371.2601
ก 382 ก
ฉ.2

การประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบเพื่อกำหนดคะแนนจุดตัดตามเทคนิคของนีเดลสกีกับวิธีการใช้กลุ่มคาบเส้น

ปริญญานิพนธ์

ของ

กัญจนา ลินทรัตนศิริกุล

ส.3 ก.ย. 2535

เสนอต่อมหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร เพื่อเป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรปริญญาการศึกษาดุษฎีบัณฑิต สาขาการวิจัยและพัฒนาหลักสูตร
กันยายน 2534

ลิขสิทธิ์เป็นของมหาวิทยาลัยศรีนครินทรวิโรฒ

180702

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อเปรียบเทียบคะแนนจุดตัด จำนวนผู้รอบรู้ ความเชื่อมั่น และความเที่ยงตรง ในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการกำหนดคะแนนจุดตัด 2 วิธี คือ วิธีการของ แวน เดอร์ ลินเดน และวิธีการของเคน จากการกำหนดเกณฑ์มาตรฐาน 2 วิธี คือการใช้เทคนิคของนีเดลสกีกับวิธีการใช้กลุ่มคาบเส้น กลุ่มตัวอย่างประกอบด้วย ผู้เชี่ยวชาญและนักเรียน ซึ่งแบ่งเป็น 2 กลุ่ม คือ กลุ่มที่มีความสามารถในระดับพอใช้และในระดับดี ข้อมูลที่ใช้ในการวิจัยประกอบด้วยข้อมูลจากโครงการตรวจสอบคุณภาพการศึกษา นักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2533 และระดับผลการเรียนของนักเรียนจากการสอบปลายภาคเรียนที่ 1 และการสอบปลายปี การวิเคราะห์ข้อมูล 1) หาคะแนนจุดตัดตามวิธีการของแวน เดอร์ ลินเดนและเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี และวิธีการใช้กลุ่มคาบเส้น 2) เปรียบเทียบจำนวนผู้รอบรู้โดยการทดสอบค่าสัดส่วน 3) หาความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้โดยวิธีการของฮวิน 4) หาความสอดคล้องหรือความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการทดสอบค่าไคสแควร์ และคอนติงเจนซีตามลำดับ ผลการวิจัยปรากฏว่า

1. เมื่อกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้ปรากฏว่า แบบทดสอบส่วนใหญ่มีคะแนนจุดตัดตามวิธีการของ แวน เดอร์ ลินเดน สูงกว่าวิธีการของเคน แต่กลุ่มนักเรียนที่มีความสามารถในระดับดีปรากฏว่า แบบทดสอบส่วนใหญ่มีคะแนนจุดตัดตามวิธีการของ แวน เดอร์ ลินเดน เท่ากับวิธีการของเคน

2. เมื่อกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้และในระดับดีพบว่า คะแนนจุดตัดตามวิธีการของ แวน เดอร์ ลินเดน สูงกว่าวิธีการของเคนในแบบทดสอบทุกฉบับ

3. คะแนนจุดตัดตามวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี ปรากฏว่า แบบทดสอบส่วนใหญ่มีคะแนนจุดตัดสูงกว่าวิธีการใช้กลุ่มคาบเส้นในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้ และคะแนนจุดตัดจากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี ปรากฏว่า แบบทดสอบส่วนใหญ่มีคะแนนจุดตัดเท่ากับวิธีการใช้กลุ่มคาบเส้นในกลุ่มนักเรียนที่มีความสามารถในระดับดี

4. คะแนนจุดตัดตามวิธีการของเคนจากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี สูงกว่าวิธีการใช้กลุ่มคาบเส้นทั้งในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้และในระดับดี

5. จำนวนผู้รอบรู้ขึ้นอยู่กับคะแนนจุดตัด กล่าวคือ วิธีการหรือเกณฑ์มาตรฐานในการกำหนดคะแนนจุดตัดที่ทำให้คะแนนจุดตัดต่ำจะมีผลทำให้จำนวนผู้รอบรู้มากกว่าวิธีการหรือเกณฑ์มาตรฐานในการกำหนดคะแนนจุดตัดสูง

6. ความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้ขึ้นอยู่กับคะแนนจุดตัด กล่าวคือ วิธีการหรือเกณฑ์มาตรฐานที่ทำให้คะแนนจุดตัดสูงจะทำให้ความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้สูง อย่างไรก็ตามในแบบทดสอบบางฉบับพบว่า ความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้จะมีค่าสูงสุด เมื่อคะแนนจุดตัดเท่ากับคะแนนเฉลี่ยของแบบทดสอบ

7. ความสอดคล้องระหว่างผลการสอบจากแบบทดสอบของโครงการตรวจสอบคุณภาพการศึกษามีความสัมพันธ์กับระดับผลการเรียนของนักเรียนในแต่ละกลุ่มวิชาในภาคเรียนที่ 2 ยกเว้นแบบทดสอบวัดคุณลักษณะทั่วไปไม่มีความสัมพันธ์กับกลุ่มสร้างเสริมลักษณะนิสัยในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้

8. ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ขึ้นอยู่กับความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้

The purposes of this research focused on comparing cut- off scores, number of mastery learners, reliability and validity in classifying mastery learners. The two methods of setting the cut-off scores used were developed by, among others, Van der Linden and Kane ; and the two techniques of setting standard criterion used were Nedelsky technique and Borderline Group method. The samples consisted of experts and students who had satisfy and good school grade. The data used composed of the scores from Nation - wide Assessment of Educational Achievement of Prathomsuksa 6 students for academic year 1990, and the achievement scores from the first and the second semester. The data was analyzed by 1) calculating the cut - off scores according to Van der Linden and Kane methods 2) comparing the number of mastery learners by using Z-test 3) Huynh's formula was used to calculate the reliability of mastery learners 4) calculating the consistency or validity of mastery learners by using χ^2 - test and contingency coefficiency respectively.

The results indicated that :

1. When setting the standard criterions to Nedelsky technique for satisfy school grade students, the cut-off scores in most of the tests determined by Van der Linden were higher than Kane method. But for good school grade students, the cut - off scores in all of the tests by Van der Linden were equal to Kane method.

2. When setting the standard criterions by Borderline Group method for satisfy and good school grade students, the cut-off scores determined by Van der Linden were higher than Kane method in all of the tests.

3. The cut - off scores determined by Van der Linden from setting the standard criterions according to Nedelsky technique; in most of the tests, the cut - off scores were higher than Borderline Group method for satisfy school grade students. But for good school grade students, the cut - off scores from setting the standard criterions according to Nedelsky technique and Borderline Group method were equal in most of the tests.

4. The cut - off scores determined by Kane from setting the standard criterions according to Nedelsky technique were higher than Borderline Group method for satisfy and good school grade students.

5. The number of mastery learners depended upon the cut- off scores. If the methods of the standard criterions that were given the low cut - off scores, the number of mastery learners would be more than the high cut - off scores.

6. The reliability of mastery learners depended upon the cut-off scores. The results further suggest that the higher the cut-off scores were, the higher the reliability of mastery learners had. However, in some tests, the reliability of mastery learners were maximized when the cut- off scores were equal to the average of the tests.

7. The scores from each tests of the Nation-wide Assessment of Educational Achievement were correlated to the scores from second semester achievement in every subject, except the scores from the Test of General Characteristics were not correlated to Habitual Characteristics Development for satisfy school grade students.

8. The Validity of mastery classification depended upon the reliability of classification examinees into achievement status which were usually taken as mastery and nonmastery.

คณะกรรมการควบคุมและคณะกรรมการสอบได้พิจารณาปริญญานิพนธ์ฉบับนี้แล้ว เห็นสมควรเป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาการศึกษาดุษฎีบัณฑิต สาขาการวิจัยและพัฒนาหลักสูตร ของมหาวิทยาลัยศรีนครินทรวิโรฒได้

คณะกรรมการควบคุม

.....ประธาน

(ศ.ดร.สาเริง บุญเรืองรัตน์)

.....กรรมการ

(ดร.ชูศักดิ์ ขัมภลิขิต)

.....กรรมการ

(ดร.สงบ ลักษณะ)

คณะกรรมการสอบ

.....ประธาน

(ศ.ดร.สาเริง บุญเรืองรัตน์)

.....กรรมการ

(ดร.ชูศักดิ์ ขัมภลิขิต)

.....กรรมการ

(ดร.สงบ ลักษณะ)

.....กรรมการที่แต่งตั้งเพิ่มเติม

(รศ.ดร.วินัย รังสินันท์)

บัณฑิตวิทยาลัยอนุมัติให้รับปริญญานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาการศึกษาดุษฎีบัณฑิต สาขาการวิจัยและพัฒนาหลักสูตร ของมหาวิทยาลัยศรีนครินทรวิโรฒ

.....คณบดีบัณฑิตวิทยาลัย

(ศ.ดร.สมพร บัวทอง)

วันที่ ๗ เดือน ตุลาคม พ.ศ. 2534

ประกาศคุณูปการ

ปริญญานิพนธ์ฉบับนี้สำเร็จได้ด้วยดี เนื่องจากผู้วิจัยได้รับความรู้ แนวคิด และคำแนะนำเป็นอย่างดียิ่งจาก ศาสตราจารย์ ดร.สำเริง บุญเรืองรัตน์ อาจารย์ ดร. ชูศักดิ์ ขัมภลิขิต อาจารย์ ดร. สงบ ลักษณะ อาจารย์ที่ปรึกษาปริญญานิพนธ์ โดยเฉพาะอย่างยิ่งอาจารย์ ดร.ชูศักดิ์ ขัมภลิขิต ได้ให้แนวคิดแก่ผู้วิจัยตั้งแต่การกำหนดปัญหาการวิจัย ผู้วิจัยขอกราบขอบพระคุณในความกรุณาของท่านด้วยความเคารพอย่างสูง

ขอกราบขอบพระคุณ รองศาสตราจารย์ ดร.วินัย รังสินันท์ รองอธิการบดีฝ่ายวางแผน มหาวิทยาลัยสุโขทัยธรรมาธิราช ซึ่งเป็นกรรมการสอบภายนอก และได้ให้คำแนะนำอันเป็นประโยชน์สำหรับการทำปริญญานิพนธ์ในครั้งนี้

ขอกราบขอบพระคุณ อาจารย์ ดร.วัลลภ กันทรัพย์ ผู้อำนวยการสำนักงานทดสอบทางการศึกษา กรมวิชาการ กระทรวงศึกษาธิการที่อนุญาตให้ผู้วิจัยใช้ข้อมูลจากการสอบของโครงการตรวจสอบคุณภาพการศึกษา นักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2533 และอำนวยความสะดวกแก่ผู้วิจัยในการเก็บรวบรวมข้อมูลตลอดระยะเวลาของการวิจัย ด้วยความเมตตาของท่านจึงทำให้ผู้วิจัยมีข้อมูลเพื่อนำมาใช้ในการวิจัยครั้งนี้

ผู้ช่วยศาสตราจารย์ ดร.นงลักษณ์ วิรัชชัย อาจารย์ที่ปรึกษาของผู้วิจัยในระหว่างที่ผู้วิจัยเป็นนิสิตระดับปริญญาเอก ถึงแม้ท่านจะไม่ได้เป็นอาจารย์ที่ปรึกษาปริญญานิพนธ์ก็ตาม แต่ท่านก็ได้ติดตามถึงความก้าวหน้าของการวิจัยด้วยความเอาใจใส่อยู่เสมอ ผู้วิจัยขอกราบขอบพระคุณในความกรุณาของท่านไว้ ณ ที่นี้

ขอขอบพระคุณผู้บริหารโรงเรียน ผู้อำนวยการโรงเรียนและอาจารย์ใหญ่ของโรงเรียนที่เป็นกลุ่มตัวอย่างที่ได้ให้ความร่วมมือเป็นอย่างดียิ่งในการกรอกข้อมูลในแบบกรอกรายชื่อ และผลการเรียนของนักเรียนที่ผู้วิจัยจัดส่งไปให้พร้อมกับจัดส่งถึงผู้วิจัยด้วยความเรียบร้อย

ขอขอบคุณ อาจารย์กนก อินทรพฤกษ์ และข้าราชการทุกท่านในฝ่ายประมวลผล สำนักงานทดสอบทางการศึกษา กรมวิชาการ กระทรวงศึกษาธิการที่ได้เขียนโปรแกรมคอมพิวเตอร์และประมวลผลข้อมูลในบางส่วนให้เป็นอย่างดียิ่ง จนงานวิจัยแล้วเสร็จ

กัญจนา ลินทรัตนศิริกุล

สารบัญ

บทที่		หน้า
1	บทนำ	1
	ความเป็นมาและความสำคัญของปัญหา	1
	วัตถุประสงค์ของการวิจัย	13
	สมมุติฐานของการวิจัย	14
	ประโยชน์ที่คาดว่าจะได้รับ	16
	ขอบเขตของการวิจัย	16
	นิยามศัพท์เฉพาะ	17
2	เอกสารและงานวิจัยที่เกี่ยวข้อง	19
	ความหมายและแนวคิดของคะแนนจุดตัด	19
	วิธีการกำหนดคะแนนจุดตัด	23
	ทฤษฎีการตอบสนองต่อข้อสอบกับการกำหนดคะแนนจุดตัด	56
	งานวิจัยที่เกี่ยวข้องกับการกำหนดคะแนนจุดตัดที่นำทฤษฎี การตอบสนองต่อข้อสอบมาประยุกต์ใช้	87
	ความเชื่อมั่นในการตัดสินจำแนกความรอบรู้	90
	โครงการตรวจสอบคุณภาพการศึกษา	105
3	วิธีดำเนินการวิจัย	109
	กลุ่มตัวอย่าง	109
	เครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูล	111
	คุณภาพของเครื่องมือ	112

บทที่ หน้า

การเก็บรวบรวมข้อมูล 113
การวิเคราะห์ข้อมูล 119

4 ผลการวิเคราะห์ข้อมูล 125
ตอนที่ 1 ผลการเปรียบเทียบคะแนนจุดตัด 125
ตอนที่ 2 ผลการเปรียบเทียบจำนวนผู้รอบรู้ 140
ตอนที่ 3 ผลการเปรียบเทียบความเชื่อมั่นในการจำแนกความรอบรู้ 156
ตอนที่ 4 ผลการเปรียบเทียบความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ 171

5 สรุป อภิปรายผลการวิจัยและข้อเสนอแนะ 219
สรุปผลการวิจัย 224
อภิปรายผลการวิจัย 234
ข้อเสนอแนะ 249

บรรณานุกรม 251
ภาคผนวก 257
ภาคผนวก ก รายชื่อโรงเรียนที่เป็นกลุ่มตัวอย่าง 258
ภาคผนวก ข หนังสือขอความอนุเคราะห์ในการเก็บรวบรวมข้อมูล 265
ภาคผนวก ค แบบกรอกรายชื่อและผลการเรียนของนักเรียน 273

ประวัติของผู้วิจัย 275

บัญชีตาราง

ตาราง หน้า

1 การคำนวณตามเทคนิคของนีเดลสกี 34

2 ผลการหาคะแนนจุดตัดจากผู้เชี่ยวชาญโดยเทคนิคของนีเดลสกี 36

3 ผลการตัดสินโดยใช้เทคนิคของแองกอฟฟ์ 39

4 ร้อยละที่คาดหวังจะสอบผ่านตามลักษณะที่เกี่ยวข้องและระดับความยาก 40

5 ตัวอย่างของการคำนวณตามเทคนิคของอีเบล 42

6 คุณภาพของแบบทดสอบแต่ละฉบับที่นำมาใช้เป็นเครื่องมือ 112

7 เกณฑ์มาตรฐานของแบบทดสอบแต่ละฉบับตามที่ผู้เชี่ยวชาญกำหนดโดยใช้เทคนิคของนีเดลสกี 116

8 จำนวนนักเรียนที่ได้ระดับผลการเรียน 1 และระดับผลการเรียน 3 จากการสอบปลายภาคเรียนที่ 1 จำแนกตามแบบทดสอบแต่ละฉบับ ... 118

9 เปรียบเทียบคะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของแวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้ 127

10 เปรียบเทียบคะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของแวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีในกลุ่มนักเรียนที่มีความสามารถในระดับดี 128

11 เปรียบเทียบคะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของแวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐาน โดยวิธีการใช้กลุ่มคาบเส้นในกลุ่มนักเรียนที่อยู่ในระดับพอใช้ 130

ตาราง หน้า

12 เปรียบเทียบคะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของแวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นในกลุ่มนักเรียนที่อยู่ในระดับดี 131

13 เปรียบเทียบคะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของแวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีของนักเรียนที่มีความสามารถในระดับพอใช้กับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้ 133

14 เปรียบเทียบคะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของแวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีของนักเรียนที่มีความสามารถในระดับดีกับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดี 135

15 เปรียบเทียบคะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีของนักเรียนที่มีความสามารถในระดับพอใช้กับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้ 137

16 เปรียบเทียบคะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคนจากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีของนักเรียนที่มีความสามารถในระดับดีกับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดี 139

ตาราง หน้า

17 เปรียบเทียบจำนวนผู้รอบรู้จากการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของแวน เดอร์ ลินเดนและเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้ 141

18 เปรียบเทียบจำนวนผู้รอบรู้จากการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของแวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีในกลุ่มนักเรียนที่มีความสามารถในระดับดี 143

19 เปรียบเทียบจำนวนผู้รอบรู้จากการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของแวน เดอร์ ลินเดน และเคนจากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นสำหรับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้ 145

20 เปรียบเทียบจำนวนผู้รอบรู้จากการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของแวน เดอร์ ลินเดน และเคนจากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นสำหรับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดี 147

21 เปรียบเทียบจำนวนผู้รอบรู้ที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของแวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีของนักเรียนที่มีความสามารถในระดับพอใช้กับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้ 149

ตาราง หน้า

22 เปรียบเทียบจำนวนผู้รอบรู้ที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของแวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีของนักเรียนที่มีความสามารถในระดับดีกับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดี 151

23 เปรียบเทียบจำนวนผู้รอบรู้ที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคนจากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีของนักเรียนที่มีความสามารถในระดับพอใช้กับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้ 153

24 เปรียบเทียบจำนวนผู้รอบรู้ที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคนจากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีของนักเรียนที่มีความสามารถในระดับดีกับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดี 155

25 เปรียบเทียบความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของแวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้ 158

26 เปรียบเทียบความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของแวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีในกลุ่มนักเรียนที่มีความสามารถในระดับดี 159

ตาราง หน้า

27 เปรียบเทียบความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของแวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นของนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้ 161

28 เปรียบเทียบความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของแวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นของนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดี 162

29 เปรียบเทียบความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของแวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีของนักเรียนที่มีความสามารถในระดับพอใช้กับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้ 164

30 เปรียบเทียบความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของแวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีของนักเรียนที่มีความสามารถในระดับดีกับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดี 166

31 เปรียบเทียบความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีของนักเรียนที่มีความสามารถในระดับพอใช้กับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้ .. 168

ตาราง หน้า

32 เปรียบเทียบความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคนจากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีของนักเรียนที่มีความสามารถในระดับดีกับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดี 170

33 ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของแวน เดอร์ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้ในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 กับกลุ่มทักษะ - ภาษาไทย 173

34 ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของแวน เดอร์ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้ในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 กับกลุ่มทักษะ - คณิตศาสตร์ 175

35 ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของแวน เดอร์ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้ในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 และแบบทดสอบวัดการคิดและแก้ปัญหากับกลุ่มสร้างเสริมประสบการณ์ชีวิต 177

36 ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของแวน เดอร์ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีในกลุ่มนักเรียนที่มีความสามารถในระดับดีในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 กับกลุ่มทักษะ - ภาษาไทย 179

ตาราง หน้า

37 ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของแวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีในกลุ่มนักเรียนที่มีความสามารถในระดับดี ในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 กับกลุ่มทักษะ - คณิตศาสตร์ 180
38 ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของแวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีในกลุ่มนักเรียนที่มีความสามารถในระดับดีในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 และแบบทดสอบวัดการคิดและแก้ปัญหากับกลุ่มสร้างเสริมประสบการณ์ชีวิต 181
39 ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของแวน เดอร์ ลินเดน และเคนจากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นในกลุ่มนักเรียนคาบเส้นที่อยู่ในระดับพอใช้ในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 กับกลุ่มทักษะ - ภาษาไทย 183
40 ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของแวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นในกลุ่มนักเรียนคาบเส้นที่อยู่ในระดับพอใช้ ในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 กับกลุ่มทักษะ - คณิตศาสตร์ 185
41 ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของแวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นในกลุ่มนักเรียนคาบเส้นที่อยู่ในระดับพอใช้ ในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 และแบบทดสอบวัดการคิดและแก้ปัญหากับกลุ่มสร้างเสริมประสบการณ์ชีวิต 187

ตาราง หน้า

42 ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของแวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นในกลุ่มนักเรียนคาบเส้นที่อยู่ในระดับดี ในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 กับกลุ่มทักษะ - ภาษาไทย 189

43 ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของแวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นในกลุ่มนักเรียนคาบเส้นที่อยู่ในระดับดี ในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 กับกลุ่มทักษะ - คณิตศาสตร์ 190

44 ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ ตามวิธีการของ แวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นในกลุ่มนักเรียนคาบเส้นที่อยู่ในระดับดีในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 และแบบทดสอบวัดการคิดและแก้ปัญหากับกลุ่มสร้างเสริมประสบการณ์ชีวิต 191

45 ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของแวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีกับวิธีการใช้กลุ่มคาบเส้นในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้ในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 กับกลุ่มทักษะ - ภาษาไทย 193

46 ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของแวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีกับวิธีการใช้กลุ่มคาบเส้นในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้ในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 กับกลุ่มทักษะ - คณิตศาสตร์ 195

ตาราง หน้า

47 ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของแวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีกับวิธีการใช้กลุ่มคาบเส้นในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้ในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 และแบบทดสอบวัดการคิดและแก้ปัญหากับกลุ่มสร้างเสริมประสบการณ์ชีวิต 197

48 ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของแวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีกับวิธีการใช้กลุ่มคาบเส้นในกลุ่มนักเรียนที่มีความสามารถในระดับดีในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 กับกลุ่มทักษะ - ภาษาไทย 199

49 ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีกับวิธีการใช้กลุ่มคาบเส้นในกลุ่มนักเรียนที่มีความสามารถในระดับดี ในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 กับกลุ่มทักษะ - คณิตศาสตร์ ... 201

50 ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของแวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีกับวิธีการใช้กลุ่มคาบเส้นในกลุ่มนักเรียนที่มีความสามารถในระดับดีในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 และแบบทดสอบวัดการคิดและแก้ปัญหา กับกลุ่มสร้างเสริมประสบการณ์ชีวิต 203

51 ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของเคนจากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีกับวิธีการใช้กลุ่มคาบเส้นในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้ในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 กับกลุ่มทักษะ - ภาษาไทย 205

ตาราง หน้า

52 ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของเคนจาก การกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี กับวิธีการใช้ กลุ่มคาบเส้นในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้ ในแบบ ทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 กับกลุ่มทักษะ - คณิตศาสตร์ 207

53 ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของเคน จาก การกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีกับวิธีการใช้ กลุ่มคาบเส้นในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้ในแบบ ทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 และแบบทดสอบวัดการคิด และแก้ปัญหากับกลุ่มสร้างเสริมประสบการณ์ชีวิต 209

54 ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของเคนจาก การกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีกับวิธีการใช้กลุ่ม คาบเส้นในกลุ่มนักเรียนที่มีความสามารถในระดับดี ในแบบทดสอบ วัดความสามารถทั่วไป ฉบับที่ 1 กับกลุ่มทักษะ - ภาษาไทย 211

55 ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีกับวิธีการ ใช้กลุ่มคาบเส้นในกลุ่มนักเรียนที่มีความสามารถในระดับดีในแบบ ทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 กับกลุ่มทักษะ - คณิตศาสตร์ 213

56 ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีกับวิธีการ ใช้กลุ่มคาบเส้นในกลุ่มนักเรียนที่มีความสามารถในระดับดีในแบบ ทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 และแบบทดสอบวัดการ คิดและแก้ปัญหากับกลุ่มสร้างเสริมประสบการณ์ชีวิต 215

ตาราง หน้า

57 เปรียบเทียบความสอดคล้องของระดับผลการเรียนในกลุ่มสร้างเสริมลักษณะนิสัยกับแบบวัดคุณลักษณะทั่วไปของนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้ 217

58 เปรียบเทียบความสอดคล้องของระดับผลการเรียนในกลุ่มสร้างเสริมลักษณะนิสัยกับแบบวัดคุณลักษณะทั่วไปของนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดี 218

บัญชีภาพประกอบ

ภาพประกอบ หน้า

1 การจัดประเภทของผู้สอบเป็นผู้ได้รับการสอนและผู้ไม่ได้รับการสอน โดยอาศัยคะแนนสอบ 27

2 ความสัมพันธ์ระหว่างการตัดสินด้านแบบทดสอบอิงเกณฑ์และเกณฑ์ภายนอก 46

3 ความสัมพันธ์ระหว่างผลจากแบบทดสอบอิงเกณฑ์และการวัดคุณค่า 51

4 ความสัมพันธ์ระหว่างคะแนนจากแบบทดสอบอิงเกณฑ์และผลที่ได้รับ 52

5 แบบจำลองของฮอฟสที 54

6 แบบจำลองของบุค 56

บทที่ 1

บทนำ

ความเป็นมาและความสำคัญของปัญหา

การวัดผลทางการศึกษาหรือการวัดผลการเรียนรู้นั้น ตัวแปรที่นักวัดผลการศึกษามุ่งให้ความสนใจหรือมุ่งวัดก็คือ ผลสัมฤทธิ์ทางการเรียน ซึ่งการวัดผลสัมฤทธิ์ทางการเรียนนั้นเป็นกระบวนการที่ใช้ในการตัดสินผลการเรียนของผู้เรียนอย่างมีระบบโดยอยู่ในรูปของการให้คะแนนหลังจากที่ได้มีการจัดดำเนินการเรียนการสอนเสร็จสิ้นไปแล้ว (Gronlund. 1982 : 1) และคะแนนดังกล่าวจะมีความหมายก็ต่อเมื่อนำคะแนนนั้นมาทำการตัดสินใจโดยอาศัยเกณฑ์อย่างใดอย่างหนึ่งที่ผู้สอนกำหนด ในอดีตการวัดผลการศึกษาจะอิงค่าร้อยละของความรู้ที่ได้จากการสุ่มข้อคำถามจากความรู้ทั้งหมดมาสอบ โดยการกำหนดเกณฑ์ไว้ เช่น ถ้าผู้สอบสามารถตอบข้อคำถามในแบบทดสอบได้ถูกต้องตั้งแต่ร้อยละ 50 ของข้อคำถามทั้งหมดในแบบทดสอบจึงจะถือว่าผู้สอบมีความรอบรู้ (Master) แต่ถ้าผู้สอบไม่สามารถตอบข้อคำถามในแบบทดสอบได้ถูกต้องตั้งแต่ร้อยละ 50 ของข้อคำถามทั้งหมดในแบบทดสอบก็จะถือว่าผู้สอบไม่มีความรอบรู้ (Non - Master) การกำหนดเช่นนี้ก็เป็นการกำหนดเกณฑ์ชนิดหนึ่งนั่นคือกำหนดที่คะแนนร้อยละ 50 ของคะแนนเต็มเป็นเกณฑ์ ส่วนการประเมินผลหรือการตัดสินผลการเรียนรู้เพื่อใช้ในการพิจารณาว่าจะให้ผู้สอบผ่านขึ้นไปเรียนในชั้นที่สูงขึ้นไปได้หรือไม่ ก็จะพิจารณาจากการนำคะแนนที่ได้จากการสอบแต่ละรายวิชามารวมกันแล้วแปลงคะแนนทั้งหมดเป็นร้อยละ ผู้ที่ได้คะแนนจากทุกรายวิชารวมกันแล้วได้คะแนนตั้งแต่ ร้อยละ 50 ขึ้นไป ก็จะตัดสินให้ผ่านสามารถขึ้นไปเรียนในระดับชั้นที่สูงขึ้นไปได้ ส่วนผู้ที่ได้คะแนนจากทุกรายวิชารวมกันแล้วไม่ถึงร้อยละ 50 ก็จะตัดสินให้ไม่ผ่านการสอบและจะต้องเรียนซ้ำชั้น เป็นต้น การกำหนดคะแนนการสอบผ่านเป็นร้อยละ 50 ดังกล่าวนั้นเป็นการกำหนดที่ไม่มีฐานที่แน่นอน กล่าวคือ ไม่ทราบว่าเป็นร้อยละ 50 ของอะไร ทั้งนี้ขึ้นอยู่กับการออกข้อสอบ การตรวจข้อสอบหรือการให้คะแนนของผู้ออกข้อสอบ

ในระยะเวลาต่อมานักวัดผลการศึกษาได้มีแนวคิดในการวัดและการประเมินผลการศึกษาที่แตกต่างไปจากเดิม คือการวัดและการประเมินผลการศึกษาเป็นรายวิชาโดยให้คะแนนในแต่ละรายวิชาออกมาในรูปของระดับคะแนน คะแนนที่ได้จากการสอบของผู้สอบจะนำไปเปรียบเทียบ

กับคะแนนของผู้สอบคนอื่น ๆ ที่สอบในรายวิชาเดียวกัน การวัดและการประเมินผลการศึกษาแบบนี้จึงเป็นการวัดที่ขึ้นอยู่กับพฤติกรรมของกลุ่มเป็นสำคัญ โดยมีความเชื่อในเรื่องความแตกต่างระหว่างบุคคลที่ว่าแต่ละบุคคลมีความสามารถไม่เท่ากัน ดังนั้นการวัดและการประเมินผลการศึกษาแบบนี้จึงเป็นการหาความแตกต่างระหว่างบุคคลว่าแต่ละบุคคลอยู่ตรงไหนของกลุ่ม และคะแนนที่ได้จะมีความหมายก็ต่อเมื่อนำไปเปรียบเทียบกับคะแนนในกลุ่มที่สอบด้วยแบบทดสอบฉบับเดียวกัน โดยนำคะแนนที่ได้มาตัดสินในรูปตำแหน่งเปอร์เซ็นไทล์หรือคะแนนมาตรฐานอื่น ๆ จะเห็นว่าการวัดและการประเมินผลการศึกษาแบบนี้เป็นการจัดอันดับความสามารถของผู้สอบว่าผู้สอบมีความรู้ในเรื่องที่เรียนมากหรือน้อยกว่าคนอื่นในกลุ่ม การแปลความหมายของคะแนนในลักษณะนี้เป็นการแปลความหมายของ คะแนนแบบอิงกลุ่ม (Norm - Referenced) ซึ่งมีจุดอ่อนคือผลที่ได้นั้นไม่สามารถบ่งบอกถึงระดับผลสัมฤทธิ์ของผู้สอบได้ว่า ผู้สอบได้เกิดพฤติกรรมตามจุดประสงค์ที่กำหนดไว้แล้วหรือไม่ เนื่องจากการวัดและการประเมินผลแบบอิงกลุ่มเป็นการมุ่งเปรียบเทียบความสามารถของบุคคลเป็นสำคัญจึงทำให้การศึกษากลายเป็นกระบวนการจำแนกคนเป็นกลุ่มเป็นระดับ ซึ่งโดยหน้าที่ที่แท้จริงแล้ว การศึกษาเป็นการพัฒนาคน เพราะฉะนั้นจุดสำคัญของการประเมินจึงควรอยู่ที่เกณฑ์การเรียนรู้เป็นสำคัญ คือมุ่งที่ประเมินว่าก่อนที่จะทำการสอนผู้เรียนย่อมไม่เกิดการเรียนรู้แต่เมื่อได้ผ่านกระบวนการเรียนการสอนไปแล้วผู้เรียนย่อมเกิดการเรียนรู้ตามจุดประสงค์ที่กำหนดไว้ (ไกวิท ประวาลพฤกษ์ และสมศักดิ์ สินธุระเวชญ์. 2523 : 11) ดังนั้นการเรียนการสอนที่เรียกว่าการเรียนเพื่อรอบรู้ (Mastery Learning) จึงเกิดขึ้น การพัฒนาทางการวัดผลจึงดำเนินไปสู่การวัดและการประเมินผลแบบอิงเกณฑ์ (Criterion - Referenced)

การวัดและการประเมินผลแบบอิงเกณฑ์ เป็นการวัดที่ไม่ได้นำผลของแต่ละบุคคลไปเปรียบเทียบกับคะแนนของคนอื่นในกลุ่ม แต่เป็นการวัดที่มุ่งไปในลักษณะที่ว่าผู้เรียนสามารถทำอะไรได้บ้างหรือไม่สามารถทำอะไรได้บ้าง หลังจากที่มีการเรียนการสอนไปแล้วโดยอาศัยเกณฑ์ ซึ่งเกณฑ์ในที่นี้ได้แก่จุดมุ่งหมายของการเรียนการสอน ระดับความสามารถที่พึงปรารถนาของผู้เรียน พฤติกรรมหลังการสอน เป็นต้น จะเห็นว่า การวัดและการประเมินผลแบบนี้เป็นการวัดและการประเมินผลเพื่อแยกผู้เรียนที่รู้แล้วและยังไม่รู้ โดยมีวัตถุประสงค์เพื่อวินิจฉัยว่า ผู้เรียนบรรลุวัตถุประสงค์ที่วางไว้หรือไม่ ดังนั้นในการวัดและการประเมินผลแบบอิงเกณฑ์นั้น สิ่งที่จะต้อง

กำหนดเพื่อใช้ในการพิจารณาตัดสินว่าจะจัดผู้เรียนเป็นผู้มีความรอบรู้หรือไม่ก็คือ การนำคะแนนจุดตัด (Cutting Score) หรือมาตรฐาน (Standard) เข้ามาเกี่ยวข้อง ซึ่งในเรื่องนี้ ลินน์ (Linn. 1978 : 302) กล่าวว่า การกำหนดคะแนนจุดตัด หรือมาตรฐานนั้นเป็นสิ่งสำคัญและเป็นปัญหากับนักวัดผลการศึกษามากโดยเฉพาะอย่างยิ่งเกี่ยวกับการวัดผลแบบอิงเกณฑ์และปัญหาในการกำหนดคะแนนจุดตัดก็ได้รับความสนใจน้อยมากด้วย ส่วนใหญ่จะมุ่งเน้นไปให้ความสนใจเกี่ยวกับการประมาณค่าความเชื่อมั่น (Reliability)ของแบบทดสอบและการพิจารณาถึงความยาวของแบบทดสอบ (Test Length) นอกจากนั้นมีงานวิจัยส่วนน้อยเท่านั้นที่สนใจเกี่ยวกับเรื่องคะแนนจุดตัดหรือมาตรฐานว่ามาจากไหนและควรใช้กระบวนการใดในการกำหนด

การกำหนดคะแนนจุดตัดได้มีการพัฒนามาตั้งแต่ปี ค.ศ. 1970 (Berk. 1986 : 137) ซึ่งคะแนนจุดตัดจะเป็นเกณฑ์หรือมาตรฐานที่ใช้จำแนกระดับความสามารถของผู้สอบออกเป็นผู้รอบรู้ และผู้ไม่รอบรู้ เพราะฉะนั้นการกำหนดคะแนนจุดตัดให้เหมาะสมจึงเป็นสิ่งจำเป็น เนื่องจากคะแนนจุดตัดจะมีผลต่อค่าความเชื่อมั่นและความเที่ยงตรงของแบบทดสอบ ยิ่งกว่านั้นคุณค่าของคะแนนจุดตัดยังมีอิทธิพลต่อการเรียนรู้ตลอดจนทัศนคติของผู้เรียนอีกด้วย (Hambleton, Swaminathan, Algina and Coulson. 1978 : 27) เพราะฉะนั้นในการกำหนดคะแนนจุดตัดหากกำหนดสูงเกินไปก็จะทำให้เกิดความคลาดเคลื่อน คือผู้ที่มีความรอบรู้จะกลายเป็นผู้ไม่มีความรอบรู้หรือถ้ากำหนดไว้ต่ำเกินไปก็จะทำให้เกิดความคลาดเคลื่อนขึ้นเช่นเดียวกัน คือผู้ที่ไม่มีความรอบรู้จะกลายเป็นผู้มีความรอบรู้ ดังนั้นความเที่ยงตรงของการกำหนดคะแนนจุดตัดจึงขึ้นอยู่กับความถูกต้องในการแยกผู้สอบเป็นผู้รอบรู้และผู้ไม่รอบรู้ (Hambleton. 1978 : 278) อย่างไรก็ตามในการกำหนดคะแนนจุดตัดนั้นก็คงจะไม่สามารถกำหนดเกณฑ์หรือมาตรฐานที่สามารถขจัดความคลาดเคลื่อนดังกล่าวให้หมดไปได้ (Zieky. 1987 : 2) เนื่องจากในการวัดผลการศึกษานั้นเป็นการวัดในสิ่งที่เป็นนามธรรมซึ่งจะต้องทำการวัดทางอ้อมโดยใช้แบบทดสอบเป็นเครื่องมือสำคัญในการเก็บรวบรวมข้อมูลที่ได้จากการวัดพฤติกรรมต่าง ๆ ตามที่ผู้สอนต้องการจะวัด ดังนั้นคะแนนที่ได้จากการสอบก็จะเป็นแต่เพียงคะแนนที่ได้จากการสุ่มข้อคำถามจากความรู้ทั้งหมดมาให้ตอบภายในช่วงระยะเวลาและเงื่อนไขที่กำหนดให้เท่านั้น นอกจากนั้นการกำหนดคะแนนจุดตัดเป็นวิธีการที่เกี่ยวข้องกับการตัดสินใจของคน เพราะฉะนั้นคะแนนจุดตัดหรือมาตรฐานที่กำหนดขึ้นจึง

เป็นแต่เพียงสิ่งที่สมมุติ (Arbitrary) ขึ้นเท่านั้น อาจจะมีความคลาดเคลื่อนหรือความไม่แน่นอนเกิดขึ้นได้ ซึ่งในเรื่องนี้ ลินน์ (Linn. 1978 : 307) ให้ข้อเสนอแนะว่าหากใช้กระบวนการที่มีระบบก็จะสามารถช่วยลดโอกาสที่จะทำให้เกิดความคลาดเคลื่อนขึ้นได้ หรือทำให้โอกาสที่จะเกิดความคลาดเคลื่อนให้เหลือน้อยที่สุดได้

วิธีการในการกำหนดคะแนนจุดตัดมีอยู่ด้วยกันหลายวิธี เช่น วิธีการนับลดจาก 100 % (Counting Backwards from 100%) ซึ่งกลาส (Glass. 1978 : 244) เสนอไว้โดยให้ผู้เชี่ยวชาญหรือครูผู้สอนเป็นผู้พิจารณาข้อสอบและจุดประสงค์เชิงพฤติกรรมว่า จุดประสงค์เชิงพฤติกรรมแต่ละข้อมีความสำคัญมากน้อยเพียงใด แล้วพิจารณากำหนดคะแนนจุดตัดของแบบทดสอบจากเกณฑ์ที่ตั้งไว้ ซึ่งแล้วแต่ว่าผู้ตั้งเกณฑ์จะลดลงมากี่เปอร์เซ็นต์วิธีการนี้เป็นการกำหนดเกณฑ์ที่ค่อนข้างจะอัตนัย (Subjective) เนื่องจากใช้ความคิดเห็นส่วนบุคคลในการพิจารณา ต่อมาจึงมีผู้พัฒนาวิธีการกำหนดคะแนนจุดตัดโดยการพิจารณาเนื้อหาของแบบทดสอบเป็นรายข้อ แล้วระบุว่าผู้สอบที่มีความสามารถหรือสมรรถภาพขั้นต่ำสุดที่จะยอมรับว่าผ่านได้นั้นควรทำแบบทดสอบได้เป็นจำนวนเท่าไร ก็จะทำให้ได้คะแนนจุดตัดที่เหมาะสมขึ้น ซึ่งวิธีการดังกล่าวนี้คือวิธีการพิจารณาตัดสินจากความสามารถต่ำสุดของผู้สอบ (Judging Minimal Competence) ตามที่กลาส (Glass. 1978 : 246 - 251) ได้เสนอไว้หรือเป็นวิธีการพิจารณาจากข้อคำถามในแบบทดสอบตามที่ลิฟวิงสตัน และซีกี (Livingston and Zieky. 1982 : 15 - 52) ได้เสนอไว้ วิธีการนี้มีหลายเทคนิคด้วยกัน คือเทคนิคของแองกอฟฟ์ (Angoff) เทคนิคของอีเบล (Ebel) และเทคนิคของนีเดลสกี (Nedelsky) ทั้ง 3 เทคนิคต่างก็มีวิธีการที่เหมือนกันคือ ผู้เชี่ยวชาญจะต้องพิจารณาถึงความรู้และทักษะของผู้สอบที่อยู่ในระดับต่ำสุดหรือที่เรียกว่า ระดับคาบเส้น (Borderline) ว่าสมควรให้ผ่านหรือไม่ ซึ่งเป็นเทคนิคที่สะดวกแก่ผู้ใช้มาก เนื่องจากสามารถกำหนดคะแนนจุดตัดได้ทั้งก่อนการสอบและหลังจากการสอบเสร็จสิ้นแล้ว โดยที่ผู้เชี่ยวชาญจะต้องพิจารณาตัดสินข้อคำถามในแบบทดสอบร่วมกัน ดังนั้น ขั้นตอนต่าง ๆ ในการนำเทคนิคทั้ง 3 นี้ไปใช้จะคล้ายคลึงกัน กล่าวคือจะเริ่มจาก1) การเลือกผู้เชี่ยวชาญ 2) การกำหนดความรู้และทักษะของผู้สอบที่อยู่ในระดับต่ำสุดที่จะยอมให้ผ่าน (Zieky. 1987 : 3) ต่อจากนั้นจึงให้ผู้เชี่ยวชาญพิจารณาตัดสินแล้วนำผลการตัดสินใจของผู้เชี่ยวชาญที่ได้มากำหนดเป็นคะแนนจุดตัด

เทคนิคของแองกอฟฟ์เป็นเทคนิคที่ง่ายและใช้ได้รวดเร็ว (Livingston and Zieky. 1982 : 54) หากกลุ่มของผู้เชี่ยวชาญพิจารณากำหนดคะแนนจุดตัดมีความเชี่ยวชาญในเนื้อหาที่จะนำมาใช้ในการเขียนข้อสอบ (Norcini, Shea and Kanya. 1988 : 57) เนื่องจากเทคนิคนี้ผู้เชี่ยวชาญจะต้องพิจารณาถึงโอกาสของผู้สอบที่มีความสามารถต่ำสุดหรือผู้ที่อยู่ในระดับคาบเส้นที่จะตอบข้อสอบในแต่ละข้อคำถามได้ถูกต้อง เพราะฉะนั้นก่อนอื่นผู้เชี่ยวชาญจะต้องให้คำนิยามหรือกำหนดความสามารถต่ำสุดของผู้สอบก่อน โดยที่ผู้เชี่ยวชาญแต่ละคนจะต้องคาดคะเนความยากในแต่ละข้อคำถามของแบบทดสอบ แล้วนำผลที่ได้จากการคาดคะเนของผู้เชี่ยวชาญมาหาค่าเฉลี่ย ค่าที่ได้ก็จะเป็นคะแนนจุดตัด (Mills. 1983 : 284)

ส่วนเทคนิคของอีเบล เป็นเทคนิคที่ผู้เชี่ยวชาญจะต้องพิจารณาถึงความยากและความสำคัญของข้อคำถามประกอบกัน ซึ่งเทคนิคนี้จะทำให้ผู้เชี่ยวชาญต้องใช้เวลาในการพิจารณามาก และคงจะไม่เหมาะสมกับแบบทดสอบที่ตอบสั้น (Zieky. 1987 : 16) นอกจากนี้เทคนิคของอีเบลไม่เหมาะกับแบบทดสอบที่มีความยาวมาก ๆ เนื่องจากผู้เชี่ยวชาญจะต้องเสียเวลาในการจัดประเภทของแบบทดสอบไปตามค่าความยากและความสำคัญ (Cross, Impara, Frary and Jaeger. 1984 : 114)

สำหรับเทคนิคของนีเดลสกี เป็นเทคนิคที่เหมาะสมกับแบบทดสอบแบบเลือกตอบเท่านั้น (Zieky. 1987 : 11) การกำหนดคะแนนจุดตัดโดยใช้เทคนิคนี้จะต้องพิจารณาถึงความยากของข้อคำถามเป็นรายตัวเลือกว่า ตัวเลือกใดที่ผู้สอบที่มีความสามารถต่ำสุดหรือผู้สอบที่อยู่ในระดับคาบเส้นสามารถทราบได้ว่าเป็นตัวเลือกที่ผิด (Plake and Milican. 1989 : 46 ; Cross, Impara, Frary and Jaeger. 1984 : 114) ซึ่งโดยความเป็นจริงแล้วผู้เชี่ยวชาญก็จะไม่มีโอกาสรู้ได้ (Zieky. 1987 : 16) ดังนั้นจะเห็นได้ว่าวิธีการกำหนดคะแนนจุดตัดโดยใช้วิธีการพิจารณาตัดสินจากความสามารถต่ำสุดของผู้สอบทั้ง 3 เทคนิค คือเทคนิคของแองกอฟฟ์ อีเบล และนีเดลสกี ดังกล่าวนั้น ในทางปฏิบัติหากเลือกเทคนิคดังกล่าวมาใช้จะต้องใช้เวลามากสำหรับให้ผู้เชี่ยวชาญในการพิจารณา โดยเฉพาะอย่างยิ่ง ถ้ามีจำนวนข้อสอบมากก็อาจจะทำให้ผู้เชี่ยวชาญเกิดความเบื่อหน่ายได้ (Norcini, Lipner, Langdon and Streker. 1987 : 56)

จากการเปรียบเทียบโดยใช้เทคนิคของแองกอฟฟ์ อีเบล และนีเดลสกี พบว่า การกำหนดคะแนนจุดตัดโดยใช้เทคนิคของแองกอฟฟ์และนีเดลสกีจะให้ผลแตกต่างกัน และยังพบว่าไม่เพียงแต่คะแนนจุดตัดที่ได้จากการกำหนดโดยเทคนิคของแองกอฟฟ์และนีเดลสกีแตกต่างกันเท่านั้น ยังพบว่ากลุ่มของผู้เชี่ยวชาญที่ใช้วิธีการที่แตกต่างกันจะได้คะแนนจุดตัดที่แตกต่างกัน และในระหว่างกลุ่มของผู้เชี่ยวชาญที่ใช้วิธีเดียวกันก็จะได้คะแนนจุดตัดที่แตกต่างกันด้วย (Halpin, Sigmon and Halpin. 1983 : 186 ; citing Behuniak, Gable and Archambault. 1982.) นอกจากนี้ แอนดรูว์และเฮชท์ (Behuniak, Archambault and Gable. 1982 : 247 ; citing Andrew and Hecht. 1976.) พบว่าการกำหนดคะแนนจุดตัดโดยใช้เทคนิคของอีเบลจะได้คะแนนจุดตัดที่สูงกว่าเทคนิคของนีเดลสกี ส่วนเบรนแนนและล็อควูด (Behuniak, Archambault and Gable. 1982 : 247 ; citing Brennan and Lockwood. 1979) สรุปว่าเทคนิคของแองกอฟฟ์จะให้คะแนนจุดตัดที่สูงกว่าและมีความแปรผันระหว่างผู้ตัดสินน้อยกว่าเทคนิคของนีเดลสกี

อย่างไรก็ตามในการกำหนดคะแนนจุดตัดทั้ง 3 เทคนิคอันได้แก่ เทคนิคของแองกอฟฟ์ อีเบลและนีเดลสกี จะเห็นว่าต่างก็เกี่ยวข้องกับค่าความยากของข้อคำถาม แต่การประมาณค่าความยากของข้อคำถามโดยใช้เทคนิคของอีเบล และนีเดลสกีก็ยังมีความชัดเจนน้อยกว่าการประมาณค่าความยากโดยใช้เทคนิคของแองกอฟฟ์ (Halpin, Sigmon and Halpin. 1983 : 188) นอกจากนี้ แวน เดอร์ ลินเดน (Berk. 1986 : 147 ; citing Van der Linden. 1984 : 15) ชี้ให้เห็นว่าเทคนิคของแองกอฟฟ์และนีเดลสกียังเป็นเทคนิคที่ไม่เหมาะสมเนื่องจากความไม่คงที่ในการระบุค่าความน่าจะเป็นของผู้ที่รอบรู้ซึ่งอาจเนื่องมาจากการเข้าใจคำว่า "การรอบรู้" แตกต่างกัน และการแปลความหมายของวัตถุประสงค์ของการเรียนแตกต่างกัน หากกล่าวโดยสรุปเมื่อเปรียบเทียบทั้ง 3 เทคนิคนี้แล้ว เบอร์ก (Berk. 1986 : 147) สรุปว่าเทคนิคของแองกอฟฟ์เป็นเทคนิคที่ดีที่สุด แต่ก็ยังไม่เหมาะกับข้อสอบแบบเลือกตอบ (Livingston and Zieky. 1982 : 24)

วิธีการกำหนดคะแนนจุดตัดโดยอาศัยความสามารถต่ำสุดของผู้สอบทั้ง 3 เทคนิค อันได้แก่ เทคนิคของแองกอฟฟ์ อีเบล และนีเดลสกีนั้น ในการที่จะนำไปใช้นั้นเทคนิคทั้ง 3 เหมาะกับ

ผู้สอบที่เป็นคนกลุ่มใหญ่ และมีความสามารถแตกต่างกันมาก ๆ เนื่องจากการใช้เทคนิคทั้ง 3 นี้ ผู้เชี่ยวชาญเพียงแต่พิจารณาจากข้อคำถามในแบบทดสอบเท่านั้น (Livingston and Zieky. 1982 : 51) เทคนิคทั้ง 3 ดังกล่าวนี้ในทางทฤษฎีเป็นสิ่งที่ดี แต่ในทางปฏิบัติยังมีปัญหาอยู่ 2 ประการด้วยกันคือ 1) ผู้เชี่ยวชาญแต่ละคนมีความคิดที่แตกต่างกันเกี่ยวกับความสามารถต่ำสุดของผู้สอบ เพราะฉะนั้นคะแนนจุดตัดที่ผู้เชี่ยวชาญกำหนดจะสะท้อนถึงความแตกต่างของผู้สอบด้วย 2) การประมาณความสามารถต่ำสุดของผู้สอบจากการพิจารณาข้อสอบจำนวนมาก เช่น จำนวน 50 ข้อ นับว่าเป็นงานที่ยาก เพราะว่าผู้สอบบางคนจะได้คะแนนสูงในบางข้อและบางคนจะได้คะแนนต่ำสุดในบางข้อ เพราะฉะนั้นการคาดคะเนจากความสามารถต่ำสุดจึงเป็นสิ่งที่เป็นไปไม่ได้ แต่ถ้าผู้สอบที่อยู่ในระดับความสามารถต่ำสุดอยู่ในระดับเดียวกัน การประมาณความสามารถต่ำสุดของผู้สอบ จากการทำแบบทดสอบทั้งฉบับจะทำได้ง่ายกว่าการพิจารณาแต่ละข้อคำถาม (Rowley. 1982 : 89 - 90) ดังนั้น โรว์เลย์ (Rowley. 1982 : 90) และซีกี (Zieky. 1987 : 12) จึงเสนอวิธีกำหนดคะแนนจุดตัดวิธีใหม่ คือการใช้วิธีกลุ่มตรงกันข้าม (Contrasting Groups Method) ซึ่งเป็นวิธีการตัดสินที่ขึ้นอยู่กับผู้สอบมากกว่าการตัดสินเกี่ยวกับข้อคำถามในแบบทดสอบวิธีการนี้มีข้อดีคือผู้ตัดสินหรือผู้เชี่ยวชาญ จะคุ้นเคยกับการพิจารณาทักษะของผู้เรียนและสามารถพิจารณาได้อย่างมีความหมาย นอกจากนี้ข้อสนเทศที่ได้จากวิธีการนี้จะสอดคล้องกับสภาพความเป็นจริงมากกว่าการพิจารณาจากข้อคำถามในแบบทดสอบ การใช้วิธีกลุ่มตรงกันข้ามในการกำหนดคะแนนจุดตัดนั้นมีหลายวิธี วิธีที่ซับซ้อนน้อยที่สุด คือการเขียนกราฟแสดงการแจกแจงของคะแนนของกลุ่มรอบรู้และกลุ่มไม่รอบรู้ แล้วพิจารณาว่าเส้นกราฟของผู้สอบทั้ง 2 กลุ่มตัดกัน (Intersection) ณ จุดใด จุดนั้นก็คือคะแนนจุดตัด (Mills. 1983 : 285)

ในการใช้วิธีกลุ่มตรงกันข้ามเพื่อกำหนดคะแนนจุดตัดนั้นปัญหาที่พบก็คือไม่สามารถแบ่งได้อย่างแน่ชัดระหว่างผู้ที่รอบรู้กับผู้ที่ไม่รอบรู้ ดังนั้น โรว์เลย์ (Rowley. 1982 : 91) และซีกี (Zieky. 1987 : 12) จึงได้เสนอวิธีการกำหนดคะแนนจุดตัดอีกวิธีหนึ่ง คือการใช้กลุ่มคาบเส้น (Borderline Group Method) ซึ่งวิธีการนี้มาจากแนวคิดที่ว่า มาตรฐานจะเป็นคะแนนที่ได้จากผู้สอบซึ่งมีความสามารถอยู่ในระดับคาบเส้น วิธีการนี้แทนที่จะให้ผู้เชี่ยวชาญเดา ถึงความสามารถของผู้สอบที่อยู่ในระดับคาบเส้นก็ใช้วิธีการให้ผู้เชี่ยวชาญระบุความรู้ความ

สามารถของผู้สอบที่อยู่ในระดับคาบเส้น และพิจารณาคะแนนจุดตัดจากผู้สอบดังกล่าว มาตรฐานที่กำหนดสำหรับการผ่านจะใช้ค่ามัธยฐาน (Median) ของคะแนนผู้สอบที่อยู่ในกลุ่มคาบเส้น ประโยชน์สำคัญของวิธีการนี้คือง่ายในการนำไปใช้และง่ายในการอธิบาย

นอกจากการกำหนดคะแนนจุดตัดโดยอาศัยความสามารถต่ำสุดของผู้สอบหรือการพิจารณาจากข้อคำถามในแบบทดสอบ และการพิจารณาจากกลุ่มผู้สอบโดยตรงแล้วก็ยังได้มีการพัฒนาโดยการนำหลักการทางสถิติเข้ามาช่วยในการพิจารณาคะแนนจุดตัดโดยใช้ทฤษฎีการตัดสินใจ (Decision - Theoretic Approaches) ซึ่งมีอยู่หลายเทคนิคด้วยกัน เช่น เทคนิคของเบส์ (Bayes) (Swaminathan, Hambleton and Algina. 1975 : 87 - 98) เทคนิคของกลาส (Glass. 1978 : 251 - 253) สำหรับเทคนิคของเบส์เป็นเทคนิคที่อาศัยคะแนนมาตรฐานซึ่งแปลงมาจากคะแนนสอบแล้วพิจารณาคะแนนจุดตัดโดยอาศัยฟังก์ชันการสูญเสียที่คาดหวัง (Expected Loss) ซึ่งเป็นค่าความสูญเสียที่เกี่ยวข้องกับความคลาดเคลื่อน เนื่องมาจากการยอมรับที่ผิด (False Positive Error) และการปฏิเสธที่ผิด (False Negative Error) ของการตัดสินใจให้ผ่านและไม่ผ่านคะแนนระดับต่าง ๆ ความสูญเสียที่คาดหวังของการตัดสินใจให้ผ่านและไม่ผ่านที่น้อยที่สุดจะเป็นตัวกำหนดคะแนนจุดตัด ส่วนเทคนิคของกลาสเป็นเทคนิคที่อาศัยความสัมพันธ์ของคะแนนจุดตัดของแบบทดสอบกับเกณฑ์ภายนอกที่สอดคล้องกันซึ่งได้มีการกำหนดไว้ก่อนจะเห็นได้ว่าวิธีการกำหนดคะแนนจุดตัดโดยใช้ทฤษฎีการตัดสินใจนี้เป็นวิธีการที่ค่อนข้างจะมีความเป็นปรนัยสูง แต่ก็นับว่าเป็นวิธีการที่ค่อนข้างจะยุ่งยากที่จะนำมาใช้กำหนดคะแนนจุดตัดในชั้นเรียน

สำหรับการกำหนดคะแนนจุดตัดของประเทศไทยได้เริ่มสนใจมาตั้งแต่ปี พ.ศ. 2521 เป็นต้นมา ส่วนใหญ่จะเป็นการสร้างแบบทดสอบอิงเกณฑ์แล้วใช้วิธีการในการกำหนดคะแนนจุดตัดด้วยวิธีการต่าง ๆ เช่น งานวิจัยของ กาญจนา วัธนสุนทร (2521) ใช้วิธีการนับลดจาก 100 % งานวิจัยของ ชมภู จันทรอมรพร (2522) จำเนียร สุขหลาย (2524) ใช้ทฤษฎีการตัดสินใจของเบส์ งานวิจัยของ ไพฑูรย์ เวทการ (2524) เกศริน บุญเกิด (2526) มัทนี เผื่อนน้อย (2526) ใช้ทฤษฎีการตัดสินใจของกลาส งานวิจัยของ รังสรรค์ มณีเล็ก (2527) ใช้วิธีการของนีเดลสกี เบอร์ก และวิธีการของเบส์มาเปรียบเทียบกัน

ส่วนในเรื่องเกี่ยวกับการนำแนวคิดและวิธีการในการกำหนดคะแนนจุดตัดไปใช้นั้น ในต่างประเทศก็ได้มีการศึกษาในแง่มุมต่าง ๆ กันดังบทความและงานวิจัยหลายชิ้น เช่น สวามินาธาน แฮมเบิลตัน และอัลไจนา (Swaninathan, Hambleton and Algina. 1975) ใช้วิธีการของเบส์ มาเลช (Malech. 1978) ใช้ทฤษฎีไบโนเมียล วิธีการของเบส์ และวิธีการตัดสินใจโดยใช้หลักความน่าจะเป็น เบฮูเนียค (Behuniak. 1981) ใช้วิธีการของนีเดลสกี แองกอฟฟ์และทฤษฎีไบโนเมียล

จะเห็นได้ว่าวิธีการในการกำหนดคะแนนจุดตัดเท่าที่ผ่านมาเป็นวิธีการที่ไม่ได้พิจารณาถึงความสัมพันธ์ระหว่างคุณลักษณะข้อสอบรายข้อ และ/หรือ คุณลักษณะแบบทดสอบทั้งฉบับกับความสามารถของผู้สอบโดยการนำแบบจำลองทางคณิตศาสตร์มาใช้แต่เป็นการพิจารณาจากคะแนนของผู้สอบเป็นหลักโดยใช้วิธีการพิจารณาของผู้เชี่ยวชาญหรือใช้วิธีการกำหนดจากเกณฑ์ภายนอก ซึ่งคะแนนจุดตัดที่ได้อาจจะไม่สอดคล้องกับความสามารถของผู้สอบ เพราะฉะนั้นคะแนนที่นำมาใช้ในการตัดสินการรอบรู้และการไม่รอบรู้ของผู้สอบก็ยังไม่น่าจะเป็นวิธีการที่ถูกต้องนัก กล่าวคือ การกำหนดคะแนนจุดตัดโดยใช้วิธีการแบบดั้งเดิม (Traditional Approach) หรือทฤษฎีการสอบแบบคลาสสิค (Classical Test Theory) นั้น การกำหนดคะแนนผ่านของแบบทดสอบจะกำหนดจากการพิจารณาจากความสามารถต่ำสุดของผู้สอบ ซึ่งตัดสินโดยผู้เชี่ยวชาญนั้นไม่ใช่เป็นวิธีที่ให้ผลดีที่สุด (Kane. 1987 : 333) เนื่องจากการกำหนดคะแนนผ่านหรือคะแนนจุดตัดเช่นนี้จะพิจารณาแต่เพียงมิติเดียว คือมิติของผลการวัดและกำหนดคะแนนจุดตัดจากมิติของผลการวัด แล้วจึงเอาคะแนนจุดตัดที่กำหนดไปเป็นเกณฑ์มาตรฐาน เพื่อพิจารณาผู้สอบเป็นผู้ที่มีความรอบรู้กับผู้ที่ไม่มีความรอบรู้ การพิจารณาเช่นนั้น แสดงว่า ความสัมพันธ์ระหว่างผลการวัดและเกณฑ์มาตรฐานมีลักษณะเป็นความสัมพันธ์เชิงเส้นตรง แต่ในความเป็นจริงแล้วความสัมพันธ์ระหว่างผลการวัดและเกณฑ์มาตรฐานไม่ได้มีลักษณะเป็นความสัมพันธ์เชิงเส้นตรง ทั้งนี้ขึ้นอยู่กับโค้งคุณลักษณะแบบทดสอบ (Test Characteristic Curve) ดังนั้นการกำหนดคะแนนจุดตัดที่เหมาะสมน่าจะพิจารณาจากเกณฑ์มาตรฐานที่กำหนดด้วยวิธีการอย่างใดอย่างหนึ่ง แล้วแปลงคะแนนที่ได้จากเกณฑ์มาตรฐานในสเกลของคะแนนโดเมน หรือคะแนนจริงให้เป็นระดับความสามารถโดยอาศัยโค้งคุณลักษณะแบบทดสอบ ซึ่งเป็นโค้งที่ได้จากค่าเฉลี่ยของโค้งคุณลักษณะของ

ข้อสอบทุกข้อในแบบทดสอบตามวิธีการของ แฮมเบิลตัน และสวามินาธาน (Hambleton and Swaminathan. 1985 : 66)

การกำหนดคะแนนจุดตัดด้วยวิธีการที่กล่าวมานั้นเป็นการกำหนดด้วยวิธีการวิเคราะห์ที่อาศัยกรอบทฤษฎีใหม่ที่ดีกว่า คือการนำทฤษฎีการตอบสนองต่อข้อสอบ (Item Response Theory) มาประยุกต์ใช้

ทฤษฎีการตอบสนองต่อข้อสอบ เป็นทฤษฎีที่มีบทบาทมากในการวัดผลการศึกษา โดยที่ทฤษฎีนี้ตั้งอยู่บนความเชื่อพื้นฐานที่ว่าในสถานการณ์ของการสอบนั้น พฤติกรรมการตอบสนองต่อข้อสอบของบุคคลสามารถอธิบายได้จากความสัมพันธ์ระหว่างคุณลักษณะภายในตัวบุคคล (Latent Trait) ของผู้สอบ ซึ่งเป็นสิ่งที่ไม่สามารถสังเกตเห็นได้กับความสามารถในการตอบสนองต่อข้อสอบ (Performance) โดยอธิบายในรูปแบบจำลองเชิงคณิตศาสตร์ ลักษณะเด่นของทฤษฎีนี้ คือค่าพารามิเตอร์ของข้อสอบไม่ขึ้นอยู่กับลักษณะกลุ่มตัวอย่างผู้ตอบข้อสอบ

นักวัดผลการศึกษาที่ได้เสนอวิธีการนำทฤษฎีการตอบสนองต่อข้อสอบไปประยุกต์ใช้กับการกำหนดคะแนนจุดตัดโดยการกำหนดเกณฑ์มาตรฐานขึ้นก่อนคือ แวน เดอร์ ลินเดน (Van der Linden. 1982 :295 - 308) และเคน (Kane. 1987 : 333 - 345) วิธีการของ แวน เดอร์ ลินเดน เป็นวิธีการกำหนดเกณฑ์มาตรฐาน โดยการพิจารณาจากความสามารถต่ำสุดของผู้สอบก่อนด้วยเทคนิคใดเทคนิคหนึ่ง เช่น แองกอฟฟ์หรือนีเดลสกี แล้วแปลงจากสเกลของคะแนนจริงไปเป็นสเกลของความสามารถโดยใช้โค้งคุณลักษณะแบบทดสอบ ส่วนวิธีการของเคนเป็นวิธีการประมาณค่าคะแนนจุดตัดจากการตัดสินความสามารถต่ำสุดของผู้สอบ และการหาค่าความน่าจะเป็นที่ผู้สอบที่มีความสามารถต่ำสุดจะตอบแต่ละข้อคำถามได้ถูกต้องจากโค้งคุณลักษณะข้อสอบ โดยการนำความแตกต่างในการตัดสินของผู้เชี่ยวชาญมาพิจารณาด้วย

การกำหนดคะแนนจุดตัดเพื่อใช้เป็นเกณฑ์การพิจารณาตัดสินความรอบรู้ของผู้สอบ เป็นผู้รอบรู้และผู้ไม่รอบรู้ สำหรับการวัดผลสัมฤทธิ์ทางการเรียน นอกจากในระดับชั้นเรียนแล้ว หน่วยงานที่รับผิดชอบทางการศึกษาของประเทศ เช่น กรมวิชาการ กระทรวงศึกษาธิการก็ได้นำคะแนนจุดตัดมาใช้ในโครงการตรวจสอบคุณภาพการศึกษาทุกระดับ การตรวจสอบคุณภาพการศึกษาได้เริ่มต้นตั้งแต่ปี พ.ศ. 2524 เป็นต้นมาโดยมีวัตถุประสงค์ที่จะประเมินผลสัมฤทธิ์ทางการเรียนของนักเรียน และจะใช้ผลการประเมินประกอบการพิจารณาปรับปรุงคุณภาพการศึกษาโดยมุ่งวัด

ด้านความรู้ ความคิด (Cognitive) ในวิชาภาษาไทย สังคมศึกษา พลานามัย คณิตศาสตร์ และวิทยาศาสตร์ ต่อมาในปี พ.ศ. 2526 - 2530 ก็ยังคงวัดคุณภาพของนักเรียนในด้านผลสัมฤทธิ์ทางวิชาการของทุกกลุ่มวิชา แต่ได้เพิ่มการวัดด้านลักษณะนิสัยและบุคลิกภาพบางส่วนขึ้น (กรมวิชาการ. ม.ป.ป. : 5) เกณฑ์ที่ใช้ในการตัดสินจะใช้เกณฑ์ตายตัวซึ่งกำหนดไว้เท่ากันสำหรับแบบทดสอบทุกฉบับ คือ 60 เปอร์เซ็นต์ของคะแนนเต็ม (กรมวิชาการ. ม.ป.ป. : 49) ต่อมาในปี พ.ศ. 2531 กรมวิชาการได้เปลี่ยนรูปแบบการตรวจสอบคุณภาพการศึกษาเป็นการประเมินคุณภาพของผู้เรียนและประเมินกระบวนการจัดการศึกษาของโรงเรียนพร้อมกันไปด้วย และเกณฑ์ที่ใช้ในการประเมินก็เปลี่ยนไป โดยใช้การหาคะแนนจุดตัดที่เหมาะสม เนื่องจากในการใช้เกณฑ์ 60 เปอร์เซ็นต์ของคะแนนเต็มนั้น พบว่า มีจุดอ่อน คือแบบทดสอบบางครั้งอาจง่ายไปหรือยากไป นอกจากนั้นยังพบว่า การใช้เกณฑ์ดังกล่าวไม่ได้สื่อความหมายถึงระดับความรู้ความสามารถของนักเรียนได้โดยตรง เพราะฉะนั้นเกณฑ์ที่ใช้ในการประเมินจึงใช้วิธีการหาคะแนนจุดตัดตามแนวคิดของนีเดลสกี โดยให้ครูผู้สอนพิจารณาแบบทดสอบและคุณสมบัติของข้อสอบเป็นรายข้อ แล้วให้ระบุว่ากลุ่มคาบเส้น 2 กลุ่ม คือนักเรียนเก่งและพอใช้ สามารถทำข้อสอบได้อย่างไร แล้วหาผลรวมโอกาสตอบถูกรายข้อของแบบทดสอบแต่ละฉบับเป็นเกณฑ์มาตรฐานหรือคะแนนจุดตัดในการจำแนกความรู้ ความสามารถของนักเรียนออกเป็น 3 กลุ่ม คือกลุ่มเรียนดี ปานกลาง และต่ำ (กรมวิชาการ. ม.ป.ป. : 7)

