

ข้อควรระวังในการใช้ทฤษฎีการเรียนรู้ของบลูม
ในการสร้างข้อสอบโอเน็ตวิชาภาษาอังกฤษ

Cautions in Using Bloom’s Taxonomy as a Framework in
Developing an English test for O-NET

บดินทร์ จินดา¹
Bordin Chinda

บทคัดย่อ

บทความนี้ศึกษาความเหมาะสมของการใช้ทฤษฎีการเรียนรู้ของบลูม (Bloom’s Taxonomy) เป็นกรอบในการออกข้อสอบโอเน็ต (O-NET) โดยการวิเคราะห์และสังเคราะห์ บทความ งานวิจัย เอกสาร ตำรา ที่เกี่ยวข้องกับข้อสอบ O-NET และทฤษฎีการวัดและประเมินผลทางภาษาอังกฤษ จากการศึกษาค้นพบว่า Bloom’s Taxonomy ไม่เหมาะต่อการนำมาใช้เป็นกรอบในการออกข้อสอบ O-NET วิชาภาษาอังกฤษเนื่องจากไม่ตรงกับทฤษฎีการวัดและประเมินผลทางภาษา

คำสำคัญ: ทฤษฎีการเรียนรู้ของบลูม, ข้อสอบภาษาอังกฤษ, โอเน็ต

Abstract

This article investigates the use of Bloom’s Taxonomy in developing an English test for the O-NET (Ordinary National Educational Test). The results were drawn from published articles, research papers, books, and related materials in language testing and assessment. It was found that it is not appropriate to use Bloom’s Taxonomy in creating an English test for the O-NET because it does comply with the language testing theories.

Key words: Bloom’s Taxonomy, English test, O-NET

บทนำ

ในการที่ประเทศไทยจะเข้าสู่ประชาคมเศรษฐกิจอาเซียน (AEC) ปี 2558 นั้น ประเทศไทยจะต้องแข่งขันในทุกๆด้านกับ 10 ประเทศสมาชิกอาเซียน สิ่งหนึ่งที่หลายฝ่ายให้ความสนใจเนื่องจากเป็นจุดอ่อนของประเทศไทยคือด้านความสามารถในการใช้ภาษาอังกฤษของพลเมืองไทยที่ดีน้อยกว่าหลายประเทศสมาชิก ซึ่งประชาคมอาเซียนได้มีมติให้ภาษาอังกฤษเป็นภาษาทำงานของอาเซียน (Association of Southeast Asian Nations, 2008, p. 29) หน่วยงานทั้งภาครัฐและเอกชนหลายหน่วยงาน เช่น กระทรวงศึกษาธิการได้มีความพยายามที่จะปรับขีดความสามารถในการใช้ภาษาอังกฤษของนักเรียนไทย นโยบายหนึ่งที่กระทรวงศึกษาธิการได้ทำคือการกำหนดให้นักเรียนทุกคน

วัดระดับภาษาอังกฤษขั้นพื้นฐาน หรือ ข้อสอบโอเน็ต (O-NET) ที่จัดทำโดยสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) หรือ สทศ. ข้อสอบ O-NET จึงกลายเป็นส่วนสำคัญมากต่อครูและนักเรียน อย่างไรก็ตามนักวิชาการหลายท่านได้วิพากษ์วิจารณ์ข้อสอบ O-NET ไว้หลายประเด็นเช่น จิรดา วุฑฒยากร (2555) เห็นว่า ข้อสอบ O-NET วิชาภาษาอังกฤษทุกระดับชั้น เป็นข้อสอบระดับชาติที่มีความสำคัญต่อการพัฒนาการเรียนการสอน แต่ตัวข้อสอบมีความลำเอียง (bias) และใช้วิธีการรายงานผลคะแนนที่ไม่เป็นไปตามมาตรฐานสากล บทความนี้จะวิพากษ์วิจารณ์ข้อสอบ O-NET วิชาภาษาอังกฤษ ชั้น ม.6 เท่านั้น เนื่องจากผู้เขียนเป็นอาจารย์สอนวิชาภาษาอังกฤษระดับอุดมศึกษา

¹ ดร., สาขาวิชาภาษาอังกฤษ ภาควิชาภาษาตะวันตก คณะมนุษยศาสตร์ มหาวิทยาลัยเชียงใหม่

ปัญหา

จากการที่ผู้เขียนได้รับเชิญให้เป็นผู้ดำเนินการสัมมนาวิชาการเชิงปฏิบัติการให้กับครูสอนภาษาอังกฤษในเขตภาคเหนือหลายครั้งภายใต้หัวข้อเรื่องการออกข้อสอบ O-NET ในการเตรียมสัมมนาวิชาการเชิงปฏิบัติการพบว่าการจัดสัมมนาเชิงปฏิบัติการทุกครั้งผู้จัดจะให้ผู้เขียนให้ความรู้และฝึกหัดครูที่เข้าร่วมสัมมนาในการแต่งข้อสอบวัดผลสัมฤทธิ์วิชาภาษาอังกฤษโดยใช้ ทฤษฎีการเรียนรู้ของ Bloom หรือ Bloom's Taxonomy เป็นกรอบในการสร้างข้อสอบโดยอิงมาตรฐานการเรียนรู้ตามหลักสูตรการศึกษาขั้นพื้นฐาน พ.ศ. 2544 (ปรับปรุง พ.ศ. 2551) เมื่อได้ไปจัดสัมมนาวิชาการเชิงปฏิบัติการดังกล่าวหลายครั้งก็ค้นพบว่าไม่สามารถช่วยครูสอนภาษาอังกฤษเหล่านั้นได้ดีเท่าที่ควร ทั้งที่พยายามใช้ประสบการณ์จากการรับราชการเป็นอาจารย์สอนภาษาอังกฤษในมหาวิทยาลัยเชียงใหม่มากกว่าสิบปี การเรียนปริญญาเอกด้านการวัดและประเมินผลภาษาอังกฤษ การเข้าร่วมสัมมนาระดับนานาชาติด้านการวัดและประเมินผลภาษาอังกฤษและการเรียนการสอนภาษาอังกฤษ รวมทั้งการเป็นกรรมการออกข้อสอบให้กับหลายสถาบัน พบว่าการออกข้อสอบภาษาอังกฤษโดยใช้ Bloom's Taxonomy เป็นกรอบนั้นไม่ได้เป็นไปตามหลักทฤษฎีการสร้างแบบวัดและประเมินผลทางภาษาอังกฤษที่เคยเรียน ผู้เขียนจึงได้พยายามไปศึกษาเพิ่มเติมเรื่องของ Bloom's Taxonomy และการสร้างข้อสอบภาษาอังกฤษเพื่อหาคำตอบว่า การใช้ Bloom's Taxonomy ในการออกข้อสอบวิชาภาษาอังกฤษในข้อสอบ O-NET นั้นเหมาะสมหรือไม่ โดยผู้เขียนวิเคราะห์และสังเคราะห์ข้อมูลเกี่ยวกับการออกข้อสอบภาษาอังกฤษเกี่ยวกับข้อสอบ O-NET ที่สามารถหาได้จากแหล่งข้อมูลต่าง ๆ เช่นจากข้อมูลที่เผยแพร่โดย สทศ. สอบถามจากผู้ที่เกี่ยวข้องกับการออกข้อสอบ O-NET และศึกษาจากเอกสารเกี่ยวกับการวัดและประเมินผลทางภาษาที่ได้รับการตีพิมพ์และเป็นที่ยอมรับในระดับนานาชาติและพบว่าการใช้ Bloom's Taxonomy มาเป็นกรอบในการสร้างข้อสอบวิชาภาษาอังกฤษนั้นไม่ตรงตามทฤษฎีการวัดและประเมินผลทางภาษาอังกฤษ

แนวทางการพัฒนาข้อสอบโอเน็ต

ข้อสอบ O-NET (Ordinary National Educational Test) คือ การทดสอบการศึกษาระดับชาตินี้ขั้นพื้นฐาน เป็นชุดข้อสอบที่ได้รับการพัฒนาโดยสถาบัน

ทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) หรือ สทศ. โดยมีทั้งหมด 8 สารวิชา สำหรับนักเรียนระดับชั้น ป.6 ม.3 และ ม.6 จากข้อมูลที่เผยแพร่โดย สทศ. (สถาบันทดสอบทางการศึกษาแห่งชาติ) เรื่องระบบการพัฒนาข้อสอบ สามารถสรุปกระบวนการออกข้อสอบ O-NET ได้ว่า การพัฒนาข้อสอบเริ่มจากการทำผังการออกข้อสอบ (test blueprint หรือ specifications) โดยอิงจากมาตรฐานการเรียนรู้ตามหลักสูตรการศึกษาขั้นพื้นฐาน พ.ศ. 2544 (ปรับปรุง พ.ศ. 2551) ขั้นตอนต่อไปคือการสังเคราะห์เนื้อหาาร่วมของแบบเรียนต่างๆที่กระทรวงศึกษาธิการให้ความเห็นชอบ หลังจากนั้นจะมีการวางระดับความยาก ง่าย และกำหนดรูปแบบของข้อสอบ รวมทั้งมีการทำรายละเอียดของผังการออกข้อสอบ 2 ผัง ต่อจากนั้นก็จะเป็นขั้นตอนการออกข้อสอบ โดยมีครูโรงเรียน 5-10 คน เป็นผู้ออกข้อสอบ O-NET ระดับ ป.6 และ ม.3 ส่วนอาจารย์มหาวิทยาลัยเป็นผู้ออกข้อสอบระดับ ม.6 เมื่อมีการออกข้อสอบแล้วก็จะมีการจัดทำข้อสอบคู่ขนานและผังการออกข้อสอบตามระดับความยากง่าย (หรือความซับซ้อน) แล้วถึงจะนำไปจัดพิมพ์ ส่วนการกลั่นกรองข้อสอบนั้น สทศ. มีขบวนการกลั่นกรอง 2 ระยะ คือ การตรวจสอบความถูกต้องทางภาษา และขั้นที่สองคือการตรวจสอบความถูกต้องของเนื้อหา นอกจากนี้ สทศ. มีการพัฒนาผู้ออกข้อสอบโดยการเผยแพร่ข้อมูลทางสถิติและมีการสัมมนากับผู้ออกข้อสอบเพื่อให้ผู้ออกข้อสอบนำไปปรับปรุงในการออกข้อสอบในครั้งต่อไป นอกจากนี้ ศาสตราจารย์ ดร. อุทุมพร จามรมาน (อดีตผู้อำนวยการ สทศ.) ได้นำเสนอข้อมูลเรื่องการออกข้อสอบ O-NET ว่า ข้อสอบของ สทศ. เน้นการคิดวิเคราะห์ และยังระบุรายละเอียดผังข้อสอบของกลุ่มสาระการเรียนรู้ภาษาต่างประเทศ (ภาษาอังกฤษ) ว่า ข้อสอบมี 70 ข้อ ประกอบไปด้วยสาระภาษาเพื่อการสื่อสาร โดยมี 3 มาตรฐานการเรียนรู้ จำนวนข้อสอบ 50 ข้อ สาระภาษาและวัฒนธรรม 2 มาตรฐานการเรียนรู้ จำนวน 15 ข้อ สาระภาษากับความสัมพันธ์กับกลุ่มสาระอื่น 1 มาตรฐานการเรียนรู้ จำนวน 3 ข้อ และ สาระภาษากับความสัมพันธ์กับชุมชนและโลก 2 มาตรฐานการเรียนรู้ จำนวน 2 ข้อ (อุทุมพร, 2553)

ข้อสอบโอเน็ตวิชาภาษาอังกฤษและทฤษฎีการเรียนรู้ของบลูม

จากการสังเกตจากข้อสอบ O-NET นั้นพบว่า สทศ.ได้ใช้ Bloom's Taxonomy เป็นกรอบในการสร้างข้อสอบโดยอิงมาตรฐานการเรียนรู้ตามหลักสูตรการศึกษาขั้นพื้นฐาน พ.ศ. 2544 (ปรับปรุง พ.ศ. 2551) Krathwohl (2002) กล่าวว่า Bloom's Taxonomy of Educational Objectives in the Cognitive Domain มีการเผยแพร่ครั้งแรกใน Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook I: Cognitive Domain โดย Benjamin S. Bloom และคณะในปี 1956 โดย Bloom ได้จำแนกจุดมุ่งหมายการเรียนรู้ออกเป็น 3 ด้าน คือ พุทธิพิสัย (Cognitive Domain) จิตพิสัย (Affective Domain) และ ทักษะพิสัย (Psychomotor Domain) ด้านที่ใช้เป็นกรอบในการสร้างแบบทดสอบผลสัมฤทธิ์การเรียนรู้หรือข้อสอบคือด้านพุทธิพิสัยหรือการเรียนรู้ โดย Bloom ได้แบ่งการเรียนรู้ออกเป็น 6 ระดับ ได้แก่ ความรู้ที่เกิดจากความจำ (knowledge) ความเข้าใจ (Comprehension) การประยุกต์ (Application) การวิเคราะห์ (Analysis) การสังเคราะห์ (Synthesis) การประเมินค่า (Evaluation) และในแต่ละระดับ ยกเว้น ประยุกต์ ก็มีการแบ่งออกเป็นระดับย่อย Krathwohl (2002) กล่าวว่า ได้มีการนำ Bloom's Taxonomy นี้ไปใช้อย่างแพร่หลายในการจัดการเรียนการสอนโดยเฉพาะการนำไปเป็นกรอบในการจัดกลุ่มวัตถุประสงค์การเรียนรู้และการออกข้อสอบ

แต่อย่างไรก็ตามในวงการการวัดผลและประเมินผลทางภาษาอังกฤษในฐานะที่เป็นภาษาที่สองหรือภาษาต่างชาติซึ่งเป็นสาขาหนึ่งของวงการการศึกษาที่มีความสำคัญและมีการพัฒนามาตลอด ไม่ได้มีการนำเอา Bloom's Taxonomy มาใช้เพื่อเป็นกรอบในการออกข้อสอบภาษาอังกฤษ จากการสำรวจบทความที่ตีพิมพ์ใน Encyclopedia of Language and Education (Volume 7 Language Testing and Assessment) (Clapham & Corson, 1997) หนังสือที่สรุปงานวิจัยด้านการวัดและประเมินผลทางภาษาอังกฤษ 2 เล่มได้แก่ Examining writing (Shaw & Weir, 2007) และ Examining reading (Khalifa & Weir, 2009) และตำราที่เกี่ยวข้องการสร้างข้อสอบทางภาษาอังกฤษที่มีการนำไปใช้เป็นเอกสารหลักในการสร้างข้อสอบภาษาอังกฤษที่เป็นที่

ยอมรับในระดับนานาชาติ เช่น Fundamental considerations in language testing (Bachman, 1990) Language test construction and evaluation (1995), Language testing in practice (Bachman & Palmer, 1996) พบว่าบทความและหนังสือต่างๆที่เกี่ยวข้องการวัดและประเมินผลภาษาอังกฤษโดยนักวิชาการชั้นนำในศาสตร์นี้ ไม่ได้กล่าวถึงการออกข้อสอบโดยใช้ Bloom's Taxonomy เป็นกรอบในการสร้างข้อสอบเลย โดยหนังสือเหล่านี้ได้ให้กรอบการสร้างข้อสอบภาษาอังกฤษที่ได้มาจากการศึกษาและวิจัยด้านการวัดและประเมินผล และการเรียนการสอนภาษาอังกฤษเป็นภาษาที่สองหรือภาษาต่างประเทศ เช่น การใช้แบบจำลองของความสามารถใช้ภาษาเชิงสื่อสาร (Model of Communicative Language Ability) ของ Bachman (1990) ที่มีการระบุถึงความรู้ความสามารถด้านต่างๆของการเรียนภาษา เช่น ความรู้ด้านไวยากรณ์ และด้านภาษาศาสตร์เชิงสังคม เป็นต้น

จากหนังสือชุด Cambridge Language Assessment Series ที่มีการเผยแพร่ 6 เล่มที่ให้แนวทางและกรอบการออกข้อสอบภาษาอังกฤษ ด้านไวยากรณ์ (Purpura, 2004) คำศัพท์ (Read, 2000) การพูด (Luoma, 2004) การเขียน (Weigle, 2002) การฟัง (Buck, 2001) และการอ่าน (Alderson, 2000) มีเพียง 1 เล่มที่มีการกล่าวถึง Bloom's Taxonomy ในการวัดและประเมินผลทางภาษาอังกฤษคือหนังสือ Assessing Reading โดย J. Charles Alderson (2000, pp. 10-11) ซึ่งตรงกับข้อมูลที่ Krathwohl (2002) ได้ให้ข้อมูลเกี่ยวกับ Bloom's Taxonomy ว่า Bloom's Taxonomy เป็นทฤษฎีที่มีอิทธิพลเป็นอย่างมากในการสร้างหลักสูตร สื่อการเรียนการสอน และข้อสอบ ส่วนในการวัดและประเมินผลทางภาษาอังกฤษนั้น Alderson ได้กล่าวว่านอกจาก Bloom's Taxonomy แล้ว ด้านทฤษฎีการเรียนรู้ภาษาอังกฤษเป็นภาษาที่สองนั้น Munby ได้แบ่งการเรียนรู้ภาษาออกเป็นทักษะย่อยต่างๆ เช่น ความเข้าใจข้อมูลจากการอ่านที่ไม่ได้กล่าวโดยตรง การตีความเรื่องที่อ่าน การอ่านแบบสำรวจ ความเข้าใจความหมายของกรอบความคิดจากเรื่องที่อ่าน ซึ่ง Munby' Taxonomy นี้ก็มีอิทธิพลต่อการสร้างหลักสูตรและสื่อการสอนรวมทั้งข้อสอบภาษาอังกฤษ (Alderson, 2000, p. 10)

แต่อย่างไรก็ตาม Alderson (2000) ได้ตั้งว่า ถึงแม้ว่า Bloom's Taxonomy หรือ Munby's Taxonomy เป็นทฤษฎีการเรียนรู้ที่มีอิทธิพลมากในการนำมาใช้เป็นกรอบในการสร้างข้อสอบแต่การใช้ต้องทำอย่างระมัดระวังการใช้ทฤษฎีการเรียนรู้เหล่านี้ ด้วยเหตุผลหลายประการด้วยกันคือ

1. ต้นกำเนิดของ ทฤษฎีการเรียนรู้เหล่านี้ ส่วนมากมาจากความคิดของนักทฤษฎีที่ไม่ได้มาจากผลการวิจัย

2. ทฤษฎีการเรียนรู้ส่วนมากมีการให้คำจำกัดความที่คลาดเคลื่อน และมีความเหลื่อมล้ำและซ้ำซ้อน

3. ผู้เชี่ยวชาญส่วนใหญ่ไม่สามารถนำทฤษฎีการเรียนรู้เหล่านี้มาใช้ระบุว่าข้อสอบแต่ละข้อวัดทักษะอะไร

ดังนั้นจึงสามารถสรุปได้ว่า การสร้างข้อสอบ O-NET วิชาภาษาอังกฤษไม่ควรใช้ Bloom's Taxonomy มาเป็นกรอบแต่ควรใช้ทฤษฎีการออกข้อสอบที่ได้รับการพัฒนามาสำหรับวิชาภาษาอังกฤษโดยเฉพาะ ซึ่งผู้ที่เกี่ยวข้องสามารถศึกษาได้จากเอกสารต่างๆที่กล่าวมาแล้วข้างต้นและจากหนังสือ ตำรา และงานวิจัยที่เผยแพร่โดยหน่วยงานที่ออกข้อสอบภาษาอังกฤษระดับนานาชาติเช่น ETS ที่สร้างข้อสอบ TOEFL และ Cambridge ESOL ที่สร้างข้อสอบ IELTS

นอกจากนี้ ในการสร้างข้อสอบ O-NET วิชาภาษาอังกฤษ สทศ.ควรให้ความสำคัญกับคุณภาพของข้อสอบให้ครบทุกด้าน เช่น คุณภาพของข้อสอบภาษาอังกฤษด้านต่างๆที่นำเสนอโดย Bachman and Palmer (1996) เนื่องจากผลจากการศึกษาคุณภาพของข้อสอบวิชาภาษาอังกฤษยังมีปัญหา ตัวอย่างเช่นจากรายงานของสิรินธร สิ้นจินดาวงศ์และคณะ (2552) จะเห็นได้ว่า ระดับความยาก ง่ายของข้อสอบวิชาภาษาอังกฤษ ของปี 2552 แตกต่างจากแผนการออกข้อสอบของ สทศ. อย่างมาก กล่าวคือ สทศ. ได้วางแผนว่าข้อสอบวิชาภาษาอังกฤษจะมีข้อสอบที่มีระดับความยาก ระดับ ยาก 20% แต่ผลการศึกษาพบว่าข้อสอบภาษาอังกฤษมีข้อสอบที่มีระดับความยาก ระดับยาก มากถึง 51% ตามแผนที่จะมีระดับความยากระดับ ปานกลาง 60% แต่ข้อสอบมี 47% และ แผนที่จะมีข้อสอบที่มีระดับความยากระดับ ง่าย 20% ข้อสอบมีจำนวนข้อสอบระดับความยากระดับง่าย เพียง 2% เท่านั้น สาเหตุที่เกิด

ความแตกต่างกันอย่างมาอาจมาจากการที่ สทศ. ไม่ได้มีระบบการทดลองข้อสอบ (trial) ก่อนที่จะมีการนำข้อสอบไปใช้จริง ข้อสอบ O-NET เป็นข้อสอบระดับชาติที่ใช้วัดผลสัมฤทธิ์การเรียนรู้ของนักเรียนทั้งประเทศควรมีวิธีการพัฒนาข้อสอบตามหลักสากลที่มีการทดลองข้อสอบก่อนที่จะมีการนำข้อสอบไปใช้จริงและ วัดความตรง (validity) ของข้อสอบก่อน นอกจากเรื่องระดับความยากง่ายของข้อสอบแล้ว อีกประเด็นหนึ่งซึ่งมีความสำคัญมากในเรื่องการออกข้อสอบคือคุณภาพด้านความเที่ยงตรงของข้อสอบ จากเอกสารของ สทศ. เรื่องระบบการพัฒนาข้อสอบสทศ. (สถาบันทดสอบทางการศึกษาแห่งชาติ) จะเห็นได้ว่าข้อสอบวิชาภาษาอังกฤษไม่ได้มีการดูแลเรื่องคุณภาพของข้อสอบโดยผู้เชี่ยวชาญด้านการวัดและประเมินผลทางภาษาในขั้นตอนที่สำคัญๆเลย ถึงแม้ว่า สทศ.ได้ใช้เครื่องมือทางสถิติมาวัดความเที่ยง (reliability) ของข้อสอบ แต่ สทศ. ไม่ได้มีการหาค่าความตรง (validity) ของข้อสอบเลย โดยเฉพาะข้อสอบทางภาษา ซึ่งถือว่า ความตรง เป็นหัวใจหลักของคุณภาพของข้อสอบ

สรุป

ในวงการการศึกษาโดยเฉพาะอย่างยิ่งการวัดและประเมินทางภาษา ได้มีการศึกษาวิจัยมาหลายทศวรรษและพบว่าข้อสอบไม่ว่าจะระดับไหนก็ตามต่างมีผลกระทบต่อชีวิตของผู้เข้าสอบและครูไม่ว่าจะเป็นผลดีหรือผลเสีย เช่นผลจากการศึกษาของ Wall (2005) ที่ศึกษาแบบสอบถามการศึกษาระดับชาติขั้นพื้นฐานของประเทศศรีลังกา และการศึกษาของ Cheng (2005) ที่ศึกษาแบบสอบถามการศึกษาระดับชาติของฮ่องกง ตัวอย่างของผลเสียที่อาจเกิดขึ้นจากข้อสอบระดับชาติเช่น ครูเน้นเรื่องของการเตรียมตัวนักเรียนให้ทำข้อสอบแทนที่จะสอนเน้นเนื้อหาหรือทักษะที่นักเรียนควรจะได้ตามที่หลักสูตรกำหนด คุณภาพของข้อสอบเป็นสิ่งหนึ่งที่จะสามารถช่วยให้ผลเสียที่อาจเกิดขึ้นนั้นลดน้อยลง ข้อสอบที่ดีของคุณภาพย่อมจะก่อให้เกิดผลเสียต่ออนาคตของผู้เข้าสอบในที่นี้คือนักเรียน แล้วใครจะเป็นคนรับผิดชอบต่ออนาคตของชาติทั้งประเทศ ถ้าไม่ใช้หน่วยงานที่รับผิดชอบในการสร้างข้อสอบโดยตรง

เอกสารอ้างอิง

- จิรดา วุฒยากร. (8 พฤษภาคม 2555). เดิมพันชีวิตผู้เรียนด้วยผลจากการวัดและประเมิน. *มติชน*. หน้า 7.
- สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) (ม.ป.ด.). *ระบบการพัฒนาข้อสอบของ สทศ. (online)*, 6 สิงหาคม 2555. เอกสารของสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน).
<http://www.niets.or.th>
- สิรินธร สิ้นจินดาวงศ์ และคณะ. (2552). *การวิเคราะห์คุณภาพของแบบทดสอบการศึกษาระดับชาติ ขั้นพื้นฐาน (O-NET)*. (online), 6 สิงหาคม 2555. โครงการวิจัยได้รับทุนจาก สทศ. <http://www.niets.or.th>
- อุทุมพร จามรمان. (2553). *รูปแบบข้อสอบของ สทศ. บรรยายในงาน 5 ปี สทศ. 11* กันยายน 2553. (online), 6 สิงหาคม 2555. เอกสารของสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน).
<http://www.niets.or.th>
- Alderson, J. C. (2000). *Assessing reading*. Cambridge: Cambridge University Press.
- Alderson, J. C., Clapham, C. and Wall, D. (1995). *Language test construction and evaluation*. Cambridge: Cambridge University Press.
- Association of Southeast Asian Nations. (2008). *The ASEAN Charter*. Jakarta: ASEAN Secretariat.
- Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Bachman, L. F. & Palmer, A. S. (1996). *Language testing in practice*. Oxford: Oxford University Press.
- Buck, G. (2001). *Assessing listening*. Cambridge: Cambridge University Press.
- Clapham, C. and Corson, D. (Eds.). (1997). *Encyclopedia of language and education*. Dordrecht, the Netherlands: Kluwer Academic Publishers.
- Cheng, L. (2005). *Changing language teaching through language testing: A washback study*. Cambridge: Cambridge University Press.
- Khalifa, H. and Weir, C. J. (2009). *Examining reading: research and practice in assessing second language reading*. Cambridge: Cambridge University Press.
- Krathwohl, D. R. (2002). *A Revision of Bloom's Taxonomy: An Overview*. *Theory into Practice*. 41(4), pp. 212-218.
- Luoma, S. (2004). *Assessing speaking*. Cambridge: Cambridge University Press.
- Purpura, J. (2004). *Assessing grammar*. Cambridge: Cambridge University Press.
- Read, J. (2000). *Assessing vocabulary*. Cambridge: Cambridge University Press.
- Shaw, S. D. and Weir, C. J. (2007). *Examining writing: research and practice in assessing second language reading*. Cambridge: Cambridge University Press.
- Wall, D. (2005). *The impact of high-stakes examinations on classroom teaching: a case study using insights from testing and innovation theory*. Cambridge: Cambridge University Press.
- Weigle, S. C. (2002). *Assessing writing*. Cambridge: Cambridge University Press.