



## บทที่ 2

### วรรณคดีและงานวิจัยที่เกี่ยวข้อง

ในการวิจัยครั้งนี้ ผู้วิจัยได้ศึกษาแนวคิด ทฤษฎี และงานวิจัยที่เกี่ยวข้องจากหนังสือ เอกสาร และงานวิจัยต่าง ๆ โดยเสนอผลการศึกษาเป็น 4 ตอน ดังต่อไปนี้

- ตอนที่ 1 ความหมาย ปรัชญา วัตถุประสงค์ของการศึกษาระดับบัณฑิตศึกษา
- ตอนที่ 2 การจัดการศึกษาระดับบัณฑิตศึกษาของมหาวิทยาลัยมหิดล
- ตอนที่ 3 การวัดและประเมินผลการศึกษา
- ตอนที่ 4 งานวิจัยที่เกี่ยวข้อง

#### ความหมาย ปรัชญา วัตถุประสงค์ของการศึกษาระดับบัณฑิตศึกษา

คำว่า บัณฑิต ตามพจนานุกรม หมายถึง ผู้มีปัญญา นักปราชญ์ คำว่า ศึกษา หมายถึง การเล่าเรียน การฝึกฝนอบรม โดยความหมายรวมคำว่า บัณฑิตศึกษา (Graduate) หมายถึง การศึกษาเล่าเรียน หรือการฝึกอบรมของผู้มีปัญญาหรือนักปราชญ์ หรือการศึกษาที่เกิดขึ้นหลังจากคนเรียนเป็นบัณฑิตแล้ว (วิจิตร ศรีสอาน, 2523) จากความหมายดังกล่าว การศึกษาระดับบัณฑิตศึกษา จึงเป็นการศึกษาอีกระดับหนึ่ง ซึ่งเป็นการศึกษาในระดับสูงที่จัดให้สำหรับผู้มีปัญญา ปราชญ์ เข้ามาศึกษา หรือกล่าวอีกนัยหนึ่งก็คือ การศึกษาที่จัดขึ้นสำหรับผู้ที่เป็นบัณฑิต ซึ่งเป็นการศึกษาที่สูงกว่าระดับปริญญา สำหรับในประเทศไทยนั้น มีการจัดการศึกษาระดับบัณฑิตศึกษาอยู่ 3 ระดับคือ (วิจิตร ศรีสอาน, 2523)

1. ระดับประกาศนียบัตร ซึ่งมีชื่อเรียกแตกต่างกัน เช่น ประกาศนียบัตรวิชาเฉพาะ ประกาศนียบัตรชั้นสูง หรือประกาศนียบัตรบัณฑิต ระยะเวลาการศึกษาตลอดหลักสูตร 1 ปี
2. ระดับปริญญาโท ระยะเวลาการศึกษาตลอดหลักสูตร 2 ปี
3. ระดับปริญญาเอก มี 2 ประเภท คือ ปริญญาเอกด้านปรัชญา (Doctor of Philosophy (Ph.D.)) และปริญญาเอกเฉพาะสาขาวิชาชีพ เช่น Doctor of Education (Ed.D.) Doctor of Science (D.Sc.) Doctor of Engineering

(D.Eng.) ระยะเวลาการศึกษาตลอดหลักสูตร 3 ปี

ในต่างประเทศ มีการจัดการศึกษาระดับบัณฑิตศึกษาอีกลักษณะหนึ่ง ซึ่งเป็นการจัดการศึกษาระดับหลังปริญญาโทแต่ไม่ถึงปริญญาเอก เนื่องจากไม่ต้องทำวิทยานิพนธ์ และเรียกชื่อแตกต่างกัน เช่น Specialist Certificate, Graduate Diploma, Professional Diploma เป็นต้น เป็นการศึกษาถึงทางระหว่างปริญญาโทและปริญญาเอก หลักสูตรมี 2 ลักษณะ คือ

1. เรียนจบแล้วไม่สามารถเรียนต่อปริญญาเอก
2. เรียนจบแล้วสามารถเรียนต่อปริญญาเอกได้ เช่น Teachers College

ใน Columbia ผู้ที่จบ Preliminary Examination และอยู่ในระหว่างการทำ Project หรือวิทยานิพนธ์นั้นจะได้รับ Professional Diploma เพื่อนำไปประกอบอาชีพก่อนได้ ซึ่งปัจจุบันสถาบันอุดมศึกษาบางแห่งในประเทศไทยได้มีการจัดการศึกษาระดับบัณฑิตศึกษาในลักษณะนี้ โดยเรียกว่า ประกาศนียบัตรวิชาชีพชั้นสูง

นอกจากนี้สถาบันอุดมศึกษาต่างประเทศบางแห่งได้มีการจัดการศึกษาระดับหลังปริญญาเอกขึ้นอีกด้วย ซึ่งเรียกว่า โปรแกรมระดับหลังปริญญาเอก (Post Doctoral Program)

ชัยพร วิชชาวุธ (2523) กล่าวถึงบัณฑิตศึกษา 3 ลักษณะคือ

1. การบัณฑิตศึกษา เป็นการศึกษาที่เพิ่มเติมต่อจากระดับปริญญาตรี เพื่อให้มีความรู้ลึกซึ้งยิ่งขึ้น และมีวิธีการศึกษาไม่แตกต่างจากระดับปริญญาตรี วิทยานิพนธ์ไม่จำเป็น
2. การบัณฑิตศึกษา เป็นการสร้างและเพิ่มพูนทักษะการประกอบอาชีพเฉพาะทาง เช่น ทักษะการเป็นนักบริหารงานบุคคล เป็นนักจิตวิทยาคลินิกเด็ก ฯลฯ การฝึกปฏิบัติมีความจำเป็นมากที่สุด วิทยานิพนธ์ไม่มีความจำเป็น ถ้าหากจะให้ทำวิทยานิพนธ์ก็เพื่อให้เรียนรู้กระบวนการเพื่อจะได้สามารถ "อ่าน" และติดตามผลงานวิจัยในวิชาชีพของตน
3. การบัณฑิตศึกษา เป็นการศึกษาเพื่อผลิตนักวิจัยและนักวิชาการในแต่ละสาขาวิชา วิทยานิพนธ์เป็นหัวใจของการบัณฑิตศึกษา เพราะเป็นการทดสอบความสามารถในการวิจัยและการวิจารณ์

ไพฑูริย์ สินลารัตน์ (2526) ได้กล่าวถึง การศึกษาระดับบัณฑิตศึกษาว่าโดยทั่ว ๆ ไป จะแบ่งเป็น 2 สายใหญ่ ๆ คือ

1. สายวิชาการ (Academic) มีเป้าหมายในการผลิตบุคลากรหลัก คือ ผลิต

ผู้สอนในมหาวิทยาลัย (College Teacher) และผลิตนักวิจัยหรือนักวิชาการ (Researcher or Scholar) ดังนั้น โปรแกรมในสาขานี้จึงเน้นหนักด้านการวิจัยเป็นสำคัญ หรือการวิจัย ผสมกับการเรียนรายวิชา ทั้งนี้เพื่อเป็นการบุกเบิกแสวงหาความรู้ใหม่ และสร้างความก้าวหน้าทางวิชาการด้านต่าง ๆ ปรึญญาที่ได้รับโดยเฉพาะปริญญาเอก คือ Doctor of Philosophy (Ph.D.)

2. สาขาวิชาชีพ (Professional) มีเป้าหมายในการผลิตผู้เชี่ยวชาญหรือผู้ชำนาญการเฉพาะทาง (Specialist) โปรแกรมในสาขานี้จะเน้นหนักด้านการฝึกฝนหาความชำนาญเฉพาะสาขา ไม่เน้นการวิจัยเหมือนสายวิชาการ ปรึญญาที่ได้รับมักจะเป็น Professional Degree เช่น ปรึญญา Ed.D. ทางการศึกษา

การจัดหลักสูตรบัณฑิตศึกษาในประเทศไทย โดยทั่วไปจัดเป็น 3 ระดับคือ

1. ระดับประกาศนียบัตร ซึ่งมีชื่อเรียกแตกต่างกันไปในแต่ละสถาบัน เช่น ประกาศนียบัตรวิชาเฉพาะ ประกาศนียบัตรชั้นสูง หรือประกาศนียบัตรบัณฑิต เป็นต้น โดยมีวัตถุประสงค์ให้เป็นหลักสูตรที่จบเบ็ดเสร็จในตัวหลักสูตรเอง ใช้เวลาประมาณ 1 ปีการศึกษา ในปัจจุบันนี้มีหลักสูตรระดับประกาศนียบัตรน้อยมาก จะยังคงมีอยู่ในส่วนของสาขาวิชาชีพเฉพาะเท่านั้น เช่น สาขาวิชาแพทยศาสตร์ ซึ่งมีเปิดขยายเพิ่มขึ้น และสาขาวิชาสังคมศาสตร์ ซึ่งยังคงมีอยู่บ้างในกลุ่มบริหารธุรกิจ เป็นต้น

2. ระดับปริญญาโท เป็นการศึกษาในระดับที่ได้รับความนิยมเป็นอย่างมาก วัฒนาการของการจัดการศึกษาระดับปริญญาโทนี้ มีความเป็นมารวดเร็วมาก ในเชิงปริมาณ ทั้งจำนวนผู้เรียน จำนวนสาขาวิชาที่เปิดและจำนวนผู้สำเร็จการศึกษา สาขาวิชาที่มีผู้เรียนมากที่สุด (หรือสามารถรับเข้าศึกษาได้มากที่สุด) ตั้งแต่ระยะแรกจนถึงปัจจุบัน คือ สาขาสังคมศาสตร์ ขณะนี้สถาบันอุดมศึกษาทั้งของรัฐและเอกชน ได้เปิดสอนระดับปริญญาโทเกือบทุกสถาบัน แม้แต่มหาวิทยาลัยเปิดทั้งสองแห่ง คือ มหาวิทยาลัยรามคำแหง และมหาวิทยาลัยสุโขทัยธรรมมาธิราชก็ได้เริ่มดำเนินการเปิดสอนระดับปริญญาโทเช่นกัน

3. ระดับปริญญาเอก การศึกษาระดับนี้ได้มีการเปิดสอนมานานพอสมควรและขยายตัวค่อนข้างเร็ว ในระยะแผนพัฒนาฯ ระยะที่ 4 และระยะที่ 5 ผู้ศึกษาในระดับนี้ จำนวนกว่าครึ่งศึกษาในสาขาวิชาแพทยศาสตร์ และสาขาวิชาที่เกี่ยวข้องกับสุขภาพ การศึกษาระดับปริญญาเอกมีสถาบันที่เปิดสอนอยู่ 5 สถาบัน คือ จุฬาลงกรณ์มหาวิทยาลัย มหาวิทยาลัยมหิดล มหาวิทยาลัยศรีนครินทรวิโรฒ สถาบันพัฒนาบริหารศาสตร์ และสถาบัน

เทคโนโลยีประจอบเกล้าเจ้าคุณทหาร ลาดกระบัง

เมื่อพิจารณาลักษณะการจัดการศึกษาในระดับบัณฑิตศึกษาจะพบว่าการจัดการศึกษาในระดับบัณฑิตศึกษามีลักษณะแตกต่างกัน แต่อย่างไรก็ตามการจัดการศึกษาในระดับนี้ก็ยังมียุทธศาสตร์ทั่วไปบางอย่างร่วมกัน ซึ่ง วิจิตร ศรีสอน (2523) ได้สรุปไว้ว่ามี 4 ลักษณะคือ

1. การศึกษาที่มุ่งเน้นความเชี่ยวชาญเฉพาะทาง (Specialization) และมีจุดมุ่งหมายหลักที่จะเสริมสร้างความรู้เชิงลึกซึ่งในสาขาใดสาขาหนึ่ง เพื่อเป็นผู้ที่มีความรู้เชี่ยวชาญในเรื่องนั้น ๆ

2. การศึกษาที่มุ่งเน้นการวิจัยค้นคว้า (Research Oriented) โดยการทำการวิจัยเรื่องใดเรื่องหนึ่ง และศึกษารายวิชาต่าง ๆ เพียงเพื่อเสริมความรู้ในการทำวิจัยนั้น

3. การศึกษาที่มุ่งเน้นเรื่องวิชาชีพโดยเฉพาะ (Professionalization) เพื่อยกฐานะทางวิชาชีพ และผลิตผู้เชี่ยวชาญหรือชำนาญการในวิชาชีพ

4. การศึกษาที่มุ่งสร้างนักวิชาการที่มีขีดความสามารถทางวิชาการสูง (Scholarship) เพื่อเป็นผู้นำทางวิชาการในสาขาของตนได้

การจัดการศึกษาในระดับนี้ จึงเป็นการจัดการศึกษาในระดับสูง สำหรับผู้ที่มีความรู้เฉพาะสาขาเข้ามาศึกษาให้มีความรู้ที่ลึกซึ้งเชี่ยวชาญยิ่งขึ้น และสามารถพัฒนาความรู้ที่มีอยู่เพื่อสร้างสรรค์ความรู้ใหม่ ๆ ได้ การจัดการศึกษาในระดับนี้จึงมีปรัชญาและวัตถุประสงค์เฉพาะ ซึ่ง กานดา ณ ถลาง (2522) ได้กล่าวถึง หลักปรัชญาทั่วไปของการศึกษาในระดับบัณฑิตศึกษาว่าไม่ควรเป็นการถ่ายทอดวิชาการแต่เพียงอย่างเดียว แต่ต้องปลุกคนขึ้นมา (Arouse the whole man) และให้พลังกำลังความคิดแก่เขา ให้เขาสนใจที่จะแสวงหาความรู้ในสาขาวิชาเฉพาะของเขา ดังนั้นความเป็นตัวของตัวเอง ความมีเอกลักษณ์ ความลึกซึ้งในการเสาะแสวงหา และความตื่นตัวทั้งหลายจึงเป็นคุณลักษณะประจำของบัณฑิตศึกษา ซึ่งสอดคล้องกับ ประชุมสุข อ้าวอ้าว (2523) ที่กล่าวว่า การศึกษาในระดับบัณฑิตศึกษาเป็นการศึกษาที่ตั้งเดี่ยว ลึกซึ้ง เป็นความชำนาญพิเศษ เป็นขบวนการถ่ายทอดความคิดสืบเสาะแสวงหาวิชาการที่บริสุทธิ์ และปลุกคนเข้าสู่สภาวะที่เป็นเลิศในวิถีทางแห่งปัญญา

นิพนธ์ ศศิธร (2524) กล่าวถึงปรัชญาของบัณฑิตศึกษาว่า ควรมีลักษณะผสมผสานระหว่างปัญญาและความรู้ที่สูงส่ง หรือคำดั่งลึกซึ้ง (Vertical) กับการมองเห็นเห็นภาพของสาขาอื่น ๆ ด้วย (Versatile) ดังที่ สุนทร อมรวิวัฒน์ (2529) ให้ความเห็นเห็นว่า ผู้ที่

ได้รับการศึกษาระดับนี้ควรจะเป็นผู้เปลี่ยนแปลงสังคมได้ เป็นนักวิชาการเป็นนักแสวงหาข้อมูล เป็นนักคิด นักทำ รวมทั้งเป็นผู้รู้แจ้ง เห็นจริงในข้อมูลต่าง ๆ

นอกจากนี้ทบวงมหาวิทยาลัยได้กำหนดปรัชญาและวัตถุประสงค์ของหลักสูตรในระดับบัณฑิตศึกษาว่า เน้นการพัฒนานักวิชาการและนักวิชาชีพที่มีความรู้ความสามารถระดับสูงในสาขาวิชาต่าง ๆ เพื่อให้สามารถบุกเบิกแสวงหาความรู้ใหม่ได้อย่างมีประสิทธิภาพ รวมทั้งมีความสามารถในการสร้างสรรค์ จรรโลงความก้าวหน้าทางวิชาการได้อย่างต่อเนื่อง (ทบวงมหาวิทยาลัย, 2533) และได้กำหนดวัตถุประสงค์ของการเปิดสอนระดับบัณฑิตศึกษาดังต่อไปนี้ (ทบวงมหาวิทยาลัย, 2533)

1. เพื่อสนองความต้องการกำลังคนระดับสูงในสาขาวิชานั้น ๆ
2. เพื่อส่งเสริมความก้าวหน้าในสาขาวิชาต่าง ๆ โดยอาศัยผลการวิจัยค้นคว้า เป็นปัจจัยสำคัญและส่งเสริมประสิทธิภาพการสอน
3. เพื่อส่งเสริมให้ผู้ศึกษาเป็นผู้มีความรู้ความสามารถทางวิชาการในระดับสูง และเป็นผู้มีจริยธรรม

จากความหมาย ปรัชญา และวัตถุประสงค์ ของการจัดการศึกษาระดับบัณฑิตศึกษาดังกล่าว สรุปได้ว่า การศึกษาระดับบัณฑิตศึกษา เป็นการจัดการศึกษาในระดับสูงสำหรับผู้ที่มีพื้นฐานความรู้ในระดับปริญญาตรี เป็นต้นไป เข้ามาศึกษาเพื่อพัฒนาความรู้ความคิดของตนในสาขาวิชาต่าง ๆ ให้มีความลุ่มลึกและลึกซึ้ง โดยใช้ระเบียบวิธีการศึกษาในลักษณะการเสาะแสวงหาการค้นคว้า การวิจัย อันเป็นกระบวนการในการแสวงหาความรู้ใหม่ และความก้าวหน้าทางวิชาการ ทั้งนี้การศึกษาในระดับบัณฑิตศึกษา มีวัตถุประสงค์เพื่อสร้างผู้นำทางความคิด มีความรู้ ทักษะ และความสามารถในระดับสูง เพื่อก้าวไปสู่ภาวะแห่งความเป็นเลิศทางวิชาการและวิชาชีพ

#### การจัดการศึกษาระดับบัณฑิตศึกษาของมหาวิทยาลัยมหิดล

บัณฑิตวิทยาลัย มหาวิทยาลัยมหิดล เป็นหน่วยงานหนึ่งในสังกัดมหาวิทยาลัยมหิดล ที่รับผิดชอบเปิดสอนหลักสูตรระดับบัณฑิตศึกษา โดยเริ่มดำเนินการตั้งแต่ปี พ.ศ. 2505 ซึ่งในขณะนั้น มหาวิทยาลัยมหิดลยังมีชื่อว่า "มหาวิทยาลัยแพทยศาสตร์" และวิชาที่สอนอยู่ในมหาวิทยาลัยแพทยศาสตร์ ก็เป็นวิชาด้านการแพทย์เป็นส่วนใหญ่ การจัดตั้งบัณฑิตวิทยาลัยขึ้นในคราวนั้นก็เพื่อต้องการแก้ไขปัญหาการขาดแคลนแพทย์ และการกระจายแพทย์ในประเทศ

ไทยด้วยการทำหน้าที่เป็นแหล่งผลิต อบรม ให้การศึกษาแก่อาจารย์ในสาขาวิทยาศาสตร์การแพทย์เพื่อที่จะสร้างและส่งเสริมสมรรถภาพของมหาวิทยาลัยในด้านการผลิตอาจารย์ประจำให้มีมาตรฐานทางวิชาการอย่างสูงภายในประเทศ และเพื่อผลิตผู้ปฏิบัติงานเกี่ยวกับวิทยาศาสตร์การแพทย์สาขาต่าง ๆ ให้มีคุณภาพที่เหมาะสม ดังนั้นบัณฑิตวิทยาลัยที่ตั้งขึ้นในคราวนั้น จึงมีหน้าที่และจุดมุ่งหมาย ดังนี้

1. ให้การศึกษอบรมและผลิตอาจารย์ที่มีคุณภาพดีและปริมาณพอแก่ความต้องการให้แก่โรงเรียนแพทย์ ทันตแพทย์ เภสัชกร ที่ตั้งขึ้นใหม่ ตลอดจนปรับปรุงส่งเสริมเพิ่มปริมาณและคุณภาพของอาจารย์ในมหาวิทยาลัยแพทยศาสตร์ โดยทั่วไปที่มีอยู่แล้วในขณะนี้ให้เหมาะสมยิ่งขึ้น

2. ช่วยเหลืออบรม ปรับปรุงคุณวุฒิ และสมรรถภาพของแพทย์ของทางราชการ และแพทย์ที่ปฏิบัติงานเป็นอิสระ รวมทั้งเภสัชกร ทันตแพทย์ ที่กำลังบริการประชาชนอยู่ในขณะนั้น ให้มีความรู้ทันสมัยและเหมาะสมเพื่อจะได้เป็นประโยชน์แก่ประชาชนให้มากยิ่งขึ้นเท่าที่จะทำได้

3. ช่วยวิจัยปัญหาการแพทย์และการสาธารณสุขต่าง ๆ โดยเฉพาะที่มีความสำคัญแก่ประเทศไทยให้มากที่สุดเท่าที่จะทำได้ (มหาวิทยาลัยมหิดล บัณฑิตวิทยาลัย, 2534)

ปัจจุบัน บัณฑิตวิทยาลัยเปิดสอนหลักสูตรสาขาวิชาต่าง ๆ ทางด้านวิทยาศาสตร์ ศึกษาศาสตร์ อักษรศาสตร์ สังคมศาสตร์ ในแต่ละระดับการศึกษาดังต่อไปนี้

ระดับประกาศนียบัตรบัณฑิต จำนวน 48 สาขาวิชา

ระดับปริญญาโท จำนวน 64 สาขาวิชา

ระดับปริญญาเอก จำนวน 19 สาขาวิชา

และในปีการศึกษา 2534 บัณฑิตวิทยาลัยได้รับนักศึกษาเข้าศึกษาในระดับบัณฑิตศึกษา จำนวน 948 คน เป็นนักศึกษาระดับปริญญาเอก 15 คน ระดับปริญญาโท 553 คน ระดับประกาศนียบัตรบัณฑิต 380 คน ปัจจุบันมีนักศึกษาระดับบัณฑิตศึกษาทั้งสิ้น 2,431 คน โดยจำแนกเป็นระดับปริญญาเอก 146 คน ระดับปริญญาโท 1,905 คน และระดับประกาศนียบัตรบัณฑิต 380 คน (มหาวิทยาลัยมหิดล บัณฑิตวิทยาลัย, 2534)

## การวัดและประเมินผลการศึกษา

### ความหมายของการวัดและประเมินผลการศึกษา

Good (1959) กล่าวถึง คำว่า การวัดผล การประเมินผล ว่าบางครั้งใช้แทนความหมายในแง่ของกระบวนการวัดและประเมินผล แต่โดยแท้จริงแล้วสองคำนี้มีความหมายแตกต่างกัน การวัดผล (Measurement) เป็นการเปรียบเทียบปริมาณเพื่อแสดงค่าเป็นตัวเลข ซึ่งแตกต่างกับการประเมินผล (Evaluation) หมายถึง กระบวนการในการตัดสินใจความสำคัญของปรากฏการณ์อย่างหนึ่งกับปรากฏการณ์อีกอย่างหนึ่ง โดยใช้มาตรฐานที่กำหนดไว้ ความหมายดังกล่าวนี้คล้ายคลึงกับที่ Guilford (1954) ได้พยายามไว้ว่า การวัดผลเป็นการจัดค่าตัวเลขให้แก่วัตถุหรือเหตุการณ์ โดยมีกฎเกณฑ์ ส่วนการประเมินผลเป็นการตัดสินค่านิยมของการกระทำ ดังนั้นการวัดผลจึงเป็นการรวบรวมข้อมูลต่าง ๆ เพื่อนำมาพิจารณาในการตัดสินใจ (Marshall, 1972 อ้างถึงใน วิวัฒน์ กระจู, 2534) และเป็นการรวบรวมและเรียบเรียงข้อมูลข้อความ หรือข่าวสารอย่างเป็นระบบ (Smith and Adams, 1966) หรือเป็นกระบวนการในการกำหนดจำนวนให้กับแต่ละสมาชิกของสิ่งของหรือบุคคล (Ebel, 1965) เพื่อบอกคุณลักษณะของคนที่ได้พัฒนาขึ้นอย่างกว้างขวางจากการให้การศึกษา โดยจะบอกคุณลักษณะได้ 2 วิธี คือ (Thorndike, 1972)

1. การบอกคุณลักษณะด้วยการทดสอบ
2. การบอกคุณลักษณะด้วยการสังเกตพฤติกรรม

จากความหมายและนิยามดังกล่าว สรุปได้ว่า การวัดและประเมินผลเป็นเรื่องของกระบวนการที่ต่อเนื่องกันและจะต้องกระทำอยู่ตลอดเวลา การวัดผลเป็นกระบวนการในการกำหนดค่าของสิ่งที่จะวัดเป็นตัวเลข สัญลักษณ์หรือเป็นการแปลงคุณลักษณะของสิ่งที่วัดออกมาในรูปปริมาณ โดยอาศัยเครื่องมือในการวัด ส่วนการประเมินผลเป็นเรื่องของการตัดสินใจ การประเมินคุณค่า หรือจัดอันดับสิ่งที่จะวัดโดยใช้เกณฑ์ในการพิจารณาอย่างมีระบบ การประเมินผลจึงมีความสัมพันธ์เกี่ยวข้องเชื่อมโยงมาจากการวัดผล และเป็นกระบวนการที่เกิดขึ้นภายหลังจากการวัดผล จากความหมายดังกล่าวนี้ การวัดและการประเมินผลจึงเกี่ยวข้องกับสิ่งสำคัญ 3 สิ่ง คือ สิ่งที่จะวัด เครื่องมือที่จะใช้ในการวัด และเกณฑ์ที่ใช้ในการพิจารณาตัดสิน

### จุดประสงค์ของการวัดและประเมินผลการศึกษา

ก่อนที่ผู้สอนจะทำการวัดและประเมินผลผู้เรียน ผู้สอนจะต้องทราบจุดประสงค์หรือจุดมุ่งหมายของการวัดก่อนว่ามีเป้าหมายอย่างไร จึงจะสามารถดำเนินการวัดและประเมินผลได้สอดคล้องกับวัตถุประสงค์นั้น ชาวาล แพร์ตกล (2508) กล่าวถึงจุดประสงค์ของการวัดผลว่ามี 5 ประการ คือ เพื่อจัดประเภท เพื่อวินิจฉัย เพื่อเปรียบเทียบ เพื่อพยากรณ์ และเพื่อประเมินค่า สำหรับการวัดผลเพื่อประเมินค่านี้ จะรวมถึงการประเมินหลักสูตรแบบเรียนและการสอนด้วย มิใช่ประเมินตัวผู้สอนเท่านั้น บัญเชิด ภิญโญนันตพงษ์ (2519) ได้กล่าวถึงจุดประสงค์ของการวัดและประเมินผลเพิ่มเติมนอกเหนือจากที่ ชาวาล แพร์ตกล กล่าวไว้ คือ เพื่อคัดเลือก เพื่อประเมินพัฒนาการ เพื่อจูงใจการเรียนและเพื่อรักษามาตรฐานสำหรับผู้สอน วัณญา วิศาลาภรณ์ (2522) กล่าวว่า การวัดผลมีจุดประสงค์เพื่อช่วยให้กระบวนการเรียนรู้เป็นไปได้อย่างสะดวกและง่ายดาย เนื่องจากการวัดผลจะช่วยวิเคราะห์ถึงสาเหตุข้อบกพร่องของการเรียนรู้ เพื่อปรับปรุงและพัฒนาการสอน เนื่องจากการวัดผลจะช่วยตรวจสอบว่าผู้เรียนเกิดประสบการณ์หรือพฤติกรรมที่พึงปรารถนาหรือไม่ เพื่อช่วยในการแนะแนว โดยการทดสอบเพื่อค้นหาความสามารถเฉพาะตัว เพื่อส่งเสริมหรือให้การช่วยเหลือและเพื่อการจัดชั้นเรียน และการศึกษาต่อ สำหรับทางด้านจิตวิทยา Thorndike and Hagen (1969) กล่าวว่า การวัดผลมีจุดประสงค์เพื่อวัดคุณลักษณะทางจิตวิทยา 2 ประการคือ

1. วัดว่าบุคคลสามารถทำอะไรได้บ้าง (What a person can do) ได้แก่ การวัดความถนัด (Aptitude) และการวัดผลสัมฤทธิ์ (Achievement)
2. วัดว่าบุคคลจะทำอะไรต่อไป (What the person will do) เป็นการวัดเพื่อพยากรณ์ หรือวัดบุคลิกภาพ ความรู้สึกของบุคคล เช่น ทศนคติ ความสนใจ

นอกจากนี้ อุตุมพร จามรมาน (2530) กล่าวถึงจุดมุ่งหมายทางการศึกษาว่า มี 3 ด้านคือ นุทธิ จิตต และพลังทักษะ และการศึกษาจะทดสอบว่า ผู้เรียนได้บรรลุจุดมุ่งหมายดังกล่าวหรือไม่ต้องมีการทดสอบ ตรวจสอบ และการวัด ซึ่งอาจจะแสดงออกมาในรูปข้อสอบหรือข้อคำถาม ดังนั้นวิธีการวัดและประเมินผลการเรียนจึงเกี่ยวพันกับจุดมุ่งหมายทางการศึกษา กิจกรรมการเรียนการสอน ผู้เรียน ผู้สอน การทดสอบ และการวัด โดยมีขั้นตอนดังนี้

1. การวางแผนร่วมกันระหว่างผู้เรียนกับผู้สอน เพื่อให้ผู้เรียนทราบวิธีการเรียนการสอน และผู้เรียนสามารถขอปรับเปลี่ยนแปลงรายการ ให้สอดคล้องกับความต้องการของตน โดยการตกลงในกลุ่มผู้เรียนและผู้สอน ทั้งนี้ผู้สอนนิยมเสนอในรูปของแนว



