Testing of English as a Second Language

โดย

David P. Harris 1969

ในส่วนที่เป็นบทนำของหนังสือเล่มนี้ผู้เขียนได้ระบุวัตถุประสงค์ไว้อย่างชัดแจ้งแล้วว่า หนังสือเล่มนี้มุ่งที่จะช่วยให้ผู้สอนสามารถปรับปรุงวิธีการวัดผลของตนเอง ขณะเดียวกันก็สามารถที่จะประเมินข้อสอบมาตรฐานใด ๆ ก็ตามที่บังเอิญต้องใช้หรือต้องให้ความคิดเห็นเกี่ยวกับข้อสอบนั้น ๆ ได้ จึงเป็นความจำเป็นที่ผู้สอนจะต้องมีเกณฑ์หรือมาตรฐานบางประการเป็นหลักยึด หนังสือเล่มนี้ได้เสนอเนื้อหาที่ครอบคลุมแทบจะทุกแง่ทุกมุมที่เกี่ยวกับกระบวนการวัดผล โดยนัยนี้จึงจัดเป็นหนังสือที่มีคุณค่ามากพอที่สมควรจะพิจารณาอ่านในรายละเอียด

เนื้อหาของหนังสือแบ่งออกเป็น 12 บทใหญ่ ใน 2 บทแรกผู้อ่านจะได้ทราบถึงวัตถุประสงค์และวิธีการวัดผลทางภาษาโดยทั่ว ๆ ไป พร้อมทั้งจะได้เห็นลักษณะสำคัญของการวัดผลใน 6 บทต่อมาผู้เขียนจะบรรยายถึงรายละเอียดของการวัดผลทักษะต่าง ๆ เช่น การอ่าน การเขียน การฟัง และองค์ประกอบที่สำคัญของภาษาอังกฤษ ซึ่งในช่วง 6 บทนี้นับเป็นหัวใจสำคัญของหนังสือ สำหรับในบทที่ 9 และ 10 นั้น ผู้เขียนได้เสนอแนะกระบวนการและขั้นตอนในการสร้างและเตรียมข้อสอบรวมทั้งการจัดสอบและใน 2 บทสุดท้าย ผู้อ่านจะได้เรียนรู้ถึงวิธีการคำนวณหาค่าทางสถิติบางประการที่เกี่ยวข้องกับการวัดผลซึ่งมีประโยชน์ในแง่ที่ช่วยให้ผู้สอนประเมินข้อสอบและความสามารถของผู้สอบได้ถูกต้องและแม่นยำยิ่งขึ้น

ประเด็นใหญ่ของเนื้อหาในช่วงแรกคือลักษณะที่ดีของข้อสอบที่ดี กล่าวคือข้อสอบนั้นจะต้องมีความน่าเชื่อถือ (reliability) และความเที่ยงตรง (validity) เพราะคุณสมบัติของข้อสอบที่ดีจะต้องวัดได้อย่างเที่ยงตรงไม่ว่าจะนำข้อสอบนั้นมาใช้เมื่อไรก็ตาม แฮริส ได้เสนอวิธีวัดความเที่ยงของข้อสอบไว้ว่า วิธีแรกอาจจะโดยการใช้ข้อสอบชุดเดียวกันให้ผู้สอบคนเดียวกันสอบ แต่

ทิ้งช่วงเวลาให้ห่างกันสักพัก แต่วิธีนี้มีข้อเสียคือ ผู้สอบอาจใช้ความจำได้ วิธีที่สองอาจจะโดยการสร้างข้อสอบคู่ขนาน และวิธีที่สามโดยการให้สอบข้อสอบทั้งชุดก่อนแล้วจึงแยกข้อสอบทั้งชุดออกเป็น 2 ส่วน (อาจจะแยกตามเลขคู่คี่) แล้วให้สอบใหม่เราก็จะได้คะแนนมาเปรียบเทียบกันได้ เป็นต้น

ในส่วนที่เกี่ยวกับความเที่ยงตรงของข้อสอบนั้น แฮริส อธิบายไว้ว่าสิ่งหนึ่งที่เราจะต้องถามตัวเองเมื่อเลือกหรือออกข้อสอบใด ๆ คือ เราต้องการวัด*อะไร* และข้อสอบนั้นจะวัดสิ่งที่เราต้องการได้*ดีเพียงไร* ตัวอย่างเช่นหากเราต้องการวัดความสามารถในการวิเคราะห์เสียงในภาษาอังกฤษ และคะแนนที่วัดได้ มีความสัมพันธ์กับความสามารถในทักษะที่เราต้องการวัด เราก็อาจจะกล่าวได้ว่า ข้อสอบนั้น ตรงกับวัตถุประสงค์ของเรา

นอกจากนี้แฮริสยังมีความละเอียดลออพอที่จะรวมความสะดวกในการใช้ (practicality) ไว้เป็นคุณสมบัติของข้อสอบที่ดีอีกประการหนึ่งด้วย โดยให้เหตุผลว่าถึงแม้ข้อสอบนั้นจะมีความเที่ยงตรงสักปานใดก็ตาม ก็อาจจัดเป็นข้อสอบที่ไม่ดีได้ ทั้งนี้ด้วยข้อสอบอาจจะไม่สะดวกแก่การใช้หรือการตรวจหรืออาจต้องใช้ต้นทุนในการผลิตสูงจนเกินไปก็ได้ เป็นต้น

ข้อที่น่านิยมของหนังสือเล่มนี้มิได้อยู่แต่เพียงครอบคลุมเนื้อหาได้ละเอียด หากแต่ยังจัดแบ่งเนื้อเรื่องได้เด่นชัด โดยแยกพูดถึงการวัดผลแต่ละทักษะเป็นบท ๆ ไป แต่ละบทสมบูรณ์ในตัวเอง โดยจะเริ่มด้วยการให้แนวทางรูปแบบของข้อสอบว่าในการวัดทักษะนั้น ๆ เราจะมีวิธีตั้งคำถามในรูปใดบ้าง เช่น จะเป็น multiple-choice ในลักษณะของ completion, sentence alternatives, sentence interpretation หรือในรูปของ supply type หรือการถามด้วยรูปภาพและอื่น ๆ เป็นต้น หลังจากนั้นก็จะเสนอข้อวิจารณ์ต่อลักษณะข้อสอบแบบนั้น ๆ ว่า มีข้อเสียตรงไหน และทุกครั้งก็จะมีข้อเสนอแนะให้ด้วย เพื่อผู้อ่านจะได้ใช้วิจารณญาณเลือกใช้หรือดัดแปลงตามที่เห็นสมควร

แฮริส พูดถึงข้อสอบที่วัดโครงสร้างทางไวยากรณ์ (grammatical structure) เป็นอันดับแรก รูปแบบของข้อสอบเท่าที่เสนอไว้ในบทที่ 3 จะมีลักษณะเป็น multiple-choice ในรูปของ completion เช่น

Mary (lives) (is living) (has lived) in New York since 1960.

หรือ Mary_____in New York since 1960.

a. lives b. is living

c. has lived

หรือในรูปของ sentence alternatives โดยเขียน item ให้เป็นประโยคที่สมบูรณ์ทั้ง 3 ตัวเลือก แต่ใช้ tense ต่างกันเป็นต้น ลักษณะอีกแบบหนึ่งก็คือ sentence interpretation ได้แก่การหา pronoun reference เช่น He ในที่นี้ refer ถึงใครเป็นต้น ส่วนข้อสอบที่เป็น scrambled sentence นั้น ข้อสอบจะสลับคำในประโยคและให้ผู้สอบเรียงลำดับคำที่ถูกต้องเช่น

When_____?

