Journal of Educational and Psychological Sciences
Volume (4), Issue (14) : 30 Apr 2020
P: 75 - 99
ISSN: 2522-3399
مجلة العلوم التربوية والنفسية
المجلد (4)، العدد (14) : 30 إبريل 2020 م
ص: 75 - 99

Using IRT in Estimate Psychometric Characteristics of the Central Science Test for the Third Intermediate Grade in Al-Madinah Al-Munawara's Schools

Easa Salamah S. Alhajory
Faculty of Education || Taibah University || KSA

Abstract: The aim of this study is to identify the psychometric characteristics of the central science test for the third grade intermediate , which was applied in 2017-2018, it consisted of (40) multiple-choice items. The analyses was according to item response theory.

For achieving the objective of the study, (1040) students' answers forms were collected, the results indicated the availability of the assumptions of (IRT) in the test, the three-parameter model was used, it was the most appropriate for the data.

The psychometric properties of the test were estimated using R software by the (ltm) package.

The results indicated that the most of items were characterized by high level of difficulty, and some items contained a high guessing. In general, the values of the items parameters were inconsistency, according to the criteria presented by the measurement literature. Values of the difficulty parameters ranged between (-0.81 to 2.22), values of the discrimination parameters ranged between (0.28 and 3.81), and guessing parameter ranged from (0.00 to 0.50.(

The test presents the greatest amount of information at above average ability level

There were some recommendations that may improve the quality of the psychometric characteristics of this test, making it an objective and accurate measurement tool in estimating examinees' abilities by using (IRT).

Keywords: (IRT), Central Tests, Achievement Test, Modern Theory of Measurement.

استخدام نظرية الاستجابة للمفردة في تقدير الخصائص السيكومترية للاختبار المركزي لمادة العلوم للصف الثالث المتوسط بالمدينة المنورة

عيسى بن سلامه بن صبيان الحجوري
كلية التربية || جامعة طيبة || المملكة العربية السعودية

الملخص: هدفت الدراسة الحالية إلى التعرف على الخصائص السيكومترية وفق نظرية الاستجابة للمفردة، للاختبار المركزي لمادة العلوم للصف الثالث المتوسط، والذي تم تطبيقه في الفصل الثاني من العام الدراسي 1438-1439هـ في المدينة المنورة، واحتوى على (40) فقرة خيار من متعدد، ولتحقيق هدف الدراسة تم جمع نماذج إجابات (1040) طالبًا، وجرى تقدير الخصائص السيكومترية للاختبار باستخدام برنامج R من خلال حزمة (ltm) وفق النموذج الثلاثي البارامتر والذي كان الأنسب لبيانات الدراسة، وأشارت النتائج بتمتع الاختبار بمستوى مرتفع من الصعوبة، ووجود بعض الفقرات التي احتوت على تخمين عالي، وبشكل عام ظهر هناك عدم اتساق بين قيم معالم المفردة للفقرات وفق المعايير التي أوردتها أدبيات القياس، حيث تراوحت قيم معالم الصعوبة بين (-0.81 إلى 2.22) بمتوسط قدره (0.59)، كما كانت قيم معْلَم التمييز بين (0.28 إلى 3.81) بمتوسط قدره (1.5)، أما معْلَم التخمين فتراوحت قيمته من ((0.00 إلى 0.50) وبمتوسط قدره (0.22). وأعطى الاختبار أكبر كمية معلومات عند مستويات القدرة فوق المتوسط. وفي ضوء النتائج تم تقديم بعض التوصيات التي قد تحسِّن من جودة الخصائص السيكومترية لهذا الاختبار، مما تجعل منه أداة قياس موضوعية ودقيقة في تقدير قدرات الأفراد وفق نظرية الاستجابة للمفردة أحادية البعد.

الكلمات المفتاحية: نظرية الاستجابة للمفردة، الاختبارات المركزية، الاختبار التحصيلي، النظرية الحديثة في القياس.

1- مقدمة:

يقاس تقدم الأمم بقوة نظامها التربوي، ومدى تحقيقه لنواتج التعلم المستهدفة، وفقاً لمعايير جودة التعليم، واعتماده على منظومة قياس وتقويم متطورة تسهم في اتخاذ قرارات موضوعية مرتكزة على أسس علمية، ذلك باعتبار أن التقويم جزء لا يتجزأ من منظومة التعليم والتعلم، لذلك فإن تحسين مستويات إتقان المتعلمين لنواتج التعلم المحددة، لا يمكن تحقيقها إلا من خلال عملية إصلاح شامل لمنظومة القياس والتقويم التربوي، تتناول فلسفته وأغراضه وأساليبه وتقنياته، ومدى تكامله مع عناصر العملية التعليمية الأخرى، فتقديم رؤية تكاملية للتقويم في إطار التعليم العام، يتطلب التفريق بين المفاهيم في هذا المجال بالإضافة إلى معرفة الشروط الواجب توافرها في أدوات التقويم، ومن ثم الوصول إلى نموذج يحتذى به في الحكم الموضوعي على الأداء وبشكل يسهم في تحسين التعليم، والارتقاء بجودة مخرجاته (القفاص، 2011، 5).

ويعد القياس والتقويم عنصران أساسيان في العملية التعليمية والتربوية، فهما يؤثران تأثيرًا كبيرًا في تقدم هذه العملية وتطوير مكوناتها ورفع كفاءتها، وتعتبر الاختبارات التربوية من أهم أدوات القياس، والتي تساعد في إعطاء قيمة عددية لأداء الفرد، حيث تزود المؤسسات التعليمية والتربوية ببيانات كمية تتيح للمتخصصين فهم الظاهرة التربوية، كما تقوم بتزويد القائمين على العملية التعليمية بمقدار التقدم في مستوى التحصيل الدراسي للطلاب، وتظهر مدى تحقيقهم للأهداف التعليمية، وبالتالي اتخاذ قرارات يعتمد على سلامتها وصحة نوع ودقة المعلومات والأساليب المستخدمة في تفسير وتحليل النتائج التي تزودنا بها الاختبارات (العنزي، 2018).

وتعتبر الاختبارات من أكثر أدوات القياس شيوعًا واستخدامًا في المؤسسات التعليمية، ليس فقط من أكثرها شيوعًا فحسب، بل من أهمها، وذلك لما تقدمه من معلومات واسعة ومؤشرات دالة ذات علاقة بالسمة المراد قياسها لدى الفرد (القضاة، 2010).

وتمثل الاختبارات التحصيلية، عنصرًا حيويًا في العملية التعليمية، لما لها من أثر فعال في العملية التربوية بأكملها، إذ عن طريقها تصدر الأحكام القيمية على مدى نجاح العملية التربوية أو فشلها في تحقيق الأهداف المرجوة. ولهذا فلا غرابة أن تعنى المؤسسات التربوية، عناية خاصة بالاختبارات التحصيلية وتوليها اهتمامًا بالغًا يكاد يضفي على مكونات العملية التربوية الأخرى حتى أصبحت الاختبارات التحصيلية تحتل المكانة الأولى في المدارس والمعاهد والجامعات (الدوسري، 2004).

وغالبًا ما تلعب الخبرة العملية دورًا هامًا في صياغة فقرات الاختبار، حيث يستفيد واضع الاختبارات من الخلل أو العيوب التي تواجهه، وبالتالي تكون أحكام الخبراء مبنية على المنطق ((Weiten,1984.

فيشير المالكي (2015) إلى أن معلّم ي العلوم في المرحلة المتوسطة يلتزمون ب (60%) من قواعد بناء أسئلة الاختيار من متعدد ذا الأربعة بدائل، والسبب في ذلك هو أن معظم المعلمين مضغوطين بالوقت ومتطلبات وظيفتهم، وعدم معرفة بعضهم بكيفية الصياغة الصحيحة للأسئلة، ونتيجة لذلك فإن اختباراتهم تكون ذات فقرات ضعيفة الصياغة.

وذلك يكون مؤشرًا لدى بعض الطلاب لما يسمى بالحكمة في الإجابة، مما قد يؤثر سلبًا على خصائص الفقرة والاختبار، والذي يؤدي إلى تضخيم درجات الطلاب بشكل مضلل، لان فقرات الاختبار بدلا من أن تقيس السمة المراد قياسها، أصبحت تقيس سمه أخرى دخيلة، وهي حكمة الاختبار أو المثابرة واستغلال الخلل للحصول على درجات مرتفعة على الاختبار (Sarnaki,1979).

لذا عند تصميم الاختبارات خاصة الاختبارات التحصيلية يتم استخدام إحدى طريقتي القياس الشهيرتين وهما: نظرية الاستجابة للمفردة (Item Response Theory) والنظرية الكلاسيكية في القياس ((Classical Test Theory التي استخدمت بكثرة في القرن العشرين لمحاولة تلافي القصور في أداة الاختبار (Traub,1997).

فنظرية الاستجابة للمفردة تعد من التطورات الحديثة في مجال القياس والتقويم التربوي والنفسي، حيث قدمت أساليب سيكومترية ذات فاعلية كبيرة في بناء المقاييس التربوية والنفسية، وكذلك طريقة تفسير درجات تلك المقاييس مقارنة بالنظرية التقليدية في القياس، وتستند هذه النظرية إلى افتراض وجود متصل للسمة بحيث يمكن تقدير احتمال الإجابة الصحيحة للفرد عن المفردة إذا علم موقعه على هذا المتصل (Warm,1978, 56).

ويرى هامبلتون وسوامنياثان (Hambleton and Swaminathan,1995) أن مستوى الطالب في الاختبار يتوقف على خصائص الطالب الذي يتعرض للاختبار، ويتأثر مستوى أداء الطالب في الاختبار بخصائص الفقرة من حيث الصعوبة والتمييز وفاعلية المموهات.

وعليه فإن الدراسة الحالية تسعى لمواكبة التطور في مجال القياس والتقويم التربوي من خلال فحص الخصائص السيكومترية للاختبار المركزي لمادة العلوم بمنطقة المدينة المنورة، وفق نظرية الاستجابة للمفردة باستخدام النموذج اللوغاريتمي المناسب.

مشكلة الدراسة:

قد يحصل بعض الطلبة على نتائج لا تعبر عن مستواهم الحقيقي في مادة ما، وقد يعود السبب في ذلك إلى خلل في الأداة المطبقة عليهم، سواء كان ذلك في تدني صدق وثبات هذه الأداة، أو كان ذلك في صعوبة أو سهولة فقرات هذا الاختبار، والتي قد تكون أعلى أو أقل من مستوى قدرتهم، مما يعطي نتائج غير دقيقة عن المستوى الحقيقي للطلبة، وذلك ما لاحظه الباحث خلال عمله كمعلم، حيث كان هناك بعض الطلبة يكونون على درجة عالية من الذكاء وعلى الرغم من مثابرتهم واجتهادهم تكون درجاتهم في بعض المواد متدنية. وعلى العكس من ذلك يحصل بعض الطلبة ذوي القدرات المنخفضة على درجة عالية في التحصيل لبعض المواد. مما يستوجب أن تكون الاختبارات على درجة ودقة عالية في خصائصها، وذلك لإصدار الحكم المناسب والصائب.

ولما للاختبارات المركزية التي طبقت على طلبة الصف الثالث المتوسط في مادة العلوم بالمدينة المنورة عام 1439/1438هـ أهمية كبرى في تحديد مدى تحقق الأهداف المرجوة من المادة، والاعتماد على نتيجة هذا الاختبار كأداة لقياس هذه الأهداف، رغب الباحث في فحص مدى جودته من خلال دراسة بعض الخصائص المتعلقة بمثل هذه الاختبارات ومدى تحقق افتراضات نظرية الاستجابة للمفردة.

أسئلة الدراسة:

ويمكن أن تلخص مشكلة الدراسة في السؤال البحثي التالي:

ما الخصائص السيكومترية المتوفرة في الاختبار المركزي لمادة العلوم للصف الثالث المتوسط وفق نظرية استجابة المفردة؟

ويتفرع من السؤال الرئيسي، الأسئلة الفرعية التالي:

1- ما مدى تحقق افتراضات نظرية استجابة المفردة في بيانات الاختبار؟
2- ما مدى مطابقة النموذج الثلاثي البارامتر من نماذج نظرية الاستجابة للمفردة أحادية البعد لبيانات الدراسة؟

3- ما قيم تقديرات معالم الفقرة للاختبار وفقًا للنموذج اللوجستي المستخدم؟
4- ما كمية المعلومات التي يقدمها الاختبار عند مستويات القدرة المختلفة؟

أهداف الدراسة:

تهدف الدراسة الحالية إلى التعرف على:

1- مدى تحقق افتراضات نظرية استجابة المفردة في بيانات الاختبار.
2- مدى مطابقة النموذج الثلاثي البارامتر من نماذج نظرية الاستجابة للمفردة أحادية البعد لبيانات الدراسة.
3- تقديرات معالم الفقرة للاختبار وفقًا للنموذج اللوجستي المستخدم.
4- كمية المعلومات التي يقدمها الاختبار عند مستويات القدرة المختلفة.

أهمية الدراسة:

تتمثل أهمية هذه الدراسة في أنها:

1- تسعى إلى تقييم الاختبارات المركزية من زاوية نظرية الاستجابة للمفردة، والتي تساعد في تطوير العملية التعليمية، حيث تقوم بالإسهام بتوضيح مدى جودة هذا الاختبار لمساعدة المختصين في الاستفادة من جوانب القوة وتلافي جوانب الضعف.
2- كما ستساعد هذه الدراسة في توفير معلومات وإرشادات تساعد المهتمين في تطوير هذا الاختبار بشكل خاص، والاختبارات الأخرى بشكل عام في تحسين طريقة اعدادها وبنائها والتحقق من مدى التزامها بمعايير الاختبار الجيد من وجهة نظر نظرية الاستجابة للمفردة، حيث أن الخلل في مثل هذه الأدوات يؤثر على شريحة كبيرة من الطلبة، مما يؤدي إلى اصدار حكم خاطئ واتخاذ قرار غير صحيح.
3- كما تسهم في تسليط الضوء على ضرورة بناء الاختبارات التربوية وفق نظرية الاستجابة للمفردة للتغلب على عيوب النظرية الكلاسيكية في القياس.

حدود الدراسة:

- الحدود الموضوعية: استخدام نظرية الاستجابة للمفردة في تقدير الخصائص السيكومترية للاختبار المركزي لمادة العلوم للصف الثالث المتوسط.
- الحدود النوعية: أسئلة الاختيار من متعدد.
- الحدود المكانية: المدينة المنورة.
- الحدود الزمانية: الفصل الثاني من العام الدراسي 1438-1439هـ.

مصطلحات الدراسة:

الأسئلة المركزية:

قامت وزارة التعليم في المملكة العربية السعودية مؤخرًا، باعتماد مشروع يهدف إلى جودة الاختبارات في التعليم العام، وأطلقت عليه ما يسمى بالأسئلة المركزية، بحيث تقوم كل إدارة تعليم باختيار مادة من مواد المرحلة المتوسطة، ووضع أسئلتها النهائية من قبل إدارة التعليم، وفي الفصل الثاني من عام 1438-1439هـ قامت إدارة

التعليم بالمدينة المنورة بوضع أسئلة لمادة العلوم للصف الثالث متوسط، وقام الباحث بجمع نماذج من إجابات الطلبة على هذه الأسئلة لغرض دراستها.

الخصائص السيكومترية:

يقصد بالخصائص السيكومترية: معالم الصعوبة والتمييز والتخمين للفقرة ومعاملي الصدق والثبات للاختبار (الفريحات، 2017).

الخصائص السيكومترية اجرائيًا: معالم الفقرة (الصعوبة، والتمييز والتخمين) ودالة المعلومات.

نظرية الاستجابة للمفردة Item Response Theory:

وتسمى أيضًا نظرية القياس الحديثة، وهي تفترض وجود علاقة منتظمة بين مستويات قدرة الفرد، واحتمالات استجابته الصحيحة لمفردات مختلفة، واستقلال خصائص المفردات، كمؤشر الصعوبة والقدرة التمييزية عن عينة الافراد التي استخدمت في تقديرها (الموسوي، 2014، 14).

معْلَم القدرة Ability Parameter:

هو مستوى القدرة أو السمة المقاسة لدى الأفراد المختبرين الذين يستجيبون على مفردات الاختبار المعد لقياس هذه السمة (Reeve,2002, 3).

معْلَمة الصعوبة:

هي نقطة على متصل القدرة تقابل احتمال (1+ci)2/ للإجابة على الفقرة ((i إجابة صحيحة (الزبون والصرايرة، 2017).

معْلَمة التمييز:

هي ميل منحنى خصائص الفقرة الذي يقابل النقطة التي تكون فيها علامة القدرة تساوي صعوبة الفقرة (القيسي، 2013).

معْلَمة التخمين:

هي احتمال إجابة فرد ذو قدرة متدنية على فقرة ما إجابة صحيحة (Baker,2001).

دالة معلومات الاختبار Test information function:

هو اقتران رياضي يعبر عن مجموع دوال المعلومات لفقرات الاختبار، وتعرف بأنها القيمة التي يتم الحصول عليها من تطبيق المعادلة التالية:

$$I(\theta) = \sum_{i-1}^{n} \frac{P_i'(\theta)^2}{P_i(\theta)Q_i(\theta)}$$

(الزبون والصرارية، 2017)

2- الإطار النظري والدراسات السابقة

أولاً- الإطار النظري:

ارتبط التقويم بالتعليم ارتباطا وثيقا وكأنهما وجهان لعملة واحدة، فهما ضروريان لتحقق العملية التعليمية أهدافها، ولكن التقارب الشديد بين التعليم والاختبار جعل من الصعوبة بمكان الفصل بينهما، فالمعْلَم لا يستطيع القيام بالعملية التعليمية لإنتاج تعليم جيد إذا لم يكن يتقن الاستفهام والاستفسار أثناء قيامه بتعليم أي جزء أو مفردة تعليمية من مقرره الدراسي، وعليه كان التعليم والتقويم متلازمين ومتداخلين ويسيران جنبا إلى جنب ويؤثر كل منهما في الآخر (عبد الرحمن، 2011).

ومفهوم التقويم أكثر اتساعا من كل من مفهومي القياس والاختبارات، حيث يعد عملية تؤكد دور الأحكام استنادا إلى أنواع مختلفة من البيانات الكمية والكيفية والتي يتم جمعها وتحليلها، وذلك للتوصل إلى أدلة تفيد في التعرف على القيمة، أو الفاعلية، أو اتخاذ قرارات معينة (علام، 2012).

وهناك ترابط بين القياس والاختبارات، لأن الاختبار في أبسط معانيه هو أحد أدوات القياس، ويعني بتحديد وتعيين أرقام للأشياء حسب قواعد معينة، تساعد مجري الاختبار من إعطاء قيمة عددية لأداء الشخص المفحوص، أما الاختبار فيعرفه بين Bean على أنه" مجموعة من المثيرات التي يمكن بواسطتها قياس بعض العمليات العقلية أو الخصائص النفسية بطريقة كمية أو كيفية (عبد الرحمن، 2011).

ويهتم المختصون في ميدان التربية وعلم النفس بالاختبارات التحصيلية، لما لها من أهمية كبرى في حياة الطالب الدراسية، فهي تعطي دلائل ومؤشرات لما يحدث في المؤسسة التعليمية من عمليات تعلم متنوعة ومتعددة للمهارات والمعارف والعلوم المختلفة والتي تدل على نشاطه العقلي (الجلالي، 2011).

ولما أنها تأخذ حيزا كبيرا عند القيام بعملية التقويم، فإنه من الضروري الاهتمام ببنائها، وتقنينها حسب المعايير المحددة لها، حيث تعد محكا رئيسا لقياس مدى نجاح المتعلم، واجتيازه لمهام التعلم التي يكون قد مر بها في مرحلة تعليمية معينة (علام، 2000).

الاختبارات التحصيلية وعلاقتها بنظريات القياس:

يرى براون (Brown,1970) أن الاختبار التحصيلي يقيس التعلم الذي يحدث نتيجة للخبرة في موقف تعلم يتم التحكم فيه مثل: صف مدرسي، أو برنامج تدريبي، وأن الإطار المرجعي الذي نهتم به هو الماضي أو الحاضر، أي الوضع الراهن لما تعلمه الطلاب.

وتعتبر الاختبارات التحصيلية من أهم الوسائل وأكثرها فاعلية بالنسبة لأطراف العملية التعليمية التعلمية جميعها، لما تقدمه من معلومات مفيدة لهم، ولكن مع تقدم العلوم التربوية والنفسية بشكل عام، وعلم القياس بشكل خاص، تبين قصور استخدام الاختبارات التحصيلية في قياس قدرات الطلبة التحصيلية، وتحديد مستوياتهم الدراسية، إذ وجهت لنظرية القياس التقليدية مجموعة من الانتقادات بسبب اعتمادها على معيار معين ووحيد في تفسير الدرجات، وهذا المعيار هو معيار الجماعة في الاختبارات جماعية المعيار، ومعيار المحك في الاختبارات محكية المرجع (علي، 2012).

ونظرا لأهمية الاختبارات بشكل عام، واختبارات التحصيل بشكل خاص، فقد اهتم الباحثون بإجراء الكثير من الدراسات والبحوث التي اهتمت ببناء فقرات الاختبارات وانتقائها بالاعتماد على هذه النظرية، وكان نتيجتها قيام الكثير من العلماء مثل ثرستون Thurstone ولورد Lord وبيرنبوم Birnbaum بتوجيه مجموعة من الانتقادات

للنظرية، بسبب قصورها في تحقيق الموضوعية في القياس، والمتمثلة في اعتماد القياس على خصائص الاختبار المستخدم (الصدق، والثبات، والموضوعية)، بمعنى اختلاف نتيجة القياس باختلاف الاختبار المستخدم، وكذلك اختلاف نتيجة القياس باختلاف مستوى الأفراد الذين يطبق عليهم الاختبار، بالإضافة إلى ذلك عدم خطية القياس؛ أي عدم وجود وحدات متساوية على مدى المتصل (القدرة) الذي يتدرج عليه المتغير موضوع القياس (علام، 2005).

نظرية الاستجابة للمفردة:

تعد نظرية الاستجابة للمفردة Item Response Theory نقطة تحول بسبب ما قدمته من طرق ذات فعالية كبيرة فيما يتعلق ببناء الاختبارات وتصحيحها وتحليل نتائجها مقارنة بما قدمته النظرية التقليدية في القياس (Mislevy & Bock,1990).

وتهدف نظرية الاستجابة للمفردة والنماذج الرياضية المرتبطة بها إلى تقدير جميع إحصاءات المفردة والقدرة، وكلما كان هناك ملائمة Fit بين النموذج المستخدم ومجموعة البيانات أدى ذلك إلى الحصول على تقديرات دقيقة لهذه الإحصاءات، حيث يتم وضع كل من المفحوصين والفقرات على مقياس القدرة، وذلك من خلال عمليات التقدير، طالما يوجد علاقة تقارب بين الاحتمالات المتوقعة للمفحوصين والاحتمالات الواقعية لأدائهم، في كل مستوى من مستويات القدرة، ويجب أن تعدل وتراجع باستمرار تقديرات احصاءات المفردة وتقديرات قدرة المفحوصين حتى يتم التوصل إلى أقصى اتساق ممكن للتنبؤ ويتم ذلك من خلال البرامج المعدة لذلك (أبو هاشم، 2006).

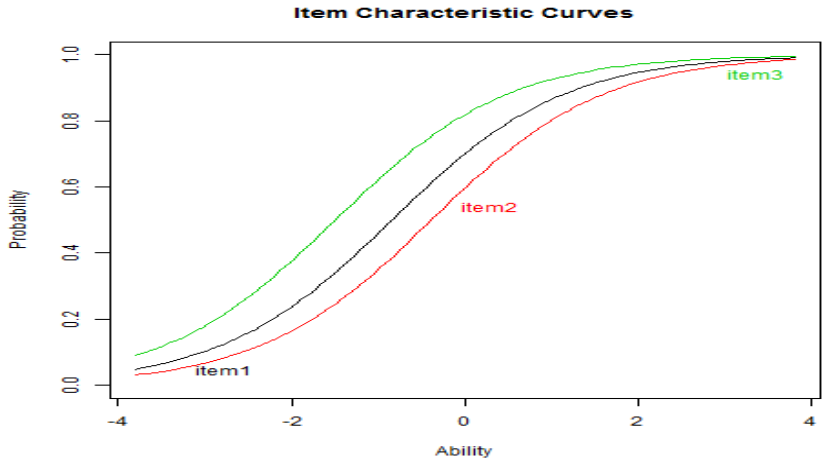
نماذج نظرية الاستجابة للمفردة أحادية البعد ثنائية التدريج:

المفردات ذات الصيغة (0.1) ثنائية الدرجات هي الأكثر استخداماً، ونماذج نظرية الاستجابة للمفردة (أحادية-ثنائية-ثلاثية) المعالم هي أيضاً من أكثر النماذج المستخدمة حالياً وهي:

1- النموذج اللوجستي أحادي المعْلَم (نموذج راش) (Rasch Model) One Parameter Logistic Model:

يعد هذا النموذج من أبسط نماذج نظرية الاستجابة أحادية البعد، ويطلق عليه نموذج المعْلَم الواحد لأنه يحتوي على معْلَم صعوبة المفردة فقط، لتمثيل الفروق بين المفردات، واستمد راش Rasch نموذجه من نقطة بسيطة، وهي أن نجاح فرد في أدائه على مفردة اختبارية معينة، يجب أن يعتمد على الفرق بين محصلة قدرة هذا الفرد وعلى صعوبة هذه المفردة الاختبارية(Lidnder,1997) .

والشكل رقم (1) يمثل (3) مفردات وفق النموذج أحادي المعْلَم حيث يفترض بأن التمييز موحد وأنه لا يوجد هناك تخمين والاختلاف فقط في معْلَم الصعوبة ويتضح بأن المفردة رقم (2) تبدو أكثر صعوبة من المفردتين الأولى والثالثة:

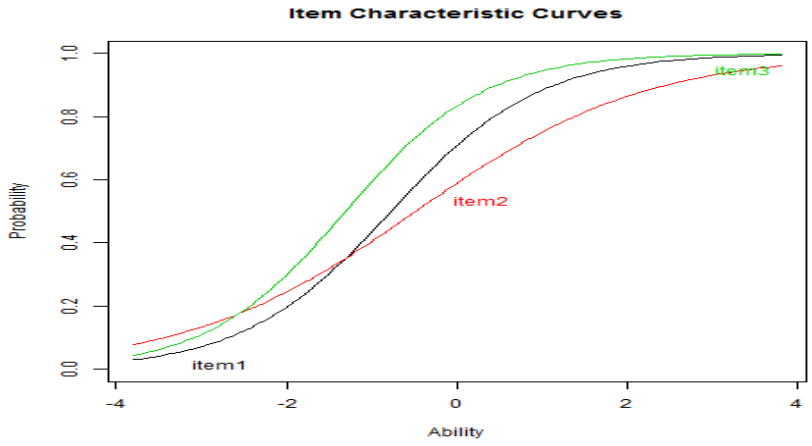


الشكل رقم (1) منحنى خصائص المفردة ل (3) مفردات وفقًا لنموذج رآش

2- النموذج اللوجستي ثنائي المعْلَم Two Parameter Logistic Model:

ينسب النموذج ثنائي المعْلَم إلى لورد (1952)، والذي استخدم المنحنى الطبيعي التراكمي، وتوصل بيرنبوم في عام (1968) إلى تعديل النسخة الرياضية لهذا النموذج مستخدما المنحنى اللوجستي وذلك بإضافة معْلَم التمييز لكل فقرة (الأشقر وعيلبوني، 2015).

وهذا النموذج يفترض أن المفردة تختلف في صعوبتها وتمييزها، كما أنه يفترض عدم تأثرها بالتخمين، والشكل رقم (2) يوضح ثلاثة مفردات مختلفة في تمييزها وصعوبتها وعدم تأثرها بالتخمين وهو ما يفترضه هذا النموذج.



الشكل رقم (2) منحنى خصائص المفردة ل (3) مفردات وفقًا للنموذج الثنائي البارامتر.

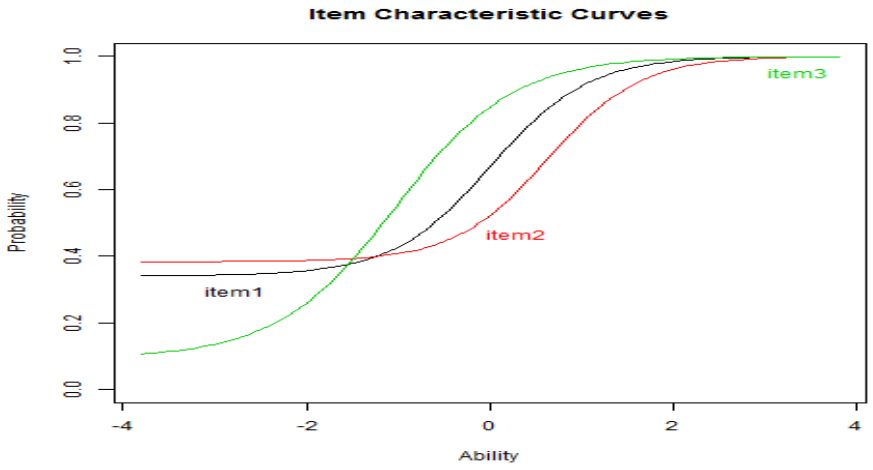
3- النموذج اللوجستي ثلاثي المعْلَم Three Parameter Logistic Model:

وهو أحد نماذج نظرية الاستجابة للمفردة الاختبارية ثنائية التدريج حيث يمكن لهذا النموذج تقدير ثلاثة معالم للمفردة الاختبارية: الصعوبة، التمييز، والتخمين (الصالح وسليمان، 2016).

كما يعد النموذج اللوجستي ثلاثي البارامتر أقل النماذج أحادية البعد تشددًا، حيث يسمح بأن تكون فقرات الاختبار مختلفة في صعوبتها وتمييزها، بخلاف نموذج راش الذي اعتبر ان الفقرات تختلف في صعوبتها فقط، ويفترض النموذج الثلاثي البارامتر أيضا أن تتأثر الفقرات بعامل التخمين والذي يسمى بمعْلَم الخط التقاربي الأدنى (Lower

Asymptote Line) والذي يقصد به أنه احتمال ان يجيب فرد منخفض القدرة على فقرة ذات صعوبة مرتفعة (الشريفين، 2012).

ويوضح الشكل رقم (3) ثلاثة مفردات مختلفة في صعوبتها وتمييزها وتأثرها بعامل التخمين:



الشكل رقم (3) منحنى خصائص المفردة ل (3) مفردات وفقًا للنموذج الثلاثي البارامتر

4- النموذج اللوجستي رباعي المعْلَم :Four Parameter Logistic Model

قدمه بارتون ولورد (Barton & Lord,1981) والذي يسمى بمعْلَم الخط التقاربي الأعلى (Upper asymptote Line) حيث (di) تمثل خط التقارب الأعلى لمنحنى خصائص الفقرة ويطلق عليه معْلَمة اللامبالاة والذي يفترض أن المفحوصين ذوي القدرات المرتفعة قد لا يستجيبون لفقرات الاختبار بإجابات صحيحة حتى على الفقرات السهلة والتي يجيب عنها ذوي القدرات المنخفضة، وقد يرجع السبب في ذلك أمور عدة منها اللامبالاة أو عدم وضوح الفقرة، لذلك أضيف هذا النموذج. (الشريفين، 2017).

الصور الرياضية لمعادلات النماذج اللوجستية (أحادي-ثنائي-ثلاثي-رباعي) المعْلَم:

1- النموذج أحادي المعْلَمة (نموذج راش)One Parameter Logistic Model (Rasch Model :

$$P_i(\theta) = \frac{e^{\theta_j - b_i}}{1 + e^{\theta_j - b_i}}$$

2- النموذج ثنائي المعْلَمة Two Parameter Logistic Model:

$$P_i(\theta) = \frac{e^{a_i(\theta_j - b_i)}}{1 + e^{a_i(\theta_j - b_i)}}$$

3- النموذج ثلاثي المعْلَمة Three Parameter Logistic Model:

$$P_i(\theta) = c_i + (1 - c_i)\frac{e^{a_i(\theta_j - b_i)}}{1 + e^{a_i(\theta_j - b_i)}}$$

4- النموذج رباعي المعْلَمة Four Parameter Logistic Model:

$$P_i(\theta) = c_i + (d_i - c_i)\ \frac{e^{Da_i(\theta - b)}}{1+e^{Da_i(\theta - b_i)}}$$

ثانياً- الدراسات السابقة:

تعددت الدراسات التي تناولت استخدام نظرية الاستجابة للمفردة في بناء الأسئلة والكشف عن خصائصها السيكومترية، وفيما يلي سيتم عرض بعض منها مرتبة ترتيبًا تصاعديًا بحسب تاريخها:

