

## การพัฒนาดัชนีความสอดคล้อง ของแบบแผนการตอบข้อสอบรายบุคคล\*

พนิชา สังข์เพชร

### บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อพัฒนาดัชนีความสอดคล้องของแบบแผนการตอบข้อสอบรายบุคคล ค่าของดัชนีที่คำนวณได้จะบอกให้รู้ว่าคะแนนรวมของผู้สอบได้มาจากการตอบข้อสอบที่จัดว่าเป็นข้อที่ง่ายกว่าความสามารถของเขาได้ถูกต้องมากน้อยเพียงใด ดัชนีจะมีค่าสูงสุดเมื่อผู้สอบมีแบบแผนการตอบข้อสอบเป็นกัตแมนสมบูรณ์ กล่าวคือผู้สอบตอบข้อสอบทุกข้อที่จัดว่าง่ายกว่าและข้อที่มีความยากเท่ากับความสามารถของเขาได้ถูกต้อง และตอบข้อสอบทุกข้อที่ยากกว่าความสามารถของเขาผิด ผลการวิจัยพบว่าดัชนีที่พัฒนาขึ้น มีคุณภาพด้านความตรง (validity) ในการบ่งชี้ระดับความสอดคล้องของแบบแผนการตอบข้อสอบที่ผู้สอบทำได้กับที่มีลักษณะเป็นกัตแมนสมบูรณ์ และความตรงในการจำแนกแบบแผนการตอบข้อสอบที่มีลักษณะสอดคล้องกับกัตแมนสมบูรณ์สูงและต่ำ

---

\* วิทยานิพนธ์ระดับดุษฎีบัณฑิต สาขาการวัดและประเมินผลการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ปีการศึกษา 2535 โดยมี ผศ.ดร.ดิเรก ศรีสุขโข และ ผศ.ดร.ทวีวัฒน์ ปิตยานนท์ เป็นคณะอาจารย์ที่ปรึกษา

## ความเป็นมาและความสำคัญของปัญหา

แบบสอบเลือกตอบ (multiple choice test) เป็นแบบสอบที่มีการนำมาใช้กันอย่างกว้างขวาง ทั้งนี้เพราะมีจุดเด่นในด้านความเป็นปรนัย ความสะดวกและรวดเร็วในการตรวจให้คะแนน โดยทั่วไปแล้วคะแนนรวมของผู้สอบที่ได้มาจากการทำแบบสอบเลือกตอบจะคำนวณได้จากผลรวมของจำนวนข้อสอบที่ผู้สอบตอบได้ถูกต้อง ยกตัวอย่างเช่นแบบสอบเลือกตอบฉบับหนึ่งมีข้อสอบจำนวน 10 ข้อ ผู้สอบคนหนึ่งได้คะแนนรวม 3 คะแนน คะแนน 3 คะแนนนี้อาจจะได้มาจากการตอบข้อสอบข้อที่ 1 ข้อที่ 2 และข้อที่ 3 ได้ถูกต้อง หรืออาจจะได้มาจากการตอบข้อสอบข้อที่ 2 ข้อที่ 4 และข้อที่ 5 ได้ถูกต้อง หรืออาจจะได้มาจากการตอบข้อสอบข้อใด ๆ ก็ได้จำนวน 3 ข้อได้ถูกต้อง ซึ่งเป็นไปได้ถึง 120 แบบ

เมื่อพิจารณาถึงประเด็นที่ผู้สอบมีคะแนนรวมเท่ากัน แต่มีรูปแบบการตอบข้อสอบที่ถูกต้องแตกต่างกันไป ร่วมกับประเด็นที่กล่าวว่าผู้ที่ตอบข้อสอบแต่ละข้อได้ถูกต้อง มีได้ถึง 3 ประเภทคือผู้ที่รู้จริง ผู้ที่มีความรู้เพียงบางส่วน และผู้ที่ไม่มีความรู้แต่เดาถูก และผู้ที่ตอบข้อสอบผิดก็มีได้ 3 ประเภทเช่นกันคือผู้ที่ไม่มีความรู้ ผู้ที่พอจะมีความรู้บ้างแต่เลือกคำตอบผิด และผู้ที่รู้จริง แต่บังเอิญเลือกคำตอบผิด จึงสรุปได้ว่าคะแนนรวมที่ผู้สอบทำได้อาจจะมีค่ามากกว่าหรือน้อยกว่าความสามารถที่แท้จริงของเขา คำถามหนึ่งที่เกิดขึ้นกับผู้ที่ต้องนำคะแนนที่ได้จากแบบสอบเลือกตอบมาใช้คือคะแนนที่ได้มานั้นเป็นตัวแทนความสามารถของผู้สอบมากน้อยเพียงใด

แนวทางหนึ่งในการพิจารณาว่าคะแนนรวมที่ผู้สอบทำได้จะมีความเป็นตัวแทนความสามารถของเขาได้เพียงใดก็คือ การพิจารณาจากคำตอบของข้อสอบแต่ละข้อของผู้สอบ แล้วคำนวณเป็นค่าดัชนีที่เรียกว่าดัชนีความสอดคล้องของแบบแผนการตอบข้อสอบรายบุคคล ดัชนีในกลุ่มนี้เริ่มมาจากความเชื่อของกัตแมน (Guttman 1954) ว่าผู้สอบที่มีความสามารถระดับหนึ่งจะตอบข้อสอบทุกข้อที่จัดว่าเป็นข้อที่ง่ายกว่าและข้อที่มีความยากเท่ากับความสามารถของเขาได้ถูกต้อง และตอบข้อสอบทุกข้อที่จัดว่ายากกว่าความสามารถของเขาผิด และเมื่อนำคำตอบของข้อสอบแต่ละข้อของผู้สอบมาเรียงลำดับจากข้อที่มีจำนวนผู้สอบตอบถูกมากที่สุด ซึ่งจัดว่าเป็นข้อที่ง่ายที่สุด ไปจนถึงข้อที่มีจำนวนผู้สอบตอบถูกน้อยที่สุด ซึ่งจัดว่าเป็นข้อที่ยากที่สุด การเรียงลำดับของคำตอบตามลักษณะดังกล่าวนี้เรียกว่า *แบบแผนการตอบข้อสอบ* ดัชนีจะมีค่าสูงสุดเมื่อผู้สอบมีแบบแผนการตอบข้อสอบที่มีลักษณะเป็น *กัตแมนสมบูรณ์* กล่าวคือผู้สอบตอบข้อสอบลำดับที่ 1 ซึ่งเป็นข้อที่ง่ายที่สุด ไปจนถึงข้อที่มีลำดับที่เท่ากับคะแนนรวมของเขาได้ถูกต้องทุกข้อ และตอบข้อสอบที่มีลำดับที่มากกว่าคะแนนรวมของเขาเป็นต้นไปผิดทุกข้อ

ตัวอย่างเช่นผู้สอบที่ได้คะแนนรวม 3 คะแนนจากการตอบข้อสอบ 10 ข้อ เมื่อเรียงลำดับข้อสอบจากข้อที่ง่ายที่สุดไปจนถึงข้อที่ยากที่สุดแล้ว ผู้สอบจะมีแบบแผนการตอบข้อสอบเป็นกัณฑ์สมบูรณ์ เมื่อเขาตอบข้อสอบลำดับที่ 1 จนถึงลำดับที่ 3 ได้ถูกต้องทุกข้อ และตอบข้อสอบลำดับที่ 4 เป็นต้นไปผิดทุกข้อ

สำหรับดัชนีความสอดคล้องของแบบแผนการตอบข้อสอบรายบุคคลที่พัฒนาขึ้น ยังคงนำแนวคิดเรื่องกัณฑ์สมบูรณ์มาเป็นหลักในการสร้างดัชนี และได้นำคะแนนที่ผู้สอบได้จากการตอบข้อสอบแต่ละข้อที่เรียกว่าคะแนนรายข้อ คะแนนรวม และลำดับที่ของข้อสอบ มาเป็นองค์ประกอบของสูตรคำนวณ ทั้งนี้เพื่อให้ดัชนีที่พัฒนาขึ้นมีคุณสมบัติครอบคลุมจุดเด่นต่าง ๆ ของดัชนีที่มีอยู่แล้ว ได้แก่

1) มีช่วงพิสัยแน่นอน ไม่ว่าผู้สอบจะได้คะแนนรวมเท่าใดหรือแบบสอบจะประกอบด้วยข้อสอบจำนวนกี่ข้อก็ตาม (ดัชนีมีค่าอยู่ระหว่าง 0 ถึง 1.0 โดยที่ดัชนีจะมีค่าเท่ากับ 1.0 เมื่อผู้สอบมีแบบแผนการตอบข้อสอบเป็นกัณฑ์สมบูรณ์ และมีค่าเท่ากับ 0 เมื่อผู้สอบมีแบบแผนการตอบข้อสอบที่มีลักษณะตรงกันข้ามกับกัณฑ์สมบูรณ์ (Harnisch and Linn 1981 : 133-146)

2) ค่าดัชนีของผู้สอบแต่ละคน จะเป็นอิสระจากแบบแผนการตอบข้อสอบของผู้สอบคนอื่น ๆ ภายในกลุ่มถ้าลำดับที่ของข้อสอบยังคงเดิม ทั้งนี้เพราะดัชนีที่พัฒนาขึ้นได้นำลำดับที่ของข้อสอบมาใช้แทนค่าความยากของข้อสอบ (Tatsuoka and Tatsuoka 1982 : 215-231)

3) ค่าของดัชนีมีความไวต่อความคลาดเคลื่อนที่เกิดจากการที่ผู้สอบตอบข้อสอบที่จัดว่าเป็นข้อที่ง่ายกว่าความสามารถของเขาผิด และที่เกิดจากการที่ผู้สอบตอบข้อสอบที่จัดว่าเป็นข้อที่ยากกว่าความสามารถของเขาได้ถูกต้อง (Kane and Brennan 1980 : 105-126)

4) ค่าของดัชนีมีความหมายตรงกับแบบแผนการตอบข้อสอบทุกกรณี โดยเฉพาะในกรณีผู้สอบมีคะแนนรวมมากกว่าครึ่งหนึ่งของจำนวนข้อสอบทั้งหมด และคะแนนรวม  $X$  คะแนนนั้นได้มาจากการตอบข้อสอบจำนวน  $X$  ข้อที่นับจากข้อสอบลำดับสุดท้ายเข้ามาถึงข้อที่ยากน้อยลง ดัชนีจะมีค่ามากกว่า 0 ซึ่งเป็นค่าต่ำสุดของดัชนีเสมอ ยกตัวอย่างเช่น ผู้สอบมีคะแนนรวม 8 คะแนน และมีแบบแผนการตอบข้อสอบเป็น 0011111111 ที่กล่าวถึงจุดเด่นข้อนี้ก็เพราะดัชนีบางตัวจะมีค่าต่ำสุดถ้าผู้สอบมีแบบแผนการตอบข้อสอบตามลักษณะดังกล่าว (Sato 1975 : 166-168)

ตามที่กล่าวมาข้างต้นจะเห็นได้ว่า ดัชนีความสอดคล้องของแบบแผนการตอบข้อ-สอบรายบุคคลที่ผู้วิจัยพัฒนาขึ้นนี้เป็นดัชนีที่ได้รวบรวมเอาคุณลักษณะเด่นของดัชนีต่าง ๆ ที่มีอยู่ไว้ ทั้งนี้เพื่อให้ได้ดัชนีที่มีประสิทธิภาพในการนำมาใช้ตัดสินว่า คะแนนรวมที่ผู้สอบทำได้นั้นมีความเป็นตัวแทนความสามารถของเขาได้มากน้อยเพียงใด ยิ่งดัชนีมีค่ามากเท่าใดก็จะช่วยให้ผู้ใช้คะแนนรวมเพื่อตัดสินความสามารถของผู้สอบมีความมั่นใจในการตัดสินใจมากเท่านั้น และดัชนีจะยังมีประโยชน์มากในกรณีที่ต้องมีการคัดเลือกผู้สอบจากกลุ่มผู้สอบที่มีคะแนนรวมเท่ากัน ซึ่งคัดเลือกได้จากผู้ที่มีค่าดัชนีสูงสุดลงมา

### วัตถุประสงค์การวิจัย

1. เพื่อพัฒนาดัชนีความสอดคล้องของแบบแผนการตอบข้อสอบรายบุคคล
2. เพื่อตรวจสอบคุณภาพด้านความตรงของดัชนีที่พัฒนาขึ้นคือ
  - 2.1 ความตรง ในการบ่งชี้ความสอดคล้องของแบบแผนการตอบข้อสอบที่ผู้สอบทำได้กับที่มีลักษณะเป็นกัตแมนสมบูรณ์
  - 2.2 ความตรงในการจำแนกแบบแผนการตอบข้อสอบที่มีความสอดคล้องกับลักษณะกัตแมนสมบูรณ์สูงและต่ำ

### สมมติฐานการวิจัย

เนื่องจากดัชนีความสอดคล้องของแบบแผนการตอบข้อสอบรายบุคคลที่พัฒนาขึ้นยังคงยึดตามหลักการของกัตแมน และการนำองค์ประกอบด้านคะแนนรวม ลำดับที่ของข้อสอบ และคะแนนรายข้อของผู้สอบมาใช้ในสูตรคำนวณได้จัดกระทำอย่างมีหลักเหตุผล เพื่อให้ได้ดัชนีที่มีคุณลักษณะที่ครอบคลุมจุดเด่นของดัชนีต่าง ๆ ที่มีอยู่ ดังนั้นดัชนีที่พัฒนาขึ้นจึงน่าจะมีคุณภาพด้านความตรงเช่นเดียวกับดัชนีที่มีอยู่ก่อนแล้ว ผู้วิจัยจึงกำหนดสมมติฐานการวิจัยดังนี้คือ

**สมมติฐานที่ 1** ดัชนีที่พัฒนาขึ้นน่าจะมี ความตรงในการบ่งชี้ความสอดคล้องของแบบแผนการตอบข้อสอบที่ผู้สอบทำได้กับที่มีลักษณะเป็นกัตแมนสมบูรณ์

**สมมติฐานที่ 2** ดัชนีที่พัฒนาขึ้นน่าจะมี ความตรงในการจำแนกแบบแผนการตอบข้อสอบที่มีความสอดคล้องกับลักษณะกัตแมนสมบูรณ์สูงและต่ำ

### ขอบเขตการวิจัย

1. ดัชนีความสอดคล้องของแบบแผนการตอบข้อสอบรายบุคคลที่พัฒนาขึ้น หมายถึงดัชนีที่บ่งชี้ระดับความสอดคล้องของแบบแผนการตอบข้อสอบที่ผู้สอบทำได้ กับลักษณะกัตแมนสมบูรณ์ โดยการนำคะแนนรายข้อของผู้สอบมาเรียงลำดับจากข้อที่มีจำนวนผู้สอบภายใน

ในกลุ่มตอบได้ถูกต้องมากที่สุด ซึ่งจัดว่าเป็นข้อที่ง่ายที่สุดไปจนถึงข้อที่มีจำนวนผู้สอบภายในกลุ่มตอบได้ถูกต้องน้อยที่สุดซึ่งจัดว่าเป็นข้อที่ยากที่สุด

2. ประชากรที่ใช้ในการตรวจสอบคุณภาพของดัชนี เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 1 ของโรงเรียนคาทอลิกระดับมัธยมศึกษา อัครสังฆมณฑลกรุงเทพฯ ในเขตกรุงเทพมหานคร ในปีการศึกษา 2534 จำนวนทั้งสิ้น 1280 คน

### **วิธีดำเนินการวิจัย**

1. การพัฒนาดัชนีความสอดคล้องของแบบแผนการตอบข้อสอบรายบุคคล มีขั้นตอนดังนี้

*ขั้นที่ 1 การรวบรวมและวิเคราะห์แนวคิดของดัชนีที่มีอยู่* ขั้นตอนนี้ประกอบด้วยการศึกษาและรวบรวมรายละเอียดของดัชนีที่มีอยู่ แล้วนำมาแบ่งเป็น 2 กลุ่มคือกลุ่มที่มีแนวคิดตามกัตแมน และกลุ่มที่มีแนวคิดตามทฤษฎีการตอบสนองรายข้อ พร้อมทั้งวิเคราะห์จุดเด่นข้อจำกัดในการนำมาใช้ และข้อมูลยืนยันด้านความตรงของดัชนีทั้ง 2 กลุ่ม

*ขั้นที่ 2 การคัดเลือกแนวคิดและตัวกำหนดความสอดคล้อง* ขั้นตอนนี้ประกอบด้วยการตัดสินใจเลือกแนวคิดของดัชนีกลุ่มกัตแมน การวิเคราะห์จุดเด่นของดัชนีแต่ละตัวเพื่อนำมาสร้างเป็นคุณลักษณะที่พึงประสงค์ของดัชนีที่ดี แล้วนำคุณลักษณะที่พึงประสงค์มาเป็นเกณฑ์ในการประเมินจุดอ่อนของดัชนีที่มีอยู่ จากนั้นหาสาเหตุของจุดอ่อนดังกล่าว ผลการประเมินพบว่าดัชนีที่มีอยู่ล้วนมีจุดอ่อนทั้งสิ้น

*ขั้นที่ 3 การสร้างดัชนีความสอดคล้อง* ขั้นตอนนี้ประกอบด้วยการกำหนดนิยามโมเดลเชิงคณิตศาสตร์ และข้อตกลงเบื้องต้นของดัชนีที่พัฒนาขึ้น แล้วสร้างสูตรคำนวณขึ้น

2. การตรวจสอบคุณภาพด้านความตรงของดัชนีที่พัฒนาขึ้น

*กลุ่มตัวอย่าง* เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 1 ของโรงเรียนคาทอลิกระดับมัธยมศึกษา อัครสังฆมณฑลกรุงเทพฯ ในเขตกรุงเทพมหานคร ปีการศึกษา 2534 โดยสุ่มโรงเรียนประเภทชาย หญิง และสหศึกษา มาอย่างละ 1 โรงเรียน

#### *เครื่องมือ*

1) แบบสอบเลือกตอบที่มี 5 ตัวเลือก วัดเนื้อหาวิชาคณิตศาสตร์ ค 101 จำนวน 45 ข้อ ใช้เวลาในการสอบ 1½ ชั่วโมง

วิธีการตอบแบบสอบฉบับนี้ ใช้เทคนิคการวัดความรู้บางส่วนมาใช้วัดระดับความรู้ในการตอบข้อสอบแต่ละข้อของผู้สอบ โดยให้เลือกคำตอบที่คิดว่าถูกต้องเพียงคำตอบเดียวแล้วให้ระบุว่า “มั่นใจ” หรือ “ไม่มั่นใจ” ที่เลือกคำตอบข้อนั้น

2) แบบสอบถาม เกี่ยวกับเหตุผลที่ผู้สอบตอบข้อสอบแต่ละข้อผิด ซึ่งแบ่งเนื้อหาหลักเป็น 2 กรณี คือ ตอบผิดเพราะมีความรู้ไม่เพียงพอ กับตอบผิดเพราะความคลาดเคลื่อนของการวัด

เกณฑ์ที่นำมาตรวจสอบ การตรวจสอบความตรงในที่นี้ แบ่งเป็น 2 เกณฑ์คือ

1) เกณฑ์ที่แสดงความตรง ในการบ่งชี้ความสอดคล้องของแบบแผนการตอบข้อสอบที่ผู้สอบทำได้กับที่มีลักษณะกัตแมนสมบูรณ์ ซึ่งพิจารณาจากค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างค่าดัชนี กับค่าสัดส่วนของจำนวนข้อสอบที่ผู้สอบตอบได้ถูกต้องด้วยความมั่นใจต่อจำนวนข้อสอบทั้งหมดที่ตอบได้ถูกต้อง และกับค่าสัดส่วนของจำนวนข้อสอบที่ผู้สอบตอบผิดเพราะมีความรู้ไม่เพียงพอต่อจำนวนข้อสอบทั้งหมดที่ตอบผิด

2) เกณฑ์ที่แสดงความตรงในการจำแนกแบบแผนการตอบข้อสอบที่มีลักษณะสอดคล้องกับกัตแมนสมบูรณ์สูงและต่ำ ซึ่งพิจารณาจากผลต่างของค่าเฉลี่ยของดัชนีที่คำนวณได้จากกลุ่มผู้สอบที่มีแบบแผนการตอบข้อสอบทั้ง 2 ประเภท โดยใช้คะแนนรวม ความมั่นใจในการตอบข้อสอบ และค่าความสามารถตามแนวทฤษฎีการตอบสนองรายข้อตามโมเดลโลจิสแบบ 3 พารามิเตอร์ มาใช้ในการจำแนกกลุ่มผู้สอบที่มีแบบแผนการตอบข้อสอบที่มีความสอดคล้องกับกัตแมนสมบูรณ์สูงและต่ำ

### ผลการวิจัย

1. **ดัชนีความสอดคล้องของแบบแผนการตอบข้อสอบรายบุคคลที่พัฒนาขึ้น**  
 ดัชนีที่พัฒนาขึ้นเรียกว่า ดัชนีเอพี (AP Index) เพราะเป็นดัชนีที่บ่งชี้ว่าแบบแผนการตอบข้อสอบที่เกิดขึ้นจริง (actual pattern) ของผู้สอบมีลักษณะที่ใกล้เคียงกับกัตแมนสมบูรณ์มากน้อยเพียงใด โดยนำสัดส่วนของผลรวมของค่าประจำข้อของคะแนนรายข้อที่มีความคลาดเคลื่อนต่อผลรวมของค่าประจำข้อของคะแนนรายข้อของข้อสอบทุกข้อ ไปลบออกจาก 1.0 ดังสูตรคำนวณต่อไปนี้

$$AP_i = 1 - \frac{\sum_{j=1}^{n_i} (1 - U_{ij})(n_i - j + 1) + \sum_{j=n_i+1}^J U_{ij}(j - n_i)}{\sum_{j=1}^{n_i} (n_i - j + 1) + \sum_{j=J-n_i+1}^J (j - n_i)}$$

โดยที่  $AP_i$  คือ ดัชนีเอพีของผู้สอบคนที่  $i$   
 $j$  คือ ลำดับที่ของข้อสอบ ;

$j = 1, 2, 3, \dots, J$  เมื่อข้อสอบลำดับที่ 1 เป็นข้อที่ง่ายที่สุด และลำดับที่  $J$  ยากที่สุด

$n_i$  คือ คะแนนรวมของผู้สอบคนที่  $i$

$U_{ij}$  คือ คะแนนรายข้อของข้อสอบลำดับที่  $j$  ของ

ผู้สอบคนที่  $i$  โดยที่  $U_{ij}$  มีค่าเท่ากับ 1 เมื่อผู้สอบตอบถูก และเท่ากับ 0 เมื่อผู้สอบตอบผิด

$$\sum_{j=1}^{n_i} (1 - U_{ij})(n_i - j + 1) \text{ คือ ผลรวมของค่าประจำข้อของคะแนนรายข้อที่มีความคลาด}$$

เคลื่อนประเภทที่ 1

$$\sum_{j=n_i+1}^J U_{ij}(j - n_i) \text{ คือ ผลรวมของค่าประจำข้อของคะแนนรายข้อที่มีความคลาด-}$$

เคลื่อนประเภทที่ 2

$$\sum_{j=1}^{n_i} (n_i - j + 1) + \sum_{j=J-n_i+1}^J (j - n_i) \text{ คือ ผลรวมของค่าประจำข้อของคะแนนราย-}$$

ข้อของข้อสอบทุกข้อ

พิสัยของดัชนีเอพี มีค่าอยู่ระหว่าง 0 ถึง 1.0 โดยที่ดัชนีนี้มีค่าเท่ากับ 0 เมื่อผู้สอบมีแบบแผนการตอบข้อสอบที่มีลักษณะตรงกันข้ามกััดแมนสมบรูณ์ และมีค่าเท่ากับ 1.0 เมื่อผู้สอบมีแบบแผนการตอบข้อสอบเป็นกััดแมนสมบรูณ์

ความคงที่ (stability) ของค่าดัชนีนี้ ผลจากการตรวจสอบด้วยข้อมูลจำลอง (simulate data) โดยเริ่มต้นจากข้อสอบ 5 ข้อ ซึ่งมีแบบแผนการตอบข้อสอบที่เป็นไปได้ทั้งหมด 32 แบบ จากนั้นจำลองแบบแผนการตอบข้อสอบทั้ง 32 แบบ ให้มีจำนวนข้อสอบเพิ่มขึ้นเป็น 10 ข้อ 20 ข้อ 30 ข้อ ไปจนถึง 120 ข้อ ผลการตรวจสอบพบว่า

- 1) ค่าของดัชนีเอพี เป็นอิสระจากคะแนนรวมที่ผู้สอบทำได้
- 2) ไม่ว่าข้อสอบจะมีจำนวนเท่าใดก็ตาม ดัชนีจะมีค่าเท่ากับ 1.0 เสมอ เมื่อผู้สอบมีแบบแผนการตอบข้อสอบเป็นกััดแมนสมบรูณ์ และมีค่าเท่ากับ 0 เสมอ เมื่อผู้สอบมีแบบแผนการตอบข้อสอบที่ตรงกันข้ามกััดแมนสมบรูณ์
- 3) สำหรับแบบแผนการตอบข้อสอบที่มีความคลาดเคลื่อน แต่ไม่รวมแบบแผนการตอบข้อสอบที่มีลักษณะที่ตรงกันข้ามกััดแมนสมบรูณ์ ดัชนีจะมีค่าคงที่เมื่อข้อสอบมีจำนวนตั้งแต่ 30 ข้อขึ้นไป

## 2. ผลการตรวจสอบคุณภาพด้านความตรงของดัชนีที่พัฒนาขึ้น

2.1 ความตรง ในการบ่งชี้ความสอดคล้องของแบบแผนการตอบข้อสอบที่ผู้สอบทำได้ กับที่มีลักษณะกีดแมนสมบูรณ์

ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างค่าดัชนีเอพี (X) กับเกณฑ์ทั้ง 2 คือ ค่าสัดส่วนของจำนวนข้อสอบที่ผู้สอบตอบถูกต้องด้วยความมั่นใจ ต่อจำนวนข้อสอบทั้งหมดที่ตอบถูก ( $y_1$ ) และค่าสัดส่วนของจำนวนข้อสอบที่ผู้สอบตอบผิดเพราะมีความรู้ไม่เพียงพอต่อจำนวนข้อสอบทั้งหมดที่ตอบผิด ( $y_2$ ) มีค่าตั้งแต่ 0.5809 ถึง 0.8796 โดยที่ทุกค่ามีนัยสำคัญทางสถิติ ดังที่ปรากฏตามตารางที่ 1

ตารางที่ 1 ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างค่าของดัชนีเอพีกับเกณฑ์

ประเภทโรงเรียน (จำนวนนักเรียน)	$r_{xy_1}$	$r_{xy_2}$
หญิง (82 คน)	.7742***	.7701***
ชาย (59 คน)	.8431***	.8796***
สหศึกษา (126 คน)	.7775***	.6531***
รวม (267 คน)	.7045***	.5809***

\*\*\*  $P < .001$

2.2 ความตรง ในการจำแนกแบบแผนการตอบข้อสอบที่มีลักษณะสอดคล้อง กับกีดแมนสมบูรณ์สูงและต่ำ

จากการทดสอบความแตกต่างของค่าเฉลี่ยของดัชนีเอพีที่ได้จากกลุ่มผู้สอบที่มีแบบแผนการตอบข้อสอบที่มีลักษณะสอดคล้องกับกีดแมนสมบูรณ์สูงและต่ำ สรุปได้ว่าโดยภาพรวมแล้วดัชนีเอพีมีความตรงในการจำแนกแบบแผนการตอบข้อสอบที่มีลักษณะสอดคล้อง กับกีดแมนสมบูรณ์สูงและต่ำ ดังรายละเอียดตามตารางที่ 2 คือ

1) เมื่อใช้คะแนนรวมในการแบ่งกลุ่มผู้สอบ ค่าเฉลี่ยของดัชนีเอพีที่ได้กลุ่มผู้สอบที่มีแบบแผนการตอบข้อสอบที่สอดคล้องกับกีดแมนสมบูรณ์สูง จะมีค่ามากกว่ากลุ่มผู้สอบที่มีแบบแผนการตอบข้อสอบที่สอดคล้องกับกีดแมนสมบูรณ์ต่ำอย่างมีนัยสำคัญ ในประเภทโรงเรียนชายและสหศึกษา



2) เมื่อใช้ความมั่นใจในการตอบข้อสอบในการแบ่งกลุ่มผู้สอบ ค่าเฉลี่ยของดัชนีเอพีที่ได้กลุ่มผู้สอบที่มีแบบแผนการตอบข้อสอบที่สอดคล้องกับกัตแมนสมบูรณ์สูง จะมีค่ามากกว่ากลุ่มผู้สอบที่มีแบบแผนการตอบข้อสอบที่สอดคล้องกับกัตแมนสมบูรณ์ต่ำอย่างมีนัยสำคัญ ในประเภทโรงเรียนหญิง และชาย

3) เมื่อใช้ค่าความสามารถตามแนวทฤษฎีการตอบสนองรายข้อตามโมเดลโลจิสแบบ 3 พารามิเตอร์ในการแบ่งกลุ่มผู้สอบ ค่าเฉลี่ยของดัชนีเอพีที่ได้กลุ่มผู้สอบที่มีแบบแผนการตอบข้อสอบที่สอดคล้องกับกัตแมนสมบูรณ์สูง จะมีค่ามากกว่ากลุ่มผู้สอบที่มีแบบแผนการตอบข้อสอบที่สอดคล้องกับกัตแมนสมบูรณ์ต่ำอย่างมีนัยสำคัญ ในทุกประเภทของโรงเรียน

ตารางที่ 2 เปรียบเทียบค่าเฉลี่ยของดัชนีเอพี ที่ได้จากกลุ่มผู้สอบที่มีแบบแผนการตอบข้อสอบที่มีลักษณะสอดคล้องกับกัตแมนสมบูรณ์สูงและต่ำ

ประเภท โรงเรียน	กลุ่มผู้สอบ	เกณฑ์ที่ใช้จำแนกกลุ่มผู้สอบ					
		คะแนนรวม		ความมั่นใจในการตอบ		ค่าความสามารถ	
		$\bar{X}$ (SD)	t	$\bar{X}$ (SD)	t	$\bar{X}$ (SD)	t
หญิง	สอดคล้องสูง	.52 (.09)	.09	.62 (.07)	6.97***	.66 (.08)	2.83**
	สอดคล้องต่ำ	.48 (.13)		.43 (.09)		.58 (.09)	
ชาย	สอดคล้องสูง	.67 (.14)	2.51*	.78 (.12)	8.66***	.74 (.10)	4.51***
	สอดคล้องต่ำ	.55 (.14)		.45 (.08)		.59 (.07)	
สห ศึกษา	สอดคล้องสูง	.76 (.08)	3.67**	.82 (.08)	7.25***	.76 (.11)	0.87
	สอดคล้องต่ำ	.65 (.15)		.61 (.14)		.74 (.09)	
รวม	สอดคล้องสูง	.65 (.16)	1.63	.76 (.12)	12.71***	.71 (.13)	0.60
	สอดคล้องต่ำ	.60 (.15)		.48 (.12)		.70 (.10)	

\*  $P < .05$

\*\*  $P < .01$

\*\*\*  $p < .001$

**อภิปรายผลการวิจัย**

ดัชนีเอพีที่ผู้วิจัยพัฒนาขึ้นนับว่าเป็นดัชนีที่มีคุณสมบัติครอบคลุมคุณลักษณะที่พึงประสงค์ของดัชนีความสอดคล้องของแบบแผนการตอบข้อสอบรายบุคคล และยังมีคุณภาพด้านความคงที่ของค่าดัชนีซึ่งเป็นคุณสมบัติสำคัญที่ White and Saltz (White and Saltz 1957 : 82) ได้กล่าวถึงไว้ รวมทั้งมีความตรงในการบ่งชี้ความสอดคล้องของแบบแผนการตอบข้อสอบที่ผู้สอบทำได้กับที่มีลักษณะกัตแมนสมบูรณ์ ซึ่งเป็นข้อมูลสำคัญในการตัดสินใจนำดัชนีเอพีไปใช้ประกอบการประเมินความสามารถของผู้สอบร่วมกับคะแนนรวมที่เขาทำได้ (Harnisch 1981 : 6) และ ความตรงในการจำแนกแบบแผนการตอบข้อสอบที่มีลักษณะสอดคล้องกับกัตแมนสมบูรณ์สูงและต่ำ ซึ่งเป็นข้อมูลที่น่ามาใช้กำหนดค่าจุดตัดของดัชนีที่ใช้จำแนกว่าแบบแผนการตอบข้อสอบที่ผู้สอบทำได้มีลักษณะที่สอดคล้องกับกัตแมนสมบูรณ์สูงหรือต่ำ ถ้ามีความสอดคล้องสูงก็แสดงว่าคะแนนรวมที่ผู้สอบทำได้มีความเป็นตัวแทนความสามารถของเขา แต่ถ้ามีความสอดคล้องต่ำก็เป็นข้อเตือนใจให้ผู้ใช้คะแนนรวมในการประเมินความสามารถของผู้สอบคนนั้นใช้ด้วยความระมัดระวัง หรืออาจต้องมาวิเคราะห์ถึงความบกพร่องในการเรียนรู้ของผู้เรียนคนนั้น

**ข้อเสนอแนะ**

ข้อเสนอแนะในการนำดัชนีเอพีไปใช้

1. เพื่อตัดสินใจความถูกต้องของคะแนนรวมที่ผู้สอบทำได้ ถ้าดัชนีมีค่ามากกว่า 0.7 แสดงว่าคะแนนรวมของผู้สอบมีความเป็นตัวแทนความสามารถของเขาได้
2. เพื่อคัดเลือกผู้สอบจากกลุ่มผู้สอบที่มีคะแนนรวมเท่ากัน โดยเลือกจากผู้ที่มีค่าดัชนีเอพีสูงสุด ในกรณีที่ต้องคัดเลือกผู้สอบที่มีคะแนนรวมเท่ากันและยังมีค่าดัชนีเอพีเท่ากันสามารถทำได้โดยพิจารณาจากดัชนีย่อยในสูตรคำนวณของดัชนีเอพี ซึ่งผู้วิจัยเรียกว่าดัชนีความคลาดเคลื่อนของการตอบข้อง่ายผิด ( $AP_{E1.E}$ ) โดยเลือกจากผู้ที่มีค่าดัชนีสูงสุดเช่นกัน

$$AP_{E1.E} = \frac{\sum_{j=1}^{n_i} (1 - U_{ij})(n_i - j + 1)}{\sum_{j=1}^{n_i} (1 - U_{ij})(n_i - j + 1) + \sum_{j=n_i+1}^J U_{ij} (j - n_i)}$$

โดยที่  $AP_{E1.E}$  คือ ดัชนีความคลาดเคลื่อนของการตอบข้อง่ายผิดของผู้สอบคนที่  $i$   
 $j$  คือ ลำดับที่ของข้อสอบ ;  $j = 1, 2, 3, \dots, J$  เมื่อข้อสอบลำดับที่ 1 เป็นข้อที่ง่ายที่สุด  
 และลำดับที่  $J$  ยากที่สุด

$n_i$  คือ คะแนนรวมของผู้สอบคนที่  $i$

$U_{ij}$  คือ คะแนนรายข้อของข้อสอบลำดับที่  $j$  ของผู้สอบคนที่  $i$  โดยที่  $U_{ij}$  มีค่าเท่ากับ  
 1 เมื่อผู้สอบตอบถูก และเท่ากับ 0 เมื่อผู้สอบตอบผิด

$$\sum_{j=1}^{n_i} (1 - U_{ij})(n_i - j + 1) \text{ คือ ผลรวมของค่าประจำข้อของคะแนนรายข้อที่มีความคลาด}$$

เคลื่อนประเภทที่ 1

$$\sum_{j=n_i+1}^J U_{ij}(j - n_i) \text{ คือ ผลรวมของค่าประจำข้อของคะแนนรายข้อที่มีความคลาดเคลื่อน}$$

ประเภทที่ 2

*ข้อเสนอแนะในการวิจัยต่อไป*

1. น่าจะมีการนำดัชนีเอพีที่พัฒนาขึ้นนี้ ไปศึกษาถึงลักษณะของการแจกแจงของค่า  
 ดัชนี เมื่อเงื่อนไขของจำนวนข้อสอบ ค่าความยากของข้อสอบ และความสามารถของกลุ่มผู้สอบ  
 แตกต่างกัน

2. น่าจะมีการนำดัชนีเอพีที่พัฒนาขึ้นไปใช้กับแบบสอบวิชาอื่น ๆ และนำไปใช้เพื่อ  
 วินิจฉัยความแตกต่างของการจัดการเรียนการสอน ความแตกต่างของภูมิหลังของผู้เรียน และ  
 จัดกลุ่มความคลาดเคลื่อนของการทำข้อสอบ

3. น่าจะมีงานวิจัยที่นำเสนอวิธีการจัดลำดับที่ของข้อสอบวิธีอื่น ๆ บ้าง โดยทั่วไปดัชนี  
 ในกลุ่มนี้จะจัดเรียงลำดับข้อสอบจากข้อที่ง่ายที่สุดไปจนถึงข้อที่ยากที่สุด ยังมีจำนวนข้อสอบ  
 มากเท่าใดก็ยังมีข้อสอบที่มีจำนวนผู้สอบภายในกลุ่มตอบได้ถูกต้องเท่า ๆ กันมากเท่านั้น ซึ่งจะ  
 เป็นปัญหาต่อการเรียงลำดับข้อสอบ

**บรรณานุกรม**

- Donlon, T.E. and Fischer, F.E. "An Index of an Individual's Agreement with Group - Determined Item Difficulties." *Educational and Psychological Measurement*. 28 (1968) : 105-113.
- Harnisch, D.L. "Analysis of Item Response Patterns : Consistency Indices and Their Application to Criterion-Referenced Tests." *Paper presented at the Annual Meeting of the American Educational Research Association*. Los Angeles, April 1981.
- Harnisch, D.L. and Linn, R.L. "Analysis of Item Response Patterns : Questionable Test Data and Dissimilar Curriculum Practices." *Journal of Educational Measurement*. 18 (Fall 1981) : 133-146.
- Harnisch, D.L. and Linn, R.L. "Identification of Aberrant Response Patterns." (Final Report on Grant No. G-80-0003.) Washington, D.C. : National Institute of Education, 1981b.
- Kane, M.T. and Brennan, R.L. "Agreement Coefficients as Indices of Dependability for Domain-Referenced Tests." *Applied Psychological Measurement*. 4 (Winter 1980) : 105-126.
- Tatsuoka, M.M. and Tatsuoka, K.K. "Detection of Aberrant Response Patterns and their Effect on Dimensionality." *Journal of Educational Statistics*. 7 (1982) : 215-231.
- Van Der Flier, H. "Environmental Factors and Deviant Response Patterns." In Y.H. Poortinga (Ed.) *Basic Problems in Cross-Cultural Psychology*. Amsterdam : Swets and Zeitlinger, B.V., 1977.
- White, B.W. and Saltz, E. "Measurement of Reproducibility." *Psychological Bulletin*. 54 (March 1957) : 81-99.