تقييم مدى مطابقة فقرات اختبار علم النفس للنظرية الكلاسيكية في القياس ونظرية الاستجابة للفقرة وفقا للنموذج اللوجستي ثنائي المعلم (دراسة مقارنة)

Evaluation of the extent of conformity of psychology test items to classical test theory and item response theory according two- Parameter Logistic Model

(Comparison Study)

عامر جبريل المرابحة[1،*] عفيف بن صالح[2]

[1]قسم طب العائلة والمجتمع، كلية الطب والعلوم الطبية، جامعة الخليج العربي، مملكة البحرين

Afifbs@agu.edu.bh[2] amerjka@agu.edu.bh[1*]

تاريخ الاستلام:15-09-2020 تاريخ القبول: 20-01-2021 تاريخ النشر : 31-03-2021

ملخص: هدفت الدراسة إلى الكشف عن مدى مطابقة فقرات اختبار علم النفس مع افتراضات النظرية الكلاسيكية في القياس ونظرية الاستجابة للفقرة باستخدام النموذج اللوجستي ثنائي المعلم، ولتحقيق أهداف الدراسة تم استخدام نتائج طلبة السنة الأولى في كلية الطب والعلوم الطبية في جامعة الخليج العربي على اختبار علم النفس للعام الدراسي 2018-2019 والبالغ عددهم (203) طالباً وطالبة، أشارت النتائج إلى أن نسبة عدد فقرات الاختبار المطابقة لافتراضات النظرية الكلاسيكية في القياس شكلت ما نسبته (79%)، ونسبة عدد الفقرات المطابقة لافتراضات النموذج اللوجستي ثنائي المعلم بلغ (87.9%)، وبلغت نسبة التوافق في عدد الفقرات المطابقة لافتراضات كلا النظريتين (72.7%)، لذا فإن مجمل النتائج أشارت إلى أن تحليل فقرات الاختبار وانتقائها باستخدام نظرية الاستجابة للفقرة يعد أفضل مقارنة بالنظرية الكلاسيكية في القياس، وتوصي الدراسة بضرورة إجراء دراسة مشابهة للدراسة الحالية باختلاف بعض المتغيرات: كنوع النموذج المستخدم، حجم العينة، وعدد فقرات الاختبار.

الكلمات المفتاحية: اللوجستي ثنائي المعلم؛ معالم الفقرة؛ قدرة الأفراد.

Abstract: This study aimed at assessment the extent of conformity of multiple-choice questions items with assumption of Classical Test Theory and Item Response Theory using Two-Parameter Logistic model. To achieve the goals of the study, the researchers use the final exam results of psychology course for year 1 students in the college of medicine and medical science in Arabian Gulf University for academic year 2018-2019. The test was administered on a sample consisting of (203) students. The results showed that (79%) of items fit the assumptions of Classical Test Theory. According the Chi-Square Goodness Fit, the results showed that (87.9%) of items fit the assumptions of Two-Parameter Logistic model, and (72.7%) of items fit both two theories. Therefore, the overall results indicate that analysis of (MCQs) items using the Item Response Theory according to the two-parameter logistic model is better compared to the Classic Test Theory. The study recommends to conducting a study like the current study according to type of logistic model, sample size, number of items.

Keywords: two Logistic model; item parameters; pearson ability.

*المؤلف المراسل:

1- مقدمة

اهتم علماء النفس منذ أن ظهرت حركة القياس النفسي، بتحقيق صدق وثبات الاختبارات والمقاييس النفسية؛ سعياً منهم لتحقيق أعلى درجة من الموضوعية في هذه الأدوات عند استخدامها في عملية القياس، ووفقاً لنظرية القياس الكلاسيكية (التقليدية) Classical Test Theory (CTT) يمكن التعبير عن قدرة الفرد من خلال الدرجة الظاهرية (المشاهدة) التي تظهر من أدائه على الاختبار، وبالتالي يختلف مستوى القدرة باختلاف مستوى صعوبة الاختبار، وقد أسفرت جهود العلماء عن ظهور بعض الاتجاهات الحديثة في مجال القياس والتقويم ومن هذه الاتجاهات نظرية الاستجابة للفقرة Item Response Theory (IRT) وحظى هذا المدخل باهتمام الباحثين؛ حيث تغلّب على الكثير من مشكلات القياس الكلاسيكية (الخياط، 2012)، وتفترض الاختبارات النفسية والتربوية بعامة أن هناك سمات أو خصائص معينة يشترك فيها الأفراد جميعهم، ولكنهم يختلفون في مقدارها، وبالرغم من أن هذه السمات غير محسوسة، إلا أنه يمكن الاستدلال على مقدارها من خلال السلوك الملاحظ للفرد المتمثل في استجاباته على فقرات الاختبار، وتتلخص الفكرة الأساسية لنظرية الاستجابة للفقرة في أنها تحاول اشتقاق قيم تقديرية للسمات الكامنة التي تنطوي عليها مجموعة من الاستجابات لمجموعة من الفقرات (Jimelo & Silvestre, 2009)، ويتضمن بناء وتطوير الاختبارات باستخدام النظرية الكلاسيكية في القياس انتقاء الفقرات وفقاً للمحتوى والخصائص المتعلقة بمعلمي الصعوبة والتمييز، وفاعلية البدائل؛ حيث يتم انتقاء الفقرات ذات الدلالة التمييزية العالية، ومستوى الصعوبة المناسب، والذي يناسب الغرض من إعداد الاختبار، وتوزيع القدرة لأفراد المجموعة التي يطبق عليها الاختبار (Wolfe & Smith, 2007).

1.1- مشكلة الدراسة وتساؤلاتها:

يعد موضوع المقارنة بين النظرية الكلاسيكية في القياس، ونظرية الاستجابة للفقرة من الموضوعات التي حازت على اهتمام الباحثين منذ ظهور نظرية الاستجابة للفقرة؛ وذلك لتبيان مدى التوافق بينهما في انتقاء فقرات مختلف الاختبارات والمقاييس، ومعرفة أي من النظريتين يمكن أن تقدم فوائد قيمة للعملية التعليمية بمختلف مراحلها، وكذلك أي من النظريتين تتمتع بقدرة أفضل على اختيار فقرات لقياس السمة التي يراد قياسها بشكل أدق، وبالتالي تكوين مقاييس أو اختبارات تتمتع بخصائص سيكومترية جيدة واكثر موضوعية، إلا أن نتاج البحوث والدراسات جاءت متناقضة، فبعضها أشار إلى تفوق مؤشر نظرية الاستجابة للفقرة على مؤشرات النظرية الكلاسيكية في القياس كدراسة (Owens,2000)، ودراسة (Kan,2006) في حين أشار بعضها الآخر إلى عدم وجود فروق بين مؤشرات النظرية الكلاسيكية في القياس ونظرية الاستجابة للفقرة، ففي ضوء هذه التناقضات بالإضافة إلى ندرة الدراسات المحلية التي بحثت في مجال المقارنة بين النظريتين الكلاسيكية ونظرية الاستجابة للفقرة باستخدام النموذج اللوجستي ثنائي المعلم الذي يعتمد على معلم الصعوبة والتمييز، وذلك في حدود علم الباحثان، لذا جاءت هذه الدراسة للمقارنة بين النظريتين من خلال استخدام اختبار تحصيلي في علم النفس؛ وذلك للتحقق من مدى التطابق أو الاختلاف في انتقاء فقرات الاختبار من حيث عددها ومستوى صعوبتها وتمييزها. ومن هنا فإنه يمكن صياغة مشكلة الدراسة من خلال الإجابة عن الأسئلة التالية:

1. ما مدى مطابقة فقرات اختبار علم النفس لافتراضات النظرية الكلاسيكية في القياس والنموذج اللوجستي ثنائي المعلم؟
2. ما مدى التوافق بين النظرية الكلاسيكية في القياس والنموذج اللوجستي ثنائي المعلم في انتقاء فقرات اختبار علم النفس؟

2.1- أهداف الدراسة:

هدفت الدراسة الحالية إلى تقييم مدى مطابقة فقرات اختبار علم النفس للنظرية الكلاسيكية في القياس ونظرية الاستجابة للفقرة، وذلك من خلال تقدير الخصائص السيكومترية لفقرات الاختبار وفقا لكلا النظريتين بالإضافة إلى الكشف عن مدى التوافق بين النظرية الكلاسيكية في القياس ونظرية الاستجابة للفقرة باستخدام النموذج اللوجستي ثنائي المعلم في انتقاء فقرات اختبار علم النفس من حيث عددها، ومستوى صعوبتها وتمييزها وتحديد الفقرات المطابقة والفقرات غير المطابقة وفقاً لافتراضات كلا النظريتين.

3.1- أهمية الدراسة:

تنبثق أهمية هذه الدراسة من جانبين: الأول نظري والثاني عملي، فمن حيث الأهمية النظرية فإن الدراسة الحالية تسهم في تقييم مدى جودة الخصائص السيكومترية لفقرات اختبار علم النفس، وذلك من خلال الكشف عن مدى مطابقة فقرات الاختبار لافتراضات النظرية الكلاسيكية في القياس، وافتراضات النموذج اللوجستي ثنائي المعلم، كما ان استخدام النموذج اللوجستي ثنائي المعلم كممثل لنظرية الاستجابة للفقرة، يضيف أهمية أخرى لهذه الدراسة في تحقيق الموضوعية والتوازن عند المقارنة، حيث تمثلت مؤشرات المقارنة الكلاسيكية في معلمتي الصعوبة والتمييز، وهي المؤشرات الموجودة في النموذج اللوجستي ثنائي المعلم، أما من حيث الأهمية العملية فإن نتائج هذه الدراسة وتوصياتها تساعد معدي الاختبارات بأفضل الطرق في تحليل فقرات الاختبارات والتي تزيد من درجة صدقها وثباتها، وتساهم في زيادة درجة الوعي لدى المهتمين بإعداد اختبارا اختيار من متعدد (MCQs)، بضرورة الوقوف على أهم شروط تحقيق الموضوعية في هذا النوع من الاختبارات، وبالتالي تساهم في تطوير أداة قياس ذات خصائص سيكومترية جيدة يمكن أن تستخدم لقياس المهارات اللازمة في علم النفس كما يمكن ان تشكّل نتائج هذه الدراسة بشكل عام النواة الأولى لبناء بنك أسئلة لمساق علم النفس.

4.1- مصطلحات الدراسة:

نظرية استجابة الفقرة: تمثل مجموعة الطرق الإحصائية التي تستخدم في حساب معالم الفقرات وقدرات الأفراد الخاصة بالاختبار والمتمثلة بمعالم الصعوبة، التمييز، التخمين، ودالة المعلومات المكونة من نماذج أحادية البعد (نموذج (راش) أحادي البارامتر، نموذج (لورد) ثنائي البارامتر، نموذج (بيرنيوم) ثلاثي البارامتر).

النموذج اللوجستي ثنائي المعلم: هو أحد نماذج نظرية استجابة الفقرة ويسمى نموذج لورد ثنائي المعلم، ويقوم بحساب معالم الفقرة (الصعوبة، التمييز)، ويفترض اختلاف الفقرات في صعوبتها وتمييزها، وغياب عامل التخمين.

