

การพัฒนาการออกข้อสอบอัตนัยของกองทัพไทย

Development of the Subjective Test of the Royal Thai Armed Forces

อนรรักษ์ โชติติกุล ทewa กาญจนชม พงศ์เทพ จันทน์แสนะ

Anuruk Chotedelok Tewa Kanjanachom Pongthep Janseana

กองการศึกษา โรงเรียนนายเรืออากาศนวมินทกษัตริยาธิราช

chotedelok@gmail.com, tawa@rtaf.mi.th, pongthep@rtaf.mi.th

บทคัดย่อ

การออกข้อสอบอัตนัยของกองทัพไทย พบว่ามีกระบวนการที่ยังไม่ครบถ้วนสมบูรณ์ บุคลากรทางการศึกษาขาดความรู้ทางสถิติ และไม่ทราบกระบวนการออกข้อสอบอัตนัยทั้งหมด บทความนี้มีวัตถุประสงค์เพื่อสร้างกระบวนการพัฒนาการออกข้อสอบแบบอัตนัยของกองทัพไทย มีประเด็นสำคัญ คือ (1) การกำหนดความมุ่งหมายของการวัดผล (2) การสร้างและพัฒนาเครื่องมือ/เกณฑ์ (3) การดำเนินการวัด (4) การวิเคราะห์และประเมินผล โดยทำการวิเคราะห์ข้อสอบทั้งในภาพรวมและวิเคราะห์ข้อสอบรายข้อ เพื่อประเมินคุณภาพข้อสอบ และ (5) การบันทึกผล กระบวนการพัฒนาการออกข้อสอบแบบอัตนัยจะทำให้บุคลากรทางการศึกษาของกองทัพไทย สามารถวิเคราะห์ และประเมินคุณภาพข้อสอบอัตนัย ซึ่งเป็นเครื่องมือที่ใช้วัดผลสัมฤทธิ์ทางการศึกษาของผู้เรียน สามารถใช้เป็นแนวทางในการพัฒนาการออกข้อสอบในครั้งต่อไป ให้มีคุณภาพดียิ่งขึ้น

คำสำคัญ : การวิเคราะห์ข้อสอบอัตนัย ข้อสอบอัตนัย ค่าความยาก ค่าอำนาจจำแนก

Abstract

It has been found that many processes of making the subjective tests of the Royal Thai Armed Forces are incomplete. Many educational personnel lack statistical knowledge and do not know the whole subjective tests processes. Hence the objectives of this study is to create processes of developing the subjective tests of the Royal Thai Armed Forces which consist of (1) determining the purpose of measurement, (2) creating and developing tools and criteria, (3) measuring (4) analysis and evaluating (analyze the items in the overview and analyze each item for evaluating the quality of the subjective tests), and (5) recording. The created processes of developing the subjective tests are expected to encourage the educational personnel to be able to analyze and evaluate the quality of the subjective test which can be used as an effective tool in measuring student achievement and practical guideline in developing further subjective tests.

Keywords : Analysis of subjective test, Subjective tests, Difficulty Index, Power of Discrimination

1. กล่าวนำ

การจัดการเรียนการสอนเพื่อให้เป็นไปตามหลักสูตรสถานศึกษาของกองทัพไทย นอกจากการจัดกิจกรรมการเรียนการสอนแล้ว ยังต้องทำการวัดและประเมินผลเพื่อจัดลำดับตามความรู้ความสามารถของผู้เรียน การวัดผลทางการศึกษาที่ดำเนินการในปัจจุบันมีหลายรูปแบบ ได้แก่ การใช้แบบทดสอบ การทำรายงาน การนำเสนอผลงานหน้าชั้นเรียน การสอบสัมภาษณ์ การสอบปฏิบัติ การทำโครงการ เป็นต้น ในบทความนี้นำเสนอเฉพาะในส่วนการวัดผลโดยใช้ข้อสอบอัตนัย (Subjective Test) เท่านั้น ซึ่งเป็นข้อสอบที่ต้องเขียนตอบตามความรู้และความเข้าใจของผู้สอบ สถานศึกษาของกองทัพไทยหลายแห่ง อาทิ วิทยาลัยป้องกันราชอาณาจักร วิทยาลัยเสนาธิการทหาร วิทยาลัยการทัพ โรงเรียนเสนาธิการทหาร วิทยาลัยพยาบาล โรงเรียนเหล่าทัพ ฯลฯ จะมีข้อสอบอัตนัยเป็นส่วนหนึ่งของการวัดผลสัมฤทธิ์ทางการศึกษาทั้งสิ้น การจัดทำข้อสอบอัตนัยแบบครบกระบวนการ สถาบันการศึกษาของกองทัพไทยส่วนใหญ่ยังไม่มี การดำเนินการที่สมบูรณ์ การจัดทำข้อสอบอัตนัยบางสถานศึกษาผู้สอนเป็นผู้จัดทำเพียงคนเดียว ขาดการตรวจสอบคุณภาพ บางสถานศึกษามีกระบวนการตรวจสอบคุณภาพของข้อสอบก่อนนำไปทดสอบกับผู้เรียน โดยผ่านการตรวจสอบจากคณะกรรมการผู้ทรงคุณวุฒิ เพื่อตรวจสอบด้านความตรงตามเนื้อหาวิชา รูปแบบการพิมพ์ การใช้ภาษา ความเหมาะสมของจำนวนข้อสอบ เวลา และการให้น้ำหนักคะแนน เป็นต้น ซึ่งเป็นการตรวจสอบคุณภาพของข้อสอบก่อนนำไปวัดผลกับผู้เรียนเท่านั้น ยังดำเนินการไม่ครบถ้วนตามกระบวนการพัฒนาคุณภาพของข้อสอบ โดยเฉพาะการวิเคราะห์ข้อสอบสอบหลังจากที่ผู้เรียนได้ทำการสอบ ทั้งนี้ เนื่องจากผู้สอนส่วนใหญ่ไม่ทราบกระบวนการที่ครบถ้วน และมีกระบวนการวิเคราะห์ข้อมูลทางสถิติ เช่น ค่าเฉลี่ยเลขคณิต ค่าเบี่ยงเบนมาตรฐาน ค่ามัธยฐาน ค่าความเที่ยง ค่าความแปร ค่าความยากง่าย ค่าอำนาจจำแนก เป็นต้น จึงทำให้ผู้สอนละเลยการปฏิบัติ อันจะนำไปสู่การตอบคำถามของผู้เรียนเกี่ยวกับคุณภาพของข้อสอบที่นำมาใช้วัดความรู้ความสามารถว่ามีคุณภาพมากน้อยเพียงใด

จากปัญหาดังกล่าวผู้เขียน เห็นว่าการสร้างกระบวนการพัฒนาการออกข้อสอบอัตนัยน่าจะเป็นประโยชน์สำหรับสถาบันการศึกษาของกองทัพไทย ที่จะสามารถตอบคำถามผู้เรียนหรือสาธารณะเกี่ยวกับคุณภาพของข้อสอบได้อย่างมีหลักการ

2. กระบวนการพัฒนาข้อสอบสอบอัตนัย

การออกข้อสอบอัตนัยมีกระบวนการ 5 ประเด็นที่สำคัญได้แก่ (1) การกำหนดความมุ่งหมายของการวัดผล (2) การสร้างและพัฒนาเครื่องมือ/เกณฑ์ (3) การดำเนินการวัดผล (4) การวิเคราะห์และประเมินผล และ (5) การบันทึกผล ซึ่งเป็นไปตามวงจรคุณภาพการปฏิบัติงาน PDCA (Plan , Do , Check and Act) กล่าวคือ P เป็นการวางแผนการออกข้อสอบ โดยการกำหนดความมุ่งหมายของการวัดผล D เป็นการดำเนินการสร้างและพัฒนาเครื่องมือและนำไปวัดผลกับผู้เรียน C เป็นการนำผลการสอบที่ได้มาทำการวิเคราะห์และประเมินผล และ A เป็นการนำผลการประเมินที่ได้มาสรุป ประชุม บันทึกผล เพื่อนำผลการบันทึกไปใช้พัฒนาการออกข้อสอบอัตนัยในครั้งต่อไป

2.1 การกำหนดความมุ่งหมายของการวัดผล

ข้อสอบอัตนัย มีจุดประสงค์เพื่อให้ผู้สอบเรียบเรียงความคิด ความรู้ ใช้ภาษาผูกประโยคให้เป็นข้อความที่ชัดเจน เขียนเป็นคำตอบให้สามารถตอบคำถามได้ตรงประเด็นตามความต้องการของคำถาม ข้อสอบแบบอัตนัย แบ่งได้ 2 แบบ คือ (1) แบบจำกัดคำตอบ (Restricted Response) เป็นแบบที่กำหนดขอบเขต รูปแบบ และเนื้อหาในการตอบคำถาม เช่น “จงอธิบายสาเหตุของการความหวาดระแวงที่จะเกิดสงครามในคาบสมุทรเกาหลี มา 3 ประการ” ข้อดี คือ ตรวจง่ายและลดปัญหาด้านความเที่ยงของการให้คะแนน ข้อเสีย คือ มีข้อจำกัดในเรื่องความคิดขั้นสูง (2) แบบไม่จำกัดคำตอบ (Extended Response) เป็นแบบที่ไม่กำหนดขอบเขต รูปแบบ และเนื้อหาในการตอบคำถาม เช่น “จงอภิปรายเกี่ยวกับการสอนวิชาคณิตศาสตร์แบบฝึกทำโจทย์กับการสอนแบบเกม โดยให้บอกถึงหลักสำคัญที่ใช้ในการสอนแต่ละวิธี และข้อดี-ข้อเสียของการสอนทั้งสองแบบ” ข้อดี คือ ผู้สอบมีอิสระในการตอบ สามารถวัดการคิดขั้นสูงได้ ข้อเสีย คือ มีปัญหาเรื่องเที่ยงของการตรวจให้คะแนน ซึ่งผู้สอนต้องมีความรอบรู้ในหลักการ [1]

Bloom กล่าวถึงจุดประสงค์ของการเรียนประกอบด้วย 6 ทักษะ ได้แก่ (1) ทักษะรู้จำ (Knowledge) (2) ทักษะความเข้าใจ (Comprehension) (3) ทักษะการนำไปใช้หรือการประยุกต์ (Application) (4) ทักษะการวิเคราะห์ (Analysis) (5) ทักษะการสังเคราะห์ (Synthesis) และ (6) ทักษะการประเมินค่า (Evaluation) [2] ข้อสอบแบบอัตนัยที่ใช้วัดผลสัมฤทธิ์ในปัจจุบัน ใช้วัดทักษะตั้งแต่ทักษะที่ 2 คือ ทักษะความเข้าใจขึ้นไปจนถึงทักษะที่ 6 ทักษะการประเมินค่า โดยมีลักษณะของข้อสอบตามทักษะ ดังนี้

ทักษะความเข้าใจ มีลักษณะของข้อสอบ คือ (1) การแปลความ เช่น การแปลสัญลักษณ์ แปลลักษณะอาการ แปลเนื้อความ แปลสุภาษิต คำพังเพย (2) การตีความ โดยยึดเนื้อความที่เป็นหลัก และ (3) การขยายความ เพื่อทำนายอนาคตโดยอาศัยข้อมูลและแนวโน้มต่าง ๆ

ทักษะการนำไปใช้หรือการประยุกต์ใช้ มีลักษณะของข้อสอบ คือ นำหลักการ หรือทฤษฎีไปใช้เปรียบเทียบกับสถานการณ์ตามบ่งการซึ่งเป็นสถานการณ์ใหม่ เท่านั้น

ทักษะการวิเคราะห์ มีลักษณะของข้อสอบ คือ (1) การเปรียบเทียบ เพื่อหาสิ่งที่ดีที่สุด (2) การแสดงความสัมพันธ์เชิงเหตุและผล และ (3) การใช้หลักการในการตอบคำถาม

ทักษะการสังเคราะห์ มีลักษณะของข้อสอบ คือ (1) การสังเคราะห์ข้อความ โดยให้เขียนเรียงความ เขียนโคลง เล่าประสบการณ์ของตนเอง วิจารณ์/อภิปรายเหตุการณ์ (2) การสังเคราะห์แผนงาน โดยให้ออกแบบโครงการ/โครงการ (3) ออกแบบระบบฐานข้อมูล วางแผนการทดลอง และ (4) การสังเคราะห์ความสัมพันธ์ โดยให้สร้างสมมติฐาน ให้ตั้งประเด็นใหม่จากเหตุการณ์ที่เกิดขึ้นหรือจากทฤษฎีต่าง ๆ

ทักษะการประเมินค่า มีลักษณะของข้อสอบ คือ (1) ประเมินโดยอาศัยข้อเท็จจริงภายใน (2) การตรวจสอบความบกพร่องของกระบวนการ (3) การวิจารณ์ความถูกต้อง ความเหมาะสมของข้อมูลหรือสถานการณ์ และ (4) การประเมินคุณค่าของผลงาน

2.2 การสร้างเครื่องมือวัด

ข้อสอบอัตนัยก่อนที่จะนำไปใช้วัด ควรมีการตรวจสอบคำถามในประเด็นต่าง ๆ เช่น เป็นคำถามที่ตรงกับระดับพฤติกรรมที่ต้องการวัด เป็นคำถามที่ต้องใช้ความคิดซึ่งยังไม่เคยถูกนำมาเฉลย

หรือไม่มีคำตอบอยู่ในตำรา เป็นคำถามที่มีความชัดเจนมีรายละเอียดและเงื่อนไขพอเพียง เป็นคำถามที่มีคำตอบถูกต้อง มีความสมเหตุสมผล มีการกำหนดคะแนนเต็มของแต่ละข้อ มีคำเฉลย/แนวคำตอบ พร้อมระบุเกณฑ์การให้คะแนนที่มีคำอธิบายพฤติกรรมในแต่ละระดับของผลสัมฤทธิ์ ประกอบด้วย 2 ลักษณะคือ (1) Analytic Rubric คือ การให้คะแนนตามองค์ประกอบย่อย มีคำอธิบายลักษณะของงานในส่วนนั้น ๆ ในแต่ละระดับไว้อย่างชัดเจน และกำหนดน้ำหนักคะแนนของแต่ละส่วนแยกจากกัน (2) Holistic Rubric คือ การให้คะแนนในภาพรวม แนวทางการให้คะแนนพิจารณาจากภาพรวมของผลงาน แนวทางการให้คะแนนอาจใช้การผสมผสานทั้ง 2 วิธี ตัวอย่าง เช่น ให้สรุปผลที่ได้จากการฟังบรรยายพิเศษในหัวข้อ “การทำวิจัยภายใต้นโยบาย Thailand 4.0” โดยกำหนดการให้คะแนนตามลักษณะของ Analytic Rubric คือ (1) การจับใจความ (2) ความถูกต้องตามหลักวิชา และ (3) การสื่อสาร ในแต่ละลักษณะใช้ Holistic Rubric ให้คะแนนในภาพรวม โดยลดระดับคะแนนตามความครบถ้วน ถูกต้อง และเหมาะสม ตัวอย่างการให้คะแนน เช่น

- (1) จับใจความมาสรุป มีคะแนน 4 คะแนน
 - จับใจความได้ดีมีหัวข้อครบถ้วน 4 คะแนน
 - จับใจความได้ดีมีหัวข้อไม่ครบ 3 คะแนน
 - จับใจความได้ปานกลาง มีหัวข้อครบ/ไม่ครบ 2 คะแนน
 - จับใจความได้น้อย มีหัวข้อครบ/ไม่ครบ 1 คะแนน
 - จับใจความไม่ได้เลย 0 คะแนน
- (2) ความถูกต้องของผลสรุป มีคะแนน 3 คะแนน
 - ผลสรุปถูกต้อง 3 คะแนน
 - ผลสรุปถูกบ้างไม่ถูกบ้าง 2 คะแนน
 - ผลสรุปถูกน้อย 1 คะแนน
 - ผลสรุปไม่ถูกเลย 0 คะแนน
- (3) การเขียนสื่อสารอ่านเข้าใจ มีคะแนน 3 คะแนน
 - สื่อความหมายง่ายต่อความเข้าใจ 3 คะแนน
 - สื่อความหมายต้องแปลความ 2 คะแนน
 - สื่อความหมายไม่ชัดเจน ต้องคาดเดา 1 คะแนน
 - สื่อความหมายไม่เข้าใจเลย 0 คะแนน

ในกรณีที่มีผู้สอนหลายคนตรวจข้อสอบข้อเดียวกัน จำเป็นต้องประชุมเพื่อพิจารณากำหนดเกณฑ์การให้คะแนนร่วมกัน

2.3 การดำเนินการวัด

การดำเนินการวัด เป็นขั้นตอนการนำข้อสอบอัตนัยไปให้ผู้เรียนสอบเพื่อวัดทักษะของผู้เรียนตามวัตถุประสงค์เชิงพฤติกรรม เป็นขั้นตอนที่สำคัญอีกขั้นตอนหนึ่งที่ผู้ควบคุมการสอบต้องชี้แจงกฎ กติกา

ข้อกำหนด เงื่อนไขการทำข้อสอบ เงื่อนไขเรื่องเวลา และโทษที่เกิดจากการทุจริตการสอบ และต้องทำหน้าที่ยืนยันการไม่ให้เกิดการทุจริตในการสอบ

2.4 การวิเคราะห์และประเมินผล

การวิเคราะห์ข้อสอบอัตนัย แบ่งเป็น 2 กรณี ดังนี้

2.4.1 การวิเคราะห์ข้อสอบอัตนัยก่อนใช้ข้อสอบ

การวิเคราะห์ข้อสอบอัตนัยก่อนใช้ข้อสอบ เป็นการตรวจสอบคุณภาพข้อสอบทั้งฉบับในเรื่องความตรง (Validity) คือความสามารถของเครื่องมือวัดที่สามารถวัดได้ในสิ่งที่ต้องการโดยจะมุ่งไปที่ความตรงตามเนื้อหาเป็นหลัก การตรวจสอบความตรงตามเนื้อหา ทำได้โดยการวิเคราะห์เนื้อหาของข้อสอบทั้งฉบับโดยอาศัยผู้ทรงคุณวุฒิ

2.4.2 การวิเคราะห์ข้อสอบอัตนัยหลังใช้ข้อสอบ แบ่งเป็น 2 ส่วน คือ การวิเคราะห์ในภาพรวม และการวิเคราะห์รายข้อ

ภายหลังจากที่กระบวนการสอบเสร็จสิ้น ผู้สอนจะทำการตรวจข้อสอบตามเกณฑ์ที่กำหนด ในกรณีที่ผู้ตรวจข้อสอบข้อเดียวกันหลายคน ควรทำการตรวจร่วมกัน เนื่องจากการตอบข้อสอบอัตนัยบางเรื่องมีความหลากหลาย เกณฑ์ที่กำหนดไว้อาจไม่ครอบคลุมทั้งหมด จึงจำเป็นต้องทำการตรวจข้อสอบร่วมกันเพื่อปรึกษาหารือเกี่ยวกับการให้คะแนนและปรับเกณฑ์ หลังจากที่ได้คะแนนสอบรายข้อของผู้สอบแต่ละคนจะเป็นกระบวนการวิเคราะห์ข้อสอบทางสถิติในภาพรวม ได้แก่ การหาคะแนนรวมของผู้สอบแต่ละคน การหาค่ากลางของข้อมูล (Central Tendency) เพื่อพิจารณาคะแนนตัวแทนของผู้สอบทั้งหมด ค่าวัดการกระจาย (Dispersion) เพื่อพิจารณาลักษณะการกระจายของคะแนนสอบ ค่าความเบ้ (Skewness) เพื่อพิจารณาลักษณะความเบ้ของของคะแนนสอบ ซึ่งจะมีผลต่อการเลือกกลุ่มผู้สอบเก่ง-อ่อนสำหรับการวิเคราะห์ข้อสอบรายข้อต่อไป และค่าความเที่ยง (Reliability) หมายถึง ความคงที่ของคะแนนที่วัดได้แต่ละครั้ง ข้อสอบที่มีค่าความเที่ยงใกล้ 1 แสดงว่า ข้อสอบมีความคงเส้นคงวาสูง โดยปกติข้อสอบที่มีค่าความเที่ยงตั้งแต่ 0.75 ขึ้นไปถือว่าข้อสอบมีค่าความเที่ยงสูง

อนุรักษ์ โชติฉิลล และคณะ อธิบายว่า ค่าสถิติต่าง ๆ ที่จำเป็นในการวิเคราะห์ข้อสอบ ได้แก่ ค่าเฉลี่ยเลขคณิต (μ) ค่ามัธยฐาน (Med.) ค่าเบี่ยงเบนมาตรฐาน (σ) ค่าความเบ้ (Skewness) และค่าความเที่ยง (Reliability) [3] มีสูตรการคำนวณ ดังนี้

$$\mu = \frac{\sum x}{N} \quad (1)$$

$$\text{Med.} = X_{\frac{N+1}{2}} \quad (2)$$

$$\sigma = \sqrt{\frac{\sum (x - \mu)^2}{N}} \quad (3)$$

$$\text{Skewness} = \frac{N}{(N-1)(N-2)} \sum \left(\frac{x - \mu}{\sigma} \right)^3 \quad (4)$$

โดยที่ X แทนคะแนน

N แทน จำนวนผู้เข้าสอบทุกคน

วิธีการหาสัมประสิทธิ์ Alpha ของ Cronbach (α) โดยมีสูตรการคำนวณ ดังนี้

$$\alpha = \frac{k}{k-1} \left(1 - \frac{\sum s_i^2}{S_t^2} \right) \quad (5)$$

เมื่อ K คือ จำนวนข้อสอบ

S_i^2 คือ ความแปรปรวนของคะแนนแต่ละข้อ

S_t^2 คือ ความแปรปรวนของคะแนนทั้งฉบับ

การหาค่าความเที่ยงโดยใช้สูตรสัมประสิทธิ์ Alpha (α) สามารถหาได้โดยใช้ผลการสอบจากแบบทดสอบชุดเดียว นำไปสอบกับผู้เข้าสอบกลุ่มเดียวและนำไปใช้ได้ โดยไม่จำกัดเฉพาะแบบทดสอบที่ให้คะแนนแบบ 1 กับ 0 การคำนวณหาค่าสถิติต่าง ๆ แสดงได้ ดังนี้

ตารางที่ 1 คะแนนการตรวจข้อสอบของผู้สอบ 10 คน

	A	B	C	D	E	F	G
	คนที่	ข้อที่ 1	ข้อที่ 2	ข้อที่ 3	ข้อที่ 4	ข้อที่ 5	รวม
1	1	5	5	5	5	5	25
2	2	5	4	4	4	4	21
3	3	4	3	3	3	3	16
4	4	4	5	4	5	4	22
5	5	4	4	4	4	4	20
6	6	3	3	3	3	3	15
7	7	3	2	3	2	3	13
8	8	2	2	2	2	2	10
9	9	1	1	1	1	1	5
10	10	2	3	2	2	2	11
รวม							158
σ_i^2		1.61	1.56	1.29	1.69	1.29	7.44
σ_t^2							34.96

จากตารางที่ 1 นำค่าคะแนนในช่องรวม (ช่อง G) จำนวนหาค่าสถิติต่าง ๆ ได้ ทั้งจำนวนเอง หรือใช้ฟังก์ชันในโปรแกรม Excel [4] ดังนี้

$$\mu = \frac{\sum X}{N} = \frac{158}{10} = 15.8 \text{ สามารถใช้คำสั่งใน Excel ดังนี้ } =\text{AVERAGE (G1:G10)}$$

อธิบายว่า ผู้สอบทั้งสิบคน แต่ละคนสอบได้คะแนน 15.8 คะแนน

$$\text{Med.} = X_{\frac{N+1}{2}} \text{ เมื่อ } X \text{ เรียงคะแนนจากมากไปน้อย}$$

5, 10, 11, 13, 15, 16, 20, 21, 22, 25

$$\text{Med.} = X_{\frac{N+1}{2}} = X_{5.5} = 15.5 \text{ สามารถใช้คำสั่งใน Excel ดังนี้ } =\text{MEDIAN(G1:G10)}$$

อธิบายว่า ผู้สอบทั้งสิบคน แต่ละคนสอบได้คะแนนในตำแหน่งกึ่งกลาง เป็น 15.5 คะแนน

$$\sigma = \sqrt{\frac{\sum (X-\mu)^2}{N}} = \sqrt{\frac{(25-15.8)^2 + \dots + (11-15.8)^2}{10}} = 5.91 \text{ สามารถใช้คำสั่งใน Excel ดังนี้}$$

=STDEVP (G1:G10)

อธิบายว่า คะแนนสอบของผู้สอบทั้งสิบคน มีการกระจายวัดได้ 5.91 คะแนน

$$\text{Skewness} = \frac{N}{(N-1)(N-2)} \sum \left(\frac{X-\mu}{\sigma} \right)^3 = \frac{10}{(9)(8)} \left(\frac{25-15.8}{5.91} \right)^3 + \dots + \left(\frac{11-15.8}{5.91} \right)^3$$

= -0.204 สามารถใช้คำสั่งใน Excel ดังนี้ =SKEW (G1:G10)

อธิบายว่า คะแนนสอบของผู้สอบทั้งสิบคน มีการแจกแจงแบบเบ้ซ้าย (ค่าติดลบ) กล่าวคือ ผู้สอบส่วนใหญ่ทำข้อสอบได้คะแนนสูง

$$\alpha = \frac{k}{k-1} \left(1 - \frac{\sum \sigma_i^2}{\sigma_t^2} \right) = \frac{5}{5-1} \left(1 - \frac{7.44}{34.96} \right) = 0.984$$

σ_i^2 เป็นค่าความแปรปรวนของคะแนนสอบรายข้อ เช่น ข้อสอบข้อที่ 1 $\sigma_1^2 = 1.61$ สามารถใช้คำสั่งใน Excel ดังนี้ =VAR.P(B1:B10)

σ_t^2 เป็นค่าความแปรปรวนของคะแนนสอบรวม $\sigma_t^2 = 34.96$ สามารถใช้คำสั่งใน Excel ดังนี้ =VAR.P (AG1:G10)

ค่า $\alpha = 0.984$ อธิบายว่า ข้อสอบอัตนัยชุดนี้มีความคงเส้นคงวาในระดับสูง (ค่าสัมประสิทธิ์ Alpha มีค่าเข้าใกล้ 1.00)

การวิเคราะห์ข้อสอบอัตนัยรายข้อ เป็นการหาค่าความยากและค่าอำนาจจำแนกของข้อสอบ ข้อสอบที่มีความยากไป (นักเรียนไม่ถึงร้อยละ 20 ทำได้) หรือข้อสอบที่มีความง่าย (นักเรียนมากกว่าร้อยละ 80 ทำได้) ส่วนข้อสอบที่มีค่าอำนาจจำแนก หมายถึง นักเรียนกลุ่มเก่งจะทำข้อสอบในข้อนั้น ๆ ได้มากกว่า นักเรียนกลุ่มอ่อน การพิจารณานักเรียนกลุ่มเก่งหรือกลุ่มอ่อน พิจารณาจากคะแนนรวมของข้อสอบทุกข้อ ที่ผู้สอบแต่ละคนทำได้ ผู้สอบที่ได้คะแนนสูงเรียกว่านักเรียนกลุ่มเก่ง ผู้สอบที่ได้คะแนนน้อยเรียกว่านักเรียนกลุ่มอ่อน

หลังจากที่ได้คะแนนรวมของผู้สอบแต่ละคน ต้องทำการแบ่งกลุ่มนักเรียนที่เข้าสอบออกเป็น 2 กลุ่ม คือกลุ่มเก่ง และกลุ่มอ่อน นภา หลิมรัตน์ (2551) กล่าวว่ากรณีที่ที่มีผู้สอบมากกว่า 1,000 คน ใช้เทคนิค ร้อยละ 27 ของจำนวนผู้สอบในการแบ่งกลุ่ม ถ้ามีจำนวนผู้สอบ 100 ถึง 300 คน ใช้เทคนิคร้อยละ 33 และ ถ้ามีจำนวนผู้เข้าสอบ 30 ถึง 100 คน ใช้เทคนิคร้อยละ 50 [5] อย่างไรก็ตาม Cureton (1972 อ้างถึงใน กังวล เทียนกัณฑ์เทศน์, 2540: 139) กล่าวว่าไว้ว่า กรณีที่คะแนนมีการแจกแจงแบบปกติ ให้ใช้เทคนิคร้อยละ 27 นั้น ถูกต้อง แต่สำหรับการแจกแจงที่ไม่เป็นปกติ ให้ใช้เทคนิคร้อยละ 33 แทน ทั้งนี้เพราะการใช้เทคนิคร้อยละ 27 จะทำให้ระดับความยาก และอำนาจจำแนกไม่ตรงกับความเป็นจริง [6] วิธีการคำนวณเพื่อวิเคราะห์คุณภาพ ของข้อสอบอัตนัย รายข้อมี 4 วิธี คือ (1) วิธีของ Whithney and Sabers (2) วิธีของ Drake (3) วิธีของบุญชม ศรีสะอาด และ (4) วิธีของ อุทุมพร จามรมาน ในบทความนี้จะกล่าวถึงวิธีของ Whithney and Sabers (1970) [7] เท่านั้น โดยมีสูตรหาดัชนีค่าความยาก และดัชนีค่าอำนาจจำแนก ดังนี้

$$(1) \text{ ดัชนีค่าความยาก (I)} \quad I = \frac{S_u + S_L - 2(N)X_{\min}}{2(N)(X_{\max} - X_{\min})} \quad (6)$$

$$(2) \text{ ดัชนีค่าอำนาจจำแนก (r)} \quad r = \frac{S_u - S_L}{N(X_{\max} - X_{\min})} \quad (7)$$

เมื่อ

S_u แทน ผลรวมของคะแนนกลุ่มเก่ง

S_L แทน ผลรวมของคะแนนกลุ่มอ่อน

N แทน จำนวนผู้เข้าสอบในกลุ่มเก่งหรือกลุ่มอ่อน

$X_{\max}$ แทน คะแนนที่นักเรียนทำได้สูงสุด

$X_{\min}$ แทน คะแนนที่นักเรียนทำได้ต่ำสุด

จากตารางที่ 1 นำคะแนนรวมมาเรียงลำดับจากคะแนนมากที่สุดลงไปหาคะแนนน้อยที่สุด เพื่อทำการแบ่งกลุ่มนักเรียนเก่ง(U) และ นักเรียนอ่อน(L) กรณีนี้นักเรียนมีจำนวนน้อยจึงใช้การแบ่งกลุ่มโดยใช้เทคนิคร้อยละ 50 ได้นักเรียนกลุ่มเก่ง 5 คน และนักเรียนกลุ่มอ่อน 5 คน ตามตารางที่ 2

ตารางที่ 2 เรียงลำดับคะแนนจำแนกกลุ่มเก่งกลุ่มอ่อน

คนที่	ข้อที่ 1	ข้อที่ 2	ข้อที่ 3	ข้อที่ 4	ข้อที่ 5	รวม	กลุ่ม
1	5	5	5	5	5	25	เก่ง(U)
4	4	5	4	5	4	22	เก่ง(U)
2	5	4	4	4	4	21	เก่ง(U)
5	4	4	4	4	4	20	เก่ง(U)
3	4	3	3	3	3	16	เก่ง(U)
6	3	3	3	3	3	15	อ่อน(L)
7	3	2	3	2	3	13	อ่อน(L)
10	2	3	2	2	2	11	อ่อน(L)
8	2	2	2	2	2	10	อ่อน(L)
9	1	1	1	1	1	5	อ่อน(L)

พิจารณาข้อสอบข้อที่ 1 ได้ค่าสถิติต่าง ๆ เพื่อประกอบสูตรในการคำนวณ ดังนี้

$$S_U = 22, S_L = 11, N = 5, X_{\max} = 5, X_{\min} = 1$$

การวิเคราะห์ข้อสอบอันดับข้อที่ 1 โดยใช้สูตรของ Whithney and Sabers

$$I = \frac{S_U + S_L - 2(N)X_{\min}}{2(N)(X_{\max} - X_{\min})} = \frac{22 + 11 - 2(5)1}{2(5)(5 - 1)} = 0.58$$

$$r = \frac{S_U - S_L}{N(X_{\max} - X_{\min})} = \frac{22 - 11}{5(5 - 1)} = 0.55$$

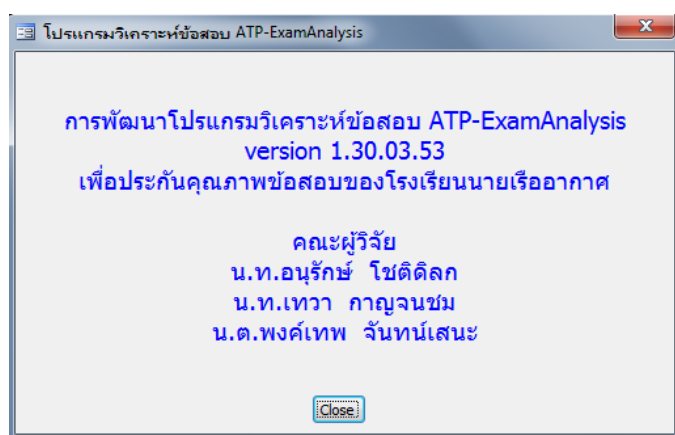
ข้อสอบอันดับข้อที่ 1 มีค่าความยากเท่ากับ 0.58 มีค่าอำนาจจำแนก เท่ากับ 0.55 ช่วงระดับความยากง่ายของปัญหาสอบ และอำนาจจำแนกแสดงได้ ตามตารางที่ 3

ตารางที่ 3 ช่วงระดับความยากง่าย และอำนาจจำแนกของปัญหาสอบ

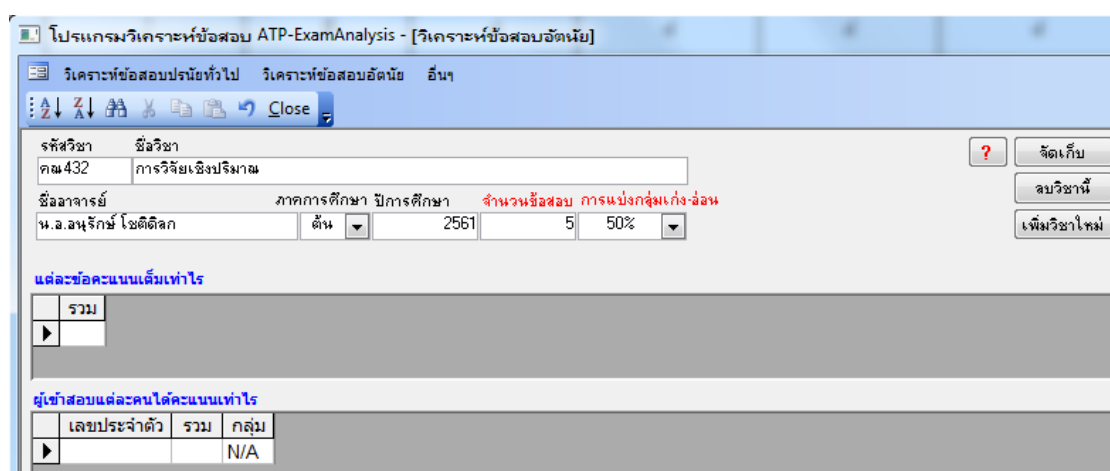
ดัชนีความยากง่าย (P)	ระดับความยากง่าย
0.00 – 0.19	ยากไป
0.20 – 0.80	ใช้ได้
0.81 – 1.00	ง่ายไป
ดัชนีค่าอำนาจจำแนก (r)	ระดับอำนาจจำแนก
น้อยกว่า 2.00	จำแนกไม่ได้
2.00 – 1.00	จำแนกได้

ดังนั้น ข้อสอบอัตนัยข้อที่ 1 สูตรของ Whithney and Sabers มีค่าความยากง่ายเท่ากับ 0.58 มีค่าอำนาจจำแนก เท่ากับ 0.55 แสดงว่า ข้อสอบอัตนัยข้อที่ 1 นี้ ผ่านเกณฑ์ความยากง่าย ($0.20 \leq I \leq 0.80$) และผ่านเกณฑ์ค่าอำนาจจำแนก ($r \geq 0.20$) สรุปว่าข้อสอบอัตนัยข้อที่ 1 มีคุณภาพ เนื่องจากผ่านทั้งเกณฑ์ความยากง่าย และเกณฑ์อำนาจจำแนก

ปัจจุบันมีโปรแกรมสำหรับวิเคราะห์ข้อสอบอัตนัยที่อำนวยความสะดวกในการวิเคราะห์ข้อมูลทางสถิติ เช่น โปรแกรม ATP-ExamAnalysis พัฒนาโดย อนุรักษ์ โชติคิลก เทวา กาญจนชม และ พงศ์เทพ จันทน์เสนะ [8] แสดงขั้นตอนการทำงานของโปรแกรมตามรูปที่ 1 ถึง รูปที่ 5 ได้ดังนี้



รูปที่ 1 ส่วนหน้าของโปรแกรมวิเคราะห์ข้อสอบ ATP-ExamAnalysis



รูปที่ 2 ส่วนรับข้อมูลเพื่อประกอบรายงานและประมวลผล

โปรแกรมวิเคราะห์ข้อสอบ ATP-ExamAnalysis - [วิเคราะห์ข้อสอบอัตโนมัติ]

วิเคราะห์ข้อสอบปรนัยทั่วไป วิเคราะห์ข้อสอบอัตนัย อื่นๆ

รหัสวิชา: คณ432 ชื่อวิชา: การวิจัยเชิงปริมาณ

ชื่ออาจารย์: น.อ.อนุรักษ์ โชติดีติก

ภาคการศึกษา: ดัน ปีการศึกษา: 2561 จำนวนข้อสอบ: 5 การแบ่งกลุ่มเก่ง-อ่อน: 50%

แต่ละข้อคะแนนเต็มเท่าไร

	ข้อ1	ข้อ2	ข้อ3	ข้อ4	ข้อ5	รวม
▶	5	5	5	5	5	25
*	0	0	0	0	0	0

ผู้เข้าสอบแต่ละคนได้คะแนนเท่าไร

เลขประจำตัว	ข้อ1	ข้อ2	ข้อ3	ข้อ4	ข้อ5	รวม	กลุ่ม
▶ 01	5	5	5	5	5	25	เก่ง
02	5	4	4	4	4	21	เก่ง
03	4	3	3	3	3	16	เก่ง
04	4	5	4	5	4	22	เก่ง
05	4	4	4	4	4	20	เก่ง
06	3	3	3	3	3	15	อ่อน
07	3	2	3	2	3	13	อ่อน
08	2	2	2	2	2	10	อ่อน
09	1	1	1	1	1	5	อ่อน
10	2	3	2	2	2	11	อ่อน
*	0	0	0	0	0	0	N/A

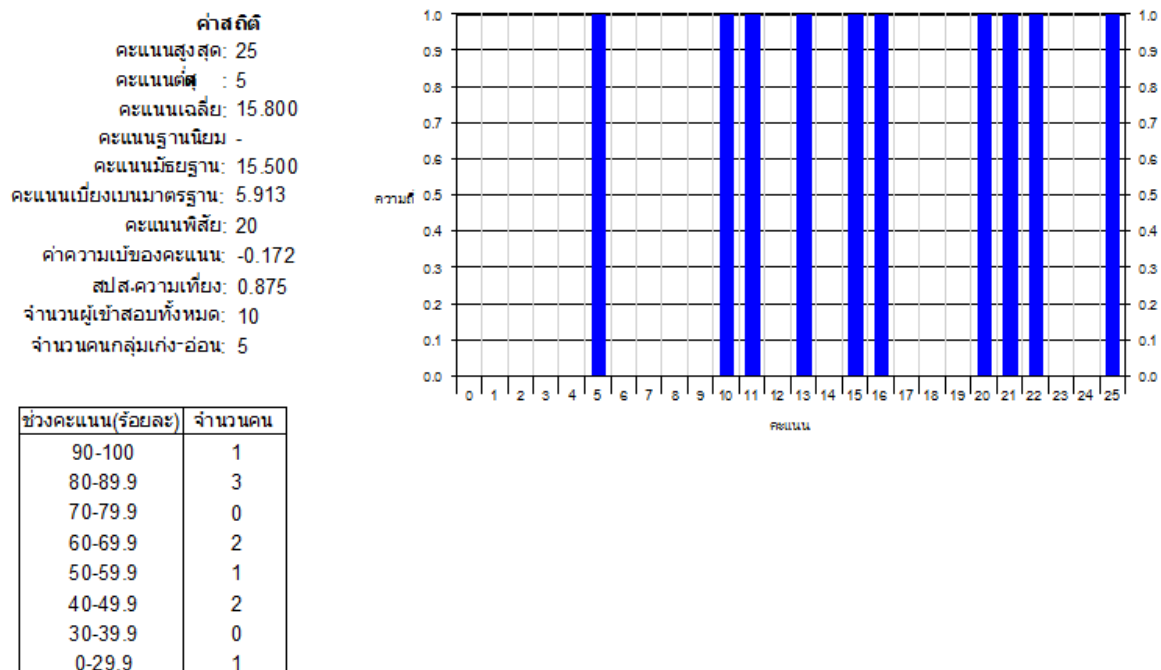
ค้นหาด้วยชื่อวิชา: การวิจัยเชิงปริมาณ/2561

1. วิเคราะห์ 2. วิเคราะห์ ผล > ภาพรวม รายชื่อ ข้อมูล Export

รูปที่ 3 ส่วนของปฏึกการทำงานเพื่อประมวลผลตามขั้นตอนการวิเคราะห์ข้อสอบอัตโนมัติ
โดยวิธี Whithney and Sabers

ผลการวิเคราะห์ข้อสอบอัตโนมัติทั่วไป ประจำภาคการศึกษา ดัน ปีการศึกษา 2561
วิชา การวิจัยเชิงปริมาณ รหัสวิชา คณ432 อาจารย์ผู้สอน น.อ.อนุรักษ์ โชติดีติก (เกณฑ์ 50%)

แผนภูมิแท่งแสดงการกระจายคะแนนสอบ



รูปที่ 4 ส่วนแสดงผลลัพธ์การวิเคราะห์ข้อสอบอัตโนมัติโดยภาพรวม

ผลการวิเคราะห์ข้อสอบอัตนัย ประจำภาคการศึกษา ต้น ปีการศึกษา 2561
วิชา การวิจัยเชิงปริมาณ รหัสวิชา คณ432 อาจารย์ผู้สอน น.อ.ณรงค์ โชติดีเล็ก (เกณฑ์)

ข้อที่	คะแนนเต็ม	SU	SL	Xmax	Xmin	I	r	ผ่านเกณฑ์	หมายเหตุ
1	5	22	11	5	1	0.58	0.55	ผ่าน	
2	5	21	11	5	1	0.55	0.50	ผ่าน	
3	5	20	11	5	1	0.53	0.45	ผ่าน	
4	5	21	10	5	1	0.53	0.55	ผ่าน	
5	5	20	11	5	1	0.53	0.45	ผ่าน	

(NH or NL = 5 คน)

สรุปข้อสอบที่ผ่านเกณฑ์คุณภาพมีจำนวน 5 ข้อ คิดเป็นร้อยละ 100

หมายเหตุ 1. ค่าความยากง่าย(I) ใช้เกณฑ์ 0.20-0.80 และค่าอำนาจจำแนก(r) ใช้เกณฑ์ ≥ 0.20

2. SU คือ ผลรวมคะแนนของคณกลุ่มสูง, SL คือ ผลรวมคะแนนของคณกลุ่มต่ำ

3. Xmax คือ คะแนนสูงสุด, Xmin คือ คะแนนต่ำสุด

4. N คือ จำนวนคนในกลุ่มสูงหรือกลุ่มต่ำ

รูปที่ 5 ส่วนแสดงผลการวิเคราะห์ข้อสอบอัตนัยรายข้อ และการสรุปผลคุณภาพของข้อสอบอัตนัย

รูปที่ 1 เป็นส่วนหน้าของโปรแกรมวิเคราะห์ข้อสอบ ATP-Exam Analysis เป็นงานสิ่งประดิษฐ์คิดค้นที่ส่งเข้าประกวด และได้รับรางวัลชมเชยในงานมหกรรมคุณภาพของกองทัพอากาศ ครั้งที่ 27 เมื่อปี พ.ศ.2533 รูปที่ 2 เป็นส่วนรับข้อมูลเพื่อประกอบส่วนหัวของเอกสารการรายงานผล ประกอบด้วย รหัสวิชา ชื่อวิชา ชื่ออาจารย์ ภาคการศึกษา ปีการศึกษา และส่วนรับข้อมูลเพื่อประมวลผล ประกอบด้วย จำนวนข้อสอบ ร้อยละการแบ่งกลุ่มนักเรียนเก่ง-อ่อน เช่น ร้อยละ 50 ร้อยละ 33 และ ร้อยละ 27 เป็นต้น คะแนนเต็มแต่ละข้อ และคะแนนรายข้อของผู้สอบแต่ละคนที่ทำได้ รูปที่ 3 เป็นส่วนของปุ่มการทำงานเพื่อประมวลผลตามขั้นตอนการวิเคราะห์ข้อสอบอัตนัยโดยวิธี Whithney and Sabers ประกอบด้วยปุ่มจัดกลุ่มนักเรียนเก่ง-อ่อน และปุ่มการวิเคราะห์ รูปที่ 4 เป็นส่วนแสดงผลการวิเคราะห์ข้อสอบอัตนัยโดยภาพรวม ประกอบด้วยค่าสถิติต่าง ๆ ได้แก่ คะแนนสูงสุด คะแนนต่ำสุด คะแนนเฉลี่ย คะแนนฐานนิยม คะแนนมัธยฐาน คะแนนเบี่ยงเบนมาตรฐาน คะแนนพิสัย ค่าความเบ้ สัมประสิทธิ์ความเที่ยง จำนวนผู้เข้าสอบทั้งหมด จำนวนผู้สอบที่เป็นกลุ่มเก่ง-กลุ่มอ่อน ค่าร้อยละ และแผนภูมิแท่งแสดงการกระจายของคะแนนสอบ รูปที่ 5 เป็นส่วนแสดงผลการวิเคราะห์ข้อสอบอัตนัยรายข้อ ประกอบด้วย ค่าความยากง่าย (I) ค่าอำนาจจำแนก (R) แปลผล และการสรุปผลคุณภาพของข้อสอบอัตนัย โดยคิดเป็นค่าร้อยละของข้อสอบที่มีคุณภาพผ่านเกณฑ์ความยากง่าย และเกณฑ์อำนาจจำแนก เทียบกับจำนวนข้อสอบทั้งหมด

2.5 การบันทึกผล

การบันทึกผล เป็นการสรุปรายงานผลการวิเคราะห์ข้อสอบอัตนัยทั้งในภาพรวมและรายข้อ โดยนำข้อสอบรายข้อที่ไม่ผ่านเกณฑ์คุณภาพ เช่น ยากเกินไป ยากเกินไป หรือไม่มีอำนาจจำแนกนักเรียนเก่ง-อ่อน มาทำการวิเคราะห์เพื่อหาสาเหตุ เมื่อได้ผลการวิเคราะห์ผู้รับผิดชอบวิชาจะนำข้อมูลที่ได้ไปใช้ในการพิจารณาการออกข้อสอบอัตนัยในครั้งต่อไป หรือนำไปใช้ตอบคำถามกรณีที่มีข้อสงสัยในเรื่องคุณภาพของข้อสอบ

3. สรุปและข้อเสนอแนะ

กระบวนการพัฒนาการออกข้อสอบอัตนัยของกองทัพไทย เพื่อให้ได้คุณภาพ ควรพิจารณาประเด็น 5 เรื่อง ได้แก่ (1) การกำหนดความมุ่งหมายของการวัดผล (2) การสร้างและพัฒนาเครื่องมือ/เกณฑ์ (3) การดำเนินการวัด (4) การวิเคราะห์และประเมินผล และ (5) การบันทึกผล ดังนี้

การกำหนดความมุ่งหมายของการวัดผลของข้อสอบอัตนัย พิจารณาทักษะของผู้เรียน 5 ทักษะ ได้แก่ (1) ทักษะเข้าใจ (2) ทักษะการนำไปใช้ (3) ทักษะการวิเคราะห์ (4) ทักษะการสังเคราะห์ และ (5) ทักษะการประเมินค่า

การสร้างและพัฒนาเครื่องมือ/เกณฑ์ ข้อสอบอัตนัยก่อนที่จะนำไปใช้วัด ควรมีการตรวจสอบประเด็นต่าง ๆ ได้แก่ คำถามตรงกับระดับพฤติกรรมที่ต้องการวัด คำถามที่ผู้ตอบต้องใช้ความคิดเป็นคำถามที่มีความชัดเจน มีรายละเอียด และเงื่อนไขพอเพียง คำถามที่มีคำตอบถูกต้องสมเหตุสมผล บอกคะแนนเต็มของแต่ละข้อ มีคำตอบ/แนวคำตอบ พร้อมระบุเกณฑ์การให้คะแนน โดยผสมผสานระหว่างเกณฑ์การให้คะแนนแบบ Analytic Rubric และแบบ Holistic Rubric ในกรณีที่ผู้สอนหลายคนตรวจข้อสอบข้อเดียวกัน จำเป็นต้องประชุมเพื่อพิจารณาเกณฑ์การให้คะแนนร่วมกัน ภายหลังจากที่สร้างข้อสอบและกำหนดเกณฑ์การให้คะแนนเรียบร้อยแล้ว ให้มีคณะกรรมการทำหน้าที่ตรวจคุณภาพของข้อสอบและเกณฑ์การให้คะแนนก่อนที่จะนำข้อสอบไปใช้ โดยมีประเด็นการตรวจ ดังนี้ ตรวจข้อสอบแต่ละข้อว่ามีความสอดคล้องกับวัตถุประสงค์เชิงพฤติกรรมที่ต้องการวัดหรือไม่ ครอบคลุมทุกเนื้อหาอย่างเป็นสัดส่วนหรือไม่ ตรวจสอบความเหมาะสมของเกณฑ์การให้คะแนน การสะกดคำ ไวยากรณ์ ความเข้าใจในประเด็นคำถาม เป็นต้น และมีอำนาจให้ผู้ออกข้อสอบชี้แจงเหตุผล หรือให้แก้ไขได้ตามข้อเสนอแนะการตรวจข้อสอบอัตนัย ผู้ตรวจควรอ่านคำตอบทุกคนอย่างคร่าว ๆ ก่อนเป็นอันดับแรก ปรับคำเฉลย และเกณฑ์การตรวจคำตอบ ควรตรวจคำตอบทีละข้อไปจนครบทุกคน ไม่ควรทราบว่าผู้ตอบคือใคร ควรมีการตรวจทานคะแนนอีกครั้ง บันทึกข้อผิดพลาดของผู้ตอบ เพื่อนำไปพิจารณาปรับปรุงการสอนต่อไป

การดำเนินการวัด เป็นขั้นตอนการนำข้อสอบอัตนัยไปให้ผู้เรียนสอบเพื่อวัดทักษะของผู้เรียนตามวัตถุประสงค์เชิงพฤติกรรม

การวิเคราะห์และประเมินผล เป็นการวิเคราะห์ก่อนนำข้อสอบไปใช้โดยผู้ทรงคุณวุฒิเพื่อหาความตรงของเนื้อหา และวิเคราะห์ภายหลังจากที่นำข้อสอบไปใช้แล้ว มาทำการวิเคราะห์ในภาพรวม โดยใช้คะแนนจากการตรวจแล้ว นำมาวิเคราะห์หาค่าสถิติต่าง ๆ ได้แก่ ค่ากลางของข้อมูล ค่าวัดการกระจาย ค่าความเบ้ และ ค่าความเที่ยง สำหรับการวิเคราะห์ข้อสอบรายข้อใช้วิธีของ Whithney and Sabers จากนั้นนำผลลัพธ์ที่ได้มาประเมินคุณภาพของข้อสอบ

การบันทึกผล เป็นการสรุปรายงานผลการวิเคราะห์ข้อสอบอัตนัยทั้งในภาพรวมและรายข้อ ผลการวิเคราะห์จะใช้ในการพิจารณาการออกข้อสอบอัตนัยในครั้งต่อไป หรือนำไปใช้ตอบคำถามกรณี que ผู้สอบมีข้อสงสัยในเรื่องคุณภาพของข้อสอบ

เอกสารอ้างอิง

- [1] วิรัช วรรณรัตน์.(2558). หลักและวิธีการสอบวัด. วารสารมนุษยศาสตร์และสังคมศาสตร์ มหาวิทยาลัยราชพฤกษ์, ปีที่ 1, เล่มที่ 2, หน้า 1-12, 2558.
- [2] Bloom, B.S. and others.(1956).Taxonomy of educational objectives Cognitive Domain. New York : Mckay.
- [3] อนุรักษ์ โขติติติก, เทวา กาญจนชม และ พงศ์เทพ จันทน์เสนะ, “การวิจัยและพัฒนาโปรแกรมประเมินผลการศึกษาของกองทัพอากาศ(RTAF-EEP),” วารสารวิชาการนายเรืออากาศ, ปีที่ 13, เล่มที่ 13, หน้า 7-18, 2560.
- [4] ปรีดาภรณ์ กาญจนสำราญวงศ์, Excel Statistics Analysis. กรุงเทพฯ: อินโฟเพลส, 2561.
- [5] นภา หลิมรัตน์, Item Analysis [Online], 2551. Available: <http://teachingresources.psu.ac.th/document/2551/limrat/9Item%20Analysis.pdf> (20 สิงหาคม 2561).
- [6] กังวล เทียนกันต์เทศน์, การวัด การวิเคราะห์ การประเมิน. กรุงเทพมหานคร: Bangkok Software Technology House, 2540.
- [7] Whitney D.R. and Sabers D.L. , Improving Essay Examinations III, Use of Item Analysis, Technical Bulletin II, (Mimeographed). Iowa City: University Evaluation and Examination Service, 1970.
- [8] อนุรักษ์ โขติติติก, เทวา กาญจนชม และ พงศ์เทพ จันทน์เสนะ, “การลดเวลาการรายงานผลการวิเคราะห์ข้อสอบ,” การนำเสนอผลงาน QCC งานมหกรรมคุณภาพกองทัพอากาศ ครั้งที่ 27, กรุงเทพฯ, 2553.