



บทที่ 2

วรรณคดีที่เกี่ยวข้อง

ในบทนี้ผู้วิจัยได้เสนอ หลักการและมโนทัศน์ที่ได้ศึกษามาจากเอกสาร คำรา บทความ และงานวิจัยต่าง ๆ ซึ่งนำมาเสนอเป็น 3 ตอน ดังนี้

ตอนที่ 1 มโนทัศน์เกี่ยวกับความสามารถในการแก้ปัญหา

ตอนที่ 2 มโนทัศน์เกี่ยวกับคุณภาพของแบบสอบ

ตอนที่ 3 งานวิจัยที่เกี่ยวข้อง

ตอนที่ 1 มโนทัศน์เกี่ยวกับความสามารถในการแก้ปัญหา

ไมเยอร์ และเฮดเจอร์เคน (Myer and Hedgerken 1962: 200) ได้ให้ความหมายของปัญหาไว้ว่า ปัญหาหมายถึง เหตุการณ์ที่ก่อให้เกิดอุปสรรคต่อการดำเนินงาน และขัดขวางมิให้บรรลุเป้าหมาย ซึ่งสอดคล้องกับ ยุพิน พิพิธกุล (2524: 24) ได้กล่าวถึงปัญหาว่า ปัญหาคือเรื่องราวที่เกี่ยวกับความสงสัย ความไม่แน่ใจ และความยาก เมื่อประสบเข้าก็ทำให้ยากแก่ปัญหานั้น

บอร์น เอกสแตรนด์ และโดมินอสกี (Borne, Ekstrand and Dominonski 1971: 9) ได้ให้แนวคิดเกี่ยวกับการแก้ปัญหาไว้ว่า การแก้ปัญหาคือเป็นกิจกรรมที่เป็นทั้งการแสดงความรู้ ความคิด จากประสบการณ์ก่อน ๆ และสถานการณ์ที่เป็นปัญหาในปัจจุบัน โดยนำมาเรียงลำดับใหม่ เพื่อผลของความสำเร็จในจุดมุ่งหมายเฉพาะอย่าง

กู๊ด (Good 1973: 518) ได้ให้ทรรศนะว่า การแก้ปัญหาคือเป็นแบบแผนหรือวิธีการ ซึ่งอยู่ในสภาวะที่มีความยากลำบาก ยุ่งยาก หรืออยู่ในสภาวะที่พยายามตรวจสอบข้อมูลที่หามาได้ ซึ่งมีความเกี่ยวข้องกับปัญหามีการตั้งสมมุติฐาน และมีการตรวจสอบสมมุติฐานภายใต้การควบคุม มีการเก็บรวบรวมข้อมูลเพื่อหาความสัมพันธ์เพื่อจะทดสอบสมมุติฐานนั้นว่าเป็นจริงหรือไม่

จากแนวคิดดังกล่าวข้างต้นจะพอจะสรุปได้ว่า ปัญหาคือสถานการณ์ที่บุคคลไม่สามารถผ่านอุปสรรคหรือความยุ่งยากไปสู่เป้าหมายได้ โดยการใช้วิธีการที่เคยเรียนรู้มาก่อน แต่จะต้องใช้กระบวนการที่เหมาะสมที่จะแก้ปัญหา นั้น ส่วนการแก้ปัญหาเป็นพฤติกรรมการเรียนรู้ เป็นกระบวนการที่ยุ่งยากซับซ้อน ต้องอาศัยความรู้ ความคิด ประสบการณ์ที่ผ่านมา เพื่อให้ได้ความรู้หรือแนวคิดใหม่ ๆ ซึ่งเป็นจุดมุ่งหมายที่ต้องการ

การแก้ปัญหาเป็นพฤติกรรมพื้นฐานของมนุษย์ ทุกขณะที่มนุษย์มีสติสัมปชัญญะ จะต้องเกี่ยวข้องกับปัญหา ทั้งนี้เพราะว่าการที่มนุษย์รู้สึกตัวอยู่นั้น สมองจะคิดอยู่ตลอดเวลา และการคิดที่ต้องการให้บรรลุเป้าหมายได้นั้น มนุษย์จะต้องมีการแก้ปัญหา (Polya 1957: 221) และ การที่บุคคลจะประสบความสำเร็จในชีวิตประจำวันได้นั้นก็จะต้องเป็นบุคคลที่มีความสามารถในการแก้ปัญหาเป็นพื้นฐาน (Tanner 1975: 25) ซึ่งธอร์นไดค์ (Thorndike 1950: 192-216) ได้กล่าวถึงสถานการณ์ที่จะเป็นปัญหาให้บุคคลแก้กันประกอบด้วย 3 ส่วน คือ

1. มีการกระตุ้นให้เกิดการศึกษาลู่ทางที่จะไปสู่เป้าหมาย
2. กระบวนการที่จะไปสู่เป้าหมายนั้นมีอุปสรรค
3. วิธีการหรือกระบวนการตอบสนองที่เคยชินของแต่ละบุคคลไม่เพียงพอที่จะผ่านอุปสรรคไปได้

บอร์น เอกสเตรน และโคมิโนสกี (Bourne, Ekstrand and Dominowoki 1971: 9) ได้แสดงความเห็นว่าปัญหาจะเกิดขึ้นก็ต่อเมื่อ

1. บุคคลพยายามให้ถึงจุดมุ่งหมายบางอย่าง หรือ เปลี่ยนสถานการณ์ของตนเองในปัจจุบันไปสู่สถานการณ์ที่ต่างออกไป
2. ความพยายามของบุคคลในขั้นแรกไม่สามารถที่จะทำให้สำเร็จได้ถึงที่สุด
3. มีทางเลือกในการกระทำมากกว่า 2 ทางขึ้นไป

นอกจากนั้น บรุคเนอร์ (Bruckner 1957: 567) ได้กล่าวไว้ว่าสิ่งที่จะเป็นปัญหาได้นั้นจะประกอบด้วย 4 ส่วน คือ

1. เป้าหมายที่ต้องการได้รับ
2. อุปสรรคที่จะไปสู่เป้าหมาย
3. การตอบสนองโดยวิธีที่เคยใช้ตามปกติไม่เพียงพอที่จะทำให้บรรลุเป้าหมายได้

4. มีการตั้งสมมุติฐาน หรือคำตอบหลาย ๆ วิธี

ดังนั้นจะเห็นได้ว่าปัญหาประกอบด้วยคำถามที่ต้องการคำตอบ แต่คำถามบางคำถาม อาจไม่เป็นปัญหาก็ได้ ทั้งนี้ขึ้นอยู่กับบุคคลที่จะตอบคำถามนั้น เช่น คำถาม "ขณะนี้เวลาเท่าไร" จะเป็นปัญหาสำหรับเด็กอายุ 5 ขวบ เพราะเขาไม่สามารถตอบคำถามนี้ได้ทันที แต่สำหรับผู้โตเป็นผู้ใหญ่มาก่อนแล้ว คำถามนี้จะไม่เป็นปัญหาสำหรับเขา สิ่งที่จะบอกได้ว่าคำถามใดเป็นปัญหาหรือไม่ก็คือ ผู้ถูกถามจะไม่สามารถตอบคำถามได้อย่างอัตโนมัติ (Cooney 1975: 241)

จากที่กล่าวมาแล้วข้างต้นอาจสรุปได้ว่า ปัญหาจะเกิดขึ้นโดยที่ผู้ถูกถามไม่สามารถตอบคำถามได้อย่างอัตโนมัติทันทีทันใด ทั้งนี้เนื่องมาจากกระบวนการหรือวิธีการที่เคยใช้ตามปกติ มาก่อนไม่เพียงพอที่จะตอบสนองปัญหาหรืออุปสรรคนั้นได้ แต่ถ้าหากผู้ถูกถามได้ศึกษารูปร่าง วิธีการแก้ไขอุปสรรค เพื่อให้บรรลุเข้าไปสู่เป้าหมายได้แล้ว ปัญหานั้นก็จะไม่มีอีกต่อไป

กระบวนการแก้ปัญหา

กระบวนการแก้ปัญหาคืออาศัยการประยุกต์ความรู้ในรูปของการวิเคราะห์ให้เข้ากับ ลักษณะของปัญหา เป็นการเลือกใช้วิธีการแก้ปัญหาที่ถูกต้อง และเหมาะสมกับปัญหาในสถานการณ์จริง ๆ นั้น ทั้งนี้จะต้องมีความเข้าใจถึงสิ่งที่เกี่ยวข้องกับปัญหาเข้าใจถึงส่วนประกอบของปัญหา เสียก่อน จึงจะสามารถนำสิ่งที่เกี่ยวข้องกับปัญหามาใช้อธิบายความหมายของปัญหาได้ (Crow 1956: 313, Wilson 1969: 240)

เกเออร์ (Gaier 1953: 138-141) กล่าวว่า ในการแก้ปัญหานั้นจำเป็นต้อง อาศัยข้อเท็จจริง แต่ข้อเท็จจริงเพียงอย่างเดียวไม่เป็นการเพียงพอต่อการแก้ปัญหา จำเป็น ต้องรู้จักสังเกต พิจารณาคัดเลือกหาแนวทางที่เป็นประโยชน์ต่อการแก้ปัญหา การสอนที่บอก แนวทางและข้อเท็จจริงในการแก้ปัญหาให้ผู้เรียนนั้นไม่สามารถช่วยให้ผู้เรียนแก้ปัญหาได้ ควร ฝึกฝนให้รู้จักสังเกต ค้นหาทางออกในการแก้ปัญหด้วยตนเอง ในเรื่องนี้ ฮิลเดรท (Hildreth 1966: 76) มีความเห็นว่าเด็กที่สามารถสรุปเรื่องราวได้อย่างสมเหตุสมผล สามารถแก้ปัญหา ได้โดยใช้ประสบการณ์และหลักเกณฑ์มาช่วยจะสามารถแก้ปัญหาในระดับสูงขึ้นไปได้อย่างรวดเร็ว และถูกต้องยิ่งขึ้น

สวนา พรพัฒน์กุล (2522: 271-272) ได้กล่าวถึงกระบวนการในการแก้ปัญหาว่า การแก้ปัญหาเป็นเรื่องสำคัญมาก มนุษย์ทุกคนเคยเผชิญกับสถานการณ์ที่เป็นปัญหามาแล้ว และจะต้องพบกับปัญหาต่าง ๆ ต่อไปอีกมากมาย ปัญหาบางประการไม่สลบซับซ้อนมากนักสามารถแก้ได้ง่าย แต่บางปัญหาก็สลบซับซ้อนมากยากแก่การที่จะแก้ไขปัญหานั้นได้สำเร็จ กระบวนการคิดเป็นองค์ประกอบที่สำคัญอย่างยิ่งต่อการแก้ปัญหา ยิ่งปัญหาสลบซับซ้อนมากก็ต้องอาศัยความคิดมากขึ้น เพราะกระบวนการดังกล่าวจำเป็นต้องอาศัยความรู้ในการพิจารณา สังเกตปรากฏการณ์ และโครงสร้างของปัญหา จึงจะสามารถทำให้บรรลุถึงจุดมุ่งหมายที่ต้องการได้

นอกจากนี้ยังมีนักจิตวิทยาบางท่านได้เน้นถึงกระบวนการแก้ปัญหามนุษย์ว่าเป็นการลองผิด-ลองถูก เช่น วูดเวอร์ค และชโลสเบิร์ก (Wood Werth and Schlooberg 1954: 818) ได้ให้ข้อสังเกตว่าในสถานการณ์ใด ๆ ที่เป็นปัญหาย่อมจะก่อให้เกิดการลองผิดลองถูกทั้งสิ้น จะต่างกันก็เพียงแต่จำนวนมากหรือน้อย และสติปัญญาสูงหรือต่ำของผู้แก้ปัญหาเท่านั้น ธอร์นไดค์ (Thorndike 1961: 40) เป็นบุคคลแรกที่ศึกษาเกี่ยวกับพฤติกรรมการลองผิดลองถูกในกระบวนการแก้ปัญหา โดยใช้แมวกับกล่องปริศนา (Puzzle box) เป็นเครื่องมือในการทดลอง แต่อย่างไรก็ตามในสมัยต่อมา มีนักจิตวิทยาหลายท่านได้ตั้งข้อสังเกตว่า สภาพการณ์ทดลองของธอร์นไดค์ เป็นการกีดกันไม่ให้เห็นพฤติกรรมอื่น ๆ ได้ นอกจากการลองผิดลองถูกเพียงอย่างเดียว

ออสูเบล (Ausubel 1968: 538) กล่าวว่า ในกระบวนการแก้ปัญหาโดยทั่วไปนั้น ต้องใช้องค์ประกอบหลายอย่างเช่น สติปัญญา และองค์ประกอบทางกระบวนการคิด เช่น ความยืดหยุ่นทางความคิด ความตั้งใจ และแบบของการคิด เป็นต้น ซึ่ง กูด (Good 1973: 518) ได้ให้ความเห็นในเรื่องนี้ว่า วิธีการทางวิทยาศาสตร์ก็คือการแก้ปัญหานั้นเอง กล่าวคือ การแก้ปัญหาเป็นแบบแผนหรือวิธีดำเนินการ ซึ่งอยู่ในสภาวะที่มีความสามารถยุ่งยากหรืออยู่ในสภาวะที่พยายามตรวจสอบข้อมูลที่หามาได้ที่มีความเกี่ยวข้องกับปัญหา มีการตั้งสมมุติฐาน ตรวจสอบสมมุติฐานภายใต้การควบคุมมีการเก็บรวบรวมข้อมูลจากการทดลอง เพื่อทดสอบสมมุติฐานว่าเป็นจริงหรือไม่

บรุนเนอร์ (Bruner 1966: 123-127) ได้สรุปขั้นตอนในการคิดแก้ปัญหาไว้ดังนี้คือ

1. รู้จักปัญหา (Problem Isolation) เป็นขั้นที่บุคคลรับรู้สิ่งเร้าที่ตนกำลังเผชิญอยู่ว่าเป็นปัญหา
2. ขั้นแสวงหาเค้าเงื่อน (Search of Cues) เป็นขั้นที่ใช้ความพยายามอย่างมากในการระลึกถึงประสบการณ์เดิม
3. ขั้นตรวจสอบความถูกต้อง (Confirmation Check) ก่อนที่จะตอบสนองในลักษณะของการจัดประเภท การแยกโครงสร้างและเนื้อหา
4. ขั้นตัดสินใจตอบสนองที่สอดคล้องเหมาะสมกับปัญหา

การแก้ปัญหาทางคณิตศาสตร์

บลูม (Bloom 1971: 86) ได้กล่าวถึงความมุ่งหมายของวิชาคณิตศาสตร์ว่า ในด้านการพัฒนาสติปัญญา (Cognitive Domain) นั้นเป็นการพัฒนาทางความรู้ (knowledge) และความสามารถทางสติปัญญา (Intellectual Ability) และทักษะ จึงได้แบ่งโครงสร้างของแบบสอบวิชาคณิตศาสตร์เป็น 5 ด้าน คือ

1. ความรู้เกี่ยวกับความคิดรวบยอด (Knowledge of Concept) เช่น กระบวนการทางคณิตศาสตร์ ความจริง และหลักการ
2. ทักษะในการคำนวณ (Computational Skill)
3. ทักษะในการแก้ปัญหา (Problem solving Skill)
4. ความรู้เกี่ยวกับโครงสร้างทางคณิตศาสตร์ และการพิสูจน์ที่มีเหตุผล
5. ทักษะในการสื่อความหมายในทางคณิตศาสตร์

ซาลิวสกี (Zalewski 1978: 2804-A) ได้ศึกษาองค์ประกอบในการแก้ปัญหาทางคณิตศาสตร์พบว่า สิ่งที่เป็นองค์ประกอบมีดังนี้

1. ความสามารถในการเข้าใจสัญลักษณ์
2. ความสามารถในการจัดกระทำ
3. ความสามารถในการอ่านและตีความ
4. ความคิดรวบยอดทางคณิตศาสตร์ (Mathematic Concepts)
5. ทักษะในการคำนวณ (Computation skill)

เกี่ยวกับขั้นตอนการแก้ปัญหาทางคณิตศาสตร์ ทบวงมหาวิทยาลัย (2524: 147) ได้เสนอแนะขั้นตอนในการแก้ปัญหาไว้ดังนี้

1. ทำความเข้าใจกับปัญหา โดยพิจารณาแยกแยะระหว่างสิ่งที่กำหนดให้และสิ่งที่ต้องการหา
2. จากสิ่งที่กำหนดให้แปลงให้อยู่ในรูปที่เห็นชัดขึ้น โดยใช้แผนภาพหรือรูปภาพประกอบ
3. จากแผนภาพที่สร้างขึ้นพิจารณาหาเงื่อนไขต่าง ๆ เพิ่มเติมโดยอาศัยนิยาม คุณสมบัติ ทฤษฎีต่าง ๆ ที่เคยเรียนรูมาแล้ว
4. ในบรรดาสิ่งที่กำหนดให้กับเงื่อนไขที่ได้เพิ่มเติม มีความเกี่ยวข้องกันอย่างไร หรือมีเงื่อนไขอะไรเพิ่มเติมอีกหรือไม่
5. คิดหาวิธีการแก้ปัญหา โดยพิจารณาถึงปัญหาที่คล้ายคลึงกันและทดลองเรียนแบบ การแก้ปัญหาหรืออาจวิเคราะห์จากสิ่งที่โจทย์ต้องการ พิจารณาว่าต้องมีเงื่อนไขใดอีกบ้างจึงจะได้สิ่งที่ต้องการ และเงื่อนไขนั้นพบได้อย่างไร
6. เรียบเรียงลำดับขั้นตอนในการแก้ปัญหา
7. ทดสอบคำตอบว่าสมเหตุสมผลหรือไม่
8. ถ้าการแก้ปัญหาทำได้หลายวิธี เลือกวิธีที่สั้นและง่ายที่สุด

โรเบิร์ต จอห์นสัน (Robert Johnson) ออกจากทบวงมหาวิทยาลัย (2524: 141) กล่าวไว้ว่า การแก้ปัญหาคงใช้กระบวนการทางสมองที่ซับซ้อน ต้องอาศัยความรู้พื้นฐาน มโนคติ และทักษะที่เกี่ยวข้องกับปัญหานั้น ๆ การที่นักเรียนมีความสามารถในการแก้ปัญหานั้น นักเรียนควรได้รับการฝึกฝนให้มีความรู้พื้นฐานดังต่อไปนี้

1. ทักษะการคิดคำนวณ เป็นความสามารถในการคิดคำนวณได้อย่างคล่องแคล่ว แม่นยำ หรือแทนสัญลักษณ์ต่าง ๆ ได้อย่างรวดเร็วถูกต้องในเวลาที่กำหนดให้ เช่น การบวก ลบ คูณ หาร และการแก้สมการอย่างง่าย ๆ ไม่ซับซ้อน และการแทนค่าตัวแปรต่าง ๆ
2. การใช้เหตุผลทางคณิตศาสตร์ เป็นความสามารถในการตีความ ขยายความ สรุปผล เปรียบเทียบ บอกขั้นตอนก่อนหลัง เช่น สามารถบอกว่าคุณลำดับใดสอดคล้องกับสมการที่กำหนดให้

๑. การแก้ปัญหามาทางคณิตศาสตร์ เป็นความสามารถในการแก้โจทย์ปัญหาต่าง ๆ ที่เคยพบเห็น และไม่เคยพบเห็นมาก่อน โดยอาศัยการคิดคำนวณ การใช้เหตุผลต่าง ๆ ประมวลเข้าด้วยกันในการแก้ปัญหานั้น ๆ ตลอดจนความสามารถในการติดตามการพิสูจน์

สำหรับการสร้างข้อสอบวิชาคณิตศาสตร์ ทฤษฎีของเทอร์สโตน (Thurstone 1947: 121) ได้กล่าวว่า ความสามารถด้านตัวเลขหรือคณิตศาสตร์นั้น คือสมรรถภาพในการที่จะเข้าใจ ความสามารถในการคำนวณตัวเลข โดยใช้พื้นฐานเบื้องต้น ผู้ที่มีความชำนาญคล่องแคล่ว แม่นยำมาก ก็เป็นผู้ที่มีความสามารถในด้านนั้นสูง ลักษณะข้อสอบประเภทนี้มีโครงสร้างดังนี้

1. ข้อสอบประเภททักษะ เป็นการวัดความคล่องแคล่วในการคำนวณ
2. ข้อสอบคณิตศาสตร์เหตุผล เป็นการวัดสมรรถภาพด้านความคิดรวบยอดทางคณิตศาสตร์เป็นการย่นหรือขยายกฎเกณฑ์ หลักการ วิธีการ และมโนภาพในความคิดแบบนามธรรม นั่นคือ เมื่อเด็กได้เรียนรู้กฎเกณฑ์ต่าง ๆ แล้วจะสามารถขยายความหมายนั้นออกไปเป็นรูปภาพ หรือสัญลักษณ์ให้กว้างขวางจากเดิมได้หรือไม่
3. ข้อสอบประเภทโจทย์ปัญหา เป็นการวัดสมรรถภาพการแก้ปัญหา แบบทดสอบประเภทนี้ประกอบด้วยคำถามโจทย์ปัญหาทำนองเดียวกับโจทย์คณิตศาสตร์ทั่วไป แต่ลักษณะของคำถามมุ่งที่จะวัดความสามารถในการแปลความ ให้ทำความเข้าใจของตัวเลข จำนวน และ การใช้เทคนิคหรือวิธีการแก้ปัญหาค้าง ๆ นอกจากนี้เป็นความสามารถในการจำสูตร และ กฎเกณฑ์ต่าง ๆ แล้วต้องเอาตัวเลขจากโจทย์ไปแทนค่าในสูตร

ชวาล แพร์คกุล (2516: 46-47) ได้ชี้ให้เห็นว่า โครงสร้างของแบบสอบคณิตศาสตร์นิยมแบ่งออกเป็น 3 ฉบับคือ แบบสอบสำหรับวัดทักษะหรือความคล่องแคล่วแม่นยำในการคำนวณ ตัวเลขแบบสอบสำหรับวัดความสามารถในการแก้โจทย์ปัญหา และแบบสอบสำหรับวัดความคิดรวบยอดหรือที่เรียกว่าแบบสอบวัดเหตุผลเชิงคณิตศาสตร์ ซึ่งมีรายละเอียดดังนี้

1. แบบสอบประเภททักษะ แบบสอบชนิดนี้ต้องการวัดความคล่องแคล่วแม่นยำในการบวก ลบ คูณ หาร ตัวเลข ซึ่งเป็นสมรรถภาพพื้นฐานในการเรียนคณิตศาสตร์ เทียบได้กับความสามารถในการสะกดคำของวิชาภาษาไทย ฉะนั้นโครงสร้างของแบบสอบชนิดนี้จึงไม่ต้องการให้มีอิทธิพลของภาษาเข้ามาเกี่ยวข้องด้วยเลย จัดว่าเป็นการวัดความคล่องแคล่วในการคำนวณตัวเลขโดยเฉพา

2. แบบสอบประเภทโจทย์ปัญหา แบบสอบชนิดนี้ประกอบด้วยคำถามประเภทโจทย์ต่าง ๆ ทำนองเดียวกับโจทย์คณิตศาสตร์ทั่วไป แต่ลีลาคำถามมุ่งวัดความสามารถในการแปลความหมาย และให้หาความสัมพันธ์ของตัวเลข จำนวน กับวิธีการแก้ปัญหาคำถามต่าง ๆ เป็นสำคัญ คำถามประเภทโจทย์ปัญหานี้มิได้หมายความว่าคำถามที่ให้นักเรียนนำตัวเลขไปแทนค่าในสูตรโดยตรง เช่น ในการหาพื้นที่ของวงกลม เพราะคำถามลักษณะนี้ไม่เป็นปัญหา แต่เป็นความสามารถในการจำสูตร และเอาตัวเลขไปแทนในสูตรเท่านั้น

3. แบบสอบประเภทเหตุผล แบบสอบชนิดนี้ต้องการวัดสมรรถภาพด้านความคิดรวบยอดเชิงคณิตศาสตร์ ซึ่งเกี่ยวข้องกับหลักการ วิธีการ และมโนภาพในความคิดแบบนามธรรม นั่นคือ ต้องการวัดว่าเมื่อเด็กได้เรียนรู้กฎเกณฑ์ต่าง ๆ แล้ว เขาจะสามารถขยายความนั้น ออกเป็นในรูปภาษา หรือสัญลักษณ์ให้กว้างขวางจากเดิมได้หรือไม่ แบบสอบประเภทนี้จะช่วยให้ครู อาจารย์ ได้ทราบว่าหลังจากเรียนเนื้อหาต่าง ๆ แล้ว นักเรียนเกิดความคิดในเรื่องความสัมพันธ์ของปริมาณเหล่านั้นหรือไม่ หรือเพียงแต่จำสูตร วิธีการตามตำราได้เท่านั้น ไม่จินตนาการในกฎเกณฑ์เหล่านั้นเลย

จากที่กล่าวมาทั้งหมดนี้จะเห็นได้ว่า องค์ประกอบของวิชาคณิตศาสตร์ และแบบสอบคณิตศาสตร์มักใช้ทักษะ 3 ประการคือ ทักษะการคำนวณ ความคิดรวบยอด และการแก้ปัญหาทางคณิตศาสตร์ เลอ บล็องค์ (Le Blane 1977: 16) ได้ให้ความเห็นเพิ่มเติมว่า ในบรรดาองค์ประกอบทั้ง 3 นั้น การแก้ปัญหาทางคณิตศาสตร์นับได้ว่าเป้าหมายสูงสุดของการเรียนการสอนวิชาคณิตศาสตร์ และเป้าหมายสูงสุดของการสอนแก้ปัญหานั้นมุ่งที่กระบวนการในการแก้ปัญหา ซึ่งตรงกับความคิดของ เวล (Whirl 1973: 551) ที่ได้กล่าวถึง การแก้ปัญหาทางคณิตศาสตร์ว่า วิธีการหรือเทคนิคในการแก้ปัญหาเป็นจุดที่สำคัญมากกว่าการได้คำตอบของโจทย์ปัญหา ส่วน กรูลิค (Krulik 1977: 550) กล่าวว่า การแก้ปัญหเป็นส่วนหนึ่งของการเรียนการสอนวิชาคณิตศาสตร์ ซึ่งครูจะต้องสร้างสถานการณ์ เพื่อให้ผู้เรียนได้เผชิญกับปัญหา เพื่อให้เขาได้ค้นหาให้ได้มาซึ่งคำตอบ บางครั้งอาจไม่ใช่คำตอบตามที่เขาหวังไว้ก็ได้ แต่นั่นไม่ใช่สิ่งที่สำคัญเท่ากับการที่ผู้เรียนได้ฝึกทักษะ ค้นหาวิธีการ ซึ่งวิธีการค้นหาคำตอบของปัญหาเป็นสิ่งสำคัญมากกว่าคำตอบในการแก้ปัญหา จะเห็นได้ว่า กระบวนการแก้ปัญหาเป็นหัวใจที่สำคัญของวิชาคณิตศาสตร์ ความจริงแล้วนักวิชาการและครูผู้สอนคณิตศาสตร์ส่วนใหญ่มองรับกันว่า คณิตศาสตร์โดยเนื้อแท้แล้วคือการแก้ปัญหานั้นเอง

ในลักษณะปัญหาคณิตศาสตร์นั้น บรูคเนอร์ (Bruckner 1957: 301) กล่าวถึงปัญหาทางคณิตศาสตร์ว่าเป็นสถานการณ์ที่เกี่ยวข้องกับปริมาณ ที่นักเรียนไม่สามารถตอบได้ทันที โดยวิธีการที่เคยชิน และสิ่งที่เป็นปัญหาของนักเรียนเมื่อวานนี้อาจจะไม่เป็นปัญหาในวันนี้ก็ได้ และ อัดัม (Adams 1977: 176) ยังได้กล่าวว่า ปัญหาคณิตศาสตร์เป็นสถานการณ์ที่เกี่ยวข้องกับปริมาณ และคำตอบที่ต้องการก็จะเกี่ยวข้องกับปริมาณ ปัญหาคณิตศาสตร์จะรวมถึงปัญหาที่เป็นภาษา (Word Problem) ปัญหาที่เป็นเรื่องราว และปัญหาที่เป็นคำพูด (Verbal Problem) นอกจากนี้เขายังกล่าวถึงความแตกต่างระหว่างปัญหากับแบบฝึกหัดว่า ในการแก้ปัญหาคงจะมีการตัดสินใจและลงมือทำ ส่วนแบบฝึกหัดไม่จำเป็นต้องมีการตัดสินใจ และในการสร้างสถานการณ์ของครูนั้น ไคลด์ (Clyde 1967: 108) ได้เสนอแนะถึงการสร้างปัญหาคณิตศาสตร์ที่น่าสนใจว่าควรมีลักษณะดังนี้

1. ให้มีความใกล้เคียงกับปัญหาในชีวิตจริง และมีความสัมพันธ์กับผู้แก้ปัญหามากที่สุด โดยอาจเป็นเรื่องราวหรือเหตุการณ์ ที่เกิดขึ้นกับผู้แก้ปัญหาในชีวิตประจำวัน หรือมักจะเกิดขึ้นกับบุคคลทั่ว ๆ ไป หรือมีลักษณะคล้ายกับสถานการณ์ในชีวิตจริง
2. สถานการณ์ที่สร้างขึ้นเป็นปัญหานั้น ควรใช้ภาษาหรือบรรยายในลักษณะที่ผู้แก้ปัญหามีประสบการณ์มาก่อน และไม่ควรเป็นปัญหาธรรมดาทั่ว ๆ ไป

นอกจากนั้น เวสต์ (West 1977: 57-58) ได้กล่าวไว้ว่า นักเรียนที่สามารถบอกคำตอบของโจทย์ปัญหาทางคณิตศาสตร์ได้ถูกต้อง อาจจะไม่สามารถบอกกระบวนการคิดทางคณิตศาสตร์ที่ถูกต้องก็ได้ กล่าวคือ นักเรียนที่คิดด้วยวิธีการที่ผิด อาจจะได้คำตอบที่ถูกต้องก็ได้ และนอกจากนี้เขายังชี้ให้เห็นถึงสาเหตุที่นักเรียนไม่สามารถจะทำข้อสอบที่เป็นปัญหาโจทย์คณิตศาสตร์ได้ถูกต้องไว้ 3 ประการ คือ

1. นักเรียนไม่เข้าใจข้อความที่เป็นโจทย์ปัญหา
2. นักเรียนไม่สามารถเปลี่ยนโจทย์ปัญหาเป็นประโยคสัญลักษณ์ทางคณิตศาสตร์ได้
3. นักเรียนไม่สามารถคำนวณตามที่โจทย์ต้องการได้

เมื่อเราไม่สามารถทราบว่าการที่นักเรียนทำข้อสอบหรือแก้ปัญหาคณิตศาสตร์ข้อนั้นผิดเนื่องมาจากสาเหตุใด ทั้งนี้เนื่องจากคำตอบผิดของนักเรียนเกิดมาจากความคลาดเคลื่อน

3 ประการ (Forsyth and Spratt 1980: 31) คือ

1. ความคลาดเคลื่อนเนื่องจากความเข้าใจในการอ่าน
2. ความคลาดเคลื่อนเนื่องมาจากการใช้กระบวนการ หรือวิธีการทางคณิตศาสตร์
3. ความคลาดเคลื่อนเนื่องมาจากการคำนวณ

นอกจากนี้ยังมีผู้รวบรวมขั้นตอนในการแก้ปัญหาทางคณิตศาสตร์ไว้มากมาย ซึ่งโดยมากแล้วจะอ้างถึงขั้นตอนในการแก้ปัญหาของ ยอร์จ โพล่า (George Polya 1957: 5-40) ซึ่งได้เสนอไว้ 4 ขั้นตอน ดังนี้

1. ทำความเข้าใจปัญหา สิ่งแรกที่จะต้องทำความเข้าใจคือ สัญลักษณ์ต่าง ๆ ในปัญหา ในขั้นนี้นักเรียนจะต้องสามารถสรุปปัญหาออกมาให้ได้ว่า ประเด็นปัญหาอยู่ที่ไหน สามารถบอกได้ว่าโจทย์ถามอะไร อะไรเป็นสิ่งที่ปัญหากำหนดให้ และอะไรคือเงื่อนไข

2. ขั้ววางแผนในการแก้ปัญหา ในขั้นนี้นักเรียนจะต้องมองเห็นความสัมพันธ์ของข้อมูลต่าง ๆ ในปัญหาอย่างชัดเจนเสียก่อน สิ่งที่ต้องการหาความสัมพันธ์กับข้อมูลให้มาอย่างไร การที่จะเกิดแนวคิดในการหาคำตอบเราต้องมีการวางแผนก่อน เทคนิคอย่างหนึ่งที่จะช่วยได้คือ พยายามที่จะนึกถึงปัญหาที่เคยแก้มาก่อน ซึ่งมีลักษณะที่ใกล้เคียงกัน

3. ขั้วลงมือทำตามแผน ขั้นนี้เป็นขั้นที่นักเรียนลงมือทำการคิดคำนวณตามแผนที่วางเอาไว้ในขั้นที่ 2 เพื่อที่จะให้ได้คำตอบของปัญหา สิ่งที่จะต้องใช้ในขั้นนี้ก็คื ทักษะในการคำนวณ การรู้จักเลือกวิธีคำนวณที่เหมาะสมมาใช้

4. ขั้วตรวจวิธีการและคำตอบ ขั้นนี้เป็นขั้นตรวจสอบ เพื่อความแน่ใจว่าถูกต้องสมบูรณ์ โดยพิจารณาและสำรวจดูผล ตลอดจนกระบวนการในการแก้ปัญหานักเรียนจะต้องรวบรวมความรู้ และพัฒนาความสามารถในการแก้ปัญหาเข้าด้วยกัน เพื่อทำความเข้าใจ และปรับปรุงคำตอบให้ดีขึ้น

ตอนที่ 2 มโนทัศน์เกี่ยวกับคุณภาพของแบบสอบ

เนื่องจากผลการสอบมีความสำคัญยิ่งต่อผู้สอบ จึงมีความจำเป็นที่ผู้สร้างแบบสอบจะต้องทราบถึงคุณลักษณะของแบบสอบที่ดี อันจะเป็นแนวทางในการสร้างแบบสอบและใช้เป็นเกณฑ์ในการวินิจฉัยว่า แบบสอบที่สร้างขึ้นนั้นมีคุณภาพดีหรือไม่อย่างไร เพื่อจะได้ปรับปรุงแก้ไข

ให้เป็นแบบสอบที่มีประสิทธิภาพยิ่งขึ้น

พวงแก้ว บุญยกนก (2521: อ้างถึงใน นิรมล บุญตะรัตน์ 2525: 15) ได้กล่าวถึงคุณลักษณะที่ดีของแบบสอบที่สามารถใช้เป็นเกณฑ์เพื่อวินิจฉัยคุณค่าของแบบสอบมี 10 ประการ คือ

1. ความตรง (Validity) คือ แบบสอบจะต้องมีความแม่นยำของคะแนนในการวัดสิ่งที่ต้องการวัด
2. ความเที่ยง (Reliability) คือ ความคงเส้นคงวาของคะแนน แบบสอบที่ตีนั้นใช้สอบซ้ำกับผู้สอบคนเดียวกัน ควรจะได้คะแนนเท่ากัน
3. ความเป็นปรนัย (Objectivity) คือ แบบสอบควรจะปราศจากความคลุมเครือของคำถาม ปราศจากอคติในการให้คะแนน
4. ความสมดุลย์ (Balance) คือ แบบสอบจะมีความตรงตามเนื้อหาได้จะต้องวัดพฤติกรรมที่เกี่ยวข้องกับเนื้อเรื่องที่สอบได้ครอบคลุมเนื้อหา และวัตถุประสงค์ สัดส่วนของคำถามเป็นไปตามรายการในตารางเฉพาะ
5. ความยุติธรรม (Fairness) คือ ผู้สอบทุกคนมีโอกาสที่จะตอบถูก ถ้ามีความรู้ในเนื้อเรื่องที่ออกข้อสอบนั้น ๆ
6. ประสิทธิภาพในการนำไปใช้ (Efficiency) คือ แบบสอบจะต้องประหยัดเวลาในการสร้าง, เวลาในการดำเนินการสอบ เวลาในการตรวจข้อสอบ ตลอดจนอุปกรณ์ต่าง ๆ ในการสร้างข้อสอบ
7. ความยาก (Difficulty) ที่เหมาะสม ซึ่งความยากที่เหมาะสมคือ คะแนนตรงกลางระหว่างคะแนนสูงสุดที่ควรจะได้กับคะแนนที่ทำได้โดยโอกาสการเดา
8. อำนาจจำแนก (Discrimination Power) คือ แบบสอบต้องจำแนกคนเก่งออกจากคนอ่อนได้
9. ความเฉพาะเจาะจง (Specific) คือ ผู้ที่มีความรู้ความสามารถเฉพาะเรื่องนั้นจริง ๆ จึงจะตอบข้อสอบนั้นได้ ซึ่งวิธีหาความเฉพาะเจาะจงสามารถทำได้โดยให้ผู้มีความสามารถในเรื่องนั้น ๆ กลุ่มหนึ่งทำแบบสอบนั้น เปรียบเทียบกับคนฉลาดที่ไม่มีความรู้ใน

เรื่องนั้น ถ้ากลุ่มแรกทำได้ดีกว่ากลุ่มหลังถือว่าแบบสอบนั้นมีความเฉพาะเจาะจง

10. ให้ความพอเหมาะ (Speededness) คือแบบสอบที่ไม่มีความจำกัดในเรื่องเวลา แม้ว่าจะต่อเวลาในการสอบนานเท่าใด คะแนนของผู้สอบก็ยังไม่เพิ่มขึ้นอยู่นั่นเอง ดังนั้นแบบสอบวัดผลสัมฤทธิ์ควรให้ความพอเหมาะพอเพียงที่เด็กเกินกว่า 90 % ของจำนวนทั้งหมดทำแบบสอบนั้นเสร็จ

วิเชียร เกตุสิงห์ (2518: 24-29) ได้สรุปถึงลักษณะที่ดีที่สำคัญของแบบสอบไว้ 10 ประการ คือ

1. ความเที่ยงตรง (Validity) หมายถึง คุณลักษณะของข้อสอบที่สามารถวัดได้ตรงตามจุดมุ่งหมายที่ตั้งไว้ ซึ่งแยกเป็น 4 ชนิด คือ

ก. ความเที่ยงตรงตามเนื้อหา (Content Validity) คือข้อสอบที่มีคำถามสอดคล้องตามเนื้อหาในหลักสูตร

ข. ความเที่ยงตรงตามโครงสร้าง (Construct Validity) หมายถึง ลักษณะของข้อสอบที่วัดสมรรถภาพด้านต่าง ๆ ตรงตามที่ได้กำหนดไว้

ค. ความเที่ยงตรงตามสภาพ (Concurrent Validity) หมายถึง ลักษณะของข้อสอบที่วัดได้ตรงตามสภาพความเป็นจริงของเด็กในขณะนั้น

ง. ความเที่ยงตรงเชิงพยากรณ์ (Predictive Validity) หมายถึง ข้อสอบที่สามารถวัดแล้วทำนายได้ว่าเด็กคนใดจะเรียนวิชาอะไรได้ดีเพียงไรในอนาคต

2. ความเชื่อมั่น (Reliability) หมายถึง แบบสอบที่สามารถวัดได้แน่นอน ไม่เปลี่ยนแปลงไปเปลี่ยนมา การวัดครั้งแรกเป็นอย่างไร เมื่อวัดซ้ำอีกครั้งหรือหลาย ๆ ครั้งก็ตาม ผลการวัดก็ยังคงเหมือนเดิม

3. อำนาจจำแนก (Discrimination) หมายถึง ลักษณะของแบบสอบที่สามารถจะแยกหรือจำแนกเด็กเก่ง-อ่อนได้ นั่นคือ เมื่อทดสอบแล้วจะบอกได้ว่าใครเก่งอ่อนอย่างไร

4. ความเป็นปรนัย (Objectivity) ข้อสอบที่จะมีความเป็นปรนัยได้นั้น ต้องประกอบด้วยคุณสมบัติ 3 ประการ คือ

- ก. มีความชัดเจนในความหมายของคำถาม ทุกคนอ่านแล้วเข้าใจตรงกัน ว่าถามว่าอะไร ส่วนใครจะตอบถูกหรือไม่ขึ้นอยู่กับความรู้ หรือความสามารถของแต่ละคน
- ข. มีความคงที่ในการตรวจให้คะแนน จะให้ใครตรวจก็คงได้คะแนนเหมือนกัน
- ค. มีความแจ่มชัดในการแปลความหมายของคะแนน

5. ความมีประสิทธิภาพ (Efficiency) หมายถึง ลักษณะข้อสอบที่มีคุณสมบัติ ที่แสดงถึงการประหยัด มีราคาถูก ง่ายในการดำเนินการสอบ พิมพ์ชัดเจน อ่านง่าย มีเนื้อหา มากแต่ใช้เวลาสอบน้อย

6. วัตถุประสงค์ที่ชัดเจน (Searching) หมายถึง ลักษณะของข้อสอบที่ถามครอบคลุม พฤติกรรมหลาย ๆ ด้าน ไม่ใช่วัดแค่ความรู้ความจำอย่างเดียว

ทั้ง 6 ประการที่กล่าวมาแล้ว เป็นหัวใจสำคัญของลักษณะข้อสอบที่ดี

7. ข้อสอบที่ดีต้องยุติธรรม (Fair) ไม่เปิดโอกาสให้เด็กคนใดคนหนึ่งได้ เปรียบคนอื่น

8. ข้อสอบที่ดีต้องเฉพาะเจาะจง (Definite) ไม่หลายแง่หลายมุมข้อเดียว พยายามถามเพียงเรื่องเดียว

9. ต้องมีความยากง่ายพอเหมาะ (Difficulty) ไม่ยากหรือง่ายเกินไป

10. ต้องมีการกระตุ้นยูแห่ (Exemplary) มีการจัดเอาข้อสอบง่าย ๆ ไว้ใน ตอนแรก ๆ แล้วจึงค่อย ๆ ถามให้ยากขึ้นตามลำดับ เป็นการเร้าให้เด็กเกิดความพยายามที่ จะทำข้อสอบให้ได้ทั้งหมด

นอกจากนี้ บุญธรรม กิจปริคาบวิสุทธิ (2524: 52-53) ได้มีความคิดเห็นในเรื่อง ว่า แบบสอบที่ดีนั้นจะต้องมีคุณสมบัติ 6 ประการ คือ

1. มีความแม่นยำตรง (Validity) คือ เป็นข้อสอบที่วัดได้ตรงประเด็นที่ต้องการวัด
2. มีความเชื่อถือได้ (Reliability) คือ ข้อสอบนั้นเมื่อใช้สอบแล้วผลที่ได้ เหมือนกันทุก ๆ ครั้งที่ใช้สอบหรือไม่ หากเหมือนกันหรือได้ผลคล้าย ๆ กันมากก็มีความ เชื่อถือ ได้มาก
3. มีความยากหรือง่ายพอเหมาะ ข้อสอบที่ดีควรมีคนทำถูกและทำผิดจำนวนพอ ๆ กัน

4. มีอำนาจจำแนก ข้อสอบที่ดีต้องสามารถแยกแยะนักเรียนเก่งกับนักเรียนไม่เก่งออกจากกันได้ กล่าวคือ จำนวนนักเรียนที่ได้คะแนนสูง ๆ ต้องตอบข้อสอบนั้นถูกมากกว่าจำนวนนักเรียนที่ได้คะแนนน้อย ๆ ตอบข้อสอบข้อนั้นถูก

5. สามารถนำไปใช้ได้ดี (Usability) ข้อสอบที่ดีต้องสามารถนำไปใช้ได้ดี คือ คำเนินการสอบได้ง่าย สะดวก ไม่ยุ่งยาก ใช้เวลาในการทดสอบพอดี ไม่สั้นหรือยาวนานจนเกินไป ให้คะแนนได้ง่าย แปลผลได้ถูกและนำไปใช้ได้สะดวก

6. มีความเป็นปรนัย (Objectivity) ข้อสอบที่ดีต้องมีความเป็นปรนัยสูง หมายถึง

6.1 ความเป็นปรนัยในเรื่องของความถูกต้องทางวิชาการ คือผู้รู้และผู้เชี่ยวชาญสาขาวิชานั้นเห็นว่ายกข้อทั้งตัวคำถามและตัวคำตอบ

6.2 ความเป็นปรนัยในการให้คะแนน จะต้องมีเกณฑ์การตรวจให้คะแนนที่แน่นอน ไม่ขึ้นอยู่กับอารมณ์ของผู้ตรวจ

6.3 ความเป็นปรนัยในด้านภาษา คือนักเรียนที่สอบทุกคนอ่านข้อสอบนั้นต้องเข้าใจความหมายของเนื้อความอย่างแจ่มแจ้งชัดเจน และถูกต้องตรงกับวัตถุประสงค์ของผู้ออกข้อสอบ

ในการวิจัยครั้งนี้ผู้วิจัยได้เลือกคุณลักษณะที่ดีของแบบสอบมาพิจารณาและ เปรียบ เทียบ คุณภาพรวม 4 ด้านคือ ค่าความยาก ค่าอำนาจจำแนก ค่าความเที่ยง และค่าความตรง ของแบบสอบ เนื่องจากคุณลักษณะดังกล่าว เป็นคุณลักษณะที่สำคัญของแบบสอบ และสามารถใช้เทคนิคการวิเคราะห์ทางสถิติได้ ดังจะกล่าวรายละเอียดต่อไปนี้

ความยากของแบบสอบ

ความยากของข้อสอบ หมายถึง สัดส่วนของนักเรียนที่ตอบข้อสอบข้อนั้นถูก

จากคำนิยามของค่าความยากจะเห็นได้ว่า ค่าความยากสามารถคำนวณได้จากจำนวนนักเรียนที่ตอบข้อสอบนั้นถูกหารด้วยจำนวนนักเรียนที่ตอบข้อสอบข้อนั้นทั้งหมด ดังนั้นค่าความยากของข้อสอบจึงมีค่ามากที่สุดเท่ากับ 1 และน้อยที่สุดเท่ากับ 0 ถ้าค่าความยากสูงหรือมีค่าใกล้ 1 แสดงว่า ข้อสอบนั้นง่ายมีนักเรียนจำนวนมากที่สามารถตอบถูก แต่ถ้าระดับค่าความยากต่ำหรือ

มีค่าใกล้ 0 แสดงว่าข้อสอบนั้นยาก มีนักเรียนน้อยคนที่สามารถตอบถูก การพิจารณาระดับความยากที่พอเหมาะนั้นต้องทราบจุดมุ่งหมายของการวัด ในการวัดผลแบบอิงกลุ่ม (Norm Reference) นั้นค่าความยากของข้อสอบที่เหมาะสมคือ .50 ทั้งนี้เพราะทำให้การกระจายของคะแนนมากที่สุด และจำแนกความสามารถของนักเรียนได้มากที่สุด แต่การจะสร้างข้อสอบในแบบสอบให้มีความยากพอเหมาะทุกข้อนั้นทำได้ยาก ดังนั้นจึงให้ใช้ข้อสอบที่มีความยากอยู่ในช่วง .20 ถึง .80 ก็นับว่าพอใช้ได้ (ชวาล แพร์ตกุล 2509: 316) ส่วนข้อสอบที่มีความยากต่ำกว่า .20 หรือสูงกว่า .80 จะเป็นข้อสอบที่ยากหรือง่ายเกินไป ข้อสอบเหล่านี้จำแนกความสามารถของนักเรียนไม่ได้หรือจำแนกได้น้อย สำหรับการวัดผลแบบอิงเกณฑ์ (Criterion Reference) ส่วนใหญ่จะเป็นข้อสอบที่มีค่าความยากสูง ทำให้การกระจายของคะแนนมีลักษณะเบ้ซ้าย กล่าวคือ คะแนนของนักเรียนส่วนใหญ่จะสูง ซึ่งจะช่วยให้ทราบว่านักเรียนคนใดบ้างที่มีความรู้และทักษะไม่ถึงเกณฑ์ที่เราคาดหวัง (อนันต์ ศรีโสภณ 2519: 257)

องค์ประกอบที่มีอิทธิพลต่อค่าความยากของแบบสอบ

องค์ประกอบที่มีอิทธิพลต่อค่าความยากของแบบสอบนี้มี 2 ประเภท (Campbell 1961: 899-913) คือ

1. องค์ประกอบภายใน (Intrinsic Factors) ประกอบด้วย

ก. เนื้อหาของแบบสอบแต่ละข้อ (Item Content) ในด้านความซับซ้อน ความเป็นนามธรรมและความแปลกใหม่

ข. โครงสร้างของข้อสอบ (Item Structure) หมายถึง วิธีในการแสดงออกซึ่งเนื้อหานั้น

2. องค์ประกอบภายนอก (Extrinsic Factors) ประกอบด้วย

ก. ความไม่คุ้นเคยต่อเนื้อหา (Unfamiliarity) คือ อยู่นอกเหนือประสบการณ์ของผู้ตอบ

ข. สิ่งที่สัมพันธ์กับข้อสอบ (Item Context) เช่น ข้อสอบที่อยู่ใกล้เคียงกัน

ค. ตัวแปรด้านบุคลิกภาพ ได้แก่ สภาพร่างกาย ลักษณะนิสัย และความตั้งใจของผู้สอบ

นอกจากองค์ประกอบที่กล่าวแล้วยังมีองค์ประกอบที่อาจจะมีผลต่อค่าความยากของแบบสอบคือ ธรรมชาติของเนื้อหา ชนิดของพฤติกรรมที่ต้องการวัดจากผู้สอบ และตัวการอื่น ๆ ที่สลับซับซ้อนมองไม่เห็นเด่นชัดเช่น ลักษณะของภาษาที่ใช้ รูปแบบของคำถาม ตลอดจนคำชี้แจงและสิ่งอื่น ๆ ซึ่งสิ่งเหล่านี้ผู้สร้างแบบสอบควรจะสนใจนำเข้ามาพิจารณาในการสร้างแบบสอบด้วย

อำนาจจำแนกของแบบสอบ

อำนาจจำแนก หมายถึง ประสิทธิภาพของข้อคำถามในการจำแนกนักเรียนออกเป็น คนเก่ง คนอ่อน อำนาจจำแนกมีค่าตั้งแต่ -1.00 ถึง 1.00 ข้อสอบที่มีค่าอำนาจจำแนกเข้าใกล้ -1.00 หรือ 1.00 แสดงว่ามีอำนาจจำแนกสูง แต่ข้อสอบที่มีคุณภาพตรงตามวัตถุประสงค์ของการสอบคือ ข้อสอบที่มีค่าอำนาจจำแนกเป็นบวก ซึ่งหมายความว่านักเรียนในกลุ่มสูงตอบได้ถูกต้องมากกว่านักเรียนในกลุ่มต่ำ และแสดงว่าข้อสอบสามารถจำแนกไปในทางเดียวกับคะแนนรวมที่ได้จากแบบสอบ ส่วนข้อสอบที่มีค่าอำนาจจำแนกเป็นทางลบถือว่าจำแนกผิดทิศทาง ไม่ตรงตามวัตถุประสงค์ที่ต้องการ และข้อสอบที่มีค่าอำนาจจำแนกเป็น 0 ถือว่าจำแนกความแตกต่างระหว่างคน 2 กลุ่มไม่ได้

ในการประเมินคุณภาพของข้อสอบโดยอาศัยค่าอำนาจจำแนก ควรถือเกณฑ์ดังนี้ (Ebel R.L. 1965: 364)

ค่าอำนาจจำแนก	คุณภาพของข้อสอบ
.40 ขึ้นไป	ดีมาก
.30 ถึง .39	ดี
.20 ถึง .29	พอใช้ได้ แต่ต้องปรับปรุง
ต่ำกว่า .19	ใช้ไม่ได้ ต้องพิจารณาปรับปรุงใหม่หรือทิ้งไปเลย

ในการประเมินคุณภาพของข้อสอบ ควรพิจารณาอำนาจจำแนกควบคู่ไปกับค่าความยากด้วย เนื่องจากข้อสอบที่มีค่าความยากเท่า ๆ กัน อาจมีความแตกต่างกันในด้าน

ค่าอำนาจจำแนกด้วย ดังที่กล่าวมาแล้วว่าข้อสอบที่มีค่าความยาก .50 อาจมีค่าอำนาจจำแนกสูงสุด หรือไม่มีเลยก็ได้ ฉะนั้นข้อสอบที่มีค่าความยาก .50 มีโอกาสที่จะทำให้ความเที่ยงของแบบสอบค่าได้ เมื่อข้อสอบนั้นมีค่าอำนาจจำแนกเป็น 0 แต่ข้อสอบที่มีค่าอำนาจจำแนกเป็น 0 แต่ข้อสอบที่มีค่าอำนาจจำแนกสูง จะมีผลทำให้ความเที่ยงของแบบสอบสูงด้วยอย่างแน่นอน เนื่องจากอำนาจจำแนกเกี่ยวข้องกับส่วนเบี่ยงเบนมาตรฐานของคะแนน ซึ่งเป็นองค์ประกอบสำคัญที่มีอิทธิพลต่อค่าความเที่ยงของแบบสอบ (Ebel 1965: 302)

องค์ประกอบที่มีอิทธิพลต่อค่าอำนาจจำแนกของแบบสอบ

องค์ประกอบที่สำคัญที่มีอิทธิพลต่อค่าอำนาจจำแนกของแบบสอบ ได้แก่

1. คุณภาพของตัวกลาง และจำนวนตัวเลือก ข้อสอบที่สร้างตัวลองได้มีความใกล้เคียงกับตัวเลือกที่เป็นคำตอบถูกหลาย ๆ ตัวเลือก ก็จะมีตัวเลือกให้เด็กมีโอกาสเลือกมากขึ้น จึงทำให้อำนาจจำแนกของข้อสอบสูงขึ้นไปด้วย (Ebel 1965: 364) เนื่องจากผู้ที่มีความรู้จริงเท่านั้นจึงจะสามารถทำข้อสอบได้ถูกต้อง แต่นักเรียนที่ไม่มีความรู้จริงจะเลือกตัวลองที่มีลักษณะใกล้เคียงกับคำตอบถูก กระจายกันออกไปทุกตัวเลือก
2. ความยากของข้อสอบ การจำแนกผลสัมฤทธิ์ระหว่างกลุ่มนักเรียนขึ้นอยู่กับความหมายของข้อสอบ (Graonlund 1976: 14) กล่าวคือ ข้อสอบที่มีค่าความยาก .50 มีโอกาสทำให้อำนาจจำแนกสูงสุดได้ เมื่อนักเรียนในกลุ่มสูงทุกคนตอบข้อสอบได้ถูกต้อง ในขณะที่นักเรียนในกลุ่มต่ำทั้งหมดตอบข้อสอบผิด และหากข้อสอบมีค่าความยากสูง หรือต่ำกว่านี้ ค่าอำนาจจำแนกของข้อสอบจะต่ำลง ซึ่งหมายความว่าข้อสอบที่ง่าย หรือยากเกินไปจะเป็นข้อสอบที่มีค่าอำนาจจำแนกต่ำนั่นเอง
3. ชนิดของแบบสอบ ลอร์ด (Lord อ้างใน อนันต์ ศรีโรภา 2524: 163) ได้ให้ข้อสังเกตในการกำหนดค่าความยากที่จะทำให้ค่าอำนาจจำแนกมากที่สุดตามชนิดของแบบสอบ ดังนี้

ชนิดของแบบสอบ	ค่าความยากที่ทำให้ค่าอำนาจ จำแนกสูงที่สุด
แบบ เติมคำและคำตอบสั้น	50
แบบ เลือกตอบชนิด 5 ตัวเลือก	70
แบบ เลือกตอบชนิด 4 ตัวเลือก	74
แบบ เลือกตอบชนิด 3 ตัวเลือก	77
แบบ ถูก-ผิด	85

ความเที่ยงของแบบสอบ

ได้มีผู้ให้ความหมายของความเที่ยงไว้มากมาย เช่น

ความเที่ยง หมายถึง ความคงที่แน่นอนของคะแนนที่ได้จากการสอบนักเรียนกลุ่มเดียวกันด้วยแบบสอบฉบับเดียวกันหลาย ๆ ครั้ง หรือด้วยแบบสอบสองฉบับที่มีลักษณะเทียบเท่ากัน หรือภายใต้สภาพการสอบที่ต่างกัน (อนาสตาซี 2519: 73)

ความเที่ยง หมายถึง ระดับ (Degree) ของความคงที่ระหว่างการวัดสิ่งเดียวกัน 2 ครั้ง (Mehrens and Lehmann 1978: 88)

ความเที่ยง หมายถึง ความคงเส้นคงวาของการวัด (Gronlund 1968: 105)

ความเที่ยง หมายถึง ความคงเส้นคงวาของชุดคะแนนจากแบบสอบที่วัดในสิ่งที่ต้องการจะวัด

จากความหมายของความเที่ยงดังกล่าวข้างต้น พอจะสรุปได้ว่า ความเที่ยงหมายถึงความคงที่ของคะแนนที่ได้จากการวัดด้วยเครื่องมือชุดเดียวกันในโอกาสต่างกัน

ทฤษฎีของความเที่ยง (Theory of Reliability)

ทฤษฎีความเที่ยงของคะแนนอธิบายได้ดังนี้ คะแนนที่วัดได้ (Observed Scores) จะประกอบด้วยคะแนน 2 ส่วนคือ คะแนนที่ได้จากสมรรถภาพที่แท้จริงเรียกว่าคะแนนจริง (True Score) และคะแนนความคลาดเคลื่อน (Error Score) ซึ่งเขียนเป็นสมการได้

ดังนี้ (Mehrens and Lehmann 1978: 90)

$$X = T + E$$

เมื่อ X แทน คะแนนที่วัดได้

T แทน คะแนนจริง

E แทน คะแนนความคลาดเคลื่อน

คะแนนจริง (True Score) หมายถึง คะแนนที่ผู้สอบทำได้จากการวัดด้วย เครื่องมือที่มีคุณลักษณะสมบูรณ์ปราศจากความคลาดเคลื่อน หรือหมายถึง คะแนนเฉลี่ยของผู้สอบที่ได้จากการทำแบบสอบเดิม หรือแบบสอบคู่ขนานหลาย ๆ ครั้ง โดยมีข้อตกลงว่า ไม่มีอิทธิพลจากการฝึกฝน ความเมื่อยล้า และการเรียนรู้ในการทดสอบซ้ำ (Brown 1976: 50)

คะแนนความคลาดเคลื่อน (Error Score) หมายถึง คะแนนที่เกิดจากความผิดพลาดในการวัด ซึ่งเป็นความคลาดเคลื่อนที่เกิดขึ้นในลักษณะสุ่ม (Random Error) คะแนนความคลาดเคลื่อนนี้อาจเกิดขึ้นจากสาเหตุหลายอย่างเช่น เกิดจากตัวแบบสอบ การบริหารการสอบ การให้คะแนน หรือเกิดจากตัวผู้เข้าสอบเอง เป็นต้น โดยทฤษฎีความคลาดเคลื่อนที่เกิดขึ้นในลักษณะสุ่มนี้อาจเป็นไปได้ทั้งทางบวกและทางลบซึ่งจะหักล้างกันหมดไป และถ้าความคลาดเคลื่อนเกิดขึ้นในลักษณะสุ่มแล้วก็จะไม่มีความสัมพันธ์กับคะแนนจริง นั่นคือ สัมประสิทธิ์สหสัมพันธ์ (Correlation) ระหว่างคะแนนจริงกับความคลาดเคลื่อนมีค่าเท่ากับ 0 ดังนั้นความแปรปรวนของคะแนนที่วัดได้สามารถเขียนเป็นสมการได้ดังนี้ (Mehrens and Lehmann 1978: 91)

$$S_x^2 = S_t^2 + S_e^2 \quad \text{----- (1)}$$

เมื่อ S_x^2 แทน ความแปรปรวนของคะแนนที่วัดได้

S_t^2 แทน ความแปรปรวนของคะแนนจริง

S_e^2 แทน ความแปรปรวนของคะแนนความคลาดเคลื่อน

โดยทฤษฎีความเที่ยง หมายถึง อัตราส่วนระหว่างความแปรปรวนของคะแนนจริง กับความแปรปรวนของคะแนนที่วัดได้ (Brown 1976: 51) จึงสามารถเขียนสมการได้เป็น

$$\gamma_{tt} = \frac{S_t^2}{S_x^2} \text{-----} (2)$$

เมื่อ γ_{tt} แทน ค่าความเที่ยง

จากสมการ (1) และ (2) จะได้สมการพื้นฐานในการหาสัมประสิทธิ์สหสัมพันธ์ของความเที่ยง ดังนี้

$$\gamma_{tt} = 1 - \frac{S_e^2}{S_x^2}$$

จะเห็นว่า ค่าความเที่ยงจะสูงขึ้นเมื่อความแปรปรวนคลาดเคลื่อนลดลง โดยกำหนดให้คะแนนผลการสอบมีค่าคงที่หรือเพิ่มขึ้น และค่าความเที่ยงจะมีค่าสูงสุดเท่ากับ 1

วิธีการประมาณค่าความเที่ยง

ในการประมาณค่าความเที่ยงสามารถทำได้หลายวิธี และในแต่ละวิธีก็มีข้อจำกัดที่แตกต่างกัน การที่จะเลือกใช้วิธีใดนั้นขึ้นอยู่กับ ความมุ่งหมาย ลักษณะของคะแนน และ ชนิดของแบบสอบ (Guilford 1978: 414) โดยทั่วไปวิธีการประมาณค่าความเที่ยงมี 4 แบบ (Mehrens and Lehmann 1978: 84) คือ

1. แบบสอบซ้ำ (Measures of Stability)
2. แบบใช้ข้อสอบคล้ายกัน (Measures of equivalence)
3. แบบใช้ข้อสอบคล้ายกันและสอบซ้ำ (Measures of equivalence and stability)
4. แบบวัดความคงที่ภายใน (Measures of internal consistency)

ซึ่งวิธีการนี้แบ่งออกเป็น

- 4.1 วิธีแบ่งครึ่งข้อสอบ (Split-half)
- 4.2 วิธีของคูเดอร์-ริชาร์ดสัน (Kuder-Richardson estimates)
- 4.3 วิธีสัมประสิทธิ์แอลฟา (Coefficient alpha)
- 4.4 วิธีวิเคราะห์ความแปรปรวนของฮอยท์ (Hoyt's analysis of variance procedure)

ในที่นี้จะกล่าวถึงเฉพาะวิธีการประมาณค่าความเที่ยงแบบคูเคอร์-ริชาร์ดสัน ซึ่งเป็นวิธีการประมาณค่าความเที่ยงในการวิจัยครั้งนี้

วิธีการประมาณค่าความเที่ยงแบบคูเคอร์-ริชาร์ดสัน

การประมาณค่าความเที่ยงแบบคูเคอร์-ริชาร์ดสัน เป็นการประมาณค่าความเที่ยงที่คูเคอร์-ริชาร์ดสัน (Kuder and Richardson) ได้เสนอไว้ในปี ค.ศ. 1934 ซึ่งเป็นที่รู้จักและนิยมใช้กันอย่างแพร่หลายคือ สูตรคูเคอร์-ริชาร์ดสัน 20 (Kuder-Richardson 20) (Brown 1976: 78)

$$Y_{tt} = \left(\frac{K}{K-1} \right) \left(\frac{S_t^2 - \sum p_i q_i}{S_t^2} \right)$$

เมื่อ	Y_{tt}	แทน	ค่าความเที่ยงของแบบสอบ
	K	แทน	จำนวนข้อในแบบสอบ
	p_i	แทน	สัดส่วนของผู้ตอบถูกในแต่ละข้อ
	q_i	แทน	สัดส่วนของผู้ตอบผิดในแต่ละข้อ
	S_t^2	แทน	ความแปรปรวนของคะแนนทั้งหมด

การประมาณค่าความเที่ยงแบบคูเคอร์-ริชาร์ดสัน เป็นการหาความสัมพันธ์กันระหว่างข้อสอบซึ่งดำเนินการสอบเพียงครั้งเดียว และใช้แบบสอบชุดเดียว ความสอดคล้องกันระหว่างข้อนี้ได้รับอิทธิพลจากแหล่งความแปรปรวนคลาดเคลื่อน 2 แหล่ง (อนาสตาซี 2519: 84) คือ

1. เนื้อหาที่สับสน
2. ความเป็นวิวิธพันธ์ของพฤติกรรมที่สับสน ยิ่งข้อสอบมีความเป็นเอกพันธ์มาก ความสอดคล้องกันระหว่างข้อยิ่งสูง (Stanley 1978: 126)

ข้อตกลงเบื้องต้นของวิธีการประมาณค่าความเที่ยงแบบคูเคอร์-ริชาร์ดสัน-20

1. การตรวจให้คะแนนในแต่ละข้อเป็น 0-1 คือ เมื่อตอบถูกให้ 1 คะแนน เมื่อตอบผิดให้ 0 คะแนน (Ebel 1965: 326)

2. ข้อสอบแต่ละข้อ ในแบบสอบจะต้องมีลักษณะเป็น เอกพันธ์ คือ วัตถุประสงค์ เดียวกัน (Ebel 1965: 318)

สูตรคูเคอร์-ริชาร์ดสัน 20 นี้ เหมาะที่จะใช้หาความเที่ยงของแบบสอบ ความสามารถ (Power test) เท่านั้นไม่เหมาะที่จะหาความเที่ยงของแบบสอบความเร็ว (Speed test) เพราะค่า p และ q ของแต่ละข้อต้องเป็นค่าที่ได้จากการที่ผู้สอบทุกคน มีโอกาสทำข้อสอบข้อนั้น เสร็จแล้ว ซึ่งเป็นไปไม่ได้สำหรับแบบสอบความเร็ว เพราะผู้สอบ ทุกคนไม่มีโอกาสทำถึงข้อสุดท้าย (Brown 1976: 326)

องค์ประกอบที่มีอิทธิพลต่อค่าความเที่ยง

นอกจากวิธีการประมาณค่าความเที่ยงของแบบสอบจะมีอิทธิพลต่อค่าความเที่ยงของแบบสอบแล้ว ยังมีองค์ประกอบอื่น ๆ ที่มีอิทธิพลต่อค่าความเที่ยงของแบบสอบด้วย (Mehrens and Lehmann 1973: 100-103) คือ

1. ความยาวของแบบสอบ (Test Length) แบบสอบที่มีจำนวนข้อสอบมากจะมีความเที่ยงสูงกว่าแบบสอบที่มีจำนวนข้อสอบน้อยกว่า เนื่องจากแบบสอบที่มีจำนวนข้อสอบมากจะประกอบด้วยกลุ่มของข้อสอบที่วัดพฤติกรรมที่ต้องการอย่างเพียงพอ คะแนนที่ได้จากอิทธิพลการเตลลดน้อยลง และเมื่อเพิ่มจำนวนข้อสอบที่มีคุณภาพดีเท่าเทียมกับข้อสอบเดิม เช่น ความยากเท่ากัน วัดพฤติกรรมเดียวกัน ผลการวัดจะคงที่ยิ่งขึ้น คะแนนที่ได้จะเป็นตัวชี้ความสามารถจริงได้มากขึ้น สูตรที่แสดงให้เห็นว่าเมื่อเพิ่มจำนวนข้อสอบในแบบสอบทำให้ค่าความเที่ยงเพิ่มขึ้นคือสูตรของสเปียร์แมน-บราวน์ (Spearman-Brown)

2. ความเร็ว (Speed) ในแบบสอบความเร็วข้อสอบจะง่าย แต่ให้เวลาน้อย จึงไม่มีใครสามารถทำได้หมดทุกข้อ ความแตกต่างของคะแนนจากการสอบจึงขึ้นอยู่กับจำนวนข้อที่ได้พยายามทำ ไม่ได้ขึ้นอยู่กับความยากของข้อสอบ การประมาณค่าความเที่ยงของแบบสอบด้วยวิธีวัดความคงที่ภายในจึงไม่เหมาะสม เพราะจะทำให้ได้ค่าความเที่ยงของแบบสอบสูงกว่าที่ควรจะเป็น

3. ความยากของข้อสอบ (Item Difficulty) เนื่องจากค่าความเที่ยงของแบบสอบนั้นประมาณจากความแปรปรวนของคะแนนผลการสอบ ถ้าข้อสอบยากเกินไปทุกคนทำผิดหมด หรือถ้าข้อสอบง่ายเกินไปจะทำให้ทุกคนทำถูกหมด ซึ่งจะ ทำให้พิสัยของคะแนนลดลง

ความแปรปรวนของคะแนนก็จะมีน้อย ซึ่งจะมีผลทำให้ความเที่ยงลดลงด้วย ดังนั้นถ้าค่าความยากของข้อสอบทำให้คะแนนกระจายมาก ๆ ก็จะทำให้ค่าความเที่ยงของแบบสอบมีค่าสูงขึ้นด้วย

4. ความเป็นเอกพันธ์ของกลุ่ม (Group Homogeneity) กลุ่มที่มีความเป็นเอกพันธ์จะมีพิสัยของคะแนนความสามารถของกลุ่มน้อยกว่ากลุ่มที่มีความเป็นวิริวิพันธ์ (Group Heterogeneous) ซึ่งจะมีผลทำให้ค่าความเที่ยงของแบบสอบจากกลุ่มที่เป็นเอกพันธ์ต่ำกว่าค่าความเที่ยงของกลุ่มที่เป็นวิริวิพันธ์

5. ความเป็นปรนัย (Objectivity) แบบสอบที่มีความเป็นปรนัยสูง เช่น ปราศจากอคติของผู้ตรวจให้คะแนน ภาษาที่ใช้ในแบบสอบถูกต้อง ทำให้ผู้สอบเข้าใจคำถาม และคำสั่งต่าง ๆ ในแบบสอบถูกต้องตรงกัน เป็นต้น ทำให้ผลการสอบตรงตามความสามารถในสิ่งที่ต้องการวัดที่ผู้สอบมีอยู่จริง การวัดนั้นก็มีความคงที่ทำให้ความเที่ยงของแบบสอบมีค่าสูง

นอกจากนี้ความคลาดเคลื่อนจากการประมาณค่าความเที่ยงอาจเกิดจากการบริหารการสอบ เช่น การจับเวลา การให้คำชี้แจง การป้องกันการทุจริต หรือองค์ประกอบที่เกี่ยวข้องกับตัวผู้สอบเอง เช่น แรงจูงใจ ความวิตกกังวล ประสบการณ์ในการสอบ การเดา เป็นต้น สิ่งเหล่านี้อาจทำให้ค่าความเที่ยงของแบบสอบเปลี่ยนแปลงได้

ความตรงของแบบสอบ

ได้มีผู้ให้ความหมายของความตรง (Validity) ของแบบสอบไว้หลายท่าน เช่น

ความตรง หมายถึง ความถูกต้องแน่นอนในการวัดสิ่งที่ต้องการวัด (Lindquist 1942: 213)

ความตรง หมายถึง คุณสมบัติของแบบสอบที่สามารถวัดในสิ่งที่ต้องการจะวัด (Ebel 1965: 273)

ความตรงของแบบสอบ หมายถึง ขอบเขตของผลการประเมินที่สอดคล้องกับจุดประสงค์นั้น (Gronlund 1976: 79)

ความตรงของแบบสอบ หมายถึง ความถูกต้องแม่นยำของคะแนนในการพยากรณ์ที่จำเพาะเจาะจง (Stanley 1978: 101)

ความตรงของแบบสอบ หมายถึง ค่าสหสัมพันธ์ระหว่างแบบสอบกับเกณฑ์
(Gullikson 1950: 88)

จากความหมายต่าง ๆ ดังกล่าวข้างต้น สามารถสรุปความหมายของความตรงได้
2 ประการ คือ

1. ความตรงของแบบสอบ หมายถึง ค่าสหสัมพันธ์ระหว่างคะแนนที่ได้จาก
แบบสอบกับคะแนนเกณฑ์
2. ความตรงของแบบสอบ หมายถึง ความสามารถของแบบสอบที่วัดได้ในสิ่งที่
ต้องการวัด หรือวัดได้สอดคล้องกับวัตถุประสงค์ และคะแนนที่ได้จากแบบสอบที่มีความตรงสูง
สามารถบอกถึงสภาพที่แท้จริง และพยากรณ์ได้ถูกต้องแม่นยำ

ชนิดของความตรง

การแบ่งประเภทของความตรงมีหลายวิธี ทั้งนี้ขึ้นอยู่กับจุดมุ่งหมายของการวัด โดย
ทั่วไปการวัดทางการศึกษาและจิตวิทยาแบ่งความตรงออกเป็น 3 ประเภท (Mehrens
and Lehmann 1973: 109) คือ

1. ความตรงตามเนื้อหา (Content Validity)
2. ความตรงตามเกณฑ์สัมพันธ์ (Criterion Related Validity)
3. ความตรงตามโครงสร้าง (Construct Validity)

การวิจัยครั้งนี้เกี่ยวข้องกับความตรงตามเนื้อหา และความตรงตามเกณฑ์สัมพันธ์
ซึ่งมีรายละเอียด ดังนี้

ความตรงตามเนื้อหา (Content Validity)

ความตรงตามเนื้อหาจะเกี่ยวข้องกับเนื้อหาสาระข้อกระทงในแบบสอบ ซึ่ง
ข้อกระทงนั้นต้องเป็นตัวแทนของเนื้อหาสาระ และครอบคลุมขอบเขตของพฤติกรรมที่ต้องการวัด
ความตรงตามเนื้อหามีความสำคัญอย่างยิ่งในแบบสอบวัดผลสัมฤทธิ์ทางการเรียน การที่จะตัดสิน
ว่าแบบสอบนี้มีความตรงตามเนื้อหามากน้อยเพียงใด จะต้องพิจารณาทั้งหัวข้อเนื้อหาวิชา และ
ชนิดของพฤติกรรมที่ต้องการวัดไปพร้อม ๆ กัน

อย่างไรก็ตาม เราไม่สามารถแสดงค่าความตรงตามเนื้อหาออกมาเป็นตัว เลขได้ โดยทั่วไปมักจะพิจารณาข้อกระทงในแบบสอบโดยเทียบกับตารางวิเคราะห์หลักสูตร (Table of Specification) แล้วประมวลออกมาว่ามีความตรงตาม เนื้อหาอย่างน้อยเพียงใด ผู้ที่จะ คัดสินความตรงตาม เนื้อหาได้ดีคือผู้เชี่ยวชาญทางสาขานั้น ๆ (Mahrens and Lehmann 1978: 111)

ความตรงตาม เกณฑ์สัมพันธ์ (Criterion-Related Validity)

ความตรงตาม เกณฑ์สัมพันธ์ของแบบสอบ แสดงถึงความสามารถของแบบสอบ ในการทำนายพฤติกรรมของบุคคลในสถานการณ์เฉพาะ หรือแสดงถึงความสัมพันธ์ระหว่าง คะแนนที่ผู้สอบกระทำได้จากแบบสอบกับเกณฑ์ภายนอก (External Criteria) ที่เป็นอิสระ โดยทั่วไปแล้วความตรงตาม เกณฑ์สัมพันธ์จะใช้เมื่อมีความต้องการที่จะประมวลจุดยืนที่น่าจะเป็น ไปได้ของผู้สอบบน เกณฑ์ที่อ้างอิงถึง (เยาวดี วิบูลย์ศรี 2526: 73)

ความตรงตาม เกณฑ์สัมพันธ์จำแนกได้ 2 ชนิด คือ

1. ความตรงตามสภาพ (Concurrent Validity) หมายถึง ความ ตรงของแบบสอบที่สามารถบอกให้เราทราบว่าสิ่งที่วัดถูกต้องตามสภาพที่เบ้จจริงในปัจจุบัน โดยการหาความสัมพันธ์ระหว่างคะแนนของแบบสอบกับคะแนน เกณฑ์ที่กำหนดขึ้นในเวลาเดียวกัน หรือ ใกล้เคียงกัน เช่น ต้องการทราบความตรงตามสภาพของแบบสอบวัดทักษะกระบวนการทาง วิทยาศาสตร์ กับตัวแปรเกณฑ์อื่น ๆ ก็อาจหาได้จากการหาความสัมพันธ์ระหว่างคะแนนที่ได้จาก แบบสอบ กับการสังเกตพฤติกรรมกระบวนการทางวิทยาศาสตร์ ในขณะที่นักเรียนทำการทดสอบ ในห้องเรียน ถ้าคะแนนที่นักเรียนทำได้จากแบบสอบได้ผลสอดคล้องกับความสามารถที่แสดงออก จริงซึ่งใช้เป็น เกณฑ์ แสดงว่าแบบสอบนั้นมีความตรงตามสภาพ

2. ความตรงเชิงทำนาย (Predictive Validity) หมายถึง ความ สามารถของแบบสอบที่จะบ่งบอกผลที่วัดในขณะนั้นได้ถูกต้องตามสภาพที่แท้จริงในอนาคต โดย อาศัยความสัมพันธ์ระหว่างคะแนนของแบบสอบกับ เกณฑ์สัมพันธ์ซึ่งจะปรากฏในอนาคต เช่น เกณฑ์สัมพันธ์ของแบบสอบความถนัดทางวิชาการเพื่อทำนายผลการเรียน ก็อาจใช้คะแนนเฉลี่ย สะสมปีสุดท้ายเป็น เกณฑ์สัมพันธ์ ซึ่งการคำนวณหาความตรง เชิงพยากรณ์จะต้องอาศัยเวลารอคอย เพราะคะแนนของแบบสอบกับ เกณฑ์สัมพันธ์ได้มาคนละ เวลา กัน ความตรงเชิงทำนายเหมาะที่จะ

ใช้กับแบบสอบถามถนัดชนิดต่าง ๆ และแบบสอบถามอื่นที่ควรระวังเพื่อการทำนาย

ข้อแตกต่างระหว่างความตรงเชิงสภาพ กับความตรงเชิงทำนายขึ้นอยู่กับวัตถุประสงค์ของการสอบ ความตรงตามสภาพ ใช้กับแบบสอบถามสำหรับวินิจฉัยสภาพที่เป็นอยู่ ส่วนความตรงเชิงทำนาย ใช้กับแบบสอบถามสำหรับพยากรณ์เหตุการณ์ในอนาคต (อนาสตาซี 2519: 105) นอกจากนี้ เมเร็นส์ และ เลแมนน์ (Mehrens and Lehmann 1973: 112) ได้กล่าวอีกว่าข้อแตกต่างระหว่างความตรง 2 ชนิดนี้คือ ช่วงระยะเวลาของการเก็บข้อมูลที่เป็นเกณฑ์ ถ้าเก็บรวบรวมข้อมูลเป็นเกณฑ์กับคะแนนการสอบในเวลาเดียวกัน หรือใกล้เคียงกัน จะเป็นความตรงตามสภาพ แต่ถ้าเก็บรวบรวมข้อมูลที่เป็นเกณฑ์ล่าช้าออกไป จะเป็นความตรงเชิงทำนาย

การวัด เกณฑ์

การที่จะได้ข้อมูลของเกณฑ์ที่เหมาะสมและถูกต้องนั้นทำได้ยาก การวัด เกณฑ์ก็เหมือนกับการวัดสิ่งอื่น ๆ คือผู้วัดจะต้องทราบคุณลักษณะของสิ่งที่จะเป็นเกณฑ์อย่างดีเสียก่อน การพิจารณาเกณฑ์นั้น เมเร็นส์ และ เลแมนน์ (Mehrens and Lehmann) เสนอว่าจะต้องพิจารณาในสิ่งต่อไปนี้

1. ความเกี่ยวข้อง (Relevance) การที่เราจะเลือกสิ่งใดมาเป็นเกณฑ์นั้น สิ่งนั้นก็ควรมีความเกี่ยวข้อง หรือสะท้อนถึงคุณลักษณะที่สำคัญของเรื่องที่เราจะศึกษา โดยทั่วไป การพิจารณาความเกี่ยวข้องของเกณฑ์นี้ขึ้นอยู่กับ การตัดสินใจคุณค่า ไม่สามารถที่จะวัดออกมาเป็นตัวเลขได้

2. ความเที่ยง (Reliability) เกณฑ์นั้นจะต้องมีความเที่ยงพอสมควร เนื่องจากความเที่ยงของแบบสอบถามมีอิทธิพลต่อความสัมพันธ์ระหว่างแบบสอบถามกับ เกณฑ์ ซึ่งทฤษฎีความสัมพันธ์กล่าวไว้ว่า ความสัมพันธ์สูงสุดระหว่างตัวแปร 2 ตัวจะเท่ากับหรือน้อยกว่ารากที่สองของผลคูณระหว่างความเที่ยงของตัวแปรทั้ง 2 นั้น นั่นคือ

$$r_{xy} \leq \sqrt{(r_{xx})(r_{yy})}$$

เมื่อ γ_{xy}	แทน	ค่าสหสัมพันธ์ระหว่างตัวทำนาย (x) กับตัวเกณฑ์ (y)
γ_{xx}	แทน	ค่าความเที่ยงของแบบสอบที่เป็นตัวทำนาย
γ_{yy}	แทน	ค่าความเที่ยงของเกณฑ์

จะเห็นได้ว่า ค่าสหสัมพันธ์ระหว่างแบบสอบกับเกณฑ์ (γ_{xy}) มีค่ามากหรือน้อยขึ้นอยู่กับความเที่ยงของแบบสอบเอง (γ_{xx}) และความเที่ยงของเกณฑ์ (γ_{yy})

๓. ความเป็นอิสระจากอคติ (Free form Bias) คือ เกณฑ์ที่ใช้ควรเป็นอิสระ และไม่มีอิทธิพลจากการสอบที่นำมาเป็นตัวทำนาย เช่น ครูนำแบบสอบความถนัดทางวิชาการ มาทำการทดสอบกับนักเรียนตอนต้นปี เพื่อพยากรณ์ผลการเรียน แต่ถ้าพบว่านักเรียนคนใดได้คะแนนความถนัดต่ำ ก็จะประเมินผลการเรียนให้นักเรียนผู้นั้นต่ำไปด้วย ซึ่งจะทำให้ตัวทำนายมีอิทธิพลต่อคะแนนเกณฑ์ อาจจะเป็นผลทำให้ค่าสหสัมพันธ์ระหว่างคะแนนสอบกับเกณฑ์สูงขึ้น อย่างไรก็ตามการป้องกันความคลาดเคลื่อนในลักษณะนี้อาจแก้ไขได้โดยอย่าให้ผู้มีหน้าที่ประเมินผลการเรียนทราบคะแนนที่ใช้เป็นตัวทำนาย

วิธีการประมาณค่าความตรงตาม เกณฑ์สัมพันธ์

การประมาณค่าความตรงตาม เกณฑ์สัมพันธ์ สามารถวิเคราะห์ได้จากสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนจากแบบสอบกับเกณฑ์ ดังนั้นสูตรที่ใช้จึงเป็นสูตรคำนวณสัมประสิทธิ์สหสัมพันธ์แบบต่าง ๆ โดยทั่วไปนิยมใช้สูตรสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน (Pearson's Product Moment Correlation Coefficient) (Guilford 1976: 140-141)

องค์ประกอบที่มีอิทธิพลต่อค่าความตรงของแบบสอบ

องค์ประกอบที่มีอิทธิพลต่อค่าความตรงของแบบสอบ มีดังนี้ (Gronlund 1976: 98-102) คือ

1. องค์ประกอบที่เกิดจากแบบสอบ มีดังต่อไปนี้

1.1 คำสั่งในการทำแบบสอบไม่ชัดเจน คำสั่งที่ไม่ชัดเจนทำให้ผู้สอบเข้าใจผิดได้ง่าย ซึ่งอาจเกิดจากไม่มีตัวอย่างประกอบ ใช้ภาษาไม่รัดกุม และพิมพ์ไม่ชัดเจน เป็นต้น

1.2 ใช้คำศัพท์และโครงสร้างทางภาษายากเกินไป อาจทำให้ผู้สอบไม่เข้าใจ ส่วนที่ไม่เกี่ยวกับสิ่งที่ต้องการสอบ แต่เสียเวลามากจนทำแบบสอบได้ไม่เสร็จทันเวลา

1.3 ข้อสอบมีความยากไม่เหมาะสม ในกรณีแบบสอบอิงกลุ่ม ข้อสอบที่ยากเกินไป หรือง่ายเกินไปทำให้แบบสอบมีค่าความตรงต่ำ และในกรณีของแบบสอบอิงเกณฑ์ ข้อสอบที่ง่ายมากเกินไปอาจไม่เหมาะสมกับจุดประสงค์ในการสอบ ดังนั้นจะทำให้แบบสอบมีความตรงต่ำด้วย

1.4 ข้อสอบไม่ดี ข้อสอบแบบ เลือกตอบถ้าสร้างไม่ดีตัวเลือกอาจแนะนำคำตอบได้ โดยผู้ออกข้อสอบไม่ได้ตั้งใจ ทำให้ข้อสอบวัดสิ่งที่ต้องการการวัดอย่างแท้จริงไม่ได้

1.5 ข้อสอบกำกวม ข้อสอบที่ใช้ภาษากำกวมอาจทำให้ผู้สอบที่เก่งทำไม่ได้ แต่ผู้สอบที่ไม่เก่งทำได้ และอาจทำให้ผู้สอบ เข้าใจความหมายของภาษามึนงงได้

1.6 ชนิดของแบบสอบที่ใช้ไม่เหมาะสม แบบสอบที่ไม่เหมาะสมอาจวัดสิ่งที่ต้องการวัดได้ไม่ดีเพียงพอ ทำให้ไม่อาจประเมินสิ่งที่ต้องการวัดได้อย่างถูกต้อง เช่น การสอบความสามารถในการเรียนรู้ควรใช้แบบสอบที่ให้ผู้สอบ เรียนมากกว่าจะให้เลือกคำตอบจากแบบสอบ เลือกตอบเท่านั้น เป็นต้น

1.7 แบบสอบสั้นเกินไป แบบสอบที่สั้นเกินไปอาจทำให้วัดสิ่งที่ต้องการวัดได้ไม่มาก และไม่ครอบคลุมเนื้อหา ทำให้การประเมินผลการสอบผิดพลาดได้ เพราะสิ่งที่วัดไม่เป็นตัวแทนของเนื้อหาที่จะวัด

1.8 การจัดเรียงข้อกระทงไม่เหมาะสม ข้อกระทงควรเรียงจากง่ายไปหายาก ถ้าเรียงจากยากไปหาง่าย จะทำให้นักเรียนต้องใช้เวลาในการทำแบบสอบนาน จนบางคนไม่มีเวลาทำข้อสอบข้อง่าย ๆ ที่อยู่ตอนท้ายของแบบสอบจึงวัดได้ไม่ตรงตามความสามารถที่แท้จริง

1.9 คำตอบที่ถูกเรียงกันอย่างมีระบบ ซึ่งจะทำให้นักเรียนเดาคำตอบได้

2. เนื้อหาและวิธีการดำเนินการสอบ ในการออกข้อสอบต้องออกให้ตรงกับเนื้อหาที่สอน นอกจากนี้ยังต้องคำนึงถึงระดับความรู้ที่ต้องการวัดด้วย เพราะถ้าต้องการวัดผลการเรียนรู้ในระดับความรู้ขั้นสูงแล้วถามในเนื้อหาที่เรียนมาแล้วจะไม่ใช่เป็นการวัดผลการเรียนรู้ในระดับความรู้ขั้นสูง จะเป็นเพียงการวัดในระดับความรู้ ความจำเท่านั้น

3. การบริหารการสอบและการให้คะแนน การบริหารการสอบที่ไม่มีมาตรฐาน เช่น การให้เวลาน้อยเกินไป การอธิบายข้อสอบให้นักเรียนบางคน การทุจริตของนักเรียน ตลอดจนการให้คะแนนที่ไม่คงที่จะมีผลทำให้ความตรงของแบบสอบลดลง

4. องค์ประกอบที่เกี่ยวกับตัวนักเรียนในการตอบแบบสอบ ในการสอบแต่ละครั้ง นักเรียนจะมีความพร้อมแตกต่างกัน ทั้งนี้เนื่องจากเกิดความวิตกกังวลเกิดการเปลี่ยนแปลงทางอารมณ์ การมีแรงจูงใจในการทำแบบสอบ การมีความสามารถในการเดา สิ่งเหล่านี้จะทำให้นักเรียนสามารถตอบแบบสอบได้แตกต่างจากความสามารถที่แท้จริงของนักเรียน

5. ธรรมชาติของกลุ่มและเกณฑ์ที่ใช้ โดยทั่วไปแบบสอบจะใช้ได้เหมาะสมกับผู้สอบกลุ่มใดกลุ่มหนึ่งโดยเฉพาะ ทั้งนี้เพราะบุคคลมีความแตกต่างกันในเรื่องอายุ เพศ ความสามารถ ภูมิหลังทางการศึกษา และวัฒนธรรม ซึ่งจะ使人ในแต่ละกลุ่มสามารถตอบแบบสอบได้แตกต่างกัน ดังนั้นการใช้แบบสอบควรคำนึงถึงลักษณะเฉพาะของกลุ่มที่ใช้หาความตรง นอกจากนั้นยังต้องคำนึงถึงเกณฑ์ที่ใช้หาความตรงของแบบสอบด้วย

ตอนที่ 3 งานวิจัยที่เกี่ยวข้อง

จากการศึกษางานวิจัยที่เกี่ยวข้องกับคุณภาพของแบบสอบ เลือกตอบที่มีรูปแบบตัวเลือกแตกต่างกัน ได้มีผู้สนใจทำการศึกษาไว้มากพอสมควรทั้งต่างประเทศและในประเทศ ซึ่งผลงานวิจัยดังกล่าวพอจะสรุปได้ดังนี้

งานวิจัยต่างประเทศ

เวสแมน และ เบนเนตต์ (Wesman and Benett 1946: 151-154)

ได้ศึกษาผลของการใช้ตัวเลือกที่ระบุว่า "ไม่มีคำตอบถูก" ในวิชาคณิตศาสตร์ และศัพท์ภาษาอังกฤษ และปรากฏว่าไม่พบความแตกต่างในด้านที่เกี่ยวกับระดับความยากของแบบสอบทั้งสองชนิด แต่ในข้อสอบบางข้อพบว่า ข้อสอบที่ระบุว่า "ไม่มีคำตอบถูก" มีประสิทธิภาพมากและผู้วิจัยทั้งสองยังได้สรุปเพิ่มเติมอีกว่า การใช้ตัวเลือก "ไม่มีคำตอบถูก" จะมีประสิทธิภาพมากน้อยเพียงใดนั้นขึ้นอยู่กับคุณภาพของตัวเลือกตัวอื่น ๆ ด้วย

บอยนตัน (Boynton 1950: 565-570) ได้ศึกษาถึงประสิทธิภาพของแบบสอบประเภทเลือกตอบที่ระบุว่า "ไม่มีคำตอบถูก" ในการสะกศค้ำพบว่าข้อสอบที่ใช้ตัวเลือก

“ไม่มีคำตอบถูก” มีระดับความยาก และอ่านาจจำแนกสูงกว่าแบบสอนที่ไม่ได้ใช้ตัวเลือกนี้

ริมแลนด์ (Rimland 1960: 533-539) ได้ศึกษาเกี่ยวกับคุณภาพของแบบสอนเลือกตอบที่ใช้ตัวเลือกแบบธรรมดา และตัวเลือกแบบ “ไม่มีคำตอบถูก” ในวิชาคณิตศาสตร์ โดยการศึกษาในกลุ่มตัวอย่าง 2 กลุ่ม ๆ ละ 3,600 คน โดยที่กลุ่มที่ 1 ทำแบบสอนที่มีรูปแบบตัวเลือกแบบธรรมดา ส่วนกลุ่มที่ 2 ทำแบบสอนที่มีรูปแบบตัวเลือกแบบ “ไม่มีคำตอบถูก” โดยให้ 1 ใน 5 ของจำนวนข้อสอบในแบบสอนที่มีตัวเลือกแบบ “ไม่มีคำตอบถูก” เป็นคำตอบถูก ผลการศึกษาในครั้งนี้ปรากฏว่าการใช้ตัวเลือกแบบ “ไม่มีคำตอบถูก” ไม่ทำให้คุณภาพของแบบสอนในด้านค่าความยาก ค่าอ่านาจจำแนก แตกต่างไปจากแบบสอนที่ใช้ตัวเลือกแบบธรรมดา

ฮิวส์ และทริมเบิล (Hughes and Trimple 1965: 117-125) ได้ทำการวิจัยเกี่ยวกับคุณภาพของแบบสอนเลือกตอบที่ใช้รูปแบบตัวเลือกแตกต่างกัน 3 รูปแบบ คือ รูปแบบตัวเลือกแบบธรรมดา แบบปลายเปิดและแบบปลายปิด และแบบผสม “ถูกทั้งข้อ 1 และข้อ 2” หรือ “ผิดทั้งข้อ 1 และข้อ 2” ในวิชาจิตวิทยาทั่วไปโดยศึกษากับนักศึกษาระดับปริญญาตรีปีที่ 1 จำนวน 63 คน โดยแบ่งออกเป็น 3 กลุ่ม ให้ทำแบบสอนแต่ละฉบับ โดยที่กลุ่มที่ 1 ทำแบบสอนที่ใช้ตัวเลือกแบบธรรมดา กลุ่มที่ 2 ทำแบบสอนที่ใช้ตัวเลือกแบบปลายเปิด และแบบปลายปิด และกลุ่มที่ 3 ทำแบบสอนที่ใช้ตัวเลือกแบบผสม โดยที่แบบสอนแต่ละชุดจะประกอบด้วยข้อสอบจำนวน 55 ข้อ แบ่งออกเป็น 3 ส่วน คือส่วนแรก ใช้เป็นข้อสอบควบคุมได้แก่ข้อสอบที่มีตัวเลือกแบบธรรมดาเหมือนกันทั้ง 3 ชุด จำนวน 15 ข้อ ส่วนที่ 2 เป็นข้อสอบที่ใช้เปรียบเทียบจริง ได้แก่ ข้อสอบที่ใช้ตัวลวงตัวสุดท้ายเป็นแบบธรรมดา หรือแบบปลายเปิด และแบบปลายปิด หรือแบบผสม ต่างกันในแต่ละฉบับ จำนวน 35 ข้อ และส่วนที่ 3 เป็นข้อสอบที่ใช้ตัวเลือกแบบต่าง ๆ เป็นคำตอบถูกจำนวน 5 ข้อ ส่วนนี้ไม่ได้นำมาเปรียบเทียบ แต่เพื่อให้เห็นความสำคัญของตัวเลือกซับซ้อนบ้างไม่ใช่เป็นตัวลวงอย่างเดียว ผลการวิจัยปรากฏว่า รูปแบบของตัวเลือกแบบต่าง ๆ มีผลต่อค่าความยาก แต่ไม่มีผลต่อค่าอ่านาจจำแนกของแบบสอน และเมื่อเปรียบเทียบคะแนนเฉลี่ยของนักเรียนปรากฏว่า แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 โดยคะแนนจากการสอบของนักเรียนที่ทำแบบสอนตัวเลือกแบบธรรมดาสูงที่สุด

มุลเลอร์ (Mueller 1975: 135-141) ได้ทำการศึกษาในลักษณะเดียวกันกับฮิวส์และทริมเบิล โดยใช้กลุ่มตัวอย่างที่เป็นนักศึกษาในมหาวิทยาลัยอินเดียนา จำนวน 2,000 คน ในวิชาเกี่ยวกับการประกอบอาชีพเป็นพนักงานขายของ ปรากฏว่า ผลจากการวิจัยที่ได้สอดคล้องกับการวิจัยของฮิวส์และทริมเบิลคือ พบว่ารูปแบบของตัวเลือกมีผลต่อค่าความยากของข้อสอบเช่นกัน โดยข้อสอบที่ใช้ตัวเลือกแบบผสมยากที่สุด รองลงมาคือ ข้อสอบที่ใช้ตัวเลือกแบบปลายเปิด และข้อสอบที่ใช้ตัวเลือกแบบธรรมดาที่ง่ายที่สุด แต่ไม่มีผลต่อค่าอำนาจจำแนกของแบบสอบ นอกจากนี้ยังได้เปรียบเทียบเมื่อตัวเลือกแบบต่าง ๆ เป็นคำตอบถูกด้วยพบว่า ข้อสอบที่ใช้ตัวเลือกแบบผสมยากที่สุด รองลงมาคือ ข้อสอบที่ใช้ตัวเลือกแบบปลายเปิด แต่ข้อสอบที่ง่ายที่สุดคือ ข้อสอบที่ใช้ตัวเลือกแบบปลายเปิด และตัวลวงที่มีประสิทธิภาพน้อยที่สุดคือตัวลวงแบบธรรมดา

ฟอร์ซิท และสเปรทท์ (Forsyth and Spratt 1980: 31-43) ได้ศึกษาถึงผลของการเลือกข้อคำถามและคุณลักษณะของแบบสอบ เลือกตอบที่ใช้วัดความสามารถในการแก้ปัญหาทางคณิตศาสตร์ โดยใช้แบบสอบใจหทัยปัญหาที่ปรับปรุงมาจากแบบสอบ ITBS (Iowa Test of Basic Skills) ของนักเรียนในระดับเกรด 7 และเกรด 8 จำนวน 20 ข้อ ที่มีตัวเลือกแตกต่างกัน 4 รูปแบบคือ รูปแบบตัวเลือกแบบธรรมดา รูปแบบตัวเลือกแบบธรรมดาที่ระบุว่า "ไม่มีคำตอบถูก" รูปแบบตัวเลือกแบบกระบวนการคิด และรูปแบบตัวเลือกแบบกระบวนการคิดที่ระบุว่า "ไม่มีคำตอบถูก" โดยใช้กลุ่มตัวอย่างนักเรียนเกรด 7 และเกรด 8 จากโรงเรียนมัธยมศึกษา 3 โรงเรียนจำนวน 988 คน ซึ่งแบ่งออกเป็น 4 กลุ่ม โดยในแต่ละกลุ่มทำแบบสอบเพียงฉบับเดียว ผลการวิจัยปรากฏว่า แบบสอบที่มีรูปแบบตัวเลือกแบบกระบวนการคิดยากกว่าแบบสอบที่มีรูปแบบตัวเลือกแบบธรรมดา และตัวเลือกที่ระบุว่า "ไม่มีคำตอบถูก" ทำให้แบบสอบยากขึ้น แบบสอบที่มีรูปแบบตัวเลือกแบบธรรมดามีค่าอำนาจจำแนกสูงกว่าแบบสอบที่มีรูปแบบตัวเลือกเป็นกระบวนการคิด และแบบสอบที่มีตัวเลือกระบุว่า "ไม่มีคำตอบถูก" ทำให้ความแตกต่างของค่าอำนาจจำแนกของแบบสอบที่มีรูปแบบตัวเลือกแบบธรรมดา และที่เป็นกระบวนการคิดลดลง นอกจากนี้ยังพบว่า แบบสอบที่มีตัวเลือกแบบธรรมดามีความตรงเชิงโครงสร้างสูงกว่าแบบสอบที่มีตัวเลือกเป็นกระบวนการคิด เมื่อใช้แบบสอบ TIBS เป็นเกณฑ์

ไวเทิน (Weiten 1982: 46-49) ได้ศึกษาเปรียบเทียบคุณภาพของแบบสอบเลือกตอบชนิดคำตอบเดียว (Single Multiple-choice) กับชนิดคำตอบซ้อน (Double

Multiple-choice) ในเนื้อหาวิชาจิตวิทยา โดยศึกษากับนักศึกษาปริญญาตรีจำนวน 47 คน โดยแบบสอบเลือกตอบแบบคำตอบเดียวก็คือ แบบสอบเลือกตอบตัวเลือกแบบธรรมดา แต่แบบสอบเลือกตอบแบบคำตอบซ้อนจะมีเงื่อนไขให้เลือก 5 ข้อ ซึ่งเหมือนและเท่ากับตัวเลือกของแบบคำตอบเดียว ตัวเลือกจะสร้างจากเงื่อนไข 1, 2, 3, 4 หรือ 5 เงื่อนไข รวมทั้งตัวเลือกถูกทั้งหมดหรือผิดทั้งหมดด้วย ในบางข้อผลการวิจัยปรากฏว่าแบบสอบชนิดคำตอบซ้อนยากซับซ้อน ทำทาย นำเท่ากับแบบคำตอบเดียว และแบบสอบทั้งสองลักษณะมีความแตกต่างกันในเรื่อง ความยาก อ่านาจนจำแนก ความเที่ยง แต่ไม่พบความแตกต่างกันในเรื่องของความตรงตามสภาพ

ออสเตอร์ฮอฟ และพามาเลา (Oosterhof and Pamela 1984: 287-293) ได้ศึกษาเปรียบเทียบคุณภาพของแบบสอบเลือกตอบที่มีตัวเลือกแตกต่างกัน 3 แบบ กับแบบสอบเติมคำในวิชาคณิตศาสตร์ที่เป็นโจทย์ปัญหา แบบสอบเลือกตอบที่มีรูปแบบตัวเลือกแตกต่างกัน 3 รูปแบบ คือ ตัวเลือกแบบธรรมดา ตัวเลือกแบบปลายเปิด "ไม่มีคำตอบถูก" และตัวเลือกเป็นแบบช่อง โดยแบบสอบแต่ละฉบับมีข้อสอบจำนวน 12 ข้อ โดยทำการศึกษากับกลุ่มตัวอย่างที่เป็นนักศึกษาธุรกิจการศึกษาจำนวน 232 คน ผลการวิจัยพบว่า ค่าความยากของแบบสอบเลือกตอบทั้ง 3 ฉบับไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ ส่วนแบบสอบเติมคำยากกว่าแบบสอบเลือกตอบทั้ง 3 ฉบับอย่างมีนัยสำคัญทั้ง 3 ฉบับ และผลจากการศึกษาครั้งนี้พบอีกว่า การใช้ตัวเลือกแบบปลายเปิด "ไม่มีคำตอบถูก" ทำให้แบบสอบยากกว่าแบบสอบที่มีตัวเลือกแบบธรรมดา และตัวเลือกแบบช่อง ถ้าใช้ตัวเลือกปลายเปิด "ไม่มีคำตอบถูก" เป็นตัวถูกทำให้ข้อสอบยากขึ้นเพื่อพิจารณาถึงความเที่ยงของแบบสอบทั้ง 4 ฉบับที่คำนวณโดยใช้สูตร KR-20 พบว่า แบบสอบเติมคำมีความเที่ยงสูงสุด ส่วนแบบสอบเลือกตอบทั้ง 3 ฉบับ มีค่าความเที่ยงใกล้เคียงกัน โดยแบบสอบที่มีตัวเลือกแบบธรรมดามีค่าความเที่ยงสูงสุด รองลงมาคือ แบบสอบที่มีตัวเลือกแบบปลายเปิด "ไม่มีคำตอบถูก" และแบบสอบที่มีตัวเลือกแบบแบบช่อง ตามลำดับ

แกรนคู (Ghandour 1985: 515-A) ได้ทำการศึกษาเปรียบเทียบคุณภาพของแบบสอบเลือกตอบที่มีลักษณะของตัวเลือกแตกต่างกัน 3 รูปแบบคือ แบบสอบที่มีตัวเลือกแบบธรรมดา แบบสอบที่มีตัวเลือกแบบปลายเปิด "ถูกทุกข้อ" และแบบสอบที่มีตัวเลือกแบบปลายเปิด "ไม่มีคำตอบถูก" ในด้านค่าอ่านาจนจำแนก ความยาก และความเที่ยง ของวิชาคณิตศาสตร์ ซึ่งแยกเนื้อหาออกเป็น แคลคูลัส แอลจีบรา (Algebra) และเรขาคณิต กลุ่มตัวอย่างเป็น

นักศึกษาจำนวน 565 คน เป็นชาย 274 คน หญิง 291 คน โดยแยกออกแบบทดลองเป็นแบบ Factorial $2 \times 3 \times 3$ แยกออกเป็น 3 แฟคเตอร์ คือ เพศ เนื้อหา และรูปแบบของตัวเลือก แบบสอบแต่ละฉบับมีข้อสอบฉบับละ 18 ข้อ ในการตั้งสมมติฐานแยกออกเป็น 3 กลุ่ม คือสมมติฐานเกี่ยวกับความเที่ยง 7 ข้อ เกี่ยวกับความยาก 7 ชุด และสมมติฐานเกี่ยวกับค่าอำนาจจำแนก 4 ข้อ ผลการวิจัยพบว่า ความเที่ยงของแบบสอบที่มีรูปแบบตัวเลือกแตกต่างกันมีค่าความแตกต่างกันทั้งในลักษณะรวม เนื้อหาและแยก เนื้อหาอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 เมื่อพิจารณาความยากปรากฏว่า มีความแตกต่างกันระหว่างเพศในกรณีแยกเนื้อหาอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 และค่าความยากที่เกิดจากปฏิสัมพันธ์ (Interaction) ระหว่างรูปแบบของตัวเลือกกับ เนื้อหาวิชา มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ส่วนค่าอำนาจจำแนกมีความแตกต่างกันระหว่างเพศในวิชาเรขาคณิต และกับรูปแบบของแบบสอบที่มีรูปแบบตัวเลือกแบบธรรมดาเท่านั้น

โทเลฟสัน (Telefson 1987: 377-383) ได้ศึกษาและเปรียบเทียบคุณภาพของแบบสอบเลือกตอบที่มีรูปแบบของตัวเลือกแตกต่างกัน 3 รูปแบบคือ ระบบสอบที่มีรูปแบบตัวเลือกแบบธรรมดา แบบสอบที่มีตัวเลือกแบบปลายเปิด "ไม่มีคำตอบถูก" เป็นตัวถูก และแบบสอบที่มีตัวเลือกแบบปลายเปิด "ไม่มีคำตอบถูก" เป็นตัวถูก และแบบสอบที่มีตัวเลือกแบบปลายเปิด "ไม่มีคำตอบถูก" เป็นตัวลวง แบบสอบทั้ง 3 ฉบับมีข้อสอบฉบับละ 12 ข้อ ซึ่งเป็นข้อสอบเกี่ยวกับวิชาสถิติ กลุ่มตัวอย่างเป็นนักเรียนระดับชั้นมัธยมศึกษาศึกษาจำนวน 82 คน ผลการวิจัยปรากฏว่า ค่าความยากของแบบสอบที่ใช้ตัวเลือกปลายเปิด "ไม่มีคำตอบถูก" ยากกว่าแบบสอบที่ใช้ตัวเลือกแบบธรรมดาอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ถ้าเปรียบเทียบระหว่างแบบสอบที่ใช้ตัวเลือกปลายเปิด "ไม่มีคำตอบถูก" เป็นตัวถูกกับแบบสอบที่ใช้ตัวเลือกแบบปลายเปิด "ไม่มีคำตอบถูก" เป็นตัวลวง ปรากฏว่า ตัวเลือกแบบปลายเปิด "ไม่มีคำตอบถูก" ที่เป็นตัวถูกยากกว่าตัวเลือกปลายเปิด "ไม่มีคำตอบถูก" ที่เป็นตัวลวง เมื่อพิจารณาค่าอำนาจจำแนกแล้วจะพบว่า แบบสอบที่มีตัวเลือกแบบธรรมดามีค่าอำนาจจำแนกสูงที่สุด แต่เมื่อทดสอบ ความแตกต่างทางสถิติไม่พบความแตกต่างอย่างมีนัยสำคัญทางสถิติ ส่วนค่าความเที่ยงของแบบสอบทั้ง 3 ฉบับที่คำนวณโดยใช้สูตร KR-20 ปรากฏว่าแบบสอบที่มีรูปแบบตัวเลือกแบบธรรมดามีค่าความเที่ยงสูงสุด

จากผลงานการวิจัยที่เกี่ยวกับคุณภาพของแบบสอบ เลือกตอบที่มีรูปแบบตัวเลือก แตกต่างกันในต่างประเทศดังกล่าวข้างต้นจะเห็นได้ว่า ส่วนใหญ่เป็นการศึกษาคุณภาพของแบบสอบในวิชาคณิตศาสตร์ ส่วนในวิชาอื่น ๆ นั้นมีน้อยมาก และจากการศึกษาข้อค้นพบเหล่านี้พบว่าคุณภาพของแบบสอบในด้านต่าง ๆ นั้นยังไม่สอดคล้องกันเช่น รีมแลนค์พบว่า การใช้ตัวเลือก "ไม่มีคำตอบถูก" ไม่ทำให้คุณภาพของแบบสอบในด้านความยาก อำนาจจำแนกต่างจากแบบสอบรูปแบบธรรมดา แต่จากการศึกษาของฮิวส์และทริมเบิล (Hughes and Trimble) พบว่า แบบสอบที่มีรูปแบบตัวเลือกแบบธรรมดา มีค่าความยากต่างกับแบบสอบที่ใช้ตัวเลือกแบบธรรมดา เป็นต้น จึงไม่อาจทำให้ได้ข้อสรุปที่แน่นอนได้

งานวิจัยภายในประเทศ

สนั่น สิทธิวัง (2512: 76) ได้ศึกษาเปรียบเทียบคุณภาพของแบบสอบ เลือกตอบที่ครูคิดสร้างตัวเลือกขึ้นเอง กับแบบสอบที่ใช้ตัวเลือกจากคำตอบของนักเรียนในวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 4 ชั้นประถมศึกษาปีที่ 7 และชั้นมัธยมศึกษาปีที่ 3 ของภาคการศึกษา 3 จำนวน 1,900 คน ผลการวิจัยพบว่า ค่าความเที่ยงของแบบสอบฉบับที่ครูได้ตัวเลือกจากนักเรียนสูงกว่าฉบับที่ครูคิดสร้างตัวเลือกขึ้นเองอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ในชั้นประถมศึกษาปีที่ 4 และมีนัยสำคัญทางสถิติที่ระดับ .01 ในชั้นประถมศึกษาปีที่ 7 และชั้นมัธยมศึกษาปีที่ 3 เมื่อพิจารณาอำนาจจำแนกของแบบสอบพบว่า แบบสอบที่ครูได้ตัวเลือกจากนักเรียนสูงกว่าฉบับที่ครูคิดสร้างตัวเลือกขึ้นเองอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ค่าความยากมาตรฐานของแบบสอบฉบับที่ครูได้ตัวเลือกจากนักเรียนยากกว่าฉบับที่ครูคิดสร้างตัวเลือกขึ้นเอง

ไพบุลย์ จิตรโค (2514: 103-105) ได้เปรียบเทียบคุณภาพของแบบสอบ เลือกตอบที่มีรูปแบบตัวเลือกต่างกัน 3 รูปแบบคือ ตัวเลือกแบบปลายเปิด "ไม่มีคำตอบถูก" แบบตัวลงเฉียง และแบบตัวลงดักกับแบบสอบแบบเติมคำ ผลการวิจัยพบว่า ค่าความตรงต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ในแบบสอบคณิตศาสตร์ทักษะ และคณิตศาสตร์ปัญหา ชั้นประถมศึกษาปีที่ 7 มีความยากต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ส่วนแบบสอบคณิตศาสตร์ทักษะและคณิตศาสตร์ปัญหาชั้นประถมศึกษาปีที่ 4 และชั้นประถมศึกษาปีที่ 7 มีค่าอำนาจจำแนกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และแบบสอบคณิตศาสตร์ปัญหาชั้นประถมศึกษาปีที่ 7 มีค่าความเที่ยงของแบบสอบทั้ง 4 แบบไม่แตกต่างกัน

นวนน้อย แต่บรรพกุล (2520: 62-65) ได้ศึกษาถึงสมรรถนะในการวัดของแบบสอบชนิดเลือกตอบที่มีรูปแบบตัวเลือกแบบต่าง ๆ ในวิชาภาษาไทย ในระดับชั้นประถมศึกษาปีที่ 7 โดยแบ่งแบบสอบออกเป็น 6 ชนิดคือ แบบสอบที่ใช้ตัวเลือกแบบธรรมดา แบบปลายเปิด แบบปลายปิด และแบบผสม ซึ่งได้แก่ "ถูกทั้ง ก และ ข" หรือ "ถูกทั้ง ก และ ค" แบบสอบแต่ละชุดยังแบ่งย่อยตามเนื้อหาเป็น 3 ฉบับคือ แบบสอบสะกดคำ หลักภาษา และความเข้าใจการอ่าน โดยแบบสอบแต่ละชุดประกอบด้วยข้อสอบจำนวน 35 ข้อ แบ่งเป็นข้อสอบที่ใช้เปรียบเทียบจริง 25 ข้อ ซึ่งใช้ตัวเลือกแบบต่าง ๆ เป็นตัวลวงหรือตัวถูกตัวสุดท้าย และอีกส่วนหนึ่งเป็นข้อสอบควบคุมหมายถึง ข้อสอบที่มีรูปแบบตัวเลือกแบบธรรมดา 5 ตัวเลือก ซึ่งเป็นข้อสอบที่ง่าย ๆ เพื่อไม่ให้มีผลต่อข้อสอบที่ใช้เปรียบเทียบจริง ผลการวิจัยครั้งนี้ปรากฏว่า รูปแบบของตัวเลือกไม่มีผลต่อค่าอำนาจจำแนก และความเที่ยงของแบบสอบ แต่มีผลต่อค่าความตรงของแบบสอบทุกฉบับ ยกเว้นแบบสอบสะกดคำ สำหรับความยากของแบบสอบแตกต่างกัน เฉพาะในแบบสอบความเข้าใจในการอ่านเท่านั้น จากผลการวิจัยครั้งนี้หากพิจารณาเฉพาะตัวเลือกแบบธรรมดา กับแบบปลายเปิด "ไม่มีคำตอบถูก" จะพบว่า ไม่แตกต่างกันทั้งในด้านความยาก อำนาจจำแนก ความเที่ยง ความตรง และคะแนนผลการสอบ

นิรมล บุญตะรันต์ (2525: 59) ได้ศึกษาเปรียบเทียบคุณภาพของแบบสอบเลือกตอบที่มีรูปแบบตัวเลือกแตกต่างกันบนวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1 โดยใช้รูปแบบตัวเลือกแตกต่างกัน 3 ฉบับคือ แบบสอบที่มีรูปแบบตัวเลือกแบบธรรมดาทั้งฉบับ แบบสอบที่มีรูปแบบที่มีรูปแบบตัวเลือกแบบปลายเปิด "ไม่มีคำตอบถูก" ทั้งฉบับ และแบบสอบที่มีรูปแบบตัวเลือกแบบผสมระหว่างแบบธรรมดาและแบบปลายเปิด "ไม่มีคำตอบถูก" จำนวน ฉบับละ 50 ข้อ ในด้านความยาก อำนาจจำแนก ความเที่ยง ความตรง ตลอดจนคะแนนเฉลี่ยจากการสอบของนักเรียนทั้ง 3 กลุ่ม กลุ่มตัวอย่างเป็นนักเรียนมัธยมศึกษาปีที่ 1 ของโรงเรียนสะเคา "ชรรค์ชัยกับพลาหนทอนุสรณ์" จังหวัดสงขลา จำนวน 200 คน โดยแต่ละคนทำแบบสอบเพียง 1 ฉบับ ผลการวิจัยพบว่า แบบสอบที่มีรูปแบบตัวเลือกแตกต่างกันทั้ง 3 ฉบับ ไม่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ ทั้งในด้านความยาก อำนาจจำแนก ความเที่ยง และความตรง ตลอดจนคะแนนเฉลี่ยจากการสอบของนักเรียน

จินดา ไตอนันต์ (2527: 61-62) ได้ศึกษาเปรียบเทียบคุณลักษณะของแบบสอบสัมฤทธิ์ผล แบบสอบเลือกตอบถูกผิดทุกตัวเลือกและแบบสอบเลือกตอบตัวเลือกเดียว

ในด้านค่าความยาก อ่านาจจำแนก และค่าความเที่ยง ของแบบสอบสัมฤทธิ์ผลวิชาวิทยาศาสตร์ (ว. 102) เรื่องสิ่งมีชีวิตและสิ่งแวดล้อม โดยการใช้แบบสอบ 2 ฉบับคือ แบบสอบเลือกตอบ ถูกผิดทุกตัวเลือก และแบบสอบเลือกตอบตัวเลือกเดียว กลุ่มตัวอย่างได้แก่ นักเรียนชั้นมัธยมศึกษาปีที่ 1 สังกัดกองการมัธยมศึกษา กรมสามัญศึกษา กระทรวงศึกษาธิการ ในกรุงเทพมหานคร จำนวน 432 คน ผลการวิจัยปรากฏว่า ในด้านความยากแบบสอบเลือกตอบตัวเลือกเดียว มีค่าอ่านาจจำแนกสูงกว่าแบบสอบเลือกตอบถูกผิดทุกตัวเลือกอย่างมีนัยสำคัญทางสถิติ และความเที่ยงของแบบสอบทั้ง 2 แบบไม่แตกต่างกัน

โกลม สรรคเสรมณี (2529: 43-44) ได้ศึกษาเปรียบเทียบคุณภาพของแบบสอบผลสัมฤทธิ์วิชาคณิตศาสตร์แบบเลือกตอบที่สร้างตัวลวงด้วยวิธีต่างกัน 3 วิธีคือ การให้ครูแต่ละคนสร้างตัวลวงเอง ให้ครูประชุมปรึกษากันสร้างตัวลวง และการใช้คำตอบผิดของนักเรียนมาเป็นตัวลวง โดยกลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 4 ของโรงเรียนบัวโรงเรียนศรีสวัสดิ์วิทยาคาร จำนวน 350 คน และครูผู้สอนจำนวน 12 คน เครื่องมือที่ใช้ในการวิจัยเป็นแบบสอบผลสัมฤทธิ์วิชาคณิตศาสตร์ (ค. 012) ชนิด 4 ตัวเลือก จำนวน 55 ข้อ โดยแบบสอบแต่ละฉบับมีค่าถ่วงเหมือนกันข้อต่อข้อ ต่างกันเฉพาะวิธีการสร้างตัวลวง ผลการวิจัยปรากฏว่า ค่าความเที่ยงของแบบสอบเลือกตอบฉบับที่ครูแต่ละคนสร้างขึ้นเองมีค่าความเที่ยงสูงสุด รองลงมาคือ ฉบับที่ครูประชุมปรึกษากันสร้างตัวลวง และฉบับที่ใช้คำตอบผิดของนักเรียนมาเป็นตัวลวง มีค่าความเที่ยงต่ำสุด สำหรับค่าความตรงและค่าอ่านาจจำแนกของแบบสอบเลือกตอบที่สร้างตัวลวงด้วยวิธีต่างกัน 3 วิธี ไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ

วรัญญา ปรามบุญจะ (2529: 44-46) ได้ศึกษาเปรียบเทียบคุณภาพของแบบสอบเลือกตอบที่มีรูปแบบตัวเลือกแตกต่างกันในวิชาวิทยาศาสตร์ ชั้นมัธยมศึกษาปีที่ 2 โดยมีแบบสอบ 3 ฉบับคือ แบบสอบเลือกตอบชนิดคำตอบเดียว แบบสอบเลือกตอบชนิดคำตอบซ้อนที่เงื่อนไขมีโอกาสได้รับเลือกเป็นตัวเลือกเท่ากัน และแบบสอบเลือกตอบชนิดคำตอบซ้อนที่เงื่อนไขมีโอกาสได้รับเลือกเป็นตัวเลือกไม่เท่ากันชุดละ 30 ข้อ โดยนำไปทดสอบกับกลุ่มตัวอย่างจำนวน 636 คน ผลการวิจัยปรากฏว่า แบบสอบชนิดคำตอบเดียวมีค่าความยากแตกต่างจากแบบสอบชนิดคำตอบซ้อนทั้ง 2 แบบ แต่ค่าความยากของแบบสอบชนิดคำตอบซ้อนทั้ง 2 แบบไม่แตกต่างกัน ส่วนความเที่ยงของแบบสอบชนิดคำตอบซ้อนที่เงื่อนไขมีโอกาสรับเลือกเป็นตัวเลือกไม่เท่ากันแตกต่างจากค่าความเที่ยงของแบบสอบชนิดคำตอบเดียว และแบบสอบชนิด

คำตอบข้อที่เงื่อนไขมีโอกาสรับเลือกเท่ากัน แต่แบบสอบชนิดคำตอบเดียว กับแบบสอบชนิดคำตอบข้อที่เงื่อนไขมีโอกาสรับเลือกเป็นตัวเลือกเท่ากัน มีค่าความเที่ยงไม่แตกต่างกัน

ไตรรงค์ เจนการ (2530: 84-85) ได้ศึกษาคุณภาพของแบบสอบวัดความสามารถในการแก้ปัญหาทางคณิตศาสตร์ ซึ่งเป็นแบบสอบที่เขียนลักษณะคำถามและคำตอบลักษณะต่าง ๆ กัน 3 แบบคือ เขียนคำถามในลักษณะแบบสอบ เอ็ม อี คิว (Modified Essay Question Test) แบบเลือกตอบและแบบผสม ในด้านค่าความยาก ค่าอำนาจจำแนก ค่าความเที่ยง และค่าความตรง ซึ่งแบบสอบมี 4 ฉบับคือ ฉบับเอ็ม อี คิว ฉบับเอ็ม อี คิว - เลือกตอบ ฉบับเลือกตอบ-เอ็ม ดี คิว และฉบับเลือกตอบ และเปรียบเทียบคุณภาพของแบบสอบทั้ง 4 ฉบับในด้านค่าความยาก ค่าอำนาจจำแนก ค่าความเที่ยง และค่าความตรง ในวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 กลุ่มตัวอย่างเป็นนักเรียนในสังกัดสำนักงานคณะกรรมการการประถมศึกษาแห่งชาติ กระทรวงศึกษาธิการ ในเขตกรุงเทพมหานคร จำนวน 550 คน ผลการวิจัยพบว่า ค่าความเที่ยงของแบบสอบเลือกตอบ และแบบสอบ เอ็ม อี คิว-เลือกตอบ มีค่าความเที่ยงต่ำกว่าแบบสอบ เอ็ม อี คิว อย่างมีนัยสำคัญทางสถิติ สำหรับค่าความยาก ค่าอำนาจจำแนก และค่าความตรงของแบบสอบทั้ง 4 ฉบับ ไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติและเมื่อพิจารณาถึงคะแนนเฉลี่ย ปรากฏว่า แตกต่างกันอย่างมีนัยสำคัญทางสถิติ

อำนวยการ ทอดวง (2531: 61-63) ได้ศึกษาผลของรูปแบบตัวเลือกของแบบสอบเลือกตอบที่มีต่อคุณภาพของแบบสอบเลือกตอบ โดยใช้แบบสอบเลือกตอบวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 ที่มีรูปแบบตัวเลือกแตกต่างกัน 4 แบบคือ แบบสอบเลือกตอบที่มีรูปแบบตัวเลือกแบบธรรมดา แบบสอบที่มีรูปแบบตัวเลือกแบบปลายเปิด แบบสอบที่มีรูปแบบตัวเลือกเป็นช่วง และแบบสอบที่มีรูปแบบตัวเลือกเป็นแบบใกล้เคียง โดยในแต่ละฉบับมีข้อสอบฉบับละ 30 ข้อ กลุ่มตัวอย่างเป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ของโรงเรียนในสังกัดสำนักงานการศึกษาจังหวัดบุรีรัมย์ จำนวน 824 คน ผลการวิจัยปรากฏว่า แบบสอบเลือกตอบที่มีรูปแบบตัวเลือกแตกต่างกันทั้ง 4 รูปแบบ มีความยากแตกต่างกันอย่างมีนัยสำคัญทางสถิติ ซึ่งแบบสอบที่มีรูปแบบตัวเลือกเป็นแบบใกล้เคียงยากกว่าแบบสอบที่มีรูปแบบตัวเลือกแบบธรรมดา ในด้านอำนาจจำแนกของแบบสอบพบว่า อำนาจจำแนกของแบบสอบเลือกตอบที่มีรูปแบบตัวเลือกแตกต่างกัน 4 รูปแบบไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ ในด้านความเที่ยงของแบบสอบพบว่า ค่าความเที่ยงของแบบสอบเลือกตอบที่มีรูปแบบตัวเลือกแตกต่างกัน 4 รูปแบบแตกต่างกันอย่าง

อย่างมีนัยสำคัญทางสถิติ โดยที่แบบสอบที่มีรูปแบบตัวเลือกแบบธรรมชาติมีค่าความเที่ยงสูงสุด

จากผลงานการวิจัยที่เกี่ยวกับคุณภาพของแบบสอบเลือกตอบที่มีรูปแบบตัวเลือกแตกต่างกันภายในประเทศดังกล่าวข้างต้นจะเห็นได้ว่า ส่วนใหญ่แล้วเป็นการศึกษาคุณภาพของแบบสอบในวิชาคณิตศาสตร์ ส่วนในวิชาอื่น ๆ นั้นมีน้อยมากเท่าที่พบก็มีวิชาวิทยาศาสตร์และภาษาไทย และจากการศึกษาข้อค้นพบเหล่านี้พบว่า คุณภาพของแบบสอบในด้านต่าง ๆ ยังไม่สอดคล้องกัน เช่น จากการศึกษาของ นิรมล บุญตะรันตน์ พบว่า แบบสอบที่มีรูปแบบตัวเลือกแตกต่างกันไม่มีความแตกต่างกันในด้าน ความยาก อ่านง่าย-ยาก ความเที่ยง ความตรง และคะแนนเฉลี่ย แต่ผลจากการศึกษาของ อำนวย ทองคัง พบว่า แบบสอบที่มีรูปแบบตัวเลือกต่างกัน ค่าความยาก และค่าความเที่ยงแตกต่างกัน เป็นต้น

จากงานวิจัยที่เกี่ยวข้องดังกล่าวมาทั้งหมดทั้งต่างประเทศและภายในประเทศ จะเห็นว่าวัตถุประสงค์ของการวิจัยใกล้เคียงกันมาก แต่ผลที่ได้รับจากการวิจัยบางส่วนยังไม่สอดคล้องกัน จึงไม่สามารถที่จะสรุปผลที่แน่นอนได้ว่ารูปแบบตัวเลือกของแบบสอบเลือกตอบจะมีผลต่อคุณภาพของแบบสอบหรือไม่ และมีตัวแปรใดบ้างที่ส่งผลต่อคุณภาพของแบบสอบร่วมกับรูปแบบของตัวเลือกแบบต่าง ๆ ซึ่งควรจะได้ทำการศึกษาต่อไป เพื่อนำผลที่ได้ไปปรับปรุงในการสร้างข้อสอบของครูต่อไป

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย