

การวิจัยทักษะการอ่านภาษาอังกฤษโดยประยุกต์โมเดลลำดับขั้นคุณลักษณะ และการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

ทิพย์ ขำอยู่ และ เสรี ชัดแจ้ง

ศูนย์การวัดผลประยุกต์ วิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา มหาวิทยาลัยบูรพา

กฤษณะ ชินสาร

คณะวิทยาการสารสนเทศ มหาวิทยาลัยบูรพา

บทคัดย่อ

การวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาวิธีการวิจัยทักษะการอ่านภาษาอังกฤษ โดยประยุกต์โมเดลลำดับขั้นคุณลักษณะและใช้การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับวินิจฉัยพื้นฐานความรู้ทักษะการอ่านภาษาอังกฤษของนักศึกษาระดับปริญญาตรี กลุ่มตัวอย่างเป็นนักศึกษาระดับปริญญาตรี มหาวิทยาลัยราชภัฏสวนดุสิต ปีการศึกษา 2555 จำนวน 1,763 คน เครื่องมือที่ใช้ในการวิจัย ได้แก่ ชุดข้อสอบวินิจฉัยทักษะการอ่านภาษาอังกฤษ และโปรแกรมการทดสอบวินิจฉัยแบบปรับเหมาะด้วยคอมพิวเตอร์

ผลการวิจัยปรากฏว่า วิธีการวินิจฉัยทักษะการอ่านภาษาอังกฤษบนโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่พัฒนาขึ้น สามารถใช้ประโยชน์จากแนวคิดโมเดลลำดับขั้นคุณลักษณะซึ่งเป็นแนวคิดของทฤษฎีทางปัญญา และทฤษฎีการตอบสนองข้อสอบสำหรับการทดสอบเพื่อวินิจฉัยพื้นฐานความรู้ด้านทักษะการอ่านภาษาอังกฤษของนักศึกษาระดับปริญญาตรีได้ ผลการประเมินโดยผู้เชี่ยวชาญและผู้ใช้งานมีความสอดคล้องกันคือ วิธีการวินิจฉัยที่พัฒนาขึ้นมีความเหมาะสม สามารถนำไปใช้ประโยชน์สำหรับการทดสอบเพื่อวินิจฉัยพื้นฐานความรู้ด้านทักษะการอ่านภาษาอังกฤษของนักศึกษาได้

คำสำคัญ: วิธีการวินิจฉัยทักษะการอ่านภาษาอังกฤษ, โมเดลลำดับขั้นคุณลักษณะ, การทดสอบวินิจฉัยแบบปรับเหมาะด้วยคอมพิวเตอร์

Diagnostic of English Reading Skills by Applying the Attribute Hierarchy Model and Computerized Adaptive Testing

Thip Khumyoo and Seree Chadcham

Centre for Applied Measurement

College of Research Methodology and Cognitive Science

Burapha University, Thailand

Krisana Chinnasarn

Faculty of Informatics, Burapha University, Thailand

Abstract

By applying the attribute hierarchy model and using computerized adaptive testing, this research aimed to develop a diagnostic English reading skills procedure allowing us to detect the knowledge state of undergraduate students. The sample used to construct the diagnostic test comprised 1,763 undergraduate students from Suan Dusit Rajabhat University (academic year 2012). Research instruments included Diagnostic English Reading Skill Tests and Computerized Adaptive Diagnostic Testing Software.

The result indicated that the development of a diagnostic procedure of English reading skills using computerized adaptive testing, make it possible to integrate the attribute hierarchy model for cognitive testing and the Item Response theory in order to detect the knowledge state of undergraduate students in English reading skills from observed response patterns.

Keywords: diagnostic English reading skill procedure, attribute hierarchy model, computerized adaptive diagnostic testing

ความนำ

การวินิจฉัย (Diagnostic) เป็นกระบวนการสำคัญในการวัดและการประเมินผลการศึกษา เป็นการเชื่อมโยงและสร้างความสอดคล้องระหว่างองค์ประกอบของการจัดการศึกษา ทั้งการบริหารจัดการการเรียนรู้และการวินิจฉัยจากการวัดและการประเมินผลการศึกษา แนวทางการวินิจฉัยแบ่งออกเป็น 2 แนวทาง แนวทางแรกคือการวัดสิ่งที่ขาดหายไป (Deficit measurement) เป็นการวินิจฉัยจุดด้อยหรือความบกพร่อง (Weakness) นิยมใช้ในการวินิจฉัยทางการแพทย์เพื่อจำแนกผู้ป่วยตามอาการ แนวทางที่สองคือการวิเคราะห์ความคลาดเคลื่อน (Error analysis) เป็นการวิเคราะห์ความบกพร่องของผู้เรียนจากแบบแผนการตอบที่ผิดพลาด เพื่อวินิจฉัยความบกพร่องทางการเรียนรู้ของผู้เรียน โดยผู้เชี่ยวชาญทางด้านเนื้อหานั้น ๆ เป็นผู้วินิจฉัย ต่อมาได้มีการนำพื้นฐานการวัดทางจิตวิทยามาช่วยในการวินิจฉัย (Bejar, 2005) และมีการนำมาใช้กันอย่างแพร่หลายถึงทุกวันนี้

การวินิจฉัยจะต้องให้สารสนเทศที่ถูกต้องและเป็นประโยชน์ต่อการบริหารจัดการการเรียนรู้และพัฒนา ศักยภาพผู้เรียนให้มีประสิทธิภาพสูงสุด การดำเนินการที่ชัดเจนและให้ผลการวินิจฉัยที่เป็นประโยชน์คือ การใช้แบบทดสอบ ซึ่งเป็นแบบทดสอบที่สร้างขึ้นเพื่อการวินิจฉัยโดยตรงเรียกว่าแบบทดสอบวินิจฉัย (Diagnostic test) โดยใช้วิธีการวิเคราะห์ความคลาดเคลื่อนจากคะแนนสอบ (Test score) และจากแบบแผนการตอบข้อสอบ (Item response pattern) ที่ผ่านมาปรากฏว่า การวินิจฉัยจากคะแนนสอบเพียงอย่างเดียวไม่เพียงพอที่จะทำให้การวินิจฉัยมีประสิทธิภาพ เพราะไม่สามารถระบุจุดเด่น จุดด้อย และสาเหตุความบกพร่องของผู้เรียนแต่ละคนได้อย่างชัดเจน จึงมีการใช้วิธีการวินิจฉัยจากแบบแผนการตอบข้อสอบ แต่การดำเนินการด้วยวิธีดังกล่าวยังมีข้อจำกัดของแบบแผนการตอบข้อสอบคือ ข้อมูลจากคะแนนสอบที่ได้เพื่อนำมาวิเคราะห์มีจำนวนมากเกินไป จึงมีการแก้ปัญหาโดยใช้วิธีการวิเคราะห์จากแบบแผนการตอบข้อสอบที่ผิดปกติแทนการวิเคราะห์จากแบบแผนการตอบข้อสอบทั้งหมด (Harnish & Linn, 2005)

ปัจจุบันนักวิชาการให้ความสนใจนำแนวคิดของทฤษฎีทางปัญญา (Cognitive theory) มาใช้กับการวินิจฉัย กว้างขวางขึ้น โดย Snow and Lohman (1989) ได้เสนอแนวคิดเรื่อง ความเกี่ยวข้องกันของจิตวิทยาทางปัญญากับการวัดผลทางการศึกษา (Implications of cognitive psychology for educational measurement) ในหนังสือการวัดผลทางการศึกษาชี้ให้เห็นว่า การจะทำให้การวัดทางจิตวิทยาเป็นที่ยอมรับในสังคมต้องใช้จิตวิทยาในการสร้างแบบทดสอบ (Psychology of test performance) เพราะจะทำให้เข้าใจลำดับโครงสร้างของคะแนน รวมทั้งการให้สารสนเทศของผลคะแนนการทดสอบที่ถูกต้อง เพราะแบบทดสอบทางการศึกษามีพื้นฐานมาจากงานการแก้ปัญหาทางปัญญา (Cognitive problem solve task)

นอกจากนี้ การวินิจฉัยที่ได้รับการยอมรับจากนักวิจัยและนักวัดผลจำนวนมากคือ การนำแนวคิดของทฤษฎีทางปัญญาใช้ในการทดสอบวินิจฉัย เป็นการเชื่อมโยงระหว่างจิตวิทยาทางปัญญา (Cognitive psychology) กับ การวัดทางจิตวิทยา (Psychometrics) โดยมีการพัฒนาเป็นโมเดลวินิจฉัยทางปัญญา (Cognitive diagnostic model: CDM) เพื่อใช้วินิจฉัยความรู้ของผู้สอบในกลุ่มที่มีการกำหนดทักษะหรือคุณลักษณะเฉพาะ ซึ่งนักวิชาการมีความเชื่อว่าจะทำให้ได้สารสนเทศจากการวินิจฉัยที่เป็นจุดเด่น จุดด้อย ปัญหา และสาเหตุที่เกิดปัญหานั้น ๆ อย่าง เฉพาะเจาะจง และสามารถนำสารสนเทศที่ได้จากการวินิจฉัยไปปรับปรุงแก้ไขหรือพัฒนาข้อบกพร่องต่าง ๆ เหล่านั้น ได้ตรงประเด็น

โมเดลวินิจฉัยทางปัญญาที่ได้รับความนิยมและนำมาใช้ในการวินิจฉัยทางการศึกษามากที่สุดโมเดลหนึ่งคือ Rule Space Model (RSM) เป็นโมเดลที่พัฒนาขึ้น โดย Tatsuoka (1997) เพื่อช่วยในการแก้ปัญหาเกี่ยวกับการวินิจฉัยมโนทัศน์ที่คลาดเคลื่อนจากการวัด โดยการแก้ไขปัญหาคือการวินิจฉัยที่ผ่านมาซึ่งไม่สามารถวินิจฉัยแหล่งมโนทัศน์ที่คลาดเคลื่อนได้ รวมทั้งสารสนเทศที่ได้จากการวินิจฉัยไม่สามารถนำมาใช้ประโยชน์ในการปรับปรุงแก้ไขจุดอ่อน หรือ

ข้อบกพร่องของผู้เรียนได้อย่างแท้จริง แต่ RSM ยังมีข้อจำกัดในเรื่องผลลัพธ์ที่ได้จากการวินิจฉัยไม่สมบูรณ์ เพราะเป็นการวินิจฉัยที่มีการดำเนินการภายหลังการสร้างแบบทดสอบแล้ว ทำให้การจำแนกคุณลักษณะที่ต้องใช้ในการตอบข้อสอบแต่ละข้อ ไม่ครอบคลุมคุณลักษณะของความรอบรู้ในแต่ละคุณลักษณะที่ผู้สอบจำเป็นต้องมี

ต่อมา Gierl, Leighton and Hunka (2000) ได้พัฒนาโมเดลวินิจฉัยทางปัญญาขึ้นใหม่ ใช้ชื่อว่าโมเดลลำดับชั้นคุณลักษณะ (Attribute Hierarchy Model: AHM) โมเดลลำดับชั้นคุณลักษณะนี้เป็นการพัฒนามาจากแนวคิดของ RSM มีลักษณะเป็นโมเดลทฤษฎีการตอบสนองข้อสอบทางปัญญา (Cognitive item response theory model) ที่แสดงถึงการเชื่อมโยงระหว่างทฤษฎีทางปัญญากับการวัดทางจิตวิทยา โดยที่นักวิชาการกลุ่มนี้มีความเชื่อว่า ในปัจจุบันการวัดทางจิตวิทยาไม่มีการนำทฤษฎีทางปัญญามาใช้อย่างแท้จริง ทำให้การพัฒนาแบบทดสอบและสารสนเทศที่ได้จากการทำแบบทดสอบขาดความชัดเจน ไม่เกิดประโยชน์ในการพัฒนาการเรียนรู้ของผู้เรียนเท่าที่ควร

โมเดลลำดับชั้นคุณลักษณะที่นำมาใช้ มีการกำหนดจำนวนข้อสอบ ลักษณะข้อสอบ และแบบแผนการตอบข้อสอบ จากโครงสร้างที่เป็นลำดับชั้นของความรู้หรือทักษะที่ต้องใช้ในการตอบข้อสอบที่เรียกว่าคุณลักษณะ (Attribute) ซึ่งแสดงถึงการทำงานที่เชื่อมโยงกันระหว่างทฤษฎีทางปัญญากับการวัดทางจิตวิทยา มีการจำแนกคุณลักษณะเนื้อหาหลัก ก่อนการพัฒนาแบบทดสอบ และมีการกำหนดคุณลักษณะเนื้อหาให้มีความสัมพันธ์กันเป็นลำดับชั้น ซึ่งเป็นขั้นตอนที่จะช่วยให้การกำหนดคุณลักษณะของข้อสอบแต่ละข้อ ให้สารสนเทศของการวินิจฉัยที่ได้จากแบบแผนการตอบข้อสอบมีความชัดเจนมากยิ่งขึ้น โดยกลุ่มผู้พัฒนาโมเดลลำดับชั้นคุณลักษณะได้กล่าวถึงการให้คุณลักษณะไว้ว่า เป็นการอธิบายถึงความรู้ และวิธีการที่ต้องใช้ในการทำงานที่มีความเฉพาะเจาะจงให้งานบรรลุเป้าหมาย คุณลักษณะเป็นตัวช่วยสร้างกลยุทธ์เพื่อใช้ในการแก้ปัญหา ชุดของคุณลักษณะที่ผู้เรียนมีความสามารถในเวลาหนึ่ง อาจไม่เหมือนกับชุดของคุณลักษณะที่ผู้เรียนมีความสามารถในเวลาหนึ่ง หากผู้เรียนมีพัฒนาการทางความรู้เพิ่มขึ้น อายุของผู้เรียนเพิ่มขึ้น หรือเกิดจากการบริหารจัดการการเรียนรู้ที่มีประสิทธิภาพมากขึ้น (Leighton, Gierl, & Hunka, 2002)

การเรียนรู้ภาษาอังกฤษสำหรับผู้เรียนที่ไม่ใช่เจ้าของภาษา ต้องใช้ทักษะการอ่านเป็นเครื่องมือในการแสวงหาความรู้จากตำรา เอกสาร งานวิจัย และบทความทางวิชาการต่าง ๆ ที่ต้องใช้ภาษาอังกฤษเป็นส่วนมาก และปัญหาโดยทั่วไปของนักศึกษาระดับปริญญาตรีที่พบคือ นักศึกษาส่วนใหญ่มีทักษะการอ่านภาษาอังกฤษอยู่ในเกณฑ์ต่ำ ต้องมีการปรับปรุงในหลายด้าน เช่น อ่านแล้วไม่เข้าใจเนื้อเรื่อง ไม่สามารถสรุปใจความสำคัญของเรื่องที่อ่านได้ หรือแปลแล้วไม่รู้เรื่อง (ประภาพร ชนะจินะศักดิ์, 2554) ซึ่งการอ่านภาษาอังกฤษให้เกิดความเข้าใจในเนื้อหาที่อ่านนั้น จำเป็นต้องมีทักษะทางภาษาอังกฤษหลายด้านที่ผู้เรียนมีข้อบกพร่องอยู่ และจำเป็นต้องได้รับการแก้ไขให้ตรงจุดที่เป็นปัญหาและสาเหตุที่แตกต่างกันเป็นรายบุคคล จึงจะสามารถเพิ่มศักยภาพของการอ่านได้อย่างแท้จริง การวินิจฉัยด้วยการจำแนกเนื้อหาในทักษะที่จำเป็นต้องใช้สำหรับการอ่านภาษาอังกฤษ และจำแนกเนื้อหาออกเป็นลำดับชั้นตามคุณลักษณะก่อนการพัฒนาแบบทดสอบวินิจฉัยถือเป็นวิธีที่ดีที่สุด ในการค้นหาสาเหตุและแก้ไขข้อบกพร่องนั้น ๆ เพราะการวินิจฉัยคุณลักษณะเนื้อหาเป็นลำดับชั้นสามารถตอบปัญหาสำคัญได้ว่า ผู้เรียนคนใดมีความบกพร่องในทักษะใด ความบกพร่องนั้นอยู่ในลำดับชั้นคุณลักษณะใด สาเหตุของความบกพร่องในการอ่านที่ตรงกับทักษะนั้นมีอะไรบ้าง จะทำการสอนซ่อมเสริมและป้องกันความบกพร่องของผู้เรียนได้อย่างไร แสดงให้เห็นว่าการวินิจฉัยโดยการจัดลำดับชั้นของเนื้อหา จะสามารถแก้ไขข้อบกพร่องที่เกิดจากการใช้ทักษะที่จำเป็นสำหรับการอ่าน และยังสามารถป้องกันทักษะที่เป็นความบกพร่องในการอ่านของผู้เรียนเป็นรายบุคคลได้อีกด้วย

นอกจากนี้ ผลงานวิจัยและการประเมินผลความสำเร็จในการอ่านภาษาอังกฤษของผู้เรียนหลายเรื่องชี้ให้เห็นว่าผู้เรียนมีผลสัมฤทธิ์ทางการเรียนอยู่ในเกณฑ์ที่ต้องปรับปรุงแก้ไข จึงจำเป็นต้องหาสาเหตุและแก้ไขให้ดีขึ้น เพื่อนำผลที่ได้ไปใช้ในพัฒนาและปรับปรุงการจัดการเรียนการสอน (สุพรรณิณี พันธุ์พริก และฉลิลิกา มหาพูนทอง,

2550; ประภาพร ชนะจีนะศักดิ์, 2554) ซึ่งการที่จะทราบว่าผู้เรียนมีข้อบกพร่องในลำดับขั้นใดของทักษะที่จำเป็นต้องใช้สำหรับการอ่านจะต้องทำการประเมินโดยการใช่วิธีการทดสอบวินิจฉัย เพราะผลที่ได้จะเป็นข้อมูลที่บ่งบอกสาเหตุและข้อบกพร่องในแต่ละลำดับขั้นของทักษะนั้น ๆ ได้

วิธีการวินิจฉัยโดยใช้แนวคิดโมเดลลำดับขั้นคุณลักษณะในการจำแนกคุณลักษณะของเนื้อหาออกเป็นลำดับขั้นก่อนการสร้างแบบทดสอบ นอกจากจะทำให้ได้สารสนเทศจากการวินิจฉัยที่ชัดเจนแล้ว วิธีการทดสอบวินิจฉัยยังมีการจำแนกแบบแผนการตอบข้อสอบระหว่างแบบแผนการตอบข้อสอบที่สังเกตได้กับแบบแผนการตอบข้อสอบที่คาดหวัง ส่งผลต่อการจำแนกที่ต้องการตัดสินคุณลักษณะที่เป็นสภาพความรู้ในตัวผู้สอบ ซึ่งเป็นหัวใจสำคัญของ การวินิจฉัย (Gierl & Wang, 2008) แต่การกำหนดลำดับขั้นคุณลักษณะหลักของเนื้อหาอาจเกิดข้อจำกัดที่ลำดับขั้นคุณลักษณะที่กำหนดไม่แสดงถึงลำดับขั้นที่เป็นทักษะหรือกระบวนการทางปัญญาที่แท้จริงในการตอบข้อสอบ อาจเป็นลำดับขั้นของทักษะหรือกระบวนการทางปัญญาที่อยู่ในลำดับขั้นเดียวกัน แต่ถูกนำมากำหนดให้อยู่ในลำดับขั้นที่สูงกว่าหรือต่ำกว่า หรืออาจมาจากสาเหตุของแบบแผนการตอบข้อสอบไม่แสดงความสามารถที่แท้จริงของผู้สอบ การที่ผู้สอบตอบข้อสอบข้อนั้นถูกต้องไม่ได้แสดงว่าผู้สอบมีความรู้หรือมีทักษะทางปัญญาในเรื่องนั้น ในขณะที่เดียวกันการที่ผู้สอบตอบข้อสอบข้อนั้นไม่ถูกต้องไม่ได้แสดงว่าผู้สอบไม่มีความรู้หรือทักษะทางปัญญานั้น แต่อาจเกิดจากตัวผู้สอบเองมีความสะเพร่า มีการเดาข้อสอบหรือคัดลอกข้อสอบจากผู้อื่นก็เป็นได้ (Gierl & Zheng, 2008)

การลดข้อจำกัดดังกล่าวมาสามารถดำเนินการได้โดยการพัฒนาวิธีการวินิจฉัยแบบใหม่คือ การวินิจฉัยโดยใช้วิธีการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (Cheng, 2009) เพราะการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์มีข้อได้เปรียบคือ เป็นวิธีการทดสอบที่ผู้สอบแต่ละคนไม่จำเป็นต้องทำข้อสอบฉบับเดียวกัน และไม่จำเป็นต้องทำข้อสอบทุกข้อ และยังลดข้อจำกัดในกรณีที่ข้อสอบมีความยากเกินไปหรือง่ายเกินไปสำหรับผู้สอบซึ่งเป็นสาเหตุที่ทำให้ผู้สอบเกิดความเบื่อหน่ายในการสอบ เนื่องจากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ใช้วิธีการประมาณค่าความสามารถของผู้สอบและคัดเลือกข้อสอบข้อถัดไปเพื่อให้ได้ข้อสอบที่สอดคล้องกับความสามารถของผู้สอบเป็นสำคัญ ผู้สอบแต่ละคนจะได้ทำข้อสอบข้อที่แตกต่างกันตามระดับความสามารถของแต่ละบุคคล นอกจากนี้การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ยังช่วยป้องกันไม่ให้ผู้สอบคัดลอกคำตอบของข้อสอบจากผู้อื่น รวมทั้งเป็นการช่วยเพิ่มประสิทธิภาพในการวัดทั้งในส่วนของการลดภาระงานการทดสอบ และเพิ่มความถูกต้องในการประมาณค่าความสามารถของผู้สอบอีกด้วย

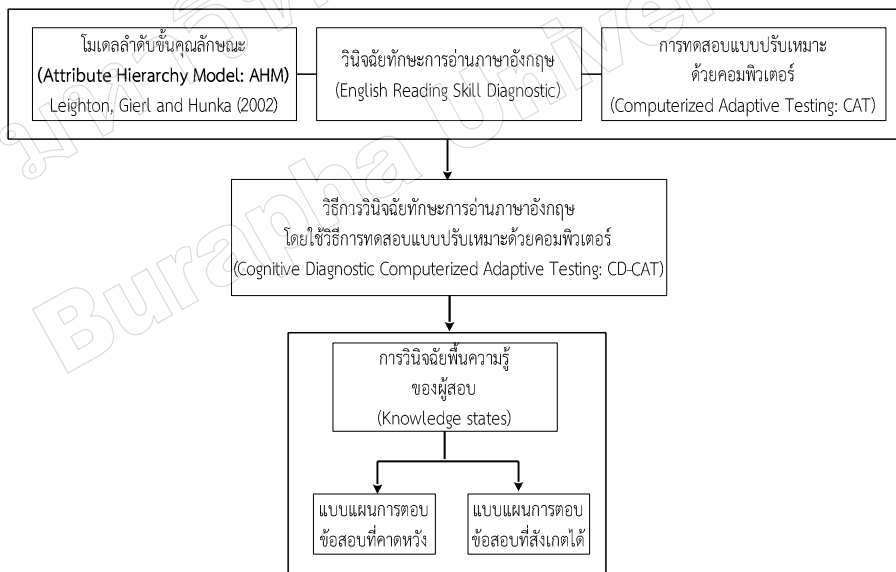
จากการศึกษาเอกสารงานวิจัยต่าง ๆ ที่เกี่ยวข้องกับสภาพปัญหาการประเมินผลความสำเร็จในทักษะการอ่านภาษาอังกฤษ แนวคิดทฤษฎีการวัดและประเมินผลการศึกษา และวิธีการวินิจฉัยดังกล่าวข้างต้น ผู้วิจัยจึงต้องการพัฒนาวิธีการวินิจฉัยทักษะการอ่านภาษาอังกฤษของนักศึกษาระดับปริญญาตรีโดยใช้แนวคิดของโมเดลลำดับขั้นคุณลักษณะในการกำหนดคุณลักษณะที่เป็นทักษะการอ่านภาษาอังกฤษให้มีคุณลักษณะเฉพาะ มีการจัดเรียงตามลักษณะความสัมพันธ์เชิงลำดับขั้นของคุณลักษณะจากคุณลักษณะที่เป็นพื้นฐานไปยังคุณลักษณะเฉพาะขั้นสูงกว่าก่อนการพัฒนาชุดข้อสอบ และจัดทำคลังข้อสอบ รวมทั้งออกแบบวิธีการทดสอบวินิจฉัยโดยใช้วิธีการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ โดยใช้แนวคิดโมเดลลำดับขั้นคุณลักษณะในการกำหนดลำดับขั้นการเลือกข้อสอบจากคลังข้อสอบ ด้วยเกณฑ์การตัดสินความรอบรู้จากแบบแผนการตอบข้อสอบในแต่ละลำดับขั้นคุณลักษณะที่แตกต่างกัน เพื่อให้สามารถจำแนกพื้นฐานความรู้ของผู้สอบเป็นรายบุคคล และได้สารสนเทศที่เป็นพื้นฐานความรู้จากวิธีการวินิจฉัยที่ชัดเจน เป็นประโยชน์สำหรับอาจารย์ผู้สอนภาษาอังกฤษสามารถนำผลการวิจัยไปใช้ในการวินิจฉัยการเรียนรู้ของนักศึกษาได้รวดเร็ว ช่วยลดเวลาในการวินิจฉัยนักศึกษาเป็นรายบุคคล รวมทั้งสามารถสรุปผลและให้ข้อมูลย้อนกลับแก่นักศึกษาในส่วนที่เป็นความรอบรู้หรือจุดบกพร่องในทันที นักศึกษาสามารถนำผลที่ได้จากการวินิจฉัยไปพัฒนาตนเองเพื่อเพิ่มศักยภาพการเรียนรู้ด้านทักษะการอ่านภาษาอังกฤษในระดับสูงต่อไป

วัตถุประสงค์การวิจัย

1. เพื่อพัฒนาแบบทดสอบวินิจฉัยทักษะการอ่านภาษาอังกฤษของนักศึกษาระดับปริญญาตรี โดยใช้โมเดลลำดับชั้นคุณลักษณะ
2. เพื่อพัฒนาวิธีการวินิจฉัยทักษะการอ่านภาษาอังกฤษ โดยใช้โมเดลลำดับชั้นคุณลักษณะและการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์
3. เพื่อวินิจฉัยพื้นฐานความรู้ของผู้สอบ (Knowledge states) ตามแบบแผนการตอบข้อสอบ โดยใช้วิธีการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่พัฒนาขึ้น

กรอบแนวคิดการวิจัย

การพัฒนาวิธีการวินิจฉัยทักษะการอ่านภาษาอังกฤษโดยการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์นี้ ผู้วิจัยใช้โมเดลลำดับชั้นคุณลักษณะ (Attribute Hierarchy Model: AHM) ของ Leighton, Gierl, and Hunka (2002) เป็นกรอบแนวคิดในการกำหนดเนื้อหาทักษะการอ่านภาษาอังกฤษให้มีคุณลักษณะเฉพาะด้วย Q-Matrix ซึ่งมีการจัดเรียงตามลักษณะความสัมพันธ์เชิงลำดับชั้นของคุณลักษณะ จากคุณลักษณะที่เป็นพื้นฐานไปยังคุณลักษณะเฉพาะขั้นสูงกว่าก่อนการพัฒนาชุดข้อสอบ ส่วนวิธีการวินิจฉัยทักษะการอ่านภาษาอังกฤษโดยใช้วิธีการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ใช้แนวคิดของโมเดลลำดับชั้นคุณลักษณะในการกำหนดลำดับชั้นการเลือกข้อสอบจากคลังข้อสอบ เพื่อใช้ในการวินิจฉัยพื้นฐานความรู้ของผู้สอบ ตามแบบแผนการตอบข้อสอบในแต่ละคุณลักษณะที่แตกต่างกัน เขียนเป็นกรอบแนวคิดการวิจัยได้ ดังภาพที่ 1



ภาพที่ 1 กรอบแนวคิดการพัฒนาวิธีการวินิจฉัยทักษะการอ่านภาษาอังกฤษ โดยใช้โมเดลลำดับชั้นคุณลักษณะ และการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

วิธีดำเนินการวิจัย

แบ่งออกเป็น 3 ขั้นตอน ดังนี้

ขั้นตอนที่ 1 การพัฒนาแบบทดสอบวินิจฉัยทักษะการอ่านภาษาอังกฤษ โดยใช้โมเดลลำดับชั้นคุณลักษณะ เป็นขั้นตอนการดำเนินการเพื่อให้ได้ชุดข้อสอบและคลังข้อสอบสำหรับวินิจฉัยพื้นฐานความรู้ด้านทักษะการอ่านภาษาอังกฤษ เพื่อใช้ทดสอบวินิจฉัยบนโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ มีขั้นตอนดังนี้ 1) การกำหนดลำดับชั้นคุณลักษณะเนื้อหา 2) การพัฒนาชุดข้อสอบ 3) การจัดทำแบบทดสอบแบบเขียนตอบบนกระดาษคำตอบ 4) การตรวจสอบคุณภาพของแบบทดสอบ 5) การจัดทำคลังข้อสอบ

ขั้นตอนที่ 2 การพัฒนาวิธีการวินิจฉัยทักษะการอ่านภาษาอังกฤษโดยใช้การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ เป็นขั้นตอนของการออกแบบวิธีการทดสอบวินิจฉัยบนโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ พัฒนาในรูปแบบของ Web Application กำหนดรายละเอียดของการทดสอบวินิจฉัย เพื่อให้โปรแกรมมีคุณลักษณะเฉพาะสำหรับใช้ในการทดสอบวินิจฉัย มีขั้นตอนดังนี้ 1) เพิ่มข้อสอบบนโปรแกรมการทดสอบ 2) กำหนดลำดับชั้นการทำข้อสอบบนโปรแกรมการทดสอบ 3) กำหนดเงื่อนไขการเลื่อนลำดับชั้นการทำข้อสอบ 4) กำหนดเงื่อนไขการยุติการทดสอบ 5) การจัดทำคู่มือการใช้งาน 6) ประเมินความเหมาะสมของวิธีการวินิจฉัยโดยผู้เชี่ยวชาญและนักศึกษาทดลองใช้

ขั้นตอนที่ 3 การวินิจฉัยพื้นฐานความรู้ด้านทักษะการอ่านภาษาอังกฤษของผู้สอบ โดยใช้วิธีการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ เป็นขั้นตอนของการนำโปรแกรมการทดสอบวินิจฉัยแบบปรับเหมาะด้วยคอมพิวเตอร์ที่พัฒนาขึ้นไปใช้เพื่อให้ได้ข้อมูลที่เป็นจุดเด่นและข้อบกพร่องของผู้สอบด้านทักษะการอ่านภาษาอังกฤษแต่ละคุณลักษณะ โดยการจำแนกจำนวนผู้สอบตามแบบแผนการตอบข้อสอบในแต่ละคุณลักษณะที่แตกต่างกัน

ประชากรและกลุ่มตัวอย่าง

ประชากรที่ใช้ในการวิจัยเป็นนักศึกษาระดับปริญญาตรี มหาวิทยาลัยราชภัฏสวนดุสิต กรุงเทพมหานคร ที่ศึกษาอยู่ในชั้นปีที่ 2 ถึงชั้นปีที่ 4 ในภาคเรียนที่ 1 ปีการศึกษา 2555 และผ่านการเรียนรายวิชาภาษาอังกฤษเพื่อทักษะการเรียนมาแล้ว จำนวน 16,361 คน

กลุ่มตัวอย่างเป็นนักศึกษาที่ผ่านการเรียนรายวิชาภาษาอังกฤษเพื่อทักษะการเรียนมาแล้ว และมีระดับผลการเรียนต่ำ (เกรดระดับ D+ และระดับ D) ได้มาโดยการสุ่มอย่างมีระบบตามรหัสนักศึกษา จาก 7 คณะวิชา ได้แก่ คณะครุศาสตร์ คณะวิทยาศาสตร์และเทคโนโลยี คณะวิทยาการจัดการ คณะมนุษยศาสตร์และสังคมศาสตร์ คณะพยาบาลศาสตร์ โรงเรียนการเรือน และโรงเรียนการท่องเที่ยวและการบริการ คณะวิชาละ 34 คน จำนวนรวมทั้งสิ้น 238 คน

การวิเคราะห์ข้อมูล

1. วิเคราะห์ข้อมูลเกี่ยวกับการพัฒนาแบบทดสอบวินิจฉัยทักษะการอ่านภาษาอังกฤษ โดยใช้โมเดลลำดับชั้นคุณลักษณะ ดังนี้

1.1 ประเมินความถูกต้องของเนื้อหา และพิจารณาคูณลักษณะที่มีค่าสัดส่วนความสอดคล้องมากกว่าหรือเท่ากับ 2 ใน 3 เปรียบเทียบลำดับชั้นคุณลักษณะด้วยค่าความถี่และร้อยละ และประเมินความถูกต้องของแผนภาพลำดับชั้นคุณลักษณะโดยตรวจสอบความถูกต้องสมบูรณ์ของข้อมูล

1.2 วิเคราะห์ข้อมูลในขั้นตอนของการพัฒนาชุดข้อสอบ การสำรวจข้อบกพร่องในทักษะการอ่านภาษาอังกฤษเพื่อจำแนกตัวลงในการสร้างแบบทดสอบวินิจฉัย ตรวจสอบความถูกต้องของข้อมูล และคำนวณค่าความถี่และร้อยละของผลการประเมิน วิเคราะห์ความสอดคล้องของข้อคำถามกับคุณลักษณะ โดยพิจารณาจากค่า

ดัชนีความสอดคล้อง ตรวจสอบคุณภาพข้อสอบโดยวิเคราะห์ค่าพารามิเตอร์ของข้อสอบตามทฤษฎีการตอบสนองข้อสอบโดยใช้โปรแกรมสำเร็จรูป Xcalibre Version 4.1.7 และวิเคราะห์ความเที่ยงของแบบทดสอบโดยใช้วิธีตรวจสอบความสอดคล้องภายใน (KR-20)

2. วิเคราะห์ข้อมูลการประเมินวิธีการวินิจฉัยทักษะการอ่านภาษาอังกฤษโดยโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ดังนี้

2.1 วิเคราะห์ความคิดเห็นของผู้เชี่ยวชาญที่มีต่อวิธีการวินิจฉัย ด้วยค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน และวิเคราะห์ข้อมูลเชิงคุณภาพโดยสรุปเนื้อหาเป็นความเรียง

2.2 วิเคราะห์ความคิดเห็นของนักศึกษากลุ่มตัวอย่าง ที่มีต่อการทดลองใช้วิธีการทดสอบวินิจฉัยบนโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์เบื้องต้น ด้วยค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน และวิเคราะห์ข้อมูลเชิงคุณภาพโดยสรุปเนื้อหาเป็นความเรียง

3. วิเคราะห์ข้อมูลการทดสอบวินิจฉัยตามแบบแผนการตอบข้อสอบโดยใช้วิธีการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ด้วยค่าความถี่ และร้อยละ

ผลการวิจัย

ผลการพัฒนาแบบทดสอบวินิจฉัยทักษะการอ่านภาษาอังกฤษโดยใช้โมเดลลำดับชั้นคุณลักษณะ

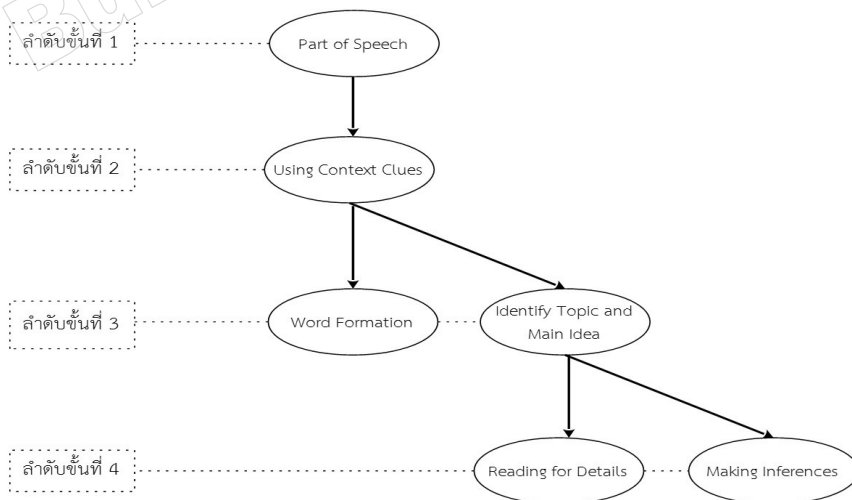
การพัฒนาแบบทดสอบวินิจฉัยทักษะการอ่านภาษาอังกฤษโดยใช้โมเดลลำดับชั้นคุณลักษณะ สรุปผลการจัดลำดับชั้นคุณลักษณะของทักษะการอ่านภาษาอังกฤษ ได้ดังนี้

ลำดับชั้นที่ 1 ได้แก่ คุณลักษณะที่ 1 Part of Speech

ลำดับชั้นที่ 2 ได้แก่ คุณลักษณะที่ 2 Using Context Clues

ลำดับชั้นที่ 3 ได้แก่ คุณลักษณะที่ 3 Word Formation และคุณลักษณะที่ 4 Identify Topics and Main Idea เป็นคุณลักษณะที่อยู่ในลำดับชั้นเดียวกัน

ลำดับชั้นที่ 4 ได้แก่ คุณลักษณะที่ 5 Reading for Details และคุณลักษณะที่ 6 Making Inferences เป็นคุณลักษณะที่อยู่ในลำดับชั้นเดียวกัน ดังภาพที่ 2



ภาพที่ 2 ผลการจัดลำดับชั้นคุณลักษณะของทักษะการอ่านภาษาอังกฤษ

ผลการวิเคราะห์ค่าพารามิเตอร์ของข้อสอบตามทฤษฎีการตอบสนองข้อสอบจำแนกตามคุณลักษณะ โดยคัดเลือกข้อสอบที่มีค่าอำนาจจำแนกของข้อสอบอยู่ระหว่าง 0.50 ถึง 2.50 ค่าความยากของข้อสอบอยู่ระหว่าง -2.50 ถึง 2.50 และค่าการเดาของข้อสอบไม่เกิน 0.30 ปรากฏว่า ได้ข้อสอบที่คัดเลือกไว้เป็นฐานข้อมูลคลังข้อสอบสำหรับใช้ทดสอบวินิจฉัยบนโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ประกอบด้วย ชุดข้อสอบคุณลักษณะที่ 1 จำนวน 120 ข้อ ชุดข้อสอบคุณลักษณะที่ 2 จำนวน 104 ข้อ ชุดข้อสอบคุณลักษณะที่ 3 จำนวน 107 ข้อ ชุดข้อสอบคุณลักษณะที่ 4 จำนวน 95 ข้อ ชุดข้อสอบคุณลักษณะที่ 5 จำนวน 103 ข้อ ชุดข้อสอบคุณลักษณะที่ 6 จำนวน 92 ข้อ รวมทั้งสิ้น 621 ข้อ ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของค่าพารามิเตอร์ของข้อสอบแสดงดังตารางที่ 1

ตารางที่ 1 ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของค่าพารามิเตอร์ของข้อสอบในคลังข้อสอบจำแนกตามคุณลักษณะของทักษะการอ่านภาษาอังกฤษ

คุณลักษณะที่	ค่าอำนาจจำแนกของข้อสอบ (a)		ค่าความยากของข้อสอบ (b)		ค่าการเดา (c)	
	M	SD	M	SD	M	SD
1	1.33	0.61	0.55	1.32	0.06	0.04
2	1.29	0.57	0.76	1.35	0.05	0.02
3	1.22	0.52	0.48	1.45	0.09	0.08
4	1.30	0.63	0.24	1.56	0.07	0.06
5	1.23	0.60	0.92	1.31	0.09	0.07
6	1.32	0.46	0.55	1.45	0.08	0.06

หมายเหตุ: M = Mean, SD = Standard Deviation

จากตารางที่ 1 แสดงค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของค่าพารามิเตอร์ของข้อสอบในคลังข้อสอบจำแนกตามคุณลักษณะ ปรากฏว่า ค่าพารามิเตอร์ของข้อสอบในคลังข้อสอบโดยเฉลี่ยอยู่ในเกณฑ์ดี พิจารณาได้จากค่าเฉลี่ยของค่าอำนาจจำแนกมีค่ามากกว่า 1.00 ทุกคุณลักษณะ ค่าเฉลี่ยของค่าการเดาน้อยกว่า 0.30 ทุกคุณลักษณะ ส่วนค่าเฉลี่ยของค่าความยากของข้อสอบมีค่าอยู่ระหว่าง 0.24 ถึง 0.92 อยู่ในเกณฑ์ง่ายถึงค่อนข้างยาก

ผลการพัฒนาวิธีการวินิจฉัยทักษะการอ่านภาษาอังกฤษโดยใช้วิธีการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

การพัฒนาวิธีการวินิจฉัยทักษะการอ่านภาษาอังกฤษโดยใช้วิธีการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ประกอบด้วยขั้นตอน ดังนี้ การพัฒนาคลังข้อสอบ การคัดเลือกข้อสอบข้อแรก การคัดเลือกข้อสอบจากคลังข้อสอบ การประมาณค่าความสามารถผู้สอบ และการรายงานผลการทดสอบตามวิธีการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ส่วนวิธีการทดสอบวินิจฉัยใช้แนวคิดของโมเดลลำดับชั้นคุณลักษณะในการกำหนดแบบแผนการทำชุดข้อสอบแต่ละคุณลักษณะ จำนวนข้อสอบที่ต้องทำในแต่ละลำดับชั้นของคุณลักษณะ เกณฑ์การเลื่อนลำดับชั้นการทำชุดข้อสอบสำหรับคุณลักษณะถัดไป รวมทั้งการยุติการทดสอบ ตามแบบแผนการตอบข้อสอบที่กำหนด 12 แบบแผน โดยกำหนดจำนวนข้อที่ให้ทำในแต่ละคุณลักษณะ ดังนี้

1) คุณลักษณะที่ 1 และคุณลักษณะที่ 2 ผู้สอบต้องทำข้อสอบทั้งหมด 10 ข้อ หากผู้สอบทำข้อสอบถูก 6 ข้อขึ้นไป สามารถผ่านไปทดสอบคุณลักษณะถัดไป แต่ถ้าทำข้อสอบถูกน้อยกว่า 6 ข้อ โปรแกรมจะยุติการทดสอบและรายงานผลการทดสอบ

แบบที่ 10 (111101) ผู้สอบสอบผ่านคุณลักษณะที่ 1 คุณลักษณะที่ 2 คุณลักษณะที่ 3 และคุณลักษณะที่ 4 ทำชุดข้อสอบคุณลักษณะที่ 5 สอบไม่ผ่าน ทำชุดข้อสอบคุณลักษณะที่ 6 สอบผ่าน โปรแกรมจะยุติการทดสอบและรายงานผลการทดสอบ

แบบที่ 11 (110111) ผู้สอบสอบผ่านคุณลักษณะที่ 1 คุณลักษณะที่ 2 สอบไม่ผ่านคุณลักษณะที่ 3 สอบผ่านคุณลักษณะที่ 4 ทำชุดข้อสอบคุณลักษณะที่ 5 สอบผ่าน ทำชุดข้อสอบคุณลักษณะที่ 6 สอบผ่าน โปรแกรมจะยุติการทดสอบและรายงานผลการทดสอบ

แบบที่ 12 (111111) ผู้สอบสอบผ่านคุณลักษณะที่ 1 คุณลักษณะที่ 2 คุณลักษณะที่ 3 และคุณลักษณะที่ 4 ทำชุดข้อสอบคุณลักษณะที่ 5 สอบผ่าน ทำชุดข้อสอบคุณลักษณะที่ 6 สอบผ่านโปรแกรมจะยุติการทดสอบและรายงานผลการทดสอบ

ผลการวิจัยพื้นฐานความรู้ของผู้สอบโดยใช้วิธีการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

ผลการวิจัยพื้นฐานความรู้ของผู้สอบโดยใช้วิธีการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ประกอบด้วย 1) ผลการวิจัยจำนวนผู้สอบจำแนกตามคุณลักษณะ 2) ผลการวิจัยจำนวนผู้สอบจำแนกตามแบบแผนการตอบข้อสอบ ปรากฏผลดังตารางที่ 2 และ 3

ตารางที่ 2 จำนวนผู้สอบจำแนกตามพื้นฐานความรู้ในแต่ละคุณลักษณะ

คุณลักษณะ	ผลการทดสอบวินิจฉัย (n=238)						รวม
	ผู้สอบผ่าน		ผู้สอบไม่ผ่าน		ไม่ได้รับการทดสอบ		
	จำนวน	ร้อยละ	จำนวน	ร้อยละ	จำนวน	ร้อยละ	
1. Part of Speech	216	90.76	22	9.24	0	0.00	238
2. Using Context Clues	180	75.63	36	15.13	22	9.24	238
3. Word Formation	110	46.22	70	29.41	58	24.37	238
4. Identify Topic and Main Idea	121	50.84	41	17.23	76	31.93	238
5. Reading for Details	37	15.55	85	35.71	116	48.74	238
6. Making Inferences	69	28.99	53	22.27	116	48.74	238

จากตารางที่ 2 แสดงจำนวนผู้สอบจำแนกตามพื้นฐานความรู้ในแต่ละคุณลักษณะ แสดงให้เห็นว่า สามารถจำแนกผู้สอบตามพื้นฐานความรู้ที่ได้จากการทดสอบวินิจฉัย ดังนี้ กลุ่มผู้สอบผ่าน เป็นกลุ่มที่มีความรู้ในทักษะการอ่านภาษาอังกฤษคุณลักษณะที่ 1 Part of Speech มากที่สุด จำนวน 216 คน คิดเป็นร้อยละ 90.76 รองลงมาคือคุณลักษณะที่ 2 Using Context Clues จำนวน 180 คน คิดเป็นร้อยละ 75.63 ส่วนคุณลักษณะที่ 5 Reading for Details มีผู้มีความรู้น้อยที่สุด

ตารางที่ 3 จำนวนผู้สอบจำแนกตามพื้นความรู้ในแต่ละแบบแผนการตอบข้อสอบ

แบบแผนการตอบ	ผลการทดสอบวินิจฉัย (n=238)						รวม (คน)
	ผู้สอบผ่าน		ผู้สอบไม่ผ่าน		ไม่ได้รับการทดสอบ		
	จำนวน	ร้อยละ	จำนวน	ร้อยละ	จำนวน	ร้อยละ	
(1) 000000	0	0.00	22	9.24	110	46.22	132
(2) 100000	36	15.13	36	15.13	144	60.50	216
(3) 110000	36	15.13	18	7.56	54	22.69	108
(4) 111000	118	49.58	42	17.65	80	33.61	240
(5) 110100	53	22.27	55	23.11	0	0.00	108
(6) 111100	83	34.87	43	18.07	0	0.00	126
(7) 110110	48	20.17	12	5.04	0	0.00	60
(8) 111110	46	19.33	8	3.36	0	0.00	54
(9) 110101	85	35.71	41	17.23	0	0.00	126
(10) 111101	131	55.04	25	10.50	0	0.00	156
(11) 110111	25	10.50	5	2.10	0	0.00	30
(12) 111111	72	30.25	0	0.00	0	0.00	72

จากตารางที่ 3 แสดงจำนวนผู้สอบตามพื้นความรู้ในทักษะการอ่านภาษาอังกฤษของผู้สอบจำแนกตามแบบแผนการตอบข้อสอบ ปรากฏว่า จำนวนผู้สอบมีแบบแผนการตอบข้อสอบแบบที่ 10 (111101) มากที่สุด จำนวน 131 คน คิดเป็นร้อยละ 55.04 รองลงมาคือ แบบแผนการตอบข้อสอบแบบที่ 4 (111000) จำนวน 118 คน คิดเป็นร้อยละ 49.58 และน้อยที่สุดคือ แบบแผนการตอบข้อสอบแบบที่ 11 (110111) จำนวน 25 คน คิดเป็นร้อยละ 10.50

อภิปรายผล

1. การพัฒนาแบบทดสอบวินิจฉัยทักษะการอ่านภาษาอังกฤษนี้ เป็นการนำแนวคิดโมเดลลำดับชั้นคุณลักษณะของ Leighton, Gierl and Hunka (2002) มาใช้ในการกำหนดลำดับชั้นคุณลักษณะทักษะการอ่านภาษาอังกฤษก่อนทำการพัฒนาชุดข้อสอบและคลังข้อสอบ สำหรับใช้ทดสอบวินิจฉัยบนโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ซึ่งจากการดำเนินการ ได้ผลการจัดลำดับชั้นคุณลักษณะของทักษะการอ่านภาษาอังกฤษ จำนวน 4 ลำดับชั้น 6 คุณลักษณะ ตามลักษณะความสัมพันธ์เชิงลำดับชั้นของคุณลักษณะ จากคุณลักษณะที่เป็นพื้นฐานไปยังคุณลักษณะเฉพาะขั้นสูงกว่าตามหลักจิตวิทยา นั่นคือ การกำหนดคุณลักษณะที่เป็นจุดเริ่มต้นซึ่งเป็นที่ต้องมีมาก่อนสำหรับคุณลักษณะทั้งหมดจากจุดเดียวกัน เพื่อช่วยให้การพัฒนาแบบทดสอบมีการตระหนักถึงคุณลักษณะเริ่มต้นที่ผู้สอบทุกคนต้องมี และได้ชุดข้อสอบในคลังข้อสอบจำนวน 6 ชุด รวมทั้งสิ้น 621 ข้อ ซึ่งเป็นไปตามข้อเสนอแนะของ Weiss (2004) ทั้งจำนวนข้อสอบในคลังข้อสอบทั้งหมดและจำนวนข้อสอบในคลังข้อสอบจำแนกตามชุดของข้อสอบส่วนด้านคำพารามิเตอร์ของข้อสอบ ผลปรากฏว่า มีความสอดคล้องกับแผนภาพลำดับชั้นคุณลักษณะเมื่อมีการจำแนกคำพารามิเตอร์ตามชุดของข้อสอบ เพราะในขั้นตอนของการกำหนดลำดับชั้นคุณลักษณะ มีการตรวจสอบความถูกต้องของการกำหนดลำดับชั้นคุณลักษณะ การจัดเรียงลำดับชั้นคุณลักษณะ และการจัดทำแผนภาพลำดับชั้นคุณลักษณะ

โดยผู้เชี่ยวชาญด้านเนื้อหาทักษะการอ่านภาษาอังกฤษ (Expert analysis) รวมทั้งการวิเคราะห์จากการศึกษาเอกสารงานวิจัยที่เกี่ยวข้อง (Document analysis)

2. วิธีการทดสอบวินิจฉัยทักษะการอ่านภาษาอังกฤษที่พัฒนาขึ้นโดยใช้แนวคิดของโมเดลลำดับชั้นคุณลักษณะในการกำหนดลำดับชั้นการคัดเลือกข้อสอบจากคลังข้อสอบ ทำให้วิธีการวินิจฉัยทักษะการอ่านภาษาอังกฤษบนโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ในงานวิจัยนี้ แตกต่างจากวิธีการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ทั่วไป คือคลังข้อสอบมีการจำแนกชุดข้อสอบตามจำนวนคุณลักษณะที่กำหนด ในขั้นตอนของการพัฒนาชุดข้อสอบ โดยชุดข้อสอบแต่ละชุดจะมีความสัมพันธ์กันเป็นลำดับชั้น การกำหนดลำดับชั้นการเลือกคุณลักษณะของชุดข้อสอบจากคลังข้อสอบมีลักษณะเฉพาะตามแผนภาพลำดับชั้นคุณลักษณะที่กำหนด จากลักษณะของวิธีการวินิจฉัยทักษะการอ่านภาษาอังกฤษบนโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ทำให้สารสนเทศของการวินิจฉัยที่อธิบายได้ด้วยทฤษฎีทางปัญญา (Cognitive theory) เกี่ยวกับพื้นฐานความรู้ที่เป็นความรู้หรือจุดเด่น และข้อบกพร่องหรือจุดด้อย ในด้านทักษะการอ่านภาษาอังกฤษของนักศึกษาที่ต้องการพัฒนาหรือปรับปรุงแก้ไข ข้อบกพร่องนั้น ๆ มีความเป็นลำดับชั้นตามแผนภาพลำดับชั้นคุณลักษณะ ซึ่งสอดคล้องกับงานวิจัยของ Wang and Gierl (2007) ที่ใช้โมเดลลำดับชั้นคุณลักษณะในการวินิจฉัยทักษะการอ่านเชิงวิจารณ์ (The SAT Critical Reading Subtest) ผลปรากฏว่า โมเดลลำดับชั้นคุณลักษณะมีประโยชน์มากกว่าวิธีการวัดทางจิตแบบเดิม (Traditional psychometric methods) ในเรื่องของผลการรายงานผลการทดสอบ ที่ให้ทั้งคะแนนรวมและรายละเอียดข้อมูลการวินิจฉัยทางปัญญาของผู้สอบ ช่วยให้ผู้สอบได้ทราบถึงจุดบกพร่องที่ควรปรับปรุงแก้ไข ในแต่ละคุณลักษณะเนื้อหาที่สอบไม่ผ่าน ดังนั้น วิธีการใช้โมเดลลำดับชั้นคุณลักษณะจึงเป็นทางเลือกหนึ่ง ที่จะพัฒนาการเรียนรู้อ่านภาษาอังกฤษด้านทักษะการอ่านภาษาอังกฤษ นอกจากนี้ การใช้โมเดลลำดับชั้นคุณลักษณะยังช่วยเพิ่มความตรงเชิงโครงสร้าง (Construct validity) ของแบบทดสอบ เพราะในลำดับชั้นของการพัฒนาชุดข้อสอบ จำเป็นต้องวิเคราะห์และสังเคราะห์เนื้อหาอย่างละเอียดก่อนการกำหนดลำดับชั้นคุณลักษณะเนื้อหาสำหรับการสร้างแบบทดสอบ

สำหรับวิธีการทดสอบวินิจฉัยโดยใช้โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แสดงให้เห็นว่าการกำหนดแผนภาพลำดับชั้นคุณลักษณะ ทำให้การดำเนินการทดสอบมีแนวทางการคัดเลือกข้อสอบจากคลังข้อสอบตามลำดับชั้นตอนที่ชัดเจน ครอบคลุมเนื้อหามากกว่าวิธีการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ทั่วไป ที่เป็นการทดสอบแบบปรับเหมาะบนพื้นฐานของเนื้อหาโดยใช้คอมพิวเตอร์ (Computerized content-based adaptive test) ที่ให้ความสำคัญกับความครอบคลุมเนื้อหา แต่ไม่มีการจัดลำดับชั้นของเนื้อหาทำให้ต้องทดสอบทุกเนื้อหา แต่วิธีการทดสอบวินิจฉัยแบบปรับเหมาะด้วยคอมพิวเตอร์ที่พัฒนาขึ้นในงานวิจัยนี้ มีการนำโมเดลวินิจฉัยทางปัญญามาใช้ในการกำหนดลำดับชั้นคุณลักษณะก่อนการพัฒนาคลังข้อสอบบนโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ทำให้ผู้สอบไม่จำเป็นต้องทดสอบทุกคุณลักษณะ เพราะคุณลักษณะที่กำหนดมีความสัมพันธ์กันเป็นลำดับชั้น ซึ่งมีการจัดเรียงลำดับชั้นคุณลักษณะจากคุณลักษณะที่เป็นพื้นฐานไปยังคุณลักษณะในลำดับชั้นที่ยากกว่า โดยในลำดับชั้นแรกที่ทำให้ทำการทดสอบจะเป็นคุณลักษณะพื้นฐานที่ง่ายซึ่งต้องมีมาก่อนคุณลักษณะในลำดับชั้นถัดไปซึ่งยากขึ้น ดังนั้นกรณีที่ผู้สอบสอบไม่ผ่านคุณลักษณะที่ต้องมีมาก่อนจะสามารถวินิจฉัยได้ว่า การทดสอบคุณลักษณะในลำดับถัดไปก็ไม่ได้ผ่านด้วยเช่นกัน

ผลการพิจารณาจำนวนข้อสอบในคลังข้อสอบแสดงให้เห็นว่า จำนวนข้อสอบในคลังข้อสอบบนโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ขึ้นอยู่กับจำนวนคุณลักษณะและจำนวนข้อสอบในแต่ละคุณลักษณะ ผลการพัฒนาปรากฏว่า ข้อสอบมีจำนวนทั้งหมด 621 ข้อ ซึ่งเป็นไปตามข้อเสนอแนะของ Weiss (2004) ที่กล่าวว่า ทั้งจำนวนข้อสอบในคลังข้อสอบทั้งหมดและจำนวนข้อสอบในคลังข้อสอบจำแนกตามชุดของข้อสอบ จะต้องไม่ต่ำกว่า 100 ถึง 200 ข้อ

สำหรับผลการประเมินวิธีการวินิจฉัยโดยผู้เชี่ยวชาญด้านการสอนภาษาอังกฤษจำนวน 20 คน ประเมินความถูกต้องเหมาะสมของวิธีการวินิจฉัย ผลปรากฏว่า วิธีการวินิจฉัยทักษะการอ่านภาษาอังกฤษบนโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ด้านการออกแบบระบบการทดสอบวินิจฉัย การจัดรูปแบบองค์ประกอบบนหน้าจอการทดสอบ และลักษณะทั่วไปของระบบการจัดการโปรแกรมการทดสอบวินิจฉัยมีความเหมาะสมในระดับดี ส่วนด้านการกำหนดคุณลักษณะการทดสอบวินิจฉัยบนโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ และคู่มือการใช้งานมีความเหมาะสมในระดับดีมาก แสดงให้เห็นว่า วิธีการวินิจฉัยทักษะการอ่านภาษาอังกฤษบนโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่พัฒนาขึ้น สามารถนำไปใช้งานกับกลุ่มเป้าหมายในสถานการณ์จริงได้ ทั้งนี้เนื่องจาก วิธีการวินิจฉัยทักษะการอ่านภาษาอังกฤษบนโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์นี้พัฒนาขึ้นอย่างมีหลักการ แนวคิด ทฤษฎีรองรับ รวมทั้งใช้เทคนิควิธีการที่หลากหลาย เริ่มตั้งแต่ขั้นตอนการพัฒนาชุดข้อสอบ โดยใช้โมเดลลำดับชั้นคุณลักษณะ (Attribute Hierarchy Model: AHM) ของ Leighton, Gierl and Hunka (2002) เป็นแนวคิดในการกำหนดและจำแนกเนื้อหาทักษะการอ่านภาษาอังกฤษ มีการระบุลักษณะและความสัมพันธ์เชิงลำดับชั้นที่มีลักษณะเฉพาะ จากคุณลักษณะพื้นฐานไปยังคุณลักษณะเฉพาะชั้นสูงกว่า ทำให้สามารถวัดกระบวนการทางปัญญาของกลุ่มผู้สอบที่มีลักษณะเฉพาะได้ชัดเจน

ส่วนการพัฒนาวิธีการวินิจฉัยบนโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ มีการใช้แนวคิดโมเดลลำดับชั้นคุณลักษณะร่วมกับทฤษฎีการตอบสนองข้อสอบ ในการกำหนดลำดับชั้นการทำข้อสอบบนโปรแกรมการทดสอบ การเลื่อนลำดับชั้นการทำข้อสอบคุณลักษณะถัดไป จนกระทั่งการยุติการทดสอบ และการรายงานผลการทดสอบ ซึ่งการใช้คอมพิวเตอร์ในการทดสอบทำให้สามารถรายงานผลการทดสอบได้ทันที ตอบโต้กับผู้สอบได้ ทำให้การทดสอบมีความเหมาะสมกับการใช้งานในสภาพปัจจุบันและกลุ่มเป้าหมาย สอดคล้องกับแนวคิดของ Parshall and others (2002) ที่กล่าวไว้ว่า ประโยชน์ที่ได้จากการทดสอบด้วยคอมพิวเตอร์คือ การนำสื่อวัตกรรมการต่าง ๆ มาผสมผสานกันในการจัดทำรูปแบบข้อสอบ การโต้ตอบระหว่างคอมพิวเตอร์กับผู้สอบ สามารถแสดงผลการทดสอบ ทำให้ผู้ใช้ได้ทราบผลคะแนนได้ในทันที ซึ่งสอดคล้องกับแนวคิดของ Cheng (2002) ที่กล่าวไว้ว่า การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ เป็นวิธีการที่ผู้สอบแต่ละคน ไม่จำเป็นต้องทำข้อสอบทุกข้อ ลดข้อจำกัดที่ข้อสอบมีความยากเกินไปหรือง่ายเกินไปสำหรับผู้สอบ เพราะการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ใช้วิธีการประมาณค่าความสามารถของผู้สอบ และคัดเลือกข้อสอบข้อถัดไปเพื่อให้ได้ข้อสอบที่สอดคล้องกับความสามารถของผู้สอบเป็นสำคัญ โดยผู้สอบแต่ละจะจะได้ทำข้อสอบในข้อที่แตกต่างกันตามความสามารถของตนเอง

ข้อเสนอแนะในการนำผลการวิจัยไปใช้

1. วิธีการวินิจฉัยทักษะการอ่านภาษาอังกฤษที่พัฒนาขึ้น ครู อาจารย์ และบุคลากรทางการศึกษาสามารถนำไปใช้วินิจฉัยสภาพความรู้ของนักศึกษาทั้งในระดับอุดมศึกษา และนักเรียนในระดับชั้นมัธยมศึกษา เนื่องจากลักษณะข้อสอบครอบคลุมความรู้ทักษะพื้นฐานการอ่านภาษาอังกฤษทั่วไป
2. กรณีต้องการให้ได้สารสนเทศของผลการวินิจฉัยที่ชัดเจน ถูกต้อง แม่นยำ ควรนำไปวินิจฉัยกับผู้เรียนที่ผ่านการเรียนในเนื้อหาทักษะการอ่านภาษาอังกฤษทั้ง 6 คุณลักษณะนี้มาก่อน
3. บุคลากรทางการศึกษาสามารถนำวิธีการวินิจฉัยที่พัฒนาขึ้นนี้ ไปใช้ในการทดสอบวินิจฉัยเพื่อจำแนกคุณลักษณะทักษะการอ่านภาษาอังกฤษ ที่เป็นความรอบรู้หรือข้อบกพร่องในตัวผู้เรียนเป็นรายบุคคล เพื่อปรับปรุงการจัดการเรียนการสอนภาษาอังกฤษได้
4. ผู้ที่จะนำวิธีการวินิจฉัยที่พัฒนาขึ้นนี้ไปใช้ ควรปรับปรุงข้อสอบในคลังข้อสอบ ตามหลักการจัดทำคลังข้อสอบของทฤษฎีการตอบสนองข้อสอบ

ข้อเสนอแนะในการวิจัยต่อไป

1. การพัฒนาวิธีการวินิจฉัยทักษะการอ่านภาษาอังกฤษโดยใช้โมเดลลำดับชั้นคุณลักษณะและการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่พัฒนาขึ้นนี้ ยังมีข้อจำกัดในเรื่องของความเป็นคุณลักษณะเฉพาะของโปรแกรมสำหรับการทดสอบวินิจฉัยเพียง 6 คุณลักษณะเท่านั้น ดังนั้นควรจะพัฒนาโปรแกรมคอมพิวเตอร์ที่ใช้ในการทดสอบวินิจฉัยแบบปรับเหมาะด้วยคอมพิวเตอร์ที่สามารถรองรับคุณลักษณะ ซึ่งสามารถนำไปใช้กับเนื้อหาที่มีมากกว่าหรือน้อยกว่า 6 คุณลักษณะได้

2. หากต้องการพัฒนาแบบทดสอบโดยใช้โมเดลลำดับชั้นคุณลักษณะในเนื้อหาวิชาอื่น การกำหนดลำดับชั้นคุณลักษณะก่อนพัฒนาแบบทดสอบจะต้องดำเนินการอย่างละเอียดตามหลักการของโมเดลลำดับชั้นคุณลักษณะ และมีการตรวจสอบความถูกต้องทุกขั้นตอน ตั้งแต่ขั้นตอนการกำหนดคุณลักษณะ การจัดเรียงลำดับคุณลักษณะ และการจัดทำแผนภาพลำดับชั้นคุณลักษณะ รวมถึงการกำหนด Q-Matrix ทั้งนี้เพื่อให้การจำแนกค่าพารามิเตอร์ของข้อสอบ โดยเฉพาะอย่างยิ่งค่าความยากของข้อสอบสอดคล้องกับแผนภาพลำดับชั้นคุณลักษณะ จะทำให้วิธีการวินิจฉัยกลุ่มผู้สอบที่มีลักษณะเฉพาะมีความชัดเจนมากยิ่งขึ้น

3. ควรมีการวิจัยเกี่ยวกับการวินิจฉัยโดยใช้โมเดลวิเคราะห์จำแนกประเภทอื่นๆ เช่น Rule-Space Method (RSM), Bayesian Inference Network (BIN), Log linear Cognitive Diagnosis Model (LCDM), General Diagnostic Model (GDM) และ Multiple Classification Latent Class Model (MCLCM) เป็นต้น เพื่อใช้เป็นแนวคิดในการพัฒนาวิธีการวินิจฉัยบนโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ และวิเคราะห์เปรียบเทียบผลของวิธีการวินิจฉัยตามโมเดลที่นำมาใช้ตั้งแต่ 2 โมเดลขึ้นไป

เอกสารอ้างอิง

- ประภาพร ชนะจันะศักดิ์. (2554). *การพัฒนาความเข้าใจในการอ่านภาษาอังกฤษโดยใช้การสอนแบบแลกเปลี่ยนบทบาท ของนักศึกษาระดับปริญญาตรี มหาวิทยาลัยราชภัฏราชชนครินทร์*. วิทยานิพนธ์หลักสูตรการศึกษามหาบัณฑิต. สาขาวิชาหลักสูตรและการสอน คณะศึกษาศาสตร์ มหาวิทยาลัยบูรพา.
- สุพรรณณี พันธุ์พริกส์ และ ฉลิลิกา มหาพูนทอง. (2550). การศึกษาความสามารถ ปัญหาและความต้องการในการใช้ภาษาอังกฤษของนักศึกษาในระดับปริญญาโทและปริญญาเอกของสถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนือ. *วารสารวิชาการพระจอมเกล้าพระนครเหนือ*, 17(3), 67-68.
- Bejar, I. I. (2005). Educational diagnostic assessment. *Journal of Educational Measurement*, 21(2), 175-189.
- Cheng, Y. (2009) When Cognitive Diagnosis Meets Computerized Adaptive Testing: CD-CAT. *Psychometrika*. 74(4), 619-632.
- Gierl, M. J., Zheng, Y. & Cui, Y. (2008), Using the Attribute Hierarchy Method to Identify and Interpret Cognitive Skills that Produce Group Differences. *Journal of Educational Measurement*, 45: 65-89. doi: 10.1111/j.1745-3984.2007.00052.x
- Gierl, M. J., Wang, C., & Zhou, J. (2008). Using the attribute hierarchy method to make diagnostic inferences about examinees' cognitive skills in algebra on the SAT[®]. *Journal of Technology, Learning*, 6, 5-47. Abstract retrieved March, 8, 2011, from: <http://www.jtla.org>.

- Gierl, M. J., Leighton, J. P., & Hunka, S. P. (2000). Exploring the logic of Tatsuoka's rule-space model for test development and analysis. *Educational Measurement: Issue and Practice*, 19(3), 34-44.
- Harnish, D. L., & Linn, R. L. (2005). Analysis of item response patterns: Questionable test data and dissimilar curriculum practices. *Journal of Educational Measurement*. 18(3), 133-146.
- Leighton, J. P., & Gierl, M. J., & Hunka, S. (2002). The attribute hierarchy model: An approach for integrating cognitive theory with assessment practice. Paper presented at the annual meeting of the National Council on Measurement in Education (NCME), New Orleans, Louisiana, USA.
- Leighton, J. P., Gierl, M. J., & Hunka, S. P. (2004). The attribute hierarchy method for cognitive assessment: A variation on Tatsuoka's rule-space approach. *Journal of Education Measurement*, 41(3), 205-237.
- Leighton, J. P., Gierl, M. J., & Hunka, S. P. (2007). Using the attribute hierarchy method to make inferences about examinees' cognitive skills. In M. J. Gierl & J. P. Leighton (Eds.), *Cognitive diagnostic assessment for education*, 242–274.
- Parshall, C. G. et al. (2002) "Innovative Item Types for Computerized Testing," in W. J. Van Der Linden and C.A.W. Glas (Eds.), *Computerized Adaptive Testing: Theory and Practice*. 129–148.
- Tatsuoka, K. K., & Tatsuoka, M. M. (1997). Computerized cognitive diagnostic adaptive testing: Effect on remedial instruction as empirical validation. *Journal of Educational Measurement*, 34(1), 3-20.
- Van der Linden, W. J. (1999). *Computerized educational testing*. In Masters, G.N. and Keeves, J. P. (eds.) *Advances in Measurement in Educational Research and Measurement*. Amsterdam: PERGAMON.
- Wang, C. & Gierl, M. J. (2007). Investigating the cognitive attribute underlying student performance on the SAT Critical Reading Subtest: An Application of the Attribute Hierarchy Method. *Annual Meeting of the National Council on Measurement in Educational (NCME)*. 1-34. Chicago.
- Weiss, D. J. (2004). Computerized adaptive testing for effective and efficient measurement in counseling and education. *Measurement and Evaluation in Counseling and Development*, 37(2), 70-84
- Snow, R. E. & Lohman, D. F. (1989). Implications of cognitive psychology for educational measurement. In R. Linn (Ed.), *Educational measurement* (3rd ed). New York: Macmillan.