

**الكشفُ عن أنماط التحِيز في الاختبارات التحصيلية لمعلمات الرياضيات
للمراحل المتوسطة في مدينة تبوك**

شاهر خالد سليمان

غادة محمد حيان العنزي

أستاذ مشارك في القياس والتقويم
كلية التربية والآداب - جامعة تبوك
المملكة العربية السعودية

ماجستير القياس والتقويم
كلية التربية والآداب - جامعة تبوك
المملكة العربية السعودية

(قدم للنشر بتاريخ ٢٤/١٢/١٤٣٩هـ، وقبل للنشر بتاريخ ٨/٢/١٤٤٠هـ)

الكشف عن أنماط التحيز في الاختبارات التحصيلية لعلمات الرياضيات للمرحلة المتوسطة في مدينة تبوك

شاهر خالد سليمان

غادة محمد حيان العنزي

أستاذ مشارك في القياس والتقويم
كلية التربية والأداب - جامعة تبوك
المملكة العربية السعودية

ماجستير القياس والتقويم
كلية التربية والأداب - جامعة تبوك
المملكة العربية السعودية

ملخص الدراسة

هدف هذا البحث إلى الكشف عن أنماط التحيز في الاختبارات التحصيلية التي أعدّتها معلمات الرياضيات لمرحلة المتوسط في مدينة تبوك، وتم استخدام المنهج الوصفي التحليلي، وتم تحليل الاختبارات إلى فقراتٍ حسب تصنيف بلوم، وتحليل الفقرات التي تتبعها كلٌّ فصل من فصول الكتاب المقرر، حيث شمل البحث المدارس الحكومية المتوسطة للبنات في مدينة تبوك، وأخذت عينة عشوائية بسيطة من اختبارات معلمات الرياضيات النهائية لمرحلة الأول متوسط للفصل الدراسي الثاني لعام (١٤٣٧ - ١٤٣٨هـ) أعدّتها معلمة رياضيات، وبلغ عدد فقرات الاختبارات (١٢٤٤) فقرة اختبارية، ونسبة الاتفاق بين المحللين بلغت ٪٨٩. أظهرت النتائج أنَّ ٥٦٪ لم يتزمنَ بنسب عدد الأسئلة حسب تصنيف بلوم للأهداف المعرفية (معرفة - استيعاب - تطبيق فما فوق)، ويتحيزُ مستوى المعرفة والاستدلال مقابل مستوى التطبيق فما فوق في فقرات اختباراتهنَّ، وأنَّ هناك فروقاً ذات دلالة إحصائية (٥٠،٥٪) بين متوسطات الفقرات الملاحظة والفقرات المتوقعة، ونحو ٣٦٪ لم يتزمنَ بتوزيع فقرات اختباراتهنَّ حسب النسب لكلٌّ فصل من فصول الكتاب المقرر. وقد تبيَّن من النتائج التحيز لمحظى فقرات الكتاب المقرر في فقرات اختباراتهنَّ للفصل الثالث والفصل الرابع، كما أنَّ هناك فروقاً ذات دلالة إحصائية (٥٠،٥٪) بين متوسطات الفقرات الملاحظة والفقرات المتوقعة.

الكلمات المفتاحية: أنماط التحيز، الاختبارات التحصيلية، المعلمات، المرحلة المتوسطة، تبوك.

Investigating the Bias Pattern of the Achievement tests by Mathematics female teachers at the intermediate schools in Tabuk

Ghadah Mohammed Alanazi

Master's degree in Measurement and Evaluation
College of Education and Arts - University of Tabuk
Kingdom of Saudi Arabia

Shaher Khaled Suliman

Associate professor of Masurement and Evaluation
College of Education and Arts - University of Tabuk
Kingdom of Saudi Arabia

Abstract

This study investigates the bias pattern of the achievement tests of Mathematics prepared by female teachers at the intermediate schools in Tabuk. To achieve this task, the researcher used analytical and descriptive analysis method. Both tests and units of the textbook were analyzed using Bloom's Taxonomy. This study focuses on the public intermediate female schools in Tabuk as the study populations and randomly selected samples of the tests presented by teachers of the first intermediate schools (36 teachers at the academic year 1437/1438). The number of the investigated items of the tests is 1244. The percentage of agreement among the researchers was 89%. The study showed that 56% of the female teachers did not follow the percentage of the number of questions according to Bloom's Taxonomy for Cognitive Objectives (Knowledge, Comprehension, Application and above).

The study also showed that there was a bias towards the questions related to knowledge and comprehension in contrast to application and above according to Bloom's Taxonomy. There were significant clear statistical differences at ($\alpha \leq 0.05$) among the actual parts of the tests and the expected parts for all the cognitive objectives of Bloom's Taxonomy levels.

The study also showed that 36% of the female teachers did not distribute the parts of the tests equally to the units of the textbook, as there was a bias towards the 3rd and 4th chapter of the textbook. The statistical results also showed the significance at ($\alpha \leq 0.05$) among the average of the actual tests and the expected tests for each unit in the textbook.

Keywords: Patterns of Bias, Achievement Tests, teachers, intermediate schools, Tabuk.

مقدمة

لقد حظي موضوع التحيز بعناية الباحثين في القياس النفسي والتربوي، وخاصة في المقاييس النفسية، والاختبارات التحصيلية التي يتم بناؤها من قبل المعلمين، والاختبارات المقننة وتطويرها وتقنيتها وإعدادها في مجتمع ما ويتم نقلها أو ترجمتها لاستخدامها في مجتمعات أخرى؛ الأمر الذي من الممكن أن يؤدي إلى ظهور مشكلة التحيز ضد ثقافة ما، أو مشكلة تحيز الفقرات لفئة دون أخرى.

يرى النبهان (٢٠٠٤: ١٢) "أن التحيز يعتبر من أخطاء القياس المنتظمة؛ بمعنى أنه يؤثر في درجات المفحوصين بمقدارِ واتجاهِ ثابتين، وتعتبر الأخطاء المنتظمة أخطاءً يسهل اكتشافها وضبطُ أثرها في درجات المفحوصين". ويشير علام (٢٠١١: ٣٣٤) "إلى ضرورة مراعاة من يقوم بإعداد الاختبارات المقننة لتبين الأفراد المختبرين في خلفياتهم الثقافية، وخصائصهم المميزة عند انتقاء مفردات الاختبار؛ لأنَّ عدم مراعاة خصائصهم وقدراتهم يؤثر، دون أدنى شكٍّ، في نتائج الاختبار ويقلل من شأن هؤلاء الأفراد بطريقه أو بأخرى".

ويعتبر عودة (٢٠١٠: ٧٣) "التحيز بجميع أنواعه من أبرز المشكلات والأثار السلبية التي تترتب عليها نتائج الاختبارات وذكر منها: تحيز عينة الأسئلة، تحيز صياغة الفقرات - فقد تكون الصياغة لصالح فئة دون أخرى -، التحيز في محتوى الفقرة".

يتربَّ على كل ذلك التزامُ لدى المعلمين يحتم عليهم تقديم أفضل تعليم ممكن لطلبتهِم، وهذا معناه أنه يجب أن تكون لدى هؤلاء المعلمين بعضُ المناهج والوسائل التي تمكّنهم بشكل فعال وصادق من تقويم فاعليَّة تدريسيهم. ويعتبر اختبار التحصيل الصفي أحد تلك الوسائل. لكن مع التزايد في استخدام الاختبارات العامة والاختبارات الأدائية لتقدير مدى تقدُّم الطالبة في المدرسة، فمن الصعب أن يصوغ المرء مفهوماً عن نظام تعليمي لا يتعرض فيه الطالب إلى الاختبارات من صنع المعلم، وعلى الرغم من أنَّ الأهداف الخاصة للاختبارات والاستخدام المنشود لنتائج الاختبارات تختلف من مدرسة لأخرى، فيحتاج المعلم أن يراعي مدى الدقة والتفصيل في التخطيط من أجل أن يتم ربط الأهداف التعليمية والخطة الاستراتيجية المستخدمة في التدريس والمادة العلمية والمنهج بشكل مُقنع ذي معنى. فعليه؛ يدرك أغلب المعلمين أهميَّة امتلاك منهج منظم للتحقق من مدى إنجاز طلبتهم لأهدافهم التعليمية. فأهمُّ الأخطاء التي يقع فيها المعلمون عند إعداد الاختبارات هي عدم تمثيلها للأهداف المحددة والمحتوى المقصود، وهذا يُعتبر من التحيز سواء كان بقصد أو بغير قصد؛ لعدم التزامهم بجدول المواصفات الذي يضمن أن تكون نسب الأسئلة متوازنة من حيث الأهداف وشموليتها لمحتوى الكتاب.

ونظراً لشيوع الاختبارات، ولأهمية القرارات التي تُبنى على نتائجها؛ تبرز تساؤلاتٌ عن مدى سلامة الاختبارات وصدق نتائجها، ودقة القرارات التي تُبنى عليها، حيث يفترض باني الاختبار، أنَّ الاختبار الذي يُعدُّ يقيس السمة المراد قياسُها من حيث تمثيل الأسئلة للأهداف التعليمية واحتواها لكافَّة مواضع المحتوى، ولكن من الممكن أنْ يقيس سمة أخرى إلى جانب السمة التي أُعدَّ لقياسها، بحيث تؤثر على علامات السمة المقاسة،

وأظهرت عدة دراساتٍ تجريبية أنَّ العلامات على الاختبارات تتأثر بعوامل غير التي يفترضها الباحث، منها تحيز الاختبار وعدم التقييد بالقواعد المتبعة في بناء الاختبار. (Mentzer, 1982). كما أوضحت ونرسما (winrsma, 1985) أنَّ للتحيز عدة معانٍ وصور، ولكنها تتفق في أنَّ التحيز خطأً منتظم في القياس. ويكون الاختبار متحيِّزاً إذا كانت علامات الأفراد أو المجموعة عليه متسيقة الارتفاع أو الانخفاض، وتتعدد مصادر التحيز، منها ظروف تطبيق الاختبار وتعليماته، ومستويات الأهداف التي تقيسها. (wotrogan, 1976)

يتضح مما سبق، أنَّ هناك عدة أنواع من التحيز في الاختبارات التحليلية العامة واختبارات المعلمين الصفيَّة، منها: التحيز لمستوى الأهداف، والتحيز في المحتوى الذي ظهر بصور متعددة منها: التحيز لموضوعات معينة دون أخرى، والتحيز في فقرات الاختبارات غير محددة المحتوى.

وجاء هذا البحث امتداداً للبحوث السابقة للكشف عن أنماط التحيز في الاختبارات التحليلية التي أعدَّتها معلمات الرياضيات للمرحلة الأولى متوسط في مدينة تبوك؛ للعمل على تطوير هذه الاختبارات، ومراعاة قدرات ومستويات التفكير لدى الطالبات.

مشكلة البحث

يعتبر التقويم الموجَّه لعملية التعلم؛ لذلك فإنَّ العنصر الأساس في العملية التربوية، ولما كانت معظم المؤسسات التربوية تعتمد على نتائج الاختبارات؛ فإنَّ الخطوة الأولى لاعتماد نتائج هذه الاختبارات يكون بتحليلها ومعرفة مدى احتواها على أنماط التحيز في مستويات الأهداف المعرفية والمحتوى. وتشير الدراسات في مجال بناء الاختبار وعمليته إلى أنَّ تفسير نتائج الاختبار يمكنُنا من التعرف على مستوى الطلاب من حيث قوتهم وضعفهم، ووضع خطة علاجية لهم؛ فالاختبارات لها أهميَّة في عملية التقويم، إذ تعطينا فكرة واضحة عن إمكانية وضع خطة علاجية للطلاب الضعاف ومستوى نشاطاتهم، ومن خلالها يمكننا وضع الأساس التي من خلالها نعدل في مستوى الأساليب وتقنيات التدريس.

ويمكن القول بأنَّ الاختبارات تُعدُّ أمراً ضرورياً في تحديد القدرات التحليلية لدى المتعلمين؛ لذا يمكن أن نعتبرها الأساس في الاختبار التحليلي. وأظهرت عدة دراساتٍ تجريبية أنَّ العلامات على الاختبارات تتأثر بعوامل غير التي يفترضها الباحث، منها: تحيز الاختبار وعدم التقييد بالقواعد المتبعة في بناء الاختبار. كما في دراسة كوفمان وكاردوس (Kauffman & Kardos, 2002) التي هدفت إلى معرفة مدى اكتساب المعلمين الجدد مهارات إعداد الاختبارات التحليلية، حيث توصلت الدراسة إلى أنَّ المعلمين يضعون أسئلة لا تقيس أهدافَ المنهج، مع ضعف في مهارات بناء الاختبارات التحليلية.

وكذلك، أجرى الخوالدة والمشاعلة (٢٠٠٧م) دراسةً هدفت لتحليل أسئلة الثانوية العامة لمبحث العلوم الإسلامية في المملكة الأردنية الهاشمية، شملت ١٤ امتحاناً وزارياً، تضمنت (٩٨٥) سؤالاً؛ لبيان درجة تمثيلها للمستويات المعرفية، وقد توصلت نتائج الدراسة إلى أنَّ الأسئلة قد تركزت في مستويات التذكر والفهم والتحليل، وقد خلَّتْ من مستويات التطبيق والتركيب والتقويم.

كما لوحظ في البحوث التجريبية السابقة أن هناك تحيزاً في محتوى فقرات اختبارات المعلمين الصفيحة، وظهر التحيز للمحتوى بعدة صور وأشكال منها: التحيز لوحداتٍ أو موضوعاتٍ محددةٍ في المحتوى دون أخرى؛ ويقصد بهذا النوع عدم تحقق صدق المحتوى؛ فالاختبار الصادق هو الذي يمثل السلوك الذي نريد قياسه بدقة، وكلما زاد تمثيلُ جوانب السلوك في الاختبار زاد الاختبار صدقاً، فإذا أراد المعلم وضع اختبار في وحدة، فعندما تكون نسبة فقرات الاختبار التي تمثل وحدة معينة أعلى من غيرها من الوحدات مع أنها متكافئة من حيث الأهمية، عندها يعتبر الاختبار متحيزاً لتلك الوحدة؛ ومن هنا كان الهدف الحصول على اختباراتٍ تحصيلية بأفضل صورة، مما دعا للبحث عن الأسباب التي أدت إلى تحيزها.

أهمية البحث

الأهمية النظرية

تبعد أهمية هذا البحث من القدرة على تشخيص واقع الاختبارات التحصيلية التي تعدّها معلمات الرياضيات؛ للوقوف على نقاط القوة وتعزيزها ومعالجة نقاط الضعف فيها. كذلك تزويد المعلمات بتوجيهات وإرشاداتٍ تساعدهنَّ في تطوير الاختبارات، وتحسين مستواها ومواقفها، بحيث تتوافق فيها مواصفاتُ الاختبار الجيد، والتعرف على الأهمية النسبية لأنواع المختلفة من الأسئلة التحصيلية التي تعدّها معلمات الرياضيات.

الأهمية التطبيقية

إن معرفة نقاط القوة ونقاط الضعف يساعد على تشخيص واقع الاختبارات التحصيلية التي تعدّها معلمات الرياضيات من خلال استخدام أشكال مختلفة من الأسئلة تحاكي الفروق الفردية عند الطلبة، ونتائج هذا البحث قد تجد عناية لدى القائمين على العملية التعليمية لأنها ستقدم للمؤولين عن التعليم معلوماتٍ عن كفاءة الاختبارات التحصيلية التي تعدّها معلمات الرياضيات بمدينة تبوك، ومدى الحاجة إلى إعادة النظر فيها؛ بغرض تحسينها وتطويرها.

أهداف البحث

تبعد أهدافُ البحث مما يلي:

- تحري التحيز في الاختبارات التحصيلية التي وضعتها معلمات الرياضيات للمرحلة الأول متوسط في مدينة تبوك؛ لضبط عملية إعداد الاختبارات بشكل عادل يعكس على نتائج الطالبات.
- التأكيد على أهمية جدول المواصفات لبناء الاختبارات؛ لتوزيع فقرات الاختبار على فصول الكتاب حسب وزنها وأهميتها؛ للتقليل من احتمال التحيز لمحتوى الأهداف والمحتوى.
- الكشف عن مدى التزام المعلمات بمواصفات الاختبارات المتفق عليها، والبحث على تطويرها، ومراعاة قدراتِ ومستوياتِ التفكير لدى الطالبات.

مصطلحات البحث

- ١- الاختبارات التحليلية: طريقة منظمة لتحديد مستوى إتقان الطالب لمعلومات ومهارات في مادة دراسية سبق أن تعلمها بصفة رسمية، وذلك من خلال إجابته على عينة من الأسئلة التي تمثل محتوى المادة الدراسية.
(عودة، ٢٠١٠: ٦٢)
- ٢- تحيز عينة الأسئلة: إعطاء بعض الوحدات أهمية أكثر من وحدات أخرى، أو بعض القدرات الفرعية أكثر من القدرات الأخرى. (سليمان، ١٩٨٨: ٧)
- ٣- تحيز في مستوى الأهداف: هو النسب المئوية لفقرات الاختبارات التي تقيس محتوى الأهداف المعرفية التالية (التذكر، الاستيعاب، التطبيق فما فوق) والتي تختلف عن النسب المئوية التي يقررها المختصون في هذا المجال. (سليمان، ١٩٨٨: ٨)
- ٤- التحيز في محتوى الاختبار: هو النسب المئوية لفقرات الاختبارات التي تمثل وحدات كتب الرياضيات المقررة للمرحلة المتوسطة، والتي تختلف عن النسب المئوية التي يقررها المختصون في هذا المجال. (سليمان، ١٩٨٨: ٨)
- ٥- مستوى التطبيق فما فوق: الأنشطة العقلية التي وضعها بلوم، وتشمل مستويات (التطبيق، التحليل، التركيب، والتقويم).
- ٦- قواعد بناء الاختبارات: مجموعة من المعارف والمهارات الالازمة لبناء الاختبارات، والتي يفترض على المعلم أن يكون على دراية بها؛ لكي يمكن من بناء اختبارات تحليلية تتمتع بمواصفات جيدة.
- ٧- تحليل المحتوى: يقصد به تفصيل المادة العلمية إلى معلومات أساسية مطلوب تعلمها من الطالب.
(طعيمة، ٢٠٠٤: ٧١)
- ٨- جدول المواصفات: جدول ذو بُعدِين، يوضح علاقة مخرجات التعلم (أهداف المقرر) بمحتوى المقرر، حيث توضع الأهداف التعليمية أفقياً والموضوعات رأسياً، ويتم تحديد الأوزان النسبية على صور نسب مئوية.
(ملحم، ٢٠١٥: ٢١٥)

حدود البحث

الحدود الزمنية: طُبّق البحث على الاختبارات النهائية لمعلمات الرياضيات في الفصل الدراسي الثاني لعام ١٤٣٧ - ١٤٣٨ هـ.

الحدود المكانية: المدارس الحكومية للبنات، للمرحلة المتوسطة، في مدينة تبوك.

الحدود الديموغرافية: معلمات الرياضيات لمرحلة أول متوسط.

الحدود الموضوعية: اقتصر البحث على نمطين من التحيز في اختبارات المعلمات النهائية للرياضيات هما: التحيز للأهداف المعرفية، والتحيز لمحتوى كتاب الأول متوسط.

الإطار النظري

التحيز

تقدّمت أساليب وطرق الكشف عن تحيز فقرات الاختبار في أواخر السبعينيات، كما تعددت الأساليب الإحصائية للكشف عن التحيز، بحيث أصبحت بحوث دراسات تحيز فقرات الاختبارات من اختصاص الباحثين في مجال القياس والتقويم؛ ومن ثمّ أصبحت طرق دراسة تحيز الاختبارات امتداداً طبيعياً لطرق التحليل التقليدية، وبمثابة إجراء أساسي في بناء هذه الاختبارات، وزادت العناية بضرورة تقييم الاختبارات مما قد يشوبها من تحيز قبل تطبيقها؛ ومن ثمّ تزايدت العناية في العقود الأخيرة بظاهرة تحيز الاختبار، أو تحيز فقراته؛ فأجريت الأبحاث، وأقيمت المؤتمرات لبحث مشكلة التحيز في الاختبارات، أو في تحيز فقراتها.

• معنى التحيز:

يعتبر التحيز من الأخطاء المنتظمة، التي تؤثر عادةً في اتجاه واحد؛ أيًّا: بالزيادة أو النقص، وهناك العديد من الأمور التي تسهم في ظهور مثل هذه المشكلة منها: عدم كفاءة المعلمين، التي تمثل في غموض صياغة بعض الفقرات، والتركيز على هدف معين، أو الاقتصار على محتوى محدد، وإعطاء معلومات مضللة، وإجاباتٍ عشوائية من قبل المفحوصين، وعدم وضوح تعليمات الإجابة عن الأسئلة.

ونظراً لتنوع مجالات التحيز ونظرة الباحثين لهذا المفهوم؛ فإنه ليس بالإمكان إعطاء تعريف واحد للتّحيز، والمتصسي لدراسات التّحيز يواجه عدداً من التعريفات التي يمكن قبول بعضها ورفض بعضها الآخر، حيث إنَّ صعوبة الحصول على تعريف عامٌ أو شامل للتّحيز تعود إلى التفاعل المعقد بين كيفية قياس اختبار ما وضع لقياسه، وما إذا كان محتوى الاختبار مناسباً لاستخدام معين.

كما أشار القاطعي (١٩٩٣: ٩) إلى أنَّ "التّحيز" بصورة مختلفة يعود إلى تفضيل مجموعة دون أخرى ليس بناءً على الخصائص التي يقيسها أو يفترض أنْ يقيسها الاختبار، إنما بناءً على جوانب ليست ذات علاقة بالسمة المقاسة، وهذه النظرة الشاملة لمفهوم التّحيز يمكن أنْ يندرج تحتها مفهوم التّحيز في الاختبار (تحيز الاختبار) ومفهوم التّحيز بالاستجابة (تحيز البنود).

ولذلك؛ يمكن أنْ تُعرَّف الباحثة التّحيز بالنسبة للأهداف بأنه: "تركيز عينة الأسئلة على مستوى معين من مستويات الأهداف المعرفية بالزيادة، دون مراعاة الأوزان النسبية المتفق عليها مسبقاً لكلٍّ مستوى من مستويات الأهداف المعرفية".

كما تُعرَّف التّحيز بالنسبة للمحتوى بأنه: "تركيز عينة الأسئلة على فصل معين من فصول الكتاب المقرر بالزيادة، دون مراعاة الأوزان النسبية المتفق عليها مسبقاً لكلٍّ فصل من فصول الكتاب المقرر".

تحيز الاختبار:

إنَّ صدق علامات الاختبار ضروريٌّ جداً لضمان صحة عمليات الاستدلال التي يمكن أنْ يتمَّ التوصل إليها من خلال علامات الأفراد على الاختبار، والاختبار الذي يفضل مجموعة عن الأخرى يوصف بأنه متّحيز، ومُضيّع للهدف الذي وضع من أجله. (McCallion and Shumacker, 2002)

ويشير أوسترلند (Osterlind, 1983) إلى أنَّ التحيز كخطأ منتظم في القياس سيكون تأثيره في صدق الاختبار أكثر من تأثيره في ثباته. ويرى جنسون (Jensen, 1981) أنَّ الاختبار يكون متحيِّزاً عندما يختلف معنى العالمة الواحدة في المجموعات الفرعية المختلفة، أمّا لورد (Lord, 1980) فيرى أنَّ الاختبار يكون متحيِّزاً عندما يختلف ما يقيسه باختلاف المجموعات الفرعية المتكافئة على السمة التي صُممَ الاختبار لقياسها.

الاختبارات التحصيلية

• مفهوم الاختبارات التحصيلية:

هي الاختبارات المستخدمة في المدارس لقياس تعلم المتعلمين وتقويمهم، وهي على الأغلب اختبارات من إعداد المعلمين أنفسهم، وهي منتشرة وشائعة، وقد أرجع كلُّ من جرين وستيجر (Green&stager, 1996) استخدامها وشيوعها لعدة أسباب أهمها: أنَّ المعلم يستطيع من خلالها قياس أنواع كثيرة من أنواع التعلم باختلاف مستوياته، وقدرتها على قياس مدى واسع من المحتوى، وسهولة إدارتها من حيث التطبيق والتصحيح والتحليل بالنسبة لغيرها من الاختبارات، وتوجه المعلمين وتركيزهم على قياس تحصيل المتعلمين وتقويمهم، حيث ركزت على مستوى تحصيل المتعلمين للمعلومات والمهارات من المحتوى المقرر، لذا فقد تميزت بأنها اختباراتٌ تحصيلية، تميزها عن غيرها من الاختبارات.

يشير علام (2011: ٢٠١٤) إلى "أنَّ الاختبارات التحصيلية تقيس مدى استيعاب الطلاب لبعض المعرف والمفاهيم والمهارات المتعلقة بالمادة الدراسية، كما يدلُّ التحصيل الدراسي على الوضع الراهن لأداء الفرد أو تعلمُه أو ما اكتسبه بالفعل من برنامج تعليمي".

وتهدف الاختبارات التحصيلية إلى قياس مدى تحقق الأهداف التعليمية، وهي تُعدُّ بواسطة المعلمين في نهاية كل فصل دراسي، وفي نهاية كل مرحلة دراسية.

دور الأهداف التربوية في عملية التقويم

على الرغم من أنَّ من يقومون ببناء الاختبارات يدعون أنهم يقيسون الأهداف نفسها أو المحتوى نفسه؛ فإنَّ هناك اختلافاتٍ بينهم في نوعية الأسئلة، وهذا قد يؤدي لظهور فجوة بين الهدف والأسئلة.

ومن هنا، أصبح التطور في مجال التعليم يأخذ أشكالاً مختلفة في تقويم نتائج التعلم، وعلى الرغم من وجود أساليب جديدة لتقويم مستوى الطلبة، لا يزال الاعتماد على نتائج الاختبارات التحصيلية في تقويم تحصيل الطلاب ومدى إتقانهم للمهارات اللازمية للانتقال للمرحلة الدراسية التي تليها. لذلك؛ فإنَّ تصميم اختبار تحصيليٌّ جيد هو ارتباطٌ مفردةٌ الاختبار بأهداف التدريس المراد قياسها، أو مخرجات التعلم المقصودة.

• تعريف الأهداف:

- يُعرف ماجر (Major) الهدف بأنه: عبارة عن مقصود يُصاغ في عبارة تصيفٌ تغييرًا مقتراحًا يراد إحداثه في سلوك المتعلم نتيجة لعملية التعلم. (الظاهروآخرون، ٢٠٠٢: ٧١)
- وعرَّف ميجر (Mejor) الأهداف بأنها: عبارة عن مجموعة من الكلمات والرموز تصنف واحدًا من مقاصدك التربوية. (عریفج وآخرون، ١٩٩٩: ١٠١)

• **تصنيف الأهداف المجال المعرفي لبلوم:**

"وتصاغ الأهداف المراد قياسها في صورة أنماط سلوكيّة يمكن تقويمها في سلوك المتعلم، ومن أهداف هذه الأهداف والتي حاول قياسها في الاختبارات، تلك الأهداف التي وضعها بلوم في المجال المعرفي". (حسن وكمال زيتون، ١٩٩٥: ٦٢ - ٨٣)

أ- المعرفة (التذكر): Knowledge

ويقصد بها: قدرة الطالب على تذكر المعلومات (المعارف) سواء بالتعرف عليها أو استدعائهما بصورتها نفسها، أو بشكل مُقارب لها، كما قدمت له أثناء عملية التعلم.

وتشمل المعرفة عدداً من المستويات الفرعية منها:

▪ معرفة المصطلحات والحقائق المرتبطة بالمادة الدراسية.

▪ معرفة طرق التعامل مع الخصوصيات من المصطلحات والرموز والتصنيفات.

ب- الاستيعاب (الفهم): Comprehension

يعتبر من أدنى درجات الفهم، هو من أكثر مستويات القدرات الذكائيّة شيوعاً في أهداف المدرسين، ويقصد به القدرة على إدراك المعاني، ويظهر ذلك بترجمة الأفكار من صورة لأخرى وشرحها وتفسيرها، أو التبيّن من خلالها بنتائج معينة بناءً على المسارات والاتجاهات المضمنة في هذه الأفكار، ويشمل عدداً من المستويات الفرعية:

▪ **الترجمة:** ويطلب من الطالب أن يعيد صياغة محتوى معين بلغة أبسط، من خلال استخدام الكلمات البسيطة، أو استخدام الأشكال والرسوم التوضيحية.

▪ **التفسير:** وهو قدرة الطالب على إدراك العلاقات الواردة في المعلومات المضمنة في مادة تعليمية، وما تتضمنه من معانٍ.

▪ **التقدير الاستقرائي:** قدرة الطالب على أن يتجاوز حدود المعلومات المعطاة في المشكلة، ويحاول تقديم مجموعة من الملاحظات، مبنية على استقراء المعلومات المعاني الجزئية، ثم يستخدم هذه الاستنتاجات في الوصول إلى توقعاتٍ محتملة.

ت- التطبيق: Application

قدرة الطالب على استخدام معلوماتٍ سابقة في حل مشكلة أو التعامل مع موقف جديد، وتكون هذه المعلومات في صورة إجراءاتٍ أو قواعد أو نظرياتٍ سبق تعلُّمها.

ث- التحليل: Analysis

قدرة الطالب على الفحص المدقق لمادة تعليمية ما وتجزئتها إلى عناصرها، والأخطاء المنطقية التي أدّت إلى استنتاجاتٍ خاطئة، أو تجزئية مشكلة إلى مركباتها وعناصرها للتمييز بين المعلومات ذات العلاقة بها.

ج- التركيب: Synthesis

يُتوقع أن يكون الطالب لديه القدرة على إنتاج فريد ومميز؛ بمعنى أن يقدم عملاً إبداعياً، وهذه القدرة العقلية تتضمن إنتاجاً فكريّاً ابتكارياً.

ح- التقويم: Evaluation

قدرة الطالب على إصدار أحكام على موضوع محل تقييم وتبير هذا الحكم، أو على حل مشكلة معينة، أو يثمن نواتج أو أفكاراً، باستخدام محكّاتٍ داخلية أو خارجية. (عوْدَة، ٢٠١٠، ٩١-٩٥). وفيما يتعلق بصدق هرميّة التصنيف، فقد أشار بلوم إلى أنَّ التصنيف يقوم على مبدأ من مبادئ التعلم، وهو الانتقالُ من السُّهُل على الصُّعب، ومن المحسوس إلى المجرد؛ وبالتالي فإنَّ الطالب لا يستطيع أنْ يستوعبه إلا إذا عرفه مسبقاً.

وقد اقترح مارشال وهاليس (Marshall & Hales, 1971, p.26) في (عوْدَة، ٢٠١٠، ١١٣:٢٠١٠) الاكتفاء بمستوياتٍ ثلاثة هي: المعرفة والاستيعاب والتطبيق، والمقصود بالتطبيق هنا: كُلُّ ما بعد الاستيعاب.

تحليل المحتوى: Content Analysis

يُعدُّ من أهمِّ الأساليب التي يتمُّ فيها الحكم على محتوى أيِّ مادة تعليميَّة، أو للحكم على مدى جودة منهج دراسيٍّ، ومدى شموليته.

كما يوضح تكامله رأسياً على مستوى مراحل التعليم ، وأفقياً على مستوى المناهج الأخرى للصف الدراسي نفسه، ومدى كفايته لتحقيق الأهداف المرجوة.

ويُستخدم تحليل المحتوى للحكم على هذا المحتوى من خلال معيار واحد أو عدة معايير، وقد أضاف تحليل المحتوى أساليب مُفصلة لمعالجة المواد المكتوبة بطريقة كميَّة. وُسُمِّيَ الأشكال الفرعية للمحتوى بوحدات تحليل المحتوى. (منصور، ٢٠٠٣:٢٠١)

خطوات تحليل المحتوى:

كما أوردها (منصور، ٢٠٠٣:١٢٥ - ١٣٠):

١- تحديد المنهج المراد تحليله.

يتمُّ فيه تحديد المادة العلميَّة التي سيتمُّ تحليلها، فمن الممكن أنْ تكون مجموعة من الأسئلة الاختباريَّة، أو قد تكون مجموعة من الكتب.

٢- تحديد فئات التحليل ووحداته.

ويتمُّ فيها تحديد العناصر الرئيسة التي تكون المادة العلميَّة التي سيتمُّ تحليلها، وتحتَلُّ فئاتُ التحليل وفقاً لطبيعة المادة العلميَّة، ومفاهيمها الخاصة بها.

وهناك بعض النماذج من فئات التحليل مثل (فئات التحليل العامة):

حيث يستندُ تحليل المحتوى إلى محكّاتٍ خارجية، تتمثلُ في محكٌ الأهداف التربويَّة الذي يعتبر من أكثرها شيوعاً في المجال التربويٍّ، ومن أهمِّ الأهداف التربويَّة التي تمثلها الأهداف المعرفية، حيث يعتمدُ أغلب الباحثين على استخدام هذه الأهداف كأداة لتحليل الكتب المدرسية والاختبارات التحصيلية؛ وذلك للكشف عن تواجدها في الكتب (المقررات) والاختبارات التحصيلية، مثل التحليل بالنسبة لتصنيف بلوم لمجالات التعليم.

ويتضمن التحليل ما يلي:

• **المجال المعرفي:** ويتضمن الحقائق، والمفاهيم، والتعليمات والقواعد، والنظريات، وحل المشكلات.

• **المجال الانفعالي:** ويتضمن الاتجاهات، والميول، والقيم، والعادات.

• **المجال النفس حركي:** ويتضمن المهارات العقلية مثل: الاستنتاج والاستقراء، كما يتضمن المهارات الحركية مثل: القراءة، والكتابة، ومهارة القياس، وإجراء التجارب.

٣- التأكيد من صدق التحليل.

يمثل صدق أداة تحليل المحتوى إذا عكست ما هو موجود بالفعل داخل السياق المراد تحليله. كما لابد أن يتوافر الثبات باعتباره يهدف لتحويل المحتوى النظري إلى محتوى كمّي قابل للقياس، بمعنى أن يحصل الباحث على النتائج نفسها عند اختلاف محلل، أو تقارب الزمن الذي تم فيه التحليل بين محلل وآخر، أو محلل مكافئ له في المستوى التعليمي نفسه.

٤- التأكيد من ثبات التحليل.

يقوم تحليل المحتوى على نوعين من الاتفاق: Consistency

• **الاتفاق بين المحللين:** Consistency among analysis

وذلك عند تحليل مادة دراسية أو جزء منها سواء وحدة أو موضوع دراسي؛ بهدف بناء اختبار تحصيلي، لذلك لابد أن يقدم المحللون عند استخدامهم لأداة التحليل نفسها، نفس نتائج التحليل للمحتوى المتفق عليه، ويتم حساب معامل الثبات حسب معادلة (Holsti, 1969)

$$\text{نسبة الاتفاق} = \frac{\text{عدد مرات الاتفاق}}{\text{عدد مرات الاتفاق} + \text{الاختلاف}} \times 100$$

• **الاتفاق خلال فترات زمنية:** through limited periods of time Consistency

لابد أن يقدم المحللون عند استخدامهم لأداة التحليل نفسها، نفس نتائج التحليل للمحتوى المتفق عليه، أو محلل واحد، ولكن في فترات زمنية مختلفة.

