المجلة الأردنية في العلوم التربوية، مجلد 16، عدد 3، 2020، 385-398

أداء مؤشر مطابقة الفرد(l_z^p) في الكشف عن الاستجابات غير الجدية للفقرات متعددة التدريج باختلاف عدد فئات التدريج

ربا جوارنة و محمود القرعان*

Doi: //10.47015/16.3.8

تاريخ تسلم البحث 2019/8/29 تاريخ قبوله 2019/10/27

ملخص: هدفت الدراسة الحالية إلى تقصي أداء المؤشر الإحصائي l_z^p كطريقة للكشف عن الاستجابات غير الجدية في المقاييس غير المعرفية باختلاف عدد فئات التدريج لفقرات ليكرت (ثلاثي، رباعي، خماسي، سباعي). ولتحقيق هدف الدراسة، تم استخدام مقياس التحيزات المعرفية المكون من (42) فقرة بعد إعداد أربع صور للمقياس، تمثل كل منها إحدى فئات التدريج الأربع. وتم تطبيق هذه الصور الأربع على عينة الدراسة الكلية المؤلفة من (1495) طالباً وطالبةً من طلبة البكالوريوس في جامعة اليرموك، المسجلين للفصل الدراسي الثاني من العام الدراسي 2019/2018. ولحساب المؤشر الإحصائي l_z^p لكل مستجيب، تم استخدام نموذج الاستجابة المتدرجة كأحد نماذج نظرية الاستجابة للفقرة. وتبين من نتائج الدراسة الحالية أن توزيع المؤشر l_z^p، بغض النظر عن عدد فئات تدريج الفقرة، ملتوٍ نحو اليسار ولا يتوزع توزيعًا طبيعيًا. كما تبين أنّ خصائص التوزيع للمؤشر الإحصائي l_z^p لا تختلف باختلاف عدد فئات التدريج للفقرة، في حين أنّ نسب الكشف عن الاستجابات غير الجدية المُكتشفة وفق المؤشر l_z^p تختلف باختلاف عدد فئات تدريج الفقرة، وكانت أعلى نسبة استجابات غير جدية كشفها المؤشر هي (19.7%) عند استخدام تدريج ليكرت الرباعي. في حين أنّ أقل نسبة استجابات غير جدية هي (10.7%) عند استخدام تدريج ليكرت الثلاثي.

(الكلمات المفتاحية: الاستجابات غير الجدية، عدد فئات تدريج المقياس، المقاييس غير المعرفية، استجابات من نوع ليكرت)

مقدمة: يفترض الباحثون في التربية وعلم النفس أنّ المشاركين في دراساتهم يمتلكون الدرجة المناسبة من الدافعية والجدية في الاستجابة على أدوات القياس المستخدمة في جمع البيانات، في حين أن بعض الدراسات قد بينت أن بعض المشاركين قد لا يقومون بإعطاء الوقت والجهد الكافيين لاختيار الاستجابة المناسبة على فقرات الاستبانات. فقد يختار بعضهم البدائل بشكل عشوائي (Random Responses)، وقد يستجيب الآخرون وفق نمط معين (Pattern Responses). وتصنف مثل هذه الاستجابات على أنها استجابات غير جدية وتهدد جودة البيانات ودقة النتائج.

وتحدث الاستجابات غير الجدية (Insufficient Effort Responding: IER) عندما لا يولي المستجيب أي انتباه لفقرات المقياس في أثناء استجابته لها (Huang, Curran, Keeney, Popski & Deshon, 2012). وقد سميت هذه الاستجابات بعدة مسميات منها: الاستجابة العشوائية (Random Responding)، والاستجابة غير الواضحة (Inattentive Responding)، وعدم الاهتمام في أثناء الاستجابة (Careless Responding)، والاستجابة غير الجدية (Insufficient Effort Responding IER) (Meade & Craig, 2012).

ويعرف باير وبالنجر وبيري وويتر (Baer, Ballenger, Berry & Wetter,1997) الاستجابات غير الجدية على أنها عدم قدرة المستجيب أو عدم رغبته في الاستجابة عن الفقرات باهتمام. وترتبط أسباب حدوث الاستجابات غير الجدية باهتمام المستجيبين، ودافعيتهم، والمشتتات البيئية، وطول المقياس؛ (Meade & Craig, 2012). وتؤثر الاستجابات غير الجدية على صدق البيانات وجودتها بشكل سلبي؛ إذ تؤثر على الثبات والنتائج المستخلصة من تلك البيانات باستخدام التحليلات الإحصائية المختلفة (McGrath, Mitchell, Kim & Hough, 2010).

The Performance of l_z^p Person-Fit Statistic in Detecting Insufficient Effort Responses According to the Likert Item Scale Points

Ruba Jawarneh & Mahmoud Alquraan, *Yarmouk University, Jordan.*

Abstract: The current study aimed at investigating the performance of l_z^p person-fit statistic in detecting insufficient effort responses in noncognitive scales according to Likert item scale points (3, 4, 5 and 7 points). To achieve the aim of this study, cognitive biases scale (DACOBS) was used to collect the data after creating four forms of the scale to represent the number of item-scale points. The four forms of the scale were administered to a sample that consisted of (1495) undergraduate students at Yarmouk University. IRT graded response model was used to estimate item and ability parameters. Then, l_z^p was calculated for every responder. The results of the current study showed that l_z^p distributions do not change over all Likert item scale points (3, 4, 5 and 7), while being not normally distributed and negatively skewed. Moreover, the results showed that the lowest detected rate of insufficient effort responses (IERs) occurred when using three-point Likert type items (10.7%), whereas the highest detection rate occurred when using four -point Likert type items no neutral option (19.7%).

(**Keywords**: Insufficient Effort Responding, Number of Scale Points, Noncognitive Scales, Likert Scale)

* جامعة اليرموك-الأردن.

© حقوق الطبع محفوظة لجامعة اليرموك، إربد، الأردن.

ويوجد عدد من المسببات التي قد تؤدي إلى وجود الاستجابات غير الجدية؛ فوجود الأسئلة الحساسة في أدوات التقدير الذاتي يؤدي أحيانًا إلى الاستجابة غير الجدية. والفقرات الحساسة هي تلك الأسئلة التي لا يرغب المستجيب بإجابتها، أو يقوم بتزوير الإجابة، أو إظهار عدم الاهتمام تجنبًا للإحراج المجتمعي، لأنها تعكس جزءًا من جوانب خاصة ومظلمة من حياته لا يرغب بإظهارها (Tourangeau & Yan, 2007). كذلك فإن استخدام مصطلحات صعبة أو غير واضحة و وجود اختصارات مبهمة وغير معروفة لدى المستجيبين عند صياغة فقرات أداة القياس، يؤديان إلى زيادة في نسب الاستجابات غير الجدية (Meade & Craig, 2012). ومن جانب آخر، فإن زيادة عدد فقرات أداة القياس من أهم أسباب وجود الاستجابات غير الجدية؛ فهي تتطلب جهدًا كبيرًا ووقتًا طويلاً من المستجيبين للإجابة عنها، مما يؤدي إلى إعطاء إجابات عشوائية أو نمطية أو غير جدية (Rupp, 2013).

ويشير زيجلست وسيجتسما (Zijlst & Sijtsma, 2011) الى أن وجود الاستجابات غير الجدية يؤثر على قيم التحليلات الإحصائية من أوساط حسابية وانحرافات معيارية، الأمر الذي يؤدي إلى التحيز.

وتؤثر مثل هذه الاستجابات أيضًا على الصدق والثبات؛ فعند حساب الصدق المرتبط بمحك، فإن وجود الاستجابات غير الجدية يؤدي إلى تضخم معامل الارتباط بين المتنبئ والمتنبأ به (Huang et al., 2012). كما أن الاستجابات غير الجدية تؤدي إلى وجود عامل جديد عند استخدام التحليل العاملي كمؤشر لصدق الأداة؛ إذ إن البناء العاملي للبيانات بوجود الاستجابات غير الجدية يختلف عنه بعد حذفها (Alquraan, 2019). كما يصعب التنبؤ بأثر الاستجابات غير الجدية على الثبات؛ فقد تؤدي إلى انخفاض معامل الثبات أو ارتفاعه (Meade & Craig, 2012). ووضح كوك وفاوست وماير (Cook, Faust & Meyer, 2016) أن الاستجابات غير الجدية تؤثر سلبًا في تشخيص الأمراض والتقارير النفسية، و بناءً عليها يتم تصنيف الأفراد في العيادات النفسية والمجالات الطبية، وتؤثر بشكل خطير على النتائج المستخلصة من التقييمات والأبحاث الطبية. كما أن الاستجابات غير الجدية تدخل في كل المجالات، سواءً كانت علمية أو طبية أو استراتيجية أو اجتماعية أو نفسية، وغيرها من المجالات التي تعتمد في قراراتها على البيانات التي تم جمعها من أدوات القياس المختلفة. لذلك لا بد من الكشف عن هذه الاستجابات حتى لا تؤثر بشكل سلبي على تلك القرارات.

وهناك نظريتان في القياس يتم استخدامهما للكشف عن الاستجابات غير الجدية، هما: النظرية الحديثة (Item Response Theory: IRT)، والنظرية التقليدية (Classical Test Theory: CTT). وتحتوي كل نظرية على مجموعة من طرق الكشف عن الاستجابات غير الجدية. فبالنسبة للطرق التي تحتويها النظرية الحديثة فهي (Meijer & VanKrimpen-Stoop,1999): طريقة اختبارات مطابقة الفرد (Person-Fit Test: PFT)، وطريقة دالة استجابة الفرد (Person Response Function: PRF)، وإحصائيات مطابقة الفرد (Person Fit Statistics: PFS) التي تشير إلى درجة المطابقة بين النموذج اللوجستي الذي تم اختياره ونمط استجابة الفرد على الفقرة. ويمكن قياس درجة المطابقة بواسطة إحصائي (مؤشر) مطابقة الفرد الذي يوضح مدى استجابة الفرد عن فقرات المقياس وفقًا لافتراضات النموذج المستخدم، ويكون نمط الاستجابة غير مطابق إذا كانت غير محتملة أو غير متوقعة في ظل النموذج المستخدم (Karabatsos, 2003). وفي هذه الدراسة، تم استخدام إحصائيات مطابقة الفرد (PFS) كطريقة للكشف عن الاستجابات غير الجدية، علماً بوجود بعض الطرق التقليدية التي تكشف عن الاستجابات غير الجدية، ومن هذه الطرق: طريقة الفقرات الزائفة (Bogus Items Method)، وسلسلة الإجابات المتكررة (Long String)، والمؤشرات البعدية (Post-hoc indices)، ومقياس مهلانوبويس (Mahalanobois).

واقترح ليفن وروبن (Levine & Rubin,1979) اقتران لوغاريتم الأرجحية العظمى (Function Log-likehood: l_o) في الكشف عن مطابقة الفرد للفقرات ثنائية التدريج الذي يعطى بالمعادلة الآتية:

$$l_o = \sum_{i=1}^{j}[X_j \ln P_j(\hat{\theta}) + (1 + X_j) \ln(1 - P_j(\hat{\theta}))] \quad \ldots\ldots (1)$$

إذ إن:

X_j: استجابة الفرد للفقرة.

j :عدد فقرات الاختبار.

$P_j(\hat{\theta})$: احتمال إجابة الفرد ذي القدرة θ عن الفقرة j إجابة صحيحة.

ولأن التوزيع الإحصائي لـِ l_o غير معروف، ولا توجد نقطة قطع محددة للإحصائي l_o، وأن تصنيف المستجيبين إلى جدي وغير جدي يعتمد على القدرة، قام دراسقو وليفن ووليمز (Drasgow, Levine & Williams, 1985) باقتراح لوغاريتم الأرجحية المعياري (Standardized Log-Likelihood Statistic l_z)، وذلك بتحويل l_o إلى صورة معيارية تشبه العلامة المعيارية الزائية (z-score)، بحيث يعطى بالمعادلة الآتية للفقرات ثنائية التدريج:

$$l_z = \frac{l - E(l_0)}{Var(l_0)^{\frac{1}{2}}} \quad \ldots\ldots\ldots\ldots (2)$$

إذ إن:

E(l_o)): القيمة المتوقعة لـِ l_o.

var (l_o): التباين في l_o.

$$E(l_0) = \sum_{j=1}^{j}\{P_j(\hat{\theta})LnP_j(\hat{\theta}) + [1 - P_j(\hat{\theta})]Ln[1 - Pj(\theta)$$ .....(3)

$$Var(l_0) = \sum_(j = 1)\text{^}j \; [\![P_j(\theta\text{^})[1 - P_j(\theta\text{^})]\{\{Ln(P_j(\theta\text{^}))/(1 - P_j(\theta\text{^}))\}\text{^}2]\!]$$ .....(4)

أما فيما يتعلق بالفقرات متعددة التدريج، فقد قام دراسقو وليفن ووليمز (Drasgow, Levine & Williams, 1985) باشتقاق إحصائي مطابقة الفرد l_z^p للبيانات متعددة الاستجابة؛ إذ بدأ الإحصائي l_o^p بتعريف المتجهات V=(V1,V2,……..,Vj) كمتجه عشوائي لبدائل الاستجابة على مجموعة من الفقرات، بحيث يكون متجه الاستجابة الملاحظ عن الفقرات هو v=(v1,v2,……..,vj)، والصيغة الرياضية للإحصائي l_o^p هي:

$$l_0^p = \sum_{j=1}^{J}\sum_{m=1}^{A}\delta_m(v_j)lnP_{jm}(\hat{\theta})$$ .....(5)

إذ أن:

M: بدائل الاستجابة (m=1,2,3,…….).

p_{jm}: نسبة المستجيبين للبديل m عن الفقرة j.

$\delta_m(v_j)$: تساوي (1) إذا كانت الاستجابة الملاحظة تساوي m، و(0) إذا كانت غير ذلك.

وبما أنّ الإحصائي l_o^p يقوم على استخدام نفس إجراءات الإحصائي l_z نفسها، فإنّ الصورة المعيارية له الخاصة بالفقرات متعددة التدريج تكون على النحو الآتي:

$$l_z^p = \frac{l_o^p - El(l_o^p)}{Var(l_o^p)^{\frac{1}{2}}}$$ .....(6)

إذ إنّ الصيغة الرياضية للقيمة المتوقعة وللتباين هي:

$$E(l_o^p) = \sum_{j=1}^{J}\sum_{m=1}^{A}P_{jm}(\hat{\theta})\ln P_{jm}(\hat{\theta})$$ .....(7)

$$Var(l_o^p) = \sum_{j=1}^{J}\left[\sum_{m=1}^{A}\sum_{k=1}^{A}P_{jm}(\hat{\theta})P_{jk}(\hat{\theta})\ln P_{jm}(\hat{\theta})\ln\left(\frac{P_{jm}(\hat{\theta})}{P_{jk}(\hat{\theta})}\right)\right]$$ (8)

وقد اقترح سنيجرز (Snijders, 2001) تصحيحًا للإحصائي l_z وأصبح يُعرف بالإحصائي l_z^* وذلك لجعل الإحصائي l_z يتوزع توزيعًا طبيعيًا عند استخدام ($\hat{\theta}$)؛ فقد قام سنيجرز بتطبيق هذا التصحيح على الفقرات ثنائية التدريج، وبعد ذلك قام سنهاري (Sinharay, 2015) بتعميم تصحيح Snijders على الإحصائي l_z^p على الفقرات متعددة التدريج.

ويتوزع الإحصائي l_z^p توزيعًا طبيعيًا فقط عندما يتم استخدام القدرة الحقيقية (θ) في المعادلة السابقة، وليس القدرة المقدرة ($\hat{\theta}$) (Molenaar& Hoijtink,1990). وبزيادة مدى عدم مطابقة الفرد، فإنّ قيمة الإحصائي l_z^p تتناقص، والقيمة السالبة للإحصائي l_z^p تشير إلى عدم المطابقة (Sinharay,2015). ومن جهة أخرى، كشفت الدراسات أن الإحصائي l_z^p ربما لا يتوزع توزيعًا طبيعيًا عندما يتم تقدير (θ) من ($\hat{\theta}$) وتحت ظروف مختلفة من خصائص الأداة والمستجيب (Nering, 1995; Schmitt, Chan, Sacco , McFarland & Jennings,1999). وبما أن l_z^p لا يتوزع توزيعًا طبيعيًا، فإن استخدام نقطة القطع (-1.64) التي تُحدد منطقة الرفض للفرضية الصفرية تحت منحنى التوزيع الطبيعي المعياري عند مستوى الدلالة (α=0.05) غير دقيق. لذا تم في الدراسة الحالية استخدام أسلوب إعادة المعاينة (Bootstrapping) لتقدير نقطة القطع على النحو الذي اقترحه سنهاري (Sinharay, 2015)، وتبين من بعض الدراسات أنه أكثر فاعلية وأكثر قدرة على الكشف عن الاستجابات غير الجدية مقارنة مع نقطة القطع (-1.64) التي تعتمد على التوزيع الطبيعي المعياري (Algazo & Alquraan, 2018).

وهناك دراسات عدة حاولت دراسة فاعلية l_z في ضوء عدد من المتغيرات، منها: طول الاختبار (Due &Reise,1991; Noonan, Boss & Gessaroli,1992; Kbar,1994; Meijer & Van, 1999; Snijders, 2001; Deng, 2007; De la Toree & Deng, 2008) وقد أجمعت نتائج هذه الدراسات على أن زيادة طول الاختبار تزيد من قوة الإحصائي في الكشف عن الأنماط غير المطابقة. ودرس دو وريس، ونيرنغ (Due & Reise, 1991; Nering, 1995) فاعلية المؤشر في ضوء مستويات التخمين، وأظهرت النتائج أنه كلما قلّ التخمين، ازدادت نسب الكشف عن الأنماط غير المطابقة للأفراد ذوي القدرات المتدنية، وأن نسب الكشف عن الأنماط غير المطابقة لم تتأثر بوجود التخمين للأفراد ذوي القدرات العليا. وتمت دراسة فاعلية المؤشر في ضوء النموذج اللوجستي المستخدم (Li & Oliejnke, 1997; Pina & Montesions, 2005) استنادًا إلى نموذج راش (Van & Meijer, 1999)، بينما استخدمت دراسات أخرى النموذج ثنائي المعلمات، واستخدمت دراسات أخرى غيرها (Due & Reise, 1991; Snijders, 2001; De la Toree & Deng, 2008; Dodeen & Darabi, 2009) النموذج ثلاثي المعلمات.

وركزت بعض الدراسات على أكثر من مؤشر إحصائي (Li & Oliejnke, 1997; Pina & Montesions, 2005). وقد تمت المقارنة بين خمسة مؤشرات معيارية هي (l_z ,$ECI4_z$,$ECI2_z$,z_u,z_w) باختلاف طول الاختبار وحجم العينة ومدى صعوبة الفقرات. وأظهرت النتائج أن الإحصائي ($ECI4_z$)

كشف عن أعلى نسبة في الأنماط غير المطابقة. وأما ميجر ونيرنغ (Meijer & Nring,1997) فدرسا فاعلية المؤشر في ضوء طريقة تقدير القدرة؛ إذ تناولت الدراسة ثلاث طرق في تقدير القدرة هي (Maximum Likelihood, BIW , Expected A Posterior)، وأظهرت نتائج الدراسة أن طريقة (*EAP*) تجعل تقدير القدرة أقل تحيزًا. وتقصى البعض الآخر من الدراسات فاعلية المؤشر الإحصائي في ضوء بعض المتغيرات الديمغرافية، (Lamprianou & Boyle, 2004) وتناولت متغيرات مثل اللغة والعرق. وتناولت دراسات أخرى (Brown & Villarreal, 2007) الجنس والعرق ومستوى الصف. كما تناولت دراسة الغزو (Alghazo, 2017) الجنس والمستوى الدراسي والكلية والتقدير الجامعي. وأظهرت نتائج تلك الدرسات أن نسب الكشف عن أنماط الاستجابة غير المطابقة تختلف باختلاف المتغيرات الديمغرافية ومستوياتها، وبنسب متفاوتة.

ويلاحظ أنّ عددًا كبيرًا من الدراسات السابقة ركزت على الكشف عن أنماط الاستجابات غير المطابقة في الاستجابات الثنائية (Dichotomous)، وقليل منها ركز على الاستجابات المتعددة (Polytomous). كما أنّ نسب الكشف عن الاستجابات غير الجدية باختلاف تدريج الفقرة للاستجابات المتدرجة لم تحظ بالاهتمام المناسب في الدراسات السابقة، وهذا ما سعت الدراسة الحالية إلى تقصيه.

مشكلة الدراسة وسؤالاها

تم في الدراسة الحالية تقصي أداء الإحصائي (l_z^p) في الكشف عن المستجيبين غير الجديين، بدلالة عدد فئات تدريج المقياس (ثلاثي، رباعي، خماسي، سباعي). وقد بينت نتائج عدد من الدراسات أن مؤشر(l_z^p) للكشف عن أنماط الاستجابة غير المطابقة والاستجابات غير الجدية من الإحصائيات الأكثر فعالية مقارنة بالطرق الأخرى (Nering, 1995; Karabatsos, 2003; Rupp, 2013; Armstrong, Stoumbos, Kung & Shi, 2007). وفي السنوات الاخيرة، تم تقصي فعالية الإحصائي (l_z^p) في ضوء عدد من المتغيرات، منها حجم العينة وعدد فقرات أداة القياس، إلا أن أداء المؤشر(l_z^p) في ضوء عدد فئات تدريج المقياس (ثلاثي، رباعي، خماسي، سباعي) لم ينل القدر الكافي من الدراسة. لذا، تسعى الدراسة الحالية لتقصي أداء المؤشر الإحصائي (l_z^p) تحت ظروف مختلفة من فئات تدريج الفقرة.

وتسعى الدراسة الحالية إلى الإجابة عن السؤالين الآتيين:

1) هل تختلف خصائص التوزيعات الإحصائية للمؤشر (l_z^p) باختلاف عدد فئات تدريج الفقرة (ثلاثي، رباعي، خماسي، سباعي)؟

2) هل تختلف نسب الاستجابات غير الجدية وفق مؤشر(l_z^p) باختلاف عدد فئات تدريج الفقرة (ثلاثي، رباعي، خماسي، سباعي)؟

أهمية الدراسة

تنبع أهمية الدراسة من تناولها لأحد أهم المؤشرات الإحصائية والأكثر انتشارًا واستخدامًا من بين إحصائيات مطابقة الفرد، وهو (l_z^p)، ومعرفة أداء هذا الإحصائي في الكشف عن أنماط الاستجابات غير الجدية باختلاف فئات تدريج الفقرات (ثلاثي، رباعي، خماسي، وسباعي). كما أن استخدام أداة قياس غير معرفية بفقرات متعددة التدريج يضيف أهمية أخرى للدراسة؛ إذ إن معظم الدراسات ركزت على الفقرات ثنائية التدريج (0،1) والمقاييس المعرفية. ومن المتوقع أن تسهم نتائج الدراسة الحالية في مساعدة الباحثين في تحديد عدد فئات التدريج للفقرات من نوع ليكرت الممكن استخدامه في دراساتهم وأبحاثهم لضمان جودة البيانات. فاختيار عدد فئات التدريج الذي يقلل من نسب الاستجابات غير الجدية في البيانات يساعد الباحثين في الحصول على بيانات ذات جودة عالية، وبالتالي يزيد من درجة الثقة بالنتائج التي يتم الحصول عليها بناءً على تلك البيانات.

الطريقة

مجتمع الدراسة

تكون مجتمع الدراسة من جميع طلبة البكالوريوس في جامعة اليرموك المسجلين للفصل الدراسي الثاني من العام الدراسي 2019/2018 والبالغ عددهم (26411) طالبًا وطالبةً بحسب إحصاءات دائرة القبول والتسجيل.

عينة الدراسة

تم اختيار عينة عشوائية عنقودية من طلبة البكالوريوس بجميع مستوياتهم الدراسية في جامعة اليرموك المسجلين للفصل الدراسي الثاني من العام الدراسي 2019/2018، وكانت وحدة الاختيار هي الشعبة الدراسية. بلغ حجم عينة الدراسة (1495) طالبًا وطالبةً، من عشر كليات مختلفة. وبلغ حجم العينة على التدريج الثلاثي (373) طالباً وطالبةً، وعلى التدريج الرباعي (375) طالباً وطالبةً، وعلى التدريج الخماسي(374) طالبًا وطالبةً، وعلى التدريج السباعي (373) طالبًا وطالبةً. وقد تم توزيع الأداة على المستجيبين بشكل عشوائي من خلال ترتيب الأداة بشكل تسلسلي من حيث فئات التدريج للفقرة (ثلاثي، رباعي، خماسي، سباعي). وعند توزيع الأداة على الطلبة، أُعطي الطالب الأول التدريج الثلاثي والثاني التدريج الرباعي والثالث التدريج الخماسي والرابع التدريج السباعي، وتكرر ذلك لجميع الطلبة بهدف ضمان تكافؤ المجموعات.

أداة الدراسة

تم استخدام مقياس دافوس لتقدير التحيزات المعرفية (Davos Assessment of Cognitive Biases Scale) الذي قام ببنائه فان دير غاج وزملاؤه (Van der Gaag et al., 2013)، وكيفهُ للبيئة الأردنية الحموري (Al-Hamouri, 2017)؛ بهدف الكشف عن مستوى التحيزات المعرفية لدى أفراد الدراسة. يتكون المقياس من (42) فقرة من نوع تدريج ليكرت السباعي ضمن ثلاثة مجالات رئيسة، تتفرع منها مجالات فرعية وعلى النحو الآتي:

1) التحيزات المعرفية (Cognitive Biases)، وتضم: (القفز إلى الاستنتاجات (Jumping to Conclusions)، وجمود المعتقدات (Belief Inflexibility)، والانتباه للمهددات (Attention for Threats)، والعزو الخارجي (External Attribution)).

2) المحددات المعرفية (Cognitive Limitations) وتضم: (المشكلات المعرفية الاجتماعية (Social Cognition Problems)، والمشكلات المعرفية الذاتية (Subjective Cognition Problems)).

3) السلوكيات الآمنة (Safety Behaviors).

هذا، ويتم تصحيح المقياس بناءً على درجة الموافقة (موافق بشدة، موافق، موافق نوعًا ما، محايد، غير موافق نوعًا ما، غير موافق، غير موافق بشدة)، بحيث تعطى الاستجابات السابقة الدرجات الآتية على التوالي (7, 6, 5, 4, 3, 2, 1).

صدق أداة الدراسة وثباتها

قام فان دير غاج وزملاؤه (Van der Gaag et al., 2013) بالتحقق من صدق أداة الدراسة بصورتها الأصلية من خلال التحليل العاملي، وقد فسرت مجالات المقياس السبعة (45%) من التباين الكلي للمقياس.

كما قاموا بالتحقق من ثبات أداة الدراسة بثلاث طرق هي: كرونباخ ألفا، والتجزئة النصفية، وطريقة الاختبار-إعادة الاختبار. وتراوحت معاملات الثبات بطريقة كرونباخ ألفا بين (0.64-0.82) للمجالات الفرعية و (0.90) للأداة ككل. في حين تراوحت معاملات الثبات بين (0.70-0.76) للمجالات و (0.92) للأداة ككل وفق طريقة التجزئة النصفية. أما بطريقة الاختبار-إعادة الاختبار فتراوحت معاملات الثبات بين (0.74-0.88) للمجالات و(0.93) للأداة ككل.

صدق الصورة الأردنية لأداة الدراسة وثباتها

قام الحموري (Al-Hamouri, 2017) بمجموعة من الإجراءات للتحقق من صدق أداة الدراسة؛ فقد قام بترجمة المقياس من اللغة الإنجليزية إلى اللغة العربية، ثم عرضه على مجموعة من المحكمين من ذوي الاختصاص لضمان صدق المحتوى.

وتم حساب معاملات الثبات للمجالات الفرعية وللأداة ككل بطريقتي الاتساق الداخلي والاستقرار، كما هو مبين في الجدول (1).

جدول (1): معاملات ثبات الصورة الأردنية لمقياس التحيزات المعرفية

اسم المجال	عدد الفقرات	معامل الاتساق الداخلي	معامل ثبات الاستقرار
القفز إلى الاستنتاجات	6	0.71	0.77
جمود المعتقدات	6	0.74	0.78
الانتباه إلى المهددات	6	0.73	0.77
العزو الخارجي	6	0.71	0.79
مشكلات المعرفة الاجتماعية	6	0.77	0.80
مشكلات المعرفة الذاتية	6	0.69	0.80
السلوكيات الآمنة	6	0.80	0.78
الدرجة الكلية	42	0.87	0.89

وقد بلغت قيمة كرونباخ ألفا للمقياس ككل (0.87) بطريقة الاتساق الداخلي و(0.89) بطريق الاختبار-إعادة الاختبار (الاستقرار). كما تم إعداد صورة جديدة لأداة الدراسة تختلف من حيث التدريج (ثلاثي، رباعي، خماسي، إضافة إلى الصورة الاصلية للمقياس وهي التدريج السباعي). وتم استخراج معاملات الثبات لها وفق النظرية التقليدية في القياس (كرونباخ ألفا) وكذلك الثبات الإمبريقي وفق نظرية الاستجابة للفقرة، الذي يعتمد على دالة معلومات الفقرة والاختبار ويأخذ بعين الاعتبار أن الخطأ المعياري في القياس يختلف باختلاف مستويات القدرة للمستجيبين، بعكس النظرية التقليدية في القياس التي تفترض تساوي الخطأ المعياري للقياس عند مستويات القدرة كافة ، كا هو مبين في الجدول (2).

جدول (2): معاملات ثبات كرونباخ ألفا والثبات الامبريقي للبيانات الحالية

عدد فئات التدريج	معامل كرونباخ ألفا	معامل الثبات الإمبريقي
ثلاثي	0.797	0.85
رباعي	0.818	0.86
خماسي	0.865	0.89
سباعي	0.887	0.91

التحقق من افتراضات النظرية الحديثة (IRT)

المؤشر الاحصائي l_z^p هو أحد تطبيقات نظرية الاستجابة للفقرة، والنموذج المناسب للبيانات هو نموذج الاستجابة المتدرجة (Graded Response Model). لذا تم التحقق من افتراضات النظرية للتمكن من حساب المؤشر (l_z^p)، وعلى النحو الآتي:

أولاً: التحقق من افتراض أحادية البعد

تم التحقق من افتراض أحادية البعد لفقرات المقياس بإعداد ملف SPSS خاص لكل فئة من فئات التدريج على حدةٍ، ثم إجراء تحليل المكونات الأساسية باستخدام (SPSS v16.0)، كما هو مبين في الجدول (3).

جدول (3): نتائج التحليل العاملي لفقرات مقياس التحيزات المعرفية

عدد فئات التدريج	الجذر الكامن للعامل الأول	الجذر الكامن للعامل الثاني	الجذر الكامن للعامل الثالث	النسبة*	النسبة**
ثلاثي	5.346	2.211	2.065	2.418	21.473
رباعي	5.516	2.672	1.932	2.064	3.843
خماسي	6.672	3.139	1.821	2.126	2.681
سباعي	7.962	2.741	1.897	2.905	6.186

النسبة*: ناتج قسمة الجذر الكامن للعامل الأول على الجذر الكامن للعامل الثاني.

النسبة**: ناتج قسمة حاصل طرح الجذر الكامن الثاني من الجذر الكامن الأول على حاصل طرح الجذر الكامن الثالث من الجذر الكامن الثاني.

يلاحظ من الجدول (3) أن التحليل العاملي لفقرات المقياس يقدم دليلاً على تحقق أحادية البعد على جميع فئات التدريج، وذلك على النحو الآتي: ناتج قسمة الجذر الكامن للعامل الأول على الجذر الكامن للعامل الثاني أكبر من(2)، ثم ناتج قسمة حاصل طرح الجذر الكامن الثاني من الجذر الكامن الأول على حاصل طرح الجذر الكامن الثالث من الجذر الكامن الثاني ذو قيمة عالية (Lord, 1980).

ثانيًا: التحقق من افتراض الاستقلال الموضعي لفقرات المقياس

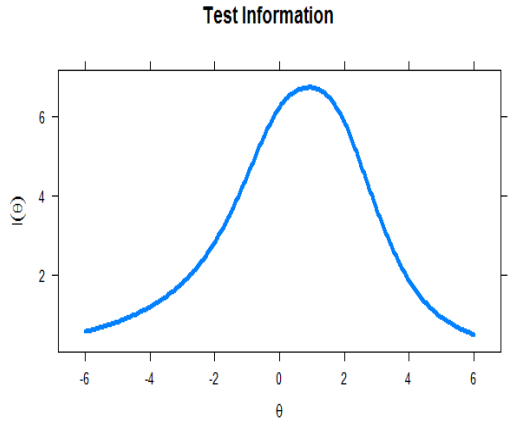
تم التحقق من افتراض الاستقلال الموضعي لفقرات المقياس بحساب قيمة (x^2) للاستقلال الموضعي المعياري (Standardized LD x^2) لكل فئة من فئات التدريج لكل زوج من أزواج فقرات المقياس التي يبلغ عددها (861) زوجًا، وذلك بضرب (42) فقرة في (41)، ثم القسمة على (2) باستخدام برنامج IRTPRO 4.2(x64). وحسب ما ذكره شِن وثيسن (Chen & Thissen, 1997)، فإن أزواج الفقرات التي تزيد قيمتها على (10) تعد انتهاكًا لافتراض الاستقلال الموضعي. وقد تم حساب عدد الأزواج التي تزيد قيمتها على (10) لكل فئة من فئات التدرج، كما هو مبين في الجدول (4).

جدول (4): العدد والنسب المئوية للاستقلال الموضعي لفقرات مقياس التحيزات المعرفية

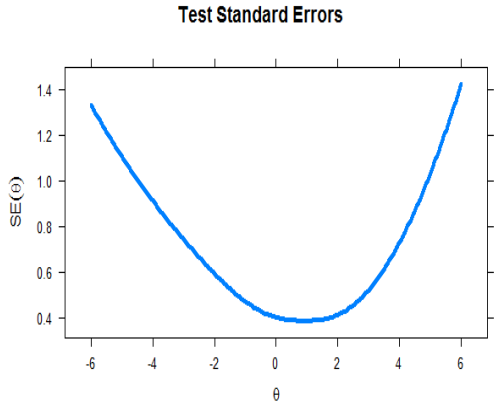
النموذج	معتمدة		مستقلة	
	العدد	النسبة	العدد	النسبة
ثلاثي	0	0.00	861	%100
رباعي	4	%0.047	857	%99.53
خماسي	4	%0.047	857	%99.53
سباعي	2	%0.023	859	%99.77

يلاحظ من الجدول (4) أنّ الاستقلال الموضعي متحقق على جميع فئات التدريج؛ فقد بلغ عدد الأزواج المستقلة على التدريج الثلاثي (861)، أي ما نسبة (100%)، وبلغ عدد الأزواج المستقلة على التدريج الرباعي والخماسي (857)، أي ما نسبته (99.53%)، وبلغ عدد الأزواج المستقلة على التدريج السباعي (859) أي ما نسبة (99.77%).

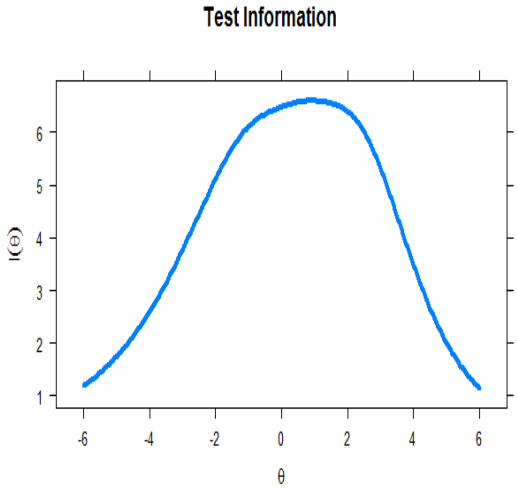
كما تم حساب دالة معلومات الاختبار لكل فئة من فئات التدريج، وكذلك تم حساب الخطأ المعياري للتعرف على مطابقة البيانات لنموذج الاستجابة المتدرجة، والأشكال الآتية توضح ذلك:



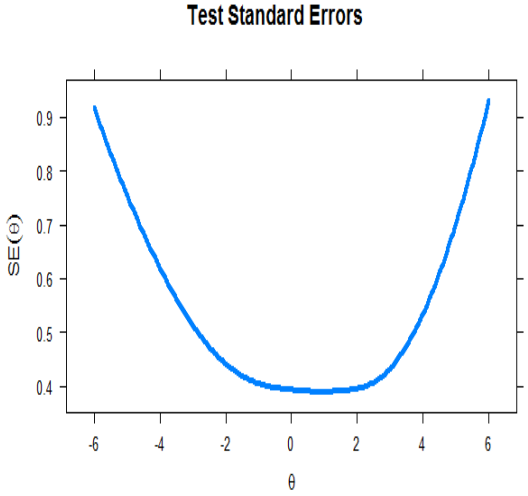
الشكل (1): دالة معلومات الاختبار للتدريج الثلاثي



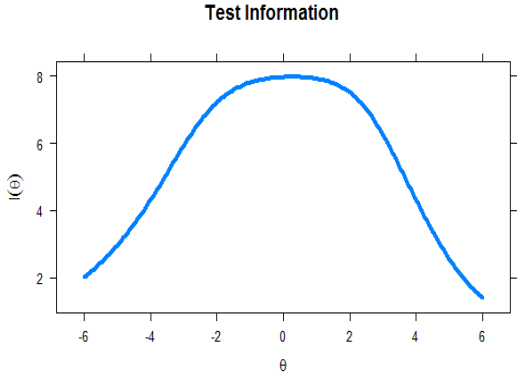
الشكل (2): الخطأ المعياري للتدريج الثلاثي



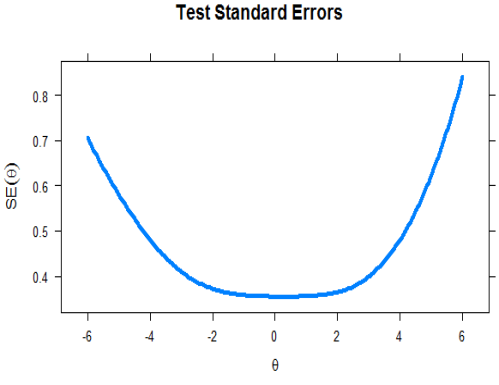
الشكل (3): دالة معلومات الاختبار للتدريج الرباعي



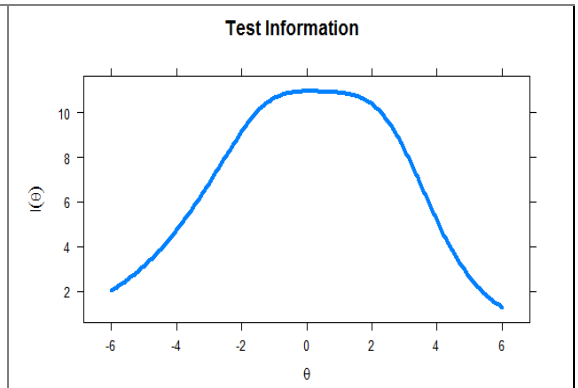
الشكل(4): الخطأ المعياري للتدريج الرباعي



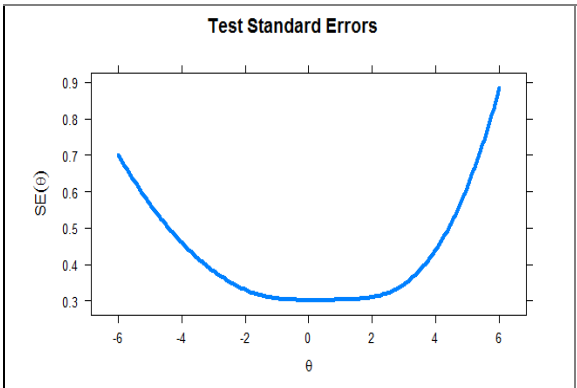
الشكل (5): دالة معلومات الاختبار للتدريج الخماسي



الشكل (6): الخطأ المعياري للتدريج الخماسي



الشكل (7): دالة معلومات الاختبار للتدريج السباعي



الشكل (8): الخطأ المعياري للتدريج السباعي

الإجراءات

لحساب مؤشر مطابقة الفرد الإحصائي (l_z^p)لأنماط الاستجابة غير الجدية ضمن نموذج الاستجابة المتدرجة على مقياس التحيزات المعرفية، تم اتباع الخطوات الآتية باستخدام برمجية (R) والرزمة الخاصة بتقدير مؤشرات المطابقة وحسابها (PerFit):

1- تم إدخال البيانات إلى ذاكرة الحاسوب، ثم إصدار أمر لعرض معالم الفقرات (معلمة التمييز وعتبات لمعلمة الصعوبة) المقدرة بطريقة (EAP).

2- إصدار الأمر الذي يحسب مؤشر مطابقة الفرد(l_z^p).

3- إصدار الأمر الذي يحسب الخطأ المعياري لمؤشر مطابقة الفرد (l_z^p).

4- إصدار أمر لتفعيل العشوائية بقيمة استهلالية قيمتها (512)، علمًا بأنها غير ملزمة؛ بهدف التمكن من حساب درجة القطع للمؤشر الإحصائي (l_z^p) لأنماط الاستجابة التي تتم بعد إعادة التعيين (Resampling) لألف مرة بشكل مسبق الإعداد في رزمة (PerFit).

5- إصدار أمر لحساب درجة القطع للمؤشر الإحصائي (l_z^p)، لأنماط الاستجابة التي تتم بعد إعادة التعيين Resampling لألف مرة بشكل مسبق الإعداد في رزمة (PerFit). ويسمى هذا الاجراء (Bootstrapping)، مع إجراء ألف مستنسخة لكل عملية إعادة تعيين. وينصح باستخدام هذه الطريقة بناءً على الدراسات التي أجريت حول هذا الموضوع (Sinharay, 2015).

6- تم حساب درجة القطع للمؤشر الإحصائي (l_z^p)، واحتمالية الخطأ لها عند مستوى الدلالة (α=0.05)، والخطأ المعياري، وفترة الثقة الخاصة بها، كما هو مبين في الجدول (5).

جدول (5): قيم درجة القطع للمؤشر الإحصائي l_z^p واحتمالية الخطأ لها والخطأ المعياري وفترة الثقة المتعلقة بها تبعًا لعدد فئات تدريج المقياس (ثلاثي، رباعي، خماسي، سباعي)

عدد فئات التدريج	حجم العينة	قيمة درجة القطع	احتمالية الخطأ	الخطأ المعياري	فترة الثقة المئينية لإعادة التعيين لدرجة القطع	
					2.5%	97.5%
ثلاثي	373	-1.351	0.1072	0.0672	-1.5121	-1.2499
رباعي	375	-1.5683	0.1973	0.1069	-1.7124	-1.3409
خماسي	374	-1.5404	0.1684	0.0816	-1.6590	-1.3797
سباعي	373	-1.448	0.1448	0.1019	-1.6673	-1.3128

يلاحظ من الجدول (5) أن فترات الثقة لدرجة القطع للمؤشر الإحصائي (l_z^p) باختلاف عدد فئات التدريج (ثلاثي، رباعي، خماسي، سباعي) كانت متقاطعة. وهذا يدل على عدم وجود فروق دالة إحصائيًا بين درجة القطع للمؤشر الإحصائي (l_z^p) لأنماط الاستجابة باختلاف تدريج المقياس. وقد كانت أصغر قيمة لدرجة القطع للمؤشر الإحصائي (l_z^p) لأنماط الاستجابة على التدريج الرباعي؛ إذ بلغت (-1.5683)، وأكبر قيمة لدرجة القطع للمؤشر الإحصائي (l_z^p) لأنماط الاستجابة على التدريج الثلاثي؛ إذ بلغت (-1.351). ويلاحظ أيضًا أنّ أصغر قيمة للخطأ المعياري لدرجة القطع للمؤشر الإحصائي (l_z^p) كانت لأنماط الاستجابة على التدريج الثلاثي؛ إذ بلغت (0.0672)، وأكبر قيمة للخطأ المعياري لدرجة

القطع للمؤشر الإحصائي (l_z^p) لأنماط الاستجابة على التدريج الرباعي؛ إذ بلغت (0.1069).

7- إصدار أمر لعرض قيم المؤشر الإحصائي (l_z^p) لكل نمط استجابة والخطأ المعياري الخاص بكل منها.

النتائج ومناقشتها

أولاً: النتائج المتعلقة بالاجابة عن السؤال الأول: *هل تختلف خصائص التوزيعات الإحصائية للمؤشر(l_z^p) باختلاف عدد فئات تدريج الفقرة (ثلاثي، رباعي، خماسي، سباعي)؟ومناقشتها.*

للإجابة عن السؤال الأول، تم حساب الإحصاءات الوصفية للمؤشر الإحصائي (l_z^p) لأنماط الاستجابة ضمن نموذج الاستجابة المتدرجة على مقياس التحيزات المعرفية، والتحقق من مطابقته للتوزيع الطبيعي باستخدام اختبار كولمجروف-سميرنوف (Kolmogorov-Smirnov) وذلك تبعًا لعدد فئات تدريج الفقرة، كما هو مبين في الجدول (6).

جدول (6): الإحصاءات الوصفية للتوزيعات العينية للمؤشر الإحصائي l_z^p لأنماط الاستجابة على مقياس التحيزات المعرفية، ونتائج اختبار كولمجروف-سميرنوف (Kolmogorov-Smirnov) وفقًا لعدد فئات التدريج (ثلاثي، رباعي، خماسي، سباعي)

عدد فئات التدريج	الوسط الحسابي	الانحراف المعياري	الخطأ المعياري للوسط	فترة الثقة لوسط المؤشر		الالتواء	التفلطح	K-S Sig.
				2.5%	97.5%			
ثلاثي	0.189	1.220	0.063	0.065	0.314	-0.876	1.236	0.000
رباعي	0.164	2.072	0.107	-4.610	0.374	-0.777	0.659	0.000
خماسي	0.179	1.940	0.100	-1.752	0.377	-0.896	1.423	0.000
سباعي	0.157	1.656	0.085	-1.097	0.326	-0.894	1.182	0.000

يلاحظ من الجدول (6) أن فترات الثقة للوسط الحسابي لقيم المؤشر الإحصائي (l_z^p) الخاصة بأنماط الاستجابة على مقياس التحيزات المعرفية لفئات التدريج الأربع متقاطعة، مما يدل على عدم وجود فروق جوهرية بين الأوساط الحسابية لقيم المؤشر الإحصائي (l_z^p). فقد تراوحت الأوساط الحسابية بين (0.157) على التدريج السباعي و (0.189) على التدريج الثلاثي. ويلاحظ أن الانحرافات المعيارية لقيم المؤشر الإحصائي (l_z^p) الخاصة بأنماط الاستجابة على مقياس التحيزات المعرفية لفئات التدريج الأربع تراوحت بين (1.220) على التدريج الثلاثي و (2.072) على التدريج الرباعي.

وتراوحت الأخطاء المعيارية للأوساط الحسابية لقيم المؤشر الإحصائي (l_z^p) باختلاف عدد فئات التدريج بين (0.063) على التدريج الثلاثي و (0.107) على التدريج الرباعي. مما يدل على فاعلية المؤشر الإحصائي (l_z^p) على التدريج الثلاثي مقارنة بفئات التدريج الأخرى. وبالتالي، فإنّه يعكس موثوقية أعلى في قيمة الوسط الحسابي للمؤشر الإحصائي. ويلاحظ أن الخطأ المعياري للمؤشر الإحصائي (l_z^p) على التدريج الرباعي قد تراجعت فاعليته بالمقارنة مع فئات التدريج الأخرى. وبالتالي، فإنّه يعكس موثوقية أقل في قيمة الوسط الحسابي للمؤشر الإحصائي.

ويلاحظ أن المدى لفترة الثقة على التدريج الرباعي تراوحت بين (4.610-) و (0.374). وهذا يدل على أن فترة الثقة الأكثر اتساعًا كانت على التدريج الرباعي، وأنه لا يتوزع توزيعًا طبيعيًا، بينما كان المدى لفترة الثقة لباقي فئات التدريج متقاربًا ولا يتوزع توزيعًا طبيعيًا.

كما يشير الجدول (6) الى أنّ الالتواء لقيم المؤشر الإحصائي (l_z^p) الخاصة بأنماط الاستجابة على مقياس التحيزات المعرفية باختلاف عدد فئات التدريج تراوحت بين (0.896-) على التدريج الخماسي و (0.777-) على التدريج الرباعي، ما يدل على أن التدريج الخماسي هو الأبعد عن التوزيع الطبيعي والتدريج الرباعي هو الأقرب إلى التوزيع الطبيعي، غير أن أيًا منهما لا يتوزع توزيعًا طبيعيًا. وهذا يتفق مع نتائج عدد من الدراسات السابقة (Noonan et al., 1992)؛ (Meijer & Van Krimpen-Stoop , 1999).

أما بالنسبة لتفلطح قيم المؤشر الإحصائي (l_z^p) الخاصة بأنماط الاستجابة على مقياس التحيزات المعرفية باختلاف عدد فئات التدريج فتراوحت بين (1.423) على التدريج الخماسي و (0.659) على التدريج الرباعي، ما يدل على أنّ التدريج الرباعي هو الأقرب إلى التوزيع الطبيعي، والتدريج الخماسي هو الأبعد عن التوزيع الطبيعي، غير أن أيًا منهما لا يتوزع توزيعًا طبيعيًا. وهذا يتفق مع نتائج دراسات سابقة (Noonan et al., 1992)؛ (Meijer & Van Krimpen-Stoop , 1999).

ويلاحظ أن قيم المؤشر الإحصائي (l_z^p) الخاصة بأنماط الاستجابة على مقياس التحيزات المعرفية باختلاف عدد فئات

التدريج قد اختلفت عن التوزيع الطبيعي بمستوى الدلالة الإحصائية ($\alpha = 0.05$) ضمن اختبار كولمجروف-سميرنوف (Kolmogorov -Smirnov).

وبناءً على نتائج الجدول (6)، فإنّ التوزيع العيني للمؤشر الإحصائي (l_z^p) ملتوٍ نحو اليسار، ولا يتوزع توزيعًا طبيعيًا. وقد اتفقت هذه النتيجة مع أغلب الدراسات التي وضحت أن المؤشر الإحصائي (l_z^p) يميل نحو الالتواء السالب والتفلطح الموجب (De la Toree & Deng, 2008 ; Reise,1995 ; Meijer &Van Krimpen-Stoop,1999). وقد يكون السبب في ذلك أن الدراسة الحالية قد استندت الى القدرة المقدرة للمستجيب ($\hat{\theta}$) وليس إلى القدرة الحقيقية (θ) في حساب قيمة المؤشر الاحصائي (l_z^p). وقد أشارات دراسة مولينار وهوجتنك (Molenaar & Hoijtink, 1990) إلى أن (l_z^p) يتوزع توزيعًا طبيعيًا فقط إذا تم استخدام القدرة الحقيقية للمستجيب وليس القدرة المقدرة.

كذلك تم حساب الإحصاءات الوصفية للخطأ المعياري للمؤشر الإحصائي (l_z^p) لأنماط استجابة الأفراد ضمن نموذج الاستجابة المتدرجة على مقياس التحيزات المعرفية، كما هو مبين في الجدول (7).

جدول (7): الإحصاءات الوصفية للتوزيعات العينية للخطأ المعياري للمؤشر الإحصائي (l_z^p) لانماط الاستجابة على مقياس التحيزات المعرفية، ونتائج اختبار كولمجروف-سميرنوف (Kolmogorov-Smirnov) وفقًا لعدد فئات التدريج (ثلاثي، رباعي، خماسي، سباعي)

عدد فئات التدريج	الوسط الحسابي	الانحراف المعياري	الخطأ المعياري للوسط	فترة الثقة لوسط المؤشر	
				2.5%	97.5%
ثلاثي	0.8521	0.22629	0.01172	0.8291	0.8752
رباعي	0.8984	0.30708	0.01586	0.8672	0.9296
خماسي	0.8957	0.28736	0.01486	0.8665	0.9249
سباعي	0.8926	0.27709	0.01435	0.8644	0.9208

يلاحظ من الجدول (7) أنّ فترات الثقة للوسط الحسابي لقيم الخطأ المعياري للمؤشر الإحصائي (l_z^p) لأنماط الاستجابة حسب عدد فئات التدريج كانت جميعها متقاطعة، مما يدل على عدم وجود فروق جوهرية بين الأوساط الحسابية لقيم الخطأ المعياري للمؤشر الإحصائي (l_z^p) لأنماط الاستجابة، ويدل على أن فاعلية تقدير المؤشر الاحصائي على فئات التدريج وموثوقيته كانتا متكافئتين. ويلاحظ أيضًا أنّ قيمة الخطأ المعياري الأصغر كانت على التدريج الثلاثي.

كما يلاحظ من الجدول (7) أنّ قيم الأوساط الحسابية للخطأ المعياري للمؤشر الإحصائي (l_z^p) لأنماط الاستجابة قد تراوحت بين (0.8521) على التدريج الثلاثي و (0.8984) على التدريج الرباعي. وفيما يتعلق بقيم الانحرافات المعيارية لقيم الخطأ المعياري للمؤشر الإحصائي (l_z^p) لأنماط الاستجابة، فقد تراوحت بين (0.22629) على التدريج الثلاثي و (0.30708) على التدريج الرباعي. كما يشير الجدول (7) إلى أن قيم الأخطاء المعيارية للأوساط الحسابية لقيم الخطأ المعياري للمؤشر الإحصائي (l_z^p) لأنماط الاستجابة تراوحت بين (0.01172) على التدريج الثلاثي و (0.01586) على التدريج الرباعي.

ولمقارنة الأوساط الحسابية لقيم الخطأ المعياري للمؤشر الإحصائي (l_z^p) لأنماط الاستجابة، تم استخدام اختبار تحليل التباين الأحادي (One-Way ANOVA)، كما هو مبين في الجدول (8).

جدول (8): تحليل التباين الأحادي للأوساط الحسابية لقيم الخطأ المعياري للمؤشر الإحصائي (l_z^p) لأنماط الاستجابة

	مجموع المربعات	درجات الحرية	متوسط المربعات	F	الدلالة الإحصائية
بين المجموعات	0.535	3	0.178	2.339	0.072
داخل المجموعات	113.678	1491	0.076		
الكلي	114.213	1494			

يلاحظ من الجدول (8) أن الأوساط الحسابية لقيم الخطأ المعياري للمؤشر الإحصائي (l_z^p) لأنماط الاستجابة غير دالة إحصائيًا عند مستوى الدلالة (0.05)، في حين يلاحظ أنها دالة إحصائيًا عند مستوى الدلالة (0.1) لصالح الوسط الحسابي الأقل، وهو للتدريج الثلاثي. فقد بلغ الوسط الحسابي لقيم الخطأ المعياري للمؤشر الإحصائي (l_z^p) لأنماط الاستجابة على التدريج

الثلاثي (0.8521)، وعلى التداريج الرباعي والخماسي والسباعي كانت القيم قريبة من القيمة (0.89)، ما يشير إلى وجود أخطاء قياس أقل عند استخدام تدريج ليكرت الثلاثي عند مقارنته مع التداريج الرباعي والخماسي والسباعي.

ثانيًا: النتائج المتعلقة بالاجابة عن السؤال الثاني: *هل تختلف نسب الاستجابات غير الجدية وفق مؤشر(l_z^p) باختلاف عدد فئات تدريج الفقرة (ثلاثي، رباعي، خماسي، سباعي)؟ ومناقشتها.*

للإجابة عن السؤال الثاني، تم تصنيف استجابات الأفراد إلى استجابات جدية واستجابات غير جدية بناءً على درجة القطع لكل تدريج وفق مؤشر (l_z^p) ضمن نموذج الاستجابة المتدرجة على مقياس التحيزات المعرفية باختلاف عدد فئات التدريج (ثلاثي، رباعي، خماسي، سباعي)، كما هو مبين في الجدول (9).

جدول (9): تصنيف استجابات الأفراد إلى استجابات جدية واستجابات غير جدية بناءً على درجة القطع لكل تدريج (ثلاثي، رباعي، خماسي، سباعي) وفق مؤشر (l_z^p)

التصنيف		التصنيف بناءً على عدد فئات التدريج				
		ثلاثي	رباعي	خماسي	سباعي	الكلي
جدي	العدد	333	301	311	319	1264
	النسبة المئوية من عدد فئات التدريج %	89.3%	80.3%	83.2%	85.5%	84.5%
غير جدي	العدد	40	74	63	54	231
	النسبة المئوية من عدد فئات التدريج %	10.7%	19.7%	16.8%	14.5%	15.5%
الكلي	العدد	373	375	374	373	1495

يلاحظ من الجدول (9) أنّ النسب الملاحظة للاستجابات غير الجدية تختلف باختلاف عدد فئات التدريج؛ فقد بلغت على التدريج الثلاثي (10.7%)، وعلى التدريج الرباعي (19.7%)، وعلى التدريج الخماسي (16.8%)، وعلى التدريج السباعي (14.5%). ونجد أن أقل نسبة من الاستجابات غير الجدية، ثم كانت على التدريج الثلاثي، ثم على التدريج السباعي، ثم التدريج الخماسي. وأخيرًا، فإن أعلى نسبة كشف عن الاستجابات غير الجدية كانت على التدريج الرباعي. ومن الملاحظ من الجدول أنّ أعلى نسبة استجابات غير جدية كشفها الإحصائي (l_z^p) هي عند تدريج ليكرت الرباعي، الذي يخلو من البديل "محايد". وتجدر الإشارة إلى أن بعض الدراسات بينت أن عدم وجود البديل " محايد " لمقياس ليكرت قد يؤدي إلى نقص المعلومات، أو عدم وجود رأي حول الموقف، أو نقص الدافعية لدى المستجيب، أو تناقض في المشاعر، أو عدم اليقين بقوة الموقف، أو عدم الاهتمام في أثناء الاستجابة، كدراسة شومان وبرسر(Schuman & Presser, 1981). وهذا قد يؤدي الى زيادة نسبة الاستجابات غير الجدية عند استخدام التدريج الرباعي لفقرات المقياس. كما وجد كالتون وآخرون (Kalton et al., 1980) وبيشوب (Bishop, 1987) في دراساتهم أن وجود البديل " محايد " له أثر على توزيع الاستجابات على البدائل المتبقية، مما قد يؤدي إلى وجود استجابات عشوائية أو غير جدية من المستجيب. واقترح البعض تسمية البديل محايد بِـ (لا أعرف)، التي تكون مقبولة اجتماعيًا للمستجيبين ولا تجبرهم على تقديم آرائهم (Schuman & Presser, 1981). وأشار فرانسس وبوش (Francis & Busch, 1975) إلى أن الاستجابة بخلو البدائل من "محايد" قد تدل على الإهمال أو الالتباس. وهذا يشير إلى الاستجابة العشوائية أو التخمين أو تقديم نمط معين في الاستجابة.

ومن الملاحظ أيضًا من الجدول (9) أنّ أقل نسبة استجابات غير جدية كشفها الإحصائي (l_z^p) هي عند تدريج ليكرت الثلاثي. وقد يفسر ذلك بأن قلة عدد الخيارات على الفقرة لا يتطلب الوقت والجهد من المستجيب لاختيار البديل المناسب؛ إذ إن السرعة في الاستجابة قد تؤدي إلى استجابات غير جدية (Huang, Liu & Bowling, 2015). وبالتالي، فإن قلة عدد الخيارات في (التدريج الثلاثي) قد تسهم في تقليل نسب الاستجابات غير الجدية الناتجة عن التسرع في الإجابة.

ولإجراء المقارنات الثنائية بين النسب المئوية الأربع بعد تصحيح مستوى الدلالة الإحصائية بالاعتماد على طريقة بنفروني (Bonferroni Correction)، ومفادها أنه عند إجراء المقارنات الثنائية يتغير مستوى الدلالة بناءً على عدد النسب؛ وذلك للمحافظة على الخطأ من النوع الأول؛ فتمت قسمة مستوى الدلالة الإحصائية على عدد المقارنات الثنائية، لتصبح قيمة مستوى الدلالة الإحصائية بعد التصحيح (0.008). وتم حساب عدد المقارنات الثنائية بين النسب، ثم تم استخدام كاي تربيع (χ^2) للمقارنة الثنائية بين النسب، كما هو مبين في الجدول (10).

جدول (10):المقارنات الثنائية بين النسب واختبار كاي تربيع χ^2

فئة التدريج	ثلاثي		رباعي		خماسي	
الثلاثي	الاحتمالية		الاحتمالية		الاحتمالية	
رباعي	11.750	0.001				
خماسي	5.886	0.015	1.045	0.307		
سباعي	2.386	0.122	3.642	0.056	0.793	0.373

يلاحظ من الجدول (10) أنّ نسب الكشف عن الاستجابات غير الجدية مرتبطة بفئة التدريج. وبناءً على مقارنة قيمة الاحتمالية بقيمة (0.008)، كانت هناك دلالة إحصائية واحدة فقط بين التدريج الرباعي والتدريج الثلاثي، مما يدل على أن أعلى نسبة كشف عن الاستجابات غير الجدية كانت على التدريج الرباعي وأقل نسبة كشف عن الاستجابات غير الجدية كانت على التدريج الثلاثي. أما باقي المقارنات الثنائية بين أزواج فئات التدريج، فلم تكن أي منها ذات دلالة إحصائية.

ويمكن تلخيص نتائج الدراسة الحالية في أن الاستجابات غير الجدية التي تم الكشف عنها باستخدام المؤشر الاحصائي (l_z^p) تتأثر بعدد فئات تدريج الفقرة وبوجود البديل "محايد". فقد اتضح أن زيادة عدد فئات التدريج قد يؤدي الى زيادة نسب الاستجابات غير الجدية. كذلك فإن عدم وجود البديل "محايد" قد يدفع بعض المستجيبين إلى عدم الجدية في الاستجابة. وقد تكون هذه النتيجة مرتبطة بشكل أساسي بالجوانب النفسية والاجتماعية للمستجيب. فقد تساهم عوامل السرعة والوقت والجهد الذي تتطلبه الاستجابة على فقرات المقياس في زيادة نسب الاستجابات غير الجدية. وبالتالي، فإن على الباحثين أخذ ذلك بعين الإعتبار عند إعداد أدوات جمع البيانات من أجل إلى الوصول الى بيانات ذات جودة عالية. كما أن حذف البديل "محايد" قد يؤدي إلى حذف البديل الذي ينطبق على المستجيب، وبالتالي دفعه لعدم الاستجابة بجدية عن فقرات المقياس.

التوصيات

في ضوء نتائج الدراسة الحالية، يوصي الباحثان بما يأتي:

1) أهمية وجود البديل " محايد " في مقياس ليكرت، كونه يقلل من نسب الاستجابات غير الجدية.

2) عدم المبالغة بزيادة عدد فئات تدريج مقياس ليكرت، ما لم يضطر الباحث لذلك أواستخدام التدريج الثلاثي للفقرات.

3) إجراء المزيد من الدراسات لمعرفة أثر البديل "محايد" على نسب الاستجابات غير الجدية بحيث يكون عدد فئات التدريج ستة بدائل.

References

Alghazo, A., & Alquraan, M. (2018). The effectiveness of bootstrapping procedure in estimating (l_z^p) person-fit statistic cutoff- score for polytomous items in noncognitive scales. *Jordan Journal of Educational Sciences* , 14(1), 1-12.

Alghazo, A. (2017). *The characteristics of person-fit index sampling distributions and the percent of miss-fit responses according to graded response model in the light of some variables.* PhD. Dissertation, Yarmouk University, Irbid, Jordan.

Al-Hamouri, F. (2017). Cognitive biases among Yarmouk University students in relation to gender and academic achievement. *Jordan Journal of Educational Sciences*, 13(1), 1-14.

Al-Quraan, M. (2019). The effect of insufficient effort responding on the validity of student evaluation of teaching. *Journal of Applied Research in Higher Education*, 11(3), 604-615.

Armstrong, R., Stoumbos, Z., Kung, M., & Shi, M. (2007). On the performance of the lz person-fit statistic. *Practical Assessment, Research & Evaluation*, 12(16), 1-15.

Baer, R., Ballenger, J., Berry, D., & Wetter, M. (1997). Detection of random responding on the MMPI-A. *Journal of Personality Assessment*, 68(1), 139-151.

Bishop, F. (1987). Experiments with the middle response alternative in survey questions. *Public Opinion Quarterly*, 51(2), 220-232.

Brown, R., & Villareal, J. (2007). Correcting for person misfit in aggregated score reporting. *International Journal of Testing*, 7(1),1-25.

Chen, W., & Thissen, D. (1997). Local dependence indices for item pairs using item response theory. *Journal of Educational and Behavioral Statistics* , 22(3), 265-289.

Cook, N., Faust, D., Meyerm J., & Faust, K. (2016). The impact of careless and random responding on juvenile forensic assessment: Susceptibility of commonly used measures and implications for research and practice. *Journal of Forensic Psychology Practice,* 16(5), 425-447.

De la Toree, J., & Deng, W.(2008). Improving person fit assessment by correcting the ability estimate and its reference distribution. *Journal of Educational Measurement,* 45(2), 159-177.

Deng, W. (2007). *An innovative use of the standardized log-likelihood statistic to evaluate person fit*. Doctoral Dissertation. The State University of New Jersey, New Brunswick, NJ.

Deu, A., & Reise, P. (1991). The influence of test characteristics on the detection of aberrant response patterns. *Applied Psychological Measurement* , 15(3), 217-226.

Dodeen,H., & Darabi, M. (2009). Person-fit: relationship with personality test in mathematics. *Research Papers in Education,* 24 (1), 115-126.

Drasgow, F., Levine, M.V.,& Williams, E. (1985). Appropriateness measurement with polychotomous item response models and standardized models. *British Journal of Mathematical and Statistical Psychology,* 38(1), 67-86.

Francis, J., & Busch, L. (1975). What we don't know about "I don't know". *Public Opinion Quarterly*, 39(2), 207-218.

Huang, J., Curran , P., Keeney, J., Poposki, E., & DeShon, R. (2012). Detecting and deterring insufficient effort responding to surveys. *Journal of Business Psychology,* 27*(*1), 99-114.

Huang, J., Liu, M., & Bowling, N. (2015). Insufficient effort responding: Examining an insidious confound in survey data. *Journal of Applied Psychology*, 100(3), 828-845.

Kalton, G., Julie, R., & Tim, H. (1980). The effects of offering a middle response opinion with opinion questions. *The Statistician*, 29(1), 65-78.

Karabatsos, G. (2003). Comparing the aberrant response detection performance of thirty – six person-fit statistics. *Applied Measurement in Eduction,*16(4), 277-298.

Kbar, F. (1994). Effectiveness of person-fit indices in detecting guessing carelessness: A simulated study. *Dissertation Abstracts International -Section A: Humanities and Social Sciences,* 55(3-A), 541.

Lamprianou , I., & Boyle , B. (2004) Accuracy of measurement in the context of mathematics national curriculum tests in England for ethic minority pupils and pupils who speak English as an additional language. *Journal of Educational Measurement* , 41(3), 239-259.

Levine, M., & Rubin, D. (1979). Measuring the appropriateness of multi-choice test scores. *Journal of Educational Statistics,* 4(4), 269-290.

Li, M., & Olejnik, S. (1997). The power of rasch person-fit statistic in detecting unusual response patterns. *Applied Psychological Measurement,* 21(3), 215-231.

Lopez, A., & Montesinos, H.(2005). Fitting Rasch model using appropriateness measure statistics. *The Spanish Journal of Psychology,* 8(1), 100-110.

Lord, F. (19980). *Application of item response to practical testing problems*. NJ: Lawrence Erlbaum Associates, Inc.

McGrath, R., Mitchell, M., Kim, B., & Hough, L. (2010). Evidence for response bias as a source of error variance in applied assessment. *Psychology Bulletin,* 136(3), 450-470.

Meade, A., & Craig, S. (2012). Identifying careless responses in survey data. *Psychological Methods,* 17(3), 437-455.

Meijer, R., & Nering. L.(1997). Trait level estimation for nonfitting response vectors. *Applied Psychological Measurement,* 21(4), 321-336.

Meijer, R., & van Krimpen-Stoop, E.(1999).The null distribution of person fit statistics for conventional and adaptive test. *Applied Psychological Measurement,* 23(4), 327-345.

Meijer, R.,& Sijtsma, K.(2001). Methodology review: Evaluating person fit. *Applied Psychological Measurement,* 25(2), 107-135.

Molenaar, I. , & Hoijtink, H. (1990). The many null distribution of person fit indices. *Psychometrika*, 55(1), 75-106.

Nering, M.(1995). The distribution of person fit using true and estimated person parameters. *Applied Psychological Measurement*, 19(2), 121-129.

Noonan, B., Boss, M., & Gessaroli, M. (1992). The effect of test length and IRT model on the distribution and stability of three appropriateness indices. *Applied Psychologcal Measurement*, 16(4), 345-352.

Noonan, B., Boss, M., & Gessaroli, M. (1992). The effect of test length and IRT model on the distribution and stability of three appropriateness indices. *Applied Psycholog-cal Measurement*, 16(4), 345-352.

Pina, J., & Montesions , M. (2005). Fitting Rasch model using appropriateness measure statistics. *The Spanish journal of Psychology*, 8(1), 100-110.

Reise, S. (1995). Scoring method and detection of person misfit in a personality assessment context. *Applied Psychological Measurement*, 19(3), 213-229.

Rupp, A. (2013). A systematic review of the methodology for person fit research in item response theory: Lessons about generalizability of inferences from the design of simulation studies. *Psychological Test and Assessment Modeling*, 55(1), 3-38.

Schuman, H., & Presser, S. (1981). *Questions and answers in attitude surveys*. New York: Academic Press.

Sinharay, S. (2015). Assessment of person fit for mixed-format tests. *Journal of Educational and Behavioral Statistics* , 40(4), 343-365.

Schmitt, N., Chan, D., Sacco, J., McFarland, L., & Jennings, D. (1999). Correlates of person-fit and effect of person-fit on test validity. *Applied Psychological Measurement*, 23(1), 41-54.

Snijders, T. (2001). Asymptotic null distribution of person fit statistics with estimated person parameters. *Psychometrika*, 66(3), 331-342.

Tourangeau, R., & Yan, T. (2007). Sensitive questions in surveys. *Psychological Bulletin*, 133(5), 859.

Van der Gaag, M., Schütz, C., Ten Napel, A., Landa, Y., Delespaul, P., Bak, M.,... & De Hert, M. (2013). Development of the Davos assessment of cognitive biases scale (DACOBS). *Schizophrenia Research*, 144(1-3), 63-71.

Zijlstra, W., Ark. , A., & Sijtsma, K. (2011). Outliers in questionnaire data: Can they detect, and should they be removed. *Journal of Educational and Behavioral Statistics*,63(2), 186-212.