

บทที่ 5

สรุป อภิปรายผล และข้อเสนอแนะ

วัตถุประสงค์ของการวิจัย

1. เพื่อศึกษาความเที่ยงตรงของการตัดสินความรอบรู้ตามจุดประสงค์การเรียนรู้ในวิชาคณิตศาสตร์ของนักเรียนชั้นประถมศึกษาปีที่ 6 ตามผลการทดสอบด้วยแบบทดสอบแบบเทเลอร์ 3 รูปแบบคือ แบบปิรามิดที่ใช้ขนาดชิ้นคงที่จำนวน 6 ชิ้นและ 10 ชิ้น แบบปิรามิดที่มีข้อสอบหลายข้อในแต่ละชิ้นจำนวน 3 ชิ้นและ 5 ชิ้น และแบบเปลี่ยนระดับจำนวน 6 ชิ้นและ 10 ชิ้น
2. เพื่อเปรียบเทียบความแตกต่างของผลการตัดสินความรอบรู้ระหว่างแบบทดสอบแบบเทเลอร์ 3 รูปแบบที่มีจำนวนข้อที่นักเรียนต้องตอบเท่ากับกับผลการตัดสินตามปกติของครู
3. เพื่อเปรียบเทียบความแตกต่างของผลการตัดสินความรอบรู้ระหว่างแบบทดสอบแบบเทเลอร์รูปแบบเดียวกันแต่มีจำนวนข้อที่นักเรียนต้องตอบต่างกัน

กลุ่มตัวอย่าง

1. กลุ่มตัวอย่างที่ใช้ในการสร้างกลุ่มข้อสอบเพื่อนำมาจัดเป็นแบบทดสอบแบบเทเลอร์ในการวิจัยนี้ เป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2529 ของจังหวัดอุดรดิตถ์ แพร่ และน่าน จำนวนรวมทั้งสิ้น 4,250 คน จำแนกเป็นนักเรียนที่ตอบแบบทดสอบฉบับ ก 1,071 คน ฉบับ ข 1,057 ฉบับ ค 1,066 คน และฉบับ ง 1,056 คน
2. กลุ่มตัวอย่างที่ใช้ในการวิจัยเป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2531 ของโรงเรียนอนุบาลอุดรดิตถ์ 1 ห้องเรียน จำนวน 33 คน ที่ได้มาจากการสุ่มแบบกลุ่ม (Cluster Sampling) โดยมีห้องเรียนเป็นหน่วยการสุ่ม

เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัยเป็นแบบทดสอบวัดผลสัมฤทธิ์วิชาคณิตศาสตร์ที่มีประถมศึกษาปีที่ 6 แบบเลือกตอบ 5 ตัวเลือกที่วัดในจุดประสงค์การเรียนรู้ที่ 4 "เขียนเลขในรูปยกกำลัง หาท.ร.ม. และ ค.ร.น. ได้" จำนวน 119 ข้อ ที่วิเคราะห์ค่าพารามิเตอร์ประจำข้อสอบแบบโลจิสติกส์ที่มี 3 พารามิเตอร์ นำมาใช้จัดเป็นแบบทดสอบแบบเทเลอร์ต่าง ๆ ดังนี้

1. แบบทดสอบแบบเทเลอร์แบบปิรามิดที่ใช้ขนาดขั้นคงที่ 2 ฉบับ คือ ขนาด 6 ขั้น และขนาด 10 ขั้น ใช้กฎการแยกทางแบบเพิ่ม 1 / ลด 1
2. แบบทดสอบแบบเทเลอร์แบบปิรามิดที่มีข้อสอบหลายข้อในแต่ละขั้น 2 ฉบับ คือ ขนาด 3 ขั้น และขนาด 5 ขั้น ใช้กฎการแยกทางแบบตอบถูกต้อง 2 ข้อ ในขั้นถัดไปเพิ่ม 1 ตอบผิด 1 หรือ 2 ข้อ ลด 1
3. แบบทดสอบแบบเทเลอร์แบบเปลี่ยนระดับ 2 ฉบับ คือ ขนาด 6 ขั้น และขนาด 10 ขั้น
4. แบบทดสอบความรอบรู้แบบเทเลอร์ (AMT : Adaptive Mastery Test) ที่ดำเนินการสอบตามชุดวิธีของเบย์ โดยใช้คอมพิวเตอร์

