คะแนนและการให้ระดับคะแนน 11

วัตถุประสงค์

เมื่อผู้อ่านได้ศึกษาเนื้อหาบทที่ 11 แล้ว ควรจะมีความสามารถดังนี้

1. บอกความหมายและหลักการของคะแนนแปลงรูปแบบต่าง ๆ ได้
2. สามารถแปลงคะแนนดิบให้เป็นคะแนนมาตรฐานในรูปแบบต่าง ๆ ได้
3. สามารถนำวิธีการให้ระดับคะแนนแบบต่าง ๆ ไปใช้ได้อย่างเหมาะสม
4. วิเคราะห์ถึงปัญหาของการให้ระดับคะแนน

เนื้อหา

11.1 คะแนนและความหมายของคะแนน

ปัญหาสำคัญของการวัดผลประการหนึ่งก็คือคะแนน ครูเป็นส่วนมากเข้าใจว่าเมื่อได้คะแนนเป็นตัวเลขมาแล้วก็เป็นอันสิ้นสุดกระบวนการในการวัดผล ซึ่งนับว่าเป็นความเข้าใจที่ไม่ถูกต้อง ครูควรจะนำเอาคะแนนผลการสอบนั้นมาจัดกระทำเพื่อให้เกิดประโยชน์แก่นักเรียน ผู้ปกครอง และผู้มีส่วนเกี่ยวข้องให้มากที่สุด เพื่อที่จะได้ใช้ผลการสอบให้คุ้มค่าสมกับความเหนื่อยยาก ผลการสอบของนักเรียนนั้นครูควรจะรายงานให้นักเรียนหรือผู้ปกครองได้ทราบอย่างง่าย ๆ โดยครูจะต้องพยายามทำให้คะแนนมีความหมายง่ายต่อการอ่านและการเข้าใจ ถ้าหากนักเรียนและผู้ปกครองไม่สนใจหรือไม่เข้าใจความหมายของคะแนนแล้ว ผลเสียย่อมอาจเกิดขึ้นกับนักเรียนได้ เช่นนักเรียนอาจเลือกเรียนผิด คือเรียนในสาขาวิชาที่ตนไม่มีความรู้หรือไม่มีความถนัดได้ ซึ่งสิ่งนี้จะไม่ก่อให้เกิดประโยชน์กับตัวนักเรียนเท่าที่ควร

คะแนนที่ได้จากการสอบวัดแต่ละครั้ง จะให้ความหมายได้มากน้อยเพียงใดย่อมขึ้นอยู่กับคุณภาพของการวัดในครั้งนั้น ๆ คะแนนที่ได้จากการสอบวัดแต่ละครั้งมักจะมีปัญหายุ่งยากในการแปลความเสมอ ทั้งนี้เพราะคะแนนเป็นเพียงตัวเลขที่บอกจำนวนผลงานที่นักเรียนทำได้ถูกต้องเท่านั้น หาได้บอกถึงปริมาณความรู้ความสามารถที่แท้จริงของนักเรียนไม่ นอกจากนี้คะแนนในการสอบวัดแต่ละครั้งมักจะมีความคลาดเคลื่อนเกิดขึ้นเสมอ นั่นคือคะแนนที่ได้จากการสอบแต่ละครั้งจะประกอบด้วยคะแนน 2 ส่วนด้วยกัน คือ คะแนนที่ได้จากความสามารถจริงของคน ๆ นั้น กับอีกส่วนหนึ่งคือคะแนนความคลาดเคลื่อน ซึ่งคะแนนส่วนหลังนี้อาจเนื่องมาจากสาเหตุหลายประการ อาทิเช่น การเดา เป็นต้น นั่นคือนักเรียนอาจได้คะแนนโดยที่นักเรียนไม่มีความรู้เลยก็ได้ หรือนักเรียนอาจไม่ได้คะแนนทั้ง ๆ ที่มีความรู้ในเรื่องนั้น ซึ่งอาจสรุปได้ว่าคะแนนที่ได้จากการสอบจะไม่เท่ากับคะแนนความสามารถจริง ๆ ของนักเรียน

จะเห็นได้ว่าปัญหาเรื่องคะแนนนั้นนับว่าเป็นปัญหาที่ยิ่งใหญ่ปัญหาหนึ่งของการวัดผลการศึกษา ดังนั้นจึงสมควรที่จะต้องมาพิจารณาในเรื่องของคะแนนให้รอบคอบและละเอียดถี่ถ้วน

11.2 คะแนนดิบไม่มีความหมาย

โดยปกติการวัดทางการศึกษาหรือจิตวิทยานั้น จะวัดได้ไม่เกินระดับ Interval scale ทั้งนี้เพราะจะไม่มีแบบทดสอบฉบับใดที่จะสามารถวัดปริมาณความรู้ที่ครบถ้วนของมนุษย์ได้

และคะแนนที่ได้จากการสอบวัดแต่ละครั้งมักจะมีความคลาดเคลื่อนเจือปนอยู่เสมอ ดังได้กล่าวมาแล้ว ข้อสอบเป็นเพียงการสุ่มเนื้อหาวิชามาถามเท่านั้น คะแนนที่ได้จากการสอบจึงมิอาจแทนจำนวนของความรู้ความสามารถที่แท้จริงของนักเรียนได้ ดังนั้นจะมาสรุปว่าคะแนนที่ได้จากการสอบเป็นมาตรฐานตายตัวเหมือนกับการชั่งน้ำหนักหรือการวัดส่วนสูงย่อมไม่ได้

ตามธรรมดาคะแนนดิบที่ได้จากการสอบจะไม่มีความหมายในตัวของมันเองจำเป็นที่จะต้องรู้ข้อมูลอย่างอื่นเพิ่มเติมอีกจึงจะทำให้คะแนนมีความหมายมากขึ้น ตัวอย่างเช่น สมมุติว่ามีนักเรียนคนหนึ่งไปบอกกับพ่อแม่ของเขาว่า เขาสอบวิชาวัดผลการศึกษาได้ 20 คะแนน ตัวเลข 20 นี้ จะไม่ช่วยให้พ่อแม่ของนักเรียนทราบอะไรเลย คือไม่รู้ว่าลูกของตนเก่งหรืออ่อน ทั้งนี้เพราะไม่รู้ว่าได้ 20 คะแนน จากคะแนนเต็มเท่าไร ถึงแม้จะทราบเพิ่มเติมอีกว่าวิชานี้ คะแนนเต็ม 40 คะแนน ก็ยังไม่สามารถสรุปได้ว่านักเรียนคนนี้เก่งหรืออ่อนในวิชาวัดผลการศึกษา เพราะว่าคะแนน 20 นี้ อาจจะเป็นคะแนนสูงสุดหรือต่ำสุดในชั้นก็ได้ หรือจะมาสรุปว่านักเรียนคนนี้มีความรู้ครึ่งหนึ่งหรือมีความสามารถปานกลาง โดยยึดหลัก 50 เปอร์เซ็นต์ก็ไม่ถูกต้อง ทั้งนี้เพราะไม่สามารถจะออกข้อสอบให้ครอบคลุมความรู้ความสามารถทั้งหมดในเรื่องนั้น ๆ ได้ และเนื่องจากข้อสอบเป็นเพียงการสุ่มตัวอย่างความรู้มาถามเท่านั้น จากตัวอย่างข้างต้นนี้อาจกล่าวได้แต่เพียงว่านักเรียนคนนี้ทำข้อสอบถูกครึ่งหนึ่ง

ถ้าหากจะสมมุติต่อไปว่า ข้อสอบฉบับนี้มี 20 ข้อ และนักเรียนคนนี้ทำถูกทั้งหมดคือได้ 20 คะแนน จะสรุปว่านักเรียนคนนี้เก่งได้ไหม คำตอบก็คือไม่แน่ เพราะข้อสอบอาจจะง่ายเกินไป นักเรียนในชั้นทุกคนอาจได้คะแนนเต็มก็ได้ หรือถ้าคะแนน 20 คะแนนนี้เป็นคะแนนสูงสุด ก็ยังจะต้องมาพิจารณาว่าเป็นคะแนนสูงสุดจากนักเรียนกี่คน 500 คน หรือ 20 คน เป็นคะแนนสูงสุดจากห้องที่เรียนอ่อนหรือเรียนเก่ง

ในทำนองเดียวกัน การที่สุรศักดิ์ สอบวัดผลได้ 0 คะแนน จะมาสรุปว่า สุรศักดิ์ไม่มีความรู้ในวิชาวัดผลย่อมไม่ได้ นั่นคือตราบใดก็ตามที่สุรศักดิ์เป็นมนุษย์ และมาเรียนหนังสือสม่ำเสมอ ย่อมเป็นไปไม่ได้ที่สุรศักดิ์เรียนแล้วจะไม่มีความรู้เพิ่มขึ้นเลย การที่สุรศักดิ์ได้ 0 คะแนนนั้น อาจเป็นเพราะข้อสอบไม่ถามในสิ่งที่สุรศักดิ์รู้ แต่ไปถามในสิ่งที่สุรศักดิ์ไม่รู้ 0 คะแนนในที่นี้มีความหมายเพียงว่าสุรศักดิ์ทำข้อสอบฉบับนี้ผิดทั้งหมด

จากตัวอย่างที่กล่าวมาข้างต้น จะเห็นได้ว่าคะแนนดิบไม่ได้ให้ความหมายอะไรมากนักในการตีความหมายของคะแนน จึงจำเป็นต้องอาศัยข้อมูลอย่างอื่น ๆ เพิ่มเติมอีกหลายประการมาประกอบการพิจารณา สาเหตุที่คะแนนดิบไม่สามารถบอกอะไรได้มากนักเพราะ

1. คะแนนดิบไม่มีความหมายในตัวของมันเอง หากไม่นำไปเปรียบเทียบกับคนอื่น หรือเกณฑ์มาตรฐานที่มีอยู่แล้ว

2. คะแนนดิบเป็นผลจากการวัดในมาตราที่ไม่มีค่าศูนย์จริง ดังนั้นค่าตัวเลขที่ได้จากการวัดจึงไม่ได้แสดงถึงปริมาณความรู้ที่แท้จริงของนักเรียนแต่อย่างใด

3. ความแตกต่างระหว่างคะแนนดิบเปรียบเทียบกันไม่ได้ เช่น 10 คะแนน ระหว่าง 80-90 จะไม่เท่ากับ 10 คะแนนระหว่าง 10-20

11.3 การรวมคะแนนดิบ 2 วิชา

การนำคะแนนดิบ 2 วิชามารวมกันนั้น เป็นการกระทำที่ไม่ถูกต้อง เพราะผิดหลักของคณิตศาสตร์ เช่น ในเรื่องการบวกเศษส่วนนั้น จะบวกเศษส่วนโดยที่ส่วนไม่เท่ากันย่อมไม่ได้ แต่ที่ผ่านมาครูก็ทำผิด ๆ เช่นนี้มาตลอดเวลา เช่น แดงสอบคณิตศาสตร์ได้ 15 คะแนน จากคะแนนเต็ม 20 คะแนน และสอบพลานามัยได้ 7 คะแนน จากคะแนนเต็ม 10 คะแนน ครูก็จะนำคะแนนทั้งสองวิชาของแดงมารวมกันเป็น $\frac{15}{20} + \frac{7}{10} = \frac{22}{30}$ ซึ่งนับได้ว่าการบวกเช่นนี้เป็นการผิดหลักของการบวกเศษส่วนที่ว่า เศษส่วนจะบวกกันได้ก็ต่อเมื่อมีส่วนเท่ากันเสียก่อน

