

การเปรียบเทียบผลการตรวจสอบความเหมาะสมของบุคคลระหว่าง
ดัชนีแอลแซดกับดัชนีดับเบิลยูวันตามทฤษฎีการตอบสนองข้อสอบ

นางสาวกฤษมา สุวรรณแก้ว



วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาครุศาสตรมหาบัณฑิต
สาขาวิชาการวัดและประเมินผลการศึกษา ภาควิชาวิจัยการศึกษา
บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2540

ISBN 974-636-916-4

ลิขสิทธิ์ของบัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย

**A COMPARISON OF THE RESULTS OF PERSON-FIT DETECTION
BETWEEN L_1 AND W_1 INDICES BASED ON ITEM RESPONSE THEORY**



MISS KUSUMA SUWANKHAW

**A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Education in Measurement and Evaluation**

Department of Educational Research

Graduate School

Chulalongkorn University

Academic Year 1997

ISBN 974-636-916-4

พิมพ์ต้นฉบับบทคัดย่อวิทยานิพนธ์ภายในกรอบสี่เหลี่ยมนี้เพียงแผ่นเดียว

กฤษมา สุวรรณแก้ว: การเปรียบเทียบผลการตรวจสอบความเหมาะสมของบุคคลระหว่างดัชนีแอลแซดกับดัชนีดับเบิลยูวันตามทฤษฎีการตอบสนองข้อสอบ (A COMPARISON OF THE RESULTS OF PERSON-FIT DETECTION BETWEEN L_2 AND W_1 INDICES BASED ON ITEM RESPONSE THEORY) อ.ที่ปรึกษา : รศ.ดร.ศิริชัย กาญจนวาสิ, 163 หน้า.
ISBN 974-636-916-4

การวิจัยครั้งนี้ มีวัตถุประสงค์เพื่อเปรียบเทียบผลการตรวจสอบความเหมาะสมของบุคคลระหว่างดัชนีซึ่งประมาณค่าด้วยวิธีที่แตกต่างกัน ได้แก่ ดัชนีแอลแซด ประมาณค่าด้วยฟังก์ชันโลดัลลิสต์ และดัชนีดับเบิลยูวัน ประมาณค่าดัชนีจากคะแนนส่วนที่เหลือ เมื่อระดับความสามารถของผู้สอบ และความยาวของแบบสอบแตกต่างกัน ดำเนินการวิจัยโดยจำลองข้อมูลด้วยโปรแกรม IRTDATA ตรวจสอบความเป็นเอกมิติของแบบสอบทุกเงื่อนไขด้วยโปรแกรม SPSSX และ SPSSPC⁺ ผลการตรวจสอบความเหมาะสมของบุคคลวัดจากคุณภาพของแบบสอบหลังการตรวจสอบความเหมาะสมของบุคคลกับโมเดลการตอบข้อสอบในด้านค่าสารสนเทศ ความเที่ยง และความตรงตามทฤษฎีของแบบสอบ

ผลการวิจัยพบว่า

1. ผลการตรวจสอบความเหมาะสมของบุคคลด้วยดัชนีแอลแซด ด้านค่าสารสนเทศของแบบสอบสูงกว่าดัชนีดับเบิลยูวัน ในทุกระดับความสามารถของผู้สอบและทุกระดับความยาวของแบบสอบ
2. ผลการตรวจสอบความเหมาะสมของบุคคลด้วยดัชนีแอลแซด และดัชนีดับเบิลยูวัน ด้านความเที่ยงและความตรงตามทฤษฎีไม่แตกต่างกัน ในทุกระดับความสามารถของผู้สอบและทุกระดับความยาวของแบบสอบ

ภาควิชา วิชาการศึกษา.....
สาขาวิชา การวัดและประเมินผลการศึกษา.....
ปีการศึกษา 2540.....

ลายมือชื่อนิติกร
ลายมือชื่ออาจารย์ที่ปรึกษา
ลายมือชื่ออาจารย์ที่ปรึกษาร่วม

C840684 : MAJOR EDUCATIONAL MEASUREMENT AND EVALUATION

KEY WORD: PERSON-FIT INDEX / ITEM RESPONSE PATTERN / ITEM RESPONSE THEORY

KUSUMA SUWANKHAW : A COMPARISON OF THE RESULTS OF PERSON-FIT

DETECTION BETWEEN L_z AND W_1 INDICES BASED ON ITEM RESPONSE THEORY.

THESIS ADVISOR : ASSO. PROF. SIRICHAI KANJANAWASEE, Ph.D. 163 pp.

ISBN 974-636-916-4

The purpose of this study was to compare the results of person-fit detection between two different statistical estimated methods, the likelihood-based approach : L_z index, and the score residual approach : W_1 index under different examinee's ability levels and test lengths. The data were simulated by the computer packages IRTDATA. The unidimensionality indicators of test in all conditions were analyzed by the computer packages SPSSX and SPSSPC⁺. The results of person-fit detection were compared in terms of the quality of tests after detected person-fit : test information function, reliability, and construct validity.

The findings were as follow :

1. The results in terms of test information function : L_z index provided higher information than W_1 index in all examinee's ability levels and test lengths.
2. The results in terms of reliability and construct validity : L_z index and W_1 index were not significantly different in all examinee's ability levels and test lengths.

ภาควิชา วิทยาการศึกษา.....

สาขาวิชา การวัดและประเมินผลการศึกษา.....

ปีการศึกษา 2540.....

ลายมือชื่อนิสิต..... + ๑๕

ลายมือชื่ออาจารย์ที่ปรึกษา.....

ลายมือชื่ออาจารย์ที่ปรึกษาร่วม.....

กิตติกรรมประกาศ



ผู้วิจัยขอขอบพระคุณ คณะจารย์ภาควิชาวิจัยการศึกษาทุกท่านที่ได้ประสิทธิ์ประสาทความรู้พื้นฐานทางการวิจัย และอบรมคุณลักษณะความเป็นนักวัดผลและนักวิจัยการศึกษาให้กับผู้วิจัยตลอดระยะเวลาที่ศึกษา

ขอกราบขอบพระคุณ รองศาสตราจารย์ ดร. ศิริชัย กาญจนวาสิ อาจารย์ที่ปรึกษาที่กรุณาให้ความเมตตา ดูแลและให้คำแนะนำด้วยความเอาใจใส่ จนวิทยานิพนธ์สำเร็จลุล่วงด้วยดี ผู้วิจัยขอปวารณาตัวยึดเป็นแบบอย่างของอาจารย์ที่ปรึกษาตลอดไป

ในรายวิชาสัมมนาและออกแบบการวิจัย ผู้วิจัยได้รับคำแนะนำ ข้อเสนอแนะที่เป็นประโยชน์ และความช่วยเหลือด้านเอกสารจาก รศ. ดร. สุวิมล ร่วงวานิช ผศ. ดร. นงลักษณ์ วิรัชชัย และ ผศ. ดร. สุวัฒน์ สุวรรณเขตนิคม จนได้แนวทางในการเสนอโครงร่างวิทยานิพนธ์ จึงขอขอบพระคุณเป็นอย่างสูง ณ โอกาสนี้

ขอขอบคุณ Dr. Michael L. Nering แห่ง American College Testing ที่กรุณาให้ความช่วยเหลือให้ด้านเอกสารและตัวอย่างโปรแกรมการวิเคราะห์ข้อมูล ที่เป็นประโยชน์สำหรับการวิจัย ขอขอบคุณ Dr. George A. Johnson แห่งมหาวิทยาลัยไอโฮโอ สหรัฐอเมริกา ผู้พัฒนาโปรแกรม IRTDATA ซึ่งผู้วิจัยใช้ในการจำลองข้อมูล

ขอขอบคุณ คณะครูโรงเรียนบ้านเขาวังทองทุกท่าน ที่เสียสละปฏิบัติหน้าที่ราชการ แทนผู้วิจัยในขณะศึกษาต่อ และเป็นกำลังใจให้ผู้วิจัยมาโดยตลอด

ขอขอบคุณ พี่ๆ น้องๆ และเพื่อนๆ ภาควิชาวิจัยการศึกษาที่ให้ความช่วยเหลือ และเป็นกำลังใจตลอดระยะเวลาที่เรียนร่วมกัน

ขอบคุณ น้องๆ และญาติมิตร ซึ่งให้การสนับสนุนในการศึกษาสำหรับผู้วิจัยมาโดยตลอด โดยเฉพาะ คุณปิยะพงษ์ หนูสังข์ ที่กรุณาให้ความช่วยเหลือในการเขียนโปรแกรม และเป็นกำลังใจที่ดียิ่ง

คุณความดีและสาระประโยชน์จากวิทยานิพนธ์เรื่องนี้ ผู้วิจัยขอมอบแต่ คุณพ่อสำราญ และ คุณแม่ฉวีวรรณ สุวรรณแก้ว เพื่อตอบแทนความรัก ความเสียสละ และการปลูกฝังการศึกษาที่ดีให้ลูกเสมอมา

กุสุมา สุวรรณแก้ว

สารบัญ

	หน้า
บทคัดย่อภาษาไทย	ง
บทคัดย่อภาษาอังกฤษ	จ
กิตติกรรมประกาศ	ฉ
สารบัญตาราง	ช
สารบัญแผนภูมิ	ฅ
บทที่ 1 บทนำ	
ความเป็นมาและความสำคัญของปัญหา	1
วัตถุประสงค์ของการวิจัย	7
สมมติฐานการวิจัย	7
ขอบเขตของการวิจัย	8
คำจำกัดความที่ใช้ในการวิจัย	9
ประโยชน์ที่คาดว่าจะได้รับ	13
บทที่ 2 วรรณคดีที่เกี่ยวข้อง	
แนวคิดเกี่ยวกับสเกลของกัตแมน	14
ดัชนีบ่งชี้ความเหมาะสมของบุคคล	16
ทฤษฎีการตอบสนองข้อสอบ	27
วิธีการตรวจสอบคุณภาพของแบบสอบ	34
รายงานการวิจัยที่เกี่ยวข้อง	43
บทที่ 3 วิธีการดำเนินการวิจัย	
ขั้นตอนการดำเนินการวิจัย	48
ตัวแปรที่ศึกษา	51
วิธีการจำลองข้อมูล	52
การวิเคราะห์ข้อมูล	55

สารบัญ (ต่อ)

หน้า

บทที่ 4 ผลการวิเคราะห์ข้อมูล

ตอนที่ 1 ผลการวิเคราะห์ข้อมูลเบื้องต้น

ผลการตรวจสอบความเป็นเอกมิติของแบบสอบ.....67

คุณภาพของแบบสอบ.....72

ค่าสถิติพื้นฐานของแบบสอบ.....73

ตอนที่ 2 ผลการวิเคราะห์บุคคลที่เหมาะสมกับโมเดลการตอบข้อสอบ

จำนวนผู้สอบที่ไม่เหมาะสม76

คุณภาพของแบบสอบหลังการวิเคราะห์.....78

ตอนที่ 3 ผลการเปรียบเทียบการตรวจสอบความเหมาะสมของบุคคล

การเปรียบเทียบค่าสารสนเทศของแบบสอบ.....87

การเปรียบเทียบค่าความเที่ยงของแบบสอบ.....89

การเปรียบเทียบความตรงตามทฤษฎีของแบบสอบ.....101

บทที่ 5 สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ

สรุปผลการวิจัย.....108

อภิปรายผลการวิจัย.....110

ข้อเสนอแนะในการนำไปใช้.....111

ข้อเสนอแนะในการทำวิจัยครั้งต่อไป.....112

รายการอ้างอิง.....114

ภาคผนวก

ภาคผนวก ก ลักษณะของโปรแกรมที่ผู้วิจัยสร้างขึ้น

แฟ้มคำสั่งของโปรแกรม.....120

สารบัญ (ต่อ)

	หน้า
ตัวอย่างรูปแบบหน้าจอของโปรแกรมส้ม	142
ตัวอย่างรูปแบบหน้าจอของโปรแกรมคำนวณค่าสถิติ L_2	143
ตัวอย่างเพิ่มผลการวิเคราะห์ข้อมูล	145
ภาคผนวก ข ตัวอย่างคำสั่งและเพิ่มผลลัพธ์ของโปรแกรมที่ใช้	
ผลการจำลองข้อมูลด้วยโปรแกรม IRTDATA	146
ตัวอย่างผลการวิเคราะห์ข้อมูลด้วยโปรแกรม SPSS	
การวิเคราะห์องค์ประกอบเชิงสำรวจ	151
ผลการทดสอบสกรี	153
เพิ่มผลลัพธ์จากโปรแกรม BILOG	
ตัวอย่างคำสั่ง	154
ค่าสถิติพื้นฐาน	155
ค่าสารสนเทศของข้อสอบ	157
เพิ่มผลลัพธ์จากโปรแกรม BIGSTEPS	158
คำสั่งและผลการวิเคราะห์เตรียมข้อมูลโปรแกรม LISRE 8.10	
คำสั่งของโปรแกรมย่อย PRELIS 2	160
เมตริกซ์ความแปรปรวน - ความแปรปรวนร่วม	161
คำสั่งการวิเคราะห์องค์ประกอบเชิงยืนยัน	162
ประวัติผู้วิจัย	163

สารบัญตาราง

ตารางที่		หน้า
1	สเกลของกัตแมน	16
2	แบบแผนการตอบข้อสอบที่ผิดปกติ	17
3	ข้อมูลที่ใช้ศึกษา	54
4	ผลการจำลองเมตริกซ์คำตอบด้วยโปรแกรม IRTDATA	65
5	ผลการสุ่มข้อมูลจากเมตริกซ์คำตอบจำแนกตามระดับความสามารถของผู้สอบ ...	66
6	ผลการตรวจสอบความเป็นเอกมิติของแบบสอบ	67
7	ค่าความยาก อำนาจจำแนก และค่าความเที่ยงเฉลี่ยของแบบสอบ	72
8	ลักษณะการแจกแจงของคะแนนของแบบสอบความยาว 20 ข้อ	73
9	ลักษณะการแจกแจงของคะแนนของแบบสอบความยาว 40 ข้อ	74
10	ลักษณะการแจกแจงของคะแนนของแบบสอบความยาว 60 ข้อ	75
11	จำนวนและร้อยละของผู้สอบที่ไม่เหมาะสมกับโมเดลการตอบข้อสอบ	76
12	ค่าสารสนเทศของแบบสอบที่ได้จากแบบสอบก่อนและการวิเคราะห์ ผู้สอบที่เหมาะสมกับโมเดลการตอบข้อสอบ	79
13	ค่าประสิทธิภาพสัมพัทธ์ของแบบสอบจากการตรวจสอบ ความเหมาะสมของบุคคลด้วยดัชนีแอลซัดและดัชนีดับเบิลยูวัน	81
14	ผลการวิเคราะห์คุณภาพของแบบสอบด้านความเที่ยง ภายหลังการวิเคราะห์บุคคลที่เหมาะสมกับโมเดลการตอบข้อสอบ	86
15	ผลการเปรียบเทียบค่าประสิทธิภาพสัมพัทธ์ของแบบสอบจากการตรวจสอบ ความเหมาะสมของบุคคลด้วยดัชนีแอลซัดและดัชนีดับเบิลยูวัน	88
16	ผลการเปรียบเทียบคุณภาพของแบบสอบด้านความเที่ยงก่อนและหลัง การวิเคราะห์ผู้สอบที่เหมาะสมกับโมเดลการตอบข้อสอบด้วยดัชนี L_2	90
17	ผลการเปรียบเทียบคุณภาพของแบบสอบด้านความเที่ยงก่อนและหลัง การวิเคราะห์ผู้สอบที่เหมาะสมกับโมเดลการตอบข้อสอบด้วยดัชนี W_1 โดยพิจารณาค่าสถิติ Infit	92

สารบัญตาราง (ต่อ)

ตารางที่	หน้า
18 ผลการเปรียบเทียบคุณภาพของแบบสอบด้านความเที่ยงก่อนและหลัง การวิเคราะห์ผู้สอบที่เหมาะสมกับโมเดลการตอบข้อสอบด้วยดัชนี W_1 โดยพิจารณาค่าสถิติ Outfit	94
19 ผลการเปรียบเทียบคุณภาพของแบบสอบด้านความเที่ยงหลัง การวิเคราะห์ผู้สอบที่เหมาะสมกับโมเดลการตอบข้อสอบ ด้วยดัชนี L_2 กับดัชนี W_1 โดยพิจารณาค่าสถิติ Infit	96
20 ผลการเปรียบเทียบคุณภาพของแบบสอบด้านความเที่ยงหลัง การวิเคราะห์ผู้สอบที่เหมาะสมกับโมเดลการตอบข้อสอบ ด้วยดัชนี L_2 กับดัชนี W_1 โดยพิจารณาค่าสถิติ Outfit	98
21 ผลการเปรียบเทียบคุณภาพของแบบสอบด้านความเที่ยงหลัง การวิเคราะห์ผู้สอบที่เหมาะสมกับโมเดลการตอบข้อสอบ ด้วยค่าสถิติในวิธีประมาณค่าดัชนีจากคะแนนส่วนที่เหลือ	100
22 ผลการเปรียบเทียบคุณภาพของแบบสอบด้านความตรงตามทฤษฎี เมื่อผู้สอบมีระดับความสามารถต่ำ	102
23 ผลการเปรียบเทียบคุณภาพของแบบสอบด้านความตรงตามทฤษฎี เมื่อผู้สอบมีระดับความสามารถปานกลาง	104
24 ผลการเปรียบเทียบคุณภาพของแบบสอบด้านความตรงตามทฤษฎี เมื่อผู้สอบมีระดับความสามารถสูง	106

สารบัญแผนภูมิ

แผนภูมิที่	หน้า
1. ผลการทดสอบสกรีนเมื่อผู้สอบมีระดับความสามารถต่ำ	69
2. ผลการทดสอบสกรีนเมื่อผู้สอบมีระดับความสามารถปานกลาง	70
3. ผลการทดสอบสกรีนเมื่อผู้สอบมีระดับความสามารถสูง	71
4. คำสารสนเทศของแบบสอบเมื่อผู้สอบมีระดับความสามารถต่ำ	83
5. คำสารสนเทศของแบบสอบเมื่อผู้สอบมีระดับความสามารถปานกลาง	84
6. คำสารสนเทศของแบบสอบเมื่อผู้สอบมีระดับความสามารถสูง	85



สถาบันวิจัยบริการ
จุฬาลงกรณ์มหาวิทยาลัย