การเก็บรวบรวมข้อมูล

มีขั้นตอนในการเก็บรวบรวมข้อมูลดังนี้

1. ดำเนินการทดสอบความรอบรู้ของนักเรียนที่เป็นกลุ่มตัวอย่างโดยใช้แบบทดสอบ AMT ณ ห้องศูนย์คอมพิวเตอร์ วิทยาลัยครูอุดรดิตถ์ เครื่องคอมพิวเตอร์จะเลือกข้อสอบให้นักเรียนทำทีละ 1 ข้อตามเงื่อนไขในชุดวิธีของเบย์ เมื่อนักเรียนตอบข้อสอบโดยการกดแป้นคอมพิวเตอร์แล้วเครื่องจะตรวจว่านักเรียนตอบถูกหรือผิดแล้วจึงประมาณค่าความสามารถและความคลาดเคลื่อน

All rights reserved

ในการประมาณค่าความสามารถ นำค่าที่ประมาณได้ไปตรวจสอบตามเงื่อนไขที่กำหนด แล้วเลือกข้อสอบข้อใหม่ให้นักเรียนทำหรือยุติการสอบแล้วแต่กรณี เมื่อการสอบเสร็จสิ้นลง จะทราบค่าความสามารถของนักเรียนเป็นรายบุคคล

2. ดำเนินการสอบโดยใช้แบบทดสอบแบบเทเลอร์แบบปิรามิดที่ใช้ขนาดชั้นคงที่ แบบปิรามิดที่มีหลายข้อในแต่ละชั้นและแบบเปลี่ยนระดับเป็นรายบุคคล โดยแบ่งนักเรียนเป็น 6 กลุ่ม หมุนเวียนกันเข้าสอบจนครบทุกรูปแบบของแบบทดสอบเทเลอร์

การวิเคราะห์ข้อมูล

1. นำผลการสอบด้วยแบบทดสอบแบบเทเลอร์แต่ละแบบ มาประมาณค่าความสามารถของนักเรียนเป็นรายบุคคล โดยใช้สูตรของเบเกิร์ตและไวส์ และเมื่อรวมผลการประมาณค่าความสามารถจากแบบทดสอบ AMT ด้วยแล้ว นักเรียนแต่ละคนจะมีค่าความสามารถ 7 ค่า

2. ตัดสินความรอบรู้ของนักเรียนแต่ละคน โดยนำค่าความสามารถที่ประมาณได้เปรียบเทียบกับค่าความสามารถที่เป็นเกณฑ์ คือ 0.6329 ซึ่งเป็นความเห็นเฉลี่ยของครูผู้สอน ถ้าค่าความสามารถสูงกว่าหรือเท่ากับเกณฑ์ตัดสินว่ารอบรู้ แต่ถ้าต่ำกว่าเกณฑ์ตัดสินว่าไม่รอบรู้ ดังนั้นนักเรียนแต่ละคน จะได้รับการตัดสินว่ารอบรู้ หรือไม่รอบรู้ 7 ครั้ง และเมื่อรวมผลการตัดสินตามปรกติของครูด้วยแล้ว นักเรียนแต่ละคนจะมีผลการตัดสินความรอบรู้ 8 ครั้ง

3. นำผลการตัดสินมาวิเคราะห์ทางสถิติตามจุดมุ่งหมายและสมมุติฐานการวิจัยดังนี้

3.1 หาค่าความเที่ยงตรงในการตัดสินความรอบรู้ของแบบทดสอบแบบเทเลอร์แบบปิรามิดที่ใช้ขนาดชั้นคงที่ แบบปิรามิดที่มีหลายข้อในแต่ละชั้น และแบบเปลี่ยนระดับ โดยหาค่าสหสัมพันธ์แบบฟาย (Phi - correlation) กับผลการตัดสินของแบบทดสอบ AMT

3.2 เปรียบเทียบผลการตัดสินความรอบรู้ของนักเรียนระหว่างแบบทดสอบเทเลอร์ที่ต่างรูปแบบกันแต่มีจำนวนข้อที่นักเรียนต้องตอบเท่ากัน แบบทดสอบ AMT และผลการตัดสินตามปรกติของครู โดยใช้สถิติทดสอบครอกแควน (Cochran Test) และทดสอบภายหลัง (Post - hoc Comparison) โดยวิธีของดัน - บอนเฟอโรน (Dunn - Bonferroni)

3.3 เปรียบเทียบผลการตัดสินความรอบรู้ของนักเรียน ระหว่างแบบทดสอบแบบ เทเลอร์รูปแบบเดียวกันแต่มีจำนวนข้อที่นักเรียนต้องตอบต่างกัน โดยใช้สถิติทดสอบซี สำหรับ สัดส่วนแบบกลุ่มสัมพันธ์

ผลการวิเคราะห์ข้อมูล

1. แบบทดสอบแบบเทเลอร์ ที่ทดลองใช้และนำมาตัดสินความรอบรู้ทุกแบบและขนาด ความยาว ให้ผลการตัดสินที่มีความเที่ยงตรงอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 โดยมีค่า สหสัมพันธ์แบบพายกับผลการตัดสินจากแบบทดสอบ AMT ที่ใช้เป็นเกณฑ์ ตามลำดับจากมากไปหาน้อย คือ

- | | |
|---|-----------------|
| 1.1 แบบเปลี่ยนระดับ 6 ชั้น | $\phi = 0.8018$ |
| 1.2 แบบปิรามิดที่มีหลายข้อในแต่ละชั้นจำนวน 3 ชั้น | $\phi = 0.7926$ |
| 1.3 แบบปิรามิดที่ใช้ขนาดชั้นคงที่ 10 ชั้น | $\phi = 0.6667$ |
| 1.4 แบบปิรามิดที่ใช้ขนาดชั้นคงที่ 6 ชั้นและแบบปิรามิดที่มีหลายข้อในแต่ละชั้น จำนวน 5 ชั้น | $\phi = 0.6500$ |
| 1.5 แบบเปลี่ยนระดับ 10 ชั้น | $\phi = 0.5129$ |

แต่เมื่อนำไปเปรียบเทียบกับค่าพายสูงสุด (ϕ_{max}) พบว่า แบบปิรามิดที่ใช้ขนาด ชั้นคงที่ 10 ชั้นให้ค่าอัตราส่วน ϕ / ϕ_{max} สูงสุดเป็น 1.00 รองลงไปได้แก่แบบเปลี่ยน ระดับ 6 ชั้น แบบปิรามิดที่มีหลายข้อในแต่ละชั้นจำนวน 3 ชั้น แบบปิรามิดที่มีหลายข้อใน แต่ละชั้นจำนวน 5 ชั้น แบบปิรามิดที่ใช้ขนาดชั้นคงที่ 6 ชั้นและแบบเปลี่ยนระดับ 10 ชั้น ให้ค่าอัตราส่วนเป็น 0.8572 0.8499 0.8125 0.8125 และ 0.5500 ตามลำดับ

2. การเปรียบเทียบผลการตัดสินความรอบรู้ของนักเรียน ระหว่างแบบทดสอบ AMT แบบปิรามิดที่ใช้ขนาดชั้นคงที่จำนวน 6 ชั้น แบบปิรามิดที่มีหลายข้อในแต่ละชั้นจำนวน 3 ชั้น แบบเปลี่ยนระดับ 6 ชั้น และผลการตัดสินตามปรกติของครูพบว่า ผลการตัดสินมีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 โดยมีค่าสถิติทดสอบครอกแคเรน $Q = 27.2941$

($\alpha = .01$ $df = 4$ $\chi^2_{\alpha} = 13.28$) จึงทดสอบภายหลังที่ละคู่พบว่า ผลการตัดสินใจของแบบทดสอบแบบเทเลอร์ทุกรูปแบบกับผลการตัดสินใจของแบบทดสอบ AMT ไม่มีความแตกต่างกัน แต่ทุกรูปแบบดังกล่าวนี้แตกต่างจากผลการตัดสินใจตามปกติของครูอย่างมีนัยสำคัญทางสถิติ

3. การเปรียบเทียบผลการตัดสินใจความรอบรู้ของนักเรียน ระหว่างแบบทดสอบ AMT แบบปิรามิดที่ใช้ขนาดชั้นคงที่ 10 ชั้น แบบปิรามิดที่สองข้อในแต่ละชั้นจำนวน 5 ชั้น แบบเปลี่ยนระดับ 10 ชั้น และผลการตัดสินใจตามปกติของครู พบว่าการทดสอบรวมมีความแตกต่างอย่างมีนัยสำคัญทางสถิติ และเมื่อทดสอบเป็นรายคู่ พบว่ามีความแตกต่างกันเพียง 2 คู่ คือ ระหว่างแบบทดสอบ AMT กับการตัดสินใจตามปกติของครู และระหว่างแบบทดสอบเปลี่ยนระดับ 10 ชั้น กับการตัดสินใจตามปกติของครู

4. การเปรียบเทียบผลการตัดสินใจความรอบรู้ระหว่างแบบทดสอบเทเลอร์รูปแบบเดียวกัน แต่มีจำนวนข้อที่นักเรียนต้องตอบต่างกัน คือ 6 ข้อกับ 10 ข้อ โดยใช้สถิติทดสอบทีสำหรับสัดส่วนแบบกลุ่มสัมพันธ์พบว่า ผลการตัดสินใจไม่แตกต่างกันทางสถิติทั้ง 3 รูปแบบ คือ ปิรามิดที่ใช้ขนาดชั้นคงที่ ปิรามิดที่มีหลายข้อในแต่ละชั้นและเปลี่ยนระดับ

อภิปรายผล

1. แบบทดสอบแบบเทเลอร์ที่ทดลองใช้ เพื่อศึกษาความเที่ยงตรงในการตัดสินใจความรอบรู้ครั้งนี้ทั้ง 3 รูปแบบ คือ แบบปิรามิดที่ใช้ขนาดชั้นคงที่ แบบปิรามิดที่มีหลายข้อในแต่ละชั้นและแบบเปลี่ยนระดับ มีความเที่ยงตรงในการตัดสินใจความรอบรู้ของนักเรียนอย่างมีนัยสำคัญทางสถิติ ซึ่งเป็นไปตามสมมุติฐานที่กำหนดไว้ ผลการวิจัยนี้สอดคล้องกับผลการวิจัยของประเวศ เวชชะ (2530) ที่ศึกษาความเที่ยงตรงในการจำแนกความรอบรู้ของแบบทดสอบแบบเทเลอร์ แบบแตกแขนง แบบตามสภาพ แบบทดลองทำ และแบบให้ตัดสินใจเลือกทำเอง พบว่าความเที่ยงตรงในการจำแนกความรอบรู้มีค่าเป็นบวกและมีนัยสำคัญที่ระดับ .01 เช่นเดียวกัน ผลการวิจัยนี้ยืนยันว่า การจัดแบบทดสอบให้เป็นแบบเทเลอร์รูปแบบปิรามิดที่ใช้ขนาดชั้นคงที่ และใช้กฎการแยกทางแบบเพิ่ม 1 / ลด 1 รูปแบบปิรามิดที่มีสองข้อในแต่ละชั้น และใช้กฎการแยกทาง

แบบตอบถูกทั้ง 2 ข้อ เติม 1 ตอนผิด 1 ข้อหรือ 2 ข้อ ลด 1 และรูปแบบเปลี่ยนระดับไม่ว่าจะเป็นขนาดความยาวของจำนวนข้อที่นักเรียนต้องตอบเป็น 6 ข้อ หรือ 10 ข้อก็ตาม เมื่อนำผลการตอบมาประมาณค่าความสามารถโดยใช้สูตรของเบทท์และไวส์ แล้วตัดสินความรอบรู้ จะให้ผลการตัดสินความรอบรู้ที่มีความเที่ยงตรงสูง

2. เมื่อเปรียบเทียบผลการตัดสินความรอบรู้จากแบบทดสอบความรอบรู้แบบเทเลอร์ตามยุทธวิธีของเบย์ การตัดสินตามปรกติของครูและการตัดสินจากแบบทดสอบเทเลอร์ 3 รูปแบบที่ทดลองใช้ ขนาดความยาว 6 ข้อ พบว่า ให้ผลการตัดสินแตกต่างจากการตัดสินตามปรกติของครูในทุกรูปแบบของแบบทดสอบเทเลอร์ ส่วนขนาดความยาว 10 ข้อ พบว่า มีความแตกต่างจากการตัดสินตามปรกติของครูในรูปแบบเปลี่ยนระดับ อาจสรุปได้ว่า ผลการตัดสินจากแบบทดสอบเทเลอร์ มีความแตกต่างจากผลการตัดสินตามปรกติของครูซึ่งเป็นไปตามสมมติฐาน ในกรณีนี้ ถ้าเราเชื่อมั่นในความถูกต้องของกระบวนการสอบตามยุทธวิธีของเบย์ ก็สามารถสรุปได้ว่าผลการตัดสินจากการสอบด้วยแบบทดสอบเทเลอร์ในการวิจัยนี้ มีความถูกต้องสูงกว่าการตัดสินตามปรกติของครู ทั้งนี้เพราะ ผลการตัดสินจากแบบทดสอบเทเลอร์ไม่แตกต่างจากผลการตัดสินตามยุทธวิธีของเบย์ ในขณะที่ผลการตัดสินตามปรกติของครู แตกต่างจากผลการตัดสินตามยุทธวิธีของเบย์ ซึ่งสอดคล้องกับผลการวิจัยของคิงส์เบอร์รี่และไวส์ (Kingsbury and Weiss, 1979 cited by Weiss and Kingsbury, 1984 : 365)

ยิ่งไปกว่านี้ ในด้านจำนวนข้อของข้อสอบที่นักเรียนต้องตอบในแบบทดสอบเทเลอร์ของการวิจัยนี้มีเพียง 6 ข้อ หรือ 10 ข้อ แต่ในการสอบตามปรกติของครู ในที่นี้ใช้ข้อสอบ 20 ข้อ ซึ่งแตกต่างกันมาก ยิ่งเป็นข้อสั้นสนับสนุนว่า แบบทดสอบแบบเทเลอร์ที่ทดลองใช้ครั้งนี้มีประสิทธิภาพผลสูงกว่าการวัดผลประเมินผลตามปรกติของครู

3. ในการเปรียบเทียบผลการตัดสินความรอบรู้ระหว่างแบบทดสอบเทเลอร์รูปแบบเดียวกัน แต่มีจำนวนข้อที่นักเรียนต้องตอบต่างกัน พบว่าให้ผลการตัดสินไม่แตกต่างกัน ซึ่งไม่เป็นไปตามสมมติฐานการวิจัย ผลการศึกษาของคิงส์เบอร์รี่และไวส์ (Kingsbury and Weiss, 1984 cited by Weiss and Kingsbury, 1984 : 365) พบว่า ในการทดสอบแบบดั้งเดิมนั้น เมื่อเพิ่มจำนวนข้อสอบให้นักเรียนตอบ จะทำให้ความเที่ยงตรงในการจำแนกสูงขึ้น

ในกรณีนี้อาจเป็นไปได้ว่า จำนวนข้อสอบที่นักเรียนต้องตอบในการสอบด้วยแบบทดสอบเทเลอร์ในการวิจัยนี้มีจำนวนข้อแตกต่างกันเพียง 4 ข้อ ซึ่งน้อยเกินไป จึงยังไม่สามารถส่งผลให้การตัดสินมีความแตกต่างกัน หรืออาจเป็นไปได้ว่าจำนวนข้อที่นักเรียนต้องตอบต่างกันในการทดสอบด้วยแบบทดสอบเทเลอร์ ส่งผลต่อการตัดสินความรอบรู้ของนักเรียน ผลการตัดสินจึงไม่แตกต่างกัน ซึ่งถ้าเป็นไปได้ตามกรณีหลัง แสดงว่าประสิทธิภาพในการจำแนกความรอบรู้ของแบบทดสอบเทเลอร์ไม่ขึ้นกับขนาดความยาวหรือจำนวนข้อสอบที่นักเรียนต้องตอบ

ข้อเสนอแนะ

1. ข้อเสนอแนะในด้านการนำไปใช้

จากผลการวิจัยนี้จะเห็นได้ว่าการทดสอบแบบเทเลอร์รูปแบบปิรามิดที่ใช้ขนาดชั้นคงที่ รูปแบบปิรามิดที่มีหลายข้อในแต่ละชั้น และรูปแบบเปลี่ยนระดับ ทั้งขนาดที่มีจำนวนข้อที่นักเรียนต้องตอบ 6 ข้อ และ 10 ข้อ สามารถนำไปใช้ในการประเมินผลการศึกษาเป็นรายจุดประสงค์การเรียนรู้ได้อย่างมีประสิทธิภาพ และเนื่องจากเป็นรูปแบบที่ไม่จำเป็นต้องใช้คอมพิวเตอร์ในการดำเนินการสอบ จึงมีความเป็นไปได้ที่จะนำไปใช้ในโรงเรียน แต่อย่างไรก็ตามสิ่งที่เป็นปัญหาใหญ่ก็คือ การพัฒนาคลังข้อสอบ (item pool) โดยการวิเคราะห์ข้อสอบด้วยรูปแบบโลจิสติกที่มีพารามิเตอร์ 3 ตัว เพราะต้องใช้กลุ่มตัวอย่างขนาดใหญ่ตั้งแต่ 1,000 คนขึ้นไป จำนวนข้อสอบในกลุ่มข้อสอบมีไม่น้อยกว่า 100 ข้อ ซึ่งต้องวัดในจุดประสงค์เดียวกันตามข้อตกลงเรื่องความเป็นมิติเดียว ดังนั้น การพัฒนาคลังข้อสอบเป็นรายจุดประสงค์อาจต้องดำเนินการในระดับจังหวัดหรือเขตการศึกษา เมื่อมีกลุ่มข้อสอบในแต่ละจุดประสงค์การเรียนรู้มากพอแล้ว จึงนำมาจัดตามรูปแบบของการทดสอบแบบเทเลอร์ต่อไป โรงเรียนต่าง ๆ ก็สามารถขอใช้แบบทดสอบไปใช้ได้ ยิ่งไปกว่านั้นถ้ามีเกณฑ์จุดตัดในระดับต่าง ๆ เช่น ระดับกลุ่มโรงเรียน ระดับอำเภอ ระดับจังหวัด หรือระดับเขต ก็จะสามารถจัดแบบทดสอบที่สอดคล้องกับเกณฑ์ไปดำเนินการได้ ผลการวัด จะช่วยให้ทราบคุณภาพการศึกษาของนักเรียนเป็นรายจุดประสงค์ที่เทียบกับระดับต่าง ๆ ได้ด้วย การดำเนินการดังกล่าวนี้ จะช่วยให้การตัดสินผลการเรียนของนักเรียนมีคุณภาพสูงขึ้น และส่งผลต่อการยกระดับคุณภาพการศึกษาได้

2. ข้อเสนอแนะในด้านการวิจัยต่อไป

จากการดำเนินงานวิจัยเรื่องนี้ ผู้วิจัยได้เกิดแนวคิดบางประการสำหรับการวิจัยที่น่าจะได้ดำเนินการต่อไปดังนี้

2.1 ผลการวิจัยที่พบว่าเมื่อมีจำนวนข้อสอบที่นักเรียนต้องตอบต่างกัน ผลการตัดสินไม่ต่างกันอาจเป็นเพราะจำนวนข้อที่นักเรียนต้องตอบยังแตกต่างกันน้อย จึงไม่พบความแตกต่างในการตัดสินดังข้ออภิปรายที่กล่าวถึงมาแล้ว เพื่อพิสูจน์หาความจริงในเรื่องนี้ น่าจะได้มีการทดลองใช้แบบทดสอบเทเลออร์ที่มีขนาดความยาวของจำนวนข้อที่นักเรียนต้องตอบแตกต่างกันหลาย ๆ ขนาดความยาว

2.2 ในการวิจัยนี้ทดลองกับแบบทดสอบเทเลออร์เพียง 3 รูปแบบเท่านั้น ยังมีรูปแบบของการทดสอบเทเลออร์ที่สามารถดำเนินการโดยไม่ต้องใช้คอมพิวเตอร์ได้ เช่น แบบปริมาตรข้างตัด แบบปริมาตรทึบตันแปรผัน แบบปริมาตรที่ให้น้ำหนักแก่ตัวเลือกของข้อสอบแยกทางเป็นต้น น่าจะได้มีการศึกษาหาความเที่ยงตรงในการตัดสินความรอบรู้ของแบบทดสอบเทเลออร์รูปแบบอื่น ๆ ด้วย

2.3 ในการวิจัยนี้ทดลองใช้การทดสอบเทเลออร์ เพื่อตัดสินความรอบรู้ในจุดประสงค์การเรียนรู้เดียว เพื่อให้สอดคล้องกับข้อตกลงเรื่องความเป็นมิติเดียว ผู้วิจัยเห็นว่าถ้ามีข้อพิสูจน์ยืนยันเรื่องความเป็นมิติเดียวของข้อสอบวิชาใด ๆ (ซึ่งอาจพิสูจน์โดยวิธีวิเคราะห์ตัวประกอบหรือวิธีอื่นใด) ก็น่าจะได้มีการทดลองใช้การทดสอบเทเลออร์ เพื่อการให้ค่าระดับผลการเรียนของนักเรียนว่ามีความเที่ยงตรงในการตัดสินหรือไม่

2.4 ควรมีการศึกษาเชิงความสัมพันธ์ระหว่างผลการประมาณค่าความสามารถที่ต่างวิธีกัน หรือระหว่างรูปแบบการทดสอบเทเลออร์ที่ต่างกัน หรือระหว่างกลุ่มข้อสอบที่วิเคราะห์แบบนารามิเตอร์ 1-2-3 ตัว ซึ่งผลการศึกษาจะช่วยให้ผู้ใช้แบบทดสอบมีแนวทางในการเลือกใช้รูปแบบของแบบทดสอบ เลือกวิธีการประมาณค่าความสามารถ หรือเลือกรูปแบบการวิเคราะห์ข้อสอบได้มากวิธีขึ้น