การสอน (Course Outline) พิมพ์แจกผู้เรียนในแต่ละภาคการศึกษา ซึ่งประกอบด้วย ชื่อวิชา จำนวนหน่วยกิต ชื่อผู้สอน วันเวลาสถานที่เรียน เนื้อหาสาระของวิชาตามที่ระบุไว้ในหลักสูตร วัตถุประสงค์ของวิชา กิจกรรมการเรียนการสอน การแบ่งเนื้อหาสาระ (Content) ออกเป็นหน่วยย่อยซึ่งระบุกิจกรรมและสื่อวัสดุช่วยการเรียนการสอน การวัดและประเมินผล การเรียนการสอน ข้อตกลงอื่น ๆ (ได้แก่ กำหนดการส่งรายงาน สอบไล่ การตัดสิ้นเกรด) และเอกสารอ้างอิงหรือตำราที่ต้องการให้ผู้เรียนค้นคว้า

2. การแปลงจุดมุ่งหมายทางการศึกษาเป็นจุดมุ่งหมายเชิงพฤติกรรม โดยผู้สอนต้องกำหนดไว้ล่วงหน้าว่าต้องการให้ผู้เรียนบรรลุจุดมุ่งหมายใด และให้ผู้เรียนแสดงพฤติกรรมอะไรบ้าง

3. การกำหนดเครื่องมือวัดผลการเรียน ซึ่งมีหลายประเภทคือ แบบทดสอบ ความรู้ความสามารถด้านสมองที่ครูสร้างขึ้น แบบวัดความรู้สึก อารมณ์ ทักษะจิต ความสนใจ บุคลิกภาพ แบบสังเกตการปฏิบัติงาน พฤติกรรมที่แสดงความประพฤติกิ แบบสัมภาษณ์ แบบสอบถามความคิดเห็น และแบบทดสอบมาตรฐาน

4. การให้การทดสอบและการรวบรวมข้อมูล การทดสอบมีหลายแบบคือ สอบรายบุคคลทีละคน (สอบปากเปล่าหรือเขียนตอบ) เพื่อวัดความสามารถเฉพาะตัว และการสอบเป็นกลุ่มเพื่อการทำงานกลุ่ม คณะหน้าที่ให้เป็นของกลุ่ม สำหรับการรวบรวมข้อมูลอื่นที่ไม่ใช่การทดสอบมี 2 ระยะคือ ระยะระหว่างการเรียนการสอน เพื่อให้เห็นปัญหาในการเรียนการสอนซึ่งจะช่วยให้แก้ไขได้ทัน และระยะสิ้นสุดการเรียนการสอน เป็นการสรุปผลการเรียนการสอน ตลอดช่วงเวลาที่มีการเรียนการสอน เป็นการขมวดผลการเรียนของนิสิต และผลการสอนของอาจารย์

5. การจัดกระทำข้อมูล เพื่อการประเมินผล 3 ลักษณะคือ การประเมินบุคคลเป็นรายๆ เช่น ความสามารถในการเรียนของนิสิตแต่ละคน การประเมินนิสิตเฉพาะกลุ่มเพื่อสรุปอ้างอิงไปสู่กลุ่มใหญ่ และเป็นข้อมูลที่บอกประเภท ลำดับที่ หรือบอกช่วง

การจัดกระทำข้อมูลสำหรับบุคคลแต่ละคน โดยเทียบกับเกณฑ์หรือกลุ่มเพื่อแสดงให้เห็นจุดอ่อน จุดแข็ง การจัดกระทำสำหรับบุคคลเป็นกลุ่ม เพื่อให้เห็นภาพรวมว่าโดยเฉลี่ยเป็นอย่างไร คณะกรรมการเกาะกลุ่มกันอย่างไร ซึ่งอาจจะใช้สถิติภาคบรรยายช่วย แต่ถ้าต้องการอ้างอิงผลจากกลุ่มตัวอย่างไปหากกลุ่มประชากร อาจใช้เทคนิคสถิติสรุปอ้างอิงตามความเหมาะสมของข้อมูล เช่น ข้อมูลที่จัดเป็นช่วงใช้มีขนิมเลขคณิต ข้อมูลบอกลำดับที่

### ใช้มาตรฐาน ข้อมูลบอกประเภทใช้ฐานนิยม

6. การตัดเกรด หลังจากรวบรวมข้อมูลต่าง ๆ เกี่ยวกับการเรียน มีการวิเคราะห์ข้อมูล และจัดกระทำคะแนนรวมของผู้เรียนแต่ละคนแล้ว จะเป็นการพิจารณาให้เกรด ซึ่งอาจจะทำเฉพาะผลการเรียนแต่ละครั้ง โดยให้เกรดในการสอบย่อยแต่ละครั้ง หรือให้เกรดคนที่ทำส่ง และการให้เกรดในตอนท้ายของบทเรียนหรือหลักสูตร

จากจุดประสงค์ดังกล่าว การวัดและประเมินผลจึงเป็นสิ่งสำคัญและจะต้องจัดกระทำอย่างเป็นขั้นตอนมีกระบวนการที่แน่นอนในการวัดและประเมินผลการศึกษา นุญธรรม กิจปริดาบรสิทธิ (2533) ได้จัดลำดับขั้นตอนของการวัดและประเมินผลการศึกษาไว้ 8 ขั้นตอน ดังนี้

1. ตั้งวัตถุประสงค์ทั่วไป และวัตถุประสงค์เฉพาะ แล้วเขียนให้อยู่ในรูปวัตถุประสงค์เชิงพฤติกรรม
2. เลือกสถานการณ์ ประสบการณ์ ปัญหา หรือกิจกรรมต่าง ๆ เพื่อให้ผู้เรียนแสดงออกตามจุดมุ่งหมาย โดยพยายามเลือกให้ครอบคลุมเนื้อหาและพฤติกรรมที่ต้องการวัด
3. เขียนข้อสอบจากสถานการณ์ ประสบการณ์ ปัญหา และกิจกรรมนั้น เพื่อกระตุ้นให้ผู้เรียนแสดงพฤติกรรมที่ต้องการตามจุดมุ่งหมาย ทั้งนี้ผู้เขียนข้อสอบต้องคัดเลือกเครื่องมือวัดหรือชนิดข้อสอบ และวิธีดำเนินการวัดให้สอดคล้อง เหมาะสมกับวัตถุประสงค์ที่ต้องการวัดให้มากที่สุด
4. ดำเนินการสอบวัด โดยกำหนดเวลาสอบให้พอเหมาะ เพื่อให้ผู้สอบได้แสดงความสามารถอย่างเต็มที่
5. กำหนดหลักเกณฑ์ในการให้คะแนนให้ชัดเจนและมีความหมาย รวมทั้งวิธีการรายงานผลว่าเป็นอย่างไร เป็นร้อยละ หรือระดับคะแนน (Grade)
6. ตรวจสอบข้อสอบและวิเคราะห์ข้อสอบ เพื่อหาคุณภาพด้านความยากและอำนาจจำแนก
7. นำผลการสอบและผลการวิเคราะห์ข้อสอบมาอภิปรายร่วมกับผู้เรียนเพื่อเป็นผลย้อนกลับ หากจุดบกพร่องของผู้สอบ และเพื่อตรวจสอบความเป็นปรนัย ความชัดเจนของข้อสอบ
8. ปรับปรุงข้อสอบ จากผลการวิเคราะห์ และการอภิปรายข้อสอบจะสามารถนำไปพิจารณาปรับปรุงข้อสอบ เพื่อให้ได้ข้อสอบที่มีคุณภาพดีไว้ใช้ประโยชน์ต่อไป

ดังนั้นการวัดและประเมินผลที่มีประสิทธิภาพ จะต้องสอดคล้องหรือตรงกับจุดมุ่งหมายของการวัดและประเมินผล สามารถวัดคุณลักษณะที่ต้องการได้ โดยมีขั้นตอนการวัดอย่างมีระบบ มีระเบียบ มีกฎเกณฑ์ที่เชื่อถือได้และในการวัดและประเมินผลการศึกษานั้น มีสิ่งสำคัญที่ควรคำนึงถึงคือ (บุญเชิด วิทยุอนันตพงษ์, 2519)

1. กำหนดจุดมุ่งหมายของการประเมินผลให้ชัดเจนที่สุด
2. เลือกเครื่องมือให้สอดคล้องและเหมาะสมกับวัตถุประสงค์ของการวัด
3. พยายามใช้เทคนิคการวัดหลาย ๆ วิธี
4. ศึกษาจุดเด่นจุดด้อยของเครื่องมือวัดแต่ละชนิด เพื่อให้ใช้ได้ถูกวิธี
5. ระมัดระวังความผิดพลาด หรือความคลาดเคลื่อนของคะแนนที่อาจเกิดขึ้นในการวัด อันเนื่องมาจากธรรมชาติของสิ่งที่จะวัดถ้าเป็นนามธรรม และมีองค์ประกอบซับซ้อน มักจะมีความคลาดเคลื่อนเกิดขึ้น คุณภาพของเครื่องมือและตัวผู้วัดเอง ถ้าขาดความชำนาญในการใช้เครื่องมือก็อาจทำให้เกิดความคลาดเคลื่อน
6. การวัดผลการศึกษาเป็นการวัดระดับมาตราอันตรภาค (Interval Scale) ซึ่งเป็นการวัดที่บ่งบอกคุณสมบัติเท่านั้น การที่ผู้เรียนได้ศูนย์คะแนน มิได้หมายความว่าไม่มีความรู้เลย แต่เป็นเพราะทำข้อสอบไม่ได้ หรือข้อสอบยากเกินไป
7. การวัดผลเป็นการวัดทางอ้อม ไม่มีเครื่องมือใด ๆ ที่จะทำการวัดได้ตรง ๆ การวัดผลจึงต้องอาศัยการนิยามพฤติกรรมของแต่ละสมรรถภาพ แล้วจึงวัดพฤติกรรมเหล่านั้นออกมาเป็นคะแนนเพื่อแปลความหมายว่า มีสมรรถภาพนั้น ๆ สูงต่ำเพียงใด
8. การวัดผลการศึกษามีความคลาดเคลื่อน ซึ่งอาจเกิดขึ้นจากการนิยามพฤติกรรมที่ไม่ถูกต้องครอบคลุมหรือเครื่องมือไม่มีคุณภาพ ดังนั้นคะแนนที่วัดได้ (Observed Score) จึงประกอบด้วย 2 ส่วนคือ คะแนนจริงกับคะแนนคลาดเคลื่อน (True and error score) และความคลาดเคลื่อนจะมีอยู่มากน้อยเพียงใดขึ้นอยู่กับสิ่งที่วัด เครื่องมือวัดและตัวผู้วัดเอง
9. การวัดผลการศึกษาเป็นการวัดในลักษณะสรุปอ้างอิง เนื่องจากการวัดผลการศึกษานั้นเราวัดทางพฤติกรรมการแสดงออก ด้วยการวิเคราะห์กลุ่มของพฤติกรรมอย่างละเอียดเป็นมวลพฤติกรรมหรือประชากรของพฤติกรรมทั้งหมด แล้วจึงสุ่มตัวอย่างของพฤติกรรมที่เป็นตัวแทนที่ได้ออกมาทำการวัดผล การวัดนี้จึงเป็นการสรุปอ้างอิงสู่มวลพฤติกรรมทั้งหมดของสมรรถภาพนั้น ๆ

### ประเภทของการประเมินผล

การประเมินผลมีรูปแบบที่แตกต่างกัน ขึ้นอยู่กับวัตถุประสงค์ของการประเมิน โดยทั่วไปจำแนกประเภทของการประเมินผลตามระบบของการวัดผลออกเป็น 2 ประเภทคือ การประเมินผลแบบอิงกลุ่ม (Norm Referenced Evaluation) และการประเมินผลแบบอิงเกณฑ์ (Criterion Referenced Evaluation) การประเมินผลแบบอิงกลุ่มเป็นการประเมินผลพฤติกรรมของผู้เรียนแต่ละคนเปรียบเทียบกับพฤติกรรมของผู้เรียน ทั้งกลุ่มที่ทดสอบด้วยเครื่องมือวัดผลชนิดเดียวกัน โกวิท ประวาลพฤษ์ (2523) กล่าวว่า การประเมินผลแบบนี้จะชี้ให้เห็นถึงความแตกต่างของระดับการสัมฤทธิ์ผลของระดับคะแนนต่าง ๆ ที่อยู่ตามเส้นโค้งของการแจกแจงปกติ สำหรับการประเมินผลแบบอิงเกณฑ์ เป็นวิธีการพิจารณาตัดสินว่าผู้เรียนแต่ละคนประสบผลสำเร็จในพฤติกรรมแต่ละอย่างตามที่ระบุไว้ในจุดมุ่งหมายของการสอนหรือไม่ โดยไม่นำคะแนนของผู้เรียนแต่ละคนไปสัมพันธ์หรือเปรียบเทียบกับคะแนนของผู้เรียนคนอื่น ๆ ในกลุ่ม ซึ่งสอดคล้องกับแนวคิดของ Gronlund (1978) ที่กล่าวว่า การประเมินผลแบบอิงกลุ่มเป็นการตัดสินความรู้ความสามารถของผู้เรียนแต่ละคน เมื่อเปรียบเทียบกับผู้เรียนคนอื่น ๆ ในกลุ่มเดียวกัน หรืออาจเปรียบเทียบได้หลาย ๆ กลุ่ม ส่วนการวัดผลแบบอิงเกณฑ์เป็นการตัดสินว่าผู้เรียนแต่ละคนมีความรู้ความสามารถเท่าใด โดยเปรียบเทียบกับเกณฑ์ที่ตั้งไว้

Bloom (1971) ได้จำแนกการประเมินผลตามบทบาทที่ต้องใช้ออกเป็น 3 ลักษณะ คือ

1. การประเมินผลย่อย (Formative Evaluation) เป็นการประเมินผลในขณะที่ดำเนินการสอนแต่ละหน่วยการสอน เพื่อวัดระดับความรู้และค้นหาคำจุดที่ผู้เรียนยังบกพร่องหรือยังไม่รอบรู้ เพื่อปรับปรุงการเรียนการสอน
2. การประเมินผลรวม (Summative Evaluation) เป็นการประเมินผลเมื่อสิ้นสุดการเรียนการสอนรายวิชานั้น หรือกลางเทอม หรือสิ้นปี การประเมินผลนี้มีจุดประสงค์เพื่อให้เกรด
3. การประเมินผลวินิจฉัย (Diagnostic Evaluation) เป็นการประเมินผลในแต่ละหน่วยการเรียนเพื่อค้นหาว่าผู้เรียนบกพร่องตรงไหน ควรจะแก้ไขซ่อมเสริมตรงไหน ก่อนที่จะเริ่มสอนในบทเรียนต่อไป

