หอสมุดกลาง
จุฬาลงกรณมหาวิทยาลัย

บทที่ 3

การตรวจสอบความเที่ยงตรงของลำดับขั้น

วิธีการจัดเนื้อหาวิชาโดยการแบ่งบทเรียนออกเป็นหน่วยทักษะย่อย ๆ หลาย ๆ หน่วยแล้วนำมาเรียงลำดับ เพื่อให้การเรียนการสอนมีเนื้อหาวิชาที่ต่อเนื่องกัน โดยอาศัยความรู้จากหน่วยทักษะที่เป็นพื้นฐานแล้วส่งถ่ายความรู้นั้นขึ้นมาตามลำดับจนกว่าจะบรรลุถึงจุดมุ่งหมายอันสุดท้ายของบทเรียนที่ตั้งไว้ วิธีการนี้ กาเยได้เสนอไว้ตั้งแต่ปี ค.ศ. 1961 ได้เป็นที่สนใจและนำไปใช้อย่างกว้างขวาง สำหรับในประเทศไทย ก็ได้เริ่มมีการนำเอาวิธีนี้ไปใช้ เช่น การปรับปรุงหลักสูตรและวิธีสอนวิชาคณิตศาสตร์ของสถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี่ เป็นต้น

การเรียงลำดับหน่วยทักษะนั้น ในขั้นต้นการวางเค้าโครงเชื่อมโยงระหว่างหน่วยอาศัยความชำนาญและการตัดสินใจจากคณะครูผู้สอนและผู้เชี่ยวชาญในสาขาวิชานั้น ซึ่งก็ยังไม่เป็นที่น่าเชื่อถือ ควรจะมีวิธีการอื่น ๆ มายืนยันความเที่ยงตรงของลำดับขั้นนั้น ๆ อีก ได้มีผู้เสนอวิธีการทดสอบไว้หลายวิธี เช่น วิธีการของกาเยและพาราไดซ์ (Gagné - Paradise) วอลเบสเซอร์และไอเซนเบริก (Walbessor - Eisenberge, 1972) ไวท์และคลาร์ก (White - Clark, 1973) สำหรับวิทยานิพนธ์นี้จะเสนอ 2 วิธีหลังไว้โดยละเอียด

วิธีการ 3 วิธีที่กล่าวถึงนี้มีลักษณะการใช้คล้าย ๆ กัน กล่าวคือ จะทำการทดสอบลำดับขั้นไปทีละคู่ ถ้าเห็นว่าทักษะ ก. ควรจะเป็นพื้นฐานของทักษะ ข. คือต้องเรียนรู้ทักษะ ก. ก่อน จึงจะสามารถเรียนรู้ทักษะ ข. ได้ ผู้ทดสอบก็จะจัดให้มีการเรียนการสอนทักษะ ก. ก่อน เมื่อจบก็ทำการทดสอบว่านักเรียนได้เกิดผลสัมฤทธิ์ในการเรียนรู้ทักษะ ก. หรือไม่ แล้วสอนทักษะ ข. และทำการทดสอบในลักษณะเดียวกัน เนื่องจาก

การเรียนรู้เป็นขบวนการที่ต่อเนื่องกันตลอดแม้ว่าแต่ละหน่วยย่อยนั้นจะมีเป้าประสงค์เป็นอิสระจากกัน ดังนั้นหากผู้ทดสอบมีความเห็นว่า ทักษะ ก. เป็นพื้นฐานของทักษะ ข. และผู้เรียนเกิดมีทักษะ ก. ขึ้นแล้ว (การตัดสินว่าผู้เรียนได้เกิดการเรียนรู้ขึ้นแล้วหรือไม่ทราบได้จากการทดสอบเมื่อจบการเรียนแต่ละทักษะ) ก็ควรที่จะสามารถเรียนรู้ทักษะ ข. ได้ไม่ยาก และหากผู้เรียนสามารถผ่านการทดสอบทักษะ ข. ได้อีก ก็แสดงว่าทักษะ ก. เป็นพื้นฐานของทักษะ ข. จริง หรือหากผู้เรียนได้ผ่านทักษะ ก. ไปแล้ว แต่ไม่อาจเรียนรู้ทักษะ ข. ได้ ทักษะ ก. ก็มิได้เป็นพื้นฐานของทักษะ ข. ควรจะได้พิจารณาลำดับขั้นของทักษะในบทเรียนนั้นใหม่ เช่นเดียวกับกรณีที่นักเรียนไม่สามารถเรียนรู้ทักษะ ก. ได้ แต่สามารถเรียนรู้ทักษะ ข. แสดงว่าทักษะ ก. นั้นไม่ได้อยู่ในลำดับที่ต่อเนื่องกันเลย

<u>วิธีการของกาเยและพาราไดซ์โดยย่อ</u>

หลังจากมีการทดสอบผลการเรียนรู้แล้ว จัดแยกนักเรียนออกเป็น 4 กลุ่ม ตามผลการทดสอบที่รวบรวมได้ทั้งหมด

1. กลุ่มนักเรียนที่สอบผ่านทักษะขั้นต่ำกว่าที่เป็นพื้นฐานทั้งหมดและสอบผ่านทักษะขั้นสุดท้าย กำหนดให้ความถี่เป็น a

2. กลุ่มนักเรียนที่ไม่ผ่านทักษะขั้นต่ำทักษะใดทักษะหนึ่ง แต่สอบผ่านทักษะขั้นสูง กำหนดให้ความถี่เป็น b

3. กลุ่มนักเรียนที่สอบผ่านทักษะขั้นต่ำกว่าที่เป็นพื้นฐานทั้งหมด แต่ไม่ผ่านทักษะขั้นสุดท้าย กำหนดให้ความถี่เป็น c

4. กลุ่มนักเรียนที่สอบไม่ผ่านทั้งทักษะที่เป็นทักษะขั้นต่ำทักษะใดทักษะหนึ่งและไม่ผ่านทักษะขั้นสุดท้ายด้วย กำหนดให้ความถี่เป็น d

เมื่อได้ความถี่ของนักเรียนทั้ง 4 กลุ่มแล้ว ก็คำนวณหาค่าดัชนีที่จะชี้บอกความเที่ยงตรงของลำดับขั้นได้ ค่าดัชนีที่ได้คือ

$$\text{proportion positive transfer} = \frac{a + d}{a + b + d}$$

ค่าที่ได้ซึ่งจะชี้ว่าลำดับขั้นการเรียนรู้ที่สร้างขึ้นนั้นมีความเที่ยงตรงหรือไม่ คือ ค่าที่ใกล้เคียงกับ 1.00 โดยมีเหตุผลว่า ลำดับขั้นที่เที่ยงตรงจริงจำนวนคนในกลุ่ม b จะต้องเท่ากับศูนย์ คือไม่มีนักเรียนที่สอบไม่ผ่านทักษะขั้นต่ำแต่ผ่านการทดสอบของทักษะขั้นสูง

วิธีการของกาเยและพาราไดซ์นี้ มีจุดด้อยอยู่ที่ใช้ค่าดัชนีเพียงตัวเดียวเป็นเครื่องตัดสินความเที่ยงตรง และค่าดัชนีที่ได้โดยวิธีนี้ก็อาจมีค่าใกล้เคียง 1.00 ได้ แม้ว่าทักษะขั้นต่ำและขั้นสูงจะไม่เกี่ยวข้องต่อกันเลย ค่าดัชนีนี้น่าจะบอกได้เพียงลักษณะความสัมพันธ์เท่านั้น ซึ่งมิใช่เงื่อนไขเพียงประการเดียวที่ใช้บอกความสัมพันธ์ด้านลำดับขั้นต่อกัน นอกจากนี้ค่าดัชนีนี้ยังขาด sampling distribution ทำให้การสรุปอ้างไปยังมวลประชากรทำได้ยาก และในกรณีที่มีข้อทดสอบของแต่ละทักษะมากกว่า 1 ข้อ ก็จะเกิดปัญหาในเรื่องการตัดสิน "ผ่าน" หรือ "ไม่ผ่าน" อีกด้วย (White, 1974 b)

<u>วิธีทดสอบของวอลเบสเซอร์และไอเซนเบิร์ก</u> (Walbessor & Eisenberg , 1972)

เป็นวิธีที่ทำการทดสอบทักษะไปทีละคู่ โดยการคำนวณหาค่าดัชนีที่สำคัญ 3 ตัว โดยใช้ความถี่ของนักเรียนกลุ่มต่าง ๆ คล้ายวิธีของกาเยและพาราไดซ์ กลุ่มนักเรียนต่าง ๆ มีความหมายดังนี้ (จากการทดสอบโดยใช้ข้อทดสอบ 2 ข้อ ต่อทักษะ)

ตารางที่ 3 การจัดกลุ่มนักเรียนตามจำนวนข้อทดสอบที่ทำได้จากแต่ละคู่ทักษะ

ทักษะขั้นต่ำกว่า (ทักษะ I) \ ทักษะขั้นสูงกว่า (ทักษะ II)	จำนวนข้อที่ทำได้ 0	1	2
0	f_{00}	f_{01}	f_{02}
1	f_{10}	f_{11}	f_{12}
2	f_{20}	f_{21}	f_{22}

จากคู่ทักษะที่อยู่ติดกัน (หรือที่เชื่อว่าจะเป็นพื้นฐานของทักษะใด) จะมีทักษะหนึ่งที่ผู้ทดสอบกำหนดเป็นทักษะขั้นต่ำกว่า และอีกทักษะหนึ่งก็จะเป็นทักษะขั้นสูงกว่า ให้เป็นทักษะ I และทักษะ II ตามลำดับ

ให้ f เป็นความถี่ของนักเรียนในกลุ่มต่าง ๆ ดังนั้น

f_{20} คือความถี่ของนักเรียนที่ทำข้อทดสอบทักษะขั้นต่ำได้ 2 ข้อ แต่ทำข้อทดสอบของทักษะขั้นสูงกว่าไม่ได้เลย

หรือ f_{11} คือความถี่ของนักเรียนที่ทำข้อทดสอบของทักษะขั้นต่ำได้ 1 ข้อ และทำข้อทดสอบของทักษะขั้นสูงกว่าได้ 1 ข้อ

และอื่น ๆ ในทำนองเดียวกัน

ค่าของดัชนีที่ใช้เป็นเกรื่องวัดมี 3 ตัว คือ

1. ความสอดคล้องกัน (Consistency Ratio, CSR) เป็นอัตราส่วนที่ใช้ทดสอบสมมติฐานที่ว่า ถ้านักเรียนเรียนรู้ทักษะขั้นสูงได้เมื่อสิ้นสุดการสอน ย่อมแสดงว่านักเรียนได้เกิดการเรียนรู้ทักษะขั้นต่ำด้วยแล้ว เป็นอัตราส่วนระหว่างจำนวนนักเรียนที่ทำข้อทดสอบของทักษะขั้นสูง (ทักษะ II) และทักษะขั้นต่ำ (ทักษะ I) ผ่านทั้ง 2 ทักษะ (กลุ่ม f_{22}) ต่อจำนวนนักเรียนที่สอบผ่านทักษะขั้นสูงและผ่านหรือไม่ผ่านทักษะขั้นต่ำ (กลุ่ม f_{22}, f_{12}, f_{02}) สูตรที่ใช้คือ

$$CSR = \frac{f_{22}}{f_{22} + f_{12} + f_{02}}$$

2. ค่าความเหมาะสม (Adequacy Ratio, ADR) ค่าดัชนีนี้ใช้ทดสอบสมมติฐานที่ว่า ถ้านักเรียนคนใดได้เรียนรู้ทักษะขั้นต่ำแล้ว ย่อมสามารถเรียนรู้ทักษะขั้นสูงได้ เมื่อสิ้นสุดการสอนแล้ว ได้จากค่าอัตราส่วนระหว่างจำนวนนักเรียนที่ทำข้อทดสอบของทักษะทั้งสองได้ถูกหมด กับจำนวนนักเรียนที่ทำข้อทดสอบของทักษะที่ I ได้ถูกหมด สูตรที่ใช้คือ

$$ADR = \frac{f_{22}}{f_{22} + f_{21} + f_{20}}$$

3. ค่าความสมบูรณ์ (Completeness Ratio, CPR) ค่าดัชนีนี้ใช้สำหรับสนับสนุนความเที่ยงตรงของลำดับขั้นการเรียนรู้ให้หนักแน่นขึ้น เพราะค่าดัชนี 2 ค่าแรกอาจมีค่าสูงโดยที่มีนักเรียนจำนวนมากไม่สามารถเรียนรู้ได้ทั้งทักษะ I และทักษะ II ซึ่งจะทำให้ไม่อาจสรุปได้ว่าลำดับขั้นนั้นมีความเที่ยงตรงโดยสมบูรณ์แล้ว สูตรที่ใช้คือ

$$CPR = \frac{f_{22}}{f_{22} + f_{00} + f_{01} + f_{10} + f_{11}}$$

ค่าดัชนีทุกค่ามีพิสัย (Range) ตั้งแต่ 0 ถึง 1 ลำดับขั้นการเรียนรู้คู่ใด ๆ ที่จะยอมรับได้ว่ามีความเที่ยงตรงจะต้องมีค่าดัชนีทั้งสามตั้งแต่ 0.85 ขึ้นไป

ในกรณีที่บทเรียนเกี่ยวกัน แต่ได้ตั้งสมมติฐานว่ามี 2 วิธีที่จะสอนให้บรรลุเป้าประสงค์ได้ และผลการตรวจสอบก็พบว่า ทั้งสองวิธีนั้นมีความเที่ยงตรงเป็นที่ยอมรับอย่างเดียวกัน การจะตัดสินว่าวิธีใดเหมาะสมกว่าให้พิจารณาเงื่อนไขของเวลาว่า วิธีใดใช้เวลาในการเรียนการสอนน้อยที่สุด ก็ให้เลือกใช้วิธีนั้น

<u>จุดเด่น</u>

วิธีการนี้ไม่ยึดค่าดัชนีค่าใดค่าหนึ่งเป็นเกณฑ์ตัดสินการยอมรับความเที่ยงตรงของลำดับขั้น แต่ต้องทั้งสามดัชนี ดังนั้นทักษะคู่ใดถ้าได้ตัดสินความเที่ยงตรงของลำดับแล้ว ก็มีเหตุผลเพียงพอที่จะยอมรับได้ นอกจากนี้ยังสามารถแปลความหมายได้จากค่าของดัชนี เช่น ถ้า CSR มากกว่า 0.85 แต่ ADR น้อยกว่า 0.85 อาจแปลความหมายได้ว่า ทักษะขั้นต่ำนั้นเป็นพื้นฐานของทักษะขั้นสูงในแผนภูมิลำดับการสอนจริง แต่ไม่ได้เป็นลำดับขั้นที่ต่อเนื่องกัน น่าจะมีทักษะอื่นมาคั่นระหว่างทักษะทั้งสองก่อน

และในกรณีที่ค่าดัชนีอื่น ๆ มากกว่า 0.85 แต่ CPR น้อยกว่า 0.85 แสดงว่ามีนักเรียนจำนวนมากที่ไม่สามารถเรียนรู้ได้ทั้งสองทักษะ จึงยังไม่อาจสรุปได้ว่า ทักษะทั้งสองนั้นเป็นพื้นฐานกัน ในกรณีเช่นนี้ลำดับขั้นการเรียนรู้ระหว่างทักษะทั้งสองอาจเป็นจริงเฉพาะนักเรียนกลุ่มที่มีผลสัมฤทธิ์ทางการเรียนค่อนข้างสูงเท่านั้น

จุดด้อย

ไม่มี Sampling distribution ในเกณฑ์การตัดสินผลของข้อทดสอบว่าผ่าน หรือไม่ผ่าน เพราะถ้าตั้งเกณฑ์ผ่านไว้สูง หรือข้อสอบยาก นักเรียนจำนวนมากสอบไม่ผ่าน ทำให้ต้องปฏิเสธสมมติฐาน ทั้ง ๆ ที่สมมติฐานนั้นเป็นที่ยอมรับของผู้สอนแล้ว

กลุ่ม f_{22} หรือ f_{33} ควรจะมีจำนวนเป็นตั้งแต่ 85% ของนักเรียนทั้งกลุ่มตัวอย่าง จึงจะมีโอกาสที่ค่าของดัชนีจะมากกว่า 0.85 ซึ่งจะยอมรับได้ว่า ทักษะ I เป็นพื้นฐานของทักษะ II จริง ถ้ากลุ่ม f_{22} หรือ f_{33} มีประมาณ 65% - 84% ของกลุ่มตัวอย่าง จะมีโอกาสที่จะต้องปฏิเสธสมมติฐานได้ เพราะมีดัชนีตั้งแต่ 1 ค่าขึ้นไปที่น้อยกว่า 0.85 และถ้ากลุ่ม f_{22} หรือ f_{33} น้อยกว่า 65% ของกลุ่มตัวอย่าง จะต้องปฏิเสธสมมติฐานเสมอ ดังนั้นความยากง่ายของข้อสอบจึงเป็นผลต่อค่าดัชนีทั้งสามด้วย การพิจารณาความยากง่ายของข้อทดสอบได้กล่าวไว้ในบทที่ 5

หอสมุดกลาง จุฬาลงกรณมหาวิทยาลัย

เมื่อมี 3 ข้อทดสอบในแต่ละทักษะ

ตารางที่ 4 การจัดกลุ่มนักเรียนตามจำนวนข้อสอบที่ทำได้ (3 ข้อทดสอบ) จากแต่ละคู่ทักษะ

ทักษะ II

ทักษะ I: จำนวนข้อที่ทำได้	0	1	2	3
0	f_{00}	f_{01}	f_{02}	f_{03}
1	f_{10}	f_{11}	f_{12}	f_{13}
2	f_{20}	f_{21}	f_{22}	f_{23}
3	f_{30}	f_{31}	f_{32}	f_{33}

ในทำนองเดียวกันกับเมื่อใช้ 2 ข้อทดสอบ จะได้ค่าดัชนีทั้งสามดังนี้

$$1.\ CSR = \frac{f_{33}}{f_{33} + f_{23} + f_{13} + f_{03}}$$

$$2.\ ADR = \frac{f_{33}}{f_{33} + f_{30} + f_{31} + f_{32}}$$

$$3.\ CPR = \frac{f_{33}}{f_{33}+f_{20}+f_{21}+f_{22}+f_{10}+f_{11}+f_{12}+f_{00}+f_{01}+f_{02}}$$

การตัดสินก็เป็นเช่นเดียวกัน คือ ทั้งสามดัชนีจะต้องมีค่ามากกว่า 0.85 จึงจะยอมรับความเที่ยงตรงของสมมติฐานของลำดับขั้นของทักษะคู่นี้

<u>วิธีการของไวท์และคลาร์ก</u> (White & Clark, 1973)

เป็นการทดสอบความเที่ยงตรงของคู่ทักษะทีละคู่เช่นเดียวกับวิธีการของวอล-เบสเซอร์และไอเซนเบิร์ก จึงจัดกลุ่มนักเรียนที่ใช้เป็นกลุ่มตัวอย่างเช่นเดียวกัน มีส่วนที่เพิ่มมาคือ การหาจำนวนรวมตามแนวตั้งและแนวนอนของตาราง เมื่อมี 2 ข้อทดสอบในแต่ละทักษะ

ตารางที่ 5 แสดงกลุ่มนักเรียนจัดตามจำนวนข้อทดสอบที่ทำได้

		ทักษะ II			
	จำนวนข้อที่ทำได้	0	1	2	$\sum_j f_{ij}$
ทักษะ I	2	f_{20}	f_{21}	f_{22}	a
	1	f_{10}	f_{11}	f_{12}	b
	0	f_{00}	f_{01}	f_{02}	c
	$\sum_i f_{ij}$	f	e	d	N

กำหนดให้ P_0 เป็นสัดส่วนของนักเรียนในมวลประชากรที่ไม่มีทักษะทั้งสอง (ไม่สามารถเรียนรู้ทักษะทั้งสองได้)

P_1 เป็นสัดส่วนของนักเรียนในมวลประชากรที่มีเฉพาะทักษะ 1

P_2 เป็นสัดส่วนของนักเรียนในมวลประชากรที่มีเฉพาะทักษะ 2

P_B เป็นสัดส่วนของนักเรียนในมวลประชากรที่มีทั้งสองทักษะ

สมมติฐานที่จะทำการทดสอบ คือ Ho : $P_2 = 0$

และ Ha : $P_2 > 0$

นั่นคือ ตั้งสมมติฐานไว้ว่า จะไม่มีนักเรียนที่มีเฉพาะทักษะ 2 โดยไม่ได้มีทักษะ 1 มาก่อน ถ้ายอมรับ Ho ปฏิเสธ Ha จะเป็นผลให้ยอมรับว่า ไม่มีนักเรียนที่จะผ่านทักษะ 2 ถ้าไม่ได้ผ่านทักษะ 1 มาก่อน (โดยที่ทักษะ 1 เป็นพื้นฐานของทักษะ 2) และถ้าปฏิเสธ Ho ยอมรับ Ha ก็จะเป็นผลให้ยอมรับว่ามีนักเรียนที่ผ่านทักษะ 2 ได้โดยไม่ต้องมีทักษะ 1 มาก่อน หรือแสดงว่าทักษะ 1 ไม่ใช่พื้นฐานของทักษะ 2

กำหนดให้ θ_a = ความน่าจะเป็นที่ผู้มีทักษะ 1 จะตอบคำถามข้อใดข้อหนึ่งที่ทดสอบทักษะ 1 ได้ถูก

θ_b = ความน่าจะเป็นที่ผู้ไม่มีทักษะ 1 จะตอบคำถามข้อใดข้อหนึ่งที่ทดสอบทักษะ 1 ได้ถูก

θ_c = ความน่าจะเป็นที่ผู้มีทักษะ 2 จะตอบคำถามข้อใดข้อหนึ่งที่ทดสอบทักษะ 2 ได้ถูก

θ_d = ความน่าจะเป็นที่ผู้ไม่มีทักษะ 2 จะตอบคำถามข้อใดข้อหนึ่งที่ทดสอบทักษะ 2 ได้ถูก

p_{02} = ความน่าจะเป็นที่จะมีนักเรียนที่ไม่มีทักษะ 1 แต่มีทักษะ 2 (กลุ่ม f_{02})

ค่าของ p_{02} เป็นค่าที่ estimate มาจากผลบวกของความน่าจะเป็น 4 กลุ่มสมมติฐาน คูณกับ conditional probability ของกลุ่มสมาชิกที่สามารถตอบคำถามได้ถูกต้อง

$$p_{02} = p_0 (1 - \theta_b)^2 \theta_d^2 + p_1 (1 - \theta_a)^2 \theta_d^2 + p_2 (1 - \theta_b)^2 \theta_c^2 + p_B (1 - \theta_a)^2 \theta_c^2 \qquad \text{---------- (1)}$$

ถ้า Q เป็นสัดส่วนของกลุ่มที่มีทักษะ 1 ต่อมวลประชากร

R เป็นสัดส่วนของกลุ่มที่มีทักษะ 2 ต่อมวลประชากร

จะได้ $Q = p_1 + p_B$

$R = p_2 + p_B$

และจากการใช้ expected marginal total พิจารณาแต่ละทักษะได้ดังนี้

ทักษะ 1

จำนวนนักเรียนที่ทำ 2 ข้อถูก $a = \{Q\,\theta_a^2 + (1-Q)^2\,\theta_b^2\}\,N$ (1.1)

จำนวนนักเรียนที่ทำ 1 ข้อถูก $b = \{2Q\,\theta_a(1-\theta_a) + 2(1-Q)\,\theta_b(1-\theta_b)\}\,N$ (1.2)

จำนวนนักเรียนที่ทำ 0 ข้อถูก $c = \{Q\,(1-\theta_a)^2 + (1-Q)(1-\theta_b)^2\}\,N$ (1.3)

ทักษะ 2

จำนวนนักเรียนที่ทำ 2 ข้อถูก $d = \{R\,\theta_c^2 + (1-R)\,\theta_d^2\}\,N$ (1.4)

จำนวนนักเรียนที่ทำ 1 ข้อถูก $e = \{2R\,\theta_c\,(1-\theta_c) + 2(1-R)\,\theta_d(1-\theta_d)\}\,N$ (1.5)

จำนวนนักเรียนที่ทำ 0 ข้อถูก $f = \{R\,(1-\theta_c)^2 + (1-R)(1-\theta_d)^2\}\,N$ (1.6)

เพื่อที่จะให้ค่า p_{02} มีค่าสูงที่สุด จึงกำหนดให้ $\theta_b = 0$, $\theta_c = 1$ หาค่า estimated ของ parameters ต่าง ๆ ดังนี้

จาก (1.1) $a = Q\,\theta_a^2\,N$

จาก (1.4) $b = 2Q\,\theta_a\,(1-\theta_a)\,N$

ดังนั้น $\frac{a}{b} = \frac{\theta_a}{2\,(1-\theta_a)}$

จะได้ $\hat{\theta}_a = \frac{2a}{2a + b}$ ——————— (2)

ในทำนองเดียวกัน จาก (1.5) และ (1.6)

และ $\hat{\theta}_d = \frac{e}{e + 2f}$

$\hat{Q} = \frac{(2a + b)^2}{4aN}$

$\hat{R} = 1 - \frac{(e + 2f)}{4fN}$

เมื่อ $\hat{Q} = \hat{p}_1 + \hat{p}_B$

$\hat{R} = \hat{p}_2 + \hat{p}_B$

เมื่อต้องการทดสอบ Ho จึงกำหนดให้ค่าของ p_2 เป็น 0

ดังนั้น จะได้ $p_1 = Q - R$

และ $p_0 = 1 - p_1 - R$

ค่าของ p_{02} จากสมการ (1) จึงหาค่าได้จาก

$$p_{02} = p_0 \theta_d^2 + p_1 (1 - \theta_a)^2 \theta_d^2 + p_B(1 - \theta_a)^2 \quad \text{——(3)}$$

$$q_{02} = 1 - p_{02}$$

จากค่าของ p_{02}, q_{02} นำมาหาค่า Binomial expansion ของ f_{02} ซึ่งมี parameters เป็น (N, p_{02}) ถ้าค่าสะสมของความน่าจะเป็น (Cumulative probability of f_{02}) ของ f_{02} ที่จะเป็น 1, 2, 3, . . . เกินกว่าค่าวิกฤติ (Critical value) ที่ตั้งไว้ ให้สรุปผลการยอมรับได้จากการเปรียบเทียบค่า f_{02} จากการคำนวณนี้ กับค่าที่เป็นจริงในตาราง ถ้า f_{02} ในตารางมากกว่า f_{02} จากการคำนวณ Ho จะถูกปฏิเสธ

การกระจายจะมีลักษณะดังนี้

$$p_{r0}(f_{02} = 0) = \binom{N}{0} p^0 q^{(N-0)} = q^N$$

$$p_{r1}(f_{02} = 1) = \binom{N}{1} p^1 q^{(N-1)} = N_{pq}^{(N-1)}$$

$$p_{r2}(f_{02} = 2) = \binom{N}{2} p^2 q^{(N-2)} = \frac{N(N-1)}{2} p^2 q^{(N-2)}$$

$$p_{rn}(f_{02} = n) = \binom{N}{n} p^n q^{(N-n)}$$

$$= \frac{N(N-1)(N-2) \cdots (N-n)}{n\ i} p^n q^{(N-n)}$$

$$p_{02n} = p_{r1} + p_{r2} + p_{r3} + \cdots p_{rn}$$

ค่าวิกฤติที่ใช้ในที่นี้คือ 0.95 ดังนั้นจึงทำการกระจายค่า p_{02} ไปเรื่อย ๆ จนกว่าผลรวมสะสมจะมากกว่าหรือเท่ากับ 0.95 ตรวจสอบ โดยถ้า f_{02} เป็น n เป็นค่าสุดท้ายที่ทำให้ค่าสะสมของความน่าจะเป็น p_{02} มากกว่า หรือเท่ากับ 0.95 นำค่า n ไปเปรียบเทียบกับจำนวน f_{02} ที่ปรากฏอยู่จริงในตาราง

ตัวอย่างของการตัดสินสมมติฐานจากการทดสอบทักษะคู่หนึ่งโดยไวท์ (White, 1971) เมื่อทำการสอนและทดสอบทักษะทั้งสองนั้นเรียบร้อยแล้วรวบรวมกลุ่มข้อมูลได้ดังนี้

a = 53

b = 13

c = 101

d = 46

e = 5

f = 116

คำนวณค่า p_{02} โดยการแทนค่า a, b, c, . . ., f ลงในสมการที่ (1)-(3) ที่กล่าวมาแล้ว หาค่า p_{02} จากสมการ (3) ได้ เป็น 0.00355 และหาค่า Binomial expansion ของแต่ละจำนวนของ f_{02} ได้ ดังนี้

$$p_r = (f_{02} = 0) = 0.552$$

$$p_r = (f_{02} = 1) = 0.329$$

$$p_r = (f_{02} = 2) = 0.097$$

จะเห็นได้ว่า ความน่าจะเป็นสะสมของการมี f_{02} เป็น 1, 2 เป็น 0.978 ถ้ากำหนดค่าวิกฤติไว้เป็น 0.95 ค่าความน่าจะเป็นสะสมจะเกินค่าวิกฤติ เมื่อ f_{02} เป็น 2 ดังนั้นถ้าจำนวนข้อมูลในกลุ่ม f_{02} จากตารางมากกว่า 2 ข้อมูล จะต้องปฏิเสธ Ho เพราะแสดงว่ามีจำนวนนักเรียนที่สามารถมีทักษะ 2 ได้โดยไม่ต้องมีทักษะ 1 มาก่อนเป็นจำนวนมากกว่าที่ควรเป็น จากการทดลองของไวท์นั้นผลปรากฏว่าค่า f_{02} ที่เกิดขึ้นจริงในตารางมีน้อยกว่า 2 ข้อมูล จึงสามารถยอมรับ Ho ได้ หรือยอมรับว่าทักษะ 1 เป็นพื้นฐานที่ควรจะเรียนรู้ก่อนที่จะเรียนทักษะ 2

เมื่อมี 3 ข้อทดสอบ จึงสรุปได้ดังนี้ (White & Clark, 1973)

ตารางที่ 6 การจัดกลุ่มนักเรียนตามจำนวนข้อทดสอบที่ทำได้ในแต่ละกลุ่ม

จำนวนคำตอบถูกของทักษะ 1 \ จำนวนคำตอบถูกของทักษะ 2	0	1	2	3	รวม
0	f_{00}	f_{01}	f_{02}	f_{03}	d
1	f_{10}	f_{11}	f_{12}	f_{13}	c
2	f_{20}	f_{21}	f_{22}	f_{23}	b
3	f_{30}	f_{31}	f_{32}	f_{33}	a
รวม	h	g	f	e	N

$$\theta_a = \frac{3a}{3a - b}$$

$$\theta_b = \frac{c}{c + 3d}$$

$$Q = \frac{\frac{\bar{X}}{3} - \theta_b}{\theta_a - \theta_b} \qquad \left(\bar{X} = \frac{(3a + 2b - c)}{N}\right.$$

$$R = \frac{\frac{\bar{Y}}{3} - \theta_d}{\theta_c - \theta_d} \qquad \left(\bar{Y} = \frac{(3e + 2f - g)}{N}\right.$$

$$\theta_c = \frac{3c}{3c + f} \quad , \quad \theta_d = \frac{g}{g + 3h}$$

ดังนั้น $p_1 = Q - R$

$p_0 = 1 - p_1 - R$

และเมื่อ $Q = p_1 + p_B$

$R = p_2 + p_B$

จาก $p_{03} = p_0 (1 - \theta_b)^3 \theta_d^3 + p_1 (1 - \theta_a)^3 \theta_d^3 + p_2 (1 - \theta_b)^3 \theta_c^3 + p_B (1 - \theta_a)^3 \theta_c^3$ ต้องการให้ค่า p_{03} เป็นค่ามากที่สุด จะได้ p_2 เป็น 0

ดังนั้น $p_B = R$ และ

$$p_{03} = p_0 (1 - \theta_b)^3 \theta_d^3 + p_1 (1 - \theta_a)^3 \theta_d^3 + p_B (1 - \theta_a)^3 \theta_c^3$$

และ $q_{03} = 1 - p_{03}$

หา Binomial expansion จะได้เป็น

$$p_r (f_{03} = 0) = q_{03}^N$$

$$p_r (f_{03} = 1) = NpQ^{N-1}$$

$$p_r\ (f_{03} = 2) = \frac{N(N-1)}{2}\ p^2 q^{N-2}$$

$$p_r\ (f_{03} = n) = \binom{N}{n}\ p^n q^{N-n} \qquad (p,q = p_{03}\ ,\ q_{03})$$

หาค่าความน่าจะเป็นสะสม (Cumulative probability) ของการมี f_{03} จำนวนต่าง ๆ จนกว่าจะมากกว่าหรือเท่ากับค่าวิกฤติที่ตั้งขึ้น แล้วเปรียบเทียบ n ที่ได้กับจำนวน f_{03} ที่มีจริงในตาราง หาก n มากกว่า f_{03} ปฏิเสธ Ho ถ้าน้อยกว่าก็สามารถยอมรับ Ho ได้

จุดเด่นของวิธีการของไวท์และกลาร์กคือ ไม่ต้องคำนึงถึงความยากง่าย หรือเกณฑ์การตัดสิน "ได้" หรือ "ตก" เหมือนกับวิธีแรก ๆ การยอมรับหรือปฏิเสธการเป็นลำดับขั้นของทักษะแต่ละคู่ขึ้นอยู่กับการแจกแจงของผลการสอบโดยตรง ประกอบกับมี Sampling distribution อยู่ด้วยซึ่งจะช่วยในการคำนวณหาค่าความน่าจะเป็นของความคลาดเคลื่อน

ในวิธีการนี้นอกจากการพิจารณาค่า f_{02} ว่ามีจำนวนคนเกินกว่าค่าวิกฤติ ซึ่งจะทำให้ต้องปฏิเสธสมมติฐานหรือไม่แล้ว ยังอาจพิจารณาค่าของ f_{20}, f_{00} ได้อีก คือแม้จะยอมรับสมมติฐานแล้ว แต่หากค่าของ f_{20} มีจำนวนมาก อาจเป็นไปได้ที่ทักษะ 1 เป็นทักษะพื้นฐานของทักษะ 2 จริง แต่อยู่ในลำดับที่ห่างกันมากกว่าที่จะอยู่ติดกันเหมือนเช่นในสมมติฐาน และสำหรับ f_{00} หากมีจำนวนมากก็อาจพิจารณาได้ว่า ความเที่ยงตรงของลำดับขั้นทั้งสองนี้น่าจะเป็นจริงสำหรับประชากรของนักเรียนที่มีผลสัมฤทธิ์ทางการเรียนในชั้นปานกลางและชั้นสูงเท่านั้น ไม่อาจพิจารณาได้ในประมวลประชากรของนักเรียนที่มีผลสัมฤทธิ์ทางการเรียนในชั้นต่ำ

การเปรียบเทียบวิธีการของไวท์-คลาร์ก กับของวอลเบสเซอร์-ไอเซนเบิร์ก

จะเห็นได้ว่า วิธีการทดสอบความเที่ยงตรงของไวท์-คลาร์กและของวอลเบสเซอร์-ไอเซนเบิร์ก นั้นมีจุดเด่นจุดด้อยต่างกันออกไป ยากที่จะตัดสินว่าควรจะใช้วิธีใด และเกณฑ์ที่จะตัดสินว่าวิธีใดดีกว่าก็มีอยู่หลายเกณฑ์ด้วยกัน เช่น ความสามารถของวิธีการในการที่จะยอมรับหรือปฏิเสธสมมติฐานได้อย่างถูกต้อง สามารถให้ความหมายได้อย่างสอดคล้องกับสภาพความเป็นจริง ผลของวิธีการนั้นทำให้ยอมรับสมมติฐานเมื่อทักษะ 2 ทักษะเป็นลำดับขั้นที่ต่อเนื่องกัน โดยทักษะขั้นต่ำกว่าเป็นพื้นฐานของทักษะขั้นสูงกว่า และจะต้องปฏิเสธสมมติฐานเมื่อทักษะทั้งสองมิได้เป็นลำดับที่ต่อเนื่องกัน แม้ว่าทักษะขั้นต่ำกว่าจะมีโอกาสได้เป็นพื้นฐานของทักษะขั้นสูงได้ในลำดับต่อ ๆ ไปก็ตาม หรือเมื่อเรียงลำดับผิดเอาทักษะขั้นต่ำที่ควรเป็นพื้นฐานไปไว้ในลำดับที่สูงกว่าทักษะขั้นสูง

อีกเกณฑ์หนึ่งที่น่าจะพิจารณาคือ ความสามารถของวิธีการที่จะใช้ได้ภายใต้เงื่อนไขต่าง ๆ กัน ตัวแปรในเรื่องจำนวนข้อสอบ ความยากง่ายของข้อสอบ จำนวนของกลุ่มตัวอย่างเหล่านี้จะต้องไม่ทำให้ผลจากการทดสอบนี้เปลี่ยนแปลงไป ซึ่งการพิจารณากฎเกณฑ์การเลือกใช้วิธีการหรือผลของตัวแปรต่าง ๆ เหล่านี้เป็นเรื่องใหญ่ที่ต้องพิจารณากันโดยเฉพาะต่อไป วิทยานิพนธ์นี้มีจุดมุ่งหมายที่จะสร้างโปรแกรมเพื่อให้สามารถคำนวณค่าดัชนีต่าง ๆ ได้ถูกต้องรวดเร็วเท่านั้น และจะกล่าวถึงข้อเปรียบเทียบระหว่างสองวิธีนี้ไว้โดยย่อ ดังนี้

1. ในกรณีที่ทักษะสองทักษะมิได้เป็นลำดับต่อเนื่องกัน กล่าวคือ ทักษะขั้นสูงอยู่ห่างทักษะขั้นต่ำมากกว่าหนึ่งลำดับ วิธีการของวอลเบสเซอร์-ไอเซนเบิร์กน่าจะมีความฉับไวต่อการที่จะปฏิเสธลำดับขั้นนี้ โดยมีค่า ADR น้อยกว่า 0.85 แต่วิธีการของไวท์-คลาร์กไม่น่าจะตรวจพบได้ เพราะพฤติกรรมขั้นสูงที่ข้ามขั้นขึ้นไปย่อมเรียนรู้ได้ยากยิ่งขึ้น นักเรียนมีโอกาสที่จะตอบถูกหมดทั้งสองทักษะน้อยมาก

ในตารางที่ 7 เป็นข้อมูลสมมติทั้ง (ก) และ (ข) เพื่อแสดงกลุ่มนักเรียนโดยคำนึงว่าทักษะขั้นต่ำอยู่ห่างจากทักษะขั้นสูงมากกว่าหนึ่งลำดับ ดังนั้นโอกาสที่นักเรียน

จะทำข้อทดสอบของทักษะขั้นต่ำกว่าได้ถูกหมด แต่ทำข้อทดสอบของทักษะขั้นสูงผิดหมดย่อมเกิดมีขึ้นได้

ตารางที่ 7 กลุ่มนักเรียน (ตัวเลขสมมติ) เมื่อลำดับทักษะต่างกัน

(ก)

จำนวนข้อถูกในทักษะขั้นสูง

จำนวนข้อถูกในทักษะขั้นต่ำ	0	1	2	
2	45	15	5	65
1	25	5	0	30
0	5	0	0	5
	75	20	5	100

(ข)

จำนวนข้อถูกในทักษะขั้นสูง

จำนวนข้อถูกในทักษะขั้นต่ำ	0	1	2	
2	45	10	5	60
1	15	10	0	25
0	10	3	2	15
	70	23	7	100

ในกรณี (ก) และ (ข) วิธีการของไวท์-คลาร์ก จะยอมรับสมมุติฐานว่าทั้งสองทักษะเป็นพื้นฐานกัน โดยมีค่า p_r (f_{02} = 0, 1) = 0.98 และ p_r (f_{02} = 0, 1, 2) = .99 ตามลำดับ แต่วิธีการของวอลเบสเซอร์-ไอเซนเบิร์ก จะปฏิเสธสมมติฐาน โดยให้ค่า CSR เป็น 1.00, 0.71 ค่า ADR เป็น 0.08, 0.09 และค่า CPR เป็น 0.13, 0.10 ตามลำดับ ในกรณีเช่นนี้จะเกี่ยวกับคุณสมบัติของความสามารถให้ความหมายสอดคล้องกับสภาพความเป็นจริงได้ ชัดเจนเพียงใด

2. ในกรณีที่ทักษะทั้งสองเป็นลำดับขั้นที่ต่อเนื่องกัน แต่มีนักเรียนจำนวนมากไม่สามารถเรียนรู้ทักษะทั้งสองได้ ซึ่งถ้าทดสอบแล้วยอมรับสมมติฐานของลำดับคู่นี้ก็อาจจะไม่เหมาะสมที่จะสรุปอ้างไปยังของประชากรทั่วไปของนักเรียนในระดับนั้นได้ วิธีการของวอลเบสเซอร์-ไอเซนเบิร์กน่าจะตรวจสอบได้ แต่วิธีของไวท์-คลาร์กจะตรวจสอบไม่ได้ พิจารณาตัวเลขสมมติในตารางที่ 8

หอสมุดกลาง จุฬาลงกรณมหาวิทยาลัย

ตารางที่ 8 กลุ่มนักเรียน (ตัวเลขสมมติ) เมื่อทักษะทั้งสองขั้นยาก

(ก)

จำนวนข้อถูกในทักษะขั้นสูง

จำนวนข้อถูกในทักษะขั้นต่ำ	0	1	2	
2	3	3	4	10
1	19	10	1	30
0	57	2	1	60
	79	15	6	100

(ข)

จำนวนข้อถูกในทักษะขั้นสูง

	0	1	2	
2	12	4	12	28
1	11	4	6	21
0	47	2	2	51
	70	10	20	100

ในกรณี (ก) และ (ข) วิธีของวอลเบสเซอร์-ไอเซนเบิร์ก จะปฏิเสธสมมติฐาน เพราะได้ค่าของดัชนีทั้งสามน้อยกว่า 0.85 แต่วิธีของไวท์-กลาร์ก จะยอมรับสมมติฐาน โดยมี p_r (f_{02} = 0, 1, 2, 3, 4, 5) = 0.97 และ p_r (f_{02} = 0, 1, 2, 3, 4) = 0.97 ตามลำดับ

3. ในด้านความยากง่ายของข้อสอบ น่าจะมีผลต่อวิธีการของวอลเบสเซอร์-ไอเซนเบิร์ก ข้อสอบยิ่งเพิ่มความยากวิธีการนี้ยิ่งมีโอกาสที่จะปฏิเสธสมมติฐานได้มาก เพราะจำนวนนักเรียนที่ทำข้อสอบไม่ได้จะมีมากกว่านักเรียนที่ทำได้ แต่ไม่น่าจะมีผลกระทบต่อวิธีการของไวท์-คลาร์กในการยอมรับสมมติฐาน และ

4. สำหรับจำนวนข้อสอบที่ใช้ทดสอบในแต่ละทักษะมากน้อยจะมีผลต่อวิธีการของวอลเบสเซอร์-ไอเซนเบิร์ก ถ้าใช้จำนวนข้อสอบมาก โอกาสที่จะปฏิเสธสมมติฐานก็ยิ่งมาก เพราะจะมีจำนวนนักเรียนที่ทำข้อสอบได้ถูกหมดน้อยมากเมื่อเปรียบเทียบกับข้อสอบที่ง่ายกว่าที่มีเกณฑ์การสอบผ่านอย่างเดียวกัน

สรุป

วิธีการวัดความเที่ยงตรงของลำดับขั้นการเรียนรู้ที่เหมาะสมคือ วิธีการของวอลเบสเซอร์-ไอเซนเบิร์ก และของไวท์-คลาร์ก ซึ่งทั้งสองวิธีนี้มีจุดด้อย จุดเด่นต่าง ๆ กัน ดังนั้นเพื่อให้สามารถแปลความหมายของ ำดัชนีและการตัดสินสมมติฐานที่ถูกต้อง ควรจะใช้ค่าดัชนีของทั้งสองวิธีประกอบกัน

เงื่อนไขต่าง ๆ ที่มีผลต่อการคำนวณในทั้งสองวิธีมีหลายประการซึ่งควรจะได้ศึกษาโดยละเอียดในโอกาสต่อไป