

วรรณคดีที่เกี่ยวข้อง



การนำเสนอวรรณคดีที่เกี่ยวข้องในบทนี้ จัดเนื้อหาเป็น 3 ส่วน คือ ความตรงของเนื้อหาของแบบสอบ ความลำเอียงของเนื้อหาของแบบสอบ และการประเมินผลการเรียนรู้ในโรงเรียนประถมศึกษา

ความตรงของเนื้อหาของแบบสอบ (Content validity)

ความตรงของแบบสอบเป็นคุณลักษณะของแบบสอบที่มีคุณภาพ ซึ่งมีผู้ให้ความหมายของความตรงของแบบสอบไว้หลายความหมาย เช่น

ความตรงของแบบสอบ คือ ความสามารถวัดในสิ่งที่ต้องการวัดได้ถูกต้อง (Lindquist 1942 : 213)

ความตรงของแบบสอบคือ ค่าสหสัมพันธ์ของแบบสอบกับเกณฑ์ (Gulliksen 1950 : 88)

ความตรงของแบบสอบ คือ ความถูกต้องแม่นยำของคะแนนในการพยากรณ์ที่เฉพาะเจาะจง (Stanley and Hopkins 1972 : 101)

จากความหมายของความตรงของแบบสอบดังกล่าว สรุปได้ว่า ความตรงของแบบสอบเป็นคุณลักษณะของแบบสอบที่สามารถวัดสิ่งที่ต้องการวัดได้ตามจุดประสงค์ของการวัด การตรวจสอบความตรงของแบบสอบจึงขึ้นอยู่กับจุดประสงค์ของการวัด ซึ่งอาจจำแนกความตรงของแบบสอบตามธรรมชาติของโครงสร้างการวัด และจุดประสงค์ของการวัด ออกเป็น 3 ประเภท (Brown 1983 : 69; Mehrens and Lehmann 1978 : 110) คือ

1. ความตรงของเนื้อหา (content validity)
2. ความตรงเชิงเกณฑ์สัมพันธ์ (criterion related validity)
3. ความตรงเชิงโครงสร้าง (construct validity)

เนื่องจากการวิจัยครั้งนี้เป็นเรื่องของความตรงของเนื้อหา จึงขอกล่าวถึงความตรงเฉพาะประเภทความตรงของเนื้อหาเท่านั้น

จุดประสงค์ที่สำคัญประการหนึ่งของการใช้แบบสอบถาม คือ เพื่อบรรยาย หรือ วัดระดับความสามารถของบุคคลที่สัมพันธ์กับเนื้อหาที่กำหนดไว้อย่างชัดเจน ในกรณีเช่นนี้ แบบสอบถามที่มีความตรงจะต้องวัดได้ครอบคลุม และเป็นตัวแทนเนื้อหาที่กำหนดไว้นั้น ประเภทของความตรงที่จำเป็น สำหรับแบบสอบถามที่ใช้ในจุดประสงค์ของการใช้แบบสอบถามแบบนี้เป็น คือ ความตรงของเนื้อหา การตรวจสอบความตรงของเนื้อหาของแบบสอบถามจึงเกี่ยวข้องกับ การตรวจสอบว่า ขอบเขตของแบบสอบถามเป็นตัวแทนที่แท้จริงของขอบเขตเนื้อหาที่ต้องการวัด หรือ อ้างอิงไปถึงมากน้อยเพียงใด ซึ่งมีวิธีการตรวจสอบได้หลายวิธี (Brown 1983 : 135 - 137) คือ

1. ใหญ่ทรงคุณวุฒิพิจารณาตัดสินว่าแบบสอบถามฉบับนั้น มีความตรงของเนื้อหาหรือไม่โดยใหญ่ทรงคุณวุฒิประเมินคุณภาพแบบสอบถามว่า ขอบเขตของแบบสอบถามเป็นตัวแทนของขอบเขตเนื้อหาที่ไคเนียมไว้อย่างชัดเจนหรือไม่ วิธีนี้มีข้อบกพร่องคือ (1) วิธีนี้ไม่มีดัชนีของความตรงของเนื้อหาในเชิงปริมาณ (2) ถ้าไม่มีการนิยามขอบเขตเนื้อหาที่ต้องการวัดให้ชัดเจน จะทำให้ใหญ่ทรงคุณวุฒิ ประเมินคุณภาพของแบบสอบถามได้ยาก (3) ผู้ตัดสินแต่ละคนอาจใช้มาตรฐานและเกณฑ์ในการประเมินแตกต่างกัน
2. ใช้การศึกษาความสัมพันธ์ระหว่างคะแนนของคนกลุ่มเดียวกันที่ได้จากแบบสอบถาม 2 ฉบับ ซึ่งสร้างจากขอบเขตเนื้อหา และคำแนะนำในการสร้างเดียวกัน แคนผู้สร้างแบบสอบถามเป็นคนละกลุ่มกัน ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนที่ได้จากแบบสอบถาม 2 ฉบับ มีค่าสูงอาจใช้เป็นข้อมูลสนับสนุนว่าแบบสอบถามมีความตรงของเนื้อหา
3. ใช้วิธีการสอบก่อน และ สอบหลัง (pretest - posttest design) โดยในขั้นแรกนำแบบสอบถามไปใช้กับกลุ่มตัวอย่างที่ไม่มีพื้นความรู้ในขอบเขตเนื้อหาของแบบสอบถาม จากนั้นจึงวัดกระบวนการเรียนการสอนในเนื้อหาเหล่านั้นให้กับกลุ่มตัวอย่าง ในขั้นสุดท้ายจึงใช้แบบสอบถามฉบับเดิมทำการสอบภายหลังอีกครั้งหนึ่ง ถ้าคะแนนที่ได้จากการสอบภายหลัง เพิ่มขึ้นอย่างมีนัยสำคัญทางสถิติ จะเป็นหลักฐานยืนยันว่าเนื้อหาของแบบสอบถามครอบคลุมเนื้อหาที่ต้องการวัด
4. ใช้การศึกษาความสัมพันธ์ระหว่างคะแนนของแบบสอบถามที่ต้องการตรวจสอบความตรงกับคะแนนที่ได้จากแบบสอบถามฉบับอื่น ที่วัดขอบเขตเนื้อหาเดียวกัน ถ้าค่าสัมประสิทธิ์สหสัมพันธ์ใดมีค่าสูง แสดงว่า แบบสอบถามทั้ง 2 ฉบับ วัดเนื้อหาเดียวกัน แต่วิธีนี้มี

ขอบปรองที่ว่า ถึงแม้ว่าแบบสอบจะวัดเนื้อหาเดียวกัน แต่ขอบเขตเนื้อหาที่ต้องการวัด อาจแตกต่างกันได้ นั่นคือ ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนที่ได้จากแบบสอบ 2 ฉบับ ที่มีข้อสมมติว่า วัดขอบเขตเนื้อหาเดียวกันไม่ได้เป็นสิ่งที่ยืนยันว่าแบบสอบมีความ ตรงของเนื้อหา

5. ใหญ่ตรงคุณวุฒิประเมินค่า (rating) คุณลักษณะต่าง ๆ ของแบบ สอบที่เกี่ยวข้องกับความตรงของเนื้อหา เช่น ความครอบคลุมของเนื้อหา การเน้นหรือ การให้ความสำคัญเนื้อหาต่าง ๆ เป็นต้น โดยในการประเมินค่านี้จะต้องมี มาตรฐาน ประเมินค่า (rating scales) ที่เป็นมาตรฐาน เมื่อผู้ทรงคุณวุฒิประเมินค่า คุณลักษณะต่าง ๆ แล้วจึงศึกษาความสอดคล้องกันของผู้ทรงคุณวุฒิ โดยศึกษาค่าความเที่ยง ระหว่างผู้ประเมินค่า (interrater reliability) ถ้าค่าความเที่ยงระหว่าง ผู้ประเมินค่ามีค่าสูงแล้ว มีดัชนีเลขคณิตของการประเมินค่าอาจใช้เป็นดัชนีของความตรง ของเนื้อหา (index of content validity) ได้

