



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

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

## The Effect of Test Length on the Characteristics of True Scores Distribution According to the Three Parameters Model

إعداد

لؤي عبد الرؤوف فلاح الذيابات

بكالوريوس معلم مجال رياضيات، جامعة اليرموك، 2001م.

قدمت هذه الدراسة استكمالاً لمتطلبات الحصول على درجة الماجستير في تخصص القياس  
والتقويم من جامعة اليرموك، اربد، الأردن.

وافق عليها

- أ.د. ضرار محمد جرادات..... رئيساً.  
أستاذ في القياس والتقويم، جامعة اليرموك.
- أ.د. أحمد سليمان عودة..... عضواً.  
أستاذ في القياس والتقويم، جامعة اليرموك.
- د. أحمد يوسف قواسمة..... عضواً.  
أستاذ مشارك في القياس والتقويم، جامعة اليرموك.
- د. أمل عبدا لله خصاونه..... عضواً.  
أستاذ مشارك في مناهج الرياضيات وأساليب تدريسها، جامعة اليرموك.

تاريخ المناقشة 18 / 3 / 2007م

بِاللّٰهِ يَا قَارِئًا كَتَبْتِي وَسَامِعَهَا      أَسْبَلُ عَلَيْهَا رِجَاءَ الْحُكْمِ وَالْكَرَمِ

وَاسْتَرِ بِلَطْفِكَ مَا تَلَقَاهُ مِنْ خَطَا      أَوْ أَطْلَعَهُ تَتَبِعْ إِنْ كُنْتِ خَا فَهَمِ

فَنَكَمِ مِنْ جَوَادِ كَبَا وَالسَّبِقِ مَحَادَتِهِ      وَكَمِ حَسَامِ نَبَا أَوْ مَحَادِ ذُو ثَلَمِ

وَكَلْنَا يَا أُخِي خَطَاءَ ذُو زَلَلِ      وَالْعَذْرِ يَقْبَلُهُ ذُو الْفَضْلِ وَالشِّيمِ

موارد الظمان 7/1

## الإهداء

إلى من شاطراني المشقة والعناء، إلى منبع صبري ومصدر نصحي وسبب توفيتي وعزيمة عطائي، إلى من أدين لهما بكل حياتي ولن أوفيهما حقهما مهما قدمت...

أمي وأبي أدامهما الله ومتعهما بالصحة والعافية وطول العمر...

إلى من يحمل همي، ويفرح لفرحي ويحزن لحزني ...

أخي الحبيب زهرة قلبي وقرّة عيني

و أخواتي العزيزات.....

إلى صاحبات القلوب الرقيقة، واللحمة الصالحة، إلى منبع الحنان ...

خالاتي العزيزات...

أهدي هذا الجهد المتواضع.....

الباحث:

لؤي الذيابات

## شكر وتقدير

الحمد لله الذي منّ علي من فضله ونعمه مالا أحصيه، وبلغني من خزائن علمه السبيل الذي أبغيه، ويسر لي إتمام هذا العمل المتواضع. وأصلي وأسلم على نبيه الأمين سيدنا محمد بن عبد الله وعلى آله وصحبه أجمعين.

أما وقد منّ الله علي بإنهاء دراستي المتواضعة، والتي أتمنى من الله العظيم أن تكون مصدر فائدة لكل من يسهم في تطوير حركة القياس والتقويم في هذا البلد المبارك، يطيب لي أن أنسب الفضل إلى أهله و أن أتقدم من أستاذي الفاضل الأستاذ الدكتور ضرار محمد جرادات بخالص الشكر والعرفان لما أحاطني به من رعاية وحفاوة واهتمام، فكان نعم المعين. وقدم لي المساعدة بكثير من الصبر وفيض من الخلق الرفيع وكان له الفضل بعد الله في إبراز هذا العمل إلى حيز الوجود، فقد منحني من فيض علمه وصدق عنايته مما يسر لي الصعوبات التي اعترضت سبيلي أثناء إعدادي للرسالة، وامثالاً لقول رسولنا الكريم صلى الله عليه وسلم " من قال لأخيه جزاك الله خيراً فقد أبلغ في الثناء" فجزاه الله عني كل الخير ومتعه بدوام الصحة والعافية.

كما يطيب لي أن أتقدم من الأستاذ الدكتور أحمد عودة على تفضله بقبول مناقشة هذه الدراسة المتواضعة، والذي نهلت من بحر علمه الفياض ما أوصلني إلى هذا المقام، كما أتوجه إلى الدكتور الفاضل أحمد يوسف قواسمة بجزيل الشكر وعظيم الامتنان على تفضله لمناقشة هذه الدراسة و إنني لمدين له بما تعلمته منه ولما أوجاه إلي بحاله ومقاله وفعاله، كما لايفوتني أن أتوجه من الدكتورة الفاضلة أمل خصاونة بعظيم الشكر والتقدير وخالص الامتنان على تفضلها لمناقشة هذه الدراسة، وأتوجه بالشكر وعظيم الامتنان إلى الدكتور إبراهيم المحاسنة على حسن تعاونه ودمائة خلقه وكرمه، فجزاهم الله جميعاً عني خير الجزاء ومتعهم بموفقور الصحة والعافية. وختاماً أسدي شكري وتقديري وعظيم امتناني لكل من أسهم في انجاز هذه الدراسة المتواضعة بطريقة مباشرة أو غير مباشرة وخصوصاً ابن خالي العزيز محمد نايف العماري .

## - فهرس المحتويات -

الموضوع	الصفحة
الإهداء.....	ب.....
شكر وتقدير .....	ج.....
فهرس المحتويات.....	د.....
فهرس الجداول.....	ز.....
فهرس الأشكال.....	ط.....
فهرس الملاحق.....	ي.....
الملخص باللغة العربية.....	ك.....
<u>الفصل الأول: مشكلة الدراسة وأسئلتها:</u>	
خلفية الدراسة.....	1 .....
مشكلة الدراسة وأسئلتها.....	6 .....
افتراضات الدراسة.....	10 .....
محددات الدراسة.....	11 .....
تعريف المصطلحات.....	11 .....
<u>الفصل الثاني: الدراسات السابقة.</u>	
الدراسات السابقة.....	13 .....
<u>الفصل الثالث: الطريقة والإجراءات :</u>	
مجتمع الدراسة.....	29 .....
عينة الدراسة.....	29 .....

الموضوع	الصفحة
أداة الدراسة.....	30
إجراءات الدراسة.....	31
إجراءات التطبيق.....	31
إجراءات التصحيح.....	33
معايرة فقرات الاختبار.....	36
تقدير احتمال إجابة المفحوص على كل فقرة من فقرات الاختبار.....	37
اختيار فقرات الاختبار حسب ما تم تحديده لأطوال الاختبار.....	38
تصميم الدراسة ومتغيراتها.....	39
المعالجة الإحصائية.....	40

#### الفصل الرابع: نتائج الدراسة.

أولاً: النتائج المتعلقة بالسؤال الأول.....	42
ثانياً: النتائج المتعلقة بالسؤال الثاني.....	55
ثالثاً: النتائج المتعلقة بالسؤال الثالث.....	57
رابعاً: النتائج المتعلقة بالسؤال الرابع.....	61
خامساً: النتائج المتعلقة بالسؤال الخامس.....	65

#### الفصل الخامس: مناقشة النتائج والتوصيات.

أولاً: مناقشة النتائج المتعلقة بالسؤال الأول.....	67
ثانياً: مناقشة النتائج المتعلقة بالسؤال الثاني.....	72
ثالثاً: مناقشة النتائج المتعلقة بالسؤال الثالث.....	77

الصفحة	الموضوع
80.....	رابعاً: مناقشة النتائج المتعلقة بالسؤال الرابع.....
81.....	خامساً: مناقشة النتائج المتعلقة بالسؤال الخامس.....
82.....	التوصيات.....
84.....	المراجع.....
88.....	الملاحق.....
95.....	الملخص باللغة الانجليزية.....



## - فهرس الجداول -

الرقم	الجدول	الصفحة
(1)	عدد الأفراد المطلوب حسب الجنس والشعبة.....	30.....
(2)	المتوسط الحسابي والتباين ودرجة الالتواء ودرجة التفلطح للعلامات الحقيقية للاختبار الكلي والأطوال المختلفة.....	43.....
(3)	المتوسط الحسابي والتباين ودرجة الالتواء ودرجة التفلطح للعلامات المشاهدة للاختبار الكلي والأطوال المختلفة.....	44.....
(4)	المتوسط الحسابي والتباين ودرجة الالتواء ودرجة التفلطح للعلامات الخطأ للاختبار الكلي والأطوال المختلفة.....	46.....
(5)	العلامات الحقيقية لمستوى القدرة (-2) ومستوى القدرة (1) بالاعتماد على منحنى خصائص الاختبار للاختبار الكلي والأطوال المختلفة.....	51.....
(6)	العلامات الحقيقية لمستوى القدرة (-2) ومستوى القدرة (1) بالاعتماد على منحنى خصائص الاختبار للاختبار الكلي والمجموعات الأربع.....	54.....
(7)	التباين الحقيقي والتباين الخطأ و الثبات للاختبار الكلي والأطوال التي تم اختيارها.....	56.....
(8)	التباين الحقيقي والتباين الخطأ والثبات لعشرة عينات من الفقرات لكل طول من الأطوال (15، 35، 55).....	57.....
(9)	معاملات ارتباط بيرسون بين العلامات الحقيقية والعلامات المشاهدة والعلامات الخطأ لكل طول مع نظيراتها في الاختبار الكلي.....	58.....
(10)	معاملات ارتباط بيرسون بين العلامات الحقيقية و العلامات المشاهدة للاختبار الكلي و الأطوال المختلفة.....	59.....

الصفحة	الجدول	الرقم
60	معاملات ارتباط بيرسون بين العلامات الحقيقية للاختبار الكلي والأطوال المختلفة...	(11)
62	معاملات ارتباط بيرسون بين العلامات الحقيقية للأطوال المختلفة مع تقديرات علامات القدرة للاختبار الكلي ولأطوال الاختبار المختلفة.....	(12)
63	معاملات ارتباط بيرسون بين العلامات المشاهدة للأطوال المختلفة مع تقديرات علامات القدرة للاختبار الكلي ولأطوال الاختبار المختلفة.....	(13)
64	معاملات ارتباط بيرسون بين العلامات الخطأ للأطوال المختلفة مع تقديرات علامات القدرة للاختبار الكلي ولأطوال الاختبار المختلفة.....	(14)
65	معاملات ارتباط بيرسون بين علامات القدرة للاختبار بطوله الكامل وعلامات القدرة للأطوال لمختلفة فيما بينها.....	(15)
73	معاملات الثبات للاختبار الكلي ولأطوال المختلفة المحسوبة باستخدام طريقتي كرنباخ- ألفا وسبيرمان براون والطريقة المتبعة في لدراسة.....	(16)
76	معاملات الارتباط بين العلامات الحقيقية والعلامات الخطأ للأطوال المختلفة والاختبار الكلي.....	(17)

## - فهرس الأشكال -

الرقم	الشكل	الصفحة
الشكل (1)	الرسم النقطي للعوامل التي أفرزها التحليل العاملي.....	34.....
الشكل (2)	منحنى خصائص الاختبار للأطوال المختلفة والاختبار الكلي.....	48.....
الشكل (3)	منحنى خصائص الاختبار للمجموعات الأربع التي تم اختيارها حسب الصعوبة والتميز وللاختبار الكلي.....	53.....
الشكل (4)	العلاقة بين طول الاختبار والثبات المحسوب باستخدام طريقتي كرنباخ- ألفا و سبيرمان براون والطريقة المستخدمة في هذه الدراسة.....	74.....

## - فهرس الملاحق -

الرقم	الملحق	الصفحة
ملحق (1)	جدول التباين المفسر الناتج للعوامل التي أفرزتها عملية التحليل العاملي.....	88
ملحق (2)	تقديرات الصعوبة والتمييز والتخمين لفقرات الاختبار الكلي.....	90
ملحق (3)	علامات القدرة والخطأ المعياري في التقدير واحتمالات الملائمة لفقرت الاختبار الكلي والتي تم الحصول عليها من البرنامج لأول 100 مفحوص.....	92

## - الملخص -

الذيات، لؤي عبدالرؤف 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 الموضعي 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 لتحقيق 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 دلّت 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 ابتداءً 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.