



วรรณคดีที่เกี่ยวข้อง

วรรณคดีที่เกี่ยวข้องกับผลของระบบประเมินผลการเรียนที่มีผลต่อผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้นมัธยมศึกษาตอนปลายแบ่งเป็น 3 ตอน คือ ตอนที่หนึ่งว่าความโน้มนัสและการประเมินผลการเรียนรู ตอนที่สองว่าควย ระบบประเมินผลการเรียน ตอนที่สามว่าควย องค์ประกอบที่มีต่อผลสัมฤทธิ์ทางการเรียน ซึ่งผู้วิจัยจะได้กล่าวในรายละเอียดแต่ละตอนดังต่อไปนี้

มโนทัศน์และการประเมินผลการเรียนรู

จากการสำรวจคำนิยามของ "การประเมินผล" (Evaluation) พบว่ามีผู้รู้หลายคนได้ให้คำนิยามของ "การประเมินผล" ไว้เช่น เอ็ม สคริฟเวน¹ (M. Scriven) ได้ให้คำนิยามของประเมินผลว่า เป็นการรวบรวมและวิเคราะห์ข้อมูลอย่างมีระบบ เพื่อบ่งชี้คุณค่าของสิ่งใดสิ่งหนึ่ง เจมส์ เอ เทอร์วิลลิเจอร์² (James A Terwilliger) ให้นิยามของการประเมินผลว่าเป็นกระบวนการตัดสินคุณค่าในด้านความสามารถ การแสดงออกหรือ

¹M.Scriven, "The Methodology of Evaluation," In R.E. Stake (Ed.), Curriculum Evaluation. American Educational Research Association. Monograph Series on Evaluation, No. 1, Chicago : Rand McNally, 1967. อ้างถึงใน สมหวัง พิธิยานุวัฒน์ "หลักการวัดและประเมินผล" คู่มืออาจารย์คานการเรียนการสอน, หน้า 177.

²James A. Terwilliger, Assigning Grades to Students. (Illinois : Scott, Foresman and Company), p. 2.

พฤติกรรมต่าง ๆ เฉพาะแต่ละวิชา จอห์น เอ กรีน¹ (John A. Green) ให้คำนิยามของการประเมินผลว่าเป็นกระบวนการตัดสินคุณค่าอย่างมีระบบของข้อมูลที่ได้ตามกระบวนการวัด นอร์แมน อี กรอนลันด์² (Norman E. Gronlund) ให้นิยามของการประเมินผลว่าเป็นกระบวนการตัดสินสัมฤทธิ์ผลอย่างมีระบบของผู้สอนตามวัตถุประสงค์ของการสอนที่ตั้งไว้ และจะมีความหมายลึกซึ้งกว่าการวัด เพราะจะรวมความถึงพฤติกรรมที่วัดได้และวัดไม่ได้ รวมทั้งการตัดสินใจของผู้สอน ลี เจ ครอนบาช³ (Lee J. Cronbach) ให้นิยามของการประเมินผลในแง่ของการวัดออกเป็น 2 ลักษณะ คือการวัดพฤติกรรมสูงสุดเพื่อตัดสินความสามารถของแต่ละบุคคล เช่น การวัดความถนัด การวัดสัมฤทธิ์ผล เป็นต้น และการวัดพฤติกรรมทั่วไป เพื่อตัดสินความสนใจ ทักษะคิด บุคลิกภาพ เป็นต้น สมหวัง พิธิยานุวัฒน์⁴ ให้นิยามของการประเมินผลโดยทั่วไปว่า เป็นกระบวนการวัดตัดสินคุณค่าของสิ่งของหรือการกระทำใด ๆ โดยเปรียบเทียบกับเกณฑ์ (Criterion) ซึ่งเกณฑ์อาจมีมาก่อนและเป็นอิสระจากพฤติกรรมของกลุ่ม หรือเป็นเกณฑ์ที่ได้จากพฤติกรรมของกลุ่ม และในแง่ของการเรียนการสอน การประเมินผลการเรียนเป็นกระบวนการที่ใช้ตัดสินว่า นักเรียนได้เปลี่ยนแปลงพฤติกรรมตามวัตถุประสงค์ที่วางไว้ตามเป้าหมายของการสอนหรือไม่ เพียงใด เช่น ในค่านพหุธิปัญญา นักเรียนเปลี่ยนจากผู้ที่ไม่สามารถคำนวณหาปริมาตรของลูกบาศก์ที่มีความถี่ เป็นผู้ที่สามารถ

¹John A. Green, Introduction to Measurement and Evaluation, (New York: Dodd, Mead, 1970), p.3.

²Norman E. Gronlund, Measurement and Evaluation in Teaching, 3rd ed. (New York: Mcmillan Co., Inc., 1976), pp. 5-18.

³Lee J. Cronbach, Essential of Psychological Testing, 3d ed. (New York : Harper and Row, 1970), p. 26.

⁴สมหวัง พิธิยานุวัฒน์, "การประเมินผลโครงการประชุม" (เอกสารประกอบการเรียนวิชา เรื่องคดีเฉพาะทางการวัดและประเมินผลการศึกษา ประจำภาคต้น ปีการศึกษา 2522), หน้า 1.

คำวาทหามัธยมศึกษิตแบบแจกแจงความดีไ้ถูกตองหรือไม และขอมูลเหล่านีเป็นเครื่อง
 ช่วยตัดสินใจประสิทธิภาพการสอนของอาจารย์ และการเรียนของนักเรียนอีกด้วย การ
 ประเมินผลการเรียนควรมุ่งเน้นที่เอ็กบุคคล ไม่ใช่ความแตกต่างระหว่างบุคคล คือควรเน้น
 ว่านักเรียนเรียนรู้อะไรบ้าง โดยเทียบกับวัตถุประสงค์ของการสอน

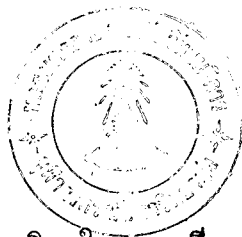
จากคำนิยามของ "การประเมินผล" ข้างต้น จะเห็นได้ว่า การประเมินผลนั้นมีความ
 หมายครอบคลุมไปถึงการวัดผลด้วย เราจะไม่สามารประเมินผลได้ก็ตามที่ต้องการ
 ถ้าไม่มีการวัดผลที่ดี แต่การวัดผลที่ดีไม่จำเป็นต้องมีการประเมินผลก็ได้ ตัวอย่างเช่น ใน
 วิชาพลศึกษา นายสิทธิวี้งแข่งระยะทาง 100 เมตร ใช้เวลา 10 วินาที ชั้นนี้เรามีเพียง
 ข้อมูลจากการวัด จนกระทั่งอาจารย์พลศึกษาประเมินความสามารถในการวิ่งของนายสิทธิ
 โดยการตีความหมายเวลาที่นายสิทธิใช้ไป 10 วินาทีต่อระยะทาง 100 เมตร ว่าเป็น
 นักวิ่งดีเท่าจัดหรือนักวิ่งลมกรด เป็นต้น ปัจจุบันจึงได้เปลี่ยนคำว่า "การวัดผลการศึกษา"
 ซึ่งเคยใช้มาแต่เดิม มาใช้เป็น "การประเมินผลการเรียน"¹

ในแง่ของการเรียนการสอนโดยเฉพาะแล้ว เจมส์ เอ เทอร์วิลลิเจอร์² (James
 A Terwilliger) กล่าวว่า การประเมินผลการเรียนหรือการตัดสินระดับสัมฤทธิ์ผล
 ของผู้เรียนนั้น ผู้สอนเป็นผู้ที่มีภาระหน้าที่รับผิดชอบโดยตรง อีกทั้งผู้บริหาร อาจารย์แนะแนว
 และเพื่อนผู้เรียนเองก็มีบทบาททางอ้อมในการตัดสินคุณค่าของระดับสัมฤทธิ์ผลของแต่ละคน
 และกล่าวว่าการประเมินผลการเรียนควรมีวัตถุประสงค์สำคัญ 3 ประการคือ

1. เพื่อบ่งชี้ว่านักเรียนมีทักษะและความรู้ที่สำคัญพอเพียงหรือไม่
2. เพื่อวินิจฉัยจุดเด่นและจุดด้อยของนักเรียน โดยสังเกตจากกระสวนของ
 สัมฤทธิ์ผลที่แสดงออกมาให้เห็น และ

¹ กระทรวงศึกษาธิการ, กรมวิชาการ. คู่มือการประเมินผลการเรียนตามหลักสูตร
 ประโยชน์มัธยมศึกษาตอนปลาย พุทธศักราช 2518. กรุงเทพฯ, 2518. หน้า 24.

² James A. Terwilliger, Assigning Grades to Students
 pp. 4-7.



3. เพื่อป้องกันระดับสัมฤทธิผลในการเรียนของนักเรียนแต่ละคน

ข้อมูลที่ได้จากการประเมินผลตามวัตถุประสงค์ ข้อ 1. และ 2. เป็นประโยชน์อย่างมากต่อการปรับปรุงการสอนให้เหมาะสมกับความต้องการของนักเรียน อีกทั้งช่วยให้นักเรียนทราบถึงจุดเด่นและจุดด้อยของตนเอง จะได้ปรับปรุงวิธีการเรียนให้ดีขึ้น ซึ่งมีประสิทธิภาพยิ่งขึ้น ซึ่งตรงกับศัพท์เทคนิคทางการประเมินผลหลักสูตรที่ สกริฟเวน¹ (Scriven) เรียกว่า การประเมินความก้าวหน้า (Formative Evaluation) ในขอบข่ายของการประเมินผลนักเรียน การประเมินความก้าวหน้า หมายถึง การประเมินระหว่างภาคเรียน หรือการประเมินขณะที่การสอนยังคงดำเนินอยู่ เพื่อมุ่งตัดสินคุณค่าเบื้องต้นของการสอนหรือการเรียน อันจะนำไปสู่การปรับปรุง เปลี่ยนแปลงวิธีการสอน วิธีการเรียน เพื่อก่อให้เกิดผลดียิ่งขึ้น ดังนั้นการประเมินความก้าวหน้า จึงมีจุดมุ่งหมายเพื่อค้นหาว่านักเรียนได้เรียนรู้ไปมากน้อยเพียงใดในเรื่องเฉพาะตอนหรือหน่วยบทเรียน และมุ่งที่จะประเมินว่าอาจารย์ได้จัดประสบการณ์เรียนรู้ให้นักเรียน และนักเรียนมีวิธีการเรียนที่เหมาะสมเพียงใด ตัวอย่างการประเมินความก้าวหน้าได้แก่ การทดสอบย่อยประจำหน่วยบทเรียน เป็นต้น

สำหรับการประเมินผลตามวัตถุประสงค์ข้อ 3 นั้น มีประโยชน์ต่อการปรับปรุงการเรียนการสอนน้อยกว่าการประเมินผลตามวัตถุประสงค์สองข้อแรก เพราะเป็นการประเมินผลนักเรียนหลังจากการสอนได้สิ้นสุดลงแล้ว การประเมินระดับสัมฤทธิผลของนักเรียนในขั้นนี้ เป็นประโยชน์ในการเลือกวิชาเรียนต่อไป หรือการจัดชั้นเรียนที่เหมาะสม เป็นต้น การประเมินผลในลักษณะนี้เป็นการตัดสินคุณค่า หรือระดับสัมฤทธิผลสะสมของนักเรียน หลังจากการสอนอย่างเป็นทางการสิ้นสุดลง ตัวอย่างเช่น การสอบไล่ประจำภาคเรียน เป็นต้น มโนทัศน์ของการประเมินผลแบบนี้เหมือนกันกับศัพท์การประเมินผลหลักสูตรที่ว่า การประเมินผลรวมสรุป (Summative Evaluation)

¹ M. Scriven, "The Methodology of Evaluation," In Taylor, F.A. and Gowley, D.M. (Ed.) Readings in Curriculum Evaluation. (Dubugue : Wm. C. Brown, 1972), pp. 28-48.

