

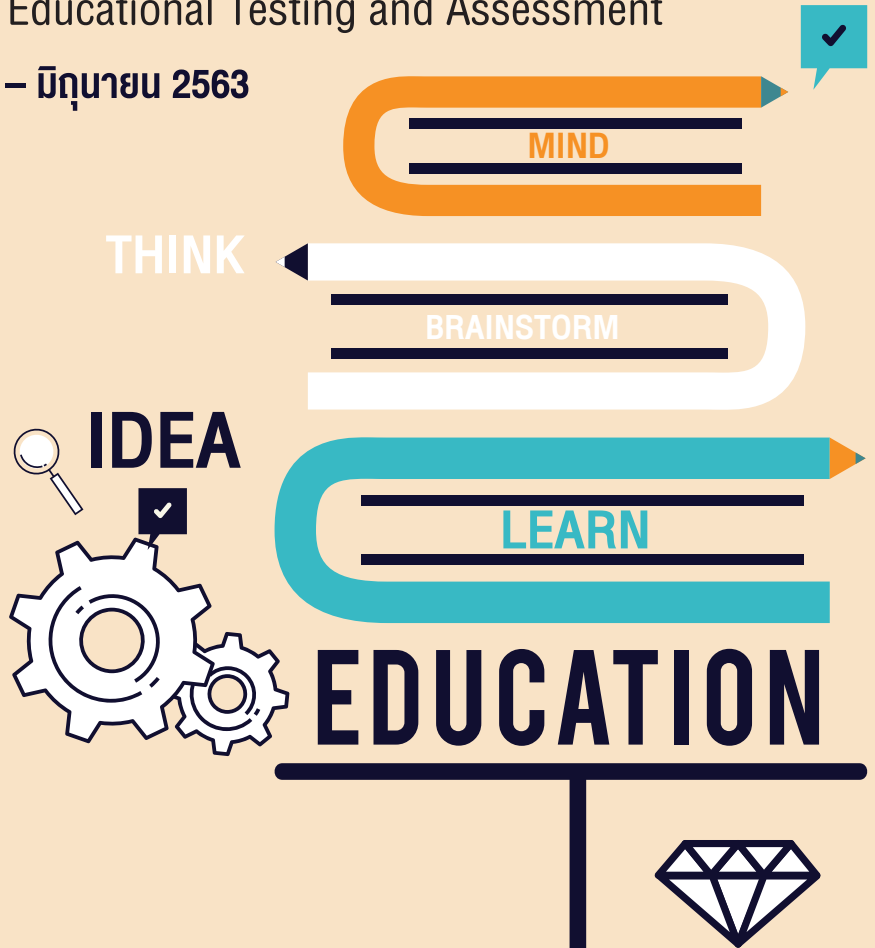
วารสาร

การทดสอบและการประเมินทางการศึกษาระดับชาติ

Journal of National Educational Testing and Assessment

ปีที่ 1 ฉบับที่ 1 มกราคม – มิถุนายน 2563

ISSN 2730-3535



สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน)

คำนิยม

สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) (สทศ.) จัดตั้งขึ้นโดยมีวัตถุประสงค์เพื่อเป็นสถาบันทดสอบและประเมินผลทางการศึกษาที่เป็นแหล่งอ้างอิงระดับชาติและนานาชาติ

“วารสารการทดสอบและประเมินทางการศึกษาระดับชาติ (Journal Of National Educational Testing and Assessment)” ฉบับนี้จัดทำขึ้นเพื่อเผยแพร่บทความวิชาการ การศึกษาวิจัย นวัตกรรมเกี่ยวกับการทดสอบทางการศึกษาเทคนิคการวัดและประเมินผลทางการศึกษา

การจัดทำวารสารฉบับนี้จะเป็นประโยชน์ต่อการยกระดับ สทศ. สู่การเป็นสถาบันวิชาการ (Center of Excellence) และศูนย์กลางข้อมูลด้านการทดสอบทางการศึกษาระดับนานาชาติ และเป็นส่วนหนึ่งในการพัฒนาการศึกษาชาติ



(นายกิตติรัตน์ มั่งคละศิริ)

ประธานกรรมการบริหาร สทศ.

คำนิยม

ขอแสดงความยินดี และชื่นชม สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) หรือ สทศ. ที่ได้จัดทำ “วารสารการทดสอบและประเมินทางการศึกษาระดับชาติ” เพื่อเผยแพร่วิทยาการด้านการทดสอบและการประเมิน เป็นช่องทางสำหรับการนำเสนอความรู้ ความคิด ที่เป็นนวัตกรรมทางการทดสอบและการประเมินทางการศึกษาสำหรับนักการศึกษาทั้งในประเทศและต่างประเทศ การดำเนินการจัดทำวารสารต้องใช้ความพยายาม และความตั้งใจจริงของฝ่ายบริหาร เพื่อทำให้เกิดวารสารนี้ ถือเป็นความพยายามและความตั้งใจที่จะทำให้วิทยาการด้านการทดสอบและการประเมินมีการพัฒนาอย่างต่อเนื่องโดย สทศ. เป็นผู้นำทางวิชาการสำหรับวิทยาการด้านนี้

ถึงแม้ว่าวิทยาการด้านการทดสอบและประเมินทางการศึกษาในประเทศไทยจะมีการจัดการเรียนการสอนถึงระดับปริญญาเอก แต่ต้นกำเนิดของวิทยาการด้านนี้ไม่ได้เกิดขึ้นในประเทศไทย การอาศัยองค์ความรู้เริ่มต้นจึงเป็นการนำความคิด ความรู้และวิทยาการต่าง ๆ ที่เกี่ยวข้องมาจากต่างประเทศ เมื่อนำองค์ความรู้นี้เข้ามาในสังคมไทย นักการศึกษาไทยได้ประยุกต์ วิจัย และปรับปรุงให้เหมาะสมกับบริบทของสังคมไทย พร้อมทั้งมีการพัฒนาต่อยอดความคิด สร้างองค์ความรู้อีกเป็นจำนวนมาก แสดงให้เห็นถึงความก้าวหน้า และพัฒนาการด้านการทดสอบและการประเมินทางการศึกษาในประเทศไทยที่ดำเนินมาอย่างต่อเนื่อง

การเผยแพร่ความรู้และความคิดต่าง ๆ ในบริบทของสังคมไทย โดยนักการศึกษาของไทยจึงสมควรได้รับการยกย่อง และนำมาเผยแพร่ให้เป็นที่ยอมรับในหมู่นักวิชาการด้านการศึกษาของไทย ผ่านช่องทางการสื่อสารต่าง ๆ และวารสารทางวิชาการ เป็นช่องทางการสื่อสารการเผยแพร่ที่สำคัญ และเป็นที่น่าเชื่อถือสำหรับการเผยแพร่วิทยาการต่าง ๆ การที่ สทศ. นำการเผยแพร่ด้วยวารสารวิชาการมาใช้เป็นสื่อสำหรับการเผยแพร่ ถือเป็น การสร้างสื่อที่มีคุณค่า น่าเชื่อถือให้กับแวดวงวิชาการทางการศึกษาของประเทศไทยเป็นอย่างมาก

เนื่องจากวารสารฉบับนี้เป็นฉบับแรก บทความที่นำมาเผยแพร่ในวารสารฉบับนี้ได้ผ่านการคัดเลือกอย่างรอบคอบเป็นอย่างดี แสดงถึง ขอบเขตเนื้อหาสาระที่วารสารนี้จะใช้เป็นแนวทางในการเผยแพร่วิทยาการด้านการทดสอบและการประเมินทางการศึกษา เมื่อผู้อ่านได้อ่านบทความในวารสารนี้แล้ว เชื่อมั่นว่าจะได้ประโยชน์และมีความเข้าใจในขอบเขตและเนื้อหาสาระที่วารสารนี้จะนำเสนอในฉบับต่อ ๆ ไป ขอเป็นกำลังใจให้ทีมงานผู้จัดทำวารสารนี้ ดำเนินการให้วารสารนี้อยู่คู่กับ สทศ. เป็นที่สำหรับการเผยแพร่ เป็นที่สำหรับการอ้างอิงทางวิชาการ และเป็นแหล่งข้อมูลทางวิชาการให้กับนักการศึกษาของไทยตลอดไป



(ศาสตราจารย์ ดร.กฤษมันต์ วัฒนานรงค์)

กรรมการบริหาร สทศ.

คำนิยม

สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) จัดตั้งขึ้นตามพระราชบัญญัติการจัดตั้งสถาบันทดสอบทางการศึกษาแห่งชาติ(องค์การมหาชน) มีวัตถุประสงค์เพื่อบริหารจัดการและดำเนินการเกี่ยวกับการศึกษาวิจัยพัฒนา และให้บริการทางการประเมินผลทางการศึกษาและทดสอบทางการศึกษาอย่างต่อเนื่อง รวมทั้งเป็นศูนย์กลางความร่วมมือด้านการทดสอบทางการศึกษาในระดับชาติและระดับนานาชาติ

การจัดทำวารสาร “การทดสอบและประเมินผลการศึกษาในระดับชาติ (Journal Of National Educational Testing and Assessment)” เป็นการนำเสนอบทความทางวิชาการและบทความวิจัย ที่เกี่ยวข้องกับด้านการวัดและประเมินทางการศึกษาระดับชาติ ที่แสดงถึงประโยชน์ทั้งเชิงทฤษฎี (Theoretical Contributions) ที่นักวิจัยหรือผู้ที่สนใจสามารถนำไปพัฒนา สร้างองค์ความรู้ใหม่ และนำไปประยุกต์ใช้ให้เกิดประโยชน์หรือต่อยอดงานวิจัยด้านการวัดและประเมินผล นอกจากนี้ ยังเป็นการบริการวิชาการแก่สังคมในการเป็นศูนย์กลางเผยแพร่ แลกเปลี่ยนความคิดเห็น ประสพการณ์ การทำวิจัยและการพัฒนาผลงานวิชาการ ของนักวิจัย นักวิชาการ คณาจารย์ ผู้บริหาร นักศึกษา และประชาชนผู้สนใจ เกี่ยวกับการจัดทำระบบ วิธีการทดสอบ การพัฒนาเครื่องมือวัดและประเมินผล การวิเคราะห์ผลการทดสอบ การติดตามและประเมินผลกระทบ รวมทั้งการวิเคราะห์งานวิจัย

วารสารของ สทศ. ฉบับนี้เป็นฉบับแรก และในอนาคตวารสารสทศ. จะเป็นเวทีกลางให้กับนักวิจัย นักวิชาการ ผู้ที่สนใจด้านการวัดและประเมินผล ในการเผยแพร่ นวัตกรรมเกี่ยวกับการทดสอบทางการศึกษาตลอดจนเผยแพร่เทคนิคการวัดและประเมินผลการศึกษา นำไปสู่การเป็นองค์กรวิชาการและวิชาชีพทางการวัดและประเมินผลและเป็น Center of Excellence ทางด้านการวัดและประเมินผลทั้งในระดับชาติและนานาชาติ



(ผู้ช่วยศาสตราจารย์ ดร.ศิริดา บุรชาติ)

ผู้อำนวยการสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน)

บรรณาธิการแถลง

เรียน ท่านผู้อ่านวารสารการทดสอบและการประเมินทางการศึกษาระดับชาติ

วารสารการทดสอบและการประเมินทางการศึกษาระดับชาติ ของ สทศ. ฉบับที่ 1 ประจำปีพุทธศักราช 2563 (เดือนมกราคม-มิถุนายน 2563) ฉบับนี้ ได้รวบรวมองค์ความรู้ด้านการวัดและประเมินทางการศึกษาระดับชาติในหลายระดับ รวมทั้งการวิเคราะห์ผลการสอบ สังเคราะห์งานวิจัยทางการศึกษา และการประเมินคุณภาพงานวิจัย

สทศ. ขอขอบพระคุณผู้ทรงคุณวุฒิ ผู้เชี่ยวชาญ เครือข่ายความร่วมมือ และผู้ทรงคุณวุฒิที่อนุเคราะห์ร่วมประเมินบทความฉบับนี้ ให้มีความสมบูรณ์ ถูกต้อง สามารถใช้อ้างอิงเป็นตำราแบบอย่าง ให้นำวัดและประเมินในยุคปัจจุบัน หรือแม้แต่ผู้ที่สนใจได้เรียนรู้ ทำความเข้าใจ ถึงการวัดและประเมินระดับชาติ และเป็นช่องทางในการแลกเปลี่ยนองค์ความรู้ทางวิชาการด้านการวัดและประเมินผลทางการศึกษาต่อไป

นางนิลรัตน์ ทิพย์ทัศนัน

บรรณาธิการ

กองบรรณาธิการ

บรรณาธิการ

นางนิลรัตน์ ทิพย์ทัศนัย

กองบรรณาธิการ

1. ผู้ช่วยศาสตราจารย์ ดร.ศิริดา บุรชาติ	ที่ปรึกษา
2. รองศาสตราจารย์ ดร.สิริพันธุ์ สุวรรณมรรคา	กองบรรณาธิการ
3. รองศาสตราจารย์ ดร.อิสรา ก้านจักร	กองบรรณาธิการ
4. ผู้ช่วยศาสตราจารย์ ดร.มาเรียม นิลพันธุ์	กองบรรณาธิการ
5. ผู้ช่วยศาสตราจารย์ ดร.สุวิมล กฤษณกุล	กองบรรณาธิการ
6. ผู้ช่วยศาสตราจารย์ ดร.สุมาลี ศรีพุทธรินทร์	กองบรรณาธิการ
7. ดร.เสกสรรค์ ทองคำบรรจง	กองบรรณาธิการ
8. นางนิลรัตน์ ทิพย์ทัศนัย	เลขานุการ

ดำเนินการโดย: สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน)
128 อาคารพญาไท พลาซ่า ชั้น 36 แขวงทุ่งพญาไท เขตราชเทวี
กรุงเทพฯ 10400
ติดต่อกอง บก. โทร. 02-217-3800 ต่อ 113 / โทรสาร: 02-219-2996
E-mail: JOURNAL.NIETS@GMAIL.COM

เจ้าของ : สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน)
วาระการออก : ราย 6 เดือน ปีงบประมาณละ 2 ฉบับ (มกราคม-มิถุนายน, กรกฎาคม-ธันวาคม)
จำนวนฉบับละ : 100 เล่ม
พิมพ์ที่ : บริษัท เปเปอร์เมท (ประเทศไทย) จำกัด
วัตถุประสงค์ : 1. เพื่อเผยแพร่ผลงานวิจัย บทความวิชาการ นวัตกรรม และเทคนิคด้านวัดและ
ประเมินผลทางการศึกษา ของบุคลากร นักวิจัย และนักวิชาการ
2. เพื่อเป็นเอกสารทางวิชาการ ที่ให้ความรู้เกี่ยวกับการวัดและประเมินผลทาง
การศึกษาที่สามารถใช้เป็นแหล่งอ้างอิงที่เชื่อถือได้

ผู้ทรงคุณวุฒิ

ผู้ทรงคุณวุฒิประเมินบทความประจำฉบับ

ศาสตราจารย์ ดร.บุญธรรม กิจปรีดาบริสุทธิ์	มหาวิทยาลัยมหาสารคาม
ศาสตราจารย์ ดร.สำเร็จ บุญเรืองรัตน์	มหาวิทยาลัยศรีนครินทรวิโรฒ
รองศาสตราจารย์ ดร.ดวงกมล ไตรวิจิตรคุณ	จุฬาลงกรณ์มหาวิทยาลัย
ผู้ช่วยศาสตราจารย์ ดร.สมพงษ์ ปั่นหุ่น	มหาวิทยาลัยบูรพา
รองศาสตราจารย์ ดร.พรทิพย์ ไชยโส	มหาวิทยาลัยเกษตรศาสตร์
ผู้ช่วยศาสตราจารย์ ดร.เรืองเดช ศิริกิจ	มหาวิทยาลัยศรีนครินทรวิโรฒ
ผู้ช่วยศาสตราจารย์ ดร.สังวรณ์ ังดกระโทก	มหาวิทยาลัยสุโขทัยธรรมธิราช
รองศาสตราจารย์ ดร.ปานใจ ธารทัศนวงศ์	มหาวิทยาลัยศิลปากร
รองศาสตราจารย์ ดร.ดุขุฎี อินทรประเสริฐ	มหาวิทยาลัยศรีนครินทรวิโรฒ
ผู้ช่วยศาสตราจารย์ ดร.กันต์ฤทัย คลังพล	มหาวิทยาลัยราชภัฏวไลยอลงกรณ์
รองศาสตราจารย์ ดร.สมคิด พรหมจู้	มหาวิทยาลัยสุโขทัยธรรมธิราช
ผู้ช่วยศาสตราจารย์ ดร.เพชรรา พิพัฒน์สันติกุล	มหาวิทยาลัยเทคโนโลยีราชมงคลกรุงเทพ
รองศาสตราจารย์ ดร.กมลวรรณ ตังชันกานนท์	จุฬาลงกรณ์มหาวิทยาลัย
รองศาสตราจารย์ ดร.บังอร เสรีรัตน์	มหาวิทยาลัยราชภัฏบ้านสมเด็จเจ้าพระยา
รองศาสตราจารย์ ดร.วรรณิ์ แกมเกตุ	จุฬาลงกรณ์มหาวิทยาลัย
ผู้ช่วยศาสตราจารย์ ดร.ณัฐภรณ์ หลาวทอง	จุฬาลงกรณ์มหาวิทยาลัย

	หน้า
1. ระบบการคัดเลือกนักเรียนเข้าศึกษาต่อในสถาบันอุดมศึกษา : เรียนจากแชมป์ เพื่อเป็นแชมป์ (Selection Systems of Students for Further Studies in Higher Education Institutions: Learning from Champions to be the Champion) สมหวัง พิธิยานุวัฒน์ และคณะ	1
2. การประเมินโครงการจัดหาตัวแทน สทศ. ประจำสนามสอบ การทดสอบ ทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) ชั้นประถมศึกษาปีที่ 6 และ ชั้นมัธยมศึกษาปีที่ 3 (The Procurement Representative of the National Institute of Educational Testing Service (Public Organization) to Ordinary 6 and Mathayom 3 Field Exam Project) พัฒนา ธนากร และคณะ	17
3. การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษ (Computerized Adaptive Testing for The Ordinary National Educational Test (O-NET) in The Subject of English Language) อ.ดร.ณภัทร ชัยมงคล	38
4. การวิเคราะห์ผลการสอบ O-NET เพื่อพัฒนาบทเรียนออนไลน์แบบเปิดเพื่อมหาชน สำหรับแก้ไขข้อบกพร่องที่คลาดเคลื่อนและข้อผิดพลาดทางคณิตศาสตร์ของนักเรียน ระดับมัธยมศึกษาตอนต้น (Analysis of O-NET Results to Develop a Massive Online Open Course for Corrections of Misconceptions and Mathematical Errors of Junior High-school Students) ผศ. ณวิวรรณ แก้วไทรฮะ และคณะ	51

	หน้า
5. การประเมินคุณภาพงานวิจัยการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนใน ศตวรรษที่ 21 ระดับมัธยมศึกษา (An evaluation of researches quality on the 21st century competencies measurement tools development for secondary school students) สร้อยญา จันทร์ชูสกุล และคณะ	74
6. การติดตามและประเมินผลกระทบของการประเมินผลการทดสอบทางการศึกษา ระดับชาติ (Monitoring and Evaluation the Impact of the Evaluation of National Educational Testing) ผศ.ดร.จตุภูมิ เขตจัตุรัส และคณะ	92
7. การสังเคราะห์งานวิจัยทางการศึกษาและการวิเคราะห์อภิมาน (Synthesis of Educational Research and Meta Analysis) ผศ.ดร.นันทิมา นาคาพงศ์ อัครวัักษ์	131

ระบบการคัดเลือกนักเรียนเข้าศึกษาต่อในสถาบันอุดมศึกษา : เรียนจากแชมป์เพื่อเป็นแชมป์

Selection Systems of Students for Further Studies in Higher Education
Institutions: Learning from Champions to be the Champion

สมหวัง พิธิยานุวัฒน์
จิราทัศน์ รัตนมณีฉัตร
แทน พิธิยานุวัฒน์

บทคัดย่อ

บทความนี้มีวัตถุประสงค์เพื่อนำเสนอระบบการคัดเลือกนักเรียนเข้าศึกษาต่อในสถาบันอุดมศึกษาของประเทศ 4 ประเทศ ซึ่งได้รับจัดอันดับว่าเป็นประเทศที่มีระบบการศึกษาชั้นมัธยมศึกษาตอนปลายที่ดีระดับมาตรฐานโลก ได้แก่ ประเทศสาธารณรัฐสิงคโปร์ (อันดับที่ 1) ประเทศสาธารณรัฐประชาธิปไตยประชาชนเกาหลี (เกาหลีใต้) (อันดับที่ 3) ประเทศฟินแลนด์ (อันดับที่ 6) และประเทศสหราชอาณาจักร (อันดับที่ 20) จากการศึกษาพบว่า 3 ใน 4 ประเทศใช้ผลการทดสอบระดับชาติเพื่อจบการศึกษามัธยมศึกษาปีที่ 6 ในการคัดเลือกบุคคลเข้าศึกษาต่อในสถาบันอุดมศึกษาโดยไม่มีการสอบคัดเลือก ตอนท้ายของบทความได้นำเสนอข้อเสนอแนะเชิงนโยบายเกี่ยวกับระบบการคัดเลือกนักเรียนเข้าศึกษาต่อในสถาบันอุดมศึกษาของไทย

คำสำคัญ: การคัดเลือกนักเรียน, การศึกษาต่อในสถาบันอุดมศึกษา, การศึกษาระดับมัธยมศึกษาตอนปลาย, ผลการสอบเพื่อจบการศึกษาระดับมัธยมศึกษาปีที่ 6, การสอบคัดเลือกระดับชาติเพื่อศึกษาต่อ

Abstract

This objective of this article is to demonstrate the selection systems of students for further studies in higher education institutions in four selected countries, which have the worldclass high schooleducation systems. These countries are the Republic of Singapore (1st rank), The Republic of Korea (South Korea) (3rd rank), Finland (6th rank) and The United Kingdom (20th rank). It was found that 3 out of 4 countries employ the National High School Graduation Examination results in student selections for further studies in higher education institutions without implementing the national entrance examination. Finally, this paper presents policy recommendations concerning the selection systems of students for further studies in Thai Higher Education institutions.

Keyword: the Selection Systems of Students, Further Studies in Higher Education Institutions, High School Education, National High School Graduation Examination, National Entrance Examination

คำนำ

ระบบการคัดเลือกบุคคลเข้าศึกษาต่อในมหาวิทยาลัยของประเทศไทย หรือ TCAS (Thai University Center Admission System) เป็นระบบที่มีความสำคัญต่ออนาคตและชีวิตของนักเรียนที่สำเร็จการศึกษาระดับมัธยมศึกษาปีที่ 6 แล้วประสงค์ที่จะศึกษาต่อในสาขาวิชาคณะวิชา และมหาวิทยาลัยชั้นนำที่พึงประสงค์เป็นที่น่าสนใจที่ว่า ทีแคสที่ผ่านมาเป็นระบบที่ให้โอกาสแก่นักเรียนแต่เอื้อให้เกิดการกักตุนที่นั่งก่อให้เกิดคะแนนสูงผิดปกติและทำให้นักเรียนหลายคนผิดหวัง ซึ่งน่าจะมีการศึกษาทบทวนเพื่อปรับปรุงระบบการศึกษาและระบบทีแคสในอนาคตให้ดียิ่งขึ้น

ในพระราชบัญญัติการศึกษาแห่งชาติ พ.ศ. 2542 แก้ไขเพิ่มเติม (ฉบับที่ 2) พ.ศ. 2545 มาตรา 26 ได้ระบุกำหนดว่า “ให้สถานศึกษาจัดการประเมินผู้เรียนโดยพิจารณาจากพัฒนาการของผู้เรียน ความประพฤติ การสังเกตพฤติกรรมการเรียน การร่วมกิจกรรม และการทดสอบความรู้ในกระบวนการเรียนการสอนตามความเหมาะสมของแต่ละระดับและรูปแบบการศึกษา” อีกทั้งในวรรคสองได้ระบุ “ให้สถานศึกษาใช้วิธีการที่หลากหลายในการจัดสรรโอกาสการเข้าศึกษาต่อ และให้นำผลการประเมินผู้เรียนตามวรรคหนึ่งมาใช้ประกอบการพิจารณาด้วย”

ศาสตราจารย์ ดร. สุวัชรวิ สุวรรณสวัสดิ์ อธิการบดี สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง ในฐานะประธานที่ประชุมอธิการบดีแห่งประเทศไทย (ทปอ.) ได้เปิดเผย (ไทยรัฐฉบับพิมพ์ 26 สิงหาคม 2562) ระบบทีแคสปีการศึกษา 2563 จะยังคงใช้ระบบเดิมเช่นเดียวกับปีการศึกษา 2562 แต่จะทำการสื่อสารและประชาสัมพันธ์ในเรื่องการรวบรวมการยืนยันสิทธิ์รอบที่ 3 และรอบที่ 4 เพื่อนำไปสู่ทีแคสใหม่ในปีการศึกษา 2564 ขณะเดียวกันจะประกาศเกณฑ์การคัดเลือกรอบแอดมิชชันปีการศึกษา 2566 และเตรียมข้อสอบแอดหรือแบบทดสอบวิชาความถนัดทั่วไปซึ่งเป็นการทดสอบศักยภาพในการเรียนในมหาวิทยาลัยให้ประสบความสำเร็จโดยการทดสอบความสามารถในการอ่าน การเขียน การคิดเชิงวิเคราะห์ การแก้โจทย์ปัญหา และความสามารถในการสื่อสารด้วยภาษาอังกฤษ แยกเป็นแบบทดสอบความถนัดทั่วไประดับชาติ (GAT : General Aptitude Test) และเตรียมข้อสอบแพต (PAT : Professional and Academic Aptitude Test) ซึ่งเป็นแบบทดสอบความทางวิชาการ และวิชาชีพซึ่งเป็นการทดสอบความรู้ที่เป็นพื้นฐานที่จะเรียนต่อในวิชาชีพนั้นๆ กับศักยภาพที่จะเรียนวิชาชีพนั้นๆ ให้ประสบความสำเร็จ ทั้งแบบทดสอบแอดและแพตชุดใหม่เป็นแบบทดสอบระดับชาติที่ยังคงรับผิดชอบโดยสถาบันทดสอบทางการศึกษาแห่งชาติ (สทศ.)

ระบบ TCAS การศึกษา 2564 จะเป็นระบบใหม่ คือ มีการสมัครสอบ 5 แบบ ยืนยันสิทธิ์ 4 รอบ และมีการเริ่มจัดสอบแอด/แพตใหม่ ระบบทีแคสปีการศึกษา 2565 จะใช้ระบบ TCAS ปีการศึกษา 2564 โดยจะมีการประชาสัมพันธ์การรวมรอบยืนยันสิทธิ์รอบ 1 และรอบ 2 และเตรียมความพร้อมการปรับข้อสอบ 9 วิชาสามัญ

ระบบ TCAS ปีการศึกษา 2566 จะใช้ระบบ TCAS ใหม่โดยให้สมัคร 5 แบบ ให้ยืนยันสิทธิ์ 3 รอบ โดยใช้เกณฑ์การคัดเลือกรอบ 3 คือ รอบแอดมิชชันใหม่ และใช้คะแนนสอบ GAT/PAT ใหม่ รวมทั้งวิชาสามัญ 9 วิชารูปแบบใหม่เพื่อให้สถาบันอุดมศึกษาต่าง ๆ นำคะแนนไปใช้ประกอบการคัดเลือกบุคคลเข้าศึกษาต่อให้สถาบันอุดมศึกษาในระบบรับตรงผ่านเคลียริงเฮาส์ (Clearinghouse)

ปฏิทินระบบ TCAS ปีการศึกษา 2563 โดยมีการสมัคร 5 รอบ ตั้งแต่วันที่ 2 ธันวาคม 2562 ถึงวันที่ 17 มิถุนายน 2563 ดังรายละเอียดโดยสรุปดังนี้

รอบ 1 รอบแฟ้มสะสมงาน (Portfolio) 2 ธันวาคม 2562-4 กุมภาพันธ์ 2563

รอบ 2 รอบโควตา 6 กุมภาพันธ์-25 เมษายน 2563

รอบ 3 รอบรับตรงร่วมกัน 17 เมษายน 2563-10 พฤษภาคม 2563

รอบ 4 รอบแอดมิชชัน 9 พฤษภาคม 2563-8 มิถุนายน 2563 และ

รอบ 5 รอบรับตรงอิสระ 30 พฤษภาคม 2563-17 มิถุนายน 2563

ประเด็นที่น่าสนใจ ก็คือระบบที่แคสที่พัฒนาและปรับปรุงอย่างต่อเนื่องเพื่อใช้คัดเลือกบุคคลเข้าศึกษาต่อในสถาบันอุดมศึกษาของไทยโดยเฉพาะในปีการศึกษา 2566 และในอนาคต มีความเหมาะสมตามเจตนารมณ์ของกฎหมายการศึกษาแห่งชาติ เป็นธรรม มีประสิทธิภาพและมีความเป็นสากลเพียงใด ในบทความนี้ จึงมีความสนใจที่จะนำเสนอระบบการคัดเลือกนักเรียนเข้าศึกษาต่อในสถาบันอุดมศึกษาของประเทศชั้นนำทางโลกที่ได้รับการยอมรับว่าเป็นประเทศที่มีการจัดการศึกษาระดับมัธยมปลายที่ดีที่สุดและมีระบบการคัดเลือกนักเรียนเข้าศึกษาต่อในสถาบันอุดมศึกษาระดับมาตรฐานโลกซึ่งได้แก่ประเทศสาธารณรัฐสิงคโปร์ ประเทศสาธารณรัฐเกาหลี (เกาหลีใต้) ประเทศฟินแลนด์ และประเทศสหราชอาณาจักร เพื่อเป็นองค์ความรู้และเป็นบทเรียนที่อาจนำไปสู่การปฏิรูปการจัดการศึกษาขั้นพื้นฐาน โดยเฉพาะระดับมัธยมศึกษาตอนปลายของประเทศ รวมทั้งระบบทดสอบและประเมินผลสัมฤทธิ์ทางการศึกษาในระดับสถานศึกษา และระดับชาติตลอดจนระบบที่แคสของไทยให้ก้าวสู่การเป็นแชมป์ในอนาคต

ประเทศที่มีระบบการคัดเลือกนักเรียนเข้าศึกษาต่อในสถาบันอุดมศึกษาดีที่สุดในโลก

1. ประเทศสาธารณรัฐสิงคโปร์ ได้รับการยอมรับว่าเป็นประเทศที่มีระบบการศึกษาที่ดีที่สุดในโลก และจากการเปิดเผยของโออีซีดี (OECD : Organization for Economic Cooperation and Development) (<https://www.oecd.com/>) ปรากฏว่าในปี พ.ศ. 2560 ประเทศสาธารณรัฐสิงคโปร์มีการจัดการศึกษาระดับมัธยมศึกษาตอนปลายที่ดีที่สุดเป็นอันดับหนึ่งของโลกจากทั้งหมด 76 ประเทศในปี พ.ศ. 2558 ประเทศสาธารณรัฐสิงคโปร์มีผลการประเมิน PISA (Programme for International Student Achievement) ด้านคณิตศาสตร์เป็นอันดับ 2 ของโลก ด้านการอ่านเป็นอันดับ 2 ของโลกและด้านวิทยาศาสตร์เป็นอันดับ 3 ของโลก และในปีเดียวกัน OECD ได้ทำการจัดอันดับโรงเรียนจากคะแนนด้านคณิตศาสตร์ และด้านวิทยาศาสตร์ของนักเรียน 76 ประเทศ พบว่าประเทศสาธารณรัฐสิงคโปร์มีโรงเรียนที่ดีที่สุดในโลก (Coughland : 2015)

นักเรียนสิงคโปร์ อายุ 3-6 ปี จะเรียนในระดับอนุบาลและในระดับประถมศึกษา (7-12 ปี) โดยแบ่งเป็นประถมศึกษาต้น 4 ปี และระดับประถมศึกษาตอนปลาย 2 ปี เมื่อเรียนจบระบบประถมศึกษาตอนต้น จะมีการพิจารณาความสามารถ และความสนใจตลอดจนความถนัด เพื่อให้เรียนต่อระดับประถมศึกษาตอนปลายใน 3 แผนการเรียน โดยแผนที่ 1 สำหรับนักเรียนที่มีความสามารถวิชาภาษาอังกฤษและคณิตศาสตร์ แผนที่ 2 สำหรับนักเรียนอยู่ในระดับมาตรฐาน และแผนการเรียนที่ 3 สำหรับนักเรียนที่ไม่ถนัดทางด้านภาษาและคณิตศาสตร์ นักเรียนที่จะสำเร็จการศึกษาระดับประถมศึกษาต้องผ่านการสอบไล่ระดับชาติ เรียกว่า PSLE หรือ Primary Leaving School Examination (สมหวัง พิทยานุวัฒน์และคณะ : 2557) ซึ่งเป็นการสอบระดับประเทศในช่วงเดือนสิงหาคมของทุกปี ผลการสอบใช้ตัดสินในการจบการศึกษา ระดับประถมศึกษา และใช้เป็นเกณฑ์ในการเลือกแผนการเรียนในระดับมัธยมศึกษา ซึ่งมี 3 แผนการเรียน คือ แผนการเรียนวิชาการเร่งรัดหรือพิเศษ แผนการเรียนสายสามัญวิชาการปกติ และแผนการเรียนสายเทคนิค ซึ่งจะเข้าเรียนแผนใดขึ้นอยู่กับผลการสอบ PSLE เมื่อจบชั้นประถมศึกษาปีที่ 6 ในแต่ละแผนการเรียนมีจุดเน้นแตกต่างกัน คือ แผนการเรียนเร่งรัดเป็นหลักสูตร 4 ปี (อายุ 13-16 ปี) เมื่อเรียนจบต้องเข้ารับการทดสอบระดับชาติ GCE 'O' Level (Singapore-Cambridge General Certificate of Education) แผนการเรียนสายสามัญและสายเทคนิค เป็นหลักสูตร 5 ปี เมื่อเรียนจบ 4 ปีแล้วต้องสอบ GCE 'N' Level ถ้ามีผลการสอบดี ก็จะมีโอกาส เรียนปีที่ 5 เพื่อสอบ GCE 'O' Level ได้ด้วย ทั้งนี้ นักเรียนอาจเปลี่ยนสายการเรียนได้ ขึ้นอยู่กับผลการสอบและการประเมินโดยครูและผู้บริหารสถานศึกษา

นักเรียนที่เรียนแผนการเรียนเร่งรัดจะเป็นนักเรียนหัวกะทิ ซึ่งเป็นผู้ได้คะแนนสูงสุด ร้อยละ 10 แรกของผลการสอบ PSLE หลังจากจบโดยการสอบ GCE 'O' Level แล้ว สามารถเข้าเรียนต่อระดับวิทยาลัยขั้นต้น อีก 2 ปี ซึ่งการเรียนต่อระดับนี้จะเป็นการเรียน A-Level ที่นิยมกัน ในประเทศอังกฤษ หลังจากจบโดยการสอบระดับชาติ GCE A-Level (Singapore-Cambridge Advance Certificate of Education 'A' Level) และสามารถนำผลการสอบไปใช้อื่นสมัครเรียนต่อมหาวิทยาลัยได้ นักเรียนต้องเลือกผลสอบ 3-4 วิชาจากจำนวนรายวิชา A-Level ที่โรงเรียนเปิดสอน ผลสอบมี 6 ระดับ คือ A* A B C D E ถ้านักเรียนต้องการเข้ามหาวิทยาลัย ที่มีชื่อเสียงต้องได้เกรด A ทุกวิชา ทั้งนี้ นักเรียนสามารถสอบแก้ตัวได้ สอบปรบเกรดได้

นักเรียนแผนการเรียนปกติใช้เวลา 5 ปี เมื่อเรียนถึงปีที่ 4 ต้องสอบการทดสอบระดับชาติ GCE 'N' Level เมื่อสอบได้จะเรียนต่อชั้นปีที่ 5 แล้วสอบระดับชาติ GCE 'O' Level แล้วเรียนต่อวิทยาลัยขั้นต้น 2 ปี หลังจากจบแล้วสามารถสอบ GCE A-Level เพื่อนำผลการสอบ GCE A-Level ไปใช้อื่นสมัครเรียนต่อมหาวิทยาลัยได้

นักเรียนที่เรียนสายเทคนิคใช้เวลา 5 ปี เมื่อเรียนจบ 4 ปี ต้องสอบ GCE ‘N’ Level แล้วจึงจะสามารถสอบ GCE ‘O’ Level เมื่อจบชั้นปีที่ 5 สำหรับนักเรียนที่มีผลการสอบ GCE ‘O’ Level ดี สามารถศึกษาต่อในวิทยาลัยชั้นต้น โพลีเทคนิคหรือสถาบันทางเทคนิคได้ ส่วนนักเรียนสอบไม่ผ่านในชั้นปีที่ 5 จะมีทางเลือกศึกษาต่อด้านเทคนิค และอาชีวศึกษาในสถาบันอาชีวศึกษาได้ แต่ถ้าใครต้องการเรียนต่อมหาวิทยาลัยก็ต้องสอบ A-Level นำไปยื่นสมัครเข้าเรียนต่อมหาวิทยาลัยได้

เป็นที่น่าสังเกตว่าประเทศสาธารณรัฐสิงคโปร์ไม่มีระบบการสอบคัดเลือกนักเรียนเข้าศึกษาในสถาบันอุดมศึกษาโดยเฉพาะ แต่ประเทศสาธารณรัฐสิงคโปร์มีระบบการทดสอบระดับชาติ คัดกรองนักเรียนตั้งแต่ระดับประถมศึกษาและมัธยมศึกษาซึ่งทำให้นักเรียนประเมินตนเองได้ว่าสนใจอะไร ชอบวิชาอะไรน่าจะเรียนต่อทางด้านไหนที่น่าสนใจเป็นอย่างยิ่งก็คือ ระบบการศึกษาของประเทศสาธารณรัฐสิงคโปร์ได้ออกแบบแผนการเรียนให้นักเรียนได้เรียนตามความถนัดและความสามารถของนักเรียน ตั้งแต่ระดับประถมศึกษาตอนปลายซึ่งมี 3 แผน และในระดับมัศึกษามีแผนการเรียน 3 แผน และมีการใช้มาตรฐานระดับนานาชาติเป็นมาตรฐานการศึกษาของประเทศสาธารณรัฐสิงคโปร์โดยนักเรียนเข้ารับการทดสอบ PISA และ TIMSS เป็นประจำตลอดจนมีการบูรณาการการสอบคัดเลือกเข้าศึกษาต่อในระดับอุดมศึกษาอยู่ในการทดสอบระดับชาติระดับการศึกษาขั้นพื้นฐาน ดังนั้นจึงถือได้ว่าประเทศสาธารณรัฐสิงคโปร์เป็นประเทศที่ระบบการคัดเลือกนักเรียนเข้าศึกษาต่อที่มีประสิทธิภาพสูงมาก (สมหวัง พิธิยานุวัฒน์ และจิราทัศน์ รัตนมณีฉัตร : 2559; <https://www.oecd.org>; <http://www.learningcurve-th.com/singapore/>; <http://seab.gov.sg/about-us/about-SEAB.html>และ<http://campus.campus-star.com/ed...>)

2. ประเทศสาธารณรัฐเกาหลี (เกาหลีใต้) เป็นอีกประเทศหนึ่งในเอเชียที่ได้รับการยอมรับว่ามีระบบการศึกษาดีเยี่ยม มีระบบการเรียนการสอนที่เข้มข้น และค่อนข้างมีความกดดันมากจากการเปิดเผยของ OECD พบว่าประเทศสาธารณรัฐเกาหลี (เกาหลีใต้) เป็นประเทศที่มีระบบการจัดการศึกษาระดับมัธยมศึกษาตอนปลายเป็นอันดับที่ 3 ของโลก (<https://www.oecd.com/>)

เด็กเกาหลีใต้จะต้องเข้าเรียนในระบบโรงเรียน เมื่ออายุ 6 ขวบถึง 15 ปี ในโรงเรียนประถมศึกษา 6 ปี มัธยมศึกษาตอนต้น 3 ปี ซึ่งเป็นการศึกษาภาคบังคับโดยมีการทดสอบระดับชาติระดับมัธยมศึกษาตอนปลาย 3 ปี (อายุ 15-18 ปี) แบ่งออกเป็นสายสามัญกับสายอาชีพ การเข้าเรียนมัธยมศึกษาตอนปลายอาจมีการ สอบเข้าหรือพิจารณาจากผลการเรียนในระดับมัธยมศึกษาตอนต้น ถ้าจบมัธยมศึกษาตอนปลาย สายอาชีพจะได้รับประกาศนียบัตรวิชาชีพ ส่วนใหญ่นักเรียนจะจบมัธยมศึกษาสายสามัญ มีส่วนน้อยเพียงร้อยละ 30 ที่เรียนสายอาชีพ เมื่อนักเรียนจบชั้นมัธยมศึกษาตอนปลายจะสอบแบบทดสอบความถนัดทางการเรียนระดับอุดมศึกษา (College Scholastic Aptitude Test : CSAT) เพื่อเข้าเรียนต่อในระดับมหาวิทยาลัย (<http://www.ncee.org/programs-affiliate/center-on-international-education-benchmarking/top-performing-countries/south-korea-overview/south-korea-instructional-systems/>)

ในประเทศเกาหลีได้มีการประเมินระดับชาติว่าด้วยสัมฤทธิ์ผลทางการศึกษา (National Assessment of Educational Achievement : NAEA) ซึ่งดำเนินการโดยสถาบันว่าด้วยหลักสูตรและการประเมินผลแห่งประเทศไทยเกาหลี (Korean Institute of Curriculum and Evaluation : KICE) เป็นการทดสอบระดับชาตินักเรียนเกรด 6 เกรด 9 และเกรด 10 ในวิชาภาษาเกาหลี สังคมศึกษา คณิตศาสตร์ วิทยาศาสตร์ และภาษาอังกฤษ ทั้งนี้มีวัตถุประสงค์เพื่อประเมินความก้าวหน้าและสัมฤทธิ์ผลทางการศึกษา เพื่อกำกับคุณภาพของการศึกษา และความเหมาะสมของหลักสูตรของชาติ นอกจากนี้มีการสอบ DTBS (Diagnostic Test for Basic Skill) เพื่อวินิจฉัยว่านักเรียนเกรด 3 มีทักษะพื้นฐานเพียงพอหรือไม่ในด้านการอ่านการเขียน และเลขคณิต นอกจากนี้ในการทดสอบระดับนานาชาตินักเรียนเกาหลีใต้จะเข้ารับการทดสอบ PISA และ TIMSS เป็นประจำ

ในการสอบคัดเลือกนักเรียนเข้าศึกษาต่อในระดับอุดมศึกษามีระบบการสอบ CSAT หรือแบบทดสอบความถนัดทางการเรียนระดับอุดมศึกษา (College Scholastic Aptitude Test) โดยผู้เข้าสอบต้องเป็นผู้สำเร็จการศึกษาระดับมัธยมศึกษาปีที่ 6 แล้วโดยสามารถเข้าทดสอบซ้ำได้ ในแต่ละปีจะมีการทดสอบ CSAT เพียงครั้งเดียวเท่านั้นโดยจะมีการสอบทุกวิชาภายในวันเดียวทั้งนี้จะจัดการสอบในเดือนพฤศจิกายนของทุกปี

การสอบ CSAT ประกอบด้วยการสอบวิชาหลัก ๆ คือเน้นวิชาบังคับได้แก่ภาษาประจำชาติ คณิตศาสตร์ ภาษาอังกฤษและประวัติศาสตร์เกาหลี ส่วนวิชาเลือกได้แก่สังคมศึกษา วิทยาศาสตร์ที่สามารถเลือกสอบได้มากที่สุดหมวดละ 2 วิชาและภาษาต่างประเทศที่สามารถเลือกสอบได้ 1 วิชา ทั้งนี้ นักเรียนจะรู้ว่าสอบเข้าเรียนต่อสาขา คณะวิชาและมหาวิทยาลัยจะได้เลือกสอบวิชาได้ถูกต้องครบถ้วนเมื่อเข้าสอบ CSAT จะรู้ผลภายในวันเดียวมีการเฉลยข้อสอบทันทีหลังสอบเสร็จและสามารถนำคะแนนไปยื่นต่อมหาวิทยาลัยได้ ถ้าสอบไม่ผ่านก็สอบได้ในปีต่อไป ทั้งการสอบ CSAT ไม่จำกัดอายุผู้เข้าสอบผลการสอบ CSAT แบ่งเป็นเกรดระดับ 1-9 ถ้าผู้สอบได้เกรดต่ำกว่า 5 ถือว่าเป็นผู้มีความถนัดในการเรียนระดับอุดมศึกษาค่อนข้างสูงโดยได้เกรด 1 หมายความว่ามีความถนัดในการเรียนระดับอุดมศึกษามากที่สุด

CSAT เป็นการสอบกลางที่มีความเป็นมาตรฐานได้รับการยอมรับเป็นการทดสอบมาตรฐานที่เข้มงวดมากที่สุดในโลกเพื่อให้เกิดความยุติธรรมและโปร่งใสมากที่สุดเรื่องน่ารู้เกี่ยวกับ CSAT ที่นักเรียนเข้าสอบพึงรู้มี 10 ประการ คือ (www.bbc.com และ www.tomatoprime.com)

1. CSAT คือการสอบคัดเลือกเข้าศึกษาต่อในมหาวิทยาลัยโดยการทดสอบความถนัดในการเรียนระดับอุดมศึกษาโดยการจัดขึ้นในเวลา 1 วัน (8.40 น.-17.40 น.) โดยต้องเข้าสอบก่อนเวลา 8.10น. และจัดสอบปีละ 1 ครั้งเท่านั้นในเดือนพฤศจิกายนของทุกปี

2. การสอบ CSAT เป็นการชี้วัดว่า นักเรียนจะได้เรียนในมหาวิทยาลัยชั้นนำหรือไม่ซึ่งจะมีผลต่อโอกาสในการทำงานบริษัทต่างๆเมื่อสำเร็จการศึกษาแล้วทั้งนี้ถือว่าการสอบ CSAT เป็นการสอบที่กำหนดชะตาชีวิตในอนาคตของนักเรียน

3. ห้ามเครื่องบินขึ้น-ลงขณะสอบ 35 นาที ทั้งนี้ต้องการให้นักเรียนที่กำลังทำข้อสอบในวิชาทักษะการฟังภาษาอังกฤษสามารถฟังภาษาอังกฤษได้อย่างชัดเจนโดยกระทรวงคมนาคมและโครงสร้างพื้นฐานทางบก ของประเทศได้ออกประกาศกำหนดเวลาห้ามบินขึ้น-ลงเป็นเวลา 35 นาที ส่วนเครื่องบินที่กำลังบินอยู่กลางอากาศในขณะนั้นไม่สามารถลดระดับเพดานเครื่องบินลงมาที่ความสูงต่ำกว่า 3 กิโลเมตรได้

4. เจ้าหน้าที่ที่ออกข้อสอบและการตรวจข้อสอบจะถูกกักตัวไว้ในสถานที่ที่ไม่สามารถเปิดเผยให้ผู้อื่นรู้ได้เพื่อเป็นการป้องกันการรั่วไหลของข้อสอบขณะถูกกักตัวไว้จะถูกตัดขาดจากการสื่อสารทุกอย่างและไม่สามารถเดินทางออกมาจากสถานที่นั้นได้

5. นักเรียนจะไม่ได้รับอนุญาตให้สอบภายหลังอย่างเด็ดขาดซึ่งหมายความว่านักเรียนทุกคนจะต้องเดินทางไปยังศูนย์สอบให้ทันกำหนดเวลาก่อนประตูห้องสอบจะปิด (ก่อนเวลา 08.10 น.)

6. เพิ่มบริการขนส่งสาธารณะในวันสอบ CSAT เพื่อให้มั่นใจว่านักเรียนทุกคนจะสามารถเดินทางไปศูนย์สอบได้ทันเวลา ทั้งนี้มีการเตรียมยานพาหนะทางภาครัฐและเอกชนเพื่อไว้ขนส่งนักเรียนในกรณีฉุกเฉิน

7. จัดเจ้าหน้าที่ตำรวจและเจ้าหน้าที่ดับเพลิงหลายหมื่นนายในการให้บริการขนส่งนักเรียนไปยังศูนย์สอบ

8. ในวันสอบ CSAT ตลาดหุ้นเปิดช้ากว่าปกติ 1 ชั่วโมงเพื่อป้องกันความวุ่นวายที่อาจเกิดขึ้นจากการเดินทางของนักเรียนในช่วงวันสอบ

9. สมาคมธนาคารเกาหลีเลื่อนเวลาเปิด-ปิดจากเดิม 9.00 น. ถึง 16.00 น. เป็น 10.00 น. ถึง 17.00 น. นอกจากนี้บริษัทเอกชนอีกหลายบริษัทเลื่อนเวลาให้พนักงานเดินทางมาทำงานสายขึ้นอีกเช่นกัน

10. การสอบ CSAT ที่สนใจของสื่อมากที่สุดในวันสอบ CSAT นอกจากมีนักเรียนรุ่นน้องมาให้กำลังใจรุ่นพี่ที่เข้าสอบแล้วยังมีผู้สื่อข่าวจากสำนักต่าง ๆ โดยมีผู้สื่อข่าวไปยืนหน้าศูนย์สอบเพื่อรอทำข่าวตั้งแต่เช้าถึงตอนเย็นพร้อมรายงานข่าวอัปเดตกันทุกนาทีที่ทุกวิชาว่าตอนนี้กำลังสอบวิชาอะไรสอบเสร็จไปกี่วิชาแล้ว วิชาไหนสอบยากขึ้นและวิชาไหนปีนี้สอบง่ายที่สุด

จะเห็นได้วาระบบการสอบคัดเลือกเข้าศึกษาต่อในมหาวิทยาลัยของประเทศสาธารณรัฐเกาหลี (เกาหลีใต้) ตามระบบ CSAT เป็นการสอบที่มีความสำคัญมีความเป็นมาตรฐานมีความร่วมมือกันเพื่อให้เด็กทุกคนสามารถเข้าสอบได้ทันเวลาและยังมีการวางแผนที่ดีมีความยุติธรรมกับผู้สมัครสอบมีการสอบเพียง 1 วันและทราบผลได้ทันทีนับว่าเป็นระบบการสอบที่มีประสิทธิภาพสูงสมกับเป็นประเทศที่ได้รับการยอมรับมีระบบการศึกษาดีเยี่ยมระดับโลก (<https://campus.campus-star.com/ed...>)

3. ประเทศฟินแลนด์ เป็นประเทศในกลุ่มสแกนดิเนเวียที่ได้รับการยอมรับว่าเป็นประเทศที่มีระบบการศึกษาและการฝึกอบรมที่ดีที่สุดในโลกและเป็นประเทศที่มีความสุขมากที่สุดในโลก (seeme.me/ch/goodmorningthailand) จากการเปิดเผยของ OECD ในปี พ.ศ. 2560 ประเทศฟินแลนด์เป็นประเทศที่มีการจัดการศึกษาระดับมัธยมศึกษาตอนปลายดีที่สุดเป็นอันดับ 6 ของโลก (www.oecd.com) ประเทศฟินแลนด์ใช้มาตรฐานนานาชาติเป็นมาตรฐานการศึกษาของชาติโดยการให้นักเรียนเข้ารับการสอบ PICA และ TIMSS เป็นประจำ

การศึกษาก่อนประถมศึกษาและการศึกษาขั้นพื้นฐานเป็นการศึกษาภาคบังคับ การศึกษาก่อนประถมศึกษาเรียน 1 ปี สำหรับเด็กต่างชาติดอายุ 6 ขวบ หรือเรียน 2 ปี สำหรับเด็กที่มีความต้องการพิเศษที่ได้เข้าเรียนตั้งแต่ 5 ขวบ การศึกษาขั้นพื้นฐานภาคบังคับ 9 ปี เป็นระบบสายเดียวจัดการเรียนการสอนในโรงเรียนแบบประสมให้ครูประจำชั้นสอนชั้นปีที่ 1-6 มีครูประจำวิชาสอนในชั้นปีที่ 7-9 ในระดับมัธยมศึกษาตอนปลายจัดการเรียนการสอนเป็นสายสามัญและสายอาชีพศึกษา

ในการจบระดับการศึกษาระดับ 2 หรือมัธยมศึกษาตอนปลาย (เกรด 10-12) นักเรียนจะต้องสอบ FME (The Finnish Matriculation Examination) (www.ylioppilastutkinto.fi) และเป็นการสอบเข้าศึกษาในระดับอุดมศึกษาในเวลาเดียวกันข้อสอบครอบคลุมเนื้อหาสาระตามหลักสูตรระดับชั้นมัธยมศึกษาตอนปลาย ดังนั้นการสอบ FME จึงเป็นการสอบเพื่อสำเร็จการศึกษามัธยมศึกษาปีที่ 6 และใช้คะแนนไปยื่นสมัครเข้าเรียนต่อในมหาวิทยาลัยโดยไม่ต้องมีการสอบคัดเลือกเข้าศึกษาต่อในระดับอุดมศึกษาอีกเป็นการสอบคุณภาพของนักเรียนเกรด 12 ที่กำลังจะจบการศึกษาระดับ 2 (เกรด 10-12 หรือ ม.4-ม.6) และจะเข้าเรียนต่อในระดับอุดมศึกษาการสอบ FME ถือเป็นการสอบครั้งสุดท้าย (Final Examination) ซึ่งนักเรียนจะมีสิทธิเข้าสอบ FME ได้นั้นจะต้องเรียนให้ครบทุกวิชาและทำคะแนนให้ผ่านเกณฑ์ของรายวิชานั้น ๆ เมื่อสอบ FME เสร็จนักเรียนก็นำคะแนนเกรดไปยื่นสมัครเรียนต่อในมหาวิทยาลัยได้โดยจะไม่นำคะแนนเกรดทุกวิชามาคิดแต้มเฉลี่ยสะสม (GPAX) เพราะมหาวิทยาลัยจะพิจารณาคะแนนเกรดในวิชาที่สาขาวิชากำหนดเท่านั้นในบางมหาวิทยาลัยก็มีการพิจารณา GPAX เฉพาะผลการสอบชั้นมัธยมศึกษาปีที่ 6 หรือเกรด 12 เท่านั้นมหาวิทยาลัยส่วนใหญ่จะมีการจัดสอบอีกรอบโดยมหาวิทยาลัยออกข้อสอบเองและสุดท้ายก็ใช้คะแนนเกรดจากการสอบ FME มารวมกับคะแนนเกรดที่มหาวิทยาลัยจัดสอบใช้ตัดสินชะตาชีวิตของนักเรียนในการเข้าเรียนในมหาวิทยาลัย อย่างไรก็ตามบางมหาวิทยาลัยไม่มีการสอบครั้งที่สองคือพิจารณาจากคะแนนเกรดจากการสอบ FME อย่างเดียว

การจัดสอบ FME จัดสอบ 2 ครั้งในช่วงฤดูใบไม้ผลิ (กุมภาพันธ์-มีนาคม) และฤดูใบไม้ร่วง (กรกฎาคม-กันยายน) ไม่มีการสอบติด ๆ กันโดยทั่วไปสอบวันเว้น 2-3 วันทำให้นักเรียนมีเวลาพักและเตรียมสอบได้อย่างเต็มที่ FME ประกอบด้วย การสอบ 4 วิชา คือ

1. ภาษาแม่ คือ ภาษาฟินแลนด์ และ ภาษาสวีเดน เป็นวิชาบังคับที่ทุกคนต้องสอบ
2. ภาษาต่างประเทศ ให้เลือกสอบ 1 วิชา ตามความถนัดส่วนใหญ่ นักเรียนมักจะเลือกสอบ วิชาภาษาอังกฤษ ฝรั่งเศส เยอรมัน อิตาลี และ ภาษาพื้นเมือง ในการสอบจะให้เลือกทำข้อสอบระดับขั้นสูง (Advance) สำหรับนักเรียนที่เชี่ยวชาญมากกับระดับปานกลาง (Intermediate) สำหรับนักเรียนที่มีพื้นฐานปานกลาง
3. คณิตศาสตร์ บังคับให้ทุกคนต้องสอบคณิตศาสตร์ แต่เลือกกระดัดการสอบได้ กล่าวคือ ถ้านักเรียนเลือกทำข้อสอบวิชาภาษาต่างประเทศระดับปานกลาง ต้องเลือกทำข้อสอบคณิตศาสตร์ระดับยาก ถ้าเลือกทำข้อสอบภาษาต่างประเทศยาก สามารถเลือกทำข้อสอบคณิตศาสตร์ระดับปานกลางได้
4. วิชาเลือก เลือกสอบได้อย่างน้อย 1 วิชา ใน 13 วิชา แล้วแต่นักเรียนต้องการศึกษาต่อในคณะวิชาอะไร วิชาที่ให้เลือกสอบ เช่น สังคมศาสตร์ ประวัติศาสตร์ จริยศาสตร์ ปรัชญา จิตวิทยา

ในปัจจุบันประเทศฟินแลนด์มีการสอบผ่านระบบคอมพิวเตอร์ที่เรียกว่า Digital Matriculation Examination ทำให้การสอบเข้ามหาวิทยาลัยของนักเรียนฟินแลนด์มีความรวดเร็วสะดวกสบายยิ่งขึ้น

ในการนี้สอบไม่ผ่านนักเรียนสามารถสอบแก้ตัวได้ผลสอบคะแนนไม่ดีสามารถสอบใหม่ได้แต่ต้องสอบให้ครบทุกวิชาภายใน 3 ฤดูกาลหรือถ้าใครรับจะสมัครสอบทุกวิชาให้เสร็จภายในครั้งเดียวก็ทำได้ อีกทั้งนักเรียนสามารถสมัครสอบเพื่อทำคะแนนเกรดใหม่ได้จนเป็นที่พอใจ (<https://yle.fi/uutiset/usasto/news/>)

นักเรียนที่ไม่ได้เรียนมัธยมศึกษาตอนปลายสายสามัญแต่เรียนสายอาชีวศึกษาไม่ต้องสอบ FME เมื่อเรียนจบจากโรงเรียนอาชีวศึกษาแล้วประสงค์จะเรียนต่อมหาวิทยาลัยก็ไปสอบข้อสอบที่มหาวิทยาลัยออกเองได้เลย ([ylioppilastutkinto.fi](https://yle.fi/uutiset/usasto/news/); www.reddit.com และ <https://campus.campus-star.com/ed...>)

4. ประเทศสหราชอาณาจักร เป็นประเทศในทวีปยุโรปที่ได้รับการยอมรับว่าเป็นประเทศที่มีการจัดการศึกษาที่มีคุณภาพและเป็นต้นแบบที่หลายประเทศได้นำไปประยุกต์ใช้ โดยเฉพาะในด้านระบบการประเมินผลสัมฤทธิ์ทางการเรียนในการสอบ ‘O’ Level และ ‘A’ Level ในระดับการศึกษาขั้นพื้นฐานและนำไปใช้ในการคัดเลือกนักเรียนเข้าศึกษาต่อในระดับอุดมศึกษาจากการเปิดเผยของ OECD ในปี พ.ศ. 2560 ประเทศสหราชอาณาจักร (อังกฤษ) เป็นประเทศที่มีระบบการศึกษาระดับมัธยมศึกษาตอนปลายที่ดีที่สุดเป็นอันดับที่ 20 ของโลก (www.oecd.com) นอกจากนี้ประเทศสหราชอาณาจักรก็มีการเทียบเคียงมาตรฐานการศึกษาของประเทศกับมาตรฐานของนานาชาติโดยการรับการทดสอบ PISA และ TIMSS ด้วย

การจัดการศึกษาในประเทศสหราชอาณาจักรจะมีความแตกต่างกันในแต่ละแคว้นในบทความนี้จะกล่าวถึงเฉพาะอังกฤษเป็นสำคัญโดยแบ่งการศึกษาออกเป็น 3 ระดับ คือ ระดับก่อนประถมศึกษา (อายุ 2-5 ปี) ระดับประถมศึกษาและมัธยมศึกษาแบ่งออกมาเป็น 4 ระดับขั้น คือ ระดับที่ 1 สำหรับเด็กอายุ 5-7 ปี ระดับที่ 2 สำหรับเด็กอายุ 7-11 ปี ระดับที่ 3 สำหรับเด็กอายุ 11-14 ปี และระดับที่ 4 สำหรับเด็กอายุ 14-16 ปี และระดับอุดมศึกษาซึ่งเป็นการจัดการศึกษาหลังมัธยมศึกษาจำแนกเป็น มหาวิทยาลัย วิทยาลัย และสถาบันวิชาชีพ

ประเทศสหราชอาณาจักรมีการสอบวัดมาตรฐานความรู้กับหน่วยสอบกลางเพื่อให้ได้ประกาศนียบัตรการเรียนจบระดับมัธยมศึกษา (GCSE) ซึ่งแบ่งระดับสามัญหรือมาตรฐาน (O-Level) และระดับสูง (A-Level) GCE O level (General Certificate of Education Ordinary level) หรือประกาศนียบัตรทางการศึกษาระดับสามัญเป็นคุณวุฒิทางการศึกษาระดับมัธยมศึกษาในระบบการศึกษาของประเทศสหราชอาณาจักรสำหรับนักเรียนเกรด 10-11 ที่มีอายุประมาณ 16 ปี ที่สอบผ่านการสอบระดับสามัญให้ได้ประกาศนียบัตรทางการศึกษาระดับสามัญในปี พ.ศ. 2531 ได้เปลี่ยนชื่อคุณวุฒิและการสอบเป็น GCSE หรือ General Certificate of Secondary Education สำหรับนักเรียนในประเทศสหราชอาณาจักรการสอบระดับสามัญเป็นการสอบรายวิชาและผู้สอบเลือกสอบได้ไม่เกิน 14 วิชา เช่น วิทยาศาสตร์ ภาษาอังกฤษ คณิตศาสตร์ เศรษฐศาสตร์ บัญชี ประวัติศาสตร์ สังคมศึกษา ศาสนา คอมพิวเตอร์และเทคโนโลยี ศิลปะ สถาปัตยกรรม ดนตรี แฟชั่น การแสดงผลการสอบแบ่งเป็นระบบตัวอักษร 7 ระดับคือ A, B, C, D, E, F, และ G ผู้สอบได้ระดับ C ครบ 5 วิชา ซึ่งถือว่าสอบผ่านในระดับ จีซีเอสอี หรือ จีซีอีโอ เลเวล ตั้งแต่ พ.ศ. 2560 มีการเปลี่ยนระบบเกรดใหม่เป็นระบบตัวเลข 9 ระดับคือ ระดับ 1 ถึง 9 และ U คือ ไม่ผ่านโดยเกรด 9 เป็นเกรดสูงสุดและเกรด 1 เป็นเกรดต่ำสุดโดยเกรด U ถือว่าไม่ผ่านจีซีเอสอี ทั้งนี้มีการสอบระดับสามัญ 2 รอบต่อปี ช่วงแรกสอบในเดือนพฤษภาคมถึงมิถุนายน และอีกช่วงสอบในเดือนตุลาคมถึงพฤศจิกายน การสอบระดับสามัญ หรือ โอเลเวล เป็นการสอบระดับชาติที่ใช้แทนการสอบระดับโรงเรียนตั้งแต่ พ.ศ. 2494 เป็นการสอบที่ถือว่าสำเร็จการศึกษาระดับมัธยมศึกษา สามารถเข้าศึกษาต่อในวิทยาลัยอาชีวศึกษา หรือ สถาบันอุดมศึกษาบางแห่ง หากต้องการศึกษาต่อในมหาวิทยาลัยส่วนใหญ่ต้องการคุณวุฒิขั้นสูง หรือ A-Level ซึ่งต้องศึกษาต่อในระดับเตรียมอุดมศึกษาอีก 2 ปี (สำนักงานราชบัณฑิตยสภา: 2561)

สำนักงานราชบัณฑิตยสภา (2559: 76-71) ได้นิยามประกาศนียบัตรทางการศึกษาระดับสูง (GCE 'A' Level หรือ General Certificate of Education Advanced Level) หมายถึง คุณวุฒิทางการศึกษาระดับเตรียมอุดมศึกษาในระบบการศึกษาของสหราชอาณาจักร สำหรับนักเรียนเกรด 12-13 ที่มีอายุประมาณ 18 ปีขึ้นไป ที่สอบผ่านการสอบขั้นสูงซึ่งเป็นการสอบระดับชาติสำหรับนักเรียนที่จบหลักสูตรเตรียมอุดมศึกษาผู้สอบสามารถเลือกสอบได้ 3-5 วิชา เนื่องจากโดยทั่วไปมหาวิทยาลัยจะกำหนดให้ผู้สมัครเข้าศึกษาต้องมีผลการสอบ A-Level

จำนวน 3 วิชาที่แตกต่างกันไปตามสาขาวิชาที่แต่ละมหาวิทยาลัยกำหนดผลการสอบ A-Level มี 6 ระดับ คือ A*, A, B, C, D และ E มหาวิทยาลัยส่วนใหญ่จะพิจารณารับนักเรียนที่มีผลการสอบระดับ C ขึ้นไปมหาวิทยาลัยชั้นนำอาจรับเฉพาะผู้ที่ได้คะแนนระดับ A*, A และ B โดยคุณสมบัติขั้นต่ำทั่วไปจะพิจารณาจากผลการสอบทั้งระดับ O-Level และ A-Level คือ พิจารณาจากผลการสอบ O-Level 3 วิชา และ A-Level 2 วิชาหรือผลการสอบวิชา O-Level 1 วิชา และเอเลเวล 3 วิชา ในกรณีมหาวิทยาลัยชั้นนำจะพิจารณาจากการรับนักเรียนเข้าศึกษาจากผลการสอบ A-Level เท่านั้น

ศูนย์ให้บริการการคัดเลือกนักศึกษาเข้าศึกษาต่อในระดับอุดมศึกษาของสหราชอาณาจักร (Universities and college Admissions Service : UCAS) (<http://en.wikipedia.org/wiki/UCAS> และ www.UCAS.com) จัดตั้งขึ้นเมื่อ พ.ศ. 2536 โดยการยุบรวมองค์กร 3 แห่ง คือ UCCA, PCAS และ SCUEUCAS เป็นองค์กรการกุศลไม่หวังผลกำไรประกอบด้วยสถาบันอุดมศึกษาที่เป็นสมาชิกทั้งหมดบริหารในรูปแบบบริษัทธุรกิจและบริหาร โดยคณะกรรมการบริหาร โดยมีผู้อำนวยการของ UCAS เป็นผู้บริหารประจำ ผลงานของ UCAS คือการคัดเลือกนักเรียนเข้าศึกษาต่อในสถาบันอุดมศึกษาในสหราชอาณาจักรซึ่งมีมากกว่า 300 สถาบัน และมากกว่า 5,000 หลักสูตร โดยที่สถาบันอุดมศึกษาแต่ละแห่งกำหนดคุณสมบัติผู้ที่จะเข้าศึกษาต่อในแต่ละหลักสูตรคุณสมบัติดังกล่าวรวมถึงคะแนนสอบคัดเลือก อาทิ BMAT (Bio Medical Admission Test) สำหรับผู้ประสงค์จะเข้าศึกษาต่อสาขาแพทยศาสตร์และสาขาสัตวแพทย์ CAT (Classic Admission Test) สำหรับผู้ประสงค์จะศึกษาต่อในมหาวิทยาลัยออกฟอร์ด HAT (History Aptitude test) สำหรับผู้ประสงค์ได้จะเข้าศึกษาในสาขาประวัติศาสตร์ในมหาวิทยาลัยออกฟอร์ด แบบทดสอบเหล่านี้อาจดำเนินการโดยมหาวิทยาลัยหรือกลุ่มของสาขาวิชาในมหาวิทยาลัยแบบทดสอบส่วนใหญ่เป็นแบบทดสอบมาตรฐานที่พัฒนาขึ้นเฉพาะกิจซึ่งมีใช้หน้าที่ของ UCAS เป็นผู้พัฒนา

UCAS ของประเทศสหราชอาณาจักรคล้ายกับ TCAS ของไทยนักเรียนที่สำเร็จการศึกษาระดับมัธยมศึกษาสามารถสมัครเข้าเรียนต่อในคณะสาขาวิชาและมหาวิทยาลัยที่ต้องการเลือกได้ 5 อันดับ ผู้สมัครส่วนใหญ่ต้องเรียนจบการศึกษาหลักสูตร A-Level โดยจะต้องแสดง

1) ใบรับรองผลการเรียน 2) ใบรับรองจากสถานศึกษา 3) คะแนนสอบ IELTS/TOEFL เลือกมาอย่างใดอย่างหนึ่ง 4) จดหมายแนะนำหรือรับรองจากอาจารย์หรือบุคคลที่เชื่อถือได้ 5) ข้อความแสดงถึงวัตถุประสงค์ในการศึกษาต่อหรือเรียงความ เพื่อให้ผู้คัดเลือกใบสมัครรู้จักผู้สมัครได้มากขึ้น 6) ประวัติของผู้สมัครโดยยื่นสมัครผ่านระบบ UCAS อปโหลดเอกสารดังกล่าวลงไปยัง website www.ucas.com

ประเทศ	อันดับโลกของการศึกษา ระดับมัธยมศึกษาตอนปลาย (จาก 76 ประเทศ)	การทดสอบระดับชาติ ระดับการศึกษาขั้นพื้นฐาน	ระบบการคัดเลือก	การพัฒนาการคัดเลือก	การสอบเพิ่มเติม
สาธารณรัฐสิงคโปร์	1 (จาก 76 ประเทศ)	PSLE, GCE 'O' Level GCE 'N' Level, GCE 'A' Level	ไม่มีระบบการคัดเลือกโดยการสอบ เข้าเรียนในมหาวิทยาลัยใช้ทดสอบ GCE 'A' Level	สถาบันอุดมศึกษาจะ พิจารณาผลการ สอบ A-Level ในวิชา ต่างๆที่สาขาวิชา กำหนด	ไม่มี
สาธารณรัฐเกาหลี (เกาหลีใต้)	3	DTBS, NAEA	มีระบบการสอบ CSAT เป็นระบบ การสอบนักเรียนที่เรียนต่อใน สถาบันอุดมศึกษาสอบปีละครั้ง สอบเสร็จเฉลี่ยข้อสอบทันทีที่รู้ผล สอบภายใน 1 วัน นำไปยื่นสมัคร เรียนต่อมหาวิทยาลัย สำหรับ ผู้จบม.6 แล้วไม่จำกัดอายุ	ผลสอบในวิชาบังคับ คือภาษาเกาหลี, คณิตศาสตร์, ภาษา อังกฤษและ ประวัติศาสตร์เกาหลี ผลสอบในวิชาเลือก ได้แก่สังคมศึกษา/ วิทยาศาสตร์เลือกได้ มากที่สุดหมวดละ 2 วิชาและภาษาต่าง ประเทศเลือกได้ 1 วิชา	ไม่มี
ประเทศ	อันดับโลกของการศึกษา ระดับมัธยมศึกษาตอนปลาย (จาก 76 ประเทศ)	การทดสอบระดับชาติ ระดับการศึกษาขั้นพื้นฐาน	ระบบการคัดเลือก	การพัฒนาการคัดเลือก	การสอบเพิ่มเติม
ฟินแลนด์	6 (จาก 76 ประเทศ)	FME	ไม่มีการคัดเลือกโดยการสอบ นักเรียนสายสามัญใช้ผลการสอบ FME	สถาบันอุดมศึกษาจะ พิจารณาจากผลการ สอบ FME และการ สอบเพิ่มเติม(ถ้ามี) ซึ่ง มหาวิทยาลัยส่วนใหญ่ ได้มีการสอบเพิ่มเติม โดยมหาวิทยาลัย เป็นผู้ออกข้อสอบเอง	ในการนี้นักเรียนจบสาย อาชีวศึกษาไม่มีคะแนน เกรด FME ต้องสอบข้อสอบ ที่มหาวิทยาลัยออกเอง
สหราชอาณาจักร	20	GCE/GCSE O-Level GCE/GCSE A-Level	ไม่มีการคัดเลือกโดยการสอบแต่มี ระบบการรับสมัครกลาง UCAS ใช้ ผลสอบ GCE/GCSE O-Level และ GCE/GCSE A-Level	สถาบันอุดมศึกษาจะ พิจารณาผลการ สอบ A-Level และ O Level กับผลการ ทดสอบเฉพาะทาง สาขาวิชาชีพ	มหาวิทยาลัยหรือกลุ่มสาขา วิชาในมหาวิทยาลัยร่วมกัน พัฒนาแบบทดสอบ มาตรฐานที่ใช้เฉพาะกิจไม่ เกี่ยวกับ UCAS

เมื่ออัปโหลดแล้วก็รอผลการพิจารณาว่ามีมหาวิทยาลัยใดตอบรับใบสมัครมาบ้าง โดยสมัครผ่านระบบ UCAS ถ้าติดทั้ง 5 อันดับผู้สมัครก็เลือกจะไปเรียนต่อที่มหาวิทยาลัยใด บางมหาวิทยาลัยก็ตอบรับแบบมีเงื่อนไขเช่นให้เรียนภาษาอังกฤษ 1 ปีก่อนแล้วจึงจะสามารถเข้าเรียนระดับปริญญาตรีได้หากยังสอบไม่ผ่านและไม่มีมหาวิทยาลัยใดตอบรับก็สามารถยื่นใบสมัครในรอบเก็บตกที่มีชื่อว่า UCAS Clearing house ได้อีกด้วย (<http://campus.campus.star.com/ed...>)

จะเห็นได้ว่า 3 ใน 4 ประเทศไม่มีระบบการสอบคัดเลือกบุคคลเข้าศึกษาต่อในสถาบันอุดมศึกษาโดยเฉพาะ คือ ประเทศสาธารณรัฐสิงคโปร์ ประเทศฟินแลนด์และประเทศสหราชอาณาจักร ทั้งนี้ทั้ง 3 ประเทศใช้ผลการทดสอบจบมัธยมศึกษาปีที่ 6 เป็นหลักในการให้สถาบันอุดมศึกษาพิจารณาผู้สมัครเข้าศึกษาต่อ อย่างไรก็ตามมหาวิทยาลัยส่วนใหญ่ในประเทศฟินแลนด์จะมีการทดสอบครั้งที่ 2 เพิ่มเติมและใช้พิจารณารวมกับผลการสอบจบมัธยมศึกษาปีที่ 6 โดย FME อีกทั้งมีการทดสอบโดยตรงจากนักเรียนที่จบสายอาชีวศึกษา ซึ่งไม่มีคะแนนเกรดจากการสอบ FME จึงต้องสอบข้อสอบที่มหาวิทยาลัยออกข้อสอบเองคือเป็นการรับตรงโดยมหาวิทยาลัยนั่นเองส่วนมหาวิทยาลัยในประเทศสหราชอาณาจักรพิจารณาผลการสอบ A-Level และในสาขาวิชาชีพเฉพาะมีการพิจารณาจากผลการทดสอบเฉพาะวิชาชีพที่มหาวิทยาลัยหรือกลุ่มสาขาร่วมกันออกข้อสอบโดยไม่เกี่ยวกับ UCAS

ประเทศสาธารณรัฐเกาหลี (เกาหลีใต้) เป็นประเทศเดียวที่มีการสอบคัดเลือกเข้ามหาวิทยาลัยคือเป็นการสอบคัดเลือกที่เข้มข้นมีมาตรฐานสูงมากมีความโปร่งใสและเป็นธรรมแก่ผู้สมัครสอบเป็นอย่างยิ่งอีกทั้งมีการบริหารจัดการสอบได้รับการยอมรับว่าดีที่สุดในโลก

ทั้ง 4 ประเทศที่ 3 ประเทศมีระบบการคัดเลือกโดยใช้ผลการสอบระดับชาติในระดับมัธยมศึกษาเพื่อสมัครเข้าศึกษาต่อในสถาบันอุดมศึกษาและมี 1 ประเทศที่มีการสอบคัดเลือกบุคคลเข้าศึกษาต่อมหาวิทยาลัยทุกประเทศได้แสดงให้เห็นถึงการให้ความสำคัญในการคัดเลือกบุคคลเข้าศึกษาต่อระดับอุดมศึกษาของแต่ละประเทศ ทั้งนี้การให้ความร่วมมือกันเพื่อให้ นักเรียนได้มีโอกาสได้รับการคัดเลือกให้เข้าศึกษาต่อในสถาบันอุดมศึกษาที่ต้องการมีการจัดระบบวางแผนการคัดเลือกที่ดีมีประสิทธิภาพ มีความเป็นมาตรฐานและมีความยุติธรรมแก่ผู้สมัคร ซึ่งเป็นเครื่องยืนยันว่าทั้ง 4 ประเทศได้รับการจัดอันดับว่ามีระบบการศึกษาที่ดีเยี่ยมระดับโลก

ข้อเสนอแนะเชิงนโยบาย

1. ประเทศไทยโดยสมาคมสถาบันอุดมศึกษาแห่งประเทศไทย กระทรวงศึกษาธิการ และกระทรวงการอุดมศึกษาวิทยาศาสตร์วิจัยและนวัตกรรม น่าจะนำการปฏิบัติที่ดีของทั้ง 4 ประเทศ ในการปรับปรุงพัฒนาระบบการคัดเลือกบุคคลเข้าศึกษาต่อในสถาบันอุดมศึกษาไทยโดยไม่ต้องมีการสอบคัดเลือกบุคคลเข้าศึกษาต่อในสถาบันอุดมศึกษาทั้งนี้ให้พิจารณาจากผลการสอบเพื่อสำเร็จการศึกษาชั้นมัธยมศึกษาปีที่ 6 เช่นในประเทศสิงคโปร์ ประเทศฟินแลนด์และประเทศสหราชอาณาจักรโดย TCAS น่าจะดำเนินการเช่นเดียวกับ UCAS ของประเทศสหราชอาณาจักร โดยสถาบันทดสอบทางการศึกษาแห่งชาติ พิจารณาปรับปรุงระบบการสอบ O-NET ชั้นมัธยมศึกษาปีที่ 6 ให้เป็นการสอบเพื่อสำเร็จการศึกษาและสามารถใช้ในการสมัครศึกษาต่อในสถาบันอุดมศึกษา เช่น ในกรณีประเทศสาธารณรัฐสิงคโปร์และประเทศสหราชอาณาจักรหรือตามแนวทางการสอบ FME ของประเทศฟินแลนด์ส่วนการทดสอบ O-NET จากกลุ่มตัวอย่างนักเรียนที่เป็นตัวแทน ของนักเรียนชั้นประถมศึกษาปีที่ 6 และชั้นมัธยมศึกษาปีที่ 3 ทั้งควรเพิ่มการทดสอบ O-NET ชั้นมัธยมศึกษาปีที่ 5 ไม่จำเป็นต้องสอบทุกคนเช่นปัจจุบัน เพื่อเป็นการประเมินพิจารณาคุณภาพ การศึกษาขั้นพื้นฐานของประเทศ และประเมินเพื่อพิจารณาความเหมาะสมของหลักสูตร การศึกษาขั้นพื้นฐาน ทำนองเดียวกับการทดสอบ NAEA ของประเทศสาธารณรัฐเกาหลี (เกาหลีใต้) อีกทั้งสถาบันการทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) น่าจะศึกษาแบบอย่าง เพื่อปรับปรุงพัฒนาการจัดการสอบให้ได้มาตรฐาน เช่น การจัดการสอบคัดเลือกเข้าศึกษาต่อสถาบัน อุดมศึกษาของประเทศสาธารณรัฐเกาหลี (เกาหลีใต้) และเพื่อสนับสนุนในการดำเนินการเรื่องนี้ ให้ประสบความสำเร็จจำเป็นต้องระบุในสาระบัญญัติพระราชบัญญัติการศึกษาแห่งชาติฉบับใหม่

2. ด้วยรัฐธรรมนูญแห่งราชอาณาจักรไทยพุทธศักราช 2560 ได้กำหนดคุณลักษณะของ คนไทยประการหนึ่ง คือ สามารถเชี่ยวชาญตามความถนัด ดังนั้นกระทรวงศึกษาธิการพึงปรับปรุง โครงสร้างหลักสูตรการศึกษาขั้นพื้นฐานโดยมีแผนการเรียนรู้ให้นักเรียนได้เรียนตามความสามารถ และความถนัดโดยมีการทดสอบผลสัมฤทธิ์ทางการเรียนทั้งเพื่อการเรียนรู้ เพื่อพัฒนานักเรียน และการประเมินผลสัมฤทธิ์ทางการเรียนที่ทำให้นักเรียนแต่ละคนรู้จักตนเองว่ามีความสนใจ มีความถนัดและความสามารถเพียงใด จะได้เริ่มเรียนในแผนการเรียนรู้ที่เหมาะสม ตลอดจน การศึกษาต่อและการประกอบอาชีพที่ตรงกับความสนใจความรู้ความสามารถและความถนัด ของตนซึ่งจะนำไปสู่การเป็นพลเมืองไทยและพลเมืองโลกที่มีคุณภาพและประสบความสำเร็จ ในการดำรงชีวิตอย่างมีความสุขเป็นที่พึงของสังคมและประเทศชาติต่อไป

เอกสารอ้างอิง

ไทยรัฐฉบับพิมพ์ 26 สิงหาคม 2562

สมหวัง พิธิยานุวัฒน์ และคณะ. (2557). รายงานผลการศึกษาโครงการวิจัยพัฒนาการออกแบบและพัฒนาาระบบติดตามและประเมินผลการศึกษาเสนอต่อสำนักงานเลขาธิการสภาการศึกษา. 29-50; 74-93 (เอกสารอัดสำเนา).

สมหวัง พิธิยานุวัฒน์ และจิราทัศน์ รัตนมณีฉัตร. (2559). กลยุทธ์การจัดการศึกษาขั้นพื้นฐานของสาธารณรัฐสิงคโปร์. วาสารราชบัณฑิตยสภา, 41(2), 254-284.

สำนักงานราชบัณฑิตยสภา.(2561).พจนานุกรมศัพท์ศึกษาศาสตร์ร่วมสมัย ชุดการประเมินการวิจัยและการประกันคุณภาพ ฉบับราชบัณฑิตยสภา.กรุงเทพฯ: สำนักงานราชบัณฑิตยสภา.76-79

Coughlan,S. (2015,May 13). Asia Topped Biggest Global School Rankings.BBC News,p.1.

Retrieved from : <http://www.bbc.com/new/business-32608772>

<http://en.wikipedia.org/wiki/ucas>

<http://www.learningcurve-th.com/singapore/>

<http://www.ncee.org/programs-affiliate/center-on-international-education-benchmarking/top-performing-countries/south-korea-overview/south-korea-instructional-systems/>

<http://seab.gov.sg/about-us/about-SEAB.html>

<https://campus.campus-star.com/ed...>

<http://www.oecd.com/>

<http://yle.fi/uutiset/osasto/news/ylioppilastutkinto.fi.reddit>

Link:seeme.me/ch/goodmorningThailand

www.reddit.com

www.ucas.com

www.ylioppilastutkinto.fi

การประเมินโครงการจัดหาตัวแทน สทศ. ประจำปีสนามสอบ การทดสอบ ทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) ชั้นประถมศึกษาปีที่ 6 และ ชั้นมัธยมศึกษาปีที่ 3

The Procurement Representative of the National Institute of Educational
Testing Service (Public Organization) to Ordinary National Educational Test
(O-NET) Primary 6 and Mathayom 3 Field Exam Project

พัฒนา ธนากร
นางสาวนพมาศ เกตุรามฤทธิ์
นางกาญจนา พวงจิตต์
นางสาวอรสา สุทธิไสย
นางสาววาสนา ประสงค์งาม

บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อประเมินโครงการจัดหาตัวแทน สทศ. ประจำปีสนามสอบ โดยใช้การประเมินรูปแบบชิป (CIPP Model) ของสตัฟเฟิลบีม (Stufflebeam) ประเมินใน 4 ด้าน คือ ด้านบริบท (Context) ด้านปัจจัยนำเข้า (Input) ด้านกระบวนการ (Process) และ ด้านผลผลิต (Product) ใช้วิธีการสุ่มตัวอย่างหลายขั้นตอน (Multi-Stage Random Sampling) เก็บรวบรวมข้อมูลด้วยแบบสอบถามจากกลุ่มตัวอย่าง คณะทำงานโครงการจัดหาตัวแทน สทศ. ประจำปีสนามสอบ จำนวน 242 คน คณะทำงานระดับศูนย์สอบและคณะกรรมการระดับสนามสอบ จำนวน 459 คน แบบทดสอบ และวิเคราะห์ เอกสารการดำเนินงาน สำหรับการวิเคราะห์ ข้อมูลจากแบบสอบถามใช้ค่าสถิติพื้นฐาน ได้แก่ ร้อยละ ค่าเฉลี่ย ($\bar{X}$) และส่วนเบี่ยงเบน มาตรฐาน (S.D) และการวิเคราะห์เนื้อหาจากเอกสารในประเด็นของการปฏิบัติงานของ คณะกรรมการระดับสนามสอบ ผลการวิจัยสรุปได้ดังนี้

1. การประเมินบริบทของโครงการ พบว่าวัตถุประสงค์ของโครงการมีความสอดคล้อง กับนโยบายของสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) การดำเนินการจัดการ ทดสอบในปัจจุบัน เป้าหมาย และกิจกรรมของโครงการ มีการกำหนดวัตถุประสงค์ เป้าหมาย และขั้นตอนของการดำเนินงานตามโครงการชัดเจนอยู่ในระดับมากที่สุด โครงการมีประโยชน์ ผู้ที่เกี่ยวข้องรับรู้และการให้ความร่วมมืออยู่ในระดับมาก

2. การประเมินปัจจัยนำเข้าของโครงการพบว่าความเหมาะสม ความเพียงพอของ บุคลากรงบประมาณ และระยะเวลาในการดำเนินงานอยู่ในระดับมาก เอกสารและคู่มือมี ความเหมาะสม และเพียงพออยู่ในระดับมากที่สุด

3. การประเมินกระบวนการของโครงการพบว่า กระบวนการของโครงการมีความเหมาะสมอยู่ในระดับมากที่สุด การดำเนินงานของคณะกรรมการโครงการตัวแทน สทศ. ประจําสนามสอบ ไม่พบปัญหาและอุปสรรค แต่มีข้อเสนอแนะเพื่อการพัฒนาว่า ควรกำชับให้ตัวแทน สทศ. ดิรุษถ่ายบนบัตรประจําตัว การระบุเวลาในการประชุมที่ชัดเจน ควรขยายระยะเวลาในการรับสมัครตัวแทน สทศ. ควรแจ้งบทบาทหน้าที่ของตัวแทน สทศ. ที่มาปฏิบัติงานในแต่ละสนามสอบให้ชัดเจน

4. การประเมินผลผลิตของโครงการ พบว่า มหาวิทยาลัยที่รับผิดชอบโครงการได้จัดส่งตัวแทน สทศ. ไปประจําสนามสอบในปีการศึกษา 2561 ได้มากกว่าร้อยละ 80 สำหรับการดำเนินการทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) ชั้นประถมศึกษาปีที่ 6 และชั้นมัธยมศึกษาปีที่ 3 ปีการศึกษา 2559 - 2560 ไม่พบการดำเนินงานที่ไม่เป็นไปตามมาตรฐานการทดสอบทางการศึกษาแห่งชาติ และความไม่โปร่งใสในการจัดการทดสอบสำหรับปีการศึกษา 2561 พบการดำเนินการจัดการทดสอบของสนามสอบที่ไม่เป็นไปตามแนวปฏิบัติในคู่มือการจัดสอบจำนวน 3 เรื่อง ได้แก่ การแจกกระดาษคำตอบ การบรรจุซองกระดาษคำตอบ และการใช้เทปกาฉีดยึดประเภทสำหรับการจัดสอบมีความโปร่งใส แต่สนามสอบบางแห่งยังมีจุดอ่อนเรื่องการกำกับดูแลให้คณะกรรมการระดับสนามสอบปฏิบัติตามแนวทางปฏิบัติในช่วงหลังการสอบ

คำสำคัญ: ประเมินโครงการ; โครงการจัดหาตัวแทน สทศ. ประจําสนามสอบ O-NET ป.6 และ ม.3

Abstract

The purpose of this research was to evaluate the Procurement Representative of the National Institute of Educational Testing Service (Public organization) to the Exam Field Project by using the CIPP Model of Stufflebeam. The evaluation covered 4 aspects : the context, the Input, the process and the product. A sample group selected by using Multi -Stage random sampling. Data were collected both questionnaires testing and analysis of operational documents. The questionnaires survey of a sample group of 242 officers of the Procurement Representative of the National Institute of Educational Testing Service (Public organization) to the Exam Field Project and 459 officers in Testing Center and Exam Field. The data from questionnaires were analyzed by percent (%), mean ($\bar{X}$) and standard deviation (S.D). The data from operational documents were analyzed by content analysis. The results of this research were as follows :

1. The context of project was the objectives of the project are consistent with the policies of the National Institute of Educational Testing Service (Public organization). Implementation of the current testing, goals and activities of the project. There are clear objectives, targets and procedures for project implementation at the highest level. The project was useful. People involved perceived and cooperated at a high level.

2. The Input of project was the appropriateness and adequacy of the personnel, the budget and the duration of the operation were at a high level. Documents and manuals are appropriate and sufficient at the highest level.

3. The process of project was at the highest level. The operation of the NIETS representative team at the exam field did not encounter any problems and obstacles. Suggestions for development that should advise the NIETS representative to attach a photograph on the identity card, specifying the meeting time clearly, should extend the period of recruitment of the NIETS representative and should clearly state the roles and responsibilities of the NIETS representatives who work in each examination field.

4. The product of project was the university responsible for the project has sent more than 80% of the NIETS representatives to the test field in the academic year 2018. The implementation of O-NET of Prathom 6 and Mattayomsuksa 3 in the academic year 2016-2017, No found operations that do not meet the national educational testing standards and non-transparency in test management. In the academic year 2018, Found that 3 of the test management operations of the exam field did not follow the guidelines in the manual. Including distribution of the answer sheet, packing answer sheet and using the wrong type of masking tape. The exam was transparent but some test grounds still have weaknesses in governance ensure that the field level committee will comply with the post-examination guidelines.

บทนำ

การทดสอบทางการศึกษาระดับชาตินิยมขั้นพื้นฐาน (O-NET) ชั้นประถมศึกษาปีที่ 6 และ ชั้นมัธยมศึกษาปีที่ 3 เป็นการทดสอบความรู้และความคิดรวบยอดของนักเรียน ตามหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551 และนำผลการทดสอบไปใช้ประโยชน์ต่าง ๆ เช่น เป็นองค์ประกอบหนึ่งในการจบการศึกษา ใช้ในการปรับปรุงคุณภาพการเรียนการสอนของโรงเรียน ใช้ในการประเมินผลการเรียนรู้ของนักเรียนระดับชาติ เป็นต้น ทำให้ในแต่ละปีการศึกษา มีผู้เข้าสอบจำนวนมาก และมีการนำผลสอบไปใช้ประโยชน์มากทั้งระดับนักเรียน สถานศึกษา และต้นสังกัด ทำให้การทดสอบทางการศึกษาระดับชาตินิยมขั้นพื้นฐาน (O-NET) ชั้นประถมศึกษาปีที่ 6 และชั้นมัธยมศึกษาปีที่ 3 มีความสำคัญ และมีผลกระทบทั้งทางตรงและทางอ้อมต่อผู้เข้าสอบและผู้ที่เกี่ยวข้องมาก (High Stake) ทำให้การจัดการทดสอบในระดับสนามสอบ จำเป็นต้องดำเนินการให้เป็นไปตามมาตรฐานการทดสอบทางการศึกษาแห่งชาติ มีความโปร่งใส ยุติธรรม และสร้างความมั่นใจและการยอมรับจากผู้เกี่ยวข้อง

โครงการจัดหาตัวแทน สทศ. ประจำสนามสอบ การทดสอบทางการศึกษาระดับชาตินิยมขั้นพื้นฐาน (O-NET) ชั้นประถมศึกษาปีที่ 6 และชั้นมัธยมศึกษาปีที่ 3 เป็นโครงการที่สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) จัดทำขึ้นตั้งแต่ปีการศึกษา 2556 จนถึงปัจจุบัน เพื่อเป็นการประกันว่าสนามสอบได้ปฏิบัติตามมาตรฐานการทดสอบที่สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) กำหนด โดยได้รับความร่วมมือจากมหาวิทยาลัยราชภัฏทั่วประเทศ มหาวิทยาลัยเทคโนโลยีราชมงคล และมหาวิทยาลัยของรัฐ ทำหน้าที่ดำเนินการตรวจสอบการทำงานของคณะกรรมการในสนามสอบที่กำหนด ซึ่งสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) ได้จัดทำคู่มือการปฏิบัติงานของตัวแทน สทศ. ประจำสนามสอบคู่มือการจัดการสอบทางการศึกษาระดับชาตินิยมขั้นพื้นฐาน (O-NET) และแบบรายงานผลการปฏิบัติงานของตัวแทน สทศ. ประจำสนามสอบ เพื่อควบคุมคุณภาพการปฏิบัติของตัวแทน สทศ. ประจำสนามสอบนอกจากนี้สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) ยังได้กำหนดกระบวนการดำเนินงาน ได้แก่

- (1) สทศ. ประชุมชี้แจงกับตัวแทนมหาวิทยาลัยที่รับผิดชอบโครงการ
- (2) ประธานคณะทำงานโครงการตัวแทน สทศ. ประชุมชี้แจงกับตัวแทน สทศ.
- (3) มหาวิทยาลัยที่รับผิดชอบโครงการสรรหาตัวแทน สทศ.
- (4) ตัวแทน สทศ. เก็บรวบรวมข้อมูลในสนามสอบ
- (5) มหาวิทยาลัยที่รับผิดชอบโครงการรายงานผลการดำเนินงาน

และสรุปค่าใช้จ่ายให้ สทศ. ทราบ รวมถึง สทศ. จัดสรรค่าใช้จ่ายให้กับหน่วยงานที่รับผิดชอบโครงการ เพื่อใช้ในการบริหารโครงการให้บรรลุตามวัตถุประสงค์

การประเมินโครงการ เป็นกระบวนการศึกษาแสวงหาความรู้ความเข้าใจเกี่ยวกับการดำเนินโครงการว่าเป็นไปตามหลักเกณฑ์และขั้นตอนต่าง ๆ ที่ได้กำหนดไว้หรือไม่ มีปัญหาและอุปสรรคอะไร และบรรลุเป้าหมายที่ต้องการหรือไม่ มีผลกระทบในแง่บวกต่าง ๆ อย่างไรที่เกิดขึ้นจากโครงการบ้าง (สุชาติ ประสิทธิ์รัฐสินธุ์. 2554 : 2) รวมถึงเพื่อตรวจสอบ หรือบ่งชี้ถึงประสิทธิภาพและประสิทธิผลของโครงการ (พิชิต ฤทธิ์จรูญ. 2557: 41) และนำผลการประเมินที่ได้ไปใช้ปรับปรุงและตัดสินผลสัมฤทธิ์ของโครงการ (สมหวัง พิธิยานุวัฒน์. 2558 : 117) นอกจากประโยชน์ที่จะเกิดกับผู้บริหารโครงการแล้ว การประเมินโครงการช่วยให้เกิดการเสริมสร้างแรงจูงใจให้กับผู้ปฏิบัติงานในโครงการเมื่อทราบผลสำเร็จ จุดเด่น หรือจุดด้อยของโครงการ โดยจะมุ่งมั่นปรับปรุงและพัฒนาการดำเนินโครงการให้มีคุณภาพและมาตรฐานสูงขึ้น ซึ่งจะเกิดคุณค่าและประโยชน์สูงสุดต่อผู้รับบริการหรือองค์กร (พิชิต ฤทธิ์จรูญ. 2557: 44) ซึ่งการศึกษาวิจัยที่เกี่ยวข้องกับการประเมินโครงการด้านการศึกษาในประเทศไทยในช่วง พ.ศ. 2555-2559 ส่วนใหญ่พบว่ามี การประเมินโครงการในด้านบริบท ด้านปัจจัยนำเข้า ด้านกระบวนการ ด้านผลผลิต

กลุ่มงานบริหารการทดสอบมีหน้าที่และความรับผิดชอบในการกำกับดูแลการดำเนินงานของโครงการจัดหาตัวแทน สทศ. ประจำปีสอบ การทดสอบทางการศึกษาระดับชาตินัยพื้นฐาน (O-NET) ชั้นประถมศึกษาปีที่ 6 และชั้นมัธยมศึกษาปีที่ 3 ให้เป็นไปตามนโยบายทั้งนี้จากการดำเนินการที่ผ่านมาพบว่ายังไม่มี การประเมินโครงการรวมถึงต้องการตรวจสอบว่าโครงการมีการดำเนินการบรรลุผลสำเร็จตามวัตถุประสงค์หรือไม่ การดำเนินงานมีความเหมาะสมมากน้อยเพียงใด รวมถึงมีปัญหาและอุปสรรคอะไรบ้าง เพื่อประกอบการตัดสินใจของผู้บริหารในการขยายหรือยุติการดำเนินงานในช่วงต่อไป รวมถึงนำข้อมูลสารสนเทศที่ได้ มาใช้ปรับปรุงพัฒนาระบบการบริหารจัดการ โครงการในครั้งต่อไปให้มีประสิทธิภาพมากยิ่งขึ้น อันจะส่งผลให้เกิดความเชื่อมั่นและการยอมรับจากผู้มีส่วนเกี่ยวข้อง กลุ่มงานบริหารการทดสอบจึงมีการประเมินโครงการจัดหาตัวแทน สทศ. ประจำปีสอบ การทดสอบทางการศึกษาระดับชาตินัยพื้นฐาน (O-NET) ชั้นประถมศึกษาปีที่ 6 และชั้นมัธยมศึกษาปีที่ 3 โดยประเมินความเหมาะสมด้านบริบท (Context) ด้านปัจจัยนำเข้า (Input) ได้แก่ทรัพยากร บุคลากร งบประมาณ ด้านกระบวนการดำเนินงาน (Process) และด้านผลผลิตของโครงการ (Product)

วัตถุประสงค์การวิจัย

เพื่อประเมินโครงการจัดหาตัวแทน สทศ. ประจำปีสอบ โดยประยุกต์ใช้การประเมินรูปแบบซิป (CIPP Model) ของสตัฟเฟิลบีม (Stufflebeam) ประเมินใน 4 ด้าน คือ ด้านบริบท ด้านปัจจัยนำเข้า ด้านกระบวนการ และด้านผลผลิต โดยมีวัตถุประสงค์เฉพาะ ดังนี้

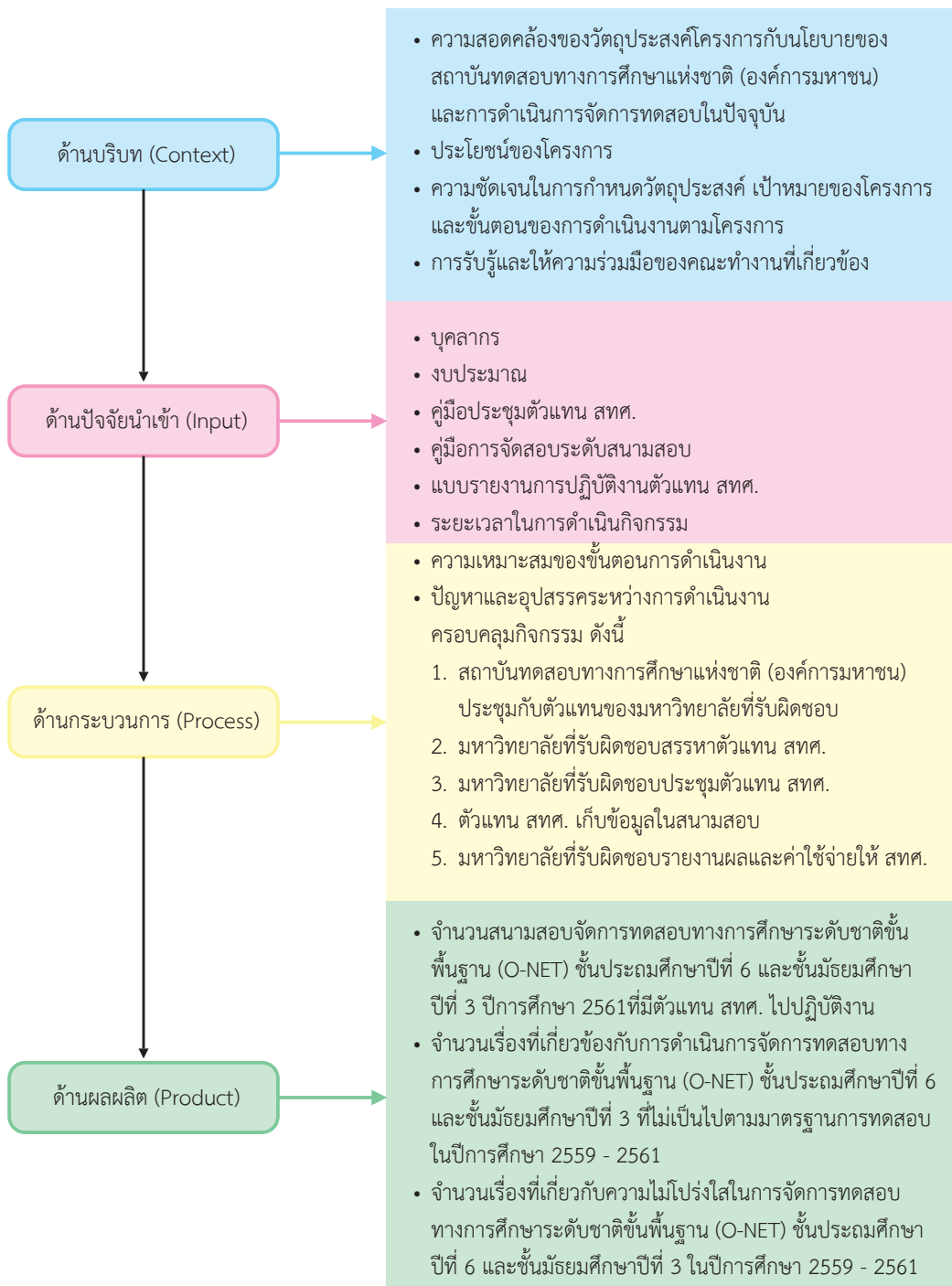
1. เพื่อประเมินบริบทของโครงการจัดหาตัวแทน สทศ. ประจำปีสอบ
2. เพื่อประเมินความเหมาะสมด้านปัจจัยนำเข้าของโครงการจัดหาตัวแทน สทศ. ประจำปีสอบ ได้แก่บุคลากร งบประมาณ เอกสารที่เกี่ยวข้องกับการดำเนินงาน และระยะเวลา

3. เพื่อประเมินความเหมาะสมของขั้นตอนการดำเนินงานของโครงการจัดหาตัวแทน สทศ. ประจำปีสามสิบ

4. เพื่อประเมินผลผลิตของ โครงการจัดหาตัวแทน สทศ. ประจำปีสามสิบ

กรอบแนวคิดการวิจัย

จากการทบทวนวรรณกรรมและงานวิจัยที่เกี่ยวข้องกับการประเมินโครงการด้านการศึกษาระดับพื้นฐานที่ใช้รูปแบบการประเมินชิป (CIPP Model) พบว่า การประเมินโครงการที่ใช้รูปแบบการประเมินชิป (CIPP Model) จะใช้กับโครงการที่เกี่ยวข้องกับการบริหารจัดการ ต้องการนำผลประเมินไปใช้ในการตัดสินใจเพื่อปรับปรุง ขยาย หรือ ยุติโครงการ เช่น งานวิจัยของชนิศกานต์ ละครวงศ์ (2556) ที่ประเมินโครงการพัฒนาโรงเรียนขนาดเล็กต้นแบบการจัดการเรียนการสอนแบบคละชั้น ของสำนักงานเขตพื้นที่การศึกษาจันทบุรี เขต 1 ในด้านบริบทปัจจัยนำเข้ากระบวนการและผลผลิตของโครงการงานวิจัยของกมลานันท์ บุญกล้า (2559) ที่ประเมินโครงการเศรษฐกิจพอเพียงของโรงเรียนบ้านหนองปลาชีว สำนักงานเขตพื้นที่การศึกษาประถมศึกษาเชิงเทรา เขต 2 ในด้านสภาวะแวดล้อมปัจจัยนำเข้า กระบวนการดำเนินงาน และผลผลิต เป็นต้น นอกจากนี้ยังมีงานวิจัยที่เกี่ยวข้องกับการประเมินศูนย์สอบและสนามสอบการทดสอบทางการศึกษาระดับชาติ ได้นำแนวคิดของการประเมินโครงการที่ใช้รูปแบบการประเมินชิป (CIPP Model) มาประยุกต์ใช้ เช่น งานวิจัยของเมธพร หิรัญวงศ์ และสุกัญญา ไกรนรา (2556) ได้ศึกษากระบวนการจัดสอบและประสิทธิภาพในด้านบริบท ด้านปัจจัยนำเข้า ด้านกระบวนการ และด้านผลลัพธ์ งานวิจัยของรุ่งฤดี กล้าหาญ (2557) ได้ประเมินมาตรฐานการบริหารการทดสอบของสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) ประจำปีการศึกษา 2557 : ศึกษากรณีศูนย์สอบมหาวิทยาลัยศรีนครินทรวิโรฒ โดยกรอบแนวคิดในการประเมินได้จากการบูรณาการรูปแบบการประเมินของ Provus (1969) Stufflebeam (1971) และ Stake (1976) โดยประเมินความเหมาะสมด้านปัจจัยนำเข้า ความเหมาะสมของกระบวนการดำเนินการบริหารการทดสอบ และผลผลิตของการบริหารการทดสอบของสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) งานวิจัยนี้จึงใช้กรอบแนวคิดในการประเมินจากรูปแบบการประเมินชิป (CIPP Model) ของสตฟเฟิลบีม (Stufflebeam) โดยประเมินด้านบริบทของโครงการ (Context) ปัจจัยนำเข้า (Input) กระบวนการ (Process) และผลผลิต (Product) ดังภาพที่ 1



ภาพที่ 1 กรอบแนวคิดการวิจัย

วิธีการดำเนินการวิจัย
การวิจัยในครั้งนี้ ใช้การวิจัยเชิงประเมิน (Evaluation Research) ซึ่งผู้วิจัยได้ดำเนินการศึกษาวิจัยตามกรอบแนวทางการประเมิน ดังนี้
ตาราง 1 กรอบแนวทางการประเมิน

วัตถุประสงค์ของ การประเมิน	ตัวชี้วัด/ประเด็นที่ศึกษา	แหล่งข้อมูล	เครื่องมือการเก็บ รวบรวมข้อมูล	การวิเคราะห์ข้อมูล	เกณฑ์การประเมิน
1. เพื่อประเมินบริบท ของโครงการ	1. ความสอดคล้องของวัตถุประสงค์ โครงการกับนโยบายของ สทศ. และ การดำเนินการจัดการทดสอบในชั้น ปี 2. ประสิทธิภาพของการจัดการ 3. ความชัดเจนในการกำหนด วัตถุประสงค์เป้าหมายของโครงการ และขั้นตอนของการดำเนินงานตาม โครงการ 4. การรับรู้และให้ความร่วมมือของ คณะทำงานที่เกี่ยวข้อง	1. คณะทำงานโครงการจัดหา ตัวแทน สทศ. ประจำ สนามสอบ 2. คณะทำงานระดับศูนย์สอบ และสนามสอบ	1. แบบสอบถาม	1. ค่าเฉลี่ย ($\bar{X}$) 2. ส่วนเบี่ยงเบน มาตรฐาน (S.D)	1. 3.50 หรือมีความ เหมาะสมตั้งแต่ ระดับมาก ขึ้นไป
2. เพื่อประเมินปัจจัย นำเข้าของโครงการ	1. บุคลากรที่เป็นตัวแทน สทศ. 1.1 ความรู้ในเรื่องบทบาทของ ตัวแทน สทศ. 1.2 ความเพียงพอของจำนวน ตัวแทน สทศ. ในสนามสอบที่ แต่ละหน่วยงานรับผิดชอบ 1.3 ความเหมาะสมของการ กำหนดตำแหน่งในคณะกรรมการ และบุคลากรที่ได้รับแต่งตั้งเป็น ตัวแทน สทศ.	1. คณะทำงานโครงการจัดหา ตัวแทน สทศ. ประจำ สนามสอบ 2. คณะทำงานระดับศูนย์สอบ และสนามสอบ	1. แบบทดสอบ 2. แบบสอบถาม	1. ค่าเฉลี่ย ($\bar{X}$) 2. ส่วนเบี่ยงเบน มาตรฐาน (S.D)	1. มีคะแนนความรู้ใน เรื่องบทบาทของ ตัวแทน สทศ. ตั้งแต่ 8 คะแนนขึ้นไป หรือในระดับมาก 2. 3.50 หรือมีความ เหมาะสมตั้งแต่ ระดับมากขึ้นไป

วัตถุประสงค์ของการประเมิน	ตัวชี้วัด/ประเด็นที่ศึกษา	แหล่งข้อมูล	เครื่องมือการเก็บรวบรวมข้อมูล	การวิเคราะห์ข้อมูล	เกณฑ์การประเมิน
	2. งบประมาณ 2.1 ความเพียงพอของงบประมาณที่หน่วยงานที่รับผิดชอบได้รับ 2.2 ความเหมาะสมของอัตราค่าใช้จ่าย 2.3 ความคุ้มค่าของงบประมาณกับผลที่ได้ 3. เอกสารที่เกี่ยวข้องกับการดำเนินงาน 3.1 ความเพียงพอของเอกสาร 3.2 ความเหมาะสมของเนื้อหาในเอกสาร 4. ความเหมาะสมของระยะเวลา การดำเนินงานตามกิจกรรมของโครงการ ได้แก่ 4.1 สทศ. ประชุมกับตัวแทนของมหาวิทยาลัยที่รับผิดชอบ 4.2 มหาวิทยาลัยที่รับผิดชอบ 4.3 มหาวิทยาลัยที่รับผิดชอบ 4.4 ตัวแทน สทศ. เก็บข้อมูลในสนามสอบ				

วัตถุประสงค์ของ การประเมิน	ตัวชี้วัด/ประเด็นที่ศึกษา	แหล่งข้อมูล	เครื่องมือการเก็บ รวบรวมข้อมูล	การวิเคราะห์ข้อมูล	เกณฑ์การประเมิน
	4.5 มหาวิทยาลัยที่รับผิดชอบ รายงานผลการดำเนินงานและ สรุป ค่าใช้จ่ายให้ สทศ.				
3. เพื่อประเมิน ผลผลิตของ โครงการตัวแทน สทศ.	1. จำนวนตัวแทน สทศ. ประจำ สนามสอบ ปีการศึกษา 2561 2. จำนวนเรื่องที่เกี่ยวข้องกับการ ดำเนินการทดสอบที่ไม่เป็นไป ตามมาตรฐานการทดสอบ ทางการศึกษาแห่งชาติของ ทดสอบทางการศึกษาระดับชาติ ขั้นพื้นฐาน (O-NET) ชั้นประถม ศึกษาปีที่ 6 และชั้นมัธยมศึกษา ปีที่ 3 ในปีการศึกษา 2559 - 2561 3. จำนวนเรื่องที่เกี่ยวข้องกับความ ไม่โปร่งใสในการจัดการทดสอบ ทางการศึกษาระดับชาติ ขั้นพื้นฐาน (O-NET) ชั้นประถมศึกษา ปีที่ 6 และ ชั้นมัธยมศึกษา ปีที่ 3 ใน ปีการศึกษา 2559 - 2561	1. รายงานการดำเนินงานโครงการ จัดทำ ตัวแทน สทศ. ประจำ สนามสอบ ปีการศึกษา 2561 2. รายงานผลการดำเนินงาน การทดสอบทางการศึกษา ระดับชาติขั้นพื้นฐาน ชั้นประถมศึกษาปีที่ 6 และ ชั้นมัธยมศึกษาปีที่ 3 ตั้งแต่ ปีการศึกษา 2559 ถึง 2561	1. แบบบันทึกข้อมูล การวิเคราะห์เอกสาร	1. วิเคราะห์เนื้อหา 2. ไม่พบเรื่อง ที่เกี่ยวข้องไม่เป็นไป ตามมาตรฐาน การทดสอบทาง การศึกษาแห่งชาติ 3. ไม่พบเรื่อง ที่เกี่ยวข้องกับความ ไม่โปร่งใสในการ จัดการทดสอบ	

ประชากรและกลุ่มตัวอย่าง

ประชากร คือ คณะทำงานโครงการจัดหาตัวแทน สทศ. จำนวน 5,659 คน คณะทำงานระดับศูนย์สอบและคณะกรรมการระดับสนามสอบ การทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) ชั้นประถมศึกษาปีที่ 6 และชั้นมัธยมศึกษาปีที่ 3 ปีการศึกษา 2561 จำนวน 22,920 คน

กลุ่มตัวอย่างได้แก่ คณะทำงานโครงการจัดหาตัวแทน สทศ. จำนวน 242 คน คณะทำงานระดับศูนย์และคณะกรรมการระดับสนามสอบ จำนวน 459 คน ซึ่งได้มาจากการสุ่มหลายขั้นตอน (Multi - Stage Random Sampling)

เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัยครั้งนี้ คือ

- (1) แบบบันทึกข้อมูลการวิเคราะห์เอกสาร
- (2) แบบสอบถาม 2 ฉบับ ได้แก่

แบบสอบถามสำหรับตัวแทน สทศ. ประจำสนามสอบมีจำนวน 2 ตอน โดยตอนที่ 1 ข้อมูลทั่วไปของผู้ตอบแบบสอบถาม รูปแบบคำถามเป็นแบบเลือกตอบเพียงคำตอบเดียว จำนวน 7 ข้อ ตอนที่ 2 ความเหมาะสมด้านบริบทปัจจัย และกระบวนการดำเนินงานของโครงการตัวแทน สทศ. รูปแบบคำถามมี 2 รูปแบบ คือ แบบเลือกตอบที่ตรงกับความคิดเห็นเพียงคำตอบเดียว จำนวนทั้งหมด 26 ข้อ และแบบเติมคำตอบ จำนวน 2 ข้อ และแบบสอบถามสำหรับศูนย์สอบและสนามสอบมีจำนวน 2 ตอน โดยตอนที่ 1 ข้อมูลทั่วไปของผู้ตอบแบบสอบถามรูปแบบคำถามเป็นแบบเลือกตอบเพียงคำตอบเดียว จำนวน 6 ข้อ ตอนที่ 2 ความคิดเห็นเกี่ยวกับโครงการตัวแทน สทศ. ประจำสนามสอบ รูปแบบคำถามมี 2 รูปแบบ คือ แบบเลือกตอบที่ตรงกับความคิดเห็นเพียงคำตอบเดียว จำนวนทั้งหมด 11 ข้อ และแบบเติมคำตอบ จำนวน 1 ข้อ (3) แบบทดสอบสำหรับตัวแทน สทศ. มีลักษณะเป็นแบบเลือกตอบ ถูก-ผิด จำนวน 10 ข้อ ข้อละ 1 คะแนน รวมทั้งรวมทั้งหมด 10 คะแนน ซึ่งเครื่องมือที่ใช้ในการวิจัยข้อที่ (2) และ (3) มีขั้นตอนการสร้างและตรวจสอบคุณภาพ ดังนี้

1. ศึกษาเอกสารและข้อมูลที่ได้จากรายงานการปฏิบัติงานของโครงการตัวแทน สทศ. ประจำสนามสอบ และงานวิจัยที่เกี่ยวข้อง เพื่อกำหนดกรอบแนวคิด โครงสร้างของตัวแปรที่ต้องการวัดและกำหนดรูปแบบและโครงสร้างของเนื้อหาของแบบสอบถาม

2. สร้างข้อคำถามในแต่ละตอนแล้วนำแบบสอบถามทั้ง 2 ฉบับที่สร้างขึ้นให้ผู้ทรงคุณวุฒิจำนวน 3 คน ตรวจสอบความตรงเชิงเนื้อหาโดยการตรวจสอบความสอดคล้องระหว่างข้อคำถามกับนิยามเชิงปฏิบัติการ (Item Objective Congruence: IOC) ซึ่งข้อคำถามทุกข้อมีค่า IOC มากกว่า 0.50 ถือว่าข้อคำถามทั้งหมดสอดคล้องกับจุดมุ่งหมาย

3. ปรับปรุงภาษาที่ใช้ในแบบสอบถามตามคำแนะนำของผู้ทรงคุณวุฒิ คือ คำถามในแบบทดสอบแบบ ถูก - ผิด ไม่ควรมีคำว่า ทุกคน เพราะจะทำให้เดาคำตอบได้ง่าย การใช้คำควรเป็นเชิงบวกและกระชับ และข้อคำถามที่มีหลายประเด็นให้แยกเป็นข้อย่อย

4. จัดพิมพ์แบบสอบถามสำหรับเก็บรวบรวมข้อมูล

5. ทดลองใช้แบบสอบถามกับกลุ่มที่คล้ายกับกลุ่มตัวอย่าง จำนวน 30 คน และวิเคราะห์ค่าความเที่ยงของแบบสอบถาม โดยใช้ค่าสัมประสิทธิ์สหสัมพันธ์ครอนบาคแอลฟา (Cronbach's alpha coefficient) ซึ่งแบบสอบถามแต่ละฉบับมีค่าความเที่ยงดังนี้ แบบสอบถามสำหรับตัวแทน สทศ. มีค่าความเที่ยงเท่ากับ 0.96 แบบสอบถามสำหรับศูนย์สอบ และสนามสอบ มีค่าความเที่ยงเท่ากับ 0.75

การเก็บรวบรวมข้อมูล

การเก็บรวบรวมข้อมูลในการวิจัยครั้งนี้ ประกอบด้วย 2 ส่วน คือ

(1) เก็บรวบรวมข้อมูลด้วยแบบสอบถามแบบทดสอบ

(2) เก็บรวบรวมข้อมูลจากเอกสารด้วยแบบบันทึกข้อมูลจากการวิเคราะห์เอกสาร

1. เก็บรวบรวมข้อมูลด้วยแบบสอบถาม มีขั้นตอนดังนี้

1.1 จัดพิมพ์และส่งหนังสือขอความอนุเคราะห์ รวมทั้งแบบสอบถามสำหรับตัวแทน สทศ. ประจำสนามสอบจำนวน 300 ฉบับ ในการเก็บรวบรวมข้อมูลไปยังมหาวิทยาลัยที่รับผิดชอบโครงการจัดหาตัวแทน สทศ. และแบบสอบถามสำหรับศูนย์สอบและสนามสอบ จำนวน 500 ฉบับ ในการเก็บรวบรวมข้อมูลไปยังศูนย์สอบและสนามสอบ โดยส่งไปเป็นจำนวนที่มากกว่ากลุ่มตัวอย่างที่ต้องการ 2 เท่า

1.2 รวบรวมแบบสอบถามสำหรับตัวแทน สทศ. ประจำสนามสอบกลับคืนจากมหาวิทยาลัยที่รับผิดชอบโครงการจัดหาตัวแทน สทศ. ประจำสนามสอบได้จำนวน 242 ฉบับ และรวบรวมแบบสอบถามสำหรับศูนย์สอบและสนามสอบกลับคืนจากศูนย์สอบและสนามสอบได้จำนวน 459 ฉบับ เพื่อนำมาวิเคราะห์ข้อมูล

2. เก็บรวบรวมข้อมูลจากเอกสารด้วยแบบบันทึกข้อมูลจากการวิเคราะห์เอกสาร มีขั้นตอนดังนี้

2.1 รวบรวมเอกสารที่เกี่ยวข้อง ได้แก่

(1) รายงานการดำเนินงานโครงการจัดหาตัวแทน สทศ. ประจำสนามสอบ ปีการศึกษา 2561

(2) รายงานผลการดำเนินงานการจัดการทดสอบทางการศึกษาระดับชาตินี้พื้นฐาน (O-NET) ชั้นประถมศึกษาปีที่ 6 และชั้นมัธยมศึกษาปีที่ 3 ตั้งแต่ปีการศึกษา 2559 ถึง 2561

2.2 บันทึกข้อมูลในประเด็นที่เกี่ยวข้องกับการปฏิบัติงานของคณะกรรมการระดับสนามสอบที่ไม่ปฏิบัติตามมาตรฐานการทดสอบทางการศึกษาแห่งชาติ และไม่โปร่งใส ลงในแบบบันทึกข้อมูลการวิเคราะห์เอกสาร

การวิเคราะห์ข้อมูล

ผู้วิจัยได้นำข้อมูลที่ได้จากแบบสอบถามและเอกสาร มาวิเคราะห์และหาค่าสถิติต่าง ๆ ตามลำดับดังต่อไปนี้

1. วิเคราะห์ค่าสถิติพื้นฐาน จากแบบสอบถามที่ได้รับรวบรวมมา ดังนี้
 - 1.1 ข้อมูลพื้นฐานของผู้ตอบแบบสอบถาม ใช้การวิเคราะห์ด้วยค่าร้อยละ
 - 1.2 ความคิดเห็นเพื่อประเมินข้อมูลด้านบริบท (Context) ด้านปัจจัยนำเข้า (Input) ด้านกระบวนการ (Process) และด้านผลผลิต (Product) ใช้การวิเคราะห์ด้วยค่าเฉลี่ย ($\bar{X}$) และส่วนเบี่ยงเบนมาตรฐาน (S.D) โดยมีเกณฑ์ตัวชี้วัดตามเกณฑ์ของ Best (1981) ดังนี้

4.50 - 5.00 หมายถึง เหมาะสมมากที่สุด

3.50 - 4.49 หมายถึง เหมาะสมมาก

2.50 - 3.49 หมายถึง เหมาะสมปานกลาง

1.50 - 2.49 หมายถึง เหมาะสมน้อย

1.00 - 1.49 หมายถึง เหมาะสมน้อยที่สุด

2. การวิเคราะห์เนื้อหา (Content Analysis) ในเอกสารเพื่อวิเคราะห์ข้อมูลในประเด็นของการปฏิบัติงานของคณะกรรมการระดับสนามสอบ ที่ไม่เป็นไปตามมาตรฐานการทดสอบทางการศึกษาแห่งชาติ ความไม่โปร่งใสในการ จัดการทดสอบทางการศึกษาระดับชาตินั้นพื้นฐาน (O-NET) ชั้นประถมศึกษาปีที่ 6 และชั้นมัธยมศึกษาปีที่ 3 ซึ่งมีขั้นตอนการวิเคราะห์ ดังนี้

2.1 วิเคราะห์ข้อมูลในประเด็นที่เกี่ยวข้องกับการปฏิบัติงานของคณะกรรมการระดับสนามสอบ โดยจำแนกข้อมูลเป็น 2 กลุ่ม คือ กลุ่มที่ 1 ไม่ปฏิบัติตามมาตรฐานการทดสอบทางการศึกษาแห่งชาติ และกลุ่มที่ 2 ไม่โปร่งใสในการจัดสอบ

2.2 สรุปจำนวนเรื่องในแต่ละกลุ่มที่จำแนกไว้

สรุปผลการวิจัย

1. การประเมินบริบทของโครงการจัดหาตัวแทน สทศ. ประจำปีสนามสอบ พบว่า วัตถุประสงค์โครงการมีความสอดคล้องกับนโยบายของสทศ. การดำเนินการจัดการทดสอบในปัจจุบัน เป้าหมาย และกิจกรรมของโครงการอยู่ในระดับมากที่สุด โครงการมีประโยชน์อยู่ในระดับมาก โดยโครงการทำให้สนามสอบมีความโปร่งใสมากขึ้น และทำให้ผู้มีส่วนเกี่ยวข้องมั่นใจในผลการทดสอบ O-NET มากขึ้น การกำหนดวัตถุประสงค์ เป้าหมายของโครงการ และการกำหนดขั้นตอนของการดำเนินงานตามโครงการ มีความชัดเจนอยู่ในระดับมากที่สุด และการรับรู้และการให้ความร่วมมือของคณะทำงานที่เกี่ยวข้อง อยู่ในระดับมาก

2. การประเมินปัจจัยนำเข้าของโครงการจัดหาตัวแทน สทศ. ประจำปีสนามสอบพบว่า

(1) คณะทำงานตัวแทน สทศ. ประจำปีสนามสอบมีความรู้เกี่ยวกับบทบาทหน้าที่ของตัวแทน สทศ. อยู่ในระดับมาก

(2) บุคลากรตัวแทน สทศ. มีความเหมาะสมอยู่ในระดับมาก ทั้งในเรื่องการกำหนดตำแหน่งของบุคลากรในโครงการความน่าเชื่อถือ ส่งเสริมภาพลักษณ์ให้กับการทดสอบและไม่มีส่วนได้ส่วนเสียกับการสอบ O-NET ป.6 และ ม.3 แต่มีข้อเสนอแนะเพิ่มเติมว่า เรื่องการแต่งตั้งบุคคลควรคำนึงถึงคุณวุฒิ วัยวุฒิ และมีประสบการณ์ในการจัดสอบประเภทอื่น ๆ ของผู้ถูกแต่งตั้งสำหรับจำนวนตัวแทน สทศ. ประจำสนามสอบเพียงพอต่อการปฏิบัติงานในสนามสอบ แต่มีข้อเสนอแนะเพิ่มเติมว่าหากเป็นสนามสอบขนาดใหญ่ควรจะเพิ่มจำนวนตัวแทน สทศ.

(3) เรื่องงบประมาณมีความเหมาะสมอยู่ในระดับมาก ทั้งเรื่องอัตราค่าใช้จ่ายที่ สทศ. กำหนด ผลของโครงการจัดหาตัวแทน สทศ. มีความคุ้มค่ากับงบประมาณ นอกจากนี้ งบประมาณที่ได้รับจาก สทศ. เพียงพอ

(4) เรื่องเอกสารและคู่มือมีความเหมาะสมอยู่ในระดับมากที่สุด ทั้งในเรื่องเนื้อหาในเอกสารและคู่มืออ่านแล้วเข้าใจง่าย และครอบคลุมกับเรื่องที่จะปฏิบัติงานสำหรับจำนวนเอกสารและคู่มือเพียงพอต่อจำนวนผู้ปฏิบัติงาน

(5) เรื่องระยะเวลาในการดำเนินงานมีความเหมาะสมอยู่ในระดับมาก

3. การประเมินกระบวนการของโครงการจัดหาตัวแทน สทศ. ประจำสนามสอบพบว่า มีความเหมาะสมอยู่ในระดับมากที่สุด โดยมีกำหนดการในแต่ละกิจกรรมชัดเจน หลักเกณฑ์การสรรหาและแต่งตั้งคณะทำงานตัวแทน สทศ. มีความชัดเจนและสามารถปฏิบัติได้ วิธีการสรรหาและแต่งตั้งคณะทำงานตัวแทน สทศ. มีความเหมาะสม มีการจัดเตรียมอุปกรณ์ ได้แก่ บัตรประจำตัว หนังสือแจ้งหัวหน้าสนามสอบ แบบรายงานการปฏิบัติงาน และคู่มือการจัดสอบระดับสนามสอบ ให้กับตัวแทน สทศ. ประจำสนามสอบครบถ้วน วิธีการกำกับติดตามการปฏิบัติหน้าที่ของตัวแทน สทศ. ประจำสนามสอบ มีความเหมาะสม ช่องทางการติดต่อเพื่อประสานงานระหว่างหน่วยงานและระหว่างบุคคลด้วยโทรศัพท์และอีเมล มีความสะดวก ติดต่อได้ง่าย การประชาสัมพันธ์โครงการจัดหาตัวแทน สทศ. ประจำสนามสอบในการประชุมชี้แจงศูนย์สอบมีความเหมาะสม นอกจากนี้มีข้อเสนอแนะเพิ่มเติมว่า ควรกำชับให้ตัวแทน สทศ. ติดรูปถ่ายบนบัตรประจำตัว ควรระบุเวลาในการประชุมที่ชัดเจน ควรขยายระยะเวลาในการรับสมัครตัวแทน สทศ. และควรมีการแจ้งบทบาทหน้าที่ของตัวแทน สทศ. ที่มาแต่ละสนามสอบให้ชัดเจน

4. การประเมินผลผลิตของโครงการจัดหาตัวแทน สทศ. ประจำสนามสอบ พบว่า ในปีการศึกษา 2561 มหาวิทยาลัยที่รับผิดชอบโครงการได้จัดส่งตัวแทน สทศ. ไปประจำสนามสอบชั้นประถมศึกษาปีที่ 6 (ร้อยละ 86.35) และจัดส่งตัวแทน สทศ. ไปประจำสนามสอบชั้นมัธยมศึกษาปีที่ 3 (ร้อยละ 93.80) ซึ่งสูงกว่าเกณฑ์ที่ตั้งไว้ (ร้อยละ 80) นอกจากนี้ การดำเนินการทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) ชั้นประถมศึกษาปีที่ 6 และชั้นมัธยมศึกษาปีที่ 3 ปีการศึกษา 2559 - 2560 ไม่พบการดำเนินงานที่ไม่เป็นไปตามมาตรฐานการทดสอบทางการศึกษาแห่งชาติ และความไม่โปร่งใสในการจัดการทดสอบ สำหรับปีการศึกษา 2561 พบการดำเนินการ

จัดการทดสอบของสนามสอบที่ไม่เป็นไปตามแนวปฏิบัติในคู่มือการจัดการทดสอบจำนวน 3 เรื่อง ได้แก่ การแจกกระดาษคำตอบไม่เป็นรูปตัว U ไม่นำซองกระดาษคำตอบสำหรับผู้เข้าสอบกรณีพิเศษบรรจุใส่กล่องกระดาษคำตอบกลับ และใช้เทปกาวติดซองกระดาษคำตอบไปติดกล่องแบบทดสอบ นอกจากนี้การจัดการทดสอบมีความโปร่งใส แต่สนามสอบบางแห่งยังมีจุดอ่อนเรื่องการกำกับดูแล ให้คณะกรรมการระดับสนามสอบปฏิบัติตามแนวทางปฏิบัติในช่วงหลังการสอบจำนวน 1 เรื่อง คือ หลังจากการสอบเสร็จสิ้น มีกรรมการคุมสอบ 1 คน ถ่ายรูปและแชร์ข้อสอบในกลุ่มไลน์โรงเรียน โดยศูนย์สอบในกำกับได้ทราบข่าวจึงได้ตั้งเตือนกรรมการคุมสอบและให้ลบข้อสอบนั้นออกไปจากกลุ่มไลน์ดังกล่าว

ในภาพรวมสามารถสรุปผลการประเมินด้านบริบท ปัจจัยนำเข้า กระบวนการของโครงการจัดหาตัวแทน สทศ. ประจำปีสนามสอบ มีความเหมาะสมเป็นไปตามเกณฑ์ที่กำหนดไว้ทุกประเด็นที่ศึกษา สำหรับผลผลิตของโครงการจัดหาตัวแทน สทศ. ประจำปีสนามสอบ ในภาพรวมเป็นไปตามเกณฑ์ที่ตั้งไว้ทุกประเด็นที่ศึกษา แต่สนามสอบบางแห่งยังมีจุดอ่อนในเรื่องการกำกับดูแล ให้คณะกรรมการระดับสนามสอบปฏิบัติตามแนวทางปฏิบัติในคู่มือการจัดการทดสอบ

อภิปรายผลการวิจัย

1. การประเมินบริบทของโครงการจัดหาตัวแทน สทศ. ประจำปีสนามสอบพบว่า ความสอดคล้องของวัตถุประสงค์โครงการกับนโยบายของสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) การดำเนินการจัดการทดสอบในปัจจุบัน และความสอดคล้องของวัตถุประสงค์ เป้าหมาย และกิจกรรมของโครงการ ความชัดเจนในการกำหนดวัตถุประสงค์ เป้าหมายของโครงการและการกำหนดขั้นตอนของการดำเนินงานตามโครงการ มีความเหมาะสมอยู่ในระดับมากที่สุดนั้น เป็นเพราะสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) มีแผนการดำเนินงานประจำปีที่ชัดเจน และมีการทบทวนแผนการปฏิบัติงานกับนโยบายและการจัดการทดสอบทุกไตรมาส ซึ่งสอดคล้องกับหลักการกำหนดจุดประสงค์โครงการที่ว่า จุดประสงค์ควรมีความเฉพาะเจาะจง วัดได้ ปฏิบัติได้ สอดคล้องกับสภาพจริง มีระยะเวลาที่แน่นอน และควรมีการทบทวนจุดประสงค์นั้น (Harvard Business Review, 2012: 6) สำหรับเรื่องประโยชน์ของโครงการ กลุ่มตัวอย่างที่เป็นคณะกรรมการระดับศูนย์สอบและคณะกรรมการระดับสนามสอบมีความคิดเห็นถึงประโยชน์และความจำเป็นของโครงการน้อยกว่ากลุ่มตัวอย่างที่เป็นคณะกรรมการโครงการตัวแทน สทศ. เนื่องจากปัจจุบันสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) มีการประชุมชี้แจงศูนย์สอบก่อนการจัดการสอบและเน้นย้ำในเรื่องมาตรฐานการทดสอบทางการศึกษาแห่งชาติ และความโปร่งใสในการจัดสอบ ทำให้ศูนย์สอบและสนามสอบได้ตระหนักถึงการปฏิบัติงานตามมาตรฐานการทดสอบทางการศึกษาแห่งชาติ และตามแนวปฏิบัติการจัดการสอบที่มีความโปร่งใสมากขึ้น สำหรับเรื่องการรับรู้และการให้ความร่วมมือของคณะกรรมการที่เกี่ยวข้องมีความเหมาะสมระดับมาก-มากที่สุดนั้น เป็นเพราะสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) มีการนำผลที่ตัวแทน

สทศ. ได้บันทึกมาจากสนามสอบ มาใช้พูดคุยในการประชุมชี้แจงกับศูนย์สอบ เพื่อหารือและช่วยกันแก้ปัญหา ดังงานวิจัยของรุ่งฤดี กล้าหาญ (2557) ทำการศึกษาวิจัยเรื่อง การประเมินมาตรฐานการบริหารการทดสอบของสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) ประจำปีการศึกษา 2557 : ศึกษากรณีศูนย์สอบมหาวิทยาลัยศรีนครินทรวิโรฒ ที่ให้แนวคิดไว้ว่า การประชุมชี้แจงระหว่างผู้รับผิดชอบกับผู้ปฏิบัติงานโดยตรง ทำให้ผู้ปฏิบัติงานรับทราบแนวทางการปฏิบัติ และหากสงสัยสามารถสอบถามได้โดยตรง

2. การประเมินปัจจัยนำเข้าของโครงการจัดหาตัวแทน สทศ. ประจำปีสนามสอบ พบว่า กลุ่มตัวอย่างที่เป็นคณะทำงานตัวแทน สทศ. ประจำปีสนามสอบมีความรู้เกี่ยวกับบทบาทหน้าที่ของตัวแทน สทศ. อยู่ในระดับมากนั้น เนื่องจากสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) มีการจัดประชุมชี้แจง จัดทำคู่มือการปฏิบัติงานและส่งให้กับหน่วยงานที่รับผิดชอบโครงการ รวมถึงคุณภาพของเอกสารและคู่มือมีความเหมาะสมและเพียงพอต่อผู้ปฏิบัติงาน ดังเห็นได้จากผลการประเมินปัจจัยเรื่องเอกสารและคู่มือมีความเพียงพอ และความเหมาะสมของเนื้อหา อยู่ในระดับมากที่สุด สำหรับเรื่องบุคลากร

ตัวแทน สทศ. พบว่า กลุ่มตัวอย่างที่เป็นคณะทำงานตัวแทน สทศ. คิดว่าจำนวนบุคลากรเพียงพอ แต่หากสนามสอบใหญ่ก็ควรเพิ่มจำนวนบุคลากรเป็นรายกรณี ซึ่งข้อเสนอแนะสอดคล้องกับอัตราการแต่งตั้งคณะกรรมการระดับสนามสอบที่จะเปลี่ยนแปลงไปตามจำนวนห้องสอบ ตามคู่มือการจัดสอบทางการศึกษาระดับชาตินั้นพื้นฐาน ชั้นประถมศึกษาปีที่ 6 และชั้นมัธยมศึกษาปีที่ 3 ปีการศึกษา 2561 สำหรับศูนย์สอบ (สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน). 2561 : 26-27) นอกจากนี้กลุ่มตัวอย่างที่เป็นคณะทำงานระดับศูนย์สอบและคณะกรรมการระดับสนามสอบมีความคิดเห็นว่า บุคลากรตัวแทน สทศ. มีความน่าเชื่อถือ และมาจากหน่วยงานที่มีความเป็นกลาง มีความเหมาะสมอยู่ในระดับมาก แต่ควรแต่งตั้งบุคคลที่มีความน่าเชื่อถือ มีคุณวุฒิและวิสัยทัศน์ และมีประสบการณ์ในการจัดสอบประเภทอื่น ๆ เนื่องจากบุคลากร ตัวแทน สทศ. ประจำปีสนามสอบ กลุ่มหนึ่งจะเป็นนักศึกษาในเรื่องงบประมาณ มีความเหมาะสมอยู่ในระดับมาก โดยคณะทำงานของโครงการตัวแทน สทศ. คิดว่างบประมาณที่ได้รับเพียงพอ แต่อัตราค่าใช้จ่ายควรจะให้เบิกแบบเหมาจ่ายได้ทุกกรณี รวมถึงเพิ่มค่าตอบแทนที่พัก และค่าเดินทางมากขึ้น ส่วนหนึ่งอาจเนื่องมาจากความห่างไกลของสนามสอบ ทำให้ผู้ปฏิบัติงานคิดว่าค่าตอบแทนที่ได้รับไม่เหมาะสม แต่กลุ่มตัวอย่างที่เป็นคณะทำงานระดับศูนย์สอบและคณะกรรมการระดับสนามสอบคิดว่าใช้งบประมาณเป็นค่าตอบแทนของตัวแทน สทศ. มาก อาจเนื่องมาจาก ศูนย์สอบและสนามสอบของการทดสอบทางการศึกษาระดับชาตินั้นพื้นฐานชั้นประถมศึกษาปีที่ 6 และชั้นมัธยมศึกษาปีที่ 3 ส่วนใหญ่ไม่มีปัญหาเกี่ยวกับความโปร่งใส ซึ่งสอดคล้องกับงานวิจัยทั้งของเมธาพร หิรัญวงศ์ และ สุกัญญา ไกรนรา (2556) และผลงานวิจัยของรุ่งฤดี กล้าหาญ (2557) ที่ว่า กระบวนการดำเนินการบริหารการทดสอบของศูนย์สอบและสนามสอบ ทั้งระยะก่อนการจัดสอบ ระหว่างการจัดสอบ และหลังเสร็จสิ้นการสอบ เป็นไปตามมาตรฐานของสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) ส่วนระยะเวลาในการดำเนินการมีความเหมาะสมอยู่ในระดับมาก เนื่องจากการทดสอบทางการศึกษาระดับชาตินั้นพื้นฐาน มีกำหนดเวลาการจัดสอบที่ชัดเจน

3. การประเมินกระบวนการของโครงการจัดหาตัวแทน สทศ. ประจำปีสามสอบพบว่า กระบวนการของโครงการในภาพรวมมีความเหมาะสมอยู่ในระดับมากที่สุด เนื่องจากสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) มีนโยบายเกี่ยวกับการจัดสอบตามมาตรฐานการทดสอบทางการศึกษาแห่งชาติ และความโปร่งใสที่ชัดเจน รวมถึงกระบวนการทำงานของคณะทำงานในโครงการไม่ซับซ้อน ซึ่งสอดคล้องกับงานวิจัยทั้งของ วิสัยพรณ์ เสรีวัฒน์ (2555) ชนิศกานต์ ละครวงศ์ (2556) และกมลนันท์ บุญกล้า (2559) ซึ่งได้ทำการประเมินโครงการด้านการศึกษาระดับการศึกษาขั้นพื้นฐาน พบว่า เมื่อการกำหนดวัตถุประสงค์ที่สอดคล้องกับวัตถุประสงค์ และนโยบาย อยู่ในระดับมาก แล้วผลของการประเมินกระบวนการอยู่ในระดับมาก เช่นกัน จากเหตุผลดังกล่าวจึงทำให้มีการกำหนดการในแต่ละกิจกรรมชัดเจน หลักเกณฑ์การสรรหาและแต่งตั้งคณะทำงานตัวแทน สทศ. มีความชัดเจน สามารถปฏิบัติได้วิธีการสรรหาและแต่งตั้งคณะทำงานตัวแทน สทศ. ที่ให้ประธานคณะทำงานเป็นผู้สรรหาที่มีความเหมาะสม มีการจัดเตรียมอุปกรณ์ ได้แก่ บัตรประจำตัว หนังสือแจ้งหัวหน้าสนามสอบ แบบรายงานการปฏิบัติงาน และคู่มือการจัดสอบระดับสนามสอบ ให้กับตัวแทน สทศ. ประจำปีสามสอบครบถ้วน วิธีการกำกับติดตามการปฏิบัติหน้าที่ของตัวแทน สทศ. ประจำปีสามสอบ โดยใช้แบบรายงานการปฏิบัติงานมีความเหมาะสมช่องทางการติดต่อเพื่อประสานงานระหว่างหน่วยงานและระหว่างบุคคล ด้วยโทรศัพท์ และอีเมล มีความสะดวก ติดต่อได้ง่าย การประชาสัมพันธ์โครงการจัดหาตัวแทน สทศ. ประจำปีสามสอบ ในการประชุมชี้แจงศูนย์สอบมีความเหมาะสม

4. การประเมินผลผลิตของโครงการจัดหาตัวแทน สทศ. ประจำปีสามสอบ พบว่ามหาวิทยาลัยที่รับผิดชอบโครงการได้จัดส่งตัวแทน สทศ. ไปประจำปีสามสอบของการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน ชั้นประถมศึกษาปีที่ 6 และชั้นมัธยมศึกษาปีที่ 3 ได้มากกว่าร้อยละ 80 เนื่องจากมีระยะเวลาในการเตรียมจัดหาบุคลากรที่เหมาะสม รวมถึงมหาวิทยาลัยที่รับผิดชอบโครงการมีเครือข่ายที่ดี สำหรับการดำเนินการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน ชั้นประถมศึกษาปีที่ 6 และชั้นมัธยมศึกษาปีที่ 3 ปีการศึกษา 2559-2560 ไม่พบการดำเนินงานที่ไม่เป็นไปตามมาตรฐานการทดสอบทางการศึกษาแห่งชาติ และความไม่โปร่งใสในการจัดการทดสอบ และปีการศึกษา 2561 พบการดำเนินการจัดการทดสอบของสนามสอบที่ไม่เป็นไปตามแนวปฏิบัติในคู่มือการจัดสอบจำนวน 3 เรื่อง นอกจากนี้การทดสอบมีความโปร่งใส แต่สนามสอบบางแห่งยังมีจุดอ่อนเรื่องการกำกับ ดูแล ให้คณะกรรมการระดับสนามสอบปฏิบัติตามแนวทางปฏิบัติในช่วงหลังการสอบ จำนวน 1 เรื่อง ซึ่งอาจเป็นเพราะบทบาทของตัวแทน สทศ. เป็นเพียงผู้สังเกตและรายงานผลการสังเกตนั้นมายังสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) ซึ่งตัวแทน สทศ. ไม่มีอำนาจตัดสินใจในสนามสอบ ดังที่ระบุไว้ในเอกสารประกอบการประชุมตัวแทน สทศ. ประจำปีสามสอบ O-NET ชั้นประถมศึกษาปีที่ 6 และชั้นมัธยมศึกษาปีที่ 3 ปีการศึกษา 2560 (สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน), 2561 : 16)

ข้อเสนอแนะจากผลการวิจัยและการนำไปใช้

1. ข้อเสนอแนะสำหรับ สทศ.

1.1 ผู้บริหารโครงการจัดหาตัวแทน สทศ. ประจําสนามสอบควรทบทวนด้านปัจจัยนำเข้า ได้แก่ เรื่องความเหมาะสมของบุคลากร ที่ควรเป็นผู้มีวิทยุติและคุณวุฒิและงบประมาณที่มีต้นทุนสูงเมื่อเทียบกับผลผลิตที่ได้รับซึ่งอาจจะส่งตัวแทน สทศ. ไปเฉพาะสนามสอบที่มีปัญหาหรือมีความเสี่ยง

1.2 สทศ.ควรมีการเน้นย้ำและกำกับคณะกรรมการระดับสนามสอบให้ปฏิบัติตามคู่มือ รวมถึงสร้างความตระหนักผ่านการประชาสัมพันธ์ในรูปแบบต่าง ๆ เช่น วิทยุทัศน์ แผ่นพับ โครงการพัฒนาบุคลากร การประชุมชี้แจงโดยศูนย์สอบ เป็นต้น

2. ข้อเสนอแนะสำหรับหน่วยงานที่รับผิดชอบโครงการจัดหาตัวแทน สทศ.

2.1 การสื่อสารเพื่อสร้างความสัมพันธ์ที่ดีระหว่างมหาวิทยาลัยที่รับผิดชอบโครงการจัดหาตัวแทน สทศ. ประจําสนามสอบกับสนามสอบจะทำให้การดำเนินงานบรรลุเป้าหมาย และลุล่วงไปด้วยดี

3. ข้อเสนอแนะสำหรับเจ้าหน้าที่ที่เกี่ยวข้อง

3.1 เจ้าหน้าที่ที่เกี่ยวข้องควรศึกษาวิธีการทำงานจากคู่มือการปฏิบัติงานให้เข้าใจก่อนที่จะเริ่มงาน เพื่อป้องกันการดำเนินงานที่ไม่เป็นไปตามแนวปฏิบัติ

ข้อเสนอแนะในการวิจัยครั้งต่อไป

1. ควรมีการใช้วิจัยเชิงคุณภาพ เพื่อให้ได้ผลเชิงลึกมาใช้ในการปรับปรุงและพัฒนาโครงการ
2. ควรเพิ่มการศึกษาผลกระทบของโครงการที่มีต่อการปฏิบัติงานของคณะทำงานทั้งหน่วยงานที่รับผิดชอบโครงการ ศูนย์สอบ และสนามสอบ

เอกสารอ้างอิง

ภาษาไทย

- กมลนันท์ บุญกล้า. (2559). การประเมินโครงการเศรษฐกิจพอเพียงของโรงเรียนบ้านหนองปลาชี
สำนักงานเขตพื้นที่การศึกษาประถมศึกษาฉะเชิงเทรา เขต 2. ปรินญาณินพนธ์ กศ.ม.
(สาขาวิชาการบริหารการศึกษา). ชลบุรี: บัณฑิตวิทยาลัย มหาวิทยาลัยบูรพา. อัดสำเนา.
- ชนิดดา บุปผามาศ.(2557). การประเมินโครงการระบบโครงข่ายการเรียนรู้ไร้พรมแดน โรงเรียน
ในสังกัดสำนักงานเขตพื้นที่การศึกษามัธยมศึกษา เขต 8. ปรินญาณินพนธ์ ศษ.ม. (สาขา
วิชาพัฒนศึกษา). กรุงเทพฯ: บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร. อัดสำเนา.
- ชนิศกานต์ ละครวงศ์.(2556).การประเมินโครงการพัฒนาโรงเรียนขนาดเล็กต้นแบบการจัด
การเรียนการสอนแบบคละชั้น ของสำนักงานเขตพื้นที่การศึกษาจันทบุรี เขต 1.
ปรินญาณินพนธ์ ค.ม. (การบริหารการศึกษา). จันทบุรี: บัณฑิตวิทยาลัย มหาวิทยาลัย
ราชภัฏรำไพพรรณี. อัดสำเนา.
- พิชิต ฤทธิ์จัญญ. (2557). เทคนิคการประเมินโครงการ. พิมพ์ครั้งที่ 2. กรุงเทพฯ: บริษัท แฮสส์ ออฟ
เคอร์มิสท์ จำกัด.
- เมธพร หิรัญวงศ์ และสุกัญญา ไกรนรา. (2556). กระบวนการจัดการสอบ V-NET ที่มีประสิทธิภาพ
ของศูนย์สอบมหาวิทยาลัยเทคโนโลยีราชมงคลศรีวิชัย วิทยาเขตตรัง. รายงานการวิจัย.
มหาวิทยาลัยเทคโนโลยีราชมงคลศรีวิชัย วิทยาเขตตรัง
- รุ่งฤดี กล้าหาญ. (2557). การประเมินมาตรฐานการบริหารการทดสอบของสถาบันทดสอบ
ทางการศึกษาแห่งชาติ (องค์การมหาชน) ประจำปีการศึกษา 2557 :ศึกษารณีสวนสอบ
มหาวิทยาลัยศรีนครินทรวิโรฒ. รายงานการวิจัยทุนอุดหนุน, สถาบันทดสอบทางการ
ศึกษาแห่งชาติ (องค์การมหาชน).
- วิไลพรรณ เสรีวัฒน์. (2555). การประเมินเชิงระบบโครงการโรงเรียนมาตรฐานสากล. ปรินญา
ณินพนธ์ ปร.ด. (สาขาวิชาการบริหารการศึกษา). ขอนแก่น: บัณฑิตวิทยาลัยมหาวิทยาลัย
ขอนแก่น. อัดสำเนา.
- สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน).(2559).รายงานผลการดำเนินการ
จัดการทดสอบทางศึกษาระดับชาติขั้นพื้นฐาน (O-NET) ชั้นประถมศึกษาปีที่ 6 และ
ชั้นมัธยมศึกษาปีที่ 3 ปีการศึกษา 2559. กรุงเทพฯ: ม.ป.พ.
- สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน). (2560).รายงานผลการดำเนินการ
จัดการทดสอบทางศึกษาระดับชาติขั้นพื้นฐาน (O-NET) ชั้นประถมศึกษาปีที่ 6 และ
ชั้นมัธยมศึกษาปีที่ 3 ปีการศึกษา 2560. กรุงเทพฯ: ม.ป.พ.
- สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน).(2560).เอกสารประกอบการประชุม
ตัวแทน สทศ. ประจำปีสอบ O-NET ชั้นประถมศึกษาปีที่ 6 และชั้นมัธยมศึกษาปีที่
3 ปีการศึกษา 2560. กรุงเทพฯ:ม.ป.พ.

- สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน). (2560). คู่มือการจัดสอบทางการศึกษา ระดับชาติขั้นพื้นฐาน ชั้นประถมศึกษาปีที่ 6 และชั้นมัธยมศึกษาปีที่ 3 ปีการศึกษา 2561 สำหรับศูนย์สอบกรุงเทพฯ: ม.ป.พ.
- สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน). (2560). คู่มือการจัดสอบทางการศึกษา ระดับชาติขั้นพื้นฐานชั้นประถมศึกษาปีที่ 6 และชั้นมัธยมศึกษาปีที่ 3 ปีการศึกษา 2561 สำหรับหัวหน้าสนามสอบและกรรมการกลาง. กรุงเทพฯ : ม.ป.พ.
- สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน). (2561). รายงานผลการดำเนินการ โครงจัดทำตัวแทน สทศ. ประจำปีสนามสอบ ปีการศึกษา 2561. กรุงเทพฯ : ม.ป.พ.
- สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน). (2561). รายงานผลการดำเนินการ จัดการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) ชั้นประถมศึกษาปีที่ 6 และ ชั้นมัธยมศึกษาปีที่ 3 ปีการศึกษา 2561. กรุงเทพฯ: ม.ป.พ.
- สุชาติ ประสิทธิ์รัฐสินธุ์. (2554). การประเมินโครงการ : หลักการและการประยุกต์. พิมพ์ครั้งที่ 2. กรุงเทพฯ: ห้างหุ้นส่วนจำกัดสามลดา.
- สมหวัง พิธิยานุวัฒน์. (2558). ความรู้พื้นฐานสำหรับการประเมินโครงการทางการศึกษา. ในรวมบทความ ทาง การประเมินโครงการชุดรวมบทความ เล่มที่ 4. สมหวัง พิธิยานุวัฒน์ (บรรณาธิการ). หน้า 101-121. พิมพ์ครั้งที่ 9. กรุงเทพฯ: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- สมหวัง พิธิยานุวัฒน์. (2559). วิธีวิทยาการประเมิน : ศาสตร์แห่งคุณค่า. พิมพ์ครั้งที่ 6. กรุงเทพฯ: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- สำราญ มีแจ้ง. (2558). การประเมินโครงการทางการศึกษา : ทฤษฎีและปฏิบัติ. กรุงเทพฯ: โรงพิมพ์แห่ง จุฬาลงกรณ์มหาวิทยาลัย.
- อาสนี นิสาและ. (2555). การประเมินโครงการระบบดูแลช่วยเหลือนักเรียนในโรงเรียนแถมจาก วิชาคุณสรณ์ สำนักงานเขตบึงกุ่ม กรุงเทพมหานคร. ปริญญาโท วิชา ม. (สาขาวิทยาการ ประเมิน). กรุงเทพฯ: บัณฑิตวิทยาลัยมหาวิทยาลัยศรีนครินทรวิโรฒ. อุดมศึกษา.

ภาษาอังกฤษ

- Harvard Business Review Press. (2016). HBR Guide to Project Management. USA: Harvard Business School Publishing Corporation.
- Stufflebeam, Daniel L. (1971). The Relevance of the CIPP Evaluation Model for Educational Accountability. Atlantic City : NJ.
- Stufflebeam, Daniel L; and Shinkfield, Anthony J. (2007). Evaluation Theory, Models, & Application. San Francisco : John Wilwy.

การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับ การทดสอบระดับชาติวิชาภาษาอังกฤษ

Computerized Adaptive Testing for The Ordinary National Educational Test
(O-NET) in The Subject of English Language

ณภัทร ชัยมงคล

บทคัดย่อ

การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์เป็นการทดสอบที่นำกระบวนการทางคอมพิวเตอร์เข้ามาช่วยในการคัดเลือกข้อสอบตามระดับความสามารถของผู้สอบแต่ละคน ซึ่งหากนำระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์มาใช้ในการทดสอบระดับชาติ จะช่วยให้ผลการทดสอบมีความแม่นยำมากยิ่งขึ้น งานวิจัยนี้จึงได้พัฒนาระบบการทดสอบแบบปรับเหมาะเพื่อนำมาใช้ในการเตรียมความพร้อมการสอบระดับชาติวิชาภาษาอังกฤษ การนำระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์เข้ามาช่วยเตรียมความพร้อมในการทดสอบระดับชาติ จะช่วยให้ผู้เรียนทราบระดับความสามารถของตนเองและนำผลที่ได้ไปใช้พัฒนาตนเอง ก่อนเข้ารับการทดสอบจริง การวิจัยนี้พัฒนาองค์ประกอบของการทดสอบแบบปรับเหมาะ 4 องค์ประกอบ ได้แก่

- 1) จุดเริ่มต้นในการทดสอบ
- 2) การประมาณค่าความสามารถของผู้สอบ
- 3) การคัดเลือกข้อสอบ
- 4) การยุติการทดสอบ และพัฒนารูปแบบรายงานผลที่สะท้อนข้อมูลป้อนกลับ

ได้แก่ รายงานคะแนนสังเกตได้และรายงานระดับความสามารถของผู้สอบ ผลการตรวจสอบคุณภาพระบบโดยผู้เชี่ยวชาญด้วยแบบประเมินระบบแบบอิงมาตรฐาน พบว่า ระบบมีคุณภาพอยู่ในระดับมากที่สุดทุกด้าน

คำสำคัญ: ระบบการทดสอบแบบปรับเหมาะ การทดสอบระดับชาติ วิชาภาษาอังกฤษ

Abstract

Computerized adaptive testing (CAT) uses computer processes to select items which match with the ability of examinees. The application of CAT in the Ordinary National Educational Test (O-NET) will make the test process more efficient. Therefore, this research developed a CAT for O-NET to prepare Thai students for the actual O-NET testing in English. The students will be aware of their abilities and then improve on them before the actual test. The results from the system development consist of 4 phases

- 1) initial stage
- 2) ability estimation
- 3) item selection and
- 4) stopping criteria.

This report will contain both the raw score and abilities of each examinee. Finally, the quality of the system was evaluated by experts that the system is of high quality in every dimension.

Keywords: Computerized Adaptive Testing, Ordinary National Educational Test, English subject

บทนำ

การทดสอบแบบดั้งเดิมใช้แบบทดสอบฉบับเดียวกับผู้สอบทุกคน ซึ่งอาจส่งผลต่อความถูกต้องแม่นยำในการวัด เนื่องจากแบบทดสอบฉบับหนึ่ง ประกอบด้วยข้อสอบหลายข้อที่มีระดับความยากแตกต่างกัน ส่งผลให้ผู้สอบบางคนได้ทำข้อสอบที่ยากหรือง่ายเกินไป ดังนั้น การทดสอบในอุดมคติจึงมีแนวความคิดว่า ข้อสอบแต่ละข้อควรมีความยากที่เหมาะสมกับความสามารถของแต่ละบุคคล และเรียกการทดสอบนั้นว่า การทดสอบแบบปรับเหมาะ (adaptive testing or tailored testing) แนวคิดเรื่องการทดสอบแบบปรับเหมาะ เริ่มนำมาใช้ในการทดสอบเขาวินิจฉัยด้วยแบบสอบของบินเน็ต (Binet) โดยผู้สอบจะได้ทำข้อสอบครั้งละหนึ่งข้อ โดยเริ่มจากข้อแรก หากข้อสอบข้อแรกยากเกินไป ข้อถัดไปก็จะง่ายลงเพื่อให้เหมาะสมกับระดับเขาวินิจฉัยของผู้สอบ และเมื่อทำข้อสอบไปเรื่อย ๆ จนถึงข้อที่ยากมากขึ้นจนผู้สอบไม่สามารถทำข้อสอบได้ถูกเป็นจำนวนหนึ่ง จึงยุติการทดสอบ (ศิริชัย กาญจนวาสี, 2555)

ต่อมา เมื่อเทคโนโลยีคอมพิวเตอร์เริ่มเข้ามามีบทบาทเพิ่มมากขึ้น David Weiss จึงได้นำคอมพิวเตอร์เมนเฟรม (mainframe computer) มาใช้กับการทดสอบแบบปรับเหมาะ และเรียกว่าการทดสอบแบบปรับเหมาะ โดยใช้คอมพิวเตอร์ (computerized adaptive testing: CAT) ซึ่งเป็นที่นิยมใช้ในการประเมินผลทางการศึกษาในประเทศสหรัฐอเมริกา เช่น การทดสอบเพื่อศึกษาต่อระดับบัณฑิตศึกษา (Graduate Record Examination: GRE) (Chang, 2004; Linden, 2008) การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ (CAT) เป็นวิธีการทดสอบที่ผสมผสานทฤษฎีการวัดและเทคโนโลยีเข้าด้วยกันเพื่อให้การวัดมีความแม่นยำมากยิ่งขึ้น และมีองค์ประกอบที่สำคัญ 3 องค์ประกอบ ได้แก่

- (1) จุดเริ่มต้นในการทดสอบ
- (2) กระบวนการคัดเลือกข้อสอบและการประมาณค่าความสามารถ
- (3) เกณฑ์ที่ใช้ในการยุติการทดสอบ

ซึ่งหลักการสำคัญของการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ คือ กระบวนการคัดเลือกข้อสอบ (item selection) ที่แตกต่างกันไปตามวิธีการออกแบบการทดสอบ โดยอาศัยคุณลักษณะของข้อสอบตามทฤษฎีการตอบสนองข้อสอบ (item response theory) ได้แก่ ค่าความยาก ค่าอำนาจจำแนก และค่าความน่าจะเป็นในการเดา (Chang, 2004, 2015; Phankokkrud, 2012; ศิริชัย กาญจนวาสี, 2555)

ปัจจุบันมีการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ (CAT) มาใช้อย่างแพร่หลาย ดังนั้น เพื่อให้การพัฒนาผู้เรียนเป็นรูปธรรมมากยิ่งขึ้น ควรมีการนำระบบการทดสอบแบบปรับเหมาะมาใช้กับทดสอบระดับชาติ เพื่อประโยชน์ต่อผู้เรียนและวงการศึกษา ซึ่ง O-NET (Ordinary National Educational Test) คือ การทดสอบทางการศึกษาระดับชาตินี้ขึ้นพื้นฐานเป็นการทดสอบวัดความรู้และความคิดของนักเรียนชั้นประถมศึกษาปีที่ 6 ชั้นมัธยมศึกษาปีที่ 3 และชั้นมัธยมศึกษาปีที่ 6 เป็นการประเมินตามมาตรฐานการเรียนรู้ในหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551 การทดสอบประกอบด้วย 5 วิชา ได้แก่ คณิตศาสตร์

วิทยาศาสตร์ ภาษาอังกฤษ ภาษาไทย และสังคมศึกษา ศาสนาและวัฒนธรรม ซึ่งการทดสอบในระดับชั้นมัธยมศึกษาปีที่ 6 ถือว่าเป็นการทดสอบที่สำคัญยิ่งต่อนักเรียน เนื่องจากนักเรียนจำเป็นต้องใช้ความรู้พื้นฐานเพื่อนำไปใช้ศึกษาต่อในระดับมหาวิทยาลัย โดยวิชาที่มีความจำเป็นต่อการศึกษาในระดับสูง คือ วิชาภาษาอังกฤษ เพราะเมื่อศึกษาต่อในระดับสูงการค้นคว้าหาความรู้ใหม่ ๆ จำเป็นต้องใช้ภาษาอังกฤษ วิชาภาษาอังกฤษจึงถือเป็นวิชาที่มีความสำคัญในการพัฒนาองค์ความรู้ในระดับสูงของผู้เรียน และมีผลต่อการพัฒนาประเทศให้ทัดเทียมกับนานาอารยประเทศ จากสถิติผลการทดสอบย้อนหลังปี 2558-2560 พบว่า คะแนนเฉลี่ยวิชาภาษาอังกฤษอยู่ในเกณฑ์ต่ำ (สถาบันทดสอบทางการศึกษาแห่งชาติ, 2562) ดังนั้นจึงควรเร่งผลักดันให้มีการพัฒนาสมรรถนะการเรียนรู้ของนักเรียนในรายวิชาภาษาอังกฤษ เพื่อให้ นักเรียนมีศักยภาพทางภาษาที่สูงขึ้นทัดเทียมกับนานาอารยประเทศ และรองรับการเติบโตขององค์ความรู้ใหม่ ๆ ในอนาคต

ด้วยเหตุนี้ ผู้วิจัยจึงมีความสนใจพัฒนาระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการทดสอบระดับชาติ วิชาภาษาอังกฤษ โดยระบบการทดสอบที่พัฒนาขึ้นนี้สามารถสะท้อนข้อมูลป้อนกลับให้กับผู้สอบได้ทันที ซึ่งมีสารสนเทศเพื่อพัฒนาการเรียนรู้ของผู้สอบในรายวิชาภาษาอังกฤษ เพื่อให้ นักเรียนนำข้อมูลป้อนกลับที่ได้รับไปใช้ในการเตรียมตัวและวางแผนการทดสอบก่อนการทดสอบจริง ซึ่งจะช่วยยกระดับมาตรฐานคะแนนการทดสอบระดับชาติให้สูงขึ้นและเป็นไปตามเกณฑ์มาตรฐาน

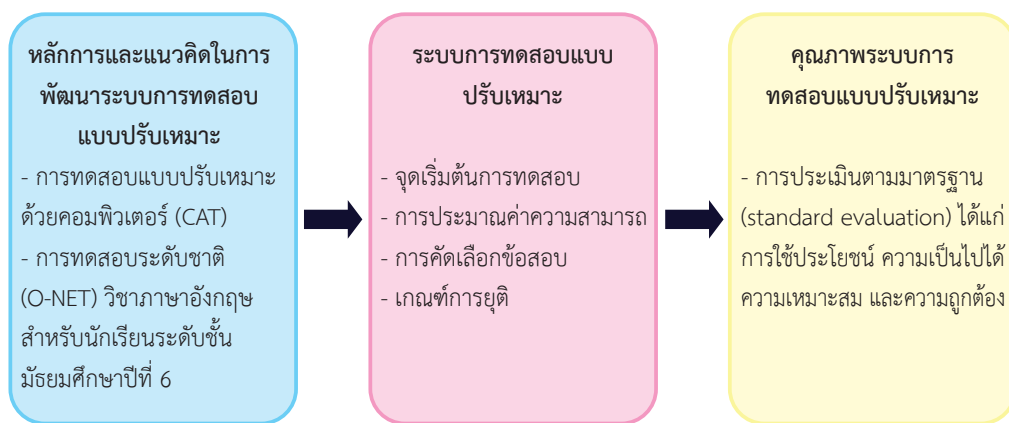
ความหมายและหลักการของการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์

การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ (computerized adaptive testing: CAT) หมายถึง การทดสอบที่ใช้วิธีการทางคอมพิวเตอร์เข้ามาช่วยในการคัดเลือกข้อสอบ และประมาณค่าความสามารถของผู้สอบ ผู้สอบแต่ละคนจึงได้รับแบบทดสอบที่มีความยากง่ายแตกต่างกันขึ้นกับระดับความสามารถของผู้สอบแต่ละคน (ทัศนศิริพันธ์ สว่างบุญ และคณะ, 2555) การใช้คอมพิวเตอร์ช่วยทำให้คะแนนการทดสอบมีความถูกต้องในเชิงสถิติเพิ่มขึ้น แนวคิดของการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์นี้ เริ่มจากการนำมาใช้กับแบบวัดเชาว์ปัญญาของ Binet-Simon ในปี 1905 โดยคอมพิวเตอร์จะคัดเลือกข้อคำถามให้เหมาะสมกับอายุสมองของผู้สอบ ซึ่งอาศัยการอ้างอิงจากอายุสมองของผู้สอบที่ได้มาจากการตอบข้อคำถามข้อก่อนหน้า และจะเลือกข้อคำถามให้ผู้สอบตอบจนเพียงพอต่อการประมาณค่าอายุสมองที่แท้จริงของผู้สอบได้ ความเป็นจริงแนวคิดการทดสอบแบบปรับเหมาะมีมานานแล้ว ดังเห็นได้จากการทดสอบปากเปล่า (oral test) ที่ผู้จัดสอบสามารถเลือกคำถามให้เหมาะสมกับความสามารถของผู้สอบได้ตามระดับความรู้จากการสอบถามและพูดคุย (Linden, 2000)

สำหรับการคัดเลือกข้อคำถามให้เหมาะสมกับความสามารถของผู้สอบนั้น มีแนวทางในการคัดเลือกข้อสอบสำหรับผู้สอบแต่ละคน โดยอ้างอิงจากผลการตอบข้อสอบข้อที่ผ่านมาในการทดสอบโดยทั่วไปผู้สอบจะได้เริ่มทำข้อสอบข้อแรกที่มีค่าความยากระดับปานกลาง จากนั้น

ระบบคอมพิวเตอร์จะทำการประมาณค่าความสามารถของผู้สอบจากผลการตอบข้อสอบแล้ว คัดเลือกข้อสอบข้อถัดไปที่มีความยากและอำนาจจำแนกที่เหมาะสมกับระดับความสามารถของผู้สอบ โดยมีหลักการคือหากผู้สอบตอบข้อสอบข้อที่ผ่านมาถูกต้อง ผู้สอบจะได้ข้อสอบข้อถัดไปที่มีค่าความยากมากขึ้น แต่หากผู้สอบตอบข้อสอบผิด ผู้สอบจะได้ข้อสอบข้อถัดไปที่มีค่าความยากลดลง จนกระทั่งการประมาณค่าความสามารถของผู้สอบมีความเชื่อถือได้คืออยู่ในเกณฑ์ของความคลาดเคลื่อนที่กำหนดจึงยุติการทดสอบ (Čisar, Radosav, Markoski, Pinter, & Čisar, 2010; Gómez, Laria, & Hernández, 2019; Harrison et. al., 2019; ทศน์ศิริรินทร์ สว่างบุญ และคณะ, 2555)

การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ (computer adaptive testing: CAT) เป็นเทคโนโลยีทางการทดสอบ ที่มีจุดแข็งในเชิงทฤษฎี คือ เป็นการทดสอบที่ผู้สอบจะได้ทำข้อสอบที่มีความยากตามระดับความสามารถของตน โดยที่ข้อสอบไม่ง่ายหรือยากจนเกินไป เหมือนกับการทดสอบแบบดั้งเดิมที่ใช้กระดาษในการทดสอบ และจำนวนข้อสอบที่ใช้ในการทดสอบก็ลดลง โดยที่ไม่สูญเสียความแม่นยำในการประมาณค่าความสามารถของผู้สอบ โดยจะเลือกข้อสอบแต่ละข้อจากคลังข้อสอบ ด้วยกระบวนการทางคอมพิวเตอร์ภายหลังการประมาณค่าความสามารถของผู้สอบบนพื้นฐานของผลการตอบข้อสอบจากข้อก่อนหน้า ซึ่งเป็นการควบคุมคุณภาพที่มีประสิทธิภาพ เนื่องจากสามารถลดความคลาดเคลื่อนที่เกิดจากคนในการจัดการข้อสอบได้ อีกทั้งยังควบคุมเงื่อนไขต่าง ๆ เช่น ความครอบคลุมของเนื้อหา (content coverage) และอัตราการเปิดเผยของข้อสอบ (item exposure) เป็นต้น (Zheng & Chang, 2015; ทศน์ศิริรินทร์ สว่างบุญ และคณะ, 2555) นอกจากนี้ การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ยังสามารถนำมาปรับใช้กับการประเมินความรู้ของนักเรียนได้ ซึ่งจากผลการวิจัยพบว่า การทดสอบแบบปรับเหมาะ โดยใช้คอมพิวเตอร์เป็นเครื่องมือที่มีประสิทธิภาพเมื่อนำมาประยุกต์ใช้ในการประเมิน (Andjelic & Cekerevac, 2014) และการทดสอบแบบปรับเหมาะ โดยใช้คอมพิวเตอร์ยังเป็นวิธีการทดสอบที่สามารถนำมาใช้ในการประเมินความก้าวหน้าของนักเรียนทั้งแบบรายบุคคลและแบบกลุ่มได้อย่างมีประสิทธิภาพ (Weiss, 2004)



กรอบแนวคิดการวิจัย

วิธีดำเนินการวิจัย

การพัฒนาระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการทดสอบระดับชาติวิชาภาษาอังกฤษใช้ระเบียบวิธีวิจัยและพัฒนา (research and development) ในการศึกษา โดยมีขั้นตอนการวิจัยทั้งสิ้น 2 ระยะ ได้แก่ ระยะที่ 1 การพัฒนาระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการทดสอบระดับชาติวิชาภาษาอังกฤษ และระยะที่ 2 การประเมินผลระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการทดสอบระดับชาติวิชาภาษาอังกฤษโดยมีรายละเอียดแต่ละระยะดังนี้

ระยะที่ 1 การพัฒนาระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการทดสอบระดับชาติ วิชาภาษาอังกฤษ

สำหรับระยะแรกจะดำเนินการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องกับการทดสอบแบบปรับเหมาะหลายขั้นตอน หลังจากนั้นเลือกองค์ประกอบที่มีความเหมาะสมสำหรับนำมาใช้ในการพัฒนาระบบ และศึกษาองค์ประกอบแต่ละองค์ประกอบให้มีความชัดเจนมากยิ่งขึ้น และดำเนินการศึกษารูปแบบการรายงานผลข้อมูลป้อนกลับที่มีความเหมาะสมกับการทดสอบระดับชาติวิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 โดยแบ่งการดำเนินงานออกเป็น 3 ตอน ได้แก่

- 1) ศึกษาองค์ประกอบของการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์
- 2) ออกแบบและพัฒนาระบบการทดสอบและรูปแบบการรายงานผล
- 3) ตรวจสอบคุณภาพของระบบการทดสอบ

ระยะที่ 2 การประเมินผลระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการทดสอบระดับชาติวิชาภาษาอังกฤษ

สำหรับระยะที่สองดำเนินการตรวจสอบคุณภาพแบบประเมินระบบด้านความตรงตามเนื้อหาด้วยค่าดัชนี IOC จากผู้เชี่ยวชาญด้านการวัดและประเมินผลการศึกษา จำนวน 5 ท่าน และปรับแก้แบบประเมินตามคำแนะนำของผู้เชี่ยวชาญ หลังจากนั้นดำเนินการส่งให้ผู้เชี่ยวชาญด้านการทดสอบและผู้เชี่ยวชาญด้านเทคโนโลยีสารสนเทศจำนวน 5 ท่าน ตรวจสอบความเหมาะสมการทำงานของระบบโดยใช้แบบประเมินระบบแบบอิงมาตรฐาน (standard evaluation) โดยระยะนี้แบ่งการดำเนินงานออกเป็น 2 ตอน ได้แก่ ได้แก่

- 1) การตรวจสอบคุณภาพแบบประเมินระบบ
- 2) การประเมินระบบการทดสอบ

ผลการวิจัย

ระยะที่ 1 การพัฒนาระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการทดสอบระดับชาติวิชาภาษาอังกฤษ

ผลการพัฒนาระบบ พบว่า องค์ประกอบของการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการทดสอบระดับชาติวิชาภาษาอังกฤษ ประกอบด้วย 4 องค์ประกอบ ได้แก่

- 1) จุดเริ่มต้นในการทดสอบ (initial stage)
- 2) การประมาณค่าความสามารถของผู้สอบ (ability estimation)
- 3) การคัดเลือกข้อสอบ (item selection)
- 4) จุดยุติการทดสอบ (stopping criteria) และรูปแบบรายงานผลที่สะท้อนข้อมูลป้อนกลับ ได้แก่ รายงานคะแนนสังเกตได้และรายงานระดับความสามารถของผู้สอบ โดยรายละเอียดของแต่ละขั้นตอนมีดังนี้

1) จุดเริ่มต้นการทดสอบ (initial stage) สำหรับการวิจัยในครั้งนี้จะทำการทดสอบจำนวน 3 ชั้น ในรูปแบบ 1-2-3 กล่าวคือ ข้อสอบในชั้นแรกประกอบด้วยข้อสอบระดับปานกลาง (Moderate: M) ข้อสอบในชั้นที่สองประกอบด้วยข้อสอบระดับยาก (Hard: H) และง่าย (Easy: E) และข้อสอบในชั้นที่สามประกอบด้วยข้อสอบระดับง่าย (E) ปานกลาง (M) และยาก (H) โดยข้อสอบในชั้นแรก มีจำนวน 25 ข้อ โดยผู้วิจัยเป็นผู้ดำเนินการจัดชุดข้อสอบไว้ จำนวน 5 ชุด ซึ่งชุดข้อสอบในชุดแรกจะถูกเรียกขึ้นมาใช้อย่างสุ่ม (random) จากทั้งหมด 5 ชุด โดยข้อสอบในชุดแรกเป็นข้อสอบที่มีค่าความยาก (b) อยู่ในระดับปานกลาง (ค่า b อยู่ระหว่าง -1 ถึง 1) ซึ่งเป็นไปตามแนวทางการทดสอบแบบปรับเหมาะหลายขั้นตอน เพื่อให้การประมาณค่าความสามารถมีความแม่นยำมากยิ่งขึ้น

สำหรับข้อสอบที่นำมาใช้ในการวิจัยนี้มาจากผลการวิเคราะห์ข้อสอบวิชาภาษาอังกฤษของการทดสอบระดับชาติสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 ซึ่งใช้ในการทดสอบระหว่างปี 2558-2561 จำนวน 320 ข้อ จากสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) และเมื่อนำมาวิเคราะห์คุณภาพข้อสอบตามเกณฑ์ คือ ค่าความยาก (b) อยู่ในช่วง -4 ถึง 4 ค่าอำนาจจำแนก (a) อยู่ในช่วง 0 ถึง α และความน่าจะเป็นในการเดา (c) ไม่เกิน 0.50 ตามทฤษฎีการตอบสนองข้อสอบ (item response theory: IRT) พบว่า มีข้อสอบที่ผ่านเกณฑ์ซึ่งสามารถนำมาใช้ในการวิจัยครั้งนี้ จำนวน 179 ข้อ

2) การประมาณค่าความสามารถของผู้สอบ (ability estimation) สำหรับการประมาณค่าความสามารถของผู้สอบจากการศึกษาพบว่า การประมาณค่าด้วยความเป็นไปได้สูงสุด (maximum likelihood estimation: MLE) เป็นวิธีการประมาณค่าซึ่งมีความแม่นยำมากที่สุด แต่จะมีความคลาดเคลื่อนสูงเมื่อผู้สอบตอบข้อสอบถูกหรือผิดหมด ซึ่งวิธีที่ประมาณค่าได้แม่นยำสำหรับกรณีผู้สอบตอบข้อสอบถูกหรือผิดหมด คือวิธีการประมาณค่าด้วยค่าคาดหวังภายหลัง (expected a posterior: EAP) ดังนั้นจึงมีการรวมวิธีจะทำการประมาณค่าทั้ง 2 วิธีเข้าด้วยกัน ได้แก่ วิธีการประมาณค่าด้วยความเป็นไปได้สูงสุด (maximum likelihood

estimation: MLE) และวิธีการประมาณค่าด้วยค่าคาดหวังภายหลัง (expected a posterior: EAP) ซึ่งวิธีการประมาณค่าด้วยค่าคาดหวังภายหลังจะใช้สำหรับกรณีผู้สอบตอบข้อสอบถูก หรือ ผิดทุกข้อ หากผู้สอบตอบข้อสอบถูกและผิดผสมกันจะใช้วิธีการประมาณค่าด้วยวิธีการประมาณค่า ด้วยความเป็นไปได้สูงสุด ดังนั้น ในการวิจัยนี้จึงใช้วิธีการประมาณค่าความสามารถทั้งสองวิธีร่วมกัน เพื่อให้ค่าความสามารถที่ได้มีความถูกต้องและแม่นยำมากยิ่งขึ้น (Bock & Mislevy, 1982; Hambleton & Swaminathan, 1985; Zheng & Chang, 2015)

วิธีการประมาณค่าด้วยความเป็นไปได้สูงสุด (maximum likelihood estimation: MLE) จะใช้แนวทางของ Hambleton & Swaminathan (1985) โดยมีสมการที่ใช้ในการคำนวณ ดังนี้

$$P(U_{ij}=1|\theta_i, a_j, b_j, c_j) = c_j + (1 - c_j) \frac{\exp(a_j(\theta_i - b_j))}{1 + \exp(a_j(\theta_i - b_j))}$$

$$L_j(\theta) = \prod_{j=1}^J [P_j(\theta)]^{x_j} [1 - P_j(\theta)]^{1 - x_j}$$

$$h_m = \frac{D[r_m - \sum P_i(\theta_m)]}{-D^2 \sum P_i(\theta_m) Q_i(\theta_m)}$$

$$\theta_{m+1} = \theta_m - h_m$$

โดยสมการที่ใช้ในการคำนวณจะไม่นำค่าปรับสเกล ($D=1.7$) มาใช้เนื่องจาก Camilli (1994) ได้ทำการศึกษาการใช้ค่าปรับสเกลพบว่าให้ผลการประมาณค่าความสามารถไม่แตกต่างจากการใช้ค่าปรับสเกล

วิธีการประมาณค่าความสามารถด้วยค่าคาดหวังภายหลัง (expected a posterior: EAP) ใช้แนวทางของ Bock and Mislevy (1982) โดยมีสมการที่ใช้ในการคำนวณดังนี้

$$\bar{\theta}_j = \frac{\sum_{k=1}^q x_k L_j(x_k) \cdot w(x_k)}{\sum_{k=1}^q L_j(x_k) \cdot w(x_k)}$$

3) การคัดเลือกข้อสอบ (item selection) สำหรับการคัดเลือกข้อสอบเป็นสิ่งที่สำคัญที่สุดในกระบวนการทดสอบแบบปรับเหมาะ ซึ่งในงานวิจัยนี้ใช้การเลือกข้อสอบตามระดับความสามารถที่สอดคล้องกับค่าความยาก นอกจากนี้ยังคำนึงถึงความครอบคลุมของเนื้อหา (content balance) และควบคุมการเปิดเผยของข้อสอบไม่ให้มีจำนวนมากจนเกินไป (item exposure control) ซึ่งการเลือกข้อสอบในคลังข้อสอบสำหรับการวิจัยครั้งนี้สามารถเลือกข้อสอบข้อเดิมได้ไม่เกินร้อยละ 20 กล่าวคือหากมีการทดสอบ จำนวน 100 ครั้ง จะสามารถเลือกข้อสอบข้อเดิมมาใช้ได้ไม่เกิน 20 ครั้ง

4) จุดยุติการทดสอบ (stopping criteria) ในการวิจัยครั้งนี้กำหนดความยาวของแบบสอบแบบคงที่ (fixed length) เนื่องจากเป็นการทดสอบแบบหลายขั้นตอนซึ่งต้องกำหนดชุดข้อสอบไว้ล่วงหน้า โดยใช้ข้อสอบทั้งสิ้นจำนวน 50 ข้อ ต่อผู้สอบ 1 คน

สำหรับตัวอย่างรายงานผลจากระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลป้อนกลับในการทดสอบระดับชาติวิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 แสดงดังภาพที่ 1



ระยะที่ 2 การประเมินผลระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการทดสอบระดับชาติวิชาภาษาอังกฤษ

ผลการตรวจสอบคุณภาพแบบประเมินระบบแบบอิงมาตรฐานที่พัฒนาขึ้นโดยผู้เชี่ยวชาญด้วยค่าดัชนี IOC พบว่า ค่าถामประเมินที่มีค่าดัชนีความสอดคล้องมากกว่า 0.5 มีจำนวน 14 ข้อ จาก 20 ข้อ ทั้งนี้ดำเนินการปรับแก้แบบประเมินตามคำแนะนำผู้เชี่ยวชาญ ก่อนนำไปใช้งาน

สำหรับการตรวจสอบคุณภาพระบบจากผู้เชี่ยวชาญด้วยแบบประเมินระบบแบบอิงมาตรฐาน (standard evaluation) โดยแบ่งประเด็นในการตรวจสอบออกเป็น 4 ด้าน ได้แก่ ด้านความมีประโยชน์ (utility) ด้านความเป็นไปได้ในการนำไปใช้ (feasibility) ด้านความเหมาะสม (propriety) และด้านความถูกต้อง (accuracy) โดยการประเมินใช้ผู้เชี่ยวชาญทางด้านการวัดและประเมินผลจำนวน 3 ท่าน และผู้เชี่ยวชาญทางด้านเทคโนโลยีสารสนเทศจำนวน 2 ท่าน

ผลการประเมิน พบว่า ผู้เชี่ยวชาญมีความเห็นว่าระบบมีความเหมาะสมทั้ง 4 ด้าน โดยอันดับหนึ่ง ได้แก่ ด้านความเป็นไปได้ในการนำไปใช้ ($M = 4.96$, $SD = 0.15$) รองลงมาคือ ด้านความถูกต้อง ($M = 4.88$, $SD = 0.18$) ด้านความมีประโยชน์ ($M = 4.84$, $SD = 0.22$) และ ด้านความเหมาะสม ($M = 4.83$, $SD = 0.14$) นอกจากนี้ ผู้เชี่ยวชาญได้ให้ข้อคิดเห็นว่า ระบบมีความชัดเจนเข้าใจง่ายและเป็นการพัฒนาระบบที่คำนึงถึงมาตรฐาน สามารถเข้าถึงผู้ใช้งาน รวมถึงให้ข้อมูลย้อนกลับเพื่อการพัฒนา ก่อให้เกิดประโยชน์ต่อส่วนรวม สำหรับส่วนที่ควรปรับปรุงเพิ่มเติม คือ ควรปรับรายละเอียดการลงทะเบียนเข้าใช้งานของผู้สอบ โดยให้ผู้สอบกรอกข้อมูลส่วนตัวก่อนเข้าสู่ระบบเพื่อให้เบ็ดเสร็จในจุดเดียว นอกจากนี้ ควรมีระบบการเรียกคืนรหัสผ่านอีเมลกรณีผู้ใช้งานลืมรหัสผ่าน และควรให้ผู้ใช้งานสามารถอัปเดตรูปภาพตนเองในระบบได้ สำหรับในอนาคตอาจเชื่อมโยงระบบเข้ากับสื่อโซเชียลมีเดียต่าง ๆ เช่น เฟซบุ๊ก ไลน์ เพื่อให้ผู้ใช้งานเข้าถึงระบบการทดสอบได้ง่ายขึ้นและอาจออกแบบระบบการทดสอบให้รองรับอุปกรณ์การใช้งานอื่นนอกจากเครื่องคอมพิวเตอร์ เช่น แท็บเล็ต สมาร์ทโฟน

สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ

สรุปผลการวิจัย มีดังนี้

1. การวิจัยนี้พัฒนาระบบการทดสอบแบบปรับเหมาะหลายด้วยคอมพิวเตอร์สำหรับการทดสอบระดับชาติวิชาภาษาอังกฤษ จำนวน 4 องค์ประกอบ ได้แก่

- 1) จุดเริ่มต้นในการทดสอบ
- 2) การประมาณค่าความสามารถของผู้สอบ
- 3) การคัดเลือกข้อสอบ
- 4) การยุติการทดสอบ

รวมถึงพัฒนารูปแบบรายงานผลที่สะท้อนข้อมูลป้อนกลับ ได้แก่ รายงานคะแนนสังเกตได้และรายงานระดับความสามารถของผู้สอบ

2. ผลการตรวจสอบคุณภาพระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการทดสอบระดับชาติวิชาภาษาอังกฤษ แบ่งออกเป็น 2 ส่วน ได้แก่

1) ผลการตรวจสอบคุณภาพแบบประเมินระบบ

2) ผลการประเมินระบบการทดสอบผลการตรวจสอบคุณภาพแบบประเมินระบบสำหรับแบบประเมินระบบแบบอิงมาตรฐานที่พัฒนาขึ้นจำนวน 20 ข้อ

พบว่า คำถามประเมินที่มีค่าดัชนีความสอดคล้องมากกว่า 0.5 มีจำนวน 14 ข้อ และผู้เชี่ยวชาญได้แนะนำให้ดำเนินการปรับแก้รายการประเมินบางส่วนซึ่งผู้วิจัยได้ดำเนินการปรับแก้ก่อนนำไปใช้งานเรียบร้อยแล้ว

ผลการประเมินคุณภาพระบบด้วยแบบประเมินระบบแบบอิงมาตรฐาน พบว่าผู้เชี่ยวชาญมีความเห็นว่าระบบมีความเหมาะสมทั้ง 4 ด้าน โดยอันดับหนึ่งได้แก่ ด้านความเป็นไปได้ในการนำไปใช้ รองลงมาคือ ด้านความถูกต้อง ด้านความมีประโยชน์ และด้านความเหมาะสมตามลำดับ

อภิปรายผลการวิจัย

จากผลการวิจัยสามารถนำมาอภิปรายได้ 2 ประเด็น ได้แก่

1) การพัฒนาระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

2) การประเมินระบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

1. การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ เป็นการทดสอบที่ผู้สอบทำแบบสอบต่างฉบับ ตามระดับความสามารถแต่ยังคงให้ความแม่นยำจากผลการสอบ โดยมีหลักการคือผู้สอบจะเริ่มทำข้อสอบจากข้อที่มีระดับความยากปานกลาง หากผู้สอบตอบข้อสอบผิดข้อสอบข้อถัดไปจะง่ายลงในทางตรงข้าม หากผู้สอบตอบข้อสอบถูกข้อสอบข้อถัดไปจะยากขึ้น ผู้สอบจะทำการสอบไปเรื่อย ๆ จนกระทั่งระดับความสามารถของผู้สอบคงที่หรืออยู่ในเกณฑ์ความคลาดเคลื่อนที่ยอมรับได้ การทดสอบจะยุติลง (Cisar et al., 2010; ศิริชัย กาญจนวาสี, 2555) จากผลการวิจัย พบว่า การ ทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการทดสอบระดับชาติวิชาภาษาอังกฤษ มีองค์ประกอบที่สำคัญ ได้แก่

1) จุดเริ่มต้นในการทดสอบ

2) การประมาณค่าความสามารถของผู้สอบ

3) การคัดเลือกข้อสอบ

4) การยุติการทดสอบ

ซึ่งสอดคล้องกับ Chaimongkol, Pasiphol, and Kanjanawasee (2016) โดยระบบการทดสอบที่พัฒนาขึ้นสามารถนำไปใช้ในการเตรียมความพร้อมของนักเรียนก่อนเข้าสอบ การทดสอบระดับชาติวิชาภาษาอังกฤษเนื่องจากข้อสอบที่นำมาใช้ในการพัฒนาระบบเป็นข้อสอบระดับชาติวิชาภาษาอังกฤษที่เคยนำมาใช้ในการทดสอบจริง นอกจากนี้ ระบบยังสามารถให้ข้อมูลป้อนกลับให้กับผู้สอบได้ทันทีเพื่อให้ผู้สอบสามารถนำข้อมูลป้อนกลับไปใช้ในการพัฒนาตนเองต่อไป

2. ประเด็นสำคัญในการพัฒนาระบบ คือ การตรวจสอบคุณภาพ ดังนั้น จึงต้องพิจารณากระบวนการประเมินให้มีความเหมาะสมกับสิ่งที่มุ่งประเมิน โดยการวิจัยครั้งนี้ได้นำวิธีการประเมินระบบแบบอิงมาตรฐาน ตามแนวคิดของ Stufflebeam โดยครอบคลุม 4 ประเด็น ได้แก่ ด้านความมีประโยชน์ ด้านความเป็นไปได้ในการนำไปใช้ ด้านความเหมาะสม และด้านความถูกต้อง (ศิริชัย กาญจนวาสี, 2554) ซึ่งจะช่วยประเมินเบื้องต้นได้ว่าระบบที่พัฒนาขึ้นสามารถนำไปใช้งานได้จริงหรือไม่

ข้อเสนอแนะในการนำผลวิจัยไปใช้

สถาบันทดสอบการศึกษาแห่งชาติ (องค์การมหาชน) ซึ่งเป็นองค์กรด้านการทดสอบระดับชาติ สามารถนำไปใช้เผยแพร่ให้ผู้เรียนเตรียมความพร้อมในการทดสอบระดับชาติ วิชาภาษาอังกฤษ เพื่อยกระดับความสามารถของผู้เรียนในวิชาภาษาอังกฤษของประเทศให้เป็นไปตามมาตรฐาน

ข้อเสนอแนะในการวิจัยครั้งต่อไป

1. การวิจัยครั้งนี้มีข้อสอบที่ผ่านเกณฑ์เพียง 179 ข้อ ซึ่งข้อสอบส่วนใหญ่เป็นข้อสอบชุด (testlets) จึงทำให้โอกาสที่ผู้สอบจะพบข้อสอบซ้ำมีอัตราสูง ด้วยเหตุนี้ งานวิจัยในอนาคตจึงควรเพิ่มจำนวนข้อสอบเพื่อบรรจุลงในคลังข้อสอบ เพื่อให้เนื้อหาในการทดสอบแบบปรับเหมาะ มีระดับความยากที่หลากหลายและครอบคลุมเนื้อหาการทดสอบระดับชาติวิชาภาษาอังกฤษ

2. การวิจัยครั้งนี้ยังไม่ได้ดำเนินการปรับเทียบคะแนนสอบของแบบสอบแต่ละฉบับ เนื่องจากข้อจำกัดของจำนวนข้อสอบที่ผ่านเกณฑ์ ดังนั้น งานวิจัยครั้งต่อไปควรมีการปรับเทียบคะแนนของแบบสอบแต่ละปีเพื่อให้คลังข้อสอบของระบบการทดสอบแบบปรับเหมาะมีความเท่าเทียมกันมากยิ่งขึ้น

3. ปัจจุบันการวินิจฉัยผู้เรียน (diagnosis) เป็นประเด็นที่เข้ามามีบทบาทสำคัญในการพัฒนาผู้เรียน ดังนั้น ในการวิจัยครั้งต่อไปควรพัฒนาระบบการทดสอบแบบปรับเหมาะเพื่อวินิจฉัยผู้เรียน เพื่อให้ผู้เรียนได้รับข้อสอบสำหรับการวินิจฉัยที่ตรงกับความสามารถกับผู้เรียน และหลังจากการทดสอบเสร็จสิ้น สามารถให้ข้อมูลป้อนกลับให้กับผู้เรียนเพื่อนำไปใช้ในการพัฒนานตนเองได้ทันที

เอกสารอ้างอิง

- ทัศนศิริพันธ์ สว่างบุญ, ศิริเดช สุชีวะ, ศิริชัย กาญจนวาสี, & Muraki, E. (2555). ประสิทธิภาพการวัดของการทดสอบแบบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์. วารสารวิจัย, 25(1), 27-45.
- ศิริชัย กาญจนวาสี. (2555). ทฤษฎีการทดสอบแนวใหม่ (Modern Test Theory) (พิมพ์ครั้งที่ 4). กรุงเทพมหานคร: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน). (2562). ค่าสถิติพื้นฐานผลการทดสอบ O-NET ชั้นมัธยมศึกษาปีที่ 6 ปีการศึกษา 2557 - 2560. สืบค้นเมื่อวันที่ 13 มิถุนายน 2562, จากเว็บไซต์ <https://www.niets.or.th/uploads/editor/files/Download/ช่วงคะแนน%20ม.6.pdf>
- Andjelic, S., & Cekerevac, Z. (2014). CAT model with personalized algorithm for evaluation of estimated student knowledge. Educational Information Technology, 2014(19), 173-191.
- Bock, R. D., & Mislevy, R. J. (1982). Adaptive EAP Estimation of Ability in a Microcomputer Environment. Applied Psychological Measurement, 6(4), 431-444.
- Čisar, S. M., Radosav, D., Markoski, B., Pinter, R., & Čisar, P. (2010). Computer Adaptive Testing of Student Knowledge. Acta Polytechnica Hungarica, 7(4), 139-152.
- Gómez, D. D, Laria, J. C., & Hernández, D. R. (2019). Computerized adaptive test and decision trees: A unifying approach. Expert Systems with Applications, 117, 358-366.
- Hambleton, R. K., & Swaminathan, H. (1985). Item Response Theory: Principles and Applications. MA, U.S.A.: Kluwer Academic Publishers.
- Harrison, C. J., Geerards, D., Ottenhof, M. J., Klassen, A. F., Riff, K. W., Swan, M. C., Pusic, A. L., & Sidey-Gibbons, C. J. (2019). Computerised adaptive testing accurately predicts CLEFT-Q scores by selecting fewer, more patient-focused questions, Journal of Plastic, Reconstructive & Aesthetic Surgery, 72(11), 1819-1824.
- Linden, W. J. v. d. (2000). Optimal assembly of tests with item sets. Applied Psychological Measurement, 24, 225-240.
- Weiss, D. J. (2004). Computerized Adaptive Testing for Effective and Efficient Measurement in Counseling and Education. Measurement and Evaluation in Counseling and Development, 37(2), 70-84.
- Zheng, Y., & Chang, H.-H. (2015). On-the-Fly Assembled Multistage Adaptive Testing. Applied Psychological Measurement, 39(2), 1-15.

การวิเคราะห์ผลการสอบ O-NET เพื่อพัฒนาบทเรียนออนไลน์แบบเปิด
เพื่อมหาชนสำหรับแก้ไขข้อบกพร่องที่คลาดเคลื่อนและข้อผิดพลาด
ทางคณิตศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 3

Analysis of O-NET Results to Develop a Massive Online Open Course for Cor-
rections of Misconceptions and Mathematical Errors of
Mathayom Suksa 3 Students

ฉวีวรรณ แก้วไพเราะ

บุญทอง บุญทวี

ลือชัย ทิพรังศรี

กนกรัตน์ คุณะสารพันธ์

พงศ์ระพี แก้วไพเราะ

บทคัดย่อ

การวิจัยครั้งนี้ได้ทำการวิเคราะห์ข้อบกพร่องที่คลาดเคลื่อนและข้อผิดพลาดทางคณิตศาสตร์ระดับมัธยมศึกษาปีที่ 3 ที่เข้าสอบ O-NET วิชาคณิตศาสตร์ในปีการศึกษา 2560 จำนวน 643,742 คน และปีการศึกษา 2561 จำนวน 645,575 คน จากการวิเคราะห์ผลการสอบพบว่าผู้เข้าสอบมีข้อบกพร่องที่คลาดเคลื่อนและข้อผิดพลาดทางคณิตศาสตร์มากกว่า 50% ทุกสาระการเรียนรู้เพราะมีคะแนนต่ำกว่า 50% ทุกสาระ จากนั้นผู้วิจัยได้นำประเด็นที่พบข้อผิดพลาดมากไปพัฒนาเป็นบทเรียนออนไลน์แบบเปิดเพื่อมหาชน (MOOC) ซึ่งประกอบไปด้วย 4 สาระการเรียนรู้วิชาคณิตศาสตร์ระดับมัธยมศึกษาตอนต้น ได้แก่ จำนวนและการดำเนินการพีชคณิต การวัดและเรขาคณิต การวิเคราะห์ข้อมูลและความน่าจะเป็น โดยทุกบทเรียนจะมีข้อสอบก่อนเรียนและหลังเรียน รวมทั้งการแนะนำการคิดแก้โจทย์ปัญหาในข้อสอบ O-NET ปีการศึกษา 2560 และ 2561 บทเรียนทั้งหมดเปิดให้เรียนบนแพลตฟอร์มของโครงการมหาวิทยาลัยไซเบอร์ไทย หรือ Thai MOOC ตั้งแต่วันที่ 15 ธันวาคม พ.ศ. 2562 ถึงวันที่ 18 มกราคม พ.ศ. 2563 โดยมีนักเรียนผู้เข้าร่วมจำนวน 68-198 คน ในแต่ละเนื้อหาหลังจากนักเรียนได้ใช้บทเรียนออนไลน์แล้วพบว่า มีจำนวนร้อยละของนักเรียนที่แก้ไขข้อบกพร่องที่คลาดเคลื่อนและข้อผิดพลาดทางคณิตศาสตร์ตั้งแต่ร้อยละ 63.64 (บทเรียนเรื่องสถิติ) ถึงร้อยละ 96.00 (บทเรียนเรื่องโจทย์ปัญหาสมการเชิงเส้นสองตัวแปร) และการใช้บทเรียนออนไลน์แบบเปิดเพื่อมหาชนมีผลต่อการพัฒนานักเรียนที่มีค่าดัชนีประสิทธิผลเท่ากับ 0.8807 แสดงว่าสามารถพัฒนานักเรียนให้มีผลสัมฤทธิ์ในการเรียนบทเรียนนี้ ร้อยละ 88.07 ซึ่งสูงกว่าเกณฑ์การยอมรับที่ต้องสูงกว่าหรือเท่ากับ 0.50 ผลการสอบของนักเรียนหลังเรียนด้วยบทเรียน

ออนไลน์แบบเปิดเพื่อมหาชนทุกบทเรียนสูงกว่าก่อนเรียนอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 สำหรับผลการวิจัยด้านความพึงพอใจพบว่า มีผู้ตอบแบบสอบถาม 180 คน มีความพึงพอใจมากที่สุดในเรื่อง ภาพคมชัด/กราฟฟิกมีขนาด สีสเหมาะสม 48.89% เนื้อหามีความทันสมัยทันต่อการเปลี่ยนแปลง 47.22% และ เสียงที่ใช้ในสื่อเหมาะสม ชัดเจน 47.22% มีความพึงพอใจมากในเรื่อง ลำดับเรื่องเข้าใจง่าย ชวนติดตาม 45.00% รูปแบบกิจกรรม/แบบทดสอบเหมาะสม ชัดเจน 44.44% นอกจากนั้นผู้เข้าเรียนมีความประสงค์ที่จะให้พัฒนาบทเรียนแบบเปิดอย่างต่อเนื่องและกำหนดช่วงเวลาของการเข้าถึงบทเรียนออนไลน์ให้มากขึ้น

คำสำคัญ: มโนทัศน์ที่คลาดเคลื่อนทางคณิตศาสตร์ บทเรียนออนไลน์แบบเปิดเพื่อมหาชน
ดัชนีประสิทธิผลโครงการมหาวิทยาลัยไซเบอร์ไทย

Abstract

This study analyzed misconceptions and mathematical errors of Mathayom Suksa 3 students from the O-NET test results of 643,742 students in the academic year 2017 and 645,575 students in academic year 2018. , The analysis of test result found that the students had misconceptions and errors in taking the tests more than 50% of the students in all learning strands with low scores less than 50% in all of strands. Based on the results of the O-NET test analysis, the researchers developed a massive open online course (MOOC) for correcting most students' misconceptions and mathematical errors consisting of 4 learning strands at junior high-school level: Numbers and Operations, Algebra, Measurement and Geometry, and Data Analysis and Probability. Every lesson has pretest and posttest including suggestions for problem-solving in the O-NET tests academic year 2017 and academic year 2018. All online lessons were open published on the platform of Thai Cyber University Project or Thai MOOC during 15 December 2019 and 18 January 2020. There were 68 up to 198 students participated in different online lessons. After the students learned the massive open online course, the results of correcting students' misconceptions and errors show that the percentage of students who correct misconceptions and errors starting from 63.64% (Lesson on Statistics) up to 96.00 % (Lesson on Two-Variable Linear Equation Problems) and using the massive open online course impacts on students' development with the effectiveness index of 0.8807, indicating that the massive open online course was able to develop students to be successful in learning this lesson 88.07 percent which is higher than the criterion for acceptance which must be higher than or equal to 0.50. The students' posttest score after learning via the massive open online course every lesson was higher than the pretest score before learning at the .05 level of statistical significance. For the research result of satisfaction was found that 180 respondents were most satisfied with 'Clear images/graphics with appropriate color and size' 48.89%, 'The content is modern catch up to the change' 47.22%, 'The sound used in the media is clear' 47.22%, 'Very satisfied with the order of the lesson is easy to understand and engaging to follow up' 45.00%, and 'The format of activities/ tests is clear and suitable' 44.44%. Also, learners wish to continually develop massive open online courses and provide more accessible.

Keywords: mathematical misconceptions, massive open online course, effectiveness index, Thai MOOC

ความเป็นมาและความสำคัญของปัญหา

การบริหารระบบการทดสอบระดับชาติของประเทศไทยได้ดำเนินการโดยสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) หรือเรียกชื่อย่อว่า สทศ. การทดสอบระดับชาติสำหรับประเมินคุณภาพการจัดการศึกษาตามหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน ได้ดำเนินการทุกสิ้นปีการศึกษา ซึ่งเรียกการทดสอบระดับชาตินี้ว่า O-NET (Ordinary National Education Test) สทศ.ได้จัดทำรายงานผลการสอบ O-NET เพื่อนำไปใช้ในการพัฒนาคุณภาพการเรียนการสอนในระดับต่าง ๆ (สถาบันทดสอบทางการศึกษาแห่งชาติ, 2560) ได้แก่

- (1) ระดับนักเรียน
- (2) ระดับผู้สอน
- (3) ระดับผู้บริหารสถานศึกษา
- (4) ระดับเขตพื้นที่การศึกษา
- (5) ระดับต้นสังกัด
- (6) ระดับประเทศ

สำหรับการวัดผลที่สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) ดำเนินการสอบคุณภาพผู้เรียนที่สิ้นสุดการเรียนของนักเรียนชั้นประถมศึกษาปีที่ 6 มัธยมศึกษาปีที่ 3 และมัธยมศึกษาปีที่ 6 ได้อิงมาตรฐานและตัวชี้วัดที่กำหนดไว้ในหลักสูตรและจัดทำรายงานให้ผู้เกี่ยวข้องทุกระดับนำไปพัฒนาการเรียนรู้ ได้แก่ นักเรียน สถานศึกษา เขตการศึกษา ผู้ปกครอง และนักการศึกษาที่สนใจทั่วไป ซึ่งผู้ที่เกี่ยวข้องสามารถนำผลไปพัฒนารูปแบบการวัดผลลักษณะต่าง ๆ ได้แก่

- (1) การวัดผลของการเรียนรู้ (Assessment of Learning)
 - (2) การวัดผลเพื่อการเรียนรู้ (Assessment for Learning)
 - (3) การวัดผลขณะเรียนรู้ (Assessment as Learning)
- ซึ่งมีลักษณะโดยรวมของการวัดผลที่แสดงในตารางที่ 1

ตารางที่ 1 ลักษณะโดยรวมของการวัดผล 3 แบบ (ของ การเรียนรู้ เพื่อ การเรียนรู้ และขณะเรียนรู้)

รูปแบบ	วัตถุประสงค์	จุดอ้างอิง	ผู้ทำหน้าที่วัดผล
การวัดผล ของ การเรียนรู้	เพื่อใช้แสดงระดับการเรียนรู้และส่งเสริมการเรียนรู้เป็นรายบุคคลเป็นต้น	เปรียบเทียบกับนักเรียนคนอื่น	ครูผู้สอน
การวัดผล เพื่อ การเรียนรู้	เพื่อได้ข้อมูลสารสนเทศสำหรับการตัดสินใจด้านการสอนของครู	มาตรฐานภายนอกหรือความคาดหวัง	ครูผู้สอน
การวัดผล ขณะ เรียนรู้	เพื่อติดตามตนเอง และแก้ไขหรือปรับปรุงตนเอง	เป้าหมายของนักเรียนและมาตรฐานภายนอก	นักเรียน

ที่มา: ปรับปรุงจาก Earl, L. (2003)

จากการพิจารณาผลการสอบรายวิชาคณิตศาสตร์ (รหัส 94) ชั้นมัธยมศึกษาปีที่ 3 ปีการศึกษา 2557 ถึง 2561 ในเบื้องต้นพบว่าผู้เข้าสอบที่มีจำนวนมากกว่า 600,000 คนในแต่ละปี มีคะแนนเฉลี่ยอยู่ที่ 26.30 ถึง 32.40 คะแนน ซึ่งอยู่ในระดับที่ต่ำกว่า 1 ใน 3 ของคะแนนเต็ม 100 คะแนน สะท้อนให้เห็นถึงปัญหาส่วนหนึ่งที่เกิดจากการที่นักเรียนมีมโนทัศน์ที่คลาดเคลื่อน หรือมีข้อผิดพลาดในกระบวนการคิดทางคณิตศาสตร์ ทำให้ไม่สามารถแก้โจทย์ปัญหาได้อย่างถูกต้อง ซึ่งกระบวนการวิเคราะห์ปัญหาจำเป็นต้องมีการวิเคราะห์ข้อสอบเป็นรายข้อโดยอาศัยข้อมูลทางสถิติ เพื่อช่วยในการพิจารณาหาข้อบกพร่องของนักเรียนที่ตอบผิด และนำไปพัฒนาเครื่องมือเพื่อช่วยแก้ไขมโนทัศน์ที่คลาดเคลื่อนและข้อผิดพลาดทางคณิตศาสตร์อย่างไรก็ตาม ผลการสอบของนักเรียนที่ได้คะแนนอยู่ระหว่าง 50.01 ถึง 100 คะแนน ในแต่ละปีไม่ถึงร้อยละ 10 ของผู้สอบทั้งหมด ตัวอย่างเช่น ผลการสอบปีการศึกษา 2560 จากรายงานของสถาบันทดสอบทางการศึกษาแห่งชาติ(องค์การมหาชน) พบว่า มีเพียง ร้อยละ 8.45 ของนักเรียนที่เข้าสอบทั้งหมด 643,720 คน ซึ่งแสดงให้เห็นถึงช่องว่างของความรู้ของนักเรียนที่เข้าสอบ O-NET ในแต่ละปี

จากการศึกษาเอกสารและงานวิจัยต่าง ๆ พบว่าปัจจัยที่ทำให้เกิดช่องว่างของความรู้ (knowledge gap) อาจเกิดจากสถานะเศรษฐกิจสังคม (socioeconomic status) ที่แต่ละคนมีโอกาสได้รับสารสนเทศที่แตกต่างกัน (sociology dictionary), และประกอบด้วยช่องว่าง 5 ด้าน (Hassan & Mahmoud, 2011) ได้แก่

- (1) ช่องว่างด้านการเตรียมตัว (preparation gap) ได้แก่ การเตรียมความรู้พื้นฐาน และทักษะที่จำเป็น
- (2) ช่องว่างด้านความเชื่อ (belief gap) ได้แก่ การให้ความสำคัญหรือเจตคติที่มีต่อความรู้ในเนื้อหา
- (3) ช่องว่างด้านเวลา (time gap) ได้แก่ การไม่มีเวลาเพียงพอในการแสวงหาหรือรับการถ่ายทอดความรู้
- (4) ช่องว่างด้านการสอน (teaching gap) ได้แก่ การใช้เทคนิคการสอนที่ไม่สอดคล้องกับความต้องการของผู้เรียน
- (5) ช่องว่างด้านผู้นำ (leadership gap) ได้แก่ ผู้บริหารหรือผู้กำหนดนโยบายด้านการศึกษา (Hassan & Mahmoud, 2011)

โครงการวิจัยครั้งนี้ได้นำผลการสอบรายวิชาคณิตศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 3 ปีการศึกษา 2560 ถึง ปีการศึกษา 2561 มาวิเคราะห์มโนทัศน์ที่คลาดเคลื่อน (Misconceptions) และข้อผิดพลาด (Errors) ในการเรียนรู้คณิตศาสตร์ (Mathematical Learning) และนำไปพัฒนาบทเรียนออนไลน์แบบเปิดเพื่อมหาชนสำหรับแก้ไขมโนทัศน์ที่คลาดเคลื่อน และข้อผิดพลาดทางคณิตศาสตร์ เพื่อให้นักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่มีอยู่เป็นจำนวนมากสามารถเรียนรู้ได้ตลอดเวลาและทุกสถานที่ และได้นำแผนยุทธศาสตร์ที่ 4 ที่เป็นกรอบแนวคิดตามแผนการศึกษาแห่งชาติ พ.ศ. 2560-2574 ที่มุ่งพัฒนาผู้เรียน ดังนี้

ยุทธศาสตร์ที่ 4: การสร้างโอกาส ความเสมอภาค และความเท่าเทียมทางการศึกษา

4.1 ผู้เรียนทุกคนได้รับโอกาสและความเสมอภาคในการเข้าถึงการศึกษาที่มีคุณภาพ มีตัวชี้วัดที่สำคัญ เช่น ดัชนีความเสมอภาคของอัตราการเข้าเรียนระดับการศึกษาขั้นพื้นฐานตามฐานะทางเศรษฐกิจและพื้นที่ลดลง ความแตกต่างระหว่างคะแนนเฉลี่ยผลการทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) ของนักเรียนระหว่างพื้นที่/ภาคการศึกษาในวิชาคณิตศาสตร์และภาษาอังกฤษลดลง เป็นต้น

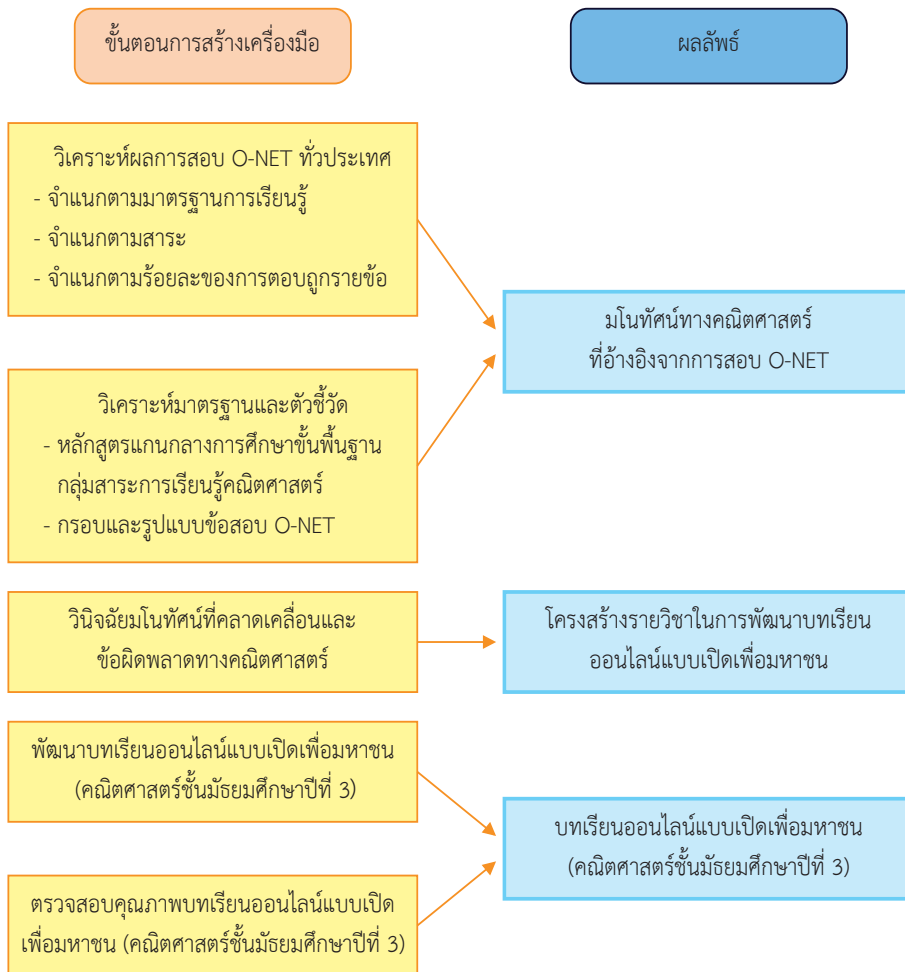
4.2 การเพิ่มโอกาสทางการศึกษาผ่านเทคโนโลยีดิจิทัลเพื่อการศึกษาสำหรับคนทุกช่วงวัย มีตัวชี้วัดที่สำคัญ เช่น มีระบบเครือข่ายเทคโนโลยีดิจิทัลเพื่อการศึกษาที่ทันสมัยสนองตอบความต้องการของผู้เรียนและผู้ให้บริการอย่างทั่วถึงและมีประสิทธิภาพ และสถานศึกษาทุกแห่งมีอินเทอร์เน็ตความเร็วสูงและมีคุณภาพ เป็นต้น

4.3 ระบบข้อมูลรายบุคคลและสารสนเทศทางการศึกษาที่ครอบคลุม ถูกต้องเป็นปัจจุบัน เพื่อการวางแผนการบริหารจัดการศึกษา การติดตามประเมิน และรายงานผล ... (กระทรวงศึกษาธิการ, 2560)

การสร้างบทเรียนออนไลน์แบบเปิดเพื่อมหาชน (Massive Online Open Course: MOOC) เป็นการสร้างโอกาสทางการเรียนรู้ผ่านเทคโนโลยีดิจิทัลที่เรียนรู้ระบบทางไกล (Distance Learning) โดยนำเสนอบทเรียนในระบบออนไลน์ ที่เปิดโอกาสให้บุคคลต่าง ๆ เข้าเรียนได้ทั่วประเทศ (หรือทั่วโลกเมื่อนำเสนอบทเรียนเป็นภาษาอังกฤษ) รายวิชาหรือบทเรียนแต่ละเรื่องจะรองรับผู้เรียนได้จำนวนมาก การพัฒนาบทเรียนออนไลน์ในระบบเปิดที่ดำเนินการโดยสำนักงานคณะกรรมการการอุดมศึกษาภายใต้โครงการ ‘Thai Cyber University Project’ ได้สร้างรายวิชาที่ออกแบบรองรับผู้เรียนที่สนใจรายวิชาต่าง ๆ โดยมีมาตรฐานการเรียนการสอน MOOC จำนวน 10 มาตรฐาน ได้แก่

- (1) โครงร่างรายวิชา
 - (2) ความพร้อมของบุคลากร
 - (3) การออกแบบการเรียนการสอน
 - (4) เนื้อหา
 - (5) สื่อการเรียนรู้
 - (6) การสื่อสาร
 - (7) ลิขสิทธิ์และทรัพย์สินทางปัญญา
 - (8) การสนับสนุนผู้เรียน
 - (9) ผลการจัดการเรียนรู้
 - (10) การปรับปรุงพัฒนา
- (ปราวินยา และ เสมอภาญจน์, 2560)

สำหรับสาระการเรียนรู้คณิตศาสตร์ที่นำมาพัฒนาบทเรียนออนไลน์แบบเปิดเพื่อมหาชนที่ใช้ในการวิจัยครั้งนี้ ได้มีขั้นตอนการพัฒนาก่อนนำไปเผยแพร่ให้นักเรียนได้เรียนรู้ดังแสดงไว้ในภาพที่ 1



ภาพที่ 1 ขั้นตอนการสร้างบทเรียนออนไลน์และผลลัพธ์

วัตถุประสงค์ของการวิจัย

1. เพื่อวิเคราะห์ข้อที่คลาดเคลื่อนและข้อผิดพลาดทางคณิตศาสตร์ระดับมัธยมศึกษาปีที่ 3 จากผลการสอบ O-NET ปีการศึกษา 2560 และ ปีการศึกษา 2561
2. เพื่อหาประสิทธิภาพของบทเรียนออนไลน์แบบเปิดเพื่อมหาชนสำหรับแก้ไขข้อที่คลาดเคลื่อนและข้อผิดพลาดทางคณิตศาสตร์ระดับมัธยมศึกษาปีที่ 3 ที่พัฒนาขึ้นโดยอ้างอิงจากผลการวิเคราะห์ข้อสอบ O-NET
3. เพื่อประเมินผลการใช้บทเรียนออนไลน์แบบเปิดเพื่อมหาชน
 - 3.1 การหาผลการแก้ไขข้อที่คลาดเคลื่อนและข้อผิดพลาด
 - 3.2 ประเมินความพึงพอใจของนักเรียนต่อการเรียนรู้จากบทเรียน

ขอบเขตของการวิจัย

ประชากร ได้แก่ นักเรียนชั้นมัธยมศึกษาปีที่ 3 ประจำปีการศึกษา 2562

กลุ่มตัวอย่าง ได้แก่ นักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่สมัครเข้าเรียนในบทเรียนออนไลน์แบบเปิดเพื่อมหาชน (Massive Online Open Course: MOOC) ที่สร้างขึ้นและเผยแพร่แบบสาธารณะ

เนื้อหาของบทเรียน ได้แก่ เนื้อหาที่อ้างอิงจากผลการวิเคราะห์ข้อสอบ O-NET ปีการศึกษา 2560 และ ปีการศึกษา 2561

ตัวแปรของการวิจัยประกอบด้วยตัวแปรดังนี้

ตัวแปรต้น คือ บทเรียนออนไลน์แบบเปิดเพื่อมหาชนที่พัฒนาจากการวิเคราะห์ผลการสอบ O-NET ของนักเรียนชั้นมัธยมศึกษาปีที่ 3 ปีการศึกษา 2560 และ ปีการศึกษา 2561

ตัวแปรตาม คือ

(1) ผลการใช้บทเรียนออนไลน์แบบเปิดเพื่อมหาชน ประกอบด้วย ร้อยละของนักเรียนที่สามารถแก้ไขข้อผิดพลาดที่คลาดเคลื่อนและข้อผิดพลาดทางคณิตศาสตร์ ที่มีคะแนนการสอบหลังเรียนตั้งแต่ 50% ขึ้นไป และประสิทธิผลของการใช้บทเรียน

(2) ความพึงพอใจของนักเรียนต่อการเรียนรู้จากบทเรียนออนไลน์แบบเปิดเพื่อมหาชน

วิธีดำเนินการวิจัย

การวิจัยครั้งนี้เป็นการวิจัยเชิงทดลอง แบบกึ่งทดลอง (Quasi-Experimental Research) ดำเนินการทดลองแบบแผนการวิจัย One-Group Pretest-Posttest Design เครื่องมือที่ใช้ทดลอง คือ บทเรียนออนไลน์แบบเปิดเพื่อมหาชน (Massive Online Open Course: MOOC) โดยอ้างอิงจากผลวิเคราะห์การสอบ O-NET รายวิชาคณิตศาสตร์ สำหรับนักเรียนมัธยมศึกษาปีที่ 3 หลังจากพัฒนาบทเรียนเสร็จแล้วได้นำไปทดลองใช้บนแพลตฟอร์มของ Thai MOOC และรวบรวมข้อมูลเพื่อหาผลการใช้และประสิทธิภาพของบทเรียน รวมทั้งประเมินผลความพึงพอใจของผู้ใช้บทเรียน

การเก็บรวบรวมข้อมูล

1) รวบรวมข้อมูลเกี่ยวกับผลการสอบ O-NET รายข้อที่นักเรียนตอบในแบบทดสอบ O-NET 2 ปีซ้อนหลัง (ปีการศึกษา 2560 และปีการศึกษา 2561) จากสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน)

2) รวบรวมข้อมูลคะแนนการสอบก่อนเรียนและหลังเรียนโดยใช้บทเรียนออนไลน์แบบเปิดเพื่อมหาชน (Massive Online Open Course: MOOC)

3) รวบรวมข้อมูลแสดงระดับความพึงพอใจในระบบออนไลน์

การวิเคราะห์ข้อมูล

การตรวจสอบคุณภาพเครื่องมือ

1) การหาคุณภาพของบทเรียนโดยใช้การหาค่าดัชนีความเที่ยงตรงของเนื้อหา (Content Validity Index: CVI) กำหนดค่าระดับความคิดเห็น 4 ระดับโดยพิจารณาจากอัตราส่วนระหว่างจำนวนข้อที่ผู้เชี่ยวชาญทั้งหมดให้ระดับ 3 หรือ 4 ต่อจำนวนข้อคิดเห็นทั้งหมด และใช้เกณฑ์การยอมรับ 0.80 (Davis, 1992)

2) การหาคุณภาพของแบบทดสอบ E-Testing โดยใช้ความเที่ยงตรงเชิงปรากฏ (Face Validity) โดยใช้การสัมภาษณ์ใน 4 ประเด็น ดังนี้

- (1) ความเหมาะสมของภาษาที่ใช้
- (2) ความเหมาะสมของมโนทัศน์ที่ต้องการทดสอบ
- (3) ความเหมาะสมของกราฟหรือภาพที่ใช้ในการทดสอบมโนทัศน์
- (4) ความเชื่อมโยงของข้อมูลที่ใช้ในการแก้ปัญหา

การหาผลของการใช้บทเรียน

1) การแก้ไขมโนทัศน์คลาดเคลื่อนและข้อผิดพลาด

ผลการแก้ไขมโนทัศน์คลาดเคลื่อนและข้อผิดพลาดทางคณิตศาสตร์คำนวณจากผลต่างของจำนวนนักเรียนที่มีคะแนนรวมต่ำกว่า 50% ของการสอบก่อนเรียนและหลังเรียน ถ้าจำนวนนักเรียนที่มีมโนทัศน์คลาดเคลื่อนและข้อผิดพลาดทางคณิตศาสตร์ของการสอบหลังเรียนลดลง แสดงว่าบทเรียนนี้สามารถแก้ไขมโนทัศน์และข้อผิดพลาดได้แล้วนำไปคิดเป็นร้อยละของนักเรียนที่สามารถแก้ไขมโนทัศน์คลาดเคลื่อนและข้อผิดพลาดก่อนเรียน โดยใช้สูตรดังนี้

ให้ N_1 แทนจำนวนนักเรียนที่มีคะแนนรวมก่อนเรียนน้อยกว่า 50%

N_2 แทนจำนวนนักเรียนที่มีคะแนนรวมหลังเรียนน้อยกว่า 50%

C แทนร้อยละของผลการแก้ไขมโนทัศน์คลาดเคลื่อนและข้อผิดพลาด

$$\text{สูตร } C = \frac{N_1 - N_2}{N_1} \times 100$$

2) หาประสิทธิภาพของบทเรียนโดยใช้ดัชนีประสิทธิผล (Effectiveness Index: E.I.) (Issel, 2000) จากคะแนนผลการสอบของนักเรียนและคะแนนเป้าหมาย

เกณฑ์การยอมรับต้องสูงกว่าหรือเท่ากับ 0.50 (เผด็จ กิจระการ, 2545) และหาผลสัมฤทธิ์ทางการเรียนของผู้เรียน โดยการทดสอบค่า t (t-test) เพื่อทดสอบความแตกต่างระหว่างก่อนเรียนและหลังเรียน

การหาระดับความพึงพอใจ

การหาระดับความพึงพอใจของจำนวนนักเรียนที่เข้าเรียนบทเรียนใช้แบบสำรวจออนไลน์ในระบบ Thai MOOC รหัสวิชา SSRU007 ชื่อวิชา คณิตพิชิต O-NET (ระดับม. ต้น) ซึ่งแบ่งเป็น 5 ระดับ คือ พึงพอใจมากที่สุด พึงพอใจ พึงพอใจปานกลาง พึงพอใจน้อย พึงพอใจน้อยที่สุด

ในการวิจัยครั้งนี้ ผู้วิจัยต้องใช้รายการในแบบสำรวจตามข้อตกลงของการนำบทเรียนขึ้นระบบออนไลน์แบบเปิด Thai MOOC ซึ่งผู้วิจัยนำข้อมูลเฉพาะรายการที่เกี่ยวข้องดังนี้

ส่วนที่ 1 แบบวัดความพึงพอใจที่มีต่อรูปแบบการจัดการเรียนการสอน ประกอบด้วย ประเด็นต่าง ๆ ได้แก่

- (1) เนื้อหามีความทันสมัย ทันต่อการเปลี่ยนแปลง
- (2) ลำดับเรื่องเข้าใจง่าย ชวนติดตาม
- (3) รูปแบบการนำเสนอสวยงาม เข้าใจง่าย
- (4) ตัวอย่าง/กรณีศึกษาเพียงพอ ชัดเจน
- (5) รูปแบบกิจกรรม/แบบทดสอบ เหมาะสม ชัดเจน
- (6) ภาพคมชัด/กราฟิกมีขนาด สีสเหมาะสม
- (7) เสียงที่ใช้ในสื่อเหมาะสม ชัดเจน

ส่วนที่ 2 ความพึงพอใจโดยรวมแบ่งเป็น 5 ระดับ คือ พึงพอใจมากที่สุด พึงพอใจ พึงพอใจปานกลาง พึงพอใจน้อย พึงพอใจน้อยที่สุด

ผลการวิจัย

ผลการวิเคราะห์การสอบ O-NET ของนักเรียนชั้นมัธยมศึกษาปีที่ 3

ผลการวิเคราะห์การสอบ O-NET ของนักเรียนชั้นมัธยมศึกษาปีที่ 3 ประกอบด้วย ประเด็นต่าง ๆ ได้แก่

- (1) ผลการสอบจำแนกตามมาตรฐานการเรียนรู้
- (2) ผลการสอบจำแนกตามสาระ
- (3) ผลการสอบจำแนกตามรายข้อ ผลการวิเคราะห์มีดังนี้

ประเด็นที่ 1 ผลการสอบจำแนกตามมาตรฐานการเรียนรู้

ผลการวิเคราะห์แต่ละมาตรฐานการเรียนรู้ที่เป็นคะแนนเฉลี่ยทั่วประเทศ พบว่า คะแนนเฉลี่ยของนักเรียนชั้นมัธยมศึกษาปีที่ 3 ในแต่ละมาตรฐานการเรียนรู้มีผลการวิเคราะห์ดังตารางที่ 2

ตารางที่ 2 ผลการสอบแยกตามมาตรฐานการเรียนรู้ของกลุ่มสาระการเรียนรู้คณิตศาสตร์

มาตรฐานการเรียนรู้	ปีการศึกษา 2560		ปีการศึกษา 2561	
	Mean	S.D.	Mean	S.D.
มาตรฐาน ค1.1 เข้าใจถึงความหลากหลายของการแสดงจำนวนและการใช้จำนวนในชีวิตจริง	17.39	28.81	24.46	32.66
มาตรฐาน ค1.2 เข้าใจถึงผลที่เกิดขึ้นจากการดำเนินการของจำนวนและความสัมพันธ์ระหว่างการดำเนินการต่างๆและสามารถใช้ในการดำเนินการในการแก้ปัญหา	28.82	28.37	25.11	30.82
มาตรฐาน ค 1.4 เข้าใจระบบจำนวนและคุณสมบัติเกี่ยวกับจำนวนไปใช้	41.46	49.26	20.64	40.47
มาตรฐาน ค 2.1 เข้าใจพื้นฐานเกี่ยวกับการวัด วัดและคาดคะเนขนาดของสิ่งที่ต้องการวัด	18.83	25.45	-	-
มาตรฐาน ค 2.2 แก้ปัญหาเกี่ยวกับการวัด	32.83	46.96	24.99	43.29

มาตรฐานการเรียนรู้	ปีการศึกษา 2560		ปีการศึกษา 2561	
	Mean	S.D.	Mean	S.D.
มาตรฐาน ค 3.1 อธิบายและวิเคราะห์รูปเรขาคณิตสองมิติและสามมิติ	32.40	35.16	65.27	35.19
มาตรฐาน ค 3.2 ใช้การนิยามภาพ (visualization) ใช้เหตุผลเกี่ยวกับปริภูมิ (spatial reasoning) และใช้แบบจำลองทางเรขาคณิต (geometric model) ในการแก้ปัญหา	23.85	24.32	27.79	26.37
มาตรฐาน ค 4.1 เข้าใจและวิเคราะห์แบบรูป (pattern) ความสัมพันธ์ และฟังก์ชัน	27.99 18.83	44.89 25.45	32.66 -	46.90
มาตรฐาน ค 4.2 ใช้นิพจน์ สมการ อสมการ กราฟ และตัวแบบเชิงคณิตศาสตร์ (mathematical model) อื่น ๆ แทนสถานการณ์ต่าง ๆ ตลอดจนแปลความหมาย และนำไปใช้แก้ปัญหา	30.45	23.09	25.34	22.03
มาตรฐาน ค 5.1 เข้าใจและใช้วิธีการทางสถิติในการวิเคราะห์ข้อมูล	30.14	28.71	32.90	27.60
มาตรฐาน ค 5.2 ใช้วิธีการทางสถิติและความรู้เกี่ยวกับความน่าจะเป็นในการคาดการณ์ได้อย่างสมเหตุสมผล	8.66	21.18	33.42	36.09
บูรณาการ	-	-	23.16	24.33

หมายเหตุ ข้อสอบที่ออกทุกข้อตามมาตรฐานเทียบเป็นคะแนนเต็ม 100 คะแนน

จากตารางที่ 2 แสดงให้เห็นว่าค่าเฉลี่ยของผลสอบของนักเรียนในปีการศึกษา 2560 มีคะแนนต่ำกว่า 50 คะแนนทุกมาตรฐาน แต่ปีการศึกษา 2561 ค่าเฉลี่ยของผลสอบในมาตรฐาน ค 3.1 อธิบายและวิเคราะห์รูปเรขาคณิตสองมิติและสามมิติ ได้ 65.27 สูงกว่า 50 คะแนน และมาตรฐาน ค 5.2 ใช้วิธีการทางสถิติและความรู้เกี่ยวกับความน่าจะเป็นในการคาดการณ์ได้อย่างสมเหตุสมผล มีคะแนนเฉลี่ยสูงกว่าผลสอบในปีการศึกษา 2560 อย่างไรก็ตาม คะแนนเฉลี่ยของผลสอบในปีการศึกษา 2561 ที่ต่ำกว่าปีการศึกษา 2560 มี 4 มาตรฐานดังนี้

มาตรฐาน ค 1.4 เข้าใจระบบจำนวนและนำเสนอสมบัติเกี่ยวกับจำนวนไปใช้

มาตรฐาน ค 2.2 แก้ปัญหาเกี่ยวกับการวัด

มาตรฐาน ค 4.2 ใช้นิพจน์ สมการ อสมการ กราฟ และตัวแบบเชิงคณิตศาสตร์ (mathematical model) อื่น ๆ แทนสถานการณ์ต่าง ๆ ตลอดจนแปลความหมาย และนำไปใช้แก้ปัญหา

สำหรับ มาตรฐาน ค 1.2 เข้าใจพื้นฐานเกี่ยวกับการวัด วัดและคาดคะเนขนาดของสิ่งที่ต้องการวัด ได้นำไปบูรณาการกับมาตรฐาน ค 2.2

ประเด็นที่ 2 ผลการสอบจำแนกตามสาระการเรียนรู้

ผลการวิเคราะห์แต่ละสาระที่เป็นคะแนนเฉลี่ยทั่วประเทศ พบว่า คะแนนเฉลี่ยของนักเรียนชั้นมัธยมศึกษาปีที่ 3 ในแต่ละสาระมีผลการวิเคราะห์ดังตารางที่ 3

ตารางที่ 3 ผลการสอบแยกตามสาระของกลุ่มสาระการเรียนรู้คณิตศาสตร์

สาระการเรียนรู้	คะแนนเต็ม	ปีการศึกษา 2560		ปีการศึกษา 2561	
		Mean	S.D.	Mean	S.D.
จำนวนและการดำเนินการ	100	32.40	35.16	65.27	35.19
การวัด	100	23.49	23.96	24.99	43.29
เรขาคณิต	100	27.27	22.47	42.78	22.94
พีชคณิต	100	30.04	22.19	26.80	21.34
การวิเคราะห์ข้อมูลและความน่าจะเป็น	100	21.54	21.27	33.11	25.33
บูรณาการ	100	-	-	23.16	24.33

จากตารางที่ 3 แสดงให้เห็นว่าค่าเฉลี่ยของผลสอบของนักเรียนในปีการศึกษา 2560 และปีการศึกษา 2561 มีคะแนนต่ำกว่า 50 คะแนนทุกสาระการเรียนรู้ สำหรับปีการศึกษา 2561 คะแนนเฉลี่ยของผลสอบในสาระการเรียนรู้ที่สูงกว่าปีการศึกษา 2560 ได้แก่ การวัด เรขาคณิต การวิเคราะห์ข้อมูลและความน่าจะเป็น สำหรับสาระการเรียนรู้ที่ต่ำกว่าปีการศึกษา 2560 ได้แก่ จำนวนและการดำเนินการ และพีชคณิต อย่างไรก็ตาม ในปีการศึกษา 2561 การออกข้อสอบได้มีการวัดความสามารถของนักเรียนในการบูรณาการสาระต่าง ๆ ซึ่งไม่มีในการสอบของปีการศึกษา 2560 และผลสอบอยู่ในระดับต่ำกว่าสาระอื่น ๆ

ประเด็นที่ 3 ผลการสอบจำแนกตามรายข้อ

ผลการวิเคราะห์จำนวนนักเรียนที่ตอบข้อสอบถูกในแต่ละข้อทั่วประเทศ พบว่า ร้อยละของนักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่ตอบถูกในแต่ละข้อ โดยเรียงลำดับจากน้อยไปหามาก เพื่อนำไปศึกษาโมโนทัศน์และข้อผิดพลาดที่จะต้องสร้างบทเรียนออนไลน์แบบเปิดเพื่อมหาชน ดังตารางที่ 4 และ 5

ตารางที่ 4 ร้อยละของนักเรียนที่ตอบถูกในแต่ละข้อทั่วประเทศเรียงจากน้อยไปมาก ปีการศึกษา 2560

ร้อยละ	ข้อที่	มโนทัศน์ที่เกี่ยวข้อง
1.58	21	พื้นที่ผิวของปริซึมฐานสี่เหลี่ยมขนมเปียกปูน
6.99	25	ความน่าจะเป็น (สุ่มหยิบบัตรที่มีจำนวนเป็นจำนวนคี่)
7.88	24	เรขาคณิต (ความคล้ายและพื้นที่)
9.36	19	เรียงลำดับเศษส่วน มีตัวเศษเป็นตัวแปร
11.31	18	ความน่าจะเป็นของการโยนเหรียญ
15.78	20	โจทย์ปัญหาเศษส่วนงานวันเด็ก
15.88	10	เรขาคณิต (วงกลมแนบในสี่เหลี่ยมย้อย)
22.90	6	อสมการ
24.15	1	การลบเศษส่วนที่ตัวส่วนเป็นเลขยกกำลัง
25.19	9	พิกัดของจุดจากการแปลงทางเรขาคณิต
25.35	17	สถิติ (ค่าเฉลี่ยและมัธยฐาน)
25.84	3	อัตราส่วน(ต่อเนื่อง)

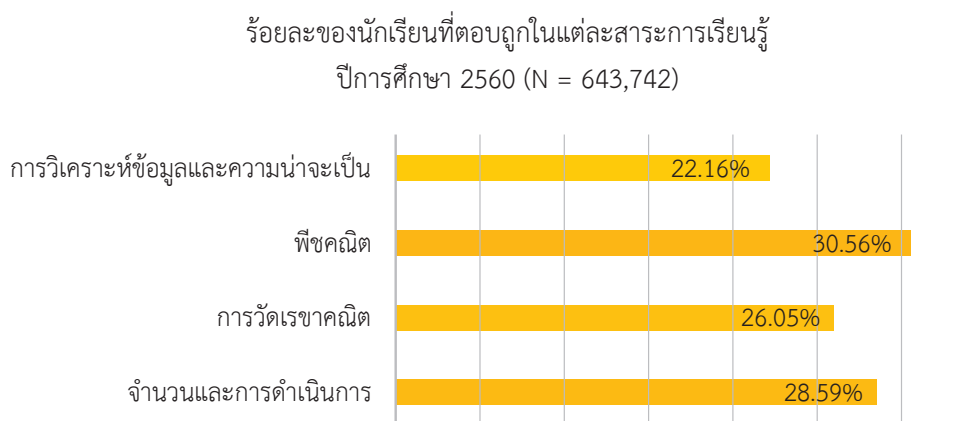
ตารางที่ 5 ร้อยละของนักเรียนที่ตอบถูกในแต่ละข้อทั่วประเทศเรียงจากน้อยไปมาก
ปีการศึกษา 2561

ร้อยละ	ข้อที่	มโนทัศน์ที่เกี่ยวข้อง
28.94	7	แบบรูป
29.77	8	หาค่าของตัวแปรในสมการเชิงเส้นที่กำหนดพิกัดของจุดผ่าน
29.87	15	สามมิติและสองมิติ (กำหนดด้านหน้า ด้านข้าง หาจำนวนลูกบาศก์)
29.93	16	สถิติ (ค่าเฉลี่ย)
32.66	11	เรขาคณิต (รูปที่ต้องใช้สำนึกปริภูมิช่วย)
35.49	14	สามมิติ (หาการมองด้านต่าง ๆ)
36.25	22	อัตราส่วน (ซื้อไข่ไก่และไข่เป็ด)
36.78	13	ปริมาตร (ทรงกระบอก)
37.22	23	สถิติ (แผนภูมิวงกลม)
40.61	5	สมการจากจำนวนเหรียญ
42.22	4	ค.ร.น.เกี่ยวกับเวลา
46.56	2	รากที่สองและรากที่สาม
49.08	12	สามเหลี่ยมคล้าย (ถามหาด้านและมุม)
5.41	24	โจทย์ปัญหาสมการสองตัวแปร
8.15	25	โจทย์ปัญหาสามเหลี่ยมคล้าย
9.31	23	สถิติ ฐานนิยม และ มัธยฐาน (หาผลบวกของข้อมูลที่เพิ่มเข้ามา)
10.67	21	โจทย์ปัญหาเศษส่วนและทศนิยม
12.13	22	โจทย์ปัญหาปริมาตรรูปทรงสี่เหลี่ยมมุมฉาก (ถังน้ำ)
16.44	2	รากที่สอง รากที่สาม
21.37	4	ห.ร.ม. และ ค.ร.น.
21.64	9	พิกัดของจุดที่เกิดจากการแปลงทางเรขาคณิต
23.71	1	เลขยกกำลัง สัญกรณ์วิทยาศาสตร์ การบวกจำนวน
24.86	7	พื้นที่และความยาวรอบวงกลมกับสี่เหลี่ยมจัตุรัส
25.89	5	โจทย์ปัญหาอัตราส่วนของ 3 สิ่ง
27.46	16	เลือกสมการเส้นตรงที่ขนานกับเส้นตรงที่กำหนด
28.74	20	ความน่าจะเป็น
31.31	18	แผนภูมิวงกลมกับร้อยละ
31.90	3	เศษส่วน เลขยกกำลัง การลบจำนวน
32.74	15	แบบรูปจำนวน
33.01	6	พื้นที่ผิวบางส่วนของปริซึม 5 เหลี่ยม
34.38	8	การแปลงทางเรขาคณิต
39.77	19	ความน่าจะเป็น
40.01	14	อสมการและความหมายของค่าเฉลี่ย
41.66	10	ความเท่ากันทุกประการและความคล้าย
46.83	13	โจทย์ปัญหาสมการหนึ่งตัวแปร
53.04	12	กำหนดสามมิติจากลูกบาศก์ที่วางซ้อนกันให้เลือกสองมิติ
60.19	17	แผนภูมิแท่ง
78.67	11	กำหนดภาพสองมิติให้เลือกสามมิติจากลูกบาศก์ที่วางซ้อนกัน

หมายเหตุ: N = 645,575

จากตารางที่ 4 และ ตารางที่ 5 แสดงให้เห็นว่าร้อยละของจำนวนนักเรียนที่ตอบถูกในปีการศึกษา 2560 น้อยกว่าร้อยละ 50 ของจำนวนผู้เข้าสอบทั่วประเทศ แต่ในปีการศึกษา 2561 ร้อยละของจำนวนนักเรียนที่ตอบได้ถูกต้องมากกว่าร้อยละ 50 ของจำนวนนักเรียนที่เข้าสอบทั่วประเทศ ได้แก่ ข้อสอบข้อที่ 12 เรื่อง กำหนดสามมิติจากลูกบาศก์ที่วางซ้อนกันให้เลือกสองมิติ (53.04%) ข้อที่ 17 แผนภูมิแท่ง (60.19%) และข้อ 11 กำหนดภาพสองมิติให้เลือกสามมิติจากลูกบาศก์ที่วางซ้อนกัน (78.67%)

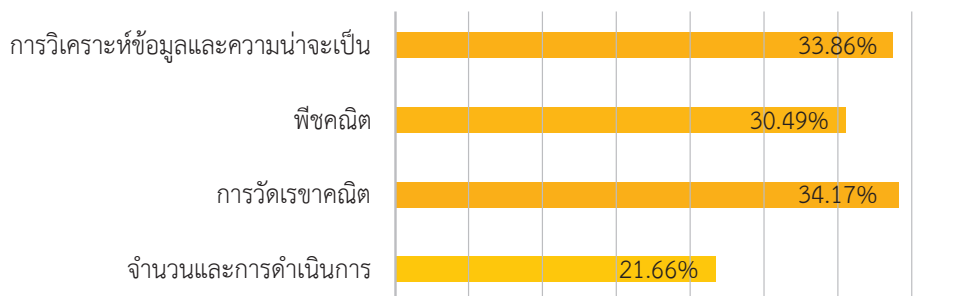
นอกจากนั้นผลการวิเคราะห์จำนวนนักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่ตอบข้อสอบ O-NET ถูกคิดเป็นร้อยละของนักเรียนที่เข้าสอบรายวิชาคณิตศาสตร์ (รหัส 94) ทั่วประเทศ จำแนกตามสาระการเรียนรู้โดยรวมสาระที่ 2 การวัด และ สาระที่ 3 เรขาคณิตเข้าด้วยกัน แสดงผลการวิเคราะห์ในปีการศึกษา 2560 (N = 643,742) (ภาพที่ 4.1) และผลการวิเคราะห์ในปีการศึกษา 2561 (N = 645,575) (ภาพที่ 2)



ภาพที่ 2 ร้อยละของนักเรียนที่ตอบถูกในกลุ่มสาระการเรียนรู้คณิตศาสตร์ ปีการศึกษา 2560

จากภาพที่ 2 แสดงให้เห็นว่าในการสอบ O-NET กลุ่มสาระการเรียนรู้คณิตศาสตร์ ปีการศึกษา 2560 นักเรียนชั้นมัธยมศึกษาปีที่ 3 ทั่วประเทศมีมีโนทัศน์คลาดเคลื่อนหรือข้อผิดพลาดในการทำข้อสอบมากกว่า 50% กระจายตามสาระการเรียนรู้จากมากไปหาน้อย ดังนี้ สาระการเรียนรู้การวิเคราะห์ข้อมูลและความน่าจะเป็น ประมาณ 77.84% สาระการเรียนรู้การวัดและเรขาคณิต ประมาณ 73.95% สาระการเรียนรู้จำนวนและการดำเนินการ ประมาณ 71.41% และสาระการเรียนรู้พีชคณิต ประมาณ 69.44%

ร้อยละของนักเรียนที่ตอบถูกในแต่ละสาระการเรียนรู้
ปีการศึกษา 2561 (N = 645,575)



ภาพที่ 3 ร้อยละของนักเรียนที่ตอบถูกในกลุ่มสาระการเรียนรู้คณิตศาสตร์ ปีการศึกษา 2561

จากภาพที่ 3 แสดงให้เห็นว่าในการสอบ O-NET กลุ่มสาระการเรียนรู้คณิตศาสตร์ ปีการศึกษา 2561 นักเรียนชั้นมัธยมศึกษาปีที่ 3 ทั่วประเทศมีมีโน้ตส์คลาดเคลื่อนหรือข้อผิดพลาดในการทำข้อสอบมากกว่า 50% เช่นเดียวกับการสอบในปีการศึกษา 2560 แต่ลำดับของกระจายตามสาระการเรียนรู้จากมากไปหาน้อยแตกต่างกัน ดังนี้สาระการเรียนรู้จำนวนและการดำเนินการ ประมาณ 78.34% สาระการเรียนรู้พีชคณิต ประมาณ 68.51% สาระการเรียนรู้การวิเคราะห์ข้อมูลและความน่าจะเป็น ประมาณ 66.14% และสาระการเรียนรู้การวัดและเรขาคณิต ประมาณ 65.83%

ผู้วิจัยได้พัฒนาบทเรียนจำแนกตามสาระ ได้แก่ จำนวนและการดำเนินการ พีชคณิต การวัดและเรขาคณิต และการวิเคราะห์ข้อมูลและความน่าจะเป็น ซึ่งแต่ละบทเรียนประกอบด้วยแบบทดสอบก่อนเรียน บทเรียนบทบทวนแบบทดสอบหลังเรียนและการเฉลยข้อสอบของปีการศึกษา 2560 และปีการศึกษา 2561 โดยก่อนการทดลองใช้ได้มีการตรวจสอบความเที่ยงตรงของเนื้อหา (Content Validity Index) จากผู้เชี่ยวชาญ 10 คน ได้ค่า CVI = 1 ผ่านเกณฑ์การยอมรับ 0.80 (Davis, 1992)

ตารางที่ 6 จำนวนผู้เชี่ยวชาญที่ให้ค่าระดับความคิดเห็นที่เกี่ยวกับเนื้อหาของบทเรียนออนไลน์แบบเปิดเพื่อมหาชนที่ประกอบด้วยแบบทดสอบก่อนเรียนบทเรียนและแบบทดสอบหลังเรียน

รายการ	ระดับความคิดเห็น*				รวมจำนวน ผู้เชี่ยวชาญประเมิน 3,4
	4	3	2	1	
แบบทดสอบก่อนเรียน					
1.มีความสอดคล้องกับวัตถุประสงค์ของการเรียนแต่ละบท	10				10
2.มีความถูกต้องตามสาระการเรียนรู้ที่กำหนด	10				10
บทเรียนออนไลน์แบบเปิดเพื่อมหาชน					
3.ความเหมาะสมของการเรียงลำดับเนื้อหา	9	1			10
4.ความถูกต้องของมโนทัศน์ตามเนื้อหาสาระที่กำหนด	9	1			10
5.ส่งเสริมให้นักเรียนเกิดการเรียนรู้ด้วยตนเองได้	8	2			10
6.ระยะเวลาที่ใช้ในการเรียนแต่ละบทเรียนมีความเหมาะสม	8	2			10
7.เนื้อหาที่น่าสนใจมีความสัมพันธ์กับตัวชี้วัดของหลักสูตร	9	1			10
8.ใช้เป็นสื่อการเรียนรู้นอกเหนือจากการเรียนในห้องเรียนได้	9	1			10
แบบทดสอบหลังเรียน					
9.มีความสอดคล้องกับวัตถุประสงค์ของการเรียนแต่ละบท	9	1			10
10.มีความถูกต้องตามสาระการเรียนรู้ที่กำหนด	10				10

หมายเหตุ*

ระดับ 4 หมายถึง มีความถูกต้อง/สอดคล้อง/เหมาะสมมากที่สุด

ระดับ 3 หมายถึง มีความถูกต้อง/สอดคล้อง/เหมาะสมมาก

ระดับ 2 หมายถึง มีความถูกต้อง/สอดคล้อง/เหมาะสมบางส่วน

ระดับ 1 หมายถึง ไม่มีความถูกต้อง/สอดคล้อง/เหมาะสม

จากข้อมูลระดับความคิดเห็นในตารางที่ 4.5 ผู้วิจัยนำไปคำนวณหาค่าความเที่ยงตรงเชิงเนื้อหาโดยใช้การหาค่าดัชนีความเที่ยงตรงของเนื้อหา (Content Validity Index : CVI) กำหนดค่าระดับความคิดเห็น 4 ระดับ

$$CVI = \frac{\text{จำนวนข้อที่ผู้เชี่ยวชาญทั้งหมดให้ระดับ 3 หรือ 4}}{\text{จำนวนข้อคิดเห็นทั้งหมด}} = \frac{10}{10} = 1$$

เกณฑ์การยอมรับ 0.80 (Davis, 1992)

การคำนวณค่าดัชนีความเที่ยงตรงของเนื้อหาได้ผลดังนี้ CVI = 1 ซึ่งสูงกว่าเกณฑ์การยอมรับ แสดงว่าบทเรียนออนไลน์แบบเปิดเพื่อมหาชนที่พัฒนาขึ้นมีคุณภาพเที่ยงตรงตามเนื้อหาสำหรับการหาคุณภาพของแบบทดสอบ E-Testing โดยใช้ ความเที่ยงตรงเชิงปรากฏ (Face Validity) ผู้วิจัยได้ใช้ข้อคำถามประกอบด้วย 5 ประเด็น ดังนี้

- (1) เวลาที่ใช้ในการทำแบบทดสอบ
- (2) ความเหมาะสมของภาษาที่ใช้
- (3) ความเหมาะสมของมโนทัศน์ที่ต้องการทดสอบ
- (4) ความเหมาะสมของกราฟหรือภาพที่ใช้ในการทดสอบมโนทัศน์
- (5) ความเชื่อมโยงของข้อมูลที่ใช้ในการแก้ปัญหา

ผลการใช้บทเรียนออนไลน์แบบเปิดเพื่อมหาชนที่พัฒนาขึ้น

ประเด็นที่ 1 การหาผลการแก้ไขข้อบกพร่องและข้อผิดพลาด

การรายงานผลในประเด็นนี้ได้คำนวณค่าร้อยละจากอัตราส่วนระหว่างผลต่างของจำนวนนักเรียนที่ทำแบบทดสอบก่อนเรียนและได้คะแนนรวมน้อยกว่า 50% (N_1) และจำนวนนักเรียนที่ทำแบบทดสอบหลังเรียนและได้คะแนนรวมน้อยกว่า 50% (N_2) เทียบกับ N_1 ซึ่งเป็นเกณฑ์การยอมรับการแก้ไขข้อบกพร่องและข้อผิดพลาดทางคณิตศาสตร์ได้ผลดังตารางที่ 7

ตารางที่ 7 ร้อยละของจำนวนนักเรียนที่แก้ไขข้อบกพร่องและข้อผิดพลาดทางคณิตศาสตร์

สาระและบทเรียน	จำนวน นักเรียน ที่เข้าเรียน (คน)	จำนวน นักเรียน ตอบถูก (%)	จำนวนนักเรียนที่มี มโนทัศน์คลาดเคลื่อน และข้อผิดพลาด (%)		จำนวน นักเรียนที่ แก้ไขได้ (%)
			ก่อน (N ₁)	หลัง (N ₁)	
1. จำนวนและการดำเนินการ					
1.1 การบวกลบจำนวนเต็ม	198	111 (56.06%)	87 (43.94%)	12 (6.06%)	75 (86.21%)
1.2 การคูณหารจำนวนเต็ม	189	142 (75.13%)	47 (24.87%)	7 (3.70%)	40 (85.11%)
1.3 ท.ร.ม. และ ค.ร.น.	185	89 (48.11%)	96 (51.89%)	20 (10.81%)	76 (79.17%)
1.4 โจทย์ปัญหา ท.ร.ม. และ ค.ร.น.	170	48 (28.24%)	122 (71.76%)	11 (6.47%)	111 (90.98%)
1.5 เศษส่วน	156	84 (53.85%)	72 (46.15%)	16 (10.26%)	56 (77.78%)
1.6 อัตราส่วนและร้อยละ	155	55 (35.48%)	100 (64.52%)	11 (7.10%)	89 (89.00%)
1.7 ทศนิยม	143	81 (56.64%)	62 (43.36%)	10 (6.99%)	52 (83.87%)
1.8 เลขยกกำลัง	143	41 (28.67%)	102 (71.33%)	19 (13.29%)	83 (81.37%)
1.9 แบบรูป	140	64 (45.71)	76 (54.29%)	8 (5.71%)	68 (89.47%)
1.10 รากที่สองและรากที่สาม	141	30 (21.28%)	111 (78.72%)	21 (14.89%)	90 (81.08%)
2. พีชคณิต					
2.1 สมการเชิงเส้นตัวแปรเดียว	118	52 (44.07%)	66 (55.93%)	17 (14.41%)	49 (74.24%)
2.2 โจทย์ปัญหาสมการเชิงเส้นตัวแปรเดียว	117	43 (36.75)	74 (63.25%)	6 (5.13%)	68 (91.89%)
2.3 สมการเชิงเส้นสองตัวแปร	106	28 (26.42%)	78 (73.58%)	6 (5.66%)	72 (92.31%)
2.4 โจทย์ปัญหาสมการเชิงเส้นสองตัวแปร	97	47 (48.45%)	50 (51.55%)	2 (2.06%)	48 (96.00%)
2.5 อสมการ	96	24 (25.00%)	72 (75.00%)	5 (5.21%)	67 (93.06%)
2.6 โจทย์ปัญหาอสมการ	94	31 (32.98%)	63 (67.02%)	7 (7.45%)	56 (88.89%)

สาระและบทเรียน	จำนวน นักเรียน ที่เข้าเรียน (คน)	จำนวน นักเรียน ตอบถูก (%)	จำนวนนักเรียนที่มี มโนทัศน์คลาดเคลื่อน และข้อผิดพลาด (%)		จำนวน นักเรียนที่ แก้ไขได้ (%)
			ก่อน (N ₁)	หลัง (N ₁)	
3. การวัดและเรขาคณิต					
3.1 รูปเรขาคณิตและการวัด	72	9 (12.50%)	63 (87.50%)	7 (9.72%)	56 (88.89%)
3.2 รูปเรขาคณิตสามมิติ	73	42 (57.53%)	31 (42.47%)	4 (5.48%)	27 (87.10%)
3.3 การแปลงทางเรขาคณิต	68	14 (20.59%)	54 (79.41%)	7 (10.29%)	47 (87.04%)
3.4 เส้นขนานและความคล้าย	79	11 (13.92%)	68 (86.08%)	27 (34.18%)	41 (60.29%)
4. การวิเคราะห์ข้อมูลและความน่าจะเป็น					
4.1 สถิติ	70	15 (21.43%)	55 (78.57%)	20 (28.57%)	35 (63.64%)
4.2 ความน่าจะเป็น	70	6 (8.57%)	64 (91.43%)	8 (11.43%)	56 (87.50%)

จากตารางที่ 7 พบว่าผลการทดสอบก่อนเรียนแต่ละบทเรียนมีโมโนทัศน์ที่คลาดเคลื่อนและข้อผิดพลาดทุกบทเรียน แต่เมื่อนักเรียนศึกษาบทเรียนและการเฉลยข้อสอบโอเน็ตที่ผู้วิจัยและทีมวิจัยสร้างขึ้น ทำให้จำนวนนักเรียนที่มีโมโนทัศน์คลาดเคลื่อนและข้อผิดพลาดลดลง และเมื่อพิจารณาจากร้อยละของนักเรียนที่มีคะแนนรวมสูงในการสอบหลังเรียนที่มีคะแนนรวมสูงกว่า 50% มีจำนวนร้อยละของนักเรียนที่แก้ไขโมโนทัศน์ที่คลาดเคลื่อนและข้อผิดพลาดทางคณิตศาสตร์ตั้งแต่ร้อยละ 63.64 (บทเรียนเรื่องสถิติ) ถึงร้อยละ 96.00 (บทเรียนเรื่องโจทย์ปัญหาสมการเชิงเส้นสองตัวแปร)

ผู้วิจัยได้หาประสิทธิภาพของบทเรียนโดยวิเคราะห์จากประสิทธิภาพในการเรียนรู้ของนักเรียน โดยคำนวณค่าดัชนีประสิทธิผล ดังตารางที่ 8

ตารางที่ 8 ดัชนีประสิทธิผลที่แสดงประสิทธิภาพของบทเรียน

รายการ	ผลรวมของคะแนนทั้งหมด	ดัชนีประสิทธิผล
คะแนนก่อนเรียน	1,915.72	$E.I. = \frac{2588.82 - 1915.72}{2680 - 1915.72}$ $= 0.8807$
คะแนนหลังเรียน	2,588.82	
คะแนนรวม	2,680	

จากตารางที่ 7 ดัชนีประสิทธิผลที่แสดงประสิทธิภาพของบทเรียนมีค่าเท่ากับ 0.8807 ซึ่งแสดงว่าบทเรียนออนไลน์แบบเปิดเพื่อมหาชนสามารถพัฒนานักเรียนให้มีผลสัมฤทธิ์ในการเรียนบทเรียนนี้ ร้อยละ 88.07 ซึ่งสูงกว่าเกณฑ์การยอมรับที่ต้องสูงกว่าหรือเท่ากับ 0.50 (เผด็จ กิจระการ, 2545)

นอกจากนั้นผลการวิเคราะห์ความแตกต่างของผลการทำแบบทดสอบก่อนเรียนและหลังเรียนของนักเรียนพบว่านักเรียนมีผลการเรียนรู้สูงกว่าก่อนการศึกษบทเรียนในทุกเรื่องอย่างมีนัยสำคัญทางสถิติ .05 ดังตารางที่ 9

ตารางที่ 9 การเปรียบเทียบผลสัมฤทธิ์ในการแก้ไขข้อบกพร่องที่คลาดเคลื่อนและข้อผิดพลาดทางคณิตศาสตร์

สาระการเรียนรู้	n	คะแนนก่อนเรียน		คะแนนหลังเรียน		t	Sig.
		Mean	S.D.	Mean	S.D.		
1. จำนวนและการดำเนินการ							
1.1 การบวก การลบจำนวนเต็ม	198	0.83	0.24	0.98	0.28	8.739*	.000
1.2 การคูณ การหารจำนวนเต็ม	189	0.90	0.19	0.99	0.07	6.306*	.000
1.3 ท.ร.ม. และ ค.ร.น.	185	0.67	0.37	0.95	0.16	9.743*	.000
1.4 โจทย์ปัญหา ท.ร.ม. และ ค.ร.น.	170	0.71	0.25	0.97	0.11	5.927*	.000
1.5 เศษส่วน	156	0.79	0.28	0.98	0.07	8.342*	.000
1.6 อัตราส่วน ร้อยละ	155	0.72	0.26	0.98	0.08	12.349*	.000
1.7 ทศนิยม	143	0.85	0.22	0.98	0.08	6.792*	.000
1.8 เลขยกกำลัง	143	0.72	0.24	0.96	0.13	10.450*	.000
1.9 แบบรูป	140	0.84	0.17	0.98	0.18	9.250*	.000
1.10 รากที่สองและรากที่สาม	141	0.68	0.25	0.94	0.17	11.661*	.000
2. พีชคณิต							
2.1 สมการเชิงเส้นตัวแปรเดียว	118	0.76	0.28	0.96	0.12	8.576*	.000
2.2 โจทย์ปัญหาสมการเชิงเส้นตัวแปรเดียว	117	0.68	0.29	0.98	0.09	11.128*	.000
2.3 สมการเชิงเส้นสองตัวแปร	106	0.71	0.25	0.98	0.11	10.224*	.000
2.4 โจทย์ปัญหาสมการเชิงเส้นสองตัวแปร	97	0.77	0.27	0.99	0.05	8.249*	.000
2.5 อสมการ	96	0.50	0.35	0.98	0.09	13.233*	.000
2.6 โจทย์ปัญหาอสมการ	94	0.71	0.26	0.97	0.11	9.867*	.000
3. การวัดและเรขาคณิต							
3.1 รูปเรขาคณิตและการวัด	72	0.33	0.35	0.96	0.14	14.816*	.000
3.2 รูปเรขาคณิตสามมิติ	73	0.81	0.27	0.97	0.13	5.044*	.000
3.3 การแปลงทางเรขาคณิต	68	0.50	0.35	0.95	0.17	9.247*	.000
3.4 เส้นขนานและความคล้าย	79	0.42	0.34	0.86	0.22	10.521*	.000
4. การวิเคราะห์ข้อมูลและความน่าจะเป็น							
4.1 สถิติ	70	0.64	0.28	0.90	0.17	7.563*	.000
4.2 ความน่าจะเป็น	70	0.50	0.25	0.95	0.16	11.621*	.000

หมายเหตุ* นัยสำคัญทางสถิติที่ระดับ .05

จากตารางที่ 9 แสดงการเปรียบเทียบความแตกต่างของค่าเฉลี่ยของคะแนนก่อนเรียนและหลังเรียนของแต่ละเรื่องที่สอบด้วยระบบ E-Testing ซึ่งกำหนดค่าของคะแนนในบทเรียนเทียบกับ 1.00 พบว่า ผลการสอบของนักเรียนหลังเรียนด้วยบทเรียนออนไลน์แบบเปิดเพื่อมหาชนทุกบทเรียนสูงกว่าก่อนเรียนอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และ p มีค่าเท่ากับ .000 ทุกบทเรียน แต่ค่า df แตกต่างกันตามจำนวนผู้เข้าเรียนที่ไม่เท่ากัน บทเรียนที่เข้าเรียนมากที่สุด คือ การบวก การลบจำนวนเต็ม $t(197) = 8.739$ บทเรียนที่เข้าเรียนน้อยที่สุด คือ การแปลงทางเรขาคณิต $t(67) = 9.247$

ประเด็นที่ 2 ผลการสำรวจความพึงพอใจของนักเรียนต่อการเรียนรู้จากบทเรียนออนไลน์แบบเปิดเพื่อมหาชน

ผู้วิจัยได้ประเมินความพึงพอใจของนักเรียนต่อการเรียนรู้จากบทเรียนออนไลน์แบบเปิดเพื่อมหาชน โดยใช้กรอบแนวคิดของการประเมินของ Thai MOOC ที่เกี่ยวข้องกับงานวิจัยครั้งนี้ ใน 2 ประเด็น ได้แก่

- (1) แบบวัดความพึงพอใจที่มีต่อรูปแบบการจัดการเรียนการสอน
- (2) แบบวัดความพึงพอใจโดยรวม ผลการประเมินความพึงพอใจแสดงไว้ในตารางที่ 10

ตารางที่ 10 ความพึงพอใจของนักเรียนต่อการเรียนรู้จากบทเรียนออนไลน์แบบเปิดเพื่อมหาชน (n = 180)

รายการ	พึงพอใจมากที่สุด	พึงพอใจปานกลาง	พึงพอใจน้อย	พึงพอใจน้อยที่สุด
ตอนที่ 1 ความพึงพอใจที่มีต่อรูปแบบการจัดการเรียนการสอน				
• เนื้อหามีความทันสมัย ทันต่อการเปลี่ยนแปลง	85 (47.22%)	76 (42.22%)	19 (10.56%)	-
• ลำดับเรื่องเข้าใจง่าย ชวนติดตาม	71 (39.44%)	81 (45.00%)	27 (15.00%)	1 (0.56%)
• รูปแบบการนำเสนอสวยงาม เข้าใจง่าย	73 (40.56%)	68 (37.78%)	36 (20.00%)	3 (1.66%)
• ตัวอย่าง/กรณีศึกษาเพียงพอ ชัดเจน	76 (42.22%)	71 (39.44%)	31 (17.22%)	2 (1.12%)
• รูปแบบกิจกรรม/แบบทดสอบ เหมาะสม ชัดเจน	74 (41.11%)	80 (44.44%)	26 (14.45%)	-
• ภาพคมชัด/กราฟิกมีขนาด สีสันเหมาะสม	88 (48.89%)	68 (37.78%)	24 (13.33%)	-
• เสียงที่ใช้ในสื่อเหมาะสม ชัดเจน	85 (47.22%)	65 (36.11%)	27 (15.00%)	3 (1.67%)
ตอนที่ 2 ความพึงพอใจโดยรวม				
• ความพึงพอใจโดยรวม	71 (39.44%)	82 (45.56%)	27 (15.00%)	-

จากตารางที่ 10 แสดงให้เห็นว่าผู้เข้าศึกษาบทเรียนออนไลน์แบบเปิดเพื่อมหาชน รายวิชาคณิตศาสตร์ระดับมัธยมศึกษาปีที่ 3 มีระดับความพึงพอใจมากที่สุด 3 รายการแรก ได้แก่ ภาพคมชัด/กราฟิกมีขนาด สีสันเหมาะสม 48.89% เนื้อหาที่มีความทันสมัย ทันต่อการเปลี่ยนแปลง 47.22% และ เสียงที่ใช้ในสื่อเหมาะสม ชัดเจน 47.22% ระดับความพึงพอใจ 3 รายการแรก ได้แก่ ลำดับเรื่องเข้าใจง่าย ชวนติดตาม 45.00% รูปแบบกิจกรรม/แบบทดสอบ เหมาะสม ชัดเจน 44.44% สำหรับความพึงพอใจโดยรวม พบว่า ความพึงพอใจโดยรวมต่อบทเรียนออนไลน์แบบเปิด ในระดับความพึงพอใจ 45.56% ระดับความพึงพอใจมากที่สุด 39.44% และระดับความพึงพอใจปานกลาง 15.00%

อภิปรายผล

สรุปผลโดยรวม นักเรียนระดับชั้นมัธยมศึกษาปีที่ 3 มีมโนทัศน์ที่คลาดเคลื่อนและข้อผิดพลาดทางคณิตศาสตร์ทุกสาระ โดยเฉพาะในปีการศึกษา 2561 สาระการเรียนรู้จำนวนและการดำเนินการมีสูงสุด 78.34% ซึ่งสาระนี้มีการเชื่อมโยงกับสาระการเรียนรู้อื่น ๆ ที่ต้องใช้แนวคิดเกี่ยวกับจำนวนและการดำเนินการไปแก้ปัญหา ดังนั้นการสร้างและพัฒนาบทเรียนในสาระที่เกี่ยวกับจำนวนและการดำเนินการมีทั้งหมด 10 บทเรียนจากจำนวนบทเรียนทั้งหมด 22 บทแบบฝึกหัดท้ายบท 4 ชุด ตัวอย่างข้อสอบ O-NET ปีการศึกษา 2560 และปีการศึกษา 2561 ท้ายบทเรียน 4 ชุด เพื่อให้ นักเรียนมีความรู้พื้นฐานที่จะนำไปเชื่อมโยงกับการแก้ปัญหาในสาระอื่น ๆ ซึ่งสอดคล้องกับ Grouws & Cebulla (2000, หน้า.4) ได้เสนอแนวทางหนึ่งที่จะช่วยเพิ่มผลสัมฤทธิ์ในการเรียนของนักเรียน (Improving Student Achievement in Mathematics) คือ เรื่อง การสอนคณิตศาสตร์ที่เน้นสำนึกเชิงจำนวน (number sense) ที่นักเรียนต้องนำไปใช้ในการแก้ปัญหาในสถานการณ์ต่าง ๆ อย่างกว้างขวาง และคณิตศาสตร์มีความสำคัญต่อการพัฒนาความคิด สำหรับมโนทัศน์ที่คลาดเคลื่อนและข้อผิดพลาดทางคณิตศาสตร์ในสาระการเรียนรู้พีชคณิต ประมาณ 68.51% สาระการเรียนรู้การวิเคราะห์ข้อมูลและความน่าจะเป็น ประมาณ 66.14% และสาระการเรียนรู้การวัดและเรขาคณิต ประมาณ 65.83% นั้นผู้วิจัยได้ให้ความสำคัญและความเชื่อมโยงสาระในการพัฒนาบทเรียนออนไลน์แบบเปิดเพื่อมหาชน สำหรับแก้ไขมโนทัศน์และความคลาดเคลื่อนทางคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 และหลังจากนักเรียนได้ใช้บทเรียนออนไลน์แล้วพบว่า มีจำนวนร้อยละของนักเรียนที่แก้ไขมโนทัศน์ที่คลาดเคลื่อนและข้อผิดพลาดทางคณิตศาสตร์ตั้งแต่ร้อยละ 63.64 (บทเรียนเรื่องสถิติ) ถึงร้อยละ 96.00 (บทเรียนเรื่องโจทย์ปัญหาสมการเชิงเส้นสองตัวแปร) และการใช้บทเรียนออนไลน์แบบเปิดเพื่อมหาชนมีผลต่อการพัฒนานักเรียนที่มีค่าดัชนีประสิทธิผลเท่ากับ 0.8807 แสดงว่าสามารถพัฒนานักเรียนให้มีผลสัมฤทธิ์ในการเรียนบทเรียนนี้ ร้อยละ 88.07 ซึ่งสูงกว่าเกณฑ์การยอมรับที่ต้องสูงกว่าหรือเท่ากับ 0.50

สำหรับผลการวิจัยด้านความพึงพอใจ พบว่า มีผู้ตอบแบบสอบถาม 180 คน มีความพึงพอใจมากที่สุดในเรื่อง ภาพคมชัด/กราฟฟิกมีขนาด สีสันเหมาะสม 48.89% เนื้อหามีความทันสมัย ทันต่อการเปลี่ยนแปลง 47.22% และ เสียงที่ใช้ในสื่อเหมาะสม ชัดเจน 47.22% มีความพึงพอใจมากในเรื่อง ลำดับเรื่องเข้าใจง่าย ชวนติดตาม 45.00% รูปแบบกิจกรรม/แบบทดสอบ เหมาะสม ชัดเจน 44.44% นอกจากนั้นผู้เข้าเรียนมีความประสงค์ที่จะให้พัฒนาบทเรียนแบบเปิดอย่างต่อเนื่องและกำหนดช่วงเวลาของการเข้าถึงบทเรียนออนไลน์ให้มากขึ้น

ข้อเสนอแนะ

ข้อเสนอแนะจากผลการวิจัยครั้งนี้

1) ผู้สอนในกลุ่มสาระการเรียนรู้คณิตศาสตร์ควรนำผลการสอบ O-NET ของนักเรียนมาวิเคราะห์หามโนทัศน์ที่คลาดเคลื่อนและความผิดพลาดในการคิดทางคณิตศาสตร์ เพื่อเป็นพื้นฐานของการเลือกใช้วิธีสอนและสื่อการเรียนรู้ที่ช่วยพัฒนานักเรียนให้มีมโนทัศน์และการคิดทางคณิตศาสตร์ที่ถูกต้อง

2) ในการเลือกใช้วิธีสอนและสื่อการสอนควรสอดคล้องกับรูปแบบการเรียนรู้ที่สร้างแรงจูงใจแก่นักเรียน เช่น การใช้เทคโนโลยีดิจิทัลเป็นส่วนหนึ่งของการเรียนรู้ เช่น การสร้างบทเรียนออนไลน์ให้นักเรียนเรียนรู้นอกห้องเรียน ซึ่งผู้สอนควรประเมินคุณภาพของสื่อเพื่อตรวจสอบความเที่ยงตรงเชิงเนื้อหาจากผู้ที่มีประสบการณ์หรือผู้ร่วมงาน

3) การจัดการเรียนรู้ควรให้ผู้เรียนได้ประเมินตนเองในด้านความก้าวหน้าของการเรียนรู้ รวมทั้งการแสดงความคิดเห็นเกี่ยวกับความพึงพอใจที่มีต่อการจัดกิจกรรมการเรียนรู้

ข้อเสนอแนะในการทำวิจัยครั้งต่อไป

ข้อจำกัดในการวิจัยครั้งนี้มีขนาดของกลุ่มตัวอย่างค่อนข้างเล็ก การมีกลุ่มตัวอย่างขนาดใหญ่จะช่วยให้สามารถวิเคราะห์ข้อมูลเชิงปริมาณได้ตรงตามวัตถุประสงค์ในข้อ 2 และ ข้อ 3 มากขึ้น

ในการวิจัยครั้งต่อไป ผู้ที่สนใจในการทำวิจัยการใช้บทเรียนออนไลน์แบบเปิดเพื่อมหาชนเพื่อการเรียนรู้อาจตั้งปัญหาของการวิจัยตามข้อเสนอแนะต่อไปนี้

1) บทเรียนออนไลน์แบบเปิดเพื่อมหาชนช่วยให้นักเรียนมีสภาพแวดล้อมในการเรียนรู้ได้ตามสภาพจริงหรือไม่

2) บทเรียนออนไลน์แบบเปิดเพื่อมหาชนช่วยให้นักเรียนได้เรียนรู้นอกเวลาเรียนที่สอดคล้องกับผลลัพธ์การเรียนรู้ที่กำหนดไว้หรือไม่

3) วิธีการใดที่จะจัดเวลาในการเรียนรู้บทเรียนออนไลน์แบบเปิดเพื่อมหาชนให้สอดคล้องกับเนื้อหาเฉพาะตามผลลัพธ์การเรียนรู้ที่กำหนด

4) อะไรเป็นการท้าทายที่สำคัญสำหรับนักเรียนให้ใช้การเรียนรู้บทเรียนออนไลน์แบบเปิดเพื่อมหาชน



เอกสารอ้างอิง

- กระทรวงศึกษาธิการ. (2560). มาตรฐานการเรียนรู้และตัวชี้วัด กลุ่มสาระการเรียนรู้คณิตศาสตร์ วิทยาศาสตร์ และสาระภูมิศาสตร์ ในกลุ่มสาระการเรียนรู้สังคมศึกษา ศาสนา และวัฒนธรรม (ฉบับปรับปรุง พ.ศ. 2560)ตามหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551. กรุงเทพมหานคร: โรงพิมพ์ชุมนุมสหกรณ์การเกษตรแห่งประเทศไทย จำกัด.
- ปราวินยา สุวรรณณัฐโชติ และ เสมอกาญจน์ ไสภณศิริรักษ์. (2560). มาตรฐานและแนวปฏิบัติ การเรียนการสอน MOOC ที่ได้รับการยอมรับระดับนานาชาติ. กรุงเทพมหานคร: โครงการมหาวิทยาลัยไซเบอร์ไทย สำนักงานคณะกรรมการการอุดมศึกษา
- เผด็จ กิจระการ. (2545) ดัชนีประสิทธิผล (Effectiveness Index: E.I.). วารสารการวัดผล การศึกษา มหาวิทยาลัยมหาสารคาม ปีที่ 8: กรกฎาคม 2545, หน้า 30-36.
- สถาบันทดสอบทางการศึกษาแห่งชาติ. (2560). รายงานประจำปี 2560. สืบค้น 10 สิงหาคม 2562 จาก www.niets.or.th
- Davis L.L. (1992). Instrument review: Getting the most from a panel of experts. *Applied Nursing Research*, 5, pp.194-197.
- Earl, L. (2003). *Assessment as learning: Using classroom assessment to maximize student learning*. Thousand Oaks, CA: Corwin.
- Grouws, D. A., & Cebulla, K. J. (2000). *Improving student achievement in mathematics*. Columbus, OH: ERIC Clearinghouse for Science, Mathematics, and Environmental Education.
- Hassan, J. A., & Mahmoud, E. (2011). *The 5 gaps*. Minneapolis, MN: Harvest Preparatory Academy and Best Academy. Retrieved from <http://www.headwatersfoundation.org/news/201110/499>
- Issel, L.M. (2000). *Health program planning and evaluation: A practical, systematic approach for community health*. Canada: Jones and Bartlett Publishers.
- Knowledge Gap. (2013). In K. Bell (Ed.), *Open education sociology dictionary*. Retrieved from <https://sociologydictionary.org/knowledge-gap/>

การประเมินคุณภาพงานวิจัยการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียน ในศตวรรษที่ 21 ระดับมัธยมศึกษา

An evaluation of researches quality on the 21st century competencies
measurement tools development for secondary school students

สร้อยญา จันทร์ชูสกุล
ทัศนศิริินทร์ สว่างบุญ

บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อประเมินคุณภาพงานวิจัยการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา ประชากรที่ใช้ในการศึกษาเป็นงานวิจัยที่ศึกษาเกี่ยวกับการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา ระหว่างปี พ.ศ. 2546-2562 จำนวน 92 เล่ม เครื่องมือที่ใช้ในการวิจัย คือ แบบประเมินคุณภาพงานวิจัย จำนวน 1 ฉบับ วิเคราะห์ข้อมูลด้วยสถิติพื้นฐาน ได้แก่ ความถี่ ร้อยละ ค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน ผลการวิจัยสรุปได้ดังนี้

1. ผลการประเมินคุณภาพ พบว่า งานวิจัยมีคุณภาพอยู่ในระดับดี จำนวน 76 เล่ม (ร้อยละ 82.61) และระดับดีมาก จำนวน 16 เล่ม (ร้อยละ 17.39)

2. คุณภาพงานวิจัยเมื่อพิจารณาจำแนกรายด้าน พบว่า ด้านผลการวิเคราะห์ข้อมูล มีค่าเฉลี่ยสูงสุด ($M=3.875$, $SD=0.382$) รองลงมาคือ ด้านการสรุป อภิปรายผล และข้อเสนอแนะ ($M=3.214$, $SD=0.473$) ด้านวิธีดำเนินการวิจัย ($M=3.046$, $SD=0.370$) ด้านการนำเสนอรายงานวิจัย ($M=2.783$, $SD=0.340$) ด้านความเป็นมาและความสำคัญของปัญหาวิจัย ($M=2.634$, $SD=0.457$) และด้านเอกสารและงานวิจัยที่เกี่ยวข้อง ($M=2.429$, $SD=0.340$) ตามลำดับ

คำสำคัญ: การประเมินคุณภาพงานวิจัย, เครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21

Abstract

The purpose of this research was to evaluate the researches quality on the 21st century competencies measurement tools development for secondary school students. The research population was 92 research papers produced in 2003-2019. The research instrument was a research quality evaluative forms. The data were analyzed by frequency, percentage, mean and standard deviation. The results of research were as follows:

1. The result of evaluate the researches quality was found that 76 papers (82.61%) were of good quality and 16 papers (17.39%) were of very good quality.

2. The researches quality when categorized in each aspect, the aspect which had the highest scores was data analysis ($M = 3.875$, $SD = 0.382$) followed by the summary, discussions and recommendations ($M = 3.214$, $SD = 0.473$), the research methodology ($M = 3.046$, $SD = 0.370$), the presentation of research ($M = 2.783$, $SD = 0.340$) the background and importance of the research problem ($M = 2.634$, $SD = 0.457$) and review literature ($M = 2.429$, $SD = 0.340$), respectively.

Key word: evaluation of research quality, 21st century competencies measurement tools

บทนำ

การเข้าสู่ศตวรรษที่ 21 และการเปลี่ยนแปลงแบบก้าวกระโดดทางเทคโนโลยี ทำให้เกิดความตื่นตัวของนักการศึกษา นักวิชาการ รวมถึงองค์กรเอกชนในการเสนอแนวคิดที่เกี่ยวกับการรอบรู้ในศตวรรษที่ 21 เพื่อวางเป้าหมายและเตรียมความพร้อมให้ผู้เรียนมีสมรรถนะที่จำเป็นในศตวรรษที่ 21 โดยเริ่มในปี ค.ศ. 2003 กลุ่ม METIRI (METIRI, 2003) ประเทศสหรัฐอเมริกา ได้ใช้เวลาในการพัฒนาและเสนอกรอบทักษะที่จำเป็นในศตวรรษที่ 21 ที่ชื่อว่า “enGauge 21st century skills” ในเวลาต่อมา Partnership for 21st Century Learning หรือ P21 (2019) ได้พัฒนารอบแนวคิดการรอบรู้ในศตวรรษที่ 21 โดยครูผู้เชี่ยวชาญทางการศึกษา และผู้นำทางภาคธุรกิจในประเทศสหรัฐอเมริกา เพื่อที่กำหนดและแสดงให้เห็นถึงความรู้และทักษะที่ผู้เรียนจำเป็นต้องมีเพื่อความสำเร็จในอาชีพและการทำงาน ด้วยเหตุนี้ทำให้ทั่วโลกเกิดการตื่นตัว และได้มีการกำหนดกรอบสมรรถนะเพื่อเป็นเป้าหมายในการพัฒนาผู้เรียนที่เหมาะสมกับบริบทในประเทศของตน

สำหรับประเทศไทยได้มีการบรรจุเรื่องการพัฒนาผู้เรียนให้มีคุณลักษณะและทักษะการเรียนรู้ ในศตวรรษที่ 21 ไว้ใน แผนการศึกษาแห่งชาติ พ.ศ. 2560-2579 โดยได้กำหนดคุณลักษณะและทักษะการเรียนรู้ในศตวรรษที่ 21 ที่เรียกตามคำย่อว่า 3Rs + 8Cs ประกอบด้วย ทักษะและคุณลักษณะต่อไปนี้ 3Rs ได้แก่ อ่านออก (Reading) เขียนได้ (WRiting) และคิดเลขเป็น (ARithmetics) 8Cs ประกอบด้วย

- 1) ทักษะด้านการคิดอย่างมีวิจารณญาณ และทักษะในการแก้ปัญหา (critical thinking and problem solving)
- 2) ทักษะด้านการสร้างสรรค์และนวัตกรรม (creativity and innovation)
- 3) ทักษะด้านความเข้าใจต่างวัฒนธรรม ต่างกระบวนทัศน์ (cross-cultural understanding)
- 4) ทักษะด้านการทำงานร่วมกัน การทำงานเป็นทีม และภาวะผู้นำ (collaboration, teamwork and leadership)
- 5) ทักษะด้านการสื่อสาร สารสนเทศ และการรู้เท่าทันสื่อ (communications, information and media literacy)
- 6) ทักษะด้านคอมพิวเตอร์และ เทคโนโลยีสารสนเทศและการสื่อสาร (computing and ICT literacy)
- 7) ทักษะอาชีพ และทักษะ การเรียนรู้ (career and learning skills)
- 8) ความมีเมตตา กรุณา มีวินัย คุณธรรม จริยธรรม (compassion) (สำนักงานเลขาธิการสภาการศึกษา, 2560)

จากความสำคัญของการพัฒนาผู้เรียนให้มีคุณลักษณะและทักษะการเรียนรู้ในศตวรรษที่ 21 จึงก่อให้เกิดงานวิจัยที่เกี่ยวข้องในหลากหลายมิติ หนึ่งในนั้นคือ การพัฒนา

เครื่องมือในการวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 สำหรับใช้ในการวัดสมรรถนะของผู้เรียน ซึ่งจะทำให้ได้ข้อมูลสารสนเทศอันเป็นประโยชน์ เช่น ทำให้ทราบถึงระดับสมรรถนะของผู้เรียนในศตวรรษที่ 21 ในด้านต่าง ๆ อันจะเป็นประโยชน์ต่อการกำหนดนโยบาย หรือวางแผนสำหรับพัฒนาผู้เรียน รวมถึงการนำมาใช้วัดพัฒนาการ หรือผลที่เกิดขึ้นเมื่อจัดกิจกรรมในการพัฒนาสมรรถนะของผู้เรียนในศตวรรษที่ 21 ไปแล้ว เป็นต้น จากการสำรวจงานวิจัยที่มีการพัฒนาเครื่องมือในการวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ในระดับมัธยมศึกษา ระหว่างปี พ.ศ. 2561-2562 พบว่า โดยส่วนใหญ่เป็นงานวิจัยประเภทวิทยานิพนธ์ และบางส่วนเป็นงานวิจัยจากสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) รวมจำนวน 92 เล่ม แต่อย่างไรก็ดี จากผลการศึกษาของ ศิวะพร ภูพันธ์ (2554) พบว่า ปัจจัยหนึ่งที่มีความสำคัญต่อการนำผลวิจัยไปใช้ประโยชน์ของอาจารย์มหาวิทยาลัย คือ คุณลักษณะของงานวิจัย ซึ่งประกอบด้วย คุณภาพของงานวิจัย ความทันสมัยของงานวิจัย และความสอดคล้องของงานวิจัยกับรายวิชาที่สอน

ธีระ อาชวเมธี (2541 อ้างถึงใน วรณีย์ แกมเกตุ, 2555) ได้เสนอ สามมิติของการวิจัยเชิงประจักษ์ (empirical research) ที่สะท้อนถึงคุณค่าของงานวิจัย ได้แก่

- 1) มิติทางความรู้ (epistemological dimension) โดยพิจารณาจากผลการวิจัย
- 2) มิติทางวิธีการ (methodological dimension) โดยพิจารณาจากวิธีการได้มาซึ่งความรู้
- 3) มิติทางประโยชน์ (practical dimension) พิจารณาว่าการวิจัยนั้นเป็นประโยชน์ต่อชุมชน สังคม หรือมนุษยชาติมากน้อยเพียงใด

ในขณะที่ Babbie (2004 อ้างถึงใน วรณีย์ แกมเกตุ, 2555) ได้เสนอชุดคำถามสำหรับการประเมินคุณภาพงานวิจัย ประกอบด้วยมิติที่เกี่ยวข้องกับ การอธิบายทฤษฎีที่ใช้ในการวิจัย การออกแบบการวิจัย การวัดตัวแปร การได้มาซึ่งกลุ่มตัวอย่าง การวิเคราะห์ข้อมูล และการเขียนรายงาน ดังนั้น ในการประเมินคุณภาพงานวิจัยควรให้ความสำคัญในประเด็นดังกล่าวด้วย

อย่างไรก็ดี สำหรับงานวิจัยเกี่ยวกับการพัฒนาเครื่องมือในการวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา กลุ่มผู้ใช้ประโยชน์จากงานวิจัย ได้แก่ นักวิชาการ อาจารย์มหาวิทยาลัย และกลุ่มสำคัญ คือ ครูในระดับมัธยมศึกษา ด้วยเหตุนี้ หากมีการประเมินคุณภาพงานวิจัยดังกล่าว จะเป็นการสร้างความมั่นใจให้แก่ผู้ใช้ประโยชน์จากงานวิจัย รวมถึง ทำให้ทราบถึงคุณภาพของงานวิจัยในมิติต่าง ๆ ซึ่งจะเป็นประโยชน์ต่อนิสิตนักศึกษาระดับบัณฑิตศึกษาที่อยู่ระหว่างการวางแผนหรือจัดทำวิทยานิพนธ์ รวมถึงอาจารย์ที่ปรึกษาวิทยานิพนธ์ และนักวิจัยสารสนเทศเกี่ยวกับคุณภาพงานวิจัยจะทำให้ นิสิตนักศึกษา อาจารย์ที่ปรึกษา และนักวิจัย ได้ทราบว่าในการทำงานวิจัยที่ผ่านมาประเด็นใดที่ยังเป็นจุดอ่อน หรือเป็นประเด็นที่ควรให้ความสำคัญในการควบคุมและพัฒนางานวิจัยให้มีคุณภาพต่อไปในอนาคต ด้วยเหตุนี้ ผู้วิจัยจึงสนใจที่จะประเมินคุณภาพงานวิจัยเกี่ยวกับการพัฒนาเครื่องมือในการวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา

วัตถุประสงค์ของการวิจัย

เพื่อประเมินคุณภาพงานวิจัยการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา

ขอบเขตการวิจัย

งานวิจัยที่ใช้ในการประเมินคุณภาพงานวิจัย เป็นงานวิจัยประเภทวิทยานิพนธ์ระดับบัณฑิตศึกษา ทั้งระดับมหาบัณฑิตและดุษฎีบัณฑิต และงานวิจัยของสถาบันต่าง ๆ โดยไม่นับรวมสารนิพนธ์หรือการค้นคว้าอิสระ โดยเป็นงานวิจัยที่มีการตีพิมพ์เผยแพร่ระหว่างปี พ.ศ. 2546-2562 เนื่องจากเป็นปีที่เริ่มมีการกำหนดกรอบแนวคิดสำหรับทักษะแห่งศตวรรษที่ 21 ภายใต้วรรณมมือระหว่าง ห้องวิจัยด้านการศึกษเขตภาคกลางตอนเหนือ ในประเทศสหรัฐอเมริกา (North Central Regional Educational Laboratory : NCREL) และกลุ่ม Metiri ในชื่อของกรอบแนวคิด “enGauge” (Metiri Group, 2003) ผู้วิจัยดำเนินการสืบค้นและคัดเลือกงานวิจัยที่สืบค้นได้จากห้องสมุดสถาบันอุดมศึกษา และฐานข้อมูลออนไลน์ เช่น เว็บไซต์โครงการเครือข่ายห้องสมุดในประเทศไทย หรือ ThaiLIS (www.thailis.or.th/tdc/), ศูนย์ข้อมูลการวิจัย Digital “วช.” (<https://dric.nrct.go.th/Index>) และเว็บไซต์ห้องสมุดของสถาบันอุดมศึกษาต่าง ๆ

สมรรถนะในศตวรรษที่ 21 ในการวิจัยครั้งนี้ได้มาจากการสังเคราะห์จากเอกสารที่นำเสนอเกี่ยวกับสมรรถนะในศตวรรษที่ 21 จากประเทศต่าง ๆ ผลการสังเคราะห์พบสมรรถนะที่สำคัญแบ่งออกเป็น 4 ด้าน คือ

- 1) ทักษะและความรู้ในเนื้อหาขั้นสูง ประกอบด้วย จิตสำนึกต่อโลก (global awareness) และความรู้เรื่องพลเมือง (civic literacy)
- 2) ทักษะการคิดขั้นสูง ประกอบด้วย การคิดอย่างมีวิจารณ์ญาณ (critical thinking), การแก้ปัญหา (solving problems) และการสร้างสรรค์และนวัตกรรม (creativity and innovation)
- 3) ทักษะต่อตนเองและสังคม ประกอบด้วย การสื่อสาร (communications literacy/skill), ทักษะด้านการทำงานเป็นทีม ความร่วมมือ (collaboration, teamwork), ภาวะผู้นำ (leadership) และความรับผิดชอบ (responsibility)
- 4) ทักษะด้านสารสนเทศ สื่อ และเทคโนโลยี ประกอบด้วย การรู้เท่าทันสื่อและสารสนเทศ (media and information literacy) และการรู้เท่าทันเทคโนโลยี (technology literacy)

วิธีดำเนินการวิจัย

ประชากร คือ งานวิจัยประเภทวิทยานิพนธ์ระดับบัณฑิตศึกษา ทั้งระดับมหาบัณฑิตและดุษฎีบัณฑิต และงานวิจัยของสถาบันต่าง ๆ ที่ศึกษาเกี่ยวกับการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา ที่มีการตีพิมพ์เผยแพร่ระหว่าง

ปีพ.ศ. 2546-2562 จำนวน 92 เล่ม แบ่งเป็น งานวิจัยประเภทวิทยานิพนธ์จาก มหาวิทยาลัยของรัฐ จำนวน 11 เล่ม มหาวิทยาลัยในกำกับของรัฐ จำนวน 43 เล่ม มหาวิทยาลัยราชภัฏ จำนวน 31 เล่ม และงานวิจัยจากสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) จำนวน 7 เล่ม

เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัยมีจำนวน 1 ฉบับ คือ แบบประเมินคุณภาพงานวิจัย และเกณฑ์การประเมินคุณภาพงานวิจัย ผู้วิจัยพัฒนาแบบประเมินคุณภาพงานวิจัย โดยปรับปรุงจากแบบประเมินคุณภาพงานวิจัยของสำนักเลขาธิการสภาการศึกษา (2552), ไชยยศ ไพวิทยศิริธรรม (2544), อาภาพรธนะ ประทุมไทย (2553) และสุวิมล เสวกสุริยวงศ์ (2553) แบบประเมินคุณภาพงานวิจัยที่พัฒนาขึ้นมีลักษณะเป็นมาตรประมาณค่า 5 ระดับ กำหนดเกณฑ์การให้คะแนนแต่ละระดับ (rubric) ตั้งแต่ 0-4 มีเนื้อหาสาระในการประเมิน 6 ด้าน ได้แก่ ด้านความเป็นมาและความสำคัญของปัญหาวิจัย, ด้านเอกสารและงานวิจัยที่เกี่ยวข้อง, ด้านวิธีดำเนินการวิจัย, ด้านผลการวิเคราะห์ข้อมูล, ด้านการสรุป อภิปรายผล และข้อเสนอแนะ และด้านการนำเสนอรายงานวิจัย จำนวน 34 ข้อ ผู้วิจัยนำแบบประเมินคุณภาพงานวิจัย และเกณฑ์การประเมินคุณภาพงานวิจัยที่พัฒนาขึ้นให้ผู้เชี่ยวชาญ จำนวน 5 ท่าน ซึ่งมีความเชี่ยวชาญในเรื่องของการวัดประเมินผลทางการศึกษา และวิจัยทางการศึกษา ตรวจสอบความตรงเชิงเนื้อหา (content validity) ผลการวิเคราะห์ค่าความสอดคล้องระหว่างเนื้อหาสาระที่ประเมินและเกณฑ์การประเมินคุณภาพงานวิจัย (IOC: index of item objective congruence) มีค่าอยู่ระหว่าง 0.80-1.00

ผู้วิจัยได้ตรวจสอบคุณภาพเครื่องมือด้วยการหาความสอดคล้องระหว่างผู้ประเมิน (inter-rater reliability) โดยคณะผู้วิจัยทั้ง 2 คน ศึกษาทำความเข้าใจเกณฑ์การให้คะแนนแต่ละระดับอย่างละเอียด จากนั้นทดลองฝึกประเมินงานวิจัยจำนวน 3 เล่ม ที่ได้จากการสุ่ม โดยประเมินงานวิจัยทีละเล่ม นำผลการประเมินของแต่ละคนมาพูดคุยเพื่อสร้างความเข้าใจให้ตรงกันก่อนมีการประเมินเล่มต่อไป และตรวจสอบความสอดคล้องในการประเมินคุณภาพงานวิจัย ตามสูตรของ Copper & Hedges (1994 อ้างถึงใน วัยญา ยิ้มวน, 2547) โดยพิจารณาจากสัดส่วนของจำนวนข้อที่ในการประเมินงานวิจัยที่สอดคล้องกันกับจำนวนข้อการประเมินงานวิจัยทั้งหมด นอกจากนี้ ผู้วิจัยได้ดำเนินการตรวจสอบความสอดคล้องภายในตัวผู้ประเมิน (intra-rater reliability) กล่าวคือ ผู้วิจัยทั้ง 2 คน ประเมินคุณภาพงานวิจัยทั้ง 3 เล่ม ซ้ำอีกครั้งหนึ่ง โดยทิ้งระยะห่างจากการประเมินครั้งแรก 2 สัปดาห์ และนำข้อมูลที่ได้จากการประเมินทั้ง 2 ครั้งมาตรวจสอบความสอดคล้องในการให้คะแนน ผลการตรวจสอบความสอดคล้องระหว่างผู้ประเมิน พบว่า มีค่าความสอดคล้องอยู่ระหว่าง 0.85-0.91 สรุปได้ว่า ผู้วิจัยทั้ง 2 คน สามารถประเมินคุณภาพงานวิจัยได้สอดคล้องใกล้เคียงกันอยู่ในระดับดีมาก ในขณะที่ผลการตรวจสอบความสอดคล้องภายในตัวผู้ประเมิน พบว่า ผู้วิจัยคนที่ 1 มีค่าความสอดคล้องอยู่ระหว่าง 0.88-0.94 และผู้วิจัยคนที่ 2 มีค่าความสอดคล้องอยู่ระหว่าง 0.88-0.97 จึงสรุปได้ว่าผู้วิจัยแต่ละคนสามารถประเมินคุณภาพงานวิจัยได้ผลการประเมินที่คงที่อยู่ในระดับดีมาก รายละเอียดดังตารางที่ 1

ตารางที่ 1 ผลการตรวจสอบความสอดคล้องระหว่างผู้ประเมินและภายในตัวผู้ประเมินจากการประเมินคุณภาพงานวิจัย

เล่มงานวิจัย	ผู้วิจัยคนที่ 1 และ 2		ผู้วิจัยคนที่ 1 ประเมินครั้งที่ 2		ผู้วิจัยคนที่ 2 ประเมินครั้งที่ 2	
	จำนวนข้อที่ประเมินตรงกัน	AR	จำนวนข้อที่ประเมินตรงกัน	AR	จำนวนข้อที่ประเมินตรงกัน	AR
เล่มที่ 1	29	0.85	30	0.88	30	0.88
เล่มที่ 2	31	0.91	32	0.94	31	0.91
เล่มที่ 3	31	0.91	32	0.94	33	0.97

(AR: Agreement Rate, ค่าความสอดคล้องในการประเมิน)

การเก็บรวบรวมข้อมูล

ผู้วิจัยทำการศึกษาทำความเข้าใจ ศึกษาเกณฑ์การประเมินคุณภาพงานวิจัย และทดลองใช้แบบประเมินคุณภาพงานวิจัย จนสามารถประเมินงานวิจัยได้ผลการประเมินที่มีความสอดคล้องกันระหว่างผู้ประเมิน จากนั้นจึงอ่านและประเมินคุณภาพงานวิจัยทั้ง 92 เล่ม

สถิติที่ใช้ในการวิจัย

การวิเคราะห์ข้อมูล เป็นการวิเคราะห์ด้วยสถิติพื้นฐาน ได้แก่ ความถี่ ค่าร้อยละ ค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน โดยมีเกณฑ์ในการแปลผลการประเมินคุณภาพวิทยานิพนธ์แบ่งออกเป็นระดับคุณภาพ 5 ระดับ ดังนี้

คะแนนประเมินอยู่ระหว่าง	3.21-4.00	หมายถึง คุณภาพดีมาก
คะแนนประเมินอยู่ระหว่าง	2.41-3.20	หมายถึง คุณภาพดี
คะแนนประเมินอยู่ระหว่าง	1.61-2.40	หมายถึง คุณภาพปานกลาง
คะแนนประเมินอยู่ระหว่าง	0.81-1.60	หมายถึง คุณภาพค่อนข้างต่ำ
คะแนนประเมินต่ำกว่า	0.80	หมายถึง คุณภาพต่ำ

สรุปผลการวิจัย

1. ผลการประเมินระดับคุณภาพงานวิจัย

งานวิจัยที่นำมาศึกษาในครั้งนี้มีจำนวน 92 เล่ม เป็นงานวิจัยในมหาวิทยาลัยในกำกับของรัฐ จำนวนมากที่สุด 43 เล่ม (ร้อยละ 46.74) รองลงมาคือ มหาวิทยาลัยราชภัฏ จำนวน 31 เล่ม (ร้อยละ 33.70) มหาวิทยาลัยของรัฐ จำนวน 11 เล่ม (ร้อยละ 11.96) และสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) จำนวน 7 เล่ม (ร้อยละ 7.61) ตามลำดับ ผลการประเมินเพื่อจัดระดับคุณภาพ พบว่า ในภาพรวมส่วนใหญ่เป็นงานวิจัยที่มีคุณภาพอยู่ในระดับดี จำนวน 76 เล่ม คิดเป็นร้อยละ 82.61 ที่เหลือเป็นงานวิจัยที่มีคุณภาพอยู่ในระดับดีมาก จำนวน 16 เล่ม (ร้อยละ 17.39) ทั้งนี้ไม่มีงานวิจัยเล่มใดที่มีคุณภาพอยู่ในระดับปานกลาง คุณภาพค่อนข้างต่ำ และคุณภาพต่ำ

เมื่อพิจารณาจำแนกตามประเภทสถาบันที่ผลิตงานวิจัย พบว่า งานวิจัยที่มีคุณภาพในระดับดีมากเป็นงานวิจัยจากมหาวิทยาลัยในกำกับของรัฐมากที่สุด จำนวน 12 เล่ม (ร้อยละ 13.04) รองลงมาคือ งานวิจัยจากสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) จำนวน 3 เล่ม (ร้อยละ 3.26) และมหาวิทยาลัยราชภัฏ จำนวน 1 เล่ม (ร้อยละ 1.09) ตามลำดับ ในขณะที่งานวิจัยที่มีคุณภาพในระดับดี เป็นงานวิจัยจากมหาวิทยาลัยในกำกับของรัฐมากที่สุด จำนวน 31 เล่ม (ร้อยละ 33.70) รองลงมาคือ งานวิจัยจากมหาวิทยาลัยราชภัฏ จำนวน 30 เล่ม (ร้อยละ 32.61) มหาวิทยาลัยของรัฐ จำนวน 11 เล่ม (ร้อยละ 11.96) และสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) จำนวน 4 เล่ม (ร้อยละ 4.35) ตามลำดับ

ทั้งนี้ เมื่อพิจารณาคุณภาพงานวิจัยตามตัวแปรประเภทงานวิจัย พบว่า งานวิจัยที่มีคุณภาพในระดับดีมากเป็นงานวิจัยประเภทวิทยานิพนธ์ในระดับปริญญาโทมากที่สุด จำนวน 10 เล่ม (ร้อยละ 10.87) รองลงมาคืองานวิจัยประเภทวิทยานิพนธ์ในระดับปริญญาเอก และรายงานวิจัย มีจำนวนเท่ากัน 3 เล่ม (ร้อยละ 3.26) ในขณะที่งานวิจัยที่มีคุณภาพในระดับดี เป็นงานวิจัยประเภทวิทยานิพนธ์ในระดับปริญญาโทมากที่สุด จำนวน 70 เล่ม (ร้อยละ 76.09) รองลงมาคือ รายงานวิจัย จำนวน 4 เล่ม (ร้อยละ 4.35) และงานวิจัยประเภทวิทยานิพนธ์ในระดับปริญญาเอก จำนวน 2 เล่ม (ร้อยละ 2.17) ตามลำดับ รายละเอียดดังตารางที่ 2

ตารางที่ 2 ระดับคุณภาพงานวิจัยจำแนกตามประเภทสถาบันที่ผลิตงานวิจัยและประเภทงานวิจัย

คุณลักษณะงานวิจัย	ดี		ดีมาก		รวม	
	ความถี่	ร้อยละ	ความถี่	ร้อยละ	ความถี่	ร้อยละ
ประเภทสถาบันที่ผลิตงานวิจัย						
1. มหาวิทยาลัยของรัฐ	11	11.96	0	0.00	11	11.96
2. มหาวิทยาลัยในกำกับของรัฐ	31	33.70	12	13.04	43	46.74
3. มหาวิทยาลัยราชภัฏ	30	32.61	1	1.09	31	33.70
4. สทศ.	4	4.35	3	3.26	7	7.61
รวม	76	82.61	16	17.39	92	100.00
ประเภทงานวิจัย						
1. วิทยานิพนธ์ ป.โท	70	76.09	10	10.87	80	86.96
2. วิทยานิพนธ์ ป.เอก	2	2.17	3	3.26	5	5.43
3. รายงานวิจัย	4	4.35	3	3.26	7	7.61
รวม	76	82.61	16	17.39	92	100.00

2. ผลการประเมินคุณภาพงานวิจัยรายด้าน

เมื่อพิจารณาคุณภาพงานวิจัยในภาพรวม พบว่า มีคุณภาพอยู่ในระดับดี ($M = 2.997$, $SD = 0.247$) เมื่อพิจารณาจำแนกรายด้าน พบว่า ด้านผลการวิเคราะห์ข้อมูล มีค่าเฉลี่ยสูงสุดมีคุณภาพอยู่ในระดับดีมาก ($M = 3.875$, $SD = 0.382$) รองลงมาคือ ด้านการสรุป อภิปรายผล และข้อเสนอแนะ ($M = 3.214$, $SD = 0.473$) ด้านวิธีดำเนินการวิจัย ($M = 3.046$, $SD = 0.370$) ด้านการนำเสนอรายงานวิจัย ($M = 2.783$, $SD = 0.340$) ด้านความเป็นมาและความสำคัญของปัญหาวิจัย ($M = 2.634$, $SD = 0.457$) และเอกสารและงานวิจัยที่เกี่ยวข้อง ($M = 2.429$, $SD = 0.340$) ตามลำดับ โดยมีคุณภาพอยู่ในระดับดีทุกด้าน ยกเว้น ด้านการสรุป อภิปรายผล และข้อเสนอแนะที่มีคุณภาพในระดับดีมาก

เมื่อพิจารณาคุณภาพงานวิจัยตามรายการประเมิน พบว่า คุณภาพงานวิจัยที่มีค่าสูงสุด 5 อันดับแรก มีคุณภาพอยู่ในระดับดีมากทุกรายการ ได้แก่ เอกสารและงานวิจัยที่เกี่ยวข้องมีปริมาณมากพอ ($M = 3.989$, $SD = 0.104$) รองลงมาคือ รูปแบบรายงานถูกต้องตามหลักวิชา ($M = 3.957$, $SD = 0.205$) ชื่อเรื่องมีความชัดเจน น่าสนใจ ($M = 3.946$, $SD = 0.228$) การแปลผลและสรุปผลการวิเคราะห์ข้อมูลมีความถูกต้องชัดเจน ($M = 3.946$, $SD = 0.310$) และลักษณะการนำเสนอการแปลผลการวิเคราะห์ข้อมูล ($M = 3.804$, $SD = 0.650$) ตามลำดับ ในขณะที่คุณภาพงานวิจัยในภาพรวมที่มีค่าต่ำสุด 5 อันดับแรก มีคุณภาพอยู่ในระดับค่อนข้างต่ำ 3 รายการ ได้แก่ เอกสารและงานวิจัยที่เกี่ยวข้องมีความทันสมัย ($M = 1.130$, $SD = 0.667$) เป็นงานวิจัยที่นำไปใช้ประโยชน์ได้ในทางวิชาการ ($M = 1.348$, $SD = 0.748$) และเหตุผลและความจำเป็นในการทำวิจัยมีความสมเหตุสมผล ($M = 1.565$, $SD = 0.816$) ในขณะที่อีก 2 รายการมีคุณภาพในระดับปานกลาง ได้แก่ การนิยามศัพท์เฉพาะมีความชัดเจน ($M = 1.728$, $SD = 1.080$) และการสร้างเครื่องมือวัดสมรรถนะในศตวรรษที่ 21 มีการกำหนดแผนผังการสร้างข้อคำถามที่ครอบคลุมเนื้อหา ($M = 2.196$, $SD = 1.361$) ตามลำดับ รายละเอียดดังตารางที่ 3

ตารางที่ 3 ผลการประเมินคุณภาพงานวิจัย

รายการประเมิน	M	SD	ระดับคุณภาพ
ความเป็นมาและความสำคัญของปัญหาวิจัย	2.634	0.457	ดี
1. ชื่อเรื่องมีความชัดเจน น่าสนใจ	3.946	0.228	ดีมาก
2. ความเป็นมาและความสำคัญของปัญหา ชัดเจน สอดคล้องกับเรื่องที่ทำวิจัย	3.022	0.851	ดี
3. ปัญหาวิจัย/วัตถุประสงค์มีความสอดคล้องกับชื่อเรื่อง ถูกต้องตามหลักการวิจัย	3.163	0.700	ดี
4. เหตุผลและความจำเป็นในการทำวิจัยมีความสมเหตุสมผล	1.565	0.816	ค่อนข้างต่ำ
5. กำหนดขอบเขตของการวิจัยกำหนดได้อย่างเหมาะสม และมีเหตุผลรองรับ	2.478	0.687	ดี
6. การนิยามศัพท์เฉพาะมีความชัดเจน	1.728	1.080	ปานกลาง
7. กรอบแนวคิดในการวิจัยมีความถูกต้องเหมาะสม ชัดเจนตามหลักการวิจัย	2.533	1.190	ดี
เอกสารและงานวิจัยที่เกี่ยวข้อง	2.429	0.340	ดี
8. เอกสารและงานวิจัยที่เกี่ยวข้องมีปริมาณมากพอ	3.989	0.104	ดีมาก
9. เอกสารและงานวิจัยที่เกี่ยวข้องมีความสอดคล้องกับปัญหา/วัตถุประสงค์ของการวิจัย	2.326	0.758	ปานกลาง
10. เอกสารและงานวิจัยที่เกี่ยวข้องมีความทันสมัย	1.130	0.667	ค่อนข้างต่ำ
11. เอกสารและงานวิจัยที่เกี่ยวข้องทั้งภาษาไทย และภาษาต่างประเทศมีสัดส่วนที่เหมาะสม	2.272	0.537	ปานกลาง
วิธีดำเนินการวิจัย	3.046	0.370	ดี
12. การออกแบบการวิจัยสอดคล้องกับปัญหาการวิจัย	3.457	0.762	ดีมาก
13. ขั้นตอนการวิจัยมีความชัดเจน	3.054	0.906	ดี
14. กำหนดประชากรและตัวอย่างได้อย่างเหมาะสมกับการวิจัย	3.380	0.850	ดีมาก
15. ขนาดตัวอย่างและวิธีการสุ่มมีความถูกต้องเหมาะสม	3.424	0.917	ดีมาก
16. การกำหนดจุดมุ่งหมายในการวัดสมรรถนะในศตวรรษที่ 21 มีความชัดเจน	2.239	1.133	ปานกลาง
17. การสร้างเครื่องมือวัดสมรรถนะในศตวรรษที่ 21 มีแนวคิดทฤษฎีเกี่ยวกับตัวแปรที่วัดและวิธีการสร้างเครื่องมือ	2.978	0.961	ดี
18. การสร้างเครื่องมือวัดสมรรถนะในศตวรรษที่ 21 มีการกำหนดแผนผังการสร้างข้อคำถามที่ครอบคลุมเนื้อหา	2.196	1.361	ปานกลาง
19. เครื่องมือวัดสมรรถนะในศตวรรษที่ 21 ที่สร้างมีความเป็นปรนัยของข้อคำถาม	3.054	1.052	ดี

รายการประเมิน	M	SD	ระดับคุณภาพ
20. ขั้นตอนในการดำเนินการวัดสมรรถนะในศตวรรษที่ 21 มีความเหมาะสม	3.054	0.942	ดี
21. วิธีการตรวจสอบคุณภาพเครื่องมือวัดสมรรถนะในศตวรรษที่ 21 มีความเหมาะสม	3.783	0.415	ดีมาก
22. มีการทดลองใช้เครื่องมือวัดสมรรถนะในศตวรรษที่ 21 (try out)	2.870	1.112	ดี
23. ตัวอย่างที่ใช้ในการทดลองเครื่องมือวัดสมรรถนะในศตวรรษที่ 21 มีความเหมาะสม	3.707	0.688	ดีมาก
24. เครื่องมือวัดสมรรถนะในศตวรรษที่ 21 ในภาพรวมมีคุณภาพเหมาะสม	2.793	0.989	ดี
25. กระบวนการเก็บรวบรวมข้อมูลมีความชัดเจนและเหมาะสม	3.043	0.982	ดี
26. สถิติที่ใช้ในการวิเคราะห์ข้อมูลมีความถูกต้องเหมาะสมกับลักษณะของข้อมูล	2.652	0.501	ดี
ผลการวิเคราะห์ข้อมูล	3.875	0.382	ดีมาก
27. ลักษณะการนำเสนอการแปลผลการวิเคราะห์ข้อมูล	3.804	0.650	ดีมาก
28. การแปลผลและสรุปผลการวิเคราะห์ข้อมูลมีความถูกต้องชัดเจน	3.946	0.310	ดีมาก
การสรุปอภิปรายผล และข้อเสนอแนะ	3.214	0.473	ดีมาก
29. สรุปผลการวิจัยอย่างถูกต้อง	3.467	0.523	ดีมาก
30. มีการอภิปรายผลการวิจัยสอดคล้องกับผลงานวิจัยในอดีต	3.272	0.786	ดีมาก
31. ข้อเสนอแนะมีความชัดเจนและเป็นประโยชน์	2.902	0.865	ดี
การนำเสนอรายงานวิจัย	2.783	0.340	ดี
32. เป็นงานวิจัยที่นำไปใช้ประโยชน์ได้ในทางปฏิบัติ	3.043	0.417	ดี
33. เป็นงานวิจัยที่นำไปใช้ประโยชน์ได้ในทางวิชาการ	1.348	0.748	ค่อนข้างต่ำ
34. รูปแบบรายงานถูกต้องตามหลักวิชา	3.957	0.205	ดีมาก
ภาพรวม	2.997	0.247	ดี

อภิปรายผลการวิจัย

ผลการประเมินคุณภาพงานวิจัย พบว่างานวิจัยมีคุณภาพอยู่ในระดับดี ร้อยละ 82.61 และดีมาก ร้อยละ 17.39 ทั้งนี้เป็นเพราะงานวิจัยที่นำมาใช้ในการวิจัยครั้งนี้เป็นงานวิจัยประเภทวิทยานิพนธ์ในระดับบัณฑิตศึกษาทั้งระดับมหาบัณฑิต และดุษฎีบัณฑิต รวมถึงเป็นรายงานการวิจัยจากสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) การที่งานวิจัยการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา มีคุณภาพอยู่ในระดับดีและดีมากนั้น เป็นเพราะงานวิจัยประเภทวิทยานิพนธ์ อยู่ภายใต้การดูแล ให้ข้อเสนอแนะ และควบคุมคุณภาพ โดยอาจารย์ที่ปรึกษาวิทยานิพนธ์ รวมถึงผ่านกระบวนการสอบจากคณะกรรมการสอบวิทยานิพนธ์ มีจำนวน 3-5 คน ซึ่งเป็นผู้เชี่ยวชาญในการพัฒนาเครื่องมือวัดจึงทำให้ได้ผลงานที่มีคุณภาพในระดับหนึ่ง ในขณะที่รายงานการวิจัยที่ได้รับทุนสนับสนุนการวิจัยจากสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) นักวิจัยผู้ได้รับทุนต่างเป็นคณาจารย์ในสาขาวิชาการวัดและประเมินผล รวมถึงวิจัยและสถิติทางการศึกษา ซึ่งเป็นผู้ที่มีความรู้ และความเชี่ยวชาญในการพัฒนา

เครื่องมือวัด ทำให้งานวิจัยในภาพรวมมีคุณภาพในระดับดีขึ้นไป ทั้งนี้ ผลการวิจัยดังกล่าว สอดคล้องกับงานวิจัยของ สุวิมล เสวกสุริยวงศ์ (2553) ที่ได้ประเมินคุณภาพงานวิจัยเกี่ยวกับการสร้างแบบสอบถามวิจัยทางคณิตศาสตร์ จำนวน 75 เล่ม ซึ่งมีการตีพิมพ์ระหว่างปี พ.ศ. 2523-2551 พบว่า ผลการประเมินคุณภาพงานวิจัยอยู่ในระดับดีมาก มีจำนวนมากที่สุด จำนวน 45 เล่ม (ร้อยละ 60.00) รองลงมาคือ มีคุณภาพในระดับดี จำนวน 29 เล่ม (ร้อยละ 38.70) และคุณภาพในระดับปานกลางมีเพียงจำนวน 1 เล่ม (ร้อยละ 1.30) รวมถึงงานวิจัยของ รัตนะ บัวสนธิ์ และคณะ (2557) ที่ได้ประเมินคุณภาพงานวิจัย ประเภทวิทยานิพนธ์ของนิสิตระดับดุขุภีบัณฑิต มหาวิทยาลัยนเรศวร ระหว่างปี พ.ศ. 2550-2554 จำนวน 30 เล่ม พบว่า ผลการประเมินคุณภาพงานวิจัยร้อยละ 53.33 อยู่ในระดับดีมาก และร้อยละ 43.33 มีคุณภาพระดับดี

เมื่อพิจารณาผลการประเมินคุณภาพรายด้าน ในด้านความเป็นมาและความสำคัญของปัญหาวิจัยมีผลการประเมินในภาพรวมอยู่ในระดับดี เมื่อพิจารณาในรายละเอียดพบว่า ผลการประเมินที่อยู่ในระดับดีมากมี 1 รายการ คือ ชื่อเรื่องมีความชัดเจน น่าสนใจ กล่าวคือ ชื่อเรื่องงานวิจัยโดยส่วนใหญ่มีการระบุถึงปัญหาวิจัย ระบุตัวแปรสำคัญที่ศึกษา และระบุตัวอย่างที่ศึกษาได้อย่างครบถ้วน อย่างไรก็ตาม มีบางประเด็นที่ควรให้ความสำคัญมากขึ้น ได้แก่ เหตุผลและความจำเป็นในการทำวิจัยมีความสมเหตุสมผล ที่มีระดับคุณภาพค่อนข้างต่ำ และการนิยามศัพท์เฉพาะมีความชัดเจน ที่มีระดับคุณภาพปานกลาง ทั้งนี้ เนื่องจากงานวิจัยที่พัฒนาเครื่องมือวัดถึงแม้จะเป็นเรื่องที่ที่น่าสนใจ แต่ผู้วิจัยควรให้เหตุผลประกอบว่าเป็นเรื่องที่ได้องค์ความรู้ใหม่อย่างไร เป็นประเด็นที่ไม่เคยมีผู้ใดศึกษามาก่อน และเป็นปัญหาเร่งด่วนที่ต้องมีการพัฒนาเครื่องมือวัดเนื่องจากเหตุผลใด สำหรับการนิยามศัพท์เฉพาะมีความชัดเจน ที่มีผลการประเมินในระดับปานกลางนั้น เนื่องมาจากงานวิจัยส่วนใหญ่มีการนิยามเชิงทฤษฎี หรือนิยามเชิงโครงสร้าง (construct definition) ที่อิงทฤษฎีได้อย่างครบถ้วน ซึ่งนิยามเชิงทฤษฎีนั้น สุวิมล ติรกันันท์ (2551) กล่าวว่า เป็นการให้ความหมายเชิงทฤษฎี และแนวคิดต่าง ๆ ที่เกี่ยวข้อง ทำให้เข้าใจถึงลักษณะและความหมายของตัวแปรมากขึ้น แต่ยังไม่สามารถสังเกตได้ หรือยังไม่เห็นส่วนประกอบอย่างชัดเจน มีลักษณะเป็นนามธรรม แต่ทั้งนี้ผลการวิจัยพบว่า ยังขาดการนิยามเชิงปฏิบัติการ (operational definition) ซึ่งเป็นการให้ความหมายเกี่ยวกับตัวแปรที่ศึกษาให้อยู่ในรูปของพฤติกรรมหรือรูปรธรรมซึ่งสามารถวัดได้ รวมทั้งต้องมีเนื้อหาสอดคล้องกับเนื้อหาในนิยามเชิงทฤษฎี (คณะกรรมการสภานโยบายแห่งชาติ สาขาศึกษา, 2555; บุญใจ ศรีสถิตยรรณกร, 2555) อย่างไรก็ตาม บุญใจ ศรีสถิตยรรณกร (2555) กล่าวว่า ตัวแปรที่ต้องมีการนิยามเชิงปฏิบัติการ คือ ตัวแปรที่มีนิยามเชิงทฤษฎีที่เป็นนามธรรม (abstract concept) บุคคลทั่วไปหรือกลุ่มตัวอย่าง ได้รับความหมายของตัวแปรหรือแปลความหมายตัวแปรไม่ตรงกัน ทั้งนี้ หากขาดการนิยามเชิงปฏิบัติการ หรือมีการเขียนนิยามเชิงปฏิบัติการที่ไม่กระจ่างชัดเจน จะทำให้เนื้อหาในนิยามเชิงปฏิบัติการไม่ถูกต้อง ย่อมส่งผลให้เครื่องมือวัดที่สร้างขึ้นขาดความตรงตามเนื้อหา อย่างไรก็ตาม ผลการประเมินดังกล่าวมีความสอดคล้องกับผลการวิจัยของ อาภาพรณ ประทุมไทย (2553) ที่ได้ประเมินคุณภาพงานวิจัยเกี่ยวกับการสร้างแบบวัดความถนัด จำนวน 61 เล่ม ที่มีการตีพิมพ์

เผยแพร่ระหว่างปี พ.ศ. 2509-2551 และผลการวิจัยของ สรัญญา จันทร์ชูสกุล (2561) ที่ได้ประเมินคุณภาพงานวิจัยนวัตกรรมการพัฒนาความคิดสร้างสรรค์ของนักเรียนระดับประถมศึกษาจำนวน 118 เล่ม ที่มีการตีพิมพ์เผยแพร่ระหว่างปี พ.ศ. 2546-2558 ผลการวิจัยพบว่า เหตุผลและความจำเป็นในการทำวิจัยมีความสมเหตุสมผล มีระดับคุณภาพค่อนข้างต่ำ และการนิยามศัพท์เฉพาะมีความชัดเจน มีระดับคุณภาพปานกลาง เช่นเดียวกัน

สำหรับผลคุณภาพด้านเอกสารและงานวิจัยที่เกี่ยวข้องในภาพรวมอยู่ในระดับดี เมื่อพิจารณาในรายละเอียดพบว่า ผลการประเมินที่อยู่ในระดับดีมาก มี 1 รายการ คือ เอกสารและงานวิจัยที่เกี่ยวข้องมีปริมาณมากพอ อย่างไรก็ตามก็ยังมีบางประเด็นที่ควรให้ความสำคัญมากขึ้น ได้แก่ เอกสารและงานวิจัยที่เกี่ยวข้องมีความทันสมัย ที่มีผลการประเมินอยู่ในระดับค่อนข้างต่ำ รวมถึงเอกสารและงานวิจัยที่เกี่ยวข้องมีความสอดคล้องกับปัญหา/วัตถุประสงค์ของการวิจัย และเอกสารและงานวิจัยที่เกี่ยวข้องทั้งภาษาไทยและภาษาต่างประเทศมีสัดส่วนที่เหมาะสม ที่มีผลการประเมินอยู่ในระดับปานกลาง สำหรับประเด็นของความทันสมัยของเอกสารและงานวิจัยนั้น พบว่า งานวิจัยโดยส่วนใหญ่มีการทบทวนเอกสารและงานวิจัยที่เกี่ยวข้องที่มีอายุระหว่าง 1-5 ปี มีปริมาณไม่ถึงร้อยละ 30 นอกจากนี้ เอกสารและงานวิจัยที่เกี่ยวข้องมีสัดส่วนเป็นภาษาไทยเป็นส่วนมาก มีสัดส่วนเอกสารและงานวิจัยภาษาต่างประเทศไม่ถึงร้อยละ 20 ทั้งนี้ ในการทบทวนเอกสารหรืองานวิจัยที่เกี่ยวข้อง นักวิจัยอาจจะอ้างอิงเอกสารต้นฉบับที่เป็นเอกสารขั้นครูได้ แต่ควรมีเอกสารที่ทันสมัยที่แสดงให้เห็นพัฒนาการจากแนวคิดนั้นด้วย โดยเฉพาะอย่างยิ่งงานวิจัยที่นำมาใช้ศึกษาอ้างอิงควรเป็นงานวิจัยใหม่ ๆ ที่ทันสมัย เนื่องจากหากเป็นการศึกษาจากงานวิจัยเก่า ๆ การนำไปใช้ประโยชน์ของนักวิจัยจะอยู่ในวงที่จำกัด เพราะงานวิจัยเป็นการทำซ้ำ ทำให้สูญเสียทรัพยากรโดยไม่จำเป็น (คณะกรรมการสภาวิจัยแห่งชาติ สาขาการศึกษา, 2555) ผลการประเมินดังกล่าวมีความสอดคล้องกับผลการวิจัยของ สุวิมล เสวกสุริยวงศ์ (2553) ที่ได้ประเมินคุณภาพงานวิจัยเกี่ยวกับการสร้างแบบสอบวินิจัยทางคณิตศาสตร์ และงานวิจัยของ อาภาพรณ ประทุมไทย (2553) ที่ได้ประเมินคุณภาพงานวิจัยเกี่ยวกับการสร้างแบบวัดความถนัด พบว่า เอกสารและงานวิจัยที่เกี่ยวข้องมีความทันสมัย มีผลการประเมินอยู่ในระดับค่อนข้างต่ำเช่นเดียวกัน

ผลการประเมินคุณภาพงานวิจัยด้านวิธีดำเนินการวิจัยในภาพรวมอยู่ในระดับดี เมื่อพิจารณาในรายละเอียดพบว่า ผลการประเมินที่อยู่ในระดับดีมาก มี 5 รายการ ได้แก่ การออกแบบการวิจัยสอดคล้องกับปัญหาการวิจัย, กำหนดประชากรและตัวอย่างได้อย่างเหมาะสมกับการวิจัย, ขนาดตัวอย่างและวิธีการสุ่มมีความถูกต้องเหมาะสม, วิธีการตรวจสอบคุณภาพเครื่องมือวัดสมรรถนะในศตวรรษที่ 21 มีความเหมาะสม และตัวอย่างที่ใช้ในการทดลองเครื่องมือวัดสมรรถนะในศตวรรษที่ 21 มีความเหมาะสม อย่างไรก็ตามก็ยังมีบางประเด็นที่ควรให้ความสำคัญมากขึ้น ได้แก่ การกำหนดจุดมุ่งหมายในการวัดสมรรถนะในศตวรรษที่ 21 มีความชัดเจน และการสร้างเครื่องมือวัดสมรรถนะในศตวรรษที่ 21 มีการกำหนดแผนผังการสร้างข้อคำถามที่ครอบคลุมเนื้อหา ซึ่งผลการประเมินที่อยู่ในระดับปานกลาง กล่าวคือ งานวิจัยบางส่วนยังไม่มี

การกำหนดจุดมุ่งหมายในการวัดสมรรถนะในศตวรรษที่ 21 หรือมีการกำหนดแต่ไม่ชัดเจน หรือไม่สอดคล้องกับการนำผลไปใช้ประโยชน์ อย่างไรก็ดี Newton (2007 อ้างถึงใน ณัฐกรณ์ หลาวทอง, 2559) กล่าวว่า การกำหนดจุดมุ่งหมายของการสร้างเครื่องมือวัดทางการวิจัย และการวัด ประเมินผลทางการศึกษา แบ่งออกเป็น 18 ประการ ได้แก่

- | | |
|-----------------------------------|---------------------------------------|
| 1) ใช้เพื่อการประเมินทางสังคม | 10) ใช้เพื่อกำหนดใบอนุญาต |
| 2) ใช้เพื่อการพัฒนา | 11) ใช้เพื่อการเลือกโรงเรียน |
| 3) ใช้เพื่อการติดตามกำกับผู้เรียน | 12) ใช้เพื่อการติดตามสถาบัน |
| 4) ใช้เพื่อการเทียบโอน | 13) ใช้เพื่อการจัดสรรทรัพยากร |
| 5) ใช้เพื่อการจัดวางตำแหน่ง | 14) ใช้เพื่อดำเนินการขององค์กร |
| 6) ใช้เพื่อการวินิจฉัย | 15) ใช้เพื่อประเมินโปรแกรม |
| 7) ใช้เพื่อการแนะแนว | 16) ใช้เพื่อการติดตามระบบ |
| 8) ใช้เพื่อวัดคุณสมบัติ | 17) ใช้เพื่อการเปรียบเทียบ/เทียบเคียง |
| 9) ใช้เพื่อการคัดเลือก | 18) ใช้เพื่อการประกันคุณภาพระดับชาติ |

สำหรับการสร้างเครื่องมือวัดสมรรถนะในศตวรรษที่ 21 มีการกำหนดแผนผังการสร้าง ข้อคำถามที่ครอบคลุมเนื้อหาที่มีผลการประเมินที่อยู่ในระดับปานกลางนั้น พบว่า มีงานวิจัย ร้อยละ 18.48 ไม่กำหนดแผนผังการสร้างข้อคำถาม ในขณะที่งานวิจัยร้อยละ 10.87 มีการกำหนด แผนผังการสร้างข้อคำถาม แต่ไม่นำเสนอในเล่มรายงานวิจัย อย่างไรก็ดี มีงานวิจัยร้อยละ 70.65 มีการกำหนดแผนผังการสร้างข้อคำถาม แต่ยังมีบางส่วนที่ขาดความเหมาะสม เช่น ไม่มีการแยก รายละเอียดตามองค์ประกอบ หรือตัวชี้วัดที่ต้องการวัด หรือขาดการระบุเลขข้อ จำนวนข้อที่วัด ตามองค์ประกอบนั้น ๆ เป็นต้น ผลการประเมินดังกล่าวมีความสอดคล้องกับผลการวิจัยของ งานวิจัยของ อภาพรรณ ประทุมไทย (2553) ที่ได้ประเมินคุณภาพงานวิจัยที่เกี่ยวกับการสร้าง แบบวัดความถนัด พบว่า การกำหนดจุดมุ่งหมายในการสร้างแบบวัดมีความชัดเจนอยู่ในระดับ ปานกลางแต่การสร้างแบบวัดมีการสร้างตารางกำหนดแผนผังการสร้างข้อสอบที่ครอบคลุม เนื้อหาที่มีคุณภาพในระดับต่ำ

ผลการประเมินคุณภาพงานวิจัยด้านผลการวิเคราะห์ข้อมูลในภาพรวมอยู่ในระดับดีมาก เมื่อพิจารณาในรายละเอียดพบว่า ผลการประเมินที่อยู่ในระดับดีมาก ทั้ง 2 รายการ ได้แก่ ลักษณะ การนำเสนอการแปลผลการวิเคราะห์ข้อมูล และการแปลผลและสรุปผลการวิเคราะห์ข้อมูลมี ความถูกต้องชัดเจน กล่าวคือ มีการนำเสนอผลการวิเคราะห์ข้อมูลที่ควรจะนำเสนออย่างครบถ้วน วิธีนำเสนอเหมาะสม เช่น นำเสนอในรูปของตาราง มีการอธิบายผลการวิเคราะห์ รวมถึงมี การแปลความหมายหรือสรุปผลการวิเคราะห์ข้อมูลครบทุกส่วน ถูกต้องและสอดคล้องกับ ผลวิเคราะห์ที่ได้ ทั้งนี้ วรณัฏฐ์ แกมเกตุ (2555) กล่าวว่า การรายงานผลการวิเคราะห์ข้อมูล ควรประกอบด้วย 3 ส่วนหลัก คือ

- 1) การวิเคราะห์ข้อมูลเบื้องต้นเกี่ยวกับกลุ่มตัวอย่างที่ใช้ในการวิจัย
- 2) ผลการวิเคราะห์ข้อมูลเพื่อตรวจสอบว่าข้อมูลเป็นไปตามข้อตกลงเบื้องต้นหรือไม่
- 3) ผลการวิเคราะห์ข้อมูลเพื่อตอบวัตถุประสงค์การวิจัย

โดยการเขียนบรรยายผลการวิเคราะห์ข้อมูลควรเขียนในลักษณะการแปลความหรือตีความค่าสถิติจากผลการวิเคราะห์ มิใช่เป็นการอ่านตาราง และในกรณีที่เป็นสถิติใหม่ ควรเขียนอธิบายความหมายของค่าสถิตินั้นก่อน แล้วนำไปสู่การตีความ นอกจากนี้ ผู้วิจัยควรออกแบบการนำเสนอตารางที่ต้องการนำเสนอในเนื้อหาในลักษณะเดียวกัน ควรออกแบบในตารางเดียวกัน เพื่อให้การนำเสนอมีความกระชับ และควรใช้รูปแบบตารางเดียวกันตลอดทั้งเล่ม ผลการวิจัยดังกล่าว สอดคล้องกับผลการศึกษาของ สุวิมล เสวกสุริยวงศ์ (2553) ที่ได้ประเมินคุณภาพงานวิจัยเกี่ยวกับการสร้างแบบสอบวินิจฉัยทางคณิตศาสตร์ ที่พบว่ามีผลการประเมินทั้ง 2 รายการอยู่ในระดับดีมาก เช่นกัน แต่ผลการศึกษาของ อาภาพรณ ประทุมไทย (2553) ที่ได้ประเมินคุณภาพงานวิจัยที่เกี่ยวกับการสร้างแบบวัดความถนัด พบว่า งานวิจัยมีลักษณะการนำเสนอ การแปลผล การวิเคราะห์ข้อมูลอยู่ในระดับดีมาก แต่มีการแปลผล และสรุปผลการวิเคราะห์ข้อมูลมีความถูกต้องอยู่ในระดับดี

ผลการประเมินคุณภาพงานวิจัยด้านการสรุป อภิปรายผล และข้อเสนอแนะ ในภาพรวมอยู่ในระดับดีมาก เมื่อพิจารณาในรายละเอียด พบว่า ผลการประเมินที่อยู่ในระดับดีมาก 2 รายการ คือ สรุปผลการวิจัยอย่างถูกต้อง และมีการอภิปรายผลการวิจัยสอดคล้องกับผลงานวิจัยในอดีต กล่าวคือ งานวิจัยส่วนใหญ่ สามารถสรุปผลวิจัยถูกต้องสอดคล้องกับวัตถุประสงค์การวิจัย ตอบปัญหาวิจัยครอบคลุมทุกหัวข้อ ใช้ข้อความกระชับ อ่านเข้าใจง่าย อ่านแล้วเข้าใจถึงประเด็นที่ต้องการสรุปในหัวข้อนั้น ๆ อย่างชัดเจน รวมถึงมีการอภิปรายผลการวิจัยที่ได้กับผลงานวิจัยในอดีตอย่างสมเหตุสมผล ครอบคลุมทุกประเด็นปัญหาวิจัย และมีผลการประเมินที่อยู่ในระดับดี 1 รายการ คือ ข้อเสนอแนะมีความชัดเจนและเป็นประโยชน์ กล่าวคือ มีการให้ข้อเสนอแนะถึงการนำผลการวิจัยไปใช้มาจากการวิจัยมีเหตุผลรองรับเพียงพอ แต่อย่างไรก็ดี สุวิมล ตีรภานันท์ (2557) กล่าวว่าสิ่งที่ผู้วิจัยจำเป็นต้องคำนึงเสมอในการเขียนข้อเสนอแนะ คือ ข้อเสนอแนะนั้นจะต้องเป็นสาระที่เกิดขึ้นจากผลการวิจัย มิใช่ความรู้สึก หรือความคิดเห็นของผู้วิจัย โดยจะต้องนำเสนอในสิ่งใหม่ ไม่ใช่เรื่องทราบโดยทั่วไป เป็นข้อเสนอที่สามารถปฏิบัติได้ขอบเขตเงื่อนไขที่กำหนดไว้โดยจะต้องเป็นข้อเสนอที่ต่อเนื่องหรือต่อยอดจากงานวิจัย ผลการวิจัยดังกล่าว สอดคล้องกับผลการศึกษาของ สุวิมล เสวกสุริยวงศ์ (2553) ที่ได้ประเมินคุณภาพงานวิจัยเกี่ยวกับการสร้างแบบสอบวินิจฉัยทางคณิตศาสตร์ และ สรัญญา จันทรชุตกุล (2561) ที่ได้ประเมินคุณภาพงานวิจัยนวัตกรรมการพัฒนาความคิดสร้างสรรค์ของนักเรียนระดับประถมศึกษา พบว่า การสรุปผลการวิจัยอย่างถูกต้องมีคุณภาพอยู่ในระดับดีมาก และข้อเสนอแนะมีความชัดเจนและเป็นประโยชน์อยู่ในระดับดีเช่นเดียวกัน ทั้งนี้ ยกเว้นรายการมีการอภิปรายผลการวิจัยสอดคล้องกับผลงานวิจัยในอดีตที่ผลการประเมินอยู่ในระดับดี

ผลการประเมินคุณภาพงานวิจัยด้านการนำเสนอรายงานวิจัย ในภาพรวมอยู่ในระดับดี เมื่อพิจารณาในรายละเอียดพบว่า ผลการประเมินที่อยู่ในระดับดีมาก 1 รายการ คือ รูปแบบรายงาน

ถูกต้องตามหลักวิชา อย่างไรก็ดี มีประเด็นที่ควรให้ความสำคัญมากขึ้น คือ เป็นงานวิจัยที่นำไปใช้ประโยชน์ได้ในทางวิชาการ ที่มีระดับคุณภาพค่อนข้างต่ำ กล่าวคือ งานวิจัยส่วนใหญ่มีประโยชน์ในด้านวิชาการระดับการประยุกต์ทฤษฎีเพื่อนำไปใช้ เนื่องจากงานวิจัยมีจุดมุ่งหมายหลักในการสร้างและตรวจสอบคุณภาพของเครื่องมือวัดที่สร้างขึ้น จากแนวคิดหรือทฤษฎีที่มีนักวิชาการเสนอไว้แล้ว หรือมีการสังเคราะห์จากหลาย ๆ แนวคิดหรือทฤษฎี เพื่อใช้เป็นองค์ประกอบหรือตัวบ่งชี้ในการสร้างเครื่องมือ แต่ยังไม่ถึงขั้นได้องค์ความรู้ใหม่ หรือปรับปรุงแนวคิดและทฤษฎี จึงทำให้ผลการประเมินมีระดับคุณภาพค่อนข้างต่ำ ซึ่งผลการประเมินดังกล่าวสอดคล้องกับงานวิจัยของ อาภาพรรณ ประทุมไทย (2553) ที่ได้ประเมินคุณภาพงานวิจัยที่เกี่ยวกับการสร้างแบบวัดความถนัด ซึ่งมีผลการประเมินไปในทิศทางเดียวกันว่างานวิจัยที่ศึกษาในบริบทของการสร้างแบบวัดมีระดับคุณภาพของการนำไปใช้ประโยชน์ได้ในทางวิชาการได้ค่อนข้างต่ำ

ข้อเสนอแนะในการวิจัย

ผลการวิจัยครั้งนี้ทำให้ทราบว่าในภาพรวมของงานวิจัยการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา มีคุณภาพอยู่ในระดับดี แต่มีบางประเด็นที่ควรให้ความสำคัญมากขึ้นโดยเฉพาะประเด็นที่มีคุณภาพอยู่ในระดับค่อนข้างต่ำและระดับปานกลาง เพื่อเพิ่มคุณค่าของงานวิจัย ดังนี้

1. สำหรับนักวิจัย หรือนิสิต นักศึกษาผู้ซึ่งอยู่ระหว่างพัฒนาหัวข้อวิจัย ควรให้ความสำคัญกับการทบทวนเอกสารและงานวิจัยที่เกี่ยวข้อง ถึงแม้ในภาพรวมจะมีผลการประเมินคุณภาพอยู่ในระดับดี แต่เมื่อพิจารณาจากค่าเฉลี่ยแล้ว เป็นด้านที่มีค่าเฉลี่ยต่ำที่สุด ประกอบผลการประเมินยังพบว่า ความทันสมัยของเอกสารและงานวิจัยที่เกี่ยวข้องอยู่ในระดับค่อนข้างต่ำ รวมถึงเอกสารและงานวิจัยที่เกี่ยวข้องมีความสอดคล้องกับปัญหา/วัตถุประสงค์ของการวิจัย และสัดส่วนของเอกสารและงานวิจัยที่เกี่ยวข้องทั้งภาษาไทยและภาษาต่างประเทศ ยังอยู่ในระดับปานกลาง ซึ่งการทบทวนเอกสารและงานวิจัยที่เกี่ยวข้องเป็นจุดเริ่มต้นของการทำวิจัยที่ได้มุมมองที่แปลกใหม่ ลดความซ้ำซ้อนกับงานวิจัยที่ผ่านมา จึงเป็นสิ่งที่นักวิจัยให้ความสำคัญตั้งแต่การเริ่มต้นทำวิจัย

2. ในการดำเนินการวิจัยเพื่อพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ผู้วิจัยควรระบุเหตุผลและความจำเป็นในการทำวิจัยมีความสมเหตุสมผล การกำหนดจุดมุ่งหมายในการวัดสมรรถนะในศตวรรษที่ 21 มีความชัดเจน สามารถนำผลที่ได้จากการวิจัยใช้ประโยชน์ได้ในทางวิชาการ เนื่องจากในการพัฒนาผู้เรียนให้มีสมรรถนะในศตวรรษที่ 21 นั้น เครื่องมือวัดถือเป็นสิ่งสำคัญในการวัดเพื่อให้ได้สารสนเทศ หรือข้อมูลพื้นฐานสำหรับการวางแผนพัฒนาส่งเสริมให้ผู้เรียนมีสมรรถนะในศตวรรษที่ 21 หรือใช้สำหรับวัดผลสำเร็จที่เกิดขึ้นการทำกิจกรรมส่งเสริมสมรรถนะดังกล่าว ดังนั้น การกำหนดจุดมุ่งหมาย และการนำเครื่องมือวัดที่ได้จากการวิจัยไปใช้ในเชิงวิชาการจึงเป็นเรื่องสำคัญที่ไม่ควรละเลย

เอกสารอ้างอิง

- คณะกรรมการการสมาวิจัยแห่งชาติ สาขาการศึกษา. (2555). หลักสูตร “มโนทัศน์ที่คาดเคลื่อนในการวิจัยการศึกษา” เอกสารในโครงการ “การพัฒนาหลักสูตรการฝึกอบรมนักวิจัยการศึกษาเพื่อเพิ่มศักยภาพด้านการวิจัย”. กรุงเทพมหานคร: ห้างหุ้นส่วนจำกัด ไอออนพรีนติ้ง.
- ไชยยศ ไพวิทยศิริธรรม. (2544). การวิเคราะห์ห่อหุ้มคุณภาพของแบบสอบหลายตัวเลือกที่มีรูปแบบของแบบสอบ บริบทของแบบสอบ และวิธีการตรวจให้คะแนนต่างกัน. วิทยานิพนธ์ปริญญาโทมหาบัณฑิต สาขาวิชาการศึกษาวัดและประเมินผลการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.
- ณัฐธรรณ หลาวทอง. (2559). การสร้างเครื่องมือการวิจัยทางการศึกษา. กรุงเทพฯ: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- บุญใจ ศรีสถิตยัณรากร. (2555). การพัฒนาและตรวจสอบคุณภาพเครื่องมือวิจัย : คุณสมบัติการวัดเชิงจิตวิทยา (Development and validation of research instruments : Psychometric properties). กรุงเทพฯ: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย
- รัตนะ บัวสนธ์ และคณะ. (2557). การประเมินคุณภาพวิทยานิพนธ์ของนิสิตระดับดุขฎีบัณฑิต มหาวิทยาลัยนเรศวร. วารสารศึกษาศาสตร์ มหาวิทยาลัยนเรศวร 16(1): 120-126.
- วรรณิ์ แกมเกตุ. (2555). วิธีวิทยาการวิจัยทางพฤติกรรมศาสตร์. พิมพ์ครั้งที่ 3. กรุงเทพมหานคร: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- วัยญา ยิ้มยวน (2547). การวิเคราะห์ห่อหุ้มของปัจจัยที่มีความสัมพันธ์กับการคิดวิจารณ์ญาณ. วิทยานิพนธ์ปริญญาโทมหาบัณฑิต ภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.
- ศิวัชร ภูพันธ์. (2554). การศึกษาและทดสอบปัจจัยเชิงสาเหตุของการใช้ประโยชน์งานวิจัยของอาจารย์มหาวิทยาลัย: การวิจัยแบบผสมวิธี. วิทยานิพนธ์ปริญญาดุขฎีบัณฑิต สาขาวิชาวิธีวิทยาการวิจัย คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.
- สรัญญา จันทรชูสกุล. (2561).การวิเคราะห์ห่อหุ้มงานวิจัยด้านนวัตกรรมการพัฒนาความคิดสร้างสรรค์ของนักเรียนระดับประถมศึกษา. นครปฐม: คณะศึกษาศาสตร์ มหาวิทยาลัยศิลปากร
- สุวิมล เสวกสุริยวงศ์. (2553). การสังเคราะห์งานวิจัยเกี่ยวกับการสร้างแบบสอบวินิจฉัยทางคณิตศาสตร์. สาขาวิชาการศึกษาวัดและประเมินผลการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.
- สุวิมล ติรกานันท์. (2551). การสร้างเครื่องมือวัดตัวแปรในการวิจัยทางสังคมศาสตร์: แนวทางสู่การปฏิบัติ. พิมพ์ครั้งที่ 2. กรุงเทพมหานคร: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

- สุวิมล ติรกานันท์. (2557). ระเบียบวิธีการวิจัยทางสังคมศาสตร์: แนวทางสู่การปฏิบัติ. พิมพ์ครั้งที่ 12. กรุงเทพมหานคร: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- สำนักงานเลขาธิการสภาการศึกษา. (2552). รายงานการสังเคราะห์งานวิจัยเกี่ยวกับคุณภาพการศึกษาไทย:การวิเคราะห์ห่อภิมาณ (Meta-analysis). กรุงเทพมหานคร: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- สำนักงานเลขาธิการสภาการศึกษา. (2560). แผนการศึกษาแห่งชาติ พ.ศ. 2560-2579. กรุงเทพมหานคร: บริษัท พรักหวานกราฟฟิค จำกัด.
- อาภาพรรณ ประทุมไทย. (2553). การวิเคราะห์ห่อภิมาณคุณภาพของแบบวัดความถนัด. วิทยานิพนธ์ปริญญาโทมหาบัณฑิตสาขาวิชากการวัดและประเมินผลการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.
- Metiri Group. (2003). enGauge 21st century skills for 21st century learners.[Online]. Retrieved from :https://www.cwasd.k12.wi.us/highschl/newsfile1062_1.pdf
- Partnership for 21st Century Learning. (2019). Framework for 21st century learning definitions.[Online]. Retrieved from: http://static.battelleforkids.org/documents/p21/P21_Framework_DefinitionsBFK.pdf

โครงการวิจัยติดตามและประเมินผลกระทบของการประเมินผล การทดสอบทางการศึกษาระดับชาติ

(Research Project for Monitoring and Evaluation of the Impact of
the Evaluation of National Educational Testing)

จตุภูมิ เขตจัตุรัส
พระฮอนด้า เข้มมา
คม พวงยะ
ดาวรรณ ฤทธิการ
ณรงค์ศักดิ์ รอบคอบ
กิตติกันตพงศ์ ศรีบัวนา
ธนัท มุลเจริญพร

บทคัดย่อ

โครงการวิจัยนี้มีวัตถุประสงค์เพื่อ

- (1) ศึกษาผลกระทบของการประเมินผลการทดสอบทางการศึกษาระดับชาติใน 5 ประเภทการทดสอบ ได้แก่ การทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) ด้านอาชีวศึกษา (V-NET) ด้านอิสลามศึกษา (I-NET) ด้านพระพุทธศาสนา (B-NET) และ ด้านการศึกษานอกระบบโรงเรียน (N-NET)
- (2) ติดตามและประเมินผลกระทบของการประเมินผลการทดสอบทางการศึกษาระดับชาติ
- (3) จัดทำข้อเสนอแนะเชิงนโยบายในการปรับปรุงและพัฒนาการทดสอบทางการศึกษาระดับชาติ

ใช้ระเบียบวิธีวิจัยแบบผสมผสานวิธี (Mixed Methods Research) แบบพหุระยะ (Multiphase design) แบ่งเป็น 3 ระยะ ได้แก่

ระยะที่ 1 การศึกษาผลกระทบของการประเมินผลการทดสอบทางการศึกษาระดับชาติ ใช้วิธีการเชิงคุณภาพเป็นหลักด้วยการสัมภาษณ์ผู้ทรงคุณวุฒิ และผู้ที่ได้รับผลกระทบ และใช้วิธีการเชิงปริมาณเสริมด้วยการประเมินความเหมาะสมและเป็นไปได้ของประเด็นผลกระทบ

ระยะที่ 2 การติดตามและประเมินผลกระทบของการประเมินผลการทดสอบทางการศึกษาระดับชาติ ใช้วิธีการเชิงปริมาณร่วมกับวิธีการเชิงคุณภาพดำเนินการพร้อมกันด้วยการสำรวจโดยแบบสอบถาม การสัมภาษณ์ การสังเกตและการศึกษาเอกสาร

ระยะที่ 3 การจัดทำข้อเสนอแนะเชิงนโยบายในการปรับปรุงและพัฒนาการทดสอบทางการศึกษาระดับชาติ ใช้วิธีการเชิงคุณภาพเป็นหลักด้วยการจัดประชุมสนทนากลุ่ม (Focus Group Discussion) ร่วมกับการสัมภาษณ์เชิงลึก (In-depth Interview)

กลุ่มตัวอย่างของงานวิจัย ประกอบด้วย ผู้เรียนที่เคยเข้ารับการทดสอบทางการศึกษาระดับชาติ จำนวน 3,370 คน ผู้บริหารสถานศึกษา จำนวน 324 คน และครูผู้สอน จำนวน 399 คน ผู้ทรงคุณวุฒิ จำนวน 56 คน เครื่องมือวิจัย ประกอบด้วย แบบสัมภาษณ์เชิงลึก แบบสอบถาม และแบบบันทึกข้อมูล การวิเคราะห์ข้อมูลเชิงปริมาณใช้สถิติ ได้แก่ ความถี่ ร้อยละ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ค่าสัมประสิทธิ์สหสัมพันธ์ การวิเคราะห์จำแนกกลุ่ม และการวิเคราะห์สมการโครงสร้าง การวิเคราะห์ข้อมูลเชิงคุณภาพใช้การสร้างข้อสรุปแบบอุปนัย (Analytic induction) สรุปผลการวิจัย ดังนี้

1. การประเมินผลการทดสอบทางการศึกษาระดับชาติได้สร้างผลกระทบทั้งทางบวกและทางลบใน 3 ระดับ รวม 9 ประเด็นผลกระทบ ดังนี้ ผลกระทบต่อนโยบาย มี 4 ประเด็นผลกระทบ คือ

- (1) คุณภาพการศึกษา
- (2) การเข้าถึงโอกาสทางการศึกษา
- (3) ความเหลื่อมล้ำทางการศึกษา
- (4) มาตรฐานของการทดสอบทางการศึกษาระดับชาติ

ผลกระทบต่อหน่วยงาน มี 2 ประเด็นผลกระทบ คือ

- (1) ประสิทธิภาพของสถานศึกษา
- (2) กลไกการพัฒนาคุณภาพการศึกษา

ผลกระทบต่อบุคคล มี 3 ประเด็นผลกระทบ คือ

- (1) แรงจูงใจในการสอบทางการศึกษาระดับชาติของผู้เรียน
- (2) การจัดการเรียนรู้และการประเมินผลของครูผู้สอน
- (3) ภาวะผู้นำทางวิชาการของผู้บริหารสถานศึกษา

2. ผลการติดตามและประเมินผลกระทบของการประเมินผลการทดสอบทางการศึกษาระดับชาติ พบว่า การทดสอบทางการศึกษาระดับชาติส่งผลกระทบต่อนโยบายอยู่ในระดับปานกลางถึงมาก ส่งผลกระทบต่อหน่วยงานอยู่ในระดับมากถึงมากที่สุด และส่งผลกระทบต่อบุคคลอยู่ในระดับปานกลางถึงมากที่สุด

3. ข้อเสนอแนะเชิงนโยบายในการปรับปรุงและพัฒนาการทดสอบทางการศึกษาระดับชาติ ประกอบด้วยเป้าหมายและวิธีการที่ครอบคลุมประเด็นผลกระทบของการประเมินผลการทดสอบทางการศึกษาระดับชาติ

คำสำคัญ: การติดตามและประเมิน ผลกระทบ การทดสอบทางการศึกษาระดับชาติ

Abstract

The purpose of this research project was

(1) to study the impact of the evaluation of national educational tests in 5 test types, namely the Ordinary National Educational Test (O-NET), Vocational National Educational Test (V-NET), Islam National Educational Test (I-NET), Buddhism National Educational Test (B-NET) and Non-formal National Educational Test (N-NET).

(2) To monitoring and evaluation of the impact of the evaluation of national educational evaluation.

(3) establishing policy recommendations for improvement and development of national educational testing.

The research methodology employed the mixed methods research as the multiphase design is divided into 3 phases as follows.

Phase 1: Study of the impact of evaluation of national educational testing, mainly use qualitative methods by interviewing qualified and affected people, and using quantitative methods supplemented by assessing the propriety and feasibility of impact issues.

Phase 2: Monitoring and evaluation of the impact of national educational test, using quantitative methods together with qualitative methods, carried out simultaneously with surveys by questionnaires, interviews, observations, and document studies.

Phase 3: establish policy recommendations for improvement and development of national educational testing, mainly used qualitative methods by organizing focus group discussions together with in-depth interviews.

The research sample consisted of 3,370 students who have previously received national educational tests. There were 324 school administrators and 399 teachers, 56 experts. The research tools consisted of in-depth interview forms, questionnaires, and record forms. The quantitative data analysis used statistics such as frequency, percentage, mean, standard deviation, correlation coefficient, discriminant analysis, and structural equation modeling, qualitative data analysis used analytic induction. Research findings as follows,

1. The evaluation of national educational testing has created both are positive and negative impact in 3 levels, including 9 issues. The impact on the policy has 4 impacts:

- (1) educational quality
- (2) access to educational opportunities
- (3) Educational inequality
- (4) National educational testing standards.

The impact on the department 2 impacts:

- (1) school effectiveness and
- (2) mechanism for the development of quality of education.

The impact on individual has 3 impacts:

- (1) national test taking motivation of learners
- (2) learning management and assessment of teachers,
- (3) academic leadership of principals.

2. The results of monitoring and evaluation of the impact of national educational testing found that the national educational testing impact on policy had medium level to high level, impact on organization had high level to highest level, and impact on individual had medium level to highest level.

3. Policy recommendations for improvement and development of national educational testing consisted of goals and approaches that cover the impact of the evaluation of national educational tests.

Keywords: Monitoring and Evaluation, Impact, National Educational Test

บทนำ

สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) หรือ สทศ. ได้จัดตั้งขึ้นโดยมีวัตถุประสงค์ เพื่อบริหารจัดการและดำเนินการเกี่ยวกับการศึกษา วิจัย พัฒนา และให้บริการทางการประเมินผลทางการศึกษาและทดสอบทางการศึกษาอย่างต่อเนื่อง รวมทั้งเป็นศูนย์กลางความร่วมมือด้านการทดสอบทางการศึกษาในระดับชาติและระดับนานาชาติ โดยมีบทบาทหน้าที่สำคัญ ดังนี้

- (1) จัดการบริหารทดสอบทางการศึกษาระดับชาติให้ได้มาตรฐาน
- (2) พัฒนาข้อสอบให้สามารถวัดได้ทุกมาตรฐานการศึกษาอย่างมีมาตรฐาน
- (3) ดำเนินการให้มีการนำผลการทดสอบไปใช้ให้เกิดประโยชน์
- (4) วิจัยและพัฒนาให้มีระบบการทดสอบทางการศึกษาให้ครบทุกมาตรฐานด้านคุณภาพผู้เรียนในทุกระดับและทุกประเภทการศึกษา
- (5) ศึกษาวิจัยเรื่องรูปแบบข้อสอบและสนับสนุนการวิจัยและพัฒนาเกี่ยวกับโปรแกรมและเทคนิคการสร้างข้อสอบและมาตรฐานการวัดรูปแบบต่าง ๆ
- (6) ประชาสัมพันธ์ให้ข้อมูลและข้อความรู้ เพื่อเพิ่มภาพลักษณ์ที่ดีต่อสาธารณชน
- (7) ทบทวนกฎหมาย ข้อบังคับและมติต่าง ๆ ให้เหมาะสมกับสภาพที่เป็นจริง ทันสมัย และเอื้อต่อการปฏิบัติงาน
- (8) สนับสนุนและส่งเสริมให้สถานศึกษาและทุกระดับสถานศึกษาและต้นสังกัดมีโปรแกรมการทดสอบและประเมินผลผู้เรียนตามมาตรา 26 แห่งพระราชบัญญัติการศึกษาแห่งชาติ พ.ศ.2542
- (9) ให้บริการการทดสอบแก่บุคลากรและหน่วยงานต่าง ๆ
- (10) เป็นศูนย์กลางข้อมูลการทดสอบทางการศึกษา สนับสนุน และให้บริการผลการทดสอบแก่หน่วยงานต่าง ๆ ตลอดจนความร่วมมือด้านการทดสอบทางการศึกษาทั้งในประเทศและต่างประเทศ ได้อย่างมีมาตรฐานระดับสากล

การจัดการทดสอบทางการศึกษาระดับชาติเป็นการทดสอบความรู้และความคิดของผู้เรียนที่สำเร็จการศึกษาในแต่ละหลักสูตรประกอบด้วย 5 การทดสอบ คือ

(1) การทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) เป็นการทดสอบความรู้และความคิดของนักเรียนเมื่อสำเร็จการศึกษาตามช่วงชั้นที่ 2 คือ ระดับชั้นประถมศึกษาปีที่ 6 ช่วงชั้นที่ 3 คือ ระดับชั้นมัธยมศึกษาปีที่ 3 และช่วงชั้นที่ 4 คือ ระดับชั้นมัธยมศึกษาปีที่ 6 โดยที่ข้อสอบจะออกครอบคลุมตามมาตรฐานและตัวชี้วัดของหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551

(2) การทดสอบทางการศึกษาระดับชาติด้านอาชีวศึกษา (V-NET) มีเป้าหมายเพื่อทดสอบความรู้และความคิดของนักศึกษาระดับประกาศนียบัตรวิชาชีพ ชั้นปีที่ 3 (ปวช.3) นักศึกษาระดับประกาศนียบัตรวิชาชีพชั้นสูง (ปวส.2) ตามหลักสูตรประกาศนียบัตรวิชาชีพ พุทธศักราช 2556 และหลักสูตรประกาศนียบัตรวิชาชีพชั้นสูง พุทธศักราช 2546 ในรายวิชา

ความรู้ด้านสมรรถนะหลักและสมรรถนะทั่วไป ประกอบด้วย ทักษะภาษาและการสื่อสาร ทักษะ การคิดและการแก้ปัญหา ทักษะทางสังคมและการดำรงชีวิต ทักษะการจัดการงานอาชีพ

(3) การทดสอบทางการศึกษาระดับชาติด้านอิสลามศึกษา (I-NET) มีเป้าหมายเพื่อ ทดสอบความรู้และความคิดรวบยอดของนักเรียนนักเรียนประถมศึกษา มัธยมศึกษาตอนต้น และมัธยมศึกษาตอนปลาย ที่ศึกษาในโรงเรียนที่ใช้หลักสูตรอิสลามศึกษาตามหลักสูตรอิสลาม ศึกษา พุทธศักราช 2546 และหลักสูตรอิสลามศึกษาตามหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551 ครอบคลุม 8 กลุ่มสาระการเรียนรู้ 14 มาตรฐานการเรียนรู้รายวิชาที่ใช้ใน การจัดสอบ ประกอบด้วย 8 รายวิชา ดังนี้ อัลกุรอาน-ฮัตตีฟซีร์ อัลหะดีษ อัลอะกีดะฮ์ อัลฟิเกห์ ฮัตตารีคอัลอ์ลาค ภาษาอาหรับ และภาษามลายู ระดับของการทดสอบระดับชาติด้านอิสลาม ศึกษา (I-NET) แบ่งเป็น 3 ระดับ ได้แก่ ระดับอิสลามศึกษาตอนต้น (อับตีดาอียะห์) ระดับอิสลาม ศึกษาตอนกลาง (มุตะวัสซึฏะฮ์) ระดับอิสลามศึกษาตอนปลาย (ชานาวียะห์)

(4) การทดสอบทางการศึกษาระดับชาติด้านพระพุทธศาสนา (B-NET) มีเป้าหมาย เพื่อทดสอบความรู้และความคิดรวบยอดของนักเรียนระดับมัธยมศึกษาตอนต้นและระดับ มัธยมศึกษาตอนปลายตามหลักสูตรโรงเรียนพระปริยัติธรรม แผนกสามัญศึกษาตามหลักสูตร การศึกษาขั้นพื้นฐาน พุทธศักราช 2554 และหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551 ในวิชาภาษาบาลี ธรรม พุทธประวัติ และวินัย

(5) การทดสอบทางการศึกษาระดับชาติด้านการศึกษานอกระบบโรงเรียน (N-NET) เป็นการทดสอบความรู้และความคิดของนักเรียนในระดับประถมศึกษา ระดับมัธยมศึกษาตอนต้น และระดับมัธยมศึกษาตอนปลายของหลักสูตรการศึกษานอกระบบระดับการศึกษาขั้นพื้นฐาน พุทธศักราช 2551 สอบในรายวิชาสาระความรู้พื้นฐาน สาระการพัฒนาสังคม สาระทักษะ การเรียนรู้ สาระการประกอบอาชีพ และสาระทักษะการดำเนินชีวิต

ผลการทดสอบทางการศึกษาระดับชาติได้ส่งผลกระทบต่อบุคคล หน่วยงาน และ การดำเนินนโยบาย ให้เกิดการปรับเปลี่ยน ปรับปรุง และพัฒนาไปในทิศทางที่พึงประสงค์ (ญาณิกา ลุนราศรี และจันทร์ทรงกลด คชเสนี, 2557; อิทธิพัทธ์ สุวทันพรกุล, 2556 ; เสตอาหมาด ศรียาน และคณะ, 2558; สัมพันธ์ พันธุ์พุกษ์, 2561) ผู้เกี่ยวข้องทุกฝ่าย ต้องช่วยการสร้างความรับผิดชอบร่วมกันพัฒนาในทุกมิติเพื่อสร้างผลกระทบที่ดีให้เกิดขึ้น โดยเฉพาะอย่างยิ่งกระบวนการตรวจสอบ ติดตามและประเมินผล (Monitoring and Evaluation) เป็นเรื่องที่ทุกหน่วยงานต้องดำเนินการ เพื่อให้สอดคล้องกับพระราชกฤษฎีกา ว่าด้วยหลักเกณฑ์ การบริหารกิจการบ้านเมืองที่ดี พ.ศ. 2546 ที่ได้กล่าวถึงการประเมินผลการปฏิบัติราชการ ที่จะต้องมุ่งเน้นผลสัมฤทธิ์ เป็นการวัดผลว่าเป็นไปตามนโยบาย แผนงาน เป้าหมาย และ วัตถุประสงค์ของงานในภาพรวม และงานแต่ละด้าน รวมทั้งสามารถนำมาใช้กำกับตรวจสอบ การใช้ทรัพยากรในการปฏิบัติงานให้เป็นไปอย่างคุ้มค่า และมีประสิทธิภาพ การติดตาม และ ประเมินผลการดำเนินงานจึงเป็นกระบวนการที่ก่อให้เกิดการปรับปรุงการบริหารและ การดำเนินงานต่อไป ไม่ว่าจะเป็นการกำหนดวัตถุประสงค์หรือมาตรฐานในการดำเนินงานที่

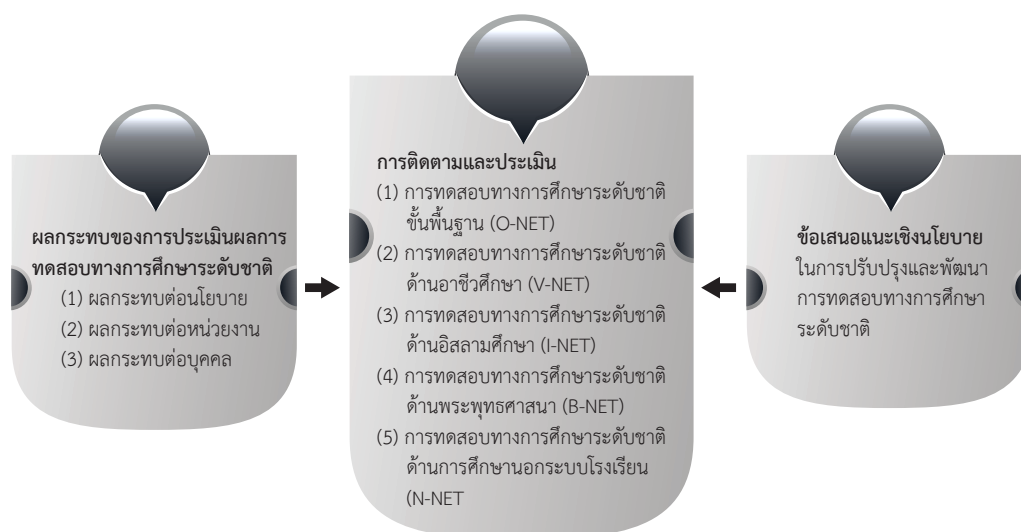
ชัดเจน การใช้ทรัพยากรให้เกิดประโยชน์ได้อย่างเต็มที่และคุ้มค่า การควบคุมคุณภาพ การตัดสินใจ ผลสัมฤทธิ์ของโครงการ การจัดการประเมินเพื่อให้ทราบว่า กิจกรรมนั้นบรรลุผลตามเป้าหมายที่กำหนดไว้เพียงใด มีปัญหาอุปสรรคอย่างไร จะได้เป็นข้อมูลในการปรับปรุงเพื่อดำเนินการต่อไป หรือยุติการดำเนินงานนั้น ระบบการติดตามและประเมินผลการดำเนินงานนั้นมีความสำคัญอย่างยิ่งต่อความสำเร็จของนโยบายและแผน กล่าวคือจะเป็นกลไกในการวัดผลสัมฤทธิ์ของการดำเนินงานโดยพิจารณาสถานะของการดำเนินงานว่าอยู่ในระดับใด ดำเนินการไปในทิศทางใด และบรรลุวัตถุประสงค์ตามเป้าหมายที่กำหนดไว้แล้วหรือไม่ ควรจะปรับปรุงแก้ไขอย่างไร

โครงการวิจัยการติดตามและประเมินผลกระทบของการประเมินผลการทดสอบทางการศึกษาระดับชาตินี้จะเป็นประโยชน์ในการนำเสนอแนะเชิงนโยบายไป พัฒนาการทดสอบทางการศึกษาระดับชาติให้มีประสิทธิภาพและเกิดประโยชน์สูงสุดต่อไป

วัตถุประสงค์ของงานวิจัย

- 1) เพื่อศึกษาผลกระทบของการประเมินผลการทดสอบทางการศึกษาระดับชาติ
- 2) เพื่อติดตามและประเมินผลกระทบของการประเมินผลการทดสอบทางการศึกษาระดับชาติ
- 3) เพื่อจัดทำข้อเสนอแนะเชิงนโยบายในการปรับปรุงและพัฒนาระบบทดสอบทางการศึกษาระดับชาติ

กรอบแนวคิดของโครงการวิจัย



แผนภาพที่ 1 กรอบแนวคิดของการวิจัย

วิธีดำเนินการวิจัย

การติดตามและประเมินผลกระทบของการประเมินผลการทดสอบทางการศึกษาแห่งชาติ ครั้งนี้เป็นการวิจัยเชิงประเมิน (Evaluative Research) ประกอบด้วย 7 ขั้นตอนที่สุดแต่แรกใน 3 ระยะของการดำเนินโครงการวิจัยที่ใช้ระเบียบวิธีวิจัยแบบผสมผสานวิธี (Mixed Methods Research) แบบพหุระยะ (Multiphase design) ทั้งวิธีการเชิงปริมาณและเชิงคุณภาพ โดยวิธีดำเนินการวิจัยนี้ได้มาจากการสังเคราะห์แนวคิดของนักวิชาการและหน่วยงานต่อไปนี้ Bamberger (2012), Creswell (2013), Kusek and Rist (2004), Rogers (2014), Stufflebeam and Coryn (2014), The World Bank. (2004) แบ่งเป็น 3 ระยะ ดังนี้

ระยะที่ 1 การศึกษาผลกระทบของการประเมินผลการทดสอบทางการศึกษาระดับชาติ ประกอบด้วย 2 ขั้นตอน ดังนี้

- (1) ศึกษาผลกระทบของการประเมินผลการทดสอบทางการศึกษาระดับชาติ
- (2) พัฒนารอบการติดตามและประเมินผลกระทบ

โดยการพัฒนาประเด็นผลกระทบ วิธีการและเครื่องมือเก็บรวบรวมข้อมูล และเกณฑ์การประเมิน วิธีดำเนินการวิจัย ใช้วิธีการเชิงคุณภาพเป็นหลักและเสริมด้วยวิธีการเชิงปริมาณ กลุ่มเป้าหมาย ประกอบด้วย ผู้ทรงคุณวุฒิ จำนวน 10 ท่าน และผู้ได้รับผลกระทบของการทดสอบทางการศึกษาระดับชาติ จำนวน 15 ท่าน เครื่องมือวิจัย ประกอบด้วย แบบสัมภาษณ์เชิงลึก แบบสอบถาม และแบบบันทึกข้อมูล การวิเคราะห์ข้อมูลเชิงปริมาณใช้สถิติได้แก่ ความถี่ ร้อยละ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ค่าสัมประสิทธิ์สหสัมพันธ์ การวิเคราะห์จำแนกกลุ่ม และการวิเคราะห์สมการโครงสร้าง การวิเคราะห์ข้อมูลเชิงคุณภาพใช้การสร้างข้อสรุปแบบอุปนัย (Analytic induction)

ระยะที่ 2 การติดตามและประเมินผลกระทบของการประเมินผลการทดสอบทางการศึกษาระดับชาติ ประกอบด้วย 3 ขั้นตอน ดังนี้

- (1) วางแผนการติดตามและประเมินผลกระทบของการประเมินผลการทดสอบทางการศึกษาระดับชาติ
- (2) เก็บรวบรวมข้อมูลทั้งเชิงปริมาณและเชิงคุณภาพจากผู้เกี่ยวข้องับประเด็นผลกระทบของการประเมินผลทดสอบทางการศึกษาระดับชาติ
- (3) วิเคราะห์ข้อมูลสารสนเทศการติดตามและประเมินผลงานเทียบกับเกณฑ์การประเมิน พร้อมจัดทำรายงานสรุปเบื้องต้น

วิธีดำเนินการวิจัย ใช้วิธีการเชิงปริมาณร่วมกับวิธีการเชิงคุณภาพดำเนินการพร้อมกัน ด้วยการเก็บรวบรวมข้อมูล ทั้งเชิงปริมาณและเชิงคุณภาพตามประเด็นผลกระทบ โดยใช้แบบสอบถาม การสัมภาษณ์ การสังเกต และการศึกษาเอกสาร จากกลุ่มตัวอย่าง ดังนี้ นักศึกษามหาวิทยาลัยชั้นปีที่ 1 ที่เคยเข้ารับการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (O-NET) ในปีการศึกษา 2561 จำนวน 2,643 คน ใช้วิธีการสุ่มแบบหลายขั้นตอน (Multi-stages random sampling) ผู้เรียนที่เคยเข้ารับการทดสอบทางการศึกษาระดับชาติด้านการอาชีวศึกษา

(V-NET) ด้านอิสลามศึกษา (I-NET) ด้านพระพุทธศาสนา (B-NET) และด้านการศึกษานอกระบบ โรงเรียน (N-NET) รวมจำนวน 727 คน ได้มาโดยการเลือกแบบเจาะจง ผู้ให้ข้อมูลหลัก ได้แก่ ผู้บริหาร จำนวน 324 คน และครูผู้สอน จำนวน 399 คน จากสถานศึกษาขั้นพื้นฐาน สถานศึกษา อาชีวศึกษา วิทยาลัยอิสลามศึกษา โรงเรียนเอกชนสอนศาสนาอิสลามควบคู่สามัญ โรงเรียน พระปริยัติธรรม และศูนย์การศึกษานอกระบบโรงเรียนและอัยยาศัย (กศน.) ได้มาโดยการเลือก แบบเจาะจง เครื่องมือวิจัย ได้แก่ แบบสอบถามผู้เรียนแบบสอบถามผู้บริหาร แบบให้ข้อมูล สัมภาษณ์ ที่ผ่านการตรวจสอบคุณภาพความตรงเชิงเนื้อหา (Content Validity) โดยผู้เชี่ยวชาญ 5 ท่าน และตรวจสอบคุณภาพความเที่ยงโดยหาค่าสัมประสิทธิ์แอลฟา (α) ของครอนบาค (Cronbach) พบว่า ทุกองค์ประกอบของแบบสอบถามผู้เรียน ครู และผู้บริหาร มีค่าสัมประสิทธิ์ ความเที่ยงอยู่ระหว่าง 0.74-0.98 อยู่ในเกณฑ์ที่ยอมรับได้ การวิเคราะห์ข้อมูลใช้สถิติเชิงบรรยาย ความถี่ ร้อยละ ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐาน

เกณฑ์การประเมินระดับผลกระทบ คะแนนเต็ม 5 มีดังนี้

4.51-5.00 หมายถึง มีผลกระทบจากการประเมินผลการทดสอบระดับชาติ อยู่ในระดับมากที่สุด

เกิดผลกระทบอย่างมากเกิดขึ้นในวงกว้าง ทำให้เกิดการเปลี่ยนแปลงอย่างยิ่ง

3.51-4.50 หมายถึง มีผลกระทบจากการประเมินผลการทดสอบระดับชาติ อยู่ในระดับมาก

เกิดผลกระทบค่อนข้างมาก ทำให้เกิดการเปลี่ยนแปลงค่อนข้างมาก

2.51-3.50 หมายถึง มีผลกระทบจากการประเมินผลการทดสอบระดับชาติ อยู่ในระดับปานกลาง

เกิดผลกระทบขึ้นบ้าง ทำให้เกิดการเปลี่ยนแปลงพอสมควร

1.51-2.50 หมายถึง มีผลกระทบจากการประเมินผลการทดสอบระดับชาติ อยู่ในระดับน้อย

เกิดผลกระทบขึ้นนาน ๆ ครั้ง ทำให้เกิดการเปลี่ยนแปลงเล็กน้อย

1.00-1.50 หมายถึง มีผลกระทบจากการประเมินผลการทดสอบระดับชาติ อยู่ในระดับน้อยที่สุด

เกิดผลกระทบขึ้นนาน ๆ ครั้ง แต่ไม่เกิดการเปลี่ยนแปลง

ระยะที่ 3 การจัดทำข้อเสนอแนะเชิงนโยบายในการปรับปรุงและพัฒนาการทดสอบ การศึกษาระดับชาติ ประกอบด้วย 2 ขั้นตอน ดังนี้

1) การประชุมสนทนากลุ่ม (Focus Group Discussion) เพื่อจัดทำข้อเสนอแนะ เชิงนโยบายในการพัฒนาการบริหารจัดการทดสอบทางการศึกษาระดับชาติ โดยใช้สารสนเทศ จากผลการติดตามและประเมินผลกระทบของการประเมินผลการทดสอบทางการศึกษาระดับชาติ

2) สรุปผลโครงการวิจัยเสนอต่อสถาบันทดสอบทางการศึกษาระดับชาติ องค์การ มหาชน (สทศ.) วิจัยดำเนินการวิจัย ใช้วิธีการเชิงคุณภาพเป็นหลักด้วยการจัดประชุมสนทนากลุ่ม (Focus Group Discussion) ร่วมกับการสัมภาษณ์เชิงลึก (In-depth Interview)

เพื่อจัดทำข้อเสนอแนะเชิงนโยบายในการปรับปรุงและพัฒนาการทดสอบทางการศึกษา ได้แก่ กลุ่มผู้เกี่ยวข้องกับการทดสอบทางการศึกษาระดับชาติด้านพื้นฐาน (O-NET) จำนวน 9 ท่าน กลุ่มผู้เกี่ยวข้องกับการทดสอบทางการศึกษาระดับชาติด้านอาชีวศึกษา (V-NET) จำนวน 12 ท่าน กลุ่มผู้เกี่ยวข้องกับการทดสอบทางการศึกษาระดับชาติด้านอิสลามศึกษา (I-NET) จำนวน 9 ท่าน กลุ่มผู้เกี่ยวข้องกับการทดสอบทางการศึกษาระดับชาติด้านพระพุทธศาสนา (B-NET) จำนวน 6 ท่าน และกลุ่มผู้เกี่ยวข้องกับการทดสอบทางการศึกษาระดับชาติด้านการศึกษานอกระบบ โรงเรียน (N-NET) จำนวน 15 ท่าน ใช้วิธีการคัดเลือกผู้เกี่ยวข้องโดยการเลือกแบบเจาะจง (purposive sampling) เครื่องมือวิจัยคือ ประเด็นการสนทนากลุ่ม วิเคราะห์ข้อมูล และสร้างข้อสรุปแบบอุปนัย (Analytic induction)

โดยสรุประยะของการดำเนินการวิจัย 3 ระยะ ได้ตั้งแผนภาพที่ 2 และกรอบการติดตามและประเมิน (Monitoring and Evaluation Framework) ดังตารางที่ 1

ระยะของการวิจัย	วิธีดำเนินการวิจัย	ขั้นตอนของการติดตามและประเมิน
ระยะที่ 1 การศึกษาผลกระทบของ การประเมินผลการทดสอบ ทางการศึกษาระดับชาติ (ระยะเวลา 3 เดือน)	วิธีการเชิงคุณภาพ เป็นหลักและเสริมด้วยวิธี การเชิงปริมาณ (Concurrent embedded design) (QUAL+quan)	1) ศึกษาผลกระทบของการประเมินผลการทดสอบทาง การศึกษาระดับชาติ 2) พัฒนารอบการติดตามและประเมินผลกระทบ โดย การพัฒนาประเด็นผลกระทบ วิธีการและเครื่องมือ เก็บรวบรวมข้อมูล และเกณฑ์การประเมิน
ระยะที่ 2 การติดตามและประเมินผล กระทบของการประเมินผล การทดสอบทางการศึกษา ระดับชาติ (ระยะเวลา 4 เดือน)	วิธีการเชิงปริมาณและวิธี การเชิงคุณภาพดำเนินการ พร้อมกัน (Concurrent triangulation design) (QUAN+QUAL)	3) วางแผนการติดตามและประเมินผลกระทบของการ ประเมินผลการทดสอบทางการศึกษาระดับชาติ 4) เก็บรวบรวมข้อมูลทั้งเชิงปริมาณและเชิงคุณภาพจากผู้ ที่เกี่ยวข้องกับผลกระทบของการประเมินผลทดสอบ ทางการศึกษาระดับชาติ 5) วิเคราะห์ข้อมูลสารสนเทศการติดตามและประเมิน ผลงานเทียบกับเกณฑ์การประเมิน พร้อมจัดทำรายงาน สรุปเบื้องต้น
ระยะที่ 3 การจัดทำข้อเสนอแนะเชิง นโยบายในการปรับปรุงและ พัฒนาการทดสอบทางการ ศึกษาระดับชาติ (ระยะเวลา 2 เดือน)	วิธีการเชิงคุณภาพเป็นหลัก (Qualitative Approach)	6) การประชุมสนทนากลุ่ม (Focus Group Discussion) เพื่อจัดทำข้อเสนอแนะเชิงนโยบายในการพัฒนาการ บริหารจัดการทดสอบทางการศึกษาระดับชาติ โดยใช้ สารสนเทศจากผลการติดตามและประเมินผลกระทบ ของการประเมินผลการทดสอบทางการศึกษาระดับชาติ 7) สรุปผลโครงการวิจัยเสนอต่อสถาบันทดสอบทางการ ศึกษาระดับชาติ องค์การมหาชน (สทศ.)

แผนภาพที่ 2 ระยะการดำเนินการวิจัยแบบผสมผสานวิธีและขั้นตอนการติดตามและประเมิน

ตารางที่ 1 กรอบการติดตามและประเมินผลกระทบของการประเมินผลการทดสอบทางการศึกษาระดับชาติ

ประเด็นผลกระทบ	ข้อมูลที่ต้องการ	ผู้ให้ข้อมูล	เครื่องมือที่ใช้เก็บรวบรวม	การวิเคราะห์ข้อมูล	เกณฑ์การประเมิน
1. คุณภาพการศึกษา	(1) ผลการประเมินคุณภาพภายนอก รอบสามของตัวบ่งชี้ในมาตรฐานด้านผลสัมฤทธิ์ของผู้เรียน (2) ผลการรับรองคุณภาพสถานศึกษา (3) ความคิดเห็นต่อผลการสอบที่ใช้บ่งชี้ถึงคุณภาพการศึกษา	(1) สำนักงานรับรองมาตรฐานและประเมินคุณภาพการศึกษา (องค์การมหาชน) หรือ สมศ. (2) ผู้บริหารสถานศึกษา	(1) แบบบันทึกข้อมูล (2) แบบสอบถามสำหรับผู้บริหาร	(1) การวิเคราะห์สถิติเชิงบรรยาย ได้แก่ ความถี่ร้อยละ ค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน (2) การวิเคราะห์สหสัมพันธ์ (3) การวิเคราะห์จำแนกกลุ่ม (Discriminant Analysis)	(1) เกณฑ์ประเมินผลกระทบ 4.51-5.00 มากที่สุด 3.51-4.50 มาก 2.5-3.50 ปานกลาง 1.51-2.50 น้อย 1.00-1.50 น้อยที่สุด (2) เกณฑ์ระดับความสัมพันธ์ 0.5 ขึ้นไป (3) เกณฑ์การวิเคราะห์จำแนกกลุ่มพบนัยสำคัญทางสถิติที่ระดับ 0.05
2. การเข้าถึงโอกาสทางการศึกษา	(1) การสำเร็จการศึกษา (2) การใช้คะแนนสอบทางการศึกษาระดับชาติในการศึกษาต่อระดับที่สูงขึ้น (3) การได้รับการเชิดชูเกียรติ (4) การได้รับทุนการศึกษา	(1) ผู้เรียนที่เคยเข้าสอบระดับชาติ (2) ครูผู้สอนที่เกี่ยวข้องกับการทดสอบ (3) ผู้บริหารสถานศึกษา (4) ศึกษานิเทศก์และผู้บริหารต้นสังกัด	(1)แบบสอบถามสำหรับผู้เรียน (2) แบบสัมภาษณ์	(1) การวิเคราะห์สถิติเชิงบรรยาย ได้แก่ ความถี่ร้อยละ ค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน (2) การวิเคราะห์สร้างข้อสรุปแบบอุปนัย (Analytic Induction)	ใช้ข้อมูลเพื่อการติดตาม ไม่ได้นำมาประเมิน จึงไม่ได้กำหนดเกณฑ์ในการตัดสิน
3. ความเหลื่อมล้ำทางการศึกษา	(1) จำนวนผู้เรียนที่มีสิทธิ์สอบแต่ละประเภทการสอบ 3 ปีย้อนหลัง (2) การเรียนกวดวิชาของผู้เรียน (3) การส่งเสริมและพัฒนาผู้เรียนในการสอบของสถานศึกษา	(1) สถาบันทดสอบทางการศึกษาแห่งชาติ(องค์การมหาชน) หรือ สทศ. (2) ผู้เรียนระดับมัธยมศึกษาตอนปลาย (3) ผู้บริหารสถานศึกษา	(1) แบบบันทึกข้อมูล (2) แบบสอบถามสำหรับผู้บริหาร (3) แบบสัมภาษณ์	(1) การวิเคราะห์สถิติเชิงบรรยาย ได้แก่ ความถี่ร้อยละ	ใช้ข้อมูลเพื่อการติดตาม ไม่ได้นำมาประเมิน จึงไม่ได้กำหนดเกณฑ์ในการตัดสิน

ประเด็นผลกระทบ	ข้อมูลที่ต้องการ	ผู้ให้ข้อมูล	เครื่องมือที่ใช้เก็บรวบรวม	การวิเคราะห์ข้อมูล	เกณฑ์การประเมิน
4 มาตรฐานของการทดสอบทางการศึกษาระดับชาติ	(1) ความคิดเห็นต่อการสร้างความเชื่อมั่น และ มีมาตรฐานของการทดสอบ (2) ลักษณะข้อสอบ (3) การบริหารจัดการสอบ (4) เป้าหมายการสอบ	(1) ผู้เรียนที่เคยเข้าสอบระดับชาติ (2) ครูผู้สอนที่เกี่ยวข้องกับการทดสอบ (3) ผู้บริหารสถานศึกษา (4)ศึกษานิเทศก์และผู้บริหารต้นสังกัด	(1) แบบสอบถามสำหรับผู้บริหาร (2) แบบสัมภาษณ์	(1) การวิเคราะห์สถิติเชิงบรรยาย ได้แก่ ความถี่ร้อยละ ค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน (2) การวิเคราะห์สร้างข้อสรุปแบบอุปนัย	เกณฑ์ประเมินผลกระทบ 4.51-5.00 มากที่สุด 3.51-4.50 มาก 2.51-3.50 ปานกลาง 1.51-2.50 น้อย 1.00-1.50 น้อยที่สุด
5. ประสิทธิภาพของสถานศึกษา	(1) ผลการประเมินคุณภาพภายนอก (2) การกำหนดเป้าหมายและผลลัพธ์ (3) การพัฒนาการจัดการเรียนรู้ในชั้นเรียน (4) ความคิดเห็นต่อการพัฒนาคุณภาพโดยใช้ผลการทดสอบระดับชาติ	(1) ครูผู้สอนที่เกี่ยวข้องกับการทดสอบ (2) ผู้บริหารสถานศึกษา (3) ศึกษานิเทศก์/นักวิชาการ (4) ผู้บริหารต้นสังกัด	(1) แบบสอบถามสำหรับผู้บริหาร (2) แบบสัมภาษณ์	(1) การวิเคราะห์สถิติเชิงบรรยาย ได้แก่ ความถี่ร้อยละ ค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน (2) การวิเคราะห์สร้างข้อสรุปแบบอุปนัย (Analytic Induction)	เกณฑ์ประเมินผลกระทบ 4.51-5.00 มากที่สุด 3.51-4.50 มาก 2.51-3.50 ปานกลาง 1.51-2.50 น้อย 1.00-1.50 น้อยที่สุด
6. กลไกการพัฒนาคุณภาพการศึกษา	(1) ผลการประเมินคุณภาพภายนอก (2) ความคิดเห็นต่อการนำผลการทดสอบระดับชาติมาใช้ในการพัฒนาคุณภาพการศึกษา (3) ความคิดเห็นต่อผลการทดสอบที่มีต่อผู้เรียน ครูผู้สอน ผู้บริหาร และผู้ปกครอง	(1) ครูผู้สอนที่เกี่ยวข้องกับการทดสอบ (2) ผู้บริหารสถานศึกษา (3) ศึกษานิเทศก์/นักวิชาการ (4) ผู้บริหารต้นสังกัด	(1) แบบสอบถามสำหรับผู้บริหาร (2) แบบสัมภาษณ์ (3) แบบบันทึกข้อมูลหัตถ์นิยม	(1) การวิเคราะห์สถิติเชิงบรรยาย ได้แก่ ความถี่ร้อยละ ค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน (2) การวิเคราะห์สร้างข้อสรุปแบบอุปนัย (Analytic Induction)	เกณฑ์ประเมินผลกระทบ 4.51-5.00 มากที่สุด 3.51-4.50 มาก 2.51-3.50 ปานกลาง 1.51-2.50 น้อย 1.00-1.50 น้อยที่สุด

ประเด็นผลกระทบ	ข้อมูลที่ต้องการ	ผู้ให้ข้อมูล	เครื่องมือที่ใช้เก็บรวบรวม	การวิเคราะห์ข้อมูล	เกณฑ์การประเมิน
7. แรงจูงใจในการสอบทางการศึกษาระดับชาติของผู้เรียน	ความคิดเห็นของผู้เรียนต่อไปนี้ (1) การเห็นคุณค่าของการสอบ (1.1) ความสำคัญของการสอบ (1.2) ประโยชน์ของการสอบ (1.3) ความสำเร็จของการสอบ (2) การพัฒนาตนเองในการสอบ (2.1) การเตรียมตัวสอบ (2.2) การพัฒนาการเรียนรู้ (2.3) การฝึกฝนทำข้อสอบ (3) แรงจูงใจในการสอบ (3.1) ความคาดหวังในการสอบ (3.2) ความวิตกกังวล	(1) นักศึกษา มหาวิทยาลัยชั้นปีที่ 1 ที่เคยเข้าสอบ O-NET (2) ผู้เรียนที่เคยเข้าสอบ V-NET, I-NET, B-NET และ N-NET	(1) แบบสอบถามผู้เรียน (2) แบบสัมภาษณ์	(1) การวิเคราะห์สถิติเชิงบรรยาย ได้แก่ ความถี่ ร้อยละ ค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน (2) การวิเคราะห์สร้างข้อสรุปแบบอุปนัย (Analytic Induction)	เกณฑ์ประเมินผลกระทบ 4.51-5.00 มากที่สุด 3.51-4.50 มาก 2.51-3.50 ปานกลาง 1.51-2.50 น้อย 1.00-1.50 น้อยที่สุด
8. การจัดการเรียนรู้และการประเมินผลของครูผู้สอน	(1) ผลการประเมินคุณภาพภายนอก (2) ความคิดเห็นต่อการจัดกิจกรรมการเรียนรู้ของครูที่สอดคล้องกับแนวทางในการทำข้อสอบ (3) ความคิดเห็นต่อการใช้เอกสาร/สื่อ/กิจกรรมของครูที่ช่วยส่งเสริมพัฒนาให้ผู้เรียนมีคะแนนสอบสูงขึ้น	(1) ครูผู้สอนที่เกี่ยวข้องกับการทดสอบ (2) ผู้บริหารสถานศึกษา (3) ศึกษาพิเศษ/นักวิชาการ (4) ผู้บริหารต้นสังกัด	(1) แบบสอบถามสำหรับผู้บริหาร (2) แบบสัมภาษณ์	(1) การวิเคราะห์สถิติเชิงบรรยาย ได้แก่ ความถี่ ร้อยละ ค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน (2) การวิเคราะห์สร้างข้อสรุปแบบอุปนัย (Analytic Induction)	เกณฑ์ประเมินผลกระทบ 4.51-5.00 มากที่สุด 3.51-4.50 มาก 2.51-3.50 ปานกลาง 1.51-2.50 น้อย 1.00-1.50 น้อยที่สุด

ประเด็นผลกระทบ	ข้อมูลที่ต้องการ	ผู้ให้ข้อมูล	เครื่องมือที่ใช้เก็บรวบรวม	การวิเคราะห์ข้อมูล	เกณฑ์การประเมิน
9. ภาวะผู้นำทางวิชาการของผู้บริหารสถานศึกษา	(1) ผลการประเมินคุณภาพภายนอก (2) ความคิดเห็นที่มีต่อคุณลักษณะของผู้บริหารในประเด็นต่อไปนี้ (2.1) การส่งเสริมสนับสนุนให้มีบรรยากาศวิชาการ (2.2) การอำนวยความสะดวกให้กับผู้เรียนในการเข้าสอบ (2.3) การส่งเสริมสนับสนุนครูผู้สอนให้พัฒนาผู้เรียน (2.4) การนิเทศติดตามการจัดกิจกรรมการเรียนรู้ของครูผู้สอน (2.5) การประชุมกำกับติดตามการดำเนินโครงการพัฒนาผู้เรียน	(1) ครูผู้สอนที่เกี่ยวข้องกับการทดสอบ (2) ผู้บริหารสถานศึกษา (3) ศึกษาพิเศษ/นักวิชาการ (4) ผู้บริหารต้นสังกัด	(1) แบบสอบถามสำหรับผู้บริหาร (2) แบบสัมภาษณ์	(1) การวิเคราะห์สถิติเชิงบรรยายได้แก่ความถี่ร้อยละ ค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน (2) การวิเคราะห์สร้างข้อสรุปแบบอุปนัย (Analytic Induction)	เกณฑ์ประเมินผลกระทบ 4.51-5.00 มากที่สุด 3.51-4.50 มาก 2.51-3.50 ปานกลาง 1.51-2.50 น้อย 1.00-1.50 น้อยที่สุด

ผลการวิจัย

1. ผลกระทบของการประเมินผลการทดสอบทางการศึกษาระดับชาติด

ผลกระทบของการประเมินผลการทดสอบทางการศึกษาระดับชาติดสำหรับ 5 ประเภทการทดสอบ ได้แก่ การทดสอบทางการศึกษาระดับชาติดขั้นพื้นฐาน (O-NET) การทดสอบทางการศึกษาระดับชาติดด้านอาชีวศึกษา (V-NET) การทดสอบทางการศึกษาระดับชาติดด้านอิสลามศึกษา (I-NET) การทดสอบทางการศึกษาระดับชาติดด้านพระพุทธศาสนา (B-NET) และการทดสอบทางการศึกษาระดับชาติดการศึกษานอกระบบโรงเรียน (N-NET) ประกอบด้วยผลกระทบ 3 กลุ่ม ได้แก่ ผลกระทบต่อนโยบาย ผลกระทบต่อหน่วยงาน และผลกระทบต่อบุคคล รวมทั้งหมด 9 ประเด็นผลกระทบ ผลการตรวจสอบคุณภาพด้านความเหมาะสมและความเป็นไปได้ของประเด็นผลกระทบเพื่อนำไปใช้ในการติดตามและประเมินผลอยู่ในระดับมากถึงมากที่สุด

ผลกระทบต่อนโยบาย เป็นผลกระทบของผลการทดสอบทางการศึกษาระดับชาติดที่มีต่อการดำเนินงานในระดับนโยบายที่เกี่ยวข้องกับคุณภาพและมาตรฐานของการจัดการศึกษามี 4 ประเด็นผลกระทบ ดังนี้

ประเด็นผลกระทบที่ 1 คุณภาพการศึกษา พบว่า มีการใช้ผลการทดสอบทางการศึกษาระดับชาติดไปประเมินคุณภาพภายนอกของสถานศึกษาในการทดสอบ O-NET V-NET และ B-NET ดังปรากฏผลดังตารางที่ 2

ตารางที่ 2 ผลการทดสอบทางสถิติระหว่างคะแนนตัวบ่งชี้ผลสัมฤทธิ์ของผู้เรียนและคะแนนประเมินคุณภาพภายนอกของสถานศึกษาที่เข้ารับการทดสอบทางการศึกษาระดับชาติด

การทดสอบ	กลุ่ม	จำนวน	ตัวบ่งชี้ ผลสัมฤทธิ์ ^a		คะแนนประเมิน คุณภาพ ภายนอก (100 คะแนน)		r_{xy}	p	Wilk's lambda	p	t	p
			M	SD	M	SD						
O-NET*	รับรอง	1,133	9.90	1.82	84.53	3.22	0.79	0.00	0.42	0.00	55.98	0.00
	ไม่ รับรอง	1,230	6.38	1.13	79.01	3.35	0.40	0.00				
	รวม	2,363	8.07	2.31	81.66	4.57	0.79	0.00				
V-NET**	รับรอง	623	3.49	0.73	83.25	5.68	0.29	0.00	0.41	0.00	12.14	0.00
	ไม่ รับรอง	162	2.03	1.47	61.23	6.59	0.33	0.00				
	รวม	785	3.19	1.10	78.76	7.12	0.56	0.00				
B-NET	รับรอง	155	10.34	1.67	83.16	2.31	0.72	0.00	0.53	0.00	23.17	0.00
	ไม่ รับรอง	250	5.95	1.95	74.95	6.04	0.02	0.00				
	รวม	405	7.63	2.82	78.09	6.36	0.54	0.00				

หมายเหตุ * การทดสอบ O-NET เป็นผลคะแนนของโรงเรียนมัธยมศึกษา

** คะแนนเต็มตัวบ่งชี้ V-NET เต็ม 5 คะแนน

M = Mean คือ ค่าเฉลี่ย SD = Standard Deviation คือ ส่วนเบี่ยงเบนมาตรฐาน
 r_{xy} = ค่าสัมประสิทธิ์สหสัมพันธ์ของเพียร์สัน Wilk's lampda คือ ค่าสถิติทดสอบที่บ่งชี้ถึง
การจำแนกกลุ่ม t = t-test คือ ค่าสถิติทดสอบความแตกต่างของคะแนนเฉลี่ย

p = p-value คือ ความน่าจะเป็นหลังผลการทดสอบใช้เปรียบเทียบกับระดับ
นัยสำคัญ (α) เพื่อตัดสินใจยอมรับหรือปฏิเสธ H_0

ผลกระทบทางบวก พบว่า ผลการทดสอบทางการศึกษาระดับชาติ มีความสัมพันธ์
ทางบวกกับผลการประเมินคุณภาพภายนอกของสถานศึกษาขั้นพื้นฐาน ($r_{xy} = 0.79$) วิทยาลัย
อาชีวศึกษา ($r_{xy} = 0.56$) และโรงเรียนพระปริยัติธรรมแผนกสามัญศึกษา ($r_{xy} = 0.54$) อย่างมี
นัยสำคัญทางสถิติที่ระดับ 0.01 ซึ่งเป็นหลักฐานหนึ่งในการสนับสนุนว่าคุณภาพการศึกษาของ
ประเทศไทยขึ้นอยู่กับผลการทดสอบทางการศึกษาระดับชาติ

ผลกระทบทางลบ พบว่า จากหลักฐานบ่งชี้ว่าผลการทดสอบทางการศึกษาระดับชาติ
เป็นตัวบ่งชี้ที่สามารถจำแนกกลุ่มของสถานศึกษาที่ได้รับการรับรองและไม่รับรองคุณภาพได้
อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01 (Wilk's lampda ของ O-NET V-NET B-NET เท่ากับ
0.42 0.51 0.53 ตามลำดับ ซึ่งหากนำผลการทดสอบทางการศึกษาระดับชาติไปใช้เป็น
องค์ประกอบในการประเมินคุณภาพการศึกษาควรคำนึงถึงผลกระทบที่อาจเกิดขึ้นตามมา
เนื่องจากมีหลากหลายปัจจัยที่ส่งผลต่อคุณภาพของคะแนนสอบทางการศึกษาระดับชาติ
ทั้งคุณภาพข้อสอบ คุณภาพของผู้เรียนในแต่ละปีการศึกษา และบริบทของสถานศึกษาที่แตกต่างกัน

ประเด็นผลกระทบที่ 2 การเข้าถึงโอกาสทางการศึกษา ผลกระทบปรากฏดังนี้

ผลกระทบทางบวก พบว่า ผู้เรียนสามารถเข้าถึงโอกาสทางการศึกษาจากการใช้
ผลการทดสอบทางการศึกษาระดับชาติในเรื่องต่าง ๆ เช่น ยื่นประกอบการสมัครเข้าศึกษาต่อ
ในระดับที่สูงขึ้น การสำเร็จการศึกษา ยื่นประกอบการสมัครเข้าทำงาน ยื่นประกอบการขอทุน
การศึกษา บอกระดับความสามารถของตนเองเพื่อแก้ไขและปรับปรุงการเรียน ผลกระทบทางลบ
พบว่า ในการทดสอบ V-NET และ I-NET ยังไม่ได้นำไปใช้กำหนดเป็นเงื่อนไขของการสำเร็จ
การศึกษา และไม่ได้นำคะแนนไปใช้เป็นองค์ประกอบในการศึกษาต่อระดับที่สูงขึ้นอย่างชัดเจน
ผลการทดสอบทางการศึกษาระดับชาติจึงยังไม่ได้นำมาใช้ให้เกิดประโยชน์สูงสุดในด้านการเข้าถึง
โอกาสทางการศึกษา อีกทั้งยังส่งผลกระทบให้ผู้เรียนขาดแรงจูงใจและไม่เห็นความสำคัญของ
การสอบ ดังตัวอย่างความคิดเห็นต่อไปนี้

*“ตอนเรียนปวช.ปี 3 ได้สอบ V-NET แต่ไม่ได้้นำคะแนนสอบมาใช้ในการ
การสมัครเรียนต่อ...ซึ่งก็เข้าใจว่าเป็นการสอบเพื่อประเมินคุณภาพวิทยาลัย
เท่านั้น..”* (สัมภาษณ์นักศึกษา, ก.ค. 62, V-NET)

*“จากมหาวิทยาลัยที่มีการเปิดสอนหลักสูตรอิสลามศึกษา ไม่ได้นำผล
I-NET มาใช้ยื่นในการสมัครสอบเข้าเรียนต่อ โดยการยื่นมาสมัครสอบหรือเรียนต่อ
ได้ใช้วุฒิของทางโรงเรียนเท่านั้น มันเลยเป็นจุดด้อยของนักศึกษาที่เข้ามาศึกษาต่อ
เพราะไม่มีพื้นฐานที่จะสามารถต่อยอดในระดับช่วงนั้นได้...”* (สัมภาษณ์อาจารย์
วิทยาลัยอิสลามศึกษา, ต.ค. 62, I-NET)

ประเด็นผลกระทบที่ 3 ความเหลื่อมล้ำทางการศึกษา ผลกระทบปรากฏดังนี้

ผลกระทบทางบวก พบว่า ผู้เรียนทุกคนได้รับความเสมอภาคและได้รับโอกาสอย่างทั่วถึงในการได้รับสิทธิเข้ารับการทดสอบมาตรฐานระดับชาติโดยสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) หรือ สทศ.อย่างเท่าเทียมกันตามระดับชั้นของแต่ละบุคคล เช่น การทดสอบด้วยแบบสอบมาตรฐาน การสอบผ่านระบบการจัดสอบอิเล็กทรอนิกส์ (E-testing) เป็นต้น รวมถึงการพัฒนาศักยภาพให้กับครูผู้สอนด้านการทดสอบ ผลกระทบทางลบ พบว่า ปรากฏการณ์ติวเพื่อสอบทางการศึกษาระดับชาติได้สร้างผลกระทบให้กับสถานศึกษาทุกประเภทต้องมีการจัดเตรียมแผนงาน จัดสรรงบประมาณ สถานที่และทรัพยากรที่เกี่ยวข้องในการจัดติวในแต่ละปีได้มีการใช้บุคลากรทั้งภายในและภายนอกสถานศึกษามาติวเนื้อหาแนวทางการทำข้อสอบให้กับนักเรียน ผู้เรียนต้องติวข้อสอบเรียนกวดวิชาเพิ่มเติม และครูผู้สอนต้องมีการหาพื้นที่ความรับผิดชอบในการติวข้อสอบ เป็นความเหลื่อมล้ำในการใช้งบประมาณ ทรัพยากรและต้นทุนในแต่ละสถานศึกษาที่มีความพร้อมแตกต่างกัน ดังตัวอย่างความคิดเห็นต่อไปนี้

“โรงเรียนต้องจัดโครงการรับรอง เพราะเป็นนโยบายของเขตพื้นที่ต้องดำเนินการยกระดับผลสัมฤทธิ์ของผู้เรียน โรงเรียนต้องมาวางแผนจัดการงบประมาณ สถานที่ ติวเตอร์ เอกสารต่าง ๆ ให้พร้อม เพราะถ้าคะแนนโรงเรียนได้ต่ำกว่ามีผลกระทบต่อโรงเรียน” (สัมภาษณ์ผู้บริหาร, ก.ค. 62, O-NET)

“สถานศึกษา ครูผู้สอน เสียเวลาในการสอนเพิ่มเติม เพื่อให้ผู้เรียนเข้าใจในเนื้อหาตามหลักสูตรเพิ่มมากขึ้น เพราะรายวิชาส่วนใหญ่อยู่ในภาคเรียนที่ 1-3 แต่ในการสอบเป็นการทดสอบในภาคเรียนที่ 6” (สัมภาษณ์ผู้บริหารอาชีวะ, ต.ค. 62, V-NET)

“การติวให้กับนักเรียน สารเนื้อหา บางอย่างไม่ได้ให้กับนักเรียน เลยใช้ช่วงเวลาติวนี้ให้กับนักเรียนเพื่อเกร็งเตรียมข้อสอบ แต่เมื่อเวลาเราเรียนจริง ๆ เราไม่ได้เอาเนื้อหานี้มาเรียน มีความต่างในหลาย ๆ อย่างในเรื่องสามัญสามารถทำได้ แต่ในเรื่องของศาสนาไม่สามารถทำได้” (สัมภาษณ์ผู้บริหาร, ต.ค. 62, I-NET)

“พระอาจารย์ท่านได้มาทบทวนความรู้เพื่อเตรียมการสอบ B-NET ส่วนวิชาสามัญที่สอบ O-NET ก็มีโยมอาจารย์มาสอนให้สำหรับสามเณร ม.6 บางรูปที่ต้องการเรียนต่อระดับมหาวิทยาลัยต้องมีการเตรียมตัวสอบ 9 วิชาสามัญเพิ่มอีกก็นับว่าหนักพอสมควร” (สัมภาษณ์สามเณร, ต.ค. 62, B-NET)

“....และก็ก่อนสอบจะมีการจัดติวโดยนำผู้รู้หรือครูในระบบหรือผู้มีความชำนาญในด้านวิชานั้นมาจัดติวให้..” (สนทนากลุ่ม, ต.ค. 62, N-NET)

ประเด็นผลกระทบที่ 4 มาตรฐานของการทดสอบทางการศึกษาระดับชาติผลกระทบ
ปรากฏดังนี้

ผลกระทบทางบวก พบว่า การทดสอบทางการศึกษาระดับชาติดำเนินการยอมรับความมีมาตรฐาน และมีความเชื่อมั่นในการนำผลการสอบไปใช้ต่อไป

- (1) ทดสอบความรู้และความคิดรวบยอดของผู้เรียนที่จะสำเร็จการศึกษา
- (2) นำผลการทดสอบไปใช้ประกอบการสำเร็จการศึกษา
- (3) นำผลการทดสอบไปใช้ในการพัฒนาคุณภาพการเรียนการสอนของสถานศึกษา
- (4) นำผลการทดสอบใช้ในการประเมินผลการเรียนรู้ของผู้เรียนระดับชาติ
- (5) นำผลไปใช้ในการประกันคุณภาพการศึกษา

ผลกระทบทางลบ พบว่า

(1) ด้านข้อสอบ พบว่าข้อสอบยังมีความยากเกินไปและไม่เหมาะสมสำหรับวัดความรู้และความคิดรวบยอดสำหรับผู้เรียนในบางประเภทของการทดสอบ

(2) ด้านรูปแบบการทดสอบ พบว่า ยังไม่เหมาะสมกับธรรมชาติของการจัดการศึกษาด้านอาชีวศึกษาที่มุ่งเน้นการปฏิบัติมากกว่าทฤษฎีส่งผลกระทบให้คะแนนสอบออกมาในระดับที่ยังไม่น่าพอใจ และไม่ได้สะท้อนศักยภาพที่แท้จริงและตรงตามความถนัดของแต่ละวิชาชีพ ดังตัวอย่างความคิดเห็นต่อไปนี้

“...อย่างที่เราทราบกันดีว่านักศึกษาอาชีวศึกษานัดจับค้อนมากกว่าจับปากกา คือพวกเขาจะนัดกับการใช้อุปกรณ์ที่เป็นทักษะ แต่พอมาสอบวัดวิชาการที่เขาเองก็เลยมาจากโรงเรียนมัธยมแล้ว กลับต้องมาเจอก็ทำให้คะแนนที่สอบได้ไม่ดี” (สนทนากลุ่มอาชีวศึกษา, ต.ค. 62, V-NET)

(3) ด้านการบริหารจัดการสอบ พบว่า เป็นภาระหน้าที่เพิ่มขึ้นของครูผู้สอนและผู้บริหารในทุกประเภทการทดสอบในการบริหารจัดการที่เกี่ยวข้องกับการสอบ ได้แก่ การเตรียมตัวสอบ การจัดเตรียมสนามสอบ การรับส่งนักเรียนไปยังสนามสอบ ช่วงเวลาการจัดสอบ เป็นต้น

(4) เป้าหมายของการสอบ ยังพบว่าไม่ได้มีกำหนดชัดเจนถึงคะแนนรายบุคคลของผู้เรียนจะนำไปใช้ประโยชน์ โดยพบในการทดสอบด้านอาชีวศึกษา ด้านอิสลามศึกษา ด้านพระพุทธศาสนา และด้านการศึกษานอกระบบส่งผลกระทบให้ผู้เรียนไม่ให้ความสำคัญกับการสอบและขาดแรงจูงใจในการสอบ ดังตัวอย่างความคิดเห็นต่อไปนี้

“เป้าหมายของการสอบ V-NET ยังมีทิศทางที่ไม่ตอบสนองความต้องการของอาชีวศึกษา นักศึกษายังไม่สามารถนำผลสอบไปใช้ในการสำเร็จการศึกษาและการรับรองมาตรฐานวิชาชีพได้ ดังนั้นการทดสอบควรบูรณาการให้ครอบคลุมไปกับการสำเร็จการศึกษา การรับรองมาตรฐานฝีมือแรงงาน เพื่อให้เกิดความคุ้มค่ากับการจัดสอบที่สุด” (สนทนากลุ่มอาชีวศึกษา, ต.ค. 62, V-NET)

“อยากให้ผลคะแนนส่วน I-NET มีบทบาทที่กระตุ้นให้นักเรียนตั้งใจเรียน เช่น ให้ผลคะแนนมีบทบาทในส่วนนำลงไปรวมกับผลสัมฤทธิ์ของนักเรียน หรือเป็นแรงจูงใจให้เด็กมีความตั้งใจที่จะนำผลการเรียนนี้ไปใช้ได้ไม่ว่าในต่างประเทศ ยื่นศึกษาต่อในต่างประเทศได้” (สัมภาษณ์ครูผู้สอน, ต.ค. 62, I-NET)

“ควรมีการนำคะแนน B-NET ไปพิจารณาเป็นองค์ประกอบหนึ่งในการคัดเลือกบุคคลเข้าเรียนต่อในระดับมหาวิทยาลัย เช่น มหาวิทยาลัยมหจุฬาลงกรณ์ราชวิทยาลัย (มจร.) มหาวิทยาลัยมหาจุฬาราชวิทยาลัย (มจร.) เป็นต้น” (สนทนากลุ่ม, ต.ค. 62, B-NET)

ผลกระทบต่อหน่วยงาน เป็นผลกระทบของผลการทดสอบทางการศึกษาระดับชาตินี้ที่มีต่อการดำเนินงานของสถานศึกษาตามนโยบายและแนวปฏิบัติของหน่วยงานต้นสังกัด มี 2 ประเด็นผลกระทบ ดังนี้

ประเด็นผลกระทบที่ 5 ประสิทธิภาพของสถานศึกษา ผลกระทบปรากฏดังนี้

ผลกระทบทางบวก พบว่า

(1) ผลการทดสอบทางการศึกษาระดับชาตินี้ส่งผลกระทบทางบวกต่อประสิทธิภาพของสถานศึกษาที่วัดจากเกณฑ์ความสำเร็จของผลการประเมินคุณภาพภายนอกของสถานศึกษา

(2) ผลการทดสอบทางการศึกษาระดับชาตินี้ใช้เป็นข้อมูลสารสนเทศเพื่อค้นหาสาระการเรียนรู้และจุดที่ควรปรับปรุงพัฒนาให้กับครูผู้สอนในการวางแผนพัฒนาคุณภาพการจัดการเรียนรู้

(3) สถานศึกษามีความตื่นตัวในการดำเนินโครงการ/กิจกรรมที่สนองต่อนโยบายการยกระดับผลการทดสอบทางการศึกษาระดับชาติของผู้เรียนให้บรรลุเป้าหมายและผลลัพธ์ที่กำหนดไว้

ผลกระทบทางลบ พบว่า ในกรณีที่สถานศึกษาจัดโครงการติวข้อสอบเพื่อมุ่งหวังให้ผลการทดสอบทางการศึกษาระดับชาติของผู้เรียนบรรลุตามเป้าหมายที่ตั้งไว้ หรือไม่น้อยกว่าปีที่ผ่านมา ได้ส่งผลกระทบต่อการจัดการเรียนการสอนปกติ การเพิ่มเวลาติวในวันหยุดเสาร์-อาทิตย์ ครูผู้สอนต้องเพิ่มภาระในการจัดเตรียมและสละเวลาในการสอนเสริมเพิ่มเติมเพื่อให้ผู้เรียนเข้าใจในเนื้อหาตามหลักสูตรและครบถ้วน รวมถึงสถานศึกษาต้องมีการจัดเตรียมแผนงาน โครงการ บุคลากรผู้รับผิดชอบ บุคลากรที่ทำหน้าที่ติวทั้งภายในและภายนอก จัดสรรงบประมาณเวลาและสถานที่ และทรัพยากรอื่น ๆ ที่เกี่ยวข้อง

ประเด็นผลกระทบที่ 6 กลไกการพัฒนาคุณภาพการศึกษา ผลกระทบปรากฏดังนี้

ผลกระทบทางบวก พบว่า ผลการทดสอบทางการศึกษาระดับชาติมีความสัมพันธ์ทางบวกกับผลการประเมินคุณภาพภายนอกรอบสาม (พ.ศ. 2554-2558) โดยสำนักงานรับรองมาตรฐานและประเมินคุณภาพการศึกษา (องค์การมหาชน) หรือ สมศ. โดยมีค่าสัมประสิทธิ์สหสัมพันธ์ เท่ากับ 0.63 อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01 และมีความสัมพันธ์ทางบวกกับตัวบ่งชี้ในการประเมินคุณภาพนอก เป็นหลักฐานหนึ่งในการสนับสนุนว่าผลการทดสอบทางการศึกษาระดับชาตินี้มีผลกระทบทางบวกต่อกลไกการพัฒนาคุณภาพตามระบบการประกันคุณภาพการศึกษาของประเทศ ผลกระทบทางลบ พบว่า การใช้ผลการทดสอบทางการศึกษาระดับชาติที่ไม่เป็นไปตามวัตถุประสงค์ของการทดสอบ และ/หรือ ไม่ได้นำไปใช้ประโยชน์ ตัวอย่างเช่น การใช้คะแนน O-NET ไปประกอบเป็นเกณฑ์พิจารณาประเมินเพื่อโยกย้ายตำแหน่งของผู้บริหาร

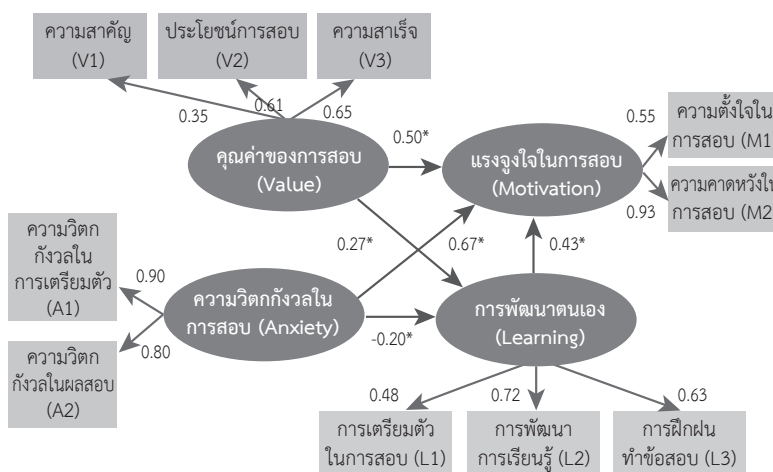
การประเมินครูผู้สอนเพื่อรับรางวัล OBEC award การนำคะแนน V-NET I-NET B-NET ไปใช้ประโยชน์ได้น้อยและไม่เป็นไปตามที่คาดหวังไว้ ส่งผลให้ผู้เรียน ครูผู้สอน ผู้บริหาร ตลอดจนผู้มีส่วนเกี่ยวข้องกับการทดสอบมีเจตคติที่ไม่ดีต่อการทดสอบทางการศึกษาระดับชาติ ขาดแรงจูงใจในการสอบ ไม่เห็นความสำคัญ รู้สึกถึงความไม่คุ้มค่า นำผลสอบใช้ประโยชน์ไม่ได้ ซึ่งเป็นผลกระทบต่อการพัฒนาที่ไม่สามารถนำผลการสอบไปใช้ในการพัฒนาคุณภาพการศึกษาของประเทศได้อย่างมีประสิทธิภาพและบรรลุตามเป้าหมายที่ตั้งไว้ ดังตัวอย่างความคิดเห็นต่อไปนี้

“...พอมีการสอบ I-NET ขึ้นมา ถามว่ามีประโยชน์ไหม ก็มีประโยชน์นะ แต่ถ้าอยากให้ประโยชน์มากกว่านี้ อยากให้นำผลการสอบ I-NET มีการนำผลไปใช้ ไม่ใช่แค่ให้นักเรียนสอบรู้ผลอยู่ในระดับใดแล้วก็จบเพียงเท่านั้น ผมเลยเห็นว่านอกจากจะเกิดประโยชน์ที่น้อยนิดแล้ว ยังเป็นการสร้างภาระงานเพิ่มให้กับพวกครูอีก ถ้ามันสามารถนำผลนี้ไปใช้ได้ผมว่ามันน่าจะมีประโยชน์มากกว่านี้นะ เพราะนี่เหมือนเป็นการสอบที่ทำให้ได้รู้ลึกๆ เป็นเพียงการสอบที่เน้นให้สอบผ่านเท่านั้น จะมีความรู้หรือไม่มีความรู้เด็กก็ไม่สนใจเท่าไร...” (สัมภาษณ์ครูผู้สอน, ต.ค. 62)

ผลกระทบต่อบุคคล เป็นผลกระทบของผลการทดสอบทางการศึกษาระดับชาติที่มีต่อผู้เรียนที่เข้ารับการทดสอบทางการศึกษาระดับชาติ ครู และผู้บริหารที่เกี่ยวข้องกับการทดสอบทางการศึกษาระดับชาติ มี 3 ประเด็นผลกระทบ ดังนี้

ประเด็นผลกระทบที่ 7 แรงจูงใจในการสอบทางการศึกษาระดับชาติของผู้เรียน ผลกระทบปรากฏดังนี้

ผลกระทบทางบวก พบว่า การเห็นคุณค่าของการสอบเป็นตัวแปรที่มีอิทธิพลบวกทางตรงต่อแรงจูงใจในการสอบซึ่งมีตัวแปรการพัฒนาตนเองของผู้เรียนเป็นตัวแปรคั่นกลาง (Mediating variable) ที่มีอิทธิพลทางตรงต่อแรงจูงใจในการสอบ อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 ข้อค้นพบเป็นหลักฐานสนับสนุนว่าการทดสอบทางการศึกษาระดับชาติมีผลต่อแรงจูงใจในการสอบของผู้เรียน ดังแผนภาพที่ 3



Chi-square=29.85, df=24, p-value=0.1897, RMSEA=0.026, ** p<0.05

แผนภาพที่ 3 โมเดลสมการโครงสร้างเชิงเส้นของปัจจัยที่ส่งผลต่อแรงจูงใจในการสอบระดับชาติ

ผลกระทบทางลบ พบว่า

- 1) ผู้เรียนขาดแรงจูงใจในการสอบ เพราะผลการทดสอบทางการศึกษาระดับชาติดังกล่าวไม่ได้มีผลต่อการสำเร็จการศึกษา การศึกษาต่อในระดับที่สูงขึ้น และเป็นภาระในการสอบ
- 2) ผลการทดสอบทางการศึกษาระดับชาติดูอยู่ในระดับต่ำ เพราะผู้เรียนขาดแรงจูงใจในการสอบ ไม่ให้ความสำคัญ ขอสอบยากเกินไปและไม่เหมาะสมกับผู้เรียน และผลการทดสอบทางการศึกษาระดับชาติไม่ได้มีผลต่อการสำเร็จการศึกษา
- 3) ความวิตกกังวลในการสอบ (Anxiety) เป็นตัวแปรกำกับ (Moderator variable) ที่มีอิทธิพลทางตรงต่อการพัฒนาตนเอง และแรงจูงใจในการสอบของผู้เรียน อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 ดังตัวอย่างความคิดเห็นต่อไปนี้

“O-NET จะมีผลกับเพื่อนที่ยื่นในรอบ Admission ถ้าเพื่อนคนไหนที่ยังไม่ยื่นคะแนนหรือได้คะแนนแล้ว ก็จะซีเรียสกับคะแนน O-NET ก็ต้องทำคะแนนให้ดี” (สัมภาษณ์นักศึกษา, ต.ค. 62, O-NET)

“ปัญหาของการนำ V-NET ไปใช้ คือ ไม่สอดคล้องกับการสอบประเภทอื่น ๆ ที่นักศึกษาต้องสอบ ทั้งการสอบมาตรฐานวิชาชีพ การสอบมาตรฐานฝีมือแรงงาน เมื่อเห็นว่า V-NET ไม่ได้ใช้ประโยชน์อะไรทำให้ไม่ได้สนใจสอบ บางแห่งมีนักศึกษาขาดสอบ แต่ส่วนนี้ก็กระทบกับการประเมินคุณภาพอาชีววะ ที่จุดหลักใหญ่ คือ นักศึกษาไม่เห็นความสำคัญของการสอบ ก็จะส่งผลกระทบต่อคะแนน” (สนทนากลุ่มอาชีวศึกษา, 2562, V-NET)

“ส่วนหนึ่งที่กระทบกับนักศึกษา คือ วิชาพื้นฐานที่สอบนักศึกษาได้เรียนตั้งแต่ชั้นปีแรก ๆ แต่การทดสอบเป็นการวัดความรู้ความคิดรวบยอดก็ถูกต้องตามหลักการวัดผลประเมินผลก่อนสำเร็จการศึกษา ที่ตัวเด็กต้องมีสมรรถนะความรู้พร้อมผลกระทบที่เกิดขึ้นจึงตกที่ทำอะไรให้เขามีสมรรถนะเหล่านั้น ธรรมชาติเขาก็เป็นนักปฏิบัติ ลงมือทำงานกันอยู่แล้ว พอต้องมาสอบลักษณะนี้ก็สอบได้แต่คะแนนอาจไม่ได้สะท้อนสิ่งที่เขามีเขาเป็นจริง ๆ ” (สนทนากลุ่มอาชีวศึกษา, 2562, V-NET)

“การสอบ I-NET ทั้ง 8 รายวิชาทำการสอบภายใน 1 วัน ถือว่าหนักพอสมควรกับเด็กนักเรียนที่ต้องสอบกับจำนวนรายวิชา...อีกทั้งยังเป็นข้อสอบที่ลึกไปถึงขั้นนำไปใช้ วิเคราะห์” (สัมภาษณ์ครูผู้สอน, ต.ค. 62, I-NET)

“ไม่ค่อยมีผลกับตัวเอง เพราะไม่ได้นำไปใช้ในการยื่นเข้ามหาวิทยาลัย” (สัมภาษณ์ผู้เรียน, ต.ค. 62, B-NET)

“ข้อสอบจะมีทั้งจำ นำไปใช้ วิเคราะห์ ส่วนนี้จะยากสำหรับผู้เรียน เพราะลึกกว่าที่ได้เรียนในชั้นเรียน ซึ่งต้องมีการเตรียมสอบในส่วนของบาลีสนามหลวงทางพระอาจารย์ก็มีการฝึก ทบทวน ตั้งโจทย์แต่งแก้กระทู้” (สัมภาษณ์ผู้บริหารโรงเรียน, ต.ค. 62, B-NET)

“กคน. นักเรียนที่มาเรียนส่วนใหญ่จะเป็นคนที่ผ่านพ้นวัยเรียนออกไปแล้ว
กำนัน ผู้ใหญ่บ้าน ผู้สูงอายุ ท่านไปดูสนามสอบไหนก็ได้ จะมีผู้สูงอายุมาเรียน 400 คน
ที่มีอายุ 60 ปีขึ้นไป แต่ข้อสอบ ให้สอบฟิลิกส์ เคมี ชีวะ ซึ่งเขาไม่ได้ใช้..” (สนทนากลุ่ม
กคน., ต.ค. 62, N-NET)

ประเด็นผลกระทบที่ 8 การจัดการเรียนรู้และการประเมินผลของครูผู้สอน ผลกระทบ
ปรากฏดังนี้

ผลกระทบทางบวก พบว่า

1) ผลการทดสอบทางการศึกษาระดับชาติใช้เป็นตัวชี้วัดคุณภาพการจัดการเรียนรู้
ของครูผู้สอน และใช้เป็นผลสะท้อนว่าครูผู้สอนได้จัดการเรียนรู้และพัฒนาผู้เรียนให้มีความรู้
ความสามารถได้ดีมากน้อยเพียงใด มีจุดใดที่ต้องปรับปรุงและพัฒนาการจัดการเรียนรู้และ
การประเมินผล

2) ผลการทดสอบทางการศึกษาระดับชาติมีความสัมพันธ์ทางบวกกับประสิทธิผล
ของการจัดการเรียนการสอนที่เน้นผู้เรียนเป็นสำคัญของครูผู้สอน

ผลกระทบทางลบ พบว่า

1) ครูผู้สอนมีภาระหน้าที่เพิ่มขึ้น เพราะต้องรับผิดชอบในการติวข้อสอบให้กับ
นักเรียน ส่งผลกระทบให้การจัดการเรียนการสอนไม่เป็นไปตามที่คาดหวังไว้ เกิดปรากฏการณ์
“ติวเพื่อสอบ” ส่งผลให้การจัดการศึกษาเกิดความไม่สมดุลและไม่สอดคล้องกันระหว่างหลักสูตร
(O) การเรียนการสอน (L) การวัดและประเมินผล (E)

2) ความไม่สอดคล้องระหว่างการทดสอบทางการศึกษาระดับชาติกับการจัดการ
เรียนรู้ในชั้นเรียนโดยพบในข้อสอบระดับชาติในด้านอิสลามศึกษาและพระพุทธศาสนาที่
มีลักษณะเน้นการคิดวิเคราะห์ เมื่อคะแนนต่ำจะเป็นตัวบ่งชี้ที่สะท้อนไปยังคุณภาพของ
การจัดการเรียนการสอนในชั้นเรียนของครูผู้สอนในแต่ละรายวิชา ส่งผลกระทบให้ครูผู้สอน
ต้องมีการทบทวนกระบวนการจัดการเรียนรู้ที่ต้องเน้นการคิดวิเคราะห์ และการวัดประเมินผล
ในชั้นเรียนที่สอดคล้องกับแนวทางการทดสอบระดับชาติ ดังตัวอย่างความคิดเห็นต่อไปนี้

“ค่อนข้างเป็นภาระเพราะต้องทั้งสอนให้ทันแล้วต้องติวข้อสอบ ทั้งเวลาเรียน
เวลาที่ต้องเพิ่มขึ้นการเชิญคนภายนอกให้มาติวให้อีก เป็นภาระกับครูเราพอสมควร
เป็นภาระความรับผิดชอบที่ต้องทำทุกปีเพื่อลูก ๆ ของเรา โรงเรียนของเรา”
(สัมภาษณ์ครูผู้สอน, ส.ค. 62, O-NET)

“ภาระหนัก ๆ เลยคือ อาจารย์ของวิทยาลัยเราต้องเดินทางพานักศึกษา
ไปสอบอีกสนามที่อยู่ไกลจากวิทยาลัย ต้องหารถไปส่ง ทั้งต้องรับผิดชอบเด็ก
ไหนจะต้องดูแลความเรียบร้อย ต้องติวข้อสอบอีก” (สัมภาษณ์ผู้บริหารอาชีวฯ,
ต.ค. 62, V-NET)

“...ก่อนจะสอบ I-NET โรงเรียนก็มีการวัดผลในระดับโรงเรียนอยู่แล้วตามปกติ พอมีการทดสอบ I-NET ขึ้นมา โรงเรียนก็ยังคงสอนเหมือนเดิมตามปกติ อาจจะมีส่วนในเรื่องของเนื้อหาเพิ่ม เช่น มีการติวเด็กนักเรียนเพิ่มขึ้น เพื่อเตรียมไปสอบ” (สัมภาษณ์ครูผู้สอน, ต.ค. 62, I-NET)

“เท่าที่ดูข้อสอบมา เราจะต้องมาเพิ่มการคิดวิเคราะห์ให้ผู้เรียนได้ฝึกวิเคราะห์มากขึ้น อาจต้องปรับในการจัดการเรียนการสอนด้วย” (สัมภาษณ์ครูผู้สอน, ต.ค. 62, B-NET)

“ครู 1 คน ต้องสอนทุกรายวิชา ต้องสอนทุกระดับ ที่ กศน. ครู 1 คน ต้องรับผิดชอบเด็ก 60 คน ถ้าน้อยกว่านี้ถือว่าตกเกณฑ์ เพราะจะต้องประเมิน อาจไม่ผ่านการประเมิน อาจไม่ต่อสัญญา ลองพิจารณาว่าคุณภาพจะขนาดไหน เรามีต้นทุนต่ำ ต้องการผลสัมฤทธิ์ที่สูงไม่ได้” (สัมภาษณ์ ผอ.กศน.อำเภอ, ต.ค. 62, N-NET)

ประเด็นผลกระทบที่ 9 ภาวะผู้นำทางวิชาการของผู้บริหารสถานศึกษา ผลกระทบปรากฏดังนี้

ผลกระทบทางบวก พบว่า

1) ผลการทดสอบทางการศึกษาระดับชาติดีผลกระทบทางบวกให้ผู้บริหารได้ใช้ความสามารถในการบริหารวิชาการ โดยที่ผู้บริหารต้องใช้กลยุทธ์การบริหารจัดการเพื่อพัฒนาผลการทดสอบทางการศึกษาระดับชาติของผู้เรียนให้บรรลุถึงเป้าหมายที่ตั้งไว้ โดยการจัดโครงการ/กิจกรรมที่จะช่วยส่งเสริมให้ผลการทดสอบทางการศึกษาระดับชาติเพิ่มสูงขึ้นกว่าปีที่ผ่านมา และเห็นแนวโน้มที่ดีเมื่อพิจารณาจาก 3 ปีย้อนหลัง

2) ผลการทดสอบทางการศึกษาระดับชาติดีมีค่าสัมประสิทธิ์สหสัมพันธ์พหุคูณ (r_{xy} เท่ากับ 0.35) เป็นความสัมพันธ์ทางบวกกับประสิทธิภาพการบริหารจัดการของผู้บริหาร และประสิทธิผลของการจัดการเรียนรู้ของครูผู้สอนอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01

ผลกระทบทางลบ พบว่า ความไม่สอดคล้องระหว่างการทดสอบทางการศึกษาระดับชาติกับหลักสูตรที่ใช้ในการจัดการศึกษามีผลกระทบต่อการบริหารจัดการหลักสูตรของผู้บริหารในบางประเภทการทดสอบ ดังนี้

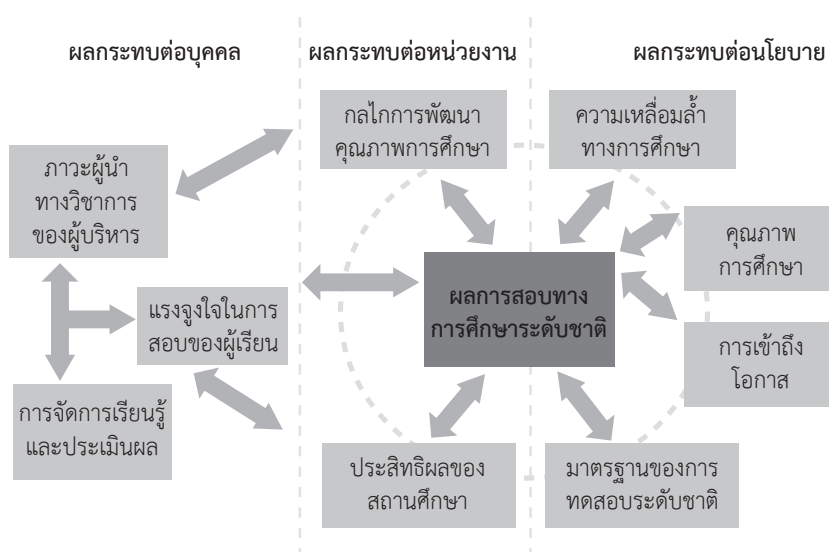
1) หลักสูตรอิสลามศึกษาในแต่ละสถานศึกษามีความแตกต่างกัน การใช้หลักสูตรมีความต่างกันไปตามบริบทของสถานศึกษา และการวัดประเมินผลผู้เรียนที่ต้องการให้อยู่ในมาตรฐานเดียวกัน ส่งผลกระทบต่อคะแนน I-NET และแรงจูงใจในการสอบของผู้เรียน

2) หลักสูตร กศน. ยังไม่สอดคล้องกับบริบทการจัดการศึกษาที่มีโครงสร้างหลักสูตรที่หลากหลายตามความต้องการของผู้เรียน และการวัดประเมินผลผู้เรียนที่ต้องการให้อยู่ในมาตรฐานเดียวกัน ส่งผลกระทบต่อคะแนน N-NET และแรงจูงใจในการสอบของนักศึกษา ดังตัวอย่างความคิดเห็นต่อไปนี้

“...คือพยายามจะปรับปรุงหลักสูตรแม้ไม่เอกภาพ แต่ผลสัมฤทธิ์ที่ได้ก็ควรเป็นในแนวทิศทางเดียวกัน” (สัมภาษณ์ผู้บริหารสถานศึกษา, ต.ค. 62, I-NET)

“หลักสูตรที่ กศน. ใช้อยู่ในปัจจุบันคือหลักสูตร 51 มันไม่สอดคล้องกับวิถีชีวิตความเป็นจริงกับนักศึกษา กศน. เลย เรียนหลายวิชามากเกินไป...ต้องนำไปใช้ใช้สอบ N-NET ที่ต้องรวบเข้า 5 สาระวิชาอีกมันเกินความสามารถของพวกเขาจริง ๆ ” (สนทนากลุ่ม กศน., ต.ค. 62, N-NET)

สรุปผลความสัมพันธ์เชื่อมโยงของประเด็นผลกระทบที่เกิดขึ้นจากการประเมินผลการทดสอบทางการศึกษาระดับชาติ ได้ตั้งแผนภาพที่ 4



แผนภาพที่ 4 ความสัมพันธ์ของประเด็นผลกระทบจากการประเมินผลการทดสอบทางการศึกษาระดับชาติ

2. ผลการติดตามและประเมินผลกระทบของการประเมินผลการทดสอบทางการศึกษาระดับชาติ

2.1 คุณภาพการศึกษา ผลการติดตามและประเมินผลกระทบต่อคุณภาพการศึกษาของการจัดการศึกษา 5 ประเภท พบว่า ผลการทดสอบทางการศึกษาระดับชาติมีผลกระทบต่อคุณภาพการศึกษาอยู่ในระดับปานกลางถึงมากค่าผลกระทบเฉลี่ยอยู่ระหว่าง 3.00-4.23 โดยที่ การทดสอบ O-NET มีค่าผลกระทบเฉลี่ยสูงที่สุดอยู่ในระดับมาก (Mean = 4.23, SD = 0.80) และการทดสอบ N-NET มีค่าผลกระทบเฉลี่ยต่ำที่สุดอยู่ในระดับปานกลาง (Mean = 3.00, SD = 1.30) สำหรับประเด็นที่มีผลกระทบสูงสุดเท่ากัน 2 ประเด็นในการทดสอบ O-NET คือ ผลการสอบใช้เป็นหลักฐานบ่งชี้คุณภาพการจัดการเรียนการสอนของครู/อาจารย์ผู้สอน และผลการสอบใช้เป็นหลักฐานบ่งชี้ความสำเร็จของการดำเนินงานภายในสถานศึกษามีผลกระทบอยู่ในระดับมาก (Mean=4.31, SD=0.79)

2.2 การเข้าถึงโอกาสทางการศึกษา ผลการติดตามผลกระทบต่อการเข้าถึงโอกาสทางการศึกษาของผู้เรียนที่ได้นำคะแนนสอบระดับชาติไปใช้ พบว่า ในทุกประเภทการทดสอบ ผู้เรียนได้นำคะแนนสอบระดับชาติไปใช้ในการยื่นประกอบการสมัครเข้าศึกษาต่อในระดับที่สูงขึ้น สูงที่สุดในการทดสอบ O-NET คิดเป็นร้อยละ 86.57 รองลงมานำไปใช้ในการยื่นประกอบการสมัครเข้าทำงานในการทดสอบ N-NET คิดเป็นร้อยละ 68.14 อย่างไรก็ตามพบว่าการนำไปยื่นประกอบการขออนุมัติการศึกษา บอกระดับความสามารถของตนเองเพื่อแก้ไขและปรับปรุงการเรียนรู้

2.3 ความเหลื่อมล้ำทางการศึกษา ผลการติดตามผลกระทบต่อการลดความเหลื่อมล้ำทางการศึกษา พบว่า ผู้เรียนทุกคนได้รับสิทธิเข้ารับการทดสอบมาตรฐานระดับชาติอย่างเท่าเทียมกันตามระดับชั้นของแต่ละบุคคล โดยที่สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) ได้จัดให้มีแบบสอบมาตรฐาน การจัดให้สอบผ่านระบบการทดสอบอิเล็กทรอนิกส์ (E-testing) สำหรับการสอบ B-NET และ N-NET การพัฒนาศักยภาพให้กับครูผู้สอนด้านการทดสอบ มีระบบอำนวยความสะดวกการจัดสอบสำหรับผู้พิการ อาทิ อักษรเบลล์ อักษรตัวขยายสำหรับผู้พิการทางสายตา

2.4 มาตรฐานของการทดสอบทางการศึกษาระดับชาติ ผลการติดตามและประเมินผลกระทบต่อมาตรฐานของการทดสอบทางการศึกษาระดับชาติ พบว่า การสร้างความเชื่อมั่นและมีมาตรฐานของการทดสอบโดยภาพรวมมีผลกระทบอยู่ในระดับมากทุกประเภทการทดสอบ มีค่าผลกระทบเฉลี่ยอยู่ระหว่าง 3.72-4.47 เมื่อพิจารณารายประเภทการทดสอบ ปรากฏผลดังนี้ ในการทดสอบ O-NET ประเด็นที่มีผลกระทบสูงสุดคือ คณะกรรมการประจำสนามสอบมีเพียงพอและมีความเหมาะสมมีผลกระทบอยู่ในระดับมากที่สุด (Mean = 4.61, SD = 0.56) ในการทดสอบ V-NET ประเด็นที่มีผลกระทบสูงสุดคือ ผลสอบมีความน่าเชื่อถือมีผลกระทบอยู่ในระดับมากที่สุด (Mean = 4.67, SD = 0.52) ในการสอบ I-NET ประเด็นที่มีผลกระทบสูงสุดคือ สนามสอบมีความพร้อมในการจัดสอบมีผลกระทบอยู่ในระดับมาก (Mean = 4.22, SD = 0.90) ในการสอบ B-NET ประเด็นที่มีผลกระทบสูงสุดคือ การประชาสัมพันธ์ข้อมูลการทดสอบมีความชัดเจนมีผลกระทบอยู่ในระดับมาก (Mean = 4.15, SD = 0.68) ในการทดสอบ N-NET มีผลกระทบสูงที่สุดในประเด็นการประชาสัมพันธ์ข้อมูลการทดสอบมีความชัดเจน สนามสอบมีความพร้อมในการจัดสอบคณะกรรมการประจำสนามสอบมีเพียงพอและมีความเหมาะสม การประกาศผลสอบของ สทศ.มีความถูกต้องรวดเร็วและแม่นยำ และผลการสอบมีความน่าเชื่อถือมีผลกระทบอยู่ในระดับมากที่สุดเท่ากัน (Mean = 4.67, SD = 0.82)

2.5 ประสิทธิภาพของสถานศึกษา ผลการติดตามและประเมินผลกระทบของการทดสอบทางการศึกษาระดับชาติต่อประสิทธิภาพของสถานศึกษาในภาพรวมของการทดสอบทั้ง 5 ประเภท พบว่า ค่าผลกระทบเฉลี่ย (Mean) มีค่าระหว่าง 3.56-4.56 ส่วนเบี่ยงเบนมาตรฐาน (SD) มีค่าระหว่าง 0.56-1.01 อยู่ในระดับผลกระทบมากถึงมากที่สุด

2.6 กลไกการพัฒนาคุณภาพการศึกษา ผลการติดตามและประเมินผลกระทบของการทดสอบทางการศึกษาระดับชาติดต่อการใช้กลไกผลการทดสอบทางการศึกษาระดับชาติในการพัฒนาคุณภาพการศึกษา พบว่าในภาพรวมสถานศึกษาได้ใช้ผลการทดสอบทางการศึกษาระดับชาติเป็นกลไกการพัฒนาคุณภาพการศึกษามีค่าผลกระทบเฉลี่ย (Mean) มีค่าระหว่าง 3.55-4.39 ส่วนเบี่ยงเบนมาตรฐาน (SD) มีค่าระหว่าง 0.69-0.97 อยู่ในระดับผลกระทบมากทุกประเภทการทดสอบ

2.7 แรงจูงใจในการสอบทางการศึกษาระดับชาติของผู้เรียน ผลการติดตามและประเมินผลกระทบของการทดสอบทางการศึกษาระดับชาติดต่อแรงจูงใจในการสอบทางการศึกษาระดับชาติของผู้เรียน พบว่า

ด้านที่ 1 การเห็นคุณค่าของการสอบทางการศึกษาระดับชาติ พบว่า ในภาพรวมทุกประเภทการทดสอบผู้เรียนมีการเห็นคุณค่าของการสอบทางการศึกษาระดับชาติมีค่าผลกระทบเฉลี่ย (Mean) อยู่ระหว่าง 3.34-3.93 ส่วนเบี่ยงเบนมาตรฐาน (SD) มีค่าระหว่าง 0.37-0.58 อยู่ในระดับผลกระทบปานกลางถึงมาก

ด้านที่ 2 การพัฒนาตนเองในการสอบทางการศึกษาระดับชาติ พบว่า ในภาพรวมทุกประเภทการทดสอบผู้เรียนมีการพัฒนาตนเองในการสอบทางการศึกษาระดับชาติ มีค่าผลกระทบเฉลี่ยอยู่ระหว่าง 3.06-3.50 ส่วนเบี่ยงเบนมาตรฐาน (SD) มีค่าระหว่าง 0.49-0.59 อยู่ในระดับผลกระทบปานกลาง ด้านที่ 3 แรงจูงใจในการสอบทางการศึกษาระดับชาติ พบว่า ในภาพรวมทุกประเภทการทดสอบผู้เรียนมีแรงจูงใจในการสอบทางการศึกษาระดับชาติมีค่าผลกระทบเฉลี่ยอยู่ระหว่าง 3.51-3.82 ส่วนเบี่ยงเบนมาตรฐาน (SD) มีค่าระหว่าง 0.39-0.50 อยู่ในระดับผลกระทบมาก

2.8 การจัดการเรียนรู้และการประเมินผลของครูผู้สอน ผลการติดตามและประเมินผลกระทบของการทดสอบทางการศึกษาระดับชาติดต่อครูผู้สอนในการพัฒนาการจัดการเรียนรู้และประเมินผลพบว่า ในภาพรวมการทดสอบทางการศึกษาระดับชาติส่งผลกระทบต่อครูผู้สอนในการพัฒนาการจัดการเรียนรู้และประเมินผลมีค่าผลกระทบเฉลี่ยอยู่ระหว่าง 3.62-4.02 ส่วนเบี่ยงเบนมาตรฐาน (SD) มีค่าระหว่าง 0.61-1.08 อยู่ในระดับผลกระทบมากทุกประเภทการทดสอบ

2.9 ภาวะผู้นำทางวิชาการของผู้บริหารสถานศึกษา ผลการติดตามและประเมินผลกระทบของการทดสอบทางการศึกษาระดับชาติดต่อภาวะผู้นำทางวิชาการของสถานศึกษาในการพัฒนาผลสอบ พบว่า ในภาพรวมการทดสอบทางการศึกษาระดับชาติส่งผลกระทบต่อผู้บริหารใช้ภาวะผู้นำทางวิชาการขับเคลื่อนพัฒนาผลสัมฤทธิ์ของผู้เรียนมีค่าผลกระทบเฉลี่ย (Mean) อยู่ระหว่าง 3.66-4.64 ส่วนเบี่ยงเบนมาตรฐาน (SD) มีค่าระหว่าง 0.50-1.03 อยู่ในระดับผลกระทบมากถึงมากที่สุด

3. ข้อเสนอแนะเชิงนโยบายในการปรับปรุงและพัฒนาการทดสอบทางการศึกษาระดับชาติ

การประเมินผลการทดสอบทางการศึกษาระดับชาติได้สร้างผลกระทบทั้งในด้านบวกและด้านลบ เกิดผลกระทบในระดับน้อยไปจนถึงระดับมากที่สุด ทั้งในระดับนโยบาย หน่วยงาน และบุคคล ดังนั้นเพื่อให้สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) หรือ สทศ. และหน่วยงานที่เกี่ยวข้องได้นำผลการทดสอบทางการศึกษาระดับชาติไปใช้ให้มีประสิทธิภาพ และเกิดประโยชน์สูงสุดในการประเมินหรือใช้ในวัตถุประสงค์อื่นของหน่วยงานต้นสังกัด จึงมีข้อเสนอแนะเชิงนโยบาย ดังตารางที่ 3

ตารางที่ 3 ข้อเสนอแนะเชิงนโยบายในการปรับปรุงและพัฒนาการทดสอบทางการศึกษาระดับชาติ

เป้าหมายที่ 1 การพัฒนานวัตกรรมทดสอบและการรายงานผล (จากผลการติดตามประเมินในประเด็นผลกระทบการเข้าถึงโอกาสทางการศึกษา)	
หน่วยงานที่รับผิดชอบ	วิธีการ/แนวทางการพัฒนา
สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน)	(1) ควรขับเคลื่อนและพัฒนานวัตกรรมการรายงานผลและแปลความหมายของคะแนนการทดสอบระดับชาติที่มีความหมายมากขึ้นให้กับสถานศึกษาได้นำสารสนเทศไปใช้ประโยชน์ในการประกันคุณภาพภายในและการประเมินคุณภาพภายนอกได้อย่างมีประสิทธิภาพยิ่งขึ้น (2) ควรสร้างและพัฒนาคลังข้อสอบมาตรฐานของการทดสอบทางการศึกษาระดับชาติ ด้านอาชีวศึกษา (V-NET) ด้านพระพุทธศาสนา (B-NET) ด้านการศึกษานอกระบบ โรงเรียน (N-NET) พร้อมกับปรับปรุงประสิทธิภาพของรูปแบบการทดสอบผ่านระบบอิเล็กทรอนิกส์ (e-testing) ที่สามารถทดสอบ วิเคราะห์ และรายงานผลได้ทันที และนำผลไปใช้ปรับปรุงพัฒนาสถานศึกษา เพื่อลดต้นทุนของการจัดสอบ และเกิดความคุ้มค่ามากที่สุด (3) ควรเปิดโอกาสการเข้าถึงข้อสอบทางการศึกษาระดับชาติด้านอาชีวศึกษา (V-NET) ด้านอิสลามศึกษา (I-NET) ด้านพระพุทธศาสนา (B-NET) ให้ผู้เรียนและอาจารย์ผู้สอนสามารถเข้าถึงได้ง่ายและทั่วถึงเพิ่มขึ้น เพื่อเพิ่มโอกาสในการเรียนรู้และพัฒนาตนเองผ่านข้อสอบมาตรฐาน
กระทรวงศึกษาธิการ	(4) ควรวิจัยและพัฒนาระบบฐานข้อมูลการติดตามและประเมินผลสัมฤทธิ์ของผู้เรียนจากคะแนนผลการทดสอบระดับชาติอย่างจริงจังในทุกประเภทการทดสอบที่เชื่อมโยงกับการประกันคุณภาพภายในสถานศึกษา และเชื่อมโยงกับการประเมินคุณภาพภายนอก
สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน (สพฐ.)	(5) ควรจัดทำระบบการส่งเสริมสนับสนุนในด้านการจัดการเรียนรู้และการวัดประเมินผลให้แก่สถานศึกษาและผู้เรียนให้มีความพร้อมในการเข้าสอบ O-NET เช่น การจัดการศึกษาทางไกลผ่านดาวเทียม DLTV การพัฒนาระบบคลังข้อสอบมาตรฐานแบบออนไลน์ เป็นต้น (6) ควรพัฒนาคลังข้อสอบตามมาตรฐานและตัวชี้วัดของหลักสูตรที่เน้นทักษะการคิดวิเคราะห์ให้บริการแก่สถานศึกษา (SIBS) รวมทั้งมีการพัฒนาคลังข้อสอบตามแนวทางการทดสอบระดับนานาชาติ PISA เพื่อเป็นการพัฒนาทักษะการคิดให้แก่ผู้เรียน

เป้าหมายที่ 2 การปฏิรูประบบการทดสอบทางการศึกษาของประเทศ (จากผลการติดตามประเมินในประเด็นผลกระทบมาตรฐานของการทดสอบทางการศึกษาระดับชาติ)	
หน่วยงานที่รับผิดชอบ	วิธีการ/แนวทางการพัฒนา
กระทรวงศึกษาธิการ ร่วมกับ สถาบันทดสอบ ทางการศึกษาแห่งชาติ (องค์การมหาชน)	(1) ควรปฏิรูปวิธีการทดสอบ วัดผล และประเมินผลทางการศึกษาระดับชาติให้มีความสอดคล้องสัมพันธ์ไปกับการจัดการเรียนรู้อยู่ใหม่ (2) ควรวิจัยและพัฒนารูปแบบการทดสอบวัดผลสัมฤทธิ์ของผู้เรียนตามหลักสูตรการศึกษาขั้นพื้นฐานในแต่ละช่วงชั้นที่มีประสิทธิภาพและเกิดประโยชน์สูงสุดต่อผู้เรียน ครูผู้สอน และผู้บริหารสถานศึกษา (3) ควรพัฒนารูปแบบและวิธีการทดสอบระดับชาติที่ครอบคลุมคุณภาพของผู้เรียนครบทั้งด้านพุทธิพิสัย ทักษะพิสัย และจิตพิสัย (K P A)
เป้าหมายที่ 3 การลดความเหลื่อมล้ำทางการศึกษาที่เกิดจากการทดสอบทางการศึกษาระดับชาติ (จากผลการติดตามประเมินในประเด็นผลกระทบความเหลื่อมล้ำทางการศึกษา)	
หน่วยงานที่รับผิดชอบ	วิธีการ/แนวทางการพัฒนา
สำนักงานคณะกรรมการ การศึกษาขั้นพื้นฐาน (สพฐ.)	(1) ควรทำศึกษาผลกระทบของการติวข้อสอบ O-NET ของสถานศึกษาในสังกัด (2) ควรมีการกำหนดข้อเสนอแนะเชิงนโยบายที่จะนำไปสู่การกำหนดมาตรการและแนวทางปฏิบัติที่เหมาะสมในการเตรียมความพร้อมสำหรับการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐานของสถานศึกษาในสังกัด
เป้าหมายที่ 4 การสร้างเครือข่ายความร่วมมือในการพัฒนาการทดสอบทางการศึกษาระดับชาติ (จากผลการติดตามประเมินในประเด็นผลกระทบมาตรฐานของการทดสอบทางการศึกษาระดับชาติ)	
หน่วยงานที่รับผิดชอบ	วิธีการ/แนวทางการพัฒนา
สถาบันทดสอบ ทางการศึกษาแห่งชาติ (องค์การมหาชน)	(1) ควรร่วมมือกับสำนักงานคณะกรรมการการอาชีวศึกษา (สอศ.) สถาบันคุณวุฒิวิชาชีพ (สคช) และกรมพัฒนาฝีมือแรงงาน การออกแบบระบบการทดสอบทางการศึกษาระดับชาติด้านอาชีวศึกษา (V-NET) ให้มีความสอดคล้องกับมาตรฐานวิชาชีพ
สำนักงานคณะกรรมการ การอาชีวศึกษา (สอศ.)	(1) ควรร่วมมือกับสถาบันทดสอบทางการศึกษาแห่งชาติ (สทศ.) สถาบันคุณวุฒิวิชาชีพ (สคช) และกรมพัฒนาฝีมือแรงงานในการสร้างระบบการทดสอบที่มีประสิทธิภาพครอบคลุมการรายงานผลทั้งการทดสอบทางการศึกษาระดับชาติด้านอาชีวศึกษา (V-NET) การทดสอบมาตรฐานวิชาชีพ และการทดสอบเพื่อรับรองมาตรฐานฝีมือแรงงาน เพื่อเพิ่มโอกาสให้นักศึกษาอาชีวศึกษาได้รับการรับรองคุณภาพที่สัมพันธ์กับค่าตอบแทนหรือเงินเดือนที่ควรจะได้รับ (2) ควรร่วมมือกับสถาบันทดสอบทางการศึกษาแห่งชาติ (สทศ.) สถาบันคุณวุฒิวิชาชีพ (สคช) และกรมพัฒนาฝีมือแรงงาน ในการพิจารณาความเชื่อมโยงของสิ่งที่วัดในการทดสอบทางการศึกษาระดับชาติด้านอาชีวศึกษา (V-NET) ที่มีความสอดคล้องกับกรอบคุณวุฒิแห่งชาติที่เป็นกรอบแนวทางการเชื่อมโยงผลลัพธ์การเรียนรู้ของระดับคุณวุฒิการศึกษาตามเกณฑ์มาตรฐานการเรียนรู้ในแต่ละระดับ และประเภทการศึกษากับระดับการปฏิบัติงานตามมาตรฐานอาชีพ เพื่อเทียบโอนคะแนนสอบ V-NET กับการรับรองมาตรฐานฝีมือแรงงานและมาตรฐานอาชีพ

หน่วยงานที่รับผิดชอบ	วิธีการ/แนวทางการพัฒนา
สำนักงานคณะกรรมการส่งเสริมการศึกษาเอกชน (สช.)	(1) ควรร่วมมือกับสถาบันทดสอบทางการศึกษาแห่งชาติ (สทศ.) ในการส่งเสริมสนับสนุนสื่อการเรียนรู้ที่เป็นมาตรฐานและสอดคล้องกับหลักสูตรอิสลามศึกษา และการทดสอบทางการศึกษาระดับชาติด้านอิสลามศึกษา (I-NET) อย่างทั่วถึง (2) ควรร่วมมือกับสถาบันทดสอบทางการศึกษาแห่งชาติ (สทศ.) ในการปรับปรุงวิธีการและกระบวนการออกข้อสอบให้มีความสอดคล้องกันระหว่างหลักสูตรอิสลามศึกษา กระบวนการจัดการเรียนรู้ที่เน้นการคิดวิเคราะห์ตามหลักสูตรอิสลามศึกษา การทดสอบและประเมินผลที่มีความเป็นมาตรฐาน (3) ควรร่วมมือกับสถาบันทดสอบทางการศึกษาแห่งชาติ (สทศ.) ในการร่วมกันออกแบบแนวทางการใช้ผลการทดสอบทางการศึกษาระดับชาติด้านอิสลามศึกษา (I-NET) สำหรับการปรับปรุงคุณภาพหลักสูตร กระบวนการจัดการเรียนรู้ และการวัดและประเมินผล ที่สอดคล้องกับระบบการประกันคุณภาพภายในสถานศึกษา
เป้าหมายที่ 5 การรักษามาตรฐานของการทดสอบระดับชาติ <i>(จากผลการติดตามประเมินในประเด็นผลกระทบมาตรฐานของการทดสอบทางการศึกษาระดับชาติ)</i>	
หน่วยงานที่รับผิดชอบ	วิธีการ/แนวทางการพัฒนา
สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน)	(1) ควรปรับปรุงทบทวนโครงสร้างการทดสอบ ขอบข่ายของเนื้อหา ความยากง่ายของแบบทดสอบรายวิชา O-NET ที่เหมาะสมและยุติธรรมกับนักเรียนทั้งที่เรียนในสายวิทย์คณิต และนักเรียนที่เรียนในสายศิลป์ (2) ควรเลือกใช้ภาษาที่เหมาะสมสำหรับทดสอบนักเรียนที่อยู่ในระดับอิสลามศึกษาตอนต้น (ระดับอิสลามศึกษาตอนต้น) ระดับอิสลามศึกษาตอนกลาง (มุตะวัซซีญะห์) และระดับอิสลามศึกษาตอนปลาย (ซานาเวียห์) (3) ควรกำหนดและจัดตารางเวลาในการสอบ 8 รายวิชา ให้เหมาะสมสำหรับการสอบ I-NET ของนักเรียนในแต่ละระดับ (4) ควรเพิ่มประสิทธิภาพในการบริหารจัดการสอบ B-NET ให้เอื้อต่อผู้เรียนที่เป็นพระและสามเณรที่เข้าสอบในแต่ละสนามสอบ (5) ควรพัฒนารูปแบบและวิธีการวัดประเมินผลสัมฤทธิ์ของผู้เรียน กศน.ที่มีความแตกต่างตามช่วงวัย (6) ควรพิจารณาให้การสอบปลายภาคของ กศน.กับการสอบ N-NET อยู่ในช่วงเวลาเดียวกัน
เป้าหมายที่ 6 การคัดเลือกบุคคลเข้าศึกษาต่อในระดับอุดมศึกษาโดยใช้ผลการทดสอบทางการศึกษาระดับชาติ <i>(จากผลการติดตามประเมินในประเด็นผลกระทบการเข้าถึงโอกาสทางการศึกษา)</i>	
หน่วยงานที่รับผิดชอบ	วิธีการ/แนวทางการพัฒนา
ที่ประชุมอธิการบดีแห่งประเทศไทย (ทปอ.) ร่วมกับกระทรวงศึกษาธิการ	(1) ควรศึกษาวิจัยผลกระทบของการนำคะแนน O-NET ไปใช้เป็นองค์ประกอบในการคัดเลือกบุคคลเข้าศึกษาต่อในระดับอุดมศึกษาในสาขาวิชาและคณะต่าง ๆ (2) ควรศึกษาแนวทางการใช้ผลคะแนนของการทดสอบทางการศึกษาระดับชาติด้านอิสลามศึกษา (I-NET) ในการเป็นองค์ประกอบหนึ่งในการคัดเลือกบุคคลเข้าศึกษาต่อในระดับมหาวิทยาลัย หรือสถาบันต่าง ๆ ของประเทศไทย และต่างประเทศได้

หน่วยงานที่รับผิดชอบ	วิธีการ/แนวทางการพัฒนา
	(3) ควรศึกษาแนวทางการใช้ผลคะแนนของการทดสอบทางการศึกษาระดับชาติด้านพระพุทธศาสนา (B-NET) ในการเป็นองค์ประกอบหนึ่งในการคัดเลือกบุคคลเข้าศึกษาต่อในระดับมหาวิทยาลัยสงฆ์ ได้แก่ มหาวิทยาลัยมหาจุฬาลงกรณราชวิทยาลัย (มจร.) มหาวิทยาลัยมหามกุฏราชวิทยาลัย (มมร.) หรือสถาบันอุดมศึกษาต่าง ๆ ได้
เป้าหมายที่ 7 การประเมินคุณภาพสถานศึกษาโดยใช้ผลการทดสอบทางการศึกษาระดับชาติ (จากผลการติดตามประเมินในประเด็นผลกระทบคุณภาพการศึกษา)	
หน่วยงานที่รับผิดชอบ	วิธีการ/แนวทางการพัฒนา
สำนักงานรับรองมาตรฐานและประเมินคุณภาพการศึกษา (องค์การมหาชน) หรือ สมศ.	(1) ควรนำเสนอสารสนเทศจากผลกระทบของการประเมินผลการทดสอบทางการศึกษาระดับชาติไปใช้ในการวางแผน กำหนดทิศทาง และพัฒนาแนวทางการประเมินคุณภาพภายนอกของสถานศึกษาในด้านคุณภาพของผู้เรียน (2) ควรศึกษาวิจัยและพัฒนาวิธีการประเมินผลสัมฤทธิ์ทางการเรียนของผู้เรียนที่ต้องนำผลการทดสอบระดับชาติมาประกอบประเมินคุณภาพให้มีความแม่นยำ เชื่อถือได้ และมีความยุติธรรมในการนำผลไปใช้ประเมินคุณภาพของสถานศึกษา
เป้าหมายที่ 8 การนำผลการทดสอบทางการศึกษาระดับชาติไปใช้ในการพัฒนาคุณภาพการศึกษา (จากผลการติดตามประเมินในประเด็นผลกระทบประสิทธิผลของสถานศึกษาและกลไกการพัฒนาคุณภาพการศึกษา)	
หน่วยงานที่รับผิดชอบ	วิธีการ/แนวทางการพัฒนา
สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน (สพฐ.)	(1) ควรร่วมมือกับสถาบันทดสอบทางการศึกษาแห่งชาติ (สทศ.) ในการออกแบบแนวทางการนำผลการทดสอบ O-NET ไปใช้พัฒนาการจัดการเรียนรู้และการวัดประเมินผลในระดับชั้นเรียนของครู (2) ควรส่งเสริมสนับสนุนให้หน่วยงานทุกระดับของ สพฐ. (ส่วนกลาง เขตพื้นที่การศึกษา และสถานศึกษา) วิเคราะห์ผลการทดสอบ O-NET ในปีการศึกษาที่ผ่านมา เพื่อค้นหากลุ่มสาระการเรียนรู้ มาตรฐานการเรียนรู้ และสาระการเรียนรู้ที่ต้องปรับปรุงเร่งด่วน นำไปสู่การวิเคราะห์หาสาเหตุและกำหนดแนวทางการยกระดับคุณภาพผู้เรียนที่สอดคล้องกับบริบทของหน่วยงานในแต่ละระดับ (3) ควรส่งเสริมสนับสนุนให้สถานศึกษาจัดการเรียนการสอนและการวัดประเมินผลตามมาตรฐานและตัวชี้วัดของหลักสูตร เพื่อให้ผู้เรียนมีความพร้อมและบรรลุตามมาตรฐานของหลักสูตรที่สอดคล้องกับการทดสอบ O-NET (4) ควรส่งเสริมสนับสนุนให้สถานศึกษานำรูปแบบข้อสอบและกระดาษคำตอบ O-NET ไปใช้ในการวัดและประเมินผลในชั้นเรียน เพื่อให้ผู้เรียนคุ้นชินกับรูปแบบของข้อสอบและกระดาษคำตอบ
สำนักงานคณะกรรมการส่งเสริมการศึกษาเอกชน (สช.)	(5) ควรพัฒนาความร่วมมือระหว่างศูนย์การศึกษาอิสลามประจำมัสยิด (ตาดีกา) กับโรงเรียนเอกชนสอนศาสนาอิสลามในด้านหลักสูตรอิสลามศึกษา กระบวนการจัดการเรียน การวัดและการประเมินผลในการส่งเสริมสนับสนุนการจัดการศึกษาอิสลามให้ไปสู่เป้าหมายร่วมกัน

หน่วยงานที่รับผิดชอบ	วิธีการ/แนวทางการพัฒนา
สำนักงานพระพุทธศาสนาแห่งชาติในส่วนงานของโรงเรียนพระปริยัติธรรมแผนกสามัญ	(6) ควรร่วมกับกระทรวงศึกษาธิการในการกำหนดแนวทางการนำคะแนน B-NET รายงานในแบบแสดงผลการเรียนรู้ของผู้ที่สำเร็จการศึกษา ในสัดส่วนของ GPAX : O-NET : B-NET ที่เหมาะสม (7) ควรร่วมมือกับสถาบันทดสอบทางการศึกษาแห่งชาติ (สทศ.) ในการส่งเสริมสนับสนุนสื่อการเรียนรู้และนวัตกรรมที่เป็นมาตรฐานและสอดคล้องกับหลักสูตรพระปริยัติธรรมและการทดสอบทางการศึกษาระดับชาติด้านพระพุทธศาสนา (B-NET) อย่างทั่วถึง (8) ควรร่วมมือกับสถาบันทดสอบทางการศึกษาแห่งชาติ (สทศ.) ในการร่วมกันออกแบบแนวทางการใช้ผลการทดสอบทางการศึกษาระดับชาติด้านพระพุทธศาสนา (B-NET) สำหรับการปรับปรุงคุณภาพหลักสูตร กระบวนการจัดการเรียนรู้ และการวัดและประเมินผลที่สอดคล้องกับระบบการประกันคุณภาพภายในสถานศึกษา
สำนักงานส่งเสริมการศึกษานอกระบบและการศึกษาตามอัธยาศัย (กศน.)	(9) ควรมีการทบทวนและปรับปรุงโครงสร้างหลักสูตร กศน. โครงสร้างรายวิชาที่มีความเหมาะสมกับธรรมชาติของนักศึกษา กศน. และสอดคล้องกับแนวทางการวัดและประเมินผลมาตรฐานของชาติ (10) ควรร่วมมือกับสถาบันทดสอบทางการศึกษาแห่งชาติ (สทศ.) ในการร่วมกันออกแบบแนวทางการใช้ผลการทดสอบทางการศึกษาระดับชาติด้านการศึกษานอกระบบโรงเรียน (N-NET) สำหรับการปรับปรุงคุณภาพหลักสูตร กระบวนการจัดการเรียนการสอน และการวัดและประเมินผล ที่สอดคล้องกับระบบการประกันคุณภาพภายใน กศน.
หน่วยงานต้นสังกัด/เขตพื้นที่การศึกษา	(11) ควรพัฒนารูปแบบการพัฒนาผลสัมฤทธิ์ของผู้เรียนเพื่อส่งเสริมสนับสนุนให้สถานศึกษาได้นำไปสู่การปฏิบัติที่จะส่งผลต่อคุณภาพของสถานศึกษา (12) สำหรับสถานศึกษาขั้นพื้นฐาน ควรพัฒนาผู้บริหารสถานศึกษาในประเด็นการนิเทศติดตามการจัดกิจกรรมการเรียนรู้ของครูผู้สอนในชั้นเรียนอย่างต่อเนื่อง (13) สำหรับสถานศึกษาอาชีวศึกษา ควรพัฒนาผู้บริหารสถานศึกษาในประเด็นการส่งเสริมสนับสนุนให้มีบรรยากาศทางวิชาการที่จะช่วยยกระดับผลการทดสอบระดับชาติด้านอาชีวศึกษา (V-NET) (14) สำหรับสถานศึกษาเอกชนสอนศาสนาอิสลาม ควรพัฒนาผู้บริหารสถานศึกษาในประเด็นการประชุมกำกับติดตามการดำเนินโครงการพัฒนาผู้เรียนให้มีคะแนนสอบระดับชาติด้านอิสลามศึกษา (I-NET) สูงขึ้น (15) สำหรับโรงเรียนพระปริยัติธรรมแผนกสามัญศึกษา ควรพัฒนาผู้บริหารสถานศึกษาในประเด็นการนิเทศติดตามการจัดกิจกรรมการเรียนรู้ของครูผู้สอนในชั้นเรียนอย่างต่อเนื่อง (16) สำหรับศูนย์การศึกษานอกระบบโรงเรียนประจำตำบล/อำเภอ/จังหวัด ควรพัฒนาผู้บริหาร กศน. ในประเด็นการนิเทศติดตามการจัดกิจกรรมการเรียนรู้ของครูผู้สอนในชั้นเรียนอย่างต่อเนื่อง

หน่วยงานที่รับผิดชอบ	วิธีการ/แนวทางการพัฒนา
สถานศึกษาในการจัดการศึกษาแต่ละประเภท	(17) ควรวางแผนพัฒนาคุณภาพการศึกษาโดยใช้สารสนเทศจากผลการประเมินผลการทดสอบทางการศึกษาระดับชาติ (18) ควรพัฒนาระบบ กลไก วิธีการ การสร้างความเข้าใจ และเจตคติที่ดีในการนำผลการทดสอบระดับชาติมาใช้ในระบบการประกันคุณภาพภายในสถานศึกษา (19) สำหรับสถานศึกษาขั้นพื้นฐาน ควรพัฒนาครูผู้สอนในประเด็นการนำผลการสอบระดับชาติ (O-NET) มาใช้จัดปรับปรุงแก้ไขให้กับผู้เรียนได้ทราบ (20) สำหรับสถานศึกษาอาชีวศึกษา ควรพัฒนาครูผู้สอนในประเด็นการนำข้อสอบระดับชาติ (V-NET) ที่ผ่านมาฝึกให้ผู้เรียนได้ทำอย่างสม่ำเสมอ (21) สำหรับสถานศึกษาเอกชนสอนศาสนาอิสลาม ควรพัฒนาครูผู้สอนในประเด็นการจัดกิจกรรมการเรียนรู้ที่สอดคล้องกับแนวทางทำข้อสอบระดับชาติ (I-NET) (22) สำหรับโรงเรียนปริยัติธรรมแผนกสามัญศึกษา ควรพัฒนาครูผู้สอนในประเด็นการวิเคราะห์คะแนนสอบระดับชาติ (B-NET) เพื่อกำหนดจุดเน้นของการปรับปรุงพัฒนาผลสัมฤทธิ์ทางการเรียนของผู้เรียน (23) สำหรับศูนย์การศึกษานอกระบบโรงเรียนประจำตำบล/อำเภอ/จังหวัด ควรพัฒนาครูผู้สอนในประเด็นการนำข้อสอบระดับชาติ (N-NET) ที่ผ่านมาฝึกให้ผู้เรียนได้ทำอย่างสม่ำเสมอ
เป้าหมายที่ 9 การพัฒนาสมรรถนะครูและผู้บริหารในด้านการทดสอบทางการศึกษา (จากผลการติดตามประเมินในประเด็นผลกระทบการจัดการเรียนรู้และการประเมินผลของครูผู้สอน ภาวะผู้นำทางวิชาการของผู้บริหารสถานศึกษา)	
หน่วยงานที่รับผิดชอบ	วิธีการ/แนวทางการพัฒนา
กระทรวงศึกษาธิการ	(1) ควรกำหนดนโยบายการพัฒนาครูในทุกประเภทการจัดการศึกษาให้มีความสมรรถนะด้านการวัดและประเมินผลที่มีความเชื่อมโยงระหว่างการทำข้อสอบมาตรฐานระดับชาติกับการจัดการเรียนรู้ในชั้นเรียนอย่างจริงจังและต่อเนื่อง (2) ควรฝึกอบรมและพัฒนาความรู้ความสามารถในการเขียนข้อสอบมาตรฐานให้กับครูในทุกประเภทการจัดการศึกษาอย่างทั่วถึง และมีความต่อเนื่อง (3) ควรกำหนดนโยบายส่งเสริมสนับสนุนและพัฒนาภาวะผู้นำทางวิชาการของผู้บริหารสถานศึกษาในทุกประเภทการจัดการศึกษาอย่างต่อเนื่องโดยเฉพาะประเด็นการนิเทศการจัดการเรียนรู้ของครูผู้สอน

เป้าหมายที่ 10 การสร้างความตระหนักและแรงจูงใจในการทดสอบทางการศึกษาระดับชาติ (จากผลการติดตามประเมินในประเด็นผลกระทบแรงจูงใจในการทดสอบทางการศึกษาระดับชาติ)	
หน่วยงานที่รับผิดชอบ	วิธีการ/แนวทางการพัฒนา
(1) สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน (สพฐ.)	(1) ควรมีการสร้างความรู้ความเข้าใจและสื่อสารสร้างความเข้าใจให้แก่สถานศึกษา/หน่วยงานในทุกระดับเกี่ยวกับความสำคัญของการทดสอบระดับชาติ
(2) สำนักงานคณะกรรมการการอาชีวศึกษา (สอศ.)	(2) สื่อสารสร้างความเข้าใจเกี่ยวกับรายละเอียดของการทดสอบระดับชาติ ให้แก่สถานศึกษา/หน่วยงานในทุกระดับ ทั้งในเรื่องของแผนผังการออกข้อสอบ (Test Blueprint) รูปแบบของข้อสอบและกระดาษคำตอบ การปฏิบัติตนของผู้เข้าสอบ เป็นต้น
(3) สำนักงานคณะกรรมการส่งเสริมการศึกษาเอกชน (สช.)	(3) ควรพัฒนาระบบ กลไก วิธีการ การสร้างความเข้าใจ และเจตคติที่ดีให้กับสถานศึกษา/หน่วยงานในสังกัดได้นำผลการทดสอบระดับชาติไปใช้อย่างเหมาะสมในระบบการประกันคุณภาพภายในของสถานศึกษา
(4) สำนักงานพระพุทธศาสนาแห่งชาติในส่วนงานของโรงเรียนพระปริยัติธรรมแผนกสามัญ	(4) ควรมีแนวทางในการสื่อสารกับสถานศึกษา/หน่วยงานในสังกัด และกำหนดมาตรการสร้างแรงจูงใจ การสร้างความเข้าใจ และความตระหนักในความสำคัญของการสอบระดับชาติ ให้กับผู้เรียนทุกระดับที่เข้าสอบ
(5) สำนักงานส่งเสริมการศึกษานอกระบบและการศึกษาตามอัธยาศัย (กศน.)	

อภิปรายผลการวิจัย

การติดตามและประเมินผลกระทบของการประเมินผลการทดสอบทางการศึกษาระดับชาติพบว่า มีผลกระทบทั้งในด้านบวกและด้านลบ เกิดผลกระทบในระดับน้อยไปจนถึงระดับมากที่สุด ทั้งในระดับนโยบาย หน่วยงานและบุคคล ทั้งนี้จากเป้าหมายที่กำหนดโดยสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) หรือ สทศ. (2562) ได้กำหนดไว้เพื่อ

- 1) ทดสอบความรู้และความคิดรวบยอดของผู้เรียนที่จะสำเร็จการศึกษา
- 2) นำผลการทดสอบไปใช้ประกอบการสำเร็จการศึกษา
- 3) นำผลการทดสอบไปใช้ในการพัฒนาคุณภาพการเรียนการสอนของสถานศึกษา
- 4) นำผลการทดสอบใช้ในการประเมินผลการเรียนรู้ของผู้เรียนระดับชาติ
- 5) นำผลไปใช้ในการประกันคุณภาพการศึกษา ประเด็นผลกระทบ 9 ประเด็นครอบคลุมผลกระทบระดับนโยบาย หน่วยงาน และบุคคล ดังนี้

(1) ผลกระทบต่อนโยบายมี 4 ประเด็นผลกระทบ ได้แก่ คุณภาพการศึกษา การเข้าถึงโอกาสทางการศึกษาการลดความเหลื่อมล้ำทางการศึกษา มาตรฐานของการทดสอบทางการศึกษาระดับชาติ ทั้งนี้ประเด็นผลกระทบต่อนโยบายในส่วนประเด็นผลกระทบด้านคุณภาพการศึกษามีความสอดคล้องกับนโยบายของของสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน

(สพฐ.) กระทรวงศึกษาธิการได้กำหนดให้มีการติดตามและประเมินผลการบริหารการศึกษาระดับพื้นฐานของสำนักงานเขตพื้นที่การศึกษาตามตัวชี้วัดแผนปฏิบัติการ ในปีงบประมาณ 2562 ตัวชี้วัดที่ 27 ร้อยละของผู้เรียนที่มีคะแนนผลการทดสอบทางการศึกษาระดับชาตินี้พื้นฐาน (O-NET) มากกว่าร้อยละ 50 ในแต่ละวิชาเพิ่มขึ้นจากปีการศึกษาที่ผ่านมา นอกจากนี้ในการประเมินคุณภาพภายนอกของสถานศึกษา รอบที่ 3 (พ.ศ. 2554-2558) โดยสำนักงานรับรองมาตรฐานและประเมินคุณภาพการศึกษา (องค์การมหาชน) (สมศ.) ได้ประเมินตัวบ่งชี้ที่ 5 ผลสัมฤทธิ์ทางการเรียนของผู้เรียน การประเมินคุณภาพของสถานศึกษาอาชีวศึกษา ได้ใช้ตัวบ่งชี้ที่ 2 ผู้เรียนมีความรู้และทักษะที่จำเป็นในการทำงาน โดยได้ใช้คะแนน V-NET เป็นส่วนประกอบหนึ่งในการคิดคะแนนของตัวบ่งชี้ การประเมินคุณภาพของโรงเรียนพระปริยัติธรรม ได้ใช้ตัวบ่งชี้ที่ 5 ผลสัมฤทธิ์ทางการเรียนของผู้เรียน ที่ได้ใช้คะแนน B-NET เป็นส่วนประกอบหนึ่งในการคิดคะแนนของตัวบ่งชี้ ค่าน้ำหนัก 8 คะแนน ใน 3 วิชา คือ

- (1) วิชาพุทธประวัติและธรรมวินัย
- (2) วิชาศาสนปฏิบัติ
- (3) วิชาภาษาบาลี

และใช้ผลสัมฤทธิ์ทางการเรียนของผู้เรียนตามหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน (O-NET) ค่าน้ำหนัก 12 คะแนน รวมคะแนนเท่ากับ 20 คะแนน

ในส่วนประเด็นผลกระทบคุณภาพการศึกษา การเข้าถึงโอกาสทางการศึกษา การลดความเหลื่อมล้ำทางการศึกษา มีลักษณะที่สอดคล้องกับตัวชี้วัดความเสมอภาคและเป็นธรรม และคุณภาพการศึกษา ที่ใช้ในโครงการพัฒนาระบบการติดตามและประเมินผลการจัดการศึกษาของประเทศ (ศิริชัย กาญจนวาสี และคณะ, 2552) รวมทั้งประเด็นผลกระทบมาตรฐานของการทดสอบทางการศึกษาระดับชาติ มีความสอดคล้องกับพันธกิจของสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) (รายงานประจำปีของสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน), 2561) พันธกิจที่ 1 ดำเนินการเกี่ยวกับการจัดทำระบบวิธีการทดสอบและพัฒนาเครื่องมือวัดและประเมินผลตามมาตรฐานการศึกษา สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) ได้จัดทำระบบ วิธีการทดสอบและพัฒนาเครื่องมือวัด และประเมินผลตามมาตรฐานการศึกษาในทุกระดับการศึกษา (ขั้นพื้นฐาน อาชีวศึกษา และในทุกประเภทการสอบ (ในระบบ นอกกระบบ การศึกษาตามอัธยาศัย) พันธกิจที่ 2 ดำเนินการเกี่ยวกับการประเมินผลการจัดการการศึกษาและการทดสอบทางการศึกษาระดับชาติ ตลอดจนให้ความร่วมมือและสนับสนุนการทดสอบทั้งระดับเขตพื้นที่การศึกษาและสถานศึกษา ตามมาตรฐานการทดสอบทางการศึกษาแห่งชาติ พันธกิจที่ 3 ดำเนินการเกี่ยวกับการทดสอบทางการศึกษา บริการสอบวัดความรู้ความสามารถและการสอบวัดมาตรฐานวิชาการและวิชาชีพ เพื่อนำผลไปใช้เป็นส่วนหนึ่งในการเทียบระดับและการเทียบโอนผลการเรียนที่มาจากการศึกษาในระบบเดียวกันหรือการศึกษาต่างระบบ

(2) ผลกระทบต่อหน่วยงาน มี 2 ประเด็นผลกระทบ ได้แก่ การยกระดับผลสัมฤทธิ์ การสอบทางการศึกษาระดับชาตินักเรียน และการใช้กลไกผลการทดสอบระดับชาติในการพัฒนาคุณภาพการศึกษา ซึ่งเป็นผลกระทบที่เกิดขึ้นจากการดำเนินงานโดยใช้ผลการทดสอบระดับชาติของสถานศึกษาตามนโยบายและแนวปฏิบัติของหน่วยงานต้นสังกัด ทั้งนี้เนื่องจากบทบาทสำคัญเขตพื้นที่การศึกษามีหน้าที่ในการสนับสนุนด้านการวัดและประเมินผลการเรียนรู้ ได้แก่ การสร้างความรู้ความเข้าใจ การส่งเสริมและสนับสนุนการเรียนรู้ การประเมินการอ่านคิดวิเคราะห์ และเขียน การประเมินคุณลักษณะที่พึงประสงค์ การส่งเสริมให้สถานศึกษาได้พัฒนาเครื่องมือ บริหารจัดการ วัดและประเมินอย่างเป็นระบบ ให้คำปรึกษาแนะนำเกี่ยวกับการวัดและประเมินผลการเรียนรู้ การประเมินคุณภาพผู้เรียน เป็นต้น ทั้งนี้กฎกระทรวงการประกันคุณภาพการศึกษา พ.ศ. 2561 ได้กำหนดให้การประกันคุณภาพการศึกษา เป็นการประเมินผลและการติดตามตรวจสอบคุณภาพ ตามมาตรฐานการศึกษาของสถานศึกษาและระดับและประเภทการศึกษา โดยมีกลไกในการควบคุม ตรวจสอบ ระบบการบริหารคุณภาพการศึกษาที่สถานศึกษาจัดขึ้น เพื่อให้เกิดการพัฒนาและสร้างความเชื่อมั่นให้แก่ผู้มีส่วนเกี่ยวข้องและสาธารณชนว่าสถานศึกษานั้นสามารถจัดการศึกษาได้อย่างมีคุณภาพตามมาตรฐานการศึกษา และบรรลุเป้าประสงค์ของหน่วยงานต้นสังกัดหรือหน่วยงานที่กำกับดูแล

นอกจากนี้ประเด็นผลกระทบการยกระดับผลสัมฤทธิ์การสอบทางการศึกษาระดับชาติของนักเรียนและการใช้กลไกผลการทดสอบระดับชาติในการพัฒนาคุณภาพการศึกษา ได้พบว่าประเด็นผลกระทบมีความสอดคล้องกับผลการวิจัยและข้อเสนอแนะที่ได้จากงานวิจัยของ อธิพัทธ์ สุวทันพรกุล (2556) ที่ได้วิจัยเกี่ยวกับการวิเคราะห์และสังเคราะห์กลยุทธ์การนำผลการทดสอบทางการศึกษาระดับชาตินักเรียนไปใช้ในการพัฒนาคุณภาพผู้เรียน สะการิยา แวโชะ และคณะ (2558) ที่ได้วิจัยเกี่ยวกับการนำผลการทดสอบทางการศึกษาระดับชาติด้านอิสลามศึกษา (I-NET) ไปใช้ในการพัฒนาคุณภาพผู้เรียนของโรงเรียนเอกชนสอนศาสนาอิสลามในโครงการตามพระราชดำริสมเด็จพระเทพรัตนราชสุดาฯ สยามบรมราชกุมารีสังกัดสำนักงานการศึกษาเอกชนจังหวัดนราธิวาส และ อุบล แก้วปิ่น และคณะ (2556) ได้ดำเนินการวิจัยเรื่อง แนวทางการนำผลการทดสอบทางการศึกษาไปใช้ในการพัฒนาคุณภาพผู้เรียนของโรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาจังหวัดร้อยเอ็ด

(3) ผลกระทบต่อบุคคล มี 3 ประเด็นผลกระทบ ได้แก่ ผู้เรียนมีแรงจูงใจในการสอบทางการศึกษาระดับชาติครูผู้สอนพัฒนาการจัดการเรียนรู้และการประเมินผล ผู้บริหารใช้ภาวะผู้นำทางวิชาการขับเคลื่อนพัฒนาผลสัมฤทธิ์ของผู้เรียน ทั้งนี้ประเด็นผลกระทบที่มีต่อผู้เรียนนั้นมีความสอดคล้องกับผลการวิจัยของ อุบล แก้วปิ่น และคณะ (2556) สุรัชย์ ไวยวรรณจิตร และคณะ (2557) ที่พบผลกระทบไปในทิศทางเดียวกันว่าผู้เรียนขาดประสบการณ์ในการทำข้อสอบตามแนวของ สทศ. ไม่เข้าใจวิธีระบายคำตอบและคำชี้แจงในกระดาษคำตอบ ทำให้ผู้เรียนขาดความมั่นใจในการทำข้อสอบ ผู้เรียนไม่เห็นความสำคัญของการทดสอบทางการศึกษา ไม่นำผลการทดสอบไปใช้ในการพัฒนาตนเอง ผู้เรียนบางคนมีปัญหาด้านการอ่านอ่านไม่คล่อง ทำให้ไม่มีความอดทนในการอ่านข้อความยาว ๆ เพื่อทำข้อสอบขาดทักษะการอ่านจับใจความสำคัญ การคิดวิเคราะห์

คิดสังเคราะห์ นอกจากนี้ยังพบว่า แรงจูงใจในการนำผลการทดสอบทางการศึกษาไปใช้ในการพัฒนาคุณภาพผู้เรียนของโรงเรียนเป็นตัวแปรสำคัญที่จะช่วยให้พัฒนาผลสัมฤทธิ์ได้ เช่น การมอบเกียรติบัตรโรงเรียนที่มีคะแนนสอบระดับชาติสูงเป็นการเสริมกำลังใจให้ผู้ตั้งใจทำงานได้ดี ให้ขวัญ กำลังใจ การเลื่อนขั้นพิเศษแก่ครูผู้ดูแลนักเรียนจนมีคะแนนสอบระดับชาติสูง เป็นต้น

สำหรับประเด็นผลกระทบต่อครูผู้สอนนั้นมีความสอดคล้องกับผลการวิจัยของ อุบล แก้วปิ่น และคณะ (2556) สุรัชย์ ไวยวรรณจิตร และคณะ (2557) ที่พบผลกระทบไปในทิศทางเดียวกันว่าครูผู้สอนยังไม่ปรับวิธีสอนของตนเองให้สอดคล้องกับการทดสอบระดับชาติ ทำให้ไม่สามารถจัดการเรียนการสอนให้สอดคล้องกับหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551 ซึ่งเป็นกรอบในการสอบระดับชาติครูผู้สอนมีภาระการทำงานอื่น ๆ มาก ทำให้มีเวลาเตรียมการสอนและทำการสอนในห้องเรียนน้อยลง รวมทั้งจำนวนครูมีน้อย และสอนไม่ตรงวิชาเอกทำให้คุณภาพการสอนลดลง ครูไม่เห็นความสำคัญของการสอบระดับชาติ O-NET จึงมอบภาระการสอบระดับชาติให้ครูประจำชั้นที่มีการสอบระดับชาติ ขาดการรวมกันทำงาน ครูละเลยไม่เอาใจใส่ให้นักเรียน ให้ผลการเรียนแบบปล่อยเกรด ทำให้นักเรียนไม่เห็นความสำคัญของการสอบและคะแนน ครูบางคนขาดทักษะการวัด และประเมินผลที่เน้นกระบวนการคิด ขาดการวิเคราะห์ข้อสอบ ไม่มีมาตรการยกระดับผลสัมฤทธิ์ทางการเรียนที่ชัดเจน และไม่ติดตามความเคลื่อนไหวของ สทศ. ทางเว็บไซต์ ทำให้ไม่เข้าใจข้อมูลการสอบทำให้เตรียมความพร้อมได้ไม่เต็มตามศักยภาพของผู้เรียน

ประเด็นผลกระทบต่อผู้บริหารนั้นมีความสอดคล้องกับผลการวิจัยของ อุบล แก้วปิ่น และคณะ (2556) ที่พบว่า ผู้บริหารสถานศึกษาให้ความสำคัญต่อการนำผลการทดสอบทางการศึกษาไปใช้ในการพัฒนาคุณภาพผู้เรียนค่อนข้างน้อยไม่เข้าใจการแปลค่าสถิติบางอย่างที่รายงาน และขาดการนิเทศติดตามผลอย่างต่อเนื่อง ในส่วนของงานวิจัย เสตอาหมาด ศรียาน และคณะ (2558) ได้วิจัยเกี่ยวกับการนำผลการทดสอบทางการศึกษาระดับชาติด้านอิสลามศึกษา (I-NET) ไปใช้ในการพัฒนาคุณภาพผู้เรียน และให้ข้อเสนอแนะว่า ผู้บริหารและหัวหน้าฝ่ายวิชาการ ศาสนามีความคิดเห็นว่าหน่วยงานต้นสังกัดควรนำผลการสอบมาวิเคราะห์เพื่อจัดโครงการแผนงาน และกิจกรรมในการช่วยเหลือโรงเรียนที่มีผลการสอบต่ำ ฝ่ายวิชาการศาสนาควรนำผลการสอบมาวิเคราะห์เพื่อหาจุดอ่อนและจุดแข็งของการสอนในแต่ละรายวิชา ครูผู้สอนควรจัดการเรียนการสอนให้สอดคล้องกับหลักสูตรอิสลามศึกษา พ.ศ. 2546 และควรนำผลการสอบมาใช้ในการทำวิจัยในชั้นเรียน

จากผลการศึกษาและติดตามประเมินผลกระทบของการประเมินผลการทดสอบทางการศึกษาระดับชาตินำมาสู่การกำหนดข้อเสนอแนะเชิงนโยบายที่ประกอบด้วยเป้าหมาย วิธีการและแนวทาง ครอบคลุมเป้าหมายสำคัญที่สอดคล้องกับประเด็นผลกระทบของการประเมินผลการทดสอบทางการศึกษาระดับชาติทั้ง 9 ประเด็นผลกระทบ ซึ่งแต่ละหน่วยงานที่เกี่ยวข้องสามารถนำข้อเสนอแนะเชิงนโยบายที่พัฒนาขึ้นไปใช้พิจารณาวางแผนพัฒนาให้การทดสอบระดับชาติได้สร้างผลกระทบทางบวก และลดผลกระทบทางลบเพื่อนำไปสู่ความเป็นมาตรฐานและสะท้อนคุณภาพของผู้เรียนอย่างแท้จริง

ข้อเสนอแนะในการนำผลการวิจัยไปใช้

การติดตามและประเมินผลกระทบของการประเมินผลการทดสอบทางการศึกษาระดับชาติดีพบว่า มีผลกระทบทั้งในด้านบวกและด้านลบ เกิดผลกระทบในระดับน้อยไปจนถึงระดับมากที่สุด ทั้งในระดับนโยบาย หน่วยงานและบุคคล ดังนั้นสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) หรือ สทศ. และหน่วยงานที่เกี่ยวข้องจึงควรได้พิจารณานำข้อเสนอแนะเชิงนโยบายที่ได้จากการวิจัยครั้งนี้ ไปใช้ในการวางแผนและพัฒนาตามประเด็นข้อเสนอแนะดังกล่าวต่อไป

ข้อเสนอแนะในการวิจัยครั้งต่อไป

(1) ควรพัฒนาระบบออนไลน์สำหรับการติดตามและประเมินผลกระทบที่สามารถรายงานข้อมูลตามประเด็นผลกระทบที่พัฒนาขึ้นจากงานวิจัยครั้งนี้

(2) ควรมีการศึกษาวิจัยระยะยาว (longitudinal Study) โดยใช้กลุ่มติดตาม (Cohort Study) กลุ่มผู้เรียนที่ได้รับผลกระทบจากการใช้ผลการทดสอบทางการศึกษาระดับชาติในการเข้าถึงโอกาสทางการศึกษา ความสำเร็จทางการศึกษา ความสำเร็จในการมีงานทำ รายได้ และการประกอบอาชีพ

(3) ควรมีการพัฒนาโมเดลเชิงสาเหตุผลกระทบของการประเมินผลการทดสอบทางการศึกษาระดับชาติ เพื่อศึกษาขนาดอิทธิพลของปัจจัยต่าง ๆ ที่ส่งผลกระทบต่อคะแนนสอบทางการศึกษาระดับชาติ

(4) ควรศึกษาวิจัยวิธีการประเมินผลสัมฤทธิ์ทางการเรียนของผู้เรียนโดยใช้ผลการทดสอบระดับชาติให้มีความแม่นยำ เชื่อถือได้ และมีความยุติธรรมในการนำผลไปใช้ประเมินคุณภาพของสถานศึกษา

(5) ควรวิจัยและพัฒนารูปแบบการทดสอบวัดผลสัมฤทธิ์ของผู้เรียนตามหลักสูตรการศึกษาขั้นพื้นฐานในแต่ละช่วงชั้นที่มีประสิทธิภาพและเกิดประโยชน์สูงสุดต่อผู้เรียน ครูผู้สอน และผู้บริหารสถานศึกษา

(6) ควรมีการวิจัยผลกระทบของการนำคะแนน O-NET ไปใช้เป็นองค์ประกอบในการคัดเลือกบุคคลเข้าศึกษาต่อในระดับอุดมศึกษาในสาขาวิชาและคณะต่าง ๆ และความสำเร็จในการศึกษาในมหาวิทยาลัย

(7) ควรมีการศึกษผลกระทบของการตีข้อสอบพร้อมทั้งกำหนดข้อเสนอแนะเชิงนโยบายที่จะนำไปสู่การกำหนดมาตรการและแนวทางปฏิบัติที่เหมาะสมในการเตรียมพร้อมสำหรับการทดสอบทางการศึกษาระดับชาติ

(8) ควรศึกษาความสัมพันธ์ระหว่างคะแนน V-NET กับผลการสอบทักษะทางวิชาชีพที่นักศึกษาอาชีวศึกษาแต่ละแผนกสาขาวิชาต้องสอบผ่านจึงจะสำเร็จการศึกษา เพื่อค้นหาหลักฐานความตรงเชิงเกณฑ์สัมพันธ์ว่าคะแนน V-NET เป็นความรู้ความคิดรวบยอดที่สามารถทำนายทักษะทางวิชาชีพของนักศึกษาอาชีวศึกษาได้

เอกสารอ้างอิง

- ญาณิกา ลุนราศรี และจันทร์ทรงกลด คชเสนี. (2557). ผลกระทบของการทดสอบทางการศึกษาระดับชาตินักเรียนชั้นพื้นฐานที่มีต่อการเรียนภาษาอังกฤษตามการรับรู้ของนักเรียนชั้นมัธยมศึกษาปีที่ 3. วารสารอิเล็กทรอนิกส์ทางการศึกษา (OJED). ปีที่ 9 ฉบับที่ 1 มกราคม - มีนาคม.
- ศิริชัย กาญจนวาสี และคณะ. (2552). รายงานการออกแบบและพัฒนาระบบข้อมูลและระบบการประมวลผลข้อมูลสำหรับการประเมินผลการจัดการศึกษาของประเทศ. กรุงเทพฯ: สำนักประเมินผลการศึกษา สำนักงานเลขาธิการสภาการศึกษา กระทรวงศึกษาธิการ.
- สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน). (2562). ข้อมูลการจัดสอบเบื้องต้น. สืบค้นจาก <http://www.niets.or.th/th/catalog/view/2950> เมื่อวันที่ 9 กุมภาพันธ์ 2562
- สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน). (2562). รายงานประจำปี 2560. สืบค้นจาก <http://www.niets.or.th/> เมื่อวันที่ 9 กุมภาพันธ์ 2562
- สุรัชย์ ไวยวรรณจิตร และคณะ. (2557). การนำผลการทดสอบทางการศึกษาแห่งชาติ (O-NET) ระดับชั้นมัธยมศึกษาปีที่ 3 และปีที่ 6 ไปใช้ในการพัฒนาคุณภาพผู้เรียน: กรณีศึกษาโรงเรียนเอกชนสอนศาสนาอิสลามในพื้นที่จังหวัดชายแดนภาคใต้. กรุงเทพฯ: สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน)
- สัมพันธ์ พันธุ์พฤกษ์. (2561). การทดสอบทางการศึกษาระดับชาตินักเรียนและอนาคตของไทย. เอกสารประกอบการประชุมสัมมนาวิชาการ การวัดผล ประเมินผล และวิจัยสัมพันธ์ แห่งประเทศไทย ครั้งที่ 26. โรงพิมพ์มหาวิทยาลัยขอนแก่น.
- สะการิยา แวโซะ และคณะ. (2558). การนำผลการทดสอบทางการศึกษาระดับชาติด้านอิสลามศึกษา (I-NET) ไปใช้ในการพัฒนาคุณภาพผู้เรียนของโรงเรียนเอกชนสอนศาสนาอิสลามในโครงการตามพระราชดำริสมเด็จพระเทพรัตนราชสุดาฯ สยามบรมราชกุมารีสังกัดสำนักงานการศึกษาเอกชนจังหวัดนครราชสีมา. กรุงเทพฯ: สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน)
- เสดอาหมาด ศรียาน และคณะ. (2558). การนำผลการทดสอบทางการศึกษาระดับชาติด้านอิสลามศึกษา (I-NET) ไปใช้ในการพัฒนาคุณภาพผู้เรียน : กรณีศึกษาโรงเรียนเอกชนสอนศาสนาอิสลามในจังหวัดสตูล. กรุงเทพฯ: สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน)
- สำนักงานรับรองมาตรฐานและประเมินคุณภาพการศึกษา. (2562). ตัวบ่งชี้และเกณฑ์การประเมินคุณภาพภายนอก รอบสาม (พ.ศ.2554 - 2558) ระดับการศึกษาขั้นพื้นฐาน. สืบค้นจาก www.onesqa.or.th
- สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน. (2554ก). การติดตามตรวจสอบคุณภาพการศึกษา. กรุงเทพฯ: โรงพิมพ์ชุมนุมสหกรณ์การเกษตรแห่งประเทศไทย.

- อิทธิพัทธ์ สุวทันพรกุล. (2556). การวิเคราะห์และสังเคราะห์กลยุทธ์การนำผลการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐานไปใช้ในการพัฒนาคุณภาพผู้เรียน: พหุกรณีศึกษา โรงเรียนที่มีผลการทดสอบทางการศึกษาในระดับดี. กรุงเทพฯ: สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน)
- อุบล แก้วปิ่น และคณะ. (2556). แนวทางการนำผลการทดสอบทางการศึกษาไปใช้ในการพัฒนาคุณภาพผู้เรียนของโรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาจังหวัดร้อยเอ็ด. กรุงเทพฯ: สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน)
- Bamberger, M. (2012). Introduction to Mixed Methods in Impact Evaluation. Impact Evaluation Notes. No.3 August 2012. Available from <https://www.interaction.org>
- Creswell, J. W. (2013). Research Design: Qualitative, Quantitative, and Mixed Methods Approaches. (2nd edition). Thousand Oaks, CA: Sage publications.
- Kusek, J. Z. and Rist, R. C. (2004). Ten Steps to a Results-Based Monitoring and Evaluation System: a Handbook for Development Practitioners. Washington, DC: World Bank.
- Rogers, P. (2014). Overview of Impact Evaluation, Methodological Briefs: Impact Evaluation1, UNICEF Office of Research, Florence. Available from https://www.unicef-irc.org/KM/IE/impact_1.php
- Stufflebeam, D. L., & Coryn, C. L. (2014). Evaluation Theory, Models, and Applications (2nd Edition). John Wiley & Sons.
- The World Bank. (2004). Monitoring & Evaluation: Some Tools, Methods & Approaches. Washington, D.C.: The International Bank for Reconstruction and Development/ The World Bank.

การสังเคราะห์งานวิจัยทางการศึกษาและการวิเคราะห์ห่อถิมาน

Synthesis of Educational Research and Meta Analysis

นันทิมาน นาคาพงศ์ อัครวรัท

บทคัดย่อ

การสังเคราะห์งานวิจัยทางการศึกษาเป็นระเบียบวิธีที่ใช้ในการศึกษาหาข้อเท็จจริงของงานวิจัยทางการศึกษาจำนวนมากที่ศึกษาปัญหาวิจัยเดียวกัน มีจุดมุ่งหมายเพื่อหาข้อสรุปความคล้ายคลึงและความแตกต่างจากงานวิจัย โดยนำคุณลักษณะงานวิจัยมาเป็นตัวแปรในการอธิบายความแตกต่างระหว่างงานวิจัยทำการวิเคราะห์ข้อมูลด้วยวิธีการทางสถิติและ/หรือการวิเคราะห์เนื้อหา การวิเคราะห์ห่อถิมานเป็นการสังเคราะห์งานวิจัยเชิงปริมาณที่มีระเบียบวิธีเป็นงานวิจัยเชิงสำรวจ มุ่งศึกษาหาข้อสรุปของผลการวิจัยจำนวนมากที่ศึกษาปัญหาเดียวกันด้วยการนำวิธีการทางสถิติมาวิเคราะห์ข้อมูล จุดมุ่งหมายมี 2 ประการ คือ เพื่อสังเคราะห์ให้ได้ข้อสรุปเกี่ยวกับดัชนีมาตรฐาน และเพื่อตรวจสอบความสัมพันธ์เชิงสาเหตุระหว่างดัชนีมาตรฐานกับตัวแปรปรับ ซึ่งก็คือตัวแปรคุณลักษณะงานวิจัย ข้อมูลสำหรับการวิเคราะห์ ได้แก่ ข้อมูลคุณลักษณะงานวิจัย และดัชนีมาตรฐานก็คือดัชนีขนาดอิทธิพลกับดัชนีสัมประสิทธิ์สหสัมพันธ์ หน่วยการวิเคราะห์ ได้แก่ งานวิจัยและ/หรือชุดการทดสอบสมมติฐาน เครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูล ได้แก่ แบบประเมินคุณภาพงานวิจัย และแบบบันทึกคุณลักษณะงานวิจัย

คำสำคัญ: สังเคราะห์งานวิจัยทางการศึกษา วิเคราะห์ห่อถิมาน

Abstract

Synthesis of educational research refers to a methodology used in the study of the facts of many educational research study which studies the same research problem. The objective is to find conclusions about the similarities and differences from the research by using research characteristics as variables to explain the differences between researches. The data is analysed by using statistical methods and/or content analysis. Meta analysis is quantitative research synthesis which has a methodology as survey research focus on finding the conclusion of many research findings that study the same problem by applying statistical methods to analyze data. Meta-analysis has two objectives: to synthesize and get a conclusion about the standard index and to examine the causal relationship between the standard index and moderator variables which is the research characteristic variables. Data for analysis included data of research characteristics and standard index such as effect size index and correlation coefficient index. Analytical units are research and/or a set of hypothesis tests. The research instruments used for data collection were research quality evaluation form and research characteristic record form.

Keywords: Synthesis of Educational Research, Meta Analysis

บทนำ

ประเทศไทยเริ่มทำการวิจัยทางการศึกษามาอย่างยาวนานหลายทศวรรษ ทำให้ยุคปัจจุบันมีข้อมูลงานวิจัยทางการศึกษาจำนวนมาก งานวิจัยทางการศึกษาที่ศึกษาในขอบข่ายของปัญหาวิจัยที่คล้ายคลึงกันหรือหมวดหมู่เดียวกันนั้น ผลการวิจัยที่ได้อาจมีความแตกต่างกันบ้างหรือคล้ายคลึงกันบ้างขึ้นอยู่กับระเบียบวิธีการที่ผู้วิจัยได้ออกแบบไว้ ถ้ามีการรวบรวมข้อมูลจากผลงานวิจัยทางการศึกษาจำนวนมากเหล่านี้มาทำการสังเคราะห์งานวิจัย ก็จะได้ข้อสรุปผลเป็นองค์ความรู้ใหม่ที่กว้างกว่า ลึกซึ้งกว่าและมีเนื้อหาครอบคลุมทุกกลุ่มตัวอย่างงานวิจัยเดิมแต่ละเรื่อง

การสังเคราะห์งานวิจัยจะใช้ระเบียบวิธีการทางวิทยาศาสตร์ในการแสวงหาความรู้ งานวิจัยที่นำมาสังเคราะห์แต่ละเรื่องต้องเป็นงานวิจัยที่ศึกษาปัญหาวิจัยเดียวกันและให้ข้อค้นพบในแง่มุมที่คล้ายคลึงกันหรือแตกต่างกันภายใต้ปรากฏการณ์ที่ศึกษา การนำผลงานวิจัยแต่ละเรื่องมาทำการสังเคราะห์จะได้ภาพรวมของปรากฏการณ์ที่ศึกษาที่ทำให้เกิดความเข้าใจในปรากฏการณ์นั้นอย่างลุ่มลึกมากกว่าความเข้าใจในงานวิจัยแต่ละเรื่อง สำหรับการสังเคราะห์งานวิจัยเชิงปริมาณรูปแบบหนึ่งที่มีความนิยมใช้ในปัจจุบันก็คือ “การวิเคราะห์ถ้อยแถลง” เป็นการสังเคราะห์ข้อมูล 2 ส่วน คือ ข้อมูลผลงานวิจัยแต่ละเรื่องที่วัดในรูปดัชนีมาตรฐานและข้อมูลคุณลักษณะงานวิจัย โดยใช้กระบวนการทางสถิติวิเคราะห์ข้อมูลโดยมีเป้าหมายหลัก 3 ประเด็น คือ เพื่อการบรรยาย (description) ว่าดัชนีมาตรฐานที่เป็นผลงานวิจัยที่นำมาสังเคราะห์ทุกเรื่องนั้นมีค่าใกล้เคียงกันหรือไม่ ถ้ามีค่าใกล้เคียงกันแล้วจะสังเคราะห์สรุปผลการวิจัยได้อย่างไร เพื่อวิเคราะห์แยกส่วน (decomposition) ว่าถ้าดัชนีมาตรฐานที่เป็นผลงานวิจัยที่นำมาสังเคราะห์ทุกเรื่องนั้นมีค่าแตกต่างกันแล้ว จะมีปริมาณมากน้อยเท่าไร จะเป็นความแปรปรวนชนิดมีระบบหรือเป็นความแปรปรวนจากความคลาดเคลื่อน (systematic or error variance) และเพื่ออธิบาย (explanation) ว่าปัจจัยใดอธิบายความแปรปรวนในดัชนีมาตรฐานได้บ้าง หรือเป็นการวิเคราะห์ว่ามีตัวแปรปรับ (moderator variables) ตัวใดบ้างที่มีอิทธิพลทำให้ดัชนีมาตรฐานนั้นมีค่าแตกต่างกัน การวิเคราะห์ถ้อยแถลงจะทำให้ได้ผลการวิจัยที่สำคัญสามประเด็นหลักข้างต้น ด้วยเหตุนี้การสังเคราะห์งานวิจัยโดยใช้วิธีการวิเคราะห์ถ้อยแถลงจึงเป็นระเบียบวิธีการที่มีประโยชน์อย่างยิ่งในยุคปัจจุบันที่เต็มไปด้วยข้อมูลงานวิจัยทางการศึกษาเป็นจำนวนมาก

บทความนี้มีขอบเขตเนื้อหาเกี่ยวกับประวัติความเป็นมาของการสังเคราะห์งานวิจัย ความหมายและขั้นตอนของการสังเคราะห์งานวิจัยทางการศึกษา รวมถึงประวัติความเป็นมา ความหมาย ขั้นตอน วิธีการและลักษณะของการวิเคราะห์ถ้อยแถลง ผู้เขียนได้เรียบเรียงจากการศึกษาสังเคราะห์แนวคิดทฤษฎีและจากประสบการณ์ตรงในการสังเคราะห์งานวิจัยทางการศึกษาโดยใช้วิธีการวิเคราะห์ถ้อยแถลงหวังเป็นอย่างยิ่งว่าผู้อ่านที่มีความสนใจในการสังเคราะห์งานวิจัยทางการศึกษาและการวิเคราะห์ถ้อยแถลงจะได้รับความรู้พื้นฐานเบื้องต้นอย่างเข้าใจ ถ่องแท้ก่อนที่จะลงมือปฏิบัติการวิจัยจริงได้อย่างมีประสิทธิภาพและเกิดองค์ความรู้ใหม่สามารถนำไปใช้ประโยชน์อย่างกว้างขวางในวงวิชาการศึกษา

ประวัติความเป็นมาของการสังเคราะห์งานวิจัย

Glass, McGaw, and Smith (1981) Hunter, Schmidt, and Jackson (1982), Hedges and Olkin (1985), Kulik and Kulik (1989) อธิบายไว้ว่า การสังเคราะห์งานวิจัยแบบดั้งเดิมที่ใช้กันมาในอดีตและยังใช้อยู่ในปัจจุบัน คือ การสังเคราะห์งานวิจัยแบบพรรณนา (narrative method) เป็นการอ่านและสรุปสาระตามความเข้าใจของนักสังเคราะห์งานวิจัยมีจุดอ่อนตรงที่เป็นวิธีการที่ไม่มีระบบ มีความเป็นอัตนัยสูง ให้ผลการสังเคราะห์ที่ขาดความชัดเจน ไม่สามารถทำซ้ำโดยให้ผลเหมือนเดิมได้และใช้ได้ดี สำหรับการสังเคราะห์งานวิจัยที่มีจำนวนไม่มากนัก

จากจุดอ่อนของการสังเคราะห์งานวิจัยแบบเดิมทำให้นักวิจัยพยายามพัฒนาวิธีการสังเคราะห์งานวิจัยให้เป็นระบบมากขึ้น ได้แก่ การนำวิธีการทางสถิติไปใช้ในการวิเคราะห์ข้อมูลเพื่อสังเคราะห์งานวิจัย วิธีการสังเคราะห์งานวิจัยเชิงปริมาณในระยะแรก ได้แก่ วิธีนับคะแนนเสียง (vote-counting method) วิธีการประมาณค่าขนาดอิทธิพล (effect size estimation) และวิธีรวมค่าความน่าจะเป็น (probability accumulation) การสังเคราะห์งานวิจัยเชิงปริมาณในระยะแรกให้ความสำคัญกับผลการทดสอบสมมติฐาน จากนั้นจึงพยายามแก้จุดอ่อนของผลการทดสอบสมมติฐานด้วยการนำระดับนัยสำคัญทางสถิติหรือค่าความน่าจะเป็นมาใช้ในการสังเคราะห์ ในระยะหลังนักวิจัยตระหนักว่าข้อค้นพบสำคัญของงานวิจัย คือ ขนาดอิทธิพล (effect size) อันเป็นค่าสถิติที่เป็นดัชนีมาตรฐาน ซึ่งมีสารสนเทศเกี่ยวกับผลการวิจัยที่สมบูรณ์ จึงได้พัฒนาวิธีการสังเคราะห์งานวิจัยโดยการรวมค่าขนาดอิทธิพลได้เป็นวิธีการอันเป็นต้นกำเนิดของการวิเคราะห่อภิมาณ ซึ่งเป็นวิธีการที่ใช้กันมากในปัจจุบัน

นอกจากนี้ Noblit and Hare (1988) ได้นำแนวคิดการวิเคราะห่อภิมาณไปสังเคราะห์งานวิจัยชาติพันธุ์วรรณาและเรียกชื่อการวิจัยประเภทนี้ว่า การวิจัยชาติพันธุ์วรรณาอภิมาณ (meta-ethnography research) ซึ่งจัดว่าเป็นนวัตกรรมทางการวิจัยเชิงคุณภาพแบบหนึ่งที่มีการประยุกต์หลักการวิเคราะห่อภิมาณให้เป็นประโยชน์ในการวิจัยชาติพันธุ์วรรณา (ethnographic research or ethnography) (นงลักษณ์ วิรัชชัย, 2542)

ความหมายของการสังเคราะห์งานวิจัยทางการศึกษา

Harries (2019) ให้ความหมายการสังเคราะห์งานวิจัยว่าเป็นการกลั่นกรองและรวบรวมข้อมูลทางการศึกษาอย่างเป็นระบบ เพื่อให้ได้ข้อสรุปที่น่าเชื่อถือมากขึ้นเกี่ยวกับประเด็นการวิจัยที่กำหนด

อุทุมพร จามรมาร (2527) ให้ความหมายการสังเคราะห์ว่าเป็นการนำหน่วยย่อย ๆ หรือ ส่วนต่าง ๆ มาประกอบให้เป็นเนื้อเรื่องเดียวกัน โดยไม่เคยนำสิ่งต่าง ๆ เหล่านี้มารวมเข้าด้วยกัน มาก่อน การสังเคราะห์จำแนกได้เป็น 3 ประเภท คือ

- 1) การสังเคราะห์จนได้ผลผลิตโดยอาศัยหน่วยย่อย ๆ ของการสื่อความหมาย
- 2) การสังเคราะห์จนได้ผลผลิตจากการวางแผน
- 3) การสังเคราะห์สิ่งที่เป็นนามธรรม นอกจากนี้การสังเคราะห์ยังจำแนกได้อีกแบบ คือ
 - (1) การสังเคราะห์เชิงบอกลักษณะ (qualitative synthesis)
 - (2) การสังเคราะห์เชิงปริมาณ (quantitative synthesis)

นงลักษณ์ วิรัชชัย (2529) ให้ความหมายของการสังเคราะห์งานวิจัยว่าเป็นระเบียบวิธีการศึกษาหาข้อเท็จจริง เพื่อตอบปัญหาใดปัญหาหนึ่งโดยการรวบรวมงานวิจัยที่เกี่ยวข้องกับการศึกษาในเรื่องเดียวกันหลายเรื่องมาวิเคราะห์ด้วยวิธีการทางสถิติหรือวิธีการวิเคราะห์เนื้อหา ให้ได้เป็นข้อสรุปของผลการวิจัยเกี่ยวกับเรื่องที่ต้องสังเคราะห์

ศิริภา พูลสุวรรณ (2555) ให้ความหมายการสังเคราะห์งานวิจัยว่าเป็นระเบียบวิธีการวิจัยแบบหนึ่งที่ศึกษาสังเคราะห์รายงานการวิจัยที่ศึกษาปัญหาเดียวกัน โดยมีเป้าหมายเพื่อหาข้อสรุปจากผลวิจัย และอธิบายความแตกต่างระหว่างงานวิจัยโดยการนำคุณลักษณะงานวิจัย มาเป็นตัวแปรในการอธิบาย สามารถใช้วิธีการสังเคราะห์เชิงคุณลักษณะโดยการพรรณนาสรุป และการสังเคราะห์เชิงปริมาณโดยใช้สถิติในการวิเคราะห์ผลวิจัยจากงานวิจัยแต่ละเรื่องที่ปรับ ให้เป็นดัชนีมาตรฐาน

สรุปได้ว่าการสังเคราะห์งานวิจัยทางการศึกษา หมายถึง ระเบียบวิธีที่ใช้ในการศึกษาหาข้อเท็จจริงของงานวิจัยทางการศึกษาจำนวนมากที่ศึกษาปัญหาวิจัยเดียวกัน โดยการวิเคราะห์ ข้อมูลด้วยวิธีการทางสถิติหรือการวิเคราะห์เนื้อหา มีเป้าหมายเพื่อหาข้อสรุปความคล้ายคลึงและความแตกต่างจากงานวิจัย โดยนำคุณลักษณะงานวิจัยมาเป็นตัวแปรในการอธิบายความแตกต่างระหว่างงานวิจัย ทำให้ได้องค์ความรู้ใหม่ทางการศึกษาที่นำไปใช้ประโยชน์อย่างกว้างขวางในวงวิชาการศึกษา

ขั้นตอนการสังเคราะห์งานวิจัยทางการศึกษา

การสังเคราะห์งานวิจัยทางการศึกษามีขั้นตอนการดำเนินงาน 5 ขั้นตอน เช่นเดียวกับขั้นตอนในการทำวิจัยทั่วไป มีรายละเอียดดังนี้

ขั้นตอนที่ 1 การกำหนดหัวข้อปัญหา การวิเคราะห์งานวิจัยเริ่มต้นจากการกำหนดปัญหาการวิจัยทางการศึกษา ซึ่งต้องเป็นปัญหาที่มีการทำวิจัยอย่างน้อยสองราย นักวิจัยมักสนใจและทำการวิจัยกับปัญหาที่มีคุณค่าและเป็นปัญหาที่ยังไม่มีคำตอบแน่ชัด การสังเคราะห์งานวิจัยจะทำได้ดีต่อเมื่อมีรายงานการวิจัยในปัญหาทางการศึกษาที่ใกล้เคียงกันและมีหลายเรื่องที่ให้ผลแตกต่างกัน จึงเหมาะสมที่จะทำการสังเคราะห์งานวิจัยทางการศึกษา Kumar (2005) เสนอแนวทางการพิจารณากำหนดปัญหาการวิจัยให้คำนึงถึงหลัก 7 ประการ ได้แก่

- 1) นักวิจัยมีความสนใจในหัวข้อปัญหาวิจัย
- 2) ขอบเขตหรือขนาดของปัญหาวิจัยมีความเฉพาะเจาะจงและชัดเจน
- 3) การวัดตัวแปรมีความชัดเจนเกี่ยวกับตัวชี้วัดและสามารถสร้างเครื่องมือวัดได้
- 4) นักวิจัยมีความเชี่ยวชาญในประเด็นปัญหาวิจัย
- 5) ประเด็นปัญหาการวิจัยมีความสัมพันธ์กับงานที่ทำ
- 6) ข้อมูลมีเพียงพอและจัดหาได้
- 7) มีจริยธรรมเกี่ยวกับข้อมูลที่ได้จากกลุ่มตัวอย่างโดยเฉพาะกลุ่มด้อยสิทธิ เด็ก สตรี

คนชรา

ขั้นตอนที่ 2 การวิเคราะห์ปัญหา เมื่อกำหนดหัวข้อปัญหาทางการศึกษาแล้ว การสังเคราะห์งานวิจัยทางการศึกษามีแนวทางการพิจารณาปัญหาอยู่ 3 ประเด็น ได้แก่

- 1) ปัญหาความขัดแย้งของผลการวิจัยที่ศึกษาปัญหาเดียวกัน
 - 2) ปัญหาความต้องการผลสรุปสุดท้ายจากงานวิจัยที่มีจำนวนมาก
 - 3) ปัญหาความต้องการพัฒนานวัตกรรมโดยการต่อยอดองค์ความรู้จากงานวิจัยที่มีอยู่เดิม
- นักวิจัยจึงต้องนิยามปัญหาทางการศึกษาให้ชัดเจน ศึกษาแนวคิด หลักการและทฤษฎีที่เกี่ยวข้อง เพื่อกำหนดระเบียบแบบแผนและสมมติฐานการวิจัย

ขั้นตอนที่ 3 การเสาะค้น คัดเลือก และรวบรวมงานวิจัย ที่ประกอบด้วยการดำเนินงาน 3 ขั้นตอนย่อย ดังนี้

3.1 การเสาะค้นหางานวิจัย นักวิจัยต้องค้นคว้าและเสาะแสวงหางานวิจัยทั้งหมดเกี่ยวกับปัญหาที่กำหนดไว้ การเสาะค้นงานวิจัยส่วนใหญ่จะหาได้จากเอกสาร เช่น รายงานการวิจัย วิทยานิพนธ์ บทความย่อ วิทยานิพนธ์ วารสาร ดัชนีค้นวารสารจากศูนย์ทรัพยากรข้อมูลทางการศึกษา เป็นต้น

3.2 การคัดเลือกงานวิจัย นักวิจัยต้องอ่าน ศึกษาและตรวจสอบงานวิจัยแต่ละเรื่องอย่างละเอียด ต้องสร้างเกณฑ์ในการคัดเลือกงานวิจัยให้ได้งานวิจัยที่มีคุณภาพดี มีความตรงภายในและความตรงภายนอกสูงตามเกณฑ์ที่กำหนดไว้

3.3 การรวบรวมผลงานวิจัย อาจใช้การจดบันทึก การถ่ายเอกสารหรือการกรอกรูปแบบฟอร์มก็ได้ ทั้งนี้ นักวิจัยต้องใช้ความระมัดระวังให้ได้ข้อมูลที่เที่ยงตรง เชื่อถือได้ และครบถ้วนสมบูรณ์

ขั้นตอนที่ 4 การวิเคราะห์เพื่อสังเคราะห์ผลการวิจัย ขั้นนี้ต้องเน้นการจัดกระทำและการวิเคราะห์ข้อมูล ซึ่งประกอบด้วยผลการวิจัย รายละเอียด ลักษณะและวิธีการวิจัยจากงานวิจัยทางการศึกษาทั้งหมด เพื่อสังเคราะห์หาข้อสรุปที่เป็นข้อยุติและทดสอบว่าสอดคล้องตามสมมติฐานที่ตั้งไว้หรือไม่ จากนั้นจึงแปลความหมายผลการวิเคราะห์เพื่อตอบปัญหาการวิจัย

ขั้นตอนที่ 5 การเสนอรายงานผลการสังเคราะห์การวิจัย การเขียนรายงานการสังเคราะห์งานวิจัยทางการศึกษา มีหลักการเช่นเดียวกับการเขียนรายงานการวิจัยโดยทั่วไป นำเสนอรายละเอียดวิธีการดำเนินงานทุกขั้นตอน พร้อมทั้งข้อสรุปข้อค้นพบและข้อเสนอแนะจากการสังเคราะห์งานวิจัย โดยใช้ภาษาให้ถูกต้องตามหลักวิชาการ อ่านเข้าใจได้ง่ายและชัดเจน

ประวัติความเป็นมาของการวิเคราะห์อภิมาน

วิธีวิเคราะห์งานวิจัยโดยการวิเคราะห์อภิมานเริ่มใช้ในสาขาวิทยาศาสตร์การเกษตรเป็นครั้งแรกในช่วงต้นคริสต์วรรษที่ 19 ต่อมาในปี ค.ศ. 1933 จึงเริ่มเข้ามาใช้ในสาขาวิชาสังคมศาสตร์ นักวิจัยตระหนักถึงความจำเป็นในการสังเคราะห์งานวิจัยซึ่งมีอยู่เป็นจำนวนมากให้ได้ข้อสรุปสุดท้ายเป็นองค์ความรู้ในศาสตร์แต่ละสาขา จึงได้พัฒนาวิธีวิทยาการสำหรับการสังเคราะห์งานวิจัยให้ดียิ่งขึ้น เทคนิคการวิเคราะห์อภิมานนับเป็นเทคนิคการสังเคราะห์เชิงปริมาณที่ได้รับการพัฒนาและแพร่หลายในปัจจุบันในต่างประเทศ (นงลักษณ์ วิรัชชัย และ สุวิมล ว่องวาณิช, 2542)

การสังเคราะห์งานวิจัยเชิงปริมาณได้เริ่มต้นมาตั้งแต่ช่วงต้นศตวรรษที่ 20 โดยในระยะเริ่มแรกจะมีจุดเน้นอยู่ที่การสังเคราะห์ผลการวิจัยจากงานวิจัยหลายเรื่องเข้าด้วยกัน ต่อมาเมื่อพบว่า การสังเคราะห์งานวิจัยมีข้อจำกัดบางประการอันเนื่องมาจากกลุ่มตัวอย่างที่ใช้ในการวิจัยทำให้การสรุปผลได้ภาพรวมที่แท้จริง โดยในระยะแรก Thorndike เป็นบุคคลแรกที่ใช้ในการวิเคราะห์อภิมานในการสังเคราะห์งานวิจัยประเภทสหสัมพันธ์ โดยการหาค่าเฉลี่ยของสัมประสิทธิ์สหสัมพันธ์ ต่อมาเมื่อปี ค.ศ. 1969 Cohen ได้เสนอการคำนวณค่าขนาดอิทธิพล (effect sizes) เพื่อใช้ในการกำหนดขนาดของกลุ่มตัวอย่างและการประมาณค่าอำนาจ (power) ต่อมาในปี ค.ศ. 1976 นับเป็นปีแรกที่มีการเริ่มต้นใช้หลักการวิเคราะห์อภิมานโดย Glass เป็นบุคคลที่เริ่มใช้ค่าขนาดอิทธิพลและสัมประสิทธิ์สหพันธ์เป็นดัชนีมาตรฐานในการสังเคราะห์งานวิจัย และบัญญัติศัพท์การวิเคราะห์อภิมาน (meta-analysis) ขึ้นโดยให้ความหมายว่าเป็นการวิเคราะห์ผลการวิเคราะห์ (analysis of analysis) ต่อจากนั้นการวิเคราะห์อภิมานได้รับการพัฒนาให้ดียิ่งขึ้นโดยบุคคลหลายท่าน เช่น Hunter, Schmidt, Rosenthal, and Hedges เป็นต้น

ในประเทศไทยมีนักวิชาการหลายท่านได้ให้ความหมายของคำว่า meta-analysis เช่น อุทุมพร จามรมาน (2527) ใช้คำว่า การวิเคราะห์เมตต้า, สุวัฒนา สุวรรณเขตนิคม (2529) ใช้คำว่า การวิเคราะห์เมตต้า, นงลักษณ์ วิรัชชัย (2529) ใช้คำว่า การวิเคราะห์อภิมาน, สุรศักดิ์ หลาบมาลา (2532) ใช้คำว่า การวิเคราะห์ผลรวม, สุพัฒน์ สุขมลสันต์ (2532) ใช้คำว่า การอภิวเคราะห์,

คุชฎี โยเหลา (2535) ใช้คำว่า การวิเคราะห์เมต้า แต่คำที่ได้รับการนิยมใช้ในปัจจุบัน คือ การวิเคราะห์อภิมาน ของนงลักษณ์ วิรัชชัย (2529) ตามข้อเสนอแนะของ ปรีชา ธรรมมา ผู้บัญญัติศัพท์ทางวิชาการเป็นที่นิยม เนื่องจากคำว่า “อภิ” หมายถึง เหนือกว่า ยิ่งกว่า ใหญ่กว่า คำว่า “มาน” หมายถึง การวัด ทาง เชิง หรือ มี เมื่อรวมคำศัพท์ว่า การวิเคราะห์อภิมานจึงหมายถึงการวิเคราะห์ที่มีการวัดหรือแบบที่เหนือกว่า สมบูรณ์กว่าการวิเคราะห์ทั่วไป ซึ่งตามความหมายนี้ การวิเคราะห์อภิมานมีความหมายตรงกับคำศัพท์ว่า meta-analysis ในความหมาย analysis of analysis ของ Glass (นงลักษณ์ วิรัชชัย, 2542)

ความหมายของการวิเคราะห์อภิมาน

Glass (1976) ให้ความหมายการวิเคราะห์อภิมานว่าเป็นการวิเคราะห์ผลการวิเคราะห์ (the analysis of analysis) ที่มีการนำงานวิจัยมาใช้เป็นแหล่งข้อมูลในการวิเคราะห์ จึงเป็นการวิเคราะห์ที่เหนือกว่าลึกซึ้งกว่าการวิเคราะห์ในงานวิจัยทั่วไป ถ้าเป็นการวิเคราะห์ข้อมูลจากแหล่งปฐมภูมิ เรียกว่า การวิเคราะห์ปฐมภูมิ (primary analysis) และถ้าเป็นการวิเคราะห์ข้อมูลจากแหล่งทุติยภูมิ เรียกว่า การวิเคราะห์ทุติยภูมิ (secondary analysis)

Glass, McGaw, and Smith (1981) ให้ความหมายการวิเคราะห์อภิมานว่าเป็นการวิจัยเชิงปริมาณเพื่อสังเคราะห์งานวิจัยหลาย ๆ เรื่องที่ศึกษาปัญหาวิจัยเดียวกันโดยใช้วิธีการทางสถิติ ซึ่งข้อมูลที่ใช้ในการวิเคราะห์อภิมานประกอบด้วยผลการวิจัยที่วัดในรูปขนาดอิทธิพล และคุณลักษณะของงานวิจัย

Hunter, Schmidt and Jackson (1982) และ Glass, McGraw, and Smith (1981) ให้ความหมายการสังเคราะห์งานวิจัยด้วยการวิเคราะห์อภิมานว่าเป็นการวิจัยเชิงปริมาณเพื่อสังเคราะห์งานวิจัยจำนวนมากที่ศึกษาปัญหาวิจัยเดียวกัน โดยใช้วิธีการทางสถิติเพื่อสังเคราะห์ให้ได้ข้อสรุปที่มีความกว้างขวางลุ่มลึกกว่าผลงานวิจัยแต่ละเรื่อง ข้อมูลสำคัญสำหรับการวิเคราะห์อภิมาน คือ ดัชนีมาตรฐาน ได้แก่ ดัชนีขนาดอิทธิพล (effect size) และดัชนีสัมประสิทธิ์สหสัมพันธ์ และข้อมูลคุณลักษณะงานวิจัย หน่วยการวิเคราะห์ คือ งานวิจัยและ/หรือชุดการทดสอบสมมติฐาน จุดมุ่งหมายในการวิเคราะห์แยกได้เป็น 2 ประการ ประการแรก คือ การสังเคราะห์ให้ได้ข้อสรุปเกี่ยวกับดัชนีมาตรฐาน ประการที่สองคือ การวิเคราะห์เพื่อตรวจสอบความสัมพันธ์เชิงสาเหตุระหว่างดัชนีมาตรฐานกับตัวแปรปรับ (moderator variables) ซึ่งได้แก่ ตัวแปรคุณลักษณะงานวิจัย

รัตนะ บัวสนธ์ (2534) ให้ความหมายการวิเคราะห์อภิมานเป็นการสังเคราะห์งานวิจัยเชิงปริมาณ (Synthesis of Quantitative Research) ประเภทหนึ่ง นอกเหนือไปจากการสังเคราะห์งานวิจัยเชิงปริมาณด้วยวิธีอื่น ๆ การวิเคราะห์อภิมานจะใช้วิธีการที่มีระบบใช้หลักการและระเบียบวิธีการทางสถิติเพื่อหาคำตอบสรุปให้กับงานวิจัยเชิงปริมาณที่มีผู้ทำการวิจัยหลายคนในเรื่องเดียวกัน มีลักษณะปัญหาการวิจัยคล้ายกัน แต่อาจจะใช้กลุ่มตัวอย่าง

ต่างกัน วิธีการทางสถิติที่ใช้วิเคราะห์ต่างกัน หรือตัวแปรบางตัวต่างกัน ซึ่งผลการวิจัยต่าง ๆ เหล่านี้ บางครั้งอาจจะให้คำตอบสอดคล้องกันหรือขัดแย้งจนยากแก่การสรุปว่าคำตอบที่แท้จริงเป็นอย่างไรกันแน่ การวิเคราะห์ห่อภิมาณได้รับการพัฒนาขึ้นมาเพื่อแก้ปัญหาดังกล่าว เมื่อพิจารณาจากความหมายและจุดมุ่งหมายของการวิเคราะห์ห่อภิมาณแล้วก็จะเห็นว่า การวิเคราะห์ห่อภิมาณมีลักษณะระเบียบวิธีเป็นงานวิจัยเชิงสำรวจ (Survey Research) นั่นเอง

นงลักษณ์ วิรัชชัย (2542) ให้ความหมายการวิเคราะห์ห่อภิมาณว่าเป็นการสังเคราะห์งานวิจัยเชิงปริมาณแบบหนึ่งทีนักวิจัยนำงานวิจัย ซึ่งศึกษาปัญหาเดียวกันมาวิเคราะห์ด้วยวิธีการทางสถิติเพื่อสังเคราะห์ให้ได้ข้อสรุปที่มีความกว้างขวางกลุ่มลึกกว่าผลงานวิจัยแต่ละเรื่อง ข้อมูลสำหรับการวิเคราะห์ห่อภิมาณ คือ ดัชนีมาตรฐาน ได้แก่ ดัชนีขนาดอิทธิพล และดัชนีสัมประสิทธิ์สหสัมพันธ์ และข้อมูลคุณลักษณะงานวิจัย หน่วยการวิเคราะห์ คือ งานวิจัยหรือการทดสอบสมมติฐาน จุดมุ่งหมายในการวิเคราะห์แยกได้เป็น 2 ประการ ประการแรก คือ การสังเคราะห์ให้ได้ข้อสรุปเกี่ยวกับดัชนีมาตรฐาน ประการที่สอง คือ การวิเคราะห์เพื่อตรวจสอบความสัมพันธ์เชิงสาเหตุระหว่างตัวแปรกับดัชนีมาตรฐาน

สรุปได้ว่า การวิเคราะห์ห่อภิมาณ หมายถึง การสังเคราะห์งานวิจัยเชิงปริมาณที่มีระเบียบวิธีเป็นงานวิจัยเชิงสำรวจที่มุ่งศึกษาหาข้อสรุปของผลการวิจัยจำนวนมากที่ศึกษาปัญหาวิจัยเดียวกันด้วยการนำวิธีการทางสถิติมาวิเคราะห์ข้อมูล การวิเคราะห์ห่อภิมาณมีจุดมุ่งหมาย 2 ประการ คือ เพื่อสังเคราะห์ให้ได้ข้อสรุปเกี่ยวกับดัชนีมาตรฐาน และเพื่อตรวจสอบความสัมพันธ์เชิงสาเหตุระหว่างดัชนีมาตรฐานกับตัวแปรปรับ (moderator variables) ซึ่งได้แก่ ตัวแปรคุณลักษณะงานวิจัย ข้อมูลสำหรับการวิเคราะห์ คือ ข้อมูลคุณลักษณะงานวิจัยและดัชนีมาตรฐาน ได้แก่ ดัชนีขนาดอิทธิพล ดัชนีสัมประสิทธิ์สหสัมพันธ์ หน่วยการวิเคราะห์ คือ งานวิจัยและ/หรือชุดการทดสอบสมมติฐาน

ขั้นตอนของการวิเคราะห์ห่อภิมาณ

การวิเคราะห์ห่อภิมาณมีขั้นตอนของการดำเนินงาน 5 ขั้นตอนเช่นเดียวกับการดำเนินงานวิจัยทั่วไป (Hunter, Schmidt and Jackson, 1982; Glass, McGraw and Smith, 1979; Rosenthal and Rosnow, 1991; นงลักษณ์ วิรัชชัย, 2542) สิ่งที่ต่างกันคือรายละเอียดในแต่ละขั้นตอนมีดังนี้

ขั้นตอนที่ 1 การกำหนดปัญหาและวัตถุประสงค์การวิจัย ปัญหาหลัก 3 ข้อของการวิเคราะห์ห่อภิมาณ ได้แก่

- 1) ปัญหาเกี่ยวกับค่าแนวโน้มส่วนกลางของดัชนีมาตรฐาน
- 2) ปัญหาเกี่ยวกับความแปรปรวนของดัชนีมาตรฐาน
- 3) ผลการวิจัยที่สังเคราะห์นั้นเกี่ยวข้องกับคุณลักษณะของงานวิจัยที่นำมาสังเคราะห์อย่างไร

ดังนั้น การกำหนดวัตถุประสงค์สำหรับการวิเคราะห์ห่อถักมามต้องตั้งเพื่อให้สามารถตอบปัญหาดังกล่าวได้อย่างสอดคล้องกับวัตถุประสงค์ที่สำคัญสำหรับการวิเคราะห์ห่อถักมาม ได้แก่

- 1) เพื่อสังเคราะห์ผลการวิจัยที่มีข้อขัดแย้งหรือมีจำนวนมากให้ได้คำตอบที่เป็นข้อสรุป
- 2) เพื่อให้ได้องค์ความรู้ใหม่ทางวิชาการเพื่อนำไปใช้ประโยชน์เชิงวิชาการ
- 3) เพื่อให้ได้ข้อค้นพบที่เป็นความสัมพันธ์แบบมีเงื่อนไขสามารถนำไปใช้ประโยชน์ในการกำหนดนโยบายและการปฏิบัติจริงได้

ขั้นตอนที่ 2 การศึกษาเอกสารที่เกี่ยวข้องกับการวิจัย เป็นการศึกษาทฤษฎีและแนวคิดที่เกี่ยวข้องกับการวิจัย การศึกษาสังเคราะห์รายงานวิจัยที่เกี่ยวข้องกับการวิจัย ซึ่งรายงานวิจัยทั้งหมดเป็นการศึกษาปัญหาเดียวกัน เพื่อนำผลที่ได้มาสร้างกรอบแนวคิดในการวิเคราะห์ห่อถักมาม

ขั้นตอนที่ 3 การรวบรวมข้อมูล ซึ่งประกอบด้วยขั้นตอนการดำเนินงาน 4 ขั้นตอนย่อย ดังนี้

3.1 การเสาะค้นงานวิจัย ซึ่ง Mullen (1989) เสนอไว้ 6 แนวทาง ได้แก่ การสืบค้นข้อมูลด้วยคอมพิวเตอร์ออนไลน์ การสืบค้นจากดัชนี การสืบค้นจากการอ่านเล่น การสืบค้นจากการขอความช่วยเหลือจากนักวิชาการหรือเพื่อนร่วมงาน การสืบค้นย้อนหลัง และการสืบค้นต่อเนื่องด้วยคำค้นดัชนี

3.2 กำหนดกลุ่มตัวอย่างงานวิจัยที่นำมาสังเคราะห์ เป็นการตัดสินใจว่าจะสังเคราะห์งานวิจัยทุกเรื่องหรือเลือกศึกษาเฉพาะบางส่วน ซึ่ง Glass, McGraw and Smith (1981) Light and Pillemer (1984) เสนอไว้ 4 แนวทาง ได้แก่ การศึกษางานวิจัยทุกเรื่องที่สืบค้นได้ การเลือกงานวิจัยโดยใช้การเลือกแบบแบ่งชั้นตามลักษณะงานวิจัย การเลือกเฉพาะงานวิจัยที่มีการพิมพ์เผยแพร่ และการเลือกเฉพาะงานวิจัยที่มีคุณภาพ

3.3 การสร้างเครื่องมือวิจัย เครื่องมือมีอยู่ 2 ประเภท ได้แก่ แบบประเมินงานวิจัยที่มีข้อความบ่งบอกคุณภาพของงานวิจัย นิยมสร้างเป็นมาตรประมาณค่า (rating scale) และแบบบันทึกข้อมูลที่อาจสร้างเป็นแบบสอบถามทั้งปลายปิดและปลายเปิดก็ได้ หลักการสร้างแบบบันทึกข้อมูลสำหรับการวิเคราะห์ห่อถักมามมี 3 ส่วน ได้แก่ ส่วนรูปเล่มและการจัดพิมพ์ ส่วนคุณลักษณะของรายงานวิจัย และส่วนผลการวิจัย เมื่อสร้างแบบบันทึกให้ครอบคลุมตัวแปรทั้งหมดและกำหนดรหัสสำหรับแต่ละตัวแปรแล้วนำไปทดลองเก็บรวบรวมข้อมูลจากงานวิจัยจำนวน 5-10 เรื่อง ถ้าผู้เก็บรวบรวมข้อมูลมีหลายคนต้องทำการตรวจสอบความสอดคล้องระหว่างผู้เก็บรวบรวมข้อมูลด้วย

3.4 การบันทึกข้อมูล เป็นการลงรหัสและเตรียมข้อมูลเพื่อการวิเคราะห์ต่อไป การบันทึกผลการวิจัยที่ต้องประมาณค่าเป็นดัชนีมาตรฐาน ซึ่ง Glass, McGraw and Smith (1981) เสนอสูตรการประมาณค่าขนาดอิทธิพลและค่าสัมประสิทธิ์สหสัมพันธ์จากงานวิจัยแต่ละเรื่อง โดยเสนอวิธีการประมาณค่า 2 วิธี ได้แก่ การคำนวณโดยตรงจากค่าสถิติที่ได้จากกลุ่มตัวอย่าง และการคำนวณจากค่าจากสถิติที่ได้จากการทดสอบนัยสำคัญ

ขั้นตอนที่ 4 การวิเคราะห์ข้อมูล ซึ่งประกอบด้วยการดำเนินงาน 4 ขั้นตอนย่อย ดังนี้

4.1 การเตรียมข้อมูลสำหรับการวิเคราะห์ เป็นการเตรียมไฟล์ข้อมูลและการตรวจสอบว่ามีตัวแปรใดมีค่าขาดหาย (missing) ถ้ามีค่าขาดหายแบบสุ่มก็สามารถประมาณค่าทดแทนได้ด้วยการใช้ค่าเฉลี่ยตัวแปรนั้น หรือใช้ค่าประมาณที่ได้จากการวิเคราะห์การถดถอยระหว่างตัวแปรที่มีข้อมูลหายขาดกับตัวแปรที่สัมพันธ์กัน (Hair et al., 1998)

4.2 การตรวจสอบลักษณะการแจกแจงของตัวแปร สำหรับ ทำการตรวจสอบลักษณะการแจกแจงของตัวแปรไม่ต่อเนื่องด้วยตารางแจกแจงความถี่ และทำการตรวจสอบลักษณะการแจกแจงของตัวแปรต่อเนื่องว่าเป็นโค้งปกติหรือไม่ ด้วยสถิติเชิงบรรยาย ได้แก่ ค่าเฉลี่ย ค่าส่วนเบี่ยงเบนมาตรฐาน ค่าความเบ้ ค่าความโด่ง

4.3 การวิเคราะห์ข้อมูลเบื้องต้น เป็นการเสนอค่าสถิติเบื้องต้นของตัวแปรที่สำคัญ อาทิเช่น การแสดงความสัมพันธ์ระหว่างตัวแปร การเสนอค่าดัชนีมาตรฐานและลักษณะการกระจายของดัชนีมาตรฐาน เป็นต้น

4.4 การวิเคราะห์เพื่อตอบคำถามการวิจัย เป็นการวิเคราะห์ผลการวิจัย ซึ่งวัดอยู่ในรูปดัชนีมาตรฐานทั้งหมดว่ามีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติหรือไม่ ถ้าไม่แตกต่างกันก็สามารถสรุปผลการสังเคราะห์ทั้งหมดเพื่อตอบปัญหาการวิจัยได้ แต่ถ้าแตกต่างกันอย่างมีนัยสำคัญทางสถิติต้องวิเคราะห์ต่อไปว่า ความแตกต่างหรือความแปรปรวนนั้นเกิดขึ้นเนื่องจากคุณลักษณะงานวิจัยด้านใดบ้าง แล้วแยกงานวิจัยเป็นกลุ่มตามลักษณะนั้น เพื่อสังเคราะห์สรุปผลการวิจัยในแต่ละกลุ่ม โดยการสังเคราะห์แบบวิธีบรรยายหรือการสังเคราะห์แบบวิธีเชิงปริมาณก็ได้

ขั้นตอนที่ 5 การสรุปและนำเสนอรายงาน เป็นการสรุปผลและอภิปรายผลที่มีการเชื่อมโยงผลการวิจัยกับความรู้ในอดีตและความรู้ทางทฤษฎี การให้ข้อเสนอแนะเชิงนโยบายและเชิงวิชาการต้องมีความกว้างโดยทั่วไป (generality) มากกว่างานวิจัยปกติ รายงานการวิเคราะห์อภิปรายจะมีบรรณานุกรมแยกเป็น 2 ส่วน ได้แก่ บรรณานุกรมของรายงานวิจัย และบรรณานุกรมของงานวิจัยที่นำมาสังเคราะห์

วิธีการวิเคราะห์อภิปราย

วิธีการวิเคราะห์อภิปรายมีเป้าหมายหลัก 3 ประการ คือ

- 1) การประมาณค่าพารามิเตอร์ของดัชนีมาตรฐาน
- 2) การอธิบายความแปรปรวนของดัชนีมาตรฐานด้วยตัวแปรคุณลักษณะงานวิจัย
- 3) การศึกษาอิทธิพลของปฏิสัมพันธ์ของตัวแปรคุณลักษณะงานวิจัยที่ส่งผลต่อ

ดัชนีมาตรฐาน

ซึ่งในการวิเคราะห์ห่อถักมานั้นไม่ว่าจะใช้วิธีการแบบใดก็ควรตอบคำถามทุกข้อที่กล่าวมาทั้ง 3 ประการ การวิเคราะห์ข้อมูลในการวิเคราะห์ห่อถักมานั้น 6 วิธีการ มีรายละเอียดดังนี้

1. วิธีการของ Glass, McGraw and Smith (1981) วิธีการนี้ได้รับความนิยมอย่างมากเป็นการสังเคราะห์งานวิจัยเชิงทดลองรวมกับงานวิจัยเชิงสหสัมพันธ์ โดยมีสูตรการคำนวณค่าขนาดอิทธิพลจากค่าสหสัมพันธ์ได้ และมีสูตรสำหรับการคำนวณค่าดัชนีมาตรฐานทั้งค่าขนาดอิทธิพลและค่าสัมประสิทธิ์สหสัมพันธ์จากงานวิจัยที่มีแบบการวิจัยแตกต่างกัน โดยมีการคำนวณจากค่าสถิติโดยตรงและคำนวณจากค่าสถิติในการทดสอบสมมติฐาน การคำนวณค่าขนาดอิทธิพลจะใช้ค่าส่วนเบี่ยงเบนมาตรฐานของกลุ่มควบคุมและค่าส่วนเบี่ยงเบนมาตรฐานรวมจุดเด่นของวิธีการนี้ คือ มีสูตรในการประมาณค่าขนาดอิทธิพลจากงานวิจัยที่มีรูปแบบการทดลองแตกต่างกันทุกแบบแผนการวิจัย และมีสูตรในการปรับเปลี่ยนสัมประสิทธิ์สหสัมพันธ์แบบอื่นมาเป็นสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน ส่วนการอธิบายความแปรปรวนของดัชนีมาตรฐานนั้นจะใช้การวิเคราะห์การถดถอย การวิเคราะห์ความแปรปรวน หรือสถิติขั้นสูงอื่นด้วย

2. วิธีการของ Hunter, Schmidt and Jackson (1982) เป็นวิธีการสังเคราะห์ที่ให้ความสำคัญกับการสังเคราะห์งานวิจัยเชิงสหสัมพันธ์มากกว่าการวิจัยเชิงทดลอง การคำนวณค่าขนาดอิทธิพลใช้ค่าส่วนเบี่ยงเบนมาตรฐานรวมตามแนวคิดของ Cohen (1969) ก่อนที่จะสังเคราะห์ค่าดัชนีมาตรฐาน ทั้งดัชนีขนาดอิทธิพลและดัชนีสหสัมพันธ์ต้องมีการปรับแก้ความคลาดเคลื่อนเนื่องจากความจำกัดของพิสัยและความคลาดเคลื่อนเนื่องจากการสุ่มตัวอย่างก่อน เมื่อปรับแก้แล้วจึงตรวจสอบความแปรปรวนมีระบบของดัชนีมาตรฐานที่ปรับแก้แล้ว ถ้าไม่มีความแปรปรวนมีระบบจึงสามารถสังเคราะห์ดัชนีมาตรฐานได้ ถ้ายังมีความแปรปรวนมีระบบอยู่จะต้องแยกกลุ่มงานวิจัยออกเป็นกลุ่มย่อย โดยใช้ตัวแปรปรับเป็นเกณฑ์ในการแบ่งกลุ่ม แล้วดำเนินการวิเคราะห์แต่ละกลุ่มตามขั้นตอนจนกว่าจะสังเคราะห์ผลการวิจัยได้ จุดเด่นของวิธีการนี้คือการปรับแก้ความคลาดเคลื่อนก่อนการสังเคราะห์ และมีข้อดีที่สามารถสังเคราะห์งานวิจัยที่ศึกษาปัญหาเดียวกันเพียง 3-4 เรื่องได้

3. วิธีการของ Hedges and Olkin (1985) เป็นวิธีการสังเคราะห์ที่ให้ความสำคัญกับคุณสมบัติทางสถิติของดัชนีมาตรฐาน โดยเฉพาะค่าขนาดอิทธิพล การคำนวณค่าอิทธิพลใช้ค่าส่วนเบี่ยงเบนมาตรฐานรวมตามแนวคิดของ Cohen (1969) และมีการปรับแก้ให้ได้ค่าประมาณไม่คลาดเคลื่อน มีการปรับแก้ความคลาดเคลื่อนเนื่องจากการวัดและความจำกัดของพิสัยตามแบบของ Hunter แต่ไม่มีการปรับแก้ความคลาดเคลื่อนจากการสุ่มตัวอย่าง มีการเสนอสูตรการคำนวณค่าเฉลี่ยโดยการถ่วงน้ำหนักและเสนอว่า การแจกแจงของค่าขนาดอิทธิพลเป็นการแจกแจงแบบไค-สแควร์ รวมทั้งให้ค่าสถิติทดสอบ Q ในการทดสอบว่าพารามิเตอร์ของขนาดอิทธิพลมีขนาดเท่ากันสำหรับทุกกลุ่มประชากรก่อนที่จะทำการสังเคราะห์

4. วิธีการของ Slavin (1987) เป็นวิธีการสังเคราะห์ที่ให้ความสำคัญกับการคัดเลือกรงานวิจัยที่นำมาสังเคราะห์ เนื่องจาก Slavin เชื่อว่าถ้างานวิจัยไม่มีคุณภาพ ผลการสังเคราะห์ย่อมไม่มีคุณภาพด้วย ดังนั้นจึงต้องมีการประเมินคุณภาพงานวิจัยและคัดเลือกเฉพาะงานวิจัยที่มีคุณภาพตามเกณฑ์มาสังเคราะห์ แต่วิธีการนี้มีข้อจำกัดในทางปฏิบัติจริงเพราะงานวิจัยที่นำมาสังเคราะห์มีจำนวนไม่มากนัก ดังนั้น เมื่อคัดงานวิจัยที่ไม่มีคุณภาพออกเหลือเฉพาะที่มีคุณภาพดีจะทำให้มีจำนวนงานวิจัยในการวิเคราะห์ห่อภิมาณน้อยลงไปอีก ทำให้การสรุปอ้างอิงผลการวิจัยมีข้อจำกัด

5. วิธีการของ Mullen (1989) เป็นวิธีการสังเคราะห์ที่ใช้การประมาณค่าขนาดอิทธิพลตามแนวคิดของ Cohen (1969) และการใช้การประมาณค่าความสัมพันธ์ระหว่างตัวแปรโดยใช้คะแนน Fisher ในการวิเคราะห์การประมาณค่าเฉลี่ยของดัชนีมาตรฐานใช้การหาค่าเฉลี่ยแบบถ่วงน้ำหนัก จุดเด่นของวิธีการนี้คือ การพัฒนาโปรแกรมคอมพิวเตอร์เพื่อใช้สำหรับการวิเคราะห์ห่อภิมาณโดยตรง โปรแกรมสามารถทดสอบความเป็นเอกพันธ์ของดัชนีมาตรฐานและตรวจสอบว่าตัวแปรปรับตัวใด สามารถอธิบายความแปรปรวนในดัชนีมาตรฐานได้

6. วิธีการของ Rosenthal (1991) เป็นวิธีการสังเคราะห์ที่แตกต่างจากการวิเคราะห์ห่อภิมาณแบบอื่น คือ การนำค่าระดับนัยสำคัญมาใช้ในการสังเคราะห์และคำนวณค่าขนาดอิทธิพล 2 แบบ คือ แบบที่ 1 คำนวณจากค่าขนาดกลุ่มตัวอย่างและค่าสถิติในการทดสอบสมมติฐาน และแบบที่ 2 คำนวณจากขนาดกลุ่มตัวอย่างและระดับนัยสำคัญ การประมาณค่าขนาดอิทธิพลใช้ตามแนวคิดของ Cohen (1969) คือ ใช้ส่วนเบี่ยงเบนมาตรฐานรวมเป็นตัวหาร สำหรับการสังเคราะห์งานวิจัยเชิงสหสัมพันธ์ต้องมีการเปลี่ยนค่าให้เป็นคะแนน Fisher's Z ก่อนที่จะมีการสังเคราะห์ วิธีการสังเคราะห์ใช้การจัดกลุ่มงานวิจัยตามตัวแปรปรับ และเมื่อตรวจสอบแล้วพบว่า ในแต่ละกลุ่มมีความแปรปรวนน้อยจึงสังเคราะห์ขนาดอิทธิพลหรือสหสัมพันธ์ได้

ลักษณะของการวิเคราะห์ห่อภิมาณ

การวิเคราะห์ห่อภิมาณมีประเด็นที่ต้องคำนึงถึงและพิจารณาให้เกิดความเข้าใจอยู่ 4 ประเด็น มีดังต่อไปนี้

ประเด็นที่ 1 งานวิจัยที่นำมาสังเคราะห์ต้องเป็นงานวิจัยที่ศึกษาปัญหาเดียวกัน เป็นประเด็นที่ค่อนข้างสำคัญเพราะถ้าไม่สามารถทำความเข้าใจประเด็นนี้ได้ถูกต้องจะมีปัญหาในการคัดเลือกรงานวิจัยที่นำมาสังเคราะห์ตามความหมายของประเด็นนี้ งานวิจัยที่นำมาสังเคราะห์ต้องมีปัญหาวิจัยตรงกัน แต่งานวิจัยอาจจะใช้แบบแผนการวิจัยต่างกัน เครื่องมือวัดต่างกัน กลุ่มตัวอย่างต่างกัน วิธีการวิเคราะห์ต่างกันก็ได้ คำว่า “ปัญหาวิจัยตรงกัน” ในที่นี้จึงขึ้นอยู่กับนิยามและการกำหนดขอบเขตของการวิจัย หากนักวิจัยกำหนดนิยามตัวแปรให้มีขอบเขตและลักษณะเฉพาะเจาะจงจะมีงานวิจัยที่จะนำมาสังเคราะห์เป็นจำนวนไม่มากเพราะ

การให้คำนิยามตัวแปรที่มีขอบเขตแคบมากจะได้งานวิจัยมาศึกษาจำนวนน้อยและการสรุปผลจะไม่ค่อยกว้างขวาง แต่ถ้านักวิจัยให้คำนิยามหลวมมีขอบเขตกว้างขวางมากเกินไปจะได้งานวิจัยมาสังเคราะห์เป็นจำนวนมาก และผลการวิจัยมีความหลากหลายมีตัวแปรปรับจำนวนมาก และทำให้การวิเคราะห์มีความซับซ้อนมากขึ้นด้วย อาทิเช่น รัตน์ บัสนันท์ และคณะ (2553) ทำการสังเคราะห์งานวิจัยด้านการคิดวิเคราะห์ของนักเรียน: การวิเคราะห์อภิमान โดยงานวิจัยที่นำมาทำการสังเคราะห์มีตัวอย่างดังนี้

เรื่องที่ 1 ผลของการจัดการเรียนรู้ที่เน้นการสร้างความรู้ด้วยตนเองต่อผลสัมฤทธิ์ทางการเรียน และความสามารถในการคิดวิเคราะห์ในวิชาวิทยาศาสตร์ของ นักเรียนช่วงชั้นที่ 3 ชั้นปีที่ 1

เรื่องที่ 2 การเปรียบเทียบความสามารถในการคิดวิเคราะห์ตามสาระการเรียนรู้ภาษาไทยของนักเรียนชั้นมัธยมศึกษาปีที่ 3 โดยเทคนิคการใช้และไม่ใช้ผังกราฟฟิก

เรื่องที่ 3 การเปรียบเทียบผลสัมฤทธิ์ทางการเรียนและความสามารถในการคิดวิเคราะห์กลุ่มสาระการเรียนรู้สังคมศึกษา ศาสนาและวัฒนธรรมของนักเรียนชั้นประถมศึกษาปีที่ 6 ระหว่างการจัดการเรียนรู้แบบโยนิโสมนสิการและการจัดการเรียนรู้แบบ 4MAT

เรื่องที่ 4 การเปรียบเทียบความสามารถการคิดวิเคราะห์และเจตคติต่อกลุ่มสาระการเรียนรู้คณิตศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 2 โดยการจัดการเรียนรู้แบบใช้คำถามตามแนวคิดแบบหวนหกกับการจัดการเรียนรู้แบบปกติ

เรื่องที่ 5 การเปรียบเทียบผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์เรื่องบทประยุกต์ ความพึงพอใจต่อการเรียนวิชาคณิตศาสตร์และความสามารถในการคิดวิเคราะห์ของนักเรียนชั้นประถมศึกษาปีที่ 6 ระหว่างการจัดกิจกรรมการเรียนรู้โดยใช้เทคนิค STAD กับการจัดกิจกรรมการเรียนรู้ตามคู่มือครู

ตัวอย่างงานวิจัยทั้ง 5 เรื่องนี้เป็นงานวิจัยเชิงทดลองทุกเรื่อง มีกรอบการศึกษาปัญหาเดียวกัน ตัวแปรต้น คือรูปแบบของวิธีการสอน ตัวแปรตามที่เหมือนกันทุกเรื่องคือความสามารถในการคิดวิเคราะห์ กลุ่มตัวอย่างเป็นนักเรียนชั้นประถมศึกษาและนักเรียนชั้นมัธยมศึกษา ซึ่งก็คือนักเรียนระดับการศึกษาขั้นพื้นฐานนั่นเอง ดังแสดงในตาราง

ตารางแสดงการวิเคราะห์ประเภท ตัวแปรต้น ตัวแปรตาม และกลุ่มตัวอย่างของตัวอย่างงานวิจัยแต่ละเรื่อง

เรื่องที่	ประเภท	ตัวแปรต้น	ตัวแปรตาม	กลุ่มตัวอย่าง
1	งานวิจัยเชิงทดลอง	การจัดการเรียนรู้ที่เน้นการสร้างความรู้ด้วยตนเอง	1. ผลสัมฤทธิ์ทางการเรียน 2. ความสามารถในการคิดวิเคราะห์	นักเรียนช่วงชั้นที่ 3 ชั้นปีที่ 1
2	งานวิจัยเชิงทดลอง	เทคนิคการใช้และไม่ใช้ผังกราฟฟิก	1. ความสามารถในการคิดวิเคราะห์	นักเรียนชั้นมัธยมศึกษาปีที่ 3
3	งานวิจัยเชิงทดลอง	การจัดการเรียนรู้แบบโยนิโสมนสิการและการจัดการเรียนรู้แบบ 4MAT	1. ผลสัมฤทธิ์ทางการเรียน 2. ความสามารถในการคิดวิเคราะห์	นักเรียนชั้นประถมศึกษาปีที่ 6
4	งานวิจัยเชิงทดลอง	การจัดการเรียนรู้แบบใช้คำถามตามแนวคิดแบบหมวกหกใบกับการจัดการเรียนรู้แบบปกติ	1. ความสามารถในการคิดวิเคราะห์ 2. เจตคติ	นักเรียนชั้นมัธยมศึกษาปีที่ 2
5	งานวิจัยเชิงทดลอง	การจัดกิจกรรมการเรียนรู้โดยใช้เทคนิค STAD กับการจัดกิจกรรมการเรียนรู้ตามคู่มือครู	1. ผลสัมฤทธิ์ทางการเรียน 2. ความพึงพอใจ 3. ความสามารถในการคิดวิเคราะห์	นักเรียนชั้นประถมศึกษาปีที่ 6

ประเด็นที่ 2 ข้อมูลสำหรับการสังเคราะห์งานวิจัย งานวิจัยเป็นข้อมูลสำหรับการสังเคราะห์งานวิจัยด้วยการวิเคราะห์ห่อภิวน ข้อมูลในที่นี้ประกอบด้วย ผลการวิจัยจากงานวิจัยแต่ละเรื่องและคุณลักษณะงานวิจัย เนื่องจากงานวิจัยแต่ละเรื่องศึกษาตัวแปรต่างกันหรือแม้จะศึกษาตัวแปรเหมือนกันแต่ก็อาจใช้เครื่องมือวัดต่างกัน ดังนั้น ผลการวิจัยจากงานวิจัยแต่ละเรื่องจึงไม่อยู่ในสเกลเดียวกัน จำเป็นต้องทำให้อยู่ในรูปดัชนีมาตรฐาน (standard indices)

ประเด็นที่ 3 หน่วยการวิเคราะห์ (unit of analysis) สำหรับการวิเคราะห์ห่อภิวน มีลักษณะแตกต่างจากหน่วยการวิเคราะห์ในงานวิจัยโดยทั่วไปทางสังคมศาสตร์ และพฤติกรรมศาสตร์ หน่วยการวิเคราะห์ คือ หน่วยตัวอย่างแต่ละเรื่องที่ให้ข้อมูลสำหรับการวิจัย งานวิจัยบางเรื่องอาจมีหน่วยการวิเคราะห์เป็นระดับองค์กร งานวิจัยบางเรื่องอาจมีหน่วยการวิเคราะห์มีทั้งระดับนักเรียน ระดับชั้นเรียน และระดับโรงเรียน แต่หน่วยการวิเคราะห์สำหรับการสังเคราะห์งานวิจัยคือ รายงานวิจัยแต่ละเรื่อง

ประเด็นที่ 4 ลักษณะเฉพาะของผลการสังเคราะห์งานวิจัย ผลการสังเคราะห์งานวิจัยมีขอบข่ายองค์ความรู้กว้างขวางลึกซึ้งมากกว่าองค์ความรู้ที่ได้จากงานวิจัยแต่ละเรื่องที่น่ามาสังเคราะห์ เพราะผลการสังเคราะห์ให้องค์ความรู้ที่มีขอบเขตกว้างขวางกว่างานวิจัยแต่ละเรื่อง และยังให้ผลการเปรียบเทียบความคล้ายคลึงและความแตกต่างระหว่างงานวิจัยแต่ละเรื่องทำให้ทราบว่ามีปัจจัยหรือเงื่อนไขใดที่ทำให้ได้ผลการวิจัยแตกต่างกัน

สรุป

การสังเคราะห์งานวิจัยทางการศึกษามีจุดมุ่งหมายเพื่อหาข้อสรุปความคล้ายคลึง และความแตกต่างจากงานวิจัย โดยนำคุณลักษณะงานวิจัยมาเป็นตัวแปรในการอธิบายความแตกต่างระหว่างงานวิจัย การดำเนินงานมี 5 ขั้นตอน ได้แก่ ขั้นการกำหนดหัวข้อปัญหา ขั้นการวิเคราะห์ปัญหา ขั้นการเสาะค้น คัดเลือกและรวบรวมงานวิจัย ขั้นการวิเคราะห์เพื่อสังเคราะห์ผลการวิจัย และขั้นการนำเสนอรายงานผลการสังเคราะห์การวิจัย

การวิเคราะห์ห่อภีมีจุดมุ่งหมาย 2 ประการ คือ เพื่อสังเคราะห์ให้ได้ข้อสรุปเกี่ยวกับดัชนีมาตรฐาน และเพื่อตรวจสอบความสัมพันธ์เชิงสาเหตุระหว่างดัชนีมาตรฐานกับตัวแปรปรับ ซึ่งได้แก่ ตัวแปรคุณลักษณะงานวิจัย ข้อมูลสำหรับการวิเคราะห์ คือ ข้อมูลคุณลักษณะงานวิจัย และดัชนีมาตรฐาน ได้แก่ ดัชนีขนาดอิทธิพล ดัชนีสัมประสิทธิ์สหสัมพันธ์ หน่วยการวิเคราะห์ คือ งานวิจัยและ/หรือชุดการทดสอบสมมติฐาน เครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูล คือ แบบประเมินคุณภาพงานวิจัยและแบบบันทึกคุณลักษณะงานวิจัย การดำเนินงานมี 5 ขั้นตอน ได้แก่ ขั้นการกำหนดปัญหาและวัตถุประสงค์การวิจัย ขั้นการศึกษาเอกสารที่เกี่ยวข้องกับการวิจัย ขั้นการรวบรวมข้อมูลที่ประกอบด้วย การเสาะค้นงานวิจัย การกำหนดกลุ่มตัวอย่างงานวิจัยที่นำมาสังเคราะห์ การสร้างเครื่องมือวิจัย การบันทึกข้อมูล ขั้นการวิเคราะห์ข้อมูลที่ประกอบด้วย การเตรียมข้อมูลสำหรับการวิเคราะห์ การตรวจสอบลักษณะการแจกแจงของตัวแปร การวิเคราะห์ข้อมูลเบื้องต้น การวิเคราะห์เพื่อตอบคำถามการวิจัย สุดท้ายขั้นการสรุปและนำเสนอรายงานวิจัย การวิเคราะห์ห่อภีมี 6 วิธีการ ได้แก่ วิธีการของ Glass, McGraw and Smit; Hunter, Schmidt and Jackson; Hedges and Olkin; Slavin; Mullen; Rosenthal นอกจากนี้ การวิเคราะห์ห่อภีมีประเด็นที่ต้องคำนึงถึงและพิจารณาให้เกิดความเข้าใจอยู่ 4 ประเด็น คือ งานวิจัยที่นำมาสังเคราะห์ต้องเป็นงานวิจัยที่ศึกษาปัญหาเดียวกัน ข้อมูลสำหรับการสังเคราะห์งานวิจัยประกอบด้วยผลการวิจัยจากงานวิจัยแต่ละเรื่องและคุณลักษณะงานวิจัย หน่วยการวิเคราะห์คือหน่วยตัวอย่างแต่ละเรื่องที่ให้ข้อมูลสำหรับการวิจัย ผลการสังเคราะห์งานวิจัยมีขอบข่ายองค์ความรู้กว้างขวางลึกซึ้งมากกว่าองค์ความรู้ที่ได้จากงานวิจัยแต่ละเรื่องที่นำมาสังเคราะห์

งานวิจัยทางการศึกษาที่มีจำนวนเพิ่มมากขึ้นอย่างรวดเร็วประกอบกับงานวิจัยหลายเรื่องเป็นการศึกษาปัญหาการวิจัยเดียวกันหรือศึกษาตัวแปรที่มีลักษณะคล้ายคลึงกัน ในอนาคตจึงทำให้การสังเคราะห์งานวิจัยและการวิเคราะห์ห่อภีจะมีบทบาทอย่างยิ่งเพื่อหาข้อสรุปอย่างเป็นระบบจากรายงานการวิจัยที่ศึกษาตัวแปรตามเดียวกัน หรือการศึกษาสถานภาพ (state of the art) ของอาณาบริเวณองค์ความรู้ในสาขาวิชาทางการศึกษา ข้อค้นพบที่ได้สามารถนำไปต่อยอดสร้างองค์ความรู้ใหม่หรือนำไปประยุกต์ใช้ประโยชน์ในเชิงวิชาการหรือข้อเสนอเชิงนโยบายทางการศึกษาได้

เอกสารอ้างอิง

- ดุขฎิ โยเหลา. (2535). การวิเคราะห์โดยใช้วิธีวิเคราะห์แบบเมตต้า (Meta-Analysis). วารสาร
ข่าวสารวิจัยการศึกษา. 16(1): 19-28.
- นงลักษณ์ วิรัชชัย. (2529). การสังเคราะห์งานวิจัย. วารสารการวัดผลการศึกษา. 8(23): 26-36.
- นงลักษณ์ วิรัชชัย. (2542). การวิเคราะห์ห่อภิมาน. กรุงเทพฯ: นิชนแอตเวอร์ไทซิงกรุ๊ป.
- นงลักษณ์ วิรัชชัย และสุวิมล ว่องวาณิช. (2542). การสังเคราะห์งานวิจัยทางการศึกษาด้วย
การวิเคราะห์ห่อภิมานและ การวิเคราะห์เนื้อหา. รายงานการวิจัย, สำนักงานคณะกรรมการ
ศึกษาแห่งชาติ สำนักนายกรัฐมนตรี.
- รัตน์ บัสนน. (2534). บางแง่มุมของการวิเคราะห์ห่อภิมาน. วารสารข่าวสารวิจัยการศึกษา. 14(5):
7-10.
- รัตน์ บัสนน และคณะ. (2553). การสังเคราะห์งานวิจัยด้านการคิดวิเคราะห์ของนักเรียน:
การวิเคราะห์ห่อภิมาน. วารสารศึกษาศาสตร์ มหาวิทยาลัยนเรศวร. 12(3): 19-32.
- ศิริภา พูลสุวรรณ. (2555). การสังเคราะห์งานวิจัย: หลักการและการประยุกต์ใช้. กรุงเทพฯ:
มิสเตอร์ก๊อปปี (ประเทศไทย).
- สุศักดิ์ หลาบมาลา. (2532). การวิเคราะห์ผลรวม. วารสารการวิจัยเพื่อการพัฒนา. 29
(พฤศจิกายน 2532): 1-2
- สุพัฒน์ สุขมลสันต์. (2532). การอภิวิเคราะห์และการสังเคราะห์งานวิจัยที่เกี่ยวข้องกับการเรียน
การสอนภาษาอังกฤษ ในระหว่างปี พ.ศ.2515-2530. วารสารวิธีวิทยาการวิจัย. 4(3),
67-78.
- สุวัฒนา สุวรรณเขตนิคม. (2529). การวิเคราะห์เมตต้าของงานวิจัย. วารสารวิธีวิทยาการวิจัย.
1,2(พฤษภาคม-สิงหาคม 2529): 16-23.
- อุทุมพร จามรมาร. (2527). การสังเคราะห์งานวิจัย: เชิงปริมาณ. กรุงเทพฯ: จุฬาลงกรณ์
มหาวิทยาลัย.
- Cohen J. (1969). Statistical Power Analysis for Behavioral Science. New York:
Academic Press.
- Glass G.V. (1976). Primary, secondary and meta-analysis of research results.
Educational Researcher. 5: 3-8.
- Glass G.V., McGaw B. & Smith, M.L. (1981). Meta-Analysis in Social Research.
Beverly Hills: Sage Publications.
- Hair J.F. et al. (1998). Multivariate data analysis. Upper Saddle River, New Jersey:
Prentice Hall.
- Harries C., Larry V.H., & Jeffrey C.V. (2019). The Handbook of Research Synthesis
and Meta-Analysis. 3rd ed. New York: Russell Sage Foundation.

- Hedges L.V. & Olkin I. (1985). *Statistical Methods for Meta-Analysis*. Orlando, Florida: Academic Press, Inc.
- Hunter J.E., Schmidt F.L. & Jackson G.B. (1982). *Meta-Analysis: Cumulating Research Findings Across Studies*. Beverly Hills: Sage Publications.
- Kulik J.A. & Kulik C.C. (1989). Meta-Analysis in Education. *International Journal of Educational Research*. 13: 223-240.
- Kumar R. (2005). *Research Methodology: A step-by-step for the beginners*. 2nd ed. Thousand Oaks: Sage Publications.
- Light R.J. & Pillemer D.B. (1984). *Summing Up: The Science of Reviewing Research*. Cambridge, Massachusetts: Harvard University Press.
- Mullen B. (1989). *Advanced BASIC Meta-Analysis*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.
- Noblit G.W. & Hare R.D. (1988). *Meta-Ethnography: Synthesizing Qualitative Studies*. Newbury Park: Sage Publications.
- Rosenthal R. (1991). *Meta-Analytic Procedures for Social Research*. Newbury Park: Sage Publications.
- Rosenthal R. & Rosnow R.L. (1991). *Essentials of Behavioral Research: Methods and Data Analysis*. 2nd ed. New York: McGraw-Hill, Inc.
- Slavin R.E. (1987). Best-evidence synthesis: why less is more. *Educational Researcher*. 16: 15-16.

รายละเอียดรูปแบบของบทความวิชาการ

การเสนอบทความเพื่อตีพิมพ์ในวารสารฯ ให้ผู้เขียนนำส่งบทความทางไปรษณีย์ หรือทางไปรษณีย์ อิเล็กทรอนิกส์ บทความที่เปิดรับให้ตีพิมพ์เผยแพร่ ประกอบด้วยบทความผลงานวิจัย บทความวิชาการ นวัตกรรม และเทคนิคด้านวัดและประเมินผลทางการศึกษา และบทความที่เป็นศาสตร์ทางการศึกษาในแขนงที่เกี่ยวข้อง อันจะก่อให้เกิดองค์ความรู้ในด้านวัดและประเมินผลทางการศึกษา

โดยบทความที่ส่งผู้เขียนต้องไม่เคยตีพิมพ์เผยแพร่ในวารสารฉบับอื่น หรือช่องทางการเผยแพร่อื่นมา ก่อน เมื่อกองบรรณาธิการได้รับบทความ ก่อนนำออกตีพิมพ์เผยแพร่ จะเสนอบทความให้ผู้ทรงคุณวุฒิ (peer review) ทำการประเมินคุณภาพอย่างน้อย 2 ท่าน โดยผู้ประเมินจะไม่ทราบชื่อเจ้าของผลงาน และเจ้าของผลงานจะไม่ทราบชื่อผู้ประเมิน

การเสนอบทความตีพิมพ์ มีแนวปฏิบัติดังนี้

1. เสนอบทความได้ทั้งภาษาไทย และภาษาอังกฤษ
2. บทความทุกเรื่อง ต้องเขียนบทคัดย่อภาษาอังกฤษ (abstract) กำกับมาด้วย หากเป็นบทความภาษาอังกฤษ ให้เขียนบทคัดย่อภาษาไทย
3. พิมพ์ด้วยโปรแกรม Microsoft word ใช้ตัวอักษร TH Sarabun New ขนาด 15pt.
4. ตั้งค่าหน้ากระดาษ A4 ขอบกระดาษทุกด้าน ๆ ละ 1 นิ้ว (2.5 เซนติเมตร) รูปแบบ 1 คอลัมน์ จำนวนหน้า ไม่เกิน 15 หน้า (รวมบรรณานุกรม)
5. บทความต้องมีรายการอ้างอิงครบถ้วน โดยยึดรูปแบบการเขียนบรรณานุกรมแบบ APA 6th edition ของ สทศ.

(รายละเอียดเพิ่มเติม : <https://www.niets.or.th/th/catalog/view/3949>)

This image shows a full page of white paper with horizontal dotted lines. The lines are evenly spaced and run across the width of the page, providing a guide for handwriting practice. There are no margins, text, or other markings on the page.



สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน)

128 อาคารพญาไท พลาซ่า ชั้น 36 แขวงทุ่งพญาไท

เขตราชเทวี กรุงเทพฯ 10400

ติดต่อกอง บก. โทร. 02-217-3800 ต่อ 113 / โทรสาร: 02-219-2996

E-mail: JOURNAL.NIETS@GMAIL.COM