การพัฒนาวิธีการกำหนดคะแนนจุดตัดเพื่อตัดสินความรอบรู้

ประภัสสร วงษ์ดี, สำราญ มีแจ้ง, รัตนะ บัวสนธ์
และ ปกรณ์ ประจันบาน
มหาวิทยาลัยนเรศวร

การวิจัยนี้ มีวัตถุประสงค์เพื่อพัฒนาวิธีการกำหนดคะแนนจุดตัดเพื่อตัดสินความรอบรู้ และเปรียบเทียบประสิทธิภาพของวิธีการกำหนดคะแนนจุดตัดเพื่อตัดสินความรอบรู้กับวิธีการกำหนดคะแนนจุดตัดของแกลสโดยอาศัยทฤษฎีการตัดสินใจ โดยการเปรียบเทียบค่าสัมประสิทธิ์ความสอดคล้องในการตัดสินความรอบรู้ ด้วยการทดสอบค่าสัมประสิทธิ์สหสัมพันธ์ฟี (r_ϕ) แล้วทำการทดสอบค่าสถิติ Z หลังจากแปลงค่าสัมประสิทธิ์สหสัมพันธ์ฟี(r_ϕ) ให้อยู่ในรูปของคะแนน ฟิชเชอร์ซี (Fisher Z - transformation) ข้อมูลที่ใช้ในการกำหนดคะแนนจุดตัดครั้งนี้เป็นข้อมูลทุติยภูมิของนักศึกษาที่เข้าศึกษาในสถาบันอุดมศึกษา ปีการศึกษา 2552 ประกอบด้วย เกรดเฉลี่ยรวม กลุ่มสาระคณิตศาสตร์และภาษาต่างประเทศ ในระดับมัธยมศึกษาตอนปลาย คะแนนสอบ O-NET วิชาคณิตศาสตร์และภาษาอังกฤษชั้นมัธยมศึกษาปีที่ 6 และเกรดเฉลี่ยสะสมชั้นปีที่ 1 ระดับปริญญาตรี

ผลการวิจัยสรุปได้ดังนี้

1. วิธีการกำหนดคะแนนจุดตัดที่พัฒนาขึ้น มีการบูรณาการโมเดลการตอบสนองข้อสอบ (IRT) แบบสามพารามิเตอร์ เพื่อวิเคราะห์ค่าความสามารถของคะแนนเกณฑ์ภายในจากการทำแบบสอบอิงเกณฑ์ก่อนนำมากำหนดคะแนนจุดตัดเกณฑ์ภายใน จำนวน 1 จุด เพื่อตัดสินและจำแนกผู้สอบออกเป็น 2 กลุ่มคือ รอบรู้ และ ไม่รอบรู้ โดยศึกษาร่วมกับเกณฑ์ภายนอกที่จำแนกผู้สอบออกเป็น 2 กลุ่มเช่นกัน โดยคะแนนเกณฑ์ภายในที่เป็นคะแนนจุดตัดที่เหมาะสม ต้องให้ค่าดัชนีความคลาดเคลื่อนในการจำแนกการเป็นผู้รอบรู้ (Error of Classification Index ; ECI) ต่ำที่สุด ซึ่งมีสูตรคำนวณ ECI ที่พัฒนาขึ้น ดังนี้

$$ECI = \frac{(P_A + P_D)^2}{[P_{(A+B)}P_{(C+D)}]+[P_{(A+C)}P_{(B+D)}]}$$

2. ผลการศึกษาส่วนใหญ่แสดงให้เห็นว่าวิธีการกำหนดคะแนนจุดตัดที่พัฒนาขึ้น มีสัมประสิทธิ์ความสอดคล้องในการตัดสินความรอบรู้ สูงกว่าวิธีการกำหนดคะแนนจุดตัดของแกลสโดยอาศัยทฤษฎีการตัดสินใจ และผลการศึกษาบางส่วนชี้ให้เห็นว่าวิธีการกำหนดคะแนนจุดตัดที่พัฒนาขึ้น มีสัมประสิทธิ์ความสอดคล้องในการตัดสินความรอบรู้ เท่ากับวิธีการกำหนดคะแนนจุดตัดของแกลสโดยอาศัยทฤษฎีการตัดสินใจ

คำสำคัญ: วิธีการกำหนดคะแนนจุดตัด, ทฤษฎีการตอบสนองข้อสอบ

มหาวิทยาลัยบูรพา Burapha University

A Proposed of Cut-Off Score Method for Mastery Judgment

Prapassorn Wongdee, Samran Mejang, Rattana Buosonte and Pakorn Prachanban

Naresuan University, Thailand

Abstract

The purposes of this study were to propose the cut–off score method for mastery judgment and to compare the efficiency of cut–off score method between the cut–off score method for mastery judgment and the cut–off score of Glass's Decision Method. The efficiency of cut-off score method considered from the consistency or validity of mastery learners by using Phi coefficiency and transformed to Fisher Z, then compared with Z–test respectively. The set of secondary data of students that were a freshman in academic year 2009; O-NET scores data at grade 12 in Mathematics and English, grade point average of high school education in Mathematics and Foreign Language, cumulative grade point average at the end of first year in Bachelor's Degree.

The research results consisted of :

1. The cut–off score method for mastery judgment integrated the Item Response Theory to analyze the ability score of a criterion–referenced test and then set a cut–off score. The cut–off score categorized persons into mastery and non-mastery and judged with the external criterion that also devided persons into mastery and non-mastery. The suitable cut–off score got the minimize Error of Classification Index (ECI) that the researcher formulated as follows:

$$ECI = \frac{(P_A + P_D)^2}{[P_{(A+B)}P_{(C+D)}]+[P_{(A+C)}P_{(B+D)}]}$$

2. Most of the research results indicated that the cut–off score method for mastery judgment had higher consistency coefficient than the cut–off score of Glass's Decision Method. Some of the research results indicated that the cut–off score method for mastery judgment and the cut–off score of Glass's Decision Method had the equal consistency coefficient.

.

***Keywords*:** cut-off score, Item Response Theory

ความนำ

การจัดการเรียนรู้เป็นหนึ่งในวิธีการพัฒนาทรัพยากรมนุษย์ให้มีคุณภาพ ซึ่งในกระบวนการจัดการเรียนรู้ต้องมีการวัดและประเมินผลการจัดการเรียนรู้ว่าผู้เรียนได้รับการพัฒนาตามที่คาดหวังไว้หรือไม่ จากผลการวัดคุณภาพนักเรียนไทย ยังพบว่า มีความขัดแย้งกันในผลการประเมินการเรียนรู้ในหลายส่วน ดังเห็นได้จากผลการวิจัยที่พบว่าผลสัมฤทธิ์ของเด็กไทยประมาณร้อยละ 70 ของประเทศ ได้เกรดเฉลี่ย 2.51 ขึ้นไป หรืออยู่ในเกณฑ์ค่อนข้างดีถึงดีมาก (สำนักงานเลขาธิการสภาการศึกษา, 2553, หน้า 38) น่าจะเป็นตัวบ่งชี้ว่าเด็กไทยส่วนใหญ่เป็นผู้ที่มีความรอบรู้ทั้งนี้เพราะเกณฑ์มาตรฐานการเรียนรู้ขั้นต่ำของสถาบันการศึกษาในระดับอุดมศึกษาทั้งระดับปริญญาตรี และระดับปริญญาโทของสถาบันการศึกษาทั้งในและต่างประเทศนั้น ส่วนใหญ่ได้กำหนดเกรดเฉลี่ยขั้นต่ำเพื่อรับเข้าทำการคัดเลือกคือ 2.50 ขึ้นไป และ 2.75 ขึ้นไปสำหรับบางสถาบัน (รวมมิตร การศึกษาต่อระดับปริญญาตรี, ออนไลน์; Liverpool International College, Online; สำนักงาน ก.พ., ออนไลน์; ศูนย์แนะแนวศึกษาต่อต่างประเทศ เอ็นริช เอ็ดดูเคชั่น, ออนไลน์) แต่ผลการศึกษาความรู้ความสามารถของเด็กไทยเมื่อเทียบกับผลการวัดความรู้ระดับโรงเรียนในระดับชาติและระดับนานาชาติแล้วยังไม่สอดคล้องกัน ดังที่พบว่า ผลสัมฤทธิ์ทางการเรียน (General Achievement Test: GAT) ของผู้เรียนส่วนใหญ่ของประเทศในช่วงชั้นที่ 2 ถึง 4 ระหว่างปีการศึกษา 2544 - 2549 มีแนวโน้มค่าเฉลี่ยผลสัมฤทธิ์ทางการเรียนค่อนข้างลดลง และค่าเฉลี่ยร้อยละของวิชาต่าง ๆ โดยส่วนใหญ่ต่ำกว่าร้อยละ 50 (สำนักงานเลขาธิการสภาการศึกษา, 2551) และจากการวัดผลสัมฤทธิ์ในระดับนานาชาติ (ผลการประเมิน PISA) ด้านการอ่าน คณิตศาสตร์และวิทยาศาสตร์ ยังพบว่าเด็กไทยได้คะแนนเฉลี่ยทั้ง 3 ด้าน ต่ำกว่าค่าเฉลี่ยนานาชาติมาตั้งแต่ปี ค.ศ. 2003 - ค.ศ. 2009 (สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี, มปท.) รวมทั้งผลการประเมินร่วมกับนานาชาติ หรือ TIMSS 2007 นักเรียนชั้นมัธยมศึกษาปีที่ 2 ปี พ.ศ. 2550 ที่พบว่าผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ และวิชาวิทยาศาสตร์ต่ำกว่าค่าเฉลี่ยนานาชาติ ทั้ง 2 วิชา (แนวหน้า, ออนไลน์) ซึ่งถือว่าเป็นความไม่สอดคล้องกันระหว่างผลการประเมินในระดับชั้นเรียน ระดับชาติ และระดับนานาชาติ รวมทั้งยังเป็นผลกระทบต่อเนื่องให้เกิดปัญหาในการจัดการเรียนการสอนในระดับการศึกษาที่สูงขึ้นอีกด้วย ดังผลการศึกษาสภาพและปัญหาการจัดการเรียนการสอนของคณาจารย์มหาวิทยาลัยบูรพาที่พบว่า ผู้เรียนมีพื้นฐานทางการเรียนไม่ดี ถือเป็นส่วนหนึ่งของปัญหาที่พบมากที่สุดที่เกิดจากตัวนิสิต และเป็นปัญหาต่อการจัดการเรียนการสอน (วิชิต สุรัตน์เรืองชัย และคณะ, 2549) แสดงว่าระบบการวัดและประเมินผลที่จะนำไปสู่การสรุปว่า ผู้เรียนได้รับการพัฒนาและมีความรอบรู้จริงหรือไม่นั้น ควรได้รับการพิจารณาว่ามีส่วนใดต้องได้รับการแก้ไข เพื่อให้สามารถประเมินความรอบรู้ของผู้เรียนได้ถูกต้อง และสอดคล้องต้องกัน

แนวทางในการจัดการเรียนรู้ เพื่อนำไปสู่การพัฒนาคุณภาพของผู้เรียนมีหลายรูปแบบ รูปแบบหนึ่งที่ได้รับความนิยมคือ **การเรียนแบบรอบรู้** (พงษ์ธารา วิจิตรเวชไพศาล, 2551) สอดคล้องกับผลการวิจัยที่พบว่าการจัดการเรียนการสอนแบบรอบรู้มีผลดีต่อผู้เรียน โดยช่วยเพิ่มความสามารถในการเรียนรู้ให้กับผู้เรียนได้ (ประยุทธ นิสภกุล, 2550; ปฐมพงษ์ จันทร์สว่าง, 2549; วรสริน พันธุ, 2550; Armacost & Armacost, 2003) ซึ่งเหมาะสมกับการประเมินผลแบบอิงเกณฑ์ ที่มีการกำหนดคะแนนจุดตัด (cut-off score) หากคะแนนจุดตัดมีความเหมาะสมจะสามารถจำแนกผู้เรียนออกเป็นผู้รอบรู้และผู้ไม่รอบรู้ (Master and Non-Mastery) ออกจากกันได้อย่างถูกต้อง (สุพัฒน์ สุกมลสันต์, 2530) การกำหนดคะแนนจุดตัดเป็นประเด็นหนึ่งของกระบวนการพัฒนาแบบสอบที่มีความท้าทาย และเป็นประเด็นที่ยังถกเถียงกันมาก ทั้งด้านระเบียบวิธี ผลกระทบของคะแนนจุดตัด และการใช้คะแนนจุดตัดเป็นประโยชน์สำหรับผู้กำหนดนโยบายทางการศึกษา (Davis-Becker & Buckeb, 2011) การกำหนดคะแนนจุดตัดที่ได้มาตรฐานควรคำนึงถึงความเข้าใจและความไว้วางใจของสาธารณะชนด้วย เพราะกว่าจะตัดสินความรอบรู้ของผู้สอบได้ ต้องใช้เวลา

งบประมาณ และความพยายามเป็นอย่างมาก ผ่านกระบวนการจัดการเรียนรู้ จึงควรมุ่งพัฒนาคะแนนจุดตัดให้บ่งบอกความสามารถได้อย่างแท้จริง (Cizek & Bunch, 2007)

วิธีกำหนดคะแนนจุดตัดเริ่มพัฒนามาตั้งแต่ปี ค.ศ.1970 และในยุคต่อมา ยังพบข้อด้อยในวิธีการกำหนดคะแนนจุดตัด (Linn, 2000 cited in Kane, 2002) ทั้งนี้เพราะนักวิจัยส่วนใหญ่มุ่งศึกษาเกี่ยวกับการประมาณค่าความเที่ยง (Reliability) ของแบบสอบ เมื่อกำหนดคะแนนจุดตัดด้วยวิธีการที่แตกต่างกันไป หรือเมื่อใช้วิธีกำหนดคะแนนจุดตัดแบบเดียวกันแต่ใช้แบบสอบต่างประเภทกัน มีงานวิจัยส่วนน้อยเท่านั้นที่สนใจเกี่ยวกับการพัฒนาวิธีกำหนดคะแนนจุดตัดที่เหมาะสม เช่น Sireci, Hambleton, and Pitoniak (2004) ได้พัฒนาวิธีการกำหนดคะแนนจุดตัดขึ้นมาที่ถือว่าใหม่ล่าสุดชื่อว่า วิธีการตรวจสอบความสอดคล้อง (The Direct Consensus Method) เพื่อแก้ปัญหาเรื่องการใช้เวลามากของวิธีการกำหนดคะแนนจุดตัดตามวิธีการของแองกอฟ (Angoff) โดยวิธีใหม่นี้ให้ผู้เชี่ยวชาญตัดสินคะแนนจุดตัดของข้อสอบเป็นด้าน ๆ (ไม่ต้องตัดสินรายข้อ) และทำให้ใช้เวลาน้อยลง (Cizek & Bunch, 2007) จะเห็นว่าวิธีการกำหนดคะแนนจุดตัดส่วนใหญ่รวมทั้งวิธีล่าสุดยังเป็นเทคนิควิธีการที่ต้องอาศัยการตัดสินใจจากผู้เชี่ยวชาญ ซึ่งถือว่ามีความไม่เที่ยงในการตัดสินแต่ละครั้ง (MacCan & Stanley, 2006 cited in MacCan & Stanley, 2010) ผู้วิจัยเห็นว่าหากคะแนนจุดตัดของการวัดผลแต่ละครั้งมีความตรงหรือถูกต้อง ก็จะนำไปสู่การให้ข้อมูลสะท้อนกลับให้แก่ผู้เรียนได้รู้ผลสัมฤทธิ์ที่แท้จริงได้ และได้ผลการประเมินที่สอดคล้องต้องกันจากการตัดสินหลายเกณฑ์ ผู้วิจัยจึงสนใจพัฒนาแนวทางการกำหนดคะแนนจุดตัดเพื่อตัดสินความรอบรู้ โดยมุ่งเปรียบเทียบประสิทธิภาพกับวิธีการกำหนดคะแนนจุดตัดของแกลสโดยอาศัยทฤษฎีการตัดสินใจ (Decision-Theory Approach of Glass's Technique)

วัตถุประสงค์ของการวิจัย

1. เพื่อพัฒนาวิธีการกำหนดคะแนนจุดตัดเพื่อตัดสินความรอบรู้
2. เพื่อเปรียบเทียบประสิทธิภาพของวิธีการกำหนดคะแนนจุดตัดที่พัฒนาขึ้น กับวิธีการของแกลสโดยอาศัยทฤษฎีการตัดสินใจ

ขอบเขตการวิจัย

1. การวิจัยนี้เป็นการศึกษาและพัฒนาวิธีการกำหนดคะแนนจุดตัด แล้วทำการเปรียบเทียบกับวิธีการกำหนดคะแนนจุดตัดของแกลสโดยอาศัยทฤษฎีการตัดสินใจ เนื่องจากมีการเปรียบเทียบความถูกต้องกับเกณฑ์ภายนอก โดยอาศัยข้อมูลเชิงประจักษ์และมีหลักการทางสถิติตรวจสอบความถูกต้องของคะแนนจุดตัดเช่นเดียวกัน

2. ข้อมูลที่ใช้ในการศึกษาเพื่อกำหนดคะแนนจุดตัดครั้งนี้เป็นข้อมูลทุติยภูมิ ได้แก่ ข้อมูลผลการเรียนของนักศึกษาที่เข้าศึกษาต่อระดับปริญญาตรีในสถาบันอุดมศึกษา ปีการศึกษา 2552 ข้อมูลกลุ่มตัวอย่างแต่ละคนประกอบด้วย

2.1 ข้อมูลที่ใช้เป็นเกณฑ์ภายใน มีดังนี้

2.1.1 คะแนนสอบ O-NET รายข้อ วิชาคณิตศาสตร์ และวิชาภาษาอังกฤษ ชั้นมัธยมศึกษาปีที่ 6 เพื่อคำนวณค่าความสามารถ ใช้กับวิธีการกำหนดคะแนนจุดตัดที่พัฒนาขึ้น

2.1.2 ร้อยละคะแนนรวมจากการสอบ O-NET วิชาคณิตศาสตร์ และวิชาภาษาอังกฤษ ชั้นมัธยมศึกษาปีที่ 6 ใช้กับวิธีการกำหนดคะแนนจุดตัดของแกลส

2.2 ข้อมูลที่ใช้เป็นเกณฑ์ภายนอก: เกณฑ์ภายนอกที่ใช้กับทั้ง 2 วิธีครั้งนี้มี 2 เกณฑ์ เพื่อศึกษาความคงที่ของคะแนนจุดตัดเมื่อเกณฑ์ภายนอกเปลี่ยนแปลงไป ประกอบด้วย 1) เกรดเฉลี่ยสะสมชั้นปีที่ 1 ระดับปริญญาตรี (GPAY1) และ 2) เกรดเฉลี่ยรวมกลุ่มสาระระดับมัธยมศึกษาตอนปลาย [กลุ่มสาระคณิตศาสตร์ (GPAM) และกลุ่มสาระภาษาต่างประเทศ (GPAE)]

มหาวิทยาลัยบูรพา Burapha University

2.3 กลุ่มเป้าหมายที่ใช้ศึกษาคะแนนจุดตัดและศึกษาประสิทธิภาพของวิธีการกำหนดคะแนนจุดตัดเพื่อศึกษาความรอบรู้จากการสอบ O-NET ในรายวิชาคณิตศาสตร์ มีจำนวน 12,497 คน และในรายวิชาภาษาอังกฤษ มีจำนวน 12,521 คน

3. ตัวแปรที่ใช้ในการตรวจสอบประสิทธิภาพของวิธีการกำหนดคะแนนจุดตัด ประกอบด้วย

3.1 ตัวแปรอิสระ ได้แก่ วิธีการกำหนดคะแนนจุดตัด ดังนี้

3.1.1 วิธีการกำหนดคะแนนจุดตัดของแกลสโดยอาศัยทฤษฎีการตัดสินใจ ซึ่งบางครั้งอาจเรียกสั้นๆ ว่า “วิธีการกำหนดคะแนนจุดตัดของแกลส”หรือ “วิธีของแกลส”

3.1.2 วิธีการกำหนดคะแนนจุดตัดเพื่อตัดสินความรอบรู้ ซึ่งเป็นวิธีการกำหนดคะแนนจุดตัดที่ผู้วิจัยพัฒนาขึ้น

3.2 ตัวแปรตาม คือประสิทธิภาพของวิธีการกำหนดคะแนนจุดตัด โดยการเปรียบเทียบค่าสัมประสิทธิ์ความสอดคล้องในการตัดสินความรอบรู้ ด้วยการทดสอบค่าสัมประสิทธิ์สหสัมพันธ์ฟี (r_{ϕ})

วิธีดำเนินการวิจัย

การดำเนินการวิจัยแบ่งเป็น 2 ขั้นตอน ดังนี้

ตอนที่ 1 การพัฒนาวิธีการกำหนดคะแนนจุดตัดเพื่อตัดสินความรอบรู้

การศึกษาและนำเสนอวิธีการกำหนดคะแนนจุดตัดเพื่อตัดสินความรอบรู้มีลำดับขั้นดังนี้

1. วิเคราะห์แนวคิด ทฤษฎี และงานวิจัยที่เกี่ยวข้องกับวิธีการกำหนดคะแนนจุดตัดเพื่อตัดสินความรอบรู้
2. ระบุข้อจำกัดของวิธีการกำหนดคะแนนจุดตัดที่มีอยู่แล้ว และพัฒนาวิธีการกำหนดคะแนนจุดตัดเพื่อตัดสินความรอบรู้ โดยพิจารณาจากประสิทธิภาพของวิธีการกำหนดคะแนนจุดตัด
3. ค้นคว้าหาแนวทางในการพัฒนาวิธีการกำหนดคะแนนจุดตัด
4. เสนอวิธีการกำหนดคะแนนจุดตัดเพื่อตัดสินความรอบรู้ที่พัฒนาขึ้น

ตอนที่ 2 การหาประสิทธิภาพของวิธีการกำหนดคะแนนจุดตัดระหว่างวิธีการที่พัฒนาขึ้น กับวิธีของแกลส

การเปรียบเทียบประสิทธิภาพของวิธีการกำหนดคะแนนจุดตัดที่พัฒนาขึ้น กับวิธีของแกลส พิจารณาจาก **ค่าสัมประสิทธิ์ความสอดคล้องในการตัดสินความรอบรู้** มีรายละเอียดดังนี้

2.1 กำหนดคะแนนจุดตัด

2.1.1 กำหนดคะแนนจุดตัดด้วยวิธีที่พัฒนาขึ้น

2.1.2 กำหนดคะแนนจุดตัดด้วยวิธีของแกลส ดังนี้

(1) เริ่มกำหนดค่าร้อยละของคะแนน O-NET (คะแนนเกณฑ์ภายใน) หรือค่า X เพื่อใช้เป็นคะแนนจุดตัด(X*)

(2) จัดจำนวนผู้สอบลงในแต่ละเซลล์ มีลักษณะดังนี้

		เกณฑ์ภายนอก เกรดเฉลี่ย(GPA) ไม่ผ่าน (ไม่ถึงGPA*)	ผ่าน (GPA* ขึ้นไป)
เกณฑ์ภายใน ร้อยละคะแนน O-NET (X)	ผ่าน (X* ขึ้นไป)	D	C
	ไม่ผ่าน (ต่ำกว่า X*)	B	A

(3) คำนวณค่าฟังก์ชันความคลาดเคลื่อน [f (C_X)] ซึ่งมีสูตรคำนวณ ดังนี้

$$f(C_X) = \frac{PA + PD}{PB + PC}$$

โดยที่ PA คือ สัดส่วนผู้สอบที่ไม่ผ่านเกณฑ์ภายใน แต่ผ่านเกณฑ์ภายนอก
PB คือ สัดส่วนผู้สอบ**ที่ไม่ผ่านทั้ง**เกณฑ์**ภายใน**และเกณฑ์**ภายนอก**
PC คือ สัดส่วนผู้สอบที่**ผ่านทั้ง**เกณฑ์**ภายใน**และเกณฑ์**ภายนอก**
PD คือ สัดส่วนผู้สอบผ่านเกณฑ์ภายใน แต่ไม่ผ่านเกณฑ์ภายนอก

(4) ทดลองเปลี่ยนค่าคะแนนเกณฑ์ภายใน (X) จากคะแนนเต็มหรือคะแนนสูงสุดจนถึงคะแนนต่ำสุดของกลุ่มผู้สอบ คะแนน X ที่ทำให้ได้ค่าฟังก์ชันความคลาดเคลื่อนน้อยที่สุด ถือเป็นคะแนนจุดตัด (X*) ที่เหมาะสม

2.2 เปรียบเทียบประสิทธิภาพของวิธีการกำหนดคะแนนจุดตัดของทั้งสองวิธี โดยพิจารณาจากค่าสัมประสิทธิ์สหสัมพันธ์ฟี (r_ϕ) แล้วแปลงค่าสัมประสิทธิ์สหสัมพันธ์ฟี (r_ϕ) ให้อยู่ในรูปของคะแนนฟิชเชอร์ซี (The Fisher Z - transformation) เพื่อทำการทดสอบความแตกต่างด้วยค่าสถิติ Z (Z -Test)

ผลการวิจัย

สรุปผลการวิจัยตามวัตถุประสงค์ของการวิจัยได้ดังนี้

ตอนที่ 1 ผลการพัฒนาวิธีการกำหนดคะแนนจุดตัดเพื่อตัดสินความรอบรู้

วิธีการกำหนดคะแนนจุดตัดที่พัฒนาขึ้น ได้บูรณาการกับโมเดลการตอบสนองข้อสอบ (IRT) แบบสามพารามิเตอร์ เพื่อวิเคราะห์ค่าความสามารถของคะแนนเกณฑ์ภายในก่อนนำมากำหนดคะแนนจุดตัด จำนวน 1 จุด เพื่อตัดสิน และจำแนกผู้สอบออกเป็น 2 กลุ่ม ประกอบด้วย “กลุ่มผ่าน (รอบรู้) และกลุ่มไม่ผ่าน (ไม่รอบรู้)” และศึกษาร่วมกับเกณฑ์ภายนอก (เกณฑ์อ้างอิง) ที่แบ่งผู้สอบออกเป็น 2 กลุ่มเช่นเดียวกัน โดยคะแนนเกณฑ์ภายในที่เป็นคะแนนจุดตัดที่เหมาะสมต้องให้ค่าดัชนีความคลาดเคลื่อนในการจำแนกการเป็นผู้รอบรู้ (ECI) ต่ำที่สุด รายละเอียดการกำหนดคะแนนจุดตัด มีดังนี้

(1) หาเกณฑ์ภายนอกและคะแนนจุดตัดของเกณฑ์ภายนอกที่เหมาะสม

คะแนนจุดตัดของเกณฑ์ภายในที่ได้จากการศึกษานี้ จะเชื่อถือได้หรือไม่ขึ้นอยู่กับการเลือกเกณฑ์ภายนอกพร้อมกับคะแนนจุดตัดของเกณฑ์ภายนอกที่น่าเชื่อถือ ซึ่งหลักการเลือกเกณฑ์ภายนอกที่เหมาะสมต้องมีค่าสัมประสิทธิ์ความตรงตามเกณฑ์ (ความตรงตามสภาพ หรือความตรงเชิงทำนาย) สูงกับตัวแปรที่เป็นเกณฑ์ภายในและกำหนดคะแนนจุดตัดของเกณฑ์ภายนอก โดยพิจารณาจากหลักฐานที่เป็นสากลหรือเป็นที่ยอมรับทั่วไป

(2) กำหนดจุดเริ่มต้นของคะแนนเกณฑ์ภายใน เพื่อทดลองหาคะแนนจุดตัด (θ*)

เนื่องจากค่าความสามารถ หรือ θ มีลักษณะการแจกแจงข้อมูลแบบโค้งปกติ จึงพิจารณาค่า $\theta = \bar{\theta} + 0.6(S.D._{\theta})$ ซึ่งเท่ากับ 0.6 เมื่อค่า θ แจกแจงแบบโค้งปกติ เพราะตามทฤษฎี IRT นิยมปรับค่า θ ให้มีค่าเฉลี่ยเป็น 0 และมีค่า S.D เป็น 1 (ศิริชัย กาญจนวาสี, 2555) ตรงกับระดับเกรด B (ผลการเรียนดี) ที่ถือว่าเป็นผู้รอบรู้ตามแนวทางตัดเกรดแบบอิงกลุ่ม และสอดคล้องกับ θ = 0.6 ที่เป็นคะแนนจุดตัดความรอบรู้ที่พบจากงานวิจัย (สิทธิชัย หาญสมบัติ, 2534) จึงทดลองให้ค่า θ = 0.6 เป็นจุดเริ่มต้นของเกณฑ์ภายใน

(3) จัดจำนวนผู้สอบลงในตารางการณ์จรขนาด 2X2 ณ จุดเริ่มต้นของเกณฑ์ภายใน ดังนี้

เกณฑ์ภายนอก

เกณฑ์ภายใน ค่าความสามารถจากการสอบ O-NET (θ)		เกรดเฉลี่ย (GPA) ไม่ผ่าน (ไม่ถึง GPA*)	เกรดเฉลี่ย (GPA) ผ่าน (GPA* ขึ้นไป)	
	ผ่าน (θ*ขึ้นไป)	A	B	A+B
	ไม่ผ่าน (ต่ำกว่า θ*)	C	D	C+D
		A+C	B+D	N

(4) คำนวณค่าดัชนีความคลาดเคลื่อนในการจำแนกผู้สอบ (Error of Classification Index; ECI) ที่พัฒนาขึ้น

(5) ทดลองเลื่อนค่า θ ขึ้นลงจนกว่าจะได้ค่า θ ที่เป็นจุดตัด (θ*) ที่เหมาะสม โดยให้ค่าสถิติ ECI น้อยที่สุด

แนวคิดในการพัฒนาสูตรคำนวณค่า ECI มีดังนี้

1) พิจารณาหลักการวิเคราะห์ค่าความคงเส้นคงวา หรือความสอดคล้องในการวัดหรือในการจำแนกผู้สอบออกเป็นผู้รอบรู้และไม่รอบรู้ จากเกณฑ์ 2เกณฑ์ ซึ่งเป็นแนวคิดในด้านการศึกษาความเที่ยง (Reliability) (Reynolds, Livingston, & Willson, 2010) ที่ใช้เกณฑ์ตัดสินผู้สอบออกเป็นผู้รอบรู้และไม่รอบรู้ ด้วยเกณฑ์การตัดสิน 2 เกณฑ์

จากแนวคิดพื้นฐานที่กล่าวว่า คะแนนสอบที่วัดได้หรือที่สังเกตได้ของแต่ละบุคคล (X) ประกอบด้วยคะแนน 2 ส่วน คือ คะแนนที่เป็นองค์ประกอบของความสามารถจริง (T) และคะแนนที่เป็นองค์ประกอบของความคลาดเคลื่อน (E) และแนวคิดดังกล่าว นำไปสู่การกำหนดสมการพื้นฐานของทฤษฎีของความเที่ยง (Theory of Reliability) (Kerlinger, 1973) ดังนี้

มหาวิทยาลัยบูรพา Burapha University

$$X = T + E \quad \text{.....(1)}$$

จากสมการ (1) สามารถเขียนสมการในรูปของความแปรปรวนได้ ดังนี้

$$V_X = V_T + V_E \quad \text{.....(2)}$$

โดยที่ V_X คือ ความแปรปรวนของคะแนนที่สังเกตได้ (หรือความแปรปรวนรวม)
V_T คือ ความแปรปรวนของคะแนนจริง หรือ ความแปรปรวนร่วมระหว่างคะแนนจริงกับเกณฑ์
V_E คือ ความแปรปรวนของคะแนนความคลาดเคลื่อน

Kerlinger (1973) ได้นิยามความเที่ยงว่า **"ความเที่ยงคือ อัตราส่วนระหว่างความแปรปรวนของคะแนนจริง กับ ความแปรปรวนของคะแนนที่วัดได้"** และยังสอดคล้องกับนิยามของ"**ความตรงตามเกณฑ์สัมพันธ์ (criterion related validity) ที่หมายถึง อัตราส่วนระหว่างความแปรปรวนร่วมของคะแนนที่วัดได้กับเกณฑ์ กับความแปรปรวนของคะแนนที่วัดหรือสังเกตได้"**

จากนิยามค่าความเที่ยง ที่สอดคล้องกับนิยามของความตรงตามเกณฑ์สัมพันธ์ (r_{tt}) ตามที่ Kerlinger ได้กล่าวไว้ข้างต้นทำให้เขียนสมการเพื่อคำนวณค่าความเที่ยงหรือความตรงตามเกณฑ์สัมพันธ์ ได้ดังนี้

$$r_{tt} = \frac{V_T}{V_X} \quad \text{.....(3)}$$

นอกจากนี้ Kerlinger (1973) ยังกล่าวว่า **"หากเครื่องมือวัดใดได้ค่าความเที่ยงเป็น 1 แสดงว่าเครื่องมือวัดนั้นมีค่าความเที่ยงอย่างสมบูรณ์ (Perfect reliability)"** ซึ่งสอดคล้องกับการนำ V_x หารทุกพจน์ ของสมการ (2) ทำให้ได้สมการ (4) ดังนี้

$$r_{tt} = 1 - \frac{V_E}{V_X} \quad \text{.....(4)}$$

แสดงว่าถ้าความคลาดเคลื่อนในการวัดตัวแปรตามมีน้อยมากเท่าไร ความแปรปรวนมีระบบทั้งส่วนที่แปรผันร่วมกับตัวแปรเกณฑ์ และส่วนที่เป็นองค์ประกอบเฉพาะจะมีค่าสูงขึ้น ซึ่งส่งผลให้ความเที่ยงของตัวแปรสูงขึ้นได้เช่นเดียวกันกับค่าความตรงตามเกณฑ์สัมพันธ์ ถ้าความแปรปรวนของตัวแปรตามในส่วนที่แปรผันร่วมกับเกณฑ์ที่ใช้ตรวจสอบความตรงมีมากเท่าไร ค่าความตรงตามเกณฑ์สัมพันธ์ของตัวแปรก็มีค่าสูงขึ้นเช่นเดียวกัน

จากสมการ (4) แสดงว่า $\frac{V_E}{V_X}$ เป็นดัชนีที่บ่งบอกความคลาดเคลื่อนในการคำนวณค่าความเที่ยงหรือค่าความสอดคล้องของการตัดสินจาก 2 เกณฑ์ ผู้วิจัยจึงสนใจนำค่า $\frac{V_E}{V_X}$ มาพัฒนาเป็นสูตรคำนวณดัชนีความคลาดเคลื่อนของการจำแนกผู้สอบ (Error of Classification Index ; ECI) ดังนี้

$$ECI = \frac{V_E}{V_X} \quad \text{.....(5)}$$

เนื่องจากสมการสำหรับคำนวณค่า r_{tt} จากคะแนนดิบ ตามแนวทางการคำนวณแบบ Pearson product-moment มีรายละเอียด (Glass & Hopkins, 1984) ดังนี้

$$r_{tt} = \frac{\sum X_i Y_i - n\overline{X}\,\overline{Y}}{\sqrt{(\sum X_i^2 - n\overline{X}^2)(\sum Y_i^2 - n\overline{Y}^2)}} \quad \text{.....(6)}$$

โดยที่

r_{tt} คือ สัมประสิทธิ์ความตรงตามเกณฑ์สัมพันธ์ หรือสัมประสิทธิ์ความเที่ยง

X คือ ข้อมูล/คะแนนที่เป็นผลจากการวัดตัวแปร X

Y คือ ข้อมูล/คะแนนที่เป็นผลจากการวัดตัวแปร Y

n คือ จำนวนกลุ่มตัวอย่างที่ให้ค่า X และ Y ครบทั้ง 2 ค่า

2) สมการ (6) เป็นสมการที่ใช้สำหรับคำนวณความสัมพันธ์หรือค่าความเที่ยงหรือความสอดคล้องในการตัดสินจากตัวแปร 2 ตัวแปร ที่เป็นตัวแปรต่อเนื่อง แต่หากใช้สมการ (6) คำนวณค่าความเที่ยงตามแนวทางการวัดผลแบบอิงเกณฑ์ที่ตัวแปรมีลักษณะไม่ใช่ตัวแปรต่อเนื่อง หรือเป็นตัวแปรจัดกลุ่มที่แบ่งเป็น 2 กลุ่ม (dichotomous variables) แล้วจะทำให้ได้ค่าความสัมพันธ์หรือค่าความเที่ยงหรือค่าความสอดคล้องมีค่าที่ต่ำมาก เพราะความแปรปรวนของข้อมูลจะถูกลดทอนลง (Popham, 1978) จึงมีการประยุกต์สูตรความสัมพันธ์ตามแนวทางการคำนวณแบบ Pearson product moment มาใช้สถิตินอนพาราเมตริกซ์แทน เช่น ใช้สัมประสิทธิ์สหสัมพันธ์ฟี (phi coefficient) เพื่อคำนวณความสัมพันธ์ของตัวแปร 2 ตัว (Glass & Hopkins, 1984)

ในกรณีที่ตัวแปร X ละ Y เป็นตัวแปรแบบแบ่ง 2 (ให้ค่าตัวแปรเป็น 0 และ 1) นั้น เมื่อนำจำนวนผู้สอบผ่านและไม่ผ่านตามเกณฑ์ตัดสินทั้ง 2 เกณฑ์ จัดลงตารางแบบ 2 X 2 ได้ดังนี้

		เกณฑ์ภายนอก (จำแนกผู้สอบจากการวัดผล 1) ไม่รอบรู้	รอบรู้	รวม
เกณฑ์ภายใน (จำแนกผู้สอบจากการวัดผล 2)	รอบรู้	A	B	A+B
	ไม่รอบรู้	C	D	C+D
	รวม	A+C	B+D	N

จากความสัมพันธ์ของข้อมูลจำนวนผู้สอบในแต่ละเซลล์ข้างต้น พบว่า สามารถจัดกลุ่มผู้สอบได้ ทั้งกลุ่มที่ได้รับการตัดสินที่สอดคล้องกันทั้ง 2 เกณฑ์ นั่นคือ รอบรู้ตรงกันทั้ง 2 เกณฑ์ และไม่รอบรู้ตรงกันทั้ง 2 เกณฑ์ (ถือว่าเป็นกลุ่มที่เป็นความแปรปรวนที่ต้องการ หรือ V_t) และกลุ่มที่ตัดสินไม่สอดคล้องกันทั้ง 2 เกณฑ์ นั่นคือ ในเกณฑ์หนึ่งตัดสินว่าเป็นผู้รอบรู้ ส่วนอีกเกณฑ์หนึ่งตัดสินว่าเป็นผู้ไม่รอบรู้ (ถือว่าเป็นกลุ่มที่เป็นความแปรปรวนที่คลาดเคลื่อนหรือ V_E) จึงนำข้อมูลสัดส่วนแต่ละกลุ่มมากำหนดค่าตัวแปร V_E และ V_X ได้ดังนี้

การพิจารณาค่า V_E ซึ่งเป็นความแปรปรวนเนื่องจากความคลาดเคลื่อนในการจำแนกผู้รอบรู้ หรือเป็นการจำแนกผู้สอบที่ได้รับการตัดสินความรอบรู้ ไม่สอดคล้องกันทั้ง 2 เกณฑ์ ซึ่งมีค่าเท่ากับค่าความแปรปรวนของผลบวกของ สัดส่วนจำนวนผู้สอบในช่องหรือเซลล์ที่ตัดสินผลไม่สอดคล้องกันนั่นคือ สัดส่วนของจำนวนผู้สอบในเซลล์ A กับจำนวนผู้สอบทั้งหมด (N) หรือ P_A รวมกับสัดส่วนจำนวนผู้สอบในช่องหรือเซลล์ D กับจำนวนผู้สอบทั้งหมด (N) หรือ P_D

นั่นคือ V_E คือ ความแปรปรวนของ P_A รวมกับ P_D

เนื่องจากค่าความแปรปรวนของความคลาดเคลื่อน(Error Variance) หมายถึง การเปลี่ยนแปลงของการวัดที่เกิดขึ้นอย่างสุ่ม ดังที่ Kerlinger กล่าวว่า “Error Variance is random variance” และ ความแปรปรวนที่เกิดขึ้นอย่างสุ่ม (Sampling variance) นั้นเป็นความแปรปรวนที่คำนวณจากกลุ่มตัวอย่าง (Samples) Kerlinger (1973) ดังนั้นการคำนวณความแปรปรวนของความคลาดเคลื่อนของผลการตัดสิน (หรือผลจากการตัดสินทั้ง 2 เกณฑ์ที่ไม่สอดคล้องกัน) ซึ่งถือว่าเกิดขึ้นอย่างสุ่ม จึงประยุกต์จากแนวคิดของสูตรคำนวณความแปรปรวนสำหรับข้อมูลที่ได้จากกลุ่มตัวอย่าง (Glass & Hopkins, 1984) ดังนี้

$$S_E^2 = \frac{SS}{df} \text{ หรือ } S_E^2 = \frac{\sum_i (X_i - \overline{X})^2}{df} \quad \text{..... (7)}$$

เมื่อ

S_E^2 คือ ความแปรปรวนของความคลาดเคลื่อน (จากการตัดสินทั้ง 2 เกณฑ์ที่ไม่สอดคล้องกัน)

SS คือ Sum of Squares หรือ ผลรวมของกำลังสองของผลต่าง = $\sum_i (X_i - \overline{X})^2$

X_i คือ ค่าหรือคะแนนที่สังเกตได้

$\overline{X}$ คือ ค่าเฉลี่ยของคะแนน/ค่าที่สังเกตได้

df คือ องศาของความเป็นอิสระ (Degrees of freedom)

สำหรับการพิจารณาความแปรปรวนของความคลาดเคลื่อน (V_E) โดยการวิเคราะห์ผลต่างระหว่างค่าที่สังเกตได้ (X_i) กับค่าเฉลี่ยของค่าที่สังเกตได้ ($\overline{X}$) หรือ X_i - $\overline{X}$ สามารถตีความได้ว่า ผลลัพธ์ของ X_i - $\overline{X}$ คือ ความคลาดเคลื่อนของการวัด ซึ่งสำหรับการตัดสินผู้สอบโดยการใช้เกณฑ์ตัดสิน 2 เกณฑ์นั้น มีค่าเท่ากับผลบวกของสัดส่วนของเซลล์ที่ตัดสินคลาดเคลื่อน ดังที่กล่าวไปแล้วข้างต้น นั่นคือ X_i - $\overline{X}$ ก็คือ $P_A + P_D$ นั่นเอง

สำหรับ df ของข้อมูลตารางแบบ 2 X 2 มีค่าเท่ากับ (C-1) X (R-1) เมื่อ C คือจำนวนแถวตามแนวตั้งหรือจำนวน Column และ R คือ จำนวนแถวตามแนวนอน หรือ จำนวน Row ดังนั้น df คือ (2-1)X(2-1) =1

เมื่อนำค่า X_i - $\overline{X}$ = $P_A + P_D$ และ df =1 แทนลงใน (7) จึงสามารถสรุปให้สมการที่ (7) เป็นการหา V_E หรือ S_E^2 จากการตัดสินด้วย 2 เกณฑ์ดังนี้

$$V_E \text{ หรือ } S_E^2 = \frac{\sum_i (X_i - \overline{X})^2}{df} = \frac{(P_A + P_D)^2}{1}$$

จะได้ $V_E = (P_A + P_D)^2$ ..... (8)

ในการพิจารณา V_X ซึ่งเป็นความแปรปรวนรวมจากการวัด โดยนำความแปรปรวนของการวัดแต่ละครั้งมารวมกัน ซึ่งความแปรปรวนของการวัดแต่ละครั้งเท่ากับผลคูณสัดส่วนจำนวนผู้สอบผ่าน กับสัดส่วนจำนวนผู้สอบไม่ผ่านของแต่ละครั้ง (หรือ $V_i = p_i X q_i$)

ที่มาของสูตรคำนวณความแปรปรวนกรณีการตัดสินให้คะแนนแบบไม่ต่อเนื่อง 2 ค่า เช่น ผ่านเกณฑ์ ให้คะแนนเท่ากับ 1 และไม่ผ่านเกณฑ์ ให้คะแนนเท่ากับ 0 (นั่นคือ X แทน 1 หรือ 0 เท่านั้น)

เริ่มจากสูตรคำนวณความแปรปรวน (V_i หรือ S_i^2) เป็นการพิจารณาผลการตัดสินความรอบรู้จากทุกเกณฑ์ (ทั้ง 2 เกณฑ์) ถือเป็นเสมือนเป็นการศึกษาจากประชากรของเกณฑ์การตัดสิน จึงอิงแนวทางการคำนวณความแปรปรวนของประชากร ซึ่งรายละเอียดสูตรมีดังนี้

$$V_i \text{ หรือ } S_i^2 = \frac{\sum(X_i - \overline{X}_i)^2}{N} \quad \text{.....(9)}$$

ถ้ากำหนดให้ สัดส่วนของผู้ตอบถูกหรือสอบผ่านเกณฑ์หรือเป็นผู้รอบรู้ = p
สัดส่วนของผู้ตอบผิดหรือสอบไม่ผ่านเกณฑ์หรือเป็นผู้ไม่รอบรู้ = q
เมื่อสัดส่วนของผู้สอบทั้งหมด = 1 นั่นคือ p+q =1
ดังนั้น สัดส่วนของผู้ตอบผิดหรือสอบไม่ผ่านเกณฑ์ (q) = 1-p

จะได้ว่า สัดส่วนผู้สอบผ่านเกณฑ์ (p_i) = $\frac{\sum X_i}{N} = \overline{X}_i$

ดังนั้นจึงนำค่า $p_i = \overline{X}_i$ แทนลงใน (9) จะได้สมการ ดังนี้

$$V_i \text{ หรือ } S_i^2 = \frac{\sum(X_i - p_i)^2}{N} = \frac{\sum X_i^2}{N} - \frac{2p_i \sum X_i}{N} + \frac{\sum p_i^2}{N} \quad \text{.....(10)}$$

เนื่องจากค่าของ X จะเป็น 1 หรือ 0 เท่านั้น ดังนั้น $X_i^2 = X_i$

ดังนั้นจะได้ $\frac{\sum X_i^2}{N} = \frac{\sum X_i}{N} = p_i$ ซึ่งเป็นค่าคงที่

และเนื่องจาก p_i เป็นค่าคงที่ จะได้ว่า $\frac{\sum p_i^2}{N} = \frac{Np_i^2}{N} = p_i^2$

จากสมการ (10) สามารถเขียนสมการใหม่ได้ดังนี้

$$V_i \text{ หรือ } S_i^2 = p_i - 2p_i p_i + p_i^2 = p_i - 2p_i^2 + p_i^2 = p_i - p_i^2 = p_i(1 - p_i) = p_i q_i$$

ดังนั้น V_i หรือ $S_i^2 = p_i q_i$ .....(11)

สรุปได้ว่า

ความแปรปรวนของการตัดสินผู้รอบรู้ – ไม่รอบรู้ ในแต่ละครั้ง (i) เท่ากับผลคูณระหว่างสัดส่วนของผู้รอบรู้ กับ สัดส่วนของผู้ไม่รอบรู้ โดยในการตัดสินความรอบรู้เมื่อทำการตัดสินจาก 2 เกณฑ์ จึงคำนวณความแปรปรวน

รวมทั้งหมดโดยนำความแปรปรวนของการตัดสินความความรอบรู้ครั้งที่ 1 (V_1) **รวมกับ**ความแปรปรวนของการตัดสินความความรอบรู้ครั้งที่ 2 (V_2)

ดังนั้น $V_x = V_1 + V_2 = p_1q_1 + p_2q_2$

โดยที่

V_x คือ ความแปรปรวนรวมจากการตัดสินผู้รอบรู้ – ไม่รอบรู้ ทั้ง 2 ครั้ง

V_1 คือ ความแปรปรวนของการตัดสินผู้รอบรู้ – ไม่รอบรู้ ครั้งที่ 1 (เกณฑ์ภายใน)

V_2 คือ ความแปรปรวนของการตัดสินผู้รอบรู้ – ไม่รอบรู้ ครั้งที่ 2 (เกณฑ์ภายนอก)

p_1 คือสัดส่วนของผู้สอบผ่านแบบสอบอิงเกณฑ์หรือเกณฑ์ภายใน เท่ากับ $P_{(C+D)}$

q_1 คือสัดส่วนของผู้สอบไม่ผ่านแบบสอบอิงเกณฑ์หรือเกณฑ์ภายใน เท่ากับ $P_{(A+B)}$

p_2 คือสัดส่วนของผู้สอบผ่านเกณฑ์ภายนอก เท่ากับ $P_{(A+C)}$

q_2 คือสัดส่วนของผู้สอบไม่ผ่านเกณฑ์ภายนอก เท่ากับ $P_{(B+D)}$

ทำให้ได้ $V_X = [P_{(C+D)}\ P_{(A+B)}] + [P_{(A+C)}\ P_{(B+D)}]$ .....(12)

นำค่า V_E จากสมการ (8) และค่า V_X จากสมการ (12) แทนลงในสมการ (5) จะได้สูตรคำนวณดัชนีความคลาดเคลื่อนของการจำแนกผู้สอบ (Error of Classification Index : ECI) ดังนี้

$$ECI = \frac{(P_A + P_D)^2}{[P_{(A+B)}P_{(C+D)}] + [P_{(A+C)}P_{(B+D)}]} \quad \text{.....(13)}$$

โดยที่

P_A คือ สัดส่วนจำนวนผู้สอบที่ผ่านเกณฑ์ภายใน แต่ไม่ผ่านเกณฑ์ภายนอก

P_D คือ สัดส่วนจำนวนผู้สอบที่ไม่ผ่านเกณฑ์ภายใน แต่ผ่านเกณฑ์ภายนอก

$P_{(C+D)}$ คือสัดส่วนของผู้สอบที่ไม่ผ่านแบบสอบอิงเกณฑ์หรือเกณฑ์ภายใน

$P_{(A+B)}$ คือสัดส่วนของผู้สอบที่ผ่านแบบสอบอิงเกณฑ์หรือเกณฑ์ภายใน

$P_{(A+C)}$ คือสัดส่วนของผู้สอบที่ไม่ผ่านเกณฑ์ภายนอก

$P_{(B+D)}$ คือสัดส่วนของผู้สอบที่ผ่านเกณฑ์ภายนอก

ตอนที่ 2 ผลการเปรียบเทียบประสิทธิภาพของวิธีการกำหนดคะแนนจุดตัดที่ผู้วิจัยพัฒนาขึ้น กับวิธีการของแกลส มีดังนี้

(1) คะแนนจุดตัดของค่าความสามารถจากการสอบ O-NET ชั้นมัธยมศึกษาปีที่ 6 รายวิชาคณิตศาสตร์และรายวิชาภาษาอังกฤษ จากวิธีที่พัฒนาขึ้น มีค่าคงที่ ทั้งให้เกรดเฉลี่ยสะสมชั้นปีที่ 1 ระดับอุดมศึกษา (GPAY1) เป็นเกณฑ์ภายนอก หรือให้เกรดเฉลี่ยกลุ่มสาระเป็นเกณฑ์ภายนอก โดยมีค่าเท่ากับ -0.45 และ -039 ตามลำดับ ส่วนคะแนนจุดตัดจากวิธีของแกลส มีค่าคงที่ในวิชาภาษาอังกฤษ คือร้อยละ 28 ทั้งสองเกณฑ์ภายนอก ส่วนวิชาคณิตศาสตร์มีค่าเท่ากับร้อยละ 28 เมื่อเกณฑ์ภายนอกเป็น GPAY1 และมีค่าเท่ากับร้อยละ 38 เมื่อเกณฑ์ภายนอกเป็นเกรดเฉลี่ยกลุ่มสาระ (รายละเอียดดังตาราง 1)

(2) ผลการเปรียบเทียบประสิทธิภาพของวิธีการกำหนดคะแนนจุดตัดที่พัฒนาขึ้น กับวิธีของแกลส พบว่า วิธีการกำหนดคะแนนจุดตัดทั้ง 2 วิธีต่างก็มีค่าสัมประสิทธิ์ความสอดคล้องในการตัดสินความรอบรู้อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ซึ่งผลการเปรียบเทียบค่าสัมประสิทธิ์สหสัมพันธ์ฟีพบว่าคะแนนจุดตัดที่ได้จากวิธีที่พัฒนาขึ้น มีค่าสัมประสิทธิ์สหสัมพันธ์ฟีสูงกว่า ค่าสัมประสิทธิ์สหสัมพันธ์ฟีจากคะแนนจุดตัดที่ได้จากวิธีของแกลส 3 คู่ และคะแนนจุดตัดที่ได้จากวิธีที่พัฒนาขึ้น มีค่าสัมประสิทธิ์สหสัมพันธ์ฟี เท่ากับ ค่าสัมประสิทธิ์สหสัมพันธ์ฟีจากคะแนนจุดตัดที่ได้จากวิธีของแกลส 1 คู่ แสดงว่าวิธีการกำหนดคะแนนจุดตัดที่ผู้วิจัยพัฒนาขึ้นเป็นวิธีที่มีประสิทธิภาพดีกว่าวิธีการกำหนดคะแนนจุดตัดของแกลส (รายละเอียดดังตาราง 1)

ตาราง 1 การเปรียบเทียบประสิทธิภาพของวิธีการกำหนดคะแนนจุดตัด ระหว่าง วิธีการที่ผู้วิจัยพัฒนาขึ้น(ศึกษาจากค่าความสามารถ) กับวิธีการของแกลสโดยอาศัยทฤษฎีการตัดสินใจ (ศึกษาจากคะแนนดิบ)

ความรอบรู้ด้าน	คู่ที่	เกณฑ์ภายนอก	วิธีที่ผู้วิจัยพัฒนาขึ้น		วิธีของแกลส		Z
			คะแนนจุดตัด	r_ϕ (Z_r)	คะแนนจุดตัด	r_ϕ (Z_r)	
คณิตศาสตร์	1	GPAY1	-0.45	0.1993** (0.2020)	28	0.1567** (0.1580)	3.48*
	3	GPAM	-0.45	0.4057** (0.4305)	38	0.4057** (0.4305)	0.00
ภาษาอังกฤษ	2	GPAY1	-0.39	0.2037** (0.2066)	28	0.1779** (0.1798)	2.12*
	4	GPAE	-0.39	0.2702** (0.2771)	28	0.2426** (0.2475)	2.34*

ทดสอบค่า r_ϕ : ** $p< 0.01$ ($t_{12495,0.01}$ = 2.326; $t_{12495,0.01}$ = 2.326)

ทดสอบค่า Z : * $p<0.05$ ($Z_{.05}$ = 1.64)

(3) ข้อค้นพบอื่น ๆ : ผู้วิจัยทำการศึกษาเพิ่มเติมที่นอกเหนือจากขอบเขตการศึกษา ได้ผลดังนี้

3.1 ผลการศึกษาจากคะแนนดิบทั้ง 2 วิธี: วิธีการกำหนดคะแนนจุดตัดที่พัฒนาขึ้น ทำให้ได้คะแนนจุดตัดคงที่ เมื่อเปลี่ยนเกณฑ์ภายนอก ในขณะที่วิธีการกำหนดคะแนนจุดตัดของแกลส ส่วนใหญ่ทำให้ได้คะแนนจุดตัดที่เปลี่ยนไป เมื่อเปลี่ยนเกณฑ์ภายนอก

3.2 ผลการศึกษาจากค่าความสามารถทั้ง 2 วิธี: วิธีการกำหนดคะแนนจุดตัดที่พัฒนาขึ้น ทำให้ได้คะแนนจุดตัดคงที่ เมื่อเปลี่ยนเกณฑ์ภายนอก ในขณะที่วิธีการกำหนดคะแนนจุดตัดของแกลส ส่วนใหญ่ทำให้ได้คะแนนจุดตัดที่เปลี่ยนไป เมื่อเปลี่ยนเกณฑ์ภายนอก

3.3 ผลการศึกษาคะแนนจุดตัดตามวิธีที่พัฒนาขึ้นเมื่อกำหนดคะแนนจุดตัดระหว่างกำหนดจากคะแนนดิบและกำหนดจากค่าความสามารถ พบว่า ให้ประสิทธิภาพใกล้เคียงหรือเท่ากัน

3.4 ผลการศึกษาคะแนนจุดตัดที่ได้จากวิธีที่ผู้วิจัยพัฒนาขึ้น เมื่อเปรียบเทียบกับที่ได้จากวิธีของแกลส ใน**กรณีที่ศึกษากับค่าความสามารถ** จำนวน 4 คู่ พบว่า ผลการเปรียบเทียบ 3 คู่ แสดงให้เห็นว่าคะแนนจุดตัดที่ได้จากวิธีที่พัฒนาขึ้นมีประสิทธิภาพดีกว่าการคะแนนจุดตัดที่ได้จากวิธีของแกลส และมี 1 คู่ที่ให้ประสิทธิภาพเท่ากัน

3.5 ผลการศึกษาคะแนนจุดตัดตามวิธีที่พัฒนาขึ้น เมื่อเปรียบเทียบกับวิธีของแกลส ใน**กรณีที่ศึกษากับคะแนนดิบ** จำนวน 4 คู่ พบว่า ผลการเปรียบเทียบ 3 คู่ แสดงให้เห็นว่าคะแนนจุดตัดจากวิธีที่พัฒนาขึ้นมีประสิทธิภาพดีกว่าคะแนนจุดตัดที่ได้จากวิธีของแกลส และมี 1 คู่ที่ให้ประสิทธิภาพเท่ากัน

อภิปรายผล

จากผลการพัฒนาวิธีการกำหนดคะแนนจุดตัดที่พบว่า วิธีการกำหนดคะแนนจุดตัดที่ผู้วิจัยพัฒนาขึ้นให้คะแนนจุดตัดที่มีค่าไม่เปลี่ยนแปลง ถึงแม้จะเปลี่ยนตัวแปรที่เป็นเกณฑ์ภายนอกทั้ง 2 รายวิชา ในขณะที่วิธีของแกลสให้คะแนนจุดตัดที่มีค่าเปลี่ยนไปเมื่อเปลี่ยนแปลงตัวแปรที่เป็นเกณฑ์ภายนอกในรายวิชาคณิตศาสตร์ และจากการเปรียบเทียบประสิทธิภาพของวิธีการกำหนดคะแนนจุดตัด ด้วยการเปรียบเทียบค่าสัมประสิทธิ์สหสัมพันธ์ฟี จำนวน 4 คู่ พบว่าวิธีการกำหนดคะแนนจุดตัดที่พัฒนาขึ้นมีประสิทธิภาพดีกว่า วิธีของแกลส จำนวน 3 คู่ และมีประสิทธิภาพเท่ากันจำนวน 1 คู่

นอกจากนี้ผู้วิจัยได้ทดลองศึกษาประสิทธิภาพ ของสูตรวิเคราะห์ค่าความคลาดเคลื่อน เพื่อการตัดสินความรอบรู้ ระหว่างสูตร ECI ที่พัฒนาขึ้น กับสูตร $f(C_x)$ ของแกลส จากการศึกษาข้อมูลชุดเดียวกัน โดยศึกษากับข้อมูลดิบ (ร้อยละของคะแนน) 4 คู่ และศึกษาจากค่าความสามารถ 4 คู่ รวมเป็น 8 คู่ พบว่า การใช้สูตร ECI ทำให้ได้คะแนนจุดตัดที่มีประสิทธิภาพดีกว่าการใช้สูตร $f(C_x)$ จำนวน 6 คู่ และมี 2 คู่ ที่ทำให้ได้คะแนนจุดตัดที่มีประสิทธิภาพเท่ากัน ซึ่งชี้ให้เห็นว่าสูตร ECI สามารถนำมากำหนดคะแนนจุดตัดได้อย่างมีประสิทธิภาพและส่วนใหญ่ได้ผลดีกว่าการใช้สูตร $f(C_x)$ ของแกลส ไม่ว่าจะวิเคราะห์ข้อมูลจากคะแนนดิบและวิเคราะห์จากค่าความสามารถก็ตาม ซึ่งจากข้อค้นพบข้างต้นสามารถอภิปรายผลการวิจัยได้ 2 ประเด็น ดังนี้

ประเด็นที่ 1 วิธีการกำหนดคะแนนจุดตัดที่ผู้วิจัยพัฒนาขึ้น มีการใช้สูตรวัดความคลาดเคลื่อนในการจำแนกผู้รอบรู้ออกจากผู้ไม่รอบรู้ (ECI) ที่ผู้วิจัยพัฒนาขึ้น ซึ่งพบว่า สามารถใช้กำหนดคะแนนจุดตัดได้อย่างมีประสิทธิภาพ ทั้งวิเคราะห์จากคะแนนดิบและจากค่าความสามารถ ทั้งนี้เนื่องมาจากการพัฒนาสูตร ECI ที่ผู้วิจัยอิงแนวคิดการคำนวณความคลาดเคลื่อนจากนิยามความเที่ยงของ Kerlinger (1973) ประกอบกับยึดแนวคิดของเครื่องมือในการวัดที่มีความเที่ยงคือเครื่องมือที่นำไปทดสอบผู้สอบซ้ำกันภายใต้สภาพการทดสอบที่เหมือนเดิม แล้วต้องให้ผลการทดสอบที่ใกล้เคียงกัน หรือได้ผลการวัดที่มีความคงเส้นคงวา (ศิริชัย กาญจนวาสี, 2544) ซึ่งมีการคำนวณความคลาดเคลื่อนที่ปะปนอยู่ในสูตรการคำนวณค่าความเที่ยง ถือว่าแนวคิดนิยามความเที่ยงของ Kerlinger เป็นแนวคิดที่มาตรฐาน ได้รับการยอมรับและใช้กันอย่างแพร่หลาย (ทั้งตามทฤษฎีการวัดผลแบบดั้งเดิม และการวัดผลแนวใหม่) ในขณะที่วิธีการประมาณค่าความคลาดเคลื่อนของการวัดความรอบรู้จากการวัดผล 2 ครั้ง ของแกลส ($f(C_X)$) ประมาณจาก ค่าสัดส่วนของความแปรปรวนของคะแนนความคลาดเคลื่อน กับ ความแปรปรวนของคะแนนจริง โดยความแปรปรวนของคะแนนจริงที่เป็นตัวหารอาจมีค่าน้อยกว่า ความแปรปรวนของคะแนนรวมผู้สอบ ทั้งนี้ หากการวัดมีความคลาดเคลื่อนเกิดขึ้น โอกาสที่ความแปรปรวนของคะแนนจริง จะมีค่าน้อยกว่า ความแปรปรวนของคะแนนรวมผู้สอบ เป็นไปได้สูง ดังนั้น ค่าความคลาดเคลื่อนตามวิธีของ $f(C_X)$ ของแกลส ที่คำนวณได้จะมีค่าสูงกว่าความเป็นจริง ในขณะที่สูตร ECI นั้นจะให้ค่าความคลาดเคลื่อนที่แม่นยำกว่า จึงทำให้ค่าความคลาดเคลื่อนในการจำแนกผู้รอบรู้จากการใช้ 2 เกณฑ์ร่วมพิจารณา ที่คำนวณจากสูตร ECI มีประสิทธิภาพมากกว่าการใช้สูตรการคำนวณความคลาดเคลื่อนของแกลส

ประเด็นที่ 2 ในการศึกษาครั้งนี้คำนึงถึงการเลือกเกณฑ์ภายนอกที่มีคุณสมบัติความตรงตามเกณฑ์สัมพันธ์ (criterion – related validity) กับคะแนนเกณฑ์ภายในที่เป็นข้อมูลจากการสอบ O-NET ระดับมัธยมศึกษาปีที่ 6 โดยผู้วิจัยกำหนดให้ GPAY1 เป็นเกณฑ์ภายนอกที่ใช้แทนข้อมูล GPAX เมื่อจบปริญญาตรี เพราะจากผลการศึกษาข้อมูลเบื้องต้นที่พบว่า ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างผลการสอบ O-NET กับ GPAY1 มีค่าสูงกว่าค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างผลการสอบ O-NET กับ GPAY2 ซึ่งมีค่าอยู่ในระดับกลาง ๆ ทั้ง 2 ค่า และเมื่อทำการตรวจสอบการใช้ GPAY1 เป็นข้อมูลตัวแทนผลการเรียนเมื่อจบปริญญาตรี (GPAX) กับข้อมูลของกลุ่มตัวอย่างอีกกลุ่มที่จบการศึกษาแล้วพบว่า GPAY1 มีสัมประสิทธิ์ความตรงเชิงทำนายอยู่ในเกณฑ์สูง

โดยประเด็นที่น่าพิจารณาคือค่าสัมประสิทธิ์ความตรงตามสภาพ และสัมประสิทธิ์ความตรงเชิงทำนายที่ปรากฏค่าอยู่ในระดับกลาง ๆ ทั้งนี้อาจเป็นเพราะกลุ่มเป้าหมายในการศึกษาเป็นกลุ่มที่มีผลการเรียนดีของประเทศคิดเป็นผู้ที่มีคะแนนสอบ O-NET สูง จึงถือว่ากลุ่มตัวอย่างครั้งนี้เป็นกลุ่มตัวอย่างคัดสรร ซึ่งการศึกษากับกลุ่มคัดสรรจะทำให้ค่าสัมประสิทธิ์ความตรงตามเกณฑ์สัมพันธ์ลดลงจากค่าที่ควรจะเป็น (ศิริชัย กาญจนวาสี, 2544) ดังนั้นหากสามารถรวบรวมข้อมูลจากตัวแทนของกลุ่มประชากรมาทำการคำนวณ ค่าสัมประสิทธิ์ความตรงตามเกณฑ์สัมพันธ์อาจทำให้ค่าสูงขึ้น จึงส่งผลให้ผลการศึกษาพบว่าคะแนนจุดตัดที่ได้จากทั้ง 2 วิธีต่างก็มีค่าสัมประสิทธิ์ความสอดคล้องในการตัดสินความรอบรู้อย่างมีนัยสำคัญทางสถิติที่ระดับ .01

ข้อเสนอแนะ

ข้อเสนอแนะในการนำผลการวิจัยไปใช้

1. วิธีการกำหนดคะแนนจุดตัดเพื่อตัดสินความรอบรู้ที่ผู้วิจัยพัฒนาขึ้นสามารถใช้กำหนดคะแนนจุดตัดได้ดีทั้งวิเคราะห์จากข้อมูลที่เป็นคะแนนดิบและจากค่าความสามารถ ดังนั้นจึงสามารถใช้วิธีการกำหนดคะแนนจุดตัดที่พัฒนาขึ้นกับการวัดผลแบบอิงเกณฑ์ได้ทุกระดับ ตั้งแต่ใช้กับกลุ่มใหญ่ ๆ จนกระทั่งไปถึงระดับห้องเรียนได้ด้วย

2. คะแนนจุดตัดของเกณฑ์ภายในที่ได้จากการศึกษาครั้งนี้ ขึ้นกับเกณฑ์ภายนอกที่มีคะแนนจุดตัดเหมาะกับกลุ่มนักศึกษาทั่วไป (GPAX = 2.50) ซึ่งหากมีการเปลี่ยนเกณฑ์ภายนอกตามคุณสมบัติที่เป็นความต้องการของแต่ละคณะวิชาแล้ว คะแนนจุดตัดของคะแนนเกณฑ์ภายในที่ได้จะมีการเปลี่ยนค่าไป ซึ่งสามารถคำนวณคะแนนจุดตัดได้ใหม่ เช่น ต้องการรับคัดเลือกผู้เรียนที่จะเข้ามหาวิทยาลัยโดยอาจกำหนดคุณสมบัติว่าเมื่อจบการศึกษาระดับปริญญาตรีต้องได้เกรดเฉลี่ยรวม (GPAX) เท่ากับ 3.50 ก็สามารถเปลี่ยนคะแนนจุดตัดของเกณฑ์ภายนอกเป็น 3.50 และเก็บข้อมูลนักศึกษาที่มี GPAX และข้อมูลย้อนหลังของการสอบ O-NET แล้วคำนวณคะแนนจุดตัดของ O-NET จากนั้นสามารถนำคะแนนจุดตัดของเกณฑ์ภายในหรือผลการสอบ O-NET ที่ได้ มากำหนดเป็นเกณฑ์การรับผู้เรียนด้านคะแนนสอบ O-NET ที่เพิ่งจบชั้นมัธยมศึกษาตอนปลายเพื่อเข้าศึกษาต่อได้

3. หน่วยงานที่ทำการทดสอบความสามารถของผู้เรียน หรือผู้สอบจำนวนมาก สามารถนำวิธีการกำหนดคะแนนจุดตัดเพื่อตัดสินความรอบรู้ที่พัฒนาขึ้นไปใช้ โดยมีเงื่อนไข ดังนี้

3.1 ให้กำหนดเกณฑ์ภายนอกที่มีคุณสมบัติของการเป็นเกณฑ์อ้างอิงที่ดี โดยต้องมีค่าสัมประสิทธิ์ความตรงตามเกณฑ์สัมพันธ์สูงกับตัวแปรเกณฑ์ภายใน (ที่ต้องการนำมากำหนดคะแนนจุดตัด)

3.2 คะแนนจุดตัดของเกณฑ์ภายนอกต้องมีแหล่งอ้างอิงหรือที่มาที่ถูกต้องและน่าเชื่อถือ

4. ผลจากการคำนวณคะแนนจุดตัดความรอบรู้จากการสอบ O-NET ชั้นมัธยมศึกษาปีที่ 6 โดยเป็นผลการศึกษาย้อนหลังของผู้สอบที่เข้าเรียนระดับปริญญาตรีแล้ว ซึ่งสามารถนำสารสนเทศคะแนนจุดตัดดังกล่าว เพื่อเสนอแนะกับผู้สอบที่มีผลการสอบ O-NET ในระดับชั้นมัธยมศึกษาปีที่ 6 ที่เพิ่งจบระดับมัธยมศึกษาตอนปลายจาก

โรงเรียนในสังกัดสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน (สพฐ.) ได้ ถ้าแบบสอบ O-NET ชั้นมัธยมศึกษาปีที่ 6 ในรายวิชาคณิตศาสตร์และภาษาอังกฤษยังอยู่ภายใต้เงื่อนไขของหลักสูตรเดียวกัน และยังสามารถใช้ผลการสอบ O-NET ทำนายผลการเรียนในระดับปริญญาตรีของผู้สอบล่วงหน้าได้ เช่น หากผู้สอบใดที่มีผลการสอบ O-NET วิชาคณิตศาสตร์และภาษาอังกฤษ ชั้นมัธยมศึกษาปีที่ 6 ตั้งแต่คะแนนจุดตัดขึ้นไป ผู้สอบนั้นสามารถศึกษาต่อในระดับปริญญาตรีในสถาบันของรัฐบาลได้และมีโอกาสประสบความสำเร็จ หรือได้ GPA ปีชั้นปีที่ 1 (GPAY1) เท่ากับ 2.50 ขึ้นไป ส่วนผู้ที่ได้คะแนนสอบ O-NET ต่ำกว่าคะแนนจุดตัดมีโอกาสได้ GPAY1 ต่ำกว่า 2.50 ซึ่งผู้เรียนกลุ่มดังกล่าวต้องขยันเพื่อความไม่ประมาทต่อการได้ GPAX ขณะจบการศึกษาระดับปริญญาตรีไม่ถึง 2.50 ทั้งนี้เพราะ GPAY1 สามารถใช้ทำนาย GPAX ขณะจบปริญญาตรีได้ด้วย เนื่องจากพบว่า GPAY1 มีความสัมพันธ์สูงกับ GPAX ขณะจบการศึกษาระดับปริญญาตรี

5. สำนักทดสอบทางการศึกษา สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน สำนักงานคณะกรรมการเขตพื้นที่การศึกษา และหน่วยงานที่เกี่ยวข้องกับการวัดผลการศึกษา ควรนำวิธีการกำหนดคะแนนจุดตัดเพื่อตัดสินความรอบรู้ที่พัฒนาขึ้นไปใช้ เพื่อวัตถุประสงค์แตกต่างกัน ดังนี้

5.1 สามารถกำหนดคะแนนจุดตัดได้จากคะแนนดิบและค่าความสามารถ โดยใช้เกณฑ์ภายนอกเป็นข้อมูลในช่วงเวลาเดียวกับข้อมูลเกณฑ์ภายใน และควรดำเนินการเป็นระยะ ๆ แบบต่อเนื่อง ในปีสุดท้ายของการเรียนแต่ละช่วงชั้น โดยควรเริ่มต้นตั้งแต่ช่วงชั้นที่ 1 เพื่อตัดสินความรอบรู้ของผู้เรียน เพื่อให้ผู้เรียนได้มีโอกาสหลายปีในการปรับปรุงตนเองในการพัฒนาความรอบรู้ให้ผ่านเกณฑ์ จนกว่าจะจบการศึกษาระดับมัธยมศึกษาตอนปลาย

5.2 สำหรับการกำหนดคะแนนจุดตัดจากค่าความสามารถนั้น โดยธรรมชาติแล้วค่าความสามารถ (Ability) ที่ได้จากการวิเคราะห์ด้วยทฤษฎี IRT จะมีลักษณะการแจกแจงแบบโค้งปกติเสมอ ดังนั้นคะแนนจุดตัดจากข้อมูลเกณฑ์ภายในที่มีการแจกแจงเข้าใกล้โค้งปกติ ถือเป็นสารสนเทศสำหรับผู้สอบที่คละความสามารถทุกระดับ จึงเป็นการให้สารสนเทศคะแนนจุดตัดกับผู้สอบที่มีความสามารถโดยเฉลี่ยอยู่ในระดับกลาง ๆ ในกลุ่มทั่ว ๆ ไป ซึ่งต้องกำหนดคะแนนจุดตัดของเกณฑ์ภายนอกให้สอดคล้องกันด้วย ในกรณีที่ต้องการหาคะแนนจุดตัดค่าความสามารถของผู้สอบเฉพาะกลุ่ม อาจเปลี่ยนคะแนนจุดตัดเกณฑ์ภายนอกให้เหมาะสมและสอดคล้องกับแต่ละกลุ่มเป็นกรณี ๆ ไป

ข้อเสนอแนะสำหรับการวิจัยครั้งต่อไป

1. ในการเลือกตัวแปรที่เป็นเกณฑ์ภายนอกควรใช้ข้อมูลทั้งหมด (ประชากร) ของตัวแปรที่จะใช้เป็นเกณฑ์ภายนอก จากผู้จบการศึกษาระดับมัธยมศึกษาตอนปลายทุกสังกัดที่ได้รับคัดเลือกเข้าศึกษาต่อระดับอุดมศึกษาได้ มาใช้ศึกษาเพื่อให้ได้ค่าสัมประสิทธิ์ความตรงตามเกณฑ์สัมพันธ์ที่แท้จริงและน่าเชื่อถือ เพื่อทำการพิจารณาตัวแปรที่จะใช้เป็นเกณฑ์ภายนอกที่เหมาะสมต่อไป

2. ควรทำการศึกษาคะแนนจุดตัดในรายวิชาพื้นฐาน เช่น คณิตศาสตร์ วิทยาศาสตร์ ภาษาไทย และภาษาอังกฤษ เป็นต้น ควรวัดความรอบรู้ของกลุ่มตัวอย่างที่เป็นตัวแทนของผู้ศึกษาต่อในระดับอุดมศึกษาจากสถาบันการศึกษาที่เป็นตัวแทนจากทุก ๆ ภูมิภาค รวมทั้งเปรียบเทียบคะแนนจุดตัดที่บอกความรอบรู้ของผู้เรียนในคณะวิชาแต่ละสาขาวิชา เช่น สาขาวิชาสังคมศาสตร์ สาขาวิชาวิทยาศาสตร์และเทคโนโลยี และสาขาวิชาวิทยาศาสตร์การแพทย์ เป็นต้น และเปรียบเทียบคะแนนจุดตัดความสามารถของผู้เรียน ในสถาบันอุดมศึกษาที่มีขนาดแตกต่างกัน

3. ควรทำการศึกษาเปรียบเทียบคะแนนจุดตัดความรอบรู้ในรายวิชาพื้นฐานของกลุ่มตัวอย่างที่เป็นตัวแทนของผู้ศึกษาต่อในระดับอุดมศึกษาในสถาบันการศึกษาของรัฐและเอกชน

4. กรณีที่เกรดเฉลี่ยสะสมระดับปริญญาตรี เป็นเกณฑ์ภายนอก ควรทำการศึกษากับข้อมูลคะแนนเกณฑ์ภายในที่เป็นคะแนนรวมจากการสอบ O-NET ระดับมัธยมศึกษาปีที่ 6 จากทุกรายวิชา

เอกสารอ้างอิง

แนวหน้า. (2554). *สสวท.ชี้ผลประเมิน TIMSS ปี 50 คณิต-วิทย์ ไทยยังรั้งอันดับกว่า 20.* ข่าวการศึกษาและวัฒนธรรม. [ออนไลน์]. สืบค้นจาก: http://www.naewna.com/news. asp?ID=137642. [2554, มีนาคม 20].

ปฐมพงษ์ จันทร์สว่าง. (2549). *ผลของการใช้กลวิธีการเรียนเพื่อรอบรู้ของบลูม ต่อผลสัมฤทธิ์ทางการเรียน ความคงทนในการเรียนรู้ เจตคติทางการเรียน และความเชื่อในความสามารถของตนในรายวิชาคณิตศาสตร์เรื่องพาราโบลาของนักเรียนชั้นมัธยมศึกษาปีที่ 3.* วิทยานิพนธ์ปริญญาศึกษาศาสตรมหาบัณฑิต มหาวิทยาลัยขอนแก่น.

ประยุทธ นิสภกุล. (2550*). การเปรียบเทียบผลสัมฤทธิ์ทางการเรียนระหว่างวิธีจัดการเรียนแบบเพื่อรอบรู้กับวิธีจัดการเรียนแบบปกติ วิชาระบบคอมพิวเตอร์และสถาปัตยกรรม หลักสูตรระดับปริญญาตรี สถาบันราชภัฏธนบุรี พ.ศ. 2543*. วิทยานิพนธ์หลักสูตรครุศาสตร์อุตสาหกรรมมหาบัณฑิต สาขาวิชาไฟฟ้า ภาควิชาครุศาสตร์ไฟฟ้า มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ.

พงศ์ธารา วิจิตเวชไพศาล. (2551). การเรียนแบบรอบรู้. *วารสารศึกษาศาสตร์.* 19(1), 17-29.

รวมมิตร การศึกษาต่อระดับปริญญาตรี, [ออนไลน์]. สืบค้นจาก: http://web.ruammid.com/ระดับปริญาตรี. [2554, กุมภาพันธ์ 20].

วรสริน พันธุ. (2550). *ผลสัมฤทธิ์ทางการเรียนวิชาวิทยาศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่1ที่ได้รับการสอนตามหลักการเรียนเพื่อรอบรู้โดยใช้ของเล่นเชิงวิทยาศาสตร์.* วิทยานิพนธ์หลักสูตรศึกษาศาสตรมหาบัณฑิต สาขาวิชาวิทยาศาสตร์ศึกษา มหาวิทยาลัยเชียงใหม่.

วิชิต สุรัตน์เรืองชัย และคณะ. (2549). การศึกษาสภาพและปัญหาการจัดการเรียนการสอนของคณาจารย์มหาวิทยาลัยบูรพา. *วารสารศึกษาศาสตร์,* 17(2), พฤศจิกายน 2548 – มีนาคม 2549, 105 -118.

นิภา ศรีไพโรจน์. (2533). *สถิตินอนพาราเมตริก.* พิมพ์ครั้งที่ 2. กรุงเทพฯ: โอ.เอส.พริ้นติ้งเฮ้าส์.

ศิริชัย กาญจนวาสี. (2544). *ทฤษฎีการทดสอบแบบดั้งเดิม.* พิมพ์ครั้งที่ 4 กรุงเทพฯ: โรงพิมพ์จุฬาลงกรณ์มหาวิทยาลัย.

ศิริชัย กาญจนวาสี. (2555). *ทฤษฎีการทดสอบแนวใหม่.* พิมพ์ครั้งที่ 4. กรุงเทพฯ:โรงพิมพ์แห่งจุฬาลงกรณ์-มหาวิทยาลัย.

ศูนย์แนะแนวศึกษาต่อต่างประเทศ เอ็นริช เอ็ดดูเคชั่น . [ออนไลน์]. สืบค้นจาก: http://www.enrich-ededucation.com/Pathway.html. [2554, มีนาคม 12].

สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี. (มปท.). *ผลการประเมิน PISA 2009 การอ่าน คณิตศาสตร์ และวิทยาศาสตร์, บทสรุปสำหรับผู้บริหาร.* กรุงเทพฯ: โครงการ PISA ประเทศไทย.

สุพัฒน์ สุกมลสันต์. (2530). การหาจุดตัดของแบบทดสอบอิงเกณฑ์และแบบทดสอบอิงปริเขต. *วารสารการวัดผลการศึกษา.* 8(24), 56-74.

สำนักงานคณะกรรมการข้าราชการพลเรือน. (2555). *ประกาศสำนักงาน ก.พ. เรื่อง รับสมัครคัดเลือกข้าราชการเพื่อรับทุนรัฐบาลไปศึกษาวิชา ณ ต่างประเทศ ประจำปี 2555.* [ออนไลน์]. สืบค้นจาก: www.ocsc.go.th/ocsc/th/files/Ann_Gov_55.pdf. [2555, เมษายน 26].

สำนักงานเลขาธิการสภาการศึกษา. (2551). *รายงานการติดตามและประเมินผลการจัดการเรียนรู้ระดับการศึกษาขั้นพื้นฐาน.* กรุงเทพฯ: บริษัท เพลิน สตูดิโอ จำกัด.

สำนักงานเลขาธิการสภาการศึกษา. (2553). *รายงานสภาพการดำเนินงานของสถานศึกษาตามมาตรฐานการศึกษาชาติ.* กรุงเทพฯ: บริษัท พริกหวานกราฟฟิค จำกัด.

สิทธิชัย หาญสมบัติ. (2534). การศึกษาประสิทธิผลของแบบทดสอบแบบเทเลอร์บางรูปแบบในการตัดสิน ความรอบรู้ในการเรียนการสอนวิชาคณิตศาสตร์ชั้นปีที่ 6. *วารสารการวัดผลการศึกษา*, 13(37), 66-78.

Armacost, R.L. & Armacost, J.P. (2003). *Using mastery-based grading to facilitate learning.* Paper presentation in 33rd ASEE/IEEE Frontiers in Education Conference. November 5 - 8, 2003: T3A-20 – T3A-25.

Cizek, G.J. & Bunch, M.B. (2007). *Standard setting : A guide to establishing and evaluating performance standards on tests.* USA: Sage Publications.

Davis-Becker, S.L. & Buckeb, C.W. (2011). *Using the results of multiple standard setting methods to inform policy.* Paper presented at 2011 annual metting of the National Council on Measurement in Education, New Orleans, LA.

Glass, G.V. & Hopkins, K.D. (1984). *Statistical methods in education and psychology* .(2nd ed.). New Jersey: Prentice-Hall.

Kane. M.T. (2002). Practice-Based Standard Setting.*The bar examiner.* August 2002.

Kerlinger, F.N. (1973). *Foundations of behavioral research.* 2nd Edition, USA: Holt, Rinehart & Winston.

Liverpool International College. *Graduate Diploma Requirements.* Available: http://www.livic.org.uk/howtoapply/Graduate-Diploma-Entry-Requirements/Pages/default.aspx. [2011, February 20].

MacCan, R.G. & Stanley, G. (2010). Extending participation in standard setting: an online jugging proposal. *Educ Asse Eval Acc*, 22:139-157.

Popham, W. J. (1978). *Criterion-referenced measurement.* USA: Prentice-Hall, Englewood Cliffs.

Reynolds, C.R., Livingston,R.B., & Willson, V. (2010). *Measurement and assessment in education.* 2nd Edition. New Jersey: Pearson.

มหาวิทยาลัยบูรพา Burapha University