



บทที่ 2

วรรณคดีและงานวิจัยที่เกี่ยวข้อง

ในบทนี้จะกล่าวถึงประเด็นใหญ่ๆที่เกี่ยวข้องกับเรื่องต่างๆต่อไปนี้

1. ทฤษฎีที่เกี่ยวข้องกับรูปแบบฟาเซท
 - 1.1 ความหมาย พัฒนาการ และรูปแบบของเทคโนโลยีการเขียนข้อสอบ
 - 1.2 ความเป็นมาและความหมายของรูปแบบฟาเซท
 - 1.3 หลักการและข้อเสนอแนะในการใช้รูปแบบฟาเซทสร้างข้อสอบและตัวเลือก
 - 1.4 ประโยชน์และข้อจำกัดของรูปแบบฟาเซท
2. หลักการสร้างและการตรวจสอบคุณภาพของแบบสอบอิง เกณฑ์
 - 2.1 หลักการสร้างแบบสอบอิง เกณฑ์
 - 2.2 การตรวจสอบคุณภาพแบบสอบอิง เกณฑ์
 - 2.3 การกำหนดเกณฑ์มาตรฐาน
 - 2.4 ความเที่ยงและความตรงของแบบสอบอิง เกณฑ์
3. งานวิจัยที่เกี่ยวข้อง

1. ทฤษฎีที่เกี่ยวข้องกับรูปแบบฟาเซท

- 1.1 ความหมาย พัฒนาการและรูปแบบของเทคโนโลยีการเขียนข้อสอบ

รอยด์ และฮาลาดินา (Roid and Haladyna, 1982) ได้กล่าวถึงความหมายของเทคโนโลยีการเขียนข้อสอบไว้ดังนี้

1. เป็นวิธีการเขียนข้อสอบที่สามารถบอกถึงความสัมพันธ์ของการเรียนการสอนและการสอบได้อย่างมีเหตุผลและเที่ยงตรง
2. เป็นวิธีการเขียนข้อสอบที่ผู้เขียนต้องมีประสบการณ์ด้านเนื้อหา แม้ว่าจะไม่มีประสบการณ์ หรือทักษะในการเขียนข้อสอบมากก็สามารถเขียนข้อสอบได้
3. เป็นวิธีการเขียนข้อสอบที่มีการกำหนดลักษณะเฉพาะ ซึ่งทำให้ง่ายสำหรับผู้

เขียนข้อสอบคนอื่นๆที่จะสามารถเขียนข้อสอบตามลักษณะเฉพาะที่กำหนดไว้แล้วได้

บุญเชิด ภิญโญนนตพงษ์ (2527) กล่าวว่า เทคโนโลยีการเขียนข้อสอบคือ เครื่องช่วยผลิตข้อสอบที่คู่ขนานกันจากประชากรข้อสอบได้จำนวนมากโดยการกำหนดลักษณะเฉพาะของมวลความรู้ ซึ่งครอบคลุมทั้งพฤติกรรมและเนื้อหาวิชาอย่างชัดเจน พร้อมทั้งสร้างกฎเกณฑ์ในการเขียนข้อสอบ

ตามคำนิยามดังกล่าว จะเห็นได้ว่า หากกำหนดเนื้อหาที่จะสอบไว้ได้อย่างละเอียด ชัดเจนและเฉพาะเจาะจงตามจุดประสงค์ของการเรียนการสอนแล้ว ก็จะช่วยให้เขียนข้อสอบได้ง่าย ตรงตามต้องการ และผลิตข้อสอบได้จำนวนมาก นับเป็นที่มาของเทคโนโลยีการเขียนข้อสอบ

ราวปี ค.ศ. 1968 ทีมมหาวิทยาลัยมินเนโซตา ไฮฟลีและคณะ (Hively and other, 1968 อ้างถึงใน ชูศักดิ์ ชัมภลชิต, 2529) ได้พัฒนาแบบสอบคู่ขนานขึ้นหลายฉบับ เพื่อใช้วิจัยควบคู่ไปกับการพัฒนาหลักสูตรวิชาคณิตศาสตร์และวิทยาศาสตร์ระดับประถมศึกษา ในการนี้ ไฮฟลีและคณะได้สร้างรูปแบบที่ช่วยในการเขียนข้อสอบชนิดหนึ่งคือ ฟอรม์ข้อสอบ (Item Form) ซึ่งถือว่าเป็นเครื่องมือชิ้นแรกของการพัฒนาการของเทคโนโลยีการเขียนข้อสอบ ในช่วงเวลาที่ใกล้เคียงกัน ทีมมหาวิทยาลัยแคลิฟอร์เนีย เมืองลอสแอนเจลิส (UCLA) ภายใต้การนำของปอปแฮม (Popham, 1978) ได้ตั้งสถาบันผู้เชี่ยวชาญทางด้านแบบสอบอิงเกณฑ์ ชื่อ Institute of Objective Exchange (IOX) สถาบัน IOX นี้ จัดตั้งขึ้นเพื่อเป็นศูนย์รวบรวมจุดประสงค์เชิงพฤติกรรม ของกระบวนการเรียนการสอนหลาย ๆ หมวดวิชาให้ครอบคลุมมากที่สุด และได้จัดพิมพ์จุดประสงค์ของแต่ละรายวิชา และตัวอย่างข้อสอบไว้บริการครู หรือผู้สนใจทั่วไป แต่สถาบัน IOX ไม่สามารถบรรลุเป้าหมายอันนี้ได้ เพราะจุดประสงค์เชิงพฤติกรรมของแต่ละข้อ เป็นข้อความที่กำหนดไว้อย่างหลวม ๆ และไม่ได้กำหนดแนวทางเฉพาะให้กับผู้เขียนข้อสอบอย่างชัดเจน สถาบัน IOX จึงดัดแปลงฟอรม์ข้อสอบของไฮฟลีและคณะ โดยสร้างรูปแบบที่ช่วยในการเขียนข้อสอบขึ้นใหม่สองอย่างคือ จุดประสงค์ขยายความ (Amplified Objectives) และลักษณะเฉพาะของข้อสอบ (Test Specifications)

บอร์มูธ (Bormuth, 1970) เป็นบุคคลแรกที่เสนอเทคนิคเก่าแก่ของ การเขียนข้อสอบอีกวิธีหนึ่ง ซึ่งเรียกว่า แบบสอบข้อความเรียง (Item for Prose Learning) ซึ่งมีหลักและ วิธีการแปลงประโยคสำคัญของเนื้อหาวิชา ให้เป็นองค์ประกอบของมวลความรู้ (Domain) แล้วเขียนข้อสอบจากมวลความรู้ดังกล่าว ชูศักดิ์ ชัมภลชิต (2529) กล่าวว่า ในระยะ 5-6 ปีที่แล้ว มีเทคโนโลยีการเขียนข้อสอบใหม่ๆ เกิดขึ้นอีกหลายแนวคิด เช่น การประยุกต์ทฤษฎีฟาเซทของกัตแมน (Guttman) มาใช้เป็นเทคนิคการเขียนข้อสอบตามแนวคิดเชิงแมป ซึ่งเรียกว่า "The Mapping Sentence Method"

รอยด์และฮาလာไดนา (Roid and Haladyna, 1982) ได้กล่าวถึงรูปแบบของเทคโนโลยีการเขียนข้อสอบไว้ 6 วิธี (แต่ละวิธีมีรายละเอียดที่ภาคผนวก) คือ

1. การใช้ข้อความเรียง (Item for Prose Learning)
2. การใช้ฟอร์มข้อสอบ (Item Form)
3. การใช้รูปแบบฟาเซทหรือการจับคู่ประโยค (Facet Design or The Mapping Sentence Method)
4. การใช้เทคนิคการวัดมโนทัศน์ (Test Item for Concept)
5. การใช้เทคนิคเชิงตรรก (Logical Operations for Generating Intended Questions)
6. การใช้เทคนิค IQI (The Instructional Quality Inventory)

1.2 ความเป็นมาและความหมายของรูปแบบฟาเซท

1.2.1 ความเป็นมาของรูปแบบฟาเซท

รูปแบบฟาเซทมีรากฐานการพัฒนามาจากทฤษฎีบุคลิกภาพ โดยการนิยามรูปแบบทางจิตวิทยาของจุง (Jung's Psychological Types) (Foa, 1965) ในปี 1965 และ 1969 กัตแมน (Guttman, 1969 quoted in Roid and Haladyna, 1982) เป็นผู้เสนอใช้ประโยชน์ของรูปแบบฟาเซท กับการวัดผลสัมฤทธิ์ และได้เสนอวิธีการจัดลำดับชั้นของวิชาต่างๆกับพฤติกรรมทางสมอง (Cognitive Behavior) ดังตาราง

ตารางที่ 1 แสดงการใช้เมตริกซ์ลำดับชั้นของกัตแมน: ในพฤติกรรมทางสมอง

	B.1 รูปภาพ (Figural)	B.2 ตัวเลข (Numerical)	B.3 ภาษา (Verbal)
A.1 การอ้างอิงกฎ (Rule Infering)	ปัญหาเกี่ยวกับเรขาคณิต(Geometric Problems)	อันดับตัวเลข (Numerical Progression)	การให้คำที่เกี่ยวข้องกัน 2 คำ เช่น car : Ford
A.2 การประยุกต์กฎ (Rule Applying)	จับคู่ลูกบาศก์ (Matching cubes)	การใช้สูตร การคำนวณ การแก้ปัญหาเลขคณิต	การกำหนดประเภท สิ่งของ เช่น โรบิน : นก
A.3 ผลสัมฤทธิ์ (Achievement)	การเรียนรู้ภาษา- สัญลักษณ์(Learning a symbolic Language)	คณิตศาสตร์ ฟิสิกส์ เคมี สถิติ	ประวัติศาสตร์ วรรณคดี

จากตารางที่ 1 เป็นการประยุกต์วิธีการของกัตแมน สำหรับพฤติกรรมทางสมอง โดยพฤติกรรมทางสมองทั้งหมดสามารถจัดลงในมิติใดมิติหนึ่ง ซึ่งประกอบด้วยรูปภาพ (B.1) ตัวเลข (B.2) หรือภาษา (B.3) และจัดลงในมิติอีกหนึ่งคือ มิติที่ประกอบด้วย การอ้างอิงกฎ (A.1) การประยุกต์กฎ (A.2) และผลสัมฤทธิ์ (A.3) ตัวอย่างเช่น วิชาคณิตศาสตร์ และฟิสิกส์เป็นวิชาที่เกี่ยวกับผลสัมฤทธิ์และตัวเลข (A.3-B.2) ประวัติศาสตร์เป็นวิชาที่เกี่ยวกับผลสัมฤทธิ์ด้านภาษา (A.3-B.3)

แบบทดสอบวัดความสามารถทางสมองหลายชนิด ที่สามารถเข้าไปในการอ้างอิงกฎ หรือการประยุกต์กฎ โดยจัดแบ่งเป็นแบบทดสอบที่เกี่ยวข้องกับ รูปภาพ ตัวเลข หรือภาษา

ที่คล้ายคลึงกับโมเดลโครงสร้างทางสมองของกิลฟอร์ด (Guilford, 1967) ดังนั้นคำอธิบายสำหรับรูปแบบฟาเซทในการสร้างแบบสอบวัดผลสัมฤทธิ์ ซึ่งเสนอโดยกัตแมนั้น เป็นลักษณะหนึ่งของการนำทฤษฎีการเรียนรู้ที่ต่อเนื่อง (Hierarchical theory of learning) มาใช้กับแบบสอบวัดผลสัมฤทธิ์ของการปฏิบัติเกี่ยวกับรูปภาพ ตัวเลขและภาษา

1.2.2 ความหมายของรูปแบบฟาเซท

กัตแมน (Guttman, 1969 quoted in Roid and Haladyna, 1982) ได้อธิบายเกี่ยวกับรูปแบบฟาเซทว่าเป็นเทคโนโลยีการเขียนข้อสอบชนิดหนึ่ง ได้พัฒนามาจากฟอร์มข้อสอบ (Item Form) ซึ่งรูปแบบฟาเซทที่ใช้ในการเขียนข้อสอบมีลักษณะดังนี้

1. มีการนิยามขอบเขตของเนื้อหาอย่างมีระบบ โดยสังเกตได้จากประโยคจับคู่
2. มีการกำหนดฟาเซทที่จะใช้ในการจับคู่ประโยคให้มีลักษณะเฉพาะเจาะจงเกณฑ์ที่ใช้ในทฤษฎีนี้มี 2 ประการ คือ (1) เนื้อหา (2) โครงสร้างทางสถิติ โดยที่เนื้อหาจะกระทำให้เฉพาะเจาะจงได้ ด้วยสถานการณ์ที่เรียกว่า ประโยคจับคู่ (Mapping Sentence) ประโยคจับคู่จะเป็นพื้นฐานของการบ่งชี้โครงสร้างทางสถิติ ดังนั้นทฤษฎีนี้จะเห็นจริงถึงสาระสำคัญของความตรงเชิงโครงสร้างในการวัดผล

ชูว์เมเกอร์ (Shoemaker, 1975) กล่าวว่า วิธีการจับคู่ประโยค (Mapping Sentence) เป็นวิธีการสำหรับกำหนดลักษณะเฉพาะของ โครงสร้างและเนื้อหาในการเรียนการสอน ซึ่งสามารถประยุกต์มาจากทฤษฎีฟาเซท (Facet Theory)

รอยด์ และ ฮาลาดินา (Roid and Haladyna, 1982) กล่าวว่า สาระสำคัญของรูปแบบฟาเซท เป็นวิธีการในการรวบรวมและนิยามโดเมน โดยการกำหนดลักษณะเฉพาะของขอบเขตและลำดับชั้นขององค์ประกอบ และส่วนย่อยอื่นๆ รูปแบบฟาเซทจะเป็นการเตรียมโครงสร้าง และขอบเขตของโดเมนโดยการรวบรวมลักษณะเฉพาะของประโยคจับคู่ ซึ่งมีส่วนคล้ายกับฟอร์มข้อสอบ และกล่าวว่า ประโยคจับคู่เป็นวิธีการในการสร้างข้อกระทางได้จำนวนมากสำหรับแบบสอบวัดผลสัมฤทธิ์ซึ่งประโยคจับคู่ จะประกอบด้วยส่วนสำคัญ

3 ส่วนคือ

1. ส่วนที่คงที่ (Fixed Part) จะมีลักษณะคล้ายกับ เปลือกหุ้มฟอร์มข้อสอบ (Item Form Shell)
2. ส่วนที่เป็นฟาเซต (Facets) เป็นแง่มุมต่างๆของตัวแปร
3. ส่วนที่เป็นสมาชิกของฟาเซต (Facet Elements) จะมีลักษณะคล้ายกับ ส่วนที่นำไปแทนที่ (Replacement) ของฟอร์มข้อสอบ

การจับคู่ประโยคจะใช้ ส่วนที่เป็นฟาเซตและสมาชิกของฟาเซต ประกอบกันเป็น รูปแบบฟาเซต ซึ่งรูปแบบฟาเซตนี้จะเป็นการกำหนดลักษณะเฉพาะของมวลความรู้ ที่จะนำไปเขียนข้อสอบ

กัตแมน (Guttman, 1969) ได้ยกตัวอย่างรูปแบบฟาเซตซึ่งสามารถใช้อธิบาย การวิเคราะห์พฤติกรรมทางสมอง ซึ่งแสดงการจับคู่ประโยคได้ดังนี้

	ฟาเซต 1		ฟาเซต 2	
ถ้านักเรียนได้รับคำถามเกี่ยวกับ	<div style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px; display: inline-block;"> รูปภาพ ตัวเลข ภาษา </div>	และต้องการ	<div style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px; display: inline-block;"> อ้างอิง ประยุกต์ </div>	
	ฟาเซต 3			ฟาเซต 4
กฎที่	<div style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px; display: inline-block;"> เหมือนกัน คล้ายกัน ไม่เหมือนกัน </div>	ที่ครูสอนในชั้นเรียนของตัวเองแล้วนักเรียนควรตอบคำถาม		<div style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px; display: inline-block;"> ถูก ผิด </div>

จากการแสดงการจับคู่ประโยคข้างบน ซึ่งมีฟาเซต 4 ฟาเซต และมีจำนวน สมาชิกในแต่ละฟาเซตเป็น 3, 2, 3 และ 2 ตามลำดับ สมาชิกของฟาเซตจะเป็นคำที่ เปลี่ยนแปลงได้ในประโยค ตัวอย่างนี้ใช้สำหรับการอธิบายพฤติกรรมของการอ้างอิงกฎ หรือการประยุกต์กฎ เพื่อใช้ในการทำนายผลสัมฤทธิ์ ฟาเซตที่ 3 ในประโยคจับคู่ดังกล่าว

เสนอแนะว่ากฎควรจะมีการจัดลำดับชั้นที่คล้ายกันในการบรรยายในการเรียนการสอน จุดสำคัญที่กัทแมน (Guttman) ต้องการชี้ให้เห็นคือ ความสัมพันธ์ระหว่างชนิดของผลสัมฤทธิ์ทางการเรียนที่กำหนดไว้ในการเรียนการสอน และความสามารถทางสมองที่จะนำไปปฏิบัติ เช่น นักเรียนคนหนึ่งมีระดับของการอ้างอิงกฎและการประยุกต์กฎอยู่ในระดับสูง พฤติกรรมของนักเรียนคนนั้นในการลำดับตัวเลขทำให้เขาสามารถนำไปปฏิบัติได้

จากประโยคจับคู่ที่แสดงให้เห็นดังกล่าวนี้ ทำให้เราสามารถสร้างประโยคได้ 36 ประโยคที่แตกต่างกัน ซึ่งมีทั้งถูกและผิด เช่น

1. ถ้านักเรียนได้รับคำถามเกี่ยวกับภาษา และ ต้องการอ้างอิงกฎที่เหมือนกัน ที่ครูสอนในชั้นเรียนของตัวเองแล้ว นักเรียนควรจะตอบคำถามผิด
2. ถ้านักเรียนได้รับคำถามเกี่ยวกับตัวเลข และต้องการประยุกต์กฎที่คล้ายกัน ที่ครูสอนในชั้นเรียนแล้ว นักเรียนควรจะตอบคำถามถูก

รูปแบบการสร้างข้อคำถามจากประโยคจับคู่ นอกจากจะใช้สร้างเป็นแบบสอบแบบถูกผิดแล้ว ยังสามารถสร้างเป็นแบบทดสอบแบบเลือกตอบ โดยการทำให้ฟาเซท 3 ฟาเซทหลังคงที่ สร้างประโยคใหม่แล้วให้ฟาเซทแรก(ภาษา ตัวเลข และรูปภาพ) จัดหาคำตอบถูก 1 ตัว และผิด 2 ตัว นอกจากนั้นเรายังสามารถสร้างแบบทดสอบแบบเติมคำได้ โดยการตั้งสมาชิกของฟาเซทใดฟาเซทหนึ่ง ออกจากประโยคทั้ง 36 ประโยค ในแง่นี้เราสามารถสร้างแบบทดสอบเติมคำได้ 144 ข้อ เพราะว่าโดยอาศัยประโยค 36 ประโยคนั้นมีฟาเซทอยู่ 4 ฟาเซท โดยใน 1 ประโยคนั้นสามารถสร้างข้อสอบเติมคำได้ 4 ข้อ ($36 \times 4 = 144$ ข้อ) ฟาเซทที่ 4 ในตารางนี้ สามารถขยายสมาชิกเป็น 4 ตัว คือ ความสามารถในการปฏิบัติสูง ค่อนข้างสูง ค่อนข้างต่ำ และต่ำ ในกรณีนี้จึงเป็นการเพิ่มจำนวนประโยคหรือข้อคำถามที่สามารถสร้างได้เป็น 72 ประโยค เพื่อให้เห็นชัด เราพบว่ายังจำนวนสมาชิกมีมาก หรือมีจำนวนฟาเซทยิ่งมาก ในประโยคจับคู่หนึ่ง ๆ ยิ่งทำให้จำนวนข้อสอบเพิ่มมากขึ้น

1.3 หลักการ และข้อเสนอแนะในการใช้รูปแบบฟาเซทสร้างแบบสอบและตัวเลือก

1.3.1 หลักการและข้อเสนอแนะในการใช้รูปแบบฟาเซทสร้างข้อสอบ

เอ็นเงล และ มาร์ทูซ่า (Engel and Martuza, 1976 quoted in Roid and Haladyna, 1982) และ เบิร์ก (Berk, 1978) ได้เสนอแนะเกี่ยวกับการสร้างแบบสอบวัดผลสัมฤทธิ์โดยใช้รูปแบบฟาเซท โดยให้ศึกษาจุดประสงค์การสอน เพื่อเป็นพื้นฐานในการสร้างข้อความสำหรับจับคู่ประโยค (Mapping Sentence) และได้กำหนดหลักการสร้างข้อสอบโดยใช้รูปแบบฟาเซท 6 ขั้นตอนดังนี้

1. คัดเลือกจุดประสงค์การสอน (Selecting an instructional objective)

จุดประสงค์ที่จะใช้สร้างประโยคจับคู่ (Mapping Sentence) จะต้องเป็นจุดประสงค์ที่สำคัญและให้มโนทัศน์ (Concept) ในการที่จะทำให้ทราบฟาเซทต่างๆของวิชาที่จะศึกษาเพื่อที่จะสร้างข้อสอบ

2. แสดงเนื้อหาการสอนที่จะใช้สร้างเครื่องมือ (List instructional material)

รายละเอียดของเนื้อหาสาระ เป็นความต้องการพื้นฐานของผู้เขียนข้อสอบ ซึ่งจะใช้สำหรับกำหนดจุดประสงค์ขยายความ (Amplifying objective) จะศึกษาได้จากหนังสือเรียน คู่มือครู ซึ่งมีรายละเอียดของเนื้อหาและจุดประสงค์

3. พัฒนาจุดประสงค์ขยายความ (Develop an amplified objective)

ในส่วนนี้จะประกอบด้วยการบรรยายทั่วไป ของเนื้อหาให้เฉพาะเจาะจงขึ้น ยกตัวอย่างข้อคำถาม บรรยายสถานการณ์และลักษณะคำตอบของข้อคำถาม

4. การสร้างประโยคจับคู่ (Generate a mapping sentence)

จุดประสงค์ขยายความจะใช้เป็นพื้นฐานในการสร้างประโยคจับคู่ ฟาเซท และสมาชิกของฟาเซท จะเป็นการสร้างขอบเขตของความเป็นไปได้สำหรับแต่ละฟาเซท

5. การสร้างโครงสร้างฟาเซทของข้อสอบ (Generate the item facet structure)

ในการอธิบายแง่มุมนี้สามารถวิเคราะห์ได้เป็น 2 ลักษณะคือ ความสมบูรณ์

ของความเป็นตัวแทนในโดเมนต่างๆ และการจัดการตัวเลือกอื่นๆ ที่ทำให้เป็นการเพิ่มจำนวนฟาสเซตต่างๆของประโยคนั้น จะเป็นการเพิ่มจำนวนข้อสอบที่สามารถผลิตออกมาได้

6. การเขียนข้อกระทง (Write the item)

ในส่วนนี้สามารถที่จะใช้คอมพิวเตอร์ผลิตข้อสอบออกมาได้ หรือจะกระทำโดยการเลือกจากสมาชิกของแต่ละฟาสเซตและการแปรเปลี่ยนไปอย่างเป็นระบบของตัวลงกระทำได้เช่นเดียวกัน เหมือนกับคำตอบถูก

1.3.2 การพัฒนาการเขียนข้อสอบจากรูปแบบฟาสเซต

รันเคล และแมคกราท (Runkel and Mcgrath, 1972 quoted in Roid and Haladyna, 1982) ได้เสนอแนะเกณฑ์สำหรับการพัฒนาการเขียนข้อสอบจากรูปแบบฟาสเซตดังนี้

1. ฟาสเซตต่างๆ ที่ถูกเลือกจะต้องสามารถอธิบายในทัศนหลักของจุดประสงค์นั้นๆได้
2. ฟาสเซตแต่ละฟาสเซตจะประกอบด้วยสมาชิกที่เป็นไปได้ในแต่ละฟาสเซต
3. สมาชิกของฟาสเซตแต่ละฟาสเซต ต้องเป็นสมาชิกที่แยกจากกัน (mutually exclusive) และถ้าจุดประสงค์ใดสามารถแบ่งออกเป็นฟาสเซตได้มากกว่าหนึ่งฟาสเซตแล้ว ฟาสเซตที่แบ่งออกมานั้นจะต้องมีความสมบูรณ์ครบถ้วนในตัวเอง
4. ความสัมพันธ์ระหว่างฟาสเซตควรมีเหตุผล และมีลักษณะเฉพาะ ถ้ามีการจัดลำดับของฟาสเซตที่แน่นอน และประโยคจับคู่สามารถเชื่อมโยงแง่มุมต่างๆให้เป็นเรื่องที่ต่อเนื่องกัน
5. สมาชิกของฟาสเซตแต่ละฟาสเซต ควรมีเหตุผลที่สามารถอธิบายได้ว่าทำไมจึงถูกเลือกเป็นสมาชิกของฟาสเซตนั้นๆ ถ้าเป็นไปได้ควรมีการลำดับชั้นให้กับสมาชิกในแต่ละฟาสเซต
6. จำนวนฟาสเซตที่เป็นไปได้ทั้งหมดจะอยู่ภายใต้ขอบเขตของเนื้อหาสาระที่สอน

1.3.3 การสร้างตัวเลือกโดยใช้รูปแบบฟาเซท

ตัวลวงเป็นส่วนหนึ่งของข้อสอบ และใช้ตัดสินระดับความยากของข้อสอบโดยเชื่อมกับส่วนคำถาม กัดแมน และชลีเซนเจอร์ (Guttman and Schlesenger, 1967) กล่าวว่า ส่วนของคำถามของข้อสอบจะไม่สามารถบอกค่าความยากของข้อสอบข้อนั้นได้เลย ถ้าตัวลวงของข้อสอบข้อนั้นไม่ได้สร้างอย่างเป็นระบบ เพราะการเขียนตัวลวงอย่างเป็นระบบจะทำให้มีข้อดีดังนี้

1. ข้อสอบข้อนั้นสามารถพยากรณ์ความสัมพันธ์ของค่าความยากของตัวลวงแต่ละตัวได้
2. ครูสามารถแปลผลของแบบทดสอบได้อย่างชัดเจน เพราะได้มีการตัดตัวแปรที่ไม่เกี่ยวข้องออก
3. ความแตกต่างของคะแนนที่ได้จากการทดสอบของนักเรียนแต่ละคน ขึ้นอยู่กับรูปแบบของความสามารถในการตัดสินใจให้เลือกของตัวลวง

นอกจากนั้นการสร้างตัวลวงอย่างมีระบบ ยังทำให้สามารถสร้างแบบทดสอบที่มีความยาวน้อยกว่าปกติโดยไม่มีผลกระทบกับค่าความเชื่อมั่นและค่าความตรงของแบบทดสอบ

การเขียนข้อสอบโดยใช้รูปแบบฟาเซทนั้น กัดแมน และชลีเซนเจอร์ (Guttman and Schlesenger, 1967) เสนอว่าใช้ได้ทั้งการสร้างข้อคำถาม (Stem) และใช้ในการสร้างตัวลวงอย่างเป็นระบบ ซึ่งสามารถใช้ในการวินิจฉัยได้ด้วย ซึ่งจะขอยกตัวอย่างของเอ็นเจล (Engel, 1976 quoted in Roid and Haladyna, 1982) ดังนี้

จุฬาลงกรณ์มหาวิทยาลัย

จุดประสงค์

1. เมื่อกำหนดค่าของคู่อันดับ x และ y ให้ สามารถคำนวณหาค่า R_{xy} ได้ถูกต้อง
2. เมื่อกำหนดค่าสหสัมพันธ์แบบ Pearson r , Point biserial, Phi และ Rank order (ρ) ผู้เรียนจะเปรียบเทียบสิ่งที่เหมือนและต่างกันได้ถูกต้อง

พาเซทที่ใช้สร้างข้อคำถาม กำหนดให้

พาเซท A
เสนอข้อมูลในรูป
1. ตาราง
2. โจทย์บรรยาย

พาเซท B
รูปแบบเนื้อหา
1. คำต่อคำ
2. มโนทัศน์

พาเซท C
เซทของคู่ลำดับเป็นค่าตัวแปร " x, y "
1. 3 คู่อันดับ
2. 4 คู่อันดับ
3. 5 คู่อันดับ

พาเซท D
เป็นตัวแปร " x "
1. เลขหลักเดียว
2. เลขสองหลัก
3. เลขสามหลัก

พาเซท E
เป็นตัวแปร " y "
1. เลขหลักเดียว
2. เลขสองหลัก
3. เลขสามหลัก

พาเซทที่ใช้สร้างตัวเลือก

นักเรียนจะเลือกคำตอบของค่าสหสัมพันธ์ $r(x, y)$ จากเซทของตัวเลือกซึ่งสร้างจาก

พาเซท F
รูปแบบของคะแนน
1. คะแนนเบี่ยงเบน
2. คะแนนดิบ

พาเซท G
เครื่องหมายของผลคูณ
1. ถูก
2. ผิด $(-)(+) = +$
$(+)(-) = +$
$(-)(-) = -$

พาเซท H
การหาร $sp(x, y)$
1. ไม่ใช่
2. ใช่

พาเซท I
รากที่สองของ $ss(x)ss(y)$
1. ใช่
2. ไม่ใช่

พาเซท J
รูปแบบของหน่วย
1. ไม่มีหน่วย
2. หน่วยเป็นเส้นตรง (Linear unit)
3. หน่วยพื้นที่ (Square unit)

จากประโยคจับคู่ข้างบน เอ็นเจล (Engel) ได้สร้างข้อสอบไว้ดังนี้

(0) ตารางข้างล่างนี้เป็นคะแนนทัศนคติก่อน-หลังการสอบของนักเรียน 3 คน ในวิชาการออกแบบ(BU 307) ให้นักเรียนหาค่า Pearson r จากข้อมูลต่อไปนี้(เมื่อสร้างจาก $A_1B_1C_1D_1E_1$)

แบบสอบถามเกี่ยวกับทัศนคติ	
ก่อนสอบ	หลังสอบ
1	6
1	3
4	3

ก. -0.50 ข. 0.67 ค. 0.50 ง. -0.17

เอ็นเจล(Engel) กล่าวว่า ถ้านักเรียนคนใดเลือกตอบ(ตัวเลือก ค) โดยอาศัยรูปแบบฟาเซท ในการสร้างตัวลวงจะทำให้ทราบความสัมพันธ์ระหว่างตัวถูก (ตัวเลือก ก) และตัวลวง (ตัวเลือก ค) ดังนี้

ก. $F_1G_1H_1I_1J_1$

ข. $F_1G_2H_1I_1J_1$

ข้อสังเกต จะเห็นได้ว่าในระหว่างตัวเลือกถูก กับตัวลวงนั้นมีความแตกต่างกันที่ฟาเซท G โดยที่สมาชิกของฟาเซท G คือ G_2 นั้นทำให้ทราบว่านักเรียนมีความบกพร่องที่การใช้เครื่องหมายของผลคูณ อันจะเป็นแนวทางในการสอนซ่อมเสริม

รอยด์และฮาลาดินา (Roid and Haladyna, 1982) เสนอตัวอย่างการเขียนข้อสอบถูกผิดจากประโยคจับคู่ดังนี้

จุดประสงค์การสอน ให้รูปแบบของยุทธวิธีการกำหนดโดเมน ดังนี้ ฟอรัมข้อสอบ, วิธีการคำนวณ, ประโยคจับคู่ และการแปลงประโยค นักเรียนสามารถเลือกเนื้อหาที่เหมาะสมกับยุทธวิธีการกำหนดโดเมนนั้น

ประโยคจับคู่ "... (ฟาเซท A) ... เหมาะสำหรับการกำหนดโดเมนของวิชา... (ฟาเซท B) ..."

ฟาเซท A มีสมาชิก 4 ตัว คือ

- 1 การแปลงประโยค (Item transformation)
- 2 ฟอรัมข้อสอบ (Item form)
- 3 วิธีคำนวณ (Algorithms)
- 4 ประโยคจับคู่ (Mapping sentence)

ฟาเซท B มีสมาชิก 5 ตัว คือ

- 1 การอ่าน
- 2 ภาษา
- 3 คณิตศาสตร์
- 4 วิทยาศาสตร์
- 5 สังคม

จากตัวอย่างนี้สามารถสร้างแบบทดสอบถูกผิดได้ 20 ข้อ ตัวอย่างเช่น

1. วิธีคำนวณเหมาะสมสำหรับการกำหนดโดเมนของวิชาคณิตศาสตร์ (สร้างจาก A_3B_3)
2. ฟอรัมข้อสอบเหมาะสมสำหรับการกำหนดโดเมนของวิชาสังคม (สร้างจาก A_2B_5)

1.4 ประโยชน์และข้อจำกัดของรูปแบบฟาเซท

เอ็นเจล และ มาร์ตูซา (Engel and Martuza, 1976 quoted in Roid and Haladyna, 1982) ได้เสนอประโยชน์ของประโยคจับคู่ไว้ดังนี้

1. ข้อกระทง คำถาม และ ตัวลวง สามารถเขียนเพิ่มขึ้นอย่างมีระบบ โดยที่ตัวลวงสามารถเลือกมาจากสมาชิกของแต่ละฟาเซทได้ รอยด์และฮาลาดินา (Roid and Haladyna, 1980) กล่าวว่า ในการเขียนตัวลวงนั้นแต่เดิมนับได้ว่าเป็นปัญหาใหญ่ปัญหาหนึ่งของผู้เขียนข้อสอบแบบเลือกตอบ และยังขึ้นกับความชำนาญของผู้เขียนข้อสอบแต่ละคน แต่กั

สามารถทำให้ง่ายขึ้นได้โดยการเขียนตัวลวงอย่างเป็นระบบของประโยคจับคู่

2. ในตัวลวงแต่ละตัวจะมีความเกี่ยวข้องกัน แต่จะมีบางตัวที่น่าจะเป็นตัวถูกมากกว่าตัวอื่น ดังนั้นการที่นักเรียนเลือกตอบตัวลวงใดตัวลวงหนึ่ง ครูย่อมวินิจฉัยได้ว่านักเรียนที่ตอบโดยเลือกตัวลวงที่มีเหตุผลน้อยกว่า ควรจะได้รับการสอนซ่อมเสริมมากกว่านักเรียนที่เลือกตัวลวงที่มีเหตุผลมากกว่า

3. ในการเขียนข้อสอบจากประโยคจับคู่ นั้น ทำให้ข้อสอบที่เขียนออกมานั้นมีความสัมพันธ์เชิงตรรกศาสตร์ (Logical Relationship) กับเนื้อหาที่ใช้ในการสอน ประโยคจับคู่เป็นวิธีการพื้นฐานของทั้งการเรียน การสอน และการสอบ

4. การสร้างแบบสอบเลือกตอบหรือแบบสอบคู่ขนานนั้น สามารถทำได้ง่าย ข้อสอบที่เลือกมาเป็นแบบสอบแต่ละฉบับนั้น สามารถสุ่มได้มาโดยปราศจากความลำเอียง และมีความสัมพันธ์กันทั้งเนื้อหา ค่าความยาก และค่าอำนาจจำแนก

5. สามารถใช้เครื่องคอมพิวเตอร์ในการเขียนข้อสอบได้ (Millman, 1980)

6. ขั้นตอนแรกของรูปแบบฟาเซทนั้น เริ่มจากการใช้จุดประสงค์การเรียนการสอน นำมาเป็นพื้นฐานในการสร้างประโยคจับคู่ ถ้านักวางแผนหลักสูตรการศึกษาสามารถรวมจุดประสงค์การเรียนการสอน แล้วแปลงให้เป็นประโยคจับคู่ วิธีการนี้จะทำให้การกำหนดขอบเขตของรูปแบบฟาเซทได้ง่ายขึ้น

บุญเชิด ภิญญอนันตพงษ์ (2527) ได้กล่าวถึงข้อดีและข้อจำกัดของรูปแบบฟาเซท

ไว้ดังนี้

ข้อดี

1. การวิเคราะห์เนื้อหาวิชาใช้การวิเคราะห์เชิงทฤษฎีหรือหลักเหตุผลพิจารณา ไม่ต้องวิเคราะห์จากผลการปฏิบัติ
2. ใช้ประโยคจับคู่เป็นตัวช่วยนิยามมวลความรู้ในเนื้อหาวิชา และเซตของข้อสอบที่เกี่ยวข้องกันที่ใช้วัดผลสัมฤทธิ์ของมวลความรู้นั้น
3. ง่ายและสะดวกกว่าการเขียนฟอร์มข้อสอบ
4. เหมาะสมสำหรับใช้เขียนข้อสอบถูกผิด หรือเติมคำมากที่สุด และสามารถดัดแปลงไปเป็นข้อสอบแบบเลือกตอบได้

ข้อจำกัด

1. ไม่มีกฎเกณฑ์ใดๆ ที่จะช่วยในการสร้างประโยคจับคู่ ทำให้ได้ประโยคที่แตกต่างกันขึ้นอยู่กับผู้เขียน ซึ่งทำให้การเขียนข้อสอบแบบนี้ ยังคงต้องอาศัยประสบการณ์ส่วนตัว
2. เหมาะสำหรับบางเนื้อหาวิชา
3. เซทของประโยคที่นำมาใช้จับคู่เพียงเซทเดียวอาจไม่เป็นตัวแทนของเนื้อหาที่ต้องการวัดทั้งหมด
4. ประโยคที่ใช้จับคู่จะบังคับให้ได้ฟอร์มข้อสอบเพียงฟอร์มเดียว ไม่สามารถประดิษฐ์ข้อสอบให้มีฟอร์มข้อสอบตามที่ต้องการได้

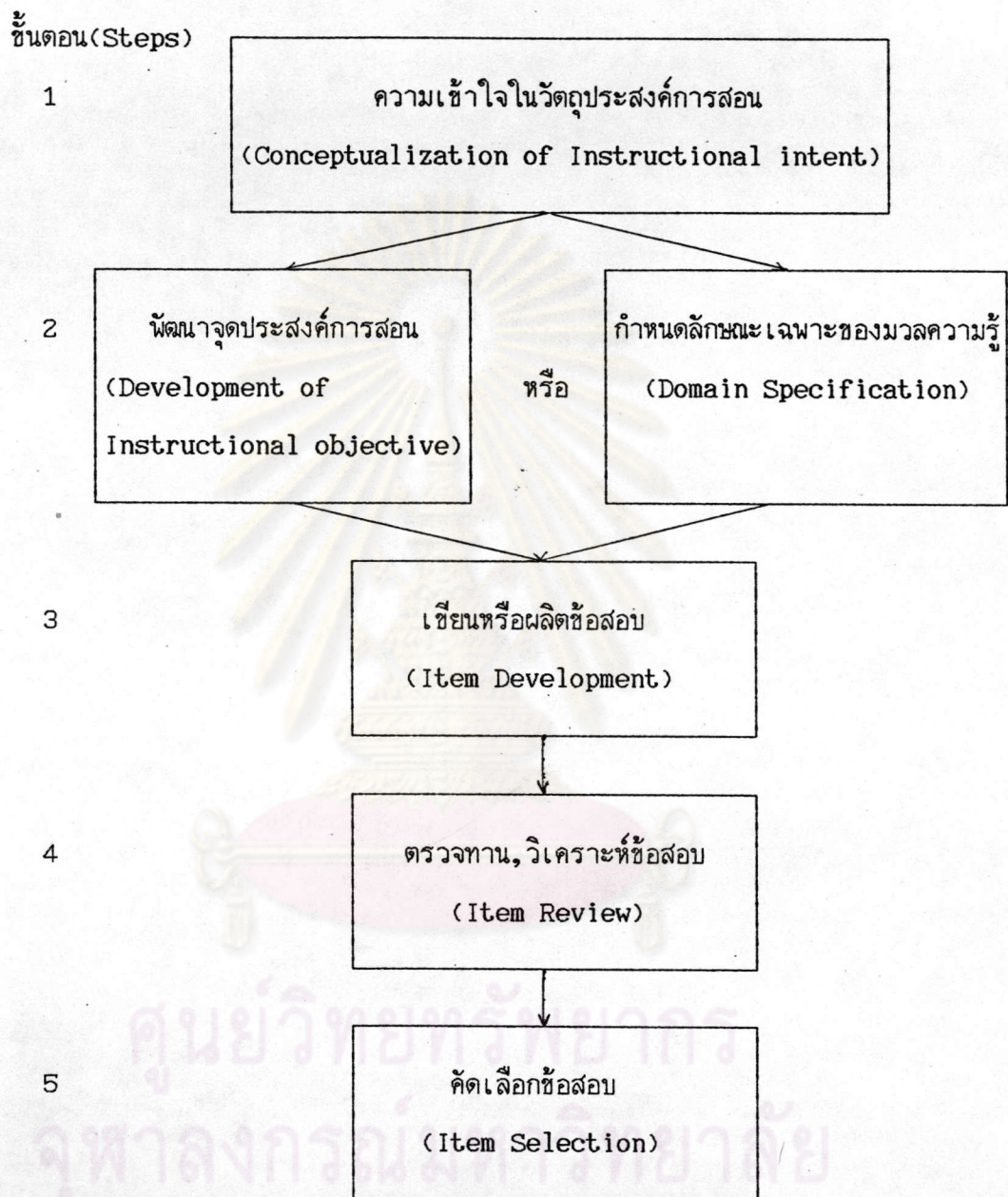
2. หลักการสร้างและการตรวจสอบคุณภาพของแบบสอบอิงเกณฑ์

2.1 หลักการสร้างแบบสอบอิงเกณฑ์

รอยด์และฮาลาดายนา (Roid and Haladyna, 1980) ได้เสนอหลักการเขียนข้อสอบเพื่อพัฒนาแบบสอบอิงเกณฑ์ ซึ่งมีสาระสำคัญ 5 ขั้นตอนพอสรุปได้ดังนี้

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ภาพที่ 1 ขั้นตอนการสร้างและพัฒนาแบบสอบอิงเกณฑ์



จากภาพที่ 1 จะกล่าวถึงรายละเอียดของแต่ละขั้นตอนดังนี้

ขั้นตอนที่ 1 ความเข้าใจในวัตถุประสงค์การสอน นับเป็นขั้นตอนแรกในการสร้างข้อสอบ ที่ผู้สร้างต้องทำความเข้าใจในจุดประสงค์ที่ต้องการวัดในหัวข้อนั้นๆ การวิเคราะห์วัตถุประสงค์หลักนี้ อาจต้องใช้วิธีการวิเคราะห์งาน (Task Analysis or

Job Analysis)

ขั้นตอนที่ 2 การพัฒนาจุดประสงค์การสอน หรือการกำหนดลักษณะเฉพาะของมวลความรู้ ขั้นตอนนี้จะเป็นการแปลงวัตถุประสงค์หลักจากขั้นตอนที่ 1 ให้อยู่ในรูปของวัตถุประสงค์เชิงพฤติกรรม ซึ่งจะบ่งบอกพฤติกรรมของผู้เรียนที่ได้จากการสอน หรือเป็นการกำหนดลักษณะเฉพาะของมวลความรู้

ขั้นตอนที่ 3 การเขียนหรือผลิตข้อสอบ เป็นขั้นตอนที่ใช้เทคนิคการเขียนข้อสอบ โดยเขียนให้สอดคล้องกับจุดมุ่งหมายเชิงพฤติกรรม หรือมวลความรู้ที่กำหนดไว้ ขั้นตอนนี้ อาจนำเทคโนโลยีการเขียนข้อสอบมาใช้ได้

ขั้นตอนที่ 4 ตรวจสอบและวิเคราะห์ข้อสอบ เป็นขั้นตอนในการตรวจสอบข้อบกพร่องของข้อสอบ นำข้อสอบไปทดลองใช้และปรับปรุงข้อสอบก่อนนำไปใช้จริง

ขั้นตอนที่ 5 คัดเลือกข้อสอบ เป็นขั้นตอนสุดท้ายในการพัฒนาแบบสอบที่จะใช้ คัดเลือกข้อสอบที่มีคุณภาพไว้ใช้ต่อไป

ตามแนวคิดของแบบสอบอิงเกณฑ์ นอกจากจะต้องกำหนดวัตถุประสงค์ และขอบเขตของเนื้อหาที่ต้องการจะสรุปพาดพิง ไปถึงให้ชัดเจนแล้ว ยังมีความจำเป็นที่จะต้องระบุถึงขอบเขตประชากรของข้อสอบให้เด่นชัดอีกด้วย แนวคิดพื้นฐานดั้งเดิมง่าย ๆ ที่จะทำให้ได้ข้อสอบเป็นจำนวนมากพอที่จะเชื่อถือได้ว่าครอบคลุมเนื้อหาทั้งหมดที่ต้องการศึกษา คือ การสร้างธนาคารข้อสอบ (Shoemaker, 1975) โดยอาจใช้วิธีระดมนักเขียนข้อสอบมาช่วยกันเขียนและเก็บรวบรวมไว้ อย่างไรก็ตามวิธีการดังกล่าวจะมีปัญหาในประเด็นที่ว่า ผู้สร้างข้อสอบต่างคนกัน แม้จะได้รับการมอบหมายให้สร้างข้อสอบที่วัตถุประสงค์เดียวกัน แต่ข้อสอบที่เขียนออกมาก็อาจจะมีความแตกต่างกันอยู่มาก จึงได้มีการพยายามพัฒนาแนวคิดการสร้างข้อสอบที่มีความเป็นปรนัยมากขึ้น ซึ่งรอยด์และฮาလာไดนาได้สรุปแนวคิดการสร้างข้อสอบในแบบสอบอิงเกณฑ์ที่มีความเป็นปรนัยไว้ ซึ่งเรียกว่า "เทคโนโลยีการเขียนข้อสอบ" ดังที่กล่าวมาแล้ว

2.2 การตรวจสอบคุณภาพของแบบสอบอิงเกณฑ์

ในการพัฒนาแบบสอบอิงเกณฑ์ สิ่งสำคัญประการแรกจะเน้นหนักในด้านความตรงเชิงเนื้อหา ซึ่งรอยด์และฮาลาดิโน (Roid and Haladyna, 1982) ได้เสนอขั้นตอนการตรวจสอบคุณภาพของแบบสอบอิงเกณฑ์ไว้ 2 ขั้นตอน คือ

(1) การตรวจสอบความสอดคล้องเชิงตรรก (Logical Review of Criterion-Referenced Test Items) เป็นการตรวจสอบคุณภาพของแบบสอบในด้านความสอดคล้อง (Consistency) และความพอเพียง (Adequacy) ของข้อสอบกับวัตถุประสงค์การสอน โดยให้ผู้เชี่ยวชาญในด้านเนื้อหาเป็นผู้ประเมิน ซึ่งมีการดำเนินการเป็น 3 ขั้นตอนดังนี้

1. ตรวจสอบความตรงเชิงเนื้อหา รายข้อ โดยการหาค่าดัชนีความสอดคล้องระหว่างข้อสอบกับจุดประสงค์ (Item-Objective Consistency) จะกระทำได้โดยผู้เชี่ยวชาญในเนื้อหานั้นๆ พิจารณาตัดสิน ซึ่งใช้วิธีการของโรวินลลีและแฮมเบิลตัน (Rovinelli and Hambleton, 1977)

2. การตรวจสอบคุณภาพเชิงเทคนิคการเขียนข้อสอบ (Technical Quality of Item Writing) ประเด็นนี้มุ่งเน้นการเขียนข้อสอบให้มีคุณภาพในแง่ของหลักการเขียนข้อสอบโดยทั่วไป (ดูตัวอย่างได้ใน Hambleton and Eignor, 1977 และใน Berk, 1980)

3. ความเป็นตัวแทนของข้อสอบในแบบสอบ (Representativeness of the Test Items) จำนวนข้อสอบที่รวมกันเข้าเป็นแบบสอบเต็มฉบับ จะต้องมีความเป็นตัวแทนของความรู้ในขอบเขตเนื้อหาทั้งหมด

(2) การตรวจสอบเชิงประจักษ์ (The Empirical Review of Test Items) ขั้นตอนนี้เป็นขั้นตอนการนำแบบสอบที่สร้างไปทดลองใช้ เพื่อวิเคราะห์หาคุณภาพข้อสอบรายข้อ (Item Analysis) ตามแนวการวัดแบบอิงเกณฑ์ จุดมุ่งหมายหลักของการวิเคราะห์ข้อสอบไม่ใช่เป็นการคัดเลือกข้อสอบ แต่เป็นการตรวจสอบคุณภาพของข้อสอบในมวลความรู้ ดังนั้น การตรวจสอบในขั้นนี้ จะทำการวิเคราะห์เพื่อ หาค่าความยาก และค่าอำนาจจำแนกของข้อสอบ เพื่อตรวจสอบข้อบกพร่องและนำไปปรับปรุงแก้ไข ดังนี้

1. ค่าความยากรายข้อ (Item Difficulty) ค่าความยากรายข้อของข้อสอบเป็นดัชนีที่บ่งบอกได้ว่าข้อสอบมีความยากง่ายเพียงใด ซึ่งคำนวณได้จากสัดส่วนของผู้สอบที่ตอบข้อนั้นถูก ในการทดสอบแบบอิงเกณฑ์ ค่าความยากหลังสอนควรมีค่าสูงซึ่งเป็นการบ่งบอกว่า ข้อสอบง่าย หลังการสอนค่าความยากควรมีค่าเป็น .75 ขึ้นไป และก่อนสอนค่าความยากควรมีค่าต่ำกว่า .40 (Roid and Haladyna, 1982) แต่โดยทั่วไป อาจถือเกณฑ์แบบอิงกลุ่มก็ได้ (.20 ถึง .80)

2. ค่าอำนาจจำแนก (Item Discrimination) ดัชนีค่าอำนาจจำแนกเป็นการบอกความสัมพันธ์ของข้อสอบ(item) กับ ผลสอบ(test performance) ถ้าพิสัยของคะแนนสอบแคบค่าอำนาจจำแนกจะต่ำ และถ้าพิสัยของคะแนนกว้างค่าอำนาจจำแนกจะสูง (Roid and Haladyna, 1974) ค่าอำนาจจำแนกในแบบสอบอิงเกณฑ์นั้นเป็นการยากที่จะแปลผล ค่าอำนาจจำแนกที่ติดลบหรือมีค่าเป็นศูนย์ ส่วนหนึ่งอาจมาจากความบกพร่องของข้อสอบ แต่ในความเป็นจริงแล้วอาจเป็นผลจากการสอน ในการทดสอบแบบอิงเกณฑ์ให้ความสำคัญกับค่าอำนาจจำแนกน้อย เนื่องจากไม่สนใจความแตกต่างระหว่างกลุ่ม แต่สนใจในการสร้างข้อสอบให้เป็นตัวแทนของมวลความรู้นั้นๆ

ในการสอบแบบอิงเกณฑ์ ได้มีการพัฒนาดัชนีประจำข้อสอบขึ้นมาใช้โดยเฉพาะตัวหนึ่งคือ "ดัชนีความไวในการสอน" (Instructional Sensitivity Index ตัวย่อ ISI) ซึ่งเป็นดัชนีชี้พัฒนาการของกลุ่มผู้เรียนในการทำข้อสอบข้อหนึ่ง ๆ ได้ถูกต้อง ซึ่งฮาลาดิโนนาและรอยด์ (Haladyna and Roid, 1981) ได้ศึกษาการคำนวณค่า ISI 4 แบบ คือ

1. แนวคิดของการสอบแบบอิงเกณฑ์ (Criterion-Referenced) มีวิธีการคำนวณ 2 วิธี คือ

- ความแตกต่างระหว่างสัดส่วนของผู้ตอบถูกหลังการสอน กับก่อนการสอน (Pretest to Posttest Difference Index, PPDI) ซึ่งเสนอโดย คอกซ์และวาร์กัส (Cox and Vargas, 1966) ซึ่งปกติค่า PPDI มีค่าอยู่ระหว่าง -1.00 ถึง +1.00 ในการกำหนดเพื่อวัดจุดประสงค์การสอน ค่า PPDI ควรมีพิสัยตั้งแต่ .10 ถึง .60 (Roid and Haladyna, 1982) ซึ่งการคำนวณโดยวิธีนี้ใช้วิธีการเช่นเดียวกันกับการหาดัชนีความไว ของคริสปินและเฟลฮูเซน (Kryspin and Felhusen, 1974 อ้างถึงใน สำเร็จบุญเรืองรัตน์, 2527) และใช้วิธีการเช่นเดียวกันกับการหาค่าดัชนีบี (Brennan Index)

- การหาค่าเปอร์เซ็นต์ที่เพิ่มขึ้น (The Percent of Possible Gain, PPG) ซึ่งคำนวณได้จากสูตร

$$PPG = \frac{\text{ค่าความยาก(หลังเรียน)} - \text{ค่าความยาก(ก่อนเรียน)}}{1 - \text{ค่าความยาก(ก่อนเรียน)}}$$

2. แนวคิดของการวิเคราะห์ข้อสอบในแนวคิดเดิม (Classical Test Theory) แนวคิดนี้ ใช้การหาค่าสัมประสิทธิ์สหสัมพันธ์แบบพอยท์ไบซีเรียล ระหว่างคะแนนจากแบบสอบ เดิมฉบับกับคะแนนประจำข้อสอบที่เป็น 0 กับ 1 โดยกลุ่มตัวอย่างที่ใช้เกิดจากการรวมกลุ่ม ผู้สอบที่ได้รับการสอน และกลุ่มผู้สอบที่ยังไม่ได้รับการสอนเข้าด้วยกัน (Combined-Samples Point Biserial Correlation, COMPBI)

3. แนวคิดของทฤษฎีการตอบสนองข้อสอบ (Item Response Theory) แนวคิดนี้ ใช้ค่าผลต่างของค่าความยากประจำข้อสอบ ที่ได้รับการปรับให้เป็นค่าคะแนน มาตรฐานใต้พื้นโค้งปกติ (Normalized difference between item difficulty, ZDIFF) เป็นดัชนี

4. แนวคิดจากทฤษฎีของเบส์ (Bayesian) ใช้การหาค่าดัชนี 3 ตัวคือ

- การหาค่าโอกาสที่ผู้ตอบข้อสอบข้อหนึ่งๆถูก จะเป็นผู้รอบรู้ที่แท้จริง (B1)
- การหาค่าโอกาสที่ผู้ตอบข้อสอบข้อหนึ่งๆผิด จะเป็นผู้รอบรู้ไม่จริง (B2)
- การหาค่าโอกาสที่จะตัดสินใจได้ถูกต้องว่าผู้สอบเป็นผู้รอบรู้หรือไม่รอบรู้ (B3)

ทั้งนี้ดัชนีแต่ละตัว มีวิธีการคำนวณคือ

$$B1 = \frac{(POSTDIFF)(COMDIFF)}{(POSTDIFF)(COMDIFF) + (PREDIFF)(1 - COMDIFF)}$$

$$B2 = \frac{(1 - PREDIFF)(1 - COMDIFF)}{(1 - PREDIFF)(1 - COMDIFF) + (1 - POSTDIFF)(COMDIFF)}$$

$$B3 = (POSTDIFF - COMDIFF) + (1 - PREDIFF) + (COMDIFF - POSTDIFF)$$

เมื่อ

POSTDIFF คือ ค่าความยากของข้อสอบหลังการสอบ

PREDIFF คือ ค่าความยากของข้อสอบก่อนการสอบ

COMDIFF คือ ค่าความยากของข้อสอบเมื่อนำผลการสอบก่อนและหลังการสอบมารวมกัน

ฮาลาไดนาและรอยด์ ได้เสนอว่า ดัชนี ISI ที่มีประสิทธิภาพสูงและคำนวณได้ง่าย คือ PFDI แต่ก็ได้กล่าวย้าว่า การตัดสินใจเลือกข้อสอบในการสร้างแบบสอบอิงเกณฑ์นั้น ไม่ควรจะพิจารณาเฉพาะค่าประจำข้อสอบเท่านั้น แต่ควรพิจารณา ความสัมพันธ์ระหว่างเนื้อหาที่ข้อสอบมุ่งวัด กับการจัดการเรียนการสอนจริงประกอบด้วย

2.3 การกำหนดเกณฑ์มาตรฐาน

แนวคิดของแบบสอบอิงเกณฑ์นั้น เป็นแบบสอบที่มีการกำหนดเกณฑ์มาตรฐาน หรือ จุดตัด (Cut-off Score) ขึ้นไว้ เพื่อใช้ในการจำแนกผู้สอบออกเป็นสองส่วน คือ ผู้สอบที่ได้คะแนนสูงกว่าเกณฑ์มาตรฐาน จะถือว่าเป็นผู้รอบรู้ (Master) และผู้ที่ได้คะแนนต่ำกว่าเกณฑ์มาตรฐาน จะถือว่าเป็นผู้ไม่รอบรู้ (Nonmaster)

การกำหนดเกณฑ์มาตรฐานนับได้ว่าเป็นส่วนสำคัญของแบบสอบอิงเกณฑ์ ได้มีผู้นำเสนอแนวคิด เทคนิค และวิธีการต่างๆ ในการกำหนดเกณฑ์มาตรฐานไว้จำนวนมาก พอสรุปได้ดังนี้

มิลแมน (Millman, 1975) ได้สรุปแนวคิดของการกำหนดเกณฑ์มาตรฐานไว้ว่า ควรจะคำนึงถึงสิ่งต่างๆ ต่อไปนี้

1. ระดับความสามารถในการทำแบบสอบของคนอื่นๆ (Performance of Others) โดยเฉพาะผลการสอบของกลุ่มที่ผ่านการยอมรับแล้ว ย่อมนำมาพิจารณาในการกำหนดเกณฑ์มาตรฐานได้

2. เนื้อหาสาระของข้อสอบในแบบสอบ (Item Content) ควรพิจารณาตัดสินใจว่า ปัญหาต่างๆ หรือจุดมุ่งหมายที่นำมาทำการทดสอบนั้น ผู้สอบที่มีความสามารถปานกลาง ควรจะทำได้ถูกต้องสักเท่าใด จึงจะอยู่ในระดับที่พอใจ

3. คำนึงถึงผลกระทบที่ตามมา (Educational Consequences) บางครั้งควรคำนึงว่า การกำหนดเกณฑ์มาตรฐานสูงเกินไป อาจทำให้ผู้เรียนไม่ผ่านโดยไม่จำเป็น

แต่ถ้าหากว่าต่ำเกินไป ก็จะทำให้มาตรฐานการศึกษาต่ำลง หรือทำให้ผู้เรียนเกิดปัญหาในการเรียนต่อไป

4. ผลทางจิตวิทยาและการลงทุนทางการศึกษา (Psychological and Financial Cost) การกำหนดเกณฑ์มาตรฐานควรคำนึงถึงผลทางจิตวิทยาที่เกิดกับผู้สอบ เช่น แรงจูงใจ ความเบื่อหน่าย การทำลายอัตรานัดขึ้น และรวมทั้งค่าใช้จ่ายในการสอนซ่อมเสริมผู้เรียนที่สอบไม่ผ่าน

5. ความคลาดเคลื่อนในการเดาและการสุ่มข้อสอบ (Error due to Chance and Item Sampling) การกำหนดเกณฑ์มาตรฐานควรคำนึงถึง โอกาสที่ผู้สอบจะเดาถูก รวมทั้งลักษณะการสุ่มข้อสอบจากประชากรข้อสอบในขอบเขตเนื้อหาที่กำหนดขึ้นด้วย

การกำหนดเกณฑ์มาตรฐานของแบบสอบอิงเกณฑ์ ได้มีผู้เสนอวิธีการต่างๆไว้อย่างกว้างขวาง กลาส (Glass, 1978) ได้รวบรวมวิธีการหาเกณฑ์มาตรฐานได้ 6 วิธี คือ

1. การพิจารณาความสามารถของผู้อื่น (Performance of Others as a Criterion) วิธีนี้ไม่แตกต่างไปจากแนวคิดของการประเมินแบบอิงกลุ่ม เพราะเป็นการนำเอาข้อมูลจากผู้สอบคนอื่น มาใช้เป็นหลักพิจารณาในการกำหนดเกณฑ์มาตรฐาน

2. การนับถอยหลัง (Counting Backward from 100%) วิธีนี้จะให้ผู้เชี่ยวชาญทางด้านเนื้อหาวิชา พิจารณาตามความสำคัญของจุดประสงค์และกำหนดค่าร้อยละตามลำดับความสำคัญ

3. การใช้เกณฑ์ภายนอกเข้ามาช่วยในการตัดสินใจ (Bootstrapping on Other Criterion Scores) วิธีนี้จะใช้ได้ก็ต่อเมื่อ ตัวเกณฑ์ภายนอกกับคะแนนจากแบบสอบอิงเกณฑ์มีความสัมพันธ์กันอย่างสมบูรณ์ มิฉะนั้นจะเกิดความคลาดเคลื่อนได้

4. การกำหนดสมรรถภาพต่ำสุด (Judging Minimal Competence) วิธีนี้ให้ผู้เชี่ยวชาญเป็นผู้ร่วมในการกำหนดเกณฑ์มาตรฐานต่ำสุดที่ยอมรับได้ เช่น วิธีการกำหนดเกณฑ์มาตรฐานของ อีเบล แองกอฟฟ์ และนี่เดลกี

5. การใช้ทฤษฎีการตัดสินใจประกอบ (Decision Theoretic Approach) วิธีการนี้มุ่งหามาตรการทางสถิติเพื่อนำมาใช้เป็นวิธีการในการกำหนดเกณฑ์มาตรฐาน เช่น การกำหนดเกณฑ์มาตรฐานที่จะทำให้เกิดความคลาดเคลื่อนในการจัดประเภทบุคคลตามระดับการรอบรู้ต่ำสุด ได้แก่วิธีการของกลาส (Glass, 1978) และวิธีการของเบิร์ก (Berk, 1976)

6. การใช้วิธีการวิจัยเชิงปฏิบัติการ (Operations Research Method) วิธีการนี้เป็น การนำเอาคุณค่าภายนอกที่เหมาะสมมาสัมพันธ์กับคะแนนที่ได้จากแบบสอบ เช่น การพิจารณาถึงทัศนคติต่อวิชาที่เรียน ซึ่งจะผันแปรไปตามเกณฑ์มาตรฐานของแบบสอบ

ในการกำหนดเกณฑ์มาตรฐานในปัจจุบันยังไม่มีวิธีใดที่ดีที่สุด คงจะต้องทำเป็น เฉพาะกรณีไป โดยคำนึงถึงข้อมูลหลาย ๆ ด้าน และเมื่อได้เกณฑ์มาตรฐานมาแล้ว ก็ควรจะใช้ เกณฑ์นั้นอย่างมีความเป็นปรนัย

2.4 ความตรง(Validity)และความเที่ยง(Reliability) ของแบบสอบอิงเกณฑ์

2.4.1 การหาความตรงของแบบสอบอิงเกณฑ์

แฮมเบิลตัน (Hambleton in Berk, 1980) ได้อภิปรายถึงประเด็นต่างๆ ที่เกี่ยวกับความตรงของแบบสอบอิงเกณฑ์ไว้ โดยมีสาระสำคัญดังนี้

ความตรงตามเนื้อหา (Content Validity) แบบสอบอิงเกณฑ์ที่มุ่งเน้นความตรงตามเนื้อหา มีเงื่อนไข 2 ประการ คือ ประการแรก ข้อสอบทั้งหมดในขอบเขตของเนื้อหาที่กำหนดไว้ได้รับการเขียนไว้ล่วงหน้า และรู้ขอบเขตอย่างเด่นชัด อีกประการหนึ่งคือ ข้อสอบทั้งหมดในแบบสอบได้มาจากการสุ่ม และเพื่อให้ได้แบบสอบที่มีคุณภาพในแง่ของความตรงเชิงเนื้อหา แบบสอบจะต้องประกอบขึ้นด้วยข้อสอบที่มีคุณภาพดีมาก่อน ซึ่งพิจารณาได้จากคุณสมบัติ 3 ประการคือ

- (1) มีความตรงของข้อสอบ (Item Validity)
- (2) มีการตรวจสอบคุณภาพเชิงเทคนิคของการเขียนข้อสอบ
- (3) มีความเป็นตัวแทนของข้อสอบในแบบสอบ

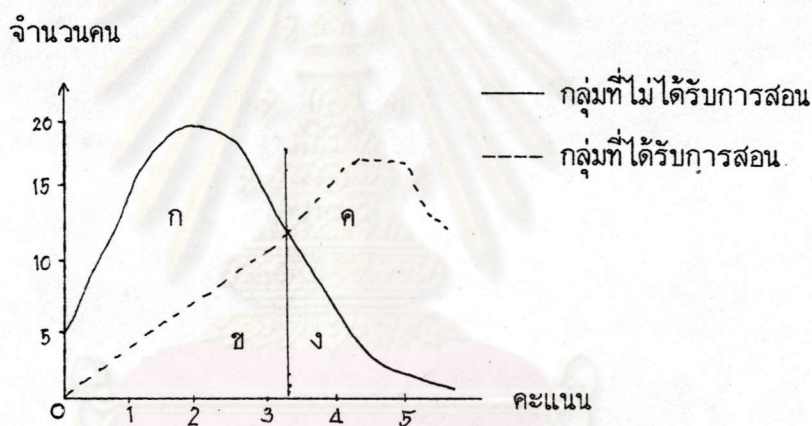
ความตรงตามเกณฑ์ของแบบสอบ (Criterion-Related Validity)

แฮมเบิลตัน ใช้คำนี้ในความหมายเช่นเดียวกับคำว่า "ความตรงอันเนื่องมาจากเกณฑ์มาตรฐานของแบบสอบ" คือ มุ่งในประเด็นที่ว่า จากจุดเกณฑ์มาตรฐานที่กำหนดและใช้เป็นจุดแบ่งผู้สอบออกเป็นผู้รอบรู้และไม่รอบรู้นั้น กลุ่มผู้เรียนที่ได้รับการแบ่งออกไปน่าจะต้องมีความแตกต่าง

ในเกณฑ์ภายนอกอื่นๆ เช่น การทำนายผลการเรียนในบทเรียนต่อไป ซึ่งเทคนิคหรือวิธีการที่จะใช้ในการศึกษา ก็จะไม่แตกต่างกันไปจากวิธีการต่างๆ ที่มีใช้กันอยู่ทั่วไป

ความตรงในการตัดสินใจ (Decision Validity) ความตรงประเภทนี้จัดเป็นความตรง เฉพาะชนิดหนึ่งของความตรงตามโครงสร้างทฤษฎี วิธีการตรวจสอบความตรงในการตัดสินใจของแบบสอบถามเชิงเกณฑ์ จะทำได้โดยการเปรียบเทียบการทำแบบสอบของกลุ่มที่ได้รับการสอนกับกลุ่มที่ไม่ได้รับการสอน ซึ่งพิจารณาได้จากภาพข้างล่างนี้

ภาพที่ 2 การพิจารณาความตรงในการตัดสินใจจากกลุ่มผู้สอบที่ได้รับการสอน และกลุ่มผู้สอบที่ไม่ได้รับการสอน



- เมื่อ ก คือ กลุ่มผู้สอบที่ไม่ได้รับการสอนและเป็นผู้ไม่รอบรู้ เป็นการตัดสินใจที่ถูกต้อง
 ข คือ กลุ่มผู้สอบที่ได้รับการสอนแต่เป็นผู้ไม่รอบรู้ เป็นการตัดสินใจที่ผิด
 ค คือ กลุ่มผู้สอบที่ได้รับการสอนและเป็นผู้รอบรู้ เป็นการตัดสินใจที่ถูกต้อง
 ง คือ กลุ่มผู้สอบที่ไม่ได้รับการสอนแต่เป็นผู้รอบรู้ เป็นการตัดสินใจที่ผิด

ความตรงในการตัดสินใจนี้ ขึ้นอยู่กับปัจจัยต่างๆคือ (1) คุณภาพของแบบสอบ (2) ความเหมาะสมของกลุ่มผู้สอบที่นำมาใช้ในการกำหนดเกณฑ์มาตรฐาน (3) ลักษณะของกลุ่มที่มุ่งจำแนกจัดประเภทระดับการรอบรู้ (4) เกณฑ์มาตรฐานต่ำสุดที่กำหนด

เครเฮิน (Crehen, 1974) ได้อ้างถึงสูตรของคาร์เวอร์ ที่ใช้ในการหาค่าความตรง ของแบบสอบถามเชิงเกณฑ์ เมื่อแบ่งผู้สอบออกเป็น 4 กลุ่ม ดังตาราง

ตารางที่ 2 แนวคิดการหาค่าความตรงของแบบสอบอิงเกณฑ์โดยใช้สูตรของคาร์เวอร์

ผลการตัดสิน	สภาวะการสอน	ไม่ได้รับการสอน	ได้รับการสอน
รอบรู้		b	a
ไม่รอบรู้		c	d

ความตรงในการตัดสินใจจัดประเภทบุคคลตามระดับการรอบรู้ จะคำนวณได้จากสูตร

$$\text{ค่าความตรง} = \frac{a+c}{a+b+c+d}$$

2.4.2 การหาค่าความเที่ยงของแบบสอบอิงเกณฑ์

เนื่องจากแบบสอบอิงเกณฑ์ไม่ได้มุ่งเน้นการจำแนกผู้สอบในกลุ่มเดียวกัน และทั้งยังเป็นไปได้ว่า ผู้สอบอาจจะได้คะแนนแตกต่างกันไม่มากนัก ซึ่งกรณีเช่นนี้ย่อมมีแนวโน้มที่ว่า แบบสอบที่สร้างขึ้นมานั้น จะมีค่าความเที่ยงที่หาโดยยึดแนวคิดทฤษฎีดั้งเดิมต่ำ ทั้งๆ ที่ความเป็นจริงแล้ว แบบสอบนั้นมีคุณภาพดี จึงได้มีผู้พยายามพัฒนาสูตรการหาค่าความเที่ยงซึ่งเหมาะสมที่จะใช้กับแบบสอบอิงเกณฑ์ขึ้นมาใหม่ สับโคเวียค (Subkoviak in Berk, 1980) ได้กล่าวถึง การหาค่าความเที่ยงตามแนวคิดในการจัดประเภทบุคคล (Reliability of Mastery Classification) ไว้ดังนี้

(1) วิธีของคาร์เวอร์ (Carver Method) มีแนวคิดที่ว่า ความเที่ยงของแบบสอบอิงเกณฑ์ ควรจะหาได้จาก ความคงที่ในการจัดประเภทบุคคลตามระดับการรอบรู้ โดยยึดเอาเกณฑ์มาตรฐานที่กำหนดเป็นจุดแบ่ง และใช้แบบสอบคู่ขนานในการทดสอบเพื่อหาความเที่ยง การพิจารณาแบบสอบคู่ขนาน 2 ฉบับที่เชื่อถือได้ จะดูจาก จำนวนผู้ผ่านเกณฑ์จากการทำแบบสอบทั้ง 2 ฉบับนั้น เท่ากัน และเป็นกลุ่มเดียวกัน ซึ่งเป็นไปได้ยาก ปัญหา

ของการหาความเที่ยง โดยวิธีนี้จะไม่มีควมไวในการหาความคงที่เกี่ยวกับการจำแนกบุคคล จึงไม่นิยมใช้กันอย่างแพร่หลาย อย่างไรก็ตามวิธีนี้ก็ยังคงเป็นแนวทางในการหาความเที่ยงของวิธีอื่นต่อไป

(2) วิธีของสวามินาธาน แยมเบิลตันและอัลจินา (Swaminathan, Hambleton and Algina Method) เป็นวิธีที่พัฒนามาจากวิธีของคาร์เวอร์ เป็นความคงที่ในการจำแนกผู้รอบรู้จากการสอบด้วยแบบสอบสองฉบับที่คู่ขนานกัน เป็นวิธีการคำนวณที่จัดความสอดคล้องโดยบังเอิญออกไป ซึ่งสวามินาธานและคณะ ได้เสนอแนวคิดในการหาความเที่ยงวิธีนี้ โดยการใช้ดัชนีแคปปา (\hat{K}) ซึ่งมีสูตรในการคำนวณดังนี้

$$\hat{K} = \frac{P_o - P_c}{1 - P_c}$$

เมื่อ \hat{P}_o คือ สัดส่วนของความสอดคล้องที่สังเกตได้จากการสอบซ้ำ คำนวณได้จาก

$$\hat{P}_o = \sum_{k=1}^m \hat{P}_{kk}$$

\hat{P}_c คือ สัดส่วนของความสอดคล้องที่คาดหวังได้จากความบังเอิญ คำนวณได้จาก

$$\hat{P}_c = \sum_{k=1}^m \hat{P}_{k.} \hat{P}_{.k}$$

เมื่อ $\hat{P}_{k.}$ และ $\hat{P}_{.k}$ คือ สัดส่วนของผู้สอบที่ถูกกำหนดลงในชั้น

ของการเรียนรู้ k ในการสอบครั้งที่ 1 และ 2 ตามลำดับ

การหาค่าดัชนีแคปปา สามารถที่จะใช้ได้ทั้งกรณีที่มีการจัดประเภทบุคคลออกไปมากกว่าสองประเภท ทั้งยังสามารถที่จะหาค่าความคลาดเคลื่อนมาตรฐาน (Standard Error) ของค่าแคปปา ได้จากสูตร

$$b_k = \sqrt{\frac{P_o(1 - P_o)}{N(1 - P_c)^2}}$$

N คือ จำนวนผู้เข้าสอบทั้งหมด

ถ้าต้องการตรวจสอบว่า ค่าเฉลี่ยที่คำนวณได้มีค่าสูงกว่า "0" อย่างมีนัยสำคัญทางสถิติหรือไม่ สามารถที่จะคำนวณได้จากสูตร

$$Z = \frac{\hat{K}}{\sqrt{\frac{\hat{P}_c}{N(1-\hat{P}_c)}}$$

(Cohen, 1960)

อย่างไรก็ตามการหาความเที่ยงจากสูตรดังกล่าว มีความจำเป็นที่จะต้องให้แบบสอบถามหรือมีการสอบซ้ำ บางครั้งอาจไม่สะดวกในการปฏิบัติ จึงได้มีผู้พยายามพัฒนาสูตรขึ้นมาทดแทนแนวคิดนี้ โดยมุ่งให้สามารถหาความเที่ยงจากแบบสอบถามโดยใช้การทดสอบเพียงครั้งเดียว บุคคลที่นำเสนอแนวคิดนี้ได้แก่ สับโคเวียค (Subkoviak), ฮวน (Huynh) และ มาร์แชลล์ (Marshall) โดยฮวนใช้หลักการของเบต้าไบนอมิเยล เป็นข้อตกลงเบื้องต้น ส่วนสับโคเวียค และมาร์แชลล์ ใช้ข้อตกลงเบื้องต้นของการแจกแจงทวินามเป็นสำคัญ ซึ่งแต่ละวิธีมีข้อดีและข้อจำกัด พอสรุปได้ดังตาราง (Berk, 1980)

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 3 ข้อดีและข้อจำกัดของการหาความเที่ยงแบบต่างๆ

วิธี	ข้อดี	ข้อจำกัด
สวามิภານและคณะ	<ul style="list-style-type: none"> - คำนวณง่าย - การประมาณค่าไม่ลำเอียง 	<ul style="list-style-type: none"> - ต้องทำการทดสอบสองครั้ง - ถ้ากลุ่มตัวอย่างมีขนาดเล็ก การประมาณค่าจะคลาดเคลื่อนมาก
ฮวน	<ul style="list-style-type: none"> - ทำการทดสอบครั้งเดียว - การประมาณค่าคลาดเคลื่อนน้อย 	<ul style="list-style-type: none"> - คำนวณยาก ใช้เวลานาน - การประมาณค่าจะลำเอียงมาก ถ้าแบบสอปสั้น
สับโคเวียค	<ul style="list-style-type: none"> - ทำการทดสอบครั้งเดียว - การประมาณค่าคลาดเคลื่อนน้อย 	<ul style="list-style-type: none"> - คำนวณยาก ใช้เวลานาน - การประมาณค่าจะลำเอียงมาก ถ้าแบบสอปสั้น
มาร์แชลล์	<ul style="list-style-type: none"> - ทำการทดสอบครั้งเดียว - การประมาณค่าคลาดเคลื่อนน้อย 	<ul style="list-style-type: none"> - คำนวณยาก ใช้เวลานาน - การประมาณค่าจะลำเอียงมาก ถ้าแบบสอปสั้น

นอกจากนี้การหาความเที่ยงของแบบสอปอิง เกณฑ์ยังยึดตามแนวคิดของการหาค่าความแตกต่างของคะแนนที่สอบได้กับคะแนนเกณฑ์มาตรฐาน โดยถือว่าระยะห่างของคะแนนที่ผู้สอบทำได้กับเกณฑ์มาตรฐานที่กำหนด มีความสำคัญในการตัดสิน ซึ่งได้แก่สูตรของ ลีฟวิงตัน (Livingston, 1972) ดังนี้

$$K^2(X, T) = \frac{6^2 T + (U-C)^2}{6^2 O + (U-C)^2}$$

เมื่อ $K^2(X, T)$ คือ ค่าความเที่ยงของแบบสอป

$6^2 T$ คือ ค่าความแปรปรวนของคะแนนที่แท้จริง คำนวณได้จากสูตร

$$6^2 T = \alpha 6^2 O$$

เมื่อ α คือ ค่าความเที่ยงของแบบสอปที่คำนวณจากสูตร KR_{20}

- 6²⁰ คือ ค่าความแปรปรวนของคะแนนที่ได้จากแบบสอบ
 U คือ ค่ามัชฌิมเลขคณิตของกลุ่มผู้เข้าสอบ
 C คือ ค่าเกณฑ์มาตรฐานที่กำหนด

3. งานวิจัยที่เกี่ยวข้อง

งานวิจัยที่เกี่ยวข้องกับการสร้างแบบสอบอิงเกณฑ์ เริ่มมีการใช้เทคโนโลยีในการเขียนข้อสอบมากขึ้น โดยเฉพาะวิธีการสร้างรูปแบบเฉพาะของข้อสอบ (Item Form) และการพัฒนาจุดประสงค์ขยายความ (Item Specification) ดังงานวิจัยต่อไปนี้

ในปี 2522 กาญจนา วัฒนสุนทร ได้ทำวิทยานิพนธ์เรื่อง การสร้างแบบสอบอิงเกณฑ์วิชาคณิตศาสตร์ เรื่องสมการ ชั้นมัธยมศึกษาปีที่ 1 ใช้วิธีการสร้างโดยกำหนดรูปแบบเฉพาะประจำข้อสอบ (Item Form) โดยทำการวิเคราะห์เนื้อเรื่องที่จะสร้างข้อสอบจากหลักสูตร นำมาสร้างโดเมนข้อสอบที่จะวัดลักษณะของหัวเรื่อง 6 โดเมนแล้วสร้างประชากรเนื้อเรื่องที่มุ่งวัดในแต่ละโดเมน โดเมนละ 40 ข้อ หาค่าความตรงของเนื้อเรื่อง โดยใช้ผู้เชี่ยวชาญตัดสิน เขียนวัตถุประสงค์เชิงพฤติกรรมเพื่อนำมาสร้างตารางวิเคราะห์หลักสูตรเพื่อกำหนดจำนวนข้อสอบ แล้วสุ่มข้อสอบจากโดเมนของข้อสอบแต่ละโดเมนเพื่อนำมาใส่ในแบบสอบย่อย 4 แบบสอบ(ที่วัดวัตถุประสงค์แต่ละข้อ) คือ ความรู้เบื้องต้นเกี่ยวกับสมการ การแทนค่าสมการ การแก้ปัญหสมการโดยใช้คุณสมบัติการบวก ลบ คูณ และหาร ความรู้ในเรื่องการเปลี่ยนใจทรงสมการเป็นสมการ และความรู้เกี่ยวกับการแก้โจทย์สมการ กลุ่มตัวอย่างเป็นนักเรียนโรงเรียนสันติราษฎร์ จำนวน 156 คน วิเคราะห์หาค่าคุณภาพรายข้อ โดยการหาค่าความยากและค่าดัชนีความไวในการสอน หาค่าความตรงโดยวิธีการของคาร์เวอร์ หาค่าความเที่ยงโดยวิธีการของลิฟวิงตัน ผลการทดสอบค่าเฉลี่ยของคะแนนหลังการเรียนและก่อนการเรียนที่ได้จากการทดสอบด้วยแบบสอบย่อย 4 ชุด แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .001 นั่นคือ แบบสอบมีความตรงตามลักษณะการวัด แบบอิงเกณฑ์

ในปี 2527 ประเทือง ทาสีแสง ได้ทำวิทยานิพนธ์เรื่อง การสร้างแบบสอบอิงโดเมนวิชาคณิตศาสตร์ เรื่องคู่อันดับและกราฟ ชั้นมัธยมศึกษาปีที่ 1 กลุ่มตัวอย่างเป็น

นักเรียนโรงเรียนประจักษ์ศิลปาคาร จำนวน 167 คน ลักษณะของการสร้างข้อสอบเป็นการสร้างขึ้นตามรูปแบบเฉพาะประจำข้อสอบที่กำหนดขึ้น แบ่งแบบสอบออกเป็น 6 ฉบับย่อย คือ ฉบับที่ 1 และ 2 เป็นแบบสอบคู่ขนานวัดเนื้อหาด้านความเข้าใจเรื่องคู่ลำดับ ฉบับที่ 3 และ 4 เป็นแบบสอบคู่ขนานวัดความเข้าใจเรื่องกราฟ ฉบับที่ 5 และ 6 เป็นแบบสอบคู่ขนานวัดเนื้อหาเรื่องการเขียนกราฟและหาคำตอบ การหาความตรงใช้วิธีการให้ผู้เชี่ยวชาญด้านเนื้อหาพิจารณา หาค่าความเที่ยงใช้วิธีการของสวามินาธานและคณะ

ในปี 2528 กำจัด เกตุสุวรรณ ได้ทำวิทยานิพนธ์เรื่อง การสร้างแบบสอบอิงโดเมนวิชาคณิตศาสตร์ เรื่องคู่ลำดับ ชั้นมัธยมศึกษาปีที่ 5 กลุ่มตัวอย่างเป็นนักเรียนโรงเรียนพิบูลวิทยาลัย ลักษณะของการสร้างข้อสอบใช้การสร้างขึ้นตามลักษณะเฉพาะประจำข้อสอบ ซึ่งสร้างเป็นแบบสอบคู่ขนาน 4 ชุด จำนวน 8 ฉบับ หาเกณฑ์ที่เหมาะสมโดยวิธีของฮวน หาค่าความเที่ยงโดยการตรวจสอบจากจุดประสงค์แต่ละข้อกับเนื้อหา หาค่าความเที่ยงโดยวิธีของคาร์เวอร์

ในปีเดียวกัน พิศิษฐ์ ตันทวนิช ได้ทำวิจัยเรื่อง การประยุกต์เทคนิคการวิเคราะห์ตามลำดับชั้นเพื่อกำหนดความยาวของแบบสอบอิงเกณฑ์ โดยมุ่งตรวจสอบหาช่วงความยาวของแบบสอบที่เกิดความคลาดเคลื่อนในการจัดประเภทของบุคคลตามระดับการรับรู้ ไม่แตกต่างกันไปจากอัตราความคลาดเคลื่อนที่กำหนดร้อยละ 5 ขอบเขตของเนื้อหาที่สร้างในหัวข้อ "การวิเคราะห์ข้อสอบ" ใช้วิธีการสร้างข้อสอบโดยการกำหนดภาระงานสอบเป็น "รูปแบบเฉพาะประจำข้อสอบ" จำนวน 14 รูปแบบ ให้ผู้มีประสบการณ์ในการสอน 18 คน ตรวจสอบความสอดคล้องระหว่างวัตถุประสงค์ ขอบเขตของเนื้อหาและรูปแบบเฉพาะประจำข้อสอบ จากนั้น ได้สร้างข้อสอบจากรูปแบบเฉพาะประจำข้อสอบ รูปแบบละ 5 ข้อ และสร้างข้อสอบจากรูปแบบเฉพาะประจำข้อสอบที่ 6 และ 7 สุ่มแทรกเพิ่มเข้าไปอีกรูปแบบละ 1 ข้อ รวมเป็นแบบสอบเต็มฉบับ 72 ข้อ กลุ่มตัวอย่างเป็นนักศึกษาระดับประกาศนียบัตรวิชาชีพการศึกษาระดับสูง จำนวน 282 คน แบบสอบที่พัฒนาขึ้นมีค่าดัชนีความไวของข้อสอบไม่ต่ำกว่า .20 เป็นแบบสอบที่มีความตรงเชิงจำแนก หาค่าสัมประสิทธิ์ความเที่ยงด้วยวิธีการของสับโคเวียคมีค่าระหว่าง 0.90-0.97 การวิเคราะห์ข้อมูลมุ่งตรวจสอบอัตราความคลาดเคลื่อนที่ปรากฏกับอัตราความคลาดเคลื่อนที่กำหนด โดยการทดสอบทวินาม เพื่อใช้คะแนนจากแบบสอบเต็ม

ฉบับเป็นตัวเกณฑ์กำหนดความถูกต้องในการจัดประเภทบุคคลตามระดับการรอบรู้ และคะแนน จากแบบสอบที่มีช่วงความยาวเป็น 10 ข้อ 14 ข้อ 20 ข้อ 25 ข้อ 28 ข้อ และ 30 ข้อ เป็นตัวอิงเกณฑ์ และหาค่าสหสัมพันธ์เคนดอลล์ ทอ ระหว่างผลการจัดประเภทบุคคลตามระดับ การรอบรู้ จากแบบสอบเต็มฉบับ และแบบสอบในช่วงความยาวต่างๆกัน พบว่า จำนวนข้อสอบ ที่มีความเหมาะสมในการจัดสอบ เพื่อจำแนกบุคคลตามระดับการรอบรู้ คือ จำนวน 20 ข้อ และมีความสัมพันธ์กับแบบสอบเต็มฉบับอยู่ในช่วง 0.53-0.89

สุนทรี ไกรกานแก้ว(2532) ได้ศึกษาคุณภาพของแบบสอบอิง เกณฑ์วิชาคณิตศาสตร์ ชนิด โจทย์ปัญหา และ โจทย์ทักษะการคิดคำนวณของครูผู้เขียนข้อสอบที่มีลักษณะต่างกันในด้าน เพศ วุฒิ และประสบการณ์ ที่เขียนข้อสอบจากลักษณะเฉพาะของแบบทดสอบที่ใช้รูปแบบ ฟาเซทในการเขียนข้อคำถามและตัวเลือก กลุ่มตัวอย่างเป็นครู 16 คน และนักเรียนชั้น ประถมศึกษาปีที่ 6 จำนวน 312 คน ผลการวิจัยพบว่า ค่าความยากก่อนสอนและหลังสอน ค่าอำนาจจำแนก และค่าความเชื่อมั่นของแบบทดสอบวิชาคณิตศาสตร์ชนิด โจทย์ปัญหาและ โจทย์ทักษะคำนวณ ที่เขียนข้อสอบโดยครูที่มีลักษณะด้านเพศ วุฒิ และประสบการณ์แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติที่ระดับ .05

รอยด์และฮาလာไดนา (Roid and Haladyna, 1982) ได้ศึกษาโดยกำหนดให้ ผู้เขียนข้อสอบ 3 ทีม ที่ได้รับความรู้เกี่ยวกับการเขียนข้อสอบ โดยใช้รูปแบบฟาเซทแตกต่างกัน เป็นผู้สร้างแบบทดสอบโดยใช้รูปแบบฟาเซทจากจุดประสงค์การสอนที่กำหนดให้สอง จุดประสงค์การสอน โดยตั้งสมมติฐานในการศึกษาครั้งนี้ไว้ 2 ข้อ คือ 1) ข้อสอบที่มีค่า ความยากเทียบเท่ากัน น่าจะถูกเขียนจากผู้เขียนข้อสอบในทีมที่ได้รับความรู้เต็มที่เกี่ยวกับ รูปแบบฟาเซท 2) ตัวลวงของข้อสอบเลือกตอบที่ผู้เขียนข้อสอบได้รับความรู้เต็มที่เกี่ยวกับรูปแบบฟาเซทจะสามารถวินิจฉัยความบกพร่องของผู้เรียน ได้ดีกว่าข้อสอบที่ผู้เขียนข้อสอบ ได้รับความรู้ในการเขียนข้อสอบโดยใช้รูปแบบฟาเซทเพียงบางส่วน ผลของการศึกษาพบว่า เมื่อพิจารณาข้อสอบคู่ขนานซึ่งจับคู่กัน โดยอาศัยหลักตรรกศาสตร์ของผู้เขียนข้อสอบในทีมที่ ได้รับความรู้เกี่ยวกับรูปแบบฟาเซทเต็มที่เหมือนกันแล้ว พบว่าค่าความยากของข้อสอบจาก ผู้เขียนข้อสอบในทีมนี้มีค่าไม่แตกต่างกัน ซึ่งสอดคล้องกับสมมติฐานที่ตั้งไว้ และเมื่อนำข้อ สอบแต่ละคู่ที่สร้างจากผู้เขียนข้อสอบแต่ละทีมเปรียบเทียบกัน พบว่ามีค่าความยากต่างกัน

นอกจากนั้นผลการศึกษายังชี้ให้เห็นว่าเมื่อผู้เขียนข้อสอบได้รับความรู้ที่แตกต่างกัน หรือผู้เขียนข้อสอบที่ได้รับอิสระในการเลือกใช้คำเขียนข้อสอบแล้วจะได้ค่าความยากของข้อสอบที่แตกต่างกัน และให้ข้อเสนอแนะว่าการสร้างตัวลอง โดยใช้รูปแบบฟาเซท จะใช้วินิจฉัยข้อบกพร่องของผู้เรียนเพื่อประโยชน์ในการสอนซ่อมเสริม

จากงานวิจัยดังกล่าว จะเห็นได้ว่างานวิจัยของกาญจนา วัฒนสุนทร ได้นำวิธีการกำหนดรูปแบบเฉพาะประจำข้อสอบ (Item Form) มาใช้ในการสร้างแบบสอบอิงเกณฑ์ นอกจากนั้นการสร้างแบบสอบอิงเกณฑ์ ยังคงยึดเอาวัตถุประสงค์เชิงพฤติกรรมเป็นหลักในการสร้าง ซึ่งมีได้นำมากล่าวในที่นี้ ต่อจากนั้นก็ยังมีบุคคลอื่นอีกหลายคนสร้างข้อสอบโดยการกำหนดรูปแบบเฉพาะประจำข้อสอบ และในปี 2532 สุนทรี ไกรกานแก้ว ได้สร้างข้อสอบโดยใช้รูปแบบฟาเซทในการใช้เป็นเครื่องมือในการวิจัย เพื่อศึกษาเปรียบเทียบลักษณะของครูในการเขียนข้อสอบแบบที่มีผลคุณภาพของเครื่องมือ มิได้เน้นในประเด็นขั้นตอนการสร้างและพัฒนาแบบสอบ ตลอดจนความเหมาะสมในการนำไปใช้ในสถานการณ์ต่างๆ เพื่อศึกษาปัญหา ประโยชน์และข้อจำกัดที่ค้นพบจากการสร้างข้อสอบโดยวิธีนี้ จึงเป็นประเด็นที่ผู้วิจัยสนใจ ได้นำมาศึกษาต่อเพื่อประโยชน์ทางการวัดผลสืบต่อไป

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย