

การตรวจสอบการทุจริตในการทดสอบด้วย DGIRTМ 2 พารามิเตอร์

Examination Cheating Detection with 2 Parameters DGIRTМ

อัมพล ชูสนุก (Ampon Shoosanuk)^{*}

ไพรัตน์ วงษ์นาม (Pairatana Wongnam)^{**}

สมพงษ์ ปั่นหูน (Sompong Panhoon)^{***}

บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์ (1) เพื่อเปรียบเทียบอัตราความคลาดเคลื่อนประเภทที่ 1 และ (2) เพื่อเปรียบเทียบอำนาจการทดสอบของการตรวจสอบการทุจริตในการทดสอบด้วย DGIRTМ 2 พารามิเตอร์ สำหรับข้อสอบแบบ ตอบถูกได้ 1 คะแนน ตอบผิดได้ 0 คะแนน การจำลองข้อมูลดังกล่าวจำลองข้อมูลผลการตอบข้อสอบภายใต้ปัจจัยแปรเปลี่ยน 6 ปัจจัย คือ ความยาวของแบบทดสอบ 2 ขนาด ระดับความแตกต่างของการแจกแจงความยากรายข้อของข้อสอบ 2 ระดับ อำนาจจำแนกรายข้อของข้อสอบ 2 ระดับ ขนาดของการทำทุจริต 2 ระดับ ขนาดของการรั่วไหลของข้อสอบ 2 ระดับ และประสิทธิผลในการทำทุจริต 3 ระดับ รวมข้อมูลทั้งหมดที่ต้องจัดกระทำเพื่อตรวจสอบการทุจริตในการทดสอบจำนวน 96 เงื่อนไข ($2 \times 2 \times 2 \times 2 \times 2 \times 3$) ในแต่ละเงื่อนไขจำลองข้อมูลวนซ้ำ 100 รอบ

ผลการวิจัย พบว่า

1. อัตราการควบคุมความคลาดเคลื่อนประเภทที่ 1 ของการตรวจสอบการทุจริตในการทดสอบด้วย DGIRTМ 2 พารามิเตอร์ ภายใต้ปัจจัยที่แตกต่างกัน 6 ปัจจัย พบว่า ทุกเงื่อนไขไม่สามารถควบคุมอัตราความคลาดเคลื่อนประเภทที่ 1 ได้ตามเกณฑ์ที่กำหนด รายละเอียดผลการวิจัย พบว่า (1) ความยาวข้อสอบมีอิทธิพลต่ออัตราความคลาดเคลื่อนประเภทที่ 1 (2) ระดับความแตกต่างของการแจกแจงความยากรายข้อของข้อสอบมีอิทธิพลต่ออัตราความคลาดเคลื่อนประเภทที่ 1 (3) ระดับอำนาจจำแนกรายข้อของข้อสอบมีอิทธิพลต่ออัตรา

^{*} นักศึกษาปริญญาเอก, สาขาวิชาวิจัย วัดผลและสถิติการศึกษา, ภาควิชาวิจัยและจิตวิทยาประยุกต์, คณะศึกษาศาสตร์, มหาวิทยาลัยบูรพา, amponsh@gmail.com

Ampon Shoosanuk, Ph.D. Student, Program in Educational Research, Measurement, and Statistics, Department of Research and Applied Psychology, Faculty of Education, Burapha University.

^{**} รองศาสตราจารย์ ดร., สาขาวิชาวิจัย วัดผลและสถิติการศึกษา, ภาควิชาวิจัยและจิตวิทยาประยุกต์, คณะศึกษาศาสตร์, มหาวิทยาลัยบูรพา

Associate Professor Pairatana Wongnam, Ph.D., Program in Educational Research, Measurement, and Statistics, Department of Research and Applied Psychology, Faculty of Education, Burapha University.

^{***} อาจารย์ ดร., สาขาวิชาวิจัย วัดผลและสถิติการศึกษา, ภาควิชาวิจัยและจิตวิทยาประยุกต์, คณะศึกษาศาสตร์, มหาวิทยาลัยบูรพา

Sompong Panhoon, Ph.D., Program in Educational Research, Measurement, and Statistics, Department of Research and Applied Psychology, Faculty of Education, Burapha University.

ความคลาดเคลื่อนประเภทที่ 1 (4) ขนาดของการทำทุจริตมีอิทธิพลต่ออัตราความคลาดเคลื่อนประเภทที่ 1 (5) ขนาดของการรั่วไหลของข้อสอบมีอิทธิพลต่ออัตราความคลาดเคลื่อนประเภทที่ 1 และ (6) ประสิทธิภาพในการทำทุจริตมีอิทธิพลต่ออัตราความคลาดเคลื่อนประเภทที่ 1

2. อำนาจการทดสอบของการตรวจสอบการทุจริตในการทดสอบด้วย DGIRTM 2 พารามิเตอร์ภายใต้ปัจจัยที่แตกต่างกัน 6 ปัจจัย พบว่า ที่ระดับประสิทธิภาพในการทำทุจริตต่ำ มีอำนาจการทดสอบต่ำกว่าเกณฑ์ที่กำหนด ส่วนที่ระดับประสิทธิภาพในการทำทุจริตปานกลาง และระดับประสิทธิภาพในการทำทุจริตสูงมีอำนาจการทดสอบสูงกว่าเกณฑ์ที่กำหนด รายละเอียดผลการวิจัย พบว่า (1) ความยาวข้อสอบมีอิทธิพลต่ออำนาจการทดสอบ (2) ระดับความแตกต่างของการแจกแจงความยากง่ายของข้อสอบมีอิทธิพลต่ออำนาจการทดสอบ (3) ระดับอำนาจจำแนกรายข้อของข้อสอบมีอิทธิพลต่ออำนาจการทดสอบ (4) ขนาดของการทำทุจริตมีอิทธิพลต่ออำนาจการทดสอบ (5) ขนาดของการรั่วไหลของข้อสอบมีอิทธิพลต่ออำนาจการทดสอบ และ (6) ประสิทธิภาพในการทำทุจริตมีอิทธิพลต่ออำนาจการทดสอบ

คำสำคัญ: การทุจริตในการทดสอบ, ข้อสอบรั่วไหล, ข้อสอบที่ถูกเรียนรู้ออกก่อนการทดสอบ

Abstract

The purposes of this research were (1) to compare type I error rates and (2) to compare the power of test of examination cheating detection with 2 parameters DGIRTM for dichotomous scored items. In this study, the data was simulated under 6-factor variant, two levels forms of test length, two levels of distribution difference of item difficulty, two levels of item discrimination, two levels of cheating size, two levels of exposed item, and three levels of effective cheater. A total 96 ($2 \times 2 \times 2 \times 2 \times 2 \times 3$) conditions were studied. The data was replicated 100 times for each condition.

It was found that:

1. Type I error rates of examination cheating detection with 2 parameters DGIRTM under 6-factor variant could not control for specified standard, details of the result were; (1) test length had influenced on type I error rate; (2) levels of distribution difference of item difficulty had influenced on type I error rate; (3) level of item discrimination had influenced on type I error rate; (4) levels of cheating size had influenced on type I error rate; (5) levels of exposed item had influenced on type I error rate; and (6) levels of effective cheater had influenced on type I error rate

2. Power of test of examination cheating detection with 2 parameters DGIRTM under 6-factor variant were lower than specified standard in the condition of low effective cheater and higher than specified standard in the condition of medium and high effective cheater,

details of the result were; (1) test length had influenced on power of test; (2) levels of distribution difference of item difficulty had influenced on power of test; (3) level of item discrimination had influenced on power of test; (4) levels of cheating size had influenced on power of test; (5) levels of exposed item had influenced on power of test; and (6) levels of effective cheater had influenced on power of test.

