

การพัฒนาเครื่องมือการประเมินนักเรียน

ดร.ชาญชัย อาจिनสมาจาร*

ปัญหา

ครูสันติเพิ่งสอนหน่วยการสอนเรื่องการปฏิวัติในฝรั่งเศสเสร็จ และต้องการทดสอบนักเรียนของเขาว่าได้เรียนรู้ไปมากน้อยแค่ไหน มีแบบทดสอบหนึ่งในคู่มือสำหรับตำราเรียนวิชาสังคมศึกษาที่เขาใช้อยู่ อย่างไรก็ตาม ขณะที่เขาทบทวนแบบทดสอบ ครูสันติสังเกตเห็นว่า แบบทดสอบทุกข้อเป็นแบบเลือกคำตอบและด้วยเหตุผลบางอย่างทำให้เขาข้องใจ เขาไม่สามารถตัดสินใจได้ว่าจะใช้แบบทดสอบดังกล่าวหรือไม่หรือเขียนแบบทดสอบใหม่ เขาจะตัดสินใจอย่างไรว่าข้อสอบในแบบทดสอบในคู่มือครูมีความเหมาะสม ถ้าไม่เหมาะสมเขาควรเขียนข้อสอบแบบไหน

ข้อมูลพื้นเพเดิม

คำว่า การทดสอบและการประเมินผลบรรจุด้วยคำที่มีความหมายแฝงอยู่ทางด้านอารมณ์ (emotional connotation) มีคนไม่กี่คนที่สนุกกับการถูกทดสอบหรือถูกประเมิน ดูเหมือนว่าไม่มีใครชอบรับผลการทดสอบเลย อย่างไรก็ตาม มีการบ่งชี้ว่าการทดสอบจะมีบทบาทที่สำคัญในการศึกษาต่อไป

ขณะที่มีการระบุสมรรถนะหรือทักษะที่นักเรียนระดับต่าง ๆ ต้องมี การเน้นในแนวทางใหม่ ๆ ของการทดสอบได้มีเพิ่มมากขึ้น แนวทางดังกล่าวเกี่ยวข้องกับการพัฒนาแบบทดสอบที่ถูกออกแบบมาเพื่อตัดสินว่า นักเรียนได้รับทักษะต่าง ๆ ที่ระบุไว้ในข้อความว่าด้วยสมรรถนะหรือไม่ ถึงแม้ว่าการพัฒนาการทดสอบในลักษณะดังกล่าว ดูเหมือนจะไม่ชอบด้วยเหตุผล แต่มันไม่ใช่แนวทางเท่าที่การทดสอบได้รับการพัฒนา แนวทางเท่าที่เกี่ยวข้องกับการสร้างข้อสอบที่ยู่งยากมากที่สามารถใช้จัดลำดับนักเรียนในรูปของความรู้ของเขาและขอบข่ายของเนื้อหา ข้อสอบไม่เคยมุ่งหมายที่จะให้เป็นตัวแทนของความรู้และทักษะทั้งหมดในขอบข่าย

เนื่องจากแนวทางใหม่ในการทดสอบโฟกัสไปยังนักเรียนว่าได้รับทักษะจำเพาะแล้วยัง มันเป็นสิ่งสำคัญมากที่ข้อสอบที่ได้รับการพัฒนาสำหรับการทดสอบดังกล่าวประเมินความสามารถในการแสดงออกในทักษะจำเพาะจริงๆ พุดอีกนัยหนึ่ง ข้อสอบต้องวัดความสามารถของนักเรียนในการแสดงออกซึ่งพฤติกรรมเดียวกันดังที่ได้รับบออย่างจำเพาะในจุดประสงค์ใดจุดประสงค์หนึ่ง สิ่งดังกล่าวมักเรียกว่า การทดสอบแบบอิงจุดประสงค์หรืออิงเกณฑ์ การทดสอบแบบอิงจุดประสงค์ยังเรียกอีกอย่างหนึ่งว่าการประเมินตามสภาพจริง (authentic assessment) ถ้าต้องการให้นักเรียนแสดงออกในทักษะในลักษณะที่จะใช้ทักษะนอกห้องเรียน ปกติการทดสอบในชั้นเรียนเป็นเพียงแค่ตัวแทนในสิ่งที่เราต้องการให้นักเรียนสามารถทำได้จริงๆ ยกตัวอย่างเช่น ถ้านักเรียนต้องเรียนรู้ทักษะ

* คณะบริหารธุรกิจ วิทยาลัยเทคโนโลยีภาคใต้ อ.ทุ่งสง จ.นครศรีธรรมราช

กล่าวภายใต้เงื่อนไขเหล่านั้น และคุณควรแน่ใจว่าคุณประเมินนักเรียนภายใต้เงื่อนไขเหล่านั้น

เครื่องมือประเมินสำหรับผลการเรียนรู้ชนิดต่าง ๆ

เรามาทบทวนมิติของจุดประสงค์และลักษณะของข้อสอบที่มักใช้บ่อย ๆ ในการประเมินจุดประสงค์ในมิติเหล่านั้น

เครื่องมือการประเมินสำหรับจุดประสงค์ทางความรู้ (knowledge objective)

จุดประสงค์ทางความรู้กำหนดให้นักเรียนจำหรือระลึกข้อมูลที่เขาเคยพบ ในที่นี้เราถามนักเรียนให้ระบุ (state) อธิบาย (describe) ทำบัญชีรายการข้อมูลหรือเลือกจากคำตอบที่เป็นไปได้ เช่นในกรณีข้อสอบแบบปรนัยเลือกคำตอบ ดังนั้น ข้อสอบความรู้กำหนดให้นักเรียนทำบัญชีรายการพรรคการเมืองในประเทศไทย หรือเลือกอาหารที่มีคาร์โบไฮเดรตสูงที่สุดจากบัญชีรายการ ในกรณีแรก เรากำหนดให้นักเรียนเขียนคำตอบให้สมบูรณ์ และในกรณีที่สอง เราเพียงแค่อ่านนักเรียนให้จำข้อมูล ทั้งสองกรณี เราสันนิษฐานว่าได้มีการสอนข้อมูลที่จำเพาะแล้ว และเป็นเรื่องการระลึกหรือจำ

เครื่องมือการประเมินจุดประสงค์ทางทักษะสติปัญญา (intellectual skills objective)

ในที่นี้ก็เช่นเดียวกัน เราใช้ข้อสอบแบบปรนัยเลือกคำตอบหรือการเติมคำในข้อสอบให้สมบูรณ์ โดยปกติเมื่อจุดประสงค์เกี่ยวข้องกับการเรียนรู้โนทัศน์ เราสามารถใช้ข้อสอบแบบปรนัยเลือกคำตอบที่นักเรียนถูกกำหนดให้เลือกตัวอย่างของมโนทัศน์ ยกตัวอย่างเช่น ถ้าเรามีการระบุจุดประสงค์ว่านักเรียนจะสามารถจำสัตว์เหล่านี้ที่เป็นสมาชิกของกลุ่มแมว เราสามารถให้คำถามที่นำเสนอชื่อหรือภาพของสัตว์ต่างๆ และเราบอกให้นักเรียนเลือก หรือระบุสัตว์ที่อยู่ในสกุลแมว

ขณะที่เราเลื่อนไปยังระดับต่อไปของทักษะสติปัญญา นั่นก็คือการใช้กฎ เรามีแนวโน้มที่จะให้นักเรียนตอบให้สมบูรณ์ แทนที่จะให้เลือกคำตอบแบบปรนัยเลือกคำตอบ แม้ว่าจะไม่เป็นอย่างนั้นเสมอไป ยกตัวอย่างเช่น สมมุติว่าเราต้องการให้นักเรียนของเราใช้กฎเพื่อตัดสินใจว่าเส้นทางเพื่อไปให้ถึงที่ใดที่หนึ่ง เมื่อเขาได้รับการกำหนดตำแหน่งปัจจุบันของเรือและทิศทางของลม ตัวเลือกตัวแรกของเราจะเป็นการบอกนักเรียนแต่ละคนในการแล่นเรือจริงๆ ภายใต้เงื่อนไขต่างๆ การกระทำดังกล่าวจะเป็นการรวมกันของทักษะการใช้กลัมนี้อีกกับทักษะทางสติปัญญา ถ้าสิ่งดังกล่าวเป็นไปได้ ดังนั้น เราต้องพิจารณาการระบุทักษะของเขาโดยตรงให้น้อยลง เราอาจประเมินความสามารถของนักเรียนในการปฏิบัติงานดังกล่าวโดยการทำบัญชีรายการตำแหน่งการแล่นเรือและบอกนักเรียนให้เลือกตำแหน่งที่ถูกต้อง หรืออธิบายตำแหน่งแล่นเรือที่ถูกต้อง

ถ้าเราดำเนินขั้นตอนต่อไปของทักษะทางสติปัญญา ซึ่งได้แก่การแก้ปัญหา ดังนั้น เราจะมีแนวโน้มในการให้นักเรียนตอบคำถามโดยการเติมคำให้สมบูรณ์ ยกตัวอย่างเช่น เราอาจต้องการให้นักเรียนสามารถอธิบายการแล่นเรือใบรอบๆ เกาะ จากด้านหนึ่งไปยังอีกด้านหนึ่ง ในความเป็นไปได้ทั้งหมด เราจะเขียนข้อสอบที่มันจะให้คำอธิบายในสถานการณ์กับเรือใบ และเขาจะถูกบอกให้อธิบายวิธีที่เขาแล่นเรือ

แล้วงานแก้ปัญหาเช่นการสร้างโครงงานวิทยาศาสตร์หรือเขียนเรียงความละ การใช้แบบตรวจสอบรายการหรือมาตราส่วนประมาณค่าในการประเมินผลผลิตเหล่านี้มีความเหมาะสม

เครื่องมือการประเมินสำหรับจุดประสงค์ของทักษะทางการใช้กลัมนี้อีก (motor skills objective)

ในหลายๆ สถานการณ์ การประเมินทักษะการใช้กลัมนี้อีกมักใช้แนวทางที่ไม่ค่อยเป็นทางการ ผู้สอนเพียงแค่สังเกตการแสดงออกในทักษะของนักเรียนและตัดสินว่า เขาประสบความสำเร็จหรือไม่ ไม่ว่าผู้สอนจะตระหนักถึงมันหรือไม่ โดยความจริงแล้ว เขาใช้แบบตรวจสอบรายการทางสมองเพื่อตัดสินว่านักเรียนประสบความสำเร็จ

หรือไม่ แนวทางที่เป็นปรนัยมากขึ้นจะเป็นเกณฑ์ในรูปของแบบตรวจสอบรายการหรือมาตราส่วนประมาณค่า สำหรับตัดสินความเหมาะสมของการแสดงออกของนักเรียนแต่ละคน แบบตรวจสอบรายการหรือมาตราส่วนประมาณค่าอาจเป็นแค่บัญชีรายการของคุณลักษณะที่คุณพิจารณาเห็นว่าเป็นการแสดงออกที่ประสบความสำเร็จ ยกตัวอย่างเช่น แบบตรวจสอบรายการสำหรับประเด็นความเหมาะสมของการเลิฟลูกเทนนิสอาจรวมถึงข้อความ เช่น “ลูกบอลกระเด็นขึ้นข้างบนอย่างถูกต้องหรือไม่” “ผู้เรียนยืนแขนของเขาออกไปอย่างเต็มที่หรือไม่ในระหว่างสวิง” “ลูกบอลข้ามตาข่ายหรือไม่” และ “ลูกบอลตกลงบนที่เหมาะสมของสนามหรือไม่”

ขณะที่พัฒนาแบบตรวจสอบรายการดังกล่าว คุณต้องแน่ใจในการทำบัญชีรายชื่อของขั้นตอนที่สำคัญ แต่ละขั้นตอนที่จำเป็นต่อการแสดงออกของทักษะการใช้กล้ามเนื้อ มันไม่ใช่ง่ายที่จะตัดสินอย่างถี่ถ้วนในสิ่งที่หมายถึงขั้นตอนที่สำคัญและเราอาจทำผิดพลาดในทิศทางที่มีขั้นตอนน้อยเกินไป หรือมากเกินไปในแบบตรวจสอบรายการของเรา อย่างไรก็ตาม ขณะที่คุณพัฒนาแบบตรวจสอบรายการเพื่อประเมินการแสดงออกของทักษะการใช้กล้ามเนื้อ สิ่งสำคัญก็คือการพยายามระบุองค์ประกอบวิกฤติ (critical component) เหล่านั้นที่มีแนวโน้มจะให้ผลลัพธ์ในความสำเร็จ หรือความล้มเหลวในการแสดงออกทั้งหมดของทักษะ

ด้านทักษะการใช้กล้ามเนื้อ จุดประสงค์ควรระบุว่า การประเมินผลกระบวนการในตัวมันเองหรือไม่ (นั่นคือการนำทักษะไปใช้) ผลลัพธ์ของการแสดงออกของทักษะ หรือทั้งสองอย่าง ยกตัวอย่างเช่น ถ้าเราสอนการดำน้ำ มันเกือบเป็นสิ่งที่แน่นอนว่า การแสดงขององค์ประกอบต่างๆ ของการดำน้ำจะถูกประเมิน ในด้านตรงกันข้าม ถ้าเราตัดสินความสามารถของนักเรียนในการทำหม้อเซรามิก เราอาจไม่สนใจขั้นตอนจำเพาะในกระบวนการทำหม้อ แต่เราจะประเมินบนพื้นฐานของคุณภาพของผลิตผลสุดท้าย นั่นคือหม้อ สิ่งเหล่านี้คือการตัดสินที่คุณในฐานะครูต้องดำเนินการขณะที่คุณพิจารณาจุดประสงค์และการประเมินสำหรับเขา

เครื่องมือการประเมินสำหรับจุดประสงค์ทางเจตคติ (attitude objective)

ดังที่ได้ระบุไว้ในตอนต้นแล้ว เราพิจารณาเจตคติเป็นความรู้สึกและความเชื่อที่สะท้อนกลับมาในแนวโน้มที่จะแสดงออกในแนวทางใดแนวทางหนึ่งในสถานการณ์ที่กำหนด ประการแรก นักเรียน มีพฤติกรรมจริง ๆ ที่เราต้องการให้เขาแสดงออก จากนั้น เราสามารถตัดสินถ้าเขาเลือกที่จะนำพฤติกรรมเหล่านั้นไปใช้ เมื่อประเมินเจตคติ เราจะต้องให้สถานการณ์แก่นักเรียนซึ่งเจตคติหรือพฤติกรรมที่สะท้อนเจตคตินั้นแสดงออกอย่างเหมาะสมปกติ เราสามารถประเมินสิ่งดังกล่าวได้ดีที่สุดโดยผ่านการสังเกตโดยตรง

เราอาจต้องการประเมินความรู้สึกของนักเรียนเกี่ยวกับสถานการณ์ที่มันสัมพันธ์กับเจตคติของเขา ยกตัวอย่างเช่น ถ้าเราพยายามที่จะสอนเจตคติเกี่ยวกับเสรีภาพทางศาสนา เราอาจใช้ข้อสอบจำนวนหนึ่งที่เกี่ยวข้องกับความรู้ของนักเรียนในความเชื่อทางศาสนาต่างๆ ที่เขาพบเห็นในเมืองต่างๆ สิ่งดังกล่าวอาจจะระบุว่าเขาเคยตระหนักถึงความเชื่อทางศาสนาที่เราต้องการให้เขามีความอดทนหรือไม่ ถ้าเราต้องการสอนเด็กหนุ่มสาวให้มีส่วนร่วม เราอาจต้องการตัดสินว่า เขาสามารถอธิบายความหมายของ “มีส่วนร่วม” หรือไม่ อย่างไรก็ตาม มันเป็นเรื่องสำคัญที่จะสังเกตว่า การประเมินความรู้เกี่ยวกับเจตคติไม่ใช่การประเมินเจตคติ การประเมินที่เป็นจริงประการเดียวก็คือการแสดงออกในความชอบในส่วนหนึ่งของนักเรียนสำหรับพฤติกรรมที่เป็นตัวแทนของเจตคติ การที่สามารถอธิบายศาสนาต่างๆ ที่อาจพบในเมืองไม่เทียบเท่ากับการนับถือและปฏิบัติศาสนกิจในบรรดาคนที่มีความเชื่อในศาสนาที่แตกต่างไปจากศาสนาของตนเอง ในสถานการณ์ดังกล่าว อะไรที่ถูกประเมินและประเมินที่ไหน หรือภายใต้เงื่อนไขใดที่ก่อให้เกิดความแตกต่างในโลกระหว่างการแสดงให้เห็นในการรอบรู้ในทัศนที่เป็นนามธรรมชุดหนึ่ง และแสดงให้เห็นในความเข้าใจจริง ๆ

วัตถุประสงค์สำหรับการทบทวนมิติต่าง ๆ ของผลลัพธ์การเรียนรู้เกี่ยวกับการทดสอบก็คือการชี้ให้เห็น

เพื่อการนำเสนอ เขาก็อาจถูกบอกให้กล่าวสุนทรพจน์ในห้อง อย่างไรก็ตาม การประเมินตามสภาพจริงกำหนดให้นักเรียนกล่าวสุนทรพจน์ต่อหน้าหอการค้าในท้องถิ่นเพื่อแสดงให้เห็นในความรอบรู้ในจุดประสงค์ดังกล่าว เช่น มาตราส่วนประมาณค่าและพอร์ท โฟลิโอ ซึ่งสามารถใช้เป็นเครื่องมือสำหรับการประเมินตามสภาพจริงได้บ่อยๆ

ในบทความนี้ เราจะครอบคลุมโมทัศน์การทดสอบที่คุ้นเคยหลายๆ โมทัศน์ เราจะโฟกัสไปยังการพัฒนาข้อสอบและเครื่องมือการประเมินอื่นๆ ที่สอดคล้องกับจุดประสงค์โดยกำหนดให้นักเรียนแสดงออกในพฤติกรรมที่ได้อธิบายไว้ในจุดประสงค์ที่ได้กำหนด เราจะสนับสนุนให้พิจารณารูปแบบทางเลือกของการประเมินที่กำหนดให้นักเรียนใช้ทักษะในลักษณะที่เหมาะสม

เราอยากจะให้ข้อสังเกตในที่นี้เกี่ยวกับการใช้คำว่าแบบทดสอบ (test) กับเครื่องมือการประเมิน (assessment tools) ปกติ เราจะคิดถึงแบบทดสอบเลือกคำตอบ (multiple choice) หรือการสอบแบบอัตนัย (essay exams) ที่ใช้ในชั้นเรียนส่วนใหญ่ อย่างไรก็ตาม เครื่องมือประเมินอาจเป็นคำบอกลักษณะที่ดีกว่าในสิ่งที่เราจะอภิปรายกัน ในบทนี้เพราะเครื่องมือการประเมินหมายถึงวิธีที่กว้างกว่าของแนวทางในการตัดสินว่านักเรียนสามารถแสดงออกหรือมีความรอบรู้ในจุดประสงค์หรือไม่ เราจะใช้คำว่าทดสอบ และเครื่องมือการประเมินโดยสลับเปลี่ยนกัน โดยอ้างถึงกรณีใดกรณีหนึ่งต่อวิธีในการตัดสินสมรรถนะของนักเรียน

โมทัศน์และตัวอย่างที่สำคัญ

มีเหตุผลจำนวนหนึ่งในการทดสอบนักเรียน

1. ให้เกรดแก่นักเรียน
2. ตัดสินว่านักเรียนรู้อะไรเพื่อให้การซ่อมเสริมที่เหมาะสม
3. ระบุส่วนที่ไม่มีประสิทธิผลของการสอน

บางทีเหตุผลที่ง่ายที่สุดสำหรับการทดสอบก็คือการให้เกรด เป็นที่สันนิษฐานกันบ่อยๆ ว่า การให้เกรดแก่นักเรียนเป็นการจูงใจเขาให้แสดงออกได้ดีที่สุด เราทุกคนคงตระหนักถึงสถานการณ์ที่เราได้พยายามเพิ่มขึ้นเพื่อให้ได้เกรดที่ดี เราไม่อาจมองข้ามความสำคัญของเกรดในฐานะองค์ประกอบของการจูงใจสำหรับนักเรียนจำนวนมาก แม้จะไม่ใช่ว่าทุกคนก็ตาม

อย่างไรก็ตามในปัจจุบัน มีเหตุผลที่สำคัญอื่นๆ สำหรับการทดสอบนักเรียน หนึ่งในคือการตัดสินความก้าวหน้าของนักเรียนสู่ผลสัมฤทธิ์ของทักษะใดทักษะหนึ่ง ภายหลังได้ทำแบบทดสอบที่ได้ออกแบบมาเพื่อวัตถุประสงค์ดังกล่าวแล้ว นักเรียนจะได้รับผลสะท้อนกลับจำเพาะที่ได้ออกแบบมา เพื่อช่วยเขาให้เรียนทักษะที่เขาไม่บรรลุอย่างรอบรู้ ยกตัวอย่างเช่น การทดสอบเกี่ยวกับการบวกเลขเศษส่วน นักเรียนจะบวกเศษกับเศษ และส่วนกับส่วน โดยคิดแบบผิดๆ ว่า $\frac{1}{4} + \frac{1}{4} = \frac{2}{8}$ สิ่งที่สำคัญกว่าในการทดสอบดังกล่าวก็คือผลสะท้อนกลับที่จำเพาะที่นักเรียนได้รับ ในตัวอย่างเกี่ยวกับเศษส่วน ควรสังเกตธรรมชาติของความผิดพลาดของนักเรียนและเขาควรได้รับการทบทวนในแนวทางที่เหมาะสมในการบวกเลขเศษส่วนของส่วนเดียวกัน ในกรณีดังกล่าวข้อสอบไม่ได้ใช้เป็นวิธีการตัดสินนักเรียน แต่เป็นวิธีช่วยนักเรียนให้บรรลุจุดอ่อนในความเข้าใจของเขาเกี่ยวกับเนื้อหา และในฐานะเป็นวิธีชี้ทิศทางให้นักเรียนที่จะเป็นประโยชน์ในการซ่อมเสริมจุดอ่อนนั้น

ในลักษณะที่คล้ายคลึงกัน ข้อมูลการทดสอบอาจนำมาใช้ช่วยเหลือครู ปกติอาจสันนิษฐานได้ว่า เมื่อนักเรียนส่วนใหญ่ในชั้นทำข้อสอบได้ไม่ดีในจุดประสงค์ใดจุดประสงค์หนึ่ง จะมีความผิดพลาดบางอย่างเกี่ยวกับการสอน บ่อยครั้งที่ข้อมูลจากข้อสอบที่ได้ออกแบบมาเป็นอย่างดีสามารถช่วยชี้ให้เห็นว่า ปัญหาเหล่านั้นคืออะไร และอาจเสนอแนะแก้ไขที่จะทำให้การสอนมีประสิทธิภาพมากขึ้นในอนาคต

การทดสอบไม่เพียงแต่ใช้จูงใจนักเรียนและให้คะแนนแก่เขาเท่านั้น แต่ยังสามารถใช้ตัดสินความก้าวหน้าของนักเรียนในการบรรลุผลสัมฤทธิ์ในทักษะบางอย่าง และให้การซ่อมเสริมได้อย่างมีประสิทธิภาพ ผลมาก นอกจากนี้ การทดสอบยังสามารถนำไปใช้ในการให้ข้อมูลแก่ครูเพื่อปรับปรุงการสอนของเขา

ใครพัฒนาแบบทดสอบ

นักวิชาชีพที่ทำงานในบริษัทที่ผลิตแบบทดสอบใช้เวลาส่วนใหญ่ของเขาเขียนและปรับปรุงข้อสอบ เขาต้องสร้างข้อสอบเป็นจำนวนมาก และตรวจสอบข้อมูลจากนักเรียนเป็นพันๆ คนที่ได้ทำแบบทดสอบ นอกจากข้อสอบที่เตรียมโดยนักวิชาชีพเหล่านี้แล้ว ข้อสอบมากมายยังเขียนโดยผู้เขียนตำราเรียนซึ่งให้ข้อสอบสำหรับบทต่างๆ ในหนังสือของเขา ปกติข้อสอบเหล่านี้ปรากฏอยู่ในคู่มือครู อย่างไรก็ตามผู้เขียนตำราเรียนไม่จำเป็นต้องเป็นผู้เชี่ยวชาญในการเขียนแบบทดสอบ และข้อสอบที่เขาผลิตขึ้นอาจจะเป็นการประเมินพฤติกรรมที่ดีหรือไม่ดีที่เขาพยายามสอน นอกจากนี้ ในชุดตำราเรียนหลายๆ ชุด ข้อสอบที่ปรากฏอยู่ในคู่มือครูอาจไม่ได้เขียนโดยผู้เขียนตำราเรียน แต่อาจเขียนโดยคนอื่นที่ไม่มีความคุ้นเคยกับเนื้อหาของตำราเรียน

อีกแหล่งหนึ่งของข้อสอบก็คือครูในชั้นเรียน โปรแกรมวิชาชีพครูได้รวมการฝึกอบรมการพัฒนาข้อสอบไว้บ้าง ครูส่วนใหญ่ซึ่งขาดการฝึกอบรมดังกล่าวมีความเข้าใจไม่เพียงพอต่อมโนทัศน์การพัฒนาแบบทดสอบที่ดี และผลที่ตามมาคือการจบลงด้วยการพัฒนาแบบทดสอบเหมือนกับครูที่เคยสอนเขา เราหวังว่า วิธีการใหม่ในการพัฒนาข้อสอบดังที่ได้อธิบายในบทความนี้ จะช่วยแก้ไขสถานการณ์ดังกล่าว และคุณ将会เห็นความสัมพันธ์ที่สำคัญระหว่างการทดสอบกับการสอน

การทำให้ข้อสอบสอดคล้องกับจุดประสงค์

ข้อสอบที่ดีคือข้อสอบที่วัดในสิ่งที่ตั้งใจจะวัด โดยพื้นฐานแล้วสิ่งดังกล่าวก็คือมโนทัศน์ของความเที่ยงตรง (validity) ของแบบทดสอบหรือข้อสอบ ถ้ามีความสอดคล้องระหว่างพฤติกรรมที่อธิบายไว้ในจุดประสงค์และพฤติกรรมที่ประเมินโดยข้อสอบที่ออกแบบมาเพื่อวัดการบรรลุจุดประสงค์นั้นของนักเรียนก็อาจกล่าวได้ว่า ข้อสอบมีความเที่ยงตรง ยกตัวอย่างเช่น ถ้าจุดประสงค์ระบุว่า “นักเรียนจะสามารถบวกเลขสองหลักได้” ดังนั้นข้อสอบที่มีความเที่ยงตรงจะเป็นข้อสอบที่กำหนดให้นักเรียนบวกเลขสองหลัก มันจะเป็นการไม่เหมาะสมถ้าเราบอกนักเรียนให้บวกเลขสามหลัก

ในระบบที่ได้นำเสนอในที่นี้ คุณเริ่มด้วยเป้าหมายประสงค์และจุดประสงค์ จากนั้นจุดประสงค์ของคุณจะถูกใช้เป็นพื้นฐานในการพัฒนาข้อสอบ กริยาในจุดประสงค์อธิบายชนิดของพฤติกรรมที่คุณคาดหวังจากนักเรียนและที่คุณจะประเมินในแบบทดสอบ ยกตัวอย่างเช่น ถ้าจุดประสงค์ระบุว่านักเรียนจะสามารถ “ทำบัญชีรายการ (list)” ดังนั้น เราจะคาดหวังว่าข้อสอบจะถามนักเรียนให้ “ทำบัญชีรายการ” ข้อมูลบางอย่าง ขณะที่มันเป็นไปไม่ได้เสมอในการใช้กริยาในจุดประสงค์ในการสร้างข้อสอบโดยตรง แต่มันให้แนวทางแก่เราว่าข้อสอบชนิดไหนมีความเหมาะสมในการวัดจุดประสงค์

สิ่งที่สำคัญเช่นเดียวกันก็คือเงื่อนไข (condition) ที่คุณรวมเข้าไว้ในข้อสอบสอดคล้องกับเงื่อนไขที่ระบุไว้ในจุดประสงค์อย่างจำเพาะในตอนต้นๆ เราใช้ตัวอย่างของการกล่าวสุนทรพจน์ เงื่อนไขที่สิ่งดังกล่าวทำสำเร็จอาจแตกต่างกันไปเป็นอย่างมากและอาจมีอิทธิพลต่อความสามารถของนักเรียนในการแสดงออกที่เห็นได้ชัดเจน การกล่าวสุนทรพจน์ในบ้านอาจง่ายกว่าการกล่าวสุนทรพจน์ต่อกลุ่มที่นักเรียนไม่เคยพบมาก่อน อย่างไรก็ตาม สุนทรพจน์มักกล่าวกับกลุ่มที่เรารู้จักคนเพียงไม่กี่คนเท่านั้น ดังนั้น จุดประสงค์ของคุณควรระบุว่าสุนทรพจน์จะ

ขั้นตอนแรกในการพัฒนาข้อสอบสำหรับจุดประสงค์ใด ๆ ก็คือ การระบุจุดประสงค์อยู่ในมิติใด เมื่อคุณได้ทำสิ่งดังกล่าวแล้วคุณก็สามารถทำให้ข้อสอบแคบเข้าซึ่งสามารถประเมินพฤติกรรมที่อธิบายไว้ในจุดประสงค์ได้ดีที่สุด เราได้สรุปข้อมูลดังกล่าวข้างล่าง

วิธีการทดสอบที่เป็นแบบอย่าง

มิติ	ชนิดของเครื่องมือที่ใช้ประเมิน
ความรู้	ข้อสอบแบบเลือกคำตอบ แบบเติมคำ
ทักษะสติปัญญา	ข้อสอบแบบเลือกคำตอบ แบบเติมคำ เรียงความหรือผลผลิต (ด้วยแบบตรวจสอบรายการหรือมาตราส่วนประมาณค่า)
ทักษะการใช้กล้ามเนื้อ	แบบตรวจสอบรายการหรือมาตราส่วนประมาณค่า
เจตคติ	การสังเกตพฤติกรรม แบบสอบถามเกี่ยวกับพฤติกรรม

เราสามารถใช้การสรุปดังกล่าวเพื่อช่วยในการพิจารณาปัญหาที่เกิดขึ้นกับครูสันติ เราทราบว่าเขาควรตรวจสอบความสัมพันธ์ระหว่างจุดประสงค์จำเพาะกับข้อสอบและตัดสินใจว่ามันมีความสอดคล้องกันหรือไม่ มันอาจมีความสอดคล้อง และข้อสอบไม่ควรถูกวิจารณ์เพียงเพราะมันอยู่ในรูปแบบเลือกคำตอบ อย่างไรก็ตามถ้าครูสันติพบว่า จุดประสงค์และข้อสอบไม่สอดคล้องกัน เขาต้องเขียนข้อสอบใหม่ เทคนิคในการทำสิ่งดังกล่าวจะได้อธิบายในตอนต่อไป

แนวทางสำหรับการพัฒนาเครื่องมือการประเมิน

ในบทความนี้ เราไม่อาจให้กฎที่สมบูรณ์ในวิธีพัฒนาแบบทดสอบและข้อสอบชนิดต่าง ๆ เราสนับสนุนให้คุณตรวจสอบแหล่งอื่น ๆ เพื่อจะได้เข้าใจกระบวนการดังกล่าวให้ละเอียดมากขึ้น อย่างไรก็ตาม มันเป็นข้อสรุปที่ได้ตระหนักถึงขั้นตอนที่สำคัญที่สุดในการสร้างข้อทดสอบในพฤติกรรมที่ได้อธิบายไว้ในจุดประสงค์และการเขียนข้อสอบที่ประเมินความสามารถของนักเรียนในการแสดงออกในพฤติกรรมนั้น อย่างไรก็ตามในการเขียนข้อสอบดังกล่าว จะต้องทราบกฎหรือแนวทางเบื้องต้นบางอย่างของการสร้างข้อสอบที่ดี

เครื่องมือการประเมินข้อเขียน

ถ้าพฤติกรรมในจุดประสงค์ของคุณระบุว่านักเรียนจะ “ระบุ” หรือ “คัดเลือก” บางสิ่งบางอย่าง ดังนั้น เราสามารถใช้ข้อสอบแบบเลือกคำตอบในการประเมินนักเรียน ในการพัฒนาข้อสอบแบบเลือกคำตอบมีกฎเบื้องต้นสองข้อที่ต้องยึด กฎข้อแรกก็คือใส่ข้อมูลให้มากที่สุดในส่วนที่นักเรียนอ่านเพื่อเลือกคำตอบ

ตัวอย่าง

ดี

ในสหรัฐอเมริกา เครื่องบินบินได้เร็วขึ้นถ้ามันบินในทิศทางใด

1. ทิศเหนือ
2. ทิศใต้
3. ทิศตะวันออก
4. ทิศตะวันตก

ไม่ได้

เครื่องบินบิน

1. ช้าลงเมื่อมุ่งหน้าไปทางทิศเหนือ
2. เร็วขึ้นเมื่อมุ่งหน้าไปทางทิศตะวันตก
3. เร็วขึ้นเมื่อบินไปทางทิศตะวันออก
4. ในอัตราเดียวกันไม่ว่าจะบินไปในทิศทางใด

กฎข้อที่สองในการพัฒนาข้อสอบแบบเลือกคำตอบก็คือการให้ตัวเลือกมีความยาวใกล้เคียงกัน โดยไม่มีตัวชี้แนะซึ่งจะทำให้ผู้ตอบเลือกคำตอบที่ถูกต้องโดยไม่ทราบจริง ๆ

ทางเลือกของเครื่องมือการประเมิน

บางที การประเมินที่ง่ายที่สุดก็คือการสังเกตนักเรียนโดยตรง ประเด็นก็คือ อะไรที่ถูกสังเกตปกติ การสังเกตเกิดขึ้นภายหลังการสอนโดยตรงได้เกิดขึ้น และนักเรียนปฏิบัติในสิ่งที่เขาได้เรียนรู้ การปฏิบัติดังกล่าวอาจแตกต่างกันไปจากเอกสารการฝึกปฏิบัติข้อมูลทางภาษาจนถึงการทำแผนที่และทักษะการใช้กล้ามเนื้อ ในชั้นเรียนพลศึกษาจนถึงพฤติกรรมที่พึงประสงค์ที่สะท้อนในจุดประสงค์ของเจตคติ วัตถุประสงค์หลักของการสังเกตของครูก็คือการให้ผลสะท้อนกลับแก่ผู้เรียนในความเหมาะสมของการแสดงออก และให้การสอนซ่อมเสริมตามที่จำเป็น ถ้านักเรียนมากเกินไป ประสบกับปัญหาเดียวกัน ครูอาจต้องใช้การสอนตรงในทักษะการตั้งคำถาม

เรายังไม่ได้ตอบคำถามว่าอะไรที่ถูกสังเกต มันควรเป็นพฤติกรรมที่อธิบายไว้ในจุดประสงค์ บางครั้งการสังเกตดังกล่าวง่ายมาก เหมือนกับคำถามข้อมูลภาษา หรือปัญหาทักษะสติปัญญา แต่การเตรียมโครงการหรือเขียนเรียงความละ ปกติครูใช้แนวทางการประเมินแบบองค์รวม (holistic assessment) ซึ่งหมายความว่า เขาได้รับความประทับใจทั้งหมดของความก้าวหน้าของนักเรียนแต่ละคนแล้วดำเนินการให้ข้อคิดเห็นที่ดูเหมือนจะมีความเหมาะสม มีปัญหาบ้างกับแนวทางที่ค่อนข้างไม่เป็นโครงสร้างดังกล่าว

โดยสาระแล้ว ขณะที่สังเกตพฤติกรรมของนักเรียนใด ๆ ครูควรมีแบบตรวจสอบรายการของคำถามเพื่อการทบทวน นักเรียนเขียนเรียงความสะท้อนคุณลักษณะอะไรในจุดประสงค์ มีการเน้นในลีลาการเขียนหรือเนื้อหา ครูควรหาความผิดพลาดในข้อเท็จจริง หรือในการสรุป หรือในการรวมย่อหน้า สิ่งเหล่านี้ดูเหมือนจะเป็นการพิจารณาที่เด่นชัด แต่มันอาจหลงทางในสิ่งสำคัญในกระบวนการสังเกตถ้าไม่ได้ทำให้พฤติกรรมที่ถูกสังเกตมีความชัดเจน

เนื่องจากการสังเกตเป็นรูปแบบปกติของการประเมินผลอย่างไม่เป็นทางการ เราควรเพิ่มเหตุผลอื่นสำหรับระบุสิ่งที่ถูกสังเกต ซึ่งได้แก่ความเป็นไปได้ของอคติ ครูส่วนใหญ่ยืนยันโดยทันทีว่าไม่มีอคติ แต่การวิจัยเสนอแนะว่า มันตอบสนองแตกต่างกันไปสำหรับนักเรียนแต่ละคน ซึ่งบางครั้งสิ่งดังกล่าวมีความชัดเจนมากที่สุดกับนักเรียนที่มีความสามารถสูงและต่ำ นักเรียนที่มีความสามารถสูงจะสร้างความประทับใจในทุกเรื่องขณะที่เด็กเรียนช้าดูเหมือนจะทำผิดอยู่เสมอ ในการสังเกต ถ้าท่านรู้จักนักเรียนเป็นอย่างดีจะก่อให้เกิดอคติดังกล่าว ซึ่งได้แก่การให้ผลสะท้อนกลับโดยตั้งอยู่บนพื้นฐานว่านักเรียนคือใคร มากกว่านักเรียนได้ทำอะไร การมีรายการพฤติกรรมหรือคุณลักษณะจะช่วยกำจัดสิ่งดังกล่าวซึ่งบ่อย ๆ เป็นรูปแบบของอคติโดยไม่ตั้งใจ

ครูควรทำอะไรถ้าเป้าหมายของหน่วยมีไว้สำหรับให้นักเรียนได้พัฒนาผลผลิตดั้งเดิมบางชนิด ไม่ว่าจะเป็นด้านสติปัญญาหรือทางกาย หรือการเน้นอยู่ที่การแสดงออกในทักษะที่เป็นตัวแทนของชุดขั้นตอนที่จะทำให้สำเร็จ

เช่นการเดินร่ำ ถ้าจุดประสงค์กำหนดให้นักเรียนได้แสดงออกในกระบวนการหรือผลผลิต มันจะเป็นการเหมาะสม
เสมอในการประเมินกระบวนการหรือผลผลิตด้วยมาตราส่วนประมาณค่า

มาตราส่วนประมาณค่า (rating scale) ประกอบด้วยองค์ประกอบที่สำคัญสององค์ประกอบ หนึ่งบัญชี
รายการขั้นตอนในกระบวนการหรือคุณลักษณะของผลผลิต และสองสเกลสำหรับประมาณค่าแต่ละขั้นตอนหรือ
คุณลักษณะคุณจะได้แนวคิดของสิ่งดังกล่าวโดยคิดถึงคำถามที่ภักดาการหรือโรงแรมที่ขอให้คุณประมาณค่า
ลักษณะต่างๆ ของบริการในรูปของดีเยี่ยม ดี ปานกลาง แย่ หรือ รับไม่ได้ นี่เป็นตัวอย่างที่ดีของมาตราส่วนประมาณ
ค่า สำหรับการประเมินผล ไม่ว่าจะเป็นบริการหรือผลผลิต ครูสามารถใช้แบบฟอร์มเดียวกันในการประเมินการ
แสดงออกของนักเรียนในจุดประสงค์ของการสอนต่างๆ

เรามาดูตัวอย่างที่เราต้องประเมินผลรายงานโครงการของนักเรียน ถ้าจุดประสงค์ของคุณระบุว่า นักเรียน
จะสามารถเขียนรายงานโครงการ คุณจะประเมินผลอย่างไร ประการแรก คุณจะอ้างถึงจุดประสงค์เพื่อตัดสินว่า
มีเกณฑ์ใดบ้างที่ระบุไว้ที่นั่น เนื้อหา ความยาว ลีลา และ/หรือการใช้การอ้างอิงมีระบุไว้ในเกณฑ์หรือไม่ ถ้ามี มาตราส่วน
ประมาณค่าของคุณจะรวมรายละเอียดเนื้อหาเหล่านั้น คุณได้ระบุไว้ในการสอนใหม่ว่า รายงานควรมีส่วนสำคัญ
ต่างๆ เช่น บทนำ การอธิบายทรัพยากรที่ใช้ การระบุประเด็นสำคัญ และการจบลงที่เชื่อถือได้ เนื้อหาเหล่านี้ควร
อยู่ในมาตราส่วนประมาณค่าของคุณเช่นเดียวกัน เราจะต้องประเมินองค์ประกอบที่สำคัญแต่ละองค์ประกอบ
ของรายงานและเนื้อหา ความยาว ลีลา และการอ้างอิงที่เหมาะสม เราจะต้องตัดสินใจสองครั้งเกี่ยวกับแต่ละ
องค์ประกอบ ซึ่งมีหรือไม่ และดีแค่ไหน ดังนั้น แต่ละรายละเอียดในสเกลต้องระบุการมีองค์ประกอบและคุณภาพ
และเราทำสิ่งดังกล่าวโดยการสร้างสเกลเพื่อประมาณค่าของละเอียด ดูภาพ 1 เป็นตัวอย่างของมาตราส่วน
ประมาณค่าเพื่อการประเมินผลรายงานโครงการที่ใช้เกณฑ์ระบุไว้

ภาพที่ 1 การประเมินผลรายงานโครงการ

ชื่อนักเรียน..... วันที่.....

	ไม่เพียงพอ	เพียงพอ	ดี	ดีเยี่ยม
	0	1	2	3
เนื้อหา				
ความยาว				
รายการอ้างอิง				
บทนำ				
การอธิบายทรัพยากร				
ประเด็นสำคัญ				
การสรุปที่เชื่อถือได้				
รวม				

ทางเลือกของเครื่องมือการประเมินผลอีกอย่างได้แก่ฟอร์ท โพลีโอของผลผลิตที่นักเรียนได้พัฒนาใน
รายวิชา ปัจจุบัน ครูจำนวนมากได้ใช้ฟอร์ท โพลีโอเป็นองค์ประกอบที่สำคัญของกระบวนการประเมินผล และเป็น

ตัวนำสำหรับการสื่อความหมายกับทั้งนักเรียนและผู้ปกครองเกี่ยวกับความก้าวหน้าของนักเรียนเรียนรายวิชา

ในรายวิชาส่วนใหญ่ นักเรียนแต่ละคนจะมีแฟ้มที่รวมตัวอย่างของงานของเขา ซึ่งอาจเป็นการทดสอบแบบดั้งเดิมหรือรายงานสิ่งที่พิมพ์ออกจากคอมพิวเตอร์ รูปภาพ หรือผลผลิตอื่นๆ ที่สะท้อนจุดประสงค์ของรายวิชา คุณสามารถตัดสินล่วงหน้าว่ารายละเอียดใดที่จะนำไปใส่ไว้ในพอร์ต โพลีโอสำหรับนักเรียนทุกคนหรือคุณอาจคุยกับนักเรียนเป็นรายบุคคลเพื่อระบุงานที่ดีที่สุดของเขาเพื่อใส่ไว้ในพอร์ต โพลีโอ

ถ้านักเรียนแต่ละคนแสวงหาเป้าประสงค์และจุดประสงค์ส่วนบุคคลในรายวิชาของคุณ ความแตกต่างเหล่านี้สามารถสะท้อนออกมาในพอร์ต โพลีโอได้โดยง่าย พอร์ต โพลีโอควรรวมงานที่ระบุในสิ่งที่นักเรียนรู้หรือสามารถทำได้เมื่อเขาเริ่มชั้นเรียนและวิธีที่ทักษะของเขาได้พัฒนาในเวลาหนึ่ง คุณค่าที่ยิ่งใหญ่ที่สุดของพอร์ต โพลีโอก็คือมันสะท้อนการเปลี่ยนแปลงในตัวนักเรียน ขณะที่ทักษะเจริญงอกงามมากขึ้นและซับซ้อนขึ้น ควรเลือกการเปรียบเทียบรายละเอียดเพื่อสะท้อนการเปลี่ยนแปลงเหล่านี้

แต่ละข่าวสารในพอร์ต โพลีโอถูกประเมินเพราะมันถูกผลิตโดยนักเรียน การประเมินผลแบบที่สองและกว้างขวางกว่าอาจเกิดขึ้นเมื่อพอร์ต โพลีโอได้รับการพิจารณาเป็นระยะเวลายาวนานกว่า เช่นหกสัปดาห์หรือหนึ่งภาคการศึกษา มันสามารถให้พื้นฐานสำหรับการอภิปรายเชิงสร้างสรรค์ระหว่างคุณกับนักเรียนและพ่อแม่ของนักเรียนเกี่ยวกับจุดแข็งของการแสดงออกของนักเรียนในรายวิชาดังกล่าว ขณะที่สามารถให้การวิพากษ์วิจารณ์เชิงสร้างสรรค์แก่นักเรียนบนพื้นฐานของผลผลิตในพอร์ต โพลีโอ จุดโฟกัสที่สำคัญสำหรับนักเรียนควรเป็นความสำเร็จที่สะท้อนโดยพอร์ต โพลีโอ ก็เหมือนกับเครื่องมือการประเมินอื่นๆ ครูสามารถเรียนรู้จากสิ่งต่างๆ ที่นักเรียนส่วนใหญ่ไม่ได้เรียนรู้ที่สะท้อนในพอร์ต โพลีโอ ข้อมูลดังกล่าวสามารถใช้ทำการสอนนักเรียนกลุ่มต่อไปให้เข้มแข็งขึ้น

สรุป

ในบทความนี้ เราได้ทบทวนความสำคัญทั้งหมดของการทดสอบในกระบวนการออกแบบการสอน เราได้เน้นข้อเท็จจริงที่ว่าข้อสอบมีความสำคัญเป็นอย่างยิ่งในการประเมินความก้าวหน้าของนักเรียนในการเรียน จุดประสงค์ และผลการทดสอบให้ข้อมูลกับครูเกี่ยวกับคุณภาพของการสอนที่เขาได้นำเสนอ ตลอดจนเสนอแนะวิธีปรับปรุงการสอนในอนาคต

ขณะที่เราใช้แนวทางที่เป็นระบบในการวางแผนการสอน คุณลักษณะที่สำคัญที่สุดของข้อสอบก็คือมันสอดคล้องกับพฤติกรรมที่อธิบายไว้ในจุดประสงค์ที่ข้อสอบตั้งใจจะทดสอบ และมันสะท้อนเงื่อนไขที่นักเรียนจะใช้ทักษะเหล่านี้ในเวลาต่อมา เราได้ให้กฎต่างๆ สำหรับพัฒนาข้อสอบแบบข้อเขียน และทางเลือกของข้อสอบการประเมินและรวมเข้าเป็นการทดสอบ เรายังได้อภิปรายคุณค่าของการสร้างพอร์ต โพลีโอของนักเรียนที่สะท้อนการพัฒนาที่ก้าวหน้าของทักษะของเขา

References

- Burke, K.(1993). **The mindful school : How to assess authentic learning**. Palatine, IL:ITI/Skylight Publishing.
- Diversity and Equity in Assessment Network (DEAN) (1993). **Guidelines for equitable assessment**. Cambridge, MA:FairTest.