a. plan b. to go

c. do d. you

แต่สำหรับวิธีนี้ แฮริส วิจารณ์ไว้ว่า จะเหมาะสมในฐานะที่เป็นแบบฝึกหัดหรือ informal test แต่หากเป็นระดับสูงขึ้นไป จะมีข้อจำกัดในตัวเองเช่นเราจะสร้าง item ที่มีความยากพอดีได้ไม่ง่ายนัก เพราะโจทย์จะง่ายไป ครั้นจะให้ item ยาวและซับซ้อนขึ้นก็จะกลายเป็นการทดสอบสติปัญญาไป

completion แบบ supply type ซึ่งเป็นลักษณะ fill–in exercise เป็นอีกลักษณะหนึ่งที่จะวัด โครงสร้างทางไวยากรณ์ได้ ในข้อสอบแบบนี้ ผู้สอบจะได้รับคำสั่ง ให้ทำประโยคให้สมบูรณ์โดยใช้คำที่ให้ไว้ในวงเล็บในรูปที่ถูกต้อง ซึ่งวิธีนี้ก็เป็นวิธีที่ใช้กันอยู่แพร่หลาย

ในหัวข้อที่เกี่ยวกับคำแนะนำในการสร้าง item แฮริสแนะว่าภาษาที่ใช้ใน item ที่เป็นบทสนทนา ควรเป็นภาษาที่ใช้ในภาษาพูดจริง ๆ ไม่ใช่เป็นภาษาเขียน คำตอบควรเป็นคำตอบที่ฟังดูเป็นธรรมชาติ ตัวลวง (distractor) ก็ไม่ควรเป็นคำตอบที่เขียนให้ผิดไวยากรณ์ในเชิงภาษาเขียน เป็นต้น

ผู้อ่านที่สนใจการวัด auditory discrimination and comprehension จะหาอ่านได้ในบทต่อมา วิธีหนึ่งก็คือ การให้ผู้อ่านเลือกวงว่า คำ หรือ เสียง ที่ได้ยินเป็นเสียงไหน โดยคำ ๆ นั้น อาจจะอยู่โดด ๆ หรือ อยู่ในปริบท (context) ก็ได้ และบางครั้งคำตอบที่ให้เลือกอาจจะมา

ในรูปของรูปภาพก็ได้ เช่นเมื่อ ได้ยินว่า 'It was a large ship.' ก็ต้องเลือกรูปเรือไม่ใช่รูปแกะเป็นต้น ส่วนการวัด ความเข้าใจ อาจจะเริ่มจากขั้นง่าย ๆ โดยการให้ทำตามคำสั่งที่ได้ยิน หรือผู้สอบอาจจะได้ยินคำถามสั้น ๆ และต้องเลือกคำตอบที่สัมพันธ์กับคำถามนั้น อีกวิธีหนึ่งที่ใช้กันก็คือ ผู้สอบจะได้ยินคำบรรยายสั้น ๆ และจะต้องเลือกภาพที่ตรงกับคำบรรยายนั้น แต่การใช้ภาพก็มีข้อเสียตรงที่ว่าจะวาดให้ชัดเจนได้ยากในกรณีที่โจทย์ยากและซับซ้อน บางครั้งตัวลวงก็อาจจะกลายเป็นคำตอบที่ใช้ได้ อีกประการหนึ่งนั้นการทดสอบด้วยภาพเช่นนี้จะทำได้ในระดับหนึ่งเท่านั้น ในระดับสูงขึ้นไปวิธีการนี้จะใช้ได้ยาก และประการสุดท้ายการทำข้อสอบลักษณะนี้จะสิ้นเปลืองค่าใช้จ่ายและเสียเวลามาก

ลักษณะการวัดผลอีกอย่างคือ ข้อสอบจะมาในรูปสนทนาของคน 2 คน และจะมีเสียงที่ 3 ถามเกี่ยวกับบทสนทนานั้น และในขั้นที่สูงขึ้นไปนั้นจะวัดความเข้าใจโดยการให้ฟัง การบรรยาย

แฮริสให้ข้อแนะนำในการเขียน item การวัดผลการฟังไว้ว่า ทั้งคำถามและคำตอบควรเป็นภาษาพูดให้มากที่สุด ศัพท์ที่ใช้ควรเป็นศัพท์ที่ใช้บ่อย เพราะเรากำลังวัดความเข้าใจประโยคหรือเนื้อเรื่องทั้งหมดไม่ได้วัดศัพท์คำตอบที่ให้เลือกก็ควรจะสั้น และท้ายสุด แต่ละประโยคที่อยู่ในเนื้อเรื่องที่ผู้สอบจะได้ยิน ควรมีความสัมพันธ์กันไม่ใช่เจตนาใส่มาเพื่อลวงหรือหันเหความสนใจของผู้ฟังเป็นต้น

บทที่นับว่าน่าสนใจอีกบทคือบทที่ 5 อันว่าด้วยการวัดศัพท์โดยเฉพาะ ข้อสอบลักษณะเก่าที่ถามเรื่องศัพท์ จะมาในรูปของการให้ศัพท์มาและให้เลือกคำจำกัดความหรือไม่ก็ในทางกลับกันให้คำจำกัดความมา และให้เลือกว่าตรงกับศัพท์ข้อใด แต่วิธีดังกล่าวนี้เมื่อนำมาเปรียบกับวิธี completion (multiple-choice) หรือ วิธี paraphrase (multiple-choice) แล้ว มีวิธีหลังนี้จะดีกว่าด้วยเหตุผลที่ว่าผู้สอบจะมีโอกาสเห็นคำนั้นเมื่ออยู่ในประโยคเลย เช่น

The old woman was too _____ to push open the heavy door.

a. feeble
b. sincere
c. deaf
d. harsh

หรือ John was astounded to hear her answer.

a. greatly amused
b. greatly relieved
c. greatly surprised
d. greatly angered

แฮริสยังได้แนะวิธีสร้าง item อีกด้วย เช่น คำจำกัดความควรจะใช้ภาษาง่าย ๆ หรือศัพท์ที่ให้มาเลือกควรอยู่ในระดับความยากง่ายพอกัน หรือหากเป็นไปได้ควรอยู่ใน area เดียวกัน เพื่อป้องกันการเดาของผู้สอบ ตัวเลือกที่ให้ยังควรระวังเรื่องตัวสะกดเพื่อไม่ให้ผู้สอบไขว้เขวและควรมีความยาวระดับเดียวกันด้วย

บทที่นับว่าสำคัญที่สุดสำหรับผู้สอนการอ่านทั้งหลายน่าจะได้แก่บทที่ว่าด้วยการวัดผลความเข้าใจในการอ่าน แฮริสเริ่มด้วยการกล่าวถึงความสามารถที่จำเป็นในการอ่าน เป็นต้นว่าต้องเข้าใจคำ (lexical items) และโครงสร้างของภาษา รวมทั้งสัญญลักษณ์ในการเขียนต่าง ๆ เช่น เครื่องหมายวรรคตอน การใช้ตัวเอนตัวหนาและอื่น ๆ นอกจากนั้นยังจะต้องเข้าใจความคิดหลักและวัตถุประสงค์ของผู้เขียนอีกด้วย เมื่ออ่านจบแล้วควรจะสามารถสรุปความคิดของผู้เขียนได้ ความสามารถเข้าถึงน้ำเสียง และลีลาการเขียนก็มีความสำคัญในการอ่านไม่ยิ่งหย่อนกว่ากัน

ในการเลือกเรื่องที่จะมาทำเป็นข้อสอบนั้น ควรจะต้องคำนึงถึงความยาวและเนื้อเรื่องเช่น ควรเป็นเรื่องที่อยู่ในขอบข่ายที่ผู้สอบจะพึงเข้าใจได้ เรื่องควรจะชัดและแจ่มแจ้งพอที่จะเข้าใจได้โดยไม่ต้องอาศัยความรู้ภายนอกมาประกอบ ขณะเดียวกันก็ไม่ใช่เรื่องที่รู้กันแพร่หลายมากถึงขั้นที่ว่าไม่ต้องอ่านก็ตอบคำถามได้ ตัวแปรที่ควรคำนึงถึงอีกอย่างก็คือ ลีลา และการดำเนินเรื่อง เรื่องที่พูดถึงประเด็นใดประเด็นหนึ่งชัดแจ้งและตรงไปตรงมาจนเกินไปอาจจะไม่เหมาะสมในแง่ที่จะเป็นข้อสอบ และทำให้คิดตั้งคำถามได้ยากด้วย ลีลาที่เหมาะสมน่าจะเป็นเรื่องหรือเหตุการณ์ที่มีลำดับขั้นตอนของเวลา หรือเป็นเรื่องที่เปรียบเทียบหรือเปรียบต่างสิ่งของ คนหรือเหตุการณ์ หรืออาจจะเป็นเรื่องที่แสดงทัศนคติและความคิดเห็นของผู้เขียนเรื่องเอง เกณฑ์ประการสุดท้ายก็คือให้ดูที่ภาษา เนื้อเรื่องไม่ควรจะเต็มไปด้วยศัพท์ที่ยากจนเกินไป หรือโครงสร้างประโยคที่ซับซ้อนจนเกินไป ฉะนั้นก่อนนำมาใช้อาจจำเป็นต้องดัดแปลงภาษาก่อน แต่ก็ไม่หมายความว่า จะทำให้ง่ายจนกระทั่งข้อสอบไม่สามารถแยกได้ระหว่างคนเก่งและคนไม่เก่ง

สิ่งที่แฮริสเน้นมากก็คือในการเขียน item ศัพท์และโครงสร้างภาษาของโจทย์ไม่ควรจะซับซ้อนเกินไป เพราะหน้าที่ของผู้สอบคือตีความเนื้อเรื่องที่อ่าน มิใช่มีภาระกับการตีความคำถามอีก นอกจากนั้นตัวคำถามส่วนที่เป็น stem ควรจะเป็นส่วนที่รวมคำถามไว้ *ไม่ใช่* มีลักษณะของคำถามที่ไม่ดี ดังต่อไปนี้

John

A. obviously liked what he heard.

B. became angry upon hearing Mary's words.

.....

หากจะเป็น item ที่ดีอาจจะต้องเขียนใหม่เป็น

It was clear from John's reactions to Mary's speech that he

A. agreed completely with what she had said.

B. found her remark highly insulting.

.....

อนึ่งในการเลือกคำตอบว่าข้อไหนถูกควรจะต้องอาศัยความเข้าใจจากการอ่านและตีความเนื้อเรื่อง ไม่ใช่ผู้สอบเลือกได้ถูกเพียงเพราะผู้ออกข้อสอบใช้คำเดียวกันกับที่ปรากฏในเนื้อเรื่อง หรือ เพียงเพราะตัวลวงข้ออื่น ๆ ขัดแย้งกับข้อที่ถูกจนเห็นได้ชัดเจนว่าผิด หรือเป็นเพราะอาศัยความรู้ข้างนอกมาตอบไม่ต้องอ่านเรื่องเลยเป็นต้น

เมื่อกล่าวถึงการวัดผลการเขียนนั้น แฮริสได้สรุปว่าองค์ประกอบสำคัญที่จะต้องใส่ใจในการตรวจการเขียนก็คือ จะต้องพิจารณาที่เนื้อหา การลำดับความคิด ไวยากรณ์ ลีลา การเลือกใช้คำ และองค์ประกอบอื่น ๆ เช่น การใช้เครื่องหมายวรรคตอน ลักษณะการวัดผลจะแบ่งออกเป็น 2 แนวคือ เป็นแบบให้เขียนเรียงความ และการทดสอบแบบปรนัย แฮริสเองไม่ได้ตัดสินลงไปว่าแบบใดดีที่สุด แต่ได้เปรียบเทียบข้อดีข้อเสียของทั้ง 2 แบบให้เราดูโดยให้ข้อคิดเห็นว่าการวัดโดยให้ผู้สอบลงมือเขียนจริง ๆ นั้น มีข้อดีตรงที่ผู้สอบต้องคิดและเขียนจริง ๆ อันจะเป็นแรงจูงใจให้ผู้สอบปรับปรุงการเขียนของตนเอง และที่สำคัญก็คือ การออกข้อสอบทำได้ง่ายกว่า แต่ในขณะเดียวกันก็มีข้อโจมตีได้ว่า การวัดเช่นนี้จะวัดได้ไม่เที่ยงตรงเพราะผู้สอบอาจจะถนัดเขียนเกี่ยวกับหัวข้อหนึ่งมากกว่าอีกหัวข้อหนึ่งก็เป็นได้ และในการเขียนข้อสอบอาจจะเลี่ยงจุดอ่อนของตัวเองได้ และแม้จะออกข้อสอบง่ายแต่การตรวจก็กินเวลานาน ฉะนั้นทางออกที่ดีน่าจะเป็นวัดรวมทั้ง 2 แบบ แต่หากจะคำนึงในแง่ความสะดวกในการตรวจแล้วการทดสอบแบบปรนัย (objective test) ก็น่าจะเป็นทางออกที่ดีกว่า

รูปแบบต่าง ๆ ของการทดสอบแบบปรนัยเท่าที่เสนอไว้ในบทจะมีแบบ error recognition คือผู้สอบจะต้องหาดูว่า part ใดในประโยคที่ผิด แต่วิธีนี้ได้รับการวิจารณ์ว่า ถึงแม้จะหาพบว่าผิดตรงไหน แต่ข้อสอบประเภทนี้ก็ไม่ได้พิสูจน์ความสามารถในการแก้ส่วนที่ผิดนั้น อีกประการหนึ่งผู้สอบอาจจะตอบถูกด้วยความบังเอิญก็เป็นได้ ฉะนั้นหากจะให้สมบูรณ์ข้อสอบประเภทนี้อาจจะต้องมีควบคู่ไปกับข้อสอบประเภท Sentence completion ซึ่งผู้สอบจะต้องเลือกจากตัวเลือกที่ให้มาเติมให้ประโยคสมบูรณ์ในแง่ของไวยากรณ์นัยความหมายและน้ำเสียง ซึ่งจะเป็นการตรวจสอบความรอบรู้ในด้านลีลาการเขียนได้ หรือไม่ก็ให้ควบคู่ไปกับข้อสอบประเภท Sentence correction โดยในข้อสอบลักษณะนี้ผู้สอบจะต้องเลือก (จากตัวเลือกที่ให้) ส่วนที่เขียนเรียบเรียงใหม่แล้วที่มีความหมายตรงกับส่วนที่ขีดเส้นใต้ไว้ในคำถามเป็นต้น สำหรับการวัดการรวบรวมลำดับความคิดนั้น อาจจะออกมาในรูปการเรียงลำดับก่อน/หลังของประโยค เรื่องเครื่องหมายวรรคตอนก็อาจจะวัดได้ในลักษณะทดสอบแบบปรนัยเช่นกันโดยถามว่าตรงไหน (ที่โจทย์ระบุไว้) ควรใช้เครื่องหมายอะไร เป็นต้น ในตอนท้ายผู้เขียนยังได้เสนอแนะวิธีเตรียมข้อสอบและการตรวจให้คะแนนไว้ด้วยเพื่อที่การวัดผลจะได้เป็นไปอย่างเที่ยงธรรมที่สุด

ทักษะที่แฮริสยอมรับว่าวัดได้ยากที่สุดก็คือ Oral production เช่นเดียวกับการเขียนซึ่งต้องประกอบด้วยความสามารถหลาย ๆ ทาง ในการวัดทักษะการอ่านออกเสียงก็ต้องพิจารณาองค์ประกอบเหล่านี้ไปด้วย กล่าวคือ การออกเสียงไวยากรณ์ ศัพท์ ความคล่อง และความเข้าใจ วิธีหนึ่งที่ใช้วัด Oral production ก็คือ การสัมภาษณ์ ซึ่งผู้สอบต้องมีสูตรการให้คะแนนของตัวเองอยู่แล้วว่าจะพิจารณาให้คะแนนตรงส่วนไหนอย่างไร

วิธีอื่น ๆ ที่แฮริสเสนอไว้เช่นกันก็คือ การให้ผู้สอบอ่าน โดยผู้สอบกำหนดไว้ว่าจะเน้นเรื่องใด เป็นต้นว่า การออกเสียงสระหรือพยัญชนะ หรือการออกเสียงหนักเบา นอกจากนี้ยังมีวิธีอื่น ๆ ได้แก่การให้ผู้สอบเปลี่ยนแปลงรูปประโยค (เช่นจากบอกเล่าเป็นปฏิเสธ) การให้ผู้สอบแต่งประโยคโดยให้สถานการณ์มาและการให้ผู้สอบบรรยายภาพ เป็นต้น

เท่าที่กล่าวมานับเป็นการเล่าอย่างคร่าว ๆ ถึงวิธีการและข้อเสนอแนะต่าง ๆ ของแฮริสในส่วนที่เกี่ยวกับการวัดทักษะต่าง ๆ ผู้ที่สนใจรายละเอียดเรื่องใดบทใดน่าจะได้อ่านจากต้นฉบับเต็มในหนังสือเล่มนี้เองจะเป็นประโยชน์อย่างยิ่ง

เพื่อที่จะให้หนังสือสมบูรณ์ยิ่งขึ้นในแง่เนื้อหา ในบทที่ 9 แฮริส จึงได้กล่าวถึงขั้นตอนของการสร้างข้อสอบ ว่าควรมีการเตรียมการอย่างไรบ้างโดยแบ่งเป็นขั้นตอนต่าง ๆ ถึง 6 ขั้นตอน ได้แก่วางแผน เตรียมคำถามและคำสั่ง ตรวจทานและแก้ไขเรียบเรียง ลอง pretest และวิเคราะห์ผลรวบรวมข้อสอบครั้งสุดท้าย และจัดพิมพ์ ในบทที่ 10 ก็พูดถึงวิธีการเตรียมการจัดสอบ

ที่น่าสนใจก็คือ แฮริส ให้ความสำคัญกับการตีความและการใช้ผลสอบมากดังจะเห็นได้ว่า ในบทที่ 11 เราจะได้ทราบถึงความสำคัญของการตีความคะแนนและการวัดแนวโน้มเข้าสู่ส่วนกลาง (measures of central tendency) ความสำคัญและการหาค่าความบ่ายเบนมาตรฐาน (Standard Deviation) ตำแหน่งเปอร์เซ็นไทล์ (Percentile Ranks) และคะแนนมาตรฐาน (Standard scores) เป็นต้น ซึ่งผู้เขียนก็ได้ให้กำลังใจมากว่าไม่ใช่เรื่องเหลือบ่ากว่าแรง แม้ผู้ที่เป็นไม้เบื่อไม้เมากับคณิตศาสตร์มากที่สุดก็ยังอาจเข้าใจและทำได้ และเพื่อให้เรื่องง่ายเข้าในบทที่ 12 จึงได้เสนอสถิติพื้นฐานที่เกี่ยวข้องกับการวัดผล เช่นการหาค่าความถี่ การหาคะแนนเฉลี่ย และอื่น ๆ

โดยสรุปแล้วหนังสือเล่มนี้จัดเป็นหนังสือที่ให้ข้อมูลและความรู้ที่เกี่ยวกับการวัดผลภาษาอังกฤษที่สมบูรณ์มากที่สุดเล่มหนึ่ง ผู้อ่านมีโอกาสที่จะได้ทำความคุ้นเคยกับรูปแบบต่าง ๆ ของข้อสอบ พร้อมทั้งรู้ข้อดีข้อเสีย ข้อจำกัดของแต่ละรูปแบบนั้น อย่างน้อยที่สุดก็ได้ความคิดว่าควรจะวัดที่จุดใดบ้างในทักษะและองค์ประกอบส่วนนั้น ๆ และประโยชน์ในชั้นสูงสุดก็คือสามารถสร้างข้อสอบที่ได้มาตรฐานเพื่อประเมินความสามารถของผู้สอบในทักษะส่วนนั้นได้ ผู้ที่สอนหนังสือคงจะเลี่ยงหน้าที่ออกข้อสอบนี้ไม่ได้ และผู้ที่จะสามารถออกข้อสอบให้ได้ดีก็คงจะพลาดหนังสือเล่มนี้ไม่ได้อีกเช่นกัน

เสาวภา จิงพัฒนาพงษ์