

สรุปผล อภิปราย และข้อเสนอแนะ

ความมุ่งหมายของการวิจัย

1. เพื่อศึกษาความเที่ยงตรงในการจำแนกความรอบรู้ จากการใช้แบบทดสอบอิงเกณฑ์ แบบทดสอบเทเลอร์แบบแตกแขนง (Branched Test) แบบทดสอบเทเลอร์แบบทดสอบตามสภาพ (Status Tailoring) แบบทดสอบเทเลอร์แบบทดลองทำ (Performance Tailoring) และแบบทดสอบเทเลอร์แบบให้ตัดสินใจเลือกเอง (Self-Tailoring) เมื่อใช้เกณฑ์ตัดสินโดยผู้เชี่ยวชาญ
2. เพื่อเปรียบเทียบความถูกต้อง ของการจำแนกความรอบรู้จากการทดสอบอิงเกณฑ์ และการทดสอบเทเลอร์ทั้ง 4 รูปแบบ
3. เพื่อเปรียบเทียบค่าความสามารถ ที่ได้จากการตอบจากการทดสอบแบบอิงเกณฑ์ และการทดสอบเทเลอร์ ทั้ง 4 รูปแบบ
4. เพื่อเปรียบเทียบ ค่าสาระประโยชน์ของข้อสอบ (Item Information Function) ของการทดสอบแบบอิงเกณฑ์และการทดสอบเทเลอร์ทั้ง 4 รูปแบบ
5. เพื่อเปรียบเทียบค่าความคลาดเคลื่อนมาตรฐาน ของค่าความสามารถจากผลการสอบการทดสอบอิงเกณฑ์และการทดสอบเทเลอร์ทั้ง 4 รูปแบบ

กลุ่มตัวอย่าง

กลุ่มตัวอย่างที่ใช้ในการสร้างเครื่องมือ คือ นักเรียนชั้นประถมศึกษาปีที่ 6 ที่ศึกษาใน

- 1) โรงเรียนสังกัดสำนักงานการประถมศึกษาจังหวัดเชียงราย ปีการศึกษา 2529 จำนวน 3 อำเภอ คือ อำเภอเมือง 8 กลุ่มโรงเรียนจำนวน 31 โรงเรียน
 - 2) อำเภอแม่สรวย 5 กลุ่มโรงเรียน จำนวน 19 โรงเรียน
 - 3) อำเภอเวียงป่าเป้า 6 กลุ่มโรงเรียนจำนวน 15 โรงเรียน
- รวมจำนวนนักเรียนที่ใช้ในการสร้างเครื่องมือทั้งสิ้น 2,400 คน

กลุ่มตัวอย่างที่ใช้ในการวิจัย เป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2529 คือกลุ่มโรงเรียนแม่สรวย สำนักงานการประถมศึกษาอำเภอแม่สรวย ในสังกัดสำนักงานการประถมศึกษาจังหวัดเชียงราย ได้แก่ โรงเรียนบ้านแม่สรวย (ศึกษาศมิตร์) โรงเรียนบ้านห้วยม่วง โรงเรียนบ้านปางท่าแม่จันทน์ และโรงเรียนบ้านโป่งปูเฟือง รวมทั้งสิ้น 115 คน

เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูลเพื่อการวิจัย ประกอบด้วยแบบทดสอบ 2 ประเภท คือ

1. แบบทดสอบอิงเกณฑ์

แบบทดสอบอิงเกณฑ์เป็นแบบทดสอบที่ผู้วิจัยสร้างขึ้นประกอบด้วย

- 1.1 แบบทดสอบอิงเกณฑ์ 10 ข้อ
- 1.2 แบบทดสอบอิงเกณฑ์ 20 ข้อ
- 1.3 แบบทดสอบอิงเกณฑ์ 30 ข้อ

2. แบบทดสอบเทเลอร์

แบบทดสอบเทเลอร์เป็นแบบทดสอบที่ผู้วิจัยสร้างขึ้นประกอบด้วย

- 2.1 แบบทดสอบเทเลอร์แบบแตกแขนง (Branched Test)
- 2.2 แบบทดสอบเทเลอร์แบบทดสอบตามสภาพ (Status Tailoring)
- 2.3 แบบทดสอบเทเลอร์แบบทดลองทำ (Performance Tailoring)
- 2.4 แบบทดสอบเทเลอร์แบบให้ตัดสินใจเลือกเอง (Self-Tailoring)

วิธีการดำเนินการเก็บรวบรวมข้อมูล

ขั้นตอนในการดำเนินการเก็บรวบรวมข้อมูลในการศึกษามีดังนี้

1. นำข้อสอบในกลุ่มข้อสอบไปทดสอบนักเรียนกลุ่มตัวอย่าง โดยวิธีเบส์ (Bayesian

Strategies) โดยทดสอบเป็นรายบุคคล เพื่อจำแนกนักเรียนแต่ละคนว่าเป็นผู้รอบรู้หรือไม่รอบรู้ที่แท้จริง

2. นำแบบทดสอบอิงเกณฑ์ทั้ง 3 ฉบับและแบบทดสอบเทเลอร์ทั้ง 4 แบบ ไปทดสอบนักเรียนทั้งกลุ่มที่รอบรู้และไม่รอบรู้ในข้อ 1.

3. นำผลการตอบในแต่ละแบบทดสอบประมาณค่าความสามารถ ค่าความคลาดเคลื่อนของค่าความสามารถ และค่าสาระประโยชน์ของข้อสอบในแต่ละรายแบบทดสอบ เพื่อใช้ในการวิเคราะห์ทางสถิติและทดสอบสมมติฐานที่ตั้งไว้ สำหรับการทดสอบเทเลอร์แบบให้ตัดสินใจเลือกเองมีการทดสอบความเหมาะสมของผู้สอบ และการควบคุมคุณภาพ (Person Fit and Quality Control) ว่าผลการตอบเป็นไปตามที่คาดหวังหรือไม่ ถ้ายอมรับผลการสอบ จึงจะนำผลการตอบไปใช้ในการวิเคราะห์ทางสถิติ

การวิเคราะห์ข้อมูล

ผู้วิจัยเสนอผลการวิเคราะห์ตามความมุ่งหมายและสมมติฐานของการวิจัยดังนี้

1. ค่าความเที่ยงตรงในการจำแนกการรอบรู้จากการใช้แบบทดสอบอิงเกณฑ์ และแบบทดสอบเทเลอร์
2. เปรียบเทียบความถูกต้องในการจำแนกการรอบรู้ของแบบทดสอบอิงเกณฑ์ และแบบทดสอบเทเลอร์
3. เปรียบเทียบค่าความสามารถจากแบบทดสอบอิงเกณฑ์และแบบทดสอบเทเลอร์
4. เปรียบเทียบค่าสาระประโยชน์ของข้อสอบจากแบบทดสอบอิงเกณฑ์ และแบบทดสอบเทเลอร์
5. เปรียบเทียบค่าความคลาดเคลื่อนมาตรฐานของค่าความสามารถ จากแบบทดสอบอิงเกณฑ์และการทดสอบเทเลอร์

ผลการวิเคราะห์ข้อมูล

จากการวิเคราะห์ข้อมูลผู้วิจัยสรุปผลการวิเคราะห์ข้อมูลดังนี้

1. ค่าความเที่ยงตรงในการจำแนกความรอบรู้จากการใช้แบบทดสอบอิงเกณฑ์ และแบบทดสอบเทเลอร์ พบว่า แบบทดสอบอิงเกณฑ์ 10, 20, และ 30 ข้อ มีความเที่ยงตรงในการจำแนกความรอบรู้ .55, .61, และ .65 ตามลำดับ แบบทดสอบเทเลอร์แบบแตกแขนง แบบทดสอบเทเลอร์แบบทดสอบตามสภาพ แบบทดสอบเทเลอร์แบบทดลองทำ และแบบทดสอบเทเลอร์แบบให้ตัดสินใจเลือกเอง มีความเที่ยงตรงในการจำแนก .71, .80, .74 และ .70 ตามลำดับ โดยทุกค่าความเที่ยงตรงในการจำแนกการรอบรู้มีนัยสำคัญทางสถิติที่ระดับ .01

เมื่อนำค่าสัมประสิทธิ์สหสัมพันธ์ฟาย (φ) ไปเทียบกับค่าสัมประสิทธิ์สหสัมพันธ์ฟายสูงสุด (ϕ_{max}) เพื่อให้ค่าฟายสูงสุด (ϕ_{max}) เป็น 1 พบว่าแบบทดสอบเทเลอร์แบบแตกแขนง (Branched Test) แบบทดสอบเทเลอร์แบบทดสอบตามสภาพ (Status Tailoring) แบบทดสอบเทเลอร์แบบทดลองทำ (Performance Tailoring) และแบบทดสอบเทเลอร์แบบให้ตัดสินใจเลือกเอง (Self-Tailoring) ให้ค่าที่มีแนวโน้มความเที่ยงตรงในการจำแนกการรอบรู้สูงคือ 1.00, .89, .76 และ .71 ตามลำดับ ในขณะที่แบบทดสอบอิงเกณฑ์ 10, 20 และ 30 ข้อ ให้ค่าความเที่ยงตรงในการจำแนกในระดับปานกลางคือ .59, .69 และ .67 ตามลำดับ

2. เปรียบเทียบความถูกต้องในการจำแนกความรอบรู้ จากแบบทดสอบอิงเกณฑ์และแบบทดสอบเทเลอร์ พบว่า แบบทดสอบอิงเกณฑ์ 10, 20 และ 30 ข้อ แบบทดสอบเทเลอร์แบบแตกแขนง แบบทดสอบเทเลอร์แบบทดสอบตามสภาพ แบบทดสอบเทเลอร์แบบทดลองทำ และแบบทดสอบเทเลอร์แบบให้ตัดสินใจเลือกเอง มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 เมื่อทดสอบค่าเฉลี่ยเป็นรายคู่พบว่า แบบทดสอบอิงเกณฑ์ 10 ข้อ ให้ความถูกต้องในการจำแนกความรอบรู้ต่ำกว่าแบบทดสอบเทเลอร์แบบทดสอบตามสภาพ อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ส่วนระหว่างคู่อื่น ๆ แตกต่างกันอย่างไม่มีนัยสำคัญ

3. เปรียบเทียบค่าความสามารถจากแบบทดสอบอิงเกณฑ์ และแบบทดสอบเทเลอร์ พบว่า ค่าความที่ประมาณได้จากแบบทดสอบอิงเกณฑ์ที่ใช้แบบทดสอบ 10, 20 และ 30 ข้อ ทดสอบเทเลอร์แบบแตกแขนง แบบทดสอบเทเลอร์แบบทดสอบตามสภาพ แบบทดสอบเทเลอร์แบบทดลองทำ และแบบทดสอบเทเลอร์แบบให้ตัดสินใจเลือกเอง มีความแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

4. เปรียบเทียบค่าสาระประโยชน์ของข้อสอบ จากการทดสอบอิงเกณฑ์และการทดสอบเทเลอร์พบว่า ค่าสาระประโยชน์ของข้อสอบจากแบบทดสอบอิงเกณฑ์ 10, 20 และ 30 ข้อ แบบทดสอบเทเลอร์แบบแตกแขนง แบบทดสอบเทเลอร์แบบทดลองทำ และแบบทดสอบเทเลอร์แบบให้ตัดสินใจเลือกเอง มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 และเมื่อทดสอบค่าสาระประโยชน์ ของข้อสอบเฉลี่ยเป็นรายคู่หลังการวิเคราะห์ความแปรปรวน พบว่า ค่าสาระประโยชน์เฉลี่ยของข้อสอบของแบบทดสอบอิงเกณฑ์ 10, 20 และ 30 ข้อ แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ และต่ำกว่าแบบทดสอบเทเลอร์แบบแตกแขนง แบบทดสอบเทเลอร์แบบทดลองทำ และแบบทดสอบเทเลอร์แบบให้ตัดสินใจเลือกทำเองอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 แบบทดสอบเทเลอร์แบบทดลองทำมีค่าสาระประโยชน์เฉลี่ยของข้อสอบต่ำกว่าแบบทดสอบเทเลอร์แบบแตกแขนง และแบบทดสอบตามสภาพอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 และ .05 ตามลำดับ แต่แตกต่างกับแบบทดสอบเทเลอร์แบบให้ตัดสินใจเลือกทำเองอย่างไม่มีนัยสำคัญ แบบทดสอบเทเลอร์แบบแตกแขนง แบบทดสอบตามสภาพและแบบให้ตัดสินใจเลือกเองแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

5. เปรียบเทียบค่าความคลาดเคลื่อนมาตรฐานของค่าความสามารถ ของแบบทดสอบอิงเกณฑ์และแบบทดสอบเทเลอร์ พบว่า ค่าความคลาดเคลื่อนมาตรฐานของค่าความสามารถจากแบบทดสอบอิงเกณฑ์ 10, 20 และ 30 ข้อ แบบทดสอบเทเลอร์แบบแตกแขนง แบบทดสอบเทเลอร์แบบทดสอบตามสภาพ แบบทดสอบเทเลอร์แบบทดลองทำ และแบบทดสอบแบบให้ตัดสินใจเลือกเอง มีความแตกต่างกันอย่างมีนัยสำคัญที่ระดับ .01 เมื่อทดสอบค่าความคลาดเคลื่อนมาตรฐานเฉลี่ยของค่าความสามารถเป็นรายคู่หลังการวิเคราะห์ความแปรปรวน พบว่า

เกือบทุกคู่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ยกเว้น 2 คู่คือ ค่าความคลาดเคลื่อนมาตรฐานเฉลี่ยของแบบทดสอบอิงเกณฑ์ 20 ข้อ กับแบบทดสอบเทเลอร์แบบให้ตัดสินใจเลือกเอง และแบบทดสอบเทเลอร์แบบทดสอบตามสภาพกับแบบทดสอบเทเลอร์แบบทดลองทำ แตกต่างกันอย่างไม่มีนัยสำคัญ ซึ่งเรื่องค่าความคลาดเคลื่อนของค่าความสามารถจากน้อยที่ได้จากแบบทดสอบน้อยไปมาก คือ 1) แบบทดสอบเทเลอร์แบบแตกแขนง 2) แบบทดสอบอิงเกณฑ์ 30 ข้อ 3) แบบทดสอบอิงเกณฑ์ 20 ข้อ และแบบทดสอบเทเลอร์แบบให้ตัดสินใจเลือกเอง 4) แบบทดสอบเทเลอร์ตามสภาพและแบบทดสอบเทเลอร์แบบทดลองทำ 5) แบบทดสอบอิงเกณฑ์ 10 ข้อ

อภิปรายผล

1. ค่าความเที่ยงตรงในการจำแนกการรอบรู้ของแบบทดสอบอิงเกณฑ์ และแบบทดสอบเทเลอร์แต่ละฉบับมีค่าเป็นบวกอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ซึ่งเป็นไปตามสมมติฐานข้อ 1 แบบทดสอบอิงเกณฑ์ 10, 20, และ 30 แบบทดสอบเทเลอร์แบบแตกแขนง (Branched Test) แบบทดสอบตามสภาพ (Status Tailoring) แบบทดลองทำ (Performance Tailoring) และแบบให้ตัดสินใจเลือกทำเอง (Self-Tailoring) มีความเที่ยงตรงในการจำแนกความรอบรู้ทุกรูปแบบ ซึ่งเมื่อพิจารณาค่าความเที่ยงตรงในการจำแนกจากค่าสัมประสิทธิ์สหสัมพันธ์ฟาย (ϕ) เปรียบเทียบกับค่าฟายสูงสุด (ϕ_{max}) จะเห็นว่ากลุ่มของการทดสอบโดยใช้แบบทดสอบเทเลอร์แนวโน้มให้ความเที่ยงตรงในการจำแนกความรอบรู้ได้สูงกว่าคือค่าฟาย (ϕ) เทียบกับฟายสูงสุด (ϕ_{max}) ของแบบทดสอบเทเลอร์แบบแตกแขนง แบบทดสอบเทเลอร์แบบตามสภาพ แบบทดสอบเทเลอร์แบบทดลองทำ และแบบทดสอบเทเลอร์แบบให้ตัดสินใจเลือกทำเอง เท่ากับ 1.00, .89, .76 และ .71 ตามลำดับในขณะที่แบบทดสอบอิงเกณฑ์ 10, 20, และ 30 ข้อ ให้ค่าฟาย (ϕ) เทียบกับฟายสูงสุดเท่ากับ .59, .69 และ .67 ตามลำดับแสดงให้เห็นแนวโน้มให้ค่าความเที่ยงตรงในการจำแนกความรอบรู้อยู่ในระดับปานกลาง ไม่สูงเท่าที่ควร ซึ่งสอดคล้องกับการศึกษา ของคิงส์เบอรีและไวส์ (Kingsbury and Weiss, 1983 cited in Weiss and Kingsbury, 1984 : 365) ซึ่งใช้แบบทดสอบแบบเดิม และการทดสอบ

เทเลอร์ที่ใช้ในการจำแนกความรอบรู้ (AMT : Adaptive Mastery Testing) พบว่าแบบทดสอบแบบเดิมที่มีความยาวของข้อสอบ 10, 20 และ 50 ข้อ โดยให้คะแนนด้วยวิธีสัดส่วนของการตอบถูกและให้คะแนนโดยวิธีเบสส์ ความเที่ยงตรงในการจำแนกความรอบรู้ในแบบทดสอบแบบเดิม มีความสัมพันธ์กันสูงกับความยาวของแบบทดสอบ คือ เมื่อเพิ่มจำนวนข้อสอบแล้วจะทำให้ความเที่ยงตรงในการจำแนกสูงขึ้น และวิธีการให้คะแนนวิธีเบสส์ก็ช่วยให้การจำแนกในการรอบรู้มีความเที่ยงตรงสูงขึ้นกว่าการให้คะแนนโดยใช้วิธีสัดส่วนการตอบถูก การทดสอบเทเลอร์ที่จำแนกการรอบรู้ด้วย AMT ให้ความเที่ยงตรงในการจำแนกสูง และใช้จำนวนข้อสอบน้อยกว่าการทดสอบแบบเดิม

2. จากการเปรียบเทียบความถูกต้อง ในการจำแนกความรอบรู้ของแต่ละแบบทดสอบ โดยความถูกต้องในการจำแนกความรอบรู้นั้น เปรียบเทียบกับสถานะการรอบรู้ที่แท้จริง ของนักเรียนว่าจำแนกได้ตรงกับสภาพความเป็นจริงหรือไม่ ซึ่งจากผลการวิเคราะห์พบว่า มีความแตกต่างกันในการจำแนกความรอบรู้อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และเมื่อทดสอบเป็นรายคู่พบว่า การทดสอบเทเลอร์แบบทดสอบตามสภาพ (Status Tailoring) ให้ความถูกต้องในการจำแนกความรอบรู้ได้สูงกว่าแบบทดสอบอิงเกณฑ์ 10, 20 และ 30 ข้อ อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ซึ่งเป็นไปตามสมมติฐานข้อที่ 2 มีข้อสังเกตประการหนึ่งของการใช้สถิติทดสอบคอกคราน (Cochran Test) ที่ใช้ในการวิเคราะห์ผลการเปรียบเทียบข้อนี้ ที่ว่าจะมีประสิทธิภาพลดลงเมื่อกลุ่มตัวอย่างมีขนาดเล็กเกินไป แต่ก็ไม่มีข้อเสนอแนะว่าควรจะมีขนาดเท่าไร (บุณผา อนันต์สุชาติกุล 2529 หน้า 156) ดังนั้นหากใช้จำนวนกลุ่มตัวอย่างมากกว่าที่ใช้ในการศึกษาครั้งนี้ อาจจะทำให้ข้อพิจารณาที่ชัดเจนมากกว่านี้ ดังพิจารณาจากค่าเฉลี่ยของความถูกต้องในการจำแนกความรอบรู้แล้ว จะเห็นว่า ในกลุ่มของแบบทดสอบเทเลอร์ แบบแตกแขนง แบบทดสอบตามสภาพ แบบทดลองทำ และแบบให้ตัดสินใจเลือกเอง มีแนวโน้มให้เห็นว่ามีความถูกต้องในการจำแนกสูงกล่าวคือ .83, .92 .88 และ .85 ตามลำดับ ในขณะที่แบบทดสอบอิงเกณฑ์ 10, 20 และ 30 ข้อ ให้ความถูกต้องในการจำแนกความรอบรู้ .77, .81 และ .82 ตามลำดับ

3. จากการเปรียบเทียบค่าระดับความสามารถ ที่ได้จากการตอบแบบทดสอบอิงเกณฑ์ และแบบทดสอบเทเลอร์ พบว่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ มีข้อสังเกตอย่างหนึ่งคือค่า F คำนวณซึ่งมีค่า 2.11 และค่า F ตาราง 2.12 จะเห็นว่าค่า F คำนวณอยู่ใกล้จุดที่เกือบจะมีนัยสำคัญทางสถิติ ดังนั้นน่าจะเป็นข้อควรพิจารณาอีกประการหนึ่ง หากมีการเพิ่มกลุ่มตัวอย่างที่ใช้ในการวิจัย หรือเปลี่ยนเกณฑ์ในการจำแนกความรู้ที่แท้จริง อาจให้ข้อพิจารณาที่ชัดเจนยิ่งขึ้น

4. จากการเปรียบเทียบค่าสาระประโยชน์ ของข้อสอบจากแบบทดสอบอิงเกณฑ์ 10, 20 และ 30 ข้อ และแบบทดสอบเทเลอร์แบบแตกแขนง แบบทดสอบตามสภาพ แบบทดลองทำ และแบบให้ตัดสินใจเลือกเอง พบว่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 และเมื่อเปรียบเทียบเป็นรายคู่พบว่า แบบทดสอบอิงเกณฑ์ทั้ง 3 ฉบับมีแสดงสาระประโยชน์เฉลี่ยแตกต่างกันอย่างไม่มีนัยสำคัญ และต่ำกว่าแบบทดสอบเทเลอร์ทั้ง 4 แบบอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ซึ่งสอดคล้องกับสมมติฐานข้อที่ 4 โดยหลักการของการทดสอบเทเลอร์ ที่มุ่งจัดข้อสอบให้เหมาะสมกับผู้สอบ หรือ ใกล้เคียงความสามารถของผู้สอบแต่ละคนมากกว่าการทดสอบแบบเดิม (Lord, 1977 : 125) เมื่อเป็นเช่นนี้จึงทำให้ค่าอินฟอร์เมชันของการทดสอบเทเลอร์สูง ดังพิจารณาได้จากฟังก์ชันของค่าสาระประโยชน์ของข้อสอบ ซึ่งค่าสาระประโยชน์จะสูงสุดเมื่อค่าความสามารถกับค่าความยากของข้อสอบได้ผลลัพธ์เท่ากับศูนย์ (Wright and Stone, 1979 : 75)

5. จากการเปรียบเทียบค่าความคลาดเคลื่อนมาตรฐานเฉลี่ย ของค่าความสามารถ จากผลการตอบแบบทดสอบอิงเกณฑ์ 10, 20 และ 30 ข้อ และแบบทดสอบเทเลอร์ทั้ง 4 รูปแบบ พบว่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 และเมื่อเปรียบเทียบรายคู่พบว่าค่าความคลาดเคลื่อนมาตรฐานเฉลี่ยของค่าความสามารถจากแบบทดสอบ เกือบทุกคู่แตกต่างกันอย่างมีนัยสำคัญที่ระดับ .01 ยกเว้น 2 คู่คือแบบทดสอบเทเลอร์แบบทดสอบตามสภาพกับแบบทดลองทำ และแบบทดสอบอิงเกณฑ์ 20 ข้อ กับแบบทดสอบเทเลอร์แบบให้ตัดสินใจเลือกทำเอง แตกแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ซึ่งเมื่อเรียงค่าความคลาดเคลื่อนมาตรฐานเฉลี่ยของค่าความสามารถจากมากไปหาน้อย คือ 1) แบบทดสอบอิงเกณฑ์ 10 ข้อ 2) แบบทดสอบเทเลอร์แบบทดสอบตาม

ตามสภาพกับแบบทดสอบเทเลอร์แบบทดลองทำ 3) แบบทดสอบอิงเกณฑ์ 20 ข้อกับแบบทดสอบเทเลอร์แบบให้ตัดสินใจเลือกเอง แบบทดสอบอิงเกณฑ์ 30 ข้อ และ 5) แบบทดสอบเทเลอร์แบบแตกแขนง ตามลำดับ ที่เป็นเช่นนั้นเนื่องจากการประมาณค่าความคลาดเคลื่อนมาตรฐานของค่าความสามารถนั้นจะขึ้นอยู่กับจำนวนข้อสอบที่ใช้ในการสอบด้วย ดังจะพิจารณาได้จากสูตร C_{max}/L ซึ่งจะเห็นว่าข้อสอบมีความยาวมากก็จะยิ่งทำให้ค่าประมาณความคลาดเคลื่อนของความสามารถลดลง ดังนั้นค่าความคลาดเคลื่อนมาตรฐานของค่าความสามารถที่ได้จากแบบทดสอบอิงเกณฑ์ และแบบทดสอบเทเลอร์จึงลดต่ำลงตามความยาวของข้อสอบที่ใช้สอบ ดังจะเห็นได้จากแบบทดสอบอิงเกณฑ์ 10 ข้อ มีค่าความคลาดเคลื่อนมาตรฐานของค่าความสามารถสูงสุด และแบบทดสอบเทเลอร์แบบแตกแขนงซึ่งแต่ละคนจะต้องตอบข้อสอบ 30 ข้อ มีค่าต่ำสุด

จากการอภิปรายผลการศึกษาก่อนหน้านี้แล้วอาจกล่าวถึงข้อดี และข้อจำกัดของแบบทดสอบแต่ละรูปแบบเพื่อประโยชน์ในการนำไปใช้ในการเรียนการสอนได้ดังนี้

1. แบบทดสอบอิงเกณฑ์ ซึ่งความยาวของแบบทดสอบเป็น 10, 20 และ 30 ข้อ การนำแบบทดสอบที่จัดในลักษณะของแบบทดสอบอิงเกณฑ์ไปใช้ในการจำแนกการรอบรู้ มีข้อควรคำนึงถึงคือ จำนวนข้อที่ใช้ในการทดสอบ ทั้งนี้เพราะถ้าใช้จำนวนข้อน้อยก็จะทำให้ความเที่ยงตรงในการจำแนกการรอบรู้ต่ำ เนื่องจากการใช้แบบทดสอบอิงเกณฑ์เพื่อจำแนกการรอบรู้ค่าความเที่ยงตรงในการจำแนกการรอบรู้จะมีความสัมพันธ์กันสูงกับความยาวของแบบทดสอบ แต่ถ้าใช้จำนวนข้อมากเกินไปก็จะทำให้มีผลต่อความเที่ยงตรงในการวัดได้เช่นกัน จากผลการศึกษาที่กล่าวมาจะเห็นว่า แบบทดสอบอิงเกณฑ์ทั้ง 3 ฉบับ แม้จะให้ความเที่ยงตรงในการจำแนกการรอบรู้อย่างมีนัยสำคัญแต่ก็อยู่ในเกณฑ์ต่ำ ซึ่งความถูกต้องในการตัดสินใจจำแนกการรอบรู้ก็ให้ผลถูกต้องแนวโน้มต่ำกว่าการทดสอบเทเลอร์ โดยเฉพาะแบบทดสอบอิงเกณฑ์ 10 ข้อ และให้ค่าสาระประโยชน์ในการสอบต่ำสุด

2. แบบทดสอบเทเลอร์แบบแตกแขนง (Branched Test) เป็นแบบทดสอบที่มีความเที่ยงตรงในการจำแนกการรอบรู้ มีความถูกต้องในการจำแนกการรอบรู้แนวโน้มสูง ค่าสาระประโยชน์ของข้อสอบสูง และให้ค่าความคลาดเคลื่อนมาตรฐานของค่าความสามารถต่ำสุด

3. แบบทดสอบเทเลอร์แบบทดสอบตามสภาพ (Status Tailoring) เป็นแบบทดสอบที่ให้ความเที่ยงตรงในการจำแนกความรู้ และมีความถูกต้องในการจำแนกความรู้ไปในแนวโน้มนสูง ให้ค่าสาระประโยชน์ของข้อสอบสูง แต่มีค่าความคลื่อนมาตรฐานของค่าความสามารถสูง

4. แบบทดสอบเทเลอร์แบบทดลองทำ (Performance Tailoring) มีความเที่ยงตรงในการจำแนกความรู้ และมีความถูกต้องในการจำแนกความรู้ไปในแนวโน้มนสูง ให้ค่าสาระประโยชน์ของข้อสอบสูง แต่มีค่าความคลาดเคลื่อนมาตรฐานของค่าความสามารถสูง

5. แบบทดสอบเทเลอร์แบบให้ตัดสินใจเลือกเอง (Self Tailoring) มีความเที่ยงตรงในการจำแนกความรู้ และมีความถูกต้องในการจำแนกความรู้ไปในแนวโน้มนสูง ให้ค่าสาระประโยชน์ของข้อสอบค่อนข้างสูง แต่มีค่าความคลาดเคลื่อนมาตรฐานของค่าความสามารถค่อนข้างสูง

จากผลการวิจัยครั้งนี้ ผู้วิจัยเห็นว่าแบบทดสอบเทเลอร์ทั้ง 4 รูปแบบมีความเหมาะสมเพื่อใช้ในการจำแนกความรู้ของนักเรียนในแต่ละราย จุดประสงค์การเรียนรู้มากกว่าแบบทดสอบอิงเกณฑ์ เนื่องจากแบบทดสอบเทเลอร์ทั้ง 4 รูปแบบให้ความเที่ยงตรงในการจำแนกความรู้สูง ให้ค่าสาระประโยชน์ของการทดสอบสูง เป็นการเพิ่มแรงจูงใจให้นักเรียนอยากทำข้อสอบมากยิ่งขึ้น เพราะทดสอบเทเลอร์ทั้ง 4 รูปแบบมุ่งจัดข้อสอบให้เหมาะสมกับทุกระดับความสามารถของนักเรียน ซึ่งในการทดสอบแบบเดิมไม่ได้คำนึง แต่ในการทดสอบตามทฤษฎีการตอบข้อคำถามคำนึงถึงในเรื่องนี้มาก

ข้อเสนอแนะ

จากผลการวิจัยนี้ ผู้วิจัยมีข้อเสนอแนะดังนี้คือ

ก. เพื่อนำผลวิจัยไปประยุกต์ใช้

1. จากการใช้แบบทดสอบเทเลอร์ทั้ง 4 รูปแบบ ผู้วิจัยเห็นว่าสามารถนำไปใช้กับการเรียนการสอนและการวัดผลที่เป็นอยู่ในขณะนี้ได้ ถ้าเป็นไปได้ เพื่อให้เกิดประโยชน์สูงสุด

และความสะดวกต่อการนำไปใช้ เมื่อจัดข้อสอบเป็นฉบับ ๆ ตามรูปแบบของแบบทดสอบเทเลอร์ แล้ว ควรประมาณค่าความสามารถเป็นตารางสำเร็จไว้เพื่อใช้เทียบจากผลการตอบในแต่ละฉบับ ได้ทันที ซึ่งได้แก่แบบทดสอบแบบแตกแขนง (Branched Test) แบบทดสอบเทเลอร์แบบทดสอบตามสภาพ (Status Tailoring) และแบบทดสอบแบบทดลองทำ

2. สำหรับแบบทดสอบแบบให้ตัดสินใจเลือกทำเอง ก่อนข้างจะมีความยุ่งยากในการดำเนินการสอบ ทั้งนี้ต้องเฉลยผลการสอบของนักเรียนแต่ละคนที่ละข้อ และเป็นรายบุคคล ทั้งนี้เพราะแต่ละคนมีจุดเริ่มต้นในการทำข้อสอบแตกต่างกัน และผลการตอบจะต้องนำไปทดสอบคุณภาพความเหมาะสมของคนในการทำข้อสอบ เพื่อควบคุมคุณภาพของการทดสอบ อย่างไรก็ตาม แม้จะมีกระบวนการที่เพิ่มเติมเข้ามานี้การทดสอบวิธีนี้ก็เป็นการเพิ่มแรงจูงใจในการทำข้อสอบมากที่สุดกว่าวิธีอื่น ๆ เพราะนักเรียนเป็นผู้เลือกจุดเริ่มต้นเอง ผู้วิจัยจึงขอเสนอแนะไว้ 2 ประการคือ

2.1 การสอบอาจไม่จำเป็นต้องให้นักเรียนสอบ 1 ข้อ แล้วก็ตรวจผลคำตอบ แต่อาจให้นักเรียนสอบทีละ 2, 3 หรือ 4 ข้อ แล้วจึงตรวจคำตอบครั้งหนึ่ง ทั้งนี้แล้วแต่ว่าครูจะกำหนด ซึ่งอาจจะพิจารณานักเรียนเป็นรายบุคคลว่านักเรียนคนใดควรทำกี่ข้อแล้วจึงตรวจ

2.2 การควบคุมคุณภาพ ผลการตอบของนักเรียน (Person Fit and Quality Control) ไม่จำเป็นต้องทดสอบผลการตอบของทุกคน ทั้งนี้อาจพิจารณาจากผลการตอบของนักเรียนเพียงอย่างเดียว ก็สามารถบอกได้ว่าผลการตอบนั้นยอมรับหรือไม่ยอมรับ เพราะโดยส่วนใหญ่แล้วในการสอบผลสัมฤทธิ์ส่วนใหญ่ นักเรียนจะใช้ความสามารถสูงสุดของเขาออกมาอยู่แล้ว เพื่อให้ได้คะแนนให้มากที่สุด ดังนั้นผลการตอบของนักเรียนส่วนใหญ่จึงเป็นไปตามที่คาดหวังของหลักการทดสอบโดยวิธีนี้ นอกจากผลการตอบของบางคนเท่านั้นที่พิจารณาเห็นว่าไม่แน่ใจว่าจะยอมรับได้หรือไม่ จึงทดสอบผลการตอบของนักเรียน (Person Fit and Quality Control)

3. การพิจารณาแบบทดสอบเทเลอร์เพื่อนำไปใช้ จำแนกความรอบรู้ของนักเรียน ผู้วิจัยขอเสนอแนะไว้ 2 วิธีที่เห็นว่าเป็นวิธีที่ดีและให้ความถูกต้องมาก โดยยึดเงื่อนไขในการนำไปใช้อยู่ 2 ประการ คือ

3.1 ถ้าครูผู้สอนหรือผู้ใช้แบบทดสอบเทเลออร์มีความใกล้ชิด คลุกคลี รู้ข้อมูลเบื้องต้นหรือระดับความสามารถของผู้เรียนดีพอ ดังเช่น การเรียนการสอนในระดับประถมศึกษา ซึ่งครูผู้สอนจะมีความใกล้ชิด กับนักเรียนมาก ก็ควรเลือกใช้แบบทดสอบเทเลออร์แบบทดสอบตามสภาพ (Status Tailoring)

3.2 ถ้าครูผู้สอนหรือผู้ใช้แบบทดสอบไม่มีความสัมพันธ์ใกล้ชิด หรือในการเรียนการสอนไม่มีปฏิสัมพันธ์ที่จะทำให้รู้ข้อมูลของนักเรียนมากพอ เช่นในระดับมัธยม หรือสูงกว่า ก็ควรเลือกใช้แบบทดสอบเทเลออร์แบบแตกแขนง (Branched Test)

ข. เพื่อการวิจัยต่อไป

1. ควรทำการศึกษาว่า แบบทดสอบเทเลออร์ทั้ง 4 แบบ เมื่อใช้ค่าพารามิเตอร์ที่วิเคราะห์โดยโมเดล 2 และ 3 พารามิเตอร์ จะให้ผลแตกต่างจากการศึกษาครั้งนี้มากน้อยเพียงใด

2. ควรทำการศึกษาว่าในกรณีที่แบบทดสอบที่มีระดับความยากใกล้เคียงกัน แต่มีค่าอำนาจจำแนกแตกต่างกัน จะให้ผลแตกต่างกันมากน้อยเพียงใด

3. ควรทำการศึกษาเปรียบเทียบ การใช้วิธีการประมาณค่าความสามารถโดยวิธีต่าง ๆ ว่า จะมีผลต่อค่าเชื่อมั่น ความเที่ยงตรง และอันดับที่ของผู้สอบ โดยอาศัยแนวความคิดของรูปแบบต่าง ๆ จากทฤษฎีความสามารถแฝงมากน้อยเพียงใด

4. ควรทำการศึกษาโดยใช้วิธีการกำหนดจุดตัดของแบบทดสอบอิงเกณฑ์วิธีต่าง ๆ เช่น วิธีประยุกต์ราชต์โมเดล วิธีเบอร์ก วิธีกำหนดเกณฑ์ผ่านระดับต่ำสุด ฯลฯ ที่ความสอดคล้องกับการจำแนกความรอบรู้ของแบบทดสอบเทเลออร์วิธีต่าง ๆ มากน้อยเพียงใด

5. การสร้างข้อสอบ เนื่องจากข้อสอบที่ใช้เป็นข้อสอบแบบเลือกตอบ จึงควรพิจารณาประสิทธิภาพของตัวลวง โดยใช้ทฤษฎีเดิม (Classical Theory) มาใช้ในการวิเคราะห์ประสิทธิภาพของตัวลวง เมื่อได้คุณภาพที่ดีแล้วจึงนำไปวิเคราะห์ค่าพารามิเตอร์ของข้อสอบ โดยทฤษฎีการตอบข้อคำถาม (Item Response Theory)

6. เปลี่ยนเนื้อหาวิชาที่ใช้ศึกษาเป็นวิชาอื่น ๆ เช่น ภาษาไทย ภาษาอังกฤษ เป็นต้น