



บทที่ 1

บทนำ

## ความเป็นมาและความสำคัญของปัญหา

แนวทางการศึกษาตั้งแต่ระดับอนุบาล จนถึงอุดมศึกษาในปัจจุบันได้มีการตื่นตัวที่จะนำเอาระบบคลังข้อสอบเข้ามาช่วยในการพัฒนาแบบสอบวัดผลสัมฤทธิ์ทางการเรียนให้มีคุณภาพ และสามารถเลือกแบบสอบได้ตามจุดมุ่งหมายของการนำไปใช้อย่างสะดวกและรวดเร็ว โดยเฉพาะระดับมัธยมศึกษาใช้การจัดทำคลังข้อสอบเป็นเกณฑ์ข้อหนึ่งในการประเมินมาตรฐานของโรงเรียน (กนก จันทร์ขจร 2538: 3) ในระดับอุดมศึกษาเองก็ได้สนับสนุนส่งเสริมให้มีการจัดทำคลังข้อสอบกันอย่างจริงจัง เช่นที่ สถาบันภาษา จุฬาลงกรณ์มหาวิทยาลัย มหาวิทยาลัยเกษตรศาสตร์ และมหาวิทยาลัยสงขลานครินทร์ เป็นต้น แต่การเลือกข้อสอบจากคลังข้อสอบที่ใช้วัดผลสัมฤทธิ์ทางการเรียนที่เป็นอยู่ในปัจจุบันได้ชีวิตกำหนดจำนวนข้อสอบจากตารางวิเคราะห์หลักสูตรตามเวลาที่ใช้ในการสอบโดยจำแนกตามเนื้อหาหรือจุดประสงค์การเรียนรู้เป็นหลัก ล้วนข้อที่เลือกจะเป็นข้อที่สุ่มได้จากข้อสอบที่มีนิสสัยของค่าคุณลักษณะของข้อสอบที่ยอมรับว่าใช้ได้โดยไม่คำนึงถึงความสามารถของผู้สอบ

เนื่องจากคลังข้อสอบเป็นที่รวมของข้อสอบซึ่งผ่านการวิเคราะห์และคัดเลือกข้อที่มีคุณภาพดีเป็นที่ยอมรับได้ไว้เป็นจำนวนมากจึงสามารถนำมาใช้ให้เป็นประโยชน์ต่าง ๆ กัน เช่นสามารถสร้างแบบสอบใหม่ได้ตามรูปแบบความต้องการและเหมาะสมกับจุดประสงค์ที่ตั้งไว้ตามจำนวนข้อสอบและเนื้อหาที่ต้องการ สามารถสร้างแบบสอบคู่ขนานซึ่งวัดสิ่งเดียวกัน โดยใช้ข้อสอบต่างชุดกันแต่มีคุณลักษณะของข้อสอบใกล้เคียงกันทำให้วัดความสามารถผู้สอบได้ในมิติเดียวกันโดยไม่ต้องใช้ข้อสอบซ้ำกัน และยังสามารถช่วยพัฒนาแบบสอบให้มีคุณภาพมากขึ้นอีกด้วย ประกอบกับปัจจุบันได้มีการนำเอาคอมพิวเตอร์เข้ามาจัดการบริหารข้อมูลคลังข้อสอบ ทำให้การจัดการเกี่ยวกับการสร้างแบบสอบเป็นไปได้ด้วยความสะดวกรวดเร็วยิ่งขึ้น จากที่

กล่าวมาจะเห็นได้ว่าคลังข้อสอบ มีประโยชน์หลายประการ แต่อย่างไรก็ตามในระบบคลังข้อสอบ ก็ยังประสบปัญหาที่สำคัญคือวิธีการเลือกข้อสอบมาประกอบเป็นแบบสอบ

วิธีการเลือกข้อสอบจากคลังข้อสอบเพื่อประกอบเป็นแบบสอบที่ใช้อยู่ในปัจจุบันเลือกตามเกณฑ์ของผู้เชี่ยวชาญแต่ละคนที่เสนอแนะว่าใช้ได้ ซึ่งในทฤษฎีการสอบแบบคลาสสิก (Classical Test Theory) ใช้ค่าความยาก ( $p$ ) อยู่ในช่วง 0.2 ถึง 0.8 และค่าอำนาจจำแนก ( $r$ ) มากกว่า .2 ในทฤษฎีการตอบสนองต่อข้อสอบ (Item Response Theory) ใช้ค่าความยาก ( $b$ ) อยู่ในช่วง -2.5 ถึง 2.5 ค่าอำนาจจำแนก ( $a$ ) อยู่ในช่วง 0.5 ถึง 2.5 และค่าการเดา ( $c$ ) อยู่ในช่วง 0 ถึง 0.3 (สัพพันธ์ สกมลสันต์ 2538: 98) วิธีการเลือกข้อสอบที่ใช้ในปัจจุบันนี้เป็นวิธีที่พิจารณาเฉพาะเกณฑ์ในการเลือกคุณลักษณะของข้อสอบในลักษณะกว้างๆ ไม่คำนึงถึงกลุ่มความสามารถกลุ่มใดกลุ่มหนึ่งโดยเฉพาะ การกำหนดเกณฑ์อย่างกว้างๆ เป็นการยากที่จะวัดความสามารถของผู้สอบที่แตกต่างกันให้ได้ข้อมูลหรือสารสนเทศมากที่สุดเหมือนกันทุกๆ กลุ่มความสามารถของผู้สอบ นั่นคือไม่สามารถประมาณความสามารถได้อย่างถูกต้อง จากการศึกษาของแฮมเบิลตัน (Hambleton 1985 : 246) พบว่าการที่เลือกข้อสอบโดยวิธีการสุ่มจะทำให้ได้แบบสอบที่ให้ค่าสารสนเทศต่ำสุด แฮมเบิลตันและโรเจอร์ (Hambleton and Roger 1983: 355-367) พบว่าวิธีการเลือกข้อสอบโดยการสุ่มจะต้องใช้จำนวนข้อสอบมากกว่าวิธีการเลือกข้อสอบตามวิธีค่าสารสนเทศสูงสุดถึง 2 เท่าตัวจึงจะได้สารสนเทศเท่าเทียมกัน และศึกษานำร่องโดยผู้วิจัยใช้ข้อมูลของโครงการสร้างแบบทดสอบมาตรฐานวิชาภาษาไทย ระดับชั้นมัธยมศึกษาปีที่ 1 ของโรงเรียนเอกชนสอนศาสนาอิสลามในจังหวัดชายแดนภาคใต้ จำนวน 2,000 คน ทำการสอบแบบสอบละ 60 ข้อ พบว่ารูปร่างของโค้งสารสนเทศของแบบสอบที่ได้จากการสุ่มข้อสอบมีลักษณะใกล้เคียงกับรูปร่างของโค้งสารสนเทศของคลังข้อสอบ ผลจากการเลือกข้อสอบที่ทำให้ได้โค้งสารสนเทศของแบบสอบดังกล่าวทำให้เกิดข้อเสียบางประการคือถ้าโค้งสารสนเทศของคลังข้อสอบมีค่าสูงสุดสำหรับผู้สอบที่มีระดับความสามารถสูง การเลือกข้อสอบแบบสุ่มข้อสอบจากช่วงค่าคุณลักษณะข้อสอบที่ยอมรับและใช้อยู่ในปัจจุบันจะได้แบบสอบที่มีค่าสารสนเทศของแบบสอบสูงสุดที่ระดับความสามารถของผู้สอบสูงด้วย ทำให้ข้อสอบที่นำไปใช้วัดความสามารถของผู้สอบได้คลาดเคลื่อนน้อยที่สุดที่ผู้มีความสามารถสูง แต่ความเป็นจริงแล้วผู้สอบที่จะนำ

แบบสอบมาตรฐานที่สร้างขึ้นไปใช้นั้น เป็นกลุ่มผู้สอบที่มีความสามารถต่ำ ที่ถูกต้องจึงควรที่จะเลือกข้อสอบที่มีค่าสารสนเทศสูงสุดที่ระดับความสามารถของผู้สอบต่ำ เพื่อที่จะได้วัดความสามารถของผู้สอบได้คลาดเคลื่อนน้อยที่สุดที่ระดับความสามารถต่ำ นั่นคือจะได้ข้อมูลหรือสารสนเทศเกี่ยวกับผู้สอบมากที่สุดตรงระดับความสามารถของผู้สอบต่ำ ดังนั้นในเรื่องคลังข้อสอบก็มีประเด็นให้หาข้อที่วิธีการเลือกข้อสอบจากคลังข้อสอบให้ได้แบบสอบที่เหมาะสมกับความสามารถของกลุ่มผู้สอบ

การเลือกข้อสอบจากคลังข้อสอบเพื่อประกอบเป็นแบบสอบให้เหมาะกับความสามารถของกลุ่มผู้สอบเป็นสิ่งสำคัญเพราะถ้าแบบสอบไม่เหมาะกับความสามารถของกลุ่มผู้สอบจะทำให้การวัดความสามารถของผู้สอบมีความคลาดเคลื่อนมาก ซึ่งวิธีการเลือกข้อสอบเพื่อประกอบเป็นแบบสอบในปัจจุบันไม่ได้คำนึงถึงว่าแบบสอบเหมาะกับความสามารถของกลุ่มผู้สอบหรือไม่ นอกจากนี้การเลือกข้อสอบเพื่อประกอบเป็นแบบสอบให้เหมาะกับความสามารถของผู้สอบยังสามารถนำไปประยุกต์ใช้ในการสอบที่มีจุดมุ่งหมายที่แตกต่างกันออกไป

การสอบที่มีจุดมุ่งหมายแตกต่างกันขึ้นอยู่กับให้นำแบบสอบไปใช้ การสอบเป็นการตรวจสอบความก้าวหน้าของพัฒนาการด้านต่างๆของผู้สอบ การสอบสามารถทำหน้าที่ได้หลายอย่าง เช่น เพื่อการคัดเลือก (selection) เพื่อการจำแนก (classification) เพื่อการบรรจุ (placement) ผู้สอบ เพื่อให้การรับรองผลการเรียน (certification) เพื่อการตัดสินใจเลื่อนชั้นหรือข้ามชั้น (summative) เพื่อการทำนายผลการเรียนต่อไป (prediction) เพื่อการรายงานผลการเรียนให้กับผู้เกี่ยวข้องเพื่อเป็นข้อมูลย้อนกลับ (feedback) ให้กับครูเพื่อปรับปรุงการสอนของตนเอง และเพื่อเป็นข้อมูลสำหรับการบริหารการจัดการและการวิจัย (อุทุมพร จามรมา 2532: 1) ถ้าเป็นแบบสอบวัดผลสัมฤทธิ์วัตถุประสงค์ในการใช้อาจแตกต่างกัน เช่น เพื่อใช้วัดความรู้ก่อน-หลังเรียน เพื่อระบุความรู้ของผู้เรียน เพื่อหาจุดอ่อน-แข็งของผู้เรียน และเพื่อประเมินความก้าวหน้าในการเรียนวิชานั้นๆ แบบสอบที่ใช้เพื่อวัดก่อนการเรียนเป็นการวัดพฤติกรรมที่มีมาก่อน ส่วนการวัดหลังเรียนเป็นการวัดพฤติกรรมภายหลังที่เกิดการเรียนรู้อแล้ว ดังนั้นข้อสอบที่วัดก่อนเรียนจึงมักจะสร้างเพื่อวัดความพร้อมที่จะเรียน เป็นข้อสอบที่ค่อนข้างง่ายและวัดพฤติกรรม

ที่จำเป็นก่อนที่จะเรียน ส่วนการวัดหลังเรียนจึงมักจะสร้างให้ยากง่ายปานกลาง และวัดได้ครอบคลุมพฤติกรรมตามที่กำหนดไว้ในจุดมุ่งหมายของวิชา แบบสอบที่วัดความรู้เป็นการวัดพฤติกรรมที่เกิดขึ้นในระดับที่ต่ำสุดซึ่งอาจรวมความรู้พื้นฐานและทักษะที่จำเป็น แบบสอบเพื่อวินิจฉัยใช้วัดจุดอ่อนแข็งของนักเรียน จึงจำเป็นต้องเป็นข้อสอบที่มีขอบเขตกว้างและมุ่งทดสอบสิ่งที่นักเรียนมีปัญหในการเรียนรู้ ดังนั้นแบบสอบนี้จึงสร้างเพื่อให้ครอบคลุมเนื้อหาวิชาไม่ใช่ลุ่มเนื้อหาวิชา แบบสอบที่ใช้วัดความก้าวหน้าทางการเรียนมักใช้เพื่อตัดสินว่านักเรียนบรรลุวัตถุประสงค์ที่ต้องการมากน้อยต่างกันหรือเพื่อตัดสินความสามารถโดยเทียบกันในกลุ่ม ลักษณะแบบสอบจึงสร้างโดยเน้นความสำคัญในเรื่องให้ข้อสอบเป็นตัวแทนของเนื้อหาวิชาและพฤติกรรมที่แสดงผลการเรียนรู้แบบสอบนี้จะต้องมี ความยากพอสมควรที่จะกระจายนักเรียนออกได้ตามพฤติกรรมผลการเรียน (อุทุมพร จามรมา 2530: 58-59) ดังนั้นจะเห็นว่าในการสอบที่มีจุดมุ่งหมายแตกต่างกันออกไปจำเป็นต้องใช้ข้อสอบที่ประกอบเป็นแบบสอบที่แตกต่างกัน ซึ่งข้อค้นพบว่าการเลือกข้อสอบเพื่อประกอบเป็นแบบสอบให้เหมาะสมกับความสามารถของกลุ่มในระดับต่างๆ สามารถที่จะประยุกต์ใช้ในการสอบที่มีจุดมุ่งหมายต่างๆได้

จากการที่กล่าวมาแล้วข้างต้นจะเห็นว่าในเรื่องคลังข้อสอบมีประเด็นปัญหาอยู่ที่วิธีการเลือกข้อสอบจากคลังข้อสอบเพื่อให้ได้แบบสอบที่เหมาะสมกับระดับความสามารถของผู้สอบ และจุดมุ่งหมายของการนำไปใช้ ซึ่งยังไม่มีเกณฑ์ในการเลือกข้อสอบมาประกอบเป็นแบบสอบที่ชัดเจน ในอดีตที่ผ่านมาก็ได้มีผู้พยายามคิดค้นหาวิธีในการพัฒนาเกณฑ์ในการเลือกข้อสอบมาประกอบเป็นแบบสอบให้มีคุณภาพยิ่ง ๆ ขึ้นไป และที่มีผลกระทบต่อการสร้างแบบสอบมาก ลอร์ด (1977b) ซึ่งได้แนวความคิดมาจาก เฮอร์นชวม (1968) (อ้างใน Hambleton and Swaminathan 1985: 229) ได้เสนอแนะแนวทางในการสร้างแบบสอบไว้ดังนี้

- 1) พิจารณาถึงฟังก์ชันสารสนเทศของแบบสอบที่ต้องการ ซึ่งลอร์ดเรียกว่า ฟังก์ชันสารสนเทศเป้าหมาย (Target Information Function)
- 2) เลือกข้อสอบที่มีค่าฟังก์ชันสารสนเทศของข้อสอบที่เมื่อนำค่าสารสนเทศที่แต่ละจุดของความสามารถมารวมกันแล้วจะได้ค่าสารสนเทศของแบบสอบใกล้เคียงกับฟังก์ชันสารสนเทศเป้าหมาย

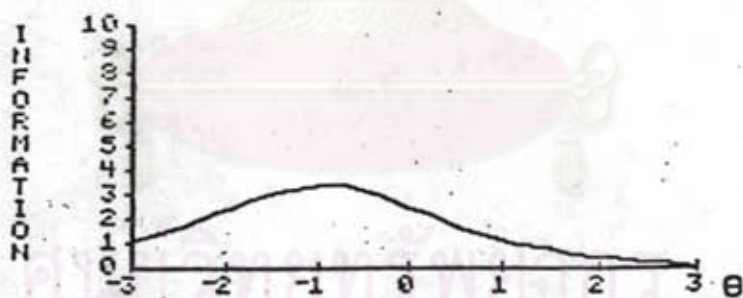
ไรท์และสโตน (1979) ได้เขียนหนังสือการออกแบบสอบที่ดีที่สุด (Best Test Design) โดยแนะนำการเลือกข้อสอบจากคลังข้อสอบมาประกอบเป็นแบบสอบ โดยเลือกข้อสอบที่มีโค้งสารสนเทศของข้อสอบ (Item Information Curve) ที่สามารถมาประกอบกันเป็นโค้งสารสนเทศของแบบสอบเป้าหมาย (Target Test Information Curve) โดยแนะนำว่าทุกครั้งที่จะออกแบบแบบสอบจะต้องกำหนดโค้งสารสนเทศเป้าหมายเสียก่อนเพราะว่าโค้งสารสนเทศเป้าหมายตามทฤษฎีการตอบสนองต่อข้อสอบเป็นค่าที่บอกกว่าที่ระดับความสามารถของผู้สอบเท่าใดแบบสอบนี้จะให้สารสนเทศเท่าใด ถ้าให้ค่าสารสนเทศสูงที่ระดับความสามารถใดก็แสดงว่าความคลาดเคลื่อนในการประมาณค่าความสามารถนั้นน้อย และค่าสารสนเทศต่ำที่ระดับความสามารถใดก็แสดงว่าความคลาดเคลื่อนในการประมาณค่าความสามารถนั้นมาก นอกจากนี้แอมเบิลตันและสวามินาทาน (1985) ได้เสนอไว้คือเมื่อกำหนดโค้งสารสนเทศของแบบสอบเป้าหมายแล้ว ก็เลือกข้อสอบเพื่อที่จะเติมเต็มพื้นที่โค้งสารสนเทศเป้าหมายที่กำหนดโดยบวกค่าสารสนเทศของข้อสอบที่แต่ละระดับความสามารถเมื่อเลือกข้อสอบเพิ่มขึ้นจนได้ผลรวมของค่าสารสนเทศใกล้เคียงกับค่าสารสนเทศของแบบสอบเป้าหมายจึงหยุด ซึ่งวิธีการนั้นนอกจากคุณลักษณะของข้อสอบแต่ละข้อที่เลือกมาประกอบกันเป็นแบบสอบให้ค่าค่าสารสนเทศของข้อสอบแตกต่างกันที่แต่ละระดับความสามารถแล้วจำนวนข้อสอบที่ประกอบกันเป็นแบบสอบมีจำนวนไม่คงที่อีกด้วย เพราะขึ้นอยู่กับค่าสารสนเทศของข้อสอบที่ถูกเลือกขึ้นมาทำให้การเติมเต็มพื้นที่โค้งสารสนเทศเป้าหมายเป็นไปได้ยาก

