สรุปคำบรรยายโครงการเสริมสร้างทักษะวิชาการและวิชาชีพครู

ของ อาจารย์สุกัญญา เหลืองไชยยะ

หัวข้อเรื่อง การตัดเกรด

เมื่อวันที่ 13 ธันวาคม 2555

ณ ห้องบรรยาย 3 อาคารอำนวยการ ชั้น 3

การตัดเกรดและการวิเคราะห์ข้อสอบ

การตัดเกรด เป็นกระบวนการขั้นสุดท้ายของการจัดการเรียนการสอน วิธีการตัดเกรดไม่มีวิธีใดที่ดีที่สุด ศาสตร์ของการวัดมีทั้งแบบอิงเกณฑ์และอิงกลุ่ม วิธีการสอนต้องจัดให้เหมาะสมและสอดคล้องกับหลักสูตร

เป้าหมายของการศึกษา คือ มุ่งพัฒนาคนให้สมบูรณ์พร้อมทั้ง 3 ด้าน ได้แก่

1. ด้านสติปัญญา
2. ความรู้สึกนึกคิด (คุณธรรม) เป็นการวัดที่ยาก
3. การปฏิบัติ (มีหลายระดับ)

การพัฒนาจะต้องเกิดกับลูกศิษย์ทุกคน ไม่ใช่เพียงบางคนเท่านั้น ครูจะต้องมีความเชื่อว่าลูกศิษย์ทุกคนสามารถพัฒนาได้ ไม่ว่าจะมีพื้นฐานเดิมมาอย่างไร

ทฤษฎีทางการวัด

จะมีความเชื่อว่า ทุกสิ่งทุกอย่างสามารถวัดได้ และต้องวัดให้ตรงตามสภาพความเป็นจริง (Assessment)

การวัดประเมิน เรียกว่า Evaluation ทฤษฎีวัดผล แบบดั้งเดิมส่วนใหญ่จะวัดทางสติปัญญาเพียงอย่างเดียว ซึ่งมีข้อจำกัดเพราะครูสอนเด็ก และวัดผลโดยข้อสอบ ศาสตร์ของการวัดผล ถ้าครูสอนและเอาใจใส่เด็กอย่างดี ทำการวัดผลตามสภาพจริงเป็นระยะ ๆ ก็จะเป็นการวัดตามสภาพความเป็นจริง ที่ไม่มีข้อจำกัด

กระบวนการวัดและประเมินผล จะต้องประกอบด้วย

1. วัตถุประสงค์การศึกษา
2. การจัดการเรียนการสอน
3. การประเมินผล

บทบาทหน้าที่สำคัญของครู ในกระบวนการจัดการเรียนการสอน ประกอบด้วย

- การวัดและประเมินผลการศึกษา
- การตัดเกรด
- การวิเคราะห์ข้อสอบ

ครูต้องวิเคราะห์ว่า กระบวนการจัดการเรียนการสอน การวัด การประเมิน และการตัดเกรด มีข้อบกพร่องอย่างไร

การสร้างตารางวิเคราะห์หลักสูตร (แบบง่าย ๆ) เพื่อวัดประเมิน

พฤติกรรม เนื้อหา	ความรู้ ความจำ	ความเข้าใจ	การนำไปใช้	การวิเคราะห์	การสังเคราะห์
สถิติพรรณา					
-การเก็บรวบรวมข้อมูล	✓	✓			
-การนำเสนอ	✓	✓	✓		
-การวิเคราะห์	✓	✓	✓		
-การวัดการกระจาย	✓	✓	✓		

การออกข้อสอบ ต้องดูว่าจะให้คะแนนข้อใดมากข้อใดน้อย (ตามน้ำหนัก) การออกข้อสอบต้องดูให้เหมาะกับเวลา การวัดโดยข้อสอบต้องวัดในสิ่งที่สอน

คุณภาพของข้อสอบ ต้องประกอบด้วย

1. ความเที่ยง/ความเชื่อมั่น (Reliability) ความคงที่คงเส้นคงวาในการวัด คือ วัดเมื่อไร ที่ไหน ผลที่ได้จะเท่ากัน ข้อสอบที่ดีจะต้องมีการทำวิเคราะห์ข้อสอบ

2. ความตรง/ความเที่ยงตรง (Validity) เป็นความสามารถในการวัดสิ่งที่ต้องการได้ถูกต้องแม่นยำ ข้อสอบถ้ามีจำนวนน้อยเกินไป ความเที่ยงตรง ความตรงก็จะมีน้อย

การออกข้อสอบแต่ละครั้ง ธรรมชาติของคะแนนทางสถิติ จะมาจากคะแนนที่แท้จริง บวกกับความคลาดเคลื่อน

การวัดและประเมินผลการศึกษา

ก่อนการตัดเกรด ต้องมีการวัดแล้วจึงนำมาประเมินผล ส่วนใหญ่ใช้ข้อสอบเป็นเครื่องมือในการวัด โดยจะต้องวัดให้ครบทั้ง 3 ด้าน (ทำการวัดหลาย ๆ ครั้ง) การนำคะแนนลูกศิษย์มาตัดเกรด จะต้องนำคะแนนนั้นมาวัดหลาย ๆ ครั้ง น้ำหนักของคะแนนในแต่ละข้อไม่จำเป็นต้องเท่ากัน

การวัดผลทางการศึกษา ประกอบด้วย

- มาตราการวัด (Scale) ได้แก่
  - มาตรานามบัญญัติ ➡ เป็นสัญลักษณ์ นำมาเป็นคะแนนไม่ได้
  - มาตราเรียงอันดับ ➡ ได้แก่การจัดอันดับที่ 1,2,3 ซึ่งคะแนนอาจจะห่างไม่เท่ากัน (ช่วงห่างไม่คงที่)
  - มาตราอันตรภาค ➡ คือไม่มีศูนย์แท้ (เป็นศูนย์สมมติ)
  - มาตราส่วน ➡ มีศูนย์แท้ (เช่น น้ำหนัก ส่วนสูง) มีความหมายเชิงปริมาณที่สมบูรณ์แบบ
- เครื่องมือ (Instrument)
- หน่วยการวัด (Unit)

การประเมินผลทางการศึกษา

- เป็นการตัดสินผล/หาทางเลือกที่เป็นไปได้หลาย ๆ ทาง
- เป็นการตัดสินความสอดคล้องระหว่างการกระทำและวัตถุประสงค์ที่ตั้งไว้
- เป็นการตัดสินภายใต้คุณธรรม จรรยาบรรณครู และวิญญาณครู

ความคลาดเคลื่อนของการตัดเกรด

ความคลาดเคลื่อน อาจจะมาจากปัจจัยอย่างใดอย่างหนึ่งหรือหลายอย่าง ได้แก่

- ความคลาดเคลื่อนของเครื่องมือ เช่น ข้อสอบ
- ความคลาดเคลื่อนจากสภาพแวดล้อม เช่น เสียงดังรบกวนในห้องสอบ
- ความคลาดเคลื่อนจากสภาพของร่างกายและจิตใจ

การวัดจึงต้องมีหลายครั้ง (ซึ่งจะเป็นผลดี) โดยเอาคะแนนมาหาค่าเฉลี่ย ซึ่งจะมีค่าเท่ากับค่าเฉลี่ยที่แท้จริง ศาสตร์ทางวัดผลจะมีค่าเฉลี่ยสอดคล้องกับค่าทางสถิติ การรวมคะแนนที่มีน้ำหนักไม่เท่ากันอาจจะเกิดการผิดพลาดได้ ดังนั้นจะต้องนำคะแนนดิบมาแปลงเป็นคะแนนมาตรฐาน

วัตถุประสงค์ของการประเมินผล

1. บ่งชี้ว่านักศึกษามีทักษะ/ความรู้ที่สำคัญพอเพียงหรือไม่ ศาสตร์ของวัดผลเชื่อว่าลูกศิษย์ทุกคนสามารถพัฒนาได้
2. วินิจฉัยจุดเด่น-จุดด้อยของนักศึกษา โดยสังเกตจากผลสัมฤทธิ์ที่แสดงออกมาให้เห็นเพื่อพัฒนา ปรับปรุงแก้ไข
3. บ่งชี้ระดับสัมฤทธิผลในการเรียนของนักศึกษาแต่ละคน

การตัดเกรด

จะต้องสอดคล้องกับวัตถุประสงค์ของหลักสูตรและจะต้องสะท้อนความสามารถของผู้เรียน

ปัญหาของการตัดเกรด

1. ขาดมาตรฐานในการตัดเกรด
2. การสอบแต่ละครั้งมีคะแนนเต็มไม่เท่ากัน
3. หน่วยของคะแนนต่างกัน
4. การกระจายของคะแนนแตกต่างกัน