ในการประเมินผลการเรียนตามหลักสูตรมัธยมศึกษาตอนปลาย พ.ศ.2518 ที่ใช้
ในปัจจุบัน ก็ได้แบ่งการประเมินผลออกเป็นประเภทใหญ่ ๆ 2 ประเภท¹ คือ

1. การประเมินผลเพื่อการเรียน (Formative Evaluation) มีจุดประสงค์
เพื่อการพัฒนาการสอนของครู การพัฒนาการเรียนของนักเรียน ตลอดจนการพัฒนาหลักสูตร
ซึ่งไม่เกี่ยวข้องกับการตัดสิน ใด ตก หรือการให้ระดับคะแนนแต่อย่างใด การประเมินผล
การเรียนแบบนี้อาจเรียกว่า การประเมินผลส่วนย่อย อาจประเมินผลเมื่อเริ่มต้นการสอน
เพื่อศึกษาว่านักเรียนคนใดมีพื้นฐานในรายวิชาที่กำลังจะเรียนนี้มากน้อยเพียงไร หรือ
ประเมินผลหลังการสอน เมื่อจบบทเรียนเพียงหนึ่งบทหรือสองบท การประเมินผลเพื่อการเรียน
นี้เพื่อต้องการเอาผลมาใช้ (Feed-back) ในการปรับปรุงการสอนของครู และปรับปรุง
การเรียนของนักเรียน ตลอดจนปรับปรุงหลักสูตรให้ดีขึ้น

2. การประเมินผลเพื่อตัดสินผลการเรียน (Summative Evaluation) ได้แก่
การประเมินผลเพื่อตัดระดับคะแนน (เกรด) หรือตัดสินใดตก การประเมินผลการเรียน
แบบนี้อาจเรียกว่า การประเมินผลรวมยอด

004945

บลูม (Bloom) ได้จัดแบ่งการประเมินผลการศึกษากออกเป็น 2 จำพวกใหญ่ ๆ²
คือ การประเมินความก้าวหน้า (Formative Evaluation) และการประเมินผลรวมสรุป
(Summative Evaluation) เช่นกัน และย้ำว่าควรจะได้ใช้ควบคู่กันไปทั้งสองจำพวก
เพื่อให้การประเมินผลการศึกษามีกระบวนการ สอดคล้องกับการเรียนการสอนให้มากที่สุด

วัตถุประสงค์ที่สำคัญของการประเมินผลการเรียน³ ในแง่ของการเรียนการสอน
มีดังนี้คือ

¹กระทรวงศึกษาธิการ, กรมวิชาการ. คู่มือการประเมินผลการเรียนตามหลักสูตร
ประโยคมัธยมศึกษาตอนปลาย พ.ศ.2518. หน้า 24-27.

²Benjamin S. Bloom; J Thomas Hastings and George F. Madaus.
Handbook of Formative and Summative Evaluation of Student Learning.
(New York: McGraw-Hill Book Company, 1971), p. 14.

1. เพื่อที่จะจัดตำแหน่ง (Placement) ซึ่ง นอร์แมน อี กรอนลันด์¹ กล่าวว่า การประเมินผลประเภทนี้มีจุดมุ่งหมายเพื่อศึกษาทักษะของนักเรียนแต่ละคนว่า มีทักษะและความรู้พอเพียงในการที่จะเริ่มบทเรียนแล้วหรือยัง ซึ่งผลที่ได้รับจะเป็นประโยชน์ในด้านการแยก² (Classification) นักเรียนแต่ละคนว่ามีพื้นฐานอย่างไร ควรจัดอยู่ในประเภท เด็กเก่ง ปานกลาง หรืออ่อน นอกจากนี้ยังมีจุดมุ่งหมายเพื่อดูความสนใจ นิสัยในการทำงาน ซึ่งจะช่วยในการคัดเลือกเรียนต่อได้อีกด้วย

2. เพื่อที่จะวินิจฉัย (Diagnostic) เทอร์วิลลิเจอร์³ กล่าวว่า จะทำให้ทราบจุดเด่น จุดด้อย ของนักเรียนแต่ละคนในแต่ละวิชา และทำให้ครูสามารถซ่อมเสริมให้แก่ นักเรียนแต่ละคนได้อย่างถูกต้อง โดยทั่วไปมักจะประเมินเมื่อจบแต่ละหน่วยเรียน⁴

3. เพื่อจะเปรียบเทียบความสามารถของตัวเด็กเอง (Assessment) โดยครู จะดูว่าในตัวนักเรียนแต่ละคนมีความมั่งงายมากน้อยเพียงใด ซึ่งจะกระทำได้โดยการทดสอบ 2 ครั้ง แล้วนำผลที่ได้จากการทดสอบมาเปรียบเทียบกัน

4. เพื่อที่จะนำผลไปใช้ทำนาย (Prediction) ผลที่ได้จากการสอบในแต่ละวิชาของนักเรียน จะเป็นแนวทางทำให้ทราบความสามารถ และความสำเร็จในการเลือกเรียนแต่ละวิชาของเด็ก

5. เพื่อที่จะนำผลไปใช้ประเมินค่า (Evaluation) การประเมินผลลักษณะนี้จะกระทำเมื่อการเรียนการสอนสิ้นสุดลง โดยนำข้อมูลที่ได้จากการวัดมาเป็นแนวทางในการให้ระดับคะแนนแก่นักเรียน⁵ ซึ่งทำให้ทราบสถานภาพของนักเรียนภายในกลุ่มว่าอยู่ระดับใด

¹Grondlund, Measurement and Evaluation in Teaching, p.525.

²เฉลิมศรี ฉิ่งถิ่น, "การวัดผลในโรงเรียน" ศูนย์ศึกษา (มีนาคม-เมษายน 2514) : 17.

³Terwilliger, Assigning Grades to Students, p. 6.

⁴Grondlund, Measurement and Evaluation in Teaching, p.526.

⁵Ibid., p.530.

นอกจากนี้ยังสามารถประเมินมาตรฐานระหว่างโรงเรียนได้อีกด้วย¹

ประโยชน์ของการประเมินผลการเรียนรู้²

1. ช่วยให้ผู้สอนทราบพฤติกรรมของผู้เรียนก่อนเข้าเรียนหรือก่อนรับการสอน เพื่อประโยชน์ในการจัดการเรียนการสอนให้เหมาะสมกับผู้เรียน อีกทั้งใช้พฤติกรรมก่อนเข้าเรียนเป็นพื้นฐานในการตรวจสอบการเปลี่ยนแปลงพฤติกรรมของผู้เรียน อันเนื่องมาจากการสอนอีกด้วย
2. ช่วยให้ผู้สอนแนะและปรับปรุงเป้าหมายตามความเป็นจริงสำหรับผู้เรียนแต่ละคน
3. ช่วยให้ผู้สอนทราบถึงประสิทธิภาพของการสอน กล่าวคือ ทราบว่าผู้เรียนบรรลุวัตถุประสงค์ของการสอนเพียงใด
4. ช่วยให้ผู้สอนระบุ ประเมิน และปรับปรุงวิธีการสอนแบบต่าง ๆ
5. ช่วยให้ผู้สอนทราบถึงระดับสัมฤทธิ์ผลของผู้เรียนในระยะต่าง ๆ ของการเรียนวิชาใดวิชาหนึ่ง
6. ช่วยให้ข้อมูลที่เป็นแนวทางเพื่อการเปรียบเทียบระดับสัมฤทธิ์ผลระหว่างผู้เรียน และเป็นแนวทางในการให้อันดับคะแนนแทนระดับสัมฤทธิ์ผลทางการเรียนของผู้เรียนด้วย
7. ใช้เป็นข้อมูลประกอบการตัดสินใจว่า ผู้เรียนมีความรู้ความสามารถพอเพียงที่จะเลื่อนชั้นสูงขึ้นหรือไม่
8. ช่วยให้ผู้เรียนทราบและมีส่วนร่วมในการตั้งเป้าหมายการสอน ซึ่งจะกระตุ้นให้ผู้เรียนตั้งใจศึกษาเล่าเรียน เพื่อบรรลุเป้าหมายของการสอนยิ่งขึ้น
9. ช่วยเพิ่มแรงกระตุ้น และเกื้อกูลให้ผู้เรียนเรียนรู้เพิ่มขึ้น

¹ เณดิมศรี ฉิ่งถิ่น, "การวัดผลในโรงเรียน" หน้า 7.

² W. Mehrens and I. Lehmann. Measurement and Evaluation in Education and Psychology. (New York: Holt, Rinehart and Wingston, 1973), pp. 7-15.

10. ช่วยส่งเสริมและปลูกฝังใฝ่เรียนเป็นผู้ที่มีนิสัยในการศึกษาอย่างมีระบบ
11. ให้ข้อมูลย้อนกลับแก่ผู้เรียน ซึ่งจะแสดงถึงจุดเด่น และจุดค้อยของแต่ละบุคคล
ในเนื้อหาเรื่องใดเรื่องหนึ่ง
12. ใช้เป็นประโยชน์ในการแนะแนวอาชีพ แนะนำการศึกษาต่อ ตลอดจนแนะ
วิธีการปัญหาส่วนตัวของผู้เรียนได้อย่างเหมาะสม
13. ใช้เป็นข้อมูลประกอบการตัดสินใจเกี่ยวกับการบริหาร เช่น การคัดเลือก
การจัดประเภท เป็นต้น
14. ใช้เป็นข้อมูลประกอบการประเมินผลหลักสูตร หรือโปรแกรมการศึกษาเพื่อ
ประโยชน์ในการปรับปรุงหลักสูตร หรือโปรแกรมการศึกษาให้เหมาะสมยิ่งขึ้น
15. เป็นประโยชน์ในการตรวจสอบสมมติฐาน และการวิจัยทางการศึกษา

ระบบการประเมินผลการศึกษา

เทอร์วิลลิเจอร์ (Terwilliger) ได้แบ่งรูปแบบในการประเมินผลสัมฤทธิ์เป็น
4 รูปแบบ¹ คือ

1. รูปแบบแบ่งประเภทหรือรูปแบบสัมบูรณ์² (The Categorical or Absolute Model) การตัดสินสัมฤทธิ์ผลของนักเรียนตามรูปแบบนี้จะตั้งบนพื้นฐานมโนทัศน์ "มาตรฐานสัมบูรณ์" (Absolute Standard) ซึ่งผู้สอนจะกำหนดตามความเหมาะสม เช่น ผู้ที่จะไคระดับคะแนน ก. ต่อเมื่อไคระคะแนน 80 % ขึ้นไป เป็นต้น ผู้ใช้ทุกวิธีวิธีการตัดสินแบบนี้ จะเน้นเนื้อหาวิชาที่เรียนเป็นสำคัญ โดยนักเรียนต้องแข่งขันกับเกณฑ์ที่ครูวางไว้ วิธีการให้ระดับคะแนนตามรูปแบบนี้กระทำได้ใน 3 ลักษณะ คือ วิธีใช้คะแนนเปอร์เซ็นต์ ซึ่งระดับคะแนนที่ได้จะกำหนดออกมาในรูปคะแนนเปอร์เซ็นต์ เช่น 60-70 เปอร์เซ็นต์ ไคระดับคะแนน C

¹James A. Terwilliger, Assigning Grades to Students, p.14.

²Ibid., pp. 26-37.

เป็นต้น วิธีใช้หน่วยธรรมชาติของการปฏิบัติ ซึ่งระดับคะแนนที่จะได้จะกำหนดออกมาในรูปจำนวนทักษะ เช่น พิมพ์ดีดได้ 120 คำต่อนาที ใ้ระดับคะแนน A เป็นต้น และวิธีใช้มโนทัศน์ของการจัดลำดับชั้น ซึ่งการใ้ระดับคะแนนจะกำหนดในรูปของความสามารถในการตอบถูกว่อกตามวัตถุประสงค์ที่วางไว้ เช่น ตอบถูก 80 เปอร์เซ็นต์ของวัตถุประสงค์ที่วางไว้ จะใ้ระดับคะแนน C เป็นต้น

2. รูปแบบพัฒนาการ หรือรูปแบบการเปลี่ยนแปลง¹ (The Growth or Change Model) การตัดสินสัมฤทธิผลของนักเรียนตามรูปแบบนี้จะตั้งบนพื้นฐานมโนทัศน์ "ความงอกงาม" (Growth) ซึ่งเกิดขึ้นก่อนและหลังการเรียนรู้ของนักเรียนแต่ละบุคคลในแต่ละภาคเรียน วิธีการใ้ระดับคะแนนตามรูปแบบนี้ นักเรียนต้องแข่งขันกับตนเอง โดยจะมีการทดสอบ 2 ครั้ง คือ ก่อนการเรียนรู้เพื่อวัดสถานภาพเริ่มต้น และหลังการเรียนรู้เพื่อวัดสถานภาพสุดท้าย แล้วนำปริมาณความงอกงามที่วัดได้มาเปรียบเทียบในลักษณะอิงเกณฑ์หรืออิงกลุ่ม สำหรับรูปแบบนี้ผู้ที่ใ้ระดับคะแนน ก. แสดงถึงผู้ที่มีพัฒนาการดีกว่าคนอื่น ซึ่งอาจจะถึงขั้นมีคุณภาพหรือไม่ถึงขั้นก็ได้

3. รูปแบบปกติวิสัยหรือรูปแบบสัมพัทธ์² (Normative or Relative Model) การตัดสินสัมฤทธิผลของนักเรียนจะยึดการเปรียบเทียบสัมฤทธิผลของนักเรียนแต่ละคนภายในกลุ่ม วิธีการใ้ระดับคะแนนตามรูปแบบนี้ นักเรียนต้องแข่งขันกับกลุ่ม ซึ่งวิธีการจะกระทำได้ใน 4 ลักษณะ คือ การใช้โค้งปกติ (Normal Curve) การใช้กลุ่มห้องเรียนเป็นปกติวิสัย (Class as the Norm Group) การใ้ที่นักเรียนที่ได้รับการฝึกฝนเหมือนกันเป็นกลุ่มปกติวิสัย และการใ้รูปแบบสัมบูรณ์และรูปแบบสัมพัทธ์รวมกัน เป็นต้น เป็นเกณฑ์ในการตัดสินระดับคะแนน

4. รูปแบบสัมฤทธิผลตามความสามารถ หรือตามความคาดหวัง³ (The

¹James A. Terwilliger, Assigning Grades to Students, pp. 38-51.

²Ibid., pp. 74-98

³Ibid., pp. 52-72.

Achievement With Respect to Ability or Expected Achievement Model) การตัดสินสัมฤทธิ์ผลของนักเรียน จะยึดการเปรียบเทียบสัมฤทธิ์ผลที่ได้จริงกับสัมฤทธิ์ผลที่คาดหวัง ซึ่งได้จากการวัดความสามารถ (Ability) ของนักเรียนแต่ละคน วิธีการให้ระดับคะแนนตามรูปแบบนี้กระทำได้ใน 2 ลักษณะ คือการใช้สมการถดถอยเชิงเส้นตรง (Linear Regression) และการใช้ระดับความสามารถทั้งหมดของนักเรียนแต่ละคน (Gross Ability Level) เป็นเกณฑ์ในการตัดสินระดับคะแนน

จากการสำรวจของเทอร์วิลลิเจอร์ ในสหรัฐอเมริกา ซึ่งใช้กลุ่มตัวอย่างเป็นอาจารย์โรงเรียนมัธยมศึกษา จำนวน 129 โรงเรียน พบว่า รูปแบบหรือระบบที่นิยมใช้ในการประเมินผลการศึกษาเรียงลำดับจากมากที่สุดไปหาน้อยที่สุดดังนี้¹

รูปแบบสัมฤทธิ์ผลตามความสามารถ	29 %
รูปแบบสัมบูรณ์	27 %
ไม่มีรูปแบบในการให้ระดับคะแนนที่แน่นอน	20 %
รูปแบบตามความสามารถของกลุ่ม	16 %
รูปแบบความมั่งคั่ง	5 %
ไม่มีคำตอบ	3 %

สำหรับผลการวิจัยเกี่ยวกับรูปแบบหรือระบบการประเมินผลการศึกษาในประเทศไทยนั้น เปรมใจ สุขสมานวงศ์² ได้ทำการวิจัยโดยใช้กลุ่มตัวอย่างเป็นอาจารย์ในวิทยาลัยครูในกรุงเทพมหานคร จำนวน 185 คน พบว่า รูปแบบการให้ระดับคะแนนที่อาจารย์วิทยาลัยครูปฏิบัติเรียงลำดับจากมากที่สุด ไปหาน้อยที่สุด ดังนี้

¹James A. Terwilliger, Assigning Grades to Students, p.10.

²เปรมใจ สุขสมานวงศ์, "รูปแบบในการให้ระดับคะแนนแก่นักศึกษาของอาจารย์ในวิทยาลัยครู" (วิทยานิพนธ์ปริญญาโทบัณฑิต ภาควิชาวิจัยการศึกษา บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย, 2523), หน้า 62.

รูปแบบอิงกลุ่ม	50.27 %
รูปแบบอิงเกณฑ์	27.57 %
รูปแบบอิงเกณฑ์และอิงกลุ่ม	16.22 %
รูปแบบอิงความงอกงาม	5.94 %

จากการศึกษาถึงระบบประเมินผลการเรียนที่นิยมใช้มากในประเทศไทยที่ควรศึกษา คือ ระบบประเมินผลการเรียนแบบอิงกลุ่ม แบบอิงเกณฑ์ และแบบอิงเกณฑ์อิงกลุ่ม ดังจะกล่าวในรายละเอียดของแต่ละระบบต่อไปนี้

1. ระบบประเมินผลการเรียนแบบอิงกลุ่ม (Norm-Referenced) เป็นวิธีการประเมินผลที่ขึ้นอยู่กับเกณฑ์มาตรฐานสัมพัทธ์¹ (a relative standard) หรือพฤติกรรมของกลุ่มเป็นสำคัญ เพื่อมุ่งระบุหาความแตกต่างระหว่างนักเรียนและเป็นแรงกระตุ้นสำคัญที่ก่อให้เกิดการเรียนรู้อย่างมีประสิทธิภาพ ระหว่างนักเรียนด้วยกันเอง แบบสอบที่สร้างขึ้นตามแนวคิดหรือปรัชญาการประเมินผลแนวนี้เรียกว่า แบบสอบอิงกลุ่ม ระดับคะแนนที่นักเรียนได้รับจะแสดงถึงระดับสัมฤทธิ์ผล และตำแหน่งสูงสุด เมื่อเทียบกับกลุ่ม คะแนนจากการสอบเมื่อครูสอนจบหน่วยบทเรียนหนึ่ง ๆ คะแนนสอบกลางภาค และคะแนนสอบปลายภาคของนักเรียนแต่ละคนในกลุ่มจะนำพิจารณาให้ระดับคะแนน ซึ่งเทคนิคการให้ระดับคะแนน อาจทำได้ 4 ลักษณะ คือ

1.1 การให้ระดับคะแนนโดยใช้โค้งปกติ ซึ่งมีข้อตกลงเบื้องต้นว่า ความสามารถของนักเรียนในกลุ่ม มีการกระจายเป็นโค้งปกติ วิธีการโดยทั่วไปคือ ครูจะนำคะแนนของนักเรียนแต่ละคนมาหาโค้งปกติ หากค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน ให้ระดับคะแนนตามแต่ละช่วงที่หาได้ในสัดส่วนที่ครูคาดหวัง เช่น²

¹Terwilliger, Assigning Grades to Students, p. 74.

²Ibid., p. 77.

<u>ระดับคะแนน</u>	<u>สัดส่วนที่คาดหวัง</u>	<u>คะแนนที่ปกติ</u>
A	สูงสุด 10 %	1.28
B	รองลงมา 25%	0.39
C	รองลงมา 45%	-0.89
D	รองลงมา 15%	-1.65
F	ต่ำสุด 5%	

หรืออาจนำคะแนนของนักเรียนทุกคนมาเข้าโค้งปกติ (Normal Distribution) แล้วให้บริเวณตรงกลางซึ่งมีคนจำนวนมาก ได้ระดับคะแนน C สำหรับพื้นที่ใต้โค้งที่สูงหรือต่ำกว่านี้จะได้ระดับคะแนน A, B และ D, F ตามลำดับ ซึ่งถ้าเปรียบเทียบออกมาจะได้สัดส่วนดังนี้¹

ระดับคะแนน A	ผู้ใดคะแนนสูงสุด	มีสัดส่วนประมาณ 7 % ของกลุ่ม
ระดับคะแนน B	ผู้ใดคะแนนรองลงมา	มีสัดส่วนประมาณ 23 % ของกลุ่ม
ระดับคะแนน C	ผู้ใดคะแนนปานกลาง	มีสัดส่วนประมาณ 40 % ของกลุ่ม
ระดับคะแนน D	ผู้ใดคะแนนพอใช้	มีสัดส่วนประมาณ 23 % ของกลุ่ม
ระดับคะแนน F	ผู้ใดคะแนนต่ำสุด	มีสัดส่วนประมาณ 7 % ของกลุ่ม

จุดเด่นของวิธีการนี้ จะทำให้นักเรียนเข้าใจสถานภาพของตนเองเมื่อเทียบกับกลุ่ม และครูก็ไม่ต้องกังวลกับเครื่องมือที่ใช้วัดมากนัก

จุดด้อยของวิธีการนี้

1. ถ้านักเรียนมีจำนวนน้อย การกระจายความถี่ของคะแนนความสามารถก็ไม่สามารถจะเป็นโค้งปกติ นอกจากนี้แบบสอบก็มีได้วางแผนอย่างดีในการที่จะได้มา ซึ่งคะแนนที่มีการกระจายเป็นโค้งปกติ²

¹Stodola Quentin and Stordahl Kalmer, Tests and Measurement, (New Delhi : Thomson Press Limited, 1972), p. 48.

²Terwilliger, Assigning Grades to Students, p. 77.

2. นักเรียนส่วนใหญ่ไม่มีความแตกต่างในความสามารถมากนัก เพราะได้ผ่านการคัดเลือกมาแล้ว และถ้าการสอนมีคุณภาพดีเหมาะสมกับสภาพความต้องการของนักเรียน นักเรียนส่วนใหญ่ก็สามารถบรรลุตามจุดมุ่งหมายของการสอน ซึ่งจะทำให้การแจกแจงไม่เป็นโค้งปกติ¹ ในกรณีเช่นนี้ การตัดสินระดับคะแนนไม่จำเป็นต้องมีทั้ง 5 ระดับคะแนน อาจจะมี 2, 3 หรือ 4 ระดับคะแนน ตามความเหมาะสม

3. วิธีการนี้เน้นการแข่งขันกัน ซึ่งไม่เหมาะในด้านทัศนการศึกษา² ซึ่งเป็นการเรียนเพื่อรอบรู้

4. มาตรฐานแปรตามระดับเขาวนปัญญาของกลุ่มซึ่งไม่คงที่ นักเรียนอาจช่วยเหลือซึ่งกันและกัน ทำให้ครูตัดสินระดับคะแนนนักเรียนแต่ละคนไม่สอดคล้องกับความสามารถที่เป็นจริง³

1.2 การให้ระดับคะแนนโดยใช้การหาค่าแห่งเปอร์เซ็นต์ไทล์ วิธีการโดยทั่วไปครูผู้สอนจะนำคะแนนทั้งหมดที่ได้มาเรียงกันเป็นช่วง จากช่วงต่าง ๆ ครูจะเลือกคะแนนมาเพียง 10-12 จำนวน (แต่ละช่วงเลือกมา 1 จำนวน ซึ่งอาจจะไม่ครบทุกช่วง) แล้วนำคะแนน 10-12 จำนวนนั้นมาหาค่าแห่งเปอร์เซ็นต์ไทล์ โดยใช้สูตร⁴

$$\text{ค่าแห่งเปอร์เซ็นต์ไทล์} = 100 \left[\frac{cf + \frac{X - LRL}{i} fw}{N} \right]$$

¹ Benjamin S. Bloom, J. Thomas Hastings, George F. Madaus, Handbook on Formative and Summative Evaluation of Student Learning, p.45.

² Quentin and Kalmer, Tests and Measurement, p. 47.

³ Green, Introduction to Measurement and Evaluation, p. 319.

⁴ Terwilliger, Assigning Grades to Students, p. 145.

เมื่อ	cf	แทน	ความถี่สะสมจากต่ำสุดถึงช่วงที่ต้องการหา
	X	แทน	คะแนนที่ต้องการหาค่าแห่งเปอร์เซ็นต์ไทล์
	LRL	แทน	ขีดจำกัดล่างแท้จริงของช่วงที่มีคะแนนที่ต้องการหาค่าแห่งเปอร์เซ็นต์ไทล์
	i	แทน	อินตรภาคชั้น
	fw	แทน	ความถี่ภายในช่วงคะแนน X
	N	แทน	จำนวนคนทั้งหมดที่ต้องการหา

นำค่าของคะแนนและค่าแห่งเปอร์เซ็นต์ไทล์ที่ได้มาเขียนกราฟ จะได้กราฟฟังก์ชันสะสม จากกราฟครูก็สามารถหาค่าแห่งเปอร์เซ็นต์ไทล์ของคะแนนอื่น ๆ ได้หมด หลังจากนั้นครูก็จะกำหนดควาถี่คส่วนของนักเรียนในแต่ละระดับคะแนนควรเป็นเท่าใด แล้วประเมินค่าออกมาว่าช่วงค่าแห่งเปอร์เซ็นต์ไทล์เท่าใด จะได้ระดับคะแนนอะไร นอกจากนี้อาจหาค่าแห่งเปอร์เซ็นต์ไทล์ที่ไม่มีการแบ่งช่วงได้จากสูตร¹

ค่าแห่งเปอร์เซ็นต์ไทล์ =

$$100 \left[\frac{\text{จำนวนคนที่ได้คะแนนต่ำกว่าคะแนนที่ต้องการหา} + \frac{1}{2} \text{ ของจำนวนคน ณ ค่าแห่งนั้น}}{N} \right]$$

เมื่อ N แทน จำนวนคนทั้งหมด

จุดเด่นของวิธีการนี้ จะแสดงให้เห็นจำนวนนักเรียนว่ามีประมาณกี่เปอร์เซ็นต์ที่อยู่ต่ำกว่าระดับคะแนนนั้นได้อย่างชัดเจน

จุดด้อยของวิธีการนี้คือ ถ้ากลุ่มมีขนาดเล็กอาจจะทำให้การแปลความหมายของระดับคะแนนผิดพลาดได้ เนื่องจากนักเรียนแต่ละคนจะต้องมีระดับคะแนนแตกต่างกัน ทั้ง ๆ

¹C. Mauritz Lindvall and Anthony J. Nitko, Measuring Pupil Achievement and Aptitude, 2d ed. (New York : Harcourt Brace Jobanovick Inc., 1975), p.75.

ที่ความสามารถอาจจะใกล้เคียงกัน

1.3 การให้ระดับคะแนน โดยใช้คะแนนมาตรฐาน ระดับคะแนนที่นักเรียนได้รับโดยวิธีการนี้ จะแสดงถึงว่ามีนักเรียนอยู่สูงหรือต่ำกว่าระดับคะแนนเฉลี่ยมากน้อยเพียงใด วิธีการนี้ครูจะกำหนดค่าเฉลี่ย 50 และส่วนเบี่ยงเบนมาตรฐาน 10 แล้วนำคะแนนทั้งหมดมาแปลงโดยใช้สูตร¹

$$T = 50 + 10 \left[\frac{\text{คะแนนดิบ} - \text{คะแนนเฉลี่ย}}{\text{ส่วนเบี่ยงเบนมาตรฐาน}} \right]$$

เมื่อแปลงเป็นคะแนนทั้งหมด ครูก็จะหาช่วงระดับคะแนนแต่ละช่วงว่าควรห่างกันเท่าใด ซึ่งอาจใช้สัดส่วนหรือหาช่วงห่างแต่ละระดับคะแนน แล้วกำหนดว่าระดับคะแนน A, B, C, D และ F จะมีคะแนนที่อยู่ในช่วงที่เท่าไรถึงเท่าไร

จุดค้อยของวิธีการนี้คือ จะต้องมีย่านนักเรียนมากพอสมควร เนื่องจากการแปลงเป็นคะแนนที่นี้เป็นที่ยอมรับว่า คะแนนของนักเรียนมีการกระจายเป็นโค้งปกติ

1.4 การให้ระดับคะแนนโดยใช้คะแนนดิบ วิธีการที่สำคัญคือครูจะให้คะแนนดิบเต็มเท่ากันทุกกลุ่ม ถ้าคะแนนดิบแต่ละกลุ่มไม่เท่ากัน จะต้องปรับคะแนนดิบให้เท่ากันโดยใช้สูตร²

$$\text{คะแนนที่ปรับแล้ว} = \frac{\text{คะแนนเต็มของคะแนนที่ต้องการปรับ}}{\text{คะแนนเต็มของคะแนนที่จะปรับ}} \times \text{คะแนนดิบ}$$

ตัวอย่าง ครูแดงทดสอบแบบสอบให้คะแนนเต็ม 200 คะแนน แต่ครูอื่น ๆ ที่สอนวิชาเดียวกันให้คะแนนเต็ม 250 คะแนน ถ้านายเขียนเรียนกับครูแดงได้คะแนน 180 คะแนน ครูแดงจะปรับคะแนนของนายเขียน ได้คะแนนที่ปรับ $= \frac{250}{100} \times 180 = 225$ คะแนน การปรับคะแนนดังกล่าวอยู่ภายใต้ข้อตกลงที่ว่า จะมีความแตกต่างในการวัดเฉพาะส่วนที่เป็นคะแนนเต็มเท่านั้น เมื่อใดคะแนนเท่ากันทุกกลุ่ม ครูจะนำคะแนนดิบมาแบ่งเป็นช่วง เพื่อหา

¹Terwilliger, Assigning Grades to Students, p. 93.

²Ibid., p. 86.

ความดีแต่ละช่วง จากนั้นครูก็จะหาคะแนนจุดตัด (Cutting Score) โดยกำหนดสัดส่วนของนักเรียนที่ควรจะได้ในแต่ละระดับคะแนน แล้วคำนวณหาช่วงคะแนนดิบที่สอดคล้องกับสัดส่วน เพื่อให้ได้ทราบระดับคะแนนของนักเรียนแต่ละคน

จุดเด่นของวิธีการนี้คือ สะดวกในการใช้สำหรับครูที่มีการวางแผนร่วมกันก่อนที่จะทำการสอน และไม่ต้องทำงานหนักถึงแม้จะสอนนักเรียนจำนวนมาก

จุดอ่อนของวิธีการนี้คือ ในกรณีที่มีครูสอนหลายคน และใช้คะแนนเต็มแตกต่างกัน การปรับคะแนนอาจทำให้การตัดสินผิดพลาด เนื่องจากข้อตกลงที่ว่าจะให้มีความแตกต่างกัน เฉพาะส่วนที่เป็นคะแนนเต็มเท่านั้น ในทางปฏิบัติทำได้ยาก

1.5 การให้ระดับคะแนนตามความสามารถของกลุ่ม ซึ่งวิธีการนี้เสนอโดย สตีวอิด¹ (Stuit) เพื่อลดปัญหาการให้ระดับคะแนนโดยใช้โค้งปกติ ซึ่งแต่ละระดับคะแนนจะมีการกำหนดสัดส่วนตายตัว ซึ่งในทางปฏิบัติจริงนักเรียนแต่ละกลุ่มจะมีความสามารถไม่เท่ากัน การที่จะให้ระดับคะแนนทุกกลุ่มเหมือนกันคือมี 5 ระดับคะแนน (A, B, C, D และ F) ความไม่ยุติธรรมก็จะเกิดขึ้น คือนักเรียนที่ได้คะแนนปานกลาง ถ้าอยู่ในกลุ่มอ่อนอาจจะได้ระดับคะแนน A หรือ B แต่ถ้าอยู่ในกลุ่มเก่งเขาอาจได้ C หรือ D ดังนั้นจึงมีการปรับระดับคะแนนตามความสามารถของกลุ่มโดยแบ่งความสามารถของกลุ่มออกเป็น 7 ระดับ โดยอาศัยแต้มเฉลี่ย (Grade Point Average) เป็นตัวกำหนดว่า กลุ่มที่ต้องการประเมินผลมีความสามารถในระดับใด แล้วหาค่ามัธยฐาน และส่วนเบี่ยงเบนมาตรฐาน ก็สามารถกำหนดคะแนนในแต่ละระดับของกลุ่มได้โดยใช้สูตร²

คะแนนต่ำสุดของระดับใด ๆ =
 มัธยฐาน + (ซีกจำกัดกลางของระดับคะแนนนั้น × ส่วนเบี่ยงเบนมาตรฐาน)
 ซึ่งค่าซีกจำกัดกลางของระดับคะแนน จะหาได้จากตารางที่ 1

¹Robert L. Ebel, Measuring Educational Achievement,
 (New Jersey: Prentice-Hall Inc., 1965), pp.426-435.

²Ibid., p. 430.

ตารางที่ 1 การแจกแจงระดับคะแนนสำหรับกลุ่มที่มีความสามารถแตกต่างกัน
7 ระดับ¹

ระดับ ความสามารถ	แต้มเฉลี่ย	%ile	ขีดจำกัดกลางของ ระดับคะแนน	รอยละของระดับคะแนน				
				A	B	C	D	F
ดีเลิศ	2.80	79	0.7	24	38	29	8	1
ดีมาก	2.60	73	0.9	18	36	32	12	2
ดี	2.40	66	1.1	14	32	36	15	3
ดีพอใช้	2.20	58	1.3	10	29	37	20	4
ปานกลาง	2.00	50	1.5	7	24	38	24	7
อ่อน	1.80	42	1.7	4	20	37	29	10
อ่อนมาก	1.60	34	1.9	3	15	36	32	14

วิธีการนี้ได้รับการรับรองจาก เอ็ดดูเคชันแนล โพลีซี คอมมิที (Educational Policy Committee) เมื่อเดือนพฤศจิกายน ค.ศ.1954

จุดค้อยของวิธีการนี้คือ

1. มีปัญหาในด้านการกำหนดความสามารถของกลุ่มว่าจะอาศัยผลการสอบวิชาใดเป็นหลัก
2. การกำหนดระดับคะแนนไว้ก่อนว่าผลการเรียนของกลุ่มจะได้ระดับคะแนนอะไร เป็นวิธีการที่ไม่เหมาะสมกับหลักการ เรียนรู้ที่จะส่งเสริมพัฒนาการของผู้เรียนแต่ละคน

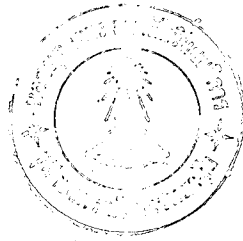
¹Robert L. Ebel, Measuring Educational Achievement,
p. 428.

2. ระบบประเมินผลการเรียนแบบอิงเกณฑ์ (Criterion-Referenced)

เป็นวิธีการประเมินผลที่ขึ้นอยู่กับเกณฑ์มาตรฐานสัมบูรณ์¹ (Absolute Standard) ที่ให้ความหมายไว้แน่ชัดล่วงหน้า เพื่อมุ่งระบุหาว่านักเรียนสามารถหรือไม่สามารถทำอะไรได้บ้าง และมุ่งผลักดันให้เกิดการเรียนรู้แบบเกื้อกูลซึ่งกันและกัน แบบสอบที่สร้างตามแนวคิดนี้เรียกว่า แบบสอบอิงเกณฑ์ ซึ่งการสร้างแบบสอบอิงเกณฑ์เน้นที่ความสอดคล้องระหว่างวัตถุประสงค์ของการสอนกับเนื้อเรื่องในแบบสอบ ทำให้คะแนนที่ได้จากแบบสอบอิงเกณฑ์มีความแปรปรวนน้อยกว่าคะแนนที่ได้จากแบบสอบอิงกลุ่ม ดังนั้นการคาดคะเนความเป็นเนื้อเดียวกัน และความคงที่ของแบบสอบควยวิธีคลาสสิก หรือวิธีประเพณีนิยม จึงมีค่าต่ำผิดปกติ และอำนาจจำแนกของแต่ละข้อสอบอิงเกณฑ์มีค่าต่ำผิดปกติ การประเมินผลแบบนี้เมื่อครูสอนจบแต่ละหน่วยบทเรียน จะมีทดสอบย่อยเมื่อควาว่านักเรียนมีความรู้พื้นฐานที่จำเป็นในขั้นที่ใช้การได้พอเพียงหรือไม่ โดยเทียบกับเกณฑ์ที่บอกให้นักเรียนทราบล่วงหน้า เช่น เกณฑ์ 80 % ของวัตถุประสงค์ของการสอนที่วัดโดยแบบสอบอิงเกณฑ์เป็นต้น ซึ่งถ้าทดสอบแล้วนักเรียนคนใดได้คะแนนไม่ถึง 80 % ก็แสดงว่านักเรียนคนนั้นยังไม่บรรลุซึ่งวัตถุประสงค์ขั้นต่ำ ครูจะทำการสอนซ่อมเสริมให้หรือมีการปรับปรุงการเรียนการสอน แล้วมีการทดสอบหน่วยบทเรียนนั้นใหม่ จนนักเรียนผ่านเกณฑ์ที่ตั้งไว้แล้วจึงเรียนหน่วยบทเรียนต่อไป ครูซึ่งใช้การประเมินผลแบบนี้จะเน้นเนื้อหาวิชาที่สอนเป็นสำคัญ และต้องตั้งวัตถุประสงค์ในการสอนอย่างชัดเจน แล้วตั้งเกณฑ์มาตรฐานของการผ่านในแต่ละหน่วยบทเรียน และเกณฑ์มาตรฐานของแต่ละระดับคะแนนในตอนปลายเทอม เพื่อบ่งชี้ระดับสัมฤทธิ์ผลในวิชานั้นซึ่งแสดงถึงระดับความสามารถและทักษะของนักเรียน จากคะแนนสอบกลางภาคและปลายภาค จึงนับได้ว่าเป็นประโยชน์มากในการเรียนเพื่อรอบรู้² การให้ระดับคะแนนนั้น อาจกระทำได้ใน 3 ลักษณะคือ

¹Terwilliger, Assigning Grades to Students, p. 26.