ดังนั้นสรุปได้ว่า การประเมินผลแบบอิงกลุ่มเป็นการประเมินผลเพื่อเปรียบเทียบคะแนนของบุคคลภายในกลุ่มเพื่อหาความแตกต่างระหว่างผู้เรียนกับกลุ่ม ส่วนการประเมินผลแบบอิงเกณฑ์เป็นการเปรียบเทียบผลหรือคะแนนของผู้เรียนแต่ละบุคคลกับเกณฑ์มาตรฐานที่กำหนดขึ้นเพื่อตัดสินว่าผู้เรียนแต่ละคนมีความรอบรู้หรือไม่ ฉะนั้นการประเมินผลแบบอิงเกณฑ์จึงเหมาะสมสำหรับการประเมินผลย่อย (Formative Evaluation) และการประเมินผลวินิจฉัย (Diagnostic Evaluation) ส่วนการประเมินผลแบบอิงกลุ่มจะเหมาะสมสำหรับการประเมินผลรวม (Summative Evaluation) เพื่อพิจารณาตัดสินให้เกรด หรือเพื่อจัดอันดับความสามารถ ทั้งนี้ในการประเมินผลแต่ละครั้ง ผู้สอนจะต้องใช้เครื่องมือในการวัดและประเมินผลให้สอดคล้องกับจุดมุ่งหมายของการวัดแต่ละรายวิชา อาทิ รายวิชาที่มีลักษณะปฏิบัติการ วิชา ๒๕๒๙ ได้อ้างแนวคิดเกี่ยวกับการประเมินผลว่า องค์ประกอบที่นำมาตัดสินคะแนนควรประกอบด้วยการวัดที่ครอบคลุม 3 ด้าน คือ

1. ทักษะพิสัย (Psychomotor Domain) เป็นการพัฒนาการเคลื่อนไหวการทำงานของกล้ามเนื้อ ทักษะในการใช้มือ ผู้สอนควรทดสอบเกี่ยวกับความสัมฤทธิ์ผลในการเรียน ทักษะของกิจกรรมประเภทต่าง ๆ สมรรถภาพทางกาย ความสามารถด้านกลไก และความสามารถในการเล่น
2. พุทธิพิสัย (Cognitive Domain) เป็นการพัฒนาด้านสติปัญญา การระลึกได้ (Recall) การจำได้หรือการรู้จัก (Recognition) การวัดด้านนี้ได้แก่ ความรู้ ความเข้าใจเกี่ยวกับกลไก หลักการเคลื่อนไหว กฎ กติกา ความปลอดภัย การสร้างสมรรถภาพตลอดจนประวัติของกิจกรรมต่าง ๆ
3. จิตพิสัย (Affective Domain) เป็นการประเมินผลเกี่ยวกับพฤติกรรมทางสังคม ซึ่งประกอบไปด้วยความสนใจ ทศนคติ ความซาบซึ้ง ความมีน้ำใจเป็นนักกีฬา ความร่วมมือ ความสามารถในการปรับตัว การเป็นผู้นำและผู้ตาม และการวัดลักษณะทางสังคมนี้ วิธีที่ดีที่สุดคือ การสังเกตของผู้สอนและผู้เรียนโดยให้ผู้เรียนประเมินคุณค่าของตนเองและเพื่อนร่วมชั้น

นอกจากองค์ประกอบ 3 ส่วนนี้แล้ว ในการประเมินผลหรือการตัดสินระดับคะแนน (เกรด) มีองค์ประกอบอื่น ๆ ที่อาจนำมาพิจารณาคือ ความพยายาม เวลาเรียน การพัฒนาความสนใจ เพื่อให้ผลการประเมินถูกต้องสามารถจำแนกผู้เรียนได้สอดคล้องกับสภาพความเป็นจริงมากที่สุด มีความผิดพลาดน้อยที่สุด

## เครื่องมือวัดและประเมินผลการศึกษา

เครื่องมือเป็นสิ่งสำคัญอย่างหนึ่งของกระบวนการวัดและประเมินผลผู้สอนต้องเลือกใช้เครื่องมือให้เหมาะสมสอดคล้องกับจุดมุ่งหมายของการจัดการเรียนการสอน อุกุมพร จามรมาน (2530) ได้จำแนกประเภทของเครื่องมือที่สร้างตามจุดมุ่งหมายทางการศึกษาแต่ละด้านดังนี้

1. นพุทธิปริเขต (Cognitive Domain) เครื่องมือที่ใช้ได้แก่ แบบทดสอบผลสัมฤทธิ์ แบบทดสอบมาตรฐาน แบบสอบถามความคิดเห็นและแบบสัมภาษณ์
2. จิตตปริเขต (Affective Domain) เครื่องมือที่ใช้ได้แก่ แบบวัดความรู้สึก ความสนใจ ทศนคติ บุคลิกภาพ
3. พลังทักษะปริเขต (Psycho-Motor Domain) เครื่องมือที่ใช้ได้แก่ แบบสังเกตและแบบบันทึกพฤติกรรม เครื่องมือที่ใช้ในการวัดประเภทนี้จะสอดคล้องกับเครื่องมือที่ใช้ในการวัดผลภาคปฏิบัติ ซึ่ง กมล สุกประเสริฐ (2529) กล่าวถึงว่าเป็นการวัดในสถานการณ์ที่ผู้เรียนปฏิบัติจริง ไม่ว่าจะการวัดนั้นจะเป็นการวัดเพื่อพัฒนาการเรียนการสอน (Formative Evaluation) หรือวัดเพื่อประเมินผลรวม (Summative Evaluation) การวัดเช่นนี้ไม่อาจหลีกเลี่ยงการใช้วิธีการสังเกตได้ และควรจะใช้ควบคู่กันกับเครื่องมือวัด เช่น แบบสำรวจรายการ (Checklist) และการจัดอันดับคุณภาพ (Rating Scales) นอกจากนี้ บุญเรียง ขจรศิลป์ (2527) ได้กล่าวถึงเครื่องมือที่ใช้ในการวัดผลการศึกษาอีก 2 ประเภท คือ เครื่องมือที่เป็นแบบสอบและเครื่องมือที่ไม่เป็นแบบสอบ

1. เครื่องมือที่เป็นแบบสอบ ได้แก่ ข้อสอบรายบุคคลหรือข้อสอบกลุ่มข้อสอบที่ครูสร้างขึ้นเองหรือข้อสอบมาตรฐาน ข้อสอบปากเปล่าหรือข้อสอบข้อเขียน ข้อสอบวัดความสามารถและทักษะ (Mastery test) หรือข้อสอบวัดความคล่องในการทำข้อสอบ (Speed test) ข้อสอบที่ใช้ในการเขียนสอบหรือข้อสอบปฏิบัติ ข้อสอบวัดความพร้อมหรือข้อสอบวินิจฉัยข้อบกพร่อง และข้อสอบจัดบุคคลเข้ากลุ่ม (Placement test) จึงเป็นข้อสอบที่ใช้ตรวจสอบทักษะเบื้องต้นของผู้เรียน

2. เครื่องมือที่ไม่เป็นแบบสอบ ได้แก่ การสัมภาษณ์ แบบสอบถาม การบันทึกพฤติกรรม เครื่องมือวัดทางสังคม การเรียงอันดับและวิธีประมาณค่า

สำหรับเครื่องมือที่เป็นแบบทดสอบ วิริยา ญุฑิย (2529) ได้ให้ข้อคิดเห็นเพิ่มเติมว่า แบบทดสอบหมายถึงแบบ (form) หรือเครื่องมือ (tool) หรือกระบวนการสำหรับวัดความสามารถ ความสัมพันธ์ หรือความสนใจของบุคคลที่แสดงออกมา แบบทดสอบนี้ใช้วัดในสิ่งที่เราไม่สามารถวัดโดยตรงได้และจะวัดได้เมื่อบุคคลแสดงผลหรือกระทำออกมา ก่อน

แบบทดสอบในการวัดผลการศึกษา นั้น สามารถแยกได้ 2 ประเภทคือ

1. แบบทดสอบที่ครูสร้างขึ้นเอง (teacher-made test) ผู้สอนจะสร้างขึ้นเองเพื่อใช้กับผู้เรียน มีลักษณะดังนี้

- 1.1 เหมาะสมกับหน่วยการสอนที่ผู้สอนกำหนดเนื้อหาและความยากง่ายไว้
- 1.2 การสร้างแบบทดสอบนั้น วิธีการ เครื่องมือ และการให้คะแนนขึ้นอยู่กับข้อกำหนดของผู้สอน โดยอาศัยความเที่ยงตรงจากหลักสูตรเป็นเกณฑ์
- 1.3 แบบทดสอบอาจจะไม่เป็นไปตามคะแนนมาตรฐานของส่วนการศึกษานั้น ๆ แต่เป็นคะแนนที่ผู้สอนรวบรวมไว้ตลอดปี และสร้างคะแนนมาตรฐานขึ้นใช้เอง
- 1.4 เป็นแบบทดสอบที่สร้างขึ้นได้เร็ว ดังนั้นวิธีการอาจจะไม่ดีเท่ากับแบบทดสอบมาตรฐาน
- 1.5 ไม่เหมาะกับการนำไปใช้สำหรับผู้สอนคนอื่น ๆ เหมาะสำหรับใช้เฉพาะในส่วนการศึกษานั้น ๆ

2. แบบทดสอบมาตรฐาน (Standardized test) หมายถึง แบบทดสอบที่มีวิธีการ เครื่องมือ และการให้คะแนนคงที่ สามารถนำไปใช้ทดสอบ ต่างสถานที่ ต่างเวลาได้ การสร้างต้องออกข้อสอบหลาย ๆ ข้อ ทำการทดสอบกับคนเป็นจำนวนมาก นำข้อสอบกลับมาวิเคราะห์ โดยเลือกเฉพาะข้อสอบที่มีคุณภาพไว้ แบบทดสอบมาตรฐานนี้จะมีวิธีการใช้ เครื่องมือและการให้คะแนนที่คงที่ มีความเที่ยงตรง (Validity) ความเชื่อถือได้ (Reliability) และมีเกณฑ์ปกติ (Norm)

นอกจากนี้ อุทุมพร จามรมาน (2530) ได้กล่าวถึงหลักในการสร้างแบบทดสอบผลสัมฤทธิ์ ดังนี้

1. ควรวัดผลการเรียนที่สอดคล้องกับจุดมุ่งหมายในการเรียนการสอนทั้ง 3 ด้าน



คือ ด้านพุทธิ (Cognitive) จิตต (Affective) และพลังทักษะ (Psycho-motor) ที่ผู้เรียนได้แสดงออกมาหลังการเรียนรู้ โดยได้รับการระบุอย่างชัดเจนและเจาะจง

2. การสร้างแบบทดสอบควรจะสัมพันธ์เนื้อหาวิชาบางหน่วยมาเขียนข้อสอบ และควรสัมพันธ์ให้มากพอที่จะเป็นตัวแทนของผลการเรียนและเนื้อหาที่กำหนด

3. แบบทดสอบผลสัมฤทธิ์ ควรสอดคล้องกับประเภทของข้อสอบและความต้องการในการวัดผลการเรียน เพื่อให้สามารถกระตุ้นหรือเร้าให้ผู้เรียน ตอบสนองออกมา ซึ่งมี 2 ประเภทคือ

3.1 ประเภทให้ผู้เรียนตอบเองได้แก่ แบบทดสอบอัตนัยสมบูรณ์ แบบทดสอบอัตนัยภายใต้ข้อกำหนด แบบตอบสั้น ๆ แบบเติมให้สมบูรณ์

3.2 ประเภทให้คำตอบมาและผู้เรียนเลือกตอบ ได้แก่ แบบผิด-ถูก จับคู่ เลือกตอบ

4. แบบทดสอบควรเหมาะสมกับการใช้ เช่น เพื่อวัดความรู้ก่อน-หลังการเรียน เพื่อระบุความรอบรู้ของผู้เรียน เพื่อหาจุดอ่อน-จุดแข็งของผู้เรียน และเพื่อประเมินความก้าวหน้าในการเรียน

5. แบบทดสอบควรมีความเที่ยงสูงและการแปลผลควรทำอย่างระมัดระวัง

นอกจากแบบทดสอบผลสัมฤทธิ์ ซึ่งผู้สอนนิยมใช้ในการวัดผลการเรียน แล้วเครื่องมือที่ผู้สอนสามารถนำมาใช้ร่วมในการวัดและประเมินผลผู้เรียน คือ เครื่องมือวัดทางด้านจิตตปริเขต ได้แก่ แบบวัดทางจิตวิทยา เพื่อวัดทัศนคติต่อวิชาและผู้สอน บุคลิกภาพ ค่านิยม ความถนัด โดยส่วนใหญ่เครื่องมือเหล่านี้จะมีผู้สร้างไว้ และเครื่องมือวัดทางด้านพลังทักษะปริเขต ได้แก่ แบบสังเกตตามพฤติกรรม และแบบบันทึกพฤติกรรม ซึ่งสอดคล้องกับแนวคิดเกี่ยวกับการประเมินผลรายวิชาที่มีลักษณะปฏิบัติการของ จรินทร์ อานันรัตน์ (2519) ได้เสนอแนะวิธีการวัดผลเพื่อการประเมินผลทางผลศึกษาว่า ควรพิจารณาจากการวัดหลาย ๆ ด้านคือ

1. วัดโดยการให้ปฏิบัติ (Performance Test) โดยให้ผู้เรียนทดสอบปฏิบัติ ในกิจกรรมที่ผู้สอนต้องการทราบ เพื่อดูความสามารถและความสัมฤทธิ์ผลทางการเรียนการสอน

2. วัดโดยการทดสอบข้อเขียน (Written Test) เพื่อทดสอบความรู้ความเข้าใจ ข้อสอบที่ใช้อาจเป็นแบบทดสอบปรนัย อัตนัย หรือทั้งสองอย่างผสมกัน

3. วัดโดยใช้เครื่องมือในการวัด ได้แก่ การทดสอบสมรรถภาพทางกาย



## (Physical Fitness Test)

4. วัดโดยการสอบปากเปล่า (Oral Test) เป็นการวัดความรู้ ความเข้าใจทัศนคติ เป็นรายบุคคล ตามหัวข้อเรื่องที่ต้องการทราบในสิ่งที่สอนไปแล้ว การสอบปากเปล่านั้น ควรเตรียมคำถามเป็นข้อ ๆ ไว้ล่วงหน้าเพื่อให้เกิดความแน่ใจ และควรเตรียมคำถามไว้จำนวนมากพอเพื่อช่วยให้มีการเปรียบเทียบความเข้าใจระหว่างผู้เรียนแต่ละคนว่า มีความเข้าใจมากน้อยเพียงใด

5. วัดโดยการสังเกต (Observation) โดยการสังเกตจากการเรียนปกติ การฝึกหัดปฏิบัติในเวลาว่าง สิ่งที่ควรสังเกตคือ ความสนใจ ความสามารถ ทักษะในการเล่น การประสานงานช่วยเหลือกัน การควบคุมอารมณ์ ความมานะพยายาม ความอดทน ความรับผิดชอบ ความเป็นผู้นำหรือผู้ตาม ทั้งนี้การสังเกตควรมีหลักเกณฑ์ในการวัดเพื่อความแน่นอน แม่นยำ และความยุติธรรม

จากประเภทของเครื่องมือดังกล่าวอาจสรุปได้ว่า ในการวัดและประเมินผลทางการศึกษาเครื่องมือที่ใช้มีหลายประเภท ถ้าเป็นการวัดข้อความรู้โดยทั่วไปจะใช้แบบทดสอบได้แก่ ข้อสอบชนิดปรนัย อัตนัย ใช้สำหรับการประเมินความรู้ความสามารถเชิงทฤษฎี ถ้าเป็นการวัดทัศนคติ ความรู้สึก โดยทั่วไปจะใช้แบบสอบถามหรือแบบทดสอบบุคลิกภาพต่าง ๆ สำหรับการประเมินผลเพื่อทำนายหรือพยากรณ์ ความสนใจ ความถนัดด้านต่าง ๆ และถ้าเป็นการวัดการปฏิบัติ โดยทั่วไปจะใช้การสังเกต โดยใช้เครื่องมือวัดควบคู่ไปกับการสังเกต เช่น แบบสำรวจรายการ การจัดอันดับคุณภาพเพื่อวัดทักษะ ความแคล่วคล่องว่องไวในการปฏิบัติงาน ฉะนั้นอาจารย์ผู้สอนในฐานะที่เป็นผู้ประเมินจะต้องเข้าใจวิธีการสร้างแบบสำรวจรายการและการจัดอันดับคุณภาพ หรือจัดทำไว้สำหรับรายวิชาที่ตนเองสอน ยิ่งไปกว่านั้นอาจทำเป็นแบบทดสอบเพื่อใช้วัดวิธีปฏิบัติในรายวิชาของตน

**เกรดและการตัดเกรดในมหาวิทยาลัย**

การตัดเกรดเป็นวิธีการประเมินผลที่ต้องอาศัยการเปรียบเทียบคะแนนของผู้เรียนกับเกณฑ์ที่กำหนด เช่น เปรียบเทียบกับคะแนนผู้เรียนทั้งชั้น หรือเปรียบเทียบกับเกณฑ์มาตรฐานที่กำหนดไว้ แล้วตัดสินออกมาเป็นสัญลักษณ์หรือตัวอักษร ที่เรียกว่า เกรดซึ่งมีหลายระดับ โดยทั่วไป พบว่าสถาบันอุดมศึกษาไทย นิยมตัดเกรดเป็น 5 ระดับคือ A, B, C, D, E หรือ F บางสถาบันตัดสินเป็น 3 ระดับ เช่น G (หมายถึงขอเดชะ) P (ผ่าน) และ

## F (ไม่ผ่าน)

การประเมินผลของมหาวิทยาลัยมหิดล มหาวิทยาลัยมหิดลได้กำหนดระดับการ  
ตัดสินเกรดเป็น 8 ระดับ ซึ่งมีแต้มประจำและความหมายดังนี้ (มหาวิทยาลัยมหิดล  
บัณฑิตวิทยาลัย, 2534)

เกรด	แต้มประจำ	ความหมาย
A	4.00	ดีมาก
B <sup>+</sup>	3.50	ค่อนข้างดีมาก
B	3.00	ดี
C <sup>+</sup>	2.50	ค่อนข้างดี
C	2.00	พอใช้
D <sup>+</sup>	1.50	ค่อนข้างอ่อน
D	1.00	อ่อน
F	0	ตก

นอกจากนี้ การประเมินผลที่มีได้จำแนกเป็นระดับแต้มประจำ ได้แก่ การ  
ประเมินผลเป็นสัญลักษณ์ S (พอใจ) U (ไม่พอใจ) I (รอการประเมินผล) P (การศึกษา  
ยังมีได้สิ้นสุด) และ AU (การศึกษาโดยไม่รับหน่วยกิต)

ทั้งนี้ตามระเบียบบัณฑิตวิทยาลัย มหาวิทยาลัยมหิดล ว่าด้วยการวัดและประเมินผล  
การศึกษา พ.ศ.2523 ได้กำหนดหลักเกณฑ์การให้เกรด ดังนี้

1. รายวิชาบังคับที่เป็นรายวิชาเอกของหลักสูตร จะต้องได้สัญลักษณ์ไม่ต่ำกว่า  
B และรายวิชาเลือก หรือรายวิชาบังคับอื่น ๆ จะต้องได้สัญลักษณ์ไม่ต่ำกว่า C จึงจะถือว่า  
อยู่ในเกณฑ์มาตรฐาน

2. การนับหน่วยกิตรวมเพื่อให้ครบตามหลักสูตร ให้นับเฉพาะหน่วยกิตของรายวิชา  
ที่ได้สัญลักษณ์ไม่ต่ำกว่าเกณฑ์มาตรฐานเท่านั้น

3. การให้ F จะกระทำได้ในกรณีดังนี้

3.1 นักศึกษาเข้าสอบ หรือมีผลงานที่ประเมินผลว่าตก

3.2 นักศึกษาขาดสอบโดยไม่ได้รับอนุญาตจากคณบดีบัณฑิตวิทยาลัย

3.3 นักศึกษาทำผิดข้อบังคับ หรือระเบียบในการสอบ และได้รับการตัดสิน

ให้สอบตก

4. การให้ S หรือ U จะกระทำได้เฉพาะรายวิชาที่ไม่มีหน่วยกิต หรือมีหน่วยกิต แต่ภาควิชาเห็นว่าไม่ควรจำแนกผลการศึกษาที่มีแต้มประจำ ดังนี้
  - 4.1 ให้ S เมื่อการประเมินผลเป็นที่พอใจ
  - 4.2 ให้ U เมื่อการประเมินผลไม่เป็นที่พอใจ
5. การให้ I จะกระทำได้ในกรณี ดังนี้
  - 5.1 นักศึกษาไม่ได้สอบ และหรือไม่ได้ส่งผลงานเพราะป่วย โดยมีใบรับรองแพทย์ จากแพทย์ประจำหน่วยสุขภาพของมหาวิทยาลัยมหิดลรับรอง แต่นักศึกษาต้องมีเวลาเรียนอย่างน้อยร้อยละ 80
  - 5.2 นักศึกษาไม่ได้สอบ และหรือไม่ได้ส่งผลงานด้วยเหตุสุดวิสัย ให้อยู่ในดุลยพินิจของอาจารย์สอนและหัวหน้าภาควิชา
  - 5.3 อาจารย์ผู้สอนและหัวหน้าภาควิชาเห็นสมควรให้รอการประเมินผล นักศึกษาจะมี I ในระหว่างศึกษาได้ไม่เกิน 2 รายวิชา ยกเว้นกรณีที่ป่วยและการเปลี่ยนสัญลักษณ์ I นักศึกษาจะต้องดำเนินการในกรณีใดกรณีหนึ่ง ดังนี้
    - 5.3.1 สอบใหม่หรือส่งผลงานเพิ่มเติมภายใน 1 เดือน นับแต่วันถัดจากวันประกาศผล
    - 5.2.3 ปฏิบัติงานเพิ่มเติม และได้รับการประเมินผลใหม่ ภายในภาคการศึกษาถัดไป
    - 5.2.3 ให้เรียนซ้ำได้อีก 1 ครั้ง ในโอกาสแรกที่วิชานั้นเปิดสอน ถ้าไม่ดำเนินการให้แล้วเสร็จตามเวลาที่กำหนดนี้สัญลักษณ์ I จะเปลี่ยนเป็น F โดยอัตโนมัติ
6. การให้ P จะกระทำได้เฉพาะรายวิชาที่มีการสอนหรือปฏิบัติการต่อเนื่องกันมากกว่า 1 ภาคการศึกษา และหรือการศึกษาในรายวิชานั้นยังไม่สิ้นสุด
7. การให้ AU จะกระทำได้เฉพาะรายวิชาที่นักศึกษาขอเข้าร่วมศึกษาโดยไม่รับหน่วยกิต และต้องมีเวลาเรียนหรือปฏิบัติการไม่น้อยกว่าร้อยละ 80 ของเวลาเรียนหรือปฏิบัติการทั้งหมดในรายวิชานั้น

#### การให้เกรดในมหาวิทยาลัยต่างประเทศ

สำหรับในมหาวิทยาลัยต่างประเทศนั้น อุตมพร จามรมา (2533) ได้กล่าวถึง การให้เกรดในมหาวิทยาลัย Yale ว่าในปี 1783 ให้เกรดเป็น 4 ลักษณะ คือ Optime,

Second Optime Inferiores และ Pejora และต่อมาได้เปลี่ยนเป็นมาตร 4 ช่วง (1-4) โดยที่ 2 หมายถึง ค่าตรงกลาง ๆ ที่มหาวิทยาลัยมิชิแกน ในปี 1838 ได้ใช้การบรรยายลักษณะโดยใส่คำคุณศัพท์เช่นเดียวกับมหาวิทยาลัย Yale ในระยะแรกที่มหาวิทยาลัย วิลเลียมส์และแมรี ใช้การบรรยายด้วยข้อความที่แสดงระดับของความสามารถ 4 ระดับคือ

1 หมายถึง เป็นที่ 1 เป็นคนมีระเบียบ ตั้งใจเรียน มีพัฒนาการที่เร็วมาก

2 หมายถึง เป็นคนมีระเบียบ มีพัฒนาการที่จัดอยู่ในระดับที่น่าพอใจ

3 หมายถึง เป็นคนมีพัฒนาการเล็กน้อย

4 หมายถึง เป็นคนที่ไม่ค่อยเรียนรู้หรือมีความรู้ น้อย เนื่องจากความไม่ขยัน

ที่มหาวิทยาลัยฮาร์วาร์ด ระยะแรกใช้มาตร 20 ช่วง ต่อมาได้ใช้คะแนนเต็ม 100 เป็นหลัก ในการจัดกลุ่มนิสิตเป็น 5 กลุ่ม คือ กลุ่มได้คะแนน 100, 75-99, 51-74, 26-50 และ ต่ำกว่า 26 ซึ่งน่าจะเป็นจุดเริ่มต้นของการให้เกรดตามกลุ่มคะแนน โดยอาศัยตัวอักษร A, B, C, D, F นอกจากนี้ยังมีการให้เกรดที่อิงโค้งปกติ โดยอาศัยแนวคิดเกี่ยวกับทฤษฎี ความน่าจะเป็นในการแบ่งกลุ่มนิสิตเป็น 3% แรกคือ 22% ถัดมาคือ ตีมาก 50% ต่อมา คือ ปานกลาง 22% ต่อมาคือ อ่อน และ 3% สุดท้ายคือ ตก

จะเห็นได้ว่า เกรดมีหลายลักษณะคือ เป็นสัญลักษณ์ตัวอักษรหรือตัวเลขเป็นแบบ บรรยายลักษณะด้วยคำคุณศัพท์ การให้เกรดมีหลายระดับ ขึ้นอยู่กับจุดมุ่งหมายของการประเมิน ซึ่งจะต้องชัดเจนว่าต้องการจัดจำแนกผู้เรียนในลักษณะใด การให้เกรดจึงเป็นเรื่องที่ยุ่งยาก และซับซ้อน และเป็นเรื่องของการประมาณ ถ้าหากไม่เข้าใจนิยามของคำว่าเกรด หรือการ ให้เกรดไม่เป็นสากล อาจทำให้ขาดมาตรฐานในการให้เกรด และความหมายของเกรดที่ให้ จะเปลี่ยนแปลงหรือผิดพลาดได้ (วัณญา วิศาลาภรณ์, 2522) และในการให้เกรดอย่าง ยุติธรรม จากการวัดผลเพียงภาคปลายครั้งเดียวอาจไม่เพียงพอโดยเฉพาะผู้ที่สอบได้คะแนน ใกล้เคียงกัน แต่ต้องถูกตัดสินให้เกรดต่างกัน การพิจารณาว่าผู้เรียนควรได้เกรดใด จึงเป็น เรื่องยาก แต่ก็มีทางที่จะประมาณได้จากการวัดหลายครั้ง หรือใช้เกณฑ์หลายอย่าง แล้วนำ ผลการวัดเหล่านั้นมาเฉลี่ยกันก็จะได้คะแนนจริง หรือความสามารถที่แท้จริงของผู้เรียน (สวัสดิ์ ปทุมราช, 2531) ซึ่งสอดคล้องกับแนวคิดของ วิริยา บุญชัย (2529) ที่ได้กล่าว ถึงหลักการสำคัญในการให้เกรดว่าควรจะมีเกณฑ์ในการพิจารณาอื่น ๆ ดังนี้

1. ต้องคำนึงถึงวัตถุประสงค์ทางการศึกษา
2. ควรสอดคล้องกับวิชาอื่น ๆ ด้วย

3. ควรสอดคล้องกับนโยบายของผู้บริหาร
  4. ควรคำนึงถึงวัตถุประสงค์หลาย ๆ ข้อ
  5. ควรสัมพันธ์อย่างใกล้ชิดกับประสบการณ์ที่ได้รับอย่างแท้จริงในชีวิต
- การให้เกรดนั้น บุญเรียง ขจรศิลป์ (2527) กล่าวว่า โดยทั่วไปที่นิยมใช้มี 2 ระบบใหญ่ ๆ คือ

1. ยึดหลักของการประเมินอิงเกณฑ์ (Criterion-referenced Evaluation) เป็นการให้เกรดโดยเทียบกับมาตรฐานที่กำหนดไว้ล่วงหน้า

2. ยึดหลักการของการประเมินอิงกลุ่ม (Norm-referenced Evaluation) เป็นการให้เกรดโดยเปรียบเทียบกับนักศึกษาทั้งชั้น วิธีนี้คนได้คะแนนสูงสุดย่อมได้เกรดสูงสุด คนได้คะแนนต่ำสุดย่อมได้เกรดต่ำสุด

ในระดับมหาวิทยาลัย บุญธรรม กิจปรีดาบริสุทธิ์ (2533) กล่าวว่า วิธีการให้เกรดที่เหมาะสมที่สุดควรเป็นวิธีการของ Dewey B. Strit เนื่องจากเป็นวิธีที่มีระบบและมีเหตุผลที่สุดสำหรับนักศึกษาระดับมหาวิทยาลัย ซึ่งเป็นผู้ผ่านการคัดเลือกมาแล้วหลายครั้ง หนัน สະເພີຍຣຊັຍ (2513) กล่าวถึง หลักการให้เกรดของ Dewey ว่า

1. เป็นการสมมติว่า A B<sup>+</sup> B C<sup>+</sup> C D<sup>+</sup> D และ F เป็น Scale-score ที่มีช่วงห่างเท่ากัน
2. ไม่ปรับคะแนนให้เป็นการแจกแจงโค้งปกติ คงใช้ตามธรรมชาติของคะแนนดิบ
3. เป็นการให้คะแนนแบบตัดสินอิงเกณฑ์และอิงกลุ่มร่วมกัน
4. เหมาะสำหรับการให้คะแนนความสามารถของกลุ่ม โดยเฉพาะในกรณีที่อาจารย์คนเดียวต้องสอนนักศึกษาหลายห้องหรือหลายกลุ่ม เพราะระดับความสามารถหรือความสัมฤทธิ์ผล หรือพื้นฐานของนักศึกษาแต่ละกลุ่มแตกต่างกัน
5. เป็นวิธีคำนวณจากคะแนนมัธยฐาน (Median) เพราะคะแนนของนักศึกษา โดยเฉพาะในระดับมหาวิทยาลัยส่วนมากจะไม่แจกแจงเป็นโค้งปกติ การใช้คะแนนมัธยฐานจะเป็นตัวแทนค่าเฉลี่ยได้ดีที่สุด แล้วใช้ค่า Standard Deviation ซึ่งเป็นตัววัดการกระจายของคะแนน สำหรับแบ่งกลุ่มเป็นระดับเกรด
6. การให้เกรดวิธีนี้ไม่จำเป็นต้องมี A หรือ F เสมอไป ขึ้นอยู่กับดุลยพินิจของผู้ประเมิน

นอกจากนี้ ยังมีวิธีการให้เกรดตามแบบ Normal Curve การให้เกรดแบบนี้ ญอธรรม กิจปรีดาวิสุทธิ (2533) กล่าวว่า วิธีนี้ถือว่าคะแนนมีการแจกแจงแบบโค้งปกติ และถ้าไม่เป็นก็จะต้องปรับให้มีการแจกแจงแบบโค้งปกติ จุดอ่อนของวิธีนี้คือ ลักษณะที่เป็นการแจกแจงโค้งปกติ จะต้องเป็นกลุ่มที่มีจำนวนมาก ๆ และต้องมาจากการสุ่ม แต่นักศึกษาของมหาวิทยาลัยไทยมาจากการสอบคัดเลือกและแต่ละวิชาที่มันักศึกษาไม่มากนัก

สำหรับการให้เกรดนั้นได้มีนักวิชาการหลายท่านได้ให้ข้อคิดเห็นเกี่ยวกับการให้เกรด อาทิ กานดา พูนลาภทวี (2526) กล่าวถึง การให้เกรดภาคปฏิบัติว่าควรจะวัดจากความสามารถและทักษะในการปฏิบัติงาน ซึ่งได้แก่ วิธีการปฏิบัติงาน (Procedure) และผลงาน (Product) และวัดจากพฤติกรรมของผู้เรียนซึ่งได้แก่ ความรับผิดชอบ ความตั้งใจ ความสนใจ ฉะนั้นการให้เกรดรายวิชา นอกจากจะวัดความรู้ความสามารถในเนื้อหาแล้ว ยังต้องวัดทักษะการปฏิบัติงานและทัศนคติ ความรู้สึกต่าง ๆ ด้วย ซึ่งสอดคล้องกับแนวคิดของ สวัสดิ์ ปทุมราช (2531) ที่กล่าวว่า การให้เกรดรายวิชาจะต้องทำการสอบหลายครั้งและต้องใช้เกณฑ์การประเมินความรู้ความสามารถหลายอย่างตามความเหมาะสมกับรายวิชานั้น ๆ เช่น การทำภาคินพนธ์ รายงาน หรือภาคปฏิบัติอื่น ๆ โดยแต่ละครั้ง แต่ละอย่างควรตัดเกรดไว้ พร้อมทั้งกำหนดน้ำหนักสำหรับการสอบและงานแต่ละชนิดไว้ แล้วนำผลการสอบและงานอื่นมาเฉลี่ยตามน้ำหนัก เพื่อตัดสินเกรดครั้งสุดท้าย

Eble (1965) ได้ชี้ให้เห็นถึงความสำคัญของการให้เกรดว่า ถ้าระบบการให้เกรดของผู้สอนไม่มีการวางนโยบายให้สอดคล้องกัน การให้เกรดนั้นจะไม่มี ความหมายและจะใช้ประโยชน์ในด้านใดมิได้ การให้เกรดมักจะประสบปัญหามากในทุกระดับการศึกษา แต่ไม่มีผู้ใดแก้ปัญหาได้เด็ดขาด ทั้งนี้เนื่องจาก

1. ข้อบกพร่องในเรื่องหน่วยที่ใช้วัด
2. การขาดความมั่นคงในเครื่องมือที่ใช้วัด
3. ความยุ่งยากซับซ้อนของมาตรการที่ใช้วัดหรือคะแนน

Dourast (1962) กล่าวถึง เกณฑ์เฉพาะที่ควรจะมีติดมือว่าเป็นวิธีการให้เกรดที่ดี คือ

1. การให้เกรดหรือให้คะแนน ต้องเป็นปรนัยมากที่สุดเท่าที่จะเป็นไปได้
2. ต้องยอมรับความแตกต่างระหว่างบุคคลในแง่ศักยภาพในการเรียน

3. คะแนนพื้นฐานของบุคคล จะต้องแสดงให้เห็นถึงสถานะของแต่ละบุคคลในกลุ่มนั้น ๆ ที่เขามีส่วนเกี่ยวข้องกับอยู่ด้วย
4. คะแนนต่าง ๆ จะต้องง่ายต่อการคิดคำนวณและการตีความหมาย
5. วิธีการให้คะแนนจะต้องสอดคล้องกับระดับที่สูงขึ้นไป
6. วิธีการให้คะแนนจะต้องสามารถวิเคราะห์ได้อย่างมีความหมาย

Wrinkle (1935 อ้างใน วารี อินทรวีชา, 2515) กล่าวถึงลักษณะสำคัญของการให้เกรดที่ดีว่า จะต้องพิจารณาพื้นฐานสำคัญ ๆ หลายประการคือ ความเชื่อมั่นของการให้เกรด ความเป็นปรนัย การเปรียบเทียบการให้เกรดระหว่างอาจารย์ต่ออาจารย์ ชั้นต่อชั้น และระหว่างคณะต่อคณะ การใช้ประโยชน์ได้อย่างสะดวก และเป็นที่น่าสนใจกันได้ดี มีลักษณะเฉพาะที่เห็นเด่นชัด มีความยืดหยุ่น และเป็นที่ยังพอใจของผู้เรียนที่จะได้รับเกรดดังกล่าว

Hambleton และ Novick (1973) กล่าวถึง ปัญหาเบื้องต้นของการวัดผลแบบอิงเกณฑ์คือการกำหนดเกณฑ์ที่ผ่านหรือคะแนนจุดตัด (Cut-off Score) หรือคะแนนระดับความรู้ (Mastery level Score) เพื่อจัดผู้เรียนออกตามความสามารถที่ทำข้อสอบในแบบทดสอบอิงเกณฑ์ ให้เป็นผู้รอบรู้ (Master) หรือผู้ไม่รอบรู้ (Non-Master) และในการกำหนดเกณฑ์การสอบผ่านหรือคะแนนจุดตัดนั้นได้มีข้อเสนอไว้แตกต่างกันหลายวิธี แต่ยังเป็นข้อถกเถียงที่หาข้อยุติไม่ได้ว่าวิธีใดจะเป็นวิธีที่ดี ซึ่ง Millman (1973) ได้เสนอว่าควรพิจารณาองค์ประกอบหลาย ๆ อย่าง ประกอบกัน ได้แก่ (1) คะแนนของผู้สอบคนอื่น ๆ โดยกำหนดคะแนนจากกลุ่มที่เคยสอบครั้งก่อนทำได้ (2) เนื้อหาของข้อสอบ โดยพิจารณาข้อสอบแต่ละข้อแล้วนำมารวมกัน เช่น ให้ผู้เชี่ยวชาญตัดสินว่า ข้อสอบแต่ละข้อควรมีผู้ทำได้เท่าไร แล้วนำสัดส่วนของทุกข้อมาพิจารณาเป็นเกณฑ์ (3) ผลการศึกษาที่ผ่านมา คือ คะแนนเกณฑ์เท่าใด จึงจะทำให้เกิดประโยชน์แก่การศึกษามากที่สุด (4) ด้านจิตวิทยาและเศรษฐศาสตร์ที่มีต่อผู้เรียน ได้แก่ ค่าใช้จ่ายในการสอนซ่อมเสริมถ้าสูงมาก ก็ควรลดเกณฑ์ให้ต่ำลง ทั้งนี้ต้องคำนึงถึงความสูญเสียไปทางการศึกษาด้วย (5) ความคลาดเคลื่อนเนื่องจากการเดา และการเลือกตัวอย่างข้อสอบ ถ้ามีปัญหารื่องนี้ ควรกำหนดเกณฑ์ให้สูงขึ้น

Hambleton (1978) เสนอแนวทางในการกำหนดเกณฑ์การสอบผ่านว่า

1. ควรกำหนดเกณฑ์ร่วมกันหลาย ๆ กลุ่ม เช่น ผู้สอน ผู้ปกครอง ผู้เชี่ยวชาญ ด้านหลักสูตร ผู้บริหารโรงเรียน และผู้เรียน

2. การใช้วิธีการของอีเบล (Ebel) หรือเนเดลสกี (Nedelsky) ควรทำการทดสอบกับเด็กหลาย ๆ กลุ่ม

3. ผู้ทำการตัดสินเกณฑ์ควรนำตารางวิเคราะห์ข้อสอบมาเสนอและอภิปรายร่วมกัน

4. พยายามจัดเวลาให้เหมาะสมในตารางวิเคราะห์ข้อสอบ

5. ผู้ตัดสินเกณฑ์จะต้องมีความรู้เกี่ยวกับการใช้แบบทดสอบ และรู้ว่าจะใช้กับเด็กกลุ่มใด

6. จุดมุ่งหมายใดที่สำคัญมากกว่าควรตั้งเกณฑ์ที่สูง

7. ควรพิจารณาตัดสินเกณฑ์หลาย ๆ กลุ่ม ควรมีการหาค่าความสอดคล้องและอภิปรายร่วมกัน

8. นำข้อมูลเกี่ยวกับการสอบครั้งก่อน ๆ ที่มีประโยชน์มาใช้พิจารณาแก้ไขในการกำหนดเกณฑ์

9. ถ้าข้อมูลที่ได้จากการสอบเหมาะสมที่จะนำมาใช้พิจารณาเปอร์เซ็นต์ของผู้รอบรู้และไม่รอบรู้ ในแต่ละวัตถุประสงค์ก็ควรนำมาศึกษา หากคะแนนที่ได้จากแบบทดสอบมีลักษณะผิดปกติ (Out of line) เราสามารถอธิบายสาเหตุได้โดยการพิจารณาข้อสอบระดับของเกณฑ์แปรเปลี่ยนของคะแนนสอบระหว่างขั้นเรียนและปริมาณเวลาที่ใช้สอน

10. เปรียบเทียบสภาพการรอบรู้ของผู้ที่ไม่ได้เรียน และผู้ที่เรียนแล้ว ทั้งนี้กลุ่มที่เรียนแล้วควรจะมีควมรอบรู้มากกว่ากลุ่มที่ยังไม่ได้เรียน หากมีข้อผิดพลาดจากนี้ จะต้องพิจารณาเกณฑ์ให้มีความเที่ยงตรงมากขึ้น

Glass (1978) กล่าวถึงคะแนนจุดตัดหรือเกณฑ์ว่า เป็นมาตรฐานการปฏิบัติหรือความสามารถขั้นต่ำของผลการสอบผ่านและไม่ผ่าน ดังนั้นการกำหนดจุดตัดหรือเกณฑ์ จึงมีความสำคัญสำหรับการวัดผลแบบอิงเกณฑ์มาก ถ้ากำหนดคะแนนจุดตัดสูงไปหรือต่ำไปจะทำให้เกิดความคลาดเคลื่อน ซึ่ง อังคณา สายยศ (2525) กล่าวถึงความคลาดเคลื่อนว่ามี 2 ประเภทคือ ประเภทที่ 1  $\beta$ -error คือ กลุ่มที่รอบรู้แต่สอบไม่ผ่าน เกิดจากการกำหนดคะแนนจุดตัดสูงเกินไป ทำให้เด็กสอบตกมาก เสียเวลาและกำลังใจ สิ้นเปลืองเวลาในการสอบซ้ำ ประเภทที่ 2  $\alpha$ -error คือกลุ่มที่ไม่รอบรู้แต่สามารถสอบผ่านได้ เกิดจากการกำหนดคะแนนจุดตัดต่ำเกินไป ทำให้เด็กที่ไม่มีคุณภาพ ไม่มีมาตรฐานสอบผ่านได้ ซึ่งจะเกิดผลเสีย



ทำให้เด็กมีความสามารถไม่เพียงพอที่จะศึกษาในจุดมุ่งหมายใหม่ หรือศึกษาในระดับสูงต่อไปไม่ได้

นอกจากนี้ Glass (1978) ได้รวบรวมวิธีการกำหนดคะแนนเกณฑ์ผ่านหรือคะแนนจุดตัดซึ่งเป็นวิธีการหลัก ๆ ดังนี้

1. กำหนดเกณฑ์โดยใช้ผลการปฏิบัติของคนอื่น ๆ เป็นเกณฑ์ (Performance of Others as a Criterion) การตั้งเกณฑ์แบบนี้ใช้ค่ามัธยฐาน (Median) ของคะแนนจากแบบทดสอบ วิธีนี้นักทฤษฎีที่เกี่ยวข้องกับแบบทดสอบอิงเกณฑ์หลายคนเห็นว่าเป็นการอิงกลุ่ม
2. ใช้วิธีนับถอยหลัง 100 เปอร์เซ็นต์ (Counting backward from 100%) เนื่องจากในการสร้างแบบทดสอบให้สอดคล้องกับจุดมุ่งหมายในการเรียนรู้ทุกคนควรทำแบบทดสอบได้ถูกต้องหมด 100% แต่ทางปฏิบัติเป็นไปได้ยาก เพราะมีข้อบกพร่องและปัญหาอื่น ๆ จึงต้องใช้วิธีลดจุดตัดลงมาจาก 100% อาจจะลดลง 5 หรือ 20% หรือมากกว่า วิธีนี้เป็นเกณฑ์ที่ตั้งขึ้นตามอำเภอใจของผู้สอน
3. ใช้วิธีการเพิ่มคะแนนเกณฑ์อื่น ๆ (Bootstrapping on Other Criterion Score) วิธีนี้เป็นการกำหนดเกณฑ์โดยอาศัยเกณฑ์จากภายนอกมาเพิ่มขึ้น เช่น ใช้เกณฑ์ผู้ที่ได้รับประกาศนียบัตรในวิชาชีพนั้น ๆ เป็นผู้มีความสามารถ ผู้ที่ยังไม่ได้รับเป็นผู้ที่ไม่มีความสามารถ หรือใช้วิธีการพิจารณาการกระจายของคะแนนที่ได้จากการทำแบบทดสอบของผู้เรียน แล้วกำหนดคะแนนเกณฑ์เพื่อแยกผู้มีความสามารถออกจากผู้ไม่มีความสามารถ วิธีนี้มีปัญหาคือ ถ้าแบบทดสอบที่เราต้องการกำหนดจุดตัดมีค่าสัมพันธ์กับเกณฑ์ภายนอกต่ำ จะไม่สามารถกำหนดจุดตัดให้สอดคล้องกับเกณฑ์ภายนอกได้อย่างสมบูรณ์ ทำให้ผู้ผ่านเกณฑ์ภายนอกบางคนอาจได้คะแนนสูงกว่าจุดตัด และถ้ากำหนดจุดตัดให้สอดคล้องกับเกณฑ์ภายนอก ก็จะมีปัญหาว่าเกณฑ์ภายนอกที่ระบุขึ้นมานั้นเลือกมาได้อย่างไร วิธีการนี้ Berk (1976) ได้นำมาประยุกต์ใช้โดยกำหนดเกณฑ์ภายนอกว่า กลุ่มที่ได้รับการสอนให้เป็นพวกรอบรู้ และกลุ่มที่ไม่ได้รับการสอนเป็นพวกไม่รอบรู้ หรืออาจใช้นักเรียนกลุ่มเดียวกัน โดยให้เป็นกลุ่มก่อนเรียนและกลุ่มหลังเรียนก็ได้ หลังจากทีนักเรียนทั้งสองกลุ่มทำแบบทดสอบแล้ว นิยามการกระจายของคะแนนทั้งสองกลุ่มจะคาบเกี่ยวกัน จุดที่ฟังก์ชันทั้งสองตัดกันนั้นคือ คะแนนพยากรณ์ที่จะแบ่งการเรียนรู้ออกเป็น 4 พวกคือ (1) พวกรอบรู้จริง (True Masters) หมายถึง ผู้ที่ได้รับการสอนและได้คะแนนเท่ากับหรือมากกว่าเกณฑ์ (2) พวกรอบรู้ไม่จริง (False



Masters) หมายถึง ผู้ที่ไม่ได้รับการสอนแต่ได้คะแนนมากกว่าเกณฑ์ ซึ่งจัดเป็นความคลาดเคลื่อนประเภทที่ 2 (3) พวกไม่รอบรู้จริง (True Nonmasters) หมายถึง ผู้ที่ไม่ได้รับการสอน และได้คะแนนต่ำกว่าเกณฑ์ (4) พวกไม่รอบรู้ไม่จริง (False Nonmasters) หมายถึง ผู้ที่ได้รับการสอนแต่ได้คะแนนต่ำกว่าเกณฑ์ จะเป็นความคลาดเคลื่อนประเภทที่ 1 คะแนนพยากรณ์จะนำมาหาค่าคะแนนเกณฑ์ ซึ่งทำได้โดยเลื่อนคะแนนพยากรณ์ไปเรื่อย ๆ จนได้จุดตัดที่ให้ความน่าจะเป็นในการตัดสินใจถูกต้องสูงสุด หรือให้ความน่าจะเป็นในการตัดสินใจผิดพลาด น จุดคะแนนนั้น จะเป็นคะแนนจุดตัดหรือเกณฑ์ที่เหมาะสมที่สุด

4. นิยามตัดสินจากความสามารถต่ำสุด (Judging Minimal Competence) วิธีนี้ให้ผู้เชี่ยวชาญศึกษาจากแบบทดสอบหรือข้อคำถาม แล้วเสนอค่าความสามารถขั้นต่ำสุดที่ผู้สอบควรทำได้ จะทำให้ได้จุดตัดที่เหมาะสม

สงบ ลักษณะ (2522) กล่าวถึงประโยชน์ของการกำหนดเกณฑ์ว่า

1. เป็นการควบคุมมาตรฐานของการเรียนการสอน ช่วยให้สามารถแยกการเรียนของผู้เรียนได้เป็น 2 ประเภทคือ ผู้เรียนรู้ได้ตามจุดประสงค์และผู้รู้ได้ไม่ครบหรือผู้สอบผ่านกับผู้สอบไม่ผ่าน

2. ช่วยให้สามารถตรวจสอบได้ว่า ผู้เรียนคนใดมีความสามารถขั้นต้น (Prerequisite) ครบถ้วน สามารถเรียนต่อได้ตามลำดับขั้นของการเรียนรู้ (Learning hierarchy)

3. ใช้แยกผู้เรียนออกเป็นประเภทย่อย ๆ ตามระดับปริมาณและคุณภาพของการเรียน เช่น ใช้เพื่อให้ระดับคะแนน A, B, C

นอกจากนี้ การกำหนดคะแนนเกณฑ์ หรือคะแนนจุดตัดไว้ล่วงหน้าตายตัว เช่น ต้องสอบได้ 80% จึงจะถือว่าผ่าน เป็นผู้รอบรู้นั้นไม่เหมาะสม เนื่องจากขาดพื้นฐานที่มาของตัวเลขว่าเพราะเหตุใด การใช้จุดตัดจุดเดียวสำหรับข้อสอบทั้งฉบับก็ไม่มีเหตุผล เนื่องจากข้อสอบฉบับหนึ่งวัดมากกว่าหนึ่งจุดมุ่งหมาย เราจะทราบได้อย่างไรว่า ผู้เรียนผ่านจุดมุ่งหมายใดบ้าง มากน้อยเพียงใด และการใช้จุดตัดเดียวกันทุกวิชานั้นไม่มีเหตุผลเพียงพอ เนื่องจากการสัมฤทธิ์ผลมีลักษณะเป็นบางระดับ บางวิชาหรือบางจุดมุ่งหมาย ผู้เรียนจำเป็นต้องมีสัมฤทธิ์ผลในระดับค่อนข้างสูงจึงจะเพียงพอต่อการศึกษา บางวิชาหรือบางจุดมุ่งหมายอาจสำคัญน้อย ไม่จำเป็นต้องมีระดับของการสัมฤทธิ์ผลสูง

ประภา แก่นเพิ่ม (2524 อ้างใน สุธิวัตร ณ พิศักดีโสภณ, 2529) กล่าวว่า การกำหนดคะแนนเกณฑ์เพื่อแยกผู้เรียนเป็นผู้รอบรู้และไม่รอบรู้นั้น ถ้าให้ผู้สอนแต่ละคน กำหนดขึ้นเอง อาจเกิดปัญหาว่าผู้สอนแต่ละคนมีมาตรฐานหรือเกณฑ์ของตนเอง อาจทำให้เกิด ความเหลื่อมล้ำต่ำสูงได้

ดังนั้น อาจสรุปได้ว่า วิธีการให้เกรดกระทำได้หลายลักษณะ และการตัดสินใจให้ เกรดสุดท้ายควรจะมาจากการวัดและการประเมินหลาย ๆ ด้าน ทั้งข้อความรู้ในภาคทฤษฎี ทักษะการปฏิบัติงาน รวมทั้งทัศนคติ ความรู้สึก ความรับผิดชอบต่าง ๆ ด้วยวิธีการวัดหลาย ครั้งหลายรูปแบบ นอกจากนี้การกำหนดเกณฑ์ผ่านหรือคะแนนจุดตัดนั้น ควรจะพิจารณาจาก องค์ประกอบหลาย ๆ ด้าน เช่น คะแนนสอบในกลุ่มที่เคยสอบครั้งก่อนทำได้ เนื้อหาของ ข้อสอบ ผลการศึกษาที่จะก่อให้เกิดประโยชน์ต่อการศึกษามากที่สุด ด้านจิตวิทยา เศรษฐศาสตร์ และความคลาดเคลื่อนที่อาจเกิดขึ้นจากการเดาหรือการเลือกตัวอย่างข้อสอบ การให้เกรด จึงจะเป็นไปด้วยความยุติธรรมที่สุดสำหรับผู้เรียน

#### งานวิจัยที่เกี่ยวข้อง

วรารณ เอี้ยวสกุล และคณะ (2529) ได้ศึกษาองค์ประกอบที่มีผลต่อการศึกษา แบบหน่วยกิตของนักศึกษา ระดับปริญญาตรี มหาวิทยาลัยมหิดล และมหาวิทยาลัยอื่น ๆ พบว่า องค์ประกอบด้านการประเมินผลที่มีผลต่อการศึกษาระบบหน่วยกิตคือ นักศึกษาสอบตกแล้วแก้ตัว ไม่ได้ เนื่องจากไม่มีเวลาให้เรียนซ้ำ

สุริมาศ ธนพศินิติ และ ทิพยา กิจวิจารณ์ (2530) ได้ศึกษาปัญหาและความ ต้องการของนักศึกษาคณะแพทยศาสตร์ มหาวิทยาลัยขอนแก่น พบว่า นักศึกษามีความคิดเห็น เกี่ยวกับการวัดและประเมินผลว่า ควรปรับปรุงอย่างยิ่ง เพราะทำให้นักศึกษาเครียดและ แดกความสามัคคี และเสนอแนะว่า อาจารย์ควรประเมินผลแบบอิงเกณฑ์ ใช้คุณธรรมในการ ประเมินผล ข้อสอบควรมีคุณภาพดี และมีความยากง่ายพอเหมาะ ไม่ควรสอบบ่อยครั้งเกินไป ควรชี้แจงระบบการวัดผลของแต่ละวิชาให้ชัดเจน ตัวข้อสอบควรเน้นการแก้ปัญหา และถาม ประเด็นที่สำคัญ ๆ ควรออกข้อสอบให้ตรงกับเนื้อหาวิชาที่สอน ครอบคลุมเนื้อหาวิชา และ ควรลดเกณฑ์การวัดผลลงบ้าง

วิทยา ศรีนิมพ์มาตย์ (2532) ได้ศึกษาทัศนะของนักศึกษาคณะแพทยศาสตร์ มหาวิทยาลัยขอนแก่น ที่สอบตกซ้ำชั้น ในองค์ประกอบที่มีผลต่อการเรียนด้านการวัดและ ประเมินผลพบว่า ควรให้เรียนซ้ำวิชาที่สอบตก และควรชี้แจงระบบการวัดและประเมินผล

แต่ละวิชาให้ชัดเจน

เปรมใจ สุขสมานวงศ์ (2522) ได้ศึกษารูปแบบในการให้ระดับคะแนนแก่นักศึกษาของอาจารย์ในวิทยาลัยครูในกรุงเทพมหานคร พบว่า อาจารย์วิทยาลัยครูนิยมให้คะแนนในระดับคะแนนตามรูปแบบต่าง ๆ แตกต่างกันไป คือ รูปแบบอิงกลุ่ม รูปแบบอิงเกณฑ์ รูปแบบอิงเกณฑ์และอิงกลุ่ม และรูปแบบอิงความงอกงาม รูปแบบที่นิยมมากที่สุดคือ รูปแบบอิงกลุ่ม และพบว่ารูปแบบแบบอิงกลุ่มที่นิยมปฏิบัติมี 3 วิธีคือ การทำคะแนนที่ การใช้โค้งปกติ และการใช้วิธีการของสตีวีก รูปแบบอิงเกณฑ์ที่นิยมปฏิบัติ คือ อิงคะแนนเปอร์เซ็นต์ รูปแบบอิงเกณฑ์และอิงกลุ่ม วิธีที่นิยมปฏิบัติคือ การทำคะแนนที่ประกอบกับการพิจารณาคะแนนเปอร์เซ็นต์ การใช้โค้งปกติประกอบกับคะแนนเปอร์เซ็นต์ และรูปแบบอิงความงอกงาม วิธีที่นิยมปฏิบัติมี 2 วิธี คือ อิงเกณฑ์คะแนนเปอร์เซ็นต์ และอิงกลุ่มโดยใช้โค้งปกติ

สุวิมล ว่องวานิช (2533) ได้ศึกษาสภาพการวัดและประเมินผลการเรียนในจุฬาลงกรณ์มหาวิทยาลัย พบว่า (1) เครื่องมือที่ผู้สอนใช้ในการวัดความรู้ ทางทฤษฎี คือ ข้อสอบ ประเภทอัตนัยหรืออัตนัยคละกับปรนัยเป็นส่วนใหญ่ (2) จุดเน้นในการวัดผลภาคปฏิบัติ คือการวัดกระบวนการปฏิบัติมากกว่าการวัดผลการปฏิบัติ และพบว่าจุดเน้นของการวัดมีความแตกต่างกันระหว่างสาขาวิชา กล่าวคือ สาขาวิทยาศาสตร์ส่วนใหญ่เน้นกระบวนการปฏิบัติมากกว่าผลงาน ส่วนสาขาสังคมศาสตร์และมนุษยศาสตร์ เน้นวัดผลงานมากกว่า ทั้งนี้ เครื่องมือที่ใช้ส่วนใหญ่เป็นรายงานการปฏิบัติ และการสอบด้วยข้อสอบ การสอบโดยจัดสถานการณ์ให้นิสิตปฏิบัติ หรือการวัดด้วยการสังเกตพฤติกรรมมีน้อย (3) วิธีการวัดและประเมินผลที่อาจารย์ปฏิบัติแต่ละสาขาวิชามีความคล้ายคลึงกัน ในเรื่องของกระบวนการสร้างข้อสอบ การสอบ การตรวจ และการตัดเกรด (4) การตัดเกรดมีลักษณะเป็นแบบอิงตนเองไม่เป็นระบบอิงเกณฑ์หรืออิงกลุ่มอย่างแท้จริง เนื่องจากมีการปรับเปลี่ยนนิยามเกณฑ์การตัดสินเป็นส่วนใหญ่และพบว่ากลุ่มสาขาวิทยาศาสตร์ ส่วนใหญ่หัวหน้าภาควิชามีบทบาทในการตัดสินผลการเรียน โดยเฉพาะขั้นตอนการตัดเกรด กลุ่มสังคมศาสตร์และมนุษยศาสตร์ หัวหน้าภาควิชามีบทบาทน้อยหรือไม่มี (5) ระดับการศึกษาที่ต่างกันคือระดับปริญญาตรี โท เอก ใช้เครื่องมือและวิธีการวัดและประเมินผลการเรียนไม่แตกต่างกัน แต่มีความแตกต่างกันที่วิธีการสอน และในกรณีที่นิสิตบัณฑิตศึกษา เรียนภายในกลุ่มระดับการศึกษาเดียวกัน ผู้สอนกำหนดเกณฑ์การตัดสินในการตรวจให้คะแนนโดยตั้งระดับความคาดหวังที่นิสิตบัณฑิตศึกษาสามารถทำได้และหากมีการเรียนด้วยกันกับปริญญาตรี ผู้สอนมีการปฏิบัติที่แตกต่างกันเป็น 2 กลุ่ม โดย

ร้อยละ 52 ตัดสินตามมาตรฐานของวิชาโดยผู้สอนไม่ได้ตั้งเกณฑ์สูงกว่านิสิตปริญญาตรี ร้อยละ 48 ตั้งเกณฑ์ สำหรับบัณฑิตศึกษาสูงกว่า

สาธิต ลิ้มขจรเดช (2524) ได้ศึกษาปัญหาการวัดและประเมินผลการเรียนการสอนวิชากิจกรรมพลศึกษาในวิทยาลัยพลศึกษา พบว่า เห็นทั้งภาคทฤษฎีและปฏิบัติ การสอบข้อเขียนวัดความรู้เกี่ยวกับการนำไปใช้ การสอบภาคปฏิบัติ ใช้วิธีการสังเกตท่าทาง ความถูกต้อง การวัดทักษะก็นำประเภทรายบุคคลภาคปฏิบัติ ทดสอบทักษะและให้แข่งขัน การวัดผลประเภทชุดภาคปฏิบัติ สังเกตทักษะเป็นรายบุคคลและเล่นเป็นชุด เกณฑ์การประเมินผลใช้ตามเกณฑ์ที่ผู้สอนกำหนดโดยมีการแจ้งแผนการสอนและระยะเวลาการทดสอบ เมื่อเริ่มต้นภาคเรียน

สมพงษ์ เรืองศรี (2530) ศึกษาการวัดและประเมินผลวิชาสุขศึกษา พบว่า ส่วนใหญ่ผู้สอนใช้แบบทดสอบปรนัยในการวัดความรู้ ใช้การสังเกตในการวัดทัศนคติ และด้านการปฏิบัติใช้การสังเกต และการปฏิบัติกิจกรรมเป็นกลุ่ม

กาญจนา คงเจริญ (2520) ได้ศึกษาปัญหาการใช้หลักสูตรของวิทยาลัยพลศึกษา ฉบับพุทธศักราช 2520 พบว่า ในด้านการวัดและประเมินผล ผู้บริหารร้อยละ 92 และอาจารย์ร้อยละ 82 เห็นว่า ในการสอบควรมีการสอบทั้งภาคทฤษฎีและปฏิบัติ และควรมีการทดสอบกลางภาคและปลายภาค แล้วนำคะแนนมารวมกัน ผู้บริหารร้อยละ 100 เห็นว่าเกณฑ์การให้คะแนนต่าง ๆ ควรดูจากสมรรถภาพทางกาย ทักษะ ความรู้ ความเข้าใจ การพัฒนาทักษะ และการมาเรียนอย่างสม่ำเสมอ สำหรับอาจารย์ร้อยละ 92 เห็นว่าคะแนนควรมาจากความรู้และความเข้าใจมากที่สุด

นอกจากนี้ได้มีนักการศึกษาชาวต่างประเทศได้ทำการศึกษาวิจัยในช่วงปี ค.ศ. 1953-1980 เกี่ยวกับการให้คะแนนหรือการให้เกรดดังนี้

1. ความเชื่อมั่นในการให้คะแนนหรือให้เกรด จากการศึกษาวิจัยพบว่า ผู้สอนที่มีมาตรฐานในการให้เกรดใกล้เคียงกัน จะมีความเชื่อมั่นในการให้เกรดสูง ซึ่งแสดงให้เห็นว่าการให้เกรดของอาจารย์สอดคล้องสัมพันธ์กัน หรือเป็นมาตรฐานเดียวกันอย่างเชื่อถือได้มาก (Bending, 1953) สำหรับครูผู้สอนที่ไม่มีมาตรฐานในการตัดสินให้คะแนน พบว่าการให้คะแนนส่วนใหญ่ขาดความเชื่อมั่น เนื่องจากผู้สอนมักจะให้คะแนนตามความคิดเห็นหรือมาตรการของตนเอง (Ebel, 1965) ผลการวิจัยนี้สอดคล้องกับผลการศึกษาของ Travers และ Gronlund (1950) ซึ่งได้ศึกษาความคิดเห็นของอาจารย์ในมหาวิทยาลัยแห่งหนึ่งในสหรัฐอเมริกาเกี่ยวกับการให้คะแนนหรือการตัดสินผลการเรียนของนิสิต พร้อมทั้งสังเกตคะแนนเฉลี่ยและมาตรฐานของการให้คะแนนของอาจารย์ พบว่า ส่วนใหญ่จะแตกต่างกันมาก ทั้งนี้ได้มีการศึกษาวิจัยที่มีผลต่อความเชื่อมั่นของการให้คะแนน พบว่า ขึ้นอยู่กับความเที่ยงตรงของเครื่องมือที่ใช้ในการทดสอบ (Orde11, 1956)

2. การกำหนดเกณฑ์ ได้มีการศึกษาวิธีการกำหนดเกณฑ์พบว่า ในด้านความรู้ความคิด (Cognitive) เกณฑ์ที่ให้ผลสูงสุดคือ 95% ส่วนในด้านความรู้สึก (Affective) เกณฑ์ที่ให้ผลสูงสุดคือ 85% (Block, 1978 อ้างใน สุกฉิวรรณ นีรคัคคีโสภณ, 2529) ผลการศึกษานี้สอดคล้องกับแนวคิดในการหาคะแนนเกณฑ์หรือคะแนนจุดตัดที่เหมาะสม จากวิธีการหาคะแนนจุดตัดสี่แบบคือแบบไมโนเมียล (Binomial Model) แบบเบย์ (Bayesian Model) แบบใช้วิธีหาความน่าจะเป็น (Probabilistic Model) และการหาลัดส่วนของความแตกต่างในการจัดประเภท โดยใช้เกณฑ์ 80% ที่นิยมกันทั่วไป พบว่า ส่วนมากการหาคะแนนเกณฑ์แบบ Nondata Dependent Model ให้คะแนนสูงกว่าแบบ Data Dependent Model และในแบบทดสอบที่มีจำนวนข้อน้อยที่สุด (5 ข้อ) พบว่า การหาเกณฑ์ทุกวิธีให้ค่าตรงกันและพบว่า การหาคะแนนเกณฑ์ทุกแบบให้ค่าคะแนนเกณฑ์กลาง ๆ มากกว่าที่จะอยู่ต่อค่อนข้างต่ำหรือค่อนข้างสูง (Malech, 1979)

นอกจากนี้ได้มีการศึกษาหลักเกณฑ์ในการวัดผลประเมินผล โดยการพิจารณาขอบข่ายของการวัด จากการศึกษาความคิดเห็นของนิสิตฝึกสอนชั้นปีสุดท้าย ในมหาวิทยาลัยของรัฐ 4 แห่ง พบว่า ส่วนใหญ่ชอบวิธีการวัดผลจากงานที่ทำตลอดปี รองลงมาชอบวิธีการวัดผลจากการสอบไล่และผลงาน ส่วนน้อยมากชอบให้ตัดสินผลจากการสอบไล่อย่างเดียว (Starr, 1970)

สำหรับแนวคิดเกี่ยวกับเกรตเพอ ได้มีผู้ทำการศึกษาวิจัยดังนี้

Summerville และ Richard (1990) ได้ศึกษาภาวะเกรตเพอในวิทยาลัยชุมชน และมหาวิทยาลัยพบว่า ภาวะเกรตเพอยังคงเกิดขึ้นอยู่ และยังไม่สามารถหาระเบียบกฎเกณฑ์ มาควบคุมได้

Weller (1986) ได้ศึกษาทัศนคติต่อภาวะเกรตเพอ โดยการสุ่มสำรวจความคิดเห็นในวิทยาลัยศิลปะ วิทยาลัยวิทยาศาสตร์ และวิทยาลัยทางการศึกษาของอเมริกา จากการเปรียบเทียบทัศนคติของคณาบดีในวิทยาลัยทางศิลปะและวิทยาลัยทางการศึกษาเกี่ยวกับ ขอบเขตและสาเหตุของการเกิดภาวะเกรตเพอ ในวิทยาลัยเขตที่รับผิดชอบ พบว่า ส่วนใหญ่ จะมีความสนใจและห่วงใยต่อภาวะเกรตเพอที่เกิดขึ้น แต่บางส่วนมีความเห็นว่าควรจะทำการตกลงกันถึงสาเหตุที่ทำให้เกิดภาวะเกรตเพอ

Millman และ Jason (1983) ได้ทำการศึกษาว่าเกรตเพอมีอิทธิพลต่อความเชื่อมั่นของเกรตเพอหรือไม่ จากการพิจารณาอิทธิพลของเกรตเพอในการสอบทั้ง 2 ครั้ง โดยการสุ่มเกรตเพอในระดับที่ต่ำกว่าและมีความเชื่อมั่นของเกรตเพอเฉลี่ย พบว่า เกรตเพอมีผลต่อค่าความเชื่อมั่นอย่างมีนัยสำคัญ ยกเว้นการศึกษาระดับบัณฑิตศึกษาซึ่งเกรตเพอเกือบทั้งหมดมีเพียง 2 ระดับ คือ A และ B ผลกระทบอื่น ๆ มีเพียงเล็กน้อยเท่านั้นที่พบในอัตราส่วนสากลที่ต่าง ๆ กัน

Rogers และ Bruce (1983) ได้ใช้วิธีการของอนุกรมเวลาในการศึกษาเกรตเพอของระดับการศึกษาที่ต่ำกว่าบัณฑิตศึกษา พบว่า ในช่วง 15 ปีที่ผ่านมาได้เกิดการเปลี่ยนแปลงรูปแบบการให้เกรตเพอในระดับอุดมศึกษา ซึ่งมีผลต่อภาวะเกรตเพอ อันเป็นปรากฏการณ์ที่สังเกตได้ชัดว่าเป็นจุดอ่อนซึ่งมีความหมายต่อเกรตเพอ แต่ด้านอื่นจะเป็นผลกระทบในทางบวกสำหรับนักศึกษาซึ่งจะช่วยโน้มน้าวการเลือกเรียนรายวิชาที่มีผลตามระดับความสามารถของนักศึกษา นักวิจัยยังคงศึกษาแนวโน้มอย่างต่อเนื่อง แต่ก็ยังไม่สามารถหารูปแบบที่จะมานำเสนอเพื่ออธิบายความแปรผัน ในทางตรงกันข้ามกับช่วงเวลาที่ย้ายออกไปในขอบเขตของตัวแปรอิสระที่สามารถจะวัดได้ การศึกษาครั้งนี้มีวัตถุประสงค์เพื่ออธิบายรูปแบบ (Model) ของช่วงระดับข้อมูลเกรตเพอเฉลี่ย โดยการใช่วิธีการของอนุกรมเวลา จากข้อมูลเกรตเพอเฉลี่ย 53 ปีของมหาวิทยาลัยในแถบ Midwestern โดยวิเคราะห์ร่วมกับตัวแปรภายนอก (ซึ่งเปลี่ยนแปลงตามความสามารถของนักศึกษา ระดับการลงทะเบียน เฟส และสถานภาพทางเศรษฐกิจ) และความล่าช้าหลังภายในตัวแปร พบว่ามีเพียงระดับการลงทะเบียนที่สามารถอธิบายส่วนของเกรตเพอที่แปรเปลี่ยนได้อย่างมีนัยสำคัญ นั่นเป็นการลงความเห็นว่าเป็นช่วงเวลาอันสั้น การศึกษา

ข้อรอยอาจจะถูกจำกัดค่า ยกเว้นตัวแปรอิสระที่สามารถจะแปรเปลี่ยนได้ตามช่วงระยะเวลา

จากการศึกษาค้นคว้าเอกสารงานและงานวิจัยที่เกี่ยวข้อง ดังที่ได้กล่าวมาแล้ว แสดงให้เห็นว่า นอกจากองค์ประกอบด้านอื่น ๆ อาทิ การสอน หลักสูตร การเรียน สุขภาพ แล้วการวัดและประเมินผล เป็นองค์ประกอบที่สำคัญอย่างหนึ่งซึ่งสัมพันธ์กับการเรียนการสอน โดยตรง และเป็นเครื่องมือที่ผู้สอนยอมรับว่าสามารถยึดถือเป็นเกณฑ์ในการตัดสินผลการ ศึกษาเพื่อบ่งชี้ถึงศักยภาพของผู้เรียน หรือเพื่อจำแนกผู้เรียนว่าเป็นผู้รอบรู้และไม่รอบรู้ ทั้งนี้ วิธีการประเมินผลมีหลายรูปแบบ ผู้สอนจะต้องเลือกรูปแบบวิธีการประเมินผลที่เหมาะสม และยุติธรรมที่สุด มีความสอดคล้องกับธรรมชาติของลักษณะวิชาที่สอนและกลุ่มผู้เรียน และการประเมินผลที่ดีย่อมจะต้องมาจากวิธีการวัดผลที่มีประสิทธิภาพ ซึ่งจะต้องอาศัยวิธีการ วัดผลหลาย ๆ รูปแบบจึงจะสามารถวัดหรือจำแนกศักยภาพทางด้านความรู้ความสามารถของ ผู้เรียนได้ครบถ้วนในทุกด้าน และครอบคลุมวัตถุประสงค์ของการเรียนรู้ได้มากที่สุด โดยเฉพาะอย่างยิ่งในระดับบัณฑิตศึกษา ซึ่งถือว่าเป็นระดับการศึกษาขั้นสูง เกณฑ์ที่ผู้สอน ประเมินผลจะต้องมีระเบียบการวัดและประเมินผล ซึ่งกำหนดแนวทางการปฏิบัติและข้อ กำหนดมาตรฐานหรือเกณฑ์ขั้นต่ำของระดับเกรดตามลักษณะวิชา เช่น วิชาบังคับหรือวิชา เลิกเพื่อถ่วงดุลองผลการศึกษานักเรียนอีกครั้งหนึ่งว่า ถึงมาตรฐานในขั้นที่จะยอมรับให้ ผ่านหรือเป็นผู้รอบรู้ได้หรือไม่เพียงใด รายวิชาใดที่มีผลการศึกษาไม่ถึงเกณฑ์มาตรฐานขั้นต่ำ ที่กำหนดไว้ ในระเบียบการวัดและประเมินผลของสถาบันการศึกษา ผู้เรียนจะต้องเรียนซ้ำ รายวิชาใหม่ จนกว่าจะรอบรู้หรือมีผลการศึกษาได้มาตรฐาน ดังนั้นผู้วิจัยจึงเห็นสมควร อย่างยิ่งที่จะศึกษาวิธีการวัดและประเมินผลการศึกษาในระดับบัณฑิตศึกษา เพื่อให้ทราบระดับ เกรดที่ผู้สอนในระดับนี้ประเมินผล ตลอดจนวิธีการวัดและประเมินผลรายวิชาที่มีลักษณะ บรรยาย ปฏิบัติการ สัมมนา และฝึกภาคสนามของอาจารย์ที่ปฏิบัติในระดับบัณฑิตศึกษา