1. قامت الحموري (2011) بإجراء دراسة هدفت إلى استقصاء إمكانية استخدام نماذج نظرية الاستجابة للمفردة في بناء اختبار تحصيلي من إعداد المعلم. حيث تم بناء اختبار تحصيلي مكون من (23) فقرة من نمط الاختيار من أربعة بدائل تقيس العدد نفسه من الأهداف المقصودة، وطبق الاختبار على (284) طالباً وطالبة، وحللت البيانات الناتجة من الاستجابة للفقرات وفق نظرية الاستجابة للمفردة، وكشفت نتائج الدراسة عدم إمكانية استخدام هذه النظرية لبناء الاختبارات ذات العدد القليل من الفقرات.
2. كما هدفت دراسة أبو جراد (2012) إلى المقارنة بين الخصائص السيكومترية لاختبار الاختيار من متعدد واختبار التكميل في ضوء نظرية الاستجابة للمفردة، حيث تم إعداد اختبارين في مادة العلوم لطلاب الصف السادس الابتدائي، أحدهما كان عبارة عن اختيار من متعدد والآخر من نوع التكميل وكلاهما مكون من (30) مفردة وتم استخدام نموذج راش لتدريج ومعايرة كلًا من الاختبارين على عينة مكونة من (80) طالباً وطالبة، وكشفت نتائج الدراسة عن حذف (6) مفردات من اختبار الاختيار من متعدد و (7) مفردات من اختبار التكميل، وأشارت النتائج إلى أن معاملات ثبات تقديرات الصعوبة للمفردات والأفراد في اختبار التكميل أعلى منها في اختبار الاختيار من متعدد.
3. وتناول النفيعي (2012) دراسة هدفت إلى إيجاد الخصائص السيكومترية ومعايير تفسير الأداء لاختبار المصفوفات المتتابعة المتقدم، وتم تحليل بيانات عينة الدراسة والبالغ عددهم (1808) فردا باستخدام النموذج اللوغاريتمي ثنائي المعْلَم، وأظهرت النتائج أنه النموذج الأكثر ملاءمة للبيانات، وذلك بعد التأكد من تحقق افتراضات نظرية الاستجابة للمفردة الاختبارية في بيانات الدراسة، والمتمثلة في أحادية البعد، والاستقلال الموضعي لأزواج الفقرات، وملائمة الأفراد والفقرات لتوقعات النموذج ثنائي المعْلَم، وأكدت النتائج تمتع الاختبار بخصائص سيكومترية جيدة (معْلَم تا التمييز والصعوبة، ودالة معلومات الاختبار) في ضوء نظرية الاستجابة للمفردة الاختبارية، وبالتالي تم إيجاد قدرة الافراد المناظرة لكل درجة كلية خام على الاختبار، والمئينيات المناظرة لها، وهي المعايير التي يفسر في ضوئها الأداء على الاختبار.
4. واستخدمت دراسة قام بها كل من: بانايدس ووالكر (Panayides & Walker, 2013) النموذج احادي البارامتر في الفحص عن الخصائص السيكومترية لمقياس قلق اللغة الأجنبية الفصلي لعينة عددها (304) من طلبة

المرحلة الثانوية، وأشارت النتائج بأن هناك بعض المفردات التي لم تطابق النموذج، ولكن المقياس تمتع بدرجة عالية من الثبات.

5. وفي دراسة الشريفين وحجازي (2014) والتي هدفت إلى تقصي الخصائص السيكومترية لاختبار المفردات بأسمائها واستعمالاتها، حسب نظرية السمات الكامنة، قاما الباحثان بتطبيق الاختبار على عينة مكونة من (246) طالباً وطالبة بين سن (5-9) سنوات تم اختيارهم عشوائيا من المدارس الحكومية في مدينة اربد، حيث استخدمت الدراسة النموذج الثنائي التدريج لتحليل الاستجابات، وتكون الاختبار في صورته النهائية من (16) فقرة كانت مطابقة للنموذج، وتراوحت قيم صعوبتها بين (2.13 و2.16) لوجيت وبمتوسط يساوي صفر لوجيت وبلغت قيم معامل الثبات للأفراد والفقرات (0.81)، (0.92) على التوالي، وتمتع المقياس بدلالات صدق متعددة.
6. وهدفت دراسة الحاج (2016) لاستخدام نموذج راش في بناء اختبار تحصيلي محكي المرجع في مقرر علم النفس التربوي، قامت الباحثة بتصميم اختبار مكون من (50) فقرة من نوع الاختيار من متعدد، بلغت قيمة معامل ثبات كرونباخ ألفا (0,87)، حيث طبقت الباحثة الاختبار على عينة مكونة من (246) طالبة من طالبات جامعة ام القرى، في الفصل الثاني من عام (1433-1434هـ) وتم استخدام برنامج (Bilog-MG3) في فحص بيانات الاختبار لنموذج راش وكذلك في حساب معالم الفقرات، وأسفرت النتائج عن تحقق افتراضات نموذج راش ومطابقة الاستجابات عن (25) فقرة من فقرات الاختبار للنموذج، بينما تم حذف (25) فقرة لم تطابق النموذج، وتراوحت قيم معالم الصعوبة بين (0,698-) و (0,570) مما يظهر أن الاختبار متوسط الصعوبة.
7. وتناول الفريحات (2017) دراسة هدفت إلى تقدير الخصائص السيكومترية لاختبار محكي المرجع في الكيمياء وفق النظرية الحديثة في القياس النفسي والتربوي، حيث تم بناء اختبار تحصيلي في مادة الكيمياء تألف من (57) فقرة من نوع الاختيار من متعدد بأربع بدائل، وتم تطبيقه على عينة بلغت (390) طالبا وطالبة من الصف الأول الثانوي العلمي في مديرية التربية والتعليم الأولى في اربد، أشارت نتائج الدراسة إلى مطابقة الاستجابات عن (47) فقرة من فقرات الاختبار لافتراضات النموذج الأحادي المغلَمة (نموذج راش) وحذفت (10) فقرات لم تطابق هذا النموذج، وبلغ معامل ثبات الأفراد (0.89) ومعامل ثبات الاختبار (0.97) كما تم التحقق من تمتع الاختبار بصدق المحتوى والصدق التلازمي، كما بينت النتائج أن قيمة المعلومات للاختبار كانت أكبر ما يمكن عندما كانت ($\theta = b$) وهذا ما يتوقعه نموذج راش، كما أن الاختبار يقدم أكبر كمية معلومات عن أداء المفحوصين عند مستويات القدرة المتوسطة، وأقل كمية من المعلومات كانت عند مستويات القدرة العالية والمنخفضة.

التعليق على الدراسات السابقة:

عند النظر إلى الدراسات السابقة، يتضح أن أكثر الدراسات مثل دراسة الحاج (2016)، الفريحات (2017)، وبانايدس ووالكر (Panayides & Walker,2013) قامت بتوظيف نموذج راش والذي يفترض تساوي التمييز واستبعاد التخمين ويقوم على مغلَمة الصعوبة فقط، وهذا نادر الحدوث خاصة مع الاختبارات التحصيلية والتي تحتوي على أسئلة اختيار من متعدد، أما النفيعي (2012) و الشريفين وحجازي (2014) فطابقت بيانات دراستهم النموذج الثنائي البارامتر، واستخدم هذا النموذج في تقصي الخصائص السيكومترية للبيانات.

وقد لاحظ الباحث أن الدراسات التي اهتمت بتصميم ودراسة خصائص الاختبارات التحصيلية باستخدام النموذج الثلاثي البارامتر ما زالت قليلة مقارنة بنموذج راش، فقد يعزى سبب ذلك بأن النموذج يتطلب عدد كبير من الأفراد وكذلك عدد كبير من الفقرات، مما يحتاج إلى الكثير من الوقت والجهد. أما دراسة اوجرندي (Ojerinde, 2013) فقد قارنت بين النظرية التقليدية في القياس ونظرية الاستجابة للمفردة، وأشارت النتائج بأن النظريتين مكملة لبعضهما البعض ومتشابهة مع تفوق نظرية الاستجابة للمفردة في دقة التقديرات لإعطائها معلومات أكثر ثراء.

وتميزت الدراسة الحالية في السبق في تقدير الخصائص السيكومترية للاختبار المركزي لمادة العلوم في المدينة المنورة، بإستخدام النموذج الثلاثي البارامتر والذي يحتوي على الثلاث معالم: الصعوبة، التمييز، التخمين.

3- منهجية الدراسة واجراءاتها

منهج الدراسة:

نظرًا لطبيعة الدراسة وبياناتها استخدم الباحث المنهج الوصفي المسحي التحليلي. وهو أسلوب من أساليب البحث العلمي، حيث يعتمد على دراسة الواقع أو الظاهرة كما توجد في الواقع، ويهتم بوصفها وصفاً دقيقاً ويعبر عنها كيفياً، من حيث وصف الظاهرة وتوضيح خصائصها، أو تعبيراً كميا من حيث إعطاء وصفاً رقمياً يوضح مقدار هذه الظاهرة أو حجمها ودرجات ارتباطها مع الظواهر المختلفة الأخرى، ولا يقتصر المنهج على وصف الظاهرة أو الواقع كما هو بل الوصول إلى استنتاجات وتعميمات تساهم في تطوير وفهم الواقع (عبيدات وعدس وكايد، 2000).

مجتمع الدراسة:

تكّون مجتمع الدراسة من مجموعة الأسئلة المركزية (الموضوعية) المستخدمة في تحديد تحصيل طلاب الصف الثالث متوسط (بنين) لمادة العلوم بالمدارس الحكومية والأهلية التابعة لإدارة التعليم بالمدينة المنورة والبالغ عددهم 12453 طالب في المدينة المنورة.

عينة الدراسة:

مجموعة الأسئلة المركزية لمادة العلوم التي طبقت في عام 1438-1439هـ، وذلك باختيار عدد نماذج إجابات (1040) طالبًا وبنسبة (8.35%) من مجتمع الدراسة البالغ 12453طالب

قام الباحث بالتأكد من توفر شروط التحليل العاملي للبيانات، حيث جرى فحص ملاءمة المعاينة للتحليل عن طريق اختبار كايزر-ماير-أولكن KMO والذي يوضح مدى كفاية حجم العين لإجراء التحليل العاملي وكانت القيمة تساوي (0.89) وكما هو واضح هذه القيمة ليست أقل من (0.5) وهو المحك الذي وضعه كايزر ((Kaiser,1974 لتحقق هذا الشرط، مما يعني تحقق شرط الملائمة، وكذلك تم اجراء اختبار بارتليت Bartlett's Test of Sphericity لفحص تجانس العينة بالنسبة لحجم العينة من خلال قيمة مربع كاي X^2 حيث بلغت (5336.14) عند درجة حرية (780) ومستوى دلالة (0.00) وهي دالة احصائياً عند (0.01) مما يؤكد تحقق هذا الشرط أيضًا.

المعالجة الإحصائية:

للإجابة عن أسئلة الدراسة استخدم الباحث البرامج التالية:

1- برنامج Excel.

- لحساب أقصى كمية معلومات تقدمها الفقرة من خلال المعادلات الخاصة بذلك.

2- برنامج SPSS.
- استخدام التحليل العاملي الاستكشافي للتحقق من أحادية البعد.

3- برنامج R.
- لمعرفة النموذج الأكثر ملاءمة من نماذج نظرية الاستجابة للمفردة أحادية البعد لبيانات الدراسة.
- لاستخراج قيم معالم مفردات الاختبار (الصعوبة، التمييز، والتخمين).

4- عرض نتائج البحث ومناقشتها

- **أولاً: للإجابة على سؤال الدراسة الأول ونصه": ما مدى تحقق افتراضات نظرية الاستجابة للمفردة في بيانات الاختبار؟**

وبعد التأكد من تحقق شروط التحليل العاملي قام الباحث بما يلي:

أولًا- افتراض أحادية البعد Unidimensionality:

استخدام الباحث التحليل العاملي الاستكشافي بطريقة المكونات الرئيسية (Principal Component Method) ومن ثم تم تدوير العوامل باستخدام طريقة Varimax وذلك بافتراض استقلالية العوامل، وأسفرت التحليلات الخاصة بهذا الأسلوب البيانات الموضحة بالجدول التالي:

جدول رقم (1) الجذور الكامنة للعوامل

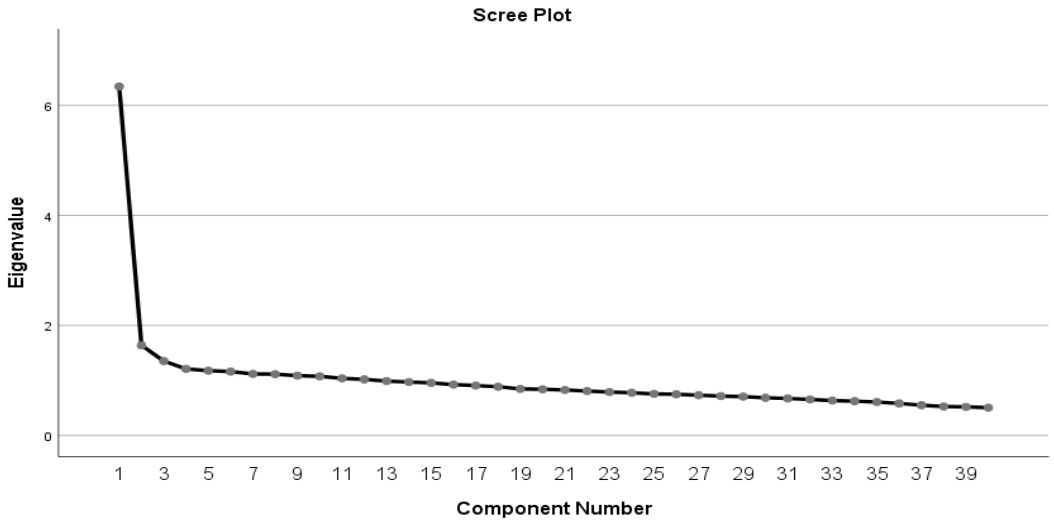
العامل	الجذر الكامن	النسبة المئوية للتباين المفسر	النسبة المئوية للتباين المفسر التراكمي
الاول	6.34	15.89	15.86
الثاني	1.64	4.09	19.95
الثالث	1.35	3.38	23.33
الرابع	1.21	3.02	26.35
الخامس	1.18	2.94	29.29
السادس	1.16	2.90	32.19
السابع	1.12	2.79	34.98
الثامن	1.11	2.78	37.76
التاسع	1.09	2.71	40.47
العاشر	1.07	2.68	43.16
الحادي عشر	1.04	2.59	45.75
الثاني عشر	1.01	2.54	48.29

أفرز التحليل (12) عاملاً وكل منها يزيد عن الواحد، ويتضح من الجدول أعلاه بأن قيمة الجذر الكامن للعامل الأول تساوي (6.34) ويفسر ما نسبته (15.86) من تباين الأداء على نموذج الاختبار وهذه القيمة تشكل ما مقداره (32.84) من التباين الكلي الذي تفسره جميع العوامل الاثني عشر مجتمعة.

وكما نلاحظه بأن قيمة الجذر الكامن للعامل الأول هي (6.34) وقيمة الجذر الكامن للعامل الثاني تساوي (1.64) وبناء على مقترح (Lord,1980) في امبرستون ورايس ((Embreston and Reise,2010، بأنه إذا كانت نسبة الجذر الكامن للعامل الأول إلى نسبة الجذر الكامن للعامل الثاني أكبر من (2) فذلك يؤكد تحقق أحادية البعد، وهذا ما نلاحظه حيث أن القيمة تساوي (3.86) وهي أكبر من المحك الذي وضعه لورد Lord.

ويذكر وورم (Warm,1978) بأن أحادية البعد تحقق في حال كان الفرق بين الجذر الكامن للعامل الأول والثاني كبيراً، والنسبة بين الجذور الكامنة المتتالية متقاربة، وهذا ما يتضح من خلال الجدول أعلاه.

ووفقًا لوهلن (Lohelin,1987) يتم التوصل إلى عدد العوامل اعتمادًا على النقطة التي يتحول فيها ميل منحنى الجذور الكامنة بسرعة من كونه منحنى يعامد تقريبًا محور السينات وتغيره إلى منحى يوازي تقريبًا هذا المحور وهذا يحصل عند العامل الأول، وهذا ما نلاحظه بالشكل رقم (7) والذي يعتبر مؤشرًا لتحقق افتراض أحادية البعد.



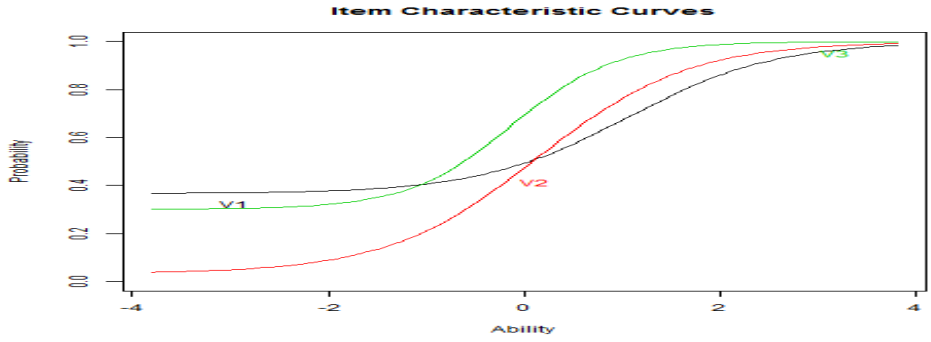
شكل رقم (7) مخطط سكري تمثيل الجذور الكامنة بيانيًا

وبناءً على ما سبق يمكن القول بأن افتراض أحادية البعد متحقق في بيانات الدراسة، أي أن الاختبار يقيس سمة واحدة.

أما بالنسبة لافتراض الاستقلال الموضعي:

فقد أشار هامبلتون وسوامينثن (Hambelton & Swaminthan, 1985) بأنه في حال تحقق افتراض أحادية البعد فذلك يعني تحقق افتراض الاستقلال الموضعي، كما أن هناك العديد من الدراسات التي أجريت للتحقق من تكافؤ الافتراضين ومنها دراسة حمادنة (2011) والتي أشارت في نتائجها بأن تحقق افتراض أحادية البعد يضمن للباحث تحقق افتراض الاستقلال الموضعي، وبناء على ذلك يمكن القول بأن هذا الافتراض متحقق بناءً على تحقق افتراض أحادية البعد.

وللتحقق من افتراض المنحنى المميز للمفردة:



شكل رقم (8) منحنى خصائص الفقرات لأول ثلاثة مفردات من الاختبار

وهي الدالة المميزة لكل مفردة والتي تصف العلاقة الوتيرية بين القدرة والأداء على المفردة، حيث يمكن التحقق من هذا الافتراض من خلال إيجاد معالم المفردات التي تحدد منحنى خصائص كل منها (الحموري، 2011).

ويمكن القول بتحقق هذا الافتراض حيث يظهر من خلال الشكل (8) دالة مميز لكل فقرة تصف العلاقة بين القدرة واحتمالية الإجابة على الفقرة، وتم التحقق من جميع المفردات للتأكد من هذا الافتراض وهذا الشكل يمثل أول ثلاثة مفردات أمكن رسم منحنى خصائص الفقرات لها.

افتراض التحرر من السرعة:

والذي ينص على أنه يجب أن لا يكون عامل السرعة عاملًا حاسمًا في الاستجابة على المفردات الاختبارية، حيث أنه من الضروري أن يكون أداء الطالب هو أقصى أداء يمكن أن يقوم به للإجابة على فقرات الاختبار، بحيث أن فشل الطالب على في الإجابة الصحيحة عن مفردة ما، تكون بسبب ضعف قدرته، وليس عدم تمكنه من الوصول إليها بسبب عامل السرعة، كما أن هذا الافتراض يعد من الافتراضات الضمنية، وذلك يعني بأنه تحقيق افتراض أحادية البعد، يعني عدم وجود عوامل أخرى (الناغي، 2011). ويتضح من نتائج افتراض أحادية البعد أن هناك عامل واحد وهو السمة الكامنة للفرد، وبذلك يمكن القول إن السرعة ليس لها تأثير، بالإضافة إلى أنه وجد أن جميع الأفراد أجابوا عن جميع فقرات الاختبار كاملة ولم يتركوا مفردات بدون إجابة ما عدا (3) طلاب وجد عندهم بعض المفردات متروكه واستبعدت أوراقهم لضمان تحقق هذا الافتراض بشكل تام، لذلك يمكن القول بأن الاختبار لم يتأثر بعامل السرعة. حيث إنه بناءً على ما ذكره هامبلتون وآخرون Hambleton et al أنه إذا كان (75%) من التلاميذ قد أكملوا الإجابة عن الاختبار و (80%) من مفردات الاختبار أجاب عليها التلاميذ فإن السرعة لن تعد في هذه الحالة عامل مؤثر في الأداء على الاختبار (الشريف وفراج، 2013).

وبناءً على ما سبق يمكن القول بأن افتراضات نظرية الاستجابة للمفردة أحادية البعد متحققة في بيانات الدراسة مما يمكن استخدام نماذجها لتحليل هذه البيانات.

- **ثانيًا- للإجابة على السؤال القائل: ما هو النموذج الأكثر ملاءمة من نماذج نظرية الاستجابة للمفردة أحادية البعد لبيانات الدراسة؟**

قام الباحث بتحليل البيانات مستخدما حزمة ltm من خلال برنامج R لمعرفة النموذج الأكثر مطابقة لبيانات الدراسة، وعند النظر لقيمة (AIC) وهو المحك الذي وضعه أكايك (Akaike,1974) والذي يشير بأن النموذج الأفضل مطابقة للبيانات هو الذي يحمل القيمة الأقل من بين النماذج الأخرى، حيث تبين بأن قيمة (AIC) كانت تساوي (38680.19) في النموذج الثلاثي البارامتر بينما كانت القيمة في النموذج الثنائي البارامتر تساوي (38838.4) أما في النموذج احادي البارامتر فكانت القيمة تساوي (39161.084) لذلك يتضح بأن القيمة الأقل ل (AIC) تكون في النموذج الثلاثي البارامتر، وأيضًا وجد بأن النموذج الثلاثي البارامتر أكثر النماذج مطابق للمفردات، حيث أتضح أن هناك (18) مفردة غير متطابقة مع النموذج الأحادي و (5) مفردات غير متطابقة مع النموذج الثنائي بينما في النموذج الثلاثي كانت هناك (4) مفردات فقط غير متطابقة، وذلك من خلال النظر لقيم مربع كاي عند مستوى دلالة (0.01) والتي تكون الفقرات غير متطابقة عندما تقل القيمة عن مستوى الدلالة المذكور، كما في دراسة (حمادنة، 2011)، وهذه الفقرات هي التي تحمل الأرقام (4، 9، 10، 24) ويمكن القول بأن النموذج الثلاثي هو النموذج الأنسب لهذه البيانات، وفيما يلي عرض لقيمة مربع كاي للنموذج الثلاثي البارامتر:

جدول رقم (2) مدى جودة مطابقة البيانات للنموذج الثلاثي البارامتر

رقم الفقرة	مربع كاي	مستوى الدلالة	رقم الفقرة	مربع كاي	مستوى الدلالة
الفقرة 1	7.639	0.366	الفقرة 21	12.148	0.096
الفقرة 2	15.792	0.027	الفقرة 22	7.767	0.354
الفقرة 3	11.846	0.106	الفقرة 23	14.047	0.050
الفقرة 4	21.195	0.003	الفقرة 24	18.944	0.008
الفقرة 5	3.320	0.854	الفقرة 25	12.891	0.075
الفقرة 6	13.142	0.069	الفقرة 26	12.269	0.092
الفقرة 7	9.628	0.211	الفقرة 27	13.254	0.066
الفقرة 8	10.915	0.142	الفقرة 28	7.682	0.361
الفقرة 9	23.281	0.002	الفقرة 29	12.378	0.089
الفقرة 10	21.149	0.004	الفقرة 30	16.207	0.023
الفقرة 11	10.945	0.141	الفقرة 31	11.977	0.101
الفقرة 12	7.444	0.384	الفقرة 32	9.053	0.249
الفقرة 13	14.687	0.040	الفقرة 33	7.044	0.424
الفقرة 14	6.769	0.453	الفقرة 34	14.850	0.038
الفقرة 15	12.000	0.101	الفقرة 35	12.813	0.077
الفقرة 16	6.303	0.505	الفقرة 36	13.002	0.072
الفقرة 17	16.726	0.019	الفقرة 37	11.038	0.137
الفقرة 18	15.045	0.035	الفقرة 38	8.084	0.325
الفقرة 19	14.112	0.049	الفقرة 39	8.392	0.299
الفقرة 20	11.793	0.108	الفقرة 40	7.348	0.394

وبالرغم من تقارب عدد المفردات الغير مطابقة في النموذج الثنائي البارامتر مع النموذج الثلاثي البارامتر، إلا أن النموذج الثلاثي حصل على القيمة الأقل بناءً على محك أكايك (Akaike) وأكثر عدد مطابقة للبيانات بالإضافة إلى أن النموذج الثلاثي البارامتر أعم وأشمل نماذج نظرية الاستجابة للمفردة الأحادية، وبناءً على ما سبق ولكون الاختبار مكوّن من أسئلة اختيار من متعدد وقد يتضمن التخمين، لذا يرى الباحث بأن النموذج الثلاثي البارامتر هو النموذج الملائم لتحليل بيانات هذه الدراسة، وذلك بسبب أن هذا النموذج أكثر واقعية كما أشار إليه هامبلتون وتروب (Hambleton & Traub,1971) وأنه الطريقة المثلى في تقدير قدرة الفرد نظرًا لاحتوائه على معالم النموذجين الأول والثاني، أي شموله على معْلَم الصعوبة والتمييز بالإضافة إلى معْلَم التخمين، مما يؤدي لتحقيق قياسات موضوعية و نتائج دقيقة.

وهذه النتيجة تتشابه مع دراسة (الشمراني، 2014) ودراسة (إبراهيم، 2009) والتي هدفت للمقارنة بين النماذج اللوغارتمية الثلاثة من حيث ملاءمتها للبيانات حيث كانت الأداة عبارة عن اختبار تحصيلي مكون من اختيار من متعدد وتكميل، ووجد أن النموذج الثلاثي أكثر النماذج مطابقًا للفقرات كما أعطى أكبر قيمة لدالة المعلومات.

وكما هو واضح عدد الفقرات غير المطابقة للنموذج الثلاثي (4) فقرات من الاختبار، لذلك تم حذفها ليتبقى (36) فقرة من الاختبار والتي تتطابق مع النموذج، لذلك يمكن القول بأن هناك حسن مطابقة بين النموذج الثلاثي وبيانات الدراسة والذي يعتبر مناسباً لاستخدامه في هذه الدراسة وإجراء التحليلات الإحصائية المختلفة.

- **ثالثًا- أما السؤال الذي ينص على " ما قيم تقديرات معالم الفقرة للاختبار وفقًا للنموذج اللوجستي المستخدم؟".**

فللإجابة عليه تم إيجاد معالم فقرات الاختبار (الصعوبة، التمييز، والتخمين) وفقًا للنموذج الثلاثي البارامتر بعد حذف الفقرات غير المطابقة للنموذج وكان عدد الفقرات الإجمالي (36) فقرة كما هو موضح بالجدول رقم (3):

جدول رقم (3) قيم معالم مفردات الاختبار (الصعوبة، التمييز، والتخمين)

المعْلَم / رقم الفقرة	الصعوبة	التمييز	التخمين
الفقرة 1	** 1.08	** 1.39	** 0.38
الفقرة 2	0.08	** 1.23	0.00
الفقرة 3	-0.15	** 1.84	0.29
الفقرة 5	** 1.20	** 1.08	** 0.38
الفقرة 6	-0.04	0.51	0.01
الفقرة 7	** 0.99	** 2.49	** 0.50
الفقرة 8	** 1.73	0.51	0.01
الفقرة 11	0.44	** 2.45	** 0.42
الفقرة 12	** 0.83	** 2.66	0.30
الفقرة 13	0.52	** 0.98	0.11
الفقرة 14	** 0.88	** 1.62	0.30
الفقرة 15	** 1.10	** 0.80	0.22
الفقرة 16	0.59	** 1.35	0.10
الفقرة 17	-0.35	** 1.14	0.00
الفقرة 18	0.12	** 1.58	0.27
الفقرة 19	-0.17	** 1.54	0.12
الفقرة 20	0.51	** 1.21	0.33
الفقرة 21	** 1.34	** 1.43	0.22
الفقرة 22	** 0.10	** 2.17	0.14
الفقرة 23	** 2.22	* 0.28	0.10
الفقرة 25	* -0.72	** 0.83	0.01
الفقرة 26	0.06	** 1.16	0.32
الفقرة 27	* -0.81	0.57	0.00
الفقرة 28	-0.40	** 1.23	** 0.46
الفقرة 29	-0.01	** 1.00	0.19
الفقرة 30	** 0.83	** 1.01	0.22
الفقرة 31	0.39	** 1.73	0.20
الفقرة 32	** 0.99	** 3.00	** 0.46
الفقرة 33	0.50	** 1.84	0.32
الفقرة 34	0.59	** 2.02	0.31
الفقرة 35	** 0.65	** 1.35	0.16
الفقرة 36	-0.11	** 0.99	0.21

المعْلَم / رقم الفقرة	الصعوبة	التمييز	التخمين
الفقرة 37	** 0.74	** 2.27	** 0.37
الفقرة 38	** 1.49	** 3.81	0.25
الفقرة 39	** 2.15	** 1.25	0.12
الفقرة 40	** 0.93	** 1.59	0.25
المتوسط	0.59	** 1.50	0.22
أعلى قيمة	** 2.22	** 3.81	** 0.50
أقل قيمة	** 0.81	* 0.28	0.00

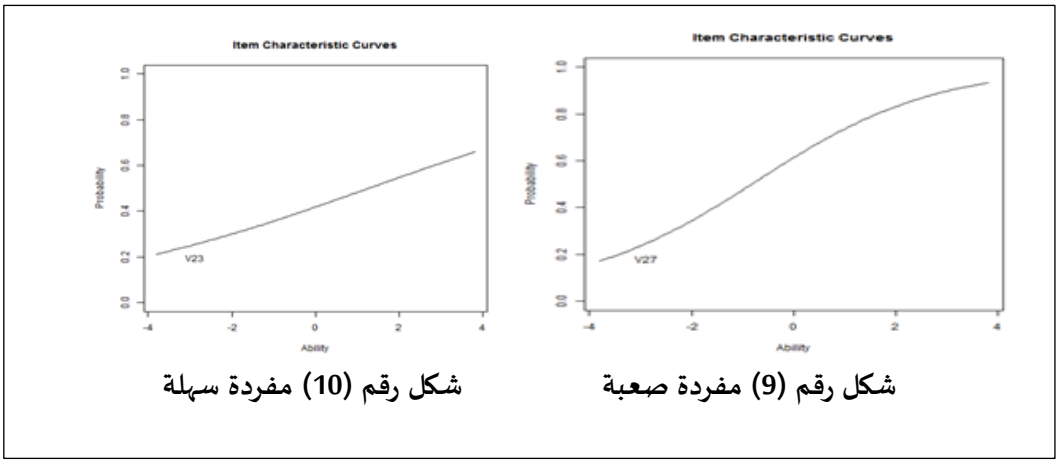
ملاحظة: في معْلَم الصعوبة *تعني منخفضة القيمة، و **تعني مرتفعة القيمة وفقًا لمعيار Choi.

وفي معْلَم التمييز والتخمين * تعني منخفضة القيمة، و **تعني مرتفعة القيمة وفقًا لمعيار Baker.

ويتضح من الجدول رقم (3) أن معْلَم الصعوبة تراوحت قيمته من (-0.81 إلى 2.22) بمتوسط قدره (0.59) بينما معْلَم التمييز تراوحت قيمته من (0.28 إلى 3.81) بمتوسط قدره (1.5) أما معْلَم التخمين تراوحت قيمته من (0.00 إلى 0.50) وبمتوسط قدره (0.22).

ووفقًا لما ذكره (حمادنة، 2009) بشأن ما أشار إليه هامبلتون (Hambleton,1989) بأنه عندما يكون معْلَم الصعوبة تتراوح قيمته من (-2 إلى 2) ومعْلَم التمييز تتراوح قيمته من (0.40 إلى 2) ومعْلَم التخمين من (0 إلى 0.25) فإن المعالم تكون متسقة وكما نلاحظ بأن متوسطات المعالم تقع بين القيم المذكورة ولكن هناك بعض الحالات التي تشذ عن تلك القيم مثل الفقرة ذات الرقم (23 و 39) واللتان قيمة صعوبتهما أكبر من (2) وكذلك هناك بعض الفقرات قيم معامل التمييز لها غير متسقة مثل الفقرات (7، 11، 12، 22، 23، 32، 34، 37، 38)، أما بالنسبة لمعْلَم التخمين الفقرات الغير متسقة هي ذوات الأرقام (7، 11، 28، 32).

أما النفيعي (2012) فيشير إلى أن تشوي (Choi,1992) وضع معياراً يعبر عن مدى صعوبة وسهولة الفقرة حيث تعتبر الفقرة سهلة إذا كانت أقل من (-0.50) وصعبة إذا كانت أكبر من (+0.50) بينما تكون متوسطة الصعوبة إذا كانت تتراوح من (-0.50إلى +0.50) وعند النظر لقيم معْلَم الصعوبة نجد أن (19) فقرة أي ما نسبته (0.53) من فقرات الاختبار أكبر من (0.50) مما يعني أنها تتمتع بالصعوبة المرتفعة، بينما الفقرات التي قلت عن (-0.50) فهما فقرتان فقط مما يعني أنها تتمتع بالسهولة، وباقي الفقرات وعددها (15) فقرة تقع بين القيمتين المشار إليها، وذلك يعني أنها متوسطة الصعوبة، ويوضح الشكلين التاليين قيمتي معْلَم الصعوبة لأكبر وأقل قيمة:



شكل رقم (9) مفردة صعبة

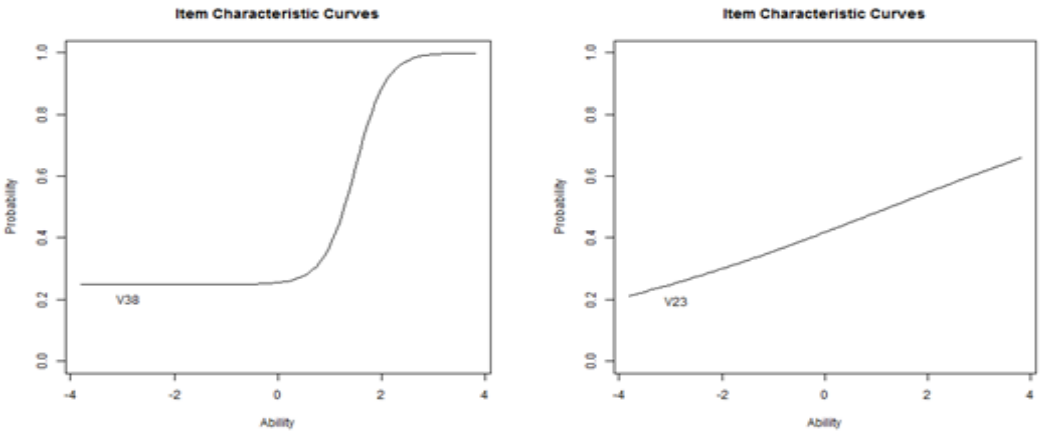
شكل رقم (10) مفردة سهلة

أما بيكر (Baker,2001) فقد وضع معيارًا لمعْلَم التمييز يمكن توضيحه من خلال الجدول التالي:

جدول رقم (4) معيار بيكر لتقييم جودة معلم التمييز للفقرات

مدى القيم	المستوى
صفر	لا يميز
0.34 -0.01	منخفض جدًا
0.64 -0.35	متوسط
1.69 – 0.65	مرتفع
أكبر من 1.7	مرتفع جدًا

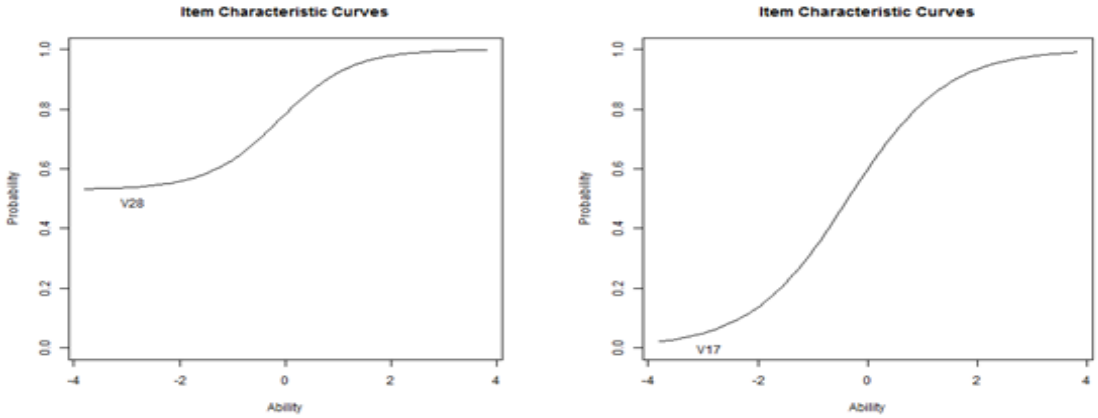
وبالنظر إلى الجدول رقم (4) نجد أن هناك فقرة واحدة تقع ضمن المستوى المنخفض جدًا و (3) فقرات ضمن المستوى المتوسط، بينما هناك (21) فقرة كانت مرتفعة التمييز أما باقي الفقرات وعددها (10) فقرات كان تمييزها أكبر من (1.7) أي مرتفع جدًا ويوضح الشكلين التاليين الفقرتين ذات أعلى قيمة تمييز وأقل قيمة في الاختبار:



شكل رقم (11) مفردة رقم (23) ذات تمييز منخفض شكل رقم (12) مفردة رقم (38) ذات تمييز مرتفع

ويتضح من الفقرة رقم (23) بأنها لا تميز بشكل مقبول بين ذوي القدرات المنخفضة وذوي القدرات المرتفعة حيث أن احتمالية الإجابة عليها بين ذوي القدرات المختلفة في هذه الفقرة متقاربة نوعًا ما والتمييز ليس كبيرا مقارنة بالفقرة المجاورة، حيث أنه بعكس الفقرة السابقة نلاحظ أن هناك تمييز كبير يتضح في الفقرة رقم (38) وذلك بين ذوي القدرات المختلفة ويظهر ذلك في مقاربة تعامد المنحنى مع محور السينات.

كما اشار بيكر (Baker,2001) إلى أن القيمة النظرية لمعْلَم التخمين تقع بين المدى (0-1)، لكن عمليًا القيمة التي أكبر من (0.35) تعتبر غير مقبولة، وعند النظر إلى جدول تقدير معالم المفردات نجد أن هناك (7) فقرات تخطت القيمة المذكورة، (3) فقرات منها كانت قريبة من (0.35) أما المفردات الأربعة المتبقية كانت قيمتهن تتراوح بين (0.423 إلى 0.498) ويبيِّن الشكلين التالين الفقرات التي يحملن أكبر وأقل قيمة تخمين:



شكل رقم (13) فقرة رقم (17) ذات تخمين منخفض جدا شكل رقم (14) فقرة رقم (28) ذات تخمين مرتفع جدا

يتضح من الشكل الذي يحمل الفقرة رقم (17) عدم وجود أو ضعف التخمين في الفقرة، بخلاف الفقرة رقم (28) والذي تظهر بأن التخمين عالي جدًا، حيث أن نصف الطلبة تقريبًا من ذوي القدرات المنخفضة أجابوا على هذه الفقرة إجابة صحيحة عن طريق التخمين، لذلك هناك خلل ما في مثل هذه الفقرة ويجب إعادة النظر فيها، حيث أنها تجاوزت القيمة المقبولة والتي ينبغي ألا تتجاوز (0.35) كما أشار لذلك بيكر (Baker).

- **رابعًا: وللإجابة على السؤال القائل ما كمية المعلومات التي يقدمها الاختبار عند مستويات القدرة المختلفة؟**

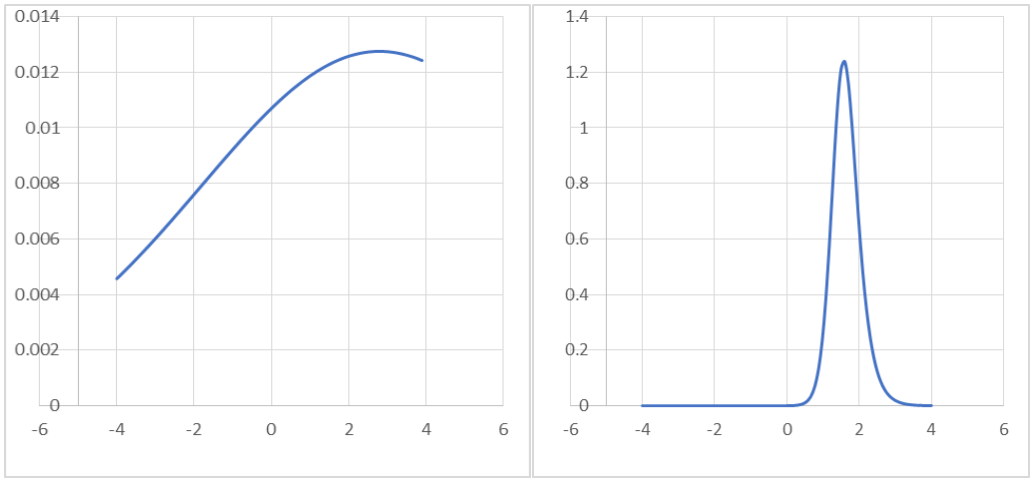
وتوضح دالة المعلومات دقة تقدير القدرة من خلال كمية المعلومات المقدمة، وقد تم حساب أقصى كمية من المعلومات والتي تقدمها كل فقرة عند ذوي القدرات المختلفة وأعطت النتائج كما هو موضح بالجدول أدناه:

جدول رقم (5) قيمة مستوى القدرة الذي تقدم عنده الفقرة أقصى كمية معلومات

رقم الفقرة	كمية المعلومات	مستوى القدرة	رقم الفقرة	كمية المعلومات	مستوى القدرة
الفقرة 1	0.09	1.4	الفقرة 22	0.67	1.1
الفقرة 2	0.38	0.1	الفقرة 23	0.01	2.8
الفقرة 3	0.24	0	الفقرة 25	0.16	0.7-
الفقرة 5	0.05	1.6	الفقرة 26	0.09	0.4
الفقرة 6	0.06	0	الفقرة 27	0.08	0.8-
الفقرة 7	0.14	1.2	الفقرة 28	0.05	0
الفقرة 8	0.06	1.8	الفقرة 29	0.11	0.2
الفقرة 11	0.22	0.6	الفقرة 30	0.10	1.1
الفقرة 12	0.49	1	الفقرة 31	0.32	0.5
الفقرة 13	0.15	0.7	الفقرة 32	0.27	1.1
الفقرة 14	0.19	1.1	الفقرة 33	0.21	0.7
الفقرة 15	0.07	1.5	الفقرة 34	0.27	0.8
الفقرة 16	0.31	0.7	الفقرة 35	0.23	0.8

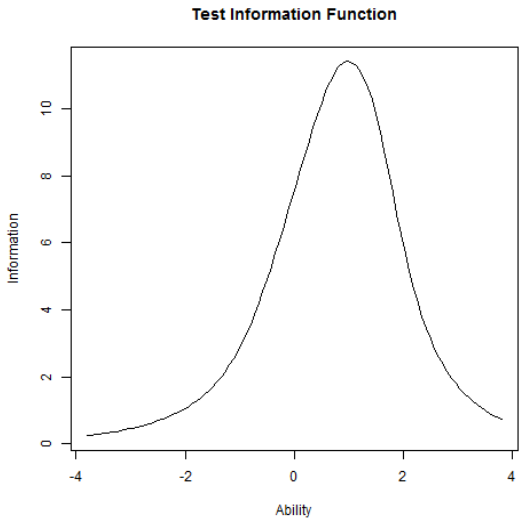
رقم الفقرة	كمية المعلومات	مستوى القدرة	رقم الفقرة	كمية المعلومات	مستوى القدرة
الفقرة 17	0.33	0.3-	الفقرة 36	0.10	0.2
الفقرة 18	0.20	0.3	الفقرة 37	0.25	0.9
الفقرة 19	0.36	0	الفقرة 38	1.24	1.6
الفقرة 20	0.09	0.8	الفقرة 39	0.24	2.3
الفقرة 21	0.20	1.5	الفقرة 40	0.22	1.1

عند مطالعة النتائج في الجدول رقم (5) نجد أن هناك تفاوت في كمية المعلومات التي تقدمها الفقرات، فإذا نظرنا إلى الفقرة التي تقدم أكبر كمية معلومات نجد أنها الفقرة رقم (38) عند مستوى (1,6) من القدرة بوحدة اللوجيت، بينما الفقرة التي تقدم أقل كمية معلومات هي الفقرة رقم (23) أي عند ذوي القدرات المرتفعة ويمكن ملاحظة ذلك من خلال منحنيات معلومات الفقرة والتي يظهره الشكلين التالين:



شكل رقم (15) منحنى دالة معلومات الفقرة رقم (23) شكل رقم (16) منحنى دالة معلومات الفقرة رقم (38)

كما تعطي دالة معلومات الاختبار الشكل التالي:



شكل رقم (17) دالة معلومات الاختبار

من الواضح من الشكل أن الاختبار يقدم أكبر قدر من المعلومات عند ذوي القدرة فوق المتوسط حيث بلغت قيمة أقصى قيمة للمعلومات التي يقدمها الاختبار ما يقارب (11.8) عند مستوى (1) من القدرة، وهذا يعني أن الاختبار يلائم الطلبة ذوي القدرات فوق المتوسط ليعطي نتائج قياس دقيقة، كما نلاحظ أن أقل كمية يمكن أن يقدمها الاختبار تكون عند ذوي مستويات القدرة المرتفعة جدًا وذوي القدرة المنخفضة، بمعنى أن الاختبار يعطي معلومات أقل عند ذوي القدرات المرتفعة جدًا والمنخفضة.

التوصيات والمقترحات:

1- ضرورة إعادة النظر في فقرات الاختبار الغير متسقة، ومراجعة الفقرات ذات التخمين المرتفع التي يوضحها جدول رقم (4)، لإعطاء الفقرات دقة أكثر في تقدير قدرة الأفراد.
2- تدريب المعلمين على كيفية صناعة الاختبارات الموضوعية.
3- توفير برامج حاسوبية مبسطة للمعلّمين تزودهم بالخصائص السيكومترية للاختبار وفق نظرية الاستجابة للفقرة.
4- القيام بدورات للتعريف والتدريب على كيفية استخدام نظرية الاستجابة للمفردة للمختصين بمجال الاختبارات.
5- إجراء دراسة للاختبار المركزي الحالي وفق النظرية الكلاسيكية والمقارنة بين نتائجها ونتائج ما توصلت له هذه الدراسة.
6- إجراء دراسة للخصائص السيكومترية لبقية الاختبارات المركزية وفق النماذج المناسبة من نماذج نظرية استجابة الفقرة.
7- استخدام نماذج نظرية الاستجابة للفقرة في بناء وتصميم الاختبارات التحصيلية لمواكبة التطورات العالمية في هذا المجال للاستفادة من مزاياها.

قائمة المراجع:

أولاً- المراجع بالعربية:

- أبو جراد، حمدي يونس. (2012). الخصائص السيكومترية لاختبار الاختيار من متعدد والتكميل: دراسة مقارنة باستخدام نظرية الاستجابة للمفردة، **مجلة العلوم التربوية والنفسية**، جامعة البحرين، 13 (3)، 375 -404.
- أبو هاشم، السيد محمد. (2006). دراسة مقارنة بين النظرية التقليدية ونموذج راش في اختبار فقرات مقياس مداخل الدراسة لدى طلاب الجامعة، **مجلة كلية التربية**، جامعة الزقازيق، 21 (52)، 1- 52.
- الأشقر، وفاء و عيلبوني سمير. (2015). فحص تحقق فرض الاستقلال الموضعي بتحقق فرض أحادية البعد باختلاف حجم العينة وطول الاختبار في نظرية استجابة الفقرة، **مجلة كلية التربية**، جامعة عين شمس، مصر، 39 (3)، 477-532.
- بيكر، فرانك. (2010). **أسس نظرية الاستجابة للمفردة، (**ترجمة: عبد الرحمن الطريري والسيد أبو هاشم و سوسن شلبي)، دار جامعة الملك سعود للنشر، الرياض.
- الجلالي، لمعان مصطفى. (2011). **التحصيل الدراسي**، دار المسيرة، عمان.
- الحاج، فريال محمود. (2016). الخصائص السيكومترية لاختبار تحصيلي محكي المرجع في مقرر علم النفس التربوي وفق نموذج راش، **المجلة التربوية**، الكويت، 120 (30)، 171- 200.

- حمادنه، إياد. (2009). استخدام نظرية الاستجابة للمفردة في بناء اختبار محكي المرجع في الرياضيات وفق النموذج اللوجستي ثلاثي المعلم، **مجلة العلوم التربوية والنفسية**، 10 (2)، 215- 238.
- الحموري، هند عبد المجيد. (2011). دراسة استكشافية لملائمة نماذج نظرية الاستجابة للمفردة في بناء اختبار تحصيلي من إعداد المعْلَم، **مجلة العلوم التربوية والنفسية**، جامعة البحرين، 2 (12)، 47- 81.
- الدوسري، راشد حماد (2004). **القياس والتقويم التربوي الحديث مبادئ وتطبيقات وقضايا معاصرة** (الطبعة الأولى). دار الفكر، عمان.
- الزبون، حابس والصرايرة، راجي. (2017). أثر ترتيب فقرات اختبار الاختيار من متعدد وفقا لمعالم الفقرة على تقدير قدرة المفحوصين ودالة المعلومات للاختبار وفقا للنموذج ثلاثي المعْلَم ة، **مجلة جامعة الحسين بن طلال للبحوث**، 1 (3)، 191- 206.
- الشريف، خالد وفراج، محمد. (2013). فاعلية استخدام كل من نظرية السمات الكامنة والنظرية الكلاسيكية في قياس التحصيل الدراسي في الجبر بالمرحلة الإعدادية: دراسة مقارنة، مجلة العلوم الإنسانية والاجتماعية، **جامعة الامام محمد بن سعود الإسلامية**، (28)، 89- 147.
- الشريفين، أحمد وحجازي، تغريد. (2014). تقصي الخصائص السيكومترية لاختبار ذكاء جمعي حسب نظرية السمات الكامنة، **مجلة الدراسات التربوية والنفسية**، كلية التربية، جامعة السلطان قابوس، 1 (8)، 1- 14.
- الشريفين، نضال كمال. (2012). أثر طريقة تقدير معالم الفقرة وقدرات الافراد على قيم معالم الفقرة والخصائص السيكومترية للاختبار في ضوء تغير حجم العينة، **المجلة التربوية**، جامعة الكويت، 26 (104)،177 - 238.
- الشريفين، نضال كمال. (2017). خصائص توزيع قدرات الافراد ومعالم فقرات الاختبار وفق نماذج نظرية الاستجابة للفقرة المعْلَمية واللامعْلَمية: دراسة مقارنة، **المجلة التربوية**، جامعة الكويت، 31 (124)، 265- 320.
- الشمراني، صالح. (2012). أثر التفاعل بين نمط الاختبار وجنس المستجيب على ثبات الاختبار وبعض الخصائص السيكومترية للفقرات، **رسالة ماجستير غير منشورة**، جامعة أم القرى، مكة المكرمة، السعودية.
- الشمراني، محمد موسى. (2014). استخدام نظرية استجابة الفقرة والنظرية التقليدية في تقدير الخصائص الإحصائية السيكومترية لاختبار مهارات التفكير والتعلم لطلبة السنة التحضيرية الجامعية، **مجلة التربية**، جامعة الأزهر، 2 (157)، 717- 802.
- الصالح، علي وسليمان، شاهر. (2016). بناء اختبار تحصيلي محكي المرجع في القياس والتقويم باستخدام النموذج اللوقارتمي ثلاثي المعْلَم لنظرية الاستجابة للمفردة الاختبارية، **مجلة التربية للبحوث التربوية والنفسية والاجتماعية**، كلية التربية، جامعة الأزهر، 168 (1)، 175- 201.
- عبد الرحمن، أحمد محمد. (2011). **تصميم الاختبارات: أسس نظرية وتطبيقات عملية**، دار أسامه للنشر، عمان.
- عبيدات، ذوقان وعدس، عبدالرحمن وعبد الحق، كايد. (2000). **البحث العلمي: (مفهومه، أدواته، أساليبه)**، دار أسامه للنشر والتوزيع، الرياض.
- علام، صلاح الدين. (2000). **القياس والتقويم التربوي والنفسي**، دار الفكر العربي، القاهرة.
- علام، صلاح الدين. (2005). **نماذج الاستجابة للمفردة الاختبارية أحادية البعد ومتعددة الابعاد وتطبيقاتها في القياس النفسي والتربوي**، دار الفكر العربي، القاهرة.
- علام، صلاح الدين. (2012). **الاختبارات والمقاييس التربوي والنفسية**، دار الفكر، عمان.

- علي، نداء. (2012). فاعلية استخدام انموذج راش في بناء اختبار تحصيلي محكي المرجع لمقرر القياس والتقويم. **رسالة دكتوراه غير منشورة**، جامعة دمشق، سوريا.
- العنزي، خالد طخيطيخ. (2018) "اثر طول الاختبار في دقة معادلة درجات صورتي اختبار في الرياضيات باستخدام نظرية استجابة المفردة، " **المجلة العربية الاجتماعية**، المؤسسة العربية للاستشارات العلمية و تنمية الموارد البشرية، مصر، 13، 19-55.
- الفريحات، أيمن محمد. (2017). الخصائص السيكومترية لاختبار محكي المرجع في الكيمياء وفق النظرية الحديثة في القياس النفسي والتربوي، **مجلة المشكاة**، جامعة الزيتونة، 15، 343-368.
- القضاة، بسمة عواد (2010) بناء اختبار تحصيلي في الرياضيات للصف التاسع الأساسي وتقدير خصائصه السيكومترية وفق النظرية الكلاسيكية ونموذج المعْلَم تين للنظرية الحديثة، **رسالة ماجستير غير منشورة**، جامعة مؤتة، الأردن.
- القفاص، وليد كمال. (2011). **التقويم والقياس النفسي والتربوي اتجاهات –معاصرة- برامج تدريبية لإعداد وتعريب الاختبارات**، المكتب الجامعي الحديث، القاهرة.
- القيسي، حسين. (2013). دقة تقدير معالم الفقرة والقدرة باستخدام نماذج نظرية استجابة الفقرة البارامترية واللابارامترية باختلاف حجم العينة وطول الاختبار، رسالة **دكتوراه غير منشورة**، جامعة اليرموك، الأردن.
- كراطي، ندى بدر. (2016). **لغة البرمجة R للحوسبة الإحصائية**، مطبعة البصرة/دار الكتب والوثائق، بغداد.
- المالكي، يحي. (2015). درجة التزام معلّم ي المرحلة المتوسطة بقواعد بناء الأسئلة الموضوعية، **رسالة ماجستير غير منشورة**، جامعة أم القرى، المملكة العربية السعودية.
- الموسوي، نعمان محمد. (2014). الخصائص السيكومترية لمقياس مهارات التواصل الالكتروني للمراهقين في ضوء نظريتي القياس التقليدية والحديثة، **مجلة الطفولة العربية**، الكويت، 39 (15)، 9- 34.
- الناغي، هبة إبراهيم. (2011). أثر عدد مفردات الاختبار على تقدير قدرات الأفراد ودالة المعلومات لاختبار تحصيلي مرجعي المحك في الرياضيات باستخدام نماذج نظرية الاستجابة للمفردة، **مجلة كلية التربية**، جامعة بورسعيد، (10)، 676- 721.
- النفيعي، عبد الرحمن عبد الله. (2012). الخصائص السيكومترية لاختبار المصفوفات المتتابعة المتقدم في ضوء نظرية الاستجابة للمفردة الاختبارية، **مجلة التربية**، جامعة الأزهر، 147 (2)، 175- 214.

ثانياً- المراجع بالإنجليزية:

- Baker, F. (2001). **The Basics of Item Response Theory**. ERIC Clearinghouse on Assessment and Evaluation. United States of America.
- Brown, F. (1978). **Principes of Educational and psychological Testing** (2nd ed.).ILL, Charles Thomas.
- Hambleton, R. & Swaminathan, H. (1985). **Item response theory principles and applications**. Boston: Kluwer Nijhoff Publishing.
- Hambleton, R. K., & Traub, R. (1971). Information Curves and Efficiency of Three Logistic Test Model**. British Journal of Mathematical and Statistical Psychology**, 24, 273-281.
- Mislevy, R. J. & Bock, R. (1990). PC-Bilog: **Item Analysis and Scoring With Binary**, Logistic Model Mooresville, IN, Scientific Software.

- Ojerinde, D. (2013). **Classical Test Theory VS Item Response Theory: an Evaluation of the comparability of item analysis results**. Nigeria, Abuja, Joint Admission and Matriculation Board (JAMB).
- Panayides, P., & Walker, M. (2013). Evaluating the psychometric properties of the foreign language classroom anxiety scale for Cypriot senior high school EFL students: the Rasch measurement approach, **Europe's Journal of Psychology**, 9 (3), 493-516.
- Reeve, B.B. (2002). **An Introduction to Modern Measurement Theory**. Applied Research Program, Division of Cancer Control and Population Sciences, 1-67.
- Traub, R. (1997). Classical test theory in historical perspective**. Educational Measurement: Issues and Practice**, 16 (4), 8-14.
- Warm, T.A. (1978). **A Primer of Item Response Theory**, Technical Report No.940279, Oklahoma City: U.S. Coast Guard Institute.
- Weiten, W. (1984). Violation of selected item construction principles in educational measurement**. Journal of Experimental Education**, 52 (3), 78-174.