وكانت عنابة مجال التربية وعلم النفس خاصةً في بناء الاختبارات التحصيلية، ونجد أكثر مواد وحدات تحليل المحتوى استخداماً لهذا الأسلوب هي الكتب المدرسية والاختبارات التحصيلية، حيث يقوم المعلم بوصف الإجراءات التي يستخدمها المعلم في وضع الأسئلة في ضوء الأهداف ومحفوظ الكتاب المقرر. (منصور، ٢٠٠٣: ١٣٠)

الدراسات السابقة

نظراً لشيوخ استخدام الاختبارات في المؤسسات التربوية وغير التربوية، وأهمية القرارات التي تُبني على نتائجها؛ كان من الطبيعي أن تُجرى عدة دراساتٍ جادة للكشف عن مدى سلامتها إعداد الاختبارات وصدق نتائجها، وسلامة القرارات التي تُبني عليها، وقد حاولت الدراسات السابقة التعرف على بعض أنماط التحيز التي قد تنتشر في الاختبارات العامة، واختبارات المعلمين، وإصدار بعض التوصيات للتقليل منها، ويمكن تصنيف هذه البحوث ضمن فئتين: البحوث التي تناولت التحيز لمستويات المجال المعرفي، والبحوث التي تناولت التحيز لمحتوى الاختبارات.

الدراسات التي تناولت النحیز لمستويات المجال المعرفي

أظهرت بحوث كثيرة في الأدب التربوي عنية متزايدة بتحييز اختبارات المعلمين لمستويات الأهداف المعرفية التي تقيسها، كما بحثت في العوامل التي تسهم في رفع مستويات الأهداف التي تقيسها الاختبارات، منها: تدريب المعلمين على وضع فقرات ذات مستويات عقلية عليا، والتقييد بقواعد بناء الاختبارات.

ففي دراسة أجراها علوة وحبيب (٢٠١٧م): هدفت الدراسة إلى تحليل وتقويم أسئلة الاختبارات النهائية التي أعدّها أعضاء هيئة التدريس في أقسام كلية التربية للبنات - جامعة الكوفة للعام الدراسي (٢٠١٤ - ٢٠١٥م)، لمعرفة درجة شمولية الأسئلة على مستويات الأهداف المعرفية وفقاً لتصنيف بلوم، حيث تألف مجتمع الدراسة من أسئلة الاختبارات النهائية لجميع المواد الدراسية، ولأربع مراحل في أقسام كلية التربية والبالغ عددها (٣٠١) مادة دراسية، وتكوّنت العينة من أسئلة (٨٠) مادة دراسية موزعة على الأقسام المختلفة، تم اختيارها بطريقة عشوائية طبقية، وأوضحت النتائج أنَّ أسئلة الاختبارات شملت جميع مستويات الأهداف المعرفية، وركّزت بشكل ملحوظ على المستويات المتقدمة للأهداف المعرفية، حيث جاء بالمرتبة الأولى مستوى الاستيعاب بنسبة ٣٠٪، يليه مستوى التذكر ٢٧٪، والتطبيق ٢٠٪، أمّا بالنسبة للمستويات العليا للأهداف المعرفية فقد مثل مستوى التحليل ٢٣٪، والتركيب ١٥٪، والتقويم ٢٪ من الأسئلة، وأوصى الباحثان بضرورة عقد دوراتٍ تدريبية في صياغة الأسئلة وبنائها.

دراسة جيوسي (٢٠١٦م): هدفت الدراسة للتحقق من مدى تحقيق أسئلة الامتحانات النهائية التي يضعها الأساتذة في جامعة فلسطين التقنية للأهداف المعرفية تبعاً لهرم بلوم، حيث حلَّ الباحثُ مجموعة من الامتحانات النهائية للفصل الدراسي الثاني (٢٠١٥ - ٢٠١٦م)، والبالغ عددها (١٢) امتحاناً، شملت (٣٢٢) سؤالاً، وركّزت الأسئلة على المجال المعرفي في أدنى مستوياته، واحتلَّ مستوى المعرفة الترتيب الأول بنسبة (٧٣.٨٪) أي: ما يعادل (٢٤٩) سؤالاً، ثم يليه مستوى الاستيعاب (٧٣) سؤالاً بنسبة (٢٣٪)، بينما أتى مستوى التطبيق والتحليل في الأخير بنسبة (١٦٪) و(٥٪) أسئلة. وعليه: أوصى الباحثُ بضرورة إنشاء مركز لتقويم الامتحانات وتطويرها.

دراسة سلمان (٢٠١٢م): عن مدى تمثيل أسئلة الاختبارات التحصيلية النهائية لمستويات الأهداف المعرفية لدى مُعلّمي العلوم الحياتية في إقليم جنوب الأردن، وهدفت الدراسة إلى الكشف عن مدى تمثيل أسئلة الامتحانات النهائية التي أعدّها معلّمو علم الأحياء في المرحلة الثانوية الأولى في المنطقة الجنوبية من الأردن لمستويات المعرفة لتصنيف بلوم في المجال المعرفي. وتكوّنت عينة الدراسة من (٦٥) معلّماً أعدوا (١٨٤٢) سؤالاً في اختبارات نهاية الفصل الدراسي ٢٠١٢ - ٢٠١٣م، وتمَّ تحليل الأسئلة باستخدام المنهج الوصفي التحليلي، وتحليل المحتوى، وتصنيف العناصر إلى مستويات مختلفة وفق تصنيف بلوم. وقد توصلت الدراسة لمجموعة من النتائج وهي:
▪ أنَّ أسئلة المعلمين وفق مستوى المعرفة والفهم، حيث بلغت (١٦٨٥) فقرة أيٌ: بنسبة (٩١.٤٨٪) من إجمالي الأسئلة.

▪ وجاءت الأسئلة في مستويات التطبيق والتحليل في الرتبة الثانية بمعدل (١٥٧) فقرة، تمثل (٨.٥٢٪)، في حين لم ترد أيٌ أسئلة على مستويات التقويم والتركيب.

الدراسات التي تناولت التحيز في محتوى الاختبارات:

لقد أُجريت عدة بحوث للكشف عن التحيز في محتوى فقرات الاختبارات، وأظهرت الدراسات صورتين من تحيز المحتوى وهما: التحيز في الاختبارات ذات المحتوى المحدد؛ كالتحيز في محتوى الاختبارات المدرسية، والتحيز في الاختبارات ذات المحتوى غير المحدد؛ كالتحيز في محتوى اختبارات الذكاء.

دراسة القواس (٢٠١٧م): هدفت إلى تقويم اختبارات القبول لمادة الرياضيات بجامعة إب، في ضوء معايير الاختبار الجيد، وتكوين العينة من (٢٤١) سؤالاً من أسئلة اختبارات القبول لمادة الرياضيات للأعوام (٢٠١٥:٢٠١٠م)، حيث استخدم الباحث المنهج الوصفي التحليلي، وأعدَّ استماره لتحليل أسئلة الاختبارات في ضوئها، وتوصَّل الباحث للنتائج التالية: أنَّ الاختبارات لا تراعي معيار تنوع الأسئلة، بالإضافة إلى عدم مراعاتها لمعيار شمول الأسئلة للمحتوى الرياضي؛ وضعف شمولها لموضوعات مقرر الرياضيات للصف الثالث الثانوي. وأوصى الباحث بعقد برنامج تدريسي لتنمية كفايات القياس والتقويم لأعضاء هيئة التدريس القائمين على إعدادها.

دراسة الناقة (٢٠١٦م): هدفت إلى تقويم أسئلة اختبارات الكيمياء للصف الثاني عشر بفلسطين، خلال السنوات (٢٠٠٧/٢٠١٥م)، شملت عينة الدراسة أسئلة اختبارات الكيمياء النهائية بمجموع (٩٠) سؤالاً، لكلٌّ سنة (١٠) أسئلة. ولتحقيق أهداف البحث؛ تمَّ تصميم أداة لتقويم الأسئلة في ضوء معيار الشمولية للمقرر الدراسي، وتمَّ حساب التكرارات للمعيار لكلٌّ سنة، وأظهرت النتائج أنَّ معيار الشمولية ككلٌّ في السنوات التسع كانت بنسبة (٧٢,٨٥٪) وهي نسبة مناسبة، وأوصى الباحثُ بالعناية بالمطابقة والشمولية عند وضع الأسئلة، مع مراعاة التوازن بين موضوعات المقرر الدراسي.

دراسة محمد (٢٠١٢م): بحث لتقويم أسئلة الرياضيات للصف الثالث المتوسط للامتحانات العامة في العراق للأعوام (٢٠١١:٢٠٠٥)، حيث بلغت العينة (٤٢) سؤالاً رئيساً بواقع (٦) أسئلة للدور الامتحاني من كلٌّ عام، أي: بواقع (١٢٠) سؤالاً فرعياً. ولتحقيق ذلك؛ اعتمد الباحث أسلوب تحليل الأسئلة، وذلك بإعداد القوائم الأولية المستمدة على معايير تحليل أسئلة الامتحانات المستند إليها، ودللت النتائج أنَّ أسئلة الامتحانات كانت شاملة لجميع الفصول المقررة وبنسب متفاوتة في عدد الأسئلة.

التعليق على البحوث السابقة

يُوضح من الدراسات السابقة: التباين في النتائج المتعلقة في كلٌّ من نسب التحيز في المستويات المعرفية التابعة لتصنيف بلوم، ونسبة التحيز في المحتوى الدراسي للكتاب المدرسي، حيث إنَّ جميع البحوث السابقة أكدَت أنَّ جميع الاختبارات التحصيلية ركزَت على مستويات التفكير العقليَّة الدنيا (تذكرة - استيعاب - تطبيق) مع قلة توافر الأسئلة ذات مستويات عقلية عليا (التحليل، والتركيب). ويتفق البحث الحالي مع الدراسات السابقة من حيث الهدف وهو تحليل أسئلة الاختبارات؛ للكشف عن أنماط التحيز بالنسبة للأهداف المعرفية والمحتوى، ويختلف في المادة الدراسية، والمرحلة، حيث طبقَ على مادة الرياضيات لمرحلة الأول متوسط، وبحدود معرفة

الباحثة لم تجر دراسة في مبحث الرياضيات في مدينة تبوك تكشف عن أنماط التحيز في الاختبارات التي أعدتها معلمات الرياضيات للمرحلة المتوسطة في المدارس التابعة لوزارة التعليم في المملكة العربية السعودية، وعليه؛ يمكن أن يكون لهذا البحث دور في لفت انتباه المعلمات والباحثين لمستويات الأسئلة التي تعدّها المعلمات وبنائها، ومدى تغطية الاختبارات للمحتوى المقرر، ومراعاتها لمستويات التفكير لدى الطالبات.

أسئلة البحث

ويسعى البحث إلى الإجابة عن التساؤلات الآتية:

- ١ - ما مدى تحيز الاختبارات التحصيلية التي أعدتها معلمات الرياضيات للصف الأول متوسط في مدينة تبوك لمستويات الأهداف المعرفية (المعرفة - الاستيعاب - التطبيق فما فوق)؟
- ٢ - ما مدى تحيز الاختبارات التحصيلية التي أعدتها معلمات الرياضيات للصف الأول متوسط في مدينة تبوك لمحظى وفصول الكتاب المقررة (الفصل الأول، الفصل الثاني، الفصل الثالث، الفصل الرابع)؟

منهجية البحث وإجراءاته

• منهجية البحث:

المنهج الوصفي التحليلي، وسيتم تحليل الاختبارات إلى فقراتٍ حسب تصنيف بلوم، وتحليل الفقرات التي تتسمى لكلٍّ فصلٍ من فصول الكتاب المقرر.

• مجتمع البحث:

سيكون مجتمع البحث من فقرات الاختبارات النهائية لمادة الرياضيات لمرحلة الأول متوسط، في كلٍّ من المدارس الحكومية للبنات في مدينة تبوك، والبالغ عددها (٥٨) مدرسة.

• عينة البحث:

سيتم اختيار عينة عشوائية بسيطة من اختبارات معلمات الرياضيات النهائية لمرحلة الأول متوسط للفصل الدراسي الثاني لعام (١٤٣٧ - ١٤٣٨هـ) التي أعدتها (٣٦) معلمة رياضيات، ويبلغ عدد فقرات الاختبارات (١٢٤٤) فقرة اختبارية.

• أدوات البحث:

قامت الباحثة بإعداد دليل لمستويات الأسئلة المعرفية، والاستجابات التي يتطلبها كلٌّ مستوى من المستويات المعرفية الستة وفق تصنيف بلوم للمجال المعرفي؛ وذلك للاعتماد عليه في تحليل وتصنيف أسئلة الاختبارات التي تعدّها معلمات الرياضيات للصف الأول متوسط في مدينة تبوك، وتم الاعتماد في إعداد الدليل على المراجع والكتب المتعلقة بهذا الموضوع مثل (ملحم، ٢٠١٥؛ عودة، ٢٠١٠؛ يوسف والرفاعي، ٢٠٠٥)، والدراسات السابقة التي تمت الاستعانة بها في هذا البحث، وتم عرضها على مجموعة من معلمات الرياضيات اللاتي أبدين بعض الملاحظات، وتم تعديلها.

- تم إجراء استطلاع للرأي في الميدان التربوي لعدد الأسئلة التي تمثل كلًّا مستوى من مستويات الأهداف المعرفية، وكذلك عدد الأسئلة التي تمثل كلًّا فصل من فصول كتاب الرياضيات للأول متوسط.
- تم إعداد أداة التحليل للأهداف المعرفية والمحتوى تألفت من استمارتين؛ تضمنت الاستماراة الأولى تصنيفاً لفقرات الاختبارات التي أعدتها معلمات الرياضيات حسب مستويات الأهداف المعرفية وفق تصنيف بلوم، وتضمنت الاستماراة الثانية تصنيفاً لفقرات الاختبارات التي أعدتها المعلمات حسب الفصول المقررة من كتاب الرياضيات للصف الأول متوسط.

إجراءات إعداد الأداة:

من البحث بعد إجراءاتٍ تمثل فيما يلي:

- إعداد دليل تحليل الأهداف حسب تصنيف بلوم للأهداف المعرفية، وتم التأكد من صحة الأمثلة المستخدمة لكل هدف معرفيٌّ، عن طريق عرضها على مشرفاتٍ ومعلماتٍ رياضياتٍ.
- زيارة إدارة التعليم في مدينة تبوك للحصول على بعض الأسئلة، وكذلك المدارس المتوسطة للبنات في مدينة تبوك.
- اختيار عينة البحث من المجتمع بطريقة عشوائية بسيطة.
- اختيار وحدة التحليل المناسبة لهذا البحث، وهي الفقرة المكونة للاختبارات التحصيلية، أمّا فئات التحليل فهي مستويات الأهداف المعرفية بلوم، وفصول الكتاب المقرر.
- قامت الباحثة بتحليل أسئلة الاختبارات التحصيلية، وفق ما يلي:
 - رصد عدد الفقرات في كلٌّ ورقة اختبارية، وكل سؤال يحتوي على تفرعاتٍ مثل: (أ، ب، ج، ...)، أو (١، ٢، ٣، ...). تعامل كل فقرة على حدة، وتصنف تبعاً للمستوى المعرفي، والفصل الذي ينتمي إليه في الكتاب المقرر.
 - تصنيف الفقرات حسب اختبار المعلمة التي أعدَّته، بحيث يتم تصنيف كلٌّ فقرة وما تقيسه من مستوى معرفيٌّ، وإذا وردَت أكثر من فكرة في الفقرة إحداها ضمن مستوى والأخر ضمن مستوى آخر؛ فإنَّ السؤال يوضع بالمستوى الأعلى، وترصد البيانات المتعلقة بالمستوى المعرفي بالاستماراة بحسب نوع الفقرة.
 - تصنيف كلٌّ فقرة وما تقيسه من محتوى الكتاب المقرر، بحيث يتم تصنيف كلٌّ فقرة على حسب الفصل الذي يمثله من فصول الكتاب المقرر، وترصد البيانات المتعلقة بفصول الكتاب بالاستماراة بحسب نوع الفقرة.
 - حساب مجموع التكرارات لكلٌّ فقرة من فقرات كلٌّ استمارة، وتحويلها إلى نسب مئوية ورصدها في جداول تسهل عملية عرض النتائج وتفسيرها.
 - إيجاد أعداد الفقرات المتوقعة والملاحظة من كلٌّ مستوى من مستويات الأهداف المعرفية.
 - إيجاد أعداد الفقرات المتوقعة والملاحظة حسب فصول الكتاب المقرر.
 - مقارنة متوازنات الفقرات المتوقعة بالملاحظة من كلٌّ مستوى من مستويات الأهداف المعرفية.
 - مقارنة متوازنات عدد الفقرات المتوقعة بمتوسطات أعداد الفقرات الملاحظة لكلٌّ فصل من فصول كتاب الرياضيات للصف الأول متوسط.

ثبات الأداة

لقياس ثبات التحليل طرق مختلفة، وأهمها وأكثرها شيوعاً في دراسات تحليل المحتوى، طريقة نسبة الاتفاق؛ أي إجراء التحليل مرتين على المادة نفسها، وتحديد درجة الاتفاق بين التحليلين لتعتبر مؤشراً لمعامل الثبات، وقد قامت الباحثة للتأكد من ثبات التحليل بما يلي:

الثبات عبر المحللين

حيث يتم تحليل المادة نفسها من قبل محللين معتمدين على القاعدة نفسها، بحيث يقوم بعملية التحليل كلّ منهم على حدة.

واختارات الباحثة عينة من أسئلة الأول متوسط للفصل الدراسي الثاني، التي أعدّتها معلمات الرياضيات لعام ١٤٣٧ - ١٤٣٨هـ) وكان عدد الفقرات (٣١٧) فقرة، وتم تحليل كلّ فقرة من قبل الباحثة ومن قبل محللين آخرين بلغ عددهم (٤) محللين من حملة الماجستير في الرياضيات والمناهج وطرق التدريس، وتم حساب عدد مرات الاتفاق وعدد مرات الاختلاف مع كلّ محلل من المحللين، بالنسبة لكلّ فقرة من حيث تصنيفها لمستويات بلوم للأهداف المعرفية.

وتم حساب معامل الثبات حسب معادلة (Holsti, 1969)

$$\text{نسبة الاتفاق} = \frac{\text{عدد مرات الاتفاق}}{\text{عدد مرات الاتفاق} + \text{الاختلاف}} \times 100$$

وقد بلغ متوسط نسبة الاتفاق بين تصنيف الباحثة والمحللين ٠٨٩، ويتبين من نتيجة حساب معامل الثبات أنَّ نسبة الاتفاق مرتفعة؛ وهو مؤشر على ثبات عملية تصنيف الأسئلة.

الأساليب الإحصائية:

استُخدم في هذا البحث الأساليب الإحصائية التالية:

- معادلة (Holsti) لحساب ثبات التحليل.
- التكرارات.
- النسب المئوية.
- اختبار Chi-square (χ^2) للاستقلالية.
- اختبار T-TEST.

نتائج البحث

للإجابة عن السؤال الأول: ما مدى تحيز الاختبارات التحصيلية التي أعدّتها معلمات الرياضيات للصف الأول متوسط في مدينة تبوك لمستويات الأهداف المعرفية (المعرفة - الاستيعاب - التطبيق وما فوق)؟

للإجابة عن هذا السؤال: تمّ تصنيف كلّ نموذج اختباري إلى مستويات الأهداف المعرفية كما قسمها بلوم، حسب المستويات (المعرفة - الاستيعاب - التطبيق وما فوق)، كما تمّ إيجاد عدد الفقرات الملاحظة وعدد الفقرات المتوقعة التي تقيس كلّ مستوى من مستويات المجال المعرفي (حسب دليل تصنيف بلوم)، ملحق رقم (١)، وكذلك حسب رأي الميدان التربوي.

وبمعرفة عدد الفقرات المتوقعة التي تقيس كلًّا مستوى من مستويات المجال المعرفي، حُسبت قيم χ^2 (Chi-square) التي تكشف التوافق بين التكرار المتوقع، والتكرار الملاحظ للفقرات باستخدام البرنامج الإحصائي SPSS، كما في الجدول (١).

جدول (١). توزيع أعداد الفقرات المتوقعة والملاحظة من اختبارات معلمات الرياضيات على مستويات الأهداف المعرفية لبلوم، وقيم χ^2 لها.

رقم المعلمة	مستوى المعرفة المتوقعة	الملاحظة المتوقعة	مستوى الاستيعاب الملاحظة المتوقعة	مستوى التطبيق بما فوق الملاحظة المتوقعة	عدد الفقرات	χ^2
١	٩,٨٩	١٥	١٢,٤٧	١٨	٤٣	٦,١٠٣
٢	٥,٩٨	٩	٧,٥٤	١١	٢٦	٣,٠٧٤
٣	٧,٨٢	١٣	٩,٨٦	١٥	٣٤	٦,٧٣٦
٤	١٢,٥٦	٢٣	١٥,٩٥	١٩	٥٥	٨,٧٣٦٨
٥	٨,٠٥	١١	١٠,١٥	١٤	٣٥	٢,٩٩٥
٦	١٠,٥٨	١٧	١٣,٣٤	١٨	٤٦	٥,٧٥٩
٧	١٠,١٢	٢٣	١٢,٧٩	١٠	٤٤	٨,٦٣٨
٨	٧,٨٢	١٥	٩,٨٦	١٠	٣٤	٤,٠٩٠
٩	١٣,٨	٢٢	١٧,٤	٢٤	٦٠	٨,٢٠٥
١٠	١٠,٥٨	١٩	١٣,٣٤	١٧	٤٦	٧,١٦٧
١١	٧,٩	١٢	٨,٧	١١	٣٠	٣,٨٤٩
١٢	١٢,٦٥	٢٦	١٥,٩٥	٢٠	٥٥	١٣,٠٣٥
١٣	٦,٢١	١١	٧,٨٣	١٠	٢٧	٤,٢٧٢
١٤	٦,٦٧	١٤	٨,٤١	١١	٢٩	٨,٣٦٣
١٥	٩,٢	١٥	١١,٦	١٣	٤٠	٣,١٢١
١٦	٨,٠٥	١١	١٠,١٥	١٨	٣٥	٨,٠٢٠
١٧	٨,٢٨	١٧	١٠,٤٤	١١	٣٦	٦,١٥١
١٨	٦,٩	٧	٨,٧	١٣	٣٠	١,٣٩٤
١٩	٦,٦٧	١٠	٨,٤١	٨	٢٩	٠,٨٨٩
٢٠	١٣,٨	٢٠	١٧,٤	٢٥	٦٠	٧,٠٣٧
٢١	١١,٥	٢٢	١٤,٥	١٥	٥٠	٦,٢٠٢
٢٢	١١,٩٦	١٧	١٥,٠٨	٢٣	٥٢	٧,١١٤
٢٣	٨,٢٨	١٢	١٠,٤٤	١٣	٣٦	٢,٤٦٣
٢٤	٨,٠٥	١٣	١٠,١٥	١٥	٣٥	٦,٦٨٨
٢٥	١٣,٨	٣٥	١٧,٤	١٤	٦٠	١٧,٣٩٠
٢٦	٦,٤٤	٨	٨,١٢	١٢	٢٨	٢,٢٥٩
٢٧	٨,٩٧	١٤	١١,٣١	١٣	٣٩	٢,٨٣٤
٢٨	٩,٢	١١	١١,٦	١٥	٤٠	١,٢٩١
٢٩	١١,٥	٢١	١٤,٥	١٧	٥٠	٦,٥٧٠
٣٠	١٣,٨	٢٥	١٧,٤	٢٠	٦٠	٧,٨٠٠

نسبة المعلمات الالاتي لم تتفق نسب توزيع فقرات اختباراتهن مع النسب التي تم افتراضها الميدان التربوي

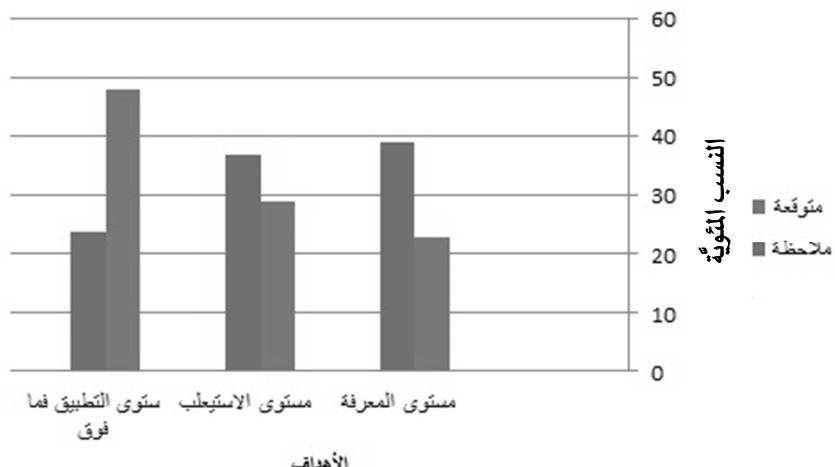
❖ ذات دلالة إحصائية عند ($p \leq 0,05$).

يظهر من الجدول (١) أنَّ ٥٦٪ من قيم χ^2 ذات دلالة إحصائية عند ($\alpha \leq 0.05$)، وهذه النسبة تمثل المعلمات الالاتي لم تتفق توزيع فقرات اختباراتهن مع التوزيع المتوقع لمستويات الأهداف المعرفية لبلوم، وقد بلغ عدد الفقرات التي تقيس مستوى المعرفة (٤٨٨) فقرة وكانت في المرتبة الأولى، وجاءت في المرتبة التي تليها الفقرات التي تقيس مستوى الاستيعاب بـ (٤٥٣) فقرة، حيث كان عدد الفقرات التي تقيس مستوى التطبيق بما فوق (٣٠٣) فقرة. ويبيّن الجدول (٢) مدى توزيع أعداد الفقرات المتوقعة والملاحظة على مستويات الأهداف المعرفية لبلوم، من خلال النسب المئوية لعدد الفقرات المتوقعة والملاحظة من اختبارات معلمات الرياضيات.

جدول (٢). "النسب المئوية لعدد الفقرات المتوقعة والملاحظة من اختبارات معلمات الرياضيات حسب كلَّ مستوى من مستويات الأهداف المعرفية لبلوم".

نوع الفقرة	مستوى المعرفة	مستوى الاستيعاب	مستوى التطبيق بما فوق	مجموع الفقرات
متوقعة	٢٨٦	٣٦١	٥٩٧	١٢٤٤
	%٢٣	%٢٩	%٤٨	%١٠٠
ملاحظة	٤٨٨	٤٥٣	٣٠٣	١٢٤٤
	%٣٩	%٣٧	%٢٤	%١٠٠

يظهر من الجدول (٢) أنَّ مستوى المعرفة أخذ العدد الأكْبَر من فقرات الاختبارات بنسبة ٣٩٪، يليه مستوى الاستدلال ٣٧٪، فيما كان العدد الأقلَّ لمستوى التطبيق بما فوق بنسبة ٢٤٪. وبالنظر إلى التوزيع المتوقع للفقرات على مستويات الأهداف المعرفية، يظهر أنَّ نسبة الفقرات بالنسبة لمستوى المعرفة ٢٣٪، وفقرات مستوى الاستيعاب ٢٩٪، ثم مستوى التطبيق بما فوق بنسبة ٤٨٪. ويوضح الشكل (١) المقارنة بين النسب المئوية للفقرات المتوقعة والملاحظة على كلَّ مستوى من مستويات الأهداف المعرفية وفق تصنيف بلوم من خلال التمثيل بالأعمدة.



شكل (١). "التمثيل بالأعمدة لتوزيع النسب المئوية للفقرات المتوقعة والملاحظة على كل مستوى من مستويات الأهداف المعرفية وفق تصنيف بلوم".

يلاحظ من الشكل (١) : أنَّ نسبة الفقراتِ الملاحظة أعلى من نسبة الفقراتِ المتوقعة في مستوى المعرفة، ومستوى الاستيعاب، فيما كانت نسب الفقراتِ الملاحظة أقلَّ من نسبة الفقراتِ المتوقعة في مستوى التطبيق فما فوق. أمّا فيما يتعلق بالكشف عن إمكانية وجود فروق ذات دلالة إحصائية ($\alpha \leq 0.05$) بين متوسطاتِ الفقراتِ الملاحظة والفقراتِ المتوقعة من مستويات الأهداف المعرفية لبلوم: المعرفة، الاستيعاب، التطبيق فما فوق؛ فقد استعملت الباحثة اختبار "ت" للكشف عن الفروق بين المتوسطات، وبيّن الجدول (٣) مدى الفروق بين متوسطاتِ الفقراتِ الملاحظة والفقراتِ المتوقعة ($\alpha \leq 0.05$) من مستويات الأهداف المعرفية لبلوم.

جدول (٣). "مقارنة بين متوسطات الفقرات الملاحظة والفقرات المتوقعة من مستويات الأهداف المعرفية لبلوم".

تطبيق فما فوق	المستوى	نوع الفقرة	
		الاستيعاب	المعرفة
١٩.٩٠	١٢.٠٢	٩.٥٣	الوسط الحسابي
٥.٣١	٣.٢١	٢.٥٤	الانحراف المعياري
١٠.١٠	١٥.١٠	١٦.٢٦	الوسط الحسابي
٢.٩٢	٤.٣٩	٦.٣٠	الانحراف المعياري
٤٨.٨٤٧	٤٣.٠٩٢	٤٥.٤٢٢	"ت" المحسوبة

* ذات دلالة إحصائية عند ($\alpha \leq 0.05$).

يظهر من جدول (٣) : أنَّ هناك فروقاً ذات دلالة إحصائية عند ($\alpha \leq 0.05$) بين متوسطاتِ الفقراتِ الملاحظة والفقراتِ المتوقعة، ولجميع مستوياتِ الأهداف المعرفية لبلوم (المعرفة، الاستيعاب، التطبيق فما فوق)؛ وذلك لأنَّ قيم "ت" المحسوبة أكبر من "ت" الحرجة ($\alpha \leq 0.05$)، ودرجات الحرية = ٥٨ (٢٠٠٠). وأنَّ متوسط الفقراتِ الملاحظة أعلى من متوسط الفقراتِ المتوقعة لـ كلٌّ من مستوى المعرفة والاستدلال، أمّا بالنسبة لمستوى التطبيق فما فوق فإنَّ متوسطاتِ الفقراتِ المتوقعة أعلى من متوسط الفقراتِ الملاحظة. للإجابة عن السؤال الثاني: ما مدى تحيز الاختبارات التحصيلية التي أعدتها معلماتُ الرياضيات للصف الأول متوسط في مدينة تبوك لمحظى وفصول الكتاب المقرر (الفصل الأول، الفصل الثاني، الفصل الثالث، الفصل الرابع)؟

للإجابة عن هذا السؤال: تمَّ تصنيف كلَّ فقرة اختبارية إلى فصول الكتاب التي تمثلها كما قسمَها الكتاب المقرر (الفصل الأول، الفصل الثاني، الفصل الثالث، الفصل الرابع)، كما تمَّ إيجاد عدد الفقراتِ الملاحظة وعدد الفقراتِ المتوقعة التي تقيس كلَّ فصول الكتاب المقرر (الفصل الأول، الفصل الثاني، الفصل الثالث، الفصل الرابع) وحسب رأي الميدان التربوي، وبمعرفة عدد الفقراتِ المتوقعة التي تمثل كلَّ فصل من فصول الكتاب المقرر، حُسبَت قيم Chi-square كـ ٢٤ (χ^2) التي تكشف التوافق بين التكرار المتوقع والتكرار الملاحظ للفقراتِ، وقد حُسبت النتائج عن طريق البرنامج الإحصائي SPSS كما في الجدول (٤).

جدول (٤). توزيع الفقرات المتوقعة والملحوظة من اختبارات المعلمات حسب الفصول المقررة من كتاب الرياضيات للصف الأول متوسط الفصل الدراسي الثاني.

كـ	عدد الفقرات	الفصل الرابع		الفصل الثالث		الفصل الثاني		الفصل الأول		رقم المعلمة
		الملاحظة	المتوقعة	الملاحظة	المتوقعة	الملاحظة	المتوقعة	الملاحظة	المتوقعة	
❖٨,٣٠٠	٤٣	٥	١٢,٩	٢٤	١٢,٠٤	٩	٩,٨٩	٥	٨,١٧	١
٣,٥٤٨	٢٦	٧	٧,٨	١٠	٧,٢٨	٨	٥,٩٨	١	٤,٩٤	٢
١,٧٠٤	٣٤	٦	١٠,٢	١٣	٩,٥٢	١٠	٧,٨٢	٥	٦,٤٦	٣
❖٨,١٧٠	٥٥	٨	١٦,٥	٢٨	١٥,٤	١٣	١٢,٦٥	٦	١٠,٤٥	٤
١,٥٢٦	٣٥	٩	١٠,٥	١١	٩,٨	١١	٨,٠٥	٤	٦,٦٥	٥
٠,٣٥٧	٤٦	١٣	١٣,٨	١٤	١٢,٨٨	١٢	١٠,٥٨	٧	٨,٧٤	٦
❖٩,٠٥٤	٤٤	٩	١٣,٢	٢٠	١٢,٣٢	٨	١٠,١٢	٧	٨,٣٦	٧
❖٨,٤٨٩	٣٤	٢	١٠,٢	١٩	٩,٥٢	٦	٧,٨٢	٧	٦,٤٦	٨
❖٨,٤٦٣	٦٠	١٤	١٨	١٤	١٦,٨	٢٧	١٣,٨	٥	١١,٤	٩
❖٨,٢٨٨	٤٦	٤	١٣,٨	١١	١٢,٨٨	١٥	١٠,٥٨	١٦	٨,٧٤	١٠
١,٣٠٩	٣٠	٩	٩	٩	٨,٤	٩	٦,٩	٣	٥,٧	١١
٢,٥٨١	٥٥	١٧	١٦,٥	٢٢	١٥,٤	٩	١٢,٦٥	٧	١٠,٤٥	١٢
٠,٣١٤	٢٧	٧	٨,١	٩	٧,٥٦	٧	٦,٢١	٤	٥,١٣	١٣
❖٧,٩٩٥	٢٩	٢	٨,٧	١٠	٨,١٢	١٤	٦,٦٧	٣	٥,٥١	١٤
٢,٠١٥	٤٠	٩	١٢	١٣	١١,٢	١٣	٩,٢	٥	٧,٦	١٥
٠,٩٨٦	٣٥	٨	١٠,٥	٩	٩,٨	١١	٨,٠٥	٧	٦,٦٥	١٦
❖٨,٤١٧	٣٦	٥	١٠,٨	٢٢	١٠,٠٨	٤	٨,٢٨	٥	٦,٨٤	١٧
٠,٠٦٠	٣٠	٦	٩	٩	٨,٤	٨	٦,٩	٧	٥,٧	١٨
٠,٢١٠	٢٩	٩	٨,٧	٩	٨,١٢	٦	٦,٦٧	٥	٥,٥١	١٩
٠,٥٨٢	٦٠	١٥	١٨	٢٢	١٦,٨	١٥	١٣,٨	٨	١١,٤	٢٠
٣,٤٣٠	٥٠	١٠	١٥	٢١	١٤	١٣	١١,٥	٦	٩,٥	٢١
❖٨,٠٢١	٥٢	٦	١٥,٦	٢٣	١٤,٥٦	٨	١١,٩٦	١٥	٩,٨٨	٢٢
١,٠٦١	٣٦	٩	١٠,٨	٨	١٠,٠٨	١٢	٨,٢٨	٧	٦,٨٤	٢٣
٠,٦٨٢	٣٥	٨	١٠,٥	١٠	٩,٨	١٠	٨,٠٥	٧	٦,٦٥	٢٤
٠,٨٣٠	٦٠	١٦	١٨	٢٣	١٦,٨	١٤	١٣,٨	٧	١١,٤	٢٥
❖٨,١١٨	٢٨	٥	٨,٤	٢	٧,٨٤	٨	٦,٤٤	١٣	٥,٣٢	٢٦
❖٨,٣٩٤	٣٩	٢٤	١١,٧	٨	١٠,٩٢	٥	٨,٩٧	٢	٧,٤١	٢٧
١,٣١٠	٤٠	٩	١٢	١٣	١١,٢	١٢	٩,٢	٦	٧,٦	٢٨
١,٧٨٩	٥٠	١١	١٥	١٨	١٤	١٤	١١,٥	٧	٩,٥	٢٩
٠,٤٣٩	٦٠	١٤	١٨	٢٠	١٦,٨	١٤	١٣,٨	١٢	١١,٤	٣٠
%٣٦		نسبة المعلمات الالاتي لم تتفق نسب توزيع فقرات اختباراتهن مع النسب التي تم افتراضها الميدان التربوي								

❖ ذات دلالة إحصائية عند ($\alpha \leq 0,05$).

يظهر من الجدول (٤) أنَّ ٣٦٪ من قيم كـ^٢ ذات دلالة إحصائية عند ($\alpha \leq 0,05$)، وهذه النسبة تمثل المعلمات اللاتي لم تتفق توزيع فقرات اختباراتهن مع التوزيع المتوقع لمحتوى فصول الكتاب المقرر. وقد بلغ عدد الفقرات التي تمثل الفصل الثالث (٤٤) فقرة وكانت في المرتبة الأولى، وجاءت في المرتبة التي تليها الفقرات التي تمثل

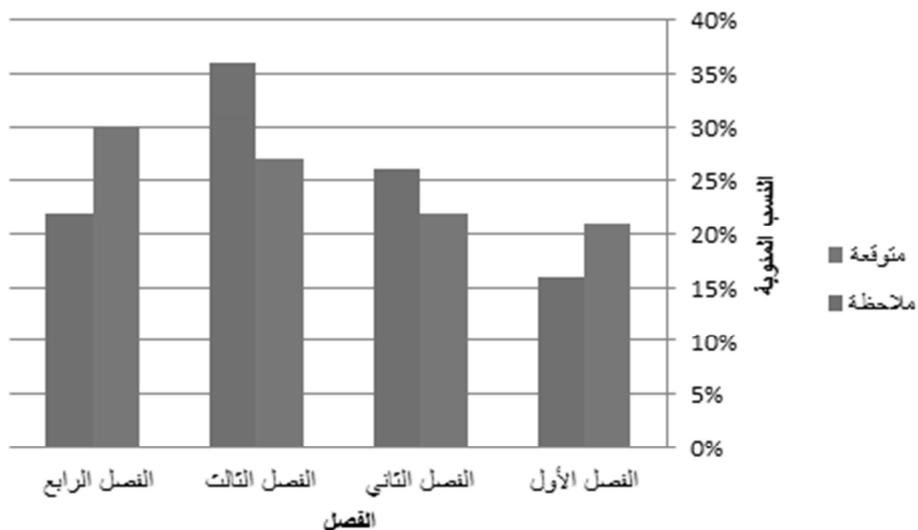
الفصل الثاني وهي (٣٢٥) فقرة، وحيث كانت عدد الفقرات التي تمثل الفصل الرابع (٢٧٦) فقرة، وجاءت في المرتبة الأخيرة عدد الفقرات التي تمثل الفصل الأول بـ(١٩٩) فقرة.

وبين الجدول (٥)؛ مدى توزيع إعداد الفقرات المتوقعة والملاحظة حسب الفصول المقررة من كتاب الرياضيات للصف الأول متوسط الفصل الدراسي الثاني، من خلال النسبة المئوية لعدد الفقرات المتوقعة والملاحظة من اختبارات معلمات الرياضيات.

جدول (٥). "النسبة المئوية لعدد الفقرات المتوقعة والملاحظة من اختبارات معلمات الرياضيات حسب الفصول المقررة من كتاب الرياضيات للصف الأول متوسط الفصل الدراسي الثاني"

نوع الفقرة	الفصل الأول	الفصل الثاني	الفصل الثالث	الفصل الرابع	مجموع الفقرات
متوقعة	٢٣٦.٣٦	٢٨٦.١٢	٣٤٨.٣٢	٣٧٣.٢	١٢٤٤
	%١٩	%٢٣	%٢٨	%٣٠	%١٠٠
الملاحظة	١٩٩	٣٢٥	٤٤٤	٢٧٦	١٢٤٤
	%١٦	%٢٦	%٣٦	%٢٢	%١٠٠

يظهر من الجدول (٥)؛ أنَّ الفصل الثالث أخذ العدد الأكْبَر من فقرات الاختبارات بنسبة ٣٦٪، يليه الفصل الثاني ٢٦٪، والفصل الرابع ٢٢٪، فيما كان العدد الأقل للفصل الأول بنسبة ١٦٪.
وبالنظر إلى التوزيع المتوقع للفقرات حسب الفصول المقررة من كتاب الرياضيات للصف الأول متوسط الفصل الدراسي الثاني، يظهر أنَّ نسبة الفقرات بالنسبة للفصل الأول ١٩٪، والفصل الثاني ٢٣٪، ثم الفصل الثالث بنسبة ٢٨٪، والفصل الرابع ٣٠٪. ويوضح الشكل (٢) المقارنة بين النسبة المئوية للفقرات المتوقعة والملاحظة على كُلِّ فصل من فصول الكتاب المقرر من خلال التمثيل بالأعمدة.



شكل (٢). "التمثيل بالأعمدة لتوزيع النسبة المئوية للفقرات المتوقعة والملاحظة على كُلِّ فصل من فصول الكتاب المقرر".

يلاحظ من الشكل (٢) أنَّ نسبة الفقراتِ الملاحظة أعلى من نسبة الفقرات المتوقعة في الفصل الثاني، والفصل الثالث، فيما كانت نسب الفقرات الملاحظة أقلَّ من نسبة الفقرات المتوقعة في الفصل الأول، والفصل الرابع.

أمّا فيما يتعلق بالكشف عن إمكانية وجود فروق ذات دلالة إحصائية ($p \leq 0.05$) بين متوسطاتِ الفقراتِ الملاحظة والفقرات المتوقعة على كلِّ فصل من فصول الكتاب المقرر، الفصل الأول والثاني والثالث والفصل الرابع؛ فقد استعملت الباحثة اختبار "ت" للكشف عن الفروق بين المتوسطات، وبيّن الجدول (٦) مدى الفروق بين متوسطاتِ الفقرات الملاحظة والفقرات المتوقعة ($p \leq 0.05$) لكلِّ فصل من فصول الكتاب المقرر.

جدول (٦). "مقارنة بين متوسطات الفقرات الملاحظة والفقرات المتوقعة لكلِّ فصل من فصول الكتاب المقرر".

الفصل الرابع	الفصل الثالث	الفصل الثاني	الفصل الأول	المستوى		نوع الفقرة
				الوسط الحسابي	متوقعة	
١٢٤٤	١١.٦١	٩.٥٣	٧.٨٧	الوسط الحسابي الانحراف المعياري	متوقعة ملاحظة	"ت" المحسوبة
٣.٣٢	٣.١٠	٢.٥٤	٢.١٠			
٩.٢٠	١٤.٨٠	١٠.٨٣	٦.٦٣			
٤.٧٤	٦.٤٤	٤.٣٤	٣.٤٢			
٤٠٦٤	٤٠٢٤٤٤	١٤١١	١.٦٩٥			

* ذات دلالة إحصائية عند ($p \leq 0.05$).

يظهر من جدول (٦) أنَّ هناك فروقاً ذات دلالة إحصائية عند ($p \leq 0.05$) بين متوسطاتِ الفقراتِ الملاحظة والفقراتِ المتوقعة، لكلِّ فصل من فصول الكتاب المقرر، وكانت الفروق بالنسبة (الفصل الثالث، الفصل الرابع)؛ وذلك لأنَّ قيم "ت" المحسوبة أكبر من "ت" الحرجة ($p \leq 0.05$)، ودرجات الحرية = ٥٨ (٢٠٠٠) وأنَّ متوسط الفقراتِ الملاحظة أعلى من متوسط الفقرات المتوقعة لكلِّ من الفصلين الثاني والثالث، أمّا بالنسبة للفصلين الأول والرابع فإنَّ متوسطاتِ الفقراتِ المتوقعة أعلى من متوسط الفقراتِ الملاحظة.

مناقشة وفسير النتائج والتوصيات والمقترنات

• مناقشة وتفسير النتائج:

فيما يتعلّق بنتائج السؤال الأول: أظهرت النتائج أنَّ ٥٦٪ من المعلمات لم يتزمن بتوزيع فقراتِ اختباراتهنَّ حسب رأي الميدان التربوي، والذي يوزع النسب على مستوياتِ الأهداف المعرفية، وأنَّ نسبة الفقرات بالنسبة لمستوى المعرفة ٢٣٪، وفقرات مستوى الاستيعاب ٢٩٪، ثم مستوى التطبيق بما فوق بنسبة ٤٨٪.

وقد تبيّن من النتائج أنَّ المعلمات يتحيّزن لمستوى المعرفة في فقراتِ اختباراتهنَّ بنسبة ٣٩٪، يليه مستوى الاستدلال ٣٧٪، فيما كان العدد الأقل لمستوى التطبيق بما فوق بنسبة ٢٤٪.

كما أظهرت نتائج من جدول (٣): أن هناك فروقا ذات دلالة إحصائية عند ($p < 0.05$) بين متوسطات الفقرات الملاحظة والفقرات المتوقعة، ولجميع مستويات الأهداف المعرفية لبلوم (المعرفة، الاستيعاب، التطبيق فما فوق)، وأن متوسط الفقرات الملاحظة أعلى من متوسط الفقرات المتوقعة لكل من مستوى المعرفة والاستيعاب. أما بالنسبة لمستوى التطبيق فما فوق؛ فإن متوسطات الفقرات المتوقعة أعلى من متوسط الفقرات الملاحظة. هذا يعني أن المعلمات يتحيزن لمستوى المعرفة والاستيعاب مقابل مستوى التطبيق فما فوق في فقرات اختباراتهن.

أختلفت النتائج مع دراسة علوة وحبيب (٢٠١٧م)، حيث جاء بالمرتبة الأولى مستوى الاستيعاب بنسبة٪٣٠، يليه مستوى التذكر٪٢٧، والتطبيق٪٢٠، وكان الاختلاف في أن مستوى الاستيعاب جاء في المرتبة الأولى، فيما تقاربت النسب بين مستويات التطبيق.

انتفقت النتائج مع دراسة جيوسي (٢٠١٦م)، حيث احتل مستوى المعرفة الترتيب الأول بنسبة٪٧٣.٨، ثم يليه مستوى الاستيعاب٪٧٣، بينما أتى مستوى التطبيق والتحليل في الأخير بنسبة٪١.٦. كذلك انتفقت النتائج مع دراسة سلمان (٢٠١٢م)، حيث إن أسئلة المعلمين كانت متحيزة وفق مستوى المعرفة والفهم، حيث بلغت (١٦٨٥) فقرة أي: بنسبة٪٩١.٤٨) من إجمالي الأسئلة. في حين أن الأسئلة لمستويات التطبيق والتحليل في المرتبة الثانية بمعدل (١٥٧) فقرة، بنسبة٪٨.٥٢). تتحيز معظم أسئلة الاختبارات لصالح مستويات الأهداف الدينية، ويمكن أن تُعزى هذه النتائج للأسباب التالية:

- تتمتع معظم الأسئلة بالسهولة؛ للتزام المعلمات لتحقيق نسبة معينة للنجاح، ولاعتقاد المعلمات بأن نجاح المعلم في مهنته يعود إلى نسبة الطلبة المتفوقين والناجحين في المقرر الذي يدرّسه.
- بعض الأسئلة لا تعطي صورة حقيقية لما تقيسه من أهداف، حيث إن الظاهر من صياغتها يختلف عما تقيسه بالفعل.
- بناءً على النتائج التي توصل إليها البحث؛ فإن الاختبارات التحصيلية التي أعدتها معلمات الرياضيات للمرحلة المتوسطة، لا تتصف بالتوازن، أي أنها لا تغطي مجالات الأهداف المعرفية بشكل وافي، وكان مستوى المعرفة وهو أدنى مستوى من مستويات الأهداف المعرفية سائداً وبشكل كبير، إذ لا يزال التركيز على الأسئلة التي تحفز الطلبة على الحفظ والاستذكار.
- عدم قدرة المعلمات على ربط الأهداف السلوكية عند صياغتها ببقية عناصر الموقف التعليمي؛ مما يعكس على أسئلة اختباراتهن، مما يجعلها ترتكز على مستوى المعرفة على حساب مستويات الأهداف المعرفية الأخرى.

فيما يتعلق بنتائج السؤال الثاني: أظهرت النتائج أن٪٣٦ من المعلمات لم يتزمن بتوزيع فقرات اختباراتهن حسب رأي الميدان التربوي، والذي توزع النسب على حسب فصول الكتاب المقرر، والذي يُظهر أن نسبة الفقرات بالنسبة للفصل الأول٪١٩، والفصل الثاني٪٢٣، ثم الفصل الثالث بنسبة٪٢٨، والفصل الرابع٪٣٠.

وقد تبيّن من النتائج أنَّ المعلمات يتحيّز لمحتوى فصول الكتاب المقرر في فقراتِ اختباراتهنَّ للفصل الأول بنسبة ١٦٪، الفصل الثاني ٢٦٪، والفصل الثالث أخذ العدد الأكْبَر من فقراتِ الاختبارات بنسبة ٣٦٪، يليه الفصل الرابع ٢٢٪.

كما أظهرت النتائج من جدول (٦): أنَّ هناك فروقاً ذات دلالة إحصائية عند (٠٥٠٥٪) بين متوسطات الفقرات الملاحظة والفقرات المتوقعة، لـكلٌّ فصل من فصول الكتاب المقرر، وكانت الفروق بالنسبة (الفصل الثالث، الفصل الرابع)، وأنَّ متوسط الفقرات الملاحظة أعلى من متوسط الفقرات المتوقعة لـكلٌّ من الفصلين الثاني والثالث. أمّا بالنسبة للفصلين الأول والرابع؛ فإنَّ متوسطات الفقرات المتوقعة أعلى من متوسط الفقرات الملاحظة. هذا يعني أنَّ المعلمات يتحيّز للفصلين الثالث والرابع، مقابل الفصلين الأول والثاني في فقراتِ اختباراتهنَّ. واتفقت نتائج البحث مع دراسة القواس (٢٠١٧م) التي أكدت أنَّ الاختبارات لا تراعي معيار تنوع الأسئلة، بالإضافة إلى عدم مراعاتها لمعيار شمول الأسئلة للمحتوى الرياضي؛ وضعف شمولها لموضوعات مقرر الرياضيات. كذلك دراسة النافع (٢٠١٦م)، التي هدفت لتقويم أسئلة اختبارات الكيمياء للصفُّ الثاني عشر بفلسطين، حيث أظهرت النتائج أنَّ معيار الشمولية كـكلٌّ في السنوات التسعة كانت بنسبة (٧٢,٨٥٪). ودراسة محمد (٢٠١٢م)، لتقويم أسئلة الرياضيات للصفُّ الثالث المتوسط للامتحانات العامة في العراق، حيث دلت النتائج أنَّ أسئلة الامتحانات كانت شاملة لجميع الفصول المقررة وبنسبة متفاوتة في عدد الأسئلة. ويمكن تفسير نتائج البحث على النحو التالي:

- ١- يرجع التحيّز في محتوى الاختبارات إلى عدم وعي المعلمات لمفهوم صدق المحتوى، وكيفية تحقيقه؛ لذلك يجب عليهنَّ التمكّن من تحقيقه؛ لتقاضي التحيّز لموضع دون آخر من موضوعات المقرر الدراسي.
- ٢- معظم الأسئلة الذي تناولها البحث دلت على شمولية الأسئلة لمحتوى المقرر الدراسي، وكان التحيّز لفصول المقرر الدراسي بنسبة ٣٦٪، وقد يعود ذلك لعدم التزام بعض المعلمات بجدول المواقف، وعدم مراعاة الأوزان النسبية لأهمية الموضوعات.
- ٣- شمولية الأسئلة لجميع فصول المقرر الدراسي وبنسبة متفاوتة، ووجود أسئلة متحيّزة لمحتوى دون آخر؛ يرجع لتركيز بعض المعلمات على الموضوعات التي تكون مبادئ الرياضيات الأساسية مثل: الهندسة والإحصاء.
- ٤- عنابة المعلمات بالفصول الأخيرة للمقرر الدراسي؛ لاعتقادهم أنَّ الوقت الذي يفصلهم عن موعد الاختبار قصير؛ لذلك تكون المعلومات والمفاهيم حاضرة في ذهن الطلبة.

النَّوْصِيَّاتُ وَالْمَقْرَرَاتُ

أولاً - النَّوْصِيَّاتُ

- ١- ضرورة تحديد الأهداف المعرفية المستهدَف قياسها قبل وضع الاختبار، مع مراعاة شمول الاختبار للمستويات المعرفية بالأوزان النسبية الناتجة عن تحليل الأهداف.

- ٢ إصدار دليل لعلمات الرياضيات، يوضح أهمية التوازن في صياغة الأسئلة من حيث قياسها للمستويات المعرفية، ومدى شموليتها للمقرر الدراسي.
- ٣ إنشاء بنوك لأسئلة الرياضيات تتوافق فيها معايير الاختبار الجيد، حتى تستطيع المعلمات الحصول على أكثر من شكل للسؤال، بحيث يحقق الهدف المرجو منه.
- ٤ تدريب الطلبة أثناء الفصل الدراسي على حل الأسئلة ذات المستويات العقلية العليا؛ حتى يتسعى للمعلمات التوسيع في مستويات الأسئلة.

ثانياً - المقترنات

- ١ إجراء بحوث تحليلية وتقديمية لاختبارات المقررات الدراسية الأخرى، خاصة التي يلاحظ تدريج مستوى الطلبة فيها، بالإضافة لمعايير أخرى لم يتناولها هذا البحث.
- ٢ إجراء بحوث للكشف عن أنماط أخرى للتحيز في الاختبارات التحصيلية، في ظل النظرية الحديثة للقياس.
- ٣ إجراء بحوث للحد من انتشار أنماط التحيز، وكيفية تجنب الواقع فيها.
- ٤ إجراء بحوث حول أساليب التقويم التي تستخدمها المعلمات، ومعرفة مدى فاعليتها.

المراجع

المراجع العربية

- جيوسى، مجدى راشد (٢٠١٦م)، مدى تحقيق أسئلة الامتحانات النهائية في جامعة فلسطين التقنية للأهداف التعليمية تبعاً لهرم بلوم في ضوء معايير الورقة الامتحانية الجيدة. المجلة التربوية الدولية المتخصصة - الجمعية الأردنية لعلم النفس -الأردن، المجلد ٥، العدد ٨، ١٥١ - ٣٧.
- الخوالدة، ناصر؛ والمشاعلة، مجدى (٢٠٠٧)، دراسة تقويمية لأسئلة امتحانات شهادة الثانوية العامة الأردنية في مبحث العلوم الإسلامية (١٩٩٧: ٥٠٠٥م) في ضوء المستويات المعرفية. مجلة جامعة النجاح للأبحاث (العلوم الإنسانية)، المجلد ٢١، العدد ٢٢، ٣٩٥ - ٤٢٠.
- ذكرى، الظاهر؛ وأخرون (٢٠٠٢م)، مبادئ القياس والتقويم في التربية. الطبعة الأولى، الدار العلمية الدولية للنشر، الأردن - عمان.
- سليمان، شاهر خالد (١٩٨٨م)، الكشف عن ثلاثة أنماط من التحبيز في اختبارات معلمي ومعلمات العلوم العامة للمرحلة الإعدادية في الأردن. رسالة ماجستير غير منشورة، جامعة اليرموك، الأردن.
- الشيخ ، تاج السر؛ ومحمد عبد الرحمن، نائل؛ وبشينة، عبد المجيد (٢٠٠٤م)، القياس والتقويم التربوي. الطبعة الثانية، الرياض: مكتبة الرشد.
- طبعمة، رشدي (٢٠٠٤م)، تحليل المحتوى في العلوم الإنسانية. مصر: القاهرة، دار الفكر العربي.
- عريفج، سامي وأخرون (١٩٩٩م)، في القياس والتقييم. الطبعة الرابعة، دار مجذولي، الأردن - عمان.
- علام، صلاح الدين محمود (٢٠١١م)، القياس والتقويم التربوي والنفسي أساسياته وتطبيقاته وتوجهاته المعاصرة. الطبعة الخامسة، القاهرة، دار الفكر.
- علوة، عبد الرزاق؛ وحبيب، عبد الحسين (٢٠١٧م)، تحليل وتقويم أسئلة الاختبارات النهائية للمواد الدراسية في الأقسام العلمية/ كلية التربية للبنات/جامعة الكوفة وفق معايير معتمدة. مجلة كلية التربية للعلوم الإنسانية، المجلد ١١، العدد ٢١، ٨٧ - ١١٤.
- عودة، أحمد (٢٠١٠م)، القياس والتقويم في العملية التدريسية. الطبعة الرابعة، عمان: دار الأمل.
- القاطعى، عبد الله (١٩٩٢م)، تحبيز بنود اختبار وكسيلر لذكاء الأطفال المعدل (الصورة السعودية) حسب الجنس. مجلة جامعة الملك سعود - العلوم التربوية والدراسات الإسلامية، السعودية، المجلد ٥، العدد ٢، ٣٧ - ٣٩٠.
- القواس، محمد أحمد مرشد (٢٠١٧م)، تقويم اختبارات القبول لمادة الرياضيات بجامعة إب في ضوء معايير الاختبار الجيد. مجلة العلوم النفسية والتربوية، المجلد ٥، العدد ١، ٢٥٧ - ٢٧٨.
- محمد، عبد الواحد محمود (٢٠١٢م)، تقويم أسئلة الرياضيات للصف الثالث المتوسط للامتحانات العامة في العراق للأعوام ٢٠٠٥: ٢٠١١م. مجلة أبحاث البصرة (العلوم الإنسانية -) العراق، المجلد ٣٧، العدد ٢، ١٣٨ - ١١٤.

ملحم، سامي محمد (٢٠١٥م)، **القياس والتقويم في التربية وعلم النفس**. الطبعة السابعة، الأردن، دار المسيرة للنشر والتوزيع.

منسي، حسن (٢٠٠٢م)، **التقويم التربوي**. الطبعة الأولى، الأردن، دار الكندي.

منصور، حسن محمد (٢٠٠٣م)، **الاختبار التحصيلي**. الطبعة الأولى، القاهرة: المركز التربوي للكمبيوتر.

الناقة، صلاح أحمد (٢٠١٦م)، **تقويم أسئلة اختبارات الكيمياء للصف الثاني عشر وفق معايير الجودة خلال السنوات من ٢٠٠٧م إلى ٢٠١٥م**. مجلة الجامعة الإسلامية للدراسات التربوية والنفسية بغزة - فلسطين، المجلد ٢٤، العدد ٣٦٤ - ٨٢.

النبهان، موسى (٢٠٠٤م)، **أساسيات القياس في العلوم السلوكية**. الطبعة الأولى، دار الشروق للنشر والتوزيع، عمان:الأردن.

يوسف، ماهر إسماعيل؛ والرافعي، محب محمود (٢٠٠٥م)، **التقويم التربوي أسلسه وإجراءاته**. الطبعة الرابعة، الرياض: مكتبة الرشد.

المراجع الأجنبية

References

- Green, K. & Stag, S.(1996), **Measuring Attitudes of teacher Toward Testing**, Measurement and Evaluation in Counseling and Development, 19,141-151.
- Jensen, A. (1981), **Bias in mental Testing**, New York, A Division of Macmillan Publishing Co. Inc.
- Kauffman ,D. & Johnson, S; & Kardos, S. (2002), **new teacher experiences with curriculum and assessment**, Teacher College Record (TCR).104,2, 273-300.
- Lord, F.M. (1980), **Application of item Response Theory to Practical Testing Problems**, Hillsdale, NJ: Erlbaum.
- Mccallon, E .L & Shumacker, R. E. (2002), **Test Bias and Differential Item Functioning**, <http://www.elmmetrics.com/articles/test>.
- Mentzer, T. L. (1982).**Response bias in multiple – choice test items bias**, Educational and Psychology measurement, 42(1), 437- 446.
- Osterlind, S. J.(1983),**Test Item Bias**, Beverly Hills: sage publication.
- Winrsma, w.& stephon, G.(1985); **Educational measurement and testing** , Boston ; Allyn and Bacon INC.
- Wotrogan, L. I.(1976).**Source of Bias in the measurement of effective teaching** , Dissertation Abstract International. 37(6) ,3532 A.