

สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ

การวัดแบบอิงเกณฑ์มีจุดมุ่งหมายที่ต้องการวัดพัฒนาการทางด้านความสามารถของนักเรียนในขอบเขตของเนื้อหาและวัตถุประสงค์ที่กำหนดไว้ และจะต้องมีการกำหนดคะแนนจุดตัดเพื่อใช้เป็นเกณฑ์ตัดสินว่านักเรียนจะผ่านหรือไม่ผ่านในขอบเขตของเนื้อหาและวัตถุประสงค์นั้นๆหรือไม่ แต่ในการกำหนดคะแนนจุดตัดนั้นยังเป็นประเด็นที่นักวัดผลโต้แย้งกันมากกว่าแต่ละวิธีที่ใช้ในการกำหนดคะแนนจุดตัดนั้นต่างมีข้อบกพร่องและมีข้อจำกัดต่างๆมากมาย จึงทำให้มีผู้ที่สนใจศึกษาวิธีการแก้ไขข้อจำกัดของการกำหนดคะแนนจุดตัดเป็นจำนวนมาก และการกำหนดคะแนนจุดตัดโดยใช้การตัดสินของผู้เชี่ยวชาญก็เป็นอีกวิธีหนึ่งที่ได้รับความนิยมสนใจเพราะมีความสะดวกในการปฏิบัติ (Berk, 1986) แต่การกำหนดคะแนนจุดตัดโดยใช้การตัดสินของผู้เชี่ยวชาญก็ยังมีข้อจำกัดในเรื่อง ของความสามารถของผู้เชี่ยวชาญในการกำหนดความสามารถขั้นต่ำของผู้สอบ และความสอดคล้องในการตัดสินคะแนนจุดตัดของผู้เชี่ยวชาญ (Glass, 1978 ; Van Der Linden, 1982) ในขณะที่การกำหนดคะแนนจุดตัดโดยอาศัยข้อมูลจากการสอบ โดยการพิจารณาความสามารถของคนอื่นแล้วนำมากำหนดคะแนนจุดตัดนั้น เป็นวิธีที่ไม่แตกต่างจากการวัดผลแบบอิงกลุ่ม (Glass, 1978) ซึ่งจะเห็นได้ว่าการกำหนดคะแนนจุดตัดทั้งสองวิธีต่างก็มีข้อจำกัด แต่เนื่องจากการกำหนดคะแนนจุดตัดโดยใช้การตัดสินของผู้เชี่ยวชาญเป็นวิธีที่เข้าใจง่ายและสะดวกในการปฏิบัติ จึงทำให้มีผู้สนใจศึกษาเพื่อแก้ไขข้อบกพร่องของวิธีนี้ เช่น เฟรดแมนและโฮ (Friedman and Ho, 1990) ได้ศึกษาความสอดคล้องในการตัดสินคะแนนจุดตัดของผู้เชี่ยวชาญโดยนำผลการตัดสินคะแนนจุดตัดครั้งแรกของผู้เชี่ยวชาญมาให้ผู้เชี่ยวชาญทั้งหมดร่วมกันอภิปรายก่อนทำการตัดสินครั้งต่อไป พบว่าทำให้ค่าความสอดคล้องในการตัดสินเพิ่มขึ้น แพทริเซีย (Patricia, 1991) เสนอว่า การให้สารสนเทศประกอบการตัดสินจะทำให้การกำหนดคะแนนจุดตัดของผู้เชี่ยวชาญมีคุณภาพยิ่งขึ้น

จากการศึกษาเอกสาร และงานวิจัยต่างๆทำให้ผู้วิจัยสนใจที่จะศึกษาความสอดคล้องหรือความเที่ยงในการตัดสินคะแนนจุดตัดของผู้เชี่ยวชาญ เมื่อให้สารสนเทศเกี่ยวกับคุณภาพของข้อสอบในเรื่องค่าความยาก อำนาจจำแนก และประสิทธิภาพตัวลวงที่วิเคราะห์ตามทฤษฎีการทดสอบแบบดั้งเดิมหรือวิธีการกำหนดคะแนนจุดตัดโดยการตัดสินของผู้เชี่ยวชาญร่วมกับข้อมูลจากการสอบของผู้เรียน (Combination Model) ซึ่งจากการประเมินวิธีกำหนดคะแนนจุดตัด 23 วิธีของเบอร์กโดยไมรวมวิธีกำหนดคะแนนจุดตัดที่มีความยุ่งยากในการคำนวณอย่างเช่น วิธีการกำหนดคะแนนจุดตัดตามทฤษฎีการตัดสินใจ (Berk, 1986) พบว่า เป็นวิธีการที่มีความเหมาะสมในการกำหนดคะแนนจุดตัด

โดยในการวิจัยครั้งนี้มีวัตถุประสงค์คือ ศึกษาคะแนนจุดตัดและความสัมพันธ์ระหว่างการตัดสินใจคะแนนจุดตัดจากการตัดสินใจของผู้เชี่ยวชาญเมื่อให้สารสนเทศเกี่ยวกับคุณภาพของข้อสอบในเรื่องค่าความยาก อำนาจจำแนก และประสิทธิภาพตัวลวงที่วิเคราะห์ตามทฤษฎีการทดสอบแบบดั้งเดิม ประกอบการตัดสินใจด้วยวิธีของแองกอฟ วิธีของนีเดลสกี และวิธีของอิมพาราและเพลคกับคะแนนความน่าจะเป็นในการตอบข้อสอบได้ถูกต้องของผู้เรียน ตลอดจนเปรียบเทียบความสอดคล้องหรือความเที่ยงในการตัดสินใจคะแนนจุดตัดของผู้เชี่ยวชาญระหว่างวิธีของแองกอฟ วิธีของนีเดลสกี และวิธีของอิมพาราและเพลค

กลุ่มตัวอย่างที่ใช้ในการวิจัย แบ่งออกเป็น 2 กลุ่ม คือ

2.1 กลุ่มตัวอย่างผู้เชี่ยวชาญ ได้แก่ ครูผู้สอนวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2541 ในโรงเรียนสังกัดสำนักงานการประถมศึกษาจังหวัดเลย จำนวน 15 คน

2.2 กลุ่มตัวอย่างนักเรียนที่ใช้ในการวิจัยเพื่อนำมาหาความสัมพันธ์กับการตัดสินใจคะแนนจุดตัดของผู้เชี่ยวชาญ คือ นักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2541 ในโรงเรียนสังกัดสำนักงานการประถมศึกษาจังหวัดเลย จำนวน 384 คน

เครื่องมือที่ใช้ในการวิจัยครั้งนี้ คือแบบสอบวัดผลสัมฤทธิ์วิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 ประเภทเลือกตอบ จำนวน 4 ตัวเลือก โดยที่ผู้วิจัยสุ่มมาจากโปรแกรมคอมพิวเตอร์พัฒนาข้อสอบทักษะการคิดคำนวณพื้นฐานวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 6 ที่พัฒนาขึ้นโดยวันเพ็ญ วิงวอน (2536) จำนวน 50 ข้อ นำไปทดลองใช้กับนักเรียน จำนวน 100 คน จากนั้นจึงนำผลการวิเคราะห์ข้อสอบมาคัดเลือกข้อสอบที่มีคุณภาพดีที่สุดในแต่ละจุดประสงค์ให้เหลือข้อสอบสำหรับใช้เป็นเครื่องมือในการวิจัยจำนวน 25 ข้อตามการออกแบบของการวิจัย

ผู้วิจัยดำเนินการเก็บรวบรวมข้อมูล โดยการนำแบบสอบที่เป็นเครื่องมือในการวิจัยจำนวน 25 ข้อ พร้อมทั้งสารสนเทศเกี่ยวกับข้อสอบซึ่งผู้วิจัยจัดพิมพ์ให้อยู่ในชุดเดียวกัน และแบบฟอร์มการตัดสินใจคะแนนจุดตัดไปให้กลุ่มตัวอย่างผู้เชี่ยวชาญใช้ดุลยพินิจตัดสินใจว่านักเรียนในระดับคาบเส้นหรือมีผลการเรียนเป็น 0 และ 1 มีโอกาสในการตอบข้อสอบแต่ละข้อถูกคิดเป็นร้อยละเท่าใดสำหรับกลุ่มตัวอย่างที่ตัดสินใจคะแนนจุดตัดด้วยวิธีของแองกอฟ และมอบหมายให้กลุ่มตัวอย่างผู้เชี่ยวชาญใช้ดุลยพินิจตัดสินใจว่านักเรียนในระดับคาบเส้นหรือมีผลการเรียนเป็น 0 และ 1 ตอบข้อสอบแต่ละข้อได้ถูกหรือผิด สำหรับกลุ่มตัวอย่างที่ตัดสินใจด้วยวิธีของอิมพาราและเพลค และมอบหมายให้กลุ่มตัวอย่างผู้เชี่ยวชาญพิจารณาตัวเลือกของข้อสอบแต่ละข้อว่ามีตัวเลือกใดที่กลุ่มนักเรียนในระดับคาบเส้นหรือมีผลการเรียนเป็น 0 และ 1 บอกได้ว่าผิดสำหรับกลุ่มตัวอย่างที่ตัดสินใจด้วยวิธีของนีเดลสกี พร้อมทั้งนำแบบสอบที่เป็นเครื่องมือในการวิจัยจำนวน 25 ข้อ ไปให้นักเรียนที่เป็นกลุ่มตัวอย่างดำเนินการสอบตามวัน เวลา ที่นัดหมายไว้ ผู้วิจัยให้เวลากลุ่มตัวอย่างผู้เชี่ยวชาญในการใช้ดุลยพินิจตัดสินใจ 1 สัปดาห์ แล้วจึงไปปรับแบบสอบและผลการพิจารณาของผู้เชี่ยวชาญคืนด้วยตนเอง หลังจากนั้นผู้วิจัยเว้นช่วงเวลา

ประมาณ 2 สัปดาห์หลังจากการเก็บรวบรวมข้อมูลครั้งที่ 1 จึงดำเนินการเก็บรวบรวมข้อมูลจากกลุ่มตัวอย่างผู้เชี่ยวชาญและนักเรียนครั้งที่ 2 ต่อไป

ผู้วิจัยนำข้อมูลที่ได้มาวิเคราะห์หาคะแนนจุดตัดของแบบสอบทั้งหมด และสัมประสิทธิ์สหสัมพันธ์ระหว่างการกำหนดคะแนนจุดตัดของผู้เชี่ยวชาญที่มีสารสนเทศประกอบการตัดสินใจกับคะแนนความน่าจะเป็นในการตอบข้อสอบได้ถูกต้องของผู้เรียนที่ตัดสินใจด้วยทฤษฎีการตัดสินใจ (Decision Theoretic Approaches) สัมประสิทธิ์สหสัมพันธ์ระหว่างการกำหนดคะแนนจุดตัดของผู้เชี่ยวชาญที่มีสารสนเทศประกอบการตัดสินใจ กับคะแนนความน่าจะเป็นในการตอบข้อสอบได้ถูกต้องของผู้เรียนตลอดจนค่าความสอดคล้องในการตัดสินใจของผู้เชี่ยวชาญโดยให้สารสนเทศประกอบ ระหว่างวิธีของแองกอฟ วิธีของนีเดลสกี และวิธีของอิมพาราและเพลคโดยการประยุกต์ตามทฤษฎีการสรุปอ้างอิง (Generalizability Theory)

สรุปผลการวิจัย

1. คะแนนจุดตัดของแบบสอบทั้งหมดที่ตัดสินใจโดยผู้เชี่ยวชาญที่มีสารสนเทศเกี่ยวกับคุณภาพข้อสอบในเรื่อง ค่าความยาก อำนาจจำแนก และประสิทธิภาพตัวรวม ซึ่งวิเคราะห์ตามทฤษฎีการทดสอบแบบดั้งเดิมประกอบการตัดสินใจด้วยวิธีของแองกอฟครั้งที่ 1 และครั้งที่ 2 เท่ากับ 11.484 และ 12.126 ตามลำดับ ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1.802 และ 2.082 ตามลำดับ คะแนนจุดตัดที่ตัดสินใจด้วยวิธีของนีเดลสกีครั้งที่ 1 และครั้งที่ 2 เท่ากับ 10.326 และ 10.517 ตามลำดับ ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1.230 และ 1.802 ตามลำดับ สำหรับคะแนนจุดตัดที่ตัดสินใจด้วยวิธีของอิมพาราและเพลคครั้งที่ 1 และครั้งที่ 2 เท่ากับ 13.60 และ 15.00 ตามลำดับ ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 2.967 และ 1.581 ตามลำดับ ผลต่างระหว่างการตัดสินใจคะแนนจุดตัดด้วยวิธีของแองกอฟ วิธีของนีเดลสกี และวิธีของอิมพาราและเพลคครั้งที่ 1 กับครั้งที่ 2 เท่ากับ 0.642 0.191 และ 1.40 ตามลำดับและค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างการตัดสินใจคะแนนจุดตัดด้วยวิธีของแองกอฟ วิธีของนีเดลสกี และวิธีของอิมพาราและเพลคครั้งที่ 1 กับครั้งที่ 2 เท่ากับ 0.995, 0.9864 และ 0.9061 ตามลำดับ

2. สัมประสิทธิ์สหสัมพันธ์ระหว่างการตัดสินใจคะแนนจุดตัดของผู้เชี่ยวชาญที่มีสารสนเทศเกี่ยวกับคุณภาพข้อสอบในเรื่อง ค่าความยาก อำนาจจำแนก และประสิทธิภาพตัวรวม ซึ่งวิเคราะห์ตามทฤษฎีการทดสอบแบบดั้งเดิมประกอบการตัดสินใจกับคะแนนความน่าจะเป็นในการตอบข้อสอบได้ถูกต้องของผู้เรียน ครั้งที่ 1 และครั้งที่ 2 เท่ากับ 0.355 และ 0.544 ตามลำดับ วิธีของนีเดลสกีครั้งที่ 1 และครั้งที่ 2 เท่ากับ 0.358 และ 0.247 และวิธีของอิมพาราและเพลคครั้งที่ 1 และครั้งที่ 2 เท่ากับ 0.421 และ 0.477 โดยที่ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างวิธีของแองกอฟและวิธีของอิมพาราและเพลคครั้งที่ 2 กับคะแนนความน่าจะเป็นในการตอบข้อสอบได้ถูกต้องของผู้เรียนมีความสัมพันธ์กันอย่างมีนัย

สำคัญทางสถิติที่ระดับ .05 สำหรับวิธีของนิเดลส์ทั้งครั้งที่ 1 และครั้งที่ 2 และวิธีของแองกอฟ และวิธีของอิมพาราและเพลคครั้งที่ 1 มีความสัมพันธ์กันอย่างไม่มีนัยสำคัญทางสถิติที่ระดับ .05

3. สัมประสิทธิ์สหสัมพันธ์ระหว่างการตัดสินใจคะแนนจุดตัดของผู้เชี่ยวชาญที่มีสารสนเทศเกี่ยวกับคุณภาพข้อสอบในเรื่องค่าความยาก อำนาจจำแนก และประสิทธิภาพตัวลง ซึ่งวิเคราะห์ตามทฤษฎีการทดสอบแบบดั้งเดิมประกอบการตัดสินใจ กับคะแนนความน่าจะเป็นในการตอบข้อสอบได้ถูกต้องของผู้เรียนในกลุ่มที่รอบรู้และไม่รอบรู้จริงตามทฤษฎีการตัดสินใจครั้งที่ 1 และครั้งที่ 2 เท่ากับ 0.555 และ 0.506 ตามลำดับ วิธีของนิเดลส์ ครั้งที่ 1 และครั้งที่ 2 เท่ากับ 0.248 และ 0.289 และวิธีของอิมพาราและเพลคครั้งที่ 1 และครั้งที่ 2 เท่ากับ 0.632 และ 0.460 โดยที่ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างการตัดสินใจคะแนนจุดตัดของผู้เชี่ยวชาญด้วยวิธีของแองกอฟครั้งที่ 1 และครั้งที่ 2 และวิธีของอิมพาราและเพลคครั้งที่ 2 กับคะแนนความน่าจะเป็นในการตอบข้อสอบได้ถูกต้องของผู้เรียนในกลุ่มที่รอบรู้และไม่รอบรู้จริงตามทฤษฎีการตัดสินใจมีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 สำหรับวิธีของนิเดลส์ทั้งครั้งที่ 1 และครั้งที่ 2 และวิธีของอิมพาราและเพลคครั้งที่ 2 มีความสัมพันธ์กันอย่างไม่มีนัยสำคัญทางสถิติที่ระดับ .05

4. ค่าความสอดคล้องหรือค่าความเที่ยงในการตัดสินใจด้วยวิธีของแองกอฟเชิงสัมพัทธ์ครั้งที่ 1 และครั้งที่ 2 เท่ากับ 0.7418 และ 0.6617 ตามลำดับ และค่าความแปรปรวนของความคลาดเคลื่อนเชิงสัมพัทธ์ เท่ากับ .0009 และ .0010 หรือคิดเป็นร้อยละ 20.455 และ 31.818 ของความแปรปรวนทั้งหมด สำหรับค่าความสอดคล้องหรือค่าความเที่ยงในการตัดสินใจด้วยวิธีของแองกอฟเชิงสัมบูรณ์ครั้งที่ 1 และครั้งที่ 2 เท่ากับ 0.5688 และ 0.4593 ตามลำดับและค่าความแปรปรวนของความคลาดเคลื่อนเชิงสัมบูรณ์ เท่ากับ .0019 และ .0024 หรือคิดเป็นร้อยละ 43.182 และ 53.545 ของความแปรปรวนทั้งหมด จากการทดสอบพบว่าค่าความสอดคล้องหรือความเที่ยงในการตัดสินใจทั้งสองครั้งไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

5. ค่าความสอดคล้องหรือค่าความเที่ยงในการตัดสินใจด้วยวิธีของนิเดลส์เชิงสัมพัทธ์ครั้งที่ 1 และครั้งที่ 2 เท่ากับ 0.3275 และ 0.3281 ตามลำดับ และค่าความแปรปรวนของความคลาดเคลื่อนเชิงสัมพัทธ์ เท่ากับ .0015 และ .0015 หรือคิดเป็นร้อยละ 57.692 และ 53.571 ของความแปรปรวนทั้งหมด สำหรับค่าความสอดคล้องในการตัดสินใจด้วยวิธีของนิเดลส์เชิงสัมบูรณ์ครั้งที่ 1 และครั้งที่ 2 เท่ากับ 0.2759 และ 0.2700 ตามลำดับ และค่าความแปรปรวนของความคลาดเคลื่อนเชิงสัมบูรณ์ เท่ากับ .0019 และ .0020 หรือคิดเป็นร้อยละ 73.077 และ 71.428 ของความแปรปรวนทั้งหมด จากการทดสอบพบว่าค่าความสอดคล้องในการตัดสินใจทั้งสองครั้งไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

6. ค่าความสอดคล้องในการตัดสินใจด้วยวิธีของอิมพาราและเพลคเชิงสัมพัทธ์ ครั้งที่ 1 และครั้งที่ 2 เท่ากับ 0.337 และ 0.384 ตามลำดับ และค่าความแปรปรวนของความคลาดเคลื่อนเชิง

สัมพัทธ์ เท่ากับ .0402 และ .040 หรือคิดเป็นร้อยละ 64.839 และ 61.633 ของความแปรปรวนทั้งหมด สำหรับค่าความสอดคล้องในการตัดสินใจด้วยวิธีของแองกอฟเชิงสัมบูรณ์ครั้งที่ 1 และครั้งที่ 2 เท่ากับ 0.329 และ 0.384 ตามลำดับ และค่าความแปรปรวนของความคลาดเคลื่อนเชิงสัมบูรณ์เท่ากับ .0416 และ .040 หรือคิดเป็นร้อยละ 67.097 และ 61.633 ของความแปรปรวนทั้งหมด จากการทดสอบพบว่า ค่าความสอดคล้องในการตัดสินใจทั้งสองครั้งไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

7. ค่าความสอดคล้องในการตัดสินใจด้วยวิธีของแองกอฟ วิธีของนีเดลสกี และวิธีของอิมพารา และเพลคเชิงสัมพัทธ์ครั้งที่ 1 เท่ากับ 0.7418, 0.3257 และ 0.337 ตามลำดับ และค่าความสอดคล้องในการตัดสินใจเชิงสัมบูรณ์ครั้งที่ 1 เท่ากับ 0.5688, 0.2579 และ 0.329 ตามลำดับ จากการทดสอบพบว่าค่าความสอดคล้องในการตัดสินใจทั้งสามวิธีไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

8. ค่าความสอดคล้องในการตัดสินใจด้วยวิธีของแองกอฟ วิธีของนีเดลสกี และวิธีของอิมพารา และเพลคเชิงสัมพัทธ์ครั้งที่ 2 เท่ากับ 0.6617, 0.3281 และ 0.384 ตามลำดับ และค่าความสอดคล้องในการตัดสินใจเชิงสัมบูรณ์ครั้งที่ 2 เท่ากับ 0.4593, 0.270 และ 0.384 ตามลำดับ จากการทดสอบพบว่าค่าความสอดคล้องในการตัดสินใจทั้งสามวิธีไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

การอภิปรายผลการวิจัย

1. ผลการตัดสินใจคะแนนจุดตัดของแบบสอบทั้งหมดโดยผู้เชี่ยวชาญที่มีสารสนเทศเกี่ยวกับคุณภาพข้อสอบในเรื่องค่าความยาก อำนาจจำแนก และประสิทธิภาพตัวลวง ซึ่งวิเคราะห์ตามทฤษฎีการทดสอบแบบดั้งเดิมประกอบการตัดสินใจด้วยวิธีของแองกอฟ วิธีของนีเดลสกี และวิธีของอิมพาราและเพลคครั้งที่ 1 เท่ากับ 11.484, 10.326 และ 13.60 และครั้งที่ 2 เท่ากับ 12.126, 10.517 และ 15.00 ตามลำดับ และผลต่างของการตัดสินใจคะแนนจุดตัดด้วยวิธีของแองกอฟ วิธีของนีเดลสกี และวิธีของอิมพาราและเพลคระหว่างครั้งที่ 1 กับครั้งที่ 2 เท่ากับ 0.642, 0.191 และ 1.40 ตามลำดับ จะเห็นว่าคะแนนจุดตัดของแบบสอบทั้งหมดและผลต่างระหว่างการตัดสินใจคะแนนจุดตัดครั้งที่ 1 กับครั้งที่ 2 จากการตัดสินใจด้วยวิธีของอิมพาราและเพลคจะสูงกว่าวิธีของแองกอฟ วิธีของนีเดลสกี และคะแนนจุดตัดของแบบสอบทั้งหมดที่ได้จากการตัดสินใจด้วยวิธีของนีเดลสกีจะมีค่าต่ำสุด ซึ่งสอดคล้องกับการศึกษาของ ชาง (Chang, 1996) ที่พบว่าคะแนนจุดตัดที่ได้จากการตัดสินใจด้วยวิธีของนีเดลสกีจะมีค่าต่ำกว่าวิธีของแองกอฟ แต่ในการวิจัยครั้งนี้ไม่ได้ทดสอบความมีนัยสำคัญทางสถิติ จึงไม่มีหลักฐานเพียงพอที่จะบอกว่าคะแนนจุดตัดที่ได้จากการตัดสินใจด้วยวิธีของแองกอฟ วิธีของนีเดลสกี และวิธีของอิมพาราและเพลคแตกต่างกันหรือไม่

2. จากการพิจารณาค่าสัมประสิทธิ์สหสัมพันธ์ ระหว่างการตัดสินใจคะแนนจุดตัดด้วยวิธีของแองกอฟ วิธีของนีเดลสกี และวิธีของอิมพาราและเพลค ครั้งที่ 1 กับครั้งที่ 2 ซึ่งมีค่าเท่ากับ

0.995, 0.9864 และ 0.9061 ตามลำดับ ซึ่งแสดงให้เห็นว่าการตัดสินคะแนนจุดตัดของผู้เชี่ยวชาญที่มีสารสนเทศเกี่ยวกับคุณภาพข้อสอบในเรื่องค่าความยาก อำนาจจำแนก และประสิทธิภาพตัวลง ซึ่งวิเคราะห์ตามทฤษฎีการทดสอบแบบดั้งเดิมประกอบการตัดสินมีความเที่ยงแบบความคงที่สูงนั่นคือผู้เชี่ยวชาญมีความคงเส้นคงวาในการตัดสินคะแนนจุดตัดในแต่ละครั้ง

3. จากการศึกษาสัมประสิทธิ์สหสัมพันธ์ระหว่างการการตัดสินคะแนนจุดตัดของผู้เชี่ยวชาญที่มีสารสนเทศเกี่ยวกับคุณภาพข้อสอบในเรื่องค่าความยาก อำนาจจำแนก และประสิทธิภาพตัวลง ซึ่งวิเคราะห์ตามทฤษฎีการทดสอบแบบดั้งเดิมประกอบการตัดสินกับคะแนนความน่าจะเป็นในการตอบข้อสอบได้ถูกต้องของผู้เรียน เพื่อตรวจสอบความตรงตามสภาพในการตัดสินคะแนนจุดตัดของผู้เชี่ยวชาญ โดยใช้คะแนนความน่าจะเป็นในการตอบข้อสอบได้ถูกต้องของผู้เรียนเป็นเกณฑ์ พบว่า การตัดสินคะแนนจุดตัดด้วยวิธีของนิตลสกีทั้งครั้งที่ 1 และ ครั้งที่ 2 วิธีของแองกอฟครั้งที่ 1 และวิธีของอิมพาราและเพลค ครั้งที่ 1 กับคะแนนที่ได้จากการทำแบบสอบของผู้เรียนทั้งหมดมีความสัมพันธ์กันอย่างไม่มีนัยสำคัญทางสถิติที่ระดับ .05 ส่วนการตัดสินคะแนนจุดตัดด้วยวิธีของแองกอฟและวิธีของอิมพาราและเพลคกับคะแนนความน่าจะเป็นในการตอบข้อสอบได้ถูกต้องของผู้เรียนครั้งที่ 2 มีค่าเท่ากับ 0.544 และ 0.477 ตามลำดับ และมีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 แสดงว่าการตัดสินคะแนนจุดตัดของผู้เชี่ยวชาญด้วยวิธีของแองกอฟและวิธีของอิมพาราและเพลค มีความถูกต้องตรงกับคะแนนความน่าจะเป็นในการตอบข้อสอบได้ถูกต้องของผู้เรียนสำหรับการเก็บข้อมูลครั้งที่ 2 แต่ครั้งที่ 1 ไม่สามารถบอกได้ว่าสัมพันธ์กันจริงหรือไม่ แสดงว่ายังไม่มีหลักฐานเพียงพอที่จะสรุปว่าผู้เชี่ยวชาญสามารถตัดสินข้อสอบรายข้อได้ถูกต้องกับคะแนนความน่าจะเป็นในการตอบข้อสอบได้ถูกต้องของผู้เรียนซึ่งตรงกับการวิจัยของอิมพาราและเพลค (Impara and Plake, 1998) ที่พบว่าผู้เชี่ยวชาญตัดสินคะแนนจุดตัดด้วยวิธีของแองกอฟได้ไม่ถูกต้องกับคะแนนความน่าจะเป็นในการตอบข้อสอบได้ถูกต้องของผู้เรียนมากนัก และ Shepard (1994) พบว่า ผู้ตัดสินจะประมาณค่าความน่าจะเป็นในการตอบข้อสอบได้ถูกต้องของผู้เรียนสูงกว่าคะแนนความน่าจะเป็นในการตอบข้อสอบได้ถูกต้องของผู้เรียนที่แท้จริงสำหรับข้อสอบที่ยาก และประมาณค่าความน่าจะเป็นในการตอบข้อสอบได้ถูกต้องของผู้เรียนได้ต่ำกว่าคะแนนความน่าจะเป็นในการตอบข้อสอบได้ถูกต้องของผู้เรียนที่แท้จริงสำหรับข้อสอบที่ง่าย ดังนั้นควรที่จะทำการศึกษาโดยใช้กลุ่มตัวอย่างหลายๆกลุ่ม หรือศึกษาซ้ำหลายๆครั้งเพื่อให้ได้ผลสรุปที่สามารถยืนยันได้ว่ามีความสัมพันธ์กันจริงหรือไม่

3. สัมประสิทธิ์สหสัมพันธ์ระหว่างการกำหนดคะแนนจุดตัดของผู้เชี่ยวชาญที่มีสารสนเทศเกี่ยวกับคุณภาพข้อสอบในเรื่องค่าความยาก อำนาจจำแนก และประสิทธิภาพตัวลงซึ่งวิเคราะห์ตามทฤษฎีการทดสอบแบบดั้งเดิมประกอบการตัดสินด้วยวิธีของแองกอฟกับคะแนนความน่าจะเป็นในการตอบข้อสอบได้ถูกต้องของผู้เรียนในกลุ่มที่รอบรู้และไม่รอบรู้จริงตามทฤษฎีการตัดสินใจครั้งที่ 1 และครั้งที่ 2 เท่ากับ 0.555 และ 0.506 ตามลำดับ และวิธีของอิมพาราและเพลคครั้งที่ 1 เท่ากับ 0.632 โดยที่

ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างการตัดสินใจคะแนนจุดตัดของผู้เชี่ยวชาญด้วยวิธีของแองกอฟครั้งที่ 1 และครั้งที่ 2 และวิธีของอิมพาราและเพลค ครั้งที่ 1 กับคะแนนความน่าจะเป็นในการตอบข้อสอบได้ถูกต้องของผู้เรียนในกลุ่มที่รอบรู้และไม่รอบรู้จริงตามทฤษฎีการตัดสินใจมีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 จะเห็นว่าวิธีของแองกอฟที่สัมพันธ์กันอย่างมีนัยสำคัญทางสถิติทั้งสองครั้ง แสดงว่าการตัดสินใจคะแนนจุดตัดของผู้เชี่ยวชาญด้วยวิธีของแองกอฟมีความถูกต้อง กับคะแนนความน่าจะเป็นในการตอบข้อสอบได้ถูกต้องของผู้เรียนในกลุ่มที่รอบรู้และไม่รอบรู้จริงตามทฤษฎีการตัดสินใจ ซึ่งเป็นกลุ่มผู้เรียนที่มีความน่าจะเป็นของการเกิดความคลาดเคลื่อนต่ำ โดยมีความสัมพันธ์กันในระดับปานกลาง ส่วนวิธีของอิมพาราและเพลคนั้นมีเพียง ครั้งที่ 1 ที่มีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติ ดังนั้นจึงยังไม่มีหลักฐานเพียงพอที่จะสรุปว่าการตัดสินใจคะแนนจุดตัดของผู้เชี่ยวชาญด้วยวิธีของอิมพาราและเพลคมีความถูกต้องกับคะแนนความน่าจะเป็นในการตอบข้อสอบได้ถูกต้องของผู้เรียนในกลุ่มที่รอบรู้และไม่รอบรู้จริงตามทฤษฎีการตัดสินใจ สำหรับวิธีของนิเดลสกีครั้งที่ 1 และครั้งที่ 2 มีความสัมพันธ์กันอย่างไม่มีนัยสำคัญทางสถิติที่ระดับ .05 แสดงว่าการตัดสินใจด้วยวิธีของนิเดลสกีไม่มีความถูกต้องกับคะแนนความน่าจะเป็นในการตอบข้อสอบได้ถูกต้องของผู้เรียนในกลุ่มที่รอบรู้และไม่รอบรู้จริงตามทฤษฎีการตัดสินใจ

4. จากการเปรียบเทียบค่าความสอดคล้องในการตัดสินใจคะแนนจุดตัด ระหว่างวิธีของแองกอฟครั้งที่ 1 กับครั้งที่ 2 วิธีของนิเดลสกี ครั้งที่ 1 กับครั้งที่ 2 และวิธีของอิมพาราและเพลคครั้งที่ 1 กับครั้งที่ 2 ทั้งเชิงสัมพัทธ์และเชิงสัมบูรณ์ ปรากฏว่า ค่าความสอดคล้องในการตัดสินใจทั้งสองครั้งของแต่ละวิธีไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 แสดงว่า ผู้เชี่ยวชาญมีความคงที่ในการตัดสินใจคะแนนจุดตัดและสามารถประมาณค่าได้ใกล้เคียงกันในการตัดสินใจแต่ละครั้ง ซึ่งทำให้เห็นถึงความสอดคล้องในการตัดสินใจคะแนนจุดตัดของแต่ละคน

5. ผลการเปรียบเทียบค่าความสอดคล้องในการตัดสินใจด้วยวิธีของแองกอฟ วิธีของนิเดลสกี และวิธีของอิมพาราและเพลคทั้งเชิงสัมพัทธ์และเชิงสัมบูรณ์ ครั้งที่ 1 และผลการเปรียบเทียบค่าความสอดคล้องในการตัดสินใจด้วยวิธีของแองกอฟ วิธีของนิเดลสกี และวิธีของอิมพาราและเพลคทั้งเชิงสัมพัทธ์และเชิงสัมบูรณ์ครั้งที่ 2 พบว่าค่าความสอดคล้องในการตัดสินใจทั้งสามวิธีไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ทั้งนี้อาจเนื่องมาจากจำนวนผู้เชี่ยวชาญที่ตัดสินใจคะแนนจุดตัดในแต่ละกลุ่มมีจำนวนน้อยเกินไป และผลการวิจัยนี้ไม่สอดคล้องกับสมมติฐานของกาววิจัย จากการศึกษาของครอสและคณะ (Cross, et. al, 1984) ที่หาความสอดคล้องในการตัดสินใจคะแนนจุดตัดด้วยวิธีของแองกอฟ วิธีของนิเดลสกี และวิธีของเจบอร์ และพบว่าค่าความสอดคล้องในการตัดสินใจคะแนนจุดตัดด้วยวิธีของแองกอฟมีค่าสูงที่สุด และวิธีของนิเดลสกีมีค่าความสอดคล้องในการตัดสินใจคะแนนจุดตัดต่ำสุดแต่ไม่ได้ทำการทดสอบความมีนัยสำคัญทางสถิติ เช่นเดียวกับ อิมพาราและเพลค (Impara and Plake, 1997) ศึกษาความสอดคล้องในการตัดสินใจคะแนนจุดตัดระหว่างวิธีของแองกอฟกับ

วิธีของอิมพาราและเพลค พบว่าค่าความสอดคล้องในการตัดสินคะแนนจุดตัดด้วยวิธีของแองกอฟสูงกว่าวิธีของอิมพาราและเพลค แต่ไม่ได้ทำการทดสอบความมีนัยสำคัญทางสถิติ แต่เมื่อพิจารณาถึงค่าความแปรปรวนความคลาดเคลื่อนซึ่งแสดงถึงความไม่สอดคล้องในการตัดสินของผู้เชี่ยวชาญด้วยวิธีของแองกอฟ วิธีของนีเดลสกี และวิธีของอิมพาราและเพลค จากการตัดสินทั้ง 2 ครั้ง พบว่า วิธีของแองกอฟมีค่าความคลาดเคลื่อนต่ำกว่าวิธีของนีเดลสกี และวิธีของอิมพาราและเพลค และวิธีของนีเดลสกีจะมีค่าความคลาดเคลื่อนมากที่สุด

ข้อเสนอแนะ

ข้อเสนอแนะสำหรับการนำผลการวิจัยไปใช้

1. การกำหนดคะแนนจุดตัดถือว่าเป็นหน้าที่ของครูผู้สอนสำหรับกาวิัดแบบอิงเกณฑ์ เพราะครูเป็นผู้ที่ใกล้ชิดและคุ้นเคยกับนักเรียนตลอดจนเข้าใจและทราบข้อบกพร่องต่างๆของนักเรียน ดังนั้นการกำหนดคะแนนจุดตัดอย่างมีหลักการและเลือกใช้วิธีที่เหมาะสมอีกทั้งยังง่ายต่อการปฏิบัติก็จะเป็นประโยชน์ต่อทั้งตัวครู และตัวนักเรียน การวิจัยครั้งนี้พบว่าค่าความสอดคล้องในการตัดสินคะแนนจุดตัดระหว่างวิธีของแองกอฟ วิธีของนีเดลสกี และวิธีของอิมพาราและเพลคไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติแสดงว่าการตัดสินคะแนนจุดตัดด้วยวิธีของแองกอฟ วิธีของนีเดลสกี และวิธีของอิมพาราและเพลคมีความสอดคล้องในการตัดสินของผู้เชี่ยวชาญไม่ต่างกัน แต่ถ้าพิจารณาค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างการตัดสินคะแนนจุดตัดของผู้เชี่ยวชาญ กับคะแนนความน่าจะเป็นในการตอบข้อสอบได้ถูกต้องของผู้เรียนซึ่งแสดงถึงความตรงตามสภาพในการตัดสินของผู้เชี่ยวชาญโดยใช้คะแนนความน่าจะเป็นในการตอบข้อสอบได้ถูกต้องของผู้เรียนเป็นเกณฑ์จะเห็นว่า การตัดสินคะแนนจุดตัดของผู้เชี่ยวชาญด้วยวิธีของแองกอฟจะให้ค่าประมาณที่ตรงกับคะแนนการทำแบบสอบแต่ละข้อของผู้เรียนมากกว่าวิธีอื่น แต่จากการสอบถามผู้เชี่ยวชาญพบว่าวิธีของอิมพาราและเพลคเป็นวิธีที่เข้าใจง่ายและมีความสะดวกในการปฏิบัติ

2. การกำหนดคะแนนจุดตัดด้วยวิธีตัดสินโดยผู้เชี่ยวชาญ ควรมีการนำสารสนเทศเกี่ยวกับผลสอบของนักเรียนในปีก่อนมาพิจารณาประกอบด้วย และมีการตรวจสอบความถูกต้องของคะแนนจุดตัดก่อนการนำไปใช้หรือให้ผู้ที่เกี่ยวข้อง เช่น ผู้บริหาร ครูและนักเรียนมีส่วนร่วมในการกำหนดคะแนนจุดตัดด้วย

ข้อเสนอแนะในการวิจัย

1. การกำหนดคะแนนจุดตัดด้วยการตัดสินใจของผู้เชี่ยวชาญครั้งนี้ ผู้วิจัยใช้สารสนเทศเกี่ยวกับคุณภาพข้อสอบในเรื่อง ค่าความยาก อำนาจจำแนก และประสิทธิภาพตัวลงที่ได้จากการวิเคราะห์ตามทฤษฎีการทดสอบแบบดั้งเดิมประกอบการตัดสินใจ พบว่า ค่าความสอดคล้องหรือความเที่ยงในการตัดสินใจได้อยู่ในระดับปานกลาง ดังนั้นจึงควรจะศึกษาความสอดคล้องในการตัดสินใจคะแนนจุดตัดโดยใช้กระบวนการกลุ่ม (Cizek and Fitzgerald,1996) และเทคนิคเดลฟาย (Delphi Technique) ซึ่งมีการนำค่าเฉลี่ยของการตัดสินใจของกลุ่มผู้เชี่ยวชาญมาประกอบในการตัดสินใจครั้งต่อไป (Cope, 1987)



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย