



245527



รายงานการวิจัย

เรื่อง

การพัฒนาแบบสอบผลสัมฤทธิ์ทางการเรียน
วิชาการวัดและประเมินผลในชั้นเรียนโดยกำหนด
คะแนนมาตรฐานด้วยวิธีปฏิกิริยา

ดร.สุริพร อนุศาสนนันท์

สนับสนุนโดย

ทุนอุดหนุนการวิจัย คณะศึกษาศาสตร์

มหาวิทยาลัยบูรพา

ประจำปี พ.ศ. 2552

คณะศึกษาศาสตร์ มหาวิทยาลัยบูรพา

2554



รายงานการวิจัย
เรื่อง



การพัฒนาแบบสอบผลสัมฤทธิ์ทางการเรียนวิชาการวัดและประเมินผล
ในชั้นเรียนโดยกำหนดคะแนนมาตรฐานด้วยวิธีบุ้คมาร์ค
A Development of Measurement and Assessment in Classroom
Achievement Test: Using Bookmark Standard Setting

ดร.สุรียพร อนุศาสนนันท์

สนับสนุนโดย
ทุนอุดหนุนการวิจัย คณะศึกษาศาสตร์ มหาวิทยาลัยบูรพา
ประจำปี พ.ศ. 2552

คณะศึกษาศาสตร์ มหาวิทยาลัยบูรพา
2554

งานวิจัยนี้ได้รับทุนอุดหนุนการวิจัย
จากคณะศึกษาศาสตร์ มหาวิทยาลัยบูรพา
ประจำปี พ.ศ. 2552

คำนำ

การวัดและประเมินทางการศึกษา (Educational Measurement and Evaluation) เป็นศาสตร์สำคัญสำหรับวงการศึกษาที่ครูทุกคนจำเป็นต้องรู้ เพราะเป็นวิธีการช่วยให้ทราบและพัฒนาศักยภาพของผู้เรียนให้เหมาะสมกับความแตกต่างระหว่างบุคคล โดยเฉพาะอย่างยิ่งการพัฒนาเครื่องมือให้มีคุณภาพ และตัดสินคะแนนด้วยความยุติธรรม

ประเทศไทยได้มีการปฏิรูปการศึกษา โดยเฉพาะการวัดและประเมินผลที่เน้นการประเมินการปฏิบัติ ดังนั้นจึงจำเป็นต้องมีวิธีการกำหนดคะแนนจุดตัดที่เหมาะสมกับการประเมินการปฏิบัติ ซึ่งวิธีการนั้นก็คือ วิธีการกำหนดมาตรฐานด้วยนิกมาร์ค ซึ่งเป็นวิธีที่นิยมใช้กันมากที่สุดในประเทศสหรัฐอเมริกา เพราะวิธีการนี้จะช่วยผู้ตัดสินในการคิดหาคะแนนจุดตัด ผ่านคู่มือการจัดเรียงข้อสอบ นอกจากนี้ วิธีนี้เหมาะกับการกำหนดคะแนนจุดตัดที่มีข้อสอบแบบเลือกตอบและข้อสอบอัตนัยอยู่ในฉบับเดียวกัน ผู้วิจัยหวังเป็นอย่างยิ่งว่าวิธีการกำหนดมาตรฐานด้วยวิธีการนิกมาร์คจะมีประโยชน์ต่อวงการศึกษาไทย

ผู้วิจัยขอขอบพระคุณฝ่ายวิจัยและบริการวิชาการ คณะศึกษาศาสตร์ มหาวิทยาลัยบูรพา ที่เห็นความสำคัญของการวิจัยนี้ และจัดสรรทุนอุดหนุนการวิจัย ประจำปี พ.ศ. 2552 สำหรับการวิจัยครั้งนี้

ผู้วิจัยขอขอบพระคุณ รองศาสตราจารย์ ดร.ไพรัตน์ วงษ์นาม เป็นอย่างยิ่ง ที่ให้คำปรึกษาด้านโปรแกรมคอมพิวเตอร์ Multilog และ Xcalibre และขอขอบคุณ ผู้ช่วยศาสตราจารย์ ดร.สมโภชน์ อเนกสุข ที่ให้คำปรึกษาด้านสถิติ และขอขอบคุณ ผู้ช่วยศาสตราจารย์ ดร.ระพีพันธ์ ฉายวิมล และดร.ดลดาว ปุณณานนท์ ที่เป็นกำลังใจในการทำวิจัยให้ผู้วิจัยด้วยดีเสมอมา และขอขอบคุณ คุณณัฐกฤตา งามมีฤทธิ์ และคุณจิรนนท์ พุ่มเหรียญ ที่มีส่วนช่วยให้งานวิจัยนี้สำเร็จลุล่วงด้วยดี และสุดท้ายขอขอบคุณ นาวาอากาศเอกปริญญา อนุศาสนนันท์ และเด็กหญิง ณัฐกาญจน์ อนุศาสนนันท์ ที่เอื้อกำลังกายกำลังใจและอยู่เคียงข้างผู้วิจัยตลอดมา

ดร.สุรพร อนุศาสนนันท์

สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	ง
บทคัดย่อภาษาอังกฤษ.....	ฉ
สารบัญ.....	ช
สารบัญตาราง.....	ญ
สารบัญภาพ.....	ฎ
บทที่	
1 บทนำ.....	1
ความเป็นมาและความสำคัญของปัญหา.....	1
วัตถุประสงค์ในการวิจัย.....	8
ขอบเขตการวิจัย.....	8
กรอบแนวคิดการวิจัย.....	9
คำจำกัดความที่ใช้ในการวิจัย.....	9
ประโยชน์ที่คาดว่าจะได้รับ.....	11
2 เอกสารและงานวิจัยที่เกี่ยวข้อง.....	12
ตอนที่ 1 มโนทัศน์ของมาตรฐาน และคะแนนจุดตัด.....	12
ตอนที่ 2 วิธีการกำหนดมาตรฐาน.....	14
ตอนที่ 3 การกำหนดคะแนนจุดตัดด้วยวิธีบีคมาร์คโดยการประยุกต์ทฤษฎี	
การตอบสนองข้อสอบ.....	37
ตอนที่ 4 แบบสอบวัดผลสัมฤทธิ์ทางการเรียน.....	44
ตอนที่ 5 งานวิจัยที่เกี่ยวข้อง.....	63
3 วิธีดำเนินการวิจัย.....	66
ขั้นตอนการดำเนินการวิจัย.....	66
เครื่องมือที่ใช้ในการวิจัย.....	75
การเก็บรวบรวมข้อมูล.....	88
การวิเคราะห์ข้อมูล.....	91

สารบัญ (ต่อ)

บทที่	หน้า
4 ผลการวิเคราะห์ข้อมูล	94
ตอนที่ 1 ผลการพัฒนาแบบสอบถามผลสัมฤทธิ์ทางการเรียนวิชาการวัดและ ประเมินผลในชั้นเรียน.....	94
ตอนที่ 2 ผลการกำหนดมาตรฐานด้วยวิธีบูคมาร์ค.....	106
5 สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ.....	129
สรุปผลการวิจัย.....	130
อภิปรายผลการวิจัย.....	132
ข้อเสนอแนะ.....	136
รายการอ้างอิง.....	138
ภาคผนวก.....	144
ภาคผนวก ก.....	145
ภาคผนวก ข.....	147
ภาคผนวก ค.....	149
ภาคผนวก ง.....	155

สารบัญตาราง

ตารางที่	หน้า
2-1 ตัวอย่างการคำนวณหาคะแนนจุดตัดด้วยวิธีแองกอฟจากเปอร์เซ็นต์การตอบ ข้อสอบแต่ละข้อของผู้สอบที่มีความสามารถขั้นต่ำ.....	16
2-2 ตัวอย่างการคำนวณหาคะแนนจุดตัดด้วยวิธีแองกอฟที่ได้รับการปรับปรุงจาก เปอร์เซ็นต์การตอบข้อสอบแต่ละข้อของผู้สอบที่มีความสามารถขั้นต่ำ.....	17
2-3 การเปรียบเทียบข้อดี และข้อบกพร่องของแบบสอบอัตนัย แบบสอบเติมคำ/ ตอบสั้น แบบสอบปรนัย.....	62
3-1 รายชื่อสาขาวิชาและจำนวนกลุ่มตัวอย่างนิสิตที่ลงทะเบียนเรียนรายวิชาการวัด และประเมินผลในชั้นเรียนที่ใช้เก็บข้อมูลเพื่อทดลองแบบสอบครั้งที่ 1 ครั้งที่ 2 และครั้งที่ 3.....	69
3-2 ตารางโครงสร้างกำหนดน้ำหนักความสำคัญของเนื้อหาและวัตถุประสงค์ การเรียนรู้ของวิชาการวัดผลและประเมินผลในชั้นเรียน.....	77
3-3 ตารางโครงสร้างกำหนดน้ำหนักความสำคัญของเนื้อหาและวัตถุประสงค์ การเรียนรู้ของวิชาการวัดผลและประเมินผลในชั้นเรียน.....	79
3-4 ผลการตรวจสอบความสอดคล้องระหว่างข้อสอบกับจุดประสงค์การเรียนรู้ ของแบบสอบกลางภาคเรียน.....	83
3-5 ผลการตรวจสอบความสอดคล้องระหว่างข้อสอบกับจุดประสงค์การเรียนรู้ ของแบบสอบกลางภาคเรียน.....	84
3-6 ค่าสถิติพื้นฐานของแบบสอบผลสัมฤทธิ์ทางการเรียนวิชาการวัดและประเมินผล การศึกษาฉบับที่ 1 แบบสอบกลางภาคเรียน ฉบับที่ 2 แบบสอบปลายภาคเรียน	86
3-7 ค่าความยาก และอำนาจจำแนกของแบบสอบผลสัมฤทธิ์ทางการเรียนวิชา การวัดและประเมินผลการศึกษาฉบับที่ 1 แบบสอบกลางภาคเรียน ฉบับที่ 2 แบบสอบปลายภาคเรียน.....	87
3-8 ข้อสอบที่ไม่มีคุณภาพตามเกณฑ์ที่กำหนด ต้องได้รับการปรับปรุงข้อสอบ.....	88
4-1 ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างระดับผลการเรียนเฉลี่ย กับคะแนนสอบของ นิสิตจากแบบสอบผลสัมฤทธิ์ทางการเรียนวิชาการวัดและประเมินผลในชั้นเรียน จำนวน 667 คน.....	95

สารบัญตาราง (ต่อ)

ตารางที่		หน้า
4-2	ค่าสถิติพื้นฐานของแบบสอบผลสัมฤทธิ์ทางการเรียนวิชาการวัดและประเมินผล การศึกษาฉบับที่ 1 แบบสอบกลางภาคเรียน ฉบับที่ 2 แบบสอบปลายภาคเรียน นิสิตจำนวน 667 คน.....	96
4-3	ค่าความยาก (b) ค่าอำนาจจำแนก (a) และค่าการเดา (c) ของแบบสอบวัดผล สัมฤทธิ์ทางการเรียนวิชาการวัดและประเมินในชั้นเรียนประเภทเลือกตอบ 4 ตัวเลือก ฉบับที่ 1 กลางภาคเรียน ฉบับที่ 2 ปลายภาคเรียน และรวมทั้งสองฉบับ	97
4-4	ค่าความยาก (δ_{ij}) ค่าอำนาจจำแนก (α) ของแบบสอบวัดผลสัมฤทธิ์ทาง การเรียนวิชาการวัดและประเมินในชั้นเรียนประเภทอัตนัย ฉบับที่ 1 กลาง ภาคเรียน จำนวน 1 ข้อ ฉบับที่ 2 ปลายภาคเรียนจำนวน 1 ข้อ	102
4-5	ผลค่าความสามารถ และลักษณะการจัดเรียงข้อสอบในคู่มือจัดเรียงข้อสอบ	106
4-6	ผลการตัดสินคะแนนจุดตัดทั้ง 7 ระดับ รอบที่ 1 ตามคะแนนความสามารถ ($\theta @ RP=.67$) และคะแนนดิบ.....	112
4-7	ผลการตัดสินคะแนนจุดตัดทั้ง 7 ระดับ รอบที่ 2 ตามคะแนนความสามารถ ($\theta @ RP=.67$) และคะแนนดิบ.....	116
4-8	ผลการตัดสินคะแนนจุดตัดทั้ง 7 ระดับ รอบที่ 3 ตามคะแนนความสามารถ ($\theta @ RP=.67$) และคะแนนดิบ.....	119
4-9	จำนวนนักเรียนผ่านแต่ละระดับมาตรฐาน ด้วยการกำหนดมาตรฐาน วิธีการบูคมาร์ค.....	123
4-10	ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างระดับผลการเรียนเฉลี่ย กับระดับผลการตัดสิน ที่กำหนดด้วยวิธีการบูคมาร์ค.....	127
ค-1	ผลการวิเคราะห์ค่าอำนาจจำแนก (a) ค่าความยาก (b) และค่าการเดา (c) ของแบบสอบฉบับที่ 1 กลางภาคเรียน จำนวน 60 ข้อ.....	150
ค-2	ผลการวิเคราะห์ค่าอำนาจจำแนก (a) ค่าความยาก (b) และค่าการเดา (c) ของแบบสอบฉบับที่ 2 กลางภาคเรียน จำนวน 50 ข้อ.....	153
ง-1	การคำนวณค่า Ability (θ) สำหรับ Specified Response Probability (RP) ของข้อสอบอัตนัย (ข้อที่ 61) ที่มีการให้คะแนนแบบ 5 คะแนน.....	156
ง-2	การคำนวณค่า Ability (θ) สำหรับ Specified Response Probability (RP) ของข้อสอบอัตนัย (ข้อที่ 112) ที่มีการให้คะแนนแบบ 5 คะแนน.....	159

สารบัญภาพ

ภาพที่	หน้า
1-1 กรอบแนวคิดในการวิจัย.....	9
2-1 คะแนนการแจกแจงผู้สอบในกลุ่มคาบเส้น	22
2-2 คะแนนการแจกแจงของผู้สอบกลุ่มที่ต่ำและสูงกว่ามาตรฐาน.....	23
2-3 คะแนนการแจกแจงกลุ่มผู้สอบคาบเส้นสำหรับระดับการปฏิบัติหลายระดับ.....	32
2-4 ตัวอย่างรูปแบบหน้าในคู่มือจัดเรียงข้อสอบ.....	37
2-5 ตัวอย่างการเรียงข้อสอบจากข้อง่ายสุดไปยังข้อยากสุดในคู่มือจัดเรียงข้อสอบ....	38
2-6 ตัวอย่างการกำหนดขีดคั่นมาร์ค (การคั่นหนังสือ) 7 ระดับ.....	40
2-7 ขั้นตอนการก่อสร้างแบบสอบผลสัมฤทธิ์ทางการเรียน.....	48
3-1 ขั้นตอนการดำเนินงานวิจัย.....	67
4-1 ค่าอำนาจจำแนก (a) ของแบบสอบกลางภาค.....	99
4-2 ค่าความยาก (b) ของแบบสอบกลางภาค.....	99
4-3 ค่าการเดา (c) ของแบบสอบกลางภาค.....	99
4-4 ค่าอำนาจจำแนก (a) ของแบบสอบปลายภาค.....	100
4-5 ค่าความยาก (b) ของแบบสอบปลายภาค.....	100
4-6 ค่าการเดา (c) ของแบบสอบปลายภาค.....	100
4-7 ค่าอำนาจจำแนก (a) ของแบบสอบรวม 2 ฉบับ.....	101
4-8 ค่าความยาก (b) ของแบบสอบรวม 2 ฉบับ.....	101
4-9 ค่าการเดา (c) ของแบบสอบรวม 2 ฉบับ.....	101
4-10 ได้ังสารสนเทศของแบบสอบกลางภาค.....	103
4-11 ได้ังลักษณะแบบสอบกลางภาค.....	103
4-12 ได้ังสารสนเทศของแบบสอบปลายภาค.....	104
4-13 ได้ังลักษณะแบบสอบปลายภาค.....	104
4-14 ได้ังสารสนเทศของแบบสอบปลายภาค.....	105
4-15 ได้ังลักษณะแบบสอบรวมกลางภาคและปลายภาค.....	105
4-16 ผลการกำหนดเลขหน้าของผู้ตัดสิน รอบที่ 1.....	121
4-17 ผลการกำหนดเลขหน้าของผู้ตัดสิน รอบที่ 2.....	122
4-18 ผลการกำหนดเลขหน้าของผู้ตัดสิน รอบที่ 3.....	122
4-19 จำนวนนิสิตในแต่ละระดับมาตรฐาน.....	124

บทคัดย่อ

ชื่อเรื่อง การพัฒนาแบบสอบผลสัมฤทธิ์ทางการเรียนวิชา การวัดและประเมินผลในชั้นเรียนโดยกำหนดคะแนนมาตรฐานด้วยวิธีบูคมาร์ค

ผู้วิจัย ดร. สุวีพร อนุศาสนนันท์

ปี พ.ศ. 2554

245527

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อ 1) สร้างแบบสอบผลสัมฤทธิ์ทางการเรียนวิชา การวัดและประเมินผลในชั้นเรียน 2) ตรวจสอบความยาก อำนาจจำแนก ความตรง และความเที่ยงของแบบสอบผลสัมฤทธิ์ทางการเรียนวิชา การวัดและประเมินผลในชั้นเรียนที่สร้างขึ้น และ 3) หาคะแนนจุดตัดของแบบสอบที่สร้างขึ้นโดยการกำหนดมาตรฐานด้วยวิธีบูคมาร์ค กลุ่มตัวอย่าง มี 2 กลุ่มคือ 1) ผู้ตัดสิน คือ อาจารย์ที่สอนรายวิชา การวัดและประเมินผลในชั้นเรียน จำนวน 6 คน 2) ผู้สอบ คือ นิสิตชั้นปีที่ 3 คณะศึกษาศาสตร์ ที่ลงทะเบียนรายวิชา 400204 การวัดและประเมินในชั้นเรียน ปีการศึกษา 2553 ใช้วิธีการสุ่มแบบยกกกลุ่ม (cluster sampling) จำนวน 667 คน เครื่องมือที่ใช้ในการวิจัย ได้แก่ 1) แบบสอบวัดผลสัมฤทธิ์ทางการเรียนวิชา การวัดและประเมินผลในชั้นเรียน แบบเลือกตอบ 4 ตัวเลือก และแบบอัตนัย จำนวน 2 ฉบับ ผลการสอบนำมาวิเคราะห์ค่าความยาก อำนาจจำแนก ความเที่ยงโดยใช้ทฤษฎีตอบสนองข้อสอบแบบ 3 พารามิเตอร์ และ Partial-Credit Model (PCM) วิเคราะห์ความตรงโดย หาสหสัมพันธ์ 2) คู่มือการจัดเรียงข้อสอบ (ordered item booklet : OIB) เพื่อใช้กำหนดมาตรฐานด้วยวิธีบูคมาร์ค

ผลการวิจัยสรุปได้ดังนี้

1) แบบสอบวัดผลสัมฤทธิ์ทางการเรียนวิชา การวัดและประเมินในชั้นเรียนประเภทเลือกตอบที่พัฒนาแล้ว ฉบับที่ 1 มีค่าความยากระหว่าง -2.50 ถึง 3.00 ค่าอำนาจจำแนกระหว่าง .49 ถึง .88 และค่าการเดาระหว่าง .11 ถึง .29 และประเภทอัตนัย ค่าอำนาจจำแนกเท่ากับ .98 และค่าความยากมีค่าระหว่าง -2.17 ถึง .47 ส่วนฉบับที่ 2 ค่าความยากระหว่าง -1.39 ถึง 3.00 ค่าอำนาจจำแนกระหว่าง .50 ถึง .90 และการเดาระหว่าง .11 ถึง .29 และประเภทอัตนัย ค่าอำนาจจำแนกเท่ากับ .47 และค่าความยากระหว่าง -0.98 ถึง .59

2) ค่าสารสนเทศของแบบสอบฉบับที่ 1 อยู่ในช่วงประมาณ 3.9 ถึง 5.0 และมีค่าสูงสุดอยู่ที่ระดับความสามารถ (θ) ประมาณ -1.0 ฉบับที่ 2 ค่าสารสนเทศของแบบสอบอยู่ในช่วงประมาณ 3.9 ถึง 12.0 และมีค่าสูงสุดอยู่ที่ระดับความสามารถ (θ) ประมาณ -0.5

3) คะแนนจุดตัดที่อยู่ในรูปคะแนนดิบ และคะแนนความสามารถ จากแบบสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาการวัดและประเมินในชั้นเรียน ซึ่งกำหนดคะแนนจุดตัดด้วยวิธีบูคมาร์ค 7 ระดับ ดังนี้ ระดับดีเยี่ยม (A) เท่ากับ 109 (1.453) ระดับดีมาก (B+) เท่ากับ 91 (1.293) ระดับดี (B) เท่ากับ 83 (1.203) ระดับดีพอใช้ (C+) เท่ากับ 65 (1.133) ระดับพอใช้ (C) เท่ากับ 49 (1.093) ระดับอ่อน (D+) เท่ากับ 30 (1.0130) และระดับอ่อนมาก (D) เท่ากับ 13 (.933)

Abstract

Title: A Development of Measurement and Assessment in Classroom Achievement Test: using Bookmark standard setting

Researcher: Dr. Sureeporn Anusasananan

Year: 2011

245522

The objectives of this study are 1) to develop achievement tests for the educational measurement and evaluation of bachelor degrees at Burapha University, 2) to investigate the item difficulty indices, the item discriminating indices, and validity and reliability in achievement tests, and 3) to investigate the cut scores on the bookmark standard setting method. The sample was divided into two groups: 1) 6 Educational Measurement and Evaluation in the Classroom lecturers assessors (acting as assessors) from the Education faculty of Burapha University, and 2) 677 undergraduate students who studied Educational Measurement and Evaluation in the Classroom, in 2010. The research instruments included: 1) both of midterm and final in educational measurement and evaluation achievement tests: multiple choice items and essay items. Their scores were analyzed to find difficulty indices, discriminating indices, the reliability—by using the IRT model; 3 PL model and the Partial-Credit Model (PCM) and the validity – by finding the correlation. 2) the ordered item booklet: the OIB used in Bookmark standard setting method.

The following are the research findings:

1. The first test, multiple choice items, indicated difficulty index were 2.50 to 3.00, the discriminating index were .49 to .88 and the guessing index were .11 to .29. The essay items indicated the discrimination index was 0.98 and the item difficulty indices were -2.17 to .47. The second test, multiple choice items, the difficulty index were -1.39 to 3.00, the discriminating index were .50 to .90 and the guessing were .11 to .29. The essay items indicated the discriminating index was 0.47 and the item difficulty indices were -.98 to .59.

2. The test information function of the first test was between 3.9 and 5.0, and the highest ability level (θ) was -1.0. For the second test, the test information function was between 3.9 to 12 and the highest ability level (θ) was -0.5

3. The cut scores in raw scores and examinee's ability (θ) of the achievement tests were divided into 7 levels ,using the Bookmark method, as follows: excellent (A) was 109 ($\theta=1.453$) , very good (B+) was 91($\theta=1.293$), good (B) was 83($\theta=1.203$), rather good (C+) was 65($\theta=1.133$), fair (C) was 49.5($\theta=1.093$), poor (D+) was 30($\theta=1.013$), and very poor (D) was 13($\theta=0.933$).