العدد الأمثل لفئات الاستجابة وترتيبها وفقاً لدالة المعلومات ومعالم المفردات والأفراد باستخدام نظرية الاستجابة للمفردة الاختبارية

The optimal number of response categories and their order according to the information function, Item and individuals' parameters using Item response theory

إعداد

بدر بن عبدالله بن صالح العيدان

Bader Abdullah Aleidan

تخصص علم النفس التربوي - مسار القياس والتقويم - كلية اللغات والعلوم الإنسانية- جامعة القصيم

أ.د/ ربيع عبده أحمد رشوان

Dr.Rabie Abdu Ahmad Rashwan

أستاذ علم النفس التربوي-كلية اللغات والعلوم الإنسانية-جامعة القصيم

Doi: 10.21608/jasep.2025.435261

استلام البحث : ٢ / ٤ /٢٠٢٥

قبول النشـــر: ٦ / ٥ / ٢٠٢٥

العيدان، بدر بن عبدالله بن صالح ورشوان، ربيع عبده أحمد (٢٠٢٥). العدد الأمثل لفئات الاستجابة وترتيبها وفقًا لدالة المعلومات ومعالم المفردات والأفراد باستخدام نظرية الاستجابة للمفردة الاختبارية. ***المجلة العربية للعلوم التربوية والنفسية***، المؤسسة العربية للتربية والعلوم والآداب، مصر، ٩(٥٠)، ٨٧ – ١٣٨.

http://jasep.journals.ekb.eg

العدد الأمثل لفئات الاستجابة وترتيبها وفقًا لدالة المعلومات ومعالم المفردات والأفراد باستخدام نظرية الاستجابة للمفردة الاختبارية

المستخلص:

هدفت إلى الكشف عن مدى اختلاف معالم المفردات ودالة المعلومات ومؤشر ثبات المقياس واختلاف تقديرات الأفراد والأخطاء المعيارية في تقديرها باختلاف عدد فئات الاستجابة وترتيبها والتفاعلات المشتركة بينهما. وقد أظهرت النتائج أنه: لا توجد فروق دالة إحصائيًا في معلم التمييز ترجع لاختلاف عدد فئات الاستجابة وترتيبها، وكانت قيمة مربع إيتا الجزئي تساوي ٠.٠٤٩ وهو حجم تأثير ضعيف، مما يؤكد أن الفروق في معلم التمييز وفقًا لتأثير عدد فئات الاستجابة وترتيبها فروق ضعيفة، بينما توجد فروق دالة إحصائياً عند مستوى دلالة ٠.٠١ في الأخطاء المعيارية في تقدير معلم التمييز ترجع لاختلاف عدد فئات الاستجابة وترتيبها، وكانت قيمة مربع إيتا الجزئي تساوي ٠.٦٦٥ وهو حجم تأثير كبير، مما يؤكد أن الفروق في الأخطاء المعيارية لتقدير معلم التمييز وفقًا لتأثير عدد فئات الاستجابة وترتيبها فروق كبيرة، أما بالنسبة لدالة المعلومات فقد أظهرت النتائج وجود فروق دالة إحصائياً عند مستوى دلالة ٠.٠١ في دوال معلومات المفردات ترجع لاختلاف عدد فئات الاستجابة وترتيبها، وكانت قيمة مربع إيتا الجزئي تساوي ٠.٦٦ وهو حجم تأثير كبير، مما يؤكد أن الفروق في دوال المعلومات وفقا لتأثير عدد فئات الاستجابة وترتيبها فروق كبيرة، وأن معاملات الثبات تزداد بزيادة عدد فئات الاستجابة، ومعاملات الثبات في حالة الترتيب التصاعدي (من السلبي إلى الإيجابي) أعلى منها في حالة الترتيب التنازلي (من الإيجابي للسلبي) عند نفس عدد فئات الاستجابة، وبينت النتائج أنه توجد فروق دالة إحصائيًا عند مستوى دلالة ٠.٠١ في قدرات الأفراد ترجع لاختلاف عدد فئات الاستجابة وترتيبها، كما أنه توجد فروق دالة إحصائيًا عند مستوى دلالة ٠.٠١ في الأخطاء المعيارية في تقدير قدرات الأفراد ترجع لاختلاف عدد فئات الاستجابة وترتيبها.

ABSTRACT:

The study investigated the optimal number and ordering of response categories based on the item information function and the parameters of both items and individuals, utilizing Item Response Theory (IRT). The primary objective was to examine the extent to which item parameters, information functions, scale reliability indices, person ability estimates, and their standard errors differ as a result of variations in the number and ordering

ISSN: 2537-0464 eISSN: 2537-0472

of response categories, as well as the interactions between these variables .The findings indicated that there were no statistically significant differences in the item discrimination parameter attributable to the number or ordering of response categories. The partial eta squared value was 0.049, reflecting a small effect size, which suggests that the variations in the discrimination parameter due to changes in category number and order are minimal. However, there were statistically significant differences at the 0.01 significance level in the standard errors of the discrimination parameter estimates, associated with changes in the number and order of response categories. The partial eta squared value for this effect was 0.665, indicating a large effect size, thus confirming that standard error estimates are substantially influenced by both the number and ordering of response categories .Regarding the item information function, results showed significant differences at the 0.01 level, with a partial eta squared of 0.66, also suggesting a large effect. This confirms that item information is significantly impacted by the number and ordering of response categories. The reliability coefficients were found to increase with the number of response categories.

مقدمة الدراسة:

تُعد المقاييس النفسية من أهم وسائل جمع البيانات، للوصول للمعلومات عن المتغيرات النفسية الكامنة، حيث يتم استخدام أدوات قياس ذات طرق مختلفة للاستجابة، وتعد محاولات قياس سمات الشخصية والاتجاهات النفسية قديمة قِدم تقنيات قياس القدرات العقلية، ومع ذلك لا يمكن الادعاء بأنها حققت نجاحًا مماثلاً، حيث توجد الكثير من الصعوبات التي يجب مواجهتها عندما يتم التعامل مع الجوانب المختلفة للسلوك (زارع، ٢٠٢٢)؛ ومن هنا تحظى جودة أداة القياس بأهمية كبيرة من قبل مصممي أدوات القياس، وكذلك من يعتمدون على نتائجها في اتخاذ القرار، كما تركز المعايير المستخدمة لتحديد جودة أداة القياس بشكل عام على مجالات تصميم أداة القياس، وتحليل وتفسير نتائجها (Murphy, 2021).

ISSN: 2537-0464 eISSN: 2537-0472

ومن المشكلات المنهجية التي تقابل علماء النفس في موقف القياس المعتمد على التقدير الذاتي، خاصة في قياس الاتجاهات وسمات الشخصية والنواحي الانفعالية، مشكلة تزييف أو تشويه الاستجابة إذ أن المستجيب في بعض الحالات يتأثر بعوامل ليس لها علاقة بمضمون المثير المقدم، فتتدخل عدة أمور في تحديد الاستجابة، كنزعة القبول في مقابل الرفض لفئة الاستجابة، أي أنه يجيب بنعم أو صواب دائماً، أو أن يجيب بلا أو خطأ دائماً، وكذلك التطرف في مقابل الاعتدال، وهو ما يعكس نزعة المستجيب للاستجابة بشكل نمطي على فقرات المقياس، وهو ما يدل على أن الاستجابات في هذه الحالات ستكون غير صادقة وتؤدي إلى نتائج غير دقيقة وغير عادلة (عوض الله، ٢٠٢١).

وعلاوة على الانتشار الواسع لاستخدام مقياس ليكرت Likert لقياس الاتجاهات والآراء والميول، إلا أنه مازال الجدل قائماً بين الباحثين في مجال منهجية البحوث والإحصاء والقياس النفسي، حول عدد فئات الاستجابة، والذي قد يؤثر في استجابة الأفراد، ومن ثم يؤثر في جودة البيانات، وفي النتائج التي يتم التوصل إليها (إسماعيل، ٢٠١٩).

وتعيين فئات الاستجابة من أكثر النقاط الجدلية بين المتخصصين في هذا الشأن، وتحديد فئات الاستجابة يعد من أكثر التحديات التي تواجه الباحثين أثناء تصميم أدوات القياس المستخدمة في العلوم الإنسانية والنفسية والتربوية، الأمر الذي يتطلب إجراء مزيداً من البحوث في هذا الصدد (Danner et al., 2016)؛ ويستمر الجدل بين الباحثين في هذا السياق حيث يذكر Asu'n et al. (2016) أنه منذ قدم ليكرت أسلوب تقدير استجابات الأفراد على بنود أدوات القياس في عام ١٩٣٢م، فإنه مازال هناك جدل بين الباحثين فيما يتعلق بالعدد الأمثل لفئات الاستجابة مع التأكيد على أن العدد الأكثر استخداما هو ٤-٧ فئات.

ومن الواضح مما سبق أن تقييم تأثير عدد فئات الاستجابة وتحديد العدد الأمثل والترتيب المناسب لها، كان منصباً على استخدام إجراءات النظرية الكلاسيكية مثل مفاهيم الثبات والصدق العاملي والاتساق الداخلي، بينما قضية العدد الأمثل لفئات الاستجابة وترتيبها لم يتم تقييمه بالشكل الكافي وفق مفاهيم وإجراءات نظرية الاستجابة للمفردة الاختبارية، وهو ما تحاول الدراسة الحالية الكشف عنه.

مشكلة الدراسة:

يؤكد علام (٢٠١٥) على تعدد مصادر الخطأ المنتظم والعشوائي في القياس السلوكي والتي منها ما يتعلق بخصائص الاختبار، ومنها ما يتعلق بخصائص الموقف الاختباري، ومنها ما يتعلق بالسمات النفسية للمفحوص، وحالته أثناء موقف القياس.

ISSN: 2537-0464 eISSN: 2537-0472

وقضية تحديد عدد فئات الاستجابات في المقياس من قبل الباحثين لا يتم وفق أسس علمية سليمة في غالبية الأحيان، وإنما يتم وفقا للآراء الشخصية أو بطريقة عشوائية وأن المتتبع للدراسات والبحوث والرسائل العلمية في التخصصات المختلفة وخاصة النفسية والتربوية المنشورة في الدوريات العلمية يجد أن طريقة ليكرت تكاد تكون الطريقة الوحيدة المستخدمة في تدريج أدوات القياس وكذلك يمكن ملاحظة التباين الواضح في عدد فئات الاستجابة التي يستخدمها كل باحث، فهناك من يستخدم استجابتين، وهناك من يستخدم ثلاث، وآخرون يزيدون في عدد فئات الاستجابة حتى سبعة أو عشرة فئات استجابة، ودافع الذين يفضلون فئات استجابة أقل هو تقليل الوقت والجهد في التصحيح والتحليل، دون علم بتأثير ذلك على الخصائص السيكومترية للأداة المستخدمة، في حين أن من يستخدم عدد فئات استجابة أكثر لا يستخدمها عن وعي بأنه يمكن أن تؤثر على الخصائص السيكومترية للمقياس (حسين، ٢٠٠٧).

وقد اتجهت بعض الدراسات إلى زيادة عدد فئات الاستجابة للمقياس حيث أنها تعطي قدر عالي من الصدق والثبات وتقلل من أخطاء القياس كما تعطي خيارات أكثر وتزيد من القدرة على إيجاد العلاقات بين المتغيرات، بينما ترى الدراسات الأخرى أن زيادة عدد فئات الاستجابة يقلل من الثبات في حين ترى دراسات أخرى أن ثبات المقياس وصدقه وخصائصه السيكومترية لا يعتمد على عدد فئات الاستجابة (الزهراني،٢٠٢١)؛ كما اختلف الباحثون حول عدد فئات الاستجابة التي يجب تضمينها في المقاييس، فلا يستحسن استعمال فئات استجابة عددها زوجي كما لا داعي أن تزيد فئات الاستجابة عن خمس فئات (Weng, 2004).

ويتضح مما سبق أن هناك العديد من الدراسات والأبحاث التي تناولت تأثير عدد فئات الاستجابة على الخصائص السيكومترية لأدوات القياس، والتي يغلب عليها أنها اعتمدت في دراستها على إجراءات الاتجاه الكلاسيكي والقليل منها طبق إجراءات النظرية الحديثة، كما وأن نتائج الدراسات تفاوتت في تحديد عدد معين وترتيب واحد لفئات الاستجابة، ولذا ونظرا لقلة الأبحاث في البيئة العربية التي تطرقت لتقييم عدد فئات الاستجابة وترتيبها وفق النظرية الحديثة للقياس، إضافة إلى أن للنظرية الحديثة نماذج تقدم معلومات أكثر دقة مقارنة بأدوات النظرية الكلاسيكية، فمن هنا دعت الحاجة للقيام بدراسة العدد الأمثل لفئات الاستجابة وترتيبها، وفقا لدالة المعلومات ومعالم المفردات والأفراد باستخدام نظرية الاستجابة للمفردة الاختبارية، وتحديداً نموذج الاستجابة المتدرجة.

أسئلة الدراسة:

تتحدد مشكلة الدراسة الحالية في الإجابة عن الأسئلة التالية:

ISSN: 2537-0464 eISSN: 2537-0472

١- هل تختلف معالم المفردات والأخطاء المعيارية في تقديرها باختلاف عدد فئات الاستجابة وترتيبها؟.
٢- هل تختلف دالة المعلومات لمفردات المقياس في تقديرها باختلاف عدد فئات الاستجابة وترتيبها؟.
٣- هل تختلف معاملات ثبات المقياس باختلاف عدد فئات الاستجابة وترتيبها؟.
٤- هل تختلف تقديرات قدرات الأفراد والأخطاء المعيارية في تقديرها باختلاف عدد فئات الاستجابة وترتيبها؟.

أهداف الدراسة:

هدفت الدراسة الحالية إلى الكشف عن:

١- مدى اختلاف معالم المفردات والأخطاء المعيارية في تقديرها باختلاف عدد فئات الاستجابة وترتيبها والتفاعلات المشتركة بينهما.
٢- مدى اختلاف دالة المعلومات للمقياس والمفردات والأخطاء المعيارية في تقديرها باختلاف عدد فئات الاستجابة وترتيبها والتفاعلات المشتركة بينهما.
٣- مدى اختلاف تقديرات مؤشرات ثبات المقياس باختلاف عدد فئات الاستجابة وترتيبها والتفاعلات المشتركة بينهما.
٤- مدى اختلاف تقديرات قدرات الأفراد والأخطاء المعيارية في تقديرها باختلاف عدد فئات الاستجابة وترتيبها والتفاعلات المشتركة بينهما.

أهمية الدراسة:

تمثلت أهمية الدراسة الحالية في التالي:

١- الأهمية النظرية:

- تزويد الجهات المختصة في الجامعات والمراكز البحثية بصورة واضحة عن العدد الأمثل لفئات الاستجابة وترتيبها.
- الوصول لصورة واضحة فيما إذا كانت الممارسات الحالية في تحديد عدد فئات الاستجابة وترتيبها تتسق مع الإجراءات التي تضمن دقة التقدير من عدمه.
- تزويد المكتبة العربية بدراسة حديثة عن تأثير عدد فئات الاستجابة وترتيبها على جودة أدوات القياس.

٢- الأهمية التطبيقية:

- قد تفيد نتائج الدراسة الحالية الباحثين والمختصين في القياس من خلال توفير المعلومات التي تساعدهم في دقة تصميم وإعداد أدوات القياس.
- نتائج الدراسة الحالية قد تفيد المؤسسات المختلفة التي تستخدم المقاييس في قطاعاتها بما يحسن من أداءها، بالاعتماد على أدوات قياس يتم تصميمها وفقاً للعدد الأمثل لفئات الاستجابة وترتيبها.

ISSN: 2537-0464 eISSN: 2537-0472

- يمكن أن تساهم نتائج الدراسة الحالية في زيادة جودة تأهيل المشرفين والقائمين على إعداد الاختبارات والمقاييس، خاصة ما يستخدم منها في الجهات الحكومية ويتوقف عليه اتخاذ قرار، وذلك بالمساهمة في توجيه تدريبهم على إجراءات النظرية المعاصرة في تحديد العدد الأمثل لفئات الاستجابة وترتيبها.

مصطلحات الدراسة:

١- عدد فئات الاستجابة Number of Response Categories:

هو عدد الاختيارات التي تكون أمام كل مفردة من مفردات المقياس، والتي يتوجب على المستجيب أن يختار منها ما يتفق مع وجهة نظره أو وفقا لما يشعر به (حسين، ٢٠٠٧)؛ ويشير عبد السميع (٢٠٢٠) إلى عدد فئات الاستجابة بأنه عدد اختيارات الاستجابة طبقاً لأسلوب ليكرت، والتي يختار المشاركون منها طبقاً لما يتفق مع وجهة نظرهم وتقتصر في الدراسة الحالية على عدد الفئات (٣، ٤، ٥، ٦، ٧) فئات للاستجابة.

٢- ترتيب فئات الاستجابة Order of Response Categories:

ويقصد به الترتيب الذي تقدم به الاستجابات، وبصفة عامة هناك الترتيب الإيجابي والترتيب السلبي، والترتيب العشوائي والترتيب المختلط (Deutsch & Strack, 2019) وستقتصر الدراسة الحالية على الترتيب التصاعدي ويقصد به من السلبيى إلى الإيجابي (غير موافق بشدة إلى موافق بشدة) والترتيب التنازلي ويقصد به من الإيجابي إلى السلبي (موافق بشدة إلى غير موافق بشدة) لفئات الاستجابة.

٣- دالة المعلومات Information Function:

تعبر عن دالة رياضية تصف العلاقة بين مستوى القدرة ومقدار المعلومات التي تقدمها المفردة أو المقياس بدلالة معالم المفردة (حرز الله، ٢٠٠٤)؛ فدالة معلومات المفردة هي درجة مساهمة المفردة في تقدير القدرة، وتمييز المفردة العالي يساهم في تأكيد دقة القياس بشكل أكبر من المفردات ذات التمييز المنخفض، وأن أكبر قيمة لدالة معلومات المفردة تكون عند قيمة على متصل القدرة تساوي قيمة صعوبة المفردة (النجار، ٢٠١٠)؛ بينما دالة معلومات الاختبار، فهي دالة اقتران رياضي تعبر عن مجموع دوال المعلومات لمفردات الاختبار (الخرشة، ٢٠١٥)؛ ودالة المعلومات تصف دقة تقدير القدرة عند مستويات القدرة المختلفة (baker, 2001).

٤- معالم المفردات Items Parameters:

معالم المفردات هي معالم الصعوبة والتمييز والتخمين ويختلف عدد المعالم أو البارامترات المطلوبة لوصف المنحنى المميز للمفردات باختلاف النموذج المستخدم، فمنها ما يشمل على معلم واحد مثل الصعوبة وهذا يحققه نموذج راش، ومنها ما يعتمد على معلمين مثل معلم الصعوبة ومعلم التمييز، ومنها ما يعتمد على

ISSN: 2537-0464 eISSN: 2537-0472

ثلاث معالم (الصعوبة، التمييز، التخمين)؛ وذلك في حالة الاختبارات ذات الأسئلة ثنائية الإجابة (صفر، ١) أما في حالة مقاييس التقدير الذاتي فإن الأمر يختلف حيث توجد نماذج تركز فقط على صعوبات المفردات أو كما تسمى عتبات الاستجابة، وهناك نماذج تركز على العتبات والتمييز لكل مفردة (النعيمي، ٢٠٠٦).

وفي الدراسة الحالية سيتم التركيز على معلم الصعوبة (العتبات) والتمييز والأخطاء المعيارية في تقديرها وفقاً لنموذج الاستجابات المتدرجة Graded Response Model.

٥- معالم الأفراد Persons Parameters:

وهي قيمة تدريجية تنتج من تطبيق الدالة الرياضية لنموذج استجابة الفقرة عند استجابة المفحوص على مجموعة من الفقرات المشكلة للمقياس، والتي تقيس مدى امتلاكه للسمة المراد قياسها، حيث يكون لكل فرد معلم يسمى معلم الفرد أو القدرة (مرشود، ٢٠٢٢).

وفي الدراسة الحالية سيتم التركيز على معلم القدرة والأخطاء المعيارية في تقديره وفقاً لنموذج الاستجابات المتدرجة Graded Response Model.

الإطار النظري

المحور الأول: طرق الاستجابة، وعدد فئات الاستجابة

أولًا: طرق الاستجابة

يشير شكل مقياس الاستجابة إلى طريقة صياغة المفردات والحصول على الإجابات وتقييمها (Furr, 2011)، و تتضمن أشكال المقياس الشائعة (Dimitrov, 2012; Barker et al., 2016): (أ) مقياس جوتمان Guttman Scaling Guttman)؛ (ب) مقياس ثورستون (Thurstone)؛ (ج) مقياس ليكرت (Likert). (DeVellis,2017).

كما ويعد نموذج القياس الكلاسيكي أكثر ملاءمة للمقاييس التي تكون فيها المفردات أداة تقييم مكافئة تقريبًا للبنية المقاسة، مثل ليكرت (Price,2016)، و بشكل عام تكون المقاييس المكونة من المفردات التي تم وضع درجات لها على سلسلة متصلة ثم يتم جمعها للحصول على درجة المقياس أكثر توافقًا مع نموذج القياس الكلاسيكي (لقياس المتغير الكامن) الذي يفترض أن المفردات هي مؤشرات قابلة للمقارنة للبنية الأساسية مقارنة بإجابة المفردة تعتبر النظرية التي تمثل منظور قياس بديل (DeVellis, 2017; Price, 2016) والحالتين (أ) و(ب) أكثر ملاءمة (Price, 2016).

ISSN: 2537-0464 eISSN: 2537-0472

لهذا السبب، سوف يتم وصف بإيجاز مقياس جوتمان وثورستون وبمزيد من التفصيل مقياس ليكرت أو عمومًا جميع المقاييس المستمرة والمرجحة بشكل متساوٍ للتقدير المباشر (DeVellis, 2017).

١. مقياس جوتمان Guttman Scaling

الهدف من مقياس جوتمان هو تطوير أدوات قياس أحادية البعد ويؤدي مقياس جوتمان إلى ترتيب إجمالي للمفردات، حيث إن الشخص الذي يجيب/يوافق بنجاح على مفردة ما يجيب/يوافق أيضًا على المفردات ذات الترتيب الأدنى (He,2016).

ومقياس جوتمان الذي يتكون من مفردات تشتمل على مستويات أعلى من إحدى السمات وتسمى أيضًا القياس الحتمي، أو القياس التراكمي (Dimitrov, 2012)؛ ويجب على المستجيب اختيار مجموعة من المفردات حتى يتجاوز مقدار السمة المقاسة التي يمتلكها المستجيب؛ عند هذه النقطة، لا ينبغي تحديد أي مفردة أخرى، وتعمل البيانات الوصفية بشكل جيد مع مقياس جوتمان، على سبيل المثال: هل تشرب؟ - "هل تشرب أكثر من كوبين في اليوم؟" إلخ، ويظهر مستوى سمة المستجيب من خلال أعلى استجابة إيجابية، ويتميز مقياس جوتمان بإمكانية تطبيق محدودة إلى حد ما مع وجود عيوب غالبًا ما تفوق المزايا لأن افتراض وجود علاقات سببية قوية متساوية بين المتغير الكامن وكل مفردة من المفردات لا ينطبق على مفردات مقياس جوتمان (Streiner et al., 2015)؛ ومن الناحية العملية، تعد أنماط الاستجابة التي تصف مقياس جوتمان المثالي نادرة (Price, 2016)؛ ويوضح جدول (١) طريقة جوتمان في إعداد فقرات المقياس.

جدول (١) طريقة جوتمان في إعداد فقرات المقياس

المقياس	العبارة	صح	خطأ
مقياس جوتمان	أنا قادر على (اختر واحدة): ١. الجري مسافة ٢٠٠ متر ٢. الجري مسافة ٤٠٠ متر ٣. الجري مسافة ٦٠٠ متر ٤. الجري مسافة كيلومتر واحد		

(Price, 2016)

٣- مقياس ثيرستون Thurstone Scaling

اقترح ثيرستون Thurstone (١٩٢٧) ثلاث طرق لتطوير مقياس أحادي البعد: وهي طريقة الفترات المتساوية في الظهور equal-appearing intervals، وطريقة الفترات المتعاقبة successive intervals، وطريقة المقارنات المزدوجة paired comparisons (Dimitrov, 2012)؛ والفكرة الأساسية في جميع الطرق الثلاثة هي أن مطور المقياس يبتكر مفردات تتوافق مع مستويات مختلفة من السمة

ISSN: 2537-0464 eISSN: 2537-0472

المقاسة (DeVellis, 2017)، ثم يقوم مجموعة من الخبراء بتقييم الدرجة التي تمثل بها المفردات السمة على مقياس من ١ (الأقل تمثيلاً) إلى ١١ = الأكثر تمثيلاً (Dimitrov, 2012). ومع ذلك، كقاعدة عامة، فإن المشكلات العملية في استخدام طريقة ثيرستون مع نموذج القياس الكلاسيكي، وعملية تطويره الصعبة بالإضافة إلى النتائج المماثلة لمقياس ليكرت غالبًا ما تقلل من مزاياه (DeVellis, 2017) ، ويوضح جدول (٢) طريقة ثيرستون في إعداد فقرات المقياس.

جدول (٢) طريقة ثيرستون في إعداد فقرات المقياس

المقياس	العبارة	اوافق	لا اوافق
مقياس ثورستون	١. النجاح بالنسبة لي شرط أساسي للسعادة ٢. الحصول على وظيفة جيدة أمر مهم ولكنه ليس ضروريا ٣. السعادة لا علاقة لها بالماديات أو العمل ٤. تحقيق النجاح يعيق الشعور بالسعادة		

(DeVellis, 2017)

٣. مقياس ليكرت Likert Scaling

ربما يكون مقياس ليكرت - أو مقياس ليكرت المعياري Likert normative scale - الذي طوره ليكرت (١٩٣٢, ١٩٥٢) - هو شكل أو تنسيق الاستجابة الأكثر شيوعًا في علم النفس (Furr, 2011; Dimitrov, 2012; Barker et al., 2016)، وهو متعدد الاستخدامات وفعال للتمييز بين مستويات القدرة أو الإنجاز (Price, 2016)، ويحتوي على جزأين: (١) المفردة و (٢) مقياس استجابة يحتوي على مجموعة من البدائل ذات الكثافة المتزايدة المشار إليها بقيمة عددية صحيحة ومواصفات لفظية تسمى المراسي (Barker et al., 2016). وبالنسبة لمقاييس ليكرت، يقرأ المشاركون نص المفردة، والذي عادة ما يكون سؤالًا أو عبارة (على سبيل المثال، "أحب الذهاب إلى الحفلات")،ثم يختارون أحد خيارات الاستجابة المتاحة (على سبيل المثال، "أوافق بشدة"). وعادة، يرتبط كل خيار استجابة متاح بقيمة كمية محددة (على سبيل المثال، ١ = لا أوافق بشدة؛ ٥ = أوافق بشدة)، والتي يتم جمعها أو حساب متوسطها عبر جميع إجابات الشخص ذات الصلة ببُعد معين (Furr, 2011). ويوضح جدول (٣) أمثلة لتدريج ليكرت.

ISSN: 2537-0464 eISSN: 2537-0472

جدول (٣) أمثلة لتدريج ليكرت (الخماسي والسباعي)

لدي الكثير في الحياة لأكون شاكراً له	إيجابى	أنا أبحث عن معنى في حياتي
١ = غير موافق بشدة ٢ = غير موافق ٣ = غير موافق قليلاً ٤ = محايد ٥ = موافق قليلاً ٦ = موافق ٧= أوافق بشدة	١= نادراً جداً أو أبداً ٢ = نادرا ٣ = أحياناً ٤ = غالبا ٥ = في كثير من الأحيان أو دائما	١ = غير صحيح على الاطلاق ٢ = غير صحيح في الغالب ٣ = غير صحيح إلى حد ما ٤ = لا أستطيع أن أقول صح أو خطأ ٥ = صحيح إلى حد ما ٦ = صحيح في الغالب ٧ = صحيح تماما
نموذج استبيان الامتنان المكون من ست مفردات (GQ-6) (McCullough, Emmons, & Tsang, 2002)	مقياس الخبرة الإيجابية والسلبية (SPANE) (Diener et al.، 2010)	استبيان معنى الحياة (MLQ) Steger et al. (2006)

(Furr, 2011)

يتم وضع التقييمات الموضحة في الجدول (٣) على سلسلة متصلة ثنائية القطب من نقاط متساوية تتراوح من الموافقة بشدة على العبارة أو المفردة إلى عدم الموافقة بشدة.

وتوفر نقطة المنتصف (المحايد) على المقياس فئات الاستجابة " في منتصف التدريج " (Price, 2016). ويمكن للمفردة الفعالة وفقا لتدريج ليكرت تقييم الآراء والاتجاهات والمعتقدات بعبارات واضحة ولكنه أكثر توافقًا مع العبارات المكتوبة بلغة صعبة لأن المفردات المعتدلة تثير اتفاقًا عامًا (DeVellis, 2017) كم وأن هناك تباين في مقياس ليكرت وهو مقياس تصنيف السلوك على سبيل المثال، يتم تصنيف سلوك الطالب في الفصل الدراسي باستخدام مفردة مثل "الطالب يسيء التصرف في الفصل" على أنه دائمًا = ٥ أبدًا = ١ (Saville & MacIver, 2017; Price, 2016).

لقد تم وضع الافتراضات التالية لمقاييس ليكرت:

١) فواصل الفئة متساوية الطول تقريبًا،.

٢) يتم تعيين تسميات الفئة بشكل موضوعي.

٣) إن مرحلة الاختبار القبلي أثناء تطوير المفردات يتبعها تحليل المفردات للاستجابات (Price, 2016)، مثلما يمكن أن يؤثر شكل السؤال على الاستجابة، كذلك يمكن أن يؤثر شكل مقياس الاستجابة (Barker et al., 2016).

ISSN: 2537-0464 eISSN: 2537-0472

من ناحية أخرى، تعتمد أساليب نظرية الاستجابة للمفردة Item Response Theory (IRT) لتحليل فقرات ليكرت نهجًا قائمًا على المفردات لتقييم المقياس. كما وأن نماذج نظرية الاستجابة للمفردة هي مجموعة من النماذج الرياضية تُستخدم لتحديد العلاقة بين مستويات سمة كامنة لدى الأفراد وخصائص المفردات، والتي تُسمى معلمات المفردات Item parameters. (Baghaei & Effatpanah, 2024).

- عدد فئات الاستجابة في مقاييس ليكرت

تتمثل حجج ليكرت حول مقاييس ليكرت في:

أ) إمكانية الحفاظ على مسافات متساوية بين الفئات.

ب) إمكانية ترتيب تسمية الفئات مسبقًا حتى لو كانت ذاتية.

ج) إمكانية تغيير أحكام المقياس المُعدّ وفقًا لتحليلات المفردات التي سيتم إجراؤها بناءً على استجابات المشاركين في المقياس (Dunn-Rankin et al., 2004).

باختصار تُقيّم حجج ليكرت بناءً على توزيع المتغيرات الحقيقية (Price, 2016)، في مقاييس ليكرت، تُرتّب فئات الاستجابة بحيث تكون مسافات تقييمها متساوية قدر الإمكان (DeVellis, 2017).

قد يكون عدد فئات الاستجابة خمسة على سبيل المثال في شكل "لا أوافق بشدة"، و "لا أوافق"، و"متردد"، و"أوافق"، و"أوافق بشدة" (Aybek et al., 2022)؛ أو قد يكون عددها ٦ فئات في شكل "لا أوافق بشدة"، و "لا أوافق"، و "لا أوافق إلى حد ما"، و "أوافق إلى حد ما"، و "أوافق"، و "أوافق بشدة" (DeVellis,2017)؛ وقد تكون هناك أيضًا فئة استجابة محايدة بين فئات الاستجابة أو لا، وهناك مقترحات لاستخدام كثيرة حول شكل أو تحديد النقطة المحايدة مثل "لا أوافق ولا أختلف" أو "أوافق وأختلف بالتساوي" لكن المناقشات حول كيفية التعبير عن هذه النقطة المحايدة لا تزال مستمرة (DeVellis,2017).

في مقاييس ليكرت، تُولَّد درجات فئوية ترتيبية من إجابات المشاركين على مفردات المقياس، تتوافق هذه الدرجات مع نطاق ثنائي القطب من "أختلف بشدة" إلى "أوافق بشدة" (Price, 2016). يجادل بعض الباحثين بأن البيانات التي تم الحصول عليها من مقياس ليكرت تكون على مستوى مقياس ترتيبي، وأنه يجب استخدام أساليب إحصائية مناسبة لمثل هذه البيانات (Jamieson, 2004). على الرغم من افتراض فترات متساوية بشكل عام لمقاييس ليكرت في الممارسة العملية (أي المسافات بين الأرقام في المقياس الترتيبي)، إلا أنه غالبًا ما لا يمكن إثبات مثل هذا الافتراض من منظور أساسيات القياس. في مواجهة هذه الاشكالية، غالبًا ما يُطرح السؤال "هل يجب معالجة البيانات على مقياس ترتيبي أم مقياس فترة متساوية؟".

ISSN: 2537-0464 eISSN: 2537-0472

أشار نورمانNorman (٢٠١٠) إلى أنه يمكن قبول مقاييس ليكرت على مستوى مقياس فترة متساوية، ويمكن استخدام التحليلات البارامترية بناءً على هذا الافتراض. في دراستهم القائمة على المحاكاة، زعم وو وليونجWu & Leung (٢٠١٧) أن زيادة عدد التقديرات في فئة الاستجابة لمقياس ليكرت من شأنه أن يؤدي إلى توزيع طبيعي وتشابه مع مقياس الفاصل الزمني.

ولقد كان الجدل حول عدد فئات الاستجابة المناسب متجذرًا بعمق في الأدبيات منذ ما يقرب من قرن من الزمان، فعلى سبيل المثال، رأى فرايد أن العدد المناسب من نقاط المقياس كان دالة لهدف المقياس ومستوى التمييز الذي يتطلبه مستخدمو الدرجات (أي المسافة بين فئات الاستجابة)، وبشكل عام اتخذ الباحثون مواقف مختلفة وقدموا توصيات مختلفة نسبيًا لعدد فئات الاستجابة الأمثل، حيث قد أوصى عدد من الباحثين بأربعة إلى سبعة فئات استجابة (Barnette, 2010; Bindak, 2013; de Winter & Dodou, 2010; Sullivan & Artino, 2013) .

وأوصى آخرون بخمسة إلى سبعة فئات للاستجابة (Abulela& Khalaf, 2024). ووفقاً لديلمان وآخرون (Dillman et al.,2014)، يجب أن يكون عدد فئات الاستجابة طويلًا بما يكفي لتغطية كامل سلسلة الاستجابات المحتملة، ولكن لا ينبغي أن يكون الفرق بين نقطتي مقياس متجاورتين صغيرًا جدًا أو لا معنى له من الناحية العملية، وفي مراجعة لتسع دراسات منشورة في المجلة الدولية لعلم النفس العيادي والصحي، وجد هارتليHartley (2014) أن أدوات القياس رباعية أو خماسية فئات الاستجابة تُستخدم بشكل شائع.

ولا يوجد إجماع بشأن العدد الأمثل لفئات الاستجابة لمقاييس ليكرت، وتشير بعض الدراسات إلى أن زيادة عدد فئات الاستجابة يمكن أن يحسّن من قوة النماذج الإحصائية المستخدمة في تحليل فقرات المقياس والكشف عن جودتها كالأداء التفاضلي للمفردات (Allahyari et al., 2016)، وتؤكد دراسات أخرى أن عدد فئات الاستجابة لا يؤثر بشكل كبير على دقة المقاييس أو موثوقيتها أو قوتها التمييزية (Chakrabartty, 2023)؛ ومع ذلك، تشير الأبحاث إلى أن خمسة فئات استجابة قد تكون مثالية في الوصول لمعلومات دقيقة حول السمة موضع الاهتمام Chen et al., (2015)؛ بالإضافة إلى ذلك، توصي الدراسات بالاحتفاظ بجميع فئات الاستجابة الممكنة ما لم يكن هناك سبب قوي لتقليصها (Ben & Nissen, 2020)، وبشكل عام، قد يعتمد العدد الأمثل لفئات الاستجابة على سياق البحث المحدد وأهداف الدراسة.

ISSN: 2537-0464 eISSN: 2537-0472

ثانيًا: ترتيب فئات الاستجابة:

نُشرت مؤخرًا مراجعة لأهم تطورات مقياس ليكرت على مدار ربع القرن الماضي (Jebb et al., 2021) ، وحدد فيها الباحثون ٤٠ دراسة تناولت المجالات التالية المتعلقة بتطوير مقياس ليكرت: (أ) مفاهيم الصدق البنائى (ب) تعريف البناء؛ (ج) بناء مفردات المقياس (د) التحقق من صدق المحتوى (هـ) إجراء دراسات تجريبية؛ (و) دقة القياس (ز) تقييم البنية العاملية؛ (ح) إنشاء نماذج مختصرة (ط) العلاقات التجريبية مع المتغيرات الأخرى. وتحت إنشاء مفردات المقياس، كانت الفئات الفرعية هي: (١) اختبارات يمكن قراءتها (٢) مقاييس حديثة يمكن قراءتها (٣) فهم المستجيب (٤) عدد خيارات الإجابة وتسمياتها (٥) شكل المفردة (٦) ثبات المفردة (٧) عرض مفردات المقياس.

قام جيب وآخرونJebb et al (٢٠٢١) بمراجعة دراسة واحدة فقط بحثت فى تنسيق المفردات (Zhang & Savalei, 2016)، على الرغم من وجود العديد من الدراسات الأخرى حول تنسيق المفردات التي لم يتم تضمينها في مراجعتهم كما واستخدم تشانغ وسافاليZhang & Savalei (٢٠١٦) تنسيق مقياس بديل استبدل كل فئة استجابة في مقياس من نوع ليكرت بجملة كاملة ووجدوا أن مثل هذه المقاييس لها بنية عاملية أفضل من المقاييس التقليدية، فحصت العديد من الدراسات تأثير تغيير تسميات فئات الاستجابة (على سبيل المثال، استخدام التسميات اللفظية فقط لنقاط نهاية المقياس أو إزالة التسميات اللفظية تمامًا Gummer & Kunz, 2021; Spratto et al., 2021).

بالإضافة إلى عمل ويجولد وآخرونWeigold et al (٢٠٢١)، درست العديد من الدراسات تأثير التوجه الأفقي مقابل الرأسي لفئات الاستجابة، وكانت نتائج هذه الدراسات غير حاسمة (Hu, 2020). كما بُحثت قضايا إضافية تتعلق بتنسيق المفردات في دراسات أخرى، مثل تأثير القطبية والتعبير اللفظي للفئة الوسطى (Menold, 2021)، وكيفية وضع وشرح فئات الاستجابة "لا أعرف" في الاستبيان (Zeglovits & Schwarzer, 2016).

قد يختار المستجيبون للمقياس فئات مختلفة عندما يتم ترتيب فئات الاستجابة بترتيب تصاعدي أو تنازلي/ إيجابي أو سلبي وقد يكون للترتيب الذي يتم به تقديم بدائل الاستجابة للمستجيبين تأثير كبير على اختياراتهم (Chyung et al., 2018)وتُعرف هذه الظاهرة بتأثيرات ترتيب الاستجابة.

وهناك نوعان شائعان من تأثيرات ترتيب الاستجابة هما تأثيرات الأسبقية والحداثة Primacy and Recency ويشير تأثير الأسبقية إلى ميل المستجيبين إلى اختيار فئات الاستجابة المقدمة في بداية قائمة فئات الاستجابة، بينما تأثير الحداثة

ISSN: 2537-0464 eISSN: 2537-0472

يمثل العكس من ذلك بمعنى الميل إلى أن يختار المستجيبون فئات الاستجابة التي يرونها في نهاية قائمة فئات الاستجابة (Holbrook, 2008).

وباستخدام بيانات من استطلاعات رأي (التقرير الذاتى) حول شرب الخمر، أثبت ماكينون وفيرث Mackinnon and Firth (٢٠١٨) تأثير الأسبقية ، حيث أظهرا أن المشاركين كانوا أكثر ميلًا لاختيار فئة استجابة "أوافق بشدة" عند ظهوره أولًا في قائمة الإجابات مقارنةً بظهوره آخرًا. ووجد مالهوتراMalhotra (٢٠٠٨) أن حجم تأثير الأسبقية يرتبط ارتباطًا إيجابيًا بدرجة تباين المستويات التعليمية للمشاركين. كما أشار إلى أن الوقت المُستثمر في الاستطلاع يتفاعل مع المستوى التعليمي في تأثير الأسبقية ، مما يعني أن المشاركين ذوي التعليم المنخفض الذين يقضون وقتًا أقل في الدراسة أكثر عرضة لهذا التحيز.

يُعد تأثير ترتيب المفردات ذا أهمية خاصة لقياس الاتجاهات، ووفقًا لتشن Chen (٢٠١٠)، استندت المحاولة الأولى لوصف هذه الظاهرة إلى الحداثة والاسبقية، إلا أن الأدبيات العلمية تحولت الآن إلى التكييف والتثبيت، يفترض التكييف والتثبيت أن الأفراد يفضلون الوصول إلى المعلومات المقدمة إليهم أولًا والحصول على تقديراتهم المعقولة من خلال التعديل بناءً على هذا التثبيت (Zhao &Linderholm, 2008). فيما يتعلق بتأثير ترتيب المفردات، فإن الردود الأولى على الخيارات تُعتبر مثبتات للإجابات اللاحقة Harrison &, McLaughlin, (1993) بمعنى آخر، يحدث التكييف والتثبيت عندما تكون الذاكرة السياقية للشخص غير كافية، مما يؤدي إلى استجابة سابقة على الأشياء تعمل كمثبتات تُغير إجابات الأمور المستقبلية اعتمادًا على هذه المثبتات(Chen, 2010).

ويحدث تأثير ترتيب فئات الاستجابة عندما يؤثر ترتيب فئات الاستجابة لمقياس التقدير على توزيع سؤال الاستبيان. وقد أوضحت التفسيرات النظرية مثل هذه التأثيرات، بما في ذلك التثبيت والتكييفAnchor-and-adjustment (Yan & Keusch,2015) ، وتحيز الذاكرة والإرضاء وقد توفر بعض الاستدلالات التفسيرية البصرية (مثل "الأعلى يعني الجيد" و"اليسار والأعلى يعني أولاً") أيضًا نظرة ثاقبة حول كيفية تأثير وضع احتمالات الاستجابة على فئات الاستجابة (Tourangeau, 2004). أجريت معظم الأبحاث السابقة حول تأثير ترتيب فئات الاستجابة في سياقات أحادية الثقافة (Harkness & Schoua-Glusberg, 1998). ومع ذلك، قد تؤثر المؤثرات "الثقافية" على وجود وحجم تأثير ترتيب فئات الاستجابة بعدة طرق(Diener et al., 2012) . أولها أن "اليسار يعني أولاً" والاستدلالات التفسيرية الأخرى قد تعمل بشكل مختلف اعتمادًا على معايير قراءة النص (على سبيل المثال، من اليمين إلى اليسار مقابل من اليسار إلى اليمين).

ISSN: 2537-0464 eISSN: 2537-0472

بالإضافة إلى ذلك، قد يمتلك الأشخاص من ثقافات ذات لغات رئيسية متعددة ومعايير قراءة متعددة أساليبَ تحديدٍ مختلفة. وأخيرًا، قد يمتلك الأشخاص في بلدان مختلفة مستويات خبرة مختلفة مع أسلوب تصميم مرئي معين(Yang et al., 2019).

يبدأ بحث إيبلي وجيلوفيتش Epley and Gilovich (٢٠٠٦) حول التثبيت والتكييف في تسليط الضوء على كيفية إسهام هذا الاختصار الذهني في تحقيق الرضا في الاستبيانات، ويفترضان أنه في الحالات التي يحدث فيها التثبيت والتكييف ، تكون التكيفات غير كافية لأن الأفراد يتوقفون عن التكييف بمجرد وصولهم إلى قيمة تقع ضمن نطاق معقول أو تبدو "قريبة بما فيه الكفاية". وبالتالي، قد يحدث التثبيت والتكييف في الاستبيانات على النحو التالي: يُقيّم المستجيب مفردة أوليًا في الاستبيان (ربما باختيار فئة الاستجابة الرابع على مقياس من خمس نقاط)، إذا كانت المفردة التالية مشابهة، فقد يبدأ المستجيب في تقييم فئة الاستجابة بتحديد ما إذا كانت الاستجابة الرابعة على المقياس معقولة، وإذا لم تكن صحيحة تمامًا، فقد ينتقل المستجيب إلى فئة الاستجابة الثالث أو الخامس. بمجرد الوصول إلى استجابة "جيدة بما فيه الكفاية"، من المرجح أن يتوقف المستجيب عن التكييف. إذا كانت التكيفات الناتجة عن هذه التثبيت غير كافية (Gehlbach et al., 2012)، فمن المرجح بشكل غير متناسب أن يجيب المستجيب على مفردة الاستبيان الثانية عند فئة الاستجابة الرابع أو بالقرب منه على المقياس. وبالتالي، على مدار الاستبيان بأكمله، من المرجح أن تكون الاستجابات لأزواج المفردات المتجاورة أكثر تشابهًا (وبالتالي تكون المسافات بينها أقصر) من الاستجابات لنفس أزواج المفردات إذا كانت في مواقع غير متجاورة . بالإضافة إلى ذلك، نظرًا لأن المستجيبين يحتاجون فقط إلى تقييم خيارات الاستجابة حتى الوصول إلى إجابة "جيدة بما فيه الكفاية"، فقد يُكمل المستجيبون الذين يقومون بالتثبيت وتكييف استبياناتهم بشكل أسرع من أولئك الذين يُقيّمون جميع فئة الاستجابة.(Gehlbach et al., 2012)

وقد قدمت عدد من الدراسات السابقة التي أجريت في استطلاعات الرأي عبر الإنترنت لتقييم الأسئلة باستخدام أنواع مختلفة من مقاييس الاستجابة الترتيبية (أحادية القطب وثنائية القطب، مع عدد مختلف من فئات الاستجابة) الدليل التجريبي على تأثيرات الأسبقية مما يعني الاحتمالية الأعلى لاختيار فئات الاستجابة الأولى من قائمة فئات الاستجابة (Krebs, 2012; Liu & Keusch, 2017; Mahon-Haft & Dillman, 2007; Toepoel et al., 2009).

قام هون لنزنر Höhne & Lenzner (٢٠١٥) بدراسة تأثيرات تسلسل الاستجابة في مقاييس التقدير الرأسية والأفقية باستخدام تتبع العين لتحديد أسبابها

ISSN: 2537-0464 eISSN: 2537-0472

وكشف وجودها. وخلصت هذه الدراسة إلى أن تأثيرات ترتيب الاستجابة في مقاييس التقدير غالبًا ما تكون ضئيلة وأكثر شيوعًا في مقاييس التقدير الرأسية منها في الأفقية.

يهدف البحث الذي أجراه هون وكريبسHöhne & Krebs (٢٠١٧) إلى معرفة المزيد حول كيفية تأثير تسلسل تقدير المفردات على النتائج. تكشف النتائج عن تأثيرات ترتيب الاستجابة ضمن صيغة "موافق/غير موافق" ولكن ليس ضمن صيغة "المفردة التى تحمل سؤالا".

واقترحت نظرية الإرضاء Satisficing Theory التفسير الأكثر انتشارًا لتأثير ترتيب فئات الاستجابة، وغالبا ما يستخدم هذا النهج النظري كتفسير محتمل لتأثيرات اتجاه عبارات المقياس (Liu & Keusch, 2017)، ووفقاً للنظرية، فبدلاً من اختيار الاستجابة المثالية للمفردة، فإن الفرد يختار الاستجابة المرضية، من أجل تقليل الجهد المعرفي (Holbrook et al., 2003)، ولذلك في النمط البصري، عندما تكون قائمة فئات الاستجابة طويلة، فإن إجابات المستجيبين التي انتقلت إلى بداية المقياس يمكن أن تكون ناجمة عن سلوك مُرضي ولا تدل على بذل الجهد في الاستجابة، وبالتالي، تتنبأ النظرية بتأثير الأسبقية كنتيجة للسلوك المرضي للمستجيب (Höhne & Lenzner, 2015; Kunz, 2015).

وتستند هذه النظرية إلى المراحل الأربع للمعالجة المعرفية للأسئلة والأجوبة. وبشكل عام، يقترح العلماء المراحل المتوازية التالية للاستجابة: (١) فهم السؤال وتفسيره (٢) استخلاص المعلومات ذات الصلة من الذاكرة (٣) تذكر أو تحديد الحكم (٤) تنفيذ وتحرير الاستجابة (Kunz, 2015)، وكل مرحلة من هذه المراحل ضرورية لجودة البيانات، وتتميز استراتيجية الاستجابة المرضية الضعيفة بالسلوك عندما يمر المستجيبون بخطوات معرفية أقل شمولاً مما يمكن ملاحظته في الموافقة وتأثير الأسبقية، ويظهر الرضا القوي عندما يتخطى المستجيبون الخطوتين الثانية و/أو الثالثة وينتج عن ذلك اختيار "لا أعرف" وعدم التمييز بين مفردات مقياس التقدير (Krosnick & Presser, 2010)، ويمكن تعزيز سلوك الرضا من خلال الشروط التالية: القدرات المعرفية ودوافع المستجيبين، وصعوبة الأسئلة، وقد تم تأكيد هذه الافتراضات النظرية في بعض الأبحاث التجريبية (Malhotra, 2008; 2009).

وأثبتت نتائج دراسة مالهوترا Malhotra (2008; 2009) تأثيرات الاسبقية على الاستجابة للمفردات البسيطة وعدم وجود تأثيرات للاسبقية في الاستجابة للمفردات المعقدة، كما أكدت نتائجها تأثيرات أكبر للأسبقية على استجابة المفحوصين الأقل في مستوى التعليم؛ وكذلك أكدت نتائج دراسة هون وكريبس Höhne and Krebs (2018) تأثيرات ترتيب الاستجابة في المفردات ذات مقاييس الموافقة - عدم

ISSN: 2537-0464 eISSN: 2537-0472

الموافقة؛ وأكدت نفس النتائج دراسة كريبس وباخنر (Krebs & Bachner, 2018).

كما وبحثت الدراسة التي أجراها تيرينتيف ومالوشونوك Terentev and Maloshonok(٢٠١٨) في تأثير ترتيب الإجابات على أسئلة التقدير المقدمة في تنسيقات كل مفردة على حدة وفي شكل شبكي، تم التنبؤ بأن تأثير الاسبقية صحيح لكلا النوعين من الاستفسارات وأن يختلف باختلاف عمر المستجيب ومستوى التعليم والتطور التكنولوجي، أُجريت تجربتان مستقلتان باستخدام بيانات من ٢٨ استبيانًا تمهيديًا خضع لها طلاب مسجلون في دورات التعلم عن بعد (عددهم ٢٢,٩١٠). تشير نتائجهما إلى أن انطباعات المشاركين عن قائمة الخيارات وأنماط استجاباتهم تتأثر بترتيب عرض بدائل الاستجابة. ويتضح تأثير الاسبقية عند إجراء استعلام لكل مفردة على حدة. ولم يُعثر على أي تأثير يُذكر للعمر أو الجنس أو نوع الجهاز على الترتيب، أما المشاركون ذوو المستوى التعليمي الأعلى، فقد واجهوا تأثير أسبقية أقل عند ترتيب المفردات.

وقام ستابلتون Stapleton (2013) بتحليل تأثيرات التفاعل بين ترتيب فئات الاستجابة ونوع الجهاز (الويب أو الهاتف المحمول) المستخدم لإكمال الاستبيان، وأظهر نتائجها تأثيرات أكبر للأولوية وبشكل ملحوظ بين المستجيبين الذين أكملوا الاستبيانات باستخدام الأجهزة المحمولة، ويمكن تفسير ذلك بحقيقة أن حجم شاشة الهاتف المحمول أقل من حجم سطح المكتب التقليدي أو الكمبيوتر المحمول، في حين أن المستجيبين أكثر عرضة لاختيار فئات الاستجابة المرئية، فإن استخدام الأجهزة المحمولة يؤدي إلى تأثيرات أكبر للأولوية.

أظهر تحليل الدراسات التي بحثت فى الارتباط بين عدد فئات الاستجابة في المقياس وصدق المقياس أن بعض الدراسات اختبرت تأثير عدد فئات الاستجابة على الصدقى البنائى ، وأن البعض الآخر سعى إلى إجابات على سؤال ما إذا كان الثبات القائم على المحك يختلف وفقًا لعدد فئات الاستجابة المستخدمة أم لا، جاءت نتائج متباينة فيما يتعلق بتأثير عدد فئات الاستجابة على الصدق في الدراسات التي أجريت. على سبيل المثال، وجد كومري ومونتانج Comrey and Montang (١٩٨٢) وكينج وآخرونKing et al (١٩٨٣) أن معدل التباين الكلي المفسر وتشبعات العوامل كان أعلى وكان البناء العاملى أكثر وضوحًا في التصنيف ذي النقاط السبع منه في التصنيف ذي النقطتين.

ووجد لوزانو وآخرونLozano et al (٢٠٠٨) أيضًا أن معدل التباين المفسر في التحليل العاملى ارتفع مع زيادة عدد فئات الاستجابة في المقياس، وتم تفسير النتيجة على أن الزيادة في عدد فئات الاستجابة أثرت على صدق تأثيرا ايجابيا. على

ISSN: 2537-0464 eISSN: 2537-0472

نحو مماثل، خلص تاركانTarkan (٢٠١٥) إلى أن الصدق العاملى يزداد بازدياد عدد فئات الاستجابة في المقياس. وعلى النقيض من هذه الدراسات، وجد كيم Kim (١٩٩٨) في الدراسة التي قارنت بين التصنيف الثلاثي والخماسي والسباعي والتاسع من حيث الصدق والثبات أن الصدق كان الأقل في التصنيف الثلاثي، ومتوسط في التصنيف السباعي، وأعلى في التصنيفين الخماسي والتاسع. ووجدت الدراسة التي أجراها مايدو-أوليفاريسMaydeu-Olivares (٢٠٠٩) أن توافق بيانات النموذج انخفض بازدياد عدد فئات الاستجابة في المقياس.

لا يوجد تداخل بين النتائج المتعلقة بالتأثير على صدق المحك كما هو الحال في النتائج المتعلقة بتأثيرات عدد فئات الاستجابة على الصدقى البنائى، في الدراسة التي أجراها تشانغChang (١٩٩٤) حيث تمت مقارنة ليكرت الرباعي والسداسى سيكومتريًا، وجد أن عدد فئات الاستجابة لم يكن له أي تأثير على صدق المحك. وبالمثل، خلص قاسم واخرونQasem et al (٢٠١٤) أيضًا إلى عدم وجود فروق ذات دلالة إحصائية بين معاملات صدق المحك للمقاييس ذات التصنيف الثنائي والثلاثي والخماسي.

من ناحية أخرى، يشير لوكين واخرون Loken (١٩٨٧) إلى أن صدق المحك يتأثر بعدد فئات الاستجابة وأن معاملات صدق المحك التي تم الحصول عليها من التصنيف ذي الإحدى عشرة نقطة أعلى من تلك المحسوبة من التصنيف ذي الثلاث أو الأربع نقاط. ومع ذلك، وجد بريستون وكولمان Preston and Colman (٢٠٠٠) أن معاملات صدق المحك للمقاييس التي تستخدم التقييم الثنائي والثلاثي والرباعي كانت أقل، وأن معاملات صدق المحك للمقاييس التي تحتوي على خمس فئات أو أكثر كانت أعلى. كما وُجد أيضًا عدم وجود فروق ذات دلالة إحصائية بين معاملات صدق المحك للمقاييس التي تحتوي على أعداد مختلفة من فئات الاستجابة.

ومع ذلك، فإن نتائج عدد من الدراسات التجريبية التي أجريت لتقييم الاستجابات باستخدام مقاييس الاستجابة الترتيبية أحادية القطب مع خمسة خيارات للاستجابة (Christian et al., 2009) وثنائية القطب وثمانية فئات للاستجابة (Hofmans et al., 2007) لم تظهر تأثير كبير لترتيب فئات الاستجابة على توزيع استجابات المستجيبين.

ومن ثم يمكن استنتاج أنه عندما يُعرض على الفرد عدة فئات للاستجابة، فإن الموقع المكاني لكل فئة على حدة فيما يتصل بترتيب المعالجة يمكن أن يكون له تأثير على الاختيار النهائي، كما تبين وجود أدلة على وجود تأثير للأسبقية، أي الميل إلى تفضيل المفهوم أو الشيء الأول الذي يواجهه المستجيب.

ISSN: 2537-0464 eISSN: 2537-0472

المحور الثاني: نظرية الاستجابة للمفردة الاختبارية

بناء وتقنين أدوات القياس وانتقاء المفردات في ضوء النظرية الكلاسيكية (CTT) واجهت بعض القصور، حيث أن الدرجة الكلية للفرد في الاختبار أو المقياس تتأثر بخصائص المفردات، والعكس صحيح أي أن خصائص المفردات تتأثر بقدرات الأفراد، كما أن التكوين العاملي للاختبار أو المقياس يتغير مع مرور الزمن، ويتأثر ثبات الاختبار بالموقف الذي يتم فيه، وتتساوى تباينات أخطاء القياس لجميع أفراد العينة، فصدق وثبات أداة القياس وفقاً للنظرية الكلاسيكية يعتمد على خصائص العينة المستخدمة في التقنين (خلف، ٢٠١٩)؛ لذلك وجه المتخصصون في القياس جهودهم لإيجاد نظام قياس أكثر موضوعية يركز على انتقاء المفردات الاختبارية بشكل أفضل ويسمح بإضافة أو حذف مفردات المقياس دون أن يتأثر المقياس ككل (الموسوي، ٢٠١٨).

وقد أسفرت هذه الجهود عن ظهور بعض الاتجاهات الحديثة لتلافي أوجه القصور في النظرية الكلاسيكية، ومن بين هذه الاتجاهات نظرية الاستجابة للمفردة IRT، والتي حظيت باهتمام الباحثين حيث ساهمت في التغلب على كثير من مشكلات نظرية القياس الكلاسيكية (الدعيس، ٢٠١٨)؛ وقد فرضت نظرية الاستجابة للمفردة نفسها منذ بدايات القرن العشرين كنموذج جديد للقياس النفسي والتربوي، وتعتمد على فرضية أساسية حيث أن القيمة الاحتمالية لاستجابة شخص ما لمفردة اختبارية تكون دالة لكل من مستوى قدرة المفحوص، وخصائص المفردة (عطيه واخرون، ٢٠١٤).

وتحقق النظرية متطلبات القياس الموضوعي والمتمثل في تحرير تدريج أدوات القياس من خصائص الأفراد وتحرير تقديرات قدرة الأفراد من خصائص المفردات المستخدمة في القياس (Kreitchmann et al., 2024).

- نموذج الاستجابة المتدرجة Graded-Response Model (GRM)

هذا النموذج يمثل العلاقة بين مستوى قدرات الأفراد واحتمال استجابتهم في كل قسم من أقسام الاستجابة، ولا يشترط في هذا النموذج أن تكون جميع مفرداته تشتمل على نفس العدد من فئات الاستجابة، ولكن يكون لكل مفردة بارامتر تمييز واحد (a) ومجموعة من بارامترات العتبات الفارقة Thresholds والتي يرمز لها بالرمز (B1,B2,…..,Bn)، وعدد هذه البارامترات أقل من عدد فئات الاستجابة بواحد صحيح، وهو عدد المنحنيات المميزة الإجرائية لأقسام هذه المفردة، فإذا كان عدد فئات الاستجابة خمسة فئات مثل مقياس ليكرت، وكانت الدرجات الخاصة بهذه الفئات هي (صفر،١ ،٢ ، ٣، ٤) فإن عدد المنحنيات المميزة الإجرائية لهذه المفردة يساوي (٤) منحنيات لكل منها نفس معامل التمييز، ولكنها تختلف في بارامترات العتبات الفارقة (B1, B2, B3, B4) (علام، ٢٠٠٥).

ISSN: 2537-0464 eISSN: 2537-0472

- نموذج التقدير الجزئي العام Generalized partial credit model (GPCM)

يقدم نموذج التقدير الجزئي العام وصفًا للنماذج أحادية البعد متعددة الاستجابات. تشمل هذه النماذج: نموذج الاستجابة الاسميةNominal Response Model (NRM) ، ونموذج التقدير الجزئيPartial Credit Model (PCM)، ونموذج مقياس التقديرRating Scale model (RSM) والتي تعد جميعها فروقا في نموذج التقدير الجزئي العام. يستخدم نموذج التقدير الجزئي نموذج راش لتحديد احتمالية النجاح. بمعنى، نموذج التقدير الجزئي هو النظير لنموذج راش الذي يقدر معلمة الصعوبة " b". بالنسبة لـنموذج التقدير الجزئي ، يمكن لحجم العينة الصغير الذي يصل إلى ٣٠٠ أن ينتج تقديرًا موثوقًا للسمة (de Ayala, 2013).

- نموذج الفئات الأسمية The nominal categories model

وقد وضعه بوك (Bock,1972) كبديل لنموذج الاستجابات المتدرجة حيث لا يتطلب هذا النموذج أن تكون الاستجابات مرتبة مسبقًا بينما يتطلب نموذج الاستجابات المتدرجة ذلك، كما يتسم هذا النموذج بأنه يسمح بتنوع معاملات التمييز بين المفردات المختلفة، مثله فى ذلك مثل نموذج الاستجابة المتدرجة (Reeve, 2004).

جدول (٤) ملخص لنماذج نظرية الاستجابة للمفردة الاختبارية ثنائية الاستجابة ومتعددة الاستجابات

النموذج	شكل المفردة	الخصائص
نموذج راش	ثنائية الاستجابة	تتساوي جميع المفردات في معلم التمييز ويساوي (١)؛ ولكل مفردة معلم للصعوبة، ولا يوجد تأثير للتخمين.
النموذج اللوجستي ثنائي المعلم		لكل مفردة معلم للتمييز، ومعلم للصعوبة، ولا يوجد تأثير للتخمين.
النموذج اللوجستي ثلاثي المعلم		لكل مفردة معلم للتميز، ومعلم للصعوبة، ومعلم للتخمين
النموذج اللوجستي رباعي المعلم		لكل مفردة معلم للتميز، ومعلم للصعوبة، ومعلم للتخمين، ومعلم اللامبالاة (سقف الاستجابة للمفردة)
نموذج الاستجابة المتدرجة	متعددة الاستجابة	امتداد للنموذج الثنائي، يُستخدم في حالة المفردات ذات الاستجابات المتعددة المرتبة، ولكل مفردة معلم للتميز، وعدد من عتبات الاستجابات التي تنقص عن عدد فئات الاستجابة بواحد
نموذج التقدير الجزئي		امتداد للنموذج الأحادي، يفترض تساوي معلم التمييز بين جميع المفردات، لكل مفردة عدد من العتبات التي تمثل صعوبة الخطوات للحل أو الاستجابة وتختلف من مفردة لمفردة، ويناسب المسائل التي تتطلب عدد من الخطوات للحل، وكذلك المفردات متعددة الاستجابات

ISSN: 2537-0464 eISSN: 2537-0472

النموذج	شكل المفردة	الخصائص
نموذج سلم التقدير		امتداد للنموذج الأحادي، يفترض تساوي معلم التمييز بين جميع المفردات، وللمفردات عدد من العتبات التي تمثل الصعوبات النسبية للاستجابة وهي متساوية لجميع المفردات، ويناسب المفردات ذات الاستجابات المرتبة
نموذج بوك الاسمي		يُستخدم للاستجابات غير المرتبة، ويسمح باختلاف التمييز عبر المفردات

(Gyamfi & Acquaye, 2023)

الدراسات السابقة

المحور الأول: الدراسات التي تناولت العدد الأمثل لفئات الاستجابة:

أجرى عبد السمع (٢٠٢٠) دراسة هدفت إلى التعرف على تأثير عدد فئات الاستجابة لليكرت (٣، ٥، ٧) وعدد المشاركين على دقة معاملي ألفا وأميجا في تقدير ثبات درجات المقياس النفسي، وتكون عدد المشاركين في الدراسة من ٢٤٣ وطالبا وطالبة موزعين على أربع مجموعات، واستخدم مقياس فاعلية الذات الوجدانية وأشارت النتائج إلى زيادة دقة معاملي ألفا وأوميجا بزيادة عدد فئات الاستجابة وخاصة في حالة العدد الخماسي والسباعي كما وأظهرت النتائج أن معامل أوميجا أدى إلى تقدير أدق لقيم معامل ثبات الدراجات مقارنة بمعامل ألفا في نسبة التباين المفسر واتساع المسافة بين حدود الثقة.

كما درس حميدة (٢٠٢٠) مدى اختلاف عدد فئات الاستجابة على كلا من معاملات الثبات والصدق لأداة القياس وتم تصميم واستخدام مقياس القلق الرقمي ذو نموذجين يختلف في عدد بدائل الاستجابات (ثلاثي وخماسي) ، وبلغ حجم العينة (١٥٠) طالباً (٧٥ ذكور، ٧٥ إناث، تتروح أعمارهم بين ١٠- ٣٠ عاماً، ٥٠ من المرحلة الأساسية، ٥٠ من المرحلة الثانوية، ٥٠ من المرحلة الجامعية) وتم اختيارهم عن طريق العينة القصدية والعشوائية الطبقية. ، وتوصلت نتائج الدراسة إلى اختلاف في الخصائص السيكومترية لأداة القياس في ضوء تباين عدد فئات الاستجابة، وأن أفضل عدد فئة استجابة هو من نوع خمس فئات استجابة.

وبحث القحطاني (٢٠٢١) تأثير مستوى الاكتئاب وعدد فئات الاستجابة على الخصائص السيكومرتية لمقياس الشعور بالوحدة النفسية، وهدفت الدراسة إلى التعرف على اختلاف عدد فئات الاستجابة (أربعة فئات – خمسة فئات) على الخصائص السيكومترية، والتفاعل بين مستويات الاكتئاب على الخصائص السيكومترية، واستخدمت الباحثة مقياس بيك للاكتئاب ومقياس الشعور لراسيل واختارت الباحثة عينة عشوائية عددها ٩٠ فردا من طالبات جامعة الإمام محمد بن

ISSN: 2537-0464 eISSN: 2537-0472

سعود الإسلامية، وخلصت النتائج إلى أنه لا يوجد اختلاف بين أربعة وخمس فئات استجابة في الخصائص السيكومترية للمقياس.

وفي دراسة Cernat et al. (2022) تعرفت على تأثير تصميم شكل الاستبيان وطول الاستبيان على جودة القياس حيث أخرج الباحثون ثلاثة تصاميم مختلفة بشكل استبيان ونمط استجابة مختلفة وكان حجم العينة (٥٣٣٨) مستجيباً، وتوصلت نتائج الدراسة إلى أن طول الاستبيان وفئات الاستجابة تؤثر على جودة البيانات المجمعة لأنه يؤدي إلى إرهاق المستجيب مما يدفع به إلى الإدلاء باستجابات غير دقيقة أو استبيانات غير مكتملة وأن الاستبيانات القصيرة وفئات الاستجابة القليلة قد لا توفر ما يكفي من المعلومات.

وفي دراسة Chakrabartty (2023) التي تقصت العدد الأمثل لفئات الاستجابة لتحويل نتائج فئات مقياس ليكرت ذات عدد مختلف (٣، ٤، ٥، ٧) إلى فترات مستمرة وموزعة بشكل طبيعي تراوح بين ١ و ٥ وتوصلت الدراسة إلى أن عدد فئات الاستجابة لم يكن له تأثير كبير على الثبات والصدق وتوصلت الدراسة إلى أنه لا يمكن تحديد عدد أمثل لفئات الاستجابة يطبق على جميع الدراسات.

المحور الثاني: الدراسات التي تناولت الترتيب الأمثل لفئات الاستجابة:

في دراسة Mackinnon and Wang (2020) التي بحثت تأثير ترتيب الاستجابة للاستبيانات باستخدام التقييم الذاتي لاستكشاف دور الترتيب في التحيز للاستجابة، وافترض الباحثون أن المشاركين من المرجح أن يختاروا فئات الاستجابة التي تم وضعها في أول قائمة فئات الاستجابة، وهي ظاهرة تسمى تأثير الأولوية وشارك في تطبيق الدراسة ٧٧٤ طالباً جامعياً من جامعة دالهوسي (Dalhousie)، وأسلوب فئة استجابة خماسي، في الأولى ترتيب تصاعدي تبدأ بفئة (لا أوافق بشدة) والأخرى ترتيب تنازلي تبدأ بفئة (أوافق بشدة)، وأظهرت النتائج أن ترتيب الاستجابات يؤثر على اختيار المستجيب للاستجابة وأن المستجيبين يفضلون اختيار الاختيارات التي ترتيبها أول فئات الاستجابة كما وكشفت الدارسة أن بعض الاستجابات تكون إجابتها في نهاية ترتيب فئات الاستجابة إلا أن بعض المشاركين انحاز إلى اختيار الفئات الأولى من الاستجابة في كلا التجربتين التصاعدي والتنازلي وأن تأثيرات تربيب الاستجابة يمكن أن تختلف تبعا لاختلاف لمفهوم ومحتوى الاستجابة.

وبحث Robie et al. (2022) تأثير ترتيب فئة الاستجابة على الخصائص السيكومترية لمقياس ليكرت وتفاعل المستجيبين، واستخدمت الدراسة مقياسين من نوع ليكرت في المقياس الأول الترتيب تصاعدي والثاني الترتيب تنازلي، وتم تبديل ترتيب فئات الاستجابة، وطبق المقياسين على عينة عشوائية من طلاب جامعة جنوب

ISSN: 2537-0464 eISSN: 2537-0472

أونتاريو عددهم ١٢٠٠ طالباً وطالبة، وتشير النتائج إلى أن ترتيب فئة الاستجابة قد يؤثر بشكل محدود على الخصائص السيكومترية للمقياس وربما ينعدم تأثيره.

وفي دراسة Kato (2023) تمثل الهدف في التعرف على تأثير ترتيب مقياس ليكرت على تقييمات المستجيبين وطبق الباحث أداة استطلاع رأي على عينة عشوائية عددها ٣٣٨ طالبا وطالبة من جامعات اليابان، لقياس مستوى الولاء الجامعي، وأظهرت النتائج أن ترتيب الاختيارات يؤثر على تقييمات المستجيبين، حيث تم الحصول على تقييمات أعلى عند استخدام ترتيب الاختيارات التصاعدي مقارنة بترتيب الاستجابات بشكل تنازلي، والاختيارات التي تظهر في الجزء الأعلى من المقياس تتلقى تقييمات أعلى وقد تم تفسير هذا التأثير بأن الاختيارات التي تظهر في الجزء العلوي من المقياس تتلقى تقييمات أعلى لأن المستجيبين يميلون إلى الاستجابة للفقرات بشكل أكثر إيجابية في البداية، وأن ترتيب فئات الاستجابة بشكل تصاعدي أفضل لتحسين جودة البيانات المجمعة وتقييمات المستجيبين.

وهدفت دراسة (العمراني والمالكي، ٢٠٢٣) إلى الكشف عن أثر طريقة ترتيب فئات التدريج (تصاعدي، تنازلي) ونموذج الاستجابة للفقرة المستخدم في دقة تقدير بارامترات الفقرات وقدرات الأفراد باستخدام نموذجي التقدير الجزئي (PCM) والاستجابات المتدرجة (GRM) ولتحقيق أهداف الدراسة تم استخدام مقياس اتخاذ القرار والذي قام بإعداده عبدون (٢٠٠٤) الصورة ب، والمكون من ٣٤ فقرة يقابلها استجابات بصورة مقياس ليكرت الخماسي. وطبق الاختبار على عينة تكونت من (٢٣٨٥) فردا من طلاب وطالبات جامعة أم القرى وقد تم استخدام برمجية (25) : SPSSو PARSCALEلتحليل نتائج الدراسة، حيث تم تقدير الأخطاء المعيارية في تقدير قدرات أفراد العينة وبارامترات صعوبة الفقرات وفق طريقتين من طرق ترتيب فئات التدريج وباستخدام نموذجي التقدير الجزئي والاستجابات المتدرجة، كما تم تقدير الأخطاء المعيارية في تقدير بارامترات تمييز الفقرات باستخدام نموذج الاستجابات المتدرجة، وأظهرت النتائج أنه: يوجد فرق دال إحصائيا بين متوسطات الأخطاء المعيارية في تقدير بارامترات صعوبة الفقرات يعزى إلى نموذج الاستجابة للفقرة المستخدم في تدريجها وكان هذا الفرق لصالح نموذج التقدير الجزئي، كما أنه لا يوجد فرق دال إحصائيا بين متوسطات الأخطاء المعيارية في تقدير بارامترات صعوبة الفقرات يعزى إلى طريقة ترتيب فئات التدريج، كما أنه لا يوجد فرق دال إحصائيا بين متوسطات الأخطاء المعيارية في تقدير بارامترات صعوبة الفقرات يعزى إلى التفاعل بين طريقة ترتيب فئات التدريج والنموذج المستخدم في التدريج، ولا يوجد فرق دال إحصائيا بين متوسطي الأخطاء المعيارية في تقدير بارامترات تمييز الفقرات باستخدام نموذج الاستجابات المتدرجة يعزى إلى طريقة ترتيب فئات

ISSN: 2537-0464 eISSN: 2537-0472

التدريج. وأسفرت النتائج أيضا عن وجود فرق دال إحصائيا بين متوسطات الأخطاء المعيارية في تقدير بارامترات قدرات الأفراد يعزى إلى نموذج الاستجابة للفقرة المستخدم في تدريجها، وكان هذا الفرق لصالح نموذج التقدير الجزئي، كما وجد فرق دال إحصائيا بين متوسطات الأخطاء المعيارية في تقدير بارامترات قدرات الأفراد يعزى إلى طريقة ترتيب فئات التدريج، كما أنه لا يوجد فرق دال إحصائيا بين متوسطات الأخطاء المعيارية في تقدير بارامترات قدرات الأفراد يعزى إلى التفاعل بين طريقة ترتيب فئات التدريج والنموذج المستخدم في التدريج.

التعقيب على الدراسات السابقة:

بناءً على هذه النتائج رأى الباحث استخدام الترتيب التنازلي مقابل الترتيب التصاعدي لفئات الاستجابة؛ للتحقق من وجود تأثير لهذا النوع من الترتيب على الخصائص السيكومترية للمقياس النفسي ومفرداته في ضوء نظرية الاستجابة للمفردة. وبصفة عامة من خلال الاطلاع على الدراسات السابقة (التي أتيح للباحث الاطلاع عليها) يلاحظ أن هناك ندرة في تطبيقات نماذج نظرية الاستجابة للمفردة الاختبارية في حالة أدوات القياس متعددة الاستجابة بعكس الدراسات التي تمت في ضوء النظرية الكلاسيكية، وعليه فإن الدراسة المقترحة ستحاول استكمال جهود الدراسات السابقة من خلال تقصي تأثير عدد فئات الاستجابة وترتيبها والتفاعل بينهما على دالة المعلومات ومعالم المفردات والخطأ المعياري في تقديرها، ففي حدود علم الباحث لا توجد دراسات توصلت إلى نتائج حاسمة للعدد الأمثل لفئات الاستجابة وترتيبها، وتضيف الدراسة في مساعدة الباحثين والدارسين لتحسين جودة الخصائص السيكومترية لأدوات القياس، ويأمل الباحث في نهاية هذه الدراسة أن يصل إلى توصيات مفيدة في الجانب التطبيقي.

منهج الدراسة وإجراءاتها

منهج الدراسة:

تم استخدام المنهج الوصفي نظراً لملائمته لإجراءات الدراسة، وذلك بالاعتماد على نموذج الاستجابات المتدرجة في نظرية الاستجابة للمفردة الاختبارية والذي يتيح تقديرات لمعالم المفردات (صعوبة وتمييز) ولمعالم القدرة والاخطاء المعيارية في تقديرها ودوال المعلومات للمقياس وللمفردات.

مجتمع الدراسة:

طلاب وطالبات جامعة القصيم لمرحلة البكالوريوس الفصل الدراسي الأول للعام الجامعي ٢٠٢٤-٢٠٢٥م، لطلاب وطالبات كلية اللغات والدراسات الإنسانية وكلية التربية وكلية العلوم التطبيقية وكلية الشريعة وكلية الهندسية وكلية العلوم الطبية.

ISSN: 2537-0464 eISSN: 2537-0472

عينة الدراسة:

تكونت عينة الدراسة من ١٠٤٠ طالباً وطالبة من طلاب جامعة القصيم في الفصل الدراسي الأول من العام الجامعي ١٤٤٦هـ – ٢٠٢٤م وتم اختيارهم بطريقة عينة عشوائية طبقية وفقا لمتغيرات الجنس والمستوى الدراسي والعمر .

أداة الدراسة:

مقياس الاستعداد الأكاديمي للطالب الجامعي (إعداد: الثبيتي، ٢٠٢١)

تم في الدراسة الحالية استخدام مقياس الاستعداد الأكاديمي للطالب الجامعي من إعداد الثبيتي (٢٠٢١)، وتم استخدام هذا المقياس لما يتحقق فيه من خصائص تناسب الدراسة الحالية، حيث أنه أحادي البعد ويتمتع بخاصية اللاتغاير للمفردات وتم التأكد من مطابقة مفرداته لنموذج الاستجابة المتدرجة، والمقياس صمم في صورته الأصلية (ملحق ١) وفقاً لتدرج ليكرت الخماسي (موافق بشدة، موافق، محايد، غير موافق، غير موافق بشدة) وتم تطويره في الدراسة الحالية بما يتناسب مع أهداف الدراسة.

ويتكون المقياس في نسخة الأصلية من (٢٢) مفردة موزعة على أربع أبعاد (بُعد أسلوب التعلم، بعد إدارة الوقت، بعد التفكير وحل المشكلات، بعد التواصل والعمل الجماعي) وقام الباحث بتطوير المقياس بما يتناسب مع أهداف الدراسة الحالية في رفع عدد فئات الاستجابة من خمس فئات استجابة إلى سبع فئات استجابة وفقا لتدرج ليكرت السباعي (موافق بشدة، موافق بدرجة متوسطة، موافق بدرجة قليلة، محايد، غير موافق بدرجة قليلة، غير موافق بدرجة متوسطة، غير موافق بشدة) وتطوير نسخة من المقياس يتم فيها عرض الاستجابات بترتيب من الإيجابي إلى السلبي (الترتيب التنازلي) (موافق بشدة، موافق بدرجة متوسطة، موافق بدرجة قليلة، محايد، غير موافق بدرجة قليلة، غير موافق بدرجة متوسطة، غير موافق بشدة) (ملحق ٣)؛ ونسخة أخرى للمقياس يتم فيها عرض الاستجابات بترتيب من السلبي إلى الإيجابي (الترتيب التصاعدي) (غير موافق بشدة، غير موافق بدرجة متوسطة، غير موافق بدرجة قليلة، محايد، موافق بدرجة قليلة، موافق بدرجة متوسطة، موافق بشدة) .

الخصائص السيكومترية لأداة الدراسة:

أ- صدق المقياس:

قام الباحث معد المقياس بصورته الأصلية الثبيتي (٢٠٢١)، بالتحقق من صدق المقياس عن طريق صدق المحكمين؛ حيث تم عرضه في صورته الأولية على عدد من المتخصصين في علم النفس، وعلم النفس التربوي، والقياس والتقويم، والإرشاد والتوجيه النفسي، بلغ عددهم (٦) محكمين، وطلب منهم الحكم على مدى ملاءمة

ISSN: 2537-0464 eISSN: 2537-0472

العبارات من حيث الصياغة اللغوية، ومناسبتها المجتمع الدراسة، وذلك في ضوء التعريفات المذكورة اصطلاحيا وإجرائيا للاستعداد الأكاديمي.

وقد جاءت آراؤهم بنسبة اتفاق تقدر ب (٨٣,٣٣) بين المحكمين الذين أقروا أن المقياس مناسب لما وضع لقياسه، وطالبوا بحذف عدد (٦) عبارات، وأصبح المقياس في صورته النهائية مكونا من (٢٢) عبارة، وهذا الإجراء أشار إلى أن أدوات الدراسة تتمتع بصدق منطقي مرتفع. ولحساب معامل صدق الاتساق الداخلي للمقياس، قام الباحث بتطبيقها على عينة استطلاعية بلغت (٦٠) طالباً وطالبة من خارج عينة الدراسة لمعرفة ما إذا كان هناك غموض في بعض العبارات أم لا، وللتأكد من سلامة التعليمات وسهولة التطبيق.

واتفق المحكمون على أن المقياس بسيطاً وغير معقداً وهو ما يتوافق مع خصائص جزء كبير من عينة الدراسة وهم الطلاب حديثي الالتحاق بالتعليم الجامعي وربما لو تم استخدام بعضا من المصطلحات الاكاديمية، قد تؤثر في فهم الطلاب لعبارات المقياس وهو ما يزيد من أخطاء القياس المرتبطة بخصائص العينة.

ب- ثبات المقياس:

وفيما يتعلق بثبات المقياس قام الثبيتي (٢٠٢١) بحساب الثبات باستخدام معامل ألفا كرونباخ؛ فقد كانت قيمة معامل ثبات ألفا كرونباخ للمقياس ككل (٠,٩٢). كما تم حساب معامل ثبات ألفا كرونباخ للأبعاد الأربعة المكونة للمقياس، حيث كانت قيمة معامل ثبات البعد الأول (٠٩٣) ، والبعد الثاني (٠٨٨) ، والبعد الثالث (٠,٩١)، والبعد الرابع .(٠,٨٧)

إضافة لذلك وخلال إجراءات الدراسة حسب الثبيتي (٢٠٢١) الثبات وفق نظرية استجابة الفقرة وباستخدام برنامج (MULTILOG7) ، فكانت قيمة معامل الثبات للمقياس ككل (٠,٨٨) ، وللبعد الأول (٠,٩٢) ، والبعد الثاني (٠,٩١) ، والبعد الثالث (٠,٨٩) ، والبعد الرابع (٠,٨٨). وقد جاءت القيم المحسوبة لمعاملات الثبات مرتفعة أعلى من (٠٨)، وتدل على أن المقياس يتمتع بدرجة عالية من الثبات ومقبولة إحصائيا، وبذلك يكون الباحث معد المقياس قد تأكد من صدق المقياس وثباته مما يجعله على ثقة بصلاحيته لتحقيق أهداف الدراسة. وفي الدراسة الحالية قام الباحث بالتحقق من ثبات المقياس من خلال الإجابة عن السؤال الثالث من أسئلة البحث، حيث تراوحت قيم معامل الثبات في المواقف العشرة المختلفة في عدد فئات الاستجابة وترتيبها بين ٠.٨٥٥ و ٠.٩٤١ وهي قيم عالية وتدل على تمتع الأداء على المقياس بدرجة عالية من الثبات.

ISSN: 2537-0464 eISSN: 2537-0472

نتائج الدراسة وتفسيرها

تم في هذا الفصل تقديم تحليل شامل لنتائج الدراسة من خلال عرض ومناقشة وتفسير النواتج التي تم التوصل إليها، وفي نهاية الفصل تم تقديم ملخصاً لأبرز النتائج، يليه عرض للتوصيات والمقترحات التي يمكن أن تساهم في تطوير الممارسات البحثية والتطبيقية مستقبلاً بناءً على ما توصلت إليه الدراسة الحالية من نتائج.

نتائج السؤال الأول:

نص السؤال الأول للدراسة الحالية على "هل تختلف معالم المفردات والأخطاء المعيارية في تقديرها باختلاف عدد فئات الاستجابة وترتيبها؟"

(أ): معلم التمييز:

للإجابة عن هذ السؤال في تم حساب المتوسطات والانحرافات المعيارية لقيم معلم التمييز للمفردات والأخطاء المعيارية في تقديرها وهو ما يتضح في جدول (١٠)، ثم الكشف عن دلالة الفروق فيها وفقاً لاختلاف عدد فئات الاستجابة وترتيبها باستخدام تحليل التباين للقياسات المتكررة، وهو ما يتضح في جدول (٥):

جدول (٥) المتوسطات الحسابية والانحرافات المعيارية لتمييز المفردات والأخطاء المعيارية في التقدير وفقا لعدد فئات الاستجابة وترتيبيها

م	عدد فئات الاستجابة وترتيبها	معلم التمييز		الأخطاء المعيارية	
		المتوسط الحسابي	الانحراف المعياري	المتوسط الحسابي	الانحراف المعياري
١	ثلاثي إيجابي	1.675	0.482	0.166	0.041
٢	ثلاثي سلبي	1.663	0.535	0.165	0.035
٣	رباعي إيجابي	1.597	0.616	0.147	0.031
٤	رباعي سلبي	1.687	0.554	0.129	0.023
٥	خماسي إيجابي	1.656	0.500	0.144	0.026
٦	خماسي سلبي	1.709	0.552	0.142	0.027
٧	سداسي إيجابي	1.630	0.534	0.117	0.019
٨	سداسي سلبي	1.681	0.550	0.116	0.024
٩	سباعي إيجابي	1.632	0.518	0.120	0.025
١٠	سباعي سلبي	1.695	0.545	0.112	0.019

تُظهر النتائج في الجدول (٥) أن هناك فروقاً ظاهرية في متوسطات معلم التمييز والأخطاء المعيارية في تقديرها، تبعاً لعدد فئات الاستجابة وترتيبها، ويلاحظ أن متوسطات معاملات التميز للفئات السلبية (ثلاثي، رباعي، خماسي، سداسي، سباعي)

ISSN: 2537-0464 eISSN: 2537-0472

كانت أعلى بشكل عام مقارنة بنظيراتها الإيجابية، وبالنسبة للأخطاء المعيارية فهي بصفة عامة تقل بزيادة عدد فئات التدريج، وتشير هذه النتائج إلى أن ترتيب الفئات (إيجابي/سلبي) قد يكون له تأثير أكبر على معاملات التميز من عدد الفئات نفسها، مع وجود ميل لارتفاع معاملات التميز في الفئات ذات الترتيب السلبي، وفيما يلي دلالة الفروق في معلم التمييز والأخطاء المعيارية في التقدير باستخدام تحليل التباين للقياسات المتكررة والاعتماد على تصحيح جرينهاوس-جيسر (Greenhouse-Geisser):

جدول (٦) دلالة الفروق في معلم التمييز والأخطاء المعيارية في تقديرها وفقاً لعدد فئات الاستجابة وترتيبها

المتغيرات	مصدر التباين	مجموع المربعات**	درجات الحرية	متوسط المربعات	F	مستوى الدلالة	مربع إيتا الجزئي
معلم التمييز	عدد فئات الاستجابة وترتيبها	0.212	9	0.024	0.979	0.459	0.049
		0.212	2.429	0.089	0.979	0.397	0.049
	الخطأ	4.114	171	0.024			
		4.114	46.151	0.089			
الأخطاء المعيارية	عدد فئات الاستجابة وترتيبها	0.072	9	0.008	37.756	0.01	0.665
		0.072	3.05	0.024	37.756	0.01	0.665
	الخطأ	0.036	171	0.000212			
		0.036	57.001	0.000636			

**في جميع الجدول المشابهة، مجموع المربعات في الصف الأول يعبر عن النتائج بدون تصحيح لدرجة الحرية، والصف الثاني يمثل النتائج بعد تصحيح درجة الحرية في حالة انتهاك افتراض الكروية.

يتضح من جدول (٦) أنه:

- لا توجد فروق دالة إحصائياً في معلم التمييز ترجع لاختلاف عدد فئات الاستجابة وترتيبها، وكانت قيمة مربع إيتا الجزئي تساوي ٠.٠٤٩ وهو حجم تأثير ضعيف، مما يؤكد أن الفروق في معلم التمييز وفقا لتأثير عدد فئات الاستجابة وترتيبها فروق ضعيفة.
- توجد فروق دالة إحصائياً عند مستوى دلالة ٠.٠١ في الأخطاء المعيارية في تقدير معلم التمييز ترجع لاختلاف عدد فئات الاستجابة وترتيبها، وكانت قيمة مربع إيتا الجزئي تساوي ٠.٦٦٥ وهو حجم تأثير كبير، مما يؤكد أن الفروق في

ISSN: 2537-0464 eISSN: 2537-0472

الأخطاء المعيارية لتقدير معلم التمييز وفقا لتأثير عدد فئات الاستجابة وترتيبها فروق كبيرة.

وبالنسبة للأخطاء المعيارية في تقدير معلم التمييز تشير النتائج إلى أن زيادة عدد فئات الاستجابة يرتبط بانخفاض الأخطاء المعيارية لمعاملات التمييز وتقليل تشتتها، مما يعني أن المقاييس ذات فئات الاستجابة الأكبر (سداسية وسباعية) تقدم تقديرات أكثر دقة واستقرارًا لمعلم التمييز مقارنة بالمقاييس ذات فئات الاستجابة الأقل (ثلاثية ورباعية)، مع ملاحظة عدم وجود فروق كبيرة بين الترتيب الإيجابي والسلبي ضمن نفس عدد فئات الاستجابة؛ وتشير النتائج إلى أن حجم تأثير عدد فئات الاستجابة وترتيبها على الأخطاء المعيارية في تقدير معلم التمييز كان مرتفعًا، وأن عدد فئات الاستجابة وترتيبها يفسر حوالي ٦٦.٥% من التباين في الأخطاء المعيارية لمعلم التمييز، وهذا يعكس تأثيرًا قويًا وذا أهمية عملية كبيرة، وجدول (٧) يوضح نتائج المقارنات الزوجية وفقا لعدد فئات الاستجابة وترتيبها باستخدام اختبار أقل فرق دال LSD كاختبار للمقارنات البعدية:

جدول (٧) المقارنات الزوجية بين الاخطاء المعيارية في تقدير معلم التمييز وفقا لعدد فئات الاستجابة وترتيبها

التدريج	ثلاثي إيجابي (م=.166)	ثلاثي سلبي (م= .165)	رباعي إيجابي (م= .147)	رباعي سلبي (م= .129)	خماسي إيجابي (م=.144)	خماسي سلبي (م= .142)	سداسي إيجابي (م= .117)	سداسي سلبي (م= .116)	سباعي إيجابي (م= .120)
ثلاثي سلبي (م= .165)	0.001								
رباعي إيجابي (م= .147)	.019**	.018**							
رباعي سلبي (م= .129)	.037**	.036**	.018**						
خماسي إيجابي (م=.144)	.022**	.021**	0.003	.015**					
خماسي سلبي (م= .142)	.024**	.023**	0.005	.013**	0.002				
سداسي إيجابي (م= .117)	.049**	.048**	.030**	.012**	.027**	.025**			
سداسي سلبي (م= .116)	.050**	.049**	.031**	.013**	.028**	.026**	0.001		
سباعي إيجابي (م= .120)	.046**	.045**	.027**	0.009	.024**	.022**	0.003	0.004	
سباعي سلبي (م= .112)	.054**	.053**	.035**	.017**	.032**	.030**	.005*	0.004	0.008

** الفروق بين المتوسطين دالة إحصائياً عند مستوى دلالة ٠.٠١

تشير نتائج المقارنات الزوجية بكما هو موضح بجدول (٧) إلى وجود فروق دالة إحصائياً عند مستوى دلالة ٠.٠١ في الأخطاء المعيارية لتقديرات معلم التمييز

ISSN: 2537-0464 eISSN: 2537-0472

باختلاف عدد فئات الاستجابة وترتيبها، حيث أظهرت النتائج أن التدريج الثلاثي الإيجابي يليه التدريج الثلاثي السلبي ثم الرباعي الإيجابي ثم الخماسي الإيجابي والخماسي السلبي ثم الرباعي السلبي لها أعلى متوسط أخطاء معيارية مقارنة بجميع الفئات الأخرى للاستجابة وترتيبها، وهذا يشير إلى أن النماذج ذات الفئات الأقل للاستجابة (الثلاثية والرباعية والخماسية) سواء كان ترتيب الاستجابة سلبي أو إيجابي تنتج أخطاءً معيارية أكبر في تقدير معلم التمييز، مقارنة بالنماذج ذات الفئات الأكثر (السداسية والسباعية)، وتظهر النتائج هنا أنه في حالة التدريج السداسي والسباعي تقل الأخطاء المعيارية في تقدير معلم التمييز تدريجياً في حالة الترتيب السبي عنه في حالة الترتيب الإيجابي، مما يدعم الاستنتاج بأن الفئات ذات المستويات الأعلى والترتيب السلبي هي الأفضل من حيث الأخطاء المعيارية في تقدير معلم التمييز.

(ب): معلم الصعوبة (العتبات):

تم حساب المتوسطات والانحرافات المعيارية لقيم معلم العتبات والأخطاء المعيارية في تقديرها، وهو ما يتضح في جدول (٨)، ثم الكشف عن دلالة الفروق فيها وفقاً لاختلاف عدد فئات الاستجابة وترتيبها باستخدام تحليل التباين للقياسات المتكررة، وهو ما يتضح في جدول (٨):

جدول (٨) متوسطات العتبات والانحرافات المعيارية في تقديرها وفقا لعدد فئات الاستجابة وترتيبيها

م	عدد فئات الاستجابة وترتيبها	معلم العتبات		الإخطاء المعيارية	
		المتوسط الحسابي	الانحراف المعياري	المتوسط الحسابي	الانحراف المعياري
١	ثلاثي إيجابي	-0.881	0.762	0.126	0.100
٢	ثلاثي سلبي	-0.649	0.793	0.161	0.288
٣	رباعي إيجابي	-0.915	0.832	0.145	0.125
٤	رباعي سلبي	-0.736	0.560	0.165	0.279
٥	خماسي إيجابي	-0.828	0.784	0.143	0.145
٦	خماسي سلبي	-0.587	0.616	0.161	0.297
٧	سداسي إيجابي	-0.853	0.774	0.161	0.138
٨	سداسي سلبي	-0.643	0.533	0.131	0.089
٩	سباعي إيجابي	-0.776	0.729	0.156	0.131
١٠	سباعي سلبي	-0.562	0.536	0.125	0.093

تظهر نتائج الجدول (٨) اختلافات في متوسطات العتبات وفقًا لعدد فئات الاستجابة وترتيبها، حيث تميل الفئات ذات الترتيب الإيجابي إلى إظهار متوسطات

ISSN: 2537-0464 eISSN: 2537-0472

عتبات أكثر سلبية (عتبات أقل) مقارنة بنظيراتها ذات الترتيب السلبي (عتبات أعلى)، مما يشير إلى أن اتجاه ترتيب فئات الاستجابة له تأثير ملموس على تقديرات العتبات. كما يلاحظ أن التدريج الرباعي بترتيب إيجابي سجل أدنى متوسط للعتبات، بينما التدريج السباعي بترتيب سلبي أظهر أعلى متوسط للعتبات،

وبالنسبة للأخطاء المعيارية في تقدير العتبات نلاحظ أن متوسطات الأخطاء المعيارية للعتبات وفقًا لعدد فئات الاستجابة وترتيبها، لم يظهر نمط ثبات للتغير، حيث كانت أقل متوسطات الأخطاء المعيارية في حالة التدريج السباعي السلبي، يليه التدريج الثلاثي الإيجابي، ثم السداسي السلبي يليه الخماسي الإيجابي، ثم الرباعي الإيجابي فالسباعي الإيجابي، يليه الخماسي السلبي، ثم الثلاثي السلبي يليه السداسي الإيجابي ثم في الأخير الرباعي السلبي، مع وجود تباين ملحوظ في الانحرافات المعيارية بين الترتيب الإيجابي والسلبي، ففي حالة الترتيب السلبي (الثلاثي، الرباعي، والخماسي) سجلت انحرافات معيارية أعلى (٠.٢٨٨، ٠.٢٧٩، ٠.٢٩٧ على التوالي) مقارنة بنظيراتها الإيجابية، بينما انعكس هذا الاتجاه في المقاييس ذات الفئات الأكبر (السداسي والسباعي)، حيث أظهرت الفئات ذات الترتيب السلبي انحرافات معيارية أقل (٠.٠٨٩، ٠.٠٩٣ على التوالي)، وهذا يشير إلى أن تأثير اتجاه الترتيب (إيجابي/ سلبي) على دقة القياس يختلف باختلاف عدد فئات الاستجابة، وبالاعتماد على تصحيح جرينهاوس-جيسر (Greenhouse-Geisser) لدرجات الحرية كانت دلالة الفروق في تقدير العتبات والأخطاء المعيارية في تقديرها كما هي موضحة بجدول (٩):

جدول (٩) دلالة الفروق في متوسطات العتبات والأخطاء المعيارية في تقديرها باختلاف عدد فئات الاستجابة وترتيبها

المتغيرات	مصدر التباين	مجموع المربعات	درجات الحرية	متوسط المربعات	F	مستوى الدلالة	مربع إيتا الجزئي
العتبات	عدد فئات الاستجابة وترتيبها	2.906	9	0.323	10.094	0.000	0.350
		2.906	1.723	1.686	10.218	0.001	0.350
	الخطأ	5.399	171	0.032			
		5.399	32.737	0.165			
الأخطاء المعيارية في تقدير العتبات	عدد فئات الاستجابة وترتيبها	0.044	9	0.005	0.625	0.802	0.030
		0.044	1.066	0.042	0.599	0.460	0.030
	الخطأ	1.423	171	0.008			
		1.423	20.257	0.070			

ISSN: 2537-0464 eISSN: 2537-0472

يتضح من جدول (٩) أنه:

- توجد فروق دالة إحصائياً عند مستوى دلالة ٠.٠١ في معلم العتبات ترجع لاختلاف عدد فئات الاستجابة وترتيبها، وكانت قيمة مربع إيتا الجزئي تساوي ٠.٣٥٠ وهو حجم تأثير كبير، مما يؤكد أن الفروق في معلم العتبات وفقا لتأثير عدد فئات الاستجابة وترتيبها فروق كبيرة.
- لا توجد فروق دالة إحصائياً في الأخطاء المعيارية في تقدير معلم العتبات ترجع لاختلاف عدد فئات الاستجابة وترتيبها، وكانت قيمة مربع إيتا الجزئي تساوي ٠.٠٣٠ وهو حجم تأثير ضعيف، مما يؤكد أن الفروق في الأخطاء المعيارية لتقدير معلم العتبات وفقا لتأثير عدد فئات الاستجابة وترتيبها فروق ضعيفة.

- نتائج السؤال الثاني:

نص السؤال الثاني للدراسة الحالية على "هل تختلف دالة المعلومات لمفردات المقياس باختلاف عدد فئات الاستجابة وترتيبها؟"

للإجابة عن هذ السؤال تم حساب المتوسطات والانحرافات المعيارية لقيم دوال المعلومات للمفردات كما يتضح من جدول (١١)، ثم الكشف عن دلالة الفروق وفقاً لاختلاف عدد فئات الاستجابة وترتيبها باستخدام تحليل التباين للقياسات المتكررة، فكانت النتائج كما هي موضحة في جدول (١٠):

جدول (١٠) متوسطات دوال المعلومات والانحرافات المعيارية وفقا لعدد فئات الاستجابة وترتيبها

م	عدد فئات الاستجابة وترتيبها	المتوسط الحسابي	الانحراف المعياري
١	ثلاثي إيجابي	0.351	0.124
٢	ثلاثي سلبي	0.349	0.132
٣	رباعي إيجابي	0.474	0.202
٤	رباعي سلبي	0.487	0.214
٥	خماسي إيجابي	0.487	0.204
٦	خماسي سلبي	0.514	0.224
٧	سداسي إيجابي	0.609	0.290
٨	سداسي سلبي	0.655	0.313
٩	سباعي إيجابي	0.630	0.292
١٠	سباعي سلبي	0.682	0.327

تُظهر نتائج جدول (١٠) متوسطات دوال المعلومات للمفردات والانحرافات المعيارية لها في كل حالة من حالات عدد فئات الاستجابة وترتيبها، ويلاحظ بصفة عامة أن هناك زيادة تدريجية ومنتظمة في متوسطات دوال المعلومات مع زيادة عدد

ISSN: 2537-0464 eISSN: 2537-0472

فئات الاستجابة، حيث كانت أعلى قيم لدوال المعلومات في حالة التدريج السباعي بينما أقلها في حالة التدريج الثلاثي، وهذا يشير إلى أن زيادة عدد فئات الاستجابة يؤدي إلى زيادة كمية المعلومات عن السمة المقاسة وهو ما يعكس الدقة في القياس.

كما يلاحظ أن المقاييس ذات الترتيب السلبي لفئات الاستجابة سجلت متوسطات أعلى مقارنة بنظيراتها ذات الترتيب الإيجابي لنفس عدد فئات الاستجابة، باستثناء التدريج الثلاثي حيث كان الفرق ضئيلاً جداً بين الترتيب الإيجابي والترتيب السلبي، وهذا النمط يشير إلى أن ترتيب الفئات بشكل سلبي يوفر دقة أكثر في قياس السمة المستهدفة مقارنة بالترتيب الإيجابي، وللتحقق من دلالة الفروق بين دوال المعلومات للمفردات باختلاف عدد فئات الاستجابة وترتيبها كانت نتائج تحليل التباين للقياسات المتكررة بالاعتماد على تصحيح جرينهاوس-جيسر **Greenhouse-Geisser** لدرجات الحرية كما هي موضحة في جدول (١١):

جدول (١١) جدول تحليل التباين ذو القياسات المتكررة للمقارنات داخل المجموعات باختلاف عدد فئات الاستجابة وترتيبها في دوال المعلومات

مصدر التباين	مجموع المربعات	درجات الحرية	متوسط المربعات	F	مستوى الدلالة	مربع إيتا الجزئي
عدد فئات الاستجابة وترتيبها	2.531	9	0.281	35.125	0.01	0.66
	2.531	1.586	1.595	37.093	0.01	0.66
الخطأ (عدد فئات الاستجابة وترتيبها	1.301	171	0.008			
	1.301	30.136	0.043			

ومن جدول (١١) يتضح أن:

- توجد فروق دالة إحصائياً عند مستوى دلالة ٠.٠١ في دوال معلومات المفردات ترجع لاختلاف عدد فئات الاستجابة وترتيبها، وكانت قيمة مربع إيتا الجزئي تساوي ٠.٦٦ وهو حجم تأثير كبير، مما يؤكد أن الفروق في دوال المعلومات وفقا لتأثير عدد فئات الاستجابة وترتيبها فروق كبيرة.

ـ نتائج السؤال الثالث:

نص السؤال الثالث للدراسة الحالية على "هل تختلف معاملات ثبات المقياس باختلاف عدد فئات الاستجابة وترتيبها؟"

للإجابة عن هذ السؤال تم استخدام الاحصائي W في مقارنة معاملات ثبات المقياس في الحالات المختلفة لعدد فئات الاستجابة وترتيبها، والإحصائي W يتبع توزيع F بدرجات حرية (ن١ -١، ن٢ – ١) حيث ن تمثل عدد الأفراد في حالة معامل الثبات الأول ومعامل الثبات الثاني، وتحسب قيمة W من خلال المعادلة التالية

ISSN: 2537-0464 eISSN: 2537-0472

كما تم توضيحها في (يعقوب وأبوفودة، ٢٠١٢، ص ٤٣٦؛ Feldt & Kim, 2006, p. 590)

$$W = \frac{1 - \propto_1}{1 - \propto_2}$$

حيث α1, α2 هما معاملي الثبات المراد المقارنة بينهما، باعتبار أن α1 هو معامل الثبات الأصغر، ووفقاً لذلك كانت الفروق بين معاملات الثبات ودلالتها الإحصائية كما هو موضح في جدول (١٢):

جدول (١٢) دلالة الفروق بين معاملات ثبات المقياس وفقاً لعدد فئات الاستجابة وترتيبها

معامل الثبات		ثلاثي إيجابي ٠.٨٥٥	ثلاثي سلبي ٠.٨٧٢	رباعي إيجابي ٠.٩١٠	رباعي سلبي ٠.٩١٨	خماسي إيجابي ٠.٩١١	خماسي سلبي ٠.٩٢١	سداسي إيجابي ٠.٩٣٦	سداسي سلبي ٠.٩٤١	سباعي إيجابي ٠.٩٣٩
ثلاثي إيجابي	٠.٨٥٥	×××								
ثلاثي سلبي	٠.٨٧٢	١.١٣٣	×××							
رباعي إيجابي	٠.٩١٠	**١.٦١١	*١.٤٢٢	×××						
رباعي سلبي	٠.٩١٨	**١.٧٦٨	**١.٥٦١	١.٠٩٨	×××					
خماسي إيجابي	٠.٩١١	**١.٦٢٩	*١.٤٣٨	١.٠١١	١.٠٨٥	×××				
خماسي سلبي	٠.٩٢١	**١.٨٣٥	**١.٦٢٠	١.١٣٩	١.٠٣٨	١.١٢٧	×××			
سداسي إيجابي	٠.٩٣٦	**٢.٢٦٦	**٢.٠٠	*١.٤٠٦	١.٢٨١	*١.٣٩١	١.٢٣٤	×××		
سداسي سلبي	٠.٩٤١	**٢.٤٥٨	**٢.١٦٩	*١.٥٢٥	*١.٣٩٠	*١.٥٠٨	١.٣٣٩	١.٠٨٥	×××	
سباعي إيجابي	٠.٩٣٩	**٢.٣٧٧	**٢.٠٩٨	*١.٤٧٥	١.٣٤٤	*١.٤٥٩	١.٢٩٥	١.٠٤٩	١.٠٣٤	×××
سباعي سلبي	٠.٩٤٤	**٢.٥٨٩	**٢.٢٨٦	**١.٦٠٧	*١.٤٦٤	**١.٥٨٩	*١.٤١١	١.١٤٣	١.٠٥٤	١.٠٨٩

* الفرق بين معاملي الثبات دال عند مستوى ٠.٠٥، ** الفرق بين معاملي الثبات دال عند مستوى ٠.٠١، قيمة ف الجدولية عند درجة حرية (١٠٣٩، ١٠٣٩) ومستوى دلالة ٠.٠٥، ٠.٠١ تساوي على الترتيب ١.٣٥٢، ١.٥٣٣

يتضح من جدول (١٢) أن:

- معاملات الثبات تزداد بزيادة عدد فئات الاستجابة، حيث كانت أدنى قيمة لمعامل الثبات عند التدريج الثلاثي، بينما كانت أعلى قيمة لمعامل الثبات عند التدريج السباعي، خاصة في حالة الترتيب السلبي، وهذا يشير إلى أن زيادة عدد فئات الاستجابة تعزز الثبات، مما يعني أن استخدام مقياس ذي عدد فئات أكبر يوفر تقديراً أكثر استقراراً ومتسقاً للسمة المقاسة.
- معاملات الثبات في حالة الترتيب التصاعدي (من السلبي إلى الإيجابي) أعلى منها في حالة الترتيب التنازلي (من الإيجابي للسلبي) عند نفس عدد فئات الاستجابة، هذه

ISSN: 2537-0464 eISSN: 2537-0472

النتيجة قد تعكس تأثير التحيز الإدراكي للمستجيبين تجاه التدريج التصاعدي (من السلبي إلى الإيجابي)، مما يجعل استجاباتهم أكثر اتساقًا عند استخدام التدريج التصاعدي.

ويتأكد مما سبق أن أعلى قيمة لمعامل الثبات كانت في حالة التدريج السباعي مع الترتيب التصاعدي، وكانت هناك فروق دالة عند مستوى ٠.٠١ بين هذا المعامل ومعاملات الثبات في حالة جميع فئات الاستجابة الأقل، وفي المقابل، أدنى قيمة لمعامل الثبات كانت في حالة التدريج الثلاثي مع الترتيب التنازلي. والنتائج في مجملها هنا تؤكد أنه كلما زاد عدد فئات الاستجابة مع الترتيب التصاعدي، زادت دقة وثبات القياس، وعلى الأقل يمكن استخدام التدريج الخماسي مع الترتيب التصاعدي (من السلبي إلى الإيجابي)، ويفضل زيادة عدد فئات الاستجابة إلى ٦ أو ٧ استجابات مع الترتيب التصاعدي إذا كان الهدف هو تحقيق ثبات مرتفع، ولكن قد يعتمد ذلك على طبيعة الظاهرة المقاسة وطريقة تفكير المستجيبين وتمييزهم بين فئات الاستجابة، وهو ما يتضح من الشكل (١٢) التالي:

- نتائج السؤال الرابع:

نص السؤال الرابع للدراسة الحالية على "هل تختلف تقديرات قدرات الأفراد والأخطاء المعيارية في تقديرها باختلاف عدد فئات الاستجابة وترتيبها؟".

للإجابة عن هذ السؤال تم حساب المتوسطات والانحرافات المعيارية لقيم معالم (قدرات) الأفراد والأخطاء المعيارية في تقديرها وهو ما يتضح في جدول (١٢)، ثم الكشف عن دلالة الفروق بينها وفقا لاختلاف عدد فئات الاستجابة وترتيبها باستخدام تحليل التباين للقياسات المتكررة وهو ما يتضح في جدول (١٣):

جدول (١٣) المتوسطات الحسابية والانحرافات المعيارية لقدرات الافراد والأخطاء المعيارية في تقديرها وفقا لعدد فئات الاستجابة وترتيبيها

عدد فئات الاستجابة وترتيبها	قدرات الأفراد		الأخطاء المعيارية	
	المتوسط الحسابي	الانحراف المعياري	المتوسط الحسابي	الانحراف المعياري
ثلاثي إيجابي	1.150	0.948	0.443	0.052
ثلاثي سلبي	0.974	1.015	0.437	0.051
رباعي إيجابي	1.122	0.924	0.402	0.051
رباعي سلبي	0.958	0.977	0.397	0.051
خماسي إيجابي	1.204	1.028	0.407	0.047
خماسي سلبي	0.987	1.106	0.402	0.048
سداسي إيجابي	0.914	0.887	0.365	0.044
سداسي سلبي	0.760	0.916	0.361	0.043
سباعي إيجابي	0.967	0.957	0.369	0.043
سباعي سلبي	0.791	0.994	0.366	0.042

ISSN: 2537-0464 eISSN: 2537-0472

تشير نتائج الجدول (١٣) وبصفة عامة إلى انخفاض المتوسط الحسابي لقدرات الأفراد بزيادة عدد فئات الاستجابة، مع الترتيب التصاعدي لفئات الاستجابة (من السلبي إلى الإيجابي) في جميع الحالات، ففي حالة نفس عدد فئات الاستجابة كانت قيمة متوسط القدرات أقل في حالة الترتيب التصاعدي عنها في حالة الترتيب التنازلي، كما أن التدريج السباعي التنازلي (الإيجابي) يسجل متوسطًا أعلى مقارنة بالتدريج السباعي التصاعدي (السلبي).

بينما في حالة الأخطاء المعيارية لتقدير قدرات الأفراد فنلاحظ أن الأخطاء المعيارية تقل تدريجياً مع زيادة عدد فئات الاستجابة، حيث كانت أعلى قيمة للأخطاء المعيارية في حالة التدريج الثلاثي والترتيب التنازلي، بينما أقل قيمة للإخطاء المعيارية كانت في حالة التدريج السباعي والترتيب التصاعدي، وهو ما يؤكد أن زيادة عدد فئات الاستجابة يقلل من الأخطاء المعيارية في تقدير القدرات، كما يُلاحظ أن الترتيب التنازلي عموماً يؤدي إلى زيادة في الأخطاء المعيارية مقارنة بالترتيب التصاعدي عند نفس عدد فئات الاستجابة، وجدول (١٤) يوضح دلالة الفروق في الأخطاء المعيارية لتقدير القدرات وفقاً لعدد فئات الاستجابة وترتيبها بالاعتماد على تصحيح جرينهاوس-جيسر Greenhouse-Geisser لدرجات الحرية:

جدول (١٤) جدول تحليل التباين ذو القياسات المتكررة للمقارنات داخل المجموعات باختلاف عدد فئات الاستجابة وترتيبها في قدرات الافراد

المتغيرات	مصدر التباين	مجموع المربعات	درجات الحرية	متوسط المربعات	F	مربع إيتا الجزئي
قدرات الأفراد	عدد فئات الاستجابة	196.003	9	21.778	**31.312	0.029
	وترتيبها	196.003	2.293	85.496	**31.312	0.029
	الخطأ	6503.754	9351	0.696		
		6503.754	2381.951	2.730		
الأخطاء المعيارية	عدد فئات الاستجابة	8.216	9	0.913	**516.531	0.332
	وترتيبها	8.216	3.290	2.498	**516.531	0.332
	الخطأ	16.527	9351	0.002		
		16.527	3417.917	0.005		

يتضح من جدول (١٤) أنه:

- توجد فروق دالة إحصائياً عند مستوى دلالة ٠.٠١ في قدرات الأفراد ترجع لاختلاف عدد فئات الاستجابة وترتيبها، وكانت قيمة مربع إيتا الجزئي تساوي ٠.٠٢٩ وهو حجم تأثير ضعيف، مما يؤكد أن الفروق في قدرات الأفراد وفقا لتأثير عدد فئات الاستجابة وترتيبها فروق ضعيفة.

ISSN: 2537-0464 eISSN: 2537-0472

- توجد فروق دالة إحصائياً عند مستوى دلالة ٠.٠١ في الأخطاء المعيارية في تقدير قدرات الأفراد ترجع لاختلاف عدد فئات الاستجابة وترتيبها، وكانت قيمة مربع إيتا الجزئي تساوي ٠.٣٣٢ وهو حجم تأثير كبير، مما يؤكد أن الفروق في الأخطاء المعيارية في تقدير قدرات الأفراد وفقا لتأثير عدد فئات الاستجابة وترتيبها فروق كبيرة.

تفسير ومناقشة النتائج:

تؤكد نتائج السؤال الأول أن معلم التمييز للمفردات يتأثر بدرجة ضعيفة باختلاف عدد فئات الاستجابة وترتيبها، وبينما تتأثر الأخطاء المعيارية في تقدير معلم التمييز بدرجة كبيرة بعدد فئات الاستجابة وترتيبها، حيث تقل الأخطاء المعيارية في تقدير معلم التمييز بزيادة عدد فئات الاستجابة والترتيب التصاعدي لها، وكانت أقل الأخطاء المعيارية في حالة التدريج السباعي والترتيب التصاعدي لفئات الاستجابة.

ونفس الأمر بالنسبة لمعلم العتبات حيث تزداد متوسطات العتبات بزيادة عدد فئات الاستجابة مع الترتيب السلبي لها، أما فيما يتعلق بالأخطاء المعيارية في تقدير العتبات، فقد أظهرت النتائج عدم تأثر الأخطاء المعيارية في تقدير العتبات بعدد فئات الاستجابة وترتيبها، وإن كانت الأفضلية لعدد فئات الاستجابة الأكبر والترتيب السلبي.

وجاءت نتائج السؤال الثاني والمتعلقة بدوال المعلومات الخاصة بالمفردات، لتؤكد النتائج التي تم التوصل إليها في السؤال الأول، حيث تؤكد النتائج على أن دوال المعلومات تزيد بزيادة عدد فئات الاستجابة مع الترتيب التصاعدي لها خاصة في حالة التدريج الخماسي والسداسي والسباعي، مما يعني أن الاعتماد على عدد أكبر لفئات الاستجابة مع الترتيب السلبي يسهم في زيادة دقة التقدير.

وفيما يتعلق بمعاملات الثبات أكدت نتائج السؤال الثالث أن معاملات الثبات تزداد بزيادة عدد فئات الاستجابة مع الترتيب التصاعدي لها، حيث كانت أدنى قيمة لمعامل الثبات في حالة التدريج الثلاثي بينما كانت أعلى قيمة لمعامل الثبات في حالة التدريج السباعي والترتيب السلبي لفئات الاستجابة.

وجاءت نتائج السؤال الرابع على قدر كبير جداً من الأهمية التطبيقية وكذلك تفسر نوعاً ما النتائج التي تم التوصل إليها في الأسئلة السابقة، حيث اتضح أن أعلى قيم لقدرات الأفراد كانت في حالة التدريج الخماسي والرباعي والثلاثي والترتيب التنازلي (تبدأ الاستجابات بالاستجابة موافق بشدة) بينما أقلها في حالة التدريج السداسي والسباعي والترتيب التصاعدي (تبدأ الاستجابات بالاستجابة غير موافق بشدة) وبالتالي فالترتيب التنازلي (الإيجابي) يؤدي إلى تضخم في تقدير القدرات بعكس الترتيب التصاعدي (السلبي)، مع الوضع في الاعتبار الأخطاء المعيارية في

ISSN: 2537-0464 eISSN: 2537-0472

تقدير القدرات والتي كانت أقل ما يمكن في حالة التدريج السباعي والسداسي والترتيب السلبي، يتأكد أن الترتيب التنازلي تزداد معه الأخطاء المعيارية في تقدير القدرات، والذي يمكن تفسيره بأن المفحوص يميل في حالة الترتيب التنازلي إلي اختيار فئة الاستجابة الأكثر إيجابية وهو ما يسهم في تضخم مستوى القدرة، بعكس الترتيب التصاعدي، فعندما تبدأ الاستجابات بالاستجابة السلبية يدفع ذلك المفحوص للتدقيق في باقي الاستجابات واختيار الاستجابة الأكثر تعبيراً عن المستوى الفعلي الذي يمتلكه من السمه.

بناءً على مجمل هذه النتائج، يمكن اقتراح استخدام عدد فئات أكبر (٥-٧ فئات) مع الترتيب التصاعدي (تبدأ الاستجابات بغير موافق) للحصول على تقديرات أكثر دقة وأقل تبايناً في نماذج الاستجابة للمفردة، مع مراعاة قدرة المفحوصين على التمييز بين الفئات المختلفة للاستجابة، كما يمكن التوصية بأن اختيار المقياس الخماسي يوفر حلاً مثالياً في معظم التطبيقات، حيث يوازن بين الدقة في التقدير والسهولة في الاستجابة.

فقد اشارت نتائج الدراسة الحالية إلى أن زيادة عدد فئات الاستجابة تساهم في تحسين دقة تقدير معالم الأفراد والمفردات، حيث لوحظ انخفاض تدريجي في الأخطاء المعيارية مع زيادة عدد الفئات حتى يصل إلى التدريج السباعي، الذي تبين أنه العدد الأمثل.

بناءً على هذه الأدلة، يمكن تفسير نتائج الدراسة الحالية بأن العدد الأمثل لفئات الاستجابة يعتمد على التوازن بين تحسين دقة التقدير وتقليل الأخطاء المعيارية، حيث يؤدي استخدام سبع فئات إلى تحقيق أعلى مستويات الدقة والاستقرار، بينما يمكن أن يؤدي استخدام عدد أقل إلى تحيز في التقديرات، واستخدام عدد أكبر إلى عدم تحسن إضافي في الخصائص السيكومترية.

تشير نتائج الدراسة الحالية ايضا إلى أن الترتيب السلبي لفئات الاستجابة يساهم في تقديم تقديرات أكثر دقة، فعندما تُرتب فئات الاستجابة ترتيباً سلبياً ، فإن هذا الترتيب يعمل كمضاد لتأثير الأسبقية ، وبعبارة أخرى، يعمل الترتيب السلبي كآلية للتحفيز المعرفي تجبر المستجيب على استكمال جميع المراحل الأربع للمعالجة المعرفية كما أشار إليها كونز - (Kunz, 2015) فهم السؤال، استخلاص المعلومات، تحديد الحكم، وتنفيذ الاستجابة - بدلاً من الاكتفاء بالمراحل الأولية فقط. هذا يقلل من "السلوك المُرضي" ويعزز "السلوك الأمثل" في الاستجابة .

من جهة أخرى، أظهرت دراسة Robie et al. (2022) أن تأثير ترتيب الفئات على الخصائص السيكومترية لمقياس ليكرت كان محدودًا أو غير مؤثر في بعض الحالات، مما يشير إلى أن تأثير الترتيب قد يكون نسبيًا ويعتمد على طبيعة

ISSN: 2537-0464 eISSN: 2537-0472

الأداة والمحتوى الذي يتم قياسه. ومع ذلك، كشفت دراسة Kato (2023) أن الترتيب التصاعدي يميل إلى رفع التقييمات مقارنة بالترتيب التنازلي، وهو ما يتماشى مع نتائج الدراسة الحالية التي أشارت إلى أن الترتيب السلبي قد يكون أكثر دقة، حيث إنه قد يمنع المستجيبين من الميل إلى إعطاء تقديرات مرتفعة لمجرد تأثير الترتيب الإدراكي للفئات.

التوصيات والمقترحات:

١. استخدام خمس أو ست أو سبع فئات استجابة عند إعداد أدوات القياس، مع الترتيب التصاعدي لفئات الاستجابة حيث يوفر ذلك دقة في تقدير معالم المفردات ومعالم الأفراد، كذلك يسهم في تحسين الثبات، ودوال المعلومات للمفردات.

٢. تجنب استخدام عدد أقل من فئات الاستجابة (ثلاثة، أربعة) عند إعداد أدوات القياس خاصة مع الأفراد الأكبر سناً كطلاب الجامعة، حيث إن ذلك يؤدي إلى زيادة الأخطاء المعيارية في التقدير، ويؤثر على الثبات، ودقة التقدير.

دراسات مستقبلية مقترحة:

١- دراسة تأثير عدد فئات الاستجابة على دقة تقدير معلمات الأفراد باستخدام نماذج مختلفة من نظرية الاستجابة للمفردة.

٢- مقارنة تأثير ترتيب فئات الاستجابة (التنازلي مقابل التصاعدي) على استجابات الأفراد في أساليب التقييم المختلفة (التطبيق الورقي مقابل التطبيق الإلكترونية).

٣- دراسة تأثير عدد فئات الاستجابة وترتيبها على الصدق البنائي لأدوات القياس بالاعتماد على التحليل العاملي التوكيدي.

٤- دراسة تأثير عدد فئات الاستجابة وترتيبها وفقا لدالة المعلومات والدقة في تقدير القدرات لدى فئات عمرية تختلف عن الفئة العمرية التي تم التطبيق عليها في الدراسة الحالية.

٥- دراسة تأثير التفاعل بين عدد فئات الاستجابة وترتيبها على تحيز الاستجابة والاستجابات النمطية.

ISSN: 2537-0464 eISSN: 2537-0472

قائمة المراجع:

أبوفودة، باسل خميس سالم. (٢٠١٤). أثر إعادة ترتيب بدائل الاستجابة في صعوبة فقرة من متعدد. *مجلة رابطة التربويين العرب*، ٥٣، ٢٦٥ – ٢٨٧.

إسماعيل، محمد عبدالرحمن. (٢٠١٨). أثر استخدام نقطة المنتصف "محايد" في مقياس ليكرت في الخصائص السيكومترية للمقياس وقياس الاتجاهات. *مجلة الإدارة العامة الصادرة عن معهد الإدارة العامة بالمملكة العربية السعودية*، ٥٩ (٣)، ٥٨٧-٦٤٢.

الثبيتي، عمر عواض عوض. (٢٠٢١). استخدام نموذج الاستجابة المتدرجة في نظرية الاستجابة للفقرة في بناء مقياس الاستعداد الأكاديمي للطالب الجامعي. *مجلة العلوم التربوية والنفسية، جامعة البحرين*، ٢٢ (٤)، ٢٠٩-٢٤١

الخرشة، هناء عوض. (٢٠١٥). أثر ترتيب فقرات الاختبار على دقة تقدير معالم الفقرات وقدرات الأفراد والخصائص السيكومترية للاختبار. *رسالة ماجستير غير منشورة*. جامعة مؤتة، الأردن.

الدعيس، محمد حاتم سعيد. (٢٠١٨). فاعلية استخدام نموذج سلم التقدير في تحليل فقرات مقياس العوامل الخمسة الكبرى للشخصية لدى طلبة جامعة تعز. *مجلة العلوم التربوية والدارسات الإنسانية*، (٤)، ٦٢-٩٢.

الزهراني، سعيد عبدالله حسين. (٢٠٢١). أثر التدريج والمرحلة الدراسية على الخصائص السيكومترية لمقياس اتجاه الطلبة نحو الرياضيات: دراسة ميدانية. *مجلة كلية الدراسات العليا للتربية بجامعة القاهرة*، ٢٩ (٢)، ٤٢٧-٤٥٤.

العمراني، عهود غازي، والمالكي، ذياب بن عايض. (٢٠٢٣). أثر ترتيب فئات التدريج ونوع النموذج المستخدم على دقة تقدير بارامترات الفقرات والأفراد. *مجلة كلية التربية جامعة طنطا*، ٨٩(٣)، ١١٢٠-١١٦٧.

القحطاني، نائلة معوض. (٢٠٢١). تأثير مستوى الاكتئاب وعدد بدائل الاستجابة على الخصائص السيكومترية لمقياس الشعور بالوحدة النفسية. *مجلة العلوم التربوية*، ٧(١)، ٢٨٩-٣٢٤.

الموسوي، نعمان محمد صالح. (٢٠١٨). توظيف نموذج راش في تطوير أداة لقياس جودة الحياة لدى طلاب المرحلة الثانوية في مملكة البحرين. *مجلة الطفولة العربية*، (٧٠)، ٤٣-٦١.

النجار، نبيل. (٢٠١٠). *القياس والتقويم منظور تطبيقي مع تطبيقات برمجية (SPSS)*. دار الثقافة.

ISSN: 2537-0464 eISSN: 2537-0472

النعيمي، عز الدين عبدالله عواد. (٢٠٠٦). أثر انتهاك الاستقلال الموضعي على التقديرات المختلفة لنظرية استجابة الفقرة. *رسالة دكتوراه غير منشورة*، جامعة اليرموك، الأردن.

حرز الله، عليه محمد. (٢٠٠٤). بناء بنك أسئلة في الرياضيات والتحقق من فاعليته في انتقاء مفردات اختبار محكي المرجع في مستوى امتحان شهادة الثانوية العامة في الأردن. رسالة دكتوراه غير منشورة]، جامعة عمان العربية، الأردن.

حسين، محمد حسين سعيد. (٢٠٠٧). أثر اتجاه المفردة وعدد بدائل الاستجابة على ثبات أدوات القياس من نوع "ليكرت". *المجلة المصرية للدراسات النفسية*، ١٧ (٥٦)، ٣٥٣-٣٩٢.

حميدة، إبراهيم عبدالرحيم إبراهيم. (٢٠٢٠). مدى اختلاف الخصائص السيكومترية لأداة القياس في ضوء تباين عدد بدائل الاستجابة: حالة مقياس القلق الرقمي. *المجلة العربية لعلم النفس*، ٥ (٢)، ٧٨ – ٩٨.

خلف، مصطفى علي. (٢٠١٩). استخدام نظرية الاستجابة للمفردة الاختبارية في بناء مقياس الاتجاه نحو مهنة التدريس. *مجلة البحث في التربية وعلم النفس بكلية التربية جامعة المنيا*، ٣٤ (١)، ١-٤٧ .

زارع، نسرين محمد سعيد. (٢٠٢٢). تأثير اختلاف طريقة التحليل الإحصائي تبعا لنوع البيانات الفترية الرئيسية والخصائص السيكومترية لمقياس ليكرت: دراسة محاكاة. *مجلة العلوم التربوية والدراسات الإنسانية جامعة تعز فرع التربية*، ٢٢، ٢٣٨-٢٦٦.

عبد السميع، محمد عبد الهادي. (٢٠٢٠). تأثير عدد فئات الاستجابة على افتراضات ومخرجات التحليل العاملي الاستكشافي والتوكيدي لبنود أدوات القياس في البحوث النفسية. *المجلة التربوية لكلية التربية بسوهاج*، ٧٦، ١١٥٢ – ١٢٢٢.

علام، صلاح الدين محمود. (٢٠١٥). *القياس والتقويم التربوي النفسي أساسياته وتطبيقاته وتوجيهاته المعاصرة*. دار الفكر العربي.

عوض الله، محمد عبدالرحيم محمد. (٢٠٢١). تزييف الاستجابة في استبيانات الشخصية. *مجلة الآداب*، ٣، ٦١-٨٥.

مرشود، محمد. (٢٠٢٢). أثر طريقة ترتيب فقرات الاختبار في تقدير معالم الأفراد والفقرات باستخدام خماسي المعلم البارامتري. *مجلة جامعة النجاح للأبحاث (العلوم الإنسانية)*، ٣٦ (٢)، ٣٦٧- ٤٠٢.

Abulela, M. A., & Khalaf, M. A. (2024). Does the Number of Response Categories Impact Validity Evidence in Self-Report

ISSN: 2537-0464 eISSN: 2537-0472

Measures? A Scoping Review. *Sage journals (SAGE Open)*, 14(1). https://doi.org/10.1177/21582440241230363

Allahyari, E., Jafari, P., & Bagheri Z. A. (2016). Simulation Study to Assess the Effect of the Number of Response Categories on the Power of Ordinal Logistic Regression for Differential Item Functioning Analysis in Rating Scales. *Computational and Mathematical Methods in Medicine*, 2016. doi: 10.1155/2016/5080826.

Asu'n, R. A., Rdz-Navarro, K., & Alvarado, J. M. (2016). Developing multidimensional Likert scales using item factor analysis: The case of four-point items. *Sociological Methods & Research*, 45(1), 109- 133.

Aybek, E. C., & Toraman, C. (2022). How Many Response Categories are Enough for Likert Type Scales? : An Empirical Study Based on Item Response Theory. *International Journal of Assessment Tools in Education*, 9(2), 534-547. Retrieved from https://ijate.net/index.php/ijate/article/view/81

Baghaei, P., & Effatpanah, F. (2024). Nonparametric Kernel Smoothing Item Response Theory Analysis of Likert Items. *Psych*, 6, 236–259. https://doi.org/10.3390/ psych6010015

Baker, F. (2001). *The Basics of Item Response Theory*. (2nd Ed.), The ERIC Clearinghouse on Assessment and Evaluation.

Barker, C., Pistrang, N., & Elliott, R. (2016). *Research Methods in Clinical Psychology: An Introduction for Students and Practitioners*. (3rd ed.), John Wiley & Sons. Ltd. https://doi.org/10.1002/9781119154082[Paper reference 12]

Barnette, B. A. (2010). *Likert scaling*. In Salkind, N. J. (2010). *Encyclopedia of research design (pp. 715–718)*. Sage Publications Inc.

Ben, V., & Nissen, J. (2020). Criteria for collapsing rating scale responses: A case study of the CLASS. *Proceedings of the Physics Education Research Conference (PERC)*, held in Provo, Utah, USA, 24-25 July 2019, edited by Ying Cao, Steven Wolf, and Michael Bennett, ISSN 1539-9028, e-ISSN 2377-2379, 2020, pp. 585-590

ISSN: 2537-0464 **eISSN: 2537-0472**

Bindak, R. (2013). Relationship between randomness and coefficient alpha: A Monte Carlo simulation study. *Journal of Data Analysis and Information Processing*, 1, 13–17. https://doi.org/10.4236/jdaip.2013.12003

Bock, R. D. (1972). Estimating item parameters and latent ability when responses are scored in two or more nominal categories. *Psychometrika*, 37(1, Pt. 1), 29–51. https://doi.org/10.1007/BF02291411

Chakrabartty, S. N. (2023). Optimum number of Response categories. *Current Psychology*, 42(7), 5590-5598. https://doi.org/10.1007/s12144-021-01866-6

Chang, L. (1994). A psychometric evaluation of 4-point and 6-point Likert-type scales in relation to reliability and validity. *Applied Psychological Measurement*, 18(3), 205-215. http://dx.doi.org/10.1177/014662169401800302

Chen, X., Yu, H., & Yu, F. (2015). What is the optimal number of response alternatives for rating scales? From an information processing perspective. *J Market Anal*,3, 69–78. https://doi.org/10.1057/jma.2015.4

Christian, L. M., Parsons, N. L., & Dillman, D. A. (2009). Designing scalar questions for web surveys. *Sociological Methods & Research*, 37(3), 393–425.

Chyung, S.Y., Kennedy, M., & Campbell, I. (2018). Evidence-Based Survey Design: The Use of Ascending or Descending Order of Likert-Type Response Options. *Performance Improvement*, 57(9), 9-16

Comrey, A. L., & Montang, I. (1982). Comparison of factor analytic results with two choice and seven choice personality item formats. *Applied Psychological Measurement*, 6(3), 285-289. http://dx.doi.org/10.1177/014662168200600304

Danner, D., Blasius, J., Breyer, B., Eifler, S., Menold, N., Paulhus, D. L., Ziegler,. (2016). Current challenges, new developments, and future directions in scale construction. *European Journal of Psychological Assessment*, 32(3),175-180

ISSN: 2537-0464 **eISSN: 2537-0472**

De Ayala, R. J. (2013). The theory and practice of item response theory. Guilford Publications.

de Winter J. C. F., & Dodou, D. (2010). Five-point Likert items: t-Test versus Mann-Whitney-Wilcoxon. *Practical Assessment, Research and Evaluation*, 15(11), 1–16.

Deutsch, B., & Strack, F. (2019). *Effects of category Range on Reported Behavior and comparative Judgments.* University of Melbourne Library

DeVellis, R. F. (2017). *Scale Development: Theory and Applications*. (4th Ed.), SAGE Publications.

Diener, E., Inglehart, R., & Tay, L. (2012). Theory and Validity of Life Satisfaction Scales. *Social Indicators Research*. 112(3): 497-527. doi: 10.1007/s11205-012-0076-y

Dillman, D. A., Smyth J. D., & Christian L. M. (2014). *Internet, phone, mail, and mixed-mode surveys: The tailored design method*. (4th ed.). Wiley.

Dimitrov, D. M. (2012). *Statistical Methods for Validation of Assessment Scale Data in Counseling and Related Fields*. American Counseling Association.

Dunn-Rankin, P., Knezek, G.A., Wallace, S., & Zhang, S. (2004). *Scaling methods.* Lawrence Erlbaum Associates, Inc.

Epley, N., & Gilovich, T. (2006). The anchoring-and-adjustment heuristic: Why the adjustments are insufficient. *Psychological Science*, 17(4), 311-318

Furr, R. M. (2011). *Scale Construction and Psychometrics for Social and Personality Psychology*. Sage Publications. https://doi.org/10.4135/9781446287866

Gehlbach, H., & Scott, B. (2012). Anchoring and Adjusting in Questionnaire Responses. *Basic and Applied Social Psychology*, 34 (5), 417-433.

Gummer, T., & Kunz, T. (2021). Using only numeric labels instead of verbal labels: Stripping rating scales to their bare minimum in web surveys. *Social Science Computer Review*, 39(5), 1003–1029. https://doi.org/10.1177/0894439320951765

ISSN: 2537-0464 **eISSN: 2537-0472**

Gyamfi, A., & Acquaye, R. (2023). Parameters and Models of Item Response Theory (IRT): A Review of Literature. *Acta Educationis Generalis*, 13(3):68-78, DOI: 10.2478/atd-2023-0022

Harkness, J., & Schoua-Glusberg A. (1998) *Questionnaires in translation*. 3, 87-126.

Harrison, D.A., & McLaughlin, M.E. (1993). Cognitive processes in self-report responses: Tests of item context effects in work attitude measures. *Journal of Applied Psychology*. 78(1): 129-140. doi: 10.1037/0021-9010.78.1.129

Hartley J. (2014). Some thoughts on Likert-type scales. *International Journal of Clinical and Health Psychology*, 14(1), 83–86. https://doi.org/10.1016/s1697-2600(14)70040-7

Hofmans, J., Theuns, P., Baekelandt, S., Mairesse, O., Schillewaert, N., & Cools, W. (2007). Bias and changes in perceived intensity of verbal qualifiers effected by scale orientation. *Survey Research Methods*, 1(2), 97– 108.

Höhne, J., & Lenzner, T. (2015). Investigating response order effects in web surveys using eye tracking. *Psihologija*, 48(4): 361-377. doi: 10.2298/psi1504361h

Höhne, J.K., & Krebs, D. (2017). Scale direction effects in agree/disagree and item-specific questions: a comparison of question formats. International *Journal of Social Research Methodology*. 21(1): 91-103. doi: 10.1080/13645579.2017.1325566

Holbrook, A. (2008). *Response order effects*. In Lavrakas, p.j. (Ed.), Encyclopedia of survey research methods (Vol. 1, pp. 754–756). SAGE. https://doi.org/10.4135/978141 2963947.n488

Holbrook, A. L., Green, M. C., & Krosnick, J. A. (2003). Telephone versus face-to-face interviewing of national probability samples with long questionnaires. *Public Opinion Quarterly*, 67, 79-125

Jamieson, S. (2004). Likert Scales: How to (Ab)use them. Medical Education, 38, 1217-1218. Joshi, A., Kale, S., Chandel, S., & Pal, D.K. (2015). Likert scale: Explored and explained. *British Journal of Applied Science &*

ISSN: 2537-0464 **eISSN: 2537-0472**

Technology (BJAST), 7(4), 396-403. https://doi.org/10.9734/BJAST/2015/14975

Jebb, A. T., Ng, V., & Tay, L. (2021). A review of key Likert scale development advances: 1995-2019. *Frontiers in Psychology*, 12, Article 63747. https://doi.org/10.3389/ fpsyg.2021.637547

Kato, T. (2023). *Descending Beats Ascending: Effects of Order on the Likert Scale on Consumer Ratings*. In Digital Marketing & eCommerce Conference (pp. 20-28), Cham: Springer Nature Switzerland

Kim, K. H. (1998). An analysis of optimum number of response categories for Korean consumers. *Journal of Global Academy of Marketing Science*, 1(1), 61-86. http://dx.doi.org/10.1080/12297119.1998.9707386

King, L. A., King, D., & Klockars, A. J. (1983). Dichotomous and multipoint scales using bipolar adjectives. *Applied Psychological Measurement*, 7(2), 173-180. http://dx.doi.org/10.1177/014662168300700205

Krebs, D. (2012). The impact of response format on attitude measurement. In Salzborn, S., Davidov, E., & Reinecke, J. (Eds.). *Methods, theories, and empirical applications in the social sciences*. Festschrift for Peter Schmidt (pp. 105–113). Wiesbaden: Springer VS

Krebs, D., & Bachner, Y. G. (2018). Effects of rating scale direction under the condition of different reading direction. Methods, Data, Analyses. *Journal for Quantitative Methods and Survey Methodology*, 12(1), 105–125.

Kreitchmann, R. S., Nájera, P., Sanz, S., & Sorrel, M. A. (2024). Enhancing content validity assessment with Item Response Theory modeling. *Psicothema*, 36(2), 145-153. https://doi.org/10.7334/psicothema2023.208

Krosnick, J. A., & Presser, S. (2010). Question and questionnaire design. In P. V. Marsden & J. D. Wrigh (Eds.). H*andbook of survey research* (pp. 263–313). Bingley: Emerald.

ISSN: 2537-0464 **eISSN: 2537-0472**

Kunz, T. (2015). Rating scales in web surveys. A test of new drag-and-drop rating procedures. *Unpublished Doctoral dissertation*, Technische Universität, Darmstadt.

Liu, M., & Keusch, F. (2017). Effects of scale direction on response style of ordinal rating scales. *Journal of Official Statistics*, 33(1), 137–154.

Loken, B., Pirie, P., Virnig, K. A., Hinkle, R. L., & Salmon, C. T. (1987). The use of 0-10 scales in telephone surveys. *Journal of the Market Research Society*, 29(3), 353-362.

Lozano, L. M., Cueto, E., & Muniz J. (2008). Effect of the Number of Response Categories on the Reliability and Validity of Rating Scales. *Methodology*, 4(2), 73–79. DOI 10.1027/1614-2241.4.2.73

Mackinnon, S. P., & Wang, M. (2020). Response-Order Effects for Self-report Questionnaires: Exploring the role of Overclaiming Accuracy and Bias. *Journal of Articles in Support of the Null Hypothesis*, 16(2), 1539–8714.

Mackinnon, S.P., & Firth, S. (2018). The effect of question structure on self-reported drinking: Ascending versus descending order effects. *Journal of Research in Personality*, 73, 21-26. doi: 10.1016/j.jrp.2017.10.004

Mahon-Haft, T., & Dillman, D. A. (2007). Isolating primacy-inducing conditions in web surveys. *Paper presented at the American association for public opinion research annual conference*, Anaheim, CA.

Malhotra N. (2008). Completion Time and Response Order Effects in Web Surveys. *Public Opinion Quarterly*, 72(5), 914-934. doi: 10.1093/poq/nfn050

Malhotra, N. (2008). Completion time and response order effects in web surveys. *Public Opinion Quarterly*, 72(5), 914–934.

Malhotra, N. (2009). Order effects in complex and simple tasks. Public Opinion Quarterly, 73(1), 180–198 Mackinnon, S. P., & Wang, M. (2020). Response-Order Effects for Self-report Questionnaires: Exploring the role of Overclaiming Accuracy and

ISSN: 2537-0464 eISSN: 2537-0472

Bias. *Journal of Articles in Support of the Null Hypothesis*, 16(2)., 113-126.

Maydeu-Olivares A., Kramp U., García-Forero C., Gallardo-Pujol, D., & Coffman, D. (2009). The effect of varying the number of response alternatives in rating scales: Experimental evidence from intraindividual effects. *Behavior Research Methods*, 41(2), 295-308. http://dx.doi.org/10.3758/BRM.41.2.295

Meijer, R. R., & Tendeiro, J. N. (2018). Unidimensional item response theory. In P. Irwing, T. Booth, & D. J. Hugh (Eds.). *The Wiley handbook of psychometric testing: A multidisciplinary reference on survey, scale and test development* (pp. 413-433). Wiley.

Menold, N. (2021). Response bias and reliability in verbal agreement rating scales: Does polarity and verbalization of the middle category matter?. *Social Science Computer Review*, 39(1), 130–147. https://doi.org/10.1177/0894439319847672

Murphy, P. (2021). *Score Interpretation Guide for Educators*. State of New Jersey, Department of Education .

Norman, G. (2010). Likert scales, levels of measurement and the "laws" of statistics. *Adv in Health Sci Educ*, 15, 625-632. https://doi.org/10.1007/s10459-010-9222-y

Pilkonis, P. A., Kim, Y., Yu, L., & Morse, J. Q. (2014). Adult Attachment Ratings (AAR): an item response theory analysis. *J. Pers. Assess*. 96, 417–425. doi: 10.1080/00223891.2013.832261

Preston, C., & Colman, A. (2000). Optimal Number of Response Categories in Rating Scales: Reliability, Validity, Discriminating Power. and Respondent Preferences. *Acta Psychologica*, 104 (1), 1-15. https://doi.org/10.1016/S0001-6918(99)00050-5

Price, L. R. (2016). *Psychometric Methods: Theory into Practice*. The Guilford Press.

Qasem, M., Almoshigah, T., & Gupta, S. (2014). The effect of number of alternatives on validity and reliability in Likert scale. International *journal of innovative research & studies*, 3(6), 324-333. http://dx.doi.org/10.13140/2.1.2237.2803

ISSN: 2537-0464 **eISSN: 2537-0472**

Reeve,B. (2004). *An Introduction To Modern Measurement Theory*. Division Of Cancer Control And Population Science, National Cancer Institute.

Saville, P., & MacIver, R. (2017). A Very Good Question? In B. Cripps (Ed.). *Psychometric Testing: Critical Perspectives* (pp. 29-42). John Wiley & Sons, Ltd.

Spratto, E. M., Leventhal, B. C., & Bandalos, D. L. (2021). Seeing the forest and the trees: Comparison of two IRTree models to investigate the impact of full versus endpoint-only response option labelling. *Educational and Psychological Measurement*, 81(1), 39–60.

Stapleton L. M. (2013). Using multilevel structural equation modeling techniques with complex sample data. In Hancock G. R., Mueller R. O. (Eds), *Structural equation modeling: A second course* (2nd ed., pp. 521–562). Charlotte, NC: Information Age Publishing.

Streiner, D. L., Norman, G. R., & Cairney, J. (2015). *Health Measurement Scales: A Practical Guide to Their Development and Use*. (5th ed.). Oxford University Press.

Sullivan, G.M., & Artino, A.R. (2013). Analyzing and interpreting data from likert-type scales. *J Grad Med Educ*, 5(4), 541-2. doi: 10.4300/JGME-5-4-18.

Tarka, P. (2015). Likert scale and change in range of response categories vs. the factors extraction in EFA model. *Folia Oeconomica*, 1(311), 27-36. http://dx.doi.org/10.18778/0208-6018.311.04

Terentev, E., & Maloshonok, N. (2018). The impact of response options ordering on respondents' answers to rating questions: results of two experiments. *International Journal of Social Research Methodology*, 22(2), 179-198. doi: 10.1080/13645579.2018.1510660

Toepoel, V., Das, M., & Soest, A. (2009). Design of web questionnaires: The effect of layout in rating scales. *Journal of Official Statistics*, 25(4), 509.

ISSN: 2537-0464 **eISSN: 2537-0472**

Tourangeau, R. (2004). Spacing, Position, and Order: Interpretive Heuristics for Visual Features of Survey Questions. *Public Opinion Quarterly*, 68(3): 368-393. doi: 10.1093/poq/nfh035

Tsang, K.K. (2012). The use of midpoint on Likert scale: the implications for educational research. *Hong Kong Teachers' center Journal*, 11, 121-130.

wang, R., & Krosnick, J. (2020). Middle alternatives and measurement validity: a recommendation or survey researchers. *International Journal of Social Research Methodology*, 23(2), 169-184.

Weigold, A., Weigold, I. K., Dykema, S. A., & Drakeford, N. M. (2021). Completing surveys with different item formats: Testing equivalence. *Social Science Computer Review*, 39(6), 1179–1202. https://doi.org/10.1177/0894439320955143

Wells, C. & Faulkner-Bond, M. (2016). *Educational measurement: from foundations to future*, Guilford Press

Weng, L. J. (2004). Impact of the number of response categories and anchor labels on coefficient α test-retest reliability. *Educational and Psychological Measurement,* 64, 956–972. https://doi.org/10.1177/0013164404268674

Wolf M. G., Nylund-Gibson K., Dowdy E., & Furlong M. (2019). An analytic approach for deciding between 4- and 6-point Likert-type response options. https://files.eric.ed.gov/fulltext/ED591440.pdf

Wu, H., & Leung, S.O. (2017). Can Likert Scales be treated as interval scales? A simulation study. *Journal of Social Service Research*, 43(4), 527-532. https://doi.org/10.1080/01488376.2017.1329775

Yan, T., & Keusch, F. (2015). The Effects of the Direction of Rating Scales on Survey Responses in a Telephone Survey. *Public Opinion Quarterly*. 79 (1): 145-165. doi: 10.1093/poq/nfu062

Yang, Y., Timpone, R., Callegaro, M., Hirschorn, M., Achimescu, V., & Natchez, M. (2019). Response Option Order Effects in Cross-Cultural Context. An experimental investigation. *Paper presented at the Conference of the European Association for Survey Research (ESRA),* Zagreb (2019)

ISSN: 2537-0464 **eISSN: 2537-0472**

Zhang, X., & Savalei, V. (2016). Improving the factor structure of psychological scales: The expanded format as an alternative to the Likert scale format. *Educational and Psychological Measurement*, 76(3), 357–386. https://doi.org/10.1177/0013164415596421

Zhao Q, &Linderholm T. (2008) Adult Metacomprehension: Judgment Processes and Accuracy Constraints. *Educational Psychology Review* 20(2): 191-206. doi: 10.1007/s10648-008-9073-8

ISSN: 2537-0464 **eISSN: 2537-0472**