ทุกคนคงจะยอมรับว่า 1 คะแนนในวิชาพลานามัย กับ 1 คะแนนในวิชาคณิตศาสตร์ ใช้ความสามารถทางสมองไม่เท่ากัน เพราะเป็นเรื่องของความคิดคนละชนิด เช่นเดียวกับการนำเงินตราต่างประเทศมารวมกับเงินไทยโดยตรงไม่ได้ ทั้งนี้หากต้องการจะรวมเป็นหน่วยเดียวกันก็จำเป็นจะต้องแปลงให้เป็นมาตราเดียวกันเสียก่อน โดยอาจแปลงเงินต่างประเทศเป็นเงินไทย หรือแปลงเงินไทยเป็นเงินต่างประเทศก็ได้ แล้วจึงนำมารวมกัน คะแนนผลการสอบแต่ละวิชาก็เช่นกัน จะนำมารวมกันโดยที่ไม่ได้แปลงให้มีหน่วยเท่ากันก็ย่อมไม่ได้ เช่นจะนำ 1 คะแนนวิชาพลานามัย มารวมกับ 1 คะแนน ในวิชาคณิตศาสตร์ ก็จะเท่ากับ 2 อะไรไม่รู้ ดุจเดียวกับการนำเอา 1 ดอลลาร์มารวมกับ 1 บาท จะเท่ากับ 2 อะไรไม่ทราบเช่นกัน

ถึงแม้ว่าทั้งสองวิชาจะมีคะแนนเต็มเท่ากัน ก็ไม่สามารถจะนำคะแนนทั้งสองวิชามารวมกันได้ ทั้งนี้เพราะลักษณะการกระจายของคะแนนในทั้งสองวิชาต่างกัน ตัวอย่างเช่น ในการสอบวิชาวิทยาศาสตร์และคณิตศาสตร์ ซึ่งทั้งสองวิชานี้มีคะแนนเต็ม 50 คะแนน แต่ความยากง่ายและการกระจายของคะแนนต่างกัน ดังนี้

จำปีได้คะแนนสูงสุดวิชาวิทยาศาสตร์ คือได้ 25 คะแนน แต่ได้คะแนนต่ำสุดวิชาคณิตศาสตร์ คือได้ 20 คะแนน ส่วนจำปาได้คะแนนสูงสุดวิชาคณิตศาสตร์ คือได้ 40 คะแนน แต่ได้คะแนนต่ำสุดในวิชาวิทยาศาสตร์ คือได้ 15 คะแนน

จะมาสรุปว่าจำปาเก่งกว่าจำปีเพราะผลรวมของคะแนนดิบเป็นเช่นนั้นย่อมไม่ได้ เพราะตามความเป็นจริงแล้ว ทั้งจำปีและจำปาต่างก็ได้คะแนนสูงสุดวิชาหนึ่งต่ำสุดวิชาหนึ่ง และทั้งสองวิชานี้มีความสำคัญเท่า ๆ กัน นักเรียนทั้งสองคนนี้ก็น่าที่จะเก่งพอ ๆ กัน แต่จากการรวมคะแนนดิบจึงทำให้สรุปผิดไป สาเหตุที่เป็นเช่นนี้เพราะการกระจายของคะแนนในวิชาคณิตศาสตร์เท่ากับ 20 ซึ่งเป็น 2 เท่าของการกระจายของคะแนนในวิชาวิทยาศาสตร์ ซึ่งเท่ากับ 10 ดังนั้นจึงทำให้น้ำหนักของวิชาคณิตศาสตร์เป็น 2 เท่าของวิชาวิทยาศาสตร์

ฉะนั้น หากต้องการให้ 2 วิชานี้มีน้ำหนักเท่ากันจริง ๆ ก็จะต้องทำให้การกระจายของคะแนนเท่ากันก่อน นั่นคือเอา 2 คูณวิชาวิทยาศาสตร์ หรือเอา 2 หารวิชาคณิตศาสตร์อย่างใดอย่างหนึ่ง แล้วรวมคะแนนใหม่จะได้ดังนี้

จำปี $(25 \times 2) + 2 = 70$

จำปา $(15 \times 2) + 40 = 70$

จากตัวอย่างดังกล่าวนี้เป็นการพยายามแปลงคะแนนให้มีหน่วยหรือมาตรฐานเดียวกันเสียก่อนจึงค่อยนำมาเปรียบเทียบกัน ซึ่งจะเห็นได้ว่าเมื่อทำให้การกระจายของคะแนนเท่ากันแล้ว จำปีและจำปา ต่างก็มีความสามารถพอ ๆ กัน คือต่างก็ได้ 70 คะแนน ซึ่งต่างกับผลการรวมคะแนนดิบในครั้งแรกที่ปรากฏผลออกมาว่าจำปีได้ 45 คะแนน ส่วนจำปาได้ 55 คะแนน ดังนั้นเพื่อไม่ให้เกิดความเข้าใจผิดดังกล่าวอีก จึงจำเป็นต้องแปลงคะแนนดิบให้อยู่ในรูปของคะแนนมาตรฐาน

11.4 **คะแนนมาตรฐาน** (Standard scores)

ดังได้กล่าวมาแล้วว่าคะแนนดิบไม่ได้ให้ความหมายอะไรมากนัก จึงได้มีผู้คิดหาวิธีการที่จะเปลี่ยนคะแนนดิบเหล่านั้นให้มีความหมายมากขึ้น และสามารถนำคะแนนทั้งหลายเหล่านั้นมาเปรียบเทียบกันได้ โดยพยายามทำให้คะแนนดิบเป็นคะแนนปรุงแต่ง (Derived scores) อันได้แก่คะแนนมาตรฐานแบบต่าง ๆ ที่จะกล่าวถึงต่อไป แต่ก่อนที่จะกล่าวถึงคะแนนมาตรฐานแบบต่าง ๆ ก็ควรที่จะได้มาทำความเข้าใจเกี่ยวกับความหมายและหลักใหญ่ ๆ ของคะแนนมาตรฐานเสียก่อน

คะแนนมาตรฐาน เป็นหน่วยของการวัดชนิดหนึ่งที่แปลงรูปมาจากคะแนนดิบ โดยทำให้คะแนนมีหน่วยเท่ากัน สามารถนำมาเปรียบเทียบกันได้อย่างมีความหมาย หลักใหญ่ของคะแนน

มาตรฐานก็คือ การนำเอาผลการสอบของแต่ละคนไปเปรียบเทียบกับกลุ่มใหญ่ หรือกลุ่มคนทั้งหมดที่อยู่ในฐานะหรือสภาพแวดล้อมเดียวกัน

จะเห็นได้ว่าคะแนนมาตรฐานเกี่ยวข้องกับคะแนนเฉลี่ย คือยึดเอารายเฉลี่ยของทุก ๆ คนในกลุ่มเป็นหลัก แล้วดูการกระจายของคะแนนแต่ละคนว่าห่างจากรายเฉลี่ยมากน้อยเท่าใด และหาค่าเฉลี่ยมาตรฐานของความแตกต่างจากจุดที่เป็นค่าเฉลี่ยโดยวิธีการทางสถิติ และเรียกค่านี้ว่าความเบี่ยงเบนมาตรฐาน (Standard deviation) ส่วนวิธีการคำนวณค่าความเบี่ยงเบนมาตรฐานนั้นได้กล่าวมาแล้วในบทที่ว่าด้วยสถิติเบื้องต้นสำหรับการวัดผลและประเมินผลในวิชา MR 303

คะแนนมาตรฐานที่นิยมใช้กันอยู่มีหลายชนิด ซึ่งสามารถแบ่งตามลักษณะการเปลี่ยนรูปคะแนนได้ 2 แบบ คือ

1. **การแปลงรูปคะแนนแบบเส้นตรง** (Linear transformation) เป็นการแปลงรูปคะแนนให้มีหน่วยใหม่ โดยลักษณะการกระจายของคะแนนไม่แตกต่างไปจากลักษณะการกระจายของคะแนนเดิม กล่าวคือ ถ้าการกระจายของคะแนนดิบเบ้เมื่อแปลงเป็นคะแนนมาตรฐานแล้วก็จะเบ้ด้วย นั่นคือคะแนนมาตรฐานและคะแนนดิบมีการกระจายขึ้นลงตามกัน ตัวอย่างคะแนนมาตรฐานที่แปลงโดยวิธีนี้ ได้แก่ Linear z-score, Linear t-score, CEEB-score และ AGCT-score เป็นต้น

คะแนนมาตรฐาน (Linear z-score)

คะแนนมาตรฐานนี้เป็นที่รู้จักกันดี และเป็นรากฐานของคะแนนมาตรฐานแบบอื่น ๆ หลักการแปลงคะแนนมาตรฐานชนิดนี้อาศัยคะแนนเฉลี่ยและความเบี่ยงเบนมาตรฐานดังได้กล่าวมาแล้ว การแปลงคะแนนดิบให้เป็นคะแนนมาตรฐาน สามารถคำนวณได้จากสูตรต่อไปนี้

$$z = \frac{X - \bar{X}}{S.D.}$$

เมื่อ X แทน คะแนนดิบ

$\bar{X}$ แทน คะแนนเฉลี่ย

S.D. แทน ความเบี่ยงเบนมาตรฐาน

ตัวอย่าง ในการสอบวิชาวัดผลการศึกษาภาคฤดูร้อน ปีการศึกษา 2523 พบว่าความเบี่ยงเบนมาตรฐานของคะแนนการสอบเท่ากับ 10 คะแนนเฉลี่ย เป็น 65 จงหาว่านักศึกษาที่สอบวัดผลการศึกษาได้ 70 คะแนน จะได้คะแนนมาตรฐานเท่าไร

วิธีทำ แทนค่าจากสูตรจะได้ $z = \frac{70 - 65}{10}$

$= \frac{5}{10}$

$= 0.5$

คะแนนมาตรฐาน z นี้ จะมีค่าเฉลี่ยเท่ากับ 0 และความเบี่ยงเบนมาตรฐานเท่ากับ 1

คะแนนมาตรฐาน t (Linear t-score)

คะแนนมาตรฐาน t เป็นคะแนนมาตรฐานชนิดหนึ่งที่ดัดแปลงมาจากคะแนนมาตรฐาน z ทั้งนี้เนื่องจากคะแนนมาตรฐาน z มีทั้งค่าบวกและค่าลบ กล่าวคือถ้าคะแนนสูงกว่าคะแนนเฉลี่ยก็จะได้ค่าบวก และถ้าคะแนนต่ำกว่าคะแนนเฉลี่ยก็จะได้ค่าลบและนอกจากนี้ค่าที่คำนวณได้ยังเป็นเลขทศนิยมไม่สะดวกต่อการนำมาใช้ ทำให้ยุ่งยากต่อการคิดและการตีความหมาย จึงได้มีผู้คิดเปลี่ยนคะแนนมาตรฐาน z ให้เป็นคะแนนมาตรฐาน t โดยหาตัวคงที่มาบวกเข้าไปเพื่อไม่ให้ค่าติดลบ ซึ่งได้สูตรในการคำนวณดังนี้

$t = 10z + 50$

คะแนนมาตรฐาน t นี้มีคุณสมบัติเหมือนคะแนน z กล่าวคือคะแนน t ที่แปลงโดยวิธีนี้จะยังคงมีทรวดทรงเหมือนคะแนนดิบทุกประการ คะแนนมาตรฐาน t จะมีค่าเฉลี่ยเท่ากับ 50 และความเบี่ยงเบนมาตรฐานเท่ากับ 10

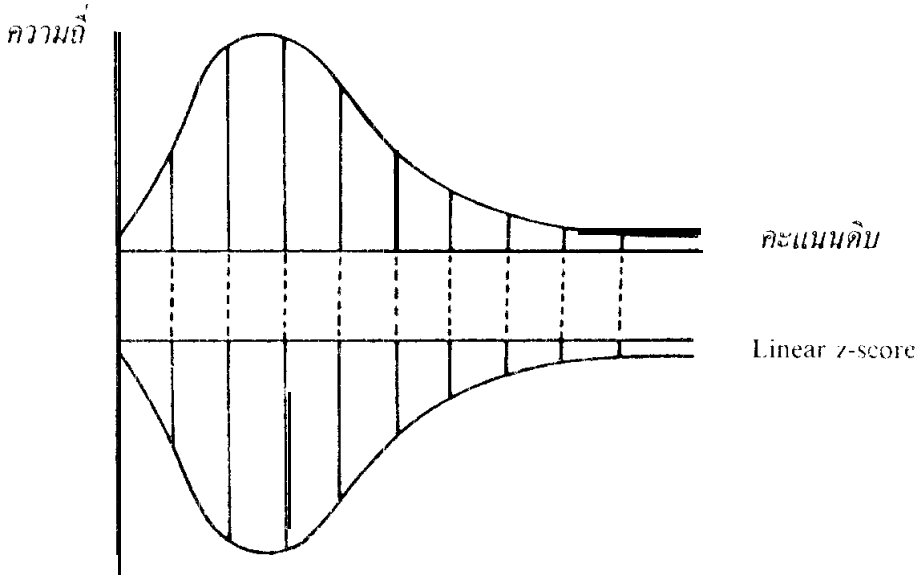
ตัวอย่าง 0.5 z เปลี่ยนเป็นคะแนนมาตรฐาน t ได้ดังนี้