จากการศึกษาที่ผ่านมาและข้อเสนอแนะของบุคคลต่างๆเกี่ยวกับการเลือกข้อสอบมาประกอบเป็นแบบสอบ ก็ยังไม่ได้ให้เกณฑ์ที่ชัดเจนในการเลือกข้อสอบ ผู้วิจัยจึงสนใจที่จะพัฒนาเกณฑ์ในการเลือกข้อสอบเพื่อเสนอแนวทางในการเลือกข้อสอบเพื่อประกอบเป็นแบบสอบให้เหมาะสมกับความสามารถของผู้สอบโดยใช้โค้งปกติมาเป็นโค้งสารสนเทศเป้าหมายเพราะลักษณะของโค้งปกติมีค่าฟังก์ชันสารสนเทศสูงและครอบคลุมระดับความสามารถที่กว้างสอดคล้องกับความสามารถของคนที่มีการแจกแจงแบบปกติ

เนื่องจากข้อสอบที่ผ่านการวิเคราะห์ตามทฤษฎีการตอบสนองต่อข้อสอบมีคุณลักษณะ

ที่แน่นอนและค่าสารสนเทศของแบบสอบที่ได้ก็จะเพิ่มขึ้นตามจำนวนข้อสอบที่เพิ่มขึ้นซึ่งเป็นไปตามหลักการที่ว่าข้อสอบเป็นตัวอย่างของโดเมนความรู้ ยิ่งสอบมากข้อก็จะยิ่งทำให้ประมาณค่าความสามารถที่แท้จริงได้ถูกต้องมากขึ้นเท่านั้น และจากแนวคิดที่ว่าลักษณะของแบบสอบที่ใช้วัดผลสัมฤทธิ์ทางการเรียนของนักเรียนได้คั้นนั้นจะต้องมีค่าสารสนเทศของแบบสอบที่เหมาะสมกับความสามารถของผู้สอบกล่าวคือค่าสารสนเทศสูงสุดที่ค่าความสามารถระดับใดก็จะวัดความสามารถของผู้สอบได้ถูกต้องแม่นยำที่สุดที่ความสามารถนั้น (Hambleton 1977: 66 อ้างใน สุนันต์ สุกมลสันต์ 2530: 55) ดังนั้นวิธีการเลือกข้อสอบที่ใช้ก็ควรจะเป็นวิธีการที่เลือกข้อสอบที่ประกอบกันเป็นแบบสอบที่ให้ค่าฟังก์ชันสารสนเทศใกล้เคียงกับ ค่าฟังก์ชันสารสนเทศเป้าหมาย

ฟังก์ชันสารสนเทศของข้อสอบ (Item Information Function) เป็นฟังก์ชัน แสดงความสัมพันธ์ของค่าสารสนเทศ (information) และความสามารถ (ability) ของผู้สอบ ดังภาพที่ 1

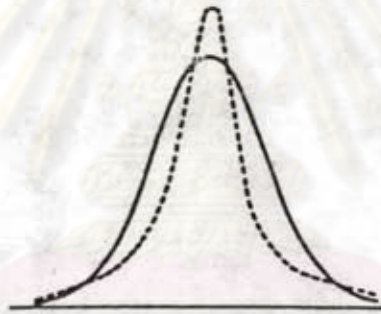


ภาพที่ 1 ฟังก์ชันสารสนเทศของแบบสอบ

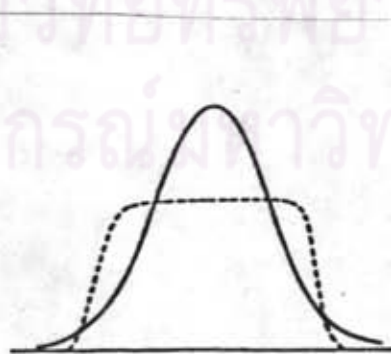
จากภาพที่ 1 จะเห็นว่าจำนวนค่าสารสนเทศสูงสุดที่ระดับความสามารถประมาณ  $-1$  และเมื่อพิจารณาความสามารถที่อยู่ระหว่าง  $-2 < \theta < 0$  จากนิสัยนี้ความสามารถของผู้สอบถูกประมาณด้วยความแม่นยำไม่เท่ากันในแต่ละระดับของความสามารถ ถ้าค่าสารสนเทศของแบบสอบสูงที่ระดับความสามารถใดแบบสอบจะประมาณค่าความสามารถนั้นด้วยความคลาดเคลื่อนต่ำ นอกเหนือจากนิสัยนี้ค่าสารสนเทศลดลงอย่างรวดเร็วการประมาณค่าความสามารถจะมีความคลาดเคลื่อนสูง

ดังนั้นฟังก์ชันสารสนเทศจะบอกเราว่าแต่ละระดับความสามารถถูกประมาณค่าความสามารถได้ดีเพียงใด นอกจากนี้ฟังก์ชันสารสนเทศไม่ขึ้นกับการแจกแจงของความสามารถของผู้สอบ ฟังก์ชันสารสนเทศในอุดมคติควรจะครอบคลุมระดับความสามารถที่กว้างและมีค่าสารสนเทศสูง แต่ในทางปฏิบัติถ้าค่าสารสนเทศสูงที่ระดับความสามารถใดก็จะครอบคลุมพิสัยของระดับความสามารถแคบ โค้งสารสนเทศที่ครอบคลุมพิสัยค่าความสามารถที่กว้างก็จะให้โค้งสารสนเทศต่ำ

ผู้วิจัยเห็นว่าลักษณะโค้งปกตินอกจากจะเป็นตัวแทนของลักษณะการแจกแจงคุณลักษณะทางจิตวิทยาแล้วยังเหมาะที่จะเป็นตัวแทนค่าสารสนเทศที่สูงขณะเดียวกันก็ครอบคลุมระดับความสามารถที่กว้าง จึงใช้โค้งปกติมาเป็นโค้งสารสนเทศเป้าหมาย ดังนี้



ภาพที่ 2ก โค้งปกติมาตรฐานเทียบกับโค้งสารสนเทศสูงครอบคลุมพิสัยแคบ



ภาพที่ 2ข โค้งปกติมาตรฐานเทียบกับโค้งสารสนเทศต่ำครอบคลุมพิสัยกว้าง

จากการที่กล่าวมาทั้งหมดจะเห็นได้ว่าคลังข้อสอบมีประโยชน์หลายอย่าง และในประเทศไทยนั้นคลังข้อสอบได้มีบทบาทสำคัญต่อการวัดและประเมินผลมากขึ้น แต่ในเรื่องคลังข้อสอบก็ยังมีประเด็นปัญหาอยู่ที่วิธีการเลือกข้อสอบจากคลังข้อสอบให้ได้แบบสอบที่เหมาะสมกับความสามารถของกลุ่มผู้สอบและจุดมุ่งหมายของการนำไปใช้ และจากการวิจัยที่ผ่านมายังไม่มีข้อสรุปในเรื่องนี้ ดังนั้นผู้วิจัยจึงสนใจที่จะสร้างเกณฑ์ในการเลือกข้อสอบที่เหมาะสมกับความสามารถของกลุ่มผู้สอบเพื่อให้ได้ข้อสรุปที่มีประโยชน์ต่อผู้ที่เกี่ยวข้องกับการวัดและประเมินผล จะได้เป็นแนวทางในการเลือกข้อสอบที่จะนำไปใช้วัดได้ถูกต้องที่สุดกับกลุ่มที่มีความสามารถต่าง ๆ กันต่อไป

### วัตถุประสงค์ของการวิจัย

การวิจัยครั้งนี้มีวัตถุประสงค์ในการพัฒนาเกณฑ์ในการเลือกข้อสอบจากค่าพารามิเตอร์ของข้อสอบในคลังข้อสอบ เพื่อให้ได้แบบสอบที่เหมาะสมกับกลุ่มความสามารถของผู้สอบ โดยมีวัตถุประสงค์เฉพาะ 2 ประการดังนี้

1. สร้างเกณฑ์ในการเลือกข้อสอบจากคลังข้อสอบโดยพิจารณาจากค่าพารามิเตอร์ของข้อสอบที่ทำให้ได้แบบสอบที่เหมาะสมกับกลุ่มผู้สอบที่มีความสามารถปานกลาง ความสามารถสูง และความสามารถต่ำ
2. ตรวจสอบเกณฑ์ในการเลือกข้อสอบที่สร้างขึ้นโดยเปรียบเทียบค่าประสิทธิภาพสัมพัทธ์ของแบบสอบระหว่างแบบสอบที่ได้จากการสุ่มข้อสอบตามเกณฑ์ที่สร้างขึ้นกับแบบสอบที่ได้จากการสุ่มข้อสอบตามเกณฑ์ที่ใช้อยู่ในปัจจุบันโดยแปรค่าจำนวนข้อสอบเป็น 20, 30, 60, และ 100 ข้อ

### สมมติฐานของการวิจัย

จากการศึกษาของแฮมเบิลตัน (Hambleton and Swaminathan 1985 : 228-229)

ได้แนะนำวิธี



การเลือกข้อสอบไว้ว่า นอกจากจะพิจารณาในแง่ของความตรงเชิงเนื้อหา (Content Validity) แล้วจะต้องเลือกข้อที่มีค่าอำนาจจำแนกสูง ส่วนค่าความยากขึ้นอยู่กับความสามารถของกลุ่มผู้สอบที่จะนำข้อสอบไปใช้โดยคาดการณ์ล่วงหน้าถึงความสามารถของกลุ่มบุคคลที่จะไปสอบ และเนื่องจากค่าการเดาในโมเดลของทฤษฎีการตอบสนองต่อข้อสอบหมายถึงโอกาสที่ผู้ที่มีความสามารถต่ำจะตอบถูก ดังนั้นค่าการเดาของข้อสอบที่เหมาะสมกับผู้สอบที่มีความสามารถสูงน่าจะต่ำ (Warr 1978: 47) ค่าการเดาของข้อสอบที่เหมาะสมกับกลุ่มผู้สอบที่มีความสามารถปานกลางน่าจะมีค่าปานกลาง และค่าการเดาของข้อสอบที่เหมาะสมกับกลุ่มผู้สอบที่มีความสามารถต่ำน่าจะมีค่าสูง

ดังนั้นผู้วิจัยจึงตั้งสมมติฐานของเกณฑ์การเลือกข้อสอบจากคลังข้อสอบมาประกอบเป็นแบบสอบเพื่อให้ได้แบบสอบที่เหมาะสมกับกลุ่มผู้สอบที่มีความสามารถต่างๆกันโดยเกณฑ์ในการแบ่งระดับของค่าพารามิเตอร์ของข้อสอบออกเป็นระดับสูง ปานกลาง และต่ำ ด้วยตำแหน่งเปอร์เซ็นต์ไทล์ที่ 25 และ 75 (รายละเอียดอยู่ในคำนิยามเชิงปฏิบัติการ) ดังต่อไปนี้

1. การเลือกข้อสอบโดยใช้ข้อสอบที่มีค่าอำนาจจำแนกอยู่ในระดับสูง ( $a > 0.85$ ) ค่าการเดาอยู่ในระดับปานกลาง ( $.15 < c < .25$ ) และค่าความยากอยู่ในระดับปานกลาง ( $-0.67 < b < 0.67$ ) น่าจะให้ค่าสารสนเทศสูงสำหรับกลุ่มผู้สอบที่มีความสามารถปานกลาง
2. การเลือกข้อสอบโดยใช้ข้อสอบที่มีค่าอำนาจจำแนกอยู่ในระดับสูง ( $a > 0.85$ ) ค่าการเดาอยู่ในระดับต่ำ ( $c < .15$ ) และค่าความยากอยู่ในระดับสูง ( $b > 0.67$ ) น่าจะให้ค่าสารสนเทศสูงสำหรับกลุ่มผู้สอบที่มีความสามารถสูง
3. การเลือกข้อสอบโดยใช้ข้อสอบที่มีค่าอำนาจจำแนกอยู่ในระดับสูง ( $a > 0.85$ ) ค่าการเดาอยู่ในระดับสูง ( $c > .25$ ) และค่าความยากอยู่ในระดับต่ำ ( $b < -0.67$ ) น่าจะให้ค่าสารสนเทศสูงสำหรับกลุ่มผู้สอบที่มีความสามารถต่ำ

4. แบบสอบที่เลือกข้อสอบตามเกณฑ์ที่สร้างขึ้นน่าจะมีประสิทธิภาพสูงกว่าแบบสอบที่สุ่มข้อสอบจากเกณฑ์ที่ใช้อยู่ในปัจจุบันสำหรับกลุ่มผู้สอบที่มีความสามารถแตกต่างกันที่ขนาดความยาวของแบบสอบจำนวน 20 30 60 และ 100 ข้อ

#### ขอบเขตการวิจัย

1. ประชากรข้อสอบที่ใช้การวิจัยครั้งนี้เป็นค่าพารามิเตอร์คุณลักษณะข้อสอบที่สร้างขึ้นด้วยฟังก์ชันทางคณิตศาสตร์โดยทำการจำลองแบบค่าที่เป็นไปได้ของข้อสอบตามแนวทฤษฎีการตอบสนองต่อข้อสอบจากโมเดลโลจิสติกสามพารามิเตอร์ ซึ่งประกอบด้วยค่าอำนาจจำแนก ค่าความยาก และค่าการเดา ในช่วงที่ยอมรับได้ว่าเป็นข้อสอบที่มีคุณภาพพอที่จะเก็บไว้ในคลังข้อสอบได้ (สัทพันธ์ สกมลลันต์ 2530 : 135) คือ

ค่าอำนาจจำแนก อยู่ระหว่าง 0.5 ถึง 2.5

ค่าความยาก อยู่ระหว่าง - 3 ถึง + 3

ค่าการเดา อยู่ระหว่าง 0 ถึง 0.8

โดยการแปรค่าที่เป็นไปได้จากค่าต่ำสุดถึงสูงสุดเพิ่มขึ้นครั้งละ 0.05 จะได้ข้อสอบจำนวน  $41 \times 121 \times 7 = 34,727$  ข้อ

2. การแบ่งกลุ่มสูง กลาง และต่ำของค่าพารามิเตอร์ของข้อสอบในคลังข้อสอบที่สร้างขึ้นใช้เปอร์เซ็นต์ไคที่ 25 และ 75 เป็นจุดแบ่งเพราะถือว่าค่าพารามิเตอร์ที่อยู่ระดับปานกลางมีโอกาสประมาณร้อยละ 50 มีค่าอยู่ในช่วงดังกล่าว ค่า  $a$  การแจกแจงไคได้เคียงกับค่าไคกำลังสอง (Chi-square) เมื่อกำหนดจุดแบ่งตามเกณฑ์ที่กำหนดแล้วจะได้ว่า  $a > 0.85$  เป็นค่าอำนาจจำแนกระดับสูง  $0.55 \leq a \leq 0.85$  เป็นค่าอำนาจจำแนกระดับปานกลาง  $a < 0.55$  เป็นค่าอำนาจจำแนกระดับต่ำ ค่า  $b$  มีการแจกแจงไคได้เคียงกับโค้งปกติมาตรฐาน เมื่อกำหนดจุดแบ่งตามเกณฑ์ที่กำหนดแล้วจะได้ว่า  $b > 0.67$  เป็นค่าความยากระดับสูง  $-0.67 \leq b \leq 0.67$  เป็นค่าความยากระดับปานกลาง  $b < -0.67$  เป็นค่าความยากระดับต่ำ และค่า  $c$  มีการแจกแจงแบบเบต้า (Beta distribution) เมื่อกำหนดจุดแบ่งตามเกณฑ์ที่กำหนดแล้วจะได้ว่า  $c > .25$  เป็นค่าการเดาระดับสูง  $.15 \leq c \leq .25$  เป็นค่าการเดาระดับปานกลาง  $c < .15$  เป็นค่าการเดาระดับต่ำ

### 3. ตัวแปรที่ศึกษาในการวิจัยนี้ มีดังต่อไปนี้

#### 3.1 ตัวแปรตาม (dependent variables) ได้แก่

- ก) ค่าสารสนเทศของข้อสอบ
- ข) ค่าสารสนเทศของแบบสอบ
- ค) ค่าประสิทธิภาพสัมพัทธ์ของแบบสอบ

#### 3.2 ตัวแปรอิสระ (independent variables) ได้แก่

- ก) ค่าอำนาจจำแนก (a)
- ข) ค่าความยาก (b)
- ค) ค่าการเดา (c)
- ง) วิธีการเลือกข้อสอบ ได้แก่ วิธีการเลือกข้อสอบตามเกณฑ์ที่สร้างขึ้น กับ การลุ่มข้อสอบตามเกณฑ์ที่ใช้กันอยู่ในปัจจุบัน
- จ) ความยาวของแบบสอบ ได้แก่ 20 30 60 และ 100 ข้อ

4. ทฤษฎีการวัดผลที่ใช้สำหรับการพัฒนาเกณฑ์ในการเลือกข้อสอบในการวิจัยครั้งนี้ คือ ทฤษฎีการตอบสนองต่อข้อสอบ เนื่องจากจุดเด่นที่สำคัญก็คือค่าคุณลักษณะข้อสอบที่วิเคราะห์ได้เป็นค่าพารามิเตอร์และสามารถนำไปคำนวณค่าสารสนเทศได้ ประกอบกับ แฮมเบิลตันและเดอกริวเตอร์ (Hambleton & De Gruijter 1983: 152-159) ได้แนะนำว่าวิธีคลาสสิก (Classical Method) ไม่สามารถให้คำตอบในการเลือกข้อสอบที่เหมาะสมกับค่าสารสนเทศเป้าหมายได้ แต่โมเดลการตอบสนองต่อข้อสอบ (Item Response Models) จะได้คำตอบที่เป็นประโยชน์มากกว่าเพราะโมเดลนี้สามารถที่จะวางตำแหน่งข้อสอบ คนและคะแนนจุดตัดในสเกลเดียวกันและค่าพารามิเตอร์ไม่ขึ้นกับกลุ่มผู้สอบ นอกจากนี้จากการศึกษาจากยูรี (Urry) (Werm 1982 : 20) พบสรุปได้ว่าโมเดลการตอบสนองต่อข้อสอบบรรยายคุณลักษณะของข้อสอบได้ชัดเจนที่สุดก็คือโมเดลสามพารามิเตอร์ และเนื่องจากการวิจัยครั้งนี้มุ่งที่จะนำไปใช้กับข้อสอบหลายตัวเลือกซึ่งโอกาสในการเดาตัวเลือกน่าจะมีส่วนสำคัญด้วย การวิจัยครั้งนี้จึงใช้โมเดลโลจิสติกสามพารามิเตอร์

### ข้อตกลงเบื้องต้นของการวิจัย

1. ค่าพารามิเตอร์ของข้อสอบที่จำลองขึ้นครอบคลุมข้อสอบที่มีคุณสมบัตินั้นเป็นที่ยอมรับได้ว่ามีคุณภาพที่จะเก็บไว้ในคลังข้อสอบเพื่อประโยชน์ในการเลือกไปประกอบเป็นแบบสอบตามจุดประสงค์การนำไปใช้ที่เห็นไปได้ทั้งหมดของโมเดลโลจิสติกสามพารามิเตอร์ การวิจัยครั้งนี้จึงถือว่าค่าพารามิเตอร์ของข้อสอบที่จำลองขึ้นเป็นตัวแทนของค่าพารามิเตอร์ของข้อสอบในคลังข้อสอบได้เป็นอย่างดี

2. ความสามารถของประชากรผู้สอบมีการแจกแจงแบบปกติ (Hambleton and Swaminathan 1985: 232) ดังนั้นผู้วิจัยจึงกำหนดให้ความสามารถของกลุ่มประชากรผู้สอบมีการแจกแจงแบบปกติ โดยแบ่งเป็น 3 กลุ่มประชากรใหญ่ ๆ คือ กลุ่มประชากรผู้สอบที่มีความสามารถสูง ปานกลาง และต่ำ และกำหนดให้โค้งปกติมาตรฐานเป็นตัวแทนการกระจายค่าความสามารถของกลุ่มผู้สอบ โดยมีค่าเฉลี่ยเป็น 1, 0 และ -1 ตามลำดับ การแบ่งกลุ่มดังกล่าวทำให้ค่าเฉลี่ยของกลุ่มทั้ง 3 แตกต่างกันอย่างมีนัยสำคัญทางสถิติ ( $p=0.15$ ) และสอดคล้องกับสภาพความเป็นจริงที่ว่าความสามารถของผู้สอบในแต่ละกลุ่มผู้สอบที่มีความสามารถแตกต่างกันจะมีส่วนที่ซ้ำซ้อนกันอยู่บ้าง และจากการศึกษานำร่องจากข้อมูลจริงจากโครงการสร้างแบบสอบมาตรฐานวิชาภาษาไทยของนักเรียนเอกชนสอนศาสนาอิสลามในจังหวัดชายแดนภาคใต้ซึ่งเป็นที่ยอมรับว่าเป็นกลุ่มประชากรที่มีความสามารถทางภาษาไทยต่ำกว่าค่าเฉลี่ยของความสามารถของกลุ่มผู้สอบเท่ากับ -1 ผู้วิจัยจึงกำหนดให้ประชากรผู้สอบที่เป็นตัวแทนของผู้มีความสามารถปานกลางมีค่าเฉลี่ยของค่าความสามารถ ( $\theta$ ) เท่ากับ 0 ประชากรผู้สอบที่เป็นตัวแทนของผู้มีความสามารถสูงมีค่าเฉลี่ยของค่าความสามารถเท่ากับ 1 ประชากรผู้สอบที่เป็นตัวแทนของผู้มีความสามารถต่ำมีค่าเฉลี่ยของค่าความสามารถเท่ากับ -1

3. โค้งสารสนเทศข้อสอบเป้าหมายกำหนดให้มีลักษณะเป็นโค้งปกติมาตรฐาน เพราะค่าฟังก์ชันสารสนเทศเป้าหมายในอุดมคติที่ต้องการคือ ฟังก์ชันสารสนเทศที่ทำให้ค่าสารสนเทศสูงและครอบคลุมระดับความสามารถที่กว้างและเพื่อให้สอดคล้องกับการแจกแจง

ความสามารถของกลุ่มความสามารถของประชากร จึงกำหนดให้ค่าสารสนเทศของกลุ่มประชากร  
ที่มีความสามารถปานกลาง สูง และต่ำ มีค่าสารสนเทศที่แต่ละค่าความสามารถเท่ากับค่าแกนนตั้ง  
(ordinate) ของโค้งปกติมาตรฐานของแต่ละกลุ่มเป้าหมาย ดังตาราง 2 (นราศรี มดุงชีวิต  
2510: 560)

ตาราง 1 แสดงค่าสารสนเทศของข้อสอบเป้าหมายตามกลุ่มความสามารถ  
ของผู้สอบและระดับความสามารถของผู้สอบ

๑	ค่าสารสนเทศข้อสอบเป้าหมายตามกลุ่มความสามารถของผู้สอบ		
	ปานกลาง	สูง	ต่ำ
-4.00	0.0001	0.0001	0.0044
-3.95	0.0002	0.0001	0.0051
-3.90	0.0002	0.0001	0.0060
-3.85	0.0002	0.0001	0.0069
-3.80	0.0003	0.0001	0.0079
-3.75	0.0004	0.0001	0.0091
-3.70	0.0004	0.0001	0.0104
-3.65	0.0005	0.0001	0.0119
-3.60	0.0006	0.0001	0.0136
-3.55	0.0007	0.0001	0.0154
-3.50	0.0009	0.0001	0.0175
-3.45	0.0010	0.0001	0.0198
-3.40	0.0012	0.0001	0.0224
-3.35	0.0015	0.0001	0.0252
-3.30	0.0017	0.0001	0.0283

ตาราง 1 (ต่อ) แสดงค่าสารสนเทศของข้อสอบเป้าหมายตามกลุ่มความสามารถ  
ของผู้สอบและระดับความสามารถของผู้สอบ

	ค่าสารสนเทศข้อสอบเป้าหมายตามกลุ่มความสามารถของผู้สอบ		
	ปานกลาง	สูง	ต่ำ
-3.25	0.0020	0.0001	0.0317
-3.20	0.0024	0.0001	0.0355
-3.15	0.0028	0.0001	0.0396
-3.10	0.0033	0.0001	0.0440
-3.05	0.0038	0.0001	0.0488
-3.00	0.0044	0.0001	0.0540
-2.95	0.0051	0.0002	0.0596
-2.90	0.0060	0.0002	0.0656
-2.85	0.0069	0.0002	0.0721
-2.80	0.0079	0.0003	0.0790
-2.75	0.0091	0.0004	0.0863
-2.70	0.0104	0.0004	0.0940
-2.65	0.0119	0.0005	0.1023
-2.60	0.0136	0.0006	0.1109
-2.55	0.0154	0.0007	0.1200
-2.50	0.0175	0.0009	0.1295
-2.45	0.0198	0.0010	0.1394
-2.40	0.0224	0.0012	0.1497
-2.35	0.0252	0.0015	0.1606
-2.30	0.0283	0.0017	0.1714

ตาราง 1 (ต่อ) แสดงค่าสารสนเทศของข้อสอบเป้าหมายตามกลุ่มความสามารถ  
ของผู้สอบและระดับความสามารถของผู้สอบ

	ค่าสารสนเทศข้อสอบเป้าหมายตามกลุ่มความสามารถของผู้สอบ		
	ปานกลาง	สูง	ต่ำ
-2.25	0.0317	0.0020	0.1826
-2.20	0.0355	0.0024	0.1942
-2.15	0.0396	0.0028	0.2059
-2.10	0.0440	0.0033	0.2179
-2.05	0.0488	0.0038	0.2299
-2.00	0.0540	0.0044	0.2420
-1.95	0.0596	0.0051	0.2541
-1.90	0.0656	0.0060	0.2661
-1.85	0.0721	0.0069	0.2780
-1.80	0.0790	0.0079	0.2897
-1.75	0.0863	0.0091	0.3011
-1.70	0.0940	0.0104	0.3123
-1.65	0.1023	0.0119	0.3230
-1.60	0.1109	0.0136	0.3332
-1.55	0.1200	0.0154	0.3429
-1.50	0.1295	0.0175	0.3521
-1.45	0.1394	0.0198	0.3605
-1.40	0.1497	0.0224	0.3683
-1.35	0.1606	0.0252	0.3752
-1.30	0.1714	0.0283	0.3814

ตาราง 1 (ต่อ) แสดงค่าสารสนเทศของข้อสอบเป้าหมายตามกลุ่มความสามารถ  
ของผู้สอบและระดับความสามารถของผู้สอบ

	ค่าสารสนเทศข้อสอบเป้าหมายตามกลุ่มความสามารถของผู้สอบ		
	ปานกลาง	สูง	ต่ำ
-1.25	0.1826	0.0317	0.3867
-1.20	0.1942	0.0355	0.3910
-1.15	0.2059	0.0396	0.3945
-1.10	0.2179	0.0440	0.3970
-1.05	0.2299	0.0488	0.3984
-1.00	0.2420	0.0540	0.3989
-0.95	0.2541	0.0596	0.3984
-0.90	0.2661	0.0656	0.3970
-0.85	0.2780	0.0721	0.3945
-0.80	0.2897	0.0790	0.3910
-0.75	0.3011	0.0863	0.3867
-0.70	0.3123	0.0940	0.3814
-0.65	0.3230	0.1023	0.3752
-0.60	0.3332	0.1109	0.3683
-0.55	0.3429	0.1200	0.3605
-0.50	0.3521	0.1295	0.3521
-0.45	0.3605	0.1394	0.3429
-0.40	0.3683	0.1497	0.3332
-0.35	0.3752	0.1606	0.3230
-0.30	0.3814	0.1714	0.3123



ตาราง 1 (ต่อ) แสดงค่าสารสนเทศของข้อสอบเป้าหมายตามกลุ่มความสามารถ  
ของผู้สอบและระดับความสามารถของผู้สอบ

	๘ ค่าสารสนเทศข้อสอบเป้าหมายตามกลุ่มความสามารถของผู้สอบ		
	ปานกลาง	สูง	ต่ำ
-0.25	0.3867	0.1826	0.3011
-0.20	0.3910	0.1942	0.2897
-0.15	0.3945	0.2059	0.2780
-0.10	0.3970	0.2179	0.2661
-0.05	0.3984	0.2299	0.2541
0.00	0.3989	0.2420	0.2420
0.05	0.3984	0.2541	0.2299
0.10	0.3970	0.2661	0.2179
0.15	0.3945	0.2780	0.2059
0.20	0.3910	0.2897	0.1942
0.25	0.3867	0.3011	0.1826
0.30	0.3814	0.3123	0.1714
0.35	0.3752	0.3230	0.1606
0.40	0.3683	0.3332	0.1497
0.45	0.3605	0.3429	0.1394
0.50	0.3521	0.3521	0.1295
0.55	0.3429	0.3605	0.1200
0.60	0.3332	0.3683	0.1109
0.65	0.3230	0.3752	0.1023
0.70	0.3123	0.3814	0.0940

ตาราง 1 (ต่อ) แสดงค่าสารสนเทศของข้อสอบเป้าหมายตามกลุ่มความสามารถ  
ของผู้สอบและระดับความสามารถของผู้สอบ

๘ ค่าสารสนเทศข้อสอบเป้าหมายตามกลุ่มความสามารถของผู้สอบ

	ปานกลาง	สูง	ต่ำ
0.75	0.3011	0.3867	0.0863
0.80	0.2897	0.3910	0.0790
0.85	0.2780	0.3945	0.0721
0.90	0.2661	0.3970	0.0656
0.95	0.2541	0.3984	0.0596
1.00	0.2420	0.3989	0.0540
1.05	0.2299	0.3984	0.0488
1.10	0.2179	0.3970	0.0440
1.15	0.2059	0.3945	0.0396
1.20	0.1942	0.3910	0.0355
1.25	0.1826	0.3867	0.0317
1.30	0.1714	0.3814	0.0283
1.35	0.1606	0.3752	0.0252
1.40	0.1497	0.3683	0.0224
1.45	0.1394	0.3605	0.0198
1.50	0.1295	0.3521	0.0175
1.55	0.1200	0.3429	0.0154
1.60	0.1109	0.3332	0.0136
1.65	0.1023	0.3230	0.0119
1.70	0.0940	0.3123	0.0104

ตาราง 1 (ต่อ) แสดงค่าสารสนเทศของข้อสอบเป้าหมายตามกลุ่มความสามารถ  
ของผู้สอบและระดับความสามารถของผู้สอบ

	ค่าสารสนเทศข้อสอบเป้าหมายตามกลุ่มความสามารถของผู้สอบ		
	ปานกลาง	สูง	ต่ำ
1.75	0.0863	0.3011	0.0091
1.80	0.0790	0.2897	0.0079
1.85	0.0721	0.2780	0.0069
1.90	0.0656	0.2661	0.0060
1.95	0.0596	0.2541	0.0051
2.00	0.0540	0.2420	0.0044
2.05	0.0488	0.2299	0.0038
2.10	0.0440	0.2179	0.0033
2.15	0.0396	0.2059	0.0028
2.20	0.0355	0.1942	0.0024
2.25	0.0317	0.1826	0.0020
2.30	0.0283	0.1714	0.0017
2.35	0.0252	0.1606	0.0015
2.40	0.0224	0.1497	0.0012
2.45	0.0198	0.1394	0.0010
2.50	0.0175	0.1295	0.0009
2.55	0.0154	0.1200	0.0007
2.60	0.0136	0.1109	0.0006
2.65	0.0119	0.1023	0.0005
2.70	0.0104	0.0940	0.0004

ตาราง 1 (ต่อ) แสดงค่าสารสนเทศของข้อสอบเป้าหมายตามกลุ่มความสามารถ  
ของผู้สอบและระดับความสามารถของผู้สอบ

๑ ค่าสารสนเทศข้อสอบเป้าหมายตามกลุ่มความสามารถของผู้สอบ

	ปานกลาง	สูง	ต่ำ
2.75	0.0091	0.0863	0.0004
2.80	0.0079	0.0790	0.0003
2.85	0.0069	0.0721	0.0002
2.90	0.0060	0.0656	0.0002
2.95	0.0051	0.0596	0.0002
3.00	0.0044	0.0540	0.0001
3.05	0.0038	0.0488	0.0001
3.10	0.0033	0.0440	0.0001
3.15	0.0028	0.0396	0.0001
3.20	0.0024	0.0355	0.0001
3.25	0.0020	0.0317	0.0001
3.30	0.0017	0.0283	0.0001
3.35	0.0015	0.0252	0.0001
3.40	0.0012	0.0224	0.0001
3.45	0.0010	0.0198	0.0001
3.50	0.0009	0.0175	0.0001
3.55	0.0007	0.0154	0.0001
3.60	0.0006	0.0136	0.0001
3.65	0.0005	0.0119	0.0001
3.70	0.0004	0.0104	0.0001

ตาราง 2 (ต่อ) แสดงค่าสารสนเทศของข้อสอบเป้าหมายตามกลุ่มความสามารถ  
ของผู้สอบและระดับความสามารถของผู้สอบ

0	ค่าสารสนเทศข้อสอบเป้าหมายตามกลุ่มความสามารถของผู้สอบ		
	ปานกลาง	สูง	ต่ำ
3.75	0.0004	0.0091	0.0001
3.80	0.0003	0.0079	0.0001
3.85	0.0002	0.0069	0.0001
3.90	0.0002	0.0060	0.0001
3.95	0.0002	0.0051	0.0001
4.00	0.0001	0.0044	0.0001

#### คำนิยามเชิงปฏิบัติการที่ใช้ในการวิจัย

1. เกณฑ์การเลือกข้อสอบ หมายถึง ช่วงค่าพารามิเตอร์ของข้อสอบในการเลือกข้อสอบที่สามารถยึดเป็นหลักในการเลือกข้อสอบเพื่อใช้ประกอบกันเป็นแบบสอบ เกณฑ์การเลือกข้อสอบสำหรับการวิจัยครั้งนี้ หมายถึง ช่วงของค่าพารามิเตอร์ของข้อสอบตามแนวทฤษฎีการตอบสนองต่อข้อสอบโมเดลโลจิสติกสามพารามิเตอร์ อันได้แก่ ค่าอำนาจจำแนก (a) ค่าความยาก (b) และค่าการเดา (c) ที่ทำให้ค่าสารสนเทศของข้อสอบใกล้เคียงกับค่าสารสนเทศเป้าหมาย สำหรับผู้มีความสามารถสูง ปานกลาง และต่ำ ตามลำดับ

2. ข้อสอบที่เหมาะสมกับความสามารถของผู้สอบ หมายถึง ข้อสอบที่สามารถวัดความสามารถที่แท้จริงของผู้สอบได้โดยมีความคลาดเคลื่อนน้อย นั่นคือ ข้อสอบที่มีค่าสารสนเทศที่ใกล้เคียงกับค่าสารสนเทศเป้าหมาย กล่าวคือ ข้อสอบที่เหมาะสมกับผู้สอบที่มีความ

สามารถปานกลาง สูง และต่ำ เป็นข้อสอบที่มีค่าสารสนเทศใกล้เคียงกับค่าสารสนเทศเป้าหมาย สำหรับผู้มีความสามารถปานกลาง สูง และต่ำ ตามลำดับ

3. ระดับความสามารถ (๑) หมายถึง คักยภาพของผู้สอบที่ประมาณได้จากการทำแบบสอบตามทฤษฎีการตอบสนองต่อข้อสอบโดยปกติแล้วมีค่า  $-3 < \theta < 3$  และระดับความสามารถเป็นตัวแปรเชิงเส้นที่มีการแจกแจงแบบปกติมาตรฐาน (Swaminathan and Gifford 1985: 350) โดย  $\theta$  มีค่าเฉลี่ยเท่ากับ 0 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1

4. วิธีการเลือกข้อสอบตามเกณฑ์ที่สร้างขึ้น หมายถึง การลุ่มข้อสอบมาประกอบเป็นแบบสอบโดยพิจารณาค่านารามิเตอร์ที่วิเคราะห์โดยทฤษฎีการตอบสนองต่อข้อสอบโมเดลโลจิสสามนารามิเตอร์ที่มีค่าอำนาจจำแนก ค่าความยาก และค่าการเดา ตามเกณฑ์ที่สร้างขึ้นที่เหมาะสมกับระดับความสามารถของกลุ่มผู้สอบ

5. วิธีการลุ่มข้อสอบตามเกณฑ์ที่ใช้อยู่ในปัจจุบัน หมายถึง การลุ่มข้อสอบมาประกอบเป็นแบบสอบโดยพิจารณาค่านารามิเตอร์ที่วิเคราะห์โดยทฤษฎีการตอบสนองต่อข้อสอบโมเดลโลจิสสามนารามิเตอร์ที่มีค่าอำนาจจำแนกอยู่ในระดับสูงกว่า 0.5 ค่าการเดาอยู่ในระดับต่ำกว่า 0.3 และค่าความยากอยู่ในระดับช่วง -2.5 ถึง 2.5

6. ประสิทธิภาพสัมพันธ์ของแบบสอบ หมายถึง อัตราส่วนระหว่างค่าสารสนเทศจากแบบสอบ เช่น แบบสอบชุด  $y$  ซึ่งมีค่าสารสนเทศของแบบสอบเท่ากับ  $I(\theta, y)$  กับแบบสอบชุด  $x$  ซึ่งมีค่าสารสนเทศของแบบสอบเท่ากับ  $I(\theta, x)$  ประสิทธิภาพสัมพันธ์ของแบบสอบ  $y$  เมื่อเทียบกับแบบสอบ  $x$  สามารถคำนวณได้จากสูตรดังนี้

$$RE(y, x) = \frac{I(\theta, y)}{I(\theta, x)}$$

โดยแบบสอบ  $x$  และ  $y$  เป็นแบบสอบที่ต่างกันแต่วัดความสามารถเดียวกัน หรือ แบบสอบ  $x$  และ  $y$  เป็นแบบสอบเดียวกันแต่ให้คะแนนคนละวิธี ประสิทธิภาพสัมพัทธ์นิยามเฉพาะเมื่อ  $\theta$  ใน  $I(\theta, y)$  เป็น  $\theta$  เดียวกันใน  $I(\theta, x)$  ในการวิจัยครั้งนี้ใช้ประสิทธิภาพสัมพัทธ์สำหรับการเปรียบเทียบแบบสอบที่สร้างขึ้นตามเกณฑ์ที่พัฒนาขึ้นกับแบบสอบที่สร้างขึ้นจากวิธีการสุ่มข้อสอบตามเกณฑ์ที่ใช้อยู่ในปัจจุบัน โดยแบบสอบทั้งสองมีจำนวน 20 30 60 และ 100 ข้อ ค่า  $RE(y, x)$  มากกว่า 1 ที่ระดับความสามารถ  $\theta$  ใด หมายความว่าแบบสอบ  $y$  ให้ค่าสารสนเทศสูงกว่าแบบสอบ  $x$  ที่ระดับความสามารถนั้น ถ้าค่า  $RE(y, x)$  น้อยกว่า 1 ที่ระดับความสามารถ  $\theta$  ใด หมายความว่าแบบสอบ  $x$  ให้ค่าสารสนเทศสูงกว่าแบบสอบ  $y$  ที่ระดับความสามารถนั้น

7. คลังข้อสอบ หมายถึงฐานข้อมูลที่รวมเอาข้อสอบที่มีค่าพารามิเตอร์ของข้อสอบ โมเดลโลจิสติกสามพารามิเตอร์ ได้แก่ ค่าอำนาจจำแนก (a) ค่าความยาก (b) และค่าการเดา (c) ในการวิจัยครั้งนี้ ประกอบด้วย คลังข้อสอบที่ได้จากการจำลองค่าพารามิเตอร์โดยกำหนดให้ค่าอำนาจจำแนกมีค่าตั้งแต่ 0.5 ถึง 2.5 ค่าความยากมีค่าตั้งแต่ -3.00 ถึง 3.00 และค่าการเดามีค่าตั้งแต่ 0.0 ถึง 0.3 โดยแปรค่าครั้งละ 0.05 รวมแล้วได้จำนวนค่าพารามิเตอร์แทนคุณลักษณะของข้อสอบที่แตกต่างกันทั้งสิ้น  $41 \times 121 \times 7 = 34,727$  ข้อ

8. ค่าอำนาจจำแนก(a)หมายถึง ค่าความชันของโค้งคุณลักษณะของข้อสอบที่จุดเปลี่ยนโค้งจึงเป็นจุดแสดงความสามารถในการจำแนก ผู้สอบที่มีความสามารถต่ำและสูง ออกจากกันได้อย่างเด่นชัด ในทางทฤษฎี  $-\alpha < a < \alpha$  แต่ในทางปฏิบัตินิยมใช้ค่า  $.5 < a < 2.5$  (Ree 1979 : 372) เนื่องจากค่าอำนาจจำแนก(a)เป็นตัวแปรเชิงสุ่ม (random variable) มีค่าที่ควรจะเป็นและนำมาใช้ได้ (Baker 1985 : 16) อยู่ระหว่าง 0 ถึง  $\alpha$  และเป็นค่าความชันของโค้งคุณลักษณะของข้อสอบที่จุดเปลี่ยนโค้ง (inflection point) ค่า  $a$  จึงมีการแจกแจงใกล้เคียงกับค่าไคกำลังสอง (Chi-square) ดังนี้

$$f(a_j | v_j, w_j) \sim a_j^{v_j-1} \exp[-a_j^2/2w_j]$$

เมื่อ  $v_j$  : องศาแห่งความเป็นอิสระ  
 $w_j$  : ค่าเสก (scale parameter)

ปกติกำหนดให้  $v_j = 10$  และ  $w_j = 0.1$  จึงได้ค่าตัวแปรเชิงสุ่ม  $a$  ที่เหมาะสมตามทฤษฎี (Swaminathan and Gifford 1985 : 356) โดยมีค่ามัธยเลขคณิตเท่ากับ ค่าองศาแห่งความเป็นอิสระและค่าความแปรปรวนเท่ากับสองเท่าขององศาแห่งความเป็นอิสระ (บุญเสริม วิสกุล 2517 : 159) ซึ่งในที่นี้จะแบ่งเป็น 3 ระดับ คือ ระดับสูง ปานกลาง และต่ำ โดยใช้ตำแหน่งเปอร์เซ็นต์ไทล์ที่ 25 และ 75 เป็นจุดแบ่งได้ดังนี้  $a > 0.85$  เป็นค่าอำนาจจำแนกระดับสูง  $0.55 \leq a < 0.85$  เป็นค่าอำนาจจำแนกระดับปานกลางและ  $a < 0.55$  เป็นค่าอำนาจจำแนกระดับต่ำ

9. ค่าความยาก (b) หมายถึงระดับความยากของข้อสอบซึ่งมีค่าเท่ากับระดับความสามารถของผู้สอบ (θ) ที่มีโอกาสทำข้อสอบนั้นถูก  $(1+c)/2$  โดยปกตินิยมใช้ ค่า  $-3 < b < 3$  เนื่องจากค่าความยาก (b) เป็นตัวแปรเชิงสุ่มมีค่าอยู่ระหว่าง  $-\alpha$  ถึง  $\alpha$  โดยปกติข้อสอบส่วนใหญ่จะมีความยากปานกลาง และลดหล่นไปตามค่าความยากที่สูงขึ้นและตามค่าความยากที่ลดลงในลักษณะสมมาตร ค่า b จึงมีการแจกแจงใกล้เคียงกับโค้งปกติมาตรฐาน (Swaminathan and Gifford 1985 : 350) โดยมีค่ามัธยเลขคณิตเท่ากับ 0 และค่าความแปรปรวนเท่ากับ 1 ซึ่งในที่นี้จะแบ่งเป็น 3 ระดับ คือ ระดับสูง ปานกลาง และต่ำ โดยใช้ตำแหน่งเปอร์เซ็นต์ไทล์ที่ 25 และ 75 เป็นจุดแบ่งได้ดังนี้  $b > 0.67$  เป็นค่าความยากระดับสูง  $-0.67 < b < 0.67$  เป็นค่าความยากระดับปานกลาง และ  $b < -0.67$  เป็นค่าความยากระดับต่ำ

10. ค่าการเดา (c) หมายถึง ค่าโอกาสที่ผู้สอบที่มีความสามารถต่ำจะสามารถทำข้อสอบข้อนั้นได้ถูกต้องอย่างน้อยเพียงใด โดยปกติค่า  $0 < c < 0.3$  เนื่องจากค่าการเดา (c) เป็นตัวแปรเชิงสุ่มมีค่าอยู่ระหว่าง 0 ถึง 1 ค่า c จึงมีการแจกแจงใกล้เคียงกับค่าเบต้า (Beta distribution) ดังนี้

$$f(c_j | s_j, t_j) \sim c_j^{s_j} (1 - c_j)^{t_j}$$



เมื่อ  $s_j$  และ  $t_j$  : ค่าเสกส (scale parameter)

ปกติกำหนดให้  $s_j = 2$  และ  $t_j = 12$  จึงได้ค่าตัวแปรเชิงลุ่ม  $c$  ที่เหมาะสมตามทฤษฎี

(Swaminathan and Gifford 1986 : 589-601) โดยมีค่ามัชฌิมเลขคณิตเท่ากับ

$$\frac{s_j+1}{s_j+t_j+2} \text{ และค่าความแปรปรวนเท่ากับ } \frac{(s_j+1)(t_j+1)}{(s_j+t_j+2)^2(s_j+t_j+3)}$$

(บุญเสริม วิสกุล 2517 : 160)

ซึ่งในที่นี้จะแบ่งเป็น 3 ระดับ คือ ระดับสูง ปานกลาง และต่ำ โดยใช้ตำแหน่งเปอร์เซ็นต์ไทล์ ที่ 25 และ 75 เป็นจุดแบ่งได้ดังนี้  $c > .25$  เป็นค่าการเคาระะดับสูง  $.15 < c < .25$  เป็นค่าการเคาระะดับปานกลาง และ  $c < .15$  เป็นค่าการเคาระะดับต่ำ

11. คำสารสนเทศของข้อสอบ หมายถึง ความแม่นยำในการประมาณค่าความสามารถจริงด้วยคะแนนจากการตอบแบบสอยซึ่งมีค่าเท่ากับอัตราส่วนกำลังสองของความชันของเส้นถดถอยบนคะแนนความสามารถจริงต่อความแปรปรวนของคะแนน โดยมีสูตรดังนี้

$$I(\theta, u_i) = [P'_i(\theta)]^2 P_i(\theta) Q_i(\theta)$$

หรือ

$$I(\theta, u_i) = \frac{(1.7a_i)^2 (1-c_i)}{(c_i + e^{1.7=1(\theta-b_i)}) (1 + e^{-1.7=1(\theta-b_i)})^2}$$

เมื่อ  $P'_i(\theta)$  คือ ความชันของโค้งคุณลักษณะของข้อสอบที่ระดับความสามารถ  $\theta$

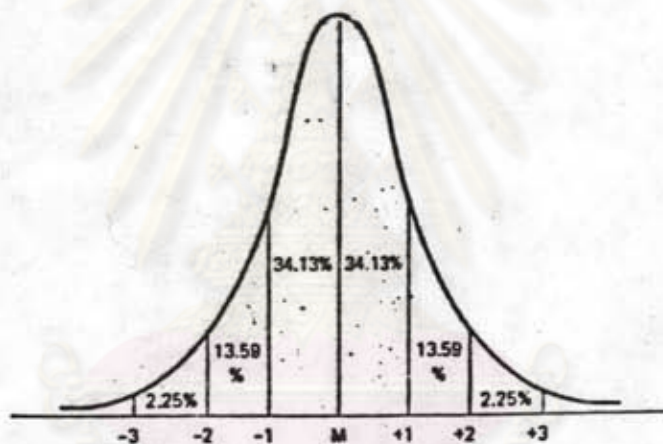
$P_i(\theta)$  คือ ความน่าจะเป็นที่ผู้สอยที่มีความสามารถ  $\theta$  จะตอบข้อสอบข้อที่  $i$  ถูก

$Q_i(\theta)$  คือ  $1 - P_i(\theta)$

12. คำสารสนเทศของแบบสอย หมายถึงผลรวมของคำสารสนเทศของข้อสอบสำหรับระดับความสามารถของผู้สอยต่างๆกัน ถ้ามีค่าสูงที่ระดับความสามารถของผู้สอยใดแสดงว่าแบบสอยนั้น สามารถจำแนกระดับความสามารถของผู้สอยได้ดี ณ ระดับความสามารถนั้น (Hambleton 1977 : 66 อ้างใน สุนันต์ สุกมลสันต์ 2530 : 56) การวิจัยนี้เป็นการรวม

ค่าสารสนเทศของข้อสอบฉบับต่างๆที่มีความยาว 20 30 60 และ 100 ข้อ ณ ระดับความสามารถจาก -3 ถึง +3

13. โค้งปกติมาตรฐาน หมายถึง ฟังก์ชันความสัมพันธ์ของตัวแปรเชิงสุ่มกับโอกาสที่จะเกิดค่าของตัวแปรเชิงสุ่มนั้น มีลักษณะเป็นรูประฆังคว่ำแบบสมมาตรตรงค่ามัธยฐานเลขคณิตเท่ากับ 0 และมีส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1 ความเ้เท่ากับ 0 และความโค้งเท่ากับ 3 และมีพื้นที่ใต้โค้ง ดังนี้



สำหรับการวิจัยครั้งนี้ใช้โค้งปกติมาตรฐานเป็นตัวแทนของโค้งสารสนเทศของข้อสอบเป้าหมาย

15. โค้งสารสนเทศของข้อสอบเป้าหมาย หมายถึง โค้งสารสนเทศของข้อสอบที่ใช้เป็นเกณฑ์ในการเปรียบเทียบกับค่าสารสนเทศของข้อสอบที่มีค่าหาค่ามัธยฐานของข้อสอบที่แตกต่างกัน ในการวิจัยครั้งนี้ได้กำหนดโค้งสารสนเทศเป้าหมายของข้อสอบ เป็น 3 รูปแบบ เพื่อให้เป็นตัวแทนของลักษณะข้อสอบที่เหมาะสมกับผู้สอบที่มีระดับความสามารถต่างกัน โดยกำหนดโค้งสารสนเทศของข้อสอบเป้าหมายที่เหมาะสมกับกลุ่มผู้สอบที่มีความสามารถปานกลาง เป็นลักษณะโค้งปกติมาตรฐาน โค้งสารสนเทศเป้าหมายของข้อสอบที่เหมาะสมกับกลุ่มผู้สอบที่มีความสามารถสูง เป็นลักษณะเหมือนโค้งปกติมาตรฐานแต่มีค่ามัธยฐานเลขคณิตเท่ากับ 1 โค้งสารสนเทศเป้าหมายของข้อสอบที่เหมาะสมกับกลุ่มผู้สอบที่มีความสามารถต่ำ เป็นลักษณะเหมือนโค้งปกติมาตรฐานแต่มีค่ามัธยฐานเลขคณิตเท่ากับ -1

16. ความคลาดเคลื่อนมาตรฐานในการประมาณค่า ( $SE(\theta)$ ) หมายถึง อัตราส่วน ผกผันของค่าสารสนเทศของแบบสอป ( $I(\theta)$ ) ณ ระดับความสามารถ ( $\theta$ ) ต่างๆ ซึ่งเขียน เป็นสูตรได้ดังนี้

$$SE(\theta) = \frac{1}{\sqrt{I(\theta)}}$$

### ความไม่สมบูรณ์ของการวิจัย

การวิจัยครั้งนี้มุ่งพัฒนาเกณฑ์การเลือกข้อสอบจึงไม่ได้ครอบคลุมไปถึงความตรงเชิง เนื้อหาของแบบสอป ดังนั้นการที่จะนำเกณฑ์ที่ได้จากการวิจัยไปใช้จะต้องทราบจำนวนข้อสอบ ที่เป็นตัวแทนแต่ละเนื้อหาหรือจุดประสงค์การเรียนรู้ตามสัดส่วนของเนื้อหาในตารางวิเคราะห์ หลักสูตรเสียก่อนจึงสุ่มข้อสอบตามจำนวนที่ต้องการจากเกณฑ์ที่พัฒนาขึ้น

### ประโยชน์ที่คาดว่าจะได้รับจากการวิจัย

1. ช่วยให้ได้เกณฑ์ในการเลือกข้อสอบหรือสุ่มข้อสอบจากคลังข้อสอบเพื่อจะได้ ตัวอย่างข้อสอบที่สามารถนำไปใช้ประกอบเป็นแบบสอปที่เหมาะสมกับการนำไปใช้กับกลุ่มเป้าหมาย ผู้สอบที่มีความสามารถแตกต่างกัน ได้แก่ กลุ่มผู้สอบที่มีความสามารถปานกลาง กลุ่มผู้สอบที่มีความสามารถสูง และกลุ่มผู้สอบที่มีความสามารถต่ำ
2. เป็นแนวทางในการบริหารคลังข้อสอบให้ครบวงจร โดยเฉพาะการบริหาร คลังข้อสอบด้วยระบบคอมพิวเตอร์ทำให้สามารถเลือกข้อสอบไปใช้ได้ได้อย่างเหมาะสมยิ่งขึ้น
3. เป็นแนวทางในการสร้างแบบสอปให้ตรงกับจุดมุ่งหมายของการสอบแต่ละครั้ง เช่น การสอบคัดเลือกบุคคล การสอบวัดผลสัมฤทธิ์ การสอบวินิจฉัย เป็นต้น
4. เพื่อเป็นแนวทางในการศึกษาถึงวิธีการคัดเลือกข้อสอบจากคลังข้อสอบ เพื่อประกอบเป็นแบบสอป ตามเป้าหมายอื่น ๆ และจุดมุ่งหมายในการนำแบบสอปไปใช้ต่อไป