Journal of Educational and Psychological Sciences
Volume (6), Issue (3): 30 Jan 2022
P: 57 - 81
ISSN: 2522-3399
مجلة العلوم التربوية والنفسية
المجلد (6)، العدد (3): 30 يناير 2022 م
ص: 57 - 81

Accuracy of Estimate of the items and the ability parameters according to the two parameters Logistic model for TIMSS 2019 Test for the science subject for the second intermediate class in Kingdom of Saudi Arabia

Sarah Ouda Saleh Albalawi
Ministry of Education || KSA
Shaher Khaled Suleiman
College of Education and Arts || Tabuk university || KSA

Abstract: The study aimed to disclose the Accuracy of Estimate of the difficulty and distinction parameter for the item and the ability for TIMSS Test for the science subject for the second intermediate class in Kingdom of Saudi Arabia according to the two parameter Logistic model for the theory of item responding. To achieve the aim of the study, the researcher used the descriptive analysis approach. She has used the data of TIMSS 2019 Test for the science subject for a sample of 5680 male and female student. They were chosen randomly from 209 school in Kingdom of Saudi Arabia. The statistical programs SPSS and BILOG- MG3 were used to analyze the data. The results showed that: the values of the difficulty parameter ranges between (1.79 –) to (4.04) with Arithmetic average (0.42). it showed that the minimum value of the standard error in assessing the difficulty parameter which reached (0.098). the difficulty of the item value reached (- 0.69), the value of the distinction parameter ranged between (0.38) to (1.50) with Arithmetic average (0.85). The minimum value of the standard error in assessing the distinction parameter which reached (0.09). the distinction of the item value reached (0.38). As for the values of assessing the ability for the individuals in the test, it ranged between (- 3.610) to (4) with Arithmetic average (0.03). The minimum value of the standard error reached (0.49) at assessing the ability (-0.58). the study concluded many recommendations, including: establishing the training courses and workshops for male and female teachers in building the tests items according to the levels of TIMSS test, engaging the teachers in the workshops which are established by the education ministry to discuss TIMSS results and making more researches concerning TIMSS test by using the other models for the respond theory for the item.

Keywords: Accuracy of Estimate, items parameters, ability, the two parameters Logistic model, TIMSS Test.

دقة تقدير معالم الفقرات والقدرة وفق النموذج اللوجستي ثنائي المعلمة لاختبار TIMSS 2019 لمادة العلوم للصف الثاني المتوسط في المملكة العربية السعودية

ساره عوده صالح البلوي
وزارة التعليم || المملكة العربية السعودية
شاهر خالد سليمان
كلية التربية والآداب || جامعة تبوك || المملكة العربية السعودية

المستخلص: هدفت الدراسة إلى الكشف عن دقة تقدير معلمة الصعوبة والتمييز للفقرة والقدرة لاختبار TIMSS لمادة العلوم للصف الثاني المتوسط وفق النموذج الثنائي المعلم لنظرية الاستجابة للفقرة. ولتحقيق هدف الدراسة استخدم الباحثان المنهج الوصفي التحليلي، حيث استخدما بيانات اختبار TIMSS 2019 لمادة العلوم لعينة عددها (5680) طالبًا وطالبة اختيروا بشكل عشوائي من (209) مدرسة في المملكة العربية السعودية، وتم استخدام البرنامج الإحصائي SPSS وBILOG-MG3 لتحليل البيانات. وأظهرت النتائج: أن قيم معلمة الصعوبة تتراوح بين (1.79-) إلى (4.04) بمتوسط حسابي قدره (0.42)، وأن أقل قيمة للخطأ المعياري في تقدير معلمة الصعوبة بلغت (0.098) حيث بلغت قيمة صعوبة الفقرة (0.69-)، وأن قيم معلمة التمييز تراوحت بين (0.38) إلى (1.50) بمتوسط حسابي قدره (0.85)، وأقل قيمة للخطأ المعياري لتقدير معلمة التمييز بلغت (0.09) حيث بلغت قيمة تمييز الفقرة (0.38)، أما قيم تقدير القدرة للأفراد في الاختبار فتراوحت ما بين (3.61-) إلى (4) بمتوسط حسابي قدره (0.03)، وأقل خطأ معياري بلغت قيمته (0.49) عند تقدير القدرة (-0.58). وتوصلت الدراسة إلى عدّة توصيات، منها: إقامة الدورات التدريبية وورش العمل للمعلمين والمعلمات في بناء فقرات الاختبارات وفق مستويات اختبار TIMSS، بالإضافة إلى إشراك المعلمين في ورش العمل التي تُقَام من قِبَل وزارة التعليم لمناقشة نتائج TIMSS، وإجراء المزيد من الأبحاث عن اختبار TIMSS باستخدام النماذج الأخرى لنظرية الاستجابة للفقرة.

الكلمات المفتاحية: دقة تقدير، معالم الفقرات، القدرة، نظرية الاستجابة للفقرة، اختبار TIMSS.

المقدمة.

تسعى الأنظمة التعليمية باستمرار إلى تتبع عمليات التطوير والتحسين، لضمان جودة مخرجاتها المبنية على المدخلات التي ينتقيها القائمون على بناء العملية التعليمية وتطويرها وفقًا للمستجدات المواكبة للتطور العلمي والتعليمي على المستوى المحلي، والعربي، والدولي. وتتطلع المملكة العربية السعودية في نظام تعليمها إلى مواكبة كل جديد وهادف يسهم في الارتقاء بالنهضة التعليمية والفكرية التي تتواءم مع طموحها وأهدافها ويجعلها في مصاف الدول العالمية (هيئة تقويم التعليم والتدريب [ETEC]، 1439).

لذا شاركت المملكة العربية السعودية في الدراسات والاختبارات الدولية بأنواعها، واتجاهاتها، ومراحلها العمرية، وفق دراسات متتالية ومتعددة، مما يدل على حرص قيادات التعليم لدينا على رفع كفاءة التعليم والوعي بالمعايير العالمية في أساليب تعليم وتقويم كل ما يخص القراءة، والرياضيات، والعلوم وفي مجالات متعددة ولأهداف متنوعة (ETEC,1439).

هذا، وتعد دراسة التوجهات الدولية في الرياضيات والعلوم «TIMSS» Trend in International Mathematics and Science Study من أوسع الدراسات وأكثرها انتشارًا، وهي دراسة تقويم دولية تتضمن اختبارًا للتحصيل المعرفي لتتبع الأداء المدرسي لطلبة الصفين: الرابع الابتدائي والثاني المتوسط في الرياضيات والعلوم، وقد بدأت هذه الدراسة في عام 1995م، وتعقد بشكل دوري كل أربع سنوات، علمًا بأن من يشرف على تنفيذ هذه الدراسة هي المنظمة الدولية للتقويم التربوي International Association for the Evaluation of Education (IEA) وتشرف على تطبيقها محليًا هيئةُ تقويم التعليم والتدريب بالشراكة مع وزارة التعليم (Martin et al., 2008؛ ETEC,1439).

ويشارك في دراسة التوجهات الدولية في الرياضيات والعلوم «TIMSS» العديد من الدول التي يتزايد عددها باستمرار في كل دورة؛ لشمولية وثراء المعلومات والبيانات التي توفرها الدراسة حول عناصر العملية التعليمية (المنهج، والطالب، والمعلم، والمدرسة)، وتمكن هذه الدراسة المقارنة بين هذه العناصر بين الدول المشاركة فيها، ومحاولة تطبيقها بما يتناسب مع البيئة المحلية لكل دولة (الشمراني وآخرون، 2016).

وذكر الشيخي (2012) أن أهمية مشاركة الدول في تلك الاختبارات الدولية ودراسة نتائج تحصيل طلبتهم فيها؛ تبرز في تحديد أوجه القوة والقصور لديهم، ومن ثم الاستفادة من تلك النتائج في إعداد الخطط

والاستراتيجيات اللازمة لتحسين وتطوير مستوى الأداء فيها. وهو ما يحدث بالفعل في الدول المهتمة بذلك، حيث يمكن أن نلحظ الاهتمام الكبير الذي توليه الدول تجاه نتائج الاختبارات الدولية TIMSS، والاحتفاء الذي يجري في حال تحقيق الدول لنتائج متميزة على الصعيد العالمي. أما في حال تحقيق الدول لنتائج متدنية، فإننا نجد أن العديد منها تسعى عبر وزارات التعليم إلى استقصاء أسباب تدني مستوى تحصيل طلبتها في تلك الاختبارات وطرق الارتقاء به، بالإضافة إلى اتخاذ العديد من الإجراءات ذات العلاقة، كعقد الندوات وورش العمل وإعداد التقارير، بالإضافة إلى الاستعانة بالخبراء والمختصين.

وتعتمد المنظمة الدولية للتقويم التربوي IEA على منهجيات إحصائية متقدمة تم تطويرها في مجال القياس النفسي والتربوي، حيث استخدمت دراسة التوجهات الدولية في الرياضيات والعلوم TIMSS نظرية استجابة الفقرة IRT من الدورة الأولى عام 1995، مبدئيًا في شكل نموذج راش Rasch. ومنذ دورة 1999م من الدراسة استخدمت TIMSS نماذج IRT أكثر عمومية.(Von Davier et al., 2020)

وتعتبر نظرية استجابة الفقرة Item Response Theory (IRT) أو نظرية السمات الكامنة Latent Traits Theory (LLT) من الاتجاهات الحديثة في القياس والتقويم، وحظي هذا المدخل باهتمام كثير من الباحثين، حيث تغلب على كثير من مشكلات القياس التقليدية. فالاختبارات النفسية والتربوية تفترض أن هناك سمات أو خصائص معينة يشترك فيها جميع الأفراد ولكنهم يختلفون في مقدارها. وعلى الرغم من أنه لا يمكننا ملاحظة هذه السمات إلا أنه يمكن الاستدلال على وجودها من مظاهر سلوكية أو تغيرات يمكن ملاحظتها، وهذا ما يبرر تسميتها بالسمات الكامنة (أبو عواد، 2018).

ونظرية استجابة الفقرة مثل النظرية الكلاسيكية تفترض أن الاستجابة على فقرات الاختبار تعزى إلى سمات كامنة، حيث تُعرّف نظرية استجابة الفقرة بأنها نظرية التقدير الإحصائي تستخدم السمات الكامنة للأفراد والفقرة كمتنبئين للاستجابات المشاهدة، وهي في جوهرها انحدار (لوجستي ثنائي أو متعدد الحدود) للاستجابات المشاهدة للفقرة على مواقع الأشخاص على المتغير الكامن والخصائص الكامنة للفقرة (De Ayala, 2009 / 2017).

ويمكن وصف العلاقة بين احتمال الاستجابة الصحيحة لفقرة من الفقرات ومقياس القدرة بالمنحنى المميز للفقرة The Item Characteristic curve (Baker, 2001). ويعد مفهوم المنحنى المميز للفقرة من المفاهيم المهمة في نماذج الاستجابة للفقرة التي تتناول سمة كامنة متصلة يتم قياسها بفقرات اختبارية من النوع الثنائي أي تتطلب إجابة صحيحة أو إجابة خاطئة. ويمثّل هذا المنحنى بدالة رياضية تربط بين احتمال نجاح المفحوص في إجابة الفقرة وبين السمة أو القدرة التي تقيسها مجموعة فقرات أو يقيسها اختبار معين، أي أنها دالة انحدار درجة الفقرة على السمة الكامنة التي يقيسها الاختبار، وهذه الدالة غير خطية No Linear (علام , 2011).

وتكمن أهمية نظرية استجابة الفقرة وتطبيقاتها في تحليل الفقرة؛ في التوصل إلى معالم فقرة ثابتة نسبيًا لا تتغير مع تغير مجموعة المفحوصين التي اسُتخدمت في تحليل الفقرة، فإذا كانت المعالم نسبيًا ثابتة يمكن تقديرها من بيانات مجموعة واحدة ثم تطبيق هذه التقديرات بثقة على أية مجموعة من المفحوصين بما فيهم المجتمع الكلي للمفحوصين. أما في النظرية الكلاسيكية فإن إحصائيات التحليل التقليدي للفقرة تتغير مع تغير المفحوصين (Crocker & Algina.2008/2017).