Keywords: Examination Cheating, Item Exposure, Item Pre-knowledge

บทนำ

การทดสอบถูกนำมาใช้เป็นเครื่องมือในการประเมินผล ช่วยในการตัดสินใจต่าง ๆ ผลการทดสอบโดยส่วนใหญ่ถูกนำมาใช้เพื่ออ้างอิงถึงความชำนาญ ความสำเร็จ ความสามารถ และความรู้ของผู้สอบ (Cizek, 1999; Deng, 2007) ดังนั้นการทดสอบจึงถูกใช้อย่างกว้างขวางสำหรับการประเมินผลการศึกษา การคัดเลือกบุคลากร การให้ใบอนุญาต การให้การรับรองวิชาชีพ และการประเมินผลโครงการ (Deng, 2007) คะแนนที่ผู้เข้าสอบได้รับจากการทดสอบที่มีผลกระทบสูงมีอิทธิพลอย่างมากต่อการตัดสินใจ รวมถึงการตัดสินใจเกี่ยวกับอนาคตซึ่งส่งผลกระทบต่อบุคคลที่เข้ารับการทดสอบ และสามารถเปลี่ยนแปลงชีวิตของผู้สอบ (Deng, 2007; Hing-fai, 2008; McLeod, 1998)

การทุจริตเกิดขึ้นในทุกระดับของการศึกษา (McManus, Lissauer, & Williams, 2005, p. 1064; Nowell & Laufer, 1997) หนึ่งในที่มาของการทุจริตในการทดสอบ คือ การรั่วไหลของข้อสอบ (Shu, Henson, & Luecht, 2013, p. 481) การทุจริตจึงเป็นพฤติกรรมที่ผู้สอบได้คำตอบในการทดสอบที่ถูกต้องมาอย่างไม่ยุติธรรม ซึ่งโดยที่จริงแล้วผู้สอบนี้ไม่มีความสามารถในการตอบข้อสอบได้ถูกต้อง (Karabatsos, 2003, p. 278) การทุจริตที่เกิดขึ้นสามารถเกิดขึ้นจากตัวผู้สอบ ผู้บริหารจัดการทดสอบ และบุคคลที่คุมสอบ (Cizek, 2001) และการทุจริตในการทดสอบที่เกิดขึ้นนั้นคุกคามต่อความตรงของผลลัพธ์ที่ได้จากการทดสอบ (McManus et al., 2005, p. 1064) ทำให้ผลการทดสอบไม่ถูกต้อง เนื่องจากผู้สอบที่มีความสามารถต่ำ หากกระทำทุจริต ทำให้สามารถผ่านการทดสอบไปได้ด้วยคะแนนที่เพิ่มขึ้นจากการทำทุจริต จะส่งผลกระทบต่อความตรงของการทดสอบอย่างมาก และในทางปฏิบัติยังส่งผลกระทบต่อระบบการศึกษา สถานศึกษาที่คนเหล่านี้เข้าไปศึกษา รวมถึงองค์การที่บุคคลเหล่านี้เข้าปฏิบัติงานหลังจากผ่านการทดสอบ เนื่องจากบุคคลเหล่านี้ไม่มีความสามารถที่แท้จริงจึงอาจส่งผลทำให้ไม่สามารถศึกษาได้สำเร็จ หรือไม่สามารถปฏิบัติหน้าที่ที่ได้รับการบรรจุได้อย่างมีประสิทธิภาพ ดังนั้นการทุจริตในการทดสอบจึงเป็นการบ่อนทำลายความน่าเชื่อถือ และความเป็นธรรมในทุกๆระดับ อันจะนำไปสู่ความอ่อนแอ ทำลายพื้นฐานคุณความดีของสังคม การทุจริตจึงเป็นปัญหาที่ควรตระหนักสำหรับพลเมืองทุกคนของประเทศ (Cole, 1998)

การรั่วไหลของข้อสอบเกิดจากการทุจริตของผู้มีส่วนเกี่ยวข้องกับข้อสอบ ตัวอย่างที่เกิดในประเทศไทย เช่น การทดสอบเพื่อบรรจุ และแต่งตั้งบุคคลเข้ารับราชการเป็นข้าราชการครู และบุคลากรทางการศึกษา ตำแหน่งครูผู้ช่วย ประจำปี 2558 ของสำนักงานเขตพื้นที่การศึกษาประถมศึกษา (สพป.) เลย เขต 2 กรมการ

คุมสอบได้ตรวจพบผู้สอบพกโพลีคาบเข้าห้องสอบ รวม 6 คน (กรุงเทพฯธุรกิจออนไลน์, 2558) นอกจากการทุจริตที่เกิดขึ้นในการทดสอบเพื่อคัดเลือกบุคลากรครู ในประเทศไทยยังพบการทุจริตในการทดสอบบุคลากรด้านอื่น ๆ อีก เช่น การทุจริตในการสอบเข้าโรงเรียนนายสิบตำรวจ ที่ทำกันเป็นขบวนการใหญ่ ครอบคลุมสนามสอบทั้ง 4 ภาคทั่วประเทศ โดยจับกุมผู้ทุจริตเข้าสอบได้ถึง 234 คน พร้อมกับแก๊งโกงข้อสอบ 5 คน ในเบื้องต้นยึดเงินสดได้ 10.5 ล้านบาท ซึ่งเป็นเพียงมัดจำที่เรียกจากผู้สอบล่วงหน้า เพียงคนละ 1-3 หมื่นบาทเท่านั้น หากสอบได้จะต้องเสียค่าใช้จ่ายอีกหัวละ 3.5-5 แสนบาท (คมชัดลึกออนไลน์, 2555) การรั่วไหลของข้อสอบเหล่านี้ส่งผลกระทบต่อระบบราชการ สังคม และประเทศชาติในระยะยาว

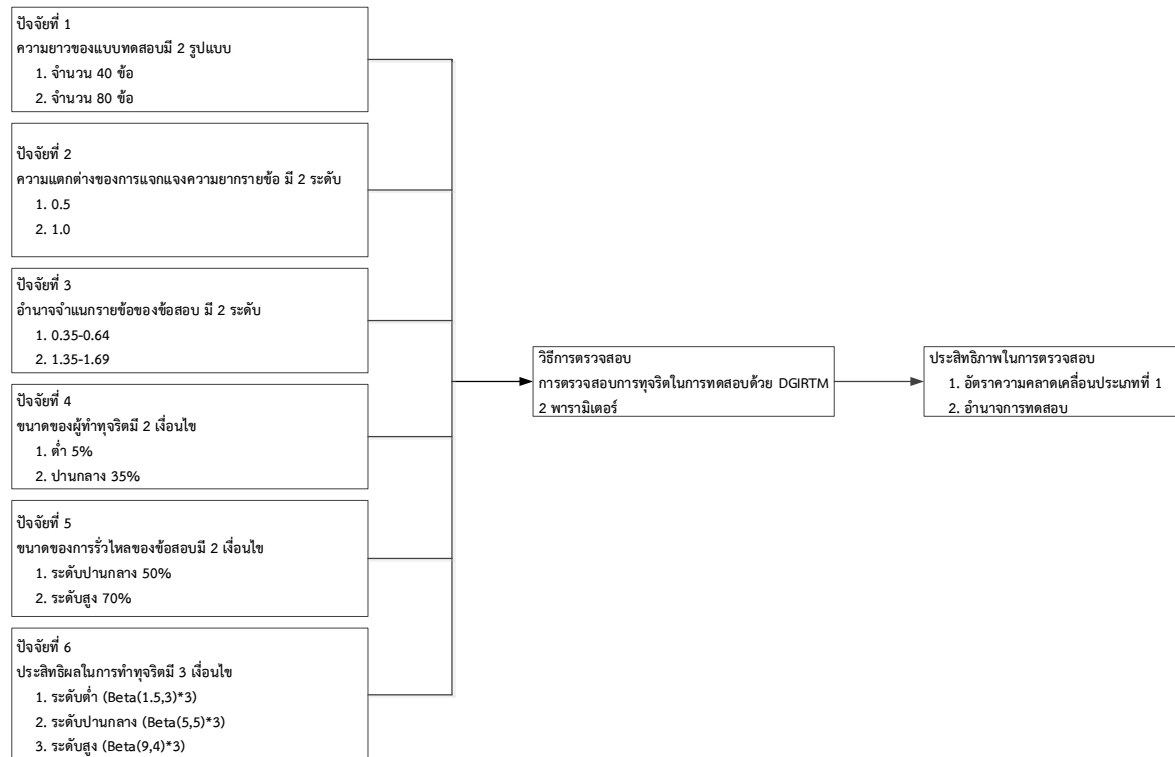
วิธีการตรวจสอบการทุจริตในรูปแบบที่แตกต่างกันจากการทดสอบ และการตรวจสอบการทุจริตได้รับการพัฒนา และการให้ความสำคัญต่อการพัฒนาวิธีการใหม่ ๆ ในการตรวจสอบการทุจริตควรถูกกระทำอย่างเร่งด่วน (Hing-fai, 2008; McLeod, 1998) องค์การที่มีส่วนร่วมในการทดสอบที่มีผลกระทบสูงส่วนใหญ่ให้ความสนใจในเรื่องของการทุจริตในการทดสอบชนิดแบบรายข้อ ที่เกิดขึ้นจากการรั่วไหลของข้อสอบ และการจดจำคำตอบ (Drasgow, Luecht, & Bennett, 2006) ซึ่งวิธีการตรวจสอบการทุจริตในการทดสอบทางสถิติที่มีอยู่ในปัจจุบันไม่สามารถตรวจสอบการทุจริตทั้งหมดได้อย่างมีประสิทธิภาพ Shu (2010) ได้นำเสนอ โมเดลการตรวจสอบการทุจริตในการทดสอบด้วย DGRTM บนพื้นฐานทฤษฎีการตอบสนองข้อสอบ 1 พารามิเตอร์ เพื่อใช้ในการตรวจสอบการทุจริตในการทดสอบสำหรับข้อสอบที่มีการรั่วไหลก่อนการทดสอบ ผู้วิจัยจึงได้ขยายขอบเขตของการวิจัยนี้โดยปรับเพิ่มเป็นโมเดลการตรวจสอบการทุจริตในการทดสอบด้วย DGRTM 2 พารามิเตอร์ โดยใช้ข้อมูลจำลอง มุ่งเน้นกรณีที่มีข้อสอบรั่วไหลอันเกิดจากการทุจริตของผู้ที่มีส่วนเกี่ยวข้องกับข้อสอบ เพื่อเติมเต็มองค์ความรู้ทางวิชาการ และเป็นประโยชน์ต่อผู้มีส่วนเกี่ยวข้องกับการทดสอบ อันได้แก่ สถานศึกษาต่าง ๆ และโดยเฉพาะอย่างยิ่งภาครัฐที่ได้รับผลกระทบจากการทุจริตในการสอบบรรจุเข้ารับข้าราชการ

วัตถุประสงค์การวิจัย

1. เพื่อศึกษาอัตราความคลาดเคลื่อนประเภทที่ 1 ของการตรวจสอบการทุจริตในการทดสอบโมเดลการตอบสนองข้อสอบด้วย DGRTM 2 พารามิเตอร์ ภายใต้ปัจจัยที่แปรเปลี่ยน 6 ปัจจัยคือ (1) ความยาวของข้อสอบ (2) ระดับความแตกต่างของการแจกแจงความยาวรายข้อของข้อสอบ (3) อำนาจจำแนกรายข้อของข้อสอบ (4) ขนาดของการทำทุจริต (5) ขนาดของการรั่วไหลของข้อสอบ และ (6) ประสิทธิภาพในการทำทุจริต
2. เพื่อศึกษาอำนาจการทดสอบของการตรวจสอบการทุจริตในการทดสอบของโมเดลการตอบสนองข้อสอบด้วย DGRTM 2 พารามิเตอร์ ภายใต้ปัจจัยที่แปรเปลี่ยน 6 ปัจจัยคือ (1) ความยาวของข้อสอบ (2) ระดับความแตกต่างของการแจกแจงความยาวรายข้อของข้อสอบ (3) อำนาจจำแนกรายข้อของข้อสอบ (4) ขนาดของการทำทุจริต (5) ขนาดของการรั่วไหลของข้อสอบ และ (6) ประสิทธิภาพในการทำทุจริต

ขอบเขตของการวิจัย

ข้อมูลที่ศึกษาเป็นข้อมูลจำลอง ใช้โมเดลการตรวจสอบการทุจริตในการทดสอบด้วย DGIRTM 2 พารามิเตอร์ ภายใต้ทฤษฎีการตอบสนองข้อสอบ จำลองข้อสอบแบบให้คำตอบ 2 ค่า การจำลองข้อมูลจำลองผล การตอบข้อสอบภายใต้ปัจจัยที่แปรเปลี่ยน 6 ปัจจัย คือ ความยาวของแบบทดสอบ 2 ขนาด ระดับความแตกต่างของการแจกแจงความยาวกรายข้อของข้อสอบ 2 ระดับ อำนาจจำแนกรายข้อของข้อสอบ 2



ภาพที่ 1 กรอบแนวความคิดการวิจัย

ระดับ ขนาดของการทำทุจริต 2 ระดับ ขนาดของการรั่วไหลของข้อสอบ 2 ระดับ และประสิทธิผลในการทำทุจริต 3 ระดับ รวมข้อมูลทั้งหมดที่ต้องจัดกระทำเพื่อตรวจสอบการทุจริตในการทดสอบจำนวน 96 เpsilon (2×2×2×2×2×3) ในแต่ละเpsilonจำลองข้อมูลซ้ำ 100 รอบ

วิธีดำเนินการวิจัย

การจัดกระทำตัวแปร

การวิจัยครั้งนี้ศึกษาสถานการณ์จำลอง ใช้ทฤษฎีการตอบสนองข้อสอบ 2 พารามิเตอร์ จำลองข้อมูล ภายใต้ปัจจัยที่แปรเปลี่ยน 6 ปัจจัย คือ ความยาวของแบบทดสอบ 2 ขนาด ระดับความแตกต่างของการแจกแจงความยาวกรายข้อของข้อสอบ 2 ระดับ อำนาจจำแนกรายข้อของข้อสอบ 2 ระดับ ขนาดของการทำทุจริต 2 ระดับ ขนาดของการรั่วไหลของข้อสอบ 2 ระดับ และประสิทธิผลในการทำทุจริต 3 ระดับ ตามลำดับ สำหรับการ จัดกระทำตัวแปรอิสระทั้ง 6 ตัวแปร เป็นดังนี้

1. **ตัวแปรต้น** มี 6 ตัวแปร คือ
 - 1.1 ความยาวของแบบทดสอบ 2 ขนาด คือ สั้น 40 ข้อ และยาว 80 ข้อ (Shu, 2010)
 - 1.2 ระดับความแตกต่างของการแจกแจงความยากรายข้อของข้อสอบ 2 ระดับ คือ ต่ำ เท่ากับ 0.5 และสูง เท่ากับ 1.0 (Shu, 2010)
 - 1.3 อำนาจจำแนกรายข้อของข้อสอบ 2 ระดับ คือ ต่ำ เท่ากับ 0.35-0.64 และสูง เท่ากับ 1.35-1.69 (Baker, 2001)
 - 1.4 ขนาดของการทำทุจริต 2 ระดับ คือ ต่ำ เท่ากับ 5% และปานกลาง เท่ากับ 35% (Shu, 2010)
 - 1.5 ขนาดของการรั่วไหลของข้อสอบ 2 ระดับ คือ ปานกลาง เท่ากับ 50% และสูงเท่ากับ 70% (Shu, 2010)
 - 1.6 ประสิทธิภาพในการทำทุจริต 3 ระดับ คือ ต่ำ เท่ากับ ความสามารถที่แท้จริง + Beta(1.5,3)*3 ปานกลาง เท่ากับ ความสามารถที่แท้จริง + Beta(5,5)*3 และสูง เท่ากับ ความสามารถที่แท้จริง + Beta(9,4)*3 (Shu, 2010)
2. **ตัวแปรตาม** มี 2 ตัวแปร คือ
 - 2.1 อัตราความคลาดเคลื่อนประเภทที่ 1
 - 2.2 อำนาจการทดสอบ

วิธีการวิเคราะห์ข้อมูล

การวิเคราะห์เพื่อตรวจสอบการทุจริตในการทดสอบ ผู้วิจัยปรับใช้โมเดลการตรวจสอบการทุจริตในการทดสอบด้วย DGIRTM 2 พารามิเตอร์ บนพื้นฐานความสามารถที่แท้จริง แสดงถึงความรู้ ความสามารถจริง ๆ ของผู้สอบ และความสามารถในการทำทุจริตในการประมาณค่าประสิทธิภาพในการทำทุจริตของผู้สอบ ส่วนของโครงสร้างของโมเดลการตรวจสอบการทุจริตในการทดสอบด้วย DGIRTM 2 พารามิเตอร์ แสดงในสมการที่ 1 ส่วนโมเดลที่ใช้วัดความสามารถ 2 ชนิด แสดงในสมการที่ 2 และ 3 ดังนี้

$$P(U_{ij}=1|\theta_{ij}, \theta_{cj}, a_i, b_i, l_i, T_j) =$$

$$P(U_{ij} = 1) = P(U_{ij} = 1 | \theta_i)^{1-T_j} * [(1-I_i) * P(U_{ij} = 1 | \theta_i) + I_i * P(U_{ij} = 1 | \theta_c)]^{T_j} \quad (1)$$

$$P(U_{ij} = 1 | \theta_{ij}) = \frac{e^{Da_i(\theta_{ij}-b_i)}}{1 + e^{Da_i(\theta_{ij}-b_i)}} \quad i = 1, 2, 3, \dots, n \quad (2)$$

$$P(U_{ij} = 1 | \theta_{cj}) = \frac{e^{Da_i(\theta_{cj}-b_i)}}{1 + e^{Da_i(\theta_{cj}-b_i)}} \quad i = 1, 2, 3, \dots, n \quad (3)$$

$$T_j = 1, \quad \text{เมื่อ} \quad \theta_{ij} < \theta_{cj}$$

เมื่อ i แทน ข้อสอบ มีค่า 1,2,3,...l

j แทน ผู้ตอบข้อสอบ หรือผู้สอบ มีค่า 1, 2,3,...J

a_i แทน ค่าอำนาจจำแนกรายข้อของข้อสอบข้อที่ i

b_i แทน ค่าความยากรายข้อของข้อสอบข้อที่ i

T_j แทน สถานะของผู้สอบที่ทำทุจริต

θ_{ij} แทน ความสามารถที่แท้จริงของผู้สอบ

θ_{cj} แทน ความสามารถในการทำทุจริตของผู้สอบ

I_i แทน ข้อสอบที่มีสถานะรั่วไหล

อัตราความคลาดเคลื่อนประเภทที่ 1

เกณฑ์การพิจารณาหากมีค่าความคลาดเคลื่อนประเภทที่ 1 ต่ำกว่า หรือเท่ากับ .05 ถือว่า ควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ดี (Atar & Kamata, 2011) นั่นคือ การตรวจสอบการทุจริตในการทดสอบนั้น ตรวจสอบผู้ที่ไม่ได้ทำทุจริตว่า เป็นผู้บริสุทธิ์ได้จริง