การจัดระดับคะแนนเป็นตัวอักษร

การตัดเกรดจะให้ความสำคัญตั้งแต่กระบวนการจัดการเรียนการสอน การวัด ประเมินผล โดยพิจารณาจาก

- ค่าระดับคะแนน ควรจะมีกี่เกรด
- กระบวนการจัดการเรียนการสอน
  * หลักสูตร
  * ประมวลการสอน
  * การวิเคราะห์หลักสูตร
- ข้อสอบ
- สภาพแวดล้อมในการสอน
- ความสามารถของผู้สอน

จัดระดับคะแนนตามธรรมชาติ

คือ การจัดให้คะแนนที่เกาะกลุ่มกันอยู่ในระดับเดียวกัน โดยกำหนดตัดตรงคะแนนที่เริ่มทิ้งช่วงห่างกัน (จะใช้ได้ดีเมื่อผู้เรียนมีจำนวนน้อย) แต่ถ้ามีผู้เรียนกลุ่มใหญ่ให้ใช้ระเบียบวิธีสถิติในการกำหนดช่วงคะแนนของแต่ละเกรด ใช้ค่าแนวโน้มเข้าสู่ส่วนกลางเป็นหลัก หรือจุดเริ่มต้นในการตั้งสเกล และใช้ค่าการกระจายของคะแนนเป็นตัวกำหนดช่วงคะแนน

ค่ากลาง ได้แก่ ค่าเฉลี่ย (Mean)
มัธยฐาน (Median) และ
ฐานนิยม (Mode)

ค่าการกระจาย ได้แก่ พิสัย (Range)
ส่วนเบี่ยงเบนมาตรฐาน (Standard deviation)

ตัวอย่างที่ 1

- นายเปรมสอบวิชาเคมีได้ 85 คะแนน จากคะแนนเต็ม 100 (อิงเกณฑ์)
- นายเปรมสอบวิชาเคมีได้เป็นที่ 1 จากผู้สอบ 160 คน (อิงกลุ่ม)

สรุป นายเปรมเรียนเคมีได้ดีมาก โดยพิจารณาทั้งอิงเกณฑ์และอิงกลุ่ม

ตัวอย่างที่ 2

นักศึกษา 2 กลุ่มเรียนวิชาเดียวกัน โดยใช้ข้อสอบที่เชื่อถือได้ชุดเดียวกัน

	กลุ่มที่ 1	กลุ่มที่ 2
คะแนนเฉลี่ย	90	90
ส่วนเบี่ยงเบนมาตรฐาน	8	10

น.ส.แอน (กลุ่มที่ 1) สอบได้ 85 คะแนน

น.ส.โอ๋ (กลุ่มที่ 2) สอบได้ 85 คะแนนเท่ากัน

คำถาม แอนและโอ๋ มีความสามารถเท่ากันหรือไม่

นำมาหาคะแนนมาตรฐาน $z = \frac{x - u}{\sigma}$

คะแนนมาตรฐานของ แอน $= \frac{85 - 90}{8} = -0.625$

คะแนนมาตรฐานของ โอ๋ $= \frac{85 - 90}{10} = -0.5$

คำตอบ โอ๋ มีความสามารถสูงกว่า แอน

การกำหนดช่วงคะแนนให้เกรดตามระเบียบวิธีสถิติ

- ใช้ค่าแนวโน้มเข้าสู่ส่วนกลางเป็นจุดเริ่มต้น ได้แก่
  - ฐานนิยม (Mode)
  - มัธยฐาน (Median)
  - ค่าเฉลี่ย (Mean)
- ใช้ค่าการกระจายของคะแนนเป็นตัวกำหนดช่วงเกรด
  - พิสัย (Range)
  - ส่วนเบี่ยงเบนควลไทล์ (Qualitle Devision)
  - ส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation)

กรณีคะแนนมีการแจกแจงเป็นโค้งปกติหรือใกล้เคียง

- จุดกลาง ใช้ Mean
- แบ่งช่วงคะแนนให้เกรดใช้ Standard Deviation

รูป 1

วิธีการและขั้นตอนการตัดเกรด

1. ประเมินความสามารถโดยเฉลี่ยของกลุ่มผู้เรียนว่าอยู่ในระดับใด (ดีเลิศ ดี พอใช้ อ่อน)
2. หาจุดกลางและค่าการกระจายของคะแนน
3. กำหนดขีดจำกัดล่างของเกรด A จากตารางของ Dewey
4. หาขีดจำกัดล่างของคะแนนที่จะให้เกรดแต่ละเกรด
5. ให้เกรดตามช่วงคะแนนที่ระบุไว้ในแต่ละเกรด

ตารางของ Dewey							
ระดับความสามารถเฉลี่ยของผู้เรียนทั้งกลุ่ม	**จุดต่ำสุดของเกรด A**	**ร้อยละของเกรด**					**ระดับความสามารถเฉลี่ย**
		A	**B**	**C**	**D**	**F**	
ดีเลิศ	0.7 SD	24	38	29	8	1	2.80
ดีมาก	0.9 SD	18	36	32	12	2	2.60
ดี	1.1 SD	14	32	36	15	3	2.40
พอใช้	1.3 SD	10	29	37	20	4	2.20
ปานกลาง	1.5 SD	7	24	38	24	7	2.00
อ่อน	1.7 SD	4	20	27	29	10	1.80
อ่อนมาก	1.9 SD	3	15	36	32	14	1.60

ตัวอย่างช่วงแบ่งเกรด

ตั้งแต่ $\overline{X}$ + 1.5 SD	ให้เกรด A	จำนวน 7%
0.5 SD – 1.5 SD	ให้เกรด B	จำนวน 24%
(-0.5) SD – 0.5 SD	ให้เกรด C	จำนวน 38%
(-1.5) SD – (-0.5) SD	ให้เกรด D	จำนวน 24%
น้อยกว่า (-1.5) SD	ให้เกรด F	จำนวน 7%

ข้อสังเกต

- การตัดเกรด ควรใช้หลาย ๆ วิธีร่วมกัน เพราะแต่ละวิธีจะมีปัญหาต่างกัน
- ไม่จำเป็นต้องตัดเกรดให้มีทุกเกรด
- ถ้าผู้เรียนกลุ่มเก่ง อาจไม่มีเกรด F
- ถ้าผู้เรียนกลุ่มอ่อน อาจไม่มีเกรด A

การกระจายของคะแนน อาจทำให้บางคนที่มีคะแนนโด่งสุด ได้เกรด A แน่นอนและคนที่มีคะแนนท้ายสุดจะได้เกรด F หรือไม่ต้องนำมาพิจารณาอีกครั้ง ดังนั้นควรแยกคน 2 ประเภทนี้ออกก่อน เนื่องจากคะแนนที่เก็บแต่ละครั้งมีการ error มาก ควรต้องย้อนมาดูที่กระบวนการเรียนการสอน

- ต้องใช้วิธีหรือระเบียบวิธีสถิติที่เหมาะสมกับชุดข้อมูลครั้งนั้น

การวิเคราะห์ข้อสอบ (Item Analysis)

การวิเคราะห์ข้อสอบเป็นการนำคำตอบของผู้เข้าสอบมาวิเคราะห์เพื่อตัดสินคุณภาพของข้อสอบว่ามีคุณลักษณะตรงตามวัตถุประสงค์ที่ผู้ออกข้อสอบต้องการหรือไม่ แล้วคัดเลือกข้อสอบที่มีคุณภาพดีเก็บไว้ใช้ต่อไป และเพื่อช่วยในการปรับปรุงข้อสอบที่มีคุณภาพไม่ค่อยดีให้เป็นข้อสอบที่ดีขึ้น

การวิเคราะห์ข้อสอบ นอกจากจะมีวัตถุประสงค์ตามที่กล่าวมาแล้ว ยังมีวัตถุประสงค์อีกประการหนึ่งที่สำคัญคือ เพื่อปรับปรุงคุณภาพของแบบสอบให้มีความเที่ยง (Reliability) และความตรง (Validity) สูงสุด

การวิเคราะห์ข้อสอบที่จะกล่าวต่อไปนี้ จำแนกเป็นสองส่วนคือ การวิเคราะห์ข้อสอบปรนัยและการวิเคราะห์ข้อสอบอัตนัย

การวิเคราะห์ข้อสอบปรนัย (ถูกให้ 1 ผิดให้ 0)

1. การหาประสิทธิภาพรายข้อ

1.1 ระดับความยาก (Level of difficulty : p)

ระดับความยากของข้อคำถามใด หมายถึง สัดส่วนของผู้ที่ตอบถูกในข้อนั้นเมื่อเทียบกับจำนวนผู้เข้าสอบทั้งหมด

$$P = \frac{n}{N}$$

โดยที่ n แทนจำนวนคนที่ตอบข้อนั้นถูก

N แทนจำนวนคนตอบทั้งหมด

ค่าระดับความยากจะมีค่าตั้งแต่ 0 ถึง 1 ยิ่งมีค่าเข้าใกล้ 1 แสดงว่าข้อสอบข้อนั้นง่าย

ตัวอย่างเกณฑ์ที่ใช้ในการแปลความหมายระดับความยากของข้อสอบ

p	ความหมาย
0.81 – 1.00	ง่ายมาก
0.61 – 0.80	ค่อนข้างง่าย
0.41 – 0.60	ง่ายพอเหมาะ
0.21 – 0.40	ค่อนข้างยาก
0.00 – 0.20	ยากมาก

1.2 อำนาจจำแนก (Power of Discrimination : r)

อำนาจจำแนกของข้อคำถามใด หมายถึง ประสิทธิภาพของข้อคำถามในการแบ่งผู้เรียนออกเป็นสองกลุ่ม คือ กลุ่มเก่งที่มีความรู้ กับกลุ่มอ่อนที่มีความรู้น้อย

$$r = \frac{n_H - n_L}{N_H} \quad \text{หรือ} \quad \frac{n_H - n_L}{N_L}$$

โดยที่ r แทนอำนาจจำแนกของข้อสอบ

n_H แทนจำนวนคนในกลุ่มสูง ที่ตอบข้อสอบข้อนั้นถูก

n_L แทนจำนวนคนในกลุ่มต่ำ ที่ตอบข้อสอบนั้นถูก

N_H แทนจำนวนคนทั้งหมดในกลุ่มสูง

N_L แทนจำนวนคนทั้งหมดในกลุ่มต่ำ

จะเห็นได้ว่าการหาอำนาจจำแนกนี้จะต้องจำแนกผู้เรียนออกเป็น 2 กลุ่ม คือ กลุ่มสูง กับ กลุ่มต่ำ ซึ่งอาจทำได้ดังนี้

1.2.1 ใช้เทคนิค 27% หมายถึงใช้ 27% ของคนที่ได้คะแนนมากเป็นกลุ่มสูง และ 27% ของคนที่ได้คะแนนน้อยเป็นกลุ่มต่ำ ทั้งนี้เมื่อแบ่งคนเป็นกลุ่มสูงและกลุ้มต่ำแล้ว จำนวนคนในแต่ละกลุ่มไม่ควรต่ำกว่า 40 คน

1.2.2 ใช้เทคนิค 50% ในกรณีนี้ใช้เมื่อผู้เข้าสอบมีจำนวนน้อย

ค่าอำนาจจำแนกสามารถหาได้อีกวิธีหนึ่งคือ หาค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนสอบของผู้สอบทั้งหมด กับคะแนนสอบของข้อที่กำลังพิจารณาโดยใช้สูตรสัมประสิทธิ์สหสัมพันธ์แบบพอยท์ ไบซีเรียล (point biserial correlation) (Guilford 1965 : 498)

$$r_{pbi} = \frac{X_p - X_q}{S} \sqrt{p(1-p)}$$

โดยที่ r_{pbi} แทนค่าอำนาจจำแนก

X_p แทนค่าเฉลี่ยเลขคณิตของคะแนนสอบของคนที่ตอบข้อนี้ถูก

X_q แทนค่าเฉลี่ยเลขคณิตของคะแนนสอบของคนที่ตอบข้อนี้ผิด

S แทนส่วนเบี่ยงเบนมาตรฐานของคะแนนสอบทั้งฉบับ

p แทนค่าระดับความยากของข้อนี้

ค่าอำนาจจำแนกจะมีค่าได้ตั้งแต่ -1 ถึง +1 ซึ่งถ้ามีค่าเป็นลบแสดงว่าเป็นข้อสอบที่ไม่ดี ควรตัดทิ้งไป หรือต้องแก้ไขปรับปรุง โดยทั่วไปข้อสอบที่ดีนั้นควรมีอำนาจจำแนกตั้งแต่ 0.20 ขึ้นไป

Ebel (1972) เสนอแนะเกณฑ์ที่ใช้พิจารณาค่าอำนาจจำแนกของข้อสอบดังนี้

ค่าอำนาจจำแนก (ของตัวถูก)	**คุณภาพของข้อสอบ**
0.40 ขึ้นไป	ดีมาก
0.30 ถึง 0.39	ดีพอสมควร อาจต้องปรับปรุงบ้าง
0.20 ถึง 0.29	พอใช้ได้ แต่ต้องปรับปรุง
ต่ำกว่า 0.20	ใช้ไม่ได้ ต้องพิจารณาปรับปรุงใหม่ หรือตัดทิ้งไปเลย

1.3 ประสิทธิภาพของตัวลวง

ตัวลวง ก็คือ ตัวเลือกที่ผิด ดังนั้นตัวลวงที่มีประสิทธิภาพหมายถึง ตัวลวงที่ดึงดูดคนในกลุ่มต่ำมาเลือก ในขณะที่คนในกลุ่มสูงไม่เลือก และตัวลวงแต่ละตัวในข้อสอบควรจะต้องทำหน้าที่ไม่เช่นนั้นเราควรจะต้องปรับปรุงแก้ไขตัวลวงใหม่

2. การหาประสิทธิภาพรวมทั้งฉบับ

ค่าความเที่ยงหรือความเชื่อมั่นของแบบทดสอบ

โดยใช้สูตร KR_{20} ของ Kuder-Richardson

$$r_{tt} = \frac{k}{k-1}\left(1 - \frac{\sum pq}{S^2}\right)$$

เมื่อ r_{tt} แทน ความเที่ยงของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน

k แทน จำนวนข้อสอบ

p แทน สัดส่วนของผู้ตอบถูกในข้อนั้น

q แทน สัดส่วนของผู้ตอบผิดในข้อนั้น หรือ 1-p

S^2 แทน ความแปรปรวนของคะแนนจากการสอบ

ค่าความเที่ยงหรือความเชื่อมั่นของแบบทดสอบ จะมีค่าได้ตั้งแต่ -1 ถึง +1

- ถ้ามีค่าเป็นลบ แสดงว่าเป็นแบบทดสอบที่ใช้ไม่ได้ (ผิดปกติ)
- ถ้ามีค่าเป็น 0 แสดงว่าไม่มีความเที่ยง
- ถ้ามีค่าเข้าใกล้ +1 มากเท่าไร แสดงว่าเป็นแบบทดสอบที่ดี มีความเที่ยงสูง

(โดยปกติค่าความเที่ยงหรือความเชื่อมั่นของแบบทดสอบมักมีค่าเป็นบวกที่มากกว่า 0)

ประโยชน์ของการวิเคราะห์ข้อสอบ

1. ทำให้รู้คุณภาพของข้อสอบในเรื่องระดับความยาก อำนาจจำแนก และประสิทธิภาพของตัวลวง
2. สามารถแก้ไข ปรับปรุงข้อสอบได้ตรงจุดมุ่งหมาย
3. ช่วยให้ทราบว่า การเรียนการสอนบรรลุตามความมุ่งหมายที่วางไว้หรือไม่ เพื่อปรับปรุงแก้ไขคุณภาพการสอน
4. ช่วยให้ครูเขียนข้อสอบได้เก่งขึ้น
5. ช่วยให้สามารถสร้างข้อสอบหลาย ๆ ฉบับในวิชาเดียวกัน เรียกว่า ข้อสอบคู่ขนาน (Paralleled Test) และอาจปรับปรุงเป็นข้อสอบมาตรฐาน (Standardized-Test)
6. เป็นแนวทางในการสร้างคลังข้อสอบ (Item Bank)

สรุป การประเมินผลหรือการตัดเกรดที่ดี

ผลที่ได้จะต้องสอดคล้องกับความสามารถที่แท้จริงของผู้เรียน และยุติธรรมต่อผู้เรียนทุกคนทุกภาคการศึกษา

1. คุณภาพของเครื่องมือวัด เช่น ข้อสอบ
2. รูปแบบของการวัด
3. ระดับความยาก-ง่าย ของเครื่องมือวัด รวมทั้งเกณฑ์และวัตถุประสงค์การเรียนรู้ที่กำหนดไว้
4. ระดับความสามารถของกลุ่มผู้เรียน
5. พื้นฐานของหลักการ และเหตุผลเกี่ยวกับทฤษฎีการวัดและการประเมิน
6. วิจารณญาณ และคุณธรรมของผู้ประเมิน