

# เกณฑ์ปกติ

## ความจำเป็นที่ต้องมีเกณฑ์ปกติ (Norm)

การสอบวัดใด ๆ จะมีความหมายก็ต่อเมื่อผลการสอบนั้นสามารถบอกถึงสถานภาพของผู้สอบว่ามีระดับความสามารถหรือมีสมรรถภาพเด่นด้อยกว่าบุคคลอื่น ๆ ภายในกลุ่มนั้นเพียงใด เพื่อครูอาจารย์จะได้นำผลนั้นไปใช้ในการประเมินค่า หรือแยกคัดจัดประเภทนักเรียนเหล่านั้นได้ถูกต้อง ว่าใครควรจะสอบได้สอบตก หรือเป็นเด็กประเภทเก่ง-อ่อนหรือควรจะรับไว้และไม่รับ เป็นต้น

จากนี้จึงอาจเห็นได้ว่าการสอบที่ดีจะต้องประกอบด้วยเครื่องมือวัดหรือแบบทดสอบที่ดี ร่วมกับวิธีการวัดและการแปลผลที่ถูกต้องด้วย จึงจะบรรลุผลตามความปรารถนา เรื่องนี้ก็ทำนองเดียวกับการหาน้ำหนักของวัตถุ ซึ่งจะต้องเริ่มด้วยการมีเครื่องชั่งที่ดี พร้อมกับวิธีการวัด และหน่วยตม่น้ำหนักตามพิภคมาตราที่กำหนดไว้จึงจะสามารถแปลความหมายของความหนัก-เบาให้ตรงกันได้

วิธีการประเมินค่าและแปลผลการสอบต่าง ๆ ที่กำหนดไว้ว่านักเรียนจะต้องสอบให้ได้คะแนนอย่างน้อย 50% หรือครึ่งหนึ่งของคะแนนเต็มจึงจะนับว่าเป็นผู้สอบได้หรือมีความรู้เข้าเกณฑ์ตามที่ต้องการนั้น ยังไม่เหมาะสมนัก ความผิดพลาดของการประเมินค่าแบบนี้ อยู่ตรงที่ความหลงผิดอย่างน้อย 3 ประการด้วยกัน คือ ในประการแรกก็จากที่นึกเอาเองว่าตัวคำถามหรือข้อสอบต่าง ๆ ที่ออกมานั้น มีคุณภาพดีเหมาะสมกับเด็กนักเรียนระดับนั้นแล้ว หรือบางทีถึงกับเชื่อว่า มีความยากง่ายเป็นมาตรฐานสำหรับชั้นนั้นแล้วก็มี โดยไม่ได้พิจารณาว่า ข้อสอบนั้นมีคุณสมบัติในด้านใด หรือมีความยากง่ายเป็นเท่าใด ฉะนั้นความคิดนี้จึงมีสภาพเป็นเพียงความเห็นที่เลื่อนลอย และยังห่างไกลกับวิธีการของวิทยาศาสตร์อยู่มาก

ความคลาดเคลื่อนประการที่สอง ก็ได้แก่การ

หลงผิดที่ไปยึดเอาตัวเลขคะแนนมาเป็นเครื่องชี้บอกปริมาณความรู้ของนักเรียน เช่นมีความคิดว่า เด็กที่สอบได้ 20 คะแนน จะมีความรู้มากเป็น 2 เท่าของเด็กที่สอบได้ 10 คะแนน ทำนองเดียวกับก้อนหิน 20 กิโลกรัม ก็นหนักเป็น 2 เท่าของหิน 10 กิโลกรัม หรือเงิน 20 บาท ก็มีค่ามากเป็น 2 เท่าของเงิน 10 บาท เป็นต้น ความเข้าใจผิดประเภทนี้เอง ที่ทำให้เกิดเกณฑ์ตัดสินได้-ตกที่ 50% กันขึ้น เพราะคิดไปว่า เด็กที่สอบได้ 50% เป็นผู้ที่มีความรู้ขนาดปานกลาง สมควรยอมให้เลื่อนชั้นได้ แต่ถ้าถามว่าคะแนน 50% นี้หมายถึงอะไร จะหมายถึงผู้ที่มีความรู้ครึ่งหนึ่งของหลักสูตร หรือครึ่งหนึ่งของครูที่สอน หรือครึ่งของอะไร ๆ แล้ว ก็ไม่สามารถชี้แจงได้อีกเช่นกัน นอกจากจะตอบว่า 50% ของคะแนนเต็มของข้อสอบฉบับนั้น เท่านั้นเอง

