



วารสารอิเล็กทรอนิกส์
ทางการศึกษา

OJED, Vol. 11, No. 4, 2016, pp. 743 - 760

OJED

An Online Journal
of Education

<http://www.edu.chula.ac.th/ojed>

การเปรียบเทียบความเที่ยงของแบบสอบวัดความสามารถในการแก้ปัญหอย่างสร้างสรรค์
ทางคณิตศาสตร์: การประยุกต์ใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด
COMPARISON OF TEST RELIABILITY FOR MEASURING MATHEMATICS CREATIVE PROBLEM SOLVING
ABILITY: AN APPLICATION OF GENERALIZABILITY THEORY

นางสาวชนิสรา สงวนไว*

Chanisara Sanguanwai

ศ.ดร.ศิริชัย กาญจนาวาสี **

Prof. Sirichai Kanjanawasee, Ph.D.

บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อ (1) ประเมินค่าความแปรปรวนขององค์ประกอบที่มีผลต่อสัมประสิทธิ์การสรุปอ้างอิง (G-Coefficient) ของแบบสอบวัดความสามารถในการแก้ปัญหอย่างสร้างสรรค์ทางคณิตศาสตร์ (2) ประเมินค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนแบบสอบวัดความสามารถในการแก้ปัญหอย่างสร้างสรรค์ทางคณิตศาสตร์ที่มีวิธีการออกแบบโดยผู้ตรวจตรวจให้คะแนนทุกข้อของผู้สอบทุกคน (pxixr) และผู้ตรวจตรวจข้อสอบเฉพาะข้อของผู้สอบทุกคน (px(i:r)) และ (3) เปรียบเทียบค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนแบบสอบวัดความสามารถในการแก้ปัญหอย่างสร้างสรรค์ทางคณิตศาสตร์ ที่มีวิธีการออกแบบฟาเซตที่ต่างกันภายใต้จำนวนข้อและจำนวนผู้ตรวจต่างกัน โดยกลุ่มตัวอย่าง เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 จำนวน 120 คน เครื่องมือที่ใช้ในการวิจัย คือ แบบสอบวัดความสามารถในการแก้ปัญหอย่างสร้างสรรค์ทางคณิตศาสตร์ สถิติที่ใช้ในงานวิจัย ได้แก่ ค่าสถิติพื้นฐาน, สัมประสิทธิ์สหสัมพันธ์ของเพียร์สัน, คุณภาพของผู้ตรวจ (intra-rater & inter-rater reliability) และ สัมประสิทธิ์การสรุปอ้างอิง (G-Coefficient)

ผลการวิจัยสรุปได้ดังนี้ (1) ผลการประมาณค่าความแปรปรวนพบว่า ความแปรปรวนของคะแนนปฏิสัมพันธ์ระหว่างผู้สอบกับข้อสถานการณ์ (σ^2_{pi}) ส่งผลต่อค่าสัมประสิทธิ์การสรุปอ้างอิงมากที่สุด (2) ผลการประมาณค่าสัมประสิทธิ์การสรุปอ้างอิงพบว่าเมื่อจำนวนข้อสถานการณ์และจำนวนผู้ตรวจเพิ่มขึ้น ส่งผลต่อค่าสัมประสิทธิ์การสรุปอ้างอิงสำหรับการตัดสินใจเชิงสัมพัทธ์ (Relative coefficient) และค่าสัมประสิทธิ์การสรุปอ้างอิงสำหรับการตัดสินใจเชิงสัมบูรณ์ (absolute coefficient) ที่สูงขึ้น (3) ผลการเปรียบเทียบค่าสัมประสิทธิ์การสรุปอ้างอิงพบว่า การออกแบบโดยผู้ตรวจตรวจเฉพาะข้อของผู้สอบทุกคนมีค่าความเที่ยงสูงกว่าผู้ตรวจตรวจทุกข้อของผู้สอบทุกคนและควรเลือกใช้สถานการณ์ pxixr ที่มีผู้ตรวจ 1 คน ตรวจให้คะแนน 7 ข้อ หรือ px(i:r) ที่มีผู้ตรวจ 2 คน ตรวจให้คะแนน 5 ข้อ เพื่อให้มีค่าสัมประสิทธิ์การสรุปอ้างอิงสำหรับการตัดสินใจเชิงสัมพัทธ์ (ρ^2_g) มากกว่า 0.80

* นิสิตมหาบัณฑิตสาขาวิชาการวัดและประเมินผลการศึกษา ภาควิชาวิจัยและจิตวิทยาการศึกษา

คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

E-mail Address: baby_fruit32@hotmail.com

** อาจารย์ประจำสาขาวิชาการวัดและประเมินผลการศึกษา ภาควิชาวิจัยและจิตวิทยาการศึกษา

คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

E-mail Address: Skanjanawasee@hotmail.com

ISSN1905-4491

Abstract

The purposes of this study were 1) to estimate variance component for mathematics creative problem solving ability test in Generalizability Study 2) to estimate G-coefficient in Two-Facet Crossed Design or pxixr design and Two-Facet Nested Design or px(i:r) design and 3) to compare the test reliability for measuring mathematics creative problem solving ability using the different design in D-Study the different number of item situation and the rater used the application to Generalizability Theory. Sample were 120 students in the ninth grade. Research instruments were mathematics creative problem solving ability test. Basic Statistics, Pearson's Product Moment Correlation intra rater & inter rater and G-Coefficient were used to analyze the data. It was found that 1) The highest of the estimated variance components is σ^2_{pi} . 2) Both of the Two-Facet Crossed Design and Two-Facet Nested Design have more G-Coefficient there were more numbers of items situation and the raters. 3) The result of the comparison of the test reliability for measuring mathematics creative problem solving ability found that the using under the different design in D-Study effected to the reliability of the test. The Two-Facet Nested Design or px(i:r) design has G-Coefficient higher than Two-Facet Crossed Design or pxixr design for example ; 1 raters for 7 items in pxixr design and 2 raters for 5 items in px(i:r) design for the least which we has the G-Coefficient for Relative Designs (ρ^2_{δ}) > 0.80 .

คำสำคัญ: ความสามารถในการแก้ปัญหาอย่างสร้างสรรค์ทางคณิตศาสตร์ / ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด

KEYWORDS: MATHEMATICS CREATIVE PROBLEM SOLVING ABILITY/ GENERALIZABILITY THEORY

บทนำ

การพัฒนาคนเป็นสิ่งสำคัญในการขับเคลื่อนประเทศให้มีความเจริญก้าวหน้าดังนั้นระบบการศึกษาที่มีคุณภาพและมีประสิทธิภาพจะช่วยพัฒนาศักยภาพและความสามารถของคนรวมถึงคุณลักษณะต่าง ๆ ให้มีคุณภาพด้วย การปลูกฝังให้เยาวชนเติบโตเป็นบุคคลที่มีคุณภาพและอยู่ในสังคมได้จะต้องมีการจัดการศึกษาที่พัฒนาให้เยาวชนรู้จักคิดและแก้ปัญหา ซึ่งทักษะการแก้ปัญหาเป็นส่วนหนึ่งของคุณลักษณะอันพึงประสงค์ที่สำคัญต่อการพัฒนาให้เกิดขึ้นกับผู้เรียนตามจุดมุ่งหมายของการศึกษาตามพระราชบัญญัติการศึกษาแห่งชาติ พ.ศ. 2542 แก้ไขเพิ่มเติม (ฉบับที่ 2) พ.ศ. 2545 และ (ฉบับที่ 3) พ.ศ. 2553ซึ่งกล่าวไว้ในมาตรา 24 ว่าให้สถานศึกษาและหน่วยงานที่เกี่ยวข้องดำเนินการฝึกทักษะ กระบวนการคิด การจัดการการเผชิญสถานการณ์ และประยุกต์ความรู้มาใช้เพื่อป้องกันและแก้ไขปัญหาจึงเห็นได้ว่าความสามารถในการแก้ปัญหาเป็นองค์ประกอบที่สำคัญของการดำเนินชีวิตอย่างมีคุณภาพของเยาวชน อีกทั้งความก้าวหน้าทั้งทางด้านเศรษฐกิจ สังคม การเมือง การปกครอง และด้านการศึกษา ความสามารถในการแก้ปัญหาแบบดั้งเดิมไม่สามารถแก้ปัญหาที่มีความเฉพาะเจาะจงหรือปัญหาที่มีความซับซ้อนมากได้ ดังนั้นความจำเป็นในการแก้ปัญหาย่างสร้างสรรค์จึงถือเป็นแนวทางการพัฒนาคน เพื่อให้มีการสร้างแนวคิดในการแก้ปัญหาได้หลากหลายแปลกใหม่

หรือมีความคิดสร้างสรรค์ในการคิดแก้ปัญหา ได้แก่ ความคิดยืดหยุ่น ความคิดริเริ่มและความคิดละเอียดลออ สามารถคิดวิเคราะห์และตัดสินใจเลือกแนวทางและวิธีการที่ดีที่สุดในการแก้ปัญหาให้มีประสิทธิภาพมากยิ่งขึ้น (ภัทรภร แสงไชย, 2551; อรรวรรณ ต้นสุวรรณรัตน์, 2552; อาพันธ์ชนิต เจนจิต, 2546)

ทฤษฎีและแนวคิดเกี่ยวกับการแก้ปัญหาย่างสร้างสรรค์ได้มีการศึกษาอย่างต่อเนื่อง ซึ่งผู้ที่มีชื่อเสียงในการคิดค้นคือ Treffinger et al.(2005) ได้เสนอองค์ประกอบและขั้นตอนในการแก้ปัญหาย่างสร้างสรรค์อย่างต่อเนื่องซึ่งปัจจุบันแผนกระบวนการแก้ปัญหาย่างสร้างสรรค์เป็นเวอร์ชันที่ 6.1 ซึ่งเป็นวิธีการแก้ปัญหามาโดยอาศัยการจินตนาการและการคิดพิจารณาอย่างละเอียดขณะเดียวกัน Treffinger et al.(2005)ได้ให้นิยามการแก้ปัญหาย่างสร้างสรรค์ว่าเป็นวิธีหนึ่งที่ใช้ในการพัฒนาศักยภาพด้านความคิดสร้างสรรค์ของนักเรียนโดยผ่านกรอบการทฤษฎีทางด้านการวิเคราะห์การแก้ปัญหา ซึ่งมีขั้นตอนโดยสรุปคือทำความเข้าใจกับปัญหา ตัดสินปัญหา หาแนวทางแก้ไข ประเมินข้อมูลและหาวิธีการแก้ปัญหาในส่วนของ Arbesman and Paccio (2001) กล่าวว่า การแก้ปัญหาย่างสร้างสรรค์เป็นการทำงานร่วมกันระหว่างความคิดสร้างสรรค์กับการคิดอย่างมีวิจารณ์ญาณ

การเรียนการสอนคณิตศาสตร์มีจุดมุ่งหมายเพื่อให้นักเรียนสามารถคิดแก้ปัญหาได้ โดยเฉพาะการแก้ปัญหาย่างสร้างสรรค์ (Creative Problem Solving) เป็นกรอบแนวคิดที่ใช้ทักษะการคิดและเครื่องมือการคิดเพื่อหาวิธีการคิดเพื่อค้นหาวิธีการหาคำตอบหรือวิธีการแก้ปัญหามากมาย มีโครงสร้างของกระบวนการที่ใช้จินตนาการ การวิเคราะห์ ความคิดสร้างสรรค์เพื่อนำไปใช้ในกระบวนการแก้ปัญหาย่างสร้างสรรค์ (อาพันธ์ชนิต เจนจิต, 2546) โดยแนวคิดของ Treffinger, Isaksen และ Dorval (2000) ได้ถูกนำมาใช้ในกระบวนการแก้ปัญหาย่างสร้างสรรค์ทางคณิตศาสตร์ (ภัทรภร แสงไชย, 2551; อรรวรรณ ต้นสุวรรณรัตน์, 2552) ซึ่งมีขั้นตอน 4 ขั้นตอน คือ ทำความเข้าใจปัญหา (Understanding the Challenge) สร้างแนวคิดที่หลากหลายในการแก้ปัญหา (Generating Ideas)เตรียมการเพื่อนำไปสู่การปฏิบัติ (Preparing for Action) และวางแผนแล้วปฏิบัติตามแนวคิด (Planning Your Approach) ซึ่งอดีตการศึกษาเกี่ยวกับความคิดสร้างสรรค์ทางคณิตศาสตร์เริ่มต้นจาก Balka (1974) ที่สร้างปัญหาทางคณิตศาสตร์ขึ้นเพื่อตอบปัญหาในสถานการณ์โลกจริง และในเวลาต่อมาได้มีการพัฒนาแบบสอบวัดความสามารถในการแก้ปัญหาย่างสร้างสรรค์ (Mathematics Creative Problem Solving Ability Test: MCPSAT) ในประเทศเกาหลี (Kim et al., 2003; Lee, Hwang, & Seo, 2003)โดยการออกแบบสอบเป็นข้อถามปลายเปิด (open-ended problems) จำนวน 5 ข้อเพื่อทดสอบความสามารถในการแก้ปัญหาย่างสร้างสรรค์ทางคณิตศาสตร์กับนักเรียนพบว่ามีค่าความเที่ยง .80 และมีการพิจารณาค่า internal validity ต่อมาKim et al. (2003)ได้พัฒนาข้อคำถามปลายเปิดจำนวน 8 ข้อ พบว่ามีการตรวจสอบคุณภาพด้านความเที่ยงของแบบสอบโดยการหาค่าความสอดคล้องภายใน (internal consistency) ด้วยวิธีสัมประสิทธิ์แอลฟาของครอนบาค (Cronbach's alpha method) มีค่า .73 โดยในการพิจารณาการตรวจให้คะแนนประกอบด้วยการให้คะแนนความคิด 3 ส่วนคือ ความคิดคล่องแคล่ว โดยพิจารณาจากจำนวนในการตอบ ความคิดยืดหยุ่น พิจารณาจากการนับประเภทของคำตอบและคิดริเริ่ม พิจารณาจากคำตอบที่แปลก

จากการที่กล่าวมาข้างต้นการทดสอบการแก้ปัญหาและความคิดสร้างสรรค์ทางคณิตศาสตร์ จำเป็นต้องให้นักเรียนแสดงความคิดและเขียนแสดงวิธีการคิด กระบวนการแก้ปัญหาที่หลากหลาย จึงจำเป็นต้องมีการสร้างข้อสอบแบบสร้างคำตอบ เพราะเป็นแบบสอบที่ให้อิสระในการแสดงความคิด ส่งเสริมการจัดระเบียบความคิด การประยุกต์ความรู้กับสถานการณ์ใหม่และความคิดสร้างสรรค์ ได้ดีกว่าข้อสอบรูปแบบอื่น อีกทั้งสร้างได้ง่าย และเป็นเครื่องมือที่มีความตรงเชิงประจักษ์ในด้านการวัดความสามารถใช้ภาษา ในการเขียนได้เป็นอย่างดี (Hopkins & Ante,1990; ศิริชัย กาญจนวาสี, 2555)การสร้างข้อสอบแบบสร้างคำตอบจึงเป็นทางเลือกหนึ่งที่ปัจจุบันนิยมนำมาใช้วัดประเมิน ซึ่งเป็นการเปิดโอกาสให้ผู้เรียนแสดงออกได้หลากหลาย ทั้งคำตอบวิธีการในการแก้ปัญหา ซึ่งจะเป็นการกระตุ้นให้นักเรียนสนใจในสถานการณ์สามารถแก้ปัญหาด้วยตนเอง โดยการพัฒนาวิธีการหาคำตอบและสื่อสารวิธีการแก้ปัญหาของตนซึ่งจะต้องมีการแสดงวิธีการได้มาซึ่งคำตอบ จึงเป็นการเปิดโอกาสให้ผู้เรียนได้แสดงทักษะการคิดคำนวณ การให้เหตุผล การสื่อสาร การแก้ปัญหา และแสดงความคิดสร้างสรรค์ แม้ข้อสอบแบบสร้างคำตอบจะเป็นแบบทดสอบที่มีคุณภาพ แต่ยังคงมีปัญหาในเรื่องของความเที่ยงต่ำ สิ้นเปลืองเวลา แรงงานและค่าใช้จ่ายในการตรวจมาก (Chui & Wolfe, 2002; Coffman, 1971)และจุดอ่อนที่สำคัญที่สุดคือการให้คะแนน (บุญเชิดภิญโญอนันตพงษ์,2547) ความเป็นอัตนัยของการให้คะแนน (subjective)คะแนนที่ได้จึงแปรเปลี่ยนไปตามลักษณะของผู้ตรวจ (Bachman, Lynch & Mason,1993; Linacre, 1993; Turner, 2003; Lunz, Wright&Linacre,1990; Smith & Kulikowich, 2004อ้างถึงในน้ำผึ้ง อินทเนตร (2554))ซึ่งสอดคล้องกับMehrens and Lehmann (1973)ที่กล่าวว่าเพื่อให้การตรวจแบบสอบมีความเที่ยงเพิ่มขึ้นผู้ตรวจควรใช้วิธีการตรวจที่เหมาะสมใช้เกณฑ์ การตรวจกับนักเรียนทุกคน มีการออกแบบกฎเกณฑ์ในการให้คะแนน (scoring rubric)และควบคุมแหล่ง ความคลาดเคลื่อนเช่น ข้อสอบ ผู้ตรวจ โดยใช้ผู้ตรวจหลายๆคนในการตรวจข้อสอบแต่ละข้อแต่ละข้อแล้วจึง หาคะแนนเฉลี่ย จากผลการวิจัยพบว่า เมื่อเพิ่มจำนวนผู้ตรวจจะให้ความเที่ยงสูงขึ้น (Brennan Gao & Colton, 1995; Sudweeks, Reeve & Bradshaw,2005;Swartz.et al,1999)แต่จำนวนผู้ตรวจที่เหมาะสม ยังไม่สามารถสรุปได้แน่ชัดเนื่องจากผลการศึกษายังพบข้อแตกต่าง

ด้วยเหตุนี้เพื่อให้คะแนนจากแบบสอบวัดความสามารถในการแก้ปัญหาย่างสร้างสรรค์ทางคณิตศาสตร์ที่ผู้วิจัยพัฒนาขึ้นสามารถแปลผลการตรวจแบบสอบความเรียงมีความเที่ยงสูงขึ้นอาจทำได้โดยการเพิ่มจำนวนข้อทดสอบให้มีมากขึ้น แต่การเพิ่มข้อสอบหรือการเพิ่มจำนวนผู้ตรวจเป็นสิ่งที่ต้องพิจารณาให้รอบคอบทั้งในแง่ทฤษฎีและปฏิบัติ และควรศึกษาเพื่อใช้เป็นสารสนเทศประกอบการตัดสินใจทางการศึกษาดังนั้นเพื่อเป็นการเสนอข้อมูลสามารถวิเคราะห์การให้คะแนนได้ตรงกับความต้องการมากที่สุด ดังนั้นทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด (Generalizability Theory)ซึ่งพัฒนาโดย Cronbach et al. (1972) อ้างถึงในศิริชัย กาญจนวาสี (2555)จึงมีความเหมาะสมในการวิเคราะห์ความแปรปรวนของคะแนน ความคลาดเคลื่อนจากแหล่งต่างๆ ซึ่งเป็นสถานการณ์ของการวัดที่เป็นไปได้ ทำให้สามารถตรวจสอบถึงแหล่งความแปรปรวนที่ไม่พึงประสงค์ และสามารถควบคุมแหล่งความคลาดเคลื่อนได้ตรงประเด็น (Shavelson & Webb, 1991) จึงมีความเหมาะสมกับการวัดประเมินที่มีความคลาดเคลื่อนจากผู้ตรวจให้คะแนนในการตรวจข้อสอบ

ความเรียง (ศิริชัย กาญจนวาสี, 2555) ในการวิเคราะห์เพื่อจะได้เป็นสารสนเทศและสามารถนำไปใช้ได้อย่างมีประสิทธิภาพมากที่สุด

วัตถุประสงค์

1. เพื่อประมาณค่าความแปรปรวนขององค์ประกอบที่มีผลต่อสัมประสิทธิ์การสรุปอ้างอิง (G-Coefficient) ของแบบสอบวัดความสามารถในการแก้ปัญหาอย่างสร้างสรรค์ทางคณิตศาสตร์
2. เพื่อประมาณค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนแบบสอบวัดความสามารถในการแก้ปัญหาอย่างสร้างสรรค์ทางคณิตศาสตร์ ที่มีวิธีการออกแบบการตรวจให้คะแนน 2 ลักษณะ คือ
 - 1) ผู้ตรวจตรวจให้คะแนนทุกข้อของผู้สอบทุกคน (Two-Facet Crossed Design)
 - 2) ผู้ตรวจตรวจข้อสอบบางข้อของผู้สอบทุกคน (Two-Facet Nested Design)
3. เพื่อเปรียบเทียบค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนแบบสอบวัดความสามารถในการแก้ปัญหาอย่างสร้างสรรค์ทางคณิตศาสตร์ มีวิธีการออกแบบฟาเซตที่ต่างกัน ภายใต้จำนวนข้อและจำนวนผู้ตรวจต่างกันตามทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด

วิธีดำเนินการวิจัย

ประชากรที่ใช้ในการวิจัยครั้งนี้ คือ นักเรียนระดับชั้นมัธยมศึกษาปีที่ 3 ที่ศึกษาอยู่ในโรงเรียนสังกัดสำนักงานเขตพื้นที่การศึกษามัธยมศึกษาเขต 1 ปีการศึกษา 2558 จำนวน 21,371 คน

ตัวอย่างที่ใช้ในการวิจัยครั้งนี้ คือ นักเรียนระดับชั้นมัธยมศึกษาปีที่ 3 ภาคเรียนที่ 2 ปีการศึกษา 2558 ของโรงเรียนสังกัดสำนักงานเขตพื้นที่การศึกษามัธยมศึกษาเขต 1 กรุงเทพมหานคร จำนวน 120 คน การกำหนดขนาดตัวอย่างจะต้องคำนึงถึงความเหมาะสม ซึ่งสามารถนำไปใช้ประมาณค่าสัมประสิทธิ์การสรุปอ้างอิงได้อย่างมั่นใจ (Smith, 1978 อ้างถึงใน สุพัฒน์ สุกมลสันต์, 2540) ได้เสนอขนาดตัวอย่างขั้นต่ำที่ใช้ในการศึกษาทั้ง 3 องค์ประกอบ คือ ควรมีอย่างน้อย 1,000 ค่า ดังนั้นผู้วิจัยจึงข้อกำหนดจำนวนผู้สอบ 120 คน จำนวนข้อสถานการณ์ 3 ข้อ และจำนวนผู้ตรวจ 3 คน ได้ขนาดของตัวอย่าง ค่าซึ่งเป็นไปตามกฎเกณฑ์ของ Smith (1978) จะได้ขนาดตัวอย่าง 120 คน โดยมีวิธีการสุ่มตัวอย่างแบบสุ่มหลายขั้นตอน (Multi-stage random sampling) เริ่มจากการสุ่มโรงเรียนจากสำนักงานเขตพื้นที่การศึกษามัธยมศึกษาเขต 1 ตามขนาดโรงเรียนโดยมาจากโรงเรียนขนาดใหญ่พิเศษ จำนวน 3 โรงเรียน ขนาดใหญ่ จำนวน 2 โรงเรียน และโรงเรียนขนาดกลาง จำนวน 1 โรงเรียนละ 20 คน

เครื่องมือที่ใช้ในการวิจัยในครั้งนี้ ประกอบด้วย 1) แบบสอบการแก้ปัญหาอย่างสร้างสรรค์ทางคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 เป็นแบบสอบความเรียง จำนวน 3 ข้อสถานการณ์ โดยแต่ละข้อสถานการณ์มีข้อความย่อยจำนวน 5 ข้อ การสร้างข้อคำถามจะสอดคล้องกับกระบวนการแก้ปัญหาอย่างสร้างสรรค์ตามแนวคิดของ Treffinger, Selby และ Isaksen (2000) โดยข้อคำถามในแต่ละสถานการณ์จะวัดกระบวนการแก้ปัญหาอย่างสร้างสรรค์ 4 ขั้นตอนคือ 1) ขั้นตอนการทำความเข้าใจปัญหา 2) ขั้นตอนการสร้าง

แนวคิดที่หลากหลายในการแก้ปัญหา3)ขั้นการเตรียมการเพื่อนำไปสู่การปฏิบัติ4)ขั้นการวางแผนแล้วปฏิบัติตามแนวคิดโดยในข้อคำถามที่ 1 วัดความสามารถในการรวบรวมข้อมูล ข้อคำถามที่ 2วัดความสามารถในการระบุปัญหา ข้อคำถามที่ 3 วัดความสามารถในการหาวิธีการที่หลากหลายในการแก้ปัญหาให้ตรงกับสาเหตุของปัญหา ข้อคำถามที่ 4 วัดความสามารถในการวางแผนในการเตรียมการเพื่อนำไปสู่การปฏิบัติ และข้อคำถามที่ 5วัดความสามารถในการวางแผนแล้วปฏิบัติตามแนวคิด 2) คู่มือการตรวจให้คะแนนแบบวิเคราะห์ย่อย (Analytic method) มีองค์ประกอบในการให้คะแนน 3 องค์ประกอบคือ ความถูกต้อง (2 คะแนน) การเข้าใจในสิ่งที่ต้องการวัด (2 คะแนน) และการจัดเรียบเรียงความคิด (1 คะแนน)ในแต่ละข้อสถานการณ์มีคะแนนเต็ม 30 คะแนน และทั้งฉบับจะมีคะแนนเต็ม 90 คะแนน

การวิเคราะห์ข้อมูล

1.การวิเคราะห์ค่าสถิติพื้นฐานคะแนนแบบสอบวัดความสามารถในการแก้ปัญหาอย่างสร้างสรรค์ทางคณิตศาสตร์ ที่ตรวจโดยผู้ตรวจจำนวน 3 คน คือ 1) ผู้วิจัย 2)ครูที่ปฏิบัติงานสอนในกลุ่มสาระการเรียนรู้คณิตศาสตร์ที่มีประสบการณ์ทางการสอนมากกว่า 10 ปีหรือครูที่มีตำแหน่งวิทยฐานะชำนาญการ (คศ.2) จำนวน 2 คน โดยผู้ตรวจทั้ง 3 คนจะทำการตรวจแบบสอบทุกฉบับอย่างเป็นอิสระต่อกัน และตรวจแบบสอบทั้งหมดจำนวน 2 ครั้ง ห่างกันครั้งละ 4 สัปดาห์และคะแนนที่ได้เกิดจากคะแนนเฉลี่ยของผู้ตรวจคนที่ 1 คนที่ 2 และคนที่ 3มาวิเคราะห์หาค่าสถิติพื้นฐาน โดยใช้สถิติบรรยาย คือ คะแนนเฉลี่ย (M)ส่วนเบี่ยงเบนมาตรฐาน (SD)คะแนนสูงสุด (Max)คะแนนต่ำสุด (Min)ของแต่ละสถานการณ์

2. การหาค่าความเที่ยงในการตรวจให้คะแนนจากผู้ตรวจคนเดียว (intra-rater reliability)คือผู้ตรวจที่ทำการตรวจให้คะแนนแบบสอบจำนวน 120 ชุด จำนวน 2 ครั้ง ห่างกันครั้งละ 4 สัปดาห์ ด้วยวิธีการตรวจให้คะแนนแบบวิเคราะห์ย่อย มาหาความสัมพันธ์โดยใช้ค่าสัมประสิทธิ์ของเพียร์สัน (Pearson's Product Moment Correlation) และหาค่าความเที่ยงในการตรวจให้คะแนนจากผู้ตรวจจำนวน 2 คน (inter-rater reliability) คือ 1) ผู้วิจัย 2) ครูที่ปฏิบัติงานสอนในกลุ่มสาระการเรียนรู้คณิตศาสตร์ที่มีประสบการณ์มากกว่า 10 ปีหรือครูที่มีตำแหน่งวิทยฐานะชำนาญการขึ้นไป (คศ.2)ซึ่งเป็นผู้ตรวจคนที่ 2 และ 3 มาจับคู่หาความสัมพันธ์โดยใช้ค่าสัมประสิทธิ์สหสัมพันธ์ของเพียร์สัน (Pearson's Product Moment Correlation)

3. การประมาณค่าความแปรปรวนขององค์ประกอบที่มีผลต่อสัมประสิทธิ์การสรุปอ้างอิง (G-Coefficient) ของแบบสอบวัดความสามารถในการแก้ปัญหาอย่างสร้างสรรค์ที่มีจำนวนข้อสถานการณ์และจำนวนผู้ตรวจต่างกันโดยใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด โดยนำคะแนนที่ได้มาวิเคราะห์ 3-way ANOVA หรือ pxiir design สำหรับ G-Study

4. การประมาณค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนแบบสอบวัดความสามารถในการแก้ปัญหาอย่างสร้างสรรค์ทางคณิตศาสตร์ เพื่อการตัดสินใจการสรุปอ้างอิง (D-Study) ที่มีวิธีการออกแบบการตรวจให้คะแนนแบบผู้ตรวจตรวจข้อสอบทุกข้อของผู้สอบทุกคน (Two-Facet Crossed Design) และวิธีการออกแบบ

การตรวจให้คะแนนแบบผู้ตรวจตรวจสอบบางข้อของผู้สอบทุกคน (Two-Facet Nested Design) โดยมีเงื่อนไขในการวัด 2 เงื่อนไข คือ จำนวนข้อสถานการณ์ ซึ่งแบบสอบแต่ละฉบับจะมีข้อสถานการณ์ ดังนี้ คือ 3, 5, 7, 9, 11, 13 และ 15 สถานการณ์ และจำนวนผู้ตรวจ 1, 2, 3 และ 4 คน

5. การเปรียบเทียบค่าสัมประสิทธิ์การสรุปอ้างอิง (G-Coefficient) ของคะแนนแบบสอบวัดความสามารถในการแก้ปัญหาอย่างสร้างสรรค์ทางคณิตศาสตร์ มีวิธีการออกแบบฟาเซตที่ต่างกัน ภายใต้จำนวนข้อสถานการณ์ (i) และจำนวนผู้ตรวจ (r) ต่างกันตามทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัดโดยเปรียบเทียบความแตกต่างค่าสัมประสิทธิ์การสรุปอ้างอิง (G-Coefficient) ด้วยวิธีของ Pittman

ผลการวิจัย

1. สรุปผลการประมาณค่าความแปรปรวนขององค์ประกอบที่มีต่อผลสัมประสิทธิ์การสรุปอ้างอิง (G-coefficient) ของคะแนนแบบสอบวัดความสามารถในการแก้ปัญหาอย่างสร้างสรรค์ทางคณิตศาสตร์ ที่มีวิธีการออกแบบ 2 ฟาเซตได้แก่ จำนวนข้อและจำนวนผู้ตรวจที่ต่างกัน ตามทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด

1.1 ผลการวิเคราะห์ค่าสถิติพื้นฐานของแบบสอบวัดความสามารถในการแก้ปัญหาอย่างสร้างสรรค์ทางคณิตศาสตร์ที่ตรวจด้วยวิธีการตรวจให้คะแนนแบบวิเคราะห์ย่อย (analytic method) ที่ได้จากคะแนนเฉลี่ยของผู้ตรวจคนที่ 1 ตรวจให้คะแนน 2 ครั้ง โดยภาพรวมพบว่า คะแนนเฉลี่ยทั้งฉบับอยู่ในระดับร้อยละ 50 ของคะแนนเต็ม และผลการวิเคราะห์ค่าสถิติพื้นฐานของแบบสอบวัดความสามารถในการแก้ปัญหาอย่างสร้างสรรค์ที่ตรวจด้วยวิธีการตรวจให้คะแนนแบบวิเคราะห์ย่อย ที่ได้จากคะแนนเฉลี่ยของผู้ตรวจทั้ง 3 คน โดยภาพรวมพบว่า แบบสอบได้คะแนนเฉลี่ย 48.40 คะแนน ซึ่งจากข้อมูลดังกล่าวอาจสรุปได้ว่า คะแนนเฉลี่ยทั้งฉบับที่นักเรียนชั้นมัธยมศึกษาปีที่ 3 ทำได้อยู่ในระดับร้อยละ 53.78 ของคะแนนเต็ม และเมื่อพิจารณาการให้คะแนนของผู้ตรวจทั้ง 3 คน พบว่า ผู้ตรวจคนที่ 1 คนที่ 2 และคนที่ 3 ให้คะแนนคิดเป็นร้อยละ 57.01 57.62 และ 46.70 ของคะแนนเต็ม

1.2 ความเที่ยงในการตรวจให้คะแนนจากผู้ตรวจคนเดียว (intra-rater reliability) คือ ผู้วิจัย ทำการตรวจให้คะแนนแบบสอบจำนวน 120 ฉบับ จำนวน 2 ครั้ง ห่างกันครั้งละ 4 สัปดาห์ ด้วยวิธีการตรวจให้คะแนนแบบวิเคราะห์ย่อย ค่าสัมประสิทธิ์สหสัมพันธ์ของคะแนนมีนัยสำคัญทางสถิติที่ระดับ .01 ทุกข้อสถานการณ์ โดยมีค่าสัมประสิทธิ์สหสัมพันธ์อยู่ในระดับปานกลาง ถึง สูง เมื่อพิจารณาค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของผู้ตรวจที่ตรวจให้คะแนนทั้ง 2 ครั้ง มีค่าใกล้เคียงกัน

1.3 ความเที่ยงในการตรวจให้คะแนนจากผู้ตรวจจำนวน 3 คน (inter-rater reliability) คือ 1) ผู้วิจัย 2) ครูที่ปฏิบัติงานสอนในกลุ่มสาระการเรียนรู้คณิตศาสตร์ที่มีประสบการณ์การสอนมากกว่า 10 ปี หรือครูที่มีวิทยฐานะชำนาญการ (คศ.2) ขึ้นไป จำนวน 2 คน ที่ตรวจแบบสอบจำนวน 120 ฉบับ ด้วยวิธีการตรวจให้คะแนนแบบวิเคราะห์ย่อยตามคู่มือการตรวจที่ผู้วิจัยพัฒนาขึ้น เมื่อจับคู่พิจารณาค่าสัมประสิทธิ์สหสัมพันธ์ของคะแนนที่ได้จากการตรวจของผู้ตรวจคนที่ 1 และผู้ตรวจคนที่ 2 ผู้ตรวจคนที่ 1 และผู้ตรวจคน

ที่ 3 และผู้ตรวจคนที่ 2 และผู้ตรวจคนที่ 3 พบว่ามีนัยสำคัญทางสถิติที่ระดับ .01 ทุกเหตุการณ์ โดยมีค่าสัมประสิทธิ์สหสัมพันธ์อยู่ในระดับสูง และเมื่อพิจารณาค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐานของผู้ตรวจทั้ง 3 คนมีค่าใกล้เคียงกัน

1.4 การประมาณค่าความแปรปรวนขององค์ประกอบที่มีต่อผลสัมประสิทธิ์การสรุปอ้างอิง (G-coefficient) ของแบบสอบวัดความสามารถในการแก้ปัญหาอย่างสร้างสรรค์ทางคณิตศาสตร์ ที่มีวิธีการออกแบบ 2 ฟาเซต ได้แก่ จำนวนข้อสถานการณ์และจำนวนผู้ตรวจที่ต่างกัน พบว่า ความแปรปรวนของคะแนนปฏิสัมพันธ์ระหว่างผู้สอบกับข้อสถานการณ์ (σ^2_{pi}) มีผลต่อค่าสัมประสิทธิ์การสรุปอ้างอิง (G-Coefficient) มากที่สุด ต่อมาคือความแปรปรวนของผู้สอบ (σ^2_p) ซึ่งเป็นความแปรปรวนจริงของคะแนน (Universe score) และความแปรปรวนของคะแนนความคลาดเคลื่อนหรือส่วนที่เหลือ ($\sigma^2_{pir, e}$) ซึ่งเป็นความแปรปรวนของความคลาดเคลื่อนที่เหลือจาก 3 แหล่ง คือ ผู้สอบ ข้อสถานการณ์และผู้ตรวจ มีผลต่อค่าสัมประสิทธิ์การสรุปอ้างอิง (G-Coefficient) ตามลำดับ และความแปรปรวนที่มีผลต่อค่าสัมประสิทธิ์การสรุปอ้างอิง (G-Coefficient) น้อยที่สุด คือ ความแปรปรวนของคะแนนปฏิสัมพันธ์ระหว่างผู้สอบกับการตรวจให้คะแนนของผู้ตรวจแต่ละครั้ง (σ^2_{pr}) ดังที่แสดงในตาราง 1

ตาราง 1 ผลการวิเคราะห์ความแปรปรวน G-Study of pxixr ของแบบสอบวัดความสามารถในการแก้ปัญหาอย่างสร้างสรรค์ทางคณิตศาสตร์ที่มีจำนวนข้อสถานการณ์ 3 ข้อ จำนวนผู้ตรวจ 3 คน

SOURCE OF VARIANCE	DF	SS	MS	ESTIMATED VARIANCE COMPONENT	% OF TOTAL VARIANCE
P	119	14244.578	119.702	9.187	29.85
I	2	2159.906	1079.953	2.768	8.99
R	2	2438.117	1219.058	3.245	10.54
PI	238	8768.983	36.844	10.856	35.27
PR	238	1058.772	4.449	0.058	0.19
IR	4	203.311	50.828	0.388	1.26
PIR, e	476	2035.133	4.275	4.275	13.89
TOTAL	1079	30908.800		30.777	100.00

2) สรุปผลการประมาณค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนแบบสอบวัดความสามารถในการแก้ปัญหาอย่างสร้างสรรค์ทางคณิตศาสตร์ ที่มีวิธีการออกแบบการตรวจให้คะแนน 2 ลักษณะ คือ 1) ผู้ตรวจตรวจให้คะแนนทุกข้อของผู้สอบทุกคน (Two-Facet Crossed Design) 2) ผู้ตรวจตรวจข้อสอบเฉพาะข้อของผู้สอบทุกคน (Two-Facet Nested Design)

2.1 การประมาณค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนแบบสอบวัดความสามารถในการแก้ปัญหาอย่างสร้างสรรค์ทางคณิตศาสตร์ ที่มีการออกแบบโดยผู้ตรวจตรวจให้คะแนนทุกข้อของผู้สอบทุกคน (Two-Facet Crossed Design) หรือ ($p \times i \times r$) จากการสรุปผลการศึกษาเพื่อการตัดสินใจการสรุปอ้างอิง (D-Study) จำนวนข้อสถานการณ์และจำนวนผู้ตรวจแบบสอบวัดความสามารถในการแก้ปัญหาอย่างสร้างสรรค์ทางคณิตศาสตร์ที่ใช้วิธีการตรวจให้คะแนนตามแบบวิเคราะห์ย่อย (analytic method) โดยใช้

การศึกษาข้อมูลและวิเคราะห์ตามหลักของทฤษฎีการสรุปอ้างอิงที่จำนวนข้อสถานการณ์ต่างกัน (3, 5, 7, 9, 11, 13 และ 15) และจำนวนผู้ตรวจต่างกัน (1, 2, 3 และ 4) พบว่า เมื่อเพิ่มจำนวนข้อสถานการณ์และผู้ตรวจจะส่งผลต่อค่าสัมประสิทธิ์การสรุปอ้างอิงสำหรับการตัดสินใจเชิงสัมพัทธ์ (relative coefficient) และสัมประสิทธิ์การสรุปอ้างอิงสำหรับการตัดสินใจเชิงสัมบูรณ์ (absolute coefficient) ที่สูงขึ้นโดยค่าสัมประสิทธิ์การสรุปอ้างอิงเชิงตัดสินใจสัมพัทธ์ (relative coefficient) มีค่ามากกว่า 0.70 ทุกกรณี ในทุกจำนวนผู้ตรวจที่ตรวจให้คะแนนจำนวนมากกว่าหรือเท่ากับ 5 ข้อสถานการณ์ขึ้นไป และค่าสัมประสิทธิ์การสรุปอ้างอิงสำหรับการตัดสินใจเชิงสัมบูรณ์ (absolute coefficient) มีค่ามากกว่า 0.70 เมื่อจำนวนผู้ตรวจเริ่มต้นที่ 2 และ 3 คน ตรวจให้คะแนนแบบสอบที่มีจำนวนข้อสถานการณ์มากกว่าหรือเท่ากับ 7 ข้อขึ้นไป และผู้ตรวจจำนวน 4 คน ตรวจให้คะแนนแบบสอบที่มีจำนวนข้อสถานการณ์มากกว่าหรือเท่ากับ 5 ข้อขึ้นไป ตามลำดับ ดังแสดงในตาราง 2

ตาราง 2 ผลการศึกษา D (D-Study of p x i x r design)

EFFECT	ESTIMATED VARIANCE COMPONENTS IN D-STUDY														
	n_r	1	1	1	1	1	1	1	2	2	2	2	2	2	2
	n_i	3	5	7	9	11	13	15	3	5	7	9	11	13	15
$\sigma^2_{p=}$	9.1872	σ^2_p	9.1872	9.1872	9.1872	9.1872	9.1872	9.1872	9.1872	9.1872	9.1872	9.1872	9.1872	9.1872	9.1872
$\sigma^2_{i=}$	2.7682	σ^2_i	0.9227	0.5536	0.3955	0.3076	0.2517	0.2129	0.1845	0.9227	0.5536	0.3955	0.3076	0.2517	0.2129
$\sigma^2_{r=}$	3.2446	σ^2_r	3.2446	3.2446	3.2446	3.2446	3.2446	3.2446	1.6223	1.6223	1.6223	1.6223	1.6223	1.6223	1.6223
$\sigma^2_{pi=}$	10.8563	σ^2_{pi}	3.6188	2.1713	1.5509	1.2063	0.9869	0.8351	0.7238	3.6188	2.1713	1.5509	1.2063	0.9869	0.8351
$\sigma^2_{pr=}$	0.0577	σ^2_{pr}	0.0577	0.0577	0.0577	0.0577	0.0577	0.0577	0.0289	0.0289	0.0289	0.0289	0.0289	0.0289	0.0289
$\sigma^2_{ir=}$	0.3879	σ^2_{ir}	0.1293	0.0776	0.0554	0.0431	0.0353	0.0298	0.0259	0.0647	0.0388	0.0277	0.0216	0.0176	0.0149
$\sigma^2_{pir=}$	4.2755	σ^2_{pir}	1.4252	0.8551	0.6108	0.4751	0.3887	0.3289	0.2850	0.7126	0.4275	0.3054	0.2375	0.1943	0.1644
ERROR	σ^2_δ	5.1016	3.0841	2.2194	1.7390	1.4333	1.2217	1.0665	4.3602	2.6277	1.8852	1.4726	1.2101	1.0284	0.8951
VARIANCE	σ^2_Δ	9.3983	6.9599	5.9149	5.3343	4.9649	4.7091	4.5215	6.9699	4.8424	3.9306	3.4241	3.1017	2.8786	2.7149
G-COEFFICIENT	ρ^2_δ	0.6430	0.7487	0.8054	0.8408	0.8650	0.8826	0.8960	0.6782	0.7776	0.8297	0.8619	0.8836	0.8993	0.9112
	ρ^2_Δ	0.4943	0.5690	0.6083	0.6327	0.6492	0.6611	0.6702	0.5686	0.6548	0.7004	0.7285	0.7476	0.7614	0.7719
$\sigma^2_{p=}$	9.1872	σ^2_p	9.1872	9.1872	9.1872	9.1872	9.1872	9.1872	9.1872	9.1872	9.1872	9.1872	9.1872	9.1872	9.1872
$\sigma^2_{i=}$	2.7682	σ^2_i	0.9227	0.5536	0.3955	0.3076	0.2517	0.2129	0.1845	0.9227	0.5536	0.3955	0.3076	0.2517	0.2129
$\sigma^2_{r=}$	3.2446	σ^2_r	1.0815	1.0815	1.0815	1.0815	1.0815	1.0815	0.8112	0.8112	0.8112	0.8112	0.8112	0.8112	0.8112
$\sigma^2_{pi=}$	10.8563	σ^2_{pi}	3.6188	2.1713	1.5509	1.2063	0.9869	0.8351	0.7238	3.6188	2.1713	1.5509	1.2063	0.9869	0.8351
$\sigma^2_{pr=}$	0.0577	σ^2_{pr}	0.0192	0.0192	0.0192	0.0192	0.0192	0.0192	0.0144	0.0144	0.0144	0.0144	0.0144	0.0144	0.0144
$\sigma^2_{ir=}$	0.3879	σ^2_{ir}	0.0431	0.0259	0.0185	0.0144	0.0118	0.0099	0.0086	0.0323	0.0194	0.0139	0.0108	0.0088	0.0075
$\sigma^2_{pir=}$	4.2755	σ^2_{pir}	0.4751	0.2850	0.2036	0.1584	0.1296	0.1096	0.0950	0.3563	0.2138	0.1527	0.1188	0.0972	0.0822
ERROR	σ^2_δ	4.1131	2.4755	1.7737	1.3838	1.1357	0.9640	0.8380	3.9895	2.3995	1.7180	1.3394	1.0985	0.9318	0.8094
VARIANCE	σ^2_Δ	6.1604	4.1366	3.2692	2.7873	2.4807	2.2684	2.1127	5.7557	3.7837	2.9385	2.4690	2.1702	1.9633	1.8116
G-COEFFICIENT	ρ^2_δ	0.6908	0.7877	0.8382	0.8691	0.8900	0.9050	0.9164	0.6972	0.7929	0.8425	0.8728	0.8932	0.9079	0.9190
	ρ^2_Δ	0.5986	0.6895	0.7375	0.7672	0.7874	0.8020	0.8130	0.6148	0.7083	0.7577	0.7882	0.8089	0.8239	0.8353

2.2 การประมาณค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนแบบสอบวัดความสามารถในการแก้ปัญหาอย่างสร้างสรรค์ทางคณิตศาสตร์ ที่มีการออกแบบโดยผู้ตรวจตรวจข้อสอบบางข้อของผู้สอบทุกคน (Two-Facet Nested Design) หรือการออกแบบแฝง ($p \times (i : r)$) จากการสรุปผลการศึกษาเพื่อการตัดสินใจการสรุปอ้างอิง (D-Study) จำนวนข้อสถานการณ์และจำนวนผู้ตรวจแบบสอบวัดความสามารถในการแก้ปัญหาอย่างสร้างสรรค์ทางคณิตศาสตร์ที่ใช้วิธีการตรวจให้คะแนนตามแบบวิเคราะห์ย่อย (analytic

method) โดยใช้การศึกษาข้อมูลและวิเคราะห์ตามหลักของทฤษฎีการสรุปอ้างอิงที่จำนวนข้อสถานการณ์ (i) และจำนวนผู้ตรวจต่างกัน (r) พบว่า เมื่อเพิ่มจำนวนข้อสถานการณ์และจำนวนผู้ตรวจจะส่งผลต่อค่าสัมประสิทธิ์การสรุปอ้างอิงเชิงตัดสินใจสัมพัทธ์ (relative coefficient) และสัมประสิทธิ์การสรุปอ้างอิงเชิงตัดสินใจสัมบูรณ์ (absolute coefficient) ที่สูงขึ้น โดยค่าสัมประสิทธิ์การสรุปอ้างอิงสำหรับการตัดสินใจเชิงสัมพัทธ์ (relative coefficient) มีค่ามากกว่า 0.70 เมื่อจำนวนผู้ตรวจเริ่มต้นที่ 1 คนตรวจให้คะแนนแบบสอบที่มีจำนวนข้อสถานการณ์มากกว่าหรือเท่ากับ 5 ข้อขึ้นไป และผู้ตรวจจำนวน 2, 3 และ 4 คนมีค่าสัมประสิทธิ์การสรุปอ้างอิงเชิงตัดสินใจสัมพัทธ์ (relative coefficient) มากกว่า 0.70 ทุกกรณี และค่าสัมประสิทธิ์การสรุปอ้างอิงเชิงตัดสินใจสัมบูรณ์ (absolute coefficient) มีค่ามากกว่า 0.70 เมื่อจำนวนผู้ตรวจเริ่มต้นที่ 2 คน ตรวจให้คะแนนแบบสอบที่มีจำนวนข้อสถานการณ์มากกว่าหรือเท่ากับ 5 ข้อขึ้นไป และผู้ตรวจจำนวน 3 คน และ 4 คน ตรวจให้คะแนนแบบสอบที่มีจำนวนข้อสถานการณ์มากกว่าหรือเท่ากับ 3 ข้อขึ้นไป ตามลำดับ ดังแสดงในตาราง 3

ตาราง 3 ผลการศึกษา D (D-Study of px(i:r) design)

EFFECT			ESTIMATED VARIANCE COMPONENTS IN D-STUDY													
px(i:r)	px(i:r)	n _r	1	1	1	1	1	1	1	2	2	2	2	2	2	2
		n _i	3	5	7	9	11	13	15	3	5	7	9	11	13	15
$\sigma^2_p =$	9.1872	$\sigma^2_p =$	9.1872	σ^2_p	9.1872	9.1872	9.1872	9.1872	9.1872	9.1872	9.1872	9.1872	9.1872	9.1872	9.1872	9.1872
$\sigma^2_R =$	3.2446	$\sigma^2_R =$	3.2446	σ^2_R	3.2446	3.2446	3.2446	3.2446	3.2446	3.2446	1.6223	1.6223	1.6223	1.6223	1.6223	1.6223
$\sigma^2_i =$	2.7682	$\sigma^2_{(iR)} =$	3.1561	$\sigma^2_{(iR)}$	1.0520	0.6312	0.4509	0.3507	0.2869	0.2428	0.2104	0.5260	0.3156	0.2254	0.1753	0.1435
$\sigma^2_{iR} =$	0.3879															
$\sigma^2_{PR} =$	0.0577	$\sigma^2_{PR} =$	0.0577	σ^2_{PR}	0.0577	0.0577	0.0577	0.0577	0.0577	0.0577	0.0289	0.0289	0.0289	0.0289	0.0289	0.0289
$\sigma^2_{pi} =$	10.8563	$\sigma^2_{(pi,R)} =$	15.1318	$\sigma^2_{(pi,R)}$	5.0439	3.0264	2.1617	1.6813	1.3756	1.1640	1.0088	2.5220	1.5132	1.0808	0.8407	0.6878
$\sigma^2_{pR} =$	4.2755															
ERROR VARIANCE			σ^2_δ	5.1016	3.0841	2.2194	1.7390	1.4333	1.2217	1.0665	2.5508	1.5420	1.1097	0.8695	0.7167	0.6108
			σ^2_Δ	9.3983	6.9599	5.9149	5.3343	4.9649	4.7091	4.5215	4.6992	3.4800	2.9574	2.6672	2.4824	2.3545
G-COEFFICIENT			ρ^2_δ	0.6430	0.7487	0.8054	0.8408	0.8650	0.8826	0.8960	0.7827	0.8563	0.8922	0.9135	0.9276	0.9377
			ρ^2_Δ	0.4943	0.5690	0.6083	0.6327	0.6492	0.6611	0.6702	0.6616	0.7253	0.7565	0.7750	0.7873	0.7960
$\sigma^2_p =$	9.1872	$\sigma^2_p =$	9.1872	σ^2_p	9.1872	9.1872	9.1872	9.1872	9.1872	9.1872	9.1872	9.1872	9.1872	9.1872	9.1872	9.1872
$\sigma^2_R =$	3.2446	$\sigma^2_R =$	3.2446	σ^2_R	1.0815	1.0815	1.0815	1.0815	1.0815	1.0815	0.8112	0.8112	0.8112	0.8112	0.8112	0.8112
$\sigma^2_i =$	2.7682	$\sigma^2_{(iR)} =$	3.1561	$\sigma^2_{(iR)}$	0.3507	0.2104	0.1503	0.1169	0.0956	0.0809	0.0701	0.2630	0.1578	0.1127	0.0877	0.0717
$\sigma^2_{iR} =$	0.3879															
$\sigma^2_{PR} =$	0.0577	$\sigma^2_{PR} =$	0.0577	σ^2_{PR}	0.0192	0.0192	0.0192	0.0192	0.0192	0.0192	0.0192	0.0144	0.0144	0.0144	0.0144	0.0144
$\sigma^2_{pi} =$	10.8563	$\sigma^2_{(pi,R)} =$	15.1318	$\sigma^2_{(pi,R)}$	1.6813	1.0088	0.7206	0.5604	0.4585	0.3880	0.3363	1.2610	0.7566	0.5404	0.4203	0.3439
$\sigma^2_{pR} =$	4.2755															
ERROR VARIANCE			σ^2_δ	1.7005	1.0280	0.7398	0.5797	0.4778	0.4072	0.3555	1.2754	0.7710	0.5548	0.4348	0.3583	0.3054
			σ^2_Δ	3.1328	2.3200	1.9716	1.7781	1.6550	1.5697	1.5072	2.3496	1.7400	1.4787	1.3336	1.2412	1.1773
G-COEFFICIENT			ρ^2_δ	0.8438	0.8994	0.9255	0.9406	0.9506	0.9576	0.9627	0.8781	0.9226	0.9430	0.9548	0.9625	0.9678
			ρ^2_Δ	0.7457	0.7984	0.8233	0.8378	0.8474	0.8541	0.8591	0.7963	0.8408	0.8614	0.8732	0.8810	0.8864

3) สรุปผลการเปรียบเทียบค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนแบบสอบวัดความสามารถในการแก้ปัญหาอย่างสร้างสรรค์ทางคณิตศาสตร์ มีวิธีการออกแบบฟาเซตที่ต่างกัน ภายใต้จำนวนข้อและจำนวนผู้ตรวจต่างกันตามทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด

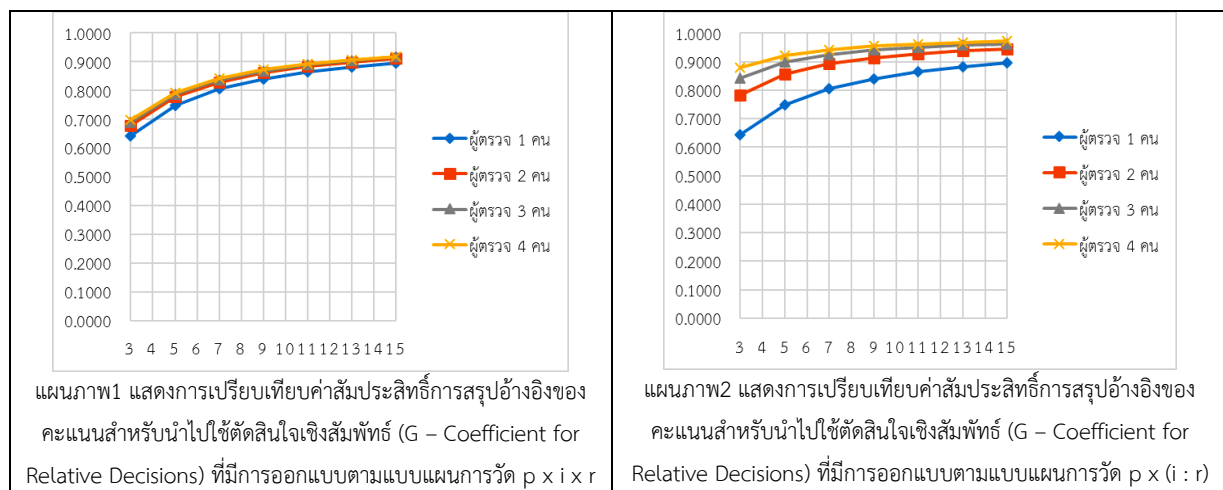
การเปรียบเทียบความแตกต่างของความเที่ยงด้วยค่าสัมประสิทธิ์การสรุปอ้างอิง (G- Coefficient) เมื่อพิจารณาเปรียบเทียบความเที่ยงของแบบสอบวัดความสามารถในการแก้ปัญหาอย่างสร้างสรรค์ทางคณิตศาสตร์ที่มีจำนวนข้อสถานการณ์และจำนวนผู้ตรวจที่เท่ากัน พบว่า การประมาณค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนแบบสอบวัดความสามารถในการแก้ปัญหาอย่างสร้างสรรค์ทางคณิตศาสตร์ ที่มีการออกแบบโดยผู้ตรวจตรวจข้อสอบบางข้อของผู้สอบทุกคน (Two-Facet Nested Design) หรือการออกแบบผสม ($p \times (i : r)$) มีค่าความเที่ยงสูงกว่าการประมาณค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนแบบสอบวัดความสามารถในการแก้ปัญหาอย่างสร้างสรรค์ทางคณิตศาสตร์ ที่มีการออกแบบโดยผู้ตรวจตรวจให้คะแนนทุก

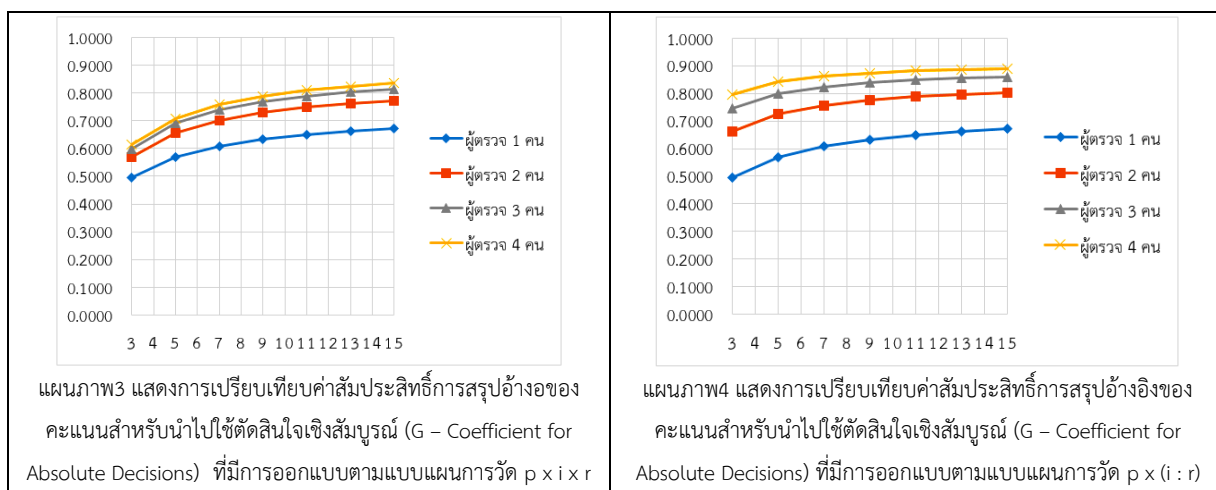
ข้อของผู้สอบทุกคน(Two-Facet Crossed Design) หรือการออกแบบไขว้ ($p \times i \times r$) ดังตาราง 4

ตาราง 4 เปรียบเทียบความเที่ยง (ค่าสัมประสิทธิ์ความน่าเชื่อถือ: G- coefficient) ของแบบสอบวัดความสามารถในการแก้ปัญหาอย่างสร้างสรรค์ทางคณิตศาสตร์ที่มีการออกแบบฟาเซตที่ต่างกัน

การออกแบบ	ค่าสัมประสิทธิ์การสรุปอ้างอิง	$r_{x_1} r_{x_2}$	t	p
$P \times (i : R) -$	ρ^2_{δ}	.807	6.530**	.000
$P \times I \times R$	ρ^2_{Δ}	.895	6.435**	.000

**p < .01





อภิปรายผลการวิจัย

จากผลการวิเคราะห์และสรุปผลการวิจัย ผู้วิจัยนำเสนอประเด็นการอภิปรายที่น่าสนใจ ดังนี้

1. ผลการประมาณค่าความแปรปรวนขององค์ประกอบที่มีต่อผลสัมประสิทธิ์การสรุปอ้างอิง (G-coefficient) ของคะแนนแบบสอบวัดความสามารถในการแก้ปัญหาอย่างสร้างสรรค์ทางคณิตศาสตร์ ที่มีวิธีการออกแบบ 2 ฟาเซตได้แก่ จำนวนข้อสถานการณ์ (i) และจำนวนผู้ตรวจ (r) ที่ต่างกัน ตามทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด สรุปผลได้ข้อค้นพบว่า ในการศึกษา G (G-study) ที่มีองค์ประกอบ 3 องค์ประกอบ ได้แก่ นักเรียนชั้นมัธยมศึกษาปีที่ 3 (p) จำนวนข้อสถานการณ์ (i) และจำนวนผู้ตรวจ (r) ที่มีการออกแบบการตรวจโดยผู้ตรวจตรวจสอบข้อสอบทุกข้อของผู้สอบทุกคน (Two-Facet Crossed Design) หรือ $p \times i \times r$ พบว่า แหล่งความแปรปรวนขององค์ประกอบที่มีผลต่อสัมประสิทธิ์การสรุปอ้างอิงคือ ความแปรปรวนของคะแนนปฏิสัมพันธ์ระหว่างผู้สอบกับข้อสถานการณ์ มีผลต่อค่าสัมประสิทธิ์การสรุปอ้างอิง (G-Coefficient) มากที่สุด ซึ่งแสดงว่าผู้สอบมีความสามารถในการทำแบบสอบวัดความสามารถในการแก้ปัญหาอย่างสร้างสรรค์ทางคณิตศาสตร์ในแต่ละข้อสถานการณ์ต่างกัน เนื่องจากในแบบสอบมีสาระในการวัด 3 สาระซึ่งแต่ละข้อสถานการณ์วัดความรู้ความเข้าใจใน 1 สาระ จึงอาจส่งผลต่อค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนแบบสอบต่อมาคือ ความแปรปรวนของผู้สอบซึ่งแสดงว่าความสามารถ ความรู้และความเข้าใจในการทำแบบสอบวัดความสามารถในการแก้ปัญหาอย่างสร้างสรรค์ทางคณิตศาสตร์ของผู้สอบหรือนักเรียนระดับชั้นมัธยมศึกษาปีที่ 3 แต่ละคนมีความแตกต่างกันจึงส่งผลต่อค่าสัมประสิทธิ์การสรุปอ้างอิงและความแปรปรวนของคะแนนความคลาดเคลื่อนหรือส่วนที่เหลือ ซึ่งเป็นความแปรปรวนของความคลาดเคลื่อนที่เหลือจาก 3 แหล่ง คือ ผู้สอบ ข้อสถานการณ์และผู้ตรวจ มีผลต่อค่าสัมประสิทธิ์การสรุปอ้างอิง (G-Coefficient) ซึ่งแสดงว่ายังมีปัจจัยอื่น ๆ นอกเหนือจากผู้สอบ จำนวนข้อสถานการณ์และจำนวนผู้ตรวจที่ส่งผลต่อค่าสัมประสิทธิ์การสรุปอ้างอิง และความแปรปรวนที่มีผลต่อค่าสัมประสิทธิ์การสรุปอ้างอิง (G-Coefficient) น้อยที่สุด คือ ความแปรปรวนของคะแนนปฏิสัมพันธ์ระหว่างผู้สอบกับการตรวจให้คะแนนของผู้ตรวจแต่ละครั้ง ซึ่งแสดงว่าผู้ตรวจมีความสามารถในการตรวจให้คะแนนแบบสอบวัดความสามารถในการแก้ปัญหาอย่างสร้างสรรค์ของ

ผู้สอบทุกคนได้ไม่แตกต่างกัน อาจเนื่องมาจากเกณฑ์การให้คะแนนมีลักษณะแบบวิเคราะห์ย่อย (analytic method) ในแต่ละประเด็นย่อยๆ ของข้อคำถาม ทำให้มีความเป็นปรนัย จากคำกล่าวของ Hopskin และ Stanley (1981) ที่ว่า การตรวจให้คะแนนโดยใช้วิธีตรวจแบบวิเคราะห์ย่อย จะให้ความเที่ยงสูงกว่าวิธีการตรวจแบบภาพรวม ทั้งนี้เนื่องจากวิธีการตรวจแบบวิเคราะห์ย่อยจะมีระบบการให้คะแนนที่ชัดเจนมีการแยกเป็นประเด็นย่อยๆ ตามองค์ประกอบ และยังมีการกำหนดน้ำหนักคะแนนในแต่ละส่วน ซึ่งข้อดีคือ คะแนนที่ได้มีความน่าเชื่อถือ การแบ่งคะแนนทำให้ง่ายต่อการตรวจทำให้ผู้ประเมินสามารถใช้เกณฑ์การให้คะแนนได้ค่อนข้างเป็นปรนัย

2. ผลการประมาณค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนแบบสอบวัดความสามารถในการแก้ปัญหาอย่างสร้างสรรค์ทางคณิตศาสตร์ ที่มีวิธีการออกแบบการตรวจให้คะแนน 2 ลักษณะ คือ ผู้ตรวจตรวจให้คะแนนทุกข้อของผู้สอบทุกคน (Two-Facet Crossed Design) และ ผู้ตรวจตรวจข้อสอบเฉพาะข้อของผู้สอบทุกคน (Two-Facet Nested Design) สรุปผลได้ข้อค้นพบดังนี้

2.1) การออกแบบฟาเซตการวัดแบบผู้ตรวจตรวจให้คะแนนทุกข้อของผู้สอบทุกคน (Two-Facet Crossed Design) หรือ $pxixr$ มีค่าสัมประสิทธิ์การสรุปอ้างอิงสำหรับการตัดสินใจเชิงสัมพัทธ์ (Relative Coefficient) มากกว่า 0.70 ขึ้นไปเมื่อผู้ตรวจจำนวน 1, 2, 3 หรือ 4 คน ตรวจให้คะแนนแบบสอบวัดความสามารถในการแก้ปัญหาอย่างสร้างสรรค์ทางคณิตศาสตร์อย่างน้อย 5 ข้อสถานการณ์ และมีค่าสัมประสิทธิ์การสรุปอ้างอิงสำหรับการตัดสินใจเชิงสัมพัทธ์ (Relative Coefficient) มากกว่า 0.80 ขึ้นไปเมื่อผู้ตรวจจำนวน 1, 2, 3 หรือ 4 คน ตรวจให้คะแนนอย่างน้อย 7 ข้อสถานการณ์ และมีค่าสัมประสิทธิ์การสรุปอ้างอิงสำหรับการตัดสินใจเชิงตัดสินใจสัมบูรณ์ (Absolute Coefficient) มากกว่า 0.70 ขึ้นไปเมื่อผู้ตรวจจำนวน 2 คน หรือ 3 คน ตรวจให้คะแนนอย่างน้อย 7 ข้อสถานการณ์ และผู้ตรวจจำนวน 4 คน ตรวจให้คะแนนอย่างน้อย 5 ข้อสถานการณ์ และค่าสัมประสิทธิ์การสรุปอ้างอิงสำหรับการตัดสินใจเชิงตัดสินใจสัมบูรณ์ (Absolute Coefficient) มากกว่า 0.80 ขึ้นไปเมื่อผู้ตรวจจำนวน 3 คน ตรวจให้คะแนนอย่างน้อย 13 ข้อสถานการณ์ และผู้ตรวจจำนวน 4 คน ตรวจให้คะแนนอย่างน้อย 11 ข้อสถานการณ์

2.2) การออกแบบฟาเซตการวัดแบบผู้ตรวจตรวจแบบสอบเฉพาะข้อของผู้สอบทุกคน (Two-Facet Nested Design) หรือ $px(i:r)$ มีค่าสัมประสิทธิ์การสรุปอ้างอิงสำหรับการตัดสินใจเชิงสัมพัทธ์ (Relative Coefficient) มากกว่า 0.70 ขึ้นไปเมื่อผู้ตรวจจำนวน 1 คน ตรวจให้คะแนนแบบสอบวัดความสามารถในการแก้ปัญหาอย่างสร้างสรรค์ทางคณิตศาสตร์อย่างน้อย 5 ข้อสถานการณ์ และผู้ตรวจจำนวน 2 คน ตรวจให้คะแนนอย่างน้อย 3 ข้อสถานการณ์ และมีค่าสัมประสิทธิ์การสรุปอ้างอิงสำหรับการตัดสินใจเชิงสัมพัทธ์ (Relative Coefficient) มากกว่า 0.80 ขึ้นไปเมื่อผู้ตรวจจำนวน 1 คน ตรวจให้คะแนนอย่างน้อย 7 ข้อสถานการณ์ ผู้ตรวจจำนวน 2 คน ตรวจให้คะแนนอย่างน้อย 5 ข้อสถานการณ์ และค่าสัมประสิทธิ์การสรุปอ้างอิงสำหรับการตัดสินใจเชิงตัดสินใจสัมบูรณ์ (Absolute Coefficient) มากกว่า 0.70 ขึ้นไปเมื่อผู้ตรวจจำนวน 2 คน ตรวจให้คะแนนอย่างน้อย 5 ข้อสถานการณ์ และค่าสัมประสิทธิ์การสรุปอ้างอิงสำหรับการตัดสินใจเชิงตัดสินใจสัมบูรณ์ (Absolute Coefficient) มากกว่า 0.80 ขึ้นไปเมื่อผู้ตรวจจำนวน 2 คน

ตรวจให้คะแนนอย่างน้อย 15 ข้อสถานการณ์และผู้ตรวจจำนวน 3 คนตรวจให้คะแนนอย่างน้อย 7 ข้อสถานการณ์ และผู้ตรวจจำนวน 4 คนตรวจให้คะแนนอย่างน้อย 5 ข้อสถานการณ์

3. ผลการเปรียบเทียบค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนแบบสอบวัดความสามารถในการแก้ปัญหาอย่างสร้างสรรค์ทางคณิตศาสตร์ มีวิธีการออกแบบฟาเซตที่ต่างกัน ภายใต้จำนวนข้อและจำนวนผู้ตรวจต่างกันตามทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด สรุปผลได้ข้อค้นพบว่า การตรวจสอบคุณภาพด้านความเที่ยงด้วยค่าสัมประสิทธิ์การสรุปอ้างอิง(G-Coefficient)ของแบบสอบวัดความสามารถในการแก้ปัญหาอย่างสร้างสรรค์ทางคณิตศาสตร์ที่มีการออกแบบฟาเซตการวัดแบบผู้ตรวจตรวจข้อสอบเฉพาะข้อของผู้สอบทุกคนมีค่าค่าสัมประสิทธิ์การสรุปอ้างอิงสำหรับการตัดสินใจเชิงสัมบูรณ์ (Absolute Coefficient) สูงกว่าการออกแบบฟาเซตการวัดแบบผู้ตรวจตรวจข้อสอบทุกข้อของผู้สอบทุกคนอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 และเมื่อเพิ่มจำนวนผู้ตรวจจะส่งผลต่อค่าสัมประสิทธิ์การสรุปอ้างอิงของแบบสอบที่สูงขึ้น ซึ่งสอดคล้องกับคำกล่าวของ Mehrens and Lehmann (1973), อ้างถึงในอุษณีย์ บัวศิริพันธ์, (2543) ซึ่งกล่าวว่า ควรตรวจให้คะแนนมากกว่า 1 คน และใช้คะแนนเฉลี่ยเป็นคะแนนความสามารถของผู้สอบ ซึ่งจะได้ผลการวัดที่มีความเที่ยงสูงขึ้น และสอดคล้องกับสุพัตน์ สุกมลสันต์ (2540) ที่กล่าวว่า การตรวจให้คะแนนข้อสอบความเรียง ผู้ตรวจที่ตรวจข้อสอบเฉพาะข้อของผู้สอบทุกคนจะทำให้ภาระงานในการตรวจน้อยกว่าการที่ต้องตรวจข้อสอบทุกข้อของทุกคน จึงมีโอกาที่จำให้ผลการตรวจมีความคงที่และสอดคล้องกันมากกว่าซึ่งสอดคล้องกับสมมติฐานที่ตั้งไว้ว่าค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนแบบสอบวัดความสามารถในการแก้ปัญหาอย่างสร้างสรรค์ทางคณิตศาสตร์ ที่มีการออกแบบโดยผู้ตรวจตรวจข้อสอบเฉพาะข้อของผู้สอบทุกคน (Two-Facet Nested Design) หรือ $px(i:r)$ จะมีค่าสูงกว่าการออกแบบการตรวจให้คะแนนแบบผู้ตรวจตรวจข้อสอบทุกข้อของผู้สอบทุกคน (Two-Facet Crossed Design) หรือ $pxixr$

ข้อเสนอแนะ

ข้อเสนอแนะในการนำผลการวิจัยไปใช้

การนำแบบสอบวัดความสามารถในการแก้ปัญหาอย่างสร้างสรรค์ทางคณิตศาสตร์ไปใช้ ผู้วิจัยเสนอแนะแนวทางไว้ ดังนี้

1. แบบสอบวัดความสามารถในการแก้ปัญหาอย่างสร้างสรรค์ที่ผู้วิจัยพัฒนาขึ้นเป็นแบบสอบที่มีความยากง่ายอยู่ในเกณฑ์ปานกลาง มีคุณภาพด้านความเที่ยงอยู่ในเกณฑ์เหมาะสม จึงเหมาะกับการนำไปใช้ในการส่งเสริมกระบวนการแก้ปัญหาอย่างสร้างสรรค์ทางคณิตศาสตร์ในนักเรียนระดับชั้นมัธยมศึกษาปีที่ 3

2. การทำแบบสอบโดยทั่วไปในรายวิชาคณิตศาสตร์ ผู้สอบจะมีความเคยชินกับรูปแบบข้อคำถามที่สอดคล้องกับการจัดการเรียนการสอนตามแนวทางการแก้ปัญหาโดยทั่วไป และบางครั้งไม่คุ้นชินกับการเขียนตอบที่เป็นลำดับขั้นตอนของกระบวนการแก้ปัญหาอย่างสร้างสรรค์ ดังนั้นเพื่อให้เกิดประสิทธิภาพในการทำแบบสอบวัดความสามารถในการแก้ปัญหาอย่างสร้างสรรค์ทางคณิตศาสตร์ ครูผู้สอนควรมีการจัดการเรียน

การสอนโดยนำแนวคิดของกระบวนการแก้ปัญหาอย่างสร้างสรรค์ทางคณิตศาสตร์ร่วมด้วย เพื่อให้นักเรียนเข้าใจวิธีการหรือรูปแบบการเขียนตอบตรงกัน

3. การตรวจให้คะแนนแบบสอบวัดความสามารถในการแก้ปัญหาอย่างสร้างสรรค์ที่ผู้วิจัยพัฒนาขึ้น ควรใช้วิธีการตรวจให้คะแนนแบบผู้ตรวจตรวจแบบสอบบางข้อของผู้สอบทุกคน (px (i:r)) เนื่องจากจะส่งผลต่อค่าความเที่ยงสูงกว่าการตรวจให้คะแนนแบบผู้ตรวจตรวจข้อสอบทุกข้อของผู้สอบทุกคน (pxixr) เนื่องจากการตรวจจำนวนข้อที่น้อยลงจะเพิ่มประสิทธิภาพการตรวจให้คะแนน และเมื่อเพิ่มจำนวนผู้ตรวจและจำนวนข้อสถานการณ์จะทำให้มีค่าความเที่ยงเพิ่มขึ้นตามลำดับ

4. การเลือกวิธีการตรวจให้คะแนนแบบผู้ตรวจตรวจข้อสอบทุกข้อของผู้สอบทุกคน (pxixr) เพื่อให้มีค่าสัมประสิทธิ์การสรุปอ้างอิงสำหรับการตัดสินใจเชิงสัมพัทธ์ (G-Coefficient for Relative Designs : ρ^2_{δ}) มากกว่า 0.80 ควรเลือกใช้ในสถานการณ์ที่มีจำนวนผู้ตรวจ 1 คน ตรวจให้คะแนนแบบสอบจำนวน 7 ข้อ สถานการณ์ (35 ข้อย่อย) และสำหรับการเลือกวิธีการตรวจให้คะแนนแบบผู้ตรวจตรวจข้อสอบเฉพาะข้อของผู้สอบทุกคน (px (i:r)) ควรเลือกใช้ในสถานการณ์ที่มีจำนวนผู้ตรวจ 2 คน ตรวจให้คะแนนจำนวน 5 ข้อ สถานการณ์ (25 ข้อย่อย) ซึ่งการพิจารณาค่าสัมประสิทธิ์การสรุปอ้างอิงสำหรับการตัดสินใจเชิงสัมพัทธ์จะทำให้สามารถพิจารณาความสามารถของผู้สอบได้ว่า ผู้สอบคนใดมีความสามารถในการแก้ปัญหาอย่างสร้างสรรค์อยู่ในระดับใด

5. การเลือกวิธีการตรวจให้คะแนนแบบผู้ตรวจตรวจข้อสอบทุกข้อของผู้สอบทุกคนเพื่อให้มีค่าสัมประสิทธิ์การสรุปอ้างอิงสำหรับการตัดสินใจเชิงสัมบูรณ์ (G-Coefficient for Relative Designs : ρ^2_{Δ}) มากกว่า 0.80 ควรเลือกใช้ในสถานการณ์ที่มีจำนวนผู้ตรวจ 3 คน ตรวจให้คะแนนแบบสอบจำนวน 13 ข้อ สถานการณ์ (65 ข้อย่อย) และสำหรับการเลือกวิธีการตรวจให้คะแนนแบบผู้ตรวจตรวจข้อสอบบางข้อของผู้สอบทุกคน ควรเลือกใช้ในสถานการณ์ที่มีจำนวนผู้ตรวจ 3 คน ตรวจให้คะแนนแบบสอบจำนวน 7 ข้อ สถานการณ์ (35 ข้อย่อย) ซึ่งการพิจารณาค่าสัมประสิทธิ์การสรุปอ้างอิงสำหรับการตัดสินใจเชิงสัมบูรณ์จะทำให้สามารถพิจารณาความสามารถของผู้สอบได้ว่า ผู้สอบคนใดมีความสามารถในการแก้ปัญหาอย่างสร้างสรรค์อยู่ในระดับสูงเพื่อใช้เป็นข้อมูลสำหรับการส่วนเข้าแข่งขันทักษะกระบวนการแก้ปัญหาอย่างสร้างสรรค์ทางคณิตศาสตร์ได้

ข้อเสนอแนะในการทำวิจัยครั้งต่อไป

1. ความสามารถในการแก้ปัญหาอย่างสร้างสรรค์เป็นส่วนหนึ่งในทักษะและกระบวนการทางคณิตศาสตร์ ดังนั้นในการศึกษาครั้งต่อไปผู้วิจัยอาจศึกษาเกี่ยวกับความสามารถในด้านอื่นๆที่มีความเกี่ยวข้องกับทักษะและกระบวนการทางคณิตศาสตร์ เช่น ความสามารถในการให้เหตุผล ความสามารถในการเชื่อมโยง เป็นต้น

2. การศึกษาครั้งนี้เป็นการศึกษาเกี่ยวกับแนวคิดความสามารถในการแก้ปัญหาอย่างสร้างสรรค์ทางคณิตศาสตร์โดยการสร้างข้อคำถามที่สอดคล้องกับกระบวนการแก้ปัญหาอย่างสร้างสรรค์ตามแนวคิดของ Treffinger, Selby and Isaksen (2000) ซึ่งในการวิจัยครั้งต่อไปอาจศึกษาโดยใช้แนวคิดหรือทฤษฎีใหม่น่าสนใจยิ่งขึ้น

3. การสร้างแบบสอบวัดความสามารถในการแก้ปัญหาอย่างสร้างสรรค์ทางคณิตศาสตร์ควรมี การศึกษาองค์ประกอบอื่นๆ ที่อาจส่งผลต่อค่าสัมประสิทธิ์การสรุปอ้างอิง เช่น ประสบการณ์ของผู้ตรวจ จำนวนมวลเนื้อหาในแต่ละสาระ จำนวนครั้งของการวัดซ้ำ หรือศึกษาเปรียบเทียบค่าสัมประสิทธิ์การสรุป อ้างอิงของการวัดที่ใช้การออกแบบในรูปแบบอื่นๆ หรือศึกษาองค์ประกอบต่างๆเพิ่มเติมยิ่งขึ้น เพื่อให้แบบ สอบวัดความสามารถในการแก้ปัญหาอย่างสร้างสรรค์ทางคณิตศาสตร์มีความเที่ยงมากยิ่งขึ้น เช่น การ เปรียบเทียบการออกแบบ $p \times i \times o$ คือ ผู้สอบทุกคนทำข้อสอบจำนวน i ข้อ จำนวน o ครั้ง กับ $p \times (i : o)$ คือ ผู้สอบทุกคนทำข้อสอบจำนวน i ข้อจำนวน o ครั้งโดยแต่ละครั้งใช้ข้อสอบต่างชุดกัน เมื่อกำหนดให้ o คือ จำนวนครั้งในการทดสอบซ้ำ

4. การศึกษาค่าสัมประสิทธิ์การสรุปอ้างอิงที่มีการออกแบบเอกภพที่มากกว่าสององค์ประกอบและมีการ ออกแบบฟาเซตที่ต่างกัน เช่น ผู้สอบ (p) จำนวนข้อ (i) จำนวนของการทดสอบซ้ำ (o) และ จำนวนผู้ตรวจ (r) ที่มีการออกแบบ $p \times (i : o) \times r$ กล่าวคือ ผู้สอบแต่ละคนทำแบบสอบ i ข้อและทำข้อสอบจำนวน o ครั้ง ซึ่ง แต่ละครั้งของการทดสอบใช้ข้อสอบต่างกัน และผู้ตรวจ r คนเป็นคนตรวจให้คะแนน เป็นต้น

รายการอ้างอิง

ภาษาไทย

- บุญเชิดภิญโญนนท์พงษ์.(2547).*การวัดและประเมินผลการเรียนรู้*. กรุงเทพฯ: คณะศึกษาศาสตร์มหาวิทยาลัยศรีนครินทรวิโรฒ.
- สำนักงานรับรองมาตรฐานและการประเมินคุณภาพการศึกษา (องค์การมหาชน). (2547).*พระราชบัญญัติ การศึกษาแห่งชาติ พ.ศ. 2542 แก้ไขเพิ่มเติม(ฉบับที่ 2)พ.ศ. 2545*. กรุงเทพฯ: พริกหวานกราฟฟิค.
- ภัทรกร แสงไชย.(2551). *การวิเคราะห์ผลของปฏิสัมพันธ์ระหว่างการสอนแบบแก้ปัญหาโดยอิงทฤษฎีสามศร กับรูปแบบการแก้ปัญหาที่มีต่อความสามารถในการแก้ปัญหาเชิงสร้างสรรค์และผลสัมฤทธิ์ทางการ เรียนคณิตศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 3*.วารสารครุศาสตร์, 4(1), 758-772.
- น้ำผึ้ง อินทะเนตร. (2554). *การศึกษาคุณลักษณะของคะแนนแบบทดสอบปลายเปิดวิชาคณิตศาสตร์เมื่อจำนวนผู้ตรวจและ รูปแบบการตรวจให้คะแนนต่างกัน โดยใช้โมเดลการสรุปอ้างอิงและโมเดลหลายองค์ประกอบของราล์ซ*. (วิทยานิพนธ์ปริญญาโทศึกษบัณฑิต). สาขาวิชาการทดสอบและวัดผลการศึกษา มหาวิทยาลัยศรีนครินทรวิโรฒ.
- ศิริชัย กาญจนวาสี. (2555). *ทฤษฎีการทดสอบแนวใหม่*. (พิมพ์ครั้งที่ 4) กรุงเทพฯ: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- สุพัฒน์ สุกมลสันต์. (2540). *การประยุกต์ใช้ทฤษฎีการอ้างอิงสรุปเพื่อให้คะแนนการสอบเรียงความและการ พัฒนาโปรแกรมคอมพิวเตอร์ใช้งาน*.สถาบันภาษา จุฬาลงกรณ์มหาวิทยาลัย.
- อรรธรณ ดันสุวรรณรัตน์. (2552). *ผลของการจัดกิจกรรมการเรียนรู้คณิตศาสตร์โดยใช้กระบวนการแก้ปัญหาเชิงสร้างสรรค์ที่ มีต่อความสามารถในการแก้ปัญหาและความคิดสร้างสรรค์ทางคณิตศาสตร์ของนักเรียนมัธยมศึกษาปีที่ 2* (วิทยานิพนธ์ปริญญาโทศึกษบัณฑิต). สาขาวิชาการศึกษาคณิตศาสตร์ บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย.
- อาพันธ์ชนิด เจนจิต (2546). *กิจกรรมการเรียนการสอนเรขาคณิตโดยใช้การแก้ปัญหาอย่างสร้างสรรค์ สำหรับ นักเรียนระดับประถมศึกษาตอนปลายที่มีความสามารถพิเศษทางคณิตศาสตร์*(วิทยานิพนธ์ปริญญาโท ศึกษบัณฑิต). สาขาวิชาคณิตศาสตร์ศึกษา บัณฑิตวิทยาลัยมหาวิทยาลัยศรีนครินทรวิโรฒ.

อุษณีย์ บัวศิริพันธุ์ (2542). การเปรียบเทียบค่าสัมประสิทธิ์การสรุปอ้างอิงของแบบทดสอบวิชาคณิตศาสตร์ที่มี การตรวจ จำนวนผู้ตรวจ และประสบการณ์ของผู้ตรวจ แตกต่างกัน. วิทยานิพนธ์ปริญญา มหาบัณฑิต สาขาการวัดผลการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ.

ภาษาอังกฤษ

- Arbesman, M.; &Puccio, G. (2001). *Enhanced quality thorough creative problem solving*. Journal of Nursing Administration, 31: 176-178.
- Balka, D, S.(1974). *The development of an instrument to measure creative ability in mathematics*, Unpublished doctoral dissertation, Univ. of Missouri, Columbia.
- Brennan, R.L.; Gao, X.; & Colton, D.A. (1995). *Generalizability Analysis Work Key Listening and Writing*. Educational and Psychological Measurement. 55(2): 157-176.
- Chui, W.T.;& Wolfe, E.W. (2002). *A Method for Analyzing Sparse Data Matrices in The Generalizability Theory Framework*. Applied Psychological Measurement. 26(3): 321-338.
- Coffman, W. E., (1971). *Essay Examinations*. In R.C. Thorndike (Ed.), Educational Measurement (2nd ed.). Washington, DC: American Council on Education.
- Hopkins, C.D., &Antes, R. (1990). *Classroom Measurement and Evaluation*. Itasca, IL (3rd ed.).
- Hopkins, K.D., &Stanley, J.C. (1981). *Educational and psychological measurement and evaluation* (6th ed.). Englewood Cliffs, NJ: Prentice Hall.
- Kim, H., Cho, S., &Ahn, D. (2003). *Development of mathematical creative problem solving ability test for identification of the gifted in math*. Gifted Educational International, 8, 164-175.
- Lee, K.S.; Hwang D.; & Seo J.J. (2003). *A Development of the Test for Mathematical Creative Problem Solving Ability*. Journal of the Korea Society of Mathematical Education Series D: Research in Mathematical Education. 7(3), 163-189.
- Mehrens, W.A.; & Lehmann, I.J. (1973). *Measurement and Evaluation in Education and Psychology*. New York; Holt, Rinehart and Winston.
- Shavelson, R. J., & Webb, N. M. (1991). *Generalizability theory: A primer* (Vol.1) Sage Publications.
- Sudwee, R.R.; Reeve, S; & Bradshaw, W.S. (2005). *A Comparison of Generalizability Theory and Many-Facet Rasch Measurement in an Analysis of College Sophomore Writing*. Assessment Writing. 9(3): 239-261.
- Swartz, C. M., et al., (1999). Using generalizability Theory to estimate the reliability of writing scores derived from holistic and analytical scoring methods. *Educational and Psychological Measurement*, 59(3), 492-506

- Treffinger, D.J., Isaksen S. G., & Dorval K.B. (2000). *Creative problem solving (CPS Version 6.1TM) A contemporary framework for managing change* [online]. Available from: <http://www.creativelearning.com/> [2015, August 18]
- Treffinger, D.J., Isaksen S. G., & Dorval K.B. (2005). *Creative problem solving: The History, Development, and Implications for Gifted Education and Talent Development* [online]. Available from: <http://gcq.sagepub.com/> [2015, August 18]