مساق علم النفس: إحدى المساقات الإجبارية التي يدرسها جميع طلبة الطب البشري في كلية الطب والعلوم الطبية في جامعة الخليج العربي في السنة الدراسية الأولى/ الفصل الدراسي الأول.

معامل صعوبة الفقرة: نسبة الطلبة الذين أجابوا عن الفقرة إجابة صحيحة من بين الطلبة الذين حاولوا الإجابة.

معامل تمييز الفقرة: قدرة الفقرة على التمييز بين أداء الطلبة ذوي الأداء المرتفع والطلبة ذوي الأداء المنخفض ويتم حسابه بإيجاد معامل الارتباط بين أداء الطلبة على كل فقرة ومقارنتها بالأداء الكلي على الاختبار.

معلمة صعوبة الفقرة (b_i) وفقا لنظرية استجابة الفقرة: تمثل نقطة على متصل السمة الكامنة يكون عندها احتمال الإجابة الصحيحة على الفقرة (i) يساوي 0.05.

معلمة تمييز الفقرة (a_i) وفقا لنظرية استجابة الفقرة: تمثل ميل منحنى خصائص الفقرة الذي يحدث عنده تغيير في اتجاه المنحنى (inflexion) الذي يقابل الصعوبة على متصل القدرة والتي يكون عندها احتمال الإجابة الصحيحة عن الفقرة (i) يساوي 0.05.

اللوجيت: هي وحدة قياس كل من معلمة الصعوبة ومعلمة القدرة للمفحوصين.

5.1- محددات الدراسة:

اقتصرت عينة الدراسة على نتائج جميع طلبة السنة الأولى المسجلين لمساق علم النفس في كلية الطب والعلوم الطبية في جامعة الخليج العربي في الفصل الدراسي الأول من العام الدراسي 2018-2019، وبالتالي فإن إمكانية تعميم النتائج تتحدد بمدى تقارب نتائج هذه العينة مع نتائج طلبة السنة الأولى في السنوات السابقة.

المفاهيم والمصطلحات المستخدمة في هذه الدراسة محددة بالتعريفات الإجرائية، وبالتالي فإن إمكانية تعميم النتائج تتحدد في ضوء هذه التعريفات.

2- الإطار النظري والدراسات السابقة

1.2-الإطار النظري:

تسيطر النظرية الكلاسيكية في القياس على بناء الكثير من أدوات القياس المختلفة وتحليلها، فقد استخدمت أسس هذه النظرية في مواقف اختبارية متنوعة تتضمن بناء مختلف أنواع الاختبارات النفسية والتربوية، وكذلك تحليل البيانات المستمدة من هذه الاختبارات، تعد طرائق تحليل بيانات الاختبارات التربوية المختلفة المبنية على النظرية الكلاسيكية للاختبار شائعة الاستخدام عند معظم الباحثين، وخاصة فيما يتعلق بتقدير ثبات المقاييس والاختبارات التربوية المختلفة، وتقدير المعالم الخاصة بالفقرة التي تتمثل بصورة رئيسة بالصعوبة والتمييز، وعلى الرغم من الألفة المتزايدة بهذه النظرية، فإن المعرفة بمدى التأثيرات التي يمكن أن تنشأ من عدم تحقق الافتراضات المختلفة لنظرية الاختبار الكلاسيكية ما زالت متواضعة (النعيمي، 2007).

وتعد أسئلة الاختيار من متعدد أكثر أنواع أسئلة الاختبارات التحصيلية شيوعاً واستخداماً، سواء كان ذلك على مستوى التعليم العام أو الجامعي (Oosterhoff, 1994)، وتمتاز بسهولة تصحيحها وموضوعيتها، وتوفر تغطية جيدة للمادة الدراسية، وتمتاز علامة الطالب عليها بدرجة عالية من الثبات، إضافة إلى أنها تحدد نتائج التعلم المقصودة بدرجة عالية، على الرغم من أنّ إعدادها يتطلب وقتاً طويلاً، وجهداً كبيراً، ومهارة فائقة من قبل معدي الاختبار، كما أن فقرات الاختيار من متعدد قادرة على قياس نواتج التعلم في المستويات العقلية العليا من المجال المعرفي بدرجة تفوق فقرات المطابقة وفقرات الصواب والخطأ، وفقرات التكميل، والإجابة القصيرة (Baghaei & Amrahi,2014)، وقد زادت الحاجة إلى دقة المعلومات التي يتم الحصول عليها من تلك الاختبارات ذات الصلة بالهدف أو الغرض الذي أعد الاختبار لأجله، وخاصة أن الاختبارات لها دور أساسي في تقويم العملية التدريسية والتي تستخدم نتائجها في اتخاذ القرارات، فعدم معرفة البعض بكيفية صياغة فقرات الاختيار من متعدد يؤدي إلى خلل في بنية الفقرة، وهذا بدوره يؤثر سلباً على فاعلية الاختبار وقدرته على قياس السمة التي صمم الاختبار لقياسها، وبالتالي قد ينعكس على الخصائص السيكومترية للفقرة ممثلة بقيم الصعوبة والتمييز وعلى صدق وثبات الاختبار وبالتالي يؤثر في درجة تقويم الطلبة (الشريفين والصبح، 2011)، ويذكر (Aiken, 2003) بأن فقرات اختبار الاختيار من متعدد ذات البناء الجيد تتمتع بخصائص سيكومترية تجعلها تقيس مدى أوسع من المخرجات التعليمية، وتقلل من تباين الأخطاء، وأخطاء القياس، وتقيس الأهداف البسيطة والمركبة، وتزيد من فهم المفحوصين وبالتالي زيادة درجاتهم على الاختبار (الشريفين، 2011).

وتعد مقارنة الخصائص السيكومترية لفقرات اختبار من نوع اختيار من متعدد وفقا للنظرية الكلاسيكية في القياس، ونظرية الاستجابة للفقرة من المواضيع الهامة لدى الباحثين في المجال التعليمي من أجل التعرف على مدى مطابقة فقرات الاختبارات وفقا لافتراضات كلا النظريتين، ووفقا لذلك أشار عودة (2010) إلى أنّ خصائص الاختبار والفقرة وفق النظرية الكلاسيكية في القياس تتأثر بموقف الاختبار وخصائص العينة، أما نظرية الاستجابة للفقرة فإنها تميز بين العلامة الحقيقية وعلامة القدرة، بمعنى أن قدرة المفحوص ثابتة في موقف معين ولا تتأثر بخصائص الفقرات، بينما تكمن الفكرة الأساسية لنظرية الاستجابة للفقرة في ربطها بين خصائص الفقرات وبين مقدار السمة التي يمتلكها الفرد واحتمال إجابته عن فقرة محددة بمستوى معين من الإجابة وفقاً لنوع الفقرة (ثنائية أو متعددة)، ويقوم هذا الرابط على تحديد موقع الفرد على مقياس السمة بصرف النظر عن مجموعة الفقرات التي يجيب عنها الفرد.

وتتفق كل من النظرية الكلاسيكية ونموذج الاستجابة للفقرة في افتراض وجود متصل للسمة وأنه يمكن تقدير احتمال إجابة فرد إجابة صحيحة عن فقرة اختبارية إذا علمنا موقعه على هذا المتصل، كما تتفق هذه النماذج في أن مقدار هذا الاحتمال يكون دالة متزايدة مطردة لمواقع الأفراد على متصل السمة، وهذا يعني أن احتمال توصل الفرد إلى الإجابة الصحيحة عن فقرة اختبارية يزداد بازدياد مقدار السمة لديه (علام،1987).

وتعد النظرية الكلاسيكية في القياس نموذج بسيط يحتوي تعريفاً لمفهوم الدرجة الظاهرة، والدرجة الحقيقية والخطأ العشوائي في القياس، وتنطلق من مجموعة من الافتراضات منها على سبيل المثال: أن الدرجة التي يحصل عليها الفرد في اختبار ما (الدرجة الملاحظة) Observe Score هي حاصل جمع الدرجة التي يستحقها الفرد (الدرجة الحقيقية) True Score مع الدرجة التي تنشأ عن أخطاء القياس (الدرجة الخطأ) Error أو ما يسمى بالخطأ المعياري للقياس، كما تفترض النظرية الكلاسيكية في القياس بأن أداء الفرد يمكن قياسه وتقديره، وأن أداء الفرد هو دالة لخصائصه، وأن الخاصية والأداء والعلاقة بينهما تختلف من فرد إلى آخر (أبو عواد والقهوجي، 2016).

وتفترض النظرية الكلاسيكية أيضا بأن الخطأ المعياري في القياس متساوي لجميع الطلبة الممتحنين، ووفقا لهذا الافتراض فإن التعبير عن قدرة الفرد يتم من خلال الدرجة الحقيقية التي تتضح من خلال أدائه على الاختبار ككل، وليس على مستوى الفقرة، وبالتالي سيتغيّر وضع قدرة الفرد وفق تغير مستوى الاختبار، وعلاوة على ذلك فإن الاختبار والفقرات تتغيّر خصائصهما بتغيّر خصائص الأفراد، كما أن خصائص الأفراد تتغيّر بتغيّر خصائص الاختبار من حيث الصعوبة والسهولة (الشريفين، 2006).

وظهرت النظرية الحديثة في القياس أو ما يسمى بنظرية الاستجابة للفقرة على أنها نظرية بديلة عن النظرية الكلاسيكية، والتي يفترض فيها تلافي عيوب النظرية الكلاسيكية، كما أنها تعد ثمرة محاولات المهتمين في القياس لتطوير مقاييس واختبارات أكثر دقة، وكمحاولة لحل المشكلات التي تعانيها النظرية الكلاسيكية، كما أنها قادرة على معالجة الكثير من القضايا المتعلقة بالقياس بشكل أكثر فاعلية من النظرية الكلاسيكية.، حيث تفترض هذه النظرية إمكانية التنبؤ بأداء الأفراد، وتفسير أدائهم في الاختبار في ضوء خاصية أو خصائص مميزة لهذا الأداء تسمى السمات (Traits)، وتحاول كذلك تقدير درجات الأفراد على هذه السمات (Hamblton & Swaminathan, 1985)

وتحدد نظرية الاستجابة للفقرة العلاقة بين الدرجة المشاهدة التي يحصل عليها الطالب عند إجابته عن فقرة ما وبين السمة الكامنة لديه والتي بناءً عليها يختار الطالب إجابته، ويتم التعبير عن هذه العلاقة من خلال دالة

رياضية، وجدير بالذكر بأن نماذج نظرية الاستجابة للفقرة تستند إلى بعض الافتراضات التي من الواجب تحققها في البيانات قبل الشروع بعمليات التحليل الأخرى لمعالم قدرات الأفراد والفقرات، ومن هذه الافتراضات: أحادية البعد (Unidimensional) بمعنى أن جميع الفقرات المكونة للاختبار ينبغي أن تقيس سمة واحدة مشتركة تسمى عادةً بقدرة الفرد(θ)، الاستقلال الموضعي (Local independence) أي أن استجابة الفرد على فقرة معينة في الاختبار لا تؤثر سلباً أو إيجاباً على استجابته على فقرات أخرى في الاختبار، عامل السرعة (Speediness) بمعنى أن إخفاق الأفراد في الاختبار يرجع إلى انخفاض قدرتهم وليس إلى تأثير عامل السرعة على الإجابة (Hamblton & Swaminathan, 1985).

وقد انبثقت عن هذه النظرية مجموعة من النماذج والتي تعرف بنماذج السمات الكامنة (Latent Trait Models) أو نماذج نظرية الاستجابة للفقرة، وتهدف جميعها إلى تحديد علاقة أداء الفرد على الاختبار وبين السمات أو القدرات التي تكمن خلف هذا الأداء وتفسره، وتهدف هذه النماذج لتحديد العلاقة بين أداء الفرد الممتحن (الطالب) في اختبار معين وبين القدرة التي تكمن وراء هذا الأداء وتفسيره، وتعد هذه النماذج دوال رياضية احتمالية تختلف باختلاف عدد معالمها، فالنموذج اللوجستي (اللوغاريتمي) أحادي المعلم One-Parameter Logistic Model (1PLM) ويسمى نموذج (راش) (Rasch model) يفترض أن جميع الفقرات لا تختلف عن بعضها إلا بعامل وصفة واحدة هي صعوبة الفقرات (b_i)، أما النموذج النموذج اللوجستي ثنائي المعلم (2PLM) Two- Parameter Logistic Model ويسمى بنموذج لورد (Lord Model) وهو النموذج المستخدم في الدراسة الحالية، ويعد هذا النموذج أكثر النماذج اتفاقا مع المؤشرات الكلاسيكية، بالإضافة إلى أفضليته في انتقاء الفقرات مقارنة بالنماذج الأخرى ويفترض هذا النموذج أن جميع فقرات الاختبار لا تختلف فيما بينها إلا بعاملي الصعوبة والتمييز ويتضح ذاك من المعادلة التالية:

$$pi(\theta_j) = \frac{e^{Dai(\theta j - bi)}}{1+e^{Dai(\theta j - bi)}} \quad \ldots\ldots\ldots\ldots\ldots (1)$$

ai: تمثل معلم تمييز الفقرة (i) (ميل منحنى خصائص الفقرة عند نقطة انعطاف المنحنى).

D: معامل التدريج المستخدم لتحويل $pi(\theta_j)$ من النموذج اللوجستي إلى النموذج الطبيعي Normal Ogive وقيمته ثابته = 1.702 وعند استخدام معامل التدريج D فإن قيمة القدرة (θ) تخضع للقياس الطبيعي وتتراوح عادة بين (±3) (Toland, 2008)، وفي النموذج اللوجستي (اللوغاريتمي) ثلاثي المعلم (3PLM) يضاف معلم التخمين (Guessing) أو ما يسمى بمعلم التقارب السفلي (Lower-asymptote) لمنحنى خصائص الفقرة.

وقد تم استخدام النموذج اللوجستي ثنائي المعلم في الدراسة الحالية؛ لكون المؤشرات الإحصائية التي يتم استخراجها وفق هذا النموذج (الصعوبة، التمييز)، تعد أكثر اتفاقا مع المؤشرات التقليدية مقارنة بالنماذج الأخرى (Hernandez, 2009)، بالإضافة إلى أفضلية هذا النموذج في انتقاء الفقرات (Pelton,2002).

ويتم تقييم حسن مطابقة النموذج المستخدم للبيانات وذلك بعد التوصل إلى القيم التقديرية لبارامترات الفقرات والأفراد، ويتم التحقق من مطابقة كل فقرة من فقرات الاختبار للنموذج المستخدم من خلال قدرة النموذج على تفسير كيفية استجابات الأفراد لإحدى الفقرات أو التنبؤ بهذه الاستجابات، ولإجراء ذلك يمكن تجميع الأفراد أثناء عملية تقدير بارامترات النموذج المستخدم في مجموعات وفقا لدرجاتهم الكلية، ومن ثم يمكن الحصول على قيمة المقياس الإحصائي (مربع كاي) (χ^2) كمؤشر لحسن مطابقة الاختبار ككل باستخدام الصيغة التالية (علام، 2005):

$$\chi_i^2 = \sum_{j=1}^{k} f_j \frac{(p_{ij}-P_{ij})^2}{p_{ij}Q_{ij}} \quad \text{..... (2)} \qquad \text{حيث}$$

p_{ij} تشير إلى نسبة عدد الإجابات الصحيحة الفعلية ويعبر عنها بإحتمالية الإجابة الصحيحة للفقرة (i) بواسطة الفرد (j). P_{ij} : النسبة المتوقعة لعدد الإجابات الصحيحة استناداً إلى تقييم القيم التقديرية للبارامترات.

f_i :عدد الأفراد عند مستوى قدرة معين. Q_{ij} : إحتمال الإجابة الخطأ للفقرة (i) بواسطة الفرد (j).

ويتم التحقق من مدى مطابقة كل فقرة من الفقرات الاختبارية للنموذج المستخدم من خلال إيجاد نسبة عدد الإجابات الصحيحة الفعلية للفقرة، وكذلك نسبة عدد الإجابات المتوقعة للفقرة نفسها في داخل كل مجموعة من مجموعات مستويات القدرة، ويتم تطبيق المعادلة (2) السابقة لاختبار الدلالة الإحصائية للبواقي.

2.2- الدراسات السابقة:

من الدراسات التي اهتمت بمدى مطابقة فقرات اختبار من نوع اختيار من متعدد وفق النظريتين الكلاسيكية والحديثة في القياس، دراسة أبو فودة (2016) التي هدفت إلى الكشف عن التوافق بين الخصائص السيكومترية للاختبار وفقراته وفق النظرية الكلاسيكية في القياس والنموذج اللوجستي ثنائي المعلم في مطابقة فقرات اختبار اختيار من متعدد÷ ولتحقيق أهداف الدراسة تم بناء اختبار محكي المرجع في مادة الرياضيات في وحدة الهندسة التحليلية لطلبة الصف العاشر الأساسي، وقد تكوّن الاختبار في صورته النهائية من (30) فقرة من نوع الاختيار من متعدد من أربعة بدائل، جرى التحقق من صدقها وثباتها.، وتم تطبيق الاختبار على عينة مكوّنة من (140) طالباً تم اختيارهم بالطريقة العشوائية العنقودية من مدارس تربية محافظة جرش خلال العام الدراسي 2012-2013م، وقد أظهرت نتائج الدراسة مطابقة (29) فقرة لكل من النظرية الكلاسيكية في القياس، والنموذج اللوجستي ثنائي المعلم، وقد بلغت قيمة معامل الصدق وفقا للنموذج اللوجستي ثنائي المعلم (0.925)، بينما بلغت قيمة معامل الصدق وفقا للنظرية الكلاسيكية في القياس (0.936)، وأشارت النتائج إلى وجود فرق ذي دلالة إحصائية بين معاملي الصدق لصالح النظرية الكلاسيكية في القياس، وبلغت قيمة معامل الثبات التجريبي وفقا للنموذج اللوجستي ثنائي المعلم (0.966)، بينما بلغت قيمة معامل الثبات وفقا للنظرية الكلاسيكية (0.955)، وأشارت النتائج إلى وجود فروق ذات دلالة إحصائية في تقدير معاملي الثبات لصالح النموذج اللوجستي ثنائي المعلم.

وأجرى حجازي والخطيب (2013) دراسة هدفت إلى الكشف عن التوافق بين النظرية الكلاسيكية في القياس والنموذج اللوجستي ثنائي المعلم من أجل التعرف على مدى مطابقة فقرات اختبار محكي المرجع في المادة النظرية لأحكام التلاوة والتجويد لافتراضات كلا النظريتين، وذلك من خلال بناء اختبار مكوّن من (41) فقرة من نوع اختيار من متعدد من أربعة بدائل، وتم تطبيقه على عينة مكونة من (404) طالباً وطالبة تم اختيارهم بالطريقة العشوائية الطبقية من (16) مركزاً قرآنياً، بواقع (9) مراكز للذكور، و(7) مراكز للإناث في العاصمة عمان، وأشارت نتائج تلك الدراسة إلى مطابقة (40) فقرة للنظرية الكلاسيكية في القياس، ومطابقة (39) فقرة للنموذج اللوجستي ثنائي المعلم، مما يشير إلى وجود (39) فقرة من فقرات الاختبار مطابقة مع افتراضات النظرية الكلاسيكية والنموذج اللوجستي ثنائي المعلم بنسبة توافق بلغت (95.12%) من العدد الكلي لفقرات الاختبار، بينما أشارت النتائج إلى وجود فقرة واحدة فقط مطابقة للنظرية الكلاسيكية في القياس وغير مطابقة للنموذج اللوجستي ثنائي المعلم وشكلت ما نسبته (2.44%) من عدد الفقرات الكلي، كذلك أشارت إلى وجود فقرة واحدة غير مطابقة لكل من النظرية الكلاسيكية في القياس، وللنموذج اللوجستي ثنائي المعلم وشكلت ما نسبته (2.44%) من عدد الفقرات الكلي، وفيما يتعلق بالخصائص السيكومترية لفقرات الاختبار

أشارت النتائج إلى أن قيمة معامل ثبات الاختبار وفقا للنظرية الكلاسيكية في القياس بلغت (0.927)، بينما بلغت وفقا للنموذج اللوجستي ثنائي المعلم (0.943).

ودراسة (أون) (Onn, 2013) التي هدفت إلى مقارنة بين النظرية الكلاسيكية في القياس ونظرية الاستجابة للفقرة وفقا للنموذج اللوجستي ثنائي المعلم، وذلك من حيث عدد الفقرات المنتقاة ومعلمة الثبات، ولتحقيق هدف الدراسة قام الباحث بإعداد اختبار في مادة الفيزياء مكوّن من (50) فقرة من نوع اختيار من متعدد، طبق على عينة مكونة من (69) طالباً وطالبة من طلبة المدارس في نيجيريا، وأسفرت نتائج الدراسة عن مطابقة (29) فقرة للنظرية الكلاسيكية، و(39) فقرة للنموذج اللوجستي ثنائي المعلم، كما أشارت النتائج إلى تدني معامل الثبات في كلا النظريتين، إلا أن قيمة معامل الثبات للنموذج اللوجستي ثنائي المعلم أعلى من معامل الثبات في النظرية الكلاسيكية؛ حيث بلغت قيمتي معامل الثبات على التوالي (0.674، 0.490).

ودراسة (Jimelo and Silvestre, 2009) التي هدفت إلى المقارنة بين النظريتين الكلاسيكية والحديثة في القياس من خلال بناء اختبار محكي المرجع في مادة الأحياء يتكون من (60) فقرة، وتم تحديد درجة قطع (26) باستخدام طريقة (أنجوف). وقد جرى تطبيق الاختبار على عينة مكونة من (326) طالباً وطالبة، وأشارت نتائج الدراسة إلى أن (33) فقرة من فقرات الاختبار طابقت النظريتين الكلاسيكية والحديثة في القياس، وقد تقاربت قيمة معامل الثبات وفقا للنظريتين، حيث بلغت قيمة معامل الثبات وفقا للنظرية الكلاسيكية في القياس (0.72) وبلغت قيمة معامل الثبات وفقاً للنظرية الحديثة في القياس (0.70)، وأظهرت نتائج التحليل عدم وجود فروق ذات دلالة إحصائية في متوسط صعوبة الفقرات بين النظريتين.

ودراسة (Kan, 2006) التي هدفت إلى المقارنة بين الخصائص السيكومترية (الصعوبة والتمييز) من خلال النظرية الكلاسيكية في القياس والنموذج ثنائي المعلم لنظرية الاستجابة للفقرة، ومن أجل ذلك تم تحليل بيانات (25) فقرة من فقرات اختبار تحديد المستوى الوطني والذي تقدم له (553105) طالب وطالبة في المدارس الثانوية، وأسفرت نتائج الدراسة عن تحقق افتراض أحادية البعد للفقرات من خلال استخدام التحليل العاملي، كما أشارت النتائج عن وجود فروق ذات دلالة إحصائية في متوسط معلمة صعوبة الفقرة بين النظريتين ولصالح النموذج ثنائي المعلم، وأشارت النتائج كذلك إلى وجود فروق ذات دلالة إحصائية في متوسط معلمة تمييز الفقرة بين النظريتين ولصالح النموذج ثنائي المعلم، بالإضافة إلى أن معاملات الارتباط بين النظرية الكلاسيكية والنموذج ثنائي المعلم في تقدير معلمة الصعوبة قوية جدا، وكذلك الحال لدى معاملات الارتباط في معلمة تمييز الفقرة.

ودراسة (Hwang, 2002) التي هدفت إلى المقارنة بين النظرية الكلاسيكية ونظرية الاستجابة للفقرة وفقاً للنماذج اللوجستية الثلاث من حيث تقديرات مؤشرات صعوبة الفقرات، وتمييزها ومستويات القدرة للطلبة، ومن حيث استقرار تقديرات الفقرات الاختبارية باختلاف مستويات قدرة الطلبة، وقد بلغ حجم عينة الدراسة (600) طالباً من طلبة الصف الثامن الأساسي، وقد تم استخدام اختبار تحصيلي في الرياضيات والذي يشتمل على (15) فقرة، وأظهرت نتائج الدراسة إلى أن معاملات الارتباط كانت مرتفعة؛ حيث كان معظمها أكبر من 0.90 بين المؤشرات الإحصائية المستمدة من النظريتين وهذا يشير إلى وجود تشابه كبير بينهما، وأظهرت النتائج وجود تشابه واضح بين النظريتين في تقدير مؤشرات صعوبة الفقرات، إلا أن هذا التشابه يقل فيما يتعلق بتقدير مؤشرات تمييز الفقرات.

ومن الدراسات التي اهتمت بتصميم الاختبارات وانتقاء فقراتها وفقاً للنظرية الكلاسيكية في القياس ونظرية الاستجابة للفقرة، أجرى (Owens, 2000) دراسة هدفت إلى المقارنة بين معايير اختيار الفقرات وفق النظرية الكلاسيكية ونظرية الاستجابة للفقرة؛ وذلك من خلال بناء خمسة اختبارات مختلفة في عدد فقراتها وفي طرق انتقاء فقراتها، وتم تقدير معاملات الصعوبة والتمييز، ثم تمت المقارنة من حيث معاملات الثبات والتباين، وأشارت نتائج الدراسة إلى وجود فروق ذات دلالة إحصائية في الثبات لصالح نظرية الاستجابة للفقرة كما أشارت النتائج إلى أن الخطأ المعياري في التقدير لفقرات تلك الاختبارات التي تم انتقائها وفق نظرية الاستجابة للفقرة كان أقل مقارنة بالفقرات التي تم انتقاؤها وفقاً للنظرية الكلاسيكية في القياس.

في ضوء ما تم عرضه من دراسات سابقة، يتضح أنه يمكن استخلاص بعض المعطيات من نتائج تلك الدراسات فيما يتعلق بمطابقة فقرات اختبار اختيار من متعدد للنظرية الكلاسيكية في القياس، ونظرية الاستجابة للفقرة باستخدام النموذج اللوجستي ثنائي المعلم، عند تتبع نتائج تلك الدراسات يتضح بأن هناك تباين في نتائج الدراسات المختلفة المتعلق بنسبة عدد الفقرات المطابقة لافتراضات كل من النظريتين، كدراسة أبو فودة (2016)، دراسة حجازي والخطيب (2013)، ودراسة (Onn, 2013)؛ إذ أشارت نتائج تلك الدراسات إلى أن نسبة عدد الفقرات المطابقة لافتراضات النظرية الكلاسيكية في القياس بلغت (96.6%، 97.3% و58%) على الترتيب، ونسبة عدد الفقرات المطابقة للنموذج اللوجستي ثنائي المعلم (95.3%، 95.1%، 78%) على الترتيب، كذلك يتضح من الدراسات السابقة وجود تباين في نسبة التوافق في عدد الفقرات التي تطابق افتراضات كلا النظريتين: كدراسة (Jimelo and Silvestre, 2009)، ودراسة فودة (2016)؛ إذ أشارت نتائج تلك الدراسات إلى أن نسبة عدد الفقرات المطابقة لإفتراضات النظرية الكلاسيكية في القياس، والنموذج اللوجستي ثنائي المعلم بلغت (55%، 93.3%) على الترتيب، ووفقا لذلك جاءت الدراسة الحالية لتكون مساندة لنتائج الدراسات السابقة والتي تهدف جميعها إلى تقييم مدى مطابقة فقرات اختبار اختيار من متعدد لافتراضات النظرية الكلاسيكية في القياس، وافتراضات النموذج اللوجستي ثنائي المعلم، وتمتاز الدراسة الحالية لكونها تطبق على اختبار عدد فقراته (100) فقرة بخمسة بدائل وهي تختلف عن معظم الدراسات السابقة التي تتضمن أربعة بدائل.

3 – الطريقة والأدوات:

1.3– منهج الدراسة:

تم استخدام المنهج الوصفي التحليلي Descriptive Analytical وذلك لمناسبته لإجراءات الدراسة الحالية.

2.3– مجتمع وعينة الدراسة:

تكوّن مجتمع الدراسة من جميع طلاب الطب البشري في جامعة الخليج العربي للفصل الدراسي الأول من العام الدراسي 2018–2019، والبالغ عددهم (1115) طالباً وطالبة وفقاً لإحصائيات وحدة القبول والتسجيل أما عينة الدراسة فقد تكوّنت من جميع طلبة السنة الدراسية الأولى في كلية الطب والعلوم الطبية في جامعة الخليج العربي الذين درسوا مساق علم النفس، وتقدموا للاختبار النهائي، والبالغ عددهم (203) طالباً وطالبة وشكلوا ما نسبته 18.21% من مجتمع الدراسة، وبلغ عدد الطلبة الذكور (57) طالباً، شكلوا ما نسبته 28.1% من أفراد العينة، و(146) طالبة شكلن ما نسبته 71.9%. والجدول (1) يبين توزيع الطلبة وفقاً لمتغير الجنس والجنسية.

جدول (1) توزيع أفراد عينة الدراسة وفقاً لمتغير الجنس والجنسية

الجنس	الجنسية					المجموع (%)
	بحريني	سعودي	كويتي	عُماني	أخرى	
الذكور	14	28	14	1	0	57 (28.1)
الإناث	29	26	66	23	2	146 (71.9)
المجموع	43	54	80	24	2	203
النسبة %	21.2	26.6	39.4	11.8	1.0	100

3.3- أداة الدراسة:

تمثل أداة الدراسة الاختبار النهائي لمساق علم النفس الذي يدرّس في كلية الطب والعلوم الطبية في جامعة الخليج العربي، ويعد هذا المساق من ضمن المساقات الإجبارية التي يدرسها طلبة الطب البشري في الفصل الدراسي الأول، ويكون عادة الاختبار من نوع الاختيار من متعدد، تم إعداد الاختبار من قبل مدرس المادة، وتم تحكيمه ومراجعته وفقا للإجراءات المتبعة في وحدة تقييم الاختبارات في الجامعة (Assessment Office)، بحيث يتم تعديل تلك الملاحظات من قبل منسق المادة، وبعد كتابة الاختبار في صورته النهائية والذي يتكون من (100) سؤال، تم تطبيقه على جميع طلبة السنة الدراسية الأولى للعام الأكاديمي 2018-2019م، وتتم الاستجابة على فقرات الاختبار من خلال اختيار بديل واحد من بين خمسة بدائل، وتعطى القيمة (1) في حالة الإجابة الصحيحة، و(0) في حالة الإجابة الخاطئة.

4.3- الخصائص السيكومترية لأداة الدراسة:

أولاً: الصدق: تم التحقق من صدق اختبار علم النفس باستخدام الصدق المحكي Criterion Validity، وذلك من خلال حساب معامل ارتباط (بيرسون) Pearson Correlation Coefficient للكشف عن العلاقة بين أداء لطلبة على اختبار علم النفس، وبين أداؤهم على اختبار علم الاجتماع كمحك خارجي، وتم اختيار مساق علم الاجتماع لكونه الأنسب والأقرب للأداء مقارنة بالبيانات المتوفرة عن أداء طلبة السنة الدراسية الأولى في كلية الطب والعلوم الطبية في جامعة الخليج العربي، وأشارت نتائج التحليل بأن قيمة معامل الارتباط بلغت (0.785) وهي قيمة دالة إحصائياً عند مستوى ($\alpha < 0.01$)، وتعد هذه القيمة مؤشراً على أن الاختبار يمتاز بدرجة صدق محكي مناسبة لأغراض الدراسة الحالية.

ثانياً: الثبات: تم التحقق من المؤشرات الأولية لثبات اختبار مساق علم النفس من خلال حساب ثبات الاتساق الداخلي باستخدام معادلة (كودر ريتشاردسون) (KR-20) بدلالة إحصائيات فقرات الاختبار، وأشارت النتائج بأن قيمة معامل الثبات بلغت (0.910)، وهي قيمة مرتفعة وتشير إلى أن فقرات الاختبار تمتاز بدرجة ثبات مناسبة، كما تم التحقق من ثبات الاختبار وفقا لنظرية الاستجابة للفقرة باستخدام النموذج اللوجستي ثنائي المعلم من خلال حساب معامل الثبات الإمبريقي Empirical Reliability أو ما يسمى بالثبات التجريبي، وذلك باستخدام برنامج (Bilog-MG3)، وأشارت النتائج بأن قيمة معامل الثبات الإمبريقي بلغت (0.908) وهي قيمة مرتفعة وتناسب أغراض الدراسة الحالية.

5.3- إجراءات الدراسة:

لغايات تحقيق أهداف الدراسة نُفذّت الإجراءات وفق الخطوات الآتية:

- أخذ الموافقة الرسمية من قبل لجنة الأخلاق الطبية في جامعة الخليج العربي من أجل البدء بإجراءات الدراسة وتزويد الباحثان بنتائج الطلبة في مساق علم النفس (PSY.121)، رقم الموافقة (E028-P1-4/19).
- التأكد من صدق وثبات بيانات الدراسة (نتائج الطلبة) لمساق علم النفس للعام الدراسي 2018-2019 من خلال استخدام صدق المرتبط بمحك، وحساب معامل ثبات الاتساق الداخلي باستخدام معادلة (كودر ريتشاردسون)، والثبات الإمبريقي وفقا لنموذج الاستجابة للفقرة باستخدام النموذج اللوجستي ثنائي المعلم.
- حساب معامل الصعوبة ومعامل التمييز لجميع فقرات الاختبار والبالغ عددها (100) فقرة والتحقق من مدى مطابقتها لافتراضات النظرية الكلاسيكية في القياس.
- التحقق من افتراض أحادية البعد لفقرات الاختبار تم إجراء التحليل العاملي وفقا لطريقة المكونات الرئيسية.
- استخدام برنامج (بايلوق) (Bilog-MG3) للتعرف على قدرات الأفراد (θ) غير المطابقة لافتراضات النموذج اللوجستي ثنائي المعلم؛ بحيث يتم حذف الأفراد الغير مطابقين لافتراضات النموذج بالإضافة إلى التعرف على مدى مطابقة تقديرات معالم فقرات الاختبار للنموذج اللوجستي ثنائي المعلم من خلال حساب الإحصائي (χ^2) لحسن مطابقة معلم الفقرة؛ بحيث يتم حذف الفقرات الغير مطابقة لافتراضات النموذج اللوجستي ثنائي المعلم.
- تصنيف فقرات الاختبار وفقا للتطابق بين النظرية الكلاسيكية في القياس والنموذج اللوجستي ثنائي المعلم وذلك من خلال التعرف على عدد ونسبة الفقرات المطابقة لافتراضات كلا النظريتين.

6.3- المعالجة الإحصائية:

للإجابة عن أسئلة الدراسة، تم استخدام برنامج (SPSS.v26) لإجراء التحليلات اللازمة وفقاً للنظرية الكلاسيكية في القياس (صعوبة الفقرة، التمييز، معامل الثبات باستخدام طريقة (كيودر ريتشاردستون) (KR20) والتحليل العاملي الاستكشافي (Exploratory Factor Analysis) باستخدام تحليل المكونات الأساسية (Principle Component Analysis). واستخدم برنامج Bilog-MG3 لإجراء التحليل اللازم وفقاً للنموذج اللوجستي ثنائي المعلم (معلمة الصعوبة ومعلمة التمييز، إحصاء مربع كاي (χ^2) لحسن المطابقة).

4- النتائج ومناقشتها:

1.4- النتائج المتعلقة بالسؤال الأول:

للإجابة عن سؤال الدراسة استخدام البرنامج الإحصائي (SPSS.v26) لحساب معالم الفقرة والمتمثلة بمعلم الصعوبة، ومعلم التمييز لفقرات اختبار علم النفس والبالغ عددها (100) فقرة، وذلك للتحقق من مدى مطابقة فقرات الاختبار لافتراضات النظرية الكلاسيكية في القياس، والجدول (2) يبين نتائج التحليل.

جدول (2) قيم معاملات الصعوبة والتمييز لفقرات الاختبار قبل حذف الفقرات غير المطابقة للنموذج

رقم الفقرة	معامل الصعوبة	معامل التمييز	رقم الفقرة	معامل الصعوبة	معامل التمييز	رقم الفقرة	معامل الصعوبة	معامل التمييز
1	0.73	0.33	35	0.94	0.22	68	0.78	0.16
2	0.70	0.50	36	0.99	0.20	69	0.81	0.43
3	0.74	0.42	37	0.99	0.20	70	0.99	0.28
4	0.71	0.51	38	0.92	0.29	71	0.89	0.22
5	1.00	0.00	39	0.98	0.06	72	0.57	0.46
6	0.71	0.26	40	0.89	0.19	73	0.93	0.20
7	0.70	0.29	41	0.86	0.40	74	0.84	0.14
8	0.71	0.36	42	0.88	0.20	75	0.95	0.11
9	0.97	0.22	43	0.88	0.31	76	0.68	0.21
10	0.99	0.04	44	0.82	0.29	77	0.50	0.20
11	0.85	0.54	45	0.83	0.47	78	0.86	0.44
12	0.86	0.18	46	0.82	0.29	79	0.79	0.20
13	0.84	0.43	47	0.96	0.36	80	0.85	0.14
14	0.89	0.20	48	0.43	0.22	81	0.81	0.25
15	0.83	0.43	49	0.55	0.21	82	0.74	0.53
16	0.93	0.12	50	0.95	0.22	83	0.70	0.37
17	0.90	0.30	51	0.77	0.36	84	0.93	0.42
18	0.78	0.40	52	0.78	0.43	85	0.97	0.02
19	0.98	0.17	53	0.78	0.40	86	0.84	0.27
20	0.54	0.15	54	0.83	0.34	87	0.91	0.20
21	0.93	0.28	55	0.62	0.36	88	0.66	0.36
22	0.65	0.57	56	0.90	0.35	89	0.35	0.20
23	0.43	0.40	57	0.86	0.31	90	0.96	0.17
24	0.81	0.43	58	0.89	0.43	91	0.83	0.46
25	0.39	0.06	59	0.82	0.22	92	0.65	0.26
26	0.78	0.42	60	0.63	0.25	93	0.32	0.38
27	0.86	0.18	61	0.81	0.38	94	0.71	0.38
28	0.92	0.26	62	0.94	0.41	95	0.84	0.24
29	0.95	0.27	63	0.47	0.42	96	0.99	0.12
30	0.62	0.42	64	0.86	0.34	97	0.60	0.35
31	0.86	0.20	65	0.90	0.20	98	0.84	0.32
32	0.65	0.31	66	0.84	0.01	99	0.96	0.21
33	0.79	0.53	67	0.52	-0.08	100	0.71	0.39
34	0.71	0.10						

يتضح من البيانات الواردة في الجدول (2) بأن قيم معامل الصعوبة تتراوح ما بين (0.32- 1.00) بمتوسط حسابي مقداره (0.79)، وانحراف معياري (1.56)، وقيم معامل التمييز تتراوح ما بين (-0.08 -0.57) بمتوسط حسابي مقداره (0.29)، وانحراف معياري (0.135). وتم اعتماد المدى المقبول لمعامل صعوبة الفقرة للدراسة الحالية من (0.20- 0.80) وفقا لافتراضات النظرية الكلاسيكية في القياس، وفيما يتعلق بمعامل التمييز تعد الفقرة جيدة إذا كان معامل تمييزها أعلى من (0.40)، ومقبولة وينصح بتحسينها إذا كانت ضمن المدى (0.20-0.40)، وضعيفة وينصح بحذفها إذا كان معامل تمييزها ضمن المدى (0.0- 0.19) بالإضافة إلى أن الفقرة سالبة التمييز يجب حذفها (Crocker & Algina, 1986)، ووفقاً لذلك فإن الفقرة التي يزيد معامل صعوبتها في هذا الاختبار عن (0.80) تعتبر فقرة سهلة يجب حذفها، بحيث تم الاحتفاظ بالفقرات السهلة التي جاء معامل تمييزها ضمن المدى المقبول، وحذف باقي الفقرات غير المطابقة والتي يزيد معامل صعوبتها عن 0.80. وأشارت النتائج وفقا للجدول (2) إلى أن قيمة معامل الصعوبة للفقرة الخامسة بلغ (1.0) مما يشير إلى أن جميع إجابات الطلبة على هذه الفقرة صحيحة، وبالتالي بلغت قيمة معامل الصعوبة (0) مما يشير إلى أنها فقرة غير قادرة على التمييز بين أداء الطلبة ذوي الأداء المرتفع، والطلبة ذوي الأداء المنخفض؛ لذا ووفقا لافتراضات النظرية الكلاسيكية في القياس فإنه تم حذف الفقرات التي أرقامها (5، 10، 12، 16، 19، 27، 29، 39، 40، 66، 74، 75، 80، 85،

90، 96)، وفيما يتعلق بالفقرات التي يقل معامل تمييزها عن 0.20، والتي تعد فقرات ذات قدرة تمييزية ضعيفة يجب حذفها بالإضافة إلى الفقرات ذات القدرة التمييزية السالبة، لذا ووفقاً لذلك فإنه تم حذف الفقرات غير المطابقة والتي أرقامها (20، 25، 34، 67، 68)، وبلغ عدد الفقرات المتبقية والمطابقة لافتراضات النظرية الكلاسيكية في القياس (79) فقرة وشكلت ما نسبته (79%) من عدد الفقرات الكلي للاختبار. وللإجابة عن الجزء الثاني من سؤال الدراسة المتعلق بمدى مطابقة فقرات الاختبار لافتراضات النموذج اللوجستي ثنائي المعلم (2PLM)، تم التحقق أولاً من افتراضات نظرية الاستجابة للفقرة (النظرية الحديثة في القياس) وفقاً لما يلي:

أولاً: افتراض أحادية البعد

تم التحقق من افتراض أحادية البعد وذلك من خلال استخدام التحليل العاملي الاستكشافي (Exploratory Factor Analysis) من الدرجة الأولى على العينة ككل، باستخدام تحليل المكونات الأساسية (Principle Component Analysis) لاستجابات الطلة على فقرات الاختبار، وتم إجراء عملية التدوير باستخدام طريقة التدوير المتعامد (Varimax Rotation) للعوامل التي كانت قيم الجذر الكامن لها أكبر من واحد صحيح والتي بلغ عددها (35) عاملاً، وتم حساب قيم الجذور الكامنة (Eigen Values)، ونسبة التباين المفسر (Explained Variance) لكل عامل من العوامل، والجدول (3) يبين نتائج التحليل.

جدول (3) قيم الجذور الكامنة ونسبة التباين المفسرة ونسبة التباين التراكمية للعوامل المستخلصة

مجموع مربعات التشبعات المستخلصة Extraction sums of squared Loadings							
العامل	الجذر الكامن	نسبة التباين المفسر	نسبة التباين المفسر التراكمي	العامل	الجذر الكامن	نسبة التباين المفسر	نسبة التباين المفسر التراكمي
1	11.84	12.20	12.08	19	1.60	1.63	52.14
2	3.24	3.33	15.38	20	1.54	1.57	53.71
3	2.85	2.91	18.30	21	1.49	1.52	55.23
4	2.77	2.83	21.12	22	1.45	1.48	56.71
5	2.59	2.65	23.77	23	1.40	1.43	58.14
6	2.46	2.52	26.29	24	1.36	1.38	59.52
7	2.32	2.38	28.67	25	1.34	1.37	60.89
8	2.23	2.28	30.95	26	1.27	1.30	62.19
9	2.22	2.26	33.22	27	1.25	1.28	63.46
10	2.10	2.20	35.42	28	1.23	1.25	64.72
11	2.00	2.08	37.49	29	1.20	1.23	65.94
12	1.94	2.00	39.50	30	1.18	1.21	67.15
13	1.92	1.96	41.46	31	1.13	1.15	68.30
14	1.88	1.92	43.38	32	1.11	1.14	69.43
15	1.84	1.88	45.26	33	1.09	1.11	70.54
16	1.82	1.86	47.11	34	1.05	1.07	71.61
17	1.68	1.72	48.83	35	1.02	1.04	72.65
18	1.65	1.68	50.51				

يتضح من البيانات الواردة في الجدول أعلاه بأن قيم الجذر الكامن لجميع العوامل الناتجة من التحليل العاملي الاستكشافي والبالغ عددها (35) عاملاً تفوق الواحد صحيح، وفسرت ما نسبته (72.65%) من التباين الكلي لفقرات الاختبار، ويتضح كذلك بأن قيمة الجذر الكامن للعامل الأول بلغت (11.84) وفسرت ما نسبته (12.20%) من التباين الكلي لفقرات الاختبار، وبلغت قيمة الجذر الكامن للعامل الثاني (3.24)، وفسرت ما نسبته (3.33%) من التباين الكلي، وعند قسمة قيمة الجذر الكامن للعامل الأول على قيمة الجذر الكامن للعامل الثاني فإن الناتج يساوي (3.7) وهي قيمة تفوق العدد (2)، ويعد ذلك مؤشراً على أحادية البعد (Hattie, 1985; Reckase, 1997)، ويلاحظ من خلال استخدام مخطط سكري (Scree Plot) للعوامل المكونة لفقرات

الاختبار مع الجذور الكامنة كما يظهر في الشكل رقم (1) أن التمثيل البياني يظهر فيه العامل الأول ينفرد بقيمة جذر كامن مرتفعة نسبياً مقارنة مع العوامل الأخرى التي تظهر بجذور كامنة صغيرة نسبياً ومتقاربة أيضاً، ويعد ذلك مؤشراً على أحادية البعد.

الشكل(1) قيم الجذور الكامنة للعوامل المكونة لفقرات الاختبار

ثانياً: افتراض الاستقلال الموضعي

وفقاً لنتائج التحليل العاملي والتي أشارت إلى تحقق افتراض أحادية البعد، فإن ذلك يعد مؤشراً على تحقق افتراض الاستقلال الموضعي، فقد أورد (Hambleton & Swaminathan, 1989) بأن التحقق من افتراض أحادي البعد يكافئ افتراض الاستقلال الموضعي لفقرات الاختبار.

ثالثاً: افتراض التحرر من السرعة في الأداء

نظرا لإجابة جميع الطلبة عن فقرات الاختبار ضمن الوقت المحدد للإجابة، فإن ذلك يعد مؤشراً على تحقق افتراض التحرر من السرعة في الأداء، وإذا وجد إخفاق لدى بعض الطلبة في الإجابة عن بعض فقرات الاختبار فإن ذلك مرده إلى عامل القدرة لدى هؤلاء الطلبة وليس لعامل الزمن.

بعد التحقق من افتراضات نظرية الاستجابة للفقرة، تم استخدام برنامج (Bilog-MG3) لفحص حسن مطابقة فقرات الاختبار للنموذج اللوجستي ثنائي المعلم، وذلك للكشف عن تقديرات معالم الفقرة، وتقدير قدرات الأفراد الغير مطابقين للنموذج اللوجستي ثنائي المعلم، أظهرت نتائج التحليل وفقا للمرحلة الثالثة من مراحل التحليل (Phase3)، بأن تقديرات قدرة الأفراد تراوحت من (-2.223) لوجيت إلى (2.433) لوجيت، بمتوسط حسابي بلغ (-0.083) لوجيتK مما يشير إلى أن تقديرات قدرة الأفراد جاءت ضمن المدى المقبول ولم يحذف أي فرد من أفراد عينة الدراسة وذلك لكون قدراتهم جاءت مطابقة لافتراضات النموذج المستخدم، وللتحقق من مدى مطابقة استجابات الطلبة على فقرات الاختبار وفقا للنموذج اللوجستي ثنائي المعلم، والذي يقوم باستبعاد الفقرات التي يجيب عنها جميع الطلبة بشكل صحيح، بالإضافة إلى استبعاد الفقرات التي يخفق جميع الطلبة بالإجابة عنها، تم تقدير معالم فقرات اختبار علم النفس وفقاً للنموذج اللوجستي ثنائي المعلم (2PLM)، والمتمثلة بمعلمة الصعوبة والتمييز، بالإضافة للخطأ المعياري لتقدير هذه المعالم، وقيمة معامل التشبع، واختبار (مربع كاي) لحسن المطابقة عند مستوى الدلالة المحدد ($\alpha < 0.05$) وفقا للمرحلة الثانية من مراحل التحليل باستخدام برنامج BILOG-MG3 (Phase2) والجدول (4) يبين نتائج التحليل.

جدول (4) معاملات الصعوبة والتمييز ومستوى دلالة جودة المطابقة وفقا للنموذج اللوجستي ثنائي المعلم

رقم الفقرة	معامل الصعوبة (b_i)	معامل التمييز (a_i)	مستوى الدلالة (χ^2)	رقم الفقرة	معامل الصعوبة (b_i)	معامل التمييز (a_i)	مستوى الدلالة (χ^2)	رقم الفقرة	معامل الصعوبة (b_i)	معامل التمييز (a_i)	مستوى الدلالة (χ^2)
1	1.29-	0.534	0.934	35	3.04-	0.596	0.779	68	2.27-	0.350	0.506
2	0.83-	0.865	0.426	36	2.93-	1.106	0.000	69	1.33-	0.923	0.964
3	1.12-	0.714	0.975	37	2.95-	1.096	0.000	70	2.90-	1.279	0.000
4	0.85-	0.972	0.856	38	2.40-	0.740	0.615	71	2.29-	0.628	0.011
6	1.39-	0.439	0.335	39	4.50-	0.561	0.000	72	0.31-	0.759	0.929
7	1.18-	0.490	0.007	40	2.98-	0.460	0.996	73	3.14-	0.532	0.850
8	0.99-	0.712	0.396	41	1.59-	0.915	0.999	74	2.59-	0.418	0.966
9	2.50-	1.053	0.000	42	2.31-	0.585	0.103	75	3.44-	0.549	0.463
10	4.66-	0.757	0.000	43	2.08-	0.699	0.463	76	1.18-	0.455	0.623
11	1.34-	1.309	0.793	44	1.83-	0.589	0.237	77	0.01	0.389	0.354
12	2.67-	0.435	0.705	45	1.31-	1.141	0.847	78	1.47-	1.125	0.996
13	1.44-	0.955	0.388	46	1.91-	0.554	0.998	79	2.26-	0.371	0.893
14	2.54-	0.547	0.739	47	2.05-	1.477	0.421	80	2.76-	0.408	0.714
15	1.32-	1.110	0.658	48	0.39	0.429	0.120	81	1.82-	0.543	0.398
16	4.09-	0.389	0.751	49	0.38-	0.370	0.615	82	0.91-	1.161	0.849
17	2.03-	0.829	0.887	50	2.76-	0.766	0.773	83	1.01-	0.619	0.922
18	1.26-	0.822	0.694	51	1.36-	0.672	0.157	84	1.85-	1.230	0.751
19	3.38-	0.813	0.000	52	1.26-	0.819	0.462	85	4.46-	0.476	0.937
20	0.31-	0.306	0.797	53	1.29-	0.783	0.994	86	2.01-	0.580	0.967
21	2.08-	1.021	0.572	54	1.67-	0.717	0.933	87	2.56-	0.609	0.721
22	0.57-	1.141	0.568	55	0.62-	0.559	0.962	88	0.82-	0.571	0.997
23	0.31	0.623	0.949	56	1.87-	0.961	0.968	89	1.11	0.339	0.988
24	1.32-	0.911	0.328	57	2.03-	0.621	0.544	90	3.30-	0.657	0.887
25	1.10	0.238	0.036	58	1.65-	1.131	0.650	91	1.39-	0.948	0.576
26	1.21-	0.880	0.690	59	2.11-	0.471	0.999	92	0.95-	0.439	0.285
27	2.77-	0.428	0.950	60	0.82-	0.442	0.684	93	0.80	0.713	0.707
28	2.31-	0.756	0.984	61	1.37-	0.847	0.448	94	1.01-	0.703	0.885
29	3.24-	0.593	0.652	62	1.76-	1.568	0.918	95	2.13-	0.522	0.950
30	0.56-	0.664	0.908	63	0.08	0.730	0.708	96	4.03-	0.915	0.000
31	2.54-	0.463	0.124	64	1.76-	0.796	0.961	97	0.49-	0.587	0.402
32	2.25-	1.072	0.241	65	2.82-	0.523	0.636	98	1.69-	0.727	0.872
33	1.06-	1.319	0.732	66	3.85-	0.266	0.074	99	2.78-	0.837	0.794
34	1.94-	0.295	0.461	67	0.30-	0.179	0.024	100	1.02-	0.661	0.877

يتضح من البيانات الواردة في الجدول (4) بأن تقديرات معلمة الصعوبة لفقرات الاختبار وفقاً للنموذج اللوجستي ثنائي المعلم تمتد من (-4.66 -1.11) لوجيت، بمتوسط حسابي مقداره (-1.767) لوجيت، وانحراف معياري بلغ (1.167) لوجيت، وأشارت نتائج التحليل بأن متوسط تقديرات القدرة للأفراد بلغت (-0.016) وهي تفوق متوسط تقديرات معلمة الصعوبة لفقرات الاختبار (-1.767)، ويعد ذلك مؤشراً على أن فقرات الاختبار كانت مناسبة لقدرات أفراد عينة الدراسة، ويتضح كذلك من البيانات بأن تقديرات قيم معلمة التمييز لفقرات الاختبار تراوحت من (0.238- 1.568)، بمتوسط حسابي مقداره (0.714)، وانحراف معياري بلغ (0.288)، وأشارت نتائج التحليل باستخدام برنامج Bilog-MG3 (Phase2) بأن قيم الخطأ المعياري لمعلم الصعوبة تمتد من (0.114– 1.562) بمتوسط مقداره (0.436)، وقيم الخطأ المعياري لمعلم التمييز تمتد من (0.047 – 0.725) بمتوسط حسابي مقداره (0.180)، وجاءت قيم متوسط الخطأ المعياري لتقديرات معلم الصعوبة والتمييز منخفضة ومناسبة، ويعد ذلك مؤشراً على دقة تقدير معالم فقرات الاختبار.

ويتضح من النتائج إلى وجود (12) فقرة من فقرات اختبار علم النفس غير مطابقة لافتراضات النموذج اللوجستي ثنائي المعلم وفقا لمؤشر المطابقة المعتمد على إحصائي (مربع كاي) لحسن المطابقة وهي الفقرات

التي أرقامها (7، 9، 10، 19، 25، 36، 37، 39، 67، 70, 71، 96)، إذ جاء مستوى الدلالة وفقا لإحصائي (χ^2) لتلك الفقرات أقل من مستوى الدلالة المحدد (0.05). لذا فإن نسبة الفقرات الغير مطابقة والتي تم استبعادها بلغت (12.1%) من عدد الفقرات الكلي، أما باقي الفقرات والبالغ عددها (87) فقرة، فقط تم الاحتفاظ بها وذلك لمطابقتها لافتراضات النموذج اللوجستي ثنائي المعلم وشكلت ما نسبته (87.9%) من عدد الفقرات، ووفقاً لنظرية الاستجابة للفقرة، فإن الفقرات التي تطابق النموذج اللوجستي ثنائي المعلم تعد فقرات مستقلة إحصائياً عن بعضها البعض ومستقلة كذلك عن عينة الطلبة الممتحنين.

أشارت نتائج التحليل المتعلقة بسؤال الدراسة الأول بأن نسبة عدد فقرات اختبار علم النفس المطابقة لافتراضات النظرية الكلاسيكية في القياس (79%) من عدد فقرات الاختبار الكلي؛ إذ بلغ عدد الفقرات الغير مطابقة (21) فقرة من فقرات الاختبار والبالغ عددها (100) فقرة، بواقع (16) فقرة جاء مؤشر معلم الصعوبة لها غير مطابق لافتراضات النظرية الكلاسيكية في القياس، وشكلت ما نسبته (16%) من عدد الفقرات و(5) فقرات جاء مؤشر معلم التمييز لها ضعيفاً، أو ذات قدرة تمييزية سالبة وشكلت هذه الفقرات ما نسبته (5%) من مجمل العدد الكلي لفقرات اختبار علم النفس، وبمقارنة هذه النتائج مع نتائج الدراسات السابقة فيما يتعلق بمدى مطابقة الفقرات لافتراضات النظرية الكلاسيكية في القياس نجد بأن هذه النسب التي أفرزتها نتائج الدراسة الحالية تختلف مع نتائج دراسة أبو فودة (2016)، ودراسة حجازي والخطيب (2013)، ودراسة (Jimelo and Silvestre, 2009)، ويمكن تبرير هذا الاختلاف إلى أنه عند استخدام قيم معاملات الصعوبة وفقا للطريقة الكلاسيكية في القياس فإن هذه القيم تتأثر بموقف الاختبار وخصائص العينة بالإضافة إلى اختلاف صيغة فقرات الاختبار ومحتوى مادة الاختبار، أما فيما يتعلق بنظرية الاستجابة للفقرة أسفرت نتائج الدراسة عن تحقق افتراض أحادية البعد لفقرات اختبار علم النفس من خلال استخدام التحليل العاملي الاستكشافي، وكذلك التحقق من افتراض الاستقلال الموضعي والتحرر من السرعة، وبالتالي التحقق من افتراضات النموذج اللوجستي ثنائي المعلم، وأظهرت النتائج بأن الفقرة الخامسة من فقرات الاختبار جاءت غير مطابقة لافتراضات النموذج اللوجستي ثنائي المعلم فتم حذفها ليصبح عدد فقرات الاختبار الكلي (99) فقرة أما فيما يتعلق بقدرات الأفراد، أشارت نتائج التحليل الإحصائي بأن جميع الأفراد جاءت قدراتهم مطابقة لافتراضات النموذج اللوجستي ثنائي المعلم ولم يتم حذف أي فرد منهم؛ إذ كشفت النتائج بأن متوسط تقديرات القدرة للأفراد جاءت أعلى من متوسط تقديرات معلمة الصعوبة لفقرات الاختبار؛ مما يشير إلى أن فقرات اختبار علم النفس جاءت مناسبة لقدرات أفراد عينة الدراسة، وكشفت النتائج عن مطابقة (87) فقرة من فقرات الاختبار لافتراضات النموذج اللوجستي ثنائي المعلم وشكلت ما نسبته (87.9%) من عدد الفقرات الكلي؛ مما يشير إلى أنها فقرات مستقلة إحصائياً عن بعضها البعض ومستقلة عن عدد الأفراد الممتحنين. ووفقا لمؤشر المطابقة المعتمد على إحصائي (مربع كاي) لحسن المطابقة أشارت النتائج إلى أن عدد الفقرات الغير مطابقة لافتراضات النموذج اللوجستي بلغ (12) فقرة من فقرات اختبار علم النفس، وشكلت هذه الفقرات ما نسبته (12.1%) من عدد الفقرات الكلي، لذا يلاحظ بوجود اختلاف بين نسبة الفقرات المطابقة لافتراضات النموذج اللوجستي ثنائي المعلم في الدراسة الحالية والبالغ (87.9%)، مع نسبة الفقرات التي تطابق نموذج الاستجابة للفقرة لكل من دراسة أبو فودة (2016)، ودراسة حجازي والخطيب (2013)، والتي بلغت نسبة الفقرات المطابقة فيها (93.3%، 95.1%) على الترتيب، وقد يعزو ذلك لكون جميع فقرات الاختبار كانت ضمن قدرات الطلبة أفراد عينة الدراسة، بالإضافة إلى أن متوسط تقدير قدرات الأفراد تفوق متوسط تقدير معلم الصعوبة لجميع الفقرات، كذلك جاءت تقديرات معلم التمييز وفقاً للنموذج اللوجستي ثنائي

المعلم موجبة ولم تكن متدنية وخالية من التقديرات السالبة، كما يمكن تفسير ذلك إلى جدية الطلبة أفراد عينة الدراسة في الإجابة عن جميع فقرات الاختبار، بالإضافة إلى طبيعة الاختبار وطريقة بناءه وعدد فقراته، بالإضافة إلى حجم العينة وعدد البدائل المستخدمة.

2.4-النتائج المتعلقة بسؤال الدراسة الثاني:

لإجابة عن سؤال الدراسة، تم التعرف على نسبة التوافق بين النظرية الكلاسيكية في القياس ونظرية الاستجابة للفقرة باستخدام النموذج اللوجستي ثنائي المعلم في انتقاء فقرات اختبار علم النفس، وذلك من خلال حصر نسبة عدد الفقرات(المطابقة/الغير مطابقة) لافتراضات كل من النظريتين، والجدول (5) يبين نتائج التحليل.

جدول (5) عدد الفقرات المطابقة/عدم المطابقة مع النظرية الكلاسيكية في القياس ونظرية الإستجابة للفقرة

نظرية الاستجابة للفقرة	النظرية الكلاسيكية في القياس						
		مطابقة		غير مطابقة			
	حالة الفقرة	عدد الفقرات	النسبة	عدد الفقرات	النسبة	عدد الفقرات	النسبة
	مطابقة	72	%72.7	15	%15.1	87	%87.8
	غير مطابقة	6	%6.1	6	%6.1	12	%12.2
	المجموع الكلي	78	%78.8	21	%21.2	99	%100

يتضح من البيانات الواردة في الجدول (5) بأن عدد فقرات الاختبار المطابقة لافتراضات النظرية الكلاسيكية في القياس، ومطابقة لافتراضات النموذج اللوجستي ثنائي المعلمة بلغ عددها (72) فقرة، وشكلت ما نسبته (72.7%) من مجمل فقرات الاختبار، كما يتضح بأن عدد الفقرات الغير مطابقة لافتراضات النظرية الكلاسيكية في القياس، وجاءت مطابقة لافتراضات النموذج اللوجستي ثنائي المعلم بلغ عددها (15) فقرة وهي الفقرات التي أرقامها (5، 12، 16، 20، 27، 29، 34، 40، 66، 68، 74، 75، 80، 85، 90) وشكلت ما نسبته (15.1%) من العدد الكلي لفقرات الاختبار، بينما بلغ عدد فقرات الاختبار الغير مطابقة لافتراضات النموذج اللوجستي ثنائي المعلم، ومطابقة لافتراضات النظرية الكلاسيكية في القياس (6) فقرات وهي الفقرات التي أرقامها (7، 9، 36، 37، 70، 71) وشكلت ما نسبته (6.1%) من عدد الفقرات الكلي، كذلك أشارت النتائج بأن عدد الفقرات الغير مطابقة لكلا النظريتين بلغ (6) فقرات وهي الفقرات التي أرقامها (10، 19، 25، 39، 67، 96) وشكلت ما نسبته (6.1%) من عدد الفقرات الكلي للاختبار.

أشارت النتائج المتعلقة بنسبة التوافق بين النظرية الكلاسيكية في القياس ونظرية الاستجابة للفقرة في انتقاء فقرات اختبار علم النفس، إلى أن نسبة اتفاق المؤشرات الإحصائية المستخدمة وفقا لافتراضات النظرية الكلاسيكية في القياس والنموذج اللوجستي ثنائي المعلم لعدد الفقرات المطابقة لافتراضات النظريتين بلغ (72.7%) من عدد الفقرات الكلي للاختبار، كما بلغت نسبة توافق المؤشرات الإحصائية المستخدمة وفقا لافتراضات النظرية الكلاسيكية في القياس والنموذج اللوجستي ثنائي المعلم في عدد الفقرات الغير مطابقة لافتراضات النظريتين (6.1%) من العدد الكلي لفقرات الاختبار، بينما أشارت النتائج بأن نسبة عدد فقرات مساق علم النفس المطابقة لافتراضات النظرية الكلاسيكية في القياس، وغير مطابقة لافتراضات النموذج اللوجستي ثنائي المعلم بلغ (6.1%) من العدد الكلي لفقرات الاختبار. ونسبة عدد فقرات الاختبار المطابقة لافتراضات النموذج اللوجستي ثنائي المعلم، وغير مطابقة لافتراضات النظرية الكلاسيكية في القياس بلغ (15.1%) من العدد الكلي لفقرات الاختبار، يتضح من ذلك بان نسبة عدد الفقرات المطابقة لافتراضات النموذج اللوجستي ثنائي المعلم يفوق نسبة عدد الفقرات المطابقة لافتراضات النظرية الكلاسيكية في القياس. وتتفق نتيجة هذه الدراسة مع دراسة (Onn, 2013) التي أسفرت نتائجها عن مطابقة (58%) من فقرات الاختبار لافتراضات للنظرية الكلاسيكية ومطابقة

(78%) من الفقرات لافتراضات النموذج اللوجستي ثنائي المعلم، لذا فإن مجمل النتائج التي تم التوصل إليها تشير إلى أن تحليل فقرات اختبار علم النفس وانتقاء فقراته باستخدام نظرية الاستجابة للفقرة يعد أفضل مقارنة بالنظرية الكلاسيكية في القياس، ويمكن تبرير ذلك إلى دقة تحليل فقرات الاختبار وفقا للنموذج اللوجستي ثنائي المعلم؛ لأن لكل فقرة من فقرات اختبار اختيار من متعدد بشكل عام خصائصها السيكومترية المتمثلة في منحنى القدرة الخاص بها، والذي بدوره يصف احتمالية الحصول على الإجابة الصحيحة عند معرفة قدرة الفرد، بالإضافة إلى أن انتقاء فقرات الاختبار من خلال تقدير الخصائص السيكومترية لتلك الفقرات باستخدام نماذج الاستجابة للفقرة يرتبط بافتراضات هذه النماذج، والتي تشير عادة إلى وجود دالة لوغاريتمية تراكمية مميزة لكل فقرة من فقرات الاختبار تمتاز بخصائص يمكن الإفادة منها في جعل ميزان درجات الاختبار خطياً، وبالتالي يكون تقدير تلك الخصائص مستقلاً عن خصائص عينة الأفراد الممتحنين. لذا في حالة نماذج الاستجابة للفقرة يكون الهدف من انتقاء فقرات الاختبار موجه نحو الحصول على تقدير صعوبة فقرات الاختبار، وتقييم حسن مطابقة درجات فقرات ذلك الاختبار للنموذج باستخدام قيم (مربع كاي) مثلاً، والإفادة من ذلك في تقدير سمات الأفراد، في حين أن الهدف من انتقاء فقرات الاختبار وفقا للنظرية الكلاسيكية في القياس هو الكشف عن مفردات الاختبار غير الصالحة أو غير المطابقة، بالإضافة إلى تقدير البارامترات (المعالم) المتعلقة بالفقرة، واستخدام القيم التقديرية لهذه البارامترات في تقدير الخصائص السيكومترية للاختبار ككل، كما أن معرفتنا للقيمة التقديرية لصعوبة فقرة اختبار بالنسبة لمجموعة معينة من الأفراد لا يفيدنا كثيراً في معرفة درجة صعوبة الفقرة بالنسبة لواحد من أفراد المجموعة؛ لأن قيم معاملات الصعوبة المشتقة بهذه الطريقة تختلف باختلاف متوسط ومدى قدرة أفراد المجموعة المستخدمة في اشتقاق تلك المعاملات.

5-الخلاصة:

هدفت هذه الدراسة إلى تحليل فقرات اختبار علم النفس، والتحقق من مدى مطابقة ومناسبة فقرات الاختبار لافتراضات النظرية الكلاسيكية في القياس، ونموذج الاستجابة للفقرة باستخدام النموذج اللوجستي ثنائي المعلم واستخدم الباحثان نظرية الاستجابة للفقرة كإطار لتحليل فقرات الاختبار، واستجابات الطلبة، وتحقيق افتراضات النموذج المستخدم كأحد نماذج الاستجابة للفقرة، بالإضافة إلى الوصول إلى اختبار موضوعي متحرر من قدرات الأفراد، ومتحرر من صعوبة الفقرات، وضمان الحصول على فقرات تمتاز بخصائص سيكومترية جيدة، وبالتالي محاولة الاقتراب من الوصول إلى العدالة والموضوعية والدقة، مما يمكننا من تشخيص جوانب الضعف أو التحقق من مدى إتقان الطلبة مما يساعد في بناء بنك أسئلة لمساق علم النفس نستطيع من خلاله أن نختار أي عدد من الفقرات بطريقة عشوائية من أجل تحقيق أهداف عملية القياس الموضوعي، كما هدفت الدراسة إلى توضيح المنهجية التي من خلالها يتم التحقق من مدى مطابقة فقرات اختبار لافتراضات النظريتين من خلال استخدام برمجيات إحصائية متنوعة، مما يساهم بفتح المجال لإمكانية تطبيق هذه المنهجية والإجراءات في بناء اختبارات تغطي كفايات تعليمية وتحصيلية أخرى، وفي ضوء ما توصلت إليه الدراسة الحالية من نتائج فإن الباحثان يوصون بـ:

- توعية معدّي الاختبارات بأهمية استخدام نظرية الاستجابة للفقرة في تحليل فقرات الاختبارات وتطويرها.
- إجراء دراسة مشابه للدراسة الحالية باستخدام النموذج اللوجستي ثلاثي المعلم.

- إجراء دراسة تهدف إلى التحقق من مدى مطابقة فقرات اختبارات ثنائية الإجابة لكل من النظرية الكلاسيكية في القياس ونظرية الاستجابة للفقرة وفقاً لاختلاف بعض المتغيرات: نوع النموذج اللوجستي المستخدم (أحادي المعلم، ثنائي المعلم، ثلاثي المعلم)، حجم العينة، وعدد فقرات الاختبار.
- الاستفادة من النتائج التي تم التوصيل إليها للدراسة الحالية من خلال تطبيق إجراءاتها عند بناء اختبارات لمساقات مختلفة.

شكر وتقدير: يتقدم الباحثان بجزيل الشكر والتقدير إلى أعضاء لجنة الأخلاق الطبية، وحدة التقويم والاختبارات في جامعة الخليج العربي، على موافقتهم بتزويدنا بالنتائج النهائية المتعلقة بمساق علم النفس للعام الدراسي 2018-2019م.

- الإحالات والمراجع:

ابو فودة، باسل خميس (2016). التوافق بين النظرية التقليدية في القياس ونظرية استجابة الفقرة في مطابقة فقرات اختبار محكي المرجع في وحدة الهندسة التحليلية. *مجلة دراسات عربية في التربية وعلم النفس*، 73 (73)، 187-215.

أبو عواد، فريال والقهوجي، أيمن (2016). تطوير اختبار في مهارات البحث العلمي لطلبة كلية التربية في الجامعات الأردنية والتحقق من خصائصه السيكومترية وفق نظريتي القياس الكلاسيكية والحديثة. *مجلة دراسات* (العلوم التربوية)، 43 (4)، 1645- 1668.

حجازي، تغريد والخطيب، عبد الله (2014). التوافق بين النظرية الكلاسيكية والنموذج ثنائي المعلمة في مطابقة فقرات اختبار محكي المرجع في أحكام التلاوة والتجويد. *مجلة جامعة النجاح للأبحاث* (العلوم الانسانية)، 28 (10)، 2239-2270.

الخياط، ماجد (2012). درجة مطابقة اختبار تحصيلي وفق نموذج راش أحادي المعلم في الكشف عن مستوى المعرفة العلمية في المهارات الرياضية لدى طلبة الصف الثامن الاساسي. *مجلة جامعة الأقصى*، 16 (1)، 87-111.

الشريفين، نضال (2006). الخصائص السيكومترية لاختبار محكي المرجع في القياس والتقويم التربوي وفق النظرية الحديثة في القياس التربوي والنفسي. *مجلة العلوم التربوية والنفسية*، كلية التربية، جامعة البحرين، **7** (4)، 80-107.

الشريفين، نضال والصبح، رانيا محمد (2011). أثر بنية فقرات الاختيار من متعدد ومستوى القدرة لدى الأفراد على دقة التقديرات لمعالم الفقرات والأفراد وفق نظرية الاستجابة للفقرة. *مجلة جامعة ام القرى للعلوم التربوية والنفسية*، 3(2)، 45-110.

علام، صلاح الدين (2005). *نماذج الاستجابة للمفردة الاختبارية أحادية البعد ومتعددة الأبعاد*. القاهرة: دار الفكر العربي.

علام، صلاح الدين (1987). دراسة موازنة ناقدة لنماذج السمات الكامنة والنماذج الكلاسيكية في القياس النفسي والتربوي. *المجلة العربية للعلوم النفسية*، 7(27)، 18-43.

عودة، أحمد سليمان (2010). *القياس والتقويم في العملية التدريسية*. (ط 4). اربد: دار الأمل للنشر والتوزيع.

النعيمي، عزالدين (2007). *نماذج نظرية استجابة الفقرة وتطبيقاتها في مجال النفس حركي*. ورقة بحثية مقدمة إلى المؤتمر العلمي الدولي الثاني "المستجدات العملية في التربية البدنية والرياضية" جامعة اليرموك، إربد، الأردن.

Aiken, L.R. (2003). Psychological Testing and Assessment (11thed.). Boston: Pearson Education Group.

Baghaei, P., and Amrahi, N. (2014). The effects of the number of options on the psychometric characteristics of multiple-choice items. Psychological Test and Assessment Modeling, 53(2), 192-211.

Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. New York: CBS College Publishing.

Jimelo, L., & Silvestre, T. (2009). Item Response Theory and Classical Test Theory: An Empirical Comparison of Item/Person Statistics in a Biological Science Test. *The International Journal of Educational and Psychological Assessment*, 1(1), 19-31.

Hambleton, R., & Swaminathan, H. (1989). *Item Response Theory "Principles & Applications*". Kluwer Nijhoff Publishing.

Hattie,J.(1985).Methodology review: Assessing Unidimensionality of tests and items. *Applied Psychological Measurement*. 9, 139-164.

Hernandez, R. (2009). Comparison of the Item Discrimination and Item Difficulty of the Quick-Mental Aptitude Test using CTT and IRT Methods. *The International Journal of Educational Employment and Psychological Assessment*, 1 (1), 12 – 18.

Hwang, D.Y. (2002). *Classical Test Theory and Item Response Theory: Analytical and Empirical Comparison*. Paper presented at the Annual Meeting of the Southwest Educational Research Association, Austin, TX. (ERIC Document Reproduction Service No. ED 466779).

Kan, A. (2006). Classical Test Theory and Latent Trait Theory Predicted by Empirical Working on Item Parameters. Mersin University *Journal of the Faculty of Education*, 2 (2), 227-235.

Onn, D. (2013). Classical test theory versus item response theory: An evaluation of the comparability of item analysis results. *Joint Admissions and Matriculation Board*, 1-23.

Oosterhoff, A. (1994). *Classroom Applications of Educational Measurement* (2nd Ed.). New York: Macmillan Publishing Company.

Owens, J. (2000). Using a Brand Loyalty Scale to Compare Item Selection Criteria and Scale Evaluation Techniques of Classical Test and Item Response Theory. Unpublished Doctoral Dissertation, Southern Illinois University.

Pelton, W. (2002). *The accuracy of unidimensional measurement models in the presence of deviations from the underlying assumptions*. Unpublished Doctoral Dissertation, Brigham Young University, U.S.A.

Reckase, M. (1997). The past and future of multidimensional item response theory. *Applied Psychological Measurement*, 21 (1), 25-36.

Toland، M. D. (2008). *Determining the Accuracy of Item Parameter Standard Error of Estimates in BILOG-MG3*. Unpublished Doctoral Dissertation, the University of Nebraska – Lincoln, AAT 3317288.

Wolfe EW, Smith Jr EV. (2007). Instrument development tools and activities for measure validation using rasch models: part II – validation activities. *Journal of Applied Measurement*, 8:204- 34.

كيفية الاستشهاد بهذا المقال حسب أسلوب APA :

المرابحة، عامر جبريل وبن صالح، عفيف (2021). تقييم مدى مطابقة فقرات اختبار علم النفس للنظرية الكلاسيكية في القياس ونظرية الاستجابة للفقرة وفقا للنموذج اللوجستي ثنائي المعلم (دراسة مقارنة). *مجلة العلوم النفسية والتربوية*. 7(2)، الجزائر: جامعة الوادي، الجزائر. 14-34.