وفي النظرية الكلاسيكية تتأثر خصائص مجموعة المفحوصين بخصائص الفقرات أي أن الدرجة الحقيقية (True Score) والدرجة الملاحظة (Observe Score) تتغيران بصعوبة الاختبار. أما نظرية استجابة الفقرة فتميز بين الدرجة الحقيقية (True Score) ودرجة القدرة (Ability Score) مما يعني أن قدرة المفحوص ثابتة في موقف معين ولا تتأثر بخصائص الفقرات (Item Free) (عودة , 2010).

ولتعزيز التكامل والتنسيق بما يحقق مستهدفات رؤية المملكة 2030 في رفع جودة التعليم والتدريب؛ اطلعت الباحثان على نتائج طلاب المملكة العربية السعودية المشاركين في الدورات السابقة، ووجدت أنهم حققوا مراكز متأخرة في جميع المشاركات، حيث حصلت المملكة على متوسط أداء أقل من متوسط المعيار الدولي لـ TIMSS (500 نقطة)، لذا يسعى هذا البحث للكشف عن تقديرات معالم الفقرة والقدرة ودقة تقديرها باستخدام نظرية استجابة الفقرة لما تتميز به من موضوعية ودقة في القياس مقارنة بالنظرية الكلاسيكية.

مشكلة الدراسة وأسئلتها:

من تطلعات الرؤية الوطنية 2030 تطوير نظم التعليم العام في المملكة العربية السعودية، وذلك لدفع عجلة الاقتصاد الوطني، وتنويعه، وتحقيق التنمية المستدامة للمجتمع السعودي. كما أن وزارة التعليم تسعى أثناء مسيرتها لتطوير التعليم إلى استيعاب مفاهيم وقيم ومتطلبات المرحلة الحالية والمستقبلية، وإلى تشخيص واقع التعليم باستخدام أدوات معيارية متكاملة تستند إلى أفضل التجارب، والممارسات العالمية (وزارة التعليم، د.ت.).

ومن منطلق التنافسية الدولية وتحقيق جودة التعليم؛ حرصت المملكة العربية السعودية على تفعيل مشاركتها في دراسة الاتجاهات الدولية في العلوم والرياضيات TIMSS 2019، وخوض غمار المنافسة الدولية مع الدول المتقدمة، وتقليص الفجوة في نتائج الدورة السابقة 2015م، وتعزيز الانتماء الوطني إلى أبنائها، وأنهم قادرون على المنافسة وإحداث التغيير المنشود (وزارة التعليم، د.ت.).

وبناء على النتائج التي نشرتها المنظمة الدولية للتقويم التربوي IEA أصدرت هيئة تقويم التعليم والتدريب تقريرًا مفصلًا حول نتائج مشاركة المملكة في اختبارTIMSS 2019 وأكدت النتائج أن هناك تحسنًا ملحوظًا في مؤشرات التحصيل الدراسي جميعها مقارنة مع TIMSS 2015، وهو تحسن يعكس جهود التحسين التي بُذلت من قبل الأطراف المختلفة المعنية بالتعليم بقيادة وزارة التعليم، لكن النتائج لاتزال منخفضة قياسًا بمعايير الأداء الدولية (ETEC,2020).

ونشرت المنظمة في موقعها الرسمي جميع البيانات الخاصة بالاختبار الدولي TIMSS حتى تتيح للباحثين والمهتمين الاطلاع عليها وتحليلها، والوقوف على نتائج تحليلهم مما يخدم أهداف بحثهم. وحرصًا من الباحثين على اكتشاف الفجوة في أداء الطلاب على اختبار TIMSS في مادة العلوم؛ جاءت هذه الدراسة لتتناول تحليل اختبار TIMSS 2019 باستخدام نظرية استجابة الفقرة للكشف عن دقة تقديرات معالم الفقرات لاختبار TIMSS 2019 والقدرة للمفحوصين وفق النموذج اللوجستي ثنائي المعلمة.

لذا تتحدد مشكلة الدراسة في الأسئلة الآتية:

1- ما دقة تقدير معلمة الصعوبة لفقرات اختبار TIMSS 2019 لمادة العلوم للصف الثاني المتوسط وفق النموذج اللوجستي ثنائي المعلمة؟
2- ما دقة تقدير معلمة التمييز لفقرات اختبار TIMSS 2019 لمادة العلوم للصف الثاني المتوسط وفق النموذج ثنائي المعلمة؟
3- ما دقة تقدير معلمة القدرة لأداء الطلاب المشاركين على اختبار TIMSS 2019 لمادة العلوم للصف الثاني المتوسط وفق النموذج اللوجستي ثنائي المعلمة؟
4- ما معاملات ثبات اختبار TIMSS 2019: (الثبات الإمبريقي، وثبات الأفراد، وثبات الفقرات) لمادة العلوم للصف الثاني المتوسط وفق النموذج اللوجستي ثنائي المعلمة؟

أهداف الدراسة:

تهدف هذه الدراسة إلى:

1- الكشف عن دقة تقدير معلمة الصعوبة لفقرات اختبار TIMSS 2019 لمادة العلوم للصف الثاني المتوسط وفق النموذج اللوجستي ثنائي المعلمة.

2- الكشف عن دقة تقدير معلمة التمييز لفقرات اختبار TIMSS 2019 لمادة العلوم للصف الثاني المتوسط وفق النموذج اللوجستي ثنائي المعلمة.

3- الكشف عن دقة تقدير معلمة القدرة لأداء الطلاب على اختبارTIMSS 2019 لمادة العلوم للصف الثاني المتوسط وفق النموذج اللوجستي ثنائي المعلمة.

4- الكشف عن معاملات ثبات اختبارTIMSS 2019: (الثبات الإمبريقي، وثبات الأفراد، وثبات الفقرات) لمادة العلوم للصف الثاني المتوسط وفق النموذج اللوجستي ثنائي المعلمة.

أهمية الدراسة:

- الأهمية النظرية:

- تسهم هذه الدراسة في تقديم بحث تحليلي ومؤشرات كمية ذات دلالة للأداء على اختبار TIMSS الدولي، وذلك باستخدام النموذج اللوجستي ثنائي المعلمة المنبثق عن نظرية الاستجابة للفقرة والتي تحظى بأهمية كبيرة في مجال القياس.
- تسهم هذه الدراسة في تسليط الضوء على ضرورة بناء الاختبارات التربوية وفق نظرية الاستجابة للفقرة للتغلب على عيوب النظرية الكلاسيكية في القياس.
- دعم القاعدة النظرية للبحوث والدراسات في مجال نظرية الاستجابة للفقرة وفي الاختبارات الدولية TIMSS.

- الأهمية التطبيقية:

- يتوقع الباحثان أن تتم الاستفادة من نتائج الدراسة في توجيه القائمين على العملية التعليمية في اتخاذ قرارات صائبة لتحسين مستوى الأداء على الاختبار.
- قد تفيد هذه الدراسة الجهات المختصة ببناء وتطوير الاختبارات في تحسين طريقة إعداد الاختبارات وبنائها، والتحقق من مدى مطابقتها لمعايير الاختبار الجيد من وجهة نظر نظرية الاستجابة للفقرة.
- قد تفيد معلمي العلوم في انتقاء فقرات الاختبارات التحصيلية أو بناء اختباراتهم بحيث تتميز بتقديرات مرتفعة لمعالم الفقرات.
- يؤمل الباحثان أن تسهم هذه الدراسة في حث مشرفي ومعلمي العلوم على الاهتمام بمستويات الأسئلة التي تطرح في القاعة الصفية ونوعيتها بطريقة تحاكي الأسئلة التي تطرح في اختبارات TIMSS.

حدود الدراسة:

تقتصر هذه الدراسة على الحدود الآتية:

- الموضوعية: دقة تقدير معالم الفقرات والقدرة وفق النموذج اللوجستي ثنائي المعلمة لاختبار TIMSS 2019 لمادة العلوم.
- البشرية: طلبة الصف الثاني المتوسط المشاركين في الاختبار الدولي TIMSS 2019.

- المكانية: المدارس المشاركة في الاختبار الدولي TIMSS 2019 في المملكة العربية السعودية.
- الزمانية: تم تطبيق هذه الدراسة خلال العام الدراسي 2020 / 2021 م.

مصطلحات الدراسة:

- دقة التقدير Accuracy of Estimate، هي: "تعبير يشير إلى جودة التقدير للمعالم والقدرة التي تميزها الاحتمالية الكبيرة في أن التقدير قريب من القيمة الحقيقية للمعلم، باستخدام الأخطاء المعيارية في التقدير" (الدرابسه وعودة، 2012: 63).
- الصعوبة difficulty of the item، هي: "نقطة على مقياس القدرة عندما يكون احتمال الاستجابة الصحيحة يساوي 0.50" (Baker, 2011: 22)
  - وتُعرّف إجرائيًا بأنها: القيمة التي تندرج على متصل السمة ويكون احتمال إجابة الفقرة عندها إجابة صحيحة مساويًا 50%.
- التمييز Discrimination of the item، هو: ميل المنحنى المميز له بالقرب من الوسط، أي: بالقرب من النقطة على مقياس القدرة التي يُتوقَّع أن يكون احتمال الاستجابة الصحيحة عن الفقرة (0.50) (Baker, 2011: 22)
  - ويعرّف التمييز إجرائيًا بأنه: قدرة الفقرة على التمييز بين الأفراد ذوي القدرة المرتفعة وذوي القدرة المنخفضة.
- القدرة (Ability): سمة كامنة يمكن من خلالها التنبؤ بأداء الفرد على فقرة من فقرات الاختبار، ويرمز لها بالرمز ثيتا θ (علام، 2005: 47).
  - وتُعرّف القدرة إجرائيًا بأنها: قيمة يتم تقديرها تعظم أرجحية استجابات الفرد عن فقرات الاختبار.
- النموذج اللوجستي ثنائي المعلمة Two-Parameter Model، هو: أحد نماذج نظرية الاستجابة للفقرة ثنائية التدريج، حيث يسمح للفقرات بأن تختلف في صعوبتها وتمييزها، بالإضافة إلى افتراضه أن الفرد لا يلجأ إلى التخمين العشوائي في إجابته عن فقرات الاختبار (علام، 2011: 700)
  - ويُعرّف إجرائيًا بأنه: نموذج ثنائي المعلمة يعمل على تقدير معلمة الصعوبة والتمييز للفقرات.
- اختبار TIMSS، هو اختبار دولي لتقييم تحصيل الطلاب بمادتي الرياضيات والعلوم للصف الرابع والثامن الأساسيين، وتشرف عليه الرابطة الدولية لتقييم التحصيل التربوي.

2- الإطار النظري والدراسات السابقة.

أولا- الإطار النظري: دراسة الاتجاهات الدولية في الرياضيات والعلوم TIMSS

تعد دراسة الاتجاهات الدولية في الرياضيات والعلوم TIMSS دراسة دولية عالمية لتقييم الاتجاهات في مدى تحصيل الطلاب والطالبات في مادتي الرياضيات والعلوم عند طلاب الصفين: الرابع الابتدائي والثامن (الثاني المتوسط). حيث تسعى معظم نظم التعليم في العالم للمشاركة في هذه الدراسات والاختبارات، إما بهدف إبراز تميز النظام التعليمي فيها، أو تقويمه للاطلاع على مستوى أدائه، رغبة منها في مواكبة الدول المتفوقة في هذا المجال (ETEC,1439).