²Grondlund, Measurement and Evaluation in Teaching p.527.



2.1 วิธีใช้คะแนนเปอร์เซ็นต์ ซึ่งอาจเป็นคะแนนเปอร์เซ็นต์จากความรอบรู้
ในเนื้อหาวิชาตามวัตถุประสงค์ที่วางไว้ หรือคะแนนเปอร์เซ็นต์ของคำตอบที่ถูกต้อง วิธีการนี้
จะกำหนดคะแนนเต็มในแต่ละงานที่มอบหมายให้ทำ หรือให้คะแนนรวมตลอดภาคเป็นร้อยละ
คะแนน และครูก็จะตั้งเกณฑ์มาตรฐานความสามารถหรือความสมบูรณ์ในการตอบของแต่ละ
ระดับคะแนน เกณฑ์ที่ยอมรับได้ (ระดับคะแนน C) ซึ่งเป็นที่ยอมรับโดยทั่วไปอยู่ในช่วง
75-84 % เพราะถือว่านักเรียนที่มีความสามารถระดับนี้จะเป็นผู้รอบรู้ในเนื้อหาวิชา¹ ระดับ
คะแนนที่ได้จะแสดงออกมาในรูปของคะแนนเปอร์เซ็นต์ เช่น 95 % ขึ้นไป = A, 85 %
ถึง 94 % = B, 75 % ถึง 84 % = C, 65 % ถึง 74 % = D, ต่ำกว่าหรือเท่ากับ 65 %
= F

จุดเด่นของวิธีการนี้คือทำให้นักเรียนมีความรอบรู้ในเนื้อหาวิชามาก
จุดด้อยของวิธีการนี้คือ ถ้าเครื่องมือที่นำไปใช้วัดมิได้วางแผนในการสร้าง
อย่างดี ผลที่ได้จะไม่มีความหมายอะไร และเกณฑ์ที่วางไว้ก็ถือว่าไม่เป็นมาตรฐาน นอกจากนี้
ถ้ากำหนดมาตรฐานความสามารถสูงหรือต่ำเกินไป ระดับคะแนนที่ได้ก็จะไม่มีความหมาย และ
การกำหนดมาตรฐานให้เหมาะสมจะต้องขึ้นกับลักษณะของนักเรียนตามความเห็นของครู²

2.2 วิธีใช้หน่วยธรรมชาติของการปฏิบัติ วิธีการนี้มักใช้กับการวัดในวิชาที่
เน้นทักษะ ซึ่งจะช่วยลดปัญหาความคลาดเคลื่อนและความลำเอียงของครู โดยครูจะนำทักษะ
มากำหนดเป็นเกณฑ์ตัดสินอย่างชัดเจน เช่น วิชาพิมพ์ดีดจะจัดออกมาเป็นจำนวนคำต่อนาที
วิชาทางพลศึกษาถ้าเป็นการวิ่งก็จะวัดออกมาเป็นระยะทางต่อนาที เป็นต้น หลักการของวิธีนี้
คือครูจะต้องตั้งเกณฑ์มาตรฐานตามทักษะที่ควรจะเป็นในแต่ละระดับอย่างเหมาะสม แล้ววัดทักษะ
ของนักเรียน ซึ่งจะทำให้ตัดสินระดับคะแนนให้นักเรียนแต่ละคนได้ ระดับคะแนนที่ได้จะ
แสดงถึงความสามารถในทักษะของนักเรียนตามเกณฑ์มาตรฐาน

¹Grondlund, Measurement and Evaluation in Teaching, p.528.

²Green, Introduction to Measurement and Evaluation, p.315.

จุดเด่นของวิธีการนี้คือ ลดปัญหาความลำเอียง และความคลาดเคลื่อนจากการประมาณควยสายตา

จุดด้อยของวิธีการนี้คือ ไม่สามารถนำมาใช้วัดทางการศึกษา เพราะเครื่องมือที่ใช้วัดทางการศึกษา เช่น ข้อสอบ แบบฝึกหัดมีความไม่คงที่¹

3. ระบบประเมินผลการเรียนแบบอิงเกณฑ์อิงกลุ่ม² (Norm and Criterion-Referenced) เป็นวิธีการประเมินผลที่พิจารณาจากเกณฑ์มาตรฐานสัมพัทธ์ และจากเกณฑ์มาตรฐานสัมบูรณ์ มุ่งกระตุ้นและส่งเสริมให้นักเรียน เกิดการเรียนรู้มากที่สุดเท่าที่จะทำได้ อันเป็นเป้าหมายที่สำคัญของการสอน คือ มีการประเมินผลบ่อยครั้ง เพื่อให้ข้อมูลย้อนกลับ (Feed-back) แก่นักเรียนและครู ในอันที่จะปรับปรุงการเรียนการสอนให้มีประสิทธิภาพยิ่งขึ้น การประเมินผลระบบนี้มุ่งสนับสนุนให้มีการเรียนรู้ทั้งแบบเกื้อกูลและแข่งขันกัน การเรียนรู้แต่ละประเภทก็มีจุดเด่นในตัว คือ การเรียนรู้แบบเกื้อกูลจะช่วยปลูกฝังให้นักเรียนรู้จักทำงานร่วมกัน เป็นผู้มีความสามารถในการประสานงานได้ดี ส่วนการเรียนรู้แบบแข่งขันปลูกฝังให้นักเรียนมีลักษณะเป็นตัวของตัวเอง ทำงานได้เป็นอิสระ ซึ่งนักเรียนระดับชั้นมัธยมศึกษาตอนปลายน่าจะควรมีลักษณะบุคลิกภาพทั้งสองประการดังกล่าว การประเมินผลวิธีนี้ให้ความยุติธรรมแก่นักเรียนคือ นักเรียนจะทราบเป็นการล่วงหน้าอย่างชัดเจนว่าตนเองจะถูกประเมินในแง่ใดบ้าง ตลอดจนแนะนำกิจกรรมการเรียนรู้ที่ปฏิบัติว่ามีอะไรบ้าง และการประเมินผลวิธีนี้พยายามเชื่อมความสัมพันธ์ระหว่างความรู้ที่นักเรียนรู้ กับเกรด และคะแนนซึ่งเป็นดัชนีชี้ระดับสัมฤทธิ์ผลควย

วิธีประเมินผลการเรียนนักเรียนแบบอิง เกณฑ์อิงกลุ่มประกอบด้วย

1. แบ่งเนื้อเรื่องของวิชาออกเป็นหน่วยบทเรียนได้อย่างมีความหมาย และสมเหตุสมผล ตลอดจนมีความเป็นเนื้อเดียวกัน

¹Terwilliger, Assigning Grades to Students, p. 30.

²สมหวัง พิธิยานุวัฒน์, ระบบประเมินการเรียนนิสิตตามแนวคิดการวัดอิงกลุ่มและการวัดอิงเกณฑ์, หน้า 3-6.

2. ตั้งวัตถุประสงค์เชิงพฤติกรรมประจำหน่วยบทเรียนนั้น ๆ
3. สร้างแบบสอบอิงเกณฑ์เพื่อตรวจสอบว่านักเรียนรู้หรือไม่รู้เนื้อเรื่องใดบ้าง หรือกล่าวอีกนัยหนึ่งหลังจากการเรียนแล้ว นักเรียนบรรลุวัตถุประสงค์ที่จำเป็นในแต่ละหน่วยบทเรียนเพียงใด
4. ประเมินว่านักเรียนมีความรู้พื้นฐานที่จำเป็นในขั้นที่ใช้การได้พอเพียงหรือไม่ โดยเทียบกับเกณฑ์ที่บอกให้นักเรียนทราบล่วงหน้า เช่น 80 เปอร์เซ็นต์ ของคะแนนเต็ม จากแบบสอบอิงเกณฑ์ เป็นต้น

หลังจากการประเมินขั้นนี้แล้วจะแบ่งนักเรียนออกเป็น 2 กลุ่ม คือ กลุ่มที่มีความรู้พื้นฐานจำเป็นอย่างพอเพียง ซึ่งถือว่าเป็นพวกที่สอบได้ กับกลุ่มที่ขาดความรู้พื้นฐานที่จำเป็น ซึ่งถือว่าเป็นพวกที่สอบตก

5. สร้างแบบสอบถามแนวความคิดรวบยอด แล้วใช้คะแนนจากแบบสอบอิงกลุ่มเป็นดัชนีถึงระดับสัมฤทธิ์ผลในการเรียนวิชานั้น ๆ ของนักเรียนที่สอบผ่านจากแบบสอบประจำหน่วยบทเรียน โดยมีความเชื่อว่านักเรียนย่อมจะมีระดับสัมฤทธิ์ผลขั้นสูงที่ซับซ้อนแตกต่างกันออกไป โดยทั่วไปในการให้แก่นักเรียนกลุ่มนี้ เราก็อาจจะพิจารณาจากการจับกลุ่มของคะแนนตามธรรมชาติเป็นหลัก หรือใช้วิธีแบ่งกลุ่มตามหลักสถิติ แล้วแต่ความเหมาะสม ในกรณีนักเรียนคาบเส้นก็ควรใช้ข้อมูลจากการสังเกต จากแบบฝึกหัดและอื่น ๆ ประกอบการพิจารณาด้วย สำหรับรายละเอียดของระบบประเมินผลอิงเกณฑ์อิงกลุ่ม จำแนกตามจุดมุ่งหมาย พร้อมทั้งวิธีการเก็บข้อมูลดังแสดงในตารางต่อไปนี้

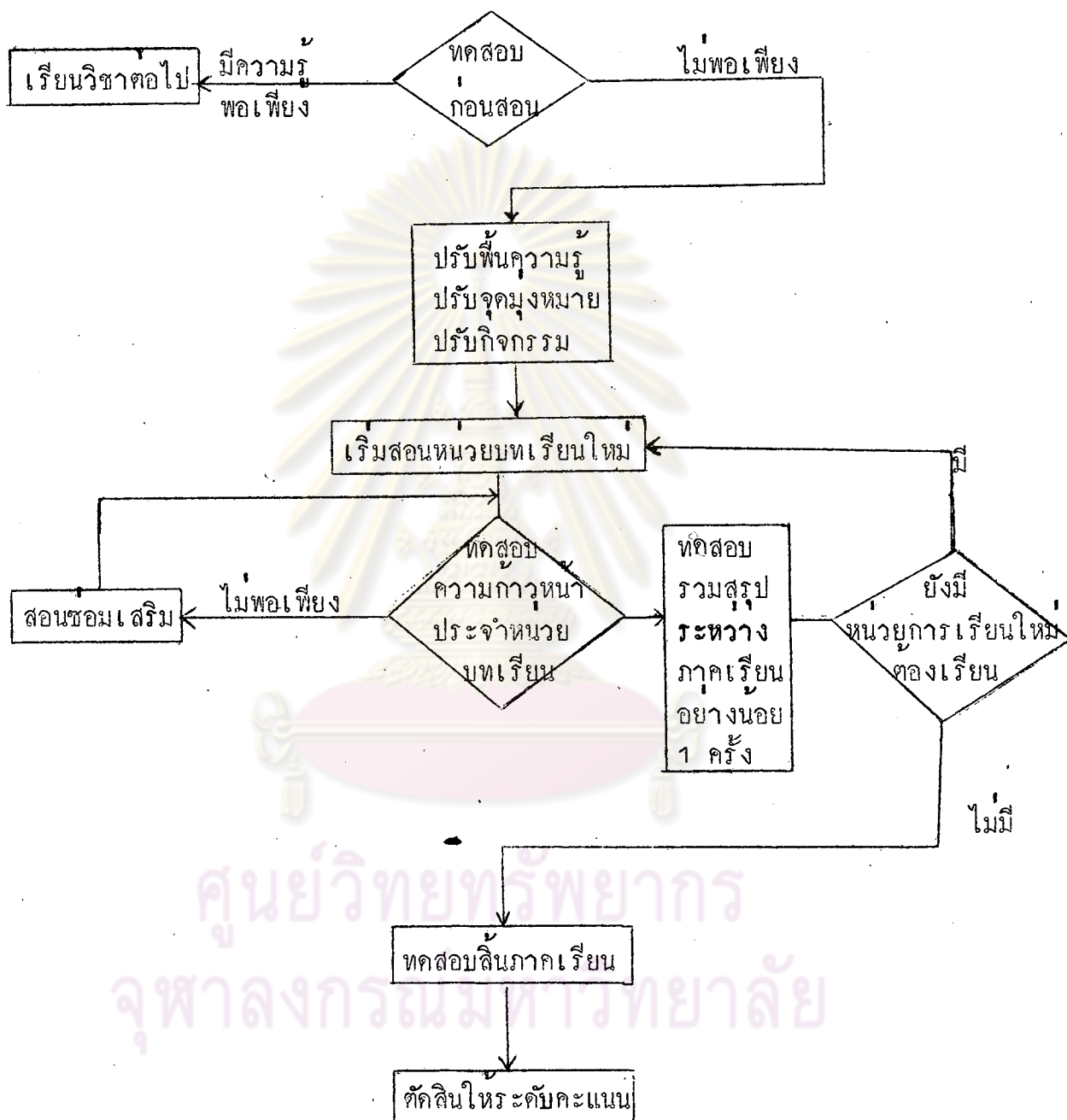
ตารางที่ 2 ระบบประเมินผลการเรียนแบบอิงเกณฑ์อิงกลุ่ม จำแนกตามจุดมุ่งหมายของการประเมิน

เพื่ออะไร อย่างไร	ก่อนสอน	ระหว่างสอน	หลังสอน
จุดมุ่งหมาย	1. บ่งชี้พื้นฐานความรู้เดิม	1. ความรู้และทักษะพื้นฐาน 2. จุดเด่นจุดด้อย	1. บ่งชี้ระดับผลสัมฤทธิ์ หรือผลการเรียนรู้
เครื่องมือหรือ วิธีเก็บข้อมูล	1. ข้อสอบ 2. การสังเกต 3. แบบสอบถาม	1. ข้อสอบอิง-เกณฑ์ 2. แบบฝึกหัด 3. แบบสอบถาม 4. การสังเกต 5. การสัมภาษณ์	1. ข้อสอบอิงกลุ่ม 2. รายงาน
การตัดสินใจ	1. ปรับปรุงจุดมุ่งหมาย ของการสอน 2. ปรับกิจกรรมการสอน	1. ปรับวิธีการสอน 2. ปรับวิธีการเรียน	1. อันคับคะแนน 2. การเลือกวิชา

เป็นที่น่าสังเกตว่า ในระบบการประเมินผลการเรียนดังกล่าว ได้เสนอแนะวิธีการเก็บข้อมูลไว้หลายวิธี เพื่อจะได้เลือกใช้ตามความเหมาะสม อีกประการหนึ่ง วิธีการเก็บข้อมูลแต่ละวิธีก็มีข้อจำกัด ตลอดจนจุดเด่นแตกต่างกันไป ในทางการศึกษาที่เรามุ่งจัดผลการเรียน หรือการเปลี่ยนแปลงในเชิงนามธรรม เราจำเป็นต้องวัดด้วยวิธีหลายวิธี และแต่ละวิธีก็ต้องใช้วัดหลายครั้ง

โดยหลักการแล้วการประเมินผลการเรียนจึงต้องสอดคล้องแก่กิจกรรมการเรียน กล่าวคือ การประเมินผลการเรียนเป็นส่วนหนึ่งของการสอน ดังนั้นการผสมผสานระบบการประเมินผลการเรียนกับการสอน ดังแสดงในแผนภาพการสอน และการประเมินผลการเรียนแบบอิงเกณฑ์อิงกลุ่มดังต่อไปนี้

แผนภาพที่ 1: แบบการสอนของการประเมินผลการเรียนแบบอิงเกณฑ์อิงกลุ่ม¹



¹สมหวัง พิธิยานุวัฒน์, และการวัดอิงเกณฑ์, หน้า 6.

ตัวอย่างการประเมินผลการเรียนแบบอิงเกณฑ์อิงกลุ่มนี้ สมหวัง พิธิยานุวัฒน์¹ ได้สอนวิชา 412 601 สถิติศาสตร์ประยุกต์ทางพฤติกรรมศาสตร์ ซึ่งเป็นวิชาในระดับปริญญาโท ประจำภาคเรียนที่ 2 ปีการศึกษา 2519 ภาควิชาวิจัยการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย โดยใช้ระบบประเมินผลการเรียนแบบอิงเกณฑ์อิงกลุ่ม ซึ่งปรากฏว่าระบบประเมินผลการเรียนแบบนี้เป็นที่พอใจแก่นิสิตผู้เรียนมาก โดยให้เหตุผลว่าระบบประเมินผลการเรียนแบบนี้ ช่วยกระตุ้นให้เรียนรู้อย่างสม่ำเสมอ โดยให้ข้อมูลย้อนกลับว่าตนเองมีความรู้หรือไม่มีความรู้ในเรื่องใ้ใ้ใ้ และนิสิตบางคนเห็นว่าการประเมินผลการเรียนแบบนี้ยุติธรรมกว่าระบบประเมินผลการเรียนแบบอื่น ๆ ที่ตนประสบมา

ประคอง กรรณสูต² ให้ความเห็นว่า การตัดสินคะแนนจะอิงเกณฑ์หรืออิงกลุ่มก็ตาม ผู้ตัดสินคะแนนจะต้องคิดให้รอบคอบ พิจารณาหาวิธีที่เหมาะสมที่สุด อย่าให้คะแนนหรือการตัดสินของท่านเป็นเครื่องบั่นทอนกำลังใจ และความสามารถของผู้เรียน ในการประเมินผล ถ้าอาจารย์ผู้สอนจะพยายามตัดสินคะแนนทั้งอิงกลุ่มและอิงเกณฑ์ ก็จะเป็นการกระตุ้นให้ผู้เรียน ทั้งแบ่งและช่วยกัน

ในทัศนะของนักการศึกษาต่างประเทศ ที่มีต่อระบบประเมินผลการเรียนแบบอิงกลุ่มและแบบอิงเกณฑ์นั้น มิลแมน³ (Millman) ได้ใช้คำว่าโคเมนแทนเกณฑ์ และกล่าวว่า การสอบอิงเกณฑ์เป็นการวัดสภาพที่เป็นปัจจุบันของนักเรียนคนใดคนหนึ่ง โดยอ้างอิงจุดของ

¹สมหวัง พิธิยานุวัฒน์, ระบบการประเมินการเรียนนิติตตามแนวคิดการวัดอิงเกณฑ์และการวัดอิงกลุ่ม, หน้า 7-9.

²ประคอง กรรณสูต, วิรัตน์ ชรรมาภรณ์. การจัดระดับคะแนนเป็นตัวอักษร. (พระนคร : โรงพิมพ์ไทยวัฒนาพานิช, 2523), หน้า 2.

³Jason Millman, "Criterion-Referenced Measurement," in Evaluation in Education, ed. W.J. Popham (Berkeley : McCutchan Publishing Corporation, 1974), p. 327.

งานที่ทำที่ได้บรรยายไว้อย่างดีแล้วที่เรียกว่าโคเมน เกลเซอร์และนิทโก¹ (Glaser and Nitko) กล่าวว่า แบบสอบอิงเกณฑ์ เป็นแบบสอบที่สร้างขึ้นโดยมีเจตนาให้เป็นการวัดที่สามารถแปลความหมายออกมาได้ตรง ๆ ในรูปของมาตรฐานการกระทำที่เฉพาะ ซึ่งในที่นี้ บ่งชี้ได้โดยให้นิยามโคเมนของงานที่นักเรียนหรือผู้สอบควรทำได้ จากนั้นนำตัวอย่างที่เป็นตัวแทนของงานจากโคเมนนี้มาจัดเป็นแบบสอบ ผลจากการวัดจะแสดงถึงพฤติกรรมของแต่ละคน เพื่อเทียบกับโคเมนที่นิยามไว้ ธอร์นไดค์² (Thorndike) ให้ข้อสังเกตไว้ว่า ในการวัดแบบอิงกลุ่มสิ่งแรกที่ต้องการทราบคือ นักเรียนทำงานที่กำหนดได้คือเพียงใด ในขณะที่สิ่งแรกที่ต้องการทราบในการวัดแบบอิงเกณฑ์คือ งานยากระดับใดที่นักเรียนสามารถทำได้ ไฮเปลี (Hively) ได้กล่าวอย่างกว้าง ๆ ว่าการวัดแบบอิงกลุ่มมีรากฐานอยู่ที่ความแตกต่างระหว่างบุคคล ให้ข้อมูลที่เป็นประโยชน์ในการทำนายและการเลือก ในขณะที่การวัดแบบอิงเกณฑ์ ซึ่งมีรากฐานบนทฤษฎีการเรียนรู้ให้ข้อมูลที่เป็นประโยชน์การประเมินความก้าวหน้า

¹R. Glaser and A.J. Nitko, "Measuring in Learning and Instruction," In Educational Measurement, ed. R.L. Thorndike (Washington : American Council on Education, 1971), p. 653.

²P.W. Airasian and George F. Madaus, "Criterion-Referenced Testing in the Classroom," in Crucial Issues in Testing, ed. R.W. Tyler and R.M. Wolf (California : McCutchan Publishing Corporation, 1974), p. 74. Citing E.L. Thorndike, "The Nature, Purposes and General Methods of Measurements of Educational Products," in The Measurement of Educational Products, Seventeenth Yearbooks of the National Society for the Study of Education, Part II (Bloomington, III.: Public School Publishing Company, 1918), pp.16-24.

ฟิตซกิบบอน (Fitzgibbon), คอกซ์ (Cox), คลีน (Klien)¹ มีความเห็นว่า วิธีการวัดและประเมินผลการเรียนแบบอิงกลุ่มและแบบอิงเกณฑ์นั้นในชั้นมโนคติพื้นฐานไม่ต่างกัน เช่น ในด้านวัตถุประสงค์ การสร้างและการพัฒนาเครื่องมือ แต่ต่างกันที่การแปลความหมายของคะแนนเท่านั้น แต่โบห์ม (Boehm), พอปแฮม (Popham), เกลเซอร์ (Glaser), ฮูเซ็ค (Husek) และนิทโก (Nitko)² มีความเห็นว่า วิธีการวัดและประเมินผลแบบอิงกลุ่มและอิงเกณฑ์ต่างกันอย่างสิ้นเชิงทั้งในด้านมโนคติพื้นฐานและการแปลความหมาย

ระบบการประเมินผลการเรียนแบบอิงกลุ่มและอิงเกณฑ์นั้นได้นิยมใช้กันแพร่หลายมานานแล้ว แต่ในระยะทศวรรษที่แล้วนี้ นักวัดและประเมินผลการศึกษาหลายท่านได้หันมานิยมใช้ระบบประเมินผลการเรียนแบบอิงเกณฑ์กันมากขึ้น โดยมีแรงจูงใจจากแนวโน้ม 4 ประการ อันเนื่องมาจากการใช้แบบสอบแบบอิงกลุ่มในการวัดและประเมินผลการศึกษา ดังที่แอร์ราเขียนและมาคอส³ กล่าวไว้คือ

1. การวิพากษ์วิจารณ์ถึง เรื่องความหมายของการวัด รวมทั้งการใช้แบบสอบมาตรฐานที่ให้ข้อมูลไม่เพียงพอและไม่เหมาะสมกับเนื้อหาบางประการที่ต้องการวัด
2. การวิพากษ์วิจารณ์เกี่ยวกับระดับคะแนน (Grade) ว่า ทำให้จุดหมายของการศึกษาเปลี่ยนไป กลายเป็นการแข่งขันกันเพื่อให้ได้ระดับคะแนนสูง ๆ และกลายเป็นสิ่งตอรองในการเข้าศึกษา การสมัครงาน ฯลฯ จนลงควยวาระระดับคะแนนไม่ให้ข้อมูลที่ เด็กทำอะไรได้บ้าง บอกแต่ว่าเขาดีกว่าหรือค้อยกว่าผู้อื่นในกลุ่มของเขาเท่านั้น
3. จากวิวัฒนาการการเคลื่อนไหวทางเทคนิคการสอนต่าง ๆ ที่ทำให้พบว่า แบบสอบอิงกลุ่มไม่เพียงพอในการที่จะใช้ประเมินพฤติกรรมเป็นรายบุคคลหรือแม้แต่วิธีการสอนที่ใช้ว่าดีหรือไม่เพียงใด

¹สมหวัง พิธิยานุวัฒน์, "หลักการวัดและประเมินผล" คู่มืออาจารย์ดำเนินการเรียนการสอน. (กรุงเทพมหานคร : หน่วยพัฒนาคณาจารย์ดำเนินการเรียนการสอน, จุฬาลงกรณ์มหาวิทยาลัย, 2520), หน้า 186.

