บทที่ 2

วรรณคดีที่เกี่ยวข้อง

ในการวิจัยครั้งนี้ผู้วิจัยได้นำเสนอวรรณคดี และงานวิจัยที่เกี่ยวข้อง โดยแบ่งออกเป็น 4 ตอน ดังนี้

ตอนที่ 1 มโนทัศน์เกี่ยวกับลำดับขั้นการเรียนรู้ด้านพุทธิพิสัย

ตอนที่ 2 งานวิจัยที่เกี่ยวข้องกับลำดับขั้นการเรียนรู้ด้านพุทธิพิสัย

ตอนที่ 3 มโนทัศน์เกี่ยวกับโมเดลโลจีสติกในทฤษฎี IRT

ตอนที่ 4 งานวิจัยที่เกี่ยวข้องกับโมเดลโลจีสติก

ตอนที่ 1 มโนทัศน์เกี่ยวกับลำดับขั้นการเรียนรู้ด้านพุทธิพิสัย

โครงสร้างของลำดับขั้นการเรียนรู้

โครงสร้างของลำดับขั้นการเรียนรู้ (Structural of Learning - Hierarchy) หมายถึงโครงสร้างของความสัมพันธ์ระหว่างสมรรถภาพในการเรียนรู้ของผู้เรียน โดยที่สมรรถภาพในการเรียนรู้ซึ่งเกิดขึ้นก่อนจะอยู่ในลำดับขั้นแรก และเป็นปัจจัยพื้นฐานของสมรรถภาพในการเรียนรู้ที่ผู้เรียนสามารถเรียนรู้ได้ในภายหลัง (Rowntree 1981 : 154 อ้างใน วรรณวิภา จัตุชัย 2530 : 3) ซึ่งในการพัฒนาหลักสูตร การจัดกิจกรรมการเรียนการสอน รวมทั้งการวัดและประเมินผลการศึกษานั้น รูปแบบโครงสร้างของลำดับขั้นการเรียนรู้มีบทบาทอย่างมาก

ในเรื่องของลำดับขั้นการเรียนรู้นี้ มีนักจิตวิทยาหลายคนที่สนใจและศึกษาเกี่ยวกับรูปแบบโครงสร้างของลำดับขั้นการเรียนรู้ ได้ให้แนวคิดเกี่ยวกับลำดับขั้นการเรียนรู้ ซึ่งจะเสนอบางแนวคิด ดังนี้

ลำดับขั้นการเรียนรู้ตามแนวคิดของกาเย่

Robert M. Gagne' (อ้างใน พรรณี ช. เจนจิต 2528 : 205-222) ได้แบ่งลำดับขั้นการเรียนรู้ออกเป็น 8 ลำดับนั้น ตามลำดับของสิ่งที่ง่ายไปหาสิ่งที่ซับซ้อน และในการเรียนรู้ลำดับขั้นต่อไปต้องอาศัยลำดับขั้นที่มีมาก่อนเป็นพื้นฐาน ดังนี้

1. Signal Learning การเรียนรู้เครื่องหมายหรือสัญญาณ เป็นการเรียนรู้ชนิดที่ง่ายที่สุด และอยู่ในระดับต่ำสุด ซึ่งผู้เรียนไม่สามารถบังคับพฤติกรรมมิให้เกิดขึ้นได้ (involuntary behavior) โดยทั่วไปจะเกี่ยวข้องกับอารมณ์และความรู้สึก เกิดขึ้นโดยกระบวนการวางเงื่อนไข หรือการทำซ้ำ ๆ

2. Stimulus-Response Learning การเรียนรู้ความสัมพันธ์ระหว่างสิ่งเร้าและการตอบสนอง เป็นการเรียนรู้อันเนื่องมาจากการเชื่อมโยงระหว่างสิ่งเร้าและการตอบสนองที่ผู้เรียนสามารถควบคุมพฤติกรรมได้ (voluntary behavior) เป็นไปอย่างตั้งใจและยอมรับในสิ่งเร้าและการตอบสนองนั้น ทั้งนี้ก็เนื่องมาจากการได้รับการเสริมแรง

3. Chaining การเรียนรู้แบบลูกโซ่ เป็นการเรียนรู้อันเนื่องมาจากการเชื่อมโยงระหว่างสิ่งเร้า และการตอบสนองที่ติดต่อกันเป็นลำดับขั้น มีลักษณะของการกระทำการเคลื่อนไหว (motor skills) เกี่ยวข้องอยู่ด้วย

4. Verbal Association การเรียนรู้โดยการเชื่อมโยงภาษาด้วยถ้อยคำ มีลักษณะเป็นลำดับขั้น และเป็นการเรียนรู้โดยใช้ภาษาที่มีต่อสิ่งเร้าหนึ่งหรือมากกว่า

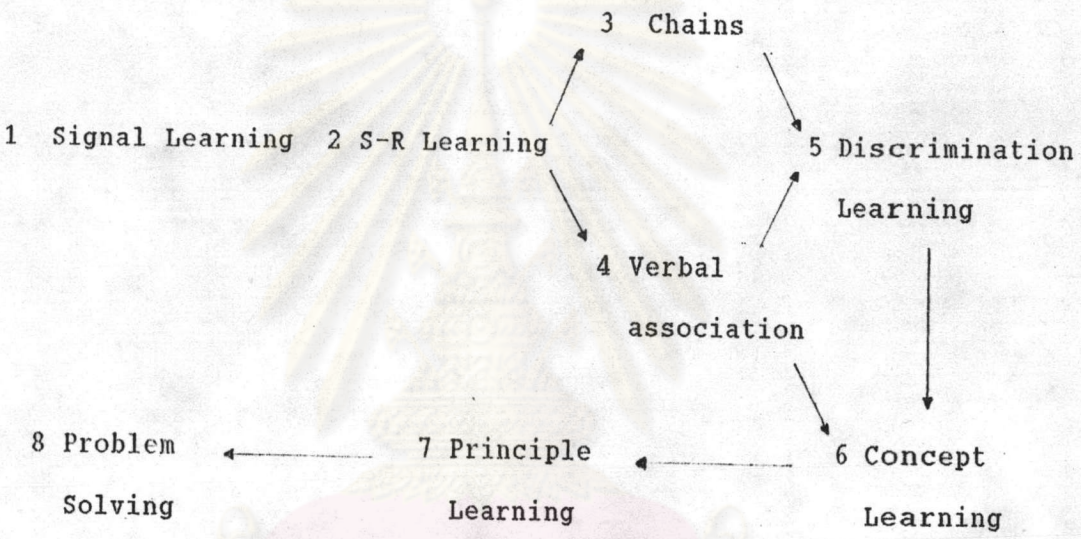
5. Discrimination Learning การเรียนรู้แบบจำแนกความแตกต่าง เป็นการเรียนรู้ที่ผู้เรียนสามารถมองเห็นความแตกต่างของสิ่งเร้าประเภทเดียวกัน ซึ่งทำให้ผู้เรียนสามารถตอบสนองต่อสิ่งที่เหมือนกัน ด้วยวิธีต่าง ๆ กัน

6. Concept Learning การเรียนรู้มโนทัศน์ เป็นการเรียนรู้ที่ผู้เรียนจำแนกสภาพความเหมือนหรือความแตกต่าง และตอบสนองต่อสิ่งเร้าต่าง ๆ ที่จัดอยู่ในกลุ่มเดียวกันได้

7. Principle Learning การเรียนรู้กฎหรือหลักการ เป็นการเรียนรู้ที่เกิดขึ้นเนื่องจากการเชื่อมโยงมโนทัศน์ตั้งแต่ 2 อย่าง หรือมากกว่าเข้าด้วยกัน และสามารถตอบสนองต่อสิ่งเร้าที่แตกต่างกัน โดยอาศัยกฎเกณฑ์ที่กำหนดขึ้น

8. Problem Solving การเรียนรู้การแก้ปัญหา เป็นการเรียนรู้ที่รวบรวมกฎเกณฑ์ต่าง ๆ เข้าด้วยกัน และนำไปใช้แก้ปัญหาในสภาพการณ์ใหม่ได้ โดยอาศัยกฎเกณฑ์ต่าง ๆ ที่เคยเรียนรู้มาแล้ว

กาเย่ ได้เสนอว่า การเรียนรู้ในลำดับขั้นสูง ๆ นั้นต้องอาศัยความรู้ในลำดับขั้นที่มีมาก่อนเป็นพื้นฐาน ซึ่งพิจารณาจากแผนผังลำดับขั้นต่อไปนี้



จากลำดับขั้นการเรียนรู้ทั้ง 8 ขั้นตอนนี้ กาเย่ ได้แบ่งประเภทของสมรรถภาพที่มนุษย์สามารถเรียนรู้ได้ 5 อย่างด้วยกัน คือ ทักษะทางด้านสติปัญญา (intelleetual skill) ยุทธศาสตร์การคิด (Cognitive Strategies) สารสนเทศทางวาจา (Verbal Information) ทักษะการเคลื่อนไหว (motor skills) และทัศนคติ (Attitude)

<u>ลำดับขั้นการเรียนรู้ตามแนวคิดของบรูเนอร์</u>

Jerome S. Bruner เชื่อว่า ความพร้อมของผู้เรียนเป็นสิ่งที่สามารถสอนให้เกิดขึ้นได้อย่างรวดเร็ว โดยการจัดสิ่งแวดล้อมให้สอดคล้องกับขั้นของพัฒนาการทางสติปัญญา

ของเขา บรูเนอร์ เสนอทฤษฎีพัฒนาการทางสติปัญญาที่เป็นการถ่ายทอดประสบการณ์ว่า มีลักษณะของลำดับขั้นการเรียนรู้ 3 ลำดับขั้น (พรรณี ช. เจนจิต 2528 : 118-120) คือ

1. Enative representation เป็นการเรียนรู้โดยการสัมผัสเริ่มตั้งแต่แรกเกิดจนถึง 2 ปี บุคคลต้องอาศัยการกระทำในอันที่จะเพิ่ม และถ่ายทอดประสบการณ์ของตน และใช้เป็นกระบวนการต่อเนื่องไปสู่การเรียนรู้ลำดับขั้นต่อไปได้ตลอดชีวิต

2. Iconic representation เป็นการเรียนรู้โดยการเห็นภาพหรือสิ่งเร้า รวมทั้งการใช้ประสาทสัมผัสต่าง ๆ เป็นการเพิ่มประสบการณ์ให้ผู้เรียนเอง แต่ก็ยังไม่สามารถจำแนกความแตกต่างอย่างมีเหตุผลได้

3. Symbolic representation เป็นการเรียนรู้โดยการใช้สัญลักษณ์ หรือภาษา ในการถ่ายทอดประสบการณ์ ผู้เรียนมีความเข้าใจสิ่งต่าง ๆ ได้ อย่างมีเหตุผล และสามารถเข้าใจสิ่งที่เป็นนามธรรมมากขึ้น ลำดับขั้นนี้ บรูเนอร์ ถือเป็นลำดับขั้นที่สูงสุดของพัฒนาการทางความรู้ ความเข้าใจ

พัฒนาการทั้ง 3 ลำดับขั้นนี้ จะดำเนินไปอย่างเป็นลำดับขั้น และต่อเนื่องกันไปตลอดชีวิต

<u>ลำดับขั้นการเรียนรู้ตามแนวคิดของ Ausubel</u>

David P. Ausubel และ Robinson (1969 : 37-38, 59-71 อ้างใน วรรณวิภา จัตุชัย 2530 : 32-33) เชื่อว่า การเรียนรู้จะเกิดขึ้นได้ ถ้าในการเรียนรู้สิ่งเร้าใหม่นั้น ผู้เรียนเคยมีพื้นฐานซึ่งเชื่อมโยงกับความรู้ใหม่ได้ ซึ่งทำให้การเรียนรู้สิ่งเร้าใหม่นั้นมีความหมาย Ausubel จำแนกลำดับขั้นการเรียนรู้ เป็น 5 ลำดับนั้น ดังนี้

1. Representational Learning เป็นการเรียนรู้ผลสรุป โดยเริ่มจากเรียนรู้ความหมายของสัญลักษณ์ต่าง ๆ แล้วเชื่อมโยงมาเป็นการเรียนรู้ภาษา ซึ่งการพิจารณาจากการเห็นความสัมพันธ์ระหว่าง ชื่อและสิ่งของ รวมทั้งสัญลักษณ์

2. Concept Learning เป็นการเรียนรู้มโนทัศน์ทั้งทางตรรกวิทยา และจิตวิทยา โดยการหาความสัมพันธ์ระหว่างสภาพการณ์ต่าง ๆ ที่กำหนดขึ้น มีกระบวนการ 2 ขั้นตอน คือ การกำหนดมโนทัศน์ (Concept formation) และการเรียนรู้ความหมายของชื่อมโนทัศน์ (Concept name) ใช้อ้างอิงไปยังมโนทัศน์อื่น ๆ ได้

3. Proposition Learning เป็นการเรียนรู้ความสัมพันธ์ระหว่างมโนทัศน์ 2 อย่างขึ้นไป โดยการเรียนรู้ความสัมพันธ์ระหว่างประโยค ชนิดของความสัมพันธ์ และนำความสัมพันธ์นั้นมากำหนดร่วมกันเป็นรูปแบบให้เห็นชัดเจน ทั้งนี้ก็เพื่อแสดงถึงความเป็นทั่วไปของมโนทัศน์นั้น

4. Application หรือ Problem Solving เป็นการเรียนรู้ที่ให้ผู้เรียนค้นพบการเรียนรู้เหล่านั้นด้วยตนเอง โดยอาศัยสภาพการเรียนรู้ในลำดับขั้นที่ผ่านมา ซึ่งผู้ตัดสินสามารถกำหนดเกณฑ์เพื่อพิจารณาวิธีการนำไปใช้แก้ปัญหาของผู้เรียนได้

5. Creativity เป็นการเรียนรู้ความสัมพันธ์ระหว่างสิ่งที่เขาเคยรู้มาก่อนและนำมาสร้างสรรค์แนวคิดใหม่ ๆ ซึ่งในการพิจารณาความคิดสร้างสรรค์ ต้องกำหนดเกณฑ์ไว้ด้วย

ลำดับขั้นการเรียนรู้ตามแนวคิดของ Ausubel และ Robinson นี้ ต้องอาศัยตัวแปรโครงสร้างของลำดับขั้นการเรียนรู้ด้านพุทธิพิสัย ซึ่งได้แก่ ลำดับขั้นของเนื้อหาวิชาที่ผู้เรียนได้ผ่านมาแล้ว ความพร้อมทางพัฒนาการต่าง ๆ มาประกอบด้วย จึงจะทำให้ลำดับขั้นการเรียนรู้ทั้ง 5 ขั้นนี้ ดำเนินไปด้วยดี

<u>ลำดับขั้นการเรียนรู้ตามแนวคิดของ บลูมและคณะ</u>

บลูมและคณะได้เสนอแนวคิดเกี่ยวกับลำดับขั้นการเรียนรู้ไว้ในหนังสือ Taxonomy of Educational Objectives Handbook I : Cognitive Domain ในปี ค.ศ. 1956 และได้รับการตีพิมพ์ถึง 17 ครั้ง ต่อมาได้เสนอหนังสือชื่อ Handbook on Formative and Summative Evaluation of Student Learning ในปี ค.ศ. 1971 ซึ่งแนวคิดของบลูมและคณะได้มีผู้นำไปใช้ในด้านของการปฏิบัติ และการวิจัยอย่างกว้างขวาง โดยเฉพาะด้านการพัฒนาหลักสูตร การจัดกิจกรรมการเรียนการสอน และการวัดและประเมินผลการศึกษา

บลูมและคณะ ได้จำแนกวัตถุประสงค์ทางการศึกษาตามหลักของ Taxonomies ออกเป็นกลุ่มใหญ่ ๆ ที่เรียกว่า โดเมน (Domains) มีอยู่ด้วยกัน 3 โดเมน ได้แก่ Cognitive Domain (กลุ่มพุทธิพิสัย) Affective Domain (กลุ่มจิตพิสัย) และ

Psychomotor Domain (กลุ่มทักษะพิสัย) โดยได้กล่าวถึงลักษณะของแต่ละกลุ่มไว้ดังนี้ (เยาวดี วิบูลย์ศรี 2526 : 24)

1. กลุ่มพุทธิพิสัย หรือกลุ่มความรู้ จะกล่าวถึงการรวมวัตถุประสงค์ต่าง ๆ ทางการศึกษาที่เกี่ยวกับความรู้ที่ระลึกได้ หรือจำได้ (Recall) และการพัฒนาซึ่งเกี่ยวข้องกับ ความสามารถทางสมองและทักษะ (Development of Intellectual Abilities and Skill) ซึ่งจะเป็นศูนย์กลางของการพัฒนาหลักสูตร และการพัฒนาแบบสอบ

2. กลุ่มจิตพิสัย หรือกลุ่มคุณค่า เป็นการรวมวัตถุประสงค์ทางการศึกษาที่เกี่ยวกับการเปลี่ยนแปลงทางด้านอารมณ์ หรือการยอมรับด้านคุณค่าต่าง ๆ เช่น ความสนใจ ทัศนคติ ค่านิยม ความซาบซึ้ง ซึ่งการเปลี่ยนแปลงในขอบเขตนี้ไม่ค่อยแน่นอน และเป็นการยากที่จะกล่าวถึงพฤติกรรมที่เหมาะสมกับวัตถุประสงค์ของกลุ่มนี้

3. กลุ่มทักษะพิสัย หรือกลุ่มพลังทักษะ เป็นการรวมวัตถุประสงค์ทางการศึกษาที่เกี่ยวข้องกับระบบการทำงานที่ต้องอาศัยพื้นฐานทางกล้ามเนื้อ หรือทักษะที่ใช้กลไกทางร่างกายในการจัดการหรือกระทำสิ่งหนึ่งสิ่งใดได้อย่างคล่องแคล่ว

ในแต่ละโดเมนดังกล่างข้างต้น ยังแบ่งออกเป็นลำดับขั้นต่าง ๆ ตามพฤติกรรมที่ง่ายไปหายาก และลักษณะประการหนึ่งของการจำแนกวัตถุประสงค์ทางการศึกษาของบลูมและคณะก็คือ การไม่คำนึงถึงเนื้อหาวิชา อายุของผู้เรียนและชนิดของการสอน (Bloom, et al 1972 : 12) โดยเฉพาะกลุ่มพุทธิพิสัยนั้น สมรรถภาพการเรียนรู้แต่ละลำดับขั้นตั้งแต่ความรู้ความจำ ความเข้าใจ การนำไปใช้ การวิเคราะห์ การสังเคราะห์ และการประเมินผล จะมีลักษณะของการสะสม และความซับซ้อนอยู่ด้วย

การจำแนกสมรรถภาพการเรียนรู้ในกลุ่มพุทธิพิสัย ตามแนวคิดของบลูมและคณะ ซึ่ง Hill (1984 : 184) ได้ให้ข้อสรุปไว้ 4 ข้อด้วยกัน ดังนี้

1. สมรรถภาพในการเรียนรู้ในกลุ่มพุทธิพิสัย แยกได้ 6 ระดับ คือ ความรู้ความจำ ความเข้าใจ การนำไปใช้ การวิเคราะห์ การสังเคราะห์ และการประเมินผล

2. สมรรถภาพต่าง ๆ ในกลุ่มพุทธิพิสัยนี้ เรียงกันอย่างเป็นลำดับขั้น (Hierarchically ordered) กล่าวคือ สมรรถภาพในลำดับขั้นที่สูงกว่า จะมีความซับซ้อนมากกว่าสมรรถภาพในลำดับขั้นที่ต่ำกว่า

3. สมรรถภาพในแต่ละลำดับขั้น มีลักษณะเป็นการสะสม (Cumulative) หมายความว่า สมรรถภาพในลำดับขั้นที่สูงกว่าจะต้องอาศัยสมรรถภาพในลำดับขั้นที่ต่ำกว่าเป็นพื้นฐาน

4. กระบวนการต่าง ๆ ภายใต้ลำดับขั้นของสมรรถภาพที่แตกต่างกันนี้ ไม่ขึ้นอยู่กับอายุของผู้เรียน วิธีการสอน และเนื้อหาวิชาที่สอน

<u>ตอนที่ 2 งานวิจัยที่เกี่ยวข้องกับลำดับขั้นการเรียนรู้ด้านพุทธิพิสัย</u>

Kropp และ Stoker (1966 cited by Seddon 1978 : 309-310, อ้างใน วรรณวิภา จัตุชัย 2530 : 49-50) ได้ศึกษาเกี่ยวกับลำดับขั้นการเรียนรู้ด้านพุทธิพิสัย กลุ่มตัวอย่างเป็นนักเรียนเกรด 9-12 ในโรงเรียนมัธยมศึกษา 10 แห่ง ในรัฐฟลอริดา เครื่องมือที่ใช้เป็นแบบสอบ 4 ฉบับ ซึ่งวัดเนื้อหาด้านวิทยาศาสตร์ 2 ฉบับ และวัดเนื้อหาด้านสังคมศาสตร์ 2 ฉบับ แต่ละฉบับมีข้อสอบ 95 ข้อ วัดสมรรถภาพด้านความรู้ความจำ ความเข้าใจ การนำไปใช้ และการวิเคราะห์อย่างละ 20 ข้อ วัดสมรรถภาพด้านการสังเคราะห์ 5 ข้อ และการประเมินผล 10 ข้อ ใช้กับกลุ่มตัวอย่างฉบับละประมาณ 1,100 - 1,500 คน ในแต่ละเกรด หาความตรงของลำดับขั้นการเรียนรู้ โดยวิธีการ 2 วิธี ได้แก่ การวิเคราะห์ซิมเพลก (Simplex analysis) และวิธีกำลังสองน้อยที่สุดของไกเซอร์ (Kaiser 's least squares) เพื่อพิจารณาการเรียนลำดับของสมรรถภาพ พบว่ามีการเรียงลำดับของสมรรถภาพ เฉพาะความรู้ความจำ ความเข้าใจ การนำไปใช้ และการวิเคราะห์เท่านั้น ส่วนสมรรถภาพด้านการสังเคราะห์ และการประเมินผลนั้นไม่เรียงลำดับกัน

Miller, Snowman และ O' Hara (1979 : 241-248) ได้นำข้อมูลของ Kropp และ Stoker มาวิเคราะห์ใหม่ โดยใช้วิธีการทางสถิติ 5 วิธี คือ การวิเคราะห์

เส้นทาง (Path Analysis) สหสัมพันธ์บางส่วน (Semi - partial Correlation) การวิเคราะห์คอมมอนนอลลิตี้ (Commonality Analysis) การวิเคราะห์ถดถอยแบบสเตปไวส์ (Stepwise Regression Analysis) และการวิเคราะห์ตัวประกอบ (Factor Analysis) พบว่าสมรรถภาพด้านการวิเคราะห์ จะพัฒนาอย่างเป็นอิสระจากสมรรถภาพด้านการสังเคราะห์ และการประเมินผล กล่าวคือรูปแบบโครงสร้างของลำดับขั้นการเรียนรู้ด้านพุทธิพิสัยตามการจำแนกของบลูมและคณะนั้น สมรรถภาพด้านการสังเคราะห์และการประเมินผลจะไม่เกี่ยวข้อง หรืออาศัยพื้นฐานจากสมรรถภาพด้านการวิเคราะห์

รุจิร์ ภู่สาระ (2514) ศึกษาเปรียบเทียบผลการใช้แบบสอบเชนิดต่าง ๆ เพื่อวัดสมรรถภาพในการเรียนรู้ขั้นต่าง ๆ ตามหนังสือ Taxonomy of Education Objectives ของบลูมและคณะ กลุ่มตัวอย่างเป็นนักเรียนชั้นประถมปีที่ 7 จำนวน 465 คน เครื่องมือที่ใช้เป็นแบบสอบวิชาภาษาไทยที่วัดสมรรถภาพในการเรียนรู้ขั้นต่าง ๆ 120 ข้อ ผลการศึกษา พบว่า ในแบบสอบเลือกตอบสมรรถภาพทั้ง 6 ลำดับขั้นมีความสัมพันธ์ซึ่งกันและกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 แต่ในแบบสอบแบบตอบสั้น ๆ สมรรถภาพด้านการสังเคราะห์ไม่สัมพันธ์กันกับด้านการวิเคราะห์ และการประเมินผล และในแบบสอบแบบผสมระหว่าง แบบเลือกตอบกับแบบตอบสั้น ๆ สมรรถภาพด้านการสังเคราะห์ไม่สัมพันธ์กับด้านการวิเคราะห์ และด้านความรู้ความจำ นอกจากนี้ยังพบว่า นักเรียนที่มีผลสัมฤทธิ์ทางการเรียนสูง มีสมรรถภาพในการเรียนรู้สูงกว่า นักเรียนที่มีผลสัมฤทธิ์ทางการเรียนต่ำ ในทุกลำดับขั้น

ชวาล แพรัตกุล (2520) ได้กล่าวยืนยันว่าตัวอย่างข้อสอบที่ได้นำเสนอในหนังสือเทคนิคการเขียนข้อทดสอบ ซึ่งวัดสมรรถภาพทั้ง 6 ขั้นนั้น ได้ผ่านการทดลองใช้และการวิเคราะห์หาคุณภาพของข้อสอบ จากกลุ่มตัวอย่างซึ่งเป็นนักเรียนชั้นประถมปีที่ 7 และชั้นมัธยมศึกษาปีที่ 3 มาแล้ว อันเป็นหลักฐานแสดงให้เห็นว่า นักเรียนชั้นประถมปีที่ 7 ถึงชั้นมัธยมศึกษาปีที่ 3 มีสมรรถภาพด้านพุทธิพิสัยครบทุกขั้นจริง

ปราณี ร่วมทอง (2527) ได้ทำการเปรียบเทียบคุณภาพแบบสอบเลือกตอบตัวเลือกธรรมดา และตัวเลือกซ้อนที่ใช้วัดลำดับขั้นการเรียนรู้ด้านพุทธิพิสัยต่างกัน โดยทดสอบกับนักเรียนที่มีผลสัมฤทธิ์ทางการเรียนต่างกัน วัดสมรรถภาพการเรียนรู้ 2 ลำดับขั้น คือ สมรรถภาพการเรียนรู้ขั้นต่ำ ได้แก่ ความรู้ความจำ ความเข้าใจ การนำไปใช้ และสมรรถภาพการเรียนรู้ขั้นสูง ได้แก่ การวิเคราะห์ การสังเคราะห์ การประเมินผล พบว่า ในการวัดสมรรถภาพการเรียนรู้ขั้นต่ำ แบบสอบเลือกตอบตัวเลือกธรรมดา มีค่าความเที่ยง และค่าอำนาจจำแนกสูงกว่าแบบสอบเลือกตอบตัวเลือกซ้อน อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ในการวัดสมรรถภาพการเรียนรู้ขั้นสูง แบบสอบทั้ง 2 ชนิด มีค่าความเที่ยงความตรง ค่าความยาก และค่าอำนาจจำแนกไม่แตกต่างกัน นักเรียนที่มีผลสัมฤทธิ์ทางการเรียนสูง และต่ำตอบแบบสอบทั้ง 2 ชนิด ไม่แตกต่างกัน แต่นักเรียนที่มีผลสัมฤทธิ์ทางการเรียนปานกลาง ตอบแบบสอบเลือกตอบตัวเลือกธรรมดา ได้คะแนนเฉลี่ยสูงกว่าแบบสอบเลือกตอบตัวเลือกซ้อนอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

วรรณวิภา จัตุชัย (2530) ได้ศึกษารูปแบบโครงสร้างของลำดับขั้นการเรียนรู้ด้านพุทธิพิสัย เพื่ออธิบายลำดับขั้นการเรียนรู้วิชาคณิตศาสตร์ของนักเรียนระดับมัธยมศึกษาตอนต้น โดยศึกษาและเปรียบเทียบการอธิบายลำดับขั้นการเรียนรู้ด้านพุทธิพิสัย จาก 3 รูปแบบคือ Simple linear Model, Miller's model และ Wright's model ซึ่งสมรรถภาพในการเรียนรู้ด้านพุทธิพิสัยมี 6 ลำดับขั้น คือ ความรู้ความจำ ความเข้าใจ การนำไปใช้ การวิเคราะห์ การสังเคราะห์ และการประเมินผล กลุ่มตัวอย่างในการวิจัย คือ นักเรียนชั้นมัธยมศึกษาปีที่ 3 จำนวน 642 คน ในเขตกรุงเทพมหานคร เครื่องมือที่ใช้ในการรวบรวมข้อมูล เป็นแบบสอบวัดผลสัมฤทธิ์วิชาคณิตศาสตร์ ระดับมัธยมศึกษาตอนต้น การวิเคราะห์ข้อมูลใช้วิธีการวิเคราะห์เส้นทาง โดยใช้เทคนิคการประมาณค่าพารามิเตอร์แบบ Partial Least Squares ในการเปรียบเทียบรูปแบบโครงสร้างของลำดับขั้นการเรียนรู้ด้านพุทธิพิสัย แต่ละรูปแบบ กับข้อมูลเชิงประจักษ์ และกับรูปแบบโครงสร้างของลำดับขั้นการเรียนรู้ด้านพุทธิพิสัยแบบเต็มรูป ผลการวิจัย พบว่า

1. Simple linear model, Miller's model และ Wright's model อธิบายลำดับขั้นการเรียนรู้ของข้อมูลเชิงประจักษ์ได้ประมาณร้อยละ 54, 55 และ 57 ตามลำดับ และทั้ง 3 รูปแบบอธิบายข้อมูลเชิงประจักษ์ได้น้อยกว่า รูปแบบโครงสร้างของลำดับขั้นการเรียนรู้ด้านพุทธิพิสัยแบบเต็มรูป อย่างมีนัยสำคัญทางสถิติ

2. รูปแบบโครงสร้างของลำดับขั้นการเรียนรู้ด้านพุทธิพิสัยแบบใหม่ อธิบายข้อมูลเชิงประจักษ์ได้ไม่แตกต่างอย่างมีนัยสำคัญทางสถิติจากรูปแบบเต็มรูป และมีลักษณะคล้ายกับรูปแบบเต็มรูป กล่าวคือสมรรถภาพระดับต่ำ ส่งผลทั้งทางตรง และทางอ้อมไปยังสมรรถภาพระดับสูง แต่ต่างกันที่สมรรถภาพปลายทางของรูปแบบใหม่ 2 ด้าน คือ สมรรถภาพด้านการสังเคราะห์ ไม่ส่งผลไปยังสมรรถภาพด้านการประเมินผล

ตอนที่ 3 มโนทัศน์เกี่ยวกับโมเดลโลจีสติกในทฤษฎี IRT

ความเป็นมาของทฤษฎี IRT

ผู้ที่ริเริ่มนำทฤษฎี IRT มาใช้ในวงการวัดและประเมินผลการศึกษา คือ Ferguson และ Lawley ในปี ค.ศ. 1942 และ 1943 โดยอาศัยหลักการว่า ผลการสอบของผู้สอบแต่ละคนในการทำแบบสอบใด ๆ ขึ้นอยู่กับคุณลักษณะแฝง (Latent trait) หรือความสามารถ (ability) หรือทักษะ (Skill) ของผู้สอบ ต่อมาในปี ค.ศ. 1952 Lord ได้เสนอทฤษฎีนี้ โดยเรียกทฤษฎีดังกล่าวว่า ทฤษฎีโค้งคุณลักษณะของข้อสอบ (Item Characteristic Curve Theory) ซึ่งต่อมาเรียกว่า Normal Ogive Model แต่ Lord ไม่ได้พัฒนาต่อเนื่องจากโมเดลนี้มีการคำนวณยุ่งยากมาก และยังไม่มีโปรแกรมคอมพิวเตอร์ที่จำเป็นต้องใช้วิเคราะห์ข้อมูล ตามแนวคิดของโมเดลนี้

ต่อมาปี ค.ศ. 1960 Rasch ได้เสนอแนวคิดหนึ่งของทฤษฎี IRT เรียกว่า Rasch Model ซึ่งเป็นโมเดลที่มีพารามิเตอร์ 1 ตัว คือ ค่าความยาก ต่อมาในปี 1965 Lord ก็ได้พัฒนาทฤษฎี IRT นี้อีกครั้งหนึ่ง เนื่องจาก Lord พบว่า ปัญหาที่พบจากการศึกษาครั้งก่อนไม่เป็นปัญหาจริง ๆ และในปี ค.ศ. 1968 Birnbaum ได้เสนอแนวคิดเกี่ยวกับ Logistic Model และ Normal Ogive Model ที่มี 2 และ 3 พารามิเตอร์ ทำให้

นักวัดผลเพิ่มความสนใจแนวคิดของทฤษฎี IRT มากขึ้น ต่อมาในปี ค.ศ.1977 Lord ได้เปลี่ยนชื่อ Item Characteristic Curve Theory มาเป็น Item Response Theory และใช้ logistic function แทน normal ogive function เนื่องจากวิธีการคำนวณทางคณิตศาสตร์สะดวกกว่า (Warm 1978 : 25)

ข้อตกลงเบื้องต้นของทฤษฎี IRT

ทฤษฎี IRT มีข้อตกลงเบื้องต้น ดังนี้

1. แบบสอบมีคุณสมบัติวัดเพียงคุณลักษณะเดียว (Unidimensionality) กล่าวคือ ข้อสอบในแบบสอบต้องมีลักษณะเป็นเอกพันธ์ (Homogeneous) หมายความว่าแบบสอบนั้นจะต้องมุ่งวัดคุณลักษณะภายใน หรือความสามารถอย่างใดอย่างหนึ่งเพียงความสามารถเดียว ส่วนการตรวจสอบว่าข้อมูลจากการสอบเป็นไปตามข้อตกลงเบื้องต้นนี้หรือไม่ทำได้โดยการวิเคราะห์ตัวประกอบ (Factor Analysis)

2. ความเป็นอิสระในการตอบข้อสอบ (Local Independence) ข้อตกลงเบื้องต้นข้อนี้กำหนดว่า การตอบสนองต่อข้อสอบต่าง ๆ กันในแบบสอบของผู้สอบคนใดคนหนึ่งมีความเป็นอิสระในทางสถิติ หมายความว่า การตอบข้อสอบข้อใดข้อหนึ่งของผู้สอบคนใดคนหนึ่งต้องไม่มีผลต่อการตอบข้อสอบข้ออื่น ๆ ในแบบสอบ Lord กล่าวว่า คุณสมบัติ Local Independence นี้จะมีโดยอัตโนมัติ เมื่อแบบสอบมีคุณสมบัติของ Unidimensionality (Lord 1980 : 19) ดังนั้นในการตรวจสอบ Local Independence ทำได้โดยอาศัยเทคนิคการวิเคราะห์ตัวประกอบ เหมือนกัน

3. โค้งลักษณะข้อสอบ (Item Characteristic Curve) โค้งลักษณะข้อสอบเป็นฟังก์ชันทางคณิตศาสตร์ ที่แสดงถึงความสัมพันธ์ระหว่าง โอกาสของการตอบข้อสอบข้อใดข้อหนึ่งได้ถูกต้อง กับ ความสามารถหรือคุณลักษณะที่วัดโดยข้อสอบข้อนั้น และเป็นฟังก์ชันการถดถอยชนิดไม่เป็นเชิงเส้นตรง (Non - linear Regression Function) ของคะแนนข้อสอบ บนความสามารถที่วัดโดยแบบสอบ โค้งลักษณะข้อสอบมีหลายรูปแบบทั้งนี้ขึ้นอยู่กับโมเดลที่ใช้อธิบายความสัมพันธ์ ดังกล่าว

ลักษณะสำคัญของทฤษฎี IRT

ทฤษฎี IRT นี้ ถ้าสามารถนำมาใช้ให้สอดคล้องกับข้อตกลงเบื้องต้นได้ ย่อมมีความเหนือกว่า (Superiority) ทฤษฎีการวัดแบบดั้งเดิม ซึ่งได้รับการวิเคราะห์ว่าเป็นทฤษฎีที่ตั้งอยู่บนข้อตกลงเบื้องต้นที่ค่อนและไม่ค่อยจะสมเหตุสมผล (Lord 1980 : 9) ดังนี้

1. ความไม่แปรเปลี่ยนของค่าพารามิเตอร์ของข้อสอบ (Invariant of Item Parameter) หมายความว่า ไม่ว่าจะประมาณค่าพารามิเตอร์ของข้อสอบ จากกลุ่มตัวอย่างที่มีความสามารถระดับใดก็ตาม ค่าพารามิเตอร์จะไม่แปรเปลี่ยนไปตามกลุ่มของผู้สอบ ซึ่งค่าพารามิเตอร์ดังกล่าวได้แก่ ค่าความยาก ค่าอำนาจจำแนก และค่าการเดาของข้อสอบ ในขณะที่ค่าสถิติต่าง ๆ ของข้อสอบที่วิเคราะห์ตามกรอบของทฤษฎีการวัดแบบดั้งเดิมนั้น มีค่าแปรเปลี่ยนไปตามกลุ่มตัวอย่างที่มีระดับความสามารถแตกต่างกัน ซึ่งค่าสถิติดังกล่าวจะมีประโยชน์ในการสร้างแบบสอบเฉพาะแต่กลุ่มประชากรผู้สอบที่มีลักษณะคล้ายคลึง กับกลุ่มตัวอย่างผู้สอบที่ใช้ในการหาค่าสถิติเหล่านั้น เท่านั้น

แต่อย่างไรก็ตาม Lord (1980 : 36) ได้อธิบายไว้ว่าการไม่แปรเปลี่ยนของค่าพารามิเตอร์นั้น ไม่ได้หมายความว่าค่าพารามิเตอร์ที่ได้จากการประมาณค่า โดยใช้กลุ่มตัวอย่างที่แตกต่างกันจะมีค่าเท่ากันเสมอ ทั้งนี้ขึ้นอยู่กับเงื่อนไขการสอบบางประการ เช่น ถ้าเลือกมาตรวัดที่มีจุดเริ่มต้นเดียวกัน และหน่วยในการวัดหน่วยเดียวกันแล้ว ค่าพารา-มิเตอร์ที่ประมาณค่าจากกลุ่มตัวอย่างที่แตกต่างกัน จะมีค่าเท่ากัน ในขณะเดียวกัน ถ้าเลือกมาตรวัดที่มีจุดเริ่มต้น และมีหน่วยในการวัดแตกต่างกันแล้ว การไม่แปรเปลี่ยนของค่าพารามิเตอร์ หมายความว่าค่าพารามิเตอร์ที่ประมาณได้จากกลุ่มตัวอย่างที่ต่างกัน ของข้อสอบชุดหนึ่งจะมีความสัมพันธ์กันในเชิงเส้นตรง

2. การประมาณค่าความสามารถของผู้สอบ ไม่จำเป็นต้องใช้แบบสอบชุดเดียวกัน หรือแบบสอบคู่ขนาน ในขณะที่ตามแนวคิดของทฤษฎีการวัดแบบดั้งเดิมจำเป็นต้องใช้ ซึ่งมโนทัศน์ของแบบสอบคู่ขนานเป็นสิ่งที่เป็นไปได้ยากในทางปฏิบัติ

3. ค่าความคลาดเคลื่อนมาตรฐานของการวัดของผู้เข้าสอบแต่ละคน สามารถประมาณได้โดยอิสระจากกัน ซึ่งทำให้การประมาณค่าความสามารถของผู้เข้าสอบแต่ละคนมีความแม่นยำสูง ในขณะที่ทฤษฎีการวัดแบบดั้งเดิม มีการสมมุติว่า ความแปรปรวนของความ

คลาดเคลื่อนมาตรฐานของการวัดมีค่าเท่ากัน สำหรับผู้สอบทุกคนซึ่งไม่สอดคล้องกับสภาพความเป็นจริง (Hambleton and Swaminathan 1985 : 2-3)

โมเดลที่สำคัญในทฤษฎี IRT

โมเดลของทฤษฎี IRT เป็นโมเดลทางคณิตศาสตร์ สำหรับโมเดลที่สำคัญ และกล่าวถึงกันมาก มีดังนี้

1. Normal Ogive Model เป็นโมเดลที่ใช้อธิบายโค้งลักษณะข้อสอบ (Item Characteristic Curve : ICC) โดยมีความเชื่อว่า ความสัมพันธ์ระหว่างโอกาสในการตอบข้อสอบถูก กับระดับความสามารถ อยู่ในรูปของฟังก์ชันการแจกแจงสะสมปกติ (Cumulative normal ogive) ซึ่งมีรูปสมการ ดังนี้

$$P_i(\theta) = \int_{-\infty}^{a_i(\theta - b_i)} \frac{1}{\sqrt{2\pi}} e^{-1/2t^2} dt \; ; \; i=1,2,3,\ldots,n$$

เมื่อ $P_i(\theta)$ คือ โอกาสที่ผู้สอบคนหนึ่ง ซึ่งมีระดับความสามารถ θ จะตอบข้อสอบข้อ i ได้ถูกต้อง

$\frac{e^{-1/2t^2}}{\sqrt{2\pi}}$ คือ normal density function

b_i คือ ค่าความยากของข้อสอบข้อ i

a_i คือ ค่าอำนาจจำแนกของข้อสอบข้อ i

θ คือ ระดับความสามารถของผู้สอบ

เส้นโค้งลักษณะข้อสอบ จะมีลักษณะคล้ายรูปตัว S ซึ่งค่า $P_i(\theta)$ จะเข้าใกล้ 0 เมื่อ θ จะมีค่าน้อย ๆ และจะมีค่าเพิ่มขึ้นเรื่อย ๆ โดยจะเข้าใกล้ 1 เมื่อ θ มีค่ามาก ๆ จากสมการดังกล่าว จะเห็นได้ว่ามีค่าพารามิเตอร์ที่แสดงถึงลักษณะของข้อสอบเพียง 2 ตัว คือ a_i กับ b_i ซึ่งเราเรียกโมเดลนี้ว่า Two - Parameter Normal Ogive Model

แต่เนื่องจากการสอบที่ใช้ข้อสอบแบบปรนัย ซึ่งผู้สอบที่มีความสามารถต่ำมาก ก็มีโอกาสตอบข้อสอบถูกโดยการเดา โมเดลที่มี 2 พารามิเตอร์ไม่เหมาะที่จะใช้อธิบายสถานการณ์การสอบนี้ได้ ทั้งนี้เพราะในกรณีดังกล่าวถ้า θ มีค่าน้อย ๆ ค่า $P_i(\theta)$ จะไม่เข้าใกล้ 0 อีกต่อไป แต่จะมีค่าเข้าใกล้ค่าที่มากกว่า 0 จึงทำให้มีการเพิ่มค่าพารามิเตอร์ของข้อสอบอีก 1 ตัว โดยกำหนดให้ เป็น c_i และมีรูปสมการ ดังนี้

$$P_i(\theta) = a_i + (1-c_i) \int_{-\infty}^{a_i(\theta-b_i)} \frac{1}{\sqrt{2\pi}} e^{-1/2t^2} dt \; ; \; i=1,2,3,\ldots,n$$

เมื่อ c_i คือ ค่าการเดาของข้อสอบข้อ i

โมเดลนี้เรียกว่า Three - Parameter Normal Ogive Model

2. Logistic Model เป็นโมเดลทางคณิตศาสตร์ที่สำคัญที่อธิบายโค้งลักษณะข้อสอบ ซึ่งได้รับการพัฒนาและมีการนำไปใช้มากที่สุดในสถานการณ์ของการสอบ เนื่องจากมีวิธีการคำนวณทางคณิตศาสตร์สะดวกกว่า Normal Ogive Model และค่าฟังก์ชัน normal ogive กับค่าฟังก์ชัน logist แตกต่างกันเล็กน้อย เพราะมีการปรับค่าตัวแปรด้วย scaling factor (มีค่า = 1.7) ดังนั้นโมเดลทั้งสองจึงสามารถใช้แทนกันได้ (Warm 1978 : 25) นอกจากนี้ ในสถานการณ์ของการสอบจริง อาจจะมีผู้ที่มีความสามารถสูงตอบข้อสอบผิดด้วยความเลินเล่อ ซึ่งกรณีดังกล่าว Logistic Model มีความแกร่งต่อข้อมูลลักษณะนี้มากกว่า Normal Ogive Model จึงทำให้ Logistic Model ได้รับความนิยมกันมากในการปฏิบัติงานจริง (Lord 1980 : 14)

Logistic Model ได้ถูกจำแนกออกเป็น 4 โมเดลตามจำนวนของค่าพารามิเตอร์ ของโค้งลักษณะข้อสอบ ดังนี้

2.1 One - Parameter Logistic Model โมเดลนี้ได้รับการพัฒนาโดย Birnbaum ในปี ค.ศ. 1968 ซึ่งบังเอิญตรงกับ Rasch Model ที่พัฒนาโดยนักคณิตศาสตร์ชาวเดนมาร์ก ชื่อ Georg Rasch ในปี ค.ศ. 1960 เป็นโมเดลที่อธิบายโค้งลักษณะ

ข้อสอบด้วยค่าพารามิเตอร์เพียงตัวเดียว คือ ค่าความยากของข้อสอบ (b_i) โดยมีความเชื่อว่า โอกาสที่ผู้สอบจะทำข้อสอบได้ถูกหรือไม่นั้น ขึ้นอยู่กับระดับความสามารถ (θ) ของตนเอง กับระดับความยากของข้อสอบ ดังนั้นจึงถือว่า ค่าการเดาของข้อสอบ (c_i) เท่ากับ 0 และค่าอำนาจจำแนกของข้อสอบ (a_i) ทุกข้อจะมีค่าเท่ากัน มีรูปสมการ ดังนี้

$$P_i(\theta) = \frac{e^{(\theta - b_i)}}{1 + e^{(\theta - b_i)}} \quad ; i=1,2,3,\ldots,n$$

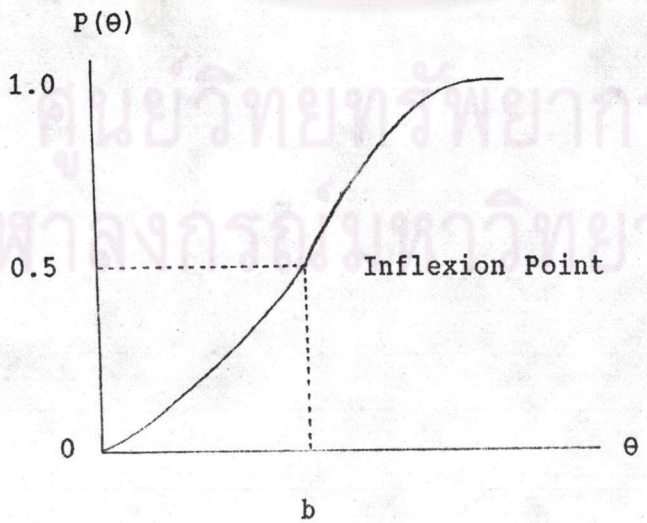
เมื่อ $P_i(\theta)$ คือ โอกาสที่ผู้สอบซึ่งมีความสามารถ θ จะทำข้อสอบข้อ i ได้ถูกต้อง

θ คือ ระดับความสามารถของผู้สอบ

b_i คือ ค่าความยากของข้อสอบ

e คือ ค่าคงที่ มีค่าเท่ากับ 2.71828...

แผนภาพที่ 1 แสดงความหมายของค่าพารามิเตอร์ของข้อสอบ ของ One - Parameter Logistic Model



One - Parameter Logistic Model ต่างจาก Rasch Model คือใช้ในการวิเคราะห์ปัญหาการวัด ในขณะที่ Rasch Model นิยมใช้ในการสร้างแบบสอบ (อุทุมพร จามรมาน 2529 : 97)

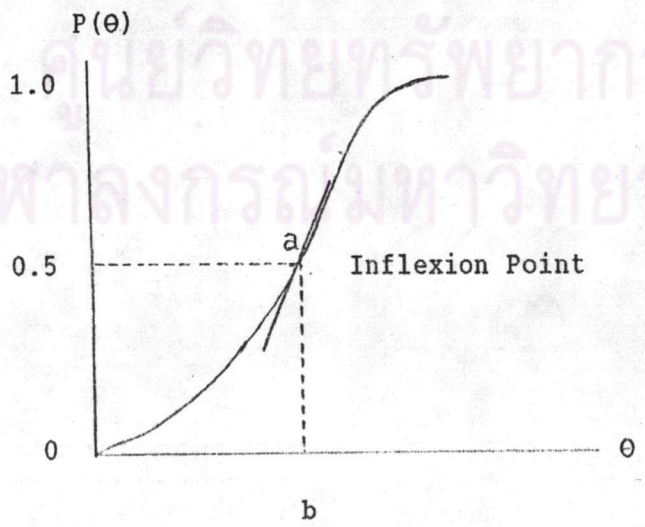
2.2 Two - Parameter Logistic Model เป็นโมเดลที่กำหนดให้ค่าการเดาของข้อสอบ (c_i) เป็น 0 รูปสมการก็จะแสดงค่าพารามิเตอร์เพียง 2 ตัว คือ ค่าอำนาจจำแนกของข้อสอบ (a_i) และค่าความยากของข้อสอบ (b_i) ดังนี้

$$P_i(\theta) = \frac{e^{Dai(\theta-bi)}}{1 + e^{Dai(\theta-bi)}} \quad ; i=1,2,3,\ldots n$$

เมื่อ a_i คือค่าอำนาจจำแนกของข้อสอบข้อ i

D คือ Scaling Factor มีค่า 1.7 เป็นค่าที่มากที่สุดในการปรับ logistic model และ normal ogive model ให้สอดคล้องกัน

แผนภาพที่ 2 แสดงความหมายของค่าพารามิเตอร์ของข้อสอบของ Two - Parameter Logistic Model



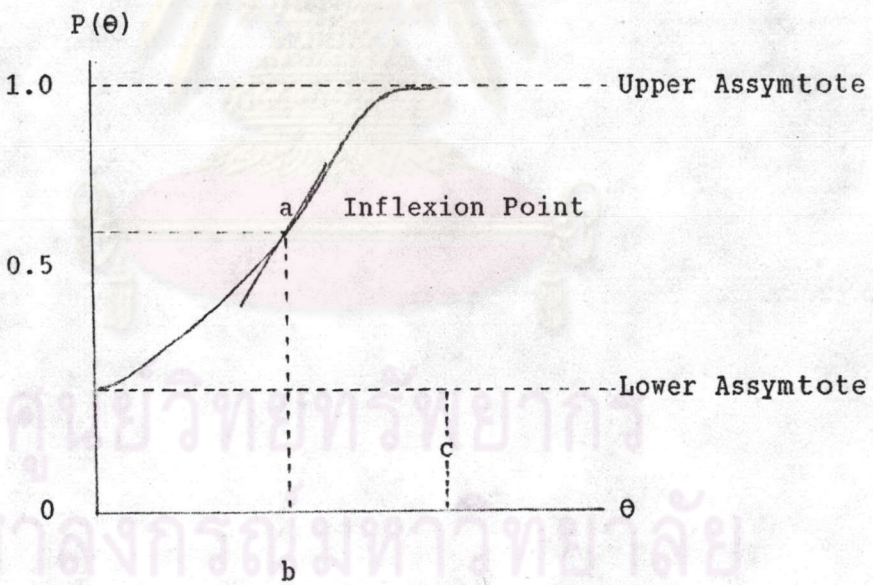
ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

2.3 Three - Parameter Logistic Model เป็นโมเดลที่พัฒนามาจาก Two - Parameter Logistic Model เพื่อให้เหมาะสมกับสถานการณ์ของการสอบที่มีอิทธิพลของการเดาเข้ามาเกี่ยวข้องอยู่ด้วย มีรูปสมการ ดังนี้

$$P_i(\theta) = c_i + (1-c_i)\ \frac{e^{Da_i(\theta-b_i)}}{1+e^{Da_i(\theta-b_i)}} \quad ; i=1,2,3,\ldots n$$

เมื่อ c_i คือ ค่าการเดาของข้อสอบข้อ i

แผนภาพที่ 3 แสดงความหมายของค่าพารามิเตอร์ของข้อสอบ
ของ Three - Parameter Logistic Model



2.4 Four - Parameter Logistic Model เป็นโมเดลที่แตกต่างจาก Three - Parameter Logistic Model คือมีพารามิเตอร์อีก 1 ตัว คือ γ_i ซึ่งมีค่าต่ำกว่า 1 เล็กน้อย ซึ่งหมายถึง ค่าที่ผู้สอบที่มีความสามารถสูง ตอบข้อสอบไม่ถูกเนื่องจากความสะเพร่า แต่โมเดลนี้ยังไม่สามารถหาวิธีปฏิบัติได้ รูปสมการเป็น ดังนี้

$$P_i(\theta) = c_i + (\gamma_i - c_i) \frac{e^{Da_i(\theta - b_i)}}{1 + e^{Da_i(\theta - b_i)}} ; \quad 1,2,3,\ldots,n$$

ค่าพารามิเตอร์ a, b, c ที่ได้จากการประมาณตาม Logistic Model

ค่าพารามิเตอร์ของข้อสอบเป็นดัชนีที่บอกถึงลักษณะของข้อสอบและทำให้รูปร่างของโค้งลักษณะข้อสอบ แต่ละข้อแตกต่างกันไป ค่าพารามิเตอร์แต่ละตัวมีความหมายดังนี้

ค่าอำนาจจำแนกของข้อสอบ (a - parameter or discrimination power) แทนด้วย a_i มีค่าเป็นสัดส่วนโดยตรงกับความชันของเส้นโค้งลักษณะข้อสอบ ณ จุดเปลี่ยนโค้ง (inflexion point) หรือจุดบนเส้น ICC ณ ที่โค้งมีความชัน (Slope) มากที่สุด (ความชัน ณ จุดเปลี่ยนโค้ง = $a_i (1-c_i) / \sqrt{2\pi}$) จุดนี้เป็นจุดที่แสดงความสามารถในการจำแนกผู้สอบที่มีความสามารถต่ำและสูงออกจากกันได้อย่างเด่นชัด ถ้าค่าอำนาจจำแนกของข้อสอบยิ่งมีค่ามาก โค้งลักษณะข้อสอบ (ICC) จะยิ่งมีความชันมาก ในทางทฤษฎีค่า a_i ที่เป็นไปได้อยู่ระหว่าง 0 ถึง $+\infty$ แต่ในทางปฏิบัติค่า a_i ของข้อสอบโดยทั่วไปจะมีค่าระหว่าง 0.50 ถึง 2.50 และค่า a_i ที่มากกว่า 0.80 นับว่าเป็นค่าที่ดีมาก (Warm 1978 : 52, Ree 1979 : 372)

เมื่อเปรียบเทียบค่า a_i กับค่าอำนาจจำแนกที่ได้จากการคำนวณตามแนวของทฤษฎีการวัดแบบดั้งเดิม แล้ว ค่า a_i มีความสัมพันธ์กับค่า r_{bis} ซึ่งใช้เป็นค่าอำนาจจำแนกของทฤษฎีการวัดแบบดั้งเดิม ดังนี้ (Lord 1980 : 33)

$$a_i \cong \frac{r_{bis}}{\sqrt{1-r_{bis}^2}}$$

และ $$r_{bis} \cong \frac{a_i}{\sqrt{1+a_i^2}}$$

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

i 10310010

เมื่อเครื่องหมาย $\cong$ แทนการเท่ากันโดยประมาณ ซึ่งความสัมพันธ์นี้จะเป็นจริงเมื่อ θ มีการแจกแจงเป็นโค้งปกติ และเมื่อไม่มีการเดา ($c = 0$)

นอกจากนี้ ยังมีผู้พบว่า ในกรณีที่มีการเดาในการตอบข้อสอบและ θ มีการแจกแจงเป็นโค้งปกติแล้ว ค่า a_i มีความสัมพันธ์กับค่า r_{pbis} ซึ่งเป็นค่าคำนวณจำแนกของทฤษฎีการวัดแบบดั้งเดิม เช่นเดียวกัน ดังนี้ (Warm 1978 : 51)

$$\bar{a}_i \cong \frac{r_{pbis}\sqrt{pq}}{\sqrt{KR_{20}(1-c_i)^2 y^2 - r_{pbis}\, pq}} \qquad i = 1,2,3,\ldots,n$$

เมื่อ r_{pbis} คือ point - biserial correlation

c_i คือ ค่าการเดาของข้อสอบข้อ i

p คือ อัตราส่วนผู้ที่ทำข้อสอบข้อ i ได้ถูกต้อง

q คือ 1 - p

y คือ ค่าพิกัดฉาก (Ordinate) ของโค้งที่มีการแจกแจงเป็นปกติ ตรงจุดแบ่งระหว่างค่า p และ q

ค่าความยากของข้อสอบ (b - parameter or difficulty power) แทนด้วย b_i เป็นค่าของระดับความสามารถ θ ระดับหนึ่งที่ทำให้เส้นโค้งลักษณะข้อสอบมีการเปลี่ยนโค้ง และเมื่อไม่มีการเดา ($c = 0$) ผู้สอบที่มี $\theta = b$ มีโอกาสทำข้อสอบนั้น ๆ ได้ถูกต้องเท่ากับ 0.5 แต่ถ้ามีการเดา ($c \neq 0$) โอกาสของการทำข้อสอบได้ถูกต้อง เท่ากับ $(1 + c) / 2$

โดยปกติค่าความยากของข้อสอบที่ได้จากทฤษฎี IRT จะมีค่าตั้งแต่ $-\infty$ ถึง $+\infty$ แต่ส่วนมากค่าความยากของข้อสอบวัดผลสัมฤทธิ์โดยทั่วไปจะมีค่าตั้งแต่ - 2.5 ถึง 2.5 โดยถือว่าค่าความยากที่น้อยกว่า - 2.5 เป็นข้อที่ง่ายมาก และค่าความยากที่มากกว่า 2.5 เป็นข้อที่ยากมาก (Warm 1978 : 52) แต่ข้อสอบที่มีค่าความยากตั้งแต่ -2 ถึง 2 ควรได้รับการคัดเลือก เพื่อนำไปใช้ (Urry 1977 : 196)

ถ้า θ มีการแจกแจงเป็นโค้งปกติ ค่าความยาก (b) นี้มีความสัมพันธ์กับค่าความยากง่าย (p - value) ของทฤษฎีการวัดแบบดั้งเดิม ดังนี้ (Warm 1978 : 52)

$$b_i \doteq \frac{yz\ (1-c_i)\ \sqrt{KR_{20}}}{r_{pbis}\ \sqrt{pq}} \qquad ; i=1,2,3,\ldots,n$$

เมื่อ b_i คือ difficulty power ของข้อสอบข้อ i

z คือ Z - score ที่จุดตัด p ของโค้งที่มีการแจกแจงปกติ

p คือ สัดส่วนของผู้ทำข้อสอบนั้น ๆ ถูก

q คือ $1 - p$

y คือ ความสูงของแกน y ณ จุดที่แบ่งค่า p และ q

c_i คือ ค่าการเดาของข้อสอบข้อ i

KR_{20} คือ ค่าความเที่ยงแบบ Kuder - Richardson สูตรที่ 20 ของแบบสอบนั้น

จากสมการ จะเห็นได้ว่า เมื่อค่าความยากง่าย (p-value) ตามทฤษฎีการวัดแบบดั้งเดิม มีค่าเพิ่มขึ้น ค่าความยากของข้อสอบ (b) ตามทฤษฎี IRT จะมีค่าลดลง ซึ่ง Hulin (1983 : 96) กล่าวว่า ค่า b กับค่า p มีความสัมพันธ์กันทางลบ (negative correlation)

<u>ค่าการเดาของข้อสอบ</u> (c - parameter of pseudo - guessing parameter or pseudo - chance score level) แทนด้วย c_i เป็นค่ากำกับเส้นโค้งที่ต่ำสุด (lower assymtode) ของโค้งลักษณะข้อสอบที่แสดงถึง ค่าโอกาสที่ผู้สอบที่มีความสามารถต่ำมาก ($\theta = -\infty$) จะทำข้อสอบถูก (Lord 1980 : 12)

โดยปกติค่า c_i จะมีค่าตั้งแต่ 0.00 ถึง 0.40 ข้อสอบที่ค่า c_i ตั้งแต่ 0.30 ขึ้นไป เป็นข้อสอบที่ไม่ดี ซึ่งข้อสอบที่ดีควรมีค่า c_i ประมาณ 0.20 หรือต่ำกว่า ค่า c_i ยิ่งต่ำแสดงว่าข้อสอบข้อนั้นยิ่งดี และถ้าค่า $c_i = 0$ ถือว่าข้อสอบนั้นดีที่สุด (Warm 1978 : 53) แต่โดยทั่วไปแล้วค่า c_i จะมีค่าประมาณ $1/A - 0.05$ เมื่อ A เป็นจำนวนตัวเลือกในข้อสอบ ดังนั้นข้อสอบที่มี 4 ตัวเลือก จึงมีค่า c_i ประมาณ 0.20

Test and Item Information Functions

ตามทฤษฎี IRT การหาคุณภาพของแบบสอบ ทำได้โดยพิจารณาจาก Test Information Function ที่ได้จากผลรวมของ Item Information Function ซึ่งจะเป็นตัวบอกความแน่นอนในการประมาณค่าความสามารถที่แท้จริง (Lord 1980 : 72) และ Item Information Function สำหรับโมเดลโลจีสติก 3 พารามิเตอร์ ในแต่ละระดับความสามารถ θ สามารถหาได้ จากสูตร ดังนี้

$$I\{\theta, U_i\} = (P_i')^2 / P_iQ_i \qquad ; i=1,2,3,\ldots,n$$

เมื่อ $I\{\theta, U_i\}$ คือ ค่า Item Information Function

P_i คือ ความชันของ ICC ที่ระดับความสามารถ θ ของข้อสอบข้อที่ i หรือค่าอนุพันธ์ (derivative) ของโอกาสในการตอบข้อที่ i ถูก

P_i คือ โอกาสที่ผู้สอบที่มีความสามารถ θ จะตอบข้อสอบข้อ i ได้ถูกต้อง

Q_i คือ $1 - P_i$

จากสมการจะเห็นว่า Item Information Function (IIF) ขึ้นอยู่กับความชันของ ICC ถ้า ICC ชันมากขึ้น ในขณะที่ความแปรปรวนของการตอบถูกยิ่งน้อยลง Item Information Curve (IIC) ที่ระดับความสามารถนั้น ก็ยิ่งสูงขึ้น ซึ่งความสูงของ IIC อยู่ที่ระดับความสามารถใด ก็แสดงว่าสามารถจำแนกระดับความสามารถของผู้สอบได้ดี ณ ระดับความสามารถนั้น (Lord 1980 : 21)

สำหรับ Three - Parameter Logistic Model จะให้ค่า Information ของข้อสอบแตกต่างกันไปตามค่าพารามิเตอร์ความยาก (b) ค่าอำนาจจำแนก (a) และค่าการเดา (c) ซึ่งพิจารณาจากตารางที่ 1

ตารางที่ 1 แสดงค่า derivatives และค่า Item Information Function ของโมเดล Logistic

Model	P_i'	$I\{\theta, U_i\}$
Three-Parameter model	$\frac{(1-c_i)Da_i \exp\{-Da_i(\theta-b_i)\}}{[1+\exp\{-Da_i(\theta-b_i)\}]^2}$	$\frac{(1-c_i)D^2a^2 \exp\{-Da_i(\theta-b_i)\}}{[1+c_i\exp\{-Da_i(\theta-b_i)\}][1+\exp\{-Da_i(\theta-b_i)\}]}$
Two-Parameter model	$\frac{Da_i \exp\{-Da_i(\theta-b_i)\}}{[1+\exp\{-Da_i(\theta-b_i)\}]^2}$	$\frac{D^2 a_i \exp\{-Da_i(\theta-b_i)\}}{[1+\exp\{-Da_i(\theta-b_i)\}]^2}$
One-Parameter	$\frac{D \exp\{-D(\theta-b_i)\}}{[1+\exp\{-D(\theta-b_i)\}]^2}$	$\frac{D^2 \exp\{-D(\theta-b_i)\}}{[1+\exp\{-D(\theta-b_i)\}]^2}$

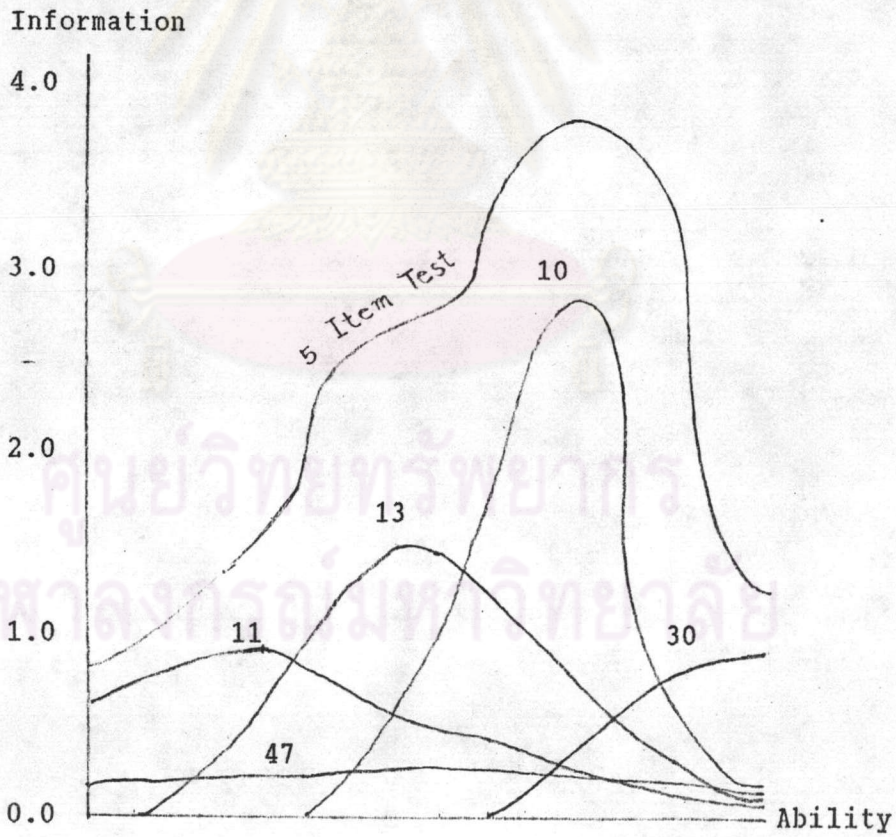
ส่วน Test Information Function (TIF) สามารถหาได้จากผลรวมของ IIF จากรูปสมการดังนี้

$$I\{\theta\} = \sum_{i=1}^{n} I\{\theta, U_i\} \qquad ; i=1,2,3,.....,n$$

จากสูตร ถ้าเรานำค่า TIF มาเขียนเป็นภาพจะได้โค้งที่เรียกว่า Test Information Curve (TIC)

ถ้ามีกลุ่มของข้อสอบอยู่ชุดหนึ่งที่สามารถทราบ Information Curve ของแต่ละข้อ เราก็สามารถสร้างแบบสอบฉบับหนึ่ง ให้มี TIC ณ ระดับหนึ่งของความสามารถที่เราต้องการได้ ซึ่งก็หมายถึงว่า เราจะสามารถสร้างแบบสอบให้เป็นไปตามจุดมุ่งหมายของการสอบได้ เช่น ถ้าต้องการได้แบบสอบคัดเลือก ก็จะต้องเลือกใช้ข้อสอบที่มีความสูงสุดของโค้งที่ระดับความสามารถสูง ๆ ทั้งนี้ก็เพื่อให้ได้ TIC สูงที่ระดับความสามารถสูง ๆ ตามที่เราต้องการ (Warm 1978 : 74-75)

แผนภาพที่ 4 แสดง Item Information Curve ของข้อสอบและ Test Information Curve ของข้อสอบ 5 ข้อ (Lord 1980 : 22)



ทุก ๆ แบบสอบเมื่อใช้ในการสอบวัดย่อมมีความคลาดเคลื่อนในการวัดทุกครั้ง ซึ่งความคลาดเคลื่อนมาตรฐานในการประมาณค่า (Stardard Error of Estimate : S.E.E) คือส่วนเบี่ยงเบนมาตรฐานในการวัดของการประมาณค่าความสามารถ (θ) ในกลุ่มผู้สอบที่มีความสามารถแตกต่างกันไป เราสามารถคำนวณ S.E.E. ได้จากสูตร ดังนี้ (Warm 1978 : 77)

$$S.E.E. = \frac{1}{\sqrt{I(\theta)}}$$

ค่า TIF จะเปลี่ยนแปลงไปตามสเกลของความสามารถซึ่งถ้าค่า TIF มาก ค่า S.E.E. จะน้อยลง และค่า S.E.E. ที่น้อย ๆ จะเป็นค่าที่เราต้องการอย่างยิ่งในการสอบวัดแต่ละครั้ง อีกประการหนึ่งค่าเฉลี่ยของ S.E.E. (ซึ่งเขียนแทนด้วย $\overline{S.E.E.}$) ของผู้สอบทั้งหมดยังมีความสัมพันธ์กับค่าความเที่ยงของทฤษฎีการวัดแบบดั้งเดิม เมื่อคะแนนเป็นมาตรฐาน มีค่าความเบี่ยงเบนในการวัดเท่ากับ 1 และสูตรในการคำนวณเป็น ดังนี้ (Warm 1978 : 77)

$$r_{tt} = 1 - \overline{S.E.E.}^{2}$$

<u>โปรแกรม Logist 5 Version 2.5</u>

การประมาณค่าพารามิเตอร์ของข้อสอบแต่ละข้อ และค่าความสามารถของผู้สอบแต่ละคน ของ Three - Parameter Logistic Model โดยทั่วไปจะใช้คอมพิวเตอร์โปรแกรม Logistic ซึ่ง Wood, Wingersky และ Lord ได้ร่วมกันพัฒนาขึ้น และในการประมาณค่านั้นจะกระทำด้วยวิธีการประมาณค่าความน่าจะเป็นสูงสุด (Maximum Likelihook Method) ซึ่งวิธีนี้มีคุณสมบัติที่สำคัญ ดังนี้ (Hambleton and Swaminathan 1985 : 88-89)

1. มีความคงที่ ในขณะที่ขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบเพิ่มขึ้น การประมาณจึงเข้าใกล้ค่าที่แท้จริง

2. ฟังก์ชันของค่าสถิติมีความเพียงพอสำหรับ Information ทั้งหมดที่เกี่ยวกับพารามิเตอร์

3. มีประสิทธิภาพ กล่าวคือวิธีการประมาณค่าความน่าจะเป็นสูงสุดมีความแปรปรวนน้อยที่สุด

4. มีลักษณะของการแจกแจงแบบโค้งปกติ

โดยทั่วไป การประมาณค่าพารามิเตอร์ของข้อสอบแต่ละข้อ และค่าความสามารถของผู้สอบแต่ละคน มี 2 ขั้นตอน คือ ขั้นตอนแรก เป็นการประมาณค่าความสามารถของผู้สอบ N คน โดยใช้วิธีการปรับค่าความสามารถในการทำซ้ำแต่ละครั้ง จนมีค่าเฉลี่ยเป็น 0 และส่วนเบี่ยงเบนมาตรฐานเป็น 1 จากนั้นก็นำค่าความสามารถดังกล่าวนี้ไปประมาณค่าพารามิเตอร์ของข้อสอบจำนวน n ข้อ ในขั้นตอนที่ 2 ต่อไป อนึ่งในการประมาณค่าพารามิเตอร์และค่าความสามารถสำหรับโปรแกรม Logist 5 Version 2.5 นี้มีการทำงาน 4 ขั้นตอน ดังนี้

1. กำหนดให้ค่าอำนาจจำแนก (a) และค่าการเดา (c) ของข้อสอบทุกข้อเท่ากับ 1 และ 1/A - 0.05 เมื่อ A คือจำนวนตัวเลือกของข้อสอบตามลำดับ แล้วจึงประมาณค่าความสามารถของผู้สอบแต่ละคน และค่าความยาก (b) ของข้อสอบแต่ละข้อ

2. นำค่าความสามารถของผู้สอบแต่ละคนจากขั้นตอนที่ 1 ไปประมาณค่าอำนาจจำแนกของข้อสอบ ค่าความยากของข้อสอบ และค่าการเดาของข้อสอบ

3. นำค่าอำนาจจำแนก และค่าการเดาของข้อสอบจากขั้นตอนที่ 2 ไปประมาณค่าความสามารถของผู้สอบแต่ละคน และค่าความยากของข้อสอบแต่ละข้อ

4. นำค่าความสามารถของผู้สอบแต่ละคนจากขั้นตอนที่ 3 ไปประมาณค่าอำนาจจำแนกของข้อสอบ ค่าความยากของข้อสอบและค่าการเดาของข้อสอบแต่ละข้อ จนกระทั่งค่า θ, a, b, c มีค่าคงที่และมีความแม่นยำในการประมาณค่ากล่าวคือจะทำให้ฟังก์ชันไลค์ลิฮูด [$L(\theta, b, a, c)$] มีค่าสูงที่สุด ซึ่งหมายถึง ภาวะน่าจะเป็นสูงสุดที่ผู้สอบที่มีความสามารถ θ มีโอกาสตอบข้อสอบถูกต้อง $P(\theta)$ นั่นเอง

การประมาณค่าความสามารถและค่าพารามิเตอร์ในแต่ละขั้นตอนแสดงได้ดังตารางที่ 2 ต่อไปนี้

ตารางที่ 2 แสดงขั้นตอนการประมาณค่าความสามารถ และค่าพารามิเตอร์ของโปรแกรม Logist 5 Version 2.5

STEP	Parameter			
	Ability	a	b	c
1	estimated	fixed	estimated	fixed
2	fixed	estimated	estimated	estimated
3	estimated	fixed	estimated	fixed
4	fixed	estimated	estimated	estimated

<u>การนำ Three - Parameter Logistic Model ไปใช้ในการสอบและวัดผลทางการศึกษา</u>

Three - Parameter Logistic Model ได้ถูกนำมาใช้อย่างกว้างขวางในการวัดผลและการสอบ โดยเฉพาะอย่างยิ่งในปัจจุบันการพัฒนาการใช้คอมพิวเตอร์โปรแกรม Logist เพื่อให้ความสะดวกและมีประสิทธิภาพ ได้ถูกออกแบบมาอย่างได้ผลดี ซึ่ง Lord กล่าวว่า ค่าความสามารถ และค่าพารามิเตอร์ จะถูกประมาณค่าโดยโปรแกรม Logist และเครื่องจะคำนึงถึงเรื่องเกี่ยวกับการเว้น การทำไม่ทัน และการเดาอย่างสุ่มของผู้เข้าสอบด้วย สำหรับการนำ Three - Parameter Logistic Model ไปใช้ประโยชน์พอสรุปได้ดังนี้

1. <u>ใช้วิเคราะห์ข้อสอบ</u> (Item Analysis) เนื่องจากการวิเคราะห์ตามทฤษฎีการวัดแบบดั้งเดิม ยังมีข้อบกพร่องจำนวนมาก ซึ่งการวิเคราะห์ตามทฤษฎี IRT ด้วย Three - Parameter Logistic Model สามารถแก้ปัญหาได้

2. ใช้ในการสร้างธนาคารข้อสอบ (Item Bank) เนื่องจากข้อสอบที่วิเคราะห์แล้วค่าพารามิเตอร์ที่ได้มีลักษณะไม่แปรเปลี่ยน (invariant) ไปตามกลุ่มตัวอย่าง เราสามารถนำข้อสอบเหล่านี้ซึ่งได้มีการบันทึกค่า Item Information Function ของแต่ละข้อไว้แล้ว มาสร้างเป็นแบบสอบชุดใหม่ ๆ ตามจุดมุ่งหมายของการวัดได้

3. ใช้ในการกำหนดเกณฑ์ของระดับการรอบรู้ (Master Level) ของแบบสอบอิงเกณฑ์ โดยที่เราสามารถรู้ถึงระดับของ minimum master level ของข้อสอบแต่ละข้อที่เหมาะสมได้ ทั้งนี้เนื่องจากคะแนนของผู้สอบจากการวิเคราะห์ข้อสอบแล้ว จะถูกแปลงให้เป็นคะแนนที่สามารถ เทียบกับคะแนนความสามารถที่เป็นเกณฑ์ที่คงที่ได้

4. ใช้วินิจฉัยความสามารถของผู้สอบ ในกรณีที่มี ICC ของข้อสอบไม่เหมาะสม (fit) กับโมเดล แสดงว่ามีบางสิ่งบางอย่างผิดปกติในตัวผู้สอบ เช่น อาจมีความสามารถอื่นแฝงเข้ามาในความสามารถที่เราต้องการวัด

5. ใช้ในการวัดระดับความสามารถของแต่ละบุคคล เราสามารถสุ่มข้อสอบที่ได้รับการวิเคราะห์ และเก็บไว้เป็นธนาคารข้อสอบแล้ว และมีระดับความยากเรียงกันตามลำดับ เพียงจำนวนข้อไม่มากมาทดสอบความสามารถของแต่ละบุคคลได้

6. ใช้ตรวจความเป็นอคติของข้อสอบ (Item bias) เมื่อมีความสามารถอื่นแฝงเข้ามาในความสามารถที่ต้องการวัด เราสามารถทำการตรวจความเป็นอคติของข้อสอบได้จากการศึกษาความเหมาะสมของ ICC ของกลุ่มบุคคลที่มีความสามารถต่าง ๆ กัน

7. ใช้หาค่าการเดา ในทฤษฎีการวัดแบบดั้งเดิม มีการใช้สูตรแก้การเดา แต่ในทางปฏิบัติจริง การเดาแบบสุ่มน่าจะเกิดขึ้นน้อยมากนอกจากเวลาในการสอบไม่เพียงพอและตัวเลือกไม่ดี ซึ่งสูตรแก้การเดาเป็นเสมือนการลงโทษผู้สอบ แต่ในทฤษฎี IRT ไม่มีการใช้สูตรแก้การเดา แต่จะใช้การพิจารณาค่าการเดา ซึ่งหาได้จากการนำโมเดลทางคณิตศาสตร์ เข้ามาช่วยสนับสนุนทำให้หลักการพิจารณาค่าการเดามีความสมเหตุสมผลยิ่งขึ้น

8. ใช้ในการเปรียบเทียบคะแนนต่างชุด (Equating Score) ในแบบสอบที่วิเคราะห์แล้ว 2 ชุด ที่ต่างกันแต่วัดในสิ่งเดียวกัน จะสามารถนำคะแนนของผู้สอบ ในแบบสอบฉบับหนึ่ง เปรียบเทียบกับคะแนนของผู้สอบอีกฉบับหนึ่งได้

9. ใช้ในการจัดชั้นเรียน (Grade - Placement Tailoring) สามารถนำค่าความสามารถของผู้สอบที่ได้จากการวิเคราะห์แล้ว มาจัดชั้นเรียนให้เหมาะกับระดับความสามารถของผู้เรียน ตามแผนการจัดการสอน หรือตามนโยบายทางวิชาการได้

ตอนที่ 4 งานวิจัยที่เกี่ยวข้องกับโมเดลโลจีสติก

เจมส์ รี (Jame Ree 1981 : 11-19) ได้ศึกษาผลของการใช้ขนาดกลุ่ม-ตัวอย่าง และจำนวนข้อสอบทั้งหมด (item pools) ต่าง ๆ กัน ในการกำหนดขนาดของการสอบ (adaptive testing) โดยใช้โมเดลโลจีสติก 3 พารามิเตอร์ ซึ่งขนาดของกลุ่มตัวอย่างเป็น 500, 1,000, 2,000 ส่วนจำนวนข้อสอบทั้งหมด เป็น 100, 200 หรือ 300 กำหนดความยาวของ adaptive testing โดยให้มีข้อสอบ 10, 15, 20, 25, 30 และ 35 ทำการศึกษาโดยวิธี Simulation 500 ครั้งสำหรับการจัดจำนวนข้อสอบและขนาดกลุ่มตัวอย่างแต่ละครั้ง พบว่า ความสัมพันธ์ระหว่าง ค่าความสามารถ และค่าความสามารถจากการประมาณ จะสูงมากถ้าการสอบมีจำนวนข้อสอบที่เพียงพอ ค่าความคลาดเคลื่อนมาตรฐานของการทำนายค่าพารามิเตอร์ของข้อสอบ และของผู้สอบ สามารถทำให้น้อยลงได้ เมื่อมีข้อสอบอย่างน้อย 200 ข้อ และกลุ่มตัวอย่าง 2,000 คน

Huttun (1981 : 4799 - A) ได้ทดสอบความเหมาะสม (fit) ระหว่างแบบสอบกับราส์ชโมเดล และกับโมเดลโลจีสติก 3 พารามิเตอร์ โดยใช้แบบสอบ 4 ชุด ๆ ละ 40 ข้อ ทำการสอบกับโมเดลทั้งสองมีถึง 80 % และความเหมาะสมของแบบที่ใช้โมเดลโลจีสติก 3 พารามิเตอร์ มากกว่าจะใช้กับราส์ชโมเดล มี 65 % และยังพบด้วยว่า ความเหมาะสมกับโมเดล กับคุณสมบัติการวัดเพียงคุณลักษณะเดียว (Unidimensionality) มีความสัมพันธ์กันอย่างมีนัยสำคัญ แต่ถ้าจำนวนข้อสอบของแต่ละชุดของแบบสอบ ลดเหลือเพียงชุดละ 20 ข้อ พบว่าการประมาณค่าความสามารถของผู้สอบ จากข้อสอบ 40 ข้อ กับข้อสอบ 20 ข้อ ด้วยเทคนิคของราส์ชโมเดล มีความสัมพันธ์กัน ด้วยค่า .923 และผลจากการใช้เทคนิคของโมเดลโลจีสติก 3 พารามิเตอร์ จะมีความสัมพันธ์กัน ด้วยค่า .866 นอกจากนี้ ยังพบอีกว่า สำหรับโมเดลโลจีสติก 3 พารามิเตอร์ หากใช้วิเคราะห์กับกลุ่มตัวอย่างขนาดเล็ก 250 คน ก็พอที่จะหาค่าความยากได้ แต่ค่าพารามิเตอร์อื่น ๆ ไม่สามารถเชื่อถือได้ ดังนั้นในการใช้โมเดลโลจีสติก 3 พารามิเตอร์ จึงควรใช้กลุ่มตัวอย่าง 1,000 คนขึ้นไป

ประดิษฐ์ เรืองตระกูล (ประดิษฐ์ เรืองตระกูล : 2528) ได้พัฒนาแบบสอบวินิจฉัยวิชาคณิตศาสตร์สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 1 โดยใช้โมเดลโลจีสติก 3 พารา-มิเตอร์ ในการหาค่าอำนาจจำแนก ค่าความยาก และค่าการเดาของข้อสอบ และใช้เกณฑ์ที่เหมาะสมของค่าพารามิเตอร์ทั้งสามในการคัดเลือกข้อสอบ เครื่องมือที่ใช้ในการวิจัย คือ แบบสอบวินิจฉัยที่ผู้วิจัยสร้างขึ้น เป็นแบบเลือกตอบ 4 ตัวเลือก จำนวน 60 ข้อ กลุ่มตัว-อย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 1 จำนวน 1,913 คน วิเคราะห์ข้อมูล โดยการตรวจสอบคุณสมบัติการวัดเพียงคุณลักษณะเดียว วิเคราะห์ข้อสอบด้วยโปรแกรมโลจีส 5 ตรวจสอบคุณภาพแบบสอบด้านความเที่ยงตามแนวอิงเกณฑ์ ด้วยสูตรแบบ Binomial และความเที่ยงตามแนวทฤษฎี IRT นอกจากนี้ผู้วิจัยยังวิเคราะห์หาความตรงของแบบสอบอีกด้วย ผลการวิจัยได้แบบสอบวินิจฉัยที่มีคุณภาพ ดังนี้

1. มีคุณสมบัติการวัดเพียงคุณลักษณะเดียว

2. มีค่าอำนาจจำแนกของข้อสอบตั้งแต่ 0.19036 ถึง 2.00000 ค่าความยากของข้อสอบตั้งแต่ -3.75288 ถึง 1.90366 และค่าการเดาของข้อสอบตั้งแต่ 0.0 ถึง 0.5

3. มีค่าความเที่ยงตามแนวอิงเกณฑ์ของทั้งฉบับเป็น .9126 และความเที่ยงตามแนวทฤษฎี IRT พบว่าแบบสอบจะมีประสิทธิภาพสูงสุด เมื่อวัดกับผู้สอบที่มีระดับความสามารถปานกลาง คือค่า θ อยู่ในช่วง -1.5 ถึง 1.5

4. มีค่าความตรงซึ่งสามารถเชื่อถือได้

จากมโนทัศน์ของลำดับขั้นการเรียนรู้ด้านพุทธิพิสัย มโนทัศน์ของโมเดลโลจีสติกในทฤษฎี IRT พร้อมทั้งจากรายงานการวิจัยที่กล่าวมาทั้งหมด จะเห็นได้ว่า การศึกษาความเป็นลำดับขั้นการเรียนรู้ของสมรรถภาพด้านต่าง ๆ ของกลุ่มพุทธิพิสัยตามแนวคิดของบลูมและคณะ โดยการประยุกต์กับทฤษฎี IRT ที่ใช้โมเดลโลจีสติก 3 พารามิเตอร์ นั้นยังไม่มีผู้ใดทำการวิจัย จะมีก็แต่เพียงการศึกษาในแนวทฤษฎีการวัดแบบดั้งเดิมเท่านั้น จากสาเหตุดังกล่าวนี้ผู้วิจัยจึงสนใจที่จะศึกษาถึงความเป็นลำดับขั้นของสมรรถภาพทั้ง 6 ด้าน ในกลุ่มพุทธิพิสัยนี้ โดยการหาความสัมพันธ์ระหว่างลำดับขั้นการเรียนรู้ กับลำดับที่ของค่าพารามิเตอร์ของข้อสอบ ซึ่งได้แก่ค่าความยาก ค่าอำนาจจำแนก และค่าการเดา ที่ประมาณตามโมเดล

โลจีสติก ในแบบสอบวิชาคณิตศาสตร์ นอกจากนี้ยังศึกษาอีกว่า ค่าสารสนเทศของกลุ่มข้อสอบที่วัดในแต่ละลำดับขั้นการเรียนรู้ด้านพุทธิพิสัยนี้ จะมีค่าสูงสุด เมื่อผู้สอบมีระดับความสามารถอย่างไร

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย