บทที่ 4

สรุปผลการวิจัย

การวิจัยนี้มีจุดมุ่งหมายเพื่อวิเคราะห์คุณภาพของแบบสอบวิชาการพยาบาลสูติศาสตร์ที่ใช้สอบนักศึกษาผดุงครรภ์ คณะพยาบาลศาสตร์ มหาวิทยาลัยมหิดล ซึ่งคณาจารย์ผู้สอนเป็นผู้สร้างขึ้น ข้อสอบมีจำนวน 200 ข้อ เป็นชนิดเลือกตอบ 4 ตัวเลือก ใช้เวลาในการสอบ 3 ชั่วโมง ใช้กระดาษคำตอบของนักศึกษาผดุงครรภ์รุ่นที่ 2 ปีการศึกษา 2517 จำนวน 95 คน

ข้อค้นพบ

1. ลักษณะการกระจายของคะแนน แบบสอบวิชาการพยาบาลสูติศาสตร์ การแจกแจงความถี่ของคะแนน มีความเบ้เป็น -.13 คือเบ้ซ้าย นั่นคือนักศึกษาผดุงครรภ์ส่วนใหญ่ได้คะแนนสูง และค่าพิสัยของคะแนนเป็น 55 ค่ามัชฌิมเลขคณิตเป็น 139.23 ส่วนเบี่ยงเบนมาตรฐานเป็น 13.755

2. การวิเคราะห์ข้อ เพื่อหาระดับความยากและอำนาจจำแนกของข้อสอบแต่ละข้อ

2.1 ผลการวิเคราะห์ข้อโดยใช้กระดาษคำตอบทั้งหมดคำนวณหาระดับความยากโดยหาสัดส่วนของผู้ตอบถูกทั้งหมดกับจำนวนคนทั้งหมด และคำนวณหาอำนาจจำแนกโดยใช้สูตรสหสัมพันธ์แบบ พอยท์ ไบซีเรียล ได้ค่าอำนาจจำแนกระหว่าง -.19 ถึง .55 และระดับความยากอยู่ระหว่างร้อยละ 10 ถึงร้อยละ 100 มีข้อสอบที่อยู่ในเกณฑ์สมควรเก็บไว้ใช้ต่อไป 55 ข้อ คิดเป็นร้อยละ 27.5 ของจำนวนข้อสอบทั้งหมด สมควรตัดทิ้งไป 22 ข้อ ข้อสอบที่เหลืออีก 123 ข้อ สมควรแก้ไขปรับปรุง

2.2 ผลการวิเคราะห์ข้อ โดยใช้เทคนิคแบ่งกลุ่มสูงและกลุ่มต่ำร้อยละ 50 แล้วคำนวณหาระดับความยากและอำนาจจำแนกโดยใช้สูตรของ จอห์นสัน และไฟน์เลย์พบว่ามีค่าอำนาจจำแนกระหว่าง -.15 ถึง .49 ระดับความยากระหว่างร้อยละ 10 ถึง ร้อยละ 100 มีข้อสอบที่อยู่ในเกณฑ์สมควรเก็บไว้ใช้ต่อไป 44 ข้อ คิดเป็นร้อยละ

22 ข้อจำนวนข้อสอบทั้งหมด สมควรตัดทิ้งไป 17 ข้อ ข้อสอบที่เหลืออีก 139 ข้อ สมควรแก้ไขปรับปรุงใหม่

2.3 ผลการวิเคราะห์ข้อ โดยใช้เทคนิคแบ่งกลุ่มสูงและกลุ่มต่ำร้อยละ 27 หาระดับความยากและอำนาจจำแนกโดยใช้ตารางวิเคราะห์ข้อสอบของ จุง เต ฟาน มีค่าอำนาจจำแนกระหว่าง -.30 ถึง .76 ระดับความยากระหว่างร้อยละ 7 ถึงร้อยละ 100 มีข้อสอบที่อยู่ในเกณฑ์สมควรเก็บไว้ใช้ต่อไป 64 ข้อ คิดเป็นร้อยละ 32 ของจำนวนข้อสอบทั้งหมด สมควรตัดทิ้งไป 19 ข้อ ข้อสอบที่เหลืออีก 117 ข้อ สมควรแก้ไขปรับปรุงใหม่

2.4 ผลการวิเคราะห์ตัวเลือก โดยใช้เทคนิคแบ่งกลุ่มสูงและกลุ่มต่ำร้อยละ 50 แล้วคำนวณหาระดับความยากและอำนาจจำแนก โดยใช้สูตรของ จอห์นสัน และไฟน์เลย์ พบว่าตัวเลือกส่วนใหญ่ยังไม่ดี สมควรแก้ไขปรับปรุง หรือตัดทิ้งไป

3. ผลการเปรียบเทียบการวิเคราะห์ข้อสอบ เมื่อใช้กระดาษคำตอบทั้งหมดและหาอำนาจจำแนกด้วยสูตรสหสัมพันธ์แบบพอยท์ ไบซีเรียล กับเมื่อใช้เทคนิคแบ่งกลุ่มสูงและกลุ่มต่ำร้อยละ 50 หาระดับความยากและอำนาจจำแนกด้วยสูตรของจอห์นสันและไฟน์เลย์ ผลปรากฏว่า ค่าระดับความยากของแต่ละข้อใกล้เคียงกันหรือเท่ากัน อำนาจจำแนกที่ได้จากการคำนวณโดยสูตรสหสัมพันธ์แบบพอยท์ ไบซีเรียลมีช่วงกว้างมากกว่าที่ได้จากการคำนวณโดยสูตรของจอห์นสันและไฟน์เลย์และจำนวนข้อสอบที่ดีสมควรเก็บไว้ใช้มีมากกว่า ทั้ง 2 วิธีนี้มีข้อดีซ้ำกัน 43 ข้อ ผลการเปรียบเทียบการวิเคราะห์ข้อสอบด้วยเทคนิคแบ่งกลุ่มสูงและกลุ่มต่ำร้อยละ 50 กับร้อยละ 27 ผลปรากฏว่า เทคนิคแบ่งกลุ่มสูงและกลุ่มต่ำร้อยละ 27 มีค่าอำนาจจำแนกช่วงกว้างกว่า ระดับความยากช่วงกว้างกว่า และมีข้อสอบที่ดีสมควรเก็บไว้ใช้มากกว่า มีข้อสอบที่ดีซ้ำกับเทคนิคแบ่งกลุ่มสูงกลุ่มต่ำร้อยละ 50 จำนวน 38 ข้อ มีข้อสอบที่ดีซ้ำกับแบบที่คำนวณด้วยสูตรสหสัมพันธ์แบบฟอยท์ ไบซีเรียล 50 ข้อ

4. ผลการคำนวณหาค่าสัมประสิทธิ์แห่งความเที่ยง ค่าสัมประสิทธิ์แห่งความเที่ยงของแบบสอบชนิดความคงที่ภายใน (Internal consistency) ใช้สูตรของคูเดอร์ ริชาร์ดสัน ที่ 20 (Kuder-Richardson 20) ได้ค่าสัมประสิทธิ์แห่งความเที่ยงของแบบสอบเท่ากับ

.8390 นับว่ามีค่าอยู่ในระดับสูง[1]

5. ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนสอบวิชาการพยาบาลสูติศาสตร์กับคะแนนเฉลี่ยภาคปฏิบัติตลอดหลักสูตร ใช้สูตรของเพียร์สัน (Pearson Product Moment) เป็น .4357 มีนัยสำคัญที่ระดับ .01

6. ความตรงตามเนื้อหา (Content Validity) ของแบบสอบเทียบกับตารางวิเคราะห์หลักสูตรวิชาการพยาบาลสูติศาสตร์ที่หัวหน้าภาควิชาการพยาบาลสูติศาสตร์ให้น้ำหนักผลปรากฏว่า จำนวนข้อสอบรวมในหน่วยที่ 2, 4 และ 5 เท่ากัน ส่วนในหน่วยที่ 1 มีข้อสอบเกิน 2 ข้อ ในหน่วยที่ 3 ขาดข้อสอบ 2 ข้อ ส่วนรายข้อตามเนื้อเรื่องและวัตถุประสงค์ในหน่วยต่าง ๆ นั้น ส่วนใหญ่ใกล้เคียงกัน นอกจากการวัดการสังเคราะห์เท่านั้นที่ไม่มีข้อสอบในแบบสอบ

อภิปรายผลการวิจัย

ผลการวิจัยเรื่องอำนาจจำแนกของข้อสอบเป็นรายข้อ พบว่าการวิเคราะห์ข้อด้วยการหาระดับความยากโดยหาสัดส่วนของผู้ตอบถูกทั้งหมดกับจำนวนคนทั้งหมด หาอำนาจจำแนกโดยใช้สูตรสหสัมพันธ์แบบพอยท์ ไบซีเรียล มีข้อสอบที่ดีตามเกณฑ์ 55 ข้อ คิดเป็นร้อยละ 27.5 ของจำนวนข้อสอบ มีข้อสอบที่สมควรตัดทิ้งไป 22 ข้อ ส่วนอีก 123 ข้อ สมควรแก้ไขปรับปรุง เมื่อคำนวณหาอำนาจจำแนกและระดับความยากด้วยสูตรของไฟน์เลย์ และจอห์นสัน มีข้อสอบที่ดีตามเกณฑ์ 44 ข้อ คิดเป็นร้อยละ 22 ของจำนวนข้อสอบ มีข้อสอบที่สมควรตัดทิ้งไป 17 ข้อ ส่วนอีก 139 ข้อ สมควรแก้ไขปรับปรุง ส่วนการวิเคราะห์ด้วยเทคนิคแบ่งกลุ่มสูงกลุ่มต่ำร้อยละ 27 มีข้อสอบที่ดีตามเกณฑ์ 64 ข้อ คิดเป็นร้อยละ 32 ของจำนวนข้อสอบทั้งหมด สมควรตัดทิ้งไป 19 ข้อ ส่วนอีก 117 ข้อ สมควรแก้ไขปรับปรุง จากการวิเคราะห์ข้อด้วยวิธีต่างกันทั้ง 3 นี้ จะเห็นว่าแบบสอบฉบับนี้มีข้อสอบที่ดีตามเกณฑ์เป็นจำนวนน้อย สมควรแก้ไขปรับปรุง สำหรับข้อที่มีอำนาจจำแนกติดลบ อาจเป็นเพราะว่ามีการเดาหรือคำถามไม่

---

[1] Loc.cit.

ชัดเจนสมควรตัดทิ้ง โดยต้องพิจารณาเนื้อหาในแบบสอบว่าขาดหายไปหรือไม่

ผลการเปรียบเทียบการวิเคราะห์ข้อสอบด้วยวิธีที่แตกต่างกันทั้ง 3 วิธี ค่าระดับความยากของทั้ง 3 วิธีใกล้เคียงกัน ค่าอำนาจจำแนกที่ได้จากเทคนิคแบ่งกลุ่มสูงกลุ่มต่ำร้อยละ 27 มีช่วงกว้างที่สุด จากสูตรสหสัมพันธ์แบบพอยท์ ไบซีเรียล และจากการคำนวณด้วยสูตรของจอห์นสัน มีช่วงกว้างน้อยลงตามลำดับ ส่วนจำนวนข้อสอบที่ดีตามเกณฑ์ จากเทคนิคแบ่งกลุ่มสูงกลุ่มต่ำร้อยละ 27 ได้ข้อสอบที่ดีมากกว่าทั้ง 2 วิธี ทั้งนี้ อาจเพราะจำนวนตัวอย่างประชากรที่ใช้มีน้อย มีเพียง 95 คน ถึงแม้ว่าคะแนนจะมีการกระจายเป็นปกติก็ตาม อนาสตาซี[2] กล่าวว่า การแบ่งกลุ่มสูงและกลุ่มต่ำร้อยละ 27 ควรใช้เมื่อตัวอย่างประชากรมีเป็นจำนวนมาก ซึ่งอย่างน้อยควรมีจำนวน 370 คน เมื่อแบ่งกลุ่มสูงและกลุ่มต่ำแล้วจะได้กลุ่มละ 100 คน จึงจะสะดวกในการคำนวณ

ผลการวิเคราะห์ตัวเลือก พบว่าตัวเลือกส่วนใหญ่ไม่ดี ไม่สามารถลวงผู้ตอบได้ ควรพิจารณาแก้ไขปรับปรุง หรือควรตัดทิ้งไป

ความเที่ยงของแบบสอบชนิดความสอดคล้องภายใน มีค่าอยู่ในระดับสูง ทั้งนี้เพราะแบบสอบชุดนี้มีข้อสอบจำนวนมาก ดังที่อนาสตาซี[3] กล่าวไว้ว่า แบบสอบยิ่งมากข้อ ความเที่ยงก็มากขึ้น เพราะการเพิ่มความยาวของแบบสอบจะช่วยเพิ่มความสม่ำเสมอในการสุ่มเนื้อเรื่อง (Content Sampling) และ บราวน์[4] (Brown) กล่าวว่า การที่แบบสอบมีความยาวมากจะเพิ่มช่วงกว้างของคะแนน ถ้าคะแนนกระจายมาก ค่าความเที่ยงจะสูงขึ้น

ความตรงตามเนื้อหาในแบบสอบส่วนใหญ่ตรงกับตารางวิเคราะห์หลักสูตร แต่เนื่องจากมีข้อสอบบางข้อที่ไม่ดี คือมีอำนาจจำแนกติดลบ สมควรตัดทิ้งไป ได้แก่ข้อที่ 9, 12, 32, 35, 40, 43, 47, 53, 57, 76, 80, 85, 88, 129, 133, 135, 146, 152, 169, 175, 186 และ

---

[2] Anastasi, op.cit., p. 169.

[3] Ibid, p. 121.

[4] Frederick G. Brown, Principal of Educational and Psychological Testing (Illinois : The Dryden Press Inc., 1970) pp.60 -69.

200 รวม 22 ข้อ เมื่อตัดข้อเหล่านี้ทิ้งไปแล้วมาพิจารณาเนื้อหาในตารางวิเคราะห์หลักสูตร ส่วนใหญ่เนื้อหาจะขาดหายไป จึงสมควรสร้างข้อสอบขึ้นทดแทน ส่วนข้อที่ 24 ซึ่งมีอำนาจจำแนกต่ำและเป็นข้อสอบที่ค่อนข้างยาก ข้อนี้อาจจะปรับปรุงแก้ไขใหม่ เพื่อไม่ให้เนื้อหาขาดหายไป หรืออาจจะตัดทิ้งไป แล้วสร้างข้อสอบทดแทนก็ได้

หอสมุดกลาง สถาบันวิทยบริการ จุฬาลงกรณ์มหาวิทยาลัย

ข้อเสนอแนะ

1. คณาจารย์ผู้สอนวิชาการพยาบาลสูติศาสตร์ ควรอาศัยผลจากการวิจัยนี้ปรับปรุงแบบสอบ ข้อใดที่ดีสมควรเก็บไว้ใช้ต่อไป ส่วนข้อใดที่ไม่ดีควรตัดทิ้งไป หรือปรับปรุงแก้ไขต่อไป ข้อใดที่ตัดทิ้งไปควรพิจารณาว่าทำให้เนื้อหาขาดหายไปหรือไม่ ถ้าขาดหายไปควรสร้างข้อสอบขึ้นทดแทน ตัวเลือกส่วนใหญ่ไม่ดี สมควรแก้ไขปรับปรุงตามการวิเคราะห์ในตารางที่ 5 ทั้งนี้เพื่อให้ได้แบบสอบที่มีคุณภาพดี สามารถวัดความสามารถได้อย่างแท้จริง

2. เมื่อคณาจารย์ผู้สอนวิชาการพยาบาลสูติศาสตร์แก้ไขปรับปรุงเนื้อหาของแบบสอบนี้ตามผลการวิจัยแล้ว ก็ควรจะได้มีการวิเคราะห์ข้อและตัวเลือกอีก แล้วปรับปรุงแก้ไขข้อสอบต่อไปจนกระทั่งได้ข้อสอบที่ดี มีระดับความยากและอำนาจจำแนกอยู่ในเกณฑ์

3. หลังจากได้แบบสอบที่ดีแล้ว ควรมีการสร้างแบบสอบคู่ขนานกันไปกับแบบสอบชุดนี้เพื่อใช้สับเปลี่ยนกัน และควรทำการวิเคราะห์คุณภาพของแบบสอบที่สร้างและปรับปรุงใหม่ด้วย

4. ในการสร้างแบบสอบควรสร้างข้อสอบให้มีเนื้อหาตรงกับตารางวิเคราะห์หลักสูตรของวิชานั้น ๆ เพื่อให้แบบสอบมีความตรงตามเนื้อหา

5. ควรมีการแก้ไขปรับปรุงการให้คะแนนภาคปฏิบัติ ให้ใช้เกณฑ์ในการให้คะแนนอันเดียวกัน เพื่อให้สามารถบอกความสามารถในการปฏิบัติงานอย่างแท้จริง

6. ผลจากการวิเคราะห์ข้อสอบโดยใช้เทคนิคที่แตกต่างกัน ได้ข้อสอบที่ดีแตกต่างกันบ้าง จึงควรใช้วิธีใดวิธีหนึ่งที่สะดวกง่าย เพื่อใช้แก้ไขปรับปรุงแบบสอบให้มีประสิทธิภาพดี