²เรื่องเดียวกัน, หน้า 186.

³Airasian, "Criterion-Referenced Testing in the Classroom," pp. 76-77.

4. ความเชื่อของนักการศึกษาส่วนมากที่ว่า คนทุกคนหรือคนส่วนมากจะสามารถเรียนรู้หรือสัมฤทธิ์ในวิชาต่าง ๆ ได้ ถ้ามีวิธีการที่ดีกว่า จึงทำให้การทดสอบเปลี่ยนจากการเปรียบเทียบในกลุ่มเป็นการตรวจสอบ และให้รางวัลการเรียนรู้ในรูปของการกระทำ (เช่น ถ้าทุกคนทำได้ถึงเกณฑ์ ทุกคนอาจได้คะแนนระดับ "ก")

กรอนลันด์¹ (Gronlund) ได้เสนอการตัดสินใจในเรื่องการใช้แบบสอบ โดยพิจารณาจากวัตถุประสงค์ของการสอบ และลักษณะของแบบสอบวัดสัมฤทธิ์ผล 4 ชนิด ดังตาราง

ตารางที่ 3 ลักษณะของแบบสอบวัดสัมฤทธิ์ผล 4 ชนิด

ชนิดของแบบสอบ	หน้าที่ของแบบสอบ	การเลือกตัวอย่างข้อสอบ	ลักษณะข้อสอบ
วางตัวบุคคล (Placement)	วัดทักษะเบื้องต้น ตัดสินใจการกระทำเริ่มแรกเกี่ยวกับวัตถุประสงค์ของวิชา	รวมไว้ด้วยพฤติกรรมที่ตรงการวัดเบื้องต้นก่อนเรียน เลือกตัวอย่างที่เป็นตัวแทนของวัตถุประสงค์ของวิชา	โดยทั่วไปจะง่ายและแปลผลแบบอิงเกณฑ์ มีความยากง่ายต่างๆ (มีพิสัยของความยากกว้าง) และอิงกลุ่ม
วัดความก้าวหน้า (Formative)	ให้ข้อมูลย้อนกลับแก่ผู้เรียนและผู้สอนในความสามารถในการเรียน	รวมไว้ด้วยทุกหน่วยของวัตถุประสงค์ที่เป็นไปได้ (หรือจัดสิ่งที่ดีที่สุด)	ข้อสอบยากหรือง่ายแล้วแต่วัตถุประสงค์ของแต่ละหน่วย และอิงเกณฑ์
วัดเพื่อวินิจฉัย (Diagnostic)	เพื่อพิจารณาสาเหตุของความซัดข้องทางการเรียนที่เกิดขึ้นซ้ำ ๆ	รวมไว้ด้วยตัวอย่างของงานซึ่งขึ้นอยู่กับแหล่งของความผิดพลาดในการเรียนที่มีร่วมกันอยู่	โดยทั่วไปจะง่ายและใช้เพื่อชี้จุดบกพร่องของนักเรียน และอิงเกณฑ์
วัดความรู้รวบยอด (Summative)	ให้ระดับคะแนนหรือรับรองความรอบรู้หลายหน่วยการสอน	เลือกตัวอย่างที่เป็นตัวแทนของวัตถุประสงค์	มีพิสัย ความยากกว้าง และอิงกลุ่ม

¹Norman E. Gronlund, Constructing Achievement Test (N.J.: Prentice-Hall, 1977), p.19.

จะเห็นได้ว่า กรอนลันด์ มีความเห็นว่า ถ้าวัตถุประสงค์ในการสอบต่างกัน แบบสอบที่ใช้ก็ต่างกันไปด้วย แต่แบบสอบที่ใช้ในโครงการ I P I¹ (Individually-Prescribed Instruction) ซึ่งเป็นแบบสอบที่ใช้วัดพัฒนาการของนักเรียนในโปรแกรมการสอนรายบุคคล วิชาคณิตศาสตร์ และวิชาอื่น ๆ ในชั้นประถมที่มีการจัดลำดับแบบสอบ ตามวัตถุประสงค์การสอน ได้ใช้แบบสอบอิงเกณฑ์เพียงแบบเดียว (หลายฉบับ) เพื่อวัตถุประสงค์ในการวัด 4 ประการคือ

1. การวางตัวบุคคล (Placement Testing) ใช้แบบสอบเพื่อบอกนักเรียนว่าเขาควรจะเริ่มเรียนชั้นใด ระดับใดของหลักสูตร โดยให้นักเรียนทำแบบสอบในตอนเริ่มปีการศึกษา

2. สอบเพื่อวินิจฉัย (Diagnostic Testing) ใช้แบบสอบก่อนเรียน (Pre-Test) ที่เตรียมไว้ให้นักเรียนทำก่อนที่จะเริ่มเรียนหน่วยการเรียนรู้ใหม่ เนื้อหาเป็นเนื้อหาที่นักเรียนถูกตัดสินให้ผ่านมาแล้ว

3. สอบหลังจากเรียน (Post-testing) นักเรียนจะเลือกโดยตัวตนเองว่าจะเรียนหน่วยใดต่อไป เมื่อผ่านแต่ละหน่วยแล้ว เช่น เมื่อจบหน่วย D อาจเรียนหน่วย E หรือ F ก็ได้ วิธีการก็คือ จะให้นักเรียนทำแบบสอบอีกครั้งหลังจากสอบไปแล้วหลายสัปดาห์ เพื่อดูว่ายังมีความรู้เหลืออยู่หรือไม่ (อาจสุ่มตัวอย่างมาให้ทำเป็นบางข้อ)

4. สอบเพื่อสำรวจความก้าวหน้า (Monitoring Progress) เพื่อตอบคำถามให้แก่เด็กนักเรียนที่ว่า "ฉันรอบรู้ทักษะนี้แล้วจริงหรือ?" จะใช้แบบสอบ CET (Curriculum Embedded Test) ให้ข้อมูลเพิ่มเติมเกี่ยวกับความรอบรู้ของนักเรียน

จะเห็นว่าโครงการ I P I ใช้แบบสอบอิงเกณฑ์วัดความรู้หลังจากเรียน (ซึ่งเป็นความรู้รวมยอดเช่นเดียวกัน) แต่กรอนลันด์ เห็นว่าควรใช้แบบสอบอิงกลุ่มวัดความรู้รวมยอด

¹C.M. Lindvall and A.J. Nitko, Measuring Pupil Achievement and Aptitude, (New York : Harcourt Brace Jovanovich, 1975), p.14.

พอปแฮมและฮูเซ็ก¹ (Popham and Husek) ได้สรุปความแตกต่างระหว่างแบบสอบอิงกลุ่ม และแบบสอบอิงเกณฑ์ไว้ดังนี้

1. การกระจายของคะแนน แมว้าแบบสอบอิงกลุ่มสร้างขึ้นมุ่งจำแนกผู้เรียนตามความสามารถหรือระดับสัมฤทธิ์ผล กล่าวคือ แบบสอบอิงกลุ่มขึ้นอยู่กับความแปรปรวน หรือการกระจายของคะแนนเป็นสำคัญ พอปแฮมและฮูเซ็ก ให้ความเห็นว่า การกระจายของคะแนนไม่เกี่ยวกับแบบสอบอิงเกณฑ์เลย ฉะนั้นวิธีการพัฒนาแบบสอบอิงกลุ่มมาตรฐาน โดยยึดหลักการกระจายของคะแนนเป็นสำคัญ จึงไม่เหมาะสมกับการสร้างหรือพัฒนาแบบสอบอิงเกณฑ์เลย

2. การสร้างแบบสอบ ความแตกต่างพื้นฐานระหว่างนักเขียนข้อสอบอิงกลุ่มกับนักเขียนข้อสอบอิงเกณฑ์นั้นอยู่ที่ว่า นักเขียนข้อสอบอิงกลุ่มมุ่งเขียนข้อสอบ เพื่อจำแนกผู้เรียนตามระดับสัมฤทธิ์ผลหรือมุ่งบอกถึงความแตกต่างระหว่างบุคคลในแง่ของสัมฤทธิ์ผลทางการเรียน แต่เป้าหมายของนักเขียนข้อสอบอิงเกณฑ์มุ่งต้องการให้ข้อสอบสะท้อนภาพให้เห็นถึงพฤติกรรมเกณฑ์ที่ระบุไว้ในวัตถุประสงค์ของการสอนเป็นสำคัญ

3. ความเที่ยง (Reliability) หรือความคงที่ในการวัดของแบบสอบพอปแฮมและฮูเซ็ก เห็นว่า แบบสอบอิงเกณฑ์ควรมีความคงที่ภายใน แต่วิธีคลาสสิกที่ใช้คำนวณค่าสัมประสิทธิ์ความเที่ยงภายในของแบบสอบอิงกลุ่มไม่เหมาะสมกับแบบสอบอิงเกณฑ์ ทั้งนี้เพราะวิธีคลาสสิกในการคำนวณค่าสัมประสิทธิ์ความเที่ยงส่วนใหญ่ขึ้นอยู่กับการกระจายของคะแนนจากแบบสอบเป็นสำคัญ แบบสอบอิงเกณฑ์สามารถมีความคงที่ในการวัดหรือความเที่ยงภายในแบบสอบสูงมาก ทั้ง ๆ ที่ค่าสัมประสิทธิ์ความเที่ยงตามประเพณีนิยม หรือตามวิธีคลาสสิกกลับมีค่าต่ำมากก็เป็นได้

4. ความตรง (Validity) ของแบบสอบเกี่ยวข้องกับเรื่องที่ว่าแบบสอบนั้นวัดอะไรและวัดได้ก็เพียงใด วิธีประเมินความตรงของแบบสอบอิงกลุ่ม ส่วนมากใช้เทคนิคการ

¹W. Popham and T. Husek, "Implications of Criterion Referenced Measurement," Journal of Educational Measurement 1965,6, pp. 1-9.

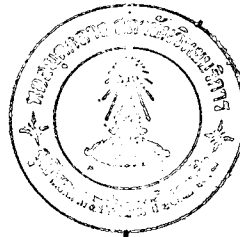
ค่านวณหาสหสัมพันธ์ ซึ่งขึ้นอยู่กับ การกระจายของคะแนนเป็นสำคัญ จึงไม่เหมาะกับการคำนวณหาความตรงของแบบสอบอิงเกณฑ์ ส่วนใหญ่ในการตรวจสอบความตรงของแบบสอบอิงเกณฑ์ อาศัยการพิจารณาว่าเนื้อเรื่องของแบบสอบอิงเกณฑ์เป็นตัวแทน หรือครอบคลุมเกณฑ์ครบถ้วนหรือไม่

5. การวิเคราะห์รายข้อ เป็นเทคนิคที่เกี่ยวข้องกับการบรรยายถึงคุณลักษณะเชิงสถิติของข้อสอบแต่ละข้อ และแล้วก็เลือกข้อสอบเพื่อว่าผลที่ได้จะทำให้ได้แบบสอบที่มีคุณลักษณะเฉพาะบางประการ เช่น แบบสอบที่สามารถจำแนกผู้เรียนเก่งออกจากผู้เรียนอ่อนได้เป็นต้น ทั้งแบบสอบอิงกลุ่มและแบบสอบอิงเกณฑ์ ต้องการข้อสอบที่สามารถจำแนกผู้เรียนรู้ออกจากผู้เรียนที่ไม่รู้อะไรได้ แต่ถาข้อสอบอิงเกณฑ์ไม่มีอำนาจจำแนกดังกล่าว เราก็ยังไม่ตัดข้อสอบข้อนั้นทิ้ง ถาข้อสอบนั้นเป็นตัวแทนของทักษะหรือความรู้สำคัญตามวัตถุประสงค์ที่มุ่งวัด

6. การแปลความหมาย และการรายงานผล ความหมายของคะแนนจากแบบสอบอิงกลุ่มจะได้จากการเปรียบเทียบคะแนนสัมฤทธิ์ผลของผู้เรียน กับคะแนนสัมฤทธิ์ผลของเพื่อนร่วมชั้นหรือกลุ่มที่คล้ายคลึง คะแนนจากแบบสอบอิงกลุ่มแสดงถึงความแตกต่างระหว่างบุคคลในมิติที่มุ่งวัด เช่น สัมฤทธิ์ผลความถนัด หรือเชาวน์ปัญญา เป็นต้น แต่คะแนนจากแบบสอบอิง เกณฑ์จะแปลความหมายโดยเปรียบเทียบกับเกณฑ์มาตรฐานเป็นสำคัญ เพื่อบอกว่าผู้เรียน เรียนรู้หรือไม่เรียนรู้เนื้อหาวิชาใดบ้าง สำหรับการรายงานผลนั้นมักนิยมรายงานคะแนนจากแบบสอบอิงกลุ่ม เป็นคะแนนเทียบกับกลุ่ม หรือคะแนนของทั้งกลุ่ม เช่น ค่ามัธยิมเลขคณิตของระดับสัมฤทธิ์ผล เป็นต้น ซึ่งไม่เหมาะสมกับรายงานคะแนนแบบสอบอิง เกณฑ์ที่มุ่งรายงานผลเป็นรายบุคคล

โบห์ม¹ (Boehm) ได้เปรียบเทียบคุณลักษณะของแบบสอบอิงกลุ่มแบบสอบอิง เกณฑ์ ใน 5 มิติที่สำคัญคือ วัตถุประสงค์ทั่วไปของการสอบ แบบสอบหรือประเภทของข้อสอบ เนื้อหาของแบบสอบ ความหมายของคะแนน และแบบของการจัดอันดับคะแนน ดังแสดงในตาราง

¹A. Boehm, "Criterion-Referenced Assessment for the Teacher" Teachers College Record, 1973, 75, pp. 117-126.