การที่เห็นแต่เพียงตัวเลขคะแนนของวิชาหนึ่ง มากกว่าอีกวิชาหนึ่ง แล้วจะไปลงสรุปว่าเด็กนั้นเก่งวิชาแรกมากกว่าวิชาหลัง เช่นนี้ ย่อมไม่ถูกต้องด้วยเหตุผลอันใดเลย เริ่มตั้งแต่คะแนนเต็มของแต่ละวิชาเหล่านั้นไม่เท่ากัน ซึ่งจะส่งผลให้นักเรียนมีโอกาสได้คะแนนสูง ๆ ในวิชาที่มีคะแนนเต็มมาก มากกว่าในวิชาที่มีคะแนนเต็มน้อย นี่เป็นประการที่หนึ่ง และในประการที่สองนั้น ถึงแม้ว่าจะนำคะแนนดิบเหล่านั้นมาทำเป็นเปอร์เซ็นต์เสียให้หมด โดยหวังจะให้มีฐานหรือคะแนนเต็มของทุกวิชาให้เป็น 100 หน่วยเท่า ๆ กันก็ตาม แต่ตัวเลขเปอร์เซ็นต์ที่ได้นั้น ก็ยังคงมีสภาพเป็นเพียงการอนุมานที่เลื่อนลอยอยู่มากเช่นเดียวกันเพราะไปหวังเอาเองว่าเด็กที่สอบเลขในใจได้คะแนน 5 จาก 10 นั้น คงจะมีความสามารถเทียบเท่ากับ 50% หรือสามารถสอบได้ 50 จาก 100 คะแนน ถ้าเราขยายคะแนนเต็มของวิชานี้ให้มากถึง 100 คะแนน เรื่องนี้ก็คล้าย ๆ กับว่า ถ้ามีคำถามอยู่ 10 ข้อ และเด็กตอบถูก 5 ข้อแล้ว ถึงเราจะขยายคำถามนั้นให้ยาวถึง 100 ข้อ เด็กก็จะ

ต้องตอบถูกถึง 50 ข้อด้วย ซึ่งมีโอกาสที่จะเป็นจริงน้อยมาก จึงเห็นได้ว่าความคิดคะแนนเป็นเปอร์เซ็นต์ก็ไม่ทำให้เกิดผลอะไรดีขึ้นกว่าคะแนนเต็มหรือคะแนนดิบ

ความผิดพลาดประการที่สาม ก็ได้แก่ความไม่รู้ในเรื่องธรรมชาติของตัวเลขคะแนน โดยนึกไปว่าคะแนนของแต่ละวิชามีหน่วยเป็นชนิดเดียวกันและเท่ากัน เพราะนับเพิ่มขึ้นทีละหนึ่ง ๆ แบบเดียวกับตัวเลขในคณิตศาสตร์ จึงสามารถนำมารวมกันได้โดยตรง ความเข้าใจผิดในเรื่องนี้ก็เทียบได้กับการที่เอาเงิน 1 ดอลลาร์ไปรวมกับ 1 บาท แล้วได้เป็น 2 ซึ่งเป็นวิธีการรวมที่ไม่มี ความหมายอะไร

สาเหตุที่ไม่สามารถนำคะแนนดิบของวิชาต่าง ๆ มาเทียบและรวมกันได้โดยตรงนั้น ก็เนื่องจากความแตกต่างในธรรมชาติและเนื้อหาของแต่ละวิชา เช่น คำถามในวิชาภาษาไทย เลขคณิต และศิลปะศึกษา ต่างต้องการความรู้ความสามารถในการตอบแตกต่างกันไป คนละชนิด ๆ ทั้งสิ้น ฉะนั้น 1 คะแนนของแต่ละวิชาเหล่านี้จึงมีคุณค่าไม่เหมือนกันหรือเท่ากัน และไม่สามารถนำมารวมกันได้โดยตรงด้วย ถึงแม้ในวิชาเดียวกันก็ตามข้อสอบที่ถามเมื่อตอนต้นภาคเรียนกับกลางภาคเรียน ก็ยังไม่สามารถนำมารวมกันได้ด้วยตัวเอง เพราะอย่างไร ๆ ฉบับหลังก็ย่อมจะยากกว่าฉบับแรกอยู่เสมอทั้ง ๆ ที่ทั้งสองฉบับนั้นต่างก็ยังมีวัดสมรรถภาพชนิดเดียวกันอยู่

ความผิดพลาดในการสอบวัดทั้ง 3 ประการที่กล่าวมานี้โดยเฉพาะอย่างยิ่งในเรื่องการตัดลินได้-ตกที่ 50% นั้น นับว่าเป็นความบกพร่องที่มีผลต่อชีวิตมนุษย์เป็นอันมาก ความผิดพลาดนี้อาจชี้ให้เห็นได้ง่าย ๆ ว่าคะแนนต่าง ๆ ที่เด็กสอบได้ในแต่ละครั้ง ๆ นั้นจะขึ้นอยู่กับความยากง่ายของข้อสอบเป็นประการแรก และเป็นประการที่สำคัญที่สุด นั่นคือ ถ้าคราวใดกรรมการออกข้อสอบยากเกินกว่าภูมิปัญญาของนักเรียนแล้ว คะแนนจากการสอบครั้งนั้นก็จะลดต่ำลงทันที นักเรียนทั่วไปจะไม่สามารถทำคะแนนได้ถึง 50% กลายเป็นสอบตกกันมาก และโดยนัยตรงข้าม ถ้าข้อสอบครั้งใดถามง่ายกว่าธรรมดา คะแนนของผู้สอบใน

ครั้งนั้นก็สูงมากกว่าปรกติ ทำให้นักเรียนสอบไล่ได้กันมากขึ้น

ด้วยประการฉะนี้ เพื่อให้เกิดความยุติธรรมแก่ทุกฝ่าย การวัดผลที่เป็นวิทยาศาสตร์ตามแบบประชาธิปไตย จึงนิยมใช้คะแนนเฉลี่ยจากการสอบเหล่านั้นมาเป็นหลักเริ่มต้นในการวัดผล แต่ทั้งนี้มิได้หมายความว่า ให้เอาคะแนนเฉลี่ยมาใช้เป็นเกณฑ์ตัดสินได้-ตก แต่หมายถึงการเอาคะแนนเฉลี่ยมาใช้เป็นจุดเริ่มต้นของการวัดเพื่อนำไปตัดสินได้-ตกอีกทีหนึ่ง หลักการใหม่นี้มีประโยชน์ตรงที่สามารถช่วยขจัดความบกพร่องของการสอบได้หลายประการ เช่น ในบางโอกาสที่ข้อสอบยากเกินไปจริง ๆ ผู้สอบต่างก็จะได้คะแนนน้อย ๆ ไปตามกัน คะแนนเฉลี่ยของข้อสอบนั้นก็จะลดต่ำลงไปด้วย ดังนั้นจุดหลักของการวัดก็จะเลื่อนต่ำลงไปเองโดยอัตโนมัติ หรือถ้าบังเอิญว่าข้อสอบนั้นง่ายเกินไป เด็ก ๆ ก็จะได้คะแนนสูงมากขึ้น ซึ่งก็เท่ากับไปเลื่อนคะแนนเฉลี่ยและจุดหลักให้สูงตามไปด้วย ผู้ที่สอบได้คะแนน 60% ก็อาจกลายเป็นคนรั้งท้ายหรือสอบตกไปเลยก็ได้ ดังนั้นจะเห็นได้ว่าการใช้คะแนนเฉลี่ยเป็นจุดหลัก ความไม่ยุติธรรมอันเนื่องมาจากข้อสอบยากหรือง่ายทั้งสองกรณีนี้ ก็จะสิ้นไปเองโดยปริยาย

การวัดผลแบบวิทยาศาสตร์ ใช้สภาพความจริงเท่าที่ปรากฏจากการสอบนั้น ๆ มาเป็นหลักในการวินิจฉัยและสำหรับตีราคาการสอบได้สอบตก ไม่พยายามใช้ความคาดคะเนของผู้ออกข้อสอบมาเกี่ยวข้องด้วยเลย ทั้งนี้โดยมีเหตุผลว่า ถึงแม้ว่าครูหรือผู้ออกข้อสอบจะมีความปรารถนาสูงส่ง อยากให้มาตรฐานการเล่าเรียนสูงขึ้นและต้องการให้นักเรียนดีอย่างนั้นและเก่งอย่างนี้ก็ตาม แต่ถ้าสิ่งนั้นเกินวิสัยเด็กธรรมดาสามัญชนแล้ว ความปรารถนาดีเหล่านั้นก็อาจกลายเป็นผลร้ายได้ง่าย ๆ ตัวอย่างจริงของเรื่องนี้ก็ได้แก่การสอบคัดเลือกต่าง ๆ ที่ปรากฏว่าสอบคัดเลือกแล้วไม่มีผู้สอบได้เลย โดยอ้างว่ายังมีมาตรฐานความรู้อยู่ต่ำเกินไปเป็นต้น พฤติการณ์นี้ชี้ชัดว่ากำหนดความหวังหรือตั้งเกณฑ์ไว้สูงเกินไปจนลึมนึกถึงสภาพของวัดดูดิบ

ที่มีอยู่ในมือ ว่ามีคุณภาพอยู่เพียงแค่นั้น หรือมีให้  
เลือกแต่เพียงเท่านั้นเอง ถึงแม้จะมีเด็กอื่นที่เก่งกว่านี้  
อยู่อีกก็จริง แต่ก็เป็นโชคร้าย ที่เด็กเก่งเหล่านั้น เขาก็  
ไม่ต้องการมาเป็นลูกศิษย์เรา เขาพากันไปสมัครสอบที่  
อื่นหรือเลือกเรียนวิชาอื่น ๆ เสียหมด ก็อาจเป็นได้

วิธีการวัดผลที่ใช้คะแนนเฉลี่ยเป็นจุดหลัก แล้ว  
จำแนกคะแนนออกเป็นชั้น ๆ ไปตามสภาพความเป็น  
จริงของนักเรียนเหล่านั้นเรียกว่าการสร้างมาตรฐาน  
เกณฑ์ปรกติ (standardized norm) ของผลการสอบ  
เหล่านั้น ซึ่งในปัจจุบันถือว่าเป็นขบวนการวัดผลที่  
ยุติธรรมและสอดคล้องกับธรรมชาติมากที่สุด หลักการ  
ง่าย ๆ ของการวัดแบบนี้มีอยู่ว่า ถ้าใครสอบได้คะแนน  
อยู่เหนือจุดหลัก ก็แปลว่าเป็นผู้มีความรู้ความสามารถ  
สูง เก่งกว่าเด็กทั่ว ๆ ไป ถ้ายังสูงห่างจากคะแนนเฉลี่ย  
มากเท่าใด ก็ยังเป็นเด็กเก่งมากขึ้นเพียงนั้น และ  
โดยนัยตรงข้าม ถ้านักเรียนคนใดสอบได้คะแนนต่ำกว่า  
คะแนนเฉลี่ยก็เป็นเด็กไม่เก่ง และถ้ายิ่งต่ำมากลงไปอีก  
ก็ยิ่งเป็นเด็กอ่อนในวิชานั้นมากขึ้นตามลำดับ

จากนี้จะเห็นว่า การสร้างเกณฑ์ปรกติ มีความ  
สำคัญตรงที่จะต้องแบ่งคะแนนออกจากจุดหลักให้เป็น  
ช่วง ๆ ให้เท่ากันอย่างมีความหมาย ทั้งตอนบนและ  
ตอนล่าง และในการนี้จะใช้คะแนนดิบหรือเปอร์เซ็นต์  
มาแบ่งไม่ได้ ด้วยเหตุผลดังที่กล่าวมาแล้ว ว่าคะแนน  
ดิบมีหน่วยไม่เท่ากันและยังไร้ความหมายอยู่มาก  
จำเป็นที่จะต้องใช้วิธีการทางสถิติมาช่วยแปลง  
คะแนนต่าง ๆ เหล่านั้น ให้มีหน่วยเทียบเท่ากันเสีย  
ก่อน จึงจะสามารถตัดคะแนนเหล่านี้ออกเป็นช่วง ๆ  
ได้ตามต้องการ จากนั้นก็ทำบัญชีสำเริงสำหรับเปลี่ยน  
มาตรฐานว่า แต่ละคะแนนดิบที่เด็กสอบได้นั้น จะมีค่า  
เทียบเท่ากับคะแนนแปลงรูปที่สร้างขึ้นมานั้นเป็นเท่าใด

การสร้างเกณฑ์ปรกติสามารถทำได้โดย  
เปลี่ยนคะแนนดิบของทุกวิชาให้เป็นคะแนน T ปรกติ  
(normalized T score) หมดทุกฉบับ การปฏิบัติเช่นนี้  
ก็ด้วยวัตถุประสงค์ ที่จะให้ผลการทดสอบต่าง ๆ  
สามารถเปรียบเทียบกันได้โดยตรงและสะดวกต่อการใช้  
ยิ่งขึ้น และอีกประการหนึ่งนั้นก็เพราะมีความเห็นว่า

คะแนน T ปรกตินี้มีหน่วยที่เท่ากัน และมีหลักการสม  
บูรณ์มากกว่าคะแนนแปลงรูปชนิดอื่น ๆ ด้วย

การสร้างเกณฑ์ปรกติของแบบทดสอบต่าง ๆ  
ให้เป็นคะแนน T ปรกตร่วมกันนี้ก็เหมือนกับที่เรา  
เปลี่ยนนาตรเงินปอนด์ ดอลลาร์ และเงินกีบ ให้เป็น  
นาตรเงินบาทเสียให้หมด นั่นเอง ฉะนั้นคะแนนใหม่  
ที่ได้จึงมีความหมายเป็นอย่างเดียวกัน และสามารถนำ  
มารวมกันได้โดยตรงอย่างถูกต้องตามหลักวิชา ยิ่งกว่า  
นั้น มาตรคะแนน T นี้ก็ง่ายใช้สะดวก มีตั้งแต่ 0 ถึง  
100 คล้าย ๆ กับสตาจค์ของเงินบาท หรือเซนติเมตร  
ของไม้เมตร โดยมีรายเฉลี่ยอยู่กึ่งกลางตรงที่ T 50 พอดี  
และในการใช้คะแนนชนิดนี้ ก็มิต้องไปกังวลว่า วิชาใด  
จะมีคะแนนเต็ม คะแนนเฉลี่ย หรือความเบี่ยงเบน  
มาตรฐานเป็นเท่าใดเลย เพราะไม่ว่าวิชานั้นจะมีค่า  
เหล่านี้เป็นเท่าใดก็ตาม เมื่อนำมาแปลงเป็นคะแนน T  
แล้ว ต่างก็จะมีคะแนนเฉลี่ยใหม่เท่ากับ 50 หน่วย และ  
ความเบี่ยงเบนก็จะเท่ากับ 10 หน่วย เหมือนกันหมด  
ทุกวิชา กลายเป็นตัวเลขมาตรฐานเดียวกันตลอด

### วิธีสร้างเกณฑ์ปรกติ

การสร้างเกณฑ์ปรกติแบบคะแนน T ต้องใช้  
เทคนิคทางสถิติอยู่บ้าง ผู้ที่สนใจอาจศึกษารายละเอียด  
และวิธีทำได้จากรายงาน*ความก้าวหน้าของโครงการ  
สร้างแบบทดสอบมาตรฐาน ฉบับที่ 2* ซึ่งสำนักทดสอบ  
ทางการศึกษาและจิตวิทยา มหาวิทยาลัยศรีนครินทร-  
วิโรฒพิมพ์เป็นเอกสารอีกชุดหนึ่งต่างหากแล้ว

### การใช้เกณฑ์ปรกติ

1. **วิธีอ่านเกณฑ์ปรกติ** เมื่อครูตรวจกระดาษ  
คำตอบว่า นักเรียนคนใดตอบถูกกี่ข้อ หรือได้คะแนน  
ดิบเท่าใดแล้วก็ไปเปิดบัญชีอ่านให้เป็นคะแนน T  
ปรกติได้เจยทันที การอ่านเกณฑ์ปรกติจะต้องระวังให้  
ถูกแถวของวิชานั้น ๆ เสมอ วิธีที่ดี ควรอ่านไปที่ละ  
วิชาจนหมดจำนวนนักเรียนที่เข้าสอบ และเมื่อจะอ่าน  
เกณฑ์ปรกติของวิชาใดก็ควรใช้กระดาษปิดตัวเลขของ  
เกณฑ์อื่น ๆ เสียก่อน เพื่อกันลายุตา

2. **วิธีบันทึกคะแนน T ปรกติ** ควรบันทึกด้วย  
หมึกหรือดินสอคนละสีกับตัวเลขคะแนนดิบ และ

ควรตกลงกันว่า ช่องที่ 1 สำหรับเขียนคะแนนดิบ ส่วนช่องที่ 2 สำหรับเขียนคะแนน T ปรกติ เป็นต้น

3. การแปลคะแนน T ปรกติ คะแนนชนิดนี้ นิยมแปลความหมายเป็นในรูปของเปอร์เซ็นต์ที่อยู่เหนือกว่าผู้อื่น ดังที่แสดงไว้ในบัญชีต่อไปนี้ *เปอร์เซ็นต์ที่อยู่เหนือ* นี้มีความหมายว่านักเรียนที่สอบได้คะแนน T 30 ก็อ่านจากตารางว่าเขาเก่งกว่า หรือมีความรู้เหนือกว่าคนอื่นอยู่ 2.28% คือใน 100 คน เขาจะมีความรู้สูงกว่าเด็กอื่น ๆ อยู่ 2.28 หรือ 2 ซึ่งก็เท่ากับเป็นเด็กจวนได้ที่โหล่นั่นเอง

ถ้าจะเทียบให้เห็นง่ายขึ้นอีกก็อาจสมมุติได้ว่า ถ้าเอาความรู้ของเด็ก 100 คนมาเรียงกันเป็นแถวบนชั้นบันไดความรู้ ตั้งแต่ 0 คือ ชั้นที่พื้นจนถึงยอด 100 ชั้น แล้วนักเรียนคนที่สอบได้คะแนน T 30 ของวิชาใด ๆ ก็ตามก็เสมือนกับเขายืนอยู่บนบันไดขั้นที่สูงกว่าคนอื่น ๆ อยู่เพียง 2 คน ซึ่ง ณ ตรงนั้น จะมีเด็กเป็นรองอยู่ได้เขา 2 คน พร้อม ๆ กับอีก 98 คน อยู่สูงกว่าเขา โดยทำนองเดียวกัน ผู้ที่สอบได้คะแนน T 50 ในวิชาใด ๆ ก็ตาม ก็อ่านจากบัญชีตรงกับ % ที่อยู่เหนือผู้อื่น 50.00 ฉะนั้นแปลว่า เด็กคนนั้นมีความรู้ระดับปานกลางพอดีคือเขายืนอยู่ตรงกึ่งกลางของบันไดความรู้ มีผู้อื่นอยู่ได้เขา 50 คน พร้อม ๆ กับอีก 50 คน อยู่เหนือเขา

**บัญชีสำหรับแปลงคะแนน T ปรกติ ให้เป็นจำนวนร้อยละที่อยู่เหนือกว่าผู้อื่น**

(จาก Henry E. Garrett ในหนังสือ *Statistics in Psychology and Education*)

คะแนน T ปรกติ	% ที่อยู่เหนือผู้อื่น	คะแนน T ปรกติ	% ที่อยู่เหนือผู้อื่น
10	.0032	16	.034
11	.0048	17	.048
12	.007	18	.069
13	.011	19	.097
14	.016	20	.13
15	.023	21	.19

คะแนน T ปรกติ	% ที่อยู่เหนือผู้อื่น	คะแนน T ปรกติ	% ที่อยู่เหนือผู้อื่น
22	.26	57	75.80
23	.35	58	78.81
24	.47	59	81.59
25	.62	60	84.13
26	.82	61	86.43
27	1.07	62	88.49
28	1.39	63	90.32
29	1.79	64	91.92
30	2.28	65	93.32
31	2.87	66	94.52
32	3.59	67	95.54
33	4.46	68	96.41
34	5.48	69	97.13
35	6.68	70	97.72
36	8.08	71	98.21
37	9.68	72	98.61
38	11.51	73	98.93
39	13.57	74	99.18
40	15.87	75	99.38
41	18.41	76	99.53
42	21.19	77	99.65
43	24.20	78	99.74
44	27.43	79	99.81
45	30.85	80	99.865
46	34.46	81	99.903
47	38.21	82	99.931
48	42.07	83	99.952
49	46.02	84	99.966
50	50.00	85	99.977
51	53.98	86	99.984
52	57.93	87	99.9890
53	61.79	88	99.9928
54	65.54	89	99.9952
55	69.15	90	99.9968
56	72.57		

ถ้าใครสอบได้คะแนน T 70 ซึ่งตรงกับเปอร์เซ็นต์ที่เหนือกว่าผู้อื่นอยู่ 97.72 หรือ 98 % ก็แปลว่าเขายืนอยู่บนบันไดขั้นที่ 98 อยู่สูงกว่าคนอื่น 98 คน และยังมีอยู่อีก 2 คนเท่านั้นที่มีความรู้สูงกว่าเขา เป็นต้น

4. การประเมินคะแนน T ประกติ เมื่อสามารถวัดว่า นักเรียนคนใดได้คะแนน T เท่าใด และเสมือนกับเขายืนอยู่บนบันไดความรู้ขั้นที่เท่าใดแล้ว ต่อไปก็เป็นการประเมินราคาของคะแนนเหล่านั้นว่ามีคุณภาพสูง-ต่ำ หรือดีเลวปานใด อันเป็นเรื่องของการตัดสินใจขาดหรือการตีราคาโดยสรุปอย่างมีหลักเกณฑ์ให้นักเรียนผู้นั้น ว่าควรจะยกย่องให้เขาเป็นผู้มีความรู้ความสามารถเทียบเท่ากับระดับใด ตามหลักสภานโยบายแบ่งคุณภาพเหล่านี้ออกเป็น 5 ระดับ ดังเช่นในชั้นอุดมศึกษาก็อาจตั้งชื่อเป็นระดับ A, B, C, D, E หรืออาจเป็นระดับ ก, ข, ค, ง และ จ ก็ได้ ซึ่งชื่อต่าง ๆ เหล่านี้ก็พอจะเทียบได้กับระดับ ดีมาก-ดี-พอใช้-ยังไม่พอใช้-อ่อน นั้นเอง

การกำหนดวิธีประเมินคุณภาพผลการทดสอบ อาจแบ่งคะแนน T ออกเป็น 5 ระดับดังนี้

ตั้งแต่ T 65 และสูงกว่า	แปลว่า ดีมาก
ตั้งแต่ T 55-65	แปลว่า ดี
ตั้งแต่ T 45-55	แปลว่า พอใช้
เฉพาะตรง T 50	แปลว่า มีความสามารถปานกลางพอดี
ตั้งแต่ T 35-45	แปลว่า ยังไม่พอใช้
ตั้งแต่ T 35 และต่ำกว่า	แปลว่า อ่อน

จะสังเกตเห็นว่า การแบ่งระดับดังข้างต้นนี้ จะมีคะแนน T บางตัวซ้ำกันที่ตรงหัวและตรงท้ายของช่วงคะแนนดังเช่นที่ T 55 เป็นต้น การที่เป็นเช่นนี้ก็เพราะตรง T 55 นั้นเป็นจุดแบ่งเขตระหว่างกลุ่ม ฉะนั้น ถ้านักเรียนคนใดได้คะแนน T ตรงจุดแบ่งเขตเหล่านั้นพอดี คือที่ T 35, 45, 55 และ 65 แล้ว ก็อาจเกิดลังเลไม่แน่ใจว่าควรจะให้เขาอยู่ในกลุ่มใด วิธีแก้ไขในเรื่องนี้ก็ให้ถือเป็นหลักไว้ว่า ให้เลื่อนนักเรียนที่คาบเส้นผู้นั้นขึ้นไปอยู่ในกลุ่มสูงที่ติดไปเสมอ เพื่อผลทางจิตวิทยา เพราะ

โอกาสที่นักเรียนคนเดียวกันจะได้คะแนนตรงนั้นซ้ำ ๆ กัน มีอยู่น้อยมาก

การประเมินค่าโดยวิธีนี้ ในการสอบทั่ว ๆ ไป จะมีนักเรียนระดับดีมากอยู่ประมาณ 7 เปอร์เซนต์ของนักเรียนทั้งหมด และจะมีอยู่ประมาณ 24 % ที่อยู่ในเกณฑ์ดี ประมาณ 38 % ที่อยู่ในเกณฑ์พอใช้หรือปานกลาง กับอีกประมาณ 24 % ยังไม่พอใช้ และที่เหลือสุดท้ายประมาณ 7 % เป็นเด็กกลุ่มอ่อน

**ประโยชน์ของเกณฑ์ปรกติ**

คุณค่าของเกณฑ์ปรกตินี้อยู่เป็นอเนกประการ แต่ในที่นี้จะกล่าวแต่เรื่องที่มีความสำคัญมากเพียง 3 ประการดังนี้

1. ใช้สำหรับเปลี่ยนคะแนนดิบให้เป็นหน่วยที่มีความหมายยิ่งขึ้น เรื่องนี้มาจากเหตุผลว่าตัวเลขคะแนนดิบที่ได้จากการสอบต่าง ๆ นั้น ยังมีหน่วยไม่เท่ากัน เป็นคนละชนิด และยิ่งแปลความหมายไม่แน่ชัดว่าหมายถึงอะไรกันแน่ ฉะนั้นในการวัดความสามารถใด ๆ จึงมีความจำเป็นที่จะต้องปรับคะแนนของแต่ละวิชาเหล่านั้นให้เป็นหน่วยเดียวกันเสียก่อน ดังเช่นเกณฑ์ปรกตินี้ จึงจะสามารถนำมาเปรียบเทียบและอธิบายความหมายของตัวเลขเหล่านั้นได้ชัดเจน และมีหลักเกณฑ์ยิ่งขึ้น

คุณประโยชน์ที่เป็นผลพลอยได้จากการใช้เกณฑ์ปรกตินี้ได้แก่การที่สามารถนำตัวเลขคะแนนเหล่านั้นมารวมหรือหารเฉลี่ยกันได้โดยตรง อย่างจำนวนธรรมดา เช่นเมื่อต้องการทราบความสามารถเฉลี่ยของนักเรียนคนใดหรือของกลุ่มนักเรียนใดก็ตาม ก็อาจกระทำได้โดยนำเอาคะแนน T ปรกติของทุก ๆ วิชาเหล่านั้นมารวมกันแล้วหารเฉลี่ย เช่น ด.ช. สำเร็จ สอบวิชาภาษาไทยได้ T 55 วิชาความงามได้ T 60 และออกแบบได้ T 40 ฉะนั้นคะแนนเฉลี่ยของทั้ง 3 วิชานี้ จึงเท่ากับ  $\frac{55+60+40}{3} = 51.66$  หรือเท่ากับ T 52 ซึ่งอาจประเมินได้ว่า เขามีความสามารถด้านศิลปะในระดับปานกลางหรือพอใช้ ถ้าอยากทราบให้แน่ชัดลงไปอีกว่าเขาเด่นด้อยขนาดใดแน่ ก็ไปอ่านบัญญัติในหน้า 20 ว่า T ตรงกับเปอร์เซ็นต์ที่เหนือผู้อื่น 57.93 หรือ

58% ซึ่งแปลได้ว่าเราสามารถปีนบันไดความรู้ได้ถึงขั้นที่ 58 ใน 100 ขั้น หรือในเด็ก 100 คน จะมีเด็กอื่นอยู่ได้เขา 58 คน หรือจะพูดกลับเสียก็ได้ว่า เขามีความรู้สูงกว่าคนอื่นอยู่ 58 คน พร้อม ๆ กับมีอีก 42 คนที่อยู่เหนือเขาหรือเก่งกว่าเขา

การที่เราสามารถนำคะแนน T ของวิชาต่าง ๆ มารวมกันได้โดยตรงนี้ ก็เพราะเราแปลงคะแนนดิบของแต่ละวิชาเหล่านั้นด้วยเกณฑ์ปรกติ ให้เป็นมาตราเดียวกันด้วยทุกหลักวิชา ซึ่งคุณสมบัติเช่นนี้จะกระทำกับคะแนนดิบและเปอร์เซ็นต์ไม่ได้เลย

2. ใช้ในการประเมินผลการศึกษา เมื่อครูอาจารย์สอนศิษย์ไประยะหนึ่งก็คงอยากทราบถึงประสิทธิภาพในการสอนของตนและคุณภาพในการเรียนของเด็กว่ามีผลสัมฤทธิ์ถึงระดับใด สมดังความปรารถนาแล้วหรือไม่เป็นต้น นี่ก็คือความต้องการที่จะวัดและตีราคาคุณภาพของการศึกษาเล่าเรียนว่ามีมาตรฐานสูงต่ำเพียงใดนั่นเอง ดังนั้นจะต้องนำเอาผลการทดสอบของแต่ละครั้งมาเปรียบเทียบกับเกณฑ์ที่กำหนดไว้ และคำตอบที่ถูกต้องนั้นย่อมจะขึ้นอยู่กับความเที่ยงตรงและความเชื่อถือได้ของการสอบวัดเป็นประการแรก เกณฑ์ปรกติที่สร้างขึ้นนี้สามารถสนองความต้องการดังกล่าวได้เป็นอย่างดี ด้วยเหตุผลดังกล่าวมาแล้วในข้อก่อน

อนึ่ง ถ้าสามารถเก็บข้อมูลเหล่านี้ไว้สัก 3-5 ปี แล้วนำแบบทดสอบชุดเดิมนั้นพร้อมกับเกณฑ์ปรกติไปทดลองกับนักเรียนรุ่นใหม่อีกครั้งหนึ่ง และนำผลมาเปรียบเทียบกันก็จะช่วยให้สามารถประเมินคุณภาพทางการศึกษาและวิวัฒนาการของท้องถิ่นนั้น ๆ ได้ อีกด้วย ว่าภายในระยะขบปีดังกล่าว สถานภาพทางการศึกษามีการเคลื่อนไหวไปในทางใด สามารถพัฒนาตรงตามทิศและระดับที่ปรารถนาไว้แล้วหรือไม่ และมีมาตรฐานสูงต่ำเป็นเช่นไรจากเดิม เป็นต้น

3. ใช้ในการแนะแนว นักเรียนในแต่ละชั้นและในแต่ละห้อง ย่อมจะมีระดับความรู้ความสามารถแตกต่างกันไปเป็นรายบุคคล เราผู้เป็นครูย่อมต้อง

การจะทราบว่าได้เด็กคนใดเก่งอ่อนในด้านใดบ้าง เพื่อจะชักจูงส่งเสริมเขาได้ตรงจุด ให้เขาร่ำเรียนตามความถนัด และมีชีวิตในงานนั้นอย่างมีความสุขสมกับอัตภาพ ซึ่งเป็นเรื่องของการแนะแนวนั่นเอง และไม่ว่าจะเป็นการแนะแนวทางการศึกษาหรืออาชีพก็ตามที ย่อมจะต้องเริ่มด้วยการวินิจฉัยศักยภาพของบุคคลหรือกลุ่มชนนั้นเสียก่อน ว่ามีความเด่นด้อยในทางใด จากนั้นจึงจะให้คำแนะนำแก่เขาได้อย่างมีหลักมีเกณฑ์แต่อย่างไรก็ดี การวินิจฉัยนี้ต้องอาศัยความแม่นยำและเที่ยงตรงของการวัด เช่นเดียวกับการประเมินค่า และจะยิ่งต้องการความแน่นอนในระดับที่สูงกว่าด้วย เพราะการวินิจฉัยความสามารถของบุคคล เป็นการเปรียบเทียบสมรรถภาพชนิดหนึ่งของเขากับความสามารถชนิดอื่น ๆ ของบุคคลผู้นั้นเอง ฉะนั้นเครื่องมือที่ใช้สำหรับวินิจฉัยสมรรถภาพเพื่อการแนะแนวนี้ จึงต้องมีความละเอียดคมมากกว่าข้อสอบแบบธรรมดาทั่ว ๆ ไป จึงจะสำเร็จ เกณฑ์ปรกติแบบคะแนน T ที่สร้างขึ้นนี้ มีความไวเพียงพอสำหรับงานด้านนี้อยู่แล้ว

ตรงนี้อาจมีผู้สงสัยว่า น่าจะใช้จำนวนคะแนนดิบหรือเปอร์เซ็นต์ที่เด็กสอบได้มาเป็นเครื่องวินิจฉัยก็ได้ ซึ่งเป็นความเข้าใจที่ผิดและคลาดเคลื่อนอยู่มาก เพราะคะแนนเหล่านั้นเป็นเพียงตัวเลขที่แสดงปริมาณงาน ที่เด็กแต่ละคนทำถูกต้องเท่านั้น ยังไม่แน่นอนเสมอไปว่าเด็กทั่วไปจะเห็นงานนั้นเป็นของง่ายหรือยากตามนั้นด้วย ฉะนั้นเพื่อให้เกิดความยุติธรรมและมั่นใจยิ่งขึ้น การสอบวัดที่ดี จึงนิยมนำแบบทดสอบไปทดลองกับเด็กอื่นอีกมาก ๆ แล้วสร้างเป็นเกณฑ์ปรกติขึ้น การวัดผลโดยใช้เกณฑ์ปรกติ ก็เสมือนกับเราเอาเด็กคนนี้ไปเปรียบเทียบกับกลุ่มชน หรือกับประชากรเด็กอื่นนับด้วยพันด้วยหมื่นคนนั่นเอง คล้ายกับเรามีห้องเรียนใหญ่มีนักเรียนนับหมื่นคนแล้วทำการทดสอบพร้อมกัน ผลการทดสอบจึงย่อมมีความเชื่อมั่นได้มากกว่า และสามารถวินิจฉัยความเด่นด้อยของเขาได้ถูกต้องกว่าด้วย เพราะเป็นการเปรียบเทียบระหว่างเด็กคนนี้กับเด็กอื่น ๆ มากมาย จึงมีโอกาสผิดพลาดได้น้อยและยังได้นำคะแนนเหล่านั้นมาปรับปรุงด้วยหลักการต่าง ๆ จน

เป็นเกณฑ์ปกติแบบคะแนน T ด้วยแล้ว ก็เป็นอันหวัง  
ได้โดยว่า เกณฑ์ปกติที่สร้างขึ้นนั้นจะอำนวยความสะดวก  
ต่อขบวนการแนะแนวเป็นอเนกประการ จนอาจกล่าว

ได้ว่าถ้าการสอบโตปราศจากเกณฑ์แบบนี้แล้ว การ  
วินิจฉัยสมรรถภาพเพื่อการแนะแนว ก็มีอาจกระทำ  
ได้อย่างเต็มภาคภูมินัก

**ชวาล แพร์ตกุล**

(หมายเหตุ เนื้อหาข้างต้นส่วนใหญ่ตัดตอนมาจากหนังสือ "การทดสอบเพื่อค้นและพัฒนาสมรรถภาพ" ของ  
ศาสตราจารย์ ดร.ชวาล แพร์ตกุล หน้า 95-104)