Journal of Educational and Psychological Sciences (JEPS) • Vol 7, Issue 35 (2023) • P: 154 - 138
https: //journals.ajsrp.com/index.php/jeps
ISSN: 2522-3399 (Online) • ISSN: 2522-3399 (Print)

Rationing the Test of Mathematic Ability (TOMA3.) on high classes primary school students (boys) at Gunfudah Educational Directorate

Mr. Khader Mohammed Al-Zailay

College of Education | Umm Al-Qura University | KSA

Received: 05/05/2023

Revised: 16/05/2023

Accepted: 16/07/2023

Published: 30/09/2023

* Corresponding author: kaaa012@hotmail.com

Citation: Al-Zailay, KH. M. (2023). Rationing the Test of Mathematic Ability (TOMA3.) on high classes primary school students (boys) at Gunfudah Educational Directorate. *Journal of Educational and Psychological Sciences, 7*(35), 138 – 154. https://doi.org/10.26389/AJSRP.L050523

2023 © AISRP • Arab Institute of Sciences & Research Publishing (AISRP) ،Palestine ،all rights reserved.

• Open Access

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY-NC) license

Abstract: This study aimed to regulate the Test of Mathematic ability (TOMA-3) on a sample of high classes primary school students (boys) at Gunfudah Educational Directorate, The study focused on recognizing the psychometric characteristics of the test after applying it on the sample of rationing and how it agreed with the characteristics of a good test and then building the performance criteria upon which the raw grades were interpreted, To achieve this goal, the researcher applied the test on a random sample chosen from the population of the study. The sample size was (1000) students distributed among the six educational areas that formed the directorate. To answer the study questions, the researcher used the descriptive approach as well as the main study tool (TOMA-3). After making a lot of statistic analysis, the researcher was able to produce the means and standard deviation of all the sample members and age groups. The psychometric characteristics (validity and reliability) of the test were also produced. The reliability of the test was found using two different methods, first, the test-retest method then by using the internal consistency method using Alfa Chrompach equation. The validity of the test was counted by finding the quantitative indicators of the validity of the premise configuration. It was assumed that the test measures a general factor by using factor analysis. The assumption of age differences and class differences was measured using the analysis of unidirectional variance. The researcher also found the performance criteria which were in the form of normative transferred grades and the corresponding percentile ranks of the age variable, Based on the statistical analysis, the following results were achieved: TOMA-3 showed high degree of reliability and high validity indicators achieved, and able to differentiate between different age groups in this study, thus agreeing with the results of the basic study conducted in an American environment.

Keywords: mathematical ability test - primary stage - students - Gunfudah Directorate / Saudi Arabia.

تقنين اختبار القدرة الرياضية TOMA-3 على طلاب الصفوف العليا بالمرحلة الابتدائية - بنين في محافظة القنفذة التعليمية

أ. خضر محمد عمر الزيلعي

كلية التربية | جامعة أم القرى | المملكة العربية السعودية

المستخلص: هدف البحث الحالي إلى تقنين اختبار القدرة الرياضيةTOMA-3 على عينة من طلاب الصفوف العليا في المرحلة الابتدائية بنين في محافظة القنفذة، والتعرف على الخصائص السيكومترية لاختبار القدرة الرياضية بعد تطبيقه على عينة التقنين ومدى اتفاقها مع خصائص الاختبار الجيد، ثم بناء معايير الأداء التي تفسر على ضوئها الدرجات الخام، وتم تطبيق الاختبار على عينة عشوائية طبقية متعددة المراحل، وبلغ حجم العينة (1000) طالبٍ موزعين على مكاتب التعليم الستة التابعة لإدارة التعليم بمحافظة القنفذة، وللإجابة على تساؤلات البحث قام الباحث باستخدام المنهج الوصفي المقارن كما استخدم أداة البحث الرئيسية المتمثلة في اختبار TOMA-3، ثم قام بإجراء العديد من التحليلات الاحصائية، حيث تم استخراج المتوسطات والانحرافات المعيارية لجميع أفراد العينة وللفئات العمرية، كما تم استخراج المؤشرات الكمية للخصائص السيكومترية (الثبات-الصدق) للاختبار، حيث تم أولاً ايجاد معاملات الثبات بطريقتين، طريقة إعادة تطبيق الاختبار، وطريقة الاتساق الداخلي باستخدام معادلة ألفاكرونباخ، كما تم ايجاد دلائل صدق الاختبار من خلال استخراج مؤشرات صدق التكوين الفرضي، حيث تم التأكد من افتراض أن الاختبار يقيس عاملاً عاماً وذلك باستخدام التحليل العاملي، وافتراضي تمايز الأعمار وتمايز الصفوف الدراسية باستخدام أسلوب تحليل التباين آحادي الاتجاه، كما تم ايجاد معايير الأداء المتمثلة في الدرجات المعيارية المحولة وما يقابلها من رتب مئينية لمتغير العمر، وبناءً على نتائج التحليلات الاحصائية تم التوصل إلى: أن اختبار TOMA-3 يتمتع بدرجة عالية من الثبات، وبدلائل صدق كافية دلت عليها مؤشرات الصدق التي تم الحصول عليها، إضافة إلى قدرة اختبار TOMA-3 على التميز بين الفئات العمرية المختلفة في هذا البحث، متفقاً في ذلك مع نتائج البحث الأصلية لتقنين الاختبار على البيئة الامريكية.

الكلمات المفتاحية: اختبار القدرة الرياضية – المرحلة الابتدائية – التلاميذ – محافظة القنفذة / السعودية.

المقدمة.

إن التقدم والتطور الذي حدث في مجال الرياضيات والذي كان مرتبطاً بالتطور العلمي والتقدم التكنولوجي الذي يعد سمة العصر، ونتيجة لذلك فق ظهرت مفاهيم جديدة وموضوعات حديثة في مجال الرياضيات اقتضتها الضرورة لحل المشكلات الحياتية الناجمة عن التقدم العلمي والتقني، وكان لابد أن ينعكس هذا التطور على المناهج المدرسية والتي يجب عليها أن تواكب هذا التطور.

ويعد مشروع تطور مناهج الرياضيات والعلوم الطبيعية في المملكة العربية السعودية من المشروعات التربوية الرائدة، ويهدف إلى التطوير الشامل لتعليم الرياضيات والعلوم من خلال تطوير المناهج والمواد التعليمية والتقويم والتعلم الإلكتروني والتطوير المهني، وذلك بالاعتماد على ترجمة ومواءمة مواد تعليمية عالمية أثبتت فاعليتها في تحسين التعليم، والاستفادة من الخبرات العالمية المتميزة في هذا المجال بما يواكب الدول المتقدمة لبناء جيل إيجابي قادر حل مشكلاته ومشكلات مجتمعة ووطنه ويسهم بشكل فاعل في بنائهما ورقيهما ومنعتهما.

ولا بد أن يصاحب هذا التطوير تطور في التقويم التربوي وذلك لأهميته في توجيه مسار العملية التعليمية وتحقيق أهدافها، فضلاً عن أنه يلعب دوراً أساسياً ومؤثراً في تحسين وتحديث المكونات التعليمية من أهداف ومناهج وطرائق تدريس واستراتيجيات تعلم وفعاليات إدارية وغيرها (أبو سل، 1999: 175).

وتعد الاختبارات أهم أدوات وأساليب التقويم شيوعا واستخداماً، وهي أنواع مختلفة منها الشفوي، والكتابي، والنظري، والعملي، والمقالي، والموضوعي، ومنها ما هو مقنن، ومنها ما هو من إعداد المعلم.

وتعتبر الاختبارات الدولية المقننة أهم أدوات القياس والتقويم، وأكثرها شيوعا في ميدان التربية وعلم النفس، وذلك لما تقدمه للمربين والتربويين من فائدة واسعة، ومؤشرات دالة ذات علاقة بسمة التحصيل عند الطلبة.

ومن ضمن الاختبارات المقننة التي أهتم بها المربون اختبارات القدرة الرياضية، حيث قام بعض الباحثين بتطوير اختبارات في القدرة الرياضية، فقد طور شطناوي (1982) مقياساً للتفكير الاستدلالي، في حين طور أبو زينة (1986) اختباراً يقيس مظاهر التفكير الرياضي لدى طلبة المرحلة الثانوية وما بعدها، وقد طور الهباهبة (1992) اختبار للقدرة الإبداعية في الرياضيات، وطور القباطي (1993) اختبار للقدرة الرياضية لطلبة المرحلة الثانوية والجامعية.

وإن هذا البحث يسعى من هذا الجانب للاهتمام بالقدرة الرياضية وذلك من خلال تقنين اختبار القدرة الرياضية الإصدار الثالث Test of Mathematical Abilities- Therd Edition المعروف اختصاراً بـ TOMA-3.

مشكلة البحث وتساؤلاته:

نظراً لقلة توفر اختبارات مقننة ملائمة تتميز بالصدق والثبات، من أجل استخدامها لقياس القدرة الرياضية، لذا يمكن القول أن قلة توفر الأدوات والمقاييس من المشكلات البارزة في مجال البحث التربوي.

وعليه فإن البحث الحالي يسعى في المساهمة في حل هذه المشكلة من خلال تقنين أحد الاختبارات العالمية وهو اختبار القدرة الرياضية Test of Mathematical Abilities - Therd Edition (TOMA-3) على طلاب الصفوف العليا بالمرحلة الابتدائية. وبصورة أكثر دقة فإن البحث الحالي يسعى للإجابة على السؤال الرئيس التالي:

1- هل يتمتع اختبار القدرة الرياضية TOMA-3 بعد تطبيقه على طلاب الصفوف العليا بالمرحلة الابتدائية- بنين في محافظة القنفذة بخصائص سيكومترية تتفق مع خصائص الاختبار الجيد؟ ما المعايير التي يفسر في ضوئها الأداء على الاختبار؟

ويتفرع عن السؤال السابق السؤالات التالية:

2- هل يتمتع اختبار القدرة الرياضية TOMA-3 بعد تطبيقه على طلاب الصفوف العليا بالمرحلة الابتدائية- بنين في محافظة القنفذة بمعاملات ثبات تتفق مع خصائص الاختبار الجيد؟

3- هل يتمتع اختبار القدرة الرياضية TOMA-3بعد تطبيقه على طلاب الصفوف العليا بالمرحلة الابتدائية- بنين في محافظة القنفذة بمؤشرات صدق تتفق مع خصائص الاختبار الجيد؟

4- ما المعايير التي يفسر في ضوئها الأداء على الاختبار؟

أهداف البحث:

يهدف هذا البحث إلى تقنيين اختبار القدرة الرياضية TOMA-3، على طلاب الصفوف العليا في المرحلة الابتدائية بنين في محافظة القنفذة، وسيتم ذلك وفق الهدفين التاليين:

1- إيجاد الخصائص السيكومترية (الثبات - الصدق) لاختبار القدرة الرياضية TOMA-3 على طلاب الصفوف العليا بالمرحلة الابتدائية- بنين في محافظة القنفذة.
2- إيجاد المعايير التي يفسر في ضوئها الأداء على الاختبار.

أهمية البحث:

بمراجعة الدراسات والبحوث التي عنيت بالقدرة الرياضية وكذلك التي عنيت بتقنين اختبار القدرة الرياضية نجد أنها قليلة ونادرة، لذلك يرى الباحث أن أهمية هذا البحث تكمن في إثراء الميدان التربوي من خلال إضافة مقاييس عالمية موثوقة يمكن أن تساعد في تطوير العملية التربوية.

مصطلحات البحث الإجرائية:

- التقنين Standardization: يعرفه النفيعي(2001، ص7) بأن العملية التي يتم خلالها التحكم في العوامل غير المناسبة والتي يمكن أن تؤثر في عملية القياس، وتخفيض أخطاء القياس إلى حدها الأدنى عن طريق اختيار عينة ممثلة لمجتمع البحث يطبق عليها اختبار تم توحيد فقراته، وإجراءات تطبيقه، وتصحيحه بشكل يوفر للاختبار خصائص سيكومترية تتفق مع خصائص الاختبار الجيد ومن ثم توفير المعايير المناسبة لتفسير الدرجات الخام.
  - ويقصد بالتقنين في هذا البحث توحيد إجراءات تطبيق وتصحيح اختبار القدرة الرياضية الإصدار الثالث TOMA-3 بعد تطبيقه على عينة من طلاب الصفوف العليا بالمرحلة الابتدائية بمحافظة القنفذة (بنين)، مع استخراج خصائصه السيكومترية ومعايير الأداء.
- القدرة الرياضية: يعرفها Werdelin في (أبو حطب، 1974: 402) بأنها قدرة التلميذ على فهم الرموز والطرق الرياضية وتذكرها وتطبيقها.
- اختبار القدرة الرياضية الإصدار الثالث TOMA-3: هو مقياس القدرة الرياضية Test of Mathematical Abilities- Therd Edition المعروف اختصاراً بـ TOMA-3، المصمم للاستخدام مع الطلبة من ثمان سنوات تماماً إلى ثمانية عشر سنة وإحدى عشر شهراً، والذي يتميز بخصائص نفسية قياسية قوية، حيث كان يعتمد على افتراض أن الاختبارات التقليدية لا تخاطب بعض المواقف الرياضية التي يعتقد بأنها تسهم بقوة في نجاح الطلاب المحتمل في تعلم العمليات والمفاهيم الرياضية.
  - ويقصد بالقدرة الرياضية في هذا البحث الدرجة الكلية التي يتحصل عليها طالب الصفوف العليا في المرحلة الابتدائية في اختبار القدرة الرياضية TOMA-3 بعد تقنينه على عينة البحث.
- **الثبات Reliability:** هو مدى التوافق أو الخلو من الخطأ بين الدرجات التي نحصل عليها من تطبيق الاختبار أو أداة ما على نفس الشخص مرتين أو أكثر لقياس سمة معينة (الدوسري، 2000، 48).
  - يشير مفهوم الثبات إلى اتساق درجات الاختبارات لمجموعة معينة من الأفراد، حيث أنها تعكس الجوانب الحقيقية للسمة أو الخاصية أو القدرة المراد قياسها ولا تعكس عوامل الخطأ والعشوائية (علام، 2014، 89).
  - ومن خلال التعريفات السابقة يمكن أن نحدد مفهوم الثبات بأنه دقت درجات الاختبار وخلوها من الأخطاء والعشوائية، أي مدي قياس الاختبار للسمة أو الخاصية أو القدرة المراد قياسها، فدرجات الاختبار تكون متسقة من قياس إلى آخر إذا طبق عليه نفس الأداة وتحت نفس الظروف.
  - ويقصد بالثبات في هذا البحث معاملات ثبات اختبار القدرة الرياضية TOMA-3 بعد تطبيقه على عينة البحث والتي سوف يتم حسابها بطريقة معامل الاتساق الداخلي لألفا كرونباخ وبطريقة إعادة تطبيق الاختبار.
- **الصدق Validity:** يعرفه مسك Messick، 1994في (الدوسري، 1999: 48) "بأنه تقييم شامل يتم من خلاله توفير الدليل المادي والمبرر النظري اللازمين لإثبات كفاية وملاءمة ومعنى أي تأويل أو فعل يبنى على درجة الاختبار."
  - ويقصد بالصدق في هذا البحث: جمع الأدلة والمؤشرات على صدق نتائج اختبار القدرة الرياضية TOMA-3 بعد تطبيقه على عينة البحث باستخدام صدق التكوين الفرضي.

حدود البحث:

يحدد هذا البحث بتقنين اختبار القدرة الرياضية TOMA-3 على طلاب الصفوف العليا بالمرحلة الابتدائية- بنين في محافظة القنفذة والذي تم تطبيقه في العام الدراسي 1435-1436 هـ

2- الإطار النظري والدراسات السابقة.

أولاً-الإطار النظري.

2-1-1-القدرة الرياضية:

القدرة الرياضية قدرة عقلية مركبة وليست بسيطة، وتبدو هذه القدرة في إجراء العمليات الحسابية بسهولة، وسرعة، ودقة، كذلك تبدو في إدراك العلاقات بين الأعداد والعلاقات بين الوحدات المختلفة في كل مسألة حسابية، والقدرة على الاستنتاج (معوض، 1994: 158).

وتعرف جمحاوي (2000) القدرة الرياضية بأنها مقدرة الطالب على فهم واستيعاب الرموز والمفاهيم والمصطلحات والتعميمات الرياضية (عناصر المعرفة الرياضية) وتذكرها وتطبيقها في مواقف رياضية.

اختبار القدرة الرياضية TOMA-3

هو مقياس القدرة الرياضية المصمم للاستخدام مع الطلبة من ثمان سنوات تماماً إلى ثمانية عشر سنة وإحدى عشر شهراً، ذو خصائص نفسية قياسية قوية، وكان يعتمد على الافتراض بأن الاختبارات التقليدية لا تخاطب بعض المعرفة والمواقف التي يعتقد بأنها تسهم بقوة في نجاح الطلاب المحتمل في تعلم العمليات والمفاهيم الرياضية، خصوصاً في العمليات المدرسية. ومنذ إعلانه مبدئياً تم استخدامه بشكل كبير في المدارس والعيادات، ولقي إشادة به في الأوراق المطبوعة الخاصة بتقييم الرياضيات.

وصف اختبار القدرة الرياضية TOMA-3

اختبار القدرة الرياضية – النسخة الثالثة هو مقياس للقدرات الرياضية الذي تم تطبيقه على عينة مكونة من (1456) طالب تم اختيارهم بشكل عشوائي ليمثلوا 21 ولاية أمريكية. وهذا الاختبار مناسب لأفراد تتراوح أعمارهم ما بين 8 سنوات تماماً و18 سنة و11 شهراً.

2-1-2-الخصائص السيكومترية

الثبات Reliability

يشير مفهوم الثبات إلى اتساق درجات الاختبارات والمقاييس لمجموعة معينة من الأفراد، وهذا الاتساق عبر الزمن، أو اتساق صيغ مختلفة من نفس الاختبار، أو اتساق مفردات الاختبار ذاته، أو الاتساق عبر أفراد مختبرين ومصححين مختلفين (علام، 2014: 89).

يقصد بمصطلح الثبات Reliability في علم القياس النفسي دقة الاختبار في القياس أو الملاحظة، وعدم تناقضه مع نفسه، واتساقه واطراده فيما يزودنا به من معلومات عن سلوك المفحوص (أبو حطب، 1997: 101)

ويعرف الثبات إحصائيا بأنه نسبة التباين الحقيقي إلى التباين الكلي، أو مربع معامل الارتباط بين الدرجة الحقيقية والدرجة الملاحظة، وبما أننا لا نعرف مطلقاً الدرجة الحقيقة، فلا يمكن حساب الثبات بهذه الطريقة، وكل ما يتوفر لدينا هو الدرجة الملاحظة، وبالتالي لا بد من الاستفادة منها بطريقة ما لتقدير الثبات، أي الحصول على مؤشر إحصائي نحكم من خلاله على دقة القياس، ويسمى هذا المؤشر بمعامل الثبات (عودة، 2010: 341).

وهناك عدة طرق لتقدير هذه المعامل ومن أهمها:

1- **طريقة إعادة الاختبار معامل ثبات الاستقرار Stability**

تقوم فكرة هذا المعامل على حساب الارتباط بين علامات مجموعة الطلبة على الاختبار عند تطبيقه وإعادة تطبيقه مرة أخرى على نفس العينة بعد فترة زمنية على التطبيق في المرة الأولى. يفترض هنا أن السمة ثابتة خلال تلك الفترة، ولذلك يكشف هذا النوع من المعاملات عن درجة ثبات السمة المقاسة خلال تلك الفترة (عودة، 2010: 342).

2- **معامل ثبات التكافؤ Equivalence Coefficient.**

تقوم فكرة هذا المعامل على استخدام صيغتين متكافئتين للاختبار نفسه، وتطبيقها على مجموعة الأفراد نفسها الواحدة تلو الأخرى.

وتستند هذه الطريقة على فرضية أن كلاً من صيغتي الاختبار تكونان في الحقيقة متكافئتين من حيث المحتوى، ونوع المفردات، وعددها، وصعوبتها، والمتوسط والانحراف المعياري لدرجات كل منهما (علام، 2014: 94).

3- **معامل ثبات استقرار التكافؤ Stability – Equivalence.**

تقوم فكرة هذا المعامل على حساب الارتباط بين علامات مجموعة الثبات على صورتين متكافئتين للاختبار مع وجود فاصل زمني طويل نسبياً بين تطبيق الصورة الأولى وتطبيق الصورة الثانية، وذلك لتقليل أثر التذكر أو أثر الحمل، أي أن هذه الطريقة عبارة عن طريقة مركبة من الطريقتين السابقتين (عودة، 2010، 344).

4- طريقة التجزئة النصفية Internal Consistency Coefficient:

تتطلب هذه الطريقة تقسيم الاختبار إلى نصفين متكافئتين إحصائياً، وهذان النصفان يكونان مستقلين ومتكافئين في المحتوى، والصعوبة، ويتساوى كل من المتوسط والانحراف لدرجاتهما. وليس الضروري تقسيم الاختبار بحسب الأسئلة أو المفردات الفردية أو الزوجية إلا إذا حقق هذا التقسيم الشروط المذكورة.

وبعد تطبيق الاختبار كوحدة واحدة وتصحيح نتائجه، ترصد درجات مفردات النصف الأول على حدة، ودرجات مفردات النصف الثاني على حدة، ثم نوجد معامل الارتباط بين درجات كل من النصفين، ويكون معامل الثبات الحقيقي للاختبار ككل والتي من الممكن الحصول عليها من الاختبار بطولة الأصلي.

وقد أوجد العلماء معادلات عديدة لتصحيح معامل ثبات نصف الاختبار ليدل على معامل ثبات الاختبار ككل ومن هذه المعادلات ما يلي(علام، 2014، 96).

1- معادلة سبيرمان – براون Spearman-Brown Formula

وهذه المعادلة تأخذ بالاعتبار الزيادة المتوقعة لمعامل الثبات بزيادة عدد فقراته، وإن استخدام هذه المعادلة لا يقتصر على التعديل لمعامل الثبات النصفي، بل يمكن استخدامها في تقدير معامل الثبات لاختبار نرغب في زيادة عدد فقراته، أو لاختبار نرغب في تقليل عدد فقراته إذا كان الزمن لا يسمح بإعطاء الوقت الأصلي المخصص للاختبار(عودة، 2010، 345).

2- معادلة رولون Rulon Formua

استعرض رولون(1939) Rulon عن حساب الارتباط وتصحيحه لأثر الطول حسب معادلة سبيرمان – براون بمعادلة أخرى هي:

$$\Gamma_{tt} = 1 - \left[\frac{S_d^2}{S_t^2}\right]$$

3- معادلة جوتمان Guttman Formua

وهذه المعادلة مثل سابقتها تأخذ في الاعتبار اختلاف تباين نصفي الاختبار ولكن تتميز عنها بأنها أسهل في عملياتها الحسابية للحصول على معامل ثبات الاختبار وقد تم اشتقاقها بواسطة العالم النفسي جتمان وأيضاً باستخدام المعادلة الأساسية للتباين والمعادلة هي:

$$\Gamma_{tt} = 2\left[1 - \frac{(S_a^2 + S_b^2)}{S_t^2}\right]$$

وتستخدم هذه الطريقة في الاختبارات التحصيلية وذلك لقدرتها على توضيح مدى الانسجام في محتوى فقرات الاختبار وأيضاً لسهولتها، وبشكل عام هذه الطريقة ذات فائدة في الاختبارات التي تكون فيها الأسئلة والعبارات متجانسة أي أنها عبارات تقيس خاصية نفسية واحدة.(عودة، 2010)

4- طريقة الاتساق الداخلي:

ظهرت هذه الطريقة لتلافي عيوب طريقة التجزئة النصفية حيث أنها تعتمد على مدى ارتباط الوحدات أو الفقرات مع بعضها البعض داخل الاختبار وأيضاً ارتباط كل فقرة مع الاختبار ككل، حيث تقوم فكرة هذه الطريقة على حساب الارتباطات بين جميع فقرات الاختبار وكأن الاختبار قُسّم إلى عدد من الأجزاء يساوي عدد فقرات الاختبار أي اعتبار أن كل فقرة تشكل اختباراً فرعياً، وبشكل متوسط معاملات الارتباط الداخلية أفضل تقدير لمتوسط معاملات الثبات على عدد كبير من مرات التقسيم للاختبار (عودة، 2010، 349).

2-1-3-معادلات معامل الثبات بطريقة الاتساق الداخلي:

1- معادلة كودرريتشاردسون (رقم 20) KR20

وهي من أولى المعادلات التي استخدمت وتهدف للتوصل إلى قيمة تقديرية لمعامل ثبات الاختبارات غير الموقوتة (اختبارات القوة) والاختبارات التي لا يمكن تقسيمها إلى نصفين متكافئين (مراد، سليمان، 2005، 346).

وتتميز هذه المعادلة بأن قيمة الثبات لا تتأثر باختلاف صعوبة البنود وبالتالي اختلاف قيم كل منp، q ولكن يعاب على هذه المعادلة صعوبة إيجاد معامل الثبات حيث تتطلب حسابات كثيرة مما يجعلها غير مناسبة للمعلمين، مما أدى إلى ظهور معادلة أخرى تأخذ في الاعتبار سهولة إيجاد معامل الثبات وهي المعادلة التالية:

2- معادلة كودر ريتشادسون (رقم21) KR21

هذه المعادلة أكثر مناسبة للمعلمين من المعادلة السابقة لسهولتها حيث لا تتطلب سوى إيجاد المتوسط والتباين للاختبار.

ولكن يعاب على هذه المعادلة أنها تفترض أن جميع بنود الامتحان في مستوى واحد من الصعوبة وبالتالي عند اختلاف صعوبة البنود أي اختلاف قيمة p، q اختلافًا كبير فإن معامل الثبات يختلف ويكون أقل مما يجب أن يكون عليه لذلك عند اختلاف صعوبة البنود فإن المعادلة المفضلة للاستخدام هي معادلةKR20 (عودة، 2010، 351).

ويجب ملاحظة أن المعادلتين السابقتين تستخدمان في حالة الاختبارات من نوع الإجابة المنتقاة (الموضوعية) وأيضًا عندما تأخذ بنود الامتحان (واحد في حالة كون البند صحيح) و(صفر في حالة كون البند خطأ) (الدوسري، 2000، 88).

لذلك كان لابد من إيجاد معادلة تستخدم لحساب معامل الثبات في حالة الاختبارات المقالية وأيضًا في الاختبارات الموضوعية عندما لا تأخذ بنود الامتحان قيم صفر وواحد بل تأخذ درجات (1، 2، 3،..... الخ) وهي:

3- معادلة ألفا كرونباخ (α) Alpha Formula

تمثل هذه المعادلة الحالة العامة حيث تعتبر المعادلتين السابقتين لكودر ريتشاردسون حالة خاصة منها وهي تستخدم في أي نوع من أنواع الأسئلة في الاختبارات سواء كانت من نوع الأسئلة الموضوعية أو الأسئلة المقالية وسواء كانت الإجابة في الأسئلة الموضوعية تأخذ القيم (صفر، واحد) أم تأخذ قيم (1، 2، 3، ...الخ) (عودة، 2010، 350):

2-1-4-الصدق:

يعرف Cronbach، 1971 الصدق بأنه العملية التي يقوم فيها مصممو الاختبار بجمع الدلائل التي تدعم الاستدلالات المستخلصة من درجات الاختبار.

ومن هذا المعنى لصدق الاختبار يمكننا القول بأن الصدق ليس سمة ذاتية في الاختبار وهذا يعني أنه لا يوجد اختبار صادق أو غير صادق، ولكنه يعتمد على عدد من العوامل وذلك حسب ظروف الاستخدام والقرارات المرتبطة بالنتائج، وهذا يعني أنه لا يمكن اختزال صدق الاختبار على شكل معامل كما هو في الثبات، بل يتطلب جمع أكثر من مؤشر يدل على صدق الاختبار مع الأخذ في الاعتبار بأن تكون هذه المؤشرات متكاملة فيما بينها.

إن عملية التحقق من صدق الاختبار لا تتم كما يحدث في التحقق من ثبات الاختبار عن طريق استخلاص مؤشر إحصائي أو معامل نسميه معامل الصدق، بل هو عملية يحتاج فيها من صمم الاختبار إلى جمع المؤشرات والدلائل لتدعيم النتائج التي سوف يتوصل اليها من خلال درجات الاختبار، أي جمع دلائل تؤكد أن الاختبار الذي تم تصميمه يقيس ما صمم لقياسه فقط ولا يقيس شيئاً آخر أو أنه يقيس ما صمم لقياسه بالإضافة إلى شيء آخر، وعملية التحقق من الصدق تشمل طرقاً مختلفة يختلف استخدامها باختلاف نوعية الدلائل التي يرغب من صمم الاختبار الوصول إليها ومن أهمها ما يلي:

أ- صدق المحتوى:

يعني صدق المحتوى "مدى تمثيل بنود الاختبار وملاءمتها للمحتوى الذي يقيسه". أي أن بنود الاختبار عبارة عن عينة ممثلة لكافة المجالات التي يغطيها المحتوى فإن بينة الصدق هنا تعتمد على مدى تمثيل العينة (الاختبار) وكفايتها لهذا المحتوى (الدوسري، 2000: 54).

ب- صدق المحك:

يعتمد هذا المفهوم على درجة علاقة درجات الاختبار بالأداء الفعلي لمحك خارجي تستخدم فيه السمة موضع الاهتمام (مراد، سليمان، 2005، 352).

ويدل الصدق المرتبط بالمحكات على قدرة الاختبار على التنبؤ بسلوك المفحوص في مواقف محددة أو تشخيص هذا السلوك.

ت- صدق التكوين (المفهوم):

يعرف صدق التكوين بأنه مدى قياس الاختبار لخاصية أو سمة صمم الاختبار أساساً لقياسها. ويعتمد هذا النوع من الصدق على وصف أوسع ويتطلب معلومات أكثر حول الخاصية السلوكية موضوع القياس نحصل عليها من مصادر مختلفة (النبهان، 2004، 294).

المعايير:

إن الدرجة الخام لا تحمل في ذاتها معنى ذا قيمة إحصائية، ولهذا تبرز الحاجة إلى معايير بسبب النقص في المعلومات التي توفرها درجة الاختبار فهذه الدرجة لا تكفي وحدها للتعرف على الأداء.

ومن هنا تبرز الحاجة إلى وجود معيار يمكن في ضوئه تفسير الدرجة الخام، وإكسابها المعنى التوضيحي المناسب لقيمتها (الدوسري، 2000: 74).

والمعايير هي مجموعة من الدرجات التي تشتق بطرق إحصائية معينة من الدرجات الخام، بحيث تأخذ في الاعتبار توزيع الدرجات الخام التي نحصل عليها نتيجة تطبيق الاختبار على عينة عشوائية ممثلة لمجتمع معين

ثانياً- الدراسات والبحوث السابقة المرتبطة:

- **قام هوفمان Hofmann (1987)** ببناء اختبار لقياس القدرة على حل المسألة الحسابية. وللكشف عن مدى صدق هذا الاختبار تم تطوير صورتين من اختبار حل المسائل ذو الأربع خطوات، والذي يزود المعلم بالمعلومات التقييمية والتشخيصية المتعلقة بأداء الطالب في كل واحدة من الخطوات الأربعة اثناء عملية حل المسألة كما تم التحقق من صدق الاختبار وثباته باستخدام نتائج الطلاب في الصفوف: السابع والثامن. وتوصل البحث إلى أن الاختبار يتمتع بدرجة مقبولة من الصدق والثبات وتصلح لقياس القدرة في حل المسألة الرياضية.
- **أما بيكر Baker (1990)** فقد قام ببناء ثلاثة مقاييس لطلاقة الرياضية وهي اختبارات لقياس حقائق الرياضيات المختلطة وحقائق الرياضيات المنظمة وحقائق الرياضيات لمستوى المرحلة (الصف). ثم التحقق ما إذا كان للاختبارات الثلاثة قدرة تنبؤية بالقدرة الرياضية، من خلال حساب معاملات الارتباط بين هذه الاختبارات واختبار كاليفورنيا للتحصيل الرياضي. أشارت نتائج البحث إلى أن اختبار حقائق الرياضيات لم يكن لها علاقة قوية مع اختبار كاليفورنيا للتحصيل في كل المستويات، في يحين وجد أن اختبارات حقائق الرياضيات تمتعت بمصداقية وصلاحية ضعيفة.
- **كما قام داونيس وروز Daunis & Rose (1990)** بإجراء بحث كان الهدف منها التحقق من مدى صدق اختبار تحديد المستوى في مادة الرياضيات لطلبة السنة الأولى في جامعة تينيسي "Tennessee " حاول البحث أن يجد مجموعة من متغيرات التنبؤ التي يمكن استخدامها لتحديد المستوى المناسب في الرياضيات لدخول الكلية. تم اختيار بعض الشعب التي تدرس الرياضيات ضمن مستوى السنة الأولى (المبتدئين) لإشراكهم في البحث. أعطى الطلاب اختبارين لتحديد المستوى. وجمعت المعلومات الأخرى ذات العلاقة بما فيها المدرسة العليا التي تخرج منها الطالب، وعدد الساعات المعتمدة " مساقات الرياضيات" التي أخذها الطالب في المدرسة العليا. توصل البحث إلى وجود ارتباطات ضعيفة بين متنبئات ومقاييس نجاح الطالب، وتم الحصول على معدلات تنبؤ كانت مسؤولة عن (20%) إلى (50%) من التباين في المستوى الرياضي. بالإضافة إلى ذلك توصل البحث إلى وجود قدر من التحيز الجنسي في اختبارات تحديد المستوى المستخدم. وهذا التحيز أدى إلى تحديد مستويات متدنية للإناث على الرغم من تفوقهن التحصيلي على الذكور في نفس مقررات وفصول الرياضيات.
- **وأجرى الهباهبة (1992)** بحثا هدف إلى بناء اختبار للقدرة الإبداعية في الرياضيات للصفوف العليا في المرحلة الأساسية للصفوف (السابع، الثامن، التاسع، العاشر). لقياس ثلاثة عوامل اساسية للقدرة الإبداعية هي: الطلاقة والمرونة والأصالة، وقد تم إعداد أربعة اختبارات فرعية هي: تمثيل العدد (4) بطرق مختلفة، وتقسيم المربعات إلى أجزاء متساوية في المساحة والشكل، واكتشاف الفروق في مجموعات مكونة من ثلاث أعداد، وإيجاد التشابهات في الأشكال الهندسية، وطبق الاختبار على عينة مكونة من (800) مفحوص موزعين بالتساوي على الأربعة صفوف من الجنسين، وقد جمعت بيانات حول صدق الاختبار بأربعة طرق: الارتباطات الداخلية، تحليل التباين الثنائي، التحليل العاملي، الارتباط مع كل من الإبداع العام والتحصيل الرياضي والذكاء، وجمعت بيانات حول ثبات الاختبار بطريقتين هما: طريقة إعادة الاختبار، طريقة كرونباخ ألفا. وقد توصل الباحث إلى أن الاختبار يتمتع بدرجات مقبولة من الصدق والثبات مما يبرر اعتماد الاختبار والوثوق به.
- **وأجرى القباطي (1993)** بحثا هدف إلى بحث النمو الحاصل في القدرة الرياضية وعلاقتها بكل من التفكير المنطقي من جهة، والتحصيل في الرياضيات من جهة أخرى لدى طلبة المرحلة الثانوية وما بعدها، تكونت عينة البحث من (774) طالبا وطالبة منهم (563) طالباً وطالبة من طلبة الأول الثانوي الأكاديمي العلمي، و(106) طالباً وطالبة في السنة الثانية تخصص رياضيات في كليتي أربد وحوارة و(105) من طلاب السنة الثالثة والرابعة تخصص رياضيات في جامعة اليرموك. استخدم لأغراض البحث مقياس القدرة الرياضية والتفكير المنطقي وهو من تطوير الباحث، وقد حسب معامل الثبات باستخدام معادلة كودر – ريتشارد سون-KR (20). ومجالات القدرة الرياضية التي حددها في الاختبار هي: القدرة الاستدلالية، والقدرة المفاهيمية، والقدرة العددية، والقدرة المكانية. وتوصل الباحث إلى أن هناك نموا في القدرة الرياضية بتقدم الطلبة في العمر، كما تبين أن أداء الطلبة على القدرة المكانية كان أعلى، يليها القدرة الاستدلالية يليها القدرة المفاهيمية ثم أخيراً القدرة العددية، وقد تبين أيضاً أن معامل الارتباط كان موجباً بين كل من القدرة الرياضية والتفكير المنطقي، وبين القدرة الرياضية والتحصيل في الرياضيات.
- **أما بحث عبد العزيز (1995)** الذي أجري من أجل تطوير ومعايرة مقياس للقدرة الرياضية للفئة العمرية من 13- 16 سنة، يقيس ثلاثة أبعاد من أبعاد القدرة الرياضية هي: القدرة العددية، القدرة الاستدلالية، والقدرة المكانية، بحيث تتوافر لهذا المقياس دلالات صدق وثبات مناسبة، وقد طبق المقياس على عينة مكونة من (1253) طالباً وطالبة من الجنسين ومن الصفوف الثلاثة: الثامن والتاسع والعاشر، وحللت فقراته حيث حسبت مستويات الصعوبة والتمييز وفعالية المموهات للفقرات، وقد تراوحت مستويات الصعوبة للفقرات من 0.27 إلى 0.92 وتراوحت القدرة التمييزية للفقرات من 0.30 إلى 0.56 كما أن المموهات تميزت بالاتجاه

المعاكس لتمييز الفقرات، وقد جمعت بيانات حول صدق المقياس بعدت طرق هي: التجانس الداخلي، وتحليل الفقرات، والتحكيم المنطقي، وتحليل التباين متعدد المتغيرات وذلك باستخدام اختبار ويلكس لامبدا وتحليل التباين الأحادي والتحليل العملي والارتباط بين الدراجات على المقياس والدراجات الفصلية في مبحث الرياضيات (الصدق المحكي)، وللحصول على بيانات حول ثبات المقياس استخدمت اربع طرق هي: طريقة كرونباخ ألفا، والتجزئة النصفية، وثبات الصور المتكافئة، وثبات إعادة تطبيق الاختبار. وتشير البيانات التي جمعت حول ثبات وصدق المقياس أنه يتمتع بدرجة مقبولة من الصدق والثبات مما يبرر اعتماد المقياس والوثوق به في عملية تقنينه على عينة أوسع وأكثر تمثيلاً، وقد تم اشتقاق معايير للأداء على المقياس منها: معايير الصف، والمعايير المئينية، والعلامات التائية المعدلة، ونسب الذكاء الإنحرافية، كما تمت معادلة الدرجات على صورتي المقياس.

- **أجرى بني عيسى (1996)** بحث هدف إلى تطور اختبار محكي المرجع في القدرة الرياضية للمرحلة الأساسية الأولى والذي يمكن من خلاله قياس مستوى اتقان الطلاب للمهارات الرياضية الأساسية المحددة باعتبارها أهم مكونات القدرة الرياضية، واستخدامه في تحديد مستوى التمكن، وتشخيص جوانب القوة والضعف لديهم. وقد طبق الاختبار على عينة مكونة من (706) طالباً وطالبة من طلبة الصف السادس الأساسي، وحللت النتائج واستخرج معامل الصعوبة ومعامل التمييز وفعالية المموهات لكل فقرة، وقد تم اشتقاق معايير للأداء في درجات الذكاء الانحرافية، وأظهرت نتائج هذه البحث تدني مستوى التحصيل عند هؤلاء الطلبة إذا ما قورن بالمستوى الذي حددته لجنة المحكمين.
- **وأجرى أبو دية (2002)** بحثا هدف إلى التوصل إلى الخصائص السيكومترية لاختبار القدرات الرياضية، واشتقاق معايير أداء طلبة الصفوف التاسع والعاشر الأساسيين والأول الثانوي الأكاديمي للطلبة الأردنيين بعد تعديله للبيئة الأردنية. تكون اختبار القدرة الرياضية من أربعة اختبارات فرعية هي: اختبار المفاهيم واختبار الحساب واختبار المعلومات العامة واختبار مسائل القصة، بالإضافة إلى اختبار خامس إضافي هو اختبار الاتجاه نحو الرياضيات، تكون الاختبار ككل من (125) فقرة تقيس مهارات القدرة الرياضية. تم تطبيق الاختبار على عينة التقنيين التي تم اختيارها بالطريقة العشوائية الطبقية والتي بلغت (1080)طالبا وطالبة. ويرى الباحث أن القيم الناتجة الصدق والثبات بأنواعها تتفق مع العديد من الدراسات السابقة. ويرى الباحث أيضاً أن الاختبار في صورته النهائية يصلح لقياس تحصيل الطلبة وتقويمهم في الرياضيات للصفوف التي تناولها البحث، لأنه يتمتع بدلالات صدق وثبات مقبولين.
- **وأجرى أبو محفوظ (2010)** بحثا هدف إلى تقنين المستوى الأول لاختبار الرياضيات من سلسلة الاختبارات المرسلة لجاري روبرتسون، يحوي الاختبار على (40) فقرة بثلاث أبعاد: (13) فقرة تقيس بعد القدرة العدية، (14) فقرة تقيس بعد فهم المفاهيم، و(13) فقرة تقيس بعد حل المسائل الرياضية. الفقرات من نوع الاختبار من متعدد، طبق الاختبار على عينة بلغت (715) طالباً وطالبة من الصفين الثاني والثالث الأساسيين في محافظة العقبة، (347) طالبا وطالبة من الصف الثاني الأساسي، و(368) طالباً وطالبة من الصف الثالث الأساسي، حللت النتائج وتم حساب معاملات الصعوبة والتمييز لفقرات الاختبار واستخراج معاملات الصدق والثبات وحساب الأوساط الحسابية والانحرافات المعيارية للصف والجنس، كما تم اشتقاق معايير الأداء للاختبار: الرتبة المئينية، والعلاقة الزائية.
- **وأجرى الفقية (1431)** بحثا هدف إلى تقنين اختبار القدرة المبكرة في الرياضيات TEMA-3 على عينة تلاميذ الصفوف الأولية بالمرحلة الابتدائية – بنين في محافظة القنفذة التعليمية. وايجاد معايير الأداء لعينة التقنين تم اختيار عينة بطريقة عشوائية متعددة المراحل، وبلغ حجم العينة (270) تلميذاً، واستخدم المنهج الوصفي، واستخدم أداة البحث الرئسية المتمثلة في اختبار TEMA-3، وأداة مساندة هي اختبار الحساب من مقياس وكسلر المعدل لذكاء الأطفال المقنن على البئية السعودية، وبناء عل نتائج التحليلات الإحصائية المستخدمة تم التوصل إلى الاستنتاجات التالية:

1- تمتع فقرات TEMA-3 بدرجة عالية من الفعالية في معاملات الصعوبة، ومعاملات التمييز، ومعاملات ارتباط الفقرات بالدرجة الكلية باستخدام معامل الارتباط المنصف للسلسلة، وباستخدام معامل ارتباط بيرسون.
2- تمتع الاختبار بدرجة عالية من الثبات، باستخدام طريقة إعادة الاختبار، وبطريقة الصور المتكافئة، وبطريقة معامل ألفا كرونباخ.
3- توفر دلائل صدق كافية للاختبار، من خلال تحليل التباين الأحادي واختبار توكي للتحليل البعدي.
4- قدرة لاختبار على التمييز بين الفئات العمرية المختلفة.

تعقيب على البحوث والدراسات السابقة:

1- قلة الدراسات التي تناولت تقنين اختبار القدرة الرياضيةTOMA بجميع نسخة الأولى والثانية والثالثة.

2- قلة الدراسات التي تناولت اختبار القدرة الرياضية TOMA على طلاب المرحلة الابتدائية وخاصة الصفوف العليا (الرابع والخامس والسادس).
3- عدم وجود دراسات محلية تناولت تقنين اختبار القدرة الرياضيةTOMA-3 على طلاب الصفوف العليا في المرحلة الابتدائية.
4- تناولت بعض الدراسات مكونات القدرة الرياضية ككل، والبعض الأخر تناول إحدى هذه المكونات.
5- تمتعت أغلب المقاييس للدراسات السابقة بمؤشرات ثبات وصدق مقبول

3- منهجية البحث وإجراءاته.

منهج البحث:

المنهج المستخدم في هذا البحث المنهج الوصفي المقارن الذي يعتمد على دراسة الواقع أو الظاهرة، ويهتم بوصفها وصفاً دقيقاً ويعبر عنها تعبيراً كيفياً أو كمياً فالتعبير الكيفي يصف الظاهرة ويوضح خصائصها والتعبير الكمي يعطي وصفاً رقمياً يوضح مقدار هذه الظاهرة وحجمها ودرجات ارتباطها مع الظواهر الأخرى.(عبيدات، وآخرون، 2000).

مجتمع البحث:

تكون مجتمع البحث من جميع طلاب الصفوف العليا (الرابع، الخامس، السادس) في المرحلة الابتدائية لإدارة التربية والتعليم بحافظة القنفذة للبنين. والذين تتراوح أعمارهم ما بين 10 سنوات و12 سنة للعام الدراسي 1435-1436 هـ حيث يبلغ حجم مجتمع البحث (6921) طالباً موزعين على مكاتب التربية والتعليم التابعة لإدارة التربية والتعليم بمحافظة القنفذة وهي: (القنفذة، القوز، المظيلف، حلي، العرضية الشمالية، العرضية الجنوبية).

عينة البحث:

تم اختيار عينة البحث باستخدام طريقة العينة العشوائية الطبقية متعددة المراحل والتي تمثل المجتمع الأصلي تمثيلاً صادقاً، حيث يتم اختيار العينة من كل فئة من فئات المجتمع بنسبة تتناسب مع حجم عددها في المجتمع الأصلي.

وقد بلغ حجم عينة البحث (1000) طالباً في صورتها النهائية، منهم (337) طالباً يمثلون طلاب الصف الرابع بنسبة 33.7% من عينة البحث، و(334) طالباً يمثلون طلاب الصف الخامس بنسبة 33.4% من عينة البحث، و(329) طالباً يمثلون طلاب الصف السادس بنسبة32.9% من عينة البحث، وجدول (3) يوضح أسماء المدارس التي طبق فيها الاختبار، وعدد الطلاب في كل صف دراسي، لكل مكتب من مكاتب التربية والتعليم.

خصائص عينة البحث:

عندما أُخذت عينة البحث أُخذ في الاعتبار المتغيرات التي تؤثر في خصائص أفراد عينة تقنين اختبار القدرة الرياضية Toma-3 وأهم هذه المتغيرات:

1- توزيع أفراد عينة البحث وفقاً لمتغير العمر، حيث انحصرت الفئات العمرية التي شملها البحث بين 10 سنوات و12سنة.
2- توزيع أفراد عينة البحث وفقاً لمتغير الصف الدراسي، حيث انحصرت الصفوف الدراسية في الصف (الرابع، الخامس، السادس) للمرحلة الابتدائية.

أداة البحث:

قام الباحث في هذا البحث باستخدام اختبار القدرة الرياضية – الإصدار الثالث -TOMA 3 لقياس القدرة الرياضية بعد ترجمته وتقنينه، وهو مصمم للاستخدام مع الطلبة من (8 سنوات-18سنة وإحدى عشر شهراً).

4- نتائج البحث ومناقشتها.

وفيما يلي عرض نتائج البحث ومناقشتها وذلك من خلال الإجابة عن تساؤلات البحث: -

إجابة السؤال الأول: "هل يتمتع اختبار القدرة الرياضية TOMA-3 بثبات يتفق مع خصائص الاختبار الجيد؟"

وللإجابة على السؤال الأول من أسئلة الدراسة والمتعلق بثبات الاختبار تم إجراء العديد من التحليلات الإحصائية، والتي هدفت الى التحقق من ثبات اختبار القدرة الرياضية TOMA-3 في هذا البحث بطريقتين هما طريقة معامل الاتساق الداخلي وطريقة إعادة الاختبار كما هو موضح في التالي:

- **أولاً- معامل الاتساق الداخلي:**

تم إيجاد معامل الاتساق الداخلي على ثلاث مستويات وهي على مستوى العينة الكلية، وعلى مستوى الفئات العمرية، وعلى مستوى الصفوف الدراسية.

1- معامل ألفا كرونباخ على مستوى العينة الكلية:

أ- تم حساب مؤشرات الثبات بطريقة الاتساق الداخلي لاختبار القدرة الرياضية TOMA-3 من خلال حساب معامل ألفا كرونباخ (α) للعينة الكلية، حيث بلغت قيمة معامل ألفا كرونباخ للعينة الكلية (0.92)، وهي قيمة عالية تدل على تجانس فقرات الاختبار وقياسها لشيء واحد.

ب- تم حساب مؤشرات الثبات بطريقة الاتساق الداخلي للاختبار الفرعي الخامس من خلال حساب معامل ألفا كرونباخ (α) للعينة الكلية، حيث بلغت قيمة معامل ألفا للعينة الكلية (086)، وهي قيمة تتمتع بدرجات ثبات عالي، مما يعني أنه يمكن الوثوق وبدرجة كبيرة في نتائجة.

2- معامل ألفا كرونبا خ على مستوى الفئات العمرية:

أ- تم حساب معامل ألفا كرونباخ لاختبار القدرة الرياضية للفئات العمرية بالنسبة للاختبار ككل والجدول (4) يوضح قيم معاملات ثبات الاختبار.

جدول(4) قيم معاملات ثبات الاختبار Toma-3 بطريقة الاتساق الداخلي على مستوى الفئات العمرية

الفئات العمرية	معامل ألفا
10	0.85
11	0.86
12	0.90

ومن خلال دراسة الجدول السابق يتضح أن قيم معامل الثبات ألفا كرونباخ تراوحت بين (0.85 -0.90) وبمتوسط مقداره (0.88)، وتعتبر هذه القيمة عالية تدل على تمتع الاختبار بدرجة ثبات عالية.

ب- تم حساب معامل ألفا كرونباخ لاختبار القدرة الرياضية للفئات العمرية بالنسبة للاختبارات الفرعية والجدول (5) يوضح قيم معاملات ثبات الاختبار.

جدول (5) قيم معاملات ثبات الاختبار Toma-3 بطريقة الاتساق الداخلي للفئات العمرية والاختبارات الفرعية

الاختبار الفرعي	الفئات العمرية			
	10	11	12	المتوسط
مفاهيم ورموز رياضية	0.70	0.79	0.78	0.76
الحساب	0.86	0.82	0.90	0.86
الرياضيات في الحياة اليومية	0.72	0.76	0.79	0.76
مسائل كلامية	0.83	0.89	0.89	0.87
الموقف من الرياضيات	0.82	0.83	0.82	0.82

ومن خلال دراسة الجدول السابق يتضح أن قيم معاملات الثبات تراوحت بين (0.70 - 0.90) ومتوسط مقداره (0.82)، وتعتبر هذه القيمة عالية تدل على تمتع الاختبار بدرجة ثبات عالية.

3- معامل ألفا كرونباخ على مستوى الصفوف الدراسية المختلفة.

أ- تم حساب معامل ألفا كرونباخ على مستوى الصف الدراسي لكل الاختبار وجدول (6) يوضح قيم معاملات الثبات.

جدول (6) قيم معاملات ثبات الاختبار Toma-3 بطريقة الاتساق الداخلي في الصفوف الدراسية المختلفة

الصف الدراسي	معامل ألفا كرونباخ
الرابع	0.88

الصف الدراسي	معامل ألفا كرونباخ
الخامس	0.87
السادس	0.91

ومن خلال دراسة الجدول السابق يتضح أن قيم معاملات الثبات تراوحت بين (0.87 - 0.91) ومتوسط مقداره (0.89)، وتعتبر هذه القيمة عالية تدل على تمتع الاختبار بدرجة ثبات عالية.

ب- تم حساب معامل ألفا كرونباخ على مستوى الصف الدراسي والاختبارات الفرعية والجدول (7) يوضح قيم معاملات الثبات.

جدول(7) قيم معاملات ثبات الاختبار بطريقة الاتساق الداخلي على مستوى الصف الدراسي والاختبارات الفرعية

الاختبار الفرعي	الصف الدراسي			
	الرابع	الخامس	السادس	المتوسط
مفاهيم ورموز رياضية	0.76	0.72	0.73	0.74
الحساب	0.86	0.81	0.90	0.86
الرياضيات في الحياة اليومية	0.71	0.74	0.79	0.75
مفاهيم ورموز رياضية	0.86	0.87	0.89	0.87
الموقف من الرياضيات	0.81	0.85	0.84	0.83

من خلال دراسة الجدول السابق يتضح أن قيم معاملات الثبات تراوحت بين (0.71 - 0.90) ومتوسط مقداره (0.81)، وتعتبر هذه القيمة عالية تدل على تمتع الاختبار بدرجة ثبات عالية.

- **ثانياً- طريقة إعادة الاختبار:**

طبقت هذه الطريقة على عينة قصدية جزيئة من العينة الكلية تكونت من (150) فرداً وبفاصل زمني قدره أسبوعان، وتم حساب معامل ارتباط بيرسون بين التطبيقين وكانت قيمة معاملا الارتباط (0.994) عند مستوى ثقة (0.000) وتم حساب ثبات الاستقرار على ثلاث مستويات:

1- ثبات الاستقرار على مستوى العينة الكلية وعلى مستوى الاختبارات الفرعية للاختبار ككل:

تم حساب معامل الاستقرار على مستوى العينة الكلية وعلى مستوى الاختبارات الفرعية والاختبار ككل وجدول (8) يوضح قيم معامل ارتباط بيرسون بين التطبيقين:

جدول (8) قيم معاملات ثبات الاستقرار على مستوى العينة الكلية ومستوى الاختبارات الفرعية للاختبار ككل.

الاختبارات الفرعية	معامل الارتباط
مفاهيم ورموز رياضية	0.98
الحساب	0.99
الرياضيات في الحياة اليومية	0.99
مسائل كلامية	0.99
الموقف من الرياضيات	0.96
الاختبار ككل	0.99

ومن خلال دراسة الجدول السابق يتضح أن قيم معاملات الارتباط بين التطبيقين تراوحت بين (0.96 - 0.99) عند مستوى ثقة (0.01) وبمتوسط مقداره (0.98)، وتعتبر هذه القيمة عالية تدل على تمتع الاختبار بدرجة ثبات عالية.

2- **ثبات الاستقرار على مستوى الفئات العمرية للاختبار:**

أ- تم حساب معامل الاستقرار للاختبار ككل للفئات العمرية وجدول (9) يوضح قيم معال ارتباط بيرسون بين التطبيقين:

جدول (9) قيم معاملات ثبات الاستقرار على مستوى الفئات العمرية للاختبار ككل.

الفئات العمرية	معامل الارتباط
10	0.98
11	0.99
12	0.99

ومن خلال دراسة الجدول السابق يتضح أن قيم معاملات الارتباط بين التطبيقين تراوحت بين (0.98 – 0.99) عند مستوى ثقة (0.01) وبمتوسط مقداره (0.99).

ب- تم حساب معامل الاستقرار للفئات العمرية للاختبارات الفرعية وجدول (10) يوضح قيم معاملات ارتباط بيرسون بين التطبيقين:

جدول (10) قيم مؤشرات ثبات الاستقرار على مستوى الفئات العمرية للاختبار الفرعية.

الاختبار الفرعي	الفئات العمرية			
	10	11	12	متوسط
	معامل الارتباط	معامل الارتباط	معامل الارتباط	معامل الارتباط
مفاهيم ورموز رياضية	0.93	0.99	0.98	0.97
الحساب	0.96	0.99	0.99	0.98
الرياضيات في الحياة اليومية	0.97	0.98	0.99	0.98
مسائل كلامية	0.99	0.98	0.99	0.99
الموقف من الرياضيات	0.99	0.92	0.99	0.97

ومن خلال دراسة الجدول السابق يتضح أن قيم معاملات الارتباط بين التطبيقين تراوحت بين (0.92 – 0.99) عند مستوى ثقة (0.01) وبمتوسط مقداره (0.98).

3- ثبات الاستقرار على مستوى الصف الدراسي:

أ- تم حساب معمل الاستقرار على مستوى الصف الدراسي والاختبار ككل وجدول (11) يوضح قيم معامل الارتباط بين التطبيقين:

جدول (11) قيم مؤشرات ثبات الاستقرار على مستوى الصف الدراسي للاختبار.

الصف الدراسي	معامل الارتباط
الرابع	0.97
الخامس	0.99
السادس	0.99

ومن خلال دراسة الجدول السابق يتضح أن قيم معاملات الارتباط بين التطبيقين تراوحت بين (0.97 - 0.99) عند مستوى ثقة (0.01) ومتوسط مقداره (0.98).

ب- تم حساب معامل الاستقرار على مستوى الصف الدراسي والاختبارات الفرعية وجدول (12) يوضح قيم معامل ارتباط بيرسون بين التطبيقين:

جدول (12) قيم مؤشرات ثبات الاستقرار على مستوى الصفوف الدراسية للاختبارات الفرعية

الاختبار الفرعي	الصف الدراسي			متوسط معامل الارتباط
	الرابع	الخامس	السادس	
	معامل الارتباط	معامل الارتباط	معامل الارتباط	
مفاهيم ورموز رياضية	0.94	0.99	0.98	0.97
الحساب	0.96	0.99	0.99	0.98
الرياضيات في الحياة اليومية	0.96	0.98	0.99	0.98
مسائل كلامية	0.99	0.98	0.99	0.99
الموقف من الرياضيات	0.99	0.90	0.99	0.96

ومن خلال دراسة الجدول السابق يتضح أن قيم معاملات الارتباط بين التطبيقين تراوحت بين (0.90- 0.99) عند مستوى ثقة (0.01) ومتوسط مقداره (0.98).

ومما سبق يتضح التالي:

- تمتع اختبار القدرة الرياضية TOMA-3 (الدرجة الكلية – الاختبارات الفرعية) بدرجة عالية من الثبات.

- دلت معاملات ثبات الاختبار الكلي المحسوبة بطريقة ألفا كرو نباخ للاختبار ككل أنه يتمتع بدرجة عالية من الثبات حيث وصلت قيمة معامل الثبات (0.92).

كما بلغت قيمة معامل ثبات الاستقرار المستخرجة بطريقة إعادة الاختبار (0.98) وهي قيمة عالية تدل على قدرة الاختبار على إعطاء نفس النتائج تقريبا إذا أعيد تطبيقه على نفس عينة الافراد وهذا يعني أن الاختبار يتمتع بدرجة عالية من الثبات.

إجابة السؤال الثاني: "هل يتمتع اختبار القدرة الرياضية TOMA-3 بدلالات صدق تتفق مع خصائص الاختبار الجيد؟"

وللإجابة على السؤال الثاني من تساؤلات الدراسة والمتعلق بصدق الاختبار تم إجراء العديد من التحليلات الإحصائية لاستخراج دلالات صدق التكوين الفرضي والذي يشتمل على عدة مؤشرات ولقد تم الحصول على دلالات الصدق في هذا البحث من خلال التالي:

1- إجراء التحليل العاملي لأداء أفراد العينة وذلك باستخدام إحدى طرق التحليل العاملي الاستكشافي وهي طريقة المكونات الأساسية (Principal Compoents)، وتم ذلك وفقاً للخطوات التالية:

- أولاً: تم التحقق من توفر الشروط التي يجب مراعاتها في مصفوفة معاملات الارتباط عند استخدام أسلوب التحليل العاملي ووضعت النتائج في جدول (13)

جدول (13) التحقق من شروط أسلوب التحليل العاملي في بيانات البحث

م	الشروط	الأسلوب المستخدم للكشف	قيمته	الدلالة
1	أن تكون القيمة المطلقة لمحدد مصفوفة معاملات الارتباط لا تساوي الصفر (أي لا تكون المصفوفة من النوع المنفرد)	محدد مصفوفة معاملات الارتباط.	0.498	دالة لأنها أكبر من 0.00001
2	أن تكون مصفوفة معاملات الارتباط ليست على صورة مصفوفة الوحدة (تمام ارتباط المتغيرات بقيمة دالة)	مربع كاي لاختبار بارتليت.	789.644	دالة عند مستوى 0.01
3	أن تكون درجة تجانس العينة والتي يتم الحصول عليها من اختبار كايز – ماير-أولكن (K.M.O) كافية.	اختبار (K.M.O).	0.678	دالة لأنها أكبر من 0.5
4	أن تكون قيم توافق العينة من خلال أزواج المتغيرات الثنائية والتي يتم الحصول عليها من اختبار كايزر (M.S.A) مقبولة.	قيم اختبار كايزر (M.S.A) والموجودة في قطر مصفوفة الارتباط الصورية.	أكبر من5	دالة لأن جميع القيم أكبر 0.5

والنتائج الواردة في الجدول السابق تؤكد تحقق شروط استخدام أسلوب التحليل العاملي في بيانات هذا البحث، وبالتالي يمكن إكمال التحليل، والوثوق بدرجة كبيرة في نتائجه.

- ثانياً: تم استخراج العوامل باستخدام طريقة المكونات الأساسية، حيث وجد عامل واحد قيمة الجذر الكامل له (2) ويفسر 50 % من تباين الأداء على الاختبار.
- ثالثا: إيجاد معاملات تشبع الاختبارات الفرعية على العامل المستخرج

2- **تحليل التباين آحادي الاتجاه one wey ANOVA.**

تشير فرضية تمايز الأعمار إلى حقيقة أن مستوى الأداء العقلي يزداد مع التقدم في العمر، وللتحقق من قدرة اختبار TOMA-3 على التمييز بين مستويات الأداء العقلي للفئات العمرية الثلاث التي تكونت منها عينة التقنين، ثم إجراء اختبار تحليل التباين آحادي الاتجاه.

وقد تم إجراء اختبار تحليل التباين وفقاً للخطوات التالية:

- أولاً: تم التحقق من افتراضات اختبار تحليل التباين الأحادي في بيانات البحث الحالي، وكانت النتائج كالتالي:

أ- افتراض الاستقلالية: ويعتبر هذا الافتراض متحقق في بيانات البحث الحالي، وذلك لأن البيانات عبارة عن درجات كل طالب في اختبار القدرة الرياضية TOMA-3، وتعتبر كل درجة مستقلة عن الأخرى، وكذلك كل مجموعة عمرية من الفئات العمرية الثلاث مستقلة عن الأخرى حيث لا يمكن أن تتكرر درجة طالب في مجموعتين عمريتين في نفس الوقت.

ب- افتراض الاعتدالية: وللتحقق من هذا الافتراض تم إجراء اختبار (Shapiro – Wilk) وذلك لتحقق من مطابقة توزيع الدرجات في المجموعات العمرية الثلاث للتوزيع الطبيعي،

ج- افتراض تجانس التباين: وللتحقق من هذه الافتراض تم إجراء اختبار ليفن (leven) وذلك للكشف عن مدى تجانس التباين للمجموعات العمرية الثلاث

ومما سبق يتضح أن افتراضات تحليل التباين الثلاثة تحققت في بيانات البحث وهي الاستقلالية واعتدالية التوزيع وتجانس التباين، وبالتالي يمكن إجراء تحليل التباين الأحادي والوثوق في صحة النتائج المتحصل عليها.

- ثانياً: تم إجراء تحليل التباين للمجموعات العمرية الثلاث، وعرضت النتائج في جدول (17)

جدول (17) تحليل التباين الأحادي لتحديد الفروق في مستوى أداء الطلاب في الاختبار وفقاً لمتغير العمر.

مصادر التباين	مجموع المربعات	درجة الحرية	متوسط المربعات	قيمة ف	مستوى الدلالة
بين المجموعات	14572.606	2	726.303	86.256	0.01
داخل المجموعات	95609.999	1132	84.461		
المجموع	110182.604	1134			

وبدراسة النتائج السابقة نلاحظ أن قيمة (ف = 86.256) دالة احصائية عند مستوى دلالة (0.01) وهذا يعني وجود فروق ذات دلالة احصائية بين أداء مجموعتين على الأقل من مجموع الفئات العمرية.

وللتعرف على المجموعات التي كان أداؤها مختلفاً عن غيرها تم اجراء التحليل البعدي (Sheffe) وعرضت النتائج في جدول (18).

جدول (18) التحليل البعدي (Sheffe) بناءاً على متغير العمر.

الفئة العمرية(i)	المتوسط	الفئة العمرية(j)	الفروق بين المتوسطات (i - j)	مستوى الدلالة
10	46.91	11	3.69040	0.01
		12	8.85138	0.01
11	50.6	10	3.69040	0.01
		12	5.16097	0.01
12	55.76	10	8.85138	0.01
		11	5.16097	0.01

ويتضح من نتائج التحليل في الجدول السابق وجود فروق دالة احصائياً بين أداء الفئات العمرية الثلاث، حيث أن متوسط الأداء للمجموعة الثالثة أكبر من متوسط الأداء للمجموعتين الأولى والثانية، كما أن متوسط الأداء للمجموعة الثانية أكبر من متوسط الأداء للمجموعة الأولى، وهذا الاختلاف يعتبر مؤشراً على قدرة الاختبار على التميز بين مستويات الأداء في الفئات العمرية المختلفة، وهذا يعتبر دليلاً من دلائل صدق التكوين الفرضي.

3- الاتساق الداخلي:

أ- تم حساب معاملات الارتباط بين درجات الاختبارات الفرعية (مفاهيم ورموز رياضية، الحساب، الرياضيات في الحياة اليومية، مسائل كلامية) والدرجة الكلية باستخدام معامل ارتباط بيرسون، وتم وضع هذه المعاملات في جدول (19).

جدول (19) معاملات الارتباط بين الاختبارات الفرعية لاختبار Toma-3 والدرجة الكلية.

الاختبار الفرعي	معامل ارتباط بيرسون بالدرجة الكلية
مفاهيم ورموز رياضية	0.81
الحساب	0.80
الرياضيات في الحياة اليومية	0.30
مسائل كلامية	0.75

يتضح من خلال دراسة الجدول السابق أن قيم معاملات الارتباط قد تراوحت بين (0.30 – 0.81) وجميع هذه القيم مرتفعة وهي دالة عند مستوى دلالة (0.01)، وسجل أعلى معامل ارتباط بين الدرجة الكلية والاختبار الفرعي المفاهيم والرموز الرياضية (0.81)، وأدنى معامل ارتباط بين الدرجة الكلية والاختبار الفرعي الرياضيات في الحياة اليومية (0.30)، وفي ذلك دلالة على أن أبعاد الاختبارات الفرعية الاربعة (مفاهيم ورموز رياضية، الحساب، الرياضيات في الحياة اليومية، مسائل كلامية) متسقة مع الدرجة الكلية للاختبار، وهذا يؤكد صدق الاختبار.

ب- تم حساب معاملات الارتباط بين درجات الاختبارات الفرعية (مفاهيم ورموز رياضية، الحساب، الرياضيات في الحياة اليومية، مسائل كلامية) والدرجة الكلية على مستوى الفئات العمرية باستخدام معامل ارتباط بيرسون، كما يبينها الجدول (20).

جدول (20) معاملات الارتباط بين الاختبارات الفرعية والدرجة الكلية على مستوى الفئات العمرية.

الاختبارات الفرعية	معامل ارتباط بيرسون بالدرجة الكلية		
	10 سنوات	11 سنة	12 سنة
مفاهيم ورموز رياضية	0.83	0.83	0.78
الحساب	0.78	0.74	0.80
الرياضيات في الحياة اليومية	0.31	0.36	0.28
مسائل كلامية	0.71	0.68	0.74

يتضح من خلال دراسة الجدول السابق أن قيم معاملات الارتباط قد تراوحت بين (0.28– 0.83) وجميع ذهذه القيم مرتفعة وهي دالة عند مستوى دلالة (0.01)، وسجل أعلى معامل ارتباط بين الدرجة الكلية والاختبار الفرعي المفاهيم والرموز الرياضية (0.83) للفئة العمرية 10 سنوات 11 سنة، وأدنى معامل ارتباط بين الدرجة الكلية والاختبار الفرعي الرياضيات في الحياة اليومية (0.28) للفئة العمرية12 سنة، وفي ذلك دلالة على أن أبعاد الاختبارات الفرعية الاربعة (مفاهيم ورموز رياضية، الحساب، الرياضيات في الحياة اليومية، مسائل كلامية) متسقة مع الدرجة الكلية للاختبار على مستوى الفئات العمرية المختلفة، وهذا يؤكد صدق الاختبار.

نتيجة السؤال الثالث: "ما معايير الأداء لاختبار TOMA-3 وفقاً لمتغير العمر؟.

حيث أظهرت النتائج السابقة تمتع اختبار القدرة الرياضية TOMA-3 بخصائص سيكومترية جيدة (ثبات- صدق)، فقد تم إيجاد معايير الأداء والتي على ضوئها يتم تفسير الدرجات الخام، وذلك من خلال حساب المتوسط والانحراف المعياري لكل مرحلة عمرية وذلك للاختبارات الفرعية ومن ثم تحويل الدرجات الخام إلى درجات معيارية محولة بمتوسط مقداره 10 وانحراف معياري مقداره 3، كما حسبت الرتب المئينية المقابلة للدرجات الخام.

ثانياً- مناقشة وتفسير نتائج التحليل الإحصائي:

- **تمتع اختبار 3- بدرجة عالية من الثبات:**

دلت عليه معاملات الثبات التي تم حسابها باستخدام طرق TOMA الثبات المختلفة، حيث بلغت قيمة معامل ثبات الاتساق الداخلي (0، 92) وهي قيمة عالية تدل على تجانس فقرات الاختبار وقياسها لشيء واحد، وبمقارنة قيمة معامل ثبات الاتساق الداخلي مع قيم معاملات الثبات للدراسات الاخرى فإنها قريبة من قيمة معامل ثبات بحث (أبو دية، 2004) والتي بلغت (0، 92)، (أبو محفوظ، 2010 م) والتي بلغت (0، 90)، وكذلك قريبة من تلك الأبحاث التي أجريت على الصورة الأصلية لاختبار TOMA-3 والتي بلغت تقريباً (0، 96) ومن هذه الأبحاث بحث (Anastasi & Urbina، 1997، 115 p).

كما أن قيمة ثبات الاستقرار المستخرجة بطريقة إعادة الاختبار بلغت (0، 99) وهي قيمة عالية تدل على استقرار الاختبار وقدرته على اعطاء نفس النتائج تقريباً إذا أعيد تطبيقه على نفس الأفراد، وبمقارنة قيم معامل ثبات إعادة تطبيق الاختبار مع قيم معامل ثبات إعادة تطبيق الاختبار للدراسات الأخرى فإنها أكبر من قيم معامل ثبات بحث (أبو دية، 2004) حيث بلغت (0، 91)، وكذلك أكبر من الدراسات التي أجريت على الصورة الأصلية لاختبار TOMA-3 والتي بلغ متوسطها (0، 89).

والنتائج السابقة تؤكد بدرجة كبيرة تمتع اختبار TOMA-3 المقنن على طلاب الصفوف العليا بالمرحلة الابتدائية بنين في محافظة القنفذة بدرجة عالية من الثبات تتفق مع خصائص الاختبار الجيد.

- **تمتع اختبارTOMA-3 بدلائل صدق كافية:**

دلت عليها المؤشرات الكمية التي تم الحصول عليها من خلال استخدام اساليب صدق التكوين الفرضي والتي استخدمت للتحقق من التالي:

1- أن الاختبار يقيس عاملاً واحداً:

دلت نتائج التحليل العاملي لفقرات الاختبار على أن الفقرات تقيس عاملاً واحداً يمكن اعتباره عامل القدرة الرياضية والذي يفسر (50 %) من تباين الأداء على الاختبار، وهذا يتفق مع الدراسات التي أجريت على الصورة الأصلية لاختبار TOMA-3.

2- تمايز الأعمار والصفوف الدراسية:

أكدت نتائج تحليل التباين الأحادي وجود فروق في الأداء بين المجموعات العمرية المختلفة ولصالح المجموعة العمرية الأكبر، وهذا يعني أن اختبار TOMA-3 يميز بين أداء الأفراد في الأعمار المختلفة، وتتفق هذه النتيجة مع بحث (أبودية، 2004، الفقية، 1431)، وكذلك تلك الأبحاث التي أجريت على الصورة الأصلية لاختبار TOMA-3 بحث (Wilkison، Robertson، 2006) وكذلك بحث (Hammill، Hreko، Ammer، Cronin، Quinby، 1998).

3- صدق التكوين الفرضي:

أكدت نتائج حساب معامل الارتباط بين الاختبارات الفرعية والدرجة الكلية باستخدام معامل ارتباط بيرسون أنها متسقة مع بعضها البعض ومع الدرجة الكلية للاختبار ومن خلال التحليل العاملي ظهر تشبع الاختبار بعامل واحد، وفي ذلك إشارة إلى تحقيق صدق التكوين الفرضي للاختبار.

وبناء على ما تقدم يمكن القول بأنه تم الحصول على عدد من المؤشرات الكمية وبطرق مختلفة يمكن اعتبارها دلائل صدق تؤكد بأن اختبار TOMA-3 والذي تم تطبيقه على طلاب الصفوف العليا بالمرحلة الابتدائية بنين في محافظة القنفذة يتمتع بدرجة عالية من الصدق تجعل منه أداة يمكن الوثوق بها.

قائمة المراجع.

أولاً- المراجع بالعربية:

- القرآن الكريم.
- أبو حطب، فؤاد(1978). *القدرات العقلية*. دار الكتب الجامعية: بيروت. لبنان.
- أبو حطب، فؤاد، عثمان، سيد أحمد، وصادق، آمال (1993). *التقويم النفسي*: القاهرة: مكتبة الأنجلو المصرية.
- أبو دية، عيد محمود(2004). *الخصائص السيكومترية لاختبار القدرات الرياضية لطلبة الصفوف التاسع والعاشر الاساسيين والأول الثانوي الاكاديمي في الأردن*. رسالة ماجستير غير منشورة، جامعة مؤته، الاردن.
- أبو زينة، فريد كامل (1986). نمو القدرة على التفكير عند الطلبة في مرحلة البحث الثانوية وما بعدها. *المجلة العربية للعلوم الانسانية 6*، (21)، (59 -71).
- أبو محفوظ، فاطمة إبراهيم(2010). *تقنين اختبار القدرة الرياضية -المستوى الأول- للصفين الثاني والثالث الأساسين في الأردن*. رسالة ماجستير غير منشورة، جامعة مؤته، الاردن.
- أبوسل، محمد عبد الكريم(1999). *مناهج الرياضيات وأساليب تدريسها في الصفوف الأولى من المرحلة الابتدائية*. عمان: دار الفرقان.
- بني عيسى، أحمد مجلي أحمد(1996). *بناء اختبار محكي المرجع في القدرة الرياضية للمرحلة الأساسية الأولى*. رسالة ماجستير غير منشورة، جامعة اليرموك، الاردن.
- جمحاوي، إيناس (2000). *مقارنة خصائص الفقرات وفق النظرية التقليدية ونظرية استجابة الفقرة في مقياس للقدرة الرياضية*. رسالة ماجستير غير منشورة، جامعة اليرموك: الاردن.
- الدوسري، إبراهيم مبارك(2000) *الإطار المرجعي للتقويم التربوي. مكتب التربية العربي لدول الخليج*. مكتب التربية العربي لدول الخليج: الرياض.
- شطناوي، فاضل(1982). *تطوير التفكير الرياضي عند طلبة المرحلة الثانوية في الاردن*. رسالة ماجستير غير منشورة، جامعة اليرموك، الاردن.
- عبد الرحمن، سعد(1998). *القياس النفسي النظرية والتطبيق*. القاهرة: دار الفكر العربي.
- عبد العزيز، نبيل عبد الله محمد(1995). *تطوير ومعايرة القدرة الرياضية للفئة العمرية من 13-16 سنة في الأردن*. رسالة ماجستير غير منشورة، الجامعة الأردنية، عمان.
- عبيدات، ذوقان، عبد الخالق، كايد، وعدس، عبد الرحمن (2000)، *البحث العلمي مفهومه وأدواته وأساليبه*. دار أسامة للنشر والتوزيع: الرياض.
- علام صلاح، صلاح الدين(2014). *الاختبارات والمقاييس التربوية والنفسية*. عمان: دار الفكر.
- عودة، أحمد سليمان(2010). *القياس والتقويم في العملية التدريسية*، دار الأمل: إربد.
- الفقية، حسن أحمد عبد الله (1431). *تقنين اختبار القدرة المبكرة في الرياضيات TEMA-3 على عينة من تلاميذ الصفوف الأولية بالمرحلة الابتدائية -بنين في محافظة القنفذة*. رسالة ماجستير غير منشورة، جامعة ام القرى: مكة المكرمة.

- القباطي، عبد السلام محمد سعيد(1993). *القدرة الرياضية وعلاقتها بالتفكير المنطقي والتحصيل في الرياضيات لدى طلبة المرحلة الثانوية وما بعدها*. رسالة ماجستير غير منشورة، جامعة اليرموك: الاردن.
- محمود، إبراهيم وجيه(1985). *القدرات العقلية خصائصها وقياسها*. دار المعارف.
- مراد، صلاح أحمد؛ سليمان، أمين علي (2005). *الاختبارات والمقاييس في العلوم النفسية والتربوية "خطوات إعدادها وخصائصها"*. القاهرة: دار الكتاب الحديث.
- معوض، خليل ميخائيل(1994). *القدرات العقلية*، الإسكندرية، دار الفكر الجامعي.
- النبهان، موسى(2004) *أساسيات القياس في العلوم السلوكية*. عمان: دار الشروق.
- النفيعي، عبد الرحمن عبد الله أحمد(2001)**.** *تقنين اختبار رافن للمصفوفات المتتابعة المتقدم على طلاب المرحلتين المتوسطة والثانوية بمنطقة مكة المكرمة*. رسالة ماجستير غير منشورة، جامعة أم القرى، مكة المكرمة.
- الهباهبة، عبد الله عيد (1992). *بناء اختبار للقدرات الإبداعية في الرياضيات للصفوف العليا من المرحلة الأساسية*. رسالة دكتوراة، الجامعة الاردنية، عمان.

ثانياً- المراجع بالإنجليزية:

- Anstaasi, A., & Urbina, S. (1997). *Psychological testing (7 th ed)***.** New York: prentice-Hall.
- Baker, J. M (1990). *Develoment and validation of mathematics fluency measure*, Doctoral Dissertation, University of Pittsburgh: Dissertation Abstracts International, 51 (5), 1572-A.
- Brown, V. L., & McEnter, E. (1984). *Test of mathematical abilities*. Austin, TX: PRO-ED.
- Cohen, L. G., & Spenciner, L. G. (1998). *Assessment of children and youth*. New York, NY: Longman.
- Cronbach, L. J. (1971). Test validation. In R. L. Thorndike (Ed.). Educational Measurement (2nd Ed.). Washington D. C. : American Council on Education.
- Daunis, M. Rose. (1990). *The validity of placement testing in freshman mathematics at the University of Tennessee*, Doctoral Dissertation, University of Tennessee.
- Davison, M. L. (1985*). Review of Test of Mathematical Abilities. In J. V. Mitchell, Jr. (ED.), Ninth mental measurements yearbook (Vol. ll, pp. 1578-1579).* Lincoln, NE: the Buros Institute of Mental Measurements.
- Hammill, D. D., Brown, L., & Bryant, B. R. (1992). *A consumer's guide to tests in print (2nd ed.).* Austin, TX: PRO-ED.
- Hills, J. (1981). *Measurement and Evaluation in the Classroom*. Ohio: A Bell and Howell Company.
- Hofmann, R. S. (1987). *Construction and Validation of a testing instrument to measure problem solving skills of students*, Doctoral Dissertation, Temple University: Dissertation Abstracts International, 48(2) 373-A.
- Howell, K. W. (1989-1990). *Test of mathematical abilities (TOMA)*. Diagnostique, 5(1-4), 210-217.
- McLoughilin, J. A., & Lewis, R, B. (1990). *Assessing special students*. New York, NY: Merrill Macmillan.
- McLoughilin, J. A., & Lewis, R, B. (2005). *Assessing students with special needs (6th ed)*. Upper Saddle River, NJ: Pearson/Merrill/Prentice-Hall.
- Messick, S. (1994). *The Interplay of Evidence and Consequences in the Validation of Performance Assessment*. Educational Researcher.
- Overton, T. (1992). *Assessment in special education: An applied approach*. New York, NY: Merrill Macmillan.
- Pierangelo, R., & Giuliani, G. A. (1998). *Special educator's complete guide to 109 diagnostic tests*. San Francisco, CA: Jossey-Bass.
- Pierangelo, R., & Giuliani, G. A. (2002). *Assessment in special education*. Boston, MA: Allyn & Bacon
- Salvia, J., & Ysseldyke, J. E. (1991). *Assessment in special education and remedial education*. Boston, MA: Houghton Mifflin.
- Salvia, J., & Ysseldyke, J. E. (2001). *Assessment (8th ed)*. Boston, MA: Houghton Mifflin.
- Schnepel, K. C. (2003). *Test of mathematical abilities, second edition (TOMA-2)*. In M. I. Friedman, C. W. Hatch, J. A. Jecobs, A. C. Lau-Dickinson, A. B. Mickerson, & K. C. Schnepel (Eds.), Educator's handbook on effective testing (pp. 446-452). Columbia, SC: Institute for Evidence-Besed Decision-Making in Education
- Taylor, R. (1989). *Assessment of exceptional students: Educational and psychological procedures*. Englewood Cliffs, NJ: Prentice-Hall.