

การประยุกต์เทคนิคการเทียบมาตราคะแนนแนวดิ่งเพื่อพัฒนาแบบสอบทักษะพื้นฐานทาง วิชาการสำหรับบุคคลที่มีความต้องการจำเป็นพิเศษ มหาวิทยาลัยเกษตรศาสตร์

An Application of Vertical Equating for development of Kasetsart Basic Academic Skills Test for Person with Special needs

ภาวิณี ศรีสุขวัฒนานนท์* ชุตินา หทัยรัตนนา** และดารณี อุทัยรัตนิก*

Pawinee Srisukvatananan Chutima Hatairatana and Daranee Utajratanakit

* หลักสูตรศึกษาศาสตรมหาบัณฑิต สาขาวิชาการศึกษาพิเศษ มหาวิทยาลัยเกษตรศาสตร์

** คณะวิทยาศาสตร์ มหาวิทยาลัยหอการค้าไทย

บทคัดย่อ

แบบสอบทักษะพื้นฐานทางวิชาการสำหรับบุคคลที่มีความต้องการจำเป็นพิเศษมหาวิทยาลัยเกษตรศาสตร์เป็นชุดแบบสอบวัดทักษะพื้นฐานทางวิชาการเพื่อวัดนักเรียนรายบุคคล ประกอบด้วยแบบสอบย่อย 4 ทักษะ ได้แก่ ทักษะการอ่านคำ ทักษะการสะกดคำ ทักษะความเข้าใจประโยค และทักษะการคำนวณทางคณิตศาสตร์ จัดเป็นชุดแบบสอบด้วยเทคนิคการเทียบมาตราคะแนนแนวดิ่งที่ทำให้แบบสอบแต่ละทักษะมีคะแนนความสามารถที่ปรับเทียบจากระดับชั้นประถมศึกษาปีที่ 1 ถึง มัธยมศึกษาปีที่ 6 ให้อยู่บนมาตราคะแนนเดียวกัน แบบสอบฉบับที่ 1 วัดทักษะชั้นประถมศึกษาปีที่ 1 ถึง ประถมศึกษาปีที่ 3 ฉบับที่ 2 วัดทักษะชั้นประถมศึกษาปีที่ 4 ถึง ประถมศึกษาปีที่ 6 ฉบับที่ 3 วัดทักษะชั้นมัธยมศึกษาปีที่ 1 ถึง มัธยมศึกษาปีที่ 3 และฉบับที่ 4 วัดทักษะชั้นมัธยมศึกษาปีที่ 4 ถึงมัธยมศึกษาปีที่ 6

กระบวนการพัฒนาแบบสอบประยุกต์เทคนิคการปรับเทียบมาตราคะแนนแนวดิ่ง ยึดโครงสร้างรูปแบบความเจริญตามระดับชั้น ที่มีข้อสอบรวมอยู่ในแบบสอบย่อยรายทักษะแต่ละฉบับ การดำเนินการประกอบด้วย 3 ขั้นตอน **ขั้นที่หนึ่ง** สร้างแบบสอบรายทักษะ ข้อสอบในแบบสอบคัดเลือกจากคลังข้อสอบที่สอดคล้องกับลักษณะเฉพาะตามค่าความยากตามทฤษฎีการตอบสนองข้อสอบ และอยู่ในข้อกำหนดด้านเนื้อหาขั้นพื้นฐานในแต่ละช่วงชั้น **ขั้นที่สอง** เลือกข้อสอบรวมที่เข้าไปทำหน้าที่เชื่อมคะแนนความสามารถภายในทักษะที่วัด 4 ช่วงชั้นให้เป็นมาตราคะแนนเดียวกันที่สอดคล้องกับรูปแบบความเจริญตามลำดับชั้น **ขั้นที่สาม** ปฏิบัติการเก็บข้อมูลกับกลุ่มตัวอย่างเทียบมาตราคะแนนแบบสอบ ซึ่งเป็นนักเรียนทุกระดับชั้นที่ได้สุ่มจากประชากรนักเรียนในกรุงเทพมหานครและจังหวัดขอนแก่น จัดกลุ่มคะแนนผู้สอบเป็น 4 กลุ่มช่วงชั้น กำหนดใช้ผลการวิเคราะห์จากกลุ่มที่ 3 (ม.1-3) เป็นฐานการปรับระบบคะแนน คำนวณ

ค่าพารามิเตอร์ของข้อสอบรวมแล้วนำค่าประมาณที่ได้ไปปรับค่าประมาณของข้อสอบรวมในกลุ่มที่ 2 และ 4 ด้วยวิธีการปรับเชิงเส้น และใช้วิธีเดียวกันนี้กับกลุ่มที่ 1 จากนั้นประมาณค่าพารามิเตอร์ของข้อสอบทั้งหมดพร้อมกันในกระบวนการทำงานครั้งเดียวของโปรแกรมไบล็อก จะได้ผลแสดงการกระจายความสามารถในแต่ละช่วงชั้น จากกระบวนการเทียบมาตราคะแนนแนวดิ่งนี้ การตอบข้อสอบในรายทักษะใดทักษะหนึ่งจะได้คะแนนความสามารถของบุคคลในทักษะนั้นและมีความหมายบนมาตราคะแนนเดียวกันของทักษะนั้นตลอด 4 ฉบับ

คุณภาพของชุดแบบสอบแสดงด้วย 1) ความตรงเชิงเนื้อหา จากการวิเคราะห์เชิงเหตุผลของผู้เชี่ยวชาญด้านหลักสูตรและการสอน และผ่านการสร้างข้อสอบอย่างพิถีพิถัน 2) ความตรงเชิงโครงสร้างรูปแบบความเจริญตามลำดับชั้น พบว่า แบบวัดทักษะสะกดคำ ความเข้าใจประโยค และการคำนวณทางคณิตศาสตร์มีความตรงเชิงโครงสร้างตามรูปแบบความเจริญตามลำดับชั้น โดยมีค่าความยากเพิ่มขึ้นตามลำดับชั้นอย่างมีนัยสำคัญทางสถิติ ส่วนแบบวัดทักษะการอ่านคำ มีความไม่สอดคล้องในช่วงชั้นที่ 3 และ 4 3) ความเที่ยงของแบบสอบทั้ง 4 ทักษะ มีค่าอยู่ระหว่าง 0.680 - 0.950

คำสำคัญ: แบบสอบมาตรฐานวัดทักษะพื้นฐานทางวิชาการ เทคนิคการเทียบคะแนนแนวดิ่ง ข้อสอบรวม โครงสร้างรูปแบบความเจริญตามช่วงชั้น บุคคลที่มีความต้องการจำเป็นพิเศษ

Abstract

The Kasetsart Basic Academic Skill Test (KBAST) is a battery test of basic academic skills which measures an individual student. It composed of four skill subtests: Word Reading,

Word Spelling, Sentence Comprehension, and Math Computation. The test was assigned in order to build a vertical scale to equate each grade's test form to all test forms for grade 1 through grade 12. The first form for individuals grades 1–3, the second form for individuals grades 4–6, the third form for individuals grade 7–9, and the fourth form for individuals grade 10–12.

There are 3 main steps of applying vertical equating techniques using grade-to-grade growth model with an internal common-item anchor-test data collection design. First, subtests were constructed over basic skill content of the curriculum in grade 1 through grade 12. The subtest forms were groups of selected items from item bank to math IRT difficulty specifications as well as content and format specification. Second, a set of common items were selected and embedded in the subtest forms to create a Θ -scale of the subtest that spans a range of grades fitted to the grade-to-grade growth model. Third, data were collected from equating sample students of grade 1 through grade 12 from the target population students in Bangkok and Khongon province. Data from individuals were grouped into four groups according to the test form they took. The third group (grade 7-9) was chosen as the base scale. The set of common items were run for item parameter estimates and used these item parameters with a linear transformation to place the item parameters and ability distributions for the group 2 and 4 examinees on the Θ -scale that was established for group 3, then used the same process for group 1. With concurrent estimation, parameters for all items were estimated in a single computer run of BILOG-MG. In this run, the grade group was identified so those ability distributions were estimated for each grade groups. By this process of vertical equating, the response on the 4 forms of each skill subtest can be identified and interpreted the ability on the same ability scale.

The quality of KBAST was shown 1) the content validity by performing of a group of content experts on the items that carefully constructed, 2) the construct validity based on the grade to grade growth model was shown the ability distributions of groups of examinees and significant tests. The results found that There were three subtests were completely matched with the growth model significantly, except the subtest of word reading, 3) The coefficient index (α) of the reliability of The four skill subtests were among ranged 0.680-0.950.

Keywords: Battery Test of Basic Academic Skills, Technique of Vertical Equating, Common Items Grade-to-Grade Growth Model, Special Need Person

บทนำ

การจัดการศึกษาสำหรับเด็กที่มีความต้องการจำเป็นพิเศษ จำเป็นต้องมีการวางแผนการจัดการศึกษารายบุคคล (Individual Education Plan: IEP) ที่ตรงกับสภาพและสอดคล้องกับความต้องการจำเป็นพิเศษของผู้เรียน ข้อมูลที่ต้องนำมาพิจารณา ได้แก่ ผลการคัดกรองความบกพร่องเบื้องต้น รายงานการประเมินทางจิตวิทยา ความพร้อมทางวิชาการ พฤติกรรมทางสังคม และลักษณะเฉพาะที่มีการติดตามให้ความช่วยเหลือจากแพทย์ ผู้ปกครองและผู้เกี่ยวข้อง สำหรับข้อมูลความพร้อมทางวิชาการ ครูผู้ทำหน้าที่ประเมิน จำเป็นต้องมีเครื่องมือที่มีคุณภาพ เพื่อให้ได้สารสนเทศที่ตรงสภาพจริงของนักเรียนแต่ละบุคคล และเพื่อให้ความช่วยเหลืออย่างต่อเนื่องต่อไป

แบบสอบทักษะพื้นฐานทางวิชาการมหาวิทยาลัยเกษตรศาสตร์ (Kasetsart Basic Academic Skills Test) เรียกชื่อย่อว่า KBAST เป็นผลผลิตของโครงการวิจัย ซึ่งได้รับทุนอุดหนุนจากสำนักงานคณะกรรมการการวิจัยแห่งชาติ (2557) เพื่อใช้กับบุคคลที่มีความต้องการจำเป็นพิเศษในระดับการศึกษาขั้นพื้นฐาน ใช้เวลาดำเนินการขั้นต้น 1 ปี และพัฒนาการนำไปใช้อีก 1 ปี หลังจากนั้นคณะผู้วิจัยได้เผยแพร่แก่สถานศึกษา และหน่วยงานที่เกี่ยวข้อง เพื่อวางแผนการให้ความช่วยเหลือบุคคลที่มีภาวะสมาธิสั้น บกพร่องทางการเรียนรู้ ออทิสซึมและบุคคลที่มีความต้องการจำเป็นประเภทอื่นๆ

บทความนี้ จะนำเสนอกระบวนการพัฒนา ด้านเทคนิคเพื่อให้ได้แบบวัดทักษะพื้นฐานทางวิชาการที่มี

มาตรฐาน โดยการประยุกต์เทคนิคการเทียบมาตราคะแนนแนวตั้ง (Vertical Equating Techniques) ครอบคลุมกลุ่มเป้าหมายบุคคลที่มีความต้องการจำเป็นพิเศษเป็นรายบุคคลในระดับการศึกษาขั้นพื้นฐาน คือชั้นประถมศึกษาปีที่ 1 ถึง มัธยมศึกษาปีที่ 6 แบบสอบชุดนี้วัดระดับทักษะที่มีช่วงยาว 12 ชั้นปี จึงแบ่งชุดแบบสอบวัดทักษะทางวิชาการพื้นฐานตามช่วงชั้น 4 ฉบับ ดังนี้ ฉบับที่ 1 ประถมศึกษาปีที่ 1-3 ฉบับที่ 2 ช่วงชั้นประถมศึกษาปีที่ 4-6 ฉบับที่ 3 ช่วงชั้นมัธยมศึกษาปีที่ 1-3 และฉบับที่ 4 ช่วงชั้นมัธยมศึกษาปีที่ 4-6 ในแต่ละฉบับมี 4 แบบสอบย่อยวัดความสามารถ 4 ทักษะ ดังนั้นแบบสอบทั้งชุดจึงต้องสร้างมาตราการให้คะแนนที่สามารถแปลความหมายตามโครงสร้างรูปแบบความเจริญตามลำดับชั้นกรอบแนวคิดการพัฒนาแบบสอบ และการดำเนินการสร้างแบบสอบ มีรายละเอียดดังนี้

กรอบแนวคิดการพัฒนาแบบสอบ

คณะผู้วิจัยพัฒนากรอบแนวคิดการพัฒนาแบบสอบ KBAST ศึกษางานวิจัยที่เกี่ยวข้องดังนี้

1. Wide Range Achievement Test 4 (WRAT 4)

แบบสอบ Wide Range Achievement Test 4 (WRAT 4) เป็นแบบสอบทักษะพื้นฐานทางวิชาการที่มีประสิทธิภาพสูง และใช้ในประเทศที่ใช้ภาษาอังกฤษ มีประวัติการพัฒนาที่ยาวนานตั้งแต่ ทศวรรษที่ 1930 โดยนักจิตวิทยา Joseph F. Jastak โดยขยายแบบวัด Wechsler-Bellevue Scales ของ David Wechsler ที่สร้างเมื่อประมาณ 70 ปีก่อน วัดทักษะพื้นฐานทางวิชาการสำคัญ 3 ด้าน คือ การอ่าน การเขียน และการคิดคำนวณ ต่อมาได้ตีพิมพ์เผยแพร่ครั้งแรก เมื่อ ค.ศ. 1946 จากนั้นมีการพัฒนาต่อเนื่อง จนถึง ปี ค.ศ. 1993 ได้เผยแพร่ฉบับ Wide Range Achievement Test: Third Edition (WRAT 3) แล้วพัฒนาต่อเนื่อง ฉบับที่ใช้อยู่ในปัจจุบัน คือ Wide Range Achievement Test: Fourth Edition (WRAT 4) ตีพิมพ์เมื่อ ค.ศ. 2006 โดยเพิ่มการวัดทักษะความเข้าใจประโยค (Sentence Comprehension) และพัฒนาการแปลความหมายคะแนนสอบด้วยคะแนนมาตรฐานเทียบตามกลุ่มอายุ (Age-based norms) จาก 5 ปี ถึง 94 ปี และมีคะแนนมาตรฐานเทียบกับระดับชั้นเรียน (grade-based norms) (Wilkinson & Robertson, 2006)

กระบวนการพัฒนาแบบสอบ WRAT ดำเนินการตามระเบียบวิธีการวิจัยและพัฒนาที่มีความเป็นมาตรฐาน เริ่มวางกรอบแนวคิดวัดทักษะที่จำเป็นต่อการเรียนรู้สาม

ด้าน คือ การอ่าน การเขียน และการคิดคำนวณ ดำเนินการสร้างเครื่องมือทุกขั้นตอนตามหลักคุณภาพพร้อมแสดงหลักฐานด้านความตรง ความเที่ยง และสร้างมาตรฐานการแปลความหมายคะแนนด้วยเกณฑ์มาตรฐานตามอายุ และระดับชั้นเรียนของผู้สอบ โดยสรุปกระบวนการและเทคนิคการพัฒนา ดังนี้

แบบสอบ WRAT 4

เริ่มด้วยการทบทวนและปรับปรุง WRAT3 ที่มีแบบสอบวัดทักษะ 3 ฉบับ แล้วสร้างแบบสอบฉบับที่ 4 Sentence Comprehension ได้แบบสอบ WRAT 4 ที่ประกอบด้วยแบบสอบดังนี้

(1) Word Reading เดิมมี 84 ข้อ หลังตัดและเพิ่มแล้ว สรุปได้ข้อสอบใน WRAT 4 จำนวน 55 ข้อ

(2) Spelling เดิมมี 80 ข้อ หลังตัดและเพิ่มแล้ว สรุปได้ข้อสอบใน WRAT 4 จำนวน 45 ข้อ

(3) Math Computation เดิมมี 80 ข้อ หลังตัดและเพิ่มแล้ว ได้ข้อสอบใน WRAT 4 จำนวน 80 ข้อ

(4) Sentence Comprehension สร้างใหม่ตามกระบวนการพัฒนาแบบสอบมาตรฐาน มีข้อสอบใน WRAT 4 จำนวน 100 ข้อ

การพัฒนามาตรฐานคะแนน ของ WRAT 4

(1) สร้าง norms 19 กลุ่มอายุตั้งแต่ 5 - 94 ปี ของประชากรสหรัฐอเมริกา

(2) สร้าง norms ระดับชั้นเรียนจากนักเรียนชั้นอนุบาล จนถึงชั้นเรียน 12

การเทียบแบบสอบ 2 ฟอรม์ (Equating of Forms)

ชุดแบบสอบ WRAT 4 มี 2 ฟอรม์ที่คู่ขนานกัน มีกระบวนการสร้างตามลักษณะเฉพาะที่กำหนด และปรับเทียบมาตราคะแนนให้อยู่ในมาตราคะแนนร่วม ตามระเบียบการเทียบมาตราคะแนน และใช้มาตรฐาน norms เดียวกัน

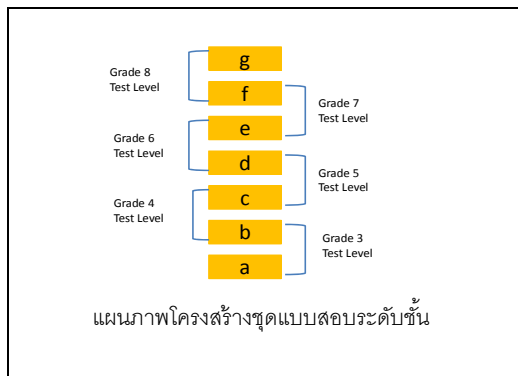
2. แนวคิดการสร้างชุดแบบสอบที่มีความยากต่างกันตามรูปแบบ Grade to Grade Growth Model

การสร้างชุดแบบสอบที่มีความยากต่างระดับกัน (multiple test levels) เพื่อใช้กับผู้สอบในระดับชั้น หรืออายุที่แตกต่างกัน ต้องออกแบบและสร้างระบบคะแนนให้มีมาตรฐานที่สื่อถึงระดับความสามารถของผู้สอบ กระบวนการนี้ทำได้โดยประยุกต์เทคนิคการเทียบมาตราคะแนนแนวตั้ง

การสร้างชุดแบบสอบที่มีความยากต่างกัน

Kolen & Brennan (2004; 372-380) แนะนำให้ใช้เทคนิคการเทียบมาตราคะแนนแนวดิ่ง ดังตัวอย่าง ชุดแบบสอบผลสัมฤทธิ์ระดับประถมศึกษา The Iowa Test of Basic Skills (ITBS) ของ Hoover และ คณะ (2003) ชุดแบบสอบวัดความถนัดระดับประถมศึกษา The Cognitive Abilities Test (CogAT) ของ Lohman และ Hagen (2002) ชุดแบบสอบดังกล่าว ประกอบด้วยจำนวน

แบบสอบย่อยหลายสาขา และใช้กับนักเรียนระดับชั้นต่างๆ แบบสอบที่วัดแต่ละระดับชั้นจะออกแบบคำถามที่สอดคล้องกับหลักสูตรของระดับนั้นๆ ข้อคำถามจะมีระดับความยากที่สอดคล้องกับระดับชั้น ทำให้การประยุกต์เทคนิคการเทียบมาตราคะแนนของชุดแบบสอบมีความเป็นมาตราเดียวกัน (common scale) ตัวอย่างการออกแบบเพื่อใช้ในแต่ละระดับชั้น ประกอบด้วยกลุ่มข้อสอบ 2 กลุ่ม ดังแผนภาพที่ 1



มีข้อสอบอยู่ 7 กลุ่ม

กลุ่ม a และ b เป็นส่วนของแบบสอบระดับชั้นที่ 3

กลุ่ม b และ c เป็นส่วนของแบบสอบระดับชั้นที่ 4

กลุ่ม c และ d เป็นส่วนของแบบสอบระดับชั้นที่ 5

กลุ่ม d และ e เป็นส่วนของแบบสอบระดับชั้นที่ 6

กลุ่ม e และ f เป็นส่วนของแบบสอบระดับชั้นที่ 7

กลุ่ม f และ g เป็นส่วนของแบบสอบระดับชั้นที่ 8

แผนภาพที่ 1 โครงสร้างชุดแบบสอบวัดระดับชั้น 3 - 8 (Kolen & Brennan, 2004: 374)

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

แบบสอบระดับ 3 ข้อสอบ a และข้อสอบร่วม b

แบบสอบระดับ 4 ข้อสอบ b และข้อสอบร่วม c

แบบสอบระดับ 5 ข้อสอบ c และข้อสอบร่วม d

แบบสอบระดับ 6 ข้อสอบ d และข้อสอบร่วม e

แบบสอบระดับ 7 ข้อสอบ e และข้อสอบร่วม f

แบบสอบระดับ 8 ข้อสอบ f และข้อสอบร่วม g

แผนภาพที่ 2 การออกแบบสร้างแบบสอบใช้ข้อสอบร่วม (Common-item design)

จากการออกแบบนี้ ข้อสอบร่วมจะทำหน้าที่เชื่อมคะแนนแต่ละระดับให้อยู่บนมาตรวัดคะแนนเดียวกัน ผู้สอบจะได้รับชุดแบบสอบที่เหมาะสมกับระดับชั้นของตนเองตามที่กำหนด ผลสอบของแต่ละคนจะบอกระดับความสามารถในชั้นของตนและระดับชั้นที่สูงขึ้นได้ ข้อมูลทั้งหมดตามการออกแบบนี้จะนำมากำหนดเป็นคะแนนของทุกระดับชั้นบนระบบคะแนนร่วม (common scale)

การคิดคะแนนแบบสอบ (Test scoring)

การคิดคะแนนแบบสอบที่วัดระดับความสามารถในแนวดิ่ง Kolen & Brennan (2010: 387-388) แนะนำให้ใช้ตัวอย่างจากการออกแบบในแผนภาพที่ 2 ดังนี้

1. วิเคราะห์ผลการตอบแบบสอบระดับชั้นที่ 3 ถึง 8 แยกวิเคราะห์ หาค่าประมาณคุณลักษณะข้อสอบ (item parameters) และการแจกแจงความสามารถ

2. ตั้งข้อกำหนดให้มาตรฐานความสามารถ (Θ) ของแบบสอบระดับชั้นที่ 3 เป็นฐานการปรับคะแนน นำค่าความยาก (b) ของข้อสอบรวมที่อยู่ในแบบสอบระดับชั้นที่ 3 กับค่าความยาก (b) ของข้อสอบรวมในแบบสอบระดับชั้นที่ 4 มาปรับแปลงคะแนนสอบของผู้สอบแบบสอบระดับชั้นที่ 4 ด้วยวิธีการเชิงเส้น ถ้าพบข้อสอบที่ปรากฏค่าผิดเพี้ยนจากระดับชั้นทั้งสองให้ตัดออกจากกลุ่มข้อสอบรวม ผลการแปลงจะนำเข้ามาประมาณค่าพารามิเตอร์ข้อสอบ (b) และค่าประมาณความสามารถ (Θ) ของผู้สอบในแบบสอบระดับชั้นที่ 4

3. ปรับมาตราคะแนนความสามารถของผู้ที่สอบแบบสอบระดับชั้นที่ 5, 6, 7, และ 8 ตามลำดับ ด้วยกระบวนการทำนองเดียวกับที่ปรับคะแนนความสามารถระดับชั้นที่ 4 ให้อยู่บนเส้นฐานมาตรฐานความสามารถระดับชั้นที่ 3 เมื่อกระบวนการเสร็จสิ้นลง จะทำให้คะแนนจากชุดแบบสอบทั้งหมด 6 ฟอรัมมีมาตรวัดความสามารถเดียวกัน การแปลความหมายผลการสอบจะให้สารสนเทศรายชั้น และตลอดทั้ง 6 ชั้นเรียนด้วย

กำหนดระดับเส้นฐานความสามารถของผู้สอบ

โดยปกติ แบบสอบที่เป็นชุด (batteries) จะประกอบด้วยจำนวนข้อสอบค่อนข้างมาก เมื่อมีการจัดวางข้อสอบเป็นลำดับตามข้อง่ายไปสู่ข้อยาก และการกำหนด

ตารางที่ 1 ผังกำหนดลักษณะแบบสอบ ทักษะพื้นฐานทางวิชาการ

แบบสอบทักษะ	เนื้อหา	ระดับชั้น	ลักษณะของข้อสอบ
การอ่านคำ	พยัญชนะ	ป.1 – ป.3	กำหนดพยัญชนะ ให้ผู้สอบอ่านออกเสียง
	คำ	ป.1 – ม.6	กำหนดคำ ให้ผู้สอบอ่านออกเสียง
การสะกดคำ	พยัญชนะ	ป.1 – ป.3	ครูอ่านพยัญชนะ ให้ผู้สอบเขียนตามคำบอก
	คำ	ป.1 – ม.6	ครูอ่านคำ ให้ผู้สอบเขียนตามคำบอก
ความเข้าใจประโยค	ความรู้จากวิชาต่างๆ	ป.1 – ม.6	กำหนดให้ผู้สอบเติมคำที่เว้นเพื่อให้ข้อความสมบูรณ์ เรื่องที่นำมาใช้ เกี่ยวกับความรู้ทั่วไปในวิชาต่างๆ ที่เหมาะสมกับระดับชั้น
การคำนวณทางคณิตศาสตร์	จำนวนเต็มบวก ลบ	ป.1 – ป.3	เติมคำตอบในโจทย์ที่กำหนดให้
	คูณ หาร	ป.1 – ม.3 ม.1-ม.6	

เมื่อผู้เชี่ยวชาญสร้างข้อสอบตามข้อกำหนดลักษณะเฉพาะ แล้วประชุมพิจารณาตรวจสอบร่วมกันด้วยหลักเหตุผล ได้ข้อสอบดังตารางที่ 2

กฎการยุติการสอบเมื่อผู้สอบทำผิดติดต่อกัน 3 ข้อ หรือ 5 ข้อ หรือ 7 ข้อ เหตุผลของการยุติการสอบเช่นนี้ ถือเป็นจุดที่ชี้ระดับเส้นฐานความสามารถของผู้ตอบ (A basal score) ผู้สอบไม่ต้องทำข้อที่ยากเกินความสามารถ เพราะอาจเกิดการเดา การกำหนดจำนวนข้อที่ผิดติดต่อกันเป็นเหตุผลของผู้สร้างแบบสอบ ในแบบสอบ WRAT 4 ใช้กฎการยุติการสอบเมื่อเกิดการทำผิดติดต่อกัน 5 ข้อในแบบสอบย่อยความเข้าใจประโยค และ 7 ข้อ ในแบบสอบย่อยการคำนวณทางคณิตศาสตร์

การพัฒนาแบบสอบ KBAST

จากการทบทวนแนวคิด ทฤษฎี และแนวทางปฏิบัติ ดังที่ได้นำเสนอ จึงกำหนดเป็นกรอบพัฒนาแบบสอบมาตรฐาน KBAST ให้เป็นแบบสอบวัดทักษะพื้นฐานทางวิชาการ 4 ทักษะ คือ การอ่านคำ (Word Reading) การสะกดคำ (Word Spelling) ความเข้าใจประโยค (Sentence Comprehension) และการคำนวณทางคณิตศาสตร์ (Math Computation) ดังนี้

1. การพัฒนาข้อสอบในชุดแบบสอบ KBAST

กำหนดหลักสูตรการสอบและสร้างผังกำหนดลักษณะแบบสอบโดยผู้วิจัยร่วมกับผู้เชี่ยวชาญ สรุปดังตารางที่ 1

ตารางที่ 2 จำนวนข้อสอบที่ผ่านกระบวนการวิเคราะห์คุณภาพเบื้องต้น

แบบสอบตามระดับ ช่วงชั้น	กลุ่มข้อสอบ			
	การอ่านคำ	การสะกดคำ	เข้าใจประโยค	การคำนวณทาง คณิตศาสตร์
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

การตรวจสอบคุณภาพของข้อสอบ 4 กลุ่มทักษะ รายข้อ ใช้กลุ่มตัวอย่างนักเรียนชั้นประถมศึกษาปีที่ 1 ถึง มัธยมศึกษาปีที่ 6 ในเขตกรุงเทพมหานคร จำแนกตาม สังกัด ระดับชั้นละ 200 คน ชุดแบบสอบย่อยเป็นกลุ่ม ข้อสอบที่จัดโดยผู้สอบในระดัชั้น แล้ววิเคราะห์ ค่าพารามิเตอร์ความยาก (b) และอำนาจจำแนก (a) ด้วย Item response theory ใช้โปรแกรม BILOG แล้วบันทึก ค่าพารามิเตอร์รายข้อในคลังข้อสอบพร้อมสร้างรหัส ข้อสอบกำกับลักษณะเฉพาะตามเนื้อหาข้อสอบ เพื่อ สะดวกกับการจัดฟอร์มแบบสอบให้มีลักษณะเนื้อหาและ ความยากตามกำหนดต่อไป

2. สร้างชุดแบบสอบตามโครงสร้างรูปแบบความเจริญ ตามระดับชั้นและข้อสอบรวม

เลือกข้อสอบจากคลังข้อสอบที่จัดเตรียมไว้ที่มี ความเหมาะสมในด้านเนื้อหาขั้นพื้นฐานตามช่วงชั้น และมีความ ยากอยู่ในกรอบที่ลดหลั่นดังผังกำหนดค่าประมาณ ความยาก เพื่อให้สอดคล้องกับการวัดทักษะที่มีความยาก ตามระดับชั้นจากน้อยไปหามาก แบบสอบย่อยแต่ละฉบับ ประกอบด้วยกลุ่มข้อสอบ 2 ส่วน คือกลุ่มเฉพาะช่วงชั้น และกลุ่มข้อสอบรวม ดังแผนภาพที่ 3

แบบสอบ	กลุ่มข้อสอบ: ประมาณค่าความยากของข้อสอบ (b) และจำนวนข้อ						
	กลุ่ม A ข้อสอบ เฉพาะ	กลุ่ม B ข้อสอบรวม	กลุ่ม C ข้อสอบ เฉพาะ	กลุ่ม D ข้อสอบรวม	กลุ่ม E ข้อสอบ เฉพาะ	กลุ่ม F ข้อสอบรวม	กลุ่ม G ข้อสอบ เฉพาะ
ฉบับที่ 1 (ป.1-3)	-2.5<b<-3 45 ข้อ	-0.25<b<-2.0 10 ข้อ					
ฉบับที่ 2 (ป.4-6)		-0.25<b<-2.0 10 ข้อ	-2.0<b<-1.5 40 ข้อ	-1.5<b<-1.0 5 ข้อ			
ฉบับที่ 3 (ม.1-3)			-	-1.5<b<-1.0 5 ข้อ	-1.0<b<1.0 35 ข้อ	1.0<b<1.5 5 ข้อ	
ฉบับที่ 4 (ม.4-6)						1.0<b<1.5 5 ข้อ	1.5>b 40 ข้อ

แผนภาพที่ 3 ผังกำหนดค่าประมาณความยาก (b) ข้อสอบเฉพาะช่วงชั้น และข้อสอบรวม

3. ประยุกต์เทคนิคการเทียบมาตราคะแนนแนวดิ่ง (Vertical Equating Technique)

กลุ่มตัวอย่างคือ นักเรียนชั้นประถมศึกษาปีที่ 1 ถึง มัธยมศึกษาปีที่ 6 จากกรุงเทพมหานคร และจังหวัด ขอนแก่นมีลักษณะคละความสามารถตามสภาพการ กระจายปกติ จำนวนทั้งสิ้น 1,248 คน ข้อมูลคะแนนจาก การสอบแต่ละฉบับในแต่ละทักษะย่อนำมาวิเคราะห์ด้วย

โปรแกรม BILOG-MG ITEM MAINTENANCE PROGRAM: LOGISTIC ITEM RESPONSE MODEL (SCIENTIFIC SOFTWARE INTERNATIONAL, INC. 2002) โดยกำหนดให้ กลุ่มผู้สอบในแบบสอบฉบับที่ 3 (ช่วงชั้นมัธยมศึกษาปีที่ 1-3) เป็นกลุ่มอ้างอิง ทั้งนี้ได้พิจารณาจากผลการวิเคราะห์มา พิจารณาความเหมาะสมของข้อสอบจากค่าความยาก (b) ค่าอำนาจจำแนก (a) และคุณลักษณะของแบบสอบ (Test

information curve: TFC) ที่อธิบายช่วงความสามารถในการวัดแต่ละแบบสอบย่อย การวิเคราะห์ดำเนินการซ้ำหลายรอบ ต้องตัดข้อสอบบางข้อ หรือข้อมูลผู้ตอบบางคนออกตามเงื่อนไขของทฤษฎีการตอบสนองข้อสอบ (Item response theory) สุดท้ายได้ชุดแบบสอบทักษะพื้นฐานทางวิชาการ

KBAST ประกอบด้วยแบบสอบ 4 ฉบับ แต่ละฉบับมีแบบสอบย่อย 4 ทักษะ สรุปจำนวนข้อในชุดแบบสอบทักษะพื้นฐานทางวิชาการ และช่วงค่าความยาก (b) ดังแสดงในตารางที่ 3 ดังนี้

ตารางที่ 3 ผลการวิเคราะห์ข้อมูลเชิงประจักษ์ จำนวนข้อ ค่าความยาก (b) สูงสุดและต่ำสุด ของแบบสอบ

แบบสอบย่อย ชุดแบบสอบ	การอ่านคำ N ($b_{\max}-b_{\min}$)	การสะกตคำ 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)

4. กำหนดระดับเส้นฐานความสามารถ (Basal level)

เมื่อผู้สอบทำข้อสอบในแบบสอบย่อยฉบับใดๆ ผิดติดต่อกัน 5 ข้อ ให้ถือเป็นข้อยุติการสอบแบบสอบนั้น การนับคะแนนการทำข้อสอบถูกต้อง จะนับจากจำนวนข้อที่อยู่เหนือระดับเส้นฐานความสามารถ ขณะเดียวกัน ข้อสอบที่อยู่ใต้เส้นฐานความสามารถ เรียกว่าระดับเพดาน (Ceiling Level) เป็นข้อสอบที่อยู่ระดับสูงเกินความสามารถของผู้สอบนั้นๆ ผู้สอบที่ทำข้อสอบถูกต้องในระดับเพดาน จะไม่ได้คะแนน

ข้อกำหนดลักษณะเฉพาะของแบบสอบ (test specification) และตัวข้อสอบ (test items) ในแต่ละทักษะนั้น ได้ดำเนินการอย่างรอบคอบในขั้นการพัฒนาแบบสอบมาตรฐาน สรุปได้แบบสอบ 4 ฉบับที่วัดทักษะย่อย การอ่านคำ การสะกตคำ ความเข้าใจประโยค และการคำนวณทางคณิตศาสตร์

2. ความตรงเชิงโครงสร้างตามรูปแบบความเจริญตามระดับชั้น

ผลการวิเคราะห์ค่าความสามารถ (θ) ของกลุ่มตัวอย่าง แสดงตามตาราง 4

ความตรงของชุดแบบสอบมาตรฐาน KBAST

1. ความตรงเชิงเหตุผลของแบบวัด

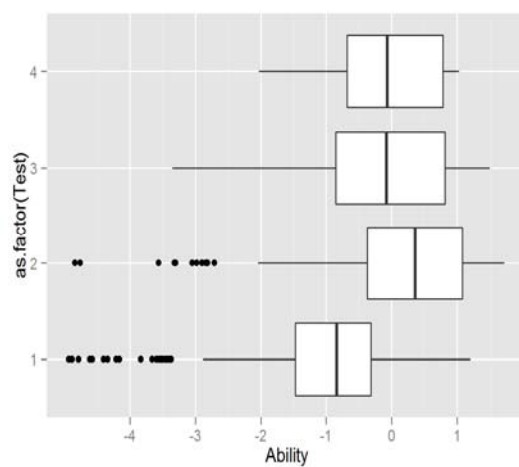
ผลการพิจารณาคูณลักษณะของแบบวัดแต่ละฉบับ โดยผู้เชี่ยวชาญร่วมกับนักวิจัยต่อความสอดคล้องของ

ตารางที่ 4 ค่าประมาณความสามารถเฉลี่ย (θ) ในแบบสอบย่อย 4 ทักษะ

แบบสอบย่อย แบบสอบช่วงชั้น	การอ่านคำ		การสะกตคำ		ความเข้าใจประโยค		การคำนวณทางคณิตศาสตร์	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
ฉบับที่ 1	-0.89	1.39	-0.87	1.39	-1.03	1.88	-2.51	1.91
ฉบับที่ 2	-0.27	1.20	-0.15	1.28	-0.16	1.39	-0.60	1.14
ฉบับที่ 3	0	1.00	0.00	1.00	-0.01	1.06	0.00	1.00
ฉบับที่ 4	-0.01	0.81	0.43	0.75	0.27	0.94	0.38	1.05

นำเสนอแผนภาพคะแนนความสามารถเฉลี่ย (θ) ตามช่วงชั้น และทดสอบความมีนัยสำคัญทางสถิติเสนอด้วยค่าสถิติทดสอบความแตกต่างของค่าความสามารถ

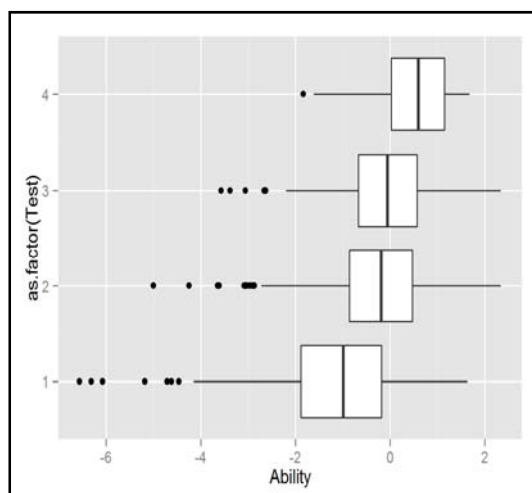
เฉลี่ยของแบบสอบย่อยทั้ง 4 ฉบับ ด้วย the Kruskal-Wallis test กำหนดระดับความมีนัยสำคัญที่ .10 ดังแสดงด้วยแผนภาพที่ 4 - 7 และตารางที่ 5 - 8 ตามลำดับดังนี้



แผนภาพที่ 4 ค่า θ เฉลี่ยแบบสอบการอ่านคำ

ตารางที่ 5 ผลทดสอบความแตกต่างทางสถิติ แบบสอบการอ่านคำ

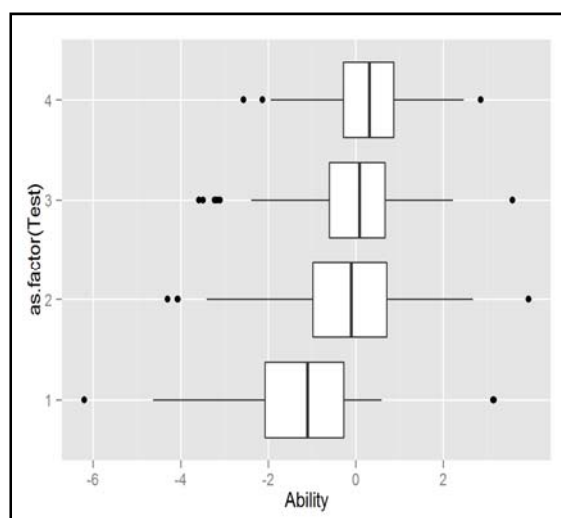
แบบสอบที่เปรียบเทียบ	Median	d.f.	P value
แบบสอบฉบับที่ 1-4	-0.15	3	5.19E-31*
ฉบับที่ 1 กับฉบับที่ 2	-0.84, 0.35	1	3.42E-26*
ฉบับที่ 1 กับฉบับที่ 3	-0.84, -0.01	1	8.37E-17*
ฉบับที่ 1 กับฉบับที่ 4	-0.84, -0.07	1	1.12E-16*
ฉบับที่ 2 กับฉบับที่ 3	0.35, -0.01	1	3.59E-6*
ฉบับที่ 2 กับฉบับที่ 4	0.35, -0.07	1	0.00001*
ฉบับที่ 3 กับฉบับที่ 4	-0.01, -0.07	1	0.914



แผนภาพที่ 5 ค่า θ เฉลี่ยแบบสอบการสะกดคำ

ตารางที่ 6 ผลทดสอบความแตกต่างทางสถิติ แบบสอบการสะกดคำ

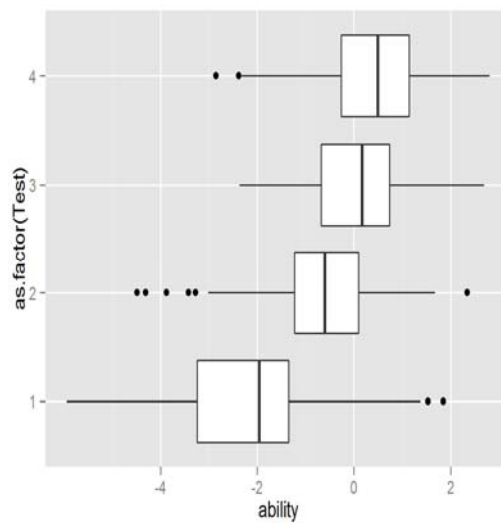
แบบสอบที่เปรียบเทียบ	Median	d.f.	P value
แบบสอบฉบับที่ 1-4	-0.08	3	6.04E-48*
ฉบับที่ 1 กับฉบับที่ 2	-0.85, -0.16	1	1.04E-11*
ฉบับที่ 1 กับฉบับที่ 3	-0.85, -0.08	1	2.78E-19*
ฉบับที่ 1 กับฉบับที่ 4	-0.85, 0.55	1	5.75E-42*
ฉบับที่ 2 กับฉบับที่ 3	-0.16, -0.08	1	0.071*
ฉบับที่ 2 กับฉบับที่ 4	-0.16, 0.55	1	3.46E-18*
ฉบับที่ 3 กับฉบับที่ 4	-0.08, 0.55	1	1.29E-14*



แผนภาพที่ 6 ค่า Θ เฉลี่ยแบบสอบความเข้าใจประโยค

ตารางที่ 7 ผลทดสอบความแตกต่างทางสถิติแบบสอบความเข้าใจประโยค

แบบสอบที่เปรียบเทียบ	Median	d.f.	P value
แบบสอบฉบับที่ 1-4	-0.10	3	1.56E-31*
ฉบับที่ 1 กับฉบับที่ 2	-1.1, -0.1	1	4.33E-13*
ฉบับที่ 1 กับฉบับที่ 3	-1.1, 0.09	1	1.29E-20*
ฉบับที่ 1 กับฉบับที่ 4	-1.1, 0.31	1	1.57E-30*
ฉบับที่ 2 กับฉบับที่ 3	-0.1, 0.09	1	0.096*
ฉบับที่ 2 กับฉบับที่ 4	-0.1, 0.31	1	0.000038*
ฉบับที่ 3 กับฉบับที่ 4	0.09, 0.31	1	0.004501*



แผนภาพที่ 7 ค่า Θ เฉลี่ยแบบสอบการคำนวณทางคณิตศาสตร์

ตารางที่ 8 ผลทดสอบความแตกต่างทางสถิติ แบบสอบการคำนวณทางคณิตศาสตร์

แบบสอบที่เปรียบเทียบ	Median	d.f	P value
แบบสอบฉบับที่ 1-4	-0.10	3	1.66E-109*
ฉบับที่ 1 กับฉบับที่ 2	-1.1, -0.1	1	5.76E-47*
ฉบับที่ 1 กับฉบับที่ 3	-1.1, 0.09	1	2.26E-66*
ฉบับที่ 1 กับฉบับที่ 4	-1.1, 0.31	1	1.14E-74*
ฉบับที่ 2 กับฉบับที่ 3	-0.1, 0.09	1	1.16E-13*
ฉบับที่ 2 กับฉบับที่ 4	-0.1, 0.31	1	8.31E-28*
ฉบับที่ 3 กับฉบับที่ 4	0.09, 0.31	1	3.45E-6*

ผลการทดสอบความตรงเชิงโครงสร้างตามรูปแบบความเจริญตามระดับชั้นของแบบสอบย่อยรายทักษะ ดังเสนอในแผนภาพที่ 4-7 และตารางที่ 5-8 สรุปผลโดยรวมพบว่าแบบสอบย่อยการอ่านคำ ฉบับที่ 3 และ 4 มีค่า Θ เฉลี่ยไม่สอดคล้องกับรูปแบบการเจริญตามช่วงชั้นและไม่พบความแตกต่างอย่างมีนัยสำคัญ ($P \text{ value} > .10$) ส่วนแบบสอบย่อยการสะกดคำ ความเข้าใจประโยค และการคำนวณทางคณิตศาสตร์มีค่า Θ เฉลี่ยสอดคล้องกับรูปแบบการเจริญตาม

ช่วงชั้น และมีความแตกต่างระหว่างแบบสอบฉบับที่ 1 - 4 ทุกคู่จากการทดสอบภายหลัง ($P \text{ value} < .10$)

ความเที่ยงของชุดแบบสอบมาตรฐาน KBAST

ความเที่ยงของชุดแบบสอบ KBAST วิเคราะห์ด้วยสัมประสิทธิ์ความคงที่ภายใน Cronbach's coefficient alpha (1951) ผลสรุปสัมประสิทธิ์ความเที่ยงและความคลาดเคลื่อนของการวัดของแบบสอบย่อยการอ่านคำ การสะกดคำ ความเข้าใจประโยค และการคำนวณทางคณิตศาสตร์ ของฉบับที่ 1 - ฉบับที่ 4 ดังนี้

ตารางที่ 9 ช่วงค่าสัมประสิทธิ์ (r) และความคลาดเคลื่อนของการวัดของแบบสอบย่อยทั้ง 4 ฉบับ

แบบสอบย่อย	ค่าสัมประสิทธิ์ความเที่ยง (r)	ค่าคลาดเคลื่อน (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

สรุปค่าความเที่ยงของชุดแบบสอบ KBAST โดยรวมอยู่ในระดับค่อนข้างสูง ($r > 0.80$) แบบสอบการคำนวณทางคณิตศาสตร์ ฉบับที่ 4 มีค่าสัมประสิทธิ์ความเที่ยงสูงสุด เท่ากับ 0.950 มีความคลาดเคลื่อนมาตรฐานของการวัดต่ำที่สุด คือเท่ากับ 0.2502 ส่วนค่าความเที่ยงของแบบสอบความเข้าใจประโยคอยู่ในระดับค่อนข้างต่ำ โดยเฉพาะฉบับที่ 4 มีค่าสัมประสิทธิ์ความเที่ยงเท่ากับ 0.680 และมีค่าความคลาดเคลื่อนมาตรฐานของการวัดสูงที่สุด คือเท่ากับ 0.5589

ผลการสร้างตารางเกณฑ์ปกติ (Norms)

เพื่อการแปลความหมายคะแนนของผู้สอบตามระดับความสามารถในรายทักษะที่อิงระดับชั้นเรียน และระดับช่วงชั้น ผู้วิจัยได้นำเสนอในรูปคะแนนมาตรฐาน (scaled score) รายชั้นประถมศึกษาปีที่ 1 ถึง มัธยมศึกษาปีที่ 6 มีค่าเฉลี่ยเท่ากับ 100 และค่าเบี่ยงเบนมาตรฐานเท่ากับ 15 เช่นเดียวกับแบบสอบมาตรฐานทั่วไป ด้วยวิธีการ Linear transformation จัดทำตารางเทียบคะแนนดิบสู่คะแนนมาตรฐาน เป็นรายฉบับของแต่ละแบบสอบย่อย เมื่อผู้สอบทำแบบสอบฉบับใด จะเทียบคะแนนดิบที่ได้เป็นคะแนนมาตรฐานของระดับชั้นและช่วงชั้นนั้นๆ ตามตารางในคู่มือการสอบ คะแนนมาตรฐานที่อ่านได้จากตารางเทียบของระดับชั้นจะให้ความหมายว่า ผู้สอบมีความสามารถในการทักษะที่สอบอยู่ในระดับชั้นใดของนักเรียนในระดับชั้นเดียวกัน ขณะเดียวกัน ถ้าอ่านตารางเทียบของช่วงชั้นก็แปลความหมายว่า ผู้สอบมีความสามารถในด้านที่สอบอยู่ในช่วงชั้นใดของนักเรียนในช่วงชั้นเดียวกัน

อภิปรายและข้อเสนอแนะ

การพัฒนาชุดแบบสอบ KBAST นี้เป็นขั้นต้นของการพัฒนาแบบสอบมาตรฐาน ที่ใช้เวลาเพียง 1 ปี และประชากรนักเรียน 2 จังหวัด มีขอบเขตการวัดทักษะของนักเรียนระดับชั้นประถมปีที่ 1 ถึง มัธยมศึกษาปีที่ 6 พบจุดเด่นและจุดที่ควรพัฒนาหลายประการ ดังนี้

จุดเด่น การประยุกต์เทคนิคการเทียบมาตรฐานคะแนนแนวดิ่งของแบบสอบทักษะพื้นฐานทางวิชาการ KBAST นี้ ได้ชุดแบบสอบ 4 ฉบับที่มีความยากตามช่วงชั้น ทำให้สามารถแปลความหมายความสามารถทักษะของผู้สอบว่าอยู่ในระดับใด เช่น นายเด่น เป็นนักเรียนพิเศษอยู่ในห้องเรียนรวมชั้นมัธยมศึกษาปีที่ 1 จากการสังเกตของครู นายเด่นว่ามีความสามารถในการคิดคำนวณทางคณิตศาสตร์อ่อนมาก ครูจึงให้นายเด่นทำแบบสอบทักษะการคำนวณทางคณิตศาสตร์ ฉบับที่ 2 (ช่วงชั้นประถมศึกษาปีที่ 4-6) พบว่า นายเด่นทำคะแนนเทียบได้กับระดับชั้นประถมศึกษาปีที่ 4 จากตัวอย่างที่ปรากฏเป็น

ประโยชน์แก่ครูในการวินิจฉัยทักษะทางวิชาการของนักเรียนที่มีความต้องการจำเป็นพิเศษ เพื่อจัดหลักสูตรเฉพาะบุคคลที่มีความสอดคล้องกับสภาพทักษะรายบุคคล

จุดที่ควรพัฒนา คุณภาพของแบบสอบมาตรฐานที่วัดทักษะพื้นฐานทางวิชาการ 4 ทักษะ 12 ระดับชั้น จำเป็นต้องดำเนินการอย่างต่อเนื่องตามกระบวนการพัฒนาแบบสอบมาตรฐาน ดังนี้

(1) พัฒนาเนื้อหาของแบบสอบ ให้ครอบคลุมสาระที่เหมาะสมกับบริบททางสังคม วัฒนธรรม ทั้งนี้ เพราะทักษะพื้นฐานทางวิชาการ ต้องสัมพันธ์กับการเรียนรู้เพื่อการสื่อสารในชีวิตประจำวัน

(2) พัฒนา norms ที่มาจากประชากรทั่วประเทศ เพื่อให้การแปลความหมายได้มาตรฐานซึ่งจะเป็นประโยชน์ในการนำไปใช้ ทั้งนี้ต้องใช้ข้อมูลเชิงประจักษ์ที่ได้จากกลุ่มตัวอย่างของประชากรของประเทศ ในช่วงเวลาเดียวกัน เพื่อให้ได้มาตรฐานความตรงอ้างอิงตามเวลา

(3) ทำโครงการวิจัยติดตามผู้ใช้แบบสอบ KBAST เพื่อพัฒนาอย่างต่อเนื่องให้มีความเป็นมาตรฐานที่อ้างอิงได้เป็นปัจจุบัน และศึกษาสภาพความมีทักษะพื้นฐานทางวิชาการ ของกลุ่มบุคคลที่มีความต้องการจำเป็นพิเศษ

References

- American Educational Research Association (AERA), American Psychological Association (APA), & National Council on Measurement in Education (NCME). (1999). *Standards for Educational and Psychological Testing*. Washington, DC: AERA.
- Jastak, J. (1946). *Wide Range Achievement Test*. Guidance Associates of Delaware, Inc. Wilmington, DE.
- Jastak, J. & Jastak, S. (1965, 1976 & 1978). *Wide Range Achievement Test*. Jastak Associates, Inc. Wilmington, DE.
- Kolen, M. J. & Brennan, R. L. (2004). *Test equating, scaling, and linking* (2nd ed.) New York: Springer.
- Zimowski, M. F., Muraki, E., Mislevy, R. J. & Darrell, B. R. (2002) *BILOG-MG for Windows*. Scientific Software International, Inc.