ตารางที่ 4 คุณลักษณะของแบบสอบอิงกลุ่มและแบบสอบอิงเกณฑ์

ชนิดแบบสอบ มิติ	แบบสอบอิงกลุ่ม (NRT)	แบบสอบอิงเกณฑ์ (CRT)
1. วัตถุประสงค์ ทั่วไป	<p>ก. มุ่งเปรียบเทียบความแตกต่างระหว่างบุคคลในแง่สัมฤทธิ์ผลหรือค่านอื่น ๆ</p> <p>ข. ต้องการตัดสินใจการคัดเลือกบุคคลเข้าโปรแกรมตามโควตาที่กำหนด</p> <p>ค. เพื่อชี้แนะว่าโปรแกรมเหมาะสมสำหรับใครบ้าง</p>	<p>ก. บ่งชี้ว่า แต่ละบุคคลสามารถทำอะไรบ้างเมื่อเทียบกับเกณฑ์</p> <p>ข. เพื่อจัดโปรแกรมให้เป็นพิเศษเป็นรายบุคคล</p> <p>ค. เพื่อบ่งชี้ว่าโปรแกรมการสอนประสบความสำเร็จหรือไม่เพียงใด</p>
2. แบบข้อสอบ	<p>ก. ข้อสอบต้องสามารถจำแนกความแตกต่างระหว่างบุคคล</p> <p>ข. ตัดข้อสอบที่ทุกคนตอบได้ถูกต้องหรือที่ทุกคนตอบไม่ได้ทั้ง</p>	<p>ก. ข้อสอบต้องสอดคล้องกับระดับของเกณฑ์</p> <p>ข. ข้อสอบต้องให้ข่าวสารอย่างชัดเจนว่าแต่ละบุคคลสามารถหรือไม่สามารถทำอะไรได้บ้าง</p>
3. เนื้อเรื่อง	<p>ก. เนื้อเรื่องในแบบสอบอาจจะไม่สอดคล้องหรือสอดคล้องกับเป้าหมายเฉพาะของการสอนก็ได้</p>	<p>ก. เนื้อเรื่องในแบบสอบต้องสอดคล้องกับวัตถุประสงค์การสอนเชิงพฤติกรรม</p>

ตารางที่ 4 (ต่อ)

ชนิดแบบสอบ มิตี	แบบสอบอิงกลุ่ม (NRT)	แบบสอบอิงเกณฑ์ (CRT)
4. คะแนน	<p>ข. สุ่มตัวอย่างเนื้อเรื่องจากมิตีของความรู้หรืองานที่ให้ทำอันกว้างขวาง</p> <p>ก. การกระจายของคะแนนเป็นคุณลักษณะที่สำคัญ</p> <p>ข. คะแนนไม่ได้เปิดเผยโดยตรงว่าบุคคลนั้นทำอะไรได้บ้าง แต่บอกถึงคนนั้นมีความสามารถอย่างไร เมื่อเทียบกับคนอื่น ๆ</p>	<p>ข. สามารถตั้งเกณฑ์ระดับต่าง ๆ ในแต่ละระดับของเนื้อเรื่องหน่วยบทเรียน และจะต้องบอกระดับความรู้ ความสามารถขั้นต่ำในแต่ละเนื้อเรื่องไว้ด้วย</p> <p>ก. การกระจายของคะแนนไม่สำคัญ</p> <p>ข. คะแนนสะท้อนภาพให้เห็นว่าอะไรบางอย่างที่บุคคลผู้นั้นสามารถทำได้และอะไรบางอย่างที่เขาไม่สามารถทำได้</p>
5. แบบของการจัดอันดับ	<p>ก. ใช้อายุปกติวิสัย ชั้นเรียนปกติ ตำแหน่งเปอร์เซ็นต์ไทล์</p>	<p>ก. ร้อยละสูงกว่าระดับเกณฑ์</p> <p>ข. ข้อมูลจากการตอบถูกหรือผิดในแต่ละข้อสอบ</p>

องค์ประกอบที่สำคัญที่มีผลต่อผลสัมฤทธิ์ทางการเรียน

เมเฮเรนส์ และเลมันน์¹ (Mehrens and Lehmann) ได้กล่าวว่า นักเรียนจะเรียนได้ดีเพียงใดนั้น ไม่ได้ขึ้นกับความสามารถหรือเชาวน์ปัญญาอย่างเดียว องค์ประกอบอื่นก็อาจมีอิทธิพลต่อผลสัมฤทธิ์ทางการเรียนได้ เช่น วุฒิภาวะ แรงจูงใจ นิสัยในการเรียน ทักษะในการเรียน และทัศนคติของนักเรียนที่มีต่อคุณค่าทางการศึกษา ต่อครู โรงเรียน และวิชาที่เรียน ทราเวอร์ส² (Travers) ให้ความเห็นว่า การทำนายผลสัมฤทธิ์ทางการเรียนด้วยตัวแปรทางด้านสติปัญญา ถึงแม้ว่าจะออกข้อสอบได้ดีเยี่ยมเพียงใด ค่าสหสัมพันธ์ระหว่างผลสัมฤทธิ์ทางการเรียนกับตัวแปรทางด้านสติปัญญา จะมีค่าสูงสุดไม่เกิน .7 หมายความว่าในการทำนายนี้จะทำนายได้ถูกต้อง 50 % ที่เหลือจะต้องใช้ตัวแปรทางด้านอื่น ที่มีใช้สติปัญญา มาวัด เช่น การปรับตัว สภาพทางบ้าน ความสนใจ ความมุ่งมั่น ทัศนคติที่มีต่อสถาบัน วิชาที่เรียน และอาจารย์ที่สอน บราวน์ และโฮลท์ซแมน³ (Brown and Holtzman) ได้ตั้งสมมติฐานขึ้นว่า นิสัยในการเรียนและทัศนคติกับแรงจูงใจในการเรียนอาจจะมีอิทธิพลต่อผลสัมฤทธิ์ทางการเรียน มากกว่าบุคลิกภาพหรือตัวแปรทางด้านจิตวิทยาอื่น ๆ และบราวน์และโฮลท์ซแมน⁴ ก็ได้สร้างแบบสำรวจนิสัย และทัศนคติในการเรียนขึ้น เพื่อวัดตัวแปรดังกล่าว

¹William A. Mehrens and Irvin J. Lehmann. Measurement and Evaluation in Education and Psychology, 2d ed. (Sydney : Holt, Rinehart and Winston, 1975), p. 584.

²Robert M.W. Travers. Educational Measurement. (New York: The Macmillan Company, 1958), p. 396.

³William F. Brown and Wayne H. Holtzman, "Use of the Survey of Study Habits and Attitudes for Counseling Students," The Personnel and Guidance Journal 35 (December 1956), pp. 214-217.

⁴William F. Brown and Wayne H. Holtzman, "Evaluating The Study Habits and Attitudes of High School Students," Journal of Educational Psychology 59 (December 1968), pp.404-409.

การทดลองใช้แบบสำรวจ ปรากฏว่า ได้ผลเป็นที่น่าพอใจ คือสามารถทำนายผลสัมฤทธิ์ทางการเรียนของนักเรียนในชั้นมัธยมศึกษาได้ สเตาเลทซ์ คอลเลจ และโพเมเร็นทซ์¹ (Schultz & College & Pomerantz) ได้ศึกษาเกี่ยวกับแรงจูงใจใฝ่สัมฤทธิ์กับสัมฤทธิ์ผลทางการเรียน โดยใช้กลุ่มตัวอย่างเป็นนักเรียนระดับชั้น 9 ที่สุ่มมาจากโรงเรียนในเมืองเล็ก ๆ แห่งหนึ่ง จำนวน 2 โรงเรียน เครื่องมือที่ใช้คือ

1. Mehrabian's Resultant Achievement Motivation Scale (R.A.M) และ Herman's Prestatic Motivate Test (PMT) เพื่อวัดแรงจูงใจใฝ่สัมฤทธิ์
2. Comprehensive Test of Basic Skill (C.T.B.S.) เพื่อประเมินความสำเร็จของผู้เรียน

ผลการศึกษาปรากฏว่า แรงจูงใจใฝ่สัมฤทธิ์และสัมฤทธิ์ผลทางการเรียน มีความสัมพันธ์กัน

รัสเซลส์² (Russell) ได้ศึกษาความสัมพันธ์ระหว่างสัมฤทธิ์ผลทางการเรียน และแรงจูงใจใฝ่สัมฤทธิ์ กับกลุ่มตัวอย่าง ที่กำหนดเรียนอยู่ในชั้นระดับ 9 โดยใช้แบบทดสอบ California Achievement Test วัดผลสัมฤทธิ์ในวิชาเลขคณิต ภาษา และการอ่าน ปรากฏว่าคะแนนผลสัมฤทธิ์กับแรงจูงใจใฝ่สัมฤทธิ์มีความสัมพันธ์กันสูง ราไฟทิพย์ ชีรนิติ³

¹C.B. Schultz, Trinity College and Michael Pomerantz, "Achievement Motivation, Locus of Control and Academic Achievement Behavior," Journal of Personality 44 (11) (March 1976), pp.39-50.

²Ivan L. Russell, "Motivation for School Achievement : Measurement and Validation," The Journal of Educational Research 62 (February 1969), pp.263-266.

³ราไฟทิพย์ ชีรนิติ, "ความสัมพันธ์ระหว่างแรงจูงใจใฝ่สัมฤทธิ์ กับสัมฤทธิ์ผลทางการเรียน" (วิทยานิพนธ์ปริญญาโทบริหารศึกษาศาสตร์ แผนกวิชาจิตวิทยา บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย, 2514), หน้า 25-27.

ได้ศึกษาความสัมพันธ์ระหว่างแรงจูงใจใฝ่สัมฤทธิ์ และผลสัมฤทธิ์ทางการเรียนเช่นกัน กลุ่มตัวอย่างเป็นนักศึกษาชั้นปีที่ 1 ที่กำลังศึกษาในวิทยาลัยครู จำนวน 240 คน ชาย 120 คน หญิง 120 คน เครื่องมือที่ใช้คือแบบสอบถามวัดแรงจูงใจใฝ่สัมฤทธิ์ของ ฮิวเบอร์ค เจ เอ็ม เออร์แมนส์ และสัมฤทธิ์ผลทางการเรียนนั้น พิจารณาจากคะแนนเฉลี่ยประจำภาคต้น ปีการศึกษา 2514 ผลการวิจัยพบว่า แรงจูงใจใฝ่สัมฤทธิ์มีความสัมพันธ์กับสัมฤทธิ์ผลทางการเรียนอย่างมีนัยสำคัญที่ระดับ .01

จึงเห็นได้ว่าระบบประเมินผลการเรียนที่ควรใช้ในการวัดสัมฤทธิ์ผลทางการเรียน ยังขัดแย้งกันอยู่ และตามระเบียบการประเมินผลการเรียนตามหลักสูตรประโยคมัธยมศึกษาตอนปลาย พุทธศักราช 2518 ก็ยังไม่ได้ตกลงไปว่าให้ใช้ระบบประเมินผลการเรียนแบบใดให้อยู่ในดุลยพินิจของครู จึงเป็นเหตุจูงใจให้ผู้วิจัยเลือกศึกษาว่าระบบประเมินผลการเรียนที่ต่างกันจะทำให้ผลสัมฤทธิ์ทางการเรียนต่างกันหรือไม่ เพียงใด โดยใช้แรงจูงใจใฝ่สัมฤทธิ์เป็นตัวแปรเกินที่นำมาศึกษาด้วย

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย