

**การประเมินความต้องการจำเป็นในการพัฒนาสมรรถนะการประเมินเพื่อการเรียนรู้ของครู
ระดับประถมศึกษาที่ได้รับผลกระทบจากการทดสอบทางการศึกษาแห่งชาติ**
**The Needs Assessment for The Development of Assessment Competency for
Learning of Elementary School Teachers Affected by National Education Tests**

กฤษฎากาญจน์ โตพิทักษ์¹

Krittayakan Topithak¹

Received: May 30, 2021; Revised: August 1, 2022; Accepted: August 3, 2022

บทคัดย่อ

การวิจัยนี้มีวัตถุประสงค์เพื่อ 1) วิเคราะห์กลุ่มแฝงของครูระดับประถมศึกษาที่ได้รับผลกระทบจากการทดสอบทางการศึกษาระดับชาติ 2) ประเมินความต้องการจำเป็นในการพัฒนาสมรรถนะการประเมินเพื่อการเรียนรู้จำแนกตามกลุ่มแฝงของครูระดับประถมศึกษาที่ได้รับผลกระทบจากการทดสอบทางการศึกษาระดับชาติ และ 3) ประเมินสมรรถนะการประเมินเพื่อการเรียนรู้จำแนกตามกลุ่มแฝงของครูระดับประถมศึกษาที่ได้รับผลกระทบจากการทดสอบทางการศึกษาระดับชาติ กลุ่มตัวอย่างเป็นครูระดับประถมศึกษาชั้นปีที่ 6 จำนวน 600 คน ในเขตพื้นที่ภาคเหนือตอนล่าง 1 ได้จากการสุ่มแบบหลายขั้นตอน เครื่องมือวิจัย ได้แก่ แบบสอบถามความต้องการจำเป็นในการพัฒนาสมรรถนะการประเมินเพื่อการเรียนรู้ แบบทดสอบเชิงสถานการณ์เพื่อวัดสมรรถนะการประเมินเพื่อการเรียนรู้ และแบบวัดผลกระทบจากการทดสอบทางการศึกษาระดับชาติ วิเคราะห์ด้วยร้อยละ การวิเคราะห์กลุ่มแฝง (Latent class analysis) ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน และดัชนี PNI แบบปรับปรุง และการวิเคราะห์ความแปรปรวนทางเดียว ผลการวิจัยพบว่า

1. กลุ่มแฝงครูระดับประถมศึกษาในภาคเหนือตอนล่างที่ได้รับผลกระทบจากการทดสอบทางการศึกษาระดับชาติแบ่งเป็น 3 กลุ่ม ได้แก่ กลุ่มที่ 1 เป็นกลุ่มที่ได้รับผลกระทบจากการทดสอบทางการศึกษาระดับชาติทั้งทางบวก และทางลบระดับต่ำ (ร้อยละ 37) กลุ่มที่ 2 เป็นกลุ่มที่ได้รับผลกระทบจากการทดสอบทางการศึกษาระดับชาติทั้งทางบวกและทางลบระดับปานกลาง (ร้อยละ 11) และกลุ่มที่ 3 เป็นกลุ่มที่ได้รับผลกระทบจากการทดสอบทางการศึกษาระดับชาติทั้งทางบวกและทางลบระดับสูง (ร้อยละ 52)

2. ครูกลุ่มแฝงที่ 2 มีความต้องการจำเป็นในการพัฒนาสมรรถนะการประเมินเพื่อการเรียนรู้สูงสุด รองลงมาเป็นครูกลุ่มแฝงที่ 1 และครูกลุ่มแฝงที่ 3 ตามลำดับ

¹รองศาสตราจารย์ คณะศึกษาศาสตร์ มหาวิทยาลัยนเรศวร

¹Associate Professor., Faculty of Education, Naresuan University

3. ผลการประเมินสมรรถนะการประเมินเพื่อการเรียนรู้ของครู พบว่า ครูกลุ่มแฝงที่ 1 มีค่าเฉลี่ย T-score สมรรถนะการประเมินเพื่อการเรียนรู้ในภาพรวมและสมรรถนะการประเมินเพื่อการเรียนรู้ด้านการใช้ผลการประเมินผลการเรียนรู้ต่ำกว่าครูกลุ่มแฝงที่ 2 และครูกลุ่มแฝงที่ 3 อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และพบว่า ครูกลุ่มแฝงที่ 1 มีค่าเฉลี่ย T-score สมรรถนะการประเมินเพื่อการเรียนรู้ด้านการบูรณาการการเรียนรู้และการประเมินต่ำกว่าครูกลุ่มแฝงที่ 3 อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ส่วนด้านอื่นไม่พบความแตกต่าง

คำสำคัญ: การประเมินความต้องการจำเป็น สมรรถนะการประเมินเพื่อการเรียนรู้ ผลกระทบจากการทดสอบทางการศึกษาแห่งชาติ

Abstract

The objectives of this research were to: 1) analyze the latent class of elementary school teachers affected by national educational tests; 2) assess the need for the development of assessment competency for learning and assessment competency for learning classified by the latent class of elementary school teachers affected by national education tests. The sample was 600 elementary school teachers in the lower northern region, selected by multi-stage sampling random. The Research tools were a questionnaires regarding the need to develop assessment competency for learning, a situation-based test to measure the assessment competency for learning, and the impact of national educational tests. The data was analyzed by percentage, latent class analysis, means, standard deviation and modified priority needs index and one-way analysis of variance. the finding for this research was as follows:

1. The elementary teacher latent class in the lower northern region which were affected by the national educational tests was divided into 3 class; the first; a class of teachers affected by low positive and negative national education tests (37%), the second; a class of teachers affected by moderate positive and negative national education tests (11%) and finally; a class of teachers affected by high positive and negative national education tests (52%).

2. The second latent class needed to develop assessment for learning at the highest level, and the first latent class and the third latent class respectively later.

3. The results of the teacher learning assessment competency assessment state that mean T-score, overall assessment competency for learning, and assessment competency for learning in the use of learning assessment results of the first latent class are statistically

significantly lower than those of the second and the third latent class at the .05 level. In addition, the mean T-score assessment competency for learning in the learning and assessment integration of the first latent class were statistically significantly lower than the third latent class at the .05 level. However, no difference was found in other areas.

Keyword: Needs assessment, Competency of assessment for learning, Washback

บทนำ

การทดสอบระดับชาติจัดเป็นการทดสอบที่มีผลกระทบสูง (High stake testing) การทดสอบลักษณะนี้ต้องการนำผลไปใช้ในการตัดสินใจที่สำคัญของผู้เรียนและผู้มีส่วนเกี่ยวข้อง ผลการทดสอบจะมีผลกระทบกับผู้สอบและผู้มีส่วนเกี่ยวข้องจากการทดสอบนั้น (Madaus, 1988) โดยเฉพาะในประเทศไทยมีการนำผลการทดสอบระดับชาติของผู้เรียนมาใช้ในการตัดสินใจคุณภาพของสถานศึกษา หรือนำมาเป็นข้อมูลเพื่อใช้เลื่อนวิทยฐานะครูหรือเลื่อนขั้นเงินเดือนครู หรือการนำผลการทดสอบเข้ามหาวิทยาลัยของผู้เรียนไปเป็นตัวชี้วัดคุณภาพของสถานศึกษา นำไปสู่การเปลี่ยนแปลงพฤติกรรมของการจัดการเรียนรู้ของครู ไปในลักษณะของการติวข้อสอบ ไม่ใช่การจัดการเรียนรู้ที่ยึดมาตรฐานและตัวชี้วัดของหลักสูตรเป็นสำคัญ จากผลการวิจัยของ วันดี สมมิตร (2552) และญานิกา ลุนราศรี (2557) พบว่า ผลการทดสอบระดับชาติทำให้ครูจัดการเรียนรู้ที่เบี่ยงเบนไปจากลักษณะที่พึงประสงค์ ครูบางส่วนเปลี่ยนแปลงวิธีการจัดการเรียนรู้โดยมุ่งเน้นเพียงให้ผู้เรียนมีคะแนนสูง โดยไม่ได้สอนเพื่อให้ผู้เรียนมีความรู้ความเข้าใจเนื้อหาวิชาอย่างแท้จริง เนื่องจากผู้สอนมีส่วนได้ส่วนเสียกับผลการทดสอบของผู้เรียน

ปรากฏการณ์ที่การทดสอบมีผลกระทบต่อการจัดการเรียนรู้ของครูจนทำให้ครูปรับตัวเปลี่ยนแปลงการจัดการเรียนรู้ เรียกว่า ปรากฏการณ์วอชแบค (Washback effect) หรือผลกระทบจากการทดสอบ (Ahmad & Rao, 2012; Ghorbani & Neissari, 2015) ปรากฏการณ์ผลกระทบจากการทดสอบสามารถเป็นไปได้ทั้งทางบวกและทางลบ โดยปรากฏการณ์ผลกระทบจากการทดสอบทางบวก (Positive washback effect) หมายถึง การจัดการเรียนรู้ของครูที่นำผลการทดสอบมาใช้ในการปรับปรุงและพัฒนาการจัดการกระบวนการเรียนรู้เพื่อให้ผู้เรียนพัฒนาตามเป้าหมายของหลักสูตร ส่วนปรากฏการณ์ผลกระทบจากการทดสอบทางลบ (Negative washback effect) หมายถึง การจัดการเรียนรู้ของครูที่มุ่งเน้นการสอนเพื่อให้ผู้เรียนได้คะแนนสูงหรือมุ่งแต่การสอนเพื่อสอบ (Teach to the test) (Pan, 2009) ซึ่งผู้เรียนอาจจะได้รับความรู้ที่ไม่เพียงพอไม่สามารถประยุกต์ใช้ความรู้ไปใช้ได้ตามเป้าหมายของหลักสูตร ปาริฉัตร ปิติสุทธิ (2558) ได้ศึกษาปรากฏการณ์ผลกระทบจากการทดสอบในประเทศไทยผ่านการวิจัยเรื่อง ผลการทดสอบระดับชาติที่มีต่อครูและนักเรียน: การวิเคราะห์ปรากฏการณ์วอชแบค พบว่า ครูมีการจัดการเรียนการสอนโดยใช้การประเมินเป็นฐานอยู่ในระดับปานกลาง มีความคิดเห็นต่อการทดสอบที่มีความสำคัญสูง (การทดสอบระดับชาติ) ในด้านประโยชน์ของการทดสอบด้านมาตรฐานของข้อสอบ ด้านความยุติธรรมจากผลสอบและด้านการใช้ผลการทดสอบโดยเฉลี่ยอยู่ในระดับปานกลาง ครูรู้สึกมีอิสระในการจัดการเรียนการสอนภายใต้แรงกดดันจาก

ผู้เกี่ยวข้องเฉลี่ยอยู่ในระดับปานกลาง ส่วนพฤติกรรมการจัดการเรียนการสอนสะท้อนผลกระทบจากการทดสอบทางบวกอยู่ในระดับค่อนข้างมาก ในขณะที่ทางลบอยู่ในระดับปานกลาง และผลที่เกิดขึ้นกับผู้เรียนตามการรับรู้ของครูอยู่ในระดับปานกลางเช่นเดียวกับผลการทดสอบความแตกต่างของตัวแปรกลุ่มครูที่อยู่ในบริบทต่างกัน พบว่า ครูสังกัดเอกชนมีการจัดการเรียนรู้โดยใช้การประเมินเป็นฐาน มีความคิดเห็นต่อการทดสอบที่มีความสำคัญสูง มีพฤติกรรมการสอนที่สะท้อนผลกระทบจากการทดสอบทางบวกและทางลบสูงกว่าครูสังกัด สพฐ. นอกจากนี้ครูในโรงเรียนที่มีชื่อเสียงจะมีอิสระในการจัดการเรียนการสอนและพฤติกรรมการสอนที่สะท้อนผลกระทบจากการทดสอบทางบวกสูงกว่าครูในโรงเรียนที่ไม่มีชื่อเสียง นอกจากนี้ข้อมูลเชิงคุณภาพยืนยันผลการวิจัยเชิงปริมาณว่า ครูที่เห็นความสำคัญของการทดสอบระดับชาติ การสอนโดยใช้แนวความคิดการประเมินเป็นฐาน และความรู้สึกว่าตนมีอิสระในการจัดการเรียนการสอนเป็นครูที่มีแนวโน้มจะมีพฤติกรรมการสอนเพื่อให้ผู้เรียนรู้มากกว่าเพื่อสอบ

จากผลการวิจัยข้างต้นมีข้อสรุปที่น่าสนใจว่า ถ้าครูเห็นความสำคัญของการทดสอบระดับชาติ การสอนโดยใช้แนวความคิดการประเมินเป็นฐาน และความรู้สึกว่าตนมีอิสระในการจัดการเรียนการสอน ครูมีแนวโน้มที่จะมีพฤติกรรมการใช้ผลการประเมินเพื่อพัฒนาการเรียนรู้ของผู้เรียนมากกว่าการตรวจสอบการจัดการเรียนรู้โดยใช้การประเมินเป็นฐานเป็นแนวคิดหนึ่งของการประเมินเพื่อพัฒนาการเรียนรู้ (Assessment for learning) ซึ่งมุ่งเน้นการนำสารสนเทศที่ได้จากการประเมินการเรียนรู้มาวินิจฉัยจุดแข็ง จุดอ่อนและนำมาออกแบบ ปรับปรุง และพัฒนาการจัดการเรียนการสอนเพื่อส่งเสริมให้ผู้เรียนบรรลุตามเป้าหมาย โดยมีแนวคิดสำคัญสามประการ ได้แก่ ประการแรกในการออกแบบการจัดการเรียนรู้ต้องมีการออกแบบการประเมินผลการเรียนรู้ก่อนที่จะพัฒนาสื่อหรือออกแบบกิจกรรมการเรียนรู้ ประการที่สอง ให้ความสำคัญกับการประเมินก่อนการจัดการเรียนรู้ และประการที่สาม ใช้วิธีการประเมินที่หลากหลายเพื่อรวบรวมข้อมูลการเรียนรู้ของผู้เรียนอย่างแท้จริง (Wilson, 2009) ซึ่งการประเมินเพื่อพัฒนาการเรียนรู้มี ยุทธวิธีสำคัญด้วยกัน 8 ยุทธวิธี (Brookhart, 2015; Chappuis, Stiggins, Chappuis & Arter, 2012 อ้างถึงในสร้อยญา จันทรชูลกุล, 2561) ได้แก่ (1) กำหนดเป้าหมายการเรียนรู้ที่ชัดเจน (2) ให้ตัวอย่างงานที่ดีและไม่ดี (3) ให้ข้อมูลย้อนกลับ (4) สอนให้ผู้เรียนประเมินตนเองและกำหนดเป้าหมาย (5) ออกแบบบทเรียนที่ตั้งอยู่บนเป้าหมายการเรียนรู้ที่สำคัญเพียงหนึ่งเดียว (6) บูรณาการการจัดการเรียนรู้และการประเมิน (7) สอนผู้เรียนให้มุ่งความสนใจในการปรับปรุงแก้ไข และ (8) ส่งเสริมผู้เรียนให้มีการสะท้อนตนเองและแบ่งปันการเรียนรู้

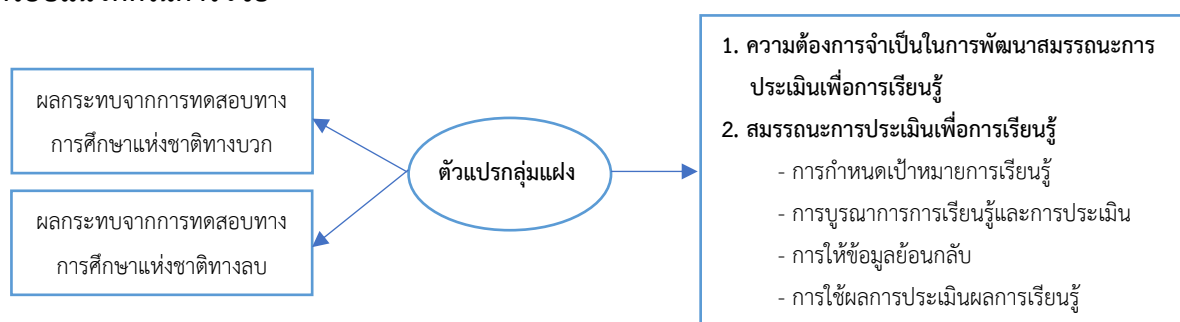
การส่งเสริมให้ครูมีสมรรถนะการประเมินเพื่อพัฒนาการเรียนรู้ซึ่งเป็นสมรรถนะทางวิชาการที่สำคัญของครูเป็นการปรับพฤติกรรมครูทั้งด้านพุทธิพิสัย จิตพิสัย และทักษะพิสัยไปพร้อมกัน ดังนั้นจึงจำเป็นต้องออกแบบวิธีการส่งเสริมครูให้มีความเชี่ยวชาญที่สามารถทำให้ครูเกิดความรู้ความเข้าใจ มีทักษะ และเกิดความมั่นใจที่จะออกแบบการจัดการเรียนรู้ การวางแผนพัฒนาครูจึงจำเป็นต้องมีสารสนเทศที่ลุ่มลึกเกี่ยวกับความต้องการจำเป็นในการพัฒนาสมรรถนะการประเมินเพื่อการเรียนรู้ของครูที่ได้รับผลกระทบจากการทดสอบทางการศึกษาแห่งชาติ การวิเคราะห์กลุ่มแฝง (Latent class analysis) เป็นวิธีการทางสถิติที่จะช่วยจัดบุคคลเข้ากลุ่มแฝง ซึ่งวัดได้จากตัวแปรสังเกตได้ โดยมีเงื่อนไขในการจัดบุคคลเข้ากลุ่มแฝงจากลักษณะร่วมกันของแต่ละบุคคลที่เป็นลักษณะที่มีความเฉพาะเจาะจงอย่างครบถ้วนสมบูรณ์ ดังนั้นการวิจัยครั้งนี้จึงเริ่มจากการ

วิเคราะห์กลุ่มแฟงของครูระดับประถมศึกษาที่ได้รับผลกระทบจากการทดสอบทางการศึกษาระดับชาติ หลังจากนั้นจึงประเมินความต้องการจำเป็นในการพัฒนาสมรรถนะการประเมินเพื่อการเรียนรู้ รวมทั้งประเมินสมรรถนะการประเมินเพื่อการเรียนรู้ของครูจำแนกตามกลุ่มแฟงการใช้วิธีการทางสถิติการวิเคราะห์กลุ่มแฟงจะช่วยจำแนกกลุ่มครูที่ได้รับผลกระทบจากการทดสอบทางการศึกษาแห่งชาติได้ชัดเจน สารสนเทศที่ได้ตามกลุ่มแฟงของครูจึงมีคุณค่ายิ่งที่จะนำไปเป็นข้อมูลในการวางแผนพัฒนาครูได้สอดคล้องกับความต้องการจำเป็นของครูในแต่ละกลุ่ม โดยมีเป้าหมายในการยกระดับการจัดการเรียนรู้เพื่อพัฒนาผู้เรียนเป็นอันดับสุดท้าย

วัตถุประสงค์

1. เพื่อวิเคราะห์กลุ่มแฟงของครูระดับประถมศึกษาที่ได้รับผลกระทบจากการทดสอบทางการศึกษาระดับชาติ
2. เพื่อประเมินความต้องการจำเป็นในการพัฒนาสมรรถนะการประเมินเพื่อการเรียนรู้จำแนกตามกลุ่มแฟงของครูระดับประถมศึกษาที่ได้รับผลกระทบจากการทดสอบทางการศึกษาระดับชาติ
3. เพื่อประเมินสมรรถนะการประเมินเพื่อการเรียนรู้จำแนกตามกลุ่มแฟงของครูระดับประถมศึกษาที่ได้รับผลกระทบจากการทดสอบทางการศึกษาระดับชาติ

กรอบแนวคิดในการวิจัย



ภาพ 1 กรอบแนวคิดในการวิจัย

วิธีดำเนินการวิจัย

การวิเคราะห์ความต้องการจำเป็นในการพัฒนาสมรรถนะการประเมินเพื่อการเรียนรู้จำแนกตามกลุ่มแฟงของครูประถมศึกษาที่ได้รับผลกระทบจากการทดสอบทางการศึกษาระดับชาติ มีวิธีดำเนินการวิจัย ดังนี้

ประชากรและกลุ่มตัวอย่าง

ประชากรในการวิจัย คือ ครูระดับประถมศึกษาชั้นปีที่ 6 ในเขตพื้นที่ภาคเหนือตอนล่าง ได้แก่ จังหวัดกำแพงเพชร พิษณุโลก พิจิตร เพชรบูรณ์ สุโขทัย อุตรดิตถ์ นครสวรรค์ อุทัยธานี และตาก โดยกลุ่มตัวอย่างเป็นครูระดับประถมศึกษาชั้นปีที่ 6 จำนวน 600 คน ในเขตพื้นที่ภาคเหนือตอนล่าง 1 ได้แก่ จังหวัดพิษณุโลก และพื้นที่ภาคเหนือตอนล่าง 2 ได้แก่ จังหวัดกำแพงเพชร ซึ่งได้มาจากการสุ่มแบบหลาย

ขั้นตอน และกำหนดขนาดกลุ่มตัวอย่างโดยคำนึงถึงวิธีการทางสถิติการวิเคราะห์กลุ่มแฝง ที่มีการประมาณค่าพารามิเตอร์ประมาณ 20-30 ตัว โดยจำเป็นต้องใช้ขนาดกลุ่มตัวอย่างใหญ่พอสมควรกำหนดเป็น 20 เท่าของพารามิเตอร์ (Hair et al., 2010) ซึ่งจะใช้กลุ่มตัวอย่าง 600 คน

เครื่องมือในการวิจัย

1. แบบสอบถามข้อมูลทั่วไปของผู้ตอบ แบบตรวจสอบรายการ สอบถามเกี่ยวกับเพศ อายุ การศึกษาสูงสุด ระดับวิทยฐานะ กลุ่มสาระการเรียนรู้ที่สอน ประสบการณ์การสอน ค่าเฉลี่ย O-NET

2. แบบสอบถามความต้องการจำเป็นในการพัฒนาสมรรถนะการประเมินเพื่อการเรียนรู้ แบบมาตราประมาณค่า 5 ระดับ แบบตอบสนองคู่ จำนวน 13 ข้อ มีความตรงเชิงเนื้อหาประเด็นความสอดคล้องเท่ากับ .80-1.00 และความชัดเจนเท่ากับ .80-1.00 ส่วนของสภาพจริงมีอำนาจจำแนกคำนวณจากสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สันเท่ากับ .274-.898 มีความเที่ยงเชิงความสอดคล้องภายในซึ่งคำนวณจากสูตรสัมประสิทธิ์แอลฟาของครอนบาคเท่ากับ .954 ส่วนของสภาพที่คาดหวังมีอำนาจจำแนกคำนวณจากสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สันเท่ากับ .499-.855 มีความเที่ยงเชิงความสอดคล้องภายในซึ่งคำนวณจากสูตรสัมประสิทธิ์แอลฟาของครอนบาคเท่ากับ .944

3. แบบทดสอบเชิงสถานการณ์เพื่อวัดสมรรถนะการประเมินเพื่อการเรียนรู้ เป็นแบบทดสอบเชิงสถานการณ์แบบเลือกตอบ ตอบถูกได้ 1 ตอบผิดได้ 0 จำนวน 15 ข้อ มีความตรงเชิงเนื้อหาประเด็นความสอดคล้องเท่ากับ .80-1.00 และชัดเจนเท่ากับ .80-1.00 มีความยากง่ายคำนวณจากสัดส่วนของคนตอบถูกเท่ากับ .433-.800 มีอำนาจจำแนกคำนวณจากผลต่างของสัดส่วนของคนตอบถูกในกลุ่มสูงกับกลุ่มต่ำเท่ากับ .288-.871 มีความเที่ยงเชิงความสอดคล้องภายในคำนวณจากสัมประสิทธิ์ความเที่ยงแบบ KR 20 เท่ากับ .763 มีความตรงเชิงโครงสร้างคำนวณด้วยการวิเคราะห์องค์ประกอบเชิงยืนยันโดยโมเดลการวัดสมรรถนะการประเมินเพื่อการเรียนรู้ของครูมีความสอดคล้องกับข้อมูลเชิงประจักษ์ โดยมีค่ามีค่าไคสแควร์เท่ากับ 1.123 p value เท่ากับ .2892

4. แบบวัดผลกระทบจากการทดสอบทางการศึกษาระดับชาติ เป็นแบบมาตราประมาณค่า 5 ระดับ เป็นผลกระทบทางบวก 8 ข้อ และทางลบ 8 ข้อมีความตรงเชิงเนื้อหาประเด็นความสอดคล้องเท่ากับ .80-1.00 และชัดเจนเท่ากับ .60-1.00 ส่วนผลกระทบทางบวกมีอำนาจจำแนกคำนวณจากสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สันเท่ากับ .381-.911 มีความเที่ยงเชิงความสอดคล้องภายในซึ่งคำนวณจากสูตรสัมประสิทธิ์แอลฟาของครอนบาคเท่ากับ .907 ส่วนผลกระทบทางลบมีอำนาจจำแนกคำนวณจากสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สันเท่ากับ .673-.793 มีความเที่ยงเชิงความสอดคล้องภายในซึ่งคำนวณจากสูตรสัมประสิทธิ์แอลฟาของครอนบาคเท่ากับ .915

การออกแบบการวิเคราะห์ข้อมูล

1. วิเคราะห์ข้อมูลทั่วไป ได้แก่ เพศ อายุ การศึกษาสูงสุด ระดับวิทยฐานะ กลุ่มสาระการเรียนรู้ที่สอน ประสบการณ์การสอน และค่าเฉลี่ย O-NET รวมทุกวิชา จำแนกตามจังหวัด ด้วยการแจกแจงความถี่ และนำเสนอด้วยคำร้อยละ นำเสนอในรูปแบบของตารางประกอบความเรียง

2. วิเคราะห์เพื่อจำแนกกลุ่มครูที่ได้รับผลกระทบจากการทดสอบทางการศึกษาระดับชาติ ทางบวก และทางลบโดยการวิเคราะห์กลุ่มแฝง (Latent class analysis) และนำเสนอในรูปตารางประกอบความเรียง

3. วิเคราะห์ความต้องการจำเป็นในการพัฒนาสมรรถนะการประเมินการเรียนรู้จำแนกตามกลุ่มแฝง ของครู โดยใช้ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน และดัชนี PNI แบบปรับปรุง จำแนกตามจังหวัดและนำเสนอใน รูปตารางประกอบความเรียง และวิเคราะห์เพื่อทดสอบความแตกต่างของค่าเฉลี่ยของกลุ่มประชากรด้วยการ วิเคราะห์ความแปรปรวนแบบทางเดียว และการวิเคราะห์แตกต่างรายคู่ด้วย Dunnett T3

ผลการวิจัย

เมื่อพิจารณาข้อมูลทั่วไปพบว่า ครูระดับประถมศึกษาชั้นปีที่ 6 ในพื้นที่ภาคเหนือตอนล่างที่ตอบ แบบสอบถามส่วนใหญ่เป็นเพศหญิง คิดเป็นร้อยละ 76.83 อายุตั้งแต่ 51 ปีขึ้นไป คิดเป็นร้อยละ 32.17 มีการศึกษาสูงสุดอยู่ในระดับปริญญาตรี คิดเป็นร้อยละ 62.50 เมื่อพิจารณาถึงระดับวิทยฐานะ พบว่า ครูส่วนใหญ่มีระดับวิทยฐานะในระดับชำนาญการพิเศษ คิดเป็นร้อยละ 32 มีเพียงร้อยละ 2.00 เท่านั้นที่มีวิทยฐานะ ระดับเชี่ยวชาญ วิชาที่สอนพบว่า ส่วนใหญ่สอนวิชาคณิตศาสตร์ คิดเป็นร้อยละ 52.50 รองลงมาคือ ภาษาไทย คิดเป็นร้อยละ 42.33 ประสบการณ์การสอนส่วนใหญ่ 26 ปีขึ้นไป คิดเป็นร้อยละ 33.33 เท่า ๆ กับ ประสบการณ์สอนน้อยกว่า 6 ปี คิดเป็นร้อยละ 32.33 เมื่อพิจารณาถึงค่าเฉลี่ย O-NET พบว่า ครูระดับประถมศึกษาชั้นปีที่ 6 ที่ตอบแบบสอบถามส่วนใหญ่สังกัดโรงเรียนที่มีค่าเฉลี่ย O-NET ต่ำกว่าระดับประเทศ คิดเป็น ร้อยละ 44.50 ผลการวิจัยนำเสนอตามวัตถุประสงค์การวิจัยดังนี้

ตารางที่ 1 ผลการวิเคราะห์กลุ่มแฝงผลกระทบการทดสอบทางการศึกษาแห่งชาติ

ผลกระทบการทดสอบทาง		พิษณุโลก (n=360)			กำแพงเพชร (n=240)			รวม (n=600)		
การศึกษาแห่งชาติ		n	$\bar{X}$	S.D.	n	$\bar{X}$	S.D.	n	$\bar{X}$	S.D.
ทางบวก	class 1	132	3.34	0.64	90.00	3.34	0.48	222	3.35	.58
	class 2	34	4.01	0.24	32.00	3.94	0.31	66	3.98	.28
	class 3	194	4.49	0.37	118.00	4.42	0.36	312	4.46	.37
ทางลบ	class 1	132	3.25	0.56	90.00	3.28	0.47	222	3.26	.52
	class 2	34	3.83	0.23	32.00	3.95	0.11	66	3.89	.19
	class 3	194	4.47	0.39	118.00	4.44	0.37	312	4.46	.38

Log likelihood=111.865 number of free parameters=10

AIC=243.730 BIC=287.700 Adj BIC = 255.952 entropy=0.998

จากตารางที่ 1 ผลการวิเคราะห์กลุ่มแฝงครูระดับประถมศึกษาชั้นปีที่ 6 ในภาคเหนือตอนล่างจำนวน 600 คน ที่ได้รับผลกระทบจากการทดสอบทางการศึกษาระดับชาติ พบว่า โมเดลที่มี 3 กลุ่ม มีความสอดคล้องกับข้อมูลเชิงประจักษ์และเป็นโมเดลที่ประหยัดที่สุด ($AIC=243.730$ $BIC=287.700$) มีค่าความน่าจะเป็นที่จำแนกกลุ่มได้ถูกต้องชัดเจน (entropy) เท่ากับ 0.998 ซึ่งมีค่าเข้าใกล้ 1 และเมื่อพิจารณาค่าความน่าจะเป็นแบบมีเงื่อนไขในการจำแนกกลุ่มครูเข้ากลุ่มแฝงผลกระทบจากการทดสอบทางการศึกษาระดับชาติ พบว่า

กลุ่มแฝงที่ 1 (class 1) ประกอบด้วยครูจำนวน 222 คน คิดเป็นสัดส่วนเท่ากับ 0.37 (ร้อยละ 37) ของจำนวนครูทั้งหมด 600 คน มีความน่าจะเป็นเฉลี่ยของกลุ่มเท่ากับ 1.00 เป็นกลุ่มที่ได้รับผลกระทบจากการทดสอบทางการศึกษาระดับชาติทั้งทางบวก และทางลบระดับต่ำ มีค่าเฉลี่ยเท่ากับ 3.35 ± 0.58 และ 3.26 ± 0.52 ตามลำดับ

กลุ่มแฝงที่ 2 (class 2) ประกอบด้วยครูจำนวน 66 คน คิดเป็นสัดส่วนเท่ากับ 0.11 (ร้อยละ 11) ของจำนวนครูทั้งหมด 600 คน มีความน่าจะเป็นเฉลี่ยของกลุ่มเท่ากับ 0.998 เป็นกลุ่มที่ได้รับผลกระทบจากการทดสอบทางการศึกษาระดับชาติทั้งทางบวกและทางลบระดับปานกลาง มีค่าเฉลี่ยเท่ากับ 3.98 ± 0.28 และ 3.89 ± 0.19 ตามลำดับ

กลุ่มแฝงที่ 3 (class 3) ประกอบด้วยครูจำนวน 312 คน คิดเป็นสัดส่วนเท่ากับ 0.52 (ร้อยละ 52) ของจำนวนครูทั้งหมด 600 คน มีความน่าจะเป็นเฉลี่ยของกลุ่มเท่ากับ 1.00 เป็นกลุ่มที่ได้รับผลกระทบจากการทดสอบทางการศึกษาระดับชาติทั้งทางบวกและทางลบระดับสูง มีค่าเฉลี่ยเท่ากับ 4.46 ± 0.37 และ 4.46 ± 0.38 ตามลำดับ

ตารางที่ 2 ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ความแตกต่างของค่าเฉลี่ย และดัชนี PNI ความต้องการจำเป็นในการพัฒนาสมรรถนะการประเมินเพื่อการเรียนรู้ จำแนกตามกลุ่มแฝงของครู ($n=600$)

ที่	สมรรถนะการประเมินเพื่อการเรียนรู้	class	สภาพที่คาดหวัง		สภาพจริง		Mean different	PNI
			$\bar{X}$	S.D.	$\bar{X}$	S.D.		
1	การกำหนดเป้าหมายการเรียนรู้	1 (222)	3.73	0.49	3.57	0.49	0.1577	0.0442
		2 (66)	3.86	0.71	3.69	0.50	0.1780	0.0483
		3 (312)	4.06	0.40	4.02	0.41	0.0457	0.0114
		รวม (600)	3.92	0.50	3.82	0.50	0.1017	0.0266
2	การบูรณาการการเรียนรู้และการประเมิน	1 (222)	3.83	0.52	3.63	0.57	0.2005	0.0552
		2 (66)	4.09	0.55	3.71	0.38	0.3788	0.1020
		3 (312)	4.28	0.44	4.15	0.47	0.1330	0.0321
		รวม (600)	4.09	0.53	3.91	0.56	0.1850	0.0473
3	การให้ข้อมูลย้อนกลับ	1 (222)	3.75	0.59	3.55	0.49	0.1982	0.0558
		2 (66)	4.12	0.52	3.58	0.43	0.5455	0.1525
		3 (312)	4.23	0.52	4.06	0.50	0.1635	0.0402
		รวม (600)	4.04	0.59	3.82	0.55	0.2183	0.0572

ตารางที่ 2 (ต่อ)

ที่	สมรรถนะการประเมิน เพื่อการเรียนรู้	class	สภาพที่คาดหวัง		สภาพจริง		Mean different	PNI
			$\bar{X}$	S.D.	$\bar{X}$	S.D.		
4	การใช้ผลการประเมินผล การเรียนรู้	1 (222)	3.89	0.63	3.64	0.60	0.2568	0.0706
		2 (66)	4.32	0.62	3.68	0.62	0.6364	0.1728
		3 (312)	4.30	0.51	4.29	0.53	0.0032	0.0007
		รวม (600)	4.15	0.60	3.98	0.65	0.1633	0.0410
รวม		1 (222)	3.79	0.44	3.59	0.45	0.1954	0.0544
		2 (66)	4.06	0.54	3.67	0.41	0.3951	0.1077
		3 (312)	4.20	0.37	4.11	0.38	0.0922	0.0224
		รวม (600)	4.03	0.46	3.87	0.48	0.1637	0.0423

ตารางที่ 2 พบว่า ครูกลุ่มแฝงที่ 2 ครูที่ได้รับผลกระทบจากการทดสอบทางการศึกษาแห่งชาติทางบกและทางลบนกลาง มีความต้องการจำเป็นในการพัฒนาสมรรถนะการประเมินเพื่อการเรียนรู้ทั้งภาพรวมและรายด้านสูงสุด รองลงมาเป็นครูกลุ่มแฝงที่ 1 ครูที่ได้รับผลกระทบจากการทดสอบทางการศึกษาแห่งชาติทางบกและทางลต่ำ และครูกลุ่มแฝงที่ 3 ครูที่ได้รับผลกระทบจากการทดสอบทางการศึกษาแห่งชาติทางบกและทางลสูง ตามลำดับ

ตารางที่ 3 ผลการวิเคราะห์ความแปรปรวนเพื่อทดสอบความแตกต่างของค่าเฉลี่ยสมรรถนะการประเมินเพื่อการเรียนรู้ของครูจำแนกตามกลุ่มแฝงของครู

ที่	องค์ประกอบสมรรถนะการ ประเมินเพื่อการเรียนรู้	class	$\bar{X}$	SD.	F	Sig
			T-score	T-score		
1	การกำหนดเป้าหมายการเรียนรู้	1 (222)	49.68	9.61	1.9029	0.1500
		2 (66)	52.26	6.70		
		3 (312)	49.75	10.79		
2	การบูรณาการการเรียนรู้และ การประเมิน	1 (222)	48.50	10.97	4.1036*	0.0170
		2 (66)	50.35	8.38		
		3 (312)	50.99	9.48		
3	การให้ข้อมูลย้อนกลับ	1 (222)	49.03	10.51	2.4749	0.0850
		2 (66)	52.00	8.36		
		3 (312)	50.26	9.90		

ตารางที่ 3 (ต่อ)

ที่	องค์ประกอบสมรรถนะการ ประเมินเพื่อการเรียนรู้	class	$\bar{X}$ T-score	SD. T-score	F	Sig
4	การใช้ผลการประเมินผลการ เรียนรู้	1 (222)	47.73	9.31	9.3536*	0.0001
		2 (66)	51.59	7.04		
		3 (312)	51.28	10.71		
รวม		1 (222)	48.13	11.03	7.1170*	0.0009
		2 (66)	52.50	5.98		
		3 (312)	50.80	9.70		

*p<.05 หมายถึง เหตุ นำเสนอด้วย T-SCORE เนื่องจากคะแนนเต็มแต่ละองค์ประกอบย่อยไม่เท่ากัน

ตารางที่ 3 พบว่า ครูกลุ่มแฟงที่ 1 ครูที่ได้รับผลกระทบจากการทดสอบทางการศึกษาแห่งชาติทางบกและทางลต่า มีค่าเฉลี่ย T-score สมรรถนะการประเมินเพื่อการเรียนรู้เท่ากับ 48.13 ± 11.03 ครูกลุ่มแฟงที่ 2 ครูที่ได้รับผลกระทบจากการทดสอบทางการศึกษาแห่งชาติทางบกและทางลปานกลาง มีค่าเฉลี่ย T-score สมรรถนะการประเมินเพื่อการเรียนรู้เท่ากับ 52.50 ± 5.98 และ ครูกลุ่มแฟงที่ 3 ครูที่ได้รับผลกระทบจากการทดสอบทางการศึกษาแห่งชาติทางบกและทางลสูง มีค่าเฉลี่ย T-score สมรรถนะการประเมินเพื่อการเรียนรู้เท่ากับ 50.80 ± 9.70 ซึ่งแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 กล่าวคือ มีค่าเฉลี่ย T-score สมรรถนะการประเมินเพื่อการเรียนรู้แตกต่างกันอย่างน้อยหนึ่งคู่ ซึ่งจะทำให้การทดสอบรายคู่ต่อไป และเมื่อพิจารณารายองค์ประกอบ พบว่า มีสมรรถนะการประเมินเพื่อการเรียนรู้ด้านการบูรณาการการเรียนรู้และการประเมิน และด้านการใช้ผลการประเมินผลการเรียนรู้ ที่พบความแตกต่าง ซึ่งจะทำให้การทดสอบรายคู่ต่อไปด้วย Dunnett T3

ตารางที่ 4 ผลการวิเคราะห์ความแปรปรวนเพื่อทดสอบความแตกต่างของค่าเฉลี่ยสมรรถนะการประเมินเพื่อการเรียนรู้ของครูจำแนกตามกลุ่มแฟงของครูรายคู่

ที่	องค์ประกอบสมรรถนะการ ประเมินเพื่อการเรียนรู้	class	$\bar{X}$ T-score	S.D. T-score	Class		
					1 (222)	2 (66)	3 (312)
2	การบูรณาการการเรียนรู้และ การประเมิน	1 (222)	48.50	10.97	2.49*		
		2 (66)	50.35	8.38			
		3 (312)	50.99	9.48			
4	การใช้ผลการประเมินผลการ เรียนรู้	1 (222)	47.73	9.31	3.86*		
		2 (66)	51.59	7.04			
		3 (312)	51.28	10.71			
รวม		1 (222)	48.13	11.03	4.37*		
		2 (66)	52.50	5.98			
		3 (312)	50.80	9.70			

*p<.05

ตารางที่ 4 พบว่า ครูกลุ่มแฝงที่ 1 ครูที่ได้รับผลกระทบจากการทดสอบทางการศึกษาแห่งชาติทางบกและทางลตต่ำ มีค่าเฉลี่ย T-score สมรรถนะการประเมินเพื่อการเรียนรู้ในภาพรวมและสมรรถนะการประเมินเพื่อการเรียนรู้ด้านการใช้ผลการประเมินผลการเรียนรู้ต่ำกว่าครูกลุ่มแฝงที่ 2 ครูที่ได้รับผลกระทบจากการทดสอบทางการศึกษาแห่งชาติทางบกและทางลตปานกลาง และกลุ่มแฝงที่ 3 ครูที่ได้รับผลกระทบจากการทดสอบทางการศึกษาแห่งชาติทางบกและทางลตสูง อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และพบว่า ครูกลุ่มแฝงที่ 1 ครูที่ได้รับผลกระทบจากการทดสอบทางการศึกษาแห่งชาติทางบกและทางลตต่ำ มีค่าเฉลี่ย T-score สมรรถนะการประเมินเพื่อการเรียนรู้ด้านการบูรณาการการเรียนรู้และการประเมินต่ำกว่าครูกลุ่มแฝงที่ 3 ครูที่ได้รับผลกระทบจากการทดสอบทางการศึกษาแห่งชาติทางบกและทางลตสูง อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

สรุปและอภิปรายผล

1. ผลการวิเคราะห์กลุ่มแฝงครูระดับประถมศึกษาชั้นปีที่ 6 ในภาคเหนือตอนล่างจำนวน 600 คน ที่ได้รับผลกระทบจากการทดสอบทางการศึกษาระดับชาติ พบว่า โมเดลที่มี 3 กลุ่ม มีความสอดคล้องกับข้อมูลเชิงประจักษ์และเป็นโมเดลที่ประหยัดที่สุด ($AIC=243.730$ $BIC=287.700$) มีค่าความน่าจะเป็นที่จำแนกกลุ่มได้ถูกต้องชัดเจน (entropy) เท่ากับ 0.998 ซึ่งมีค่าเข้าใกล้ 1 และเมื่อพิจารณาค่าความน่าจะเป็นแบบมีเงื่อนไขในการจำแนกกลุ่มครูเข้ากลุ่มแฝงผลกระทบจากการทดสอบทางการศึกษาระดับชาติ พบว่า

กลุ่มแฝงที่ 1 (class 1) ประกอบด้วยครูจำนวน 222 คน คิดเป็นสัดส่วนเท่ากับ 0.37 (ร้อยละ 37) ของจำนวนครูทั้งหมด 600 คน มีความน่าจะเป็นเฉลี่ยของกลุ่มเท่ากับ 1.00 เป็นกลุ่มที่ได้รับผลกระทบจากการทดสอบทางการศึกษาระดับชาติทั้งทางบกและทางลตระดับต่ำ มีค่าเฉลี่ยเท่ากับ 3.35 ± 0.58 และ 3.26 ± 0.52 ตามลำดับ

กลุ่มแฝงที่ 2 (class 2) ประกอบด้วยครูจำนวน 66 คน คิดเป็นสัดส่วนเท่ากับ 0.11 (ร้อยละ 11) ของจำนวนครูทั้งหมด 600 คน มีความน่าจะเป็นเฉลี่ยของกลุ่มเท่ากับ 0.998 เป็นกลุ่มที่ได้รับผลกระทบจากการทดสอบทางการศึกษาระดับชาติทั้งทางบกและทางลตระดับปานกลาง มีค่าเฉลี่ยเท่ากับ 3.98 ± 0.28 และ 3.89 ± 0.19 ตามลำดับ

กลุ่มแฝงที่ 3 (class 3) ประกอบด้วยครูจำนวน 312 คน คิดเป็นสัดส่วนเท่ากับ 0.52 (ร้อยละ 52) ของจำนวนครูทั้งหมด 600 คน มีความน่าจะเป็นเฉลี่ยของกลุ่มเท่ากับ 1.00 เป็นกลุ่มที่ได้รับผลกระทบจากการทดสอบทางการศึกษาระดับชาติทั้งทางบกและทางลตระดับสูง มีค่าเฉลี่ยเท่ากับ 4.46 ± 0.37 และ 4.46 ± 0.38 ตามลำดับ

2. ผลการประเมินความต้องการจำเป็นในการพัฒนาสมรรถนะการประเมินเพื่อการเรียนรู้ด้านการกำหนดเป้าหมายการเรียนรู้ สรุปได้ดังนี้

ครูกลุ่มแฝงที่ 2 ครูที่ได้รับผลกระทบจากการทดสอบทางการศึกษาแห่งชาติทางบกและทางลตปานกลาง มีความต้องการจำเป็นในการพัฒนาสมรรถนะการประเมินเพื่อการเรียนรู้ทั้งภาพรวมและรายด้านสูงสุด รองลงมาเป็นครูกลุ่มแฝงที่ 1 ครูที่ได้รับผลกระทบจากการทดสอบทางการศึกษาแห่งชาติทางบกและ

ทางลบต่ำ และครูกลุ่มแฟงที่ 3 ครูที่ได้รับผลกระทบจากการทดสอบทางการศึกษาแห่งชาติทางบกและทางลบสูง ตามลำดับ

3. ผลการประเมินสมรรถนะการประเมินเพื่อการเรียนรู้ของครู สรุปได้ว่า ครูกลุ่มแฟงที่ 1 ครูที่ได้รับผลกระทบจากการทดสอบทางการศึกษาแห่งชาติทางบกและทางลบต่ำ มีค่าเฉลี่ย T-score สมรรถนะการประเมินเพื่อการเรียนรู้เท่ากับ 48.13 ± 11.03 ครูกลุ่มแฟงที่ 2 ครูที่ได้รับผลกระทบจากการทดสอบทางการศึกษาแห่งชาติทางบกและทางลบปานกลาง มีค่าเฉลี่ย T-score สมรรถนะการประเมินเพื่อการเรียนรู้เท่ากับ 52.50 ± 5.98 และ ครูกลุ่มแฟงที่ 3 ครูที่ได้รับผลกระทบจากการทดสอบทางการศึกษาแห่งชาติทางบกและทางลบสูง มีค่าเฉลี่ย T-score สมรรถนะการประเมินเพื่อการเรียนรู้เท่ากับ 50.80 ± 9.70 ซึ่งแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 โดยครูกลุ่มแฟงที่ 1 มีค่าเฉลี่ย T-score สมรรถนะการประเมินเพื่อการเรียนรู้ในภาพรวม และสมรรถนะการประเมินเพื่อการเรียนรู้ด้านการใช้ผลการประเมินผลการเรียนรู้ต่ำกว่าครูกลุ่มแฟงที่ 2 และครูกลุ่มแฟงที่ 3 อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และพบว่า ครูกลุ่มแฟงที่ 1 มีค่าเฉลี่ย T-score สมรรถนะการประเมินเพื่อการเรียนรู้ด้านการบูรณาการการเรียนรู้และการประเมินต่ำกว่าครูกลุ่มแฟงที่ 3 อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

โดยมีประเด็นที่น่าสนใจอภิปรายผลการวิจัยดังนี้

1. กลุ่มแฟงครูระดับประถมศึกษาในภาคเหนือตอนล่างที่ได้รับผลกระทบจากการทดสอบทางการศึกษาระดับชาติแบ่งเป็น 3 กลุ่ม กลุ่มแฟงที่ 3 เป็นกลุ่มที่ได้รับผลกระทบจากการทดสอบทางการศึกษาระดับชาติทั้งทางบกและทางลบระดับสูงมีจำนวนครูมากที่สุดคิดเป็นร้อยละ 52 รองลงมาเป็นกลุ่มแฟงที่ 1 เป็นกลุ่มที่ได้รับผลกระทบจากการทดสอบทางการศึกษาระดับชาติทั้งทางบกและทางลบระดับต่ำคิดเป็นร้อยละ 37 และกลุ่มที่ 2 เป็นกลุ่มที่ได้รับผลกระทบจากการทดสอบทางการศึกษาระดับชาติทั้งทางบกและทางลบระดับปานกลาง มีเพียงร้อยละ 11 ซึ่งกลุ่มครูแต่ละกลุ่มต่างได้รับผลกระทบทั้งทางบกและทางลบไปพร้อมกัน สอดคล้องกับผลการวิจัยของปาริฉัตร ปิติสุทธิ (2558) ที่พบว่า กลุ่มครูเอกชนครูมีพฤติกรรมการสอนที่สะท้อนวอชแบบคหรือผลกระทบจากการทดสอบทางบกสูง และขณะเดียวกันก็มีพฤติกรรมการสอนที่สะท้อนวอชแบบคหรือผลกระทบจากการทดสอบทางลบสูงเช่นเดียวกัน ซึ่งในการจัดการเรียนการสอนของครูจะได้รับผลกระทบจากการทดสอบที่มีความสำคัญสูง (High stage test) ซึ่งในทางทฤษฎีและการศึกษาเชิงประจักษ์เชื่อว่าผลกระทบจากการทดสอบเกิดขึ้นได้ทั้งทางบกและทางลบ (Glover, 2014; Yu & Jie, 2013; Munoz & Alvarez, 2010) และนอกจากนี้ Kennedy & Lui (2013) ยังพบว่าผลกระทบจากการทดสอบอาจจะเป็นทางบกหรือทางลบ หรืออาจไม่เกิดก็ได้หากผู้มีส่วนเกี่ยวข้องมองว่าการสอบไม่ได้มีผลต่อการจัดการเรียนการสอนและการเรียนรู้ของนักเรียน และจากผลการวิจัย พบว่า กลุ่มแฟงที่ 3 เป็นกลุ่มที่ได้รับผลกระทบจากการทดสอบทางการศึกษาระดับชาติทั้งทางบกและทางลบระดับสูงมีจำนวนครูมากที่สุดคิดเป็นร้อยละ 52 สอดคล้องกับงานวิจัยของปาริฉัตร ปิติสุทธิ (2558) เช่นกันที่พบว่าครูมีพฤติกรรมการจัดการเรียนการสอนสะท้อนผลกระทบจากการทดสอบทางบกอยู่ในระดับค่อนข้างมากเนื่องจากการจัดการเรียนการสอนตามพระราชบัญญัติการศึกษาแห่งชาติ พุทธศักราช 2542 เน้นให้จัดการเรียนการสอนที่มุ่งเน้นผู้เรียนเป็นสำคัญ และนับตั้งแต่มีการประกาศใช้พระราชบัญญัติทางการศึกษาแห่งชาติ

ดังกล่าว กว่าจะสองทศวรรษที่วงการการศึกษาของประเทศไทยขับเคลื่อนการเรียนการสอนเชิงรุก (Active learning) การจัดการศึกษาที่เน้นผู้เรียนเป็นสำคัญ ในขณะที่เดียวกันก็พบว่าครูกลุ่มเดียวกันนั้นอาจจะได้รับผลกระทบจากการทดสอบทางลบสูงเช่นเดียวกัน ซึ่งจากการตรวจสอบผลกระทบของการสอบพบว่านโยบายที่สร้างแรงกดดันต่อครูในการสอบที่มีความสำคัญสูงทำให้ครูเรียนรู้ที่จะเปลี่ยนแปลงการจัดการเรียนการสอนเพื่อให้สอดคล้องกับ พฤติกรรมการเรียนรู้และเนื้อหาที่เกี่ยวข้องกับการสอบมากยิ่งขึ้น ครูจึงมีการสอนเพื่อสอบ (Teach to the test) โดยการเพิ่มความสม่ำเสมอและความเอาใจจริงเอาใจในการฝึกนักเรียนทำข้อสอบ เกิดการบิดเบือนการปฏิบัติในการจัดการเรียนการสอน สอนไม่ครอบคลุมหลักสูตร นักเรียนไม่ได้รับประสบการณ์ในการเรียนรู้ ซึ่งล้วนแล้วแต่เป็นผลกระทบจากการทดสอบทางลบทั้งสิ้น ความขัดแย้งในเชิงแนวคิดการจัดการเรียนการสอนและการปฏิบัติจึงเกิดขึ้นในหน่วยปฏิบัติ ทั้งที่หน่วยงานต้นสังกัดก็ส่งเสริมให้ครูจัดการเรียนรู้ที่เน้นผู้เรียนเป็นสำคัญด้วย (Au, 2011; Madaus, 1988; Polesel, Rice & Dulfer, 2014)

2. กลุ่มแฝงที่ 2 ครูที่ได้รับผลกระทบจากการทดสอบทางการศึกษาแห่งชาติทางบกและทางลบนานกลาง มีความต้องการจำเป็นในการพัฒนาสมรรถนะการประเมินเพื่อการเรียนรู้สูงสุด ในขณะที่มีค่าเฉลี่ย T-score สูงกว่าทั้งสองกลุ่ม ครูกลุ่มนี้มีเพียงร้อยละ 11 จากจำนวนครู 600 คน ซึ่งหากครูมีความรู้ในหลักการประเมินครูเหล่านี้ก็จะขัดขวางการเกิดขึ้นของปรากฏการณ์ผลกระทบจากการทดสอบทางลบทั้งที่เกิดกับการจัดการเรียนการสอนของครูและผลกระทบทางลบที่จะเกิดกับผู้เรียน ครูมีบทบาทสำคัญอย่างยิ่งในการลดผลกระทบทางลบที่มีต่อการจัดการเรียนการสอน (Chiekam, 2014) ส่วนกลุ่มแฝงที่ 1 ครูที่ได้รับผลกระทบจากการทดสอบทางการศึกษาแห่งชาติทางบกและทางลบต่ำจะมีความต้องการจำเป็นรองลงมาจากกลุ่มแฝงที่ 2 ในขณะที่ครูกลุ่มนี้มีค่าเฉลี่ย T-score สมรรถนะการประเมินเพื่อการเรียนรู้จากการทดสอบในภาพรวมและบางด้านที่ต่ำกว่ากลุ่มแฝงทั้งสองอย่างมีนัยสำคัญทางสถิติ ครูกลุ่มนี้มีจำนวนมากพอสมควรคิดเป็นร้อยละ 37 ซึ่งนับเป็นเรื่องที่ดีที่ครูกลุ่มนี้ได้รับผลกระทบทางลบต่ำหากแต่ก็มีผลกระทบทางลบต่ำเช่นกัน ครูกลุ่มนี้อาจจะเป็นกลุ่มที่มีความคิดเห็นว่าการสอบไม่ได้มีผลต่อการจัดการเรียนการสอนและการเรียนรู้ของนักเรียน (Kennedy & Lui, 2013) จึงพบว่า ผลกระทบทั้งทางบวกและทางลบต่ำ

ข้อเสนอแนะ

ข้อเสนอแนะเชิงนโยบาย

หน่วยงานต้นสังกัดควรนำสารสนเทศเกี่ยวกับปรากฏการณ์ผลกระทบจากการทดสอบระดับชาติ ไปผลักดันให้มีการขับเคลื่อนกระบวนการพัฒนาครูที่มุ่งเน้นให้เกิดผลกระทบทางบวก รวมทั้งการกำหนดนโยบายการพัฒนาครูที่สอดคล้องกับความต้องการจำเป็นสำหรับครูกลุ่มต่าง ๆ ที่ต่างก็มีความต้องการที่หลากหลาย ทั้งนี้เพื่อให้เกิดผลในเชิงปฏิบัติที่มีประสิทธิภาพและยั่งยืน

ข้อเสนอแนะเชิงวิชาการ

1. ควรมีการศึกษาความสัมพันธ์เชิงโครงสร้างที่ส่งผลต่อสมรรถนะการประเมินเพื่อการเรียนรู้ของครูที่ได้รับผลจากการทดสอบทางการศึกษาแห่งชาติ
2. ควรมีการวิจัยเชิงคุณภาพค้นหาปัจจัยที่ทำให้ครูได้รับผลกระทบจากการทดสอบทั้งทางบวกและทางลบสูง หรือต่ำ
3. ควรมีการวิจัยเชิงออกแบบกระบวนการพัฒนาสมรรถนะการประเมินเพื่อการเรียนรู้ของครูที่ได้รับผลกระทบจากการทดสอบทางการศึกษาแห่งชาติ โดยคำนึงถึงการออกแบบที่อิงความต้องการจำเป็นที่แตกต่างกัน

ข้อเสนอแนะในการนำผลการวิจัยไปใช้ประโยชน์

ควรนำสารสนเทศที่ได้จากการวิจัยในประเด็นความแตกต่างของความต้องการของครูในแต่ละกลุ่มไปวางแผนการพัฒนาครูให้ตรงกับความต้องการจำเป็น กล่าวคือ

กลุ่มแฟงที่ 1 เป็นกลุ่มครูที่มีสมรรถนะทางการประเมินเพื่อการเรียนรู้ต่ำกว่ากลุ่มอื่นแต่มีความต้องการพัฒนาสมรรถนะการประเมินเพื่อการเรียนรู้รองลงมาจากกลุ่มแฟงที่ 2 ดังนั้นสำหรับครูกลุ่มนี้ควรส่งเสริมให้มีสมรรถนะทางการประเมินเพื่อพัฒนาการเรียนรู้สูงขึ้น ซึ่งจะทำให้ผลจากการทดสอบทางบวกมากขึ้นตามไปด้วย

กลุ่มแฟงที่ 2 เป็นกลุ่มครูที่มีสมรรถนะทางการประเมินเพื่อการเรียนรู้สูงกว่ากลุ่มอื่น ซึ่งแม้ว่ามีเพียงร้อยละ 11 จากจำนวน 600 คน แต่ควรตระหนักถึงความต้องการในการพัฒนาสมรรถนะทางการประเมินเพื่อพัฒนาการเรียนรู้ของครูที่มีระดับสูงสุด ดังนั้นควรหากระบวนการส่งเสริมให้ทันท่วงทีเพื่อให้ครูเหล่านี้มีพฤติกรรมการเรียนการสอนที่เป็นไปทางบวกมากยิ่งขึ้น และลดผลกระทบทางลบลงไปด้วย

กลุ่มแฟงที่ 3 เป็นกลุ่มครูที่มีสมรรถนะทางการประเมินเพื่อการเรียนรู้สูงไม่แตกต่างจากกลุ่มแฟงที่ 2 หากแต่มีความต้องการในการพัฒนาสมรรถนะทางการประเมินเพื่อพัฒนาการเรียนรู้ต่ำกว่ากลุ่มอื่น ๆ ครูกลุ่มนี้มีสูงถึงร้อยละ 52 ควรเป็นประเด็นเร่งด่วนในการพัฒนาครูให้มีพฤติกรรมการจัดการเรียนการสอนทางลบให้น้อยที่สุด

เอกสารอ้างอิง

- ญาณิกา ลุนราศรี. (2557). ผลกระทบของการทดสอบทางการศึกษาระดับชาติที่มีต่อการเรียนการสอนภาษาอังกฤษตามการรับรู้ของนักเรียนชั้นมัธยมศึกษาปีที่ 3. *OJED*, 9(4), 226-240.
- ปาริฉัตร ปิติสุทธิ. (2558). ผลการสอบระดับชาติที่มีต่อครูและนักเรียน: การวิเคราะห์ปรากฏการณ์วอชแบค. วิทยานิพนธ์ครุศาสตรมหาบัณฑิต สาขาวิชาวิธีวิทยาการวิจัยการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.
- วันดี สมมิตร. (2552). ผลกระทบของการสอบระดับชาติขั้นพื้นฐานที่มีต่อพฤติกรรมการสอนของครูในเขตกรุงเทพมหานคร: การวิจัยแบบผสม. วิทยานิพนธ์ปริญญาโทมหาบัณฑิต สาขาวิชาวิธีวิทยาการวิจัยการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.
- สร้อยญา จันทร์ชูสกุล. (2561). “แนวคิด หลักการ และยุทธวิธีในการประเมินผลเพื่อการเรียนรู้ Concepts, Principles and Strategies of Assessment for Learning”. *วารสารการวัดผลการศึกษามหาวิทยาลัยมหาสารคาม*, 24(1), 14-28.
- Ahmad, S., & Rao, C. (2012). Examination Washback Effect: Syllabus, Teaching Methodology and the Learners' Communicative Competence. *Journal of Education and Practice*, 3(15), 173-183.
- Au, W. (2011). Teaching under the new Taylorism: High stakes testing and the standardization of the 21st century curriculum. *Journal of Curriculum Studies*, 43(1), 25-45.
- Chiekem, E. (2014). Teachers perception of washback effect of high-stakes tests on post university tertiary matriculation examination in Delta State University, Abraka, Delta State. *Journal of Education and Policy Review*, 6(1), 18-25.
- Ghorbani, M. R., & Neissari, M. (2015). Washback Effect of the Iranian Concours on Senior High School Students' EFL Learning Activities. *Iranian Journal of Language Testing*, 5(1), 1-28.
- Glover, P. (2014). Do language examinations influence how teachers teach. *International Online Journal of Education and Teaching*, 1(3), 197-214.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis*(7 ed.). NJ: Pearson Prentice Hall.
- Kennedy, S., & Lui, R. (2013). Washback of a high-stakes English test in China: Student and teacher perceptions. *Concordia Working Papers in Applied Linguistics*, 4, 22-29.
- Madaus, G. F. (1988). The distortion of teaching and testing: High stakes testing and instruction. *Peabody Journal of Education*, 65(3), 29-46.

- Pan, Y. C. (2009). A review of washback and its pedagogical implications. *VNU Journal of Science, Foreign Languages*, 25, 257-263.
- Muñoz, A. P., & Álvarez, M. E. (2010). Washback of an oral assessment system in the EFL classroom. *Language testing*, 27(1), 33-49.
- Polesel, J., Rice, S., & Dulfer, N. (2014). The impact of high-stakes testing on curriculum and pedagogy: a teacher perspective from Australia. *Journal of Education Policy*, 29(5), 640-657.
- Wilson, A. T. (2009). *Assessment-based instruction applied to a course and lab in digital signal processing*. In ASEE Southeast Section Conference, USA.
- Yu, W., & Jie, B. (2013). *Washback Effect on Second Language Acquisition Test: A Case Study of College English Entrance Exam*. In The International Academic Workshop on Social Science (IAW-SC 2013), AMS.