

บทที่ 4

การพัฒนาแบบสอบทักษะพื้นฐานทางวิชาการเพื่อการประเมินและวินิจฉัยบุคคลที่มีภาวะสมาธิสั้น บกพร่องทางการเรียนรู้ และออทิซึม

บทนำ

ตามที่โครงการวิจัย “การพัฒนาศักยภาพนักเรียนที่มีภาวะสมาธิสั้น บกพร่องทางการเรียนรู้ และออทิซึม : ระยะที่ 1 การสร้างกระบวนการและเครื่องมือที่เป็นมาตรฐานสำหรับการคัดกรองนักเรียนที่มีภาวะสมาธิสั้น บกพร่องทางการเรียนรู้ และออทิซึม” ได้รับทุนอุดหนุนการวิจัยจากสำนักงานคณะกรรมการวิจัยแห่งชาติ ปีงบประมาณ 2547 เพื่อดำเนินการสร้างและพัฒนาเครื่องมือนักเรียนที่มีภาวะสมาธิสั้น บกพร่องทางการเรียนรู้ และออทิซึม ระดับชั้นประถมศึกษาปีที่ 1 – 6 อายุระหว่าง 6 – 13 ปี 11 เดือน ที่เป็นมาตรฐานสำหรับประเทศไทย นั้น ผลผลิตที่ได้จากโครงการวิจัยนี้คือ “แบบคัดกรองนักเรียนที่มีภาวะสมาธิสั้น บกพร่องทางการเรียนรู้ และออทิซึม” มีชื่อภาษาอังกฤษว่า “KUS-SI Rating Scales: ADHD/ LD/ Autism (PDDs)” มีคุณภาพ ประสิทธิภาพ และมีเกณฑ์มาตรฐานระดับประเทศ (National Norms) สำหรับแปลความหมายของคะแนนจากแบบคัดกรอง ขณะนี้ศูนย์การศึกษาพิเศษ เขตการศึกษาต่างๆ และ บุคลากรที่เกี่ยวข้องกับการจัดการศึกษาพิเศษ ได้นำแบบคัดกรองนี้ไปใช้อย่างกว้างขวาง และมีความชำนาญมากขึ้น

ผลการคัดกรองนักเรียนที่มีภาวะสมาธิสั้น บกพร่องทางการเรียนรู้ และออทิซึม เป็นข้อมูลสำคัญ 1 ใน 3 ในกระบวนการจัดทำแผนการศึกษาเฉพาะบุคคลอย่างถูกต้องตามหลักวิชาการ ยังมีข้อมูลอีก 2 ส่วน คือ ผลการประเมินทักษะพื้นฐานทางวิชาการ และผลการประเมินสติปัญญา ประกอบกับข้อมูลนักเรียนจากครูประจำชั้น และผู้ปกครอง สำหรับข้อมูล 2 ส่วนดังกล่าว ต้องอาศัยเครื่องมือที่มีคุณภาพตามมาตรฐานของแบบวัด ดังนั้น คณะผู้วิจัยจึงได้เสนอแผนพัฒนาแบบสอบทักษะพื้นฐานทางวิชาการ ด้วยการสร้างและพัฒนาแบบสอบทักษะพื้นฐานทางวิชาการให้เป็นแบบสอบมาตรฐาน (Standardized Tests) ระดับประเทศ เพื่อให้บริการครูและหน่วยงานที่ให้ความช่วยเหลือทางการศึกษา นักเรียนที่มีความต้องการจำเป็นพิเศษเป็นรายบุคคล มีข้อมูลสำคัญอีกด้านหนึ่งเสริมความพร้อมที่จะจัดการศึกษาอย่างมีมาตรฐานตามแผนการจัดการศึกษาเฉพาะบุคคล (Individual Education Plan)

แนวคิดการพัฒนาแบบสอบมาตรฐานเพื่อประเมินทักษะพื้นฐานทางวิชาการ

ปัจจุบันนี้ พบว่าแบบสอบทักษะพื้นฐานทางวิชาการที่มีประสิทธิภาพสูง และใช้อย่างแพร่หลายในประเทศที่ใช้ภาษาอังกฤษ คือ แบบทดสอบ Wide Range Achievement Test 4 (WRAT4) มีประวัติการพัฒนาที่ยาวนานตั้งแต่ ทศวรรษที่ 1930 โดยนักจิตวิทยา Joseph F. Jastak โดยขยายแบบวัด Wechsler-Bellevue Scales ของ David Wechsler ที่สร้างเมื่อประมาณ 70 ปีก่อน Wide Range

Achievement Test ที่สร้างขึ้นวัดพื้นฐานทางวิชาการสำคัญ 3 ด้าน คือ การอ่าน การเขียน และการคิดคำนวณ ต่อมาได้ตีพิมพ์เผยแพร่ครั้งแรก เมื่อ ค.ศ. 1946 จากนั้นมีการพัฒนาต่อเนื่องจนถึง ปี ค.ศ. 1993 ได้เผยแพร่ฉบับ Wide Range Achievement Test : Third Edition (WRAT3) แล้วพัฒนาต่อเนื่อง ฉบับที่ใช้อยู่ในปัจจุบัน คือ Wide Range Achievement Test : Fourth Edition (WRAT4) ตีพิมพ์เมื่อปี ค.ศ. 2006 โดยเพิ่มการวัดความสามารถในการเข้าใจประโยค (Sentence Comprehension) และพัฒนาการแปลความหมายคะแนนสอบด้วยคะแนนมาตรฐานเทียบตามกลุ่มอายุ (Age-based norms) จาก 5 ปี ถึง 94 ปี และมีคะแนนมาตรฐานเทียบกับระดับชั้นเรียน (grade-based norms) (Wilkinson, G. S. & Robertson, G. J.;2006)

กระบวนการพัฒนาแบบสอบ WRAT เป็นไปตามหลักการวิจัยและพัฒนา เริ่มกรอบแนวคิดวัดทักษะที่มีความจำเป็นต่อการเรียนรู้สามด้าน คือ การอ่าน การเขียน และ การคิดคำนวณ การพัฒนาต่อมาเพิ่มด้านความเข้าใจประโยค เพื่อขยายผลการประเมินให้เกิดประโยชน์มากขึ้น ในด้านกระบวนการสร้างเครื่องมือแสดงหลักฐานของคุณภาพทั้งด้านความตรง ความเที่ยง และเน้นประโยชน์ในการนำไปใช้อย่างกว้างขวาง มีการสร้างมาตรฐานการแปลความหมายคะแนนด้วยเกณฑ์มาตรฐานตามอายุ และระดับชั้นเรียนของผู้สอบ นอกจากนี้ มีวิธีการบริหารการสอบที่สะดวก และแนวทางการใช้ผลสอบที่ง่าย

จากการศึกษาทั้งในด้านหลักทฤษฎีของการประเมินความสามารถทางวิชาการพื้นฐาน และระเบียบวิธีพัฒนาแบบสอบตามมาตรฐานของแบบสอบ WRAT ผู้วิจัยพบแนวทางที่จะนำมาสร้างแบบแผนการพัฒนาแบบสอบมาตรฐานที่จะใช้กับการประเมินทักษะทางวิชาการพื้นฐานของเด็กไทยที่ต้องการความช่วยเหลือพิเศษ แต่อย่างไรก็ตาม ด้วยเงื่อนไขที่แตกต่างกับการพัฒนาแบบสอบ WRAT หลายประการ ผู้วิจัยได้ศึกษาทฤษฎีการวัดความสามารถที่เหมาะสมภายใต้เงื่อนไขของการพัฒนาเป็นระยะๆ สำหรับระยะแรก ด้วยเงื่อนไขของเวลาที่จำกัด 1 ปี และได้รับอนุมัติขอเขตการสร้างมาตรฐานคะแนนเพียง 2 จังหวัด จึงได้กำหนดกรอบเนื้อหาของทักษะพื้นฐานทางวิชาการ 4 เรื่อง คือ การอ่าน และการเขียนสะกดคำ ความเข้าใจประโยค และการคิดคำนวณทางคณิตศาสตร์ และใช้ระเบียบการสร้างแบบวัดเพื่อให้ได้สารสนเทศที่อธิบายความสามารถของประชากรนักเรียนตั้งแต่ระดับชั้นประถมศึกษาปีที่ 1 ถึงมัธยมศึกษาปีที่ 6 ด้วยการใช้ Item Response Theory Methods (Kolen, M. J., & Brennan, R. L. 2004; 155-227) และสร้างแบบแผนการแปลผลคะแนนมาตรฐานตามระดับชั้น (grade-based norms) จัดทำตารางชั้นประถมศึกษาปีที่ 1 ถึง ชั้นมัธยมศึกษาปีที่ 6 รวม 12 ชั้นปี และจัดทำอีก 4 ตารางของ 4 ช่วงชั้น คือ ช่วงชั้นที่ 1 (ชั้นประถมศึกษาปีที่ 1-3) ช่วงชั้นที่ 2 (ชั้นประถมศึกษาปีที่ 4-6) ช่วงชั้นที่ 3 (ชั้นมัธยมศึกษาปีที่ 1-3) และช่วงชั้นที่ 4 (ชั้นมัธยมศึกษาปีที่ 4-6)

ด้วยแนวคิดดังกล่าว คณะผู้วิจัยให้ชื่อแบบสอบที่จะพัฒนาขึ้นนี้ว่า **แบบสอบทักษะพื้นฐานทางวิชาการ** ชื่อภาษาอังกฤษว่า **Kasetsart Basic Academic Skills Test** และใช้ตัวย่อว่า **KBAST** เนื้อหาที่นำมาสร้างเพื่อวัดทักษะพื้นฐานนี้ กำหนดความยากให้อยู่ในกรอบของความรู้ขั้นพื้นฐานที่เป็น

ระดับอ่านออกเขียนได้ที่ใช้ในชีวิตประจำวัน ซึ่งจะมีกลุ่มผู้ทรงคุณวุฒิเป็นผู้พิจารณา ทั้งนี้เพื่อใช้เป็นเครื่องมือหาข้อมูลพื้นฐานทางวิชาการที่จำเป็นเพื่อการวางแผนการให้ความช่วยเหลือแก่นักเรียนบุคคลที่มีภาวะสมาธิสั้น บกพร่องทางการเรียนรู้ และออทิซึม

วัตถุประสงค์ของการวิจัย

เพื่อพัฒนาแบบสอบทักษะพื้นฐานทางวิชาการเพื่อการประเมินและวินิจฉัยบุคคลที่มีภาวะสมาธิสั้น บกพร่องทางการเรียนรู้ และออทิซึม โดยมีวัตถุประสงค์เฉพาะดังนี้

1. พัฒนาแบบสอบมาตรฐานวัดทักษะพื้นฐานทางวิชาการ ซึ่งประกอบด้วยแบบวัดทักษะ 4 ด้าน คือ แบบสอบการอ่านคำ (word reading) แบบสอบการสะกดคำ (word spelling) แบบสอบความเข้าใจประโยค (sentence comprehension) และแบบสอบการคิดคำนวณ และแก้โจทย์ปัญหาอย่างง่าย (math computation)
2. สร้างตารางเกณฑ์ปกติ (Norms) เพื่อการแปลความหมายคะแนนของผู้สอบตามระดับความสามารถในรายทักษะที่อิงระดับชั้นเรียน และระดับช่วงชั้น ตามมาตรฐานของกลุ่มประชากรจังหวัดกรุงเทพมหานครและจังหวัดขอนแก่น
3. จัดทำคู่มือบริหารการสอบแบบสอบทักษะพื้นฐานทางวิชาการ เพื่อการนำไปใช้

ขอบเขตของการวิจัย

ขอบเขตของการพัฒนาแบบสอบมาตรฐานเพื่อประเมินทักษะพื้นฐานทางวิชาการ ประกอบด้วย

1. เป็นแบบสอบด้านพุทธิพิสัยที่วัดระดับความจำ และการนำไปใช้ในระดับพื้นฐาน สร้างหลักสูตรการประเมินทักษะพื้นฐานทางวิชาการพัฒนาจากการคัดเลือกเนื้อหาจาก มาตรฐานการเรียนรู้ และตัวชี้วัดในกลุ่มสาระการเรียนรู้ภาษาไทยและคณิตศาสตร์ ระดับชั้นประถมศึกษาปีที่ 1 ถึงมัธยมศึกษาปีที่ 6 ของหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551 ในส่วนที่เป็นทักษะที่ต้องฝึกฝนจนเกิดความชำนาญในการใช้ภาษาเพื่อสื่อสารได้อย่างมีประสิทธิภาพและมีทักษะการคิดคำนวณพื้นฐานที่นำไปใช้ในชีวิตจริง ซึ่งประกอบด้วย ทักษะการอ่าน การอ่านออกเสียงคำ การอ่านในใจ เพื่อสร้างความเข้าใจที่ใช้ในชีวิตประจำวัน ทักษะการเขียน การเขียนสะกดตามอักขรวิธี การเขียนสื่อสารในรูปแบบประโยค และทักษะการคิดคำนวณทางคณิตศาสตร์ ทั้งนี้ คณะผู้เชี่ยวชาญได้พิจารณาและลงความเห็นในรายละเอียดเพื่อความเหมาะสมของความจำเป็นของการประเมินกลุ่มบุคคลที่มีภาวะสมาธิสั้น บกพร่องทางการเรียนรู้ และออทิซึม

2. ลักษณะของแบบสอบมาตรฐานวัดทักษะพื้นฐานทางวิชาการ เป็นชุดแบบสอบ (battery) ประกอบด้วยแบบสอบย่อยวัดทักษะ 4 ด้าน คือ แบบสอบการอ่านคำ (word reading) แบบสอบการ

สะกดคำ (word spelling) แบบสอบความเข้าใจประโยค (sentence comprehension) และแบบสอบการคิดคำนวณทางคณิตศาสตร์ (math computation) แบบสอบย่อยรายทักษะหนึ่งๆ จะมีแบบสอบ 4 ฉบับที่มีความยากของข้อสอบแตกต่างกันตามระดับช่วงชั้น ฉบับที่ 1 มีความยากของข้อสอบตามหลักสูตรช่วงชั้นที่ 1 (ประถมศึกษาปีที่ 1 ถึง 3) ฉบับที่ 2 มีความยากตามช่วงชั้นที่ 2 (ประถมศึกษาปีที่ 4 ถึง 6) ฉบับที่ 3 มีความยากตามช่วงชั้นที่ 3 (มัธยมศึกษาปีที่ 1 ถึง 3) และฉบับที่ 4 มีความยากตามช่วงชั้นที่ 4 (มัธยมศึกษาปีที่ 4 ถึง 6) ทั้งนี้เป็นโครงสร้างกำหนดขึ้นตามรูปแบบการเพิ่มพูนทักษะการเรียนรู้ชั้นต่อชั้น (Grade to Grade Growth Model)

3. กระบวนการพัฒนาแบบสอบมาตรฐาน ใช้ระเบียบวิธีการพัฒนาที่แสดงลักษณะเฉพาะของแบบสอบและข้อสอบ แสดงความตรงของแบบวัด และความเที่ยงของระบบคะแนน จัดทำเกณฑ์ปกติจากกลุ่มตัวอย่างของประชากรมาตรฐานนักเรียนชั้นประถมปีที่ 1 ถึงชั้นมัธยมศึกษาปีที่ 6 ของกรุงเทพมหานครและจังหวัดขอนแก่น ในปีการศึกษา 2555

นิยามศัพท์

ทักษะพื้นฐานทางวิชาการ (Basic Academic Skills) หมายถึง ทักษะการอ่านคำ การสะกดคำ ความเข้าใจประโยค และการคำนวณทางคณิตศาสตร์ ที่มีขอบเขตเนื้อหาตามหลักสูตรแกนกลางระดับชั้นประถมศึกษาปีที่ 1 ถึงมัธยมศึกษาปีที่ 6 ที่มีระดับความยากที่จำเป็นต่อการเรียนรู้ในชีวิตประจำวัน (literacy levels) ที่เหมาะสมกับการนำไปใช้ประเมินและวินิจฉัยบุคคลที่มีภาวะสมาธิสั้นบกพร่องทางการเรียนรู้ และออทิซึมทั้งนี้ เป็นผลการพิจารณาของคณะผู้เชี่ยวชาญด้านการสอนและการพัฒนาหลักสูตร และมีประสบการณ์การสอนในวิชานั้นและระดับชั้นนั้นๆ

ชุดแบบสอบ (Test Batteries) หมายถึง แบบสอบทักษะพื้นฐานทางวิชาการ ที่สร้างขึ้นประกอบด้วย 4 แบบสอบย่อยวัดรายทักษะ ภายในแบบสอบย่อยที่วัดแต่ละทักษะ จะถูกจัดเป็น 4 ฉบับลำดับความยากในทักษะนั้นอิงเนื้อหาตามระดับช่วงชั้น ดังนี้ แบบสอบฉบับที่ 1 วัดระดับช่วงชั้นที่ 1 (ประถมศึกษาปีที่ 1-3) แบบสอบฉบับที่ 2 วัดระดับช่วงชั้นที่ 2 (ประถมศึกษาปีที่ 4-6) แบบสอบฉบับที่ 3 วัดระดับช่วงชั้นที่ 3 (มัธยมศึกษาปีที่ 1-3) และแบบสอบฉบับที่ 4 วัดระดับช่วงชั้นที่ 4 (มัธยมศึกษาปีที่ 4-6)

ความตรงของแบบวัดเชิงเหตุผล (Logical Validity) หมายถึง ผลการพิจารณาคูณลักษณะของแบบวัดแต่ละฉบับโดยผู้เชี่ยวชาญต่อความสอดคล้องของข้อกำหนดลักษณะเฉพาะของข้อสอบ (test specification) กับวัตถุประสงค์ของชุดแบบสอบทักษะพื้นฐานทางวิชาการ และความสอดคล้องของข้อสอบ (test items) กับข้อกำหนดลักษณะเฉพาะของข้อสอบ (test specification) นั้น ซึ่งเป็น

หลักฐานความตรงภายใน (Internal Evidence of validity) ที่เกิดขึ้นในกระบวนการพัฒนาแบบสอบมาตรฐาน

ความตรงเชิงโครงสร้างของแบบสอบรายทักษะ (Construct Validity) หมายถึง ค่าความสามารถเฉลี่ย () จากการใช้แบบสอบย่อยรายทักษะกับกลุ่มตัวอย่างมาตรฐานการพัฒนาแบบสอบ มีลักษณะสอดคล้องกับรูปแบบการเพิ่มพูนทักษะการเรียนรู้ขั้นต่อขั้น (Grade to Grade Growth Model) กล่าวคือ ค่าความสามารถเฉลี่ย แบบสอบฉบับที่ 1 (μ_1) น้อยกว่าค่าความสามารถเฉลี่ย แบบสอบฉบับที่ 2 (μ_2) ค่าความสามารถเฉลี่ยแบบสอบฉบับที่ 2 (μ_2) น้อยกว่าค่าความสามารถเฉลี่ยแบบสอบฉบับที่ 3 (μ_3) ค่าความสามารถเฉลี่ยแบบสอบฉบับที่ 3 (μ_3) น้อยกว่าค่าความสามารถเฉลี่ยแบบสอบฉบับที่ 4 (μ_4) [$\mu_1 < \mu_2 < \mu_3 < \mu_4$] หลักฐานเชิงสถิติ ใช้การทดสอบความแตกต่างของผลสอบของกลุ่มตามระดับช่วงชั้น ด้วยสถิติทดสอบ : the Kruskal-Wallis test ความตรงเชิงโครงสร้างของแบบสอบรายทักษะนี้เป็นหลักฐานความตรงจากการตรวจสอบคุณลักษณะทางการวัด (Measurement Properties) ของแบบสอบย่อย ที่เกิดจากการใช้ทฤษฎีการตอบสนองข้อสอบ ตรวจสอบสมบัติทางการวัดผลที่สำคัญของรายข้อ และชุดแบบสอบผ่านกระบวนการเทียบมาตราคะแนนของแบบสอบ (test equating) เพื่อสร้างชุดแบบสอบที่วัดความสามารถ () ตามรูปแบบการวัดพัฒนาการของผู้สอบ (grade to grade growth model)

ความเที่ยง (Reliability) หมายถึง คุณลักษณะของแบบสอบแต่ละฉบับที่ทำหน้าที่วัดความสามารถเดียวกันในทักษะนั้นๆ แสดงด้วยค่าสัมประสิทธิ์ความคงที่ภายในของแบบสอบ (Cronbach's Alpha) และความคลาดเคลื่อนมาตรฐานของการวัด (SEM) จากผลการใช้แบบสอบ (คะแนน) กับกลุ่มตัวอย่างมาตรฐานของการพัฒนาชุดแบบสอบทักษะพื้นฐานทางวิชาการ ค่าความเที่ยงที่สูงแสดงถึงคะแนนผลการวัดรายบุคคลมีความน่าเชื่อถือ

คะแนนมาตรฐานของแบบสอบ (Scaled score) หมายถึง คะแนนมาตรฐานที่เกิดจากการแปลงคะแนนดิบรายทักษะตามระดับชั้นเป็นคะแนนมาตรฐานปกติด้วยวิธีการทางสถิติ กำหนดให้ค่าเฉลี่ยเท่ากับ 100 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 15

คะแนนเกณฑ์ปกติ (Norms) หมายถึง ชุดคะแนนมาตรฐานของแบบสอบที่ได้จากกลุ่มตัวอย่างของประชากรมาตรฐานที่กำหนดในการพัฒนาแบบสอบมาตรฐาน KBAST ซึ่งเป็นประชากรนักเรียนชั้นประถมศึกษาปีที่ 1 ถึง มัธยมศึกษาชั้นปีที่ 6 ของกรุงเทพมหานครและจังหวัดขอนแก่น เพื่อแปลความหมายผลคะแนนดิบที่ได้จากการแจกแจงปกติของเกณฑ์ปกติที่พัฒนาขึ้น

ตารางคะแนนเกณฑ์ปกติของแบบสอบตามระดับชั้น (Grade Norms) หมายถึง ตารางที่นำเสนอชุดของคะแนนดิบที่ได้จากการสอบรายทักษะที่มีค่าตรงกับคะแนนเกณฑ์ปกติของระดับชั้นนั้นๆ เพื่อใช้แสดงผลการประเมินระดับทักษะของรายบุคคลจากคะแนนดิบที่ได้ เทียบกับระดับและช่วงชั้น

การทบทวนวรรณกรรมที่เกี่ยวข้อง

การจัดการศึกษาสำหรับเด็กที่มีความต้องการพิเศษ จำเป็นต้องมีการวางแผน ที่ตรงกับสภาพและ สอดคล้องกับความต้องการของผู้เรียน ข้อมูลเกี่ยวกับนักเรียนที่ต้องนำมาพิจารณาได้แก่ รายงานประเมิน ทางจิตวิทยา ความพร้อมทางวิชาการ พฤติกรรมทางสังคม และ ลักษณะเฉพาะที่มีการติดตามให้ความ ช่วยเหลือจากแพทย์ ผู้ปกครองและผู้เกี่ยวข้อง สำหรับข้อมูลด้านสติปัญญา และการประเมินความพร้อม ทางวิชาการ ครูผู้ทำหน้าที่ประเมินจำเป็นต้องมีเครื่องมือที่มีคุณภาพ เพื่อให้ได้สารสนเทศที่ตรงสภาพของ นักเรียน ส่วนข้อมูลด้านอื่นๆ สามารถรวบรวมได้จากการติดตามแหล่งข้อมูลต่างๆ เช่น ข้อมูลด้านประวัติ ทางสุขภาพ จากบันทึกการรักษาจากแพทย์หรือสถานพยาบาล ข้อมูลเกี่ยวกับผลการเรียนรู้ในอดีตจากรายงานแผนและผลงานการเรียนรู้ตลอดจนสัมภาษณ์ครูหรือผู้ให้ความช่วยเหลือเด็กในเวลาที่ผ่านมา สัมภาษณ์ผู้ปกครองและผู้เกี่ยวข้องกับนักเรียนเกี่ยวกับพฤติกรรมที่เป็นบุคลิก การปรับตัว เป็นต้น

จากการสำรวจของคณะวิจัยโครงการนี้ พบว่าเครื่องมือมาตรฐานสำหรับประเมินความพร้อมทาง วิชาการโดยเฉพาะทักษะขั้นพื้นฐานทางวิชาการยังไม่มี ครูหรือผู้มีหน้าที่จัดการศึกษาให้เด็กที่มีภาวะ สมาธิสั้น บกพร่องทางการเรียนรู้ และออทิซึมจะใช้การสังเกตและพิจารณาตามประสบการณ์ของตนเอง คณะวิจัยจึงได้เสนอโครงการพัฒนาแบบวัดมาตรฐานขึ้น เป็นการสนับสนุนการทำงานเพิ่มประสิทธิภาพ การจัดการศึกษาให้เด็กที่มีภาวะสมาธิสั้น บกพร่องทางการเรียนรู้ และออทิซึม

ดังนั้นการนำเสนอวรรณกรรมที่เกี่ยวข้องในบทนี้จึงเป็นสาระของหลักการ แนวทาง และเทคนิคที่ นำมาออกแบบพัฒนาแบบสอบทักษะพื้นฐานทางวิชาการของโครงการวิจัยนี้ กำหนดไว้ 4 ตอน คือ

1. แบบสอบมาตรฐาน WRAT : การพัฒนา และการนำไปใช้
2. การออกแบบชุดแบบสอบที่มีความยากต่างกันตามรูปแบบ grade to grade growth model
3. การตรวจสอบคุณภาพแบบสอบมาตรฐาน
4. สรุปรอบแนวคิดการพัฒนาแบบสอบทักษะพื้นฐานทางวิชาการ KBAST

ตอนที่ 1 แบบสอบมาตรฐาน Wide Range Achievement Test 4 (WRAT4)

แบบทดสอบ Wide Range Achievement Test 4 (WRAT 4) พัฒนาขึ้นโดย Wilkinson, G. S. & Robertson, G. J. ปี 2006 เป็นแบบทดสอบรายบุคคล (Individual Test) ใช้เพื่อประเมินทักษะ พื้นฐานทางวิชาการที่จำเป็นต่อการเรียนรู้ การสื่อสาร และการคิดอย่างมีประสิทธิภาพ ซึ่งประกอบด้วย การอ่าน การสะกดคำ และการคิดคำนวณทางคณิตศาสตร์ แบบทดสอบ WRAT 4 ใช้สำหรับบุคคลอายุ 5 – 94 ปี ประกอบด้วย 4 แบบทดสอบย่อย ดังนี้

1. แบบทดสอบด้านการอ่าน ประเมินทักษะการอ่านและรู้จักคำ
2. แบบทดสอบด้านความเข้าใจประโยค ประเมินความเข้าใจความหมายของคำและประโยค

3. แบบทดสอบด้านการสะกดคำ ประเมินความสามารถในการเขียนจากเสียงที่ได้ยิน
4. แบบทดสอบด้านการคำนวณทางคณิตศาสตร์ ประเมินความสามารถเบื้องต้นทางคณิตศาสตร์ การนับ ระบุตัวเลข แก้โจทย์ปัญหาอย่างง่าย และคำนวณตัวเลข

แบบทดสอบ WRAT 4 มีการดำเนินการทดสอบ และการแปลผลที่ไม่ยุ่งยาก มีกระบวนการสร้างแบบทดสอบที่เป็นมาตรฐาน มีความเที่ยงและความตรงสูง คณะผู้วิจัยจะใช้เป็นแนวทางพัฒนาแบบสอบทักษะพื้นฐานทางวิชาการเพื่อการประเมินและวินิจฉัยบุคคลที่มีภาวะสมาธิสั้น บกพร่องทางการเรียนรู้ และออทิซึม สำหรับนักเรียนไทยต่อไป

1.1 การพัฒนาเนื้อหาแบบสอบ (WRAT4)

คณะทำงานของ WRAT4 ทบทวนแบบสอบ WRAT3 ซึ่งมีผลการวิจัยและพัฒนาแบบสอบย่อย (1) Reading (2) Spelling (3) Arithmetic อย่างรอบด้าน ได้ศึกษาถึงกระบวนการพัฒนาที่บ่งบอกจุดแข็ง จุดอ่อนของ ทั้ง 3 แบบสอบย่อย แล้วจึงปรับเพียงเล็กน้อย แล้วสร้างแบบสอบย่อยที่ 4 Sentence Comprehension ซึ่งเป็นของใหม่ และทำการพัฒนาข้อสอบและทดลอง (tryout) อย่างเป็นระบบ เพื่อให้สามารถกำหนดเนื้อหาที่เหมาะสม ขั้นสุดท้ายได้จัดชุดแบบสอบ WRAT 4 ประกอบด้วยแบบสอบย่อย 4 แบบสอบ และจัดทำเป็น 2 ฟอรม คือ ฟอรมสีน้ำเงิน และฟอรมสีเขียว มีลักษณะและการดำเนินการดังนี้

แบบสอบย่อย Word Reading

ผลการวิเคราะห์ (editorial review) ข้อสอบ 84 ข้อ ใน WRAT3 ฉบับสีน้ำเงิน และฉบับสีน้ำตาล (ฉบับละ 42 ข้อ) พิจารณาตัดออก 3 ข้อ จากฉบับสีน้ำตาล แล้วเพิ่มจำนวนข้อใน WRAT4 เป็น 55 ข้อ ทั้งสองฟอรม เป็นข้อสอบ คำใหม่ 29 ข้อ จัดไว้ในฟอรมสีน้ำเงิน 13 ข้อ ไว้ในฟอรมสีเขียว 16 ข้อ คำใหม่เหล่านี้ เลือกรจาก the EDL Core Vocabularies in Reading, Mathematics, Science and Social Science (Taylor, et al., 1989)¹ สุ่มคำจากหลายระดับชั้นเพื่อให้ได้ตัวแทนคำของระดับชั้น

แบบสอบย่อย Spelling

การวิเคราะห์ (editorial review) ข้อสอบ 80 ข้อ ใน WRAT3 ฉบับสีน้ำเงิน และฉบับสีน้ำตาล (ฉบับละ 40 ข้อ) พิจารณาตัดคำออก 3 ข้อ จากฉบับสีน้ำตาล และตัดคำออกอีก 3 ข้อ จากฉบับสีน้ำเงิน แล้วเพิ่มจำนวนข้อใน WRAT4 เป็น 45 ข้อ ทั้งสองฟอรม เป็นข้อสอบ คำใหม่ 16 ข้อ (ไว้ใน blue form 8 ข้อ ไว้ใน green form 8 ข้อ) คำใหม่เหล่านี้ เลือกรจาก the EDL Core Vocabularies in Reading, Mathematics, Science and Social Science (Taylor, et al., 1989) และ A Writing Vocabulary of Elementary Children (Hillerich, 1978)

¹ เป็นหนังสือรวบรวมคำที่ใช้ตรวจสอบความเหมาะสมของคำที่ใช้ในการพัฒนาข้อสอบของ ISAT และมีคำแนะนำสำหรับครูในการสร้างแบบประเมินในระดับโรงเรียนหรือท้องถิ่น

แบบสอบย่อย Math Computation

ผลการวิเคราะห์ (editorial review) ข้อสอบ 80 ข้อ ใน WRAT3 ฉบับสีน้ำเงิน และฉบับสีน้ำตาล (ฉบับละ 40 ข้อ) พิจารณาตัดข้อสอบออก 11 ข้อ จากฉบับสีน้ำตาล และตัดออกอีก 13 ข้อ จากฉบับสีน้ำเงิน จำนวนข้อใน WRAT4 คงไว้เหมือนเดิมคือ 40 ข้อ ทั้งสองฟอร์ม การสร้างข้อสอบใหม่แทนข้อที่ตัดทิ้งไปในแต่ละฟอร์ม สรุปในแบบสอบแต่ละฟอร์มมีรายละเอียดของเนื้อหา ดังตารางที่ 4.1

ตารางที่ 4.1 จำนวนข้อสอบและร้อยละจำแนกตามเนื้อหาในแบบสอบย่อย Math Computation

Content category	Blue form		Green form	
	Number of items	% of total items	Number of items	% of total items
Whole numbers	21	52.5	20	50.0
Addition, Subtraction	12	30.0	12	30.0
Multiplication, Division	7	17.5	7	17.5
Counting/ Rounding	2	5.0	1	2.5
Fractions	9	22.5	7	17.5
Comparing, Addition, subtraction	6	15.0	4	10.0
Multiplication, Division	3	7.5	3	7.5
Decimals	4	10.0	3	7.5
Comparing, Addition, Subtraction	1	2.5	1	2.5
Multiplication, Division	3	7.5	2	5.0
Percent	2	5.0	4	10.0
Algebra/ Statistics	4	10.0	6	15.0
Total	40	100.0	40	100.0

แบบสอบย่อย Sentence Comprehension

ขั้นตอนการพัฒนาข้อสอบ มีดังนี้

1. พัฒนาเนื้อหา (content development) ลักษณะข้อสอบ (item type) เป็นประโยคเดี่ยว หรือประโยคสั้น 2 ประโยค ที่เว้นคำสำคัญ (key word) การตัดสินใจว่าผู้สอบมีความเข้าใจในประโยคที่อ่านหรือไม่ ดูจากการเติมคำสำคัญในช่องว่างที่เว้นได้ถูกต้อง เมื่อได้เลือกแบบของข้อสอบ (item type) แล้ว ก็เขียนรายละเอียดของลักษณะจำเพาะของข้อสอบ (item specification) ให้เป็นแนว

ทางการสร้างข้อสอบ สรุปแนวทางสำคัญดังต่อไปนี้ หลีกเลี่ยงการใช้เนื้อหาที่เป็นความรู้เฉพาะที่อาจส่งผลให้ผู้สอบบางคนได้เปรียบเพราะมีความรู้เฉพาะนั้นๆ

2. ข้อสอบมีเนื้อหาที่ดึงมาจากด้านต่างๆ อาทิ ความรู้ทั่วไป สังคมวิทยา ธรรมชาติวิทยา ด้านมนุษยศาสตร์ (เช่น ศิลปะ วรรณคดี ดนตรี) งานอดิเรก กีฬา และ สารระควมบันเทิง ดินฟ้าอากาศ อาหาร การท่องเที่ยว

3. ข้อสอบต้องเหมาะกับผู้สอบทั้งเพศหญิงและชาย กลุ่มเชื้อชาติต่างๆ ไม่มีความลำเอียงในกลุ่มผู้สอบที่มีสังคมสิ่งแวดล้อมที่ต่างกัน

4. เนื้อหาในข้อสอบต้องไม่ผูกกับเงื่อนไขเวลา (timeless) คือไม่ล้าสมัยในช่วงเวลาสั้นๆ

5. ข้อสอบต้องหลีกเลี่ยงสิ่งที่จะเกิดผลกระทบทางลบ เช่น ประเด็นทางการเมืองหรือศาสนา และสถานการณ์ที่สร้างอารมณ์ทางลบที่รุนแรง

6. ข้อสอบมีระดับความยากที่สอดคล้องกับการกำหนดระดับชั้น โดยใช้ EDL Core Vocabularies in Reading, Mathematics, Science and Social Science (Taylor, et al., 1989)

7. ข้อสอบควรวัด part of speech ที่หลากหลาย

8. ข้อสอบควรมีตัวเฉลยหรือคำตอบที่ถูกต้องน้อยที่สุดเท่าที่จะทำได้

1.2 การเลือกข้อสอบเข้าสู่ชุดแบบสอบฉบับสุดท้าย

ทบทวนค่าลักษณะเฉพาะของข้อสอบ (item characteristics) ก่อนพิมพ์แบบสอบชุดสีน้ำเงิน และชุดสีเขียว ทั้งค่าสถิติแบบดั้งเดิม (classical) และแบบตอบสนองข้อสอบ 1 พารามิเตอร์ (IRT one parameter item calibration data) ซึ่งได้จากคะแนนทั้งหมดของกลุ่มตัวอย่างมาตรฐานบนพื้นฐานของการวิเคราะห์ข้อสอบ ในส่วนของแบบสอบย่อยชุด word reading และ Math computation ไม่เปลี่ยนแปลง สำหรับแบบสอบย่อย word spelling ตัดออกไปฟอร์มละ 3 ข้อ ชุดแบบสอบย่อย Sentence Comprehension ตัดข้อสอบออกไป 20 ข้อ เป็นของฟอร์มสีน้ำเงินและสีเขียว ฟอร์มละ 10 ข้อ สรุปจำนวนข้อในแต่ละ type of content ของแบบสอบย่อย Sentence Comprehension ในตารางที่ 4.2

ตารางที่ 4.2 ข้อสอบ WRAT4 Sentence Comprehension จำแนกตามประเภทของเนื้อหา ในแบบสอบฟอร์มสีน้ำเงินและสีเขียว

Type of content	Blue Form		Green Form	
	จำนวนข้อ	%	จำนวนข้อ	%
General Interest	8	16	14	28
Quantitative Comparison	5	10	2	4
Social Science	10	20	5	10
Natural Science	22	44	21	42
Humanities (Art, Music, Literature)	2	4	6	12
Sports/Recreation/Entertainment)	3	6	2	4
Total	50	100	50	100

1.3 การพัฒนามาตรฐาน

เป้าหมายสำคัญของการพัฒนาความเป็นมาตรฐานของ WRAT4 คือ การสร้าง norms ที่เป็นตัวแทนของสหรัฐอเมริกา รายกลุ่มอายุตั้งแต่ 5 – 94 ปี เป้าหมายข้อสอง คือ สร้าง norms จากนักเรียนชั้นอนุบาล จนถึงชั้นเรียน 12 กลุ่มตัวอย่างได้จากการสุ่มแบบแบ่งชั้นกำหนดจำนวน จากฐานข้อมูลประชากรของประเทศ (a stratified quota-based sampling procedure with proportionate allocation based on U.S. Census data) จำนวนตัวอย่างที่ใช้ในการคำนวณประมาณ 3,000 คน แต่ละระดับอายุใช้ 100-150 คน มีทั้ง 19 กลุ่มอายุ เป้าหมายของกระบวนการสุ่มตัวอย่าง คือ ต้องการตัวอย่างที่มีจำนวนผู้เข้าสอบมีสัดส่วนในตัวแปรเพศ เชื้อชาติ ระดับการศึกษา และถิ่นฐานตามเขตภูมิศาสตร์ ใกล้เคียงกับข้อมูลสำมะโนของประเทศที่ดำเนินการเมื่อเดือน มีนาคม 2001

1.4 การเทียบมาตรฐานคะแนนแบบสอบ 2 ฟอร์ม (Equating of Forms)

แบบสอบที่สร้างขึ้นในโปรแกรมนี้มี 2 ฟอร์ม เป้าหมายสำคัญคือ การทำให้แบบสอบทั้งสองฟอร์มมี norms ที่ให้ความหมายเทียบกันได้ในเวลาเดียวกัน (concurrent norms) ตามลักษณะของแบบสอบคู่ขนานที่ออกแบบไว้ การสร้าง norms ที่เท่าเทียมกัน กระทำโดยให้ผู้สอบรับการสอบจากฟอร์มใดฟอร์มหนึ่งโดยสุ่มในจำนวนที่เท่าๆ กันและทำให้เกิดความสมดุลของข้อมูล

การ equate แบบสอบ มีการออกแบบวิจัยหลายแบบต่างๆกัน โปรแกรมการพัฒนาแบบสอบมาตรฐาน WRAT นี้ได้เลือก แบบผู้สอบร่วม (Wright & Stoe, 1979) ที่กำหนดให้ผู้สอบทุกคนสอบฟอร์มสีเขียว หรือ สีน้ำเงินทำครบทุก subtest และทั้งสองกลุ่มจะทำอีก 1 subtest -ของอีกฟอร์มหนึ่งโดยสุ่ม การวิจัยโดยแบบผู้สอบร่วม (common person equating design) จะเกิดปฏิบัติการ 8 คู่ ดังนี้

1. ฟอร์มสีน้ำเงิน ทุกแบบสอบย่อย กับ ฟอร์มสีเขียว แบบสอบย่อย การอ่านคำ

2. รูปแบบสื่อน้ำเงิน ทุกแบบสอบย่อย กับ รูปแบบสีเขียว แบบสอบย่อย เข้าใจประโยค
3. รูปแบบสื่อน้ำเงิน ทุกแบบสอบย่อย กับ รูปแบบสีเขียว แบบสอบย่อย สะกดคำ
4. รูปแบบสื่อน้ำเงิน ทุกแบบสอบย่อย กับ รูปแบบสีเขียว แบบสอบย่อย การคำนวณ
5. รูปแบบสีเขียว ทุกแบบสอบย่อย กับ รูปแบบสื่อน้ำเงิน แบบสอบย่อย การอ่านคำ
6. รูปแบบสีเขียว ทุกแบบสอบย่อย กับ รูปแบบสื่อน้ำเงิน แบบสอบย่อย เข้าใจประโยค
7. รูปแบบสีเขียว ทุกแบบสอบย่อย กับ รูปแบบสื่อน้ำเงิน แบบสอบย่อย สะกดคำ
8. รูปแบบสีเขียว ทุกแบบสอบย่อย กับ รูปแบบสื่อน้ำเงิน แบบสอบย่อย การคำนวณ

แบบการวิจัย นี้ เป็น randomized blocked research design มีจุดเด่นหลายประการที่สำคัญ คือ

1. ใช้เวลาในการบริหารการสอบน้อยกว่าการใช้แบบวิจัยที่ให้ผู้สอบทำข้อสอบทั้งสองฟอร์ม และทุกแบบสอบย่อย และ
2. การจัดให้ทำสอบโดยสุ่มฟอร์มใดฟอร์มหนึ่ง สะดวกในการจัดการและควบคุม

1.5 ขั้นตอนการควบคุมคุณภาพ (Quality Control Procedures)

ผู้สร้างแบบสอบ ได้ตรวจกระดาษคำตอบอย่างถี่ถ้วน รอบคอบ ก่อนที่จะนำข้อมูลมาวิเคราะห์เพื่อสร้าง Norms ได้ออกแบบรหัสระบุเขตพื้นที่

การตรวจสอบความถูกต้องของแบบสอบย่อยความเข้าใจประโยค ตรวจโดยผู้สร้างแบบสอบเอง ในส่วนแบบสอบย่อยการสะกดคำ และ การคิดคำนวณ มีการตรวจสอบซ้ำอย่างรอบคอบ จากนั้นจึงบันทึกข้อมูลของผู้สอบรายบุคคลในแฟ้มข้อมูล แล้วทำการตรวจทานข้อมูล แก้ไขในกรณีที่มีความซ้ำซ้อนและความผิดพลาดใดๆ ที่เกิดจากการป้อนข้อมูล เมื่อข้อมูลถูกต้องเรียบร้อยแล้ว จึงดำเนินการวิเคราะห์ต่อไป

1.6 Item-Calibration

ใช้โปรแกรมคอมพิวเตอร์ของ Wright and Linacre's BIGSTEPS (1995) วิเคราะห์ตามหลัก one-parameter item-calibration โปรแกรมนี้ให้ผลการวิเคราะห์ค่าประมาณความยากของข้อสอบตาม latent trait model กระบวนการวิเคราะห์รายงานค่าประมาณที่เหมาะสมของข้อสอบรายข้อ ส่วนข้อสอบที่ไม่สามารถประเมินค่าประมาณที่เหมาะสมผู้วิจัยจะตัดออก ผลการวิเคราะห์ครั้งนี้พบว่า ตัดออก 26 ข้อ เป็นข้อสอบจากแบบทดสอบย่อย ชุดความเข้าใจประโยค 20 ข้อ จากแบบทดสอบย่อยชุดสะกดคำ 6 ข้อ จากนั้น ได้วิเคราะห์อีกครั้งกับข้อสอบที่เหลือ ผลการวิเคราะห์สุดท้ายแสดงค่าประมาณความยากของข้อสอบ (item -difficulty) และ ค่าประมาณความสามารถของผู้สอบรายบุคคล (person-ability) คะแนนของรายบุคคลในแต่ละแบบสอบย่อย เรียกว่า คะแนน W-Ability คำนวณโดย linear transformation ใช้สมการ $W\text{-Ability} = 9.1024 (\text{logit}) + 100$ คะแนน W-Ability จากแบบสอบพอร์มน้ำเงิน เป็นคะแนนที่ปรับเทียบกับคะแนนที่ได้จากพอร์มสีเขียวจากกระบวนการ

equate ตามหลักการที่แนะนำมาของ Wright & Stne (1979) ค่าความต่างเฉลี่ยของ 2 ฟอรั่ม นำมาปรับให้คะแนน W-Ability สุดท้าย ซึ่งทำให้ คะแนน W-Ability ของสองฟอรั่มในขั้นสุดท้ายที่นำไปใช้ มีความเท่าเทียมกัน

1.7 ความเที่ยงและความตรงของแบบสอบ WRAT4

หลักฐานความเที่ยง (Reliability) และความตรง (Validity) ของแบบสอบ ดำเนินการตามมาตรฐานแบบสอบทางการศึกษาและจิตวิทยา อ้างถึง American Educational Research Association et al., (1999)

เทคนิคการหาหลักฐานความเที่ยงที่เสนอ ประกอบด้วยสารสนเทศจากการวิเคราะห์ความเที่ยงตามทฤษฎีการวัดผลแบบดั้งเดิม (Classical test reliability theory) ได้แก่ ความคงที่ภายใน (internal consistency) ความคงเส้นคงวา (test-retest stability) ความคลาดเคลื่อนของการวัด (standard error of measurement) ช่วงความเชื่อมั่นของคะแนนมาตรฐาน (standard score confidence intervals) และค่าสถิติจากการประยุกต์ IRT

หลักฐานความตรงนำเสนอขึ้นการพัฒนาเนื้อหาแบบสอบและโครงสร้างของชุดแบบสอบ (structure of test battery) ค่าความสัมพันธ์ระหว่างแบบสอบ WRAT4 กับ แบบสอบวัดผลสัมฤทธิ์ที่มีการใช้อย่างแพร่หลาย และศึกษาการใช้ผลแบบสอบ WRAT4 กับกลุ่มผู้สอบที่มีลักษณะเฉพาะต่างๆ

ผลการวิเคราะห์ความเที่ยง

(1) ความคงที่ภายใน (internal Consistency Reliability)

ค่าความคงที่ภายในแสดงถึงความเป็นเอกภาพของข้อสอบภายในชุดที่ทำหน้าที่วัดความสามารถเดียวกับในครั้งที่น่าแบบสอบนั้นไปใช้ เป็นการวิเคราะห์จากผลการตอบข้อสอบแต่ละข้อของกลุ่มผู้สอบ Cronbach's (1951) Coefficient alpha เป็นวิธีการวิเคราะห์ความเที่ยงกับแบบวัดทั่วไป ชุดแบบสอบ WRAT4 ได้แสดงค่าสัมประสิทธิ์แอลฟาของแต่ละแบบสอบย่อย จัดแบ่งตามกลุ่มอายุ ได้ค่าสัมประสิทธิ์อยู่ในช่วง .87 ถึง .93 ทั้ง 4 แบบสอบย่อย สำหรับแบบสอบย่อย Math Computation มีค่าต่ำกว่าแบบสอบย่อยอื่นเล็กน้อย (ได้ .89 ในฉบับสีน้ำเงิน และ .87 ในฉบับสีเขียว) ค่ามัธยฐานของสัมประสิทธิ์ความเที่ยงของชุดสีน้ำเงินกับชุดสีเขียวคล้ายคลึงมาก ผู้พัฒนาได้สรุปว่า จากผลการวิเคราะห์แสดงถึงแบบสอบที่ออกแบบเป็นฟอรั่มสีน้ำเงินและสีเขียวมีค่าความคงที่สูง (Sufficient consistency) พอที่จะทำให้คะแนนผลการวัดรายบุคคลน่าเชื่อถือได้

(2) ความเที่ยงแบบ Alternate-form

การหาค่าความเที่ยง จากการหาความสัมพันธ์ระหว่างผลการวัดจากแบบวัด 2 ฉบับ มีสิ่งที่พึงระวังอย่างยิ่งคือ เวลาของการให้การทดสอบ ถ้าไม่สามารถจัดการได้เหมาะสมจะเกิดความคลาดเคลื่อนได้ การวิเคราะห์ของ WRAT4 จึงได้ออกแบบศึกษาผลกระทบต่อค่าความเที่ยง และแสดงผล

ในแต่ละแบบดังเช่น แบบ Alternate-form Reliability Immediate Retest แบบ Alternate-form Reliability-Delayed Retest โดยมีวัตถุประสงค์เพื่อศึกษาถึงตัวแปรกระทบจากการฝึก (practice effect) ที่เกิดขึ้นกับแต่ละระดับชั้น จากการศึกษาได้มีข้อสรุปว่า ผลกระทบจากการฝึก (practice effects) มีเพียงเล็กน้อย ถ้าให้ผู้สอบทำแบบสอบฟอร์มที่แตกต่างกันในเวลาประมาณ 30 วัน

(3) ความคลาดเคลื่อนมาตรฐานของการวัด

โดยธรรมชาติของคะแนนสอบใดๆ มีความคลาดเคลื่อนปะปนอยู่เสมอไม่มากก็น้อย ค่าความคลาดเคลื่อนมาตรฐานของการวัด (SEM) จะแสดงถึงความคลาดเคลื่อนที่เกิดขึ้นว่ามีมากน้อยเพียงไร จะเป็นประโยชน์ต่อการแปลผลคะแนนสอบรายบุคคล ผลการวิเคราะห์ค่าความคลาดเคลื่อนมาตรฐานการวัดของ 4 แบบทดสอบย่อยในชุดแบบสอบ WRAT4 พบว่ามีขนาดใกล้เคียงกันทุกแบบสอบ

(4) ค่าสถิติจาก Rasch IRT

Test Information Function :TIF เป็นการแสดงถึงความคมชัด (precision of measurement) ณ จุดต่างๆ ของเส้นความสามารถ (the latent trait continuum) TIF เป็นค่าที่เกิดจากสมการทางคณิตศาสตร์ที่มีความสัมพันธ์ของคะแนนบนมาตราวัดความสามารถของแบบสอบนั้นกับความสามารถของผู้สอบ แผนภาพ TIF ของแต่ละแบบทดสอบย่อยใน WRAT4 จะใช้ประมาณความคมชัดในการวัดตามคะแนนความสามารถของบุคคล (the person ability) ได้ ขอเสนอตัวอย่างที่ได้อธิบายในรายงานของ WRAT4

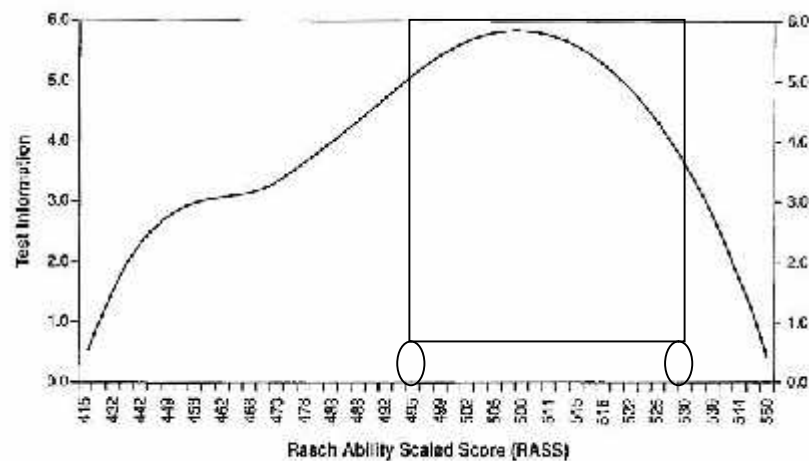


Figure 5.2. Test Information Function for Sentence Comprehension, Green Form.

จากตัวอย่าง TIF ของแบบทดสอบย่อย ความเข้าใจประโยค พบว่า เมื่อนำไปใช้กับคนที่มีความสามารถ (RASS) 488-525 คะแนน ที่บุคคลได้จะมีความถูกต้องสูง การพิจารณาความเที่ยงของแบบวัดจาก TIF จึงเป็นการให้สารสนเทศที่ชัดเจนมากขึ้น

ผลการวิเคราะห์ความตรง

ความตรงของแบบวัด เป็นคุณสมบัติที่แสดงถึงการวัดในสิ่งที่กำหนดไว้ในแบบวัดนั้น เป็นตัวชี้วัดคุณภาพที่สำคัญมาก ในชุดแบบสอบ WRAT4 ได้แสดงคุณสมบัตินี้ไว้ ด้วยหลักฐานต่างๆ เช่น

(1) หลักฐานความตรงภายใน (Internal Evidence of Validity) เป็นหลักฐานที่เกิดขึ้นในขั้นตอนแรก เมื่อวางแผนการพัฒนาเป็นกระบวนการที่มีความละเอียด รอบคอบ ตรวจสอบทบทวนหลายครั้งจนมีความชัดเจนในการกำหนดเนื้อหาที่สอดคล้องกับวัตถุประสงค์ของโครงการสอบ และคุณลักษณะที่ต้องการวัดในกลุ่มประชากรเป้าหมายที่กำหนด สรุปเป็นแผนกำหนดลักษณะของข้อสอบ

(2) หลักฐานของความตรงจากตัวแบบสอบ (Evidence Based on Test Content) เป็นหลักฐานที่เกิดจากการสร้างและพัฒนาแบบสอบ แสดงให้เห็นถึงแหล่งที่เกิดของเนื้อหาข้อสอบว่ามีความตรงในแบบวัดอย่างไร จุดสำคัญคือมีแหล่งที่มาอย่างถูกต้อง และมีกระบวนการตรวจสอบความถูกต้องด้วยกระบวนการที่น่าเชื่อถือ ตรวจสอบได้ ดังแสดงไว้ในขั้นตอนของการพัฒนา

(3) หลักฐานความตรงจากการตรวจสอบคุณลักษณะทางการวัด (Measurement Properties) ของแบบสอบย่อย คุณสมบัติทางการวัดผลที่สำคัญที่ต้องตรวจสอบชุดแบบสอบ WRAT4 คือ ความถูกต้องเที่ยงตรงในแต่ละแบบสอบย่อยในการวัดช่วงความสามารถที่กำหนด ตัวชี้วัดที่สำคัญ คือ ค่าสัมประสิทธิ์ความเที่ยง ซึ่งสรุปความคมชัดในการวัดกลุ่มความสามารถต่างๆ กลุ่มระดับชั้นต่างๆ ข้อมูลขั้นนี้ได้นำเสนอมาแล้วในหัวข้อสัมประสิทธิ์ความเที่ยงของแบบสอบย่อย โดยรวมๆ อยู่ที่ .80-.90 อยู่ในระดับที่น่าพอใจ นอกจากนี้การแสดงผลการวิเคราะห์ค่าสารสนเทศของแบบสอบ (TIF) ในแต่ละแบบสอบย่อยได้เพิ่มหลักฐานความชัดเจนของการวัดระดับความสามารถต่างๆ โดยทั่วไปในระดับความสามารถที่สูงและต่ำมาก มักพบว่ามีสารสนเทศจากแบบวัดที่น้อยกว่า ซึ่งจะนำไปสู่การอภิปรายหาแนวทางการเพิ่มเติม ข้อสอบ หรือปรับปรุงรูปแบบข้อสอบให้เหมาะสมยิ่งขึ้น

(4) หลักฐานภายนอกของความตรง (External Evidence of Validity) เป็นหลักฐานที่มาจากการศึกษาความสัมพันธ์ระหว่างแบบสอบย่อยใน WRAT4 กับแบบสอบอื่นๆ เป็นวิธีที่แนะนำให้ทำเมื่อพบว่าแบบสอบที่สร้างขึ้นมีแบบแผนของความสัมพันธ์ (Convergent หรือ divergent) กับแบบสอบหรือเครื่องมือวัดอื่นที่จะใช้เป็นเกณฑ์ เป็นความตรงประเภท Concurrent Validity การขยายการศึกษาเช่นนี้ต้องออกแบบลักษณะของกลุ่มตัวอย่างทั้งสองกลุ่มที่สอดคล้องกัน และ จัดเก็บข้อมูลด้วยแบบสอบแต่ละคู่ ค่าสหสัมพันธ์จะแสดงถึง คุณสมบัติของ WRAT 4 ในการอธิบายความสามารถในด้านที่ตรวจสอบว่าเป็นลักษณะใด WRAT4 ได้แสดงผลการศึกษา Concurrent Validity จากการใช้แบบสอบชนิดต่างๆ เป็นเกณฑ์ เช่น ศึกษาความสัมพันธ์กับ แบบสอบผลสัมฤทธิ์ทางการเรียน Wechsler Individual Achievement Test Second Edition (WIAT-II), Woodcock Johnson III Tests of Achievement (W-J III), the Kaufman Test of Educational Achievement : Second Edition –Comprehensive Form (KTEA-II) เป็นต้น ศึกษาความสัมพันธ์กับ แบบสอบที่วัดด้านพุทธิปัญญา เช่น the Wide Range

Intelligence Test (WRIT) , Wechler Intelligence Scale for Children Fourth Edition (WISC-IV), Stanford Binet Intelligence Scales, Fifth Edition (SB-5) เป็นต้น

(5) ศึกษาผลการใช้ WRAT4 กับกลุ่มบุคคลพิเศษที่มีความสามารถในการเรียนรู้ 2 กลุ่ม คือ กลุ่มความสามารถสูง และกลุ่มความสามารถต่ำ เปรียบเทียบผลการสอบในแต่ละแบบสอบย่อยกับกลุ่มควบคุม ด้วยสถิติทดสอบ t และดูขนาดของ effect ด้วยวิธีการทางสถิติที่แนะนำโดย Cohen (1988) ผลการศึกษาพบว่า ผู้มีความสามารถต่ำ 49 คน ได้คะแนนเฉลี่ยในแบบสอบย่อยแตกต่างน้อยกว่ากลุ่มควบคุมอย่างมีนัยสำคัญ ($p < .01$) และมี Effect size .74-.80 ส่วนการศึกษาในกลุ่มความสามารถสูง จำนวน 46 คน ได้คะแนนเฉลี่ยสูงกว่ากลุ่มควบคุมอย่างมีนัยสำคัญที่ระดับ .01 3 แบบสอบย่อย ยกเว้นแบบสอบย่อย Math Computation จากผลวิเคราะห์แสดงให้เห็นว่าแบบสอบ WRAT4 ใช้เพื่อระบุความสามารถของบุคคลที่มีความสามารถต่ำหรือบกพร่องในการเรียนรู้ได้

1.8 สรุปผลการพัฒนาแบบสอบ WRAT4

ผู้พัฒนาแบบสอบกล่าวว่า เป็นความรับผิดชอบที่ผู้พัฒนาต้องแสดงหลักฐานของการนำแบบสอบไปใช้ มีข้อมูลจำนวนมากทั้งที่เกี่ยวข้องกับเนื้อหาข้อสอบ (item content) การวิเคราะห์แบบต่างๆ ของความเที่ยง ความตรงภายใน ความตรงภายนอกและการวัดในกลุ่มต่างๆ และ การตรวจสอบข้อมูลที่ได้จากแหล่งต่างๆ

กล่าวโดยสรุป ชุดแบบสอบ WRAT4 มีผลการศึกษาที่อธิบายคุณภาพหลายแง่มุม ใช้ทฤษฎีการวัดความสามารถแบบดั้งเดิม (Classical Theory) และทฤษฎีการวัดความสามารถแฝง (Latent Trait Theory) เทคนิคการสร้างแบบสอบเพื่อให้ได้คุณลักษณะของความเป็นมาตรฐานของแบบสอบ ตรวจสอบคุณสมบัติด้วยหลักฐานกระบวนการอย่างมีเหตุผล และ ข้อมูลเชิงประจักษ์ ศึกษาคุณสมบัติด้านความตรงที่หลากหลายด้วยการออกแบบการวิจัยที่จะตอบคำถามความตรงแบบต่างๆ และ สร้างคู่มือสำหรับการนำแบบสอบ WRAT4 ไปใช้เพื่อการแปลผลคะแนนที่ถูกต้องอย่างมีมาตรฐาน นับว่าเป็นแบบสอบที่ใช้เวลามาก ไม่น้อยกว่า 70 ปี และ พัฒนาอย่างต่อเนื่อง จนสามารถใช้แบบสอบเพื่อพัฒนาคุณภาพการวัดกลุ่มบุคคลต่างๆ เสนอแนวทางการใช้ผลการวัดเพื่อพัฒนาการเรียนการสอน และ อื่นๆ อีกมาก นับเป็นงานที่สร้างเสริมคุณค่าทางวิชาการและเป็นแหล่งเรียนรู้ที่มีค่าอย่างยิ่ง

ตอนที่ 2 การออกแบบชุดแบบสอบ (Test battery) ที่มีความยากต่างกัน

การนำเสนอสาระในตอนที่ 2 นี้ สรุปจากแนวคิดทฤษฎีที่ Michael J. Kolen และ Robert L. Brennan เขียนไว้ในตำราชื่อ Test Equating, Scaling, and Linking :Methods and Practices . (Kolen, Michael J.-2nd ed. 2004, pp.329-417) มี 3 หัวข้อดังต่อไปนี้

2.1 ลักษณะชุดแบบสอบ (Test batteries)

ชุดแบบสอบมี 2 ลักษณะ ลักษณะที่หนึ่ง เป็นชุดแบบสอบย่อยจำนวนหนึ่งที่มีความสัมพันธ์ซึ่งกันและกันอย่างเป็นระบบ (system of interrelated tests) วัดความสามารถหรือเนื้อหาหลายด้าน (across subject areas) การคิดคะแนนแต่ละด้านหรือแต่ละแบบสอบย่อยแยกจากกัน อาจจะมีการคิดคะแนนรวมทุกแบบสอบย่อย (composite score) ตามที่ออกแบบในชุดแบบทดสอบ กระบวนการสร้างแบบสอบ (test construction process) การให้คะแนน (scaling) และการสร้างเกณฑ์มาตรฐานของการแปลคะแนน (norming) จะดำเนินการอย่างเดียวกันทุกแบบสอบย่อยในชุดแบบสอบนั้น คะแนนสอบของผู้สอบรายบุคคล สามารถคำนวณและเปรียบเทียบระหว่างกันแบบสอบย่อยเดียวกัน เมื่อได้กำหนดเกณฑ์ (norms) แปลความหมายคะแนนของชุดแบบสอบ การสร้างเกณฑ์คะแนน (norms) พัฒนาจากกลุ่มผู้สอบกลุ่มเดียวกันในทุกแบบสอบย่อย ลักษณะที่สอง ชุดแบบสอบที่สร้างขึ้นเพื่อแปลความหมายคะแนนอธิบายความสามารถระดับชั้น โดยมีเกณฑ์ปกติที่สร้างเพื่อการแปลความหมายเปรียบเทียบกลุ่มความสามารถหลายๆ ระดับโดยตรง ตัวอย่างที่ใช้มากในสหรัฐอเมริกา ได้แก่ชุดแบบสอบวัดผลสัมฤทธิ์ทางการเรียน ระดับอนุบาลถึงมัธยมศึกษาตอนปลาย (K-12) กระบวนการสร้างแบบสอบและเกณฑ์ปกติเพื่อการแปลความหมายตามการออกแบบสร้างชุดแบบสอบเป็นระเบียบวิธีการสร้างแบบสอบมาตรฐาน ซึ่งยึด Standards for educational and psychological testing ของ American Educational Research Association, American Psychological Association. & National Council on Measurement in Education (1999) ใช้ข้อมูลเชิงปริมาณที่ต้องออกแบบการเก็บข้อมูลจากกลุ่มประชากรและขนาดตัวอย่างที่สอดคล้องกับระเบียบวิธีเชิงปริมาณ และทฤษฎีที่ใช้สร้างมาตรฐานวัดความสามารถจากคะแนนสอบที่ได้พัฒนาบนเงื่อนไขของกระบวนการต่างๆ จนสำเร็จเป็นแบบวัดมาตรฐาน Michael J. Kolen และ Robert L. Brennan ได้บรรยายไว้ในบทที่ 9 ว่าด้วย Score Scales (p.329-417) ผู้วิจัยจะนำเสนอเพื่อประกอบการพัฒนาแบบวัด KBAST ครั้งนี้ต่อไป

2.2 การสร้างชุดแบบสอบเพื่อสอบกลุ่มบุคคลที่มีความสามารถหลายระดับ

แบบสอบที่วัดคุณสมบัติอย่างเดียวกันกับกลุ่มบุคคลที่มีความสามารถแตกต่างกัน คะแนนที่ได้ถ้าสามารถแปลความหมายให้เห็นระดับความสามารถที่แตกต่างกัน โดยเฉพาะแบบสอบที่ใช้ในการประเมินระดับการเรียนรู้ตามหลักสูตรการศึกษา ก็จะสะท้อนถึงความสามารถของแบบสอบในด้านความตรงตามเป้าหมายโครงสร้างที่ได้ออกแบบไว้ ความแตกต่างที่อธิบายได้จะให้ความหมายถึงระดับความเจริญเติบโต หรือพัฒนาการความสามารถในด้านนั้นๆ ของบุคคล Kolen และ Brennan (p.376) ให้นิยามของ Growth ว่า เป็นผลการสอบที่แสดงถึงความสามารถในเนื้อหาหรือวิชาที่มีการจัดการเรียนการสอนที่สัมพันธ์กับการเรียนรู้ที่เพิ่มขึ้นทุกชั้นปี เรียกว่าเป็น grade to grade growth

การออกแบบให้ได้แบบสอบที่มีคุณสมบัติตามนิยามที่กล่าวมา ที่สำคัญคือ ความครอบคลุมและสอดคล้องของเนื้อหาของข้อสอบในวิชานั้นๆ ทุกระดับชั้น เพื่อวัดความสามารถของผู้สอบแต่ละ

ระดับชั้น เมื่อมีการนำข้อสอบไปใช้จะนำคะแนนดิบมาคำนวณแล้วแปลงเป็นคะแนนมาตรฐานที่กำหนด (score scale) และรายงานผลการสอบ คะแนนที่รายงานจะเปรียบเทียบว่าใครมีความสามารถสูงต่ำเพียงไร แบบสอบที่จะบอกพัฒนาการระดับชั้นหนึ่งไปต่ออีกระดับชั้นหนึ่ง จะเป็นข้อสอบที่ยาวมาก และมีความยากที่ผู้สอบที่อยู่ระดับชั้นที่ต่ำมากจะทำได้ ขณะที่ผู้สอบที่อยู่ระดับชั้นที่สูง ก็จะต้องเสียเวลาทำข้อสอบที่ง่ายเกินไป จึงจำเป็นต้องออกแบบสร้างเป็นชุดแบบสอบและสร้างคะแนนมาตรฐานวัดความสามารถที่ต้องสัมพันธ์ และให้ความหมายที่ถูกต้อง Kolen และ Brennan จึงได้แนะนำการสร้างและการออกแบบการเก็บข้อมูลด้วยการสร้างมาตรฐานวัดความสามารถแบบแนวตั้ง (vertical Scaling) หลายรูปแบบ รายงานนี้จะนำเสนอการออกแบบที่ใช้ข้อสอบร่วม (p.377-381)

2.3 การสร้างมาตรฐานแบบสอบแนวตั้ง (vertical scoring)

ระบบการสร้างมาตรฐานแบบสอบแนวตั้ง (vertical scoring) เป็นการสร้างมาตรฐานจากแบบสอบย่อยที่วัดในสิ่งเดียวกันแต่มีความยากแตกต่างกัน พบว่ามีการนำไปใช้กับแบบสอบวัดผลสัมฤทธิ์ในระดับประถมศึกษา ในการประเมินความถนัดทางวิชาการ หรือผลสัมฤทธิ์ทางการเรียนของนักเรียน ผลการประเมินจะให้สารสนเทศที่สำคัญในการอธิบายพัฒนาการของนักเรียนจากปีหนึ่งไปอีกปีหนึ่ง และต่อไปจนจบหลักสูตร การอธิบายความเจริญหรือพัฒนาการของนักเรียนน่าจะต้องเกิดจากการใช้แบบสอบเดียวกัน แต่ต่างระดับความยากตามปีที่เรียน คะแนนสอบสามารถนำมาสร้างแผนภาพของการพัฒนาปีต่อปี หลายๆ ปีจนจบหลักสูตร

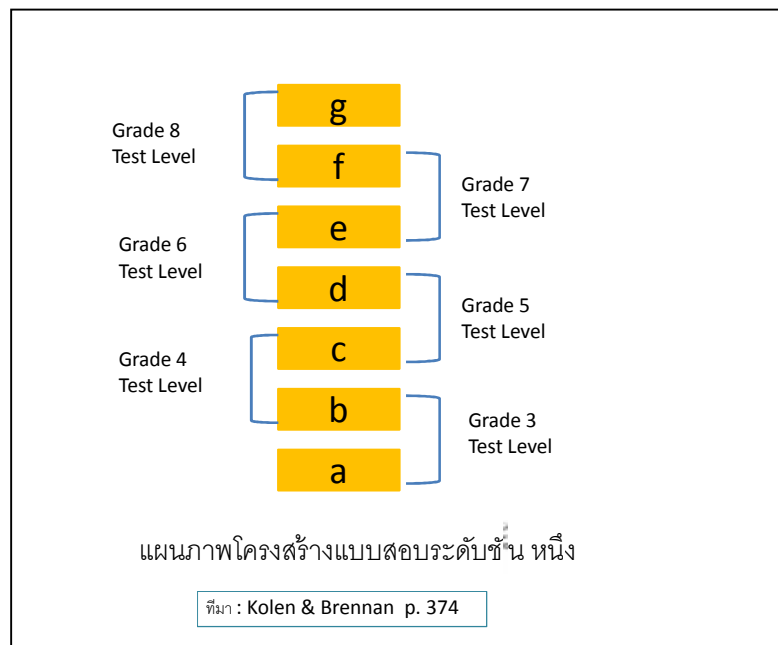
การสร้างชุดแบบสอบเพื่อประเมินพัฒนาการของนักเรียนแบบ grade to grade growth เป็นการสร้างชุดแบบสอบย่อยที่วัดเนื้อหาที่มีความยากต่างระดับกัน (multiple test levels) แต่ละแบบสอบจะออกแบบสร้างเพื่อวัดผู้สอบในระดับที่เหมาะสม ปกติจะกำหนดด้วยระดับชั้น หรือ อายุ การใช้ผลการวัดเพื่อบอกระดับความสามารถจะสร้างระบบคะแนนให้มีความหมายที่สื่อถึงพัฒนาการ กระบวนการสร้างมาตรฐานลักษณะนี้ เรียกว่า การสร้างมาตรฐานแนวตั้ง (vertical scaling)

ตัวอย่างการใช้มาตรฐานแนวตั้ง (vertical scaling) พบในแบบสอบผลสัมฤทธิ์ในระดับประถมศึกษา เช่น ITBS (Hoover et al., 2003) ในชุดแบบสอบความถนัด เช่น The Cognitive Abilities Test (CogAT) แบบสอบเหล่านี้จะมีแบบสอบย่อยวัดด้านต่างๆ และใช้กับนักเรียนหลายระดับชั้น

โครงสร้างของชุดแบบสอบวัดผลสัมฤทธิ์ที่ใช้สอบนักเรียนแต่ละระดับชั้น ในแบบสอบย่อยจะมีคำถามที่มีเนื้อหาของระดับชั้นนั้น โดยมีระดับการวัดที่ขยับจากระดับชั้นก่อนชั้นที่สอบ คำถามจะยากขึ้น ถ้ามคำถามที่ลึกลับมากขึ้น ข้อคำถามในแต่ละระดับมีความซ้ำกันอยู่ ทั้งนี้เพราะหลักสูตรการสอนปกติจะมีเนื้อหาที่ซ้ำ บางครั้งใช้ข้อสอบซ้ำ เพื่อใช้เป็นตัวกลางเชื่อมแบบสอบต่างระดับกัน ตัวอย่างในรูปภาพ 9.6 เป็นโครงสร้างของชุดแบบสอบหนึ่งที่วัดระดับชั้นที่ 3 ถึง 8 มีข้อสอบอยู่ 7 บล็อก ให้ข้อบล็อก a ถึง g ข้อสอบบล็อก a และ b เป็นส่วนของแบบสอบระดับชั้นที่ 3

บล็อก b, c เป็นส่วนของแบบสอบระดับชั้นที่ 4
 บล็อก c,d เป็นส่วนของแบบสอบระดับชั้นที่ 5
 บล็อก d, e เป็นส่วนของแบบสอบระดับชั้นที่ 6
 บล็อก e,f เป็นส่วนของแบบสอบระดับชั้นที่ 7
 บล็อก f, g เป็นส่วนของแบบสอบระดับชั้นที่ 8

โครงสร้างแบบนี้ จะมีข้อสอบร่วม การสร้างชุดแบบสอบ จะสร้าง 2 ฟอรัม เพื่อให้ผู้สอบคนเดียวสอบครั้งต่อไป ได้รับการทดสอบด้วยข้อสอบที่ต่างออกไป ดังแผนภาพที่ 4.1 (Kolen & Brennan, 2010: 374)



แผนภาพที่ 4.1 โครงสร้างแบบสอบระดับชั้น หนึ่ง

2.4 การออกแบบเก็บข้อมูลเพื่อสร้างมาตรฐานแบบสอบแนวตั้ง (vertical scoring)

จากการออกแบบโครงสร้างให้มีความซ้ำในระดับชั้นที่ติดกันดังที่กล่าวมาในตัวอย่างของแบบสอบสัมฤทธิ์ผลทางการเรียน จะได้แบบใช้ข้อสอบร่วม (common items) ในรอยต่อของแต่ละระดับ เพื่อใช้ระเบียบวิธีการเทียบมาตรวัดของแบบสอบในชุด (battery) มาตรวจสอบความตรงของชุดแบบสอบว่าเป็นไปตามโครงสร้างความเพิ่มพูนของความสามารถตามช่วงชั้นหรือไม่ เทคนิคนี้เป็นการเทียบมาตราคะแนนแนวตั้ง (Vertical Equating) ที่มีข้อสอบร่วม (Common Items) เป็นตัวเชื่อมในการวิเคราะห์ผลการตอบจากกลุ่มตัวอย่างระดับชั้นที่ต่างกัน ดังแผนภาพที่ 4.2 (Kolen & Brennan, 2010: 378)

ระดับชั้น	Item block						
	a	b	c	d	e	f	g
3	a	b	-	-	-	-	-
4	-	b	c	-	-	-	-
5	-	-	c	d	-	-	-
6	-	-	-	d	e	-	-
7	-	-	-	-	e	f	-
8	-	-	-	-	-	f	g

Illustration of a common-item design

ที่มา : Kolen & Brennan p. 378

แผนภาพที่ 4.2 การออกแบบการใช้ข้อสอบร่วม

จากรูปแบบโครงสร้างที่กำหนดนี้ ข้อสอบกลุ่ม a ที่คัดมาจากคลังข้อสอบ a จะมีระดับความยากน้อยกว่าข้อสอบกลุ่ม c ทำนองเดียวกันข้อสอบกลุ่ม c จะมีระดับความยากน้อยกว่าข้อสอบกลุ่ม e และข้อสอบกลุ่ม e ก็มีระดับความยากน้อยกว่าข้อสอบกลุ่ม g ส่วนข้อสอบกลุ่ม b d และ f เป็นข้อสอบร่วมที่อยู่ใน 2 ฉบับ เพื่อให้สารสนเทศเชื่อมต่อแบบสอบ 2 ฉบับ ด้วยกระบวนการวิเคราะห์ของโปรแกรม BILOG จะทำให้ค่าประมาณความยากรายข้อ ในแบบสอบย่อยทักษะนั้นมีการเชื่อมโยงอยู่ในมาตรวัดเดียวกัน ทำให้ค่าประมาณความสามารถของผู้สอบจากการสอบแบบสอบต่างฉบับสามารถนำมาเปรียบเทียบกันได้

ระบบการสร้างมาตรคะแนนแบบสอบแนวตั้ง (vertical scoring) เป็นการสร้างมาตรคะแนนจากแบบสอบย่อยที่วัดในสิ่งเดียวกันแต่มีความยากแตกต่างกัน พบว่ามีการนำไปใช้กับแบบสอบวัดผลสัมฤทธิ์ในระดับประถมศึกษา

ในการประเมินความถนัดทางวิชาการ หรือ ผลสัมฤทธิ์ทางการเรียนของนักเรียน ผลการประเมินจะให้สารสนเทศที่สำคัญในการอธิบายพัฒนาการของนักเรียนจากปีหนึ่งไปอีกปีหนึ่ง และต่อไปจนจบหลักสูตร การอธิบายความเจริญหรือพัฒนาการของนักเรียนน่าจะต้องเกิดจากการใช้แบบสอบเดียวกัน แต่ต่างระดับความยากตามปีที่เรียน คะแนนสอบสามารถนำมาสร้างแผนภาพของการพัฒนาปีต่อปี หลายๆ ปีจนจบหลักสูตร

การคิดคะแนนแบบสอบ (Test scoring)

การคิดคะแนนแบบสอบที่วัดระดับความสามารถในแนวตั้ง ปกติคิด 2 ชั้น ชั้นแรกคิดคะแนนดิบตามจำนวนข้อที่ทำถูกต้อง ชั้นสองแปลงคะแนนดิบเป็นมาตราคะแนนของแบบสอบนั้น สำหรับแบบสอบที่ใช้ทฤษฎี IRT ผลการตอบข้อสอบของแต่ละคนจะวิเคราะห์ผ่านโปรแกรมการประมาณค่าความสามารถ () ของผู้สอบ การแปลงคะแนนดิบเป็นคะแนนตามมาตราของแบบสอบ มีรูปแบบการแปลงเชิงเส้นและแบบอื่น ในที่นี้จะสรุปวิธีการที่ใช้ค่าสถิติของ IRT เมื่อออกแบบโครงสร้างแบบสอบที่ใช้ข้อสอบร่วม (Common-item design) (Kolen & Brennan, 2010: 387-388)

ข้อมูลที่ใช้สร้างมาตราคะแนนแบบ IRT เป็นข้อมูลการตอบข้อสอบรายข้อในชุดแบบสอบของกลุ่มตัวอย่างนักเรียนที่กำหนด ข้อมูลจะผ่านการประมาณค่าโดยโปรแกรมของ IRT ข้อสอบที่เป็นข้อสอบร่วมจะมีคะแนนในสองระดับชั้นเพื่อทำหน้าที่เชื่อมมาตราคะแนนของแบบสอบสองระดับชั้นนั้นให้เป็นมาตราคะแนนเดียวกัน การวิเคราะห์จะทำที่ละแบบสอบย่อย หรือวิเคราะห์สองระดับไปในเวลาเดียวกันก็ได้ วิธีการที่ตามมาก็จะแตกต่างกัน

ในกรณีที่เราวิเคราะห์หนึ่งแบบสอบระดับชั้น ค่าพารามิเตอร์ที่ได้จะเป็นค่าในแบบสอบระดับนั้นๆ และถือว่ามาตราคะแนนความสามารถ () จะวางฐานไว้ที่แบบสอบระดับนั้น ชั้นต่อไปจะนำค่าพารามิเตอร์ของข้อสอบร่วม ใส่เป็นค่ากำหนดในแบบสอบระดับชั้นที่มีข้อสอบร่วมชุดเดียวกัน แล้วทำการวิเคราะห์ จะได้ค่าประมาณความสามารถของผู้ที่สอบแบบสอบระดับชั้นที่สอง ค่าประมาณความสามารถของกลุ่มผู้สอบที่สองจะกระจายบนมาตราคะแนนฐานเดียวกัน แต่อย่างไรก็ตามอาจพบว่ามีพฤติกรรมตอบข้อสอบร่วมบางข้อมีลักษณะที่แตกต่างกับการเชื่อมระหว่างระดับชั้น ก็จำเป็นต้องตัดออก ทั้งนี้ ควรใช้เส้นโค้งแสดงภาพลักษณะเฉพาะของข้อสอบ (item characteristics curve: ICC) จากการวิเคราะห์มาประกอบการตัดสินใจ ขั้นตอนที่จะแนะนำในกระบวนการวิเคราะห์ เมื่อใช้ตัวอย่างการออกแบบในรูปภาพที่ 2 มีดังนี้

1. วิเคราะห์ผลการตอบแบบสอบระดับชั้น 3 ถึง 8 แยกวิเคราะห์ หาค่าประมาณคุณลักษณะข้อสอบ (item parameters) และการแจกแจงความสามารถ (ability distribution)

2. ตั้งข้อกำหนดให้ใช้มาตราความสามารถ (the -scale) ของแบบสอบระดับชั้น 3 เป็นฐาน ใช้ข้อสอบร่วมที่อยู่ในแบบสอบระดับชั้น 3 กับแบบสอบระดับชั้น 4 แปลงคะแนนสอบของผู้สอบแบบสอบระดับชั้น 4 ด้วยวิธีการเชิงเส้น (ด้วยการใช้วิธีทดสอบเส้นโค้ง ค่าพารามิเตอร์ของข้อสอบ) ข้อสอบที่ปรากฏค่าที่ผิดเพี้ยนจากระดับชั้นทั้งสองให้ตัดออกจากกลุ่มข้อสอบร่วม ผลการแปลงจะนำเข้ามาประมาณค่าพารามิเตอร์ข้อสอบ และค่าประมาณความสามารถของผู้สอบแบบสอบระดับชั้น 4

3. ชั้นต่อไปเป็นการปรับมาตราคะแนนความสามารถของผู้ที่สอบแบบสอบระดับชั้น 5 6 7 และ 8 ตามลำดับ ด้วยกระบวนการทำนองเดียวกับที่ปรับคะแนนความสามารถ 4 ให้อยู่บนเส้นฐานมาตราความสามารถของระดับชั้น 3 เมื่อกระบวนการเสร็จสิ้นลง จะทำให้คะแนนจากชุดแบบสอบทั้งหมด 6

ฉบับมีมาตรฐานวัดความสามารถเดียวกัน การแปลความหมายผลการสอบจะให้สารสนเทศรายชั้น และตลอดทั้ง 6 ชั้นเรียน

2.6 กลุ่มตัวอย่างมาตรฐาน

แบบสอบมาตรฐานเป็น norm-referenced tests เป็นเครื่องมือที่ให้ความหมายของผลการวัดอิงกลุ่มประชากรที่กำหนด จึงต้องใช้กลุ่มตัวอย่างประชากรที่กำหนดไว้แต่แรก (Normative sample) ผลการสอบจะนำไปพัฒนาระบบคะแนน และสร้างตารางคะแนน norm

2.7 กำหนดระดับเส้นฐานความสามารถของผู้สอบ

โดยปกติ แบบสอบที่เป็นชุด (batteries) จะประกอบด้วยจำนวนข้อสอบค่อนข้างมาก เมื่อมีการจัดวางข้อสอบเป็นลำดับตามข้อง่ายไปสู่ข้อยาก มีการกำหนดกฎการข้อยุติการสอบเมื่อผู้สอบทำผิดติดต่อกัน 3 ข้อ หรือ 5 ข้อ หรือ 7 ข้อ จะเป็นเหตุผลของการยุติการสอบ ถือว่าเป็นจุดที่ชี้ระดับเส้นฐานความสามารถของผู้สอบ (A basal score) ผู้สอบไม่ต้องทำข้อที่เกินความสามารถ เพราะอาจเกิดการเดา การกำหนดจำนวนข้อที่ผิดติดต่อกันเป็นเหตุผลของผู้สร้างแบบสอบ ในแบบสอบ WRAT4 ใช้กฎการยุติการสอบเมื่อเกิดการทำผิดติดต่อกัน 5 ข้อในแบบสอบย่อยความเข้าใจประโยค และ 7 ข้อ ในแบบสอบย่อยการคำนวณทางคณิตศาสตร์

ตอนที่ 3 การตรวจสอบคุณสมบัติของแบบสอบ

การตรวจสอบคุณสมบัติแบบสอบ เป็นข้อกำหนดที่สำคัญของการพัฒนาแบบสอบมาตรฐาน ที่ใช้กำหนดตาม *Standards for Educational and Psychological Testing*. (American Educational Research Association (AERA), American Psychological Association (APA), & National Council on Measurement in Education (NCME): 1999) ที่สำคัญ ได้แก่

ความเที่ยง (reliability) มีหลายแบบ ที่ใช้กันมาก ได้แก่ Internal consistency reliability แสดงถึงข้อสอบแต่ละข้อในแบบสอบทำหน้าที่วัดความสามารถเดียวกันในแบบสอบฉบับที่นำไปดำเนินการสอบครั้งนั้นๆ การประมาณค่าความเที่ยงแบบความคงที่ภายในนี้ วิเคราะห์จากผลการตอบข้อสอบแต่ละข้อของผู้สอบ Cronbach's (1951) coefficient alpha เป็นวิธีที่ใช้ประมาณค่าความคงที่ ได้กับการวัดทั่วไป แบบวัดบุคลิกภาพ และมาตราประมาณค่า

ความตรง (validity)

(1) หลักฐานความตรงภายใน (Internal Evidence of Validity) เป็นหลักฐานที่เกิดขึ้นในขั้นตอนแรก เมื่อวางแผนการพัฒนาเป็นกระบวนการที่มีความละเอียด รอบคอบ ตรวจสอบทบทวนหลาย

ครั้งจนมีความชัดเจนในการกำหนดเนื้อหาที่สอดคล้องกับวัตถุประสงค์ของโครงการสอบ และคุณลักษณะที่ต้องการวัดในกลุ่มประชากรเป้าหมายที่กำหนด สรุปเป็นแผนกำหนดลักษณะของข้อสอบ

(2) หลักฐานของความตรงจากตัวแบบสอบ (Evidence Based on Test Content) เป็นหลักฐานที่เกิดจากการสร้างและพัฒนาแบบสอบ แสดงให้เห็นถึงแหล่งที่เกิดของเนื้อหาข้อสอบว่ามีความตรงในแบบวัดอย่างไร จุดสำคัญคือมีแหล่งที่มาอย่างถูกต้อง และมีกระบวนการตรวจสอบความถูกต้องด้วยกระบวนการที่น่าเชื่อถือ ตรวจสอบได้ ดังแสดงไว้ในขั้นตอนของการพัฒนา

(3) หลักฐานภายนอกของความตรง (External Evidence of Validity) เป็นหลักฐานที่มาจาก การหาความสัมพันธ์ระหว่างแบบสอบกับแบบสอบอื่นๆ ที่มีแบบแผนของความสัมพันธ์ (convergent หรือ divergent) กับแบบสอบหรือเครื่องมือวัดอื่น เป็นความตรงประเภท Concurrent Validity การขยายการศึกษาเช่นนี้ ต้องออกแบบลักษณะของกลุ่มตัวอย่างทั้งสองกลุ่มที่สอดคล้องกัน และจัดเก็บข้อมูลด้วยแบบสอบแต่ละคู่ ค่าสหสัมพันธ์จะแสดงถึงคุณสมบัติของแบบสอบ ในการอธิบายความสามารถในด้านที่ตรวจสอบว่าเป็นลักษณะใด

ตอนที่ 4 สรุปกรอบแนวคิดการพัฒนาแบบสอบทักษะพื้นฐานทางวิชาการ KBAST

จากการทบทวนแนวคิด ทฤษฎี และแนวทางปฏิบัติตั้งที่รายงานมา จึงได้กำหนดเป็นกรอบพัฒนาแบบสอบมาตรฐาน KBAST ให้เป็นแบบสอบวัดทักษะพื้นฐานทางวิชาการ 4 ทักษะ คือ ด้านการอ่านคำ (Word Reading) ด้านการสะกดคำ (Word Spelling) ด้านความเข้าใจประโยค (Sentence Comprehension) และด้านการคิดคำนวณทางคณิตศาสตร์ (Math Computation) เพื่อการประเมินและวินิจฉัยบุคคลที่มีภาวะสมาธิสั้น บกพร่องทางการเรียนรู้ และออทิซึม ดังนี้

1. การพัฒนาเนื้อหาแบบสอบ KBAST ให้ตรงวัตถุประสงค์และถูกต้องตามหลักวิชาการ โดยกำหนดนิยามและลักษณะของการวัดโดยผู้เชี่ยวชาญ

2. สร้างข้อสอบที่มีคุณภาพ โดยการตรวจสอบด้วยหลักเหตุผล และข้อมูลเชิงประจักษ์ ใช้เทคนิคการวิเคราะห์ตามทฤษฎี IRT ในการคัดเลือกข้อสอบตามเกณฑ์คุณภาพ

3. ออกแบบชุดแบบสอบที่ประกอบด้วยแบบสอบย่อย 4 ทักษะตามเป้าหมายของแบบสอบ แต่ละทักษะจัดแบ่งเป็นฉบับตามระดับช่วงชั้น คือ ฉบับที่ 1 มีความยากของข้อสอบตามหลักสูตรช่วงชั้นที่ 1 (ประถมศึกษาปีที่ 1 ถึง 3) ฉบับที่ 2 มีความยากตามช่วงชั้นที่ 2 (ประถมศึกษาปีที่ 4 ถึง 6) ฉบับที่ 3 มีความยากตามช่วงชั้นที่ 3 (มัธยมศึกษาปีที่ 1 ถึง 3) และฉบับที่ 4 มีความยากตามช่วงชั้นที่ 4 (มัธยมศึกษาปีที่ 4 ถึง 6) การออกแบบนี้กำหนดโครงสร้างของชุดแบบสอบแบบ Grade to Grade Growth Model

4. คัดเลือกข้อสอบที่มีคุณภาพตรงตามข้อกำหนดเฉพาะของแบบสอบ จัดเรียงข้อสอบภายในชุดตามลำดับค่าความยากของข้อสอบ (b) ตรวจสอบความถูกต้องเหมาะสมด้วยหลักเหตุผล

5. สร้างระบบคะแนนและมาตราคะแนนแบบสอบด้วยวิธีสถิติ IRT และเทคนิคการเทียบมาตราคะแนนแบบ vertical equating technique
6. กลุ่มตัวอย่างผู้สอบได้จากการสุ่มให้ตัวแทนของนักเรียนชั้นประถมศึกษาปีที่ 1 ถึงมัธยมศึกษาปีที่ 6 จาก กรุงเทพมหานคร และจังหวัดขอนแก่น เป็นกลุ่มตัวอย่างพัฒนามาตราคะแนน และสร้างตารางเกณฑ์มาตรฐาน
7. แปลงคะแนนดิบเป็นคะแนนมาตรฐาน มีค่าเฉลี่ยเท่ากับ 100 และมีค่าเบี่ยงเบนมาตรฐานเท่ากับ 15 และสร้างตารางเกณฑ์มาตรฐานคะแนน (norms) เป็นเกณฑ์ระดับชั้น และช่วงชั้น แยกรายทักษะย่อยทั้ง 4 ทักษะ
8. แสดงคุณภาพของความตรงด้วยการบรรยายกระบวนการที่สร้างความตรงเชิงเนื้อหา (content validity)
9. แสดงความตรงเชิงโครงสร้างด้วยค่าสถิติของค่าประมาณความสามารถของกลุ่มผู้สอบแบบสอบย่อย และทดสอบความแตกต่างของค่าเฉลี่ยความสามารถระหว่างกลุ่มที่ทำแบบสอบต่างระดับช่วงชั้นด้วยสถิติทดสอบ the Kruskal-Wallis test
10. แสดงคุณภาพด้านความเที่ยงด้วยค่าสัมประสิทธิ์ความคงที่ภายใน และความคลาดเคลื่อนมาตรฐานของการวัด
11. กำหนดกฎการยุติการสอบเมื่อผู้สอบทำผิดติดต่อกัน 5 ข้อ ด้วยเหตุผลของการใช้แบบสอบกับนักเรียนที่มีความสามารถต่างกัน โดยเฉพาะการใช้กับนักเรียนที่มีภาวะบกพร่องทางการเรียนรู้ สมาธิสั้น และออทิซึม การนับคะแนนข้อที่อยู่เหนือจุดยุติจะถือว่าผิด เพราะเป็นข้อสอบที่อยู่เหนือความสามารถของผู้สอบนั้น
12. สร้างคู่มือสำหรับการใช้ชุดแบบสอบ KBAST เพื่อให้มีการนำไปใช้ได้ถูกต้องแบบข้อกำหนดของแบบสอบมาตรฐาน

ระเบียบวิธีดำเนินการวิจัย

การพัฒนาแบบสอบทักษะพื้นฐานทางวิชาการ KBAST เพื่อการประเมินและวินิจฉัยบุคคลที่มีภาวะสมาธิสั้น บกพร่องทางการเรียนรู้ และออทิซึม กำหนดผลผลิตของการวิจัย ดังนี้

1. ชุดแบบสอบ (Test battery) วัดทักษะทางวิชาการ 4 ทักษะ คือ แบบสอบการอ่านคำ (Word Reading) แบบสอบการสะกดคำ (Word Spelling) แบบสอบความเข้าใจประโยค (Sentence Comprehension) และ แบบสอบการคิดคำนวณ และแก้โจทย์ปัญหาอย่างง่าย (Math Computation) โดยจัดชุดแบบสอบเป็น 4 ฉบับ ด้วยรูปแบบการวัดพัฒนาทักษะตามช่วงชั้น (Grade-to-Grade Growth) กล่าวคือ แต่ละฉบับประกอบด้วยแบบสอบย่อย 4 ทักษะ ซึ่งมีระดับความยากของข้อสอบที่จัดเรียงตามลำดับจากง่ายไปยากที่สอดคล้องกับช่วงชั้นตั้งแต่ ช่วงชั้นที่ 1 (ระดับชั้นประถมศึกษาปีที่ 1 ถึง 3) ช่วงชั้นที่สอง (ระดับชั้นประถมศึกษาปีที่ 3 ถึง 6) ช่วงชั้นที่สาม (ระดับชั้นมัธยมศึกษาปีที่ 1 ถึง 3) และ ช่วงชั้นที่สี่ (ระดับชั้นมัธยมศึกษาปีที่ 3 ถึง 6)

2. ตารางเกณฑ์ปกติ (Norms) เพื่อแปลความหมายคะแนนสอบ อ้างอิงกลุ่มประชากรนักเรียน ชั้นประถมศึกษาปีที่ 1 ถึงชั้นมัธยมศึกษาปีที่ 6 จากกลุ่มตัวอย่างนักเรียนในกรุงเทพมหานคร และจังหวัดขอนแก่น ตารางเกณฑ์ปกติมีสองแบบ คือ ตารางแปลงคะแนนดิบสู่คะแนนมาตรฐานตามเกณฑ์ปกติช่วงชั้น และตารางแปลงคะแนนดิบสู่คะแนนมาตรฐานตามเกณฑ์ปกติระดับชั้น

3. คู่มือการใช้ชุดแบบสอบ KBAST ซึ่งเป็นเอกสารประกอบการดำเนินการสอบ การรายงานผลการสอบ และข้อแนะนำการนำไปใช้ประกอบการประเมินและวินิจฉัยบุคคลที่มีภาวะสมาธิสั้น บกพร่องทางการเรียนรู้ และออทิซึม

การดำเนินการพัฒนาแบบสอบทักษะพื้นฐานทางวิชาการ

การดำเนินการมี 4 ขั้นตอน ประกอบด้วย

1. **ขั้นการสร้างข้อสอบ** เป็นเป้าหมายแรกของกระบวนการพัฒนาแบบสอบมาตรฐาน ข้อสอบเป็นหน่วยวัดที่ต้องมีคุณภาพเพื่อให้ได้แบบสอบที่เป็นมาตรวัดสามารถวัดในสิ่งที่ต้องการวัดได้ สาระของขั้นนี้มีลำดับงานที่สำคัญ 6 ขั้นตอนย่อย คือ

1.1 พัฒนาหลักสูตรการวัดทักษะขั้นพื้นฐานทางวิชาการเพื่อการประเมินและวินิจฉัยบุคคลที่มีภาวะสมาธิสั้น บกพร่องทางการเรียนรู้ และออทิซึม

1.2 สร้างผังกำหนดลักษณะแบบสอบทักษะพื้นฐานทางวิชาการ เพื่อเป็นผังกำกับกรเขียนข้อสอบของกลุ่มผู้เชี่ยวชาญ

1.3 วิเคราะห์คุณภาพข้อสอบโดยใช้เหตุผล (logical review)

1.4 ทบทวนคุณภาพข้อสอบด้วยข้อมูลเชิงประจักษ์ขั้นเบื้องต้น

1.5 จัดชุดแบบสอบ ประกอบด้วยแบบสอบวัดทักษะย่อย 4 ชุด ที่มีความยากลดหลั่นเป็น 4 ระดับ แล้วนำมาประกอบเป็น 4 ฉบับ เพื่อวัดทักษะผู้สอบตามช่วงชั้น

1.6 สร้างระบบคลังข้อสอบ เพื่อจัดเก็บข้อสอบเป็นกลุ่ม (item pools) ตามทักษะที่วัด

2. ขั้นการสร้างมาตรฐานการวัดของแบบสอบ (Standardization) เป็นขั้นตอนที่ใช้กระบวนการพัฒนาแบบสอบตามทฤษฎีการวัดความสามารถในการวัดทักษะที่ต่างระดับช่วงชั้น (ชั้นประถมศึกษาปีที่ 1 ถึงมัธยมศึกษาปีที่ 6) เทคนิคที่ใช้คือ การพัฒนาแบบสอบย่อยทั้ง 4 ชุด ให้ผลการสอบมีความหมายแสดงถึงทักษะของผู้สอบตามระดับชั้น ทฤษฎีที่ใช้ คือ Item-response theory ออกแบบตามกระบวนการทางสถิติด้วยเทคนิคการปรับเทียบคะแนนต่างระดับชั้น (vertical equating)

3. ขั้นการพัฒนาตารางเกณฑ์มาตรฐาน (norms development) เป็นขั้นตอนที่นำแบบสอบไปให้กลุ่มตัวอย่างประชากรตามขอบเขตการวิจัย นำผลการสอบมาใช้ในกระบวนการแปลงคะแนนดิบสู่คะแนนมาตรฐาน แล้วสร้างตารางเกณฑ์ปกติ (Norms) เพื่อการแปลความหมายคะแนนสอบในโอกาสที่นำแบบสอบไปใช้กับตัวอย่างผู้สอบในกรอบของประชากรที่พัฒนาแบบสอบ

4. สร้างคู่มือบริหารการสอบชุดแบบสอบ KBAST เป็นคู่มือที่ให้รายละเอียดแก่ผู้นำชุดแบบสอบไปใช้ ตามข้อกำหนดที่เป็นมาตรฐานของชุดแบบสอบ

รายละเอียดแต่ละขั้นตอนมีดังต่อไปนี้

1. การสร้างข้อสอบวัดทักษะพื้นฐานทางวิชาการ

การดำเนินการขั้นนี้ มีเป้าหมายเพื่อให้ได้แบบวัดที่มีความตรงเชิงเนื้อหา (content validity) มีคุณภาพที่วัดทักษะพื้นฐานทางวิชาการของบุคคลที่มีภาวะสมาธิสั้น บกพร่องทางการเรียนรู้ และออทิสซึม ในวิชาภาษาไทยและคณิตศาสตร์ การปฏิบัติการครั้งนี้จำเป็นต้องออกแบบการสอบ และกำหนดสาระในแบบสอบ รวมทั้งเขียนข้อสอบที่ตรงตามแบบที่ตรวจสอบแล้ว ผู้เชี่ยวชาญในคณะทำงานสร้างแบบสอบนี้ ประกอบด้วยบุคคลที่มีประสบการณ์การสอนวิชาภาษาไทยและคณิตศาสตร์ในระดับการศึกษาขั้นพื้นฐาน และมีประสบการณ์จัดการศึกษาแก่นักเรียนที่มีความบกพร่องการเรียนรู้ รวม 13 คน (ภาคผนวก ก.) ปฏิบัติงานร่วมกับคณะผู้วิจัยตั้งแต่การออกแบบการสอบ การพิจารณาสาระในแต่ละทักษะ การเขียนข้อสอบ การพิจารณาปรับปรุงข้อสอบ การคัดเลือกข้อสอบ และพิจารณาความสอดคล้องของข้อสอบกับระดับการเรียนรู้ด้วยเหตุผลทางวิชาการและบริบทของสังคมวัฒนธรรม รายละเอียดมีลำดับดังนี้

1.1 จากความหมาย และขอบเขตเนื้อหาของแบบสอบทักษะพื้นฐานทางวิชาการ KBAST เพื่อการประเมินและวินิจฉัยบุคคลที่มีภาวะสมาธิสั้น บกพร่องทางการเรียนรู้ และออทิสซึม ที่กำหนดว่าเป็นแบบสอบที่วัดทักษะที่จำเป็นในการเรียนรู้ขั้นพื้นฐานและการสื่อสารในชีวิตประจำวัน ซึ่ง

ประกอบด้วยทักษะทางภาษา ได้แก่ ทักษะการอ่านคำ ทักษะการเขียนสะกดคำ ทักษะการอ่านเข้าใจ ประโยค และทักษะทางคณิตศาสตร์ ได้แก่ ทักษะการคำนวณ และทักษะการแก้โจทย์ปัญหาในชีวิตประจำวัน ผู้เชี่ยวชาญสร้างหลักสูตรการวัดทักษะขั้นพื้นฐานทางวิชาการ 2 หมวด คือ ภาษาไทยและการคิดคำนวณทางคณิตศาสตร์ ครอบคลุมทักษะ 4 ด้าน คือ การอ่านคำ การสะกดคำ ความเข้าใจประโยค และการคิดคำนวณทางคณิตศาสตร์ โดยยึดนิยามทักษะพื้นฐานที่กำหนด ว่าเป็นทักษะการอ่านออกเขียนได้ขั้นต่ำที่ใช้เพื่อการสื่อสารและเรียนรู้ของนักเรียนระดับการศึกษาขั้นพื้นฐาน (ชั้นประถมศึกษาปีที่ 1 ถึงมัธยมศึกษาปีที่ 6) เอกสารที่ใช้ประกอบการพิจารณาได้แก่หลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551 สรุปผลการวิเคราะห์หลักสูตรการศึกษาขั้นพื้นฐาน (สำนักวิชาการและมาตรฐานการศึกษา : 2546) เอกสารสาระการเรียนรู้คณิตศาสตร์ ตามหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2553 ที่จัดทำโดยโรงเรียนต่างๆ เอกสารวิชาการที่เกี่ยวข้องการจัดการการศึกษาพิเศษ ศูนย์วิจัยการศึกษาเพื่อเด็กที่ต้องการความช่วยเหลือพิเศษ โรงเรียนสาธิตแห่งมหาวิทยาลัยเกษตรศาสตร์ ศูนย์วิจัยและพัฒนาการศึกษา และประสบการณ์จากการให้ความช่วยเหลือนักเรียนในการพัฒนาทักษะดังกล่าว

1.2 สร้างฝั่งกำหนดลักษณะแบบสอบทักษะพื้นฐานทางวิชาการ เป็นพิมพ์เขียวของการพัฒนาแบบสอบครั้งนี้ ประกอบด้วยขอบเขตเนื้อหาในแต่ละช่วงชั้น และลักษณะของข้อสอบ โดยผู้เชี่ยวชาญร่วมกับนักทดสอบ (ผู้วิจัย) จากนั้นผู้เชี่ยวชาญแบ่งการพิจารณาเป็นกลุ่มย่อยตามความเชี่ยวชาญ และกำหนดลักษณะข้อสอบในรายละเอียดเพื่อกำกับการเขียนข้อสอบอีกครั้ง

1.3 วิเคราะห์คุณภาพข้อสอบด้วยเหตุผล ผู้เชี่ยวชาญเขียนข้อสอบ ตามลักษณะเฉพาะที่กำหนด แล้วประชุมพิจารณาตรวจสอบความตรงตามเนื้อหาด้วยวิธีเชิงเหตุผลโดยผู้เชี่ยวชาญร่วมกับนักทดสอบ ปรับปรุงแก้ไข ได้กลุ่มข้อสอบ (item pools) ตามทักษะที่ต้องการวัด

1.4 ทบทวนคุณภาพข้อสอบด้วยข้อมูลเชิงประจักษ์ขั้นเบื้องต้น เพื่อคัดเลือกข้อสอบที่มีความยากพอเหมาะและที่คุ้นเคยในชีวิตของนักเรียน เป็นการขจัดความลำเอียงตามสภาพสังคมเบื้องต้นด้วยข้อมูลเชิงประจักษ์ ตรวจสอบความเหมาะสมเป็นรายชื่อว่ามีความยาก และเป็นคำหรือข้อความ หรือปัญหาที่นักเรียนในระดับชั้นนั้นๆ มีความคุ้นเคยเพียงไร ทีมวิจัยนำข้อสอบไปทดลองใช้กับนักเรียนในโรงเรียนขนาดใหญ่ 1 โรง โรงเรียนเอกชน 1 โรง โรงเรียนในสังกัดกรุงเทพมหานคร 1 โรง โดยจัดกลุ่มข้อสอบรายวิชาตามระดับชั้นเป็นฉบับย่อยๆ ฉบับละ 30 - 40 ข้อ แล้วให้นักเรียนแต่ละระดับชั้น ขณะทำการทดสอบ นักวิจัยสังเกตพฤติกรรมการทำแบบสอบ เวลาที่ใช้ แล้วนำมาหาค่าร้อยละของความถูกต้อง คัดข้อสอบที่มีความยากมากและง่ายมากออกจากกลุ่มข้อสอบ (ใช้ระดับร้อยละ 90 - 100 และร้อยละ 0 - 25 ของระดับชั้นนั้นๆ) นำเสนอให้ที่ประชุมพิจารณาเพื่อคัดเลือกข้อสอบอีกครั้ง จนได้ชุดยุดิของคุณภาพข้อสอบที่วัดตรงตามข้อกำหนดในหลักสูตรการทดสอบ

1.5 ตรวจสอบคุณภาพของข้อสอบด้วยข้อมูลเชิงประจักษ์ ข้อสอบแต่ละกลุ่มจะนำไปหาค่าพารามิเตอร์ประมาณความยาก (b) และอำนาจจำแนก (a) ตาม Item response theory แบบ 2 – parameter ด้วยโปรแกรม BILOG (Kolen & Brennan, 2010: 155-218) การดำเนินการขั้นนี้ มีดังนี้

1.5.1 จัดชุดข้อสอบเป็นแบบสอบย่อยรายทักษะ และแบ่งเป็นฉบับตามช่วงชั้น จัดพิมพ์เอกสารแบบสอบให้พร้อมใช้ เขียนคู่มือการบริหารการสอบ

1.5.2 ตรวจสอบค่าเฉลี่ยเพื่อพร้อมในการตรวจคำตอบ และออกแบบการบันทึกข้อมูล เพื่อนำมาวิเคราะห์ในโปรแกรม BILOG

1.5.3 สุ่มตัวอย่างนักเรียนชั้นประถมศึกษาปีที่ 1 ถึงมัธยมศึกษาปีที่ 6 ในเขตกรุงเทพมหานคร จำแนกตามสังกัด ระดับชั้นละ 200 คน รวมประมาณ 2,400 คน

1.5.4 ประสานงานกับโรงเรียน ชี้แจงการบริหารจัดการในระดับโรงเรียน และอบรมครูผู้ทำหน้าที่ให้การทดสอบในห้องเรียน

1.5.5 คณะผู้วิจัยรวบรวมผลการสอบ และวิเคราะห์ค่าพารามิเตอร์ข้อสอบรายข้อค่าความยาก (b) ค่าอำนาจจำแนก (a) ลักษณะโค้งข้อสอบ (ICC) และสารสนเทศข้อสอบ (item information) แล้วคัดเลือกข้อสอบที่ให้สารสนเทศในพฤติกรรมคำตอบข้อสอบในระดับชั้นที่เหมาะสม

1.6 สร้างระบบคลังข้อสอบ เพื่อจัดเก็บข้อสอบเป็นกลุ่ม (item pools) ตามทักษะที่วัด ข้อสอบทุกข้อมีรหัสประจำข้อ (item ID) กำหนดดังนี้ หลักที่ 1 เป็นรหัสทักษะ (1 คือ การอ่านคำ 2 คือ การสะกดคำ 3 คือ ความเข้าใจประโยค และ 4 คือ การคิดคำนวณทางคณิตศาสตร์) หลักที่ 2 เป็นรหัสช่วงชั้น (1 คือ ช่วงชั้นที่ 1 2 คือ ช่วงชั้นที่ 2 3 คือ ช่วงชั้นที่ 3 4 คือ ช่วงชั้นที่ 4) ส่วน y เป็นรหัสกำหนดหมวดหมู่เนื้อหาในทักษะนั้นๆ x เป็นรหัสแสดงลำดับข้อในแต่ละหมวดหมู่เนื้อหา ดังผังการจัดเก็บข้อสอบ ดังตารางที่ 4.3

ตารางที่ 4.3 ผังการจัดเก็บกลุ่มข้อสอบ จำแนกตามทักษะที่วัด และช่วงชั้น

ระดับช่วงชั้น	คลังข้อสอบ			
	การอ่านคำ	การสะกดคำ	ความเข้าใจประโยค	การคิดคำนวณทางคณิตศาสตร์
ช่วงชั้นที่ 1 ป.1-ป.3	11yx	21yx	31yx	41yx
ช่วงชั้นที่ 2 ป.4-ป.6	12yx	22yx	32yx	42yx
ช่วงชั้นที่ 3 ม.1-ม.3	13yx	23yx	33yx	43yx
ช่วงชั้นที่ 4 ม.4-ม.6	14yx	24yx	34yx	44yx

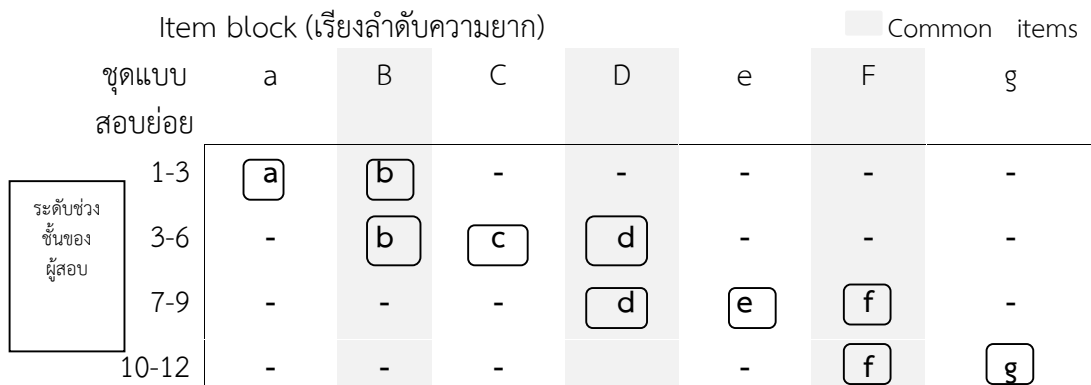
1.7 จัดเก็บข้อสอบ พร้อมข้อมูลข้อสอบที่สำคัญ ได้แก่ รหัสประจำข้อ (Item ID) จัดเรียงลำดับตามค่าความยากแต่ละหมวดหมู่เนื้อหา คำบรรยายลักษณะเฉพาะของข้อสอบ ตัวข้อสอบ และค่าประมาณความยากเพื่อประมาณระดับความสามารถ () ของผู้สอบ

2. การพัฒนามาตรฐานชุดแบบสอบ KBAST

2.1 การออกแบบชุดแบบสอบมาตรฐานวัดทักษะพื้นฐานทางวิชาการ KBAST เลือกรูปแบบโครงสร้างความเพิ่มพูนของความสามารถตามช่วงชั้น (Grade-to-Grade Growth model) (Kolen & Brennan, 2010: 376-377) สร้างชุด (Battery) วัดทักษะพื้นฐานทางวิชาการของประชากรนักเรียนชั้นประถมศึกษาปีที่ 1 ถึงมัธยมศึกษาปีที่ 6 ใน 4 ทักษะ กำหนดใช้แบบสอบฉบับที่ 1 2 3 และ 4 กับนักเรียนที่เรียนอยู่ช่วงชั้นที่ 1 2 3 และ 4 ตามลำดับ ชุดแบบสอบ KBAST จึงประกอบด้วย

1. แบบสอบย่อยการอ่านคำ ฉบับที่ 1 2 3 และ 4
2. แบบสอบย่อยการสะกดคำ ฉบับที่ 1 2 3 และ 4
3. แบบสอบย่อยความเข้าใจประโยค ฉบับที่ 1 2 3 และ 4
4. แบบสอบย่อยการคำนวณทางคณิตศาสตร์ ฉบับที่ 1 2 3 และ 4

2.2 สร้างแบบสอบย่อยรายทักษะ เลือกข้อสอบจากคลังข้อสอบลงในฉบับที่ 1 ฉบับที่ 2 ฉบับที่ 3 และฉบับที่ 4 ให้ตรงตามโครงสร้างความเพิ่มพูนของความสามารถตามช่วงชั้น (Grade-to-Grade Growth model) ซึ่งต้องพิจารณาทั้งความครอบคลุมของเนื้อหา และระดับความยากที่ต้องลดหลั่นตามช่วงชั้น พร้อมทั้งเลือกข้อสอบที่จะทำหน้าที่เป็นข้อสอบร่วม (common items) ในรอยต่อของแต่ละฉบับ เพื่อใช้ระเบียบวิธีการเทียบมาตรวัดของแบบสอบในชุด (battery) มาตรวจสอบความตรงของชุดแบบสอบว่าเป็นไปตามโครงสร้างความเพิ่มพูนของความสามารถตามช่วงชั้นหรือไม่ เทคนิคนี้เป็นการเทียบมาตรคะแนนแนวตั้ง (Vertical Equating) ที่มีข้อสอบร่วม (Common Items) เป็นตัวเชื่อมในการวิเคราะห์ผลการตอบจากกลุ่มตัวอย่างระดับชั้นที่ต่างกัน ดังแผนภาพที่ 4.3 (Kolen & Brennan, 2010: 378)



แผนภาพที่ 4.3 โมเดลโครงสร้างแบบสอบ เพื่อวัดทักษะ 1 ทักษะ

จากรูปแบบโครงสร้างที่กำหนดนี้ ข้อสอบกลุ่ม a ที่คัดมาจากคลังข้อสอบ จะมีระดับความยากน้อยกว่าข้อสอบกลุ่ม c ทำนองเดียวกันข้อสอบกลุ่ม c จะมีระดับความยากน้อยกว่าข้อสอบกลุ่ม e และข้อสอบกลุ่ม e ก็มีระดับความยากน้อยกว่าข้อสอบกลุ่ม g ส่วนข้อสอบกลุ่ม b d และ f เป็นข้อสอบรวมที่อยู่ใน 2 ฉบับ เพื่อให้สารสนเทศเชื่อมต่อแบบสอบ 2 ฉบับ ด้วยกระบวนการวิเคราะห์ของโปรแกรม BILOG จะทำให้ค่าประมาณความยากรายข้อในแบบสอบย่อยที่กษะนั้นมีการเชื่อมโยงในมาตรวัดเดียวกัน ทำให้ค่าประมาณความสามารถของผู้สอบจากการสอบแบบสอบต่างฉบับสามารถนำมาเปรียบเทียบกันได้

2.3 ออกแบบการใช้ข้อมูลเชิงปริมาณตามเทคนิคการเทียบมาตราคะแนนแบบแนวตั้ง (Vertical Equating Technique) กลุ่มตัวอย่างนักเรียนได้จากการสุ่มหลายชั้นตอน เป็นตัวแทนนักเรียนชั้นประถมศึกษาปีที่ 1 ถึงมัธยมศึกษาปีที่ 6 ของโรงเรียนในกรุงเทพมหานครและจังหวัดขอนแก่น ตามข้อกำหนดของโครงการวิจัย จำนวนทั้งสิ้น 1,248 คน กลุ่มตัวอย่างนักเรียนดังกล่าว จะได้รับการทดสอบจากแบบสอบ KBAST และแบบสอบสติปัญญาที่ไม่ใช้ภาษา

ข้อมูลคะแนนจากการสอบนำมาวิเคราะห์ด้วยโปรแกรม BILOG-MG ITEM MAINTENANCE PROGRAM: LOGISTIC ITEM RESPONSE MODEL (SCIENTIFIC SOFTWARE INTERNATIONAL, INC. 2002) กำหนดให้กลุ่มผู้สอบในแบบสอบฉบับที่ 3 เป็นกลุ่มอ้างอิง นำผลการวิเคราะห์มาพิจารณาความเหมาะสมของข้อสอบจาก ICC ซึ่งแสดงด้วยค่าความยาก (b) ค่าอำนาจจำแนก (a) และคุณลักษณะของแบบสอบ (Test information curve : TFC) ที่อธิบายช่วงความสามารถในมาตรวัดแต่ละแบบสอบย่อย

การปฏิบัติการวิเคราะห์ทำหลายรอบ เริ่มพิจารณาผลการวิเคราะห์ครั้งที่หนึ่ง ว่าการวัดทักษะย่อยแต่ละทักษะด้วยแบบสอบฉบับที่ 1 2 3 และ 4 มีคุณลักษณะสอดคล้องกับโมเดลความเพิ่มพูนของความสามารถตามช่วงชั้น (Grade-to-Grade Growth) หรือไม่ แล้วปรับปรุงโดยตัดข้อมูลของบุคคล (Person) ที่มีพฤติกรรมตอบไม่สอดคล้องกับทฤษฎีการตอบสนองข้อสอบ (IRT) ออกจากชุดข้อมูล ขณะเดียวกันพิจารณาที่ข้อสอบ (Item) ข้อสอบที่มีค่าความยาก (b) อยู่นอกขอบเขตการวัดความสามารถในช่วงชั้น ถือว่าไม่สอดคล้องกับโมเดลความเพิ่มพูนของความสามารถตามช่วงชั้น จะถูกตัดออก ปรับเพิ่มข้อมูลแล้ววิเคราะห์หาค่าพารามิเตอร์ของข้อสอบและแบบสอบทั้งชุด ครั้งที่สองดำเนินการพิจารณาผลการวิเคราะห์และปรับปรุงเพิ่มข้อมูลอีกครั้ง ผู้วิจัยจะใช้กระบวนการวิเคราะห์ทำนองนี้หลายครั้ง จนกว่าได้ชุดแบบสอบมีความสอดคล้องกับโมเดลความเพิ่มพูนของความสามารถตามช่วงชั้น

2.4 ตรวจสอบความสอดคล้องของชุดแบบทดสอบตามโมเดลความเพิ่มพูนของความสามารถตามช่วงชั้น ด้วยการทดสอบความมีนัยสำคัญของความแตกต่างของค่าประมาณความสามารถของผู้สอบในแต่ละช่วงชั้น สถิติที่ใช้คือ the Kruskal-Wallis test สถิติทดสอบคือ "Adjusted H" ซึ่งปรับจำนวนที่มีซ้ำ ข้อมูลคือค่ามัธยฐานของค่าประมาณความสามารถ () ของผู้สอบ

ตามผลการวิเคราะห์ BILOG ใน Phase 3 สมมติฐานของการทดสอบทางสถิติกำหนด Null Hypothesis: Median Ability₁ = Median Ability₂ = Median Ability₃ = Median Ability₄ ส่วนกำหนดไว้ว่า Alternative Hypothesis: at least a pair of Median ability is different. ถ้าค่า p-value < α = 0.05, แสดงว่ามีความแตกต่างอย่างมีนัยสำคัญเกิดขึ้นในค่ามัธยฐานของค่าความสามารถของกลุ่มที่ทำการทดสอบนั้นๆ หลังจากนั้นทดสอบทีละคู่ เพื่อดูว่าคู่ไหนต่างกัน

ลักษณะความแตกต่างนี้อธิบายถึงความเพิ่มพูนของความสามารถตามช่วงชั้นของกลุ่มผู้สอบ (operationalizing growth) จากช่วงชั้นที่ 1 2 3 และ 4 ในลักษณะที่เป็นขั้นบันได (Kolen & Brennan, 2010: 376-381) ซึ่งจะนำไปสู่ข้อสรุปว่าชุดแบบสอบที่ออกแบบพัฒนาเป็น 4 ฉบับ เพื่อวัดบุคคลตามช่วงชั้นนั้น มีคุณสมบัติของการระบุความสามารถได้ตามโมเดลความเพิ่มพูนของความสามารถตามช่วงชั้น

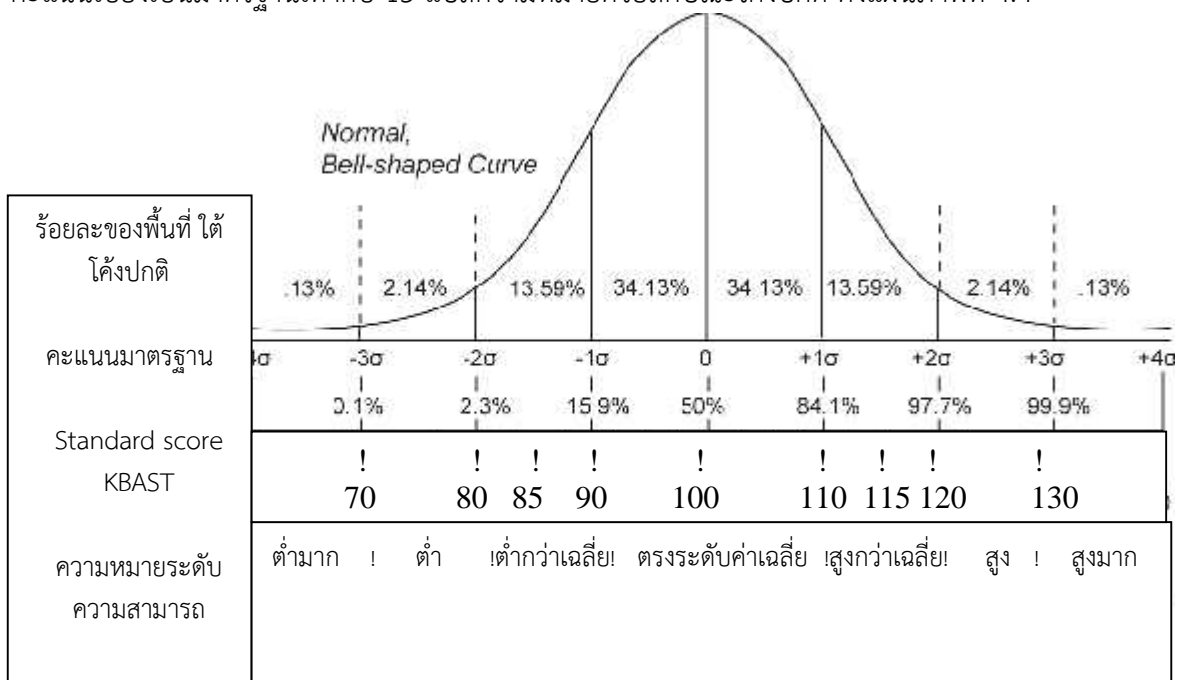
2.5 หาค่าประมาณความเที่ยง (reliability) ตามที่กำหนดในมาตรฐานการทดสอบทางการศึกษาและจิตวิทยา (the Standards for Educational and Psychological Testing : American Educational Research Association et al., 1999) ค่าความเที่ยงได้จากการวิเคราะห์ตามทฤษฎีการทดสอบแบบดั้งเดิม (classical test theory) เป็นค่าความคงที่ภายใน (internal consistency) ค่าความคลาดเคลื่อนของการวัด (standard error of measurement) และช่วงความเชื่อมั่นของคะแนนมาตรฐาน

3. การพัฒนาเกณฑ์มาตรฐาน

การวิจัยพัฒนาแบบสอบมาตรฐาน KBAST ได้ออกแบบพัฒนาให้เป็นชุดแบบสอบ (Test Battery) วัดทักษะขั้นพื้นฐานทางวิชาการทางภาษาไทย 3 ทักษะ และคณิตศาสตร์ 1 ทักษะ กลุ่มเป้าหมายของการนำแบบสอบนี้ไปใช้คือ บุคคลที่มีภาวะสมาธิสั้น บกพร่องทางการเรียนรู้ และออทิซึม เพื่อนำผลการสอบไปประเมินและวินิจฉัยระดับความสามารถของบุคคลดังกล่าว อันจะนำไปสู่การออกแบบการให้ความช่วยเหลือที่เหมาะสม กรอบเนื้อหาการอ้างอิงระดับความสามารถของทักษะขั้นพื้นฐานจึงใช้กรอบการเรียนรู้ระดับการอ่านออกและคิดคำนวณที่ใช้ในชีวิตประจำวัน ส่วนกรอบประชากรที่ใช้ในการอ้างอิงคือ บุคคลในระบบการศึกษาตั้งแต่ชั้นประถมศึกษาปีที่ 1 ถึงมัธยมศึกษาปีที่ 6 ที่บุคคลที่มีภาวะสมาธิสั้น บกพร่องทางการเรียนรู้ และออทิซึมรวมอยู่ด้วย ชุดแบบสอบ KBAST จึงเป็นแบบสอบมาตรฐานอิงกลุ่มผู้เรียนที่เป็นตัวแทนประชากร (Norm-referenced Test) กระบวนการพัฒนาเกณฑ์มาตรฐาน (norms development) ดำเนินการดังนี้

3.1 กลุ่มตัวอย่างผู้สอบได้จากการสุ่มตัวแทนประชากรนักเรียนชั้นประถมศึกษาปีที่ 1 ถึงชั้นมัธยมศึกษาปีที่ 6 ของกรุงเทพมหานครและจังหวัดขอนแก่น ดังได้กล่าวมาแล้ว เมื่อผ่านกระบวนการวิเคราะห์ BILOG ได้ตัดผู้สอบที่มีพฤติกรรมคำตอบข้อสอบไม่สอดคล้องกับเงื่อนไขของทฤษฎีการตอบสนองข้อสอบออกไปจำนวนหนึ่ง นำคะแนนสอบของนักเรียนมาประมวลผลเป็นคะแนนรวม

รายบุคคล หาค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐานเป็นกลุ่มตามระดับชั้น และช่วงชั้น แล้วนำมาแปลงคะแนน แต่ละกลุ่มระดับชั้นและช่วงชั้นให้เป็นคะแนนมาตรฐาน (scaled score) มีค่าเฉลี่ยเท่ากับ 100 และคะแนนเบี่ยงเบนมาตรฐานเท่ากับ 15 แปลความหมายด้วยลักษณะโค้งปกติ ดังแผนภาพที่ 4.4



แผนภาพที่ 4.4 แสดงความหมายของผลการสอบที่แปลงเป็นคะแนนมาตรฐาน 7 ระดับความสามารถ

3.2 จัดทำตารางเทียบคะแนนดิบสู่คะแนนมาตรฐานมีทั้ง 18 ตาราง ได้แก่ ตารางเทียบคะแนนดิบสู่คะแนนมาตรฐานรายชั้นประถมศึกษาปีที่ 1-มัธยมศึกษาปีที่ 6 (Raw Score Subtest Standard Score Conversion by Grade) 12 ตาราง ตารางเทียบคะแนนดิบสู่คะแนนมาตรฐานรายฉบับ (ฉบับที่ 1- 4) (Raw Score Standard Score Conversion by Test Form A-D) 4 ตาราง

3.3 จัดทำคู่มือการแปลความหมายคะแนน แล้วนำไปให้ครูที่มีนักเรียนที่มีความบกพร่องในการเรียนรู้ในความรู้ความรับผิดชอบนำไปใช้เก็บข้อมูลด้วยการสัมภาษณ์ นำมาปรับปรุงและสร้างข้อเสนอแนะในการใช้แบบสอบ KBAST ในสถานการณ์ที่เป็นจริง

4. สร้างคู่มือบริหารการสอบชุดแบบสอบ KBAST

4.1 คู่มือเป็นส่วนประกอบสำคัญของชุดแบบสอบที่พัฒนาขึ้น โดยมีวัตถุประสงค์เพื่อให้รายละเอียดแก่ผู้นำชุดแบบสอบไปใช้ ปฏิบัติได้ถูกต้องตามข้อกำหนดที่เป็นมาตรฐานของชุดแบบสอบ (Professional Manual) ผู้ใช้ชุดแบบสอบจำเป็นต้องได้รับการฝึก หรือเป็นผู้ที่มีประสบการณ์ด้านนี้ หัวข้อที่บรรจุในคู่มือ ประกอบด้วย

1. บทนำ : แนวคิดการพัฒนา วัตถุประสงค์ และการใช้ชุดแบบสอบ
2. การบริหารการสอบและการตรวจคะแนน
3. การแปลความหมายผลการสอบ
4. การพัฒนาชุดแบบสอบ

4.2 การตรวจสอบคุณภาพของคู่มือ ใช้วิธีการสำรวจความเห็นของผู้ใช้แบบสอบ 3 ประเด็น คือความง่ายในการบริหารการสอบ ความชัดเจนในการตรวจให้คะแนน และความสะดวกในการใช้ประโยชน์จากการสอบ เก็บข้อมูลด้วยการสัมภาษณ์ และความเพียงพอของการนำไปวางแผนให้ความช่วยเหลือแก่นักเรียนที่รับการสอบ

ผลการวิจัย

ผลการพัฒนาชุดแบบสอบ KBAST

จากการกำหนดวัตถุประสงค์ของการพัฒนาแบบสอบทักษะพื้นฐานทางวิชาการ เพื่อนำไปใช้ในการประเมินผู้ที่มีความบกพร่องทางการเรียนรู้ ซึ่งเป็นประโยชน์ต่อผู้ทำหน้าที่ให้ความช่วยเหลือบุคคลดังกล่าว ที่จะนำผลการประเมินเป็นข้อมูลประกอบการสร้างแผนให้ความช่วยเหลือ (IEP) แก่ผู้ที่มีความบกพร่องทางการเรียนรู้ที่อยู่ในความรับผิดชอบ การพัฒนาแบบสอบดังกล่าวจึงได้วางแผนและออกแบบพัฒนาให้เป็นชุดแบบสอบมาตรฐาน กำหนดให้ชื่อแบบสอบว่า แบบสอบทักษะขั้นพื้นฐานทางวิชาการ (Kasetsart Basic Academic Skills Test : KBAST) มีขอบเขตวัดทักษะภาษาไทย 3 ทักษะ และคณิตศาสตร์ 1 ทักษะ พัฒนาการวัดระดับการเรียนรู้ขั้นต่ำของนักเรียนชั้นประถมศึกษาปีที่ 1 ถึงมัธยมศึกษาปีที่ 6 รวม 12 ชั้นปี และสร้างมาตรฐานการวัดแบบอิงกลุ่มที่จะให้สามารถแปลความหมายผลการสอบของบุคคลใดบุคคลหนึ่งที่มีภาวะสมาธิสั้น บกพร่องทางการเรียนรู้ และออทิสซึม ว่ามีทักษะในระดับใดเมื่อพิจารณาจากกลุ่มอ้างอิงนักเรียนชั้นประถมศึกษาปีที่ 1 ถึงมัธยมศึกษาปีที่ 6 จะช่วยให้การออกแบบเพื่อพัฒนาและจัดกิจกรรมให้ความช่วยเหลือได้เหมาะสมยิ่งขึ้น ผลการพัฒนาแบบสอบได้ดำเนินการตามแบบแผนของการพัฒนาแบบสอบมาตรฐาน นำเสนอเป็น 5 ตอน ดังนี้

1. การสร้างข้อสอบวัดทักษะพื้นฐานทางวิชาการ KBAST
2. การสร้างมาตรฐานแบบสอบ KBAST
3. ความตรงของแบบสอบมาตรฐาน
4. ความเที่ยงของแบบสอบมาตรฐาน
5. การพัฒนาเกณฑ์มาตรฐาน

ตอนที่ 1 การสร้างข้อสอบวัดทักษะพื้นฐานทางวิชาการ

การดำเนินการขั้นนี้มีเป้าหมายเพื่อให้ได้แบบวัดที่มีความตรงเชิงเนื้อหา (content validity) มีคุณภาพที่วัดทักษะพื้นฐานทางวิชาการของบุคคลที่มีภาวะสมาธิสั้น บกพร่องทางการเรียนรู้ และออทิสซึมในวิชาภาษาไทยและคณิตศาสตร์ การออกแบบการสอบครั้งนี้จำเป็นต้องกำหนดนิยามทักษะพื้นฐานและกรอบเนื้อหา แล้วเขียนข้อสอบที่สามารถวัดได้ตรง คณะผู้เชี่ยวชาญประกอบด้วย บุคคลที่มีประสบการณ์การสอนวิชาภาษาไทยและคณิตศาสตร์ในระดับการศึกษาขั้นพื้นฐาน และมีประสบการณ์จัดการศึกษาแก่นักเรียนที่มีความบกพร่องทางการเรียนรู้ รวม 13 คน (ภาคผนวก ก.) ปฏิบัติงานร่วมกันกับคณะผู้วิจัย ตั้งแต่การออกแบบการสอบ การพิจารณาสาระในแต่ละทักษะ การเขียนข้อสอบ การพิจารณาปรับปรุง

ข้อสอบ การคัดเลือกข้อสอบ และพิจารณาความสอดคล้องของข้อสอบกับระดับการเรียนรู้ด้วยเหตุผลทางวิชาการและบริบทของสังคมวัฒนธรรม ผลการพัฒนาในแต่ละขั้นตอนมีดังนี้

1.1 ขอบเขตเนื้อหาของแบบสอบทักษะพื้นฐานทางวิชาการ ประกอบด้วยทักษะทางภาษาไทย และทักษะทางคณิตศาสตร์ หลังจากทบทวนทฤษฎีและผู้เชี่ยวชาญด้านหลักสูตรและการสอนได้ตรวจสอบความสอดคล้องของวัตถุประสงค์ของแบบวัดและองค์ประกอบของทักษะพื้นฐานแล้ว กำหนดเป็นหลักสูตรการสอบและแสดงเป็นผังกำหนดลักษณะแบบสอบอย่างสังเขป ดังตารางที่ 4.4

ตารางที่ 4.4 ผังกำหนดลักษณะแบบสอบ ทักษะพื้นฐานทางวิชาการ

แบบสอบย่อย	เนื้อหา	ระดับชั้น	ลักษณะของข้อสอบ
การอ่าน	พยัญชนะ	ป.1 – ป.3	กำหนดพยัญชนะให้ผู้สอบอ่านออกเสียง
	คำ	ป.1 – ป.6 ม.1 – ม.6	กำหนดคำให้ผู้สอบอ่านออกเสียง
การสะกด	พยัญชนะ	ป.1 – ป.3	ครูอ่านพยัญชนะให้ผู้สอบเขียนตามคำบอก
	คำ	ป.1 – ป.6 ม.1 – ม.6	ครูอ่านคำ ผู้สอบเขียนตามคำบอก
ความเข้าใจ ประโยค	สาระความรู้จาก วิชาต่างๆ	ป.1 – ป.6 ม.1 – ม.6	กำหนดให้ผู้สอบเติมคำที่เว้นเพื่อให้ใจความสมบูรณ์ เรื่องที่น่าสนใจ เกี่ยวกับความรู้ทั่วไปในวิชาต่างๆ ที่เหมาะสมกับระดับชั้น
การคำนวณ ทาง คณิตศาสตร์	จำนวนเต็ม	ป.1 – ป.3	เติมคำตอบในโจทย์ที่กำหนดให้
	บวก ลบ คูณ หาร	ป.1 – ป.6 ม.1-ม.6	เติมคำตอบในโจทย์ที่กำหนดให้
	การนับ จำนวน	ป.1 – ป.6	เติมคำตอบในโจทย์ที่กำหนดให้
	เต็ม เศษส่วน โจทย์ปัญหา	ม.1 - ม.6 ม.1 - ม.6	ส่วนโจทย์ปัญหาที่มีลักษณะที่เกี่ยวข้องกับ ชีวิตประจำวันที่ใช้การคำนวณที่ไม่ซับซ้อน

1.2 ผู้เชี่ยวชาญเขียนข้อสอบตามลักษณะเฉพาะที่กำหนด แล้วตรวจสอบความตรงตามเนื้อหาในองค์ประกอบความครอบคลุมระดับความยาก-ง่าย สัดส่วนเนื้อหา และความเหมาะสมด้านบริบททางวัฒนธรรม การปรับปรุงแก้ไขได้ดำเนินการหลายขั้นตอน ทั้งวิธีเชิงเหตุผล และใช้ข้อมูลเชิงประจักษ์ เกณฑ์การประเมินคุณภาพขั้นการสร้างข้อสอบ ได้แก่ ความสอดคล้องของข้อสอบกับเนื้อหาในกรอบ

หลักสูตรขั้นพื้นฐาน โดยการวิเคราะห์เชิงเหตุผลของผู้เชี่ยวชาญ ความสอดคล้องของข้อสอบกับความสามารถของนักเรียนในแต่ละระดับชั้น โดยใช้ข้อมูลร้อยละของการทำถูกจากตัวอย่างนักเรียนในแต่ละระดับ ความเหมาะสมทางสังคมวัฒนธรรมในเนื้อหาของข้อสอบ โดยใช้ร้อยละการทำถูกในกลุ่มผู้สอบในโรงเรียนในเมืองและนอกเมือง ผลการสร้างข้อสอบจัดเป็นกลุ่มข้อสอบ (item pools) แล้วจัดเก็บข้อสอบเพื่อพัฒนาเป็นคลังข้อสอบต่อไป สรุปจำนวนข้อสอบในแต่ละกลุ่มที่ผ่านกระบวนการวิเคราะห์คุณภาพเบื้องต้น ดังตารางที่ 4.5

ตารางที่ 4.5 จำนวนข้อสอบในกลุ่มข้อสอบที่ผ่านกระบวนการวิเคราะห์คุณภาพเบื้องต้น

แบบสอบ ตาม ระดับช่วงชั้น	กลุ่มข้อสอบ			
	การอ่านค่า	การสะกดคำ	เข้าใจประโยค	การคิดคำนวณ
ป.1-ป.3	55	65	33	47
ป.4-ป.6	65	69	36	50
ม.1-ม.3	48	60	41	45
ม.4-ม.6	51	60	46	45
รวม	219	254	155	187

1.3 ตรวจสอบคุณภาพของข้อสอบ 4 กลุ่มทักษะรายข้อ ด้วยวิธีการเชิงประจักษ์ ในขั้นนี้กลุ่มตัวอย่างเป็นตัวอย่างนักเรียนชั้นประถมศึกษาปีที่ 1 ถึงมัธยมศึกษาปีที่ 6 ในเขตกรุงเทพมหานคร จำแนกตามสังกัด ระดับชั้นละ 200 คน รวมประมาณ 2,400 คน ข้อสอบอยู่ในชุดแบบสอบย่อยเป็นกลุ่มข้อสอบที่จัดโดยสุ่มภายในระดับชั้น แล้ววิเคราะห์ค่าพารามิเตอร์ความยาก (b) และอำนาจจำแนก (a) ด้วย Item response theory 2 - parameter ใช้โปรแกรม BILOG แล้วคัดเลือกข้อสอบจากการพิจารณา ค่าประมาณความยาก (b) และค่าอำนาจจำแนก (a)

1.4 กำหนดรหัสประจำข้อสอบที่ผ่านการคัดเลือกแล้ว บันทึกลงในคลังข้อสอบตามผังคลังข้อสอบ (ตารางที่ 4.1) เป็นขั้นสุดท้ายของการพัฒนาข้อสอบรายทักษะ มีคำบรรยายลักษณะเฉพาะของข้อสอบและเฉลยคำตอบ สรุปจำนวนและเนื้อหาข้อสอบในแต่ละทักษะย่อยดังตารางที่ 4.6 ถึง 4.9 และสรุปค่าประมาณความยากสูงสุด-ต่ำสุด ในแต่ละกลุ่มทักษะย่อย ในตารางที่ 4.10

ตารางที่ 4.6 เนื้อหาและจำนวนข้อสอบ ทักษะการอ่านคำ

แบบสอบ ช่วงชั้น	เนื้อหา						รวม
	พยัญชนะ	คำโดด	ประ-ไม่ประ วิสรรชนีย์	คำควบกล้ำ /ห นำ (3,4)	คำ ใช้ รร คำ มีการ์นต์ 6	คำวิชาการ/ ต่างประเทศ	
ป.1-ป.3	12	17	2	5	-	-	36
ป.4-ป.6	8	8	1	10	3	-	38
ม.1-ม.3	-	10	4	10	6	5	35
ม.4-ม.6	-	8	10	9	5	3	35

ตารางที่ 4.7 เนื้อหาและจำนวนข้อสอบ ทักษะการสะกดคำ

แบบสอบ ช่วงชั้น	เนื้อหา						รวม
	พยัญชนะ	คำโดด 1,2	ประ-ไม่ประ วิสรรชนีย์	คำควบกล้ำ /ห นำ (3,4)	คำ ใช้ รร 5 คำ มีการ์นต์	คำวิชาการ/ ต่างประเทศ	
ป.1-ป.3	12	12	2	4	-	-	30
ป.4-ป.6	8	7	2	5	5	3	30
ม.1-ม.3	-	10	5	5	6	4	30
ม.4-ม.6	-	4	4	8	6	8	30

ตารางที่ 4.8 เนื้อหาและจำนวนข้อสอบ ทักษะความเข้าใจประโยค

แบบสอบ ช่วงชั้น	เนื้อหาวิชา				รวม	ชนิดของคำ				รวม
	สุขศึกษา ปลอดภัย	สังคมศึกษา วัฒนธรรม	ธรรมชาติ วิทยาศาสตร์	ความรู้ ทั่วไป		คำนาม	กริยา	วิเศษณ์	บุพบท สันธาน	
ป.1-ป.3	5	2	4	5	16	7	2	4	3	16
ป.4-ป.6	2	8	5	5	20	8	5	5	2	20
ม.1-ม.3	5	5	7	5	22	8	7	4	3	22
ม.4-ม.6	5	7	5	3	20	7	7	5	1	20

ตารางที่ 4.9 เนื้อหาและจำนวนข้อสอบ ทักษะการคิดคำนวณทางคณิตศาสตร์

แบบสอบ ช่วงชั้น	เนื้อหา					รวม
	จำนวน	เศษส่วน	ทศนิยม	เปอร์เซ็นต์	พีชคณิต	
ป.1-ป.3	20	-	-	-	-	20
ป.4-ป.6	4	10	11	-	-	25
ม.1-ม.3	-	10	8	4	6	28
ม.4-ม.6	-	5	2	8	15	30

ตารางที่ 4.10 จำนวนข้อสอบ ค่าประมาณความยากสูงสุด-ต่ำสุด ในแต่ละกลุ่มทักษะย่อย

ช่วงชั้น	การอ่านคำ	การสะกดคำ	ความเข้าใจประโยค	การคิดคำนวณทาง คณิตศาสตร์
ป.1-ป.3	55 (-4.879 -0.303)	65 (-2.625 0.174)	33 (-3.888 -1.017)	47 (-2.175 1.517)
ป.4-ป.6	65 (-4.208 -0.27)	69 (-2.885 0.586)	36 (-3.511 1.597)	50 (-2.232 1.369)
ม.1-ม.3	48 (-4.932 0.938)	60 (-3.716 0.461)	41 (-2.498 1.597)	45 (-2.067 1.216)
ม.4-ม.6	51 (-5.092 0.701)	60 (-4.381 -0.267)	46 (-2.685 3.062)	45 (-0.841 1.645)

ตอนที่ 2 การสร้างมาตรฐานแบบสอบ KBAST

ขั้นการสร้างมาตรฐานการวัดของแบบสอบ (Standardization) เป็นขั้นตอนที่ใช้กระบวนการพัฒนาแบบสอบตามทฤษฎีการวัดความสามารถในการวัดทักษะที่ต่างระดับช่วงชั้น (ชั้นประถมศึกษาปีที่ 1 ถึงมัธยมศึกษาปีที่ 3) เทคนิคที่ใช้คือ การพัฒนาแบบสอบย่อยทั้ง 4 ชุด ให้ผลการสอบมีความหมายแสดงถึงทักษะของผู้สอบตามระดับชั้น ตามรูปแบบ Grade to Grade Growth Model ทฤษฎีที่ใช้พัฒนาข้อสอบ คือ Item-response theory เทคนิคการปรับเทียบคะแนนต่างระดับชั้น (vertical equating) ดำเนินการสร้างแบบสอบมาตรฐาน แล้วตรวจสอบความตรงตาม Model ดังนี้

2.1 ออกแบบชุดแบบสอบมาตรฐาน วัดทักษะพื้นฐานทางวิชาการ KBAST ตามรูปแบบโครงสร้างความเพิ่มพูนของความสามารถตามช่วงชั้น (Grade-to-Grade Growth Model) ของประชากรนักเรียนชั้นประถมศึกษาปีที่ 1 ถึงมัธยมศึกษาปีที่ 6 ใน 4 ทักษะ กำหนดใช้แบบสอบฉบับที่ 1 2 3 และ 4 กับนักเรียนที่เรียนอยู่ช่วงชั้นที่ 1 2 3 และ 4 ตามลำดับ แบบสอบย่อยวัดแต่ละทักษะเป็นผลการจัดข้อสอบที่มีคุณภาพจากคลังข้อสอบ วางลงอย่างมีแบบแผนตาม Grade-to-Grade Growth model

และเลือกข้อสอบร่วม (common items) ที่ทำหน้าที่เป็นตัวเชื่อมในการประมาณค่าในแบบสอบย่อย ฉบับที่ 1 ถึงฉบับที่ 4 เพื่อการอธิบายความสามารถผู้สอบตามระดับช่วงชั้นด้วยกระบวนการของ vertical equating ผลการออกแบบแบบสอบย่อยปรากฏดังแผนภาพที่ 4.5 ถึง 4.8 ดังนี้

แผนภาพที่ 4.5 ฝั่งกำหนดจำนวนข้อสอบในแบบสอบย่อยการอ่านคำ ฉบับที่ 1 ถึงฉบับที่ 4

ข้อสอบช่วงชั้น	ป. 1-3	ข้อสอบ ร่วม	ป. 4-6	ข้อสอบ ร่วม	ม.1-ม.3	ข้อสอบ ร่วม	ม.4-ม.6
ฉบับที่1 (รวม 40)							
พยัญชนะ	10	5					
คำ	35	5					
ฉบับที่2 (รวม 45)							
พยัญชนะ		5	5	-			
คำ		5	35	5			
ฉบับที่3 (รวม 45)							
คำ				5	35	5	
ฉบับที่4 (รวม 45)							
คำ						5	40

แผนภาพที่ 4.6 ฝั่งกำหนดจำนวนข้อสอบในแบบสอบย่อยการสะกดคำ ฉบับที่ 1 ถึงฉบับที่ 4

ข้อสอบช่วงชั้น	ป. 1-3	ข้อสอบ ร่วม	ป. 4-6	ข้อสอบ ร่วม	ม.1-ม.3	ข้อสอบ ร่วม	ม.4-ม.6
ฉบับที่1 (รวม 40)							
พยัญชนะ	10	5					
คำ	20	5					
ฉบับที่2 (รวม 45)							
พยัญชนะ		5	5	5			
คำ		5	20				
ฉบับที่3 (รวม 45)							
คำ				5	25	5	
ฉบับที่4 (รวม 45)							
คำ						5	35

แผนภาพที่ 4.7 ฝั่งกำหนดจำนวนข้อสอบในแบบสอบย่อยความเข้าใจประโยค ฉบับที่ 1 ถึงฉบับที่ 4

ข้อสอบช่วงชั้น	ป. 1-3	ข้อสอบ รวม	ป. 4-6	ข้อสอบ รวม	ม.1-ม.3	ข้อสอบ รวม	ม.4-ม.6
ฉบับที่1 (รวม 20) ประโยค	15	5					
ฉบับที่2 (รวม 25) ประโยค		5	15	5			
ฉบับที่3 (รวม 25) ประโยค				5	15	5	
ฉบับที่4 (รวม 30) ประโยค						5	25

แผนภาพที่ 4.8 ฝั่งกำหนดจำนวนข้อสอบในแบบสอบย่อยการคำนวณทางคณิตศาสตร์ ฉบับที่ 1 ถึงฉบับที่ 4

ข้อสอบช่วงชั้น	ป. 1-3	ข้อสอบ รวม	ป. 4-6	ข้อสอบ รวม	ม.1-ม.3	ข้อสอบ รวม	ม.4-ม.6
ฉบับที่1 (รวม 25)	20	5					
ฉบับที่2 (รวม 30)		5	20	5			
ฉบับที่3 (รวม 30)				5	20	5	
ฉบับที่4 (รวม 30)						5	25

2.2 ตรวจสอบความเรียบร้อยของข้อสอบทุกข้อในคลังแบบสอบย่อยทุกระดับ จัดเข้าชุดเป็นฉบับตามระดับช่วงชั้น จัดเรียงลำดับข้อสอบตามค่าประมาณความยาก (b) ตรวจสอบเฉลยคำตอบ และเขียนคู่มือการบริหารการสอบ

2.3 จัดอบรมผู้ช่วยนักวิจัยที่ทำหน้าที่ให้การสอบเพื่อให้การสอบเป็นไปตามมาตรฐานของการพัฒนาแบบสอบมาตรฐาน เมื่อวันที่ 5 -6 มกราคม พ.ศ. 2556 จำนวน 23 คน

2.4 นำแบบสอบไปเก็บข้อมูลการตอบข้อสอบของกลุ่มตัวอย่างนักเรียน และวิเคราะห์ตามเทคนิคการเทียบมาตราคะแนนแบบแนวตั้ง (Vertical Equating Technique) กลุ่มตัวอย่างนักเรียนได้

จากการสุ่มหลายขั้นตอน เป็นตัวแทนนักเรียนชั้นประถมศึกษาปีที่ 1 ถึงมัธยมศึกษาปีที่ 6 ของโรงเรียนใน กรุงเทพมหานครและจังหวัดขอนแก่น ที่กำลังศึกษาภาคปลาย ปีการศึกษา 2555 ตามข้อกำหนดของ โครงการวิจัย จำนวนทั้งสิ้น 1,248 คน

2.5 ข้อมูลคะแนนจากการสอบนำมาวิเคราะห์ด้วยโปรแกรม BILOG-MG ITEM MAINTENANCE PROGRAM: LOGISTIC ITEM RESPONSE MODEL (SCIENTIFIC SOFTWARE INTERNATIONAL, INC. 2002) กำหนดให้กลุ่มผู้สอบในแบบสอบฉบับที่ 3 (ช่วงชั้นมัธยมศึกษาปีที่ 1-3) เป็นกลุ่มอ้างอิง นำผลการวิเคราะห์มาพิจารณาความเหมาะสมของข้อสอบจากค่าความยาก (b) ค่าอำนาจ จำแนก (a) และคุณลักษณะของแบบสอบ (Test information curve : TFC) ที่อธิบายช่วงความสามารถ ในการวัดแต่ละแบบสอบย่อย

2.6 การพัฒนาแบบสอบด้วยเทคนิคของการปรับแบบสอบแต่ละระดับให้เป็นไปตามโครงสร้าง Grade-to-Grade Growth Model โดยอาศัยระเบียบวิธีการ vertical equating พิจารณาผลการ วิเคราะห์ข้อสอบหลายรอบ ต้องตัดข้อสอบบางข้อ หรือข้อมูลการตอบของคนบางคนตามความจำเป็น เพื่อเข้า เจ็อนไขของทฤษฎีการตอบสนองข้อสอบ (Item response theory) สุดท้ายได้ชุดแบบสอบทักษะ พื้นฐานทางวิชาการ KBAST ประกอบด้วยแบบสอบย่อย 4 ทักษะ แต่ละทักษะจัดแบ่งตามระดับความ ยากตามช่วงชั้น เป็นแบบสอบ 4 ฉบับ สรุปจำนวนข้อในชุดแบบสอบทักษะพื้นฐานทางวิชาการ ดังตาราง ที่ 4.11 และ ค่าความยาก (b) ดังแสดงในตารางที่ 4.12

ตารางที่ 4.11 จำนวนข้อสอบในชุดแบบสอบ KBAST จำแนกตามแบบสอบย่อย ในแต่ละฉบับ

แบบสอบย่อย	จำนวน รวมทั้งฉบับ	การอ่านคำ	การสะกดคำ	ความเข้าใจประโยค	การคำนวณ-
ชุดแบบสอบ		Word Reading	Spelling	Sentence Comprehension	คณิตศาสตร์ Math Computation
ฉบับที่ 1	102	36	30	16	20
ฉบับที่ 2	113	38	30	20	25
ฉบับที่ 3	115	35	30	22	28
ฉบับที่ 4	118	35	33	20	30

ตารางที่ 4.12 จำนวนข้อสอบ ค่าความยากสูงสุดและค่าความยากต่ำสุด ของข้อสอบ

แบบสอบย่อย ชุดแบบสอบ	การอ่านคำ N ($b_{\max}-b_{\min}$)	การสะกดคำ N ($b_{\max}-b_{\min}$)	เข้าใจประโยค N ($b_{\max}-b_{\min}$)	การคิดคำนวณ N ($b_{\max}-b_{\min}$)
ฉบับที่ 1	36 (-0.415, -7.49)	30 (-0.791, -8.357)	16 (-1.116, -4.265)	20 (1.517, -2.175)
ฉบับที่ 2	38 (-0.133,-7.49)	30 (2.996, -8.357)	20 (1.592, -3.811)	25 (1.369, -2.232)
ฉบับที่ 3	35 (-0.163, -3.761)	30 (2.996, -3.917)	22 (1.592, -2.632)	28 (1.216, -2.067)
ฉบับที่ 4	35 (-0.408, -4.926)	30 (-0.582, -2.764)	20 (3.186, -2.686)	30 (1.645, -0.841)

2.7 ข้อสอบในชุดแบบสอบ KBAST แต่ละแบบสอบย่อย ได้จัดเรียงตามลำดับความยาก เพื่อสร้างกฎการยุติการสอบและนับคะแนนการตอบ ดังนี้ กำหนดระดับเส้นฐานความสามารถ (Basal level) เมื่อผู้สอบทำข้อสอบในแบบสอบย่อยของฉบับใดๆ ผิดติดต่อกัน 5 ข้อ ให้ถือเป็นข้อยุติการสอบแบบสอบย่อยฉบับนั้นๆ การนับคะแนนการทำข้อสอบถูกต้อง จะนับจากจำนวนข้อที่อยู่เหนือระดับเส้นฐานความสามารถ ขณะเดียวกันข้อสอบที่อยู่ใต้เส้นฐานความสามารถ เรียกว่าระดับเพดาน (Ceiling Level) เป็นข้อสอบที่อยู่ระดับสูงเกินความสามารถของผู้สอบนั้นๆ การนับคะแนนข้อสอบที่ผู้สอบนั้นทำในระดับเพดาน จะไม่ได้คะแนน ถือว่าคำตอบไม่ถูกต้อง

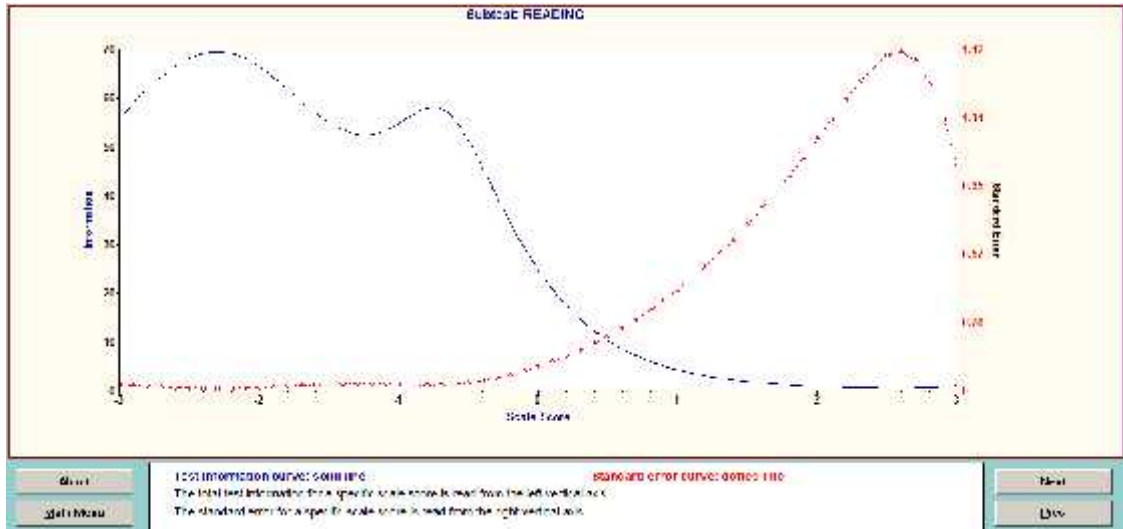
2.8 ผลการนำแบบสอบไปใช้กับกลุ่มตัวอย่าง

ได้ค่าประมาณความสามารถ () ของกลุ่มตัวอย่างพัฒนาแบบสอบมาตรฐานทักษะพื้นฐาน แต่ละแบบสอบย่อย โดยการวิเคราะห์ด้วยโปรแกรม BILOG ได้ค่าสถิติ ดังตารางที่ 4.13 ถึง 4.16 และแผนภาพแสดงลักษณะเส้นโค้ง Test Information ของแต่ละแบบสอบย่อย ดังแผนภาพที่ 4.9 ถึง 4.12 ดังนี้

ตารางที่ 4.13 ค่าสถิติของค่าประมาณความสามารถ () ในแบบสอบย่อยทักษะการอ่านคำ

แบบสอบ	Mean	SD	Median	Skewness	Kurtosis	N
รวม	-0.16	1.20	-0.15	-0.75	1.24	1,239
ฉบับที่ 1	-0.89	1.39	-0.84	-0.52	0.48	307
ฉบับที่ 2	-0.27	1.20	-0.35	-1.06	1.86	312
ฉบับที่ 3	0	1.00	-0.01	.005	-.06	311
ฉบับที่ 4	-0.01	0.81	-0.07	-0.17	-1.16	309

แผนภาพที่ 4.9 Test information function ของแบบย่อยสอบทักษะการอ่านคำ

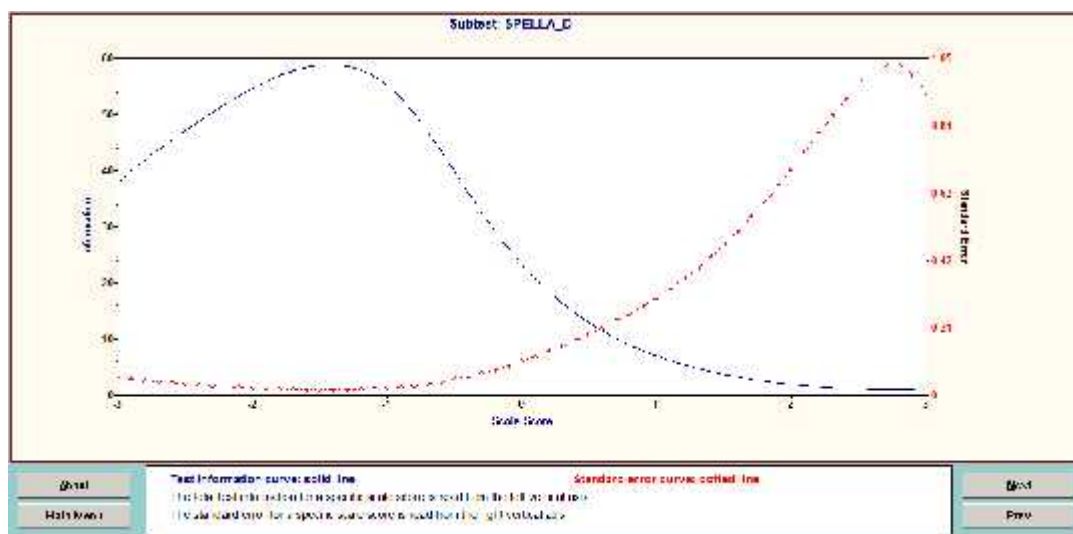


จากเส้นโค้งลักษณะของ Test information ของแบบสอบทักษะการอ่านคำ แสดงถึงช่วงความเชื่อมั่นของการวัดบุคคลที่มีความสามารถ () อยู่ระหว่าง -3 ถึง -0.5 โดยประมาณ และจะมีความคลาดเคลื่อนสูงสุดเมื่อวัดบุคคลที่มีความสามารถ สูงถึง +2.5 โดยประมาณ

ตารางที่ 4.14 ค่าสถิติของค่าประมาณความสามารถ () ในแบบสอบย่อยทักษะการสะกดคำ

แบบสอบ	Mean	SD	Median	Skewness	Kurtosis	N
รวม	-0.15	1.22	-0.08	-0.64	1.11	1236
ฉบับที่ 1	-0.87	1.39	-0.85	-0.52	0.50	308
ฉบับที่ 2	-0.15	1.28	-0.16	-0.09	0.39	311
ฉบับที่ 3	0.00	1.00	-0.08	0.13	0.34	310
ฉบับที่ 4	0.43	0.75	0.55	-0.87	0.05	307

แผนภาพที่ 4.10 Test information function ของแบบย่อยสอบทักษะการสะกดคำ

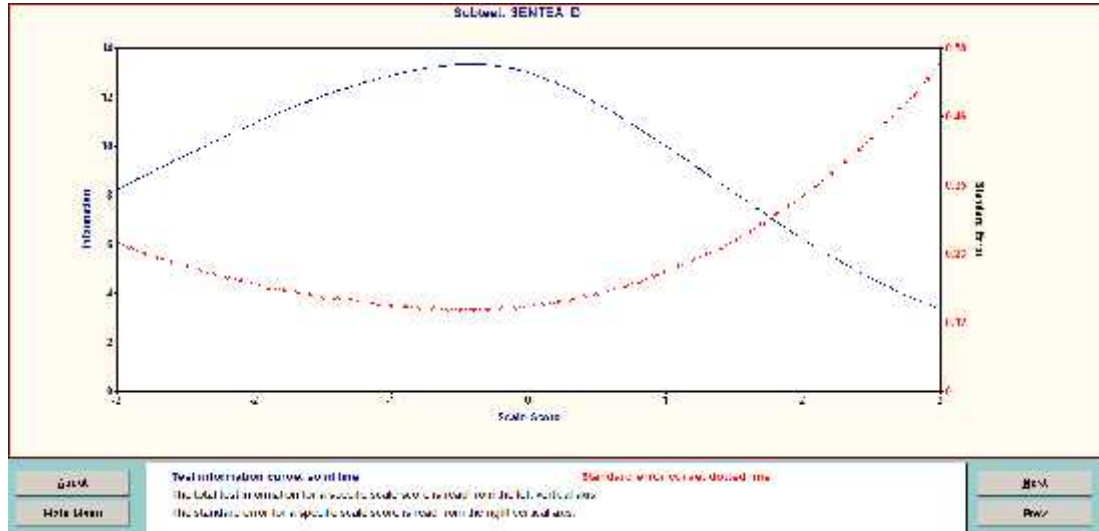


จากเส้นโค้งลักษณะของ Test information ของแบบสอบทักษะการสะกดคำ แสดงถึงช่วงความเชื่อมั่นเมื่อการวัดบุคคลที่มีความสามารถ () อยู่ระหว่าง -2.55 ถึง -0.5 โดยประมาณ และจะมีความคลาดเคลื่อนสูงสุดเมื่อวัดบุคคลที่มีความสามารถ สูง +2.5 ถึง +3.0 โดยประมาณ

ตารางที่ 4.15 ค่าสถิติของค่าประมาณความสามารถ () ในแบบสอบย่อยทักษะความเข้าใจประโยค

แบบสอบ	Mean	SD	Median	Skewness	Kurtosis	N
รวม	-0.21	1.43	-0.10	-0.43	1.53	1191
ฉบับที่ 1	-1.03	1.88	-1.1	0.26	0.92	271
ฉบับที่ 2	-0.16	1.39	-0.1	0.05	0.69	304
ฉบับที่ 3	-0.01	1.06	0.09	-0.48	1.06	309
ฉบับที่ 4	0.27	0.94	0.31	-0.08	-0.11	307

แผนภาพที่ 4.11 Test information function ของแบบย่อยสอบทักษะความเข้าใจประโยค

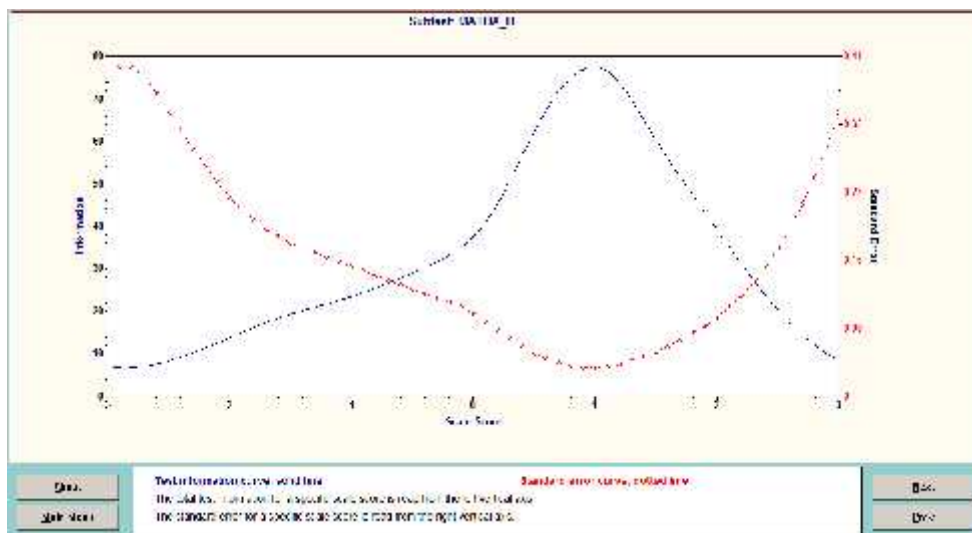


จากเส้นโค้งลักษณะของ Test information ของแบบสอบความเข้าใจประโยค แสดงถึงช่วงความเชื่อมั่นเมื่อการวัดบุคคลที่มีความสามารถ () อยู่ระหว่าง -1.0 ถึง +1.0 โดยประมาณ และจะมีความคลาดเคลื่อนสูงที่สุดเมื่อวัดบุคคลที่มีความสามารถ สูง ถึง +3.0 โดยประมาณ ขณะเดียวกันความคลาดเคลื่อนก็กระจายอยู่ตลอดช่วงความสามารถ

ตารางที่ 4.16 ค่าสถิติของค่าประมาณความสามารถ () ในแบบสอบทักษะการคำนวณทางคณิตศาสตร์

แบบสอบ	Mean	SD	Median	Skewness	Kurtosis	N
รวม						1239
ฉบับที่ 1	-2.51	1.91	1.97	-0.54	-0.16	311
ฉบับที่ 2	-0.60	1.14	0.60	0.12	1.18	311
ฉบับที่ 3	0.00	1.00	0.17	-0.38	-0.39	310
ฉบับที่ 4	0.38	1.05	0.49	-0.52	0.00	307

แผนภาพที่ 4.12 Test information function ของแบบย่อยสอบทักษะการคำนวณทางคณิตศาสตร์



จากเส้นโค้งลักษณะของ Test information ของแบบสอบทักษะการคำนวณทางคณิตศาสตร์ แสดงถึงช่วงความเชื่อมั่นเมื่อการวัดบุคคลที่มีความสามารถ () อยู่ระหว่าง +.5 ถึง +1.5 โดยประมาณ และจะมีความคลาดเคลื่อนสูงสุดเมื่อวัดบุคคลที่มีความสามารถ ต่ำกว่า 0 และสูงกว่า +2.5 โดยประมาณ

ตอนที่ 3 ความตรงของแบบสอบมาตรฐาน

3.1 ความตรงเชิงเนื้อหา (Content validity)

หลักฐานความตรงเชิงเนื้อหา มีขั้นตอนการพัฒนาอย่างเป็นระบบดังได้แสดงในตอนที่ 1 ผลการสร้างข้อสอบวัดทักษะพื้นฐานทางวิชาการ มีหลักฐานสำคัญ ดังนี้

1. นิยามและข้อกำหนดลักษณะเฉพาะของข้อสอบในแต่ละทักษะได้ผ่านการพิจารณาของผู้เชี่ยวชาญว่ามีความเหมาะสมและสอดคล้องกับเป้าหมายของชุดแบบสอบ KBAST
2. ข้อสอบทุกข้อผ่านการตรวจสอบคุณภาพของข้อสอบ โดยการวิเคราะห์เชิงเหตุผลและการใช้ข้อมูลเชิงประจักษ์
3. ข้อสอบมีค่าสถิติ IRT แสดงลักษณะเฉพาะ (a, b) และมีค่าเฉลี่ยที่ถูกต้องและได้จัดเก็บข้อสอบในระบบคลังข้อสอบ พร้อมสร้างชุดแบบสอบตามโครงสร้างชุดแบบสอบ

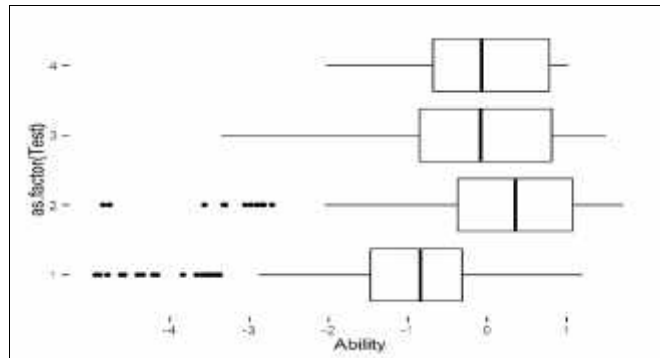
3.2 ความตรงเชิงโครงสร้าง (Construct validity)

จากการออกแบบโครงสร้างชุดแบบสอบ KBAST ซึ่งประกอบด้วยแบบสอบย่อย 4 ทักษะ แต่ละทักษะจัดข้อสอบตามระดับความยากให้สอดคล้องกับระดับช่วงชั้นที่ 1 ถึง 4 เป็นฉบับที่ 1 – 4 ตามลำดับ ดังนั้น ถ้าค่าประมาณความสามารถของผู้สอบในแต่ละฉบับมีปริมาณที่ลดหลั่นลงไปอย่างสอดคล้อง สมเหตุสมผล จะแสดงว่าชุดแบบสอบมีโครงสร้างที่สอดคล้องกับ Grade to grade growth model

การตรวจสอบความตรงจากหลักฐานเชิงประจักษ์นี้วิเคราะห์ข้อมูล 2 แบบ คือ 1) พิจารณาจากภาพลักษณะการกระจายของค่าความสามารถ () ของกลุ่ม 2) ทดสอบความมีนัยสำคัญทางสถิติ Kruskal-Wallis ดังนี้

แบบสอบย่อยทักษะด้านการอ่านคำ (Reading Subtest)

แผนภาพที่ 4.13 ภาพลักษณะการกระจายของค่าความสามารถ () ของกลุ่มที่สอบแบบสอบย่อย ทักษะการอ่านคำ



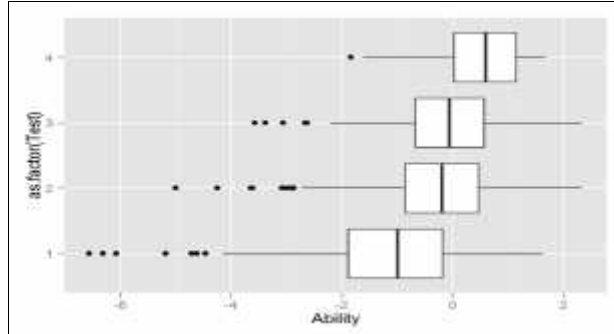
ตารางที่ 4.17 Reading Subtest : the Kruskal-Wallis test

แบบสอบที่เปรียบเทียบ	ค่ามัธยฐาน Median	ค่าสถิติ Adjusted H	d.f.	P value
แบบสอบฉบับที่ 1-4	-0.15	143.999	3	5.19E-31*
ฉบับที่ 1 กับฉบับที่ 2	-0.84, 0.35	112.085	1	3.42E-26*
ฉบับที่ 1 กับฉบับที่ 3	-0.84, -0.01	69.321	1	8.37E-17*
ฉบับที่ 1 กับฉบับที่ 4	-0.84, -0.07	68.743	1	1.12E-16*
ฉบับที่ 2 กับฉบับที่ 3	0.35, -0.01	21.475	1	3.59E-6*
ฉบับที่ 2 กับฉบับที่ 4	0.35, -0.07	19.434	1	0.00001*
ฉบับที่ 3 กับฉบับที่ 4	-0.01, -0.07	0.012	1	0.914

จากการพิจารณาค่าเฉลี่ยและแผนภาพลักษณะการกระจายของค่าความสามารถ () ของกลุ่มที่สอบแบบสอบย่อยทักษะการอ่านคำแต่ละกลุ่ม พบว่า ค่าเฉลี่ย () ของแบบสอบฉบับที่ 2 มีมากกว่าฉบับที่ 3 และ 4 ดังตารางที่ 4.17 เมื่อทดสอบด้วยสถิติพารามเมตริก Kruskal-Wallis พบความแตกต่างอย่างมีนัยสำคัญระหว่างแบบสอบย่อยทุกฉบับ แต่ทั้งนี้ เมื่อดูจากค่ามัธยฐานคู่ของฉบับที่ 2 กับ 3 และ 2 กับ 4 เป็นค่าที่สลับกัน ส่วนคู่ฉบับ 3 และ 4 ไม่พบนัยสำคัญของความแตกต่าง จึงทำให้ไม่สามารถสรุปว่า แบบสอบย่อยทักษะด้านการอ่านคำมีความตรงเชิงโครงสร้างอย่างสมบูรณ์ การพัฒนาแบบสอบครั้งนี้จำเป็นต้องทบทวนกระบวนการอย่างรอบคอบและพิจารณาค่าสถิติ IRT ต่อไป

แบบสอบย่อยทักษะด้านการสะกดคำ (Spelling Subtest)

แผนภาพที่ 4.14 ภาพลักษณะการกระจายของค่าความสามารถ () ของกลุ่มที่สอบแบบสอบย่อยทักษะการสะกดคำ



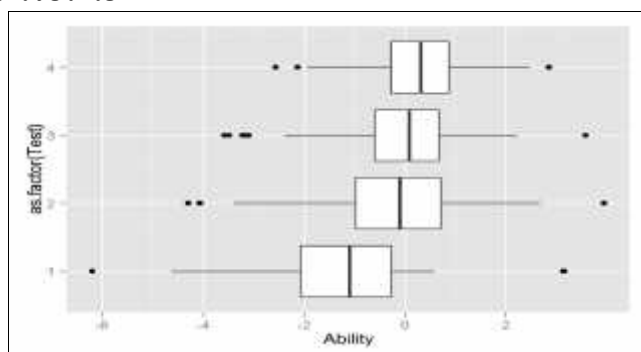
ตารางที่ 4.18 Spelling Subtest : the Kruskal-Wallis test

แบบสอบที่เปรียบเทียบ	ค่ามัธยฐาน Median	ค่าสถิติ Adjusted H	d.f.	P value
แบบสอบฉบับที่ 1-4	-0.08	222.414	3	6.04E-48*
ฉบับที่ 1 กับฉบับที่ 2	-0.85, -0.16	46.26	1	1.04E-11*
ฉบับที่ 1 กับฉบับที่ 3	-0.85, -0.08	80.586	1	2.78E-19*
ฉบับที่ 1 กับฉบับที่ 4	-0.85, 0.55	184.242	1	5.75E-42*
ฉบับที่ 2 กับฉบับที่ 3	-0.16, -0.08	3.257	1	0.071
ฉบับที่ 2 กับฉบับที่ 4	-0.16, 0.55	75.605	1	3.46E-18*
ฉบับที่ 3 กับฉบับที่ 4	-0.08, 0.55	59.392	1	1.29E-14*

จากการพิจารณาค่าเฉลี่ยและแผนภาพการกระจายความสามารถ () ของกลุ่มผู้สอบแต่ละกลุ่มพบว่า มีปริมาณของความสามารถที่เพิ่มขึ้นจากฉบับที่ 1 ถึง 4 เมื่อพิจารณาการวิเคราะห์ความแตกต่างด้วยสถิติ непараметрический Kruskal-Wallis พบความแตกต่างอย่างมีนัยสำคัญระหว่างฉบับ ยกเว้นแบบสอบย่อยฉบับที่ 2 และ 3 มีความแตกต่างเล็กน้อย ($p=0.071$) จึงสรุปได้ว่า แบบสอบย่อยทักษะด้านการสะกดคำมีความตรงเชิงโครงสร้างเกือบสมบูรณ์ การตรวจสอบรายละเอียดภายในแบบสอบฉบับที่ 2 และ 3 อาจได้คำอธิบายเพื่อพัฒนาต่อไป

แบบสอบย่อยทักษะด้านความเข้าใจประโยค (Sentence Comprehension Subtest)

แผนภาพที่ 4.15 ภาพลักษณะการกระจายของค่าความสามารถ () ของกลุ่มที่สอบแบบสอบความเข้าใจประโยค



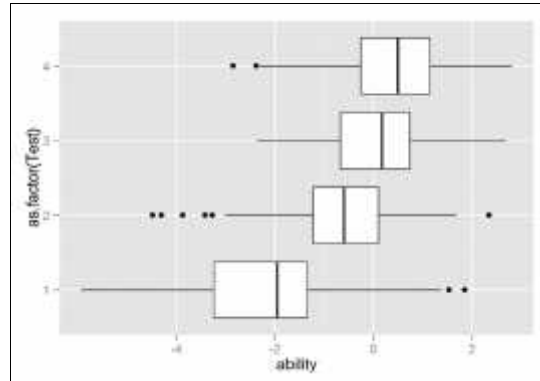
ตารางที่ 4.19 Sentence Comprehension Subtest : the Kruskal-Wallis test

แบบสอบที่เปรียบเทียบ	ค่ามัธยฐาน Median	ค่าสถิติ Adjusted H	d.f.	P value
แบบสอบฉบับที่ 1-4	-0.10	146.42	3	1.56E-31*
ฉบับที่ 1 กับฉบับที่ 2	-1.1, -0.1	52.489	1	4.33E-13*
ฉบับที่ 1 กับฉบับที่ 3	-1.1, 0.09	86.653	1	1.29E-20*
ฉบับที่ 1 กับฉบับที่ 4	-1.1, 0.31	131.909	1	1.57E-30*
ฉบับที่ 2 กับฉบับที่ 3	-0.1, 0.09	2.763	1	0.096
ฉบับที่ 2 กับฉบับที่ 4	-0.1, 0.31	16.972	1	0.000038*
ฉบับที่ 3 กับฉบับที่ 4	0.09, 0.31	8.07	1	0.004501*

จากการพิจารณาค่าเฉลี่ยและแผนภาพการกระจายความสามารถ () ของกลุ่มผู้สอบแต่ละกลุ่มพบว่า มีปริมาณของความสามารถที่เพิ่มขึ้นจากฉบับที่ 1 ถึง 4 เมื่อพิจารณาการวิเคราะห์ความแตกต่างด้วยสถิติพหุพารามetric Kruskal-Wallis พบความแตกต่างอย่างมีนัยสำคัญระหว่างฉบับ ยกเว้นแบบสอบย่อยฉบับที่ 2 และ 3 มีความแตกต่างเล็กน้อย ($p=0.096$) จึงสรุปได้ว่า แบบสอบย่อยทักษะด้านความเข้าใจประโยคมีความตรงเชิงโครงสร้างเกือบสมบูรณ์ ควรตรวจสอบรายละเอียดภายในแบบสอบฉบับที่ 2 และ 3 อาจได้คำอธิบายเพื่อพัฒนาต่อไป

แบบสอบย่อยทักษะด้านการคำนวณทางคณิตศาสตร์ (Math Computation Subtest)

แผนภาพที่ 4.16 ค่าประมาณความสามารถเฉลี่ยของแบบสอบย่อยทักษะด้านการคำนวณทางคณิตศาสตร์



ตารางที่ 4.20 Math Computation Subtest : the Kruskal-Wallis test

แบบสอบที่เปรียบเทียบ	ค่ามัธยฐาน Median	ค่าสถิติ Adjusted H	d.f.	P value
แบบสอบฉบับที่ 1-4		506.734	3	1.66E-109*
ฉบับที่ 1 กับฉบับที่ 2	1.97, 0.60	207.148	1	5.76E-47*
ฉบับที่ 1 กับฉบับที่ 3	1.97, 0.17	296.157	1	2.26E-66*
ฉบับที่ 1 กับฉบับที่ 4	1.97, 0.49	334.243	1	1.14E-74*
ฉบับที่ 2 กับฉบับที่ 3	0.60, 0.17	55.083	1	1.16E-13*
ฉบับที่ 2 กับฉบับที่ 4	0.60, 0.49	119.458	1	8.31E-28*
ฉบับที่ 3 กับฉบับที่ 4	0.17, 0.49	21.549	1	3.45E-6*

จากการพิจารณาค่าเฉลี่ยและแผนภาพการกระจายความสามารถ () ของกลุ่มผู้สอบแต่ละกลุ่ม พบว่า มีปริมาณของความสามารถที่เพิ่มขึ้นเป็นลำดับ สอดคล้องกับโครงสร้างแบบสอบ เมื่อพิจารณาการวิเคราะห์ความแตกต่างด้วยสถิติ непараметрич Kruskal-Wallis พบว่าแต่ละคู่ของแบบสอบมีความแตกต่างอย่างมีนัยสำคัญ จึงสรุปได้ว่า แบบสอบย่อยทักษะด้านการคำนวณทางคณิตศาสตร์มีความตรงเชิงโครงสร้างอย่างสมบูรณ์

สรุปผลการวิเคราะห์ความตรงเชิงโครงสร้างของแบบสอบย่อย

5. แบบสอบย่อยการอ่านคำ มีความตรงเชิงโครงสร้างไม่สมบูรณ์ เนื่องด้วยค่าเฉลี่ยของแบบสอบฉบับที่ 2 มีค่าสูงกว่าฉบับที่ 3 และ 4 การพัฒนาแบบสอบครั้งนี้จำเป็นต้องทบทวนกระบวนการอย่างรอบคอบ ค้นหาจุดอ่อนเพื่อปรับปรุงต่อไป

6. แบบสอบย่อยการสะกดคำ มีความตรงเชิงโครงสร้างเกือบสมบูรณ์ ควรตรวจสอบรายละเอียดภายในแบบสอบฉบับที่ 2 และ 3 เพื่อค้นหาจุดอ่อนเพื่อปรับปรุงต่อไป

7. แบบสอบย่อยความเข้าใจประโยค มีความตรงเชิงโครงสร้างเกือบสมบูรณ์ ควรตรวจสอบรายละเอียดภายในแบบสอบฉบับที่ 2 และ 3 เพื่อค้นหาจุดอ่อนเพื่อปรับปรุง

8. แบบสอบย่อยการคำนวณทางคณิตศาสตร์ มีความตรงเชิงโครงสร้างอย่างสมบูรณ์

ตอนที่ 4 ความเที่ยงของแบบสอบมาตรฐาน (Reliability)

ความเที่ยงแบบความคงที่ภายใน แสดงถึงข้อสอบภายในแบบสอบชุดหนึ่งวัดความสามารถด้านหนึ่ง ผลการวัดจากข้อสอบเหล่านั้นให้ผลการวัดตามที่แบบสอบกำหนดได้อย่างคงเส้นคงวาเพียงไร ค่าประมาณนี้ได้จากการวิเคราะห์ผลการตอบข้อสอบแต่ละข้อของกลุ่มตัวอย่างที่ทำแบบสอบชุดนั้นๆ การวิเคราะห์นี้ใช้ค่าสัมประสิทธิ์ความเที่ยง Cronbach's coefficient alpha (1951) ผลการวิเคราะห์ค่าความเที่ยงของชุดแบบสอบ KBAST แสดงในตารางที่ 4.21 จำแนกตามแบบสอบย่อยการอ่าน การสะกดคำ ความเข้าใจประโยค และการคำนวณ ซึ่งแต่ละแบบสอบย่อย ประกอบด้วยแบบสอบที่วัดตามระดับช่วงชั้น ฉบับที่ 1 (ระดับชั้นประถมปีที่ 1-3) ฉบับที่ 2 (ระดับชั้นประถมปีที่ 4-6) ฉบับที่ 3 (ระดับชั้นมัธยมศึกษาปีที่ 1-3) และฉบับที่ 4 (ระดับชั้นมัธยมศึกษาปีที่ 4-6)

ตารางที่ 4.21 สัมประสิทธิ์ความคงที่ภายใน (r) และค่าความคลาดเคลื่อนมาตรฐาน (SEM) ของการวัด

แบบสอบฉบับที่ 1-4		การอ่านคำ Word Reading	การสะกดคำ Spelling	ความเข้าใจประโยค Sentence Comprehension	การคำนวณทาง คณิตศาสตร์ Math Computation
ฉบับที่	r	.939	.883	.790	.917
	SEM	0.4789	0.4557	0.6993	0.3729
1	จำนวนข้อ	36	30	16	20
	จำนวนคน	307	310	310	311
ฉบับที่	r	.928	.914	.839	.915
	SEM	0.5019	0.3960	0.5599	0.2698
2	จำนวนข้อ	38	30	20	25
	จำนวนคน	312	311	311	311
ฉบับที่	r	.848	.864	.774	.944
	SEM	0.3819	0.4169	0.5271	0.2331
3	จำนวนข้อ	35	30	22	28
	จำนวนคน	311	311	310	310
ฉบับที่	r	.808	.858	.680	.950
	SEM	0.3472	0.4745	0.5589	0.2502
4	จำนวนข้อ	35	30	20	30
	จำนวนคน	309	307	307	307

สรุปผลการวิเคราะห์ค่าสัมประสิทธิ์ความเที่ยงของแต่ละแบบสอบย่อย ดังตารางที่ 4.22

ตารางที่ 4.22 สัมประสิทธิ์ความคงที่ภายใน (r) ของแต่ละแบบสอบย่อย ฉบับที่ 1 – ฉบับที่ 4

แบบสอบย่อย	ค่าความเที่ยง (r)
การอ่านคำ (Word Reading)	0.808 - 0.939
การสะกดคำ (Spelling)	0.858 - 0.914
ความเข้าใจประโยค (Sentence Comprehension)	0.680 - 0.839
การคำนวณทางคณิตศาสตร์ (Math Computation)	0.915 - 0.950

สรุปค่าความเที่ยงของแบบสอบ KBAST โดยรวมอยู่ในระดับค่อนข้างสูง (> 0.80) แบบสอบทักษะการคำนวณทางคณิตศาสตร์ ฉบับที่ 4 มีค่าสัมประสิทธิ์ความเที่ยงสูงสุด เท่ากับ 0.950 มีความคลาดเคลื่อนมาตรฐานของการวัดต่ำที่สุด คือเท่ากับ 0.2502

ส่วนแบบสอบทักษะความเข้าใจประโยคอยู่ในระดับค่อนข้างต่ำ โดยเฉพาะฉบับที่ 4 มีค่าสัมประสิทธิ์ความเที่ยงเท่ากับ 0.680 และ มีความคลาดเคลื่อนมาตรฐานของการวัดสูงที่สุด คือเท่ากับ 0.5589

ตอนที่ 5 การพัฒนาเกณฑ์มาตรฐาน

กระบวนการพัฒนาเกณฑ์ปกติของชุดแบบวัด KBAST ได้ออกแบบเพื่อผลิตคะแนนมาตรฐาน และ percentile ranks จำแนกตามระดับชั้น และช่วงชั้นของแต่ละแบบสอบย่อย ฉบับที่ 1 ถึง 4

การเก็บข้อมูลการสอบใช้กลุ่มตัวอย่างสร้างเกณฑ์มาตรฐานที่สุ่มจากประชากรนักเรียนชั้นประถมศึกษาปีที่ 1 ถึงชั้นมัธยมศึกษาปีที่ 6 จากกรุงเทพมหานครและจังหวัดขอนแก่น ตามที่ได้กำหนดไว้ในโครงการวิจัย จากการสุ่มตัวอย่างด้วยวิธีสุ่มหลายขั้นตอน ได้ตัวอย่างนักเรียนดังตารางที่ 4.23 ดังนี้

ตารางที่ 4.23 จำนวนตัวอย่างนักเรียนจากการสุ่มหลายขั้นตอน

ระดับชั้น	กทม.	ขอนแก่น	รวม	ระดับชั้น	กทม.	ขอนแก่น	รวม
ชั้นประถมศึกษาปีที่ 1	48	56	104	ชั้นมัธยมศึกษาปีที่ 1	48	56	104
ชั้นประถมศึกษาปีที่ 2	48	56	104	ชั้นมัธยมศึกษาปีที่ 2	48	56	104
ชั้นประถมศึกษาปีที่ 3	48	56	104	ชั้นมัธยมศึกษาปีที่ 3	48	56	104
ชั้นประถมศึกษาปีที่ 4	48	56	104	ชั้นมัธยมศึกษาปีที่ 4	48	56	104
ชั้นประถมศึกษาปีที่ 5	48	56	104	ชั้นมัธยมศึกษาปีที่ 5	48	56	104
ชั้นประถมศึกษาปีที่ 6	48	56	104	ชั้นมัธยมศึกษาปีที่ 6	48	56	104
รวม	288	336	624	รวม	288	336	624

จากการตรวจสอบความสมบูรณ์ของชุดแบบสอบบรมผู้ช่วยนักวิจัย ในการดำเนินการสอบแล้ว นำผลการสอบมาตรวจให้คะแนน คำตอบที่ถูกให้ 1 คะแนน และคำตอบที่ผิดให้ 0 คะแนน ตรวจทานข้อมูลซ้ำอย่างรอบคอบ จึงนำไปเข้ากระบวนการวิเคราะห์คุณภาพข้อสอบ และความสามารถผู้สอบด้วยโปรแกรมคอมพิวเตอร์ BILOG เช่นเดียวกับการสร้างแบบสอบ ผลการวิเคราะห์พบว่า การประมาณความสามารถของผู้สอบบางคนไม่สามารถประมาณได้ ด้วยเหตุที่มาจากผลการตอบรายข้อไม่ได้เป็นไปตามข้อกำหนดของทฤษฎีการตอบสนองข้อสอบ จึงได้ตัดผู้สอบที่มีพฤติกรรมการตอบข้อสอบไม่สอดคล้องกับเงื่อนไขของทฤษฎีการตอบสนองข้อสอบออกไปจำนวนหนึ่ง นำคะแนนสอบของนักเรียนมา

ประมวลผลเป็นคะแนนรวมรายบุคคล หาค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐานเป็นกลุ่มตามระดับชั้น และช่วงชั้น ผลการประมวลผลคะแนน แสดงในตารางที่ 4.24

ตารางที่ 4.24 ค่าสถิติของกลุ่มตัวอย่างสร้างเกณฑ์มาตรฐาน

แบบ สอบ	ค่าสถิติ	การอ่านคำ Word Reading	การสะกดคำ Spelling	ความเข้าใจประโยค Sentence Comprehension	การคำนวณทาง คณิตศาสตร์ Math Computation
ฉบับที่ 1	N-items	36	30	16	20
	Mean	30.67	26.51	12.31	12.01
	SD.	6.707	4.201	3.154	5.449
ฉบับที่ 2	N-items	38	30	20	25
	Mean	33.97	22.86	12.12	12.66
	SD.	5.800	5.951	4.393	6.330
ฉบับที่ 3	N-items	35	30	22	28
	Mean	30.44	24.98	12.43	14.43
	SD.	4.287	4.387	3.983	8.331
ฉบับที่ 4	N-items	35	33	20	30
	Mean	31.19	30.01	11.55	15.43
	SD.	3.714	3.815	3.154	9.061

จากนั้นแปลงคะแนนแต่ละกลุ่มระดับชั้นและช่วงชั้นให้เป็นคะแนนมาตรฐาน (scaled score) มีค่าเฉลี่ยเท่ากับ 100 และคะแนนเบี่ยงเบนมาตรฐานเท่ากับ 15 แล้วสร้างตารางเทียบคะแนนดิบสู่คะแนนมาตรฐานรายชั้น ประถมศึกษาปีที่ 1 ถึงมัธยมศึกษาปีที่ 6 (Raw score Subtest Standard Score Conversion by Grade) ทั้งหมด 16 ตาราง นำเสนอในภาคผนวก ข ชื่อตารางมี ดังนี้

ตารางเกณฑ์ปกติที่ 1	แบบสอบฉบับที่ 1 (Form A)	ชั้นประถมศึกษาปีที่ 1
ตารางเกณฑ์ปกติที่ 2	แบบสอบฉบับที่ 1 (Form A)	ชั้นประถมศึกษาปีที่ 2
ตารางเกณฑ์ปกติที่ 3	แบบสอบฉบับที่ 1 (Form A)	ชั้นประถมศึกษาปีที่ 3
ตารางเกณฑ์ปกติที่ 4	แบบสอบฉบับที่ 2 (Form B)	ชั้นประถมศึกษาปีที่ 4
ตารางเกณฑ์ปกติที่ 5	แบบสอบฉบับที่ 2 (Form B)	ชั้นประถมศึกษาปีที่ 5
ตารางเกณฑ์ปกติที่ 6	แบบสอบฉบับที่ 2 (Form B)	ชั้นประถมศึกษาปีที่ 6
ตารางเกณฑ์ปกติที่ 7	แบบสอบฉบับที่ 3 (Form C)	ชั้นมัธยมศึกษาปีที่ 1
ตารางเกณฑ์ปกติที่ 8	แบบสอบฉบับที่ 3 (Form C)	ชั้นมัธยมศึกษาปีที่ 2
ตารางเกณฑ์ปกติที่ 9	แบบสอบฉบับที่ 3 (Form C)	ชั้นมัธยมศึกษาปีที่ 3
ตารางเกณฑ์ปกติที่ 10	แบบสอบฉบับที่ 4 (Form D)	ชั้นมัธยมศึกษาปีที่ 4
ตารางเกณฑ์ปกติที่ 11	แบบสอบฉบับที่ 4 (Form D)	ชั้นมัธยมศึกษาปีที่ 5
ตารางเกณฑ์ปกติที่ 12	แบบสอบฉบับที่ 4 (Form D)	ชั้นมัธยมศึกษาปีที่ 6
ตารางเกณฑ์ปกติที่ 13	แบบสอบฉบับที่ 1 (Form A)	ช่วงชั้นประถมศึกษาปีที่ 1 -3
ตารางเกณฑ์ปกติที่ 14	แบบสอบฉบับที่ 2 (Form B)	ช่วงชั้นประถมศึกษาปีที่ 4 -6
ตารางเกณฑ์ปกติที่ 15	แบบสอบฉบับที่ 3 (Form C)	ช่วงชั้นมัธยมศึกษาปีที่ 1 -3
ตารางเกณฑ์ปกติที่ 16	แบบสอบฉบับที่ 4 (Form D)	ช่วงชั้นมัธยมศึกษาปีที่ 4 -6

สรุปผลการพัฒนาแบบสอบทักษะพื้นฐานทางวิชาการ KBAST

จากวัตถุประสงค์ของการวิจัยที่กำหนดไว้ว่าเพื่อพัฒนาแบบสอบทักษะพื้นฐานทางวิชาการ เพื่อการประเมินและวินิจฉัยบุคคลที่มีภาวะสมาธิสั้น บกพร่องทางการเรียนรู้ และออทิซึม กระบวนการพัฒนาได้นำเสนอในระเบียบวิธีดำเนินการวิจัย และผลการพัฒนาได้นำเสนอในผลการวิจัย จึงขอสรุปผลการพัฒนาแบบสอบทักษะพื้นฐานทางวิชาการ เป็น 3 ตอน คือ

ตอนที่ 1 ผลผลิตของการพัฒนาแบบสอบ KBAST

ตอนที่ 2 หลักฐานคุณภาพแบบสอบ KBAST

ตอนที่ 3 คู่มือการใช้แบบสอบ KBAST

ตอนที่ 1 ผลผลิตของการพัฒนาแบบสอบ KBAST

1.1 ลักษณะแบบสอบมาตรฐานวัดทักษะพื้นฐานทางวิชาการ

(1) แบบสอบทักษะพื้นฐานทางวิชาการ KBAST เป็นชุดแบบสอบ (Batteries) ที่สร้างขึ้นเพื่อประเมินทักษะพื้นฐานทางวิชาการของผู้เรียนระดับประถมศึกษาปีที่ 1 ถึง มัธยมศึกษาปีที่ 6 ที่เป็นพื้นฐานของการอ่านออกเขียนได้ และคิดเลขได้ (Literacy) ชุดแบบสอบ KBAST ประกอบด้วยแบบสอบย่อยวัดทักษะ 4 ด้าน คือ

แบบสอบการอ่านคำ (word reading) เป็นการวัดความสามารถของบุคคลในการจำตัวอักษร พยัญชนะ และ คำ ผู้สอบจะอ่านออกเสียงคำที่กำหนดต่อหน้าผู้ให้การทดสอบ

แบบสอบการสะกดคำ (word spelling) เป็นการวัดความสามารถของบุคคลในการเปลี่ยนเสียงคำที่ได้ยินไปสู่การเขียน ผู้สอบจะฟังเสียง แล้วเขียนตามทั้งตัวอักษรและคำ

แบบสอบความเข้าใจประโยค (sentence comprehension) เป็นการวัดความสามารถในการรู้ความหมายคำและเข้าใจความคิดและข้อมูลที่อยู่ในประโยค ด้วยการเติมคำในช่องว่างเพื่อให้ประโยคมีใจความครบสมบูรณ์

แบบสอบการคิดคำนวณทางคณิตศาสตร์ (math computation) เป็นการวัดความสามารถของบุคคลในการคำนวณด้านคณิตศาสตร์พื้นฐาน ด้วยวิธีการนับ ระบุตัวเลข แก้ปัญหาอย่างง่าย ด้วยการเขียนคำตอบ

(2) การจัดชุดแบบสอบย่อยรายทักษะ ตามรูปแบบโครงสร้างความเพิ่มพูนของความสามารถตามช่วงชั้น (Grade-to-Grade Growth model) ชุดแบบสอบ (Test Batteries) วัดทักษะพื้นฐานทาง

วิชาการของประชากรนักเรียนชั้นประถมศึกษาปีที่ 1 ถึง มัธยมศึกษาปีที่ 6 ใน 4 ทักษะ กำหนดใช้แบบสอบฉบับที่ 1 2 3 และ 4 กับนักเรียนที่เรียนอยู่ช่วงชั้นที่ 1 2 3 และ 4 ตามลำดับ ชุดแบบสอบ KBAST จึงประกอบด้วยแบบสอบย่อย 4 ทักษะ ดังนี้

- แบบสอบย่อยการอ่านคำ ฉบับที่ 1 ฉบับที่ 2 ฉบับที่ 3 และฉบับที่ 4
- แบบสอบย่อยการสะกดคำ ฉบับที่ 1 ฉบับที่ 2 ฉบับที่ 3 และฉบับที่ 4
- แบบสอบย่อยความเข้าใจประโยค ฉบับที่ 1 ฉบับที่ 2 ฉบับที่ 3 และฉบับที่ 4
- แบบสอบย่อยการคำนวณทางคณิตศาสตร์ ฉบับที่ 1 ฉบับที่ 2 ฉบับที่ 3 และฉบับที่ 4

แบบสอบย่อยทั้ง 4 ทักษะจะจัดเข้าเป็นฉบับตามระดับชั้นของผู้เรียน ชุดแบบสอบแต่ละฉบับจึงประกอบด้วยแบบสอบย่อยรายทักษะทั้ง 4 ในระดับเดียวกัน ดังนี้

- แบบสอบฉบับที่ 1 สำหรับผู้สอบช่วงชั้นที่ 1 (ระดับชั้นประถมศึกษาปีที่ 1 ถึง 3)
- แบบสอบฉบับที่ 2 สำหรับผู้สอบช่วงชั้นที่ 2 (ระดับชั้นประถมศึกษาปีที่ 4 ถึง 6)
- แบบสอบฉบับที่ 3 สำหรับผู้สอบช่วงชั้นที่ 3 (ระดับชั้นมัธยมศึกษาปีที่ 1 ถึง 3)
- แบบสอบฉบับที่ 4 สำหรับผู้สอบช่วงชั้นที่ 4 (ระดับชั้นมัธยมศึกษาปีที่ 4 ถึง 6)

เอกสารแบบสอบ ออกแบบการพิมพ์เพื่อสะดวกในการนำไปใช้ ดังเช่นภาพที่แสดงข้างล่าง

นี้

ภาพหน้าแรกของแบบสอบทักษะพื้นฐานทางวิชาการ KBAST



หน้าแรก เป็นข้อมูลผู้สอบ : ชื่อ นามสกุล เพศ วันเดือนปีเกิด
ข้อมูลเกี่ยวกับการสอบ : วันที่ทดสอบ ชื่อผู้ให้การ
ทดสอบ

ผลการสอบ : คะแนนดิบรายทักษะย่อย และ
คะแนนมาตรฐานที่เทียบจากตาราง
คะแนนเกณฑ์ปกติ

หน้าที่สอง แบบสอบย่อย การสะกดคำ

หน้าที่สาม แบบสอบย่อย ความเข้าใจประโยค

หน้าที่สี่ แบบสอบย่อย การคิดคำนวณทางคณิตศาสตร์

สำหรับแบบสอบย่อยการอ่าน เป็นเอกสารแยก ที่ผู้ดำเนินการสอบจะนำมาให้ผู้สอบอ่านในเวลาและจัดเก็บทันที

แบบสอบทั้ง 4 ฉบับ ได้จัดพิมพ์ ด้วยสีที่แตกต่างเพื่อความสะดวกในการใช้

แบบสอบฉบับที่ 1 สีเขียว

แบบสอบฉบับที่ 2 สีแดง

แบบสอบฉบับที่ 3 สีฟ้า

แบบสอบฉบับที่ 4 สีม่วง

(3) มีตารางคะแนนเกณฑ์ปกติของแบบสอบตามระดับชั้น (Grade Norms) เป็นตารางที่นำเสนอชุดของคะแนนดิบ (Raw Scores) ที่ได้จากการสอบรายทักษะที่มีค่าตรงกับคะแนนมาตรฐานเกณฑ์ปกติ (Standard Scaled Score) ของระดับชั้นนั้นๆ ซึ่งมีค่าเฉลี่ยเท่ากับ 100 และคะแนนเบี่ยงเบนมาตรฐานเท่ากับ 15 ความหมายของคะแนนมาตรฐานที่เทียบได้ แปรตามมาตรฐานของการแจกแจงปกติ ดังนี้ คะแนนมาตรฐานเกณฑ์ปกติ ตั้งแต่ 69 ลงไป ต่ำมาก, 70-79 ต่ำ, 80-89 ต่ำกว่าค่าเฉลี่ย, 9-109 อยู่ระดับค่าเฉลี่ย, 110-119 สูงกว่าค่าเฉลี่ย, 120-129 สูง, และตั้งแต่ 130 ขึ้นไป สูงมาก

เมื่อนักเรียนคนหนึ่งๆ ได้รับการทดสอบ ผู้ดำเนินการสอบจะตรวจให้คะแนนตามค่าเฉลี่ย นับรวมคะแนนที่ทำได้ แล้วนำคะแนนดิบมาเทียบคะแนนเกณฑ์ปกติของระดับชั้นนั้นๆ ซึ่งบอกระดับความสามารถเมื่อเทียบกับเกณฑ์ปกติของระดับชั้นนั้น นอกจากนี้ นำคะแนนดิบไปเทียบในตารางเกณฑ์ปกติของช่วงชั้นได้ ตัวอย่าง นักเรียนชั้นประถมศึกษาปีที่หนึ่งคนหนึ่ง ทำแบบสอบ ฉบับที่ 1 ได้คะแนนแบบสอบย่อย การอ่าน 22 คะแนน เมื่อเทียบคะแนนเกณฑ์มาตรฐานในตารางเกณฑ์ปกติที่ 1 (ภาคผนวก ข) ได้คะแนนมาตรฐานตามระดับชั้นประถมศึกษาปีที่ 1 เท่ากับ 84 แสดงว่า นักเรียนคนนี้มีความสามารถในด้านการอ่านอยู่ในระดับต่ำกว่าคะแนนเฉลี่ยของกลุ่มนักเรียนระดับชั้นประถมศึกษาปีที่ 1 และเมื่อนำคะแนนของนักเรียนคนนี้ไปเทียบกับตารางเกณฑ์ปกติที่ 13 ของช่วงชั้นที่ 1 (ประถมศึกษาปีที่ 1 ถึง ปีที่ 3) ได้ 75 แสดงว่า นักเรียนคนนี้มีความสามารถในด้านการอ่านอยู่ในระดับต่ำกว่าคะแนนเฉลี่ยของกลุ่มนักเรียนช่วงชั้นที่ 1

ตารางคะแนนเกณฑ์ปกติของแบบสอบที่จัดทำมีทั้งหมด 16 ตาราง แสดงในภาคผนวก ข ดังเช่นตัวอย่าง คะแนนจากแบบสอบฉบับที่หนึ่ง ใช้ตารางเกณฑ์ปกติที่ 1 2 หรือ 3 และ 13 ดังนี้

ตารางเกณฑ์ปกติที่ 1 เทียบคะแนนดิบสู่คะแนนมาตรฐาน แบบสอบฉบับที่ 1 ชั้นประถมศึกษาปีที่ 1

Word Reading		Spelling		Sentence Comprehension		Math Computation	
Raw score	Scaled score	Raw score	Scaled score	Raw score	Scaled score	Raw score	Scaled score
0	55	0	55	0	55	0	64
1	55	1	55	1	75	1	76
2	55	2	55	2	79	2	84
3	55	3	55	3	83	3	91
4	70	4	55	4	87	4	98
5	70	5	55	5	91	5	103
6	71	6	70	6	95	6	108
7	71	7	71	7	98	7	113
8	71	8	72	8	101	8	117
9	71	9	73	9	104	9	121
10	72	10	73	10	107	10	125
11	72	11	75	11	110	11	128
12	73	12	75	12	113	12	128
13	73	13	78	13	116	13	135
14	74	14	80	14	119	14	138
15	75	15	80	15	122	15	141
16	75	16	84	16	124	16	145
17	75	17	86			17	145
18	78	18	88			18	145
19	79	19	90			19	145
20	81	20	93			20	145
21	82	21	95				
22	84	22	98				
23	85	23	100				
24	85	24	103				
25	89	25	106				
26	91	26	109				
27	94	27	113				
28	96	28	116				
29	98	29	119				
30	101	30	123				
31	104						
32	107						
33	110						
34	113						
35	117						
36	120						

ตารางเกณฑ์ปกติที่ 2 เทียบคะแนนดิบสู่คะแนนมาตรฐาน แบบสอบฉบับที่ 1 ชั้นประถมศึกษาปีที่ 2

Word Reading		Spelling		Sentence Comprehension		Math Computation	
Raw score	Scaled score	Raw score	Scaled score	Raw score	Scaled score	Raw score	Scaled score
0	55	0	55	0	55	0	66
1	55	1	55	1	70	1	66
2	55	2	55	2	71	2	66
3	55	3	55	3	72	3	81
4	55	4	69	4	74	4	84
5	55	5	69	5	76	5	88
6	66	6	69	6	79	6	91
7	66	7	69	7	82	7	95
8	66	8	69	8	85	8	98
9	66	9	69	9	88	9	101
10	66	10	69	10	92	10	104
11	67	11	69	11	97	11	106
12	67	12	70	12	101	12	109
13	67	13	70	13	106	13	112
14	68	14	71	14	111	14	114
15	68	15	71	15	117	15	117
16	68	16	72	16	123	16	119
17	68	17	73			17	121
18	70	18	73			18	124
19	70	19	73			19	126
20	70	20	77			20	128
21	71	21	79				
22	72	22	81				
23	74	23	84				
24	76	24	87				
25	77	25	91				
26	80	26	95				
27	82	27	99				
28	85	28	105				
29	88	29	111				
30	92	30	117				
31	96						
32	101						
33	107						
34	113						
35	119						
36	145						

ตารางเกณฑ์ปกติที่ 3 เทียบคะแนนดิบสู่คะแนนมาตรฐาน แบบสอบฉบับที่ 1 ชั้นประถมศึกษาปีที่ 3

Word Reading		Spelling		Sentence Comprehension		Math Computation	
Raw score	Scaled score	Raw score	Scaled score	Raw score	Scaled score	Raw score	Scaled score
0	55	0	55	0	55	0	55
1	55	1	55	1	55	1	55
2	55	2	55	2	69	2	76
3	55	3	55	3	69	3	81
4	55	4	55	4	69	4	81
5	62	5	55	5	70	5	88
6	62	6	55	6	71	6	88
7	62	7	55	7	73	7	95
8	62	8	55	8	75	8	98
9	62	9	55	9	78	9	101
10	62	10	55	10	81	10	104
11	62	11	69	11	85	11	106
12	62	12	69	12	90	12	109
13	62	13	69	13	96	13	112
14	62	14	69	14	103	14	114
15	63	15	69	15	111	15	117
16	63	16	69	16	121	16	119
17	63	17	69			17	121
18	63	18	69			18	124
19	63	19	70			19	126
20	63	20	70			20	128
21	63	21	71				
22	66	22	73				
23	67	23	74				
24	67	24	77				
25	67	25	80				
26	67	26	84				
27	67	27	89				
28	76	28	95				
29	79	29	104				
30	82	30	115				
31	86						
32	91						
33	96						
34	103						
35	110						
36	118						

ตารางเกณฑ์ปกติที่ 13 เทียบคะแนนดิบสู่คะแนนมาตรฐาน แบบสอบฉบับที่ 1 ช่วงชั้นประถมศึกษา
ปีที่ 1-3

Word Reading		Spelling		Sentence Comprehension		Math Computation	
Raw score	Scaled score	Raw score	Scaled score	Raw score	Scaled score	Raw score	Scaled score
0	55	0	55	0	55	0	66
1	55	1	55	1	72	1	71
2	55	2	55	2	73	2	76
3	55	3	55	3	75	3	81
4	68	4	70	4	76	4	84
5	68	5	70	5	79	5	88
6	68	6	70	6	81	6	91
7	68	7	70	7	84	7	95
8	68	8	71	8	87	8	98
9	68	9	71	9	90	9	101
10	68	10	71	10	94	10	104
11	68	11	71	11	98	11	106
12	68	12	72	12	102	12	109
13	68	13	72	13	106	13	112
14	69	14	73	14	111	14	114
15	69	15	74	15	116	15	117
16	69	16	75	16	121	16	119
17	70	17	76			17	121
18	71	18	78			18	124
19	72	19	79			19	126
20	72	20	81			20	128
21	73	21	83				
22	75	22	85				
23	76	23	88				
24	77	24	91				
25	79	25	94				
26	81	26	98				
27	83	27	102				
28	86	28	107				
29	89	29	112				
30	92	30	117				
31	95						
32	99						
33	103						
34	108						
35	113						
36	119						

1.2 เวลาในการบริหารการสอบ

เนื่องจากวัตถุประสงค์ของการใช้แบบสอบ KBAST เพื่อประเมินทักษะพื้นฐานทางวิชาการของรายบุคคลที่เลือกมา เพื่อใช้ผลการประเมินเป็นข้อมูลประกอบการวางแผนการจัดการศึกษาให้บุคคลนั้น เวลาที่ใช้ในการบริหารการสอบจึงไม่ได้กำหนดตายตัว แต่จะขึ้นอยู่กับลักษณะผู้สอบที่มี อายุ ทักษะ และรูปแบบพฤติกรรมที่แตกต่างกัน โดยทั่วไป จะใช้เวลาประมาณ 30 – 45 นาที แต่อาจต้องใช้เวลามากหรือน้อยแตกต่างกันตามแต่กรณี เช่น กรณีที่ผู้สอบเป็นบุคคลที่มีภาวะบกพร่องทางการเรียนรู้ สมาธิสั้น หรือออทิสซึม จำเป็นต้องทำความเข้าใจผู้สอบก่อน เวลาที่ใช้อาจมากขึ้น เป็นต้น

1.3 การนำแบบสอบทักษะพื้นฐานทางวิชาการ KBAST ไปใช้

ตามเจตนารมณ์ของการพัฒนาแบบสอบทักษะพื้นฐานทางวิชาการ KBAST ก็เพื่อให้มีเครื่องมือที่มีคุณภาพสนับสนุนการทำงานช่วยเหลือนักเรียนที่มีภาวะสมาธิสั้น บกพร่องทางการเรียนรู้ และออทิสซึม จึงเชิญชวนผู้เชี่ยวชาญ และผู้เกี่ยวข้องนำแบบสอบมาตรฐานนี้ไปใช้ นำผลการประเมินร่วมพิจารณากับข้อมูลทางการประเมินทางจิตวิทยาอื่นๆ ประวัติภูมิหลัง ผลการสังเกตพฤติกรรม และการประเมินอย่างไม่เป็นทางการอื่นๆ จะช่วยในการระบุและวินิจฉัยความบกพร่องของกรณีที่ต้องให้ความช่วยเหลือได้ละเอียดถูกต้องยิ่งขึ้น นอกจากนี้ ยังให้ประโยชน์อีกหลายด้าน เช่น เป็นเครื่องมือสำหรับนักจิตวิทยา นักการศึกษา และนักประเมินในการเก็บข้อมูลเบื้องต้นเพื่อพัฒนาโครงการให้ความช่วยเหลือต่างๆ ใช้การสอบเพื่อคัดกรอง เพื่อค้นหารายบุคคลที่ต้องรับการประเมินที่ละเอียดมากขึ้นเพื่อแก้ปัญหาทางด้านวิชาการ ประเมินซ้ำเพื่อการวินิจฉัยรายบุคคลในความบกพร่องทางการเรียนรู้ และ/หรือ ทางความคิด ใช้เป็นเครื่องมือประเมิน ความบกพร่องของผลสัมฤทธิ์ทางการเรียนเพื่อระบุความบกพร่องทางการเรียนรู้ เป็นต้น

ตอนที่ 2 หลักฐานคุณภาพแบบสอบ KBAST

ขั้นการสร้างมาตรฐานการวัดของแบบสอบ (Standardization) เป็นขั้นตอนที่ใช้กระบวนการพัฒนาแบบสอบตามทฤษฎีการวัดความสามารถ การตรวจสอบคุณสมบัติแบบสอบ เป็นข้อกำหนดที่สำคัญของการพัฒนาแบบสอบมาตรฐาน มาตรฐานที่ใช้กำหนด ตาม *Standards for Educational and Psychological Testing*. (American Educational Research Association (AERA), American Psychological Association (APA), & National Council on Measurement in Education (NCME): 1999) ที่สำคัญ ได้แก่ ความตรง (Validity) และ ความเที่ยง (Reliability) ในการพัฒนาแบบสอบมาตรฐาน KBAST นี้ ได้ดำเนินการภายในกรอบปฏิบัติการที่จะให้เกิดความเป็นมาตรฐานและนำเสนอหลักฐาน ดังนี้

2.1 ความตรง

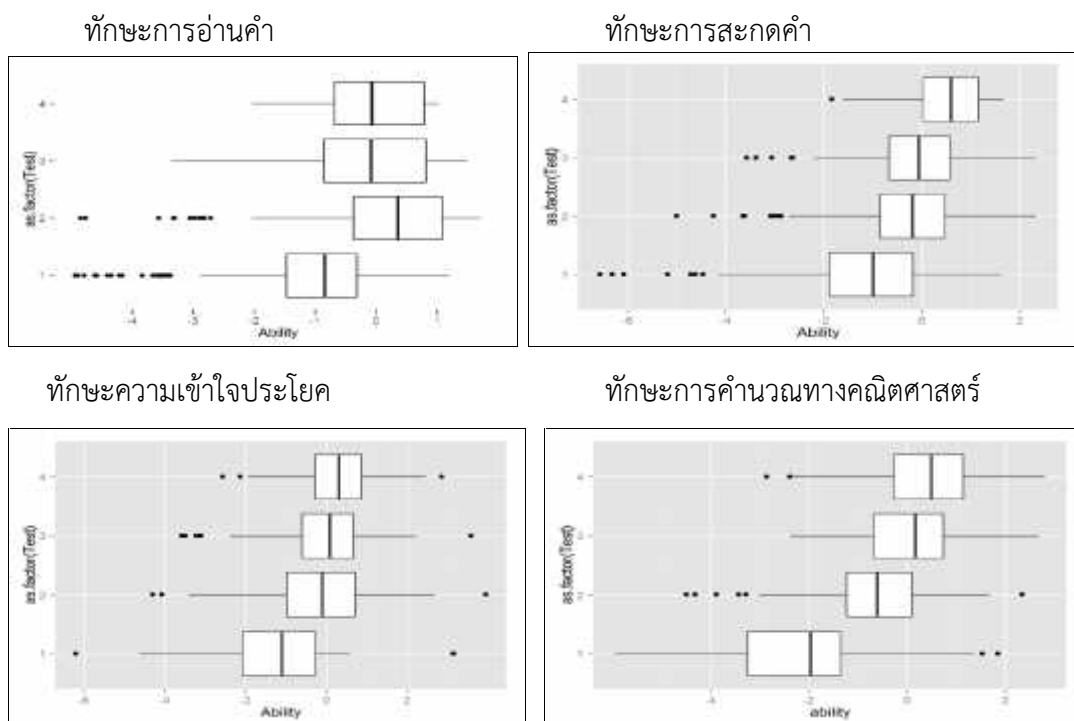
(1) หลักฐานภายในของความตรง (Internal Evidence of validity) เป็นหลักฐานที่เกิดขึ้นตั้งแต่ขั้นตอนแรก เมื่อวางแผนการพัฒนาเป็นกระบวนการที่มีความละเอียด รอบคอบ ตรวจสอบทบทวนหลายครั้งจนมีความชัดเจนในการกำหนดเนื้อหาที่สอดคล้องกับวัตถุประสงค์ของโครงการสอบ และคุณลักษณะที่ต้องการวัดในกลุ่มประชากรเป้าหมายที่กำหนด สรุปลงเป็นแผนกำหนดลักษณะของข้อสอบแล้วดำเนินการเป็นขั้นเป็นตอน แสดงหลักฐานความตรงเชิงเนื้อหา (Content Validity) ดังนี้ มีนิยามและข้อกำหนดลักษณะเฉพาะของข้อสอบในแต่ละทักษะที่ได้ผ่านการพิจารณาของผู้เชี่ยวชาญว่ามีความเหมาะสมและสอดคล้องกับเป้าหมายของชุดแบบสอบ KBAST และพิจารณาความตรงเชิงเหตุผลของข้อสอบและแบบสอบย่อยแต่ละชุดที่ออกมา ดังตารางที่ 4.1 ฝั่งกำหนดลักษณะแบบสอบ ทักษะพื้นฐานทางวิชาการ และ ตารางที่ 4.2 จำนวนข้อสอบในกลุ่มข้อสอบที่ผ่านกระบวนการวิเคราะห์คุณภาพเบื้องต้น และ สรุปลงจำนวนและเนื้อหาข้อสอบในแต่ละทักษะย่อยดังตารางที่ 4.3 ถึง 4.6 นอกจากนี้ข้อสอบทุกข้อได้ผ่านกระบวนการวิเคราะห์เชิงปริมาณด้วย IRT แสดง ลักษณะเฉพาะ (a, b) และได้จัดเก็บข้อสอบในระบบคลังข้อสอบ พร้อมสร้างชุดแบบสอบตามโครงสร้างชุดแบบสอบ

(2) หลักฐานภายนอกของความตรง (External Evidence of validity) เป็นหลักฐานที่มาจาก การหาความสอดคล้องของการวัดภายในชุดแบบสอบว่ามีคุณลักษณะตามโครงสร้างที่ออกแบบไว้หรือไม่ สำหรับชุดแบบสอบ KBAST ประกอบด้วยแบบสอบย่อย 4 ทักษะ แต่ละทักษะจัดแบ่งเป็นฉบับตามระดับช่วงชั้น คือ ฉบับที่ 1 มีความยากของข้อสอบตามหลักสูตรช่วงชั้นที่ 1 (ประถมศึกษาปีที่ 1 ถึง 3) ฉบับที่ 2 มีความยากตามช่วงชั้นที่ 2 (ประถมศึกษาปีที่ 4 ถึง 6) ฉบับที่ 3 มีความยากตามช่วงชั้นที่ 3 (มัธยมศึกษาปีที่ 1 ถึง 3) และ ฉบับที่ 4 มีความยากตามช่วงชั้นที่ 4 (มัธยมศึกษาปีที่ 4 ถึง 6) การออกแบบนี้ กำหนดโครงสร้างความเพิ่มพูนของความสามารถตามช่วงชั้น (Grade to Grade Growth Model) กระบวนการพัฒนาให้ได้ข้อสอบที่ fit กับ model และ ตรวจสอบความสอดคล้องของแบบสอบแต่ละฉบับ ได้ใช้ทฤษฎี Item-response theory การสร้างมาตราคะแนนมาตรฐาน (Scaling) ให้เป็นคะแนนปกติมาตรฐาน (Standard Scaled Scores) และเทคนิคการปรับเทียบคะแนนต่างฉบับตามระดับชั้น ด้วย Vertical Equating Techniques เมื่อสิ้นสุดกระบวนการ ได้ตรวจสอบความสอดคล้องภายในแบบสอบแต่ละทักษะดั่งนิยามปฏิบัติการที่กำหนด

นิยาม ความตรงเชิงโครงสร้างของแบบสอบรายทักษะ (Construct Validity)
หมายถึง ค่าความสามารถเฉลี่ย (\bar{X}) จากการใช้แบบสอบย่อยรายทักษะกับกลุ่มตัวอย่างมาตรฐานการพัฒนาแบบสอบ มีลักษณะสอดคล้องกับรูปแบบการเพิ่มพูนทักษะการเรียนรู้ขั้นต่อขั้น (Grade to Grade Growth Model) กล่าวคือ ค่าความสามารถเฉลี่ย แบบสอบฉบับที่ 1 (\bar{X}_1) น้อยกว่าค่าความสามารถเฉลี่ย แบบสอบฉบับที่ 2 (\bar{X}_2) ค่าความสามารถเฉลี่ยแบบสอบฉบับที่ 2 (\bar{X}_2) น้อยกว่าค่าความสามารถเฉลี่ยแบบสอบฉบับที่ 3 (\bar{X}_3) ค่าความสามารถเฉลี่ยแบบสอบฉบับที่ 3 (\bar{X}_3) น้อยกว่าค่าความสามารถ

เฉลี่ยแบบสอบฉบับที่ 4 (4) [$\mu_1 < \mu_2 < \mu_3 < \mu_4$] หลักฐานเชิงสถิติ ใช้การทดสอบความแตกต่างของผลสอบของกลุ่มตามระดับช่วงชั้น ด้วยสถิติทดสอบ : the Kruskal-Wallis test ความตรงเชิงโครงสร้างของแบบสอบรายทักษะนี้เป็นหลักฐานความตรงจากการตรวจสอบคุณลักษณะทางการวัด (Measurement Properties) ของแบบสอบย่อย ที่เกิดจากการใช้ทฤษฎีการตอบสนองข้อสอบ ตรวจสอบคุณสมบัติทางการวัดผลที่สำคัญของรายข้อ และชุดแบบสอบ ผ่านกระบวนการเทียบมาตรฐานของแบบสอบ (test equating) เพื่อสร้างชุดแบบสอบที่วัดความสามารถ () ตามรูปแบบการวัดพัฒนาการของผู้สอบ (grade to grade growth model)

สรุปผลการวิเคราะห์ความตรงเชิงโครงสร้างของแบบสอบย่อย 4 ทักษะ แสดงด้วยแผนภาพค่าประมาณความสามารถเฉลี่ยของแบบสอบรายทักษะ ดังนี้



แบบสอบย่อยทักษะการอ่าน แผนภาพลักษณะการกระจายของค่าความสามารถ () แบบสอบฉบับที่ 2 มีความไม่สอดคล้องอย่างชัดเจน จึงทำให้ไม่สามารถสรุปว่า แบบสอบย่อยทักษะด้านการอ่านคำมีความตรงเชิงโครงสร้างอย่างสมบูรณ์

แบบสอบย่อยทักษะด้านการสะกดคำ แผนภาพลักษณะการกระจายของค่าความสามารถ () พบว่า มีปริมาณของความสามารถเพิ่มขึ้นจากฉบับที่ 1 ถึง 4 ตามลำดับ จึงสรุปได้ว่า แบบสอบย่อยทักษะด้านการสะกดคำมีความตรงเชิงโครงสร้างสมบูรณ์

แบบสอบย่อยทักษะความเข้าใจประโยค แผนภาพลักษณะการกระจายของค่าความสามารถ () พบว่า มีปริมาณของความสามารถที่เพิ่มขึ้นจากฉบับที่ 1 ถึง 4 ตามลำดับ จึงสรุปได้ว่า แบบสอบย่อยทักษะความเข้าใจประโยคมีความตรงเชิงโครงสร้างสมบูรณ์

แบบสอบย่อยทักษะการคำนวณทางคณิตศาสตร์ แผนภาพลักษณะการกระจายของค่าความสามารถ () พบว่า มีปริมาณของความสามารถที่เพิ่มขึ้นจากฉบับที่ 1 ถึง 4 ตามลำดับ จึงสรุปได้ว่า แบบสอบย่อยทักษะการคำนวณทางคณิตศาสตร์มีความตรงเชิงโครงสร้างสมบูรณ์

2.2 ความเที่ยง

ความเที่ยงแบบความคงที่ภายใน แสดงถึงข้อสอบภายในแบบสอบชุดหนึ่งวัดความสามารถด้านหนึ่ง ผลการวัดจากข้อสอบเหล่านั้นให้ผลการวัดตามที่แบบสอบกำหนดได้อย่างคงเส้นคงวาเพียงไร ค่าประมาณนี้ได้จากการวิเคราะห์ผลการตอบข้อสอบแต่ละข้อของกลุ่มตัวอย่างที่ทำแบบสอบชุดนั้นๆ การวิเคราะห์นี้ใช้ค่าสัมประสิทธิ์ความเที่ยง Cronbach's coefficient alpha (1951) ผลการวิเคราะห์ค่าความเที่ยงของชุดแบบสอบ KBAST สรุป สัมประสิทธิ์ความเที่ยงและความคลาดเคลื่อนของการวัดของแบบสอบย่อยการอ่าน การสะกดคำ ความเข้าใจประโยค และการคำนวณ ซึ่งแต่ละแบบสอบย่อยประกอบด้วยแบบสอบที่วัดตามระดับช่วงชั้น ฉบับที่ 1 (ระดับชั้นประถมปีที่ 1-3) ฉบับที่ 2 (ระดับชั้นประถมปีที่ 4-6) ฉบับที่ 3 (ระดับชั้นมัธยมปีที่ 1-3) และฉบับที่ 4 (ระดับชั้นมัธยมปีที่ 4-6) ได้ดังนี้

ตารางที่ 4.25 สัมประสิทธิ์ความคงที่ภายใน (r) และความคลาดเคลื่อนของการวัดของแบบสอบย่อย

แบบสอบย่อย	ค่าสัมประสิทธิ์ความเที่ยง (r)	ความคลาดเคลื่อนของการวัด (SEM)
การอ่านคำ	0.808 - 0.939	0.3472 - 0.5019
การสะกดคำ	0.858 - 0.914	0.3960 - 0.4745
ความเข้าใจประโยค	0.680 - 0.839	0.5271 - 0.6993
การคำนวณทางคณิตศาสตร์	0.915 - 0.950	0.2331 - 0.3729

สรุปค่าความเที่ยงของแบบสอบ KBAST โดยรวมอยู่ในระดับค่อนข้างสูง (> 0.80) แบบสอบทักษะการคำนวณทางคณิตศาสตร์ ฉบับที่ 4 มีค่าสัมประสิทธิ์ความเที่ยงสูงสุด เท่ากับ 0.950 มีความคลาดเคลื่อนมาตรฐานของการวัดต่ำที่สุด คือเท่ากับ 0.2502 ส่วนแบบสอบทักษะความเข้าใจประโยคอยู่ในระดับค่อนข้างต่ำ โดยเฉพาะฉบับที่ 4 มีค่าสัมประสิทธิ์ความเที่ยงเท่ากับ 0.680 และมีค่าความคลาดเคลื่อนมาตรฐานของการวัดสูงที่สุด คือเท่ากับ 0.5589

ตอนที่ 3 คู่มือบริหารการสอบแบบสอบ KBAST

คู่มือเป็นส่วนประกอบสำคัญของชุดแบบสอบที่พัฒนาขึ้น โดยมีวัตถุประสงค์เพื่อให้รายละเอียดแก่ผู้นำชุดแบบสอบไปใช้ ปฏิบัติได้ถูกต้องตามข้อกำหนดที่เป็นมาตรฐานของชุดแบบสอบ การจัดทำคู่มือการใช้แบบสอบทักษะพื้นฐานทางวิชาการ เป็นข้อกำหนดสำคัญของแบบสอบมาตรฐาน เพื่อให้การนำแบบสอบไปใช้ดำเนินการภายใต้เงื่อนไขที่เป็นมาตรฐาน เช่น นักเรียนที่รับการทดสอบทุกคนจะต้องได้รับการบริหารการสอบแบบสอบที่มีคุณลักษณะ และคุณภาพที่เปรียบเทียบกันได้ ได้รับคำสั่งในการทำแบบสอบที่เหมือนกัน ทำในกรอบเวลาเดียวกัน การตรวจให้คะแนนมีความเป็นปรนัย และการรายงานผลมีความตรงตามผลการประเมิน นอกจากนี้เงื่อนไขของการทำข้อสอบแล้ว ในส่วนตัวผู้ให้การทดสอบ ก็ต้องมีความพร้อมทั้งร่างกาย และอารมณ์ และจัดสภาพแวดล้อมให้เหมาะกับการใช้ความคิดของผู้สอบ ดังนั้น ผู้ใช้ชุดแบบสอบจำเป็นต้องได้รับการฝึก หรือเป็นผู้ที่มีประสบการณ์ด้านนี้ หัวข้อที่บรรจุในคู่มือ ประกอบด้วย

1. บทนำ : แนวคิดการพัฒนา วัตถุประสงค์ และการใช้ชุดแบบสอบ
2. การบริหารการสอบ และการตรวจคะแนน
3. การแปลความหมายผลการสอบ
4. การพัฒนาชุดแบบสอบ

1. บทนำ : แนวคิดการพัฒนา วัตถุประสงค์ และการใช้ชุดแบบสอบ

แบบสอบทักษะพื้นฐานทางวิชาการ KBAST ที่พัฒนาขึ้นเป็นชุดแบบสอบ (Batteries) เพื่อประเมินทักษะพื้นฐานทางวิชาการของผู้เรียนระดับประถมศึกษาปีที่ 1 ถึง มัธยมศึกษาปีที่ 6 ชุดแบบสอบ KBAST ประกอบด้วยแบบสอบย่อยวัดทักษะ 4 ด้าน คือ

แบบสอบการอ่านคำ (word reading) เป็นการวัดความสามารถของบุคคลในการจำตัวอักษร พยัญชนะ และ คำ ผู้สอบจะอ่านออกเสียงคำที่กำหนดต่อหน้าผู้ให้การทดสอบ

แบบสอบการสะกดคำ (word spelling) เป็นการวัดความสามารถของบุคคลในการเปลี่ยนเสียงคำที่ได้ยินไปสู่การเขียน ผู้สอบจะฟังเสียง แล้วเขียนตามทั้งตัวอักษรและคำ

แบบสอบความเข้าใจประโยค (sentence comprehension) เป็นการวัดความสามารถในการรู้ความหมายคำและเข้าใจความคิดและข้อมูลที่อยู่ในประโยค ด้วยการเติมคำในช่องว่างเพื่อให้ประโยคมีใจความครบ สมบูรณ์

แบบสอบการคิดคำนวณทางคณิตศาสตร์ (math computation) เป็นการวัดความสามารถของบุคคลในการคำนวณด้านคณิตศาสตร์พื้นฐาน ด้วยวิธีการนับ ระบุตัวเลข แก้ปัญหาอย่างง่าย ด้วยการเขียนคำตอบ

2. การบริหารการสอบ และการตรวจคะแนน

บทนี้ กล่าวถึง เอกสารและอุปกรณ์ที่จำเป็นต่อการบริหารการสอบ ชุดแบบสอบ KBAST คุณสมบัติของผู้ใช้ คำชี้แจงการบริหาร และการตรวจให้คะแนนการสอบ ผู้ให้การทดสอบจำเป็นต้องสร้างความคุ้นเคยกับวิธีการสอบในแต่ละแบบสอบย่อยก่อนที่จะลงมือดำเนินการสอบ

3. การแปลความหมายผลการสอบ

บทนี้ กล่าวถึง หลักการแปลความหมายผลการสอบ วิธีแปลความหมายผลการสอบ และ ตัวอย่าง

4. การพัฒนาชุดแบบสอบ

บทนี้ กล่าวถึง การพัฒนาเนื้อหาของแบบสอบ ผลการพัฒนาเนื้อหาของแบบสอบ และข้อมูลทางเทคนิค ความเที่ยงแบบความคงที่ภายใน และการสร้างมาตรฐานแบบสอบ KBAST

เอกสารชุดแบบสอบ KBAST

ประกอบด้วย

คู่มือบริหารการสอบ

คู่มือนี้จัดทำขึ้นเพื่อให้ผู้ให้การทดสอบที่มีความเข้าใจสาระสำคัญเกี่ยวกับการดำเนินการสอบ การตรวจให้คะแนน และการแปลความหมายคะแนนดิบจากการสอบในชุดแบบสอบนี้ ประกอบด้วยหลักการพื้นฐานของการพัฒนาชุดแบบสอบ และคุณลักษณะของแบบสอบด้านเทคนิค (technical characteristics) รวมทั้งหลักฐานของความเที่ยง (Reliability) และความตรง (Validity) ของแบบสอบ

เอกสารแบบสอบ สำหรับผู้ให้การทดสอบ

แบบสอบ KBAST ประกอบด้วย แบบสอบย่อย 4 ทักษะ ในแต่ละแบบสอบย่อยจะมีแบบสอบ 4 ฉบับ แยกตามระดับความยากสำหรับผู้รับการทดสอบในช่วงชั้นที่ 1 2 3 และ 4 ผู้ให้การทดสอบจะถือเอกสารเป็นแผ่นข้อสอบเฉพาะฉบับที่ตรงกับรายบุคคลของผู้รับการทดสอบเท่านั้น ในแผ่นข้อสอบ มีชื่อ รหัสผู้รับการทดสอบ มีคำชี้แจงวิธีการสอบ ช่องสำหรับบันทึกคะแนนการตอบรายข้อ และคะแนนรวมของผู้รับการทดสอบ

เอกสารข้อสอบ สำหรับผู้รับการทดสอบ

ผู้รับการทดสอบเมื่อได้รับการกำหนดแล้วว่าจะสอบแบบสอบย่อยใด ในแบบสอบฉบับใดแล้ว ผู้ให้การทดสอบจะมอบเอกสารข้อสอบที่มีการเว้นที่ว่างสำหรับเติมคำในแบบสอบย่อยการเขียนสะกดคำ แบบสอบย่อยความเข้าใจประโยค และแบบสอบย่อยการคิดคำนวณ การเขียนคำตอบลงในเอกสารจะกำหนดตามขั้นตอนดำเนินการที่ผู้ให้การทดสอบจะบริหารการสอบตามคู่มือต่อไป

แผ่นรายการคำเพื่อทดสอบการอ่านคำ

สำหรับแบบสอบย่อยการอ่านคำ จัดทำเป็นแผ่นกระดาษแข็งมีลักษณะเป็นบัตรรายการคำสำหรับผู้รับการทดสอบอ่าน 1 แผ่น ผู้รับการทดสอบจะได้รับคำสั่งให้อ่านไปตามลำดับทีละคำ ขณะที่ผู้ดำเนินการให้การสอบ จะถือแผ่นข้อสอบเดียวกันแต่มีที่ว่างเพื่อระบุชื่อผู้สอบและบันทึกผลการอ่านของผู้สอบนั้นๆ ไปขณะที่ดำเนินการ

แผ่นรายการคำเพื่อทดสอบการเขียนสะกดคำ

แผ่นรายการคำของแบบสอบสะกดคำ จัดทำเป็นแผ่นรายการคำ สำหรับผู้ให้การทดสอบ เพื่ออ่านออกเสียงคำให้ผู้รับการทดสอบฟังและเขียนคำนั้นลงในกระดาษ ที่ให้ไว้ ดำเนินการให้เขียนตามคำบอกที่คู่มือกำหนด

ข้อปฏิบัติในการสอบ

การจัดสภาพแวดล้อมทางกายภาพเพื่อการทดสอบ

ดำเนินการทดสอบในห้องที่เงียบ ปราศจากสิ่งรบกวนสมาธิของผู้รับการทดสอบ ห้องสอบควรมีอากาศที่ถ่ายเท แสงสว่างเพียงพอ และใหญ่พอสำหรับจัดโต๊ะ 1 ตัวและเก้าอี้ 2 ตัว ผู้รับการทดสอบจะไม่ได้รับอนุญาตให้เห็น แบบสอบ ผู้ให้การทดสอบจะดำเนินการตามขั้นตอนในคู่มือการจัดการสอบ

การสร้างความคุ้นเคยให้ผู้รับการทดสอบ

ผู้ให้การทดสอบควรสร้างบรรยากาศที่เป็นมิตรกับผู้รับการทดสอบก่อนที่จะเริ่มการทดสอบ กล่าวคำทักทายที่ทำให้รู้สึกผ่อนคลาย ชี้แจงถึงวัตถุประสงค์ของการทดสอบอย่างคร่าวๆ แก่ผู้รับการทดสอบ เพื่อให้เข้าใจ หลีกเลี่ยงการทดสอบที่เร่งรีบ

ลำดับการบริหารการทดสอบ

ประมาณความสามารถเบื้องต้นของผู้รับการทดสอบ เพื่อกำหนดแบบสอบที่เหมาะสม เนื่องจากแต่ละแบบสอบย่อย จัดข้อสอบตามระดับความยากเป็น 4 ระดับตามช่วงชั้น ดังนั้น การประมาณความสามารถผู้สอบเบื้องต้น จึงประมาณจากระดับชั้นที่ผู้รับการทดสอบเรียนอยู่ในขณะนั้น กล่าวคือ ให้ผู้รับการทดสอบที่เรียนในระดับชั้นใดรับการทดสอบด้วยแบบสอบที่มีความยากในช่วงชั้นนั้น ดังนี้ ผู้รับการทดสอบเป็นนักเรียนระดับชั้นประถมศึกษาปีที่ 1 ถึง 3 ให้รับการทดสอบด้วยแบบสอบฉบับที่ 1 ผู้รับการทดสอบเป็นนักเรียนระดับชั้นประถมศึกษาปีที่ 4 ถึง 6 ให้รับการทดสอบด้วยแบบสอบฉบับที่ 2 ผู้รับการทดสอบเป็นนักเรียนระดับชั้นมัธยมศึกษาปีที่ 1 ถึง 3 ให้รับการทดสอบด้วยแบบสอบฉบับที่ 3 และ ผู้รับการทดสอบเป็นนักเรียนระดับชั้นมัธยมศึกษาปีที่ 4 ถึง 6 ให้รับการทดสอบด้วยแบบสอบฉบับที่ 4

ในกรณีที่ผู้รับการทดสอบเป็นผู้ที่มีความบกพร่องทางการเรียนรู้หรือเป็นผู้ที่ต้องการความช่วยเหลือพิเศษ ผู้ให้การทดสอบจะหาข้อมูลแวดล้อมเพื่อวินิจฉัย และประมาณความสามารถขั้นเบื้องต้นและกำหนดให้ทำแบบทดสอบฉบับที่เหมาะสมกับระดับความสามารถ เช่น นักเรียนพิเศษคนหนึ่งเรียนอยู่ในชั้นประถมศึกษาปีที่ 5 จากข้อมูลที่ได้จากครูผู้สอน พบว่านักเรียนมีความสามารถในการอ่านต่ำกว่านักเรียนทั่วไปของระดับชั้นประถมศึกษาปีที่ 3 แต่ทักษะการคิดคำนวณ ใกล้เคียงกับนักเรียนชั้นประถมศึกษาปีที่ 4 ดังนั้น การจัดการทดสอบ ควรใช้แบบสอบฉบับที่ 1 ในแบบสอบย่อยการอ่านคำ การสะกดคำ และ ความเข้าใจประโยค สำหรับ แบบสอบย่อยการคำนวณ ควรใช้ฉบับที่ 2 อย่างไรก็ตาม การจัดการสอบสำหรับนักเรียนที่ต้องการความช่วยเหลืออาจมีเงื่อนไขที่ต้องพิจารณาในรายละเอียดเฉพาะตัว ผู้ดำเนินการสอบควรบันทึกเหตุการณ์และรายละเอียดเฉพาะกรณีเพื่อประกอบการแปลผลการสอบและให้ข้อเสนอแนะให้ผู้เกี่ยวข้อง ให้ความช่วยเหลือที่เหมาะสมเป็นกรณี ๆ ไป

กำหนดเวลาในการสอบ

การดำเนินการสอบใช้เวลาทั้งหมดประมาณ 60 นาที เพื่อเวลาสำหรับการดำเนินการ ประมาณ 10 นาที (การแจกแบบสอบและการชี้แจงการสอบ) ส่วนเวลาในการสอบทักษะแต่ละแบบสอบย่อย ประมาณเวลา ดังนี้

ชื่อแบบสอบย่อย	เวลา (นาที/คน) โดยประมาณ
การอ่านคำ	10
การสะกดคำ	10
ความเข้าใจประโยค	15
การคำนวณทางคณิตศาสตร์	25
การดำเนินการ	10

สำหรับผู้รับการทดสอบที่มีความบกพร่องทางการเรียนรู้ หรือสมาธิสั้น มักพบว่าต้องใช้เวลามากขึ้น หรือยุติการสอบได้ก่อนเวลา ทั้งนี้ ให้ผู้ดำเนินการทดสอบบริหารการสอบตามคู่มือ และบันทึกเหตุการณ์เพื่อให้คะแนนมีความถูกต้องนำไปสู่การแปลความหมายที่ให้คำอธิบายตามบริบทการสอบของรายบุคคล

กฎการยุติการสอบ

เนื่องจากการออกแบบการพัฒนาแบบสอบชุดนี้ ได้จัดลำดับความยากในแต่ละแบบสอบย่อยออกเป็น 4 ระดับ เรียงตามระดับความยากจาก ง่ายที่ 1 2 3 และ 4 ตามลำดับ การจัดเรียงความยากกำหนดจากผลการเทียบมาตรฐานคะแนนแบบสอบเป็นช่วงชั้น (Vertical test equating technique) (รายละเอียดกล่าวไว้ในบทที่ 4) นอกจากนี้ ภายในแต่ละฉบับ ได้จัดเรียงข้อสอบตามความยาก (item parameter :b) ที่ผ่านกระบวนการวิเคราะห์ด้วยโปรแกรม BILOG ตามทฤษฎีการตอบสนองข้อสอบ (IRT) จึงกำหนดกฎการยุติการสอบในกรณีที่ผู้รับการทดสอบไม่สามารถตอบข้อสอบได้ **ถูกต้องติดต่อกันตามลำดับ 5 ข้อ** ผู้ดำเนินการสอบจะให้การสอบนั้นยุติลง ณ จุดนี้ถือเป็นระดับสูงสุด (ceiling level) ของผู้รับการทดสอบนั้น ผู้ให้การทดสอบจะพูดกับผู้รับการทดสอบว่า “นักเรียนทำข้อสอบแค่นี้ พอแล้ว ข้อถัดไป เป็นข้อสอบที่ยากเกินไป นักเรียนไม่ต้องทำ”

การจัดการทดสอบเป็นกลุ่ม

แบบสอบย่อยการอ่านคำ ต้องดำเนินการทดสอบเป็นรายบุคคล ผู้ดำเนินการสอบจะต้องบันทึกคะแนนผลการอ่านขณะเวลาที่อ่านออกเสียงเท่านั้น ส่วนอีกสามแบบสอบย่อย ได้แก่ แบบสอบย่อยการสะกดคำ แบบสอบย่อยความเข้าใจประโยค และแบบสอบย่อยการคิดคำนวณ การดำเนินการทดสอบกับผู้รับการทดสอบปกติทำได้ทั้งรายบุคคล

หรือเป็นกลุ่ม กลุ่มไม่เกิน 5 คน เพื่อให้การทดสอบสามารถติดตามการตอบของรายบุคคลและใช้กฎการยุติการสอบดังกล่าวมาแล้ว สำหรับผู้รับการทดสอบที่มีลักษณะบกพร่องการเรียนรู้ หรือผู้ที่ต้องการความช่วยเหลือพิเศษ ควรดำเนินการสอบเป็นรายบุคคลทุกแบบสอบ และใช้กฎการยุติการสอบดังกล่าวมาแล้วเช่นกัน

สรุปการเตรียมการดำเนินการสอบ

1. อ่านคำแนะนำในคู่มือนี้อย่างละเอียดและครบถ้วน
2. ทำความคุ้นเคยกับแบบสอบโดยการทดลองทำการสอบ
3. เตรียมนักเรียนเพื่อรับการสอบ นักเรียนจำนวนมากยังไม่มีประสบการณ์ในการสอบด้วยแบบสอบมาตรฐาน ผู้ให้การสอบควรให้คำแนะนำแก่นักเรียน
4. ตรวจสอบให้แน่ใจว่าสภาพแวดล้อมทางกายภาพของห้องสอบเป็นที่น่าพอใจ มีแสงสว่างเพียงพอ
5. ถ้าจำเป็น แขนงป้าย “กำลังทดสอบ ห้ามรบกวน” ที่ประตูห้อง
6. ถ้าทดสอบนักเรียนเป็นกลุ่ม ควรมีผู้ช่วย
7. ตรวจสอบให้แน่ใจว่าจัดแบบสอบและกระดาษคำตอบที่ถูกต้องให้นักเรียนแต่ละคน

คำชี้แจงในการให้การทดสอบ

แบบสอบการอ่านคำ

แบบสอบการอ่านคำเป็นแบบสอบความจำ ไม่ใช่แบบสอบการพูดหรือการเลือกคำมาใช้พูดหรือเขียนการออกเสียงแบบภาษาพูด การออกเสียงท้องถิ่น หรือข้อบกพร่องในการพูด ยอมรับได้ว่าถูกต้อง คำอ่านออกเสียงที่ถูกต้องจะอยู่ในเอกสารแบบสอบฉบับที่ให้ผู้ให้การสอบถือไว้ ดังนั้นผู้ทดสอบสามารถตรวจสอบการอ่านออกเสียงอย่างระมัดระวังและให้คะแนนระหว่างการทดสอบ ส่วนเอกสารแบบสอบฉบับที่ให้ผู้รับการสอบอ่านออกเสียงจะวางตรงหน้าผู้รับการทดสอบ

แบบสอบการอ่านคำต้องทดสอบเป็นรายบุคคลเท่านั้น แบบสอบการอ่านคำประกอบด้วยคำอ่านพยัญชนะและคำ โดยมีจำนวนของพยัญชนะและคำแตกต่างกันตามชุดแบบสอบฉบับที่ 1 ถึง 4 ดังนี้

แบบสอบ	ส่วนที่ 1 การอ่านพยัญชนะ	ส่วนที่ 2 การอ่านคำ	รวม
ฉบับที่ 1	12 ตัว	24 คำ	36 ข้อ
ฉบับที่ 2	8 ตัว	30 คำ	38 ข้อ
ฉบับที่ 3	--	35 คำ	35 ข้อ
ฉบับที่ 4	--	35 คำ	35 ข้อ

ผู้ให้การสอบเตรียมบัตรรายการคำสำหรับให้ผู้รับการทดสอบอ่าน 1 แผ่น พร้อมกระดาษแข็ง และ แผ่นข้อสอบเดียวกันแต่มีที่ว่างเพื่อระบุชื่อผู้สอบและบันทึกผลการอ่านของผู้สอบนั้นๆ ขณะที่ดำเนินการ ชี้แจงข้อปฏิบัติการทำแบบสอบ ดังนี้
ผู้ให้การสอบวางแผ่นรายการคำที่ให้อ่านตรงหน้านักเรียนพร้อมนำกระดาษแข็งวางปิดให้เห็นคำในแถวแรก

1. ผู้ให้การสอบพูด “ให้นักเรียนดูพยัญชนะ (คำ) ในแถวนี้ทีละตัว (คำ)” ผู้ให้การสอบชี้ไปที่คำในแผ่นรายการคำของนักเรียน ผู้ให้การสอบพูด “ให้นักเรียนอ่านออกเสียงพยัญชนะ (คำ) ทีละตัว (คำ) ตามลำดับจากซ้ายไปขวา อ่านเสียงดังฟังชัด ไม่ต้องรีบร้อน เมื่ออ่านแถวแรกจบให้อ่านแถวที่สอง และอ่านไปเรื่อยๆ จนหมดหน้ากระดาษ หรือเมื่อครูบอกให้หยุด”
2. ผู้ให้การสอบตรวจสอบความพร้อมของนักเรียน แล้วบอกว่า “นักเรียนเริ่มอ่านได้”
3. ผู้ให้การสอบให้นักเรียนใช้เวลาอ่านแต่ละคำประมาณ 5 – 8 วินาที ถ้านักเรียนหยุดนิ่งเกิน 8 วินาที (ผู้ให้การสอบนับในใจ 1 – 8) ผู้ให้การสอบบอก “เอาละ ลองอ่านคำต่อไป”
4. ถ้าผู้ให้การสอบได้ยินไม่ชัด ผู้ให้การสอบบอก “ครูไม่ได้ยินคำที่นักเรียนอ่าน ให้นักเรียนอ่านคำนั้นอีกครั้ง อ่านอย่างที่อ่านเมื่อครู”
5. เมื่อนักเรียนอ่านครบทุกคำในแถวแรก ผู้ให้การสอบเลื่อนกระดาษแข็งลงมาให้เห็นแถวที่ 2 “นักเรียนอ่านพยัญชนะ (คำ) ในแถวนี้ต่อไป” ผู้ให้การสอบชี้นำสายต่านักเรียนตรงพยัญชนะ (คำ) ที่อ่าน
6. ผู้ให้การสอบดำเนินการให้นักเรียนอ่านทีละคำในแต่ละแถว จนหมดคำที่กำหนด เมื่อนักเรียนอ่านจบ หรือสิ้นสุดการทดสอบ “นักเรียนอ่านจบแล้ว นักเรียนทำได้ดีมาก ครูขอบคุณในความร่วมมือ”

แบบสอบการสะกดคำ

แบบสอบการสะกดคำในฉบับที่ 1 และ ฉบับที่ 2 มี 2 ส่วน ส่วนที่ 1 เขียนพยัญชนะตามที่บอก ฉบับที่ 1 ให้เขียน 12 ตัว ฉบับที่ 2 ให้เขียน 8 ตัว ส่วนที่ 2 เขียนคำตามคำบอก ฉบับที่ 1 ให้เขียน 18 คำ ฉบับที่ 2 ให้เขียน 22 คำ สำหรับฉบับที่ 3 และ 4 เป็นการเขียนคำตามคำบอกเพียงอย่างเดียว ฉบับที่ 3 ให้เขียน 30 คำ ฉบับที่ 4 ให้เขียน 30 คำ

เมื่อผู้ให้การสอบแจกแบบสอบทักษะทางวิชาการ ตรงตามชื่อของผู้รับการสอบแล้วชี้แจงข้อปฏิบัติการทำแบบสอบดังนี้

1. ผู้ให้การสอบ “นักเรียนได้รับแจกแบบสอบแล้ว ดูนักเรียนว่าเป็นชื่อตนเองหรือไม่ เมื่อตรงแล้ว เปิดหน้าต่อไป พบให้เห็นหน้าแบบสอบการสะกดคำ” (ผู้ให้การสอบทำหู หรือ ทำให้นักเรียน)
2. ผู้ให้การสอบ “ครูจะอ่านพยัญชนะ (คำ) ให้นักเรียนฟัง แล้วให้นักเรียนเขียนพยัญชนะ (คำ) ตามที่บอกลงในช่องบรรทัด ให้ตรงเลขที่ข้อ ครูจะอ่านพยัญชนะ (คำ) ที่ให้เขียน 2 ครั้ง ขอให้นักเรียนตั้งใจฟังแล้วเขียน”
3. ผู้ให้การสอบตรวจสอบความพร้อมของนักเรียน แล้วเริ่มบอกคำแรก (ฉบับที่ 1 และ 2)

“ส่วนที่ 1 ให้นักเรียนเขียนพยัญชนะ ข้อ 1 ยอ ยัก ฟังอีกครั้ง ยอ ยัก ลงมือเขียนได้” ผู้ให้การสอบดูนักเรียนเขียนให้ตรงเลขที่ข้อตามคำบอก เมื่อนักเรียนเขียนเสร็จผู้ให้การสอบบอกคำต่อไป ในกรณีนี้ที่นักเรียนเขียนช้ามาก เพราะคิดไม่ออก เขียนไม่ถูก เมื่อใช้เวลานานเกิน 10 วินาที (ผู้ให้การสอบนับในใจ 1- 10) ก็ให้นักเรียนเว้นข้อนั้นแล้วบอกพยัญชนะตัวต่อไป จนจบส่วนที่ 1

“ส่วนที่ 2 ให้นักเรียนเขียนคำ ข้อ 1 หู ฟังอีกครั้ง หู ลงมือเขียนได้” (ให้เวลาประมาณ 10 วินาที) “ข้อ 2

4. (ฉบับที่ 3 และ 4)

“ให้นักเรียนเขียนคำ ข้อ 1 รถไฟฟ้า ฟังอีกครั้ง รถไฟฟ้า ลงมือเขียนได้ (ให้เวลาประมาณ 10 วินาที) “ข้อ 2

5. เมื่อนักเรียนทำถึงข้อสุดท้าย หรือยุติการสอบด้วยเงื่อนไขเขียนสะกดคำไม่ถูกต้องกัน 5 ข้อ ผู้ให้การสอบพูด “นักเรียนทำข้อสอบเขียนคำเสร็จแล้ว ครูจะให้นักเรียนทำข้อสอบหน้าต่อไป” ผู้ให้การสอบแนะนำให้นักเรียนเปลี่ยนอิริยาบถ เช่น ลุกนั่งๆ 2-3 รอบ หรือ ปรบมือให้กับงานที่ทำเสร็จ หรือ อื่นๆ ตามความ

เหมาะสมของวัย และธรรมชาติของนักเรียนผู้รับการสอบ แล้วทำการสอบแบบสอบวัดทักษะด้านความเข้าใจประโยค ต่อไป

แบบสอบความเข้าใจประโยค

1. แบบสอบวัดความเข้าใจประโยค ประกอบด้วย ข้อความที่กำหนดให้ผู้สอบเติมคำลงในช่องที่เว้นว่าง เพื่อให้เป็นประโยคสมบูรณ์
2. เมื่อผู้ให้การสอบตรวจความความพร้อมของแบบสอบ ว่าตรงตามชื่อของผู้รับการสอบ และเปิดตรงหน้า แบบสอบความเข้าใจประโยค แล้วชี้แจงข้อปฏิบัติการทำแบบสอบดังนี้ “ให้นักเรียนอ่านข้อสอบข้อ 1 อ่านในใจเอง นักเรียนจะเห็นว่า ข้อความที่อ่านนั้น ยังไม่สมบูรณ์ ให้นักเรียนเติมคำตอบในช่องว่างที่ทำให้ประโยคมีใจความสมบูรณ์ถูกต้อง นักเรียนจะเข้าไปที่ละข้อตามลำดับ ตั้งแต่ ข้อ 1 จนจบ”
3. ผู้ให้การสอบตรวจความพร้อมของนักเรียน แล้วบอก “นักเรียนลงมือทำข้อ 1 ได้แล้ว”
4. ผู้ให้การสอบดูนักเรียน และกระตุ้นให้นักเรียนทำต่อ “นักเรียนทำข้อต่อๆ ไปโดยไม่ต้องรอครูสั่งทีละข้อ”
5. เมื่อนักเรียนทำจนข้อสุดท้าย หรือสั่งให้ยุติตามเงื่อนไขที่กำหนด ผู้ให้การสอบพูด “นักเรียนทำข้อสอบส่วนความเข้าใจประโยค เสร็จแล้ว ครูจะให้นักเรียนทำข้อสอบหน้าต่อไป เกี่ยวกับตัวเลข” ผู้ให้การสอบแนะนำนักเรียนเปลี่ยนอิริยาบถ เช่น ลุกนั่งๆ 2-3 รอบ หรือ ปรบมือให้กับงานที่ทำเสร็จ หรือ อื่นๆ ตามความเหมาะสมของวัย และธรรมชาติของนักเรียนผู้รับการสอบ

แบบสอบการคำนวณทางคณิตศาสตร์

ประกอบด้วย โจทย์คณิตศาสตร์ จำนวน 20 – 30 ข้อ กำหนดให้ผู้สอบคิดหาคำตอบและเติมลงในช่องว่าง ให้ได้คำตอบที่ถูกต้อง

1. เมื่อผู้ให้การสอบตรวจความความพร้อมของแบบสอบการคำนวณทางคณิตศาสตร์ ว่าตรงตามชื่อของผู้รับการสอบ และเปิดตรงหน้า แบบสอบการคำนวณทางคณิตศาสตร์ แล้ว ชี้แจงข้อปฏิบัติการทำแบบสอบดังนี้ “นักเรียนดูแบบสอบการคำนวณทางคณิตศาสตร์ จะเห็นข้อสอบ ... ข้อ” (จำนวนข้อ ดูจากเอกสารที่ให้นักเรียน ฉบับที่ 1 มี 20 ข้อ ฉบับที่ 2 มี 25 ข้อ ฉบับที่ 3 มี 28 ข้อ และฉบับที่ 4 มี 30 ข้อ)

2. “ให้นักเรียนหาคำตอบและเขียนคำตอบในที่ว่างของโจทย์ข้อนั้นๆ อ่าน คิด และเขียนคำตอบทีละข้อตามลำดับตั้งแต่ข้อ 1 จนถึงข้อสุดท้าย” ถ้า นักเรียนต้องการทดเลข ให้เขียนลงในที่ว่างในกระดาษข้อสอบได้ นักเรียน เข้าใจแล้ว ลงมือทำข้อที่ 1 ได้ และทำข้อต่อไป
3. ผู้ให้การสอบดูนักเรียนทำให้เสร็จ และยุติการสอบตามข้อกำหนด
4. ผู้ให้การสอบกล่าวชมนักเรียน “นักเรียนทำแบบสอบครบแล้ว ครูขอ ขอบใจที่ให้ความร่วมมืออย่างดี คุณครูที่ดูแลนักเรียน จะบอกผลสอบ ของนักเรียนภายหลังต่อไป”

การตรวจคะแนน

การบันทึกคะแนน

1. การบันทึกคะแนนในการสอบการอ่านคำ ผู้ให้การสอบเป็นผู้บันทึกคะแนน ขณะทำการสอบ ผู้ให้การสอบต้องตั้งใจฟังและบันทึกคะแนนอย่างถูกต้อง ผู้ให้ การสอบอาจจดบันทึกการออกเสียงของผู้รับการสอบตามที่ได้ยินเพื่อช่วย พิจารณาการให้คะแนน ถ้าผู้รับการสอบเปลี่ยนการออกเสียง ให้ตัดสินผลการ อ่านตามการออกเสียงครั้งสุดท้าย ขั้นตอนการให้คะแนนมีดังนี้
 - การบันทึกคะแนนให้จดบันทึกลงใน “แบบสอบการอ่านคำ แบบบันทึกคะแนน การอ่าน” เท่านั้น
 - เมื่อผู้รับการสอบอ่านพยัญชนะ/คำ ถูกต้อง ให้ 1 คะแนนต่อ 1 พยัญชนะ/คำ
 - สำหรับการออกเสียงอ่านได้ถูกต้อง ให้ทำรอยคะแนน “1” ส่วนการออกเสียงอ่านไม่ถูกต้อง ให้ทำรอยคะแนน “0”
 - ให้ทำรอยคะแนนที่ด้านหลังของพยัญชนะหรือคำนั้นๆ
 - ให้ทำรอยคะแนนในทุกข้ออย่างครบถ้วนที่ดำเนินการสอบ เมื่อสิ้นสุดการสอบ ตามเงื่อนไข จะสิ้นสุดการทำรอยคะแนน
 - ถ้าผู้รับการสอบอ่านผิดติดต่อกัน 5 คำ ให้ยุติการสอบตามเงื่อนไข
 - เมื่อเสร็จสิ้นการสอบ ผู้ให้การสอบตรวจสอบความเรียบร้อยและสรุปคะแนนลง ในช่องคะแนนรวมที่กำหนดให้
2. การบันทึกคะแนนในการสอบการสะกดคำ ความเข้าใจประโยค และการ คำนวณทางคณิตศาสตร์ ผู้ให้การสอบจะเทียบคำตอบที่ถูกต้องจาก “เฉลย” ของแต่ละฉบับ ให้ทำรอยคะแนน “1” เมื่อคำตอบถูก และ ทำรอยคะแนน

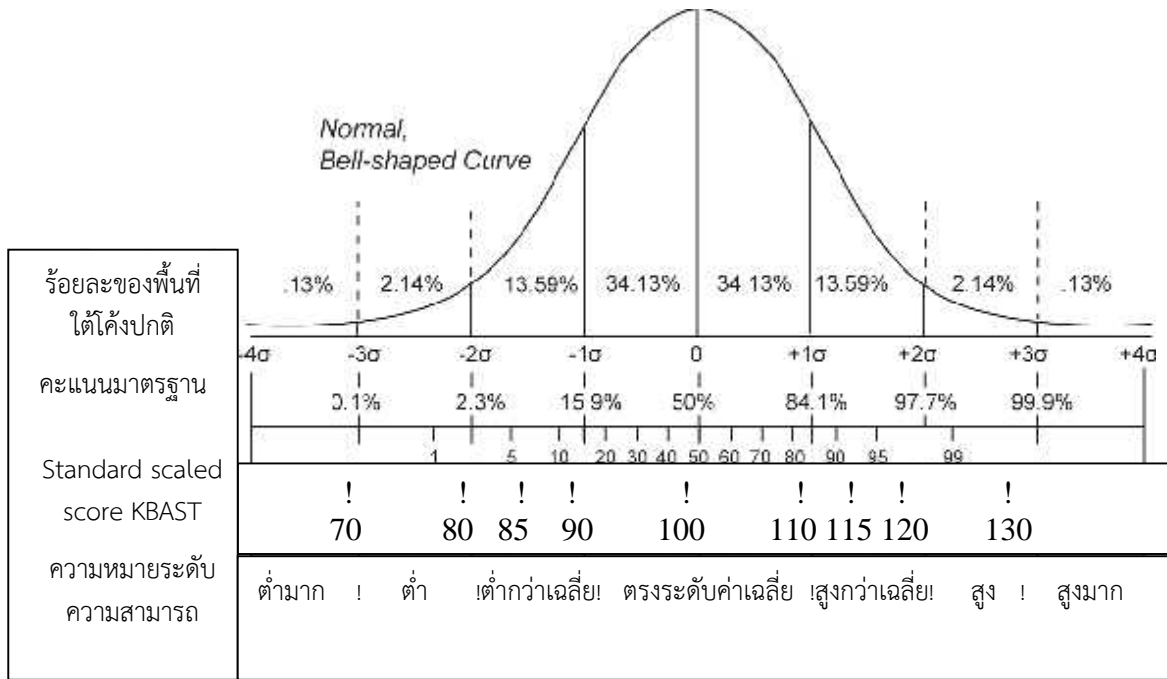
“0” เมื่อคำตอบไม่ถูกต้อง แล้วรวมคะแนนทั้งหมดที่มุมขวาล่างของแบบสอบ มีประเด็นที่ต้องพิจารณาเพิ่มเติมดังนี้

- ในกรณีที่ผู้รับการทดสอบได้รับคำสั่งให้ยุติการสอบ ผู้ให้การสอบจะให้คะแนน “0” หลังจากข้อที่ให้ยุติ ถึงแม้จะพบว่าผู้รับการทดสอบได้เขียนคำตอบที่ถูกต้อง
 - ในกรณีที่ตรวจให้คะแนนแบบสอบความเข้าใจประโยค ผู้ตรวจพบคำตอบที่ผิดจากเฉลย แต่มีความไม่แน่ใจว่าจะเป็นคำตอบที่ไม่ถูกต้อง ให้ขอความเห็นจากฝ่ายบริหารแบบสอบมาตรฐาน KBAST โดยตรง หรือ บันทึกในเอกสารส่งกลับถึง ฝ่ายบริหารแบบสอบมาตรฐาน KBAST “ศูนย์ศึกษาและพัฒนาการจัดการศึกษาพิเศษ “คุณพุ่ม” มหาวิทยาลัยเกษตรศาสตร์”
3. การบันทึกคะแนนรวมทุกแบบสอบย่อยของผู้สอบนั้นลงในช่องคะแนน ที่อยู่ในหน้าแรกของชุดแบบสอบ แล้วนำคะแนนดิบที่ได้ ไปเปิดเทียบคะแนนมาตรฐาน (Standard Scaled Score) จากตารางเทียบคะแนนมาตรฐานในภาคผนวก ข เพื่อเขียนรายงานต่อหน่วยงาน หรือคณะกรรมการที่ร่วมจัดการศึกษาเฉพาะนักเรียนผู้นั้นต่อไป

3. การแปลความหมายผลการสอบ

การพัฒนาแบบสอบมาตรฐานที่อาศัยหลักการแปลผลคะแนนแบบอิงกลุ่มประชากรที่กำหนด ผลการสอบของกลุ่มตัวอย่างประชากร จะนำมาแปลงคะแนนดิบให้เป็นคะแนนมาตรฐาน โดยอ้างอิงในการกระจายระดับความสามารถของประชากรภายใต้พื้นที่ของโค้งปกติ การพัฒนามาตราของคะแนนของชุดแบบสอบ KBAST ก็ได้ใช้หลักการนี้ โดยแปลงให้มีค่าเฉลี่ยเท่ากับ 100 และมีส่วนเบี่ยงเบนมาตรฐานเท่ากับ 15 เพื่อให้สอดคล้องกับการแปลความหมายของแบบสอบมาตรฐานที่เป็นสากล ทั้งนี้ จะเป็นประโยชน์ต่อการพิจารณากรณีบุคคลที่มีการวัดหรือประเมินคุณลักษณะหลายด้าน

หลักการอธิบายระดับความสามารถภายใต้พื้นที่ของโค้งปกติ เมื่อนำมาประยุกต์กับการสอบจาก KBAST จะแปลงคะแนนดิบเป็นคะแนนมาตรฐานตามระเบียบวิธีการแปลง ซึ่งผลการพัฒนาแบบสอบนี้ ได้จัดทำตารางแปลง (ภาคผนวก ข) เพื่อให้สะดวกแก่การใช้ เมื่ออ่านผลการสอบที่เป็นคะแนนดิบ จะเปิดตารางเทียบคะแนนมาตรฐาน แล้วแปลความหมายตามพื้นที่ใต้โค้ง ดังแสดงในแผนภาพที่ 4.17



แผนภาพที่ 4.17 แสดงความหมายของผลการสอบที่แปลงเป็นคะแนนมาตรฐาน 7 ระดับความสามารถ

ตารางที่ 4.26 การแปลความหมายคะแนนตามคะแนนมาตรฐาน

คำอธิบายระดับคุณภาพ	ช่วงคะแนนมาตรฐาน	ร้อยละของผู้สอบภายใต้พื้นที่ใต้โค้งปกติ	ร้อยละโดยประมาณของตัวอย่างผู้สอบ KBAST ที่กระจายภายใต้พื้นที่ใต้โค้ง
สูงมาก	ตั้งแต่ 130 ขึ้นไป	2	2
สูง	120-129	7	7
สูงกว่าค่าเฉลี่ย	110-119	16	15
ค่าเฉลี่ย	90-109	50	50
ต่ำกว่าค่าเฉลี่ย	80-89	16	15
ต่ำ	70-79	7	7
ต่ำมาก	ตั้งแต่ 69 ลงไป	2	2

4. การใช้ผลการสอบ

การวินิจฉัยบุคคลที่มีภาวะสมาธิสั้น บกพร่องทางการเรียนรู้ และออทิซึมเป็นเป้าหมายหลักของการวิจัยพัฒนาแบบสอบมาตรฐาน KBAST เพื่อให้ได้ข้อมูลด้านทักษะพื้นฐานทางวิชาการของบุคคลที่มีความบกพร่องในด้านการเรียนรู้หลากหลายประเภท ข้อมูลจากการสอบนี้ จะเป็นสารสนเทศหนึ่งที่มีความสำคัญต่อการนำไปร่วมพิจารณากับสารสนเทศด้านอื่น ตามความจำเป็นและซับซ้อนในบุคคล เป้าหมายที่ต้องให้ความช่วยเหลือ ดังนั้น เพื่อให้เกิดการใช้ข้อมูลนี้เป็นประโยชน์ต่อนักเรียนอย่างจริงจัง จึงเป็นหน้าที่ของฝ่ายปฏิบัติที่ต้องนำไปใช้

อย่างไรก็ดี การพัฒนาแบบสอบมาตรฐานเป็นงานที่ต้องมีการพัฒนาปรับปรุงอยู่เสมอ และจะทำได้ดีเมื่อมีข้อมูลป้อนกลับจากผู้ใช้ ที่ให้ข้อสังเกต และข้อมูลเชิงประจักษ์เป็นรายกรณีถึง “ศูนย์ศึกษาและพัฒนาการจัดการศึกษาพิเศษ “คุณพุ่ม” มหาวิทยาลัยเกษตรศาสตร์”

การพัฒนาต่อจากนี้จะเป็นการเรียนรู้ร่วมกันระหว่างผู้สร้าง คือ “ศูนย์ศึกษาและพัฒนาการจัดการศึกษาพิเศษ “คุณพุ่ม” มหาวิทยาลัยเกษตรศาสตร์” และ ผู้ใช้ คือ บุคลากร นักวิชาการ และหน่วยงานที่ทำหน้าที่ให้ความช่วยเหลือบุคคลที่ต้องการความช่วยเหลือทางการศึกษาหลากหลายรูปแบบ

สรุปผลการวิจัยและอภิปรายผล

จากการดำเนินการพัฒนาแบบสอบทักษะพื้นฐานทางวิชาการ KBAST ด้วยเทคนิคของการพัฒนาแบบสอบมาตรฐาน ได้ผลผลิตชุดแบบสอบตามทีออกแบบไว้ เป็นชุดแบบสอบ (Test battery) วัดทักษะทางวิชาการ 4 ทักษะ คือ แบบสอบการอ่านคำ (Word Reading) แบบสอบการสะกดคำ (Word Spelling) แบบสอบความเข้าใจประโยค (Sentence Comprehension) และ แบบสอบการคำนวณทางคณิตศาสตร์ (Math Computation) โดยจัดชุดแบบสอบเป็น 4 ฉบับ ด้วยรูปแบบการวัดพัฒนาทักษะตามช่วงชั้น (Grade-to-Grade Growth) กล่าวคือ แต่ละฉบับประกอบด้วยแบบสอบย่อย 4 ทักษะ ซึ่งมีระดับความยากของข้อสอบที่จัดเรียงตามลำดับจากง่ายไปยากที่สอดคล้องกับช่วงชั้นตั้งแต่ ช่วงชั้นที่ 1 (ระดับชั้นประถมศึกษาปีที่ 1 ถึง 3) ช่วงชั้นที่สอง (ระดับชั้นประถมศึกษาปีที่ 4 ถึง 6) ช่วงชั้นที่สาม (ระดับชั้นมัธยมศึกษาปีที่ 1 ถึง 3) และช่วงชั้นที่สี่ (ระดับชั้นมัธยมศึกษาปีที่ 4 ถึง 6)

บทนี้ จะนำเสนอ 2 ตอน ตอนที่หนึ่ง สรุปผลการวิจัย 2 ส่วน ส่วนที่หนึ่ง ผลผลิตและการนำไปใช้ ส่วนที่สอง สรุปคุณภาพเชิงเทคนิค ตอนที่สอง อภิปรายผล มี 3 ส่วน ส่วนที่หนึ่งข้อจำกัดของการวิจัย ส่วนที่สองข้อเสนอเพื่อการพัฒนาต่อเนื่องให้ได้คุณภาพในเชิงลึก ส่วนที่สาม การสร้างเครือข่ายการพัฒนาแบบสอบมาตรฐานในเชิงกว้าง

ตอนที่ 1 สรุปผลการวิจัย

ผลผลิต

การพัฒนาชุดแบบสอบมาตรฐานเพื่อประเมินทักษะขั้นพื้นฐานทางวิชาการ KBAST นี้ ได้ผลผลิตชุดแบบสอบ สามารถให้ครูและผู้เกี่ยวข้องที่ทำหน้าที่จัดการศึกษาให้แก่แก่นักเรียนที่มีความบกพร่องทางการเรียนรู้ สมารถสืบ และมีภาวะอหิซึม นำไปใช้ประเมินทักษะพื้นฐานทางวิชาการ แล้วนำผลมาประกอบเป็นสารสนเทศร่วมกับข้อมูลจากการประเมินทางสติปัญญา และด้านประวัติทางสุขภาพ แล้วตัดสินใจกำหนดแผนการช่วยเหลือนักเรียนเป็นรายบุคคล ชุดแบบสอบ KBAST มีวัสดุประกอบดังนี้



ภาพตัวอย่าง
แบบสอบ

1. ชุดแบบสอบออกแบบจัดเป็นฉบับๆ ตามระดับช่วงชั้นของผู้สอบ ผู้สอบที่เรียนอยู่ในช่วงชั้นที่ 1 ครูจะให้ทำแบบสอบฉบับที่ 1 ผู้สอบที่เรียนอยู่ในช่วงชั้นที่ 2 ครูจะให้ทำแบบ

สอบฉบับที่ 2 ผู้สอบที่เรียนอยู่ในช่วงชั้นที่ 3 และ 4 ครูก็จะจัดให้นักเรียนสอบด้วยแบบสอบฉบับที่ 3 และ 4 ตามลำดับ แต่ละฉบับจะมีแบบสอบย่อย 4 ทักษะ คือ ทักษะการอ่านคำ ทักษะการสะกดคำ ความเข้าใจประโยค และ ทักษะการคำนวณทางคณิตศาสตร์

2. แบบเฉลยคู่กับชุดแบบสอบ ผู้บริหารการสอบจะใช้เมื่อตรวจให้คะแนน
3. ตารางเกณฑ์คะแนนมาตรฐานปกติ การแปลความหมายจากคะแนนดิบเป็นคะแนนมาตรฐาน เพื่อให้ผลการวัดใช้กลุ่มมาตรฐานเป็นเกณฑ์ในการแปลความหมายตามกลุ่มอ้างอิง (รวมไว้ในคู่มือ)
4. คู่มือบริหารการสอบชุดแบบสอบ KBAST กำหนดการดำเนินการให้เป็นมาตรฐานตามกำหนดของแบบสอบ

การนำไปใช้

1. การนำชุดแบบสอบ KBAST ไปใช้ ผู้ใช้ในฐานะผู้ให้การทดสอบ ต้องศึกษาคู่มือการใช้ อย่างเข้าใจ และฝึกซ้อมการบริหารการสอบอย่างมั่นใจก่อนใช้ โดยเฉพาะการใช้กับนักเรียนที่มีความบกพร่องแบบต่างๆ ซึ่งผู้ใช้ต้องศึกษาทำความเข้าใจนักเรียนก่อนดำเนินการสอบ ผู้ใช้แบบสอบจะเป็นผู้ตรวจให้คะแนนการตอบของนักเรียน ต้องศึกษาวิธีการ และปฏิบัติด้วยความรอบคอบ โดยเฉพาะแบบสอบย่อยความเข้าใจประโยค เพราะบางข้อมีค่าเฉลยมากกว่า 1 คำตอบ
2. ผลการตรวจให้คะแนน เมื่อนำมาเทียบกับตารางแปลงคะแนนดิบที่นักเรียนทำได้ จะเป็นคะแนนมาตรฐาน แล้วแปลความหมายคะแนนตามคะแนนมาตรฐานเทียบกับระดับชั้น หรือ ช่วงชั้นที่ใช้แบบสอบฉบับนั้นๆ สรุปเป็นผลการประเมินทักษะนั้นๆ
3. ในกรณีการให้นักเรียนที่มีความบกพร่องในการเรียนรู้ สมาธิสั้น หรือผู้ที่มีภาวะออทิซึม ผู้ทำหน้าที่ให้การทดสอบต้องให้การทดสอบเป็นรายบุคคลเท่านั้น และสังเกตความพร้อมในการทำข้อสอบ ถ้าจำเป็นต้องพักเพื่อให้มีความพร้อม ก็ควรทำ เพราะการสอบเมื่อผู้สอบพร้อม จึงจะให้ข้อมูลความสามารถของเขาตามจริงได้ และจะทำให้ผลการประเมินมีคุณค่า แต่อย่างไรก็ดี ผู้ให้การทดสอบควรบันทึกเหตุการณ์ เพื่อนำข้อมูลเหล่านี้ไปพิจารณาประกอบการปรับปรุงแบบสอบ และการสอบครั้งต่อไป

4. ในกรณีที่ผู้สอบทำผิดติดต่อกัน 5 ข้อ ถือว่าการสอบได้สิ้นสุดลง ผู้บริหารการสอบสังเกตปฏิกิริยาของผู้สอบ และบอกการยุติเมื่อผู้สอบมีสมาธิรับรู้การยุติ
5. การนำผลการสอบไปใช้ เป็นเป้าหมายของสร้างและพัฒนาแบบสอบ จากการชี้แจงกับครูผู้รับผิดชอบการให้ความช่วยเหลือแก่นักเรียนว่าผลการประเมินเป็นสารสนเทศเบื้องต้นของการเข้าใจระดับความพร้อมทางวิชาการในทักษะการอ่าน การเขียนสะกดคำ ความเข้าใจประโยค และ ทักษะการคำนวณทางคณิตศาสตร์ ครูจะใช้ผลการประเมินร่วมกับผลการประเมินด้านจิตวิทยา ด้านทักษะทางสังคม และ อื่นๆ ที่จะทำให้ครูเข้าใจผู้เรียนมากที่สุด ทำให้ครูวางแผนพัฒนาเด็กได้อย่างเหมาะสมกับเด็กรายบุคคลนั้นๆ

ตอนที่ 2 อภิปราย

ข้อจำกัดของการวิจัย

การพัฒนาแบบสอบมาตรฐานอย่างเคร่งครัด เพื่อให้ได้คุณภาพ และเกิดคุณค่าต่อการนำไปใช้เป็นงานที่ต้องการปัจจัยทั้งด้านความรู้ โครงสร้างพื้นฐานในการผลิต อันได้แก่ คอมพิวเตอร์ โปรแกรม หรือระบบการวิเคราะห์คุณภาพ จำเป็นต้องตรวจสอบกระบวนการผลิตทุกขั้นตอน เวลา และงบประมาณที่เพียงพอ การทำวิจัยครั้งนี้ พบว่าปัจจัยที่กล่าวมา มีความจำกัด และที่ขาดมาก จนเป็นสิ่งที่ไม่เพียงพอ คือ เวลา และงบประมาณ แต่อย่างไรก็ดี งานนี้ได้มองเห็นความสำเร็จในระดับของการเริ่มต้น เสมือนหนึ่งเป็นโครงการนำร่องของการพัฒนาชุดแบบสอบทักษะพื้นฐานทางวิชาการ ที่มีความจำเป็นในการเป็นเครื่องมือสำคัญ จากการศึกษาการพัฒนาของชุดแบบสอบ Wide Range Achievement Test 4 (WRAT 4) พบว่า เป็นผลงานที่มีการพัฒนามากกว่า 70 ปี บัดนี้ คุณภาพ และคุณค่าของแบบสอบชุดนี้ มีทั้งด้านวิชาการที่ก่อให้เกิดความรู้หลายสาขา ด้านการใช้เพื่อการให้ความช่วยเหลือนักเรียนเป็นรายบุคคล และใช้ในการวางนโยบายเพื่อการจัดการศึกษารูปแบบต่างๆ แก่กลุ่มเป้าหมายที่ด้อยโอกาสทางการศึกษา และด้อยโอกาสการเรียนรู้ เช่นกลุ่มที่มีความบกพร่องทางการเรียนรู้ สมาธิสั้น และบุคคลที่มีภาวะออทิซึมดังที่โครงการวิจัยนี้กำหนดเป็นกลุ่มบุคคลเป้าหมาย

ข้อเสนอเพื่อการพัฒนาต่อเนื่องให้ได้คุณภาพในเชิงลึก

จากผลการพัฒนาชุดแบบสอบ KBAST พบข้อจำกัดหลายประการ ซึ่งจำเป็นต้องดำเนินการอย่างต่อเนื่องตามกระบวนการพัฒนาแบบสอบมาตรฐาน ดังนี้

1. พัฒนาเนื้อหาของแบบสอบ ให้มีเนื้อหาครอบคลุมสาระที่กว้างขวางมากขึ้น เพื่อนำไปพัฒนาข้อสอบที่มีคุณภาพ เก็บเป็นคลังข้อสอบ เป็นทางเลือกที่จะสร้างข้อสอบคู่ขนานที่เหมาะสมกับบริบททางสังคม วัฒนธรรมของกลุ่มสังคมชุมชน เช่น กลุ่มนักเรียนที่อยู่

ในภูมิภาคที่แตกต่างกัน อยู่ในสังคมที่มีประเพณีที่แตกต่าง เป็นต้น ทั้งนี้ เพราะเป้าหมายของการวัดคือ ความสามารถในทักษะพื้นฐาน ต้องสัมพันธ์กับการเรียนรู้เพื่อการสื่อสารในชีวิตประจำวัน

2. ออกแบบวิจัยเพื่อพัฒนาคุณภาพด้านความตรง โดยเฉพาะ ความตรงภายใน (Internal Evidence of Validity) เป็นขั้นตอนแรกของการพัฒนา ที่มีกระบวนการที่มีความละเอียด ต้องรอบคอบ เป็นงานที่ต้องตรวจสอบทบทวนความสอดคล้องระหว่างเนื้อหาสาระที่ต้องการวัดกับวัตถุประสงค์ของโครงการสอบ และต้องอาศัยข้อมูลเชิงประจักษ์ที่จะแสดงคุณลักษณะของข้อสอบที่สัมพันธ์กับความสามารถของผู้ตอบตามทฤษฎีการวัด ความสามารถแฝงในกลุ่มประชากรเป้าหมายที่กำหนด เป็นงานที่ต้องการเวลา และมีปัจจัยสนับสนุนการขับเคลื่อนงานอย่างเป็นระบบ
3. การพัฒนาความตรงภายนอก (External Evidence of Validity) เป็นหลักฐานที่มาจากการหาความสัมพันธ์ระหว่างคะแนนจากแบบสอบ KBAST กับ คะแนนจากเครื่องมือที่วัดในคุณลักษณะ (traits) เดียวกัน หรือ โกล้เคียงกัน เช่น แบบวัดทักษะต่างๆ ที่กรรมการการศึกษาาร่วมกันสร้าง ซึ่งจะทำให้การสร้าง/พัฒนาเครื่องมือที่วัดในระดับท้องถิ่นได้มากขึ้น
4. ทำต้นแบบการศึกษาหาความสัมพันธ์ระหว่างผลการสอบย่อยแต่ละทักษะ ของนักเรียนกลุ่มที่สนใจ โดยเฉพาะ กลุ่มที่มีความบกพร่องทางการเรียนรู้ สมาธิสั้น หรือ ที่มีภาวะออทิซึม กับผลการประเมินด้านอื่นๆ กับแบบสอบอื่นๆ เป็นวิธีที่จะให้ความเข้าใจในการนำผลการสอบไปอธิบายพฤติกรรมมิติต่างๆ และวางแผนให้ความช่วยเหลือด้วย

การสร้างเครือข่ายการพัฒนาแบบสอบมาตรฐานในเชิงกว้าง

ตัวอย่างกิจกรรม เช่น

1. สนับสนุนให้โรงเรียน หรือศูนย์ฯ ที่มีหน้าที่จัดการเรียนการสอนให้เด็กที่ต้องการความช่วยเหลือพิเศษ นำชุดแบบสอบ KBAST ไปใช้ และร่วมกันศึกษาเพื่อหาความรู้หลายแง่มุมจากการใช้
2. หน่วยงานที่เกี่ยวข้องดังกล่าวข้างต้น ควรได้รับงบประมาณสนับสนุนให้บุคลากรเข้าอบรมและปฏิบัติการเป็นเครือข่ายในโครงการพัฒนาแบบสอบมาตรฐาน KBAST

บรรณานุกรม

- กระทรวงศึกษาธิการ. (2551). *พระราชบัญญัติการจัดการศึกษาสำหรับคนพิการ พ.ศ. 2551*. กรุงเทพมหานคร : โรงพิมพ์ชุมนุมสหกรณ์การเกษตรแห่งประเทศไทย จำกัด.
- กระทรวงศึกษาธิการ. (2552). *ประกาศกระทรวงศึกษาธิการ เรื่อง หลักเกณฑ์และวิธีการจัดทำแผนการจัดการศึกษาเฉพาะบุคคล ระดับการศึกษาขั้นพื้นฐาน พ.ศ. 2552*. กรุงเทพมหานคร : ศุภสภา.
- กระทรวงศึกษาธิการ. (2531). *บัญชีคำพื้นฐานที่ใช้ในการเรียนการสอนภาษาไทย*. กรุงเทพมหานคร : ศุภสภา.
- สำนักวิชาการและมาตรฐานการศึกษา สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน
กระทรวงศึกษาธิการ. (2551). *ตัวชี้วัดและสาระการเรียนรู้แกนกลางกลุ่มสาระการเรียนรู้ภาษาไทยตามหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช ๒๕๕๑*. กรุงเทพมหานคร : ศุภสภา
- สำนักวิชาการและมาตรฐานการศึกษา สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน
กระทรวงศึกษาธิการ. (2551). *ตัวชี้วัดและสาระการเรียนรู้แกนกลาง กลุ่มสาระการเรียนรู้คณิตศาสตร์ ตามหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช ๒๕๕๑*. กรุงเทพมหานคร : ศุภสภา
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for Educational and Psychological Testing*. Washington, DC: AERA.
- Association of American Publishers. (2001). *Education assessment: A primer*. Washington, DC: Author.
- Hegde, M. N. (2008). *Assessment of communication disorders in children: Resources and protocols*. San Diego, CA: Plural Publishing, Incorporated.
- Jastak, J. (1946). *Wide Range Achievement Test*. Wilmington, DE : Guidance Associates of Delaware, Inc.
- Jastak, J., & Jastak, S. (1965, 1976 & 1978). *Wide Range Achievement Test*. Wilmington, DE: Jastak Associates, Inc.

- Kolen, M. J., & Brennan, R. L. (2004). *Test equating, scaling, and linking* (2nd ed.). New York: Springer.
- Michele F., Zimowski, Eiji Muraki, Robert J. Mislevy, R. Darrell Bock. (2014). *BILOG-MG for Windows*. Scientific Software International, Inc.
- Test Service Notebook, 27. U.S. Department of Education. (2003). ***No Child Left Behind***, accountability and AYP. Presentation at the Student Achievement and School Accountability Conference, October 2002. Retrieved November 11, 2003, from <http://www.ed.gov/admins/lead/account/ayp/edlite-index.html>
- Wall, J. E., & Walz, G. R. (Eds.). (2004). *Measuring up: Assessment issues for teachers, counselors, and administrators*. Greensboro, NC: CAPS Press.
- Wright, Benjamin & Stone, Mark. (1999). *Measurement essentials* (2nd ed.). Wilmington, Delaware : Wide Range, INC.

ภาคผนวก ก.

รายนามผู้เชี่ยวชาญกำหนดหลักสูตรการสอบ และสร้างข้อสอบ
แบบสอบทักษะพื้นฐานทางวิชาการ (Kasetsart Basic Academic Skills Test)
เพื่อการประเมินและวินิจฉัยบุคคลที่มีภาวะสมาธิสั้น บกพร่องทางการเรียนรู้ และออทิซึม

- | | |
|----------------------------|--|
| 1. อ.ดร.ธัชทฤต เทียมธรรม | หัวหน้าทีมผู้เชี่ยวชาญวิชาคณิตศาสตร์ |
| 2. ผศ.นฤกร รุจิเรข | หัวหน้าทีมผู้เชี่ยวชาญวิชาภาษาไทยประถม |
| 3. ผศ.สุมาลี เกษรวิชวัฒนา | หัวหน้าทีมผู้เชี่ยวชาญวิชาภาษาไทยมัธยม |
| 4. อ.สมนึก พงศ์กุหลาบ | ผู้เชี่ยวชาญวิชาคณิตศาสตร์ |
| 5. อ.ชญานิน ภูเจริญ | ผู้เชี่ยวชาญวิชาคณิตศาสตร์ |
| 6. อ.พรรณทิภา คำพรหม | ผู้เชี่ยวชาญวิชาคณิตศาสตร์ |
| 7. อ.ณรงค์ฤทธิ์ ฉายา | ผู้เชี่ยวชาญวิชาคณิตศาสตร์ |
| 8. อ.จิตติมา พิศาภาค | ผู้เชี่ยวชาญวิชาคณิตศาสตร์ |
| 9. อ.นิชยา กมลพัฒนะ | ผู้เชี่ยวชาญวิชาคณิตศาสตร์ |
| 10. ผศ.สมใจ บุญอุรพิภิญโญ | ผู้เชี่ยวชาญวิชาภาษาไทย |
| 11. ผศ.พรรัตน์ สุวรรณรัตน์ | ผู้เชี่ยวชาญวิชาภาษาไทย |
| 12. อ.สุวิกรม มาประณีต | ผู้เชี่ยวชาญวิชาภาษาไทย |
| 13. อ.มลฤดี ฐิติชัยมงคล | ผู้เชี่ยวชาญวิชาภาษาไทย |

ภาคผนวก ข.

ตารางเทียบคะแนนดิบสู่คะแนนมาตรฐานรายชั้น ประถมศึกษาปีที่ 1-มัธยมศึกษาปีที่ 6 (Raw score Subtest Standard Score Conversion by Grade)

ตารางเกณฑ์ปกติที่ 1	แบบสอบฉบับที่ 1 (Form A)	ชั้นประถมศึกษาปีที่ 1
ตารางเกณฑ์ปกติที่ 2	แบบสอบฉบับที่ 1 (Form A)	ชั้นประถมศึกษาปีที่ 2
ตารางเกณฑ์ปกติที่ 3	แบบสอบฉบับที่ 1 (Form A)	ชั้นประถมศึกษาปีที่ 3
ตารางเกณฑ์ปกติที่ 4	แบบสอบฉบับที่ 2 (Form B)	ชั้นประถมศึกษาปีที่ 4
ตารางเกณฑ์ปกติที่ 5	แบบสอบฉบับที่ 2 (Form B)	ชั้นประถมศึกษาปีที่ 5
ตารางเกณฑ์ปกติที่ 6	แบบสอบฉบับที่ 2 (Form B)	ชั้นประถมศึกษาปีที่ 6
ตารางเกณฑ์ปกติที่ 7	แบบสอบฉบับที่ 3 (Form C)	ชั้นมัธยมศึกษาปีที่ 1
ตารางเกณฑ์ปกติที่ 8	แบบสอบฉบับที่ 3 (Form C)	ชั้นมัธยมศึกษาปีที่ 2
ตารางเกณฑ์ปกติที่ 9	แบบสอบฉบับที่ 3 (Form C)	ชั้นมัธยมศึกษาปีที่ 3
ตารางเกณฑ์ปกติที่ 10	แบบสอบฉบับที่ 4 (Form D)	ชั้นมัธยมศึกษาปีที่ 4
ตารางเกณฑ์ปกติที่ 11	แบบสอบฉบับที่ 4 (Form D)	ชั้นมัธยมศึกษาปีที่ 5
ตารางเกณฑ์ปกติที่ 12	แบบสอบฉบับที่ 4 (Form D)	ชั้นมัธยมศึกษาปีที่ 6

ตารางเทียบคะแนนดิบสู่คะแนนมาตรฐานรายฉบับ (ฉบับที่ 1- 4) (Raw score Standard Score Conversion by Test Form A-D)

ตารางเกณฑ์ปกติที่ 13	แบบสอบฉบับที่ 1 (Form A)	ช่วงชั้นประถมศึกษาปีที่ 1 –3
ตารางเกณฑ์ปกติที่ 14	แบบสอบฉบับที่ 2 (Form B)	ช่วงชั้นประถมศึกษาปีที่ 4 –6
ตารางเกณฑ์ปกติที่ 15	แบบสอบฉบับที่ 3 (Form C)	ช่วงชั้นมัธยมศึกษาปีที่ 1 –3
ตารางเกณฑ์ปกติที่ 16	แบบสอบฉบับที่ 4 (Form D)	ช่วงชั้นมัธยมศึกษาปีที่ 4 –6

Form A Standard Score Conversion

ตารางเกณฑ์ปกติที่ 1 เทียบคะแนนดิบสู่คะแนนมาตรฐาน แบบสอบฉบับที่ 1 ชั้นประถมศึกษาปีที่ 1

Word Reading		Spelling		Sentence Comprehension		Math Computation	
Raw score	Scaled score	Raw score	Scaled score	Raw score	Scaled score	Raw score	Scaled score
0	55	0	55	0	55	0	64
1	55	1	55	1	75	1	76
2	55	2	55	2	79	2	84
3	55	3	55	3	83	3	91
4	70	4	55	4	87	4	98
5	70	5	55	5	91	5	103
6	71	6	70	6	95	6	108
7	71	7	71	7	98	7	113
8	71	8	72	8	101	8	117
9	71	9	73	9	104	9	121
10	72	10	73	10	107	10	125
11	72	11	75	11	110	11	128
12	73	12	75	12	113	12	128
13	73	13	78	13	116	13	135
14	74	14	80	14	119	14	138
15	75	15	80	15	122	15	141
16	75	16	84	16	124	16	145
17	75	17	86			17	145
18	78	18	88			18	145
19	79	19	90			19	145
20	81	20	93			20	145
21	82	21	95				
22	84	22	98				
23	85	23	100				
24	85	24	103				
25	89	25	106				
26	91	26	109				
27	94	27	113				
28	96	28	116				
29	98	29	119				
30	101	30	123				
31	104						
32	107						
33	110						
34	113						
35	117						
36	120						

Form A Subtest Standard Score Conversion

ตารางเกณฑ์ปกติที่ 2 เทียบคะแนนดิบสู่คะแนนมาตรฐาน แบบสอบฉบับที่ 1 ชั้นประถมศึกษาปีที่ 2

Word Reading		Spelling		Sentence Comprehension		Math Computation	
Raw score	Scaled score	Raw score	Scaled score	Raw score	Scaled score	Raw score	Scaled score
0	55	0	55	0	55	0	66
1	55	1	55	1	70	1	66
2	55	2	55	2	71	2	66
3	55	3	55	3	72	3	81
4	55	4	69	4	74	4	84
5	55	5	69	5	76	5	88
6	66	6	69	6	79	6	91
7	66	7	69	7	82	7	95
8	66	8	69	8	85	8	98
9	66	9	69	9	88	9	101
10	66	10	69	10	92	10	104
11	67	11	69	11	97	11	106
12	67	12	70	12	101	12	109
13	67	13	70	13	106	13	112
14	68	14	71	14	111	14	114
15	68	15	71	15	117	15	117
16	68	16	72	16	123	16	119
17	68	17	73			17	121
18	70	18	73			18	124
19	70	19	73			19	126
20	70	20	77			20	128
21	71	21	79				
22	72	22	81				
23	74	23	84				
24	76	24	87				
25	77	25	91				
26	80	26	95				
27	82	27	99				
28	85	28	105				
29	88	29	111				
30	92	30	117				
31	96						
32	101						
33	107						
34	113						
35	119						
36	145						

Form A Subtest Standard Score Conversion

ตารางเกณฑ์ปกติที่ 3 เทียบคะแนนดิบสู่คะแนนมาตรฐาน แบบสอบฉบับที่ 1 ชั้นประถมศึกษาปีที่ 3

Word Reading		Spelling		Sentence Comprehension		Math Computation	
Raw score	Scaled score	Raw score	Scaled score	Raw score	Scaled score	Raw score	Scaled score
0	55	0	55	0	55	0	55
1	55	1	55	1	55	1	55
2	55	2	55	2	69	2	76
3	55	3	55	3	69	3	81
4	55	4	55	4	69	4	81
5	62	5	55	5	70	5	88
6	62	6	55	6	71	6	88
7	62	7	55	7	73	7	95
8	62	8	55	8	75	8	98
9	62	9	55	9	78	9	101
10	62	10	55	10	81	10	104
11	62	11	69	11	85	11	106
12	62	12	69	12	90	12	109
13	62	13	69	13	96	13	112
14	62	14	69	14	103	14	114
15	63	15	69	15	111	15	117
16	63	16	69	16	121	16	119
17	63	17	69			17	121
18	63	18	69			18	124
19	63	19	70			19	126
20	63	20	70			20	128
21	63	21	71				
22	66	22	73				
23	67	23	74				
24	67	24	77				
25	67	25	80				
26	67	26	84				
27	67	27	89				
28	76	28	95				
29	79	29	104				
30	82	30	115				
31	86						
32	91						
33	96						
34	103						
35	110						
36	118						

Form B Subtest Standard Score Conversion

ตารางเกณฑ์ปกติที่ 4 เทียบคะแนนดิบสู่คะแนนมาตรฐาน แบบสอบฉบับที่ 2 ชั้นประถมศึกษาปีที่ 4

Word Reading		Spelling		Sentence Comprehension		Math Computation	
Raw score	Scaled score	Raw score	Scaled score	Raw score	Scaled score	Raw score	Scaled score
0	55	0	55	0	55	0	67
1	55	1	55	1	65	1	72
2	55	2	55	2	69	2	77
3	67	3	55	3	72	3	81
4	67	4	55	4	76	4	85
5	67	5	71	5	79	5	88
6	67	6	71	6	83	6	91
7	67	7	71	7	86	7	95
8	67	8	73	8	90	8	98
9	67	9	74	9	94	9	101
10	67	10	75	10	97	10	104
11	67	11	75	11	101	11	107
12	67	12	78	12	104	12	109
13	67	13	79	13	108	13	112
14	68	14	81	14	112	14	115
15	68	15	83	15	115	15	117
16	68	16	84	16	119	16	120
17	68	17	87	17	123	17	122
18	68	18	89	18	126	18	125
19	68	19	91	19	130	19	127
20	68	20	94	20	145	20	129
21	68	21	96			21	129
22	68	22	99			22	129
23	74	23	102			23	129
24	76	24	105			24	129
25	77	25	109			25	141
26	79	26	112				
27	81	27	116				
28	83	28	120				
29	85	29	124				
30	87	30	128				
31	90						
32	93						
33	96						
34	100						
35	104						
36	108						
37	112						
38	117						

Form B Subtest Standard Score Conversion

ตารางเกณฑ์ปกติที่ 5 เทียบคะแนนดิบสู่คะแนนมาตรฐาน แบบสอบฉบับที่ 2 ชั้นประถมศึกษาปีที่ 5

Word Reading		Spelling		Sentence Comprehension		Math Computation	
Raw score	Scaled score	Raw score	Scaled score	Raw score	Scaled score	Raw score	Scaled score
0	55	0	55	0	55	0	64
1	55	1	55	1	70	1	70
2	55	2	55	2	73	2	74
3	55	3	55	3	76	3	78
4	55	4	55	4	79	4	82
5	55	5	55	5	82	5	85
6	55	6	55	6	86	6	88
7	55	7	72	7	89	7	91
8	55	8	73	8	92	8	94
9	55	9	74	9	95	9	97
10	55	10	75	10	98	10	99
11	67	11	76	11	101	11	102
12	67	12	77	12	104	12	104
13	67	13	77	13	107	13	107
14	67	14	80	14	110	14	109
15	67	15	82	15	113	15	111
16	68	16	84	16	116	16	114
17	68	17	86	17	119	17	116
18	68	18	88	18	122	18	118
19	68	19	90	19	125	19	120
20	68	20	92	20	128	20	122
21	68	21	95			21	124
22	68	22	97			22	126
23	68	23	100			23	126
24	72	24	103			24	130
25	72	25	106			25	132
26	74	26	109				
27	76	27	112				
28	78	28	116				
29	80	29	119				
30	82	30	123				
31	85						
32	88						
33	92						
34	96						
35	101						
36	106						
37	113						
38	119						

Form B Subtest Standard Score Conversion

ตารางเกณฑ์ปกติที่ 6 เทียบคะแนนดิบสู่คะแนนมาตรฐาน แบบสอบฉบับที่ 2 ชั้นประถมศึกษาปีที่ 6

Word Reading		Spelling		Sentence Comprehension		Math Computation	
Raw score	Scaled score	Raw score	Scaled score	Raw score	Scaled score	Raw score	Scaled score
0	55	0	55	0	55	0	55
1	55	1	55	1	55	1	55
2	55	2	55	2	55	2	66
3	55	3	55	3	66	3	68
4	58	4	65	4	66	4	68
5	58	5	65	5	69	5	72
6	58	6	65	6	72	6	74
7	58	7	65	7	74	7	77
8	58	8	66	8	77	8	79
9	58	9	66	9	80	9	81
10	58	10	66	10	83	10	83
11	58	11	66	11	86	11	86
12	58	12	66	12	90	12	88
13	58	13	66	13	94	13	91
14	58	14	66	14	98	14	93
15	58	15	66	15	103	15	96
16	58	16	66	16	107	16	99
17	58	17	72	17	112	17	101
18	58	18	74	18	118	18	104
19	58	19	76	19	123	19	107
20	59	20	78	20	129	20	109
21	59	21	81			21	112
22	59	22	83			22	115
23	59	23	87			23	118
24	59	24	90			24	121
25	61	25	95			25	124
26	61	26	99				
27	61	27	105				
28	61	28	110				
29	65	29	117				
30	65	30	124				
31	65						
32	73						
33	77						
34	82						
35	89						
36	96						
37	105						
38	116						

Form C Subtest Standard Score Conversion

ตารางเกณฑ์ปกติที่ 7 เทียบคะแนนดิบสู่คะแนนมาตรฐาน แบบสอบฉบับที่ 3 ชั้นมัธยมศึกษาปีที่ 1

Word Reading		Spelling		Sentence Comprehension		Math Computation	
Raw score	Scaled score	Raw score	Scaled score	Raw score	Scaled score	Raw score	Scaled score
0	55	0	55	0	55	0	68
1	55	1	55	1	65	1	74
2	55	2	55	2	68	2	78
3	55	3	66	3	71	3	81
4	55	4	66	4	73	4	84
5	55	5	66	5	77	5	87
6	55	6	66	6	80	6	90
7	55	7	66	7	83	7	92
8	55	8	66	8	87	8	94
9	55	9	66	9	90	9	96
10	55	10	66	10	94	10	98
11	55	11	69	11	98	11	100
12	55	12	69	12	102	12	102
13	67	13	69	13	106	13	104
14	67	14	71	14	110	14	106
15	67	15	73	15	114	15	107
16	67	16	74	16	118	16	109
17	67	17	74	17	122	17	110
18	67	18	78	18	127	18	112
19	70	19	80	19	131	19	113
20	70	20	82	20	145	20	115
21	70	21	85	21	145	21	116
22	70	22	88	22	145	22	117
23	70	23	91			23	119
24	75	24	95			24	120
25	75	25	99			25	121
26	78	26	103			26	122
27	80	27	108			27	124
28	82	28	113			28	145
29	84	29	119				
30	87	30	124				
31	89						
32	92						
33	95						
34	98						
35	102						
36	106						
37	110						
38	115						
39	119						

Form C Subtest Standard Score Conversion

ตารางเกณฑ์ปกติที่ 8 เทียบคะแนนดิบสู่คะแนนมาตรฐาน แบบสอบฉบับที่ 3 ชั้นมัธยมศึกษาปีที่ 2

Word Reading		Spelling		Sentence Comprehension		Math Computation	
Raw score	Scaled score	Raw score	Scaled score	Raw score	Scaled score	Raw score	Scaled score
0	55	0	55	0	55	0	72
1	55	1	55	1	66	1	76
2	55	2	55	2	68	2	79
3	55	3	55	3	68	3	82
4	55	4	55	4	73	4	85
5	55	5	65	5	76	5	87
6	55	6	65	6	79	6	90
7	55	7	65	7	82	7	92
8	55	8	65	8	85	8	94
9	55	9	65	9	89	9	95
10	55	10	66	10	93	10	97
11	55	11	66	11	97	11	99
12	55	12	66	12	101	12	101
13	55	13	66	13	105	13	102
14	55	14	66	14	110	14	104
15	55	15	70	15	114	15	106
16	55	16	70	16	119	16	107
17	55	17	72	17	124	17	109
18	55	18	74	18	129	18	110
19	55	19	76	19	134	19	111
20	55	20	79	20	145	20	113
21	55	21	81	21	145	21	114
22	55	22	85	22	145	22	115
23	70	23	88			23	117
24	70	24	92			24	118
25	70	25	97			25	119
26	74	26	103			26	121
27	75	27	108			27	122
28	77	28	115			28	145
29	79	29	122				
30	81	30	131				
31	84						
32	87						
33	90						
34	93						
35	97						
36	102						
37	107						
38	112						
39	118						

Form C Subtest Standard Score Conversion

ตารางเกณฑ์ปกติที่ 9 เทียบคะแนนดิบสู่คะแนนมาตรฐาน แบบสอบฉบับที่ 3 ชั้นมัธยมศึกษาปีที่ 3

Word Reading		Spelling		Sentence Comprehension		Math Computation	
Raw score	Scaled score	Raw score	Scaled score	Raw score	Scaled score	Raw score	Scaled score
0	55	0	55	0	55	0	55
1	55	1	55	1	55	1	74
2	55	2	55	2	55	2	77
3	55	3	55	3	55	3	80
4	55	4	65	4	67	4	82
5	55	5	65	5	69	5	85
6	55	6	65	6	72	6	87
7	55	7	65	7	72	7	89
8	55	8	65	8	78	8	91
9	55	9	65	9	81	9	93
10	55	10	65	10	84	10	95
11	55	11	65	11	88	11	97
12	55	12	65	12	91	12	99
13	55	13	65	13	95	13	101
14	55	14	66	14	99	14	102
15	55	15	66	15	103	15	104
16	55	16	68	16	107	16	106
17	55	17	68	17	112	17	107
18	55	18	68	18	116	18	109
19	55	19	72	19	121	19	110
20	55	20	74	20	126	20	112
21	55	21	76	21	126	21	113
22	55	22	79	22	135	22	115
23	55	23	83			23	116
24	55	24	87			24	118
25	71	25	91			25	119
26	73	26	97			26	120
27	75	27	103			27	122
28	77	28	110			28	123
29	80	29	118				
30	83	30	127				
31	86						
32	89						
33	92						
34	96						
35	100						
36	105						
37	109						
38	114						
39	120						

Form D Subtest Standard Score Conversion

ตารางเกณฑ์ปกติที่ 10 เทียบคะแนนดิบสู่คะแนนมาตรฐาน แบบสอบฉบับที่ 4 ชั้นมัธยมศึกษาปีที่ 4

Word Reading		Spelling		Sentence Comprehension		Math Computation	
Raw score	Scaled score	Raw score	Scaled score	Raw score	Scaled score	Raw score	Scaled score
0	55	0	55	0	55	0	74
1	55	1	55	1	55	1	83
2	55	2	55	2	55	2	89
3	55	3	55	3	69	3	93
4	55	4	55	4	72	4	96
5	55	5	55	5	76	5	99
6	55	6	55	6	80	6	101
7	55	7	69	7	84	7	103
8	55	8	69	8	88	8	105
9	55	9	69	9	92	9	107
10	55	10	69	10	96	10	109
11	55	11	69	11	101	11	110
12	55	12	69	12	106	12	112
13	55	13	69	13	111	13	113
14	55	14	69	14	116	14	114
15	55	15	69	15	121	15	115
16	55	16	71	16	126	16	116
17	55	17	72	17	145	17	116
18	55	18	73	18	145	18	118
19	68	19	73			19	119
20	69	20	76			20	119
21	69	21	77			21	121
22	73	22	80			22	122
23	74	23	82			23	122
24	77	24	85			24	124
25	79	25	89			25	124
26	82	26	93			26	124
27	85	27	98			27	124
28	88	28	103			28	127
29	92	29	110			29	145
30	96	30	117				
31	100						
32	105						
33	110						
34	116						
35	122						

Form D Subtest Standard Score Conversion

ตารางเกณฑ์ปกติที่ 11 เทียบคะแนนดิบสู่คะแนนมาตรฐาน แบบสอบฉบับที่ 4 ชั้นมัธยมศึกษาปีที่ 5

Word Reading		Spelling		Sentence Comprehension		Math Computation	
Raw score	Scaled score	Raw score	Scaled score	Raw score	Scaled score	Raw score	Scaled score
0	55	0	55	0	55	0	72
1	55	1	55	1	55	1	75
2	55	2	55	2	65	2	78
3	55	3	55	3	65	3	81
4	55	4	55	4	69	4	83
5	55	5	55	5	71	5	85
6	55	6	55	6	74	6	87
7	55	7	55	7	77	7	89
8	55	8	55	8	81	8	91
9	55	9	55	9	85	9	93
10	55	10	55	10	89	10	95
11	55	11	55	11	94	11	96
12	55	12	55	12	99	12	98
13	55	13	67	13	104	13	100
14	55	14	67	14	110	14	100
15	55	15	67	15	116	15	103
16	55	16	67	16	122	16	105
17	55	17	67	17	129	17	106
18	55	18	68	18	136	18	108
19	55	19	68			19	109
20	71	20	69			20	111
21	71	21	70			21	112
22	73	22	70			22	114
23	74	23	70			23	115
24	75	24	75			24	117
25	77	25	78			25	118
26	79	26	83			26	120
27	81	27	88			27	121
28	84	28	96			28	122
29	87	29	105			29	124
30	90	30	117				
31	94						
32	99						
33	105						
34	111						
35	119						

Form D Subtest Standard Score Conversion

ตารางเกณฑ์ปกติที่ 12 เทียบคะแนนดิบสู่คะแนนมาตรฐาน แบบสอบฉบับที่ 4 ชั้นมัธยมศึกษาปีที่ 6

Word Reading		Spelling		Sentence Comprehension		Math Computation	
Raw score	Scaled score	Raw score	Scaled score	Raw score	Scaled score	Raw score	Scaled score
0	55	0	55	0	55	0	74
1	55	1	55	1	55	1	81
2	55	2	55	2	55	2	85
3	55	3	55	3	55	3	89
4	55	4	55	4	55	4	92
5	55	5	55	5	55	5	95
6	55	6	55	6	70	6	97
7	55	7	55	7	75	7	99
8	55	8	55	8	79	8	101
9	55	9	55	9	84	9	103
10	55	10	55	10	89	10	104
11	55	11	55	11	94	11	106
12	55	12	55	12	99	12	107
13	55	13	55	13	105	13	108
14	55	14	55	14	110	14	110
15	55	15	55	15	116	15	111
16	55	16	55	16	122	16	112
17	55	17	55	17	128	17	113
18	55	18	55	18	145	18	114
19	55	19	69			19	115
20	71	20	69			20	116
21	71	21	69			21	117
22	71	22	70			22	118
23	73	23	70			23	119
24	74	24	72			24	120
25	75	25	74			25	121
26	77	26	74			26	121
27	78	27	81			27	122
28	81	28	88			28	123
29	84	29	99			29	124
30	87	30	115				
31	91						
32	96						
33	102						
34	109						
35	117						

Form A Subtest Standard Score Conversion

ตารางเกณฑ์ปกติที่ 13 เทียบคะแนนดิบสู่คะแนนมาตรฐาน แบบสอบฉบับที่ 1 ช่วงชั้นประถมศึกษาปีที่ 1-3

Word Reading		Spelling		Sentence Comprehension		Math Computation	
Raw score	Scaled score	Raw score	Scaled score	Raw score	Scaled score	Raw score	Scaled score
0	55	0	55	0	55	0	66
1	55	1	55	1	72	1	71
2	55	2	55	2	73	2	76
3	55	3	55	3	75	3	81
4	68	4	70	4	76	4	84
5	68	5	70	5	79	5	88
6	68	6	70	6	81	6	91
7	68	7	70	7	84	7	95
8	68	8	71	8	87	8	98
9	68	9	71	9	90	9	101
10	68	10	71	10	94	10	104
11	68	11	71	11	98	11	106
12	68	12	72	12	102	12	109
13	68	13	72	13	106	13	112
14	69	14	73	14	111	14	114
15	69	15	74	15	116	15	117
16	69	16	75	16	121	16	119
17	70	17	76			17	121
18	71	18	78			18	124
19	72	19	79			19	126
20	72	20	81			20	128
21	73	21	83				
22	75	22	85				
23	76	23	88				
24	77	24	91				
25	79	25	94				
26	81	26	98				
27	83	27	102				
28	86	28	107				
29	89	29	112				
30	92	30	117				
31	95						
32	99						
33	103						
34	108						
35	113						
36	119						

Form B Subtest Standard Score Conversion

ตารางเกณฑ์ปกติที่ 14 เทียบคะแนนดิบสู่คะแนนมาตรฐาน แบบสอบฉบับที่ 2 ช่วงชั้นประถมศึกษาปีที่

4-6

Word Reading		Spelling		Sentence Comprehension		Math Computation	
Raw score	Scaled score	Raw score	Scaled score	Raw score	Scaled score	Raw score	Scaled score
0	55	0	55	0	55	0	65
1	55	1	55	1	67	1	71
2	55	2	55	2	70	2	75
3	65	3	55	3	72	3	79
4	65	4	70	4	75	4	83
5	65	5	70	5	78	5	87
6	65	6	70	6	81	6	90
7	65	7	71	7	84	7	93
8	65	8	71	8	87	8	96
9	65	9	72	9	90	9	99
10	65	10	73	10	93	10	101
11	65	11	74	11	97	11	104
12	65	12	75	12	100	12	107
13	65	13	76	13	103	13	109
14	65	14	77	14	107	14	112
15	66	15	79	15	110	15	114
16	66	16	80	16	114	16	116
17	66	17	82	17	118	17	118
18	66	18	84	18	121	18	121
19	66	19	86	19	125	19	123
20	67	20	89	20	129	20	125
21	67	21	91			21	127
22	67	22	94			22	129
23	69	23	97			23	129
24	70	24	100			24	133
25	71	25	104			25	135
26	72	26	107				
27	74	27	111				
28	76	28	115				
29	78	29	119				
30	80	30	124				
31	83						
32	86						
33	90						
34	94						
35	99						
36	104						
37	110						
38	117						

Form C Subtest Standard Score Conversion

ตารางเกณฑ์ปกติที่ 15 เทียบคะแนนดิบสู่คะแนนมาตรฐาน แบบสอบฉบับที่ 3 ช่วงชั้นมัธยมศึกษาปีที่ 1-

3

Word Reading		Spelling		Sentence Comprehension		Math Computation	
Raw score	Scaled score	Raw score	Scaled score	Raw score	Scaled score	Raw score	Scaled score
0	55	0	55	0	55	0	70
1	55	1	55	1	64	1	73
2	55	2	55	2	66	2	76
3	55	3	65	3	69	3	78
4	55	4	65	4	71	4	81
5	55	5	65	5	74	5	83
6	55	6	65	6	77	6	85
7	55	7	65	7	80	7	87
8	55	8	65	8	84	8	89
9	55	9	65	9	87	9	92
10	55	10	66	10	91	10	94
11	55	11	67	11	94	11	96
12	55	12	67	12	98	12	97
13	65	13	67	13	102	13	99
14	65	14	69	14	106	14	101
15	65	15	70	15	110	15	103
16	65	16	71	16	114	16	105
17	65	17	72	17	118	17	107
18	65	18	74	18	122	18	109
19	67	19	76	19	127	19	110
20	67	20	78	20	131	20	112
21	67	21	81	21	131	21	114
22	67	22	84	22	140	22	116
23	71	23	88			23	117
24	72	24	92			24	119
25	74	25	96			25	121
26	75	26	101			26	122
27	77	27	106			27	124
28	79	28	113			28	145
29	81	29	119				
30	84	30	127				
31	86						
32	89						
33	92						
34	96						
35	100						
36	104						
37	109						
38	114						
39	119						

Form D Subtest Standard Score Conversion

ตารางเกณฑ์ปกติที่ 16 เทียบคะแนนดิบสู่คะแนนมาตรฐาน แบบสอบฉบับที่ 4 ช่วงชั้นมัธยมศึกษาปีที่ 4-

6

Word Reading		Spelling		Sentence Comprehension		Math Computation	
Raw score	Scaled score	Raw score	Scaled score	Raw score	Scaled score	Raw score	Scaled score
0	55	0	55	0	55	0	73
1	55	1	55	1	55	1	78
2	55	2	55	2	64	2	82
3	55	3	55	3	67	3	86
4	55	4	55	4	69	4	89
5	55	5	55	5	72	5	91
6	55	6	55	6	76	6	93
7	55	7	68	7	79	7	96
8	55	8	68	8	83	8	97
9	55	9	68	9	87	9	99
10	55	10	68	10	92	10	101
11	55	11	68	11	96	11	103
12	55	12	68	12	101	12	104
13	55	13	68	13	106	13	106
14	55	14	68	14	112	14	107
15	55	15	68	15	117	15	109
16	55	16	68	16	123	16	110
17	55	17	68	17	129	17	111
18	55	18	69	18	136	18	112
19	70	19	69			19	114
20	70	20	70			20	115
21	70	21	71			21	116
22	72	22	73			22	117
23	74	23	75			23	118
24	75	24	77			24	119
25	77	25	80			25	120
26	79	26	85			26	121
27	81	27	90			27	122
28	84	28	97			28	123
29	88	29	105			29	124
30	91	30	116				
31	95						
32	100						
33	106						
34	112						
35	119						

ภาคผนวก ค

สรุปจำนวนนักเรียนที่ทำแบบทดสอบ (เก็บข้อมูล norm) จำแนกตาม วิชา ระดับชั้น โรงเรียน

การอ่านคำ

ระดับชั้น	จำนวนนักเรียน (ตามรหัสโรงเรียน)											รวม จำนวน นักเรียน
	1101	1102	1103	1104	2111	2112	2113	2114	2119	2120	2121	
ป.1	12	12	12	11	8	8	8	8	8	8	8	103
ป.2	12	12	12	11	8	8	8	8	8	8	8	103
ป.3	12	12	9	12	8	8	8	8	8	8	8	101
ป.4	12	12	12	12	8	8	8	8	8	8	8	104
ป.5	12	12	12	12	8	8	8	8	8	8	8	104
ป.6	12	12	12	12	8	8	8	8	8	8	8	104
รวม	72	72	69	70	48	48	48	48	48	48	48	619

ระดับชั้น	จำนวนนักเรียน (ตามรหัสโรงเรียน)										รวมจำนวน นักเรียน
	1205	1206	1207	1208	2215	2216	2217	2218	2222	2223	
ม.1	12	12	12	12	8	8	12	12	8	8	104
ม.2	12	12	12	12	8	8	12	12	8	8	104
ม.3	12	12	12	11	8	8	12	12	8	8	103
ม.4	12	12	12	12	8	8	12	12	8	8	104
ม.5	12	12	12	10	8	8	12	12	8	8	102
ม.6	12	12	12	11	8	8	12	12	8	8	103
รวม	72	72	72	68	48	48	72	72	48	48	620

การสะกตคำ

ระดับชั้น	จำนวนนักเรียน (ตามรหัสโรงเรียน)											รวม จำนวน นักเรียน
	1101	1102	1103	1104	2111	2112	2113	2114	2119	2120	2121	
ป.1	12	12	12	11	8	8	8	8	8	8	8	103
ป.2	12	12	12	11	8	8	8	8	8	8	8	103
ป.3	12	12	12	12	8	8	8	8	8	8	8	104
ป.4	12	12	12	12	8	8	8	8	8	8	8	104
ป.5	12	12	12	12	8	8	8	8	8	8	8	104
ป.6	12	12	12	11	8	8	8	8	8	8	8	103
รวม	72	72	72	69	48	48	48	48	48	48	48	621

ระดับชั้น	จำนวนนักเรียน (ตามรหัสโรงเรียน)										รวมจำนวน นักเรียน
	1205	1206	1207	1208	2215	2216	2217	2218	2222	2223	
ม.1	12	12	12	12	8	8	12	12	8	8	104
ม.2	12	12	12	12	8	8	12	12	8	8	104
ม.3	11	12	12	12	8	8	12	12	8	8	103
ม.4	12	12	12	12	8	8	12	12	8	8	104
ม.5	12	12	12	10	8	8	12	12	8	8	102
ม.6	12	11	12	10	8	8	12	12	8	8	101
รวม	71	71	72	68	48	48	72	72	48	48	618

ความเข้าใจประโยค

ระดับชั้น	จำนวนนักเรียน (ตามรหัสโรงเรียน)											รวม จำนวน นักเรียน
	1101	1102	1103	1104	2111	2112	2113	2114	2119	2120	2121	
ป.1	12	12	12	11	8	8	8	8	8	8	8	103
ป.2	12	12	12	11	8	8	8	8	8	8	8	103
ป.3	12	12	12	12	8	8	8	8	8	8	8	104
ป.4	12	12	12	12	8	8	8	8	8	8	8	104
ป.5	12	12	12	12	8	8	8	8	8	8	8	104
ป.6	12	12	12	11	8	8	8	8	8	8	8	103
รวม	72	72	72	69	48	48	48	48	48	48	48	621

ระดับชั้น	จำนวนนักเรียน (ตามรหัสโรงเรียน)										รวมจำนวน นักเรียน
	1205	1206	1207	1208	2215	2216	2217	2218	2222	2223	
ม.1	12	12	12	12	8	8	12	12	8	8	104
ม.2	12	12	12	12	8	8	12	12	8	8	104
ม.3	11	12	12	11	8	8	12	12	8	8	102
ม.4	12	12	12	12	8	8	12	12	8	8	104
ม.5	12	12	12	10	8	8	12	12	8	8	102
ม.6	12	11	12	10	8	8	12	12	8	8	101
รวม	71	71	72	67	48	48	72	72	48	48	617

การคิดคำนวณ

ระดับชั้น	จำนวนนักเรียน (ตามรหัสโรงเรียน)											รวม จำนวน นักเรียน
	1101	1102	1103	1104	2111	2112	2113	2114	2119	2120	2121	
ป.1	12	12	12	12	8	8	8	8	8	8	8	104
ป.2	12	12	12	11	8	8	8	8	8	8	8	103
ป.3	12	12	12	12	8	8	8	8	8	8	8	104
ป.4	12	12	12	12	8	8	8	8	8	8	8	104
ป.5	12	12	12	12	8	8	8	8	8	8	8	104
ป.6	12	12	12	11	8	8	8	8	8	8	8	103
รวม	72	72	72	70	48	48	48	48	48	48	48	622

ระดับชั้น	จำนวนนักเรียน (ตามรหัสโรงเรียน)										รวมจำนวน นักเรียน
	1205	1206	1207	1208	2215	2216	2217	2218	2222	2223	
ม.1	12	12	12	12	8	8	12	12	8	8	104
ม.2	12	12	12	12	8	8	12	12	8	8	104
ม.3	11	12	12	11	8	8	12	12	8	8	102
ม.4	12	12	12	12	8	8	12	12	8	8	104
ม.5	12	12	12	10	8	8	12	12	8	8	102
ม.6	12	11	12	10	8	8	12	12	8	8	101
รวม	71	71	72	67	48	48	72	72	48	48	617