ความลำเอียงของเนื้อหาของแบบสอบ (content bias)

ในส่วนนี้กล่าวถึงมโนภาพทั่วไปของความลำเอียงของแบบสอบ ความหมาย ของความลำเอียงของแบบสอบ ความลำเอียงของเนื้อหาของแบบสอบวัดผลสัมฤทธิ์ทาง การเรียน ชนิดของความลำเอียงของเนื้อหา การประมาณค่าความลำเอียงของแบบสอบ สารบบจำแนกเนื้อหา และงานวิจัยที่เกี่ยวข้องกับความลำเอียงของเนื้อหา

1. มโนภาพทั่วไปของความลำเอียงของแบบสอบ

การวัดคุณลักษณะต่าง ๆ ทางการศึกษา และจิตวิทยาของบุคคลไม่สามารถ วัดได้โดยตรง เช่น การวัดคุณลักษณะทางกายภาพ คะแนนที่ได้จากการวัดทางการศึกษา และจิตวิทยาจึงเป็นเพียงค่าประมาณของคะแนนจริงของคุณลักษณะที่ต้องการวัดนั้น คะแนน ที่วัดได้จึงมีความคลาดเคลื่อนไปจากคะแนนจริง ซึ่งความคลาดเคลื่อนที่เกิดขึ้น มี 2 ชนิด คือ ความคลาดเคลื่อนอย่างสุ่ม (random error) และความคลาดเคลื่อนอย่างเป็น ระบบ (systematic error) ดังที่ ไวเลย์ แฮร์เทล และ ฮาร์นีสฟีเกอร์ (Wiley, Haertel and Harnischfeger 1981, quoted in Schmidt 1983 : 165)

ไค้แสดงความแตกต่างระหว่างคะแนนจริงของบุคคล และคะแนนที่วัดไค้ด้วยเครื่องมือวัด Y ในรูปของความคลาดเคลื่อน ดังนี้

$$(Y_i - Z_i) = (Y_i - E(Y_i)) + (E(Y_i) - Z_i) \quad (1)$$

เมื่อ Y_i คือ คะแนนที่วัดไค้ของบุคคลที่ i
 Z_i คือ คะแนนจริงที่ต้องการประมาณค่าของบุคคลที่ i
 $E(Y_i)$ คือ ค่าคาดหวังของคะแนนที่วัดไค้ของบุคคลที่ i

จากสมการที่ (1) นิพจน์ $(Y_i - Z_i)$ เป็นความแตกต่างระหว่างคะแนนที่วัดไค้ด้วยเครื่องมือ Y และคะแนนจริงที่ต้องการประมาณค่าของบุคคลที่ i ซึ่งมีความแตกต่างนี้แสดงถึงความคลาดเคลื่อนในการประมาณค่าคะแนนจริง จากการไค้เครื่องมือ Y

นิพจน์ $(Y_i - E(Y_i))$ เป็นความแตกต่างระหว่างคะแนนที่วัดไค้ด้วยเครื่องมือ Y และค่าความคาดหวังของคะแนนที่วัดไค้ของบุคคลที่ i ที่วัดไค้ด้วยเครื่องมือ Y ดังนั้น $(Y_i - E(Y_i))$ จึงเป็นคะแนนของความคลาดเคลื่อนในการวัดไค้ด้วยเครื่องมือ Y ซึ่งเป็นความคลาดเคลื่อนอย่างสุ่ม ซึ่งตามทฤษฎีคลาสสิกของการวัดผลแล้ว

ความคลาดเคลื่อน $(Y_i - E(Y_i))$ เป็นตัวบ่งชี้อย่างฉงนของความเที่ยงของการวัด

นิพจน์ $(E(Y_i) - Z_i)$ เป็นความแตกต่างระหว่างค่าคาดหวังของคะแนนที่วัดไค้ด้วยเครื่องมือ Y และคะแนนจริงของบุคคลที่ i ความแตกต่าง $(E(Y_i) - Z_i)$ นี้ เป็นความคลาดเคลื่อนอย่างเป็นระบบที่แสดงให้เห็นว่าคะแนนที่ไค้จากการวัดเป็นตัว

ประมาณค่าที่ลำเอียง (bias estimator) ของคะแนนจริง ซึ่ง ไวเลย์ และคณะ

เรียกความแตกต่าง $(E(Y_i) - Z_i)$ นี้ว่า ความไม่ตรง (invalidity) หรือ

ความลำเอียง (bias) ของการวัด ถ้าขนาดของความแตกต่าง $(E(Y_i) - Z_i)$

มีค่าเป็นศูนย์ กล่าวไค้ว่าการวัดไค้โดยเครื่องมือ Y หรือคะแนนที่ไค้จากการวัดปราศจาก

ความลำเอียง นั่นคือการวัดไค้มีความตรง แต่ขนาดของความแตกต่าง $(E(Y_i) - Z_i)$

มีค่ายิ่งห่างจากศูนย์มากเท่าไค้ การวัดไค้ด้วยเครื่องมือ Y หรือคะแนนที่ไค้จากการวัด

ควยเครื่องมือ Y ยิ่งมีความลำเอียง หรือยิ่งมีความไม่ตรงมากขึ้นเท่านั้น

2. ความหมายของความลำเอียงของแบบสอบถาม

จัสแมน (Justman 1971, in Gronlund 1968 : 289) ได้ให้ความหมายของความลำเอียงของแบบสอบถามว่า ความลำเอียงของแบบสอบถาม หมายถึง อิทธิพลขององค์ประกอบที่ไม่เกี่ยวข้องกับจุดประสงค์ของการวัดที่มีต่อคะแนนที่ได้จากแบบสอบถาม

เจนเซน (Jensent 1980 : 375) ได้กล่าวถึงความหมายของความลำเอียงไว้ 2 นัย คือ (1) ในทางสถิติความลำเอียงหมายถึง การที่ค่าสถิติที่ใช้เป็นตัวประมาณค่า ($\hat{\theta}$) ของพารามิเตอร์ (θ) มีค่าความคาดหวังที่ต่างจากค่าของพารามิเตอร์ (นั่นคือ $E(\hat{\theta}) \neq \theta$) (2) ในทางจิตวิทยาความลำเอียงของการวัด หมายถึง ความคลาดเคลื่อนอย่างเป็นระบบในความตรงเชิงพยากรณ์ และความตรงเชิงโครงสร้างของคะแนนของบุคคล ซึ่งความคลาดเคลื่อนนี้สัมพันธ์กับความเป็นสมาชิกของกลุ่มบุคคล

จากความหมายของความลำเอียงดังกล่าว เมื่อพิจารณาความลำเอียงของแบบสอบถามจึงน่าจะหมายถึง สภาพการณ์ที่คะแนนที่ได้จากการใช้แบบสอบถามวัดความสามารถหรือ คุณลักษณะของบุคคล เป็นตัวประมาณค่าที่ไม่ถูกต้องของคะแนนจริงของบุคคล

3. ความลำเอียงของเนื้อหาของแบบสอบถามวัดผลสัมฤทธิ์ทางการเรียน

จุดประสงค์ประการหนึ่งของการใช้แบบสอบถามวัดผลสัมฤทธิ์ทางการเรียน คือ การวัดความสามารถของบุคคลเพื่อตรวจสอบผลสัมฤทธิ์ หรือผลการเรียนรู้ในสิ่งที่ได้รับการเรียนการสอน สิ่งสำคัญที่แบบสอบถามวัดผลสัมฤทธิ์ทางการเรียนต้องวัดจึงควรเป็นขอบเขตเนื้อหาที่กำหนดไว้อย่างแน่นอน ฉะนั้นความตรงที่แบบสอบถามวัดผลสัมฤทธิ์ทางการเรียนจำเป็นต้องมี คือ ความตรงของเนื้อหาซึ่งเป็นความตรงที่เกี่ยวข้องกับสภาพความเป็นตัวแทนของเนื้อหาของแบบสอบถามขอบเขตเนื้อหาที่ตรงการวัดนั้น ในกรณีสิ่งที่ต้องการวัดอ้างอิงอยู่กับขอบเขตเนื้อหาที่แน่นอน จะนิยามคะแนนจริงของบุคคล (Z_i) ได้ว่า เป็นความรู้ความสามารถของบุคคลในขอบเขตเนื้อหาที่ต้องการวัดนั้น และค่าคาดหวังของคะแนนที่วัดได้ ($E(Y_i)$) คือพารามิเตอร์ที่อธิบายค่าเฉลี่ยของการแจกแจงทางทฤษฎี ซึ่งเป็นผลมาจากการสอบโดยวิธีสอบจำนวนไม่จำกัด ถ้าเนื้อหาในแบบสอบถามไม่เป็น

ตัวแทนของขอบเขตเนื้อหาที่ต้องการวัด [อาจเนื่องมาจากข้อกระทงที่สร้างขึ้นวัดไม่ตรงกับเนื้อหาและพฤติกรรมที่ต้องการวัด และ/หรือ นำหนักความสำคัญของเนื้อหาอยู่ในแบบสอบไม่ตรงตามที่ควรจะเป็น] จะทำให้เกิดความแตกต่างระหว่างค่าคาดหวังของคะแนนที่วัดได้ และคะแนนจริงของแต่ละบุคคล $(E(Y_i) - Z_i)$ ซึ่งความแตกต่างนี้จะแสดงถึงความลำเอียงของแบบสอบอันเนื่องมาจากเนื้อหา (Schmidt 1983 : 166) จึงอาจกล่าวได้ว่าความลำเอียงของเนื้อหาของแบบสอบเป็นสภาพที่เนื้อหาของแบบสอบไม่เป็นตัวแทนที่ถูกต้องของเนื้อหาที่ต้องการวัด ซึ่งสภาพการณ์นี้ทำให้คะแนนที่ได้จากการวัดเป็นตัวประมาณค่าที่ลำเอียงของคะแนนจริงในขอบเขตเนื้อหาที่ต้องการวัด (นั่นคือเมื่อ $E(Y_i) \neq Z_i$ แบบสอบที่ใช้จะเป็นแบบสอบที่ลำเอียงในเนื้อหา ฉะนั้นความลำเอียงเท่ากับ $E(Y_i) - Z_i$)

4. ชนิดของความลำเอียงของเนื้อหา

ในการประมาณค่าความลำเอียงของเนื้อหาของแบบสอบจำเป็นต้องทราบว่าการสอบครั้งนั้น มีจุดประสงค์ที่จะประมาณค่าคะแนนจริงของบุคคล สัมพันธ์กับเนื้อหาใด ซึ่ง ชมิตท์ (Schmidt 1983 : 167 - 168) ได้จำแนกขอบเขตเนื้อหาที่ต้องการวัดไว้ 3 ชนิด คือ

1. บุพปริเขต (a priori domain) เป็นขอบเขตเนื้อหาที่นักเรียนทุกคนควรจะได้เรียนรู้ เช่น ขอบเขตเนื้อหาที่ใช้ในการประเมินมาตรฐานชาติ
2. ปริเขตเนื้อหาตามหลักสูตร (curricular domains) เป็นขอบเขตเนื้อหาที่กำหนดโดยจุดมุ่งหมายของหลักสูตร หรือเนื้อหาสาระของเอกสารหลักสูตร เช่น โปรแกรมการศึกษา และแบบเรียนเป็นต้น เช่น ขอบเขตเนื้อหาที่ใช้ในการประเมินระดับท้องถิ่น หรือ ระดับโรงเรียน
3. ปริเขตเนื้อหาตามการสอน (instructional domains) เป็นขอบเขตเนื้อหาที่สอนจริงในชั้นเรียน เป็นขอบเขตเนื้อหาที่จำเพาะเจาะจงในแต่ละชั้นเรียน และครูแต่ละคนเท่านั้น

จากการที่ขอบเขตเนื้อหาที่แบบสอบวัดผลสัมฤทธิ์ของการวัด อาจจำแนกได้เป็น 3 ชนิด ดังกล่าวแล้ว การศึกษาความลำเอียงของเนื้อหาของแบบสอบ จึงอาจอ้างอิง

กับขอบเขตเนื้อหาใดขอบเขตเนื้อหาหนึ่ง ความลำเอียงของเนื้อหาของแบบสอบ จึงมี 3 ชนิด ตามขอบเขตเนื้อหาที่ต้องการวัด คือ (1) ความลำเอียงของเนื้อหาของแบบสอบจากบุปรีเขต (2) ความลำเอียงของเนื้อหาของแบบสอบจากปรีเขตเนื้อหาตามหลักสูตร และ (3) ความลำเอียงของเนื้อหาของแบบสอบจากปรีเขตเนื้อหาตามที่สอน

5. การประมาณค่าความลำเอียงของเนื้อหาของแบบสอบ

จากสมการ (1) ความลำเอียงของเนื้อหาของแบบสอบวัดได้จากความแตกต่าง $(E(Y_i) - Z_i)$ แต่ในความเป็นจริงแล้วไม่สามารถทราบค่าของ Z_i ได้ จึงไม่สามารถวัดความลำเอียงของเนื้อหาของแบบสอบได้จากความแตกต่าง $(E(Y_i) - Z_i)$ โดยตรง สมิตท์ (Schmidt 1983 : 171) ได้กล่าวว่า "ในการนิยาม Z_i และ $E(Y_i)$ จำเป็นต้องอาศัยสารบบจำแนกเนื้อหา (content taxonomy) ที่ครอบคลุมขอบเขตเนื้อหาที่ต้องการวัด" สารบบจำแนกเนื้อหานี้จะประกอบด้วยเซลล์ต่าง ๆ ซึ่งแต่ละเซลล์จะแทนเนื้อหาย่อยที่ประกอบกันขึ้นเป็นขอบเขตเนื้อหาที่ต้องการวัด การประกอบกันเป็นขอบเขตเนื้อหาที่ต้องการวัดของเนื้อหาย่อย ๆ หรือเซลล์ย่อย ๆ เหล่านี้ จะประกอบกันขึ้นด้วยน้ำหนักของแต่ละเนื้อหาย่อยที่ต่างกันตามความสำคัญของเนื้อหาย่อยนั้น ซึ่งน้ำหนักของแต่ละเนื้อหาย่อยหรือแต่ละเซลล์จะแทนความลำเอียงของคะแนนจริงของแต่ละเนื้อหาย่อย (Z_{ij}) ที่เป็นส่วนประกอบในคะแนนจริงของขอบเขตเนื้อหาที่ต้องการวัด (Z_i) ดังนั้นคะแนนจริงของขอบเขตเนื้อหาที่ต้องการวัดจึงแสดงได้ในรูปแบบการรวมเชิงเส้นตรงของคะแนนจริงของเนื้อหาย่อย ๆ ตามน้ำหนักที่เหมาะสม

จึงสมการ (2)

$$Z_i = \sum^J W_j^D Z_{ij} \quad (2)$$

เมื่อ Z_i คือ คะแนนจริงของบุคคลที่ i

W_j^D คือ น้ำหนักของเนื้อหาย่อยที่ j ตามที่ควรวัด (j th domain)

Z_{ij} คือ คะแนนจริงของเนื้อหาย่อยที่ j ของบุคคลที่ i

J คือ จำนวนเนื้อหาย่อยทั้งหมด

ดังนั้นถ้าสร้างแบบสอบถาม Y โดยให้มีขอบเขตเนื้อหาครอบคลุมตามสารบบจำแนกเนื้อหาขอบเขตเนื้อหาของแบบสอบถามจะประกอบด้วยเนื้อหาของ J ของสารบบจำแนกเนื้อหาตามน้ำหนักความสำคัญของแต่ละเนื้อหาของที่ปรากฏในแบบสอบถาม (จำแนกข้อกระทงในแต่ละเนื้อหาของจะแทนน้ำหนักของเนื้อหาของนั้น) ดังนั้นค่าคาดหวังของคะแนนที่วัดได้โดยแบบสอบถาม Y แสดงได้ในรูปการรวมเชิงเส้นตรงของค่าคาดหวังของคะแนนที่วัดได้ในแต่ละเนื้อหา ตามน้ำหนักที่ปรากฏในแบบสอบถาม ดังสมการ (3)

$$E(Y_i) = \sum^J W_j^T E(Y_{ij}) \quad (3)$$

เมื่อ $E(Y_i)$ คือ ค่าคาดหวังของคะแนนที่วัดได้ของบุคคลที่ i

W_j^T คือ น้ำหนักของเนื้อหาของที่ j ที่ปรากฏในแบบสอบถาม

$E(Y_{ij})$ คือ ค่าคาดหวังของคะแนนที่วัดได้ในเนื้อหาของที่ j

จากแนวคิดของ ไวลีย์ และคณะ ที่ว่าความลำเอียงของแบบสอบถามอธิบายได้ด้วยความแตกต่าง $(E(Y_i) - Z_i)$ เมื่อแทนค่า Z_i และ $E(Y_i)$ จากสมการ (2) และ (3) จะได้

$$E(Y_i) - Z_i = \sum^J (W_j^T E(Y_{ij}) - W_j^D Z_{ij}) \quad (4)$$

แต่ถ้าสารบบจำแนกเนื้อหาที่มีความครอบคลุมและชัดเจนและข้อกระทงที่สร้างขึ้นเพื่อวัดคะแนนจริงในแต่ละเนื้อหาของมีคุณภาพดี และวัดได้ตรงกับเนื้อหา และพฤติกรรมในแต่ละเนื้อหาของแล้ว ค่าคาดหวังของคะแนนที่วัดได้ในเนื้อหาของที่ j ของแต่ละบุคคล ($E(Y_{ij})$) จะเท่ากับคะแนนจริงของเนื้อหาของที่ j ของบุคคลนั้น (Z_{ij}) ดังนั้น

$$E(Y_i) - Z_i = \sum^J Z_{ij} (W_j^T - W_j^D) \quad (5)$$

จากสมการ (5) จะเห็นว่าความลำเอียงของเนื้อหาที่มีส่วนประกอบ 2 ส่วน คือ (1) คะแนนจริงของแต่ละเนื้อหาของของบุคคล (Z_{ij}) และ (2) ความแตกต่างระหว่างน้ำหนักของเนื้อหาที่ปรากฏในแบบสอบถาม และน้ำหนักของเนื้อหาที่ต้องการวัด ในแต่ละเนื้อหาของ ซึ่ง ชมิดท์ (Schmidt 1983 : 171) ได้กล่าวว่า

"ถ้าระบบจำแนกเนื้อหาที่มีความสมบูรณ์พอที่จะแทนเนื้อหาของแบบสอบ และเนื้อหาที่ต้องการวัดแล้ว คะแนนจริงของแต่ละเนื้อหาย่อยของบุคคล (Z_{ij}) จะไม่เปลี่ยนแปลง แต่น้ำหนักของแต่ละเนื้อหาย่อย (W_j^T และ W_j^D) อาจเปลี่ยนแปลงได้" ดังนั้นถ้าระบบจำแนกเนื้อหาที่ใช้ในการกำหนดหรือจำแนกขอบเขตเนื้อหาที่ต้องการวัดหรือขอบเขตเนื้อหาของแบบสอบ มีความชัดเจนและครอบคลุม และขอกระทั่งสร้างขึ้นมีคุณภาพและวัดได้ตรงกัเนื้อหาและพฤติกรรมในแต่ละเนื้อหาย่อยแล้วความลำเอียงของเนื้อหาของแบบสอบจึงเป็นฟังก์ชัน ของน้ำหนักเท่านั้น ความลำเอียงของเนื้อหาของแบบสอบจึงประมาณได้ด้วยความแตกต่างระหว่างน้ำหนักของเนื้อหาที่ปรากฏในแบบสอบและน้ำหนักของเนื้อหาที่ต้องการวัด ดังสมการ (6)

$$b_j = W_j^T - W_j^D \quad (6)$$

เมื่อ b_j คือ คำนวณวัดความลำเอียงของเนื้อหาของแบบสอบในเนื้อหาย่อยที่ j การแปลความหมายของค่าคณันวัดความลำเอียงของเนื้อหาในเนื้อหาย่อยที่ j มีดังนี้

ถ้าค่าคณันวัดความลำเอียงของเนื้อหาในเนื้อหาย่อยใดเป็นลบ แสดงว่าแบบสอบมีเนื้อหาที่วัดในเนื้อหาอยุ่้นน้อยกว่าที่ควรวัด ซึ่งกล่าวไว้ในเนื้อหาอยุ่้นแบบสอบมีความลำเอียงของเนื้อหาขาดไป หรือมากไป

ถ้าค่าคณันวัดความลำเอียงของเนื้อหาในเนื้อหาย่อยใดเป็นบวก แสดงว่าแบบสอบมีเนื้อหาที่วัดในเนื้อหาอยุ่้นมากกว่าที่ควรวัด ซึ่งกล่าวไว้ในเนื้อหาอยุ่้นแบบสอบมีความลำเอียงของเนื้อหาแบบเกินไป หรือมากไป

ถ้าค่าคณันวัดความลำเอียงของเนื้อหาในเนื้อหาย่อยใดเป็นศูนย์ แสดงว่าแบบสอบมีเนื้อหาที่วัดในเนื้อหาอยุ่้นตรงตามที่ควรวัด ซึ่งกล่าวไว้ในเนื้อหาอยุ่้นแบบสอบไม่มีความลำเอียงของเนื้อหา

เป็นที่น่าสังเกตว่าความไม่ตรง หรือความลำเอียงในการวัดเนื้อหาของเครื่องมือหนึ่ง ๆ นั้น $(E(Y_j) - Z_j)$ เป็นผลรวมของความไม่ตรงหรือความลำเอียงในการวัดเนื้อหาอยุ่้น ๆ โดยเครื่องมือหนึ่ง

จากที่กล่าวมาแล้วว่า ความตรงของแบบสอบเป็นคุณลักษณะของแบบสอบที่สามารถวัดสิ่งที่ต้องการวัดได้ตามจุดประสงค์ของการวัด นั่นคือ ถ้าใช้แบบสอบที่มีความตรงวัดความสามารถหรือคุณลักษณะของบุคคลแล้ว คะแนนที่ได้จากการวัดจะมีความถูกต้องในการประมาณคะแนนจริงของบุคคลนั้น ซึ่งในทางสถิติจะเรียกได้ว่า คะแนนที่ได้จากการวัดเป็นตัวประมาณค่าที่ไม่ลำเอียงของคะแนนจริง ในทางตรงกันข้าม ถ้าใช้แบบสอบที่ไม่มีความตรง คะแนนที่ได้จะผิดพลาดไปจากคะแนนจริง ซึ่งเรียกได้ว่าคะแนนที่ได้จากการวัดเป็นตัวประมาณค่า ที่ลำเอียงของคะแนนจริง การที่คะแนนที่ได้จากการวัดเป็นตัวประมาณค่าที่ลำเอียงของคะแนนจริง ตามแนวคิดของ ซมิคท์ กล่าวได้ว่า แบบสอบมีความลำเอียง จะเห็นได้ว่าความลำเอียงของแบบสอบมีความสัมพันธ์กับความตรงของแบบสอบในเชิงบวก นั่นคือ ถ้าระดับความลำเอียงของแบบสอบมีค่ามากขึ้น ระดับความตรงของแบบสอบจะลดลง หรือ อาจกล่าวได้ว่า ความลำเอียง คือ ความไม่ตรงของแบบสอบ ดังนั้น วิธีหนึ่งในการตรวจสอบความตรงของ เนื้อหาของแบบสอบกระทำได้โดยการประมาณค่าความลำเอียงของเนื้อหาของแบบสอบ นั่นคือ ถ้าพบว่า ความลำเอียงของเนื้อหาของแบบสอบที่ศึกษามีค่าสูง จะเป็นการบ่งชี้ว่า ความตรงของเนื้อหาของแบบสอบต่ำ ในทางตรงกันข้าม ถ้าพบว่า ความลำเอียงของเนื้อหาของแบบสอบมีค่าต่ำ จะเป็นการบ่งชี้ว่า ความตรงของเนื้อหาของแบบสอบมีค่าสูง

6. สารบบจำแนกเนื้อหา (taxonomy)

ดังที่กล่าวแล้วว่า ในการประมาณค่าความลำเอียงของเนื้อหาจำเป็นต้องใช้สารบบจำแนกเนื้อหา เพื่อจำแนก หรือ กำหนดเนื้อหาของแบบสอบและเนื้อหาที่ต้องการวัด ซึ่งตามปกติแล้ว เนื้อหาของรายวิชา หรือ หลักสูตร มีส่วนประกอบที่สำคัญ 2 ส่วน คือ เนื้อหาวิชา (subject - matter content) ซึ่งเป็นหัวข้อที่รายวิชา หรือหลักสูตรครอบคลุมอยู่ และส่วนที่เป็นจุดประสงค์ในการสอน (instructional objectives) ซึ่งเป็นพฤติกรรมที่ต้องการให้มีการเปลี่ยนแปลงในตัวนักเรียน (Gronlund 1971 : 77) ฉะนั้นโดยทั่วไปสารบบจำแนกเนื้อหาจะมีลักษณะเป็นตาราง 2 มิติ คือ มิติเนื้อหาวิชา และ มิติจุดประสงค์ หรือพฤติกรรม ในแต่ละมิติจะมีรายละเอียดย่อยลงไป รายละเอียดในมิติเนื้อหาวิชามักจะเป็นหัวข้อสำคัญของเนื้อหาวิชา ซึ่งได้จากการวิเคราะห์เนื้อหาวิชาของหลักสูตร ส่วนรายละเอียดในมิติจุดประสงค์ หรือ

พฤติกรรมนั้นมักจะกำหนดเป็นระดับชั้นต่าง ๆ ของพฤติกรรมการเรียนรู้ ซึ่งจะแตกต่างกันไปในแต่ละรายวิชา ระดับชั้น และแนวความคิดของนักการศึกษา สำหรับวิชาคณิตศาสตร์นั้น วิลสัน (Wilson, in Bloom, Hasting and Madaus 1971 : 645 - 648) ได้จำแนกจุดมุ่งหมายในการสอนคณิตศาสตร์ระดับมัธยมศึกษา คำนวณสัญญา เป็น 4 ชั้น คือ การคำนวณ (computation) ความเข้าใจ (comprehension) การนำไปใช้ (application) และการวิเคราะห์ (analysis)

สำหรับพฤติกรรมการวัดในวิชาคณิตศาสตร์ ตามหลักสูตรประถมศึกษา พุทธศักราช 2521 ได้แบ่งพฤติกรรมการวัดเป็น 3 พฤติกรรม (กรมวิชาการ 2521 : 10 - 13) คือ

1. ทักษะ ได้แก่ทักษะต่าง ๆ ของสิ่งที่เรียนโดยเน้นที่ความถูกต้องแม่นยำ รวดเร็ว
2. ความเข้าใจ ได้แก่การเกิดความคิดรวบยอดของสิ่งต่าง ๆ
3. การแก้ไขข้อปัญหา ได้แก่ ความสามารถในการแก้ไขข้อปัญหาเกี่ยวกับเรื่องต่าง ๆ โดยเน้นในเรื่องกระบวนการต่อเนื่อง การนำความเข้าใจและทักษะต่าง ๆ มาใช้จนได้คำตอบที่ถูกต้อง

คูส และคณะ (Kuks, et al. 1979, quoted in Freeman, et al. 1979 : 3 - 4) ได้พัฒนาสารบบจำแนกเนื้อหาวิชาคณิตศาสตร์ ระดับประถมศึกษา ซึ่งมีลักษณะเป็นตาราง 3 มิติ คือ

1. มิตินำเสนอปัญหา (mode of presentation) เป็นมิติที่จะจำแนกว่าข้อกระทงหนึ่ง ๆ นั้นถามคำถามอย่างไร แบ่งออกเป็น 3 ด้าน คือ (1) แผนภูมิ รูปภาพ ตาราง หรือวัสดุ (2) การกระทำที่เฉพาะเจาะจง และ (3) การกระทำที่ไม่เฉพาะเจาะจง (โจทย์ปัญหา)
2. มิติธรรมชาติของเนื้อหาสาระ (nature of material) เป็นมิติที่จะจำแนกว่าข้อกระทงหนึ่ง ๆ นั้น ใช้จำนวนหรือพจน์ทางคณิตศาสตร์อย่างไร แบ่งออกเป็น 13 ด้าน เช่น ทศนิยม ประโยคสัญลักษณ์และรูปเรขาคณิต เป็นต้น

3. มติการจักกระทำ (operation) เป็นมติที่จะจำแนกว่าข้อกระทง
หนึ่ง ๆ นั้นใช้การกระทำทางคณิตศาสตร์อย่างไร แบ่งออกเป็น
12 คำน เช่น การบวก การคูณ และการจัดหมู่ เป็นต้น

มติดัง ๆ ทั้ง 3 มตินี้จะทำให้เกิดเขตต่าง ๆ 468 เขต เขตเหล่านี้ จะแทนเนื้อหาที่
เฉพาะเจาะจงที่ประกอบกันขึ้นเป็นขอบเขตเนื้อหาทั้งหมด

ชมิทท์ (Schmidt 1983 : 173) ได้กล่าวถึงสารบบจำแนกเนื้อหาวิชา
คณิตศาสตร์ ที่มี 3 มติ คล้ายกับสารบบจำแนกเนื้อหาที่ คูส และคณะ พัฒนาขึ้น แต่มีราย
ละเอียดบางประการแตกต่างกันออกไป มติทั้ง 3 มติ ของสารบบจำแนกเนื้อหาที่ ชมิทท์
กล่าวถึง คือ

1. มติจุดประสงค์ทั่วไปของบทเรียน (general intent of the lesson) แบ่งออกเป็น 3 คำน คือ (1) ความเข้าใจในภาพ (2) ทักษะ และ (3) การนำไปใช้
2. มติธรรมชาติของเนื้อหา (nature of material) แบ่งออกเป็น 14 คำน คล้ายกับสารบบจำแนกเนื้อหาที่ คูส และคณะ พัฒนาขึ้น
3. มติการจักกระทำ (operation) แบ่งออกเป็น 15 คำน คล้ายกับสารบบจำแนกเนื้อหาที่ คูส และคณะ พัฒนาขึ้น

จะเห็นได้ว่าสารบบจำแนกเนื้อหาที่ คูส และคณะ พัฒนาขึ้น และสารบบจำแนก
เนื้อหาที่ ชมิทท์ กล่าวถึง มีความคล้ายคลึงกันในมิติธรรมชาติของเนื้อหา และมติการจัก
กระทำ แต่มีความแตกต่างกันที่สารบบจำแนกเนื้อหาที่ คูส และคณะ พัฒนาขึ้น มีมิติวิธีนำ
เสนอปัญหา ซึ่งมีลักษณะเป็นพฤติกรรมการเรียนรู้ และลักษณะของคำถามที่เสนอให้นักเรียน
ประกอบกัน แต่สารบบจำแนกเนื้อหาที่ ชมิทท์ กล่าวถึง มีมิติจุดมุ่งหมายทั่วไปของบทเรียน
ซึ่งมีลักษณะเป็นพฤติกรรมการเรียนรู้ที่ต่อการวัด เมื่อพิจารณามิติธรรมชาติของเนื้อหา และ
มติการจักกระทำทั้งในสารบบจำแนกเนื้อหาที่ คูส และคณะ พัฒนาขึ้น และสารบบจำแนก
เนื้อหาที่ ชมิทท์ กล่าวถึงจะเห็นได้ว่ามิตินี้ 2 มติ นี้ จะประกอบกันเป็นเนื้อหาสาระของ
รายวิชา ฉะนั้นโดยสรุปแล้วสารบบจำแนกเนื้อหาเป็นตารางหลายมิติที่ใช้เพื่อจำแนกเนื้อหา
ของแบบสอบ หรือกำหนดเนื้อหาที่ต้องการวัดสารบบจำแนกเนื้อหานี้ มีส่วนประกอบที่สำคัญ
2 ส่วน คือ เนื้อหาของรายวิชา และจุดประสงค์ในการสอน หรือพฤติกรรมที่ต้องการวัด

ซึ่งแต่ละส่วนสามารถแยกแยะรายละเอียดย่อย ๆ ลงไปได้อีก

7. งานวิจัยที่เกี่ยวข้องกับความลำเอียงของเนื้อหา

ฟรีแมน และคณะ (Freeman, et al. 1979 : 5) ได้เปรียบเทียบ เนื้อหาของแบบสอบมาตรฐาน ระดับ 4 จำนวน 4 ฉบับ คือ เทอะ สแตนฟอร์ด อะชีฟเมนต์ เทส (the Stanford Achievement test : SAT) เทอะ ไอโวา เทส ออฟ เบสิค ซคิล (the Iowa test of Basic Skills : Iowa) เทอะ เมโทรโพลิตันอะชีฟเมนต์ เทส (the Metropolitan Achievement test : MAT) และ เทอะ คอมพรีเฮนซีฟ เทส ออฟ เบสิค ซคิล (the Comprehensive test of Basic Skills : CTBS) โดยจำแนกเนื้อหาของแบบสอบทั้ง 4 ฉบับ ตามสารบบจำแนกเนื้อหาที่พัฒนาขึ้น พบว่าแบบสอบมาตรฐานทั้ง 4 ฉบับ มีจำนวนข้อ กระทบที่วัดในแต่ละเนื้อหาแตกต่างกัน เช่น ในเนื้อหาเรื่องการคูณจำนวนเฉพาะ แบบสอบ CTBS มีจำนวนข้อกระทบที่วัดในเนื้อหานี้มากที่สุด ในจำนวนแบบสอบทั้ง 4 ฉบับ คือ มี จำนวนข้อกระทบร้อยละ 13.3 ของจำนวนข้อกระทบทั้งหมด รองลงไปคือแบบสอบ SAT (ร้อยละ 12.5) แบบสอบ MAT (ร้อยละ 7.8) และแบบสอบ Iowa (ร้อยละ 5.2) จากความแตกต่างกันของเนื้อหาของแบบสอบทั้ง 4 ฉบับนี้ แสดงให้เห็นว่าผู้สร้างแบบสอบที่ต่างกันมีความคาดหวังต่อความรู้และระดับการเรียนรู้ที่นักเรียนได้เรียนรู้ต่างกัน นอกจากนี้จากการศึกษาของ ฟรีแมน และคณะ ยังพบว่า แบบสอบมาตรฐานทุกฉบับที่ศึกษา มีจำนวนข้อกระทบฉบับละ 1 ข้อ ที่วัดเนื้อหาเกี่ยวกับการหารไม่ลงตัว แต่ในความเป็นจริงแล้ว แบบเรียนคณิตศาสตร์ระดับ 4 มีแบบฝึกหัดในเนื้อหาการหารไม่ลงตัวเป็นจำนวนมาก (Porter 1978 : 5) ซึ่งถึงกำหนดว่าจำนวนแบบฝึกหัดในแบบเรียนคือสิ่งที่ กำหนดน้ำหนักความสำคัญของเนื้อหาในหลักสูตรที่ควรวัด ตามข้อเสนอของ ชมิดท์ (Schmidt 1983 : 172) แล้ว แสดงว่าแบบสอบมาตรฐานทั้ง 4 ฉบับ มีความลำเอียงของเนื้อหาในบางเนื้อหา และความลำเอียงของเนื้อหาของแบบสอบแต่ละฉบับ แตกต่างกัน

สำหรับในประเทศไทย เจตนา แคงอินทวัฒน์ (2526 : 70 - 73) ได้ศึกษาความสัมพันธ์ระหว่างหลักสูตรการสอนและการวัดสัมฤทธิ์ผลในค่านเนื้อหาวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 4 ในโรงเรียนสังกัดกรุงเทพมหานคร โดยการศึกษา

ความสัมพันธ์ระหว่างเนื้อหาในหลักสูตร ซึ่งได้จากการประเมินระดับความสำคัญของเนื้อหา จากผู้ทรงคุณวุฒิด้านหลักสูตร จำนวน 10 คน เนื้อหาในการสอน ซึ่งได้จากการประเมินระดับความสำคัญ โดยครูผู้สอน วิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 4 จำนวน 119 คน และเนื้อหาในการวัดสัมฤทธิ์ผล ซึ่งได้จากการวิเคราะห์แบบสอบวัดผลสัมฤทธิ์ วิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 4 ประจำภาคเรียนที่ 1 ปีการศึกษา 2525 ผลการศึกษาในด้านความสัมพันธ์ระหว่างเนื้อหาในการวัดสัมฤทธิ์ผล และเนื้อหาในหลักสูตร พบว่า ไม่มีความสัมพันธ์กันระหว่างระดับความสำคัญของเนื้อหาในการวัดสัมฤทธิ์ผล และระดับความสำคัญของเนื้อหาในหลักสูตร ซึ่งถ้ากำหนดว่า เนื้อหาที่แบบสอบวัดผลสัมฤทธิ์วัด ควรเป็น เนื้อหาที่หลักสูตรกำหนด การศึกษาของ เจตนา บังชีวะ เนื้อหาของแบบสอบวัดผลสัมฤทธิ์ และเนื้อหาที่ควรวัดตามหลักสูตรไม่สอดคล้องกัน แสดงว่าแบบสอบมีความลำเอียงของเนื้อหา

การประเมินผลการเรียนในโรงเรียนประถมศึกษา

ในส่วนนี้ได้กล่าวถึงหลักเกณฑ์การประเมินผลการเรียนในโรงเรียนประถมศึกษา และงานวิจัยที่เกี่ยวข้องกับสภาพและปัญหาในการวัด และประเมินผลการศึกษา

1. หลักเกณฑ์การประเมินผลการเรียนในโรงเรียนประถมศึกษา

การประเมินผลการเรียนตามหลักสูตรประถมศึกษา พุทธศักราช 2521 ได้กำหนดให้โรงเรียนประเมินผลการเรียนทุกชั้น เพื่อปรับปรุงการเรียนการสอน และตัดสินผลการเรียน ดังนี้ (กรมวิชาการ 2525 : 185)

1. การประเมินผลการเรียนเพื่อปรับปรุงการเรียนการสอน

1.1 การตรวจสอบความรู้พื้นฐาน และทักษะเบื้องต้นของนักเรียนที่สอน คือการทดสอบความพร้อมก่อนเรียน เพื่อทราบพื้นฐานทางกาย อารมณ์ สังคม ความรู้ และความคิดของผู้เรียนแต่ละคน จะได้นำไปใช้ให้เป็นประโยชน์ในการจัดการเรียนการสอนที่จะดำเนินต่อไปได้เหมาะสม

1.2 การประเมินความรู้ตามจุดประสงค์การเรียนรู้ที่กำหนดไว้ในสมุดประจำชั้น เป็นการประเมินผลในลักษณะสอบไป สอบไป เพื่อดูว่านักเรียนมีความสามารถ และพฤติกรรมตามที่กำหนดหรือไม่ ทั้งนี้เพื่อประโยชน์ในการ

ปรับปรุงการเรียนรู้ของนักเรียน และกำรสอนของครู โดยเลือกใช้วิธีการ และ เครื่องมือที่เหมาะสมหลาย ๆ วิธี เช่น สั้งเกต สัมภาษณ์ ตรวจผลงาน และทดสอบ

1.3 การประเมินผลปลายภาคเรียนเป็นการประเมินผลเพื่อตรวจสอบ คุณภาพการสอนของครู และความก้าวหน้าในการเรียนของนักเรียนในแต่ละกลุ่ม ประสมการณ และเพื่อเป็นการสรุปผลการเรียนรายงานใหญ่ปกครองทราบคอย

2. การประเมินผลเพื่อตัดสินผลการเรียน

2.1 การผ่านจุดประสงค์การเรียนรู้ ผู้สอนต้องดำเนินการตรวจสอบ จุดประสงค์ที่กำหนดไว้ ในสมุดประจำชั้นว่านักเรียนผ่านจุดประสงค์นั้น ๆ หรือไม่ ให้เสร็จสิ้นก่อนการสอบครั้งสุดท้าย 2 สัปดาห์ ซึ่งแต่ละกลุ่ม ประสมการณ นักเรียนจะต้องผ่านจุดประสงค์ไม่น้อยกว่าร้อยละ 60

2.2 การประเมินผลปลายปี เป็นการประเมินเพื่อใช้เป็นเกณฑ์ในการพิจารณาเลื่อนชั้น ในชั้นประถมศึกษาปีที่ 2, 4, 6

จากระเบียบการประเมินผลการเรียนดังกล่าวข้างต้น จะเห็นได้ว่าการตรวจสอบความรู้พื้นฐาน และทักษะเบื้องต้นของนักเรียนที่สอน เป็นการประเมินผลในลักษณะ การประเมินผลก่อนสอน (pre evaluation) ซึ่งเป็นการตรวจสอบความรู้เดิม และความพร้อมของนักเรียน และเป็นการตรวจสอบว่านักเรียนได้เรียนรู้ในสิ่งที่ต้องเรียน มามากน้อยเพียงใด การประเมินความรู้ตามจุดประสงค์การเรียนรู้ที่กำหนดไว้ในสมุด ประจำชั้น เป็นการประเมินผลในลักษณะการประเมินผลย่อย หรือการประเมินความ ก้าวหน้า (formative evaluation) ซึ่งเป็นการประเมินผลระหว่างทำการสอน เพื่อวัดระดับความรอบรู้ และประเมินขอบกว้างของการเรียนการสอน ส่วนการประเมิน ผลปลายภาค และปลายปี เป็นการประเมินผลในลักษณะการประเมินผลรวม (summative evaluation) เป็นการประเมินผลหลังจากสิ้นสุดการเรียนการสอนแล้ว เพื่อเป็น ข้อมูลสำหรับตัดสินความสามารถของนักเรียน

ในส่วนที่เกี่ยวกับการประเมินผลปลายปี ระเบียบกระทรวงศึกษาธิการว่าด้วย การประเมินผลการเรียนระดับประถมศึกษา ได้กำหนดให้นำจุดประสงค์ที่สำคัญในแต่ละ มวลประสบการณ์ที่ทำการสอนในภาคเรียนที่ 1 บางจุดประสงค์มาทำการสอบร่วมกับ จุดประสงค์ที่สอนในภาคเรียนที่ 2 (กรมวิชาการ 2523 : 265) จึงอาจกล่าวได้ว่าการวัดและประเมินผลปลายปีตามหลักสูตรประถมศึกษา มีจุดมุ่งหมายเพื่อวัดผลสัมฤทธิ์ทางการ เรียนของนักเรียน ตามเนื้อหาและจุดประสงค์ที่กำหนดในหลักสูตร สำหรับจุดประสงค์

ปลายทางที่กำหนดไว้ในหลักสูตรคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 4 มี 20 จุดประสงค์ คือ

1. เรียงลำดับจำนวนจากจำนวนมากไปจำนวนน้อย และจากจำนวนน้อยไปจำนวนมากได้
2. อ่านและเขียนบ้านเลขที่ หมายเลขโทรศัพท์ ทะเบียนยานพาหนะ เลขที่หนังสือราชการ ปีพุทธศักราช และคริสต์ศักราช ได้
3. ทำโจทย์ปัญหาเกี่ยวกับการบวกและลบระคนให้เป็นประโยคสัญลักษณ์ และหาคำตอบได้
4. ทำโจทย์ปัญหาเกี่ยวกับความจุ และน้ำหนัก ได้
5. บอกได้ว่าเส้นตรงบนระนาบเดียวกันคู่ใดขนานกัน
6. สร่างรูปวงกลมได้
7. ใช้รูปเรขาคณิตประติรูปคล้าย หรือภาพได้
8. อ่านแผนรูปภาพ แผนภูมิแท่ง ตาราง และเขียนแผนรูปภาพ แผนภูมิแท่ง ได้
9. ทำโจทย์ปัญหาความยาว หรือ ระยะทาง ได้
10. ทำโจทย์การคูณ หาร ระคนได้ (จำนวนไม่เกิน 2 หลัก)
11. หาพื้นที่รูปสี่เหลี่ยมมุมฉากที่กำหนดความยาวของด้านให้ได้
12. ทำโจทย์ประโยคสัญลักษณ์ และโจทย์ปัญหาการคูณเศษส่วนกับจำนวนเต็มได้
13. ทำโจทย์ปัญหา บวก ลบ คูณ หาร เกี่ยวกับการเงินได้
14. ทำบันทึก อ่าน และเขียนแผนภูมิรูปภาพ แผนภูมิแท่งเกี่ยวกับรายรับ รายจ่าย ส่วนตัว และครอบครัวได้
15. ทำโจทย์ระคน และโจทย์ปัญหาการคูณ การหารได้ (จำนวนมากกว่า 2 หลัก)
16. ทำโจทย์ปัญหาเกี่ยวกับเวลาได้
17. อ่านตารางเวลาหรือแผนภูมิเวลาได้
18. อ่านและเขียนทศนิยมหนึ่งตำแหน่งได้

19. เปรียบเทียบค่าศนิยมหนึ่งตำแหน่งสองจำนวนและใช้เครื่องหมาย = หรือ \neq หรือ $>$ หรือ $<$ แสดงการเปรียบเทียบได้

20. ทำโจทย์ปัญหาร้อยละได้

สำหรับการจัดการเรียนการสอนของโรงเรียนสังกัดกรุงเทพมหานครนั้น ได้จัดการเรียนการสอนเป็น 2 ภาคเรียน ซึ่งตามหลักสูตรแล้ว ได้กำหนดให้จัดการเรียนการสอน จุดประสงค์ที่ 1 ถึงจุดประสงค์ที่ 10 ในภาคเรียนที่ 1 และจัดการเรียนการสอน จุดประสงค์ที่ 11 ถึงจุดประสงค์ที่ 20 ในภาคเรียนที่ 2 (สำนักการศึกษา กรุงเทพมหานคร มปป. : 62 - 63) และในการประเมินผลปลายปีนั้น แบบสอบที่ใช้ส่วนใหญ่มีเนื้อหา ซึ่งเป็นเนื้อหาที่สอนในภาคเรียนที่ 1 ร้อยละ 20 และเป็นเนื้อหาที่สอนในภาคเรียนที่ 2 ร้อยละ 80

2. งานวิจัยที่เกี่ยวข้องกับสภาพและปัญหาในการวัดและประเมินผลการศึกษา

จันทร์คัน โศครคำ (2523 : 67-71) ได้ศึกษาความคิดเห็นของครูประถมศึกษา สังกัดกรุงเทพมหานคร ที่เคยผ่านการอบรมวัดผลหลักสูตรระยะสั้นและระยะยาว ที่มีต่อการปฏิบัติงาน ปัญหาและอุปสรรคในการปฏิบัติงาน ความต้องการการอบรมวัดผล โดยใช้เวลาส่วนประเมินค่า 4 ระดับ จากการศึกษาความสัมพันธ์ของคะแนนความคิดเห็น พบว่า ในด้านความคิดเห็นที่มีต่อการปฏิบัติงานด้านการวัดผล ผู้เข้ารับการอบรมมีความคิดเห็นว่า หลังจากเข้ารับการอบรมแล้ว ตนปฏิบัติในด้านการวิเคราะห์จุดมุ่งหมายและเนื้อหาก่อนออกข้อสอบ ด้านการยึดโครงการสอนเป็นหลัก ในการออกข้อสอบด้านการออกข้อสอบได้ครอบคลุมเนื้อหาที่สอน ด้านการออกข้อสอบสอดคล้องกับจุดประสงค์ของเนื้อหา ด้านการออกข้อสอบวัดความรู้ความจำได้ ด้านการออกข้อสอบวัดความเข้าใจ และการนำไปใช้ได้ และด้านการออกข้อสอบวัดการวิเคราะห์ได้ ในระดับน้อย (ค่าสัมประสิทธิ์ อยู่ระหว่าง 1.56 ถึง 2.55) และในด้านการออกข้อสอบวัดพฤติกรรมระดับต่าง ๆ นั้น ผู้เข้ารับการอบรมมีความคิดเห็นว่า ตนออกข้อสอบวัดความเข้าใจ และการนำไปใช้ และออกข้อสอบวัดการวิเคราะห์ได้ในระดับน้อยกว่า การออกข้อสอบวัดความรู้ความจำ สำหรับในค่านปัญหา และอุปสรรคในการปฏิบัติงานวัดผล ด้านวิชาการ ผู้เข้ารับการอบรมมีความคิดเห็นว่า หลังจากเข้ารับการอบรมแล้วตนมีปัญหาด้านการวิเคราะห์เนื้อหา และพฤติกรรมของหลักสูตร และการออกข้อสอบวัดความเข้าใจและการนำไปใช้ในระดัปลานกลาง



สมปอง ฎวารางกูร (2525 : 89 - 90) ได้ศึกษาความคิดเห็นของครู
ประถมศึกษาในโรงเรียนประถมศึกษาจังหวัดกระบี่ทุกสังกัด โดยใช้มาตราส่วนประเมินค่า
5 ระดับ พบว่า ในส่วนที่เกี่ยวกับปัญหาของการวัดผลการเรียน ครูประถมศึกษาใน
จังหวัดกระบี่ มีความคิดเห็นว่า มีปัญหาปานกลางในด้านการกำหนดจุดประสงค์ที่จะวัด
สำหรับครูประจำชั้นประถมปีที่ 1 และ 2 ในโรงเรียนประถมศึกษาสังกัดองค์การบริหาร
ส่วนจังหวัดกาญจนบุรี มีความคิดเห็นว่าตนมีปัญหามาก ในการจัดสอบปลายภาคเรียนให้
ครอบคลุมจุดประสงค์การเรียนรู้ (วิไลภา อารีรัตน์, 2523 : 62 - 63)

จากระดับการปฏิบัติ และระดับความเป็นปัญหาเกี่ยวกับการวัดผลการศึกษาของ
ครูประถมศึกษา ดังกล่าวข้างต้น จะเห็นได้ว่า ระดับการปฏิบัติ และระดับความเป็น
ปัญหา มีความเกี่ยวข้องกัน กล่าวคือ เมื่อครูประถมศึกษา มีปัญหาและอุปสรรค ใน
ด้านการกำหนดเนื้อหา และจุดประสงค์ที่จะวัด การจัดการสอบให้ครอบคลุมจุดประสงค์
การเรียนรู้ และการออกข้อสอบวัดพฤติกรรมระดับสูง อาจทำให้มีการปฏิบัติ ในด้าน
การวิเคราะห์หลักสูตร การออกข้อสอบให้สอดคล้อง และครอบคลุมจุดประสงค์ และ
การออกข้อสอบวัดพฤติกรรมระดับสูง ใ้คนอยลง ซึ่งสิ่งเหล่านี้อาจมีผลกระทบทำให้แบบ
สอบที่สร้างขึ้น มีความลำเอียงของเนื้อหา และขาดความตรงของเนื้อหาได้

จากที่กล่าวมาทั้งหมดจะเห็นได้ว่า ความตรงที่แบบสอบวัดผลสัมฤทธิ์ทางการ
เรียน จำเป็นต้องมี คือ ความตรงของเนื้อหา ซึ่งความตรงของเนื้อหานี้ มีความ
เกี่ยวข้องกับระดับความเป็นตัวแทนของเนื้อหาของแบบสอบ ต่อเนื้อหาที่ต้องการวัด
ถ้าแบบสอบที่ขาดความตรงของเนื้อหา คะแนนที่ได้จากการใช้แบบสอบวัดความรู้ความ
สามารถของบุคคล จะเป็นตัวประมาณค่าที่ลำเอียงของคะแนนจริงที่สัมพันธ์กับขอบเขต
เนื้อหาที่ต้องการวัด ซึ่งกล่าวได้ว่า แบบสอบมีความลำเอียงของเนื้อหา ดังนั้น
ความตรงของแบบสอบ และความลำเอียงของแบบสอบมีความสัมพันธ์กันแบบผกผัน จึง
สามารถตรวจสอบความตรงของเนื้อหาของแบบสอบได้ โดยการประเมินค่าความลำเอียง
ของเนื้อหาของแบบสอบ โดยอ้างอิงกับขอบเขตเนื้อหาที่ต้องการวัด ซึ่งขึ้นอยู่กับจุดประสงค์
ของการสอบนั้น ๆ สำหรับการวัด และประเมินผลปลายปี ตามหลักสูตรประถมศึกษา
พุทธศักราช 2521 ได้กำหนดให้เป็นการประเมินผล เพื่อตรวจสอบผลการเรียนขั้น
สุดท้าย ตามเนื้อหา และจุดประสงค์ที่กำหนดในหลักสูตร ฉะนั้นการตรวจสอบความตรง

ของเนื้อหา และความลำเอียงของเนื้อหาของแบบสอบถามผลสัมฤทธิ์ปลายปี ความหลักสูตร
ประถมศึกษา พุทธศักราช 2521 จึงต้องอ้างอิงอยู่กับเนื้อหาและจุดประสงค์ที่กำหนด
ในหลักสูตร