

บทที่ 2

วรรณคดีที่เกี่ยวข้อง

เนื่องจากการศึกษาในครั้งนี้ เป็นการศึกษาการตรวจสอบความตรงเชิงเนื้อหา ของแบบสอบในบริบทของการสร้างแบบสอบเพื่อใช้คัดเลือกบุคคลเข้าทำงาน การศึกษาคำรา เอกสาร และงานวิจัย ที่นำเสนอในบทนี้จึงมุ่งเน้นเรื่องที่เกี่ยวข้องกับบริบทดังกล่าว โดยแบ่งการนำเสนอเป็น 3 ตอน คือ

ตอนที่ 1 ความตรงของแบบสอบ

ตอนที่ 2 การตรวจสอบความตรงเชิงเนื้อหาของแบบสอบ

2.1 ตามกรอบแนวคิดของ Fitzpatric (1983)

2.2 ตามแนวคิดที่เสนอโดย Lawshe (1975)

ตอนที่ 3 ปัจจัยที่มีผลต่อการตรวจสอบค่าความตรงเชิงเนื้อหาในเชิงปริมาณ มีรายละเอียด ดังนี้

ตอนที่ 1 ความตรงของแบบสอบ

การตรวจสอบคุณภาพแบบสอบก่อนการนำแบบสอบไปใช้ เป็นสิ่งที่มีความสำคัญมากที่สุดที่ผู้สร้างแบบสอบต้องคำนึงถึง และให้ความเป็นธรรมต่อผู้สอบ วิธีการตรวจสอบแบบสอบว่าสามารถวัดในสิ่งที่ต้องการวัดได้หรือหรือไม่ ในศาสตร์ทางการวัดผล เรียกว่า การตรวจสอบความตรง(validation) ของแบบสอบนั่นเอง (Brown, 1983)

ความตรง (validity) เป็นคุณสมบัติที่สำคัญที่สุดของเครื่องมือในการวัดทุกชนิด เพราะเป็นสิ่งที่บ่งบอกว่าเครื่องมือนั้นสามารถวัดได้ถูกต้องและครบถ้วนตามความมุ่งหมายที่ต้องการวัด แบบสอบนับว่าเป็นเครื่องมือในการวัดชนิดหนึ่งที่ช่วยให้ทราบตัวอย่างพฤติกรรมของแต่ละบุคคล (Allen and Yen, 1979) ซึ่งต้องมีการออกแบบและสร้างอย่างเป็นระบบ รวมทั้ง ตรวจสอบจนแน่ใจว่ามีความตรงจึงจะนำไปใช้ (Brown, 1983)

Messick (1989 cited in Linn, 1989) ได้นิยามความตรงว่า เป็นการรวบรวมหลักฐานของการตัดสินใจและประเมินแบบสอบ ซึ่งอาจเป็นหลักฐานในเชิงประจักษ์หรือความเป็นเหตุเป็นผลเชิงทฤษฎีที่จะสนับสนุนผลที่ได้จากการสอบหรือวิธีการอื่น ๆ ในการประเมินอย่างเพียงพอและเหมาะสม การแบ่งประเภทของความตรง ตามหลักฐานที่สนับสนุนว่าแบบสอบมีความตรงเป็น 3 กลุ่ม (American Educational Research Association [AERA] American Psychological Association [APA] and National Council on Measurement in Education [NCME], 1985) คือ

1. หลักฐานที่แสดงความสัมพันธ์กับโครงสร้างเชิงทฤษฎี (construct related evidence) ซึ่งเน้นที่การพิจารณาว่าคุณลักษณะทางทฤษฎีอย่างไรที่แบบสอบฉบับนั้นมุ่งวัด โดยกำหนดระดับของแนวคิดหรือทฤษฎีที่ใช้ในการอธิบายเพื่อการพัฒนาแบบสอบ จนกระทั่งการแสดงให้เห็นถึงความสัมพันธ์เชิงประจักษ์ของคะแนนที่ได้จากแบบสอบกับตัวแปรตัวอื่นๆ ที่มีความหมายของคะแนน ที่ได้จากแบบสอบอย่างชัดเจน ซึ่งต้องมีการศึกษาอย่างต่อเนื่อง มีข้อมูลยืนยันจากแหล่งข้อมูลต่าง ๆ มีการตรวจสอบจนแน่ใจก่อนที่จะลงข้อสรุปถึงความหมายและสาระที่ได้จากแบบสอบ

2. หลักฐานที่แสดงความสัมพันธ์กับเนื้อหา (content related evidence) เป็นหลักฐานที่แสดงว่าเนื้อหาที่เป็นตัวอย่างในข้อสอบเป็นตัวแทนของเอกภพ หรือมวลเนื้อหาในเรื่องนั้น ได้ดีเพียงไร เป็นหลักฐานที่สำคัญที่ชี้ให้เห็นถึงกระบวนการในการพัฒนาแบบสอบมีเนื้อหาสอดคล้องตามสาระของเนื้อหาที่ต้องการวัด

3. หลักฐานที่แสดงความสัมพันธ์กับเกณฑ์ (criterion related evidence) เป็นหลักฐานที่เน้นการเปรียบเทียบคะแนนจากแบบสอบกับเกณฑ์ในการวัด การเลือกเกณฑ์ให้ถูกต้องจึงเป็นสิ่งสำคัญในการยืนยันความหมายของผลของคะแนนที่ได้จากแบบสอบ ส่วนใหญ่แล้วเกณฑ์ที่ใช้มี 2 ลักษณะ คือ เกณฑ์เพื่อการทำนาย และเกณฑ์เพื่อศึกษาสภาพปัจจุบัน การจะเลือกใช้เกณฑ์ประเภทใดจึงขึ้นกับความต้องการในการแปลความหมายของคะแนนที่นำไปใช้ และต้องพิจารณาถึงประโยชน์ในบริบทที่ต้องการด้วย

จากความหมายและประเภทของความตรงที่กล่าวมาข้างต้น เป็นผลให้วิธีการตรวจสอบความตรงของแบบสอบมีความหลากหลาย เพื่อให้เป็นหลักฐานในการยืนยันว่าแบบสอบมีความตรง และถ้ายังมีหลักฐานหลายประเภทก็ยังคงเป็นการรับรองถึง

คุณภาพของแบบสอบว่าสามารถนำไปใช้ได้ตามความตั้งใจ หรือไม่ (AERA APA and NCME, 1985) อย่างไรก็ตาม ในกระบวนการสร้างแบบสอบหลักฐานที่นับว่ามีความสำคัญในการยืนยันว่า แบบสอบที่สร้างขึ้นเป็นตัวแทนสิ่งที่ต้องการวัดหรือไม่ นั่นคือหลักฐานที่แสดงถึงความสัมพันธ์ของเนื้อหาในแบบสอบและเนื้อหาที่ต้องการวัด ซึ่งอาจเรียกประเภทของหลักฐานชนิดนี้ว่า ความตรงเชิงเนื้อหา (Ebel, 1986) เพราะขั้นตอนเริ่มแรกของการพัฒนาแบบสอบ คือ การระบุเอกภพของเนื้อหา (universe of content) อย่างเพียงพอ เพื่อกำหนดให้เนื้อหาที่เป็นตัวแทนในแบบสอบตามจุดมุ่งหมายของการใช้แบบสอบหรือเมื่อผู้ใช้แบบสอบตัดสินใจใช้แบบสอบที่หาได้ ผู้ใช้ต้องพิจารณาวัตถุประสงค์ดั้งเดิมของผู้สร้างเพื่อตัดสินใจ เกี่ยวกับความเหมาะสมของนิยามโดเมนที่กำหนดไว้ว่า สอดคล้องกับสถานการณ์ที่ตั้งใจจะนำแบบสอบไปใช้หรือไม่ (AERA APA and NCME, 1985)

การตรวจสอบความตรงเชิงเนื้อหาของแบบสอบ มีความสัมพันธ์กับเอกภพหรือโดเมนของเนื้อหาที่กำหนดขึ้น ถ้าสามารถกำหนดโดเมนของเนื้อหาที่ต้องการวัดได้ชัดเจนเท่าไร การตรวจสอบความตรงเชิงเนื้อหา ก็มีความถูกต้องมากขึ้นเท่านั้น เมื่อพิจารณาที่มาของเนื้อหาในแบบสอบต่าง ๆ แล้ว แบบสอบคัดเลือกคนเข้าทำงาน นับว่าเป็นแบบสอบที่มีที่มาของเนื้อหาที่หลากหลาย เนื่องจากมีจุดมุ่งหมายในการสร้างขึ้นเพื่อเป็นเครื่องมือในการเสาะหาบุคคลที่มีความสามารถเหมาะสมกับตำแหน่งหน้าที่กิจกรรมในงานที่ต้องการ และนโยบายขององค์กร (Ebel, 1977) ซึ่งต้องใช้หลักฐานเชิงเกณฑ์สัมพันธ์ในการทำนายค่าของคะแนนของผู้สมัครว่า มีความสัมพันธ์กับเกณฑ์ที่บ่งบอกว่าผู้นั้นจะประสบความสำเร็จในการทำงานหรือไม่ อย่างไรก็ตาม ถ้าคะแนนที่ได้จากแบบสอบมีความสัมพันธ์กับเกณฑ์สูงแต่ไม่สามารถยืนยันได้ว่าแบบสอบสร้างตามโดเมนของเนื้อหาตามนิยามของสิ่งที่มุ่งวัด ค่าคะแนนที่ได้จากแบบสอบนั้นก็ไม่มี ความหมายในเชิงทำนายความสำเร็จของผู้ที่ได้รับการคัดเลือก

ตอนที่ 2 ความตรงเชิงเนื้อหาของแบบสอบ

ความตรงเชิงเนื้อหาเป็นหลักฐานที่สำคัญในการยืนยันว่า แบบสอบที่สร้างขึ้นนั้นมีเนื้อหาที่เป็นตัวแทนของโดเมนในเอกภพที่ต้องการวัด หรือเป็นตัวแทนของพฤติกรรมที่ต้องการวัด และส่งผลให้มั่นใจได้ว่าเนื้อหาในแบบสอบนั้นสอดคล้องกับสิ่งที่ต้องการวัดจริง ๆ (Kintop, 1977) มีการแบ่งความตรงเชิงเนื้อหาตามแนวคิดที่มีผู้นำเสนอไว้ 2 แนวคิด คือ

2.1 ตามแนวคิดของ Fitzpatric (1983)

Fitzpatric (1983) ได้รวบรวมแนวคิดเกี่ยวกับความหมายของความตรงเชิงเนื้อหาของแบบสอบจากทัศนะผู้เชี่ยวชาญด้านการวัดผล เช่น Ebel (1956), Cronbach (1971), Linn (1974), Anatasi (1976) และ Brown (1976) เป็นต้น โดยสรุปความหมายของความตรงเชิงเนื้อหาไว้ 4 กรอบแนวคิด ดังนี้

1. กรอบแนวคิดที่ได้จากการสุ่มโดเมนเนื้อหาที่ต้องการวัด (the concept of domain sampling) ซึ่งอธิบายความหมายของความตรงเชิงเนื้อหาของแบบสอบว่าเป็นแบบสอบที่เป็นตัวแทนของโดเมนเนื้อหาที่ต้องการวัดได้ แนวคิดเกี่ยวกับการสุ่มโดเมนเนื้อหาที่ต้องการวัดมี 2 ลักษณะ คือ

1.1 แบบสอบเป็นตัวอย่างของเนื้อหา (the test as a content sample)
โดยจะพิจารณาจาก

1) ข้อสอบแต่ละข้อในแบบสอบ ต้องมีความสอดคล้องกับเนื้อหาสาระที่มีการอธิบายไว้ในโดเมนเนื้อหาได้ดีเพียงใด

2) แบบสอบมีความเป็นตัวแทนของโดเมนเนื้อหาที่ต้องการวัดหรือไม่ โดยมีผู้เชี่ยวชาญเป็นผู้พิจารณาตัดสินว่ามีความเป็นตัวแทนที่ดีเพียงใด

1.2 แบบสอบเป็นตัวอย่างของพฤติกรรม (the test as a behavioral sample) โดยพิจารณาความตรงเชิงเนื้อหาของแบบสอบว่า คำตอบในแบบสอบเป็นตัวอย่างของพฤติกรรมที่ต้องการวัดซึ่งจะให้แนวคิดเกี่ยวกับการแปลความหมายของคำตอบในแบบสอบ และบอกให้ทราบว่าผู้สอบมีพฤติกรรมอะไรบ้าง การตัดสินแบบสอบว่าเป็นตัวอย่างของพฤติกรรมนั้น พิจารณาโดยผู้เชี่ยวชาญ

2. กรอบแนวคิดที่สัมพันธ์กับโดเมนในเนื้อหาที่ต้องการวัด (the concept of domain relevance) ซึ่งอธิบายความหมายของความตรงเชิงเนื้อหาของแบบสอบว่าแบบสอบมีความสัมพันธ์กับประเด็นของโดเมนในเอกภพของเนื้อหาที่ต้องการวัด จำแนกเป็น 2 ลักษณะ ดังนี้

2.1 เป็นประเด็นที่สัมพันธ์กับเนื้อหาของแบบสอบ(the relevance of test content) กล่าวคือ เนื้อหาในแบบสอบจะต้องครอบคลุมสาระสำคัญของเอกภพ (universe) ของเนื้อหา ตามเป้าหมายที่ผู้สร้างแบบสอบกำหนดขึ้น การพิจารณาตัดสินความสอดคล้องของเนื้อหา มักจะกระทำโดยผู้เชี่ยวชาญ ผู้เชี่ยวชาญนี้อาจเป็นอาจารย์ ผู้สอนที่มีประสบการณ์ อาจารย์ ผู้จัดหลักสูตร หรือผู้เขียนตำรา เป็นต้น

2.2 เป็นประเด็นที่สัมพันธ์กับคำตอบของแบบสอบ(the relevance of test responses) โดยพิจารณาจากคำตอบที่ได้จากแบบสอบมีความสัมพันธ์กับพฤติกรรมที่สอดคล้องกับเนื้อหาที่เป็นประเด็นในสิ่งที่ต้องการวัด

3. กรอบแนวคิดที่มีความชัดเจนของโดเมนเนื้อหาที่ต้องการวัด (the concept of domain clarity) ตามแนวคิดนี้ พิจารณาที่โดเมนของเนื้อหาที่ต้องการวัด สามารถให้คำนิยามในสิ่งที่ต้องการวัดได้อย่างชัดเจน โดยพิจารณาที่

3.1 ความชัดเจนของตัวแปรในเนื้อหาที่ต้องการวัด สามารถกำหนดวิธีการที่เป็นระบบในการเก็บรวบรวมข้อมูลเพื่อการตัดสินใจ

3.2 การอธิบายสาระสำคัญของเนื้อหาที่ต้องการวัด และนำไปสู่การปฏิบัติได้

3.3 การให้คำจำกัดความที่มีการอธิบายถึงรายละเอียดของเนื้อหาและแนวทางในการตรวจให้คะแนนข้อสอบที่สร้างขึ้น รวมทั้ง การแปลความคะแนนจากการสอบว่าใช้วัดพฤติกรรมใดที่แบบสอบตั้งใจจะวัดด้วย

4. กรอบแนวคิดของเทคนิคการสร้างข้อสอบที่มีคุณภาพ (the concept of technical quality in test items) ผู้สร้างแบบสอบต้องศึกษาเทคนิคในการสร้างข้อสอบที่มีคุณภาพ โดยเน้นการใช้กระบวนการเชิงประจักษ์ และความสมเหตุสมผลในการสร้าง ระหว่างดำเนินการควรมีผู้เชี่ยวชาญให้คำชี้แนะ และต้องคำนึงถึงเสมอว่าความตรงเชิงเนื้อหาของแบบสอบเป็นสิ่งที่จำเป็น เทคนิคการสร้างข้อสอบที่ผู้เขียนได้

เสนอแนะไว้ คือ จะต้องมีการวิเคราะห์ทั้งเนื้อหาของข้อสอบในแบบสอบ และคะแนนของคำตอบจากแบบสอบด้วย

จากความหมายของความตรงเชิงเนื้อหาทั้ง 4 กรอบแนวคิด ที่นำเสนอมานั้น Fitzpatric ยังได้วิเคราะห์ถึง ความหมายของความตรงเชิงเนื้อหาของแบบสอบคัดเลือกบุคคลเข้าทำงาน ว่าเป็นไปตามกรอบแนวคิดในประเด็นที่ 2 คือ กรอบแนวคิดที่สัมพันธ์กับโดเมนของเนื้อหาที่ต้องการวัด (the concept of domain relevance) โดยให้เหตุผลว่า เนื้อหาในแบบสอบคัดเลือกบุคคลเข้าทำงานกำหนดได้จากกระบวนการวิเคราะห์งาน (job analysis) ซึ่งจะได้เพียงเอกภพของประเด็นในเนื้อหาที่เกี่ยวข้องกับกิจกรรมในการปฏิบัติงานทั้งหมดหรือเป็นเนื้อหาที่ได้จากการบ่งชี้ของผู้มีความรู้ความชำนาญในลักษณะงานนั้น ๆ ว่า เนื้อหาส่วนใดที่มีความสำคัญต่อการทำงานให้ประสบความสำเร็จเพื่อกำหนดเป็นพฤติกรรมที่ต้องการวัดต่อไป นอกจากนี้ Fitzpatric ยังได้อ้างถึงผู้เชี่ยวชาญด้านการสอบ เช่น Brown (1976), Anatasi (1976) Thronidike & Hagen (1977) รวมทั้ง Lawshe (1975) และ Guion (1978) ว่าเป็นผู้ที่มีความคิดเห็นที่สอดคล้องกับแนวคิดดังกล่าวด้วย

2.2 แนวคิดที่เสนอโดย Lawshe (1975)

Lawshe (1975 cited in Pitiyanuwat, 1981) ได้เขียนบทความที่ชื่อว่า "A quantitative approach to content validity" เนื้อหาในบทความเกี่ยวข้องกับการตรวจสอบความตรงเชิงเนื้อหาในเชิงปริมาณ ซึ่งมีรายละเอียด ดังนี้

เนื่องจาก ความจำเป็นด้านของกฎหมายว่าด้วยสิทธิมนุษยชนของประเทศสหรัฐอเมริกาในการคัดเลือกคนเข้าทำงาน ส่งผลต่อ การประยุกต์การตรวจสอบความตรงเชิงเนื้อหาของแบบสอบวัดผลสัมฤทธิ์ทางวิชาการมาใช้ในแบบทดสอบคัดเลือกบุคคลเข้าทำงานในทางธุรกิจ และทางอุตสาหกรรม เพราะกฎหมายกำหนดให้แสดงค่าความตรงของแบบสอบเป็นหลักฐาน ประกอบกับข้อจำกัดในการตรวจสอบความตรงเชิงเกณฑ์สัมพันธ์ซึ่งต้องมีขั้นตอนและวิธีการที่ยุงยากว่า ผู้สร้างแบบสอบจึงใช้การตรวจสอบความตรงเชิงเนื้อหาในการตรวจสอบแบบสอบของคน และค่อย ๆ เปลี่ยนแปลงมาเป็นรูปแบบของ "การยอมรับของผู้เชี่ยวชาญ" ในทางปฏิบัติ ซึ่งอาจไม่ใช่ความตรง

เชิงเนื้อหาตามความหมายที่ควรจะเป็น การหาความสัมพันธ์แบบใหม่เป็นการกระทำ โดยการใช้ผู้เชี่ยวชาญในทางอุตสาหกรรม ซึ่งไม่มีการสอดแทรกเรื่องของการทำนาย หรือการประจุมได้เถียงกันของผู้เชี่ยวชาญ เหตุผลที่สำคัญ คือ ยังไม่มีการศึกษาเกี่ยวกับการตรวจสอบความตรงเชิงเนื้อหาของแบบสอบคัดเลือกคนเข้าทำงาน โดยเฉพาะสำหรับหน่วยงานของรัฐหรือการคัดเลือกคนเข้ารับราชการ แนวคิดของการ "สอบได้ หรือสอบตก" จึงเป็นกรณีเฉพาะที่เป็นข้อจำกัดในการนำไปประยุกต์ใช้การคัดเลือกบุคคลเข้าทำงานด้วย และอีกประการหนึ่ง คือ ยังไม่มีการศึกษาความสอดคล้องในเรื่องของความคิดเห็นของผู้เชี่ยวชาญและผู้ปฏิบัติงาน จึงกลายเป็นปัญหาเฉพาะหน้าที่ต้องพยายามแก้ไข ว่าผลการพิจารณาและการยอมรับการตัดสินใจของผู้เชี่ยวชาญนั้น เป็นองค์ประกอบของการตรวจสอบความตรงเชิงเนื้อหาได้อย่างไร มีความเกี่ยวข้องมากต่อการตัดสินใจของศาลในเรื่องของการรับคนเข้าทำงานได้ตามบทบาทที่กำหนด จากเหตุผลที่กล่าวมานี้ Lawshe จึงนำเสนอบทความนี้เพื่อให้แนวคิด ดังต่อไปนี้

1. ก่อให้เกิดความชัดเจนของกรอบแนวคิดเกี่ยวกับการตรวจสอบความตรงเชิงเนื้อหา
2. เสนอวิธีการตรวจสอบความตรงเชิงเนื้อหาของแบบสอบด้วยวิธีการเชิงปริมาณ

การกำหนดกรอบแนวคิดความคิดเกี่ยวกับการตรวจสอบความตรงเชิงเนื้อหา มีประเด็นสำคัญ 4 ประเด็น ดังนี้

ประเด็นที่ 1. การจำแนกความแตกต่างระหว่างการตรวจสอบความตรงเชิงเนื้อหาของหลักสูตร และการปฏิบัติงาน ในเรื่องของหลักสูตรเพื่อสร้างแบบสอบวัดผลสัมฤทธิ์ทางวิชาการ การตรวจสอบความตรงเชิงเนื้อหาของแบบสอบนี้ จึงตรวจสอบในเรื่องของ

- 1) การนิยามเอกภพของหลักสูตร (เรียกว่าโดเมนของเนื้อหา) หรือ
- 2) ความเพียงพอของตัวอย่างในแบบสอบที่เป็นตัวแทนของเอกภพของหลักสูตร

แต่ในการตรวจสอบความตรงเชิงเนื้อหาของการปฏิบัติงาน ซึ่งไม่มีการนิยามเอกภพของการปฏิบัติงานไว้เลย ผู้สร้างแบบสอบต้องศึกษาผลจากการวิเคราะห์งานอย่างระมัดระวัง สิ่งที่สามารถทำได้ คือ การแบ่งงานออกเป็น ส่วน ๆ ในเอกภพของการปฏิบัติงานทั้งหมด Lawshe ได้เสนอให้นิยามโดเมนของการปฏิบัติงาน โดยแบ่งงาน

เป็นส่วน ๆ เพื่อให้สามารถกำหนดกรอบของตรวจสอบความตรงเชิงเนื้อหาของการปฏิบัติงานในเรื่องของ

- 1) การนิยามการปฏิบัติงานที่กำหนดขึ้นจากผลการวิเคราะห์งาน และ
- 2) การอ้างถึงสิ่งที่ต้องปฏิบัติในงาน

นอกจากนี้ ความยากของงาน คือการกำหนดลักษณะงาน ที่จำแนกเป็นงานเดี่ยว เช่น งานพิมพ์คิดเป็นการปฏิบัติงานที่ใช้ความรู้ความสามารถเดี่ยวคือการพิมพ์งานจากข้อความที่ร่างไว้ หรือ งานพหุ ที่มีหลายงานประกอบกัน เช่น งานเลขานุการเป็นงานที่ต้องใช้การปฏิบัติงานหลายด้าน หรือหลายมวลเนื้อหา คือ นอกจากความรู้ในการพิมพ์งานจากกระดาษที่ร่างไว้ด้วย ยังต้องมีความรู้ในเรื่องของมนุษย์สัมพันธ์ การต้อนรับ และความรู้ด้านภาษาที่ใช้ อีกด้วย ความแตกต่างนี้ทำให้การกำหนดนิยามและตัวอย่างที่ครบถ้วนในเรื่องของงานเหมือนการวัดผลสัมฤทธิ์ทางวิชาการทำได้ยาก ฉะนั้น ในเรื่องของการปฏิบัติงานเรากำหนดมวลเนื้อหาของตัวอย่างการปฏิบัติงานโดยประมาณจากเอกภพของการปฏิบัติงานทั้งหมด ซึ่งยังไม่ใช้ทั้งหมดทุกส่วน จึงยังต้องแบ่งเป็นส่วนต่าง ๆ ของการปฏิบัติงานที่มีการระบุถึงความเฉพาะของตำแหน่งงานแต่ละตำแหน่งด้วย แต่บทความนี้ไม่ได้กล่าวถึง การกำหนดมาตรฐานในการปฏิบัติงาน

ประเด็นที่ 2. การตัดสินของผู้เชี่ยวชาญ Adkins (1973, cited in Lawshe, 1975) เสนอแนวคิดเกี่ยวกับการตัดสินความตรงเชิงเนื้อหาของแบบสอบ ว่า

"ในแบบสอบวัดผลสัมฤทธิ์ทางวิชาการ การตัดสินว่าแบบสอบมีเนื้อหาอย่างไร และมีกระบวนการคิดทางสมองอย่างไร มีความสัมพันธ์กับวัตถุประสงค์ในการสอน

แต่ในแบบสอบวัดการเข้าทำงาน การตรวจสอบความตรงเชิงเนื้อหาจะถูกตัดสินเกี่ยวกับความสามารถที่วัดโดยแบบสอบว่าวัดความสามารถที่จำเป็นสำหรับการทำงานให้ประสบความสำเร็จได้หรือไม่ "

คำถามที่สำคัญ ก็คือ "ใครเป็นผู้ตัดสิน" ในแบบสอบวัดผลสัมฤทธิ์ เราใช้การตัดสินของผู้เชี่ยวชาญตามนิยามเอกภพของหลักสูตรและนำมาร่างมวลเนื้อหา แล้วให้ผู้เชี่ยวชาญกำหนดน้ำหนักของแบบสอบในแต่ละส่วน ความตรงเชิงเนื้อหาจึงเป็น

หลักฐานที่แสดงให้เห็นว่า ข้อสอบในแบบสอบมีความเหมาะสมที่จะเป็นตัวแทนของ มวลเนื้อหาที่เรียน หรือไม่ และถ้าผู้เชี่ยวชาญมีความเชี่ยวชาญในเนื้อหาวิชานั้นจริง จะ ทำให้ได้การผลตรวจสอบความตรงเชิงเนื้อหาที่เหมาะสม

ในแบบสอบคัดเลือกบุคคลเข้าทำงาน ผู้เชี่ยวชาญน่าจะเป็นผู้ที่ เป็นหัวหน้างาน หรือผู้ควบคุมงานที่รู้เกี่ยวกับงานนั้น หรือนักจิตวิทยา หรือผู้เชี่ยวชาญอื่น ๆ ที่ได้รับการ คาดหวังว่ามีความเข้าใจในบุคลิกลักษณะของบุคคลในองค์การและเข้าใจอย่างลึกซึ้งว่า แบบสอบนั้นวัดอะไร จึงเป็นเหตุผลสำคัญของการกำหนดมวลเนื้อหาในการปฏิบัติงาน และคุณลักษณะของคนที่ต้องการให้ได้ เพื่อใช้เป็นแนวทางในการคัดเลือกผู้เชี่ยวชาญ

ประเด็นที่ 3. ขอบเขตของการศึกษางานที่ต้องปฏิบัติ องค์ประกอบด้าน เนื้อหาของงานที่ต้องปฏิบัติ ศึกษาได้จากทั้งพฤติกรรมการทำงานที่สามารถสังเกตได้ โดยตรง และพฤติกรรมที่เขียนเป็นรายงานได้ จนถึงพฤติกรรมที่มีความเป็นนามธรรม สูง หรืองานที่ต้องอาศัยความชำนาญ เช่น การคิดเลขหรือการพิมพ์คิดไปจนถึงการใช้ กระบวนการคิดในระดับสูง เช่น การอุปนัยและนิรนัยอย่างสมเหตุสมผล การ เปรียบเทียบพฤติกรรมของผู้สมัครโดยใช้แบบสอบค้นหาความจริงในการวัดพฤติกรรมที่ ต้องการในการปฏิบัติงานนั้น ยากที่จะอ้างอิงถึงกันได้ ซึ่งอาจต้องใช้วิธีการสังเกตอย่าง ต่อเนื่องโดยตรง อย่างไรก็ตาม พฤติกรรมที่มีระดับความเป็นนามธรรมสูง จะมีความ เสี่ยงในการอ้างอิงถึงการตรวจสอบความตรงเชิงเนื้อหาเพียงวิธีเดียวแต่ต้องการวิธี การอื่น ๆ นอกจากความสัมพันธ์กับเกณฑ์ด้วย ตัวอย่างเช่น การกล่าวว่า "มวลเนื้อหา ในการปฏิบัติงานนี้เป็นการบวกเลขทั้งหมด แบบสอบ A วัดความสามารถในการบวก เลข ฉะนั้น แบบสอบ A มีความตรงเชิงเนื้อหาที่เป็นตัวแทนคนที่มีศักยภาพด้านนี้" หรือในแบบสอบอื่น ๆ เช่น "มวลเนื้อหาของการทำงานนี้ ใช้การนิรนัยที่สมเหตุสมผล แบบสอบ B มีวัตถุประสงค์ในการวัดการนิรนัยที่สมเหตุสมผล ฉะนั้นแบบสอบ B มีความตรงที่จะวัดความสามารถของการทำงานตามที่กำหนด" การใช้วิธีการ สังเกตและสรุป อาจไม่ได้ความจริงของพฤติกรรมที่ต้องการก็ได้ ผู้ตัดสินที่กำลังปฏิบัติ งานนั้นอยู่ ผู้ควบคุมงาน หรือคนอื่น ๆ ที่มีความรู้เกี่ยวกับงาน น่าจะเป็นผู้ค้นหา พฤติกรรมที่ประมาณโดยแบบสอบจากตัวอย่างของงานในขอบเขตของงานที่ต้องปฏิบัติ ได้ดีกว่า การทดสอบความรู้ความสามารถเกี่ยวกับงานโดยแบบสอบที่ได้รับการประเมิน

ความตรงเชิงเนื้อหาที่พิจารณาจากบุคคลที่มีความรู้เกี่ยวกับงานหรือในประเด็นที่อยู่ในขอบเขตของงานที่ต้องปฏิบัติ น่าจะได้รับการยอมรับ

เมื่อพฤติกรรมที่วัดเป็นพฤติกรรมที่มีระดับความเป็นนามธรรมสูง การที่จะอ้างอิงถึง "การลงข้อสรุปที่เพิ่มขึ้น" ได้อย่างมีนัยสำคัญ ผู้ปฏิบัติงานหรือผู้ควบคุมงานที่ไม่เข้าใจงานอย่างลึกซึ้งจะไม่สามารถบอกถึงพฤติกรรมที่ต้องการได้ เมื่อเงื่อนไขถูกกำหนดขึ้น การเชื่อมโยงระหว่างความตรงเชิงเนื้อหากับความตรงเชิงโครงสร้างจึงเป็นสิ่งจำเป็น ตัวอย่างเช่น การทดสอบความสามารถในการนิรนัยอย่างสมเหตุสมผล ต้องอาศัยผู้เชี่ยวชาญที่เป็นนักจิตวิทยา จึงจะสามารถตัดสินได้ว่าแบบสอบนั้นวัดการนิรนัยอย่างสมเหตุสมผลหรือไม่ คือ การถูกวัดโดยแบบสอบ และการกำหนดประเด็นของมวลดเนื้อหาในงานที่ต้องปฏิบัติ ต้องเชื่อมั่นว่าได้รับการศึกษามาแล้วเป็นอย่างดี ซึ่งต้องอ้างอิงอย่างมีมาตรฐานด้วย หลักฐานที่เกี่ยวข้องกับความตรงเชิงโครงสร้าง (construct validity) ไม่ได้อยู่ที่การศึกษาเพียงครั้งเดียว การตัดสินความตรงเชิงโครงสร้างขึ้นกับผลการวิจัยที่ได้รับการตั้งสมมาในเรื่องนั้น

ประเด็นที่ 4. การนิยามเชิงปฏิบัติการความตรงเชิงเนื้อหา เป็นเรื่องของการตัดสินความคาบเกี่ยวกัน(overlap) ระหว่าง

- 1) การสร้างแบบสอบที่ได้รับการพิจารณาเป็นอย่างดีแล้ว
- 2) ความสามารถในภาระหน้าที่ที่กำหนดจากขอบเขตของงานที่ต้องปฏิบัติ

กรอบแนวคิดความคิดเกี่ยวกับการตรวจสอบความตรงเชิงเนื้อหาเป็นการวิเคราะห์เนื้อหาอย่างทันท่วงที และมีการตรวจสอบอย่างละเอียดในขอบเขตของงานที่ต้องปฏิบัติ โดยผู้ที่มีความรู้เกี่ยวกับงานสามารถตัดสินว่าเป็นความรู้ความสามารถหรือพฤติกรรมที่ต้องการ แต่ถ้าพฤติกรรมที่ต้องการนั้นเป็นพฤติกรรมเชิงนามธรรมต้องใช้ความตรงเชิงโครงสร้างเป็นตัววัดซึ่งผู้ปฏิบัติงานหรือผู้ควบคุมงานธรรมดาจะไม่สามารถตัดสินได้ นิยามเชิงปฏิบัติการของความตรงเชิงเนื้อหา คือ การที่จำนวนของสมาชิกในคณะผู้ประเมินรับรู้ถึงความคาบเกี่ยวกัน(overlap)ระหว่างแบบสอบกับขอบเขตของงานที่ต้องปฏิบัติและแบบสอบถูกจำกัดอยู่เฉพาะ 1)แบบสอบวัดความชำนาญหรือทักษะที่ง่าย ๆ 2) แบบสอบวัดความรู้เกี่ยวกับงาน 3) แบบสอบวัดตัวอย่างงาน

การเสนอวิธีการตรวจสอบความตรงเชิงเนื้อหาของแบบสอบด้วยวิธีการเชิงปริมาณ มีประเด็นสำคัญ 4 ประเด็น ดังนี้

ประเด็นที่ 1. ขนาดของส่วนที่คาบเกี่ยวกันระหว่างแบบสอบที่ได้รับการพิจารณาเป็นองค์แล้ว และ ความสามารถในการระบุหน้าที่ที่กำหนดจากขอบเขตของงานที่ต้องปฏิบัติ คณะกรรมการประเมินเนื้อหาจะสามารถกำหนดความคาบเกี่ยวกันระหว่างขอบเขตของงานที่ต้องปฏิบัติ กับแบบสอบที่สร้างขึ้นเฉพาะนั้น ได้หรือไม่? ในส่วนนี้ การใช้คณะกรรมการประเมินเนื้อหาที่ประกอบด้วยบุคคลที่รู้เกี่ยวกับงาน ผลที่ดีที่สุดที่ควรจะได้รับเมื่อคณะกรรมการประเมินเนื้อหาที่มีจำนวนพอ ๆ กัน ระหว่างผู้ปฏิบัติงานและผู้ควบคุมงาน การที่คณะกรรมการประเมินเนื้อหาได้รับข้อสอบที่ต้องการประเมิน และให้มีอิสระในการประเมินข้อสอบแต่ละข้อ ดังนี้

ทักษะหรือความรู้ที่วัดโดยข้อสอบข้อนี้ เป็นทักษะหรือความรู้ที่

- ก. จำเป็น
- ข. มีประโยชน์แต่ไม่จำเป็น
- ค. ไม่จำเป็น

ในการปฏิบัติงาน

ผลการประเมินของคณะกรรมการประเมินเนื้อหาทั้งหมด จะได้จำนวนของข้อสอบที่ดูซึ่งมีความจำเป็นในการวัดทักษะหรือความรู้ในการปฏิบัติงานระดับหนึ่ง

ประเด็นที่ 2. ความตรงของการตัดสินใจ เมื่อคณะกรรมการประเมินเนื้อหาได้ตัดสินใจประเมินข้อสอบแต่ละข้อแล้ว ถ้าคณะกรรมการประเมินเนื้อหา มีความเห็นเหมือนกัน สามารถสรุปได้ว่าความเห็นของคณะกรรมการประเมินเนื้อานั้น "ถูกต้องทั้งหมด" หรือ "ผิดทั้งหมด" เนื่องจากคณะกรรมการนั้นกำลังปฏิบัติงานอยู่ หรือการเป็นผู้ควบคุมงานโดยตรง จึงไม่สามารถโต้แย้งในความเห็นที่ลงรอยอย่างหนักแน่นเช่นนั้น จึงถือว่าคณะกรรมการประเมินเนื้อหา มีความตรงในการตัดสินใจ

ประเด็นที่ 3. ปริมาณของความสอดคล้องของความคิดเห็น ถ้าผลการตัดสินใจ เมื่อคณะกรรมการประเมินเนื้อหา มีความเห็นไม่ลงรอยกันว่าข้อสอบข้อนั้น จำเป็น สำหรับการวัดความรู้หรือทักษะในการปฏิบัติงานหรือไม่ เพราะ เมื่อคณะกรรมการได้ตัดสินใจ

ว่าข้อสอบวัดความรู้หรือทักษะที่ "จำเป็น" หรือ "ไม่จำเป็น" ทั้งหมด เราสามารถเชื่อมั่นได้ว่าความรู้และทักษะนั้นเป็นความจำเป็นที่แท้จริงหรือไม่ ซึ่งแสดงถึงความหนักแน่นของความคิดเห็นที่เป็นเชิงเดียวกัน แต่ถ้าข้อสอบนั้นได้รับการตัดสินจากคณะกรรมการประเมินเนื้อหาโดยมีคณะกรรมการฯ ครั้งหนึ่งเห็นว่า ข้อสอบนั้นมีความ "จำเป็น" และ อีกครั้งหนึ่งเห็นว่า "ไม่จำเป็น" จะตัดสินอย่างไร

Lawshe(1975) จึงตั้งสมมติฐาน 2 ข้อ เพื่อให้มีการตรวจสอบตามหลักการ ดังนี้

1) ข้อสอบทุก ๆ ข้อที่ได้รับการตัดสินว่า จำเป็น ในการวัดเนื้อหาเกี่ยวกับงานลักษณะนั้น ๆ ความเห็นของคณะกรรมการฯ มากกว่าครึ่งหนึ่งของจำนวนคณะกรรมการฯ ก็จะมีความตรงเชิงเนื้อหาในระดับหนึ่ง

2) ถ้าคณะกรรมการ ฯ มีความเห็นด้วย มากกว่าครึ่งหนึ่งของจำนวนคณะกรรมการฯ ทั้งหมด ยิ่งมากเท่าไร ยิ่งแสดงว่าข้อสอบนั้น จำเป็น ในการวัดเนื้อหาเกี่ยวกับงานลักษณะนั้น ๆ ก็จะมีระดับของความตรงเชิงเนื้อหาที่มากกว่า

จากสมมติฐานที่ตั้งขึ้นนั้น Lawshe จึงได้เสนอสูตรในการคำนวณสัดส่วนของความตรงเชิงเนื้อหา (CVR) ว่า

$$CVR = \frac{n_c - 1/2 N}{1/2 N}$$

เมื่อ n_c คือ จำนวนคณะกรรมการประเมินเนื้อหาที่ชี้ว่าข้อกระทงนั้นจำเป็นในการวัดเนื้อหาเกี่ยวกับงานลักษณะนั้น ๆ

N คือ จำนวนคณะกรรมการประเมินเนื้อหาทั้งหมด

โดย กำหนดให้ ค่า CVR ที่คำนวณได้เป็นโมเดลเชิงเส้น คำนวณจากความถี่ของผู้ที่ตอบว่าข้อสอบข้อนั้น "จำเป็น" ซึ่งมีลักษณะ ดังนี้

1. เมื่อค่ามากกว่าครึ่งหนึ่งของคณะกรรมการฯ เห็นว่าข้อสอบข้อนั้น "จำเป็น" ค่า CVR จะมีค่าเป็นลบ

2. เมื่อครึ่งหนึ่งของคณะกรรมการฯ เห็นว่าข้อสอบข้อนั้น "จำเป็น" และอีกครึ่งหนึ่งของคณะกรรมการฯ เห็นว่าข้อสอบข้อนั้น "ไม่จำเป็น" ค่า CVR จะมีค่าเป็น 0

3. ถ้าคณะกรรมการฯ ทั้งหมดเห็นว่า ข้อสอบข้อนั้น "จำเป็น" ค่า CVR จะมีค่าเป็น 1

4. ถ้าคณะกรรมการเห็นว่าข้อสอบข้อนั้น "จำเป็น" มากกว่าครึ่งหนึ่งของคณะกรรมการประเมินเนื้อหาทั้งหมด ค่าของ CVR จะอยู่ระหว่าง 0 ถึง .99

ประเด็นที่ 4. การคัดเลือกข้อสอบ ในการตรวจสอบความตรงเชิงเนื้อหาของแบบสอบเมื่อกำหนดค่า CVR ของข้อสอบแต่ละข้อแล้ว การคัดเลือกข้อสอบซึ่งคณะกรรมการประเมินเนื้อหาเห็นว่า "ไม่จำเป็น" ในการวัดเนื้อหาที่เกี่ยวข้องกับงานออกไปนั้น Schipper (1975 cited in Pitayanuwat, 1981) ได้เตรียมตารางค่าวิกฤตขั้นต่ำสุดของ CVR ที่ยอมรับ ได้ว่าข้อสอบข้อนั้นมีความตรงเชิงเนื้อหา หรือไม่

วิธีการคำนวณค่า CVR ที่เป็นค่าวิกฤต ได้คำนวณตามแนวคิดที่ Lawshe (1975) เสนอไว้ว่า จะยอมรับว่าข้อสอบมีความตรงเชิงเนื้อหา ก็ต่อเมื่อ จำนวนคณะกรรมการอย่างน้อยที่สุดครึ่งหนึ่งเห็นตรงกันว่า ข้อสอบนั้นมีความจำเป็นในการวัดความรู้ หรือ พฤติกรรมในการปฏิบัติงาน

Schipper (1975 cited in Pitayanuwat, 1981) เสนอสูตรการคำนวณค่าวิกฤตของ CVR โดยใช้การประมาณค่าการแจกแจงแบบทวินามด้วยการแจกแจงแบบปกติ เพราะการตัดสินใจแต่ละครั้งให้ผล 2 แบบ คือ การยอมรับหรือไม่ยอมรับว่าข้อสอบข้อนี้ใช้ความรู้ความสามารถเกี่ยวกับงานนี้ได้หรือไม่ โดยมีสูตรในการคำนวณ ดังนี้

เมื่อจำนวนผู้เชี่ยวชาญที่ใช้ในการตัดสินใจน้อยกว่า 10 คน จะใช้สูตรการประมาณค่าการแจกแจงความน่าจะเป็นชนิดต่อเนื่องแบบทวินามโดยตรง

เมื่อจำนวนผู้เชี่ยวชาญที่ใช้ในการตัดสินใจมากกว่า 10 คน จะใช้สูตรการประมาณค่าโดยคำนวณจากสูตรของลอว์ซี เป็นขั้นตอน ดังนี้

$$\text{จากสูตร} \quad CVR = \frac{n_c - 1/2 N}{1/2 N}$$

ถ้ากำหนดตัวแปรใหม่ว่า $N_a = n_c$ คือ จำนวนคณะกรรมการประเมินเนื้อหาที่เห็นว่าข้อสอบนั้นมีความจำเป็นในการวัดเนื้อหาเกี่ยวกับงานลักษณะนั้น ๆ

$$\text{กำหนดให้} \quad N_b = N - N_a$$

$$\begin{aligned}
 \text{เขียนเป็นสูตรใหม่ได้ว่า } CVR &= \frac{N_a - (N_a + N_b)/2}{1/2 N} \\
 &= \frac{N_a/2 - N_b/2}{1/2 N} \\
 &= \frac{N_a - N_b}{N} \text{ ----- สมการที่ 1}
 \end{aligned}$$

จาก สูตรการประมาณค่าการแจกแจงแบบทวินามด้วยการแจกแจงปกติ

$$Z = \frac{(N_a - NP) - .5}{\sqrt{NPNQ}}$$

สถานการณ์ กำหนดให้มีคณะกรรมการประเมินเนื้อหาในการตัดสินข้อสอบ 20 คน จากค่าวิกฤตของ Z ทางสถิติ(one tails) ที่ระดับนัยสำคัญ .05 มีค่า 1.645 แทนค่า

$$1.645 = \frac{[N_a - (20)(.50)] - .5}{\sqrt{(20)(.50)(.50)}}$$

แก้สมการเพื่อหาค่า N_a

$$N_a = 14.178$$

สรุปค่าต่ำสุดของ CVR ตามสมการที่ 1 เป็น

$$\frac{14.178 - 5.822}{20} = .42$$

การคำนวณค่าวิกฤต CVR ขึ้นตอนดังกล่าว แปรตามจำนวนคณะกรรมการประเมินเนื้อหา จึงได้ตารางค่าวิกฤต CVR คือ ระดับของค่า CVR ต่ำสุดที่ยอมรับได้ว่าข้อสอบข้อนั้นมีความตรงเชิงเนื้อหา ดังแสดงในตารางที่ 1

ตารางที่ 1 ค่าวิกฤตของ CVR ที่ยอมรับว่าข้อสอบมีความตรงเชิงเนื้อหา ที่ระดับนัยสำคัญ .05 แปรตามจำนวนคณะกรรมการประเมินเนื้อหา

| จำนวนคณะกรรมการ ประเมินเนื้อหา | ค่า CVRI ต่ำสุด |
|-----------------------------------|--------------------|
| 5 | 0.99 |
| 6 | 0.99 |
| 7 | 0.99 |
| 8 | 0.78 |
| 9 | 0.78 |
| 10 | 0.62 |
| 11 | 0.59 |
| 12 | 0.56 |
| 13 | 0.54 |
| 14 | 0.51 |
| 15 | 0.49 |
| 20 | 0.42 |
| 25 | 0.37 |
| 30 | 0.33 |
| 35 | 0.31 |
| 40 | 0.29 |

ตัวอย่างการนำตารางที่ 1 ไปใช้ ถ้ากำหนดให้มีคณะกรรมการตัดสินแบบสอบ 15 คน ตัดสินความจำเป็นของข้อสอบในการวัดเนื้อหาเกี่ยวกับงานลักษณะหนึ่ง คำนวณค่า CVR ของข้อสอบได้ค่าหนึ่ง เมื่อนำมาเทียบกับค่าในตารางที่ 1 มีค่า CVR ต่ำสุดที่ยอมรับได้ของผู้ตัดสิน 15 คน คือ .49 ถ้าข้อสอบที่มีค่า CVR สูงกว่า .49 จึงจะเป็นข้อสอบที่น่าพอใจ และถูกคัดเลือกให้ไว้ในแบบสอบ

มีข้อที่พึงระวังในการนำค่า CVR ไปใช้ ค่า CVR ของข้อสอบนั้นไม่เกี่ยวข้อง กับค่าอำนาจจำแนกหรือการวิเคราะห์คุณภาพข้อสอบแบบอื่น ๆ ด้วย

ดัชนีบ่งชี้ความตรงเชิงเนื้อหาของแบบสอบ(content validity Index) ค่า CVR เป็นค่าสถิติประจำข้อสอบที่มีประโยชน์ต่อการคัดเลือกข้อสอบไว้ในแบบสอบ นำจำนวนข้อสอบที่คัดเลือกไว้ในแบบสอบขั้นสุดท้ายมาคำนวณค่าดัชนีที่บ่งชี้ความตรงเชิงเนื้อหาของแบบสอบ(CVI) เป็นค่าเฉลี่ยของค่า CVR ของข้อสอบทั้งหมดที่คงอยู่ในแบบสอบ ซึ่งเป็นจุดเน้นสำคัญอยู่ที่ว่าดัชนีความตรงเชิงเนื้อหาของข้อกระทง(CVI) ไม่ใช่ค่าสัมประสิทธิ์สหสัมพันธ์

ค่า CVI หมายถึง การยอมรับส่วนที่คาบเกี่ยวกันระหว่างความสามารถในการปฏิบัติงานในหน้าที่ ซึ่งนิยามจากขอบเขตของงานที่ต้องปฏิบัติ กับการสร้างแบบสอบที่ได้รับการตรวจสอบแล้ว ค่า CVI นี้จึงเป็นค่าเฉลี่ยของเปอร์เซ็นต์ที่คาบเกี่ยวกันระหว่างข้อสอบในแบบสอบ และขอบเขตของมวลเนื้อหาในการปฏิบัติงาน

ตัวอย่างของการนำการตรวจสอบความตรงเชิงเนื้อหาที่คำนวณโดยใช้ค่า CVR ไปใช้

ตัวอย่างที่ 1 การนำไปใช้ในเบื้องต้น

ในการพัฒนางานทางด้านอุตสาหกรรมหนักทางการเกษตรที่ต้องการผู้มีความชำนาญงานทางช่าง มีวัตถุประสงค์ในการหาค่าความตรงของแบบสอบสำหรับคัดเลือกผู้สมัครที่ควรได้รับการพิจารณามารวมฝึกฝีมือทางเครื่องจักรและเครื่องกล

การวิเคราะห์งาน ชี้ให้เห็นว่าความรู้ด้านคณิตศาสตร์มีความจำเป็นสำหรับการปฏิบัติงานในหน้าที่ Lawshe(1975) ได้ทดลองนำแบบวัดความรู้ทางคณิตศาสตร์ คือแบบสอบ SRA Arithmetic Index มาเป็นแบบทดสอบให้คณะกรรมการประเมินเนื้อหาตัดสินความเกี่ยวข้องกับงาน

คณะกรรมการประเมินเนื้อหา มีจำนวน 175 คน จากโรงงาน 17 แห่ง ซึ่งประกอบด้วย หัวหน้าช่าง ช่าง บุคคลที่ฝึกเป็นช่าง ครูผู้สอน ผู้สมัครเข้าร่วมในโครงการฝึก เป็นต้น คณะกรรมการฯ จะได้รับแบบสอบ SRA Arithmetic Index จำนวน 45 ข้อ และแบบประเมินเนื้อหาในแบบสอบซึ่งมีตัวเลือก คือ ข้อสอบวัดสิ่งที่ "จำเป็นต่อการปฏิบัติงาน" หรือ "มีประโยชน์แต่ไม่จำเป็น" หรือ "มีประโยชน์แต่สามารถเรียนรู้ได้จากการปฏิบัติงาน"

ผลการประเมินในเชิงปริมาณ จำนวนข้อสอบที่ได้รับการประเมินว่า "จำเป็นต่อการปฏิบัติงาน" จะนำมาคำนวณค่า CVR ค่าความถี่ของ ค่า CVR ที่คำนวณได้ ดังแสดงในตารางที่ 2

ตารางที่ 2 การแจกแจงของค่า CVR ของแบบสอบ SRA Arithmetic Index ในการคัดเลือกช่างจำนวน 54 ชื่อ

| CVR _i | f |
|------------------|----|
| 90-94 | 13 |
| 85-89 | 12 |
| 80-84 | 0 |
| 75-79 | 3 |
| 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)

ผลจากการตัดสินเนื้อหาของข้อสอบนำไปเปรียบเทียบกับเกณฑ์ขั้นต่ำของ Schipper แล้วนำค่า CVR ของข้อสอบทุกข้อมาคำนวณค่าดัชนีความตรงเชิงเนื้อหาของแบบสอบ(CVI)เท่ากับ .67 และเมื่อคัดข้อสอบที่ไม่ผ่านเกณฑ์ของชิปเปอร์ออกไป 1 ข้อ ค่า CVI ที่คำนวณได้ มีค่าเป็น .69

ตัวอย่างที่ 2 การปรับให้เหมาะกับการนำไปใช้

ความสามารถในการประยุกต์ใช้ จากตัวอย่างที่ 1 เป็นตัวอย่างของการนำไปใช้ในงานที่มีลักษณะงานที่วัดโดยแบบสอบที่มีเพียงเนื้อหาเดียว ค่า CVI รวม จึงสามารถคำนวณได้ ในตัวอย่างที่ 2 นี้ เป็นการขยายการนำไปใช้ในลักษณะของการตั้งเคราะห์ค่าความตรง (Lawshe and Steinberg, 1995 cited in Pitayanuwat, 1981) โดยระบุว่า

- 1) เมื่อขอบเขตของงานที่ต้องปฏิบัติตามโดยจำนวนข้อความที่ใช้เฉพาะงาน
- 2) เมื่อแบบสอบที่สร้างขึ้นภายใต้การตรวจสอบอย่างดี และ/หรือ มีความเที่ยง

แบบสอบคล่องภายในสูง

การศึกษาครั้งนี้ใช้แบบรายการงานด้านเสมียน(Clerical Task Inventory; CTI) ซึ่งมีข้อความงานด้านเสมียนที่เป็นมาตรฐาน 139 ข้อความ เรียงตามความถี่ของงานที่ต้องทำงานหน่วยงานต่าง ๆ เป็นต้นว่า

ข้อความที่ 16. งานเรียงพิมพ์ งานเขียนจดหมายโต้ตอบ หรืองานบันทึก เป็นงานที่ต้องทำเป็นประจำ

ข้อความที่ 38. การตรวจสอบหรือการพิสูจน์ข้อมูลที่เป็นตัวเลข โดยการคำนวณซ้ำจากต้นฉบับ ทั้งการใช้เครื่องคำนวณช่วยและไม่ใช้เครื่องคำนวณช่วย

ฯลฯ

แบบสอบที่นำมาใช้ คือ แบบสอบ Purdue Clerical Adaptability Test (PCAT) ซึ่งวัดความสามารถใน 6 ด้าน คือ การสะกดคำ การคำนวณ การตรวจสอบ คำศัพท์ การคัดลอก และการมีเหตุผลเชิงคณิตศาสตร์

ประเด็นของค่า CVR_t คณะกรรมการประเมินเนื้อหา มี 14 คนที่กำลังทำงานด้านนี้ และจำนวนครึ่งหนึ่งเป็นผู้ที่มีหน้าที่ในงานนี้โดยตรง อีกครึ่งหนึ่งเป็นผู้ควบคุมงาน คณะกรรมการประเมินเนื้อหาจะได้รับ แบบรายการงานด้านเสมียน (CTI) เป็นขอบเขตของงานที่ต้องปฏิบัติ และ แบบสอบ PCAT เป็นแบบสอบที่จะนำมาประเมินว่า คอบ

เกี่ยวกับงานที่ต้องปฏิบัติอย่างไร ผลการศึกษาได้ค่า CVR_t ที่คำนวณจากแบบรายการงานทั้ง 139 รายการ เทียบกับ แบบสอบ PCAT ที่วัดความสามารถในแต่ละด้าน รวม 6 ด้าน ข้อสอบที่คัดเลือกไว้เป็นข้อสอบที่คณะกรรมการฯ เห็นว่า "จำเป็น" สำหรับการ ทำงาน มีค่า CVR ผ่านเกณฑ์ของชิปเปอร์

การบรรยายลักษณะงาน สมมติให้มีลักษณะงานด้านเสมือนของบริษัทแห่งหนึ่ง ให้ชื่อว่า งาน A ตั้งทีมในการวิเคราะห์งาน ประกอบด้วย นักวิเคราะห์งาน 8 คน เสมียน 4 คน ผู้ตรวจการ 4 คน โดยคณะกรรมการฯ เหล่านี้จะพิจารณาแบบรายการงาน เสมียน CTI แล้วตรวจสอบว่ารายการใดเป็นงานที่บริษัททำ นอกจากนี้ ยังให้เดิมงานที่บริษัทต้องทำแต่ไม่ได้กำหนดไว้ในรายการด้วย ผลการวิเคราะห์งาน พบว่า งาน A มีขอบเขตของงานที่ต้องทำอยู่ 47 รายการ

ดัชนีความตรงเชิงเนื้อหา มีผล 2 ส่วนที่ได้จากการศึกษาในครั้งนี้

ส่วนแรก เป็นตารางค่า CVR_t ของรายการเสมือนในงาน A ต้องใช้แบบสอบด้านต่าง ๆ ทั้ง 6 ด้านในแบบสอบ PCAT ในการทดสอบ เช่น ตารางที่ 3 แสดงภาระงานที่กำหนด ไว้ในงาน A ที่ใช้แบบสอบ PCAT ด้านการคำนวณในการทดสอบ

ตารางที่ 3 ภาระงานเสมือนที่กำหนดไว้ในงาน A สำหรับค่า CVI ของแบบสอบ PCAT ด้านการคำนวณ

| เนื้อหาที่ | ภาระงานเสมือน | CVR_t |
|--|--|---------|
| 1 | การคำนวณง่าย ๆ เช่น การบวก หรือ การลบ | .99 |
| 3 | การคำนวณที่เป็นขั้นคองมากกว่า 1 ขั้น เช่น การคูณ การหาร | .99 |
| 4 | การจับบันทึกข้อกระทง หรือ จำนวนช่วงเวลาทำการ โดยไม่ใช้เครื่องมือช่วย | .99 |
| 45 | กำหนดอัตรา ราคา จำนวน และการจัดประเภทข้อมูล ฯลฯ | .99 |
| 9 | การคำนวณที่ซับซ้อน เช่น การถอดราก เลขยกกำลัง ฯลฯ | .83 |
| 38 | การคำนวณซ้ำเพื่อแก้ไขหรือตรวจสอบข้อมูลที่เป็นตัวเลข | .83 |
| 39 | บอกความคลาดเคลื่อนที่พบ การบันทึกข้อมูล ฯลฯ | .83 |
| ค่าดัชนีความตรงเชิงเนื้อหา CVI (ค่าเฉลี่ยของ CVR_t) | | .92 |

ส่วนที่ 2 แสดงค่าดัชนีความตรงเชิงเนื้อหา(CVI)ของแบบสอบ PCAT ทั้ง 6 ด้าน ของงาน A และจำนวนรายงานของภาระงาน A ที่ถูกกำหนดไว้ในการสอบแต่ละด้าน ดังแสดงใน ตารางที่ 4

ตารางที่ 4 ค่าดัชนีความตรงเชิงเนื้อหา(CVI)ของแบบสอบ PCAT ทั้ง 6 ด้านของงาน A และ จำนวนภาระงาน A ที่ถูกกำหนดไว้ในการสอบแต่ละด้าน

| ส่วนของแบบสอบ | จำนวนของงานที่กำหนด | CVI | ค่าเฉลี่ยของน้ำหนัก |
|---------------------------|---------------------|-----|---------------------|
| การสะกดคำ | 4 | .87 | .87 |
| การคำนวณ | 7 | .92 | .95 |
| การตรวจสอบ | 5 | .73 | .79 |
| คำศัพท์ | 3 | .72 | .71 |
| การคัดลอก | 3 | .94 | .96 |
| การมีเหตุผลเชิงคณิตศาสตร์ | 4 | .87 | .89 |

ความต้องการขั้นต่ำสุด เป็นส่วนสำคัญที่ความต้องการจะเน้นให้เห็นว่า จำนวนของงานอยู่ในแต่ละแบบสอบถูกรับรู้ว่า "จำเป็น" ซึ่งไม่ตรงกับปัญหาที่ความต้องการ เมื่อค่า CVR_t สำหรับแบบสอบถูกคำนวณสำหรับแต่ละงานซึ่งเป็นองค์ประกอบของงานที่ต้องปฏิบัติ การพัฒนาดัชนีความตรงเชิงเนื้อหา สำหรับขอบเขตของงานทั้งหมดที่ต้องปฏิบัติ ต้องคำนึงถึงความจริงตามธรรมชาติของงานด้วย

กล่าวคือ ถ้าส่วนของงานที่ต้องการมีทักษะเดียว และทักษะนี้ต้องการบุคคลที่มีลักษณะเฉพาะ ส่วนของทักษะอื่น ๆ ที่ความต้องการไม่เหมือนงานนี้จะนำไปใช้ไม่ได้

ถ้าส่วนของงานที่ต้องทำมีมากกว่าทักษะที่ต้องการ หลักฐานของทักษะสำหรับงานทั้งหมด คือ ค่า CVR_t ที่มีค่าสูง ๆ ในงานที่ต้องการ ในปัญหานี้ ไม่มีค่า CVR ที่แท้จริงสำหรับแต่ละงาน มีเพียงค่า CVR_t ที่เป็นค่าโดยประมาณ ที่ตัดสินโดยคณะกรรมการประเมินเนื้อหา การนำค่า CVR_t ของข้อสอบทั้งหมดมาใช้จะให้ผลที่ต่ำมาก ถ้าไม่ได้พิจารณาความเที่ยงของคณะกรรมการประเมินเนื้อหาด้วย

ประเด็นในเรื่องของน้ำหนัก โดยเฉพาะในเรื่องของการวิเคราะห์งาน การให้น้ำหนักความสำคัญของงานเป็นประเด็นที่น่าพิจารณา โดยเฉพาะในเรื่องของความตรงเชิงเนื้อหา มีข้อควรพิจารณาในเรื่องของน้ำหนักความสำคัญทั้งในเชิงเหตุผลและข้อมูลเชิงประจักษ์ ดังนี้

ในเชิงเหตุผล เนื่องจากขอบเขตของงานที่ต้องปฏิบัติต้องประมาณจากเอกภพของงานที่ต้องปฏิบัติทั้งหมด การให้เหตุผลสำหรับการประเมินที่ไม่สอดคล้องกับการให้น้ำหนักความสำคัญ อย่างไรก็ตาม ถ้าการนิยามขอบเขตของงานที่ต้องปฏิบัติในบทความนี้ได้ยอมรับผลการตัดสินของคณะกรรมการประเมินเนื้อหาที่ว่า ข้อสอบนี้เป็นข้อสอบที่วัดเนื้อหาที่ "จำเป็น" ในการปฏิบัติงาน น่าจะสำคัญมากกว่าที่จะใช้มาตรฐานประมาณค่า (Rating) ประมาณน้ำหนักความสำคัญของข้อสอบ เพราะ ค่าที่ได้จาก CVR เป็นค่าที่ไม่ต่อเนื่อง ในขณะที่ค่าที่ได้จากมาตรฐานประมาณค่าเป็นค่าที่ต่อเนื่อง

ในข้อมูลเชิงประจักษ์ การหาเหตุผลสนับสนุนกับการทดลองเพื่อหาน้ำหนักความสำคัญไม่ใช่สิ่งที่ทำร่วมกันได้ ในการออกแบบเพื่อการศึกษาในครั้งนี้ อย่างไรก็ตามในตัวอย่างที่ 2 ในตารางที่ 4 เมื่อทีมการวิเคราะห์งานได้ทดลองประเมินน้ำหนักความสำคัญของงาน โดยเรียงน้ำหนักความสำคัญของภาระงานที่สำคัญกับงาน เทียบกับค่า CVR ที่คำนวณแล้ว ให้ผลที่ไม่ต่างกัน

หลังจากที่ Lawshe (1975) ได้เสนอการตรวจสอบความตรงเชิงเนื้อหาในเชิงปริมาณแล้ว มีผู้สนับสนุนการนำไปใช้และผู้คัดค้านมากมาย (Kleiman และ Faley, 1978) ผู้วิจัย สรุปเป็นประเด็นที่เป็นข้อสนับสนุน และข้อกำหนดในการนำการตรวจสอบความตรงเชิงเนื้อหาไปใช้ ดังนี้

1. ข้อสนับสนุนในการนำการตรวจสอบความตรงเชิงเนื้อหาไปใช้ในแบบสอบคัดเลือกบุคคลเข้าทำงาน

บทความของ Ebel (1977) ได้สนับสนุนว่า การใช้การตรวจสอบความตรงเชิงเนื้อหาในแบบสอบคัดเลือกบุคคลเข้าทำงานที่ให้ค่าในเชิงปริมาณนี้ ทำให้ความตรงเชิงเนื้อหากลายเป็นแนวปฏิบัติที่เป็นประโยชน์ที่เคียวในการพิจารณาว่าข้อสอบวัดในสิ่งที่ตั้งใจจะวัด หรือไม่ และยังพิจารณาได้ว่า ข้อสอบแต่ละข้อเป็นข้อสอบที่ตั้งใจจะวัดในสิ่งที่จะวัดเป็นเชิงปริมาณได้อย่างเป็นอิสระจากกันด้วย นอกจากนี้ Ebel ยัง

กล่าวว่ ความตรงเชิงเนื้อหาเป็นพื้นฐานของการตรวจสอบความตรงทุกชนิด และถ้าไม่มีความตรงชนิดนี้ก็ไม่มีความตรงชนิดอื่น ๆ เช่นกัน สิ่งสำคัญ คือ ความเหมาะสมของกรอบที่ใช้เป็นเกณฑ์ในการตัดสินใจตรวจสอบความตรงของแบบสอบร่วมกับ การตัดสินใจอย่างมีเหตุผล ซึ่งยังไม่มีข้อมูลใด ๆ จะทำหน้าที่แทนสำหรับการตัดสินใจที่ดีได้ ผลของการตัดสินใจเป็นเหมือนพื้นฐานของการกำหนดความตรงของแบบสอบที่เป็นที่ยอมรับได้ ต่อมา ในปี ค.ศ.1978 Uniform Guidelines on Employee Selection Procedures (cited in Sackett, 1987) ได้สนับสนุนการใช้ความตรงเชิงเนื้อหาในแบบสอบคัดเลือกบุคคลเข้าทำงานว่า เป็นกลยุทธ์ที่มีความเหมาะสมในการตรวจสอบเกี่ยวกับตัวอย่างของงาน เนื่องจากข้อจำกัดของกลยุทธ์ในการตรวจสอบความตรงโดยวิธีอื่น เช่น การใช้ความตรงเชิงเกณฑ์สัมพัทธ์ที่ต้องศึกษาเกณฑ์จากกลุ่มตัวอย่างขนาดใหญ่ และมีความเที่ยงสูงซึ่งทำได้ยาก และการใช้ความตรงเชิงโครงสร้างก็ยากเกินไปที่คนทั่วไปจะทำความเข้าใจได้ แนวคิดนี้ยังสอดคล้องกับการศึกษาของ Kleiman และ Faley (1978) ที่ศึกษากรณีการฟ้องร้องต่อศาลในเรื่องของการประเมินความตรงของการสอบเขียนตอบ 31 กรณี ที่เป็นผลจากการใช้กลยุทธ์ตัดสินใจความตรงเชิงประจักษ์ เช่น ความตรงเชิงทำนาย หรือความตรงร่วมสมัย เมื่อใช้ตัดสินผลการสอบคัดเลือกคนเข้าทำงานแล้วมีการตีความที่ผิด ๆ ก่อให้เกิดคดีฟ้องร้องกันขึ้น ซึ่งส่งผลให้คนเห็นความสำคัญในการประเมินความตรงเชิงเนื้อหา โดยเฉพาะเมื่อ Lawshe ได้พัฒนาวิธีการตัดสินใจความตรงเชิงเนื้อหาเป็นปริมาณขึ้น ซึ่งง่ายต่อการแปลความหมาย และการนำไปใช้ จึงเกิดเป็นแนวความคิดว่า การใช้การตัดสินใจความตรงเชิงเนื้อหาน่าจะเป็นทางออกที่ดี นอกจากนี้งานวิจัยของ Hirsh และคณะ(1986) รวมทั้ง Carrier และคณะ(1990) ที่ได้้นำการตรวจสอบความตรงเชิงเนื้อหาไปทดลองใช้เพื่อเปรียบเทียบกับตรวจสอบความตรงเชิงเกณฑ์สัมพัทธ์ พบว่า มีความสอดคล้องกัน

2. ข้อกำหนดในการนำการตรวจสอบความตรงเชิงเนื้อหาไปใช้ในแบบสอบคัดเลือกบุคคลเข้าทำงาน

Ebel (1977) ได้ให้ข้อเสนอแนะเกี่ยวกับการนำตรวจสอบความตรงเชิงเนื้อหาไปใช้ในแบบสอบคัดเลือกบุคคลเข้าทำงานว่า สามารถนำมาใช้เมื่อต้องการความถูกต้องในเรื่องของความตรง แต่มีข้อจำกัดของสถานการณ์ว่า สิ่งที่ยั่งยืนต้องสามารถ

สังเกตได้จากพฤติกรรมที่ต้องการ และมีความสัมพันธ์กับสิ่งที่อ้างถึง ซึ่งหมายความว่า การอ้างถึงพฤติกรรมของบุคคลที่ใช้ในการตรวจสอบความตรงเชิงเนื้อหา ต้องเป็น พฤติกรรมที่สังเกตได้โดยตรง และมีความสัมพันธ์กับขอบเขตของงานที่ต้องปฏิบัติโดย แท้จริง และผลการศึกษาของ Kleiman และ Faley (1978) ที่ให้ฟังระวิงว่า ผู้สร้างแบบ สอบได้กล่าวอ้างถึงว่าได้พิจารณาความตรงเชิงเนื้อหาแล้ว แต่แท้ที่จริงทำเพียงแค่การ มองดูแบบสอบอย่างผิวเผิน "face validity" ของผู้เชี่ยวชาญเท่านั้น

ผลการศึกษาของ Kleiman and Faley (1978) เพื่อกำหนดมาตรฐานของวิธีการ นำการตรวจสอบความตรงเชิงเนื้อหาไว้ 3 ประการ คือ

1. แบบสอบที่สร้างขึ้นนั้นต้องสร้างจากการกำหนดโดเมนของเนื้อหาของงาน หรือขอบเขตของงานที่ต้องปฏิบัติ

2. การตัดสินใจความตรงให้ตัดสินส่วนที่คาบเกี่ยวกัน(overlap) ระหว่างประเด็น เนื้อหาในแบบสอบกับขอบเขตของงานที่ปฏิบัติ ยิ่งคาบเกี่ยวกันมากเท่าไรก็จะมี ความตรงเชิงเนื้อหามากเท่านั้น ทั้งนี้ ต้องให้ความสำคัญกับขั้นตอนในการพัฒนาแบบ ทดสอบด้วย

3. ในบางกรณีอาจไม่มีความจำเป็นต้องใช้ผลการวิเคราะห์งานเพื่อกำหนดเนื้อ หา เพราะลักษณะงานบางอย่างนั้นสามารถเรียนรู้ได้จากการทำงาน เช่น งานโรงงานที่มี การกำหนดขั้นตอนของงานที่ต้องทำไว้อย่างชัดเจนแล้ว ฉะนั้น การวัดก็ควรวัดความ พร้อมในการเรียนรู้งานได้อย่างรวดเร็ว ซึ่งอาศัยการวัดทางด้านจิตวิทยาแทน เช่น การ วัดเชาว์ปัญญา หรือ การวัดบุคลิกภาพ เป็นต้น ซึ่งต้องอาศัยหลักฐานการตรวจสอบ ความตรงแบบอื่น

ในเรื่องของประเภทของแบบสอบที่ใช้การตรวจสอบความตรงเชิงเนื้อหา Lawshe (1975) กล่าวว่าจำกัคอยู่เฉพาะ แบบสอบวัดความชำนาญ หรือทักษะที่ง่าย ๆ หรือ แบบสอบวัดความรู้เกี่ยวกับงาน หรือแบบสอบที่วัดตัวอย่างงาน โดยผู้ปฏิบัติงาน หรือผู้ควบคุมลักษณะงานนั้นสามารถตัดสินพฤติกรรมนั้นได้ แต่ถ้าพฤติกรรมที่ต้องการ นั้นเป็นพฤติกรรมเชิงนามธรรมระดับสูง เช่น ความสามารถในการนิรนัยอย่างมีเหตุผล ต้องใช้ความตรงเชิงโครงสร้าง (construct validity) เป็นตัวชี้วัด ซึ่งในเรื่องนี้สอดคล้อง กับงานวิจัยที่ Wollack (1976) ได้ศึกษาว่าควรนำการตรวจสอบความตรงเชิงเนื้อหาไป

ใช้ในแบบสอบประเภทใดบ้าง พบว่า ถ้าเป็นแบบสอบที่วัดเกี่ยวกับทักษะของงาน (job skill) ความรู้หรือพฤติกรรม (knowledge or behavior) เกี่ยวกับงานแล้ว การตรวจสอบความตรงเชิงเนื้อหาไม่มีความเหมาะสม แต่ถ้าเป็นการวัดเกี่ยวกับ trait หรือ constructs เช่น การวัดเข่าปัญญา หรือบุคลิกภาพ นั้น การตรวจสอบความตรงเชิงเนื้อหาไม่เหมาะสม

ตอนที่ 3 ปัจจัยที่มีผลต่อการตรวจสอบความตรงเชิงเนื้อหาในเชิงปริมาณ

จากแนวคิดของการตรวจสอบความตรงเชิงเนื้อหาของแบบสอบคัดเลือกบุคคลเข้าทำงานที่ Lawshe(1975) เสนอแนะ เป็นการตรวจสอบที่สามารถนำมาใช้ได้ง่าย และยังได้ค่าของระดับความตรงของแบบสอบในเชิงปริมาณซึ่งเป็นผลจากการสรุปความเห็นของคณะกรรมการประเมินเนื้อหาที่เกี่ยวข้องกับงานลักษณะนั้น ว่าเป็นข้อสอบที่มีเนื้อหาที่ "จำเป็น" ต่อการวัดลักษณะงานนั้น ๆ สูตรที่ลอว์ชีเสนอได้ก่อให้เกิดผลในทางปฏิบัติ มีการศึกษาวิจัยที่เกี่ยวข้องกับผลการนำการตรวจสอบความตรงเชิงเนื้อหาของแบบสอบคัดเลือกบุคคลเข้าทำงานไปใช้ และพบประเด็นสำคัญที่น่าสนใจเกี่ยวกับตัวแปรที่เป็นส่งผลต่อการนำไปใช้ ดังนี้

1) **ตารางแทนที่เป็นผลจากการวิเคราะห์งาน** ผลการวิเคราะห์งานเพื่อกำหนดโดเมนของเนื้อหาในการปฏิบัติงาน มีความจำเป็นต่อการตรวจสอบความตรงเชิงเนื้อหาตามหลักการของ American Psychological Association[APA] (1975 cited in Kleiman และ Faley (1978)) กล่าวว่าการวิเคราะห์งานจะต้องได้รับการพิจารณาเพื่อเป็นข้อมูลในการตัดสินใจเกี่ยวกับตัวบุคคล ขอบเขตของงานที่ต้องปฏิบัติทั้งในงานที่ต้องกระทำและความรับผิดชอบเกี่ยวกับงาน ลักษณะการวิเคราะห์งานที่ดีต้องมองให้ถึงแก่น (core) ของงาน แหล่งข้อมูลควรได้จากการสัมภาษณ์ผู้เชี่ยวชาญเกี่ยวกับงาน การศึกษาขั้นตอนและรายละเอียดของการปฏิบัติงานจริง รวมทั้งแหล่งข้อมูลจากภายนอกอื่น ๆ ที่เกี่ยวข้อง การประเมินแบบสอบเพื่อเปรียบเทียบระหว่างเนื้อหาในแบบสอบกับงานที่เป็นผลจากการวิเคราะห์งานเพื่อให้ได้รายละเอียดของงาน ถ้าสอดคล้องกันก็จะถือว่าแบบสอบมีความตรงเชิงเนื้อหา

มีงานวิจัยที่สนับสนุนการใช้ผลการวิเคราะห์ลักษณะงานเพื่อกำหนดเนื้อหาของงาน คือ งานวิจัยของ Krzystofiak, Newman และ Anderson (1979) ได้ศึกษา

เพื่อยืนยันผลของการวิเคราะห์งานว่า ให้สารสนเทศที่สำคัญในการตัดสินใจเกี่ยวกับเนื้อหาของงาน และการตัดสินใจในการบริหารงานบุคคล ในงานวิจัยเรื่องนี้ยังได้อ้างถึงงานวิจัยของ Hemphill (1959) และงานวิจัยของ Pinto และ Tomow (1975) ที่ศึกษาในทำนองเดียวกัน เพื่อชี้ให้เห็นถึงความสำคัญของการวิเคราะห์งานเพื่อให้ได้ลักษณะงานและนำมาใช้เป็นสารสนเทศที่สำคัญในการสร้างแบบสอบคัดเลือกบุคคลเข้าทำงาน รวมทั้งการตัดสินใจที่เกี่ยวข้องกับการบริหารงาน นอกจากนี้ ยังมีงานวิจัยของ Schwarz (1977) ที่ศึกษาถึงการใช้ความตรงเชิงเนื้อหาในการสร้างแบบสอบเพื่อคัดเลือกบุคคลในระบบคุณธรรม (merit system) โดยใช้เทคนิคที่เกี่ยวข้องกับการสุ่มตัวอย่างงาน (job sampling) ว่าให้สารสนเทศที่จำเป็นสำหรับประเมินกระบวนการในการคัดเลือก และเป็นหลักการในการตรวจสอบความตรงเชิงเนื้อหาของแบบสอบคัดเลือกบุคคลเข้าทำงาน งานวิจัยของ Prien (1977) ได้ศึกษาถึง หน้าที่ของการวิเคราะห์ลักษณะงานในการตรวจสอบความตรงเชิงเนื้อหาของแบบสอบคัดเลือกคนเข้าทำงานว่า เป็นวิธีการเพื่อให้ได้ลักษณะงานที่วิเคราะห์เป็นตัวกำหนดเนื้อหาของแบบสอบ ต่อมา งานวิจัยของ Hughes และ Prien (1989) ที่ชี้ให้เห็นถึงความสำคัญของการประเมินทักษะที่ใช้ในงานที่สามารถเชื่อมโยงสู่การตัดสินใจเพื่อพัฒนาการจัดทำกรอบเนื้อหาของแบบสอบคัดเลือกบุคคลเข้าทำงาน ในงานวิจัยเรื่องนี้ยังชี้ให้เห็นถึงลักษณะงานที่ต่างกันจะให้สารสนเทศที่ต่างกัน ซึ่งจะส่งผลต่อความยากง่ายในการตัดสินใจ และความเชี่ยวชาญของผู้ประเมินด้วย

จะเห็นได้ว่า จากผลการวิจัยที่กล่าวมาข้างต้น เป็นการยืนยันถึงผลการวิเคราะห์งานให้สารสนเทศที่จำเป็นต่อการกำหนดเนื้อหาที่ใช้ในการสร้างแบบสอบคัดเลือกบุคคลเข้าทำงาน และนำไปใช้ในการตัดสินใจของผู้ประเมินในการประเมินความตรงเชิงเนื้อหาของแบบสอบด้วย

2) ระดับความเชี่ยวชาญของผู้ประเมินความตรงเชิงเนื้อหา

ในปี ค.ศ. 1986 Hirsh และ คณะ ได้ศึกษาการประมาณค่าความตรงของแบบสอบคัดเลือกคนเข้าทำงานจากผลการตัดสินใจของผู้เชี่ยวชาญที่มีประสบการณ์น้อย ซึ่งพบว่า มีความคลาดเคลื่อนอย่างเป็นระบบมากกว่าผู้เชี่ยวชาญที่มีประสบการณ์มาก ขณะเดียวกันผลการศึกษาของ Carrier และคณะ (1990) ที่ศึกษาการสอบคัดเลือก

ตัวแทนขายประกัน โดยใช้ผู้เชี่ยวชาญมากกว่า 300 คน คัดสินแบบสอบ โดยใช้ค่าสัดส่วนความตรงเชิงเนื้อหา(CVR) ในการประมาณค่าความตรงเชิงเนื้อหา ไม่พบความแตกต่างในการตัดสินความตรงเชิงเนื้อหาของผู้เชี่ยวชาญที่มีความเชี่ยวชาญในด้านประสบการณ์การทำงานที่ต่างกัน

จากผลการวิจัย 2 เรื่องที่ศึกษา ยังไม่อาจสรุปถึงความแตกต่างของผลการตัดสินความตรงเชิงเนื้อหาของแบบสอบเมื่อผู้ประเมินมีระดับความเชี่ยวชาญต่างกัน ได้ชัดเจนนัก ซึ่งเป็นตัวแปรที่ต้องศึกษาต่อไป

8) ตัวแปรที่เกี่ยวข้องกับการตัดสินใจ

งานวิจัยของ Newman และคณะ (1987) ได้ศึกษาปัจจัยที่มีอิทธิพลต่อกระบวนการในการตัดสินใจแล้วพบว่า ตัวแปรด้านคุณลักษณะของผู้ตัดสินใจ และสารสนเทศที่น่าเสนอให้แก่ผู้ตัดสินใจ เป็นปัจจัยที่มีอิทธิพลต่อการตัดสินใจที่ถูกต้อง และงานวิจัยของ อังฉรียา ปราบอริพ่าย (2531) ที่ศึกษาผลการให้สารสนเทศแก่ผู้เชี่ยวชาญต่อการกำหนดจุดคัดของแบบสอบผลสัมฤทธิ์วิชาคณิตศาสตร์ พบว่า ผลการตัดสินใจของผู้เชี่ยวชาญที่ไม่มีสารสนเทศและผู้เชี่ยวชาญที่มีสารสนเทศแตกต่างกันอย่างมีนัยสำคัญทางสถิติ ในปี พ.ศ. 2538 งานวิจัยของ ชูดี วัฒนานนท์ ที่ได้ศึกษาเงื่อนไขที่เหมาะสมสำหรับการตรวจสอบความตรงเชิงเนื้อหาของแบบสอบผลสัมฤทธิ์ โดยใช้แนวคิดของ Lawshe (1975) และวิธีการวิเคราะห์สัดส่วนความสอดคล้องของเนื้อหาภายใต้สถานการณ์ 3 แบบ คือ ผู้ตัดสินเนื้อหาของแบบสอบที่มีความเชี่ยวชาญต่างกัน 2 ระดับ คือระดับดีเยี่ยม และระดับดี ต่างกันที่ประสบการณ์ในการทำงาน และ รูปแบบของการให้ข้อมูลเกี่ยวกับเนื้อหาที่ต้องการวัดต่างกัน 3 แบบ ที่มีข้อมูลเพิ่มขึ้นจากรูปแบบที่ 1 ไปยังรูปแบบที่ 3 และ จำนวนของผู้ตัดสินเนื้อหาที่ต่างกัน 9 แบบ คือ มีจำนวนผู้ตัดสินเนื้อหา ตั้งแต่ 2 คน ถึง 10 คน ผลการศึกษา พบว่า เมื่อผู้เชี่ยวชาญระดับดีได้รับข้อมูลเกี่ยวกับเนื้อหาที่ต้องการวัดรูปแบบที่ 3 จะมีค่าความคงที่ของค่าดัชนีมากกว่าความคงที่ของค่าดัชนีจากกลุ่มผู้เชี่ยวชาญระดับดีที่ได้รับข้อมูลในรูปแบบอื่น ๆ อย่างมีนัยสำคัญทางสถิติ แต่ในกลุ่มผู้เชี่ยวชาญระดับดีเยี่ยมไม่ว่าจะได้รับข้อมูลในรูปแบบใดก็ไม่มีความแตกต่างของความคงที่ของค่าดัชนี และเมื่อเปรียบเทียบความคงที่ของค่าดัชนีในกลุ่มผู้เชี่ยวชาญระดับดีเยี่ยมที่ได้รับข้อมูลรูปแบบที่ 1 กับ ผู้เชี่ยวชาญ

ระดับคิดที่ได้รับข้อมูลในรูปแบบที่ 3 พบว่ามีความแตกต่างอย่างมีนัยสำคัญทางสถิติ รวมทั้ง เมื่อเปรียบเทียบความคงที่ของค่าดัชนีใน 2 กลุ่มดังกล่าว โดยใช้จำนวนผู้เชี่ยวชาญเป็นเกณฑ์ในการเปรียบเทียบพบว่า ความคงที่ของค่าดัชนีที่ได้จากกลุ่มผู้เชี่ยวชาญที่มีจำนวนมากมีความคงที่มากกว่าความแปรปรวนของค่าดัชนี ที่ได้จากกลุ่มผู้เชี่ยวชาญที่มีจำนวนน้อยอย่างมีนัยสำคัญทางสถิติ

จากผลการวิจัยตัวแปรที่เกี่ยวข้องกับผลการตัดสินใจ พบว่า ทั้งความเชี่ยวชาญของผู้ตัดสินใจและสารสนเทศที่ให้แก่มือผู้เชี่ยวชาญในการตัดสินใจน่าจะส่งต่อผลการตัดสินใจที่แตกต่างกัน

การกำหนดกรอบแนวคิดในการวิจัย

จากผลการศึกษาเกี่ยวกับความตรงเชิงเนื้อหาในเชิงปริมาณโดยใช้ค่าสัดส่วนความตรงเชิงเนื้อหา (CVR) ของ Lawshe(1975) ที่ผู้วิจัยนำเสนอดังกล่าวข้างต้น ทำให้ผู้วิจัยคิดว่า ข้อมูลที่ศึกษาเกี่ยวกับตัวแปรที่ส่งต่อการตัดสินใจความตรงเชิงเนื้อหาของแบบสอบถามคัดเลือกคนเข้าทำงาน คือ การวิเคราะห์งานเพื่อใช้เป็นกรอบเนื้อหาในการสร้างแบบสอบถามและนำเสนอให้ผู้เชี่ยวชาญใช้ประเมินเนื้อหาของแบบสอบถาม ซึ่งยังไม่ได้ออกันพบว่าปริมาณข้อมูลแบบใดที่ส่งผลให้ผู้ประเมินใช้ตัดสินใจได้ดีที่สุด และระดับความเชี่ยวชาญของผู้ตัดสินใจน่าจะมีผลต่อการตัดสินใจความตรงเชิงเนื้อหาแต่ผลการศึกษาที่ผ่านมาซึ่งไม่มีข้อสรุปที่ชัดเจนในเรื่องนี้ รวมทั้งผลการศึกษาถึงตัวแปรที่เกี่ยวข้องกับการตัดสินใจที่แตกต่างกันเมื่อได้รับสารสนเทศที่ต่างกัน และผู้ประเมินเนื้อหาที่มีระดับความเชี่ยวชาญที่ต่างกัน ซึ่งยังไม่มีการศึกษาที่อธิบายถึงความสัมพันธ์ระดับความเชี่ยวชาญและปริมาณข้อมูลให้ผู้เชี่ยวชาญต่อการตัดสินใจความตรงเชิงเนื้อหาของแบบสอบถามเพื่อคัดเลือกบุคคลเข้าทำงาน ข้อสงสัยในประเด็นต่าง ๆ เหล่านี้ ส่งผลให้ผู้วิจัยสนใจที่จะศึกษาปฏิสัมพันธ์ของระดับความเชี่ยวชาญและปริมาณข้อมูลให้ผู้เชี่ยวชาญต่อการตัดสินใจความตรงเชิงเนื้อหาของแบบสอบถาม เพื่อตอบคำถามในประเด็นต่าง ๆ ที่ยังเป็นปัญหา กล่าวคือ ในกระบวนการสร้างแบบสอบถามคัดเลือกบุคคลเข้าทำงาน หลักฐานที่แสดงถึงความตรงเชิงเนื้อหาของแบบสอบถาม ควรกำหนดคุณลักษณะของผู้ประเมินอย่างไร และควรให้ปริมาณข้อมูลแก่ผู้ประเมินในรูปแบบใด และที่สำคัญการตัดสินใจความตรงเชิงเนื้อหาของแบบสอบถามขึ้นกับตัวแปรทั้งสองในลักษณะใด