بدأ تطبيق دراسة الاتجاهات الدولية في الرياضيات والعلوم TIMSS عام 1995م، حيث تعقد بشكل دوري كل أربع سنوات، وشارك في الدورة الأولى 26 دولة في الصف الرابع الابتدائي و41 دولة في الصف الثامن، ثم انعقدت دورتها الثانية في عام 1999م، وشارك فيها 38 دولة في الصف الثامن، ولم يتم تطبيق الدراسة على الصف الرابع

الابتدائي. وكانت الدورة الثالثة للدراسة في عام 2003م، وشارك فيها 25 دولة في الصف الرابع الابتدائي، و46 دولة في الصف الثامن. والدورة الرابعة كانت في عام 2007م، وشارك فيها 36 دولة في الصف الرابع الابتدائي، و48 دولة في الصف الثامن. ودورتها الخامسة في عام 2011، وشارك فيها 50 دولة في الصف الرابع الابتدائي، و42 دولة للصف الثامن. وفي عام 2015م انعقدت الدورة السادسة، وشارك فيها 57 دولة أساسية، و7 مقاطعات ومدن للمقارنة (Benchmarking)، منها 49 دولة شاركت في الرياضيات و47 دولة شاركت في العلوم للصف الرابع الابتدائي، و39 دولة شاركت في الرياضيات والعلوم للصف الثامن. أما دورتها الحالية TIMSS 2019 فهي الدورة السابعة، وشارك فيها 64 دولة و8 مقاطعات ومدن للمقارنة المرجعية (Benchmarking). شاركت 58 دولة و6 مقاطعات ومدن للمقارنة المرجعية للصف الرابع الابتدائي، و39 دولة و7 مقاطعات ومدن للمقارنة المرجعية للصف الثامن (الشمراني وآخرون، 2016؛ Mullis et al., 2020).

نموذج المنهج في TIMSS:

تبنت دراسة التوجهات الدولية في الرياضيات والعلوم TIMSS نموذجًا للمنهج يركز بشكل رئيس على كيفية توفير الفرص التعليمية للطلاب، والعوامل التي تؤثر على كيفية استخدام الطلاب لهذه الفرص. ويحتوي نموذج منهج TIMSS على ثلاثة جوانب وهي:

- المنهج المقصود Intended Curriculum: يقصد به المنهج الذي يحدد على المستوى الوطني للدولة والذي يراعي الهوية والمحتوى الاجتماعي والتعليمي.
- المنهج المنفذ Implemented Curriculum: يقصد به المنهج الذي يطبق في المدرسة من قبل المعلم حيث يراعي المحتوى المدرسي والبيئة المدرسية.
- المنهج المكتسب Attained Curriculum: يقصد به المنهج الذي يتعلمه الطالب ويكتسب من خلاله بعض الصفات. (الفهيدي،1433هـ ؛Mullis & Martin, 2017)

يتبين من خلال نموذج منهج TIMSS أنها دراسة شاملة تعكس مدى الجودة في النظام التعليمي للدول المشاركة، من خلال المقارنة بين المأمول في المنهج المستهدف، والواقع الفعلي في المنهج المكتسب، بالإضافة إلى المقارنة بينها وبين أنظمة الدول المتقدمة. وكيفية هذه الأنظمة للوصول إلى تحقيق أهداف النظام التربوي والتعليمي، من خلال معرفة مواطن القوة وتعزيزها، ومعرفة مواطن الضعف والعمل على معالجتها (السليماني،2016).

مجالات TIMSS 2019:

تم تحديث أطر تقييم TIMSS لعام 2019 من تلك المستخدمة في عام 2015. حيث تم تحديد مجالين لدراسة TIMSS 2019 هما: بعد المحتوى وبعد التفكير (العمليات المعرفية). وستعرض الباحثان في السطور القادمة المجالات التي تخص العلوم للصف الثامن (الثاني المتوسط) كما ورد في الإطار النظري TIMSS Science Framework) (2019 لدراسة التوجهات الدولية في الرياضيات والعلوم TIMSS 2019 وهي كما يلي:

أولًا: مجالات المحتوى المعرفي Content Domains: تحدد أربعة مجالات محتوى رئيسية للمحتوى العلمي لتقييم الصف الثامن في TIMSS Science وهي: علم الأحياء، والكيمياء، والفيزياء، وعلوم الأرض. يتضمن كل مجال من مجالات المحتوى العديد من مجالات الموضوعات الرئيسية، وكل مجال فرعي بدوره يشمل موضوعًا واحدًا أو أكثر. ويتم وصف كل موضوع بمزيد من التفاصيل من خلال أهداف محددة تمثل المعرفة، والقدرات، والمهارات المتوقعة من الطلاب، والتي تم تقييمها في كل موضوع. وقد تم تقسيم مجالات موضوعات المحتوى في كتاب الصف

الثامن إلى نسب مئوية، ويوضح الجدول (1) النسب المئوية المستهدفة لكل مجال من مجالات المحتوى الأربعة في التقييم العلمي TIMSS 2019. (Mullis & Martin, 2017).

جدول (1): النسبة المئوية لمجالات موضوعات المحتوى المستهدفة في دراسة TIMSS 2019

مجالات موضوعات المحتوى للصف الثامن	النسبة المئوية
الأحياء	%35
الكيمياء	%20
الفيزياء	%25
علم الأرض	%20

ثانيًا: مجال العمليات المعرفية Cognitive Domains: ينقسم مجال العمليات المعرفية إلى ثلاثة مجالات فرعية تصف عمليات التفكير التي يُتوقع من الطلاب المشاركة فيها للإجابة بشكل صحيح عن الفقرات التي يغطيها اختبار TIMSS 2019. ويوضح الجدول (2) النسب المئوية المستهدفة لتقييم العمليات المعرفية الثلاث في الصف الثامن. (Mullis & Martin, 2017).

جدول (2): النسبة المئوية لمجالات العمليات المعرفية المستهدفة في دراسة TIMSS 2019

مجال العمليات المعرفية للصف الثامن	النسبة المئوية
المعرفة	%35
التطبيق	%35
الاستدلال	%30

أدوات دراسة التوجهات الدولية في الرياضيات والعلوم TIMSS 2019:

تتضمن الدراسة عدة أدوات خاصة بها كما ورد في (حسان، 2013؛ السليماني، 2016؛ Mullis & Martin, 2017؛ ETEC, n.d.) وهي على الشكل الآتي:

أ- كراسات الاختبار: يتراوح عدد كل نوع من الكتيبات بين (7-14) نموذجًا من الكتيبات المكافئة توزع على الطلبة بطريقة عشوائية عن طريق البرمجيات الخاصة بهذه الدراسة التي تحدد اسم الطالب ورقم الكتيب الخاص به. بحيث يشتمل كل كتيب على عدد من أسئلة الرياضيات والعلوم، (70%) من هذه الأسئلة: نوع الاختيار من متعدد و (30%) أسئلة الإجابة القصيرة المعتمدة على استنتاج الحل، ومدة الاختبار (90) دقيقة موزعة بين جلستين (الرياضيات – العلوم).

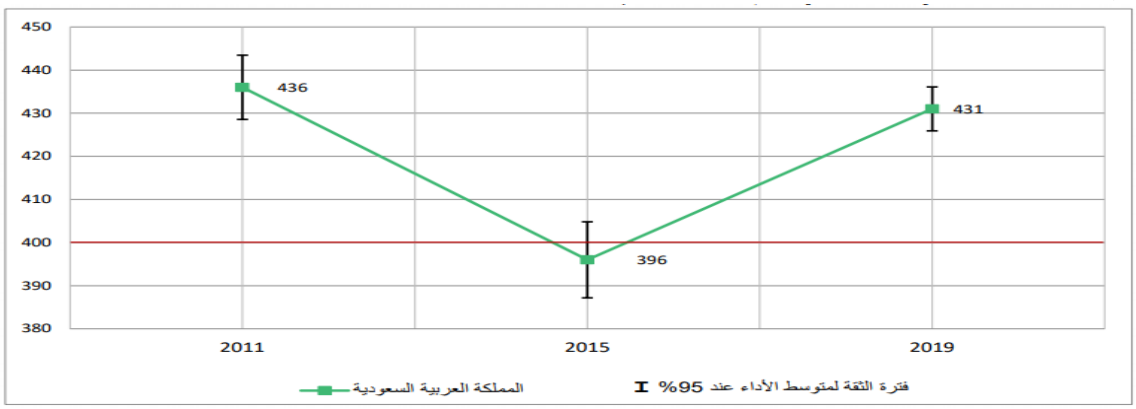
ب- استبانات الدراسة: تفيد استبانات الدراسة التي تسهم فيها الدول المشاركة في تقديم معلومات مهمة لموسوعة TIMSS حول سياساتها التعليمية، ومناهجها الدراسية في مادتي الرياضيات والعلوم؛ بهدف تطوير الطالب أكاديميًّا ومهنيًّا، بالإضافة إلى تطوير تعليم الرياضيات والعلوم وتعليمهما وتحسينه في الدول المشاركة من خلال الأنظمة التربوية والتعليمية التي تقدمها الدول المشاركة. وتشمل استبانات الدراسة ما يلي: استبانة الطالب، استبانة المعلم، استبانة المدرسة، استبانة المنزل، واستبانة المنهج. (السليماني، 2016).

مشاركة المملكة العربية السعودية في دراسة التوجهات الدولية في الرياضيات والعلوم TIMSS:

تستهدف دراسة التوجهات الدولية في الرياضيات والعلوم الصفين: الرابع والثامن (الثاني المتوسط) الأساسيين في الدول المشاركة، حيث تقوم كل دولة مشاركة باختيار أحد الصفين أو كليهما معًا للمشاركة. وكانت أول

مشاركة للملكة العربية السعودية في دراسة التوجهات الدولية في الرياضيات والعلوم TIMSS في الدورة الثالثة عام (2003م) حيث شاركت بالصف الثامن (الثاني المتوسط)، وشاركت أيضًا في الدورة TIMSS 2007 بالصف الثامن. أما في الدورة الخامسة TIMSS 2011 والدورة السادسة TIMSS 2015 والدورة السابعة TIMSS 2019 فقد شاركت المملكة العربية السعودية بالصفين: الرابع الابتدائي والثامن (الثاني المتوسط) (السليماني، 2016).

وهنا سنستعرض مستوى أداء طلاب المملكة العربية السعودية في الدورات التي شاركت فيها، ففي الدورة الثالثة TIMSS 2003 كأول مشاركة للمملكة العربية السعودية جاء ترتيب الطلاب في الموقع (39) من بين (46) دولة مُشاركة في الدراسة، بمتوسط أداء مقداره (398) بأقل بـ (76) نقطة من المتوسط الدولي (474)، وفي الدورة الرابعة TIMSS 2007 شاركت في الدراسة (48) دولة، جاء ترتيب طلاب الصف الثاني متوسط (الثامن) في المملكة في الموقع (44)، بمتوسط أداء مقداره (403) بأقل بـ (79) نقطة من المتوسط الدولي المُقدر بـ (500)، أما في الدورة الخامسة TIMSS 2011 فجاء ترتيب الطلاب في المرتبة (31) من بين (45) دولة مُشاركة، بمتوسط أداء مقداره (436) بأقل بـ (64) نقطة من المتوسط الدولي المُقدر بـ (500)، وفي الدورة السادسة TIMSS 2015 جاء ترتيب طلاب المملكة العربية السعودية في المرتبة (35) من بين (39) دولة مشاركة في الدراسة، بمتوسط أداء مقداره (396) بأقل من المتوسط الدولي (500) بـ (104) نقطة (الرويلي،2019). وفي الدورة السابعة TIMSS 2019 جاء ترتيب طلاب المملكة العربية السعودية في المرتبة (35) من بين (39) دولة مشاركة في الدراسة، بمتوسط أداء مقداره (431) بأقل من المتوسط الدولي (69) نقطة. ويوضح الشكل البياني التالي (1) متوسط درجات طلاب الصف الثاني المتوسط في المملكة في العلوم وفق TIMSS من2011 إلى 2019 (ETEC,2020):



شكل (1): التغيرات في متوسط الأداء في العلوم لطلبة المملكة العربية السعودية حسب اختبار TIMSS

ثانياً- الدراسات السابقة:

فيما يلي عرضٌ لمجموعة من الدراسات السابقة ذات الصلة بموضوع هذه الدراسة، من خلال دراسات تناولت استخدام نظرية الاستجابة للفقرة متمثلة في النموذج اللوجستي ثنائي المعلمة في بناء الاختبارات وتحليلها، ودراسات تناولت دراسة الاتجاهات الدولية في الرياضيات والعلوم TIMSS، وقد تم ترتيب الدراسات ترتيبًا زمنيًا بدءًا من الأحدث وانتهاءً بالأقدم.

- **دراسة جراح (2020):** هدفت هذه الدراسة إلى التعرف على أنماط الاستجابة في الاختبار الدولي (TIMSS, 2015) في مبحثي العلوم والرياضيات لطلبة الصف الثامن في المملكة العربية السعودية ولطلبة الصف الثامن من دولة سنغافورة باستخدام مؤشرات مطابقة الشخص (مؤشر التمكن، ومؤشر سوء الفهم، ومؤشر التخمين، ومؤشر عدم المبالاة). واستخدم الباحث المنهج الوصفي التحليلي، وبلغت العينة (270) طالبًا وطالبة من المملكة العربية السعودية ومن دولة سنغافورة (270) طالبًا وطالبة، وأظهرت نتائج التحليل أنماط

الاستجابة لبيانات دولة سنغافورة لمبحث العلوم والرياضيات؛ أن 92% و94% على التوالي صنفت كأنماط استجابة طبيعية، وأن 8% و6% على التوالي صنفت كأنماط استجابة غير طبيعية. بينما أظهرت نتائج تحليل أنماط الاستجابة لبيانات طلبة الدولة السعودية لمبحث العلوم والرياضيات؛ أن 85% و86% على التوالي صنفت كأنماط استجابة طبيعية، وأن 15% و14% على التوالي صنفت كأنماط استجابة غير طبيعية، وأن نسبة أنماط الاستجابة غير الطبيعية لبيانات مبحث الرياضيات لطلبة الدولة السعودية (12%) كانت أكثر منها لمبحث العلوم (11%)، وأن نسبة أنماط الاستجابة غير الطبيعية لبيانات مبحث العلوم لدولة سنغافورة (8%) أكثر منها لمبحث الرياضيات (6%).

- **دراسة الرويلي (2019):** هدفت هذه الدراسة إلى تحليل مقارن لمحتوى مقرر العلوم للصف الثاني المتوسط (الثامن) في المملكة العربية السعودية والإمارات العربية المتحدة في ضوء متطلبات دراسة التوجهات الدولية للرياضيات والعلوم (TIMSS 2015). واستخدم الباحث المنهج الوصفي التحليلي الذي يتمثَّل في تحليل المحتوى. ولتحقيق أهداف هذه الدراسة قام الباحث بإعداد أداة الدراسة المتمثِّلة في قائمة بمتطلبات TIMSS 2015 وذلك استنادًا إلى الأبعاد التي حددتها الجمعية الدولية لتقييم التحصيل التربوي (IEA)، وتم تحويلها إلى بطاقة تحليل محتوى، تكون مجتمع الدراسة وعينتها من جميع الدروس لكتاب الطالب وكراسة التجارب العملية لمقرر العلوم للصف الثاني متوسط في المملكة العربية السعودية للفصلين الدراسيين: (الأول والثاني) في طبعتهما لعام 1438هـ/2017م، وكتاب العلوم للصف الثامن في الإمارات العربية المتحدة بفصوله الثلاثة: (الأول، والثاني، والثالث) في طبعته 2016م. وأظهرت النتائج أن درجة تحقق مجالات بُعد المحتوى في مقرر العلوم في المملكة العربية السعودية؛ جاءت ضعيفة. أما مقرر العلوم للصف الثامن في الإمارات العربية المتحدة، فجاء بدرجة تحقق متوسطة. أما متطلباتTIMSS 2015 المتعلقة بتحقق مجالات بُعد التفكير في مقرر العلوم للصف الثاني متوسط في المملكة العربية السعودية، فجاءت بدرجة متوسطة وبنسبة تضمين 69.96%، في حين جاءت بدرجة تحقق كبيرة وبنسبة تضمين 97.55% في مقرر العلوم للصف الثامن في الإمارات العربية المتحدة.
- **دراسة العنزي (2019):** هدفت هذه الدراسة إلى التعرف على أثر شكل فقرات الاختيار من متعدد والصواب والخطأ على معالم الفقرة والخصائص السيكومترية للاختبار وفق نموذج ثنائي المعلم لنظرية استجابة الفقرة، ولتحقيق هذا الهدف استخدمت الدراسة المنهج شبه التجريبي وتمثلت الأداة في بناء شكلين للاختبار في مادة الحاسب الآلي: الاختيار من متعدد والصواب والخطأ، حيث تكون كل شكل من شكلي الاختبار في صورته النهائية من (24) فقرة، وتم تطبيق الاختبار على (421) طالبًا. كما تم استخدام البرنامج الإحصائي (SPSS) وبرنامج (Bilog-Mg3) لتحليل البيانات. وأظهرت النتائج عدم وجود فروق ذات دلالة إحصائية في صعوبة الفقرة بين شكلي الاختبار، إلا أن النتائج كشفت عن وجود فروق ذات دلالة إحصائية في معامل الثبات التجريبي بين شكلي الاختبار.
- **دراسة المرواني (2019):** هدفت هذه الدراسة إلى الكشف عن أثر موقع البديل الصحيح في اختبار اختيار من متعدد ذي أربعة بدائل على دقة تقدير معلمة الصعوبة والتمييز للفقرات، وعلى تقديرات معاملات الثبات التجريبي ومعاملات ثبات الأفراد، وفق النموذج ثنائي المعلمة لنظرية الاستجابة للفقرة. ولتحقيق أهداف الدراسة واستخدمت الدراسة المنهج شبه التجريبي وتمثلت الأداة في بناء اختبار تحصيلي من نوع اختيار من متعدد ذي أربعة بدائل في مادة العلوم مكون في صورته النهائية من (31) فقرة، ومن ثلاثة نماذج لها نفس المحتوى والمتغير هو موقع البديل الصحيح. وطبقت النماذج الثلاثة على عينة عشوائية عنقودية بلغت (500)

طالبة من طالبات الصف الثالث متوسط في مدينة ينبع. وتم استخدام البرنامج الإحصائي (SPSS) وبرنامج (Bilog-Mg3) لتحليل البيانات. وأظهرت النتائج عدم وجود فروق دالة إحصائيًا بين متوسطات تقديرات معاملات ثبات الأفراد تعزى إلى موقع البديل الصحيح.

- **دراسة السلامي (2018):** هدفت هذه الدراسة إلى بناء اختبار محكي المرجع في الرياضيات باستخدام نظرية الاستجابة للفقرة وفق النموذج اللوجيستي ثنائي المعلم لطلاب الصف الخامس الابتدائي، وذلك باستخدام المنهج الوصفي التحليلي. وجاءت الأدوات متمثلة في بناء اختبار محكي المرجع في مادة الرياضيات للصف الخامس الابتدائي، وتم تطبيقه على عينة من طلاب الصف الخامس الابتدائي التابعة لإدارة تعليم تبوك وبلغ عددهم (400) طالب. وأظهرت النتائج باستخدام برنامج (BILOG-MG3) تحقق افتراضات النموذج اللوجستي ثنائي المعلم ومطابقة (32) فقرة. وتم حذف الفقرات الستة غير المطابقة للنموذج اللوجيستي ثنائي المعلم، وتتمتع بخصائص سيكومترية مقبولة ضمن المحكات التي أوردتها أدبيات القياس التربوي.
- **دراسة دبوس (2016):** هدفت هذه الدراسة إلى استخدام نظرية الاستجابة للفقرة في بناء تجمع فقرات اختبار محكي المرجع في الرياضيات وفق النموذج اللوجستي ثنائي المعلم. واستخدمت الدراسة المنهج الوصفي المسحي، وتمثلت الأداة في إعداد تجمع من الفقرات ثنائية التدريج مكون من (50) فقرة، وتجمع من الفقرات متعددة التدريج مكون من (10) فقرات. وتكونت عينة الدراسة من (502) من طلاب وطالبات الصف الثامن الأساسي، واستخدم الباحث برنامج MULTILOG.7 لتحليل البيانات. وأظهرت النتائج: تَحقُّق افتراضات نظرية الاستجابة للفقرة، كما بيَّنَت نتائج التحليل مطابقة جميع الفقرات ثنائية التدريج للنموذج الثنائي المعلم، وأن الفقرات المتعددة التدريج كانت مطابقة لنماذج الاستجابة المتدرجة CRM.
- **دراسة Shawn (2012):** هدفت هذه الدراسة إلى إجراء تقييم مستقل لعينة ممثلة من فقرات اختبار TIMSS في العلوم. واستخدمت الدراسة المنهج الوصفي وتمثلت الأداة في بيانات اختبار TIMSS2007 الكتيب الثاني حيث تم تطبيقه على 16009 طالب في الصف الثامن من 50 دولة. يتكون التقييم من ثلاثة مكونات: أولًا: استخدام نموذج راش، الذي يركز على الفقرات عالية الجودة، لتقييم الفقرات من الناحية السيكومترية (الصدق، والثبات، والصعوبة) واستخدم الباحث برنامج Winsteps. ثانيًا: تم استخدام تحليلات المقروئية والفقرات لتقييم صياغة الفقرات للتأكد من أنها كانت مفهومة للطلاب. ثالثًا: تم استخدام إرشادات تطوير الفقرات من قبل مجموعة مركزة من معلمي العلوم لتقييم الفقرات في ضوء إطار تقييم TIMSS، والذي حدد التنسيق والمحتوى والمجالات المعرفية للفقرات. وأشارت مكونات التقييم إلى أن غالبية الفقرات كانت ذات جودة عالية، مما يسهم في صحة درجات اختبار TIMSS. كان لهذه الفقرات خصائص سيكومترية جيدة، وتميزت بسهولة القراءة، ووضوح الفقرات، والتوافق مع إطار التقييم. بشكل عام، كانت الفقرات التي تميل إلى أن تكون صعبة: فقرات ذات الإجابة القصيرة التي تقيم مستوى الاستدلال أو التطبيق، بينما كانت فقرات الاختيار من متعدد التي تقيم المعرفة أقل صعوبة. كذلك قام المعلمون بمراجعة بعض فقرات العينة لتحسين وضوح المحتوى وإيجاز الصياغة والتوافق مع مواصفات التنسيق. بالنسبة لاختبار TIMSS، تشير النتائج إلى أن بعض الفقرات التي لم يتم أخذ عينات منها قد تحتاج إلى مراجعة أيضًا.
- **دراسة البريكيت (2008):** هدفت هذه الدراسة إلى استخدام نموذج راش في انتقاء بنود اختبار TIMSS للمهارات الرياضية لطلاب الصف الثاني المتوسط، إضافة إلى التحقق من مطابقة نموذج راش للبيانات التي جمعت من عينة الدراسة. واستخدمت الدراسة المنهج الوصفي وتمثلت الأداة في بناء اختبار من خلال انتقاء فقرات من اختبار TIMSS وترجمتها وإعدادها بالصورة النهائية، وتكونت عينة الدراسة من طلبة الصف الثاني

المتوسط، حيث تم تطبيق الاختبار على (670) طالبًا بمدينة تبوك. ولتحليل نتائج الدراسة استخدم الباحث برنامج SPSS لإيجاد الخصائص السيكومترية للفقرات وفق النظرية الكلاسيكية، واستخدام برنامج (Rumm-2010) لإيجاد معالم الفقرات وفق نموذج راش. ودلّت نتائج الدراسة على تمتع الاختبار بدرجات مقبولة من الصدق والثبات وفق النظرية الكلاسيكية بالإضافة إلى تمتع فقرات الاختبار بخصائص سيكومترية ملائمة من حيث مستوى الصعوبة والتمييز. وتمت مطابقة بيانات الاختبار إلى نموذج راش وأفرز التحليل (18) فقرة مطابقة لنموذج راش أحادي المعلمة.

التعليق على الدراسات السابقة:

تبيّن من خلال عرض الدراسات السابقة ذات الصلة الاهتمام الكبير بنظرية الاستجابة للفقرة؛ أنّ لها دقةً وموضوعية في بناء المقاييس والاختبارات وأنها تتلاءمُ مع تحليل فقرات الاختبار، والتحقق من صدقها وثباتها وقياس السمة الكامنة لدى الأفراد. وتتفق الدراسة الحالية مع الدراسات السابقة في استخدام نظرية الاستجابة للفقرة متمثلة في النموذج اللوجستي ثنائي المعلمة مثل: دراسات؛ دبوس، 2016؛ والسلامي، 2018؛ والعنزي، 2019؛ والمرواني، 2019.

أما بالنسبة لمنهج الدراسة، فتتفق الدراسة الحالية مع معظم الدراسات السابقة في المنهج الوصفي التحليلي.

كذلك، تباينت الأدوات المستخدمة في الدراسات السابقة، فبعضهم استخدم اختبار TIMSS كأداة للدراسة مثل: دراسة البريكيت، 2008؛ وShawn, 2012؛ ودراسة جراح، 2020. بينما استخدمت بعض الدراسات اختبار محكي المرجع كما في دراسة دبوس، 2016؛ ودراسة السلامي، 2018؛ واستخدمت دراسة العنزي، 2019؛ ودراسة المرواني، 2019 نماذج مختلفة من اختبار تحصيلي. بينما استخدم الرويلي (2019) بطاقة تحليل المحتوى أداة للدراسة. كما تتفق الدراسة الحالية مع دراسة البريكيت، 2008؛ Shawn, 2012؛ ودراسة جراح، 2020 في استخدام اختبار TIMSS أداة للدراسة.

أما بالنسبة للعينة، فكانت عينة أغلب الدراسات السابقة من فئة الطلاب، وتباينت من حيث المرحلة الدراسية والعمرية، مثل: دراسة البريكيت، 2008؛ Shawn, 2012؛ ودراسة دبوس، 2016؛ ودراسة السلامي، 2018؛ ودراسة العنزي، 2019؛ ودراسة المرواني، 2019؛ ودراسة جراح، 2020. أما دراسة الرويلي (2019)، فكانت عينة الدراسة من كتب الطالب، وكراسة التجارب العملية للصف الثاني المتوسط. أيضًا تتفق الدراسة الحالية في عينة الدراسة "طلاب الصف الثامن (الثاني المتوسط)" مع دراسة البريكيت، 2008؛ Shawn, 2012؛ ودراسة دبوس، 2016؛ ودراسة جراح، 2020.

كذلك، تتفق الدراسة الحالية مع بعض الدراسات السابقة في البرنامج الإحصائي (BILOG-MG3) المستخدم في تحليل فقرات الاختبار وتدريجها، مثل: دراسة السلامي، 2018؛ ودراسة العنزي، 2019؛ ودراسة المرواني، 2019. واختلفت مع دراسة البريكيت، 2008؛Shawn, 2012؛ ودراسة دبوس، 2016.

أجمعت معظم الدراسات السابقة على أن نماذج نظرية الاستجابة للفقرة تستند إلى افتراضات قوية ترتكز بدورها على أساس إحصائي سليم؛ مما يُثبت فعالية النماذج لبناء الاختبارات والمقاييس النفسية والتربوية.

وتتفرد الدراسة الحالية عن الدراسات السابقة التي تناولت اختبار TIMSS، أنّ الحالية استخدمت النموذج الثنائي المعلم الذي يأخذ بالاعتبار معامل الصعوبة والتمييز للفقرات، أيضًا اقتصرت على مادة العلوم للصف الثاني المتوسط.

وبالنسبة إلى أوجه استفادة الدراسة الحالية من الدراسات السابقة، فقد تمثَّلَت في: تحديد موضوع الدراسة، وإثراء وتدعيم الأدب النظري للدراسة الحالية، واختيار منهج الدراسة الحالية، وإجراءات الحصول على البيانات، وتفسير النتائج التي تم استخلاصها.

3- منهجية الدراسة وإجراءاتها.

منهج الدراسة:

نظرًا لطبيعة الدراسة وبياناتها استخدم الباحثان المنهج الوصفي التحليلي، وهو أسلوب من أساليب المنهج العلمي، ويعتمد على دراسة الظاهرة كما توجد في الواقع ويهتم بوصفها وصفًا دقيقًا من خلال التعبير النوعي الذي يصف الظاهرة، ويوضح خصائصها، أو التعبير الكمي الذي يعطي وصفًا رقميًا يوضح مقدار وحجم الظاهرة (عباس وآخرون، 2007).

مجتمع الدراسة:

تكون مجتمع الدراسة من جميع طلبة الصف الثاني المتوسط المشاركين في اختبار TIMSS 2019 في المملكة العربية السعودية والبالغ عددهم 5680 طالبًا وطالبة اختيروا بشكل عشوائي من 209 مدرسة (ETEC, 2020).

عينة الدراسة:

اقتصرت عينة الدراسة على الطلاب الذين استجابوا للنموذج الثالث من كراسات الاختبار وعددهم 405 طالبا وطالبة.

أداة الدراسة:

يشكل الاختبار الدولي TIMSS 2019 لمادة العلوم للصف الثاني المتوسط مصدرًا للبيانات، وهو الذي اعتمدت عليه هذه الدراسة.

إجراءات الحصول على البيانات وتحليلها:

تم الحصول على البيانات وفق الخطوات الآتية:

1- تنزيل البيانات من الموقع الإلكتروني التابع المنظمة الدولية للتقويم التربوي (IEA): https://timss2019.org/international-database/

2- اختيار ملف إجابات طلاب المملكة العربية السعودية على اختبار TIMSS في مادة العلوم التي تحمل رمز BSASAUM7.

3- فرز البيانات حسب رقم النموذج وعلى فقرات مادة العلوم، حيث اختير النموذج الثالث من كراسات الاختبار، وذلك لاحتوائه على أكبر عدد من فقرات الاختيار من متعدد والبالغ عددها (22) فقرة.

4- استبعاد الفقرات ذات الإجابة القصيرة والاقتصار فقط على الفقرات ذات الاختيار من متعدد.

5- تصحيح النموذج من خلال ملف T19_G8_Item Information.

وقبل الإجابة عن أسئلة الدراسة تم التحقق أولًا من افتراضات نظرية الاستجابة للفقرة، إضافة إلى التحقق من مدى ملاءمة النموذج اللوجستي ثنائي المعلمة للبيانات المستمدة من اختبار TIMSS 2019 لمادة العلوم.

أولًا: التحقق من افتراضات نظرية الاستجابة للفقرة كما يلي:

1- افتراض أحادية البعد Unidimensionality:

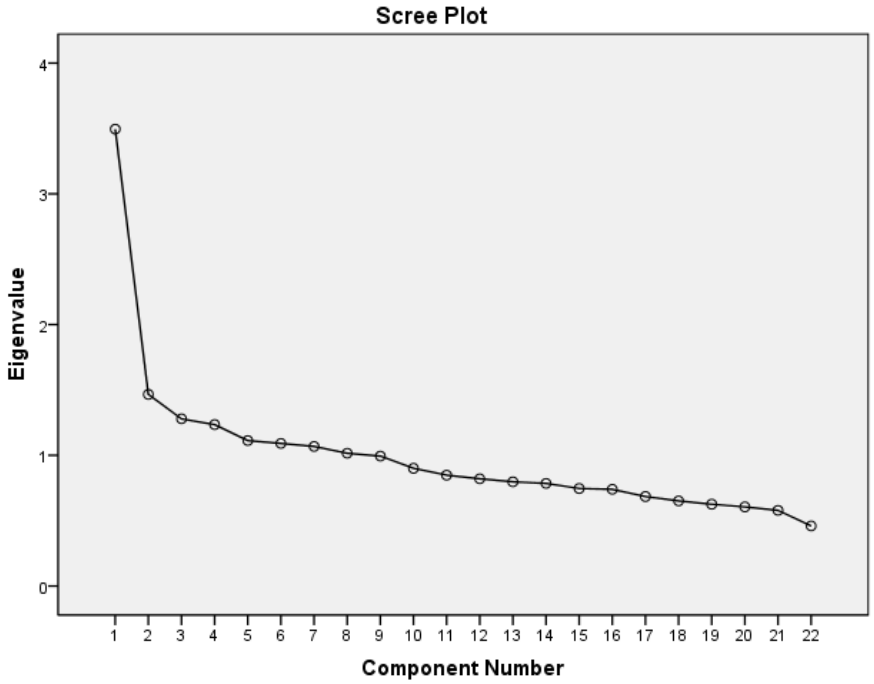
من افتراضات نظرية الاستجابة للفقرة أن تكون استجابة الفرد على الفقرة تقيس عاملًا واحدًا وليس عدة عوامل، أي تقيس سمة معينة بذاتها. وتم التحقق من افتراض أحادية البعد باستخدام برنامج SPSS، من خلال التحليل العاملي الاستكشافي (Factor Analysis) بطريقة المكونات الرئيسية (Principal Component) للنموذج الثالث من كراسات الاختبار. ويظهر في الجدول (3) قيمة الجذر الكامن (Eigen Value) ونسبة التباين المفسر (Explained Variance) وكذلك التباين المفسر التراكمي لكل عامل من العوامل:

جدول (3): قيم الجذور الكامنة ونسبة التباين المفسر والتراكمي لنموذج الاختبار.

العامل	الجذر الكامن	نسبة التباين المفسر	نسبة التباين المفسر التراكمي	الجذر الكامن الأول / الجذر الكامن الثاني
1	3.50	15.89	15.89	2.38
2	1.47	6.66	22.55	
3	1.28	5.81	28.37	
4	1.24	5.61	33.98	
5	1.11	5.06	39.04	
6	1.10	4.96	43.99	
7	1.07	4.85	48.85	
8	1.02	4.62	53.47	

يتضح من الجدول رقم (3) أن العامل الأول يفسر أكبر نسبة تباين في درجات الأفراد مقارنة ببقية العوامل. بالإضافة إلى أن ناتج قسمة الجذر الكامن للعامل الأول على الثاني في نموذج الاختبار؛ كان أكبر من (2). مما يدل على وجود سمة سائدة للاختبار، وبالتالي يمكن القول بتحقق افتراضية أحادية البعد للاختبار (Hattie, 1985).

ويلاحظ من خلال مخطط سكري (Scree Plot) للعوامل مع الجذور الكامنة كما يظهر في الشكل رقم (2) أن التمثيل البياني يظهر فيه العامل الأول ينفرد بقيمة جذر كامن مرتفعة نسبيًّا مقارنة مع العوامل الأخرى التي تظهر بجذور كامنة صغيرة ونسبيًا متقاربة، وهذا يعتبر أيضًا مؤشرًا على أحادية البعد.



شكل (2): مخطط سكري لقيم الجذور الكامنة للعوامل في نموذج الاختبار.

2- افتراض الاستقلال الموضعي Local Independence:

يتضمن هذا الافتراض أن إجابة الطالب عن الفقرة لا تتأثر بإجابته عن فقرة أخرى والعامل الوحيد المؤثر على احتمال الإجابة عن الفقرة هو مقدار القدرة (θ) وخصائص الفقرة. وهذا الافتراض مكافئ لافتراض أحادية البعد، أي أن تَحقُّق افتراض أحادية البُعد لفقرات الاختبار يُعَدّ دليلًا على تَحقُّق افتراض الاستقلال الموضعي (Hambleton & Swaminathan, 1985).

3- التحرر من السرعة في الأداء:

يمكن التأكد من أن الاختبار ليس اختبار سرعة من خلال فحص نسبة الطلاب الذين أكملوا الاختبار، وكذلك فحص الفقرات التي لم يجب عنها الطلاب. ووفقًا لما يراه هامبلتون وآخرون (Hambelton et al., 1991) فإنه إذا كان (75%) من الطلاب قد أكملوا الإجابة عن الاختبار، وإذا كان (80%) من فقرات الاختبار قد تمت الإجابة عنها من قبل الطلاب، فإن السرعة لن تعد في هذه الحالة عاملًا مهمًا في الأداء على الاختبار. وحيث إن نسبة الأفراد الذين أكملوا هذا الاختبار هي (100%)، وأن نسبة الفقرات التي تمت الإجابة عنها أعلى من (90%) فإن هذا يعني أن الاختبار يقيس القوة وليس السرعة.

ثانيًا- مطابقة الأفراد والفقرات للنموذج اللوجستي الثنائي المعلمة:

استخدم برنامج (Bilog-MG3)؛ للتأكد من مطابقة الأفراد (Person-Fit)، ومطابقة الفقرات (Item-Fit) للنموذج اللوجستي ثنائي المعلمة باستخدام اختبار مربع كاي (χ^2) عند مستوى دلالة ($\alpha = 0.01$)، وقد أظهرت نتائج التحليل عدم مطابقة استجابات فرد واحد للنموذج ثنائي المعلمة، حيث كانت القيم الاحتمالية له أقل من (0.01)؛ لذا تم حذف استجابته وأصبح حجم العينة في صورتها النهائية (404) أفراد.

وفيما يتعلق بفحص مطابقة الفقرات للنموذج المستخدم أُعيد التحليل باستخدام برنامج (Bilog-MG3) بعد حذف استجابة الفرد غير المطابق للنموذج بناءً على مؤشر مربع كاي (χ^2) عند مستوى دلالة ($\alpha = 0.01$)؛ وكانت النتائج كما يلي:

جدول (4): مطابقة فقرات الاختبار للنموذج الثنائي المعلمة

رقم الفقرة	قيمة χ^2	P-Value	رقم الفقرة	قيمة χ^2	P-Value
1	2.2	0.95	12	1.5	0.98
2	3.3	0.86	13	9	0.33
3	5.5	0.7	14	4.9	0.43
4	14.2	0.08	15	5.5	0.48
5	10	0.12	16	6.2	0.29
6	4.6	0. 80	17	14.9	0.02
7	9.5	0.3	18	8.8	0.36
8	3.5	0.83	19	11.9	0.06
9	5.3	0.62	20	4.6	0.8
10	10.2	0.25	21	7.5	0.48
11	4.9	0.77	22	5.8	0.67

يوضح الجدول (4) قيم اختبار مربع كاي χ^2 ودلالتها الإحصائية عند مستوى ($\alpha = 0.01$)، ويتضح مطابقة جميع فقرات الاختبار للنموذج ثنائي المعلمة، حيث كانت القيمة الاحتمالية أكبر من (0.01) لجميع الفقرات.

الأساليب الإحصائية:

من أجل الإجابة عن أسئلة الدراسة تم استخدام البرامج الإحصائية الآتية:

1- البرنامج الإحصائي spss لحساب التحليل العاملي الاستكشافي للتحقق من افتراض أحادية البعد للاختبار.

2- البرنامج الإحصائي Bilog-Mg3 لحساب الآتي:

- اختبار مربع كاي χ^2 لحسن المطابقة للكشف عن مدى مطابقة الأفراد والفقرات للنموذج اللوجستي ثنائي المعلمة.
- معامل صعوبة الفقرة والتمييز وفق النموذج ثنائي المعلمة.
- دقة تقدير معلمة صعوبة وتمييز الفقرة والقدرة وفق النموذج ثنائي المعلمة.
- ثبات الاختبار الإمبريقي.

4- نتائج الدراسة ومناقشتها.

- **النتائج المتعلقة بالإجابة عن السؤال الأول: ما دقة تقدير معلمة الصعوبة لفقرات اختبار TIMSS 2019 لمادة العلوم للصف الثاني المتوسط وفق النموذج اللوجستي ثنائي المعلمة؟**

للإجابة عن هذا السؤال تم استخدام برنامج (Bilog-Mg3) لحساب الخطأ المعياري في دقة تقدير معلمة الصعوبة لفقرات الاختبار TIMSS 2019، كما هو موضَّح في الجدول (5)، والذي يوضِّح كذلك تقديرات معلمة

الصعوبة وحساب المتوسط الحسابي والانحراف المعياري لمعلمة الصعوبة والخطأ المعياري لدقة تقدير معلمة الصعوبة.

جدول (5): تقديرات قيم معلمة الصعوبة والأخطاء المعيارية في تقدير معلمة الصعوبة.

رقم الفقرة	معلمة الصعوبة	الخطأ المعياري	رقم الفقرة	معلمة الصعوبة	الخطأ المعياري
1	-1.79	0.39	12	4.04	1.08
2	0.81	0.15	13	1.51	0.4
3	0.54	0.29	14	-1.19	0.15
4	2	0.457	15	-0.69	0.098
5	-1.35	0.218	16	-0.72	0.11
6	1.71	0.41	17	-0.67	0.11
7	0. 90	0.298	18	0.86	0.19
8	0.25	0.115	19	-0.59	0.11
9	-0.17	0.133	20	0.07	0.14
10	0.71	0.186	21	0.59	0.16
11	2.27	0.485	22	0.11	0.13
المتوسط الكلي	0.42	0.26	الانحراف الكلي	1.36	0.22

يتضح من الجدول (5) أن فقرات الاختبار تتّسم بدرجات متفاوتة من السهولة والصعوبة، حيث تراوحت قيم معلمة الصعوبة بين (1.79-) إلى (4.04) بمتوسط حسابي قدره (0.42)؛ ويذكر دي أيالا (De Ayala, / 2017 2009) أن الفقرات التي تقع قيمتها تحت الصفر (مثلا -2) توصف بأنها فقرات سهلة، وتوصف الفقرات التي تقع فوق الصفر بأنها فقرات صعبة (مثلًا +2)، أما الفقرات الواقعة نحو الصفر فتُعتبر فقرات متوسطة الصعوبة. وبناءً على ما تقدَّم نلاحظ أن المتوسط الحسابي لقيم معلمة الصعوبة يقترب من الصفر؛ وهذا يدل على أن فقرات الاختبار متوسطة الصعوبة. وبلغت قيمة الانحراف المعياري 1.36 والسبب في هذا التشتت ربما يعود إلى محتوى الفقرة (موضح في الملحق رقم 1 نص فقرات الاختبار ومجال المحتوى) حيث يتفاوت فهم الأفراد للفقرة، أيضا تختلف الفقرات في القدرات المطلوبة للإجابة عنها بعض الفقرات تحتاج قدرة منخفضة للإجابة عنها خاصة عندنا يكون معامل الصعوبة بالسالب، وبعض الفقرات تحتاج إلى قدرة متوسطة من الإفراد عندما تتمركز معاملات الصعوبة حول الصفر، وبعض الفقرات تحتاج إلى قدرة عالية للإجابة عنها خاصة عندما يكون معامل الصعوبة للفقرة موجب.

وقد حصلت الفقرة رقم (12) وهي في مجال (الفيزياء-تحولات الطاقة ونقلها) على أعلى قيمة للصعوبة (4.04)، ووفقًا لدراسة الرويلي (2019) فإن متطلب تحولات الطاقة ونقلها تَحقَّق بدرجة كبيرة في كتاب العلوم للصف الثاني المتوسط. وتتفق هذه النتيجة مع دراسة شاون (Shawn, 2012) التي توصَّلَت إلى أن أصعب الفقرات هي التي تُقيم مستوى الاستدلال والتطبيق. في حين حصلت الفقرة رقم (1) وهي في مجال (الكيمياء-التغير الكيميائي) على أقل قيمة للصعوبة (1.79-)، ووفقًا لدراسة الرويلي (2019) فإن متطلب التغير الكيميائي قد تحقَّق بدرجة ضعيفة في كتاب العلوم للصف الثاني المتوسط. وترى الباحثان أن حصول هذه الفقرة على أقل قيمة لمعلمة الصعوبة على الرغم من تَحقُّقها بدرجة ضعيفة في مقرر العلوم للصف الثاني المتوسط بسبب أن مقرر العلوم

للصف الأول المتوسط يحوي وحدة (طبيعية المادة) التي تتضمَّن فصل (المادة وتغيُّراتها). أيضا تتفق هذه النتيجة مع دراسة شاون (Shawn, 2012) التي توصَّلَت إلى أن الفقرات الأقل صعوبة هي التي تُقيم مستوى المعرفة.

أما ما يتعلق بالخطأ المعياري لتقدير معلمة الصعوبة فهو دلالة على دقة تقدير معلمة الصعوبة، وهو أدق من تقدير معلمة الصعوبة نفسها. ويتبيَّن من الجدول (5) أن الفقرة (15) تبلغ فيها قيمة الخطأ المعياري لتقدير معلمة الصعوبة (0.098)، وهي أقل قيمة في الجدول، وبالتالي تُعَدّ هذه الفقرة أكثر دقة في تقدير معلمة الصعوبة. وهي في مجال (علم الأرض- استخدام موارد الأرض وحفظها)، وقد تحقَّق هذا المطلب بدرجة كبيرة في مقرر العلوم للصف الثاني المتوسط كما أشار الرويلي (2019) في دراسته. أما الفقرة (12) فتبلغ فيها قيمة الخطأ المعياري لتقدير معلمة الصعوبة (1.08)، وهي أكبر قيمة في الجدول، وبالتالي تُعَدّ هذه الفقرة أقل دقة في تقدير معلمة الصعوبة.

- **النتائج المتعلقة بالإجابة عن السؤال الثاني: ما دقة تقدير معلمة التمييز لفقرات اختبار TIMSS 2019 لمادة العلوم للصف الثاني المتوسط وفق النموذج اللوجستي ثنائي المعلمة؟**

للإجابة عن هذا السؤال تم استخدام برنامج (Bilog-Mg3) لحساب الخطأ المعياري في دقة تقدير معلمة التمييز لفقرات الاختبار TIMSS 2019، كما هو موضَّح في الجدول (6) والذي يوضح كذلك تقديرات معلمة التمييز وحساب المتوسط الحسابي والانحراف المعياري لدقة تقدير معلمة التمييز.

جدول (6): تقديرات قيم معلمة التمييز والأخطاء المعيارية في تقدير معلمة التمييز.

رقم الفقرة	معلمة التمييز	الخطأ المعياري	رقم الفقرة	معلمة التمييز	الخطأ المعياري
1	0.61	0.14	12	0.39	0.1
2	1.06	0.16	13	0.46	0.1
3	0.38	0.09	14	1.43	0.25
4	0.51	0.11	15	1.5	0.24
5	0.95	0.18	16	1.36	0.23
6	0.49	0.11	17	1.39	0.21
7	0.45	0.1	18	0.77	0.13
8	1.07	0.16	19	1.26	0.18
9	0.82	0.15	20	0.77	0.14
10	0.73	0.13	21	0.79	0.13
11	0.55	0.12	22	0.88	0.14
المتوسط الكلي	0.85	0.15	الانحراف الكلي	0.36	0.05

يتّضح من الجدول (6) أن قيم معلمة التمييز تراوحت بين (0.38) إلى (1.50) بمتوسط حسابي قدره (0.85)؛ ويذكر دي أيالا (De Ayala, 2009 /2017) أن الفقرات التي تتراوح قيمتها ما بين (0.80) إلى (2.5) تُعتبر جيدة التمييز. وبناءً على ما تقدَّم نلاحظ أن المتوسط الحسابي لقيم معلمة التمييز (0.85) أكبر من (0.80)؛ وهذا يدل على أن فقرات الاختبار جيدة التمييز.

أما ما يتعلق بالخطأ المعياري لتقدير معلمة التمييز فيتضح من الجدول (6) أن الفقرة (3) تبلغ فيها قيمة الخطأ المعياري لتقدير معلمة التمييز (0.09)، وهي أقل قيمة في الجدول، وبالتالي تُعَدّ هذه الفقرة أكثر دقة في تقدير

معلمة التمييز. وهي في مجال (الفيزياء-الضوء والصوت)، وقد تَحقَّق هذا المطلب بدرجة كبيرة في مقرر العلوم للصف الثاني المتوسط كما أشار الرويلي (2019) في دراسته.

- **النتائج المتعلقة بالإجابة عن السؤال الثالث: ما دقة تقدير القدرة لأداء الطلاب المشاركين على اختبار TIMSS 2019 لمادة العلوم للصف الثاني المتوسط وفق النموذج اللوجستي ثنائي المعلمة؟**

للإجابة عن هذا السؤال تم تقدير القدرة لأداء الطلاب المشاركين على اختبار TIMSS 2019 وتحديد الخطأ المعياري في تقدير القدرة باستخدام برنامج (Bilog-Mg3) مخرجات (PHASE-3)، وتراوحت قيم تقدير القدرة للأفراد في الاختبار بين (3.61-) إلى (4) بمتوسط حسابي قدره (0.03) وانحراف معياري (1.26)، وأقل خطأ معياري بلغت قيمته (0.49) عند تقدير القدرة (0.58-). ويُبيِّن الجدول (7) المتوسط الحسابي والانحراف المعياري لتقدير القدرة والخطأ المعياري في تقديرها.

جدول (7): قيم المتوسط الحسابي والانحراف المعياري لتقدير قدرة الأفراد والخطأ المعياري في تقدير القدرة.

المتوسط/ الانحراف	تقدير القدرة	الخطأ المعياري
المتوسط الحسابي	0.03	0.49
الانحراف المعياري	1.26	0.13

وذكر دبوس (2016) في دراسته أنه إذا كان متوسط تقديرات القدرة للأفراد أكبر من متوسط تقديرات صعوبة الفقرات يمكن القول إن الفقرات كانت في متناول قدرات أفراد العينة، أما إذا كان متوسط تقديرات صعوبة الفقرات أكبر من متوسط تقديرات القدرة للأفراد يمكن القول إن الفقرات كانت صعبة إلى حدٍّ ما. ويُلاحَظ أنّ متوسط صعوبة الفقرات الذي بلغ (0.42) أكبر من متوسط تقدير قدرة الأفراد الذي بلغ (0.03)؛ مما يدل على أن فقرات اختبار TIMSS 2019 كانت صعبة إلى حدٍّ ما ولم تكن في متناول قدرات أفراد العينة. أما قيمة الانحراف المعياري التي بلغت 1.26 والسبب في التشتت يرجع إلى اختلاف قدرات الأفراد وفهمهم لمحتوى الفقرات ومراعاة لمبدأ الفروق الفردية بين المفحوصين.

وتم حساب تكرارات الأفراد حسب درجات الاختبار، ومتوسط تقدير القدرة، ومتوسط الخطأ المعياري لتقدير قدرة الأفراد، كما هو مُوَضَّح في الجدول (8).

جدول (8): تكرارات الأفراد ومتوسط تقدير القدرة ومتوسط الخطأ المعياري.

الدرجة	التكرارات	متوسط تقدير القدرة	متوسط الخطأ المعياري
1	1	-3.61	1.26
2	4	-3.01	0.98
3	5	-2.03	0.67
4	15	-1.74	0.6
5	22	-1.47	0.56
6	21	-1.28	0.54
7	30	-.95	0.51
8	40	-.63	0.5
9	35	-.41	0.5
10	39	-.13	0.51

الدرجة	التكرارات	متوسط تقدير القدرة	متوسط الخطأ المعياري
11	36	0.12	0.53
12	31	0.28	0.54
13	28	0.69	0.58
14	24	0.97	0.62
15	29	1.27	0.66
16	19	1.67	0.73
17	7	2.21	0.84
18	8	2.77	0.97
19	6	3.27	1.11
20	4	3.96	1.27

يتضح من الجدول (8) أن مدى الدرجات يمتدّ من (1) إلى (20) درجة، ويُلاحَظ أن الطالب الذي قدرته (-3.61) حصل على درجة (1) بالاختبار، ويعتبر من فئة الأفراد ذوي القدرة المنخفضة الذين تمتدّ القدرة لديهم من (-3.61) إلى (-1.49). والطالب الذي قدرته (0.12) حصل على درجة (11) بالاختبار، ويعتبر من فئة الأفراد ذوي القدرة المتوسطة الذين تمتد القدرة لديهم من (-1.5) إلى (1.5). أما الطالب الذي قدرته (3.96) فحصل على درجة (20) بالاختبار، ويُعتبر من فئة الأفراد ذوي القدرة المرتفعة الذين تمتد القدرة لديهم من (1.51، 4).

- **النتائج المتعلقة بالإجابة عن السؤال الرابع: ما معاملات ثبات اختبار TIMSS 2019: (الثبات الإمبريقي، وثبات الأفراد، وثبات الفقرات) لمادة العلوم للصف الثاني المتوسط وفق النموذج اللوجستي ثنائي المعلمة؟**

أ- الثبات الإمبريقي:

تم حساب معامل الثبات الإمبريقي باستخدام برنامج BIOLG-MG3 من مخرجات PHASE-3، حيث بلغت قيمة معامل الثبات الإمبريقي (0.78)، وهي قيمة مقبولة وتُعَدّ مؤشرًا جيدًا للاختبار.

ب- ثبات الأفراد:

بلغت قيمة المتوسط الحسابي لتقدير قدرات الأفراد (0.0324) بانحراف معياري (1.256)، كما بلغت قيمة المتوسط الحسابي للخطأ المعياري في تقدير القدرة (0.5873) بانحراف معياري (0.1336).

معامل الفصل بين الأفراد G_p = الانحراف المعياري لقيم تقدير القدرة / المتوسط الحسابي للخطأ المعياري في تقدير القدرة = $\frac{1.26}{0.59} = 2.14$

معامل الثبات للأفراد:R_p

$$R_p = \frac{G^2}{1+G^2} = 0.82$$

ونلاحظ أن معامل الفصل بين الأفراد بلغت قيمته (2.14)، وهي أكبر من (2)؛ مما يدل على أن عينة الأفراد كافية للفصل بين الفقرات، ومعامل الثبات للأفراد يساوي (0.82)، وهي قيمة مقبولة، وتدل على أن عينة الأفراد كانت كافية للفصل بين الفقرات.

ج- معامل ثبات الفقرات:

بلغت قيمة المتوسط الحسابي لتقدير صعوبة الفقرات (0.42) بانحراف معياري (1.36)، كما بلغت قيمة المتوسط الحسابي للخطأ المعياري في تقدير القدرة (0.26) بانحراف معياري (0.22).

معامل الفصل بين الفقرات G_i= $\frac{\text{الانحراف المعياري لقيم معلم الصعوبة للفقرات}}{\text{المتوسط الحسابي للخطأ المعياري في تقدير معلم الصعوبة}}$ = $\frac{1.36}{0.26}$ = 5.23

معامل الثبات للفقرات R_i:

$$R_i = \frac{G^2}{1+G^2} = 0.96$$

ونلاحظ أن معامل الفصل بين الفقرات بلغت قيمته (5. 23) وهي أكبر من (2)؛ مما يدل على أن عينة الفقرات كافية للفصل بين الأفراد. وبلغت قيمة معامل الثبات للفقرات (0.96)، وهي قيمة مرتفعة تدل على أن فقرات الاختبار كانت كافية للتمييز بين قدرات الأفراد.

مناقشة النتائج:

هدفت هذه الدراسة إلى الكشف عن دقة تقدير معلمة الصعوبة والتمييز للفقرة والقدرة لاختبار TIMSS لمادة العلوم للصف الثاني المتوسط وفق النموذج الثنائي المعلم لنظرية الاستجابة للفقرة باستخدام برنامج BILOG-MG3.

مناقشة النتائج المتعلقة بالسؤال الأول حيث أظهرت النتائج إن الفقرات تتمتع بمؤشرات متوسطة الصعوبة حيث تراوحت قيم صعوبة الفقرات ما بين (1.79-) إلى (4.04) مما يشير إلى أن الاختبار يصنف بشكل عام على أنه متوسط الصعوبة، وتتفق هذه النتيجة مع دراسة البريكيت (2008)، ودبوس (2016)، والسلامي (2018) وأيضا تتفق هذه النتيجة مع دراسة شاون (Shawn, 2012) التي توصَّلَت إلى أن الفقرات الأقل صعوبة هي التي تُقيم مستوى المعرفة. وأن أصعب الفقرات هي التي تُقيم مستوى الاستدلال والتطبيق. واختلفت هذه النتيجة مع دراسة الرويلي (2019) حيث أن محتوى الفقرة الأكثر صعوبة قد تَحقَّق بدرجة كبيرة في كتاب العلوم للصف الثاني المتوسط، وأن محتوى الفقرة الأقل صعوبة قد تحقَّق بدرجة ضعيفة في كتاب العلوم للصف الثاني المتوسط.

مناقشة النتائج المتعلقة بالسؤال الثاني: حيث أظهرت النتائج أنّ قِيَم معلمة تمييز الفقرات تراوحت ما بين (0.38) إلى (1.50)، بمتوسط مقداره (0.85)، أي إن الفقرات تتمتع بمؤشرات تمييز جيدة، وتتفق هذه النتيجة مع دراسة البريكيت (2008)، ودبوس (2016)، والسلامي (2018) والرويلي (2019).

مناقشة النتائج المتعلقة بالسؤال الثالث: حيث أظهرت النتائج أن قيم تقدير القدرة للأفراد تراوحت ما بين (3.61-) إلى (4)، بمتوسط حسابي مقداره (0.03). ويشير إلى أن فقرات اختبار TIMSS 2019 كانت صعبة إلى حدٍّ ما ولم تكن في متناول قدرات أفراد العينة. واختلفت هذه النتيجة مع دراسة البريكيت (2008) ودبوس (2016) حيث أشارت نتائج دراستهم على مناسبة فقرات الاختبار لقدرات العينة المختارة بمعنى أن الفقرات كانت في متناول أفراد العينة.

مناقشة النتائج المتعلقة بالسؤال الرابع: أظهرت النتائج أنّ قيمة معامل الثبات الإمبريقي تساوي (0.78)، أما معامل الثبات للأفراد فقد بلغت قيمته (0.82)، وبلغت قيمة معامل الثبات للفقرات (0.96)، وهذه القيم لمعاملات الثبات تعتبر قيم مقبولة وتُعَدّ مؤشرًا جيدًا للاختبار. واتفقت هذه النتيجة مع دراسة البريكيت (2008) والعنزي (2019) والمرواني (2019).

توصيات الدراسة ومقترحاتها:

بناءً على ما أسفرت عنه نتائج الدراسة، يوصي الباحثان ويقترحان ما يلي:

1- إقامة إدارة التدريب التربوي والابتعاث الدورات التدريبية وورش العمل للمعلمين والمعلمات في بناء فقرات الاختبارات وفق مستويات اختبار TIMSS.

2- إشراك المعلمين في ورش العمل التي تقام من قِبَل وزارة التعليم لمناقشة نتائج TIMSS.

3- تحقيق التكامل بين درجة توافر متطلبات TIMSS في كتب العلوم وبين إلمام المعلمين والمعلمات بها وتدريبهم عليها، وتحسين مهاراتهم وقدرتهم على تنميتها لدى الطلاب.

4- تدريب المعلمين والمعلمات على تصميم أنشطة تعليمية تتوافق مع متطلبات TIMSSتحت إشراف إدارات التدريب التربوي.

5- استخدام إدارة الاختبارات والقبول نماذج الاستجابة للفقرة في بناء أدوات واختبارات جديدة والتحقق من الخصائص السيكومترية لها.

6- كما تقترح الباحثان ما يلي:

1. تحليل بيانات الدراسة وفق النموذج الأحادي والثلاثي ومقارنتها بنتائج هذه الدراسة.
2. دراسة تقدير معالم الفقرة والقدرة للأفراد المشاركين في اختبار TIMSS لمادة الرياضيات وفق نظرية الاستجابة للفقرة.
3. دراسة مقارنة لتقدير معالم الفقرة والقدرة للأفراد المشاركين في اختبار TIMSS بين المملكة العربية السعودية وبعض الدول المتقدِّمة.
4. دراسة بناء اختبار محكي يقيس المهارات العلمية لمادة العلوم للمرحلة المتوسطة.
5. دراسة بناء مقياس الكفايات العلمية لمتطلبات TIMSS لمعلمي العلوم للمرحلة المتوسطة.
6. إجراء المزيد من الأبحاث عن اختبار TIMSS باستخدام النماذج الأخرى لنظرية الاستجابة للفقرة.

قائمة المراجع

أولاً- المراجع بالعربية:

- أبو عواد، فريال محمد عثمان. (2018). استقصاء تقديرات معالم الفقرات والقدرة ودالة المعلومات لاختبار القدرات المعرفية باستخدام النموذج اللوجستي ثلاثي المعلمة. مجلة دراسات نفسية وتربوية، جامعة قاصدي مرباح، 11 (1)، 1-17.
- البريكيت، ممدوح سليمان. (2008). استخدام نموذج راش في انتقاء بنود اختبار تيمس (TIMSS) للمهارات الرياضية لطلاب الصف الثاني المتوسط [رسالة ماجستير غير منشورة]. جامعة الملك سعود.
- الجراح، بندر نواف. (2020). الكشف عن أنماط الاستجابة في اختباري العلوم والرياضيات الدوليين لبيانات (TIMSS, 2015) لدى عينة من طلبة الدولة السعودية ودولة سنغافورة باستخدام مؤشرات مطابقة الشخص. مجلة العلوم التربوية: جامعة الملك سعود - كلية التربية، 32(2)، 299-320.
- حسان، محمود عبد اللطيف محمود. (2013). تقويم محتوى المناهج في ضوء المتطلبات المعرفية لمشروع التيمز TIMSS. مكتبة القانون والاقتصاد.

- دبوس، محمد طالب. (2016). استخدام نظرية الاستجابة للفقرة في بناء فقرات اختبار محكي المرجع في الرياضيات بفقرات ثنائية التدريج ومتعددة التدريج وفق النموذج اللوجستي ثنائي المعلم. مجلة جامعة النجاح الوطنية، 30 (7)، 1453-1480.
- الدرابسه، رياض أحمد صالح؛ وعودة، أحمد سليمان. (2012). أثر طريقة تقدير القدرة وطريقة التعامل مع القيم المفقودة على دقة تقدير معالم الفقرات والأفراد [رسالة دكتوراه، جامعة اليرموك]. قاعدة معلومات دار المنظومة. https://search-mandumah-com.sdl.idm.oclc.org/Record/739433
- دي إيالا، ر.ج (2017). النظرية والتطبيق في نظرية الاستجابة للفقرة. (عبد الله الكيلاني وإسماعيل البرصان، ترجمة). دار جامعة الملك سعود للنشر. (العمل الأصلي نشر في 2009).
- الرويلي، فارس. (2019). تحليل محتوى مقرر العلوم للصف الثاني المتوسط في المملكة العربية السعودية في ضوء متطلبات دراسة التوجهات الدولية للرياضيات والعلوم TIMSS 2015 [ماجستير غير منشورة]. جامعة الجوف.
- السلامي، عبد الله محمد. (2018). بناء اختبار محكي المرجع في الرياضيات باستخدام نظرية الاستجابة للمفردة وفق النموذج اللوجستي ثنائي المعلم لطلاب الصف الخامس (المرحلة الابتدائية-الصف الخامس). دراسات في الإرشاد النفسي والتربوي: جامعة أسيوط - كلية التربية - مركز الإرشاد النفسي والتربوي، 1(1)، 111-137.
- السليماني، نايف محمد (2016) تقويم محتوى مناهج العلوم بالمرحلتين الابتدائية والمتوسطة في ضوء متطلبات دراسة التوجهات الدولية السادسة للرياضيات والعلوم (TIMSS) [رسالة دكتوراه غير منشورة]. جامعة الملك خالد.
- الشمراني، صالح؛ الشمراني، سعيد؛ البرصان، اسماعيل؛ والدوراني، بكيل. (2016). اضاءات حول نتائج دول الخليج في دراسة التوجهات الدولية للعلوم والرياضيات TIMSS 2015. جامعة الملك سعود.
- الشيخي، هاشم سعيد. (2012). إستراتيجية مقترحة لتحسين مستوى تحصيل طلبة المملكة العربية السعودية في الرياضيات في المسابقة الدولية (TIMSS). الجامعة الأردنية- عمادة البحث العلمي، 39 (1)، 33-60.
- عباس، محمد؛ نوفل، محمد؛ العبسي، محمد؛ وأبو عواد، فريال. (2007). مدخل إلى مناهج البحث في التربية وعلم النفس. دار المسيرة للنشر والتوزيع.
- علام، صلاح الدين محمود. (2005). نماذج الاستجابة للمفردة الاختبارية أحادية البعد ومتعددة الأبعاد وتطبيقاتها في القياس النفسي والتربوي. دار الفكر العربي.
- علام، صلاح الدين محمود. (2011). القياس والتقويم التربوي والنفسي: أساسياته وتطبيقاته وتوجهاته المعاصرة (ط.5). دار الفكر العربي.
- العنزي، محمد عويض؛ وعبدالعال، صبري محمد. (2019). أثر شكل فقرة الاختيار من متعدد والصواب والخطأ على الخصائص السيكومترية وفق نظرية استجابة الفقرة لاختبار مادة الحاسب الآلي للصف الأول ثانوي مدينة تبوك. المجلة الدولية للدراسات التربوية والنفسية، 6(1)، 1-17. https://doi.org/DOI:10.31559/EPS2019.6.1.1
- عودة، أحمد سليمان. (2010). القياس والتقويم في العملية التدريسية (ط.4). دار الأمل.

- الفهيدي، هذال عبيد. (1433هـ). تقويم محتوى مقررات العلوم المطورة بالمرحلة الابتدائية في المملكة العربية السعودية في ضوء متطلبات دراسة التوجهات الدولية للرياضيات والعلوم (TIMSS,2011) [رسالة دكتوراة غير منشورة] جامعة أم القرى.
- كروكر، ليندا ؛ وألجينا، جيمس (2017). مدخل إلى نظرية القياس التقليدية والمعاصرة. (هند الحموري وزينات دعنا، ترجمة؛). دار الفكر (العمل الأصلي نشر في 2008).
- المرواني، أشواق ضيف الله؛ وسليمان، شاهر خالد. (2019). أثر موقع البديل الصحيح في اختبار اختيار من متعدد على دقة تقدير معالم الفقرة وفق النموذج ثنائي المعلمة لنظرية الاستجابة للفقرة. المجلة الدولية للدراسات التربوية والنفسية، 6(1)، 66-88. https://doi.org/DOI:10.31559/EPS2019.6.1.5
- هيئة تقويم التعليم والتدريب ETEC. (1439). نبذة حول الاختبارات الدولية TIMSS PIRLS PISA TALIS. https://www.etec.gov.sa/ar/Media/Publications/%d8%a7%d9%84%d9%86%d8%b4%d8%b1%d8%a7%d8%aa%201439
- هيئة تقويم التعليم والتدريب ETEC. (2020). تقرير تيمز 2019 نظرة أولية في تحصيل طلبة الصفين الرابع والثاني المتوسط في الرياضيات والعلوم بالمملكة العربية السعودية في سياق دولي. https://www.etec.gov.sa/ar/ Researchers/Research-Studies/Documents/TIMSS%202019.pdf
- هيئة تقويم التعليم والتدريب ETEC. (د.ت.) الدليل الارشادي للاختبارات الدولية TIMSS 2019. https://www.etec.gov.sa/ar/productsandservices/Qiyas/internationaltests/TIMSS/Documents/TIMSS_International_Testing_Guides.pdf
- وزارة التعليم. (د.ت.) دليل اختبارات TIMSS مادتي العلوم والرياضيات. https://edu.moe.gov.sa/Mokhwah/ Departments/mdir/jodh/Documents

ثانياً- المراجع بالإنجليزية:

- Baker, Frank b. (2001). The Basics of Item Response Theory. (2nd.ed). ERIC Clearinghouse on Assessment and Evaluation.
- Hambleton, R., & Swaminathan, H. (1985). Item Response Theory: Principles and Applications. Kluwer· Nijh off Publishing.
- Hattie, J (1985). Methodology Review: Assessing Unidimensionality of Tests and Items. Psychological Measurement, 9(2), 139-164.
- Mullis, Ina. Martin, Michael. (2017). TIMSS 2019 assessment frameworks. Boston College, TIMSS& PIRLS international study center.
- Mullis, Ina. Martin, Michael. Foy, Pierre. Kelly, Dana and Fishbein, Bethany. (2020). TIMSS 2019 International Results in Mathematics and Science. Boston College, TIMSS& PIRLS international study center.
- Shawn M. Glynn. (2012). International Assessment: A Rasch Model and Teachers' Evaluation of TIMSS Science Achievement Items. Journal of Research in Science Teaching, 49(10), 1321–1344.

- Von Davier, M., Gonzalez, E., & Schulz, W. (2020). Ensuring validity in international comparisons using state-of-the-art psychometric methodologies. In H. Wagemaker (Ed.), Reliability and validity of international large-scale assessments (pp. 187-219).

ملحق (1) ترجمة لفقرات النموذج الثالث لاختبار TIMSS2019 لمادة العلوم

الرقم	نص الفقرة	مجال محتوى الفقرة
1	عملية امتصاص الطاقة الضوئية	تغير كيميائي
2	نماذج ثاني أكسيد الكربون	تكوين المادة
3	صوت تستعمله الحيوانات	الضوء والصوت
4	تسخين الماء على الموقد	تحويل الطاقة ونقلها
5	تغيير الطاقة في الانزلاق	تحويل الطاقة ونقلها
6	مصدر المياه لتحلية المياه	هيكل الأرض والميزات الفيزيائية
7	سلسلة جبال على حافة المحيط	عمليات الأرض ودوراتها وتاريخها
8	خصائص غير عادية من النضناض	الخصائص والعمليات الحياتية للكائنات
9	أصل خلايا جديدة مع نمو الكائنات الحية	الخلايا ووظائفها
10	تحديد الحيوانات المفترسة في الرسم التخطيطي	النظم البيئية
11	كيف يدخل الكربون في الخشب	النظم البيئية
12	نقل الطاقة أثناء تغير درجة الحرارة	تحويل الطاقة ونقلها
13	أين يجب أن تقف ميغان	الضوء والصوت
14	تقليل أو إعادة تدوير - العناصر التي تحتوي على القليل من التعبئة والتغليف	موارد الأرض واستخدامها والمحافظة عليها
15	إعادة التدوير - إعادة تعبئة زجاجات المياه البلاستيكية	موارد الأرض واستخدامها والمحافظة عليها
16	تقليل البطاريات أو إعادة تدويرها - إعادة شحن البطاريات	موارد الأرض واستخدامها والمحافظة عليها
17	قم بتقليل أو إعادة تدوير - كرتون جديد من الصناديق القديمة	موارد الأرض واستخدامها والمحافظة عليها
18	التقليل أو إعادة التدوير - التسميد	موارد الأرض واستخدامها والمحافظة عليها
19	التقليل أو إعادة التدوير - باستخدام ألواح الطاقة الشمسية	موارد الأرض واستخدامها والمحافظة عليها
20	تقليل أو إعادة تدوير - النوافذ من الزجاجات	موارد الأرض واستخدامها والمحافظة عليها
21	لماذا تتعرض المنطقة للفيضانات في كثير من الأحيان بعد إزالة الغابات	موارد الأرض واستخدامها والمحافظة عليها
22	كيف تنتج الشمس الضوء	الأرض في النظام الشمسي والكون