فاعلية النماذج اللوجستية ثنائية التدريج في تحليل البيانات المختلطة وفقا لنظرية الاستجابة للفقرة

د. هيثم فؤاد درويش[i]

تاريخ الاستلام
2023/11/25
تاريخ القبول
2024/1/3

الملخص:

هدفت الدراسة الحالية إلى الكشف عن النموذج اللوجستي الأكثر فاعلية وقدرة على التعامل مع تعدد عدد معالم الفقرة في الاختبار الواحد؛ ولتحقيق ذلك تم اتباع المنهج التجريبي في إجراء الدراسة، حيث تم استخدام بيانات افتراضية تحاكي البيانات المختلطة على اختبار مكون من (60) فقرة افتراضية منها (20) تتبع النموذج أحادي المعلمة، و(20) فقرة تتبع النموذج ثنائي المعلمة، و(20) فقرة تتبع النموذج ثلاثي المعلمة، تمت الاستجابة لها من خلال (1000) مفحوص افتراضي تم توليدهم باستخدام برمجية (WinGen) بحيث تتبع قدراتهم التوزيع الطبيعي بمتوسط حسابي (0) وانحراف معياري (1)، إذ جرى تحليل تلك الاستجابات وفق كل من النماذج اللوجستية الثلاث، وقد بينت النتائج أن نسبة الفقرات المطابقة لصالح النموذج ذي المعالم الأكثر، أي أن نسبة مطابقة الفقرات في علاقة إيجابية مع عدد معالم النموذج، وقد أوصت الدراسة بالاعتماد على النماذج ذات المعالم الأكثر عند تحليل بيانات تتبع نماذج متعددة تظهر بصورة بيانات مختلطة.

الكلمات المفتاحية: النماذج اللوجستية، البيانات المختلطة، نظرية الاستجابة للفقرة، ثنائية التدريج.

[i] الأونروا/الأردن

The Effectiveness of Polytomous Logistic Models in Analyzing Mixed Data According to Item Response Theory

Abstract:

The study aimed to reveal the most effective logistic model and its ability to deal with multiple Item parameters in one test. To achieve this, the experimental approach was followed in conducting the study. Hypothetical data simulating mixed data was used on a test consisting of (60) hypothetical items, including (20) items that follow the one-parameter model, (20) items that follow the two-parameter model, and (20) items that follow the three-parameter model. (1000) virtual subjects who were generated using (WinGen) software answered it. So that their abilities follow a normal distribution with an arithmetic mean (0) and a standard deviation (1), these responses were analyzed according to each of the three logistic models. The results showed that the percentage of matching Items favors the model with the most parameters, that is, the percentage of matching Items is positively correlated with the number of model parameters. The study recommended relying on models with more parameters when analyzing data that follows multiple models that appear as mixed data.

Keywords: Logistic models, mixed data, item response theory, polytomous.

المقدمة:

يعد علم القياس من أبرز العلوم التربوية والنفسية وأهمها؛ فهو العلم الذي يهتم بالاختبارات التربوية والمقاييس النفسية، ويقيّم الخصائص السيكومترية لأدوات القياس فاستخدام أدوات بشكل عشوائي وغير مضبوط يهدد البحث العلمي، ويهدد مصداقية النتائج التي قد يبنى عليها قرارات مهمة ومصيرية.

إذ تختلف الأدوات المستخدمة في البحث التربوي والنفسي وفق الغرض الذي أُعدَّ البحث لأجله ووفقًا للتصميم المستخدم ووفق الأسئلة البحثية التي تترجم مشكلة البحث؛ لذلك تم العناية بالأدوات من خلال اتباع خطوات ثابته لبناء الأداة البحثية كما تم الاهتمام بالفقرات التي تشكل الأداة البحثية واختيارها بعناية، ويتم ذلك من خلال الخصائص السيكومترية للفقرات التي تتمثل في معالمها (Clark & Watson, 2016).

كما تُعدُّ المعاملات والمعالم من الخصائص المميزة للفقرات التي تشير لصلاحيتها ضمن المقياس أو تعديلها أو حتى حذفها، وقد تم الاهتمام باستخراج المعاملات، ومن أهمها الصعوبة والتمييز والتخمين في النظرية الكلاسيكية كخصائص رئيسة للفقرة، وبتطور علم القياس وظهور نظرية الاستجابة للفقرة تم العناية بمعالم الفقرة التي تعدُّ هوية ثابتة للفقرة لا تتغير بتغير عينة الدراسة وفقا للمبدأ المسمى بـ "تحرر الفقرة" (Baker & Kim, 2017).

وبما أن معالم الفقرات وفق نظرية الاستجابة للفقرة تتسم بالثبات، وعدم التغير فقد قدمت نظرية الاستجابة للفقرة نماذج مختلفة للتعامل مع مجموعة الفقرات التي يتضمنها المقياس فظهر النموذج أحادي المعلمة (1PLM: one parametric logistic model) الذي يفترض وجود معلمة واحدة للفقرات وهي الصعوبة، ويفترض قيمة ثابتة لمعلمة التمييز لجميع فقرات المقياس تبلغ (1) ويفترض انعدام التخمين لجميع فقرات المقياس. في حين يفترض النموذج ثنائي المعلمة (2PLM: two parametric logistic model) عدم ثبات معلمة التمييز لفقرات المقياس واختلافها، وفي النموذج ثلاثي المعلمة (3PLM: three parametric logistic model) تختلف قيم معلمة التخمين عن الصفر (Zickar Gibby & Robie, 2004).

وبما أن شكل البيانات ونوعها يختلف من اختبار إلى آخر بالتزامن مع معرفة، وقدرة باني الاختبار على اتباع أحد النماذج اللوجستية الخاصة بنظرية استجابة الفقرة، إذ إن كل نموذج من هذه النماذج يعد صالحًا للتعامل مع نوع معين من البيانات، فعند استخدام النموذج اللوجستي أحادي المعلمة، فإن المعلمة التي لا بد من الاهتمام بها بشكل جذري معلمة الصعوبة، إذ إن الاستجابة ونمطها تتأثر بمعلمة واحدة فقط ألا وهي معلمة الصعوبة، في حين أن الاستجابات التي تتبع النموذج ثنائي المعلمة تتأثر بمعلمتي الصعوبة والتمييز بشكل متزامن وهكذا (Hambleton & Swaminathan, 2013).

كما أن المبدأ الذي تم من خلاله الانتقال من النماذج اللوجستية الأحادية إلى النماذج اللوجستية الثنائية وحتى الثلاثية، قام على افتراض أن البيانات التي تتبع أحد هذه النماذج متوافقة مع عدد معالم هذا النموذج، وبخلاف ذلك فإن اختيار النموذج غير المناسب للبيانات يظهر نسبة عالية من عدم المطابقة للنموذج بالنسبة للفقرات، وفي وجود عدد قليل من الفقرات غير المتوافقة للنموذج فإنه ستظهر كفقرات غير مطابقة سيتم حذفها بالاستناد إلى اختبارات المطابقة للنموذج، وفي حال كان عدد الفقرات غير المطابقة كبيرًا فإن ذلك سيؤدي إلى ظهور خلل في بنية الاختبار من حيث الثبات، ومدى تمثيل الفقرات للمجال الذي أخذت منه (Hambleton, Swaminathan & Rogers, 2006).

ومن هنا فإن احتمالية وجود بيانات مختلطة أي بيانات ناتجة عن فقرات تتبع نماذج مختلفة واردة، فقد تكون تلك البيانات تتبع النموذج أحادي المعلمة أو ثنائي المعلمة أو ثلاثي المعلمة، وهو الأمر الذي يؤكد على ضرورة التحقق من مطابقة الفقرات للنموذج الذي سيتم استخدامه في تقدير معالم الفقرات بأقل خطأ معياري ممكن (حبشي، 2022).
وفي ضوء أن خصائص الفقرات كثيرا ما تختلف عن افتراضات النموذج المستخدم؛ مما يخلق فقرات غير مطابقة للنموذج فوجود فقرات بثلاث معالم ضمن فقرات ثنائية المعلمة يشكل انتهاك لافتراضات النموذج المستخدم؛ وأعني هنا النموذج ثنائي المعلمة، وفي الواقع يعدُّ وجود خليط من الفقرات وفق عدد المعالم وضعا طبيعيا ومن هنا يظهر السؤال الأبرز هل سيتم عزل الفقرات غير المطابقة للنموذج حتى لو شكلت نصف فقرات المقياس (خمنو، 2023).

نظرية الاستجابة للفقرة (Item Response Theory):
هي نظرية إحصائية تستخدم لتحليل نتائج الاختبارات من خلال تحليل الاستجابة للفقرة، وهي تستند إلى فكرة أن احتمال استجابة الفرد بشكل صحيح لفقرة معينة يعتمد على قدرة الفرد وخصائص الفقرة نفسها، إذ جاءت هذه النظرية لمعالجة كثير من أوجه القصور في النظرية الكلاسيكية، باعتمادها على عدد من الافتراضات التي حدت من الأخطاء المنتظمة التي قد تتداخل مع العلامة الحقيقية للفرد (Hambleton & Swaminathan, 2013).
كما عملت هذه النظرية على تحقيق مبدأ تحرر الفقرات (item free) من قدرات الأفراد بافتراض وجود مجموعة كبيرة نسبياً من فقرات الاختبار التي تقيس السمة نفسها، وكذلك بافتراض وجود مجتمع كبير من الأفراد، إذ تكون تقديرات الخصائص السيكومترية للفقرات (مثل معاملات الصعوبة والتمييز) مستقلة عن عينة الأفراد التي استخدمت في تقدير هذه الخصائص؛ أي لا تتأثر معالم الفقرات بالعينة المختارة من المجتمع (Person-free)، كما مكنت هذه النظرية من الحصول على إحصائي مثل (الخطأ المعياري في التقدير) لتقدير درجة الدقة في قياس القدرة لكل فرد الذي قد يختلف من فرد إلى آخر (Reise & Revicki, 2014).

افتراضات نظرية الاستجابة للفقرة:
تستند هذه النظرية على العديد من الافتراضات التي جعلت منها نظرية ذات بناء رياضي منيع وهذه الافتراضات هي:
1. افتراض أحادية البعد (Unidimensionality): يعني أن الفقرات المكونة للاختبارأو المقياس تقيس قدرة واحدة فقط وتكون مسؤولة عن تفسير أداء الفرد في الاختبار، أيضا إذا كان التباين بين المفحوصين في الأداء على الاختبار يعود إلى عامل واحد فقط وهو القدرة المقيسة (Swaminathan, Hambleton & Rogers, 2006).
فاذا افترضنا أن هناك مجموعة من الفقرات عددها (n) صممت للاستخدام في مجتمعات فرعية من المفحوصين، على سبيل المثال مجموعات عرقية متعددة (أ، ب، ج) وافترض توزيعات شرطية لعلامات الاختبار عند مستوى قدرة معين في واحد من مجتمعات المفحوصين الفرعية، وكانت تلك التوزيعات الشرطية عبر المجتمعات الفرعية متطابقة، كان الاختبار أحادي البعد حيث لا توجد فروق في الأداء على الاختبار فيما يتعلق بالقدرة عند مستويات مختلفة (Hambleton & Swaminathan, 2013).

2. الاستقلال الموضعي (Local Independence): يعني أن استجابات المفحوصين على فقرات الاختبار تكون مستقلة إحصائيا (Statistically Independent) عند مستوى قدرة معينة؛ أي أن احتمال الإجابة الصحيحة عن فقرة ما من فقرات المقياس لا يرتبط باحتمال الإجابة الصحيحة عن أي فقرة أخرى، هذا ولا يعني افتراض الاستقلال الموضعي أنه لا توجد أية علاقة ارتباطية بين الفقرات فقد أوضح (سوميثينان وهاميتلون وروجرز) (Swaminathan, Hambleton & Rogers, 2006) أن الاستقلال الموضعي يمكن تحقيقه رغم ارتباط هذه الفقرات عند المستويات المتباينة من السمة أو القدرة المقيسة.
ويرى (كروكر والجانيا) (Crocker & Algina, 1986) أن الاستقلال الموضعي بين فقرتين يتم إذا كان احتمال الإجابة الصحيحة عن كلا المفردتين P (+؛ +) مساوياً لحاصل ضرب احتمال الإجابة الصحيحة عن الفقرة الأولى P1 (+) في حاصل احتمال الإجابة الصحيحة عن الفقرة الثانية P2(+) وكذلك أن يكون احتمال الإجابة الخطأ عن كلاهما P (-,-) مساويا لحاصل ضرب احتمال الإجابة الخطأ عن الفقرة الأولى P1 (-) في احتمال الإجابة الخطأ عن الفقرة الثانية P2(-).
3. منحنى خصائص الفقرة (Item Characteristic Curve): هو اللبنة الأساسية في نظرية استجابة الفقرة حيث أن كافة عناصرها تعتمد عليه، إذ يمثل هذا المنحنى خصائص الفقرة من خلال دالة رياضية تربط بين احتمال الإجابة الصحيحة عن الفقرة (P^{θ}) وقدرة المفحوص (θ) التي يتم قياسها بمجموعة من الفقرات في الاختبار الذي تم بناؤه لتلك الغاية ($Pi\theta$) وهو احتمال أن يجيب فرد قدرته θ إجابة صحيحة عن الفقرة i (Hambleton & Swaminathan, 2013).
4. التحرر من السرعة (Speediness): ينص هذا الافتراض على أن عامل السرعة لا يلعب دورًا في الإجابة عن فقرات الاختبار بمعنى أن اخفاق الأفراد في الاختبار يرجع إلى انخفاض قدراتهم وليس إلى تأثير عامل السرعة في الإجابة، أي أن قدرة المفحوص أقل من صعوبة الفقرة، وهناك طرق عديدة للتحقق من هذا الافتراض، وهو احصاء المفحوصين الذين لم يكملوا الاختبار كنسبة مئوية ممن تقدموا للاختبار ككل (Swaminathan, Hambleton & Rogers, 2006).

نماذج نظرية الاستجابة للفقرة (IRT Models)

نتيجة الاستخدام الواسع لنظرية الاستجابة للفقرة (IRT) تم تطوير نماذج مختلفة تصف العلاقة بين أداء المفحوصين على الاختبار والسمات الكامنة وراء هذا الأداء باقتران رياضي لوغاريتميي، ومن هذه النماذج: النماذج اللوجستية، والنماذج التي تستخدم المنحنى الطبيعي التراكمي، إذ تختلف النماذج اللوجستية باختلاف عدد المعالم في كل نموذج، وأشهر هذه النماذج هي (Hambloteton & Swminthan, 1985):

1. النموذج اللوجستي أحادي المعلمة (one-parameter Logistic model): وهو وذلك النموذج الذي وضعه العالم الدنماركي راش (Rasch) عام (1960م)، وطوّعه للتطبيق العملي العالم الأمريكي (رايت) (Wright)، وهو من أهم نماذج السمات الكامنة وأبسطها، وكان الهدف منه تحقيق خاصية الموضوعية (Objectivity) في القياس النفسي والتربوي، إذ تعني الموضوعية في هذا النموذج المقارنة بين نتيجة التفاعل بين قدرتي فردين مع صعوبة فقرة مناسبة، كما تعني كذلك المقارنة بين صعوبة فقرتين استجاب لهما مفحوص مناسب (Wright & Stone, 1970).

ويوصف بالمعادلة الرياضية الآتية:

$$P_i(\theta) = \frac{1}{1 + e^{-D(\theta_i - b_i)}} \quad \text{..... .... (1)}$$

2. النموذج اللوجستي ثنائي المعلمة (Tow-parameter Logistic model): وهو ذلك النموذج الذي يفترض اختلاف كل من معلمة الصعوبة والتخمين للفقرات، مع اقتراب معلمة التخمين من الصفر، وهو الأمر الذي يلاحظ على كافة الاختبارات؛ إذ من الصعوبة بمكان إيجاد مجموعة من الفقرات ذات معلمة تمييز بدرجة واحدة، تميز بين مستويات السمة أو القدرة التي يقيسها الاختبار، كما يقوم هذا النموذج على افتراض عدم تأثر إجابات المفحوصين عن فقرات الاختبار بعامل التخمين، إذ يسمى هذا النموذج بنموذج "بيرنيوم"؛ ويوصف بالمعادلة الرياضية الآتية (Hambleton, 1990):

$$P_i(\theta) = \frac{1}{1 + e^{-D a_i(\theta_i - b_i)}} \quad \text{..... ....(2)}$$

3. النموذج اللوجستي ثلاثي المعلمة (Three-parameter Logistic model): وهو ذلك النموذج الذي يفترض اختلاف كل من معلمة الصعوبة والتمييز والتخمين للفقرات ويسمى بنموذج "لورد"، وهو بذلك يعد الحالة العامة للنموذجين السابقين الأحادي والثنائي، ويصلح لتدريج الفقرات المختلفة في معالمها مع الأخذ بعين الاعتبار التغلب على مشكلة التخمين من قبل المفحوصين في الاختبارات الموضوعية كالصواب والخطأ والاختيار من متعدد (Birnbaum, 1968).

حيث أضاف "لورد" عليه معلم الخط التقاربي الأدنى (Lower Asymptotic Line)، إذ يحدد هذا المعلم احتمال أن يجيب المفحوص ذو قدره منخفضة إجابة صحيحة عن غالبية الفقرات في الاختبار عن طريق التخمين، ويوصف بالمعادلة الرياضية الآتية (Lord & Nonvick, 1980):

$$P_i(\theta) = c_i + (1 - c_i)\frac{1}{1 + e^{-D a_i(\theta_i - b_i)}} \quad \text{..... .....(3)}$$

الدراسات السابقة:

تناولت العديد من الدراسات البيانات المختلطة:

في دراسة سوبرياتي وفلاني ومولانا (Supriyati, Falani, & Maulana, 2021) تمت المقارنة بين تدريج الأسئلة المختلطة في اختبار للفيزياء تم تطبيقه عينة مكونة من (300) طالب من طلبة المدارس. حيث تم استخدام نماذج نظرية الاستجابة للفقرة الثنائية والمتعددة لتدريج الفقرات بطريقتين: في الطريقة الأول تم التدريج للفقرات التابعة لكل نموذج بشكل منفصل وفي

الطريقة الثانية تم تدريج الفقرات المختلطة. وقد أظهرت نتائج الدراسة ان تدريج الفقرات المختلطة يقدر بشكل أدق من الفقرات المنفصلة من خلال دالة معلوماتية الاختبار.
كما هدفت دراسة الزعابي والمحرزي وعبد الحميد (2020) إلى مقارنة طريقة معادلة الدرجات الحقيقية وطريقة معادلة الدرجات المشاهدة لإجراء معادلة صور الاختبارات المكونة من فقرات مختلطة: (ثنائية الاستجابة ومتعددة الاستجابة)، وذلك من خلال استخدام تصميم المجموعة الواحدة، وكذلك تصميم المجموعات غير المتكافئة ذات الجذع المشترك. حيث تم استخدام اختبار قياس مهارات التقدم في القراءة (PIRLS) لعام 2011، والمكون من 146 فقرة اختبارية (منها 74 فقرة ثنائية الاستجابة، و72 فقرة متعددة الاستجابة)، وتشكل بمجموعها 13 كُتيبًا بحيث يشمل كل كُتيب على نص أدبي ونص علمي، وقد تم تطبيق المقياس على 13 مجموعة، من طلبة الصف الرابع الأساسي بسلطنة عمان، وقد بلغ حجم العينة (10394) طالبًا وطالبة. وتم تقدير معالم الفقرات وقدرات الافراد وفق النموذج المتدرج الاستجابة ثم وضعها على تدريج مشترك، ثم إجراء معادلة للصور من خلال طرق المعادلة وفق نظرية الاستجابة للفقرة، وتحديدا (طريقة معادلة الدرجات الحقيقية، وطريقة معادلة الدرجات المشاهدة)، وقد أظهرت نتائج وجود فروق في الإحصاءات الوصفية لصور اختبار (PIRLS 2011) قبل المعادلة، بينما تقاربت الإحصاءات الوصفية بعد المعادلة بالطريقتين. كما أظهرت نتائج اختبار "ت" للعينات المرتبطة وجود فروق ذات دلالة إحصائية عند مستوى دلالة (0.05)، بين متوسطات القيم المعادلة بطريقتي الدرجات الحقيقية، والدرجات المشاهدة لصالح المعادلة بالدرجات المشاهدة في جميع صور الاختبار، ماعدا الصورة (10)، والتي كانت لصالح المعادلة بالدرجات الحقيقية، بينما لم تظهر الصور (4)، و(5) أي فروق دالة إحصائيًا بين طريقتي المعادلة.
وأجرى حبشي (2022) دراسة للمقارنة بين النماذج أحادية البعد ومتعددة الابعاد من حيث تدريج فقرات الاختبارات المختلطة حيث جمع بين نوعين من الأسئلة: أسئلة الاختيار من متعدد والأسئلة المقالية داخل اختبار واحد فقد هدفت الدراسة إلى المقارنة بين كفاءة النماذج أحادية البعد ومتعددة الأبعاد من حيث تدريج الاختبارات ذات الفقرات المختلطة. حيث تم تطبيق اختبار لقياس القدرة الرياضة على 738 طالبًا وطالبة من طلبة الصف السادس الابتدائي حيث تكون الاختبار من 10 أسئلة من نوع الاختيار من متعدد وخمسة أسئلة مقالية أي أن طول الاختبار بلغ 15 سؤال. ولتدريج أسئلة الاختيار من متعدد تم استخدام النموذج ثنائي المعلم، وتم استخدام نموذج التقدير الجزئي العام لتدريج الأسئلة مفتوحة النهايات بالاعتماد على برمجيتي (R) و (PARSCALE). وقد أظهرت النتائج أن النماذج متعددة الأبعاد قادرة على تدريج الاختبار بفاعلية أكثر من النماذج الأحادية بالرجوع إلى قيم دالة معلوماتية الفقرة وجودة مطابقة النموذج للبيانات.

التعقيب على الدراسات السابقة:
يلاحظ ان الدراسات السابقة قارنت بين قدرة نماذج نظرية الاستجابة للفقرة على تدريج الفقرات بوجود فقرات مختلطة مثل دراسة (سوبرياتي وفلاني ومولانا) (Supriyati, Falani, & Maulana, 2021) ودراسة حبشي (2022). أو أنها قارنت بين طرق معادلة صور الاختبار بوجود فقرات مختلطة، مثل: دراسة الزعابي، والمحرزي، وعبد الحميد (2020). أو أنها قارنت بين نتائج التحليل بفصل الفقرات ودمجها إلا أنها وفي حدود علم الباحث واطلاعه لم تجرِ مقارنة بين فاعلية النماذج اللوجستية في التعامل مع البيانات المختلطة المدمجة.

مشكلة الدراسة وأسئلتها:
عند تطبيق مقاييس مكونة من عدد من الفقرات وبحسب نظرية الاستجابة للفقرة؛ فإنه في الواقع من الصعب أن تتبع جميع الفقرات نموذجا واحدًا فمن جهة يمكن أن تتبع الفقرات نماذج ثنائية، ونماذج متعددة أو من الممكن أن تتبع الفقرات ثنائية التدريج نماذج ثنائية مختلفة أي أنه من الصعب توفير مقياس تتبع جميع فقراته نموذج لوجستي موحد ففي الواقع يحتمل وجود فقرات غير مطابقة للنموذج، ويعدَّ الحل الأمثل هو حذفها من المقياس عندما يكون عددها قليل لكن عندما يكون عدد الفقرات غير المطابقة للنموذج كبيرًا؛ أي عندما يكون المقياس خليط من الفقرات التي تتبع النماذج المختلفة فإنه لا بد من اختيار النموذج الأكثر مناسبة والذي يحقق الحالة الأعم ومن هنا جاءت هذه الدراسة للكشف عن النموذج اللوجستي الأكثر فاعلية، وقدرة على التعامل مع تعدد عدد معالم الفقرة، وتحديدًا فإن هذه الدراسة تجيب عن التساؤل الرئيس والذي ينص على:
سؤال الدراسة: هل تختلف نسبة مطابقة الفقرات للنموذج اللوجستي، وله ثلاثة مستويات (أحادي المعلمة، ثنائي المعلمة، ثلاثي المعلمة) للبيانات المختلطة؟

أهمية الدراسة:
تأتي أهمية هذه الدراسة على وجهين أولهما الجانب النظري حيث إن هذه الدراسة تقارن بين منعة النماذج اللوجستية اتجاه وجود خليط من الفقرات متعددة النماذج كما أنها تزود الأدب المتعلق بنظرية الاستجابة للفقرة بمؤشرات حول صلاحية النماذج اللوجستية ومدى الوثوق بها في ضوء مطابقة الفقرات، وثانيهما الجانب العملي حيث إن هذه الدراسة تزود الباحثين، ومطبقي الاختبارات بمؤشرات لاختيار النموذج اللوجستي الأكثر ملائمة؛ لتحليل بياناتهم عند وجود نسب عالية من عدم المطابقة في فقرات المقياس.

أهداف الدراسة:
تهدف هذه الدراسة الى تقييم فاعلية كل من نماذج نظرية الاستجابة للفقرة اللوجستية من خلال قدرتها على ملائمة البيانات المختلطة بالاعتماد على نسبة الفقرات المطابقة للنموذج. كما تهدف هذه الدراسة إلى المفاضلة بين النماذج اللوجستية بوجود فقرات مختلطة.

المصطلحات المفاهيمية والإجرائية:
النموذج اللوجستي: "معادلة تربط بين قدرة الفرد واحتمالية الاستجابة للفقرة بشكل صحيح في ضوء عدد معالم الفقرة، وتحقق افتراضات نظرية الاستجابة للفقرة" (Verhelst & Glas, 1995: p222).
البيانات المختلطة: البيانات الناتجة من استجابات الأفراد على فقرات تتبع النماذج المختلفة لنظرية استجابة الفقرة باختلاف شكل الاستجابة لتلك الفقرات سواء أكانت اختيار من متعدد أو فقرات ذات نهاية مفتوحة، أو الفقرات التي يتم تصحيحها بطرق مختلفة أو الفقرات المختلفة في عدد معالمها (Rijmen et al., 2003: p188).
فاعلية النموذج اللوجستي (اصطلاحا): قدرة النموذج اللوجستي على التوفيق بين مستوى القدرة واحتمالية الاستجابة الصحيحة لكل من فقرات الاختبار (Verhelst & Glas, 1995: p231).

فاعلية النموذج اللوجستي (إجرائيا): في هذه الدراسة تم تقييم فاعلية النموذج اللوجستي من خلال نسبة مطابقة الفقرات للنموذج.

محددات الدراسة:

تتحدد نتائج هذه الدراسة بالمحددات الآتي:

- تقييم فاعلية النماذج اللوجستية ثناية التدريج فقط.
- محاكاة اختبار افتراضي بطول (1000) مفحوص.
- محاكاة عينة افتراضية بطول (60) فقرو.

الطريقة والإجراءات

منهجية الدراسة: تم اتباع المنهج التجريبي في إجراء البحث، حيث تم استخدام بيانات افتراضية تحاكي البيانات المختلطة والتي تتكون من استجابات الأفراد على اختبار يتكون من فقرات تتبع نماذج معلميه متعددة. وتحليل تلك الاستجابات وفق كل من النماذج اللوجستية الثلاث.

عينة الدراسة:

تكونت عينة الدراسة من (1000) مفحوص افتراضي تم توليدهم باستخدام برمجية (WinGen) بحيث تتبع قدراتهم التوزيع الطبيعي بمتوسط حسابي (0) وانحراف معياري (1).

أداة الدراسة:

تكونت أداة الدراسة من (60) فقرة افتراضية تحاكي اختبار يتكون من فقرات مختلطة (20) منها تتبع النموذج أحادي المعلمة، و(20) فقرة تتبع النموذج ثنائي المعلمة، و(20) فقرة تتبع النموذج ثلاثي المعلمة.

إجراءات الدراسة:

للإجابة عن أسئلة الدراسة تم اتباع الإجراءات الآتية:

1. توليد عينة افتراضية من المستجيبين حجمها (1000) فرد تتبع قدرتهم التوزيع الطبيعي بمتوسط حسابي (0) وانحراف معياري (1)، باستخدام برمجية (WinGen).
2. توليد اختبار افتراضي بطول (60) فقرة بحيث تتبع (20) فقرة للنموذج الأحادي المعلمة بصعوبة تتبع التوزيع الطبيعي المعياري (0، 1) باستخدام برمجية (WinGen). وتتبع (20) فقرة للنموذج الثنائي المعلمة بصعوبة تتبع التوزيع الطبيعي المعياري (0، 1)، وتمييز يتبع التوزيع اللوجستي الطبيعي (Log-Normal) بمتوسط (1) وانحراف (1.7) باستخدام برمجية (WinGen)، وتتبع (20) فقرة للنموذج الثلاثي المعلمة بصعوبة تتبع التوزيع الطبيعي المعياري (0، 1) وتمييز يتبع التوزيع اللوجستي الطبيعي (Log-Normal) بمتوسط (1) وانحراف (1.7)، وتخمين يتبع التوزيع المنتظم ضمن المدى (0 – 0.5) باستخدام برمجية (WinGen).
3. توليد استجابات للأفراد في الخطوة (1) على الاختبار في الخطوة (2)؛ لتتكون مصفوفة الاستجابات الافتراضية، باستخدام برمجية (WinGen).

4. التحقق من افتراض أحادية البعد من خلال التحليل العاملي الاستكشافي باستخدام برمجية (SPSS).
5. التحقق من افتراض الاستقلال الموضعي باستخدام برمجية (LDID).
6. التحقق من افتراض منحنى خصائص الفقرة باستخدام برمجية BILOG.
7. تحليل مصفوفة استجابات الافراد في الخطوة السابقة باستخدام برمجية (بايلوج) (BILOG) وفقا لكل من النماذج اللوجستية الثلاثة.
8. حساب نسبة مطابقة الفقرات لكل من النماذج اللوجستية.
9. فحص اختلاف نسبة مطابقة الفقرات باختلاف النموذج اللوجستي باستخدام برمجية (SPSS).

التحقق من افتراضات نظرية الاستجابة للفقرة كالاتي:

أولًا: التحقق من افتراض أحادية البعد (Unidimensional)

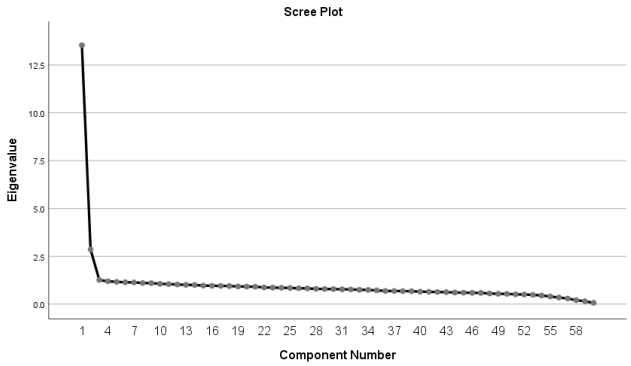
تم التحقق من هذا الافتراض من خلال إجراء اختبار التحليل العاملي الاستكشافي بالاعتماد على تحليل المكونات الرئيسة (Principal component analysis)، لإيجاد قيم الجذور الكامنة (Eigen value)، ونسب التباين المفسر لكل من العاملين الأول والثاني لمصفوفة الاستجابات الافتراضية. كما يأتي:

جدول رقم (1): التحليل العاملي الاستكشافي للمقياس

من حيث	الجذر الكامن	نسبة التباين المفسر
العامل الأول	13.535	%23
العامل الثاني	2.854	%5

تؤكد البيانات الواردة في الجدول رقم (1) الناتجة من عملية التحليل العاملي الاستكشافي، أن البيانات المقياس تحقق افتراض أحادية البعد، حيث إن نسبة الجذر الكامن للعامل الأول إلى نسبة الجذر الكامن للعامل الثاني قد بلغت (4.7)، وهي أكبر من (2) كما أن نسبة التباين المفسر من خلال العامل الأول أكبر من (20%)، وهي مؤشرات تؤكد على تحقق افتراض أحادية البعد (Hatti, 1985)، وهذا ما يمكن ملاحظته أيضًا من خلال التمثيل البياني للجذور الكامنة (Scree Plot) الآتي:

شكل رقم (1): التمثيل البياني للجذور الكامنة (Scree Plot)



يظهر لنا من الشكل رقم (1) والخاص بالتمثيل البياني للجذور الكامنة (Scree Plot) للمقياس الذي تم توليده، أن هناك وجود انعطاف حاد عند وجود عاملين وهو الأمر الذي يدل على وجود سمة واحدة فقط سائدة في المقياس.

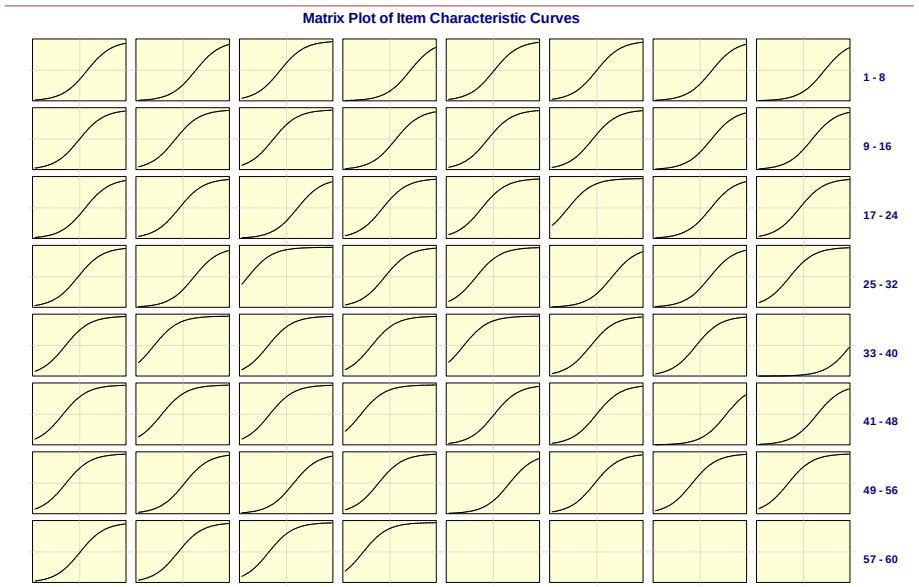
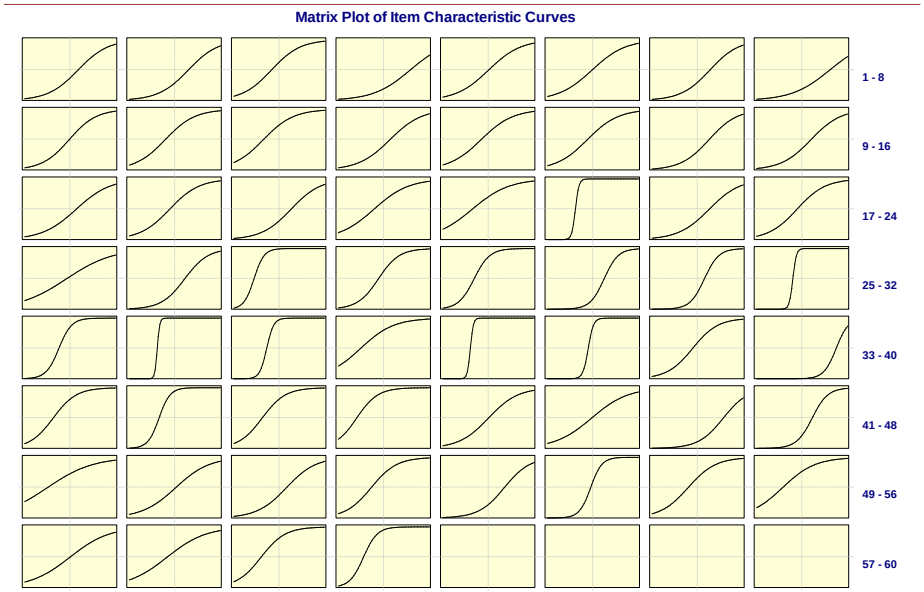
ثانيًا: التحقق من افتراض الاستقلال الموضعي للمقياس

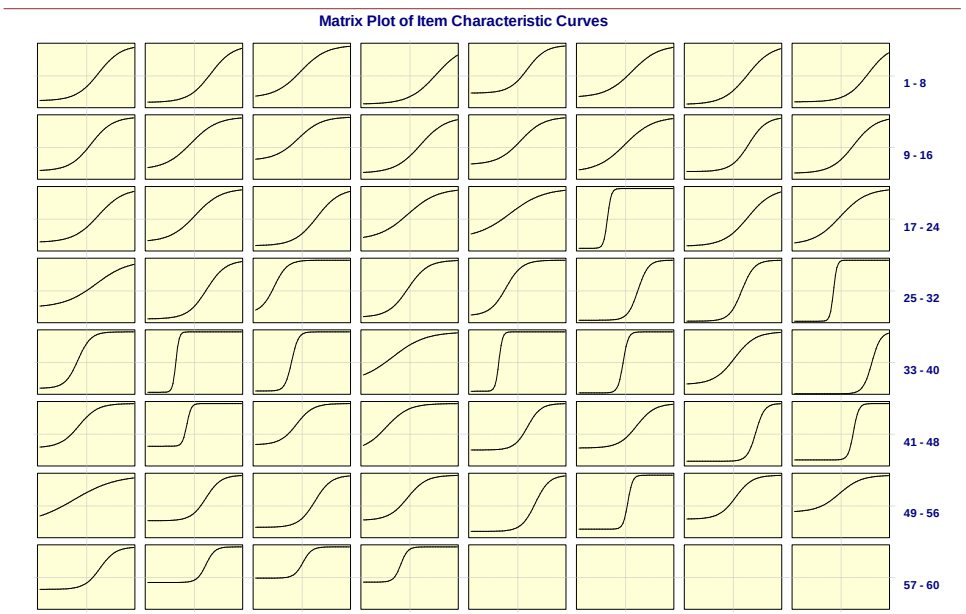
تم التحقق من هذا الافتراض بالاعتماد على اختبار (Z_{Q3})، حيث يعمل هذا الاختبار على الكشف عن أزاوج الفقرات المرتبطة من خلال قيم معاملات ارتباط البواقي لها، حيث تبين أن عدد الأزواج الكلي يساوي (1770) زوجًا، منها (59، 41، 29) زوجًا غير مستقل عند التحليل وفقا للنماذج اللوجستية الثلاثة (أحادي، ثنائي، ثلاثي) على الترتيب بنسبة مئوية بلغت (3.3%، 2.3%، 1.6%) من مجمل الأزواج الكلية، وهذا يبين ان عدد أزاوج الفقرات التي حققت افتراض الاستقلال الموضعي أعلى من أربعة أمثال عدد الأزواج التي حققت التبعية الموضعية والتي بلغت (29، 42، 60) ضعف عند التحليل وفقا للنماذج الثلاثة (أحادي، ثنائي، ثلاثي) على الترتيب، وهذا مؤشر على تحقق افتراض الاستقلال الموضعي (Kim, Cohen & Lin, 2005).

ثالثًا: التحقق من افتراض منحنى خصائص الفقرة

تم التحقق من خلال رسم منحنى خصائص الفقرة باستخدام برمجية BILOG لتحليل البيانات وفقا لكل من النماذج اللوجستية الثلاثة كما هو مبين في الشكل (2).

الشكل (2): منحنى خصائص الفقرة وفقا للنماذج اللوجستية.

النموذج	منحنى خصائص الفقرة
الأحادي	
الثنائي	

الثلاثي	

من الشكل (2) يتبين أن الافتراض متحقق؛ إذ يشير منحنى خصائص الفقرة الى علاقة مطّردة خطية بين القدرة واحتمالية الإجابة عن الفقرة بشكل صحيح تشابه الحرف (S).

نتائج الدراسة ومناقشتها:

النتائج المتعلقة بسؤال الدراسة: **"هل تختلف نسبة مطابقة الفقرات للنموذج اللوجستي وله ثلاثة مستويات (أحادي المعلمة، ثنائي المعلمة، ثلاثي المعلمة) للبيانات المختلطة؟"**

للإجابة عن هذا السؤال تم تحليل مصفوفة الاستجابات الافتراضية باستخدام برمجية BILOG وفقا للنماذج اللوجستية الثلاثة، واستخراج نسبة مطابقة فقرات الاختبار الافتراضي في لكل منها كما هو في الجدول (2):

الجدول (2): نسبة مطابقة الفقرات باختلاف النموذج اللوجستي المستخدم.

	النموذج اللوجستي												
	1PLM				2PLM				3PLM				
الفقرة	مربع كاي	القيمة الاحتمالية	درجة الحرية	المطابقة	مربع كاي	القيمة الاحتمالية	درجة الحرية	المطابقة	مربع كاي	القيمة الاحتمالية	درجة الحرية	المطابقة	
1	11.7	0.166	8	مطابق	21.9	0.0091	9	غير مطابق	22.1	0.0047	8	غير مطابق	
2	5.7	0.6826	8	مطابق	5.4	0.7983	9	مطابق	4	0.8538	8	مطابق	
3	12.3	0.1949	9	مطابق	9.6	0.3817	9	مطابق	10.7	0.2979	9	مطابق	
4	17.9	0.022	8	غير مطابق	9.8	0.2789	8	مطابق	8.2	0.4122	8	مطابق	
5	17.5	0.0421	9	غير مطابق	16.5	0.0565	9	مطابق	15.3	0.0828	9	مطابق	
6	12.6	0.1831	9	مطابق	2.3	0.9848	9	مطابق	3.1	0.9589	9	مطابق	
7	5.9	0.6593	8	مطابق	5.4	0.8022	9	مطابق	7	0.5377	8	مطابق	
8	17.9	0.0222	8	غير مطابق	5.9	0.6637	8	مطابق	9.9	0.2722	8	مطابق	
9	2.6	0.978	9	مطابق	5.6	0.7749	9	مطابق	11.2	0.2596	9	مطابق	
10	17.7	0.0389	9	غير مطابق	17.7	0.0384	9	غير مطابق	7.7	0.5687	9	مطابق	
11	15.8	0.0716	9	مطابق	5.8	0.7549	9	مطابق	10	0.3492	9	مطابق	
12	7.4	0.4903	8	مطابق	12.7	0.1768	9	مطابق	9.5	0.3943	9	مطابق	
13	14.9	0.0937	9	مطابق	4.6	0.866	9	مطابق	3.1	0.9605	9	مطابق	
14	14.3	0.1117	9	مطابق	13.1	0.1602	9	مطابق	9.2	0.4152	9	مطابق	
15	7.3	0.5007	8	مطابق	4.3	0.8909	9	مطابق	4.5	0.811	8	مطابق	
16	4.6	0.8017	8	مطابق	11.2	0.2627	9	مطابق	5.6	0.7766	9	مطابق	
17	16.7	0.033	8	غير مطابق	6.1	0.728	9	مطابق	4.7	0.8616	9	مطابق	
18	7.1	0.5251	8	مطابق	5.5	0.7864	9	مطابق	3.5	0.9422	9	مطابق	
19	9.6	0.2938	8	مطابق	7.6	0.5788	9	مطابق	7.7	0.4657	8	مطابق	
20	27.1	0.0013	9	غير مطابق	23.1	0.006	9	غير مطابق	11.4	0.2461	9	مطابق	
21	24.4	0.0037	9	غير مطابق	4.6	0.8666	9	مطابق	5.1	0.8261	9	مطابق	
22	145	0	3	غير مطابق	1.2	0.5447	2	مطابق	3.5	0.1722	2	مطابق	
23	5.6	0.6927	8	مطابق	0.6	0.9999	9	مطابق	8.2	0.5171	9	مطابق	
24	7.3	0.5086	8	مطابق	13.1	0.1574	9	مطابق	11.2	0.2592	9	مطابق	
25	56.1	0	9	غير مطابق	10.1	0.3457	9	مطابق	9.8	0.3672	9	مطابق	
26	6.6	0.5851	8	مطابق	5	0.755	8	مطابق	5.1	0.7493	8	مطابق	
27	44.4	0	6	غير مطابق	3.7	0.2933	3	مطابق	2	0.7322	4	مطابق	
28	12.7	0.1218	8	مطابق	7.1	0.6258	9	مطابق	6.3	0.7062	9	مطابق	
29	25.5	0.0013	8	غير مطابق	6.1	0.6384	8	مطابق	3.1	0.9248	8	مطابق	
30	69.2	0	7	غير مطابق	4.8	0.6813	7	مطابق	3.4	0.8465	7	مطابق	
31	62.8	0	6	غير مطابق	4.7	0.6981	7	مطابق	6.5	0.3682	6	مطابق	
32	180.7	0	3	غير مطابق	11	0.0041	2	غير مطابق	4.6	0.098	2	مطابق	
33	55.9	0	6	غير مطابق	5.5	0.4866	6	مطابق	7.6	0.3657	7	مطابق	
34	172.8	0	3	غير مطابق	0	0	0	غير مطابق	0.2	0	0	غير مطابق	

35	156.5	0	4	غير مطابق		2.3	0.6759	4	مطابق		7.2	0.2073	5	مطابق
36	21.4	0.0061	8	غير مطابق		8.7	0.4655	9	مطابق		22.6	0.0071	9	غير مطابق
37	158.5	0	3	غير مطابق		0.7	0.7173	2	مطابق		4	0.2573	3	مطابق
38	238.2	0	4	غير مطابق		12.1	0.0164	4	غير مطابق		19.3	0.0007	4	غير مطابق
39	6.3	0.7061	9	مطابق		5	0.8311	9	مطابق		2.5	0.9804	9	مطابق
40	26.1	0	4	غير مطابق		2.3	0.5067	3	مطابق		5.8	0.2155	4	مطابق
41	6.1	0.6313	8	مطابق		7.5	0.5819	9	مطابق		10.7	0.2167	8	مطابق
42	44.5	0	5	غير مطابق		34.6	0	4	غير مطابق		2.9	0.7129	5	مطابق
43	6.3	0.6163	8	مطابق		5.9	0.6583	8	مطابق		2.6	0.9589	8	مطابق
44	3.2	0.861	7	مطابق		3.3	0.8516	7	مطابق		6.5	0.5905	8	مطابق
45	51.7	0	9	غير مطابق		23.4	0.0054	9	غير مطابق		4.5	0.8765	9	مطابق
46	44	0	9	غير مطابق		11.7	0.233	9	مطابق		3.4	0.9488	9	مطابق
47	37.2	0	7	غير مطابق		38.7	0	8	غير مطابق		4	0.8595	8	مطابق
48	79.1	0	7	غير مطابق		77.7	0	7	غير مطابق		5.4	0.6141	7	مطابق
49	43.4	0	9	غير مطابق		3.2	0.9545	9	مطابق		6.3	0.713	9	مطابق
50	33.2	0.0001	9	غير مطابق		20.1	0.017	9	غير مطابق		4.6	0.8681	9	مطابق
51	14.3	0.0746	8	مطابق		10	0.3527	9	مطابق		2.7	0.9757	9	مطابق
52	6	0.7417	9	مطابق		14.7	0.1009	9	مطابق		11.5	0.2444	9	مطابق
53	13.9	0.0832	8	مطابق		8.6	0.373	8	مطابق		4.9	0.7677	8	مطابق
54	136.7	0	7	غير مطابق		54	0	5	غير مطابق		9.3	0.1577	6	مطابق
55	15.3	0.0835	9	مطابق		12.7	0.1758	9	مطابق		5.3	0.8061	9	مطابق
56	18.2	0.0197	8	غير مطابق		6.5	0.6871	9	مطابق		5.1	0.8287	9	مطابق
57	40.8	0	9	غير مطابق		13.5	0.1394	9	مطابق		4.5	0.8738	9	مطابق
58	81.1	0	9	غير مطابق		52.7	0	9	غير مطابق		4.9	0.7706	8	مطابق
59	23.8	0.0013	7	غير مطابق		34	0	7	غير مطابق		11.8	0.1058	7	مطابق
60	22.4	0.0004	5	غير مطابق		24.5	0.0002	5	غير مطابق		6.6	0.3582	6	مطابق
عدد الفقرات المطابقة			26					45					56	
نسبة الفقرات المطابقة			%43					%75					%93	

من قيم نسب المطابقة يلاحظ وجود اختلاف في نسب المطابقة باختلاف النموذج اللوجستي المستخدم في التحليل حيث يلاحظ أن (26) فقرة فقط مطابقة للنموذج اللوجستي الأحادي المعلمة، وأصبح عدد الفقرات المطابقة (45) فقرة عندما تم استخدام النموذج ثنائي المعلمة، وارتفع عدد الفقرات الى (56) فقرة عند استخدام النموذج الثلاثي المعلمة وبنسبة مطابقة بلغت (93%). ولفحص دلالة الاختلاف بين النسب الثلاث تم استخدام اختبار مربع (كاي) لفحص الفروق بين نسبتين وأكثر والجدول (3) يبين النتائج.

الجدول (3): مربع كاي لاختبار الفروق بين نسب المطابقة.

مربع كاي	درجة الحرية	القيمة الاحتمالية
36.957	2	0.000

يبين اختبار مربع (كاي) وجود اختلاف جوهري ذي دلالة إحصائية في نسبة مطابقة الفقرات باختلاف النموذج اللوجستي المستخدم، ولمعرفة لصالح من كانت الفروق تم اختبار الاختلاف بين أزواج النسب كمقارنة بعدية باستخدام اختبار z لفحص الاختلاف بين النسب من خلال المعادلة الرياضية الآتية:

$$Z = \frac{\hat{P}_1 - \hat{P}_2}{\sqrt{\hat{P}.\hat{q}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$$

حيث أن:
$\hat{P}1$: نسبة الفقرات المطابقة في المقياس الأصلي.
$\hat{P}2$: نسبة الفقرات المطابقة بعد عملية الفقد والمعالجة.
$\hat{q}$: نسبة الفقرات غير المطابقة في المقياس الأصلي والمقياس المعالج.
$\hat{P}$: نسبة الفقرات المطابقة في صورتي المقياس الأصلي وبعد المعالجة.
n_1: عدد الفقرات في المقياس الأصلي.
n_2: عدد الفقرات في المقياس المعالج.

وقد كانت النتائج كما يأتي:

الجدول (4): المقارنات البعدية لاختلاف نسب مطابقة الفقرات باختلاف النموذج اللوجستي.

	النموذج اللوجستي		
*	1PLM	2PLM	3PLM
1PLM	-	-3.529	-5.887
2PLM	0.000	-	-2.751
3PLM	0.000	0.006	-

* القيم فوق القطر تمثل اختبار z والقيم أسفل القطر تمثل الدلالة الإحصائية لاختبار z.

من خلال الاطلاع على المقارنات البعدية يتبين أن نسبة الفقرات المطابقة لصالح النموذج ذي المعالم الأكثر، أي أن نسبة مطابقة الفقرات في علاقة إيجابية مع عدد معالم النموذج، ويمكن تفسير ذلك من خلال قدرة النموذج على التعامل مع المعالم التي تلعب دورا في كيفية الإجابة عن تلك الفقرات، فالنموذج أحادي المعلمة يتضمن معلمة الصعوبة وتُعدُّ الإجابة عن الفقرة مبنية على هذه المعلمة فقط، ووجود معالم أخرى تنتهك النموذج، وتصبح الدالة اللوجستية المرتبطة بهذا النموذج غير قادرة على التنبؤ بكيفية الاستجابة، حيث إن المعالم الإضافية تلعب دورا غير متوقع. وعليه لم يكن النموذج أحادي المعلمة فعّالًا في التعامل مع الفقرات المختلطة التي تتضمن معالم أخرى: مثل

التمييز والتخمين، في حين كان النموذج ثلاثي المعلمة قادرًا على التعامل مع الفقرات المتوافقة معه ومع الفقرات المتوافقة مع كل من النموذجين الأحادي المعلمة والثنائي المعلمة؛ إذ إن هذا النموذج مبني على وجود ثلاث معالم هي: الصعوبة التي تعتبر المعلمة المشتركة في كل النماذج ثنائية التدريج بالإضافة إلى معلمة التمييز والتخمين. أما بالنسبة لمعلمة التمييز فإنها ضمن الفقرات التي تتبع النموذج الأحادي ومن وجهة نظر النموذج الثلاثي موجودة وتتوزع بوسط حسابي (1)، وانحراف معياري (0)، وبإضافتها إلى فقرات تتوافق من النموذجين الثنائي والثلاثي يصبح توزيع معلمة التمييز لجميع الفقرات مناسب للنموذج الثلاثي والثنائي، وهذا ما جعل النموذج الثنائي أكثر فعالية من النموذج الأحادي. بالإضافة إلى اعتبار أن التخمين يتوزع بوسط حسابي (0) في النموذجين الأحادي والثنائي من وجهة نظر النموذج الثلاثي وبجمع الفقرات يصبح توزيع معلمة التخمين يختلف عن الصفر، وهذا متوافق مع النموذج الثلاثي بمقابل النموذجين الأحادي والثنائي؛ وهذا ما جعل من النموذج الثلاثي أكثر النماذج فعالية في التعامل مع البيانات المختلطة وجعلة النموذج الأكثر عمومية والأكثر تهيئة للتعامل مع البيانات المختلطة بشرط وجود فقرات تتوافق مع هذا النموذج حيث بلغت نسبة الفقرات التي تتبع لكل نموذج 0.33 من مجموع الفقرات.

التوصيات:

وفي ضوء النتائج التي تم التوصل لها فان الباحث يوصي بما يأتي:

- الاعتماد على النماذج ذات المعالم الأكثر عند تحليل بيانات تتبع نماذج متعددة تظهر بصورة بيانات مختلطة.
- التحقق من مدى مطابقة البيانات للنموذج كمؤشر على مدى مطابقة النموذج للبيانات قبل الاعتماد على نموذج معين في تقدير معالم الفقرات وقدرات الأفراد.
- إجراء دراسات مشابهة باستخدام النماذج متعددة التدريج وصيغها المعممة وبنسب مطابقة مختلفة.

المصادر والمراجع:

المراجع العربية:

حبشي، حسين (2022). تدريج الاختبارات مختلطة الفقرات باستخدام نظرية الاستجابة للمفردة مقارنة بين النماذج الأحادية ومتعددة الأبعاد. المجلة المصرية للدراسات النفسية، 117(32)، 23-40.

خمنو، دنيا (2023). الخصائص السيكومترية لأدوات القياس في علم النفس. مجلة الروائز، 7(1)، 21-40.

الزعابي، أمل والمحرزي، راشد وعبد الحميد، حسن (2020). معادلة كتيبات اختبار الدراسة الدولية لقياس مهارات القراءة (PIRLS 2011) بسلطنة عمان باستخدام نظرية الاستجابة للمفردة. مجلة العلوم التربوية جامعة قطر، 15(15)، 68-93.

المراجع الأجنبية:

Baker, F. B., & Kim, S. H. (2017). The basics of item response theory using R (pp. 17-34). New York: Springer.

Birnbaum, A. (1968). Some latent trait models and their use in inferring an examinee's ability. In F. M. Lord & M. R. Novick (Eds.), Statistical theories of mental test scores. Reading, MA: Addison-Wesley.

Clark, L. A., & Watson, D. (2016). Constructing validity: Basic issues in objective scale development.

Crocker, L., & Algina, J. (1986). Introduction to classical and modern test theory. Holt, Rinehart and Winston, 6277 Sea Harbor Drive, Orlando, FL 32887.

Hambleton, R. (1990). Item response theory: introduction and bibliography. Psicothema, 2(1), 97-107.

Hambleton, R. K., & Swaminathan, H. (2013). Item response theory: Principles and applications. Springer Science & Business Media.

Lord, F. & Nonvick, M. (1980). Application of item Response Theory to practical Testing problem. Hillsdale, NJ: Lawrence Erlbaum Associates.

Reise, S. P., & Revicki, D. A. (Eds.). (2014). Handbook of item response theory modeling: Applications to typical performance assessment. Routledge.

Rijmen, F., Tuerlinckx, F., De Boeck, P., & Kuppens, P. (2003). A nonlinear mixed model framework for item response theory. Psychological methods, 8(2), 185-205.

Supriyati, Y., Falani, I., & Maulana, S. (2021, March). The information function of mixed-format test of physics learning outcomes measurement. In AIP Conference Proceedings (Vol. 2320, No. 1). AIP Publishing.

Swaminathan, H., Hambleton, R. K., & Rogers, H. J. (2006). 21 Assessing the Fit of Item Response Theory Models. Handbook of statistics, 26, 683-718.

Verhelst, N. D., & Glas, C. A. (1995). The one parameter logistic model. In Rasch models: Foundations, recent developments, and applications (pp. 215-237). New York, NY: Springer New York.

Wright, D. & Stone, M. (1979). Best Test Design a handbook for Rasch measurement. Chicago: MESA Press.

Zickar, M. J., Gibby, R. E., & Robie, C. (2004). Uncovering faking samples in applicant, incumbent, and experimental data sets: An application of mixed-model item response theory. Organizational Research Methods, 7(2), 168-190.