



جامعة اليرموك  
كلية التربية  
قسم علم النفس الإرشادي والتربوي

# تأثير طول الاختبار على خصائص توزيع العلامات الحقيقية حسب النموذج ثلاثي المعلم

إعداد

لؤي عبد الرؤوف فلاح الذيابات

إشراف

أ.د. ضرار محمد جراحات

حقل التخصص - القياس والتقويم

1428هـ / 2007 م

## - الملخص -

الذيات، لؤي عبدالرؤوف فلاح. أثر التغير في طول الاختبار على خصائص توزيع العلامات الحقيقية حسب النموذج الثلاثي المعلم. رسالة ماجستير بجامعة اليرموك. 2007 (المشرف: أ.د. ضرار محمد جرادات).

تهدف هذه الدراسة للكشف عن تأثير طول الاختبار على خصائص توزيع العلامات الحقيقية للمفحوصين Domain Scores لعينة من طلبة الصفوف الأساسية (الرابع، الخامس و السادس) في مديرية تربية الرمثا حجمها (2100) طالب وطالبة، على اختبار أوتيس- لينون للقدرة العقلية للصورة k المعربة للبيئة الأردنية حسب النموذج ثلاثي المعلم باستخدام البرنامج BILOG-MG3.

وتم تقدير العلامات الحقيقية لمجتمع المفحوصين باستخدام النموذج الثلاثي المعلم، وذلك بعد التأكد من تحقق افتراضات النموذج المستخدم إذ قام الباحث بالتأكد من تحقق افتراض الاستقلال الموضوعي باستخدام الإحصائي Q3، كما تم التأكد من أحادية البعد باستخدام التحليل العاملي، وأظهرت النتائج مائة 79 فقرة من أصل 80 فقرة للاختبار.

وقد حاولت هذه الدراسة الإجابة عن الأسئلة التالية:

- السؤال الأول: هل تختلف خصائص توزيع العلامات الحقيقية Domain

Scores والعلامات المشاهدة Observed Domain Scores والعلامات

الخطأ Residuals باختلاف طول الاختبار؟

- السؤال الثاني: هل توجد فروق بين تقديرات ثبات الاختبار تعزى إلى التغير في

طول الاختبار؟

- السؤال الثالث: ما هو التغير في تقديرات معاملات الارتباط بين تقديرات كل من العلامات الحقيقية Domain Scores والعلامات المشاهدة Observed Domain Scores والعلامات الخطأ Residuals للمفحوصين تحت أطوال الاختبار المختلفة مع نظيراتها في الاختبار الكلي حسب النموذج الثلاثي المعلم؟
- السؤال الرابع: ما هو التغير في تقديرات معاملات الارتباط بين تقديرات كل من العلامات الحقيقية Domain Scores والعلامات المشاهدة Observed Domain Scores والعلامات الخطأ Residuals للمفحوصين تحت أطوال الاختبار المختلفة مع تقديرات علامات القدرة  $\theta$  للاختبار الكلي ولأطوال الاختبار المختلفة حسب النموذج الثلاثي المعلم؟
- السؤال الخامس: ما هو أقل عدد من الفقرات لتحقق خاصية اللاتغير Invariant في علامات القدرة  $\theta$ ؟
- وللإجابة عن هذه الأسئلة تم تقدير خصائص توزيع كل من العلامات الحقيقية والمشاهدة والخطأ من حيث التباين، المتوسط الحسابي وشكل التوزيع لـ 12 مجموعة من الفقرات تم اختيارها عشوائياً من مجموعة الفقرات الكلية (5، 10، 15، 20، 25، 30، 35، 40، 45، 50، 55، 60)، حيث تمثل كل مجموعة من هذه المجموعات طول مختلف للاختبار. وكشفت نتائج هذه الدراسة إلى أن خصائص توزيع العلامات الحقيقية تتأثر بتغير طول الاختبار، إذ دلت النتائج إلى أن التباين الحقيقي يزداد بزيادة طول الاختبار وقيمته قريبة من بعضها بعضاً و المتوسطات الحسابية للعلامات الحقيقية أيضاً كانت قريبة من بعضها البعض، كما بينت النتائج بأن خصائص توزيع العلامات الحقيقية تقترب من خصائص توزيع العلامات الحقيقية للاختبار الكلي بزيادة طول الاختبار، وأن

خصائص العلامات الحقيقية كانت مستقرة نسبياً و أقل تأثراً بطول الاختبار من خصائص العلامات المشاهدة. وأظهرت النتائج بأن الثبات يزداد بزيادة طول الاختبار وأن قيم الثبات باستخدام نسبة التباين الحقيقي إلى التباين الكلي اعتماداً على النموذج الثلاثي المعلم كانت مرتفعة جداً وقريبة من بعضها البعض ومتسقة وجميعها يصلح للاستخدام حيث كانت ضمن المدى (0.832 – 0.9586)، كما كشفت هذه الدراسة بأن علامات القدرة تقترب من علامات الاختبار الكلي ابتداء من الطول (15)، وأخيراً توفر هذه الدراسة تطبيقاً عملياً لتقدير العلامات الحقيقية وعلامات القدرة للمفحوصين باستخدام إجاباتهم على أي مجموعة من الفقرات يتم اختيارها عشوائياً من المجموعة الكلية لفقرات الاختبار، تسهم في تحديد أفضل الفقرات التي تناسب قدرة المفحوص باستخدام دالة خصائص الاختبار والذي يوفر إمكانية استخدام العلامات الحقيقية في التفسير المحكي المرجع.

**الكلمات المفتاحية: طول الاختبار، العلامات الحقيقية، خصائص توزيع، النموذج ثلاثي المعلم.**

## ABSTRACT

**Al Thiabat, Loai Abdel-raoof.(2007). The Effect of Test Length on the Characteristics of True Scores Distribution According to the Three Parameters Model . Master Thesis. Yarmouk University. (Supervisor: Prof. Derar Mohammad Jaradat).**

This study aimed at investigating the effect of test length on the characteristics of true scores distribution according to the three parameter logistic model (3PLM) in IRT using the form K of Otis – Lennon Mental Test which was developed to suite Jordanian primary schools environment. The sample of the study consists of (2100) male and female pupils grades 4<sup>th</sup>, 5<sup>th</sup> and 6<sup>th</sup> in Ramtha Education department.

True scores were estimated according to the 3PLM using BILOG-MG3 program after checking the assumption of Unidimensionality by Principal Component Factor Analysis and the assumption of Local Independence by Q3 statistic. Analyses found that 79 items of 80 fit the 3PLM.

The study attempt to answer the following questions:

- 1- Do the characteristics of the distribution of Domain scores, Observed scores and Error scores differ due to the variation in test length?
- 2- Is there any difference in the test reliability estimates due to the variation in test length?
- 3- What are the differences in correlation coefficients between Domain scores, Observed scores and Error scores of different test lengths and their counterparts in the full test length according to the 3PLM?

- 4- What are the differences in correlation coefficients between Domain scores, Observed scores and Error scores with ability scores for different test lengths and full test length according to the 3PLM?
- 5- What is the least number of items that achieves invariance of ability scores?

To answer the study questions, means, variances and distribution shapes (Skewness and Kurtosis) were computed to investigate the characteristics of domain scores distribution for 12 selected random groups of items with lengths (5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60) from the whole group.

The results show that the test length affect the characteristics of domain scores distribution. The true variance increases due to increases in test length and its values are close together. The means of domain scores are also close together.

The results also show that true scores distribution for the 12 selected lengths approaches to the true scores distribution of the full test length as the full test length increases. The results indicate that the characteristics of true scores distribution are less affected by test length than the characteristics of observed scores distribution.

The results also show that the reliability estimates for the deferent test lengths are high and close together and consistence. These values are in the range of (0.832 – 0.9586). The ability scores of different test lengths approach to the ability scores of the full test length starting from the length 15.

Finally the over all study provides an application for estimating true scores and ability scores of examinees by using their answers to any selected random group of items from the whole group of test items. This contributes to determining the best items that fit the examinees ability by using the test

characteristics curve TCC, It also provides the possibility of using true scores in criterion referenced interpretations.

**Key words:** Test Length, True Scores, Distribution Characteristics, Three Parameters Model, IRT, Domain Scores.