$t = 10(0.5) + 50$

$= 55$

และในกรณีที่จะคิดคะแนนมาตรฐาน t จากคะแนนดิบ ก็สามารถทำได้ดังนี้

$t = \frac{10 (X - \bar{X})}{S.D.} + 50$



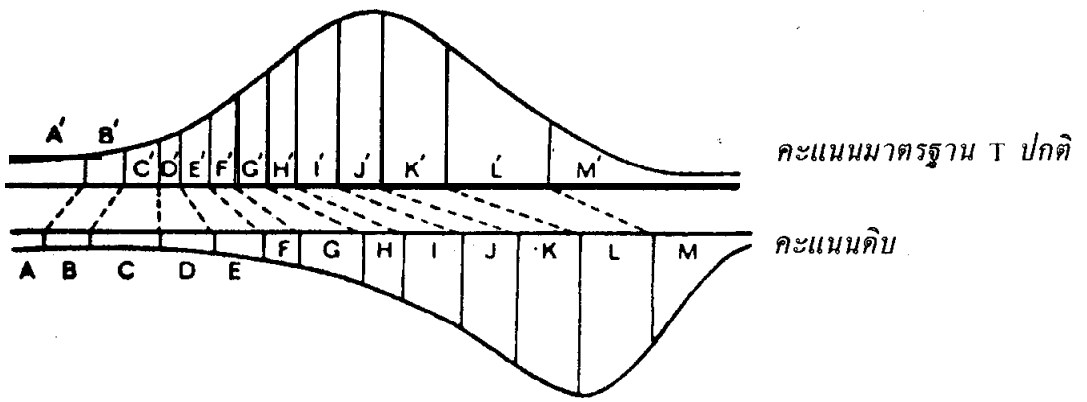
ภาพ 11-1 แสดงการแปลงคะแนนแบบเส้นตรง Linear z-score

นอกจากนี้ยังสามารถแปลงเป็นคะแนนมาตรฐานรูปอื่น ๆ ได้อีก เช่น

CEEB	=	100z + 500
AGCT	=	20z + 100
Wechsler scale	=	3z + 10
Deviation IQ	=	15z + 100

2. การแปลงรูปคะแนนโดยอาศัยพื้นที่ใต้โค้ง (Area transformation) เป็นการแปลงคะแนนดิบให้อยู่ในมาตราใหม่ โดยอาศัยพื้นที่ใต้โค้งเป็นเกณฑ์ในการแปลงรูปคะแนน โดยทั่วไปคะแนนจากผลการสอบแต่ละครั้งมีการแจกแจงเป็นรูปทรงต่าง ๆ กันออกไปหลายชนิด ไม่จำเป็นจะต้องแจกแจงเป็นโค้งปกติ เมื่อการแจกแจงของคะแนนมีรูปทรงแตกต่างกัน ถึงแม้จะทำคะแนนดิบให้เป็นคะแนนมาตรฐานแบบเส้นตรง ก็ยังไม่สามารถนำมาเปรียบเทียบกันได้ จึงต้องหาวิธีการเปลี่ยนโค้งรูปทรงต่าง ๆ ให้อยู่ในรูปโค้งปกติ (Normal curve) โดยเอาตำแหน่ง Percentile ของโค้งนั้น ๆ ไปเทียบกับ Percentile ของโค้งปกติ ก็จะได้ตำแหน่งของคะแนนมาตรฐานที่อยู่ในรูปทรงของการแจกแจงโค้งปกติ ซึ่งเรียกว่าการแปลงคะแนนโดยอาศัยพื้นที่ใต้โค้งเป็นหลัก ลักษณะของการแจกแจงของคะแนนมาตรฐานชนิดนี้จะเป็นโค้งปกติเสมอ ไม่ว่าการแจกแจงของคะแนนดิบเดิมจะมีลักษณะอย่างไร ซึ่งต่างกับการเปลี่ยนคะแนนดิบให้เป็นคะแนนมาตรฐาน

t โดยแปลงต่อจากคะแนนมาตรฐาน z นั้น การแจกแจงยังคงเหมือนคะแนนดิบทุกประการ ซึ่งการแจกแจงนี้อาจจะเป็นโค้งปกติหรือไม่ก็ได้ ถ้าหากการแจกแจงไม่เป็นโค้งปกติ จะนำคะแนน t นั้นมาเปรียบเทียบกันย่อมไม่ได้ แต่ถ้าเป็นการแปลงรูปโดยอาศัยพื้นที่ใต้โค้ง เช่น คะแนนมาตรฐาน T ปกติ (Normalized T-score) แล้วสามารถนำมาเปรียบเทียบกันได้ การแปลความหมายของคะแนนก็แปลในรูปของพื้นที่ใต้โค้งปกติ เช่น แดงสอบได้คะแนนสูงกว่าเพื่อน 84 คน ในจำนวนนักเรียน 100 คน



ภาพ 11-2 แสดงการแปลงคะแนนโดยยึดพื้นที่ใต้โค้งเป็นหลัก

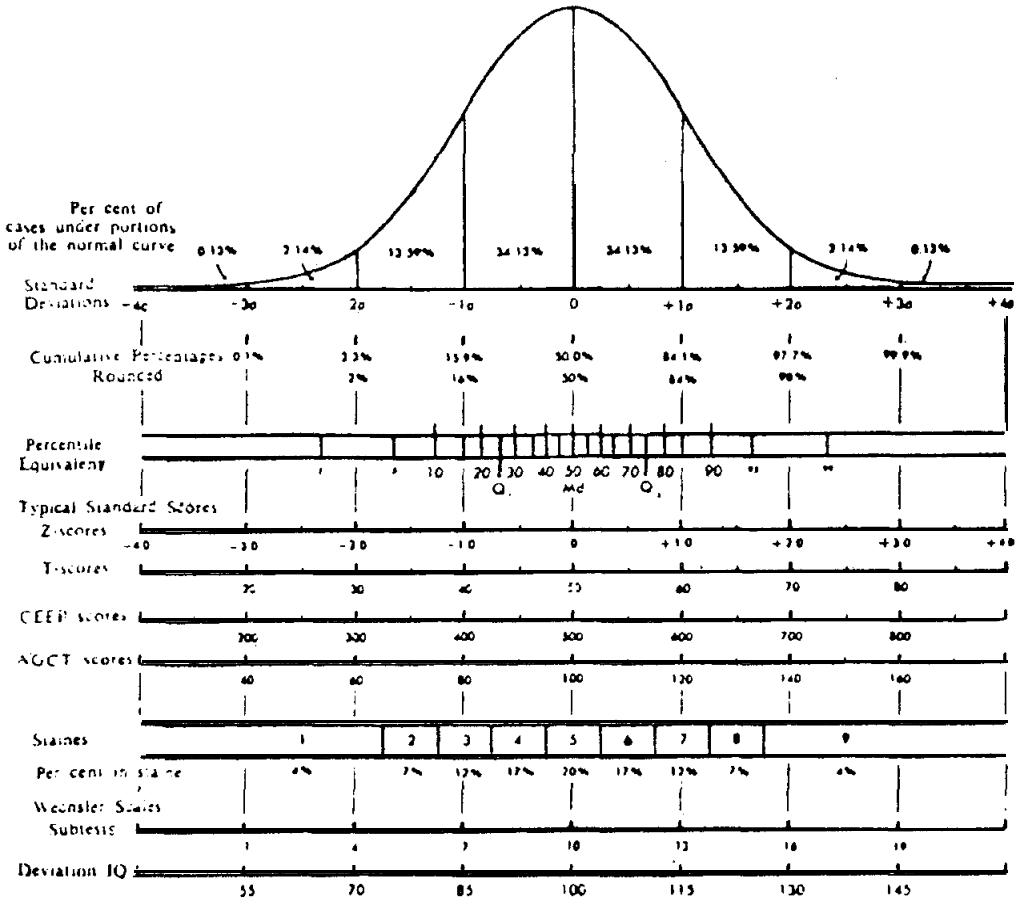
คะแนนมาตรฐาน T ปกติ (Normalized T-score)

เป็นคะแนนมาตรฐานที่นิยมใช้กันมาก ซึ่งมีวิธีการแปลงคะแนนดิบให้เป็นคะแนนมาตรฐาน T ปกติ เป็นลำดับขั้นดังต่อไปนี้

1. เรียงคะแนนดิบจากมากไปหาน้อย
2. หาความถี่ (f) ของคะแนนแต่ละคะแนน
3. หาความถี่สะสม (cf) โดยรวมความถี่ของคะแนนจากชั้นที่มีคะแนนต่ำสุดขึ้นมาหาชั้นที่มีคะแนนสูงสุด จำนวนสุดท้ายจะเท่ากับจำนวนนักเรียนที่เข้าสอบทั้งหมด
4. หา $(cf + \frac{1}{2} f)$
5. หาค่าเปอร์เซ็นต์ไทล์ (Percentile) โดยการเอา $\frac{100}{N}$ ไปคูณกับ $(cf + \frac{1}{2} f)$
6. นำผลจากข้อ 5 ไปหาค่า T ปกติ จากตารางในภาคผนวก

ตาราง 11-1 แสดงการแปลงคะแนนดิบให้เป็นคะแนน T ปกติ

X	f	cc	$cf + \frac{1}{2}f$	% ile	T
30	1	45	44.50	98.8888	73
29	0	43	44.00	97.7777	70
28	0	44	44.00	97.7777	70
27	1	44	43.50	96.6666	68
26	2	43	42.00	93.3333	65
25	1	41	40.50	69.9999	63
24	2	40	39.00	86.6666	61
23	2	38	37.00	82.2222	59
22	3	36	34.50	76.6666	57
21	2	33	32.00	71.1111	56
20	2	31	30.00	66.6466	54
19	3	29	27.50	61.1111	53
18	4	26	24.00	53.3333	51
17	4	22	20.00	44.4444	49
16	3	18	16.50	36.6666	47
15	3	15	13.50	29.9999	45
14	2	12	11.00	24.4444	43
13	3	10	8.50	18.8888	41
12	2	7	6.00	13.3333	39
11	1	5	4.50	9.9999	37
10	2	4	3.00	6.6666	35
9	0	2	2.00	4.4444	33
8	1	2	1.50	3.3333	32
7	0	1	1.00	2.1222	30
6	1	1	.50	1.1111	27



ภาพ 11-3 แสดงการแจกแจงของคะแนนภายใต้โค้งปกติ

คะแนน T ปกติ มีความหมายดีกว่าคะแนนชนิดอื่น ดังตัวอย่างความหมายของคะแนน T ปกติ ดังนี้

คะแนน T 30 หมายถึง ชนะเพื่อนอยู่ 2 คน ใน 100 คน

คะแนน T 40 หมายถึง ชนะเพื่อนอยู่ 16 คน ใน 100 คน

คะแนน T 50 หมายถึง ชนะเพื่อนอยู่ 50 คน ใน 100 คน

คะแนน T 60 หมายถึง ชนะเพื่อนอยู่ 84 คน ใน 100 คน

คะแนน T 70 หมายถึง ชนะเพื่อนอยู่ 98 คน ใน 100 คน

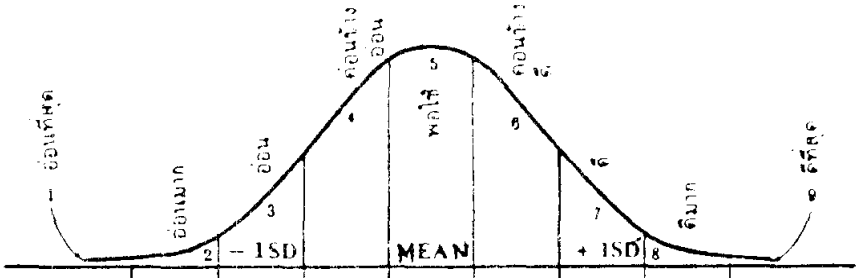
คะแนนมาตรฐาน 9 อันดับ (Stanine score)

คะแนนมาตรฐานชนิดนี้เกิดขึ้นเมื่อสมัยสงครามโลกครั้งที่ 2 โดยนักจิตวิทยากลุ่มหนึ่ง

คำว่า Stanine ย่อมาจาก Standard score of nine units คะแนนมาตรฐานชนิดนี้จะแบ่งสัดส่วนการกระจายของพื้นที่ใต้โค้งปกติออกเป็น 9 ช่วง เป็นคะแนนมาตรฐานที่มีค่าเฉลี่ยเท่ากับ 5 และความเบี่ยงเบนมาตรฐานเท่ากับ 2 ช่วงกว้างในแต่ละ Stanine จาก Stanine ที่ 2 ถึง Stanine ที่ 8 จะมีความกว้างเท่ากับ $\frac{1}{2}$ ของความเบี่ยงเบนมาตรฐาน เขียนเป็นสูตรแสดงความสัมพันธ์ระหว่าง z-score และ Stanine ได้ดังนี้

$$\text{Stanine} = 2z + 5$$

ในแต่ละ Stanine จะครอบคลุมพื้นที่ใต้โค้งปกติดังนี้คือ Stanine ที่ 1 และ 9 เท่ากับ 4.01% Stanine ที่ 2 และ 8 เท่ากับ 6.55% Stanine ที่ 3 และ 7 เท่ากับ 12.10% Stanine ที่ 4 และ 6 เท่ากับ 17.47% Stanine ที่ 5 เท่ากับ 19.74% ดังแสดงในรูป



เปอร์เซ็นต์ที่คาด Stanine	4	7	12	17	20	17	12	7	4
เปอร์เซ็นต์สะสมจากต่ำไปหาสูง	4	11	23	40	60	77	89	96	100
ช่วง Percentile	4 ลงมา	5–11	12–23	24–40	41–60	61–77	78–89	90–96	97 ขึ้นไป
ช่วง SD	−1.75 ลงไป	−1.75 ถึง −1.25	−1.25 ถึง −.75	−.75 ถึง −.5	−.5 ถึง +.5	+.5 ถึง +.75	+.75 ถึง +1.25	+1.25 ถึง +1.75	+1.75 ขึ้นไป
ระดับคะแนน	0	1	1	2	2	3	3	4	4
ระดับคะแนน	0	1		2			3		4

ภาพ 11-4 แสดงการแบ่งสัดส่วนของพื้นที่ใต้โค้งปกติออกเป็น 9 ช่วง

การแปลงคะแนนดิบให้เป็น Stanine ทำได้ดังนี้

1. เรียงคะแนนจากน้อยไปหามาก

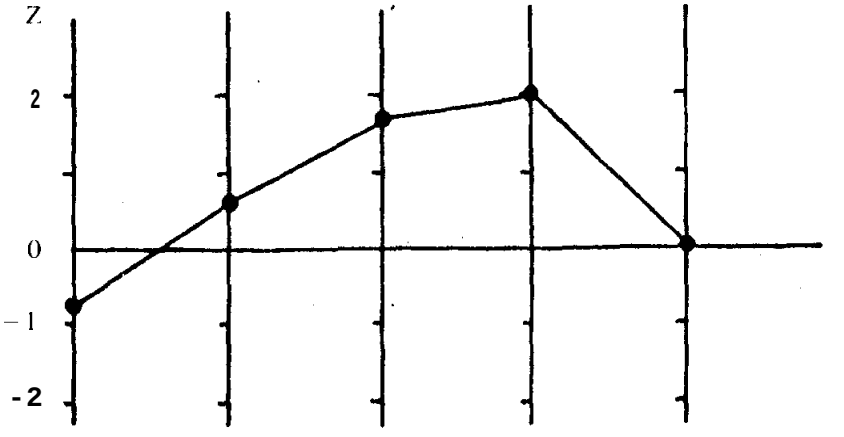
2. คำนวณหาจำนวนคนที่จะตกอยู่ในแต่ละ Stanine โดยคิดจากจำนวนผู้เข้าสอบ เช่น มีผู้เข้าสอบ 400 คน จำนวนคนที่จะอยู่ใน Stanine ที่ 1 และ 9 เท่ากับ $\frac{4}{100} \times 400 = 16$ คน

3. ตัดเป็นช่วง Stanine ตามจำนวนคนที่คำนวณได้ในข้อ 2

11.5 ประโยชน์ของคะแนนมาตรฐาน

1. ใช้เปรียบเทียบความสามารถของนักเรียนในกลุ่ม คะแนนมาตรฐานจะช่วยให้ทราบว่านักเรียนแต่ละคนมีความสามารถขนาดใด เก่งหรืออ่อนกว่าผู้อื่นเป็นสัดส่วนมากน้อยเพียงใดจากจำนวนคนที่เข้าสอบทั้งหมด

2. ใช้เปรียบเทียบความสามารถในด้านต่าง ๆ ของนักเรียน โดยใช้เส้นภาพ (Profile) ทั้งนี้จะช่วยให้เห็นสภาพความสามารถที่แท้จริงในแต่ละด้านของนักเรียนดีขึ้น อันจะเป็นประโยชน์ต่อการแนะแนว เช่น



ภาพ 11-5 เส้นภาพแสดงการเปรียบเทียบความสามารถด้านต่าง ๆ ของนักเรียน

3. ทำให้คะแนนสามารถรวมกันได้ เพราะมีหน่วยเท่ากัน

11.6 การตัดเกรดหรือการให้ระดับคะแนน (Grading)

การวัดผลการศึกษาเป็นเรื่องยุ่งยากซับซ้อนดังได้กล่าวมาแล้ว ปัญหาที่ยุ่งยากของการวัดผลเริ่มตั้งแต่การสร้างข้อสอบ การแปลคะแนนผลการสอบ และการนำผลการสอบไปใช้ให้เกิดประโยชน์มากที่สุด ปัญหาการให้ระดับคะแนนอย่างยุติธรรม นับได้ว่าเป็นปัญหาที่มีความยุ่งยากและซับซ้อนเช่นกัน ทั้งนี้เพราะการให้เกรดจะต้องอาศัยวิจารณญาณและคุณธรรมประกอบการพิจารณา โดยพยายามให้ความเป็นธรรมและขจัดความลำเอียงหรือความคิดส่วนตัวออกไป ซึ่งครูจำเป็นต้องพิจารณาอย่างรอบคอบถี่ถ้วน เพราะการให้เกรดมีผลกระทบโดยตรงต่ออนาคตของนักเรียน ดังนั้นจึงเป็นสิ่งที่ครูทุกคนควรคำนึงถึงไว้ให้มาก

การตัดเกรด เป็นวิธีการสรุปผลการเรียนขั้นสุดท้าย เพื่อประเมินผลและกำหนดระดับความสามารถในการเรียนของนักเรียนว่า ผ่าน-ไม่ผ่าน เก่ง อ่อน ระดับใด (A B C D E หรือ F) การตัดเกรดจึงเป็นการนำผลการสอบวัดทุกชนิดมาเป็นหลักในการประเมินผล ในการตัดเกรดเพื่อให้เกิดความถูกต้องและเหมาะสมนั้นจะต้องคำนึงถึงองค์ประกอบ 3 ประการต่อไปนี้

1. **ผลของการวัด** (measurement) การตัดเกรดที่ดีจะต้องอาศัยผลการวัดที่ถูกต้อง แม่นยำ มีความเที่ยงตรง และเชื่อมั่นได้ ถ้าผลการวัดเชื่อมั่นไม่ได้หรือขาดความเที่ยงตรง เมื่อนำผลที่ได้จากการวัดนั้นไปตัดเกรดก็ย่อมทำให้เกิดความคลาดเคลื่อนได้

2. **เกณฑ์การพิจารณา** (criteria) เป็นมาตรฐานที่ใช้เป็นหลักในการเปรียบเทียบ ถือว่าเป็นเครื่องตัดสินชี้ขาดระดับความสามารถของนักเรียน

3. **วิจารณญาณและคุณธรรม** (value judgement) ผลที่ได้จากการวัดเป็นเพียงข้อมูลส่วนหนึ่งที่เกี่ยวกับตัวนักเรียนเท่านั้น การประเมินผลที่เที่ยงตรงจำต้องอาศัยดุลยพินิจ หรือการพิจารณาอย่างรอบคอบถี่ถ้วนของผู้สอนประกอบ โดยพยายามให้ความเป็นธรรม ขจัดความลำเอียง หรือคติส่วนตัว และควรคำนึงถึงความงอกงามของนักเรียนในด้านต่าง ๆ เช่น ความสนใจ ความตั้งใจในการเรียน ประกอบการตัดสินให้ระดับคะแนนด้วย ซึ่งจะช่วยให้การให้ระดับคะแนนนั้นมีความเหมาะสมและยุติธรรม

การพยายามตัดเกรดหรือให้ระดับคะแนนโดยใช้โค้งปกติ (Normal curve) และคะแนน T ปกตินั้น เป็นที่นิยมมากในประเทศไทย สิ่งเหล่านี้หากครูนำเอามาใช้ด้วยความรู้ความเข้าใจที่แท้จริงก็ย่อมจะเกิดประโยชน์ แต่หากนำมาใช้ด้วยความไม่เข้าใจอาจเกิดผลเสียหายขึ้นได้ ดังตัวอย่างเช่น การตัดเกรดโดยใช้คะแนน T และใช้มาตรฐานเดียวกันหมด คือถ้าได้คะแนน T ต่ำกว่า 30 ถือว่าสอบตก หากมองเรื่องนี้อย่างผิวเผิน จะเห็นว่ามีความเป็นธรรมดี แต่ถ้าพิจารณาให้รอบคอบจะเห็นว่าในชั้นที่มีนักเรียนตั้งแต่ 21 คนลงมา จะไม่มีนักเรียนคนใดให้คะแนน T ต่ำกว่า 30 เลย ไม่ว่านักเรียนคนที่อ่อนที่สุดจะได้คะแนนเท่าไรก็ตาม นั่นคือหากวิชาใดมีคนเรียนไม่เกิน 21 คน และใช้วิธีตัดเกรดดังกล่าวมาแล้วข้างต้น วิชานั้นจะไม่มีคนสอบตกเลย

การตัดเกรดโดยใช้คะแนน T นั้น มีบางท่านเสนอว่า เมื่อมี A ก็จะต้องมี E (หรือ F) ดร.เสริมศักดิ์ วิศาลาภรณ์ และ เอนกกุล กรีแสง ได้ให้ข้อคิดเกี่ยวกับเรื่องนี้ โดยยกตัวอย่างประกอบ 2 เรื่อง ดังนี้

เรื่องที่ 1 ในห้องเรียนวิชาฝรั่งเศส มีนักเรียน 7 คนด้วยกัน ชั่วโมงแรกที่อาจารย์เข้าสอน อาจารย์ก็บอกว่าการตัดเกรดในเทอมนี้ก็เหมือนเทอมก่อน คือใช้คะแนน T และโค้งปกติ เมื่อ

เทอมที่แล้วมีคนได้ A และก็มีคนได้ E ด้วย แล้วอาจารย์ผู้นั้นก็บรรยายไปจนหมดชั่วโมง หลังจากอาจารย์ออกจากห้องไปแล้ว นิสิตคนหนึ่งคิดขึ้นมาได้ว่าการตัดเกรดโดยคะแนน T นั้น เมื่อมี A ก็ต้องมี E จึงพูดกับเพื่อน ๆ ว่า "ผมเชื่อเหลือเกินว่าในเทอมนี้จะต้องมีพวกเราคนหนึ่งคนใดตกแน่ ๆ อย่ากระนั้นเลย ผมจะช่วยพวกเราโดยผมจะไปขอถอนวิชาเรียนที่แผนกทะเบียนเดี๋ยวนี้" นิสิตอีกคนหนึ่งทำท่าจะรู้จักคะแนน T ดีกว่าคนอื่นค้านว่า "คุณทำอย่างนั้นไม่ได้ เพราะเรายังไม่รู้เลยว่าใครจะเป็นผู้สอบตก"

เรื่องที่ 2 หลังจากสงครามโลกครั้งที่ 2 ยุติลง ทหารอเมริกันทั้งหลายก็กลับมาบ้าน ทหารผ่านศึกเหล่านี้ได้รับสิทธิพิเศษให้เรียนต่อในมหาวิทยาลัย ทหารผ่านศึกเหล่านี้พยายามให้ภรรยาของเขาเข้าเรียนด้วย วิชาใดที่ยาก ๆ ทหารผ่านศึกเหล่านี้จะให้ภรรยาลงทะเบียนวิชาร่วมกับพวกตน บรรดาภรรยาทหารผ่านศึกไม่ได้ตั้งใจเรียนเพื่อจะเอาความรู้ เวลาสอบก็ขีด ๆ เตา ๆ ไป เมื่อตัดเกรดโดยใช้คะแนน T และโค้งปกติ ภรรยาของทหารผ่านศึกเหล่านี้จะได้เกรด D กับ E ส่วนทหารผ่านศึกก็จะได้เกรด C^- ไปอย่างสบาย ๆ ทั้ง ๆ ที่ไม่มีความรู้อะไรมากนัก ทั้งนี้เพราะมีผู้ได้คะแนน T ต่ำกว่าบรรดาทหารผ่านศึกอยู่มากพอสมควร

จากตัวอย่าง 2 เรื่องที่กล่าวมานี้ ช่วยชี้ให้เห็นว่าการตัดเกรดโดยใช้คะแนน T และโค้งปกติ หากขาดความเข้าใจที่แท้จริงอาจเกิดผลเสียหายขึ้นได้

จะเห็นได้ว่าการให้เกรดนั้นมีปัญหามากมาย ซึ่งอาจจะมากกว่าปัญหาในการเขียนข้อสอบที่ดีด้วยซ้ำไป วัญญา วิศาลาภรณ์ ได้สรุปปัญหาที่สำคัญ ๆ ของการให้เกรดไว้ดังนี้

1. การให้เกรดเป็นเรื่องยุ่งยาก และซับซ้อน เพราะการให้เกรดเป็นเรื่องของ การประมาณ เราไม่สามารถวัดความรู้ความสามารถที่แท้จริงของนักเรียนได้ สาเหตุที่เราไม่สามารถวัดสมรรถภาพที่แท้จริงของมนุษย์ได้ อาจเป็นเพราะ

1.1 ไม่มีความแน่นอน ในหน่วยที่เราจะวัด

1.2 ข้อเท็จจริงที่เราจะวัดนั้นไม่คงที่ เปลี่ยนแปลงอยู่เสมอ

1.3 สิ่งที่จะวัดมีความซับซ้อนมาก

ด้วยเหตุนี้สิ่งที่เราวัดจึงเป็นการประมาณ และการประเมินผลก็ตั้งอยู่บนรากฐานของการประมาณ

2. โดยธรรมชาติครูเป็นผู้คอยช่วยเหลือ แนะแนว แนะนำ และสั่งสอนแต่ในการให้เกรดนั้น ครูต้องทำหน้าที่ตัดสิน เป็นเสมือนผู้พิพากษา จึงอาจขัดกับความรู้สึกในวิญญาณของความเป็นครู เป็นการลดมิตรภาพระหว่างศิษย์กับครูให้ลดลง ทำให้ครูบางคน ไม่กล้าที่จะให้

เกรดต่ำ ๆ จึงทำให้ความเชื่อถือได้ของเกรดลดลง

3. ความเข้าใจในการให้เกรดยังไม่เป็นสากล ทั้งนี้เพราะนิยามของคำว่าเกรดยังไม่กระจ่างชัด ซึ่งเป็นผลทำให้มาตรฐานของการให้เกรด และความหมายของการให้เกรดเปลี่ยนไป การให้เกรดของแต่ละโรงเรียน มีมาตรฐานไม่เหมือนกัน เกรดเดียวกันแต่จากคนละโรงเรียน ย่อมมีความหมายไม่เหมือนกัน เมื่อเราใช้เกรดเป็นสัญลักษณ์เพื่อการสื่อความหมายและชณะเดียวกัน สัญลักษณ์ที่เราใช้สื่อความหมาย ก็นิยามไม่ชัดเจนเป็นผลให้การสื่อความหมาย ผิดพลาดได้ง่าย

4. หลักฐานที่จะใช้เป็นพื้นฐานในการให้เกรด ยังมีไม่พอเพียง เช่นการสอบเพียงสองครั้ง หรือการทำรายงานเพียง 3 ฉบับ สิ่งเหล่านี้ยังไม่เพียงพอในการตัดสินให้เกรดจึงเป็นผลทำให้เกรดขาดความเชื่อถือได้

5. เป็นความเชื่อของครูบางคนที่ว่าเมื่อสอนแล้ว ผู้เรียนจะต้องได้ความรู้บ้าง ครูประเภทนี้จึงไม่เคยให้ใครต่ำกว่า C

6. ลักษณะและธรรมชาติของผู้เรียนมีผลต่อการให้เกรดของครู Ebel ได้อ้างจากงานวิจัยพบว่าในขณะที่ความสามารถเท่ากันนักเรียนหญิงมีแนวโน้มที่จะได้เกรดสูงกว่านักเรียนชาย

7. ยังมีอีกหลายอย่างที่ยังหาข้อยุติเกี่ยวกับเกรดไม่ได้ เช่น

7.1 นักเรียนที่ควรจะได้เกรดนั้น ๆ ควรจะมีคุณลักษณะอย่างไร

7.2 ควรจะใช้หลักฐานอะไรบ้างเพื่อเป็นพื้นฐานในการให้เกรด

11.7 การตัดเกรดของมหาวิทยาลัยในประเทศไทย

การตัดเกรดที่นิยมและแพร่หลายกันมากในเมืองไทย คือการให้ระดับคะแนนที่แบ่งเป็น 5 กลุ่ม หรือ 5 เกรด โดยแบ่งตามสัดส่วนของพื้นที่ใต้โค้งปกติ การให้เกรดระบบนี้อยู่บนความเชื่อพื้นฐานว่าคะแนนความสามารถในการเรียนของนักเรียนมีการกระจายเป็นโค้งปกติ กล่าวคือคนที่เก่งและอ่อนจะมีน้อยคนส่วนใหญ่จะมีความสามารถปานกลาง ซึ่งถ้าหากเขียนโค้งแทนความสามารถของนักเรียนก็จะได้เป็นโค้งรูประฆังคว่ำหรือโค้งปกติ การตัดเกรดตามแนวนี้ใช้กันในมหาวิทยาลัยเกือบทุกแห่งของประเทศไทย

ส่วนการตัดเกรดของมหาวิทยาลัยรามคำแหงนั้นใช้ระบบ 3 เกรด คือ C, P และ F เอนก เพียรอนุกูลบุตรได้กล่าวว่า แต่เดิมนั้นต้องการให้เพียง 2 เกรด คือผ่านกับไม่ผ่าน ทั้งนี้โดยมีเหตุผลว่า มหาวิทยาลัยรามคำแหงเป็นมหาวิทยาลัยเปิดการวัดผลไม่ครอบคลุมไม่ถี่ถ้วน และยาจมีความคลาดเคลื่อนเจือปนอยู่มาก จึงควรลดความคลาดเคลื่อนให้เหลือน้อย ๆ โดยตัดสินเพียงผ่าน-ไม่ผ่านเท่านั้น และยังสอดคล้องกับการวัดผลที่ยึดหลักการเรียนเพื่อรอบรู้ (Mastery learning)

โดยแจ้งเกณฑ์และจุดประสงค์ที่เฉพาะเจาะจงในการเรียนกระบวนวิชานั้น ๆ ไว้ล่วงหน้ามาก ๆ ข้อ เช่น 400-500 ข้อ แล้วก็ออกข้อสอบวัดตามเกณฑ์นั้น โดยไม่สนใจว่าใครจะเรียนรู้มาโดยวิธีใด เพราะส่วนใหญ่ต้องศึกษาด้วยตนเอง แล้วตัดสินผลการสอบตามเกณฑ์นั้นจึงจะยุติธรรม แต่จะด้วยเหตุผลกลใดก็ตามได้มีเกรด G เพิ่มขึ้นมาอีก 1 เกรด สำหรับคนที่ทำคะแนนได้ P อ้วน ๆ หรือ P แก่ ๆ

11.8 แนวคิดบางประการในการให้เกรด

วัญญา วิศาลาภรณ์ ได้สรุปแนวคิดของนักวัดผลที่สำคัญ ๆ เกี่ยวกับการให้เกรดไว้ดังนี้

1. การให้เกรดควรจะสมบูรณ์ (absolute) หรือสัมพัทธ์ (relative) การให้เกรดที่สมบูรณ์ก็คือการให้เกรดที่เทียบกับมาตรฐานบางอย่างที่กำหนดให้ เช่น จะสอบผ่านได้ต้องสอบได้อย่างน้อย 50% หรือได้เกรด A ควรจะได้ 90% ขึ้นไป การให้เกรดสัมพัทธ์คือ การให้เกรดที่เทียบกับกลุ่มของผู้เรียนเดียวกัน คนได้ A คือคนที่เก่งของกลุ่มนั้น

2. การให้เกรดควรจะตั้งอยู่บนรากฐานของสัมฤทธิ์ผลของการเรียนหรือตั้งอยู่บนรากฐานของทัศนคติและความพยายาม ครูบางคนใช้องค์ประกอบอื่นที่นอกเหนือไปจากสัมฤทธิ์ผลทางการเรียนเป็นเครื่องช่วยในการให้เกรด เช่น ใช้เกรดเป็นเครื่องปกครองชั้นเรียน ใช้ความมานะพยายาม ใช้การไม่ขาดเรียนเป็นพื้นฐานในการให้เกรด โดยทั่ว ๆ ไปเกรดควรเป็นเครื่องชี้บ่งว่านักเรียนบรรลุถึงจุดมุ่งหมายของการสอนมากน้อยเพียงไร ดังนั้นถ้าจุดมุ่งหมายของการสอนเพื่อการพัฒนาทัศนคติและความพยายาม แล้วการให้เกรดโดยอาศัยทัศนคติและความพยายามของผู้เรียนเป็นพื้นฐานในการให้เกรดย่อมเป็นสิ่งที่กระทำได้

3. เกรดควรจะเป็นเครื่องบอกสถานภาพของผลสัมฤทธิ์หรือควรจะเป็นเครื่องบอกระดับความก้าวหน้าของผลสัมฤทธิ์ ครูบางคนพยายามที่จะให้เกรดตั้งอยู่บนรากฐานของการพัฒนาหรือความก้าวหน้ามากกว่าที่จะตั้งอยู่บนรากฐานของระดับผลสัมฤทธิ์ที่นักเรียนบรรลุถึง ครูประเภทนี้จะใช้การสอบครั้งแรกเป็นเครื่องวัดสถานภาพเดิม และความแตกต่างระหว่างการสอบครั้งแรกกับครั้งหลังเป็นเครื่องวัดการพัฒนา เช่น นักเรียนคนหนึ่งสอบครั้งแรกพิมพ์ได้ 20 คำต่อนาที ทดสอบครั้งที่สองพิมพ์ได้ 35 คำต่อนาที ผลของการพัฒนาคือ 15 คำต่อนาที นักเรียนอีกคนหนึ่งทดสอบครั้งแรกพิมพ์ได้ 30 คำต่อนาที ทดสอบครั้งที่สองพิมพ์ได้ 35 คำต่อนาที ผลของการพัฒนาคือ 5 คำต่อนาที นักเรียนสองคนนี้มีผลของการพัฒนาต่างกัน และระดับสถานภาพของผลสัมฤทธิ์เท่ากัน นักเรียนสองคนนี้ควรจะได้เกรดเท่ากันหรือไม่

4. การให้เกรดนั้นแต่ละวิชาควรจะให้เกรดเดียวหรือหลายเกรด ถ้าหากเราคิดว่าการให้เกรดเป็นการประเมินผลโดยรวบยอด การให้เกรดก็น่าจะมีเกรดเดียว แต่ในบางครั้งการให้เกรดเดียวก็อาจทำให้เกิดการเข้าใจผิดในเกรดนั้น ๆ ได้ จากตัวอย่างของนักเรียนพิมพ์ดีดในข้อ 3 ถ้าจะให้เกรดนักเรียนสองคนเท่ากัน กับการให้สองเกรดดังนี้

นักเรียน	สถานภาพของผลสัมฤทธิ์	การพัฒนา
คนแรก	B	B
คนสอง	B	C

จะเห็นได้ว่าการให้เกรดสองเกรด จะช่วยให้สามารถเลือกคนพิมพ์ดีดได้ถูกต้องกว่า

5. เกรดที่ให้ควรจะมีน้อยเกรดหรือมากเกรด ในปัจจุบันมีการให้น้อยเกรดบ้างมากเกรดบ้าง เช่น บางสถาบันให้เพียงสองเกรดเท่านั้นคือ ผ่าน กับ ไม่ผ่าน เกรดจึงมี P กับ F บางสถาบันให้พอใจกับไม่พอใจ เกรดจึงมี S กับ U แต่ในขณะที่อีกหลายสถาบันให้ 5 เกรดคือ A, B, C, D, E หรือ F และค่าของเกรดเหล่านี้ บางแห่งกำหนดค่าระหว่าง 0 ถึง 4 บางแห่งกำหนดค่าระหว่าง 0 ถึง 5 การให้เกรดเพียงสองเกรดนั้นก็มีส่วนช่วยลดความแก่งแย่งชิงดีกันระหว่างผู้เรียน แต่ขณะเดียวกันก็ทำให้ความหมายของเกรดลดน้อยลง นักวัดผลได้สรุปว่า การให้เกรดมากเกรดจะทำให้ความเชื่อมั่นของการให้เกรดสูงขึ้น

6. จะให้เกรดโดยใช้อะไรเป็นฐาน การให้เกรดอาจยึดเอาเปอร์เซ็นต์เป็นฐาน หรืออาจยึดเอาความเบี่ยงเบนมาตรฐาน คะแนน T หรือ stanine สิ่งเหล่านี้ย่อมมีทั้งลักษณะเด่นและลักษณะด้อย

7. ควรจะมีการกำหนดปริมาณของแต่ละเกรดหรือไม่ การให้เกรดนั้น เกรด A มีจำนวนน้อย B มีจำนวนมากขึ้น และ C มีจำนวนมากที่สุดหรือไม่ และปัญหาที่ควรจะพิจารณาต่อไปคือ ควรมี A กี่เปอร์เซ็นต์ บางสถาบันกำหนดไว้เลยว่า ครูจะให้ A ได้ 7 เปอร์เซ็นต์ บางแห่งกำหนดว่าจะให้ A ได้ไม่เกิน 10 เปอร์เซ็นต์ สิ่งเหล่านี้ยังหาข้อยุติไม่ได้ว่าในแต่ละเกรดนั้นควรจะกำหนดปริมาณของเกรดหรือไม่

Noll ได้เสนอหลักในการให้เกรดไว้ดังนี้.-

1. เกรดควรจะเป็นผลจากการเปรียบเทียบกับนักเรียนในกลุ่มด้วย กล่าวคือนักเรียนที่เก่งที่สุดควรจะได้เกรดสูงสุด

2. การให้เกรดเป็นตัวอักษรนั้นดีกว่าการบอกผลเป็นเปอร์เซ็นต์

3. การให้เกรดควรพิจารณาว่าผู้เรียนบรรลุถึงจุดมุ่งหมายที่ตั้งไว้มากน้อยเพียงใด

4. เกรดควรจะเป็นตัวแทนของผลสัมฤทธิ์มากกว่าเป็นตัวแทนของสิ่งอื่น กล่าวคือ นักเรียนที่ได้เกรด B ในวิชาคณิตศาสตร์ ควรจะหมายถึงเขาบรรลุถึงผลสัมฤทธิ์วิชาคณิตศาสตร์ในระดับ B มิใช่ว่าเขามีความพยายามในการเรียนคณิตศาสตร์ระดับ B หรือถ้ามีนักเรียนสองคนเรียนคณิตศาสตร์ นักเรียนคนแรกมีผลสัมฤทธิ์ทางการเรียนในระดับ A แต่มีความพยายามในระดับ C ส่วนนักเรียนอีกคนหนึ่งมีผลสัมฤทธิ์ทางการเรียนในระดับ C แต่มีความพยายามในระดับ A การที่ครูให้เกรดนักเรียนสองคนนี้เท่ากันคือให้ B คงจะไม่ถูกต้องนัก ทั้งนี้เพราะเกรดควรจะเป็นเครื่องชี้บ่งระดับผลสัมฤทธิ์ในเนื้อหาวิชา (subject) มากกว่า

5. ในการตัดเกรดเพื่อให้มีความยุติธรรมนั้น ควรจะได้มาจากการวัดหลาย ๆ ครั้ง และหลาย ๆ ทาง

11.9 วิธีการตัดเกรด

นักเรียนทุกคนย่อมต้องผ่านการสอบมามากมายหลายครั้ง คะแนนที่ได้จากการสอบแต่ละครั้งเป็นเพียงค่าประมาณเท่านั้น มิใช่ค่าจริงของความรู้ความสามารถของนักเรียน เมื่อได้คะแนนมาแล้วครูก็จะนำเอาคะแนนเหล่านี้มาบวกลบคูณหาร คิดเป็นร้อยละ คิดเป็นคะแนน T ในกระบวนการดังกล่าวนี้ความผิดพลาดย่อมเกิดขึ้นได้ จากนั้นจึงตัดสินว่าได้หรือตก หรือได้เกรดอะไร การให้ระดับคะแนนถือว่าเป็นการประเมินผล ผู้ประเมินจะต้องนำข้อมูลต่าง ๆ ที่ได้จากการวัดผลมาพิจารณาด้วยความระมัดระวัง ใช้ดุลพินิจและคุณธรรมในการตัดสินใจให้ระดับคะแนน การให้ระดับคะแนนไม่มีกฎเกณฑ์ตายตัวแน่นอน บางครั้งอาจมี 5 เกรด 4 เกรด หรือ 3 เกรดก็ได้ หรือแม้แต่เกรดเดียวก็สามารถทำได้ การทำคะแนนดิบให้เป็นคะแนนมาตรฐานก็บอกไม่ได้ว่าคะแนนมาตรฐานเท่าใดควรจะได้ระดับคะแนนใด และต้องไม่ยึดหลักว่ามี A แล้วจะต้องมี E หรือ F ทุกครั้งไป ครูผู้สอนจะเป็นผู้ที่ประเมินให้ระดับคะแนนได้ดีที่สุด เพราะเป็นผู้ใกล้ชิดนักเรียน รู้จักและเข้าใจปัญหาของนักเรียน หลังจากให้ระดับคะแนนเรียบร้อยแล้ว ครูควรพิจารณาระดับคะแนนที่ให้คุณให้โทษแก่นักเรียน เช่น ระดับคะแนน A(4) และระดับคะแนน E(0) อีกครั้งหนึ่งว่าเหมาะสมหรือยัง ถ้ายังไม่เหมาะสมอาจพิจารณาให้ระดับคะแนนใหม่ตามความเหมาะสมได้

Robert L. Ebel (1965) ได้ให้แนวคิดในการตัดเกรด โดยแบ่งความสามารถของนักเรียนออกเป็น 7 ระดับ โดยอาศัยค่าเฉลี่ยของคะแนนทั้งชั้นในภาคเรียนที่แล้ว หรือในวิชาที่ผ่านมาเป็น

เกณฑ์ในการเปรียบเทียบ ซึ่ง Ebel ได้สรุปเป็นตารางจำนวนร้อยละของผู้ที่ได้ระดับคะแนนต่าง ๆ ไว้ในหนังสือ Measuring Educational Achievement ดังนี้

ตารางที่ 11-2 กำหนดร้อยละของผู้เข้าสอบที่ได้ระดับคะแนนต่าง ๆ จำแนกตามระดับความสามารถของกลุ่ม

ระดับความสามารถของนักเรียนทั้งชั้น	ระดับเฉลี่ยของกลุ่ม	จำนวนเปอร์เซ็นต์ของแต่ละระดับคะแนน					ขีดจำกัดล่างของ A	% ile
		A(4)	B(3)	C(2)	D(1)	E(0)		
ดีเลิศ	2.80	24	38	29	8	1	0.7	79
ดีมาก	2.60	18	36	32	12	2	0.9	73
ดี	2.40	14	32	36	15	3	1.1	66
ดีพอใช้	2.20	10	29	37	20	4	1.3	58
ปานกลาง	2.00	7	24	38	26	7	1.5	50
อ่อน	1.80	4	20	37	29	10	1.7	42
อ่อนมาก	1.60	3	15	36	32	14	1.9	34

การตัดเกรดโดยอาศัยตารางของ Ebel นี้ ต้องดู GPA (Grade point average) ของนักเรียนก่อนว่าเป็นเท่าไร เช่น นักเรียนชั้นหนึ่งมีค่าเฉลี่ยของ GPA เป็น 2.50 ก็สรุปว่าระดับความสามารถของนักเรียนชั้นนี้อยู่ในขั้นดี จากนั้นก็คำนวณหามัธยฐาน ความเบี่ยงเบนมาตรฐานของผลการสอบวิชานั้น ทั้งนี้เพื่อหาขีดจำกัดล่างของเกรด A ในการหาขีดจำกัดล่างให้เอามัธยฐานบวกกับผลคูณของความเบี่ยงเบนมาตรฐาน กับขีดจำกัดล่างของ A (จากตาราง) ช่วงของคะแนนในแต่ละเกรดจะมีค่าเท่ากับความเบี่ยงเบนมาตรฐาน

ตัวอย่าง ผลการสอบคณิตศาสตร์ของนักเรียนในโรงเรียนแห่งหนึ่ง ปรากฏว่าระดับความสามารถในวิชาคณิตศาสตร์ของนักเรียนในโรงเรียนนี้อยู่ในกลุ่มดี จากคะแนนผลการสอบหามัธยฐานได้เท่ากับ 95 ความเบี่ยงเบนมาตรฐานเท่ากับ 15 การให้เกรดจะเป็นดังนี้

ตาราง 11-3 แสดงการตัดเกรดโดยอาศัยตารางของ Ebel

เกรด	ขีดจำกัดล่างของเกรด	ช่วงคะแนนของเกรด
A	95 + (15 x 1.1) = 111.5	คะแนนตั้งแต่ 112 ขึ้นไป
B	111.5 • 15 = 96.5	คะแนนระหว่าง 97-111
c	96.5 15 = 81.5	คะแนนระหว่าง 82-96
D	81.5 • 15 = 66.5	คะแนนระหว่าง 67-81
E	66.5 15 = 51.5	คะแนนต่ำกว่า 66 ลงมา

การให้ระดับคะแนนเท่าที่ผ่านมาใช้วิธีการต่าง ๆ กัน และยังไม่สามารถสรุปได้ว่าวิธีใดดีที่สุด ในที่นี้ผู้เขียนขอเสนอแนะวิธีการให้ระดับคะแนนที่นิยมใช้กันอยู่ทั่ว ๆ ไป 4 วิธี คือ

1. **การตัดเกรดโดยการจัดกลุ่มตามธรรมชาติ** วิธีนี้จะนำคะแนนของนักเรียนมาเรียงกันจากมากไปหาน้อย หรือจากน้อยไปหามากก็ได้ แล้วพิจารณาว่ามีช่องว่าง (gap) ตรงไหน ให้ถือว่าตรงนั้นเป็นจุดแบ่งระดับคะแนน จุดอ่อนของวิธีนี้คือ ถ้ามีนักเรียนเข้าสอบจำนวนมาก ๆ โอกาสที่จะมีช่องว่างเป็นไปได้ยาก

2. **การตัดเกรดโดยใช้เกณฑ์ที่คาดหวัง** การตัดเกรดวิธีนี้จะใช้คะแนนดิบหรือเปอร์เซ็นต์ที่นักเรียนสอบได้เป็นหลักในการตัดเกรด โดยผู้สอนกำหนดไว้ก่อนว่าแต่ละระดับต้องอยู่ในช่วงเปอร์เซ็นต์ช่วงใด ดังตัวอย่าง

ระดับคะแนน				
A	หรือ ก	หรือ 4	นักเรียนจะต้องได้	90% ขึ้นไป
B	„ ข	„ 3		75-89%
C	„ ค	„ 2		60-74%
D	„ ง	„ 1		45-59%
E หรือ F	„ จ	„ 0		44% ลงมา

การให้ระดับคะแนนแบบนี้ ได้แก่เกณฑ์การตัดสินผลการเรียนตามระเบียบการวัดผลของกระทรวงศึกษาธิการ ระดับประถมศึกษา พุทธศักราช 2520 และมัธยมศึกษา พ.ศ. 2521 ซึ่งใช้เกณฑ์ดังนี้

ตาราง 11-4 แสดงเกณฑ์การตัดสินผลการเรียนตามระเบียบการวัดผลของกระทรวงศึกษาธิการ

ระดับ	ความหมาย			ประถม	มัธยม
4	ดีมาก	หมายถึง	คะแนนตั้งแต่ร้อยละ	80 ขึ้นไป	80-100
3	ดี	"	"	70-79	70-79
2	ค่อนข้างดี	"	"	60-69	60-69
1	พอใช้	"	"	40-59	50-59
0	ต้องแก้ไข	"	"	ต่ำกว่าร้อยละ 40	0-49

การตัดเกรดในระบบนี้มีจุดอ่อนตรงที่ใช้เปอร์เซ็นต์เป็นหลัก โดยไม่ใช้วิจารณญาณของผู้สอนเข้ามาเกี่ยวข้องเลย ซึ่งที่จริงแล้วเปอร์เซ็นต์หรือคะแนนที่นักเรียนได้นั้นขึ้นอยู่กับความยากง่ายของข้อสอบ กล่าวคือถ้าข้อสอบง่ายนักเรียนก็จะได้คะแนนสูงและถ้าหากข้อสอบยากนักเรียนก็จะได้คะแนนต่ำ นอกจากนี้ 50% ของแต่ละโรงเรียน ยังบอกไม่ได้ว่าเท่ากันไหม ถ้าหากไม่เท่ากันแล้วอะไรจะเกิดขึ้น ถ้า 55% ของโรงเรียนแห่งหนึ่งเท่ากับ 45% ของอีกโรงเรียนหนึ่ง นักเรียนที่เรียนโรงเรียนหลังก็จะเป็นผู้โชคร้ายคือสอบตก และในทางตรงกันข้ามถ้าหากเขาไปเรียนในโรงเรียนแรกก็จะสอบได้ ทำนองเดียวกัน เกรด B โรงเรียนวัดลิงขบ อาจเท่ากับเกรด D ของโรงเรียนเตรียมอุดมศึกษาก็ได้

3. การตัดเกรดโดยใช้อัตราส่วน วิธีนี้จะยึดเปอร์เซ็นต์ที่กำหนดไว้เป็นหลักซึ่งอาจเปลี่ยนแปลงได้ตามความเหมาะสม หรือจะกำหนดตายตัวเลยก็ได้ เช่น อาจกำหนดอัตราส่วนในแต่ละเกรด ดังต่อไปนี้

ระดับคะแนน A	มี	10%
ระดับคะแนน B	มี	20%
ระดับคะแนน C	มี	40%
ระดับคะแนน D	มี	20%
ระดับคะแนน E	มี	10%

หรืออาจกำหนดเปอร์เซ็นต์ตามการแจกแจงปกติก็ได้ เช่น

ระดับคะแนน A	มี	2%
ระดับคะแนน B	มี	14%
ระดับคะแนน C	มี	68%
ระดับคะแนน D	มี	14%
ระดับคะแนน E	มี	2%

นอกจากนี้อาจกำหนดช่วงในแต่ละระดับคะแนนให้ผู้ประเมินยืดหยุ่นได้ตามความเหมาะสม ดังตัวอย่าง

ระดับคะแนน A	มีช่วงระหว่าง 0-15%
ระดับคะแนน B	มีช่วงระหว่าง 20-30%
ระดับคะแนน C	มีช่วงระหว่าง 40-50%
ระดับคะแนน D	มีช่วงระหว่าง 10-20%
ระดับคะแนน E	มีช่วงระหว่าง 0-10%

การตัดเกรดวิธีนี้เหมาะสำหรับผู้ที่ไม่ชอบคำนวณเลขมาก ๆ

4. การตัดเกรดโดยการเปรียบเทียบกับนักเรียนในกลุ่ม วิธีนี้เป็นการนำเอาคะแนนมาตรฐานภายในกลุ่มมาเปรียบเทียบกัน โดยแปลงคะแนนที่นักเรียนสอบได้ให้เป็นคะแนนมาตรฐานเสียก่อน จึงนำมาเปรียบเทียบกัน แล้วจึงใช้วิจารณญาณของผู้สอนกำหนดเกณฑ์การพิจารณาตามสภาพของกลุ่มนั้น ดังนั้นเกรดของนักเรียนจะขึ้นอยู่กับตำแหน่งที่เขาชนะเพื่อนได้มากน้อยเพียงใด มิได้ขึ้นอยู่กับคะแนนในการสอบของเขา การตัดเกรดตามแนวนี้ยึดหลักการกระจายของคะแนนตามโค้งปกติ ถ้าหากจะนำวิธีการตัดเกรดแบบนี้ไปใช้กับกลุ่มที่มีลักษณะเฉพาะ เช่นเด็กเก่งทั้งกลุ่มหรืออ่อนทั้งกลุ่ม ก็อาจเกิดความไม่เหมาะสมขึ้นได้ การตัดเกรดแบบนี้ยังมีข้อเสียอีกอย่างหนึ่ง คือเป็นการส่งเสริมให้นักเรียนเกิดความเห็นแก่ตัว คือพยายามเอาตัวรอดตามลำพัง ทั้งนี้เพราะเกรดที่จะได้ขึ้นอยู่กับว่าเขาทำคะแนนได้ดีกว่าเพื่อน ๆ หรือไม่ และโดยวิธีการตัดเกรดแบบนี้จะทำให้นักเรียนไม่อยากเรียนร่วมกับพวกเก่ง ๆ การตัดเกรดตามวิธีการดังกล่าวนี้มีหลายวิธีด้วยกันแต่ในที่นี้จะขอเสนอเพียง 3 วิธี คือ

4.1 การตัดเกรดโดยใช้คะแนนมาตรฐาน z (Linear z-score)

การตัดเกรดโดยวิธีนี้ เหมาะสำหรับวิชาที่มีการสอบเพียงครั้งเดียวแล้วนำคะแนนที่ได้มาตัดเกรดหรือให้ระดับคะแนนเลย โดยเริ่มต้นจากการแปลงคะแนนดิบให้อยู่ในรูปของคะแนนมาตรฐาน z ดังสูตร

$$z = \frac{X - \bar{X}}{S.D.}$$

ตามธรรมดาคะแนนมาตรฐาน z จะมีค่าอยู่ระหว่าง – 3 S.D. ถึง + 3 S.D. มีคะแนนเฉลี่ยเป็น 0 ความเบี่ยงเบนมาตรฐานเป็น 1 ส่วนการที่จะให้กี่ระดับนั้นขึ้นอยู่กับครูผู้สอน สมมุติว่าต้องการให้ 5 ระดับคะแนน อาจให้ดังนี้

เกรด A หรือ 4 สำหรับผู้ได้คะแนนมาตรฐานตั้งแต่ +2 ขึ้นไป มีประมาณ 2%

เกรด B หรือ 3 สำหรับผู้ได้คะแนนมาตรฐานตั้งแต่ +1 ถึง +2 มีประมาณ 14%

เกรด C หรือ 2 สำหรับผู้ได้คะแนนมาตรฐานตั้งแต่ -1 ถึง +1 มีประมาณ 68%

เกรด D หรือ 1 สำหรับผู้ได้คะแนนมาตรฐานตั้งแต่ -2 ถึง -1 มีประมาณ 14%

เกรด E หรือ 0 สำหรับผู้ได้คะแนนมาตรฐานตั้งแต่ -2 ลงมา มีประมาณ 2%

หรืออาจคิดจากความยาวฐานของโค้งปกติที่มีค่าเท่ากับ 6 S.D. ถ้าต้องการตัด 5 เกรด ก็จะต้องแบ่งคะแนนเป็น 5 ช่วง แต่ละช่วงมีความยาวเท่ากับ 1.2 S.D. คะแนน C จะอยู่ตรงกลางโค้งปกติ คือเริ่มจาก – .6 S.D. ถึง + .6 S.D. เมื่อเป็นเช่นนี้คะแนน B จะเริ่มจาก + .6 S.D. จนถึง 1.8 S.D. และคะแนน A จะเริ่มจาก + 1.8 S.D. เป็นต้นไป ในทำนองเดียวกันคะแนน D จะเริ่มจาก .6 S.D. จนถึง – 1.8 S.D. และคะแนน E จะเริ่มจาก – 1.8 S.D. เป็นต้นไป ดังนั้น

เกรด A คือคนที่ได้คะแนนมาตรฐานตั้งแต่ + 1.8 S.D. ขึ้นไป

เกรด B คือคนที่ได้คะแนนมาตรฐานระหว่าง + 0.6 S.D. กับ + 1.8 S.D.

เกรด C คือคนที่ได้คะแนนมาตรฐานระหว่าง – 0.6 S.D. กับ + 0.6 S.D.

เกรด D คือคนที่ได้คะแนนมาตรฐานระหว่าง – 1.8 S.D. กับ – 0.6 S.D.

เกรด E คือคนที่ได้คะแนนต่ำกว่า – 1.8 S.D.

หรือถ้าครูผู้สอนต้องการให้เพียง 4 หรือ 3 เกรด ก็อาจกำหนดได้ดังนี้

คะแนน					
ระดับคะแนน	– 2	– 1	0	1	2
	D(7%)	C(43%)	B(43%)	A(7%)	แบบ 4 เกรด
	C(10.5%)	B(79%)		A(10.5%)	แบบ 3 เกรด

โดยปกติแล้วเกรด E มีค่าเท่ากับศูนย์ แต่ไม่ได้หมายความว่าคนที่สอบได้เกรด E จะไม่มีความรู้เลย เป็นแต่เพียงมีความรู้ยังไม่ถึงขีดต่ำสุดที่กำหนดไว้เท่านั้น

4.2 การตัดเกรดโดยใช้คะแนนมาตรฐาน T ปกติ (Normalized T-score)

การให้ระดับคะแนนโดยวิธีนี้ ทำโดยการเปลี่ยนคะแนนดิบให้อยู่ในรูปของคะแนนมาตรฐาน T ปกติ แล้วกำหนดการให้ระดับคะแนนดังนี้

วิธีที่ 1 คะแนน T ที่ 71 ขึ้นไป ได้ระดับคะแนน A

คะแนน T ที่ 60-70 ได้ระดับคะแนน B

คะแนน T ที่ 40-59 ได้ระดับคะแนน C

คะแนน T ที่ 30-39 ได้ระดับคะแนน D

คะแนน T ที่ 29 ลงมา ได้ระดับคะแนน E

วิธีที่ 2 ถ้าต้องการให้ระดับคะแนนน้อยกว่า 5 ระดับ ก็อาจดำเนินการตามลำดับขั้น ดังนี้

1. แปลงคะแนนดิบของนักเรียนแต่ละคนเป็นคะแนน T ปกติ แล้วหาความแตกต่างระหว่างคะแนน T สูงสุด และคะแนน T ต่ำสุด ซึ่งค่าความแตกต่างนี้ก็คือพิสัย สามารถคำนวณได้โดยใช้สูตรดังนี้

พิสัย = T สูงสุด – T ต่ำสุด

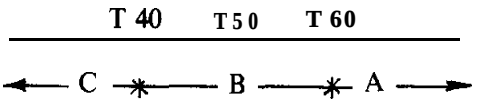
ตัวอย่าง คะแนน T สูงสุดเท่ากับ 70 คะแนน T ต่ำสุดเท่ากับ 10

พิสัย = 70 - 10 = 60

2. พิจารณาว่าจะให้กี่ระดับคะแนน และมีระดับเกรดใดบ้าง สมมุติจะให้ 3 ระดับคะแนน ก็ให้เอา 3 ไปหารพิสัย ค่าที่ได้จะเป็นช่วงของแต่ละเกรด

ช่วงของเกรด = $\frac{60}{3}$ = 20

3. ตัดเกรดเป็นช่วง ๆ โดยอาจเริ่มจากที่ใดก็ได้ เช่น อาจเริ่มจากสูงลงมา หรือจากต่ำขึ้นไป หรือเริ่มจากตรงกลางก็ได้ แต่โดยทั่วไปนิยมเริ่มที่จุดกึ่งกลาง ถ้าต้องการตัด 3 เกรด หรือ 5 เกรด จะต้องให้กลุ่มกลางคลุมคะแนน T 50 ทุกครั้ง โดยแบ่งช่วงของเกรดออกเป็น 2 ส่วน แล้วนับออกไปจากคะแนน T 50 ข้างละ 10 คะแนน ที่เหลือเป็นเกรด A และ C



4. ในกรณีที่เป็นเกรดคู่ เช่น 2 หรือ 4 เกรด ก็ให้นับเริ่มจาก T 50 ไปทางซ้ายและขวาได้เลย โดยไม่ต้องคร่อม T 50

ตาราง 11-5 ตัวอย่างการตัดเกรด

คะแนน T	กำหนดให้ 3 เกรด เป็น B, C, D,	กำหนดให้ 4 เกรด เป็น A, B, C, D	กำหนดให้ 5 เกรด เป็น A, B, C, D, E
75			A
70	B	A	
64			
61			B
57		B	
55			
52			
เฉลี่ย (50)	C		C
48			
46			
44		C	
41			D
39			
36			
33	D		
32		D	
30			E
25			
Range คะแนน	75 25 = 50	75 25 = 50	75 25 = 50
Range ของเกรด	$\frac{50}{3}$ = 16.6	$\frac{50}{4}$ = 12.5	$\frac{50}{5}$ 10
กำหนดเกรด	B — 58.3 7 5 C — 41.7 58.3 D — 2 5 41.7	A — 62.5 7 5 B — 5 0 62.5 C — 35.7 5 0 D — 2 5 37.5	A - 65 75 B — 55 65 C — 45 55 D — 35 45 E — 25 35

4.3 การตัดเกรดโดยใช้คะแนนมาตรฐาน 9 อันดับ (Stanine score)

ระบบการให้เกรดแบบ Stanine มีวิธีการเช่นเดียวกับการให้เกรด โดยใช้คะแนน Tปกติ แต่การให้เกรดแบบนี้ให้คุณค่ามากกว่าการให้เกรดแบบห้าตัวอักษร เพราะเพิ่มช่วงในการวัดมากขึ้น การให้เกรดแบบ Stanine จะเริ่มต้นที่ค่ามัธยฐาน (Median) โดย Stanine จะแบ่งช่วงของคะแนนออกเป็น 9 ช่วง แทนที่จะเป็น 5 ช่วง ซึ่งมีวิธีในการทำดังนี้

1. เรียงคะแนนจากสูงไปหาต่ำ
2. หาความถี่ของแต่ละคะแนน
3. หามัธยฐานของคะแนนชุดนั้น เช่น ถ้ามีนักเรียน 45 คน มัธยฐานของคะแนนคือคะแนนของคนที่ 23 ถ้ามีนักเรียน 46 คน มัธยฐานของคะแนนคือคะแนนของคนที่ 23 รวมกับคะแนนของคนที่ 24 แล้วหารด้วย 2
4. หาจำนวนคนที่อยู่ในแต่ละ Stanine จากจำนวนคนทั้งหมดตามทฤษฎี
5. พิจารณา Stanine ที่ 5 ก่อน แล้วพยายามปรับแต่ละ Stanine ให้ใกล้เคียงกับการจัดกลุ่มตามทฤษฎี โดยยึดหลักว่าคนที่ได้คะแนนเท่ากันจะต้องอยู่ใน Stanine เดียวกัน
6. ให้ระดับคะแนน

ตัวอย่าง

ในการสอบครั้งหนึ่งมีผู้เข้าสอบ 20 คน คะแนนเรียงตามลำดับแล้วได้ดังนี้

30 27 26 26 26 24 24 24 23 22

22 22 21 20 20 19 18 17 16 15

ตามทฤษฎีจะมีจำนวนคนอยู่ในแต่ละ Stanine ดังนี้

Stanine ที่	1	2	3	4	5	6	7	8	9	
จำนวนคน	1	1	2	4	4	4	2	1	1	

หรือคะแนนที่สอดคล้องกับแต่ละ Stanine เป็นดังนี้

คะแนน	30	27	26 26	26 24 24 24	23 22 22 22	21 20 20 19	18 17	16	15
Stanine ที่	9	8	7	6	5	4	3	2	1

แต่ตามหลักการของ Stanine คนที่ได้คะแนนเท่ากันต้องอยู่ Stanine เดียวกัน จึงต้องปรับใหม่ดังนี้

Stanine	9	8	7	6	5	4	3	2	1
คะแนน	30	27	26 26 2(	24 24 24 23 22 22 22 21 20 2(			19 16	7 16	15
ระดับคะแนน	A	B		C			D		E

ในกรณีที่มีคนสอบจำนวนมากให้ทำตารางแจกแจงความถี่เสียก่อน ดังนี้

ตาราง 11-6 แสดงการให้ระดับคะแนนโดยใช้ Stanine

คะแนน	ความถี่	cf	Stanine	จำนวนคนตามทฤษฎี	จำนวนคนที่เป็นจริง	ระดับคะแนน
30	1	45	9’			
27	1	44		2	2	ก
26	2	43	8			
25	1	41		3	3	ข
24	2	40				
23	2	38	7	5	7	
22	3	36				
21	2	33				
20	2	31	6	8	7	
19	3	29				ค
18	4	26	5	9	8	
17	4	22				
16	3	18				
15	3	15	4	8	8	
14	2	12				
13	3	10	3	5	5	
12	2	7				ง
11	1	5	2	3	3	
10	2	4				
8	1	2	1	2	2	จ
6	1	11				
	45					

ดร.อนันต์ ศรีโสภา ได้เสนอวิธีการให้เกรดแบบ Stanine สำหรับชั้นเรียนที่มีระดับความสามารถแตกต่างกัน 7 ระดับ ดังนี้

ตาราง 11-7 แสดงการให้เกรดแบบ Stanines สำหรับชั้นเรียนที่มีระดับความสามารถแตกต่างกัน 7 ระดับ

ระดับความสามารถเฉลี่ยของนักเรียนทั้งชั้น	ขีดจำกัดล่างสุดของ Stanine ที่ 9	จำนวนเปอร์เซ็นต์ในแต่ละ Stanine									ระดับความสามารถเฉลี่ย	
		9	8	7	6	5	4	3	2	1	Mean Stanine	Percentile
เยี่ยมยอด	1.90	18	16	19	21	13	8	4	1	0	6.60	79
ดีมาก	2.30	13	14	18	20	16	11	6	1	1	6.20	73
ดี	2.70	9	11	17	21	17	13	7	3	2	5.80	66
พอใช้	3.10	6	9	14	20	19	15	10	4	3	5.40	58
ปานกลาง	3.50	4	7	12	17	20	17	12	7	4	5.00	50
อ่อน	3.90	3	4	10	15	19	20	14	9	6	4.60	42
อ่อนมาก	4.30	2	3	7	13	17	21	17	11	9	4.20	34

สรุปเนื้อหาบทที่ 11

1. คะแนนดิบเป็นคะแนนที่ไม่มีความหมายในตัวของมันเอง จำเป็นต้องอาศัยข้อมูลอื่นเพิ่มเติม คะแนนจึงจะมีความหมาย
2. การรวมคะแนนดิบ 2 วิชาเป็นสิ่งที่ทำไม่ได้เนื่องจากคะแนนจากแต่ละวิชาวัดสมรรถภาพต่างกัน ดังนั้นหน่วยของคะแนนในแต่ละวิชาจึงไม่เท่ากัน
3. คะแนนมาตรฐานเป็นหน่วยของการวัดชนิดหนึ่งที่แปลงรูปมาจากคะแนนดิบ โดยทำให้คะแนนมีหน่วยเท่ากัน
4. หลักใหญ่ของคะแนนมาตรฐานก็คือการนำเอาคะแนนผลการสอบของแต่ละคนไปเปรียบเทียบกับกลุ่มใหญ่ หรือกลุ่มคนทั้งหมดที่อยู่ในฐานะหรือสภาพแวดล้อมเดียวกัน
5. คะแนนที่ปกติจะสูงหรือต่ำนั้นไม่ได้ขึ้นอยู่กับคะแนนดิบ แต่จะขึ้นอยู่กับว่าคะแนนของคน ๆ นั้นอยู่เหนือหรือสูงกว่าคะแนนของคนอื่นมากน้อยเพียงใด
6. การแปลความหมายของคะแนนทีปกติจะแปลในลักษณะที่ว่าคน ๆ นั้นชนะคนอื่นมากี่คนในคน 100 คน ตัวอย่างเช่น คะแนน T60 หมายความว่า คน ๆ นั้นชนะเพื่อนมา 84 คนใน 100 คน
7. ในการตัดเกรดเพื่อให้เกิดความถูกต้องและเหมาะสม ผู้ตัดเกรดควรจะต้องคำนึงถึงองค์ประกอบต่อไปนี้คือผลการวัด เกณฑ์การพิจารณา วิจารณญาณ และคุณธรรม
8. การตัดเกรดโดยใช้โค้งปกตินั้นเป็นเรื่องที่ผู้ตัดเกรดจะต้องมีความรู้ความเข้าใจอย่างแท้จริง หากขาดความเข้าใจที่แท้จริงอาจจะเกิดผลเสียหายขึ้นได้
9. การตัดเกรดนั้น ผู้ตัดเกรดจะต้องนำข้อมูลต่าง ๆ ที่ได้จากการวัดมาพิจารณาด้วยความระมัดระวัง ใช้ดุลพินิจและคุณธรรมในการตัดสินให้เกรด การตัดเกรดไม่มีกฎเกณฑ์ตายตัวแน่นอน บางครั้งอาจมี 5 เกรด 4 เกรด 3 เกรด หรือแม้แต่เกรดเดียวก็สามารถทำได้

คำถามท้ายบทที่ 11

1. แดงสอบภาษาไทยได้ 90 คะแนน จากคะแนนเต็ม 100 และสอบคณิตศาสตร์ได้ 40 คะแนน จากคะแนนเต็ม 100 คะแนนทีปกติของแดงในวิชาภาษาไทยจะต้องสูงกว่าคะแนนทีปกติของเขาในวิชาคณิตศาสตร์ใช่หรือไม่ เพราะเหตุใด
2. ทำไมคะแนนดิบจึงไม่มีความหมาย
3. การนำคะแนนดิบ 2 วิชามารวมกันนั้นจะทำได้หรือไม่ได้ ถ้าทำได้จะทำในกรณีใด ถ้าทำไม่ได้เป็นเพราะเหตุใด อธิบายพร้อมยกตัวอย่างประกอบ
4. คะแนนมาตรฐานคืออะไร วิธีการในการแปลงคะแนนดิบให้เป็นคะแนนมาตรฐานมีกี่วิธี อะไรบ้าง
5. การตัดเกรดของมหาวิทยาลัยในประเทศไทยปัจจุบันเป็นอย่างไร
6. การตัดเกรดของโรงเรียนมัธยมศึกษา ปัจจุบันใช้วิธีการตัดเกรดแบบใด