บทที่ 1

บทนำ

ความเป็นมาของปัญหา

การวัดผลการศึกษาเป็นเครื่องมืออันหนึ่งที่จะช่วยพัฒนาคุณภาพของการศึกษาในระดับการศึกษาต่างๆ เพราะผลจากการวัดจะเป็นพื้นฐานในการตัดสินใจของครูและนักการศึกษา เพื่อใช้ปรับปรุงวิธีการสอน การแนะแนว การประเมินผลหลักสูตร แบบเรียน การใช้อุปกรณ์การสอน ตลอดจนการจัดระบบบริหารทั่วไปของโรงเรียนและนอกจากนี้ยังช่วยปรับปรุงการเรียนการสอนของนักเรียนให้เรียนถูกวิธียิ่งขึ้น เช่น ผลการสอบรวมทั้งชั้นจะแสดงถึงการสอนไม่ดีของครู และถ้าพิจารณาผลการสอบรวมทั้งโรงเรียนก็จะแสดงถึงความบกพร่องในทางบริหารโรงเรียนของครูใหญ่ยิ่งกว่านั้นถ้าเราพิจารณาผลการสอบรวมทั้งประเทศก็จะแสดงถึงความบกพร่องในการจัดระบบการบริหารของนักการศึกษาระดับประเทศอีกด้วย ดังนั้นการวัดผลการศึกษาจึงนับว่ามีความสำคัญยิ่งในการพัฒนาคุณภาพของการศึกษา (อนันต์ ศรีโสภา. 2524. หน้า 1)

ในการวัดผลทางการศึกษาและจิตวิทยานั้น สิ่งที่ต้องการจะวัดและตรวจสอบมักเป็นคุณลักษณะ (Traits) ต่างๆ ที่เป็นคุณลักษณะภายใน (Latent traits) ที่ไม่สามารถสังเกตได้โดยตรง เช่น ความสามารถในการคำนวณ ทักษะในการอ่านจับใจความ เป็นต้น ข้อสอบเป็นเพียงสิ่งเร้าที่กระตุ้นให้บุคคลแสดงพฤติกรรมออกมาโดยการตอบข้อสอบ (Response) เหล่านั้น แล้วพิจารณาว่าบุคคลนั้นทำข้อสอบได้ถูกต้องกี่ข้อได้กี่คะแนนต่อจากนั้นก็จะอนุมาน (Infer) จากจำนวนข้อที่ถูกโยงไปหาความสามารถที่แท้จริงซึ่งเป็นคุณลักษณะภายในของผู้สอบ (สงบ ลักษณะ. 2525. หน้า 47)

แบบสอบที่ใช้วัดคุณลักษณะภายในของผู้สอบเท่าที่ผ่านมามักจะยึดทฤษฎีการวัดแบบดั้งเดิม (Classical Test Theory : CTT) แบบทดสอบจะประกอบด้วยข้อทดสอบจำนวนมาก เพื่อให้การทดสอบมีความแม่นตรงและความเชื่อมั่นสูง ข้อสอบเหล่านั้นจะถูกเรียงลำดับไว้ก่อนแล้วตามเนื้อหาที่ต้องการทดสอบ และมักจะประกอบด้วยข้อทดสอบที่มีระดับความยากในระดับปานกลางเพื่อให้เหมาะสมกับระดับความสามารถทั่วไปของผู้สอบ ซึ่งการกระทำดังกล่าวนี้ทำให้เกิดความคลาดเคลื่อนในการวัดสำหรับผู้สอบที่มีความรู้ในระดับเก่งหรืออ่อน เพราะข้อสอบที่ง่ายเกินไปจะไม่ท้าทายความสามารถในการสอบของผู้ที่เก่งทำให้ผู้สอบเกิดความเบื่อหน่ายในการ

ตอบ ทำให้ผู้สอบเกิดความเบื่อหน่ายในการตอบอย่างไม่เต็มความสามารถ ส่วนข้อสอบที่ยากเกินไปสำหรับผู้สอบที่อ่อนก็จะทำให้ผู้สอบรู้สึกท้อถอยและใช้วิธีเดา สิ่งต่างๆ เหล่านี้ทำให้ความถูกต้องและแม่นยำในการทดสอบลดลงและทำให้การประมาณค่าความสามารถของผู้สอบคลาดเคลื่อน ดังนั้นการทดสอบแบบดั้งเดิมจึงใช้ได้ผลดีเฉพาะกรณีที่ผู้สอบมีความสามารถระดับปานกลางเท่านั้น และความคลาดเคลื่อนจะมีมากขึ้นเมื่อความสามารถของผู้สอบเบี่ยงเบนจากความสามารถปานกลางมากขึ้น (สุพัฒน์ สุกมลสันต์. 2538. หน้า 197)

การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ (Adaptive Testing) มีความแตกต่างที่สำคัญจากการทดสอบแบบดั้งเดิมตรงที่การทดสอบแบบดั้งเดิมเป็นการสร้างแบบสอบขึ้นมาชุดเดียวให้ครอบคลุมองค์ประกอบของคุณลักษณะหรือเนื้อหาที่ต้องการวัด มีค่าความยากของข้อสอบที่หลากหลาย โดยไม่มีการกำหนดสัดส่วนของค่าความยากที่แน่นอน โดยทั่วไปนิยมให้มีค่าความยาก (p) อยู่ระหว่าง 0.2-0.8 สำหรับนำไปใช้สอบกับผู้สอบกลุ่มใดก็ได้ ไม่ว่าผู้สอบจะมีความสามารถสูง ปานกลาง หรือต่ำ ทุกคนจะต้องทำข้อสอบชุดเดียวกันนั้น มีจำนวนข้อเท่ากันและเหมือนกันหมดทุกข้อ แต่การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบเป็นการใช้ข้อสอบจากคลังข้อสอบมาสร้างแบบสอบ แบบสอบแต่ละชุดมีการออกแบบให้เหมาะสมกับความสามารถของผู้สอบแต่ละคน การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบจะช่วยลดจำนวนข้อสอบและเวลาของการสอบ นอกจากนี้ยังช่วยผ่อนคลายความเครียดของการสอบอีกด้วย เช่น ในกรณีของผู้สอบที่มีความสามารถต่ำไม่จำเป็นต้องเครียดกับการทำข้อสอบที่ยากเกินความสามารถในขณะเดียวกันผู้สอบที่มีความสามารถสูงก็ไม่จำเป็นต้องทำข้อสอบที่ง่ายเกินไปจนอาจมีผลให้เกิดการขาดความระมัดระวังในการทำข้อสอบข้อที่ยากขึ้น เป็นต้น ข้อสอบข้อที่ง่ายเกินไปหรือยากเกินไปเมื่อเทียบกับระดับความสามารถของผู้สอบย่อมไม่ค่อยให้สารสนเทศ (information) ที่เป็นประโยชน์มากนักสำหรับผู้สอบ ในขณะที่การคัดเลือกข้อสอบที่มีระดับความยากพอเหมาะกับผู้สอบตามทฤษฎีการตอบสนองข้อสอบ (IRT) ย่อมให้สารสนเทศที่สำคัญเกี่ยวกับความสามารถที่แท้จริงของผู้สอบ ดังนั้นข้อสอบแต่ละข้อจากคลังข้อสอบที่นำมาใช้แต่ละคนจึงถูกเลือกหรือเสมือนเป็นการปรับให้เหมาะสมกับความสามารถของผู้สอบขณะที่ทำการทดสอบ ผู้สอบแต่ละคนจึงทำข้อสอบแตกต่างกัน และจำนวนข้อของการสอบไม่จำเป็นต้องเท่ากัน ในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ หลักการคัดเลือกข้อสอบสำหรับแต่ละบุคคลอยู่บนพื้นฐานของผลการตอบข้อสอบข้อที่ผ่านมาของผู้สอบนั้น เมื่อผู้สอบทำข้อสอบข้อเริ่มต้นหรือชุดแรก (ขึ้นอยู่กับการออกแบบ) จากคลังข้อลอบแล้ว จะมีการวิเคราะห์ระดับความสามารถหรือประเมินความสามารถของผู้สอบเบื้องต้น เพื่อคัดเลือกข้อสอบข้อต่อไปที่มีค่าความยากและ

อำนาจจำแนกเหมาะสมที่จะใช้วัดระดับความสามารถของผู้สอบ ประมาณระดับความสามารถของผู้สอบใหม่ จากนั้นก็จะเลือกข้อสอบที่เหมาะสมข้อต่อไป โดยอาศัยหลักที่ว่า ถ้าการทำข้อสอบที่ผ่านมาถูก ข้อถัดไปจะยากขึ้น แต่ถ้าทำข้อสอบที่ผ่านมาผิด ข้อถัดไปจะง่ายลง กระบวนการนี้จะดำเนินการต่อไปเรื่อยจนสามารถประมาณระดับความสามารถของผู้สอบได้อย่างน่าเชื่อถือ (ความคลาดเคลื่อนต่ำ) ตามเกณฑ์ที่กำหนดไว้ การทดสอบก็จะยุติลง (ศิริชัย กาญจนวาสี. 2545. หน้า 177)

การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบมีความยุ่งยากในการดำเนินการสอบเมื่อใช้กับการทดสอบที่ใช้กระดาษกับดินสอ (Paper and Pencil Test) จึงได้มีการนำเอาคอมพิวเตอร์มาประยุกต์ใช้กับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ ซึ่งเรียกว่า การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบโดยใช้คอมพิวเตอร์หรือคอมพิวเตอร์ไรซ์อแดปตีป เทสติ้ง (Computerized Adaptive Testing : CAT) โดยให้คอมพิวเตอร์เป็นผู้ประมาณค่าความสามารถของผู้สอบและจัดทำข้อสอบให้เหมาะสมกับระดับความสามารถของผู้สอบ (ต่าย เชี่ยงฉี. 2534. หน้า 4 อ้างอิงจาก Weiss. 1982. p 473-492) นอกจากนั้นยังได้นำเอาทฤษฎีการตอบสนองข้อสอบ (Item Response Theory : IRT) มาใช้ในการประมาณค่าความสามารถของผู้สอบจึงทำให้การประมาณค่าความสามารถของผู้สอบได้เที่ยงตรง (Validity)และเชื่อถือได้ (Reliability) ยิ่งขึ้น

การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบโดยใช้คอมพิวเตอร์ (CAT) ให้ประสิทธิภาพสูงกว่าการทดสอบโดยการเขียนตอบแบบเดิมหลายประการ ดังนี้

1. ใช้เวลาในการสอบน้อย เนื่องจากการทดสอบเทเลอร์ใช้จำนวนข้อสอบน้อยกว่าการทดสอบแบบดั้งเดิม ดังผลวิจัยของ นันทิยา พึ่งคำ (2531) พบว่า การทดสอบโดยใช้การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบโดยใช้คอมพิวเตอร์ (CAT) สามารถลดจำนวนข้อสอบลงได้เกินกว่าครึ่งหนึ่งโดยยังคงให้ค่าความตรงเท่าเดิมซึ่งก็สอดคล้องกับผลการวิจัยของไวส์ (Weiss.1982. pp 473-492) และโมรีโนและคณะ (Moreno et al. 1984. pp 155-163)

2. สามารถรายงานผลได้ทันที ผู้สอบเพียงแต่นั่งหน้าจอคอมพิวเตอร์เมื่อคำถามขึ้นที่จอคอมพิวเตอร์แล้วผู้สอบก็ตอบคำถามลงไปทางแป้นพิมพ์ เครื่องคอมพิวเตอร์จะตรวจสอบว่าคำตอบนั้นถูกต้องหรือไม่แล้วก็ประมาณค่าความสามารถของผู้สอบพร้อมทั้งเลือกข้อสอบข้อต่อไปที่เหมาะสมกับระดับความสามารถของผู้สอบออกมาทางจอคอมพิวเตอร์เพื่อให้ผู้สอบทำต่อไป การทดสอบนี้จะสิ้นสุดลงเมื่อมีความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบถึงระดับที่กำหนดไว้ (Urry. 1977. p 194 ; Linden & Zwarts. 1989. p 182 ; ฉัตรนภา

พรหมมา. 2525 หน้า 46) เครื่องคอมพิวเตอร์จะใช้เวลาน้อยมากในการตรวจคำตอบ การประมาณค่าความสามารถของผู้สอบและการค้นหาข้อสอบที่เหมาะสมกับระดับความสามารถของผู้สอบ หลังจากสิ้นสุดการทดสอบแล้วเครื่องคอมพิวเตอร์ยังสามารถที่จะรายงานผลการสอบได้ทันทีอีกด้วย

3. เครื่องคอมพิวเตอร์สามารถบันทึกผลการสอบได้โดยในขณะที่ผู้สอบทำข้อสอบอยู่นั้น เครื่องคอมพิวเตอร์สามารถที่จะเก็บบันทึกผลการสอบไว้ให้ได้ด้วย

4. ผู้สอบที่มีความสามารถสูงไม่เกิดความเบื่อหน่ายเพราะไม่ต้องตอบข้อสอบที่ง่ายๆ และผู้สอบที่มีความสามารถต่ำก็ไม่เกิดความเคร่งเครียดเพราะไม่ต้องพบกับข้อสอบที่ยากๆ จึงทำให้ผู้สอบมีทัศนคติที่ดีต่อการสอบเทเลอร์ ซึ่งจะเห็นจากผลการวิจัยของ งามนิตย์ ธาตุทอง และปรีชา เครือวรรณ (2527) และ ลี (Lee. 1987 : 3742-A) ที่พบว่า ผู้สอบส่วนใหญ่มีทัศนคติที่ดีต่อการทดสอบเทเลอร์

5. เป็นการสอบที่ยั่วยุให้ผู้ตอบคิดเพราะ มีการปรับความยากของข้อสอบให้เหมาะกับความสามารถของผู้สอบอยู่ตลอดเวลา

6. สามารถทดสอบกับนักเรียนเป็นรายบุคคลได้โดยไม่ต้องรอสอบพร้อมๆ กัน นักเรียนคนใดพร้อมที่จะสอบก็เปิดเครื่องคอมพิวเตอร์แล้วทดสอบได้ทันที ซึ่งจะเป็นประโยชน์อย่างยิ่งต่อการจัดการเรียนการสอนในด้านการปรับปรุงการเรียนการสอน

7. ช่วยรักษาความลับและป้องกันการทุจริตในการสอบเป็นอย่างดี เพราะข้อสอบทั้งหมดถูกเก็บไว้ในลักษณะของคลังข้อสอบ ซึ่งไม่สามารถเปิดดูได้อย่างง่ายๆ เหมือนการพิมพ์ข้อสอบไว้เป็นเอกสารเหมือนแต่เดิม สำหรับกรณีการทุจริตในการสอบนั้นสามารถป้องกันได้ เพราะผู้สอบไม่สามารถลอกคำตอบจากเพื่อนได้ถึงแม้ว่าจะนั่งใกล้กันก็ตามทั้งนี้เพราะข้อสอบที่หน้าจอคอมพิวเตอร์นั้นไม่ใช่ข้อสอบข้อเดียวกัน

8. ประหยัดค่าใช้จ่าย การทดสอบลักษณะนี้จะมีค่าใช้จ่ายสูงในตอนเริ่มแรก กล่าวคือจำเป็นต้องมีเครื่องคอมพิวเตอร์เพื่อใช้เก็บข้อสอบและดำเนินการสอบ แต่เมื่อเปรียบเทียบระยะเวลากันนานๆ แล้ว การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบโดยใช้คอมพิวเตอร์จะเสียค่าใช้จ่ายน้อยกว่าเพราะไม่จำเป็นต้องสร้างข้อสอบใหม่ทุกครั้ง แต่การทดสอบแบบเดิมนั้นต้องสร้างแบบทดสอบขึ้นมาใหม่เกือบทุกครั้งที่มีการทดสอบ

9. สามารถกำหนดความถูกต้องในการประมาณค่าความสามารถของผู้สอบ โดยกำหนดค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบเป็นเกณฑ์ยุติการสอบ ถ้ากำหนดค่าความคลาดเคลื่อนมาตรฐานไว้สูงความถูกต้องในการประมาณค่าก็จะต่ำ แต่ถ้า

กำหนดค่าความคลาดเคลื่อนมาตรฐานไว้ต่ำความถูกต้องในการประมาณค่าก็จะสูง ซึ่งต่างจากแบบทดสอบแบบเดิมที่ผู้ดำเนินการสอบไม่สามารถกำหนดค่าความถูกต้องในการวัดล่วงหน้าก่อนการทดสอบได้

10. เพิ่มค่าความตรงและความเที่ยงของการทดสอบได้ เพราะการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบจะตัดข้อสอบที่ไม่เหมาะสมกับความสามารถของผู้สอบออกไป ซึ่งการทดสอบด้วยข้อสอบที่มีความเหมาะสมกับความสามารถของผู้สอบจะช่วยให้ความตรงและความเที่ยงของการวัดมีค่าสูงขึ้น (Larkin & Weiss, 1974. p. 1 ; Weiss, 1984. p. 362)

จากที่กล่าวมาจะเห็นว่า การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบโดยใช้คอมพิวเตอร์ (CAT) สามารถนำมาใช้สอบแทนการทดสอบแบบเดิมที่ใช้กระดาษกับดินสอ (Paper and Pencil Test) ได้เป็นอย่างดีและมีประสิทธิภาพสูงกว่าการทดสอบแบบเดิมหลายประการดังกล่าวแล้ว แต่การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบโดยใช้คอมพิวเตอร์ (CAT) ยังเป็นแบบเครื่องเดียว (Stand Alone) ซึ่งจะทำงานเป็นอิสระต่อกันเกือบจะโดยสิ้นเชิง การนำข้อมูลป้อนเข้าเครื่องคอมพิวเตอร์จะต้องป้อนข้อมูลกับคอมพิวเตอร์ทุกเครื่องทำให้ต้องใช้แรงงานเพิ่มขึ้นสำหรับผู้ที่เกี่ยวข้องกับคลังข้อสอบและกระทำได้ในขณะที่ไม่มีผู้สอบเท่านั้นทำให้เสียเวลาและไม่สะดวกในการปรับปรุงข้อมูล ในส่วนของผู้เข้าสอบสามารถเข้าสอบได้เฉพาะวันทำการและข้อมูลของแต่ละคนจะถูกจัดเก็บหรือพิมพ์ออกมาสำหรับผู้สอบเป็นรายบุคคลเท่านั้นไม่สามารถเห็นผลการสอบรวมได้ จากผลงานวิจัยที่เกี่ยวข้องยังไม่พบว่ามีผู้ใดศึกษาเรื่องการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบโดยใช้คอมพิวเตอร์แบบมีระบบเครือข่ายอินเทอร์เน็ต (Internet) มาก่อน ด้วยเหตุนี้ผู้วิจัยจึงได้สนใจที่จะพัฒนาโปรแกรมคอมพิวเตอร์สำหรับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบบนระบบเครือข่ายอินเทอร์เน็ต

จุดมุ่งหมายของการวิจัย

ในการวิจัยครั้งนี้ ผู้วิจัยมีวัตถุประสงค์หลักเพื่อพัฒนาโปรแกรมคอมพิวเตอร์สำหรับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบบนระบบเครือข่ายอินเทอร์เน็ต

โดยมีวัตถุประสงค์ย่อย ดังนี้

1. เพื่อสร้างและหาประสิทธิภาพของโปรแกรมคอมพิวเตอร์สำหรับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบบนระบบเครือข่ายอินเทอร์เน็ต

2. เพื่อทดลองใช้โปรแกรมคอมพิวเตอร์สำหรับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบบนระบบเครือข่ายอินเทอร์เน็ต

3. เพื่อประเมินโปรแกรมคอมพิวเตอร์สำหรับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบบนระบบเครือข่ายอินเทอร์เน็ต

ความสำคัญของการวิจัย

ในการวิจัยครั้งนี้ทำให้ได้

1. โปรแกรมคอมพิวเตอร์สำหรับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบที่ใช้ได้กับคอมพิวเตอร์แบบเครื่องเดียว (Stand Alone) ระบบเครือข่ายเฉพาะบริเวณ (LAN) และระบบเครือข่ายอินเทอร์เน็ต (Internet)

2. เป็นคลังข้อสอบเพื่อใช้ในการวัดและประเมินผลความสามารถของผู้สอบ เพื่อเป็นการเตรียมตัวก่อนที่จะใช้ข้อสอบกลางของสถาบันทดสอบทางการศึกษาแห่งชาติ

3. เป็นแนวทางในการวัดและประเมินผลการศึกษาในอนาคต เพราะสามารถพิจารณาผลการสอบรวมระดับชั้นเรียน ผลการสอบรวมทั้งโรงเรียน และผลการสอบรวมทั้งประเทศได้อย่างรวดเร็วและสามารถตรวจสอบคุณภาพการศึกษาได้ตลอดเวลา

ขอบเขตของการวิจัย

ในการวิจัยครั้งนี้ได้กำหนดขอบเขตในการศึกษาไว้ดังนี้

1. กลุ่มตัวอย่างที่ใช้ในการทดลองใช้และประเมินโปรแกรม เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 1 ภาคเรียนที่ 2 ปีการศึกษา 2548 โรงเรียนเทศบาล 3 (ชาญวิทยา) และโรงเรียนหล่มเก่าพิทยาคม จำนวน 193 คน และ 204 คน ตามลำดับ รวมทั้งสิ้น 397 คน เป็นครูโรงเรียนละ12 คน รวมทั้งสิ้น 24 คน

2. การวิจัยครั้งนี้ใช้ภาษา ASP.NET ในการพัฒนาโปรแกรมคอมพิวเตอร์สำหรับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบบนระบบเครือข่ายอินเทอร์เน็ต และคลังข้อสอบใช้ฐานข้อมูล Microsoft Access

3. คลังข้อสอบ ใช้ข้อสอบวัดความเข้าใจ เรื่องเศษส่วน ชั้นประถมศึกษาปีที่ 6 จำนวน 200 ข้อ สร้างโดย รังสรรค์ มณีเล็ก (2540) ข้อสอบดังกล่าวมีค่าพารามิเตอร์ของข้อสอบตามทฤษฎีการตอบสนองข้อสอบ

นิยามศัพท์เฉพาะ

1. **การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบโดยใช้คอมพิวเตอร์** (Computerized Adaptive Testing : CAT) **หมายถึง การทดสอบที่ให้คอมพิวเตอร์เป็นผู้ประมาณ**ค่าความสามารถของผู้สอบและคัดเลือกข้อสอบให้เหมาะสมกับความสามารถของผู้สอบแต่ละคน โดยอาศัยทฤษฎีการตอบสนองข้อสอบ (Item Response Theory : IRT) มาใช้เป็นพื้นฐานของการดำเนินการสอบ และยุติการทดสอบเมื่อค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบต่ำกว่าเกณฑ์ที่กำหนด

2. โปรแกรมคอมพิวเตอร์สำหรับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบบนระบบเครือข่ายอินเทอร์เน็ต (Computerized Adaptive Testing On Internet) หมายถึง โปรแกรมคอมพิวเตอร์สำหรับการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบบนระบบเครือข่ายอินเทอร์เน็ต ใช้รูปแบบแยกทางแบบแปรผัน (Variable Branching Model) เป็นรูปแบบการคัดเลือกข้อสอบที่ไม่แน่นอนขึ้นอยู่กับการคำนวณค่าความสามารถของผู้สอบ ซึ่งประมาณค่าความสามารถของผู้สอบโดยวิธีของเบส์ เมื่อผู้สอบตอบข้อสอบแต่ละข้อแล้วจะทำการคัดเลือกข้อสอบที่มีความยากใกล้เคียงกับค่าความสามารถของผู้สอบที่คำนวณได้ครั้งล่าสุดซึ่งจะทำให้ได้ข้อสอบที่มีความเหมาะสมกับความสามารถของผู้สอบมากยิ่งขึ้น โดยการทดสอบจะดำเนินไปเรื่อยๆ จนค่าความคลาดเคลื่อนมาตรฐานในการทดสอบที่คำนวณได้เท่ากับหรือน้อยกว่าเกณฑ์ที่กำหนดคือ 0.3 จึงยุติการทดสอบหรือกดปุ่มจบการทดลอบ โปรแกรมนี้ติดตั้งไว้ที่เว็บไซต์ http://www.adapttest-nu.com

3. ทฤษฎีการตอบสนองข้อสอบ (Item Response Theory : IRT) หมายถึง ทฤษฎีที่อธิบายความสัมพันธ์ระหว่างความสามารถของผู้สอบกับลักษณะของข้อสอบ 3 ประการ คือ ค่าอำนาจจำแนก (a) ค่าความยาก (b) และค่าการเดา (c)

4. ค่าอำนาจจำแนก (a - parameter) หมายถึง คุณลักษณะของข้อสอบในการจำแนกผู้สอบที่มีความสามารถแตกต่างกัน ซึ่งสามารถประมาณค่าได้จากความชัน (slope) ของโค้ง**คุณลักษณะข้อสอบ ณ จุดเปลี่ยนโค้ง** (inflexion point)

5. **ค่าความยาก** (b- parameter) **หมายถึง คุณลักษณะของข้อสอบที่แสดงถึงความยากง่ายของข้อสอบ เป็นค่าที่แสดงถึงระดับความสามารถที่จุดเปลี่ยนโค้ง ซึ่งสามารถประมาณค่าได้จากระดับความสามารถของผู้สอบที่มีความน่าจะเป็นของการตอบข้อสอบข้อนั้นถูก 50 เปอร์เซ็นต์ สำหรับโมเดลหนึ่งและสองพารามิเตอร์ และ (1+c)/2 เปอร์เซ็นต์ สำหรับโมเดลสามพารามิเตอร์**

6. ค่าการเดา (c- parameter) หมายถึง คุณลักษณะของข้อสอบที่แสดงถึงความน่าจะเป็นของผู้สอบที่มีความสามารถต่ำมากจะมีโอกาสตอบข้อสอบนั้นได้ถูกต้อง ซึ่งสามารถประมาณค่าได้จากค่ากำกับต่ำสุดของเส้นโค้งคุณลักษณะข้อสอบ

7. ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถ [Standard Error :SE(θ)] หมายถึง ค่าเบี่ยงเบนที่ได้จากการประมาณค่า ในการวิจัยครั้งนี้สามารถคำนวณได้จากการถอดรากที่สองของค่าความแปรปรวนในการทดสอบ โดยกำหนดให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบเท่ากับหรือน้อยกว่า 0.30 จึงยุติการสอบ

8. ประสิทธิภาพของโปรแกรม หมายถึง โปรแกรมสามารถทำงานได้ถูกต้องตามหลักการทำงานของการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบโดยใช้คอมพิวเตอร์ พิจารณาได้จากการทดลองสร้างแบบแผนการตอบข้อสอบของผู้สอบ 3 แบบ คือ แบบถูกทุกข้อ แบบผิดทุกข้อ และแบบถูก-ผิด สถานการณ์ไม่คงที่ และพิจารณาว่า โปรแกรมสามารถคัดเลือกข้อสอบข้อแรกที่มีความยากปานกลางหรือไม่ โปรแกรมสามารถประมาณค่าความสามารถของผู้สอบในกรณีตอบถูกหรือตอบผิดได้ถูกต้องหรือไม่ และยุติการทดสอบเมื่อค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบเท่ากับหรือน้อยกว่า 0.30 หรือไม่ โดยการเปรียบเทียบผลการคำนวณด้วยมือกับการคำนวณที่ได้จากโปรแกรม

9. คุณภาพของโปรแกรม หมายถึง ความถูกต้องในการทำงานของโปรแกรมย่อยผู้ดูแลระบบโปรแกรมย่อยการทดสอบ และโปรแกรมย่อยผลการสอบ พิจารณาได้จาก ทุกโปรแกรมย่อยมีความทนทานต่อความผิดพลาดและมีระบบป้องกันความผิดพลาดที่อาจจะเกิดขึ้น โปรแกรมสามารถทำงานต่อไปได้หลังจากแก้ไขความผิดพลาดแล้ว รวมถึงมีคำแนะนำให้ผู้ทดลองใช้สามารถใช้งานได้ถูกต้องด้วย