สำหรับในปีการศึกษา 2533 กรมวิชาการ กระทรวงศึกษาธิการก็มีโครงการตรวจสอบคุณภาพการศึกษาทุกระดับชั้น โดยได้ทำการปรับปรุงแบบทดสอบจากที่เคยใช้ในปีการศึกษา 2531 ใหม่ เพื่อให้สอดคล้องกับหลักสูตรฉบับปรับปรุง พ.ศ. 2533 ด้วย ดังนั้นการที่จะนำผลการสอบเพื่อจัดจำแนกผู้สอบเป็นกลุ่มเรียนดี ปานกลาง และต่ำนั้น จะต้องทำการกำหนดคะแนนจุดตัดใหม่ซึ่งสอดคล้องกับแนวคิดของ แฮมเบิลตัน (Hambleton. 1978 : 287) ที่ว่า ถ้ามีการเปลี่ยนแปลงหลักสูตรก็ควรจะพิจารณากำหนดคะแนนจุดตัดใหม่ นอกจากนี้จากการตรวจสอบคุณภาพการศึกษาเท่าที่ผ่านมาก็ไม่ได้นำผลการตรวจสอบไปสัมพันธ์กับผลสัมฤทธิ์ทางการเรียนของนักเรียนว่า นักเรียนที่มีความรอบรู้ในรายวิชาต่าง ๆ ของโครงการตรวจสอบคุณภาพการศึกษาจะมีความรอบรู้ในรายวิชาต่าง ๆ ที่ศึกษาอยู่ในโรงเรียนตามปกติหรือไม่โดยทำการเปรียบเทียบในรายวิชาที่มีเนื้อหาวิชาใกล้เคียงกัน

จากที่ได้อภิปรายมาทั้งหมดนี้ ผู้วิจัยสนใจที่จะศึกษาการกำหนดคะแนนจุดตัดโดยกำหนดเกณฑ์มาตรฐานขึ้นก่อนแล้วจึงนำทฤษฎีการตอบสนองต่อข้อสอบมาประยุกต์ใช้ตามแนวคิดของ แวน เดอร์ ลินเดน (Van der Linden. 1982 : 295 - 308) และเคน (Kane. 1987 : 333 - 345) เพราะทั้ง 2 แนวคิดนี้ต่างก็หาคะแนนจุดตัดด้วยกันทั้งสิ้น จึงเป็นปัญหาว่าคะแนนจุดตัดที่ได้จากทั้ง 2 แนวคิดจะแตกต่างกันหรือไม่ การหาคะแนนจุดตัดด้วยการกำหนดเกณฑ์มาตรฐานขึ้นก่อนเช่นนี้น่าจะเหมาะสมกว่าการพิจารณาคะแนนจุดตัดโดยการกำหนดเกณฑ์มาตรฐานแต่เพียงอย่างเดียว สำหรับข้อมูลที่จะนำมาใช้ศึกษา คือข้อมูลจากโครงการตรวจสอบคุณภาพการศึกษานักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2533 เพราะฉะนั้นวิธีการและแนวทางในการศึกษาครั้งนี้ผู้วิจัยจะศึกษาวิธีการกำหนดเกณฑ์มาตรฐาน 2 วิธีคือ วิธีการพิจารณาจากความสามารถต่ำสุดของผู้สอบโดยใช้เทคนิคของนีเดลสกี กับวิธีการใช้กลุ่มคาบเส้น ทั้งนี้เนื่องจากวิธีการพิจารณาจากความสามารถต่ำสุดของผู้สอบโดยใช้เทคนิคของนีเดลสกีเป็นวิธีการที่ลิฟวิงสตัน และซีกี (Livingston and Zieky. 1982 : 54) กล่าวว่า เป็นวิธีการที่เหมาะสมกับผู้สอบที่เป็นคนกลุ่มใหญ่ และมีความสามารถแตกต่างกันมาก ๆ รวมทั้งเหมาะสมกับข้อสอบแบบเลือกตอบด้วย วิธีการนี้จึงสอดคล้องกับข้อมูลของโครงการตรวจสอบคุณภาพการศึกษาเนื่องจากเป็นการทดสอบกับนักเรียนทั่วประเทศและข้อสอบที่ใช้ก็เป็นข้อสอบแบบเลือกตอบด้วย ส่วนการนำวิธีการใช้กลุ่มคาบเส้นมาใช้ศึกษาควบคู่กันเนื่องจากการใช้วิธีการพิจารณาจากความสามารถต่ำสุดของผู้สอบตามเทคนิคของนีเดลสกี ผู้เชี่ยวชาญจะต้องใช้วิธีการคาดคะเนหรือเดาว่าผู้สอบที่มีความสามารถต่ำสุดจะมีโอกาสตอบแต่ละข้อคำถามได้ถูกต้องมากน้อยเพียงใด การคาดคะเนเช่นนี้อาจมีความคลาดเคลื่อนเกิดขึ้นได้หากผู้เชี่ยวชาญแต่ละคนมีความคิดเกี่ยวกับความสามารถต่ำสุดของผู้สอบแตกต่างกัน วิธีการใช้กลุ่มคาบเส้นนี้เป็นการพิจารณาจากความสามารถของผู้สอบแทนที่จะพิจารณาจากข้อคำถามในแบบทดสอบ การพิจารณาเช่นนี้จะใกล้เคียงกับสภาพความเป็นจริงของผู้สอบมาก (Zieky. 1987 : 12) หากผู้เชี่ยวชาญสามารถระบุได้อย่างแน่ชัดว่าผู้สอบคนใดเป็นผู้ที่มีความสามารถต่ำสุด ในการวิจัยครั้งนี้ผู้วิจัยได้ศึกษาการกำหนดเกณฑ์มาตรฐานทั้ง 2 วิธีนี้ ก็เพื่อที่จะทราบว่าเกณฑ์มาตรฐานที่ได้จากทั้ง 2 วิธีนี้ วิธีใดมีความสอดคล้องกับสภาพความสามารถของผู้สอบมากที่สุด ต่อจากนั้นจึงหาคะแนนจุดตัดโดยการนำทฤษฎีการตอบสนองต่อข้อสอบมาประยุกต์ใช้ตามแนวคิดของ แวน เดอร์ ลินเดน

และเคน เพื่อจะได้เปรียบเทียบว่าคะแนนจุดตัดที่ได้นี้จะแตกต่างกันหรือไม่วิธีใดจะมีความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้ได้ดีกว่ากัน นอกจากนั้นจะนำผลที่ได้จากการศึกษาไปพิจารณาถึงความสอดคล้องกับผลสัมฤทธิ์ทางการเรียนของนักเรียนเป็นรายวิชาซึ่งมีเนื้อหาวิชาใกล้เคียงกัน ว่า นักเรียนที่มีความรอบรู้ในรายวิชาต่าง ๆ ของโครงการตรวจสอบคุณภาพการศึกษาจะมีความรอบรู้ในรายวิชาต่าง ๆ ที่เรียนในโรงเรียนซึ่งมีเนื้อหาวิชาใกล้เคียงกันหรือไม่ เพื่อสะท้อนให้เห็นสภาพการเรียนการสอนตามหลักสูตรที่กำหนดไว้โดยมีวัตถุประสงค์ของการวิจัย ดังนี้

<u>วัตถุประสงค์ของการวิจัย</u>

1. เพื่อเปรียบเทียบคะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน และเคน โดยการพิจารณาจากความสามารถต่ำสุดของผู้สอบตามเทคนิคของนีเดลสกี

2. เพื่อเปรียบเทียบคะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน และเคน โดยวิธีการใช้กลุ่มคาบเส้น

3. เพื่อเปรียบเทียบคะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน โดยการพิจารณาจากความสามารถต่ำสุดของผู้สอบตามเทคนิคของนีเดลสกี กับวิธีการใช้กลุ่มคาบเส้น

4. เพื่อเปรียบเทียบคะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีของ เคน โดยการพิจารณาจากความสามารถต่ำสุดของผู้สอบตามเทคนิคของนีเดลสกีกับวิธีการใช้กลุ่มคาบเส้น

5. เพื่อศึกษาผลการตัดสินจำนวนผู้รอบรู้และไม่รอบรู้จากการใช้คะแนนจุดตัดที่ได้จากการคำนวณในข้อ 1, 2, 3 และ 4

6. เพื่อศึกษาความเชื่อมั่นในการจำแนกความรอบรู้จากวิธีการที่ใช้คะแนนจุดตัดที่ได้จากการคำนวณในข้อ 1, 2, 3 และ 4

7. เพื่อศึกษาความเที่ยงตรงในการจำแนกความรอบรู้จากวิธีการใช้คะแนนจุดตัดที่ได้จากการคำนวณในข้อ 1, 2, 3 และ 4

สมมุติฐานของการวิจัย

วิธีการกำหนดคะแนนจุดตัดโดยการกำหนดเกณฑ์มาตรฐานขึ้นก่อน 2 วิธี คือวิธีการพิจารณาจากความสามารถต่ำสุดของผู้สอบตามเทคนิคของนีเดลสกีกับวิธีการใช้กลุ่มคาบเส้นนั้น วิธีการใช้กลุ่มคาบเส้นเป็นวิธีการที่น่าจะใกล้เคียงกับสภาพความเป็นจริงของผู้สอบมาก (Zieky. 1987 : 12) เนื่องจากเป็นวิธีการพิจารณาจากความสามารถของผู้สอบแทนที่จะพิจารณาจากข้อคำถามในแบบทดสอบ ต่อจากนั้นจึงนำทฤษฎีการตอบสนองต่อข้อสอบมาประยุกต์ใช้ตามแนวคิดของ แวน เดอร์ ลินเดน และเคน ทั้ง 2 แนวคิดนี้ต่างก็หาคะแนนจุดตัดโดยอาศัยโค้งคุณลักษณะแบบทดสอบ แต่วิธีการของเคนเป็นวิธีการที่นำความแตกต่างในการตัดสินของผู้เชี่ยวชาญมาพิจารณาด้วย ดังนั้นผู้วิจัยจึงกำหนดสมมุติฐานการวิจัยดังนี้

1. คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคนน่าจะมีค่าต่ำกว่าคะแนนจุดตัดที่ได้จากวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานโดยการพิจารณาจากความสามารถต่ำสุดของผู้สอบตามเทคนิคของนีเดลสกี

2. คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคนน่าจะมีค่าต่ำกว่าคะแนนจุดตัดที่ได้จากวิธีการของแวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้น

3. คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของแวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นน่าจะมีค่าต่ำกว่าวิธีการพิจารณาจากความสามารถต่ำสุดของผู้สอบตามเทคนิคของนีเดลสกี

4. คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นน่าจะมีค่าต่ำกว่าวิธีการพิจารณาจากความสามารถต่ำสุดของผู้สอบตามเทคนิคของนีเดลสกี

5. จำนวนผู้รอบรู้จากการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคน มีจำนวนมากกว่าวิธีการของแวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานโดยการพิจารณาจากความสามารถต่ำสุดของผู้สอบตามเทคนิคของนีเดลสกี

6. จำนวนผู้รอบรู้จากการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคนมีจำนวนมากกว่าวิธีการของแวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้น

7. จำนวนผู้รอบรู้จากการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของแวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้น มีจำนวนมากกว่าวิธีการพิจารณาจากความสามารถต่ำสุดของผู้สอบตามเทคนิคของนีเดลสกี

8. จำนวนผู้รอบรู้จากการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นมีจำนวนมากกว่าวิธีการพิจารณาจากความสามารถต่ำสุดของผู้สอบตามเทคนิคของนีเดลสกี

9. ความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคนน่าจะมีความเชื่อมั่นสูงกว่าวิธีการของแวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานโดยการพิจารณาจากความสามารถต่ำสุดของผู้สอบตามเทคนิคของนีเดลสกี

10. ความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคนน่าจะมีความเชื่อมั่นสูงกว่าวิธีการของแวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้น

11. ความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของแวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นน่าจะมีความเชื่อมั่นสูงกว่าวิธีการพิจารณาจากความสามารถต่ำสุดของผู้สอบตามเทคนิคของนีเดลสกี

12. ความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคนจากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นน่าจะมีความเชื่อมั่นสูงกว่าวิธีการพิจารณาจากความสามารถต่ำสุดของผู้สอบตามเทคนิคของนีเดลสกี

13. ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคนน่าจะมีความเที่ยงตรงสูงกว่าวิธีการของแวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานโดยการพิจารณาจากความสามารถต่ำสุดของผู้สอบตามเทคนิคของนีเดลสกี

14. ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคนน่าจะมีความเที่ยงตรงสูงกว่าวิธีการของแวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้น

15. ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นน่าจะมีความเที่ยงตรงสูงกว่าวิธีการพิจารณาจากความสามารถต่ำสุดของผู้สอบตามเทคนิคของนีเดลสกี

16. ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคนจากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นน่าจะมีความเที่ยงตรงสูงกว่าวิธีการพิจารณาจากความสามารถต่ำสุดของผู้สอบตามเทคนิคของนีเดลสกี

<u>ประโยชน์ที่คาดว่าจะได้รับ</u>

1. ทำให้ได้วิธีการกำหนดคะแนนจุดตัดซึ่งจะเป็นประโยชน์ต่อการตัดสินใจเลือกใช้ให้เหมาะสมกับโครงการตรวจสอบคุณภาพการศึกษาและอื่น ๆ

2. ข้อค้นพบจากการวิจัยนี้จะเป็นข้อมูลย้อนกลับในการปรับปรุงคุณภาพการเรียนการสอนแก่หน่วยงานที่มีหน้าที่รับผิดชอบทางการศึกษาต่อไป

<u>ขอบเขตของการวิจัย</u>

1. วิธีการที่นำมาใช้ในการกำหนดเกณฑ์มาตรฐาน มี 2 วิธีคือ 1) วิธีการพิจารณาจากความสามารถต่ำสุดของผู้สอบตามเทคนิคของนีเดลสกี 2) วิธีการใช้กลุ่มคาบเส้น ส่วน

วิธีการที่นำมาใช้ในการกำหนดคะแนนจุดตัดโดยการนำทฤษฎีการตอบสนองต่อข้อสอบมาประยุกต์ใช้มี 2 วิธีคือ 1) วิธีการของแวน เดอร์ ลินเดน 2) วิธีการของเคน

2. ประชากรที่ใช้ในการวิจัยครั้งนี้เป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ภาคเรียนที่ 2 ปีการศึกษา 2533

3. คะแนนที่นำมาใช้ในการกำหนดคะแนนจุดตัด เป็นคะแนนที่ได้จากแบบทดสอบวัดความสามารถทั่วไป การคิดและแก้ปัญหา และคุณลักษณะทั่วไปของโครงการตรวจสอบคุณภาพการศึกษา นักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2533 ดังนี้

3.1 แบบทดสอบวัดความสามารถทั่วไปฉบับที่ 1

3.2 แบบทดสอบวัดความสามารถทั่วไปฉบับที่ 2

3.3 แบบทดสอบวัดความสามารถทั่วไปฉบับที่ 3

3.4 แบบทดสอบวัดการคิดและแก้ปัญหา

3.5 แบบทดสอบวัดคุณลักษณะทั่วไป

4. ผลสัมฤทธิ์ทางการเรียนที่จะนำมาศึกษาความสอดคล้องกับผลการสอบของนักเรียนในโครงการตรวจสอบคุณภาพการศึกษา เป็นระดับผลการเรียนของนักเรียนที่ได้จากการสอบปลายปี ปีการศึกษา 2533 ของนักเรียนชั้นประถมศึกษาปีที่ 6 ซึ่งมี 4 กลุ่ม คือ

4.1 กลุ่มทักษะ - ภาษาไทย

4.2 กลุ่มทักษะ - คณิตศาสตร์

4.3 กลุ่มสร้างเสริมประสบการณ์ชีวิต

4.4 กลุ่มสร้างเสริมลักษณะนิสัย

5. วิธีการที่นำมาใช้ในการศึกษาความเชื่อมั่นในการจำแนกความรอบรู้ คือวิธีของฮวิน

<u>นิยามศัพท์เฉพาะ</u>

1. คะแนนจุดตัด หมายถึง คะแนนที่ใช้ในการตัดสินพฤติกรรมขั้นต่ำสุดของผู้สอบที่ยอมรับว่าเป็นผู้รอบรู้

2. ผู้รอบรู้ หมายถึง นักเรียนที่สอบได้คะแนนเท่ากับหรือมากกว่าคะแนนจุดตัดที่กำหนดจากแบบทดสอบของโครงการตรวจสอบคุณภาพการศึกษาของนักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2533

3. แบบทดสอบของโครงการตรวจสอบคุณภาพการศึกษาของนักเรียนชั้นประถมศึกษาปีที่ 6 หมายถึง แบบทดสอบที่ใช้วัดความสามารถทั่วไป แบบทดสอบวัดการคิดและแก้ปัญหา และแบบทดสอบวัดคุณลักษณะทั่วไป

4. วิธีการพิจารณาจากความสามารถต่ำสุดของผู้สอบตามเทคนิคของนีเดลสกี หมายถึง วิธีการกำหนดเกณฑ์มาตรฐานจากการพิจารณาของผู้เชี่ยวชาญที่ใช้วิธีการคาดคะเนว่าผู้สอบที่มีความสามารถต่ำสุดจะมีโอกาสตอบข้อคำถามในแต่ละข้อถูกได้มากน้อยเพียงใด

5. วิธีการใช้กลุ่มคาบเส้น หมายถึง วิธีการกำหนดเกณฑ์มาตรฐานจากคะแนนที่ได้จากการสอบของผู้สอบที่ได้ระดับผลการเรียน 1 และระดับผลการเรียน 3 จากการสอบปลายภาคเรียนที่ 1 ของนักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2533

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

การวิจัยครั้งนี้ ผู้วิจัยจะเสนอเอกสารและงานวิจัยที่เกี่ยวข้อง ดังนี้คือ

1. ความหมายและแนวคิดของคะแนนจุดตัด
2. วิธีการกำหนดคะแนนจุดตัด
3. ทฤษฎีการตอบสนองต่อข้อสอบกับการกำหนดคะแนนจุดตัด
4. งานวิจัยที่เกี่ยวข้องกับการกำหนดคะแนนจุดตัดที่นำทฤษฎีการตอบสนองต่อข้อสอบมาประยุกต์ใช้
5. ความเชื่อมั่นในการตัดสินจำแนกความรอบรู้
6. โครงการตรวจสอบคุณภาพการศึกษา

<u>ความหมายและแนวคิดของคะแนนจุดตัด</u>

คำว่า "คะแนนจุดตัด" (Cut-off Scores) นั้นมีชื่อเรียกต่าง ๆ กัน เช่น เกณฑ์ (Criteria) มาตรฐาน (Standard) คะแนนผ่าน (Passing Score) ระดับความรอบรู้ (Mastery Level) หรือความสามารถต่ำสุด (Minimal Competence) สุดแล้วแต่นักวิชาการแต่ละท่านจะกำหนดขึ้น โดยมีความหมายที่แตกต่างกัน ดังนี้

แฮมเบิลตัน (Hambleton. 1978 : 279) ใช้คำว่า "มาตรฐาน" (Standard) แทนซึ่งหมายถึง คะแนนที่ได้จากการสอบที่ใช้ในการแยกผู้สอบออกเป็น 2 กลุ่มตามระดับความสามารถที่แตกต่างกัน กล่าวคือ เป็นกลุ่มที่รอบรู้ (Masters) และกลุ่มที่ไม่รอบรู้ (Non-masters)

แอนดรูว์ และเฮชท์ (Halpin, Sigmon and Halpin. 1983 : 185 ; citing Andrew and Hecht. 1976) ได้กล่าวว่า "มาตรฐาน" (Standard) หมายถึง ระดับการผ่าน - ไม่ผ่านในแบบทดสอบที่กำหนดให้ผู้สอบสอบ

ไปแฟม (Halpin, Sigmon and Halpin. 1983 : 185 ; citing Popham. 1978) ได้ให้นิยามคำว่า "มาตรฐาน" ว่าเป็นการวัดความสามารถของผู้สอบตามจุดประสงค์ที่กำหนดไว้

ซีกี (Zieky. 1983 : 1) ได้กล่าวถึง คำว่า "มาตรฐาน" หรือคะแนนผ่าน (Passing Score) หรือคะแนนจุดตัด (Cut Score) หรือจุดได้ - ตก (Pass-Fail Point) ว่าจะอยู่ในรูปแบบของการตัดสินที่นำคะแนนมาใช้ คนที่ได้คะแนนเท่ากับหรือสูงกว่ามาตรฐานจะถูกจัดอยู่ในกลุ่มหนึ่ง และคนที่ได้คะแนนต่ำกว่ามาตรฐานก็จะจัดไว้ในอีกกลุ่มหนึ่ง

เบอร์ก (Berk. 1986 : 138) ได้ให้ความหมายของ "คะแนนจุดตัด" ว่าหมายถึง จุดที่ใช้แบ่งการแจกแจงของคะแนนออกเป็น 2 ส่วน ส่วนหนึ่งเป็นคะแนนของผู้สอบที่จัดว่าเป็นผู้รอบรู้ อีกส่วนหนึ่งเป็นคะแนนของผู้สอบที่จัดว่าเป็นผู้ไม่รอบรู้ เนื่องจากคะแนนจุดตัดเป็นค่าที่ได้จากการสังเกต เพราะฉะนั้นการกำหนดคะแนนจุดตัดจึงย่อมมีความคลาดเคลื่อนมาตรฐาน (Standard Error) สำหรับประเภทของความคลาดเคลื่อนนี้ เบอร์ก แบ่งออกเป็น 2 ประเภท คือ

1. ผู้สอบที่มีความรอบรู้อย่างแท้จริง แต่ถูกจัดเป็นประเภทไม่มีความรอบรู้ ความคลาดเคลื่อนนี้ เกิดจากการกำหนดคะแนนจุดตัดสูงเกินไป จึงทำให้ผู้สอบที่มีความรอบรู้สอบไม่ผ่านเกณฑ์ที่กำหนด

2. ผู้สอบที่ไม่มีความรอบรู้อย่างแท้จริง แต่ถูกจัดเป็นประเภทมีความรอบรู้ ความคลาดเคลื่อนนี้ เกิดจากการกำหนดคะแนนจุดตัดต่ำเกินไป จึงทำให้ผู้สอบที่ไม่มีความรอบรู้สอบผ่านเกณฑ์ที่กำหนด

ส่วน ลิฟวิงสตัน และซีกี (Livingston and Zieky. 1982 : 10)ได้ให้ความเห็นว่า คำว่า "มาตรฐาน" เป็นการตอบคำถามว่า " มากเท่าไรจึงจะเพียงพอ" ซึ่งสามารถเป็นได้ทั้งมาตรฐานเชิงสัมบูรณ์ (Absolute) หรือมาตรฐานเชิงสัมพัทธ์ (Relative) มาตรฐานเชิงสัมพัทธ์แตกต่างจากมาตรฐานเชิงสัมบูรณ์คือ มาตรฐานเชิงสัมพัทธ์ เป็นมาตรฐานที่ขึ้นอยู่กับการเปรียบเทียบระหว่างผู้สอบด้วยกัน ซึ่งเป็นการพยายามที่จะตอบคำถามว่า "ดีเท่าไรจึงจะดีเพียงพอ" คำว่า "ดีเพียงพอ" นั้น หมายถึง คะแนนของผู้สอบ กล่าวคือ คะแนนของผู้สอบแต่ละคนจะนำมาเปรียบเทียบกับคะแนนของผู้สอบคนอื่น ๆ ถ้าผู้สอบคนอื่น ๆ ได้คะแนนสูงมาตรฐานก็จะสูงตามไปด้วย แต่มาตรฐานเชิงสัมบูรณ์ เป็นมาตรฐานที่ไม่ต้องขึ้นอยู่กับคะแนนของผู้สอบคนอื่น ๆ ในการที่จะทราบว่าคะแนนจุดตัดเป็นมาตรฐานเชิงสัมบูรณ์ หรือมาตรฐานเชิงสัมพัทธ์ ก็จะต้องรู้ว่าคะแนนที่ได้จากการสอบอยู่ในลักษณะใด เช่น ถ้ากล่าวว่า คะแนนจุดตัดเท่ากับ 60 คะแนน จาก

100 คะแนน ข้อความดังกล่าวก็จะไม่บอกอะไรมากนัก ถ้าไม่ทราบความหมายของคำว่า "คะแนน 60" แต่ถ้าความหมายของ "คะแนน 60" หมายความว่า ร้อยละ 60 ของการตอบข้อคำถามได้ถูกต้อง คะแนนจุดตัดที่กำหนดก็จะเป็นมาตรฐานเชิงสัมบูรณ์ แต่ถ้า "คะแนน 60" หมายความว่า ดีกว่าร้อยละ 60 ของผู้สอบทั้งหมด หรือต่ำกว่าคะแนนเฉลี่ยของผู้สอบ การกำหนดคะแนนจุดตัดเช่นนี้ เป็นการกำหนดด้วยมาตรฐานเชิงสัมพัทธ์

หากกล่าวโดยสรุปจะเห็นได้ว่า คะแนนจุดตัดก็คือ คะแนนที่เป็นเกณฑ์หรือมาตรฐานที่ใช้ในการตัดสินขั้นต่ำสุด เพื่อที่จะใช้บอกว่าผู้สอบเป็นผู้รอบรู้หรือไม่รอบรู้

จากความหมายของคะแนนจุดตัดดังกล่าวข้างต้น จะเห็นว่าการกำหนดคะแนนจุดตัดเป็นสิ่งจำเป็น ซึ่งสืบเนื่องมาจากการวัดผลแบบอิงเกณฑ์จะต้องทำการประเมินให้สอดคล้องกับวัตถุประสงค์การเรียนการสอน (Objectives - Based Instructional Programs) ที่กำหนดขึ้น หากพิจารณาในเชิงทฤษฎีจะเห็นว่า ในการที่จะปรับปรุงคุณภาพของการเรียนการสอนจะต้องเริ่มจาก 1) การกำหนดวัตถุประสงค์ของหลักสูตร 2) การจัดการเรียนการสอนให้สอดคล้องกับวัตถุประสงค์ 3) การประเมินผลการเรียนการสอน และจากการศึกษาพบว่าการจัดการเรียนการสอนตามวัตถุประสงค์ที่กำหนดจะทำให้ผลสัมฤทธิ์ทางการเรียนของผู้เรียนสูงกว่าการจัดการเรียนการสอนที่ไม่เป็นไปตามวัตถุประสงค์ เพราะฉะนั้นในการที่จะประเมินผลสัมฤทธิ์ทางการเรียนจึงมีความจำเป็นที่จะต้องกำหนดคะแนนจุดตัด (Hambleton. 1978 : 280 - 281)

สำหรับในการกำหนดคะแนนจุดตัดนั้น จะต้องพิจารณาถึงปัจจัยหลาย ๆ อย่างประกอบกัน (Hambleton. 1978 : 281) ดังนี้คือ

1. ความตรงของเนื้อหาการสอบวัด (Item Content)
2. ผลกระทบต่อเนื่องทางการศึกษา (Educational Consequence)
3. ผลทางจิตวิทยาและผลกระทบทางการลงทุน (Psychological and Financial Costs)
4. ผลกระทบต่องานของผู้อื่น (Performance of Others)
5. ขนาดของความคลาดเคลื่อนซึ่งขึ้นอยู่กับการเดาและการสุ่มข้อคำถาม (Error due to Guessing and Item Sampling)

180702

นอกจากนี้แล้ว แฮมเบิลตัน (Hambleton. 1978 : 288 - 289) ยังได้ให้ข้อเสนอแนะเพิ่มเติมในการกำหนดคะแนนจุดตัด ดังนี้

1. ในการกำหนดคะแนนจุดตัดควรจะกำหนดร่วมกันหลาย ๆ ฝ่าย เช่น ครู ผู้ปกครอง ผู้เชี่ยวชาญด้านหลักสูตร ผู้บริหารโรงเรียน และนักเรียน จำนวนคนในแต่ละกลุ่มขึ้นอยู่กับความสำคัญของแบบทดสอบที่นำมาพิจารณาและตารางวิเคราะห์ข้อสอบหรือมวลความรู้ที่จะสอบวัด (Domain Specification)

2. การใช้วิธีการของ อีเบล หรือนีเดลสกีนั้น จะต้องได้รับการฝึกฝนจากการกระทำจากกลุ่มตัวอย่างหลาย ๆ ครั้ง หากพบความแตกต่างที่เกิดขึ้นจากกลุ่มควรจะได้นำมาอภิปราย และแก้ปัญหาร่วมกัน

3. ผู้ทำการตัดสินเกณฑ์ ควรจะได้ทำความเข้าใจถึงมวลความรู้ที่เป็นจุดประสงค์ของการวัดร่วมกันอย่างถ่องแท้ก่อน

4. ควรจะได้กำหนดเวลาให้เหมาะสมในการพิจารณาตารางวิเคราะห์ข้อสอบ หากตารางวิเคราะห์ข้อสอบบางตารางเป็นตารางที่ซับซ้อนหรือมีความสำคัญมากควรจะจัดเวลาไว้สำหรับการพิจารณาให้มากขึ้น

5. ผู้ทำการตัดสินเกณฑ์ควรจะต้องรู้ว่า แบบทดสอบจะใช้อย่างไรและใช้กับผู้สอบกลุ่มใด

6. ในบางจุดประสงค์ของตารางวิเคราะห์ข้อสอบ หากพบว่าเป็นจุดประสงค์ที่จำเป็นต้องมีหรือเป็นจุดประสงค์ที่ต้องมีก่อนจุดประสงค์อื่น ๆ ก็ควรจะบันทึกไว้และในการกำหนดเกณฑ์ก็ควรจะกำหนดเกณฑ์ให้สูงกว่าจุดประสงค์อื่น ๆ

7. ถ้าเป็นไปได้ควรจะแบ่งกลุ่มทำการกำหนดคะแนนจุดตัดออกเป็น 2 กลุ่ม หรือมากกว่า 2 กลุ่ม และควรจะมีการหาความสอดคล้องของการกำหนดคะแนนจุดตัด หากพบว่ามีความแตกต่างกันก็ควรจะอภิปรายร่วมกัน

8. ผลกระทบต่อเนื่องทางการศึกษา หากเห็นว่ายังมีประโยชน์ก็น่าจะนำมาใช้ในการพิจารณากำหนดคะแนนจุดตัด

9. ถ้าข้อมูลที่ได้จากการสอบเหมาะสมที่จะนำมาใช้พิจารณาเปอร์เซ็นต์ของผู้รอบรู้และผู้ไม่รอบรู้ในแต่ละจุดประสงค์ก็ควรนำมาศึกษา หากคะแนนที่ได้จากแบบทดสอบมีลักษณะผิดปกติก็สามารถอธิบายสาเหตุได้โดยการพิจารณาข้อสอบ ระดับของคะแนนจุดตัดจะแปรเปลี่ยนไปตามคะแนนการสอบระหว่างชั้นเรียน เวลาที่ใช้ในการสอบและอื่น ๆ

10. ถ้าเป็นไปได้ก็ควรจะทำการเปรียบเทียบสภาพการรอบรู้ของกลุ่มผู้ที่ไม่ได้เรียนและที่เรียนแล้ว โดยกลุ่มผู้ที่เรียนแล้วควรมีจำนวนผู้รอบรู้มากกว่ากลุ่มที่ยังไม่ได้เรียน

11. ถ้ามีการเปลี่ยนแปลงหลักสูตรหรือเปลี่ยนแปลงวิธีการสอนก็ควรจะพิจารณาการกำหนดคะแนนจุดตัดใหม่

<u>วิธีการกำหนดคะแนนจุดตัด</u>

การกำหนดคะแนนจุดตัดได้แบ่งออกเป็นหลายลักษณะตามแนวคิดของนักวิชาการแต่ละท่านดังนี้

กลาส (Glass. 1978 : 243 - 257) ได้แบ่งวิธีการกำหนดคะแนนจุดตัดออกเป็น 6 วิธี คือ

1. การใช้คะแนนของคนอื่น ๆ (Performance of Others)
2. การใช้วิธีนับถอยหลังจาก 100% (Counting Backwards from 100%)
3. การปรับคะแนนเกณฑ์อื่น ๆ (Bootstrapping on Other Criterion Scores)
4. การตัดสินใจโดยการพิจารณาจากความสามารถต่ำสุดของผู้สอบ (Judging Minimal Competence)
5. การใช้ทฤษฎีการตัดสินใจ (Decision-Theoretic Approaches)
6. การใช้วิธีการวิจัยเชิงปฏิบัติการ (Operations Research Methods)

แฮมเบิลตัน และไอก์เนอร์ (Berk. 1980 : 103 - 107 ; citing Hambleton and Eigner. 1979) ได้แบ่งวิธีการกำหนดคะแนนจุดตัดออกเป็น 3 วิธี คือ

1. การพิจารณาโดยอาศัยการตัดสินของผู้เชี่ยวชาญ (Judgmental)

2. การพิจารณาโดยอาศัยข้อมูลเชิงประจักษ์ (Empirical)

3. การพิจารณาโดยอาศัยวิธีการผสมผสาน (Combination) ระหว่างการตัดสินของผู้เชี่ยวชาญ และอาศัยข้อมูลเชิงประจักษ์

ลิฟวิงสตัน และซีกี (Livingston and Zieky. 1982 : 15 - 52) ได้แบ่งวิธีการกำหนดคะแนนจุดตัดออกเป็น 3 วิธี คือ

1. วิธีการพิจารณาจากข้อคำถามในแบบทดสอบ (Methods Based on Judgments About Test Questions)

2. วิธีการพิจารณาจากผู้สอบเป็นรายบุคคล (Methods Based on Judgments About Individual Test-Takers)

3. วิธีการพิจารณาจากกลุ่มผู้สอบ (Methods Based on Judgments About a Group of Test-Takers)

ครอส อิมพารา ฟรารีย์ และแจเกอร์ (Cross, Impara, Frary and Jaeger. 1984 : 114) ได้แบ่งวิธีการกำหนดคะแนนจุดตัดออกเป็น 2 วิธีคือ

1. การกำหนดโดยใช้การตัดสินจากเนื้อหาของแบบทดสอบ

2. การกำหนดโดยใช้การพิจารณาจากผลการสอบของกลุ่มผู้สอบที่รู้แน่ชัดแล้ว (Known Group)

ถึงแม้ว่าแนวคิดในการกำหนดคะแนนจุดตัดของนักวิชาการแต่ละท่านจะไม่เหมือนกันก็ตาม แต่รายละเอียดในแต่ละวิธีการจะคล้ายกัน เช่น การกำหนดคะแนนจุดตัดโดยวิธีการตัดสินจากการพิจารณาความสามารถต่ำสุดของผู้สอบตามแนวคิดของกลาส การพิจารณาโดยอาศัยการตัดสินของผู้เชี่ยวชาญตามแนวคิดของ แฮมเบิลตัน และไอก์เนอร์ กับวิธีการพิจารณาจากข้อคำถามในแบบทดสอบ ตามแนวคิดของ ลิฟวิงสตัน และซีกี ต่างก็เป็นวิธีการเดียวกัน และในทำนองเดียวกัน การกำหนดคะแนนจุดตัดจากการพิจารณาโดยอาศัยวิธีการผสมผสานระหว่างการตัดสินของผู้เชี่ยวชาญ และอาศัยข้อมูลเชิงประจักษ์ตามแนวคิดของ แฮมเบิลตัน และไอก์เนอร์ กับวิธีการพิจารณาจาก

ผู้สอบเป็นรายบุคคลตามแนวคิดของ ลิฟวิงสตัน และซีกี ก็เป็นวิธีการเดียวกัน

สำหรับวิธีการกำหนดคะแนนจุดตัดในการวิจัยครั้งนี้ ผู้วิจัยจะเสนอวิธีการกำหนดคะแนนจุดตัด โดยการสรุปจากแนวคิดของนักวิชาการแต่ละท่าน อันได้แก่ กลาส แฮมเบิลตัน และไอก์เนอร์ ลิฟวิงสตันและซีกี กับครอส อิมพารา ฟรารีย์ และแจเกอร์ โดยแบ่งวิธีการกำหนดคะแนนจุดตัดออกเป็น 8 วิธี ดังนี้คือ

1. การใช้คะแนนของคนอื่น
2. การใช้วิธีนับถอยหลังจาก 100%
3. การปรับคะแนนเกณฑ์อื่น ๆ
4. การพิจารณาจากข้อคำถามในแบบทดสอบ หรือการตัดสินโดยการพิจารณาจากความสามารถต่ำสุดของผู้สอบ
5. การพิจารณาจากผู้สอบเป็นรายบุคคล
6. การพิจารณาจากกลุ่มผู้สอบ
7. การใช้ทฤษฎีการตัดสินใจ
8. การใช้วิธีการวิจัยเชิงปฏิบัติการ

1. <u>การใช้คะแนนของคนอื่น</u>

การกำหนดระดับเกณฑ์แบบนี้สร้างขึ้นโดยอ้างอิงค่าพารามิเตอร์ของประชากรผู้สอบ ดังนั้นเกณฑ์จึงอาจกำหนดด้วยคะแนนมัธยฐาน (Median) ของคะแนนจากแบบทดสอบหรืออาจจะกำหนดเป็นเปอร์เซ็นไทล์ (Percentile) ซึ่งวิธีการนี้ มิลล์แมน และแฮมเบิลตัน (Glass. 1978 : 244 ; citing Millman. 1973 and Hambleton. 1978) มีความเห็นว่า ยังเป็นวิธีการอิงกลุ่มอยู่เพราะฉะนั้นจึงเป็นวิธีการที่ไม่ค่อยเหมาะสมนัก

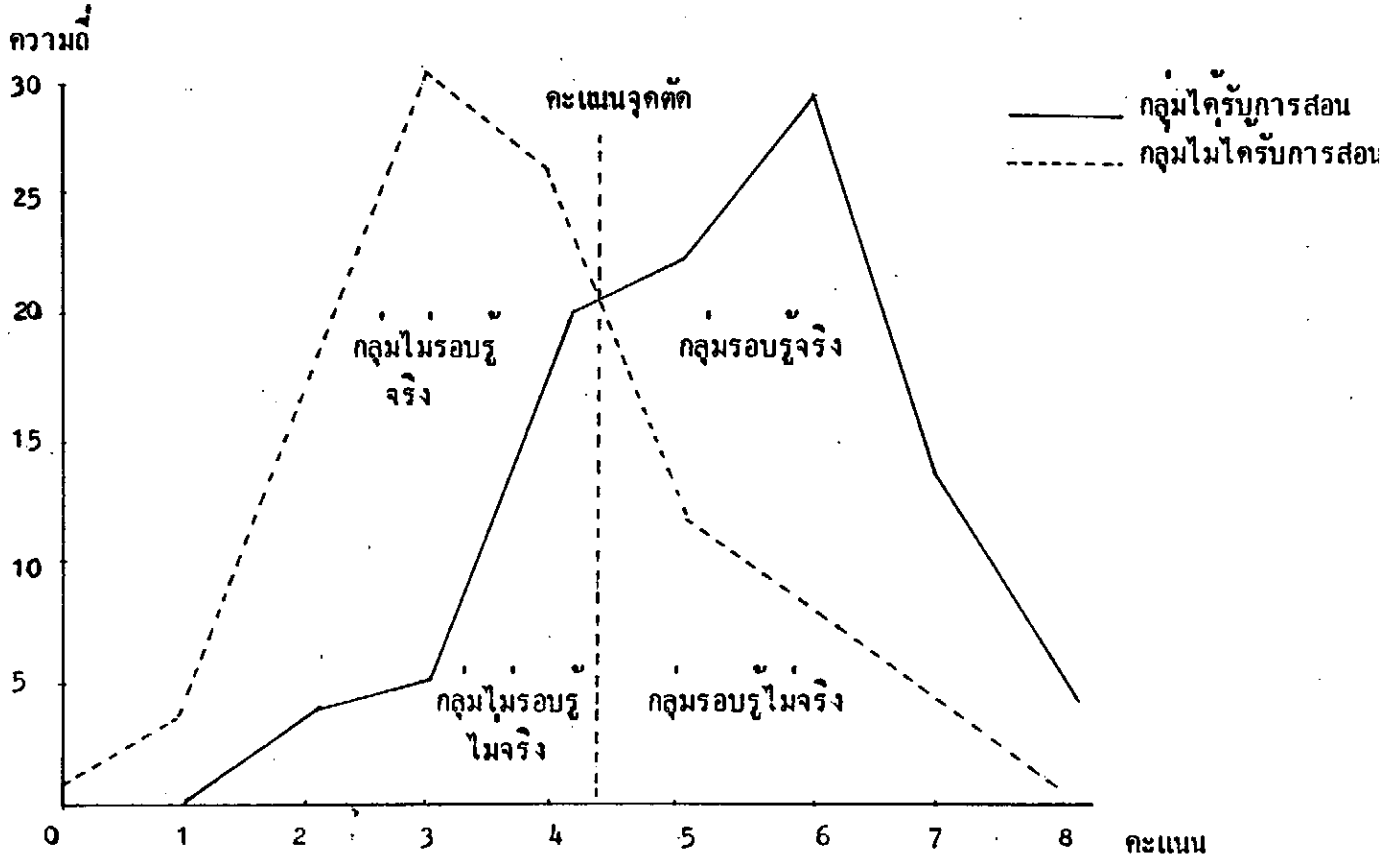
2. <u>การใช้วิธีนับถอยหลังจาก 100 %</u>

โดยปกติในการสร้างแบบทดสอบนั้น จะต้องสร้างให้สอดคล้องกับวัตถุประสงค์ ซึ่งผู้ออกข้อสอบจะถือว่าวัตถุประสงค์เป็นสิ่งสำคัญ เพราะฉะนั้นผู้ออกข้อสอบก็จะคาดหวังว่าทุก

คนควรจะตอบข้อสอบได้ ดังนั้นในการกำหนดระดับคะแนนจึงกำหนดไว้เป็น 100 % แต่ในทางปฏิบัติก็จะเป็นไปไม่ได้ เนื่องจากอาจมีความคลาดเคลื่อนเกิดขึ้น เช่น ความไม่ตั้งใจในการสอบ การพิมพ์ข้อสอบผิดพลาด เพราะฉะนั้นเกณฑ์ที่ตั้งไว้ 100 % อาจลดลงมาเป็น 95 % หรือ 90 % เป็นต้น วิธีการนี้เบอร์ก (Berk. 1986 : 164 - 165) มีความเห็นว่าเหมาะกับระดับชั้นประถมศึกษา (Primary Grades) ในการวัดทักษะพื้นฐานบางอย่างเป็นการวัดในระหว่างดำเนินการสอน (Formative Test) เพื่อพิจารณาถึงความก้าวหน้าของผู้เรียนเป็นสำคัญ

3. <u>การปรับคะแนนเกณฑ์อื่น ๆ</u>

เป็นการกำหนดโดยอาศัยเกณฑ์ภายนอก เพื่อใช้ในการตัดสินว่า "สำเร็จ" หรือ "รอบรู้" ซึ่งกลาส (Glass. 1978 : 246) กล่าวว่า วิธีการที่ใช้แนวคิดนี้ได้แก่ วิธีการของเบอร์ก (Berk. 1976 : 4 - 9) เกณฑ์ที่ใช้คือ ได้รับการสอน และไม่ได้รับการสอน โดยแบ่งผู้สอบออกเป็น 2 กลุ่ม คือ กลุ่มที่ได้รับการสอนจะจัดเป็นกลุ่มรอบรู้ และกลุ่มที่ไม่ได้รับการสอนจะจัดเป็นกลุ่มไม่รอบรู้ ดังนั้นในการหาคะแนนจุดตัดก็คือ นำคะแนนของผู้สอบทั้ง 2 กลุ่ม นี้มาแจกแจง การแจกแจงของคะแนนทั้ง 2 กลุ่มตัดกันที่จุดใด จุดนั้นคือคะแนนจุดตัด ดังภาพประกอบ 1



ภาพประกอบ 1 การจัดประเภทของผู้สอบเป็นผู้ได้รับการสอน และผู้ไม่ได้รับการสอนโดยอาศัยคะแนนสอบ

จากภาพประกอบ 1 เมื่อเลือกคะแนนจุดตัดขึ้นมาแล้ว ก็จะแบ่งผู้สอบออกเป็น 2 กลุ่ม คือ กลุ่มที่ได้คะแนนเท่ากับหรือมากกว่าคะแนนจุดตัด จะเป็นพวกรอบรู้โดยการทำนาย (Predicted Master : PM) และกลุ่มที่ได้คะแนนน้อยกว่าคะแนนจุดตัดจะเป็นพวกไม่รอบรู้โดยการทำนาย (Predicted-Nonmaster : PN) ดังนั้น เมื่อนำเกณฑ์มาผสมกับคะแนนจุดตัดแล้วสามารถเขียนในรูปตาราง 2 X 2 ได้ดังนี้

	ประเภทของเกณฑ์	
	ได้รับการสอน	ไม่ได้รับการสอน
รอบรู้โดยการทำนาย (PN = TM+FM)	กลุ่มรอบรู้จริง (True Master : TM)	กลุ่มรอบรู้ไม่จริง (False Master : FM)
คะแนนจุดตัด ไม่รอบรู้โดยการทำนาย (PN = FN+TN)	กลุ่มไม่รอบรู้ไม่จริง (False Nonmaster : FN	กลุ่มไม่รอบรู้จริง (True Nonmaster : TN
	รอบรู้ (M = TM+FN)	ไม่รอบรู้ (N = FM+TN)

จากตารางแบ่งออกเป็น

1. กลุ่มรอบรู้จริง (True Master : TM) คือกลุ่มผู้สอบที่ได้รับการสอนและได้คะแนนสูงกว่าคะแนนจุดตัด

2. กลุ่มรอบรู้ไม่จริง (False Master : FM) หรือความคลาดเคลื่อนในการแยกประเภทผิดแบบที่ 2 คือกลุ่มผู้สอบที่ไม่ได้รับการสอนแต่ได้คะแนนสูงกว่าคะแนนจุดตัด

3. กลุ่มไม่รอบรู้จริง (True Nonmaster : TN) คือกลุ่มผู้สอบที่ไม่ได้รับการสอนแต่ได้คะแนนต่ำกว่าคะแนนจุดตัด

4. กลุ่มไม่รอบรู้ไม่จริง (False Nonmaster : FN) หรือความคลาดเคลื่อนในการแยกประเภทผิด แบบที่ 1 คือกลุ่มผู้สอบที่ได้รับการสอน แต่ได้คะแนนต่ำกว่าคะแนนจุดตัด

สำหรับความน่าจะเป็นของกลุ่มผู้สอบแต่ละกลุ่ม สามารถหาได้ดังนี้

$P(TM) = TM/(M+N)$

$P(FM) = FM/(M+N)$

$P(TN) = TN/(M+N)$

$P(FN) = FN/(M+N)$

นอกจากการพิจารณาค่าความน่าจะเป็นดังกล่าวแล้ว ยังอาจพิจารณาจากสัมประสิทธิ์ความแม่นตรงของคะแนนจุดตัด ซึ่งเป็นระดับความถูกต้องของคะแนนจุดตัดที่จะพยากรณ์ผู้สอบว่าเป็นไปตามเกณฑ์มากน้อยเพียงใด โดยหาจากสหสัมพันธ์แบบ ฟี (ϕ_{vc}) ดังนี้

$$\phi_{vc} = \frac{P(TM) - BR(SR)}{\sqrt{BR(1-BR)\ SR(1-SR)}}$$

เมื่อ $BR = P(FN) + P(TM)$

$SR = P(TM) + P(FM)$

ถ้า $BR = SR$ แล้ว จะทำให้สัมประสิทธิ์ความแม่นตรงคือ ϕ_{vc} มีค่าเท่ากับ 1.00

4. <u>การพิจารณาจากข้อคำถามในแบบทดสอบ หรือการตัดสินโดยการพิจารณาจากความสามารถต่ำสุดของผู้สอบ</u>

วิธีการตัดสินที่ขึ้นอยู่กับข้อคำถามในแบบทดสอบ วิธีนี้ กลาส (Glass. 1976 : 247 - 251) เรียกว่า เป็นวิธีการพิจารณาจากความสามารถต่ำสุดของผู้สอบ เป็นวิธีการที่อาศัยแนวคิดเกี่ยวกับผู้สอบที่มีความรู้และทักษะอยู่ในระดับคาบเส้น (Borderline) ระหว่างกลุ่มที่ได้คะแนนสูงกับกลุ่มที่ได้คะแนนต่ำ คะแนนของผู้สอบที่มีความรู้และทักษะอยู่ในระดับคาบเส้นก็จะเป็นคะแนนจุดตัด การตัดสินโดยอาศัยวิธีการนี้จะต้องพิจารณาจากข้อคำถามในแบบทดสอบ วิธีการนี้

สามารถกำหนดคะแนนจุดตัดได้ทั้งก่อนและหลังจากการนำแบบทดสอบไปทำการดำเนินการสอบแล้ว หากกล่าวโดยสรุปจะเห็นว่า กระบวนการของวิธีนี้จะมุ่งพิจารณาที่เนื้อหาของแบบทดสอบเป็นสำคัญ เพราะฉะนั้นผู้ตัดสินหรือผู้เชี่ยวชาญก็จะต้องพิจารณาว่า ผู้สอบที่อยู่ในระดับคาบเส้นจะสามารถตอบข้อคำถามแต่ละข้อในแบบทดสอบได้อย่างไร

การกำหนดคะแนนจุดตัดโดยวิธีนี้มีขั้นตอนพื้นฐานดังนี้

1. เลือกผู้ตัดสินหรือผู้เชี่ยวชาญ
2. กำหนดความรู้และทักษะของผู้สอบที่อยู่ในระดับคาบเส้น
3. ฝึกฝนวิธีการตัดสินในแต่ละวิธีการที่จะเลือกใช้
4. เก็บรวบรวมผลการตัดสิน
5. นำผลการตัดสินมารวมกันเพื่อกำหนดคะแนนจุดตัด

ในการเลือกผู้ตัดสินหรือผู้เชี่ยวชาญนั้นถือเป็นขั้นตอนแรกของวิธีการนี้ ซึ่งผู้ตัดสินต้องเป็นผู้มีความสามารถที่จะตัดสินถึงระดับความรู้หรือทักษะที่แบบทดสอบมุ่งวัด เช่น แบบทดสอบวัดความสามารถในการอ่านในระดับมัธยมศึกษาตอนปลาย ผู้ตัดสินก็จะต้องพิจารณาว่าผู้ที่จะได้รับประกาศนียบัตรระดับมัธยมศึกษาตอนปลาย ควรเป็นผู้มีความรู้ระดับใด การกำหนดจำนวนผู้ตัดสินอาจจะกำหนดเพียง 2 - 3 คน ก็ได้ ถ้าผู้ตัดสินดังกล่าวเป็นผู้มีความสามารถ แต่บางกรณีก็จะต้องกำหนดเป็นจำนวนมาก อย่างไรก็ตาม ถ้าผู้ตัดสินมีจำนวนน้อย ก็จะมีผลต่อการกำหนดเกณฑ์มาตรฐานคือ มาตรฐานอาจสูงเกินไปหรือต่ำเกินไป เพราะฉะนั้นในการกำหนดจำนวนผู้ตัดสินถ้ามีจำนวนผู้ตัดสินมากเท่าไรก็จะยิ่งดี

ในทางปฏิบัติอาจจะแบ่งผู้ตัดสินเป็นกลุ่มย่อย ๆ แต่ในครั้งแรกควรมีการประชุมร่วมกันก่อนเพื่อที่จะให้ผู้ตัดสินนิยามคำว่า "ความรู้และทักษะของผู้สอบที่อยู่ในระดับคาบเส้น" ต่อจากนั้นจะต้องมีการฝึกผู้ตัดสินเพื่อจะได้เข้าใจสิ่งที่แบบทดสอบวัดและการนำคะแนนสอบไปใช้

วิธีการตัดสินที่ขึ้นอยู่กับข้อคำถามในแบบทดสอบ แบ่งออกเป็น 3 เทคนิค คือ

1. เทคนิคของนีเดลสกี
2. เทคนิคของแองกอฟฟ์
3. เทคนิคของอีเบล

1. เทคนิคของนีเดลสกี

นีเดลสกีได้เสนอเทคนิคนี้ในปี ค.ศ. 1954 ซึ่งเป็นเทคนิคที่เหมาะกับข้อสอบแบบเลือกตอบ วิธีการก็คือผู้เชี่ยวชาญจะต้องพิจารณาคำถามและระบุว่าผู้สอบที่อยู่ในระดับคาบเส้นจะทราบว่าตัวเลือกใดเป็นตัวเลือกที่ผิด เพราะฉะนั้นในการรวบรวมผลการตัดสินของผู้เชี่ยวชาญอาจจะให้ผู้เชี่ยวชาญแต่ละบุคคลพิจารณาเองหรือจะพิจารณาร่วมกันเป็นกลุ่มก็ได้ ถ้าจำนวนของผู้เชี่ยวชาญไม่มากก็อาจจะพิจารณาร่วมกัน ซึ่งวิธีการนี้จะให้ผลการตัดสินที่เที่ยงตรงกว่าเนื่องจากผู้เชี่ยวชาญได้มีโอกาสแลกเปลี่ยนความคิดเห็นซึ่งกันและกัน ดังนั้นหากกล่าวโดยสรุปเทคนิคของนีเดลสกี มีดังนี้

1. ให้ผู้เชี่ยวชาญแต่ละคนพิจารณาคำถามทั้งหมด และใช้ดินสอขีดลงไปที่ตัวเลือกที่ผิดที่เห็นว่า ผู้สอบที่อยู่ในระดับคาบเส้นจะไม่เลือก

2. อภิปรายถึงคำถามแต่ละข้อ ตามแนวทางดังนี้

2.1 ให้ผู้เชี่ยวชาญพิจารณาคำตอบที่ผิดก่อน และดูว่ามีผู้เชี่ยวชาญที่ท่านที่คิดว่าผู้สอบที่อยู่ในระดับคาบเส้นจะไม่เลือก และมีกี่ท่านไม่เห็นด้วย

2.2 ถ้าผู้เชี่ยวชาญมีความเห็นไม่ตรงกัน จะต้องให้แต่ละคนอธิบายเหตุผลแล้วถามผู้เชี่ยวชาญที่ไม่เห็นด้วย ซึ่งตอนนี้ยังไม่ต้องมีการสรุปจนกว่าผู้เชี่ยวชาญจะแน่ใจว่าได้เข้าใจความหมายของคำว่า ผู้สอบที่มีความสามารถต่ำสุดได้ตรงกัน

2.3 พิจารณาตัวเลือกที่เป็นตัวผิดตัวต่อไป

3. หลังจากที่ได้มีการอภิปรายคำถามทุกข้อในลักษณะนี้แล้ว ขั้นตอนต่อมาคือจะต้องให้ผู้เชี่ยวชาญทบทวนถึงการตัดสินโดยที่ผู้เชี่ยวชาญจะต้องทำเครื่องหมายกากบาท ลงไปบนตัวเลือกที่เขาแน่ใจว่า ผู้สอบที่มีความสามารถต่ำสุดจะไม่เลือก

4. รวบรวมผลการตัดสิน แล้วดำเนินการเช่นเดียวกันในคำถามข้อต่อไป

เทคนิคของนีเดลสกีนี้ อาจให้ผู้เชี่ยวชาญแต่ละคนตัดสินเองก็ได้ โดยปราศจากการอภิปรายร่วมกัน แต่ปัญหาที่จะตามมาก็คือ ผู้เชี่ยวชาญจะยังไม่คุ้นเคยกับวิธีการตัดสิน เพราะฉะนั้นถ้าเป็นไปได้ก็ควรจะให้โอกาสแก่ผู้เชี่ยวชาญในการฝึกตัดสินใจ จากตัวอย่างของคำถามและอภิปรายซึ่งกันและกันก่อนที่จะมีการพิจารณาคำถามที่เหลือด้วยตนเองต่อไป

รูปแบบของข้อสอบแบบเลือกตอบบางครั้งก็เป็นปัญหาแก่การใช้เทคนิคของนีเดลสกี คือคำถามในลักษณะดังต่อไปนี้

1. คำถามในรูปปฏิเสธ เช่น

(0) อาหารในข้อใดต่อไปนี้ ไม่มีวิตามินซี

ก. นม ข. น้ำส้ม ค. กระหล่ำปลี ง. มันฝรั่ง

ในการพิจารณาตัดสินคำตอบที่ถูกผู้เชี่ยวชาญจะต้องจำไว้ว่า อาหารอะไรที่ให้วิตามินซีดีกว่า และต้องรู้ว่าผู้สอบที่อยู่ในระดับคาบเส้นจะจำว่าเป็นคำตอบที่ผิด

2. แบบสอบเลือกตอบประเภทถูกผิดทุกตัวเลือก (Multiple True-False) ดังตัวอย่างต่อไปนี้

(0) ในระหว่างสงครามโลกครั้งที่ 2 ประเทศสหรัฐอเมริกาสู้รบกับประเทศใด

1. เยอรมัน 2. รัสเซีย 3. อิตาลี 4. ญี่ปุ่น

ก. ข้อ 1 เท่านั้น

ข. ข้อ 2 เท่านั้น

ค. ข้อ 1 และข้อ 4 เท่านั้น

ง. ข้อ 1, 3 และ ข้อ 4

จ. ข้อ 1, 2, 3 และข้อ 4

จากตัวอย่างข้างต้น ผู้เชี่ยวชาญจะต้องพิจารณาจำนวนตัวเลือกที่ผู้สอบที่อยู่ในระดับคาบเส้นจะระบุว่า ตัวเลือกใดเป็นตัวเลือกที่ถูกต้อง ตัวเลือกใดไม่ถูกต้องและตัวเลือกใดที่ไม่แน่ใจ แล้วผู้เชี่ยวชาญก็จะต้องพิจารณาว่า ตัวเลือกใดที่ผู้สอบที่อยู่ในระดับคาบเส้นจะไม่เลือกหรือจะกำจัดทิ้งไป

ในกรณีที่แบบทดสอบเป็นคำถามปฏิเสธ หรือแบบทดสอบเลือกตอบแบบถูกผิดทุกตัวเลือก ดังตัวอย่างข้างต้น ถ้าใช้เทคนิคของนีเดลสกี ก็ต้องให้เวลาแก่ผู้เชี่ยวชาญในการฝึกก่อนที่ผู้เชี่ยวชาญแต่ละคนจะทำการพิจารณาตัดสินด้วยตนเอง หลังจากที่ผู้เชี่ยวชาญพิจารณาเสร็จแล้วก็จะต้องให้อธิบายเหตุผลในการตัดสินของเขา เพื่อให้เกิดความแน่ใจในสิ่งที่ผู้เชี่ยวชาญได้พิจารณา

3. คำถามในลักษณะของการคำนวณทางคณิตศาสตร์ ตัวเลือกของคำตอบที่ผิดจะเป็นผลมาจากการคิดคำนวณผิด ซึ่งก็ไม่ได้เด่นชัดนัก เพราะฉะนั้นผู้เชี่ยวชาญจะต้องใช้เวลามากในการพิจารณาแม้แต่ผู้เชี่ยวชาญที่มีความเชี่ยวชาญมาก ๆ ก็ต้องใช้เวลาในการคิดคำนวณถึงวิธีการที่จะนำไปสู่การตอบผิด เพื่อที่จะหลีกเลี่ยงปัญหานี้ก็ควรจะให้แบบทดสอบแก่ผู้ตัดสินอีก 1 ฉบับ และแสดงให้เห็นถึงรูปแบบของการคิดคำนวณที่ผิด ซึ่งจะนำไปสู่การเลือกคำตอบที่ผิด

ดังนั้นประเด็นสำคัญของการนำเทคนิคของนีเดลสกีไปใช้ก็คือ จะต้องให้คำเฉลยแก่ผู้เชี่ยวชาญ ซึ่งจะทำให้การพิจารณาคำถามง่ายขึ้น แต่ถ้าไม่ให้เฉลยหรือบอกตัวเลือกที่เป็นคำตอบที่ถูกแก่ผู้เชี่ยวชาญ ผู้เชี่ยวชาญจะรู้สึกว่าเขากำลังถูกทดสอบและอาจลืมไปว่าเขากำลังจะต้องพิจารณาตัดสินผู้สอบที่อยู่ในระดับคาบเส้น ยิ่งกว่านั้นกระบวนการตัดสินใจจะใช้เวลานาน ถ้าผู้เชี่ยวชาญจะต้องหาคำตอบที่ถูกต้องในแต่ละข้อคำถามเอง

ต่อไปนี้เป็นตัวอย่างการใช้เทคนิคของนีเดลสกี

สมมุติว่า แบบทดสอบเลือกตอบฉบับหนึ่งมีคำถาม 10 ข้อ เป็นแบบทดสอบเลือกตอบ 5 ตัวเลือกคือ ก. ข. ค. ง. และ จ. ดังนี้

ตาราง 1 การคำนวณตามเทคนิคของนีเดลสกี

ข้อคำถาม	ตัวเลือก	จำนวนของตัวเลือกที่ตัดทิ้ง	จำนวนของตัวเลือกที่เหลือ	โอกาสของการตอบถูก
1	ก (ข) ~~ค~~ ~~ง~~ ~~จ~~	3	2	1/2 = .50
2	~~ก~~ ~~ข~~ ~~ค~~ ~~ง~~ (จ)	4	1	1/1 = 1.00
3	~~ก~~ ~~ข~~ ค (ง) ~~จ~~	3	2	1/2 = .50
4	ก ~~ข~~ ค (ง) ~~จ~~	2	3	1/3 = .33
5	(ก) ~~ข~~ ~~ค~~ ~~ง~~ ~~จ~~	4	1	1/1 = 1.00
6	ก ข (ค) ง จ	0	5	1/5 = .20
7	ก ข ค ~~ง~~ (จ)	1	4	1/4 = .25
8	(ก) ข ~~ค~~ ง จ	1	4	1/4 = .25
9	ก (ข) ค ง จ	0	5	1/5 = .20
10	ก (ข) ค ง จ	0	5	1/5 = .20
			รวม	4.43

ตัวเลือกที่วงกลม คือตัวเลือกที่เป็นคำตอบที่ถูก ส่วนตัวเลือกที่ผู้เชี่ยวชาญจะคิดว่าผู้สอบที่มีความสามารถต่ำสุดจะตัดทิ้งไปหรือไม่เลือก ผู้เชี่ยวชาญก็จะเขียนเครื่องหมายกากบาททับลงไปบนตัวเลือกนั้น สำหรับในการคำนวณหาคะแนนจุดตัดนั้น คะแนนในแต่ละข้อจะมีค่าเท่ากับ1 คะแนน ดังนั้นในแต่ละข้อจะต้องหาว่า มีจำนวนตัวเลือกที่ผู้สอบที่อยู่ในระดับความสามารถต่ำสุดจะตัดทิ้งเป็นจำนวนเท่าไร แล้วจึงหาโอกาสของการตอบถูก (Probability of Correct Answer) ซึ่งเท่ากับ 1 หารด้วย จำนวนตัวเลือกทั้งหมด ลบด้วยจำนวนตัวเลือกที่ตัดทิ้ง หากเขียนในรูปสูตรทั่วไปจะได้ว่า

$$\text{โอกาสของการตอบถูก} = \frac{1}{\text{จำนวนตัวเลือกทั้งหมด} - \text{จำนวนตัวเลือกที่ตัดทิ้ง}}$$

ดังนั้น สำหรับข้อสอบแบบเลือกตอบที่มี 5 ตัวเลือก

ถ้าตัวเลือกทั้งหมดตัดทิ้งไม่ได้เลย

$$\text{โอกาสของการตอบถูก} = \frac{1}{5-0} = \frac{1}{5} = 0.20$$

ถ้าตัวเลือกทั้งหมดตัดทิ้งได้ 1 ตัว

$$\text{โอกาสของการตอบถูก} = \frac{1}{5-1} = \frac{1}{4} = 0.25$$

ถ้าตัวเลือกทั้งหมดตัดทิ้งได้ 2 ตัว

$$\text{โอกาสของการตอบถูก} = \frac{1}{5-2} = \frac{1}{3} = 0.33$$

ถ้าตัวเลือกทั้งหมดตัดทิ้งได้ 3 ตัว

$$\text{โอกาสของการตอบถูก} = \frac{1}{5-3} = \frac{1}{2} = 0.50$$

ถ้าตัวเลือกทั้งหมดตัดทิ้งได้ 4 ตัว

$$\text{โอกาสของการตอบถูก} = \frac{1}{5-4} = \frac{1}{1} = 1.00$$

ตัวอย่างในการหาคะแนนจุดตัดตามวิธีการของนีเดลสกี ดังในตาราง 1 นั้น เป็นการหาคะแนนจากผู้เชี่ยวชาญ 1 คน ในกรณีที่มีผู้เชี่ยวชาญมากกว่า 1 คน จะต้องนำคะแนนของผู้เชี่ยวชาญทั้งหมดมาคิดในรูปของคะแนนเฉลี่ย (Mean) มัธยฐาน (Median) หรือ Trimmed Mean แล้วแต่ความเหมาะสมดังตัวอย่างในตาราง 2

ตาราง 2 ผลการหาคะแนนจุดตัดจากผู้เชี่ยวชาญ โดยเทคนิคของนีเดลสกี

ผู้เชี่ยวชาญคนที่	คะแนน
1	92.50
2	77.25
3	67.00
4	66.67
5	65.33
รวม	368.75

จากตาราง 2 จะได้คะแนนเฉลี่ยเท่ากับ 368.75/5 = 73.75

จะได้มัธยฐานเท่ากับ 67.00

จะได้ Trimmed Mean เท่ากับ 210.92/3 = 70.31 ซึ่งการหาโดยวิธีนี้จะตัดคะแนนสูงสุดคือ 92.50 และคะแนนต่ำสุดคือ 65.33 ออกไป

การหาคะแนนจุดตัดตามเทคนิคของนีเดลสกี ดังกล่าวข้างต้นเป็นวิธีการตามแนวคิดของลิฟวิงสตัน และซีกี (Livingston and Zieky. 1982 : 17 - 24)

2. <u>เทคนิคของแองกอฟฟ์</u>

แองกอฟฟ์ได้เสนอเทคนิคของเขา ในปี ค.ศ. 1971 ซึ่งเป็นเทคนิคที่มีลักษณะคล้ายกับเทคนิคของนีเดลสกี แต่เทคนิคนี้จะใช้กับแบบทดสอบที่ไม่ใช่แบบเลือกตอบ คะแนนจุดตัดที่หาโดยเทคนิคของแองกอฟฟ์จะหาจากคะแนนที่ได้จากแต่ละข้อคำถาม โดยการพิจารณาในภาพรวมจากการพิจารณาถึงโอกาสที่ผู้สอบที่อยู่ในระดับคาบเส้นจะตอบแต่ละข้อคำถามได้ถูกต้อง ซึ่งโอกาสที่ผู้เชี่ยวชาญใช้ในการตัดสินจะอยู่ระหว่าง 0.00 ถึง 1.00 กล่าวคือ ถ้าเป็นข้อคำถามที่ง่ายโอกาสที่ผู้สอบที่อยู่ในระดับคาบเส้น จะตอบถูกจะมีค่าใกล้ 1.00 แต่ถ้าเป็นข้อคำถามที่ยาก โอกาสที่ผู้สอบที่อยู่ในระดับคาบเส้นจะตอบถูกจะมีค่าใกล้ 0 อย่างไรก็ตาม เทคนิคของแองกอฟฟ์ หากนำมาใช้กับแบบทดสอบแบบเลือกตอบ โอกาสที่ผู้สอบที่อยู่ในระดับคาบเส้นจะตอบแต่ละข้อคำถามถูกไม่ควรต่ำกว่าโอกาสการตอบถูกโดยการเดา หากกล่าวโดยสรุปเทคนิคของแองกอฟฟ์ มีดังนี้

1. ให้ผู้เชี่ยวชาญแต่ละคนพิจารณาตัดสินข้อคำถามประมาณ 2 - 3 ข้อแรก แล้วอภิปรายร่วมกัน ดังนี้

1.1 ให้ผู้เชี่ยวชาญบอกความน่าจะเป็นในการตัดสินแต่ละข้อคำถามแล้วเขียนลงบนกระดานดำ

1.2 ให้ผู้เชี่ยวชาญที่ให้ค่าความน่าจะเป็นสูงสุด และต่ำสุดในแต่ละข้อคำถามชี้แจงเหตุผล

1.3 ผู้เชี่ยวชาญสามารถเปลี่ยนแปลงการตัดสินได้ ถ้าต้องการเปลี่ยนแปลง แต่อย่างไรก็ตาม ผู้เชี่ยวชาญจะต้องมั่นใจว่าได้เข้าใจ คำว่า "ความสามารถต่ำสุดของผู้สอบ" ตรงกัน

2. หลังจากที่มีการอภิปรายข้อคำถาม 2 - 3 ข้อแรกแล้ว คำถามข้อที่เหลือ ก็ดำเนินการเช่นเดียวกัน

3. เก็บรวบรวมผลการตัดสิน

ส่วนในการคำนวณหาคะแนนจุดตัดสามารถทำได้ 3 วิธี (Zieky. 1987 : 5 - 6) คือวิธีที่ 1 ใช้คะแนนเฉลี่ย (Mean) วิธีที่ 2 ใช้ค่ามัธยฐาน (Median) วิธีที่ 3 ใช้ทั้งค่าเฉลี่ยและค่ามัธยฐานร่วมกันโดยตัดค่าสูงสุดและค่าต่ำสุดออก วิธีการดังกล่าวนี้เรียกว่า Trimmed Mean ดังตัวอย่างในตาราง 3 สมมุติว่ามีข้อคำถามอยู่ 10 ข้อ และให้ผู้เชี่ยวชาญ 5 คน ตัดสินได้ผลดังนี้

ตาราง 3 ผลการตัดสินโดยใช้เทคนิคของแองกอฟฟ์

ข้อคำถาม	ผู้ตัดสิน				
	A	B	C	D	E
1	.95	.90	1.00	.85	.90
2	.80	.80	1.00	.80	.85
3	.90	.85	1.00	.80	.85
4	.60	.65	.95	.70	.70
5	.75	.70	.90	.75	.75
6	.40	.60	.90	.65	.65
7	.50	.65	.90	.60	.70
8	.25	.30	.85	.45	.55
9	.25	.25	.80	.30	.50
10	.40	.30	.75	.25	.45
รวม	5.80	6.00	9.05	6.15	6.90

ค่าเฉลี่ยเท่ากับ 6.78

ค่ามัธยฐานเท่ากับ 6.15

Trimmed Mean เท่ากับ 6.35 (โดยตัดค่าสูงสุดและค่าต่ำสุดออก)

3. เทคนิคของอีเบล

อีเบลได้กล่าวถึงจุดอ่อนของการนิยามคะแนนจุดตัดหรือคะแนนผ่านโดยการพิจารณาจากร้อยละของคะแนนรวมทำให้ขาดรายละเอียดอื่น ๆ ที่จะนำมาใช้ในการพิจารณาด้วย เนื่องจากในบางครั้งข้อสอบอาจจะมีความยากมากหรือง่ายเกินไป หรือมีอำนาจจำแนกต่ำมากเกินกว่าที่ผู้ออกข้อสอบคาดหวังไว้ ดังนั้นการที่ผู้สอบจะสอบผ่านหรือไม่จึงขึ้นอยู่กับคำถามในแบบทดสอบมากกว่าการพิจารณาจากระดับความสามารถของผู้สอบ อีเบลจึงเสนอวิธีการแก้ไขจุดอ่อนดังกล่าวโดยการวิเคราะห์ความเกี่ยวข้อง (Relevance) หรือความสำคัญ (Importance) ของแต่ละข้อคำถามในแบบทดสอบ สำหรับในเรื่องของความเกี่ยวข้อง อีเบลได้แบ่งออกเป็น 4 ประเภทคือ มีความจำเป็น (Essential) มีความสำคัญ (Important) เป็นที่ยอมรับได้ (Acceptable) และไม่แน่ใจ (Questionable) ส่วนในเรื่องของระดับความยาก อีเบลได้แบ่งเป็น 3 ระดับ คือ ง่าย (Easy) ปานกลาง (Medium) และยาก (Hard) และกำหนดร้อยละที่คาดหวังที่จะให้ผ่านข้อคำถามในแต่ละกลุ่ม ซึ่งร้อยละที่คาดหวังดังกล่าวนี้ถือว่าเป็นบริมาณความคาดหวังของผู้ที่มีความสามารถต่ำสุดจะสอบผ่าน ดังรายละเอียดในตาราง 4

ตาราง 4 ร้อยละที่คาดหวังจะสอบผ่านตามลักษณะที่เกี่ยวข้องและระดับความยาก

ลักษณะที่เกี่ยวข้อง	ระดับความยาก		
	ง่าย	ปานกลาง	ยาก
มีความจำเป็น	100%	-	-
มีความสำคัญ	90	70	-
เป็นที่ยอมรับได้	90	60	40%
ไม่แน่ใจ	70	50	30%

เทคนิคของอีเบล นั้น สามารถสรุปได้ดังนี้

1. ให้ผู้เชี่ยวชาญแต่ละคนพิจารณาข้อคำถาม แล้วจัดแยกเป็น 12 ประเภท (คือ ลักษณะที่เกี่ยวข้องมี 4 ลักษณะ และระดับความยากมี 3 ระดับ)

2. อภิปรายถึงแต่ละข้อคำถามดังนี้

2.1 ถามผู้เชี่ยวชาญว่ามีกี่คนที่จัดคำถามอยู่ในระดับง่าย ปานกลาง และยาก ถ้าผู้เชี่ยวชาญจัดไม่เหมือนกันก็ควรจะให้อธิบายเหตุผล

2.2 ถามผู้เชี่ยวชาญว่ามีกี่คนที่แยกประเภทของข้อคำถามออกเป็น มีความจำเป็น มีความสำคัญ เป็นที่ยอมรับได้ และไม่แน่ใจ ถ้าผู้เชี่ยวชาญแยกประเภทของข้อคำถามไม่เหมือนกัน ก็ควรให้อธิบายเหตุผลเช่นเดียวกัน

2.3 ให้โอกาสแก่ผู้เชี่ยวชาญจัดประเภทของข้อคำถามใหม่ ถ้าต้องการเปลี่ยนแปลง

3. ให้ผู้เชี่ยวชาญกำหนดร้อยละของผู้สอบที่คาดหวังว่า จะตอบในแต่ละข้อคำถามได้ถูกต้องทั้ง 12 ประเภท

4. รวบรวมผลการตัดสิน

วิธีการหาคะแนนจุดตัดของอีเบลก็คือ นำจำนวนข้อสอบในแต่ละลักษณะคูณกับค่าร้อยละที่คาดหวัง แล้วนำผลคูณที่ได้มาหาค่าเฉลี่ย ค่าเฉลี่ยที่ได้ก็จะเป็นคะแนนจุดตัด หรือคะแนนที่ผู้มีความสามารถต่ำสุดจะสอบผ่าน ดังตัวอย่างในตาราง 5

ตาราง 5 ตัวอย่างของการคำนวณตามเทคนิคของอีเบล

ลักษณะของข้อคำถาม	จำนวนของข้อคำถาม	ร้อยละที่คาดหวัง	จำนวนข้อคำถามคูณร้อยละที่คาดหวัง
มีความจำเป็น	94	100	9400
มีความสำคัญ			
ง่าย	106	90	9540
ปานกลาง	153	70	10710
เป็นที่ยอมรับได้			
ง่าย	24	80	1920
ปานกลาง	49	60	2940
ยาก	52	40	2080
ไม่แน่ใจ			
ง่าย	4	70	280
ปานกลาง	11	50	550
ยาก	7	30	210
รวม	500		37130

จากตาราง 5 จะได้คะแนนจุดตัดเท่ากับ 37130/500 = 74.26 หรือ 74%

5. การพิจารณาจากผู้สอบเป็นรายบุคคล

วิธีการนี้จะต้องอาศัยข้อมูลจากผู้สอบเป็นรายบุคคลอันได้แก่ คะแนนของผู้สอบและการพิจารณาถึงความรู้และทักษะขั้นต่ำของผู้สอบ วิธีการนี้แบ่งออกเป็น 2 วิธี คือ

1. วิธีการใช้กลุ่มคาบเส้น (Borderline-Groups Method)
2. วิธีการใช้กลุ่มตรงกันข้าม (Contrasting - Groups Method)

1. วิธีการใช้กลุ่มคาบเส้น

วิธีการนี้ขึ้นอยู่กับแนวคิดที่ว่า คะแนนจุดตัดควรจะเป็นคะแนนที่ได้จากผู้สอบที่มีความสามารถอยู่ในระดับคาบเส้น ข้อดีของวิธีการนี้ คือง่ายในการนำไปใช้และง่ายในการอธิบาย แต่ข้อจำกัดของวิธีการนี้คือค่อนข้างจะยากในการที่จะตัดสินว่า ผู้สอบอยู่ในกลุ่มที่มีลักษณะเป็นกลุ่มคาบเส้นจริง ๆ อย่างไรก็ตามหากจะนำวิธีการนี้ไปใช้ มีขั้นตอนดังนี้

1. เลือกผู้ตัดสิน
2. นิยามความหมายของคำว่า ระดับความรู้และทักษะของผู้สอบที่อยู่ในระดับคาบเส้น
3. ระบุกลุ่มผู้สอบที่เป็นกลุ่มคาบเส้น
4. เก็บรวบรวมคะแนนที่ได้จากการสอบของผู้สอบในกลุ่มคาบเส้น
5. กำหนดคะแนนมัธยฐานของผู้สอบในกลุ่มคาบเส้นให้เป็นคะแนนจุดตัด

2. วิธีการใช้กลุ่มตรงกันข้าม

วิธีการนี้ขึ้นอยู่กับแนวคิดที่ว่าจะต้องแบ่งผู้สอบออกเป็น 2 กลุ่มที่มีลักษณะตรงกันข้ามกัน คือกลุ่มที่ได้คะแนนสูงกับกลุ่มที่ได้คะแนนต่ำโดยมีขั้นตอนในการประยุกต์เพื่อนำไปใช้ดังนี้

1. เลือกผู้ตัดสิน
2. ระบุความรู้และทักษะของผู้สอบที่คิดว่าเพียงพอและไม่เพียงพอ
3. เลือกกลุ่มผู้สอบที่มีความรู้และทักษะที่ระบุในข้อ 2
4. เก็บรวบรวมคะแนนที่ได้จากการสอบ และการตัดสินโดยอย่าให้ผู้ตัดสินรู้คะแนนของผู้สอบ

5. แบ่งผู้สอบออกเป็น 2 กลุ่ม คือกลุ่มที่ได้คะแนนสูง และกลุ่มที่ได้คะแนนต่ำ ต่อจากนั้นจึงคำนวณหาร้อยละของผู้สอบในแต่ละระดับคะแนน

6. ใช้วิธีการที่เรียกว่า Smoothing เพื่อหาคะแนนจุดตัด

สำหรับวิธีการใช้กลุ่มตรงกันข้ามนี้ สามารถหาคะแนนจุดตัดได้อีกวิธีหนึ่ง คือ นำคะแนนของผู้สอบที่จัดว่าอยู่ในกลุ่มรอบรู้และกลุ่มไม่รอบรู้มาแจกแจง จุดที่คะแนนการแจกแจงของผู้สอบทั้ง 2 กลุ่มตัดกันก็จะเป็นคะแนนจุดตัด ซึ่งจะเห็นว่าวิธีการใช้กลุ่มตรงกันข้ามนี้คล้ายกับวิธีการของเบอร์ก (Berk. 1976 : 4 - 9) เพียงแต่ว่าวิธีการของเบอร์กนั้นจะใช้กลุ่มที่ได้รับการสอนกับกลุ่มที่ไม่ได้รับการสอน และเสนอวิธีการทางสถิติในการแปลความหมายของผลที่ได้ สำหรับในการนำไปใช้นั้น วิธีการของเบอร์กจะมีปัญหาคือ ไม่สามารถหากลุ่มที่มีความสามารถเหมือนกันทั้งกลุ่มที่ได้รับการสอนและกลุ่มที่ไม่ได้รับการสอนได้ (Berk. 1980 : 107)

6. <u>การพิจารณาจากกลุ่มผู้สอบ</u>

วิธีการนี้ขึ้นอยู่กับการพิจารณาตัดสินเกี่ยวกับกลุ่มผู้สอบ ซึ่งกลุ่มนี้จะเรียกว่า กลุ่มอ้างอิง (Reference Group) วิธีการที่ง่ายที่สุดของวิธีนี้คือ การกำหนดคะแนนจุดตัดตามจำนวนคนที่ยอมให้ผ่านของผู้สอบในกลุ่มอ้างอิง เช่น ถ้าผู้ตัดสินมีเหตุผลที่จะเชื่อว่า ร้อยละ 85 ของผู้สอบเมื่อปีที่ผ่านมาเป็นพวกที่มีคุณภาพ (Qualified) ก็อาจจะใช้คะแนนจุดตัดจากผู้สอบในปีที่ผ่านมาเป็นเกณฑ์ แต่ถ้ากรณีที่ได้มีการเปลี่ยนแปลงข้อสอบทุกปีก็จะต้องหาคะแนนของการสอบในปีนี้ให้เทียบเท่ากับร้อยละ 85 ของผู้สอบเมื่อปีที่ผ่านมาโดยใช้วิธีการทางสถิติที่เรียกว่า Equating เข้าช่วย อย่างไรก็ตามหากจะนำวิธีการนี้ไปประยุกต์ใช้มีขั้นตอนดังนี้

1. เลือกผู้ตัดสินขึ้นมาจำนวนหนึ่ง

2. ให้ผู้ตัดสินพิจารณาว่า กลุ่มผู้สอบที่เคยสอบข้อสอบฉบับเดียวกันในปีที่ผ่านมา (หรือที่เรียกว่า กลุ่มอ้างอิง ซึ่งจะต้องเป็นกลุ่มที่มีขนาดใหญ่พอสมควร) ที่เป็นผู้มีระดับความรู้และทักษะที่จัดว่าอยู่ในระดับที่เพียงพอ (Adequate) มีจำนวนร้อยละเท่าไร

3. หาคะแนนของกลุ่มผู้สอบที่มีความรู้และทักษะในระดับที่เพียงพอของปีที่ผ่านมาว่ามีคะแนนโดยเฉลี่ยเป็นเท่าไร คะแนนที่ได้ก็จะเป็นคะแนนจุดตัด แต่ถ้าได้มีการเปลี่ยนแปลงแบบทดสอบใหม่ก็จะต้องใช้วิธีการ Equating คะแนนทั้ง 2 ชุด เพื่อกำหนดคะแนนจุดตัด

7. <u>การใช้ทฤษฎีการตัดสินใจ</u>

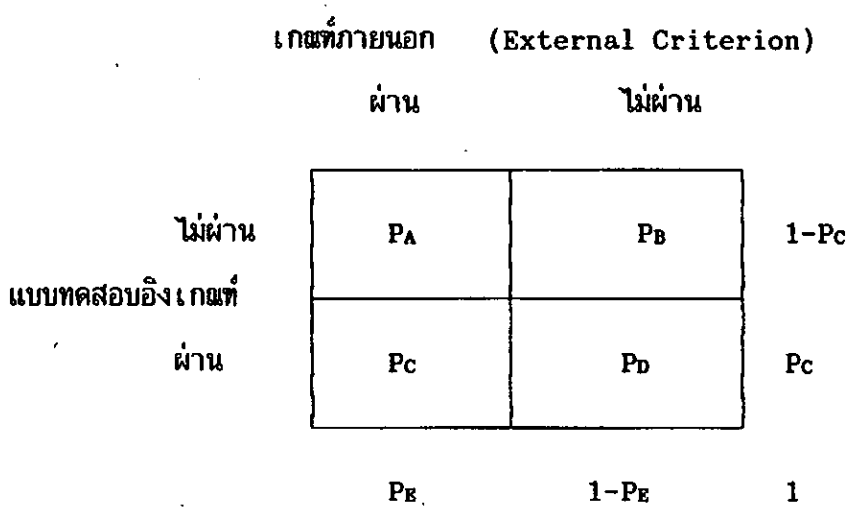
การใช้ทฤษฎีการตัดสินใจในการกำหนดคะแนนจุดตัดมีดังนี้

7.1 เทคนิคของกลาส

7.2 เทคนิคของเบส์

7.1 <u>เทคนิคของกลาส</u>

กลาส (Glass. 1978 : 251 - 253) ได้เสนอวิธีการใช้ทฤษฎีการตัดสินใจในการกำหนดคะแนนจุดตัดด้วยการเริ่มจาก การแบ่งผู้สอบออกเป็น 2 กลุ่ม โดยอาศัยเกณฑ์ภายนอกบางอย่างที่เกี่ยวข้องกับที่เราสนใจจะศึกษาเป็นหลัก เช่น ผู้ที่จบการศึกษาจากวิทยาลัย และผู้ที่ไม่จบการศึกษาจากวิทยาลัย ผู้ที่ได้รับการจ้าง และไม่ได้รับการจ้าง เป็นต้น สัดส่วนของผู้สอบทั้ง 2 กลุ่มนี้แทนด้วย P_E และ $1-P_E$ แล้วนำแบบทดสอบอิงเกณฑ์มาทดสอบผู้สอบทั้ง 2 กลุ่ม โดยกำหนดให้ C_x เป็นคะแนนเกณฑ์หรือคะแนนจุดตัดที่แบ่งผู้สอบออกเป็นกลุ่มผู้สอบผ่าน และกลุ่มผู้สอบไม่ผ่านก็จะได้สัดส่วนของจำนวนคน ดังภาพประกอบ 2



ภาพประกอบ 2 ความสัมพันธ์ระหว่างการตัดสินด้วยแบบทดสอบอิงเกณฑ์และเกณฑ์ภายนอก

จากภาพประกอบ 2 P_A แทนสัดส่วนของผู้ที่ไม่ผ่านแบบทดสอบอิงเกณฑ์ แต่ผ่านเกณฑ์ภายนอก ซึ่งเป็นการจำแนกผิดทางลบ (False Negative)

P_D แทนสัดส่วนของผู้ที่ผ่านแบบทดสอบอิงเกณฑ์ แต่ไม่ผ่านเกณฑ์ภายนอก ซึ่งเป็นการจำแนกผิดทางบวก (False Postive)

P_B แทนสัดส่วนของผู้ที่ไม่ผ่านทั้งแบบทดสอบอิงเกณฑ์ และเกณฑ์ภายนอก ซึ่งเป็นการจำแนกที่ถูก

P_C แทนสัดส่วนของผู้ที่ผ่านทั้งแบบทดสอบอิงเกณฑ์ และเกณฑ์ภายนอก ซึ่งเป็นการจำแนกที่ถูก

คะแนนจุดตัดที่กำหนดโดยเกณฑ์ภายนอกจะไม่เปลี่ยนแปลง แต่คะแนนจุดตัดของแบบทดสอบอิงเกณฑ์จะแปรผันไปตามค่า P_A, P_B, P_C และ P_D เพราะฉะนั้นการหาคะแนนจุดตัดจะต้องแทนค่าต่าง ๆ ลงในสมการ

$$f(C_X) = (P_A + P_D) / (P_B + P_C).....(1)$$

ค่า $f(C_X)$ ที่น้อยที่สุดก็จะเป็นคะแนนจุดตัด ซึ่งในสมการ (1) มีข้อตกลงว่าค่าของการจำแนกผิดทางบวก และการจำแนกผิดทางลบมีความสำคัญเท่ากัน แต่ถ้าให้ความสำคัญของการจำแนกทั้งสองไม่เท่ากัน ค่าฟังก์ชันจะเปลี่ยนแปลงไป ดังนั้นจึงต้องมีการปรับแก้โดยกำหนดค่าถ่วงน้ำหนักสองค่าขึ้นมา คือให้ α เป็นค่าความสำคัญของการจำแนกผิดทางลบ และ β เป็นค่าความสำคัญของการจำแนกผิดทางบวก ดังนั้นจะได้ฟังก์ชันใหม่ดังนี้

$$f(C_X) = (\alpha P_A + \beta P_D) / (P_B + P_C).....(2)$$

ค่า α และ β นี้ จะกำหนดขึ้นจากการใช้ดุลพินิจของผู้ตัดสินว่า จะให้ความสำคัญมากน้อยเพียงใด เพื่อให้ $f(C_X)$ มีความถูกต้องต่อการกำหนดคะแนนจุดตัด

7.2 <u>เทคนิคของเบส์</u>

สวามินาธาน แฮมเบิลตัน และอัลไจนา (Swaminathan, Hambleton and Algina. 1975 : 87 - 98) ได้เสนอการกำหนดคะแนนจุดตัดโดยใช้เทคนิคของเบส์ ดังนี้

1. แปลงคะแนนที่ได้จากการสอบของผู้สอบ (X_i) ให้เป็น g_i โดยแปลงแบบอาร์คไซน์ (Arc Sine) เพื่อให้การแจกแจงเป็นปกติ ซึ่งจะช่วยให้การคำนวณสะดวกขึ้น โดยใช้สูตร

$$g_i = \sin^{-1} \{(X_i + 3/8) / (n + 3/4)\}^{1/2}$$

เมื่อ n คือจำนวนข้อคำถามของแบบทดสอบ

การแจกแจงนี้มีค่าเฉลี่ยเท่ากับ γ_i และความแปรปรวนเท่ากับ ν โดยที่

$$\gamma_i = \sin^{-1}\sqrt{\pi_i}$$

$$\nu = (4n + 2)^{-1}$$

การประมาณนี้จะใช้ได้ดีเมื่อ π มีค่าระหว่าง .15 ถึง .85 และจำนวนข้อคำถาม (n) ไม่ต่ำกว่า 8 ข้อ เนื่องจากการแจกแจงของ g_i เราทราบค่าความแปรปรวน แต่เราไม่ทราบค่า π_i การแจกแจงของ g_i จึงเป็นการแจกแจงที่มีเงื่อนไข คือ

$$g_i / \gamma_i \sim N(\gamma_i, \nu)$$

เมื่อ $N(\gamma_i, \nu)$ แทนการแจกแจงปกติ โดยมีค่าเฉลี่ยเท่ากับ γ_i และความแปรปรวนเท่ากับ ν การแจกแจงของ g_i จึงขึ้นอยู่กับ γ_i เพื่อให้ได้การแจกแจงภายหลังของ γ_i ต้องมีข้อตกลงว่า γ_i เป็นตัวอย่างที่ได้จากการสุ่ม โดยสมมุติว่าการแจกแจงของ γ_i เป็นปกติด้วยค่าเฉลี่ย α และความแปรปรวน ϕ ซึ่งเราไม่ทราบทั้งค่า α และ ϕ ดังนั้นการแจกแจงของ γ_i จึงขึ้นอยู่กับ α และ ϕ อย่างไรก็ตาม ไนวิคและคนอื่น ๆ (Glass. 1975 : 91 ; citing Novick and others. 1973) แนะนำว่าเราสามารถหา ϕ ได้จากการแจกแจงแบบอินเวอร์ส ไค-สแคร์ (InverseChi-Square) ที่มีองศาแห่งความเป็นอิสระเท่ากับ ν และสเกลพารามิเตอร์ λ

นั่นคือ $\bar{\phi} = \lambda / (\nu - 2)$

$$\lambda = (\nu - 2)\,\bar{\phi}$$

ค่า ν ที่เหมาะสมคือ 8 ส่วนการประมาณค่า $\bar{\phi}$ จะต้องมีข้อมูลเกี่ยวกับ π ดังนั้น จึงต้องกำหนดจำนวนข้อคำถาม (t) เพื่อให้ทราบข้อมูลเกี่ยวกับค่า π ค่าประมาณที่ได้จากการแปลงค่า π จากแบบทดสอบจำนวน t ข้อ จะมีการแจกแจงแบบปกติด้วยความ

แปรปรวนเท่ากับ $(4t + 2)^{-1}$ ดังนั้นเราสามารถใช้ค่า $(4t + 2)^{-1}$ เป็นค่าประมาณของ ϕ และ λ ได้

เลวิส และคนอื่น ๆ (Glass. 1978 : 91 ; citing Lewis and others. 1973) ได้แสดงให้เห็นว่า การแจกแจงของ γ_i มีการแจกแจงเป็นปกติดังนี้

$$\gamma_i/\text{Data} \sim N(\mu_i, \sigma_i^2)$$

เมื่อ

$$\mu_i = \bar{g} + \rho^*(g_i - \bar{g})$$

และ

$$\sigma_i^2 = \frac{1 + (m - 1)\rho^* + (g_i - \bar{g})^2\sigma^{*2}}{(4n + 2)m}$$

$$\bar{g} = m^{-1} \sum_{i=1}^{m} g_i$$

เมื่อ m คือ จำนวนผู้เข้าสอบ

n คือ จำนวนข้อคำถามในแบบทดสอบ

สำหรับค่า ρ^* และ σ^2 ได้จากตารางแวง (Wang) เมื่อทราบค่า ν และ λ

2. กำหนดคะแนนจุดตัด (π_o) เพื่อแบ่งผู้สอบเป็นกลุ่มผู้รอบรู้และไม่รอบรู้ แล้วแปลง π_o เป็น γ_o จากสูตร

$$\gamma_o = \sin^{-1}\sqrt{\pi_o}$$

3. หาคะแนนมาตรฐานปกติจากสูตร

$$Z_{oi} = \frac{\gamma_o - \mu_1}{\sigma_1}$$

4. หาค่าความน่าจะเป็นของการสอบผ่านและไม่ผ่านของคะแนนที่ได้จากผู้สอบซึ่งสามารถหาได้จากตารางพื้นที่ใต้โค้งปกติดังนี้

Prob [Z<Z_{oi}/Data] แทนความน่าจะเป็นในการสอบไม่ผ่าน

Prob [Z≥Z_{oi}/Data] แทนความน่าจะเป็นในการสอบผ่าน

5. หาความสูญเสียที่คาดหวังของการตัดสินให้ผ่านและไม่ผ่านของคะแนนจากผู้สอบ ดังนี้

l_{12} Prob [Z<Z_{oi}/Data] แทนความสูญเสียที่คาดหวังของการตัดสินให้ผ่าน

l_{21} Prob [Z≥Z_{oi}/Data] แทนความสูญเสียที่คาดหวังของการตัดสินให้ไม่ผ่าน

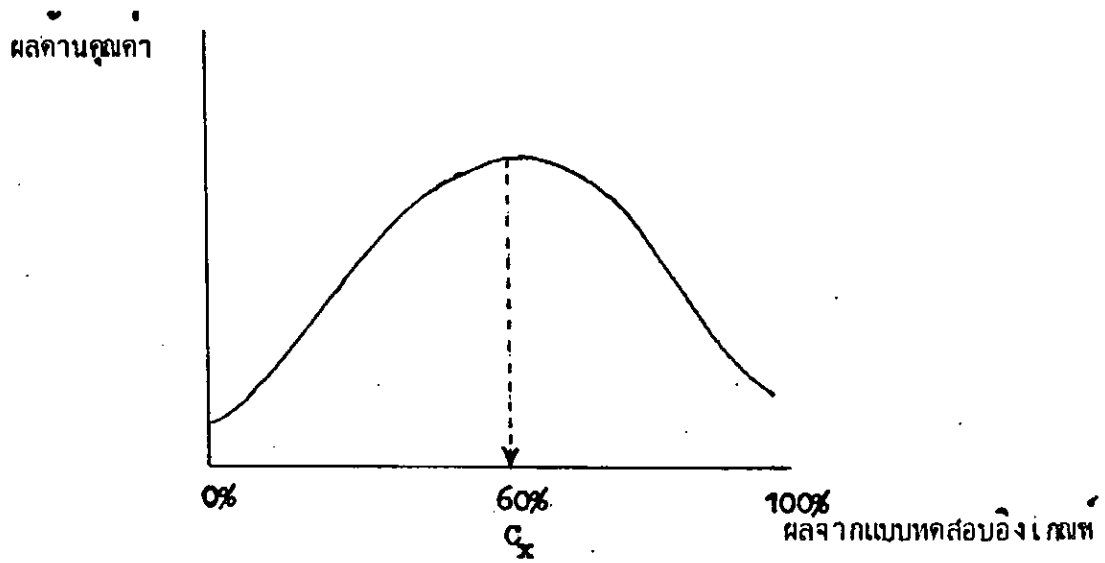
เมื่อ l_{12} แทนความสูญเสียที่เกิดจากความคลาดเคลื่อนในการจำแนกผิดทางบวก

l_{21} แทนความสูญเสียที่เกิดจากความคลาดเคลื่อนในการจำแนกผิดทางลบ

6. เลือกคะแนนที่น้อยที่สุดที่ทำให้ค่า l_{21} Prob [Z≥Z_{oi}/Data] มากกว่าค่า l_{12} Prob [Z<Z_{oi}/Data] เป็นคะแนนจุดตัด

8. <u>การใช้วิธีการวิจัยเชิงปฏิบัติการ</u>

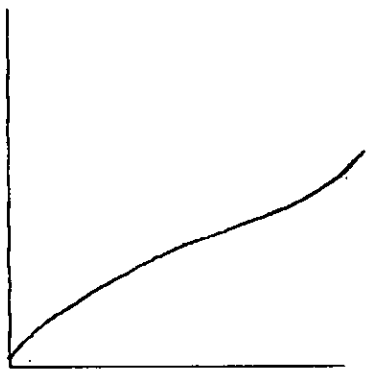
การใช้วิธีนี้ในการกำหนดคะแนนเกณฑ์โดยการพิจารณาจากลักษณะกราฟ คะแนนที่ตรงกับจุดสูงสุดของกราฟจะถือเป็นคะแนนเกณฑ์ ซึ่งบล็อค (Glass. 1978 : 253 ; citing Block. 1972) ได้ใช้วิธีนี้ในการกำหนดคะแนนเกณฑ์ของแบบทดสอบอิงเกณฑ์ โดยเขาสุ่มกลุ่มที่สอนที่มีลักษณะเท่าเทียมกันจนมีผลสัมฤทธิ์จากการทำแบบทดสอบอิงเกณฑ์ในระดับต่าง ๆ เช่น 10 %, 15 %, 20 %,...,95 %, 100 % แล้ววัดผลด้านคุณค่าอย่างหนึ่งอย่างใดที่สัมพันธ์กับคะแนนจากแบบทดสอบอิงเกณฑ์ของแต่ละกลุ่มที่มีผลสัมฤทธิ์ระดับต่าง ๆ กันนั้น นำผลที่ได้มาเขียนกราฟแสดงความสัมพันธ์เพื่อพิจารณาค่าของคะแนนเกณฑ์ ถ้ากราฟเป็นนอนโมโนโทนิค (Non-Monotonic) คะแนนเกณฑ์จะอยู่ระหว่าง 0 % และ 100 % ถ้ากราฟที่ได้มีลักษณะเป็นโมโนโทนิค คือไม่มีจุดที่กราฟจะโค้งกลับ คะแนนเกณฑ์จะเป็น 100 % ดังภาพประกอบ 3



ภาพประกอบ 3 ความสัมพันธ์ระหว่างผลจากแบบทดสอบอิงเกณฑ์และการวัดคุณค่า

จากภาพประกอบ 3 มักจะเป็นไปไม่ได้ เนื่องจากเส้นกราฟจะเป็น นอนโมโนโทนิคนั้นเป็นไปได้ยากเว้นแต่ว่าทั้งการสอบแบบอิงเกณฑ์และผลจากการวัดคุณค่าเป็นการวัดด้านความรู้ บล็อคได้เสนอวิธีแก้ปัญหาเช่นนี้คือ ให้ใช้ผลด้านคุณค่าด้านอื่นประกอบการพิจารณา ซึ่งอาจจะเป็นด้านความรู้สึก (Affective) เช่น ความสนใจหรือทัศนคติที่มีต่อวิชานั้น ๆ ดังภาพประกอบ 4

คุณค่าด้านความรู้

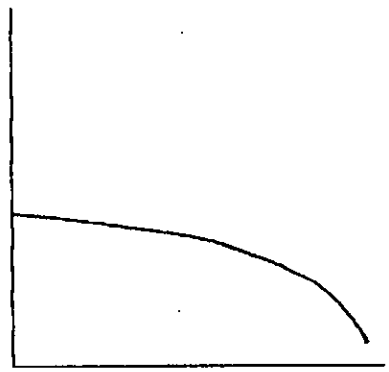


0 % 100 %

คะแนนจากแบบทดสอบอิงเกณฑ์

(CRT Score)

คุณค่าด้านความรู้สึก



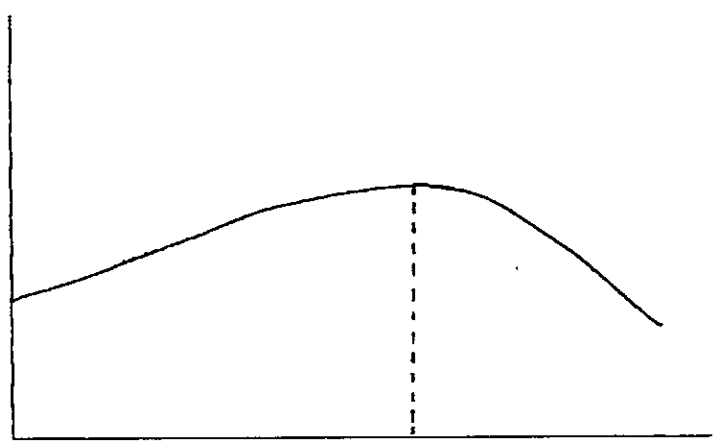
0 % 100 %

คะแนนจากแบบทดสอบอิงเกณฑ์

(CRT Score)

ผลจากด้านความรู้

รวมกับความรู้สึก



0 % Cx 100 %

คะแนนจากแบบทดสอบอิงเกณฑ์

(CRT Score)

ภาพประกอบ 4 ความสัมพันธ์ระหว่างคะแนนจากแบบทดสอบอิงเกณฑ์และผลที่ได้รับ

จากภาพประกอบ 4 การพิจารณาคุณค่ารวม หรือผลจากด้านความรู้รวมกับความรู้สึกนั้น ควรต้องมีการกำหนดน้ำหนักให้กับผลของคุณค่าทั้งในด้านความรู้และความรู้สึกด้วย โดยการกำหนดน้ำหนัก α ให้กับผลด้านความรู้ และกำหนดน้ำหนัก β ให้กับผลด้านความรู้สึก ดังนั้น

ผลลัพธ์ด้านคุณค่ารวม = α (ผลด้านความรู้) + β (ผลด้านความรู้สึก)

นอกจากการกำหนดคะแนนจุดตัดด้วยวิธีดังกล่าวแล้ว ยังมีวิธีการกำหนดคะแนนจุดตัดโดยการผสมผสาน (Compromise) กันระหว่างเกณฑ์เชิงสัมพัทธ์ (Relative) กับเกณฑ์เชิงสัมบูรณ์ (Absolute) ซึ่งบุค (Beuk. 1984 : 147) ได้ให้ความหมายเกณฑ์เชิงสัมพัทธ์และเกณฑ์เชิงสัมบูรณ์ดังนี้

1. เกณฑ์เชิงสัมพัทธ์ คะแนนจุดตัดหรือมาตรฐานที่ได้จากวิธีการนี้จะขึ้นอยู่กับการแจกแจงของคะแนนผู้สอบ กล่าวคือ หากคะแนนสอบสูงคะแนนจุดตัดก็จะสูงด้วย

2. เกณฑ์เชิงสัมบูรณ์ คะแนนจุดตัดหรือมาตรฐานที่ได้จากวิธีการนี้จะแปลความหมายโดยอาศัยเกณฑ์ภายนอก (External Criterion) กล่าวคือ คะแนนจุดตัดจะไม่ขึ้นอยู่กับคะแนนที่ได้จากผู้สอบ

เด กรุจเตอร์ (De Gruijter. 1985 : 263 - 269) ได้เสนอแบบจำลองในการกำหนดคะแนนจุดตัด โดยใช้วิธีการผสมผสานระหว่างเกณฑ์เชิงสัมพัทธ์และเกณฑ์เชิงสัมบูรณ์ดังนี้

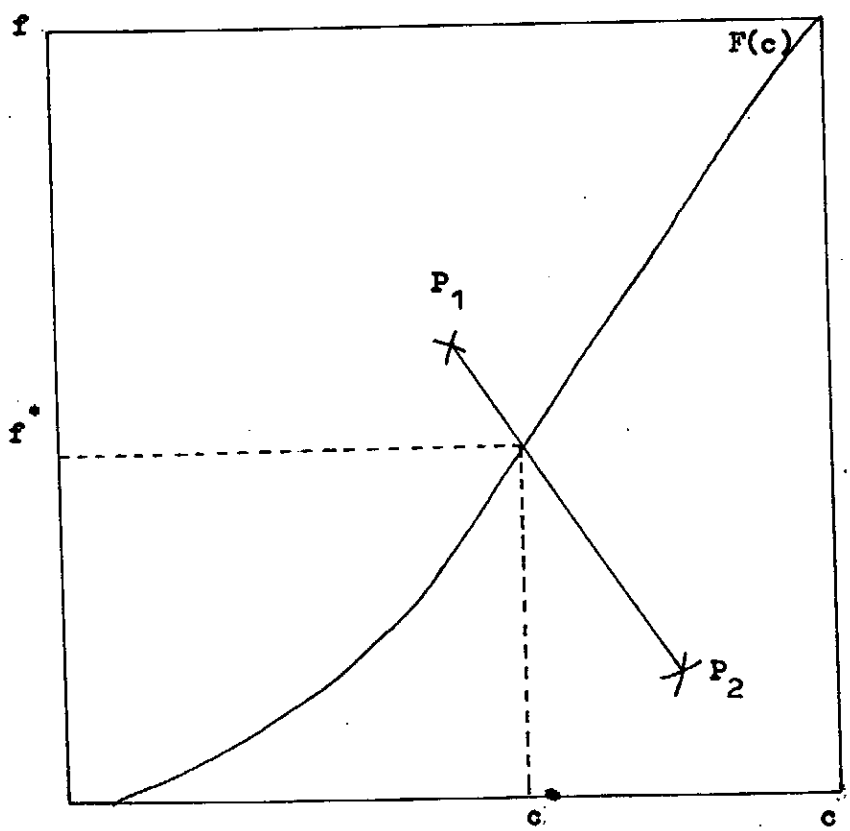
1. <u>แบบจำลองของฮอฟสที</u> (Hofstee's Model)

ฮอฟสที ได้เสนอแบบจำลองในการหาคะแนนจุดตัดโดยให้แกน X แทนคะแนนจุดตัด (c) หรือร้อยละของการตอบแบบทดสอบถูก และแกน Y แทนร้อยละของการไม่ผ่านในการพิจารณาคะแนนจุดตัดจะต้องหาค่าต่าง ๆ ดังนี้

1. คะแนนผ่านสูงสุด (C_{max})
2. คะแนนผ่านต่ำสุด (C_{min})
3. ร้อยละของการยอมรับว่าไม่ผ่านสูงสุด (f_{max})
4. ร้อยละของการยอมรับว่าไม่ผ่านต่ำสุด (f_{min})

จุด P_1 (C_{min}, f_{max}) และ P_2 (C_{max}, f_{min}) จะแทนการรวมค่า c และ f ที่ยอมรับได้

ความสัมพันธ์เชิงประจักษ์ระหว่าง c และร้อยละของการไม่ผ่าน เขียนได้ว่า $f = F(c)$ ซึ่งเป็นโค้งความถี่สะสม เพราะฉะนั้นค่าที่ได้จากจุดตัดของโค้ง $F(c)$ กับเส้น P_1 P_2 ก็จะเป็นคะแนนจุดตัด (c^*) ซึ่งเป็นเกณฑ์มาตรฐานที่ได้ระหว่างวิธีการเชิงสมบูรณ์และเชิงสัมพัทธ์ ดังภาพประกอบ 5



ภาพประกอบ 5 แบบจำลองของฮอฟสที

2. <u>แบบจำลองของบุค</u> (Beuk's Model)

บุค (Beuk. 1984 : 147 - 152) ได้เสนอแบบจำลองในการหาคะแนนจุดตัด ดังนี้

ให้ k คือ ระดับความรู้ขั้นต่ำของผู้สอบ ซึ่งกำหนดในรูปร้อยละของคะแนนดิบที่ผู้สอบจะสอบผ่าน

S_k คือ ส่วนเบี่ยงเบนมาตรฐานของการตัดสินคะแนนผ่าน

$\bar{k}$ คือ ค่าเฉลี่ยของการตัดสินการผ่าน

v คือ อัตราการผ่านซึ่งกำหนดในรูปของร้อยละ

S_v คือ ส่วนเบี่ยงเบนมาตรฐานของอัตราการผ่าน

$\bar{v}$ คือ ค่าเฉลี่ยของอัตราการผ่าน

ดังนั้น ความสัมพันธ์ระหว่าง v และ k สามารถเขียนได้ดังนี้

$$v = (S_v/S_k)(k-\bar{k}) + \bar{v}$$

ให้ $M(\bar{k},\bar{v})$ แทนค่า $v = (S_v/S_k)(k-\bar{k}) + \bar{v}$ ซึ่งเป็นคะแนนผ่านที่ตัดสินโดยผู้เชี่ยวชาญ

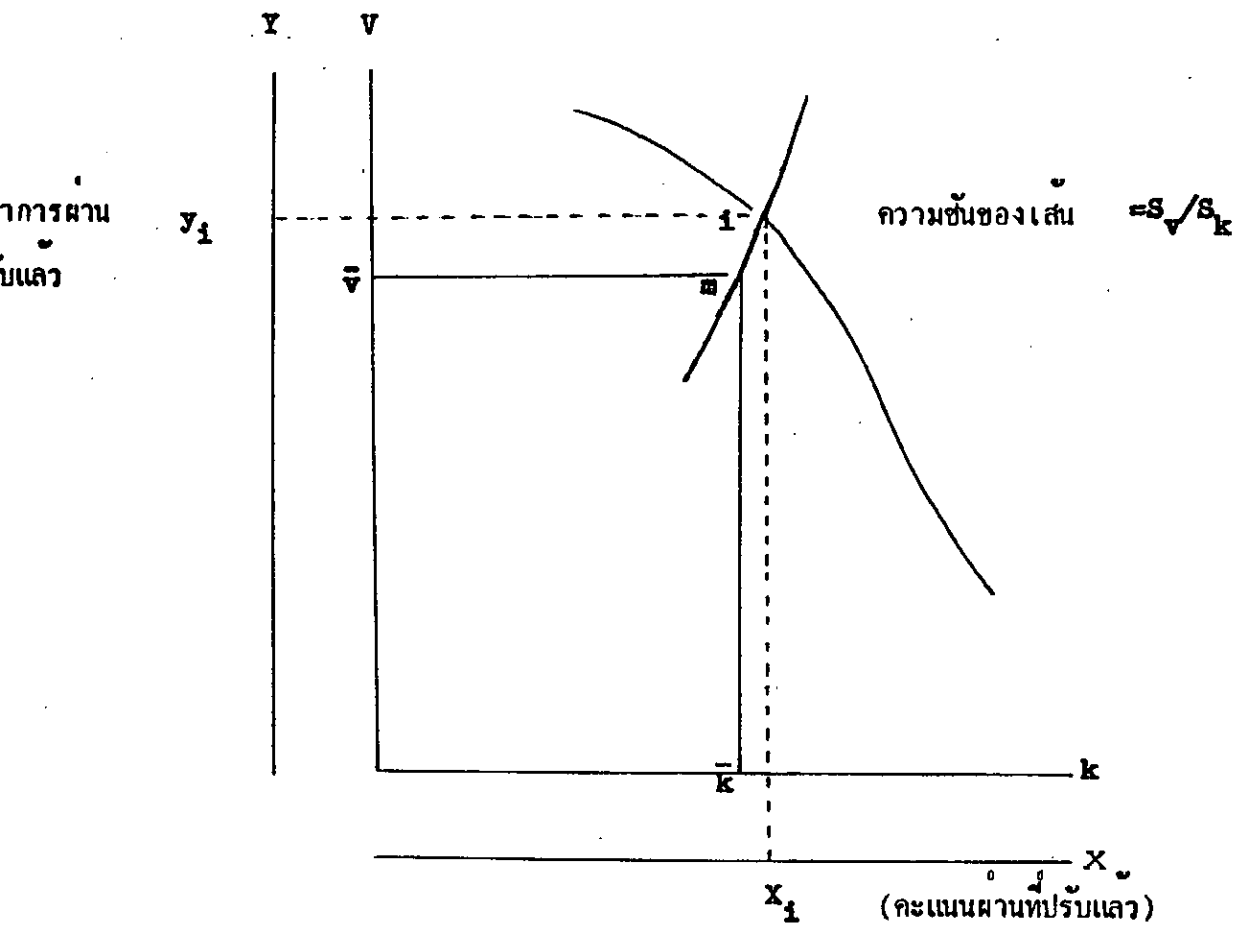
หากกล่าวโดยสรุปแบบจำลองของบุค มีดังนี้

1. ให้แกน X แทนคะแนนผ่าน และแกน Y แทนอัตราการผ่าน และจุด M แทนค่าเฉลี่ยการตัดสินของผู้เชี่ยวชาญในการพิจารณาระดับความรู้ขั้นต่ำ และค่าเฉลี่ยการตัดสินของร้อยละที่คาดหวังของการผ่าน

2. ลากเส้นผ่าน M โดยมีค่าความชันเท่ากับ S_v/S_k

3. พล็อทจุดของอัตราการผ่าน (Y) และคะแนนการผ่าน (X) ก็จะได้โค้งดังภาพ

4. จุดที่โค้งในข้อ 3 และเส้นที่ผ่านจุด M ในข้อ 2 ตัดกันคือ คะแนนจุดตัด (i) ซึ่งเป็นจุดที่แสดงความสอดคล้องที่ดีที่สุดระหว่างเกณฑ์มาตรฐานเชิงสัมบูรณ์กับเกณฑ์มาตรฐานเชิงสัมพัทธ์ดังภาพประกอบ 6



ภาพประกอบ 6 แบบจำลองของบุค

<u>ทฤษฎีการตอบสนองต่อข้อสอบกับการกำหนดคะแนนจุดตัด</u>

ในเรื่องเกี่ยวกับทฤษฎีการตอบสนองต่อข้อสอบกับการกำหนดคะแนนจุดตัดผู้วิจัยจะเสนอรายละเอียดตามหัวข้อดังนี้

1. แนวคิดของทฤษฎีการตอบสนองต่อข้อสอบ
2. ข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองต่อข้อสอบ (Assumptions of Item Response Theory)
3. แบบจำลองของการตอบสนองต่อข้อสอบ (Item Response Model)

4. ความสัมพันธ์ของคะแนนความสามารถ (Ability Scores) กับคะแนนโดเมน (Domain Scores)

5. ความสัมพันธ์ระหว่างคะแนนโดเมนที่ได้จากการสังเกตและความสามารถ

6. การนำทฤษฎีการตอบสนองต่อข้อสอบมาประยุกต์ใช้ในการกำหนดคะแนนจุดตัด

7. การใช้วิธีการมินิแมกซ์ (Minimax) และวิธีการของราสช์ในการกำหนดคะแนนจุดตัดของแบบทดสอบย่อย

1. <u>แนวคิดของทฤษฎีการตอบสนองต่อข้อสอบ</u>

ทฤษฎีการตอบสนองต่อข้อสอบได้มีการพัฒนามาเป็นเวลานานแต่ก็ได้รับความสนใจน้อย เนื่องจากเป็นเรื่องที่ต้องอาศัยคณิตศาสตร์ขั้นสูง เน้นทฤษฎีมากกว่าการนำไปใช้ นอกจากนี้ยังไม่มีโปรแกรมคอมพิวเตอร์เพื่อใช้ในการประมาณค่าพารามิเตอร์ (Hambleton and Cook. 1977 : 75 - 76) บุคคลแรกที่กล่าวถึงทฤษฎีการตอบสนองต่อข้อสอบ คือ ริชาร์ด (Hambleton and Swaminathan. 1985 : 4 ; citing Richard. 1936) โดยกล่าวถึงความสัมพันธ์ระหว่างพารามิเตอร์ของแบบจำลองทฤษฎีการตอบสนองต่อข้อสอบและพารามิเตอร์จากแบบจำลองทฤษฎีคลาสสิค ต่อมาลอว์เลย์ (Hambletion and Swaminathan. 1985 : 4 ; citing Lawley. 1943, 1944) ได้เสนอวิธีการใหม่ในการประมาณค่าพารามิเตอร์ หลังจากนั้นจนถึงปี ค.ศ. 1952 ลอร์ด (Lord) ได้เสนอในวิทยานิพนธ์ระดับปริญญาเอกของเขา และขณะนั้นเขาเรียกทฤษฎีการตอบสนองต่อข้อสอบว่าทฤษฎีโค้งคุณลักษณะข้อสอบ (Item Characteristic Curve Theory) ต่อมาในปี ค.ศ. 1977 ลอร์ด ได้เปลี่ยนชื่อจาก ทฤษฎีโค้งคุณลักษณะข้อสอบมาเป็นทฤษฎีการตอบสนองต่อข้อสอบ (Warm. 1978 : 19 - 21)

ทฤษฎีการตอบสนองต่อข้อสอบหรือบางทีเรียกว่า ทฤษฎีคุณลักษณะแฝง (Latent Trait Theory) เป็นทฤษฎีที่อธิบายถึงความสัมพันธ์ระหว่างลักษณะหรือความสามารถที่มีอยู่ภายในตัวบุคคลซึ่งไม่สามารถจะสังเกตได้โดยตรงกับพฤติกรรมการตอบสนองต่อข้อสอบในลักษณะของแบบจำลองทางคณิตศาสตร์

ถ้าให้ Y คือ คะแนนที่ได้จากการตอบข้อสอบ ซึ่งเป็นพฤติกรรมการ ตอบสนองต่อข้อสอบ

θ คือ ความสามารถภายในตัวบุคคล

f คือ ฟังก์ชัน (Function)

ดังนั้นฟังก์ชันการถดถอยของ Y บน θ จะเป็นฟังก์ชันทางคณิตศาสตร์ที่แสดงว่าพฤติกรรมการตอบสนองต่อข้อสอบขึ้นอยู่กับระดับความสามารถภายในตัวบุคคล เขียนในรูปแบบจำลองทางคณิตศาสตร์ได้ว่า

$$Y = f(\theta)$$

จากแบบจำลองทางคณิตศาสตร์ดังกล่าว ลอร์ด (Lord. 1980 : 12) เรียกว่า ฟังก์ชันการตอบสนองต่อข้อสอบ (Item Response Function) หรือฟังก์ชันคุณลักษณะข้อสอบ (Item Characteristic Function) (Hamblelton and Swaminathan. 1985 : 13) สำหรับในการพิจารณาแบบจำลองทางคณิตศาสตร์ของทฤษฎีการตอบสนองต่อข้อสอบเริ่มจากการพิจารณาถึงลักษณะเส้นโค้งประจำข้อสอบโดยอธิบายด้วยเส้นโค้งที่เรียกว่า Normal Ogive ซึ่งเป็นเส้นโค้งในลักษณะของความถี่สะสมสามารถหาได้จากสมการดังนี้ (Warm. 1978 : 23 - 25)

$$\int_{-\infty}^{Z} N(0, 1) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{Z} e^{-Z^2/2} dZ$$

จากเส้นโค้งที่เรียกว่า Normal Ogive นั้น วอร์ม (Warm. 1978 : 24 - 26) แสดงให้เห็นว่าอาจใช้เส้นโค้งที่เรียกว่า Logistic Ogive แทนได้ เนื่องจากเส้นโค้ง Normal Ogive และ Logistic Ogive มีค่าใกล้เคียงกันมาก สำหรับเส้นโค้ง Logistic Ogive สามารถหาได้จากสมการดังนี้

$$\int_{-\infty}^{Z} L(0, 1.7) = \frac{1}{1 + e^{-1.7Z}}$$

จากสมการ Logistic Ogive จะเห็นได้ว่าทางด้านขวามือของสมการไม่มีเครื่องหมายอินทิเกรท (Integrate : $\int$) จึงคำนวนได้ง่ายกว่าการใช้สมการ Normal Ogive ดังนั้น วอร์ม (Warm. 1978 : 26) จึงเสนอให้ใช้เส้นโค้ง Logistic Ogive แทนเส้นโค้ง Normal Ogive

2. ข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองต่อข้อสอบ (Assumptions of Item Response Theory)

ข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองต่อข้อสอบมีดังนี้

2.1 ความเป็นอิสระในการตอบสนองต่อข้อสอบ (Local Independent) หมายความว่า การตอบสนองต่อข้อสอบข้อต่าง ๆ ในแบบทดสอบของผู้สอบมีความเป็นอิสระต่อกันในเชิงสถิติ กล่าวคือ การตอบสนองต่อข้อสอบข้อใดข้อหนึ่งของผู้สอบคนหนึ่ง ไม่มีผลต่อการตอบสนองต่อข้อสอบข้ออื่น ๆ ในแบบทดสอบ เช่น เนื้อหาของข้อสอบข้อหนึ่งต้องไม่ชี้แนะในการตอบข้อสอบข้ออื่นในแบบทดสอบ

คำว่า "ความเป็นอิสระในการตอบสนองต่อข้อสอบ" มิได้หมายความว่าข้อสอบแต่ละข้อมีสหสัมพันธ์กันเป็นศูนย์ แต่หมายความว่า ข้อสอบแต่ละข้อไม่มีสหสัมพันธ์กันถ้าผู้สอบแต่ละคนมีความสามารถ θ เหมือนกัน (Warm. 1978 : 99 - 100)

ถ้าการตอบสนองต่อข้อสอบมีความเป็นอิสระกันจริง ความน่าจะเป็นของคะแนนแต่ละข้อของผู้สอบจะเท่ากับผลคูณของความน่าจะเป็นของคะแนนแต่ละข้อ ตัวอย่าง เช่น ความน่าจะเป็นของการตอบข้อสอบ 5 ข้อ มีรูปแบบดังนี้ (Hambleton and Swaminathan. 1985 : 23)

$$U = (1\ 0\ 1\ 1\ 0)$$

โดยที่ 1 แทนการตอบถูก และ 0 แทนการตอบผิด เพราะฉะนั้นความน่าจะเป็นของการตอบข้อสอบ 5 ข้อเท่ากับ P_1, $(1-P_2)$, P_3, P_4, $(1-P_5)$ เมื่อ P_i คือความน่าจะเป็นในการตอบข้อ i ถูก และ $1-P_i$ คือความน่าจะเป็นในการตอบข้อ i ไม่ถูก

ในทำนองเดียวกัน ถ้ากำหนดให้ U_i แทนจำนวนข้อสอบที่มี n ข้อ โดยที่ i = 1, 2,..., n และเป็นข้อสอบที่มีคำตอบ 2 คำตอบ คือ ถ้าตอบถูกจะได้ 1 คะแนน ตอบผิดจะได้ 0 คะแนน P_i แทนความน่าจะเป็นของการตอบข้อสอบข้อที่ i ถูก Q_i = $1-P_i$แทนความน่าจะเป็นของการตอบข้อสอบข้อที่ i ผิด เพราะฉะนั้นความน่าจะเป็นที่ผู้สอบที่มีความสามารถเท่ากับ θ ทำข้อสอบจำนวน n ข้อ คือ

$$\begin{aligned}\text{Prob}[U_1 = u_1, U_2 = u_2, \ldots U_n = u_n \mid \theta] &= P_1(\theta)^{u_1} Q_1(\theta)^{1-u_1} P_2(\theta)^{u_2} Q_2(\theta)^{1-u_2} \ldots P_n(\theta)^{u_n} Q_n(\theta)^{1-u_n} \\ &= \prod_{i=1}^{n} P_i(\theta)^{u_i} Q_i(\theta)^{1-u_i}\end{aligned}$$

2.2 ความเป็นมิติเดียวของแบบทดสอบ (Unidimensionality) หมายความว่า ข้อสอบทุกข้อในแบบทดสอบจะต้องวัดความรู้หรือความสามารถในลักษณะเดียวกัน คำว่า "ความเป็นมิติเดียวในแบบทดสอบ" ไม่ได้หมายความว่า ข้อสอบแต่ละข้อจะต้องมีสหสัมพันธ์ในเชิงบวกเท่านั้น อาจจะมีสหสัมพันธ์กันในเชิงลบก็ได้ ตัวอย่างของแบบทดสอบที่มีความเป็นมิติเดียวที่เห็นได้ชัดเจนคือ แบบวัดความสามารถ เช่น ความสามารถทางภาษา คำนวณ มิติสัมพันธ์ ความเข้าใจเชิงจักรกล นอกจากนี้แล้วข้อสอบทุกข้อที่วัดความรู้ที่ได้จากการเรียนรู้ก็ถือว่ามีความเป็นมิติเดียวได้ อย่างเช่น ข้อสอบปลายภาค เป็นต้น จากการศึกษาของบีจาร์ วีสส์ และคิงสเบอรี่ (Warm. 1978 : 102 ; citing Bejar, Weiss and Kingsbury. 1977) ในรายวิชาชีววิทยาเบื้องต้น ซึ่งประกอบด้วยเนื้อหาสาระ 3 ส่วน คือ เคมี เซลและพลังงาน ก็พบว่า เป็นแบบทดสอบที่มีความเป็นมิติเดียวซึ่งแสดงให้เห็นว่าเนื้อหาวิชาเคมีเป็นสิ่งจำเป็นที่ทำให้เข้าใจในเรื่องเกี่ยวกับ เซล และเนื้อหาเกี่ยวกับเซลทำให้เข้าใจในเรื่องเกี่ยวกับพลังงาน ดังนั้นข้อสอบที่วัดความรู้อย่างเป็นเหตุเป็นผลกันและเป็นเนื้อหาที่มีความสัมพันธ์กันอย่างต่อเนื่องก็จะถือว่ามีความเป็นมิติเดียวกันได้ วิธีการที่จะตรวจสอบว่าแบบทดสอบมี

ความเป็นมิติเดียวกันหรือไม่สามารถตรวจสอบได้โดยการใช้วิธีการวิเคราะห์องค์ประกอบ (Factor Analysis) ความเป็นมิติเดียวของแบบทดสอบเป็นเงื่อนไขที่สำคัญของความเป็นอิสระในการตอบสนองต่อข้อสอบ ถ้าพบว่าแบบทดสอบมีความเป็นมิติเดียวก็แสดงว่า แบบทดสอบจะต้องมีความเป็นอิสระในการตอบสนองต่อข้อสอบด้วย (Warm. 1978 : 101 - 107)

2.3 โค้งคุณลักษณะข้อสอบ (Item Characteristic Curves : ICC)

โค้งคุณลักษณะข้อสอบ หรือฟังก์ชันคุณลักษณะข้อสอบ (Item Characteristic Function) เป็นฟังก์ชันเชิงคณิตศาสตร์ที่แสดงความสัมพันธ์ระหว่างความน่าจะเป็นในการตอบข้อสอบถูกกับความสามารถที่ต้องการวัด ความสัมพันธ์ดังกล่าวจะอยู่ในรูปการถดถอยของคะแนนรายข้อ (Item Score) บนความสามารถ เพราะฉะนั้นถ้ามีการวัดเพียงความสามารถเดียวการถดถอยของคะแนนนั้น เรียกว่า Item Characteristic Curve หรือ ICC แต่ถ้ามีการวัดความสามารถในหลาย ๆ ด้าน การถดถอยของคะแนนนี้จะเรียกว่า Item Characteristic Function การถดถอยของคะแนนรายข้อบนความสามารถจะไม่เป็นเส้นตรง (Nonlinear) ความน่าจะเป็นของผู้สอบแต่ละคนในการที่จะตอบข้อสอบถูกเป็นรายข้อขึ้นอยู่กับโค้งคุณลักษณะข้อสอบซึ่งความน่าจะเป็นนี้ เป็นการแจกแจงความสามารถของผู้สอบที่เป็นอิสระกัน เพราะฉะนั้นความน่าจะเป็นของการตอบถูกเป็นรายข้อของผู้สอบจึงไม่ขึ้นอยู่กับจำนวนของผู้สอบที่มีระดับความสามารถเดียวกันคุณสมบัติของการไม่แปรผัน (Invariance) ของโค้งคุณลักษณะข้อสอบในกลุ่มผู้สอบนี้เป็นลักษณะเด่นของแบบจำลองการตอบสนองต่อข้อสอบ โค้งคุณลักษณะข้อสอบมีหลายรูปแบบขึ้นอยู่กับแบบจำลองของการตอบสนองต่อข้อสอบว่า แบบจำลองนี้มีพารามิเตอร์กี่ตัว เช่น อาจมีตัวเดียว สองตัว หรือสามตัว (Hambleton and Swaminathan. 1985 : 25 - 27)

3 <u>แบบจำลองของการตอบสนองต่อข้อสอบ</u> (Item Response Model)

แบบจำลองของการตอบสนองต่อข้อสอบมีหลายลักษณะดังนี้

3.1 Two-Parameter Normal Ogive Model เป็นแบบจำลองที่ลอร์ด (Lord) เสนอในปี ค.ศ. 1952 (Hambleton and Swaminathan. 1985 : 35 - 36) ซึ่งมีรูปแบบของสมการเชิงคณิตศาสตร์ดังนี้

$$P_i(\theta) = \int_{-\infty}^{a_i(\theta - b_i)} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz$$

เมื่อ $P_i(\theta)$ คือ โอกาส (Probability) ในการตอบข้อสอบข้อที่ i ถูกของคนที่มีความสามารถในระดับ θ

Z คือ ค่าเบี่ยงเบนปกติจากการแจกแจง (Normal Deviate From a Distribution) ซึ่งมีค่าเฉลี่ยเท่ากับ b_i ส่วนเบี่ยงเบนมาตรฐานเท่ากับ $1/a_i$

b_i คือ ระดับความยากของข้อสอบข้อที่ i ซึ่งแทนจุดบนสเกลของความสามารถ (Ability) ของผู้สอบที่มีโอกาสในการตอบข้อสอบข้อที่ i ถูก 50เปอร์เซ็นต์

a_i คือ ค่าอำนาจจำแนกของข้อสอบข้อที่ i ซึ่งเป็นสัดส่วนกับค่าความชันของ $P_i(\theta)$ ที่จุด $\theta = b_i$

เมื่อคะแนนความสามารถของกลุ่มได้เปลี่ยน (Transform) ไปเพื่อทำให้ค่าเฉลี่ยเท่ากับ 0 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1 แล้ว ค่าของ b จะมีค่าอยู่ระหว่าง -2.0 ถึง +2.0 ค่า b ที่ใกล้ -2.0 แสดงว่าข้อสอบนั้นง่าย แต่ถ้า b ใกล้ +2.0 แสดงว่าข้อสอบนั้นยากสำหรับกลุ่มผู้สอบ ส่วนค่าอำนาจจำแนก (a_i) โดยทฤษฎีแล้วมีค่าอยู่ระหว่าง $-\infty$ ถึง $+\infty$ แต่โดยปกติแล้วช่วงของค่าอำนาจจำแนกจะมีค่าอยู่ในช่วง (0,2)

3.2 Two-Parameter Logistic Model เป็นแบบจำลองที่เบิร์นบอม (Birnbaum) เสนอในรูปของ Two-Parameter Logistic Distribution Function ขึ้นก่อนโดยมีรูปแบบของสมการเชิงคณิตศาสตร์ดังนี้ (Hambleton and Swaminathan. 1985 : 36 - 52)

$$P_i(\theta) = \frac{e^{Da_i(\theta - b_i)}}{1 + e^{Da_i(\theta - b_i)}} \qquad (i = 1, 2, \ldots, n)$$

เมื่อ $P_i(\theta)$ คือ ความน่าจะเป็นที่ผู้สอบที่มีความสามารถ θ จะตอบคำถามข้อ i ได้ถูกต้อง

a_i และ b_i คือ พารามิเตอร์ของข้อคำถาม i ($i = 1, 2, \ldots n$)

n คือ จำนวนข้อคำถามในแบบทดสอบ

b_i คือ ดัชนีของค่าความยากซึ่งแทนจุดบนสเกลความสามารถ

a_i คือ อำนาจจำแนกของข้อคำถามซึ่งเป็นสัดส่วนกับความชันของ $P_i(\theta)$ ที่จุด $\theta = b_i$

D คือ ค่าคงที่มีค่า = 1.7

เบิร์นบอมได้นำแบบจำลอง Two-Parameter Logistic ที่เขาพัฒนาขึ้นไปใช้แทนแบบจำลอง Two-Parameter Normal Ogive ปรากฏว่าเส้นโค้งแบบ Logistic ใช้ได้ง่ายกว่าโค้งแบบ Normal Ogive เนื่องจากโค้งแบบ Logistic ไม่ต้องคำนวณโดยการอินทิเกรทอย่างโค้งแบบ Normal Ogive

สำหรับค่า $P_i(\theta)$, b_i, a_i มีความหมายเช่นเดียวกับแบบจำลองของ Normal Ogive ส่วนค่า D เป็นค่าคงที่มีค่าเท่ากับ 1.7 ค่าของ $P_i(\theta)$ ที่คำนวณจากแบบจำลอง Two-Parameter Normal Ogive เมื่อเปรียบเทียบกับการคำนวณจากแบบจำลอง Two-Parameter Logistic ปรากฏว่ามีค่าแตกต่างกันเพียง .01

3.3 Three-Parameter Logistic Model เป็นแบบจำลองที่พัฒนามาจากแบบจำลอง Two-Parameter Logistic เพียงแต่เพิ่มค่าพารามิเตอร์อีก 1 ตัวคือ c_i ซึ่งมีรูปแบบของสมการเชิงคณิตศาสตร์ดังนี้

$$P_i(\theta) = c_i + (1 - c_i)\frac{e^{Da_i(\theta - b_i)}}{1 + e^{Da_i(\theta - b_i)}} \qquad (i = 1, 2, \ldots, n)$$

เมื่อ $P_i(\theta)$ คือ โอกาสในการตอบข้อสอบข้อที่ i ถูกของคนที่มีความสามารถในระดับ θ

b_i คือ ค่าพารามิเตอร์ของความยาก

a_i คือ ค่าพารามิเตอร์ของอำนาจจำแนก

D คือ มีค่าเท่ากับ 1.7

c_i คือ ค่าพารามิเตอร์ของการเดา (Guessing Parameter)

3.4 One-Parameter Logistic Model (Rasch Model)

ผู้ที่พัฒนาแบบจำลอง One-Parameter Logistic ขึ้นมาคือ Georg Rasch นักคณิตศาสตร์ชาวเดนมาร์ก แบบจำลองนี้มีข้อตกลงเบื้องต้นว่า ข้อสอบทุกข้อมีค่าอำนาจจำแนกเท่ากันและค่าการเดาเป็นศูนย์ ดังนั้นสมการของแบบจำลองนี้จึงมีค่าพารามิเตอร์เพียงตัวเดียวคือ b_i ดังนั้น สมการของโค้งคุณลักษณะข้อสอบของแบบจำลอง One-Parameter Logistic สามารถเขียนได้ดังนี้

$$P_i(\theta) = \frac{e^{D\bar{a}(\theta - b_i)}}{1 + e^{D\bar{a}(\theta - b_i)}} \qquad (i = 1, 2, \ldots, n)$$

เมื่อ $\bar{a}$ เป็นค่าที่ไม่มีการนิยาม ซึ่งถือว่าเป็นค่าอำนาจจำแนกของทุกข้อ จากสมการข้างต้นบางครั้งจะเขียนในรูป

$$P_i(\theta) = \frac{e^{\theta' - b_i'}}{1 + e^{\theta' - b_i'}}$$

โดยถือว่าค่า $\theta' = D\bar{a}\,\theta$ และ $b' = D\bar{a}b_i$

3.5 Four - Parameter Logistic Model

แบบจำลองนี้คล้ายกับ Three - Parameter Logistic Model เพียงแต่เพิ่มค่าพารามิเตอร์อีก 1 ตัว คือความสะเพร่าของผู้สอบ (γ_i) เนื่องจากในการสอบบางครั้งผู้สอบที่มีความสามารถสูงจะไม่ตอบข้อคำถามในแบบทดสอบได้อย่างถูกต้องเนื่องจากมีความสะเพร่าเกิดขึ้น แต่ถ้าค่า $\gamma_i = 1$ แบบจำลองนี้จะคล้ายกับ Three - Parameter Logistic Model สมการของแบบจำลอง Four - Parameter Logistic สามารถเขียนได้ดังนี้

$$P_i(\theta) = c_i + (\gamma_i - c_i)\frac{e^{Da_i(\theta - b_i)}}{1 + e^{Da_i(\theta - b_i)}}$$

3.6 Nominal Response Model

แบบจำลอง One-Parameter Logistic, Two-Parameter Logistic และ Three-Parameter Logistic ดังกล่าวมาข้างต้นเป็นแบบจำลองที่ใช้กับข้อคำถามในแบบทดสอบที่มีการให้คะแนน 2 ลักษณะ (Dichotomous) คือ ตอบถูกได้ 1 คะแนน และตอบผิดได้ 0 คะแนน ส่วนแบบจำลอง Nominal Response นี้ เป็นแบบจำลองที่ใช้กับข้อคำถามที่มีการให้คะแนนมากกว่า 2 ลักษณะ (Multichotomous) จุดประสงค์ของแบบจำลองนี้คือ เพื่อเพิ่มความถูกต้องในการประมาณความสามารถของผู้สอบโดยใช้ประโยชน์จากข้อสนเทศที่ได้จากแต่ละตัวเลือกของข้อคำถาม ซึ่งแต่ละตัวเลือกจะมีโค้งคุณลักษณะข้อสอบประจำตัวเลือก (Item Option Characteristic Curve) แม้แต่ตัวเลือกที่เป็นตัวผิดก็ตามสำหรับตัวเลือกที่เป็นตัวถูกโค้งจะมีลักษณะเป็นการเพิ่มขึ้นโดยตลอด ส่วนตัวเลือกที่เป็นตัวผิดลักษณะโค้งจะขึ้นอยู่กับผู้สอบที่มีความสามารถในระดับแตกต่างกันที่จะเลือกตัวเลือกนั้น ดังนั้นโอกาสที่ผู้สอบที่มีความสามารถระดับ θ จะเลือกตัวเลือก k จากตัวเลือกทั้งหมด m ตัวเลือกในข้อคำถามที่ i สามารถเขียนในรูปแบบจำลองเชิงคณิตศาสตร์ได้ดังนี้

$$P_{ik}(\theta) = \frac{e^{b^*_{ik}+a^*_{ik}\theta}}{\sum_{h=1}^{m} e^{b^*_{ih}+a^*_{ih}\theta}} \quad (i = 1, 2, \ldots, n;\ k = 1, 2, \ldots, m)$$

ที่ระดับความสามารถ θ ใด ๆ ผลรวมของโอกาสการเลือกแต่ละตัวเลือก m จะเท่ากับ 1 ส่วน b^*_{ik} และ a^*_{ik} คือพารามิเตอร์ของตัวเลือก k ในกรณีที่ m = 2 แบบจำลองนี้จะเหมือนกับ Two-Parameter Logistic

แฮมเบิลตัน และสวามินาธาน (Hambleton and Swaminathan. 1985 : 53) ได้กล่าวถึงจุดประสงค์ของทฤษฎีการตอบสนองต่อข้อสอบว่า เพื่อใช้ในการพยากรณ์คาดคะเนหรืออ้างอิงถึงความสามารถของผู้สอบที่ได้จากการใช้แบบทดสอบวัด ซึ่งวิธีการประมาณค่าคะแนนความสามารถนั้นจะประมาณจากโค้งคุณลักษณะข้อสอบ (Item Characteristic Curve : ICC) ในรูปแบบที่แตกต่างกัน เช่น อาจมีโค้งโลจิสติคเป็นแบบ 1, 2, 3 พารามิเตอร์ (Hambleton and Cook. 1977 : 84) สำหรับการประมาณค่าความสามารถและพารามิเตอร์ของข้อสอบในปัจจุบันได้มีโปรแกรมคอมพิวเตอร์ เช่น LOGIST พัฒนาโดยวูด และลอร์ด (Wood and Lord) ใช้กับแบบจำลอง Three-Parameter Logistic โปรแกรม BICAL พัฒนาโดย ไรท์ และมีด (Wright and Mead) ใช้กับแบบจำลอง ราสช์ (Hambleton and Cook. 1977 : 89 - 90)

4. <u>ความสัมพันธ์ของคะแนนความสามารถ (Ability Scores) กับคะแนนโดเมน (Domain Scores)</u>

ถ้านำแบบทดสอบ 2 ฉบับที่ไม่ใช่แบบทดสอบคู่ขนานกันอย่างแท้จริงแต่เป็นแบบทดสอบที่วัดความสามารถเดียวกันไปสอบกับผู้สอบกลุ่มเดียวกันจะพบว่า การแจกแจงของคะแนนที่ได้จากแบบทดสอบทั้ง 2 ฉบับนั้น จะแตกต่างกัน ทั้งนี้เนื่องจากความยากของแบบทดสอบทั้ง 2 ฉบับ จากการแจกแจงของคะแนนสอบก็จะไม่ให้ข้อมูลเกี่ยวกับความสามารถของผู้สอบ

เนื่องจากคะแนนจากแบบทดสอบแต่ละฉบับไม่เท่ากัน ถึงแม้ว่าคะแนนสอบของผู้สอบจะเปลี่ยนแปลงไปตามแบบทดสอบที่ไม่ใช่คู่ขนานกันก็ตาม แต่คะแนนความสามารถของผู้สอบจะมีลักษณะเดียวกันในแบบทดสอบแต่ละฉบับ

แนวคิดของคะแนนจริงหรือคะแนนโดเมนเป็นสิ่งสำคัญในทฤษฎีการสอบแบบคลาสสิค การนิยามคะแนนจริงจะนิยามในรูปของค่าคาดหวังของคะแนนที่ได้จากการสอบดังนี้

ถ้าให้ t คือ คะแนนจริง

X คือ คะแนนที่ได้จากการสอบ

E คือ ค่าคาดหวัง

ดังนั้น $t = E(X)$

คะแนนจริงหรือคะแนนโดเมนเป็นคะแนนที่มีความสัมพันธ์กับคะแนนความสามารถของผู้สอบ (θ)

ถ้าให้ r คือ คะแนนทั้งหมดที่ได้จากการสังเกตของผู้สอบ

n คือ จำนวนข้อสอบ

U_i คือ การตอบสนองต่อข้อสอบข้อที่ i ที่ให้คะแนนแบบตอบถูกได้ 1 ตอบผิดได้ 0

ดังนั้น $$r = \sum_{i=1}^{n} U_i \qquad \text{.....(1)}$$

ถ้าให้ $\hat{\theta}$ คือ ค่าประมาณของคะแนนโดเมนของผู้สอบ (θ) ดังนั้นจะได้ว่า

$$\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n} U_i \qquad \text{.....(2)}$$

จากนิยามของคะแนนจริงที่ว่า E ($\hat{\pi}$) = π ดังนั้น จะได้ว่า

$$E(\hat{\pi}) = \pi = \frac{1}{n}\sum_{i=1}^{n} E(U_i) \qquad \text{.....(3)}$$

สำหรับผู้สอบที่มีความสามารถ θ จากสมการ (3) สามารถเขียนได้ว่า

$$\hat{\pi}|\theta = \frac{1}{n}\sum_{i=1}^{n} (U_i|\theta) \qquad \text{.....(4)}$$

และ $$E(\hat{\pi}|\theta) = \pi|\theta \qquad \text{.....(5)}$$

$$= \frac{1}{n}\sum_{i=1}^{n} E(U_i|\theta) \qquad \text{.....(6)}$$

U_i เป็นตัวแปรเชิงสุ่มซึ่งมีค่า 1 หรือ 0 ดังนั้น

$$E(U_i/\theta) = (U_i=1)\ P_i\ (U_i=1/\theta) + (U_i=0)\ P_i(U_i=0/\theta)$$

$$= 1.\ P_i\ (U_i = 1/\theta) + 0.P_i\ (U_i=0/\theta)$$

$$= P_i\ (U_i = 1/\theta)$$

$$= P_i(\theta) \qquad \text{.....(7)}$$

เพราะฉะนั้นฟังก์ชันการตอบสนองต่อข้อสอบคือ

$$\pi|\theta = \frac{1}{n}\sum_{i=1}^{n} P_i(\theta) \qquad \text{.....(8)}$$

จากสมการ (8) เพื่อทำให้ง่ายขึ้นจะใช้ π แทน $\hat{\pi}$ /θ ซึ่งฟังก์ชันนี้คือค่าเฉลี่ยของฟังก์ชันการตอบสนองต่อข้อสอบสำหรับข้อสอบจำนวน n ข้อ หรือที่รู้จักกันว่า คือฟังก์ชันคุณลักษณะแบบทดสอบ (Test Characteristic Function) หรือโค้งคุณลักษณะแบบทดสอบ (Test Characteristic Curve)

คะแนนโดเมน ($\hat{\pi}$) และคะแนนความสามารถ (θ) มีความสัมพันธ์กันในทางเดียวกัน (Monotonically) โดยสัมพันธ์กันในรูปของฟังก์ชันคุณลักษณะแบบทดสอบแนวคิดของ π และ θ เหมือนกันยกเว้นสเกลของการวัดแตกต่างกัน กล่าวคือ คะแนนโดเมนมีสเกลของการวัดอยู่ในช่วง [0, 1] แต่คะแนนความสามารถมีสเกลของการวัดอยู่ในช่วง $(-\infty, \infty)$ ความแตกต่างที่สำคัญมากที่สุดระหว่าง θ และ π ก็คือ θ เป็นสเกลความสามารถซึ่งไม่ขึ้นอยู่กับคำถาม ส่วน π เป็นสเกลของคะแนนโดเมนซึ่งขึ้นอยู่กับคำถามในแบบทดสอบ

5. <u>คะแนนความสัมพันธ์ระหว่างคะแนนโดเมนที่ได้จากการสังเกต และความสามารถ</u>

คะแนนโดเมนที่ได้จากการสังเกต ($\hat{\pi}$) นิยามในรูปของ

$$\hat{\pi}|\theta = \frac{1}{n}\sum_{i=1}^{n}(U_i|\theta)$$

และ $E(\hat{\pi}/\theta) = \pi$ ดังนั้น

$$\pi = \frac{1}{n}\sum_{i=1}^{n}P_i(\theta)$$

ที่ระดับความสามารถ θ ความแปรปรวน $\hat{\pi}/\theta$ มีค่าดังนี้

$$\mathrm{Var}(\hat{\pi}|\theta) = \mathrm{Var}\{\frac{1}{n}\sum_{i=1}^{n}(U_i|\theta)\}$$

โดยนิยามของ Local Independence จะได้

$$\mathrm{Var}\frac{1}{n}\sum_{i=1}^{n}(U_i|\theta)=\frac{1}{n^2}\sum_{i=1}^{n}\mathrm{Var}(U_i|\theta)$$

เนื่องจาก U_i/θ เป็นตัวแปรทวินาม (Binomial Variable)

$$\mathrm{Var}(U_i|\theta)=P_i(\theta)Q_i(\theta)$$

ดังนั้น

$$\mathrm{Var}\frac{1}{n}(U_i|\theta)=\frac{1}{n^2}P_i(\theta)Q_i(\theta)$$

$$\mathrm{Var}(\hat{\pi}|\theta)=\frac{1}{n^2}\sum_{i=1}^{n}P_i(\theta)Q_i(\theta)$$

จากสมการข้างต้นแสดงว่า ค่าเฉลี่ยแบบมีเงื่อนไขของ ($\hat{\pi}/\theta$) อยู่บนโค้งคุณลักษณะแบบทดสอบ ความสัมพันธ์ระหว่าง $\hat{\pi}/\theta$ และ θ จะไม่มีความสัมพันธ์แบบ 1 ต่อ 1 (One-to-one) ถ้าแบบทดสอบมีความยาวมากขึ้นโดยการเพิ่มจำนวนข้อที่ไม่จำกัดซึ่งมีลักษณะเป็นข้อสอบที่คู่ขนานกัน เช่น $n \longrightarrow \infty$, $\hat{\pi}/\theta \longrightarrow \pi$ และ $\mathrm{Var}(\hat{\pi}/\theta \longrightarrow \theta$ ดังนั้น $\hat{\pi}$ จะเป็นค่าประมาณคงที่ของ π ซึ่งเป็นคะแนนโดเมน

จากที่กล่าวมาข้างต้นถึงค่าเฉลี่ยและความแปรปรวนของ $\hat{\pi}$ ที่ระดับ θ ที่กำหนดนั้นเป็นการกล่าวที่ตั้งอยู่บนข้อตกลงเบื้องต้นว่าค่า θ นั้นเรารู้ค่า แต่ในทางปฏิบัติแล้วเราไม่รู้ค่า θ เราจะต้องใช้วิธีการประมาณค่า θ ซึ่งจะแทนด้วย $\hat{\theta}$ ดังนั้นการประมาณค่าของคะแนนโดเมนที่ระดับความสามารถ $\hat{\theta}$ จะได้ว่า

$$\hat{\pi}\,|\,\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n} P_i(\hat{\theta})$$

6. <u>การนำทฤษฎีการตอบสนองต่อข้อสอบมาประยุกต์ใช้ในการกำหนดคะแนนจุดตัด</u>

นักวัดผลการศึกษาที่ได้เสนอวิธีการนำทฤษฎีการตอบสนองต่อข้อสอบมาใช้ในการกำหนดคะแนนจุดตัดนั้นเริ่มจาก แวน เดอร์ ลินเดน (Van der Linden. 1982) ต่อมา เคน (Kane. 1987) ได้นำแนวความคิดของ แวน เดอร์ ลินเดน ไปใช้และเสนอวิธีการกำหนดคะแนนจุดตัด 3 วิธี

สำหรับ แวน เดอร์ ลินเดน (Van der Linder. 1982 : 296 - 307) ได้เสนอวิธีการกำหนดคะแนนจุดตัดและวิธีการหาความคงที่ของการตัดสินดังนี้

1. <u>การกำหนดคะแนนจุดตัด</u>

1.1 ใช้เทคนิคของแองกอฟฟ์หรือนีเดลสกีในการประมาณคะแนนจุดตัด (τ_c) ซึ่งอยู่บนสเกลของคะแนนจริงดังนี้

1.1.1 วิธีการของแองกอฟฟ์

1.1.1.1 หาความน่าจะเป็นที่ผู้สอบที่มีความสามารถต่ำสุดจะตอบข้อคำถาม i ถูกด้วยระดับความสามารถ θ เขียนได้ว่า $P_i\ (+/\theta_c)$

1.1.1.2 คะแนนจุดตัดของแบบทดสอบหาได้จากสูตร

$$\tau_c = \sum_{i=1}^{n} P_i(+/\theta_c)$$

1.1.2 วิธีการของนีเดลสกี

1.1.2.1 หาความน่าจะเป็นที่ผู้สอบที่มีความสามารถต่ำสุดจะตอบข้อคำถาม i ถูก ซึ่งจะเท่ากับส่วนกลับของตัวเลือกที่เหลือ นั่นคือ

$$P_i(+|\theta_c) = [q_i - k_i^{(c)}]^{-1}$$

เมื่อ q_i คือ จำนวนของตัวเลือกของข้อคำถาม i

$k_i(c)$ คือ จำนวนของตัวเลือกที่ผู้เชี่ยวชาญพิจารณาว่าผู้สอบที่มีความสามารถต่ำสุด (θ_c) รู้ว่าเป็นตัวเลือกที่ผิด

1.1.2.2 คะแนนจุดตัดของแบบทดสอบจะเท่ากับผลบวกของจำนวนตัวเลือกที่เหลือ นั่นคือ

$$\tau_c = \sum_{i=1}^{n} [q_i - k_i^{(c)}]^{-1}$$

1.2 แปลงคะแนนจุดตัด (τ_c) ซึ่งอยู่บนสเกลของคะแนนจริงให้เป็นสเกลของความสามารถ โดยใช้โค้งคุณลักษณะแบบทดสอบ

2. <u>การหาความคงที่ของการตัดสิน</u>

การหาความคงที่ของการตัดสินมีขั้นตอนดังนี้

1. หาค่าความน่าจะเป็นของการตอบข้อคำถามแต่ละข้อได้ถูกซึ่งแทนด้วย $p_i(s)$ ตามเทคนิคของนีเดลสกีหรือ แองกอฟฟ์

2. หาความน่าจะเป็นในการตอบข้อคำถามแต่ละข้อได้ถูก จากโค้งคุณลักษณะข้อสอบซึ่งแทนด้วย P_i

3. หาผลต่างของ $P_i(s)$ และ P_i สำหรับแต่ละข้อโดยใช้สูตร

$$e = P_i(s) - P_i$$

เมื่อ $P_i(s)$ คือ ค่าที่ได้จากการพิจารณาของผู้เชี่ยวชาญ

P_i คือ ค่าที่ได้จากการคำนวณโดยใช้โค้งคุณลักษณะข้อสอบสำหรับในกรณีที่ใช้แบบจำลอง 3 พารามิเตอร์

$$P_i = c_i + (1-c_i)\{1+\exp[-a_i(\theta-b_i)]\}^{-1}$$

4. หาผลต่างโดยเฉลี่ยระหว่าง $P_i(s)$ และ P_i ของแบบทดสอบทั้งฉบับจากสูตร

$$E = \sum_{i=1}^{n} \left| P_i(s) - P_i \right| / n$$

เมื่อ n คือ จำนวนข้อคำถามในแบบทดสอบ

5. หาความคงที่ของการตัดสิน ดังนี้

$$C_1 = \frac{M - E}{M}$$

เมื่อ $$M = \sum_{i=1}^{n} e_i^{(u)}/n$$

$$e_i^{(u)} = \max \{p_i, 1 - p_i\}$$

สำหรับเทคนิคของนีเดลสกี สามารถหาความถูกต้องของการตัดสินได้ดังนี้

$$\lambda = C_2 - C_1$$

$$C_2 = \frac{M - E}{M - m}$$

$$m = \sum_{i=1}^{n} e^{(p)}/n$$

$$e^{(p)} = |[q_i - k_i^*]^{-1} - p_i|$$

เคน (Kane. 1987 : 333 - 345) ได้เสนอวิธีการกำหนดคะแนนจุดตัด (θ^*) โดยการกำหนดเกณฑ์มาตรฐานก่อนจากการให้ผู้เชี่ยวชาญแต่ละคนพิจารณาความสามารถต่ำสุดของผู้สอบที่จะตอบแต่ละข้อคำถามได้ถูกต้องด้วยเทคนิคใดเทคนิคหนึ่ง เช่น เทคนิคของแองกอฟฟ์ อีเบลและนีเดลสกี ซึ่งมีทั้งหมด 3 วิธี คือ

วิธีที่ 1 ใช้โค้งคุณลักษณะข้อสอบในการประมาณค่า θ^* ของข้อคำถามแต่ละข้อแล้วจึงนำค่า θ^* จากทุกข้อคำถามมาหาค่าเฉลี่ย ซึ่งมีขั้นตอนดังนี้

1. คำนวณหาค่าเฉลี่ยของผู้สอบที่มีความสามารถต่ำสุดที่จะตอบข้อคำถาม i ได้ถูกจากการตัดสินโดยผู้เชี่ยวชาญทั้งหมด ดังนี้

$$M_{iR} = \frac{\sum M_{ir}}{k}$$

เมื่อ M_{ir} คือ ค่าความสามารถต่ำสุดของผู้สอบจากการตัดสินโดยผู้เชี่ยวชาญคนที่ r ในข้อคำถาม i

k คือ จำนวนผู้เชี่ยวชาญทั้งหมด

M_{iR} คือ ค่าความสามารถต่ำสุดโดยเฉลี่ยของข้อคำถาม i จากการตัดสินโดยผู้เชี่ยวชาญ k คน

2. กำหนดให้ M_{iR} เป็นค่าประมาณของความน่าจะเป็นที่ผู้สอบมีความสามารถต่ำสุดจะตอบข้อคำถาม i ได้ถูกต้องโดยการใช้โค้งคุณลักษณะข้อสอบ ประมาณค่า $\hat{\theta}^*_{iR}$ บนสเกลความสามารถ (θ) นั่นคือ

$$M_{iR} \doteq P_i (\hat{\theta}^*_{iR})$$

3. หาค่าความแปรปรวนของ $\hat{\theta}^*_{iR}$ ของข้อคำถามที่ i ดังนี้

ให้ $P\acute{i} (\hat{\theta}^*_{iR})$ เป็นความชันของโค้งคุณลักษณะข้อสอบ สำหรับข้อคำถาม i ที่ $\hat{\theta}^*_{iR}$

ΔM_{iR} แทน พิสัยของค่า M_{iR}

ความชันของโค้งคุณลักษณะข้อสอบที่ i ที่ความสามารถ $\hat{\theta}^*_{iR}$ คือ

$$P'_i(\hat{\theta}^*_{iR}) \simeq \frac{\Delta M_{iR}}{\Delta \hat{\theta}^*_{iR}} \qquad \text{.....(1)}$$

ถ้าแทน ΔM_{iR} ด้วย $\sigma_i(M_{iR})$ และ $\Delta \hat{\theta}^*_{iR}$ ด้วย $\sigma_i(\hat{\theta}^*_{iR})$ จากสมการ (1) จะได้

$$P'_i(\hat{\theta}^*_{iR}) \simeq \frac{\sigma_i(M_{iR})}{\sigma_i(\hat{\theta}^*_{iR})}$$

$$\frac{P'_i(\hat{\theta}^*_{iR})}{\sigma_i(M_{iR})} \simeq \frac{1}{\sigma_i(\hat{\theta}^*_{iR})} \qquad \text{.....(2)}$$

จากสมการ (2) ยกกำลัง 2

$$\frac{[P'_i(\hat{\theta}^*_{iR})]^2}{\sigma^2_i(M_{iR})} \simeq \frac{1}{\sigma^2_i(\hat{\theta}^*_{iR})} \qquad \text{.....(3)}$$

4. ประมาณค่า θ^* จากการหาค่าเฉลี่ยของ $\hat{\theta}^*_{iR}$ ดังนี้

$$\hat{\theta}^* = \frac{1}{n}\sum_i \hat{\theta}^*_{iR} \qquad \text{.....(4)}$$

ถ้า $\hat{\theta}^*_{iR}$ แต่ละค่าเป็นอิสระซึ่งกันและกัน ดังนั้น ความแปรปรวนของ $\hat{\theta}^*$ คือ

$$\sigma^2(\hat{\theta}^*) = \frac{1}{n^2}\sum_i \sigma_i^2(\hat{\theta}^*_{iR})$$

แต่ถ้า $\hat{\theta}^*_{iR}$ แต่ละค่าไม่เป็นอิสระซึ่งกันจะต้องใช้ค่าถ่วงน้ำหนัก ดังนั้น สมการ (4) จึงเขียนได้ว่า

$$\hat{\theta}^*_w = \sum_i w_i \hat{\theta}^*_{iR} \qquad \text{.....(5)}$$

โดยที่ $$\sum_i w_i = 1 \qquad \text{.....(6)}$$

ค่า w_i จะทำให้ความแปรปรวนใน $\hat{\theta}^*_w$ มีค่าน้อยที่สุด ความแปรปรวนของ $\hat{\theta}^*_w$ สามารถคำนวณได้จากสูตร

$$\sigma^2(\hat{\theta}^*_w) = \sum_i w_i^2 \sigma_i^2(\hat{\theta}^*_{iR}) \qquad \text{.....(7)}$$

ค่า w_i ที่ทำให้ความแปรปรวนในสมการ (7) มีค่าน้อยที่สุดและสอดคล้องกับสมการ (4) ซึ่งสามารถหาได้โดยใช้ Lagrange Multiplier ดังนี้

$$w_i = \frac{\dfrac{1}{\sigma_i^2(\hat{\theta}^*_{iR})}}{\displaystyle\sum_i \frac{1}{\sigma_i^2(\hat{\theta}^*_{iR})}} \qquad \text{.....(8)}$$

การประมาณค่าโดยการถ่วงน้ำหนักนี้ จะทำให้ลดค่าความแปรปรวนเนื่องจากความคลาดเคลื่อนในการประมาณค่า $\hat{\theta}^*$ โดยการถ่วงน้ำหนักมากที่สุดแก่ค่า $\hat{\theta}^*_{iR}$ ในกรณีที่ผู้ตัดสินมีความเห็นตรงกันมากที่สุด และจะมีการถ่วงน้ำหนักน้อยแก่ค่า $\hat{\theta}^*_{iR}$ ในกรณีที่ผู้ตัดสินมีความเห็นไม่ตรงกัน

จากสมการ (8) ค่า $\sigma_i^2(\hat{\theta}^*_{iR})$ สามารถหาได้จากสมการ (3)

$$\sigma_i^2(\hat{\theta}^*_{iR}) = \frac{\sigma_i^2(M_{iR})}{[P_i'(\hat{\theta}^*_{iR})]^2} \qquad \text{.....(9)}$$

ส่วนการประมาณค่า $\sigma_i^2(M_{iR})$ สามารถหา ได้ดังนี้

$$\hat{\sigma}_i^2(M_{ir}) = \frac{\sum_r (M_{ir} - M_{iR})^2}{k-1}$$

โดยที่ k คือ จำนวนผู้เชี่ยวชาญทั้งหมด

M_{ir} คือ ความสามารถต่ำสุดของผู้สอบที่จะตอบคำถามข้อ i ถูกจากการตัดสินของผู้เชี่ยวชาญคนที่ r

M_{iR} คือ ค่าความสามารถต่ำสุดของผู้สอบที่จะตอบคำถามข้อ i ถูกโดยเฉลี่ยจากการตัดสินของผู้เชี่ยวชาญ k คน

ดังนั้น $\hat{\sigma}_i^2(M_{iR}) = \hat{\sigma}_i^2(M_{ir})/k$ .....(10)

เพราะฉะนั้น หากนำสมการ (9) และ (10) ไปแทนในสมการ (8) ก็จะหาค่า w_i ได้จากค่า w_i ที่ได้นำไปแทนในสมการ (5) ก็จะหาค่า $\hat{\theta}^*_w$ ซึ่งเป็นคะแนนจุดตัดที่ได้จากการถ่วงน้ำหนัก

วิธีที่ 2 เป็นวิธีที่ใช้กันมากที่สุดในการกำหนดคะแนนจุดตัด ซึ่งมีวิธีการดังนี้ คือ

1. คำนวณหาค่าเฉลี่ยของผู้สอบที่มีความสามารถต่ำสุดที่จะตอบข้อคำถาม i ได้ถูกจากการตัดสินของผู้เชี่ยวชาญทั้งหมด ดังนี้

$$M_{iR} = \frac{\sum M_{ir}}{k}$$

เมื่อ M_{ir} คือ ค่าความสามารถต่ำสุดของผู้สอบจากการตัดสินโดยผู้เชี่ยวชาญคนที่ r ในข้อคำถาม i

k คือ จำนวนผู้เชี่ยวชาญทั้งหมด

M_{iR} คือ ค่าความสามารถต่ำสุดโดยเฉลี่ยของข้อคำถาม i จากการตัดสินโดยผู้เชี่ยวชาญ k คน

2. หาผลรวมของ M_{iR} จากทุกข้อคำถามในแบบทดสอบก็จะได้เป็นคะแนนจุดตัดบนสเกลของคะแนนจริง

3. ใช้โค้งคุณลักษณะแบบทดสอบ (Test Characteristic Curve : TCC) ในการประมาณค่าความสามารถ (θ) ซึ่งจะเท่ากับผลรวมของ M_{iR} ในขั้นที่ 2 นั่นคือ

$$\sum_i M_{iR} = \sum_i P_i(\hat{\theta}^*) \qquad \text{.....(11)}$$

จากสมการ (11) สามารถเขียนใหม่ได้ว่า

$$\sum_i \left[M_{iR} - P_i(\hat{\theta}^*)\right] = 0 \qquad \text{.....(12)}$$

จากสมการ (12) แทนค่า $M_{iR} = P_i(\hat{\theta}^*_{iR})$ จะได้

$$\sum_i [P_i(\hat{\theta}^*_{iR}) - P_i(\hat{\theta}^*)] = 0$$

ถ้า $\hat{\theta}^*_{iR}$ มีค่าใกล้กับ θ^* ความสัมพันธ์นี้สามารถประมาณได้ด้วย Linear Function นั่นคือ

$$\sum P'_i(\hat{\theta}^*_{iR})\,(\hat{\theta}^*_{iR} - \hat{\theta}^*) \simeq 0 \qquad \text{.....(13)}$$

$$\hat{\theta}^* \simeq \frac{1}{\sum_i P'_i(\hat{\theta}^*_{iR})} \sum_i P'_i(\hat{\theta}^*_{iR})\,\hat{\theta}^*_{iR} \qquad \text{.....(14)}$$

จากสมการ (2) จะได้

$$P'_i(\hat{\theta}^*_{iR}) \simeq \frac{\sigma_i(M_{iR})}{\sigma_i(\hat{\theta}^*_{iR})} \qquad \text{.....(15)}$$

แทนค่าสมการ (15) ในสมการ (14) จะได้

$$\hat{\theta}^* \simeq \frac{1}{\sum_i \sigma_i(M_{iR})/\sigma_i(\hat{\theta}^*_{iR})} \sum_i \frac{\sigma_i(M_{iR})}{\sigma_i(\hat{\theta}^*_{iR})}\,\hat{\theta}^*_{iR} \qquad \text{.....(16)}$$

ค่า $\hat{\theta}^*$ ที่ได้จากสมการ (14) และ (16) ยังไม่สามารถแปลความหมายได้อย่างชัดเจน เนื่องจากน้ำหนักที่ได้ไม่เท่ากับน้ำหนักสูงสุดของการประมาณค่า θ^* จาก θ^*_{iR} ในสมการ (8) ดังนั้นจากสมการ (14) และ (16) จะต้องปรับโดยการถ่วงน้ำหนักด้วยค่า $1/\sigma_i(M_{iR})$ เพราะฉะนั้น จากสมการ (14) จะได้ว่า

$$\hat{\theta}^*_w \simeq \frac{1}{\sum_i [P'_i(\hat{\theta}^*_{iR})/\sigma_i(M_{iR})]} \sum_i \frac{P'_i((\hat{\theta}^*_{iR})}{\sigma_i(M_{iR})} \hat{\theta}^*_{iR} \quad \text{.....(17)}$$

หรือการใช้ค่าประมาณในสมการ (15) แทนค่าในสมการ (17) จะได้

$$\hat{\theta}_w{}^* \simeq \frac{1}{\sum_i [(1/\sigma_i(\hat{\theta}^*_{iR})]} \sum_i \frac{1}{\sigma_i(\hat{\theta}^*_{iR})} \hat{\theta}^*_{iR} \quad \text{.....(18)}$$

วิธีที่ 3 เป็นวิธีการประมาณค่าคะแนนจุดตัด ($\hat{\theta}^*$) ที่ดีที่สุดระหว่างการตัดสินจากความสามารถต่ำสุดของผู้สอบโดยผู้เชี่ยวชาญ และการหาค่าความน่าจะเป็น $P_i(\hat{\theta}^*)$ จากโค้งคุณลักษณะข้อสอบ ซึ่งมีขั้นตอนดังนี้

1. คำนวณหาค่าเฉลี่ยของผู้สอบที่มีความสามารถต่ำสุดที่จะตอบข้อคำถาม i ได้ถูกจากการตัดสินของผู้เชี่ยวชาญทั้งหมด ดังนี้

$$M_{iR} = \frac{\sum M_{ir}}{k}$$

เมื่อ M_{ir} คือ ค่าความสามารถต่ำสุดของผู้สอบจากการตัดสินโดยผู้เชี่ยวชาญคนที่ r ในข้อคำถาม i

k คือ จำนวนผู้เชี่ยวชาญทั้งหมด

M_{iR} คือ ค่าความสามารถต่ำสุดโดยเฉลี่ยของข้อคำถาม i จากการตัดสินโดยผู้เชี่ยวชาญ k คน

2. ประมาณค่า $\sigma_i^2(M_{iR})$ โดยใช้สมการ (10)

3. ประมาณค่า θ^* ที่ทำให้ผลรวมของส่วนเบี่ยงเบนกำลังสองระหว่าง M_{iR} และ $P_i(\hat{\theta}^*)$ มีค่าน้อยที่สุด ดังนี้

$$\sum_i \frac{(M_{iR} - P_i(\hat{\theta}^*))^2}{\sigma_i^2(M_{iR})} \simeq \sum_i \frac{[P_i(\hat{\theta}^*_{iR}) - P_i(\hat{\theta}^*)]^2}{\sigma_i^2(M_{iR})}$$

แต่ถ้า $\hat{\theta}^*_{iR}$ มีค่าใกล้กับค่า $\hat{\theta}^*$ และการใช้ค่าประมาณในสมการ (13) จะได้

$$\sum_i \frac{(M_{iR} - P_i(\hat{\theta}^*))^2}{\sigma_i^2(M_{iR})} \simeq \sum_i \frac{1}{\sigma_i^2(M_{iR})} [P_i'(\hat{\theta}^*_{iR})\,(\hat{\theta}^*_{iR} - \hat{\theta}^*)]^2$$

$$\simeq \sum_i \frac{P_i'^2(\hat{\theta}^*_{iR})}{\sigma_i^2(M_{iR})} (\hat{\theta}^*_{iR} - \hat{\theta}^*)^2$$

.....(19)

จากสมการ (19) ถ้า

$$\sum_i \frac{P_i'^2(\hat{\theta}^*_{iR})}{\sigma_i^2(M_{iR})} (\hat{\theta}^*_{iR} - \hat{\theta}^*) \simeq 0$$

จะได้

$$\hat{\theta}^{*} \simeq \frac{1}{\sum_{i} [P_i'^2(\hat{\theta}^{*}_{iR})/\sigma_i^2(M_{iR})]} \sum_{i} \frac{P_i'^2(\hat{\theta}^{*}_{iR})}{\sigma_i^2(M_{iR})} \hat{\theta}^{*}_{iR}$$

หรือการใช้ค่าประมาณในสมการ (2) จะได้

$$\hat{\theta}^{*} \simeq \frac{1}{\sum_{i} [1/\sigma_i^2(\hat{\theta}^{*}_{iR})]} \sum_{i} \frac{1}{\sigma_i^2(\hat{\theta}^{*}_{iR})} \hat{\theta}^{*}_{iR}$$

หากกล่าวโดยสรุปจะเห็นว่า วิธีการกำหนดคะแนนจุดตัดของแบบทดสอบทั้ง 3 วิธี เป็นวิธีการกำหนดเกณฑ์มาตรฐานด้วยวิธีการพิจารณาจากความสามารถต่ำสุดของผู้สอบ ซึ่งจะไม่เหมือนกับวิธีการดั้งเดิมที่กำหนดคะแนนจุดตัดบนสเกลของคะแนนจริง เคน กล่าวว่าวิธีการดังกล่าวนั้นเป็นวิธีการที่ได้โดยการหาผลรวมจากการตัดสินแต่ละข้อคำถามโดยเฉลี่ยซึ่งเป็นวิธีการประมาณคะแนนจุดตัดที่ไม่มีเหตุผลและการกำหนดคะแนนจุดตัดทั้ง 3 วิธีที่เคน เสนอนี้ยังไม่มีการศึกษาในเชิงประจักษ์

สำหรับในการวิจัยครั้งนี้ ผู้วิจัยได้เลือกศึกษาวิธีที่ 3 ซึ่งเป็นวิธีที่เคน กล่าวว่าเป็นวิธีที่ดีที่สุดที่จะนำไปใช้ในการพิจารณาคะแนนจุดตัด (Kane. 1987 : 341)

ในการกำหนดคะแนนจุดตัดโดยการนำทฤษฎีการตอบสนองต่อข้อสอบไปประยุกต์ใช้ นอกจากตามวิธีการที่ แวน เดอร์ ลินเดน และเคนเสนอนั้น หากแบบทดสอบได้มีการแบ่งเป็นแบบทดสอบย่อยก็ยังสามารถหาคะแนนจุดตัดของแบบทดสอบย่อยแต่ละฉบับได้ด้วยวิธีการที่เรียกว่า มินิแมกซ์ (Minimax) และวิธีการของราสช์ ดังจะกล่าวต่อไป

7. การใช้วิธีการมินิแมกซ์ (Minimax) และวิธีการของราสช์ (Rasch) ในการกำหนดคะแนนจุดตัดของแบบทดสอบย่อย

ฮวิน และแคสตีล (Huynh and Casteel. 1985 : 334 - 344) ได้เสนอวิธีการในการกำหนดคะแนนจุดตัดของแบบทดสอบย่อย 2 วิธี คือ วิธีการมินิแมกซ์ และวิธีการของราสช์ ทั้ง 2 วิธีนี้ จะต้องทราบคะแนนจุดตัดของแบบทดสอบทั้งฉบับก่อน วิธีการของมินิแมกซ์เป็นการพยายามทำให้ความสอดคล้องระหว่างการตัดสินได้/ตก ในแบบทดสอบย่อยและแบบทดสอบทั้งฉบับมีค่าสูงสุด ส่วนวิธีการของราสช์เป็นการพิจารณาถึงสเกลความสามารถและพยายามหาคะแนนจุดตัดของแบบทดสอบย่อยที่มีค่าใกล้เคียงกับคะแนนจุดตัดทั้งหมดบนสเกลของความสามารถรายละเอียดของทั้ง 2 วิธี สามารถสรุปได้ดังนี้

1. วิธีการมินิแมกซ์

ถ้าให้ Y แทนคะแนนที่ได้จากแบบทดสอบ และแบบทดสอบแบ่งออกเป็น k แบบทดสอบย่อยแต่ละแบบทดสอบย่อยจะประกอบไปด้วยคะแนน $X_1, X_2, \ldots, X_k$ ดังนั้น $Y = X_1 + X_2 + \ldots + X_k$ โดยที่ c คือคะแนนจุดตัดของแบบทดสอบทั้งฉบับ $r = (r_1, r_2, r_3, \ldots, r_k)$ เป็นคะแนนจุดตัดของแบบทดสอบย่อยซึ่งสอดคล้องกับคะแนนจุดตัดของแบบทดสอบทั้งฉบับ (c) เพราะฉะนั้นผลรวมของคะแนนจุดตัดจากแบบทดสอบย่อยจะเท่ากับคะแนนจุดตัดของแบบทดสอบทั้งฉบับ นั่นคือ $r_1 + r_2, \ldots, + r_k = c$ ซึ่งจะทำให้มั่นใจได้ว่าผู้สอบใด ๆ ที่ผ่านแต่ละวัตถุประสงค์ก็จะผ่านแบบทดสอบทั้งฉบับด้วย

ขั้นตอนของวิธีการมินิแมกซ์

ให้ $P_i(r_i)$ เป็นสัดส่วนของผู้สอบซึ่งถูกจัดประเภทแบบเดียวกับแบบทดสอบทั้งฉบับและแบบทดสอบย่อย i หรือกล่าวอีกนัยหนึ่งคือ P_i แทนผู้สอบซึ่งมีคะแนน $Y < c$ และ $X_i < r_i$ หรือ $Y \geq c$ และ $X_i \geq r_i$ ดังนั้น

$$P_i(r_i) = P_r(Y < c, X_i < r_i) + P_r(Y \geq c, X_i \geq r_i)$$

สำหรับแต่ละเซทของคะแนนจุดตัด $r = (r_1, r_2, \ldots, r_k)$ โดยที่ $P_{min}(r)$ คือค่าที่น้อยที่สุดของ P_i สำหรับ k แบบทดสอบย่อย นั่นคือ

$$P_{min}(r) = \min [P_1(r_1), P_2(r_2), \ldots, P_k(r_k)]$$

คะแนนจุดตัดที่สูงสุด คือ $r = (r_1, r_2, \ldots, r_k)$ ของแบบทดสอบย่อยจะสอดคล้องกับเวคเตอร์ $r^o = (r_1^o, r_2^o, \ldots, r_k^o)$ ซึ่งความน่าจะเป็นที่น้อยที่สุด ของ $P_{min}(r^o)$ จะมีค่ามากที่สุดระหว่างความน่าจะเป็น $P_{min}(r)$ ที่คำนวณจากค่า r ที่เป็นไปได้ทั้งหมด ดังนั้นวิธีการมินิแมกซ์ คือการพยายามทำให้ความน่าจะเป็นที่น้อยที่สุดของความสอดคล้องในการจัดประเภทระหว่างคะแนนทั้งหมดในแบบทดสอบ และคะแนนจากแบบทดสอบย่อยมีค่าสูงสุด (นั่นคือ เป็นการพยายามลดความน่าจะเป็นสูงสุด ของความไม่สอดคล้องในการแยกประเภทระหว่างคะแนนจากแบบทดสอบทั้งฉบับและแบบทดสอบย่อยแต่ละฉบับ

วิธีการมินิแมกซ์ สามารถนำไปใช้ได้ต่าง ๆ กัน คือ เมื่อคะแนนจากแบบทดสอบย่อยมีค่าแตกต่างกัน และจำนวนของแบบทดสอบย่อยมีไม่มาก เราสามารถหาค่า $r = (r_1, r_2, \ldots, r_k)$ ซึ่ง $r_1 + r_2 + \ldots + r_k = c$ แล้วคำนวณค่า $P_{min}(r)$ ที่ค่า r แต่ละค่า ต่อจากนั้นจึงหาค่า r^o ซึ่งเป็นค่าความน่าจะเป็นที่มีค่ามากที่สุด

2. <u>วิธีการของราสช์</u>

วิธีการนี้เป็นการกำหนดคะแนนจุดตัดโดยใช้แบบจำลองคุณลักษณะแฝง ซึ่งมีวิธีการดังนี้ คือ ถ้าแบบทดสอบประกอบด้วยจำนวนข้อ L ข้อ แต่ละข้อจะมีโค้งคุณลักษณะประจำข้อในรูปของ $P(\theta)$ ซึ่งหมายถึงความน่าจะเป็นที่ผู้สอบที่มีความสามารถ θ จะตอบข้อคำถามได้ถูกต้อง ดังนั้นโค้งคุณลักษณะข้อสอบจากแบบทดสอบทั้งฉบับ จะเท่ากับค่าคาดหวังของจำนวนข้อสอบที่ผู้สอบมีความสามารถ θ จะตอบได้ถูกต้อง นั่นคือ

$$E(\theta) = \sum_{j=1}^{L} P_j(\theta)$$

เมื่อ $P_j(\theta)$ คือ โค้งคุณลักษณะข้อสอบข้อที่ j

c คือ คะแนนจุดตัดของแบบทดสอบทั้งฉบับ

θ_c คือ ความสามารถที่ระดับคะแนนจุดตัด

ดังนั้นจะได้

$$E(\theta_c) = c$$

ถ้ากำหนดให้ k เป็นแบบทดสอบย่อยประกอบด้วยจำนวนข้อดังนี้ L_1, L_2 ,..., L_k เพราะฉะนั้น ฟังก์ชันคุณลักษณะแบบทดสอบย่อย i คือ

$$E_i(\theta) = \sum_j P_j(\theta)$$

ที่ระดับความสามารถเท่ากับคะแนนจุดตัด θ_c เพราะฉะนั้น ค่าที่คาดหวังของการตอบแบบทดสอบย่อยถูก คือ $E_i(\theta_c)$ ซึ่งสามารถสรุปได้ว่าผลบวกของ $E_1(\theta_c + E_2(\theta_c) +,...,+ E_k(\theta_c)$ เป็นคะแนนจุดตัด (c) ของแบบทดสอบทั้งฉบับ

เนื่องจากฟังก์ชันคุณลักษณะแบบทดสอบ $E_1(\theta)$ เป็นการเพิ่มทางเดียวโดยตลอด ดังนั้นการใช้คะแนนจุดตัด c แสดงให้เห็นว่า ผู้สอบจะผ่านแบบทดสอบทั้งฉบับ เมื่อความสามารถจริง (θ) มีค่าน้อยที่สุดเท่ากับ θ_c เพราะฉะนั้นผู้สอบจะสอบผ่านแต่ละแบบทดสอบย่อยจะต้องมีเงื่อนไขดังนี้ คือ $\theta \geq \theta_c$ ซึ่งแสดงให้เห็นว่า คะแนนจุดตัดสำหรับแบบทดสอบย่อย i จะมีค่าเท่ากับ $E_i(\theta_c)$ และคะแนนจุดตัดสำหรับแต่ละแบบทดสอบย่อย $E_i(\theta_c)$ จะมีค่าใกล้เคียงกับ r ภายใต้เงื่อนไขที่ว่า $r_1 + r_2 +,..., + r_k = c$ นั่นคือ ผู้สอบที่ผ่านแต่ละแบบทดสอบย่อยก็จะผ่านแบบทดสอบทั้งฉบับด้วย

งานวิจัยที่เกี่ยวข้องกับการกำหนดคะแนนจุดตัดที่นำทฤษฎีการตอบสนองต่อข้อสอบมาประยุกต์ใช้

งานวิจัยที่เกี่ยวข้องกับการกำหนดคะแนนจุดตัดที่นำทฤษฎีการตอบสนองต่อข้อสอบมาประยุกต์ใช้ยังมีค่อนข้างน้อยและเพิ่งจะได้รับความสนใจเพียงไม่กี่ปี จากการศึกษางานวิจัยในลักษณะนี้ ปรากฏว่าเริ่มได้รับความสนใจตั้งแต่ปี พ.ศ. 2528 เป็นต้นมา คืองานวิจัยของ ชวลิต ไพธิ์นคร (2528 : 91 - 100) ซึ่งได้ทำการเปรียบเทียบ การกำหนดคะแนนจุดตัดแบบทดสอบอิงเกณฑ์ตามแนวทฤษฎีการวัดแบบคลาสสิค 2 วิธีคือ วิธีของเบอร์ก และวิธีกำหนดเกณฑ์ผ่านระดับต่ำสุดกับทฤษฎีการตอบสนองต่อข้อสอบโดยการประยุกต์ราสช์โมเดล ผลการวิจัยปรากฏว่า คะแนนจุดตัดของแบบทดสอบอิงเกณฑ์ที่คำนวณโดยวิธีกำหนดเกณฑ์ผ่านระดับต่ำสุดได้เปอร์เซ็นต์ของคะแนนจุดตัดอยู่ระหว่าง 75 - 83.33 เปอร์เซ็นต์ เมื่อคำนวณโดยวิธีของเบอร์กได้เปอร์เซ็นต์ของคะแนนจุดตัดอยู่ระหว่าง 50 - 83.33 เปอร์เซ็นต์ และเมื่อคำนวณโดยวิธีการประยุกต์ราสช์โมเดล พบว่า ได้เปอร์เซ็นต์ของคะแนนจุดตัดเท่ากับ 50 เปอร์เซ็นต์ ซึ่งจะเห็นได้ว่า คะแนนจุดตัดที่ได้จากการคำนวณโดยวิธีกำหนดเกณฑ์ผ่านระดับต่ำสุดมีเปอร์เซ็นต์สูงกว่าวิธีอื่น ๆ นอกจากนี้ยังพบว่า คะแนนผลสัมฤทธิ์รวมของกลุ่มรอบรู้ที่ได้จากการกำหนดคะแนนจุดตัดโดยวิธีของเบอร์ก วิธีประยุกต์ราสช์โมเดล และวิธีกำหนดเกณฑ์ผ่านระดับต่ำสุด มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ต่อมาในปี พ.ศ. 2531 ศุภลักษณ์ สินธนา (2531 : 58 - 61) ได้ทำการเปรียบเทียบผลการใช้คะแนนจุดตัดที่กำหนดโดยการประยุกต์ไลจิสติกโมเดล แบบพารามิเตอร์ตัวเดียว พารามิเตอร์ 2 ตัว และพารามิเตอร์ 3 ตัว ในแบบทดสอบอิงเกณฑ์วิชาคณิตศาสตร์ในระดับชั้นประถมศึกษาปีที่ 6 โดยการนำความยากของข้อสอบทุกข้อมาหาค่าความยากเฉลี่ย ของข้อสอบทั้งฉบับ ($\bar{b}$) แล้วนำค่าความยากเฉลี่ยของข้อสอบไปเทียบกับค่าความสามารถของนักเรียน (θ) ซึ่งได้มาจากการวิเคราะห์โดยวิธีเดียวกัน แล้วพิจารณาว่าความสามารถของนักเรียนที่มีค่าเท่ากับค่าความยากเฉลี่ยของข้อสอบนั้นได้มาจากคะแนนในการสอบเป็นเท่าไร คะแนนที่ได้จากการสอบนี้จะถือเป็นคะแนนจุดตัดของแต่ละวิธี ผลการวิจัยปรากฏว่า คะแนนจุดตัดที่ได้จากการวิเคราะห์

ด้วยโลจิสติกโมเดล แบบพารามิเตอร์ตัวเดียว พารามิเตอร์ 2 ตัว และพารามิเตอร์ 3 ตัว ไม่แตกต่างกันแต่จำนวนคนรอบรู้จากการใช้คะแนนจุดตัดทั้ง 3 วิธี แตกต่างกันอย่างมีนัยสำคัญที่ระดับ .01

อัจฉริยา ปราบอริพ่าย (2532 : 101 - 106) ได้เปรียบเทียบคะแนนจุดตัดของแบบทดสอบผลสัมฤทธิ์วิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1 และค่าดัชนีความคงที่ของการตัดสินคะแนนจุดตัด ซึ่งตัดสินโดยผู้เชี่ยวชาญที่ไม่มีสารสนเทศและผู้เชี่ยวชาญที่มีสารสนเทศ โดยใช้วิธีการของแองกอฟฟ์ กลุ่มตัวอย่างผู้เชี่ยวชาญเป็นครูผู้มีประสบการณ์สอนวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1 ในโรงเรียนมัธยมศึกษา สังกัดกรมสามัญศึกษา ในเขตกรุงเทพมหานคร ปีการศึกษา 2531 จำนวน 96 คน ซึ่งแบ่งเป็น 2 กลุ่ม ๆ ละ 48 คน ให้กลุ่มหนึ่งตัดสินคะแนนจุดตัดโดยไม่มีสารสนเทศ อีกกลุ่มหนึ่งตัดสินคะแนนจุดตัดโดยมีสารสนเทศ (อันได้แก่ ค่าอำนาจจำแนก ค่าความยาก ค่าการเดา ความถี่ของคะแนนสอบ สัดส่วนของผู้สอบที่ตอบข้อสอบแต่ละข้อได้ถูก และค่าความน่าจะเป็นที่จะตอบข้อสอบแต่ละข้อได้ถูก) ส่วนกลุ่มตัวอย่างนักเรียน คือนักเรียนชั้นมัธยมศึกษาปีที่ 1 ปีการศึกษา 2531 ในโรงเรียนมัธยมศึกษา สังกัดกรมสามัญศึกษา ในเขตกรุงเทพมหานคร จำนวน 2,341 คน ซึ่งใช้ในการสอบแบบทดสอบเพื่อหาค่าพารามิเตอร์ของข้อสอบ (อันได้แก่ ค่าอำนาจจำแนก ค่าความยาก และค่าการเดา) ผลจากการวิจัย พบว่าคะแนนจุดตัดที่ใช้เป็นเกณฑ์ตัดสินความสามารถขั้นต่ำที่ยอมรับได้ของผู้สอบ ซึ่งอยู่บนสเกลของคะแนนโดเมนหรือคะแนนจริงที่ได้จากการตัดสินของผู้เชี่ยวชาญที่ไม่มีสารสนเทศ และผู้เชี่ยวชาญที่มีสารสนเทศ เท่ากับ 0.68 และ 0.53 ตามลำดับ ซึ่งแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ส่วนในการหาค่าดัชนีความคงที่ของการตัดสินคะแนนจุดตัดซึ่งตัดสินโดยผู้เชี่ยวชาญที่ไม่มีสารสนเทศ และผู้เชี่ยวชาญที่มีสารสนเทศนั้น ผู้วิจัยได้แปลงคะแนนจุดตัดในสเกลของคะแนนโดเมนให้เป็นระดับความสามารถโดยอาศัยโค้งคุณลักษณะแบบทดสอบ ปรากฏว่าดัชนีความคงที่ของการตัดสินเท่ากับ 0.78 และ 0.76 ตามลำดับซึ่งไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

นอกจากนี้ มาลี จิตติวุฒิการ (2533 : 96 - 112) ได้ทำการศึกษาคะแนนจุดตัดที่คำนวณจากทฤษฎีการตอบสนองต่อข้อสอบ (IRT) และทฤษฎีตามแนวคลาสสิค โดยกำหนดคะแนนโดเมนต่างกัน คือ .5, .6, .7 และ .8 และความยาวแบบทดสอบต่างกันคือ 20 ข้อ และ

30 ข้อ ว่าให้ผลต่างกันเพียงใด และจะมีผลต่อค่าความน่าจะเป็นในการจำแนกผิดพลาดในลักษณะใด เพื่อจะใช้จำแนกกลุ่มผู้รอบรู้และไม่รอบรู้และศึกษาว่าผลของคะแนนจุดตัดที่มีต่อความเชื่อมั่นในการตัดสินใจที่คำนวณตามแนวคิด ทฤษฎีต่างกันแนวไม้มต่างกันอย่างไร ๓ จุดตัดที่ใช้ในการจำแนก กลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 สังกัดกรมสามัญศึกษา กระทรวงศึกษาธิการ เขตกรุงเทพมหานคร จำนวน 1,005 คน เครื่องมือที่ใช้ในการวิจัย เป็นแบบทดสอบอิงเกณฑ์ลักษณะอิงโดเมนแบบเลือกตอบ 5 ตัวเลือก จำนวน 120 ข้อ ผลการวิจัยปรากฏว่า คะแนนโดเมนที่กำหนดคือ .5, .6, .7 และ .8 คะแนนจุดตัดตามแนวคลาสสิค คือ 10, 12, 14 และ 16 ตามลำดับสำหรับความยาวแบบทดสอบ 20 ข้อ และคะแนนจุดตัด 15, 18, 21 และ 24 ตามลำดับสำหรับความยาวแบบทดสอบ 30 ข้อ แต่เมื่อนำแบบทดสอบดังกล่าวมาหาคะแนนจุดตัดตามแนวทฤษฎีการตอบสนองต่อข้อสอบที่สอดคล้องกับคะแนนโดเมนที่กำหนด คะแนนจุดตัดฉบับความยาวแบบทดสอบ 20 ข้อ คือ 12, 13, 12 และ 13 คะแนน ส่วนความยาวแบบทดสอบ 30 ข้อ คะแนนจุดตัดคือ 18, 16, 18 และ 19 คะแนน สำหรับค่าความน่าจะเป็นในการจำแนกความรอบรู้ผิดพลาดฉบับความยาว 20 ข้อ และ 30 ข้อ ที่คิดคำนวณคะแนนจุดตัดตามแนวคลาสสิค กับไบโนเมียลโมเดล จะให้ค่าคงที่และมีค่าเพิ่มขึ้นตามการกำหนดคะแนนโดเมนที่เพิ่มขึ้นส่วนที่คิดคำนวณคะแนนจุดตัดตามแนวทฤษฎีการตอบสนองต่อข้อสอบความน่าจะเป็นในการจำแนกความรอบรู้ผิดพลาดจะแปรเปลี่ยนไปตามคะแนนโดเมนและคะแนนจุดตัด โดยที่ความยาวแบบทดสอบ 20 ข้อ และ 30 ข้อ จะให้ค่าความน่าจะเป็นในการจำแนกความรอบรู้ผิดพลาดประมาณ 18% และ 15% ตามลำดับ นอกจากนั้นยังพบว่า ค่าความเชื่อมั่นในการจำแนกความรอบรู้ ถ้าใช้วิธีตามแนวทฤษฎีคลาสสิคที่สอดคล้องกับคะแนนโดเมนและความยาวแบบทดสอบที่กำหนด ค่าที่ได้จะประมาณ .5 - .6 ส่วนค่าความเชื่อมั่นที่ใช้ตามแนวทฤษฎีการตอบสนองต่อข้อสอบจะให้ค่าสูงมากประมาณ .94 - .95

จากงานวิจัยดังกล่าวจะเห็นว่า งานวิจัยของ ชวลิต ไพธิ์นคร (2528) และงานวิจัยของ ศุภลักษณ์ สินธนา (2531) นั้นได้นำทฤษฎีการตอบสนองต่อข้อสอบมาใช้ในการกำหนดคะแนนจุดตัดโดยตรง แต่งานวิจัยของ อัจฉริยา ปราบอริพ่าย (2532) และงานวิจัยของ มาลี จิตติวุฒิการ (2533) นั้นได้กำหนดเกณฑ์มาตรฐานด้วยวิธีการอย่างใดอย่างหนึ่งก่อนแล้วจึงนำทฤษฎีการตอบสนองต่อข้อสอบไปประยุกต์ใช้กล่าวคือ งานวิจัยของ อัจฉริยา ปราบอริพ่าย จะใช้วิธีการกำหนดเกณฑ์มาตรฐานด้วยวิธีการให้ผู้เชี่ยวชาญพิจารณาจากความสามารถต่ำสุดของ

ผู้สอบตามเทคนิคของแองกอฟฟ์แล้วแปลงคะแนนจากคะแนนไดเมน เป็นระดับความสามารถ ส่วนงานวิจัยของ มาลี จิตติวุฒิการ ได้ใช้วิธีการกำหนดเกณฑ์มาตรฐานก่อนเช่นเดียวกันโดยกำหนดในรูปของคะแนนไดเมนแล้วนำทฤษฎีการตอบสนองต่อข้อสอบมาประยุกต์ใช้โดยการแปลงจากคะแนนไดเมนเป็นระดับความสามารถ

<u>ความเชื่อมั่นในการตัดสินจำแนกความรอบรู้</u>

ความเชื่อมั่นในการตัดสินจำแนกความรอบรู้ เป็นการศึกษาสัมประสิทธิ์ความสอดคล้องในการตัดสินความรอบรู้ - ไม่รอบรู้ ซึ่งมีอยู่ด้วยกัน 5 วิธี (Berk. 1980 : 129 - 185) คือ

1 วิธีของคาร์เวอร์ (Carver)

2 วิธีของสวามินาธาน แฮมเบิลตัน และอัลไจนา (Swaminathan. Hambleton and Algina)

3 วิธีของ ฮวิน (Huynh)

4 วิธีของสับโคเวียค (Subkoviak)

5 วิธีของมาร์แชลล์ และเฮอร์เทล (Marshall and Haertel)

1 <u>วิธีของคาร์เวอร์</u>

คาร์เวอร์ได้เสนอวิธีการประมาณค่าความเชื่อมั่นในการตัดสินจำแนกความรอบรู้ออกเป็น 2 วิธี คือวิธีแรกใช้แบบทดสอบ 2 ฉบับ ที่เป็นแบบทดสอบคู่ขนานกัน และนำไปสอบกับกลุ่มตัวอย่างกลุ่มเดียวกันแล้วนำค่าร้อยละของผู้สอบที่จัดอยู่ในกลุ่มรอบรู้มาเปรียบเทียบกัน ถ้าค่าร้อยละจากการสอบของแบบทดสอบทั้ง 2 ฉบับ ใกล้เคียงกัน แสดงว่า แบบทดสอบมีความเชื่อมั่นในการตัดสินจำแนกความรอบรู้ วิธีที่ 2 เป็นการใช้แบบทดสอบเพียงฉบับเดียวกับผู้สอบ 2 กลุ่มที่จะนำมาเปรียบเทียบกันโดยการนำค่าร้อยละของผู้สอบที่จัดอยู่ในกลุ่มรอบรู้ทั้ง 2 กลุ่ม มาเปรียบเทียบกัน

2 วิธีของสวามินาธาน แฮมเบิลตัน และอัลไจนา

สวามินาธาน แฮมเบิลตันและอัลไจนา ได้เสนอวิธีการประมาณค่าความเชื่อมั่นในการตัดสินจำแนกความรอบรู้ในรูปแบบของสัดส่วนของความสอดคล้องในการจำแนกผู้สอบเป็นผู้รอบรู้และไม่รอบรู้ ด้วยการใช้แบบทดสอบ 2 ฉบับ โดยใช้สูตรสัมประสิทธิ์แคปปา (Kappa) ของโคเฮน (Cohen) ดังนี้

$$\hat{K} = \frac{\hat{P}_o - \hat{P}_c}{1 - \hat{P}_c}$$

เมื่อ $\hat{K}$ คือ สัดส่วนของความสอดคล้องในการจำแนกผู้สอบเป็นผู้รอบรู้จากการสอบด้วยข้อสอบคู่ขนาน 2 ฉบับ

$\hat{P}_o$ คือ สัดส่วนของความสอดคล้องที่ตัดสินว่ารอบรู้/รอบรู้ หรือไม่รอบรู้/ไม่รอบรู้ในแบบทดสอบทั้ง 2 ฉบับ

$\hat{P}_c$ คือ สัดส่วนของความสอดคล้องที่คาดหวังซึ่งอาจเกิดขึ้นโดยบังเอิญ

โดยที่

$$\hat{P}_o = \sum_{k=1}^{m} \hat{P}_{kk}$$

$$\hat{P}_c = \sum_{k=1}^{m} \hat{P}_{k.} \hat{P}_{.k}$$

เมื่อ $\hat{P}_{kk}$ คือ สัดส่วนของความสอดคล้องในการแยกผู้สอบเป็นประเภทที่ k จากแบบทดสอบทั้ง 2 ฉบับ

$\hat{P}_{k.}$ และ $\hat{P}_{.k}$ คือ สัดส่วนของผู้สอบที่จัดอยู่ในประเภทที่ k จากการสอบแบบทดสอบฉบับที่ 1 และ 2 ตามลำดับ

ต่อไปนี้เป็นตัวอย่างการหาค่าความเชื่อมั่นตามวิธีการของ สวามินาธาน แฮมเบิลตัน และอัลไจนา

สมมุติว่าจากการใช้แบบทดสอบคู่ขนานกัน 2 ฉบับ ไปสอบกับนักเรียนกลุ่มหนึ่งจำนวน 30 คน มีผู้รอบรู้และไม่รอบรู้จากการสอบแบบทดสอบแต่ละฉบับดังนี้

แบบทดสอบฉบับที่ 1	แบบทดสอบฉบับที่ 2		รวม
	ผู้รอบรู้	ผู้ไม่รอบรู้	
ผู้รอบรู้	1	1	2
ผู้ไม่รอบรู้	0	28	28
รวม	1	29	30

$$\hat{P}_o = \sum_{k=1}^{m} \hat{P}_{kk}$$

$$= \frac{1}{30} + \frac{28}{30}$$

$$= \frac{29}{30}$$

$$= .97$$

$$\hat{P}_c = \sum_{k=1}^{m} \hat{p}_{k\bullet}\ \hat{p}_{\bullet k}$$

$$= \frac{2}{30} \times \frac{1}{30} + \frac{28}{30} \times \frac{29}{30}$$

$$= \frac{2}{900} + \frac{812}{900}$$

$$= .90$$

$$\hat{K} = \frac{\hat{P}_o - \hat{P}_c}{1 - \hat{P}_c}$$

$$= \frac{.97 - .90}{1 - .90}$$

$$= .70$$

3 <u>วิธีของฮวิน</u>

ฮวิน ได้เสนอวิธีการหาความเชื่อมั่นในการตัดสินจำแนก ความรอบรู้จากการสอบ 2 ครั้ง หรือครั้งเดียวก็ได้ โดยการใช้สัมประสิทธิ์แคปปา (K) และตั้งข้อสมมุติว่ากระบวนการสุ่มความน่าจะเป็นในการเลือกข้อคำถามจากข้อคำถามทั้งหมดที่มีอยู่ได้นิยามไว้อย่างชัดเจน ถ้าแบบทดสอบประกอบด้วยข้อคำถาม n ข้อ ผู้สอบที่มีความสามารถ θ มีคะแนนสอบเท่ากับ X และคะแนนเหล่านี้มีการกระจายแบบ Binomial Density ดังนี้

$$f(x \mid \theta) = \binom{n}{x} \theta^x (1 - \theta)^{n-x},\ x = 0, 1, \ldots, n$$

จากแบบจำลองไบโนเมียลนี้สมมุติว่าข้อคำถามทั้งหมดในจักรวาลของข้อสอบสามารถใช้แทนกันได้ นั่นแสดงว่าการแจกแจงของคะแนนสอบของข้อคำถาม n ข้อ ไม่ขึ้นอยู่กับข้อคำถามใดที่เลือกขึ้นมาโดยเฉพาะซึ่งชี้ให้เห็นว่าข้อคำถามมีความยากเหมือน ๆ กัน

จาก Beta Density นี้จะใช้พารามิเตอร์ α และ β เป็นความสามารถแท้จริง (θ) ถ้าขนาดของกลุ่มตัวอย่างมาก จะสามารถ หาα, β ได้จากการหาความเชื่อมั่นโดยใช้ KR - 21 ดังนี้

$$\alpha_{21} = \frac{n}{n-1}\left[1 - \frac{\mu(n-\mu)}{n\sigma^2}\right]$$

ดังนั้น $\alpha = (-1 + 1/\alpha_{21})\mu$

และ $\beta = -\alpha + n/\alpha_{21} - n$

จาก Beta Density ข้างต้น จะได้ Marginal Density ของคะแนน X อยู่ในรูปของ Negative Hypergeometric (หรือ Beta - Binomial) ดังนี้

$$f(x) = \binom{n}{x} B(\alpha + x, n + \beta - x)/B(\alpha,\beta)$$

เมื่อ β คือ Beta Function สำหรับ $X \longrightarrow 0$ f(X) สามารถหาได้ดังนี้

$$f(0) = \prod_{i=1}^{n} \frac{n + \beta - i}{n + \alpha + \beta - i}$$

และ

$$f(x+1) = f(x)\cdot\frac{(n-x)(\alpha+x)}{(x+1)(n+\beta-x-1)}, x = 0, 1, \ldots, n-1.$$

ถ้า $X \longrightarrow n$ จะได้ดังนี้

$$f(n) = \prod_{i=1}^{n} \frac{n + \alpha - i}{n + \alpha + \beta - i}$$

และ $$f(x-1) = f(x)\cdot\frac{x(n+\beta-x)}{(n-x+1)(\alpha+x-1)}, x = 1,\ldots,n.$$

ถ้า X และ Y เป็นคะแนนที่ได้จากการสอบฟอร์ม X และ Y ภายใต้เงื่อนไขของความเป็นอิสระในการตอบสนองต่อข้อสอบ (Local Independence) แล้ว คะแนน X และ Y จะมีการแจกแจงแบบ Bivariate Beta - Binomial (หรือ NegativeHypergeometric) ด้วย Joint Density ดังนี้

$$f(x,y) = \frac{\binom{n}{x}\binom{n}{y}}{B(\alpha,\beta)} B(\alpha+x+y, 2n+\beta-x-y)$$

สมการนี้จะสมมาตรถ้า f(X,Y) = f(Y,X) สำหรับค่า X และ Y มีค่าใกล้ 0 ดังนั้น f(X,Y) สามารถหาได้ดังนี้

$$f(0,0) = \prod_{i=1}^{2n}\frac{2n+\beta-i}{2n+\alpha+\beta-i} = f(0)\cdot\prod_{i=1}^{n}\frac{2n+\beta-i}{2n+\alpha+\beta-i}$$

และ $$f(x+1,y) = f(x,y)\cdot\frac{(n-x)(\alpha+x+y)}{(x+1)(2n+\beta-x-y-1)}.$$

ถ้า X และ Y มีค่าใกล้ n สูตรในการหาจะเป็นดังนี้

$$f(n,n) = \prod_{i=1}^{2n}\frac{2n+\alpha-i}{2n+\alpha+\beta-i} = f(n)\cdot\prod_{i=1}^{n}\frac{2n+\alpha-i}{2n+\alpha+\beta-i}$$

และ $$f(x-1,y) = f(x,y)\cdot\frac{x(2n+\beta-x-y)}{(n-x+1)(\alpha+x+y-1)}.$$

การประเมินค่า K โดยการใช้แบบทดสอบคู่ขนาน 2 ฟอร์ม

ถ้าการสอบที่ใช้แบบทดสอบคู่ขนาน 2 ฟอร์ม คือฟอร์ม X และ Y โดยมีคะแนนจุดตัด c ถ้าผู้สอบที่ได้คะแนนสอบมากกว่าคะแนนจุดตัดจะถือว่าเป็นผู้รอบรู้ และผู้สอบที่ได้คะแนนต่ำกว่าคะแนนจุดตัดจะถือว่าเป็นผู้ไม่รอบรู้ซึ่งสามารถเขียนได้ดังนี้

ฟอร์ม Y / ฟอร์ม X	ไม่รอบรู้	รอบรู้	ค่าสัดส่วน
ไม่รอบรู้	P_{oo}		P_o
รอบรู้		P_{11}	P_1
ค่าสัดส่วน	P_o	P_1	

ถ้าคะแนนจุดตัดมีค่าใกล้ n สูตรสัมประสิทธิ์ความเชื่อมั่นของแคปปาหาได้ดังนี้

$$K = (P_{11} - P_1^2)/(P_1 - P_1^2)$$

$$P_{11} = \sum_{x,y=c}^{n} f(x,y)$$

$$P_1 = \sum_{x=c}^{n} f(x)$$

ถ้าคะแนนจุดตัดมีค่าน้อยหรือใกล้ 0 สูตรในการหาสัมประสิทธิ์ความเชื่อมั่นของแคปปาหาได้ดังนี้

$$K = (P_{oo} - P_o^2) / (P_o - P_o^2)$$

$$\text{เมื่อ} \quad P_{oo} = \sum_{x,y=0}^{c-1} f(x,y)$$

$$\text{และ} \quad P_{o} = \sum_{x=0}^{c-1} f(x)$$

<u>การประมาณค่า K จากการสอบเพียงครั้งเดียว</u>

ในการประมาณค่า K จากการสอบเพียงครั้งเดียวมีขั้นตอนดังนี้

1. คำนวณค่าเฉลี่ย ($\bar{X}$) และส่วนเบี่ยงเบนมาตรฐาน (S) ของคะแนนที่ได้จากการสอบ

2. ประมาณค่า α และ β โดยแทนค่า μ ด้วย $\bar{X}$ และ σ ด้วย S ดังนี้

$$\alpha_{21} = \frac{n}{n-1}\left[1 - \frac{\bar{x}(n-\bar{x})}{ns^2}\right]$$

$$\alpha = (-1 + 1/\alpha_{21})\bar{x}$$

$$\beta = -\alpha + n/\alpha_{21} - n$$

โดยที่ n คือ จำนวนข้อคำถามในแบบทดสอบ

3. การประมาณค่า K มี 2 กรณี คือ

3.1 ถ้าคะแนนจุดตัด (c) มีค่าน้อยหรือใกล้ 0 คำนวณ P_o และ P_{oo} ดังนี้

$$P_o = f(0) + \ldots + f(c-1)$$

$$P_{oo} = f(0,0) + f(0,1) +...+ f(c-1,c-1)$$

ดังนั้น

$$K = (P_{oo} - P_o^2) / (P_o - P_o^2)$$

3.2 ถ้าคะแนนจุดตัด (c) มีค่าบวก คำนวณ P_1 และ P_{11}

ดังนี้

$$P_1 = f(n) +,...,+ f(c)$$

$$P_{11} = f(n,n) + f(n,n-1) + f(n-1,n) +,...,+ f(c,c)$$

ดังนั้น

$$K = (P_{11} - P_1^2)/(P_1 - P_1^2)$$

<u>การประมาณค่า K เมื่อแบบทดสอบมีจำนวนข้อคำถามมาก</u>

ถ้าแบบทดสอบประกอบด้วยจำนวนข้อคำถามมาก เช่น มีจำนวนมากกว่า 10 ข้อ การประมาณค่า K ดังที่กล่าวไว้ข้างต้นค่อนข้างจะยุ่งยากมาก ดังนั้น ฮวินจึงเสนอวิธีการตามที่คีทส์และลอร์ด (Huynh. 1976 : 258 ; citing Keats and Lord. 1962) เสนอไว้ดังนี้

1. แปลงคะแนนที่ได้จากการสอบ จากสูตร

$$X' = \sin^{-1}\sqrt{X/n}$$

ซึ่งมีการกระจายเป็นปกติโดยมีค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานดังนี้

$$\mu_{x'} = \sin^{-1}\sqrt{\mu/n}$$

$$\sigma_{x'} = [(\alpha_{21} + 1)/(\alpha + n)]^{1/2}.$$

ถ้า X' และ Y' เป็นคะแนนที่ได้จากการแปลงจากคะแนน X และ Y จากแบบทดสอบคู่ขนาน 2 ฟอร์มแล้ว ค่าสหสัมพันธ์ของคะแนนทั้ง 2 จะมีค่าดังนี้

$$\rho = \alpha_{21}[(n - 1)/(n + \alpha_{21})]^{1/2}$$

ไนวิค เลวิส และแจ็คสัน (Huynh. 1976 : 258 ; citing Novick, Lewis and Jackson. 1973) ได้ใช้วิธีการแปลงคะแนนแบบเดียวกันนี้กับตัวแปรทวินาม (Binomial Variable) และพบว่า การประมาณค่าจะสมเหตุสมผลดีเมื่อ μ /n มีค่าอยู่ระหว่าง .15 และ .85 และจำนวนข้อสอบอย่างน้อย 8 ข้อ

เนื่องจากฟังก์ชัน Arcsin มีลักษณะเป็นการเพิ่มไปในทางเดียวกัน การกำหนดผู้รอบรู้ ก็ต่อเมื่อ $X' \geq c'$ โดยที่ $c' = \sin^{-1}\sqrt{(c - .5)/n}$ และใช้ตารางของกุปตา (Huynh. 1976 : 259 ; citing Gupta. 1963) ก็จะสามารถประมาณค่า P_o และ P_{oo} ได้

เพื่อให้เห็นภาพชัดเจนขึ้น ต่อไปนี้จะกล่าวโดยสรุปถึงการประมาณค่าแคปปา ดังนี้

1. คำนวณค่า $\bar{X}$ และ S ของคะแนนที่ได้จากการสอบ
2. ประมาณค่า $\mu_{x'}$, $\sigma_{x'}$ และ ρ โดยแทน μ ด้วย $\bar{X}$ และ σ ด้วย S ดังนี้

$$\alpha_{21} = \frac{n}{n - 1}\left[1 - \frac{\bar{X}(n - \bar{X})}{nS^2}\right]$$

$$\alpha = (-1 + 1/\alpha_{21})\bar{X}$$

$$\beta = -\alpha + n/\alpha_{21} - n$$

$$\mu_{x'} = \sin^{-1}\sqrt{\bar{X}/n}$$

$$\sigma_{x'} = [(\alpha_{21} + 1) / (\alpha + n)]^{1/2}$$

$$\rho = \alpha_{21}[(n - 1)/(n + \alpha_{21})]^{1/2}$$

เมื่อ n คือ จำนวนข้อคำถามในแบบทดสอบ

3. คำนวณอัตราส่วน $z = (c' - \mu_{x'})/\sigma_{x'}$

4. หาค่าความน่าจะเป็นของตัวแปรปกติมาตรฐานที่มีค่าน้อยกว่า z เพื่อ ประมาณค่า P_o

5. หาค่าความน่าจะเป็นของตัวแปรปกติมาตรฐาน 2 ค่าที่มีค่าสหสัมพันธ์ ρ น้อยกว่า z เพื่อประมาณค่า P_{oo}

6. ประมาณค่า K จากสูตร

$$K = (P_{oo} - P_o^2)/(P_o - P_o^2)$$

4. <u>วิธีของสับไคเวียค</u>

สับไคเวียค ได้เสนอวิธีการหาความเชื่อมั่นของการตัดสินเพื่อจำแนกผู้สอบเป็นผู้รอบรู้และไม่รอบรู้จากการสอบเพียงครั้งเดียว โดยมีวิธีการในการหาดังนี้

$$P_C = \frac{\sum_{i=1}^{N} P_C^{(i)}}{N}.$$

เมื่อ P_c คือ ความน่าจะเป็นของการตัดสิน

N คือ จำนวนผู้สอบทั้งหมด

$P_c^{(i)}$ คือ สัมประสิทธิ์ของความสอดคล้องของคนที่ i เมื่อกำหนดคะแนนจุดตัดเท่ากับ c

โดยที่ $$P_c^{(i)} = [P(x_i \geq c)]^2 + [1 - P(x_i \geq c)]^2$$

$$P(X_i \geq C) = \sum_{X_i = C}^{n} \binom{n}{X_i} p_i^{X_i} (1 - p_i)^{n - X_i}$$

เมื่อ P_i คือ ความน่าจะเป็นที่แท้จริงในการตอบข้อคำถามถูกของคนที่ i ค่าของ P_i สามารถหาได้ดังนี้

1. ถ้าแบบทดสอบมีจำนวนข้อสอบมากกว่า 40 ข้อ

$$\hat{p}_i = X_i / n$$

เมื่อ X_i คือ จำนวนข้อที่ตอบถูก

n คือ จำนวนข้อสอบทั้งหมดในแบบทดสอบ

2. ถ้าแบบทดสอบมีจำนวนข้อสอบน้อย

$$\hat{p}_i = \alpha_{21/x} \left(\frac{X_i}{n}\right) + (1 - \alpha_{21/x}) \left(\frac{M_x}{n}\right)$$

โดยที่ $$\alpha_{21/x} = \frac{n}{n - 1} \left[1 - \frac{M_x(n - M_x)}{nS_x^2}\right]$$

เมื่อ $\alpha_{21/x}$ คือ สัมประสิทธิ์ความเชื่อมั่นของแบบทดสอบโดยวิธี Kuder Richardson Formula 21

M_x คือ คะแนนเฉลี่ยที่ได้จากการสอบ

S_x^2 คือ ความแปรปรวนของคะแนนที่ได้จากการสอบ

5. วิธีของมาร์แชลล์ - เฮอร์เทล

มาร์แชลล์ - เฮอร์เทล ได้เสนอวิธีการหาความเชื่อมั่นในการตัดสินจำแนกความรอบรู้จากการสอบเพียงครั้งเดียว โดยมีข้อตกลงเบื้องต้นว่า ถ้าผู้สอบแต่ละคนได้รับการทดสอบซ้ำ การแจกแจงของคะแนนจะมีลักษณะเป็นไบโนเมียล วิธีการของมาร์แชลล์ - เฮอร์เทล คือ แบ่งครึ่งแบบทดสอบออกเป็น 2 ฉบับ เพื่อหาสัดส่วนของความคงที่ในการแบ่งแบบทดสอบออกเป็น 2 ฉบับ $(\hat{P}_o)$ ดังนี้

1. กำหนดให้ X แทนคะแนนที่ได้จากการสอบ N_x แทนความถี่ของคะแนนที่ได้จากการสอบแบบทดสอบฉบับที่ 1 และ N_w แทนความถี่ของคะแนน (w) จากการสอบด้วยแบบทดสอบที่มีความยาวเป็น 2 เท่าของแบบทดสอบฉบับที่ 1 ดังนั้น

$$N_w = \sum_{x=0}^{n} N_x \binom{2n}{w} \left(\frac{x}{n}\right)^w \left(1 - \frac{x}{n}\right)^{2n-w}$$

2. หาความน่าจะเป็นของความสอดคล้องในการจำแนกผู้สอบด้วยการใช้แบบทดสอบที่แบ่งครึ่งออกเป็น 2 ฉบับ ซึ่งมีคะแนนสอบเป็น w จากการสอบด้วยแบบทดสอบที่มีความยาวเป็น 2 เท่า ดังนี้

$$\Phi_w(a,b) = \sum_{j=a}^{b} \frac{\binom{w}{j}\binom{2n-w}{n-j}}{\binom{2n}{n}}$$

โดยที่ a = w - [c - 1]

b = c - 1

เมื่อ c คือ คะแนนจุดตัด

3. คำนวณค่าความเชื่อมั่นในการตัดสินความรอบรู้จากการสอบ ดังนี้

$$\hat{p}_o = \frac{1}{N}\left[\sum_{w=0}^{c-1} N_w + \sum_{w=c}^{2c-2} N_w \cdot \Phi_w(w-[c-1], c-1) + \sum_{w=2c}^{n+c-1} N_w \cdot \Phi_w(c, w-c) + \sum_{w=n+c}^{2n} N_w\right]$$

เมื่อ N คือ จำนวนผู้สอบ

n คือ จำนวนข้อคำถาม

c คือ คะแนนจุดตัด

N_w คือ ความถี่ของคะแนน w

$\phi_w(a,b)$ คือ ความน่าจะเป็นของความสอดคล้องในการจำแนก

การเปรียบเทียบวิธีการของ สวามินาธาน ฮวิน สับโคเวียค และมาร์แชลล์

เบอร์ก (Berk. 1980 : 150) ได้สรุปถึงข้อดีและข้อจำกัดของการหาความเชื่อมั่นจากการใช้วิธีการของสวามินาธาน ฮวิน สับโคเวียค และมาร์แชลล์ ดังนี้

วิธีการ	ข้อดี	ข้อจำกัด
สวามินาธาน	1. คำนวณง่าย 2. เป็นการประมาณที่ไม่มีความลำเอียง	1. ต้องใช้การสอบ 2 ครั้ง 2. การประมาณค่าจะมีความคลาดเคลื่อนมาก เมื่อกลุ่มตัวอย่างมีขนาดเล็ก
ฮวิน	1. ดำเนินการทดสอบเพียงครั้งเดียว 2. การประมาณค่าจะมีความคลาดเคลื่อนน้อย	1. คำนวณยุ่งยาก 2. การประมาณค่าจะมีความลำเอียงสำหรับแบบทดสอบที่สั้น
สับโคเวียค	1. ดำเนินการทดสอบเพียงครั้งเดียว 2. การประมาณค่าจะมีความคลาดเคลื่อนน้อย	1. คำนวณยุ่งยาก 2. การประมาณค่าจะมีความลำเอียงสำหรับแบบทดสอบที่สั้น
มาร์แชลล์	1. ดำเนินการสอบเพียงครั้งเดียว 2. การประมาณค่าจะมีความคลาดเคลื่อนน้อย	1. คำนวณยุ่งยาก 2. การประมาณค่าจะมีความลำเอียงสำหรับแบบทดสอบที่สั้น

จากวิธีการทั้ง 4 ดังกล่าวข้างต้น เบอร์ก (Berk. 1980 : 147 - 150) ได้สรุปถึงการศึกษาในเชิงประจักษ์กับนักเรียนจำนวน 1,586 คน โดยแบ่งนักเรียนออกเป็น 50 กลุ่ม ๆ ละ 30 คน ด้วยการใช้วิธีการสุ่มแบบแทนที่ ให้ทำแบบทดสอบคู่ขนานที่มีจำนวน 10, 30 และ 50 ข้อ แบบทดสอบแต่ละฉบับได้กำหนดเกณฑ์การรอบรู้เป็น 4 ระดับ คือ 50 เปอร์เซ็นต์ 60 เปอร์เซ็นต์ 70 เปอร์เซ็นต์ และ 80 เปอร์เซ็นต์ ผลจากการศึกษาปรากฏว่า การใช้วิธีการของสวามินาธานจะได้ค่าความคลาดเคลื่อนมาตรฐาน (Standard Error) สูงกว่าวิธีของฮวิน สับไคเวียค และมาร์แชลล์ นอกจากนี้ยังพบว่า จากการใช้วิธีการทั้ง 4 ค่าความคลาดเคลื่อนมาตรฐานจะลดลงเมื่อแบบทดสอบมีความยาวเพิ่มขึ้น หากกล่าวโดยสรุปจะพบว่าข้อดีและข้อจำกัดสำหรับวิธีของ ฮวิน สับไคเวียค และมาร์แชลล์ จะเหมือนกัน กล่าวคือ ดำเนินการสอบเพียงครั้งเดียว การประมาณค่าจะมีความคลาดเคลื่อนน้อย โดยเฉพาะอย่างยิ่งวิธีการของฮวิน มีแนวโน้มที่จะให้ค่าเฉลี่ยที่ต่ำกว่าวิธีอื่น ๆ สำหรับแบบทดสอบที่มีความยาว 10 ข้อ หากพิจารณาในภาพรวมแล้ว วิธีการของ ฮวิน ดีกว่าวิธีอื่น

ดังนั้นในการวิจัยครั้งนี้ ผู้วิจัยได้เลือกวิธีการศึกษาความเชื่อมั่นในการตัดสินจำแนกความรอบรู้ตามวิธีของฮวิน เนื่องจากเป็นวิธีที่เหมาะสมมากกว่าวิธีอื่น ตามที่เบอร์ก (Berk. 1980 : 150) สรุปไว้ ถึงแม้ว่าวิธีการนี้จะคำนวณค่อนข้างยุ่งยากก็ตาม แต่ในปัจจุบันสามารถใช้คอมพิวเตอร์ช่วยในการคิดคำนวณได้ซึ่งปัญหาดังกล่าวก็จะหมดสิ้นไป

ในการวิจัยครั้งนี้ ผู้วิจัยจะนำข้อมูลของโครงการตรวจสอบคุณภาพการศึกษามาใช้ ดังนั้นผู้วิจัยจึงขอกล่าวถึงโครงการตรวจสอบคุณภาพการศึกษาโดยสรุปดังนี้ (กรมวิชาการ. ม.ป.ป. : 5 - 6, 46 - 50)

<u>โครงการตรวจสอบคุณภาพการศึกษา</u>

โครงการตรวจสอบคุณภาพการศึกษาเป็นโครงการที่กระทรวงศึกษาธิการมอบให้กรมวิชาการตรวจสอบคุณภาพการศึกษาทุกระดับภายหลังจากที่กระทรวงศึกษาธิการได้ยกเลิกการสอบรวมแล้วมอบให้แต่ละโรงเรียนดำเนินการวัดและประเมินผลการศึกษาเอง การตรวจสอบคุณภาพการศึกษาเริ่มเป็นครั้งแรกในปีการศึกษา 2524 โดยมีวัตถุประสงค์ในการประเมินผลสัมฤทธิ์

ทางการเรียนของนักเรียนและใช้ผลการประเมินประกอบการพิจารณาปรับปรุงคุณภาพการศึกษา สำหรับในปีการศึกษา 2524 กรมวิชาการทำการตรวจสอบคุณภาพการศึกษาเฉพาะนักเรียนชั้นมัธยมศึกษาปีที่ 3 ต่อมาในปีการศึกษา 2525 กรมวิชาการจึงทำการตรวจสอบคุณภาพการศึกษานักเรียนชั้นมัธยมศึกษาปีที่ 6 ในด้านความรู้ และความคิด (Cognitive) จากการเรียนวิชาภาษาไทย สังคมศึกษา พลานามัย คณิตศาสตร์ และวิทยาศาสตร์ ในปีการศึกษา 2526 จนถึงปีการศึกษา 2530 การตรวจสอบคุณภาพการศึกษาก็ยังคงมุ่งวัดคุณภาพของนักเรียนด้านผลสัมฤทธิ์ทางวิชาการในทุกกลุ่มวิชาหรือทุกกลุ่มประสบการณ์เช่นกัน แต่เพิ่มการประเมินด้านลักษณะนิสัยและบุคลิกภาพบางส่วนขึ้น และทำการตรวจสอบคุณภาพนักเรียนชั้นมัธยมศึกษาปีที่ 6 ทุกปีการศึกษา ส่วนนักเรียนชั้นประถมศึกษาปีที่ 6 และมัธยมศึกษาปีที่ 3 สอบสลับปีกันจะเห็นได้ว่าการตรวจสอบคุณภาพการศึกษาตั้งแต่ปีการศึกษา 2524จนถึงปีการศึกษา 2530 กรมวิชาการทำการประเมินคุณภาพการศึกษาจากผู้เรียนอย่างเดียวมิได้ตรวจสอบองค์ประกอบอื่น ๆ อันได้แก่ สภาพของโรงเรียน วิธีการจัดการเรียนการสอน สภาพของผู้บริหาร อุปกรณ์การสอน ตลอดจนการบริหารหลักสูตรจึงทำให้ข้อมูลที่ได้ไม่สามารถนำไปพัฒนาการศึกษาด้านต่าง ๆ ได้

ในปีการศึกษา 2531 กรมวิชาการทำการประเมินทั้ง 3 ระดับ พร้อมกันคือ ประถมศึกษาปีที่ 6 มัธยมศึกษาปีที่ 3 และมัธยมศึกษาปีที่ 6 และได้เปลี่ยนรูปแบบการตรวจสอบคุณภาพการศึกษาไปจากเดิมคือ มีการประเมินคุณภาพของผู้เรียนและประเมินกระบวนการจัดการศึกษาของโรงเรียนพร้อมกัน ในการประเมินคุณภาพของผู้เรียนนั้นใช้เครื่องมือวัดทั้ง ความรู้ความสามารถทางวิชาการ ความสามารถในการคิดและแก้ปัญหา เจตคติที่ดีต่อการศึกษา ตลอดจนการมีพฤติกรรมที่ส่งเสริมการเรียนรู้ ส่วนการประเมินกระบวนการจัดการศึกษาของโรงเรียนนั้นประเมินในด้านสภาพพื้นฐานของนักเรียน สภาพทั่วไปของนักเรียน บุคลากรในโรงเรียนและสภาพการนำหลักสูตรไปใช้ นอกจากนี้สิ่งสำคัญที่มีการเปลี่ยนแปลงคือ การกำหนดเกณฑ์การผ่านใช้วิธีการหาคะแนนจุดตัดตามแนวคิดของนีเดลสกี แทนการใช้เกณฑ์ 60 เปอร์เซ็นต์ของคะแนนเต็มหรือ 80 เปอร์เซ็นต์ ของคะแนนเต็ม

สำหรับในปีการศึกษา 2533 กรมวิชาการก็มีนโยบายตรวจสอบคุณภาพการศึกษาของนักเรียนทุกระดับชั้นอีกเช่นเดียวกันโดยเปลี่ยนชื่อเป็นโครงการติดตามประเมินการใช้หลักสูตรของสถานศึกษา การประเมินการใช้หลักสูตรในปีการศึกษา 2533 นี้ ประเมินการใช้หลักสูตร 2 ด้าน

คือด้านผลสัมฤทธิ์ของนักเรียน และด้านการจัดการศึกษาของโรงเรียน (กรมวิชาการ. ม.ป.ป. : 1 - 24) ดังนี้

1. ผลสัมฤทธิ์ของนักเรียน

ในการประเมินผลสัมฤทธิ์ของนักเรียนแต่ละระดับชั้นพิจารณาจากกรอบแนวคิดพื้นฐานในการพัฒนาหลักสูตรแบ่งผลสัมฤทธิ์ของนักเรียนเป็น คุณลักษณะเพื่อพัฒนาตน คุณลักษณะเพื่อพัฒนาอาชีพ และคุณลักษณะเพื่อพัฒนาสังคม

2. การจัดการศึกษาของโรงเรียน

ในการประเมินการจัดการศึกษาของโรงเรียนใช้แนวคิดในการพัฒนางานเชิงระบบโดยวิเคราะห์งานโรงเรียนเป็น 3 ด้าน คือด้านการเตรียมความพร้อม ด้านการบริหารหลักสูตร และด้านการเรียนการสอน

ส่วนเครื่องมือที่ใช้ในโครงการตรวจสอบคุณภาพการศึกษาในปีการศึกษา 2533 แบ่งออกเป็นเครื่องมือที่ใช้ในการประเมินผลสัมฤทธิ์ของนักเรียน และเครื่องมือการประเมินการจัดการศึกษาของโรงเรียน สำหรับเครื่องมือที่ใช้ในการประเมินผลสัมฤทธิ์ของนักเรียนมีดังนี้

1. แบบวัดความสามารถทั่วไป ฉบับที่ 1 เป็นการวัดความรู้พื้นฐานในการใช้ภาษาเพื่อการสื่อสาร

2. แบบวัดความสามารถทั่วไป ฉบับที่ 2 เป็นการวัดความรู้พื้นฐานในการคิดคำนวณความสามารถนำไปใช้ในชีวิตประจำวันได้

3. แบบวัดความสามารถทั่วไป ฉบับที่ 3 เป็นการวัดความเข้าใจในธรรมชาติและสิ่งแวดล้อม

4. แบบวัดการคิดและแก้ปัญหา เป็นการวัดการเสนอแนวทางในการแก้ปัญหาของตนเองและครอบครัว

5. แบบวัดคุณลักษณะทั่วไป เป็นการวัดสมรรถภาพเกี่ยวกับการมีนิสัยไม่เอาเปรียบผู้อื่น มีนิสัยรักการอ่าน และแสวงหาความรู้อยู่เสมอ และมีสุขนิสัยที่ดีในการบำรุงรักษาสุขภาพอนามัยของตน

6. แบบวัดคุณลักษณะเกี่ยวกับการพัฒนาอาชีพ เป็นการวัดสมรรถภาพเกี่ยวกับการรู้และเข้าใจ กระบวนการทำงาน ความสามารถทำงานร่วมกับผู้อื่นและรักการทำงาน

7. แบบวัดคุณลักษณะเกี่ยวกับการพัฒนาสังคม เป็นการวัดสมรรถภาพเกี่ยวกับการรู้และเข้าใจสภาพและการเปลี่ยนแปลงของสังคมที่บ้าน สามารถปฏิบัติตนตามฐานะบทบาทหน้าที่ในฐานะที่ดีของบ้าน และเห็นคุณค่าของการอนุรักษ์สิ่งแวดล้อม ศิลปวัฒนธรรมในชุมชนรอบ ๆ ด้าน

ส่วนเครื่องมือการประเมินการจัดการศึกษาของโรงเรียนมีดังนี้

1. แบบประเมินการเตรียมความพร้อมของโรงเรียน เป็นการประเมินเกี่ยวกับการประชาสัมพันธ์หลักสูตร การประสานงานกับหน่วยงานต่าง ๆ การเตรียมพร้อมของบุคลากร การจัดการบริหารงาน การจัดระบบข้อมูลพื้นฐาน สื่อ วัสดุ อุปกรณ์และเอกสารที่เกี่ยวข้อง ความมั่นใจของคณะครู ผู้บริหาร

2. แบบประเมินการบริหารหลักสูตร เป็นการประเมินเกี่ยวกับการวางแผนทั่วไป การจัดทำแผนด้านวิชาการ การดำเนินงานตามแผน การประเมินผลการดำเนินงาน

3. แบบประเมินการเรียนการสอน เป็นการประเมินเกี่ยวกับการจัดและพัฒนาการเรียนการสอนให้เหมาะสมกับจุดหมายหลักสูตร การสอนเน้นกระบวนการ การสอนเน้นผู้เรียนเป็นศูนย์กลาง การสอนเน้นกระบวนการกลุ่ม การสอนซ่อมเสริมเพื่อพัฒนาผู้เรียน

บทที่ 3

วิธีดำเนินการวิจัย

การวิจัยเรื่อง การประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบเพื่อกำหนดคะแนนจุดตัดตามเทคนิคของนีเดลสกีกับวิธีการใช้กลุ่มคาบเส้น ผู้วิจัยดำเนินการดังนี้

1. กลุ่มตัวอย่าง
2. เครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูล
3. คุณภาพของเครื่องมือ
4. การเก็บรวบรวมข้อมูล
5. การวิเคราะห์ข้อมูล

<u>กลุ่มตัวอย่าง</u>

กลุ่มตัวอย่างที่ใช้ในการวิจัยครั้งนี้แบ่งเป็น 2 กลุ่ม คือ

1. กลุ่มตัวอย่างผู้เชี่ยวชาญที่ใช้ในการพิจารณาแบบทดสอบเพื่อกำหนดคะแนนจุดตัด ได้แก่ ผู้ที่มีความรู้ด้านหลักสูตร ด้านการวัดผลการศึกษา ศึกษานิเทศก์และครูผู้สอนประจำวิชาตามที่สำนักงานทดสอบทางการศึกษา กรมวิชาการ กระทรวงศึกษาธิการสุ่มไว้แล้วจำนวน 20 คนต่อแบบทดสอบ 1 ฉบับ

2. กลุ่มตัวอย่างนักเรียนที่ใช้ในการทำแบบทดสอบจำนวน 1,577 คน ได้จากตารางกำหนดขนาดกลุ่มตัวอย่างของฮินเคิล โอลิเวอร์ และฮินเคิล (Hinkle, Oliver and Hinkle. 1985 : 275) เมื่อให้ α เท่ากับ 0.05 อำนาจในการทดสอบทางสถิติ (Power of the Statistical Test) เท่ากับ 0.99 และขนาดอิทธิพล (Effect Size) เท่ากับ 0.10 ทั้งนี้เพื่อให้ได้กลุ่มตัวอย่างขนาดใหญ่ประมาณ 1,000 คน (Wood and Others. 1976 : 5) สำหรับกลุ่มตัวอย่างดังกล่าวผู้วิจัยเลือกจากโครงการตรวจสอบคุณภาพการศึกษา ระดับชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2533 จำนวน 112,039 คน การเลือกกลุ่มตัวอย่างใช้วิธีการสุ่มอย่างง่าย ดังนี้

2.1 สุ่มจังหวัดจากทุกเขตการศึกษา เขตการศึกษาละ 1 - 2 จังหวัด

2.2 ในแต่ละจังหวัดสุ่มโรงเรียน จังหวัดละ 5 - 25 โรง

จากกลุ่มตัวอย่างโรงเรียนดังกล่าว ผู้วิจัยใช้วิธีการสุ่มแบบเฉพาะเจาะจง (Purposive Sampling) โดยเลือกเฉพาะนักเรียนที่ได้ระดับผลการเรียน 1 และระดับผลการเรียน 3 จากการสอบปลายภาคเรียนที่ 1 ใน 4 กลุ่มอันได้แก่ กลุ่มทักษะ - ภาษาไทย กลุ่มทักษะ - คณิตศาสตร์ กลุ่มสร้างเสริมประสบการณ์ชีวิต และกลุ่มสร้างเสริมลักษณะนิสัย เนื่องจากในการวิจัยครั้งนี้ ผู้วิจัยศึกษากลุ่มนักเรียนที่อยู่ในระดับคาบเส้น 2 กลุ่ม คือ กลุ่มนักเรียนที่อยู่ในระดับคาบเส้นที่มีความสามารถในระดับพอใช้ และกลุ่มนักเรียนที่อยู่ในระดับคาบเส้นที่มีความสามารถในระดับดี ดังนั้น กลุ่มตัวอย่างนักเรียนคาบเส้นประกอบด้วย

1. กลุ่มตัวอย่างนักเรียนคาบเส้นที่มีความสามารถในระดับพอใช้ ได้แก่ นักเรียนที่ได้ระดับผลการเรียน 1 จากการสอบปลายภาคเรียนที่ 1 ในกลุ่มต่อไปนี้ คือ

1.1 กลุ่มทักษะ - ภาษาไทย

1.2 กลุ่มทักษะ - คณิตศาสตร์

1.3 กลุ่มสร้างเสริมประสบการณ์ชีวิต

1.4 กลุ่มสร้างเสริมลักษณะนิสัย

2. กลุ่มตัวอย่างนักเรียนคาบเส้นที่มีความสามารถในระดับดี ได้แก่ นักเรียนที่ได้ระดับผลการเรียน 3 จากการสอบปลายภาคเรียนที่ 1 ในกลุ่มต่อไปนี้ คือ

2.1 กลุ่มทักษะ - ภาษาไทย

2.2 กลุ่มทักษะ - คณิตศาสตร์

2.3 กลุ่มสร้างเสริมประสบการณ์ชีวิต

2.4 กลุ่มสร้างเสริมลักษณะนิสัย

สำหรับรายละเอียดของเขตการศึกษา และรายชื่อโรงเรียนที่สุ่มมาเป็นกลุ่มตัวอย่าง ปรากฏในภาคผนวก ก

เครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูล

เครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูลมีดังนี้

1. แบบทดสอบที่ใช้วัดความสามารถทั่วไปมี 3 ฉบับ แบบทดสอบวัดการคิดและแก้ปัญหามี 1 ฉบับ และแบบทดสอบวัดคุณลักษณะทั่วไปมี 1 ฉบับ ลักษณะเป็นแบบทดสอบแบบเลือกตอบชนิด 4 ตัวเลือก ในแต่ละฉบับมีรายละเอียดดังนี้

1.1 แบบทดสอบวัดความสามารถทั่วไปฉบับที่ 1 เป็นแบบทดสอบที่มีจุดประสงค์ในการวัดการใช้ภาษาเพื่อการสื่อสารโดยเน้นสมรรถภาพที่สำคัญคือ การมีความรู้พื้นฐานในการใช้ภาษาเพื่อการสื่อสาร ประกอบด้วยข้อคำถามจำนวน 30 ข้อ ใช้เวลาทำ 30 นาที

1.2 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 เป็นแบบทดสอบที่มีจุดประสงค์ในการวัดการคิดคำนวณโดยเน้นสมรรถภาพที่สำคัญคือ การมีความรู้พื้นฐานและทักษะในการคิดคำนวณ ความสามารถในการนำไปใช้ในชีวิตประจำวัน ประกอบด้วยข้อคำถามจำนวน 30 ข้อ ใช้เวลาทำ 50 นาที

1.3 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 เป็นแบบทดสอบที่มีจุดประสงค์ในการวัดการเข้าใจธรรมชาติและสิ่งแวดล้อมโดยเน้นสมรรถภาพที่สำคัญคือ การมีความเข้าใจในธรรมชาติและสิ่งแวดล้อม ประกอบด้วยข้อคำถามจำนวน 30 ข้อ ใช้เวลาทำ 40 นาที

1.4 แบบทดสอบวัดการคิดและแก้ปัญหา เป็นแบบทดสอบที่มีจุดประสงค์ในการวัดความสามารถในการคิดและแก้ปัญหาโดยเน้นสมรรถภาพที่สำคัญคือ เสนอแนวทางในการแก้ปัญหาของตนเองและครอบครัวได้ ประกอบด้วยข้อคำถามจำนวน 30 ข้อ ใช้เวลาทำ 50 นาที

1.5 แบบทดสอบวัดคุณลักษณะทั่วไป เป็นแบบทดสอบที่มีจุดประสงค์ในการวัดค่านิยมในการเสียสละ การมุ่งพัฒนาและการรักษาสุขภาพกายใจโดยเน้นสมรรถภาพที่สำคัญคือ การมีนิสัยไม่เอาเปรียบผู้อื่น การมีนิสัยรักการอ่านและแสวงหาความรู้อยู่เสมอ และมีสุขนิสัยที่ดีในการบำรุงสุขภาพอนามัยของตน ประกอบด้วยข้อคำถามจำนวน 30 ข้อ ใช้เวลาทำ 40 นาที

2. แบบกรอกรายชื่อและผลการเรียนของนักเรียน เป็นเครื่องมือที่ผู้วิจัยสร้างขึ้นเพื่อให้ครูผู้สอนกรอกรายชื่อและคะแนนหรือระดับผลการเรียนของนักเรียน ชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2533 จากการสอบปลายภาคเรียนที่ 1 และการสอบปลายปีใน 4 กลุ่ม อันได้แก่

2.1 กลุ่มทักษะ - ภาษาไทย

2.2 กลุ่มทักษะ - คณิตศาสตร์

2.3 กลุ่มสร้างเสริมประสบการณ์ชีวิต

2.4 กลุ่มสร้างเสริมลักษณะนิสัย

คุณภาพของเครื่องมือ

เนื่องจากแบบทดสอบที่ใช้วัดความสามารถทั่วไป การคิดและแก้ปัญหา และคุณลักษณะทั่วไป ของโครงการตรวจสอบคุณภาพการศึกษา นักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2533 แต่ละฉบับสร้างตามสมรรถภาพโดยครูผู้สอน ศึกษานิเทศก์ นักวัดผลการศึกษา และนักวิชาการสอบ จึงถือได้ว่าเป็นแบบทดสอบที่มีความเที่ยงตรงเชิงเนื้อหา และความเที่ยงตรงเชิงโครงสร้างนอกจากนี้ยังได้นำไปทดลองใช้กับนักเรียนทั่วประเทศ ทั้งในส่วนกลางและส่วนภูมิภาคเพื่อหาค่าความยากของข้อสอบเป็นรายข้อ และค่าความเชื่อมั่นของแบบทดสอบทั้งฉบับซึ่งค่าสถิติของแบบทดสอบแต่ละฉบับปรากฏผลดังตาราง 6

ตาราง 6 คุณภาพของแบบทดสอบแต่ละฉบับที่นำมาใช้เป็นเครื่องมือ

แบบทดสอบ	พิสัยของค่าความยาก	ค่าความเชื่อมั่น
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1	0.1817 - 0.8499	0.6764
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2	0.1962 - 0.8551	0.7875
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3	0.1724 - 0.8693	0.6990
แบบทดสอบวัดการคิดและแก้ปัญหา	0.1149 - 0.9268	0.4595
แบบทดสอบวัดคุณลักษณะทั่วไป	-	0.6784

การเก็บรวบรวมข้อมูล

การเก็บรวบรวมข้อมูล ผู้วิจัยดำเนินการดังนี้

1. ข้อมูลที่ได้จากผู้เชี่ยวชาญซึ่งได้แก่ การกำหนดเกณฑ์มาตรฐานโดยใช้เทคนิคของนีเดลสกี ผู้วิจัยเก็บรวบรวมข้อมูลร่วมกับสำนักงานทดสอบทางการศึกษา กรมวิชาการ กระทรวงศึกษาธิการ ดังนี้

1.1 จัดเตรียมแบบทดสอบพร้อมกับเฉลยให้แก่ผู้เชี่ยวชาญพิจารณากำหนดเกณฑ์มาตรฐาน (Criteria : C) สำหรับการกำหนดเกณฑ์มาตรฐานนั้นแบ่งออกเป็น 2 เกณฑ์ คือ

1.1.1 เกณฑ์มาตรฐานขั้นต่ำ (C_1) เป็นเกณฑ์มาตรฐานที่แบ่งนักเรียนระหว่างกลุ่มที่มีความสามารถระดับที่ต้องปรับปรุงแก้ไขกับนักเรียนกลุ่มที่มีความสามารถระดับพอใช้หรือระดับปานกลาง

1.1.2 เกณฑ์มาตรฐานขั้นสูง (C_2) เป็นเกณฑ์มาตรฐานที่แบ่งนักเรียนระหว่างกลุ่มที่มีความสามารถระดับพอใช้หรือระดับปานกลางกับนักเรียนกลุ่มที่มีความสามารถระดับดี

1.2 ให้ผู้เชี่ยวชาญพิจารณาตัวเลือกที่เป็นตัวลวงในแบบทดสอบแต่ละข้อว่านักเรียนที่มีความสามารถต่ำสุด ซึ่งในที่นี้หมายถึง นักเรียนที่มีความสามารถระดับพอใช้ สามารถรู้ได้ว่าตัวเลือกใดเป็นตัวเลือกที่ผิดก็ให้เขียนตัวอักษร "ร" (หมายถึง "รู้") หรือขีดเครื่องหมายกากบาท ลงบนตัวเลือกนั้น แต่ถ้าตัวเลือกใดที่ผู้เชี่ยวชาญไม่แน่ใจว่านักเรียนที่มีความสามารถระดับพอใช้จะรู้ว่าเป็นตัวเลือกที่ผิด ผู้เชี่ยวชาญก็ไม่ต้องเขียนตัวอักษร "ร" หรือขีดเครื่องหมายกากบาทลงบนตัวเลือกนั้นแต่ให้ปล่อยว่างไว้ ในการพิจารณาตัวเลือกที่เป็นตัวลวงในแบบทดสอบจะดำเนินการเช่นนี้เรื่อยไปจนครบทุกข้อคำถามดังตัวอย่าง

(0) ถ้าส่วนเบี่ยงเบนมาตรฐานของข้อมูลชุดหนึ่งมีค่าเท่ากับ 9 ความแปรปรวนมีค่าเท่าไร

ก. 3
ข. 8
ค. 72
ง. 81

ก.
ข.
ค.
C
ง.

จากตัวอย่างคำตอบที่ถูกต้องคือ ตัวเลือก ง. ดังนั้นในแบบทดสอบที่ให้ผู้เชี่ยวชาญพิจารณานั้นจะเขียนตัวอักษร "C" ซึ่งหมายถึงเป็นคำตอบที่ถูกต้อง (Corect) ไว้ให้ ส่วนตัวเลือกที่เหลือผู้เชี่ยวชาญจะต้องพิจารณาว่า นักเรียนที่มีความสามารถระดับพอใช้จะคิดว่าตัวเลือกใดเป็นตัวเลือกที่ผิด สมมุติว่าเมื่ออ่านตัวเลือก ก. แล้วผู้เชี่ยวชาญคิดว่านักเรียนที่มีความสามารถระดับพอใช้ คิดว่าตัวเลือก ก. เป็นตัวเลือกที่ผิด ผู้เชี่ยวชาญก็จะต้องขีดเครื่องหมายกากบาทลงไปบนตัวเลือก ก. ส่วนตัวเลือก ข. กับ ค. ถ้าผู้เชี่ยวชาญไม่แน่ใจว่านักเรียนที่อยู่ในระดับคาบเส้นจะรู้ว่าเป็นตัวเลือกที่ผิดหรือไม่ ผู้เชี่ยวชาญก็ไม่ต้องขีดเครื่องหมายกากบาทลงไปบนตัวเลือก ข. กับ ค. แต่ให้ปล่อยว่างไว้ ดังนั้นในการพิจารณาคำถามข้อนี้ ผลการพิจารณาจะได้ดังนี้

(0) ถ้าส่วนเบี่ยงเบนมาตรฐานของข้อมูลชุดหนึ่งมีค่าเท่ากับ 9 ความแปรปรวนมีค่าเท่าไร

ก.	3	ก.	X
ข.	8	ข.	
ค.	72	ค.	
ง.	81	ง.	C

1.3 ให้ผู้เชี่ยวชาญแต่ละคนหาโอกาสของการตอบถูกในแต่ละข้อคำถาม ดังนี้

$$\text{โอกาสของการตอบถูก} = \frac{1}{\text{จำนวนตัวเลือกทั้งหมด} - \text{จำนวนตัวเลือกที่ตัดทิ้ง}}$$

$$\text{หรือ} \quad = \left[q_i - k_i(c) \right]^{-1}$$

(Van der Linder. 1982 : 297)

ดังนั้นโอกาสของการตอบถูกของข้อคำถามข้างต้น = $\frac{1}{4-1}$ = .33

1.4 นำค่าโอกาสของการตอบถูกของแต่ละข้อคำถามของผู้เชี่ยวชาญแต่ละคนมาหาค่าเฉลี่ย (Mean) มัธยฐาน (Median) ฐานนิยม (Mode) หรือค่าเฉลี่ยปรับ (Trimmed Mean) แล้วแต่ความเหมาะสม

1.5 นำค่าเฉลี่ย มัธยฐาน ฐานนิยม หรือค่าเฉลี่ยปรับของทุกข้อคำถามมารวมกันแล้วหารด้วยจำนวนข้อคำถามทั้งหมดก็จะได้เกณฑ์มาตรฐานและเกณฑ์มาตรฐานดังกล่าวเป็นเกณฑ์มาตรฐานขั้นต่ำ (C_1) ซึ่งเป็นเกณฑ์ในการพิจารณาแบ่งนักเรียนระหว่างกลุ่มที่มีความสามารถระดับที่ต้องปรับปรุงหรือระดับปานกลางกับนักเรียนกลุ่มที่มีความสามารถระดับดี

ส่วนการหาเกณฑ์มาตรฐานขั้นสูง (C_2) ก็ดำเนินการเช่นเดียวกับการหาเกณฑ์มาตรฐานขั้นต่ำ (C_1) เพียงแต่ผู้เชี่ยวชาญจะต้องพิจารณาว่านักเรียนที่มีความสามารถระดับดีรู้ว่าตัวเลือกใดในแต่ละข้อคำถามเป็นตัวเลือกที่ผิดก็ให้เขียนตัวอักษร "ร" หรือขีดเครื่องหมายกากบาทลงบนตัวเลือกตัวนั้น ต่อจากนั้นก็ดำเนินการเช่นเดียวกับการหาเกณฑ์มาตรฐานขั้นต่ำ (C_1) คือดำเนินการตามข้อ 1.3 ถึงข้อ 1.5

ในการวิจัยครั้งนี้แบบทดสอบที่นำมาหาเกณฑ์มาตรฐานโดยใช้เทคนิคของนีเดลสกีมี 4 ฉบับ คือแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 แบบทดสอบวัดการคิดและแก้ปัญหา เกณฑ์มาตรฐานของแบบทดสอบแต่ละฉบับซึ่งได้จากการคำนวณในข้อ 1.5 ปรากฏผลดังตาราง 7

ตาราง 7 เกณฑ์มาตรฐานของแบบทดสอบแต่ละฉบับตามที่ผู้เชี่ยวชาญกำหนดโดยใช้เทคนิคของนีเดลสกี

แบบทดสอบ	คะแนนเต็ม	เกณฑ์มาตรฐาน	
		ขั้นต่ำ (C_1)	ขั้นสูง (C_2)
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1	30	12	17
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2	30	10	18
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3	30	12	20
แบบทดสอบวัดการคิดและแก้ปัญหา	30	10	17

สำหรับแบบทดสอบวัดคุณลักษณะทั่วไปเป็นแบบทดสอบที่ให้คะแนนตามระดับความรู้สึกเป็น 4 ระดับ เกณฑ์การประเมินแบ่งช่วงคะแนนเป็น 5 ช่วง ดังนี้

102 - 120 หมายถึง ดีมาก

84 - 101 หมายถึง ดี

66 - 83 หมายถึง ปานกลาง

48 - 65 หมายถึง ต่ำ

30 - 47 หมายถึง ต่ำมาก

2. ข้อมูลของนักเรียนที่มีความสามารถในการเรียนอยู่ในระดับคาบเส้นและข้อมูลผลสัมฤทธิ์ทางการเรียนของนักเรียน ผู้วิจัยเก็บรวบรวมข้อมูลโดยการทำหนังสือขอความร่วมมือในการวิจัยจากบัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร ไปติดต่อกับสำนักงานคณะกรรมการการประถมศึกษาแห่งชาติ สำนักงานคณะกรรมการการศึกษาเอกชน และสำนักการศึกษา เพื่อขอหนังสือราชการอำนวยความสะดวกในการขอความร่วมมือจากโรงเรียนที่เป็น

กลุ่มตัวอย่างพร้อมกับส่งหนังสือขอความร่วมมือในการวิจัยของผู้วิจัย แบบกรอกรายชื่อและผลการเรียนของนักเรียน พร้อมกับจัดซองติดแสตมป์จ่าหน้าซองถึงผู้วิจัยแนบไปโดยส่งทางไปรษณีย์ให้ผู้บริหารโรงเรียน ผู้อำนวยการโรงเรียนหรืออาจารย์ใหญ่เพื่อกรอกรายชื่อนักเรียนและผลการเรียนของนักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2533 ในการสอบปลายภาคเรียนที่ 1 และการสอบปลายปีแล้วส่งคืนผู้วิจัยทางไปรษณีย์ในระยะเวลาที่กำหนด ต่อจากนั้นผู้วิจัยจึงสุ่มรายชื่อนักเรียนแบบเฉพาะเจาะจงจากแบบบันทึกรายชื่อ และผลการเรียนของนักเรียนเฉพาะนักเรียนที่ได้ระดับผลการเรียน 1 และระดับผลการเรียน 3 จากการสอบปลายภาคเรียนที่ 1 นักเรียนที่ได้ระดับผลการเรียน 1 ถือว่าเป็นนักเรียนซึ่งอยู่ในระดับคาบเส้นที่มีผลการเรียนในระดับพอใช้ ส่วนนักเรียนที่ได้ระดับผลการเรียน 3 ถือว่าเป็นนักเรียนซึ่งอยู่ในระดับคาบเส้นที่มีผลการเรียนในระดับดี ต่อจากนั้นผู้วิจัยจึงนำรายชื่อของนักเรียนที่ได้ระดับผลการเรียน 1 และระดับผลการเรียน 3 ในแต่ละกลุ่มไปหากระดาษคำตอบจากการสอบของโครงการตรวจสอบคุณภาพการศึกษาหรือโครงการติดตามประเมินการใช้หลักสูตรของสถานศึกษา ปีการศึกษา 2533 ดังนี้

2.1 นำรายชื่อของนักเรียนที่ได้ระดับผลการเรียน 1 และระดับผลการเรียน 3 ในกลุ่มทักษะ - ภาษาไทย ไปหากระดาษคำตอบจากการสอบในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1

2.2 นำรายชื่อของนักเรียนที่ได้ระดับผลการเรียน 1 และระดับผลการเรียน 3 ในกลุ่มทักษะ - คณิตศาสตร์ ไปหากระดาษคำตอบจากการสอบในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2

2.3 นำรายชื่อของนักเรียนที่ได้ระดับผลการเรียน 1 และระดับผลการเรียน 3 ในกลุ่มสร้างเสริมประสบการณ์ชีวิตไปหากระดาษคำตอบจากการสอบในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 และแบบทดสอบวัดการคิดและแก้ปัญหา

2.4 นำรายชื่อของนักเรียนที่ได้ระดับผลการเรียน 1 และระดับผลการเรียน 3 ในกลุ่มสร้างเสริมลักษณะนิสัยไปหากระดาษคำตอบจากการสอบในแบบทดสอบวัดคุณลักษณะทั่วไป

จากรายชื่อของนักเรียนที่ได้ระดับผลการเรียน 1 และระดับผลการเรียน 3 ในแต่ละกลุ่ม เมื่อนำไปหากระดาษคำตอบจากแบบทดสอบแต่ละฉบับ ปรากฏผลดังตาราง 8

ตาราง 8 จำนวนนักเรียนที่ได้ระดับผลการเรียน 1 และระดับผลการเรียน 3 จากการสอบปลายภาคเรียนที่ 1 จำแนกตามแบบทดสอบแต่ละฉบับ

แบบทดสอบ	จำนวนนักเรียน	
	กลุ่มระดับผลการเรียน 1	กลุ่มระดับผลการเรียน 3
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1	1359	2300
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2	2149	1619
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3	1240	2520
แบบทดสอบวัดการคิดและแก้ปัญหา	1240	2520
แบบทดสอบวัดคุณลักษณะทั่วไป	322	1680

อนึ่ง ข้อมูลที่ได้จากการสอบของนักเรียน เนื่องจากกรมวิชาการได้ขอความร่วมมือจากหน่วยงานระดับจังหวัดเพื่อดำเนินการเกี่ยวกับการสอบ การจัดเก็บแบบทดสอบและการตรวจให้คะแนนตลอดจนบันทึกคะแนนเฉพาะในส่วนที่เป็นคะแนนรวมของนักเรียนเป็นรายบุคคลจำแนกตามวิชา แล้วจัดส่งกระดาษคำตอบพร้อมแบบบันทึกผลให้กรมวิชาการ ดังนั้นในการเก็บรวบรวมข้อมูล ผู้วิจัยจึงขออนุญาตใช้กระดาษคำตอบของนักเรียนชั้นประถมศึกษาปีที่ 6 จากสำนักงานทดสอบทางการศึกษา กรมวิชาการ กระทรวงศึกษาธิการ เพื่อนำผลการตอบข้อสอบรายข้อของผู้สอบเป็นรายบุคคลในแบบทดสอบ 4 ฉบับ อันได้แก่ แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 แบบทดสอบวัดการคิดและแก้ปัญหาไปวิเคราะห์โดยกำหนดรหัส ดังนี้ (1) หมายถึง การตอบถูกต้อง (0) หมายถึง การตอบไม่ถูกต้อง และ (9) หมายถึง การไม่ตอบหรือเว้นว่างไว้ ส่วนแบบทดสอบวัดคุณลักษณะทั่วไป ผู้วิจัยจะใช้เฉพาะคะแนนรวมของนักเรียนเป็นรายบุคคล

สำหรับข้อมูลผลสัมฤทธิ์ทางการเรียน ผู้วิจัยใช้ระดับผลการเรียนของนักเรียนจากการสอบปลายปีใน 4 กลุ่ม คือ ทักษะ - ภาษาไทย ทักษะ - คณิตศาสตร์ สร้างเสริมประสบการณ์ชีวิต และสร้างเสริมลักษณะนิสัยโดยศึกษาเฉพาะกลุ่มนักเรียนที่ได้ระดับผลการเรียน 1 และระดับผลการเรียน 3 จากการสอบปลายภาคเรียนที่ 1

<u>การวิเคราะห์ข้อมูล</u>

การวิเคราะห์ข้อมูลผู้วิจัยดำเนินการดังนี้

1. หาเกณฑ์มาตรฐานของแบบทดสอบแต่ละฉบับตามเทคนิคของนีเดลสกี ดังนี้

1.1 หาความน่าจะเป็นที่ผู้สอบที่มีความสามารถในระดับพอใช้ และในระดับดีที่จะตอบแต่ละข้อคำถามได้ถูกจากสูตร (Van der Linden. 1982 : 297)

$$P_i(+/\theta_c) = \left[q_i - k_i^{(c)} \right]^{-1}$$

เมื่อ $P_i(+/\theta_c)$ คือ ความน่าจะเป็นที่ผู้สอบที่มีความสามารถระดับพอใช้และระดับดีที่จะตอบแต่ละข้อคำถามได้ถูก

q_i คือ จำนวนของตัวเลือกข้อที่ i

$k_i^{(c)}$ คือ จำนวนของตัวเลือกซึ่งผู้เชี่ยวชาญตัดสินว่าผู้สอบที่มีความสามารถระดับพอใช้ และระดับดีสามารถรู้ว่าเป็นตัวเลือกที่ผิด

1.2 หาเกณฑ์มาตรฐานของแต่ละข้อคำถามของแบบทดสอบจากการตัดสินโดยผู้เชี่ยวชาญทั้งหมดโดยการนำค่าในข้อ 1.1 จากการตัดสินโดยผู้เชี่ยวชาญแต่ละคนมาหาค่าเฉลี่ยมัธยฐาน ฐานนิยม หรือค่าเฉลี่ยปรับแล้วแต่ความเหมาะสม

1.3 หาเกณฑ์มาตรฐานของแบบทดสอบทั้งฉบับโดยการนำค่าในข้อ 1.2 มาหาค่าเฉลี่ย ก็จะได้เกณฑ์มาตรฐานโดยใช้เทคนิคของนีเดลสกีในกลุ่มผู้สอบที่มีความสามารถระดับพอใช้ ซึ่งถือว่าเป็นเกณฑ์มาตรฐานขั้นต่ำ (C_1) และกลุ่มผู้สอบที่มีความสามารถระดับดี ซึ่งถือว่าเป็นเกณฑ์มาตรฐานขั้นสูง (C_2)

2. นำคะแนนที่ได้จากแบบทดสอบแต่ละฉบับไปวิเคราะห์โดยใช้แบบจำลอง Three-Parameter Logistic เพื่อประมาณค่าความสามารถของผู้สอบ และค่าพารามิเตอร์ของข้อสอบ อันได้แก่ ค่าความยาก ค่าอำนาจจำแนกและค่าการเดาด้วยเครื่องคอมพิวเตอร์โปรแกรมไลจิส (LOGIST)

3. หาคะแนนจุดตัดตามวิธีการของ แวน เดอร์ ลินเดน โดยการแปลงเกณฑ์มาตรฐาน ในสเกลของคะแนนโดเมนให้เป็นระดับความสามารถโดยอาศัยโค้งคุณลักษณะแบบทดสอบ (Test Characteristic Curve) ที่ได้จากค่าเฉลี่ยของโค้งคุณลักษณะข้อสอบทุกข้อในแบบทดสอบจากสูตร (Hambleton and Swaminathan. 1985 : 66)

$$\pi_c / \theta^* = \frac{1}{n} \sum_{i=1}^{n} P_i(\theta^*)$$

โดยที่ $P_i(\theta^*) = c_i + (1 - c_i) \dfrac{e^{Da_i(\theta - b_i)}}{1 + e^{Da_i(\theta - b_i)}}$

เมื่อ $P_i(\theta^*)$ คือ โอกาสในการตอบข้อสอบข้อที่ i ถูกของคนที่มีความสามารถในระดับ θ^*

n คือ จำนวนข้อสอบ

b_i คือ ค่าพารามิเตอร์ของความยาก

a_i คือ ค่าพารามิเตอร์ของอำนาจจำแนก

D คือ ค่าเท่ากับ 1.7

c_i คือ ค่าพารามิเตอร์ของการเดา

θ^* คือ ระดับความสามารถที่เป็นเกณฑ์มาตรฐานแบ่งเป็น

1) เกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี

θ^* คือ ค่าความสามารถต่ำสุดของผู้สอบโดยเฉลี่ยที่จะตอบแต่ละข้อคำถามได้ถูกจากการตัดสินโดยผู้เชี่ยวชาญ k คน

2) เกณฑ์มาตรฐานตามวิธีการใช้กลุ่มคาบเส้น

θ^* คือ ระดับความสามารถของผู้สอบโดยเฉลี่ย

4. หาคะแนนจุดตัดตามวิธีการของ เคน จากสูตร (Kane. 1987 : 333 - 345)

$$\hat{\theta}^* = \frac{1}{\sum\left[P_i^2(\theta^*_{iR}) / \sigma_i^2(M_{iR})\right]} \sum_i \frac{P_i^2(\theta^*_{iR})}{\sigma_i^2(M_{iR})} \cdot \hat{\theta}^*_{iR}$$

โดย $P_i^2(\hat{\theta}^*_{iR}) = D^2 a_i^2 Q_i^2(\hat{\theta}^*_{iR}) \left[P_i(\hat{\theta}^*_{iR}) - c_i\right]^2 / (1 - c_i)^2$

$$Q_i^2(\hat{\theta}^*_{iR}) = \left[1 - P_i(\hat{\theta}^*_{iR})\right]^2$$

$$P_i(\hat{\theta}^*_{iR}) = c_i + (1 - c_i) \frac{e^{Da_i(\hat{\theta}^*_{iR} - b_i)}}{1 + e^{Da_i(\hat{\theta}^*_{iR} - b_i)}}$$

เมื่อ $\hat{\theta}^*_{iR}$ คือ ระดับความสามารถที่เป็นเกณฑ์มาตรฐาน แบ่งเป็น

1) เกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี

$\hat{\theta}^*_{iR}$ คือ ค่าความสามารถต่ำสุดของผู้สอบโดยเฉลี่ยที่จะตอบแต่ละข้อคำถามได้ถูกจากการตัดสินโดยผู้เชี่ยวชาญ k คน

นั่นคือ $$M_{iR} = \frac{\sum M_{ir}}{k}$$

และ $$\sigma_i^2(M_{ir}) = \frac{\sum_r (M_{ir} - M_{iR})^2}{k - 1}$$

$$\sigma_i^2(M_{iR}) = \sigma_i^2(M_{ir}) / k$$

2) เกณฑ์มาตรฐานตามวิธีการใช้กลุ่มคาบเส้น

$\hat{\theta}^*_{iR}$ คือ ระดับความสามารถของผู้สอบแต่ละคน

ดังนั้น $$\sigma_i^2(M_{iR}) = \sigma_\theta^2 = \frac{\sum(\theta - \bar{\theta})^2}{n - 1}$$

อนึ่ง ในการหาคะแนนจุดตัดตามวิธีการของเคน โดยกำหนดเกณฑ์มาตรฐานตามวิธีการใช้กลุ่มคาบเส้นนั้น ได้นำจำนวนผู้เข้าสอบไปหารอีกครั้งหนึ่ง เนื่องจาก $\hat{\theta}^*_{iR}$ เป็นการประมาณจากระดับความสามารถของผู้สอบแต่ละคน

5. แปลงคะแนนจุดตัดในสเกลของความสามารถให้เป็นคะแนนโดเมนโดยอาศัยโค้งคุณลักษณะแบบทดสอบหรือโดยใช้สูตร

$$\text{คะแนนจุดตัดในสเกลของคะแนนโดเมน} = \sum_{i=1}^{m} P_i(\theta)$$

6. หาความเชื่อมั่นของการจำแนกผู้สอบเป็นผู้รอบรู้ - ไม่รอบรู้โดยใช้วิธีการของ ฮวิน (Huynh. 1976 : 258 - 259) ดังนี้

6.1 คำนวณค่า μ , σ^2 จากคะแนนสอบที่ได้

6.2 คำนวนค่า α_{21}, α ดังนี้

$$\alpha_{21} = \frac{n}{n-1}\left[1 - \frac{\mu(n-\mu)}{n\sigma^2}\right]$$

$$\alpha = \left[-1 + 1/\alpha_{21}\right]\mu$$

6.3 คำนวณค่า μ_x', σ_x' และ ρ ดังนี้

$$\mu_x' = \sin^{-1}\sqrt{\mu/n}$$

$$\sigma_x' = \left[(\alpha_{21} + 1) / (\alpha + n)\right]^{1/2}$$

$$\rho = \left[\alpha_{21} \ (n-1)/n \ \alpha_{21}\right]^{1/2}$$

6.4 คำนวณอัตราส่วน $Z = (c' - \mu_{x'})/\sigma_{x'}$ เมื่อ c' คือคะแนนจุดตัด

6.5 ประมาณค่า P_o โดยใช้ค่าความน่าจะเป็นของการกระจายแบบปกติมาตรฐานที่น้อยกว่า Z ในการคำนวณ

6.6 ประมาณค่า P_{oo} โดยการเปิดตาราง ซึ่งเป็นค่าความน่าจะเป็นของตัวแปรปกติ มาตรฐานสองค่าที่จะมีค่าสหสัมพันธ์ ρ น้อยกว่า Z

6.7 หาค่า K จากสูตร

$$K = \left[P_{oo} - P_o \right]^2 / \left[P_o - P_o^2 \right]$$

7. เปรียบเทียบความแตกต่างของจำนวนผู้รอบรู้จากการสอบด้วยแบบทดสอบแต่ละฉบับโดยการใช้คะแนนจุดตัดแต่ละวิธีด้วยการทดสอบค่าสัดส่วน (Downie and Heath. 1970 : 189)

8. หาความเที่ยงตรงหรือความสอดคล้องของคะแนนจากแบบทดสอบแต่ละฉบับกับผลสัมฤทธิ์ทางการเรียนในแต่ละกลุ่มด้วยการทดสอบค่า ไคสแควร์ (Downie and Heath. 1970 : 198) และสัมประสิทธิ์คอนติงเจนซี (Contingency Coefficient) (Ferguson. 1981 : 214) ดังนี้

8.1 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 กับกลุ่มทักษะ - ภาษาไทย

8.2 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 กับกลุ่มทักษะ - คณิตศาสตร์

8.3 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 แบบทดสอบวัดการคิดและแก้ปัญหา กับกลุ่มสร้างเสริมประสบการณ์ชีวิต

8.4 แบบทดสอบวัดคุณลักษณะทั่วไป กับกลุ่มสร้างเสริมลักษณะนิสัย

บทที่ 4

ผลการวิเคราะห์ข้อมูล

ในการวิจัยครั้งนี้ ผู้วิจัยแบ่งการเสนอผลการวิเคราะห์ข้อมูล ออกเป็น 4 ตอน ดังนี้

ตอนที่ 1 ผลการเปรียบเทียบคะแนนจุดตัด

ตอนที่ 2 ผลการเปรียบเทียบจำนวนผู้รอบรู้

ตอนที่ 3 ผลการเปรียบเทียบความเชื่อมั่นในการจำแนกความรอบรู้

ตอนที่ 4 ผลการเปรียบเทียบความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้

<u>ตอนที่ 1 ผลการเปรียบเทียบคะแนนจุดตัด</u>

การเปรียบเทียบคะแนนจุดตัด ผู้วิจัยแบ่งการเสนอผลดังนี้

1. การเปรียบเทียบคะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี

2. การเปรียบเทียบคะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้น

3. การเปรียบเทียบคะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีกับวิธีการใช้กลุ่มคาบเส้น

4. การเปรียบเทียบคะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี กับวิธีการใช้กลุ่มคาบเส้น

สำหรับรายละเอียดของการเปรียบเทียบมีดังนี้

1. การเปรียบเทียบคะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี

การกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี เป็นการกำหนดเกณฑ์มาตรฐานโดยการให้ผู้เชี่ยวชาญพิจารณาว่า นักเรียนที่มีความสามารถต่ำสุดจะมีโอกาสตอบแต่ละข้อคำถามได้ถูกต้องมากน้อยเพียงใด ในการวิจัยครั้งนี้ผู้วิจัยได้แบ่งความสามารถต่ำสุดของนักเรียนออกเป็น 2 ระดับ คือ

1.1 นักเรียนที่มีความสามารถในระดับพอใช้

1.2 นักเรียนที่มีความสามารถในระดับดี

1.1 นักเรียนที่มีความสามารถในระดับพอใช้

ผลการเปรียบเทียบคะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้ แยกตามแบบทดสอบแต่ละฉบับ ปรากฏผลดังตาราง 9

ตาราง 9 เปรียบเทียบคะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของแวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี ในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้

แบบทดสอบ	คะแนนเต็ม	คะแนนจุดตัด	
		วิธีการของ แวน เดอร์ ลินเดน	วิธีการของ เคน
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1	30	20	19
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2	30	18	17
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3	30	20	19
แบบทดสอบวัดการคิดและแก้ปัญหา	30	18	18

จากตาราง 9 จะเห็นว่าเมื่อกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี ของกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้ คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน สูงกว่าวิธีการของเคนในแบบทดสอบทุกฉบับ ยกเว้นแบบทดสอบวัดการคิดและแก้ปัญหาจะได้คะแนนจุดตัดจากทั้ง 2 วิธี เท่ากัน

1.2 นักเรียนที่มีความสามารถในระดับดี

ผลการเปรียบเทียบคะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีในกลุ่มนักเรียนที่มีความสามารถในระดับดี แยกตามแบบทดสอบแต่ละฉบับ ปรากฏผลดังตาราง 10

ตาราง 10 เปรียบเทียบคะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของแวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี ในกลุ่มนักเรียนที่มีความสามารถในระดับดี

แบบทดสอบ	คะแนนเต็ม	คะแนนจุดตัด	
		วิธีการของ แวน เดอร์ ลินเดน	วิธีการของ เคน
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1	30	20	20
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2	30	18	18
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3	30	20	20
แบบทดสอบวัดการคิดและแก้ปัญหา	30	19	18

จากตาราง 10 จะเห็นว่าเมื่อกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีของกลุ่มนักเรียนที่มีความสามารถในระดับดี คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน และเคน มีค่าเท่ากันในแบบทดสอบทุกฉบับ ยกเว้นแบบทดสอบวัดการคิดและแก้ปัญหาจะได้คะแนนจุดตัดตามวิธีการของ แวน เดอร์ ลินเดน สูงกว่าวิธีการของเคน

2. การเปรียบเทียบคะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบ ตามวิธีการของ แวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้น

การกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้น เป็นการกำหนดเกณฑ์มาตรฐานโดยการพิจารณาจากผลการเรียนของนักเรียนในการสอบปลายภาคเรียนที่ 1 ดังนี้

2.1 นักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้ หมายถึง นักเรียนที่สอบปลายภาคเรียนที่ 1 ได้ระดับผลการเรียน 1 ในแต่ละกลุ่ม คือ ทักษะ - ภาษาไทย ทักษะ - คณิตศาสตร์ สร้างเสริมประสบการณ์ชีวิต และสร้างเสริมลักษณะนิสัย

2.2 นักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดี หมายถึง นักเรียนที่สอบปลายภาคเรียนที่ 1 ได้ระดับผลการเรียน 3 ในแต่ละกลุ่ม คือ ทักษะ - ภาษาไทย ทักษะ - คณิตศาสตร์ สร้างเสริมประสบการณ์ชีวิต และสร้างเสริมลักษณะนิสัย

2.1 นักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้

ผลการเปรียบเทียบคะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการ แวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นในกลุ่มนักเรียนที่อยู่ในระดับพอใช้ แยกตามแบบทดสอบแต่ละฉบับปรากฏผลดังตาราง 11

ตาราง 11 เปรียบเทียบคะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นในกลุ่มนักเรียนที่อยู่ในระดับพอใช้

แบบทดสอบ	คะแนนเต็ม	คะแนนจุดตัด	
		วิธีการของ แวน เดอร์ ลินเดน	วิธีการของ เคน
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1	30	19	14
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2	30	17	11
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3	30	19	14
แบบทดสอบวัดการคิดและแก้ปัญหา	30	18	15

จากตาราง 11 จะเห็นว่าเมื่อกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นของนักเรียนที่มีความสามารถในระดับพอใช้ คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน สูงกว่าวิธีการของเคนในแบบทดสอบทุกฉบับ

2.2 นักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดี

ผลการเปรียบเทียบคะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นในกลุ่มนักเรียนที่อยู่ในระดับดี แยกตามแบบทดสอบแต่ละฉบับปรากฏผลดังตาราง 12

ตาราง 12 เปรียบเทียบคะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นในกลุ่มนักเรียนที่อยู่ในระดับดี

แบบทดสอบ	คะแนนเต็ม	คะแนนจุดตัด	
		วิธีการของ แวน เดอร์ ลินเดน	วิธีการของ เคน
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1	30	20	19
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2	30	18	17
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3	30	20	19
แบบทดสอบวัดการคิดและแก้ปัญหา	30	18	17

จากตาราง 12 จะเห็นว่าเมื่อกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นของนักเรียนที่มีความสามารถในระดับดี คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน สูงกว่าวิธีการของเคนในแบบทดสอบทุกฉบับ

3. การเปรียบเทียบคะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีกับวิธีการใช้กลุ่มคาบเส้น

การกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี ในการวิจัยครั้งนี้ผู้วิจัยได้แบ่งความสามารถต่ำสุดของนักเรียนออกเป็น 2 ระดับ คือ นักเรียนที่มีความสามารถในระดับพอใช้กับนักเรียนที่มีความสามารถในระดับดี

ส่วนการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้น ในการวิจัยครั้งนี้ผู้วิจัยได้แบ่งนักเรียนกลุ่มคาบเส้นออกเป็น 2 กลุ่ม คือ นักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้กับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดี

ดังนั้นในการเปรียบเทียบคะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน นั้น จะเปรียบเทียบดังนี้

3.1 นักเรียนที่มีความสามารถในระดับพอใช้ ตามเทคนิคของนีเดลสกีกับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้

3.2 นักเรียนที่มีความสามารถในระดับดี ตามเทคนิคของนีเดลสกีกับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดี

3.1 นักเรียนที่มีความสามารถในระดับพอใช้ ตามเทคนิคของนีเดลสกีกับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้

ผลการเปรียบเทียบคะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีของนักเรียนที่มีความสามารถในระดับพอใช้ กับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้ แยกตามแบบทดสอบแต่ละฉบับ ปรากฏผลดังตาราง 13

ตาราง 13 เปรียบเทียบคะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีของนักเรียนที่มีความสามารถในระดับพอใช้กับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้

แบบทดสอบ	คะแนนเต็ม	คะแนนจุดตัดจากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน	
		นักเรียนที่มีความสามารถในระดับพอใช้ตามเทคนิคของนีเดลสกี	นักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1	30	20	19
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2	30	18	17
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3	30	20	19
แบบทดสอบวัดการคิดและแก้ปัญหา	30	18	18

จากตาราง 13 จะเห็นว่า คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีของนักเรียนที่มีความสามารถในระดับพอใช้ จะได้คะแนนจุดตัดสูงกว่าการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นของนักเรียนที่อยู่ในระดับพอใช้ในแบบทดสอบทุกฉบับ ยกเว้นแบบทดสอบวัดการคิดและแก้ปัญหาที่มีคะแนนจุดตัดเท่ากัน

3.2 นักเรียนที่มีความสามารถในระดับดีตามเทคนิคของนีเดลสกีกับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดี

ผลการเปรียบเทียบคะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีของนักเรียนที่มีความสามารถในระดับดีกับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดี แยกตามแบบทดสอบแต่ละฉบับปรากฏผลดังตาราง 14

ตาราง 14 เปรียบเทียบคะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีของนักเรียนที่มีความสามารถในระดับดี กับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดี

แบบทดสอบ	คะแนนเต็ม	คะแนนจุดตัดจากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน	
		นักเรียนที่มีความสามารถในระดับดีตามเทคนิคของนีเดลสกี	นักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดี
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1	30	20	20
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2	30	18	18
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3	30	20	20
แบบทดสอบวัดการคิดและแก้ปัญหา	30	19	18

จากตาราง 14 จะเห็นว่าคะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีของนักเรียนที่มีความสามารถในระดับดี จะได้คะแนนจุดตัดเท่ากับการกำหนดเกณฑ์มาตรฐานโดยวิธีการ

ใช้กลุ่มคาบเส้นของนักเรียนที่อยู่ในระดับดี ในแบบทดสอบทุกฉบับ ยกเว้นแบบทดสอบวัดการคิดและแก้ปัญหาที่มีคะแนนจุดตัดจากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสก็สูงกว่าการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้น

4. การเปรียบเทียบคะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสก็กับวิธีการใช้กลุ่มคาบเส้น

ในการเปรียบเทียบคะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคนนั้นจะเปรียบเทียบ ดังนี้

4.1 นักเรียนที่มีความสามารถในระดับพอใช้ ตามเทคนิคของนีเดลสก็ กับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้

4.2 นักเรียนที่มีความสามารถในระดับดี ตามเทคนิคของนีเดลสก็กับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดี

4.1 นักเรียนที่มีความสามารถในระดับพอใช้ตามเทคนิคของนีเดลสก็กับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้

ผลการเปรียบเทียบคะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ เคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสก็ของนักเรียนที่มีความสามารถในระดับพอใช้กับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้ แยกตามแบบทดสอบแต่ละฉบับปรากฏผลดังตาราง 15

ตาราง 15 เปรียบเทียบคะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีของนักเรียนที่มีความสามารถในระดับพอใช้กับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้

แบบทดสอบ	คะแนนเต็ม	คะแนนจุดตัดจากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ เคน	
		นักเรียนที่มีความสามารถในระดับพอใช้ตามเทคนิคของนีเดลสกี	นักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1	30	19	14
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2	30	17	11
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3	30	19	14
แบบทดสอบวัดการคิดและแก้ปัญหา	30	18	15

จากตาราง 15 จะเห็นว่า คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีของนักเรียนที่มีความสามารถในระดับพอใช้ จะได้คะแนนจุดตัดสูงกว่าการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นของนักเรียนที่อยู่ในระดับพอใช้ในแบบทดสอบทุกฉบับ

4.2 นักเรียนที่มีความสามารถในระดับดี ตามเทคนิคของนีเดลสกีกับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดี

ผลการเปรียบเทียบคะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีของนักเรียนที่มีความสามารถในระดับดีกับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดี แยกตามแบบทดสอบแต่ละฉบับ ปรากฏผลดังตาราง 16

ตาราง 16 เปรียบเทียบคะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีของนักเรียนที่มีความสามารถในระดับดี กับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดี

แบบทดสอบ	คะแนนเต็ม	คะแนนจุดตัดจากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ เคน	
		นักเรียนที่มีความสามารถในระดับดีตามเทคนิคของนีเดลสกี	นักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดี
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1	30	20	19
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2	30	18	17
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3	30	20	19
แบบทดสอบวัดการคิดและแก้ปัญหา	30	18	17

จากตาราง 16 จะเห็นว่า คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีของนักเรียนที่มีความสามารถในระดับดี จะได้คะแนนจุดตัดสูงกว่า การกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นของนักเรียนที่มีความสามารถในระดับดีในแบบทดสอบทุกฉบับ

ตอนที่ 2 ผลการเปรียบเทียบจำนวนผู้รอบรู้

การเปรียบเทียบจำนวนผู้รอบรู้ ผู้วิจัยได้แบ่งการเสนอผล ดังนี้

1. การเปรียบเทียบจำนวนผู้รอบรู้จากการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี

2. การเปรียบเทียบจำนวนผู้รอบรู้ จากการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้น

3. การเปรียบเทียบจำนวนผู้รอบรู้ จากการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีกับวิธีการใช้กลุ่มคาบเส้น

4. การเปรียบเทียบจำนวนผู้รอบรู้ จากการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ เคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีกับวิธีการใช้กลุ่มคาบเส้น

สำหรับรายละเอียดของการเปรียบเทียบมีดังนี้

1. การเปรียบเทียบจำนวนผู้รอบรู้จากการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี

การกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี ผู้วิจัยได้แบ่งความสามารถต่ำสุดของนักเรียนออกเป็น 2 ระดับ คือ

1.1 นักเรียนที่มีความสามารถในระดับพอใช้

1.2 นักเรียนที่มีความสามารถในระดับดี

1.1 นักเรียนที่มีความสามารถในระดับพอใช้

ผลการเปรียบเทียบจำนวนผู้รอบรู้จากการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี ในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้ แยกตามแบบทดสอบแต่ละฉบับ ปรากฏผลดังตาราง 17

ตาราง 17 เปรียบเทียบจำนวนผู้รอบรู้จากการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้

แบบทดสอบ	จำนวนผู้รอบรู้ในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้		Z
	วิธีการของ แวน เดอร์ ลินเดน	วิธีการของ เคน	
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1	166	252	4.5910**
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2	333	440	4.2586**
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3	196	277	4.1547**
แบบทดสอบวัดการคิดและแก้ปัญหา	303	303	-

**p<.01

จากตาราง 17 จะเห็นว่า จำนวนผู้รอบรู้ที่ได้จากการกำหนดคะแนนจุดตัดโดยวิธีการของเคนมีมากกว่าวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้ในแบบทดสอบทุกฉบับ ยกเว้นแบบทดสอบวัดการคิดและแก้ปัญหาที่มีจำนวนผู้รอบรู้เท่ากัน

แต่ถ้าพิจารณาถึงความแตกต่างของจำนวนผู้รอบรู้ในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 และแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 ปรากฏว่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

1.2 นักเรียนที่มีความสามารถในระดับดี

ผลการเปรียบเทียบจำนวนผู้รอบรู้จากการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีในกลุ่มนักเรียนที่มีความสามารถในระดับดี แยกตามแบบทดสอบแต่ละฉบับ ปรากฏผลดังตาราง 18

ตาราง 18 เปรียบเทียบจำนวนผู้รอบรู้จากการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีในกลุ่มนักเรียนที่มีความสามารถในระดับดี

แบบทดสอบ	จำนวนผู้รอบรู้ในกลุ่มนักเรียนที่มีความสามารถในระดับดี		Z
	วิธีการของ แวน เดอร์ ลินเดน	วิธีการของ เคน	
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1	1062	1062	-
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2	801	801	-
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3	1054	1054	-
แบบทดสอบวัดการคิดและแก้ปัญหา	859	1208	10.0979**

** $p<.01$

จากตาราง 18 จะเห็นว่า จำนวนผู้รอบรู้จากการกำหนดคะแนนจุดตัดโดยวิธีการของ แวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีในกลุ่มนักเรียนที่มีความสามารถในระดับดี มีจำนวนเท่ากันในแบบทดสอบทุกฉบับ ยกเว้นแบบทดสอบวัดการคิดและแก้ปัญหาที่มีจำนวนผู้รอบรู้จากการกำหนดคะแนนจุดตัดตามวิธีการของเคนมากกว่าวิธีการของ แวน เดอร์ ลินเดน

แต่ถ้าพิจารณาถึงความแตกต่างของจำนวนผู้รอบรู้ในแบบทดสอบวัดการคิดและแก้ปัญหา ปรากฏว่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

2. การเปรียบเทียบจำนวนผู้รอบรู้จากการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้น

การกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้น ผู้วิจัยได้แบ่งกลุ่มคาบเส้นออกเป็น 2 ระดับ คือ

2.1 นักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้

2.2 นักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดี

2.1 นักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้

ผลการเปรียบเทียบจำนวนผู้รอบรู้จากการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นสำหรับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้ แยกตามแบบทดสอบแต่ละฉบับ ปรากฏผลดังตาราง 19

ตาราง 19 เปรียบเทียบจำนวนผู้รอบรู้จากการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นสำหรับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้

แบบทดสอบ	จำนวนผู้รอบรู้ในกลุ่มนักเรียนคาบเส้นที่อยู่ในระดับพอใช้		Z
	วิธีการของ แวน เดอร์ ลินเดน	วิธีการของ เคน	
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1	252	855	26.3852**
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2	440	1459	35.6211**
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3	277	788	22.7916**
แบบทดสอบวัดการคิดและแก้ปัญหา	303	736	18.8368**

**p<.01

จากตาราง 19 จะเห็นว่า จำนวนผู้รอบรู้จากการกำหนดคะแนนจุดตัดตามวิธีการของเคนมีมากกว่าวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้น สำหรับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้ในแบบทดสอบทุกฉบับ

แต่ถ้าพิจารณาถึงความแตกต่างของจำนวนผู้รอบรู้ในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 และแบบทดสอบวัดการคิดและการแก้ปัญหา ปรากฏว่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

2.2 นักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดี

ผลการเปรียบเทียบจำนวนผู้รอบรู้จากการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นสำหรับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดี แยกตามแบบทดสอบแต่ละฉบับ ปรากฏผลดังตาราง 20

ตาราง 20 เปรียบเทียบจำนวนผู้รอบรู้จากการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการกลุ่มคาบเส้นสำหรับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดี

แบบทดสอบ	จำนวนผู้รอบรู้ในกลุ่มนักเรียนคาบเส้นที่อยู่ในระดับดี		Z
	วิธีการของ แวน เดอร์ ลินเดน	วิธีการของ เคน	
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1	1062	1307	7.2698**
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2	801	916	4.0606**
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3	1054	1339	8.0927**
แบบทดสอบวัดการคิดและแก้ปัญหา	1208	1576	10.5389**

**p<.01

จากตาราง 20 จะเห็นว่าจำนวนผู้รอบรู้ที่ได้จากการกำหนดคะแนนจุดตัดตามวิธีการของเคน มีมากกว่า วิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นสำหรับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดีในแบบทดสอบทุกฉบับ

แต่ถ้าพิจารณาถึงความแตกต่างของจำนวนผู้รอบรู้ในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 และแบบทดสอบวัดการคิดและแก้ปัญหา ปรากฏว่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

3. การเปรียบเทียบจำนวนผู้รอบรู้จากการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี กับวิธีการใช้กลุ่มคาบเส้น

การเปรียบเทียบจำนวนผู้รอบรู้จากการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน นั้น จะเปรียบเทียบดังนี้

3.1 นักเรียนที่มีความสามารถในระดับพอใช้ตามเทคนิคของนีเดลสกี กับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้

3.2 นักเรียนที่มีความสามารถในระดับดี ตามเทคนิคของนีเดลสกี กับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดี

3.1 นักเรียนที่มีความสามารถในระดับพอใช้ ตามเทคนิคของนีเดลสกี กับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้

ผลการเปรียบเทียบจำนวนผู้รอบรู้ที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีของนักเรียนที่มีความสามารถในระดับพอใช้กับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้แยกตามแบบทดสอบแต่ละฉบับ ปรากฏผลดังตาราง 21

ตาราง 21 เปรียบเทียบจำนวนผู้รอบรู้ที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีของนักเรียนที่มีความสามารถในระดับพอใช้กับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้

แบบทดสอบ	จำนวนผู้รอบรู้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน		Z
	นักเรียนที่มีความสามารถในระดับพอใช้ตามเทคนิคของนีเดลสกี	นักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้	
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1	166	252	4.5910**
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2	333	440	4.2586**
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3	196	277	4.1547**
แบบทดสอบวัดการคิดและแก้ปัญหา	303	303	-

** $p<.01$

จากตาราง 21 จะเห็นว่า จำนวนผู้รอบรู้จากการกำหนดคะแนนจุดตัดโดยวิธีการของ แวน เดอร์ ลินเดน เมื่อกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นมีจำนวนผู้รอบรู้มากกว่าการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีในแบบทดสอบทุกฉบับ สำหรับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้ ยกเว้นแบบทดสอบวัดการคิดและแก้ปัญหาที่มีจำนวนผู้รอบรู้เท่ากัน

แต่ถ้าพิจารณาถึงความแตกต่างของจำนวนผู้รอบรู้ในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 และแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 ปรากฏว่า แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

3.2 นักเรียนที่มีความสามารถในระดับดี ตามเทคนิคของนีเดลสกี กับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดี

ผลการเปรียบเทียบจำนวนผู้รอบรู้ที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีของนักเรียนที่มีความสามารถในระดับดี กับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดี แยกตามแบบทดสอบแต่ละฉบับ ปรากฏผลดังตาราง 22

ตาราง 22 เปรียบเทียบจำนวนผู้รอบรู้ที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีของนักเรียนที่มีความสามารถในระดับดีกับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดี

แบบทดสอบ	จำนวนผู้รอบรู้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน		Z
	นักเรียนที่มีความสามารถในระดับดี ตามเทคนิคของนีเดลสกี	นักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดี	
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1	1062	1062	-
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2	801	801	-
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3	1054	1054	-
แบบทดสอบวัดการคิดและแก้ปัญหา	859	1208	10.0979**

**P<.01

จากตาราง 22 จะเห็นว่า จำนวนผู้รอบรู้จากการกำหนดคะแนนจุดตัดโดยวิธีการของแวน เดอร์ ลินเดน เมื่อกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีกับวิธีการใช้กลุ่มคาบเส้น สำหรับนักเรียนที่มีความสามารถในระดับดี มีจำนวนผู้รอบรู้เท่ากันในแบบทดสอบทุกฉบับ ยกเว้นแบบทดสอบวัดการคิดและแก้ปัญหาจากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นมีจำนวนผู้รอบรู้มากกว่าการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี

แต่ถ้าพิจารณาถึงความแตกต่างของจำนวนผู้รอบรู้ในแบบทดสอบวัดการคิดและแก้ปัญหา ปรากฏว่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

4. การเปรียบเทียบจำนวนผู้รอบรู้จากการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีกับวิธีการใช้กลุ่มคาบเส้น

การเปรียบเทียบจำนวนผู้รอบรู้จากการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคนนั้นจะเปรียบเทียบ ดังนี้

4.1 นักเรียนที่มีความสามารถในระดับพอใช้ตามเทคนิคของนีเดลสกี กับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้

4.2 นักเรียนที่มีความสามารถในระดับดีตามเทคนิคของนีเดลสกี กับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดี

4.1 นักเรียนที่มีความสามารถในระดับพอใช้ตามเทคนิคของนีเดลสกีกับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้

ผลการเปรียบเทียบจำนวนผู้รอบรู้ที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อ ข้อสอบตามวิธีการของเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี ของนักเรียนที่มีความสามารถในระดับพอใช้กับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้ แยกตามแบบทดสอบแต่ละฉบับ ปรากฏผลดังตาราง 23

ตาราง 23 เปรียบเทียบจำนวนผู้รอบรู้ที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีของนักเรียนที่มีความสามารถในระดับพอใช้กับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้

แบบทดสอบ	จำนวนผู้รอบรู้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ เคน		Z
	นักเรียนที่มีความสามารถในระดับพอใช้ตามเทคนิคของนีเดลสกี	นักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้	
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1	252	855	26.3852**
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2	440	1459	35.6211**
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3	277	788	22.7916**
แบบทดสอบวัดการคิดและแก้ปัญหา	303	736	18.8368**

** $p < .01$

จากตาราง 23 จะเห็นว่า จำนวนผู้รอบรู้จากการกำหนดคะแนนจุดตัดโดยวิธีการของเคน เมื่อกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นมีจำนวนผู้รอบรู้มากกว่า การกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีสำหรับนักเรียนที่มีความสามารถในระดับพอใช้ในแบบทดสอบทุกฉบับ

แต่ถ้าพิจารณาถึงความแตกต่างของจำนวนผู้รอบรู้ในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 แบบทดสอบวัดการคิดและแก้ปัญหา ปรากฏว่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

4.2 นักเรียนที่มีความสามารถในระดับดี ตามเทคนิคของนีเดลสกีกับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดี

ผลการเปรียบเทียบจำนวนผู้รอบรู้ที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ เคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีของนักเรียนที่มีความสามารถในระดับดีกับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดี แยกตามแบบทดสอบแต่ละฉบับ ปรากฏผลดังตาราง 24

ตาราง 24 เปรียบเทียบจำนวนผู้รอบรู้ที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีของนักเรียนที่มีความสามารถในระดับดีกับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดี

แบบทดสอบ	จำนวนผู้รอบรู้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ เคน		Z
	นักเรียนที่มีความสามารถในระดับดีตามเทคนิคของนีเดลสกี	นักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดี	
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1	1062	1307	7.2698**
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2	801	916	4.0606**
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3	1054	1339	8.0927**
แบบทดสอบวัดการคิดและแก้ปัญหา	1208	1576	10.5389**

** $p<.01$

จากตาราง 24 จะเห็นว่า จำนวนผู้รอบรู้จากการกำหนดคะแนนจุดตัดโดยวิธีการของเคน เมื่อกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้น มีจำนวนผู้รอบรู้มากกว่าการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี สำหรับนักเรียนที่มีความสามารถในระดับดี ในแบบทดสอบทุกฉบับ

แต่ถ้าพิจารณาถึงความแตกต่างของจำนวนผู้รอบรู้ในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 แบบทดสอบวัดการคิดและแก้ปัญหา ปรากฏว่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

ตอนที่ 3 ผลการเปรียบเทียบความเชื่อมั่นในการจำแนกความรอบรู้

การเปรียบเทียบความเชื่อมั่นในการจำแนกความรอบรู้ ผู้วิจัยได้แบ่งการเสนอผล ดังนี้

1. การเปรียบเทียบความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี

2. การเปรียบเทียบความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้น

3. การเปรียบเทียบความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีกับวิธีการใช้กลุ่มคาบเส้น

4. การเปรียบเทียบความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีกับวิธีการใช้กลุ่มคาบเส้น

สำหรับรายละเอียดของการเปรียบเทียบมีดังนี้

1. การเปรียบเทียบความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี

การกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี ผู้วิจัยได้แบ่งความสามารถต่ำสุดของนักเรียน ออกเป็น 2 ระดับ คือ

1.1 นักเรียนที่มีความสามารถในระดับพอใช้

1.2 นักเรียนที่มีความสามารถในระดับดี

1.1 นักเรียนที่มีความสามารถในระดับพอใช้

ผลการเปรียบเทียบความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี ในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้แยกตามแบบทดสอบแต่ละฉบับ ปรากฏผลดังตาราง 25

ตาราง 25 เปรียบเทียบความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้ โดยการใช้คะแนน จุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้

แบบทดสอบ	ความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้ในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้	
	วิธีการของ แวน เดอร์ ลินเดน	วิธีการของ เคน
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1	0.2741	0.2461
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2	0.3823	0.4725
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3	0.4157	0.3571
แบบทดสอบวัดการคิดและแก้ปัญหา	0.2044	0.2044

จากตาราง 25 จะเห็นว่า เมื่อกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี ในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้ การกำหนดคะแนนจุดตัดตามวิธีการของ แวน เดอร์ ลินเดน มีความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้สูงกว่า วิธีการของเคนในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 และแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 แต่แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 ปรากฏว่าความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้จากการกำหนดคะแนนจุดตัดตามวิธีการของเคน สูงกว่าวิธีการของ แวน เดอร์ ลินเดน ส่วนแบบทดสอบวัดการคิดและแก้ปัญหา ไม่ว่าจะกำหนดคะแนนจุดตัดตามวิธีการของ แวน เดอร์ ลินเดน หรือเคน ก็ตามพบว่ามีความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้ได้เท่ากัน

1.2 นักเรียนที่มีความสามารถในระดับดี

ผลการเปรียบเทียบความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีในกลุ่มนักเรียนที่มีความสามารถในระดับดี แยกตามแบบทดสอบแต่ละฉบับ ปรากฏผลดังตาราง 26

ตาราง 26 เปรียบเทียบความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีในกลุ่มนักเรียนที่มีความสามารถในระดับดี

แบบทดสอบ	ความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้ในกลุ่มนักเรียนที่มีความสามารถในระดับดี	
	วิธีการของ แวน เดอร์ ลินเดน	วิธีการของ เคน
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1	0.2988	0.2988
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2	0.3848	0.3848
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3	0.2956	0.2956
แบบทดสอบวัดการคิดและแก้ปัญหา	0.2007	0.1937

จากตาราง 26 จะเห็นว่า เมื่อกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีในกลุ่มนักเรียนที่มีความสามารถในระดับดี ปรากฏว่าการกำหนดคะแนนจุดตัดตามวิธีการของ แวน เดอร์ ลินเดน และเคน มีความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้ได้เท่ากันในแบบทดสอบทุกฉบับ

ยกเว้น แบบทดสอบวัดการคิด และแก้ปัญหาพบว่าวิธีการกำหนดคะแนนจุดตัดตามวิธีการของ แวน เดอร์ ลินเดน มีความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้สูงกว่าวิธีการของเคน

2 การเปรียบเทียบความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้น

การกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้น ผู้วิจัยได้แบ่งกลุ่มคาบเส้นออกเป็น 2 ระดับ คือ

2.1 นักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้

2.2 นักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดี

2.1 นักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้

ผลการเปรียบเทียบความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นสำหรับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้ แยกตามแบบทดสอบแต่ละฉบับ ปรากฏผลดังตาราง 27

ตาราง 27 เปรียบเทียบความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นของนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้

แบบทดสอบ	ความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้ในกลุ่มนักเรียนคาบเส้นที่อยู่ในระดับพอใช้	
	วิธีการของ แวน เดอร์ ลินเดน	วิธีการของ เคน
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1	0.2461	0.3741
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2	0.4725	0.4005
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3	0.3571	0.2873
แบบทดสอบวัดการคิดและแก้ปัญหา	0.2044	0.1215

จากตาราง 27 จะเห็นว่า เมื่อกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้ ปรากฏว่าการกำหนดคะแนนจุดตัดตามวิธีการของ แวน เดอร์ ลินเดน มีความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้สูงกว่าวิธีการของเคนในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 และแบบทดสอบวัดการคิดและแก้ปัญหา แต่แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 วิธีการของเคนมีความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้สูงกว่าวิธีการของ แวน เดอร์ ลินเดน

2.2 นักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดี

ผลการเปรียบเทียบความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นสำหรับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดี แยกตามแบบทดสอบแต่ละฉบับ ปรากฏผลดังตาราง 28

ตาราง 28 เปรียบเทียบความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นของนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดี

แบบทดสอบ	ความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้ในกลุ่มนักเรียนคาบเส้นที่อยู่ในระดับดี	
	วิธีการของ แวน เดอร์ ลินเดน	วิธีการของ เคน
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1	0.2988	0.2869
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2	0.3848	0.4765
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3	0.2956	0.1386
แบบทดสอบวัดการคิดและแก้ปัญหา	0.1937	0.2158

จากตาราง 28 จะเห็นว่าเมื่อกำหนดเกณฑ์มาตรฐาน โดยวิธีการใช้กลุ่มคาบเส้นในกลุ่มนักเรียนที่มีความสามารถในระดับดี ปรากฏว่าการกำหนดคะแนนจุดตัดตามวิธีการของ แวน เดอร์ ลินเดน มีความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้สูงกว่าวิธีการของเคนในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 และแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 แต่แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 และแบบทดสอบวัดการคิดและแก้ปัญหาวิธีการของเคนมีความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้สูงกว่าวิธีการของ แวน เดอร์ ลินเดน

3. การเปรียบเทียบความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีกับวิธีการใช้กลุ่มคาบเส้น

การเปรียบเทียบความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน นั้นจะเปรียบเทียบดังนี้

3.1 นักเรียนที่มีความสามารถในระดับพอใช้ตามเทคนิคของนีเดลสกีกับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้

3.2 นักเรียนที่มีความสามารถในระดับดีตามเทคนิคของนีเดลสกีกับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดี

3.1 นักเรียนที่มีความสามารถในระดับพอใช้ตามเทคนิคของนีเดลสกีกับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้

ผลการเปรียบเทียบความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีของนักเรียนที่มีความสามารถในระดับพอใช้กับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้แยกตามแบบทดสอบแต่ละฉบับ ปรากฏผลดังตาราง 29

ตาราง 29 เปรียบเทียบความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีของนักเรียนที่มีความสามารถในระดับพอใช้กับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้

แบบทดสอบ	ความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน	
	นักเรียนที่มีความสามารถในระดับพอใช้ ตามเทคนิคของนีเดลสกี	นักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1	0.2741	0.2461
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2	0.3823	0.4725
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3	0.4157	0.3571
แบบทดสอบวัดการคิดและแก้ปัญหา	0.2044	0.2044

จากตาราง 29 จะเห็นว่า เมื่อกำหนดคะแนนจุดตัดตามวิธีการของ แวน เดอร์ ลินเดน ในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้ ปรากฏว่า การกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี มีความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้สูงกว่าการกำหนดเกณฑ์มาตรฐานโดยวิธี

การใช้กลุ่มคาบเส้นในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 และแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 แต่แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 ปรากฏว่า ความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้จากการกำหนดเกณฑ์มาตรฐาน โดยวิธีการใช้กลุ่มคาบเส้นสูงกว่าการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี ส่วนแบบทดสอบวัดการคิดและแก้ปัญหาไม่ว่าจะกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีหรือวิธีการใช้กลุ่มคาบเส้นก็ตาม พบว่ามีความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้ได้เท่ากัน

3.2 นักเรียนที่มีความสามารถในระดับดีตามเทคนิคของนีเดลสกี กับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดี

ผลการเปรียบเทียบความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีของนักเรียนที่มีความสามารถในระดับดีกับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดี แยกตามแบบทดสอบแต่ละฉบับ ปรากฏผลดังตาราง 30

ตาราง 30 เปรียบเทียบความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีของนักเรียนที่มีความสามารถในระดับดีกับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดี

แบบทดสอบ	ความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน	
	นักเรียนที่มีความสามารถในระดับดีตามเทคนิคของ นีเดลสกี	นักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดี
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1	0.2988	0.2988
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2	0.3848	0.3848
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3	0.2956	0.2956
แบบทดสอบวัดการคิดและแก้ปัญหา	0.2007	0.1937

จากตาราง 30 จะเห็นว่า เมื่อกำหนดคะแนนจุดตัดตามวิธีการของ แวน เดอร์ ลินเดน ในกลุ่มนักเรียนที่มีความสามารถในระดับดี ปรากฏว่า การกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี และวิธีการใช้กลุ่มคาบเส้น มีความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้ได้เท่ากันในแบบทดสอบทุกฉบับ ยกเว้นแบบทดสอบวัดการคิดและแก้ปัญหา พบว่า การกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีมีความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้สูงกว่าวิธีการใช้กลุ่มคาบเส้น

4. การเปรียบเทียบความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีกับวิธีการใช้กลุ่มคาบเส้น

การเปรียบเทียบความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคนนั้นจะเปรียบเทียบดังนี้

4.1 นักเรียนที่มีความสามารถในระดับพอใช้ตามเทคนิคของนีเดลสกีกับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้

4.2 นักเรียนที่มีความสามารถในระดับดีตามเทคนิคของนีเดลสกีกับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดี

4.1 นักเรียนที่มีความสามารถในระดับพอใช้ตามเทคนิคของนีเดลสกีกับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้

ผลการเปรียบเทียบความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีของนักเรียนที่มีความสามารถในระดับพอใช้ กับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้ แยกตามแบบทดสอบแต่ละฉบับ ปรากฏผลดังตาราง 31

ตาราง 31 เปรียบเทียบความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีของนักเรียนที่มีความสามารถในระดับพอใช้กับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้

แบบทดสอบ	ความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ เคน	
	นักเรียนที่มีความสามารถในระดับพอใช้ตามเทคนิคของนีเดลสกี	นักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1	0.2461	0.3741
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2	0.4725	0.4005
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3	0.3571	0.2873
แบบทดสอบวัดการคิดและแก้ปัญหา	0.2044	0.1215

จากตาราง 31 จะเห็นว่า เมื่อกำหนดคะแนนจุดตัดตามวิธีการของเคน ในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้ ปรากฏว่า การกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี มีความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้สูงกว่าการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 และแบบทดสอบวัดการคิดและแก้ปัญหา แต่แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 พบว่า การกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นมีความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้สูงกว่าการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี

4.2 นักเรียนที่มีความสามารถในระดับดีตามเทคนิคของนีเดลสกี กับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดี

ผลการเปรียบเทียบความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีของนักเรียนที่มีความสามารถในระดับดีกับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดี แยกตามแบบทดสอบแต่ละฉบับ ปรากฏผลดังตาราง 32

ตาราง 32 เปรียบเทียบความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีของนักเรียนที่มีความสามารถในระดับดีกับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดี

แบบทดสอบ	ความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ เคน	
	นักเรียนที่มีความสามารถในระดับดีตามเทคนิคของนีเดลสกี	นักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดี
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1	0.2988	0.2869
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2	0.3848	0.4765
แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3	0.2956	0.1386
แบบทดสอบวัดการคิดและแก้ปัญหา	0.1937	0.2158

จากตาราง 32 จะเห็นว่า เมื่อกำหนดคะแนนจุดตัดตามวิธีการของเคนในกลุ่มนักเรียนที่มีความสามารถในระดับดี ปรากฏว่า การกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีมีความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้สูงกว่า การกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 และแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 แต่แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 และแบบทดสอบวัดการคิดและแก้ปัญหา พบว่า การกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นมีความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้สูงกว่าการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี

ตอนที่ 4 ผลการเปรียบเทียบความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้

การเปรียบเทียบความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้นั้น ผู้วิจัยศึกษาโดยการเปรียบเทียบจำนวนนักเรียนที่สอบได้ในแต่ละระดับคะแนนจากการสอบปลายปีในแต่ละกลุ่ม กับจำนวนนักเรียนที่สอบผ่านและไม่ผ่านในแบบทดสอบแต่ละฉบับ ดังนี้

1. กลุ่มทักษะ - ภาษาไทย กับแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1

2. กลุ่มทักษะ - คณิตศาสตร์ กับแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2

3. กลุ่มสร้างเสริมประสบการณ์ชีวิต กับแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 และ แบบทดสอบวัดการคิดและแก้ปัญหา

4. กลุ่มสร้างเสริมลักษณะนิสัย กับแบบทดสอบวัดคุณลักษณะทั่วไป

สำหรับในการเสนอผลการวิจัย ผู้วิจัยได้แบ่งการเสนอผลการวิจัยดังนี้

1. การเปรียบเทียบความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน และเคนจากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี

2. การเปรียบเทียบความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน และเคนจากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้น

3. การเปรียบเทียบความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีกับวิธีการใช้กลุ่มคาบเส้น

4. การเปรียบเทียบความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ เคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี กับวิธีการใช้กลุ่มคาบเส้น

สำหรับรายละเอียดของการเปรียบเทียบมีดังนี้

1. การเปรียบเทียบความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน และเคนจากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี

การกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี ผู้วิจัยได้แบ่งความสามารถต่ำสุดของนักเรียน ออกเป็น 2 ระดับ คือ

1.1 นักเรียนที่มีความสามารถในระดับพอใช้

1.2 นักเรียนที่มีความสามารถในระดับดี

1.1 นักเรียนที่มีความสามารถในระดับพอใช้

ผลการเปรียบเทียบความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้แยกตามแบบทดสอบแต่ละฉบับและแต่ละกลุ่ม ปรากฏผลดังตาราง 33, 34 และ 35

ตาราง 33 ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ ตามวิธีการของ แวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้ในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 กับกลุ่มทักษะ - ภาษาไทย

การกำหนดคะแนนจุดตัด	ผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1	ระดับผลการเรียนในกลุ่มทักษะ - ภาษาไทย				χ^2	C
		3	2	1	รวม		
วิธีการของแวน เดอร์ ลินเดน	ผ่าน	56	73	37	166	46.9103**	0.1827
	ไม่ผ่าน	168	541	484	1193		
	รวม	224	614	521	1359		
วิธีการของเคน	ผ่าน	71	123	58	252	45.5018**	0.1800
	ไม่ผ่าน	153	491	463	1107		
	รวม	224	614	521	1359		

**P<.01

จากตาราง 33 เมื่อพิจารณาความสอดคล้องระหว่างผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 กับระดับผลการเรียนในกลุ่มทักษะ - ภาษาไทย จากการทดสอบค่าไคสแควร์ ปรากฏว่า มีนัยสำคัญทางสถิติที่ระดับ .01 แสดงว่า ผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 มีความสัมพันธ์กับระดับผลการเรียนของนักเรียนในกลุ่มทักษะ - ภาษาไทยแต่ถ้าพิจารณาค่าคอนติงเจนซีแล้วจะเห็นว่า ค่าคอนติงเจนซีจากการกำหนดคะแนนจุดตัดตามวิธีการของ แวน เดอร์ ลินเดน สูงกว่าค่าคอนติงเจนซี จากการกำหนดคะแนนจุดตัดตามวิธีการของ เคน เพียงเล็กน้อย แสดงว่าความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ ตามวิธีการของ แวน เดอร์ ลินเดน สูงกว่าวิธีการของเคนเพียงเล็กน้อย เมื่อกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเคลสกีในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้

ตาราง 34 ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ ตามวิธีการของ แวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้ในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 กับกลุ่มทักษะ - คณิตศาสตร์

การกำหนดคะแนนจุดตัด	ผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2	ระดับผลการเรียนในกลุ่มทักษะ - คณิตศาสตร์						χ^2	C
		4	3	2	1	0	รวม		
วิธีการของแวน เดอร์ ลินเดน	ผ่าน	43	58	112	109	11	333	154.4697**	0.2590
	ไม่ผ่าน	52	163	372	999	230	1816		
	รวม	95	221	484	1108	241	2149		
วิธีการของเคน	ผ่าน	45	75	145	158	17	440	146.4748**	0.2526
	ไม่ผ่าน	50	146	339	950	224	1709		
	รวม	95	221	484	1108	241	2149		

** P<.01

จากตาราง 34 เมื่อพิจารณาความสอดคล้องระหว่างผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 กับระดับผลการเรียนในกลุ่มทักษะ - คณิตศาสตร์ จากการทดสอบค่าไคสแควร์ ปรากฏว่า มีนัยสำคัญทางสถิติที่ระดับ .01 แสดงว่า ผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 มีความสัมพันธ์กับระดับผลการเรียนของนักเรียนในกลุ่มทักษะ - คณิตศาสตร์ แต่ถ้าพิจารณาค่าคอนติงเจนซีแล้ว จะเห็นว่าค่าคอนติงเจนซี จากการกำหนดคะแนนจุดตัดตามวิธีการของแวน เดอร์ ลินเดน สูงกว่าค่าคอนติงเจนซี จากการกำหนดคะแนนจุดตัดตามวิธีการของเคนเพียงเล็กน้อย แสดงว่า ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของ แวน เดอร์ ลินเดน สูงกว่าวิธีการของเคนเพียงเล็กน้อยเมื่อกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี ในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้

ตาราง 35 ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของ แวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีในกลุ่มนักเรียนที่มีความสามารถระดับพอใช้ในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 และแบบทดสอบวัดการคิดและแก้ปัญหา กับกลุ่มสร้างเสริมประสบการณ์ชีวิต

การกำหนดคะแนนจุดตัด	ผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 และแบบทดสอบวัดการคิดและแก้ปัญหา	ระดับผลการเรียนในกลุ่มสร้างเสริมประสบการณ์ชีวิต					χ^2	C
		4	3	2	1	รวม		
วิธีการของแวน เดอร์ ลินเดน	ผ่าน	8	47	50	61	166	36.0514**	0.1681
	ไม่ผ่าน	18	140	360	556	1074		
	รวม	26	187	410	617	1240		
วิธีการของเคน	ผ่าน	10	50	64	81	205	28.6103**	0.1502
	ไม่ผ่าน	16	137	346	536	1035		
	รวม	26	187	410	617	1240		

** P<.01

จากตาราง 35 เมื่อพิจารณาความสอดคล้องระหว่างผลการสอบ จากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 และแบบทดสอบวัดการคิดและแก้ปัญหากับระดับผลการเรียนในกลุ่มสร้างเสริมประสบการณ์ชีวิต จากการทดสอบค่าไคสแควร์ ปรากฏว่า มีนัยสำคัญทางสถิติที่ระดับ .01 แสดงว่าผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 และแบบทดสอบวัดการคิดและแก้ปัญหามีความสัมพันธ์กับระดับผลการเรียนของนักเรียนในกลุ่มสร้างเสริมประสบการณ์ชีวิต แต่ถ้าพิจารณาค่าคอนติงเจนซีแล้วจะเห็นว่า ค่าคอนติงเจนซีจากการกำหนดคะแนนจุดตัดตามวิธีการของแวน เดอร์ ลินเดน สูงกว่าค่าคอนติงเจนซีจากการกำหนดคะแนนจุดตัดตามวิธีการของเคน แสดงว่า ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของ แวน เดอร์ ลินเดน สูงกว่าวิธีการของเคน เมื่อกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้

1.2 นักเรียนที่มีความสามารถในระดับดี

ผลการเปรียบเทียบความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีในกลุ่มนักเรียนที่มีความสามารถในระดับดี แยกตามแบบทดสอบแต่ละฉบับ ปรากฏผลดังตาราง 36, 37 และ 38

ตาราง 36 ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของ แวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีในกลุ่มนักเรียนที่มีความสามารถในระดับดีในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 กับกลุ่มทักษะ - ภาษาไทย

การกำหนดคะแนนจุดตัด	ผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1	ระดับผลการเรียนในกลุ่มทักษะ - ภาษาไทย					χ^2	C
		4	3	2	1	รวม		
วิธีการของแวน เดอร์ ลินเดน	ผ่าน	399	529	124	10	1062	283.1453**	0.3311
	ไม่ผ่าน	163	603	408	64	1238		
	รวม	562	1132	532	74	2300		
วิธีการของเคน	ผ่าน	399	529	124	10	1062	283.1453**	0.3311
	ไม่ผ่าน	163	603	408	64	1238		
	รวม	562	1132	532	74	2300		

**P<.01

จากตาราง 36 เมื่อพิจารณาความสอดคล้องระหว่างผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 กับระดับผลการเรียนในกลุ่มทักษะ - ภาษาไทย จากการทดสอบค่าไคสแควร์ ปรากฏว่ามีนัยสำคัญทางสถิติที่ระดับ .01 แสดงว่า ผลการสอบจากแบบทดสอบวัดความสามารถทั่วไปฉบับที่ 1 มีความสัมพันธ์กับระดับผลการเรียนของนักเรียนในกลุ่มทักษะ - ภาษาไทย แต่ถ้าพิจารณาค่าคอนติงเจนซีแล้วจะเห็นว่าค่าคอนติงเจนซีจากการกำหนดคะแนนจุดตัดตามวิธีการของ แวน เดอร์ ลินเดน เท่ากับวิธีการของเคน แสดงว่าความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของ แวน เดอร์ ลินเดน เท่ากับวิธีการของเคน เมื่อกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีในกลุ่มนักเรียนที่มีความสามารถในระดับดี

ตาราง 37 ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของ แวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีในกลุ่มนักเรียนที่มีความสามารถในระดับดีในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 กับกลุ่มทักษะ - คณิตศาสตร์

การกำหนดคะแนนจุดตัด	ผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2	ระดับผลการเรียนในกลุ่มทักษะ - คณิตศาสตร์					χ^2	C
		4	3	2	1	รวม		
วิธีการของ	ผ่าน	340	259	163	39	801		
แวน เดอร์	ไม่ผ่าน	84	234	245	255	818	330.8660**	0.4119
ลินเดน	รวม	424	493	408	294	1619		
วิธีการของ	ผ่าน	340	259	163	39	801		
เคน	ไม่ผ่าน	84	234	245	255	818	330.8660**	0.4119
	รวม	424	493	408	294	1619		

**P<.01

จากตาราง 37 เมื่อพิจารณาความสอดคล้องระหว่างผลการสอบจากแบบทดสอบวัดความสามารถทั่วไปฉบับที่ 2 กับระดับผลการเรียนในกลุ่มทักษะ - คณิตศาสตร์ จากการทดสอบค่าไคสแควร์ ปรากฏว่า มีนัยสำคัญทางสถิติที่ระดับ .01 แสดงว่า ผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 มีความสัมพันธ์กับระดับผลการเรียนของนักเรียนในกลุ่มทักษะ - คณิตศาสตร์ แต่ถ้าพิจารณาค่าคอนติงเจนซีแล้ว จะเห็นว่า ค่าคอนติงเจนซีจากการกำหนดคะแนนจุดตัดตามวิธีการของ แวน เดอร์ ลินเดน เท่ากับวิธีการของเคน แสดงว่า ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของ แวน เดอร์ ลินเดน เท่ากับวิธีการของเคน เมื่อกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีในกลุ่มนักเรียนที่มีความสามารถในระดับดี

ตาราง 38 ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของ แวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีในกลุ่มนักเรียนที่มีความสามารถในระดับดี ในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 และแบบทดสอบวัดการคิดและแก้ปัญหากับกลุ่มสร้างเสริมประสบการณ์ชีวิต

การกำหนดคะแนนจุดตัด	ผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 และแบบทดสอบวัดการคิดและแก้ปัญหา	ระดับผลการเรียนในกลุ่มสร้างเสริมประสบการณ์ชีวิต					χ^2	C
		4	3	2	1	รวม		
วิธีการของแวน เดอร์ ลินเดน	ผ่าน	239	342	162	35	778	206.4219**	0.2752
	ไม่ผ่าน	180	720	646	196	1742		
	รวม	419	1062	808	231	2520		
วิธีการของเคน	ผ่าน	272	419	234	42	967	195.4536**	0.2683
	ไม่ผ่าน	147	643	574	189	1553		
	รวม	419	1062	808	231	2520		

**P<.01

จากตาราง 38 เมื่อพิจารณาความสอดคล้องระหว่างผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 และแบบทดสอบวัดการคิดและแก้ปัญหากับระดับผลการเรียนในกลุ่มสร้างเสริมประสบการณ์ชีวิต จากการทดสอบค่าไคสแควร์ ปรากฏว่า มีนัยสำคัญทางสถิติที่ระดับ .01 แสดงว่าผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 และแบบทดสอบวัดการคิดและแก้ปัญหามีความสัมพันธ์กับระดับผลการเรียนของนักเรียนในกลุ่มสร้างเสริมประสบการณ์ชีวิต แต่ถ้าพิจารณาค่าคอนติงเจนซีแล้วจะเห็นว่า ค่าคอนติงเจนซีจากการกำหนดคะแนนจุดตัดตามวิธีการของ แวน เดอร์ ลินเดน สูงกว่าค่าคอนติงเจนซีจากการกำหนดคะแนนจุดตัดตามวิธีการของเคน แสดงว่า ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของ แวน เดอร์ ลินเดน สูงกว่าวิธีการของเคน เมื่อกำหนดเกณฑ์มาตรฐานตามเทคนิคชองนีเดลสกีในกลุ่มนักเรียนที่มีความสามารถในระดับดี

2. การเปรียบเทียบความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน และเคนจากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้น

การกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้น ผู้วิจัยได้แบ่งกลุ่มคาบเส้นออกเป็น 2 ระดับ คือ

2.1 นักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้

2.2 นักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดี

2.1 นักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้

ผลการเปรียบเทียบความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้น สำหรับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้ แยกตามแบบทดสอบแต่ละฉบับและแต่ละกลุ่ม ปรากฏผลดังตาราง 39, 40 และ 41

ตาราง 39 ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของ แวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นในกลุ่มนักเรียนคาบเส้นที่อยู่ในระดับพอใช้ในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 กับกลุ่มทักษะ - ภาษาไทย

การกำหนดคะแนนจุดตัด	ผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1	ระดับผลการเรียนในกลุ่มทักษะ - ภาษาไทย				χ^2	C
		3	2	1	รวม		
วิธีการของแวน เดอร์ ลินเดน	ผ่าน	71	123	58	252	45.5018**	0.1800
	ไม่ผ่าน	153	491	463	1107		
	รวม	224	614	521	1359		
วิธีการของเคน	ผ่าน	176	417	262	855	65.7156**	0.2148
	ไม่ผ่าน	48	197	259	504		
	รวม	224	614	521	1359		

**P<.01

จากตาราง 39 เมื่อพิจารณาความสอดคล้องระหว่างผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 กับระดับผลการเรียนในกลุ่มทักษะ - ภาษาไทย จากการทดสอบค่าไคสแควร์ ปรากฏว่า มีนัยสำคัญทางสถิติที่ระดับ .01 แสดงว่าผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 มีความสัมพันธ์กับระดับผลการเรียนของนักเรียนในกลุ่มทักษะ - ภาษาไทย แต่ถ้าพิจารณาค่าคอนติงเจนซีแล้วจะเห็นว่าค่าคอนติงเจนซีจากการกำหนดคะแนนจุดตัดตามวิธีการของเคน สูงกว่าวิธีการของ แวน เดอร์ ลินเดน แสดงว่าความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของเคนสูงกว่าวิธีการของ แวน เดอร์ ลินเดน เมื่อกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้

ตาราง 40 ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของ แวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นในกลุ่มนักเรียนคาบเส้นที่อยู่ในระดับพอใช้ในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 กับกลุ่มทักษะ - คณิตศาสตร์

การกำหนดคะแนนจุดตัด	ผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2	ระดับผลการเรียนในกลุ่มทักษะ - คณิตศาสตร์						χ^2	C
		4	3	2	1	0	รวม		
วิธีการของแวน เดอร์ ลินเดน	ผ่าน	45	75	145	158	17	440	146.4748**	0.2526
	ไม่ผ่าน	50	146	339	950	224	1709		
	รวม	95	221	484	1108	241	2149		
วิธีการของเคน	ผ่าน	82	176	384	695	122	1459	104.4123**	0.2153
	ไม่ผ่าน	13	45	100	413	119	690		
	รวม	95	221	484	1108	241	2149		

**P<.01

จากตาราง 40 เมื่อพิจารณาความสอดคล้องระหว่างผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 กับระดับผลการเรียนในกลุ่มทักษะ - คณิตศาสตร์ จากการทดสอบค่าไคสแควร์ ปรากฏว่ามีนัยสำคัญทางสถิติที่ระดับ .01 แสดงว่า ผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 มีความสัมพันธ์กับระดับผลการเรียนของนักเรียนในกลุ่มทักษะ - คณิตศาสตร์ แต่ถ้าพิจารณาค่าคอนติงเจนซีแล้วจะเห็นว่าค่าคอนติงเจนซีจากการกำหนดคะแนนจุดตัดตามวิธีการของแวน เดอร์ ลินเดน สูงกว่าวิธีการของเคน แสดงว่า ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของ แวน เดอร์ ลินเดน สูงกว่าวิธีการของเคน เมื่อกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้

ตาราง 41 ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของ แวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นในกลุ่มนักเรียนคาบเส้นที่อยู่ในระดับพอใช้ในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 และแบบทดสอบวัดการคิดและแก้ปัญหา กับกลุ่มสร้างเสริมประสบการณ์ชีวิต

การกำหนดคะแนนจุดตัด	ผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 และแบบทดสอบวัดการคิดและแก้ปัญหา	ระดับผลการเรียนในกลุ่มสร้างเสริมประสบการณ์ชีวิต					χ^2	C
		4	3	2	1	รวม		
วิธีการของแวน เดอร์ ลินเดน	ผ่าน	10	50	64	81	205	28.6102**	0.1502
	ไม่ผ่าน	16	137	346	536	1035		
	รวม	26	187	410	617	1240		
วิธีการของเคน	ผ่าน	-	165	253	334	752	36.4049**	0.1293
	ไม่ผ่าน	-	48	157	283	488		
	รวม	-	213	410	617	1240		

**P<.01

จากตาราง 41 เมื่อพิจารณาความสอดคล้องระหว่างผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 และแบบทดสอบวัดการคิดและแก้ปัญหากับระดับผลการเรียนในกลุ่มสร้างเสริมประสบการณ์ชีวิต จากการทดสอบค่าไคสแควร์ ปรากฏว่ามีนัยสำคัญทางสถิติที่ระดับ .01 แสดงว่าผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 และแบบทดสอบวัดการคิดและแก้ปัญหามีความสัมพันธ์กับระดับผลการเรียนของนักเรียนในกลุ่มสร้างเสริมประสบการณ์ชีวิต แต่ถ้าพิจารณาค่าคอนติงเจนซีแล้วจะเห็นว่าค่าคอนติงเจนซีจากการกำหนดคะแนนจุดตัดตามวิธีการของ แวน เดอร์ ลินเดน สูงกว่าวิธีการของเคน แสดงว่า ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของ แวน เดอร์ ลินเดน สูงกว่าวิธีการของเคน เมื่อกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้

2.2 นักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดี

ผลการเปรียบเทียบความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นสำหรับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดีแยกตามแบบทดสอบแต่ละฉบับและแต่ละกลุ่ม ปรากฏผลดังตาราง 42, 43 และ 44

ตาราง 42 ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของ แวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นในกลุ่มนักเรียนคาบเส้นที่อยู่ในระดับดีในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 กับกลุ่มทักษะ - ภาษาไทย

การกำหนดคะแนนจุดตัด	ผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1	ระดับผลการเรียนในกลุ่มทักษะ - ภาษาไทย					χ^2	C
		4	3	2	1	รวม		
วิธีการของแวน เดอร์ ลินเดน	ผ่าน	399	529	124	10	1062	283.1453**	0.3311
	ไม่ผ่าน	163	603	408	64	1238		
	รวม	562	1132	532	74	2300		
วิธีการของเคน	ผ่าน	451	664	170	22	1307	283.5019**	0.3313
	ไม่ผ่าน	111	468	362	52	993		
	รวม	562	1132	532	74	2300		

**P<.01

จากตาราง 42 เมื่อพิจารณาความสอดคล้องระหว่างผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 กับระดับผลการเรียนในกลุ่มทักษะ - ภาษาไทย จากการทดสอบค่าไคสแควร์ ปรากฏว่ามีนัยสำคัญทางสถิติที่ระดับ .01 แสดงว่า ผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 มีความสัมพันธ์กับระดับผลการเรียนของนักเรียนในกลุ่มทักษะ - ภาษาไทย แต่ถ้าพิจารณาค่าคอนติงเจนซีแล้วจะเห็นว่าค่าคอนติงเจนซีจากการกำหนดคะแนนจุดตัดตามวิธีการของ แวน เดอร์ ลินเดน เท่ากับวิธีการของเคน แสดงว่าความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของแวน เดอร์ ลินเดน เท่ากับวิธีการของเคน เมื่อกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นในกลุ่มนักเรียนที่มีความสามารถในระดับดี

ตาราง 43 ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของ แวน เดอร์ ลินเดน และ เคน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นในกลุ่มนักเรียนคาบเส้นที่อยู่ในระดับดี ในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 กับกลุ่มทักษะ - คณิตศาสตร์

การกำหนดคะแนนจุดตัด	ผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2	ระดับผลการเรียนในกลุ่มทักษะ - คณิตศาสตร์					χ^2	C
		4	3	2	1	รวม		
วิธีการของแวน เดอร์ ลินเดน	ผ่าน	340	259	163	39	801	330.8660**	0.4119
	ไม่ผ่าน	84	234	245	255	818		
	รวม	424	493	408	294	1619		
วิธีการของเคน	ผ่าน	356	300	202	58	916	303.8908**	0.3975
	ไม่ผ่าน	68	193	206	236	703		
	รวม	424	493	408	294	1619		

**P<.01

จากตาราง 43 เมื่อพิจารณาความสอดคล้องระหว่างผลการสอบ จากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 กับระดับผลการเรียนในกลุ่มทักษะ - คณิตศาสตร์ จากการทดสอบค่าไคสแควร์ ปรากฏว่ามีนัยสำคัญทางสถิติที่ระดับ .01 แสดงว่า ผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 มีความสัมพันธ์กับระดับผลการเรียนของนักเรียนในกลุ่มทักษะ - คณิตศาสตร์ แต่ถ้าพิจารณาค่าคอนติงเจนซีแล้วจะเห็นว่า ค่าคอนติงเจนซีจากการกำหนดคะแนนจุดตัดตามวิธีการของ แวน เดอร์ ลินเดน สูงกว่าวิธีการของ เคน แสดงว่าความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของ แวน เดอร์ ลินเดน สูงกว่าวิธีการของเคน เมื่อกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นในกลุ่มนักเรียนที่มีความสามารถในระดับดี

ตาราง 44 ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของ แวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นในกลุ่มนักเรียนคาบเส้นที่อยู่ในระดับดีในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 และแบบทดสอบวัดการคิดและแก้ปัญหากับกลุ่มสร้างเสริมประสบการณ์ชีวิต

การกำหนดคะแนนจุดตัด	ผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 และแบบทดสอบวัดการคิดและแก้ปัญหา	ระดับผลการเรียนในกลุ่มสร้างเสริมประสบการณ์ชีวิต					χ^2	C
		4	3	2	1	รวม		
วิธีการของแวน เดอร์ ลินเดน	ผ่าน	272	419	234	42	967	195.4536**	0.2683
	ไม่ผ่าน	147	643	574	189	1553		
	รวม	419	1062	808	231	2520		
วิธีการของเคน	ผ่าน	323	610	375	73	1381	160.1653**	0.2445
	ไม่ผ่าน	96	452	433	158	1139		
	รวม	419	1062	808	231	2520		

** $p<.01$

จากตาราง 44 เมื่อพิจารณาความสอดคล้องระหว่างผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 และแบบทดสอบวัดการคิดและแก้ปัญหากับระดับผลการเรียนในกลุ่มสร้างเสริมประสบการณ์ชีวิต จากการทดสอบค่าไคสแควร์ปรากฏว่ามีนัยสำคัญทางสถิติที่ระดับ .01 แสดงว่า ผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 และแบบทดสอบวัดการคิดและแก้ปัญหามีความสัมพันธ์กับระดับผลการเรียนของนักเรียนในกลุ่มสร้างเสริมประสบการณ์ชีวิต แต่ถ้าพิจารณาค่าคอนติงเจนซีแล้วจะเห็นว่าค่าคอนติงเจนซีจากการกำหนดคะแนนจุดตัดตามวิธีการของ แวน เดอร์ ลินเดน สูงกว่าค่าคอนติงเจนซีจากการกำหนดคะแนนจุดตัดตามวิธีการของเคน แสดงว่าความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของ แวน เดอร์ ลินเดน สูงกว่าวิธีการของเคนเมื่อกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นในกลุ่มนักเรียนที่มีความสามารถในระดับดี

3. การเปรียบเทียบความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีกับวิธีการใช้กลุ่มคาบเส้นจะเปรียบเทียบดังนี้

3.1 นักเรียนที่มีความสามารถในระดับพอใช้ตามเทคนิคของนีเดลสกีกับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้

3.2 นักเรียนที่มีความสามารถในระดับดีตามเทคนิคของนีเดลสกีกับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดี

3.1 นักเรียนที่มีความสามารถในระดับพอใช้ตามเทคนิคของนีเดลสกี กับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้

ผลการเปรียบเทียบความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีของนักเรียนที่มีความสามารถในระดับพอใช้กับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้แยกตามแบบทดสอบแต่ละฉบับและแต่ละกลุ่มปรากฏผลดังตาราง 45, 46 และ 47

ตาราง 45 ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีกับวิธีการใช้กลุ่มคาบเส้นในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้ในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 กับกลุ่มทักษะ - ภาษาไทย

เกณฑ์มาตรฐานในการกำหนดคะแนนจุดตัด	ผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1	ระดับผลการเรียนในกลุ่มทักษะ - ภาษาไทย				χ^2	C
		3	2	1	รวม		
เทคนิคของนีเดลสกี	ผ่าน	56	73	37	166		
	ไม่ผ่าน	168	541	484	1193	46.9103**	0.1827
	รวม	224	614	521	1359		
วิธีการใช้กลุ่มคาบเส้น	ผ่าน	71	123	58	252		
	ไม่ผ่าน	153	491	463	1107	45.5018**	0.1800
	รวม	224	614	521	1359		

**P<.01

จากตาราง 45 เมื่อพิจารณาความสอดคล้องระหว่างผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 กับระดับผลการเรียนในกลุ่มทักษะ - ภาษาไทย จากการทดสอบค่าไคสแควร์ ปรากฏว่ามีนัยสำคัญทางสถิติที่ระดับ .01 แสดงว่า ผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 มีความสัมพันธ์กับระดับผลการเรียนของนักเรียนในกลุ่มทักษะ - ภาษาไทย แต่ถ้าพิจารณาค่าคอนติงเจนซีแล้วจะเห็นว่า ค่าคอนติงเจนซีจากการกำหนดคะแนนจุดตัดตามวิธีการของแวน เดอร์ ลินเดน เมื่อกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี มีค่าคอนติงเจนซีสูงกว่าวิธีการใช้กลุ่มคาบเส้นเพียงเล็กน้อย แสดงว่าความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของ แวน เดอร์ ลินเดน เมื่อกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีมีค่าความเที่ยงตรงสูงกว่าวิธีการใช้กลุ่มคาบเส้นเพียงเล็กน้อยในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้

ตาราง 46 ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีกับวิธีการใช้กลุ่มคาบเส้นในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้ในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 กับกลุ่มทักษะ - คณิตศาสตร์

เกณฑ์มาตรฐานในการกำหนดคะแนนจุดตัด	ผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2	ระดับผลการเรียนในกลุ่มทักษะ - คณิตศาสตร์						χ^2	C
		4	3	2	1	0	รวม		
เทคนิคของนีเดลสกี	ผ่าน	43	58	112	109	11	333		
	ไม่ผ่าน	52	163	372	999	230	1816	154.4697**	0.2590
	รวม	95	221	484	1108	241	2149		
วิธีการใช้กลุ่มคาบเส้น	ผ่าน	45	75	145	158	17	440		
	ไม่ผ่าน	50	146	339	950	224	1709	146.4748**	0.2526
	รวม	95	221	484	1108	241	2149		

**P<.01

จากตาราง 46 เมื่อพิจารณาความสอดคล้องระหว่างผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 กับระดับผลการเรียนในกลุ่มทักษะ - คณิตศาสตร์ จากการทดสอบค่าไคสแควร์ปรากฏว่ามีนัยสำคัญทางสถิติที่ระดับ .01 แสดงว่า ผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 มีความสัมพันธ์กับระดับผลการเรียนของนักเรียนในกลุ่มทักษะ - คณิตศาสตร์ แต่ถ้าพิจารณาค่าคอนติงเจนซีแล้วจะเห็นว่า ค่าคอนติงเจนซีจากการกำหนดคะแนนจุดตัดตามวิธีการของ แวน เดอร์ ลินเดน เมื่อกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีมีค่าคอนติงเจนซี สูงกว่าวิธีการใช้กลุ่มคาบเส้นเพียงเล็กน้อย แสดงว่าความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของ แวน เดอร์ ลินเดน เมื่อกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี มีค่าความเที่ยงตรงสูงกว่าวิธีการใช้กลุ่มคาบเส้นเพียงเล็กน้อยในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้

ตาราง 47 ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีกับวิธีการใช้กลุ่มคาบเส้นในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้ในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 และแบบทดสอบวัดการคิดและแก้ปัญหากับกลุ่มสร้างเสริมประสบการณ์ชีวิต

เกณฑ์มาตรฐานในการกำหนดคะแนนจุดตัด	ผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 และแบบทดสอบวัดการคิดและแก้ปัญหา	ระดับผลการเรียนในกลุ่มสร้างเสริมประสบการณ์ชีวิต					X^2	C
		4	3	2	1	รวม		
เทคนิคของนีเดลสกี	ผ่าน	8	47	50	61	166	36.0514**	0.1681
	ไม่ผ่าน	18	140	360	556	1074		
	รวม	26	187	410	617	1240		
วิธีการใช้กลุ่มคาบเส้น	ผ่าน	10	50	64	81	205	28.6103**	0.1502
	ไม่ผ่าน	16	137	346	536	1035		
	รวม	26	187	410	617	1240		

** $P<.01$

จากตาราง 47 เมื่อพิจารณาความสอดคล้องระหว่างผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 และแบบทดสอบวัดการคิดและแก้ปัญหา กับระดับผลการเรียนในกลุ่มสร้างเสริมประสบการณ์ชีวิต จากการทดสอบค่าไคสแควร์ ปรากฏว่า มีนัยสำคัญทางสถิติที่ระดับ .01 แสดงว่า ผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 และแบบทดสอบวัดการคิดและแก้ปัญหา มีความสัมพันธ์กับระดับผลการเรียนของนักเรียนในกลุ่มสร้างเสริมประสบการณ์ชีวิต แต่ถ้าพิจารณาค่าคอนติงเจนซีแล้วจะเห็นว่าค่าคอนติงเจนซีจากการกำหนดคะแนนจุดตัดตามวิธีการของ แวน เดอร์ ลินเดน เมื่อกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีมีค่าคอนติงเจนซีสูงกว่าวิธีการใช้กลุ่มคาบเส้น แสดงว่า ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของ แวน เดอร์ ลินเดน เมื่อกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี มีค่าความเที่ยงตรงสูงกว่าวิธีการใช้กลุ่มคาบเส้นในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้

3.2 นักเรียนที่มีความสามารถในระดับดีตามเทคนิคของนีเดลสกีกับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดี

ผลการเปรียบเทียบความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีของนักเรียนที่มีความสามารถในระดับดีกับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดี แยกตามแบบทดสอบแต่ละฉบับและแต่ละกลุ่ม ปรากฏผลดังตาราง 48, 49 และ 50

ตาราง 48 ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีกับวิธีการใช้กลุ่มคาบเส้นในกลุ่มนักเรียนที่มีความสามารถในระดับดี ในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 กับกลุ่มทักษะ - ภาษาไทย

เกณฑ์มาตรฐานในการกำหนดคะแนนจุดตัด	ผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1	ระดับผลการเรียนในกลุ่มทักษะ - ภาษาไทย					χ^2	C
		4	3	2	1	รวม		
เทคนิคของนีเดลสกี	ผ่าน	399	529	124	10	1062		
	ไม่ผ่าน	163	603	408	64	1238	283.1453**	0.3311
	รวม	562	1132	532	74	2300		
วิธีการใช้กลุ่มคาบเส้น	ผ่าน	399	529	124	10	1062		
	ไม่ผ่าน	163	603	408	64	1238	283.1453**	0.3311
	รวม	562	1132	532	74	2300		

**P<.01

จากตาราง 48 เมื่อพิจารณาความสอดคล้องระหว่างผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 กับระดับผลการเรียนในกลุ่มทักษะ - ภาษาไทย จากการทดสอบค่าไคสแควร์ ปรากฏว่า มีนัยสำคัญทางสถิติที่ระดับ .01 แสดงว่า ผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 มีความสัมพันธ์กับระดับผลการเรียนของนักเรียนในกลุ่มทักษะ - ภาษาไทย แต่ถ้าพิจารณาค่าคอนติงเจนซีแล้วจะเห็นว่าค่าคอนติงเจนซีจากการกำหนดคะแนนจุดตัดตามวิธีการของแวน เดอร์ ลินเดน เมื่อกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี มีค่าคอนติงเจนซีเท่ากับวิธีการใช้กลุ่มคาบเส้น แสดงว่า ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของ แวน เดอร์ ลินเดน เมื่อกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี มีค่าความเที่ยงตรงเท่ากับวิธีการใช้กลุ่มคาบเส้นในกลุ่มนักเรียนที่มีความสามารถในระดับดี

ตาราง 49 ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีกับวิธีการใช้กลุ่มคาบเส้นในกลุ่มนักเรียนที่มีความสามารถในระดับดีในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 กับกลุ่มทักษะ - คณิตศาสตร์

เกณฑ์มาตรฐานในการกำหนดคะแนนจุดตัด	ผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2	ระดับผลการเรียนในกลุ่มทักษะ - คณิตศาสตร์					χ^2	C
		4	3	2	1	รวม		
เทคนิคของนีเดลสกี	ผ่าน	340	259	163	39	801	330.8660**	0.4119
	ไม่ผ่าน	84	234	245	255	818		
	รวม	424	493	408	294	1619		
วิธีการใช้กลุ่มคาบเส้น	ผ่าน	340	259	163	39	801	330.8660**	0.4119
	ไม่ผ่าน	84	234	245	255	818		
	รวม	424	493	408	294	1619		

**P<.01

จากตาราง 49 เมื่อพิจารณาความสอดคล้องระหว่างผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 กับระดับผลการเรียนในกลุ่มทักษะ - คณิตศาสตร์ จากการทดสอบค่าไคสแควร์ ปรากฏว่า มีนัยสำคัญทางสถิติที่ระดับ .01 แสดงว่า ผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 มีความสัมพันธ์กับระดับผลการเรียนของนักเรียนในกลุ่มทักษะ - คณิตศาสตร์ แต่ถ้าพิจารณาค่าคอนติงเจนซีแล้วจะเห็นว่าค่าคอนติงเจนซีจากการกำหนดคะแนนจุดตัดตามวิธีการของ แวน เดอร์ ลินเดน เมื่อกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีมีค่าคอนติงเจนซีเท่ากับวิธีการใช้กลุ่มคาบเส้น แสดงว่า ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของ แวน เดอร์ ลินเดน เมื่อกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี มีค่าความเที่ยงตรงเท่ากับวิธีการใช้กลุ่มคาบเส้นในกลุ่มนักเรียนที่มีความสามารถในระดับดี

ตาราง 50 ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีกับวิธีการใช้กลุ่มคาบเส้นในกลุ่มนักเรียนที่มีความสามารถในระดับดี ในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 และแบบทดสอบวัดการคิดและแก้ปัญหา กับกลุ่มสร้างเสริมประสบการณ์ชีวิต

เกณฑ์มาตรฐานในการกำหนดคะแนนจุดตัด	ผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 และแบบทดสอบวัดการคิดและแก้ปัญหา	ระดับผลการเรียนในกลุ่มสร้างเสริมประสบการณ์ชีวิต					χ^2	C
		4	3	2	1	รวม		
เทคนิคของนีเดลสกี	ผ่าน	239	342	162	35	778	206.4219**	0.2752
	ไม่ผ่าน	180	720	646	196	1742		
	รวม	419	1062	808	231	2520		
วิธีการใช้กลุ่มคาบเส้น	ผ่าน	272	419	234	42	967	195.4536**	0.2683
	ไม่ผ่าน	147	643	574	189	1553		
	รวม	419	1062	808	231	2520		

**P<.01

จากตาราง 50 เมื่อพิจารณาความสอดคล้องระหว่างผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 และแบบทดสอบวัดการคิดและแก้ปัญหากับระดับผลการเรียนในกลุ่มสร้างเสริมประสบการณ์ชีวิต จากการทดสอบค่าไคสแควร์ปรากฏว่ามีนัยสำคัญทางสถิติที่ระดับ .01 แสดงว่าผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 และแบบทดสอบวัดการคิดและแก้ปัญหามีความสัมพันธ์กับระดับผลการเรียนของนักเรียนในกลุ่มสร้างเสริมประสบการณ์ชีวิต แต่ถ้าพิจารณาค่าคอนติงเจนซีแล้ว จะเห็นว่าค่าคอนติงเจนซีจากการกำหนดคะแนนจุดตัดตามวิธีการของแวน เดอร์ ลินเดน เมื่อกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีมีค่าคอนติงเจนซีสูงกว่าวิธีการใช้กลุ่มคาบเส้น แสดงว่าความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของแวน เดอร์ ลินเดน เมื่อกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีมีค่าความเที่ยงตรงสูงกว่าวิธีการใช้กลุ่มคาบเส้นในกลุ่มนักเรียนที่มีความสามารถในระดับดี

4. การเปรียบเทียบความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีกับวิธีการใช้กลุ่มคาบเส้น จะเปรียบเทียบ ดังนี้

4.1 นักเรียนที่มีความสามารถในระดับพอใช้ตามเทคนิคของนีเดลสกีกับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้

4.2 นักเรียนที่มีความสามารถในระดับดีตามเทคนิคของนีเดลสกีกับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดี

4.1 นักเรียนที่มีความสามารถในระดับพอใช้ตามเทคนิคของนีเดลสกีกับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้

ผลการเปรียบเทียบความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีของนักเรียนที่มีความสามารถในระดับพอใช้กับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้ แยกตามแบบทดสอบแต่ละฉบับและแต่ละกลุ่ม ปรากฏผล ดังตาราง 51, 52 และ 53

ตาราง 51 ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีกับวิธีการใช้กลุ่มคาบเส้นในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้ในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 กับกลุ่มทักษะ - ภาษาไทย

เกณฑ์มาตรฐานในการกำหนดคะแนนจุดตัด	ผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1	ระดับผลการเรียนในกลุ่มทักษะ - ภาษาไทย				χ^2	C
		3	2	1	รวม		
เทคนิคของนีเดลสกี	ผ่าน	71	123	58	252	45.5018**	0.1800
	ไม่ผ่าน	153	491	463	1107		
	รวม	224	614	521	1359		
วิธีการใช้กลุ่มคาบเส้น	ผ่าน	176	417	262	855	65.7156**	0.2148
	ไม่ผ่าน	48	197	259	504		
	รวม	224	614	521	1359		

**P<.01

จากตาราง 51 เมื่อพิจารณาความสอดคล้องระหว่างผลการสอบ จากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 กับระดับผลการเรียนในกลุ่มทักษะ - ภาษาไทย จากการทดสอบค่าไคสแควร์ ปรากฏว่า มีนัยสำคัญทางสถิติที่ระดับ .01 แสดงว่า ผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 มีความสัมพันธ์กับระดับผลการเรียนของนักเรียนในกลุ่มทักษะ - ภาษาไทย แต่ถ้าพิจารณาค่าคอนติงเจนซีแล้วจะเห็นว่าค่าคอนติงเจนซีจากการกำหนดคะแนนจุดตัดตามวิธีการของเคน เมื่อกำหนดเกณฑ์มาตรฐานตามวิธีการใช้กลุ่มคาบเส้นมีค่าคอนติงเจนซีสูงกว่าการใช้เทคนิคของนีเดลสกี แสดงว่าความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของเคน เมื่อกำหนดเกณฑ์มาตรฐานตามวิธีการใช้กลุ่มคาบเส้นมีค่าความเที่ยงตรงสูงกว่าการใช้เทคนิคของนีเดลสกีในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้

ตาราง 52 ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีกับวิธีการใช้กลุ่มคาบเส้นในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้ในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 กับกลุ่มทักษะ - คณิตศาสตร์

เกณฑ์มาตรฐานในการกำหนดคะแนนจุดตัด	ผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2	ระดับผลการเรียนในกลุ่มทักษะ - คณิตศาสตร์						χ^2	C
		4	3	2	1	0	รวม		
เทคนิคของนีเดลสกี	ผ่าน	45	75	145	158	17	440		
	ไม่ผ่าน	50	146	339	950	224	1709	146.4748**	0.2526
	รวม	95	221	484	1108	241	2149		
วิธีการใช้กลุ่มคาบเส้น	ผ่าน	82	176	384	695	122	1459		
	ไม่ผ่าน	13	45	100	413	119	690	104.4123**	0.2153
	รวม	95	221	484	1108	241	2149		

** $P<.01$

จากตาราง 52 เมื่อพิจารณาความสอดคล้องระหว่างผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 กับระดับผลการเรียนในกลุ่มทักษะ - คณิตศาสตร์ จากการทดสอบค่าไคสแควร์ ปรากฏว่า มีนัยสำคัญทางสถิติที่ระดับ .01 แสดงว่า ผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 มีความสัมพันธ์กับระดับผลการเรียนของนักเรียนในกลุ่มทักษะ - คณิตศาสตร์ แต่ถ้าพิจารณาค่าคอนติงเจนซีแล้วจะเห็นว่าค่าคอนติงเจนซีจากการกำหนดคะแนนจุดตัดตามวิธีการของเคน เมื่อกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีมีค่าคอนติงเจนซีสูงกว่าวิธีการใช้กลุ่มคาบเส้น แสดงว่าความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของเคนเมื่อกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีมีค่าความเที่ยงตรงสูงกว่าวิธีการใช้กลุ่มคาบเส้นในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้

ตาราง 53 ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีกับวิธีการใช้กลุ่มคาบเส้นในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้ในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 และแบบทดสอบวัดการคิดและแก้ปัญหา กับกลุ่มสร้างเสริมประสบการณ์ชีวิต

เกณฑ์มาตรฐานในการกำหนดคะแนนจุดตัด	ผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 และแบบทดสอบวัดการคิดและแก้ปัญหา	ระดับผลการเรียนในกลุ่มสร้างเสริมประสบการณ์ชีวิต					X^2	C
		4	3	2	1	รวม		
เทคนิคของนีเดลสกี	ผ่าน	10	50	64	81	205	28.6103**	0.1502
	ไม่ผ่าน	16	137	346	536	1035		
	รวม	26	187	410	617	1240		
วิธีการใช้กลุ่มคาบเส้น	ผ่าน	-	165	253	334	752	36.4049**	0.1293
	ไม่ผ่าน	-	48	157	283	488		
	รวม	-	213	410	617	1240		

**p<.01

จากตาราง 53 เมื่อพิจารณาความสอดคล้องระหว่างผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 และแบบทดสอบวัดการคิดและแก้ปัญหากับระดับผลการเรียนในกลุ่มสร้างเสริมประสบการณ์ชีวิตจากการทดสอบค่าไคสแควร์ ปรากฏว่า มีนัยสำคัญทางสถิติที่ระดับ .01 แสดงว่าผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 และแบบทดสอบวัดการคิดและแก้ปัญหามีความสัมพันธ์กับระดับผลการเรียนของนักเรียนในกลุ่มสร้างเสริมประสบการณ์ชีวิต แต่ถ้าพิจารณาค่าคอนติงเจนซีแล้วจะเห็นว่าค่าคอนติงเจนซีจากการกำหนดคะแนนจุดตัดตามวิธีการของเคนเมื่อกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี มีค่าคอนติงเจนซีสูงกว่าวิธีการใช้กลุ่มคาบเส้น แสดงว่า ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของเคน เมื่อกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีมีค่าความเที่ยงตรงสูงกว่าวิธีการใช้กลุ่มคาบเส้นในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้

4.2 นักเรียนที่มีความสามารถในระดับดีตามเทคนิคของนีเดลสกีกับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดี

ผลการเปรียบเทียบความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีของนักเรียนที่มีความสามารถในระดับดี กับนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดีแยกตามแบบทดสอบแต่ละฉบับ และแต่ละกลุ่ม ปรากฏผลดังตาราง 54, 55 และ 56

ตาราง 54 ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของเคนจากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีกับวิธีการใช้กลุ่มคาบเส้นในกลุ่มนักเรียนที่มีความสามารถในระดับดี ในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 กับกลุ่มทักษะ - ภาษาไทย

เกณฑ์มาตรฐานในการกำหนดคะแนนจุดตัด	ผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1	ระดับผลการเรียนในกลุ่มทักษะ - ภาษาไทย					χ^2	C
		4	3	2	1	รวม		
เทคนิคของนีเดลสกี	ผ่าน	399	529	124	10	1062	283.1453**	0.3311
	ไม่ผ่าน	163	603	408	64	1238		
	รวม	562	1132	532	74	2300		
วิธีการใช้กลุ่มคาบเส้น	ผ่าน	451	664	170	22	1307	283.5019**	0.3313
	ไม่ผ่าน	111	468	362	52	993		
	รวม	562	1132	532	74	2300		

**P<.01

จากตาราง 54 เมื่อพิจารณาความสอดคล้องระหว่างผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 กับระดับผลการเรียนในกลุ่มทักษะ - ภาษาไทย จากการทดสอบค่าไคสแควร์ ปรากฏว่ามีนัยสำคัญทางสถิติที่ระดับ .01 แสดงว่า ผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 มีความสัมพันธ์กับระดับผลการเรียนในกลุ่มทักษะ - ภาษาไทย แต่ถ้าพิจารณาค่าคอนติงเจนซีแล้วจะเห็นว่าค่าคอนติงเจนซีจากการกำหนดคะแนนจุดตัดตามวิธีการของเคน เมื่อกำหนดเกณฑ์มาตรฐานตามวิธีการใช้กลุ่มคาบเส้น มีค่าคอนติงเจนซีเท่ากับการใช้เทคนิคของนีเดลสกี แสดงว่าความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของเคน เมื่อกำหนดเกณฑ์มาตรฐานตามวิธีการใช้กลุ่มคาบเส้นมีค่าความเที่ยงตรงเท่ากับการใช้เทคนิคของนีเดลสกี ในกลุ่มนักเรียนที่มีความสามารถในระดับดี

ตาราง 55 ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีกับวิธีการใช้กลุ่มคาบเส้นในกลุ่มนักเรียนที่มีความสามารถในระดับดี ในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 กับกลุ่มทักษะ - คณิตศาสตร์

เกณฑ์มาตรฐานในการกำหนดคะแนนจุดตัด	ผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2	ระดับผลการเรียนในกลุ่มทักษะ - คณิตศาสตร์					χ^2	C
		4	3	2	1	รวม		
เทคนิคของนีเดลสกี	ผ่าน	340	259	163	39	801	330.8660**	0.4119
	ไม่ผ่าน	84	234	245	255	818		
	รวม	424	493	408	294	1619		
วิธีการใช้กลุ่มคาบเส้น	ผ่าน	356	300	202	58	916	303.8908**	0.3975
	ไม่ผ่าน	68	193	206	236	703		
	รวม	424	493	408	294	1619		

**P<.01

จากตาราง 55 เมื่อพิจารณาความสอดคล้องระหว่างผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 กับระดับผลการเรียนในกลุ่มทักษะ - คณิตศาสตร์ จากการทดสอบค่าไคสแควร์ ปรากฏว่ามีนัยสำคัญทางสถิติที่ระดับ .01 แสดงว่า ผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 มีความสัมพันธ์กับระดับผลการเรียนในกลุ่มทักษะ - คณิตศาสตร์ แต่ถ้าพิจารณาค่าคอนติงเจนซีแล้วจะเห็นว่าค่าคอนติงเจนซีจากการกำหนดคะแนนจุดตัดตามวิธีการของเคน เมื่อกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี มีค่าคอนติงเจนซีสูงกว่าวิธีการใช้กลุ่มคาบเส้นแสดงว่าความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของเคน เมื่อกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี มีค่าความเที่ยงตรงสูงกว่าวิธีการใช้กลุ่มคาบเส้นในกลุ่มนักเรียนที่มีความสามารถในระดับดี

ตาราง 56 ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีกับวิธีการใช้กลุ่มคาบเส้นในกลุ่มนักเรียนที่มีความสามารถในระดับดี ในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 และแบบทดสอบวัดการคิดและแก้ปัญหากับกลุ่มสร้างเสริมประสบการณ์ชีวิต

เกณฑ์มาตรฐานในการกำหนดคะแนนจุดตัด	ผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 และแบบทดสอบวัดการคิดและแก้ปัญหา	ระดับผลการเรียนในกลุ่มสร้างเสริมประสบการณ์ชีวิต					χ^2	C
		4	3	2	1	รวม		
เทคนิคของนีเดลสกี	ผ่าน	272	419	234	42	967		
	ไม่ผ่าน	147	643	574	189	1553	195.4536**	0.2683
	รวม	419	1062	808	231	2520		
วิธีการใช้กลุ่มคาบเส้น	ผ่าน	323	610	375	73	1381		
	ไม่ผ่าน	96	452	433	158	1139	160.1653**	0.2445
	รวม	419	1062	808	231	2520		

** P<.01

จากตาราง 56 เมื่อพิจารณาความสอดคล้องระหว่างผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 และแบบทดสอบวัดการคิดและแก้ปัญหากับระดับผลการเรียนในกลุ่มสร้างเสริมประสบการณ์ชีวิตจากการทดสอบค่าไคสแควร์ ปรากฏว่า มีนัยสำคัญทางสถิติที่ระดับ .01 แสดงว่าผลการสอบจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 และแบบทดสอบวัดการคิดและแก้ปัญหา มีความสัมพันธ์กับระดับผลการเรียนในกลุ่มสร้างเสริมประสบการณ์ชีวิต แต่ถ้าพิจารณาค่าคอนติงเจนซีแล้วจะเห็นว่าค่าคอนติงเจนซีจากการกำหนดคะแนนจุดตัดตามวิธีการของเคน เมื่อกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีมีค่าคอนติงเจนซีสูงกว่าวิธีการใช้กลุ่มคาบเส้น แสดงว่า ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของเคน เมื่อกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีมีค่าความเที่ยงตรงสูงกว่าวิธีการใช้กลุ่มคาบเส้นในกลุ่มนักเรียนที่มีความสามารถในระดับดี

สำหรับแบบทดสอบวัดคุณลักษณะทั่วไปไม่ได้มีการกำหนดคะแนนจุดตัดเหมือนกับแบบทดสอบฉบับอื่น ๆ แต่กำหนดเกณฑ์การประเมินโดยการแบ่งคะแนนออกเป็น 5 ช่วง คือ คะแนน 102 - 120 หมายถึง ดีมาก คะแนน 84 - 101 หมายถึง ดี คะแนน 66 - 83 หมายถึงปานกลาง คะแนน 48 - 65 หมายถึง ต่ำ และคะแนน 30 - 47 หมายถึง ต่ำมาก

ดังนั้นในการหาความเที่ยงตรงหรือความสอดคล้องของคะแนนจากแบบทดสอบวัดคุณลักษณะทั่วไป กับกลุ่มสร้างเสริมลักษณะนิสัย ผู้วิจัยหาความสอดคล้องของคะแนนจากนักเรียน 2 กลุ่ม คือ

1. นักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้ หมายถึง นักเรียนที่สอบปลายภาคเรียนที่ 1 ได้ระดับผลการเรียน 1 ในกลุ่มสร้างเสริมลักษณะนิสัย

2. นักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดี หมายถึง นักเรียนที่สอบปลายภาคเรียนที่ 1 ได้ระดับผลการเรียน 3 ในกลุ่มสร้างเสริมลักษณะนิสัย

1. นักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้

ผลการเปรียบเทียบความสอดคล้องของคะแนนผลสัมฤทธิ์ทางการเรียนในกลุ่มสร้างเสริมลักษณะนิสัยจากการสอบปลายปี กับแบบวัดคุณลักษณะทั่วไป ปรากฏผลดังตาราง 57

ตาราง 57 เปรียบเทียบความสอดคล้องของระดับผลการเรียนในกลุ่มสร้างเสริมลักษณะนิสัยกับแบบวัดคุณลักษณะทั่วไปของนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้

คะแนนจากแบบทดสอบวัดคุณลักษณะทั่วไป	ระดับผลการเรียนในกลุ่มสร้างเสริมลักษณะนิสัย				χ^2	C
	3	2	1	รวม		
84 - 120	54	48	44	146		
48 - 83	48	68	60	176	3.4981	0.1037
รวม	102	116	104	322		

P<.05

จากตาราง 57 เมื่อพิจารณาความสอดคล้องระหว่างคะแนนจากแบบทดสอบวัดคุณลักษณะทั่วไป กับระดับผลการเรียนในกลุ่มสร้างเสริมลักษณะนิสัยจากการทดสอบค่าไคสแควร์ ปรากฏว่าไม่มีนัยสำคัญทางสถิติ แสดงว่าผลการสอบจากแบบทดสอบวัดคุณลักษณะทั่วไปไม่มีความสัมพันธ์กับระดับผลการเรียนในกลุ่มสร้างเสริมลักษณะนิสัยในกลุ่มนักเรียนคาบเส้นที่อยู่ในระดับพอใช้

2. นักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดี

ผลการเปรียบเทียบความสอดคล้องของคะแนนผลสัมฤทธิ์ทางการเรียนในกลุ่มสร้างเสริมลักษณะนิสัยจากการสอบปลายปี กับแบบวัดคุณลักษณะทั่วไปปรากฏผลดังตาราง 58

ตาราง 58 เปรียบเทียบความสอดคล้องของระดับผลการเรียนในกลุ่มสร้างเสริมลักษณะนิสัยกับแบบวัดคุณลักษณะทั่วไปของนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดี

คะแนนจากแบบทดสอบวัดคุณลักษณะทั่วไป	ระดับผลการเรียนในกลุ่มสร้างเสริมลักษณะนิสัย				χ^2	C
	3	2	1	รวม		
84 - 120	308	531	17	1008		
30 - 83	160	318	22	672	36.3721**	0.1456
รวม	468	849	39	1680		

** $p<.01$

จากตาราง 58 เมื่อพิจารณาความสอดคล้องระหว่างคะแนนจากแบบทดสอบวัดคุณลักษณะทั่วไป กับระดับผลการเรียนในกลุ่มสร้างเสริมลักษณะนิสัยจากการทดสอบค่าไคสแควร์ ปรากฏว่ามีนัยสำคัญทางสถิติที่ระดับ .01 แสดงว่าผลการสอบจากแบบทดสอบวัดคุณลักษณะทั่วไปมีความสัมพันธ์กับระดับผลการเรียนในกลุ่มสร้างเสริมลักษณะนิสัยในกลุ่มนักเรียนคาบเส้นที่อยู่ในระดับดี ด้วยขนาดความสัมพันธ์ 0.1456

บทที่ 5

สรุป อภิปรายผลการวิจัย และข้อเสนอแนะ

การวิจัยนี้มีวัตถุประสงค์ของการวิจัยดังนี้

1. เพื่อเปรียบเทียบคะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน และเคน โดยการพิจารณาจากความสามารถต่ำสุดของผู้สอบตามเทคนิคของนีเดลสกี

2. เพื่อเปรียบเทียบคะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน และเคน โดยวิธีการใช้กลุ่มคาบเส้น

3. เพื่อเปรียบเทียบคะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน โดยการพิจารณาจากความสามารถต่ำสุดของผู้สอบตามเทคนิคของนีเดลสกีกับวิธีการใช้กลุ่มคาบเส้น

4. เพื่อเปรียบเทียบคะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคน โดยการพิจารณาจากความสามารถต่ำสุดของผู้สอบตามเทคนิคของนีเดลสกีกับวิธีการใช้กลุ่มคาบเส้น

5. เพื่อศึกษาผลการตัดสินจำนวนผู้รอบรู้และไม่รอบรู้จากการใช้คะแนนจุดตัดที่ได้จากการคำนวณในข้อ 1, 2, 3 และ 4

6. เพื่อศึกษาความเชื่อมั่นในการจำแนกความรอบรู้จากวิธีการใช้คะแนนจุดตัดที่ได้จากการคำนวณในข้อ 1, 2, 3 และ 4

7. เพื่อศึกษาความเที่ยงตรงในการจำแนกความรอบรู้จากวิธีการใช้คะแนนจุดตัดที่ได้จากการคำนวณในข้อ 1, 2, 3 และ 4

<u>สมมุติฐานของการวิจัย</u>

1. คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคนน่าจะมีค่าต่ำกว่าคะแนนจุดตัดที่ได้จากวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานโดยการพิจารณาจากความสามารถต่ำสุดของผู้สอบตามเทคนิคของนีเดลสกี

2. คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคนน่าจะมีค่าต่ำกว่าคะแนนจุดตัดที่ได้จากวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้น

3. คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นน่าจะมีค่าต่ำกว่าวิธีการพิจารณาจากความสามารถต่ำสุดของผู้สอบตามเทคนิคของนีเดลสกี

4. คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคนจากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นน่าจะมีค่าต่ำกว่าวิธีการพิจารณาจากความสามารถต่ำสุดของผู้สอบตามเทคนิคของนีเดลสกี

5. จำนวนผู้รอบรู้จากการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคนมีจำนวนมากกว่าวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานโดยการพิจารณาจากความสามารถต่ำสุดของผู้สอบตามเทคนิคของนีเดลสกี

6. จำนวนผู้รอบรู้จากการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคนมีจำนวนมากกว่าวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้น

7. จำนวนผู้รอบรู้จากการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นมีจำนวนมากกว่าวิธีการพิจารณาจากความสามารถต่ำสุดของผู้สอบตามเทคนิคของนีเดลสกี

8. จำนวนผู้รอบรู้จากการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นมีจำนวนมากกว่าวิธีการพิจารณาจากความสามารถต่ำสุดของผู้สอบตามเทคนิคของนีเดลสกี

9. ความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคนน่าจะมีความเชื่อมั่นสูงกว่าวิธีการของแวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานโดยการพิจารณาจากความสามารถต่ำสุดของผู้สอบตามเทคนิคของนีเดลสกี

10. ความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคนน่าจะมีความเชื่อมั่นสูงกว่าวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้น

11. ความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นน่าจะมีความเชื่อมั่นสูงกว่าวิธีการพิจารณาจากความสามารถต่ำสุดของผู้สอบตามเทคนิคของนีเดลสกี

12. ความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคนจากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นน่าจะมีความเชื่อมั่นสูงกว่าวิธีการพิจารณาจากความสามารถต่ำสุดของผู้สอบตามเทคนิคของนีเดลสกี

13. ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคนน่าจะมีความเที่ยงตรงสูงกว่าวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานโดยการพิจารณาจากความสามารถต่ำสุดของผู้สอบตามเทคนิคของนีเดลสกี

14. ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคนน่าจะมีความเที่ยงตรงสูงกว่าวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้น

15. ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นน่าจะมีความเที่ยงตรงสูงกว่าวิธีการพิจารณาจากความสามารถต่ำสุดของผู้สอบตามเทคนิคของนีเดลสกี

16. ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคนจากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นน่าจะมีความเที่ยงตรงสูงกว่าวิธีการพิจารณาจากความสามารถต่ำสุดของผู้สอบตามเทคนิคของนีเดลสกี

วิธีดำเนินการวิจัย

กลุ่มตัวอย่างที่ใช้ในการวิจัยครั้งนี้มี 2 กลุ่ม กลุ่มแรกคือ กลุ่มตัวอย่างผู้เชี่ยวชาญที่ใช้ในการพิจารณาแบบทดสอบประมาณ 20 คนต่อแบบทดสอบ 1 ฉบับ เพื่อกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี กลุ่มที่สองคือ กลุ่มตัวอย่างนักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2533 ที่มีผลการเรียนอยู่ในระดับคาบเส้นโดยแบ่งเป็น กลุ่มนักเรียนที่ได้ระดับผลการเรียน 1 และกลุ่มนักเรียนที่ได้ระดับผลการเรียน 3 จากการสอบปลายภาคเรียนที่ 1 ใน 4 กลุ่ม อันได้แก่ กลุ่มทักษะ - ภาษาไทย กลุ่มทักษะ - คณิตศาสตร์ กลุ่มสร้างเสริมประสบการณ์ชีวิต และกลุ่มสร้างเสริมลักษณะนิสัย การสุ่มตัวอย่างนักเรียนใช้วิธีการสุ่มอย่างง่ายและแบบเฉพาะเจาะจง

การเก็บรวบรวมข้อมูล

ในการเก็บรวบรวมข้อมูลเพื่อใช้ในการวิจัยนั้น ผู้วิจัยดำเนินการดังนี้

1. ข้อมูลที่ได้จากผู้เชี่ยวชาญในการกำหนดเกณฑ์มาตรฐานของแบบทดสอบแต่ละฉบับตามเทคนิคของนีเดลสกีนั้น ผู้วิจัยเก็บรวบรวมข้อมูลร่วมกับสำนักงานทดสอบทางการศึกษา กรมวิชาการ กระทรวงศึกษาธิการ

2. ข้อมูลของนักเรียนที่มีความสามารถในการเรียนอยู่ในระดับคาบเส้น และข้อมูลผลสัมฤทธิ์ทางการเรียนของนักเรียน ผู้วิจัยเก็บรวบรวมข้อมูลโดยการทำหนังสือขอความร่วมมือจากหน่วยงานที่เกี่ยวข้องอันได้แก่ สำนักงานคณะกรรมการการประถมศึกษาแห่งชาติ สำนักงานคณะกรรมการการศึกษาเอกชน และสำนักการศึกษา เพื่อขอหนังสือราชการในการขอความร่วมมือ

จากโรงเรียนที่เป็นกลุ่มตัวอย่างโดยการส่งแบบกรอกรายชื่อและผลการเรียนของนักเรียนไปให้ผู้บริหารโรงเรียนทางไปรษณีย์ เพื่อให้ผู้บริหารโรงเรียนกรอกรายชื่อและคะแนนหรือระดับผลการเรียนของนักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2533 ในการสอบปลายภาคเรียนที่ 1 และการสอบปลายปี เมื่อกรอกเสร็จเรียบร้อยแล้วให้ส่งคืนผู้วิจัยทางไปรษณีย์ หลังจากนั้นผู้วิจัยจึงเลือกเฉพาะรายชื่อนักเรียนที่ได้ระดับผลการเรียน 1 และระดับผลการเรียน 3 จากการสอบปลายภาคเรียนที่ 1 ใน 4 กลุ่ม เพื่อนำรายชื่อของนักเรียนไปหากระดาษคำตอบจากการสอบของโครงการตรวจสอบคุณภาพการศึกษา

เมื่อได้กระดาษคำตอบของนักเรียนแล้วนำไปวิเคราะห์โดยกำหนดรหัสดังนี้ (1) หมายถึง การตอบถูกต้อง (2) หมายถึง การตอบไม่ถูกต้อง และ (9) หมายถึงการไม่ตอบหรือเว้นว่างไว้

ส่วนข้อมูลผลสัมฤทธิ์ทางการเรียนของนักเรียน ผู้วิจัยใช้ผลการเรียนจากการสอบปลายปีใน 4 กลุ่มคือ กลุ่มทักษะ - ภาษาไทย กลุ่มทักษะ - คณิตศาสตร์ กลุ่มสร้างเสริมประสบการณ์ชีวิต และกลุ่มสร้างเสริมลักษณะนิสัยโดยศึกษาเฉพาะผลสัมฤทธิ์ทางการเรียนของนักเรียนที่ได้ระดับผลการเรียน 1 และระดับผลการเรียน 3 ในแต่ละกลุ่มจากการสอบปลายภาคเรียนที่ 1

<u>การวิเคราะห์ข้อมูล</u>

ในการวิเคราะห์ข้อมูล ผู้วิจัยดำเนินการดังนี้

1. หาเกณฑ์มาตรฐานของแบบทดสอบแต่ละฉบับตามเทคนิคของนีเดลสกี

2. นำคะแนนที่ได้จากแบบทดสอบแต่ละฉบับไปวิเคราะห์โดยใช้แบบจำลอง Three - Parameter Logistic เพื่อประมาณค่าความสามารถของผู้สอบและค่าพารามิเตอร์ของข้อสอบ

3. หาคะแนนจุดตัดตามวิธีการของ แวน เดอร์ ลินเดน โดยการแปลงเกณฑ์มาตรฐานในสเกลของคะแนนโดเมนให้เป็นระดับความสามารถโดยอาศัยโค้งคุณลักษณะแบบทดสอบ

4. หาคะแนนจุดตัดตามวิธีการของเคน

5. แปลงคะแนนจุดตัดตามวิธีการของ แวน เดอร์ ลินเดน และเคน จากสเกลของความสามารถเป็นสเกลของคะแนนโดเมน

6. หาความเชื่อมั่นของการจำแนกผู้สอบเป็นผู้รอบรู้ - ไม่รอบรู้โดยใช้วิธีการของฮวิน

7. เปรียบเทียบความแตกต่างของจำนวนผู้รอบรู้จากการสอบด้วยแบบทดสอบแต่ละรายวิชาโดยการใช้คะแนนจุดตัดแต่ละวิธีด้วยการทดสอบค่าสัดส่วน

8. หาความเที่ยงตรงหรือความสอดคล้องของคะแนนจากแบบทดสอบแต่ละฉบับกับผลสัมฤทธิ์ทางการเรียนในแต่ละกลุ่มด้วยการทดสอบค่าไคสแควร์ และค่าคอนติงเจนซี

<u>สรุปผลการวิจัย</u>

1. จากการเปรียบเทียบคะแนนจุดตัด สรุปได้ดังนี้

1.1 คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี พบว่า

1.1.1 นักเรียนที่มีความสามารถในระดับพอใช้ คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน สูงกว่าวิธีการของเคนในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 และแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 แต่แบบทดสอบวัดการคิดและแก้ปัญหา ปรากฏว่าทั้งสองวิธีมีคะแนนจุดตัดเท่ากัน

1.1.2 นักเรียนที่มีความสามารถในระดับดี คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน และเคน มีค่าเท่ากันในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 และแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 แต่แบบทดสอบวัดการคิดและแก้ปัญหา ปรากฏว่าคะแนนจุดตัดตามวิธีการของ แวน เดอร์ ลินเดน สูงกว่าวิธีการของเคน

1.2 คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้น พบว่า

1.2.1 นักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้ คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน สูงกว่าวิธีการของเคนในแบบทดสอบทุกฉบับอันได้แก่ แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 และแบบทดสอบวัดการคิดและแก้ปัญหา

1.2.2 นักเรียนกลุ่มคาบเส้นที่อยู่ในระดับดี คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน สูงกว่าวิธีการของเคนในแบบทดสอบทุกฉบับอันได้แก่ แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 และแบบทดสอบวัดการคิดและแก้ปัญหา

1.3 คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีกับวิธีการใช้กลุ่มคาบเส้นพบว่า

1.3.1 การกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีของนักเรียนที่มีความสามารถในระดับพอใช้จะได้คะแนนจุดตัดสูงกว่าการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นของนักเรียนที่อยู่ในระดับพอใช้ในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 และแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 แต่แบบทดสอบวัดการคิดและแก้ปัญหา ปรากฏว่า การกำหนดเกณฑ์มาตรฐานทั้งสองวิธีมีคะแนนจุดตัดเท่ากัน

1.3.2 การกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีของนักเรียนที่มีความสามารถในระดับดีจะได้คะแนนจุดตัดเท่ากับการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นของนักเรียนที่อยู่ในระดับดีในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 และแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 แต่แบบทดสอบวัดการคิดและแก้ปัญหา ปรากฏว่าการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีมีคะแนนจุดตัดสูงกว่าการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้น

1.4 คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีกับวิธีการใช้กลุ่มคาบเส้น พบว่า

1.4.1 การกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีของนักเรียนที่มีความสามารถในระดับพอใช้จะได้คะแนนจุดตัดสูงกว่าการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นของนักเรียนที่มีความสามารถในระดับพอใช้ในแบบทดสอบทุกฉบับ อันได้แก่ แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 และแบบทดสอบวัดการคิดและแก้ปัญหา

1.4.2 การกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีของนักเรียนที่มีความสามารถในระดับดีจะได้คะแนนจุดตัดสูงกว่าการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นของนักเรียนที่มีความสามารถในระดับดีในแบบทดสอบทุกฉบับ อันได้แก่ แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 และแบบทดสอบวัดการคิดและแก้ปัญหา

2. จากการเปรียบเทียบจำนวนผู้รอบรู้ สรุปได้ดังนี้

2.1 จำนวนผู้รอบรู้จากการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี พบว่า

2.1.1 จำนวนผู้รอบรู้จากการกำหนดคะแนนจุดตัดโดยวิธีการของเคนมีมากกว่าวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้ ในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 และแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 เมื่อทดสอบความแตกต่างของจำนวนผู้รอบรู้ ปรากฏว่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

2.1.2 จำนวนผู้รอบรู้จากการกำหนดคะแนนจุดตัดโดยวิธีการของ แวน เดอร์ ลินเดน มีจำนวนเท่ากับวิธีการของเคนจากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีในกลุ่มนักเรียนที่มีความสามารถในระดับดีในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1

แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 และแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 แต่แบบทดสอบวัดการคิดและแก้ปัญหา ปรากฏว่าจำนวนผู้รอบรู้จากการกำหนดคะแนนจุดตัดตามวิธีการของเคนมีมากกว่าวิธีการของ แวน เดอร์ ลินเดน เมื่อทดสอบความแตกต่างของจำนวนผู้รอบรู้ปรากฏว่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

2.2 จำนวนผู้รอบรู้จากการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้น พบว่า

2.2.1 จำนวนผู้รอบรู้จากการกำหนดคะแนนจุดตัดตามวิธีการของเคนมีมากกว่าวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นในกลุ่มนักเรียนคาบเส้นที่อยู่ในระดับพอใช้ในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 และแบบทดสอบวัดการคิดและแก้ปัญหา เมื่อทดสอบความแตกต่างของจำนวนผู้รอบรู้ ปรากฏว่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

2.2.2 จำนวนผู้รอบรู้จากการกำหนดคะแนนจุดตัดตามวิธีการของเคนมีมากกว่าวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นในกลุ่มนักเรียนคาบเส้นที่อยู่ในระดับดีในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 และแบบทดสอบวัดการคิดและแก้ปัญหา เมื่อทดสอบความแตกต่างของจำนวนผู้รอบรู้ปรากฏว่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

2.3 จำนวนผู้รอบรู้จากการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีกับวิธีการใช้กลุ่มคาบเส้น พบว่า

2.3.1 จำนวนผู้รอบรู้จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นมีมากกว่าการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีในกลุ่มนักเรียนคาบเส้นที่อยู่ในระดับพอใช้ในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 และแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 เมื่อทดสอบความแตกต่างของจำนวนผู้รอบรู้ ปรากฏว่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

2.3.2 จำนวนผู้รอบรู้จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นมีมากกว่าการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีในกลุ่มนักเรียนคาบเส้นที่อยู่ในระดับดีในแบบทดสอบวัดการคิดและแก้ปัญหา เมื่อทดสอบความแตกต่างของจำนวนผู้รอบรู้ปรากฏว่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

2.4 จำนวนผู้รอบรู้จากการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคนจากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีกับวิธีการใช้กลุ่มคาบเส้น พบว่า

2.4.1 จำนวนผู้รอบรู้จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นมีมากกว่าการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้ในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 และแบบทดสอบวัดการคิดและแก้ปัญหา เมื่อทดสอบความแตกต่างของจำนวนผู้รอบรู้ปรากฏว่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

2.4.2 จำนวนผู้รอบรู้จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นมีมากกว่าการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีในกลุ่มนักเรียนที่มีความสามารถในระดับดีในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 และแบบทดสอบวัดการคิดและแก้ปัญหา เมื่อทดสอบความแตกต่างของจำนวนผู้รอบรู้ ปรากฏว่า แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

3. จากการเปรียบเทียบความเชื่อมั่นในการจำแนกความรอบรู้ สรุปได้ดังนี้

3.1 ความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี พบว่า

3.1.1 ในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้ การกำหนดคะแนนจุดตัดตามวิธีการของ แวน เดอร์ ลินเดน มีความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้สูงกว่าวิธีการของเคนในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 และแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 ส่วนแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 วิธีการของเคนมีความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้สูงกว่าวิธีการของ แวน เดอร์ ลินเดน สำหรับแบบทดสอบวัดการคิดและแก้ปัญหา ปรากฏว่ามีความเชื่อมั่นในการจำแนกผู้สอบได้เท่ากันไม่ว่าจะกำหนดคะแนนจุดตัดตามวิธีการของ แวน เดอร์ ลินเดน หรือเคนก็ตาม

3.1.2 ในกลุ่มนักเรียนที่มีความสามารถในระดับดี การกำหนดคะแนนจุดตัดตามวิธีการของ แวน เดอร์ ลินเดน และเคนมีความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้ได้เท่ากันในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 และแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 แต่แบบทดสอบวัดการคิดและแก้ปัญหาปรากฏว่าการกำหนดคะแนนจุดตัดตามวิธีการของ แวน เดอร์ ลินเดน มีความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้สูงกว่าวิธีการของเคน

3.2 ความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้น พบว่า

3.2.1 ในกลุ่มนักเรียนคาบเส้นที่อยู่ในระดับพอใช้ ความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้จากการกำหนดคะแนนจุดตัดตามวิธีการของเคนต่ำกว่าวิธีการของ แวน เดอร์ ลินเดน ในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 และแบบทดสอบวัดการคิดและแก้ปัญหา แต่แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 วิธีการของเคนมีความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้สูงกว่าวิธีการของ แวน เดอร์ ลินเดน

3.2.2 ในกลุ่มนักเรียนคาบเส้นที่อยู่ในระดับดี ความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้จากการกำหนดคะแนนจุดตัดตามวิธีการของ แวน เดอร์ ลินเดน สูงกว่าวิธีการของเคนในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 และแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 แต่แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 และแบบทดสอบวัดการคิดและแก้ปัญหา ปรากฏว่าวิธีการของเคนมีความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้สูงกว่าวิธีการของ แวน เดอร์ ลินเดน

3.3 ความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีกับวิธีการใช้กลุ่มคาบเส้น พบว่า

3.3.1 นักเรียนที่มีความสามารถในระดับพอใช้ การกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี มีความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้สูงกว่าการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 และแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 แต่แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 ปรากฏว่าการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นมีความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้สูงกว่าการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี ส่วนแบบทดสอบวัดการคิดและแก้ปัญหาปรากฏว่ามีความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้ได้เท่ากันไม่ว่าจะกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี หรือวิธีการใช้กลุ่มคาบเส้นก็ตาม

3.3.2 นักเรียนที่มีความสามารถในระดับดี การกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีและวิธีการใช้กลุ่มคาบเส้นมีความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้ได้เท่ากันในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 และแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 แต่แบบทดสอบวัดการคิดและแก้ปัญหา ปรากฏว่าการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีมีความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้สูงกว่าวิธีการใช้กลุ่มคาบเส้น

3.4 ความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคนจากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีกับวิธีการใช้กลุ่มคาบเส้น พบว่า

3.4.1 นักเรียนที่มีความสามารถในระดับพอใช้ การกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีมีความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้สูงกว่าการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 และแบบทดสอบวัดการคิดและแก้ปัญหา แต่แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 ปรากฏว่า การกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นมีความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้สูงกว่าการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี

3.4.2 นักเรียนที่มีความสามารถในระดับดี การกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีมีความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้สูงกว่าการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 และแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 แต่แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 และแบบทดสอบวัดการคิดและแก้ปัญหา ปรากฏว่าการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นมีความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้สูงกว่าการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี

4. จากการเปรียบเทียบความสอดคล้องระหว่างผลการสอบจากแบบทดสอบของโครงการตรวจสอบคุณภาพการศึกษากับผลสัมฤทธิ์ทางการเรียนหรือความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ สรุปได้ดังนี้

จากการศึกษาความสัมพันธ์ระหว่างผลการสอบจากโครงการตรวจสอบคุณภาพการศึกษากับผลสัมฤทธิ์ทางการเรียนของนักเรียน พบว่า แบบทดสอบทุกฉบับมีความสัมพันธ์กับผลสัมฤทธิ์ทางการเรียนในกลุ่มวิชาต่าง ๆ อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ยกเว้นแบบทดสอบวัดคุณลักษณะทั่วไปไม่มีความสัมพันธ์กับระดับผลการเรียนในกลุ่มสร้างเสริมลักษณะนิสัยโดยเฉพาะนักเรียนกลุ่มคาบเส้นที่อยู่ในระดับพอใช้

ในการวิจัยครั้งนี้ ผู้วิจัยได้ศึกษาต่อไปอีกว่า หากผลการสอบสอดคล้องกับผลสัมฤทธิ์ทางการเรียนของนักเรียนแล้ว การกำหนดคะแนนจุดตัดวิธีใดและการกำหนดเกณฑ์มาตรฐานแบบใดที่จะทำให้ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ได้สูงกว่ากัน ซึ่งผลการศึกษามีดังนี้

4.1 ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี พบว่า

4.1.1 ในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้ ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของ แวน เดอร์ ลินเดน สูงกว่าวิธีการของเคนเมื่อกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 และแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 กับแบบทดสอบวัดการคิดและแก้ปัญหา

4.1.2 ในกลุ่มนักเรียนที่มีความสามารถในระดับดี ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของ แวน เดอร์ ลินเดน เท่ากับวิธีการของเคนเมื่อกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 แต่แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 กับแบบทดสอบวัดการคิดและแก้ปัญหา ปรากฏว่า ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของ แวน เดอร์ ลินเดน สูงกว่าวิธีการของเคน

4.2 ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้น พบว่า

4.2.1 ในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้ ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ ตามวิธีการของเคนสูงกว่าวิธีการของ แวน เดอร์ ลินเดน เมื่อกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 แต่แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 และแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 กับ

แบบทดสอบวัดการคิดและแก้ปัญหา ปรากฏว่าความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของ แวน เดอร์ ลินเดน สูงกว่าวิธีการของเคน

4.2.2 ในกลุ่มนักเรียนที่มีความสามารถในระดับดี ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของ แวน เดอร์ ลินเดน เท่ากับวิธีการของเคน เมื่อกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 แต่แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 และแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 กับแบบทดสอบวัดการคิดและแก้ปัญหา ปรากฏว่า ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของ แวน เดอร์ ลินเดน สูงกว่าวิธีการของเคน

4.3 ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการกำหนดคะแนนจุดตัดจากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีกับวิธีการใช้กลุ่มคาบเส้น พบว่า

4.3.1 ในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้ ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีมีความเที่ยงตรงสูงกว่าการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 และแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 กับแบบทดสอบวัดการคิดและแก้ปัญหา

4.3.2 ในกลุ่มนักเรียนที่มีความสามารถในระดับดี ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีมีความเที่ยงตรงเท่ากับ การกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 แต่แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 กับแบบทดสอบวัดการคิดและแก้ปัญหา ปรากฏว่าการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีมีความเที่ยงตรงสูงกว่าการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้น

4.4 ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคนจากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี กับวิธีการใช้กลุ่มคาบเส้น พบว่า

4.4.1 ในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้ ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของเคนจากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นมีความเที่ยงตรงสูงกว่าการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี ในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 แต่แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 และแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 กับแบบทดสอบวัดการคิดและแก้ปัญหา ปรากฏว่าการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีมีความเที่ยงตรงสูงกว่าการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้น

4.4.2 ในกลุ่มนักเรียนที่มีความสามารถในระดับดี ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของเคนจากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีมีความเที่ยงตรงเท่ากับ การกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 แต่แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 และแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 กับแบบทดสอบวัดการคิดและแก้ปัญหา ปรากฏว่าการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีมีความเที่ยงตรงสูงกว่าการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้น

<u>อภิปรายผลการวิจัย</u>

การประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบเพื่อกำหนดคะแนนจุดตัดตามเทคนิคของนีเดลสกีกับวิธีการใช้กลุ่มคาบเส้น ปรากฏผลที่สำคัญและได้นำมาอภิปรายดังต่อไปนี้

1. จากการเปรียบเทียบคะแนนจุดตัด พบว่า

1.1 คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี ปรากฏว่าในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้คะแนนจุดตัดตามวิธีการของ แวน เดอร์ ลินเดน

สูงกว่าวิธีการของเคนในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 และแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 ส่วนกลุ่มนักเรียนที่มีความสามารถในระดับดีปรากฏว่า คะแนนจุดตัดตามวิธีการของ แวน เดอร์ ลินเดน สูงกว่าวิธีการของเคนในแบบทดสอบวัดการคิดและแก้ปัญหา ซึ่งสอดคล้องกับสมมุติฐานที่ตั้งไว้ว่า คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคนน่าจะมีค่าต่ำกว่าคะแนนจุดตัดที่ได้จากวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานโดยการพิจารณาจากความสามารถต่ำสุดของผู้สอบตามเทคนิคของนีเดลสกี

ส่วนคะแนนจุดตัดจากแบบทดสอบวัดการคิดและแก้ปัญหาของนักเรียนที่มีความสามารถในระดับพอใช้ และคะแนนจุดตัดจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 ของนักเรียนที่มีความสามารถในระดับดีไม่สอดคล้องกับสมมุติฐานที่ตั้งไว้ ทั้งนี้อาจเป็นเพราะการใช้เทคนิคของนีเดลสกีนั้นเป็นการพิจารณาจากการคาดคะเนของผู้เชี่ยวชาญ ซึ่งผู้เชี่ยวชาญแต่ละคนอาจมีภาพความคิดเกี่ยวกับความสามารถต่ำสุดของผู้สอบแตกต่างกัน กล่าวคือ ผู้เชี่ยวชาญที่มาจากโรงเรียนซึ่งมีนักเรียนที่มีความสามารถในการเรียนดีเป็นส่วนใหญ่หรือเกือบทั้งหมดก็จะมีภาพความคิดเกี่ยวกับความสามารถต่ำสุดของผู้สอบที่จะตอบแต่ละข้อคำถามได้ถูกต้องค่อนข้างสูง แต่ถ้าผู้เชี่ยวชาญที่มาจากโรงเรียนที่นักเรียนมีความสามารถในการเรียนอยู่ระดับปานกลางหรือต่ำก็จะมีภาพความคิดเกี่ยวกับความสามารถต่ำสุดของผู้สอบที่จะตอบแต่ละข้อคำถามได้ถูกต้องค่อนข้างต่ำ เพราะฉะนั้นการที่แบบทดสอบวัดการคิดและแก้ปัญหาในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้พบว่า คะแนนจุดตัดที่ได้จากทั้ง 2 วิธีเท่ากันหรือการที่แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 ในกลุ่มนักเรียนที่มีความสามารถในระดับดีที่พบว่า คะแนนจุดตัดจากทั้ง 2 วิธีเท่ากัน อาจเนื่องมาจาก ผู้เชี่ยวชาญมีภาพความคิดเกี่ยวกับความสามารถต่ำสุดของผู้สอบในแบบทดสอบดังกล่าวไม่แตกต่างกันมากนัก จึงทำให้ความแปรปรวนในการประมาณค่าความสามารถต่ำสุดของผู้สอบที่จะตอบแต่ละข้อคำถามได้ถูกต้องไม่แตกต่างกัน ซึ่งจะส่งผลให้คะแนนจุดตัดที่ได้จากทั้ง 2 วิธีไม่แตกต่างกัน นอกจากนี้แบบทดสอบวัด

การคิดและแก้ปัญหาเป็นแบบทดสอบที่มีข้อคำถามบางข้อมีคำตอบที่ถูกต้องมากกว่า 1 ตัวเลือก ดังนั้นจึงเป็นการยากสำหรับผู้เชี่ยวชาญที่จะคาดคะเนว่า นักเรียนที่มีความสามารถในระดับพอใช้ จะทราบว่า ตัวเลือกใดเป็นตัวเลือกที่ผิดซึ่งลักษณะของข้อสอบดังกล่าวคล้ายคลึงกับลักษณะของข้อสอบที่เรียกว่า แบบสอบเลือกตอบประเภทถูกผิดทุกตัวเลือก (Multiple True False) ซึ่งเป็นแบบทดสอบที่ ลิฟวิงสตัน และซีกี (Livingston and Zieky. 1982 : 20) กล่าวว่าเป็นข้อสอบที่ค่อนข้างยากสำหรับผู้เชี่ยวชาญในการตัดสินใจว่าตัวเลือกใดเป็นตัวเลือกที่ผู้สอบที่มีความสามารถต่ำสุดจะตัดทิ้ง

1.2 คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน และเคน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้น ปรากฏว่าในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้และในระดับดี คะแนนจุดตัดตามวิธีการของแวน เดอร์ ลินเดน สูงกว่าวิธีการของเคนในแบบทดสอบทุกฉบับ ซึ่งสอดคล้องกับสมมุติฐานที่ตั้งไว้ว่าคะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคนน่าจะมีค่าต่ำกว่าคะแนนจุดตัดที่ได้จากวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้น

1.3 คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีกับวิธีการใช้กลุ่มคาบเส้น ปรากฏว่าในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้ คะแนนจุดตัดจากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีสูงกว่าการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 และแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 ส่วนในกลุ่มนักเรียนที่มีความสามารถในระดับดีปรากฏว่าคะแนนจุดตัดจากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีสูงกว่าการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นในแบบทดสอบวัดการคิดและแก้ปัญหา ซึ่งสอดคล้องกับสมมุติฐานที่ตั้งไว้ว่า คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นน่าจะมีค่าต่ำกว่าวิธีการพิจารณาจากความสามารถต่ำสุดของผู้สอบตามเทคนิคของนีเดลสกี

สำหรับแบบทดสอบวัดการคิดและแก้ปัญหาในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้ และแบบทดสอบวัดความสามารถทั่วไป ทั้ง 3 ฉบับ คือ ฉบับที่ 1 ฉบับที่ 2 และฉบับที่ 3 ของนักเรียนที่มีความสามารถในระดับดีไม่เป็นไปตามสมมุติฐาน ทั้งนี้อาจเนื่องมาจากการคาดคะเนความสามารถของผู้สอบตามเทคนิคของนีเดลสกี และวิธีการใช้กลุ่มคาบเส้นใกล้เคียงกันจึงทำให้คะแนนจุดตัดตามวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี และวิธีการใช้กลุ่มคาบเส้นไม่แตกต่างกัน

1.4 คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีกับวิธีการใช้กลุ่มคาบเส้น ปรากฏว่านักเรียนที่มีความสามารถในระดับพอใช้และในระดับดี จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีจะได้คะแนนจุดตัดสูงกว่าการกำหนดเกณฑ์มาตรฐานตามวิธีการใช้กลุ่มคาบเส้นในแบบทดสอบทุกฉบับ ซึ่งสอดคล้องกับสมมุติฐานที่ตั้งไว้ว่า คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคนจากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นน่าจะมีค่าต่ำกว่าวิธีการพิจารณาจากความสามารถต่ำสุดของผู้สอบตามเทคนิคของนีเดลสกี

2. จากการเปรียบเทียบจำนวนผู้รอบรู้ พบว่า

2.1 จำนวนผู้รอบรู้จากการกำหนดคะแนนจุดตัดตามวิธีการของเคนมีมากกว่าวิธีการของ แวน เดอร์ ลินเดน เมื่อกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีในแบบทดสอบวัดความสามารถทั่วไปทั้ง 3 ฉบับ คือ ฉบับที่ 1 ฉบับที่ 2 และฉบับที่ 3 สำหรับนักเรียนที่มีความสามารถในระดับพอใช้ และแบบทดสอบวัดการคิดและแก้ปัญหา สำหรับนักเรียนที่มีความสามารถในระดับดี ซึ่งสอดคล้องกับสมมุติฐานที่ตั้งไว้ว่า จำนวนผู้รอบรู้จากการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคนมีจำนวนมากกว่าวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานโดยการพิจารณาจากความสามารถต่ำสุดของผู้สอบตามเทคนิคของนีเดลสกี ซึ่งปรากฏการณ์เช่นนี้อธิบายได้ว่า คะแนนจุดตัดที่ได้ตามวิธีการของเคนต่ำกว่าวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี จึงทำให้โอกาสที่ผู้สอบจะจัดอยู่ในกลุ่มรอบรู้ จากการกำหนดคะแนนจุดตัดตามวิธีการของเคน ย่อมมีจำนวนมากกว่า การกำหนดคะแนนจุดตัดตามวิธีการของ แวน เดอร์ ลินเดน

2.2 จำนวนผู้รอบรู้ที่ได้จากการกำหนดคะแนนจุดตัดตามวิธีการของเคนมีมากกว่าวิธีการของ แวน เดอร์ ลินเดน เมื่อกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นในแบบทดสอบทุกฉบับไม่ว่าจะเป็นนักเรียนที่มีความสามารถในระดับพอใช้หรือระดับดีก็ตาม ซึ่งสอดคล้องกับสมมุติฐานที่ว่า จำนวนผู้รอบรู้จากการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคน มีจำนวนมากกว่าวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้น ซึ่งปรากฏการณ์เช่นนี้อธิบายได้ว่า คะแนนจุดตัดที่ได้ตามวิธีการของเคนต่ำกว่าคะแนนจุดตัดตามวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นในแบบทดสอบทุกฉบับ จึงทำให้โอกาสที่ผู้สอบที่จะจัดอยู่ในกลุ่มรอบรู้จากการกำหนดคะแนนจุดตัดตามวิธีการของเคน ย่อมมีจำนวนมากกว่าการกำหนดคะแนนจุดตัดตามวิธีการของ แวน เดอร์ ลินเดน

2.3 จำนวนผู้รอบรู้จากการกำหนดคะแนนจุดตัดตามวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นมีมากกว่า การกำหนดเกณฑ์

มาตรฐานตามเทคนิคของนีเดลสกีในแบบทดสอบวัดความสามารถทั่วไปทั้ง 3 ฉบับ คือ ฉบับที่ 1 ฉบับที่ 2 และฉบับที่ 3 สำหรับนักเรียนที่มีความสามารถในระดับพอใช้ และแบบทดสอบวัดการคิดและแก้ปัญหาสำหรับนักเรียนที่มีความสามารถในระดับดี ซึ่งสอดคล้องกับสมมุติฐานที่ว่า จำนวนผู้รอบรู้จากการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของแวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้น มีจำนวนมากกว่าวิธีการพิจารณาจากความสามารถต่ำสุดของผู้สอบตามเทคนิคของนีเดลสกี ซึ่งปรากฏการณ์เช่นนี้อธิบายได้ว่าคะแนนจุดตัดตามวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นต่ำกว่าการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี จึงทำให้โอกาสที่ผู้สอบที่จะจัดอยู่ในกลุ่มรอบรู้จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นย่อมมีจำนวนมากกว่าการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี

2.4 จำนวนผู้รอบรู้จากการใช้คะแนนจุดตัดตามวิธีการของเคนจากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นมีมากกว่าการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีในแบบทดสอบทุกฉบับไม่ว่าจะเป็นนักเรียนที่มีความสามารถในระดับพอใช้หรือในระดับดีก็ตาม ซึ่งสอดคล้องกับสมมุติฐานที่ว่า จำนวนผู้รอบรู้จากการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคนจากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นมีจำนวนมากกว่าวิธีการพิจารณาจากความสามารถต่ำสุดของผู้สอบตามเทคนิคของนีเดลสกี ซึ่งปรากฏการณ์เช่นนี้อธิบายได้ว่า คะแนนจุดตัดตามวิธีการของเคน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นต่ำกว่าการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี จึงทำให้โอกาสที่ผู้สอบที่จะจัดอยู่ในกลุ่มรอบรู้จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นมีจำนวนมากกว่าการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี

3. จากการเปรียบเทียบความเชื่อมั่นในการจำแนกความรอบรู้ พบว่า

3.1 ความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้โดยวิธีการกำหนดคะแนนจุดตัดตามวิธีการของ แวน เดอร์ ลินเดน สูงกว่าวิธีการของเคนจากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 แบบทดสอบวัดความสามารถ

ทั่วไป ฉบับที่ 3 สำหรับนักเรียนที่มีความสามารถในระดับพอใช้ และแบบทดสอบวัดการคิดและแก้ปัญหาสำหรับนักเรียนที่มีความสามารถในระดับดี ซึ่งผลการวิจัยนี้ไม่สอดคล้องกับสมมุติฐาน ทั้งนี้เนื่องจากว่าคะแนนจุดตัดในแบบทดสอบดังกล่าวตามวิธีการของ แวน เดอร์ ลินเดน สูงกว่าวิธีการของเคนจากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี จึงมีผลทำให้ความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้สูงด้วย ซึ่งจากการศึกษาของฮวิน (Huynh. 1976 : 260) พบว่า ความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้สูงเป็นผลมาจากการกำหนดคะแนนจุดตัดสูงโดยที่ฮวินได้ทำการทดลองกำหนดคะแนนจุดตัดที่ระดับคะแนนต่าง ๆ กัน คือ เมื่อกำหนดคะแนนจุดตัดเท่ากับ 9 จะมีความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้เท่ากับ 0.277 เมื่อกำหนดคะแนนจุดตัดเท่ากับ 11 จะมีความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้เท่ากับ 0.339 และเมื่อกำหนดคะแนนจุดตัดเท่ากับ 13 จะมีความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้เท่ากับ 0.361 แต่อย่างไรก็ตามความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้นี้จะมีค่าเพิ่มขึ้นตามคะแนนจุดตัดที่กำหนดจนถึงค่าสูงสุดแล้วต่อจากนั้นจะลดลงมาตามลำดับ จากการศึกษาของฮวินเมื่อคะแนนจุดตัดเท่ากับ 15, 17 และ 19 ปรากฏว่าความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้ลดลงมาเป็น 0.331, 0.243 และ 0.107 ตามลำดับ

สำหรับแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 ในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้ เมื่อกำหนดคะแนนจุดตัดตามวิธีการของเคนปรากฏว่ามีความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้สูงกว่าวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี ซึ่งสอดคล้องกับสมมุติฐานที่ว่า ความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคน น่าจะมีความเชื่อมั่นสูงกว่าวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานโดยการพิจารณาจากความสามารถต่ำสุดของผู้สอบตามเทคนิคของนีเดลสกี ทั้งนี้เนื่องจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 เป็นไปตามปรากฏการณ์การทดลองของฮวิน กล่าวคือ เมื่อคะแนนจุดตัด เท่ากับ 18 มีความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้เท่ากับ 0.3823 เมื่อคะแนนจุดตัดเท่ากับ 17 มีความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้เท่ากับ 0.4725 และเมื่อคะแนนจุดตัดเท่ากับ 11 มีความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้ เท่ากับ 0.4005 แสดงให้เห็นว่าความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้จะสูงสุด เมื่อคะแนนจุดตัดเท่ากับ 17 และเมื่อคะแนนจุดตัด

เท่ากับ 18 ความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้จะลดลงไม่ว่าจะกำหนดคะแนนจุดตัดโดยวิธีใดก็ตาม

3.2 ความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้โดยวิธีการกำหนดคะแนนจุดตัดตามวิธีการของ แวน เดอร์ ลินเดน สูงกว่าวิธีการของเคนจากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 แบบทดสอบวัดการคิดและแก้ปัญหา สำหรับนักเรียนที่มีความสามารถในระดับพอใช้ และแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 สำหรับนักเรียนที่มีความสามารถในระดับดี ซึ่งไม่สอดคล้องกับสมมุติฐาน ทั้งนี้เนื่องจากว่าคะแนนจุดตัดในแบบทดสอบดังกล่าว ตามวิธีการของ แวน เดอร์ ลินเดน สูงกว่าวิธีการของเคนจากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นจึงมีผลทำให้ความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้สูงด้วย ซึ่งผลการวิจัยนี้สอดคล้องกับการศึกษาของฮวิน (Huynh. 1976 : 260) ดังได้กล่าวมาแล้วในข้อ 3.1

ส่วนแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 ในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้ และแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 กับแบบทดสอบวัดการคิดและแก้ปัญหาในกลุ่มนักเรียนที่มีความสามารถในระดับดี ปรากฏว่าความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้จากการกำหนดคะแนนจุดตัดตามวิธีการของเคน สูงกว่าวิธีการของ แวน เดอร์ ลินเดน เมื่อกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นซึ่งสอดคล้องกับสมมุติฐานที่ว่า ความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคนน่าจะมีความเชื่อมั่นสูงกว่าวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้น ซึ่งความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้นี้ สอดคล้องกับการศึกษาของ ฮวิน (Huynh. 1976 : 260) กล่าวคือความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้จะมีค่าสูงสุดเมื่อคะแนนจุดตัดมีค่าใกล้กับคะแนนเฉลี่ยของแบบทดสอบ ซึ่งแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 ในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้ มีคะแนนเฉลี่ยเท่ากับ 14.831 เมื่อกำหนดคะแนนจุดตัดเท่ากับ 20, 19 และ 14 จะมีความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้เท่ากับ 0.2741, 0.2461 และ 0.3741 ตามลำดับ ในทำนองเดียวกัน แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 ในกลุ่มนักเรียนที่มีความสามารถในระดับดี

มีคะแนนเฉลี่ยเท่ากับ 17.717 เมื่อกำหนดคะแนนจุดตัดเท่ากับ 18 และ 17 จะมีความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้ได้เท่ากับ 0.3848 และ 0.4765 ตามลำดับ ส่วนแบบทดสอบวัดการคิดและแก้ปัญหาในกลุ่มนักเรียนที่มีความสามารถในระดับดีมีคะแนนเฉลี่ยเท่ากับ 17.243 เมื่อกำหนดคะแนนจุดตัดเท่ากับ 19, 18 และ 17 จะมีความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้เท่ากับ 0.2007, 0.1937 และ 0.2158 ตามลำดับ นั่นแสดงให้เห็นว่า ฉ คะแนนเฉลี่ยของแบบทดสอบจะจำแนกผู้สอบเป็นผู้รอบรู้ได้ดีที่สุด

3.3 ความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้โดยวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี มีความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้สูงกว่าการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 สำหรับนักเรียนที่มีความสามารถในระดับพอใช้ และแบบทดสอบวัดการคิดและแก้ปัญหาสำหรับนักเรียนที่มีความสามารถในระดับดี ซึ่งไม่สอดคล้องกับสมมุติฐาน ทั้งนี้เนื่องจากว่าคะแนนจุดตัดในแบบทดสอบดังกล่าวจากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีสูงกว่าการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นจึงทำให้ความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีสูงกว่าการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้น ซึ่งผลการวิจัยนี้สอดคล้องกับการศึกษาของ ฮวิน (Huynh. 1976 : 260) ดังได้กล่าวมาแล้วในข้อ 3.1

ส่วนแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 ในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้ ปรากฏว่าความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐาน โดยวิธีการใช้กลุ่มคาบเส้นสูงกว่าการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีซึ่งผลการวิจัยนี้สอดคล้องกับสมมุติฐานที่ว่า ความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นน่าจะมีความเชื่อมั่นสูงกว่าวิธีการพิจารณาจากความสามารถต่ำสุดของผู้สอบตามเทคนิคของนีเดลสกี ทั้งนี้อาจเนื่องมาจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 เมื่อกำหนดคะแนนจุดตัด 18, 17 และ 11 จะมีความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้ เท่ากับ 0.3823 , 0.4725 และ

0.4005 แสดงว่า ความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้จะสูงสุดเมื่อคะแนนจุดตัดเท่ากับ 17 และเมื่อคะแนนจุดตัดเท่ากับ 18 ความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้จะลดลงไม่ว่าจะกำหนดคะแนนจุดตัดโดยวิธีใดก็ตาม ซึ่งปรากฏการณ์นี้สอดคล้องกับการศึกษาของ ฮวิน (Huynh. 1976 : 260) ดังได้กล่าวมาแล้วในข้อ 3.1

3.4 ความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้ โดยวิธีการของเคนจากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี มีความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้สูงกว่าการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 แบบทดสอบวัดการคิดและแก้ปัญหาสำหรับนักเรียนที่มีความสามารถในระดับพอใช้ และแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 สำหรับนักเรียนที่มีความสามารถในระดับดี ซึ่งผลการวิจัยนี้ไม่สอดคล้องกับสมมุติฐาน ทั้งนี้เนื่องจากว่าคะแนนจุดตัดในแบบทดสอบดังกล่าวจากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีสูงกว่าการกำหนดเกณฑ์มาตรฐาน โดยวิธีการใช้กลุ่มคาบเส้นจึงทำให้ความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีสูงกว่าการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้น ซึ่งผลการวิจัยนี้สอดคล้องกับการศึกษาของ ฮวิน (Huynh. 1976 : 260) ดังได้กล่าวมาแล้วในข้อ 3.1

ส่วนแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 ในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้ และแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 กับแบบทดสอบวัดการคิดและแก้ปัญหาในกลุ่มนักเรียนที่มีความสามารถในระดับดี ปรากฏว่าความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของเคน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นสูงกว่าการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี ซึ่งสอดคล้องกับสมมุติฐานที่ว่าความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคนจากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นน่าจะมีความเชื่อมั่นสูงกว่าวิธีการพิจารณาจากความสามารถต่ำสุดของผู้สอบตามเทคนิคของนีเดลสกี ทั้งนี้อาจเนื่องมาจากแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 มีคะแนนเฉลี่ยของแบบทดสอบเท่ากับ 14.831 และเมื่อคะแนนจุดตัดเท่ากับ 20, 19 และ 14 จะมีความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้เท่ากับ 0.2741, 0.2461 และ 0.3741 ตามลำดับ สำหรับแบบทดสอบวัดความ

สามารถทั่วไป ฉบับที่ 2 มีคะแนนเฉลี่ยเท่ากับ 17.717 เมื่อคะแนนจุดตัดเท่ากับ 18 และ 17 จะมีความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้เท่ากับ 0.3848 และ 0.4765 ตามลำดับ ส่วนแบบทดสอบวัดการคิดและแก้ปัญหา มีคะแนนเฉลี่ยเท่ากับ 17.243 เมื่อคะแนนจุดตัดเท่ากับ 19, 18 และ 17 จะมีความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้เท่ากับ 0.2007, 0.1937 และ 0.2158 ตามลำดับซึ่งแสดงให้เห็นว่า แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 ในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้ แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 และแบบทดสอบวัดการคิดและแก้ปัญหาในกลุ่มนักเรียนที่มีความสามารถในระดับดีมีความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้ได้สูงสุดเมื่อคะแนนจุดตัดมีค่าใกล้เคียงกับคะแนนเฉลี่ยของแบบทดสอบไม่ว่าคะแนนจุดตัดจะกำหนดจากเกณฑ์มาตรฐานด้วยวิธีใดก็ตาม ซึ่งผลการวิจัยนี้สอดคล้องกับการศึกษาของ ฮวิน (Huynh. 1976 : 260) ดังได้กล่าวมาแล้วในข้อ 3.2

4. การเปรียบเทียบความสอดคล้องระหว่างผลการสอบจากแบบทดสอบของโครงการตรวจสอบคุณภาพการศึกษากับผลสัมฤทธิ์ทางการเรียน พบว่า แบบทดสอบทุกฉบับมีความสัมพันธ์กับผลสัมฤทธิ์ทางการเรียนของนักเรียนในกลุ่มวิชาต่าง ๆ แสดงว่า นักเรียนที่สอบผ่านแบบทดสอบต่าง ๆ ของโครงการตรวจสอบคุณภาพการศึกษาก็จะสอบผ่านแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนในกลุ่มวิชาที่มีเนื้อหาใกล้เคียงกันด้วย ซึ่งแสดงให้เห็นว่าแบบทดสอบของโครงการตรวจสอบคุณภาพการศึกษาได้วัดสมรรถภาพต่าง ๆ ตามที่ระบุไว้ในหลักสูตรที่กำหนดในแต่ละกลุ่มวิชา ยกเว้น แบบทดสอบวัดคุณลักษณะทั่วไปไม่มีความสัมพันธ์กับผลการเรียนของนักเรียนในกลุ่มสร้างเสริมลักษณะนิสัย โดยเฉพาะนักเรียนในกลุ่มคาบเส้นที่อยู่ในระดับพอใช้ ทั้งนี้อาจเป็นเพราะนักเรียนในกลุ่มคาบเส้นที่อยู่ในระดับพอใช้ยังไม่สัมฤทธิ์ผลตามสมรรถภาพต่าง ๆ ที่ระบุไว้ในหลักสูตรในกลุ่มสร้างเสริมลักษณะนิสัยถึงแม้ว่าแบบทดสอบได้วัดตามสมรรถภาพต่าง ๆ ตามหลักสูตรแล้วก็ตาม

ส่วนการเปรียบเทียบความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ พบว่า

4.1 ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ ตามวิธีการของ แวน เดอร์ ลินเดน สูงกว่าวิธีการของเคนจากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีในแบบทดสอบทุกฉบับในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้ และแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 กับแบบทดสอบวัดการคิดและแก้ปัญหา ในกลุ่มนักเรียนที่มีความสามารถในระดับดี ซึ่งผลการวิจัยนี้ไม่สอดคล้องกับสมมุติฐาน ทั้งนี้น่าจะเป็นผลมาจากการกำหนดคะแนนจุดตัด กล่าว

คือ ในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 กับแบบทดสอบวัดการคิดและแก้ปัญหา ในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้ และแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 กับแบบทดสอบวัดการคิดและแก้ปัญหา ในกลุ่มนักเรียนที่มีความสามารถในระดับดี ปรากฏว่า การกำหนดคะแนนจุดตัดตามวิธีการของแวน เดอร์ ลินเดน สูงกว่าวิธีการของเคนจึงมีผลทำให้ความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้ได้สูง ยกเว้นแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 เมื่อความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้สูงก็จะทำให้ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้สูงตามไปด้วย เนื่องจากความเชื่อมั่นและความเที่ยงตรงมีความสัมพันธ์กัน (Gulliksen. 1967 : 33)

4.2 ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของ แวน เดอร์ ลินเดน สูงกว่าวิธีการของเคนจากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 กับแบบทดสอบวัดการคิดและแก้ปัญหาในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้ และแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 กับแบบทดสอบวัดการคิดและแก้ปัญหา ในกลุ่มนักเรียนที่มีความสามารถในระดับดี ซึ่งผลการวิจัยนี้ไม่สอดคล้องกับสมมุติฐาน ทั้งนี้น่าจะเป็นผลมาจากการกำหนดคะแนนจุดตัด กล่าวคือในแบบทดสอบดังกล่าวปรากฏว่า การกำหนดคะแนนจุดตัดตามวิธีการของ แวน เดอร์ ลินเดน สูงกว่าวิธีการของเคนจึงมีผลทำให้ความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้สูง ยกเว้นแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 ในกลุ่มนักเรียนที่มีความสามารถในระดับดี ดังนั้นเมื่อความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้สูงจึงมีผลทำให้ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้สูงด้วย

สำหรับแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 ในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้ ปรากฏว่า ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของเคนสูงกว่าวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้น ซึ่งผลการวิจัยนี้สอดคล้องกับสมมุติฐานที่ว่า ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบ ตามวิธีการของเคนน่าจะมีความเที่ยงตรงสูงกว่าวิธีการของแวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้น ทั้งนี้น่าจะเป็นผลมาจากการกำหนดคะแนนจุดตัด กล่าวคือ ในแบบทดสอบ

วัดความสามารถทั่วไป ฉบับที่ 1 เมื่อคะแนนจุดตัดเท่ากับ 14 จะมีความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้ได้สูงสุดเนื่องจากคะแนนจุดตัดมีค่าใกล้เคียงกับคะแนนเฉลี่ยของแบบทดสอบ ซึ่งเท่ากับ 14.831 ดังนั้นเมื่อความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้สูงจึงมีผลทำให้ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ด้วย

4.3 ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ตามวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี สูงกว่าการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 กับแบบทดสอบวัดการคิดและแก้ปัญหาในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้ และแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 กับแบบทดสอบวัดการคิดและแก้ปัญหาในกลุ่มนักเรียนที่มีความสามารถในระดับดี ซึ่งผลการวิจัยนี้ไม่สอดคล้องกับสมมุติฐาน ทั้งนี้น่าจะเป็นผลมาจากการกำหนดคะแนนจุดตัด กล่าวคือในแบบทดสอบดังกล่าว ปรากฏว่า การกำหนดคะแนนจุดตัดตามวิธีการของ แวน เดอร์ ลินเดน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีสูงกว่าวิธีการใช้กลุ่มคาบเส้น จึงมีผลทำให้ความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้สูง ยกเว้นแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 ดังนั้นเมื่อความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้สูงก็จะทำให้ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้สูงด้วย

4.4 ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการกำหนดคะแนนจุดตัดตามวิธีการของเคน จากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี สูงกว่าการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นในแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 กับแบบทดสอบวัดการคิดและแก้ปัญหา ในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้ และแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 แบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 3 กับแบบทดสอบวัดการคิดและแก้ปัญหาในกลุ่มนักเรียนที่มีความสามารถในระดับดี ซึ่งผลการวิจัยนี้ไม่สอดคล้องกับสมมุติฐาน ทั้งนี้เนื่องจากการกำหนดคะแนนจุดตัดในแบบทดสอบดังกล่าวโดยวิธีการของเคนจากการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกีสูงกว่าการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้น จึงมีผลทำให้ความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้สูง ยกเว้นแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 2 ในกลุ่ม

นักเรียนที่มีความสามารถในระดับดี และเมื่อความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้สูง ก็จะทำให้ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้สูงด้วย

สำหรับแบบทดสอบวัดความสามารถทั่วไป ฉบับที่ 1 ในกลุ่มนักเรียนที่มีความสามารถในระดับพอใช้ปรากฏว่า ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการกำหนดคะแนนจุดตัดตามวิธีการของเคน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นสูงกว่าการกำหนดเกณฑ์มาตรฐานตามเทคนิคของนีเดลสกี ซึ่งสอดคล้องกับสมมุติฐานที่ว่า ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้โดยการใช้คะแนนจุดตัดที่ได้จากการประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบตามวิธีการของเคน จากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้นน่าจะมีความเที่ยงตรงสูงกว่าวิธีการพิจารณาจากความสามารถต่ำสุดของผู้สอบตามเทคนิคของนีเดลสกี ทั้งนี้น่าจะเป็นผลมาจากคะแนนจุดตัดของแบบทดสอบดังกล่าว มีผลทำให้ความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้ได้สูงสุด เนื่องจากคะแนนจุดตัดของแบบทดสอบมีค่าใกล้เคียงกับคะแนนเฉลี่ยของแบบทดสอบ ดังนั้นเมื่อความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้สูงก็จะส่งผลทำให้ความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้สูงด้วย ดังที่ได้กล่าวมาแล้วในข้อ 4.2

อย่างไรก็ตามจากผลการวิจัยนี้ไม่อาจสรุปได้ว่าวิธีการกำหนดคะแนนจุดตัดวิธีใดและการกำหนดเกณฑ์มาตรฐานแบบใดที่จะทำให้ความเชื่อมั่นและความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ได้สูงกว่ากัน ทั้งนี้เนื่องจากความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้นั้นขึ้นอยู่กับองค์ประกอบหลาย ๆ ประการ ประการแรกคือ คะแนนจุดตัด (Cutoff Score) ถ้ากำหนดคะแนนจุดตัดสูงก็จะทำให้ความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้สูง ความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้จะมีค่าสูงสุดตามคะแนนจุดตัดที่เพิ่มขึ้น และเมื่อเพิ่มขึ้นถึงค่าสูงสุดแล้วต่อจากนั้นค่าความเชื่อมั่นจะลดลงตามลำดับซึ่งค่าความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้นั้นจะเพิ่มสูงสุดเมื่อคะแนนจุดตัดมีค่าใกล้เคียงกับคะแนนเฉลี่ยของแบบทดสอบ ประการที่สองคือ ความแปรผันของคะแนนสอบ (Test Score Variability) ถ้าคะแนนจากการสอบมีการกระจายมาก นั่นคือผู้สอบมีความสามารถแตกต่างกันมากหรือที่เรียกว่า กลุ่มวิวิธพันธ์ (Heterogeneous Group) ก็จะทำให้ความแปรปรวนของคะแนนที่ได้จากการสอบสูง เมื่อความแปรปรวนมีค่าสูงก็จะมีผลทำให้ความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้สูงด้วย ในทำนองเดียวกันถ้าคะแนนที่ได้จากการสอบมีการกระจายน้อย นั่นคือผู้สอบมีความสามารถใกล้เคียงกันมากหรือที่เรียกว่า กลุ่มเอกพันธ์ (Homogeneous Group) ก็จะทำให้ความแปรปรวนของคะแนนที่ได้จากการสอบมีค่าต่ำ เมื่อความ

แปรปรวนมีค่าต่ำก็จะมีผลทำให้ความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้ได้ต่ำด้วย ประการที่สาม คือความยาวของแบบทดสอบ (Test Length) กล่าวคือ แบบทดสอบที่มีความยาวมากก็จะมีผลทำให้ความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้ได้สูงกว่าแบบทดสอบที่มีความยาวน้อย (Huynh. 1976 : 258 - 261)

สำหรับในการวิจัยครั้งนี้องค์ประกอบที่ไม่มีความสัมพันธ์กับความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้ ก็คือ องค์ประกอบเกี่ยวกับความยาวของแบบทดสอบ เนื่องจากแบบทดสอบทุกฉบับที่นำมาศึกษานั้นมีความยาวเท่ากัน คือมีจำนวน 30 ข้อ ทุกฉบับ ส่วนองค์ประกอบเกี่ยวกับความแปรผันของคะแนนสอบน่าจะมีความสัมพันธ์กับความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้น้อยมาก เนื่องจากในการวิจัยครั้งนี้ ผู้วิจัยศึกษานักเรียนทั้ง 2 กลุ่มคือ นักเรียนที่มีความสามารถในระดับพอใช้ และนักเรียนที่มีความสามารถในระดับดี ซึ่งแต่ละกลุ่มที่ศึกษาก็จะมีลักษณะของกลุ่มเป็นแบบเอกพันธ์ กล่าวคือความสามารถของผู้สอบภายในกลุ่มจะใกล้เคียงกันมาก หรือมีความแตกต่างกันค่อนข้างน้อย เพราะฉะนั้นองค์ประกอบที่จะส่งผลต่อความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้สูงหรือไม่ จากการวิจัยครั้งนี้น่าจะมีองค์ประกอบมาจาก การกำหนดคะแนนจุดตัดเป็นสำคัญ กล่าวคือ แบบทดสอบบางฉบับสามารถเพิ่มค่าความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้ได้เมื่อคะแนนจุดตัดเท่ากับคะแนนเฉลี่ยของแบบทดสอบ แต่แบบทดสอบบางฉบับการเพิ่มค่าความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้ จะขึ้นอยู่กับวิธีการและเกณฑ์มาตรฐานในการกำหนดคะแนนจุดตัด

ส่วนความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้จะขึ้นอยู่กับความเชื่อมั่นในการจำแนกผู้สอบเป็นผู้รอบรู้ด้วย ทั้งนี้เนื่องจากว่าความเที่ยงตรงและความเชื่อมั่นมีความสัมพันธ์กัน ซึ่งเขียนในรูปแบบจำลองทางคณิตศาสตร์จะได้ว่า $r_{xt} = \sqrt{r_{gh}}$ (Gulliksen. 1967 : 33) เมื่อ r_{xt} คือ ค่าความเที่ยงตรงและ r_{gh} คือ ค่าความเชื่อมั่น จากความสัมพันธ์ของความเชื่อมั่นกับความเที่ยงตรงนี้จะเห็นได้ว่า ค่าความเที่ยงตรงจะไม่มีโอกาสสูงไปกว่าค่าความเชื่อมั่น

จากการวิจัยครั้งนี้มีข้อที่น่าสังเกตคือ ไม่สามารถศึกษาความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ของแบบทดสอบวัดคุณลักษณะทั่วไปได้ทั้งนี้เนื่องจากแบบทดสอบวัดคุณลักษณะทั่วไปมีลักษณะเป็นแบบมาตราส่วนประมาณค่า ซึ่งแบบวัดในลักษณะนี้ยังไม่เคยมีการศึกษาถึงคะแนนจุดตัดว่าจะกำหนดคะแนนจุดตัดด้วยวิธีการใดจึงจะเหมาะสม

ข้อเสนอแนะ

จากผลการวิจัยในครั้งนี้ ผู้วิจัยใคร่เสนอแนะสำหรับผู้ที่จะนำผลการวิจัยไปใช้และผู้ที่คิดจะค้นคว้าวิจัยเรื่องทำนองนี้ต่อไป ดังนี้

1. ข้อเสนอแนะสำหรับผู้ที่จะนำผลการวิจัยไปใช้

1.1 ในระดับชั้นเรียนหรือในระดับโรงเรียน ซึ่งเป็นระดับที่ครูผู้สอนรู้จักนักเรียนดีที่สุด จากสภาพการเรียนการสอนว่าในแต่ละวิชานักเรียนในชั้นเรียนคนใดมีความสามารถในระดับใด โดยเฉพาะอย่างยิ่งนักเรียนที่มีความสามารถต่ำสุดที่จะยอมให้ผ่านในแต่ละระดับควรมีความรู้ความสามารถเพียงใดในเนื้อหาวิชาที่เรียน หากครูผู้สอนรู้จักนักเรียนของตนเป็นอย่างดีแล้ว การกำหนดเกณฑ์มาตรฐานเพื่อจะหาคะแนนจุดตัดที่นักเรียนจะสอบผ่านในแต่ละระดับ น่าจะใช้เทคนิคของนีเดลสกีเพราะเป็นเทคนิคที่ง่ายที่ครูผู้สอนในโรงเรียนสามารถจะนำไปใช้ได้ ถ้าในโรงเรียนมีนักเรียนในระดับเดียวกันหลาย ๆ ห้องเรียน และมีครูผู้สอนในวิชาเดียวกันหลายคน ครูผู้สอนก็อาจจะร่วมกันออกข้อสอบฉบับเดียวกันแล้วกำหนดเกณฑ์มาตรฐานที่จะให้นักเรียนที่มีความสามารถต่ำสุดสอบผ่านตามเทคนิคของนีเดลสกีโดยดำเนินการตามลำดับดังได้กล่าวในรายละเอียดเกี่ยวกับเทคนิคของนีเดลสกีไว้แล้วในบทที่ 2 เพราะฉะนั้นหลังจากที่ครูผู้สอนได้เกณฑ์มาตรฐานแล้ว ต่อจากนั้นจึงหาคะแนนจุดตัดตามวิธีการของ แวน เดอร์ ลินเดน หรือเคนก็ได้ เนื่องจากทั้ง 2 วิธีการนี้คะแนนจุดตัดที่ได้จะใกล้เคียงกันมาก ยิ่งกลุ่มนักเรียนที่มีความสามารถในระดับดีแล้ว จะพบว่าคะแนนจุดตัดตามวิธีการของ แวน เดอร์ ลินเดน และเคน เท่ากัน นอกจากนี้แล้วความเชื่อมั่นและความเที่ยงตรงในการจำแนกผู้สอบเป็นผู้รอบรู้ก็ไม่สามารถสรุปได้ว่าวิธีการใดสูงกว่ากัน ซึ่งขึ้นอยู่กับคะแนนจุดตัดเป็นสำคัญ

1.2 ในระดับกลุ่มโรงเรียนหรือในระดับประเทศ เนื่องจากมีนักเรียนเป็นจำนวนมากและนักเรียนในแต่ละโรงเรียนก็มีความสามารถไม่เหมือนกัน กล่าวคือโรงเรียนที่มีนักเรียนส่วนใหญ่มีผลการเรียนดีหรือเป็นนักเรียนค่อนข้างเก่ง ความสามารถต่ำสุดของนักเรียนที่จะยอมให้สอบผ่านในแต่ละระดับก็ค่อนข้างสูง ในทำนองเดียวกันโรงเรียนที่มีนักเรียนส่วนใหญ่มีผลการเรียนอยู่ในระดับปานกลางหรือค่อนข้างอ่อน ความสามารถต่ำสุดของนักเรียนที่จะยอมให้สอบผ่านในแต่ละระดับก็ค่อนข้างต่ำ เมื่อเป็นเช่นนี้แล้ว หากใช้เทคนิคของนีเดลสกีในการพิจารณาความ

สามารถต่ำสุดของผู้สอบที่จะผ่านนั้นก็จะเป็นการยากเนื่องจากครูผู้สอนแต่ละคนมาจากโรงเรียนที่ต่างกันจึงทำให้การคาดคะเนเกี่ยวกับความสามารถต่ำสุดของผู้สอบที่จะยอมให้ผ่านต่างกัน ซึ่งการคาดคะเนเกี่ยวกับความสามารถต่ำสุดของผู้สอบที่จะให้ผ่านนั้น ย่อมเกิดความคลาดเคลื่อนขึ้นได้ หากครูผู้สอนส่วนใหญ่มาจากโรงเรียนที่มีนักเรียนอยู่ในระดับเก่งหรือค่อนข้างอ่อนมาก ๆ ดังนั้นวิธีการที่เหมาะสมที่สุดในการกำหนดเกณฑ์มาตรฐานก็คือ วิธีการใช้กลุ่มคาบเส้น โดยการให้ครูผู้สอนระบุว่านักเรียนคนใดอยู่ในกลุ่มคาบเส้นจากการพิจารณาถึงผลการเรียนในรายวิชาเดียวกันจากภาคเรียนหรือปีการศึกษาที่ผ่านมา แล้วประมาณค่าความสามารถของผู้สอบโดยการใช้แบบจำลองทางคณิตศาสตร์ อาจเป็น One - Parameter Logistic, Two - Parameter Logistic หรือ Three - Parameter Logistic แล้วแต่ความเหมาะสม ต่อจากนั้นจึงหาคะแนนจุดตัดตามวิธีการของ แวน เดอร์ ลินเดน หรือเคนก็ได้ แต่ถ้าจะนำไปใช้กับกลุ่มนักเรียนที่มีความสามารถแตกต่างกันคือ นักเรียนที่มีความสามารถในระดับพอใช้และนักเรียนที่มีความสามารถในระดับดี วิธีการและเกณฑ์มาตรฐานที่เหมาะสมน่าจะใช้วิธีการของเคนจากการกำหนดเกณฑ์มาตรฐานโดยวิธีการใช้กลุ่มคาบเส้น

2. ข้อเสนอแนะสำหรับผู้ที่คิดจะทำวิจัยในทำนองนี้ต่อไป จากการวิจัยในครั้งนี้ชี้ให้เห็นว่า แบบวัดที่อยู่ในลักษณะมาตราส่วนประมาณค่าซึ่งเป็นลักษณะของแบบวัดที่ใช้มากในการวัดเกี่ยวกับความรู้สึก ความคิดเห็น ตลอดจนเจตคติต่าง ๆ หรือกล่าวอีกนัยหนึ่งคือแบบวัดด้านเจตพิสัย (Affective Domain) ซึ่งแบบวัดในลักษณะนี้ยังไม่มีการศึกษาถึงคะแนนจุดตัดว่าควรจะกำหนดคะแนนจุดตัดในลักษณะอย่างไรจึงจะเหมาะสมกับมาตรวัดนั้น ๆ ซึ่งเป็นเรื่องที่น่าศึกษา ค้นคว้า และทำวิจัยเพื่อหาข้อสรุปต่อไป

บรรณานุกรม

บรรณานุกรม

กาญจนา วัธนสุนทร. การสร้างแบบทดสอบอิงเกณฑ์วิชาคณิตศาสตร์. วิทยานิพนธ์ ค.ม. กรุงเทพฯ : จุฬาลงกรณ์มหาวิทยาลัย, 2521. อัดสำเนา.

เกศริน บุญเกิด. การสร้างแบบทดสอบอิงเกณฑ์วิชาภาษาไทยเรื่องการเขียนสะกดคำ ชั้นประถมศึกษาปีที่ 4. ปริญญานิพนธ์ กศ.ม. กรุงเทพฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2526. อัดสำเนา.

ไกวิท ประวาลพฤกษ์ และสมศักดิ์ สินธุระเวชญ์. การประเมินในชั้นเรียน. กรุงเทพฯ : วัฒนาพานิช, 2523.

จำเนียร สุขหลาย. การใช้ทฤษฎีการตัดสินใจของเบส์กำหนดจุดตัดแบบทดสอบอิงเกณฑ์เพื่อตรวจสอบคุณภาพการวัดและประเมินผลของนักศึกษาครู. รวมบทคัดย่องานวิจัยและวิทยานิพนธ์ สาขาวิชาการศึกษาของข้าราชการ กรมการฝึกหัดครู, 2524. อัดสำเนา.

ชมภู จันทรอมรพร. การใช้ทฤษฎีการตัดสินใจของเบส์กำหนดคะแนนจุดตัดของแบบทดสอบอิงเกณฑ์. วิทยานิพนธ์ ค.ม. กรุงเทพฯ : จุฬาลงกรณ์มหาวิทยาลัย, 2523. อัดสำเนา.

ชวลิต ไพธิ์นคร. การเปรียบเทียบผลของการกำหนดจุดตัดแบบทดสอบอิงเกณฑ์ชนิดเลือกตอบโดยวิธีของ เบอร์ก วิธีการประยุกต์ราชค์ไมเคล และวิธีกำหนดเกณฑ์ผ่านระดับต่ำสุด. วิทยานิพนธ์ ศษ.ม. เชียงใหม่ : มหาวิทยาลัยเชียงใหม่, 2528. อัดสำเนา.

ทดสอบทางการศึกษา, สำนักงาน. รายงานการประเมินผลการนำหลักสูตรไปใช้ ปีการศึกษา 2531 ระดับประถมศึกษา. เอกสารประกอบการประชุม หมายเลข 5. อัดสำเนา.

_____. คู่มือการดำเนินงานโครงการติดตามประเมินการใช้หลักสูตรของสถานศึกษา ปีการศึกษา 2533. ม.ป.ท., ม.ป.ป. อัดสำเนา.

ไพฑูรย์ เวทการ. การสร้างแบบทดสอบอิงเกณฑ์วิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1 (ม.1). ปริญญานิพนธ์ กศ.ม. กรุงเทพฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2524. อัดสำเนา.

มัทนี เผื่อนน้อย. การสร้างแบบทดสอบอิงเกณฑ์วิชาคณิตศาสตร์เรื่องร้อยละชั้นประถมศึกษาปีที่ 4. ปริญญานิพนธ์ กศ.ม. กรุงเทพฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2526. อัดสำเนา.

มาลี จิตติวุฒิการ. ผลของวิธีการหาคะแนนจุดตัด คะแนนโดเมน และความยาวของแบบสอบที่มีต่อค่าความน่าจะเป็นในการจำแนกความรอบรู้และความเที่ยงในการตัดสินใจ. วิทยานิพนธ์ ค.ม. กรุงเทพฯ : จุฬาลงกรณ์มหาวิทยาลัย, 2533. อัดสำเนา.

รังสรรค์ มณีเล็ก. การเปรียบเทียบความเที่ยงของแบบสอบอิงเกณฑ์โดยการกำหนดจุดตัดต่างกัน. วิทยานิพนธ์ ค.ม. กรุงเทพฯ : จุฬาลงกรณ์มหาวิทยาลัย, 2527. อัดสำเนา.

วิชาการ, กรม กระทรวงศึกษาธิการ. สรุปรายงานการประชุมทางวิชาการ เรื่องการขยายผลโครงการประเมินผลการนำหลักสูตรไปใช้ ปีการศึกษา 2531 วันที่ 9 - 20 กันยายน 2532 ณ. โรงแรมเวียงใต้. ม.ป.ป. อัดสำเนา.

ศุภลักษณ์ สินธนา. การเปรียบเทียบผลการใช้จุดตัดที่กำหนดโดยการประยุกต์ไลจิสติกโมเดลแบบพารามิเตอร์ตัวเดียว พารามิเตอร์ 2 ตัว และพารามิเตอร์ 3 ตัวในแบบทดสอบอิงเกณฑ์วิชาคณิตศาสตร์ชั้นประถมศึกษาปีที่ 6. วิทยานิพนธ์ ศษ.ม. เชียงใหม่ : มหาวิทยาลัยเชียงใหม่, 2531. อัดสำเนา.

อัจฉริยา ปราบอริพ่าย. ผลของการใช้สารสนเทศแก่ผู้เชี่ยวชาญต่อการกำหนดคะแนนจุดตัดของแบบสอบผลสัมฤทธิ์วิชาคณิตศาสตร์. วิทยานิพนธ์ ค.ม. กรุงเทพฯ : จุฬาลงกรณ์มหาวิทยาลัย, 2532. อัดสำเนา.

Behuniak, P.JB. "An Investigation of Standard Setting Methods for Objective-Referenced Proficiency Tests," Dissertation Abstracts International. 41 : 3998A - 3999A; March, 1981.

Berk, R.A. "Determination of Optimal Cutting Scores in Criterion - Referenced Measurement," Journal of Experimental Education. 45(2) : 4 - 9; Winter, 1976.

_____. Criterion - Referenced Measurement : The State of the Art. Maryland : The Johns Hopkins University Press, 1980.

_____. "A Consumer's Guide to Setting Performance Standards on Criterion-Referenced Tests," Review of Educational Research. 56(1) : 137 - 172 ; Spring, 1986.

Beuk, C.H. "A Method for Reaching a Compromise between Absolute and Relative Standards in Examinations," Journal of Educational Measurement. 21(2) : 147 - 152; Summer, 1984.

Behuniak. P., F.X. Archambault., and R.K. Gable. "Angoff and Nedelsky Standard Setting Procedures : Implications for the Validity of Proficiency Test Score Interpretation," Educational and Psychological Measurement. 42(1) : 247 - 255 : Spring, 1982.

Cross and others. "A Comparison of Three Methods for Establishing Minimum Standards on the National Teacher, Examinations," Journal of Educational Measurement. 21(20) : 113 - 129; Summer, 1984.

De Gruijter, D.N.M. "Compromise Models for Establishing Examination Standard," Journal of Educational Measurement. 22(4) : 263 - 269; Winter, 1985.

Downie, N.M. and R.W. Heath. Basic Statistical Methods. New York : Harper and Row Publishers, 1970.

Ferguson, G.A. Statistical Analysis in Psychology and Education. Tokyo : Kosaido Printing Co Ltd., 1981.

Glass, G.V. "Standards and Criteria," Journal of Educational Measurement. 15(4) : 237 - 261; Winter, 1978.

Gronlund, N.E. Constructing Achievement Test. 3 rd ed. Englewood Cliffs, New Jersey : Prentice Hall, 1982.

Gulliksen, H. Theory of Mental Tests. New York : John Wiley & Sons, Inc., 1967.

Halpin, G., G. Sigmon, and G. Halpin. "Minimum Competency Standards Set by Three Divergent Groups of Raters Using Three Judgmental Procedures : Implications for Validity," Educational and Psychological Measurement. 43(1) : 185 - 197; Spring, 1983.

Hambleton, R.K. "On the Use of Cut-off Scores with Criterion-Referenced Tests in Instructional Settings," Journal of Educational Measurement. 15(4) : 278 - 290; Winter, 1978.

Hambleton, R.K. and L.L. Cook. "Latent Trait Models and Their Use in the Analysis of Educational Test Data," Journal of Educational Measurement. 14(2) : 75 - 96; Summer, 1977.

Hambleton, R.K. and H. Swaminathan. Item Response Theory : Principles and Applications. Boston : Kluwer Nijhoff Publishing, 1985.

Hambleton, R.K. and others. "Criterion-Referenced Testing and Measurement : A Review of Technical Issues and Developments," Review of Educational Research. 48(1) : 1 - 47; Winter, 1978.

Hinkle, D.E. J.D. Oliver and C.A. Hinkle. "How large should the sample be ? Part II - The One - Sample Case for Survey Research," Educational and Psychological Measurement. 45(2) : 271 - 280; Summer, 1985.

Huynh, H. "On the Reliability of Decision in Domain - Referenced Testing," Journal of Educational Measurement. 13(4) : 253 - 264; Winter, 1976.

Huynh, H. and J. Casteel. "A Comparison of the Minimax and Rasch Approaches to Set Simultaneous Passing Scores for Subtests," Journal of Educational Statistics. 10(4) : 334 - 344; Winter, 1985.

Kane, M.T. "On the Use of IRT Models with Judgmental Standard Setting Procedures," Journal of Educational Measurement. 24(4) : 333 - 345; Winter, 1987.

Linn, R.L. "Demands, Cautions, and Suggestions for Setting Standards," Journal of Educational Measurement. 15(4) : 301 - 308; Winter, 1978.

Livingston, S.A. and M.J. Zieky. Passing Scores : A Manual for Setting Standards of Performance on Educational and Occupational Tests. New Jersey : Educational Testing Service, 1982.

Lord, F.M. Application of Item Response Theory to Practical Testing Problems. Hillsdale, New Jersey : Lawrence Erlbaum Associates, 1980.

Malech, A.R. "A Comparative Analysis of Four Classification Procedures for Use with Criterion-Referenced Test," Dissertation Abstracts International. 40 : 810-A; August, 1979.

Mills, C.N. "A Comparison of Three Methods of Establishing Cut - off Scores on Criterion - Referenced Tests," Journal of Educational Measurement. 20 (3) : 283 - 292; Fall, 1983.

Norcini, J.J. and others. "A Comparison of Three Variations on a Standard-Setting Method," Journal of Educational Measurement. 24(1) : 56 - 64; Spring, 1987.

Norcini, J.J. J.A. Shea and D.T. Kanya. "The Effects of Various Factors on Standard Setting," Journal of Educational Measurement. 25(1) : 57 - 65; Spring, 1988.

Plake, B.S. and G.J. Melican. "Effects of Item Context on Intrajudge Consistency of Expert Judgments Via the Nedelsky Standard Setting Method," Educational and Psychological Measurement. 49(1) : 45 - 51; Spring, 1989.

Rowley, G.L. "Historical Antecedents of the Standard-Setting Debate : An Inside Account of the Minimal - Beardedness Controversy," Journal of Educational Measurement. 19(2) : 87 - 95; Summer, 1982.

Smith, R.L. and J.K. Smith, "Differential Use of Item Information by Judges Using Angoff and Nedelsky Procedures," Journal of Educational Measurement. 25(4) : 259 - 274; Winter, 1988.

Swaminathan, H., R.K. Hambleton and J. Algina. "A Bayesian Decision-Theoretic Procedures for Use with Criterion - Referenced Tests," Journal of Educational Measurement. 12(2) : 87 - 98; Summer, 1975.

Van der Linden, W.J. "Latent Trait Method for Determining Intrajudge-Inconsistency in the Angoff and Nedelsky Techniques of Standard Setting," Journal of Educational Measurement. 19(4) : 295 - 307; Winter, 1982.

Warm, T.A. A Primer of Item Response Theory. Oklahoma : V.S.A. Coast Guard Institute, 1978.

Wood, R.L., M.S. Wingersky and F.M. Lord. "Logist a Computer Program for Estimation Examinee Ability and Item Characteristic Curve Parameters," Research Memorandum. No. 79. New Jersey : Educational Testing Service, 1976.

Zieky, M. Methods of Setting Standards of Performance on Criterion Referenced Tests. Paper Presented at the 13 th International Conference of International Association for Educational Assessment, Bangkok, Thailand, November, 1987.

ภาคผนวก

รายชื่อโรงเรียนที่เป็นกลุ่มตัวอย่าง

เขตการศึกษา	จังหวัด	ชื่อโรงเรียน
1	นนทบุรี	เขมพิทยา
		สุเหร่าปากคลองลาวี
		ชลประทานสงเคราะห์
		วัดผาสุกมณีจักร
2	ยะลา	เทศบาล 4 (ธนวิถี)
		บ้านตลาดลาใหม่
		บ้านท่าสาป
		บ้านสะเอะ
		บันนังสตาอินทรฉัตรฯ
		บ้านโกตาบารู
		บ้านพะปูเงาะ
3	สงขลา	จุลสมัย
		ครุฑศึกษา
		ชุมชนบ้านด่าน
		บ้านกลาง
		บ้านน้ำกระจาย
		วัดบ่อแดง
		บ้านนา
		ชุมชนนิคมสร้างตนเองเทพา

เขตการศึกษา	จังหวัด	ชื่อโรงเรียน
3	สงขลา	บ้านเขาพระ
		ชุมชนบ้านปาดัง
		บ้านคลองหวะ (ทวีรัตน์ราษฎร์บำรุง)
		วัดแม่เปียะ
		บ้านควนเนียง
3	นครศรีธรรมราช	วัดท่าช้าง
		วัดท้ายสำเภา
		วัดมุขธารา
		บ้านห้วยหาร
		วัดประทุมทายการาม
5	ราชบุรี	เทศบาล 2 (วัดช่องลม)
		เทศบาล 3 (เทศบาลสงเคราะห์)
		ดุสิตวิทยา
		บ้านจอมบึง (วาปีพร้อมประชาศึกษา)
		ชุมชนวัดประสาทสิทธิ์
		วัดคงคาราม (อินทราชผดุงศิลป)
6	พระนครศรีอยุธยา	ราษฎร์บำรุงศิลป์
		วัดพระญาติการาม
		วัดกำแพง
		วัดลาดชะโด

เขตการศึกษา	จังหวัด	ชื่อโรงเรียน
6	ลพบุรี	วัดรัตนาราม
		บ้านบัวชุม
7	นครสวรรค์	เทศบาลวัดไทรใต้
		เทศบาลวัดวรนาถบรรพต
		วัดดอนใหญ่
		วัดบ้านมะเกลือ
		บ้านไคกเดื่อ (ไคกเดื่อประชาสรรค์)
		วัดวังม้า (ชลประสิทธิ์วิท)
7	พิษณุโลก	วัดศรีสุทธาราม (วิไลราษฎร์อุปถัมภ์)
		บ้านพร้าว
		บ่อวิทยบางระกำ
		บ้านปลักแรด
9	อุดรธานี	บ้านเชียงยืน
		บ้านเชียงแหว
		บ้านสงเปลือยดงสามสิบ
		บ้านม่วงประชาบำรุง
		บ้านตูม
		บ้านนาจาน
		บ้านวังทอง
		ชุมชนบ้านนาคำไฮ

เขตการศึกษา	จังหวัด	ชื่อโรงเรียน
9	อุดรธานี	ชุมชนบ้านแวง
11	นครราชสีมา	ชุมชนเทศบาล 2 (รัฐราษฎ์บำรุง)
		บ้านจอหอ
		อนุบาลนครราชสีมา
		รัฐราษฎร์รังสรรค์
		ไทยวัฒนาประชารัฐ
		นิยมมิตรวิทยาคาร
		ไนนกุ่มมิตรภาพ 210
		จักราชราษฎร์สามัคคี
		บ้านปากช่อง
		บ้านหนองสาหร่าย
		เสมาอุปถัมภ์
		ไชคชัยพรหมบุตรบริหาร
		บ้านด่านเกวียน
		บ้านหาญ
		ชุมชนบ้านหนองบัวลาย
		บ้านเมืองปัก (ชุณหะวัณ)
12	ชลบุรี	เมืองพัทยา 4 (วัดหนองใหญ่)
		เชนต์ปอลคอนแวนต์
		วัดราษฎร์ศรัทธา
		วัดอุทยานนที

เขตการศึกษา	จังหวัด	ชื่อโรงเรียน
13	กรุงเทพฯ	วัดปลูกศรัทธา
		วัดทรัพย์สโมสร
		ประชาราษฎ์บำเพ็ญ

สำหรับเขตการศึกษา 4, 8 และ 10 ผู้วิจัยได้สุ่มจังหวัดภูเก็ต เชียงใหม่ และอุบลราชธานี ตามลำดับ แต่ข้อมูลที่ได้รับไม่สมบูรณ์ ดังนั้นจึงมิได้สุ่มนักเรียนมาเป็นกลุ่มตัวอย่าง ในการศึกษาครั้งนี้

ที่ ทม. 1007/0002

บัณฑิตวิทยาลัย
มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร
สุขุมวิท 23 กรุงเทพฯ 10110

4 มกราคม 2534

เรื่อง ขอความอนุเคราะห์

เรียน เลขาธิการสำนักงานคณะกรรมการการประถมศึกษาแห่งชาติ

บัณฑิตวิทยาลัยขอรับรองว่า นางสาวกัญจนา ลินทรัตนศิริกุล เป็นนิสิตระดับปริญญาเอก สาขาการวิจัยและพัฒนาหลักสูตร ของมหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร

นิสิตผู้นี้มีความประสงค์จะมาติดต่อขอความสะดวกในการศึกษาค้นคว้า เพื่อทำปริญญานิพนธ์ เรื่อง การประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบเพื่อกำหนดคะแนนจุดตัดตามเทคนิคของนีเดลสกีกับวิธีการใช้กลุ่มคาบเส้น

ทั้งนี้อยู่ในความควบคุมดูแลของ

ศ.ดร.สำเริง	บุญเรืองรัตน์	ประธาน
อ.ดร.ชูศักดิ์	ขัมภลิขิต	กรรมการ
อ.ดร.สงบ	ลักษณะ	กรรมการ

สิ่งที่นิสิตใคร่ขอความอนุเคราะห์ คือขอหนังสือราชการเพื่ออำนวยความสะดวกในการขอความร่วมมือจากผู้อำนวยการสำนักงานคณะกรรมการการประถมศึกษาจังหวัด โดยให้โรงเรียนต่าง ๆ ในจังหวัดตามรายชื่อโรงเรียนที่แนบกรอกคะแนนหรือระดับผลการเรียนของนักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2533 ในการสอบปลายภาคเรียนที่ 1 และการสอบปลายปี ใน 4 วิชา คือ ภาษาไทย คณิตศาสตร์ สร้างเสริมประสบการณ์ชีวิตและสร้างเสริมลักษณะนิสัย ในระหว่างเดือนมีนาคม 2534 เพื่อเป็นข้อมูลในการทำปริญญานิพนธ์

บัณฑิตวิทยาลัยหวังเป็นอย่างยิ่งว่าท่านจะกรุณาให้ความร่วมมือในครั้งนี้ และขอขอบคุณในการช่วยเหลืออนุเคราะห์ใด ๆ ที่ท่านจะโปรดให้แก่นิสิตผู้นี้ด้วย

ขอแสดงความนับถือ

(นายจรูญ ไกมลบุญย์)
รักษาราชการแทนคณบดีบัณฑิตวิทยาลัย

บัณฑิตวิทยาลัย
โทร. 2584119

ที่ ทม. 1007/0011

บัณฑิตวิทยาลัย
มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร
สุขุมวิท 23 กรุงเทพฯ 10110

4 มกราคม 2534

เรื่อง ขอความอนุเคราะห์

เรียน เลขาธิการสำนักงานคณะกรรมการการศึกษาเอกชน

บัณฑิตวิทยาลัยขอรับรองว่า นางสาวกัญจนา ลินทรัตนศิริกุล เป็นนิสิตระดับปริญญาเอก สาขาการวิจัยและพัฒนาหลักสูตร ของมหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร

นิสิตผู้นี้มีความประสงค์จะมาติดต่อขอความสะดวกในการศึกษาค้นคว้าเพื่อทำปริญญานิพนธ์ เรื่อง การประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบเพื่อกำหนดคะแนนจุดตัดตามเทคนิคของนีเดลสกี กับวิธีการใช้กลุ่มคาบเส้น

ทั้งนี้อยู่ในความควบคุมดูแลของ

ศ.ดร.สาเริง	บุญเรืองรัตน์	ประธาน
อ.ดร.ชูศักดิ์	บัมภลิขิต	กรรมการ
อ.ดร.สงบ	ลักษณะ	กรรมการ

สิ่งที่นิสิตใคร่ขอความอนุเคราะห์ คือขอหนังสือราชการเพื่ออำนวยความสะดวกในการขอความร่วมมือจากผู้บริหารโรงเรียนตามรายชื่อโรงเรียนที่แนบให้กรอกคะแนนหรือระดับผลการเรียนของนักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2533 ในการสอบปลายภาคเรียนที่ 1 และการสอบปลายปี ใน 4 วิชา คือ ภาษาไทย คณิตศาสตร์ สร้างเสริมประสบการณ์ชีวิต และสร้างเสริมลักษณะนิสัย ในระหว่างเดือนมีนาคม 2534 เพื่อเป็นข้อมูลในการทำปริญญานิพนธ์

บัณฑิตวิทยาลัยหวังเป็นอย่างยิ่งว่าท่านจะกรุณาให้ความร่วมมือในครั้งนี้ และขอขอบคุณในความช่วยเหลืออนุเคราะห์ใด ๆ ที่ท่านจะโปรดให้แก่นิสิตผู้นี้ด้วย

ขอแสดงความนับถือ

(นายจรูญ โกมลบุณย์)
รักษาราชการแทนคณบดีบัณฑิตวิทยาลัย

บัณฑิตวิทยาลัย
โทร. 2584119

ที่ ทม. 1007/0012

บัณฑิตวิทยาลัย
มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร
สุขุมวิท 23 กรุงเทพฯ 10110

4 มกราคม 2534

เรื่อง ขอความอนุเคราะห์

เรียน ผู้อำนวยการสำนักการศึกษา

บัณฑิตวิทยาลัยขอรับรองว่า นางสาวกัญจนา ลินทรัตนศิริกุล เป็นนิสิตระดับปริญญาเอก สาขาการวิจัยและพัฒนาหลักสูตร ของมหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร

นิสิตผู้นี้มีความประสงค์จะมาติดต่อขอความสะดวกในการศึกษาค้นคว้าเพื่อทำปริญญานิพนธ์ เรื่อง การประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบเพื่อกำหนดคะแนนจุดตัดตามเทคนิคของนีเดลสกี กับวิธีการใช้กลุ่มคาบเส้น

ทั้งนี้อยู่ในความควบคุมดูแลของ

ศ.ดร.สำเริง	บุญเรืองรัตน์	ประธาน
อ.ดร.ชูศักดิ์	ขัมภลิขิต	กรรมการ
อ.ดร.สงบ	ลักษณะ	กรรมการ

สิ่งที่นิสิตใคร่ขอความอนุเคราะห์คือ ขอหนังสือราชการเพื่ออำนวยความสะดวกในการขอความร่วมมือจากผู้บริหารโรงเรียนตามรายชื่อโรงเรียนที่แนบให้กรอกคะแนนหรือระดับผลการเรียนของนักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2533 ในการสอบปลายภาคเรียนที่ 1 และการสอบปลายปีใน 4 วิชา คือ ภาษาไทย คณิตศาสตร์ สร้างเสริมประสบการณ์ชีวิต และสร้างเสริมลักษณะนิสัย ในระหว่างเดือนมีนาคม 2534 เพื่อเป็นข้อมูลในการทำปริญญานิพนธ์

บัณฑิตวิทยาลัยหวังเป็นอย่างยิ่งว่าท่านจะกรุณาให้ความร่วมมือในครั้งนี้ และขอขอบคุณในความช่วยเหลืออนุเคราะห์ใด ๆ ที่ท่านจะโปรดให้แก่นิสิตผู้นี้ด้วย

ขอแสดงความนับถือ

(นายจรูญ ไกมลบุญย์)
รักษาราชการแทนคณบดีบัณฑิตวิทยาลัย

บัณฑิตวิทยาลัย
โทร. 2584119

ที่ ศธ 1404/607

สำนักงานคณะกรรมการการประถมศึกษาแห่งชาติ
กระทรวงศึกษาธิการ กทม. 10300

8 มกราคม 2534

เรื่อง ขอความอนุเคราะห์

เรียน ผู้อำนวยการการประถมศึกษาจังหวัด

สิ่งที่ส่งมาด้วย รายชื่อโรงเรียน 1 ฉบับ

เนื่องด้วย นางสาวกัญจนา ลินทรัตนศิริกุล นิสิตปริญญาเอก สาขาการวิจัย และพัฒนาหลักสูตร มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร ได้ขอความร่วมมือในการศึกษาค้นคว้าเพื่อทำปริญญานิพนธ์ เรื่อง การประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบเพื่อกำหนดคะแนนจุดตัดตามเทคนิคของนีเดลสกีกับวิธีการใช้กลุ่มคาบเส้น ทั้งนี้อยู่ในความควบคุมดูแลของศาสตราจารย์ ดร.สำเริง บุญเรืองรัตน์ ประธาน อาจารย์ ดร.ชูศักดิ์ ขัมภลิขิต และอาจารย์ ดร.สงบ ลักษณะ กรรมการ โดยให้โรงเรียนต่าง ๆ ในจังหวัดนี้กรอกคะแนนหรือระดับผลการเรียนของนักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2533 ในการสอบปลายภาคเรียนที่ 1 และผลการสอบปลายปีใน 4 วิชา คือ ภาษาไทย คณิตศาสตร์ สร้างเสริมประสบการณ์ชีวิต และสร้างเสริมลักษณะนิสัย ในระหว่างเดือนมีนาคม 2534 เพื่อเป็นข้อมูลในการทำปริญญานิพนธ์ สำหรับแบบกรอกคะแนนหรือระดับผลการเรียน นิสิตจะส่งโดยตรงไปยังโรงเรียนที่จะเก็บข้อมูล

สำนักงานคณะกรรมการการประถมศึกษาแห่งชาติพิจารณาแล้วเห็นควรให้ความอนุเคราะห์ให้เก็บข้อมูลในการศึกษาค้นคว้า เพื่อประโยชน์ทางวิชาการ

จึงเรียนมาเพื่อทราบและดำเนินการต่อไป

ขอแสดงความนับถือ

(นายกมล ภู่ประเสริฐ)
รองเลขาธิการปฏิบัติราชการแทน
เลขาธิการคณะกรรมการการประถมศึกษาแห่งชาติ

กองวิชาการ
โทร. 2810588

หมายเหตุ

ผอ. ปจ.	นครปฐม	นนทบุรี	ยะลา	สงขลา	นครศรีธรรมราช
	ภูเก็ต	สุพรรณบุรี	ราชบุรี	พระนครศรีอยุธยา	ลพบุรี
	นครสวรรค์	เพชรบูรณ์	พิษณุโลก	เชียงใหม่	ขอนแก่น
	อุดรธานี	อุบลราชธานี	นครราชสีมา	ชลบุรี	และผอ.ป.กทม.

ที่ ศธ 1007/1131

สำนักงานคณะกรรมการการศึกษาเอกชน
กระทรวงศึกษาธิการ กทม. 10300

9 มกราคม 2634

เรื่อง ขอความร่วมมือในการวิจัย

เรียน ผู้รับใบอนุญาต ผู้จัดการหรือครูใหญ่โรงเรียน

ด้วยนางสาวกัญจนา ลินทรัตนศิริกุล นิสิตระดับปริญญาเอก สาขาการวิจัยและพัฒนาหลักสูตร ของบัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร กำลังทำวิจัยเพื่อเรียบเรียงวิทยานิพนธ์ เรื่อง "การประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบเพื่อกำหนดคะแนนจุดตัดตามเทคนิคของนีเดลสกีกับวิธีการใช้กลุ่มคาบเส้น" มีความประสงค์จะขอความร่วมมือจากโรงเรียนในการกรอกคะแนนหรือระดับผลการเรียนของนักเรียน ชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2533 ในการสอบปลายภาคเรียนที่ 1 และการสอบปลายปี 4 วิชา คือ ภาษาไทย คณิตศาสตร์ สร้างเสริมประสบการณ์ชีวิต และสร้างเสริมลักษณะนิสัย ลงในแบบฟอร์มที่แนบ

จึงเรียนมาเพื่อขอความร่วมมือจากท่าน โปรดให้ความอนุเคราะห์แก่นิสิตด้วย ทั้งนี้เพื่อประโยชน์ทางการศึกษา

ขอแสดงความนับถือ

(นายเช็น แก้วยศ)
ผู้อำนวยการกองนโยบายและแผน

กองนโยบายและแผน
โทร. 2828654

ที่ กท 3004/258 สำนักการศึกษา

9 มกราคม 2534

เรื่อง ขอความร่วมมือในการกรอกคะแนนนักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2533

เรียน ผู้บริหารโรงเรียน

ด้วยนางสาวกัญจนา ลินทรัตนศิริกุล นิสิตชั้นบริญญาดุษฎีบัณฑิต มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร สาขาการวิจัยและพัฒนาหลักสูตร มีความประสงค์จะมาติดต่อขอความสะดวกในการศึกษาค้นคว้าเพื่อทำบริญญานิพนธ์ เรื่อง "การประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบเพื่อกำหนดคะแนนจุดตัดตามเทคนิคของนีเดลสกีกับวิธีการใช้กลุ่มคาบเส้น" ซึ่งจะต้องขอความร่วมมือจากผู้บริหารโรงเรียนให้กรอกคะแนนหรือระดับผลการเรียนของนักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2533 ในการสอบปลายภาคเรียนที่ 1 และการสอบปลายปีใน 4 วิชา คือ ภาษาไทย คณิตศาสตร์ สร้างเสริมประสบการณ์ชีวิต และสร้างเสริมลักษณะนิสัย

สำนักการศึกษาใคร่ขอความร่วมมือจากท่านในการกรอกคะแนน เพื่อเป็นข้อมูลในการทำบริญญานิพนธ์ ในการนี้ นางสาวกัญจนา ลินทรัตนศิริกุล จะส่งแบบกรอกรายชื่อและผลการเรียนของนักเรียนพร้อมหนังสือชี้แจงรายละเอียดการกรอกคะแนนไปยังผู้บริหารโรงเรียนทางไปรษณีย์

จึงเรียนมาเพื่อขอความอนุเคราะห์ และให้ความร่วมมือในการกรอกคะแนน ดังกล่าวต่อไปด้วย จักขอบคุณยิ่ง

ขอแสดงความนับถือ

(นางอุไรวรรณ ธรรมกรบัญญัติ)
รองผู้อำนวยการสำนักการศึกษา
ปฏิบัติราชการแทนผู้อำนวยการสำนักการศึกษา

กองวิชาการ
โทร. 4372047

บัณฑิตวิทยาลัย
มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร

มกราคม 2534

เรื่อง ขอความร่วมมือในการทำวิจัย

เรียน

ด้วยข้าพเจ้า นางสาวกัญจนา ลินทรัตนศิริกุล นิสิตปริญญาเอก สาขาการวิจัยและพัฒนาหลักสูตร มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร กำลังทำวิจัยเพื่อเสนอเป็นปริญญานิพนธ์ เรื่อง "การประยุกต์ทฤษฎีการตอบสนองต่อข้อสอบเพื่อกำหนดคะแนนจุดตัดตามเทคนิคของนีเดลสกี กับวิธีการใช้กลุ่มคาบเส้น" ซึ่งการทำวิจัยครั้งนี้ ได้ใช้ข้อมูลจากโครงการติดตามการใช้หลักสูตรของสถานศึกษา ปีการศึกษา 2533 ของสำนักงานทดสอบทางการศึกษา กรมวิชาการ กระทรวงศึกษาธิการ ดังนั้นเพื่อให้สามารถนำข้อมูลจากโครงการดังกล่าวมาใช้ได้ จึงมีความจำเป็นจะต้องใช้ข้อมูลบางประการจากโรงเรียนของท่านคือ

1. รายชื่อนักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2533 พร้อมกับคะแนนหรือระดับผลการเรียนของนักเรียนจากการสอบปลายภาคเรียนที่ 1 ใน 4 วิชา คือ ภาษาไทย คณิตศาสตร์ สร้างเสริมประสบการณ์ชีวิต และสร้างเสริมลักษณะนิสัย

2. รายชื่อนักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2533 พร้อมกับคะแนน หรือระดับผลการเรียนของนักเรียนจากการสอบปลายปี ใน 4 วิชา คือ ภาษาไทย คณิตศาสตร์ สร้างเสริมประสบการณ์ชีวิต และสร้างเสริมลักษณะนิสัย

อนึ่ง เพื่อให้งานวิจัยครั้งนี้สำเร็จด้วยดีจึงใคร่ขอความกรุณาจากท่านกรอกผลการเรียนของนักเรียนตามแบบฟอร์มที่แนบมานี้ หรือท่านจะถ่ายสำเนาเอกสารผลการเรียนของนักเรียนจาก ป.02/6 แล้วแต่ท่านจะสะดวก และขอให้ท่านโปรดส่งผู้วิจัยตามชื่อ และที่อยู่ที่ได้แนบมาให้ภายในเดือนมีนาคม 2534

หวังเป็นอย่างยิ่งว่าคงได้รับความร่วมมือจากท่านเป็นอย่างดี และขอขอบพระคุณเป็นอย่างสูงมา ณ โอกาสนี้ด้วย

ขอแสดงความนับถือ

(นางสาวกัญจนา ลินทรัตนศิริกุล)

แบบกรอกรายชื่อและผลการเรียนของนักเรียน
ชั้นประถมศึกษาปีที่ 6/ ห้อง.....

โรงเรียน.....จังหวัด.....

เลขที่	ชื่อ-นามสกุลของนักเรียน	ผลการสอบปลายภาคเรียนที่ 1 (คะแนนหรือระดับผลการเรียน)				ผลการสอบปลายปี (คะแนนหรือระดับผลการเรียน)			
		ภาษาไทย	คณิตศาสตร์	สปช.	สลน.	ภาษาไทย	คณิตศาสตร์	สปช.	สลน.
....	.....	.....	.....	....	....	.....	.....	....	...
....	.....	.....	.....	....	....	.....	.....	....	...
....	.....	.....	.....	....	....	.....	.....	....	...
....	.....	.....	.....	....	....	.....	.....	....	...
....	.....	.....	.....	....	....	.....	.....	....	...
....	.....	.....	.....	....	....	.....	.....	....	...
....	.....	.....	.....	....	....	.....	.....	....	...
....	.....	.....	.....	....	....	.....	.....	....	...
....	.....	.....	.....	....	....	.....	.....	....	...
....	.....	.....	.....	....	....	.....	.....	....	...
....	.....	.....	.....	....	....	.....	.....	....	...
....	.....	.....	.....	....	....	.....	.....	....	...
....	.....	.....	.....	....	....	.....	.....	....	...
....	.....	.....	.....	....	....	.....	.....	....	...
....	.....	.....	.....	....	....	.....	.....	....	...
....	.....	.....	.....	....	....	.....	.....	....	...
....	.....	.....	.....	....	....	.....	.....	....	...
....	.....	.....	.....	....	....	.....	.....	....	...
....	.....	.....	.....	....	....	.....	.....	....	...
....	.....	.....	.....	....	....	.....	.....	....	...
....	.....	.....	.....	....	....	.....	.....	....	...
....	.....	.....	.....	....	....	.....	.....	....	...
....	.....	.....	.....	....	....	.....	.....	....	...

หมายเหตุ วิชาใดที่ทางโรงเรียนไม่ได้หาระดับผลการเรียนไว้ ขอให้กรอกคะแนนดิบ

ประวัติย่อผู้วิจัย

ประวัติของผู้วิจัย

ชื่อ - นามสกุล	กัญจนา ลินทรัตนศิริกุล
วัน เดือน ปีเกิด	28 กันยายน 2493
สถานที่เกิด	7 ซอยมาตี - ไปตี ถนนสุขุมวิท 53 อำเภอพระโขนง กรุงเทพฯ 10100
ที่อยู่ปัจจุบัน	13/3 ซอยสมปรารถนา ถนนประชาสงเคราะห์ ดินแดง กรุงเทพฯ 10400
ตำแหน่งหน้าที่การงานปัจจุบัน	ผู้ช่วยศาสตราจารย์ ระดับ 8
สถานที่ทำงานปัจจุบัน	มหาวิทยาลัยสุโขทัยธรรมาธิราช ถนนเจ้งวัฒนะ ตำบลบางพูด อำเภอปากเกร็ด นนทบุรี 11120
ประวัติการศึกษา	พ.ศ. 2509 มัธยมศึกษาปีที่ 3 จากโรงเรียนมัธยมวัดธาตุทอง
	พ.ศ. 2511 มัธยมศึกษาปีที่ 5 แผนกวิทยาศาสตร์ จากโรงเรียนสมาคมโรงเรียนราษฎร์ ปัจจุบันคือโรงเรียนมักกะสันพิทยา
	พ.ศ. 2515 ค.บ. (ชีววิทยา - คณิตศาสตร์) จากจุฬาลงกรณ์มหาวิทยาลัย
	พ.ศ. 2518 ค.ม. (การวัดและประเมินผลการศึกษา) จากจุฬาลงกรณ์มหาวิทยาลัย
	พ.ศ. 2534 กศ.ด. (การวิจัยและพัฒนาหลักสูตร) จากมหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร