่อ) ผู้วิจัยต้องการพัฒนาแบบสอบให้เนื้อหาของข้อสอบวัดได้ครอบคลุมในขอบเขตของเนื้อหาที่ต้องการวัด มีการศึกษาเปรียบเทียบการสุ่มเนื้อหาที่ใช้สร้างข้อสอบอยู่ 2 วิธี คือ แบบ Random Sampling กับ แบบ Stratified Random Sampling จากข้อสอบที่มีอยู่ทั้งหมด และมีการกำหนดขนาดความยาวของแบบสอบเป็น จำนวน 100, 50, 25, 12, และ 6 ข้อ พบว่าปริมาณของความตรงเชิงเนื้อหาของแบบสอบจะเพิ่มขึ้นเมื่อจำนวนข้อสอบลดลง

ครอว์ฟอร์ด (Crawford, 1991 บทคัดย่อ) เป้าหมายของการวิจัย คือ การทำให้เกิดความตรงเชิงเนื้อหาของตัวอย่างข้อสอบจำนวน 75 ข้อ ที่คัดลอกมาจากคลังข้อสอบวิชากระบวนการพยาบาล 1 ผู้วิจัยตรวจสอบความตรงเชิงเนื้อหาของแบบสอบ โดยการประยุกต์ใช้เทคนิคเดลฟาย การคัดเลือกผู้เชี่ยวชาญได้จากคำแนะนำของหัวหน้าหน่วยงานการสอนของภาควิชา Technical and Adult Education คณะผู้เชี่ยวชาญต้องทำการประเมินเนื้อหา 2 รอบ ไม่ได้ระบุจำนวนผู้เชี่ยวชาญ ผลการวิจัยพบว่าข้อสอบทั้ง 75 ข้อใช้วัดในเนื้อหาที่ต้องการวัด ไม่มีตัวเลขของระดับความตรงเชิงเนื้อหา และข้อสอบมีจำนวนไม่เพียงพอที่จะใช้วัดได้ครอบคลุมทุกโดเมนของเนื้อหา ต้องคัดเลือกข้อสอบเพิ่มขึ้น

แวนซินิ (Vancini, 1991 บทคัดย่อ) การศึกษาครั้งนี้ผู้วิจัยต้องการพัฒนาและทดสอบมาตรที่ใช้วัด "การสาธิตการพยาบาลเกี่ยวกับภาวะสุขภาพของบุคคล" ทำการตรวจสอบความตรงเชิงเนื้อหาโดยให้ผู้เชี่ยวชาญจำนวน 6 คน ประเมินเนื้อหา แล้วคำนวณหาสัดส่วนความเกี่ยวข้องกันเป็นค่าดัชนีที่บอกถึงความตรงเชิงเนื้อหาของมาตรวัด ผลจากการวิจัยพบว่า ข้อคำถามจำนวน 15 ข้อของมาตรวัด มีค่าดัชนีความตรงเชิงเนื้อหาอยู่ระหว่าง 0.62 ถึง 0.98

คลาร์ค (Clark, 1991 บทคัดย่อ) เป้าหมายของการศึกษาครั้งนี้คือการพัฒนาแบบสอบให้ทั้งความตรงและความเที่ยง โดยเฉพาะทางด้านความตรงเชิงเนื้อหาได้มีการวิเคราะห์วัตถุประสงค์การวัดและขอบเขตของเนื้อหาที่ต้องการวัด แล้วสร้างข้อสอบให้ครอบคลุมเนื้อหา แล้วตรวจสอบความตรงเชิงเนื้อหาโดยใช้เทคนิคเดลฟาย ไม่ระบุผู้เชี่ยวชาญและจำนวนผู้เชี่ยวชาญ ไม่มีดัชนีบอกความตรงเชิงเนื้อหา

เพรสเลย์ (Presley, 1991 บทคัดย่อ) เป้าหมายของการศึกษาครั้งนี้คือ พัฒนาเครื่องมือที่ใช้วัดระดับแอลกอฮอล์ในเด็กวัยรุ่น ให้ทั้งความเที่ยงและความตรง มีการตรวจสอบ Logical content Validity โดยให้ผู้เชี่ยวชาญพิจารณาความสอดคล้องของเนื้อหาของข้อสอบกับตารางวิเคราะห์เนื้อหา ใช้เกณฑ์ร้อยละ 70 ของความคิดเห็นที่สอดคล้องกันได้ข้อสอบที่ผ่านเกณฑ์จำนวน 107 ข้อ

ซาง (Zhang, 1994 บทคัดย่อ) มีเป้าหมายของการศึกษาคือ ต้องการพัฒนาแบบสอบความรู้ เพื่อออกไปรับรองสำหรับครูฝึกวอลเลย์บอลล์ ได้พัฒนาคุณภาพด้านความตรงเชิงเนื้อหาของแบบสอบ โดยมีการวิเคราะห์งานและทำตารางวิเคราะห์เนื้อหาที่ต้องการวัด เมื่อสร้างข้อสอบให้มีจำนวนครอบคลุมเนื้อหาที่ต้อวัด จึงนำไปให้ผู้เชี่ยวชาญตรวจสอบความลำเอียงของเนื้อหาของข้อสอบ และตรวจดูความชัดเจนของภาษาที่ใช้ด้วย

จากทฤษฎีและแนวคิดต่างๆที่เกี่ยวข้องกับความตรงเชิงเนื้อหาของแบบสอบที่ศึกษาค้นคว้ามา ผู้วิจัยได้นำมาใช้เป็นแนวทางของการวิจัยเพื่อพัฒนาวิธีการตรวจสอบความตรงเชิงเนื้อหาของแบบสอบต่อไป