

การเปรียบเทียบดัชนีความเป็นคู่ขนานของข้อสอบและแบบสอบ
ที่ได้จากการคัดเลือกข้อสอบโดยกลุ่มผู้เชี่ยวชาญที่มีลักษณะแตกต่างกัน:
การประยุกต์ใช้ทฤษฎีฉันทามติทางวัฒนธรรมแบบมีรูปแบบ



ว่าที่ร้อยเอกอนุพงษ์ กันธิวงศ์

จุฬาลงกรณ์มหาวิทยาลัย

บทคัดย่อและแฟ้มข้อมูลฉบับเต็มของวิทยานิพนธ์ตั้งแต่ปีการศึกษา 2554 ที่ให้บริการในคลังปัญญาจุฬาฯ (CUIR)
เป็นแฟ้มข้อมูลของนิสิตเจ้าของวิทยานิพนธ์ ที่ส่งผ่านทางบัณฑิตวิทยาลัย

The abstract and full text of theses from the academic year 2011 in Chulalongkorn University Intellectual Repository (CUIR)
are the thesis authors' files submitted through the University Graduate School.

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาครุศาสตรมหาบัณฑิต

สาขาวิชาการวัดและประเมินผลการศึกษา ภาควิชาวิจัยและจิตวิทยาการศึกษา

คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2560

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

COMPARISON OF ITEM AND TEST PARALLEL INDEX BETWEEN TEST FORMS
THAT SELECTED BY DIFFERENT CHARACTERISTICS OF EXPERTS:
APPLICATION OF FORMAL CULTURAL CONSENSUS THEORY

Acting Captain Anupong Kantiwong



A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Education Program in Educational Measurement and
Evaluation

Department of Educational Research and Psychology

Faculty of Education

Chulalongkorn University

Academic Year 2017

Copyright of Chulalongkorn University

หัวข้อวิทยานิพนธ์

การเปรียบเทียบดัชนีความเป็นคู่ขนานของข้อสอบและ
แบบสอบที่ได้จากการคัดเลือกข้อสอบโดยกลุ่มผู้เชี่ยวชาญ
ที่มีลักษณะแตกต่างกัน: การประยุกต์ใช้ทฤษฎีฉันทามติ
ทางวัฒนธรรมแบบมีรูปแบบ

โดย

ว่าที่ร้อยเอกอนุพงษ์ กันธิวงศ์

สาขาวิชา

การวัดและประเมินผลการศึกษา

อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก

รองศาสตราจารย์ ดร.โชติกา ภาษีผล

คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้รับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่ง
ของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรบัณฑิต

..... คณบดีคณะครุศาสตร์
(รองศาสตราจารย์ ดร.ศิริเดช สุชีวะ)

คณะกรรมการสอบวิทยานิพนธ์

..... ประธานกรรมการ
(ศาสตราจารย์ ดร.ศิริชัย กาญจนวาสี)

..... อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก
(รองศาสตราจารย์ ดร.โชติกา ภาษีผล)

..... กรรมการภายนอกมหาวิทยาลัย
(ผู้ช่วยศาสตราจารย์ ดร.สุวิมล กฤษศยาสา)

อนุพงษ์ กันธิวงศ์ : การเปรียบเทียบดัชนีความเป็นคู่ขนานของข้อสอบและแบบสอบที่ได้จากการคัดเลือกข้อสอบโดยกลุ่มผู้เชี่ยวชาญที่มีลักษณะแตกต่างกัน: การประยุกต์ใช้ทฤษฎีฉันทามติทางวัฒนธรรมแบบมีรูปแบบ (COMPARISON OF ITEM AND TEST PARALLEL INDEX BETWEEN TEST FORMS THAT SELECTED BY DIFFERENT CHARACTERISTICS OF EXPERTS: APPLICATION OF FORMAL CULTURAL CONSENSUS THEORY) อ.ที่ปรึกษาวิทยานิพนธ์หลัก: รศ. ดร.โชติกา ภาษิมผล, 145 หน้า.

การวิจัยในครั้งนี้มีจุดประสงค์ 1) เพื่อศึกษาระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญ 2) เพื่อเปรียบเทียบดัชนีความเป็นคู่ขนานของแบบสอบที่ได้จากการคัดเลือกโดยกลุ่มผู้เชี่ยวชาญที่มีลักษณะแตกต่างกันและ 3) เพื่อเปรียบเทียบดัชนีความเป็นคู่ขนานของข้อสอบที่ได้จากการคัดเลือกโดยกลุ่มผู้เชี่ยวชาญที่มีลักษณะแตกต่างกัน ตัวอย่างที่ใช้ในการศึกษาครั้งนี้ คือ ผู้เชี่ยวชาญที่มีประสบการณ์สอนนักศึกษาแพทยวิทยาลัยแพทยศาสตร์พระมงกุฎเกล้าจากการเลือกแบบเจาะจง 100 คน โดยจัดกลุ่มเพื่อคัดเลือกข้อสอบสำหรับสร้างแบบสอบคู่ขนานตามระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบ เครื่องมือที่ใช้ในการวิจัย ประกอบด้วยแบบวัดระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบและข้อสอบสำหรับคัดเลือกเพื่อสร้างแบบสอบคู่ขนานสำหรับกลุ่มผู้เชี่ยวชาญ ผลการวิจัยสรุปได้ดังนี้ 1) ผู้เชี่ยวชาญที่มีประสบการณ์สอนน้อยกว่า 10 ปี ส่วนใหญ่มีระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบในช่วง 0.40-0.59, 0.60-0.79 และ 0.80-1.00 ตามลำดับ แตกต่างจากผู้เชี่ยวชาญที่มีประสบการณ์สอนตั้งแต่ 10 ปีขึ้นไปที่ 0.60-0.79, 0.40-0.59 และ 0.80-1.00 ตามลำดับ 2) แบบสอบคู่ขนานที่สร้างขึ้นจากการคัดเลือกข้อสอบโดยกลุ่มผู้เชี่ยวชาญที่มีระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบระดับ 0.80-1.00 และมีประสบการณ์สอนตั้งแต่ 10 ปีขึ้นไปมีค่าความเป็นคู่ขนานของแบบสอบสูงกว่ากลุ่มที่อยู่ในระดับต่ำกว่า ($RMSD = 0.283$ [$MSG = 0.020$] และ $MRD = 0.042$ [$MRIG = 0.003$]) และ 3) ในทุกกลุ่มผู้เชี่ยวชาญที่ผ่านเกณฑ์ขั้นต่ำของการตรวจสอบความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบเป็นโดยอาศัยทฤษฎีฉันทามติทางวัฒนธรรมแบบมีรูปแบบมีค่าเฉลี่ยของ $RMSD$ ไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 ($F=0.582$, $p=0.755$)

ภาควิชา วิจัยและจิตวิทยาการศึกษา ลายมือชื่อนิสิต

สาขาวิชา การวัดและประเมินผลการศึกษา ลายมือชื่อ อ.ที่ปรึกษาหลัก

ปีการศึกษา 2560

5883863027 : MAJOR EDUCATIONAL MEASUREMENT AND EVALUATION

KEYWORDS: PARALLEL TEST / EXPERT / CULTURAL CONSENSUS THEORY

ANUPONG KANTIWONG: COMPARISON OF ITEM AND TEST PARALLEL INDEX BETWEEN TEST FORMS THAT SELECTED BY DIFFERENT CHARACTERISTICS OF EXPERTS: APPLICATION OF FORMAL CULTURAL CONSENSUS THEORY. ADVISOR: ASSOC. PROF.SHOTIGA PASIPHOL, Ph.D., 145 pp.

The purposes of this research were 1) to determine consensus competence of the expert about item parallelism 2) to comparison of test parallel index between tests form that selected by different characteristics of experts 3) to comparison of item parallel index between tests form that selected by different characteristics of experts. The sample consisted of 100 instructors of Phramongkutkloa College of Medicine using purposive sampling for devised the expert group by level of consensus competence. The research instruments were consensus competence test for the expert about the item parallelism and item pool for constructed parallel test form. The research findings were as follows: 1) Experts with less than 10 years of teaching experience are most likely to have the consensus competence of the expert about item parallelism of the test in the range of 0.40-0.59, 0.60-0.79 and 0.80-1.00, respectively, unlike experts with higher 10 years of teaching experience to have in the range of 0.60-0.79, 0.40-0.59 and 0.80-1.00, respectively. 2) The parallel tests created by the selection of the items by a group of experts with the consensus competence of the expert about item parallelism of the test in the range of 0.80-1.00 and teaching experience above 10 years has the parallelism index of the test more than lower competences (RMSD = 0.283 [MSG = 0.020] และ MRD = 0.042 [MRIG = 0.003]) and 3) it was found that all groups of experts who passed the minimum criterion of the consensus competence of the expert about item parallelism of the test were not significantly different of the RMSD means ($F=0.582$, $p=0.755$).

Department: Educational Research and Student's Signature

Psychology

Advisor's Signature

Field of Study: Educational Measurement
and Evaluation

Academic Year: 2017

กิตติกรรมประกาศ

วิทยานิพนธ์เล่มนี้สำเร็จลุล่วงได้เป็นอย่างดี โดยได้รับความกรุณาและเมตตาอย่างสูงจากท่านรองศาสตราจารย์ ดร.โชติกา ภาษีผล อาจารย์ที่ปรึกษาวิทยานิพนธ์ ผู้ถ่ายทอดองค์ความรู้ด้านวิธีวิทยาการวิจัย จุดประเด็นเกี่ยวกับงานวิจัยที่ต่อยอดจากการเรียนรู้อันเข้มข้นเรียนประสบการณ์การดำเนินงานในด้านต่างๆ ให้แก่ผู้วิจัยอย่างเต็มที่ รวมถึงสละเวลาอันมีค่ายิ่งในการให้คำแนะนำและคำปรึกษา ตลอดจนแก้ไขข้อบกพร่องต่างๆ แก่ผู้วิจัย รวมทั้งให้กำลังใจ ดูแลเอาใจใส่และกำกับติดตามการจัดทำวิทยานิพนธ์ด้วยความปรารถนาดีแก่ศิษย์อย่างเต็มที่ จนทำให้วิทยานิพนธ์เล่มนี้สำเร็จลุล่วงได้ ผู้วิจัยขอขอบพระคุณท่านอาจารย์เป็นอย่างสูง มา ณ โอกาสนี้

ขอขอบพระคุณคณาจารย์ภาควิชาวิจัยและจิตวิทยาการศึกษาทุกท่านที่ได้ประสาทความรู้ในด้านต่างๆ ที่สำคัญแก่ผู้วิจัยเสมอมา โดยเฉพาะอย่างยิ่ง ศาสตราจารย์ ดร. ศิริชัย กาญจนวาสี รองศาสตราจารย์ ดร.โชติกาภาษีผล รองศาสตราจารย์ ดร.ศิริเดช สุชีวะ รองศาสตราจารย์ ดร.กมลวรรณ ตังธนากานนท์ และผู้ช่วยศาสตราจารย์ ดร.ณัฐภรณ์ หลาวทอง ทั้งนี้ ผู้วิจัยขอขอบพระคุณศาสตราจารย์ ดร. ศิริชัย กาญจนวาสี และผู้ช่วยศาสตราจารย์ ดร.สุวิมล กฤษคฤหาสน์ ที่ให้ความกรุณาเป็นกรรมการสอบวิทยานิพนธ์และได้ให้คำชี้แนะที่เป็นประโยชน์ในการปรับปรุงวิทยานิพนธ์ให้มีความสมบูรณ์มากยิ่งขึ้น

ขอขอบพระคุณ ครู อาจารย์ นักศึกษาแพทย์ เจ้าหน้าที่หน่วยแพทยศาสตรศึกษา วิทยาลัยแพทยศาสตร์พระมงกุฎเกล้าทุกท่านที่ให้ความอนุเคราะห์ในการเก็บข้อมูล อำนวยความสะดวกในการดำเนินการวิจัยเป็นอย่างดี

ขอขอบคุณเพื่อนๆ พี่ๆ น้องๆ สาขาการวัดและประเมินผลการศึกษา ที่คอยช่วยเหลือในการเรียนและการวิจัยตลอดเวลาที่ศึกษาในที่แห่งนี้ และขอขอบคุณอีกหลายท่านที่ยังไม่ได้กล่าวมาที่มีส่วนช่วยให้วิทยานิพนธ์ประสบความสำเร็จได้เป็นอย่างดี

ท้ายนี้ ผู้วิจัยขอขอบพระคุณ คุณแม่ยุพินพร ณะราชา น้องอนุสรรา กันธิวงศ์ และ พี่ภคสร อมรวงศ์ ที่คอยให้กำลังใจ ให้ความช่วยเหลือและสนับสนุนในทุกด้านเสมอมา

สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	ง
บทคัดย่อภาษาอังกฤษ.....	จ
กิตติกรรมประกาศ.....	ฉ
สารบัญ.....	ช
สารบัญตาราง.....	ฅ
สารบัญภาพ	ฎ
บทที่ 1 บทนำ.....	1
ความเป็นมาและความสำคัญของปัญหา	1
คำถามวิจัย	9
วัตถุประสงค์	9
สมมติฐานการวิจัย	9
ขอบเขตการวิจัย.....	10
นิยามศัพท์เฉพาะ.....	10
ประโยชน์ที่ได้รับ.....	12
บทที่ 2 เอกสารและงานวิจัยที่เกี่ยวข้อง.....	13
ตอนที่ 1 แบบสอบถาม.....	14
ตอนที่ 2 ทฤษฎีฉันทามติทางวัฒนธรรม	28
ตอนที่ 3 กรอบแนวคิดในการทำวิจัย	39
บทที่ 3 วิธีดำเนินการวิจัย	41
ประชากรและตัวอย่าง	41
การเก็บรวบรวมข้อมูล	44
เครื่องมือที่ใช้.....	46

การวิเคราะห์ข้อมูล.....	50
บทที่ 4 ผลการวิเคราะห์ข้อมูล.....	53
ตอนที่ 1 เพื่อศึกษาระดับความสามารถในการตัดสินใจด้านความเป็นคู่ขนานของข้อสอบของ ผู้เชี่ยวชาญ.....	53
ตอนที่ 2 เพื่อเปรียบเทียบดัชนีความเป็นคู่ขนานของแบบสอบคู่ขนานที่ได้จากการคัดเลือกโดย กลุ่มผู้เชี่ยวชาญที่มีระดับความสามารถในการตัดสินใจด้านความเป็นคู่ขนานของข้อสอบและ ประสบการณ์สอนนักศึกษาแพทย์ที่แตกต่างกัน.....	65
ตอนที่ 3 เพื่อเปรียบเทียบดัชนีความเป็นคู่ขนานของข้อสอบคู่ขนานที่คัดเลือกโดยกลุ่ม ผู้เชี่ยวชาญที่มีระดับความสามารถในการตัดสินใจด้านความเป็นคู่ขนานของข้อสอบและ ประสบการณ์สอนนักศึกษาแพทย์ที่แตกต่างกัน.....	82
บทที่ 5 สรุปผล อภิปรายและข้อเสนอแนะ.....	86
สรุปผลการวิจัย.....	87
อภิปรายผลการวิจัย.....	91
ข้อเสนอแนะ.....	96
รายการอ้างอิง.....	98
ภาคผนวก.....	102
ภาคผนวก ก เครื่องมือที่ใช้ในการวิจัย.....	103
ภาคผนวก ข ค่าสถิติพื้นฐาน.....	119
ภาคผนวก ค ผลการวิเคราะห์.....	121
ประวัติผู้เขียนวิทยานิพนธ์.....	145

สารบัญตาราง

ตารางที่ 2.1	เปรียบเทียบลักษณะของการสร้างแบบสอบคู่ขนานก่อนและหลังนำไปใช้สอบ	17
ตารางที่ 2.2	การศึกษาดัชนีความเป็นคู่ขนานของแบบสอบ	25
ตารางที่ 2.3	ความน่าจะเป็นของการตอบคำถามถูกต้องของรูปแบบการตอบคำถามของ ผู้เชี่ยวชาญ 2 คน	35
ตารางที่ 2.4	การกำหนดขนาดตัวอย่างและค่าความตรงในระดับความสามารถในการเห็นพ้อง ร่วมกันที่แตกต่างกัน.....	36
ตารางที่ 3.1	การกำหนดขนาดตัวอย่างและค่าความตรงในระดับความสามารถในการตัดสินใจที่ แตกต่างกัน.....	42
ตารางที่ 3.2	การแบ่งระดับความสามารถความสามารถในการตัดสินใจเป็นคู่ขนานของ ข้อสอบของผู้เชี่ยวชาญ	43
ตารางที่ 3.3	นักศึกษาที่ใช้เป็นกลุ่มผู้ให้ข้อมูลจำแนกชั้นปีและชุดข้อสอบที่สร้างขึ้นโดยกลุ่ม ผู้เชี่ยวชาญ.....	43
ตารางที่ 3.4	ลักษณะของข้อสอบจำแนกระบบตามเกณฑ์ความรู้ความสามารถในการประเมิน เพื่อรับใบอนุญาต เป็นผู้ประกอบวิชาชีพเวชกรรม พ.ศ. 2555	49
ตารางที่ 4.1	ข้อมูลภูมิหลังของผู้เชี่ยวชาญที่ทำแบบวัดระดับความสามารถในการตัดสินใจเป็น คู่ขนานของข้อสอบ.....	55
ตารางที่ 4.2	ผลการแบ่งกลุ่มผู้เชี่ยวชาญจำแนกตามระดับความสามารถในการตัดสินใจเป็น คู่ขนานของข้อสอบ	57
ตารางที่ 4.3	ผลการแบ่งกลุ่มผู้เชี่ยวชาญจำแนกตามระดับความสามารถในการตัดสินใจเป็น คู่ขนานของข้อสอบ	57
ตารางที่ 4.4	ผลการแบ่งกลุ่มผู้เชี่ยวชาญจำแนกตามระดับความสามารถในการตัดสินใจเป็น คู่ขนานของข้อสอบ	58
ตารางที่ 4.5	ผลการตรวจสอบระดับความสามารถในการตัดสินใจเป็นคู่ขนานของผู้เชี่ยวชาญ..	60
ตารางที่ 4.6	ผลการแบ่งกลุ่มผู้เชี่ยวชาญจำแนกตามระดับความสามารถในการตัดสินใจเป็น คู่ขนานของข้อสอบ	61

ตารางที่ 4.7 ผลการแบ่งกลุ่มผู้เชี่ยวชาญจำแนกตามระดับความสามารถในการตัดสินความเป็น คู่ขนานของข้อสอบ	62
ตารางที่ 4.8 ข้อมูลภูมิหลังของกลุ่มผู้เชี่ยวชาญที่เป็นตัวอย่างในการคัดเลือกข้อสอบเพื่อสร้าง แบบสอบคู่ขนาน.....	63
ตารางที่ 4.9 ผลการตอบแบบวัดความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบโดย ผู้เชี่ยวชาญ.....	64
ตารางที่ 4.10 ผลการคัดเลือกข้อสอบคู่ขนานของกลุ่มผู้เชี่ยวชาญที่มีลักษณะแตกต่างกัน	66
ตารางที่ 4.11 ผลการตรวจสอบความสอดคล้องของข้อสอบที่คัดเลือกโดยผู้เชี่ยวชาญกับโมเดล ตามทฤษฎีการตอบสนองข้อสอบแบบ 2 พารามิเตอร์ (2PL) โดยใช้สถิติทดสอบไค แอสควร์ (χ^2).....	68
ตารางที่ 4.12 ผลการตรวจสอบความเป็นคู่ขนานของแบบสอบโดยอาศัยทฤษฎีการตอบสนอง ข้อสอบ	72
ตารางที่ 4.13 ผลการวิเคราะห์เปรียบเทียบค่าเฉลี่ยของ MRD ที่ได้จากแบบสอบทั้งฉบับจำแนก ตามกลุ่มผู้เชี่ยวชาญที่มีลักษณะแตกต่างกัน 6 กลุ่ม	80
ตารางที่ 4.14 ผลการวิเคราะห์ความแปรปรวนของ MRD จำแนกตามระดับความสามารถในการ ตัดสินความเป็นคู่ขนานของข้อสอบและประสบการณ์สอน	81
ตารางที่ 4.15 ผลการตรวจสอบความเป็นคู่ขนานของข้อสอบและแบบสอบโดยอาศัยทฤษฎีการ ตอบสนองข้อสอบ	83
ตารางที่ 4.16 ผลการวิเคราะห์เปรียบเทียบค่าเฉลี่ยของ RMSD ที่ได้จากแบบสอบทั้งฉบับ จำแนกตามกลุ่มผู้เชี่ยวชาญที่มีลักษณะแตกต่างกัน 6 กลุ่ม	84

สารบัญภาพ

ภาพที่ 2.1	โค้งลักษณะข้อสอบแบบ 2 พารามิเตอร์.....	19
ภาพที่ 2.2	โค้งฟังก์ชันสารสนเทศของแบบสอบ (TIF) และโค้งลักษณะของแบบสอบ (TCC).....	21
ภาพที่ 2.3	ตัวอย่างการเปรียบเทียบโค้งฟังก์ชันสารสนเทศของแบบสอบ (TIF) และโค้งลักษณะของแบบสอบ (TCC) ของแบบสอบคู่ขนาน 5 ฉบับ	24
ภาพที่ 2.4	กรอบแนวคิดในการทำการวิจัย	40
ภาพที่ 4.1	เปรียบเทียบค่า RMSD และ MSG ระหว่างแบบสอบคู่ขนานที่คัดเลือกข้อสอบ โดยกลุ่มผู้เชี่ยวชาญ (กำหนดเกณฑ์การแปลผลแบบสอบคู่ขนาน; ค่า RMSD และ MSG ต่ำกว่า 0.50).....	73
ภาพที่ 4.2	เปรียบเทียบค่า MRD และ MRIG ระหว่างแบบสอบคู่ขนานที่คัดเลือกข้อสอบโดยกลุ่มผู้เชี่ยวชาญ (กำหนดเกณฑ์การแปลผลแบบสอบคู่ขนาน; ค่า MRD และ MRIG ต่ำกว่า 0.10).....	73
ภาพที่ 4.3	โค้งสารสนเทศของแบบสอบ (Test information) เปรียบเทียบระหว่างแบบสอบคู่ขนาน ชุด 1A และ 1B ที่ระดับความสามารถของผู้สอบ 15 ระดับ.....	74
ภาพที่ 4.4	โค้งสารสนเทศของแบบสอบ (Test information) เปรียบเทียบระหว่างแบบสอบคู่ขนาน ชุด 2A และ 2B ที่ระดับความสามารถของผู้สอบ 15 ระดับ.....	75
ภาพที่ 4.5	โค้งสารสนเทศของแบบสอบ (Test information) เปรียบเทียบระหว่างแบบสอบคู่ขนานชุด 3A และ 3B ที่ระดับความสามารถของผู้สอบ 15 ระดับ	76
ภาพที่ 4.6	โค้งสารสนเทศของแบบสอบ (Test information) เปรียบเทียบระหว่างแบบสอบคู่ขนาน ชุด 4A และ 4B ที่ระดับความสามารถของผู้สอบ 15 ระดับ.....	77
ภาพที่ 4.7	โค้งสารสนเทศของแบบสอบ (Test information) เปรียบเทียบระหว่างแบบสอบคู่ขนาน ชุด 5A และ 5B ที่ระดับความสามารถของผู้สอบ 15 ระดับ.....	78
ภาพที่ 4.8	โค้งสารสนเทศของแบบสอบ (Test information) เปรียบเทียบระหว่างแบบสอบคู่ขนาน ชุด 6A และ 6B ที่ระดับความสามารถของผู้สอบ 15 ระดับ.....	79
ภาพที่ 4.9	กราฟเปรียบเทียบค่าเฉลี่ยของ MRD จำแนกตามระดับความสามารถ ในการตัดสินความเป็นคู่ขนานของข้อสอบ.....	81

ภาพที่ 4.10 เปรียบเทียบค่าเฉลี่ย RMSD ระหว่างแบบสอบคู่ขนานที่คัดเลือกข้อสอบโดย
ผู้เชี่ยวชาญ 6 กลุ่ม (กำหนดเกณฑ์การแปลผลแบบสอบคู่ขนาน; ค่าเฉลี่ย RMSD ต่ำ
กว่า 0.50)..... 84



บทที่ 1

บทนำ

ความเป็นมาและความสำคัญของปัญหา

จากพระราชบัญญัติการศึกษาแห่งชาติ ในหมวด 4 มาตรา 26 ได้กำหนดให้สถานศึกษาจัดการประเมินผู้เรียนโดยพิจารณาจากพัฒนาการของผู้เรียน ความประพฤติ การสังเกตพฤติกรรม การเรียน การร่วมกิจกรรมและการทดสอบควบคู่ไปในกระบวนการเรียนการสอนตามความเหมาะสมของแต่ละระดับและรูปแบบการศึกษา แสดงให้เห็นถึงความสำคัญของการวัดและประเมินผลทั้งในระหว่างเรียน (Formative evaluation) รวมถึงการวัดและประเมินผลรวบยอด (Summative evaluation) ซึ่งมีการตรวจสอบความก้าวหน้าของผู้เรียนว่าบรรลุวัตถุประสงค์การเรียนรู้หรือไม่ เพื่อพิจารณาปรับปรุงข้อบกพร่องโดยการให้ข้อมูลย้อนกลับเพื่อส่งเสริมการเรียนรู้อย่างมีประสิทธิภาพ และสามารถพัฒนาการเรียนรู้ให้มีผลสัมฤทธิ์ทางการเรียนได้อย่างที่ตั้งเป้าหมายไว้ในแต่ละระดับชั้น (Hattie, 2009; โชติกา ภาชีผล, 2559) สอดคล้องกับจุดมุ่งหมายของการวัดและประเมินผลเพื่อเปรียบเทียบพัฒนาการของผู้เรียน ว่ามีความเจริญงอกงามเพียงใด โดยเปรียบเทียบเป็นระยะๆ คนละช่วงเวลา อาจจะเป็นการวัดและประเมินผลช่วงก่อนเรียน ระหว่างเรียน หรือหลังเรียน แต่ต้องคำนึงถึงเนื้อหาที่วัดเพื่อเปรียบเทียบพัฒนาการของแต่ละบุคคลควรเป็นเนื้อหาเดียวกัน (Gronlund, 2009; โชติกา ภาชีผล, 2559) ดังนั้น การวัดและประเมินผลในแต่ละครั้งจะต้องคำนึงถึงคุณภาพและความคงเส้นคงวาของแบบสอบเพื่อให้การนำผลการสอบไปใช้ได้อย่างถูกต้องและมีความเชื่อมั่นในข้อมูลเชิงประจักษ์ว่าผลการสอบถูกนำไปใช้เปรียบเทียบพัฒนาการการเรียนรู้ที่เกิดขึ้นผ่านการวัดและประเมินผลที่เท่าเทียมกัน หรือใช้แบบสอบเพื่อวัดและประเมินผลรวบยอดในกลุ่มผู้สอบหลายกลุ่มซึ่งมีเกณฑ์การตัดสินผลเหมือนกันจะต้องใช้แบบสอบที่มีความเท่าเทียมกันเพื่อให้การตัดสินผลมีความโปร่งใสและเป็นธรรม จึงมีการนำแบบสอบคู่ขนานมาใช้เป็นเครื่องมือสำหรับวัดและประเมินผลอย่างแพร่หลาย

เมื่อมีการนำแบบสอบคู่ขนานมาใช้เพื่อวัดและประเมินผลทางการศึกษาเป็นจำนวนมาก โดยส่วนใหญ่จะมุ่งเน้นให้แบบสอบทุกฉบับมีความเป็นคู่ขนานกันทั้งในด้านเนื้อหาและการตรวจสอบโดยอาศัยค่าทางสถิติ ดังนั้น การออกแบบการสร้างแบบสอบคู่ขนาน จะต้องมีการกำหนดรูปแบบของแบบสอบ สร้างแผนผังการสอบ (Test blueprint) และสร้างตารางวิเคราะห์ข้อสอบ (Table of specifications) ซึ่งเป็นขั้นตอนที่สำคัญที่ทำให้ได้ข้อสอบที่วัดได้ตรงตามจุดประสงค์และพฤติกรรมการเรียนรู้ และสามารถใช้เป็นแนวทางในการพัฒนาแบบสอบคู่ขนานได้ (โชติกา ภาชีผล, 2559) แบบสอบคู่ขนาน ได้ถูกนำมาใช้ในการสอบทั้งการสอบขนาดใหญ่ที่มีการสอบหลายครั้ง เช่น

การทดสอบวิชาความถนัดทั่วไป (General Aptitude Test) การทดสอบวิชาความถนัดทางวิชาการ และวิชาชีพ (Professional and Academic Aptitude Test) เช่น ใ้วิชาชีพแพทย์มีการใช้แบบสอบ คู่ขนานในการสอบเพื่อประเมินและรับรองความรู้ความสามารถในการประกอบวิชาชีพเวชกรรม 2 ขั้นตอน ตามเกณฑ์ที่ศูนย์ประเมินรับรองความรู้ความสามารถในการประกอบวิชาชีพเวชกรรม (ศ.ร.ว.) กำหนดไว้ คือ ขั้นตอนที่ 1 สอบประเมินความรู้ทางด้านวิทยาศาสตร์การแพทย์พื้นฐาน และ ขั้นตอนที่ 2 สอบประเมินความรู้ทางด้านวิทยาศาสตร์การแพทย์คลินิก เพื่อเป็นการรับรองมาตรฐาน ของแพทย์ที่จบการศึกษาแพทยศาสตรบัณฑิตจากทั้งในและต่างประเทศให้มีมาตรฐานเดียวกัน โดยมีการจัดสอบ 2 ครั้งต่อปี ในเดือนพฤศจิกายนและมิถุนายน จึงมีการใช้แบบสอบคู่ขนานโดยมีเป้าหมาย ให้แบบสอบทุกระดับมีความเป็นคู่ขนานกันทั้งในด้านเนื้อหาและการตรวจสอบโดยอาศัยค่าทางสถิติ (ประกาศแพทยสภาที่ 12/2555 เรื่อง เกณฑ์ความรู้ความสามารถในการประเมินเพื่อรับใบอนุญาต เป็นผู้ประกอบวิชาชีพเวชกรรม พ.ศ. 2555) สอดคล้องกับการใช้แบบสอบคู่ขนานในการสอบประเมิน เพื่อรับใบอนุญาตเป็นผู้ประกอบวิชาชีพเวชกรรมของประเทศแคนาดาซึ่งประกอบด้วยข้อสอบจำนวน 168 ข้อโดยมีการสร้างข้อสอบคู่ขนานจากคลังข้อสอบโดยอาศัยทฤษฎีการตอบสนองข้อสอบ แบบ 2 พารามิเตอร์ (Boughton & Gierl, 2000) นอกจากนี้ มีการใช้แบบสอบคู่ขนานเพื่อการประเมินผล ความก้าวหน้าและประเมินผลรวบยอดในรายวิชาต่างๆ รวมถึงแบบสอบก่อนและหลังเรียนในหลาย รายวิชา เช่น เกสซ์วิทยา (Carolina, 2014) ศัลยศาสตร์ (Ramos, 2013) และสรีรวิทยา (Hettiaratchi, 1978) เป็นต้น สำหรับวิทยาลัยแพทยศาสตรพระมงกุฎเกล้า มีการสอนรายวิชาเกสซ์ วิทยา 1 และ 2 สำหรับนักศึกษาแพทย์ชั้นปีที่ 3 โดยมีการใช้แบบสอบคู่ขนานสำหรับการ ประเมินผลความก้าวหน้าในแต่ละการสอบซึ่งจัดสอบจำนวน 10 ครั้ง โดยแบ่งตามรายวิชาเป็น 6 ครั้ง และ 4 ครั้งตามลำดับ ทั้งนี้จัดทำการสอบก่อนเรียนและหลังเรียนโดยใช้แบบสอบที่เป็นคู่ขนานกัน โดยอาศัยการตัดสินใจตรงตามเนื้อหาโดยอาศัยผู้เชี่ยวชาญจำนวน 3 คน โดยกำหนดให้ข้อสอบแต่ละฉบับมีจำนวน 20-40 ข้อ และสร้างข้อสอบโดยคัดเลือกข้อสอบจากคลังข้อสอบตั้งแต่ปีการศึกษา 2554-2558 และมีการใช้แบบสอบคู่ขนานในการจัดสอบประเมินความรู้ทางด้านวิทยาศาสตร์ การแพทย์พื้นฐาน เพื่อใช้ประเมินผลรวบยอดในรายวิชาพื้นฐานชั้นปรีคลินิก โดยจัดสอบปีละ 2 ครั้ง ในเดือนมิถุนายนและพฤศจิกายนของแต่ละปี นอกจากนี้ ในการเรียนชั้นคลินิกของนักศึกษาแพทย์ ได้มี การแบ่งกลุ่มเพื่อปฏิบัติงานในหอผู้ป่วยและเรียนในแต่ละสาขาวิชาเป็นกลุ่มย่อยวนไปตามภาควิชา ต่างๆ จึงมีการใช้แบบสอบคู่ขนานสำหรับการสอบในแต่ละกลุ่มเพื่อให้การวัดและประเมินผลเท่าเทียม กัน แต่พบว่า แบบสอบคู่ขนานที่สร้างขึ้น มีการตรวจสอบคุณภาพของแบบสอบและข้อสอบในด้าน ความตรงของแบบสอบและการวิเคราะห์ข้อสอบรายข้อและแบบสอบทั้งฉบับ เพื่อนำสารสนเทศไป ปรับปรุงพิจารณาข้อสอบสำหรับประยุกต์ใช้ในการสอบครั้งต่อไปเท่านั้น ยังขาดการตรวจสอบความ

เป็นคู่ขนานของข้อสอบและแบบสอบโดยอาศัยการทดสอบทางสถิติเพื่อยืนยันความเป็นคู่ขนานของแบบสอบที่สร้างขึ้น

เมื่อพิจารณาหลักการสร้างแบบสอบคู่ขนาน สามารถแบ่งตามการบริหารการสอบได้เป็น 2 รูปแบบ รูปแบบที่ 1 การสร้างแบบสอบคู่ขนานหลังการสอบ เป็นการสร้างข้อสอบจากคลังข้อสอบที่มีการเก็บรวบรวมข้อมูลค่าพารามิเตอร์หรือค่าสถิติของข้อสอบรายชื่อจากการสอบครั้งก่อนเพื่อใช้เป็นเงื่อนไขหรือสารสนเทศในการคัดเลือกข้อสอบเพื่อสร้างแบบสอบคู่ขนาน สามารถสร้างแบบสอบคู่ขนานโดยการใช้โปรแกรมคอมพิวเตอร์เพื่อสร้างชุดข้อสอบตามเงื่อนไขที่กำหนดขึ้นตามค่าสถิติที่ใช้ตรวจสอบความเป็นคู่ขนาน โดยกำหนดเนื้อหา จำนวนข้อสอบ และโครงสร้างในการวัดของข้อสอบตามผังข้อสอบ และทำการตรวจสอบความเป็นคู่ขนานที่สร้างขึ้นโดยอาศัยทฤษฎีการทดสอบแบบดั้งเดิมหรือทฤษฎีการทดสอบแนวใหม่ ข้อดีของแบบสอบคู่ขนานหลังการสอบ คือ มีการตรวจสอบความเป็นคู่ขนานโดยอาศัยค่าทางสถิติ ทำให้มีความน่าเชื่อถือและมีความเป็นคู่ขนานของข้อสอบและแบบสอบตามเกณฑ์ความเป็นคู่ขนานที่ชัดเจน สามารถสร้างชุดแบบสอบคู่ขนานได้จำนวนหลายชุดขึ้นกับขนาดของคลังข้อสอบ และไม่ขึ้นกับการตัดสินใจของผู้เชี่ยวชาญ แต่ข้อเสียของการสร้างแบบสอบคู่ขนานวิธีนี้ คือ ต้องมีคลังข้อสอบและการวิเคราะห์ข้อสอบทางสถิติเพื่อหาค่าพารามิเตอร์ โดยต้องมีจำนวนผู้สอบเพียงพอสำหรับตรวจสอบความเป็นคู่ขนานโดยเฉพาะการหาค่าพารามิเตอร์และการตรวจสอบความเป็นคู่ขนานของแบบสอบโดยอาศัยทฤษฎีการตอบสนองข้อสอบ นอกจากนี้การประยุกต์ใช้ในสภาพจริงค่อนข้างยากเนื่องจากต้องอาศัยโปรแกรมคอมพิวเตอร์ที่ต้องการความเชี่ยวชาญในการใช้งานโปรแกรมสร้างชุดข้อสอบจากคลังข้อสอบ จึงเหมาะสำหรับใช้ในการสอบขนาดใหญ่ที่ผู้สอบจำนวนมากและมีการใช้แบบสอบคู่ขนานหลายชุดหรือมีการสอบหลายครั้ง ซึ่งต้องอาศัยการพัฒนาโปรแกรมการสร้างชุดแบบสอบอัตโนมัติจากคลังข้อสอบและพัฒนาระบบฐานข้อมูลที่ต้องอาศัยความเชี่ยวชาญสูง เพื่อตรวจสอบคุณภาพของแบบสอบในคลังข้อสอบก่อนนำไปใช้ในการสอบ (Lin, 2008) รูปแบบที่ 2 การสร้างแบบสอบคู่ขนานก่อนการสอบ เป็นการสร้างแบบสอบคู่ขนานโดยผู้ออกข้อสอบสร้างข้อคำถามและตัวเลือกตามแผนผังการสอบและตารางวิเคราะห์ข้อสอบ ที่กำหนดจุดประสงค์การเรียนรู้และโครงสร้างการวัดเหมือนกัน จากนั้นจึงตรวจสอบความเป็นคู่ขนานโดยอาศัยผู้เชี่ยวชาญ ทั้งนี้สามารถตรวจสอบความเป็นคู่ขนานของแบบสอบรายชื่อโดยประยุกต์ใช้การตรวจสอบความตรงเชิงเนื้อหา ซึ่งเป็นวิธีที่ให้ผู้เชี่ยวชาญตัดสินใจข้อคำถามที่สร้างขึ้นว่าเป็นไปตามเนื้อหาในแต่ละวัตถุประสงค์ที่ต้องการวัดหรือไม่ มีความสอดคล้องของข้อคำถามกับเนื้อหาและพฤติกรรมการเรียนรู้ และพิจารณาถึงความเป็นคู่ขนานของแบบสอบเป็นรายข้อคำถาม โดยการระบุน้ำหนักคะแนนและน้ำหนักที่ได้แต่ละข้อมาคำนวณความสอดคล้องระหว่างข้อคำถามกับวัตถุประสงค์ (Objective Congruency Index: IOC) โดยมีเกณฑ์สำหรับพิจารณาค่าความสอดคล้องควรจะมีค่าตั้งแต่ 0.50 ขึ้นไป ข้อดีของการสร้างแบบสอบคู่ขนานก่อนการสอบและการตรวจสอบโดย

ผู้เชี่ยวชาญ คือ สามารถประยุกต์ใช้ในการสร้างข้อสอบในสภาพจริงได้ง่ายเนื่องจากใช้ผู้เชี่ยวชาญที่มีอยู่ในสาขาที่เชี่ยวชาญนั้นๆ และสามารถสร้างข้อสอบใหม่ได้ตามโครงสร้างและจุดประสงค์ที่กำหนด โดยไม่ต้องอาศัยค่าสถิติหรือพารามิเตอร์ของข้อสอบเดิมจากคลังข้อสอบ แต่ข้อเสียของวิธีนี้ คือ การพิจารณาคุณภาพของแบบสอบโดยผู้เชี่ยวชาญ ขึ้นอยู่กับระดับความสามารถในการตัดสินใจรายบุคคลของผู้เชี่ยวชาญที่แตกต่างกัน (โชติกา ภาชีผล, 2559) ดังนั้น จะต้องมีการตรวจสอบความเป็นคู่ขนานของแบบสอบและข้อสอบโดยอาศัยการทดสอบทางสถิติเพื่อให้ได้หลักฐานเชิงประจักษ์สำหรับผู้สร้างข้อสอบและผู้มีส่วนได้ส่วนเสียในการวัดและประเมินผลโดยใช้แบบสอบคู่ขนานนั้นๆ ไปใช้ตัดสินผลและปรับปรุงพัฒนาแบบสอบและข้อสอบคู่ขนานให้เป็นไปตามมาตรฐาน โปร่งใส และเป็นธรรม

เมื่อสร้างแบบสอบคู่ขนาน โดยมีการกำหนดรูปแบบของแบบสอบ สร้างแผนผังการสอบและตารางวิเคราะห์ข้อสอบแล้ว สามารถตรวจสอบความเป็นคู่ขนานเชิงเนื้อหาก่อนการสอบโดยผู้เชี่ยวชาญพิจารณาเป็นรายข้อคำถาม และพิจารณาค่าความสอดคล้องระหว่างข้อคำถามกับวัตถุประสงค์และทำการตรวจสอบความเป็นคู่ขนานของแบบสอบหลังการสอบโดยอาศัยการทดสอบทางสถิติต่างๆ พบว่า การตรวจสอบความเป็นคู่ขนานของแบบสอบโดยอาศัยข้อตกลงเบื้องต้นของทฤษฎีการทดสอบแบบดั้งเดิม กำหนดให้แบบสอบคู่ขนาน 2 ฉบับจะต้องมีคะแนนจริงของผู้สอบแต่ละคนและค่าความคลาดเคลื่อนมาตรฐานของประชากรที่สอบทั้ง 2 ฉบับเท่ากัน (Lord, 1980; Lord & Novick, 1968) โดยคะแนนจริง มีค่าเท่ากับคะแนนเฉลี่ยของคะแนนที่ได้จากการวัดซ้ำหลายๆ ครั้ง ซึ่งเป็นการพิจารณาค่าสถิติหลังจากการสอบ (ค่าเฉลี่ยรวม ค่าความคลาดเคลื่อนมาตรฐานและค่าความเที่ยงของแบบสอบ) หากจะนำแบบสอบคู่ขนานนี้ไปใช้ในการสอบครั้งต่อไป ต้องใช้แบบสอบซ้ำและความเที่ยงของแบบสอบคู่ขนานอาจเปลี่ยนไปตามประชากรที่แตกต่างกัน ซึ่งเป็นจุดอ่อนของการตรวจสอบความเป็นคู่ขนานของแบบสอบโดยอาศัยทฤษฎีการทดสอบแบบดั้งเดิม ซึ่งขาดสารสนเทศในการพิจารณาความเป็นคู่ขนานของข้อสอบรายข้อด้วย (Dorans, Pommerich, & Holland, 2007) นอกจากนี้ การตรวจสอบความเป็นคู่ขนานของแบบสอบทั้งฉบับโดยอาศัยทฤษฎีการทดสอบแบบดั้งเดิมที่มีหลักการพิจารณาจากค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของข้อสอบคู่ขนานแต่ละคู่มิค่าใกล้เคียงกัน แต่จะพบว่า ค่าเฉลี่ยรวมของแบบสอบคู่ขนานทั้งฉบับอาจจะแตกต่างกันอย่างมีนัยสำคัญ ซึ่งอาจเกิดจากข้อจำกัดของคะแนนความคลาดเคลื่อนและค่าพารามิเตอร์ของข้อสอบ ทั้งนี้ แบบสอบคู่ขนานที่ใช้สอบจะมีลักษณะเฉพาะขึ้นอยู่กับกลุ่มผู้ตอบและเปลี่ยนไปตามการสอบแต่ละครั้งจึงเป็นข้อจำกัดของการตรวจสอบความเป็นคู่ขนานของแบบสอบโดยอาศัยทฤษฎีการทดสอบแบบดั้งเดิม (ศิริชัย กาญจนวาสี, 2555)

ต่อมาจึงมีการนำทฤษฎีการตอบสนองข้อสอบมาประยุกต์ใช้ในการสร้างแบบสอบคู่ขนาน และตรวจสอบความเป็นคู่ขนานของแบบสอบนั้นๆ โดยอาศัยค่าพารามิเตอร์ของข้อสอบที่ได้จาก

การสอบ ซึ่งประกอบด้วย ค่าพารามิเตอร์ความยากของข้อสอบ (b) ค่าพารามิเตอร์อำนาจจำแนกของข้อสอบ (a) และค่าพารามิเตอร์โอกาสในการเดาข้อสอบได้ถูก (c) โดยสามารถจัดชุดแบบสอบคู่ขนาน โดยกำหนดเงื่อนไขของพารามิเตอร์ข้อสอบที่แตกต่างกัน (Ali & Rijn, 2016; Lin, 2008) และตรวจสอบความเป็นคู่ขนานของแบบสอบ โดยอาศัยโค้งลักษณะข้อสอบ (Item characteristic curve; ICC) พบว่า แบบสอบคู่ขนาน 2 ฉบับจะต้องมีความน่าจะเป็นที่ผู้ตอบข้อสอบถูก ($P_i(\theta)$) เท่ากันในทุกความสามารถที่มีอยู่ในผู้ตอบ (θ) และมีรูปแบบการกระจายไม่แตกต่างกัน (Linden & Luecht, 1998; Spray & Miller, 1992) นอกจากนี้ มีการใช้ฟังก์ชันสารสนเทศของแบบสอบ (Test information; TIF) โดยพบว่า แบบสอบคู่ขนานจะต้องมีฟังก์ชันสารสนเทศของแบบสอบเหมือนกัน (Linden & Adema, 1998; Veldkamp, Matteucci, & De Jong, 2013) สอดคล้องกับการศึกษาของ Ali and Rijn (2016) พบว่า การวิเคราะห์หัดดัชนีความเป็นคู่ขนานของแบบสอบโดยอาศัยทฤษฎีการตอบสนองข้อสอบ จะพิจารณาจากฟังก์ชันสารสนเทศของแบบสอบเป็นหลัก เพราะให้สารสนเทศที่ถูกต้องและน่าเชื่อถือมากกว่าการใช้ ICC หรือ TCC เนื่องจากอาจพบว่ามี ความแตกต่างของค่าพารามิเตอร์ที่ส่งผลต่อคะแนนความสามารถของผู้สอบให้มีความคลาดเคลื่อนตามระบบได้ (Systematic error) ซึ่งจำเป็นต้องแก้ไขโดยการปรับเทียบคะแนนของผู้สอบ (Equated score) จึงทำให้การตรวจสอบความเป็นคู่ขนานของแบบสอบโดยอาศัยโค้งลักษณะข้อสอบ (TCC) มีข้อจำกัดหากมีความแตกต่างของคะแนนระหว่างผู้สอบเกิดขึ้น จึงนิยมใช้ค่าสารสนเทศของการสอบ (TIF) ในการตรวจสอบความเป็นคู่ขนานของการสอบมากกว่า (Debeer, Ali, & Rijn, 2017) จากการศึกษาเปรียบเทียบความเป็นคู่ขนานของแบบสอบที่ได้จากการจัดชุดแบบสอบอัตโนมัติโดยอาศัยค่าพารามิเตอร์ พบว่า เมื่อตรวจสอบโดยอาศัยทฤษฎีการทดสอบแบบดั้งเดิมจะเน้นคุณภาพของแบบสอบทั้งฉบับมากกว่าคุณภาพข้อสอบรายข้อ และทฤษฎีการตอบสนองข้อสอบสามารถพิจารณาความเป็นคู่ขนานโดยอาศัยค่าสารสนเทศของข้อสอบและแบบสอบมาประยุกต์ใช้ได้ง่ายกว่าเนื่องจากสามารถตรวจสอบได้ทั้งความเป็นคู่ขนานของข้อสอบรายข้อและความเป็นคู่ขนานของข้อสอบทั้งฉบับ (Ali & Rijn, 2016; Boughton & Gierl, 2000; Sanders & Verschoor, 1998) โดยพบว่าค่าพารามิเตอร์ที่ได้จากการวิเคราะห์ข้อสอบมีลักษณะความไม่แปรเปลี่ยน จึงสามารถสร้างแบบสอบได้อย่างมีประสิทธิภาพ นอกจากนี้ ค่าสารสนเทศที่ได้จากข้อสอบ เมื่อนำมารวมกันเป็นสารสนเทศของแบบสอบจะสะท้อนความถูกต้องแม่นยำของการประมาณค่าความสามารถของผู้สอบ จึงทำให้ได้หลักประกันสำหรับการพัฒนาแบบสอบคู่ขนานให้ตรงกับเป้าหมายการนำไปใช้ได้อย่างมีประสิทธิภาพ (โชติกา ภาชีผล, 2559; ศิริชัย กาญจนวาสี, 2555)

แม้ว่าการสร้างแบบสอบคู่ขนานโดยอาศัยค่าพารามิเตอร์ความยากและอำนาจจำแนกตามทฤษฎีการตอบสนองข้อสอบ จะมีข้อดีคือ สามารถเลือกข้อสอบให้เหมาะสมกับระดับความสามารถ

ของผู้สอบโดยนำค่าพารามิเตอร์ไปใช้สร้างแบบสอบคู่ขนานตามเงื่อนไขที่กำหนดขึ้นโดยอาศัยโปรแกรมคอมพิวเตอร์ (Automated test assembly) (Ali & Rijn, 2016; Lin, 2008) แต่ค่าพารามิเตอร์และสารสนเทศของข้อสอบจะคำนวณได้หลังจากการสอบเท่านั้น จึงเป็นข้อด้อยที่สำคัญ คือ การเปิดเผยและรับรู้ข้อสอบของผู้สอบ ซึ่งในสภาพจริงของการสอบที่มีความสำคัญสูงจึงนำไปประยุกต์ใช้ค่อนข้างยาก จึงจำเป็นต้องมีคลังข้อสอบเป็นจำนวนมากพอที่จะลดความซ้ำซ้อนของข้อสอบที่เลือกใช้ระหว่างการสอบแต่ละครั้ง รวมถึงปัญหาด้านความตรงตามเนื้อหาของข้อสอบที่ลดลงเนื่องจากต้องอาศัยการตัดสินใจที่มีความตรงที่มีรายละเอียดแต่ละข้อที่แตกต่างกันและต้องใช้ความเชี่ยวชาญสูง จากการศึกษาของ Wanichsan, Panjaburee, Laosinchai, Triampo, and Chookaew (2012) และ Panjaburee, Hwang, Triampo, and Shih (2010) พบว่าการสร้างแบบสอบโดยกลุ่มผู้เชี่ยวชาญ สามารถเพิ่มระดับความสัมพันธ์ระหว่างเนื้อหาหรือจุดประสงค์การเรียนรู้กับแบบสอบได้ดีขึ้น ซึ่งลดข้อด้อยของการสร้างแบบสอบโดยใช้โปรแกรมคอมพิวเตอร์ เนื่องจากมีความซับซ้อนของรายละเอียดในแต่ละเนื้อหา ทำให้การประมาณค่าความสามารถของผู้สอบมีความตรงมากขึ้น และสามารถให้ข้อมูลย้อนกลับได้อย่างมีประสิทธิภาพสำหรับข้อสอบวินิจฉัยที่ต้องการความสัมพันธ์ระหว่างข้อสอบกับเนื้อหาค่อนข้างสูง สอดคล้องกับสภาพจริงที่การคัดเลือกข้อสอบจากคลังข้อสอบเพื่อสร้างเป็นแบบสอบคู่ขนาน จะใช้กลุ่มที่ประกอบด้วยผู้เชี่ยวชาญอย่างน้อย 3 คน โดยพิจารณาจากเกณฑ์มาตรฐาน จุดประสงค์การเรียนรู้ และเนื้อหาที่กำหนดตามแผนผังการสอบและตารางวิเคราะห์ข้อสอบเพื่อให้แบบสอบที่สร้างขึ้นมีความเที่ยงและความตรงสูง แต่พบว่าปัญหาของการตัดสินใจโดยอาศัยผู้เชี่ยวชาญ คือ ระดับความสามารถในการตัดสินใจของผู้เชี่ยวชาญแต่ละคนมีความแตกต่างกัน จึงส่งผลให้ความตรงและความถูกต้องของผลการตัดสินใจมีค่าที่เบี่ยงเบนไปจากค่าที่แท้จริง ทั้งนี้ ขึ้นอยู่กับประสบการณ์ในการสอนรายวิชานั้นๆ รวมถึงประสบการณ์และความเชี่ยวชาญในการออกข้อสอบ คัดเลือกข้อสอบ รวมถึงการตัดสินใจความตรงตามเนื้อหาของข้อสอบตามจุดประสงค์การเรียนรู้รายวิชานั้นๆ ของแต่ละผู้เชี่ยวชาญด้วย (Romney, H., & C., 1986; Weller, 2007) จึงเป็นที่มาของการนำทฤษฎีฉันทามติทางวัฒนธรรม (Cultural consensus theory) มาประยุกต์ใช้เพื่อตรวจสอบความสามารถในการตัดสินใจเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญสำหรับการศึกษาในครั้งนี้

ฉันทามติ เป็นกระบวนการตัดสินใจที่กลุ่มคนต้องการให้เกิดการมีส่วนร่วมและความเห็นพ้องร่วมกันมากที่สุด การสรุปผลการตัดสินใจขึ้นกับข้อกำหนดที่ตกลงกันได้ ตามน้ำหนักและขอบเขตของการเห็นพ้องร่วมกัน โดยมีการประยุกต์ใช้ทฤษฎีฉันทามติทางวัฒนธรรม สำหรับวิเคราะห์ผลจากการตัดสินใจหรือการลงฉันทามติ เพื่อประมาณค่าระดับความสามารถในการเห็นพ้องร่วมกันของแต่ละบุคคลว่ามีความสามารถในการตัดสินใจในเรื่องนั้นๆ มากน้อยเพียงใด โดยการหาความสอดคล้องของคำตอบระหว่างผู้เชี่ยวชาญในกลุ่มของแต่ละข้อคำถาม ซึ่งเป็นประมาณค่าระดับความสามารถใน

การเห็นพ้องร่วมกันว่าคำตอบของผู้เชี่ยวชาญมีความถูกต้องมากน้อยเพียงใด และสามารถตรวจสอบคำตอบที่ถูกต้องของข้อคำถามนั้นๆ โดยอาศัยทฤษฎีฉันทามติทางวัฒนธรรมแบบมีรูปแบบ ทำให้เพิ่มความเชื่อมั่นของหรือการลงฉันทามติโดยผู้เชี่ยวชาญมีความน่าเชื่อถือมากขึ้นและสามารถถ่วงน้ำหนักระดับความสามารถในการเห็นพ้องร่วมกันของผู้เชี่ยวชาญแต่ละคนในกลุ่มได้ ดังนั้น ผู้วิจัยจึงประยุกต์ใช้ทฤษฎีฉันทามติทางวัฒนธรรมแบบมีรูปแบบเพื่อศึกษาผลของระดับความสามารถในการตัดสินใจเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญ เพื่อลดจุดอ่อนของการสร้างแบบสอบคู่ขนานก่อนการสอบโดยกลุ่มผู้เชี่ยวชาญ ในด้านความเป็นปัจเจกบุคคลที่ผู้เชี่ยวชาญแต่ละคนมีระดับความสามารถในการตัดสินใจเป็นคู่ขนานของข้อสอบที่แตกต่างกัน ทั้งนี้ โมเดลฉันทามติทางวัฒนธรรมแบ่งตามรูปแบบการประมาณค่าความสามารถในการเห็นพ้องร่วมกันได้ 2 โมเดล (Romney et al., 1986) ได้แก่ 1. โมเดลฉันทามติทางวัฒนธรรมแบบไม่มีรูปแบบ (Informal cultural consensus model) เป็นวิธีการวิเคราะห์เพื่อระบุระดับความสอดคล้องของการตอบข้อคำถามของผู้เชี่ยวชาญแต่ละคนกับฉันทามติของกลุ่ม แต่จะไม่เน้นการหาคำตอบที่ถูกต้องของข้อคำถาม แต่จะเน้นความสอดคล้องของฉันทามติของบุคคลต่อฉันทามติของกลุ่ม เพราะมีแนวคิดที่ว่าฉันทามติเกิดขึ้นมาจากข้อคำถามหรือเนื้อหาที่ยังไม่ทราบคำตอบที่แน่ชัดและคำตอบที่แท้จริงของข้อคำถามทางสังคมวิทยามักจะแตกต่างกันตามความหลากหลายทางวัฒนธรรม ดังนั้น จึงเป็นโมเดลที่เน้นการหาค่าถ่วงน้ำหนักฉันทามติของผู้เชี่ยวชาญต่อฉันทามติของกลุ่มว่าควรถ่วงน้ำหนักแต่ละมากน้อยเพียงใดตามความสามารถในการเห็นพ้องร่วมกัน 2. โมเดลฉันทามติทางวัฒนธรรมแบบมีรูปแบบ (Formal cultural consensus model) เป็นวิธีการประมาณค่าความสามารถในการเห็นพ้องร่วมกันของแต่ละผู้เชี่ยวชาญ โดยได้จากความเห็นพ้องร่วมกันของคำตอบแต่ละข้อคำถามจากฉันทามติของผู้เชี่ยวชาญ โดยอาศัยเมทริกซ์ความสัมพันธ์หรือเมทริกซ์ของความแปรปรวนร่วมของฉันทามติของผู้เชี่ยวชาญทุกคู่ในกลุ่มผู้เชี่ยวชาญทั้งหมดที่ลงฉันทามติหรือตอบข้อคำถามนั้นๆ และสามารถคำนวณความน่าจะเป็นของความถูกต้องสำหรับคำตอบที่ได้จากฉันทามติของผู้เชี่ยวชาญโดยอาศัยทฤษฎีสถิติของเบย์มาใช้ในการคำนวณ ภายใต้ข้อตกลงเบื้องต้นอย่างเคร่งครัดตามทฤษฎีฉันทามติทางวัฒนธรรมแบบมีรูปแบบ คือ ผู้เชี่ยวชาญแต่ละคนต้องเป็นอิสระต่อกัน ไม่มีการปรึกษากันและต้องไม่ใช้การสัมภาษณ์กลุ่มย่อย ข้อคำถามที่ใช้เป็นเอกมิตี โดยต้องมีความเท่าเทียมกันและเป็นข้อคำถามที่คำตอบมีความสอดคล้องกันสูง จึงจะสามารถนำมาใช้ตรวจสอบความสามารถในการเห็นพ้องร่วมกันได้

ในการสร้างแบบสอบคู่ขนานเพื่อวัดและประเมินผลในรายวิชาเภสัชวิทยาสำหรับนักศึกษาแพทย์ชั้นปีที่ 3 มีการออกแบบการสร้างแบบสอบโดยใช้รูปแบบหลักการตัดสินใจแบบอิงเกณฑ์สอดคล้องตามเกณฑ์การสอบเพื่อประเมินและรับรองความรู้ความสามารถในการประกอบวิชาชีพเวชกรรม โดยมีลักษณะข้อสอบเป็นแบบเลือกตอบชนิดหลายตัวเลือกตามแผนผังการสอบและตาราง

วิเคราะห์ข้อสอบ โดยกำหนดให้ผู้เชี่ยวชาญคัดเลือกข้อสอบจากคลังข้อสอบตามจุดประสงค์ที่กำหนดให้ และนำแบบสอบไปใช้ จากนั้นจึงวิเคราะห์คุณภาพของข้อสอบรายข้อและแบบสอบทั้งฉบับ และสร้างคลังข้อสอบสำหรับสร้างแบบสอบคู่ขนานเพื่อใช้ในการสอบประเมินผลในรายวิชาเภสัชวิทยาต่อไป พบว่าผู้เชี่ยวชาญที่คัดเลือกข้อสอบนั้นมีคุณลักษณะที่หลากหลายแตกต่างกัน ทั้ง รัดับความสามารถในการคัดเลือกข้อสอบเพื่อสร้างแบบสอบคู่ขนาน รวมถึงประสบการณ์ในการสอนนักศึกษาแพทย์ การสร้างและตัดสินความตรงตามเนื้อหาของข้อสอบ ฯลฯ ซึ่งเป็นตัวแปรที่ส่งผลต่อความเป็นคู่ขนานของแบบสอบ ทั้งนี้ ผู้ออกข้อสอบเป็นผู้นำหน้าทีสร้างข้อสอบตามมาตรฐานที่กำหนด รวมทั้งสร้างเกณฑ์การให้คะแนนตามชนิดของข้อสอบที่ออก ผู้ออกข้อสอบต้องมีคุณสมบัติที่สำคัญดังนี้ มีความรู้ความเข้าใจในเนื้อหาสาระที่จะออกข้อสอบเป็นอย่างดี มีประสบการณ์ในการสอนในเนื้อหาสาระดังกล่าวเป็นอย่างดี มีความสามารถในการทำความเข้าใจกับกรอบมาตรฐานหลักสูตรและกรอบมาตรฐานการออกข้อสอบเป็นอย่างดี มีประสบการณ์ในการสร้างเครื่องมือวัดผลตามประเภทที่กำหนดไว้ในมาตรฐานการสร้างข้อสอบ เป็นผู้ใจกว้างยอมรับฟังความคิดเห็นผู้อื่นต่อข้อสอบที่ตนเองได้สร้างไว้ สามารถให้ข้อโต้แย้งคำแนะนำได้อย่างเหมาะสม และสามารถปรับแก้ข้อสอบได้ตามที่ผู้พิจารณาข้อสอบให้คำแนะนำ (สถาบันทดสอบทางการศึกษาแห่งชาติ, 2556) ผู้วิจัยจึงมีความสนใจศึกษาเปรียบเทียบดัชนีความเป็นคู่ขนานของแบบสอบคู่ขนานที่ระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญและประสบการณ์สอนที่แตกต่างกัน เพื่อให้ได้แบบสอบคู่ขนานที่มีคุณภาพสูงและนำไปใช้ได้ในสภาพจริง การวิจัยครั้งนี้จึงใช้วิธีการสร้างแบบสอบคู่ขนาน โดยอาศัยผู้เชี่ยวชาญในการคัดเลือกข้อสอบที่มีตัวแปรอิสระ คือระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญเพื่อตรวจสอบระดับความสามารถของผู้เชี่ยวชาญ มุ่งเน้นลดจุดอ่อนในเรื่องความเป็นปัจเจกบุคคลและระดับความถูกต้องในการตัดสินใจที่แตกต่างกันของผู้เชี่ยวชาญแต่ละบุคคล และประสบการณ์สอนเพื่อเป็นสารสนเทศในการตัดสินใจคัดเลือกผู้เชี่ยวชาญเพื่อสร้างแบบสอบคู่ขนาน ซึ่งการใช้การคัดเลือกข้อสอบคู่ขนานโดยกลุ่มผู้เชี่ยวชาญสามารถนำไปประยุกต์ใช้ในการสร้างแบบสอบคู่ขนานในสภาพจริงได้ สามารถตรวจสอบความเป็นคู่ขนานของแบบสอบเมื่อนำไปใช้ในการสอบโดยอาศัยการทดสอบทางสถิติและนำข้อสอบไปใช้ในการสร้างคลังข้อสอบสำหรับคัดเลือกข้อสอบเพื่อสร้างแบบสอบคู่ขนานที่มีคุณภาพสูงในครั้งต่อไป

คำถามวิจัย

1. ผู้เชี่ยวชาญมีระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบเป็นอย่างไร
2. แบบสอบคู่ขนานที่ได้จากการคัดเลือกโดยกลุ่มผู้เชี่ยวชาญที่มีระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบและประสบการณ์สอนที่แตกต่างกันจะมีดัชนีความเป็นคู่ขนานของแบบสอบแตกต่างกันหรือไม่ อย่างไร
3. ข้อสอบคู่ขนานที่ได้จากการคัดเลือกโดยกลุ่มผู้เชี่ยวชาญที่มีระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบและประสบการณ์สอนที่แตกต่างกันจะมีดัชนีความเป็นคู่ขนานของข้อสอบแตกต่างกันหรือไม่ อย่างไร

วัตถุประสงค์

1. เพื่อศึกษาระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญ
2. เพื่อเปรียบเทียบดัชนีความเป็นคู่ขนานของแบบสอบคู่ขนานที่ได้จากการคัดเลือกโดยกลุ่มผู้เชี่ยวชาญที่มีระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบและประสบการณ์สอนที่แตกต่างกัน
3. เพื่อเปรียบเทียบดัชนีความเป็นคู่ขนานของข้อสอบคู่ขนานที่ได้จากการคัดเลือกโดยกลุ่มผู้เชี่ยวชาญที่มีระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบและประสบการณ์สอนที่แตกต่างกัน

สมมติฐานการวิจัย

การสร้างแบบสอบคู่ขนาน โดยกำหนดรูปแบบของแบบสอบ สร้างแผนผังการสอบและสร้างตารางวิเคราะห์ข้อสอบ เป็นขั้นตอนที่สำคัญที่ทำให้ได้ข้อสอบที่วัดได้ตรงตามจุดประสงค์และพฤติกรรมการเรียนรู้ (โชติกา ภาชีผล, 2559) และจากการวิจัยของ Wanichsan et al. (2012) และ Panjaburee et al. (2010) พบว่าการสร้างแบบสอบโดยกลุ่มผู้เชี่ยวชาญ สามารถเพิ่มระดับความสัมพันธ์ระหว่างเนื้อหาหรือจุดประสงค์การเรียนรู้กับแบบสอบได้ดีขึ้น แต่พบว่าคุณลักษณะที่แตกต่างกันของผู้เชี่ยวชาญ เช่น ประสบการณ์ในการสอนนักศึกษาแพทย์ การสร้างและตัดสินความตรงตามเนื้อหาของข้อสอบ ส่งผลให้ระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญแต่ละคนมีความแตกต่างกัน จึงมีการนำทฤษฎีฉันทามติทางวัฒนธรรมแบบมีรูปแบบมาประยุกต์ใช้เพื่อให้การตัดสินใจมีความตรงและถูกต้องมากขึ้น (Romney, Batchelder, & Weller, 1987; Weller, 2007) ส่วนดัชนีความเป็นคู่ขนานที่สร้างขึ้นสามารถใช้ทฤษฎีการตอบสนองข้อสอบโดยพิจารณาจากค่าสารสนเทศของข้อสอบและแบบสอบซึ่งมีการประยุกต์ใช้ได้สอดคล้องกับสภาพ

จริงมากกว่าทฤษฎีการทดสอบแบบดั้งเดิม (Ali & Rijn, 2016; Boughton & Gierl, 2000; Gell, 1998) ผู้วิจัยจึงตั้งสมมติฐานการวิจัยในครั้งนี้คือ แบบสอบคู่ขนานที่ได้จากการคัดเลือกโดยกลุ่มผู้เชี่ยวชาญที่มีระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบและประสบการณ์สอนนักศึกษาแพทย์สูงน่าจะคัดเลือกข้อสอบเพื่อสร้างแบบสอบคู่ขนานที่มีดัชนีความเป็นคู่ขนานสูงกว่ากลุ่มผู้เชี่ยวชาญที่มีระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบในระดับต่ำกว่าและประสบการณ์สอนนักศึกษาแพทย์น้อยกว่า

ขอบเขตการวิจัย

ตัวแปรที่ศึกษา ประกอบด้วย ตัวแปรอิสระ 2 ตัวแปรได้แก่ 1) ระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญ แบ่งเป็น 3 ระดับ คือ ระดับที่ 1 (สูง), ระดับที่ 2 (ปานกลาง) และ ระดับที่ 3 (ต่ำ) และ 2) ประสบการณ์สอนนักศึกษาแพทย์ ได้แก่ ประสบการณ์สอนตั้งแต่ 10 ปีขึ้นไปและประสบการณ์สอนน้อยกว่า 10 ปี ตัวแปรตาม 2 ตัวแปร คือ 1) ดัชนีความเป็นคู่ขนานของแบบสอบทั้งฉบับ พิจารณาจากรากที่สองของส่วนเบี่ยงเบนค่าเฉลี่ยกำลังสองของสารสนเทศของแบบสอบ ($RMSD_{TF}$) และค่าเฉลี่ยสัดส่วนที่แตกต่างกันของสารสนเทศของแบบสอบ (MRD) และ 2) ความเป็นคู่ขนานของข้อสอบรายข้อ พิจารณาจากรากที่สองของส่วนเบี่ยงเบนค่าเฉลี่ยกำลังสองของสารสนเทศของข้อสอบ ($RMSD_{IF}$)

เนื้อหาที่ใช้ในการคัดเลือกเพื่อสร้างแบบสอบคู่ขนานตามเกณฑ์ความรู้ความสามารถในการประเมินเพื่อรับใบอนุญาตเป็นผู้ประกอบวิชาชีพเวชกรรม พ.ศ. 2555 ประกอบด้วย 5 หมวด ได้แก่ หมวด A = ยาลดความดันโลหิต หมวด B = ยารักษาโรคหัวใจ หมวด C = ยารักษาแผลในกระเพาะอาหาร หมวด D = ยาที่เกี่ยวข้องกับการบีบตัวของลำไส้ และ หมวด E = ยารักษาโรคเบาหวาน โดยมีข้อสอบสำหรับคัดเลือกทั้งหมด 200 ข้อ จำแนกเป็นหมวดละ 40 ข้อ เพื่อให้ผู้เชี่ยวชาญคัดเลือกข้อสอบเพื่อสร้างแบบสอบคู่ขนาน หมวดละ 5 ข้อ รวมทั้งสิ้นแบบสอบละ 25 ข้อ

นิยามศัพท์เฉพาะ

แบบสอบคู่ขนาน หมายถึง แบบสอบที่มีโครงสร้างทางเนื้อหา โคลงลักษณะของข้อสอบ และฟังก์ชันสารสนเทศของแบบสอบเหมือนกัน มีความน่าจะเป็นที่ผู้ตอบข้อสอบถูก ($P_i(\theta)$) เท่ากันในทุกความสามารถที่มีอยู่ในผู้ตอบ (θ) และรูปแบบการกระจายไม่แตกต่างกัน ตามโมเดลทฤษฎีการตอบสนองข้อสอบแบบ 2 พารามิเตอร์

ดัชนีความเป็นคู่ขนานของแบบสอบโดยอาศัยค่า $RMSD_{TF}$ หมายถึง ค่าสถิติที่บ่งชี้ว่าแบบสอบคู่ขนาน 2 ฉบับมีส่วนเบี่ยงเบนค่าเฉลี่ยกำลังสองของสารสนเทศของแบบสอบ ($RMSD_{TF}$) สำหรับ

ตรวจสอบการรูปแบบและการกระจายของฟังก์ชันสารสนเทศของแบบสอบ 2 ฉบับที่วัดโครงสร้างทางเนื้อหาเดียวกันในระดับความสามารถของผู้ตอบข้อสอบทุกระดับแตกต่างกันไม่เกิน 0.50

ดัชนีความเป็นคู่ขนานของแบบสอบโดยอาศัยค่า MRD หมายถึง ค่าสถิติที่บ่งชี้ว่าแบบสอบคู่ขนาน 2 ฉบับมีค่าเฉลี่ยสัดส่วนที่แตกต่างกันของสารสนเทศของแบบสอบ (MRD) สำหรับตรวจสอบการรูปแบบและการกระจายของฟังก์ชันสารสนเทศของแบบสอบ 2 ฉบับที่วัดโครงสร้างทางเนื้อหาเดียวกันในระดับความสามารถของผู้ตอบข้อสอบทุกระดับแตกต่างกันไม่เกิน 0.10

ดัชนีความเป็นคู่ขนานของข้อสอบรายข้อโดยอาศัยค่า $RMSD_{IF}$ หมายถึง ค่าสถิติที่บ่งชี้ว่าข้อสอบคู่ขนาน 2 ข้อมีส่วนเบี่ยงเบนค่าเฉลี่ยกำลังสองของสารสนเทศของข้อสอบ ($RMSD_{IF}$) สำหรับตรวจสอบการรูปแบบและการกระจายของฟังก์ชันสารสนเทศของข้อสอบ 2 ข้อที่วัดโครงสร้างทางเนื้อหาเดียวกันในระดับความสามารถของผู้ตอบข้อสอบทุกระดับแตกต่างกันไม่เกิน 0.50

ผู้เชี่ยวชาญ หมายถึง ผู้สอนนักศึกษาแพทยวิทยาลัยแพทยศาสตร์พระมงกุฎเกล้าที่มีคุณวุฒิการศึกษาสาขาเภสัชศาสตร์บัณฑิต แพทยศาสตร์บัณฑิต หรือสาขาวิชาชีพทางสาธารณสุขอื่นๆ

ระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญ หมายถึง ค่าความสามารถของผู้เชี่ยวชาญในการตัดสินความเป็นคู่ขนานของข้อสอบที่ได้จากการวิเคราะห์ฉันทามติของผู้เชี่ยวชาญทั้งหมด โดยอาศัยทฤษฎีฉันทามติทางวัฒนธรรมแบบมีรูปแบบ แบ่งออกเป็น 3 ระดับ ได้แก่

ระดับที่ 1 หมายถึง ผู้เชี่ยวชาญที่มีระดับความสามารถสูงในการคัดเลือกข้อสอบคู่ขนาน ตั้งแต่ 0.80 ถึง 1.00

ระดับที่ 2 หมายถึง ผู้เชี่ยวชาญที่มีระดับความสามารถปานกลางในการคัดเลือกข้อสอบคู่ขนานตั้งแต่ 0.60 ถึง 0.79

ระดับที่ 3 หมายถึง ผู้เชี่ยวชาญที่มีระดับความสามารถต่ำในการคัดเลือกข้อสอบคู่ขนานตั้งแต่ 0.40 ถึง 0.59

ประสบการณ์สอนนักศึกษาแพทย์ หมายถึง จำนวนปีที่ผู้เชี่ยวชาญผ่านการสอนนักศึกษาแพทย์ ทั้งการบรรยายในชั้นเรียน การสอนข้างเตียง และการสอนในรูปแบบอื่นๆ เกี่ยวกับการเลือกใช้ยาในระบบต่างๆ

ประโยชน์ที่ได้รับ

1. ทำให้ได้แนวทางในการตรวจสอบความเป็นคู่ขนานของแบบสอบโดยใช้ระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญ
2. ทำให้ทราบว่าระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญมีผลต่อการคัดเลือกข้อสอบเพื่อสร้างแบบสอบให้มีความเป็นคู่ขนานกันหรือไม่ เพื่อใช้เป็นสารสนเทศในการคัดเลือกผู้เชี่ยวชาญสำหรับการสร้างแบบสอบคู่ขนาน
3. ได้สารสนเทศถึงประสบการณ์สอนนักศึกษาแพทย์ของผู้เชี่ยวชาญส่งผลต่อการคัดเลือกข้อสอบเพื่อสร้างแบบสอบคู่ขนานหรือไม่ เพื่อกำหนดนโยบายในการคัดเลือกกลุ่มผู้เชี่ยวชาญในการสร้างแบบสอบคู่ขนาน



บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

การวิจัยครั้งนี้เป็นการวิจัยเพื่อเปรียบเทียบดัชนีความเป็นคู่ขนานของข้อสอบและแบบสอบที่ได้จากการคัดเลือกข้อสอบโดยกลุ่มผู้เชี่ยวชาญที่มีลักษณะแตกต่างกัน: การประยุกต์ใช้ทฤษฎีฉันทามติทางวัฒนธรรมแบบมีรูปแบบผู้วิจัยได้เสนอเอกสารและงานวิจัยที่เกี่ยวข้องโดยแบ่งออกเป็น 3 ตอน ดังนี้

ตอนที่ 1 แบบสอบคู่ขนาน

- 1.1 นิยามและการสร้างแบบสอบคู่ขนาน
- 1.2 ลักษณะของแบบสอบคู่ขนานตามทฤษฎีการตอบสนองข้อสอบ
- 1.3 ดัชนีความเป็นคู่ขนานของแบบสอบและเกณฑ์การแปลผล

ตอนที่ 2 ทฤษฎีฉันทามติทางวัฒนธรรม

- 2.1 แนวคิดและข้อตกลงเบื้องต้นของทฤษฎีฉันทามติทางวัฒนธรรม
- 2.2 โมเดลฉันทามติทางวัฒนธรรม
- 2.3 การวิเคราะห์ทางสถิติสำหรับทฤษฎีฉันทามติทางวัฒนธรรมแบบมีรูปแบบ

ตอนที่ 3 กรอบแนวคิดในการวิจัย

ตอนที่ 1 แบบสอบคู่ขนาน

แบบสอบคู่ขนาน ถูกนำมาใช้ในการสอบต่างๆอย่างแพร่หลาย โดยเฉพาะการสอบที่มีการจัดสอบหลายครั้งหรือมีกลุ่มผู้สอบหลายกลุ่ม เช่น การสอบวัดความสามารถทางภาษาอังกฤษต่างๆ การสอบเพื่อขอรับใบประกอบวิชาชีพต่างๆ การสอบเพื่อคัดเลือกบุคคลเข้าศึกษาต่อในระดับต่างๆ รวมถึงการสอบเพื่อวัดและประเมินผลในชั้นเรียน ทั้งการวัดและประเมินผลความก้าวหน้าระหว่างเรียน (formative evaluation) และรวบยอด (summative evaluation) โดยมีเป้าหมายให้แบบสอบทุกฉบับมีความเป็นคู่ขนานกันทั้งในด้านเนื้อหาและการตรวจสอบโดยอาศัยค่าทางสถิติ ในตอนที่ 1 ผู้วิจัยจึงได้แบ่งการเสนอเอกสารและงานวิจัยที่เกี่ยวข้องกับแบบสอบคู่ขนานเป็น 3 ส่วน ได้แก่ นิยามและการสร้างแบบสอบคู่ขนาน ลักษณะของแบบสอบคู่ขนานตามทฤษฎีการตอบสนองข้อสอบ และดัชนีความเป็นคู่ขนานของแบบสอบและเกณฑ์การแปลผล โดยอธิบายในแต่ละส่วน ดังนี้

1.1 นิยามและการสร้างแบบสอบคู่ขนาน

แบบสอบคู่ขนาน หมายถึง แบบสอบตั้งแต่ 2 ฉบับที่สามารถใช้ทดแทนกันได้ เนื่องจากมีโครงสร้างการวัดเนื้อหาหรือจุดประสงค์การเรียนรู้เหมือนกัน มีค่าความยากของแบบสอบหรือค่าเฉลี่ยส่วนเบี่ยงเบนมาตรฐานและความเที่ยงเท่ากัน ตามข้อตกลงเบื้องต้นของทฤษฎีการทดสอบแบบดั้งเดิม ซึ่งจะส่งผลให้คะแนนจริงของผู้สอบทั้ง 2 ฉบับมีค่าเท่ากันและความคลาดเคลื่อนมาตรฐานของประชากรที่ทำแบบสอบทั้ง 2 ฉบับมีค่าเท่ากัน ตามหลักการตรวจสอบความเป็นคู่ขนานของแบบสอบจึงพิจารณาจากค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของข้อสอบคู่ขนานแต่ละคู่มิค่าใกล้เคียงกัน แต่จะพบว่า ค่าเฉลี่ยรวมของแบบสอบคู่ขนานทั้งฉบับอาจจะแตกต่างกันอย่างมีนัยสำคัญ ซึ่งอาจเกิดจากข้อจำกัดของคะแนนความคลาดเคลื่อนและค่าพารามิเตอร์ของข้อสอบ ทั้งนี้ แบบสอบคู่ขนานที่ใช้สอบจะมีลักษณะเฉพาะขึ้นอยู่กับกลุ่มผู้ตอบและเปลี่ยนไปตามการสอบแต่ละครั้งจึงเป็นข้อจำกัดของการตรวจสอบความเป็นคู่ขนานของแบบสอบโดยอาศัยทฤษฎีการทดสอบแบบดั้งเดิม (ศิริชัย กาญจนวาสี, 2555) ส่วนในทฤษฎีการตอบสนองข้อสอบ แบบสอบคู่ขนานคือ แบบสอบที่มีโครงสร้างทางเนื้อหา ไค่งลักษณะของข้อสอบ และฟังก์ชันสารสนเทศของแบบสอบเหมือนกัน มีความน่าจะเป็นที่ผู้ตอบข้อสอบถูก ($P_i(\theta)$) เท่ากันในทุกความสามารถที่มีอยู่ในผู้ตอบ (θ) และรูปแบบการกระจายไม่แตกต่างกัน ดังนั้น การออกแบบการสร้างแบบสอบคู่ขนาน จะต้องมีการกำหนดรูปแบบของแบบสอบ สร้างแผนผังการสอบและสร้างตารางวิเคราะห์ข้อสอบซึ่งเป็นขั้นตอนที่สำคัญที่ทำให้ข้อสอบที่วัดได้ตรงตามจุดประสงค์และพฤติกรรมการเรียนรู้ที่กำหนดขึ้น นอกจากนี้ ค่าพารามิเตอร์ที่ได้จากการวิเคราะห์ข้อสอบมีลักษณะความไม่แปรเปลี่ยนของข้อสอบและแบบสอบ จึงสามารถสร้างแบบสอบได้อย่างมีประสิทธิภาพ โดยค่าสารสนเทศที่ได้จากข้อสอบเมื่อนำมารวมกันเป็นสารสนเทศ

ของแบบสอบจะสะท้อนความถูกต้องแม่นยำของการประมาณค่าความสามารถของผู้สอบ จึงทำให้ได้หลักประกันสำหรับการพัฒนาแบบสอบคู่ขนานให้ตรงกับเป้าหมายการนำไปใช้อย่างมีประสิทธิภาพ (โชติกา ภาชีผล, 2559; ศิริชัย กาญจนวาสี, 2555) โดยการสร้างแบบสอบคู่ขนานสามารถจำแนกออกเป็น 2 รูปแบบตามการบริหารการสอบ ดังแสดงรายละเอียดต่อไปนี้

1. การสร้างแบบสอบคู่ขนานหลังการสอบ เป็นการสร้างข้อสอบจากคลังข้อสอบที่มีการเก็บรวบรวมข้อมูลค่าสถิติหรือพารามิเตอร์ของข้อสอบรายข้อ เช่น ค่าพารามิเตอร์ความยากของข้อสอบ (b) ค่าพารามิเตอร์อำนาจจำแนกของข้อสอบ (a) และค่าพารามิเตอร์โอกาสในการเดาข้อสอบได้ถูก (c) รวมถึงค่าสารสนเทศของข้อสอบและแบบสอบที่ได้จากการวิเคราะห์ข้อสอบโดยอาศัยทฤษฎีการตอบสนองข้อสอบ จากการสอบครั้งก่อนเพื่อใช้กำหนดเงื่อนไขในการคัดเลือกข้อสอบเพื่อสร้างแบบสอบคู่ขนาน สามารถสร้างแบบสอบคู่ขนานโดยใช้โปรแกรมคอมพิวเตอร์เพื่อสร้างชุดแบบสอบคู่ขนานที่ต้องการ โดยมีการกำหนดขอบเขตเนื้อหา จุดประสงค์การเรียนรู้ที่ต้องการวัดและประเมินจำนวนข้อสอบ และรูปแบบของข้อสอบตามผังข้อสอบ จากนั้นจึงนำแบบสอบคู่ขนานไปใช้สอบในสภาพจริงและทำการตรวจสอบความเป็นคู่ขนานที่สร้างขึ้นโดยพิจารณาดัชนีความเป็นคู่ขนานของแบบสอบและข้อสอบตามหลักทฤษฎีการทดสอบแบบดั้งเดิมหรือทฤษฎีการตอบสนองข้อสอบที่ผู้บริหารจัดการสอบเลือกใช้ให้สอดคล้องกับการสร้างแบบสอบคู่ขนาน

2. การสร้างแบบสอบคู่ขนานก่อนการสอบ เป็นการสร้างแบบสอบคู่ขนานโดยมีการกำหนดรูปแบบของแบบสอบ สร้างแผนผังการสอบและสร้างตารางวิเคราะห์ข้อสอบตามจุดประสงค์และพฤติกรรมการเรียนรู้ การสร้างข้อสอบคู่ขนานมีความจำเป็นในการทดสอบต่างๆ วัตถุประสงค์หลักคือเพื่อลดการใช้ข้อสอบซ้ำที่ได้รับการเปิดเผยไปแล้ว ดังนั้น ผู้สร้างข้อสอบนอกจากจะต้องมีความเข้าใจในเนื้อหาเป็นอย่างดีและควรจะต้องเรียนรู้เทคนิควิธีการสร้างข้อสอบคู่ขนาน 2 ฉบับ ให้มีความเป็นคู่ขนานกัน เช่น สลับคำตอบมาเป็นคำถาม เปลี่ยนคำถามตรงกันข้ามหรือเป็นปฏิเสธ ใช้คำถามเหมือนกันแต่ตัวเลือกแตกต่างกัน ใช้คำถามต่างกันแต่ตัวเลือกเหมือนกันและเปลี่ยนถ้อยคำในคำถามและคำตอบแต่ถามในประเด็นเดียวกัน (อนุพงษ์ กันธิวงศ์ และ โชติกา ภาชีผล, 2560) โดยผู้ออกข้อสอบเขียนข้อคำถามและตัวเลือกตามจุดประสงค์การเรียนรู้และโครงสร้างการวัดที่กำหนดให้มีความเท่าเทียมกัน และตรวจสอบความเป็นคู่ขนานของแบบสอบที่สร้างขึ้นโดยอาศัยผู้เชี่ยวชาญ (โชติกา ภาชีผล, 2559) ทั้งนี้ สามารถตรวจสอบความเป็นคู่ขนานของแบบสอบรายข้อโดยประยุกต์ใช้การตรวจสอบความตรงเชิงเนื้อหา ซึ่งเป็นวิธีที่ให้ผู้เชี่ยวชาญตัดสินข้อคำถามที่สร้างขึ้นว่าเป็นไปตามเนื้อหาละวัตถุประสงค์ที่ต้องการวัดหรือไม่ มีความสอดคล้องของข้อคำถามกับเนื้อหาและพฤติกรรมการเรียนรู้ และพิจารณาถึงความเป็นคู่ขนานของแบบสอบเป็นรายข้อคำถาม โดยการระบุน้ำหนักคะแนนและนำคะแนนที่ได้แต่ละข้อมาคำนวณความสอดคล้องระหว่างข้อคำถามกับวัตถุประสงค์ (Item Objective Congruency Index: IOC) โดยมีเกณฑ์สำหรับพิจารณาค่าความสอดคล้องควรจะ

มีค่าตั้งแต่ 0.50 ขึ้นไป ข้อดีของการสร้างแบบสอบคู่ขนานก่อนการสอบและการตรวจสอบโดยผู้เชี่ยวชาญ คือ สามารถประยุกต์ใช้ในการสร้างข้อสอบในสภาพจริงได้ง่ายและสร้างข้อสอบใหม่ได้ตามโครงสร้างและจุดประสงค์ที่กำหนด แต่ข้อเสียของวิธีนี้ คือ การพิจารณาคุณภาพของแบบสอบโดยผู้เชี่ยวชาญ ขึ้นอยู่กับระดับความสามารถในการตัดสินใจรายบุคคลของผู้เชี่ยวชาญที่แตกต่างกัน (โชติกา ภาชีผล, 2559) จากการศึกษาของ Wanichsan et al. (2012) และ Panjaburee et al. (2010) พบว่า การสร้างแบบสอบโดยกลุ่มผู้เชี่ยวชาญ สามารถเพิ่มระดับความสัมพันธ์ระหว่างเนื้อหาหรือจุดประสงค์การเรียนรู้กับแบบสอบได้ดีขึ้น ทำให้การประมาณค่าความสามารถของผู้สอบมีความตรงมากขึ้น สอดคล้องกับสภาพจริงที่การคัดเลือกข้อสอบจากคลังข้อสอบเพื่อสร้างเป็นแบบสอบคู่ขนาน จะใช้กลุ่มที่ประกอบด้วยผู้เชี่ยวชาญอย่างน้อย 3 คน โดยพิจารณาจากเกณฑ์มาตรฐานจุดประสงค์การเรียนรู้ และเนื้อหาที่กำหนดตามแผนผังการสอบและตารางวิเคราะห์ข้อสอบ แต่พบว่าปัญหาของการตัดสินใจโดยอาศัยผู้เชี่ยวชาญ คือ ระดับความสามารถในการตัดสินใจของผู้เชี่ยวชาญแต่ละคนมีความแตกต่างกัน จึงส่งผลให้ความตรงและความถูกต้องของผลการตัดสินใจมีค่าที่เบี่ยงเบนไปจากค่าที่แท้จริง (Romney et al., 1986; Weller, 2007)

ตารางที่ 2.1 เปรียบเทียบลักษณะของการสร้างแบบสอบคู่ขนานก่อนและหลังนำไปใช้สอบ

รายการ	การสร้างแบบสอบคู่ขนานก่อนนำไปใช้สอบ	การสร้างแบบสอบคู่ขนานหลังนำไปใช้สอบ
1. ที่มาของข้อสอบ	ครู ผู้สอน อาจารย์ หรือ ผู้ทรงคุณวุฒิในเนื้อหานั้น เป็นผู้ ออกข้อสอบเมื่อต้องการใช้แบบ สอบสำหรับวัดและประเมินผล	คัดเลือกจากคลังข้อสอบในเนื้อหา นั้นๆ ที่ประกอบด้วยข้อสอบจากการ สอบและมีผลวิเคราะห์ข้อสอบแล้ว
2. เนื้อหาที่ใช้สร้างข้อสอบ	ตามแผนผังการสอบและตาราง วิเคราะห์ข้อสอบ	ตามแผนผังการสอบและตาราง วิเคราะห์ข้อสอบ
3. ช่วงเวลาในการสร้าง ข้อสอบ	ก่อนนำข้อสอบไปใช้สอบ	หลังนำข้อสอบไปใช้สอบ
4. วิธีการคัดเลือกข้อสอบ	คัดเลือกข้อสอบโดยผู้ทรงคุณวุฒิ หลังการสร้างข้อสอบในแต่ละครั้ง	คัดเลือกข้อสอบตามจุดประสงค์จาก คลังข้อสอบที่วิเคราะห์ข้อสอบแล้ว
5. จำนวนข้อสอบที่ใช้	ใกล้เคียงกับจำนวนข้อสอบตาม ตารางวิเคราะห์ข้อสอบ	มากกว่าจำนวนข้อสอบตามตาราง วิเคราะห์ข้อสอบ
7. ระยะเวลาการสร้าง ข้อสอบ	ใช้เวลาค่อนข้างมาก	ใช้เวลาสั้นกว่า
8. การเปิดเผยข้อสอบ	ไม่มีการนำข้อสอบไปใช้ก่อนการ สร้างแบบสอบคู่ขนาน	ข้อสอบถูกนำไปใช้สอบก่อนการสร้าง แบบสอบคู่ขนาน
9. การนำไปใช้	<ul style="list-style-type: none"> - การสอบที่ต้องการความปลอดภัย ของข้อมูลในระดับสูง - การสอบเพื่อวัดและประเมินผล ในชั้นเรียน ที่ไม่มีคลังข้อสอบ 	<ul style="list-style-type: none"> - การสอบขนาดใหญ่ที่มีคลังข้อสอบ จำนวนมาก - การสอบเพื่อวัดและประเมินผลใน รูปแบบ formative evaluation ที่ ต้องการความปลอดภัยของข้อมูลใน ระดับปานกลาง
10. ข้อจำกัด	<ul style="list-style-type: none"> - ไม่ทราบค่าสถิติและผลการ วิเคราะห์ข้อสอบก่อนการสร้าง แบบสอบคู่ขนาน - ความลำเอียงในการสร้างและ คัดเลือกข้อสอบของผู้ออกข้อสอบ 	<ul style="list-style-type: none"> - ข้อสอบบางจุดประสงค์อาจไม่ เพียงพอต่อการสร้างแบบสอบคู่ขนาน - ระดับความสามารถของผู้ทรงคุณวุฒิ ในการคัดเลือกข้อสอบ คู่ขนาน แตกต่างกัน

หลังจากสร้างแบบสอบคู่ขนานแล้ว เมื่อนำแบบสอบคู่ขนานไปใช้สอบกับผู้สอบกลุ่มเดียวกัน โดยแบบสอบคู่ขนาน 2 ฉบับมีความเป็นอิสระจากกัน จะต้องคำนึงถึงความตรงภายใน กล่าวคือ การใช้แบบสอบ A สอบก่อนเรียนและแบบสอบ B สอบหลังเรียนจะมีความตรงภายในสูงกว่าใช้แบบสอบ 2 ฉบับที่เหมือนกันมาใช้สอบซ้ำ นอกจากนี้ถ้ามีการออกแบบไขว้ชุดแบบสอบ โดยสุ่มลำดับของแบบสอบที่แต่ละคนจะได้รับ เช่น ใช้แบบสอบ A หรือ B ในการสอบก่อนเรียน และสลับเมื่อสอบหลังเรียน เพื่อลดความคลาดเคลื่อนที่เป็นผลมาจากลำดับชุดของแบบสอบได้ (R. L. Johnson, 2009) จากนั้นจึงวิเคราะห์ข้อสอบและแบบสอบเพื่อตรวจสอบความเป็นคู่ขนานของข้อสอบและแบบสอบ พบว่าจุดอ่อนของการตรวจสอบความเป็นคู่ขนานของแบบสอบโดยอาศัยทฤษฎีการทดสอบแบบดั้งเดิมคือขาดสารสนเทศในการพิจารณาความเป็นคู่ขนานของข้อสอบรายข้อ และเมื่อนำแบบสอบคู่ขนานนี้ไปใช้ในการสอบครั้งต่อไป การใช้แบบสอบซ้ำในประชากรที่แตกต่างกัน ค่าพารามิเตอร์ของข้อสอบและความเที่ยงของแบบสอบคู่ขนานอาจเปลี่ยนแปลงไปตามลักษณะของประชากร จึงมีการนำทฤษฎีการตอบสนองข้อสอบมาใช้เพื่อลดจุดอ่อนของทฤษฎีการวัดแบบดั้งเดิม โดยสามารถตรวจสอบความเป็นคู่ขนานของข้อสอบได้ในระดับรายข้อและระดับแบบสอบทั้งฉบับ โดยอยู่ภายใต้ข้อตกลงเบื้องต้นคือความไม่แปรเปลี่ยนของพารามิเตอร์ข้อสอบและความสามารถของผู้สอบซึ่งเป็นข้อได้เปรียบของทฤษฎีการตอบสนองข้อสอบ

1.2 ลักษณะของแบบสอบคู่ขนานตามทฤษฎีการตอบสนองข้อสอบ

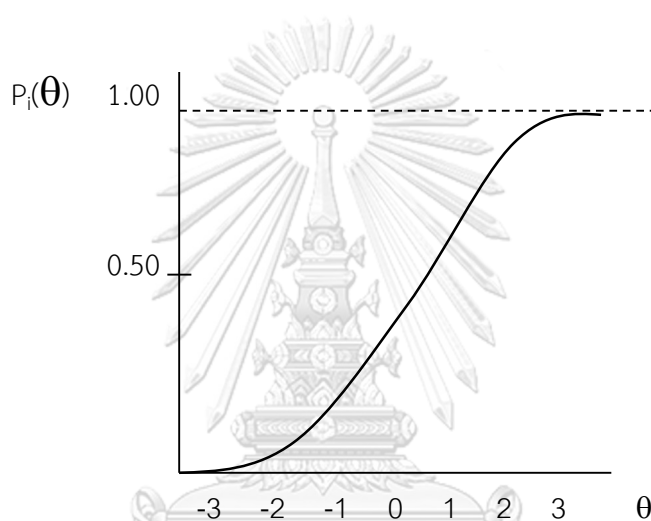
ความเป็นคู่ขนานของแบบสอบ หมายถึง แบบสอบที่มีโครงสร้างทางเนื้อหา ใ้คงลักษณะของข้อสอบ และฟังก์ชันสารสนเทศของแบบสอบเหมือนกัน มีความน่าจะเป็นที่ผู้ตอบข้อสอบถูก ($P_i(\theta)$) เท่ากันในทุกความสามารถที่มีอยู่ในผู้ตอบ (θ) และรูปแบบการกระจายไม่แตกต่างกัน ตามโมเดลทฤษฎีการตอบสนองข้อสอบแบบ 2 พารามิเตอร์ สามารถแบ่งความเป็นคู่ขนานของแบบสอบเป็น 2 ระดับ คือ ความเป็นคู่ขนานของแบบสอบทั้งฉบับ โดยตรวจสอบว่าแบบสอบมีค่าฟังก์ชันสารสนเทศของแบบสอบเท่าเทียมกันหรือมีการกระจายเหมือนกัน และความเป็นคู่ขนานของข้อสอบรายข้อ โดยตรวจสอบเพื่อระบุว่าข้อสอบมีค่าฟังก์ชันสารสนเทศของข้อสอบเท่าเทียมกันหรือมีความสอดคล้องกันในระดับความสามารถของผู้ตอบข้อสอบทุกระดับ

การตรวจสอบความเป็นคู่ขนานของแบบสอบโดยอาศัยโมเดลการตอบสนองข้อสอบแบบ 2 พารามิเตอร์ โดยข้อสอบจะต้องมีคุณสมบัติความเป็นเอกมิติและมีไ้คงลักษณะข้อสอบรูปแบบฟังก์ชันโลจิส (ศิริชัย กาญจนวาสี, 2555) ดังสมการ

$$P_i(\theta) = \frac{1}{1 + e^{-Da_i(\theta - b_i)}}$$

- เมื่อ $P_i(\theta)$ = ค่าความน่าจะเป็นที่ผู้ตอบซึ่งมีความสามารถ θ จะตอบข้อที่ i ได้ถูกต้อง
- b = ค่าพารามิเตอร์ความยากของข้อสอบข้อที่ i ซึ่งเป็นค่าที่แสดงตำแหน่งของ ICC ณ จุด θ ที่มีโอกาสตอบข้อสอบถูก 0.50
- a = ค่าพารามิเตอร์อำนาจจำแนกของข้อสอบข้อที่ i ซึ่งเป็นค่าความชันของ ICC ณ ตำแหน่ง b_i
- e = 2.718
- D = 1.70

จากสมการสามารถเขียนโค้งลักษณะข้อสอบได้ ดังภาพที่ 2.1



ภาพที่ 2.1 โค้งลักษณะข้อสอบแบบ 2 พารามิเตอร์

โค้งลักษณะข้อสอบ (Item characteristic curve; ICC) สำหรับโมเดลการตอบสนองข้อสอบแบบ 2 พารามิเตอร์ มีลักษณะที่สำคัญ คือ ค่า a และ b มีค่าแปรเปลี่ยนตามลักษณะของข้อสอบแต่ละข้อ โดยมีค่า c เท่ากับ 0 ซึ่งสามารถใช้โค้งลักษณะข้อสอบในการตรวจความเป็นคู่ขนานของข้อสอบรายข้อได้โดย พบว่า แบบสอบคู่ขนาน 2 ฉบับจะต้องมีความน่าจะเป็นที่ผู้ตอบข้อสอบถูก ($P_i(\theta)$) เท่ากันในทุกความสามารถที่มีอยู่ในผู้ตอบ (θ) และมีรูปแบบการกระจายไม่แตกต่างกัน (Linden & Luecht, 1998; Spray & Miller, 1992)

จากข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองข้อสอบ ในด้านความเป็นอิสระระหว่างข้อสอบและผู้สอบ (Local independence) ซึ่งเชื่อมโยงมาจากความเป็นเอกมิติของแบบสอบ พบว่า เมื่อมีการควบคุมความสามารถ (θ) ที่ส่งผลต่อการตอบข้อสอบให้คงที่แล้ว ผลการตอบแต่ละข้อจะต้องเป็นอิสระจากกัน หรือโมเดลการตอบสนองข้อสอบมีเพียง θ ปัจจัยเดียวเท่านั้นที่มีอิทธิพลต่อ

การตอบรายชื่อ ดังนั้นผลรวมของฟังก์ชันการตอบสนองข้อสอบจึงแสดงถึงโค้งลักษณะของแบบสอบ (Test characteristic curve; TCC, $T(\theta)$) โดยหาค่าได้จากสมการ (Ali & Rijn, 2016; ศิริชัย กาญจนวาสิ, 2555)

$$T(\theta) = \sum_{i=1}^n P_i(\theta)$$

จากการสังเคราะห์งานวิจัย พบว่า มีงานวิจัยบางส่วนที่ใช้โค้งลักษณะของแบบสอบในการตรวจสอบความเป็นคู่ขนานของแบบสอบที่สร้างขึ้น โดยกำหนดให้ $T(\theta)$ เทียบเท่ากับคะแนนจริงในทฤษฎีการตอบสนองข้อสอบได้ (Ali & Rijn, 2016; Samejima, 1977b)

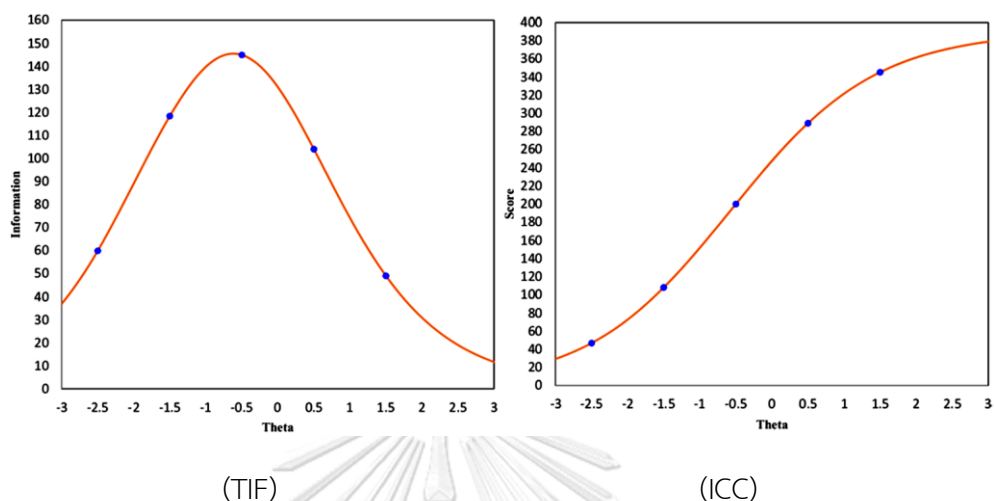
ฟังก์ชันสารสนเทศของแบบสอบ (Test information; $I(\theta)$) เป็นผลรวมเชิงพีชคณิตของค่าฟังก์ชันสารสนเทศของข้อสอบแต่ละข้อเข้าด้วยกันทั้งฉบับ ณ ตำแหน่ง θ เดียวกัน ดังสูตร

$$I(\theta) = \sum_{i=1}^k I_i(\theta) = \sum_{i=1}^k a_i^2 P_i(\theta)(1 - P_i(\theta))$$

เมื่อ $I(\theta) =$ ค่าฟังก์ชันสารสนเทศที่ได้รับจากแบบสอบสำหรับผู้ตอบที่มีความสามารถ θ
 $i = 1, 2, \dots, k$

แบบสอบคู่ขนานตามทฤษฎีการตอบสนองข้อสอบ สามารถตรวจสอบความเป็นคู่ขนานของแบบสอบ โดยฟังก์ชันสารสนเทศของแบบสอบ (Test information; TIF) โดยพบว่า แบบสอบคู่ขนานจะต้องมีฟังก์ชันสารสนเทศของแบบสอบเหมือนกัน (Linden & Adema, 1998; Veldkamp et al., 2013) สอดคล้องกับการศึกษาของ Ali and Rijn (2016) พบว่า การวิเคราะห์ดัชนีความเป็นคู่ขนานของแบบสอบโดยอาศัยทฤษฎีการตอบสนองข้อสอบ จะพิจารณาจากฟังก์ชันสารสนเทศของแบบสอบซึ่งให้สารสนเทศที่ถูกต้องและน่าเชื่อถือมากกว่าการใช้โค้งลักษณะของแบบสอบ ทั้งนี้ การวิเคราะห์ดัชนีความเป็นคู่ขนานของแบบสอบโดยอาศัยทฤษฎีการตอบสนองข้อสอบ จะพิจารณาจากฟังก์ชันสารสนเทศของแบบสอบเป็นหลัก เพราะให้สารสนเทศที่ถูกต้องและน่าเชื่อถือมากกว่าการใช้ ICC หรือ TCC เนื่องจากอาจพบว่ามี ความแตกต่างของค่าพารามิเตอร์ที่ส่งผลต่อคะแนนความสามารถของผู้สอบให้มีความคลาดเคลื่อนตามระบบได้ (Systematic error) ซึ่งจำเป็นต้องแก้ไขโดยการปรับเทียบคะแนนของผู้สอบ (Equated score) จึงทำให้การตรวจสอบความเป็นคู่ขนานของแบบสอบโดยอาศัย

โค้งลักษณะข้อสอบ (TCC) มีข้อจำกัดหากมีความแตกต่างของคะแนนระหว่างผู้สอบเกิดขึ้น จึงนิยมใช้ค่าสารสนเทศของการสอบ (TIF) ในการตรวจสอบความเป็นคู่ขนานของการสอบมากกว่า (Debeer et al., 2017)



ภาพที่ 2.2 โค้งฟังก์ชันสารสนเทศของแบบสอบ (TIF) และโค้งลักษณะของแบบสอบ (TCC)

Linden and Luecht (1998) พบว่า เมื่อแบบสอบมีฟังก์ชันสารสนเทศของแบบสอบเท่าเทียมกันจะมีค่าความแปรปรวนของคะแนนจริงเท่ากับค่าความแปรปรวนของคะแนนสังเกตได้ ซึ่งถือว่าเป็นลักษณะของแบบสอบคู่ขนานโดยอาศัยทฤษฎีการตอบสนองข้อสอบ เมื่อเปรียบเทียบกับทฤษฎีการวัดแบบดั้งเดิม ที่แบบจะเป็นคู่ขนานกันเมื่อมีคะแนนจริงของผู้สอบแต่ละคนและค่าความคลาดเคลื่อนมาตรฐานของประชากรที่สอบทั้ง 2 ฉบับเท่ากัน (Lord, 1980; Lord & Novick, 1968) โดยคะแนนจริง มีค่าเท่ากับคะแนนเฉลี่ยของคะแนนที่ได้จากการวัดซ้ำหลายๆ ครั้ง ซึ่งเป็นการพิจารณาค่าสถิติหลังจากการสอบ (ค่าเฉลี่ยรวม ค่าความคลาดเคลื่อนมาตรฐานและค่าความเที่ยงของแบบสอบ) ฟังก์ชันการตอบสนองข้อสอบ (Item response function) และ ฟังก์ชันลักษณะข้อสอบ (Test characteristic functions) เป็นค่าที่แสดงถึงคะแนนจริง ส่วนฟังก์ชันสารสนเทศของแบบสอบ (Test information function) จะแสดงถึง ค่าความแปรปรวนของความคลาดเคลื่อน ซึ่งเป็นค่าที่คงที่โดยไม่ขึ้นกับชุดข้อสอบตามข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองข้อสอบ ดังนั้น หากมีการตรวจสอบความเป็นคู่ขนานของแบบสอบโดยทฤษฎีการตอบสนองข้อสอบแล้วมีดัชนีบ่งชี้ว่าข้อสอบ 2 ฉบับมีความเป็นคู่ขนานกันระหว่างแบบสอบ จะสามารถบอกได้ว่าแบบสอบทั้งคู่มีความเป็นคู่ขนานตามทฤษฎีการวัดแบบดั้งเดิมได้ด้วย จึงมีการศึกษาค่าสถิติหรือดัชนีต่างๆ ที่นำมาใช้ตรวจสอบความเป็นคู่ขนานของแบบสอบโดยอาศัยทฤษฎีการตอบสนองข้อสอบมากขึ้น

1.3 ดัชนีความเป็นคู่ขนานของแบบสอบและเกณฑ์การแปลผล

การตรวจสอบความเป็นคู่ขนานของข้อสอบและแบบสอบโดยอาศัยทฤษฎีการตอบสนองของข้อสอบ จะทำการตรวจสอบหลังจากนำแบบสอบคู่ขนานไปใช้ โดยอาศัยค่าดัชนีความเป็นคู่ขนานซึ่งพิจารณาจากค่าสถิติที่ได้จากการวิเคราะห์เพื่อเปรียบเทียบความแตกต่างกันระหว่างข้อสอบแต่ละคู่และแบบสอบ 2 ฉบับ สามารถจำแนกเป็น 2 ประเภท แสดงรายละเอียดดังต่อไปนี้

1.3.1 ดัชนีความเป็นคู่ขนานของข้อสอบ

จากคำจำกัดความของข้อสอบคู่ขนานในทฤษฎีการตอบสนองของข้อสอบ ถ้าฟังก์ชันการตอบสนองของข้อสอบ หากพิจารณาค่า $P_i(\theta)$ ที่แสดงถึงโอกาสในการตอบข้อสอบถูกต้องของผู้สอบที่มีระดับความสามารถที่แตกต่างกัน พบว่ามีการกระจายของค่าดังกล่าวในรูปแบบเดียวกันระหว่างข้อสอบที่อยู่ในชุดข้อสอบที่แตกต่างกันในทุกๆระดับความสามารถของผู้สอบ จะถือว่าเป็นการตรวจสอบความเป็นคู่ขนานของข้อสอบรายข้อ (Item parallel test) จึงใช้ค่าไค์ลักษณะข้อสอบเป็นดัชนีในการตรวจสอบความเป็นคู่ขนานดังกล่าวเรียกว่า Item characteristic curve parallel (ICCP) (McDonald, 1999) โดยพิจารณาจากร้อยละของอัตราการร่วมกันของไค์ลักษณะข้อสอบ โดยแสดงถึงการกระจายของโอกาสในการตอบข้อสอบถูกต้องของผู้สอบที่ระดับความสามารถแตกต่างกัน เมื่อเปรียบเทียบร้อยละของอัตราการร่วมกันของไค์ลักษณะข้อสอบ หากแบบสอบที่มีการร่วมกันของไค์ลักษณะข้อสอบมากจะมีค่าร้อยละของอัตราการร่วมกันของไค์ลักษณะข้อสอบสูง (Spray & Miller, 1992)

นอกจากนี้ มีการพิจารณาค่าดัชนีการตรวจสอบเพื่อบอกระดับความเป็นคู่ขนานรายข้อระหว่างแบบสอบ 2 ชุด จากส่วนเบี่ยงเบนค่าเฉลี่ยกำลังสอง (Mean square deviation; MSD) ของฟังก์ชันสารสนเทศของข้อสอบ (MSD_{IF}) ดังสมการ ดังต่อไปนี้

$$MSD_{IF} = \frac{\sum_{j=1}^q [i(\theta_j) - i'(\theta_j)]^2}{q}$$

เมื่อ q = จำนวนระดับความสามารถที่ใช้ในการตรวจสอบฟังก์ชัน

$i(\theta)$ = ค่าสารสนเทศของข้อสอบที่ระดับความสามารถของผู้สอบ θ

ค่าดังกล่าวนำมาพิจารณาเพื่อใช้ในการคำนวณร้อยละของจำนวนข้อสอบที่มีส่วนเบี่ยงเบนค่าเฉลี่ยกำลังสองของสารสนเทศของข้อสอบ (MSD_{IF}) น้อยกว่า 0.05

การใช้ MSD_{IF} เป็นดัชนีสำหรับตัดสินความเป็นคู่ขนานของข้อสอบรายข้อ พบว่ามีความเข้มงวดของระดับความเป็นคู่ขนานค่อนข้างสูง เหมาะสำหรับการใช้สร้างข้อสอบคู่ขนานจากคลังข้อสอบ โดยการกำหนดเงื่อนไขค่าความคลาดเคลื่อนระหว่างข้อสอบให้ไม่เกินร้อยละ 5 แต่เมื่อนำมาใช้ในการตรวจสอบความเป็นคู่ขนานในสภาพจริงพบว่า มีความแตกต่างกันระหว่างสารสนเทศของข้อสอบค่อนข้างมาก จึงได้มีการเสนอดัชนีในการตรวจสอบความเป็นคู่ขนานรายข้อเพิ่มเติม โดยประยุกต์ใช้ในสภาพจริงทั้งข้อสอบที่มีรูปแบบเลือกตอบหลายตัวเลือกหรือรูปแบบผสมที่มีทั้งรูปแบบเลือกตอบหลายตัวเลือกและเติมคำตอบในข้อสอบฉบับเดียวกัน (Debeer et al., 2017)

การตรวจสอบความเป็นคู่ขนานของข้อสอบในปัจจุบันจะใช้ ดัชนีความเป็นคู่ขนานของข้อสอบรายข้อโดยอาศัยค่า Root mean square deviation of item information ($RMSD_{TIF}$) ซึ่งเป็นค่าสถิติที่บ่งชี้ว่าข้อสอบคู่ขนาน 2 ข้อมีส่วนเบี่ยงเบนค่าเฉลี่ยกำลังสองของสารสนเทศของข้อสอบสำหรับตรวจสอบการรูปแบบและการกระจายของฟังก์ชันสารสนเทศของข้อสอบ 2 ข้อที่วัดโครงสร้างทางเนื้อหาเดียวกันในระดับความสามารถของผู้ตอบข้อสอบทุกระดับ โดยกำหนดให้ข้อสอบคู่ขนานจะต้องมีค่า $RMSD_{IF}$ แตกต่างกันไม่เกิน 0.50 ทั้งนี้สามารถคำนวณได้จากสมการดังต่อไปนี้

$$RMSD_{IF} = \sqrt{\frac{\sum_{j=1}^q [i(\theta_j) - i'(\theta_j)]^2}{q}}$$

เมื่อ q = จำนวนระดับความสามารถที่ใช้ในการตรวจสอบฟังก์ชัน

$i(\theta)$ = ค่าสารสนเทศของข้อสอบที่ระดับความสามารถของผู้สอบ θ

CHULALONGKORN UNIVERSITY

ทั้งนี้การตรวจสอบความเป็นคู่ขนานของข้อสอบควรใช้ระดับความสามารถในการตรวจสอบฟังก์ชันอย่างน้อย 6 ระดับ จึงจะทำให้การตรวจสอบมีความน่าเชื่อถือและคลาดเคลื่อนน้อย (Debeer et al., 2017)

1.3.2 ดัชนีความเป็นคู่ขนานของแบบสอบ

การตรวจสอบความเป็นคู่ขนานของแบบสอบทั้งฉบับ จะใช้ดัชนีการตรวจสอบเพื่อบอกระดับความเป็นคู่ขนานระหว่างแบบสอบ 2 ชุด โดยพบว่า ถ้าแบบสอบมีโครงสร้างของข้อสอบหรือคลังฟังก์ชันสารสนเทศของแบบสอบเหมือนกัน ถือว่าเป็นแบบสอบที่มีความเป็นคู่ขนานกัน โดยจะต้องพิจารณาทุกๆค่าระดับความสามารถของผู้สอบของแบบสอบคู่ขนานนั้นๆมาเปรียบเทียบกัน (Ali & Rijn, 2016; McDonald, 1999) โดยดัชนีความเป็นคู่ขนานของแบบสอบทั้งฉบับ สามารถวิเคราะห์ได้

จากดัชนี 2 ค่า คือ โค้งลักษณะของแบบสอบ (TCC-parallel) และ ฟังก์ชันสารสนเทศของแบบสอบ (TIF-parallel) โดยพิจารณาจาก ส่วนเบี่ยงเบนค่าเฉลี่ยกำลังสอง (Mean square deviation; MSD) แบ่งเป็น 2 ค่าได้แก่ 1. ส่วนเบี่ยงเบนค่าเฉลี่ยกำลังสองของโค้งลักษณะแบบสอบ (MSD_{TCC}) 2. ส่วนเบี่ยงเบนค่าเฉลี่ยกำลังสองของฟังก์ชันสารสนเทศของแบบสอบ (MSD_{TIF}) ซึ่งคำนวณได้จากสมการดังต่อไปนี้

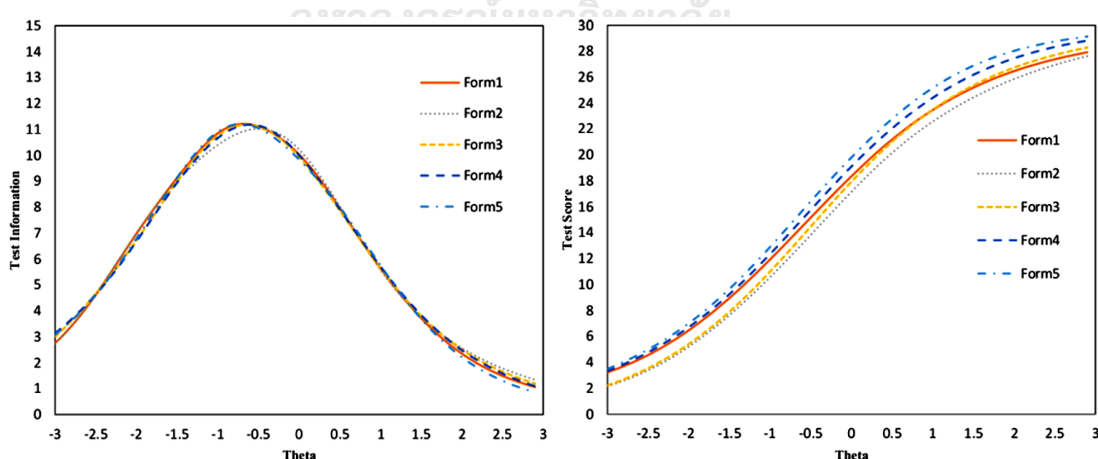
$$MSD_{TCC} = \frac{\sum_{j=1}^q [T(\theta_j) - T'(\theta_j)]^2}{q}$$

$$MSD_{TIF} = \frac{\sum_{j=1}^q [I(\theta_j) - I'(\theta_j)]^2}{q}$$

เมื่อ q = จำนวนระดับความสามารถที่ใช้ในการตรวจสอบฟังก์ชัน

$I(\theta)$ = ค่าสารสนเทศของแบบสอบที่ระดับความสามารถของผู้สอบ θ

เกณฑ์ที่ใช้ในการตัดสินผลดัชนีความเป็นคู่ขนานของแบบสอบทั้งฉบับ พิจารณาจากแบบสอบที่มีความเป็นคู่ขนานกันจะต้องมีค่า MSD_{TCC} หรือ MSD_{TIF} ไม่เกิน 0.05 (Luecht and Nunggestor, 1998)



ภาพที่ 2.3 ตัวอย่างการเปรียบเทียบโค้งฟังก์ชันสารสนเทศของแบบสอบ (TIF) และโค้งลักษณะของ แบบสอบ (TCC) ของแบบสอบคู่ขนาน 5 ฉบับ

จากการสังเคราะห์งานวิจัย (ตารางที่ 2.2) พบว่า มีการใช้ฟังก์ชันสารสนเทศของแบบสอสำหรับตรวจสอบความเป็นคู่ขนานของแบบสอทั้งฉบับมากกว่าโค้งลักษณะของแบบสอ เนื่องจากสามารถพิจารณาประสิทธิภาพสัมพัทธ์ซึ่งเป็นการเปรียบเทียบค่าฟังก์ชันสารสนเทศระหว่างแบบสอที่วัดคุณลักษณะเดียวกันได้ และสามารถใช้อ้างอิงความคลาดเคลื่อนมาตรฐานจากการประมาณค่าได้ จึงเป็นดัชนีที่ใช้ตรวจสอบความเป็นคู่ขนานของแบบสอทั้งฉบับที่มีความถูกต้องแม่นยำสูง

ตารางที่ 2.2 การศึกษาดัชนีความเป็นคู่ขนานของแบบสอ

การศึกษาดัชนีความเป็นคู่ขนานของแบบสอ	ดัชนีความเป็นคู่ขนานของแบบสอทั้งฉบับ		ดัชนีความเป็นคู่ขนานของข้อสอบรายข้อ ICC-parallel
	TCC-parallel	TIF-parallel	
Samejima (1977)	✓		
Luecht and Nungester (1998)		✓	
McDonald (1999)		✓	✓
Linden (2005)		✓	
Lin (2008)		✓	✓
Chen (2012)		✓	
Ali (2016)	✓	✓	
Debeer, Ali and Rijn (2017)	✓	✓	

ทั้งนี้การใช้ MSD_{TIF} เป็นดัชนีสำหรับตัดสินความเป็นคู่ขนานของแบบสอคู่ขนานมีความเข้มงวดของระดับความเป็นคู่ขนานของแบบสอค่อนข้างสูง นำมาใช้ในการตรวจสอบความเป็นคู่ขนานในสภาพจริงได้ยากเนื่องจากมีความไวต่อความแตกต่างกันระหว่างสารสนเทศของแบบสอค่อนข้างมาก จึงได้มีการเสนอดัชนีในการตรวจสอบความเป็นคู่ขนานของแบบสอเพิ่มเติม ที่สามารถใช้ในสภาพจริงทั้งแบบสอที่ประกอบด้วยข้อสอบที่มีรูปแบบเลือกตอบหลายตัวเลือกหรือรูปแบบผสมที่มีทั้งรูปแบบเลือกตอบหลายตัวเลือกและเติมคำตอบในข้อสอบฉบับเดียวกัน (Debeer et al., 2017) การตรวจสอบความเป็นคู่ขนานของแบบสอในปัจจุบันจะพิจารณาจากดัชนีความเป็นคู่ขนานของข้อสอบรายข้อโดยอาศัยค่า $RMSD_{TIF}$ และ MRD แสดงรายละเอียดดังต่อไปนี้

1.3.2.1 ดัชนีความเป็นคู่ขนานของแบบสอบโดยอาศัยค่า $RMSD_{TIF}$ หมายถึง ค่าสถิติที่บ่งชี้ว่าแบบสอบคู่ขนาน 2 ฉบับมีส่วนเบี่ยงเบนค่าเฉลี่ยกำลังสองของสารสนเทศของแบบสอบ (Root mean square deviation of test information; $RMSD_{TIF}$) สำหรับตรวจสอบการรูปแบบและการกระจายของฟังก์ชันสารสนเทศของแบบสอบ 2 ฉบับที่วัดโครงสร้างทางเนื้อหาเดียวกันในระดับความสามารถของผู้ตอบข้อสอบทุกระดับแตกต่างกันไม่เกิน 0.50 ซึ่งคำนวณได้จากสมการ ดังต่อไปนี้

$$RMSD_{TIF} = \sqrt{\frac{\sum_{j=1}^q [I(\theta_j) - I'(\theta_j)]^2}{q}}$$

เมื่อ q = จำนวนระดับความสามารถที่ใช้ในการตรวจสอบฟังก์ชัน
 $I(\theta)$ = ค่าสารสนเทศของแบบสอบที่ระดับความสามารถของผู้สอบ θ

ทั้งนี้ การพิจารณาดัชนีความเป็นคู่ขนานของแบบสอบโดยอาศัยค่า $RMSD_{TIF}$ ควรพิจารณา ร่วมกับค่าผลต่างระหว่างสารสนเทศของแบบสอบสูงสุดที่ระดับความสามารถของผู้สอบเดียวกัน (Maximum score gap; MSG) โดยค่า MSG ของแบบสอบคู่ขนานจะไม่เกิน 0.50 ซึ่งคำนวณได้จากสมการ ดังต่อไปนี้

$$MSG = \max(\theta) \{ \max I(\theta) - \min I'(\theta) \}$$

เมื่อ $I(\theta)$ = ค่าสารสนเทศของแบบสอบที่ระดับความสามารถของผู้สอบ θ

1.3.2.2 ดัชนีความเป็นคู่ขนานของแบบสอบโดยอาศัยค่า MRD หมายถึง ค่าสถิติที่บ่งชี้ว่าแบบสอบคู่ขนาน 2 ฉบับมีค่าเฉลี่ยสัดส่วนที่แตกต่างกันของสารสนเทศของแบบสอบ (MRD) สำหรับตรวจสอบการรูปแบบและการกระจายของฟังก์ชันสารสนเทศของแบบสอบ 2 ฉบับที่วัดโครงสร้างทางเนื้อหาเดียวกันในระดับความสามารถของผู้ตอบข้อสอบทุกระดับแตกต่างกันไม่เกิน 0.10 ซึ่งคำนวณได้จากสมการ ดังต่อไปนี้

$$MRD = \frac{\sum_{j=1}^q \left| \frac{I(\theta_j) - I'(\theta_j)}{I'(\theta_j)} \right|}{q}$$

เมื่อ q = จำนวนระดับความสามารถที่ใช้ในการตรวจสอบฟังก์ชัน
 $I(\theta)$ = ค่าสารสนเทศของแบบสอบที่ระดับความสามารถของผู้สอบ θ

ทั้งนี้ การพิจารณาดัชนีความเป็นคู่ขนานของแบบสอบโดยอาศัยค่า MRD ควรพิจารณา ร่วมกับค่าผลต่างสัมพัทธ์ระหว่างสารสนเทศของแบบสอบสูงสุดที่ระดับความสามารถของผู้สอบ เดียวกัน (Maximum relative information gap; MRIG) โดยค่า MRIG ของแบบสอบคู่ขนานจะไม่ เกิน 0.10 ซึ่งคำนวณได้จากสมการ ดังต่อไปนี้

$$MRIG = \max(\theta) \left\{ \frac{\max I(\theta) - \min I'(\theta)}{I'(\theta)} \right\}$$

เมื่อ $I(\theta)$ = ค่าสารสนเทศของแบบสอบที่ระดับความสามารถของผู้สอบ θ

ทั้งนี้การตรวจสอบความเป็นคู่ขนานของแบบสอบในทุกดัชนีความเป็นคู่ขนานของแบบสอบ ควรใช้ระดับความสามารถในการตรวจสอบฟังก์ชันอย่างน้อย 6 ระดับ จึงจะทำให้การตรวจสอบมี ความน่าเชื่อถือและคลาดเคลื่อนน้อย สอดคล้องกับงานวิจัยของ Debeer et al. (2017) พบว่า การ ใช้ระดับความสามารถของผู้สอบเพียง 3 ระดับ (-0.4, 0.5 และ 1.4) จะทำให้เกิดความคลาดเคลื่อน ในการคำนวณค่า MRD และ MRIG สูงกว่าการใช้ระดับความสามารถของผู้สอบ 6 ระดับ (-1.0, -0.4, 0.2, 0.8, 1.4 และ 2.0) และเมื่อพิจารณาค่า RMSD และ MSG จะพบความคลาดเคลื่อนได้มาก เช่นเดียวกับค่า MRD และ MRIG ในแบบสอบที่มีรูปแบบข้อสอบแบบหลายตัวเลือกสอดคล้องกับ งานวิจัยของ Linden and Adema (1998) ที่เสนอให้ใช้อย่างน้อย 5 ระดับความสามารถของผู้สอบ ขึ้นไป ดังนั้น การตรวจสอบความเป็นคู่ขนานของแบบสอบจึงต้องใช้ระดับความสามารถของผู้สอบ อย่างน้อย 6 ระดับขึ้นไปเพื่อเพิ่มความน่าเชื่อถือและลดความคลาดเคลื่อนดังกล่าว

ตอนที่ 2 ทฤษฎีฉันทามติทางวัฒนธรรม

ในตอนที 2 ผู้วิจัยได้เสนอเอกสารและงานวิจัยที่เกี่ยวข้องกับแบบสอบถาม โดยแบ่งออกเป็น 3 ส่วน ได้แก่ แนวคิดและข้อตกลงเบื้องต้นของทฤษฎีฉันทามติทางวัฒนธรรม โมเดลฉันทามติทางวัฒนธรรม และการวิเคราะห์ทางสถิติสำหรับทฤษฎีฉันทามติทางวัฒนธรรมแบบมีรูปแบบ โดยอธิบายในแต่ละส่วน ดังนี้

2.1 แนวคิดและข้อตกลงเบื้องต้นของทฤษฎีฉันทามติทางวัฒนธรรม

การสร้างแบบสอบถามโดยกลุ่มผู้เชี่ยวชาญ สามารถเพิ่มระดับความสัมพันธ์ระหว่างเนื้อหาหรือจุดประสงค์การเรียนรู้กับแบบสอบถามได้ดีขึ้น ซึ่งลดข้อดีของการสร้างแบบสอบถามโดยใช้โปรแกรมคอมพิวเตอร์เนื่องจากมีความซับซ้อนของรายละเอียดในแต่ละเนื้อหา ทำให้การประมาณค่าความสามารถของผู้สอบมีความตรงมากขึ้น และสามารถให้ข้อมูลย้อนกลับได้อย่างมีประสิทธิภาพสำหรับข้อสอบวินิจฉัยที่ต้องการความสัมพันธ์ระหว่างข้อสอบกับเนื้อหาค่อนข้างสูง สอดคล้องกับสภาพจริงที่การคัดเลือกข้อสอบจากคลังข้อสอบเพื่อสร้างเป็นแบบสอบถาม จะใช้กลุ่มที่ประกอบด้วยผู้เชี่ยวชาญอย่างน้อย 3 คน โดยพิจารณาจากเกณฑ์มาตรฐาน จุดประสงค์การเรียนรู้ และเนื้อหาที่กำหนดตามแผนผังการสอบและตารางวิเคราะห์ข้อสอบเพื่อให้แบบสอบถามที่สร้างขึ้นมีความเที่ยงและความตรงสูง แต่พบว่า ปัญหาของการตัดสินโดยอาศัยผู้เชี่ยวชาญ คือ ระดับความสามารถในการตัดสินของผู้เชี่ยวชาญแต่ละคนมีความแตกต่างกัน จึงส่งผลให้ความตรงและความถูกต้องของผลการตัดสินมีค่าที่เบี่ยงเบนไปจากค่าที่แท้จริง ทั้งนี้ ขึ้นอยู่กับประสบการณ์ในการสอนรายวิชานั้นๆ รวมถึงประสบการณ์และความเชี่ยวชาญในการออกข้อสอบ คัดเลือกข้อสอบ รวมถึงการตัดสินความตรงตามเนื้อหาของข้อสอบตามจุดประสงค์การเรียนรู้รายวิชานั้นๆ ของแต่ละผู้เชี่ยวชาญด้วย (Romney et al., 1986; Weller, 2007) จึงเป็นที่มาของการนำทฤษฎีฉันทามติทางวัฒนธรรม (Cultural consensus theory) มาประยุกต์ใช้เพื่อตรวจสอบความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญ ทั้งนี้ ฉันทามติ เป็นกระบวนการตัดสินใจที่กลุ่มคนต้องการให้เกิดการมีส่วนร่วมและมีความเห็นพ้องร่วมกันมากที่สุด การสรุปผลการตัดสินใจขึ้นกับข้อกำหนดที่ตกลงกันไว้ ตามน้ำหนักและขอบเขตของการเห็นพ้องร่วมกัน ซึ่งถูกออกแบบมาให้มีความแตกต่างจากการตัดสินใจโดยใช้เสียงข้างมาก (Majority decision) โดย ฉันทามติจะไม่เน้นที่การลงคะแนนเสียง เพราะการลงคะแนนเสียงอาจทำให้เสียงข้างน้อย (Minorities) ถูกกลืนหายไป แต่จะเน้นที่กระบวนการในการอภิปรายถกเถียง รับฟังความคิดเห็นเพื่อหาข้อตกลงร่วมกันระหว่างทุกฝ่าย ทำให้เป็นรูปแบบการตัดสินใจที่ป้องกันเผด็จการเสียงข้างมาก (Majority tyranny) ได้

การตัดสินใจโดยใช้ฉันทามติ เป็นกระบวนการตัดสินใจของกลุ่มคนจำนวนมากอาจเห็นไม่ตรงกัน แต่มีการสรุปผลการตัดสินใจตามข้อกำหนดที่ตกลงกันไว้ล่วงหน้า โดยน้ำหนักและขอบเขตของการเห็นพ้องร่วมกันมีหลายขนาด เช่น ทุกคนเห็นพ้องร่วมกัน เสียงข้างมาก เสียงสองในสาม หรือมากกว่า ครึ่งหนึ่ง บางกรณีอาจใช้การตัดสินใจของคณะกรรมการที่ตั้งขึ้นก็ได้ หากพิจารณาใช้เสียงเห็นพ้องร่วมกันอย่างสมบูรณ์ (Unanimity) ผู้ที่ไม่เห็นด้วยเพียงหนึ่งเสียงก็สามารถยับยั้งการตัดสินใจนั้นได้ ขึ้นอยู่กับกติกาที่กำหนดไว้ (Hartnett, 2011) ต่อมามีการเสนอทฤษฎีฉันทามติทางวัฒนธรรม (Cultural consensus theory; CCT) โดย Romney et al. (1986) ซึ่งเป็นเทคนิคในการวิเคราะห์เกี่ยวกับการคัดเลือกหรือการลงฉันทามติเพื่อตัดสินใจคำตอบ โดยใช้การประมาณค่าระดับความสามารถในการเห็นพ้องร่วมกันของบุคคลว่ามีความสามารถในการตัดสินใจหรือตอบข้อคำถามในเรื่องนั้นๆ มากน้อยเพียงใด และสามารถประมาณค่าความถูกต้องของคำตอบหรือฉันทามติของบุคคลต่อคำตอบของกลุ่มหรือคณะผู้ลงฉันทามติ โดยความน่าจะเป็นของคำตอบสามารถประมาณค่าได้จากความสัมพันธ์ระหว่างฉันทามติภายในกลุ่มต่อข้อคำถามนั้นๆ โดยอาศัยทฤษฎีและการวิเคราะห์ทางสถิติในรูปแบบต่างๆ ทั้งนี้มีการนำทฤษฎีฉันทามติทางวัฒนธรรมไปใช้อย่างแพร่หลายในสาขาสังคมศาสตร์และมานุษยวิทยา ต่อมาจึงมีการประยุกต์ใช้ในสาขาอื่นๆ เช่น นิเทศศาสตร์ จิตวิทยา แพทยศาสตร์ และศึกษาศาสตร์ โดยแนวคิดของทฤษฎีฉันทามติทางวัฒนธรรม จึงอยู่บนพื้นฐานของการวิเคราะห์ 2 ขั้นตอน คือ 1. ประมาณค่าความสามารถในการเห็นพ้องร่วมกันของบุคคล 2. ถ่วงน้ำหนักคำตอบของบุคคลด้วยระดับความสามารถในการเห็นพ้องร่วมกัน เพื่อให้การตัดสินใจมีความถูกต้องมากขึ้น ทั้งนี้การประมาณค่าความน่าจะเป็นของคำตอบที่ถูกต้องของข้อคำถามเพื่อใช้ตรวจสอบระดับความสามารถในการเห็นพ้องร่วมกันจะต้องเป็นชุดข้อคำถามที่ศึกษาในเรื่องเดียวกัน โดยผลการตอบหรือฉันทามติของบุคคลจะต้องไม่มีการให้คะแนนแยกรายบุคคลโดยไม่เทียบกับฉันทามติจากกลุ่มที่ศึกษา ไม่มีการเปลี่ยนแปลงหรือแก้ไขข้อความหรือคำตอบ และต้องมีการแปลผลตามระดับหรือเกณฑ์ที่กำหนดขึ้น เนื่องจากทฤษฎีฉันทามติทางวัฒนธรรมจะใช้คำตอบหรือฉันทามติที่เกิดขึ้นโดยตรง ไม่สามารถดัดแปลงฉันทามติที่บุคคลหรือผู้เชี่ยวชาญได้ให้ไว้จึงจะสามารถประมาณค่าความสามารถในการเห็นพ้องร่วมกันได้ถูกต้องที่สุด (Weller, 2007) จึงได้มีการกำหนดข้อตกลงเบื้องต้นของทฤษฎีฉันทามติทางวัฒนธรรม ประกอบด้วย 3 ข้อ (S. Borgatti, 1996) ดังนี้ 1. ความเป็นอิสระต่อกันของผู้ตอบ (Conditional independence) โดยกำหนดให้ผู้เชี่ยวชาญตอบคำถามแบบวัดระดับความสามารถในตัดสินใจความเป็นคู่ขนานของข้อสอบโดยไม่สามารถสอบถามความคิดเห็นระหว่างกัน ไม่มีการปรึกษากันและไม่ใช้การสัมภาษณ์กลุ่มย่อย 2. ความเป็นเอกมิติของแบบวัด (Homogeneity) ซึ่งข้อคำถามที่ใช้ต้องมีความเท่าเทียมกันและวัดเพียงลักษณะเดียวเท่านั้น และ 3. ความถูกต้อง (Common truth) โดยกำหนดให้คำตอบของแบบวัดระดับความสามารถในตัดสินใจ

ความเป็นคู่ขนานของข้อสอบมีเพียงคำตอบเดียว โดยมุ่งเน้นเพื่อประเมินระดับของฉันทามติ ไม่ใช่การสร้างฉันทามติจากกลุ่มผู้เชี่ยวชาญ

2.2 โมเดลฉันทามติทางวัฒนธรรม

โมเดลฉันทามติทางวัฒนธรรม ถูกนำมาประยุกต์ใช้อย่างแพร่หลาย เพื่อตอบจุดประสงค์ที่สำคัญในการหาข้อยุติของคำถามหรือปัญหาที่ยังไม่สามารถยืนยันคำตอบที่แน่ชัดได้ โดยมีหลักการพื้นฐานที่ว่า ความเชื่อของแต่ละบุคคลอาจจะไม่สามารถระบุได้ชัดเจนว่าคำตอบของคำถามได้อย่างไร ถูกต้องแน่นอน แต่สามารถนำไปเปรียบเทียบกับความเชื่อหรือการตอบสนองต่อคำถามในกลุ่มคนที่อยู่ในสถานะแวดล้อมหรือวัฒนธรรมเดียวกันเพื่อหาความน่าจะเป็นของความเห็นพ้องต้องกันหรือฉันทามติได้ ซึ่งจะนำมาถึงการบ่งชี้ระดับความสามารถในการตัดสินใจหรือการลงฉันทามติของบุคคลนั้นๆ ต่อฉันทามติโดยส่วนรวม เรียกว่า ความสามารถในการตัดสินใจหรือความสามารถในการเห็นพ้องร่วมกัน (Cultural competence) ความสามารถนี้จะถูกนำมาใช้เพื่อหาคำตอบที่มีความถูกต้องบนพื้นฐานของการวิเคราะห์ฉันทามติ (Consensus analysis) สำหรับเป็นสารสนเทศในการตัดสินใจในสังคมนั้นๆ ทั้งนี้ สามารถจำแนกโมเดลฉันทามติทางวัฒนธรรมโดยการพิจารณาตามรูปแบบการประมาณค่าความสามารถในการเห็นพ้องร่วมกัน ได้ 2 โมเดล แสดงรายละเอียดดังต่อไปนี้ (Romney et al., 1986)

2.2.1 โมเดลฉันทามติทางวัฒนธรรมแบบไม่มีรูปแบบ (Informal cultural consensus model)

เป็นวิธีการวิเคราะห์เพื่อระบุระดับความสอดคล้องของการตอบข้อคำถามของผู้เชี่ยวชาญแต่ละคนกับฉันทามติของกลุ่ม โดยไม่เน้นการหาคะแนนความถูกต้องของฉันทามติของบุคคลต่อฉันทามติของกลุ่ม โดยมีแนวคิดที่ฉันทามติเกิดขึ้นมาจากข้อคำถามหรือเนื้อหาที่ยังไม่ทราบคำตอบที่แน่ชัด ดังนั้นจึงเน้นการหาค่าถ่วงน้ำหนักฉันทามติของผู้เชี่ยวชาญต่อฉันทามติของกลุ่มว่าควรถ่วงน้ำหนักแต่ละผู้เชี่ยวชาญมากน้อยเพียงใดตามความสอดคล้องของฉันทามติของกลุ่ม โดยประยุกต์ใช้การวิเคราะห์องค์ประกอบและพิจารณาจากค่าสัมประสิทธิ์สหสัมพันธ์ของเพียร์สัน โมเดลฉันทามติทางวัฒนธรรมแบบไม่มีรูปแบบนิยมใช้ในการหาฉันทามติที่ไม่ทราบคำตอบที่แน่ชัด เช่น การหาฉันทามติของรูปแบบการให้การสนับสนุนทางสังคม (Dressler, McBalieiro, & Dos Santos, 1997) การศึกษาฉันทามติของนักเรียนต่อสาเหตุสูงสุดและต่ำสุดของการเสียชีวิต (Romney et al., 1987) การศึกษาฉันทามติเพื่อเรียงลำดับความสำคัญของอาชีพ (Magana, Burton, & Ferreira-Pinto, 1995) เป็นต้น

2.2.2 โมเดลฉันทามติทางวัฒนธรรมแบบมีรูปแบบ (Formal cultural consensus model)

เป็นวิธีการประมาณค่าความสามารถในการเห็นพ้องร่วมกันของแต่ละบุคคล โดยได้จากความเห็นพ้องต้องกันหรือเปรียบเทียบคำตอบระหว่างผู้เชี่ยวชาญภายในกลุ่ม ข้อสำคัญของโมเดลนี้คือ จะต้องมีการประมาณค่าความสามารถในการเห็นพ้องร่วมกันโดยเปรียบเทียบกับคำตอบที่ถูกต้องของฉันทามติกลุ่มภายใต้ข้อตกลงเบื้องต้นอย่างเคร่งครัดตามรูปแบบของทฤษฎีฉันทามติทางวัฒนธรรม โดยวิเคราะห์ความถูกต้องของคำตอบต่อข้อคำถามโดยการประมาณค่าจากคำตอบที่ได้จากกลุ่มที่เกิดจากความแตกต่างกันของการตอบข้อคำถาม โดยมีแนวคิดเบื้องต้นคือ คำตอบส่วนใหญ่ที่ได้จากฉันทามติของกลุ่มเป็นความเชื่อหรือองค์ความรู้ที่กลุ่มคาดว่าจะถูกต้อง (D'Andrade, 1987) เช่น การคำนวณร้อยละของการตอบข้อคำถามจำกัดคำตอบ 2 คำตอบ ถ้ามีความเห็นพ้องร่วมกันสูง (Very strong agreement) ในกรณีตัวอย่าง ร้อยละ 99 ตอบว่าใช่ และมีเพียง ร้อยละ 1 ที่ตอบว่าไม่ใช่ คำตอบดังกล่าวแสดงให้เห็นถึงคำตอบที่ใช้อ้างอิงว่าเป็นฉันทามติที่ถูกต้องของกลุ่ม แต่บางครั้งพบว่า สัดส่วนของฉันทามติไม่สามารถแบ่งแยกได้อย่างชัดเจน เช่น ร้อยละ 51 ตอบว่าใช่ และร้อยละ 49 ตอบว่าไม่ใช่ จะไม่สามารถตอบได้เลยว่าความเห็นพ้องร่วมกันหรือฉันทามติของกลุ่มใดถูกต้อง ดังนั้นจึงมีการนำสถิติทดสอบมาใช้เพื่อตรวจสอบว่าฉันทามติที่มีคำตอบ 2 คำตอบมีค่าแตกต่างจากสัดส่วน 50:50 หรือไม่ โดยใช้ binomial test เพื่อใช้บอกข้อยู่ว่าฉันทามติใดเป็นฉันทามติที่ถูกต้องของกลุ่มอย่างมีนัยสำคัญทางสถิติ จึงนำมาสู่การพัฒนาแนวคิดและมีการนำการวิเคราะห์ทางสถิติมาประยุกต์ใช้เพื่อตรวจสอบและประมาณค่าความสามารถในการเห็นพ้องร่วมกันและตรวจสอบคำตอบที่ถูกต้องของข้อคำถามต่างๆในทฤษฎีฉันทามติทางวัฒนธรรม

2.3 การวิเคราะห์ทางสถิติสำหรับทฤษฎีฉันทามติทางวัฒนธรรมแบบมีรูปแบบ

เป็นการประมาณค่าคะแนนความสามารถในการเห็นพ้องร่วมกันของผู้เชี่ยวชาญ โดยกำหนดให้ พารามิเตอร์ D_i คือ คะแนนความสามารถในการเห็นพ้องร่วมกันของผู้เชี่ยวชาญคนที่ i ซึ่งเป็นค่าที่บอกถึงความน่าจะเป็นที่ผู้เชี่ยวชาญคนที่ i จะตอบได้ถูกต้องในข้อคำถามต่างๆ พบว่า $0 \leq D_i \leq 1$ และความน่าจะเป็นที่ผู้เชี่ยวชาญตอบข้อคำถามนั้นไม่ถูกต้องมีค่าเท่ากับ $1 - D_i$ นอกจากนี้ ความน่าจะเป็นในการเดาคำตอบได้ถูกต้องโดยผู้เชี่ยวชาญ มีค่าเท่ากับ $\frac{1}{L} (1 - D_i)$ เมื่อ L คือจำนวนตัวเลือกของข้อคำถามนั้นๆ

ตัวอย่าง กำหนดให้ผู้เชี่ยวชาญมีคะแนนความสามารถในการเห็นพ้องร่วมกันเท่ากับ 0.70 ($D_i = 0.70$) ในการตอบข้อคำถามที่มีตัวเลือกทั้งหมด 5 ตัวเลือก ($L = 5$) จะพบว่าความน่าจะเป็นในการ

เดาคำตอบได้ถูกต้องมีค่าเท่ากับ $\frac{1}{5}(1 - 0.70) = 0.06$ ดังนั้น ผู้เชี่ยวชาญจึงมีความน่าจะเป็นทั้งหมดที่จะตอบข้อคำถามได้ถูกต้องเท่ากับ $0.70 + 0.06 = 0.76$

ซึ่งสามารถเขียนสมการเพื่อประมาณค่าความน่าจะเป็นที่ผู้เชี่ยวชาญจะตอบคำถามได้ถูกต้อง ดังนี้

$$\Pr(Y_{ik} = 1) = D_i + \frac{1 - D_i}{L}$$

เมื่อ $Y_{ik} \begin{cases} 1 = \text{เมื่อผู้เชี่ยวชาญตอบคำถามข้อที่ } k \text{ ถูกต้อง} \\ 0 = \text{เมื่อผู้เชี่ยวชาญตอบคำถามข้อที่ } k \text{ ไม่ถูกต้อง} \end{cases}$

และสมการการประมาณค่าความน่าจะเป็นที่ผู้เชี่ยวชาญจะตอบข้อคำถามไม่ถูกต้อง ดังนี้

$$\Pr(Y_{ik} = 0) = \frac{(1 - D_i)(L - 1)}{L}$$

โดยในทางปฏิบัติพบว่า คะแนนความสามารถในการเห็นพ้องร่วมกันที่แท้จริง ไม่สามารถหาค่าได้โดยตรงเพราะมีค่าความน่าจะเป็นในการเดาคำตอบได้ถูกต้องรวมอยู่ด้วย แต่สามารถประมาณค่าได้โดยการปรับแก้ ดังสมการ

$$\hat{D}_i = \frac{(LT_i - 1)}{(L - 1)}$$

เมื่อ $T_i =$ สัดส่วนของการตอบข้อคำถามได้ถูกต้องของผู้เชี่ยวชาญคนที่ i

ในกรณีที่ไม่ทราบคำตอบของข้อคำถามและต้องการประมาณค่าความสามารถในการเห็นพ้องร่วมกัน จะไม่สามารถใช้สมการข้างต้นในการประมาณค่า D_i ได้ จึงมีการใช้ความสอดคล้องของการตอบระหว่างผู้เชี่ยวชาญมาใช้ในการประมาณค่า โดยกำหนดให้ผู้เชี่ยวชาญ 2 คนคือ คนที่ i และ j มีความน่าจะเป็นในการตอบข้อคำถามได้ถูกต้อง ดังสมการ

สำหรับผู้เชี่ยวชาญคนที่ i

$$\Pr(\text{correct}) = D_i + \frac{1 - D_i}{L}$$

สำหรับผู้เชี่ยวชาญคนที่ j

$$\Pr(\text{correct}) = D_j + \frac{1 - D_j}{L}$$

เมื่อพิจารณาจากการจับคู่ ผลการตอบข้อความจากผู้เชี่ยวชาญทั้ง 2 คน ได้ผลดังนี้

1. ความน่าจะเป็นที่ผู้เชี่ยวชาญทั้ง 2 คน ตอบคำถามได้ถูกต้อง มีค่าเท่ากับ $D_i D_j$
2. ความน่าจะเป็นที่ผู้เชี่ยวชาญคนใดคนหนึ่งตอบได้ถูกต้อง (ปรับค่าตามความน่าจะเป็นในการเดาคำตอบได้ถูกต้อง) มีค่าเท่ากับ $\frac{D_i(1-D_i)}{L}$ หรือ $\frac{D_j(1-D_j)}{L}$
3. ความน่าจะเป็นที่ผู้เชี่ยวชาญทั้ง 2 คน ตอบคำถามไม่ถูกต้อง มีค่าเท่ากับ $\frac{(1-D_i)(1-D_j)}{L}$

จากความน่าจะเป็นทั้ง 4 รูปแบบ สามารถเขียนสมการได้ ดังนี้

$$\begin{aligned} \Pr(M_{ij,k}) &= D_i D_j + \frac{D_i(1-D_i)}{L} + \frac{D_j(1-D_j)}{L} + \frac{(1-D_i)(1-D_j)}{L} \\ &= D_i D_j + \frac{(1-D_i D_j)}{L} \end{aligned}$$

$$\text{เมื่อ } M_{ij,k} = \begin{cases} 1 & \text{เมื่อ } i \text{ และ } j \text{ ตอบสอดคล้องกันในข้อความที่ } k \\ 0 & \text{เมื่อ } i \text{ และ } j \text{ ตอบไม่สอดคล้องกันในข้อความที่ } k \end{cases}$$

สามารถประมาณค่า $D_i D_j$ ได้ดังสมการ

$$\widehat{D_i D_j} = \frac{LM_{ij} - 1}{L - 1}$$

การประมาณค่าความสามารถในการเห็นพ้องร่วมกัน โดยอาศัยความสอดคล้องของการตอบระหว่างผู้เชี่ยวชาญทุกคู่ สามารถประมาณค่าได้ดังสมการ

$$M_{ij}^* = \frac{LM_{ij} - 1}{L - 1}$$

จากสมการแสดงความน่าจะเป็นของความสอดคล้องในการตอบข้อความของผู้เชี่ยวชาญคนที่ i และ j ที่ถูกปรับแก้แล้วตามข้อตกลงเบื้องต้น สามารถเขียนอยู่ในรูปของเมทริกซ์ได้ ดังนี้

$$\begin{pmatrix} D_1^2 & M_{12}^* & \cdot & \cdot & M_{1j}^* & \cdot & \cdot & M_{1N}^* \\ M_{21}^* & D_2^2 & \cdot & \cdot & M_{2j}^* & \cdot & \cdot & M_{2N}^* \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ M_{i1}^* & M_{i2}^* & \cdot & \cdot & M_{ij}^* & \cdot & \cdot & M_{iN}^* \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ M_{N1}^* & M_{N2}^* & \cdot & \cdot & M_{Nj}^* & \cdot & \cdot & D_N^2 \end{pmatrix} = \begin{pmatrix} D_1 \\ D_2 \\ \vdots \\ D_i \\ \vdots \\ D_N \end{pmatrix} (D_1, D_2, \dots, D_j, \dots, D_N)$$

เมื่อ $M_{ij}^* = M_{ji}^*$ ในทุกคู่ของผู้เชี่ยวชาญคนที่ i และ j

จากเมทริกซ์ดังกล่าวสามารถใช้การวิเคราะห์หึ่งค์ประกอบ เพื่อประมาณค่าความสามารถในการเห็นพ้องร่วมกันได้ แต่การตรวจสอบความถูกต้องของคำตอบต่อข้อคำถามไม่สามารถทำได้ด้วยวิธีดังกล่าว จึงมีการประยุกต์ใช้ทฤษฎีของเบย์ (Bayes' Theorem) มาหาค่าความน่าจะเป็นของการตอบคำถามถูกต้อง (Posteriori probability) โดยอาศัยความน่าจะเป็นก่อนเกิดเหตุการณ์ (Priori probability) และข้อมูลการตอบข้อคำถามของผู้เชี่ยวชาญ (Evidence of the informants' response)

กำหนดให้ X_1 และ X_2 คือ การตอบข้อคำถามของผู้เชี่ยวชาญ คนที่ 1 และ 2

$\Pr(T)$ และ $\Pr(F)$ คือ ความน่าจะเป็นก่อนเกิดเหตุการณ์

$\Pr(T | <X_1, X_2>)$, $\Pr(F | <X_1, X_2>)$, คือ ความน่าจะเป็นหลังเกิดเหตุการณ์

และสามารถคำนวณ ค่าความน่าจะเป็นของการตอบคำถามถูกต้อง ได้ดังสมการ

$$\Pr(T | <X_1, X_2>) = \frac{\Pr(<X_1, X_2> | T) \Pr(T)}{\Pr(<X_1, X_2> | T) \Pr(T) + \Pr(<X_1, X_2> | F) \Pr(F)}$$

โดย $\Pr(T | <X_1, X_2>)$ คือ ความน่าจะเป็นของการตอบข้อคำถามนั้นได้ถูกต้อง

ตัวอย่าง กำหนดให้ ผู้เชี่ยวชาญ 2 คนมีคะแนนความสามารถในการวินิจฉัยได้ตรงกับฉันทามติ เท่ากับ 0.80 และ 0.20 ($D_1 = 0.80$ และ $D_2 = 0.20$) ถ้าไม่ทราบการตอบข้อคำถามจะพบว่าความน่าจะเป็นก่อนเกิดเหตุการณ์ กรณีข้อคำถามประเภทถูกผิด ได้ 2 กรณีคือ ความน่าจะเป็นในการตอบคำถามถูกต้อง เท่ากับ 0.50 และความน่าจะเป็นในการตอบคำถามไม่ถูกต้อง เท่ากับ 0.50 (กำหนดให้ 1 = ตอบข้อคำถามถูกต้อง และ 0 = ตอบข้อคำถามไม่ถูกต้อง)

จากสมการ $\Pr(\text{correct}) = D_i + \frac{1-D_i}{L}$ สามารถคำนวณความน่าจะเป็นของการตอบคำถามถูกต้องของผู้เชี่ยวชาญทั้ง 2 คน ได้ค่า $\Pr(<X_1, X_2> | T) = \left(0.80 + \frac{1-0.80}{2}\right) \left(0.20 + \frac{1-0.20}{2}\right) = 0.54$

และคำนวณความน่าจะเป็นของการตอบข้อคำถามนั้นว่าเป็นคำตอบที่ถูกต้อง ได้ดังสมการ

$$\Pr(T I < X1, X2 >) = \frac{0.54 \times 0.50}{0.54 \times 0.50 + 0.40 \times 0.50} = 0.931$$

เมื่อพิจารณารูปแบบการตอบคำถามของผู้เชี่ยวชาญ 2 คน จะคำนวณความน่าจะเป็นของการตอบคำถามถูกต้อง แสดงดังตาราง (Romney et al., 1986)

ตารางที่ 2.3 ความน่าจะเป็นของการตอบคำถามถูกต้องของรูปแบบการตอบคำถามของผู้เชี่ยวชาญ 2 คน

รูปแบบการตอบข้อ คำถาม	ความน่าจะเป็นของการตอบคำถาม		ค่าความน่าจะเป็นหลังเกิดเหตุการณ์	
	ถูกต้อง	ไม่ถูกต้อง	คำตอบถูกต้อง	คำตอบไม่ถูกต้อง
1 1	0.54	0.04	0.931	0.069
1 0	0.36	0.06	0.857	0.143
0 1	0.06	0.36	0.143	0.857
0 0	0.04	0.54	0.069	0.931

การกำหนดขนาดตัวอย่างหรือจำนวนผู้เชี่ยวชาญที่เหมาะสมสำหรับแบบสอบที่ตรวจสอบโดยอาศัยทฤษฎีฉันทามติทางวัฒนธรรม พบว่า หากมีความสัมพันธ์ระหว่างระดับความสามารถในการเห็นพ้องร่วมกันของแต่ละบุคคลกับความเห็นพ้องต้องกันภายในกลุ่มสูง ความตรงของการตอบข้อคำถามจะสูงด้วยเช่นกัน ถ้าต้องการค่าความตรง 0.95 โดยมีค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างระดับความสามารถของแต่ละบุคคลกับความเห็นพ้องต้องกันภายในกลุ่มเท่ากับ 0.25 จะต้องกำหนดขนาดตัวอย่างอย่างน้อย 28 คน (Weller, 2007) ในกรณีที่ไม่มีทราบค่าสัมประสิทธิ์สหสัมพันธ์ดังกล่าว พิจารณาใช้เกณฑ์ขั้นต่ำของระดับความสามารถที่ 0.50 จะต้องกำหนดขนาดตัวอย่างอย่างน้อย 23 คน ดังตารางที่ 3 (Romney et al., 1986)

ตารางที่ 2.4 การกำหนดขนาดตัวอย่างและค่าความตรงในระดับความสามารถในการเห็นพ้องร่วมกันที่แตกต่างกัน

ความสามารถในการเห็นพ้องร่วมกัน	ขนาดตัวอย่างที่เหมาะสม จำแนกตามค่าความตรง (ระดับความเชื่อมั่น 0.99)				
	0.80	0.85	0.90	0.95	0.99
0.50	15	12	21	23	30
0.60	10	10	12	14	20
0.70	5	7	7	9	13
0.80	4	5	5	7	8
0.90	4	4	4	4	6

ทั้งนี้สามารถประยุกต์ใช้กับฉันทามติที่มีรูปแบบเลือกตอบหลายตัวเลือกได้ แต่จะต้องมีตัวเลือกที่ระบุคำตอบชัดเจน ไม่สามารถใช้วิเคราะห์ฉันทามติที่ตัดสินไม่ได้ เช่น ตัวเลือกไม่แน่ใจหรือไม่สามารถระบุได้ เนื่องจากตัวเลือกดังกล่าวไม่สามารถนำไปใช้ประมาณค่าความถูกต้องของฉันทามติได้ นอกจากนี้ แบบสอบที่ใช้การตรวจสอบโดยอาศัยทฤษฎีฉันทามติทางวัฒนธรรม ควรมีอัตราส่วนของข้อสอบที่เป็นคำถามเชิงบวกต่อคำถามเชิงลบประมาณ 70:30 หรือสัดส่วนของคำตอบใช่ต่อไม่ใช่ประมาณ 70:30 เช่นเดียวกัน (Romney et al., 1986) และจะต้องมีข้อคำถามอย่างน้อย 20 ข้อคำถามจึงจะประมาณค่าความสามารถได้ ลักษณะของข้อคำถามจะต้องมีคำตอบที่ถูกต้องเพียงคำตอบเดียวเท่านั้น (Weller, 2007) ในการตรวจสอบฉันทามติที่ถูกต้องสามารถวิเคราะห์ทางสถิติโดยใช้โปรแกรมสำเร็จรูปในการคำนวณระดับความสามารถในการตัดสินใจหรือความสามารถในการเห็นพ้องร่วมกันของผู้เชี่ยวชาญได้ นิยมใช้โปรแกรม UCINET ซึ่งถูกพัฒนาโดย S. P. Borgatti, Everett, and Freeman (1992) โดยมีขั้นตอนการวิเคราะห์ฉันทามติ แสดงรายละเอียดดังต่อไปนี้

ขั้นตอนที่ 1 การตรวจสอบข้อตกลงเบื้องต้นของทฤษฎีฉันทามติทางวัฒนธรรม แบ่งเป็น 3 ส่วน ดังนี้

1.1 ความเป็นเอกมิติของแบบวัด (Homogeneity) ผลการตรวจสอบข้อตกลงเบื้องต้นด้านความเป็นเอกมิติของแบบวัดระดับความสามารถในตัดสินใจหรือการเห็นพ้องร่วมกันโดยอาศัยทฤษฎีฉันทามติทางวัฒนธรรมแบบมีรูปแบบ โดยพิจารณาจากสัดส่วนของ Eigenvalue สูงสุดต่อลำดับที่ 2 จะต้องมีสัดส่วนมากกว่า 3 เท่าขึ้นไป

1.2 ความเป็นอิสระต่อกันของผู้ตอบ (Conditional independence) โดยกำหนดให้ผู้เชี่ยวชาญตอบคำถามแบบวัดระดับความสามารถในตัดสินใจหรือการเห็นพ้องร่วมกันโดยไม่สามารถสอบถามความคิดเห็นระหว่างกัน ไม่มีการปรึกษากันและไม่ใช้การสัมภาษณ์กลุ่มย่อย

1.3 ความถูกต้อง (Common truth) โดยกำหนดให้คำตอบของแบบวัดระดับความสามารถในตัดสินใจหรือการเห็นพ้องร่วมกันโดยมีเพียงคำตอบเดียวเท่านั้น

ขั้นตอนที่ 2 วิเคราะห์ระดับความสามารถในตัดสินใจหรือการเห็นพ้องร่วมกันโดยพิจารณาจากค่าน้ำหนักองค์ประกอบ (Factor loading)

ขั้นตอนที่ 3 วิเคราะห์คำตอบที่ถูกต้องของคำถามจากการลงมติโดยบุคคลผู้อยู่ในวัฒนธรรมเดียวกันโดยอาศัยสถิติเบย์ (Bayesian statistic) ที่ใช้หลักความน่าจะเป็นสูงสุด (Maximize probability) ซึ่งออกแบบโดย Romney et al. (1986)

ตัวอย่างงานวิจัยที่เกี่ยวข้อง

ทฤษฎีฉันทามติทางวัฒนธรรมถูกนำไปใช้อย่างแพร่หลายในงานวิจัยต่างๆ จากหลากหลายสาขาวิชา ซึ่งมักมีข้อค้นพบ หลักการประยุกต์ใช้ศาสตร์ต่างๆ และแนวทางการแก้ปัญหาจากการวิเคราะห์ฉันทามติที่มีประโยชน์แสดงรายละเอียดดังต่อไปนี้

การใช้ทฤษฎีฉันทามติทางวัฒนธรรมแบบมีรูปแบบในการตัดสินใจสำหรับนักเรียนที่ลงทะเบียนเรียนวิชาสังคมศาสตร์จำนวน 101 คน ซึ่งแบ่งเป็น 2 ตอนเรียนและอาจารย์แตกต่างกัน ในมหาวิทยาลัยเซาท์แคโรไลนา สหรัฐอเมริกา ซึ่งใช้แบบสอบถามหลายตัวเลือกจำนวน 58 ข้อ โดยแต่ละข้อมีคำตอบที่ถูกต้องเพียงคำตอบเดียว และตัดสินผลโดยใช้การประเมินผลแบบอิงเกณฑ์ 9 เกรด (A ถึง F) พบว่า การตรวจสอบข้อตกลงเบื้องต้นของทฤษฎีฉันทามติทางวัฒนธรรมพบว่า การตัดสินใจทั้ง 101 คนมีความเป็นเอกมิตี โดยพิจารณาจากสัดส่วนของ Eigenvalue สูงสุดต่อลำดับที่ 2 มากกว่า 3 เท่า (16.871) โดยที่แต่ละเกรดที่พิจารณาจากคะแนนสอบโดยอาจารย์มีความเป็นอิสระต่อกัน พบว่า การตัดสินใจของอาจารย์สอดคล้องกับการตัดสินใจโดยพิจารณาความสามารถในการตอบข้อสอบของนักเรียนที่ค่าความสัมพันธ์ของเพียร์สัน 0.937 โดยที่ระดับความสามารถในการตอบข้อสอบของนักเรียนอยู่ในช่วง 0.35 ถึง 0.89 (Maher, 1987) สอดคล้องกับงานวิจัยอื่นๆ ที่ประยุกต์ใช้ทฤษฎีฉันทามติทางวัฒนธรรมแบบมีรูปแบบและข้อมูลมีค่าสถิติเป็นไปตามข้อตกลงเบื้องต้นในด้านความเป็นเอกมิตีของแบบวัด อาทิ การศึกษาของ Chavez (1995) ซึ่งประยุกต์ใช้ในการสอบถามปัจจัยเสี่ยงต่อมะเร็งเต้านมและมะเร็งปอดลูกโดยเปรียบเทียบการรับรู้ระหว่างผู้ป่วยและแพทย์ที่ดูแลการรักษา พบว่า ทฤษฎีฉันทามติสามารถประยุกต์ใช้ในแบบสอบถามที่มีมาตราวัดรูปแบบอันดับ (Ordinal scale) ได้ไม่แตกต่างกับตัวแปรต่อเนื่องในงานวิจัยอื่นๆ นอกจากนี้ยังมีการประยุกต์ใช้ในงานวิจัยเพื่อจัดอันดับคุณภาพของทีมอเมริกันฟุตบอลทั้งหมด 15 ทีม โดยสำนักข่าวต่างๆ จำนวน 10 สำนักข่าวซึ่งพบว่ามีสัดส่วนของ Eigenvalue สูงสุดต่อลำดับที่ 2 มากกว่า 3 เท่า (10.490) และระดับความสามารถในการจัดอันดับของสำนักข่าวอยู่ในช่วง 0.53 ถึง

0.68 โดยสารสนเทศที่นำไปใช้สำหรับผู้มีส่วนได้ส่วนเสีย คือจัดอันดับเพื่อพัฒนาคุณภาพของทีมตั้งแต่ระดับที่ 1 ถึง 6 ตามคำตอบที่ได้จากการวิเคราะห์ฉันทามติ (S. P. Borgatti & Halgin, 2011)

แม้ว่างานวิจัยส่วนใหญ่ข้อมูลที่นำมาใช้วิเคราะห์จะเป็นไปตามข้อตกลงเบื้องต้นของทฤษฎีฉันทามติทางวัฒนธรรมแบบมีรูปแบบ แต่ S. P. Borgatti and Halgin (2011) ได้มีข้อเสนอแนะหากพบว่าแบบวัดหรือแบบสอบถามที่ใช้วัดระดับความสามารถในการเห็นพ้องร่วมกัน ไม่มีความเป็นเอกมิติ กล่าวคือมีสัดส่วนของ Eigenvalue สูงสุดต่อลำดับที่ 2 น้อยกว่า 3 เท่า อาจเกิดจากกลุ่มผู้ที่มีความสามารถในการเห็นพ้องร่วมกันในเรื่องนั้นๆค่อนข้างต่ำ (Low competence) และแตกต่างจากกลุ่มคนที่มีความฉันทามติร่วมกัน จึงควรพิจารณาคัดกลุ่มนั้นออกและทำการวิเคราะห์ฉันทามติอีกครั้ง สอดคล้องกับการวิเคราะห์ฉันทามติในงานวิจัยของ J. C. Johnson (1990); Krackhardt (1987) และ Weller (2007) ที่วิเคราะห์ฉันทามติหลังจากตัดกลุ่มที่มีระดับความเห็นพ้องร่วมกันต่ำออกและแสดงถึงความเป็นเอกมิติของกลุ่มที่อยู่ในวัฒนธรรมเดียวกันมากขึ้นอย่างชัดเจน นอกจากนี้ S. P. Borgatti and Halgin (2011) ยังได้ให้ข้อเสนอแนะเพิ่มเติมว่า การเปลี่ยนกลุ่มตัวอย่างข้ามวัฒนธรรม อาจส่งผลต่อฉันทามติที่ต้องการศึกษาได้ ซึ่งการวิเคราะห์ฉันทามติทางวัฒนธรรมแบบมีรูปแบบ จะต้องเก็บและวิเคราะห์ข้อมูลแยกกันเพราะมีความเสี่ยงที่จะละเมิดข้อตกลงเบื้องต้นในด้านเอกมิติได้ค่อนข้างมาก และคำตอบที่ได้จากการวิเคราะห์ฉันทามติอาจจะไม่ถูกต้องเสมอไปโดยต้องคำนึงถึงบริบททางสังคมหรือวัฒนธรรมนั้นๆ ไม่สามารถใช้ได้ทั่วไปเนื่องจากมีความจำเพาะเจาะจงต่อกลุ่มประชากรในวัฒนธรรมที่ศึกษาเท่านั้น การอ้างอิงจึงควรพิจารณาอย่างถี่ถ้วนก่อนนำสารสนเทศที่ได้รับไปใช้ แต่สามารถอ้างอิงถึงกลุ่มประชากรขนาดใหญ่ได้ถ้าวางแผนในการเก็บตัวอย่างที่เป็นตัวแทนของประชากรได้อย่างมีประสิทธิภาพ (Weller, 2007)

ตอนที่ 3 กรอบแนวคิดในการทำวิจัย

ในการวิจัยครั้งนี้เป็นการเปรียบเทียบดัชนีความเป็นคู่ขนานของข้อสอบและแบบสอบที่ได้จากการคัดเลือก ข้อสอบโดยกลุ่มผู้เชี่ยวชาญที่มีลักษณะแตกต่างกันโดยประยุกต์ใช้ทฤษฎีฉันทามติทางวัฒนธรรมแบบมีรูปแบบซึ่งแบ่งเป็น 2 ส่วน โดยส่วนแรกเป็นการประมาณค่าระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญโดยอาศัยโมเดลฉันทามติทางวัฒนธรรมแบบมีรูปแบบ โดยใช้แบบวัดระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบจำนวน 20 ข้อ เพื่อใช้ในการแบ่งกลุ่มผู้เชี่ยวชาญตามตัวแปรอิสระ 2 ตัวแปรได้แก่ 1. ระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญแบ่งเป็น 3 ระดับ (ระดับที่ 1 [0.80-1.00], ระดับที่ 2 [0.60-0.79] และระดับที่ 3 [0.40-0.59]) 2. ประสบการณ์สอนนักศึกษาแพทย์ (ตั้งแต่ 10 ปีขึ้นไปและน้อยกว่า 10 ปี) ซึ่งสามารถแบ่งกลุ่มผู้เชี่ยวชาญเป็น 6 กลุ่ม ได้แก่

กลุ่มที่ 1 มีความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบในระดับที่ 1 และประสบการณ์สอน ≥ 10 ปี

กลุ่มที่ 2 มีความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบในระดับที่ 1 และประสบการณ์สอน < 10 ปี

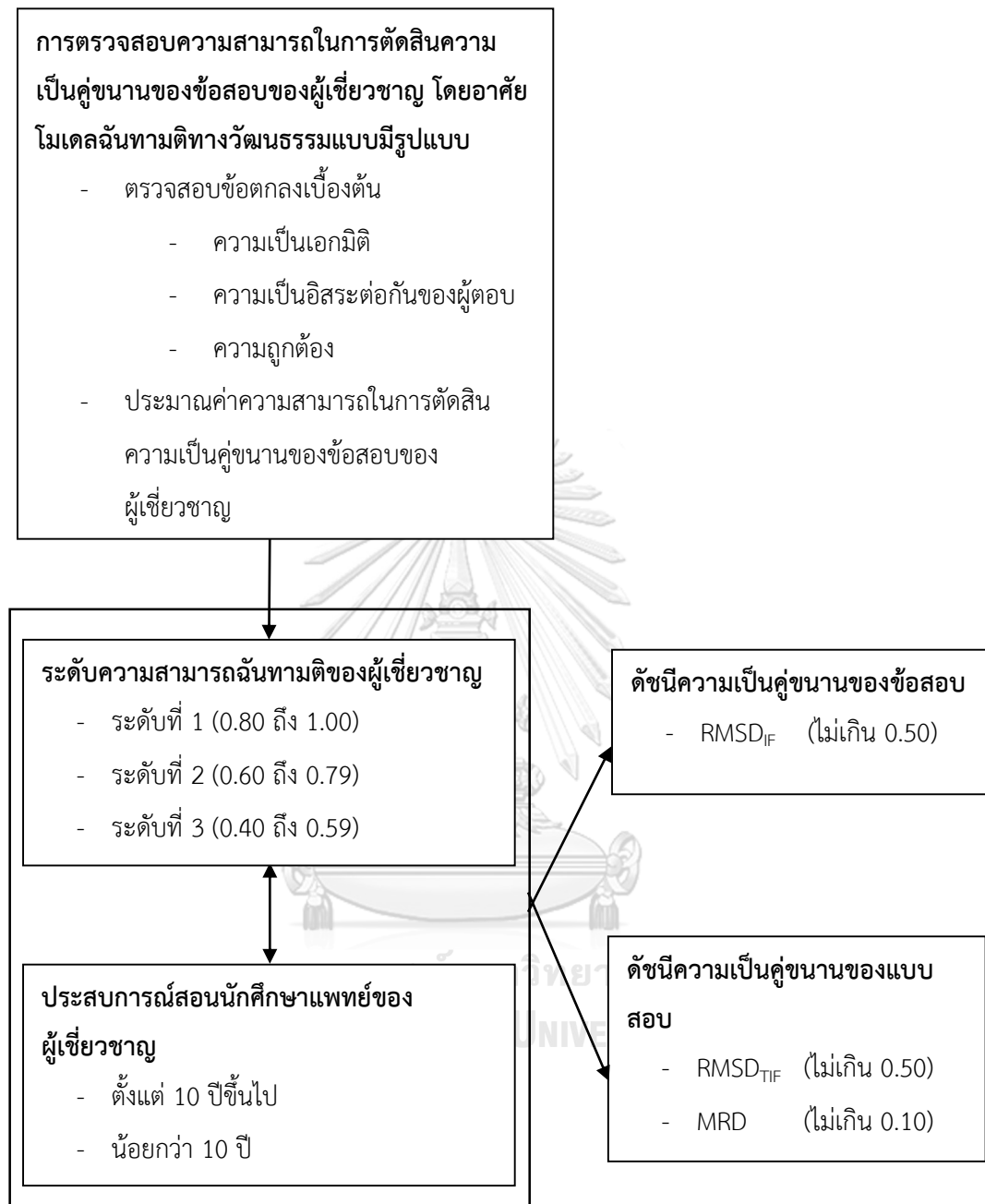
กลุ่มที่ 3 มีความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบในระดับที่ 2 และประสบการณ์สอน ≥ 10 ปี

กลุ่มที่ 4 มีความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบในระดับที่ 2 และประสบการณ์สอน < 10 ปี

กลุ่มที่ 5 มีความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบในระดับที่ 3 และประสบการณ์สอน ≥ 10 ปี

กลุ่มที่ 6 มีความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบในระดับที่ 3 และประสบการณ์สอน < 10 ปี

ตัวแปรตามคือความเป็นคู่ขนานของข้อสอบและแบบสอบ จัดเป็นการวิจัยในส่วนที่ 2 ผู้วิจัยเลือกใช้การตรวจสอบความเป็นคู่ขนานของข้อสอบและแบบสอบโดยอาศัยทฤษฎีการตอบสนองข้อสอบ (โมเดล 2 พารามิเตอร์; 2PL) ซึ่งพิจารณาดัชนีความเป็นคู่ขนานของข้อสอบโดยอาศัยค่า $RMSD_{IF}$ (ไม่เกิน 0.50) ส่วนดัชนีความเป็นคู่ขนานของแบบสอบพิจารณาจากค่า $RMSD_{TF}$ (ไม่เกิน 0.50) และ MRD (ไม่เกิน 0.10) จากนั้นจึงเปรียบเทียบความเป็นคู่ขนานของแบบสอบที่ได้มาจากการคัดเลือกข้อสอบโดยผู้เชี่ยวชาญที่แตกต่างกันทั้ง 6 กลุ่มโดยใช้สถิติทดสอบ one-way ANOVA เพื่อเปรียบเทียบค่าเฉลี่ยของ $RMSD_{TF}$ และ MRD ดังแสดงในภาพที่ 4



ภาพที่ 2.4 กรอบแนวคิดในการทำการวิจัย

บทที่ 3 วิธีดำเนินการวิจัย

การวิจัยในครั้งนี้เป็นการเปรียบเทียบดัชนีความเป็นคู่ขนานของแบบสอบที่ได้จากการคัดเลือกข้อสอบโดยกลุ่มผู้เชี่ยวชาญที่ระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบ และประสพการณ์สอนนักศึกษาแพทย์ที่แตกต่างกันโดยประยุกต์ใช้ทฤษฎีฉันทามติทางวัฒนธรรมแบบมีรูปแบบ โดยมีวัตถุประสงค์ 3 ประการ ได้แก่ 1) เพื่อศึกษาระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญ 2) เพื่อเปรียบเทียบดัชนีความเป็นคู่ขนานของแบบสอบคู่ขนานที่ได้จากการคัดเลือกโดยกลุ่มผู้เชี่ยวชาญที่มีระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบและประสพการณ์สอนที่แตกต่างกัน และ 3) เพื่อเปรียบเทียบดัชนีความเป็นคู่ขนานของข้อสอบคู่ขนานที่ได้จากการคัดเลือกโดยกลุ่มผู้เชี่ยวชาญที่มีระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบและประสพการณ์สอนที่แตกต่างกัน ซึ่งมีวิธีการดำเนินงานและเก็บรวบรวมข้อมูลแสดงรายละเอียดดังต่อไปนี้

ประชากรและตัวอย่าง

ประชากรในการวิจัยครั้งนี้ ได้แก่ ผู้เชี่ยวชาญในการสอนนักศึกษาแพทย์ สังกัดวิทยาลัยแพทยศาสตร์พระมงกุฎเกล้าที่มีคุณวุฒิการศึกษาสาขาเภสัชศาสตร์บัณฑิต แพทยศาสตรบัณฑิต หรือสาขาวิชาชีพทางสาธารณสุขอื่นๆ จำนวน 225 คน ตัวอย่างที่ใช้สำหรับการวิจัยครั้งนี้ได้มาจากการเลือกตัวอย่างแบบเจาะจงผู้เชี่ยวชาญที่มีประสพการณ์สอนเกี่ยวกับการเลือกใช้ยาสำหรับนักศึกษาแพทย์ที่แตกต่างกัน แสดงรายละเอียด ดังนี้

ขั้นตอนที่ 1 ศึกษาข้อมูลจำนวนผู้เชี่ยวชาญในการสอนนักศึกษาแพทย์ สังกัดวิทยาลัยแพทยศาสตร์พระมงกุฎเกล้าที่มีคุณวุฒิการศึกษาสาขาเภสัชศาสตร์บัณฑิต แพทยศาสตรบัณฑิต หรือสาขาวิชาชีพทางสาธารณสุขอื่นๆ พบว่ามีจำนวนประชากรจำนวน 225 คน จากนั้นกำหนดขนาดตัวอย่างผู้เชี่ยวชาญสำหรับตรวจสอบระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบโดยอาศัยทฤษฎีฉันทามติทางวัฒนธรรมแบบมีรูปแบบ พบว่า ในกรณีที่ไม่ทราบค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างระดับความสามารถของแต่ละบุคคลกับฉันทามติภายในกลุ่ม ให้พิจารณาใช้เกณฑ์ขั้นต่ำของระดับความสามารถในการตัดสินใจที่ 0.50 พบว่า จะต้องกำหนดขนาดตัวอย่างอย่างน้อย 23 คน (Romney et al., 1986) ดังแสดงในตารางที่ 4 โดยการวิจัยครั้งนี้ ใช้ตัวอย่างผู้เชี่ยวชาญจำนวน 100 คนจากประชากรทั้งหมด 225 คน เพื่อศึกษาระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบเพื่อให้มีจำนวนเพียงพอต่อการวิเคราะห์ฉันทามติ (Consensus analysis) สำหรับ

แบ่งระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญที่จะใช้ในขั้นตอนต่อไปและมุ่งเน้นให้แบบวัดมีคุณภาพต่อการตรวจสอบฉันทามติ ในด้านความตรงต่อการประมาณค่าความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบเป็นไปตามที่กำหนด

ตารางที่ 3.1 การกำหนดขนาดตัวอย่างและค่าความตรงในระดับความสามารถในการตัดสินใจที่แตกต่างกัน

ความสามารถในการตัดสินใจ	ขนาดตัวอย่างที่เหมาะสม จำแนกตามค่าความตรง (ระดับความเชื่อมั่น 0.99)				
	0.80	0.85	0.90	0.95	0.99
0.50	15	12	21	23	30
0.60	10	10	12	14	20
0.70	5	7	7	9	13
0.80	4	5	5	7	8
0.90	4	4	4	4	6

ขั้นตอนที่ 2 จากการตรวจสอบความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญจากขั้นตอนที่ 1 สามารถแบ่งกลุ่มผู้เชี่ยวชาญจำนวน 18 คนเป็น 6 กลุ่มๆละ 3 คน ตามระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบ โดยคัดเลือกอย่างเจาะจงจากผู้เชี่ยวชาญที่มีค่าความสามารถสูงสุด 3 อันดับแรกของแต่ละระดับ (ตารางที่ 3.1) และประสบการณ์สอนนักศึกษาแพทย์ ซึ่งจัดกลุ่มเป็น 6 กลุ่ม ดังต่อไปนี้

กลุ่มที่ 1 มีความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบในระดับที่ 1 และประสบการณ์สอน ≥ 10 ปี

กลุ่มที่ 2 มีความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบในระดับที่ 1 และประสบการณ์สอน < 10 ปี

กลุ่มที่ 3 มีความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบในระดับที่ 2 และประสบการณ์สอน ≥ 10 ปี

กลุ่มที่ 4 มีความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบในระดับที่ 2 และประสบการณ์สอน < 10 ปี

กลุ่มที่ 5 มีความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบในระดับที่ 3 และประสบการณ์สอน ≥ 10 ปี

กลุ่มที่ 6 มีความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบในระดับที่ 3 และ ประสิทธิภาพสอน < 10 ปี

ตารางที่ 3.2 การแบ่งระดับความสามารถความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญ

กลุ่ม	ระดับความสามารถในการตัดสิน ความเป็นคู่ขนานของข้อสอบ	ค่าความสามารถในการตัดสินความ เป็นคู่ขนานของข้อสอบ
1	ระดับที่ 1	0.80-1.00
2	ระดับที่ 2	0.60-0.79
3	ระดับที่ 3	0.40-0.59

กลุ่มผู้ให้ข้อมูล

นักศึกษาแพทย์วิทยาลัยแพทยศาสตร์พระมงกุฎเกล้า ชั้นปีที่ 3-6 ปีการศึกษา 2560 จำนวน 400 คน สำหรับทำแบบสอบคู่ขนานที่กลุ่มผู้เชี่ยวชาญสร้างขึ้นเพื่อใช้ในการหาดัชนีความเป็นคู่ขนานของแบบสอบและดัชนีความเป็นคู่ขนานของข้อสอบ โดยกลุ่มผู้ให้ข้อมูลได้มาจากการนักศึกษาแพทย์ทั้งหมด ดังตาราง 3.3

ตารางที่ 3.3 นักศึกษาที่ใช้เป็นกลุ่มผู้ให้ข้อมูลจำแนกชั้นปีและชุดข้อสอบที่สร้างขึ้นโดยกลุ่มผู้เชี่ยวชาญ

กลุ่มผู้เชี่ยวชาญ	แบบสอบที่สร้างขึ้น โดยกลุ่มผู้เชี่ยวชาญ	ชั้นปีของนักศึกษาแพทย์ที่เป็นกลุ่มผู้ให้ข้อมูล				รวม
		ชั้นปีที่ 3	ชั้นปีที่ 4	ชั้นปีที่ 5	ชั้นปีที่ 6	
1	ชุด 1A และ 1B	50	50	50	50	200
2	ชุด 2A และ 2B					
3	ชุด 3A และ 3B					
4	ชุด 4A และ 4B	50	50	50	50	200
5	ชุด 5A และ 5B					
6	ชุด 6A และ 6B					
รวม		100	100	100	100	400

การเก็บรวบรวมข้อมูล

การเก็บรวบรวมข้อมูลในการวิจัยในครั้งนี้ แบ่งตามวัตถุประสงค์ของการวิจัยเป็น 3 ตอน ดังนี้

ตอนที่ 1 เพื่อศึกษาระดับความสามารถในการตัดสินใจเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญ

1. ดำเนินการเก็บรวบรวมข้อมูลกับผู้เชี่ยวชาญในการสอนนักศึกษาแพทย์ สังกัดวิทยาลัยแพทยศาสตร์พระมงกุฎเกล้าที่มีคุณวุฒิการศึกษาสาขาเภสัชศาสตร์บัณฑิต แพทยศาสตรบัณฑิต หรือสาขาวิชาชีพทางสาธารณสุขอื่นๆ อย่างจำเพาะเจาะจง จำนวน 100 คน โดยใช้แบบวัดระดับความสามารถในการตัดสินใจเป็นคู่ขนานของข้อสอบจำนวน 20 ข้อเพื่อประมาณค่าความสามารถในการตัดสินใจเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญและวิเคราะห์ภูมิหลังของผู้เชี่ยวชาญ

2. ดำเนินการวิเคราะห์ทางสถิติ เพื่อหาระดับความสามารถในการตัดสินใจเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญ โดยอาศัยโมเดลฉันทามติทางวัฒนธรรมแบบมีรูปแบบ (Formal cultural consensus model) เพื่อระบุระดับความสอดคล้องของการตอบข้อคำถามของผู้เชี่ยวชาญแต่ละคนในกลุ่ม โดยใช้โปรแกรม UCINET

3. ตรวจสอบข้อตกลงเบื้องต้นของแบบวัดระดับความสามารถในการตัดสินใจเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญ ดังต่อไปนี้

3.1 ความเป็นเอกมิติของแบบวัด (Homogeneity) ผลการตรวจสอบข้อตกลงเบื้องต้นด้านความเป็นเอกมิติของแบบวัดระดับความสามารถในการตัดสินใจหรือการเห็นพ้องร่วมกันโดยอาศัยทฤษฎีฉันทามติทางวัฒนธรรมแบบมีรูปแบบ โดยพิจารณาจากสัดส่วนของ Eigenvalue สูงสุดต่อลำดับที่ 2 จะต้องมีส่วนมากกว่า 3 เท่าขึ้นไป ในกรณีที่สัดส่วนน้อยกว่า 3:1 แสดงถึงการขาดความเป็นเอกมิติ จะพิจารณาตัดผู้เชี่ยวชาญที่มีระดับความสามารถในการตัดสินใจเป็นคู่ขนานของข้อสอบที่ต่ำกว่า 0.40 ออก และทำการวิเคราะห์ฉันทามติเพื่อตรวจสอบข้อตกลงเบื้องต้นอีกครั้ง

3.2 ความเป็นอิสระต่อกันของผู้ตอบ (Conditional independence) โดยกำหนดให้ผู้เชี่ยวชาญตอบคำถามแบบวัดระดับความสามารถในการตัดสินใจหรือการเห็นพ้องร่วมกันโดยไม่สามารถสอบถามความคิดเห็นระหว่างกัน ไม่มีการปรึกษากันและไม่ใช้การสัมภาษณ์กลุ่มย่อย

3.3 ความถูกต้อง (Common truth) โดยกำหนดให้คำตอบของแบบวัดระดับความสามารถในการตัดสินใจหรือการเห็นพ้องร่วมกันโดยมีเพียงคำตอบเดียวเท่านั้น

4. แบ่งกลุ่มผู้เชี่ยวชาญตามระดับความสามารถในการตัดสินใจเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญออกเป็น 3 กลุ่ม ได้แก่ ระดับที่ 1 [0.80-1.00], ระดับที่ 2 [0.60-0.79] และระดับที่ 3 [0.40-0.59]

ในตอนนี้ การได้มาของกลุ่มผู้เชี่ยวชาญ สรุปเป็นขั้นตอน ดังนี้ ประชากรผู้เชี่ยวชาญ 225 คน เลือกตัวอย่างผู้เชี่ยวชาญ จำนวน 100 คน เพื่อวัดระดับความสามารถในการตัดสินใจตัดสินใจของข้อสอบ จากนั้นคัดเลือกผู้เชี่ยวชาญ 18 คน โดยจำแนกตามระดับความสามารถในการตัดสินใจตัดสินใจของข้อสอบและประสบการณ์สอนเป็น 6 กลุ่ม กลุ่มละ 3 คน เพื่อคัดเลือกข้อสอบเพื่อสร้างแบบสอบคู่ขนานในตอนต่อไป

ตอนที่ 2 เพื่อเปรียบเทียบดัชนีความเป็นคู่ขนานของแบบสอบคู่ขนานที่ได้จากการคัดเลือกโดยกลุ่มผู้เชี่ยวชาญที่มีระดับความสามารถในการตัดสินใจตัดสินใจของข้อสอบและประสบการณ์สอนที่แตกต่างกัน มีขั้นตอนดังนี้

1. คัดเลือกข้อสอบจากคลังข้อสอบภาควิชาเภสัชวิทยา สำหรับให้กลุ่มผู้เชี่ยวชาญคัดเลือกข้อสอบในการสร้างแบบสอบคู่ขนาน โดยเนื้อหาที่ใช้ในการคัดเลือกเพื่อสร้างแบบสอบคู่ขนานตามเกณฑ์ความรู้ความสามารถในการประเมินเพื่อรับใบอนุญาตเป็นผู้ประกอบวิชาชีพเวชกรรม พ.ศ. 2555 ประกอบด้วย 5 หมวด ได้แก่ หมวด A = ยาลดความดันโลหิต หมวด B = ยารักษาโรคหัวใจ หมวด C = ยารักษาแผลในกระเพาะอาหาร หมวด D = ยาที่เกี่ยวข้องกับการบีบตัวของลำไส้ และ หมวด E = ยารักษาโรคเบาหวาน โดยมีข้อสอบสำหรับคัดเลือกทั้งหมด 200 ข้อ จำแนกเป็นหมวดละ 40 ข้อ เพื่อให้ผู้เชี่ยวชาญคัดเลือกข้อสอบเพื่อสร้างแบบสอบคู่ขนาน หมวดละ 5 ข้อ รวมทั้งสิ้นแบบสอบละ 25 ข้อ

2. ดำเนินการเก็บรวบรวมข้อมูลกับผู้เชี่ยวชาญที่เป็นตัวอย่าง จำนวน 18 คน แบ่งออกเป็น 6 กลุ่มๆละ 3 คนตามระดับความสามารถในการตัดสินใจตัดสินใจของข้อสอบของผู้เชี่ยวชาญและประสบการณ์สอนนักศึกษาแพทย์ โดยให้กลุ่มผู้เชี่ยวชาญ คัดเลือกข้อสอบเพื่อสร้างแบบสอบคู่ขนานจำนวนกลุ่มละ 2 แบบสอบๆ ละ 25 ข้อ (จุดประสงค์ละ 5 ข้อ) ตามแผนผังการสอบและตารางวิเคราะห์ข้อสอบที่กำหนดให้ตามเกณฑ์ความรู้ความสามารถในการประเมินเพื่อรับใบอนุญาตเป็นผู้ประกอบวิชาชีพเวชกรรม พ.ศ. 2555

3. ดำเนินการเก็บรวบรวมข้อมูลกลุ่มผู้ให้ข้อมูล ได้แก่ นักศึกษาแพทย์วิทยาลัยแพทยศาสตร์พระมงกุฎเกล้า ชั้นปีที่ 3-6 ปีการศึกษา 2560 จำนวน 400 คน โดยทำการสอบแบบสอบคู่ขนานละ 200 คน โดยนักศึกษาแต่ละคนจะทำข้อสอบคนละ 3 ชุดแบบสอบ (ชุดที่ 1-2-3 หรือ ชุดที่ 4-5-6) โดยแบ่งนักศึกษาเป็น 2 กลุ่มกลุ่มละ 100 คน โดยกลุ่มที่ 1 ทำแบบสอบฉบับ A ก่อนและจึงทำแบบสอบฉบับ B ส่วนกลุ่มที่ 2 ทำแบบสอบฉบับ B ก่อนฉบับ A โดยกำหนดเวลาทำข้อสอบแบบสอบละ 30 นาที และทำแบบสอบทั้ง 3 ชุดต่อเนื่องกันเพื่อลดความคลาดเคลื่อนที่เกิดจากผู้สอบทราบข้อสอบและการสอบซ้ำ

4. ดำเนินการวิเคราะห์ทางสถิติโดยใช้โปรแกรม IRTpro เพื่อหาค่าสารสนเทศของแบบสอบ และพารามิเตอร์ของข้อสอบโดยอาศัยโมเดลการตอบสนองข้อสอบแบบ 2 พารามิเตอร์ และตรวจสอบความสอดคล้องกลมกลืนของข้อสอบกับโมเดล 2PL โดยอาศัยสถิติทดสอบไคสแควร์

5. ดำเนินการวิเคราะห์ทางสถิติโดยใช้โปรแกรม SPSS เพื่อเปรียบเทียบค่าดัชนีความเป็นคู่ขนานของแบบสอบที่สร้างขึ้น โดยพิจารณาค่า $RMSD_{TIF}$ และ MRD จากนั้นเปรียบเทียบค่าเฉลี่ยของ $RMSD_{TIF}$ และ MRD โดยใช้สถิติทดสอบ one-way ANOVA

ตอนที่ 3 เพื่อเปรียบเทียบดัชนีความเป็นคู่ขนานของข้อสอบคู่ขนานที่ได้จากการคัดเลือกโดยกลุ่มผู้เชี่ยวชาญที่มีระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบและประสบการณ์สอนที่แตกต่างกัน

1. ดำเนินการวิเคราะห์ทางสถิติโดยใช้โปรแกรม IRTpro เพื่อหาค่าสารสนเทศของข้อสอบ และพารามิเตอร์ของข้อสอบโดยอาศัยโมเดลการตอบสนองข้อสอบแบบ 2 พารามิเตอร์ คือ ค่าพารามิเตอร์ความยากของข้อสอบ (b) ค่าพารามิเตอร์อำนาจจำแนกของข้อสอบ (a)

2. ดำเนินการวิเคราะห์ทางสถิติโดยใช้โปรแกรม SPSS เพื่อเปรียบเทียบค่าดัชนีความเป็นคู่ขนานของข้อสอบที่สร้างขึ้น โดยพิจารณาค่า $RMSD_{IF}$ และ ค่าเฉลี่ย $RMSD_{IF}$

เครื่องมือที่ใช้

เครื่องมือที่ใช้สำหรับการวิจัยในครั้งนี้มี 3 ฉบับ ประกอบด้วย แบบวัดระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบ คลังข้อสอบสำหรับใช้สร้างแบบสอบคู่ขนาน และแบบบันทึกการคัดเลือกข้อสอบคู่ขนานโดยกลุ่มผู้เชี่ยวชาญ มีรายละเอียด ดังนี้

ฉบับที่ 1 แบบวัดระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบ แบ่งออกเป็น 2 ตอน

ตอนที่ 1 ข้อมูลภูมิหลัง ประกอบด้วย เพศ สังกัดภาควิชา ระดับการศึกษาสูงสุด สาขาที่จบการศึกษา การอบรมด้านแพทยศาสตรศึกษาและประสบการณ์สอนนักศึกษาแพทย์

ตอนที่ 2 แบบวัดระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญ จำนวน 20 ข้อ โดยให้ผู้เชี่ยวชาญพิจารณาความเป็นคู่ขนานของข้อสอบ ซึ่งมีสัดส่วนความน่าจะเป็นของคำตอบระหว่างข้อสอบคู่ขนานกันและไม่เป็นคู่ขนานกันเท่ากับ 14:6 ข้อ (อัตราส่วน 7:3) (Batchelder & Romney, 1988)

ตัวอย่าง แบบวัดระดับความสามารถในการตัดสินใจเป็นคู่ขนานของข้อสอบ

แบบวัดระดับความสามารถในการตัดสินใจเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญ

การศึกษาครั้งนี้เป็นส่วนหนึ่งของวิทยานิพนธ์ หัวข้อ การเปรียบเทียบดัชนีความเป็นคู่ขนานของข้อสอบและแบบสอบที่ได้จากการคัดเลือกข้อสอบโดยกลุ่มผู้เชี่ยวชาญที่มีลักษณะแตกต่างกัน: การประยุกต์ใช้ทฤษฎีฉันทามติทางวัฒนธรรมแบบมีรูปแบบ โดยมีจุดประสงค์เพื่อศึกษาระดับความสามารถในการเห็นพ้องร่วมกันต่อความเป็นคู่ขนานของแบบสอบของผู้เชี่ยวชาญและใช้เป็นข้อมูลในการจัดกลุ่มผู้เชี่ยวชาญเพื่อคัดเลือกข้อสอบสำหรับสร้างแบบสอบคู่ขนาน ผู้วิจัยจึงขออนุญาตสอบถามเกี่ยวกับชื่อ-สกุล เพื่อใช้ประสานงานในการศึกษาขั้นต่อไป ทั้งนี้ข้อมูลทั้งหมดจะเก็บเป็นความลับและไม่เผยแพร่ข้อมูลที่ระบุถึงบุคคลได้

ผู้เข้าร่วมวิจัย

ยินยอมให้ใช้ข้อมูลได้

ไม่ยินยอมให้ใช้ข้อมูล

ลงชื่อ..... ผู้เข้าร่วมวิจัย

ส่วนที่ 1 ข้อมูลภูมิหลัง

ชื่อ-สกุล.....

เพศ ชาย หญิง

ท่านเป็นอาจารย์ ชั้นปรีคลินิก ภาควิชา.....

ชั้นคลินิก ภาควิชา.....

ระดับการศึกษาสูงสุด ปริญญาตรี ปริญญาโท ปริญญาเอก

สาขาที่จบการศึกษา แพทยศาสตร์ เกษตรศาสตร์ อื่นๆ ระบุ.....

ประสบการณ์การสอนนักศึกษาแพทย์ ปี

ท่านเคยเข้ารับการอบรมด้านแพทยศาสตรศึกษาในช่วง 5 ปีที่ผ่านมาหรือไม่

เคย ไม่เคย

ชุดที่

ส่วนที่ 2 แบบฟอร์มการตัดสินความเป็นคู่ขนานของแบบสอบโดยผู้เชี่ยวชาญ

ประเด็นที่พิจารณา: ความเป็นคู่ขนานกันของข้อสอบแต่ละคู่ จำนวนทั้งหมด 20 คู่ ต่อไปนี้

- ถ้าข้อสอบคู่ใดมีความเป็นคู่ขนานกัน โปรดทำเครื่องหมาย ✓ ในช่อง ข้อสอบมีความเป็นคู่ขนานกัน

- ถ้าข้อสอบคู่ใดไม่มีความเป็นคู่ขนานกัน โปรดทำเครื่องหมาย ✓ ในช่อง ข้อสอบไม่มีความเป็นคู่ขนานกัน

***กรุณาตัดสินความเป็นคู่ขนานของข้อสอบทุกข้อ

ข้อสอบคู่ที่ 1	
ข้อ 1A	ข้อ 1B
หญิงอายุ 25 ปี มีอาการหอบเหนื่อยมา 3 วัน จากผลตรวจร่างกายและ CXR ได้รับการวินิจฉัยเป็น Community acquire pneumonia หากแพทย์ต้องการสั่งใช้ยา ceftriaxone ควรเขียนใบสั่งยาอย่างไร	ชายอายุ 25 ปี มีแผลหนองบริเวณเท้าซ้ายมา 3 วัน จากผลตรวจร่างกาย ได้รับการวินิจฉัยเป็น Infected wound at Lt. foot หากแพทย์ต้องการสั่งใช้ยา cefazolin ควรเขียนใบสั่งยาอย่างไร
<ol style="list-style-type: none"> 1. Ceftriaxone 1 g IV 2. Ceftriaxone 1 amp (V) 3. Ceftriaxone 1 amp IV 4. Ceftriaxone 1000 mg (V) 5. Ceftriaxone 1000 mg IV 	<ol style="list-style-type: none"> 1. Cefazolin 1 g IV 2. Cefazolin 1 amp (V) 3. Cefazolin 1 amp IV 4. Cefazolin 1000 mg (V) 5. Cefazolin 1000 mg IV
ผลการตัดสินความเป็นคู่ขนานของแบบสอบ	
<input type="checkbox"/> ข้อสอบมีความเป็นคู่ขนานกัน <input type="checkbox"/> ข้อสอบไม่มีความเป็นคู่ขนานกัน	

ฉบับที่ 2 คลังข้อสอบเพื่อใช้สร้างแบบสอบคู่ขนานโดยการคัดเลือกข้อสอบสำหรับกลุ่มผู้เชี่ยวชาญ

คลังข้อสอบภาควิชาเภสัชวิทยา สำหรับให้กลุ่มผู้เชี่ยวชาญคัดเลือกข้อสอบในการสร้างแบบสอบคู่ขนาน โดยเนื้อหาที่ใช้ในการคัดเลือกเพื่อสร้างแบบสอบคู่ขนานตามเกณฑ์ความรู้ความสามารถในการประเมินเพื่อรับใบอนุญาตเป็นผู้ประกอบวิชาชีพเวชกรรม พ.ศ. 2555 หมวดที่ 2.3 หลักการรักษและหลักการจ่าย โดยเฉพาะยาในบัญชียาหลักแห่งชาติ โดยใช้เนื้อหาใน 3 ระบบ ประกอบด้วย ยาในระบบหลอดเลือดและหัวใจ (B 7.3) ระบบทางเดินอาหาร (B 9.3) และระบบต่อมไร้ท่อ (B 11.3) รวมข้อสอบทั้งหมดประกอบด้วย แบบสอบละ 25 ข้อ แบ่งเป็น 5 จุดประสงค์ๆละ 5 ข้อ โดยคัดเลือกจากข้อสอบในแต่ละจุดประสงค์ซึ่งประกอบด้วยข้อสอบจำนวนทั้งหมด 40 ข้อ (ดังตารางที่ 3.4)

ตารางที่ 3.4 ลักษณะของข้อสอบจำแนกระบบตามเกณฑ์ความรู้ความสามารถในการประเมินเพื่อรับใบอนุญาต เป็นผู้ประกอบวิชาชีพเวชกรรม พ.ศ. 2555

จุดประสงค์	ระบบ	จำนวน (ข้อ)
1	ยาในระบบหลอดเลือดและหัวใจ (ยาลดความดันโลหิต)	40
2	ยาในระบบหลอดเลือดและหัวใจ (ยารักษาโรคหัวใจ)	40
3	ยาในระบบทางเดินอาหาร (ยารักษาแผลในทางเดินอาหาร)	40
4	ยาในระบบทางเดินอาหาร (ยาที่เกี่ยวข้องกับบีบตัวของลำไส้)	40
5	ยาในระบบต่อมไร้ท่อ (ยารักษาโรคเบาหวาน)	40
	รวม	200

ฉบับที่ 3 แบบบันทึกการคัดเลือกข้อสอบคู่ขนานโดยกลุ่มผู้เชี่ยวชาญ แบ่งเป็น 2 ตอน

ตอนที่ 1 ข้อมูลระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบและประสบการณ์สอนของกลุ่มผู้เชี่ยวชาญที่คัดเลือกอย่างเจาะจงเพื่อให้มีคุณสมบัติตามที่กำหนด

ตอนที่ 2 แบบบันทึกการคัดเลือกข้อสอบคู่ขนานโดยกลุ่มผู้เชี่ยวชาญ ประกอบด้วยข้อสอบที่ต้องคัดเลือก จำนวน 25 คู่ โดยแบ่งเป็นชุด A และชุด B สำหรับผู้เชี่ยวชาญบันทึกเมื่อพิจารณาข้อสอบร่วมกันภายในกลุ่มที่กำหนดไว้ ดังแสดงตัวอย่างเครื่องมือในภาคผนวก ก

การวิเคราะห์ข้อมูล

จากการเก็บรวบรวมข้อมูลด้วยเครื่องมือทั้ง 2 ฉบับในการวิจัย สามารถแบ่งการดำเนินการวิเคราะห์ข้อมูลตามวัตถุประสงค์ของการวิจัย โดยมีรายละเอียด ดังนี้

1. เพื่อศึกษาระดับความสามารถในการตัดสินใจเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญ

1.1 การวิเคราะห์ค่าสถิติพื้นฐานของตัวแปรภูมิหลังของผู้เชี่ยวชาญ โดยใช้สถิติบรรยาย ได้แก่ ความถี่ ร้อยละ โดยใช้โปรแกรมวิเคราะห์สถิติสำเร็จรูป SPSS เพื่อตารางแบบไขว้ (Crosstab table)

1.2 ดำเนินการวิเคราะห์ทางสถิติ เพื่อประมาณค่าความสามารถในการตัดสินใจเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญ โดยอาศัยโมเดลฉันทามติทางวัฒนธรรมแบบมีรูปแบบ (Formal cultural consensus model) เพื่อระบุระดับความสอดคล้องของการตอบข้อคำถามของผู้เชี่ยวชาญแต่ละคนในกลุ่ม โดยใช้โปรแกรม UCINET ทั้งนี้ การประมาณค่าความสามารถในการตัดสินใจเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญสามารถพิจารณาจากค่าน้ำหนักองค์ประกอบ (Factor loading) ที่ได้จากการวิเคราะห์ฉันทามติ ทั้งนี้ต้องผ่านการตรวจสอบข้อตกลงเบื้องต้นด้านความเป็นเอกมิติโดยพิจารณาจากการเปรียบเทียบสัดส่วนของค่าไอเกนลำดับที่ 1 และไอเกนลำดับที่ 2 จะต้องมากกว่า 3:1

1.3 ตรวจสอบความถูกต้องของคำตอบต่อข้อคำถามโดยประยุกต์ใช้ทฤษฎีของเบย์ (Bayes' Theorem) มาหาค่าความน่าจะเป็นของการตอบคำถามถูกต้อง (Posteriori probability) โดยอาศัยความน่าจะเป็นก่อนเกิดเหตุการณ์ (Priori probability) และข้อมูลการตอบข้อคำถามของผู้เชี่ยวชาญ (Evidence of the informants' response) โดยใช้การวิเคราะห์ฉันทามติผ่านโปรแกรม UCINET

2. เพื่อเปรียบเทียบดัชนีความเป็นคู่ขนานของแบบสอบคู่ขนานที่ได้จากการคัดเลือกโดยกลุ่มผู้เชี่ยวชาญที่มีระดับความสามารถในการตัดสินใจเป็นคู่ขนานของข้อสอบและประสบการณ์สอนที่แตกต่างกัน

2.1 ดำเนินการวิเคราะห์ทางสถิติโดยใช้โปรแกรม IRTpro เพื่อหาค่าสารสนเทศของแบบสอบและพารามิเตอร์ของข้อสอบโดยอาศัยโมเดลการตอบสนองข้อสอบแบบ 2 พารามิเตอร์ และตรวจสอบความสอดคล้องกลมกลืนของข้อสอบกับโมเดล 2PL โดยอาศัยสถิติทดสอบไคสแควร์

2.2 ดำเนินการวิเคราะห์ทางสถิติโดยใช้โปรแกรม SPSS เพื่อเปรียบเทียบค่าดัชนีความเป็นคู่ขนานของแบบสอบที่สร้างขึ้น โดยอาศัยทฤษฎีการตอบสนองข้อสอบซึ่งพิจารณาสถิติ ดังต่อไปนี้

2.2.1 ดัชนีความเป็นคู่ขนานของแบบสอบโดยอาศัยค่า $RMSD_{TF}$ หมายถึง ค่าสถิติที่บ่งชี้ว่าแบบสอบคู่ขนาน 2 ฉบับมีส่วนเบี่ยงเบนค่าเฉลี่ยกำลังสองของสารสนเทศของแบบสอบ (Root

mean square deviation of test information; $RMSD_{TIF}$) สำหรับตรวจสอบการรูปแบบและการกระจายของฟังก์ชันสารสนเทศของแบบสอบ 2 ฉบับที่วัดโครงสร้างทางเนื้อหาเดียวกันในระดับความสามารถของผู้ตอบข้อสอบทุกระดับแตกต่างกันไม่เกิน 0.50 ซึ่งคำนวณได้จากสมการ ดังต่อไปนี้

$$RMSD_{TIF} = \sqrt{\frac{\sum_{j=1}^q [I(\theta_j) - I'(\theta_j)]^2}{q}}$$

เมื่อ q = จำนวนระดับความสามารถที่ใช้ในการตรวจสอบฟังก์ชัน
 $I(\theta)$ = ค่าสารสนเทศของแบบสอบที่ระดับความสามารถของผู้สอบ θ

ทั้งนี้ การพิจารณาดัชนีความเป็นคู่ขนานของแบบสอบโดยอาศัยค่า $RMSD_{TIF}$ ควรพิจารณา ร่วมกับค่าผลต่างระหว่างสารสนเทศของแบบสอบสูงสุดที่ระดับความสามารถของผู้สอบเดียวกัน (Maximum score gap; MSG) โดยค่า MSG ของแบบสอบคู่ขนานจะไม่เกิน 0.50 ซึ่งคำนวณได้จากสมการ ดังต่อไปนี้

$$MSG = \max(\theta) \{ \max I(\theta) - \min I'(\theta) \}$$

เมื่อ $I(\theta)$ = ค่าสารสนเทศของแบบสอบที่ระดับความสามารถของผู้สอบ θ

2.2.2 ดัชนีความเป็นคู่ขนานของแบบสอบโดยอาศัยค่า MRD หมายถึง ค่าสถิติที่บ่งชี้ว่าแบบสอบคู่ขนาน 2 ฉบับมีค่าเฉลี่ยสัดส่วนที่แตกต่างกันของสารสนเทศของแบบสอบ (MRD) สำหรับตรวจสอบการรูปแบบและการกระจายของฟังก์ชันสารสนเทศของแบบสอบ 2 ฉบับที่วัดโครงสร้างทางเนื้อหาเดียวกันในระดับความสามารถของผู้ตอบข้อสอบทุกระดับแตกต่างกันไม่เกิน 0.10 ซึ่งคำนวณได้จากสมการ ดังต่อไปนี้

$$MRD = \frac{\sum_{j=1}^q \left| \frac{I(\theta_j) - I'(\theta_j)}{I'(\theta_j)} \right|}{q}$$

เมื่อ q = จำนวนระดับความสามารถที่ใช้ในการตรวจสอบฟังก์ชัน
 $I(\theta)$ = ค่าสารสนเทศของแบบสอบที่ระดับความสามารถของผู้สอบ θ

ทั้งนี้ การพิจารณาดัชนีความเป็นคู่ขนานของแบบสอบโดยอาศัยค่า MRD ควรพิจารณาร่วมกับค่าผลต่างสัมพัทธ์ระหว่างสารสนเทศของแบบสอบสูงสุดที่ระดับความสามารถของผู้สอบเดียวกัน (Maximum relative information gap; MRIG) โดยค่า MRIG ของแบบสอบคู่ขนานจะไม่เกิน 0.10 ซึ่งคำนวณได้จากสมการ ดังต่อไปนี้

$$MRIG = \max(\theta) \left\{ \frac{\max I(\theta) - \min I'(\theta)}{I'(\theta)} \right\}$$

เมื่อ $I(\theta)$ = ค่าสารสนเทศของแบบสอบที่ระดับความสามารถของผู้สอบ θ

2.3 เปรียบเทียบค่าเฉลี่ยของ $RMSD_{TF}$ และ MRD โดยใช้สถิติทดสอบ one-way ANOVA

3. เพื่อเปรียบเทียบดัชนีความเป็นคู่ขนานของข้อสอบคู่ขนานที่ได้จากการคัดเลือกโดยกลุ่มผู้เชี่ยวชาญที่มีระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบและประสบการณ์สอนที่แตกต่างกัน

3.1 ดำเนินการวิเคราะห์ทางสถิติโดยใช้โปรแกรม IRTpro เพื่อหาค่าสารสนเทศของข้อสอบและพารามิเตอร์ของข้อสอบโดยอาศัยโมเดลการตอบสนองข้อสอบแบบ 2 พารามิเตอร์ คือ ค่าพารามิเตอร์ความยากของข้อสอบ (b) ค่าพารามิเตอร์อำนาจจำแนกของข้อสอบ (a)

3.2 ดำเนินการวิเคราะห์ทางสถิติโดยใช้โปรแกรม SPSS เพื่อเปรียบเทียบค่าดัชนีความเป็นคู่ขนานของข้อสอบที่สร้างขึ้น โดยพิจารณาค่า $RMSD_{IF}$ และ ค่าเฉลี่ย $RMSD_{IF}$ มีรายละเอียดดังนี้

ดัชนีความเป็นคู่ขนานของข้อสอบรายข้อโดยอาศัยค่า Root mean square deviation of item information ($RMSD_{TF}$) ซึ่งเป็นค่าสถิติที่บ่งชี้ว่าข้อสอบคู่ขนาน 2 ข้อมีส่วนเบี่ยงเบนค่าเฉลี่ยกำลังสองของสารสนเทศของข้อสอบ สำหรับตรวจสอบการรูปแบบและการกระจายของฟังก์ชันสารสนเทศของข้อสอบ 2 ข้อที่วัดโครงสร้างทางเนื้อหาเดียวกันในระดับความสามารถของผู้ตอบข้อสอบทุกระดับ โดยกำหนดให้ข้อสอบคู่ขนานจะต้องมีค่า $RMSD_{IF}$ แตกต่างกันไม่เกิน 0.50 ทั้งนี้สามารถคำนวณได้จากสมการดังต่อไปนี้

$$RMSD_{IF} = \sqrt{\frac{\sum_{j=1}^q [i(\theta_j) - i'(\theta_j)]^2}{q}}$$

เมื่อ q = จำนวนระดับความสามารถที่ใช้ในการตรวจสอบฟังก์ชัน

$i(\theta)$ = ค่าสารสนเทศของข้อสอบที่ระดับความสามารถของผู้สอบ θ

บทที่ 4

ผลการวิเคราะห์ข้อมูล

ผลการดำเนินการวิจัยเพื่อเปรียบเทียบความเป็นคู่ขนานของข้อสอบและแบบสอบที่ได้มาจากการคัดเลือกข้อสอบโดยกลุ่มผู้เชี่ยวชาญที่มีลักษณะแตกต่างกัน แบ่งการนำเสนอการวิเคราะห์ข้อมูลตามวัตถุประสงค์ออกเป็น 3 ตอน ได้แก่ ตอนที่ 1 เพื่อศึกษาระดับความสามารถในการตัดสินใจเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญ ตอนที่ 2 เพื่อเปรียบเทียบดัชนีความเป็นคู่ขนานของแบบสอบคู่ขนานที่ได้จากการคัดเลือกโดยกลุ่มผู้เชี่ยวชาญที่มีระดับความสามารถในการตัดสินใจของข้อสอบและประสบการณ์สอนนักศึกษาแพทย์ที่แตกต่างกันและตอนที่ 3 เพื่อเปรียบเทียบดัชนีความเป็นคู่ขนานของข้อสอบคู่ขนานที่คัดเลือกโดยกลุ่มผู้เชี่ยวชาญที่มีระดับความสามารถในการตัดสินใจเป็นคู่ขนานของข้อสอบและประสบการณ์สอนนักศึกษาแพทย์ที่แตกต่างกัน ซึ่งในแต่ละตอนมีรายละเอียดดังต่อไปนี้

ตอนที่ 1 เพื่อศึกษาระดับความสามารถในการตัดสินใจเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญ

การตัดสินใจเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญมีความสำคัญในการคัดเลือกข้อสอบเพื่อสร้างแบบสอบคู่ขนาน เนื่องจากความแตกต่างของระดับความสามารถในการตัดสินใจของผู้เชี่ยวชาญส่งผลต่อประสิทธิภาพในการคัดเลือกข้อสอบเพื่อสร้างแบบสอบคู่ขนาน ทั้งคุณภาพในด้านความตรงและความเที่ยงของแบบสอบ รวมถึงทำให้เกิดความเหลื่อมล้ำของการตัดสินใจในกลุ่มผู้เชี่ยวชาญได้ ดังนั้น การวิเคราะห์ฉันทามติเพื่อวัดระดับความสามารถในการตัดสินใจเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญจะให้สารสนเทศในคัดเลือกเพื่อจัดกลุ่มผู้เชี่ยวชาญสำหรับการคัดเลือกข้อสอบได้อย่างเหมาะสม ผู้วิจัยจึงแบ่งการวิเคราะห์ข้อมูลออกเป็น 4 ส่วน ดังนี้ ส่วนที่ 1 ข้อมูลภูมิหลังของผู้เชี่ยวชาญที่ทำแบบวัดระดับความสามารถในการตัดสินใจเป็นคู่ขนานของข้อสอบส่วนที่ 2 ผลการตรวจสอบระดับความสามารถในการตัดสินใจเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญโดยอาศัยทฤษฎีฉันทามติทางวัฒนธรรมแบบมีรูปแบบ ส่วนที่ 3 ผลการแบ่งกลุ่มผู้เชี่ยวชาญจำแนกตามระดับความสามารถในการตัดสินใจเป็นคู่ขนานของข้อสอบและประสบการณ์สอน และส่วนที่ 4 ผลการตัดสินใจคำตอบที่ถูกต้องจากการวิเคราะห์ฉันทามติจากการตัดสินใจเป็นคู่ขนานของข้อสอบโดยผู้เชี่ยวชาญ แสดงรายละเอียดดังต่อไปนี้

ส่วนที่ 1 ข้อมูลภูมิหลังของผู้เชี่ยวชาญที่ทำแบบวัดระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบ

จากการเก็บรวบรวมข้อมูลโดยให้ผู้เชี่ยวชาญทำแบบวัดระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบซึ่งมีผู้เชี่ยวชาญเข้าร่วมทำแบบวัดจำนวน 63 คนจากผู้เชี่ยวชาญที่เป็นอาจารย์สังกัดวิทยาลัยแพทยศาสตร์พระมงกุฎเกล้าที่เป็นตัวอย่างทั้งหมด 100 คน คิดเป็นร้อยละ 63 ข้อมูลภูมิหลังของผู้เชี่ยวชาญที่ทำแบบวัดระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบ พบว่า ผู้เชี่ยวชาญส่วนใหญ่เป็นเพศหญิง (ร้อยละ 53.968) มากกว่าเพศชาย (ร้อยละ 46.032) เมื่อจำแนกตามประสบการณ์สอนจะเป็นเพศหญิงมากกว่าเพศชายในกลุ่มประสบการณ์สอนตั้งแต่ 10 ปีขึ้นไป เป็นเพศหญิงร้อยละ 20.635 และเพศชายร้อยละ 12.698 แต่มีจำนวนเพศชายและหญิงเท่ากันในกลุ่มประสบการณ์สอนน้อยกว่า 10 ปี (ร้อยละ 33.33)

ผู้เชี่ยวชาญส่วนใหญ่สังกัดภาควิชาคลินิก (ร้อยละ 52.381) มากกว่าปรีคลินิก (ร้อยละ 47.619) เมื่อจำแนกตามประสบการณ์สอนจะสังกัดภาควิชาคลินิกมากกว่าปรีคลินิกในกลุ่มประสบการณ์สอนตั้งแต่ 10 ปีขึ้นไป สังกัดภาควิชาคลินิกร้อยละ 19.048 และปรีคลินิกร้อยละ 12.698 แต่มีสังกัดภาควิชาคลินิกและปรีคลินิกเท่ากันในกลุ่มประสบการณ์สอนน้อยกว่า 10 ปี (ร้อยละ 33.33)

ผู้เชี่ยวชาญส่วนใหญ่มีระดับการศึกษาสูงสุดในระดับปริญญาเอก (ร้อยละ 49.206) รองลงมาคือ ปริญญาโท (ร้อยละ 44.444) และปริญญาตรี (ร้อยละ 6.349) ตามลำดับ พบว่า ผู้เชี่ยวชาญส่วนใหญ่ในกลุ่มประสบการณ์สอนตั้งแต่ 10 ปีขึ้นไป มีระดับการศึกษาสูงสุดในระดับปริญญาเอก (ร้อยละ 20.635) แตกต่างจากกลุ่มประสบการณ์สอนน้อยกว่า 10 ปี ที่ส่วนใหญ่มีระดับการศึกษาสูงสุดในระดับปริญญาโท (ร้อยละ 31.746) โดยสาขาที่จบการศึกษาส่วนใหญ่ผู้เชี่ยวชาญจบการศึกษาในสาขาแพทยศาสตร์ (ร้อยละ 61.905) รองลงมา คือสาขาเภสัชศาสตร์ (ร้อยละ 20.635) และ สาขาอื่นๆ (ร้อยละ 17.460) ตามลำดับ

ในภาพรวมของการเข้าอบรมด้านแพทยศาสตรศึกษาของผู้เชี่ยวชาญใน 5 ปีที่ผ่านมา ผู้เชี่ยวชาญส่วนใหญ่เข้าอบรมร้อยละ 73.016 สอดคล้องกันทั้งกลุ่มประสบการณ์สอนตั้งแต่ 10 ปี ขึ้นไป (ร้อยละ 23.810) และน้อยกว่า 10 ปี (ร้อยละ 49.206) แสดงรายละเอียดดังตาราง 4.1

ตารางที่ 4.1 ข้อมูลภูมิหลังของผู้เชี่ยวชาญที่ทำแบบวัดระดับความสามารถในการตัดสินใจ
เป็นคู่ขนานของข้อสอบ

รายละเอียด	ประสบการณ์สอน		รวม
	ประสบการณ์สอน \geq 10 ปี จำนวน (ร้อยละ)	ประสบการณ์สอน $<$ 10 ปี จำนวน (ร้อยละ)	
เพศ			
ชาย	8 (12.698)	21 (33.333)	29 (46.032)
หญิง	13 (20.635)	21 (33.333)	34 (53.968)
รวม	21(33.333)	42 (66.667)	63 (100.000)
สังกัดภาควิชา			
ปริคณีก	9 (14.286)	21 (33.333)	30 (47.619)
คลินิก	12 (19.048)	21 (33.333)	33 (52.381)
รวม	21 (33.333)	42 (66.667)	63 (100.000)
ระดับการศึกษาสูงสุด			
ปริญญาตรี	0 (0.000)	4 (6.349)	4 (6.349)
ปริญญาโท	8 (12.698)	20 (31.746)	28 (44.444)
ปริญญาเอก	13 (20.635)	18 (28.571)	31 (49.206)
รวม	21 (33.333)	42 (66.667)	63 (100.000)
สาขาที่จบการศึกษา			
แพทยศาสตร์	14 (22.222)	25 (39.683)	39 (61.905)
เภสัชศาสตร์	2 (3.175)	11 (17.460)	13 (20.635)
สาขาอื่นๆ	5 (7.937)	6 (9.524)	11 (17.460)
รวม	21 (33.333)	42 (66.667)	63 (100.000)
การอบรมด้านแพทยศาสตรศึกษา			
เคยเข้าอบรมใน 5 ปีที่ผ่านมา	15 (23.810)	31 (49.206)	46 (73.016)
ไม่เคยเข้าอบรมใน 5 ปีที่ผ่านมา	6 (9.524)	11 (17.460)	17 (26.984)
รวม	21 (33.333)	42 (66.667)	63 (100.000)

ส่วนที่ 2 ผลการตรวจสอบระดับความสามารถในตัดสินความเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญโดยอาศัยทฤษฎีฉันทามติทางวัฒนธรรมแบบมีรูปแบบ

การวิจัยในครั้งนี้ประยุกต์ใช้ทฤษฎีฉันทามติทางวัฒนธรรมแบบมีรูปแบบ ในการตรวจสอบระดับความสามารถในตัดสินความเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญ โดยใช้แบบวัดระดับความสามารถในตัดสินความเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญ จำนวน 20 ข้อ ทั้งนี้ก่อนการวิเคราะห์ฉันทามติ (Consensus analysis) จะต้องมีการตรวจสอบข้อตกลงเบื้องต้นแล้วจึงตัดสินค่าความสามารถในตัดสินความเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญ แสดงรายละเอียดดังต่อไปนี้

2.1 ผลการตรวจสอบข้อตกลงเบื้องต้นของแบบวัดระดับความสามารถในตัดสินความเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญโดยอาศัยทฤษฎีฉันทามติทางวัฒนธรรมแบบมีรูปแบบ แบ่งออกเป็น 3 ด้าน ดังนี้

2.1.1 ความเป็นเอกมิติของแบบวัด (Homogeneity) ผลการตรวจสอบข้อตกลงเบื้องต้นด้านความเป็นเอกมิติของแบบวัดระดับความสามารถในตัดสินความเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญโดยอาศัยทฤษฎีฉันทามติทางวัฒนธรรมแบบมีรูปแบบ ทำการวิเคราะห์โดยใช้โปรแกรม UCINET เพื่อวิเคราะห์องค์ประกอบของการลงฉันทามติเกี่ยวกับการตัดสินความเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญโดยพิจารณาจากสัดส่วนของ Eigenvalue สูงสุดต่อลำดับที่ 2 พบว่า มีสัดส่วนน้อยกว่า 3 เท่า ดังนั้น แบบวัดระดับความสามารถในตัดสินความเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญขาดความเป็นเอกมิติของแบบวัดและไม่เป็นไปตามข้อตกลงเบื้องต้นของทฤษฎีฉันทามติทางวัฒนธรรมแบบมีรูปแบบ แสดงรายละเอียดดังตารางที่ 4.2

2.1.2 ความเป็นอิสระต่อกันของผู้ตอบ (Conditional independence) โดยกำหนดให้ผู้เชี่ยวชาญตอบคำถามแบบวัดระดับความสามารถในตัดสินความเป็นคู่ขนานของข้อสอบโดยไม่สามารถสอบถามความคิดเห็นระหว่างกัน ดังนั้น แบบวัดนี้จึงมีความเป็นอิสระต่อกันของผู้ตอบตามข้อตกลงเบื้องต้นของทฤษฎีฉันทามติทางวัฒนธรรมแบบมีรูปแบบ

2.1.3 ความถูกต้อง (Common truth) โดยกำหนดให้คำตอบของแบบวัดระดับความสามารถในตัดสินความเป็นคู่ขนานของข้อสอบมีเพียงคำตอบเดียว คือ ความเป็นคู่ขนานกันของข้อสอบหรือไม่มีความเป็นคู่ขนานกันของข้อสอบ ดังนั้น แบบวัดนี้จึงมีคุณสมบัติด้านความถูกต้องตามข้อตกลงเบื้องต้นของทฤษฎีฉันทามติทางวัฒนธรรมแบบมีรูปแบบ

ตารางที่ 4.2 ผลการแบ่งกลุ่มผู้เชี่ยวชาญจำแนกตามระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบ

ค่าสถิติ	ผลการวิเคราะห์
จำนวนผู้เชี่ยวชาญ	63 คน
Eigenvalue สูงสุด	20.109
Eigenvalue ลำดับที่ 2	9.111
สัดส่วนของ Eigenvalue สูงสุดต่อลำดับที่ 2	2.207

2.2 ผลการตรวจสอบระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญ

จากการตรวจสอบข้อตกลงเบื้องต้นพบว่า ไม่เป็นไปตามข้อตกลงเบื้องต้นด้านความเป็นเอกมิติ ซึ่งเกิดจากความแตกต่างของฉันทามติของผู้เชี่ยวชาญหรือขาดความสอดคล้องกันของฉันทามติทำให้ผู้เชี่ยวชาญมีลักษณะของความเป็นอันหนึ่งเดียวกันน้อย โดยจะมีผลต่อการตรวจสอบระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบได้ จึงต้องพิจารณาระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญโดยการวิเคราะห์ฉันทามติ โดยกำหนดเกณฑ์สำหรับจำแนกผู้เชี่ยวชาญตามระดับความสามารถในการคัดเลือกข้อสอบคู่ขนานได้เป็น 2 กลุ่ม ได้แก่

2.2.1 กลุ่มที่มีความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบผ่านเกณฑ์ (ตั้งแต่ 0.40 ถึง 1.00)

2.2.2 กลุ่มที่มีความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบไม่ผ่านเกณฑ์ (ตั้งแต่ 0.00 ถึง 0.39) แสดงรายละเอียดดังตารางที่ 4.3

ตารางที่ 4.3 ผลการแบ่งกลุ่มผู้เชี่ยวชาญจำแนกตามระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบ

ระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบ	ช่วงคะแนนความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบ	การแปลผล	จำนวน	ร้อยละ
ผ่านเกณฑ์	0.40-1.00	ผ่านเกณฑ์	43	68.25
ไม่ผ่านเกณฑ์	0.00-0.39	ไม่ผ่านเกณฑ์	20	31.75
รวม			63	100

เมื่อพิจารณาข้อตกลงเบื้องต้นของทฤษฎีฉันทามติทางวัฒนธรรมแบบมีรูปแบบ พบว่า ชาติความเป็นเอกมิตินี้ เนื่องจากผู้เชี่ยวชาญที่มีระดับความสามารถในการตัดสินใจด้านความเป็นคู่ขนานของข้อสอบที่ไม่ผ่านเกณฑ์มีผลให้เกิดระดับความสามารถที่แตกต่างกันของกลุ่มประชากร จึงต้องทำการวิเคราะห์ฉันทามติอีกครั้งเพื่อตรวจสอบระดับความสามารถในการตัดสินใจด้านความเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญโดยอาศัยทฤษฎีฉันทามติทางวัฒนธรรมแบบมีรูปแบบ ทั้งนี้จะกำหนดเงื่อนไขสำหรับการวิเคราะห์ทางสถิติโดยตัดผู้เชี่ยวชาญที่ไม่ผ่านเกณฑ์ออกเพื่อทำให้กลุ่มผู้เชี่ยวชาญที่นำมาวิเคราะห์ฉันทามติมีการตอบสนองต่อการตัดสินใจด้านความเป็นคู่ขนานของข้อสอบไปในทิศทางเดียวกันซึ่งแสดงถึงความเป็นเอกมิติตามข้อตกลงเบื้องต้น มีรายละเอียดการวิเคราะห์ฉันทามติเมื่อไม่พิจารณาผู้เชี่ยวชาญที่ไม่ผ่านเกณฑ์ ดังต่อไปนี้

2.3 ผลการตรวจสอบข้อตกลงเบื้องต้นของการตรวจสอบระดับความสามารถในการตัดสินใจด้านความเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญโดยอาศัยทฤษฎีฉันทามติทางวัฒนธรรมแบบมีรูปแบบเมื่อไม่พิจารณาผู้เชี่ยวชาญที่ไม่ผ่านเกณฑ์ แบ่งออกเป็น 3 ด้าน ดังต่อไปนี้

2.3.1 ความเป็นเอกมิตินี้ของแบบวัด (Homogeneity) ผลการตรวจสอบข้อตกลงเบื้องต้นด้านความเป็นเอกมิตินี้ของแบบวัดระดับความสามารถในการตัดสินใจด้านความเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญโดยอาศัยทฤษฎีฉันทามติทางวัฒนธรรมแบบมีรูปแบบ โดยพิจารณาจากสัดส่วนของ Eigenvalue สูงสุดต่อลำดับที่ 2 พบว่า มีสัดส่วนมากกว่า 3 เท่า ดังนั้น แบบวัดระดับความสามารถในการตัดสินใจด้านความเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญจึงมีความเป็นเอกมิตินี้ของแบบวัดและเป็นไปตามข้อตกลงเบื้องต้นของทฤษฎีฉันทามติทางวัฒนธรรมแบบมีรูปแบบสามารถวิเคราะห์ระดับความสามารถในการตัดสินใจด้านความเป็นคู่ขนานของข้อสอบได้ แสดงรายละเอียดดังตาราง 4.4

ตารางที่ 4.4 ผลการแบ่งกลุ่มผู้เชี่ยวชาญจำแนกตามระดับความสามารถในการตัดสินใจด้านความเป็นคู่ขนานของข้อสอบ

ค่าสถิติ	ผลการวิเคราะห์
จำนวนผู้เชี่ยวชาญ	43 คน
Eigenvalue สูงสุด	18.294
Eigenvalue ลำดับที่ 2	5.878
สัดส่วนของ Eigenvalue สูงสุดต่อลำดับที่ 2	3.112

2.3.2 ความเป็นอิสระต่อกันของผู้ตอบ (Conditional independence) โดยกำหนดให้ผู้เชี่ยวชาญตอบคำถามแบบวัดระดับความสามารถในตัดสินความเป็นคู่ขนานของข้อสอบโดยไม่สามารถสอบถามความคิดเห็นระหว่างกัน ดังนั้น แบบวัดนี้จึงมีความเป็นอิสระต่อกันของผู้ตอบตามข้อตกลงเบื้องต้นของทฤษฎีฉันทามติทางวัฒนธรรมแบบมีรูปแบบ

2.3.3 ความถูกต้อง (Common truth) โดยกำหนดให้คำตอบของแบบวัดระดับความสามารถในตัดสินความเป็นคู่ขนานของข้อสอบมีเพียงคำตอบเดียว คือ ความเป็นคู่ขนานกันของข้อสอบหรือไม่มีความเป็นคู่ขนานกันของข้อสอบ ดังนั้น แบบวัดนี้จึงมีคุณสมบัติด้านความถูกต้องตามข้อตกลงเบื้องต้นของทฤษฎีฉันทามติทางวัฒนธรรมแบบมีรูปแบบ

2.4 ผลการตรวจสอบระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญ

เมื่อวิเคราะห์ฉันทามติแล้วเป็นไปตามข้อตกลงเบื้องต้นของทฤษฎีฉันทามติทางวัฒนธรรมแบบมีรูปแบบแล้วจึงทำการตรวจสอบระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญ (ตารางที่ 4.5) เพื่อใช้ในการแบ่งกลุ่มผู้เชี่ยวชาญตามระดับความสามารถในการคัดเลือกข้อสอบคู่ขนานได้เป็น 3 ระดับ ตามเกณฑ์ดังต่อไปนี้

ระดับที่ 1 ผู้เชี่ยวชาญที่มีระดับความสามารถในการคัดเลือกข้อสอบคู่ขนานตั้งแต่ 0.80 ถึง 1.00

ระดับที่ 2 ผู้เชี่ยวชาญที่มีระดับความสามารถในการคัดเลือกข้อสอบคู่ขนานตั้งแต่ 0.60 ถึง 0.79

ระดับที่ 3 ผู้เชี่ยวชาญที่มีระดับความสามารถในการคัดเลือกข้อสอบคู่ขนานตั้งแต่ 0.40 ถึง 0.59

แสดงรายละเอียดดังตารางที่ 4.6

ตารางที่ 4.5 ผลการตรวจสอบระดับความสามารถในการตัดสินความเป็นคู่ขนานของผู้เชี่ยวชาญ

ผู้เชี่ยวชาญ	ความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบ	ระดับ
1	0.942	1
2	0.918	1
3	0.860	1
4	0.853	1
5	0.834	1
6	0.821	1
7	0.803	1
8	0.788	2
9	0.781	2
10	0.742	2
11	0.741	2
12	0.737	2
13	0.735	2
14	0.719	2
15	0.694	2
16	0.688	2
17	0.668	2
18	0.660	2
19	0.631	2
20	0.630	2
21	0.629	2
22	0.618	2
23	0.601	2
24	0.598	3
25	0.597	3
26	0.586	3
27	0.581	3
28	0.579	3
29	0.579	3
30	0.557	3
31	0.551	3
32	0.538	3

ตารางที่ 4.5 ผลการตรวจสอบระดับความสามารถในการตัดสินความเป็นคู่ขนานของผู้เชี่ยวชาญ
(ต่อ)

ผู้เชี่ยวชาญ	ความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบ	ระดับ
33	0.530	3
34	0.522	3
35	0.521	3
36	0.499	3
37	0.483	3
38	0.451	3
39	0.451	3
40	0.424	3
41	0.423	3
42	0.414	3
43	0.401	3

ตารางที่ 4.6 ผลการแบ่งกลุ่มผู้เชี่ยวชาญจำแนกตามระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบ

ระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบ	ช่วงคะแนนความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบ	การแปลผล	จำนวน	ร้อยละ
ระดับที่ 1	0.80-1.00	ผ่านเกณฑ์	7	16.28
ระดับที่ 2	0.60-0.79		16	37.21
ระดับที่ 3	0.40-0.59		20	46.51
รวม			43	100

จากการประมาณค่าระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญ โดยอาศัยทฤษฎีฉันทามติทางวัฒนธรรมแบบมีรูปแบบ พบว่า ผู้เชี่ยวชาญส่วนใหญ่มีระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบอยู่ในระดับที่ 3 มากกว่าระดับที่ 2 และ 1 ตามลำดับ ทั้งนี้จำนวนตัวอย่างผู้เชี่ยวชาญในแต่ละกลุ่มเพียงพอสำหรับการจำแนกผู้เชี่ยวชาญสำหรับการคัดเลือกข้อสอบเพื่อสร้างแบบสอบคู่ขนานได้ นอกจากนี้พบว่า ระดับความสามารถในการตรวจสอบความเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญมีค่าสูงสุดเท่ากับ 0.942 และค่าต่ำสุดเท่ากับ 0.401 โดยมีค่าเฉลี่ยเท่ากับ 0.637

ส่วนที่ 3 ผลการแบ่งกลุ่มผู้เชี่ยวชาญจำแนกตามระดับความสามารถในการตัดสินความเป็น คู่ขนานของข้อสอบและประสบการณ์สอน

จากการตรวจสอบระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญสามารถแบ่งกลุ่มผู้เชี่ยวชาญตามระดับความสามารถในการคัดเลือกข้อสอบคู่ขนานได้เป็น 3 ระดับ แสดงรายละเอียดดังตาราง 4.1 และพิจารณาจำแนกกลุ่มผู้เชี่ยวชาญตามประสบการณ์สอนออกเป็น 2 กลุ่ม (ตารางที่ 4.7) ได้แก่ ประสบการณ์สอนตั้งแต่ 10 ปีขึ้นไปและน้อยกว่า 10 ปี จึงสามารถแบ่งกลุ่มผู้เชี่ยวชาญที่มีลักษณะแตกต่างกันเป็น 6 กลุ่ม โดยพบว่า ผู้เชี่ยวชาญส่วนใหญ่มีระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบอยู่ในระดับที่ 3 มากกว่าระดับที่ 2 และ 1 โดยเฉพาะในกลุ่มผู้เชี่ยวชาญที่มีประสบการณ์สอนนักศึกษาแพทย์น้อยกว่า 10 ปี แต่ในกลุ่มผู้เชี่ยวชาญที่มีประสบการณ์สอนตั้งแต่ 10 ปีขึ้นไปจะมีระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบในระดับที่ 2 จำนวนสูงสุด ทั้งนี้กลุ่มผู้เชี่ยวชาญในแต่ละกลุ่มมีจำนวนเพียงพอต่อการเลือกผู้เชี่ยวชาญอย่างจำเพาะสำหรับคัดเลือกข้อสอบเพื่อสร้างแบบสอบคู่ขนานในขั้นตอนถัดไป คือ ใช้ตัวอย่างผู้เชี่ยวชาญจำนวน ทั้งหมด 18 คนโดยแบ่งเป็น 6 กลุ่มๆละ 3 คน แสดงรายละเอียดดังตารางที่ 4.8

ตารางที่ 4.7 ผลการแบ่งกลุ่มผู้เชี่ยวชาญจำแนกตามระดับความสามารถในการตัดสินความเป็น คู่ขนานของข้อสอบ

ลักษณะของผู้เชี่ยวชาญ		ระดับความสามารถในการตัดสิน ความเป็นคู่ขนานของข้อสอบ			รวม
		ระดับที่ 1 (0.80-1.00)	ระดับที่ 2 (0.60-0.79)	ระดับที่ 3 (0.40-0.59)	
ประสบการณ์ สอน	ตั้งแต่ 10 ปีขึ้นไป	4	10	5	19
	น้อยกว่า 10 ปี	3	6	15	24
รวม		7	16	20	43

ตารางที่ 4.8 ข้อมูลภูมิหลังของกลุ่มผู้เชี่ยวชาญที่เป็นตัวอย่างในการคัดเลือกข้อสอบเพื่อสร้างแบบสอบคู่ขนาน

กลุ่ม	ลำดับ	เพศ	สังกัด ภาควิชา	ระดับ การศึกษา สูงสุด	สาขาที่จบ การศึกษา	ประสบการณ์ การสอน นักศึกษา แพทย์	การอบรม แพทยศาสตร ศึกษา	ระดับความสามารถ ในการตัดสินความ เป็นคู่ขนานของ ข้อสอบ
1	1	หญิง	คลินิก	ปริญญาเอก	พบ.	18	ไม่เคย	0.86
	2	หญิง	คลินิก	ปริญญาเอก	พบ.	15	เคย	0.853
	3	หญิง	ปริคlinik	ปริญญาเอก	อื่นๆ	10	เคย	0.821
2	1	หญิง	ปริคlinik	ปริญญาเอก	อื่นๆ	1	เคย	0.942
	2	หญิง	คลินิก	ปริญญาโท	พบ.	9	เคย	0.918
	3	ชาย	ปริคlinik	ปริญญาโท	พบ.	1	เคย	0.834
3	1	หญิง	ปริคlinik	ปริญญาเอก	อื่นๆ	20	เคย	0.742
	2	ชาย	คลินิก	ปริญญาโท	พบ.	10	เคย	0.741
	3	ชาย	ปริคlinik	ปริญญาเอก	อื่นๆ	11	เคย	0.737
4	1	ชาย	ปริคlinik	ปริญญาโท	พบ.	6	เคย	0.781
	2	ชาย	ปริคlinik	ปริญญาตรี	พบ.	2	เคย	0.688
	3	หญิง	ปริคlinik	ปริญญาเอก	พบ.	7	เคย	0.66
5	1	หญิง	คลินิก	ปริญญาโท	อื่นๆ	10	เคย	0.579
	2	หญิง	ปริคlinik	ปริญญาเอก	อื่นๆ	10	เคย	0.551
	3	หญิง	ปริคlinik	ปริญญาโท	ภบ.	10	ไม่เคย	0.538
6	1	ชาย	คลินิก	ปริญญาโท	พบ.	5	เคย	0.598
	2	หญิง	ปริคlinik	ปริญญาเอก	อื่นๆ	2	ไม่เคย	0.586
	3	หญิง	คลินิก	ปริญญาตรี	อื่นๆ	1	ไม่เคย	0.579

ส่วนที่ 4 ผลการตัดสินคำตอบที่ถูกต้องจากการวิเคราะห์ฉันทามติจากการตัดสินความเป็นคู่ขนานของข้อสอบโดยผู้เชี่ยวชาญ

การวิเคราะห์ฉันทามติ นอกจากใช้ประมาณค่าความสามารถในการตัดสินความเป็นคู่ขนานกันของข้อสอบของผู้เชี่ยวชาญแล้ว ยังสามารถวิเคราะห์ทางสถิติเพื่อตัดสินคำตอบที่ถูกต้องจากการวิเคราะห์ฉันทามติจากการตัดสินความเป็นคู่ขนานของข้อสอบโดยผู้เชี่ยวชาญ ในครั้งนี้เมื่อทำการวิเคราะห์ทางสถิติโดยอาศัยทฤษฎีฉันทามติทางวัฒนธรรมแบบมีรูปแบบ พบว่า ข้อสอบมีความเป็นคู่ขนานกันจำนวน 15 คู่ และข้อสอบไม่มีความเป็นคู่ขนานกันจำนวน 5 คู่ คิดเป็นอัตราส่วนระหว่างข้อคำถามที่มีคำตอบเชิงบวกต่อข้อคำถามที่มีคำตอบเชิงลบเท่ากับร้อยละ 75 ต่อ 25 ซึ่งใกล้เคียงกับสัดส่วนที่เหมาะสมสำหรับการวิเคราะห์ฉันทามติที่กำหนดว่าควรมีอัตราส่วนของคำตอบเชิงบวกต่อเชิงลบที่ร้อยละ 70 ต่อ 30 แสดงรายละเอียดดังตารางที่ 4.9

ตารางที่ 4.9 ผลการตอบแบบวัดความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบโดยผู้เชี่ยวชาญ

ข้อ	จำนวนผลการตัดสินความเป็นคู่ขนานของข้อสอบโดยผู้เชี่ยวชาญ		คำตอบที่ถูกต้องจากการวิเคราะห์ฉันทามติ
	ข้อสอบเป็นคู่ขนานกัน	ข้อสอบไม่เป็นคู่ขนานกัน	
1	33	10	ข้อสอบเป็นคู่ขนานกัน
2	35	8	ข้อสอบเป็นคู่ขนานกัน
3	7	36	ข้อสอบไม่เป็นคู่ขนานกัน
4	42	1	ข้อสอบเป็นคู่ขนานกัน
5	34	9	ข้อสอบเป็นคู่ขนานกัน
6	32	11	ข้อสอบเป็นคู่ขนานกัน
7	32	11	ข้อสอบเป็นคู่ขนานกัน
8	40	3	ข้อสอบเป็นคู่ขนานกัน
9	36	7	ข้อสอบเป็นคู่ขนานกัน
10	6	37	ข้อสอบไม่เป็นคู่ขนานกัน
11	19	24	ข้อสอบไม่เป็นคู่ขนานกัน
12	40	3	ข้อสอบเป็นคู่ขนานกัน
13	24	19	ข้อสอบเป็นคู่ขนานกัน
14	41	2	ข้อสอบเป็นคู่ขนานกัน
15	34	9	ข้อสอบเป็นคู่ขนานกัน
16	34	9	ข้อสอบเป็นคู่ขนานกัน
17	17	26	ข้อสอบไม่เป็นคู่ขนานกัน
18	39	8	ข้อสอบเป็นคู่ขนานกัน
19	15	28	ข้อสอบไม่เป็นคู่ขนานกัน
20	40	3	ข้อสอบเป็นคู่ขนานกัน

ตอนที่ 2 เพื่อเปรียบเทียบดัชนีความเป็นคู่ขนานของแบบสอบคู่ขนานที่ได้จากการคัดเลือกโดยกลุ่มผู้เชี่ยวชาญที่มีระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบและประสบการณ์สอนนักศึกษาแพทย์ที่แตกต่างกัน

เมื่อทำการแบ่งกลุ่มผู้เชี่ยวชาญตามระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบและประสบการณ์สอนนักศึกษาแพทย์ที่แตกต่างกันแล้ว ขั้นตอนต่อไปกลุ่มผู้เชี่ยวชาญทำการคัดเลือกข้อสอบเพื่อสร้างแบบสอบคู่ขนาน และตรวจสอบความเป็นคู่ขนานของข้อสอบและแบบสอบโดยอาศัยทฤษฎีการตอบสนองข้อสอบ โมเดล 2 พารามิเตอร์โดยใช้สอบกับกลุ่มผู้ให้ข้อมูลที่เป็นนักศึกษาแพทย์ชั้นปีที่ 3 ถึง 6 จำนวน 400 คน โดยนักศึกษาแพทย์แต่ละคนทำการสอบด้วยแบบสอบคู่ขนานทั้งหมด 3 คู่ รวมทั้งสิ้นคนละ 6 ฉบับ แบ่งเป็นแบบสอบคู่ขนานชุด 1A, 1B, 2A, 2B, 3A และ 3B จำนวน 200 คนและแบบสอบคู่ขนานชุด 4A, 4B, 5A, 5B, 6A และ 6B จำนวน 200 คน จากนั้นจึงวิเคราะห์ทางสถิติเพื่อเปรียบเทียบดัชนีความเป็นคู่ขนานของแบบสอบ ผู้วิจัยได้แบ่งการนำเสนอเป็น 2 ส่วน ได้แก่ ส่วนที่ 1 ผลการคัดเลือกข้อสอบคู่ขนานของกลุ่มผู้เชี่ยวชาญที่มีลักษณะแตกต่างกันและส่วนที่ 2 ผลการตรวจสอบความเป็นคู่ขนานของแบบสอบและข้อสอบโดยอาศัยทฤษฎีการตอบสนองข้อสอบ แสดงรายละเอียดดังนี้

ส่วนที่ 1 ผลการคัดเลือกข้อสอบคู่ขนานของกลุ่มผู้เชี่ยวชาญที่มีลักษณะแตกต่างกัน

จากการคัดเลือกข้อสอบเพื่อสร้างแบบสอบคู่ขนานของกลุ่มผู้เชี่ยวชาญที่มีลักษณะแตกต่างกันโดยจำแนกตามระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบและประสบการณ์สอนเป็น 6 กลุ่ม และคัดเลือกข้อสอบสำหรับสร้างแบบสอบคู่ขนานจำนวน 25 คู่ (แบบสอบละ 25 ข้อ) จากทั้งหมด 5 หมวด (หมวด A, B, C, D และ E) ซึ่งแต่ละหมวดประกอบด้วยข้อสอบ 40 ข้อ (เช่น หมวด A ตั้งแต่ A1 ถึง A40 เป็นต้น) โดยแบ่งตามจุดประสงค์การเรียนรู้ ดังนี้ หมวด A = ยาลดความดันโลหิต หมวด B = ยารักษาโรคหัวใจ หมวด C = ยารักษาแผลในทางเดินอาหาร หมวด D = ยาที่เกี่ยวข้องกับบิบัติตัวของลำไส้ และหมวด E = ยารักษาโรคเบาหวาน พบว่า ส่วนใหญ่กลุ่มผู้เชี่ยวชาญทั้ง 6 กลุ่มคัดเลือกข้อสอบเพื่อสร้างแบบสอบคู่ขนานที่แตกต่างกัน มีเพียงบางส่วนที่สอดคล้องกันระหว่างกลุ่ม เช่น ข้อ A2 และ A18, A4 และ A7, C2 และ C10 เป็นต้น โดยผลการคัดเลือกข้อสอบคู่ขนานของกลุ่มผู้เชี่ยวชาญที่มีลักษณะแตกต่างกันแสดงรายละเอียดดังตารางที่ 4.10

ตารางที่ 4.10 ผลการคัดเลือกข้อสอบคู่ขนานของกลุ่มผู้เชี่ยวชาญที่มีลักษณะแตกต่างกัน

กลุ่ม		1		2		3		4		5		6	
ระดับความสามารถ ในการตัดสินความ เป็นคู่ขนานของ ข้อสอบ		ระดับที่ 1		ระดับที่ 1		ระดับที่ 2		ระดับที่ 2		ระดับที่ 3		ระดับที่ 3	
ประสบการณ์สอน		≥ 10 ปี		< 10 ปี		≥ 10 ปี		< 10 ปี		≥ 10 ปี		< 10 ปี	
ชุดข้อสอบ		1	2	1	2	1	2	1	2	1	2	1	2
หมวด A	ข้อ 1	A2	A18	A2	A18	A1	A17	A2	A18	A2	A18	A2	A18
	ข้อ 2	A3	A37	A3	A37	A4	A7	A4	A7	A4	A7	A5	A19
	ข้อ 3	A4	A7	A4	A7	A8	A32	A5	A19	A5	A19	A4	A7
	ข้อ 4	A5	A19	A5	A19	A9	A19	A7	A14	A6	A8	A8	A32
	ข้อ 5	A15	A22	A11	A23	A25	A27	A28	A32	A17	A40	A10	A11
หมวด B	ข้อ 6	B1	B21	B6	B35	B6	B35	B2	B3	B5	B6	B6	B11
	ข้อ 7	B5	B17	B7	B8	B7	B36	B6	B11	B7	B8	B7	B36
	ข้อ 8	B6	B11	B17	B33	B13	B24	B7	B8	B13	B24	B13	B24
	ข้อ 9	B9	B15	B21	B24	B14	B31	B9	B15	B20	B25	B17	B33
	ข้อ 10	B22	B32	B22	B32	B22	B32	B20	B23	B17	B36	B22	B32
หมวด C	ข้อ 11	C5	C19	C2	C10	C2	C10	C2	C10	C7	C14	C2	C10
	ข้อ 12	C6	C9	C6	C13	C5	C12	C6	C9	C12	C19	C5	C15
	ข้อ 13	C8	C36	C15	C29	C7	C14	C7	C14	C15	C29	C6	C13
	ข้อ 14	C26	C32	C24	C30	C8	C36	C16	C25	C24	C28	C8	C36
	ข้อ 15	C28	C30	C26	C32	C26	C32	C26	C32	C27	C31	C34	C38
หมวด D	ข้อ 16	D2	D9	D1	D37	D4	D8	D2	D9	D4	D11	D2	D9
	ข้อ 17	D4	D11	D2	D14	D7	D13	D4	D11	D13	D15	D4	D11
	ข้อ 18	D5	D12	D4	D8	D23	D25	D5	D18	D19	D28	D12	D18
	ข้อ 19	D7	D13	D7	D15	D29	D30	D7	D15	D23	D25	D13	D15
	ข้อ 20	D10	D28	D23	D25	D35	D36	D19	D24	D29	D30	D29	D30
หมวด E	ข้อ 21	E4	E12	E2	E28	E1	E38	E2	E14	E3	E9	E6	E33
	ข้อ 22	E7	E15	E3	E9	E2	E19	E5	E27	E4	E12	E10	E17
	ข้อ 23	E10	E17	E4	E12	E7	E21	E9	E10	E5	E27	E18	E32
	ข้อ 24	E9	E25	E5	E27	E8	E40	E18	E32	E20	E37	E22	E29
	ข้อ 25	E18	E32	E17	E24	E10	E24	E22	E29	E24	E38	E24	E31

ส่วนที่ 2 ผลการตรวจสอบความเป็นคู่ขนานของแบบสอบและข้อสอบโดยอาศัยทฤษฎีการตอบสนองข้อสอบ

เมื่อกลุ่มผู้เชี่ยวชาญแต่ละกลุ่มทำการคัดเลือกข้อสอบเพื่อสร้างแบบสอบคู่ขนานแล้วจึงดำเนินการเก็บข้อมูลกับกลุ่มนักศึกษาแพทย์ผู้ให้ข้อมูล จากนั้นจึงวิเคราะห์ทางสถิติโดยใช้โปรแกรม IRTpro เพื่อหาค่าสารสนเทศของแบบสอบและพารามิเตอร์ของข้อสอบ ซึ่งการตรวจสอบความเป็นคู่ขนานของแบบสอบและข้อสอบโดยอาศัยทฤษฎีการตอบสนองข้อสอบ จะทำการพิจารณาความสอดคล้องของข้อสอบที่คัดเลือกโดยผู้เชี่ยวชาญกับโมเดลตามทฤษฎีการตอบสนองข้อสอบแบบ 2 พารามิเตอร์ (2PL) โดยอาศัยสถิติทดสอบไคสแควร์ เมื่อข้อสอบมีความสอดคล้องกับโมเดลแล้วจึงตรวจสอบความเป็นคู่ขนานของแบบสอบและข้อสอบตามลำดับ มีรายละเอียดดังต่อไปนี้

2.1 ผลการตรวจสอบความสอดคล้องของข้อสอบที่คัดเลือกโดยผู้เชี่ยวชาญกับโมเดลตามทฤษฎีการตอบสนองข้อสอบแบบ 2 พารามิเตอร์ (2PL)

เมื่อกลุ่มผู้ให้ข้อมูลนักศึกษาแพทย์ชั้นปีที่ 3 ถึง 6 ทำแบบสอบคู่ขนานทั้งหมด 12 ชุด ประกอบด้วย 1A, 1B, 2A, 2B, 3A, 3B, 4A, 4B, 5A, 5B, 6A และ 6B แล้วจึงทำการตรวจสอบความสอดคล้องของโมเดล 2 พารามิเตอร์ (2PL) โดยใช้สถิติทดสอบไคสแควร์ (χ^2) กำหนดระดับนัยสำคัญทางสถิติที่ 0.05 พบว่า ส่วนใหญ่แบบสอบที่ได้จากการคัดเลือกข้อสอบโดยกลุ่มผู้เชี่ยวชาญสอดคล้องกับโมเดล 2PL ซึ่งแบบสอบจำนวน 9 ฉบับ (ร้อยละ 75.000) ที่ข้อสอบมีความสอดคล้องกับโมเดล 2PL อยู่ในช่วงร้อยละ 90-100 และแบบสอบ 2 ฉบับที่ข้อสอบมีความสอดคล้องกับโมเดล 2PL อยู่ในช่วงร้อยละ 50-89 (ข้อสอบชุด 2B และ 6B) แต่พบว่ามีแบบสอบ 1 ฉบับที่มีข้อสอบสอดคล้องกับโมเดล 2PL ค่อนข้างต่ำ (ร้อยละ 28) ได้แก่ ข้อสอบชุด 3A นอกจากนี้ จากการวิเคราะห์ พบว่า ค่าเฉลี่ยของพารามิเตอร์ความยากของข้อสอบของทุกฉบับอยู่ในช่วง 0.00 ถึง 1.00 ซึ่งอยู่ในระดับปานกลาง และค่าเฉลี่ยพารามิเตอร์อำนาจจำแนกของข้อสอบส่วนใหญ่อยู่ระหว่าง 0.00 ถึง 1.50 ซึ่งถือว่าแบบสอบเหมาะสมสำหรับผู้สอบที่มีความสามารถในระดับปานกลางถึงค่อนข้างสูง แสดงรายละเอียดดังตาราง 4.11

ตารางที่ 4.11 ผลการตรวจสอบความสอดคล้องของข้อสอบที่คัดเลือกโดยผู้เชี่ยวชาญกับโมเดลตามทฤษฎีการตอบสนองข้อสอบแบบ 2 พารามิเตอร์ (2PL) โดยใช้สถิติทดสอบไคแอสควร์ (χ^2)

กลุ่ม		1				2			
ระดับความสามารถในการตัดสินใจ ความเป็นคู่ขนานของข้อสอบ		ระดับที่ 1				ระดับที่ 1			
ประสบการณ์สอน		≥ 10 ปี				< 10 ปี			
ชุดข้อสอบ		1A		1B		2A		2B	
สถิติทดสอบ		χ^2	p	χ^2	p	χ^2	p	χ^2	p
ข้อ	1	19.51	0.243	14.88	0.461	20.59	0.112	11.52	0.403
	2	11.01	0.530	14.58	0.409	16.59	0.346	13.03	0.222
	3	21.31	0.127	15.04	0.450	14.67	0.259	24.38	0.007
	4	13.86	0.610	18.03	0.261	18.22	0.250	28.66	0.007
	5	12.60	0.634	13.73	0.395	18.45	0.239	22.37	0.022
	6	17.67	0.280	19.45	0.193	9.54	0.732	33.31	0.001
	7	11.55	0.714	16.04	0.310	9.01	0.831	15.59	0.112
	8	13.13	0.361	6.44	0.843	14.19	0.512	20.08	0.028
	9	21.89	0.146	6.71	0.979	15.17	0.232	12.13	0.145
	10	13.87	0.461	12.11	0.599	14.29	0.356	10.52	0.486
	11	12.83	0.617	14.81	0.540	16.50	0.283	21.52	0.028
	12	28.32	0.020	13.67	0.476	18.17	0.199	10.58	0.307
	13	16.16	0.240	10.32	0.740	15.27	0.362	22.42	0.033
	14	15.79	0.200	14.35	0.278	18.57	0.182	15.86	0.146
	15	14.66	0.260	18.62	0.098	12.40	0.415	29.19	0.001
	16	9.12	0.872	14.52	0.561	19.00	0.213	11.84	0.543
	17	11.28	0.589	11.32	0.503	17.32	0.239	14.52	0.150
	18	23.87	0.092	17.12	0.145	21.58	0.119	28.07	0.009
	19	10.40	0.846	23.88	0.067	17.19	0.190	23.98	0.020
	20	18.48	0.185	17.73	0.219	15.53	0.275	14.64	0.145
	21	24.40	0.081	22.21	0.074	9.58	0.793	27.80	0.006
	22	12.02	0.800	13.17	0.515	13.19	0.357	38.29	0.000
	23	24.19	0.043	10.53	0.723	17.62	0.128	16.21	0.094
	24	21.73	0.115	20.02	0.029	14.59	0.336	16.93	0.110
	25	28.78	0.011	11.7	0.632	14.67	0.403	22.11	0.054
ค่าความยากเฉลี่ย		0.86		0.85		0.73		0.68	
ค่าอำนาจจำแนกเฉลี่ย		0.51		0.88		0.44		1.42	
ร้อยละของข้อสอบที่สอดคล้องกับโมเดล 2PL		92.00		96.00		100.00		52.00	

ตารางที่ 4.11 ผลการตรวจสอบความสอดคล้องของข้อสอบที่คัดเลือกโดยผู้เชี่ยวชาญกับโมเดลตาม
ทฤษฎีการตอบสนองข้อสอบแบบ 2 พารามิเตอร์ (2PL) โดยใช้สถิติทดสอบไคแอสควร์ (χ^2) (ต่อ)

กลุ่ม		3				4			
ระดับความสามารถในการตัดสิน ความเป็นคู่ขนานของข้อสอบ		ระดับที่ 2				ระดับที่ 2			
ประสบการณ์สอน		≥ 10 ปี				< 10 ปี			
ชุดข้อสอบ		3A		3B		4A		4B	
สถิติทดสอบ		χ^2	p	χ^2	p	χ^2	p	χ^2	p
ข้อ	1	24.88	<u>0.024</u>	11.17	0.515	15.77	0.329	9.53	0.658
	2	17.21	0.142	14.96	0.312	9.13	0.871	6.98	0.904
	3	22.45	0.070	11.70	0.554	20.29	0.088	14.89	0.247
	4	34.40	<u>0.002</u>	15.17	0.296	17.77	0.274	9.05	0.770
	5	33.47	<u>0.004</u>	11.81	0.622	13.70	0.550	9.01	0.831
	6	22.09	0.077	8.97	0.776	15.94	0.388	8.28	0.826
	7	41.75	<u>0.000</u>	9.85	0.707	25.81	<u>0.011</u>	26.49	<u>0.009</u>
	8	23.37	<u>0.037</u>	4.12	0.847	13.29	0.274	10.30	0.328
	9	22.67	<u>0.031</u>	7.72	0.808	16.36	0.430	10.10	0.609
	10	34.64	<u>0.001</u>	9.68	0.645	13.29	0.581	9.51	0.798
	11	22.15	0.053	10.15	0.682	18.83	0.277	10.16	0.682
	12	36.05	<u>0.002</u>	15.88	0.255	23.45	0.053	12.84	0.462
	13	16.49	0.169	10.31	0.670	22.15	0.104	11.66	0.474
	14	26.29	<u>0.016</u>	17.26	0.242	22.63	<u>0.046</u>	7.25	0.841
	15	33.94	<u>0.001</u>	9.70	0.643	17.83	0.271	13.71	0.397
	16	26.79	<u>0.030</u>	17.25	0.188	22.15	0.138	7.59	0.939
	17	19.53	0.076	12.34	0.580	17.65	0.347	12.68	0.475
	18	26.69	<u>0.009</u>	12.89	0.537	10.77	0.824	16.99	0.256
	19	29.79	<u>0.008</u>	14.27	0.218	21.16	0.097	9.45	0.739
	20	23.98	<u>0.046</u>	17.29	0.241	24.40	0.058	19.93	0.097
	21	20.31	0.061	19.14	0.119	13.76	0.545	6.97	0.936
	22	25.92	<u>0.011</u>	6.49	0.889	14.13	0.292	14.57	0.337
	23	32.45	<u>0.002</u>	21.17	0.097	12.62	0.633	19.51	0.108
	24	43.80	<u>0.000</u>	8.84	0.785	20.19	0.063	16.60	0.345
	25	33.43	<u>0.001</u>	9.98	0.534	17.18	0.376	14.70	0.328
ค่าความยากเฉลี่ย		0.72		0.74		0.90		0.76	
ค่าอำนาจจำแนกเฉลี่ย		0.57		1.42		0.84		0.43	
ร้อยละของข้อสอบที่สอดคล้องกับ โมเดล 2PL		28.00		100.00		92.00		96.00	

ตารางที่ 4.11 ผลการตรวจสอบความสอดคล้องของข้อสอบที่คัดเลือกโดยผู้เชี่ยวชาญกับโมเดลตาม
ทฤษฎีการตอบสนองข้อสอบแบบ 2 พารามิเตอร์ (2PL) โดยใช้สถิติทดสอบไคแอสควร์ (χ^2) (ต่อ)

กลุ่ม		5				6			
ระดับความสามารถในการตัดสิน ความเป็นคู่ขนานของข้อสอบ		ระดับที่ 3				ระดับที่ 3			
ประสบการณ์สอน		≥ 10 ปี				< 10 ปี			
ชุดข้อสอบ		5A		5B		6A		6B	
สถิติทดสอบ		χ^2	p	χ^2	p	χ^2	p	χ^2	p
ข้อ	1	16.85	0.155	11.95	0.369	11.50	0.648	15.83	0.147
	2	13.69	0.474	10.52	0.398	8.53	0.744	11.71	0.387
	3	20.71	<u>0.036</u>	17.17	0.103	16.55	0.348	17.44	0.134
	4	18.67	0.097	10.58	0.566	8.87	0.783	21.93	<u>0.025</u>
	5	12.28	0.506	10.50	0.573	10.20	0.808	15.90	0.196
	6	7.23	0.843	14.96	0.133	11.79	0.381	22.35	<u>0.034</u>
	7	8.40	0.754	10.02	0.693	13.55	0.408	22.69	<u>0.030</u>
	8	15.72	0.264	13.61	0.255	13.42	0.495	15.17	0.231
	9	16.27	0.296	13.16	0.283	7.12	0.930	22.75	<u>0.045</u>
	10	19.39	0.150	9.80	0.635	8.55	0.807	10.23	0.597
	11	8.89	0.782	16.10	0.186	9.72	0.783	30.40	<u>0.001</u>
	12	15.19	0.295	9.89	0.704	9.22	0.817	9.76	0.714
	13	6.34	0.786	9.01	0.532	13.61	0.404	15.40	0.220
	14	12.67	0.475	20.50	0.058	8.97	0.776	21.48	<u>0.044</u>
	15	15.34	0.286	10.96	0.363	9.01	0.703	15.95	0.193
	16	15.03	0.305	10.69	0.638	8.78	0.845	18.08	0.113
	17	8.33	0.872	16.59	0.218	18.35	0.244	23.56	<u>0.035</u>
	18	6.16	0.802	6.65	0.880	6.06	0.870	27.65	<u>0.010</u>
	19	6.97	0.802	9.16	0.425	15.46	0.279	9.95	0.621
	20	8.57	0.805	11.25	0.509	10.04	0.692	22.52	<u>0.048</u>
	21	4.18	0.383	13.72	0.321	6.79	0.913	26.95	<u>0.013</u>
	22	9.02	0.702	4.99	0.959	15.28	0.226	22.60	<u>0.020</u>
	23	8.83	0.639	21.77	<u>0.040</u>	13.70	0.474	12.01	0.446
	24	16.43	0.226	8.76	0.792	12.08	0.674	13.66	0.400
	25	8.19	0.916	15.93	0.194	17.55	0.175	22.58	<u>0.031</u>
ค่าความยากเฉลี่ย		0.61		0.58		0.71		0.49	
ค่าอำนาจจำแนกเฉลี่ย		0.67		0.68		0.32		0.48	
ร้อยละของข้อสอบที่สอดคล้องกับ โมเดล 2PL		96.00		96.00		100.00		52.00	

2.2 ผลการตรวจสอบความเป็นคู่ขนานของแบบสอบโดยอาศัยทฤษฎีการตอบสนอง ข้อสอบ

ผลการตรวจสอบความเป็นคู่ขนานของแบบสอบโดยอาศัยทฤษฎีการตอบสนองข้อสอบสามารถจำแนกการพิจารณาตามดัชนีความเป็นคู่ขนานของแบบสอบ มีรายละเอียดดังต่อไปนี้

2.2.1 ผลการตรวจสอบความเป็นคู่ขนานของแบบสอบโดยใช้ RMSD และ MSG

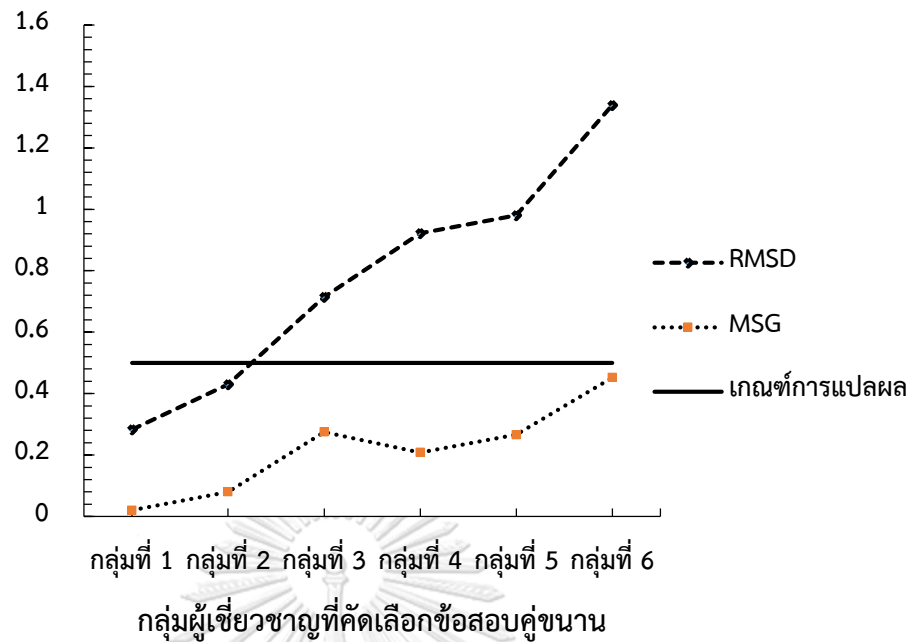
จากการตรวจสอบความเป็นคู่ขนานของแบบสอบโดยใช้ RMSD และ MSG ซึ่งกำหนดเกณฑ์สำหรับแบบสอบคู่ขนานต้องมีค่า RMSD และ MSG ต่ำกว่า 0.50 พบว่า แบบสอบคู่ขนานที่คัดเลือกโดยกลุ่มผู้เชี่ยวชาญที่มีความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบในระดับที่ 1 ทั้งกลุ่มที่มีประสบการณ์สอน ≥ 10 ปี (กลุ่มที่ 1) และประสบการณ์สอน < 10 ปี (กลุ่มที่ 2) มีความเป็นคู่ขนานของแบบสอบโดยมีค่า RMSD (MSG) เท่ากับ 0.283 (0.020) และ 0.430 (0.080) ตามลำดับ ส่วนแบบสอบที่คัดเลือกโดยกลุ่มผู้เชี่ยวชาญกลุ่มอื่นๆ ไม่มีความเป็นคู่ขนานของแบบสอบ แสดงรายละเอียดดังตารางที่ 4.12 นอกจากนี้ พบว่า แนวโน้มของค่า RMSD และ MSG มีค่าสูงขึ้นตามระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบของกลุ่มผู้เชี่ยวชาญที่ลดลงและมีประสบการณ์สอน < 10 ปี ดังภาพที่ 4.1

2.2.2 ผลการตรวจสอบความเป็นคู่ขนานของแบบสอบโดยใช้ MRD และ MRIG

จากการตรวจสอบความเป็นคู่ขนานของแบบสอบโดยใช้ MRD และ MRIG ซึ่งกำหนดเกณฑ์สำหรับแบบสอบคู่ขนานต้องมีค่า MRD และ MRIG ต่ำกว่า 0.10 (น้อยกว่าร้อยละ 10) พบว่าแบบสอบคู่ขนานที่คัดเลือกโดยกลุ่มผู้เชี่ยวชาญที่มีความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบในระดับที่ 1 ที่มีประสบการณ์สอน ≥ 10 ปี (กลุ่มที่ 1) มีความเป็นคู่ขนานของแบบสอบโดยมีค่า MRD (MRIG) เท่ากับ 0.042 (0.003) ส่วนแบบสอบที่คัดเลือกโดยกลุ่มผู้เชี่ยวชาญกลุ่มอื่นๆ ไม่มีความเป็นคู่ขนานของแบบสอบ แสดงรายละเอียดดังตาราง 4.12 นอกจากนี้ พบว่า แนวโน้มของค่า MRD และ MRIG มีค่าสูงขึ้นตามระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบของกลุ่มผู้เชี่ยวชาญที่ลดลงและมีประสบการณ์สอน < 10 ปี ดังภาพที่ 4.2

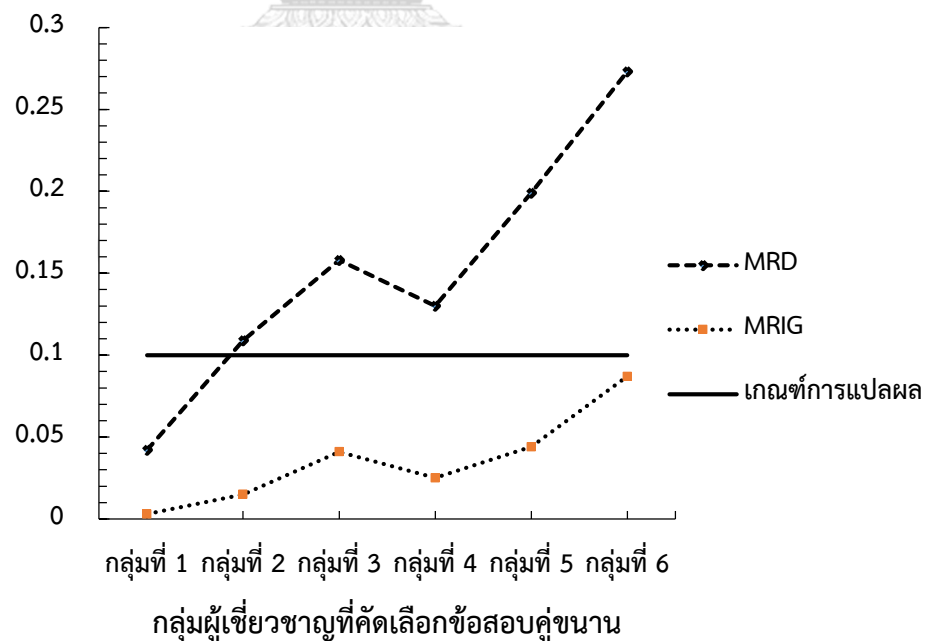
ตารางที่ 4.12 ผลการตรวจสอบความเป็นคู่ขนานของแบบสอบโดยอาศัยทฤษฎีการตอบสนอง
ข้อสอบ

กลุ่ม	1		2		3		4		5		6		
ระดับ ความสามารถ ในการตัดสิน ความเป็น คู่ขนานของ ข้อสอบ	ระดับที่ 1		ระดับที่ 1		ระดับที่ 2		ระดับที่ 2		ระดับที่ 3		ระดับที่ 3		
ประสบการณ์ สอน	≥ 10 ปี		< 10 ปี		≥ 10 ปี		< 10 ปี		≥ 10 ปี		< 10 ปี		
ชุดข้อสอบ	1A	1B	2A	2B	3A	3B	4A	4B	5A	5B	6A	6B	
T(θ)	θ												
	-2.80	1.77	1.73	1.85	1.56	1.95	1.58	1.80	1.71	2.20	1.61	1.98	1.54
	-2.40	2.09	2.05	2.18	1.79	2.32	1.76	2.12	1.96	2.47	1.81	2.39	1.68
	-2.00	2.56	2.55	2.64	2.13	2.81	2.02	2.59	2.30	2.86	2.12	2.95	1.85
	-1.60	3.23	3.29	3.26	2.62	3.43	2.40	3.34	2.78	3.47	2.56	3.63	2.07
	-1.20	4.14	4.35	4.01	3.28	4.12	2.98	4.46	3.49	4.33	3.14	4.35	2.31
	-0.80	5.26	5.63	4.77	4.06	4.75	3.94	5.99	4.52	5.27	3.78	4.91	2.57
	-0.40	6.37	6.67	5.33	4.77	5.15	5.46	7.51	5.76	5.96	4.28	5.15	2.82
	0.00	6.95	6.81	5.49	5.05	5.20	6.63	8.11	6.42	6.06	4.45	5.03	3.02
	0.40	6.63	6.07	5.21	4.83	4.90	5.98	7.42	5.99	5.56	4.24	4.64	3.15
	0.80	5.69	5.09	4.63	4.36	4.37	4.75	6.05	5.08	4.75	3.81	4.10	3.17
	1.20	4.67	4.26	3.96	3.84	3.78	3.90	4.73	4.24	3.93	3.31	3.53	3.08
	1.60	3.83	3.61	3.34	3.37	3.22	3.33	3.73	3.59	3.22	2.85	3.01	2.90
	2.00	3.18	3.08	2.82	2.95	2.76	2.89	3.04	3.09	2.67	2.47	2.57	2.68
	2.40	2.69	2.65	2.41	2.60	2.39	2.55	2.56	2.70	2.25	2.17	2.22	2.44
2.80	2.31	2.31	2.10	2.30	2.10	2.28	2.21	2.39	1.95	1.94	1.96	2.21	
RMSD	0.283		0.430		0.714		0.922		0.980		1.339		
MSG	0.020		0.080		0.275		0.208		0.266		0.452		
MRD	0.042		0.109		0.158		0.130		0.199		0.273		
MRIG	0.003		0.015		0.041		0.025		0.044		0.087		



ภาพที่ 4.1 เปรียบเทียบค่า RMSD และ MSG ระหว่างแบบสอบคู่ขนานที่คัดเลือกข้อสอบ
โดยกลุ่มผู้เชี่ยวชาญ

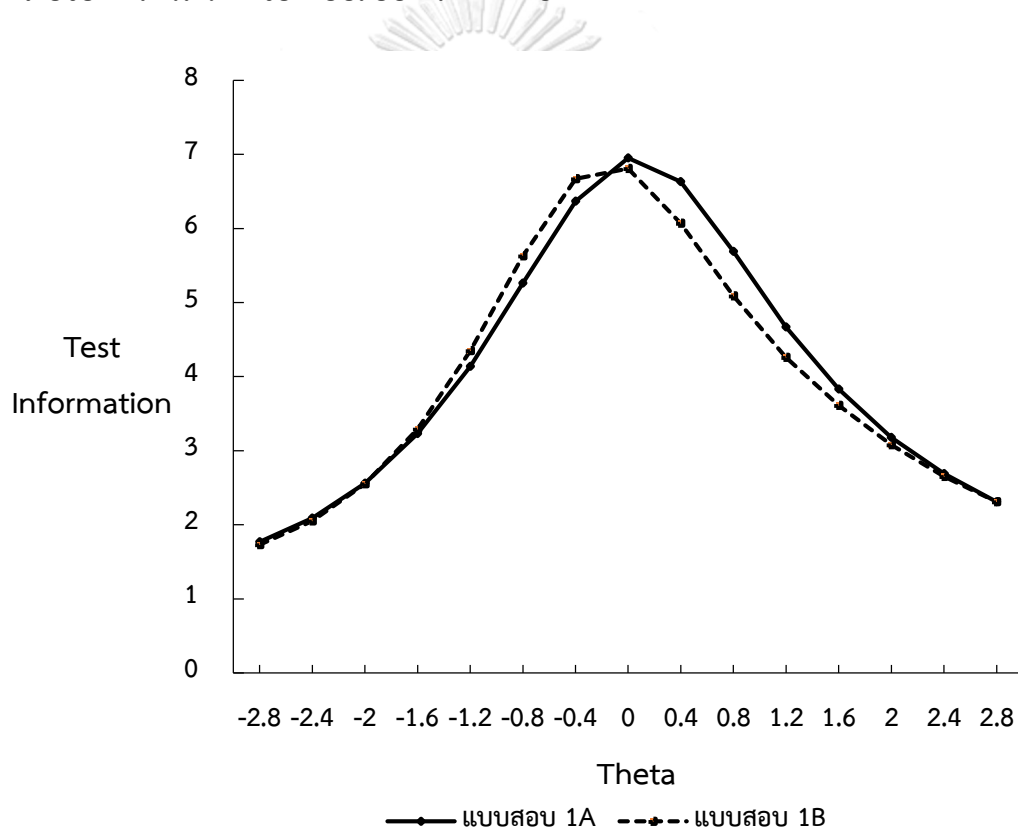
(กำหนดเกณฑ์การแปลผลแบบสอบคู่ขนาน; ค่า RMSD และ MSG ต่ำกว่า 0.50)



ภาพที่ 4.2 เปรียบเทียบค่า MRD และ MRIG ระหว่างแบบสอบคู่ขนานที่คัดเลือกข้อสอบโดย
กลุ่มผู้เชี่ยวชาญ (กำหนดเกณฑ์การแปลผลแบบสอบคู่ขนาน; ค่า MRD และ MRIG ต่ำกว่า 0.10)

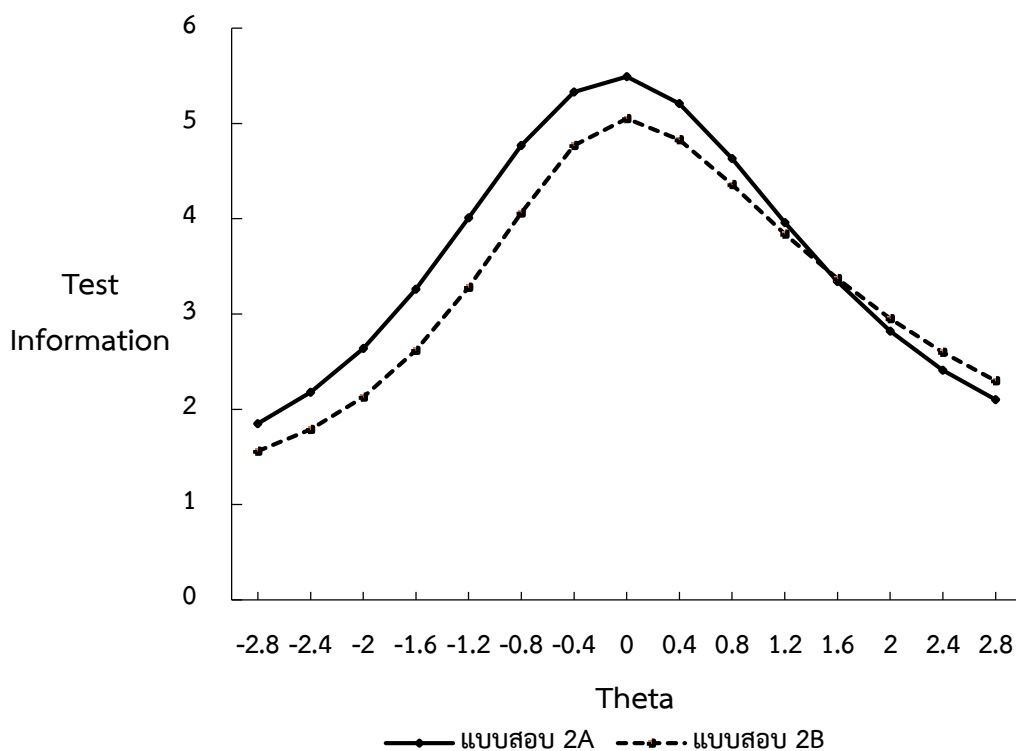
2.2.3 ผลการตรวจสอบความเป็นคู่ขนานของแบบสอบจำแนกตามกลุ่มผู้เชี่ยวชาญที่มีลักษณะแตกต่างกัน

3.1 ผลการตรวจสอบความเป็นคู่ขนานของแบบสอบที่ได้จากการคัดเลือกข้อสอบโดยผู้เชี่ยวชาญในระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบที่ 1 และประสบการณ์สอน ≥ 10 ปี (แบบสอบชุดที่ 1A และ 1B) พบว่า แบบสอบมีความเป็นคู่ขนานกันในทุกดัชนีความเป็นคู่ขนานของแบบสอบ ได้แก่ $RMSD = 0.283$ (ต่ำกว่า 0.50), $MSG = 0.020$ (ต่ำกว่า 0.50), $MRD = 0.042$ (ต่ำกว่า 0.10) และ $MRIG = 0.003$ (ต่ำกว่า 0.10) ณ ตำแหน่งระดับความสามารถของการตอบข้อสอบ $\theta = 0$ ที่ค่าสารสนเทศของข้อสอบสูงสุดเท่ากับ 6.95 โดยมีการกระจายของค่าสารสนเทศของแบบสอบดังภาพที่ 4.3



ภาพที่ 4.3 โด่งสารสนเทศของแบบสอบ (Test information) เปรียบเทียบระหว่างแบบสอบคู่ขนาน ชุด 1A และ 1B ที่ระดับความสามารถของผู้สอบ 15 ระดับ

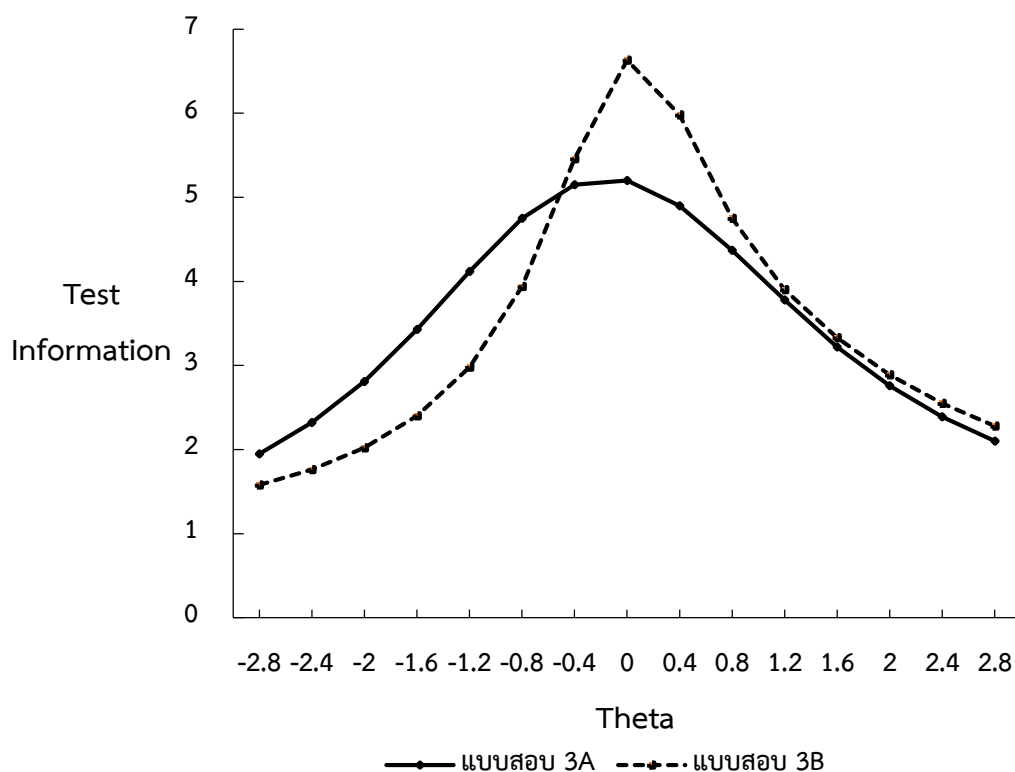
3.2 ผลการตรวจสอบความเป็นคู่ขนานของแบบสอบที่ได้จากการคัดเลือกข้อสอบ โดยผู้เชี่ยวชาญในระดับความสามารถในการตัดสินใจตัดสินความเป็นคู่ขนานของข้อสอบที่ 1 และ ประสบการณ์สอน < 10 ปี (แบบสอบชุดที่ 2A และ 2B) พบว่า แบบสอบมีความเป็นคู่ขนานกันในระดับนี้ความเป็นคู่ขนานของแบบสอบ ได้แก่ $RMSD = 0.430$ (ต่ำกว่า 0.50), $MSG = 0.080$ (ต่ำกว่า 0.50) แต่พบว่า แบบสอบไม่มีความเป็นคู่ขนานกันเมื่อพิจารณาค่า $MRD = 0.109$ (สูงกว่า 0.10) แม้ว่าแบบสอบมีค่า $MRIG = 0.015$ (ต่ำกว่า 0.10) ณ ตำแหน่งระดับความสามารถของการตอบข้อสอบ $\theta = 0$ ที่ค่าสารสนเทศของข้อสอบสูงสุดเท่ากับ 5.49 โดยมีการกระจายของค่าสารสนเทศของแบบสอบดังภาพที่ 4.4



ภาพที่ 4.4 โด่งสารสนเทศของแบบสอบ (Test information) เปรียบเทียบระหว่างแบบสอบ คู่ขนาน ชุด 2A และ 2B ที่ระดับความสามารถของผู้สอบ 15 ระดับ

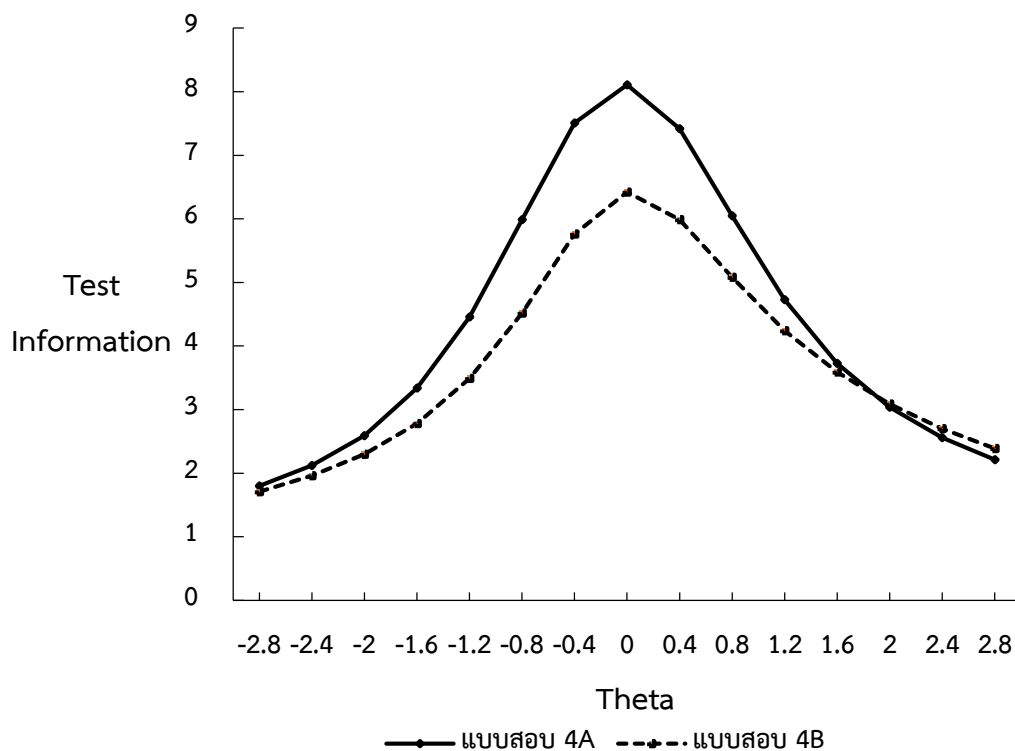
3.3 ผลการตรวจสอบความเป็นคู่ขนานของแบบสอบที่ได้จากการคัดเลือกข้อสอบ โดยผู้เชี่ยวชาญในระดับความสามารถในการตัดสินใจตัดสินความเป็นคู่ขนานของข้อสอบที่ 2 และ ประสบการณ์สอน ≥ 10 ปี (แบบสอบชุดที่ 3A และ 3B) พบว่า แบบสอบไม่มีความเป็นคู่ขนานกันในระดับนี้ความเป็นคู่ขนานของแบบสอบ ได้แก่ $RMSD = 0.714$ (สูงกว่า 0.50) แม้ว่าแบบสอบจะมีค่า $MSG = 0.275$ (ต่ำกว่า 0.50), ค่า $MRD = 0.158$ (สูงกว่า 0.10) แม้ว่าแบบสอบจะมีค่า

MRIG = 0.041 (ต่ำกว่า 0.10) ณ ตำแหน่งระดับความสามารถของการตอบข้อสอบ $\theta = 0$ ที่ค่าสารสนเทศของข้อสอบสูงสุดเท่ากับ 6.63 โดยมีการกระจายของค่าสารสนเทศของแบบสอบดังภาพที่ 4.5



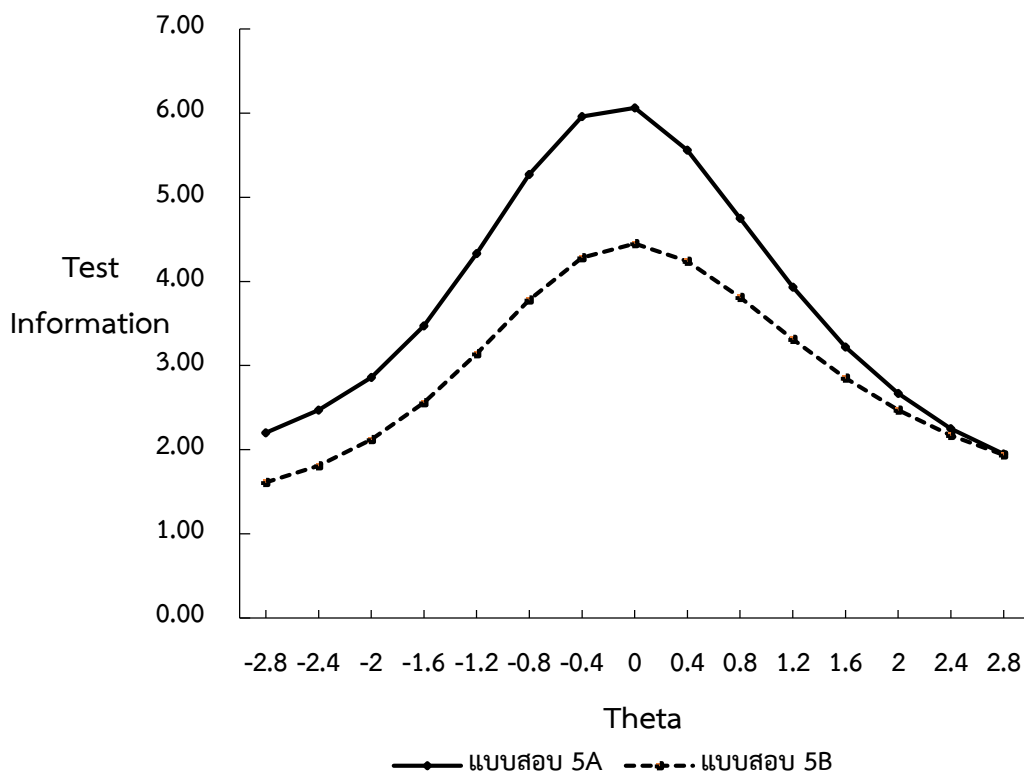
ภาพที่ 4.5 โค้งสารสนเทศของแบบสอบ (Test information) เปรียบเทียบระหว่างแบบสอบคู่ขนานชุด 3A และ 3B ที่ระดับความสามารถของผู้สอบ 15 ระดับ

3.4 ผลการตรวจสอบความเป็นคู่ขนานของแบบสอบที่ได้จากการคัดเลือกข้อสอบโดยผู้เชี่ยวชาญในระดับความสามารถในการตัดสินใจความเป็นคู่ขนานของข้อสอบที่ 2 และประสบการณ์สอน < 10 ปี (แบบสอบชุดที่ 4A และ 4B) พบว่า แบบสอบไม่มีความเป็นคู่ขนานกันในด้านนี้ความเป็นคู่ขนานของแบบสอบ ได้แก่ RMSD = 0.922 (สูงกว่า 0.50) แม้ว่าแบบสอบจะมีค่า MSG = 0.208 (ต่ำกว่า 0.50), ค่า MRD = 0.130 (สูงกว่า 0.10) แม้ว่าแบบสอบจะมีค่า MRIG = 0.025 (ต่ำกว่า 0.10) ณ ตำแหน่งระดับความสามารถของการตอบข้อสอบ $\theta = 0$ ที่ค่าสารสนเทศของข้อสอบสูงสุดเท่ากับ 8.11 โดยมีการกระจายของค่าสารสนเทศของแบบสอบดังภาพที่ 4.6



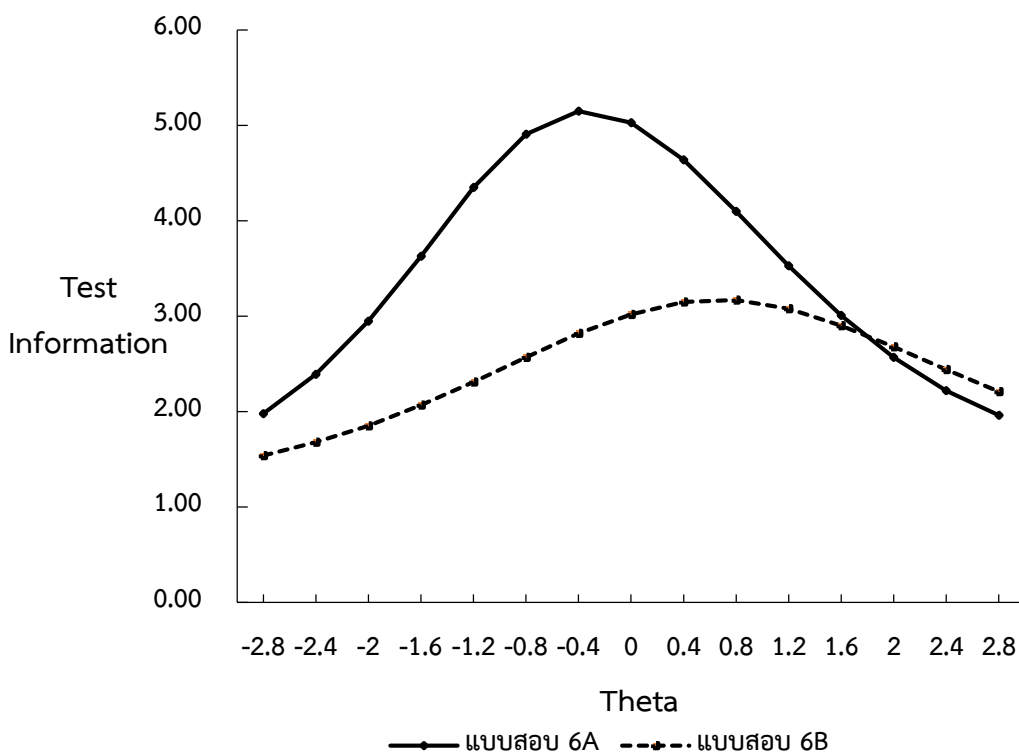
ภาพที่ 4.6 โคน้สารสนเทศของแบบสอบ (Test information) เปรียบเทียบระหว่างแบบสอบ
คู่ขนาน ชุด 4A และ 4B ที่ระดับความสามารถของผู้สอบ 15 ระดับ

3.5 ผลการตรวจสอบความเป็นคู่ขนานของแบบสอบที่ได้จากการคัดเลือกข้อสอบ โดยผู้เชี่ยวชาญในระดับความสามารถในการตัดสินใจความเป็นคู่ขนานของข้อสอบที่ 3 และ ประสบการณ์สอน ≥ 10 ปี (แบบสอบชุดที่ 5A และ 5B) พบว่า แบบสอบไม่มีความเป็นคู่ขนานกัน ใน ดัชนีความเป็นคู่ขนานของแบบสอบ ได้แก่ RMSD = 0.980 (สูงกว่า 0.50) แม้ว่าแบบสอบจะมีค่า MSG = 0.266 (ต่ำกว่า 0.50), ค่า MRD = 0.199 (สูงกว่า 0.10) แม้ว่าแบบสอบจะมีค่า MRIG = 0.044 (ต่ำกว่า 0.10) ณ ตำแหน่งระดับความสามารถของการตอบข้อสอบ $\theta = 0$ ที่ค่าสารสนเทศ ของข้อสอบสูงสุดเท่ากับ 6.06 โดยมีการกระจายของค่าสารสนเทศของแบบสอบดังภาพที่ 4.7



ภาพที่ 4.7 โค้งสารสนเทศของแบบสอบ (Test information) เปรียบเทียบระหว่างแบบสอบ
คู่ขนาน ชุด 5A และ 5B ที่ระดับความสามารถของผู้สอบ 15 ระดับ

3.6 ผลการตรวจสอบความเป็นคู่ขนานของแบบสอบที่ได้จากการคัดเลือกข้อสอบ โดยผู้เชี่ยวชาญในระดับความสามารถในการตัดสินใจเป็นคู่ขนานของข้อสอบที่ 3 และ ประสิทธิภาพสอน < 10 ปี (แบบสอบชุดที่ 6A และ 6B) พบว่า แบบสอบไม่มีความเป็นคู่ขนานกันในด้านนี้ความเป็นคู่ขนานของแบบสอบ ได้แก่ $RMSD = 1.339$ (สูงกว่า 0.50) แม้ว่าแบบสอบจะมีค่า $MSG = 0.452$ (ต่ำกว่า 0.50), ค่า $MRD = 0.273$ (สูงกว่า 0.10) แม้ว่าแบบสอบจะมีค่า $MRIG = 0.087$ (ต่ำกว่า 0.10) ณ ตำแหน่งระดับความสามารถของการตอบข้อสอบ $\theta = -0.4$ ที่ค่าสารสนเทศของข้อสอบสูงสุดเท่ากับ 5.15 โดยมีการกระจายของค่าสารสนเทศของแบบสอบดังภาพที่ 4.8



ภาพที่ 4.8 โควงสารสนเทศของแบบสอบ (Test information) เปรียบเทียบระหว่างแบบสอบ
 คู่ขนาน ชุด 6A และ 6B ที่ระดับความสามารถของผู้สอบ 15 ระดับ

2.3 ผลการเปรียบเทียบดัชนีความเป็นคู่ขนานของแบบสอบและข้อสอบที่คัดเลือกโดย
 กลุ่มผู้เชี่ยวชาญที่มีลักษณะแตกต่างกัน ใช้สถิติวิเคราะห์โดยการเปรียบเทียบค่าเฉลี่ยของ MRD ที่
 ได้จากแบบสอบ

จากการวิเคราะห์เปรียบเทียบค่าเฉลี่ยของ MRD ที่ได้จากแบบสอบทั้งฉบับจำแนกตามกลุ่ม
 ผู้เชี่ยวชาญที่มีลักษณะแตกต่างกัน 6 กลุ่ม พบว่า ค่าเฉลี่ยของ MRD ที่ได้จากแบบสอบที่คัดเลือกโดย
 กลุ่มผู้เชี่ยวชาญทั้ง 6 กลุ่มมีอย่างน้อย 1 คู่ที่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05

เมื่อทำการวิเคราะห์เปรียบเทียบรายคู่ พบว่า กลุ่มผู้เชี่ยวชาญในระดับความสามารถในการ
 ตัดสินความเป็นคู่ขนานของข้อสอบที่ 1 และประสบการณ์สอน ≥ 10 ปี มีค่าเฉลี่ยของ MRD ต่ำกว่า
 กลุ่มผู้เชี่ยวชาญอื่นๆทุกกลุ่ม และกลุ่มผู้เชี่ยวชาญในระดับความสามารถในการตัดสินความเป็น
 คู่ขนานของข้อสอบที่ 1 และประสบการณ์สอน < 10 ปี มีค่าเฉลี่ยของ MRD ต่ำกว่ากลุ่มผู้เชี่ยวชาญ
 ในระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบที่ 3 และประสบการณ์สอน < 10 ปี
 ดังตารางที่ 4.13

ตารางที่ 4.13 ผลการวิเคราะห์เปรียบเทียบค่าเฉลี่ยของ MRD ที่ได้จากแบบสอบทั้งฉบับจำแนกตามกลุ่มผู้เชี่ยวชาญที่มีลักษณะแตกต่างกัน 6 กลุ่ม

ค่าเฉลี่ยของ MRD จำแนกตามกลุ่มผู้เชี่ยวชาญ						ค่าสถิติทดสอบความแตกต่างของค่าเฉลี่ย				ผลการเปรียบเทียบรายคู่	p
กลุ่ม 1	กลุ่ม 2	กลุ่ม 3	กลุ่ม 4	กลุ่ม 5	กลุ่ม 6	Levene Statistic	p	F	p		
0.042	0.109	0.158	0.130	0.199	0.273	11.801	0.001	10.087	0.001	กลุ่ม 1 < กลุ่ม 2	0.021
										กลุ่ม 1 < กลุ่ม 3	0.009
										กลุ่ม 1 < กลุ่ม 4	0.010
										กลุ่ม 1 < กลุ่ม 5	0.001
										กลุ่ม 1 < กลุ่ม 6	0.001
										กลุ่ม 2 < กลุ่ม 6	0.020

หมายเหตุ

กลุ่มที่ 1 คือ ผู้เชี่ยวชาญในระดับความสามารถในการตัดสินใจด้านความเป็นคู่ขนานของข้อสอบที่ 1 และประสบการณ์สอน ≥ 10 ปี

กลุ่มที่ 2 คือ ผู้เชี่ยวชาญในระดับความสามารถในการตัดสินใจด้านความเป็นคู่ขนานของข้อสอบที่ 1 และประสบการณ์สอน < 10 ปี

กลุ่มที่ 3 คือ ผู้เชี่ยวชาญในระดับความสามารถในการตัดสินใจด้านความเป็นคู่ขนานของข้อสอบที่ 2 และประสบการณ์สอน ≥ 10 ปี

กลุ่มที่ 4 คือ ผู้เชี่ยวชาญในระดับความสามารถในการตัดสินใจด้านความเป็นคู่ขนานของข้อสอบที่ 2 และประสบการณ์สอน < 10 ปี

กลุ่มที่ 5 คือ ผู้เชี่ยวชาญในระดับความสามารถในการตัดสินใจด้านความเป็นคู่ขนานของข้อสอบที่ 3 และประสบการณ์สอน ≥ 10 ปี

กลุ่มที่ 6 คือ ผู้เชี่ยวชาญในระดับความสามารถในการตัดสินใจด้านความเป็นคู่ขนานของข้อสอบที่ 3 และประสบการณ์สอน < 10 ปี

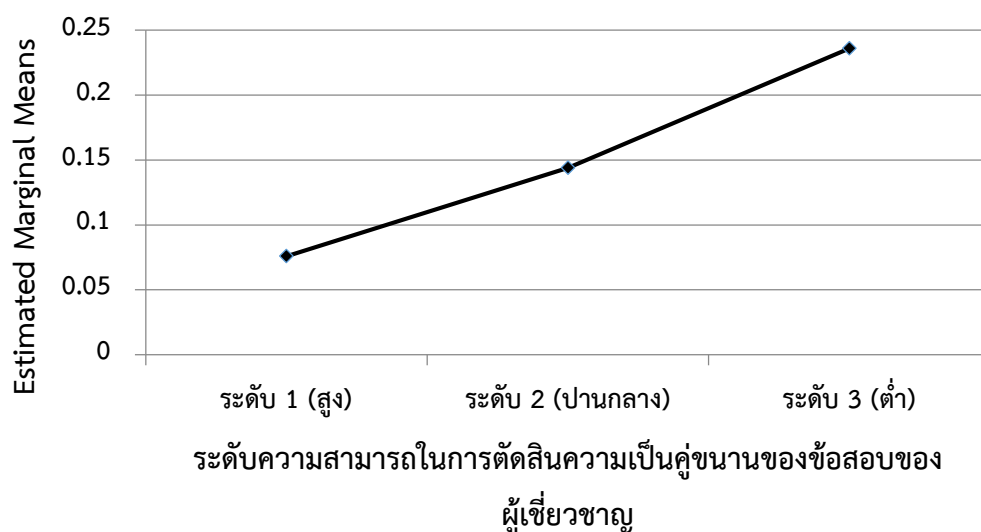
เมื่อทำการวิเคราะห์เปรียบเทียบค่าเฉลี่ย โดยการวิเคราะห์ความแปรปรวนของ MRD จำแนกตามระดับความสามารถในการตัดสินใจด้านความเป็นคู่ขนานของข้อสอบและประสบการณ์สอน พบว่าปฏิสัมพันธ์ระหว่างระดับความสามารถในการตัดสินใจด้านความเป็นคู่ขนานของข้อสอบและประสบการณ์สอนไม่มีอิทธิพลต่อค่า MRD อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 หมายความว่า ระดับความสามารถในการตัดสินใจด้านความเป็นคู่ขนานของข้อสอบและประสบการณ์สอนไม่มีอิทธิพลร่วมกันในการส่งผลต่อค่า MRD ที่แตกต่างกัน ทั้งนี้พบว่า ระดับความสามารถในการตัดสินใจด้านความเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญสูงจะมีค่าเฉลี่ยของค่า MRD ที่น้อยกว่าระดับต่ำกว่า ปรากฏดังภาพที่ 4.9

ตารางที่ 4.14 ผลการวิเคราะห์ความแปรปรวนของ MRD จำแนกตามระดับความสามารถในการตัดสินใจเป็นคู่ขนานของข้อสอบและประสบการณ์สอน

แหล่งความแปรปรวน	Type III Sum of Squares	df	Mean Square	F	p	Levene's test
โมเดล	0.470	5	0.094	10.087	0.001*	F=11.801 df1=5 df2 = 84 p =0.001
ระดับความสามารถในการตัดสินใจเป็นคู่ขนานของข้อสอบ	0.390	2	0.195	20.909	0.001*	
ประสบการณ์สอน	0.032	1	0.032	3.452	0.067	
ระดับความสามารถในการตัดสินใจเป็นคู่ขนานของข้อสอบ*ประสบการณ์สอน	0.048	2	0.024	2.583	0.082	
ความคลาดเคลื่อน	0.783	84	0.009			
รวม	3.332	90				

*p<0.05

Estimated Marginal Means of MRD



ภาพที่ 4.9 กราฟเปรียบเทียบค่าเฉลี่ยของ MRD จำแนกตามระดับความสามารถในการตัดสินใจเป็นคู่ขนานของข้อสอบ

ตอนที่ 3 เพื่อเปรียบเทียบดัชนีความเป็นคู่ขนานของข้อสอบคู่ขนานที่คัดเลือกโดยกลุ่มผู้เชี่ยวชาญที่มีระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบและประสบการณ์สอนนักศึกษาแพทย์ที่แตกต่างกัน

การตรวจสอบความเป็นคู่ขนานรายข้อ ให้สารสนเทศที่สำคัญต่อการพัฒนาและปรับปรุงคลังข้อสอบ และเป็นการตรวจสอบความสอดคล้องระหว่างกลุ่มผู้เชี่ยวชาญที่มีระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบกับดัชนีความเป็นคู่ขนานของข้อสอบ ผู้วิจัยได้แบ่งการนำเสนอผลการวิเคราะห์ข้อมูลออกเป็น 2 ส่วนได้แก่ ส่วนที่ 1 ผลการตรวจสอบความเป็นคู่ขนานของข้อสอบโดยอาศัยทฤษฎีการตอบสนองข้อสอบและส่วนที่ 2 ผลการเปรียบเทียบดัชนีความเป็นคู่ขนานของข้อสอบที่นำมาสร้างแบบสอบจากการคัดเลือกโดยกลุ่มผู้เชี่ยวชาญที่มีลักษณะแตกต่างกัน แสดงรายละเอียดดังต่อไปนี้

ส่วนที่ 1 ผลการตรวจสอบความเป็นคู่ขนานของข้อสอบโดยอาศัยทฤษฎีการตอบสนองข้อสอบ

ผลการตรวจสอบความเป็นคู่ขนานของข้อสอบโดยอาศัยทฤษฎีการตอบสนองข้อสอบสามารถจำแนกการพิจารณาตามดัชนีความเป็นคู่ขนานของแบบสอบ มีรายละเอียดดังต่อไปนี้

1.1 ผลการตรวจสอบความเป็นคู่ขนานของข้อสอบโดยใช้ RMSD และ MSG

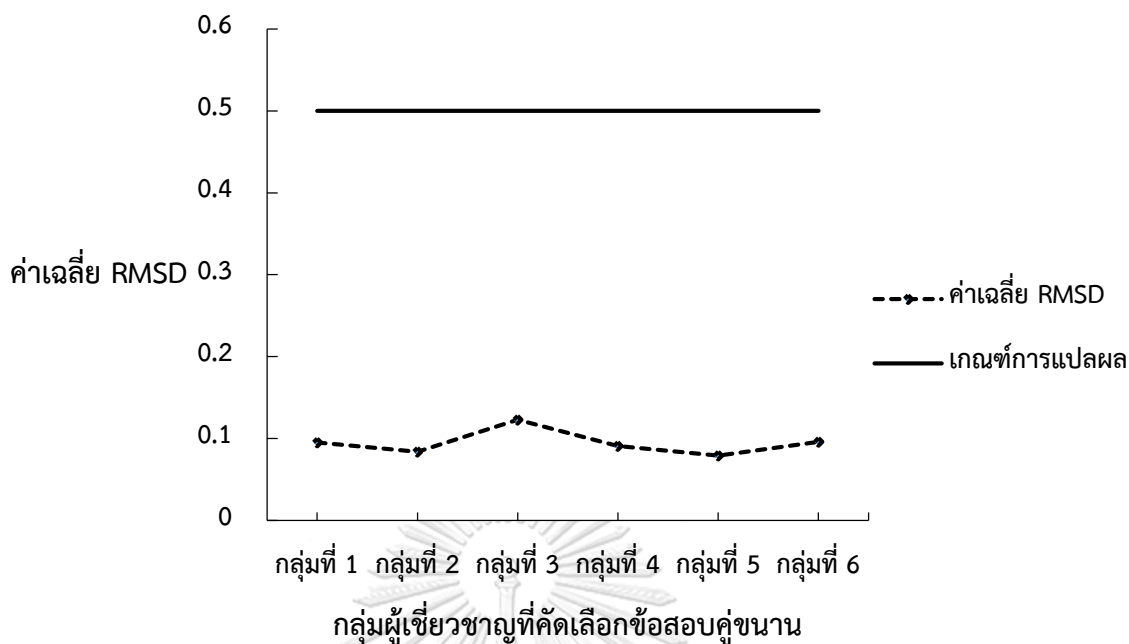
จากการตรวจสอบความเป็นคู่ขนานของข้อสอบโดยใช้ RMSD และ MSG ซึ่งกำหนดเกณฑ์สำหรับแบบสอบคู่ขนานต้องมีค่า RMSD และ MSG ต่ำกว่า 0.50 พบว่า แบบสอบคู่ขนานที่คัดเลือกโดยกลุ่มผู้เชี่ยวชาญส่วนใหญ่มีความเป็นคู่ขนานกันของข้อสอบมากกว่าร้อยละ 95 โดยทุกกลุ่มมีความเป็นคู่ขนานของข้อสอบร้อยละ 100 จะมีเพียงกลุ่มที่ 3 เท่านั้นที่มีความเป็นคู่ขนานของข้อสอบร้อยละ 98 ดังตารางที่ 4.15

1.2 ผลการตรวจสอบความเป็นคู่ขนานของแบบสอบโดยใช้ค่าเฉลี่ย RMSD

จากการตรวจสอบความเป็นคู่ขนานของแบบสอบโดยใช้ ค่าเฉลี่ย RMSD กำหนดเกณฑ์สำหรับแบบสอบคู่ขนานต้องมีค่าเฉลี่ย RMSD ต่ำกว่า 0.50 พบว่า แบบสอบคู่ขนานที่คัดเลือกโดยกลุ่มผู้เชี่ยวชาญทุกกลุ่มมีค่าเฉลี่ย RMSD ต่ำกว่า 0.50 (ค่าเฉลี่ย RMSD อยู่ในช่วง 0.079 ถึง 0.123) แสดงว่า ข้อสอบที่ถูกคัดเลือกโดยกลุ่มผู้เชี่ยวชาญเมื่อนำมาสร้างเป็นแบบสอบแล้วในภาพรวมมีความเป็นคู่ขนานกันในทุกแบบสอบ ดังตารางที่ 4.15 และแสดงการเปรียบเทียบค่าเฉลี่ย RMSD ระหว่างกลุ่มดังภาพที่ 4.10

ตารางที่ 4.15 ผลการตรวจสอบความเป็นคู่ขนานของข้อสอบและแบบสอบโดยอาศัยทฤษฎีการตอบสนองข้อสอบ

กลุ่ม	1	2	3	4	5	6	
ชุดแบบสอบ	1A-1B	2A-2B	3A-3B	4A-4B	5A-5B	6A-6B	
ระดับความสามารถในการตัดสิน ความเป็นคู่ขนานของข้อสอบ	ระดับที่ 1	ระดับที่ 1	ระดับที่ 2	ระดับที่ 2	ระดับที่ 3	ระดับที่ 3	
ประสบการณ์สอน	≥ 10 ปี	< 10 ปี	≥ 10 ปี	< 10 ปี	≥ 10 ปี	< 10 ปี	
RMSD	ข้อ (คู่ที่)						
	1	0.026	0.009	0.000	0.034	0.044	0.008
	2	0.408	0.010	0.099	0.100	0.035	0.134
	3	0.043	0.124	0.033	0.152	0.233	0.024
	4	0.060	0.016	0.044	0.045	0.010	0.039
	5	0.199	0.032	0.023	0.066	0.075	0.034
	6	0.056	0.146	0.039	0.034	0.057	0.291
	7	0.019	0.083	0.125	0.097	0.158	0.045
	8	0.343	0.103	0.824	0.151	0.090	0.071
	9	0.058	0.220	0.244	0.064	0.006	0.006
	10	0.076	0.088	0.118	0.050	0.013	0.099
	11	0.012	0.008	0.028	0.058	0.024	0.023
	12	0.007	0.219	0.136	0.068	0.079	0.031
	13	0.073	0.064	0.206	0.083	0.084	0.153
	14	0.118	0.028	0.020	0.082	0.046	0.143
	15	0.044	0.051	0.047	0.052	0.009	0.144
	16	0.000	0.035	0.018	0.090	0.092	0.035
	17	0.063	0.039	0.189	0.015	0.027	0.008
	18	0.042	0.051	0.230	0.005	0.311	0.297
	19	0.027	0.135	0.055	0.173	0.065	0.180
	20	0.034	0.070	0.050	0.038	0.018	0.149
	21	0.143	0.067	0.131	0.043	0.208	0.035
	22	0.036	0.272	0.093	0.335	0.078	0.139
	23	0.061	0.041	0.052	0.031	0.186	0.158
	24	0.367	0.081	0.050	0.369	0.028	0.018
	25	0.074	0.093	0.208	0.037	0.003	0.144
ร้อยละของข้อสอบคู่ขนาน	100.00	100.00	98.00	100.00	100.00	100.00	
ค่าเฉลี่ย RMSD (\overline{RMSD})	0.095	0.084	0.123	0.091	0.079	0.096	



ภาพที่ 4.10 เปรียบเทียบค่าเฉลี่ย RMSD ระหว่างแบบสอบคู่ขนานที่คัดเลือกข้อสอบโดยผู้เชี่ยวชาญ 6 กลุ่ม (กำหนดเกณฑ์การแปลผลแบบสอบคู่ขนาน; ค่าเฉลี่ย RMSD ต่ำกว่า 0.50)

ส่วนที่ 2 ผลการเปรียบเทียบดัชนีความเป็นคู่ขนานของข้อสอบที่นำมาสร้างแบบสอบจากการคัดเลือกโดยกลุ่มผู้เชี่ยวชาญที่มีลักษณะแตกต่างกัน

ผลการเปรียบเทียบดัชนีความเป็นคู่ขนานของแบบสอบและข้อสอบที่คัดเลือกโดยกลุ่มผู้เชี่ยวชาญที่มีลักษณะแตกต่างกัน ใช้สถิติวิเคราะห์โดยการเปรียบเทียบค่าเฉลี่ยของ RMSD ที่ได้จากรMSD ของข้อสอบรายข้อ (**RMSD**) มีรายละเอียดดังต่อไปนี้

จากการวิเคราะห์เปรียบเทียบค่าเฉลี่ยของ RMSD ที่ได้จากแบบสอบทั้งฉบับจำแนกตามกลุ่มผู้เชี่ยวชาญที่มีลักษณะแตกต่างกัน 6 กลุ่ม พบว่า แบบสอบที่ได้จากการคัดเลือกข้อสอบโดยกลุ่มผู้เชี่ยวชาญทั้ง 6 กลุ่มมีค่าเฉลี่ยของ RMSD ไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 ดังตารางที่ 4.16

ตารางที่ 4.16 ผลการวิเคราะห์เปรียบเทียบค่าเฉลี่ยของ RMSD ที่ได้จากแบบสอบทั้งฉบับจำแนกตามกลุ่มผู้เชี่ยวชาญที่มีลักษณะแตกต่างกัน 6 กลุ่ม

ค่าเฉลี่ยของ RMSD จำแนกตามกลุ่มผู้เชี่ยวชาญ						ค่าสถิติทดสอบความแตกต่างของค่าเฉลี่ย			
กลุ่ม	กลุ่ม	กลุ่ม	กลุ่ม	กลุ่ม	กลุ่ม	Levene Statistic	p	F	p
1	2	3	4	5	6				
0.095	0.084	0.123	0.091	0.079	0.096	1.145	0.340	0.528	0.755

หมายเหตุ

กลุ่มที่ 1 คือ ผู้เชี่ยวชาญในระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบที่ 1 และประสบการณ์สอน ≥ 10 ปี

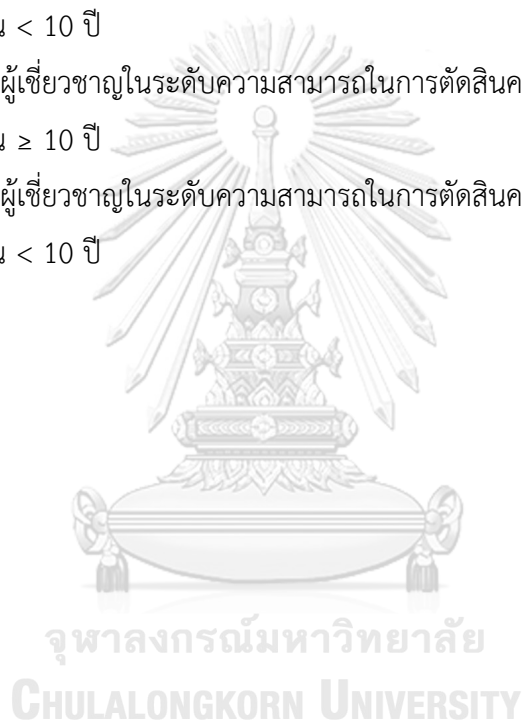
กลุ่มที่ 2 คือ ผู้เชี่ยวชาญในระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบที่ 1 และประสบการณ์สอน < 10 ปี

กลุ่มที่ 3 คือ ผู้เชี่ยวชาญในระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบที่ 2 และประสบการณ์สอน ≥ 10 ปี

กลุ่มที่ 4 คือ ผู้เชี่ยวชาญในระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบที่ 2 และประสบการณ์สอน < 10 ปี

กลุ่มที่ 5 คือ ผู้เชี่ยวชาญในระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบที่ 3 และประสบการณ์สอน ≥ 10 ปี

กลุ่มที่ 6 คือ ผู้เชี่ยวชาญในระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบที่ 3 และประสบการณ์สอน < 10 ปี



บทที่ 5

สรุปผล อภิปรายและข้อเสนอแนะ

การวิจัยในครั้งนี้มีจุดประสงค์เพื่อ 1) เพื่อศึกษาระดับความสามารถในการตัดสินใจเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญ 2) เพื่อเปรียบเทียบดัชนีความเป็นคู่ขนานของแบบสอบคู่ขนานที่ได้จากการคัดเลือกโดยกลุ่มผู้เชี่ยวชาญที่มีระดับความสามารถในการตัดสินใจเป็นคู่ขนานของข้อสอบและประสบการณ์สอนนักศึกษาแพทย์ที่แตกต่างกันและ 3) เพื่อเปรียบเทียบดัชนีความเป็นคู่ขนานของข้อสอบคู่ขนานที่คัดเลือกโดยกลุ่มผู้เชี่ยวชาญที่มีระดับความสามารถในการตัดสินใจเป็นคู่ขนานของข้อสอบและประสบการณ์สอนนักศึกษาแพทย์ที่แตกต่างกัน

การดำเนินการวิจัยแบ่งเป็น 3 ส่วน ได้แก่ **ส่วนที่ 1** ศึกษาความสามารถในการตัดสินใจเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญ เป็นการประมาณค่าระดับความสามารถในการตัดสินใจเป็นคู่ขนานของข้อสอบซึ่งประยุกต์ใช้ทฤษฎีฉันทามติทางวัฒนธรรมแบบมีรูปแบบและการวิเคราะห์ฉันทามติ แหล่งข้อมูลสำหรับการวิจัยในส่วนนี้เป็นตัวอย่างผู้เชี่ยวชาญที่มีประสบการณ์สอนนักศึกษาแพทย์ วิทยาลัยแพทยศาสตร์พระมงกุฎเกล้าจำนวน 63 คนที่ได้จากการเลือกแบบเจาะจงและตอบรับในการให้ความร่วมมือในการเก็บรวบรวมข้อมูล เครื่องมือที่ใช้ ได้แก่ แบบวัดระดับความสามารถในการตัดสินใจเป็นคู่ขนานของข้อสอบที่ประกอบด้วย ข้อสอบจำนวน 20 คู่โดยให้ผู้เชี่ยวชาญพิจารณาตัดสินใจเป็นคู่ขนานของข้อสอบแต่ละข้ออย่างอิสระโดยไม่ถามความคิดเห็นระหว่างกันและไม่เป็นการประชุมกลุ่มย่อยตามข้อตกลงเบื้องต้นของทฤษฎีฉันทามติทางวัฒนธรรม การวิเคราะห์ข้อมูลแบ่งออกเป็น 4 ส่วน ดังนี้ 1) ข้อมูลภูมิหลังของผู้เชี่ยวชาญที่ทำแบบวัดระดับความสามารถในการตัดสินใจเป็นคู่ขนานของข้อสอบ ใช้สถิติบรรยายโดยนำเสนอข้อมูลเป็นจำนวนและร้อยละในรูปแบบตารางไขว้ 2) ตรวจสอบระดับความสามารถในการตัดสินใจเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญโดยอาศัยทฤษฎีฉันทามติทางวัฒนธรรมแบบมีรูปแบบ โดยตรวจสอบข้อตกลงเบื้องต้นของทฤษฎีฉันทามติทางวัฒนธรรม 3 ด้าน คือ ความเป็นเอกมิติของแบบวัดความเป็นอิสระต่อกันของผู้ตอบและความถูกต้อง และประมาณค่าความสามารถในการตัดสินใจเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญจากการวิเคราะห์ฉันทามติโดยอาศัยโปรแกรม UCINET 3) แบ่งกลุ่มผู้เชี่ยวชาญจำแนกตามระดับความสามารถในการตัดสินใจเป็นคู่ขนานของข้อสอบและประสบการณ์สอนนักศึกษาแพทย์ และ 4) ตัดสินคำตอบที่ถูกต้องจากการวิเคราะห์ฉันทามติจากการตัดสินใจเป็นคู่ขนานของข้อสอบโดยผู้เชี่ยวชาญ

ส่วนที่ 2 เปรียบเทียบดัชนีความเป็นคู่ขนานของแบบสอบคู่ขนานที่ได้จากการคัดเลือกโดยกลุ่มผู้เชี่ยวชาญที่มีระดับความสามารถในการตัดสินใจเป็นคู่ขนานของข้อสอบและประสบการณ์

สอนที่แตกต่างกัน ตัวอย่างที่ใช้ คือ ผู้เชี่ยวชาญจำนวน 18 คน แบ่งออกเป็น 6 กลุ่มๆละ 3 คนตามระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญ (ระดับที่ 1, 2 และ 3) และประสบการณ์สอนนักศึกษาแพทย์ (ตั้งแต่ 10 ปีขึ้นไปและน้อยกว่า 10 ปี) โดยให้กลุ่มผู้เชี่ยวชาญคัดเลือกข้อสอบเพื่อสร้างแบบสอบคู่ขนาน 1 ชุดๆละ 2 แบบสอบๆละ 25 ข้อ โดยใช้เนื้อหาการเลือกใช้ใน 3 ระบบ ประกอบด้วย ยาในระบบหลอดเลือดและหัวใจ ระบบทางเดินอาหาร และระบบต่อมไร้ท่อ โดยจำแนกเป็น 5 จุดประสงค์ๆละ 5 ข้อ ซึ่งในแต่ละจุดประสงค์ประกอบด้วยข้อสอบจำนวนทั้งหมด 40 ข้อ กลุ่มผู้ให้ข้อมูลสำหรับตรวจสอบความเป็นคู่ขนานของแบบสอบและข้อสอบ ได้แก่ นักศึกษาแพทย์วิทยาลัยแพทยศาสตร์พระมงกุฎเกล้า ชั้นปีที่ 3-6 ปีการศึกษา 2560 จำนวน 400 คน โดยทำการสอบแบบสอบคู่ขนานละ 200 คน วิเคราะห์ข้อสอบโดยใช้โปรแกรม IRTpro เพื่อหาค่าสารสนเทศของแบบสอบและพารามิเตอร์ของข้อสอบโดยอาศัยโมเดลการตอบสนองข้อสอบแบบ 2 พารามิเตอร์ และตรวจสอบความสอดคล้องกลมกลืนของข้อสอบกับโมเดล 2PL โดยอาศัยสถิติทดสอบไคสแควร์ และเปรียบเทียบค่าดัชนีความเป็นคู่ขนานของแบบสอบที่สร้างขึ้น โดยพิจารณาค่า $RMSD_{TF}$ และ MRD จากนั้นเปรียบเทียบค่าเฉลี่ยของ $RMSD_{TF}$ และ MRD โดยใช้สถิติทดสอบ one-way ANOVA

ส่วนที่ 3 เปรียบเทียบดัชนีความเป็นคู่ขนานของข้อสอบคู่ขนานที่คัดเลือกโดยกลุ่มผู้เชี่ยวชาญที่มีระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบและประสบการณ์สอนนักศึกษาแพทย์ที่แตกต่างกัน โดยพิจารณาค่า $RMSD_{IF}$ และ ค่าเฉลี่ย $RMSD_{IF}$

สรุปผลการวิจัย

การสรุปผลการวิจัย แยกนำเสนอตามวัตถุประสงค์การวิจัยเป็น 3 ข้อ ได้แก่ 1) ผลการศึกษา ระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญ 2) ผลการเปรียบเทียบดัชนีความเป็นคู่ขนานของแบบสอบคู่ขนานที่ได้จากการคัดเลือกโดยกลุ่มผู้เชี่ยวชาญที่มีระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบและประสบการณ์สอนนักศึกษาแพทย์ที่แตกต่างกัน และ 3) ผลการเปรียบเทียบดัชนีความเป็นคู่ขนานของข้อสอบที่คัดเลือกโดยกลุ่มผู้เชี่ยวชาญที่มีระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบและประสบการณ์สอนนักศึกษาแพทย์ที่แตกต่างกัน ซึ่งในแต่ละข้อมีรายละเอียดดังต่อไปนี้

1. ผลการศึกษาระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญ สามารถสรุปได้ 3 ประเด็นสำคัญ ดังนี้

ประเด็นที่ 1 การตรวจสอบข้อตกลงเบื้องต้นของแบบวัดระดับความสามารถในตัดสินความเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญโดยอาศัยทฤษฎีฉันทามติทางวัฒนธรรมแบบมีรูปแบบ

การตรวจสอบระดับความสามารถในตัดสินความเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญ โดยใช้แบบวัดระดับความสามารถในตัดสินความเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญ พบว่า แบบวัดนี้มีความเป็นอิสระต่อกันของผู้ตอบ (Conditional independence) และความถูกต้อง (Common truth) ตามข้อตกลงเบื้องต้นของทฤษฎีฉันทามติทางวัฒนธรรมแบบมีรูปแบบ แต่ไม่ปฏิบัติตามข้อตกลงเบื้องต้นในด้านความเป็นเอกมิตินี้ของแบบวัด (Homogeneity) จึงพิจารณาตัดสินผู้เชี่ยวชาญที่มีความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบในระดับต่ำออกเนื่องจากส่งผลให้ลักษณะของฉันทามติแบ่งออกเป็นมากกว่า 1 กลุ่ม ซึ่งมีผลทำให้การประมาณค่าระดับความสามารถของผู้เชี่ยวชาญมีความคลาดเคลื่อนเพราะการวิเคราะห์ฉันทามติทางวัฒนธรรมผู้เชี่ยวชาญจะต้องมีความสามารถในการตัดสินใจไปในทิศทางเดียวกันและมีความสอดคล้องกันสูงในกลุ่ม ดังนั้น จึงต้องพิจารณาคัดกลุ่มที่มีความแตกต่างของฉันทามติออกเมื่อตรวจสอบข้อตกลงเบื้องต้นแล้วขาดความเป็นเอกมิตินี้ของแบบวัด เมื่อวิเคราะห์ฉันทามติโดยไม่พิจารณาผู้เชี่ยวชาญที่มีระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบต่ำแล้ว พบว่า แบบวัดระดับความสามารถในตัดสินความเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญมีคุณสมบัติด้านความเป็นเอกมิตินี้ของแบบวัดเป็นไปตามข้อตกลงเบื้องต้นของทฤษฎีฉันทามติทางวัฒนธรรมแบบมีรูปแบบ จึงสามารถนำข้อมูลไปใช้ในการวิเคราะห์ระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบได้

ประเด็นที่ 2 ระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญ

จากการประมาณค่าระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญ โดยอาศัยทฤษฎีฉันทามติทางวัฒนธรรมแบบมีรูปแบบ พบว่า ผู้เชี่ยวชาญส่วนใหญ่มีระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบในระดับที่ 3 มากที่สุด รองลงมาเป็นระดับที่ 2 และระดับที่ 1 ตามลำดับ เมื่อนำมาจัดกลุ่มผู้เชี่ยวชาญโดยพิจารณาจำแนกตามประสบการณ์สอนนักศึกษาแพทย์ พบว่า ผู้เชี่ยวชาญที่มีประสบการณ์สอนน้อยกว่า 10 ปี ส่วนใหญ่มีระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบในระดับที่ 3 มากที่สุด รองลงมาเป็นระดับที่ 2 และระดับที่ 1 ตามลำดับ แตกต่างจากผู้เชี่ยวชาญที่มีประสบการณ์สอนตั้งแต่ 10 ปีขึ้นไปที่มีระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบในระดับที่ 2 มากที่สุด รองลงมาเป็นระดับที่ 3 และระดับที่ 1 ตามลำดับ

ประเด็นที่ 3 คำตอบที่ถูกต้องจากการวิเคราะห์ฉันทามติจากการตัดสินความเป็นคู่ขนานของข้อสอบโดยผู้เชี่ยวชาญ

ผลการตัดสินคำตอบที่ถูกต้องจากการวิเคราะห์ฉันทามติจากการตัดสินความเป็นคู่ขนานของข้อสอบโดยผู้เชี่ยวชาญ เมื่อวิเคราะห์ทางสถิติโดยอาศัยทฤษฎีฉันทามติทางวัฒนธรรมแบบมีรูปแบบ พบว่า ข้อสอบมีความเป็นคู่ขนานกันมากกว่าข้อสอบไม่มีความเป็นคู่ขนานกันในอัตราส่วน 3 ต่อ 1 ซึ่งเป็นไปตามข้อกำหนดของแบบวัดที่เหมาะสมสำหรับการวิเคราะห์ฉันทามติ ซึ่งคำตอบที่ถูกต้องทั้ง 20 ข้อนี้สามารถใช้เป็นสารสนเทศในการประมาณค่าระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญได้

2. ผลการเปรียบเทียบดัชนีความเป็นคู่ขนานของแบบสอบคู่ขนานที่ได้จากการคัดเลือกโดยกลุ่มผู้เชี่ยวชาญที่มีระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบและประสบการณ์สอนนักศึกษาแพทย์ที่แตกต่างกัน สามารถสรุปได้ 4 ประเด็นสำคัญ ดังนี้

ประเด็นที่ 1 ผลการตรวจสอบความเป็นคู่ขนานของแบบสอบโดยใช้ RMSD และ MSG

เมื่อกำหนดให้ดัชนีความเป็นคู่ขนานของแบบสอบโดยใช้ RMSD และ MSG สำหรับแบบสอบคู่ขนานต้องมีค่า RMSD และ MSG ต่ำกว่า 0.50 พบว่า แบบสอบคู่ขนานที่คัดเลือกโดยกลุ่มผู้เชี่ยวชาญที่มีความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบในระดับที่ 1 ทั้งกลุ่มที่มีประสบการณ์สอนตั้งแต่ 10 ปีขึ้นไป (กลุ่มที่ 1) และประสบการณ์สอนน้อยกว่า 10 ปี (กลุ่มที่ 2) มีความเป็นคู่ขนานของแบบสอบ ส่วนแบบสอบที่คัดเลือกโดยกลุ่มผู้เชี่ยวชาญกลุ่มอื่นๆ ไม่มีความเป็นคู่ขนานของแบบสอบ โดยมีแนวโน้มของค่า RMSD และ MSG สูงขึ้นตามระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบของกลุ่มผู้เชี่ยวชาญที่ลดลงและมีประสบการณ์สอนนักศึกษาแพทย์ที่น้อยกว่า 10 ปี

ประเด็นที่ 2 ผลการตรวจสอบความเป็นคู่ขนานของแบบสอบโดยใช้ MRD และ MRIG

เมื่อกำหนดเกณฑ์สำหรับแบบสอบคู่ขนานต้องมีค่า MRD และ MRIG ต่ำกว่า 0.10 (น้อยกว่าร้อยละ 10) พบว่า แบบสอบคู่ขนานที่คัดเลือกโดยกลุ่มผู้เชี่ยวชาญที่มีความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบในระดับที่ 1 ที่มีประสบการณ์สอน ≥ 10 ปี (กลุ่มที่ 1) มีความเป็นคู่ขนานของแบบสอบ ส่วนแบบสอบที่คัดเลือกโดยกลุ่มผู้เชี่ยวชาญกลุ่มอื่นๆ ไม่มีความเป็นคู่ขนานของแบบสอบ โดยมีแนวโน้มของค่า MRD และ MRIG มีค่าสูงขึ้นตามระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบของกลุ่มผู้เชี่ยวชาญที่ลดลงและมีประสบการณ์สอนนักศึกษาแพทย์น้อยกว่า 10 ปี

ประเด็นที่ 3 การเปรียบเทียบดัชนีความเป็นคู่ขนานของแบบสอบและข้อสอบที่คัดเลือกโดยกลุ่มผู้เชี่ยวชาญที่มีลักษณะแตกต่างกัน ใช้สถิติวิเคราะห์โดยการเปรียบเทียบค่าเฉลี่ยของ MRD ที่ได้จากแบบสอบ

เมื่อเปรียบเทียบค่าเฉลี่ยของ MRD ที่ได้จากแบบสอบที่คัดเลือกโดยกลุ่มผู้เชี่ยวชาญทั้ง 6 กลุ่มมีอย่างน้อย 1 คู่ที่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 เมื่อทำการวิเคราะห์เปรียบเทียบรายคู่ พบว่า กลุ่มผู้เชี่ยวชาญในระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบที่ 1 และประสบการณ์สอนตั้งแต่ 10 ปีขึ้นไปมีค่าเฉลี่ยของ MRD ต่ำกว่ากลุ่มผู้เชี่ยวชาญอื่นๆ ทุกกลุ่ม และกลุ่มผู้เชี่ยวชาญในระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบที่ 1 และประสบการณ์สอนน้อยกว่า 10 ปี มีค่าเฉลี่ยของ MRD ต่ำกว่ากลุ่มผู้เชี่ยวชาญในระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบที่ 3 และประสบการณ์สอนน้อยกว่า 10 ปี

ประเด็นที่ 4 การเปรียบเทียบดัชนีความเป็นคู่ขนานของแบบสอบและข้อสอบที่คัดเลือกโดยกลุ่มผู้เชี่ยวชาญที่มีลักษณะแตกต่างกัน โดยการวิเคราะห์ความแปรปรวนของ MRD จำแนกตามระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบและประสบการณ์สอน

จากการวิเคราะห์ความแปรปรวนของ MRD พบว่า ปฏิสัมพันธ์ระหว่างระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบและประสบการณ์สอนไม่มีอิทธิพลต่อค่า MRD อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 แสดงว่า ระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบและประสบการณ์สอนไม่มีอิทธิพลร่วมกันในการส่งผลต่อค่า MRD ที่แตกต่างกัน แต่จากการวิเคราะห์แสดงให้เห็นถึงระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญสูงจะมีค่าเฉลี่ยของค่า MRD ที่น้อยกว่าระดับต่ำกว่าอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05

จุฬาลงกรณ์มหาวิทยาลัย

3. ผลการเปรียบเทียบดัชนีความเป็นคู่ขนานของข้อสอบที่คัดเลือกโดยกลุ่มผู้เชี่ยวชาญที่มีระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบและประสบการณ์สอนนักศึกษาแพทย์ที่แตกต่างกัน สามารถสรุปได้ว่า เมื่อกำหนดเกณฑ์สำหรับแบบสอบคู่ขนานต้องมีค่า RMSD และ MSG ต่ำกว่า 0.50 พบว่า แบบสอบคู่ขนานที่คัดเลือกโดยกลุ่มผู้เชี่ยวชาญส่วนใหญ่มีความเป็นคู่ขนานกันของข้อสอบสอดคล้องกับการตรวจสอบความเป็นคู่ขนานของแบบสอบโดยใช้ค่าเฉลี่ย RMSD แบบสอบคู่ขนานที่คัดเลือกโดยกลุ่มผู้เชี่ยวชาญทุกกลุ่มมีค่าเฉลี่ย RMSD ต่ำกว่า 0.50 จึงสรุปได้ว่า ข้อสอบที่ถูกคัดเลือกโดยกลุ่มผู้เชี่ยวชาญเมื่อนำมาสร้างเป็นแบบสอบแล้วในภาพรวมมีความเป็นคู่ขนานกันของข้อสอบในทุกแบบสอบ เมื่อวิเคราะห์เปรียบเทียบค่าเฉลี่ยของ RMSD ที่ได้จากแบบสอบจำแนกตามกลุ่มผู้เชี่ยวชาญที่มีลักษณะแตกต่างกัน 6 กลุ่ม พบว่า แบบสอบที่ได้จากการคัดเลือกข้อสอบโดยกลุ่มผู้เชี่ยวชาญทั้ง 6 กลุ่มมีค่าเฉลี่ยของ RMSD ไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05

อภิปรายผลการวิจัย

จากผลการเปรียบเทียบดัชนีความเป็นคู่ขนานของข้อสอบและแบบสอบที่ได้จากการคัดเลือกข้อสอบโดยกลุ่มผู้เชี่ยวชาญที่มีลักษณะแตกต่างกันโดยประยุกต์ใช้ทฤษฎีฉันทามติทางวัฒนธรรมแบบมีรูปแบบ มีประเด็นสำคัญที่นำมาอภิปราย 4 ประเด็น ได้แก่ ประเด็นที่ 1 การตรวจสอบข้อตกลงเบื้องต้นของแบบวัดระดับความสามารถในการตัดสินใจความเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญโดยอาศัยทฤษฎีฉันทามติทางวัฒนธรรมแบบมีรูปแบบ ประเด็นที่ 2 การตรวจสอบระดับความสามารถในการตัดสินใจความเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญโดยอาศัยทฤษฎีฉันทามติทางวัฒนธรรมแบบมีรูปแบบ ประเด็นที่ 3 การเปรียบเทียบดัชนีความเป็นคู่ขนานของแบบสอบคู่ขนานที่ได้จากการคัดเลือกโดยกลุ่มผู้เชี่ยวชาญที่มีระดับความสามารถในการตัดสินใจความเป็นคู่ขนานของข้อสอบและประสบการณ์สอนนักศึกษาแพทย์ที่แตกต่างกันและ ประเด็นที่ 4 การเปรียบเทียบดัชนีความเป็นคู่ขนานของข้อสอบคู่ขนานที่คัดเลือกโดยกลุ่มผู้เชี่ยวชาญที่มีระดับความสามารถในการตัดสินใจความเป็นคู่ขนานของข้อสอบและประสบการณ์สอนนักศึกษาแพทย์ที่แตกต่างกัน

ประเด็นที่ 1 การตรวจสอบข้อตกลงเบื้องต้นของแบบวัดระดับความสามารถในการตัดสินใจความเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญโดยอาศัยทฤษฎีฉันทามติทางวัฒนธรรมแบบมีรูปแบบ

การตรวจสอบระดับความสามารถในการตัดสินใจความเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญโดยอาศัยทฤษฎีฉันทามติทางวัฒนธรรมแบบมีรูปแบบ มีข้อตกลงเบื้องต้นที่ต้องตรวจสอบก่อนการวิเคราะห์ฉันทามติจำนวน 3 ข้อ ได้แก่ ความเป็นเอกมิติของแบบวัด (Homogeneity) ความเป็นอิสระต่อกันของผู้ตอบ (Conditional independence) และความถูกต้อง (Common truth) จากการวิจัยในครั้งนี้พบว่า ข้อตกลงเบื้องต้นในด้านความเป็นอิสระต่อกันของผู้ตอบและความถูกต้อง เป็นข้อตกลงที่สามารถวางแผนและควบคุมการวัดให้เป็นไปตามข้อตกลงได้โดยกำหนดให้ผู้เชี่ยวชาญตอบคำถามแบบวัดระดับความสามารถในการตัดสินใจความเป็นคู่ขนานของข้อสอบโดยไม่สามารถสอบถามความคิดเห็นระหว่างกัน ไม่มีการปรึกษากันและไม่ใช้การสัมภาษณ์กลุ่มย่อย ซึ่งบริหารการเก็บรวบรวมข้อมูลอย่างเป็นระบบและป้องกันการละเมิดข้อตกลงได้ รวมถึงการสร้างข้อคำถามโดยกำหนดให้คำตอบของแบบวัดระดับความสามารถในการตัดสินใจความเป็นคู่ขนานของข้อสอบมีเพียงคำตอบเดียวเท่านั้น คือ ข้อสอบมีความเป็นคู่ขนานกันหรือข้อสอบไม่มีความเป็นคู่ขนานกัน จึงทำให้มีคุณสมบัติด้านความถูกต้องเป็นไปตามข้อตกลงเบื้องต้นของทฤษฎีฉันทามติทางวัฒนธรรมแบบมีรูปแบบ ทั้งนี้เหมาะสำหรับการประยุกต์ใช้ในสภาพจริงของการสร้างแบบวัดหรือแบบสอบโดยไม่สามารถทราบคำตอบที่แท้จริงก่อนการนำไปใช้ทดสอบแต่มีลักษณะของคำตอบที่ชัดเจนเพียงคำตอบเดียวเป็นไปตามคุณสมบัติด้านความถูกต้อง สอดคล้องกับงานวิจัยของ S. P. Borgatti and Halgin (2011);

Chavez (1995); Maher (1987) ที่พบว่าข้อตกลงเบื้องต้น 2 ข้อดังกล่าวจะต้องทำการควบคุมโดยผู้วิจัยหรือผู้มีส่วนเกี่ยวข้องในการบริหารการสอบได้เพื่อลดการละเมิดข้อตกลงเบื้องต้นดังกล่าว แต่พบว่า ข้อตกลงเบื้องต้นด้านความเป็นเอกมิตของแบบวัด เป็นข้อตกลงเบื้องต้นที่สำคัญเนื่องจากบ่งบอกถึงความสอดคล้องกันของผู้ลงฉันทามติในประชากรหรือตัวอย่างที่ทำการให้ข้อมูลและเป็นปัจจัยสำคัญที่ส่งผลต่อการวิเคราะห์ระดับความสามารถในการตัดสินใจหรือความเห็นพ้องร่วมกัน เพราะหากมีระดับความสามารถดังกล่าวแบ่งออกเป็น 2 กลุ่มหรือมากกว่า 2 กลุ่มจะทำให้ฉันทามตินั้นเกิดการโต้แย้งระหว่างกลุ่มได้และผลของการตัดสินใจคำตอบนั้นๆโดยการวิเคราะห์ฉันทามติจะขาดความน่าเชื่อถือ ในการวิจัยครั้งนี้พบว่า เมื่อวิเคราะห์ระดับความสามารถในการตัดสินใจเป็นคู่ขนานของแบบสอบของผู้เชี่ยวชาญที่เป็นตัวอย่างทั้งหมดไม่เป็นไปตามข้อตกลงเบื้องต้นด้านความเป็นเอกมิตของแบบวัด โดยพบว่ามีกลุ่มผู้เชี่ยวชาญที่มีระดับความสามารถในการตัดสินใจเป็นคู่ขนานของข้อสอบค่อนข้างต่ำถึงต่ำมาก (0.00 ถึง 0.40) ซึ่งอาจเกิดจากการขาดองค์ความรู้ในด้านความเป็นคู่ขนานของข้อสอบหรือให้ความร่วมมือในการทำแบบวัดค่อนข้างน้อย ทำให้ฉันทามตินั้นแตกต่างจากกลุ่มผู้เชี่ยวชาญที่มีระดับความสามารถในการตัดสินใจเป็นคู่ขนานของข้อสอบสูงกว่า เมื่อพิจารณาตัดผู้เชี่ยวชาญที่มีระดับความสามารถในการตัดสินใจเป็นคู่ขนานของข้อสอบต่ำกว่าเกณฑ์ออกและวิเคราะห์ฉันทามติใหม่พบว่า แบบวัดเป็นไปตามข้อตกลงเบื้องต้นของทฤษฎีฉันทามติทางวัฒนธรรมแบบมีรูปแบบ สอดคล้องกับงานวิจัยของ J. C. Johnson (1990); Krackhardt (1987) และ Weller (2007) ที่วิเคราะห์ฉันทามติหลังจากตัดกลุ่มที่มีระดับความเห็นพ้องร่วมกันต่ำออกและแสดงถึงความ เป็นเอกมิตของกลุ่มที่อยู่ในวัฒนธรรมเดียวกันมากขึ้นอย่างชัดเจน

การตรวจสอบข้อตกลงเบื้องต้นของทฤษฎีฉันทามติทางวัฒนธรรมแบบมีรูปแบบ มีความสำคัญที่แบบสอบหรือแบบวัดจะต้องมีความเป็นเอกมิตเพราะคำตอบที่ต้องการทราบเกิดจากฉันทามติรวมของกลุ่มที่มีคุณลักษณะเดียวกันจึงจะมีความน่าเชื่อถือและเป็นประโยชน์ในการเป็นสารสนเทศในการตัดสินใจในเรื่องที่วัดและประเมินผลนั้นๆ หากพบว่ามีกรณีการละเมิดข้อตกลงเบื้องต้นข้อนี้ควรพิจารณาวิเคราะห์ฉันทามติเฉพาะกลุ่มตัวอย่างหรือประชากรที่มีความสอดคล้องกลมกลืนเป็นฉันทามติไปในทิศทางเดียวกัน หรือวิเคราะห์ฉันทามติแยกกลุ่มวัฒนธรรม (Cross-cultural consensus analysis) (Hartnett, 2011; Weller, 2007) ส่วนข้อตกลงเบื้องต้นด้านความเป็นอิสระต่อกันของผู้ตอบหรือลงฉันทามติและด้านความถูกต้องของคำตอบ จะต้องวางแผน ออกแบบและบริหารจัดการวัดหรือการสอบให้รัดกุมและเป็นไปตามข้อตกลงเบื้องต้นอย่างเคร่งครัด

ประเด็นที่ 2 การตรวจสอบระดับความสามารถในการตัดสินใจเลือกความเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญโดยอาศัยทฤษฎีฉันทามติทางวัฒนธรรมแบบมีรูปแบบ

จากการประมาณค่าระดับความสามารถในการตัดสินใจเลือกความเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญ โดยอาศัยทฤษฎีฉันทามติทางวัฒนธรรมแบบมีรูปแบบ พบว่า ผู้เชี่ยวชาญส่วนใหญ่มีระดับความสามารถในการตัดสินใจเลือกความเป็นคู่ขนานของข้อสอบอยู่ในระดับที่ 3 มากกว่าระดับที่ 2 และ 1 โดยเฉพาะในกลุ่มผู้เชี่ยวชาญที่มีประสบการณ์สอนนักศึกษาแพทย์น้อยกว่า 10 ปี แต่ในกลุ่มผู้เชี่ยวชาญที่มีประสบการณ์สอนตั้งแต่ 10 ปีขึ้นไปจะมีระดับความสามารถในการตัดสินใจเลือกความเป็นคู่ขนานของข้อสอบในระดับที่ 2 มากที่สุด ซึ่งแสดงให้เห็นว่าประสบการณ์ในการสอนนักศึกษาแพทย์มีส่วนช่วยในการตัดสินใจเลือกความเป็นคู่ขนานของแบบสอบได้ดีขึ้น โดยคุณสมบัติที่สำคัญของผู้เชี่ยวชาญที่ควรพิจารณาสำหรับคัดเลือกผู้เชี่ยวชาญในการสร้างแบบสอบคู่ขนานควรมีประสบการณ์ในการสอนรายวิชานั้นๆ รวมถึงประสบการณ์และความเชี่ยวชาญในการออกข้อสอบ คัดเลือกข้อสอบ รวมถึงการตัดสินใจตามเนื้อหาของข้อสอบตามจุดประสงค์การเรียนรู้รายวิชานั้นของแต่ละผู้เชี่ยวชาญด้วย (Romney et al., 1986; Weller, 2007) ทั้งนี้การตรวจสอบระดับความสามารถในการตัดสินใจเลือกความเป็นคู่ขนานของข้อสอบทำให้ได้สารสนเทศในการคัดเลือกผู้เชี่ยวชาญที่เหมาะสมในการสร้างแบบสอบคู่ขนานให้มีคุณภาพความตรงและความถูกต้องสูงขึ้น รวมถึงเพิ่มระดับความสัมพันธ์ระหว่างเนื้อหาหรือจุดประสงค์การเรียนรู้กับแบบสอบได้ดีขึ้น (Panjaburee et al., 2010; Wanichsan et al., 2012)

ประเด็นที่ 3 การเปรียบเทียบดัชนีความเป็นคู่ขนานของแบบสอบคู่ขนานที่ได้จากการคัดเลือกโดยกลุ่มผู้เชี่ยวชาญที่มีระดับความสามารถในการตัดสินใจเลือกความเป็นคู่ขนานของข้อสอบและประสบการณ์สอนนักศึกษาแพทย์ที่แตกต่างกัน

งานวิจัยในครั้งนี้พิจารณาดัชนีความเป็นคู่ขนานของแบบสอบโดยใช้ RMSD และ MSG พบว่า แบบสอบคู่ขนานที่คัดเลือกโดยกลุ่มผู้เชี่ยวชาญที่มีความสามารถในการตัดสินใจเลือกความเป็นคู่ขนานของข้อสอบในระดับที่ 1 ทั้งกลุ่มที่มีประสบการณ์สอนตั้งแต่ 10 ปีขึ้นไป ซึ่งผลการตรวจสอบความเป็นคู่ขนานของแบบสอบเป็นไปในทิศทางเดียวกันกับการพิจารณาดัชนีความเป็นคู่ขนานของแบบสอบโดยใช้ ค่า MRD และ MRIG โดยมีแนวโน้มของค่า ดัชนีทั้ง 4 ค่าสูงขึ้นตามระดับความสามารถในการตัดสินใจเลือกความเป็นคู่ขนานของข้อสอบของกลุ่มผู้เชี่ยวชาญที่ลดลงและมีประสบการณ์สอนนักศึกษาแพทย์ที่น้อยกว่า 10 ปี ซึ่งดัชนีดังกล่าวมีความเหมาะสมในการใช้ตรวจสอบความเป็นคู่ขนานของแบบสอบสอดคล้องกับงานวิจัยของ Debeer et al. (2017) ที่ได้ปรับเกณฑ์การตัดสินใจเลือกความเป็นคู่ขนานของแบบสอบที่ใช้สอบในสภาพจริง เนื่องจากการใช้ MSD ซึ่งเป็นดัชนีในการตรวจสอบความเป็นคู่ขนานของแบบสอบในงานวิจัยของ Luecht and Nungester (1998); McDonald (1999) และ

Ali and Rijn (2016) เหมาะสำหรับการใช้สร้างแบบสอบคู่ขนานจากคลังข้อสอบโดยอาศัยการกำหนดเงื่อนไขดังกล่าวก่อนเพื่อสร้างแบบสอบคู่ขนานหลายชุดโดยโปรแกรมคอมพิวเตอร์ซึ่งเป็นดัชนีที่ค่อนข้างเข้มงวดและใช้ได้ในสภาพจริงค่อนข้างยาก นอกจากนี้ การเลือกใช้สารสนเทศของแบบสอบ (Test information) ในการคำนวณดัชนีความเป็นคู่ขนานของแบบสอบทั้ง RMSD, MSG, MRD และ MRIG สามารถใช้ตรวจสอบความเป็นคู่ขนานของแบบสอบได้ดี เนื่องจากไม่มีปัจจัยด้านการทำหน้าที่ต่างกันของข้อสอบ (Differential item functioning; DIF) เหมือนการใช้โค้งลักษณะของแบบสอบ ซึ่งจำเป็นต้องมีการปรับเทียบคะแนนสอบของผู้สอบ สอดคล้องกับงานวิจัยที่เกี่ยวกับการตรวจสอบความเป็นคู่ขนานของแบบสอบอื่นๆ ที่นิยมใช้ค่าสารสนเทศของแบบสอบมากกว่าโค้งลักษณะของแบบสอบ (Chen, Chang, & Wu, 2012; Debeer et al., 2017; Lin, 2008; Luecht & Nungester, 1998; McDonald, 1999; Van Der Linden, 2005)

จากการวิจัยในครั้งนี้ พบว่า เมื่อพิจารณา MSG และ MRIG ซึ่งเป็นการเปรียบเทียบค่าสารสนเทศของแบบสอบ ณ ระดับความสามารถของผู้สอบสูงสุดเพียงจุดเดียว สามารถใช้เป็นดัชนีการตรวจสอบความเป็นคู่ขนานเบื้องต้นได้แต่ควรพิจารณาร่วมกับค่า RMSD และ MRD เสมอ เนื่องจากแบบสอบคู่ขนานจะต้องมีโครงสร้างทางเนื้อหา โค้งลักษณะของข้อสอบ และฟังก์ชันสารสนเทศของแบบสอบเหมือนกัน มีความน่าจะเป็นที่ผู้ตอบข้อสอบถูก ($P_i(\theta)$) เท่ากันในทุกความสามารถที่มีอยู่ในผู้ตอบ (θ) และรูปแบบการกระจายไม่แตกต่างกัน ตามโมเดลทฤษฎีการตอบสนองข้อสอบแบบ 2 พารามิเตอร์ (Linden & Adema, 1998; Veldkamp et al., 2013) ในการวิจัยครั้งนี้ ใช้การตรวจสอบความเป็นคู่ขนานของแบบสอบโดยพิจารณาระดับความสามารถของผู้สอบจำนวนทั้งสิ้น 15 ระดับ ซึ่งแสดงให้เห็นว่าแบบสอบมีความเป็นคู่ขนานตลอดโค้งสารสนเทศของแบบสอบ สอดคล้องกับงานวิจัยของ Linden and Adema (1998) เสนอให้ใช้อย่างน้อย 5 ระดับความสามารถของผู้สอบขึ้นไป และงานวิจัยของ Debeer et al. (2017) พบว่า การใช้ระดับความสามารถของผู้สอบเพียง 3 ระดับ จะทำให้เกิดความคลาดเคลื่อนในการคำนวณค่า MRD และ MRIG สูงกว่าการใช้ระดับความสามารถของผู้สอบ 6 ระดับ

จากการวิเคราะห์เปรียบเทียบค่าเฉลี่ยของ MRD ที่ได้จากแบบสอบทั้งฉบับจำแนกตามกลุ่มผู้เชี่ยวชาญที่มีลักษณะแตกต่างกัน 6 กลุ่ม พบว่า ค่าเฉลี่ยของ MRD ที่ได้จากแบบสอบที่คัดเลือกโดยกลุ่มผู้เชี่ยวชาญทั้ง 6 กลุ่มพบว่า กลุ่มผู้เชี่ยวชาญในระดับความสามารถในการตัดสินใจความเป็นคู่ขนานของข้อสอบที่ 1 และประสบการณ์สอนตั้งแต่ 10 ปีขึ้นไป มีค่าเฉลี่ยของ MRD ต่ำกว่ากลุ่มผู้เชี่ยวชาญอื่นๆทุกกลุ่ม และกลุ่มผู้เชี่ยวชาญในระดับความสามารถในการตัดสินใจความเป็นคู่ขนานของข้อสอบที่ 1 และประสบการณ์สอน น้อยกว่า 10 ปี มีค่าเฉลี่ยของ MRD ต่ำกว่ากลุ่มผู้เชี่ยวชาญในระดับความสามารถในการตัดสินใจความเป็นคู่ขนานของข้อสอบที่ 3 และประสบการณ์สอนน้อยกว่า 10 ปี ซึ่ง

แสดงให้เห็นว่า กลุ่มผู้เชี่ยวชาญที่มีระดับความสามารถในการตัดสินระดับความเป็นคู่ขนานของข้อสอบที่สูงสามารถคัดเลือกข้อสอบให้มีดัชนีความเป็นคู่ขนานของแบบสอบได้ดีกว่ากลุ่มผู้เชี่ยวชาญที่มีระดับความสามารถในการตัดสินระดับความเป็นคู่ขนานของข้อสอบที่ต่ำกว่า สอดคล้องกับการวิเคราะห์ความแปรปรวนของค่า MRD แม้ว่าปฏิสัมพันธ์ระหว่างระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบและประสบการณ์สอนไม่มีอิทธิพลต่อค่า MRD แต่ระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบของผู้เชี่ยวชาญสูงจะมีค่าเฉลี่ยของค่า MRD ที่น้อยกว่าระดับต่ำกว่าอย่างชัดเจน ดังนั้น การตรวจสอบระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบ จึงให้สารสนเทศที่สำคัญต่อการคัดเลือกกลุ่มผู้เชี่ยวชาญที่ทำหน้าที่พิจารณาคัดเลือกข้อสอบเพื่อสร้างแบบสอบคู่ขนานในสภาพจริง

ประเด็นที่ 4 การเปรียบเทียบดัชนีความเป็นคู่ขนานของข้อสอบที่คัดเลือกโดยกลุ่มผู้เชี่ยวชาญที่มีระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบและประสบการณ์สอนนักศึกษาแพทย์ที่แตกต่างกัน

งานวิจัยในครั้งนี้ พบว่า แบบสอบคู่ขนานที่คัดเลือกโดยกลุ่มผู้เชี่ยวชาญส่วนใหญ่มีความเป็นคู่ขนานกันของข้อสอบสอดคล้องกับการตรวจสอบความเป็นคู่ขนานของแบบสอบโดยใช้ $RMSD_{IF}$ และค่าเฉลี่ย $RMSD_{IF}$ แบบสอบคู่ขนานที่คัดเลือกโดยกลุ่มผู้เชี่ยวชาญทุกกลุ่มให้ผลสอดคล้องกัน คือ ข้อสอบที่ถูกคัดเลือกโดยกลุ่มผู้เชี่ยวชาญเมื่อนำมาสร้างเป็นแบบสอบแล้วในภาพรวมมีความเป็นคู่ขนานกันของข้อสอบในทุกแบบสอบ และเมื่อวิเคราะห์เปรียบเทียบค่าเฉลี่ยของ $RMSD_{IF}$ พบว่าแบบสอบที่ได้จากการคัดเลือกข้อสอบโดยกลุ่มผู้เชี่ยวชาญทั้ง 6 กลุ่มมีค่าเฉลี่ยของ $RMSD_{IF}$ ไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 ทั้งนี้ เมื่อพิจารณาระดับความสามารถของผู้เชี่ยวชาญในการตัดสินความเป็นคู่ขนานของข้อสอบพบว่า กลุ่มผู้เชี่ยวชาญมีความเป็นเอกมิติและอยู่ในกลุ่มฉันทามติที่มีทิศทางการตัดสินใจอยู่ในทิศทางเดียวกัน จากการวิเคราะห์สถิติตามทฤษฎีฉันทามติทางวัฒนธรรมแบบมีรูปแบบหลังจากตัดกลุ่มผู้เชี่ยวชาญที่มีระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบต่ำกว่า 0.40 ดังนั้น ผู้เชี่ยวชาญส่วนใหญ่ จึงมีระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบผ่านเกณฑ์ที่กำหนดซึ่งเป็นผลให้การคัดเลือกข้อสอบมีดัชนีความเป็นคู่ขนานของข้อสอบมีประสิทธิภาพ สอดคล้องกับงานวิจัยของ Spray and Miller (1992) และ Debeer et al. (2017) ที่ให้ความสำคัญต่อความเป็นคู่ขนานรายข้อ เพราะจะนำไปสู่ผลรวมของความเป็นคู่ขนานของแบบสอบทั้งฉบับ ซึ่งจะพบว่า แบบสอบ 2 ฉบับที่ประกอบด้วยข้อสอบคู่ขนานรายข้อในสัดส่วนที่สูงมีโอกาสที่แบบสอบคู่นั้นจะคู่ขนานกันสูงขึ้น ทั้งนี้การตรวจสอบความเป็นคู่ขนานของข้อสอบควรใช้ระดับความสามารถของผู้สอบอย่างน้อย 6 ระดับ จึงจะทำให้การตรวจสอบมีความ

นำเชื้อถื้อและคลาดเคลื่อนน้อย (Debeer et al., 2017) โดยงานวิจัยในครั้งนี้นำใช้ระดับความสามารถของผู้สอบ 15 ระดับซึ่งให้ผลการตรวจสอบความเป็นคู่ขนานของข้อสอบสอดคล้องกันเป็นอย่างดี

ข้อเสนอแนะ

ข้อเสนอแนะในการนำผลการวิจัยไปใช้

1. จากผลการวิจัย พบว่า กลุ่มผู้เชี่ยวชาญที่มีระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบโดยอาศัยทฤษฎีฉันทามติทางวัฒนธรรมแบบมีรูปแบบอยู่ในระดับสูง สามารถคัดเลือกข้อสอบเพื่อสร้างแบบสอบคู่ขนานให้มีดัชนีความเป็นคู่ขนานของแบบสอบได้สูงกว่ากลุ่มผู้เชี่ยวชาญที่อยู่ในระดับต่ำกว่า ซึ่งให้สารสนเทศในการพิจารณาคัดเลือกกลุ่มผู้เชี่ยวชาญในการสร้างแบบสอบคู่ขนาน ควรมีการตรวจสอบระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบก่อนเพื่อให้ได้แบบสอบคู่ขนานที่มีคุณภาพทั้งความตรงและดัชนีความเป็นคู่ขนานของแบบสอบสูงขึ้น

2. จากผลการวิจัย พบว่า ประสบการณ์ในการสอนนักศึกษาแพทย์มีส่วนช่วยเพิ่มค่าดัชนีความเป็นคู่ขนานของแบบสอบที่คัดเลือกโดยกลุ่มผู้เชี่ยวชาญที่มีระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบในระดับสูง ดังนั้น การคัดเลือกกลุ่มผู้เชี่ยวชาญในการคัดเลือกข้อสอบเพื่อสร้างแบบสอบคู่ขนานควรพิจารณาประสบการณ์การสอนนักศึกษาในรายวิชาหรือเนื้อหาที่เกี่ยวข้องกับข้อสอบร่วมด้วย

3. จากผลการวิจัย พบว่า ผู้เชี่ยวชาญมีความเป็นเอกมิตีและอยู่ในกลุ่มฉันทามติที่มีทิศทางตัดสินใจอยู่ในทิศทางเดียวกัน จากการวิเคราะห์สถิติตามทฤษฎีฉันทามติทางวัฒนธรรมแบบมีรูปแบบหลังจากตัดกลุ่มผู้เชี่ยวชาญที่มีระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบต่ำ ข้อสอบที่ถูกคัดเลือกโดยกลุ่มผู้เชี่ยวชาญทุกกลุ่มเมื่อนำมาสร้างเป็นแบบสอบแล้วในภาพรวมมีดัชนีความเป็นคู่ขนานของข้อสอบผ่านเกณฑ์ในทุกแบบสอบ ดังนั้น การตรวจสอบระดับความสามารถในการตัดสินความเป็นคู่ขนานของข้อสอบจะให้สารสนเทศที่สำคัญถึงคุณสมบัติขั้นต่ำที่ใช้เป็นการคัดกรองผู้เชี่ยวชาญเบื้องต้นสำหรับคัดเลือกผู้เชี่ยวชาญในการสร้างแบบสอบคู่ขนานในสภาพจริง

4. จากผลการวิจัย พบว่า การตรวจสอบความเป็นคู่ขนานของข้อสอบและแบบสอบโดยพิจารณาระดับความสามารถของผู้สอบจำนวนทั้งสิ้น 15 ระดับ แสดงให้เห็นว่าข้อสอบและแบบสอบมีความเป็นคู่ขนานตลอดไค้งสารสนเทศของข้อสอบและแบบสอบ แสดงให้เห็นว่า การตรวจสอบความเป็นคู่ขนานของข้อสอบและแบบสอบควรพิจารณาระดับความสามารถของผู้สอบในหลายระดับตลอดไค้งสารสนเทศของข้อสอบและแบบสอบ

ข้อเสนอแนะในการทำวิจัยครั้งต่อไป

1. จากผลการวิจัยในครั้งนี้ พบว่า แบบสอบคู่ขนานที่สร้างขึ้นจากการคัดเลือกข้อสอบโดยกลุ่มผู้เชี่ยวชาญที่มีระดับความสามารถในการตัดสินใจเลือกข้อสอบในระดับเดียวกันอยู่ในกลุ่มเดียวกันซึ่งกลุ่มที่อยู่ในระดับสูงสามารถคัดเลือกข้อสอบเพื่อสร้างแบบสอบคู่ขนานดีกว่ากลุ่มที่อยู่ในระดับต่ำกว่า แต่พบว่า ในทุกกลุ่มผู้เชี่ยวชาญที่ผ่านเกณฑ์ขั้นต่ำของการตรวจสอบความเป็นคู่ขนานของข้อสอบโดยอาศัยทฤษฎีฉันทนิยามสามารถคัดเลือกข้อสอบให้มีความเป็นคู่ขนานของข้อสอบรายข้อได้ดีไม่แตกต่างกัน ดังนั้นจึงควรศึกษาการจัดกลุ่มผู้เชี่ยวชาญที่มีสัดส่วนระดับความสามารถในการตัดสินใจเลือกข้อสอบที่แตกต่างกันส่งผลต่อดัชนีความเป็นคู่ขนานของข้อสอบและแบบสอบหรือไม่ เพื่อเป็นสารสนเทศสำหรับการจัดกลุ่มผู้เชี่ยวชาญในสภาพจริงซึ่งอาจจะมีจำนวนผู้เชี่ยวชาญในการคัดเลือกข้อสอบคู่ขนานผู้เชี่ยวชาญที่มีระดับความสามารถในการตัดสินใจเลือกข้อสอบระดับสูงค่อนข้างจำกัด ดังแสดงค่าสถิติที่ได้จากการวิเคราะห์ภูมิหลังของผู้เชี่ยวชาญในงานวิจัยครั้งนี้

2. การวิจัยในครั้งนี้ ทำการตรวจสอบความเป็นคู่ขนานของข้อสอบและแบบสอบโดยพิจารณาค่า RMSD และค่า MRD ของแบบสอบที่ได้จากการคัดเลือกข้อสอบโดยกลุ่มผู้เชี่ยวชาญและวิเคราะห์ดัชนีความเป็นคู่ขนานของข้อสอบและแบบสอบโดยอาศัยทฤษฎีการตอบสนองข้อสอบโมเดล 2 พารามิเตอร์สำหรับข้อสอบรูปแบบเลือกตอบหลายตัวเลือกได้เป็นอย่างดี ดังนั้น ควรประยุกต์ใช้ดัชนีการตรวจสอบความเป็นคู่ขนานของข้อสอบและแบบสอบนี้สำหรับตรวจสอบความเป็นคู่ขนานของแบบสอบในรูปแบบอื่นๆ เช่น ข้อสอบเติมคำตอบ หรือรูปแบบผสมที่มีทั้งข้อสอบแบบเลือกตอบหลายตัวเลือกและเติมคำตอบในสัดส่วนที่แตกต่างกัน โดยพิจารณาจากค่าสารสนเทศของแบบสอบตามโมเดลต่างๆที่เหมาะสมโดยอาศัยทฤษฎีการตอบสนองข้อสอบ

3. การวิจัยในครั้งนี้ ทำการตรวจสอบความเป็นคู่ขนานของข้อสอบและแบบสอบโดยอาศัยทฤษฎีการตอบสนองข้อสอบ โดยอาศัยค่าสารสนเทศของข้อสอบและแบบสอบเพื่อคำนวณดัชนีความเป็นคู่ขนานของข้อสอบและแบบสอบเป็นการมุ่งเน้นพิจารณาความเป็นคู่ขนานของข้อสอบหรือแบบสอบเพื่อใช้ตัดสินผลในการวัดและประเมินผลทางการศึกษา โดยไม่พิจารณาความเป็นคู่ขนานของข้อสอบและแบบสอบในแง่มุมของค่าพารามิเตอร์ต่างๆ เช่น ค่าอำนาจจำแนก (b) และค่าความยากของข้อสอบ (a) ซึ่งจะตรวจสอบได้เมื่อใช้ไค้ลักษณะของข้อสอบและแบบสอบ แต่มีข้อจำกัดคือ ต้องเปรียบเทียบคะแนนของผู้สอบที่มีการทำหน้าที่ต่างกันของข้อสอบ ซึ่งจะต้องใช้ขนาดตัวอย่างค่อนข้างมากในการวิเคราะห์ทางสถิติ โดยเหมาะสำหรับการทดสอบขนาดใหญ่ที่มีผู้มีส่วนได้ส่วนเสียจำนวนมากและต้องการคุณภาพของแบบสอบคู่ขนานที่อยู่ในระดับสูง จึงควรมีการศึกษาวิจัยในการทำ การตรวจสอบความเป็นคู่ขนานของข้อสอบและแบบสอบโดยใช้ทั้งค่าสารสนเทศและไค้ลักษณะของข้อสอบและแบบสอบหากทำการวิจัยสำหรับการสอบขนาดใหญ่และมีผู้เข้าสอบเป็นจำนวนมาก

รายการอ้างอิง

- Ali, U. S., & Rijn, P. W. (2016). An evaluation of different statistical target for assembling parallel form in item response theory. *Applied psychological measurement, 40*(3), 163-179.
- Batchelder, W. H., & Romney, A. K. (1988). Test theory without an answer key. *Psychometrika, 53*(1), 71-92.
- Borgatti, S. (1996). *ANTHROPAC 4.0*. Columbia, SC: University of South Carolina Press.
- Borgatti, S. P., Everett, M. G., & Freeman, L. C. (1992). UCINET IV network analysis software. *Connections, 15*(1), 12-15.
- Borgatti, S. P., & Halgin, D. S. (2011). 10 Consensus Analysis. *A companion to cognitive anthropology, 171*.
- Boughton, K. A., & Gierl, M. (2000). Running Head: PARALLEL FORMS CONSTRUCTION Automated Test Assembly Procedures for Criterion-Referenced Testing Using Optimization Heuristics.
- Carolina, J. P. W. e. a. (2014). A comparison of medical and pharmacy students' knowledge and skills of pharmacology and pharmacotherapy. *British Journal of Clinical Pharmacology*.
- Chavez, L. (1995). Comparative Views of Breast and Cervical Cancer Risk Factors among Latinas. *Anglo Women, and Physicians*.
- Chen, P. H., Chang, H. H., & Wu, H. (2012). Item selection for the development of parallel forms from an IRT-based seed test using a sampling and classification approach. *Education and psychological measurement*.
- D'Andrade, R. G. (1987). Modal responses and cultural expertise. *American Behavioral Scientist*.
- Debeer, D., Ali, U. S., & Rijn, P. W. (2017). Evaluating Statistical Targets for Assembling Parallel Mixed-Format Test Forms. *Journal of Educational Measurement, 54*(2), 218-242.
- Dorans, N. J., Pommerich, J., & Holland, P. (2007). *Linking and aligning scores and scales*. New York, NY: Springer-Verlag.

- Dressler, W. W., McBalieiro, M. C., & Dos Santos, J. E. (1997). The cultural construction of social support in Brazil: Associations with health outcomes. *Culture, Medicine, and Psychology*.
- Gell, A. (1998). *Art and agency: an anthropological theory*. Clarendon Press.
- Gronlund, N. E. (2009). *Assessment of Student Achievement* (9th ed.): Pearson.
- Hartnett, T. (2011). *Consensus-oriented decision-making: the codm model for facilitating groups to widespread agreement*: New society publishers.
- Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. London: Routledge.
- Hettiaratchi, E. S. G. (1978). A comparison of student performance in two parallel physiology tests in multiple choice and short answer forms. *Medical education*.
- Johnson, J. C. (1990). *Selecting Ethnographic Informants*. Newbury Park: Sage Publications.
- Johnson, R. L. (2009). *Assessing performance: designing, scoring, and validating performance tasks*: The Guilford Press.
- Krackhardt, D. (1987). Cognitive social structures. *Social networks*, 9(2), 109-134.
- Lin, C. J. (2008). Comparisons between classical theory and item response theory in automated assembly of parallel test forms. *The journal of technology, learning and assessment*.
- Linden, W. J., & Adema, J. J. (1998). Simultaneous assembly of multiple test forms. *Journal of Educational Measurement*.
- Linden, W. J., & Luecht, R. M. (1998). Observed-score equating as a test assembly problem. *Psychometrika*.
- Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. Hillsdale, NJ: Erlbaum.
- Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading, Massachusetts: Addison-Wesley.
- Luecht, R. M., & Nungester, R. J. (1998). Some Practical Examples of Computer-Adaptive Sequential Testing. *Journal of Educational Measurement*.

- Magana, J. R., Burton, M., & Ferreira-Pinto, J. (1995). Occupational cognition in three nations. *Journal of Quantitative Anthropology*.
- Maher, K. M. (1987). *A multiple choice model for aggregating group knowledge and estimating individual competencies*. (Ph.D. Dissertation), University of California, Irvine.
- McDonald, R. P. (1999). *Test theory : a unified treatment*. Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc.
- Panjaburee, P., Hwang, G. J., Triampo, W., & Shih, B. Y. (2010). A multi-expert approach for developing testing and diagnostic system based on the concept-effect model. *Computer and education*.
- Ramos, R. C. (2013). *Use of Item Response Theory in Medical Surgical Nursing Achievement Examination*. Paper presented at the World Academy of Science, Engineering and Technology.
- Romney, A. K., Batchelder, W. H., & Weller, S. C. (1987). Recent applications of cultural consensus. *American Behavioral Scientist*.
- Romney, A. K., H., B. W., & C., W. S. (1986). Culture as Consensus: A Theory of Culture and Informant Accuracy. *American Anthropologist*.
- Samejima, F. (1977b). Weakly parallel tests in latent trait theory with some criticisms of classical test theory. *Psychometrika*.
- Sanders, P. F., & Verschoor, A. J. (1998). Parallel test construction using classical item parameters. *Applied psychological measurement*.
- Spray, J. A., & Miller, T. R. (1992). *Performance of the Mantel-Haenszel statistic and the standardized difference in proportions correct when population ability distributions are incongruent*. Retrieved from Iowa City, Iowa:
- Van Der Linden, W. J. (2005). A Comparison of Item-Selection Methods for Adaptive Tests with Content Constraints. *Journal of Educational Measurement*, 42(3), 283-302.
- Veldkamp, B. P., Matteucci, M., & De Jong, M. G. (2013). Uncertainties in the item parameter estimates and robust automated test assembly. *Applied psychological measurement*.

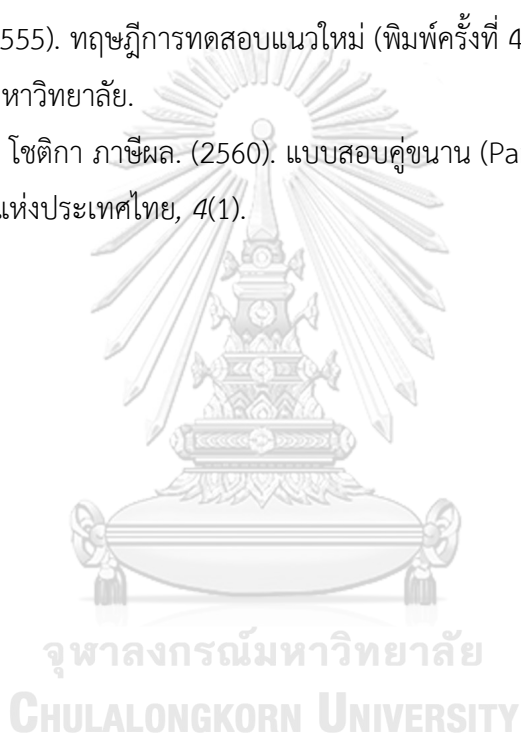
Wanichsan, D., Panjaburee, P., Laosinchai, P., Triampo, W., & Chookaew, S. (2012). A majority-density approach to developing testing and diagnostic systems with the cooperation of multiple experts based on an enhanced concept-effect relationship model. *Expert systems with application*.

Weller, S. C. (2007). Cultural consensus theory: Applications and frequently asked questions. *Field methods*, 19(4), 339-368.

โชติกา ภาชีผล. (2559). การวัดและประเมินผลการเรียนรู้. กรุงเทพฯ: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

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

อนุพงษ์ กันธิวงศ์ และ โชติกา ภาชีผล. (2560). แบบสอบคู่ขนาน (Parallel test). สารสมาคมวิจัยสังคมศาสตร์แห่งประเทศไทย, 4(1).





ภาคผนวก

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

ภาคผนวก ก
เครื่องมือที่ใช้ในการวิจัย

แบบวัดระดับความสามารถในการตัดสินใจเป็นคู่ขนานของแบบสอบโดยผู้เชี่ยวชาญ

การศึกษาครั้งนี้เป็นส่วนหนึ่งของวิทยานิพนธ์ หัวข้อ การเปรียบเทียบดัชนีความเป็นคู่ขนานของข้อสอบและแบบสอบที่ได้จากการคัดเลือกข้อสอบโดยกลุ่มผู้เชี่ยวชาญที่มีลักษณะแตกต่างกัน: การประยุกต์ใช้ทฤษฎีฉันทามติทางวัฒนธรรมแบบมีรูปแบบ โดยมีจุดประสงค์เพื่อศึกษาระดับความสามารถในการเห็นพ้องร่วมกันต่อความเป็นคู่ขนานของแบบสอบของผู้เชี่ยวชาญและใช้เป็นข้อมูลในการจัดกลุ่มผู้เชี่ยวชาญเพื่อคัดเลือกข้อสอบสำหรับสร้างแบบสอบคู่ขนาน ผู้วิจัยจึงขออนุญาตสอบถามเกี่ยวกับชื่อ-สกุลเพื่อใช้ประสานงานในการศึกษาขั้นต่อไป ทั้งนี้ข้อมูลทั้งหมดจะเก็บเป็นความลับและไม่เผยแพร่ข้อมูลที่ระบุถึงบุคคลได้

ผู้เข้าร่วมวิจัย

ยินยอมให้ใช้ข้อมูลได้

ไม่ยินยอมให้ใช้ข้อมูล

ลงชื่อ..... ผู้เข้าร่วมวิจัย

ส่วนที่ 1 ข้อมูลทั่วไป

ชื่อ-สกุล.....

ชุดที่

เพศ ชาย หญิง

ท่านเป็นอาจารย์ ชั้นปรีคลินิก ภาควิชา.....

ชั้นคลินิก ภาควิชา.....

ระดับการศึกษาสูงสุด ปริญญาตรี ปริญญาโท ปริญญาเอก

สาขาที่จบการศึกษา แพทยศาสตร์ เกษตรศาสตร์ อื่นๆ ระบุ.....

ประสบการณ์การสอนนักศึกษาแพทย์ ปี

ท่านเคยเข้ารับการอบรมด้านแพทยศาสตรศึกษาในช่วง 5 ปีที่ผ่านมาหรือไม่

เคย ไม่เคย

ส่วนที่ 2 แบบฟอร์มการตัดสินใจเป็นคู่ขนานของแบบสอบโดยผู้เชี่ยวชาญ

ประเด็นที่พิจารณา: ความเป็นคู่ขนานกันของข้อสอบแต่ละคู่ จำนวนทั้งหมด 20 คู่ ต่อไปนี้

- ถ้าข้อสอบคู่ใดมีความเป็นคู่ขนานกัน โปรดทำเครื่องหมาย ✓ ในช่อง ข้อสอบมีความเป็นคู่ขนานกัน
- ถ้าข้อสอบคู่ใดไม่มีความเป็นคู่ขนานกัน โปรดทำเครื่องหมาย ✓ ในช่อง ข้อสอบไม่มีความเป็นคู่ขนานกัน

***กรุณาตัดสินใจเป็นคู่ขนานของข้อสอบทุกข้อ

ข้อสอบคู่ที่ 1	
ข้อ 1A	ข้อ 1B
หญิงอายุ 25 ปี มีอาการหอบเหนื่อยมา 3 วัน จากผลตรวจร่างกายและ CXR ได้รับการวินิจฉัยเป็น Community acquire pneumonia หากแพทย์ต้องการสั่งใช้ยา ceftriaxone ควรเขียนใบสั่งยาอย่างไร	ชายอายุ 25 ปี มีแผลหนองบริเวณเท้าซ้ายมา 3 วัน จากผลตรวจร่างกาย ได้รับการวินิจฉัยเป็น Infected wound at Lt. foot หากแพทย์ต้องการสั่งใช้ยา cefazolin ควรเขียนใบสั่งยาอย่างไร
<ol style="list-style-type: none"> 1. Ceftriaxone 1 g IV 2. Ceftriaxone 1 amp (V) 3. Ceftriaxone 1 amp IV 4. Ceftriaxone 1000 mg (V) 5. Ceftriaxone 1000 mg IV 	<ol style="list-style-type: none"> 1. Cefazolin 1 g IV 2. Cefazolin 1 amp (V) 3. Cefazolin 1 amp IV 4. Cefazolin 1000 mg (V) 5. Cefazolin 1000 mg IV
ผลการตัดสินความเป็นคู่ขนานของแบบสอบ <input type="checkbox"/> ข้อสอบมีความเป็นคู่ขนานกัน <input type="checkbox"/> ข้อสอบไม่มีความเป็นคู่ขนานกัน	

ข้อสอบคู่ที่ 2	
ข้อ 2A	ข้อ 2B
ข้อมูลในข้อใดที่ไม่จำเป็นต้องเขียนในใบสั่งยาของผู้ป่วยที่มารับยาต่อเนื่อง	ข้อมูลในข้อใดที่จำเป็นต้องเขียนในใบสั่งยาของผู้ป่วยที่มารับยาต่อเนื่อง
<ol style="list-style-type: none"> 1. เลขประจำตัวผู้ป่วยนอก 2. ประวัติการแพ้ยาของผู้ป่วย 3. ผลข้างเคียงที่พบบ่อยของยา 4. ชื่อสามัญของยาที่ต้องการสั่งใช้ 5. ชื่อ-นามสกุลของผู้ป่วย 	<ol style="list-style-type: none"> 1. ผลข้างเคียงที่พบบ่อยของยา 2. ประวัติการแพ้ยาในครอบครัว 3. ชื่อสามัญของยาที่ต้องการสั่งใช้ 4. เลขประจำตัวประชาชนของผู้ป่วย 5. ประวัติการตั้งครรภ์ของผู้ป่วย
ผลการตัดสินความเป็นคู่ขนานของแบบสอบ <input type="checkbox"/> ข้อสอบมีความเป็นคู่ขนานกัน <input type="checkbox"/> ข้อสอบไม่มีความเป็นคู่ขนานกัน	

ข้อสอบคู่ที่ 3	
ข้อ 3A	ข้อ 3B
<p>หญิงอายุ 70 ปี โรคประจำตัว cerebro-vascular thrombosis มา 1 ปี รับประทานยา warfarin ทุกวัน 5 วันก่อน มีอาการปวดท้องบริเวณลิ้นปี่จึงไปซื้อยา cimetidine มารับประทานเอง ต่อมาพบว่า มีปัสสาวะเป็นเลือด ข้อความต่อไปนี้ข้อใดถูก</p> <ol style="list-style-type: none"> 1. Cimetidine มี therapeutic index แคบ 2. Warfarin กระตุ้นการทำงานของ CYP2C9 3. Cimetidine ยับยั้งการทำงานของ CYP3A4 4. Warfarin จะถูกเปลี่ยนแปลงเป็น active metabolite มากขึ้น 5. Cimetidine และ warfarin เป็นสารตั้งต้นของเอนไซม์ CYP450 ชนิดเดียวกัน 	<p>หญิงอายุ 70 ปี มีโรคประจำตัวเป็น ischemic stroke และ AF รับประทานยา warfarin ทุกวัน 5 วันก่อน มีอาการปวดท้องบริเวณลิ้นปี่ และคลื่นไส้อาเจียน แพทย์วินิจฉัยเป็น stress-induced gastritis ควรให้การรักษาผู้ป่วยรายนี้ด้วยยาใด</p> <ol style="list-style-type: none"> 1. Omeprazole 2. Ranitidine 3. Lansoprazole 4. Cimetidine 5. Domperidone
<p>ผลการตัดสินความเป็นคู่ขนานของแบบสอบ</p> <p><input type="checkbox"/> ข้อสอบมีความเป็นคู่ขนานกัน</p> <p><input type="checkbox"/> ข้อสอบไม่มีความเป็นคู่ขนานกัน</p>	

ข้อสอบคู่ที่ 4	
ข้อ 4A	ข้อ 4B
<p>Esomeprazole ซึ่งเป็น PPI ตัวใหม่ล่าสุดมีคุณสมบัติเด่นด้านใดเมื่อเทียบกับ omeprazole</p> <ol style="list-style-type: none"> 1. Bioavailability มากกว่า 2. Free form ในกระแสเลือดมากกว่า 3. Affinity ต่อ ATPase มากกว่า 4. ไม่ต้องเปลี่ยนแปลงโดยเอนไซม์ในตับ 5. Adverse effect น้อยกว่า 	<p>เพราะเหตุใด esomeprazole จึงออกฤทธิ์เร็วและนานกว่า omeprazole</p> <ol style="list-style-type: none"> 1. เป็น active drug 2. จับกับโปรตีนในเลือดมากกว่า 3. ถูกดูดซึมได้ดีในกระเพาะอาหาร 4. ยับยั้ง H⁺/K⁺ ATPase อย่างถาวร 5. มี affinity ต่อ H⁺/K⁺ ATPase มากกว่า
<p>ผลการตัดสินความเป็นคู่ขนานของแบบสอบ</p> <p><input type="checkbox"/> ข้อสอบมีความเป็นคู่ขนานกัน</p> <p><input type="checkbox"/> ข้อสอบไม่มีความเป็นคู่ขนานกัน</p>	

ข้อสอบคู่ที่ 5	
ข้อ 5A	ข้อ 5B
<p>ยาใดมีคุณสมบัติต่อไปนี้</p> <ul style="list-style-type: none"> - เป็น serotonin receptor agonists - ไม่มี central depressant effects - กระตุ้นการหลั่ง ACh ใน enteric nerves <ol style="list-style-type: none"> 1. <i>Cisapride</i> 2. Ondansetron 3. Domperidone 4. Metoclopramide 5. Diphenhydramine 	<p>ข้อใดเป็นคุณสมบัติทางเภสัชวิทยาของ cisapride</p> <ol style="list-style-type: none"> 1. <i>เพิ่มการหลั่ง ACh ใน enteric nerves</i> 2. ออกฤทธิ์เป็น μ-opioid receptor agonist 3. ใช้รักษา inflammatory bowel disease 4. ใช้เตรียมลำไส้ก่อนการทำ colonoscopy 5. ผลข้างเคียงที่สำคัญ คือ EPS symptoms
<p>ผลการตัดสินความเป็นคู่ขนานของแบบสอบ</p> <p><input type="checkbox"/> ข้อสอบมีความเป็นคู่ขนานกัน</p> <p><input type="checkbox"/> ข้อสอบไม่มีความเป็นคู่ขนานกัน</p>	

ข้อสอบคู่ที่ 6	
ข้อ 6A	ข้อ 6B
<p>หญิงอายุ 20 ปี ไม่มีโรคประจำตัว มาพบแพทย์เพื่อปรึกษาเลือกใช้อยาคคุมกำเนิด โดยต้องการคุมกำเนิดเป็นเวลา 1 ปี ควรเลือกใช้อยาในข้อใด</p> <ol style="list-style-type: none"> 1. <i>DMPA</i> 2. Etonorgestrel 3. Levonorgestrel 4. Low dose COCs 5. Standard dose COCs 	<p>หญิงอายุ 25 ปี ไม่มีโรคประจำตัว มาพบแพทย์เพื่อปรึกษาเลือกใช้อยาคคุมกำเนิด โดยต้องการคุมกำเนิดเป็นเวลา 3 ปี ควรเลือกใช้อยาในข้อใด</p> <ol style="list-style-type: none"> 1. DMPA 2. <i>Etonorgestrel</i> 3. Levonorgestrel 4. Low dose COCs 5. Standard dose COCs
<p>ผลการตัดสินความเป็นคู่ขนานของแบบสอบ</p> <p><input type="checkbox"/> ข้อสอบมีความเป็นคู่ขนานกัน</p> <p><input type="checkbox"/> ข้อสอบไม่มีความเป็นคู่ขนานกัน</p>	

ข้อสอบคู่ที่ 7	
ข้อ 7A	ข้อ 7B
หญิงอายุ 25 ปี ไม่เคยมีบุตร มีโรคประจำตัวเป็นไมเกรน มาปรึกษาเลือกให้ยาคุมกำเนิดชนิดรับประทาน ควรเลือกให้ยาใดเป็นอันดับแรก	หญิงอายุ 35 ปี เพิ่งคลอดบุตรมา 2 สัปดาห์ ไม่มีโรคประจำตัวใดๆ มาปรึกษาเลือกให้ยาคุมกำเนิดชนิดรับประทาน ควรเลือกให้ยาใดเป็นอันดับแรก
1. DMPA 2. <i>Norgestrel</i> 3. Etonorgestrel 4. Low dose COCs 5. Standard dose COCs	1. DMPA 2. <i>Norgestrel</i> 3. Etonorgestrel 4. Low dose COCs 5. Standard dose COCs
ผลการตัดสินความเป็นคู่ขนานของแบบสอบ <input type="checkbox"/> ข้อสอบมีความเป็นคู่ขนานกัน <input type="checkbox"/> ข้อสอบไม่มีความเป็นคู่ขนานกัน	

ข้อสอบคู่ที่ 8	
ข้อ 8A	ข้อ 8B
แพทย์ผู้สั่งยาจะต้องเขียนรายละเอียดที่ส่วนใดของใบสั่งยา จึงจะบอกความแตกต่างของการใช้ยาต่อไปนี้ได้อย่างชัดเจน	แพทย์ผู้สั่งยาจะต้องเขียนรายละเอียดที่ส่วนใดของใบสั่งยา จึงจะบอกความแตกต่างของการใช้ยาต่อไปนี้ได้อย่างชัดเจน
<div style="display: flex; justify-content: space-around;"> <div style="border: 1px solid black; padding: 5px; width: 45%;"> <p>Klacid® MR 500 mg Clarithromycin</p> </div> <div style="border: 1px solid black; padding: 5px; width: 45%;"> <p>Klacid® 500 mg Clarithromycin</p> </div> </div>	<div style="display: flex; justify-content: space-around;"> <div style="border: 1px solid black; padding: 5px; width: 45%;"> <p>Klacid® MR 500 mg Clarithromycin</p> </div> <div style="border: 1px solid black; padding: 5px; width: 45%;"> <p>Klacid® 500 mg Clarithromycin</p> </div> </div>
1. Subscription และ Inscription 2. <i>Inscription และ Transcription</i> 3. Superscription และ Subscription 4. Superscription และ Transcription	ก. Superscription ข. Inscription ค. Subscription ง. Transcription 1. ก. และ ค. 2. ก. และ ง. 3. ข. และ ค. 4. <i>ข. และ ง.</i>
ผลการตัดสินความเป็นคู่ขนานของแบบสอบ <input type="checkbox"/> ข้อสอบมีความเป็นคู่ขนานกัน <input type="checkbox"/> ข้อสอบไม่มีความเป็นคู่ขนานกัน	

ข้อสอบคู่ที่ 9	
ข้อ 9A	ข้อ 9B
ยาใดที่ ห้ามใช้ ในหญิงตั้งครรภ์	ยาใดที่ ห้ามใช้ ในหญิงตั้งครรภ์
<ol style="list-style-type: none"> 1. Psyllium 2. Ranitidine 3. <i>Omeprazole</i> 4. Dimenhydrinate 5. Aluminum hydroxide 	<ol style="list-style-type: none"> 1. Psyllium 2. Ranitidine 3. <i>Loperamide</i> 4. Dimenhydrinate 5. Aluminum hydroxide
ผลการตัดสินความเป็นคู่ขนานของแบบสอบ	
<input type="checkbox"/> ข้อสอบมีความเป็นคู่ขนานกัน <input type="checkbox"/> ข้อสอบไม่มีความเป็นคู่ขนานกัน	

ข้อสอบคู่ที่ 10	
ข้อ 10A	ข้อ 10B
Nitrous oxide มีฤทธิ์ทำให้ผู้ป่วยหลับเร็ว เนื่องจากคุณสมบัติข้อใด	Nitrous oxide เป็นยาดมสลบที่ควรระวังในการผ่าตัดอะไร
<ol style="list-style-type: none"> 1. ไม่จับรวมกับเนื้อเยื่อต่างๆ ในร่างกาย 2. โมเลกุลขนาดเล็ก 3. ถูกขับถ่ายทางปอดได้เร็ว 4. <i>การละลายในเลือดน้อย</i> 5. ค่า MAC สูงมาก 	<ol style="list-style-type: none"> 1. ผ่าตัดต่อมทอนซิล 2. ผ่าตัดกระดูกหัก 3. <i>ผ่าตัดแก้วหูทะลุ</i> 4. ผ่าตัดไส้ติ่ง 5. ผ่าตัดไส้เลื่อน
ผลการตัดสินความเป็นคู่ขนานของแบบสอบ	
<input type="checkbox"/> ข้อสอบมีความเป็นคู่ขนานกัน <input type="checkbox"/> ข้อสอบไม่มีความเป็นคู่ขนานกัน	

ข้อสอบคู่ที่ 11	
ข้อ 11A	ข้อ 11B
<p>ชายอายุ 20 ปี อาชีพทหาร ปฏิบัติงานในชายแดนภาคใต้ หลังเกิดการปะทะ มีเพื่อนร่วมค่ายเสียชีวิต ทำให้รู้สึกเครียด หดหู่ ไม่ยอมพูดคุยกับใคร เป็นเวลาหลายเดือน ควรให้การรักษาด้วยยาใด</p> <ol style="list-style-type: none"> 1. Lithium 2. <i>Fluoxetine</i> 3. Haloperidol 4. Risperidone 5. Olanzapine 	<p>ชายอายุ 25 ปี กลัวบุคคลแปลกหน้า ชอบอยู่คนเดียว หูแว่ว จิตแพทย์วินิจฉัยเป็น schizophreniform disorder ควรให้การรักษาด้วยยาใด</p> <ol style="list-style-type: none"> 1. Lithium 2. Fluoxetine 3. Olanzapine 4. <i>Risperidone</i> 5. Haloperidol
<p>ผลการตัดสินความเป็นคู่ขนานของแบบสอบ</p> <p><input type="checkbox"/> ข้อสอบมีความเป็นคู่ขนานกัน</p> <p><input type="checkbox"/> ข้อสอบไม่มีความเป็นคู่ขนานกัน</p>	

ข้อสอบคู่ที่ 12	
ข้อ 12A	ข้อ 12B
<p>ยาในข้อใดที่มีผลลด total peripheral resistance น้อยที่สุด</p> <ol style="list-style-type: none"> 1. Enalapril 2. Losartan 3. Amlodipine 4. Hydrochlorothiazide 5. <i>Atenolol</i> 	<p>ยาในข้อใดที่มีผลลด total peripheral resistance มากที่สุด</p> <ol style="list-style-type: none"> 1. Verapamil 2. Diltiazem 3. <i>Amlodipine</i> 4. Furosemide 5. Atenolol
<p>ผลการตัดสินความเป็นคู่ขนานของแบบสอบ</p> <p><input type="checkbox"/> ข้อสอบมีความเป็นคู่ขนานกัน</p> <p><input type="checkbox"/> ข้อสอบไม่มีความเป็นคู่ขนานกัน</p>	

ข้อสอบคู่ที่ 13	
ข้อ 13A	ข้อ 13B
การติดตามระดับ serum electrolyte มีความจำเป็นในผู้ป่วยที่ใช้ยาลดความดันโลหิตในข้อใดมากที่สุด <i>Enalapril</i> Verapamil Amlodipine Prazosin Atenolol	การติดตามระดับ serum electrolyte มีความจำเป็นในผู้ป่วยที่ใช้ยารักษาภาวะหัวใจวายในข้อใดมากที่สุด <i>Furosemide</i> Dobutamine Epinephrine Dopamine Metoprolol
ผลการตัดสินความเป็นคู่ขนานของแบบสอบ <input type="checkbox"/> ข้อสอบมีความเป็นคู่ขนานกัน <input type="checkbox"/> ข้อสอบไม่มีความเป็นคู่ขนานกัน	

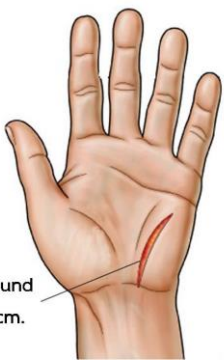
ข้อสอบคู่ที่ 14	
ข้อ 14A	ข้อ 14B
หญิงอายุ 50 ปี โรคประจำตัว T2DM, HT ต่อมามีอาการหอบเหนื่อย นอนราบไม่ได้ แพทย์วินิจฉัยเป็น Acute HFrEF ยาใดควรหลีกเลี่ยงในผู้ป่วยรายนี้ <i>Metformin</i> Insulin Glipizide Linagliptin Canagliflozin	หญิงอายุ 50 ปี โรคประจำตัว T2DM, HT ต่อมามีอาการหอบเหนื่อย นอนราบไม่ได้ แพทย์วินิจฉัยเป็น Acute HFrEF ยาใดควรหลีกเลี่ยงในผู้ป่วยรายนี้ <i>Propranolol</i> Amlodipine Enalapril Losartan Hydrochlorothiazide
ผลการตัดสินความเป็นคู่ขนานของแบบสอบ <input type="checkbox"/> ข้อสอบมีความเป็นคู่ขนานกัน <input type="checkbox"/> ข้อสอบไม่มีความเป็นคู่ขนานกัน	

ข้อสอบคู่ที่ 15	
ข้อ 15A	ข้อ 15B
ยาใดไม่ควรรับประทานพร้อมกับ Antacids 1. <i>Ciprofloxacin</i> 2. Metronidazole 3. Cephalexin 4. Amoxicillin 5. Roxithromycin	ยาใดไม่ควรรับประทานพร้อมกับ Ciprofloxacin 1. <i>Antacids</i> 2. Omeprazole 3. Ranitidine 4. Domperidone 5. Metoclopramide
ผลการตัดสินความเป็นคู่ขนานของแบบสอบ <input type="checkbox"/> ข้อสอบมีความเป็นคู่ขนานกัน <input type="checkbox"/> ข้อสอบไม่มีความเป็นคู่ขนานกัน	

ข้อสอบคู่ที่ 16	
ข้อ 16A	ข้อ 16B
หากใช้ยาในข้อใดเป็นเวลานาน อาจพบภาวะ down-regulation ของ receptor ได้ 1. Atenolol 2. <i>Salbutamol</i> 3. Doxazosin 4. Loratadine 5. Amlodipine	หากใช้ยาในข้อใดเป็นเวลานาน อาจพบภาวะ up-regulation ของ receptor ได้ 1. Fenofibrate 2. Salbutamol 3. Tramadol 4. <i>Loratadine</i> 5. Nitroglycerine
ผลการตัดสินความเป็นคู่ขนานของแบบสอบ <input type="checkbox"/> ข้อสอบมีความเป็นคู่ขนานกัน <input type="checkbox"/> ข้อสอบไม่มีความเป็นคู่ขนานกัน	

ข้อสอบคู่ที่ 17	
ข้อ 17A	ข้อ 17B
<p>ข้อใดเป็นข้อห้ามใช้ของ amoxicillin</p> <ol style="list-style-type: none"> 1. แพ้ยา <i>penicillin</i> 2. ตั้งครรภ์ 3. ผู้สูงอายุ 4. QT-prolongation 5. รับประทานยา warfarin ทุกวัน 	<p>ข้อใดเป็นข้อห้ามใช้ของ moxifloxacin</p> <ol style="list-style-type: none"> 1. แพ้ยา penicillin 2. ภาวะ CKD 3. ผู้สูงอายุ 4. <i>QT-prolongation</i> 5. รับประทานยา aspirin ทุกวัน
<p>ผลการตัดสินความเป็นคู่ขนานของแบบสอบ</p> <p><input type="checkbox"/> ข้อสอบมีความเป็นคู่ขนานกัน</p> <p><input type="checkbox"/> ข้อสอบไม่มีความเป็นคู่ขนานกัน</p>	

ข้อสอบคู่ที่ 18	
ข้อ 18A	ข้อ 18B
<p>เด็กชายอายุ 3 ปี มีไข้ ไอ เจ็บคอ มา 4 วัน BW 25 kg ได้รับการวินิจฉัยเป็น acute pharyngitis แพทย์สั่งยาปฏิชีวนะเป็น amoxicillin ในขนาด 30 mg/kg/day ควรเติมข้อความใด ในช่องว่างจึงจะเหมาะสม</p> <p>Amoxicillin suspension (250 mg/5mL)</p> <p>Sig. 1 tsp ◉ <input type="text"/></p> <ol style="list-style-type: none"> 1. tid ac 2. <i>tid pc</i> 3. bid ac 4. bid pc 5. qid ac 	<p>เด็กชายอายุ 1 ปี มีไข้ ไอ เจ็บคอ 4 วัน BW 12.5 kg ได้รับการวินิจฉัยเป็น acute pharyngitis แพทย์สั่งยาปฏิชีวนะเป็น amoxicillin ในขนาด 30 mg/kg/day ควรเติมข้อความใด ในช่องว่างจึงจะเหมาะสม</p> <p>Amoxicillin suspension (250 mg/5mL)</p> <p>Sig. <input type="text"/> ◉ tid pc</p> <ol style="list-style-type: none"> 1. <i>2.5 mL</i> 2. 5 mL 3. 10 mL 4. 1 tsp 5. 2 tsp
<p>ผลการตัดสินความเป็นคู่ขนานของแบบสอบ</p> <p><input type="checkbox"/> ข้อสอบมีความเป็นคู่ขนานกัน</p> <p><input type="checkbox"/> ข้อสอบไม่มีความเป็นคู่ขนานกัน</p>	

ข้อสอบคู่ที่ 19	
ข้อ 19A	ข้อ 19B
<p>ข้อใดเป็นผลทางเภสัชวิทยาของ atropine</p> <ol style="list-style-type: none"> 1. เป็น potent enzyme inducer 2. เป็น potent enzyme inhibitor 3. ออกฤทธิ์เป็น <i>competitive inhibitor</i> 4. ออกฤทธิ์เป็น non-competitive inhibitor 5. ออกฤทธิ์เป็น partial agonist 	<p>ข้อใดเป็นผลทางเภสัชวิทยาของ roxithromycin</p> <ol style="list-style-type: none"> 1. เป็น potent enzyme inducer 2. เป็น <i>potent enzyme inhibitor</i> 3. ออกฤทธิ์เป็น competitive inhibitor 4. ออกฤทธิ์เป็น non-competitive inhibitor 5. ออกฤทธิ์เป็น partial agonist
<p>ผลการตัดสินความเป็นคู่ขนานของแบบสอบ</p> <p><input type="checkbox"/> ข้อสอบมีความเป็นคู่ขนานกัน</p> <p><input type="checkbox"/> ข้อสอบไม่มีความเป็นคู่ขนานกัน</p>	
ข้อสอบคู่ที่ 20	
ข้อ 20A	ข้อ 20B
<p>ชาย 25 ปี ถูกมีดบาด พบบาดแผลฉีกขาดบริเวณฝ่ามือซ้าย ดังภาพ แพทย์วางแผนการรักษาโดยการเย็บปิดบาดแผล ควรเลือกใช้ยาชาเฉพาะที่ในข้อใด</p>  <p>Lacerated wound 0.5 x 8 x 0.5 cm.</p> <ol style="list-style-type: none"> 1. Pethidine 2. Procaine with epinephrine 3. Procaine 4. <i>Lidocaine HCl with epinephrine</i> 5. Lidocaine HCl 	<p>ชาย 25 ปี ถูกมีดบาด พบบาดแผลฉีกขาดบริเวณฝ่ามือซ้ายขนาด 0.5x8x0.5 cm แพทย์วางแผนการรักษาโดยการเย็บปิดบาดแผล ควรเลือกใช้ยาชาเฉพาะที่ในข้อใด</p> <ol style="list-style-type: none"> 1. Pethidine 2. Procaine with epinephrine 3. Procaine 4. <i>Lidocaine HCl with epinephrine</i> 5. Lidocaine HCl
<p>ผลการตัดสินความเป็นคู่ขนานของแบบสอบ</p> <p><input type="checkbox"/> ข้อสอบมีความเป็นคู่ขนานกัน</p> <p><input type="checkbox"/> ข้อสอบไม่มีความเป็นคู่ขนานกัน</p>	

- ขอขอบพระคุณอาจารย์ทุกท่านที่ให้ความร่วมมือในการทำแบบวัดครั้งนี้ -

ข้อสอบรายวิชาเภสัชวิทยาสำหรับคัดเลือกเพื่อสร้างแบบสอบคู่ขนานโดยกลุ่มผู้เชี่ยวชาญ

ข้อสอบต่อไปนี้สำหรับสร้างแบบสอบคู่ขนานประกอบแบ่งหัวข้อของยาตามระบบ ดังต่อไปนี้

1. Cardiovascular System	จำนวน	จำนวนข้อสอบที่ต้องการ	
จุดประสงค์การเรียนรู้	ข้อสอบ	สำหรับสร้างแบบสอบคู่ขนาน	
มีความรู้เรื่องหลักการรักษาและหลักการใช้ยา โดยเฉพาะยาในบัญชียาหลักแห่งชาติของโรค กลุ่มอาการและภาวะผิดปกติ รวมถึงยาในกลุ่มต่อไปนี้	สำหรับ	(ข้อ)	
	คัดเลือก	แบบสอบ	แบบสอบ
	(ข้อ)	ชุดที่ 1	ชุดที่ 2
A-Antihypertensive drugs	40	5	5
B-Cardiac drugs	40	5	5
รวม	80	10	10
2. Gastrointestinal System	จำนวน	จำนวนข้อสอบที่ต้องการ	
จุดประสงค์การเรียนรู้	ข้อสอบ	สำหรับสร้างแบบสอบคู่ขนาน	
มีความรู้เรื่องหลักการรักษาและหลักการใช้ยา โดยเฉพาะยาในบัญชียาหลักแห่งชาติของโรค กลุ่มอาการ และภาวะผิดปกติ รวมถึงยาในกลุ่มต่อไปนี้	สำหรับ	(ข้อ)	
	คัดเลือก	แบบสอบ	แบบสอบ
	(ข้อ)	ชุดที่ 1	ชุดที่ 2
C-Drugs for dyspepsia and ulcer-healing drugs	40	5	5
D-Antispasmodics, drugs altering gut motility and Laxatives	40	5	5
รวม	80	10	10
3. Endocrine System	จำนวน	จำนวนข้อสอบที่ต้องการ	
จุดประสงค์การเรียนรู้	ข้อสอบ	สำหรับสร้างแบบสอบคู่ขนาน	
มีความรู้เรื่องหลักการรักษาและหลักการใช้ยา โดยเฉพาะยาในบัญชียาหลักแห่งชาติของโรค กลุ่มอาการ และภาวะผิดปกติ รวมถึงยาในกลุ่มต่อไปนี้	สำหรับ	(ข้อ)	
	คัดเลือก	แบบสอบ	แบบสอบ
	(ข้อ)	ชุดที่ 1	ชุดที่ 2
E-Drugs used in diabetes (insulins and oral antidiabetic drugs)	40	5	5
รวม	40	5	5

แบบบันทึกการคัดเลือกข้อสอบคู่ขนานโดยกลุ่มผู้เชี่ยวชาญ กลุ่มที่.....

รายชื่อผู้เชี่ยวชาญ

- 1.
- 2.
- 3.

คำสั่ง: เลือกแบบสอบที่มีความเป็นคู่ขนานกันในแต่ละจุดประสงค์การเรียนรู้ที่กำหนดให้

จุดประสงค์การเรียนรู้	แบบสอบชุดที่ 1	แบบสอบชุดที่ 2
1. Cardiovascular System		
หมวด A-Antihypertensive drugs		
ข้อ 1	A	A
ข้อ 2	A	A
ข้อ 3	A	A
ข้อ 4	A	A
ข้อ 5	A	A
หมวด B-Cardiac drugs		
ข้อ 1	B	B
ข้อ 2	B	B
ข้อ 3	B	B
ข้อ 4	B	B
ข้อ 5	B	B
2. Gastrointestinal System		
หมวด C-Drugs for dyspepsia and ulcer-healing drugs		
ข้อ 1	C	C
ข้อ 2	C	C
ข้อ 3	C	C
ข้อ 4	C	C
ข้อ 5	C	C
หมวด D-Antispasmodics, drugs altering gut motility and Laxatives		
ข้อ 1	D	D
ข้อ 2	D	D
ข้อ 3	D	D
ข้อ 4	D	D
ข้อ 5	D	D

จุดประสงค์การเรียนรู้	แบบสอบชุดที่ 1	แบบสอบชุดที่ 2
3. Endocrine System		
หมวด E-Drugs used in diabetes (insulins and oral antidiabetic drugs)		
ข้อ 1	E	E
ข้อ 2	E	E
ข้อ 3	E	E
ข้อ 4	E	E
ข้อ 5	E	E

ตัวอย่างข้อสอบรายวิชาเภสัชวิทยาสำหรับคัดเลือกเพื่อสร้างแบบสอบคู่ขนานโดยกลุ่มผู้เชี่ยวชาญ

A-Antihypertensive drugs

ข้อ	A-Antihypertensive drugs
A1	<p>ยาใดที่เลือกใช้เป็นอันดับแรกในผู้ป่วยสูงอายุที่มีความดันโลหิตสูงร่วมกับภาวะ metabolic syndromes</p> <ol style="list-style-type: none"> 1. HCTZ 2. Prazosin 3. Propanolol 4. Enalapril 5. <u>Amlodipine</u>

B-Cardiac drugs

ชื่อ	B- Cardiac drugs
B1	<p>ชายอายุ 65 ปี มีประวัติโรคความดันโลหิตสูง เจ็บแน่นหน้าอกซ้ายร้าวไปกราม 30 นาที PTA เหงื่อออกมาก ใจสั่น แต่ไม่มีเวียนหรือหมดสติ ตรวจร่างกาย BP 150/96 mmHg PR 80 bpm RR 16 bpm O₂ sat. 98% ระหว่างรอผลตรวจ EKG-12 leads และ cardiac enzyme</p> <p>ควรให้การรักษาเบื้องต้นอย่างไร</p> <ol style="list-style-type: none"> 1. Aspirin 162 mg P.O. 2. <u>Aspirin 325 mg P.O. (chewed)</u> 3. Nitroglycerine 0.4 mg P.O. 4. Nitroglycerine 0.4 mg P.O. (chewed) 5. Clopidogrel 600 mg P.O.

C-Gastrointestinal System

ชื่อ	C-Drugs for dyspepsia and ulcer-healing drugs
C18	<p>หญิงอายุ 70 ปี โรคประจำตัว cerebro-vascular thrombosis มา 1 ปี รับประทานยา warfarin ทุกวัน 5 วันก่อน มีอาการปวดท้องบริเวณลิ้นปี่จึงไปซื้อยา cimetidine มารับประทานเอง ต่อมาพบว่ามีปัสสาวะเป็นเลือด ข้อความต่อไปนี้อธิบายเหตุการณ์ดังกล่าวได้ถูกต้องที่สุด</p> <ol style="list-style-type: none"> 1. Cimetidine มี therapeutic index แคบ 2. Warfarin กระตุ้นการทำงานของ CYP1A2 และ CYP2C9 3. <u>Cimetidine ยับยั้งการทำงานของ CYP3A4 และ CYP2D6</u> 4. Warfarin จะถูกเปลี่ยนแปลงเป็น active metabolite มากขึ้น 5. Cimetidine และ warfarin เป็นสารตั้งต้นของเอนไซม์ CYP450 ชนิดเดียวกัน

D-Antispasmodics, drugs altering gut motility and Laxatives

ชื่อ	D-Antispasmodics, drugs altering gut motility and Laxatives
D34	<p>ชายอายุ 75 ปีได้รับยาระงับอาเจียนชนิดหนึ่งติดต่อกัน 1 สัปดาห์ หลังจากนั้นเกิดอาการเคลื่อนไหวของร่างกายผิดปกติ ยาที่ได้รับคือข้อใด</p> <ol style="list-style-type: none"> 1. Domperidone 2. Diphenhydramine 3. Metoclopramide 4. Dimenhydrinate 5. Hydroxyzine

E-Endocrine System

ชื่อ	E-Drugs used in diabetes (Insulins and oral antidiabetic drugs)
E24	<p>หญิงอายุ 45 ปี เป็นโรคประจำตัวความดันโลหิตสูง มาตรฐานร่างกายประจำปี น้ำหนัก 89 กิโลกรัม สูง 150 เซนติเมตร อาการปกติดี ระดับน้ำตาลในเลือด (FBS) = 170 mg/dL, LDL 165 mg/dL ควรให้การรักษาผู้ป่วยรายนี้อย่างไร</p> <ol style="list-style-type: none"> 1. เลือกใช้ glipizide เป็นอันดับแรก 2. เลือกใช้ metformin เป็นอันดับแรก 3. เลือกใช้ pioglitazone เป็นอันดับแรก 4. เลือกใช้การฉีด basal insulin เป็นอันดับแรก 5. ยังไม่ให้ยาลดระดับน้ำตาลในเลือด และนัดมาตรวจ FBS ซ้ำ

ภาคผนวก ข
ค่าสถิติพื้นฐาน

ค่าสถิติพื้นฐาน จำแนกตามภูมิหลังของผู้เชี่ยวชาญ แสดงรายละเอียดดังต่อไปนี้

เพศ

เพศ	จำนวน	ร้อยละ
ชาย	29	46.032
หญิง	34	53.968
รวม	63	100.000

สังกัดภาควิชา

ภาควิชา	จำนวน	ร้อยละ
ปรีคลินิก	30	47.619
คลินิก	33	52.381
รวม	63	100.000

ระดับการศึกษา

ระดับการศึกษา	จำนวน	ร้อยละ
ปริญญาตรี	4	6.350
ปริญญาโท	28	44.444
ปริญญาเอก	31	49.206
รวม	63	100.000

สาขาที่จบการศึกษา

สาขาที่จบการศึกษา	จำนวน	ร้อยละ
แพทยศาสตร์	39	61.905
เภสัชศาสตร์	13	20.635
สาขาอื่นๆ	11	17.460
รวม	63	100.000

ประสบการณ์การสอนนักศึกษาแพทย์

ประสบการณ์การสอน	จำนวน	ร้อยละ
ตั้งแต่ 10 ปีขึ้นไป	23	36.508
น้อยกว่า 10 ปี	40	63.492
รวม	63	100.000

การเข้าร่วมอบรมด้านแพทยศาสตรศึกษาในช่วง 5 ปีที่ผ่านมา

การเข้าร่วมอบรมด้านแพทยศาสตรศึกษา ในช่วง 5 ปีที่ผ่านมา	จำนวน	ร้อยละ
เคยเข้าร่วม	46	73.016
ไม่เคยเข้าร่วม	17	26.984
รวม	63	100

ภาคผนวก ค
ผลการวิเคราะห์

ผลการวิเคราะห์ข้อสอบและแบบสอบคู่ขนานโดยอาศัยทฤษฎีการตอบสนองข้อสอบของแบบสอบชุด 1A

2PL Model Item Parameter Estimates for Group 1, logit: $a\theta + c$ or $a(\theta - b)$ ([Back to TOC](#))

Item	<i>a</i>	<i>s.e.</i>	<i>c</i>	<i>s.e.</i>	<i>b</i>	<i>s.e.</i>	χ^2	<i>df.</i>	Probability
1	0.47	0.20	-1.58	0.19	3.37	1.40	19.51	16	0.2427
2	2.16	0.40	-0.13	0.21	0.06	0.10	11.01	12	0.5297
3	1.08	0.23	-1.11	0.19	1.03	0.22	21.31	15	0.1267
4	0.65	0.18	-0.41	0.15	0.62	0.26	13.86	16	0.6104
5	-0.17	0.18	-1.23	0.16	-7.44	8.28	12.60	15	0.6339
6	0.90	0.20	-0.48	0.16	0.53	0.19	17.67	15	0.2797
7	0.88	0.20	-0.12	0.15	0.14	0.17	11.55	15	0.7137
8	1.45	0.28	0.48	0.18	-0.33	0.13	13.13	12	0.3612
9	0.64	0.18	-0.26	0.14	0.41	0.24	21.89	16	0.1464
10	0.89	0.20	-0.59	0.16	0.67	0.21	13.87	14	0.4610
11	0.80	0.20	-0.88	0.16	1.10	0.30	12.83	15	0.6165
12	0.58	0.17	0.06	0.14	-0.11	0.25	28.32	15	0.0196
13	1.16	0.24	0.37	0.16	-0.31	0.15	16.16	13	0.2397
14	1.78	0.33	-0.15	0.19	0.08	0.11	15.79	12	0.2004
15	1.52	0.29	-0.06	0.18	0.04	0.12	14.66	12	0.2597
16	0.13	0.15	-0.26	0.14	2.09	2.80	9.12	15	0.8716
17	1.34	0.26	0.44	0.17	-0.33	0.14	11.28	13	0.5887
18	0.44	0.17	-0.49	0.14	1.12	0.51	23.87	16	0.0922
19	0.41	0.17	-0.81	0.15	1.97	0.85	10.40	16	0.8455
20	0.78	0.19	-0.40	0.15	0.51	0.22	18.48	14	0.1854
21	0.30	0.17	-0.73	0.15	2.46	1.41	24.40	16	0.0810
22	0.35	0.18	-1.15	0.16	3.32	1.71	12.02	17	0.7997
23	0.94	0.21	-0.58	0.16	0.61	0.19	24.19	14	0.0434
24	0.82	0.20	-0.84	0.16	1.02	0.28	21.73	15	0.1147
25	1.13	0.23	-0.04	0.16	0.03	0.14	28.78	14	0.0112

Likelihood-based Values and Goodness of Fit Statistics ([Back to TOC](#))

Statistics based on the loglikelihood

-2loglikelihood: 6544.65

Akaike Information Criterion (AIC): 6644.65

Bayesian Information Criterion (BIC): 6813.42

Item Information Function Values for Group 1 at 15 Values of θ from -2.8 to 2.8[\(Back to TOC\)](#)

Item	θ :														
	-2.8	-2.4	-2.0	-1.6	-1.2	-0.8	-0.4	-0.0	0.4	0.8	1.2	1.6	2.0	2.4	2.8
1	0.01	0.01	0.02	0.02	0.02	0.02	0.03	0.03	0.04	0.04	0.04	0.05	0.05	0.05	0.05
2	0.01	0.02	0.05	0.12	0.27	0.54	0.92	1.16	1.02	0.65	0.34	0.16	0.07	0.03	0.01
3	0.02	0.03	0.04	0.06	0.09	0.12	0.17	0.22	0.26	0.29	0.29	0.26	0.22	0.18	0.13
4	0.04	0.05	0.06	0.07	0.08	0.09	0.10	0.10	0.11	0.11	0.10	0.10	0.09	0.08	0.07
5	0.01	0.01	0.01	0.01	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
6	0.04	0.05	0.07	0.09	0.12	0.15	0.17	0.19	0.20	0.20	0.19	0.16	0.14	0.11	0.08
7	0.05	0.07	0.09	0.11	0.14	0.16	0.18	0.19	0.19	0.18	0.16	0.13	0.11	0.08	0.06
8	0.06	0.10	0.16	0.25	0.36	0.47	0.53	0.50	0.40	0.29	0.19	0.11	0.07	0.04	0.02
9	0.04	0.05	0.06	0.07	0.08	0.09	0.10	0.10	0.10	0.10	0.10	0.09	0.08	0.07	0.06
10	0.03	0.05	0.06	0.08	0.11	0.13	0.16	0.18	0.19	0.20	0.19	0.17	0.14	0.11	0.09
11	0.03	0.03	0.05	0.06	0.08	0.09	0.11	0.13	0.15	0.16	0.16	0.16	0.14	0.12	0.10
12	0.05	0.06	0.06	0.07	0.08	0.08	0.08	0.09	0.08	0.08	0.07	0.07	0.06	0.05	0.04
13	0.07	0.10	0.15	0.20	0.26	0.31	0.34	0.33	0.28	0.23	0.17	0.12	0.08	0.05	0.03
14	0.02	0.04	0.07	0.14	0.27	0.45	0.66	0.79	0.74	0.54	0.34	0.19	0.10	0.05	0.02
15	0.03	0.05	0.10	0.16	0.26	0.39	0.51	0.57	0.53	0.42	0.29	0.18	0.11	0.06	0.03
16	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
17	0.06	0.10	0.16	0.23	0.32	0.41	0.45	0.43	0.36	0.27	0.18	0.12	0.07	0.04	0.03
18	0.02	0.03	0.03	0.03	0.04	0.04	0.04	0.05	0.05	0.05	0.05	0.05	0.05	0.04	0.04
19	0.02	0.02	0.02	0.03	0.03	0.03	0.03	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04
20	0.04	0.05	0.07	0.08	0.10	0.12	0.14	0.15	0.15	0.15	0.14	0.13	0.11	0.09	0.08
21	0.01	0.01	0.01	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
22	0.01	0.01	0.01	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.03	0.03	0.03	0.03	0.03
23	0.03	0.05	0.06	0.09	0.12	0.15	0.18	0.20	0.22	0.22	0.21	0.18	0.15	0.12	0.09
24	0.03	0.04	0.05	0.06	0.08	0.10	0.12	0.14	0.16	0.17	0.17	0.16	0.14	0.12	0.10
25	0.05	0.07	0.11	0.15	0.20	0.26	0.30	0.32	0.31	0.27	0.21	0.16	0.11	0.08	0.05
Test Information	1.77	2.09	2.56	3.23	4.14	5.26	6.37	6.95	6.63	5.69	4.67	3.83	3.18	2.69	2.31
Expected s.e.:	0.75	0.69	0.63	0.56	0.49	0.44	0.40	0.38	0.39	0.42	0.46	0.51	0.56	0.61	0.66

Marginal Reliability for Response Pattern Scores: 0.80

ผลการวิเคราะห์ข้อสอบและแบบสอบคู่ขนานโดยอาศัยทฤษฎีการตอบสนองข้อสอบของแบบสอบชุด 1B

2PL Model Item Parameter Estimates for Group 1, logit: $a\theta + c$ or $a(\theta - b)$ [\(Back to TOC\)](#)

Item		<i>a</i>	<i>s.e.</i>	<i>c</i>	<i>s.e.</i>	<i>b</i>	<i>s.e.</i>	χ^2	<i>df.</i>	Probability
1	²	0.21	0.17	¹ -0.99	0.15	4.60	3.67	14.88	15	0.4614
2	⁴	0.97	0.21	³ -0.75	0.17	0.77	0.21	14.58	14	0.4092
3	⁶	0.93	0.21	⁵ -1.01	0.17	1.09	0.27	15.04	15	0.4503
4	⁸	0.36	0.17	⁷ -0.87	0.15	2.41	1.16	18.03	15	0.2605
5	¹⁰	1.14	0.24	⁹ -1.02	0.18	0.90	0.20	13.73	13	0.3945
6	¹²	0.68	0.18	¹¹ -0.39	0.15	0.57	0.25	19.45	15	0.1934
7	¹⁴	0.94	0.21	¹³ -0.03	0.15	0.04	0.16	16.04	14	0.3098
8	¹⁶	2.37	0.46	¹⁵ 0.37	0.23	-0.16	0.10	6.44	11	0.8432
9	¹⁸	0.33	0.17	¹⁷ -0.65	0.14	1.96	1.04	6.71	16	0.9785
10	²⁰	0.61	0.20	¹⁹ -1.21	0.17	1.99	0.61	12.11	14	0.5986
11	²²	0.81	0.20	²¹ -0.69	0.16	0.85	0.25	14.81	16	0.5400
12	²⁴	0.59	0.17	²³ -0.08	0.14	0.13	0.25	13.67	14	0.4764
13	²⁶	0.90	0.21	²⁵ 0.29	0.15	-0.33	0.18	10.32	14	0.7395
14	²⁸	1.44	0.27	²⁷ -0.10	0.18	0.07	0.12	14.35	12	0.2784
15	³⁰	1.38	0.26	²⁹ 0.05	0.17	-0.04	0.13	18.62	12	0.0978
16	³²	-0.04	0.16	³¹ -0.69	0.14	-19.51	88.40	14.52	16	0.5612
17	³⁴	1.52	0.30	³³ 0.75	0.19	-0.49	0.14	11.32	12	0.5029
18	³⁶	0.13	0.23	³⁵ -1.95	0.21	15.35	27.60	17.12	12	0.1448
19	³⁸	0.20	0.20	³⁷ -1.52	0.18	7.70	7.58	23.88	15	0.0669
20	⁴⁰	0.64	0.18	³⁹ -0.62	0.15	0.96	0.33	17.73	14	0.2185
21	⁴²	0.98	0.21	⁴¹ 0.03	0.15	-0.03	0.16	22.21	14	0.0741
22	⁴⁴	0.55	0.19	⁴³ -1.02	0.16	1.84	0.63	13.17	14	0.5148
23	⁴⁶	0.73	0.20	⁴⁵ -0.94	0.16	1.28	0.36	10.53	14	0.7234
24	⁴⁸	1.91	0.38	⁴⁷ 0.89	0.24	-0.46	0.11	20.02	10	0.0289
25	⁵⁰	0.90	0.21	⁴⁹ -0.50	0.15	0.56	0.20	11.70	14	0.6318

Likelihood-based Values and Goodness of Fit Statistics [\(Back to TOC\)](#)

Statistics based on the loglikelihood	
-2loglikelihood:	6362.58
Akaike Information Criterion (AIC):	6462.58
Bayesian Information Criterion (BIC):	6631.34

Item Information Function Values for Group 1 at 15 Values of θ from -2.8 to 2.8[\(Back to TOC\)](#)

Item	θ :														
	-2.8	-2.4	-2.0	-1.6	-1.2	-0.8	-0.4	-0.0	0.4	0.8	1.2	1.6	2.0	2.4	2.8
1	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
2	0.03	0.04	0.06	0.08	0.11	0.14	0.17	0.21	0.23	0.23	0.22	0.20	0.17	0.13	0.10
3	0.02	0.03	0.04	0.06	0.08	0.11	0.14	0.17	0.19	0.21	0.21	0.20	0.18	0.15	0.12
4	0.01	0.02	0.02	0.02	0.02	0.02	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03
5	0.02	0.03	0.04	0.07	0.10	0.14	0.20	0.25	0.30	0.32	0.31	0.28	0.22	0.17	0.12
6	0.04	0.05	0.06	0.07	0.08	0.09	0.10	0.11	0.11	0.11	0.11	0.10	0.09	0.08	0.07
7	0.05	0.07	0.10	0.13	0.16	0.19	0.21	0.22	0.22	0.20	0.17	0.14	0.10	0.08	0.06
8	0.01	0.03	0.07	0.17	0.40	0.82	1.29	1.35	0.93	0.48	0.21	0.09	0.03	0.01	0.01
9	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.03	0.03	0.03	0.03	0.03	0.03	0.03
10	0.02	0.02	0.03	0.03	0.04	0.05	0.06	0.07	0.07	0.08	0.09	0.09	0.09	0.09	0.09
11	0.03	0.04	0.05	0.07	0.09	0.11	0.13	0.15	0.16	0.17	0.16	0.15	0.13	0.11	0.09
12	0.04	0.05	0.06	0.07	0.07	0.08	0.08	0.09	0.09	0.08	0.08	0.07	0.07	0.06	0.05
13	0.07	0.09	0.12	0.15	0.17	0.19	0.20	0.20	0.18	0.16	0.13	0.10	0.08	0.06	0.04
14	0.03	0.06	0.10	0.16	0.25	0.36	0.46	0.52	0.49	0.40	0.29	0.19	0.11	0.07	0.04
15	0.04	0.07	0.11	0.18	0.27	0.37	0.45	0.48	0.44	0.35	0.25	0.16	0.10	0.06	0.04
16	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
17	0.07	0.11	0.19	0.30	0.44	0.54	0.57	0.50	0.37	0.25	0.15	0.09	0.05	0.03	0.02
18	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
19	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
20	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.09	0.10	0.10	0.10	0.10	0.09	0.08	0.07
21	0.06	0.08	0.11	0.14	0.18	0.21	0.23	0.24	0.23	0.20	0.17	0.13	0.10	0.07	0.05
22	0.02	0.02	0.03	0.03	0.04	0.05	0.05	0.06	0.07	0.07	0.07	0.08	0.08	0.07	0.07
23	0.02	0.03	0.04	0.05	0.06	0.08	0.09	0.11	0.12	0.13	0.13	0.13	0.12	0.11	0.10
24	0.04	0.09	0.18	0.34	0.58	0.82	0.91	0.75	0.49	0.27	0.14	0.07	0.03	0.02	0.01
25	0.04	0.05	0.07	0.09	0.11	0.14	0.17	0.19	0.20	0.20	0.18	0.16	0.14	0.11	0.08
Test Information:	1.73	2.05	2.55	3.29	4.35	5.63	6.67	6.81	6.07	5.09	4.26	3.61	3.08	2.65	2.31
Expected s.e.:	0.76	0.70	0.63	0.55	0.48	0.42	0.39	0.38	0.41	0.44	0.48	0.53	0.57	0.61	0.66

Marginal Reliability for Response Pattern Scores: 0.80

ผลการวิเคราะห์ข้อสอบและแบบสอบคู่ขนานโดยอาศัยทฤษฎีการตอบสนองข้อสอบของแบบสอบชุด 2A

2PL Model Item Parameter Estimates for Group 1, logit: $a\theta + c$ or $a(\theta - b)$ [\(Back to TOC\)](#)

Item	a	s.e.	c	s.e.	b	s.e.	χ^2	df.	Probability
1	² -0.24	0.22	¹ -1.70	0.19	-6.98	6.15	20.59	14	0.1123
2	⁴ 0.23	0.18	³ 1.01	0.16	-4.32	3.27	16.59	15	0.3459
3	⁶ 1.39	0.28	⁵ 0.21	0.17	-0.15	0.12	14.67	12	0.2592
4	⁸ 0.09	0.18	⁷ -1.05	0.15	12.22	24.92	18.22	15	0.2504
5	¹⁰ 0.63	0.19	⁹ -0.47	0.15	0.75	0.29	18.45	15	0.2391
6	¹² 1.21	0.25	¹¹ -0.71	0.18	0.58	0.16	9.54	13	0.7320
7	¹⁴ 0.65	0.18	¹³ -0.14	0.15	0.22	0.23	9.01	14	0.8306
8	¹⁶ 0.15	0.16	¹⁵ -0.07	0.14	0.49	1.03	14.19	15	0.5123
9	¹⁸ 1.45	0.29	¹⁷ 0.18	0.17	-0.13	0.12	15.17	12	0.2318
10	²⁰ 0.96	0.22	¹⁹ -0.80	0.17	0.83	0.22	14.29	13	0.3559
11	²² 0.61	0.20	²¹ -0.94	0.16	1.53	0.49	16.50	14	0.2830
12	²⁴ 0.68	0.19	²³ -0.27	0.15	0.39	0.23	18.17	14	0.1985
13	²⁶ 0.64	0.18	²⁵ 0.20	0.14	-0.32	0.24	15.27	14	0.3616
14	²⁸ -0.02	0.18	²⁷ -1.07	0.15	-48.85	393.38	18.57	14	0.1815
15	³⁰ 1.44	0.28	²⁹ -0.02	0.17	0.02	0.12	12.40	12	0.4154
16	³² 0.50	0.18	³¹ -0.50	0.15	0.99	0.42	19.00	15	0.2131
17	³⁴ 0.64	0.19	³³ 0.33	0.15	-0.51	0.26	17.32	14	0.2388
18	³⁶ 0.49	0.18	³⁵ -0.50	0.15	1.01	0.44	21.58	15	0.1189
19	³⁸ 1.22	0.26	³⁷ 0.54	0.17	-0.44	0.15	17.19	13	0.1902
20	⁴⁰ 0.59	0.18	³⁹ -0.47	0.15	0.79	0.33	15.53	13	0.2746
21	⁴² 0.45	0.22	⁴¹ -1.61	0.19	3.56	1.61	9.58	14	0.7925
22	⁴⁴ 1.49	0.30	⁴³ 0.78	0.19	-0.52	0.14	13.19	12	0.3573
23	⁴⁶ 1.32	0.27	⁴⁵ -0.07	0.16	0.05	0.12	17.62	12	0.1275
24	⁴⁸ 0.89	0.21	⁴⁷ -0.66	0.16	0.75	0.22	14.59	13	0.3355
25	⁵⁰ 0.85	0.21	⁴⁹ 0.24	0.15	-0.28	0.18	14.67	14	0.4030

Likelihood-based Values and Goodness of Fit Statistics [\(Back to TOC\)](#)

Statistics based on the loglikelihood

-2loglikelihood: 6577.73

Akaike Information Criterion (AIC): 6677.73

Bayesian Information Criterion (BIC): 6846.50

Item Information Function Values for Group 1 at 15 Values of θ from -2.8 to 2.8
[\(Back to TOC\)](#)

Item	θ :														
	-2.8	-2.4	-2.0	-1.6	-1.2	-0.8	-0.4	-0.0	0.4	0.8	1.2	1.6	2.0	2.4	2.8
1	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.00	0.00
2	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
3	0.05	0.08	0.13	0.20	0.29	0.40	0.47	0.48	0.42	0.32	0.22	0.14	0.09	0.05	0.03
4	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
5	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.09	0.10	0.10	0.10	0.09	0.09	0.08	0.07
6	0.02	0.04	0.06	0.09	0.14	0.19	0.26	0.32	0.36	0.36	0.32	0.26	0.19	0.13	0.09
7	0.05	0.05	0.06	0.08	0.09	0.09	0.10	0.10	0.10	0.10	0.09	0.09	0.08	0.07	0.06
8	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
9	0.04	0.07	0.12	0.20	0.30	0.42	0.51	0.52	0.46	0.35	0.23	0.15	0.09	0.05	0.03
10	0.03	0.04	0.05	0.07	0.10	0.13	0.17	0.20	0.22	0.23	0.22	0.20	0.17	0.14	0.11
11	0.02	0.03	0.03	0.04	0.05	0.06	0.07	0.08	0.08	0.09	0.09	0.09	0.09	0.09	0.08
12	0.04	0.05	0.06	0.07	0.09	0.10	0.11	0.11	0.11	0.11	0.11	0.10	0.09	0.07	0.06
13	0.06	0.07	0.08	0.09	0.10	0.10	0.10	0.10	0.10	0.09	0.08	0.07	0.06	0.05	0.04
14	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
15	0.03	0.06	0.10	0.17	0.26	0.37	0.48	0.52	0.48	0.38	0.27	0.17	0.11	0.06	0.04
16	0.03	0.03	0.04	0.04	0.05	0.05	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.05
17	0.06	0.07	0.08	0.09	0.10	0.10	0.10	0.10	0.09	0.09	0.08	0.07	0.06	0.05	0.04
18	0.03	0.03	0.04	0.04	0.05	0.05	0.05	0.06	0.06	0.06	0.06	0.06	0.06	0.05	0.05
19	0.08	0.11	0.17	0.23	0.30	0.35	0.37	0.34	0.29	0.22	0.16	0.11	0.07	0.04	0.03
20	0.03	0.04	0.05	0.05	0.06	0.07	0.08	0.08	0.09	0.09	0.09	0.08	0.08	0.07	0.06
21	0.01	0.01	0.01	0.02	0.02	0.02	0.03	0.03	0.03	0.04	0.04	0.04	0.05	0.05	0.05
22	0.07	0.12	0.20	0.31	0.43	0.53	0.55	0.48	0.36	0.24	0.15	0.09	0.05	0.03	0.02
23	0.04	0.06	0.10	0.16	0.23	0.32	0.40	0.43	0.41	0.34	0.26	0.18	0.12	0.07	0.04
24	0.03	0.04	0.06	0.08	0.10	0.13	0.15	0.18	0.19	0.20	0.19	0.17	0.15	0.12	0.09
25	0.07	0.09	0.11	0.13	0.15	0.17	0.18	0.18	0.16	0.15	0.12	0.10	0.08	0.06	0.05
Test Information:	1.85	2.18	2.64	3.26	4.01	4.77	5.33	5.49	5.21	4.63	3.96	3.34	2.82	2.41	2.10
Expected s.e.:	0.73	0.68	0.62	0.55	0.50	0.46	0.43	0.43	0.44	0.46	0.50	0.55	0.60	0.64	0.69

Marginal Reliability for Response Pattern Scores: 0.77

ผลการวิเคราะห์ข้อสอบและแบบสอบคู่ขนานโดยอาศัยทฤษฎีการตอบสนองข้อสอบของแบบสอบชุด 2B

2PL Model Item Parameter Estimates for Group 1, logit: $a\theta + c$ or $a(\theta - b)$ [\(Back to TOC\)](#)

Item	a	s.e.	c	s.e.	b	s.e.	χ^2	df.	Probability
1	² -0.05	0.21	¹ -1.51	0.18	-33.12	148.82	11.52	11	0.4027
2	⁴ -0.02	0.23	³ -1.82	0.20	-93.49	1080.46	13.03	10	0.2216
3	⁶ 1.23	0.26	⁵ -0.81	0.19	0.66	0.16	24.38	10	0.0066
4	⁸ 0.27	0.17	⁷ -0.75	0.15	2.81	1.84	28.66	13	0.0073
5	¹⁰ 0.64	0.24	⁹ -1.74	0.21	2.72	0.91	22.37	11	0.0216
6	¹² 0.64	0.19	¹¹ -0.41	0.15	0.64	0.27	33.31	12	0.0009
7	¹⁴ -0.05	0.22	¹³ -1.75	0.19	-33.79	144.59	15.59	10	0.1116
8	¹⁶ 0.78	0.20	¹⁵ 0.00	0.15	-0.00	0.19	20.08	10	0.0284
9	¹⁸ 2.07	0.44	¹⁷ 0.39	0.21	-0.19	0.10	12.13	8	0.1450
10	²⁰ 0.61	0.20	¹⁹ -1.01	0.16	1.65	0.53	10.52	11	0.4856
11	²² 0.60	0.19	²¹ -0.68	0.15	1.14	0.40	21.52	11	0.0283
12	²⁴ 1.43	0.32	²³ 1.20	0.21	-0.83	0.17	10.58	9	0.3072
13	²⁶ 0.83	0.22	²⁵ -1.22	0.18	1.48	0.38	22.42	12	0.0329
14	²⁸ 0.41	0.21	²⁷ -1.50	0.18	3.63	1.76	15.86	11	0.1459
15	³⁰ 1.59	0.32	²⁹ -0.13	0.18	0.08	0.12	29.19	10	0.0012
16	³² 0.31	0.19	³¹ -1.15	0.16	3.73	2.25	11.84	13	0.5425
17	³⁴ 0.69	0.20	³³ -0.95	0.16	1.38	0.40	14.52	10	0.1503
18	³⁶ 0.12	0.17	³⁵ -0.72	0.15	5.94	8.30	28.07	13	0.0088
19	³⁸ 0.81	0.21	³⁷ -0.60	0.16	0.75	0.24	23.98	12	0.0204
20	⁴⁰ -0.03	0.24	³⁹ -1.99	0.21	-71.33	622.74	14.64	10	0.1454
21	⁴² 0.72	0.20	⁴¹ -0.59	0.15	0.82	0.29	27.80	12	0.0059
22	⁴⁴ 0.47	0.18	⁴³ -0.82	0.15	1.75	0.71	38.29	12	0.0001
23	⁴⁶ 1.31	0.28	⁴⁵ -0.44	0.17	0.34	0.14	16.21	10	0.0936
24	⁴⁸ 1.16	0.27	⁴⁷ -1.38	0.21	1.19	0.23	16.93	11	0.1095
25	⁵⁰ 0.39	0.17	⁴⁹ -0.06	0.14	0.15	0.36	22.11	13	0.0536

Likelihood-based Values and Goodness of Fit Statistics [\(Back to TOC\)](#)

Statistics based on the loglikelihood

-2loglikelihood: 6058.92

Akaike Information Criterion (AIC): 6158.92

Bayesian Information Criterion (BIC): 6327.68

Item Information Function Values for Group 1 at 15 Values of θ from -2.8 to 2.8[\(Back to TOC\)](#)

Item	θ :														
	-2.8	-2.4	-2.0	-1.6	-1.2	-0.8	-0.4	-0.0	0.4	0.8	1.2	1.6	2.0	2.4	2.8
1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
3	0.02	0.03	0.05	0.08	0.13	0.18	0.25	0.32	0.37	0.37	0.34	0.27	0.20	0.14	0.09
4	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
5	0.01	0.01	0.02	0.02	0.03	0.04	0.04	0.05	0.06	0.07	0.08	0.09	0.10	0.10	0.10
6	0.04	0.04	0.05	0.06	0.07	0.08	0.09	0.10	0.10	0.10	0.10	0.09	0.09	0.08	0.07
7	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
8	0.06	0.07	0.09	0.10	0.12	0.14	0.15	0.15	0.15	0.14	0.12	0.10	0.09	0.07	0.06
9	0.02	0.04	0.10	0.21	0.42	0.73	1.02	1.03	0.75	0.43	0.22	0.10	0.05	0.02	0.01
10	0.02	0.03	0.03	0.04	0.05	0.06	0.06	0.07	0.08	0.09	0.09	0.09	0.09	0.09	0.08
11	0.03	0.03	0.04	0.05	0.06	0.07	0.07	0.08	0.09	0.09	0.09	0.09	0.08	0.08	0.07
12	0.11	0.18	0.27	0.39	0.48	0.51	0.47	0.37	0.26	0.16	0.10	0.06	0.03	0.02	0.01
13	0.02	0.03	0.03	0.05	0.06	0.08	0.10	0.12	0.14	0.16	0.17	0.17	0.16	0.15	0.13
14	0.01	0.01	0.01	0.02	0.02	0.02	0.02	0.03	0.03	0.03	0.03	0.04	0.04	0.04	0.04
15	0.03	0.05	0.09	0.15	0.26	0.40	0.55	0.63	0.59	0.46	0.31	0.19	0.11	0.06	0.03
16	0.01	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
17	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.10	0.11	0.11	0.12	0.12	0.11	0.11	0.09
18	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
19	0.03	0.04	0.06	0.07	0.09	0.11	0.13	0.15	0.16	0.16	0.16	0.15	0.13	0.11	0.09
20	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
21	0.03	0.04	0.05	0.07	0.08	0.09	0.11	0.12	0.13	0.13	0.13	0.12	0.11	0.09	0.08
22	0.02	0.02	0.03	0.03	0.03	0.04	0.04	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
23	0.03	0.04	0.07	0.12	0.18	0.26	0.34	0.41	0.43	0.39	0.32	0.23	0.16	0.10	0.06
24	0.01	0.02	0.03	0.05	0.07	0.11	0.16	0.22	0.27	0.32	0.34	0.32	0.27	0.21	0.16
25	0.03	0.03	0.03	0.03	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.03	0.03	0.03
Test Information:	1.56	1.79	2.13	2.62	3.28	4.06	4.77	5.05	4.83	4.36	3.84	3.37	2.95	2.60	2.30
Expected s.e.:	0.80	0.75	0.68	0.62	0.55	0.50	0.46	0.44	0.45	0.48	0.51	0.55	0.58	0.62	0.66

Marginal Reliability for Response Pattern Scores: 0.75

ผลการวิเคราะห์ข้อสอบและแบบสอบคู่ขนานโดยอาศัยทฤษฎีการตอบสนองข้อสอบของแบบสอบชุด 3A

2PL Model Item Parameter Estimates for Group 1, logit: $a\theta + c$ or $a(\theta - b)$

[\(Back to TOC\)](#)

Item	a	s.e.	c	s.e.	b	s.e.	χ^2	df.	Probability
1	² -0.05	0.20	¹ -1.58	0.18	-32.15	132.23	24.88	13	0.0239
2	⁴ 1.24	0.25	³ 0.11	0.16	-0.08	0.13	17.21	12	0.1416
3	⁶ -0.23	0.18	⁵ -1.11	0.15	-4.80	3.72	22.45	14	0.0697
4	⁸ 0.52	0.19	⁷ -1.08	0.16	2.10	0.76	34.40	14	0.0018
5	¹⁰ -0.05	0.17	⁹ -0.85	0.15	-16.86	56.28	33.47	15	0.0040
6	¹² 0.63	0.20	¹¹ -1.21	0.17	1.93	0.60	22.09	14	0.0765
7	¹⁴ 0.93	0.21	¹³ -0.49	0.15	0.52	0.19	41.75	14	0.0001
8	¹⁶ 0.94	0.21	¹⁵ 0.40	0.16	-0.42	0.18	23.37	13	0.0373
9	¹⁸ 1.29	0.25	¹⁷ 0.35	0.16	-0.27	0.13	22.67	12	0.0305
10	²⁰ 0.90	0.21	¹⁹ -0.35	0.15	0.39	0.18	34.64	13	0.0010
11	²² 0.83	0.21	²¹ 0.87	0.16	-1.05	0.28	22.15	13	0.0529
12	²⁴ 0.37	0.19	²³ -1.30	0.17	3.54	1.82	36.05	15	0.0017
13	²⁶ 1.19	0.24	²⁵ 0.53	0.16	-0.44	0.15	16.49	12	0.1693
14	²⁸ 0.28	0.21	²⁷ -1.58	0.18	5.54	3.95	26.29	13	0.0155
15	³⁰ 0.70	0.23	²⁹ -1.57	0.19	2.26	0.67	33.94	12	0.0007
16	³² 0.35	0.16	³¹ -0.34	0.14	0.98	0.58	26.79	15	0.0304
17	³⁴ 1.31	0.26	³³ 0.55	0.17	-0.42	0.14	19.53	12	0.0764
18	³⁶ 1.33	0.26	³⁵ 0.21	0.16	-0.15	0.12	26.69	12	0.0085
19	³⁸ 0.52	0.18	³⁷ -0.56	0.14	1.07	0.43	29.79	14	0.0081
20	⁴⁰ 0.26	0.16	³⁹ -0.43	0.14	1.70	1.17	23.98	14	0.0460
21	⁴² 1.03	0.22	⁴¹ 0.01	0.15	-0.01	0.15	20.31	12	0.0612
22	⁴⁴ 1.23	0.25	⁴³ 0.32	0.16	-0.26	0.13	25.92	12	0.0110
23	⁴⁶ 0.87	0.20	⁴⁵ -0.35	0.15	0.40	0.18	32.45	13	0.0021
24	⁴⁸ 0.71	0.19	⁴⁷ -0.33	0.14	0.47	0.22	43.80	13	0.0001
25	⁵⁰ 0.91	0.21	⁴⁹ -0.20	0.15	0.22	0.17	33.43	12	0.0008

Likelihood-based Values and Goodness of Fit Statistics [\(Back to TOC\)](#)

Statistics based on the loglikelihood

-2loglikelihood: 6658.42

Akaike Information Criterion (AIC): 6758.42

Bayesian Information Criterion (BIC): 6928.78

Item Information Function Values for Group 1 at 15 Values of θ from -2.8 to 2.8[\(Back to TOC\)](#)

Item	θ :														
	-2.8	-2.4	-2.0	-1.6	-1.2	-0.8	-0.4	-0.0	0.4	0.8	1.2	1.6	2.0	2.4	2.8
1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2	0.05	0.08	0.12	0.18	0.25	0.32	0.37	0.39	0.35	0.29	0.22	0.15	0.10	0.06	0.04
3	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
4	0.02	0.02	0.03	0.03	0.03	0.04	0.04	0.05	0.06	0.06	0.06	0.07	0.07	0.07	0.06
5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
6	0.02	0.02	0.03	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.09	0.10	0.10	0.10	0.09
7	0.04	0.05	0.07	0.09	0.12	0.15	0.18	0.21	0.22	0.21	0.20	0.17	0.14	0.11	0.08
8	0.08	0.10	0.13	0.17	0.19	0.21	0.22	0.21	0.19	0.16	0.13	0.10	0.07	0.05	0.04
9	0.06	0.09	0.15	0.21	0.30	0.37	0.41	0.40	0.35	0.27	0.19	0.13	0.08	0.05	0.03
10	0.04	0.06	0.08	0.10	0.13	0.15	0.18	0.20	0.20	0.20	0.18	0.15	0.12	0.10	0.07
11	0.11	0.13	0.15	0.16	0.17	0.17	0.16	0.14	0.12	0.10	0.08	0.06	0.05	0.04	0.03
12	0.01	0.01	0.01	0.02	0.02	0.02	0.02	0.02	0.02	0.03	0.03	0.03	0.03	0.03	0.03
13	0.08	0.11	0.17	0.23	0.29	0.34	0.35	0.33	0.28	0.21	0.15	0.11	0.07	0.05	0.03
14	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.02
15	0.01	0.02	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.11	0.11	0.12	0.12	0.12
16	0.02	0.02	0.02	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03
17	0.07	0.11	0.17	0.25	0.33	0.40	0.43	0.40	0.33	0.24	0.16	0.11	0.07	0.04	0.02
18	0.05	0.08	0.13	0.20	0.28	0.37	0.43	0.44	0.39	0.30	0.22	0.14	0.09	0.06	0.03
19	0.03	0.03	0.04	0.04	0.05	0.05	0.06	0.06	0.06	0.07	0.07	0.07	0.06	0.06	0.06
20	0.01	0.01	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
21	0.05	0.08	0.11	0.14	0.19	0.23	0.25	0.26	0.25	0.22	0.18	0.14	0.11	0.08	0.05
22	0.06	0.09	0.14	0.20	0.28	0.34	0.38	0.37	0.32	0.25	0.19	0.13	0.08	0.05	0.03
23	0.04	0.06	0.07	0.10	0.12	0.15	0.17	0.18	0.19	0.18	0.17	0.15	0.12	0.10	0.07
24	0.04	0.05	0.06	0.08	0.09	0.10	0.11	0.12	0.12	0.12	0.12	0.11	0.09	0.08	0.07
25	0.05	0.06	0.09	0.11	0.14	0.17	0.19	0.20	0.21	0.19	0.17	0.14	0.11	0.09	0.07
Test Information:	1.95	2.32	2.81	3.43	4.12	4.75	5.15	5.20	4.90	4.37	3.78	3.22	2.76	2.39	2.10
Expected s.e.:	0.72	0.66	0.60	0.54	0.49	0.46	0.44	0.44	0.45	0.48	0.51	0.56	0.60	0.65	0.69

Marginal Reliability for Response Pattern Scores: 0.77

ผลการวิเคราะห์ข้อสอบและแบบสอบคู่ขนานโดยอาศัยทฤษฎีการตอบสนองข้อสอบของแบบสอบชุด 3B

2PL Model Item Parameter Estimates for Group 1, logit: $a\theta + c$ or $a(\theta - b)$ [\(Back to TOC\)](#)

Item	a	s.e.	c	s.e.	b	s.e.	χ^2	df.	Probability
1	² 0.12	0.22	¹ -1.79	0.19	14.86	26.66	11.17	12	0.5154
2	⁴ 1.10	0.23	³ -0.84	0.17	0.76	0.19	14.96	13	0.3120
3	⁶ 0.46	0.18	⁵ -0.73	0.15	1.58	0.62	11.70	13	0.5536
4	⁸ 0.69	0.19	⁷ -0.37	0.15	0.54	0.24	15.17	13	0.2962
5	¹⁰ 0.36	0.19	⁹ -1.25	0.16	3.50	1.80	11.81	14	0.6223
6	¹² 0.75	0.19	¹¹ -0.63	0.15	0.84	0.26	8.97	13	0.7756
7	¹⁴ 0.41	0.21	¹³ -1.60	0.18	3.94	1.95	9.85	13	0.7072
8	¹⁶ 3.14	0.74	¹⁵ -0.05	0.27	0.02	0.09	4.12	8	0.8467
9	¹⁸ 0.01	0.21	¹⁷ -1.68	0.18	116.74	1684.68	7.72	12	0.8075
10	²⁰ 1.31	0.26	¹⁹ -0.19	0.17	0.15	0.13	9.68	12	0.6446
11	²² 0.73	0.19	²¹ 0.42	0.15	-0.57	0.24	10.15	13	0.6823
12	²⁴ 0.97	0.22	²³ -0.72	0.16	0.74	0.20	15.88	13	0.2551
13	²⁶ 0.28	0.19	²⁵ -1.40	0.17	4.92	3.29	10.31	13	0.6696
14	²⁸ 0.41	0.18	²⁷ -0.92	0.15	2.25	0.98	17.26	14	0.2419
15	³⁰ 0.71	0.19	²⁹ -0.13	0.14	0.18	0.21	9.70	12	0.6432
16	³² 0.49	0.20	³¹ -1.36	0.17	2.78	1.07	17.25	13	0.1878
17	³⁴ 0.62	0.18	³³ -0.55	0.15	0.88	0.32	12.34	14	0.5801
18	³⁶ 0.42	0.18	³⁵ -1.04	0.16	2.47	1.05	12.89	14	0.5365
19	³⁸ -0.19	0.24	³⁷ -2.00	0.21	-10.55	13.21	14.27	11	0.2180
20	⁴⁰ 0.60	0.19	³⁹ -1.13	0.17	1.89	0.59	17.29	14	0.2406
21	⁴² 0.59	0.22	⁴¹ -1.60	0.19	2.69	0.91	19.14	13	0.1185
22	⁴⁴ 1.33	0.26	⁴³ -0.54	0.18	0.41	0.14	6.49	12	0.8894
23	⁴⁶ 0.66	0.19	⁴⁵ -0.62	0.15	0.93	0.31	21.17	14	0.0970
24	⁴⁸ 0.91	0.20	⁴⁷ -0.24	0.15	0.26	0.17	8.84	13	0.7853
25	⁵⁰ 1.60	0.30	⁴⁹ -0.05	0.18	0.03	0.11	9.98	11	0.5339

Likelihood-based Values and Goodness of Fit Statistics [\(Back to TOC\)](#)

Statistics based on the loglikelihood	
-2loglikelihood:	6277.14
Akaike Information Criterion (AIC):	6377.14
Bayesian Information Criterion (BIC):	6547.50

Item Information Function Values for Group 1 at 15 Values of θ from -2.8 to 2.8
[\(Back to TOC\)](#)

Item	θ :														
	-2.8	-2.4	-2.0	-1.6	-1.2	-0.8	-0.4	-0.0	0.4	0.8	1.2	1.6	2.0	2.4	2.8
1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2	0.02	0.04	0.05	0.08	0.11	0.16	0.21	0.26	0.29	0.30	0.29	0.25	0.20	0.15	0.11
3	0.02	0.03	0.03	0.03	0.04	0.04	0.04	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
4	0.04	0.05	0.06	0.07	0.08	0.10	0.11	0.11	0.12	0.12	0.11	0.10	0.09	0.08	0.07
5	0.01	0.01	0.01	0.02	0.02	0.02	0.02	0.02	0.02	0.03	0.03	0.03	0.03	0.03	0.03
6	0.03	0.04	0.05	0.07	0.08	0.10	0.11	0.13	0.14	0.14	0.14	0.13	0.12	0.10	0.09
7	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.02	0.03	0.03	0.03	0.03	0.04	0.04	0.04
8	0.00	0.00	0.02	0.06	0.21	0.66	1.66	2.47	1.75	0.71	0.23	0.07	0.02	0.01	0.00
9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
10	0.03	0.06	0.09	0.14	0.21	0.30	0.38	0.43	0.42	0.36	0.28	0.19	0.13	0.08	0.05
11	0.07	0.09	0.10	0.12	0.13	0.13	0.13	0.13	0.12	0.10	0.09	0.07	0.06	0.05	0.04
12	0.03	0.04	0.06	0.08	0.11	0.14	0.18	0.21	0.23	0.24	0.22	0.20	0.17	0.13	0.10
13	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.02	0.02
14	0.02	0.02	0.02	0.02	0.03	0.03	0.03	0.03	0.04	0.04	0.04	0.04	0.04	0.04	0.04
15	0.05	0.06	0.07	0.09	0.10	0.11	0.12	0.12	0.12	0.12	0.11	0.10	0.08	0.07	0.06
16	0.01	0.02	0.02	0.02	0.03	0.03	0.03	0.04	0.04	0.05	0.05	0.06	0.06	0.06	0.06
17	0.03	0.04	0.05	0.06	0.07	0.07	0.08	0.09	0.09	0.10	0.10	0.09	0.09	0.08	0.07
18	0.02	0.02	0.02	0.02	0.03	0.03	0.03	0.03	0.04	0.04	0.04	0.04	0.04	0.04	0.04
19	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
20	0.02	0.02	0.03	0.04	0.04	0.05	0.06	0.07	0.07	0.08	0.09	0.09	0.09	0.09	0.08
21	0.01	0.02	0.02	0.02	0.03	0.03	0.04	0.05	0.06	0.07	0.07	0.08	0.08	0.09	0.09
22	0.02	0.04	0.07	0.11	0.17	0.25	0.34	0.41	0.44	0.41	0.34	0.25	0.17	0.11	0.07
23	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.10	0.11	0.11	0.11	0.10	0.10	0.09	0.08
24	0.05	0.06	0.08	0.11	0.14	0.16	0.19	0.20	0.20	0.19	0.17	0.15	0.12	0.09	0.07
25	0.03	0.05	0.09	0.16	0.27	0.42	0.57	0.64	0.58	0.45	0.30	0.18	0.10	0.06	0.03
Test Information:	1.58	1.76	2.02	2.40	2.98	3.94	5.46	6.63	5.98	4.75	3.90	3.33	2.89	2.55	2.28
Expected s.e.:	0.80	0.75	0.70	0.65	0.58	0.50	0.43	0.39	0.41	0.46	0.51	0.55	0.59	0.63	0.66

Marginal Reliability for Response Pattern Scores: 0.76

ผลการวิเคราะห์ข้อสอบและแบบสอบคู่ขนานโดยอาศัยทฤษฎีการตอบสนองข้อสอบของแบบสอบชุด 4A

2PL Model Item Parameter Estimates for Group 1, logit: $a\theta + c$ or $a(\theta - b)$ [\(Back to TOC\)](#)

Item		<i>a</i>	<i>s.e.</i>	<i>c</i>	<i>s.e.</i>	<i>b</i>	<i>s.e.</i>	χ^2	<i>df.</i>	Probability
1	²	-0.51	0.20	¹ -1.34	0.17	-2.65	0.99	15.77	14	0.3293
2	⁴	0.16	0.18	³ -1.25	0.16	7.92	8.87	9.13	15	0.8711
3	⁶	1.40	0.25	⁵ -0.12	0.17	0.09	0.12	20.29	13	0.0880
4	⁸	0.80	0.19	⁷ -0.66	0.16	0.82	0.25	17.77	15	0.2741
5	¹⁰	0.99	0.20	⁹ -0.51	0.16	0.52	0.18	13.70	15	0.5495
6	¹²	0.89	0.19	¹¹ -0.39	0.15	0.44	0.18	15.94	15	0.3878
7	¹⁴	1.86	0.32	¹³ 0.03	0.19	-0.02	0.10	25.81	12	0.0114
8	¹⁶	2.29	0.40	¹⁵ 0.44	0.22	-0.19	0.10	13.29	11	0.2743
9	¹⁸	0.81	0.19	¹⁷ -0.25	0.15	0.31	0.19	16.36	16	0.4298
10	²⁰	0.75	0.19	¹⁹ -0.72	0.16	0.95	0.27	13.29	15	0.5814
11	²²	0.54	0.17	²¹ -0.66	0.15	1.22	0.44	18.83	16	0.2768
12	²⁴	0.82	0.19	²³ 0.16	0.15	-0.19	0.18	23.45	14	0.0531
13	²⁶	0.78	0.19	²⁵ -0.53	0.15	0.67	0.23	22.15	15	0.1036
14	²⁸	1.62	0.28	²⁷ -0.19	0.19	0.11	0.11	22.63	13	0.0461
15	³⁰	0.46	0.18	²⁹ -1.05	0.16	2.27	0.86	17.83	15	0.2708
16	³²	0.72	0.18	³¹ 0.23	0.14	-0.33	0.21	22.15	16	0.1380
17	³⁴	0.30	0.16	³³ -0.48	0.14	1.61	0.94	17.65	16	0.3467
18	³⁶	0.17	0.16	³⁵ -0.66	0.14	3.95	3.74	10.77	16	0.8238
19	³⁸	1.24	0.24	³⁷ 0.55	0.17	-0.44	0.15	21.16	14	0.0972
20	⁴⁰	0.69	0.20	³⁹ -1.18	0.17	1.70	0.46	24.40	15	0.0584
21	⁴²	0.87	0.19	⁴¹ -0.45	0.15	0.52	0.19	13.76	15	0.5452
22	⁴⁴	1.98	0.35	⁴³ -0.45	0.20	0.23	0.11	14.13	12	0.2916
23	⁴⁶	0.58	0.18	⁴⁵ -0.75	0.15	1.30	0.43	12.62	15	0.6327
24	⁴⁸	1.70	0.30	⁴⁷ 0.42	0.19	-0.25	0.11	20.19	12	0.0634
25	⁵⁰	0.50	0.16	⁴⁹ -0.16	0.14	0.32	0.29	17.18	16	0.3758

Likelihood-based Values and Goodness of Fit Statistics [\(Back to TOC\)](#)

Statistics based on the loglikelihood

-2loglikelihood: 6706.64

Akaike Information Criterion (AIC): 6806.64

Bayesian Information Criterion (BIC): 6977.00

Item Information Function Values for Group 1 at 15 Values of θ from -2.8 to 2.8[\(Back to TOC\)](#)

Item	θ :														
	-2.8	-2.4	-2.0	-1.6	-1.2	-0.8	-0.4	-0.0	0.4	0.8	1.2	1.6	2.0	2.4	2.8
1	0.06	0.06	0.06	0.06	0.06	0.05	0.05	0.04	0.04	0.03	0.03	0.02	0.02	0.02	0.01
2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.01
3	0.03	0.06	0.09	0.15	0.24	0.34	0.44	0.49	0.47	0.39	0.28	0.19	0.12	0.07	0.04
4	0.03	0.04	0.05	0.07	0.09	0.11	0.13	0.14	0.16	0.16	0.16	0.15	0.13	0.11	0.09
5	0.03	0.05	0.07	0.10	0.13	0.16	0.20	0.23	0.24	0.24	0.22	0.19	0.15	0.11	0.08
6	0.04	0.05	0.07	0.10	0.12	0.15	0.17	0.19	0.20	0.19	0.18	0.15	0.13	0.10	0.08
7	0.02	0.04	0.08	0.16	0.31	0.53	0.76	0.86	0.75	0.51	0.29	0.15	0.08	0.04	0.02
8	0.01	0.03	0.08	0.19	0.43	0.84	1.25	1.25	0.85	0.44	0.20	0.08	0.03	0.01	0.01
9	0.05	0.06	0.08	0.09	0.11	0.13	0.15	0.16	0.16	0.16	0.14	0.13	0.11	0.09	0.07
10	0.03	0.04	0.05	0.06	0.08	0.09	0.11	0.12	0.14	0.14	0.14	0.13	0.12	0.11	0.09
11	0.03	0.03	0.04	0.04	0.05	0.06	0.06	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.06
12	0.06	0.08	0.10	0.12	0.14	0.16	0.17	0.17	0.16	0.14	0.12	0.10	0.08	0.06	0.05
13	0.04	0.05	0.06	0.08	0.09	0.11	0.13	0.14	0.15	0.15	0.15	0.13	0.12	0.10	0.08
14	0.02	0.04	0.08	0.14	0.25	0.40	0.56	0.65	0.63	0.49	0.33	0.20	0.11	0.06	0.03
15	0.02	0.02	0.02	0.03	0.03	0.03	0.04	0.04	0.04	0.05	0.05	0.05	0.05	0.05	0.05
16	0.06	0.08	0.09	0.11	0.12	0.13	0.13	0.13	0.12	0.11	0.10	0.08	0.07	0.06	0.04
17	0.01	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
18	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
19	0.07	0.11	0.17	0.24	0.31	0.37	0.39	0.36	0.30	0.22	0.16	0.11	0.07	0.04	0.03
20	0.02	0.03	0.03	0.04	0.05	0.06	0.07	0.09	0.10	0.11	0.12	0.12	0.12	0.11	0.10
21	0.04	0.05	0.07	0.09	0.11	0.14	0.16	0.18	0.19	0.19	0.17	0.15	0.13	0.10	0.08
22	0.01	0.02	0.05	0.10	0.21	0.40	0.68	0.93	0.95	0.72	0.43	0.23	0.11	0.05	0.02
23	0.03	0.03	0.04	0.04	0.05	0.06	0.07	0.07	0.08	0.08	0.08	0.08	0.08	0.08	0.07
24	0.04	0.07	0.13	0.24	0.40	0.58	0.71	0.69	0.54	0.36	0.21	0.12	0.06	0.03	0.02
25	0.04	0.04	0.05	0.05	0.05	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.05	0.05	0.04
Test Information:	1.80	2.12	2.59	3.34	4.46	5.99	7.51	8.11	7.42	6.05	4.73	3.73	3.04	2.56	2.21
Expected s.e.:	0.74	0.69	0.62	0.55	0.47	0.41	0.36	0.35	0.37	0.41	0.46	0.52	0.57	0.63	0.67

Marginal Reliability for Response Pattern Scores: 0.82

ผลการวิเคราะห์ข้อสอบและแบบสอบคู่ขนานโดยอาศัยทฤษฎีการตอบสนองข้อสอบของแบบสอบชุด 4B

2PL Model Item Parameter Estimates for Group 1, logit: $a\theta + c$ or $a(\theta - b)$ [\(Back to TOC\)](#)

Item	a	s.e.	c	s.e.	b	s.e.	χ^2	df.	Probability
1	² -0.29	0.23	¹ -1.85	0.20	-6.47	5.10	9.53	12	0.6582
2	⁴ 0.76	0.20	³ -0.94	0.16	1.25	0.33	6.98	13	0.9035
3	⁶ 0.91	0.21	⁵ -0.54	0.16	0.59	0.20	14.89	12	0.2466
4	⁸ 0.62	0.18	⁷ -0.85	0.15	1.37	0.43	9.05	13	0.7699
5	¹⁰ 0.72	0.19	⁹ -0.33	0.15	0.45	0.22	9.01	14	0.8309
6	¹² 1.01	0.22	¹¹ -0.26	0.15	0.26	0.16	8.28	13	0.8255
7	¹⁴ 1.84	0.34	¹³ -0.48	0.20	0.26	0.11	26.49	12	0.0091
8	¹⁶ 2.61	0.55	¹⁵ 0.24	0.23	-0.09	0.09	10.30	9	0.3282
9	¹⁸ 1.05	0.22	¹⁷ -0.35	0.16	0.34	0.15	10.10	12	0.6087
10	²⁰ 0.56	0.19	¹⁹ -1.30	0.17	2.32	0.77	9.51	14	0.7979
11	²² 0.81	0.20	²¹ -0.84	0.16	1.03	0.27	10.16	13	0.6821
12	²⁴ 1.08	0.22	²³ 0.16	0.15	-0.15	0.14	12.84	13	0.4619
13	²⁶ 0.98	0.23	²⁵ 0.66	0.16	-0.67	0.19	11.66	12	0.4744
14	²⁸ 1.40	0.27	²⁷ -0.36	0.17	0.26	0.13	7.25	12	0.8409
15	³⁰ 0.71	0.20	²⁹ -1.13	0.17	1.59	0.43	13.71	13	0.3966
16	³² 0.18	0.16	³¹ -0.32	0.14	1.77	1.70	7.59	15	0.9392
17	³⁴ 0.43	0.18	³³ -1.16	0.16	2.73	1.14	12.68	13	0.4746
18	³⁶ -0.26	0.18	³⁵ -1.09	0.16	-4.16	2.88	16.99	14	0.2562
19	³⁸ 0.63	0.18	³⁷ -0.72	0.15	1.14	0.36	9.45	13	0.7388
20	⁴⁰ 0.55	0.20	³⁹ -1.49	0.18	2.73	0.95	19.93	13	0.0967
21	⁴² 0.68	0.18	⁴¹ -0.30	0.14	0.45	0.23	6.97	14	0.9362
22	⁴⁴ 0.97	0.22	⁴³ 0.30	0.15	-0.31	0.16	14.57	13	0.3366
23	⁴⁶ 0.42	0.17	⁴⁵ -0.77	0.15	1.84	0.78	19.51	13	0.1079
24	⁴⁸ 0.01	0.17	⁴⁷ -0.93	0.15	107.35	2060.87	16.60	15	0.3453
25	⁵⁰ 0.67	0.20	⁴⁹ -1.20	0.17	1.78	0.50	14.70	13	0.3281

Likelihood-based Values and Goodness of Fit Statistics [\(Back to TOC\)](#)

Statistics based on the loglikelihood

-2loglikelihood: 6560.46

Akaike Information Criterion (AIC): 6660.46

Bayesian Information Criterion (BIC): 6830.81

Item Information Function Values for Group 1 at 15 Values of θ from -2.8 to 2.8[\(Back to TOC\)](#)

Item	θ :														
	-2.8	-2.4	-2.0	-1.6	-1.2	-0.8	-0.4	-0.0	0.4	0.8	1.2	1.6	2.0	2.4	2.8
1	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
2	0.02	0.03	0.04	0.05	0.07	0.08	0.10	0.12	0.13	0.14	0.14	0.14	0.13	0.12	0.10
3	0.03	0.05	0.07	0.09	0.11	0.14	0.17	0.19	0.20	0.20	0.19	0.17	0.14	0.11	0.09
4	0.03	0.03	0.04	0.05	0.05	0.06	0.07	0.08	0.09	0.09	0.10	0.10	0.09	0.09	0.08
5	0.04	0.05	0.06	0.08	0.09	0.11	0.12	0.13	0.13	0.13	0.12	0.11	0.10	0.08	0.07
6	0.04	0.06	0.09	0.12	0.16	0.19	0.23	0.25	0.25	0.24	0.21	0.17	0.13	0.09	0.07
7	0.01	0.02	0.05	0.10	0.20	0.37	0.59	0.80	0.83	0.67	0.43	0.25	0.13	0.06	0.03
8	0.01	0.02	0.05	0.13	0.34	0.80	1.45	1.68	1.15	0.55	0.22	0.08	0.03	0.01	0.00
9	0.04	0.06	0.08	0.11	0.15	0.20	0.24	0.27	0.27	0.26	0.23	0.18	0.14	0.10	0.07
10	0.02	0.02	0.02	0.03	0.03	0.04	0.05	0.05	0.06	0.07	0.07	0.07	0.08	0.08	0.08
11	0.03	0.04	0.05	0.06	0.08	0.10	0.12	0.14	0.15	0.16	0.16	0.16	0.14	0.12	0.10
12	0.06	0.09	0.12	0.17	0.21	0.26	0.28	0.29	0.27	0.23	0.18	0.13	0.10	0.07	0.04
13	0.09	0.13	0.16	0.20	0.23	0.24	0.24	0.22	0.19	0.15	0.11	0.08	0.06	0.04	0.03
14	0.03	0.05	0.08	0.13	0.20	0.30	0.40	0.47	0.48	0.42	0.33	0.22	0.14	0.09	0.05
15	0.02	0.03	0.03	0.04	0.05	0.07	0.08	0.09	0.11	0.12	0.12	0.13	0.12	0.12	0.11
16	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
17	0.01	0.02	0.02	0.02	0.02	0.03	0.03	0.03	0.04	0.04	0.04	0.04	0.04	0.04	0.05
18	0.02	0.02	0.02	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
19	0.03	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.09	0.10	0.10	0.10	0.09	0.08	0.08
20	0.01	0.02	0.02	0.02	0.03	0.03	0.04	0.04	0.05	0.06	0.06	0.07	0.07	0.07	0.07
21	0.04	0.05	0.06	0.07	0.09	0.10	0.11	0.11	0.12	0.12	0.11	0.10	0.09	0.08	0.06
22	0.07	0.10	0.13	0.16	0.20	0.22	0.23	0.23	0.21	0.18	0.14	0.11	0.08	0.06	0.04
23	0.02	0.02	0.02	0.03	0.03	0.03	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04
24	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
25	0.02	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.10	0.11	0.11	0.11	0.11	0.10
Test Information:	1.71	1.96	2.30	2.78	3.49	4.52	5.76	6.42	5.99	5.08	4.24	3.59	3.09	2.70	2.39
Expected s.e.:	0.76	0.71	0.66	0.60	0.54	0.47	0.42	0.39	0.41	0.44	0.49	0.53	0.57	0.61	0.65

Marginal Reliability for Response Pattern Scores: 0.78

ผลการวิเคราะห์ข้อสอบและแบบสอบคู่ขนานโดยอาศัยทฤษฎีการตอบสนองข้อสอบของแบบสอบชุด 5A

2PL Model Item Parameter Estimates for Group 1, logit: $a\theta + c$ or $a(\theta - b)$

[\(Back to TOC\)](#)

Item	a	s.e.	c	s.e.	b	s.e.	χ^2	df.	Probability
1	² -0.54	0.20	¹ -1.25	0.18	-2.31	0.82	16.85	12	0.1549
2	⁴ -0.51	0.21	³ 1.50	0.18	2.95	1.16	13.69	14	0.4742
3	⁶ 1.62	0.31	⁵ 0.36	0.19	-0.22	0.12	20.71	11	0.0364
4	⁸ 0.00	0.18	⁷ -1.12	0.16	318.85	15995.27	18.67	12	0.0967
5	¹⁰ -0.00	0.18	⁹ -1.20	0.16	-2212.64	736667.42	12.28	13	0.5062
6	¹² 1.24	0.25	¹¹ -0.53	0.17	0.43	0.15	7.23	12	0.8426
7	¹⁴ 1.10	0.25	¹³ -1.04	0.19	0.95	0.22	8.40	12	0.7538
8	¹⁶ 1.18	0.25	¹⁵ -0.75	0.17	0.64	0.17	15.72	13	0.2641
9	¹⁸ 0.29	0.19	¹⁷ -1.33	0.17	4.65	3.09	16.27	14	0.2963
10	²⁰ 0.59	0.18	¹⁹ -0.64	0.15	1.09	0.38	19.39	14	0.1500
11	²² 0.59	0.18	²¹ 0.12	0.14	-0.20	0.25	8.89	13	0.7819
12	²⁴ 0.79	0.20	²³ -0.40	0.15	0.51	0.22	15.19	13	0.2950
13	²⁶ 1.70	0.33	²⁵ 0.20	0.19	-0.12	0.12	6.34	10	0.7864
14	²⁸ 1.15	0.24	²⁷ -0.02	0.16	0.02	0.14	12.67	13	0.4752
15	³⁰ 0.19	0.19	²⁹ -1.24	0.16	6.35	6.02	15.34	13	0.2858
16	³² 0.73	0.19	³¹ 0.17	0.15	-0.23	0.21	15.03	13	0.3048
17	³⁴ 0.33	0.16	³³ -0.35	0.14	1.04	0.65	8.33	14	0.8720
18	³⁶ 1.65	0.32	³⁵ 0.91	0.21	-0.55	0.14	6.16	10	0.8021
19	³⁸ 1.54	0.31	³⁷ 0.54	0.18	-0.35	0.13	6.97	11	0.8024
20	⁴⁰ 0.37	0.17	³⁹ -0.33	0.14	0.89	0.54	8.57	13	0.8051
21	⁴² -1.39	0.56	⁴¹ -3.94	0.65	-2.83	0.80	4.18	4	0.3828
22	⁴⁴ 0.76	0.20	⁴³ -0.40	0.15	0.53	0.22	9.02	12	0.7016
23	⁴⁶ 1.50	0.29	⁴⁵ 0.13	0.18	-0.09	0.12	8.83	11	0.6390
24	⁴⁸ 0.40	0.17	⁴⁷ -0.65	0.15	1.64	0.74	16.43	13	0.2259
25	⁵⁰ 0.05	0.16	⁴⁹ -0.65	0.14	14.21	50.08	8.19	15	0.9161

Likelihood-based Values and Goodness of Fit Statistics [\(Back to TOC\)](#)

Statistics based on the loglikelihood

-2loglikelihood: 6264.25

Akaike Information Criterion (AIC): 6364.25

Bayesian Information Criterion (BIC): 6533.01

Item Information Function Values for Group 1 at 15 Values of θ from -2.8 to 2.8[\(Back to TOC\)](#)

Item	θ :														
	-2.8	-2.4	-2.0	-1.6	-1.2	-0.8	-0.4	-0.0	0.4	0.8	1.2	1.6	2.0	2.4	2.8
1	0.07	0.07	0.07	0.07	0.07	0.06	0.06	0.05	0.04	0.04	0.03	0.03	0.02	0.02	0.02
2	0.01	0.02	0.02	0.02	0.02	0.03	0.03	0.04	0.04	0.05	0.05	0.06	0.06	0.06	0.06
3	0.04	0.07	0.13	0.23	0.37	0.53	0.64	0.63	0.51	0.35	0.22	0.12	0.07	0.04	0.02
4	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
6	0.03	0.04	0.07	0.11	0.16	0.23	0.30	0.36	0.39	0.37	0.31	0.24	0.17	0.11	0.07
7	0.02	0.03	0.04	0.07	0.10	0.13	0.18	0.23	0.28	0.30	0.30	0.27	0.22	0.17	0.12
8	0.02	0.04	0.06	0.09	0.13	0.18	0.24	0.30	0.34	0.34	0.31	0.26	0.19	0.14	0.09
9	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.02	0.02	0.02
10	0.03	0.03	0.04	0.05	0.06	0.06	0.07	0.08	0.08	0.09	0.09	0.08	0.08	0.07	0.07
11	0.05	0.06	0.07	0.07	0.08	0.09	0.09	0.09	0.09	0.08	0.07	0.07	0.06	0.05	0.04
12	0.04	0.05	0.07	0.08	0.10	0.12	0.14	0.15	0.16	0.15	0.14	0.13	0.11	0.09	0.08
13	0.03	0.06	0.11	0.20	0.34	0.52	0.68	0.71	0.60	0.41	0.25	0.14	0.08	0.04	0.02
14	0.05	0.07	0.11	0.15	0.21	0.27	0.31	0.33	0.31	0.27	0.21	0.16	0.11	0.08	0.05
15	0.00	0.00	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
16	0.06	0.08	0.09	0.10	0.12	0.13	0.13	0.13	0.13	0.12	0.10	0.09	0.07	0.06	0.05
17	0.02	0.02	0.02	0.02	0.02	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03
18	0.06	0.12	0.21	0.35	0.52	0.65	0.67	0.56	0.39	0.24	0.14	0.07	0.04	0.02	0.01
19	0.05	0.09	0.16	0.26	0.40	0.53	0.59	0.55	0.43	0.29	0.18	0.11	0.06	0.03	0.02
20	0.02	0.02	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03
21	0.48	0.44	0.35	0.25	0.16	0.10	0.06	0.04	0.02	0.01	0.01	0.00	0.00	0.00	0.00
22	0.04	0.05	0.06	0.08	0.10	0.11	0.13	0.14	0.14	0.14	0.13	0.12	0.11	0.09	0.07
23	0.04	0.07	0.11	0.19	0.30	0.43	0.53	0.56	0.49	0.37	0.25	0.15	0.09	0.05	0.03
24	0.02	0.02	0.02	0.03	0.03	0.03	0.03	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04
25	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Test Information:	2.20	2.47	2.86	3.47	4.33	5.27	5.96	6.06	5.56	4.75	3.93	3.22	2.67	2.25	1.95
Expected s.e.:	0.67	0.64	0.59	0.54	0.48	0.44	0.41	0.41	0.42	0.46	0.50	0.56	0.61	0.67	0.72

Marginal Reliability for Response Pattern Scores: 0.78

ผลการวิเคราะห์ข้อสอบและแบบสอบคู่ขนานโดยอาศัยทฤษฎีการตอบสนองข้อสอบของแบบสอบชุด 5B

2PL Model Item Parameter Estimates for Group 1, logit: $a\theta + c$ or $a(\theta - b)$ [\(Back to TOC\)](#)

Item	a	s.e.	c	s.e.	b	s.e.	X^2	df.	Probability
1	² 0.30	0.24	¹ -1.94	0.21	6.43	5.01	11.95	11	0.3691
2	⁴ -0.51	0.26	³ -1.96	0.22	-3.85	1.83	10.52	10	0.3975
3	⁶ 1.09	0.25	⁵ -1.12	0.19	1.03	0.24	17.17	11	0.1028
4	⁸ 0.21	0.17	⁷ -0.62	0.14	2.90	2.37	10.58	12	0.5661
5	¹⁰ 0.61	0.19	⁹ -0.36	0.15	0.59	0.29	10.50	12	0.5734
6	¹² 1.29	0.28	¹¹ -0.01	0.17	0.01	0.13	14.96	10	0.1331
7	¹⁴ 0.40	0.17	¹³ -0.39	0.14	0.98	0.53	10.02	13	0.6928
8	¹⁶ 1.35	0.29	¹⁵ -0.09	0.17	0.07	0.13	13.61	11	0.2548
9	¹⁸ 0.22	0.19	¹⁷ -1.11	0.16	5.05	4.28	13.16	11	0.2825
10	²⁰ 0.58	0.21	¹⁹ -1.13	0.17	1.94	0.66	9.80	12	0.6349
11	²² 0.67	0.20	²¹ -0.54	0.15	0.81	0.31	16.10	12	0.1863
12	²⁴ 0.42	0.18	²³ -0.66	0.15	1.57	0.72	9.89	13	0.7040
13	²⁶ 1.45	0.31	²⁵ 0.04	0.17	-0.03	0.12	9.01	10	0.5320
14	²⁸ 0.98	0.23	²⁷ -0.02	0.15	0.02	0.16	20.50	12	0.0581
15	³⁰ 0.13	0.22	²⁹ -1.62	0.18	11.97	19.06	10.96	10	0.3626
16	³² 0.23	0.17	³¹ -0.44	0.14	1.92	1.51	10.69	13	0.6379
17	³⁴ 0.11	0.18	³³ -0.87	0.15	7.82	12.42	16.59	13	0.2183
18	³⁶ 0.54	0.19	³⁵ -0.48	0.15	0.89	0.38	6.65	12	0.8799
19	³⁸ 1.73	0.38	³⁷ 0.56	0.19	-0.32	0.12	9.16	9	0.4246
20	⁴⁰ 0.27	0.21	³⁹ -1.57	0.18	5.84	4.58	11.25	12	0.5086
21	⁴² 0.89	0.22	⁴¹ -0.46	0.16	0.51	0.20	13.72	12	0.3207
22	⁴⁴ 0.41	0.21	⁴³ -1.33	0.17	3.24	1.57	4.99	12	0.9586
23	⁴⁶ 0.86	0.21	⁴⁵ 0.05	0.15	-0.06	0.18	21.77	12	0.0400
24	⁴⁸ 0.16	0.18	⁴⁷ -0.98	0.15	6.02	6.70	8.76	13	0.7915
25	⁵⁰ 0.15	0.19	⁴⁹ -1.21	0.16	7.91	9.93	15.93	12	0.1938

Likelihood-based Values and Goodness of Fit Statistics [\(Back to TOC\)](#)

Statistics based on the loglikelihood

-2loglikelihood: 6304.95

Akaike Information Criterion (AIC): 6404.95

Bayesian Information Criterion (BIC): 6573.71

Item Information Function Values for Group 1 at 15 Values of θ from -2.8 to 2.8[\(Back to TOC\)](#)

Item	θ :														
	-2.8	-2.4	-2.0	-1.6	-1.2	-0.8	-0.4	-0.0	0.4	0.8	1.2	1.6	2.0	2.4	2.8
1	0.00	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.02	0.02
2	0.06	0.06	0.05	0.05	0.04	0.04	0.03	0.03	0.02	0.02	0.02	0.01	0.01	0.01	0.01
3	0.02	0.03	0.04	0.06	0.09	0.13	0.17	0.22	0.26	0.29	0.29	0.27	0.23	0.18	0.13
4	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
5	0.04	0.04	0.05	0.06	0.07	0.08	0.09	0.09	0.09	0.09	0.09	0.09	0.08	0.07	0.06
6	0.04	0.07	0.11	0.16	0.24	0.32	0.39	0.42	0.39	0.33	0.24	0.17	0.11	0.07	0.04
7	0.02	0.03	0.03	0.03	0.03	0.03	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.03
8	0.04	0.06	0.10	0.16	0.24	0.33	0.41	0.45	0.43	0.36	0.27	0.18	0.12	0.07	0.04
9	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
10	0.02	0.02	0.03	0.03	0.04	0.05	0.05	0.06	0.07	0.08	0.08	0.08	0.08	0.08	0.08
11	0.03	0.04	0.05	0.06	0.07	0.08	0.10	0.10	0.11	0.11	0.11	0.10	0.10	0.08	0.07
12	0.02	0.02	0.03	0.03	0.03	0.03	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04
13	0.04	0.06	0.11	0.18	0.28	0.39	0.49	0.53	0.48	0.38	0.26	0.17	0.10	0.06	0.03
14	0.05	0.08	0.10	0.14	0.17	0.21	0.23	0.24	0.23	0.21	0.18	0.14	0.11	0.08	0.06
15	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
16	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
17	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
18	0.03	0.04	0.04	0.05	0.05	0.06	0.06	0.07	0.07	0.07	0.07	0.07	0.07	0.06	0.06
19	0.04	0.08	0.15	0.27	0.44	0.63	0.74	0.69	0.52	0.33	0.19	0.10	0.05	0.03	0.01
20	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.02
21	0.04	0.05	0.07	0.09	0.12	0.14	0.17	0.19	0.20	0.20	0.18	0.16	0.13	0.11	0.08
22	0.01	0.01	0.02	0.02	0.02	0.02	0.03	0.03	0.03	0.03	0.04	0.04	0.04	0.04	0.04
23	0.06	0.08	0.10	0.12	0.15	0.17	0.18	0.18	0.18	0.16	0.14	0.12	0.09	0.07	0.05
24	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
25	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01
Test Information:	1.61	1.81	2.12	2.56	3.14	3.78	4.28	4.45	4.24	3.81	3.31	2.85	2.47	2.17	1.94
Expected s.e.:	0.79	0.74	0.69	0.62	0.56	0.51	0.48	0.47	0.49	0.51	0.55	0.59	0.64	0.68	0.72

Marginal Reliability for Response Pattern Scores: 0.72

ผลการวิเคราะห์ข้อสอบและแบบสอบคู่ขนานโดยอาศัยทฤษฎีการตอบสนองข้อสอบของแบบสอบชุด 6A

2PL Model Item Parameter Estimates for Group 1, logit: $a\theta + c$ or $a(\theta - b)$

[\(Back to TOC\)](#)

Item	a	s.e.	c	s.e.	b	s.e.	χ^2	df.	Probability
1	² -0.11	0.18	¹ -1.05	0.15	-9.29	14.55	11.50	14	0.6477
2	⁴ 1.35	0.27	³ 0.16	0.17	-0.12	0.12	8.53	12	0.7435
3	⁶ 0.00	0.19	⁵ -1.32	0.17	295.00	12616.05	16.55	15	0.3481
4	⁸ 0.95	0.22	⁷ -0.60	0.16	0.63	0.19	8.87	13	0.7830
5	¹⁰ 0.04	0.16	⁹ -0.23	0.13	5.22	18.98	10.20	15	0.8076
6	¹² 1.56	0.32	¹¹ 1.15	0.21	-0.74	0.15	11.79	11	0.3813
7	¹⁴ 0.72	0.20	¹³ -0.75	0.16	1.05	0.32	13.55	13	0.4078
8	¹⁶ 0.68	0.19	¹⁵ -0.31	0.14	0.45	0.24	13.42	14	0.4950
9	¹⁸ 0.26	0.17	¹⁷ -0.84	0.15	3.19	2.11	7.12	14	0.9301
10	²⁰ 0.68	0.19	¹⁹ 0.47	0.15	-0.69	0.27	8.55	13	0.8068
11	²² 0.78	0.20	²¹ -0.77	0.16	0.98	0.29	9.72	14	0.7830
12	²⁴ 0.61	0.18	²³ -0.34	0.14	0.56	0.27	9.22	14	0.8174
13	²⁶ 0.99	0.23	²⁵ 1.02	0.17	-1.02	0.25	13.61	13	0.4039
14	²⁸ 0.02	0.21	²⁷ -1.65	0.18	72.12	658.66	8.97	13	0.7761
15	³⁰ 1.08	0.24	²⁹ -0.14	0.16	0.13	0.14	9.01	12	0.7027
16	³² 0.45	0.17	³¹ 0.33	0.14	-0.74	0.40	8.78	14	0.8451
17	³⁴ 0.17	0.16	³³ -0.15	0.13	0.93	1.20	18.35	15	0.2440
18	³⁶ 1.47	0.31	³⁵ 1.21	0.21	-0.82	0.17	6.06	11	0.8697
19	³⁸ 1.35	0.28	³⁷ 0.48	0.17	-0.36	0.13	15.46	13	0.2785
20	⁴⁰ 1.13	0.24	³⁹ 0.47	0.16	-0.42	0.15	10.04	13	0.6917
21	⁴² 0.72	0.23	⁴¹ -1.58	0.19	2.20	0.64	6.79	13	0.9127
22	⁴⁴ 1.33	0.28	⁴³ -0.69	0.17	0.52	0.15	15.28	12	0.2259
23	⁴⁶ 0.23	0.17	⁴⁵ -0.90	0.15	3.86	2.88	13.70	14	0.4738
24	⁴⁸ 0.30	0.17	⁴⁷ -0.71	0.14	2.41	1.41	12.08	15	0.6737
25	⁵⁰ 0.92	0.22	⁴⁹ 0.42	0.15	-0.46	0.18	17.55	13	0.1747

Likelihood-based Values and Goodness of Fit Statistics [\(Back to TOC\)](#)

Statistics based on the loglikelihood

-2loglikelihood: 6711.84

Akaike Information Criterion (AIC): 6811.84

Bayesian Information Criterion (BIC): 6982.20

Item Information Function Values for Group 1 at 15 Values of θ from -2.8 to 2.8[\(Back to TOC\)](#)

Item	θ :														
	-2.8	-2.4	-2.0	-1.6	-1.2	-0.8	-0.4	-0.0	0.4	0.8	1.2	1.6	2.0	2.4	2.8
1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2	0.05	0.08	0.12	0.19	0.28	0.37	0.44	0.45	0.40	0.32	0.22	0.15	0.09	0.06	0.03
3	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
4	0.03	0.05	0.06	0.09	0.11	0.15	0.18	0.21	0.22	0.22	0.21	0.18	0.15	0.12	0.09
5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
6	0.09	0.16	0.26	0.40	0.54	0.61	0.57	0.44	0.30	0.19	0.11	0.06	0.03	0.02	0.01
7	0.03	0.04	0.05	0.06	0.07	0.09	0.10	0.11	0.12	0.13	0.13	0.12	0.11	0.10	0.09
8	0.04	0.05	0.06	0.07	0.09	0.10	0.11	0.11	0.12	0.11	0.11	0.10	0.09	0.08	0.06
9	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.02	0.02	0.02
10	0.07	0.08	0.10	0.11	0.11	0.12	0.12	0.11	0.10	0.09	0.08	0.07	0.06	0.04	0.04
11	0.03	0.04	0.05	0.06	0.08	0.10	0.12	0.13	0.14	0.15	0.15	0.14	0.13	0.11	0.10
12	0.04	0.04	0.05	0.06	0.07	0.08	0.08	0.09	0.09	0.09	0.09	0.08	0.08	0.07	0.06
13	0.12	0.16	0.20	0.23	0.25	0.24	0.22	0.19	0.16	0.12	0.09	0.06	0.04	0.03	0.02
14	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
15	0.05	0.07	0.10	0.14	0.18	0.23	0.27	0.29	0.29	0.26	0.21	0.16	0.12	0.09	0.06
16	0.04	0.04	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.04	0.04	0.04	0.03	0.03
17	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
18	0.11	0.18	0.28	0.40	0.50	0.54	0.49	0.38	0.26	0.17	0.10	0.06	0.03	0.02	0.01
19	0.06	0.10	0.16	0.24	0.34	0.42	0.46	0.43	0.36	0.26	0.18	0.11	0.07	0.04	0.02
20	0.08	0.11	0.16	0.21	0.27	0.31	0.32	0.30	0.26	0.21	0.15	0.11	0.07	0.05	0.03
21	0.01	0.02	0.02	0.03	0.04	0.05	0.06	0.07	0.09	0.10	0.11	0.12	0.13	0.13	0.12
22	0.02	0.03	0.06	0.09	0.15	0.22	0.31	0.39	0.44	0.43	0.36	0.27	0.19	0.12	0.08
23	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
24	0.01	0.01	0.01	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
25	0.08	0.10	0.13	0.16	0.19	0.21	0.21	0.20	0.18	0.15	0.12	0.10	0.07	0.05	0.04
Test Information:	1.98	2.39	2.95	3.63	4.35	4.91	5.15	5.03	4.64	4.10	3.53	3.01	2.57	2.22	1.96
Expected s.e.:	0.71	0.65	0.58	0.52	0.48	0.45	0.44	0.45	0.46	0.49	0.53	0.58	0.62	0.67	0.72

Marginal Reliability for Response Pattern Scores: 0.76

ผลการวิเคราะห์ข้อสอบและแบบสอบคู่ขนานโดยอาศัยทฤษฎีการตอบสนองข้อสอบของแบบสอบชุด 6B

2PL Model Item Parameter Estimates for Group 1, logit: $a\theta + c$ or $a(\theta - b)$ [\(Back to TOC\)](#)

Item	a	s.e.	c	s.e.	b	s.e.	χ^2	df	Probability
1	² -0.21	0.25	¹ -1.80	0.19	-8.46	9.84	15.83	11	0.1473
2	⁴ 1.12	0.28	³ -0.97	0.19	0.87	0.21	11.71	11	0.3872
3	⁶ 0.33	0.19	⁵ -0.80	0.15	2.42	1.37	17.44	12	0.1335
4	⁸ 0.85	0.24	⁷ -1.02	0.18	1.20	0.32	21.93	11	0.0249
5	¹⁰ 0.40	0.19	⁹ -0.66	0.15	1.65	0.80	15.90	12	0.1955
6	¹² 0.49	0.19	¹¹ -0.63	0.15	1.30	0.55	22.35	12	0.0337
7	¹⁴ 0.50	0.19	¹³ -0.31	0.14	0.64	0.36	22.69	12	0.0304
8	¹⁶ 0.36	0.22	¹⁵ -1.50	0.18	4.20	2.52	15.17	12	0.2314
9	¹⁸ 0.25	0.20	¹⁷ -1.13	0.16	4.49	3.53	22.75	13	0.0446
10	²⁰ 1.03	0.26	¹⁹ -0.45	0.16	0.44	0.17	10.23	12	0.5972
11	²² 0.68	0.22	²¹ -0.66	0.15	0.97	0.34	30.40	11	0.0014
12	²⁴ 0.45	0.20	²³ -0.65	0.15	1.44	0.66	9.76	13	0.7143
13	²⁶ 0.19	0.18	²⁵ 0.51	0.14	-2.72	2.72	15.40	12	0.2197
14	²⁸ 0.91	0.24	²⁷ -0.16	0.15	0.17	0.17	21.48	12	0.0437
15	³⁰ 0.47	0.20	²⁹ -0.47	0.14	0.99	0.47	15.95	12	0.1929
16	³² 0.21	0.18	³¹ -0.32	0.14	1.49	1.35	18.08	12	0.1131
17	³⁴ 0.29	0.20	³³ -1.02	0.15	3.46	2.27	23.56	13	0.0353
18	³⁶ -0.11	0.19	³⁵ -0.85	0.15	-7.38	12.14	27.65	13	0.0101
19	³⁸ 0.74	0.22	³⁷ -0.48	0.15	0.65	0.25	9.95	12	0.6211
20	⁴⁰ 0.57	0.21	³⁹ -0.69	0.15	1.21	0.46	22.52	13	0.0477
21	⁴² 0.73	0.22	⁴¹ -0.56	0.15	0.77	0.28	26.95	13	0.0126
22	⁴⁴ 0.89	0.25	⁴³ -1.19	0.19	1.33	0.34	22.60	11	0.0201
23	⁴⁶ 1.01	0.26	⁴⁵ 0.19	0.15	-0.19	0.16	12.01	12	0.4463
24	⁴⁸ 0.09	0.19	⁴⁷ -0.96	0.15	10.75	22.71	13.66	13	0.3999
25	⁵⁰ -0.08	0.20	⁴⁹ -1.22	0.16	-14.48	34.53	22.58	12	0.0314

Likelihood-based Values and Goodness of Fit Statistics [\(Back to TOC\)](#)

Statistics based on the loglikelihood

-2loglikelihood: 6771.83

Akaike Information Criterion (AIC): 6871.83

Bayesian Information Criterion (BIC): 7042.19

Item Information Function Values for Group 1 at 15 Values of θ from -2.8 to 2.8[\(Back to TOC\)](#)

Item	θ :														
	-2.8	-2.4	-2.0	-1.6	-1.2	-0.8	-0.4	-0.0	0.4	0.8	1.2	1.6	2.0	2.4	2.8
1	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00
2	0.02	0.03	0.05	0.07	0.10	0.14	0.20	0.25	0.29	0.31	0.30	0.26	0.21	0.16	0.12
3	0.01	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.03	0.03	0.03	0.03	0.03	0.03
4	0.02	0.03	0.04	0.06	0.07	0.09	0.12	0.14	0.16	0.18	0.18	0.18	0.16	0.14	0.12
5	0.02	0.02	0.02	0.03	0.03	0.03	0.03	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04
6	0.02	0.03	0.03	0.04	0.04	0.05	0.05	0.05	0.06	0.06	0.06	0.06	0.06	0.06	0.05
7	0.03	0.04	0.04	0.05	0.05	0.05	0.06	0.06	0.06	0.06	0.06	0.06	0.05	0.05	0.05
8	0.01	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.02	0.02	0.02	0.03	0.03	0.03	0.03
9	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.02
10	0.04	0.05	0.07	0.10	0.14	0.18	0.22	0.25	0.27	0.26	0.23	0.19	0.15	0.11	0.08
11	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.10	0.11	0.12	0.12	0.11	0.10	0.09	0.08
12	0.02	0.03	0.03	0.03	0.04	0.04	0.04	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
13	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
14	0.05	0.07	0.09	0.11	0.14	0.17	0.19	0.21	0.20	0.19	0.17	0.14	0.11	0.09	0.06
15	0.03	0.03	0.04	0.04	0.04	0.05	0.05	0.05	0.05	0.06	0.06	0.05	0.05	0.05	0.05
16	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
17	0.01	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
18	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
19	0.04	0.05	0.06	0.07	0.09	0.10	0.12	0.13	0.14	0.14	0.13	0.12	0.11	0.09	0.08
20	0.03	0.03	0.04	0.05	0.05	0.06	0.07	0.07	0.08	0.08	0.08	0.08	0.08	0.07	0.07
21	0.03	0.04	0.06	0.07	0.08	0.10	0.11	0.12	0.13	0.13	0.13	0.12	0.11	0.10	0.08
22	0.02	0.03	0.04	0.05	0.07	0.09	0.12	0.14	0.17	0.19	0.20	0.20	0.18	0.16	0.13
23	0.06	0.09	0.12	0.16	0.20	0.23	0.25	0.25	0.23	0.20	0.16	0.12	0.09	0.07	0.05
24	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
25	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Test Information:	1.54	1.68	1.85	2.07	2.31	2.57	2.82	3.02	3.15	3.17	3.08	2.90	2.68	2.44	2.21
Expected s.e.:	0.81	0.77	0.73	0.70	0.66	0.62	0.60	0.58	0.56	0.56	0.57	0.59	0.61	0.64	0.67

Marginal Reliability for Response Pattern Scores: 0.64

ประวัติผู้เขียนวิทยานิพนธ์

ร้อยเอก อนุพงษ์ กันธิวงศ์ เกิดเมื่อวันที่ 22 พฤศจิกายน พ.ศ. 2528 ที่จังหวัดเชียงใหม่ สำเร็จการศึกษาแพทยศาสตรบัณฑิต จากวิทยาลัยแพทยศาสตร์พระมงกุฎเกล้า เมื่อปีการศึกษา 2554 ต่อมารับราชการในตำแหน่ง แพทย์ใช้ทุน สังกัดโรงพยาบาลค่ายจักรพงษ์ จากนั้น ดำรงตำแหน่งอาจารย์ภาควิชาเภสัชวิทยา วิทยาลัยแพทยศาสตร์พระมงกุฎเกล้า และปฏิบัติหน้าที่ในหน่วยแพทยศาสตรศึกษา วิทยาลัยแพทยศาสตร์พระมงกุฎเกล้า และในปีการศึกษา 2558 ได้เข้าศึกษาต่อในระดับปริญญาโท สาขาการวัดและประเมินผลการศึกษา ภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ได้รับทุน “90 ปี จุฬาลงกรณ์มหาวิทยาลัย” จากกองทุนรัชดาภิเษกสมโภช จุฬาลงกรณ์มหาวิทยาลัยรุ่นที่ 37

