

บทที่ 1

บทนำ



ความเป็นมาและความสำคัญของปัญหา

ทฤษฎีการวัดแบบดั้งเดิม (Classical Test Theory) เป็นทฤษฎีการทดสอบที่มีการพัฒนาเป็นเวลานาน และบุคคลที่มีอิทธิพลมากในการนำเสนอทฤษฎีนี้ เมื่อ ค.ศ. 1950 คือ กุลลิคเซน (Gulliksen) (Warm, 1978: 15) ทฤษฎีนี้มุ่งตรวจสอบความสัมพันธ์ ระหว่างคะแนนที่สังเกต (observe score) กับคะแนนที่แท้จริง (true score) และหาคุณภาพโดยรวมของเครื่องมือที่ใช้วัดบุคคลในสถานการณ์ทดสอบที่เฉพาะเจาะจง ตั้งอยู่บนข้อตกลงเบื้องต้นที่ว่า ความคลาดเคลื่อนที่เกิดจากการวัด มีแบบแผนที่คล้ายคลึงกันสำหรับทุกคนในการตอบ จึงไม่สนใจที่จะสรุปอ้างอิง โดยเชื่อว่าคะแนนจากการสังเกตที่วัดได้นั้น ส่วนหนึ่งเป็นคะแนนความสามารถที่แท้จริง และอีกส่วนหนึ่งเป็นคะแนนความคลาดเคลื่อน ซึ่งถือว่าคะแนนความคลาดเคลื่อน หรือความคลาดเคลื่อนของการวัด เป็นความคลาดเคลื่อนแบบสุ่ม และเป็นคะแนนรวมของความคลาดเคลื่อนจากทุกแหล่งเข้าด้วยกัน ในลักษณะเป็นความคลาดเคลื่อนแหล่งเดียวเท่านั้น ไม่สามารถแยกแหล่งของความคลาดเคลื่อนออกเป็นส่วน ๆ ได้ ดังนั้นความแปรปรวนของคะแนนจะประกอบด้วยความแปรปรวนเพียง 2 แหล่งเท่านั้น คือความแปรปรวนของคะแนนที่แท้จริง และความแปรปรวนของคะแนนความคลาดเคลื่อน เมื่อนำไปสู่การประมาณค่า จึงสามารถเพียงประมาณค่าความคลาดเคลื่อนในแหล่งเดียวเท่านั้น โดยไม่ได้คำนึงถึงสภาพแวดล้อมหรือเงื่อนไขการวัดที่อาจแตกต่างกัน ทั้งนี้ก็เพราะคะแนนของบุคคลนั้น มักจะไม่คงเส้นคงวาเสมอไปหรือมีความผันแปรของสถานการณ์วัดเกิดขึ้น ทำให้คะแนนของบุคคลมีความแตกต่างเกิดขึ้นอันเนื่องจากปัจจัยต่าง ๆ เป็นต้นว่า ความแตกต่างของรูปแบบของแบบสอบ ความยาวของแบบสอบ และผู้บริหารการสอบ (Shevelson, 1991: 2) อันนำไปสู่ปัญหาของการประมาณค่าความเที่ยงและความตรง บนพื้นฐานของการประมาณค่าตามทฤษฎีการวัดแบบดั้งเดิม

จากการที่ทฤษฎีการวัดแบบดั้งเดิมมุ่งเน้นที่ความคลาดเคลื่อนเชิงสุ่ม โดยให้ความสนใจกับความคลาดเคลื่อนเชิงระบบค่อนข้างน้อยมาก ทั้งนี้อาจเป็นเพราะมุ่งเน้นไปที่ข้อตกลงเบื้องต้นซึ่งค่อนข้างเข้มงวด ในเรื่องการวัดที่เป็นคู่ขนาน (strictly parallel measurement) ซึ่งเป็นการกำหนดว่าค่าเฉลี่ย และความแปรปรวนของคะแนนในแต่ละข้อกระทงเท่ากัน จึงเป็น

การยอมรับว่าไม่มีความแปรปรวนที่เนื่องมาจากข้อกระทง (Shevelson, 1991: 94) ดังนั้น ทฤษฎีการวัดแบบดั้งเดิมที่ยังใช้กันอยู่จึงเป็นการประมาณค่าความเที่ยงของวิธีวัด ที่ใช้กับเฉพาะ กลุ่มบุคคล ภายใต้งื่อนไขของการทดสอบที่เฉพาะ ได้แก่ การทดสอบที่ผู้สอบทุกคนต้องทำข้อสอบ ทุกข้อ ผลการตอบได้รับการให้คะแนนโดยผู้ตรวจคนเดียว และแบบสอบที่ใช้ตั้งอยู่บนข้อตกลง เบื้องต้นของความเป็นคู่ขนานระหว่างแบบสอบ แต่ถ้าสถานการณ์ของการทดสอบแตกต่างไปจาก ที่กล่าวมา การประมาณค่าความเที่ยงตามทฤษฎีการวัดแบบดั้งเดิมก็ไม่สามารถนำมาประยุกต์ใช้ได้ เช่นกรณีของการทดสอบที่มีผู้ตรวจหลายคน ให้คะแนนข้อสอบความเรียงของผู้สอบทุกคน หรือกรณี การทดสอบที่ใช้ผู้ตรวจหลายคน ให้คะแนนผลการตอบแบบสอบที่มีความยาวต่างกัน หรือกรณีผู้ตรวจ หลายคนให้คะแนนผลการตอบแบบสอบที่มีรูปแบบและความยาวแตกต่างกัน (ศิริชัย, 2535: 109) ดังนั้น เมื่อผู้วิจัยต้องการประมาณค่าความเที่ยง โดยมุ่งศึกษาตั้งแต่ 2 องค์ประกอบขึ้นไปพร้อมกัน เช่น องค์ประกอบด้านผู้ตรวจและจำนวนครั้งของการตรวจ ในการทดสอบด้วยข้อสอบแบบความ เรียงแล้ว การวิเคราะห์ตามทฤษฎีการวัดแบบดั้งเดิม ไม่สามารถพิจารณาร่วมกันในองค์ประกอบ ทั้งสองด้านไปพร้อม ๆ กัน แต่ต้องพิจารณาครั้งละองค์ประกอบเดียวเท่านั้น เป็นต้นว่า เมื่อต้อง การประมาณค่าความเที่ยงที่พิจารณาจากองค์ประกอบด้านผู้ตรวจ ก็จะต้องทำการประมาณค่าจาก องค์ประกอบด้านผู้ตรวจ โดยไม่สามารถทำการพิจารณาองค์ประกอบด้านจำนวนครั้งของการตรวจ ไปพร้อม ๆ กับองค์ประกอบด้านผู้ตรวจ ลักษณะนี้จึงเป็นการประมาณค่าที่ใช้การพิจารณาแยกจาก กันระหว่างผู้ตรวจและจำนวนครั้งของการตรวจ การประมาณค่าจึงเป็นการพิจารณาแหล่งความ แปรปรวนที่ต้องแยกจากกันทีละส่วน ไม่สามารถพิจารณาแหล่งความแปรปรวนที่เป็นไปได้ทั้งหมด ในการวิเคราะห์เพียงครั้งเดียวได้ ทฤษฎีนี้จึงมีข้อจำกัดที่ไม่สามารถศึกษาในรายละเอียดของ แหล่งความคลาดเคลื่อนจากการวัดต่าง ๆ ได้เพียงพอ

ด้วยข้อจำกัดของทฤษฎีการวัดแบบดั้งเดิมในด้านการวัดที่เป็นคู่ขนานและการวัดที่ถือว่า มี แหล่งความคลาดเคลื่อนรวมแหล่งเดียว ทำให้ค่าที่ประมาณได้นั้น มีอาจแบ่งแยกได้ นอกจากนั้น ยังเป็นทฤษฎีที่มุ่งเน้นการวิเคราะห์สภาพการสอบแบบองค์ประกอบเดียว ซึ่งส่วนมากจะเป็น ความ ยาวของข้อกระทง เช่นศึกษาว่า ถ้าเพิ่มจำนวนข้อกระทงแล้ว ค่าประมาณความเที่ยงจะเพิ่มขึ้น เท่าไร โดยไม่ได้มุ่งพิจารณาจากหลายองค์ประกอบไปพร้อม ๆ กัน เช่น ผู้ตรวจ จำนวนครั้งของ การตรวจ ฯลฯ อันทำให้เกิดข้อคิดขึ้นมาว่า ค่าประมาณความเที่ยงที่ได้จากทฤษฎีนี้ จะเป็นค่า ประมาณความเที่ยงที่แท้จริง ในทฤษฎีการวัดของการวัดที่ต้องการศึกษาหรือไม่ เนื่องจากไม่ได้ พิจารณาจากแหล่งความคลาดเคลื่อนที่เป็นไปได้ทั้งหมด และเมื่อประมาณค่าความเที่ยงที่อาศัยจาก แหล่งความคลาดเคลื่อนเพียงแหล่งเดียวเช่นนี้แล้ว น่าจะเป็นข้อควรพิจารณาอย่างยิ่ง ต่อการ นำไปใช้ในการวัดผลการศึกษา ในอันที่จะยอมรับวิธีการประมาณค่าความเที่ยง และยิ่งเมื่อต้อง

การศึกษาความตรงควบคู่ไปกับความเที่ยงด้วยแล้ว ก็ยังต้องให้ความสำคัญกับการประมาณค่าความเที่ยงเพิ่มมากยิ่งขึ้นด้วย ทั้งนี้เนื่องจากความตรงและความเที่ยงมีความสัมพันธ์กัน กล่าวคือ กระบวนการวัดที่จะมีความตรงได้นั้น จะต้องมีความเที่ยงด้วย หากความเที่ยงที่ได้จากการวัดยังมีข้อเคลือบแคลงใจอยู่เช่นนี้ จะทำให้ยอมรับความตรงที่ได้จากทฤษฎีนี้ได้หรือไม่ ทั้งนี้ก็เพราะในกระบวนการวัดนั้น คุณลักษณะที่สำคัญก็คือความเที่ยงและความตรง ในขณะที่ความเที่ยงเกี่ยวข้องกับความแม่นยำ (precision) นั้น ความตรงจะเกี่ยวข้องกับความถูกต้อง (accuracy) (Stalling & Gilmore, 1971 cited by Kane, 1982) ความตรงจะเกี่ยวข้องกับการแปลความคะแนนสังเกต ในลักษณะของการเป็นตัวแทนของคุณสมบัติภายนอก และความเที่ยงจะเกี่ยวข้องกับความคงเส้นคงวาระหว่างคะแนนสังเกต แม้ว่าความเที่ยงจะมีความสำคัญต่อกระบวนการวัดก็ตาม แต่ถ้าความเที่ยงนั้นให้ความคงเส้นคงวาในระดับสูง จากการวัดคุณลักษณะที่ไม่ถูกต้องแล้ว ค่าความคงเส้นคงวาในระดับสูงเช่นนี้ ก็จะมีประโยชน์น้อยกว่าความคงเส้นคงวาในระดับที่ต่ำกว่า ที่เกิดจากการวัดคุณลักษณะในสิ่งที่ต้องการวัดอย่างแท้จริง ดังนั้นจึงนับได้ว่า ความตรงมีความสำคัญมากกว่าความเที่ยง (Kane, 1980:5)

ดังนั้น การที่ทฤษฎีการวัดแบบดั้งเดิมถือว่ามิแหล่งความคลาดเคลื่อนแหล่งเดียว ทำให้ค่าความเที่ยงที่ประมาณได้มีค่ามากกว่าที่ควรจะเป็น นอกจากนี้ ยังมุ่งเน้นที่การวิเคราะห์องค์ประกอบเดียว ซึ่งส่วนมากได้แก่ข้อกระทง เช่นศึกษาว่า ถ้าเพิ่มจำนวนข้อกระทงแล้ว ค่าประมาณความเที่ยงจะเพิ่มขึ้นเท่าไร โดยไม่มุ่งพิจารณาจากอีกหลายองค์ประกอบที่ควรทำการศึกษา เช่น ผู้ตรวจ จำนวนครั้งของการตรวจ ฯลฯ ทำให้ทฤษฎีการวัดแบบดั้งเดิม สามารถนำไปประยุกต์ใช้ในการวัดผลการศึกษาได้ในวงจำกัด ทำให้ผลที่ได้จากการวิเคราะห์ ไม่สามารถนำไปใช้ได้อย่างกว้างขวาง และไม่สามารถสรุปอ้างอิงได้

ต่อมาใน ค.ศ. 1963 ครอนบาค เกลสเซอร์ และราชารัทนัม (Cronbach, Gleser and Rajaratnam, 1963) ได้พัฒนาทฤษฎีการอ้างอิงสรุป (Generalizability Theory) ขึ้นเป็นครั้งแรก และได้มีการพัฒนาทฤษฎีนี้ต่อมาเป็นระยะ ๆ เป็นทฤษฎีที่ใช้เพื่อประมาณความเที่ยงทั่วไปของแบบสอบในสถานการณ์การสอบต่าง ๆ ได้ปรับปรุงมาจากทฤษฎีการวัดแบบดั้งเดิม ในลักษณะของการวิเคราะห์ความแปรปรวนจากหลาย ๆ แหล่ง (multifacet analysis of variance) อันครอบคลุมเงื่อนไขการวัดหรือแหล่งความคลาดเคลื่อนต่าง ๆ ด้วย โดยเชื่อว่าความแปรผันของผลการวัด เกิดขึ้นได้จากแหล่งต่าง ๆ หลายแหล่งด้วยกันที่เป็นไปได้ทั้งหมด ทั้งจากผลหลัก (main effect) และผลของผลร่วมหรือปฏิสัมพันธ์ (interaction effect) ด้วย ทฤษฎีนี้ไม่ได้ขึ้นกับข้อตกลงเบื้องต้นอย่างเข้มงวดของการวัดที่ต้องเป็นแบบคู่ขนาน แต่เป็น

ทฤษฎีที่ยินยอมให้นักวิจัยทำการออกแบบวิธีการเก็บรวบรวมข้อมูล ให้เหมาะสมสำหรับแต่ละปัญหาของการวัดผลภายใต้การพิจารณาต่าง ๆ กันได้ (Schroeder and Hakstian, 1990: 429) ทฤษฎีนี้จึงได้รับการยอมรับในทางการวัดผลการศึกษา เนื่องจากครอบคลุมเงื่อนไขของการวัด ซึ่งเป็นประโยชน์ต่อการวัดผลการศึกษา โดยเฉพาะในการพัฒนาแบบสอบ เพื่อให้ได้ความเที่ยงที่ต้องการ ภายใต้สถานการณ์การสอบที่ต้องการนำแบบสอบไปใช้ อันจะทำให้เกิดความถูกต้อง ความแม่นยำ และได้สารสนเทศที่เพียงพอ สำหรับการตัดสินใจเลือกแบบการวัดที่มีประสิทธิภาพ

กระบวนการวิเคราะห์ในทฤษฎีการอ้างอิงสรุป ประกอบด้วย 2 ขั้นตอน คือ ขั้นตอนการศึกษาเพื่อการอ้างอิง หรือการศึกษา G (Generalizability Study Design or G Study) และการศึกษาเพื่อการตัดสินใจ หรือการศึกษา D (Decision Study Design or D Study) โดยขั้นการศึกษาเพื่อการอ้างอิง หรือ G Study นั้น เป็นขั้นตอนของการรวบรวมข้อมูล เพื่อที่จะประมาณค่าความแปรปรวนจากแหล่งต่างๆ ในการทำ G Study จะประกอบด้วย การกำหนดฟาเซต (facet) หรือองค์ประกอบต่าง ๆ พิจารณาความสัมพันธ์ระหว่างฟาเซต กำหนดเอกภพของการสังเกตที่ยอมรับได้ (universe of admissible observation) โดยใช้การวิเคราะห์ความแปรปรวน (analysis of variance) เพื่อประมาณค่าความแปรปรวนขององค์ประกอบต่าง ๆ (variance component)

สำหรับขั้นการศึกษาเพื่อการตัดสินใจ หรือ D Study นั้น เป็นการรวบรวมข้อมูลเพื่อทำการตัดสินใจหรือลงข้อสรุป ในการทำ D Study นั้น จะต้องกำหนดเอกภพของการอ้างอิง (universe of generalizability) หาความแปรปรวนของคะแนนเอกภพ (universe score variance) ความแปรปรวนของส่วนที่เป็นความคลาดเคลื่อน (error variance) แล้วหาค่าสัมประสิทธิ์การอ้างอิงสรุป (generalizability coefficient) ซึ่งได้จากการปรับแบบแผนการวัด เพื่อให้ได้ค่าสัมประสิทธิ์การอ้างอิงสรุปที่เหมาะสม

ในส่วนของการแปลความของสัมประสิทธิ์การอ้างอิงสรุปนั้น จะคล้ายกับการแปลความของสัมประสิทธิ์ความเที่ยงในทฤษฎีการวัดแบบดั้งเดิม โดยจะเป็นดัชนีที่บอกให้ทราบว่า ถ้าทำการสังเกตหรือวัดภายใต้เงื่อนไขอื่น ๆ ในเอกภพของการอ้างอิงเดียวกันแล้ว ผลการวัดจะมีความคงเส้นคงวาเพียงใด ดังนั้นจะเห็นได้ว่า ในทฤษฎีการอ้างอิงสรุปนั้นมีความเกี่ยวข้องกับการวิเคราะห์ความคงเส้นคงวาของการวัดจากการวิเคราะห์หลายฟาเซต โดยนำความแปรปรวนในคะแนนสังเกต ไปสัมพันธ์กับเงื่อนไขของการวัดต่าง ๆ ที่ได้มาโดยการสุ่ม ทฤษฎีนี้จึงให้กรอบทั่วไปในการพิจารณาความเชื่อถือได้ของการวัด (dependability of measurement)

แม้ว่าทฤษฎีนี้ จะให้กรอบการพิจารณาค่าสัมประสิทธิ์การสรุปร่างอิง ในลักษณะของสัมประสิทธิ์ความเที่ยง (a reliability-like coefficient) (Cronbach, 1972: 5) โดยไม่ได้มุ่งเน้นในเรื่องของความตรงมากนัก ทั้ง ๆ ที่ความตรงและความเที่ยง ต่างก็มีความสัมพันธ์ต่อกัน ความเที่ยงได้บ่งชี้ถึง ความคงเส้นคงวาของชุดคะแนน จากการวัดในสิ่งที่ต้องการวัด จึงเกี่ยวข้องกับความแม่นยำของการวัด และสะท้อนถึงระดับของความคงเส้นคงวา (Ebel, 1979: 310) ในขณะที่ความตรงบ่งชี้ถึงความถูกต้องของการวัดในสิ่งที่ต้องการวัด ซึ่งเกี่ยวข้องกับความหมายของการวัด ความตรงจึงเป็นสิ่งที่เกี่ยวข้องกับผลของแบบสอบ หรือเครื่องมือที่ใช้ประเมิน ไม่ใช่ความตรงของเครื่องมือโดยตรง เป็นความตรงของผลการสอบ หรือความตรงของการแปลความหมายที่ได้จากการสอบที่บ่งชี้ในลักษณะระดับ (Gronlund, 1976: 80) โดยไม่ควรตัดสินว่า "มี" หรือ "ไม่มี" ความตรง แต่ควรตัดสินว่ามีความตรงในลักษณะระดับมากน้อย และขึ้นอยู่กับกระบวนการวัด (Kane, 1982:134)

เมื่อพิจารณาในประเด็นของความตรงแล้ว จะเห็นได้ว่าเป็นการพิจารณา ทั้งในด้านความแม่นยำของการวัด และพิจารณาความตรงในลักษณะที่ว่า วัดอะไร และจะวัดได้แม่นยำได้อย่างไร การที่จะวัดให้แม่นยำได้นั้น จึงเป็นการระบุความเที่ยงของคะแนนที่ได้จากการวัดนั้นด้วยความเที่ยงจึงเป็นเงื่อนไขที่จำเป็นของความตรง (Ebel and Frisbie, 1986: 89-90) ดังนั้น การวัดที่มีความตรง จำเป็นที่ต้องมีความเที่ยงด้วย แต่การวัดที่มีความเที่ยงสูง อาจมีค่าความตรงต่ำก็ได้เพราะอาจไม่ได้วัดในสิ่งที่มุ่งทำการวัดก็ได้ (Ebel, 1979: 444) ความเที่ยงจึงมีส่วนสำคัญในการกำหนดความตรง โดยความตรงจะมีค่าไม่เกินรากที่สองของความเที่ยง ซึ่งแสดงว่าความตรงจะถูกจำกัดโดยค่าความเที่ยง (Brown, 1976)

ดังนั้นการที่ทฤษฎีการอ้างอิงสรุปร่างอิงได้ใช้การวิเคราะห์ความแปรปรวน มาศึกษาแหล่งของความแปรปรวนต่าง ๆ เพื่อประมาณค่าองค์ประกอบของความแปรปรวน และนำไปสู่การอ้างอิงสรุปร่าง เพื่อหาค่าสัมประสิทธิ์การอ้างอิงสรุปร่างอิงสรุปร่างอิงในลักษณะของความเที่ยงนั้น ค่าสัมประสิทธิ์การอ้างอิงสรุปร่างอิงตามทฤษฎีการอ้างอิงสรุปร่างอิง ก็น่าจะพัฒนาไปสู่ดัชนีของความตรงได้เช่นกัน แต่ก็ยังมีการนำทฤษฎีนี้ไปใช้ในการศึกษาความตรงน้อยมาก ทั้ง ๆ ที่ ความตรงและความเที่ยงเป็นคุณสมบัติที่มีความสัมพันธ์ต่อกัน และต่างก็เป็นลักษณะที่สำคัญของการวัด เมื่อทฤษฎีการอ้างอิงสรุปร่างอิงสามารถนำไปประยุกต์ใช้เพื่อศึกษาความเที่ยงได้ ก็น่าจะศึกษาแนวทางประยุกต์ใช้เพื่อวิเคราะห์ความตรงได้

อย่างไรก็ตาม การศึกษาแนวคิดเกี่ยวกับความตรงตามทฤษฎีการอ้างอิงสรุป ก็ได้มีความพยายามในการดำเนินการอยู่ ผู้ที่มีบทบาทสำคัญในการศึกษาประเด็นความตรงตามทฤษฎีการอ้างอิงสรุป ก็คือ ไมเคิล เคน (Michael Kane) (อ้างใน Brennan, 1983: 77)

เคน (Michael Kane, 1982: 125-160) ได้ศึกษารอบแนวคิดทางสถิติของทฤษฎีการอ้างอิงสรุป เพื่อใช้ในการศึกษาสัมประสิทธิ์ความตรง โดยเคนพิจารณาว่า คະแนนที่เกิดจากการวัดหรือคะแนนที่สังเกตได้นั้น เกิดจากหลาย ๆ แหล่งที่เป็นไปได้ทั้งหมด ทั้งจากผลหลักและผลของปฏิสัมพันธ์ ทั้งนี้ก็เพราะทฤษฎีการอ้างอิงสรุป มุ่งเน้นที่ขนาดของความแปรปรวนเชิงสัมพันธ์เนื่องมาจากองค์ประกอบของแหล่งต่าง ๆ อันเป็นการให้ค่าประมาณของขนาดความคลาดเคลื่อนของการวัด ในลักษณะขององค์ประกอบของความแปรปรวน ดังนั้นสัมประสิทธิ์การอ้างอิงสรุปจะสะท้อนความเที่ยงว่าคะแนนที่สังเกตได้ จะมีความคล้ายคลึงหรือแตกต่างกันเพียงใด ภายใต้สถานการณ์การวัดอื่น ๆ จากกรอบแนวคิดของทฤษฎีนี้ จึงสามารถนำไปใช้ในการประมาณขนาดความแปรปรวนของคะแนนที่สังเกตได้ อันเนื่องมาจากการสุ่มวิธีวัด ซึ่งเป็นที่มาของการใช้ทฤษฎีการอ้างอิงสรุปในการศึกษาความตรง

การศึกษาความตรงตามทีเคนได้เสนอไว้ โดยใช้ทฤษฎีการอ้างอิงสรุปก็ยังมีข้อจำกัดในลักษณะของการนำไปประยุกต์ใช้ เนื่องจากเป็นการนำเสนอ ในลักษณะของแนวคิดมากกว่าประเด็นทางด้านเทคนิค โดยมุ่งเสนอผลในรูปขององค์ประกอบของความแปรปรวน จึงยังขาดวิธีที่จะนำไปใช้ในการปฏิบัติอย่างแท้จริง ปี ค.ศ. 1990 ซุน (Suen, 1990) กล่าวถึงการศึกษาความตรง เมื่อใช้วิธีวัดหลาย ๆ วิธีในเนื้อหาเดียวกัน ซึ่งเป็นการชี้ให้เห็นความตรงในการวัด โดยใช้ทฤษฎีการอ้างอิงสรุปตามแนวคิดของเคน โดยกล่าวไว้ในหนังสือ Principles of Test Theories

ในปี ค.ศ. 1993 เซนเวลสัน และคณะ (Shavelson, Boxtor and Gao, 1993: 215-232) ได้ใช้แนวคิดของเคนเป็นส่วนหนึ่งของการหาความตรงของการศึกษาการประเมินการปฏิบัติงานตามทฤษฎีการอ้างอิงสรุป เป็นการศึกษาสัมประสิทธิ์การอ้างอิงสรุป ในลักษณะของดัชนีความตรง (validity coefficient) อันประกอบด้วยเอกภพที่ซับซ้อน ที่เนื่องมาจากองค์ประกอบการวัด และองค์ประกอบวิธีวัด โดยองค์ประกอบการวัดนั้น ประกอบด้วยองค์ประกอบของงาน โอกาส และผู้ประเมิน เซนเวลสันและคณะได้ใช้ทฤษฎีนี้ในการศึกษาความเที่ยงและความตรง ประกอบด้วยการอ้างอิงสรุปไปสู่ความเที่ยงของผู้ประเมิน (Interrater Reliability) ความเที่ยงของงาน (Internal Reliability) และความเที่ยงจากการ

วัดซ้ำ (Retest Reliability) และใช้ทฤษฎีการอ้างอิงสรุปในการวัดความตรงที่มาจากวิธีวัด 4 วิธี ประกอบด้วย 1) การสังเกตการปฏิบัติงานของนักเรียนจากงานที่มอบหมาย 2) สมุดบันทึกของนักเรียน 3) สถานการณ์จำลองทางคอมพิวเตอร์ และ 4) การตอบคำถามสั้น ๆ กล่าวได้ว่า งานของเซพเวลสันและคณะ เป็นงานวิจัยที่มีการประยุกต์ใช้ทฤษฎีการอ้างอิงสรุปให้ครอบคลุมกระบวนการวัด ทั้งการวัดความเที่ยงและการศึกษาความตรงในลักษณะการอ้างอิงสรุป

เมื่อพิจารณาจากการศึกษาและงานวิจัยเกี่ยวกับทฤษฎีการอ้างอิงสรุปโดยทั่วไปนั้น ยังมีการศึกษาทฤษฎีนี้ไม่กว้างขวางนัก ซึ่งอาจจะเป็นเพราะความสลับซับซ้อนทางด้านคณิตศาสตร์ เป็นส่วนสำคัญส่วนหนึ่ง ที่อาจทำให้ทฤษฎีนี้ยังไม่แพร่หลายเท่าทฤษฎีการวัดแบบดั้งเดิม ในขณะที่งานวิจัยในประเทศไทยที่ใช้ทฤษฎีการอ้างอิงสรุป ก็ยังเป็นการศึกษาเพื่อนำไปสู่ การสรุปอ้างอิงในลักษณะของความเที่ยงเท่านั้น เช่น แดง กลางท่าได้ (2531) ได้ประยุกต์ใช้ทฤษฎีการอ้างอิงสรุปในการหาความเที่ยงของการประเมินความตรงเชิงเนื้อหา โดยมีวัตถุประสงค์เพื่อใช้ทฤษฎีนี้ในการค้นหาแหล่งความแปรปรวน ที่มีอิทธิพลต่อการหาค่าความเที่ยงของการประเมินความตรงเชิงเนื้อหาของแบบสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์โดยผู้เชี่ยวชาญ และหารูปแบบการวัดที่ให้ค่าความเที่ยงของการประเมินสูง พร้อมทั้งหาขนาดของกลุ่มตัวอย่างที่จะเป็นตัวแทนของประชากรในองค์ประกอบที่ศึกษา 3 องค์ประกอบ คือ ข้อสอบ ผู้เชี่ยวชาญ และโรงเรียน และยังมีการศึกษาของไพรัตน์ วงษ์นาม (2533) ที่ได้ศึกษาสัมประสิทธิ์การอ้างอิงสรุปของแบบสอบความเรียง โดยมีวัตถุประสงค์เพื่อศึกษา และเปรียบเทียบผลของวิธีตรวจ การชี้แจง และการรู้ผลการเรียนรู้ของผู้ตอบ ที่มีต่อค่าสัมประสิทธิ์การอ้างอิงสรุป ในลักษณะของความเที่ยงที่ได้จากการใช้แบบสอบความเรียง

การศึกษาเป็นสิ่งสำคัญและมีความจำเป็นต่อชีวิต การศึกษาช่วยให้บุคคลสามารถใช้เป็นเครื่องมือในการพัฒนาศักยภาพของตนเองได้ อันจะนำไปสู่การพัฒนาประเทศต่อไป ดังนั้นการจัดการศึกษาให้กับบุคคลจึงเป็นสิ่งจำเป็นต่อการพัฒนากำลังคน โดยต้องอาศัยกระบวนการต่าง ๆ เพื่อมุ่งที่จะพัฒนาบุคคลให้มีความเจริญงอกงาม ทั้งทางด้านร่างกาย สติปัญญา อารมณ์ อันเป็นที่ยอมรับของสังคม การจัดการศึกษาทุกระดับจึงมีความสำคัญต่อการพัฒนาประเทศ โดยเฉพาะอย่างยิ่งการจัดการศึกษาในระดับประถมศึกษา นับได้ว่ามีความสำคัญมากต่อการพัฒนากำลังคนส่วนใหญ่ของประเทศ ทั้งนี้เนื่องจากการจัดการศึกษาในระดับประถมศึกษานั้น เป็นการจัดการศึกษาในระดับภาคบังคับที่จัดให้กับเยาวชนทุกคน เป็นการจัดการศึกษาขั้นพื้นฐาน ที่มุ่งให้ผู้เรียนมีความรู้ ความสามารถ ทักษะ และเจตคติในระดับพื้นฐาน สำหรับการดำรงชีวิตอยู่ในสังคม และในหลักสูตรประถมศึกษา พุทธศักราช 2521 (ฉบับปรับปรุง 2533) นั้น กำหนดให้วิชาคณิตศาสตร์

เป็นวิชาบังคับ และอยู่ในกลุ่มทักษะที่เป็นเครื่องมือการเรียนรู้ โดยกำหนดจุดประสงค์ของหลักสูตร ที่มุ่งให้ผู้เรียนได้พัฒนาความสามารถในการคิดคำนวณ สามารถนำคณิตศาสตร์ไปใช้เป็นเครื่องมือ ในการเรียนรู้สิ่งต่าง ๆ และในการดำรงชีวิตให้มีคุณภาพ

การวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ ซึ่งมุ่งวัดผลสัมฤทธิ์ทางการเรียนของ นักเรียน ตามจุดมุ่งหมายที่กำหนดไว้ในหลักสูตร สามารถกระทำได้หลายรูปแบบ เป็นต้นว่า ใช้ แบบสอบ การสังเกต การตรวจงาน แต่ที่นิยมใช้กันมากในระบบการเรียนการสอน ก็คือการวัด ผลด้วยการใช้แบบสอบ แบบสอบที่ใช้อาจเป็นแบบสอบที่เป็นปรนัย (Objectivity) หรือแบบ สอบที่เป็นอัตนัย (Subjectivity) แต่ละแบบก็มีทั้งจุดเด่นและจุดด้อยที่แตกต่างกันไป แบบสอบ ที่เป็นปรนัย เป็นแบบสอบที่มีความคลาดเคลื่อนน้อยกว่า ส่วนแบบสอบที่เป็นอัตนัย เป็นแบบ สอบที่มีความคลาดเคลื่อนมากกว่า โดยเฉพาะความคลาดเคลื่อน ที่สืบเนื่องมาจากการตรวจให้ คะแนน และการแปลผลคะแนน (อุทุมพร, 2535: 18)

การวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ ด้วยแบบสอบที่เป็นอัตนัย มีข้อจำกัดหลาย ประการในการนำไปใช้ เช่น ผู้ตอบต้องใช้เวลามากในการตอบปัญหาแต่ละข้อ ถามคำถามได้น้อย ข้อ ทำให้การวัดไม่ครอบคลุมในสิ่งที่ต้องการวัด การตอบคำถามแต่ละข้อ การตรวจให้คะแนน นับว่าเป็นปัญหาอย่างยิ่งในเรื่องของความเป็นปรนัย อันจะมีผลให้การสอบขาดความเที่ยงและ ความตรง การตรวจให้คะแนนต้องใช้ผู้ที่มีความรู้ในเนื้อหานั้น ๆ เป็นผู้ตรวจ และต้องใช้เวลา มาก คะแนนที่ได้มักขาดความคงที่แน่นอน

การวัดผลสัมฤทธิ์ทางการเรียนด้วยการใช้แบบสอบที่เป็นปรนัย จึงเป็นที่นิยมใช้มากกว่า แบบสอบประเภทนี้ ให้นักเรียนเขียนตอบน้อยที่สุด ถ้าหากต้องเขียนตอบ ก็ให้เขียนตอบเฉพาะคำ ตอบที่ต้องการ อย่างเฉพาะเจาะจงเท่านั้น แบบสอบจะประกอบด้วยชุดของคำถามหรือข้อความ ที่มุ่งลดความลำเอียงในการสอบ การคุมสอบ การตรวจให้คะแนน อีกทั้งสามารถวัดได้ครอบคลุม เนื้อหา การตรวจให้คะแนนมีความเป็นปรนัยสูง กล่าวคือการตรวจให้คะแนนในลักษณะของผลที่ คงที่ ไม่ว่าใครจะเป็นผู้ตรวจ และทำได้รวดเร็ว แม้ว่าจะมีผู้เข้าสอบมากน้อยเพียงใดก็ตาม ตัวอย่างของแบบสอบแบบเป็นปรนัย ได้แก่ แบบสอบแบบเลือกตอบ แบบสอบแบบถูกผิด และแบบ สอบแบบเติมคำตอบ

แบบสอบแบบเลือกตอบเป็นแบบสอบที่ได้รับความนิยมมาก ทั้งนี้เพราะมีจุดเด่นในการ ตรวจและการให้คะแนน เพราะสามารถวัดเนื้อหาได้ครอบคลุม สะดวกรวดเร็วในการตรวจ และ

ไม่ว่าใครเป็นผู้ตรวจ ก็ให้ผลที่ไม่แตกต่างกัน และยิ่งเหมาะต่อการวิเคราะห์เพื่อนำข้อสอบไปปรับปรุงต่อไป แต่ปัญหาที่พบมากที่สุดของแบบสอบชนิดนี้ ก็คือ การเดา เนื่องจากแบบสอบแบบเลือกตอบมีคำตอบไว้ให้เลือก ผู้ตอบสามารถเลือกคำตอบที่ถูกได้โดยที่ไม่มีความรู้ในเรื่องนั้น อันเนื่องมาจากการเดา ในขณะที่แบบสอบแบบเติมคำตอบก็เป็นแบบสอบแบบปรนัย ที่สามารถจัดการเดาได้ โดยผู้ตอบจะต้องมีความรู้ จึงจะสามารถตอบคำตอบได้ถูกต้อง เนื่องจากไม่มีคำตอบที่ถูกไว้ให้เลือก และจากการศึกษาของ กรองทอง เทพศิริอำนวย (2525) ในเรื่องการเปรียบเทียบคุณภาพของแบบสอบวิชาคณิตศาสตร์ประเภทเลือกตอบและประเภทเติมคำตอบ พบว่า แบบสอบเลือกตอบ ประกอบด้วยข้อสอบที่ง่ายกว่าแบบสอบแบบเติมคำตอบ และมีค่าความเที่ยงที่ต่ำกว่าจากการใช้แบบสอบแบบเติมคำตอบ

แม้ว่าวิธีวัดโดยให้นักเรียนเขียนตอบสั้นเฉพาะคำตอบ หรือเลือกตอบโดยการเลือกจากตัวเลือกที่กำหนดให้ เป็นวิธีที่นิยมนำมาใช้ในการวัดผลสัมฤทธิ์ทางการเรียนในวิชาคณิตศาสตร์กันอย่างกว้างขวางก็ตาม การวัดผลสัมฤทธิ์ทางการเรียนก็ไม่ควรมุ่งที่การใช้แบบสอบเท่านั้น การประเมินผลสัมฤทธิ์ทางการเรียนของนักเรียนโดยผ่านครูผู้สอนวิชาคณิตศาสตร์ ก็น่าจะเป็นวิธีวัดที่ควรนำมาใช้ในการเป็นสารสนเทศ ที่แสดงถึงผลสัมฤทธิ์ทางการเรียนของนักเรียนได้ด้วย ทั้งนี้เนื่องจากครูผู้สอนเป็นผู้ที่มีบทบาทในการเรียนการสอน มีปฏิสัมพันธ์กับนักเรียน ได้สังเกตและติดตามผลการเรียนรวมทั้งผลงานอย่างสม่ำเสมอ การประเมินผลสัมฤทธิ์ของนักเรียนโดยผ่านครูคณิตศาสตร์ ก็น่าที่จะใช้เป็นวิธีวัดผลสัมฤทธิ์ทางการเรียนได้อีกทางหนึ่งด้วย

ดังนั้นในการวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ในระดับประถมศึกษา ควรได้มีการศึกษาว่า วิธีวัดแบบใดที่สามารถจะใช้ร่วมกันได้ดี สำหรับการวัดผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ในระดับประถมศึกษา และสามารถวัดในสิ่งที่ต้องการจะวัดได้หรือไม่ ถ้าคะแนนการสอบที่ได้จากการวัดหนึ่ง ๆ ไม่เปลี่ยนแปลงไป เมื่อไปใช้กับการวัดอีกวิธีหนึ่งแล้ว แสดงว่าวิธีวัดนั้น ๆ ต่างก็สามารถวัดคุณลักษณะเดียวกันได้ ซึ่งกล่าวได้ว่าจะมีความตรงลู่เข้าสูงในการวัดคุณลักษณะเดียวกัน และคุณลักษณะที่ต้องการในที่นี้ก็คือผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์

ตามทฤษฎีการวัดแบบดั้งเดิมนั้น ได้มีการศึกษาในเรื่องของความตรงลู่เข้า โดยใช้การพิจารณาจากสหสัมพันธ์ของผลการสอบด้วยวิธีวัดที่วัดลักษณะเดียวกัน จากวิธีวัดที่แตกต่างกัน ในลักษณะของการพิจารณาสหสัมพันธ์เป็นรายคู่ ซึ่งถ้ามีความตรงลู่เข้าแล้ว จะต้องมีความสัมพันธ์สูง อันแสดงให้เห็นว่าวิธีวัดมีการลู่เข้าหากันในการวัดคุณลักษณะเดียวกัน แต่ถ้าค่าสหสัมพันธ์ต่ำ แสดงว่ายังไม่ยอมให้การสนับสนุนว่ามีความตรงลู่เข้า สำหรับทฤษฎีการอ้างอิงสรุปนั้น ได้ให้

แนวคิดเกี่ยวกับความตรงลู่เข้าในลักษณะของการอ้างอิง กล่าวคือ ถ้าคะแนนสอบนั้นไม่แปรเปลี่ยนไปในระหว่างวิธีวัดที่แตกต่างกันแล้ว คะแนนสอบที่ได้จากการวัดในกลุ่มเงื่อนไขหนึ่งก็จะสามารถใช้ในการอ้างอิงไปสู่ภายใต้กลุ่มเงื่อนไขการวัดอื่น ๆ ได้ (Messick, 1989) นั่นก็คือ คะแนนสอบนั้นไม่เพียงแต่จะเป็นตัวบ่งชี้ที่มีความเชื่อถือได้ของคะแนนแท้จริง ภายในกลุ่มเงื่อนไขการวัดเดียวเท่านั้น (a single set of measurement conditions) แต่เป็นตัวบ่งชี้ของคะแนนเอกภพระหว่างเงื่อนไขการวัดที่แตกต่างกันออกไปด้วย (different measurement conditions) (Suen, 1990: 153) ดังนั้น การสรุปอ้างอิงระหว่างวิธีวัดที่แตกต่างกัน ก็จะเป็นตัวบ่งชี้ความตรงของคะแนนสอบ ค่าสัมประสิทธิ์การอ้างอิงสรุปในที่นี้ จะเป็นสัมประสิทธิ์ความตรง (validity coefficient) นั่นก็คือเป็นการให้ข้อมูลการวัดในลักษณะของค่าเฉลี่ยความตรงลู่เข้าในคู่ของวิธีวัด (Kane, 1980: 37)

นอกจากวิธีวัดจะเป็นองค์ประกอบที่สำคัญประการหนึ่งของความตรงแล้ว จำนวนข้อกระทงก็มีความสำคัญ แต่ถ้าสร้างข้อกระทงเป็นจำนวนมากเกินไป ก็จะทำให้สิ้นเปลืองค่าใช้จ่ายในการดำเนินการวัดรวมทั้งเวลาในการสอบ ถ้ามีจำนวนข้อกระทงน้อยเกินไป อาจวัดได้ไม่ครอบคลุมเนื้อหาที่ต้องการวัด ทำให้การประเมินผลคลาดเคลื่อน เพราะสิ่งที่วัด ไม่เป็นตัวแทนของเนื้อหาที่จะวัด (Guilford, 1954: 407)

วิธีวัดและความยาวของข้อกระทง เป็นองค์ประกอบที่มีความสำคัญต่อความตรงของการวัดผลสัมฤทธิ์ทางการเรียนมากน้อยเพียงใด ความยาวของข้อกระทงและวิธีวัดในลักษณะใด ที่นำไปสู่การเกิดความตรงของการวัดผลสัมฤทธิ์ทางการเรียนได้สูงสุด ในขณะที่ทฤษฎีการอ้างอิงสรุปที่ผ่านมามักเป็นการใช้เพื่อศึกษาความเที่ยง แต่ความเที่ยงและความตรง เป็นคุณสมบัติที่มีความสัมพันธ์ต่อกัน ดังนั้นเมื่อใช้ทฤษฎีการอ้างอิงสรุปในการศึกษาความเที่ยงได้ ก็น่าที่จะนำทฤษฎีการอ้างอิงสรุปมาประยุกต์ใช้ในการศึกษาความตรงได้เช่นกัน ผู้วิจัยจึงเห็นว่าทฤษฎีการอ้างอิงสรุปมีกรอบแนวคิดที่เป็นไปได้และน่าสนใจต่อการศึกษาความตรง และในประเทศไทยก็ยังไม่มีการได้ศึกษาด้านความตรงตามทฤษฎีการอ้างอิงสรุป ดังนั้น เพื่อเป็นการเสริมสร้างความรู้เชิงทฤษฎี และผลประโยชน์ในการนำไปใช้ ผู้วิจัยจึงมีความสนใจที่จะศึกษาวิจัย เพื่อวิเคราะห์ความตรงลู่เข้าของการวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ เมื่อมีวิธีวัดและความยาวของข้อกระทงแตกต่างกัน โดยใช้ทฤษฎีการอ้างอิงสรุปในการศึกษา

วัตถุประสงค์ของการวิจัย

การวิจัยครั้งนี้มีวัตถุประสงค์ทั่วไป เพื่อใช้ทฤษฎีการอ้างอิงสรุป สำหรับการวิเคราะห์ความตรงลู่เข้าของการวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ โดยมีวัตถุประสงค์เฉพาะ ดังนี้

1. เพื่อวิเคราะห์และเปรียบเทียบความตรงลู่เข้า จากสูตรที่พัฒนาตามทฤษฎีการอ้างอิงสรุป ในองค์ประกอบของข้อกระทบ และวิธีวัด
2. เพื่อตรวจสอบความตรงข้ามกลุ่ม (cross-validation) จากสูตรความตรงลู่เข้าที่พัฒนาขึ้นตามทฤษฎีการอ้างอิงสรุป

ขอบเขตของการวิจัย

1. เอกภพของการวิจัย
 - 1.1 นักเรียน เป็นนักเรียนชั้นประถมศึกษาปีที่ 6 สังกัดสำนักงานคณะกรรมการการศึกษาเอกชน กรุงเทพมหานคร ปีการศึกษา 2537 จำนวน 39,836 คน
 - 1.2 วิธีวัด เป็นวิธีที่ใช้ในการวัดผลสัมฤทธิ์ทางการเรียนในวิชาคณิตศาสตร์ ในชั้นประถมศึกษาปีที่ 6 ในที่นี้เป็นการเลือกวิธีวัด 3 วิธี คือ วัดโดยใช้แบบสอบผลสัมฤทธิ์ทางการเรียนแบบเลือกตอบ แบบสอบผลสัมฤทธิ์ทางการเรียนแบบเติมคำตอบ และแบบประเมินผลสัมฤทธิ์ทางการเรียนโดยครูผู้สอนคณิตศาสตร์
 - 1.3 ข้อกระทบแบบเลือกตอบ เป็นคำถามที่วัดความรู้ความสามารถในวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 ตามหลักสูตรของกรมวิชาการ โดยประกอบด้วยข้อความที่เป็นคำถาม และคำตอบให้เลือกจาก 4 ตัวเลือก ซึ่งถือว่าจำนวนข้อกระทบมีจำนวนอนันต์ และลุ่มข้อกระทบให้ครอบคลุมเนื้อหา เพื่อใช้ในการศึกษา 6 จำนวน 30 ข้อ
 - 1.4 ข้อกระทบแบบเติมคำตอบ เป็นคำถามที่วัดความรู้ความสามารถในวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 ตามหลักสูตรของกรมวิชาการ โดยการให้เติมคำตอบลงในที่เว้นว่างไว้ให้ถูกต้อง ซึ่งถือว่าจำนวนข้อกระทบมีจำนวนอนันต์ และลุ่มข้อกระทบให้ครอบคลุมเนื้อหา เพื่อใช้ในการศึกษา 6 จำนวน 30 ข้อ
 - 1.5 ข้อกระทบแบบประเมินค่า เป็นคำถามที่วัดความรู้ความสามารถในวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 ตามหลักสูตรของกรมวิชาการ ให้ครูคณิตศาสตร์เป็นผู้ประเมินผลสัมฤทธิ์ของนักเรียนลงในแบบประเมินค่า ซึ่งถือว่าจำนวนข้อกระทบมีจำนวนอนันต์ และลุ่มข้อกระทบให้ครอบคลุมเนื้อหา เพื่อใช้ในการศึกษา 6 จำนวน 30 ข้อ

2. ทฤษฎีที่ใช้ในการวิเคราะห์ความตรงลู่เข้า

ใช้การประยุกต์ทฤษฎีการอ้างอิงสรุป (Generalizability Theory) สำหรับวิเคราะห์ความตรงลู่เข้า

สมมติฐานของการวิจัย

แบบสอบแบบเลือกตอบเป็นแบบสอบที่สะดวกในการนำไปใช้ สามารถใช้ในการดำเนินการสอบกับผู้สอบได้ครั้งละมาก ๆ การตรวจให้คะแนนก็กระทำได้เสร็จสิ้นภายในระยะเวลาสั้นๆ แต่ก็มีข้อจำกัดในการนำแบบสอบลักษณะนี้ไปใช้ เนื่องจากมีคำตอบให้เลือก โดยคำตอบที่ถูกต้องรวมอยู่ในตัวเลือกที่กำหนดให้ ผู้สอบที่ไม่มีความรู้ในข้อสอบนั้น ๆ ก็ยังมีโอกาสที่จะเดาคำตอบได้ถูกต้องถึง $1/4$ ในกรณีที่ เป็นข้อสอบแบบ 4 ตัวเลือก ทำให้ผู้สอบมีโอกาสได้คะแนนสูงขึ้นกว่าที่ควรจะเป็น ดังนั้น ผู้ที่ไม่มีความรู้ความสามารถที่แท้จริง อาจใช้การเดาจากตัวเลือกที่กำหนดให้ ทำให้ได้คะแนนที่อาจจะมากกว่าหรือน้อยกว่าที่ควรจะได้รับจริง นั่นคือมีความคลาดเคลื่อนในทางเชิงบวก หรือความคลาดเคลื่อนในเชิงลบเกิดขึ้นได้ ในขณะเดียวกัน การเดาก็มีผลต่อความเที่ยง อันเป็นแหล่งที่ทำให้เกิดความคลาดเคลื่อนในการวัด (Magnunson, 1966: 101) สำหรับแบบสอบแบบเติมคำตอบนั้น ลักษณะของข้อสอบจะไม่มีคำตอบที่ถูกไว้ให้ผู้ตอบมีโอกาสเลือก ผู้ตอบต้องคิดหาคำตอบด้วยตนเอง สามารถที่จะนำไปใช้ในการวัดความรู้ความสามารถได้ทุกกระดับการเรียนรู้ เพราะผู้ตอบมีโอกาสในการเดาคำตอบที่ถูกได้น้อยกว่าแบบสอบแบบเลือกตอบ

การที่แบบสอบใดมีโอกาสการเดามาก ความคงเส้นคงวาของคะแนนจากการสอบก็จะน้อย อันส่งผลต่อความเที่ยงด้วย นอกจากนี้ จากการศึกษาของสมหวัง พิธิยานุวัฒน์ และเพ็ญศิริ ด่านชนะ (2524) ได้ศึกษาเปรียบเทียบคุณภาพของแบบสอบ พบว่าแบบสอบแบบตอบสั้นมีแนวโน้มที่จะมีความเที่ยงสูงกว่าแบบสอบแบบเลือกตอบ ทำนองเดียวกับการศึกษาของกรรทอง เทพศิริอำนวย (2525) พบว่าแบบสอบแบบเลือกตอบมีความเที่ยงน้อยกว่าแบบสอบแบบเติมคำตอบแบบสอบทั้ง 2 ชนิดนี้ ต่างก็ใช้เป็นวิธีวัดเพื่อวัดผลสัมฤทธิ์ทางการเรียนกับผู้เรียนได้โดยตรง ในขณะที่การประเมินผลสัมฤทธิ์ทางการเรียน ก็เป็นวิธีวัดผลสัมฤทธิ์ทางการเรียนได้เช่นกัน แต่เป็นลักษณะของการวัดผลสัมฤทธิ์ทางอ้อม (Magnunson, 1966: 97) เป็นวิธีวัดที่ครูเป็นผู้ประเมินผลสัมฤทธิ์ทางการเรียนของนักเรียน วิธีนี้เป็นวิธีวัดอีกแบบหนึ่งที่เป็นลักษณะของการคาดคะเนความรู้ความสามารถของนักเรียน อันเนื่องมาจากครูเป็นผู้ใกล้ชิดกับนักเรียนมาก จึงมีโอกาสที่จะสังเกตพฤติกรรมทางการเรียนของนักเรียนได้อย่างสะดวก เพราะการเรียนการสอนส่วนใหญ่เกิดขึ้นในชั้นเรียน แต่การประเมินผลสัมฤทธิ์โดยครูผู้สอนก็อาจมีความคลาดเคลื่อนเกิดขึ้นได้ ถ้าผู้

ประเมินมีความลำเอียง อาจจะเป็นความคลาดเคลื่อนในลักษณะการประเมินที่สูงหรือต่ำกว่าความเป็นจริงหรือโน้มเข้าสู่ส่วนกลาง นอกจากนี้อาจมีความคลาดเคลื่อน ในลักษณะของผลจากการมีปฏิสัมพันธ์มาก่อนที่เรียกว่า Halo Effect อันเนื่องมาจากความรู้จักเป็นส่วนตัวของผู้ประเมิน โดยอาศัยความประทับใจหรือความรู้จักเป็นส่วนตัวกับผู้ประเมินในการตัดสินที่มีความลำเอียง ทำให้การประเมินมีความตรงน้อย (Guilford, 1954: 279) นั่นคือเมื่อผู้ประเมินทำการประเมินนั้น ลักษณะที่เด่นหรือด้อยบางอย่างในตัวผู้ถูกประเมิน อาจจะไปมีอิทธิพลต่อการประเมินในความรู้ความสามารถที่ต้องการอย่างแท้จริงได้ เช่น นักเรียนที่ครูพิจารณาว่าเป็นผู้ที่มีปัญหาทางด้านความประพฤติ อาจจะได้รับคะแนนจากการประเมินด้านผลสัมฤทธิ์ทางการเรียน ต่ำกว่านักเรียนที่ครูพิจารณาว่าไม่มีปัญหาทางด้านความประพฤติก็ได้ ซึ่งผลการประเมินไม่ได้มาจากความรู้ความสามารถอย่างแท้จริง

วิธีวัดทั้ง 3 วิธีต่างก็มุ่งวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ ดังนั้นการที่ใช้วิธีวัดใด ๆ ร่วมกัน น่าจะมีความตรงสูงเข้า โดยเฉพาะอย่างยิ่งความตรงสูงเข้าระหว่างวิธีวัดด้วยแบบสอบถามเลือกตอบกับแบบสอบถามเติมคำตอบ น่าจะสูงกว่าความตรงสูงเข้าระหว่างวิธีวัดด้วยวิธีวัดอื่น ๆ ร่วมกัน ดังนั้น จากวัตถุประสงค์การวิจัยเพื่อเปรียบเทียบความสามารถในการสรุปอ้างอิงของความตรงสูงเข้าระหว่างวิธีวัดต่าง ๆ ผู้วิจัยจึงกำหนดสมมติฐานการวิจัยต่อไปนี้

สมมติฐานที่ 1 ค่าความตรงสูงเข้าระหว่างวิธีวัดด้วยแบบสอบถามเลือกตอบ กับแบบสอบถามเติมคำตอบ มีค่าสูงกว่าค่าความตรงสูงเข้าระหว่างวิธีวัดด้วยแบบสอบถามเลือกตอบ กับแบบประเมิน

สมมติฐานที่ 2 ค่าความตรงสูงเข้าระหว่างวิธีวัดด้วยแบบสอบถามเลือกตอบ กับแบบสอบถามเติมคำตอบ มีค่าสูงกว่าค่าความตรงสูงเข้าระหว่างวิธีวัดด้วยแบบสอบถามเติมคำตอบ กับแบบประเมิน

สมมติฐานที่ 3 ค่าความตรงสูงเข้าระหว่างวิธีวัดด้วยแบบสอบถามเลือกตอบ แบบสอบถามเติมคำตอบ และแบบประเมิน มีค่าสูงกว่าค่าความตรงสูงเข้าระหว่างวิธีวัดด้วยแบบสอบถามเลือกตอบกับแบบประเมิน

ตามทฤษฎีการอ้างอิงสรุป ซึ่งเป็นโมเดลเชิงลุ่ม โดยใช้การลุ่มนั้น การตรวจสอบทานผลในลักษณะของการตรวจสอบความตรงข้ามกลุ่ม (cross-validation) จากกลุ่มตัวอย่างอีกกลุ่มหนึ่ง ที่มาจากประชากรเดียวกัน น่าจะให้ความตรงสูงเข้าไม่แตกต่างกัน ดังนั้น ผู้วิจัยจึงกำหนดสมมติฐานการวิจัยดังนี้

สมมติฐานที่ 4 ค่าความตรงสูงเข้าที่ได้จากกลุ่มตัวอย่างและกลุ่มตัวอย่างสอบทานผล ไม่มีความแตกต่างกัน

ข้อตกลงเบื้องต้น

1. นักเรียนชั้นประถมศึกษาปีที่ 6 ที่เข้าร่วมในการวิจัย ได้ตั้งใจทำแบบสอบถามอย่างเต็มความสามารถ
2. เพศของนักเรียนที่เข้าร่วมในการวิจัย ไม่มีผลต่อดัชนีความตรง
3. เนื้อหาวิชาคณิตศาสตร์ชั้นประถมศึกษาปีที่ 6 ที่ใช้ในการสอบในภาคปลาย ปีการศึกษา 2537 ถือว่าสามารถสร้างจำนวนข้อกระทงได้จำนวนอนันต์ จึงเป็นตัวแทนเชิงลุ่ม

คำจำกัดความและนิยามศัพท์

1. แบบสอบแบบเลือกตอบ หมายถึง แบบสอบวัดผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ที่ให้นักเรียนเลือกตอบ ในเนื้อหาที่ลุ่มมาจากหลักสูตรคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 ของกรมวิชาการ กระทรวงศึกษาธิการ
2. แบบสอบแบบเติมคำตอบ หมายถึง แบบสอบวัดผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ที่ให้นักเรียนเขียนตอบลงในช่องว่างที่กำหนดให้ ในเนื้อหาที่ลุ่มมาจากหลักสูตรคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 ของกรมวิชาการ กระทรวงศึกษาธิการ
3. แบบประเมินผลสัมฤทธิ์ทางการเรียน หมายถึง แบบประเมินผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ที่ให้ครุคณิตศาสตร์เป็นผู้ประเมินผลนักเรียน ในการวัดความรู้ความสามารถทางวิชาคณิตศาสตร์ ในเนื้อหาที่ลุ่มมาจากหลักสูตรวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 ของกรมวิชาการ กระทรวงศึกษาธิการ
4. นักเรียน หมายถึง นักเรียนชั้นประถมศึกษาปีที่ 6 ประจำปีการศึกษา 2537 ในกรุงเทพมหานคร สังกัดสำนักงานคณะกรรมการการศึกษาเอกชน
5. เงื่อนไขการวัด หมายถึง สภาพการณ์ของการสอบ สำหรับการวิจัยนี้ ได้แก่ ข้อกระทงแต่ละข้อ และวิธีวัดแต่ละวิธีที่ลุ่มมาใช้
6. ฟาเซ็ต หมายถึง องค์ประกอบหรือลักษณะร่วมกัน ของกลุ่มเงื่อนไขการวัดชนิดเดียวกัน ในการวิจัยครั้งนี้ มี 2 ฟาเซ็ต ได้แก่ ข้อกระทง และวิธีวัด
7. ประชากร หมายถึง สิ่งที่มีงวัด ได้แก่ นักเรียนชั้นประถมศึกษาปีที่ 6 จากโรงเรียนสังกัดสำนักงานคณะกรรมการการศึกษาเอกชน ปีการศึกษา 2537 ในกรุงเทพมหานครทุกคน
8. เอกภพ หมายถึง เงื่อนไขที่เป็นไปได้ทั้งหมดในการที่จะวัด สำหรับการวิจัยนี้ ได้แก่ ข้อกระทงที่วัดเกี่ยวกับผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ทุกข้อ และวิธีวัดที่เกี่ยวกับผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ทุกวิธี

9. เอกภพการสังเกตที่ยอมรับได้ หมายถึง กลุ่มเงื่อนไขการวัดที่เป็นไปได้ทั้งหมด ได้แก่ นักเรียนทุกคน ข้อกระทงทุกข้อ และวิธีวัดทุกวิธี ที่อยู่ในเอกภพของแต่ละฟาเซ็ท ซึ่งผู้วิจัยสามารถสังเกตหรือวัดค่าได้

10. เอกภพของการอ้างอิงสรุป หมายถึง กลุ่มเงื่อนไขของฟาเซ็ท ที่ผู้วิจัยต้องการสรุปอ้างอิงผลการวัด สำหรับการวิจัยนี้ ได้แก่ ข้อกระทงและวิธีวัด

11. ความตรงลู่เข้า หมายถึง สัมประสิทธิ์การอ้างอิงสรุปตามทฤษฎีการอ้างอิงสรุป ซึ่งเป็นความตรงลู่เข้าที่ได้จากวิธีวัดมากกว่า 1 วิธีในการวัดคุณลักษณะเดียวกัน สำหรับการวิจัยนี้ จะมีความตรงลู่เข้า 4 ชนิด คือ ความตรงลู่เข้าระหว่างวิธีวัดด้วยแบบสอบผลสัมฤทธิ์ทางการเรียนแบบเลือกตอบกับแบบสอบผลสัมฤทธิ์ทางการเรียนแบบเติมคำตอบ ความตรงลู่เข้าระหว่างวิธีวัดด้วยแบบสอบผลสัมฤทธิ์ทางการเรียนแบบเลือกตอบ กับแบบประเมินผลสัมฤทธิ์ทางการเรียนแบบเติมคำตอบ และความตรงลู่เข้าระหว่างวิธีวัดด้วยแบบสอบผลสัมฤทธิ์ทางการเรียนแบบเลือกตอบ กับแบบประเมินผลสัมฤทธิ์ทางการเรียน และความตรงลู่เข้าระหว่างวิธีวัดด้วยแบบสอบผลสัมฤทธิ์ทางการเรียนแบบเลือกตอบ กับแบบประเมินผลสัมฤทธิ์ทางการเรียน

ประโยชน์ที่คาดว่าจะได้รับ

7.1 ประโยชน์ในเชิงทฤษฎี

ทำให้ได้กระบวนการ แนวคิด และสูตรสำหรับการวิเคราะห์ความตรงลู่เข้าของการวัดผลสัมฤทธิ์ทางการเรียน

7.2 ประโยชน์ในการนำไปใช้

1. ทำให้ได้วิธีการคัดเลือกเครื่องมือในการวัดผลสัมฤทธิ์ทางการเรียน ในลักษณะของความตรงลู่เข้า อันจะนำไปสู่การขยายแนวคิดของการศึกษาความตรง ให้ครอบคลุมความตรงตามโครงสร้าง นอกเหนือจากการพิจารณาโดยเน้นที่ความตรงเชิงเนื้อหาเท่านั้น หากวิธีการคัดเลือกเครื่องมือเพื่อให้เกิดความตรงลู่เข้า เป็นการศึกษาที่ไม่ซับซ้อนต่อการตรวจสอบความตรงแล้ว ก็จะเป็นแนวทางนำไปสู่การศึกษาความตรงเชิงโครงสร้าง

2. เป็นแนวทางในการประยุกต์ทฤษฎีการอ้างอิงสรุป ในการศึกษาปัญหาทางการวัดผลการศึกษาในด้านอื่น ๆ อันจะนำไปสู่การพัฒนาการวัดผลการศึกษาให้มีประสิทธิภาพยิ่งขึ้น