อำนาจการทดสอบ

เกณฑ์การพิจารณาหากมีค่าอำนาจการทดสอบมีค่าเฉลี่ยตั้งแต่ 0.80 ขึ้นไป ถือว่า มีอำนาจการทดสอบเพียงพอ หากต่ำกว่า 0.80 ถือว่า การตรวจสอบการทุจริตได้ไม่ดี (Atar & Kamata, 2011)

ผลการวิจัย

1. ผลการวิเคราะห์ประสิทธิภาพการควบคุมความคลาดเคลื่อนประเภทที่ 1 ของการตรวจสอบการทุจริตในการทดสอบโมเดลการตอบสนองข้อสอบด้วย DGIRTM 2 พารามิเตอร์ ภายใต้ปัจจัยที่แตกต่างกัน 6 ปัจจัย ได้แก่ ความยาวของข้อสอบ ระดับความแตกต่างของการแจกแจงความยากรายข้อของข้อสอบ อำนาจจำแนกรายข้อของข้อสอบ ขนาดของการทำทุจริต ขนาดของการรั่วไหลของข้อสอบ และประสิทธิผลในการทำทุจริต ผลการวิเคราะห์ พบว่า ทุกเงื่อนไขไม่สามารถควบคุมอัตราความคลาดเคลื่อนประเภทที่ 1 ได้ตามเกณฑ์ที่กำหนด สรุปผลตามรายละเอียดเงื่อนไขดังต่อไปนี้

1.1 เมื่อพิจารณาภายใต้เงื่อนไขภายใต้เงื่อนไขความยาวของข้อสอบ สรุปได้ว่า ความยาวของข้อสอบมีอิทธิพลต่ออัตราความคลาดเคลื่อนประเภทที่ 1 อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 โดยความยาวของข้อสอบ 80 ข้อ มีค่าเฉลี่ยของอัตราความคลาดเคลื่อนประเภทที่ 1 ต่ำกว่าความยาวของข้อสอบ 40 ข้อ

1.2 เมื่อพิจารณาภายใต้เงื่อนไขระดับความแตกต่างของการแจกแจงความยากรายข้อของข้อสอบ สรุปได้ว่า ระดับความแตกต่างของการแจกแจงความยากรายข้อของข้อสอบมีอิทธิพลต่ออัตราความคลาดเคลื่อนประเภทที่ 1 อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 โดยที่ระดับความแตกต่างของการแจกแจงความยากรายข้อของข้อสอบต่ำมีอัตราความคลาดเคลื่อนประเภทที่ 1 ต่ำกว่าระดับความแตกต่างของการแจกแจงความยากรายข้อของข้อสอบสูง

1.3 เมื่อพิจารณาภายใต้เงื่อนไขอำนาจจำแนกรายข้อของข้อสอบ สรุปได้ว่า อำนาจจำแนกรายข้อของข้อสอบมีอิทธิพลต่ออัตราความคลาดเคลื่อนประเภทที่ 1 อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 โดยที่ระดับอำนาจจำแนกรายข้อของข้อสอบสูงมีอัตราความคลาดเคลื่อนประเภทที่ 1 ต่ำกว่าอำนาจจำแนกรายข้อของข้อสอบต่ำ

1.4 เมื่อพิจารณาภายใต้เงื่อนไขขนาดของการทำทุจริต สรุปได้ว่า ขนาดของการทำทุจริตมีอิทธิพลต่ออัตราความคลาดเคลื่อนประเภทที่ 1 อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 โดยที่ขนาดของการทำทุจริตสูงมีอัตราความคลาดเคลื่อนประเภทที่ 1 ต่ำกว่าขนาดของการทำทุจริตต่ำ

1.5 เมื่อพิจารณาภายใต้เงื่อนไขขนาดของการรั่วไหลของข้อสอบ สรุปได้ว่า ขนาดการรั่วไหลของข้อสอบมีอิทธิพลต่ออัตราความคลาดเคลื่อนประเภทที่ 1 อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 โดยที่ขนาดการรั่วไหลของข้อสอบสูงมีอัตราความคลาดเคลื่อนประเภทที่ 1 สูงกว่าขนาดการรั่วไหลของข้อสอบต่ำ

1.6 เมื่อพิจารณาภายใต้เงื่อนไขประสิทธิผลในการทำทุจริต สรุปได้ว่า ประสิทธิผลในการทำทุจริตมีอิทธิพลต่ออัตราความคลาดเคลื่อนประเภทที่ 1 อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 โดยที่ระดับประสิทธิผลในการทำทุจริตสูงมีอัตราความคลาดเคลื่อนประเภทที่ 1 น้อยที่สุด รองลงมา คือ ประสิทธิผลในการทำทุจริตปานกลาง และมากที่สุด คือ ประสิทธิผลในการทำทุจริตต่ำ

2. ผลการวิเคราะห์ประสิทธิภาพอำนาจการทดสอบของการตรวจสอบการทุจริตในการทดสอบด้วย DGRTM 2 พารามิเตอร์ ภายใต้ปัจจัยที่แตกต่างกัน 6 ปัจจัย ได้แก่ ความยาวของข้อสอบ ระดับความแตกต่างของการแจกแจงความยากรายข้อของข้อสอบ อำนาจจำแนกรายข้อของข้อสอบ ขนาดของการทำทุจริต ขนาดของการรั่วไหลของข้อสอบ และประสิทธิผลในการทำทุจริต สรุปได้ว่า ที่ระดับประสิทธิผลในการทำทุจริตต่ำ มีอำนาจการทดสอบต่ำกว่าเกณฑ์ที่กำหนด ส่วนที่ระดับประสิทธิผลในการทำทุจริตปานกลาง และระดับประสิทธิผลในการทำทุจริตสูงมีอำนาจการทดสอบสูงกว่าเกณฑ์ที่กำหนด สรุปผลตามรายละเอียดเงื่อนไขดังต่อไปนี้

2.1 เมื่อพิจารณาภายใต้เงื่อนไขความยาวของข้อสอบ สรุปได้ว่า ความยาวของข้อสอบมีอิทธิพลต่ออำนาจการทดสอบอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 โดยที่ระดับอำนาจจำแนกรายข้อของข้อสอบต่ำ และที่ระดับประสิทธิผลในการทำทุจริตต่ำ ความยาวของข้อสอบ 80 ข้อมีค่าอำนาจการทดสอบต่ำกว่าความยาวของข้อสอบ 40 ข้อ นอกนั้นที่ความยาวของข้อสอบ 80 ข้อมีค่าอำนาจการทดสอบสูงกว่าความยาวข้อสอบ 40 ข้อ ในทุกเงื่อนไข

2.2 เมื่อพิจารณาภายใต้เงื่อนไขระดับความแตกต่างของการแจกแจงความยากรายข้อของข้อสอบ สรุปได้ว่า ระดับความแตกต่างของการแจกแจงความยากรายข้อของข้อสอบมีอิทธิพลต่ออำนาจการทดสอบ อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 โดยที่ระดับความแตกต่างของการแจกแจงความยากรายข้อของข้อสอบสูงมีค่าอำนาจการทดสอบสูงกว่าระดับความแตกต่างของการแจกแจงความยากรายข้อของข้อสอบต่ำ

2.3 เมื่อพิจารณาภายใต้เงื่อนไขอำนาจจำแนกรายข้อของข้อสอบ สรุปได้ว่า อำนาจจำแนกรายข้อของข้อสอบมีอิทธิพลต่ออำนาจการทดสอบ อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 โดยที่อำนาจจำแนกรายข้อของข้อสอบสูงมีค่าอำนาจการทดสอบสูงกว่าอำนาจจำแนกรายข้อของข้อสอบต่ำ

2.4 เมื่อพิจารณาภายใต้เงื่อนไขขนาดของการทำทุจริต สรุปได้ว่า ขนาดของการทำทุจริตมีอิทธิพลต่ออำนาจการทดสอบอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 โดยที่ขนาดของการทำทุจริตสูงมีอำนาจการทดสอบสูงกว่าขนาดของการทำทุจริตต่ำ

2.5 เมื่อพิจารณาภายใต้เงื่อนไขขนาดของการรั่วไหลของข้อสอบ สรุปได้ว่า ขนาดการรั่วไหลของข้อสอบมีอิทธิพลต่ออำนาจการทดสอบอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 โดยเมื่อพิจารณาที่อำนาจจำแนกรายข้อของข้อสอบร่วมกับขนาดการรั่วไหลของข้อสอบ พบว่า ที่ระดับอำนาจจำแนกรายข้อของข้อสอบต่ำ ขนาดการรั่วไหลของข้อสอบต่ำมีอำนาจการทดสอบสูงกว่าขนาดการรั่วไหลของข้อสอบสูง ในทางกลับกันเมื่อพิจารณาที่อำนาจจำแนกรายข้อของข้อสอบร่วมกับขนาดการรั่วไหลของข้อสอบ พบว่า ที่ระดับอำนาจจำแนกรายข้อของข้อสอบสูง ขนาดการรั่วไหลของข้อสอบสูงมีอำนาจการทดสอบสูงกว่าขนาดการรั่วไหลของข้อสอบต่ำ

2.6 เมื่อพิจารณาภายใต้เงื่อนไขประสิทธิผลในการทำทุจริต สรุปได้ว่า ประสิทธิผลในการทำทุจริตมีอิทธิพลต่ออำนาจการทดสอบ อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 โดยประสิทธิผลในการทำทุจริตสูงมีค่าเฉลี่ยอำนาจการทดสอบมากที่สุด รองลงมา คือ ประสิทธิผลในการทำทุจริตปานกลาง และน้อยที่สุด คือ ประสิทธิผลในการทำทุจริตต่ำ

อภิปรายผลการวิจัย

การวิจัยครั้งนี้มีเป้าหมายเพื่อตอบคำถามใน 2 ประเด็นหลัก คือ (1) อัตราความคลาดเคลื่อนประเภทที่ 1 ของการตรวจสอบการทุจริตในการทดสอบด้วย DGIRTM 2 พารามิเตอร์ ภายใต้ปัจจัยที่แปรเปลี่ยน 6 ปัจจัย คือ ความยาวของข้อสอบ ระดับความแตกต่างของการแจกแจงความยาวรายข้อของข้อสอบ อำนาจจำแนกรายข้อของข้อสอบ ขนาดของการทำทุจริต ขนาดของการรั่วไหลของข้อสอบ และประสิทธิผลในการทำทุจริต และ (2) อำนาจการทดสอบของการตรวจสอบการทุจริตในการทดสอบด้วย DGIRTM 2 พารามิเตอร์ ภายใต้ปัจจัยที่แปรเปลี่ยน 6 ปัจจัย คือ ความยาวของข้อสอบ ระดับความแตกต่างของการแจกแจงความยาวรายข้อของข้อสอบ อำนาจจำแนกรายข้อของข้อสอบ ขนาดของการทำทุจริต ขนาดของการรั่วไหลของข้อสอบ และประสิทธิผลในการทำทุจริต จากการสรุปผลการวิจัยข้างต้น มีประเด็นสำคัญที่นำมาอภิปรายผล ดังนี้

1. ปัจจัยความยาวของข้อสอบ

อัตราความคลาดเคลื่อนประเภทที่ 1 ของการตรวจสอบการทุจริตในการทดสอบด้วย DGIRTM 2 พารามิเตอร์ ภายใต้ปัจจัยที่แปรเปลี่ยน 6 ปัจจัย พบว่า ความยาวของข้อสอบมีอิทธิพลต่ออัตราความคลาดเคลื่อนประเภทที่ 1 อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 โดยความยาวของข้อสอบ 80 ข้อมีค่าเฉลี่ยของอัตราความคลาดเคลื่อนประเภทที่ 1 ต่ำกว่าความยาวของข้อสอบ 40 ข้อ

ผลการวิจัยดังกล่าว สอดคล้องกับผลการวิจัยของ Shu (2010) ที่ศึกษาการตรวจสอบการทุจริตในการทดสอบด้วย DGIRTM 1 พารามิเตอร์ พบว่า ข้อสอบแบบสั้นให้ค่าความจำเพาะของโมเดล (Model Specificity) สูงกว่าข้อสอบแบบยาว ทั้งนี้ ค่าความจำเพาะของโมเดล คือ สัดส่วนของผู้สอบที่ไม่ได้ทำทุจริต และถูกแยกแยะได้ถูกต้องว่า ไม่ใช่เป็นผู้ที่ทำทุจริต นั่นคือ เทียบเคียงได้ว่า ข้อสอบแบบยาวให้ผลค่าเฉลี่ยอัตราความคลาดเคลื่อนประเภทที่ 1 ต่ำกว่า ข้อสอบแบบสั้น (Shu, 2010, pp. 75-78) ซึ่งคล้ายกับผลการศึกษาของ Deng (2007) ที่พบว่า ยิ่งข้อสอบมีขนาดความยาวมากขึ้น ก็ยิ่งส่งผลให้ค่าอัตราความคลาดเคลื่อนประเภทที่ 1 ลดลง ทั้งนี้ ความยาวของข้อสอบเป็นสารสนเทศปฐมภูมิ (Primary Information) สารสนเทศปฐมภูมิมิค่ามากขึ้นเมื่อความยาวของข้อสอบมากขึ้น ซึ่งส่งผลให้ค่าอัตราความคลาดเคลื่อนประเภทที่ 1 สามารถแยกแยะได้

ถูกต้องมากขึ้นเมื่อมีสารสนเทศอย่างเพียงพอ และมีความเที่ยงภายในข้อมูลที่ทำการทดสอบ (Deng, 2007, p. 111)

อำนาจการทดสอบของการตรวจสอบการทุจริตในการทดสอบด้วย DGIRTM 2 พารามิเตอร์

ภายใต้ปัจจัยที่แปรเปลี่ยน 6 ปัจจัย พบว่า ความยาวของข้อสอบมีอิทธิพลต่ออำนาจการทดสอบอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 โดยที่ระดับอำนาจจำแนกรายข้อของข้อสอบต่ำ และที่ระดับประสิทธิผลในการทำทุจริตต่ำ ความยาวของข้อสอบ 80 ข้อมีค่าอำนาจการทดสอบต่ำกว่าความยาวของข้อสอบ 40 ข้อ นอกนั้นที่ความยาวของข้อสอบ 80 ข้อมีค่าอำนาจการทดสอบสูงกว่าความยาวข้อสอบ 40 ข้อ ในทุกเงื่อนไข

ผลการวิจัยดังกล่าว สอดคล้องกับผลการวิจัยของ Shu (2010) ที่ศึกษาการตรวจสอบการทุจริตในการทดสอบด้วย DGIRTM 1 พารามิเตอร์ พบว่า ข้อสอบแบบยาวให้ค่าความไวของโมเดล (Model Sensitivity) สูงกว่าข้อสอบแบบสั้น (Shu, 2010, pp. 75-78) ทั้งนี้ ค่าความไวของโมเดล คือ สัดส่วนของผู้สอบที่ทำทุจริต และถูกแยกแยะได้ถูกต้องว่า เป็นผู้ที่ทำทุจริต ที่แท้จริง ซึ่งก็คือ อำนาจการทดสอบนั่นเอง ผลการวิจัยตรงกับผลการวิจัยของ Deng (2007, p. 113) ที่พบว่า ข้อสอบที่ยาวขึ้นสัมพันธ์กับอำนาจการทดสอบที่สูงขึ้น ความยาวของข้อสอบเป็นปัจจัยที่สำคัญมากที่เป็นตัวกำหนดสารสนเทศของการทดสอบ ซึ่งส่งผลต่ออำนาจการทดสอบของโมเดล และความถูกต้องของการประมาณค่า (Shu, 2010, p. 111)

2. ปัจจัยระดับความแตกต่างของการแจกแจงความยาวรายข้อของข้อสอบ

อัตราความคลาดเคลื่อนประเภทที่ 1 ของการตรวจสอบการทุจริตในการทดสอบด้วย DGIRTM 2 พารามิเตอร์ ภายใต้ปัจจัยที่แปรเปลี่ยน 6 ปัจจัย พบว่า ระดับความแตกต่างของการแจกแจงความยาวรายข้อของข้อสอบมีอิทธิพลต่ออัตราความคลาดเคลื่อนประเภทที่ 1 อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 โดยที่ระดับความแตกต่างของการแจกแจงความยาวรายข้อของข้อสอบต่ำมีอัตราความคลาดเคลื่อนประเภทที่ 1 ต่ำกว่าระดับความแตกต่างของการแจกแจงความยาวรายข้อของข้อสอบสูง

ผลการวิจัย ที่ระดับความแตกต่างของการแจกแจงความยาวรายข้อของข้อสอบต่ำจะให้ค่าสารสนเทศที่สูง (Shu, 2010, p. 61) ส่งผลต่อระดับอัตราความคลาดเคลื่อนประเภทที่ 1 มีค่าลดลง ในทางกลับกันที่ระดับความแตกต่างของการแจกแจงความยาวรายข้อของข้อสอบสูงจะให้ค่าสารสนเทศที่ต่ำ (Shu, 2010, p. 61) ส่งผลต่อระดับอัตราความคลาดเคลื่อนประเภทที่ 1 มีค่าสูงกว่าที่ระดับความแตกต่างของการแจกแจงความยาวรายข้อของข้อสอบต่ำ ทำให้สามารถสรุปได้ว่า ระดับความแตกต่างของการแจกแจงความยาวรายข้อของข้อสอบต่ำมีค่าเฉลี่ยอัตราความคลาดเคลื่อนประเภทที่ 1 ต่ำกว่าที่ระดับความแตกต่างของการแจกแจงความยาวรายข้อของข้อสอบสูง

อำนาจการทดสอบของการตรวจสอบการทุจริตในการทดสอบด้วย DGIRTM 2 พารามิเตอร์ ภายใต้ปัจจัยที่แปรเปลี่ยน 6 ปัจจัย พบว่า ระดับความแตกต่างของการแจกแจงความยาวรายข้อของข้อสอบมีอิทธิพลต่ออำนาจการทดสอบ อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 โดยที่ระดับความแตกต่างของการแจกแจงความยาวรายข้อของข้อสอบสูงมีค่าอำนาจการทดสอบสูงกว่าระดับความแตกต่างของการแจกแจงความยาวรายข้อของข้อสอบต่ำ

ผลการวิจัยขัดแย้งกับผลการวิจัยของ Shu (2010, p. 89) ที่พบว่า ระดับความแตกต่างของการแจกแจงความยากรายข้อของข้อสอบต่ำ (High Informative) ให้ผลลัพธ์อำนาจการทดสอบที่สูงกว่า ระดับความแตกต่างของการแจกแจงความยากรายข้อของข้อสอบสูง งานวิจัยนี้ได้ทำการเพิ่มปัจจัย อำนาจจำแนกรายข้อของข้อสอบเข้ามาในการศึกษา ส่งผลร่วมกับระดับความแตกต่างของการแจกแจงความยากรายข้อของข้อสอบ ทำให้ผลการวิจัยแตกต่างจากผลการวิจัยเดิม คือ ที่ระดับอำนาจจำแนกรายข้อของข้อสอบสูง และระดับความแตกต่างของการแจกแจงความยากรายข้อของข้อสอบสูงมีค่าอำนาจการทดสอบสูงกว่าระดับความแตกต่างของการแจกแจงความยากรายข้อของข้อสอบต่ำ

3. ปัจจัยอำนาจจำแนกรายข้อของข้อสอบ

อัตราความคลาดเคลื่อนประเภทที่ 1 ของการตรวจสอบการทุจริตในการทดสอบด้วย DGIRTM 2 พารามิเตอร์ ภายใต้ปัจจัยที่แปรเปลี่ยน 6 ปัจจัย พบว่า อำนาจจำแนกรายข้อของข้อสอบมีอิทธิพลต่ออัตราความคลาดเคลื่อนประเภทที่ 1 อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 โดยที่ระดับอำนาจจำแนกรายข้อของข้อสอบสูงมีอัตราความคลาดเคลื่อนประเภทที่ 1 ต่ำกว่าอำนาจจำแนกรายข้อของข้อสอบต่ำ

ผลการวิจัยดังกล่าว สอดคล้องกับผลการวิจัยของ Hing-fai (2008) ที่พบว่า เมื่อค่าอำนาจจำแนกรายข้อของข้อสอบสูงขึ้น จะส่งผลทำให้โมเดลในการตรวจสอบการทุจริตดีขึ้น ในขณะที่ Han (2006, p. 104) พบว่า ที่อำนาจจำแนกรายข้อของข้อสอบเท่ากับ 1.6 ให้ค่าอัตราความคลาดเคลื่อนประเภทที่ 1 ต่ำกว่าที่ระดับอำนาจจำแนกรายข้อของข้อสอบเท่ากับ 0.7 อำนาจจำแนกรายข้อของข้อสอบ (ai) มีค่าสัดส่วนโดยตรงของค่าความชันของ ICC ที่ตำแหน่ง bi ค่า ai ที่สูงสามารถจำแนกผู้สอบที่มีความสามารถแตกต่างกันได้ดี (Baker, 2001; ศิริชัย กาญจนวาสิ, 2555) ดังนั้นจึงส่งผลทำให้ข้อสอบที่มีอำนาจจำแนกรายข้อสูง เกิดอัตราความคลาดเคลื่อนประเภทที่ 1 ต่ำกว่าที่ระดับอำนาจจำแนกรายข้อต่ำ

อำนาจการทดสอบของการตรวจสอบการทุจริตในการทดสอบด้วย DGIRTM 2 พารามิเตอร์ ภายใต้ปัจจัยที่แปรเปลี่ยน 6 ปัจจัย พบว่า อำนาจจำแนกรายข้อของข้อสอบมีอิทธิพลต่ออำนาจการทดสอบ อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 โดยที่อำนาจจำแนกรายข้อของข้อสอบสูงมีค่าอำนาจการทดสอบสูงกว่าอำนาจจำแนกรายข้อของข้อสอบต่ำ

ผลการวิจัยดังกล่าว สอดคล้องกับผลการวิจัยของ Han (2006, p. 104) ที่พบว่า ข้อสอบที่มีค่าอำนาจจำแนกรายข้อ 1.2 มีค่าอำนาจการทดสอบสูงกว่าข้อสอบที่มีค่าอำนาจจำแนกรายข้อ 0.7 และ 0.4 ตามลำดับ ทั้งเป็นผลมาจากอำนาจจำแนกรายข้อของข้อสอบ (ai) มีค่าสัดส่วนโดยตรงของค่าความชันของ ICC ที่ตำแหน่ง bi ค่า ai ที่สูงสามารถจำแนกผู้สอบที่มีความสามารถแตกต่างกันได้ดี (Baker, 2001; ศิริชัย กาญจนวาสิ, 2555) ดังนั้นจึงส่งผลทำให้ข้อสอบที่มีอำนาจจำแนกรายข้อสูง มีอำนาจการทดสอบสูงกว่าข้อสอบที่มีค่าอำนาจจำแนกรายข้อต่ำ

4. ปัจจัยขนาดของการทำทุจริต

อัตราความคลาดเคลื่อนประเภทที่ 1 ของการตรวจสอบการทุจริตในการทดสอบด้วย DGIRTM 2 พารามิเตอร์ ภายใต้ปัจจัยที่แปรเปลี่ยน 6 ปัจจัย พบว่า ขนาดของการทำทุจริตมีอิทธิพลต่ออัตราความคลาดเคลื่อนประเภทที่ 1 อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 โดยที่ขนาดของการทำทุจริตสูงมีอัตราความคลาดเคลื่อนประเภทที่ 1 ต่ำกว่าขนาดของการทำทุจริตต่ำ

ผลการวิจัยดังกล่าว สอดคล้องกับผลการวิจัยของ Shu, Henson, and Luecht (2013) ทั้งนี้เป็นผลมาจากการที่จำนวนผู้ทำทุจริตมีจำนวนมากขึ้น ส่งผลทำให้โมเดลมีประสิทธิภาพในการตรวจสอบผู้ทำทุจริตได้ดีขึ้น ทำให้โอกาสในการตรวจสอบผู้ที่ไม่ได้ทำทุจริตว่าเป็นผู้ที่ทำทุจริตลดลง ส่งผลให้อัตราความคลาดเคลื่อนประเภทที่ 1 ต่ำลง ทำให้สามารถสรุปได้ว่า ขนาดของการทำทุจริตสูงมีอัตราความคลาดเคลื่อนประเภทที่ 1 ต่ำกว่าขนาดของการทำทุจริตต่ำ

อำนาจการทดสอบของการตรวจสอบการทุจริตในการทดสอบด้วย DGIRTM 2 พารามิเตอร์ ภายใต้ปัจจัยที่แปรเปลี่ยน 6 ปัจจัย พบว่า ขนาดของการทำทุจริตมีอิทธิพลต่ออำนาจการทดสอบอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 โดยที่ขนาดของการทำทุจริตสูงมีอำนาจการทดสอบสูงกว่าขนาดของการทำทุจริตต่ำ

ผลการวิจัยดังกล่าว สอดคล้องกับผลการวิจัยของ Shu, Henson, and Luecht (2013) ทั้งนี้ จำนวนผู้ทำทุจริตในการทดสอบยิ่งมาก จะส่งผลให้ออกาสในการตรวจพบผู้ที่ทำทุจริต ยังมีโอกาสมากขึ้น เป็นเพราะจำนวนผู้ทำทุจริตยิ่งมากขึ้นก็จะมีโอกาสสามารถตรวจพบได้มากขึ้นนั่นเอง ทำให้สามารถสรุปได้ว่า ขนาดของการทำทุจริตสูงมีอำนาจการทดสอบสูงกว่าขนาดของการทำทุจริตต่ำ

5. ปัจจัยขนาดของการรั่วไหลของข้อสอบ

อัตราความคลาดเคลื่อนประเภทที่ 1 ของการตรวจสอบการทุจริตในการทดสอบด้วย DGIRTM 2 พารามิเตอร์ ภายใต้ปัจจัยที่แปรเปลี่ยน 6 ปัจจัย พบว่า ขนาดการรั่วไหลของข้อสอบมีอิทธิพลต่ออัตราความคลาดเคลื่อนประเภทที่ 1 อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 โดยที่ขนาดการรั่วไหลของข้อสอบสูงมีอัตราความคลาดเคลื่อนประเภทที่ 1 สูงกว่าขนาดการรั่วไหลของข้อสอบต่ำ

ผลการวิจัยดังกล่าว สอดคล้องกับผลการวิจัยของ Shu (2010) ที่พบว่า ขนาดของการรั่วไหลของข้อสอบมีอิทธิพลต่ออัตราความคลาดเคลื่อนประเภทที่ 1 กล่าวคือ ที่ความยาวของข้อสอบ 40 ข้อ และมีขนาดการรั่วไหลของข้อสอบสูงมีอัตราความคลาดเคลื่อนประเภทที่ 1 สูงกว่าที่ความยาวของข้อสอบ 40 ข้อ และมีขนาดการรั่วไหลของข้อสอบต่ำ ในขณะที่ความยาวของข้อสอบ 80 ข้อ พบว่า ขนาดการรั่วไหลของข้อสอบสูงมีอัตราความคลาดเคลื่อนประเภทที่ 1 สูงกว่าขนาดการรั่วไหลของข้อสอบต่ำ (Shu, 2010, p. 77) ซึ่งเมื่อผู้วิจัยทำการเพิ่มปัจจัยอำนาจจำแนกรายข้อของข้อสอบเข้าไปในการวิจัย ให้ผลในทำนองเดียวกับผลการวิจัยของ Shu (2010) เพียงแต่ให้รายละเอียดที่เพิ่มเติมดังที่กล่าวมาข้างต้น

อำนาจการทดสอบของการตรวจสอบการทุจริตในการทดสอบด้วย DGIRTM 2 พารามิเตอร์ ภายใต้ปัจจัยที่แปรเปลี่ยน 6 ปัจจัย พบว่า ขนาดการรั่วไหลของข้อสอบมีอิทธิพลต่ออำนาจการทดสอบอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 โดยที่ระดับอำนาจจำแนกรายข้อของข้อสอบต่ำ ขนาดการรั่วไหลของข้อสอบต่ำมีอำนาจการทดสอบสูงกว่าขนาดการรั่วไหลของข้อสอบสูง ในทางกลับกันเมื่อพิจารณาที่ อำนาจจำแนกรายข้อของ

ข้อสอบร่วมกับขนาดการรั่วไหลของข้อสอบ พบว่า ที่ระดับอำนาจจำแนกรายข้อของข้อสอบสูง ขนาดการรั่วไหลของข้อสอบสูงมีอำนาจการทดสอบสูงกว่าขนาดการรั่วไหลของข้อสอบต่ำ

ผลการวิจัยดังกล่าว สอดคล้องกับผลการวิจัยของ Shu (2010) ที่พบว่า ที่ขนาดการรั่วไหลของข้อสอบ 70% มีอำนาจการทดสอบสูงกว่าที่ขนาดการรั่วไหลของข้อสอบ 30% (Shu, 2010, p. 86) แต่ในกรณีที่ข้อสอบรั่วไหลมีขนาดเท่ากับข้อสอบที่ไม่รั่วไหลที่ 50% คือ สมดุล จะให้ค่าอำนาจการทดสอบสูงกว่าขนาดการรั่วไหลของข้อสอบ 30% และ 70% ในกรณีของงานวิจัยนี้ ได้เพิ่มปัจจัย อำนาจจำแนกรายข้อของข้อสอบเข้าไป พบว่า ที่อำนาจจำแนกรายข้อของข้อสอบสูง และขนาดการรั่วไหลของข้อสอบสมดุล ให้ค่าอำนาจการทดสอบต่ำกว่า ขนาดการรั่วไหลของข้อสอบ 70% ซึ่งเป็นผลการวิจัยที่แตกต่างไปจากเดิม อันเนื่องมาจากปัจจัยที่เพิ่มขึ้น

6. ปัจจัยประสิทธิผลในการทำทุจริต

อัตราความคลาดเคลื่อนประเภทที่ 1 ของการตรวจสอบการทุจริตในการทดสอบด้วย DGIRTM 2 พารามิเตอร์ ภายใต้ปัจจัยที่แปรเปลี่ยน 6 ปัจจัย พบว่า ประสิทธิภาพในการทำทุจริตมีอิทธิพลต่ออัตราความคลาดเคลื่อนประเภทที่ 1 อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 โดยที่ระดับประสิทธิผลในการทำทุจริตสูงมีอัตราความคลาดเคลื่อนประเภทที่ 1 น้อยที่สุด รองลงมา คือ ประสิทธิภาพในการทำทุจริตปานกลาง และมากที่สุด คือ ประสิทธิภาพในการทำทุจริตต่ำ

ผลการวิจัยดังกล่าว สอดคล้องกับผลการวิจัยของ Shu (2010) ที่ศึกษา การตรวจสอบการทุจริตในการทดสอบ พบว่า ประสิทธิภาพในการทำทุจริตมีอิทธิพลต่ออัตราความคลาดเคลื่อนประเภทที่ 1 โดยอัตราความคลาดเคลื่อนประเภทที่ 1 สำหรับประสิทธิผลในการทำทุจริตสูงมีค่าต่ำกว่าอัตราความคลาดเคลื่อนประเภทที่ 1 สำหรับประสิทธิผลในการทำทุจริตต่ำ (Shu, pp. 75-76) เนื่องจาก ประสิทธิภาพในการทำทุจริตนั้น ส่งผลต่อสารสนเทศของการทดสอบ ผู้สอบที่มีประสิทธิผลในการทำทุจริตสูง จะได้คะแนนเพิ่มขึ้นมากกว่า ผู้เข้าสอบที่มีประสิทธิผลในการทำทุจริตต่ำ และผู้เข้าสอบที่ไม่ทำทุจริต ส่งผลทำให้มีสารสนเทศที่สูงกว่า สารสนเทศของข้อสอบนั้นส่งผลโดยตรงต่ออัตราความคลาดเคลื่อนประเภทที่ 1 สารสนเทศที่สูงจะให้อัตราความคลาดเคลื่อนประเภทที่ 1 ต่ำ ทำให้สามารถสรุปได้ว่า หากประสิทธิผลในการทำทุจริตสูงขึ้นจะส่งผลให้อัตราความคลาดเคลื่อนประเภทที่ 1 ต่ำลง

อำนาจการทดสอบของการตรวจสอบการทุจริตในการทดสอบด้วย DGIRTM 2 พารามิเตอร์ ภายใต้ปัจจัยที่แปรเปลี่ยน 6 ปัจจัย พบว่า ประสิทธิภาพในการทำทุจริตมีอิทธิพลต่ออำนาจการทดสอบ อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 โดยประสิทธิผลในการทำทุจริตสูงมีค่าเฉลี่ยอำนาจการทดสอบมากที่สุด รองลงมา คือ ประสิทธิภาพในการทำทุจริตปานกลาง และน้อยที่สุด คือ ประสิทธิภาพในการทำทุจริตต่ำ

ผลการวิจัยดังกล่าว สอดคล้องกับผลการวิจัยของ Shu (2010) ที่ศึกษา การตรวจสอบการทุจริตในการทดสอบ พบว่า ประสิทธิภาพในการทำทุจริตมีอิทธิพลต่ออำนาจการทดสอบ โดยอำนาจการทดสอบสำหรับประสิทธิผลในการทำทุจริตสูงมีค่าสูงกว่าอำนาจการทดสอบสำหรับประสิทธิผลในการทำทุจริตต่ำ (Shu, 2010, pp. 87-88) เนื่องจาก ประสิทธิภาพในการทำทุจริตนั้น ส่งผลต่อสารสนเทศของการทดสอบ ผู้สอบที่มีประสิทธิผลในการทำทุจริตสูง จะได้คะแนนเพิ่มขึ้นมากกว่า ผู้เข้าสอบที่มีประสิทธิผลในการทำทุจริตต่ำ และผู้เข้าสอบที่ไม่ทำทุจริต ส่งผลทำให้มีสารสนเทศที่สูงกว่า สารสนเทศของข้อสอบนั้นส่งผลโดยตรงต่ออำนาจการ

ทดสอบ สารสนเทศที่สูงจะให้อำนาจการทดสอบที่สูง ทำให้สามารถสรุปได้ว่า ประสิทธิภาพในการทำทุจริตที่สูงขึ้น จะส่งผลให้อำนาจการทดสอบที่สูงขึ้นตามไปด้วย

ข้อเสนอแนะ

เพื่อการนำไปใช้

1. ผลการศึกษาครั้งนี้ พบว่า ความยาวของข้อสอบแบบยาว (80 ข้อ) ให้ผลลัพธ์อัตราความคลาดเคลื่อนประเภทที่ 1 และอัตราอำนาจการทดสอบสูงกว่า ความยาวของข้อสอบแบบสั้น (40 ข้อ) ดังนั้น โมเดลการตรวจสอบการทุจริตในการทดสอบด้วย DGIRTM 2 พารามิเตอร์ มีความเหมาะสมสำหรับความยาวของข้อสอบแบบยาว

2. ผลการศึกษาครั้งนี้ พบว่า ระดับความแตกต่างของการแจกแจงของความยากรายข้อของข้อสอบต่ำ ให้ผลลัพธ์อัตราความคลาดเคลื่อนประเภทที่ 1 สูงกว่า ระดับความแตกต่างของการแจกแจงของความยากรายข้อของข้อสอบสูง ดังนั้นโมเดลการตรวจสอบการทุจริตในการทดสอบด้วย DGIRTM 2 พารามิเตอร์ มีความเหมาะสมสำหรับระดับความแตกต่างของการแจกแจงของความยากรายข้อของข้อสอบสูง

3. ผลการศึกษาครั้งนี้ พบว่า ระดับอำนาจจำแนกรายข้อของข้อสอบสูงให้ผลลัพธ์อัตราความคลาดเคลื่อนประเภทที่ 1 สูงกว่า ระดับอำนาจจำแนกรายข้อของข้อสอบต่ำ ดังนั้น โมเดลการตรวจสอบการทุจริตในการทดสอบด้วย DGIRTM 2 พารามิเตอร์ มีความเหมาะสมสำหรับระดับอำนาจจำแนกรายข้อของข้อสอบสูง

4. ผลการศึกษาครั้งนี้ พบว่า ขนาดของการทำทุจริตสูงให้ผลลัพธ์อัตราอำนาจการทดสอบสูงกว่า ขนาดของการทำทุจริตต่ำ ดังนั้น โมเดลการตรวจสอบการทุจริตในการทดสอบด้วย DGIRTM 2 พารามิเตอร์ มีความเหมาะสมสำหรับขนาดของการทำทุจริตสูง

5. ผลการศึกษาครั้งนี้ พบว่า ประสิทธิภาพในการทำทุจริตสูง ให้ผลลัพธ์อัตราอำนาจการทดสอบสูงกว่า ประสิทธิภาพในการทำทุจริตต่ำ ดังนั้น โมเดลการตรวจสอบการทุจริตในการทดสอบด้วย DGIRTM 2 พารามิเตอร์ มีความเหมาะสมสำหรับประสิทธิภาพในการทำทุจริตสูง

เพื่อการวิจัยครั้งต่อไป

1. ควรทำการศึกษาในลักษณะเดียวกันกับการศึกษาในครั้งนี้ โดยทำการศึกษาปัจจัยอื่น ๆ เพิ่มเติม เช่น ขนาดของการรั่วไหลของข้อสอบต่ำ ๆ ที่ 30% ขนาดของการทำทุจริตสูงที่ 70% เป็นต้น

2. ควรทำการศึกษาในลักษณะเดียวกันกับการศึกษาในครั้งนี้ โดยทำการเพิ่มพารามิเตอร์โอกาสการเดาเข้าไปเป็นโมเดลการตรวจสอบการทุจริตในการทดสอบด้วย DGIRTM 3 พารามิเตอร์

3. ควรทำการศึกษาในลักษณะเดียวกันกับการศึกษาในครั้งนี้ โดยปรับเปลี่ยนเงื่อนไขเพิ่มเติม เช่น ตำแหน่งความยากรายข้อของข้อสอบให้สอดคล้องกับลักษณะการทำทุจริต เพื่อเน้นให้เกิดสารสนเทศสูงสุด อันจะส่งผลต่อ อัตราความคลาดเคลื่อนประเภทที่ 1 และอำนาจการทดสอบที่ดีขึ้น

เอกสารอ้างอิง

ภาษาไทย

กรุงเทพมหานครออนไลน์. (2558, 24 เมษายน). “สอบครูผู้ช่วย พบข้อทุจริตใน สพป.เลย เขต 2 ผู้เข้าสอบพกโปยเข้าห้องสอบ”. ค้นเมื่อ 5 มิถุนายน 2560, จาก <http://www.kruwande.com/news-id28047.html>

คมชัดลึกออนไลน์. (2555). “โกงเข้าตำรวจ”. ค้นเมื่อ 5 มิถุนายน, 2560, จาก <http://www.komchadluek.net/detail/20120614/132748/โกงเข้าตำรวจ.html>

ศิริชัย กาญจนวาสี. (2555). “ทฤษฎีการทดสอบแนวใหม่” (พิมพ์ครั้งที่ 4). กรุงเทพฯ: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

ภาษาต่างประเทศ

Atar, B., & Kamata, A. (2011). "Comparison of IRT likelihood ratio test and logistic regression DIF detection procedures". *Hacettepe University Journal of Education*, 41(36-47).

Baker, F. B. (2001). *The basics of item response theory*. ERIC Clearinghouse on Assessment and Evaluation, University of Maryland, College Park, MD.

BangkokBiz Online. (2015, 24 April). “Teacher assistant exam. Potential cheat found in the District 2 of Loei, the exam candidates carry answer into the examination room”. Retrieved 5 June 2017, from <http://www.kruwande.com/news-id28047.html>

Cizek, G. J. (1999). *Cheating on tests: How to do it, detect it, and prevent it*. Mahway, NJ: Lawrence Erlbaum Associates.

Cizek, G. J. (2001). "An overview of issues concerning cheating on large-scale tests". Paper presented at the annual meeting of the National Council on Measurement in Education, April, 2001, Seattle, Washington.

Cole, N. (1998). "Teen cheating hurts all". *USA Today*, pp. A-24.

Deng, W. (2007). *An innovative use of the standardized log-likelihood statistic to evaluate person fit*. Unpublished Doctoral Dissertation, Rutgers, The State University of New Jersey, New Jersey, USA.

Dragow, F., Luecht, R. M., & Bennett, R. (2006). "Technology and testing". In R. L. Brennan (Ed.), *Educational measurement* (4 ed., pp. 471-515). Washington, DC: American Council on Education/Praeger Publishers.

Han, N. (2006). *Detecting exposed items in computer-based testing*. University of Massachusetts Amherst, Massachusetts, USA.

- Hing-fai, H. (2008). "*Stability and sensitivity of a model-based person-fit index in detecting item pre-knowledge in computerized adaptive test*". The Chinese University of Hong Kong, Hong Kong.
- Karabatsos, G. (2003). "Comparing the aberrant response detection performance of thirty-six person-fit statistics". *Applied Measurement in Education*, 16, 277-298.
- McLeod, L. D. (1998). "*Alternative methods for the detection of item preknowledge in computerized adaptive testing*". Unpublished Doctoral Dissertation, University of North Carolina at Chapel Hill, North Carolina.
- McManus, I. C., Lissauer, T., & Williams, S. E. (2005). "Detecting cheating in written medical examinations by statistical analysis of similarity of answers: pilot study". *British Medical Journal*, 330(7499), 1064-1066.
- Komchadluek Online. (2012). "*Police Exam cheating*". Retrieved 5 June 2017, from <http://www.komchadluek.net/detail/20120614/132748/โกงเข้าตำรวจ.html>
- Nowell, C., & Laufer, D. (1997). "Undergraduate student cheating in the fields of business and economics". *The Journal of Economic Education*, 28(1), 3-12.
- Shu, Z. (2010). "*Detecting test cheating using a deterministic, gated item response theory model*". Ph.D. dissertation, The University of North Carolina at Greensboro, USA.
- Shu, Z., Henson, R., & Luecht, R. (2013). "Using deterministic, gated item response theory model to detect test cheating due to item compromise". *Psychometrika*, 78(3), 481-497.
- Sirichai Kanjanawasee. (2012). "*Modern test theory (4th Ed)*". Bangkok: Chulalongkorn Univeristy Press.