أثر الاشكال المختلفة للاستجابة في الخصائص السيكومترية لمقياس الشخصية القائم على شبكة الإنترنت باستخدام مقياس ليكرت، ومقياس التناظر البصري والمقياس ثنائي التفرع

إعداد دكتورة/ شيرين فاروق محمد طنطاوي

مدرس علم النفس الإحصائي بكلية الاداب – جامعة الفيوم

المستخلص

هدفت الدراسة إلى البحث في أثر الاشكال المختلفة للاستجابة في الخصائص السيكومترية لمقياس الشخصية القائم على شبكة الإنترنت باستخدام مقياس ليكرت، ومقياس التناظر البصري والمقياس ثنائي التفرع. تستخدم الدراسة الحالية تصميم القياسات المتكررة (تسمح بتحليل الصدق التقاربى وتحليل الانحدار بين المقاييس) بما في ذلك المقياس ثنائي التفرع وأنواع ليكرت من ٥ نقاط، ومقياس التناظر البصري لمقارنة أشكال الاستجابة فيما يتعلق بالخصائص السيكومترية والفترات الزمنية الفاصلة للمقياس. من أجل الحصول على معلومات حول استخدام مقياس التناظر البصري في استبيان الشخصية، يتم استخدام المفردات الثمانية التي تقيس مظهر "الإخلاص" لبعد "الضمير" في مقياس العوامل الكبرى للشخصية NEO-PI-R1 (Berth, Goldschmidt, Ostendorf, & Angleitner, 2006). تتم مقارنة صدق المحك ومعاملات الثبات بين أشكال الاستجابة المختلفة. يتم تقدير الصدق التقاربى من خلال ارتباطات المتغيرات الكامنة للمقاييس الثلاثة. أيضًا، تتم مقارنة الفترات الزمنية المدركة والنقاط المركزية لمقياس التناظر البصري بالمقياس ثنائي التفرع والمكون المكون من ٥ نقاط للحصول على معلومات حول قابلية المقارنة بين عتبات ومراكز مقاييس الاستجابة. تكونت عينة الدراسة الحالية في صورتها النهائية من (٣٠٠) طالب/طالبة في كلية الآداب جامعة الفيوم، في الفرق الدراسية الأولى إلى الرابعة (١٩٠ طالبة، ٦٣,٣ %، ١١٠ طالب ،٣٦,٧%). تتراوح أعمارهم من ١٩ الى ٢١ عامًا (متوسط= ٢٠,٤ ، انحراف معياري = ٦,١٢). تتكون العينة من ١١٠ طالب/طالبة (٣٦,٧%) من الفرقة الاولى، ٩٠ طالب/طالبة (٣٠%) من الفرقة الثانية، ٧٠ طالب/طالبة (٢٣,٣%) من الفرقة الثالثة،٣٠ طالب/طالبة (١٠%) من الفرقة الرابعة. أشارت النتائج الى أنه فيما يتعلق

بتقديرات الثبات، فقد تبين أن ألفا كرونباخ ازداد مع عدد بدائل الاستجابة للمقاييس، بينما ظلت معاملات أوميجا ومعاملات الصدق ثابتة. توفر أوميجا تقديرًا أكثر ملاءمة للاتساق الداخلي. أشارت تحليلات الانحدار الخطي المعممة إلى أن المشاركين حكموا على الفترات الزمنية بين مستويات مختلفة من الإخلاص على مقياس التناظر البصري بشكل مختلف عن المقياسين الآخرين. وخلصت الدراسة إلى أن اختيار شكل الاستجابة لا ينبغي أن يعتمد حصريًا على الخصائص السيكومترية المرغوبة بل على اعتبارات عملية.

الكلمات المفتاحية: أشكال الاستجابة، الخصائص السيكومترية، مقياس الشخصية، مقياس ليكرت، ومقياس التناظر البصري، المقياس ثنائي التفرع

المقدمة

يعتبر إجراء البحث عبر الإنترنت ثورة ثانية في البحث السلوكي بعد ثورة الكمبيوتر في أواخر الستينيات وأوائل السبعينيات والتي جلبت العديد من المزايا مقارنة بإجراءات الورق والقلم المستخدمة على نطاق واسع (على سبيل المثال، القياس الآلي، دقة أفضل). أصبحت الاستبيانات والاختبارات القائمة على الإنترنت أكثر شيوعًا في علم النفس والعلوم الاجتماعية الأخرى (Reips, 2002; Reips & Birnbaum, 2011). عند تصميم استطلاع عبر الإنترنت، يعتمد السؤال المتعلق بكيفية تصميم العناصر وخيارات الاستجابة التي يجب تقديمها على مجموعة مختلفة من القواعد عن الاستطلاعات غير المتصلة بالإنترنت (Callegaro, Lozar Manfreda, & Vehovar, 2015; Hewson, Vogel, & Laurent, 2015) عند قياس درجة التوافق مع العبارة، غالبًا ما يتم تقديم خيارات الاستجابة في مقياس ليكرت Likert-type scales مع عدد معين من خيارات الإجابة الترتيبية (سامي محمد ،٢٠٠٩؛ صلاح الدين علام، ٢٠٠٦). تعتبر مقاييس التناظر البصرى visual analogue scales (VAS) بديلاً لمقاييس ليكرت. تقدم هذه المقاييس عددًا من المزايا مقارنة بالمقاييس من نوع ليكرت فيما يتعلق بخصائص القياس النفسي وهي سهلة التنفيذ عبر الإنترنت (Reips & Funke,2008). أظهر كل من سميث, Smith (1995)، وجينكينز وديلمان Jenkins and Dillman (1995)، وكوبر، تورانجو، كونراد، كروفورد Couper, Tourangeau, Conrad, CRAWFORD (2004) ، تويبويل ، داس ، فان سوست Toepoel, Das , Van Soest (2009)، وتويبويل وديلمان Toepoel and Dillman (2011) أن أشكال الاستجابة Response Formats تؤثر على الاستجابات. أظهر كوبر واخرون Couper et al (2004) أن الاستجابات تعتمد على العرض البصري المرئي للأسئلة.

مقياس ليكرت، الذي قدمه رينين ليكرت في عام ١٩٣٢، لا يزال حاليًا أكثر أشكال الاستجابة شيوعًا واستخدامًا. مقياس ليكرت النموذجي هو مقياس من ٥-٧ نقاط (فئة) يستخدم بشكل متكرر في استبيانات استطلاعات الرأي أو المسوح حيث يشير المشاركون إلى مستوى موافقتهم أو عدم موافقتهم مع عبارة ما. تجعل مقاييس ليكرت المستفتي يحدد خبرته عن طريق اختيار فئة متدرجة واحدة من خيارات الفئة (حجاج غانم أحمد ،٢٠١١)، ومع ذلك، هناك آراء مختلفة حول العدد الأمثل لبدائل

الاستجابة(Dourado et al.,2021). على الرغم من أنه من الأسهل والأبسط في الاستخدام، والأكثر فهما للباحث والمستجيب، خاصة الأطفال، وأن الترميز والتفسير أكثر سهولة ويسر، إلا أنه يحتوي على بعض القيود المتأصلة. أحد القيود هو أن الفئات الوصفية المصطنعة ليست كافية للتعرف على ظاهرة مستمرة. لا يمكن افتراض أن الفئات متباعدة بشكل متساوٍ (Alkadi et al.2022) .

لبناء استبيان، توجد مجموعة كبيرة ومتنوعة من أشكال الاستجابة المختلفة. يمكن استخدام أشكال الاستجابة العددية أو اللفظية أو الثنائية أو المرئية والعديد من الأشكال الأخرى. العدد الأمثل لفئات الاستجابة سؤال تتم مناقشته غالبًا في الأدبيات. قام مايديو أوليفاريس ، كرامب ، فوريرو ، جالاردو بوجول وكوفمان Maydeu-Olivares, Kramp, Forero, Gallardo-Pujol and Coffman (٢٠٠٩) بدراسة هذه المشكلة في بحثهم، حيث سألوا الطلاب ملء استبيان من فئتين وثلاث وخمس فئات للاستجابة. لتقييم النتائج استخدم المؤلفون نظريات مختلفة (نظرية الاختبار الكلاسيكية، التحليل العاملي للمفردة، نظرية الاستجابة للمفردة). وأظهرت النتائج، التي استندت إلى نظريات مختلفة، بشكل متطابق أن زيادة عدد فئات الاستجابة أسفرت عن تقديرات ثبات أعلى، خاصة فيما يتعلق بالاتساق الداخلي. أظهرت النتائج أيضًا أن أشكال الاستجابات المختلفة ليس لها أي تأثير على الصدق التمييزي والتقاربى، والتي تم دراستها من خلال استبيانات أخرى. ومع ذلك، لا يزال المؤلفون يعتقدون أن هناك تأثيرًا على الصدق، ولكن فقط للاستبيانات الأطول التي تحتوي على أكثر من ٢٠ مفردة.

كان أوستر (Oaster,1989) مهتمًا أيضًا بالعدد الأمثل لفئات الاستجابة والتأثيرات على الخصائص السيكومترية للاستبيان، حيث أعطى الطلاب استبيانًا من ثلاث إلى تسع فئات للاستجابة. أشارت النتائج إلى أن تقديرات الثبات قد زادت بشكل ملحوظ مع عدد فئات الاستجابة. قارن بيركيتBirkett (1986) أشكال الاستجابة الثنائية ومقاييس الاستجابة من ستة وأربعة عشر فئة استجابة. أفاد المؤلف أن مقياس النقاط الست كان شكل الاستجابة الأكثر ثباتًا. قارن بريستون وكولمان Preston and Colman (2000) أيضًا الاشكال المختلفة للاستجابة. استخدم الباحثان مقاييس الاستجابة من فئتين إلى إحدى عشرة فئة بالإضافة إلى شكل استجابة مع أكثر من مائة فئة. اكتشف الباحثان أن أشكال الاستجابة ذات الفئات الأقل أسفرت عن معاملات أقل من حيث الثبات والصدق. وقد أثر هذا أيضًا على القوة التمييزية. مع زيادة فئات الاستجابة من ست إلى سبع نقاط، زاد الثبات باستخدام طريقة إعادة

الاختبار والصدق التقاربي والتمييزي. تبين أن معاملات أشكال الاستجابة ذات المزيد من نقاط القياس متسقة، مما يعني أنه لم يكن هناك مزيد من النمو. تعامل هانكوك وكلوكارس Hancock and Klockars (1991)مع الصدق المرتبط بالمعيار وقارنا أشكال الاستجابة من خمس وتسع نقاط للمقياس. كان على الطلاب تقييم تكرارات سلوك الأفراد في مقاطع فيديو قصيرة على مقاييس تصنيف مختلفة. أشارت النتائج إلى أن الصدق يزداد مع زيادة عدد فئات الاستجابة.

وتعد استبيانات الشخصية واحدة من أكبر المصادر في علم النفس للحصول على معلومات عن الأفراد. من أجل الحصول على أداة قياس يمكن الاعتماد عليها، يجب أن يفي الاختبار بمعايير محددة للجودة. يجب مراعاة الكثير من الجوانب المختلفة أثناء إنشاء الاختبار، على سبيل المثال اختيار تنسيق الاستجابة. لهذا السبب تفحص الدراسة الحالية تأثيرات تنسيقات الاستجابة المختلفة على الخصائص السيكومترية لاستبيان الشخصية.

مشكلة الدراسة

لأكثر من نصف قرن تم استخدام استطلاع الرأي باستخدام خيارات الاستجابة بمقياس ليكرت على نطاق واسع في المدارس لاستخلاص استنتاجات التي تتعلق بالطلاب. حتى الآن، ومع ذلك، لم يفحص الباحثون التربويون ما إذا كان مقياس مختلف - مثل المقياس الذي يستخدم شكل مستمر للاستجابة - سيكون له تأثير مماثل على استجابات الطلاب أو يؤدي إلى نتائج مختلفة عن تلك التي تم الحصول عليها من استبيان مقياس ليكرت في المراحل التعليمية المختلفة. هذا النقص في المقارنات المباشرة بين مقياس ليكرت والمقاييس الأخرى (على سيبل المثال: مقياس التناظر البصري والمقياس ثنائي التفرع) يترك أفضل طريقة لبناء خيارات الاستجابة في أبحاث المسح التربوي في المراحل التعليمية المختلفة غير واضحة

نظرًا لوجود ندرة في الأدلة التجريبية لدعم اختيار مقياس على آخر، فقد يكون الباحثون التربويون أقل ميلًا للتمييز بين المقاييس لخيارات استجابة المسح وأكثر ميلًا لاختيار ما هو مألوف أكثر (على سبيل المثال، مقياس ليكرت) أو أسهل في البناء (أو وضع الدرجات أو التطبيق) بدلاً من القياس الأكثر ملاءمة (على سبيل المثال، استنادًا إلى عمر العينة، أو سياق الدراسة، أو البناء الذي يتم قياسه) أو ما سيحقق أكثر النتائج دقة. من خلال مدخل القياس التقريبي في القياس، يمكن لـمقياس

ليكرت LS إحداث تحيزات إحصائية يمكن أن تكون ذات عواقب كبيرة لأنها يمكن أن "تزيد بشكل مصطنع" أو تخفف من أحجام التأثير ومعاملات الارتباط والثبات (Hasson & Arnetz,2005)

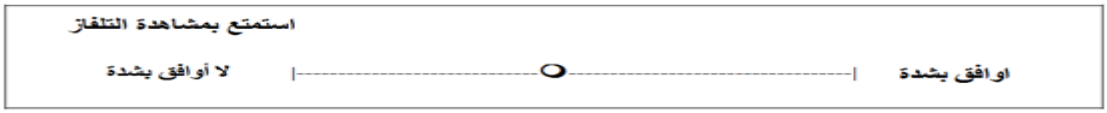
بالإضافة إلى التحيزات الإحصائية، يرى النقاد بأن مفردات مقياس ليكرت LS تحد من قدرة المستجيب على التعبير عن آرائه بدقة، وبالتالي فهي غير قادرة على تقديم أدلة غير متحيزة حول درجات معينة من الاتفاق أو الخلاف لأنها تفشل في توضيح الفروق الدقيقة في تعبير الفرد (Flynn, van Schaik, & Middlesorough, 2004).

في الواقع، إن ما يحاول مقياس ليكرت القيام به هو نقل البناء المستمر إلى نظام رقمي ترتيبي. وبالتالي، من خلال إجبار المستجيبين على الاختيار من بين مجموعة من الاستجابات "المقترحة أو المقدمة" - والتي قد تعكس أو لا تعكس بدقة ما يشعرون به حقًا أو ما يفكرون به حقًا - يمكن أن تكون النتائج التي تم الحصول عليها من مفردات مقياس ليكرت متحيزة لتعكس فقط الدرجات المحدودة للموافقة (على سبيل المثال، موافق، لا أوافق بشدة) المقدمة من الشخص الذي قام ببناء المقياس بدلاً من أن يعكس الاستجابات المدركة أو المقصودة من المستجيبين أنفسهم (Bowling, 1998; Brunier & Graydon, 1996). تحقيقًا لهذه الغاية، يبدو أن مقياس ليكرت قادر على تزويد الباحثين فقط بتقريب متجانس لمواقف المستجيبين بسبب التصنيف الخام لاستجابات الأفراد. نتيجة لذلك، قد يكون مقياس ليكرت غير قادر على تقديم انعكاس أكثر دقة للظاهرة المقاسة لأنه يجمع المستجيبين في مجموعات متميزة بشكل مصطنع (على سبيل المثال، المستجيبون الذين لا يوافقون بشدة مقابل أولئك الذين يوافقون ولا يوافقون)، مما يشير إلى تطابق الافراد في استجابة واحدة وموحدة. باختصار، يمكن أن يقدم مقياس ليكرت لتصنيف الاستجابات، في أحسن الأحوال، حساسية محدودة للمقياس فقط، والتي يمكن أن تثقل بشكل مباشر قدرة الباحث على الحصول على أكثر النتائج دقة.

وبذلك يصبح المطلوب في بحث استطلاع الآراء التربوية هو بديل لخيار استجابة مقياس ليكرت الذي يمكن أن يوفر للمستجيبين مزيدًا من الحرية لجعل استجاباتهم أكثر تأثرًا بشخصيتهم وله القدرة على تحقيق تقدير أكثر دقة للبنية المقاسة. أحد البدائل الممكنة لمقياس ليكرت هو مقياس التناظر البصري، والذي يمنح المستجيبين القدرة على التعبير عن آرائهم الشخصية بشكل أكثر دقة (Byrom et al.,2022)

وهو قادر على توفير حساسية أكبر للمقياس حتى يتمكن الباحثون من الحصول على نتائج أكثر دقة من الناحية النظرية.

مقياس التناظر البصريThe visual analogue scale (VAS) هو مقياس أحادي البعد - يعني فقط قدرة أو سمة أو بُعدًا واحدًا يتم قياسه في كل مرة (Bond & Fox,2001) - وغالبًا ما يتم تقديمه كخط أفقي منفرد مثبت على الجانب الأيسر بواسطة سمة سلبية أو العبارة الأكثر سلبية وعلى الجانب الأيمن من خلال سمة إيجابية أو العبارة الأكثر إيجابية. يُطلب من المستجيبين عادةً تحديد نقطة على طول السلسلة المتصلة بين الطرفين والتي تتطابق بشكل أفضل مع درجة التوافق أو قوة الاتفاق مع بعض العبارات.



بالمقارنة، لا تقدم خيارات استجابة مقياس ليكرت سوى جزء بسيط من خيارات الاستجابة الممكنة لمقياس التناظر البصري VAS، والتي لا تقتصر على مجموعة منفصلة من الاستجابات المحددة مسبقًا (على سبيل المثال: أوافق بشدة وأوافق وأعارض وأعارض بشدة) التي تقدمها مفردات مقياس ليكرت. كشكل من أشكال الاستجابة للمفردة، يمكن تطبيق مقياس التناظر البصري بنفسه (على سبيل المثال، باستخدام استراتيجية مفردة واحدة لأداة القياس) أو بالاشتراك مع مقياس التناظر البصري أخر لقياس البناءات المتعددة على أدوات متعددة المفردات (Dutheil et al.,2022).

وبالتالي، تسعى هذه الدراسة إلى المساهمة في الأدبيات من خلال معالجة الأسئلة البحثية التالية:

١- هل توجد فروق في متوسطات وتباينات الدرجات باختلاف أشكال الاستجابة (مقياس ليكرت، ومقياس التناظر البصري والمقياس ثنائي التفرع)؟

٢-هل يغير شكل الاستجابة البنية العاملية للمقياس؟

٣- هل توجد علاقة بين أشكال الاستجابة وثبات الاختبار؟

٤- هل الأشكال المختلفة للاستجابة (مقياس ليكرت، ومقياس التناظر البصري والمقياس ثنائي التفرع) تعبر عن نفس المتغير الكامن؟

٥- هل يرتبط العمر ارتباطًا موجبًا ودالًا بمظهر الإخلاص على مستوى مجموع الدرجات وكذلك على مستوى المتغيرات الكامنة؟

٦- هل توجد فروق ذات دلالة إحصائية لجميع أشكال الاستجابة ($p > .05$) من خلال اختبارات t- بين الإناث والذكور؟

هدف الدراسة

هدفت الدراسة إلى البحث في أثر الأشكال المختلفة للاستجابة في الخصائص السيكومترية لمقياس الشخصية القائم على شبكة الإنترنت باستخدام مقياس ليكرت، ومقياس التناظر البصري والمقياس ثنائي التفرع.

مبررات الدراسة

لذلك، تستخدم الدراسة الحالية تصميم القياسات المتكررة (تسمح بتحليل الصدق التقاربى وتحليل الانحدار بين المقاييس) بما في ذلك المقياس ثنائي التفرع، وأنواع ليكرت من ٥ نقاط، ومقياس التناظر البصري لمقارنة أشكال الاستجابة فيما يتعلق بالخصائص السيكومترية والفترات الزمنية الفاصلة للمقياس. من أجل الحصول على معلومات حول استخدام مقياس التناظر البصري في استبيان الشخصية، يتم استخدام المفردات الثمانية التي تقيس مظهر "الإخلاص" لبعد "الضمير" في مقياس العوامل الكبرى للشخصيةNEO-PI-R1 (Berth, Goldschmidt, Ostendorf, & Angleitner, 2006). تتم مقارنة صدق المعيار ومعاملات الثبات بين أشكال الاستجابة المختلفة. يتم تقدير الصدق المتقاربى من خلال ارتباطات المتغيرات الكامنة للمقاييس الثلاثة. أيضًا، تتم مقارنة الفترات الزمنية المدركة والنقاط المركزية لمقياس التناظر البصري بالمقياس ثنائي التفرع ومقياس ليكرت المكون من ٥ نقاط للحصول على معلومات حول قابلية المقارنة بين عتبات ومراكز مقاييس الاستجابة.

أهمية الدراسة

أهمية المعلومات الدقيقة أمر حتمي في جميع مجالات البحث، والبحوث الاستقصائية التربوية ليست استثناء. مع الطلب على اتخاذ القرارات القائمة على البيانات في البيئة

التعليمية عالية المخاطر اليوم، من الضروري أن تكون الأدوات المستخدمة في جمع البيانات دقيقة ومفيدة قدر الإمكان.

مجموعات الاستجابة مهمة جدًا لاختيار أشكال الاستجابة الأمثل. فقط تخيل أن شخصين قد وصلا إلى نفس الدرجات في الاستبيان. ربما يفضل أحدهما فئات الاستجابة في المنتصف بينما يميل الشخص الآخر إلى اختيار الفئات المتطرفة. على الرغم من أن كلاهما وصل إلى نفس الدرجة، إلا أن أنماط استجابتهما تختلف اختلافًا كبيرًا. ومن ثم تتمثل أهمية الدراسة الحالية في أن جودة أي دراسة بحثية تعتمد بشكل كبير على قدرة الباحث على جمع وتفسير بيانات صحيحة وموثوق بها.

كما تعتمد الدراسة الحالية على التطبيق عبر شبكة الانترنت، حيث أصبحت استطلاعات الويب بشكل عام جذابة بشكل متزايد للباحثين اليوم لأن بنائها غير مكلف نسبيًا، ويمكن أن يكون التطبيق سريعاً وسهلاً، ويمكن الحصول على النتائج في جزء صغير من الوقت الذي تستغرقه الاستطلاعات الورقية التقليدية.

مصطلحات الدراسة

مقياس ليكرت A Likert scale هو مقياس سيكومترى مشترك بشكل شائع في البحث الذي يستخدم الاستبيانات(Wuensch, 2005, P.6) .

مقياس التناظر البصري (VAS) هو مقياس استجابة سيكومترى يمكن استخدامه في الاستبيانات، وهو أداة قياس للخصائص أو الاتجاهات الشخصية التي لا يمكن قياسها بشكل مباشر(Funke& Reips,2012, P.311).

المقياس ثنائي التفرع A dichotomous scale هو مقياس من نقطتين يقدم خيارات معاكسة تمامًا لبعضها البعض. هذا النوع من خيارات الرد لا يمنح المستفتي فرصة ليكون محايدًا في إجابته على سؤال(Simms et al.,2019, P.2) .

الإطار النظري والدراسات السابقة

في علم النفس، تعد الاستبيانات هي الإجراء الأكثر استخدامًا للحصول على معلومات حول الاختلافات الشخصية وداخل الشخصية. يتم الرد عليها بشكل شائع من خلال مقاييس الاستجابة، والتي يتم تحديدها بواسطة أنواع مختلفة من تنسيقات الاستجابة. من بين العديد من الجوانب، تعد أشكال الاستجابة عنصرًا مهمًا يجب مراعاته أثناء

بناء الاستبيانات النفسية. لذلك، خضعت أشكال الاستجابة بشكل عام لأبحاث مكثفة منذ بداية الاستخدام الواسع للاستبيانات في علم النفس Simms, Zelazny, Williams& Bernstein,2019).

على وجه الخصوص، تم البحث بدقة في العدد الأمثل لفئات الاستجابة (Dawes, 2008; Maydeu-Olivares, Kramp, García-Forero, Gallardo-Pujol, & Coffman, 2009; Preston & Colman, 2000) لأنه ينتج عنه آثار مهمة تتعلق بالخصائص السيكومترية للاستبيان، حيث تبين أن الثبات (Weng,2004) وكذلك الصدق (Hancock & Klockars)، 1991 يتأثران بشكل كبير بعدد البدائل المعطاة على مقياس الاستجابة. أيضًا، يجب أن تكون مقاييس الاستجابة مناسبة لموضوع البحث المحدد وكذلك للمجموعة المستهدفة، إذ يمكن أن يؤدي المقياس الذي يفرق بين مجموعة مستهدفة إلى اتجاهات استجابة غير كافية (Briggs & Closs, 1999; Faulbaum, Prüfer, & Rexroth, 2009). عادةً ما يتم تعريف مقاييس الاستجابة التي تشتمل على عدد ثابت من فئات الاستجابة على أنها مقاييس من نوع ليكرت.

تشتمل الاختبارات النفسية عادةً على مقياس استجابة الغرض منه تنظيم وتقييد الخيارات المتاحة للمستجيبين وتسهيل تسجيل النقاط. أحد مقاييس الاستجابة هو مقياس ليكرت، والذي تم تقديمه في البداية ليكون له شكل محدد مكون من ٥ نقاط. في الممارسة العملية، تختلف هذه المقاييس اختلافًا كبيرًا في طبيعة وعدد خيارات الاستجابة. ومع ذلك، يوجد إجماع ضئيل نسبيًا بشأن العديد من الأسئلة التي ظهرت فيما يتعلق باستخدام مفردات من نوع ليكرت. أولاً، هل هناك عدد مثالي لخيارات الاستجابة من الناحية السيكومترية؟ ثانيًا، هل من الأفضل تضمين عدد زوجي أم فردي من خيارات الاستجابة؟ أخيرًا، هل تقدم مفردات مقياس التناظر البصري أي مزايا على المفردات من نوع ليكرت؟ درس سيمز ، زيلازني ، ويليامز ، وبرنشتاين Simms, Zelazny, Williams, & Bernstein (2019)هذه الأسئلة على عينة من ١,٣٥٨ طالبًا جامعيًا تم تعيينهم عشوائيًا في مجموعات للإجابة على استبيان الشخصية باستخدام مقاييس استجابة تتراوح من ٢ إلى ١١ خيارًا، وحالة التناظر البصري. أظهرت النتائج دقة سيكومترية ضعيفة لمقاييس الاستجابة مع ٢ إلى ٥ خيارات استجابة. ومن المثير للاهتمام، مع ذلك، أن نتائج صدق المعيار لم تتبع هذا النمط. أيضًا، لم يتم الكشف عن المزايا السيكومترية لأي مقاييس استجابة تتجاوز ٦ خيارات، بما في ذلك التناظر البصري.

أتاحت دراستان (Jones& Loe,2014) في عيادة / مختبر جامعي الفرصة لاستكشاف تأثير تقليل عدد بدائل الاستجابة في مقياس يقيس سمات الشخصية المهنية. في دراسة محاكاة، قدم التطبيق المعتمد على الكمبيوتر مقياسًا رقميًا لكل مفردة يتراوح من ٠ إلى ١٠. ثم تم إعادة الاختبارات لمحاكاة تأثير ثلاث اختيارات فقط. بالنسبة لدراسة المتابعة، تم إنشاء نسختين من المقياس، أحدهما بخيارين للاستجابة والآخر بستة خيارات للاستجابة، وتم تخصيصهما عشوائيًا للمشاركين. كانت العلاقات النموذجية واضحة بين سمات الشخصية المهنية والنتائج على المقاييس القياسية لسمات الشخصية الأساسية مع تأثير ضئيل لتقليل عدد بدائل الاستجابة. لم يكن لقدرة المشاركين ولا التوتر لاستخدام التقرير الذاتي تأثير قابل للقياس على فائدة النتائج مع تقليل بدائل الاستجابة.

هدفت دراسة القاضي واخرين (Alkadi et al.,2022) إلى تقييم ما إذا كانت استجابات مقياس التناظر البصري تحظى بدرجة عالية من الثبات مثل استجابات ليكرت من منظور طلاب طب الأسنان(جامعة الملك سعود بن عبد العزيز). تم أيضًا تقييم تفضيل كل مقياس وتأثير الجنس ومستوى التعليم على الشكل المفضل. لم يثبت أن ايا من نوعى مقياس الاستجابة أفضل من الاخر . وكانت من توصيات هذه الدراسة عند اختيار نوع مقياس الاستجابة، من المهم مراعاة هدف الدراسة والسياق، بما في ذلك نوع مجتمع الدراسة، والبنية التي يتم تقييمها أثناء تطوير أو تعديل الأدوات وتصميم الدراسة.

الخصائص السيكومترية لأشكال الاستجابة بفواصل زمنية ثابتة

وجد وينج Weng (2004) زيادة مطردة في الاتساق الداخلي من فئتين إلى تسع فئات استجابة، بما يتماشى مع الدراسات السابقة المختلفة (Birkett, 1986). تم اقتراح ذلك أيضًا من خلال دراسة محاكاة مونت كارلو بواسطة اجراها Lozano, García-Cueto, and Muñiz (2008). ومع ذلك، فإن ارتفاع درجة ثبات المقاييس مع زيادة إمكانيات الاستجابة ليس لها حدود معينة. فقد أشارت بيركيت Birkett إلى مقياس مكون من ٦ نقاط لتحقيق أعلى اتساق داخلي عند مقارنته بالمقياس ثنائي التفرع (DS) ومقياس مكون من١٤ نقطة.

قام بريستون، كولمان Preston and Colman (2000) بالبحث في تأثير بناء المقياس على الصدق التقاربى والصدق المرتبط بالمعيار. في حين أن الصدق

المرتبط بالمعيار لم يختلف اختلافًا دالاً بين المقياسين، فإن الصدق التقاربى يميل إلى أن يكون أكبر مع زيادة عدد احتمالات الاستجابة، لتصل إلى ذروتها عند سبع نقاط. من أجل البحث في تأثير أشكال الاستجابة على الصدق المرتبط بالمعيار، قارن هانكوك، كلوكارس Hancock and Klockars (1991) الاشكال بمقياسين من ٥ نقاط و٩ نقاط، مع ملاحظة معاملات أعلى للأخير.

إلا أن فيراندو Ferrando (2003) لم يجد علاقة ذات دلالة إحصائية بين عدد فئات الاستجابة والثبات داخل الصف أو الصدق المرتبط بالمعيار. ومع ذلك، في دراسته، قارن مقاييس الاستجابة بين نقطتين و١٩ نقطة، مما أدى إلى ١٨ مجموعة مختلفة، تضم كل مجموعة ٢٠ شخصًا فقط. لذلك من المحتمل أن يكون سبب عدم وجود اختلاف دال في الثبات أو الصدق بين المجموعات هو نقص القوة الإحصائية.

ربما أجرى مايادو واخرون Maydeu- Olivares et al. (2009)- على الأرجح - التحليل الأكثر شمولاً من الناحية السيكومترية للعلاقة بين أشكال الاستجابة والخصائص السيكومترية للمقاييس، حيث قارن عواقب استخدام أشكال استجابة مختلفة ضمن أطر نظرية القياس الكلاسيكية(CTT) ، والتحليل العاملي للمفردة (IFA)، ونظرية الاستجابة للمفردة .(IRT) في هذه الدراسة، استخدم هؤلاء الباحثون استبيانات مع فئتين وثلاث وخمس فئات للاستجابة. في إطار نظرية القياس الكلاسيكية، تم استخدام معامل ألفا، وهو التقدير الأكثر تطبيقًا للاتساق الداخلي (Graham,2006)، لتقدير الاتساق الداخلي وأظهر تقديرات متزايدة مع فئات استجابة أكثر. في الإطارين الآخرين، تم استخدام معامل أوميجا Coefficient Omega ولم يظهر أي فرق دال مرتبط بعدد فئات الاستجابة، في كل من التحليل العاملي للمفردة ونظرية الاستجابة للمفردة. بالإضافة إلى ذلك، ظل الصدق التمييزي والتقاربي غير متأثرة على نطاق واسع بعدد بدائل الاستجابة في جميع الأطر.

يبدو أن العدد الأمثل لفئات الاستجابة - فيما يتعلق بالخصائص السيكومترية للمقياس - يكمن حول سبعة. قد تكون الصورة المخططة للنتائج نتاجًا للتأثيرات التي لا حصر لها على الخصائص السيكومترية للاستبيانات والفروق في أخطاء القياس، والتي تميل إلى أن تكون أعلى بالنسبة لأشكال الاستجابة مع العديد من إمكانيات الاستجابة (Goggin & Stoker, 2014; Shulman & Boster, 2014)

الخصائص السيكومترية لمقاييس الاستجابة بدون فترات زمنية ثابتة

بالإضافة إلى أنواع ليكرت مع أعداد متفاوتة من الفئات، يتم استخدام مقاييس التناظر البصري (VASs) لتحديد الاستجابات في الاستبيانات. ومع ذلك، فقد حظيت أشكال الاستجابة باهتمام أقل بكثير، وهذا ليس مفاجئًا لأن أنواع ليكرت هي أشكال الاستجابة الأكثر استخدامًا في الاستبيانات (Bühner,2011) وتستخدم مقاييس التناظر البصري في الغالب في التقييم الإكلينيكي ((Wewers & Lowe,1990 والدراسات الاستقصائية عبر الإنترنت (Couper, Tourangeau, Conrad, & Singer, 2006). في واحدة من الدراسات القليلة بما في ذلك أنواع ليكرت ومقاييس التناظر البصري، قارن فلاين وفان، وريسك Flynn, van Schaik, and van Wersch (٢٠٠٤) مقياسًا للتناظر البصري، ٦٥ ملم ونوع ليكرت من ٧ نقاط لاستبيان عن المجابهة النفسية Coping. أظهرت النتائج معاملات ثبات مُرضية لكلا المقياسين، بينما كان الصدق البنائي أعلى بالنسبة لمقياس التناظر البصري. ومن المثير للاهتمام، أن المقياس المكون من ٧ نقاط أدى إلى مستويات أعلى بكثير من متوسط المجابهة الوظيفية مقارنة بمقياس التناظر البصري، ربما بسبب نقص خبرة المشاركين في شكل الاستجابة، كما توقع المؤلفون. وقد أوضح Grant et al. (1999) أن مقياس التناظر البصري متفوقًا على نوع ليكرت المكون من ٥ نقاط فيما يتعلق بالثبات بطريقة إعادة الاختبار. كما وجدجيسك، سينجر، جويات Jaeschke, Singer, and Guyatt (1990)تقاربًا بين الصدق المتعلق بالمعيار لنوع ليكرت المكون من ٧ نقاط ومقياس التناظر البصري. ومع ذلك، كما لاحظ هؤلاء الباحثون، فإن حجم العينة الصغير المكون من ٢٠ مريضًا فقط يحد من قابلية تعميم هذه النتيجة. وتجدر الإشارة إلى أن أشكال الاستجابة ثنائية التفرع، حسب علم الباحثة، لم تتم مقارنتها بمقاييس التناظر البصري.

وبالتالي، تظل النتائج المتعلقة بمقياس التناظر البصري متنوعة، كما هو الحال بالنسبة للجوانب المنهجية للدراسات المذكورة أعلاه: المجابهة النفسية (Flynn et al.,2004)، والحالات الطبية (Grant et al.,1999)، وجودة الحياة المدركة (Jaeschke et al., 1990) ليست سوى بعض المقاييس المختلفة التي تم البحث فيها في الدراسات السابقة التي سبق ذكرها. أيضًا، استخدم العديد من المؤلفين مفردة واحدة فقط لمقياس التناظر البصري (Grant et al.,1999)، وبالتالي توفير إمكانية مقارنة محدودة للغاية للأدوات متعددة المفردات .(Wewers & Lowe,1990)

شدد معظم المؤلفين على أهمية التدريب والشرح الصحيح لاستخدام مقاييس التناظر البصري: قد يؤدي عدم وجودها إلى ميول المستجيبين لاستخدام مقاييس التناظر البصري بنفس الطريقة التي يستخدمون بها أنواع ليكرت وعدم الاستفادة من النطاق الكامل من السابق (Couper et al., 2006)

المقاييس على شبكة الإنترنت

أصبحت استطلاعات الويب بشكل عام جذابة بشكل متزايد للباحثين اليوم لأن بناءها غير مكلف نسبيًا، ويمكن أن يكون التطبيق سريع وسهل، ويمكن الحصول على النتائج في جزء صغير من الوقت الذي تستغرقه المقاييس الورقية التقليدية .ومع ذلك، تتطلب المقاييس على الويب الجيدة (كما تم قياسها من خلال المؤشرات المقبولة لجودة المقاييس) بذل القليل من الجهد الإضافي لأن الباحثين يجب أن يأخذوا في الاعتبار ليس فقط الافراد الذين يحاولون الوصول إليهم ولكن أيضًا القضايا الخاصة بالنمط المتعلق بأشكال خيارات الاستجابة، أنواع الأسئلة التي يتم طرحها، وكذلك عملية جمع البيانات (Tucker-Seeley,2008) .

إجراءات الدراسة

عينة الدراسة

تكونت عينة الدراسة الحالية في صورتها النهائية (من عينة مبدئية قوامها ٣٤٠ طالبا من الجنسين مما اجابوا على الاستبيان ، حيث تم استبعاد المقاييس غير المكتملة) من (٣٠٠) طالب/ طالبة في كلية الآداب جامعة الفيوم، في الفرق الدراسية الأولى إلى الرابعة (١٩٠ طالبة، ٦٣,٣ %، ١١٠ طالب ،٣٦,٧%). تتراوح اعمارهم ١٩ الى ٢١ عاما (متوسط= ٢٠,٤ ، انحراف معياري = ٦,١٢). تتكون العينة من ١١٠ طالب/ طالبة (٣٦,٧%) من الفرقة الأولى، ٩٠ طالب/طالبة (٣٠%) من الفرقة الثانية، ٧٠ طالب/ طالبة (٢٣,٣%) من الفرقة الثالثة،٣٠ طالب/ طالبة (١٠%) من الفرقة الرابعة. تم استخدام أسلوب العينة العشوائية فى اختيار العينة(انظر جدول ١).

جدول (١) خصائص العينة وتوزيعها

الفرقة	العدد	بنين	بنات	النسبة المئوية من العينة الكلية
الاولى	110	56	54	%36.7
الثانية	90	24	66	%30
الثالثة	70	20	50	%23.3
الرابعة	30	10	20	%10
العينة الكلية	300			

الأدوات

استبانة الشخصية: مظهر الإخلاص في بُعد" الضمير" NEO-PI-R (Berth et al., 2006)

نسخة معدلة من الأسئلة الثمانية التي تقيس مظهر الاخلاص في بُعد" الضمير" NEO-PI-R (Berth et al., 2006). تم اختيار مظهر الاخلاص في بُعد" الضمير" لأنه ذو محتوى محايد وسهل الفهم. تم إعادة صياغة بعض المفردات الأصلية من أجل الوضوح. على سبيل المثال: (١) "في بعض الأحيان أنا لست الشخص الذي يمكن أن يُعتمد عليه كما ينبغي" تم إعادة صياغة هذه العبارة إلى "أنا شخص يمكن الاعتماد عليه" و (٢) تم تغيير عبارة "أنا أفي بالتزاماتي المالية بالكامل وفي أسرع وقت ممكن" إلى "أنا أفي بالتزاماتي". تم إرسال ثلاث نسخ من الاستبيان مسبقًا على ثلاث أوراق منفصلة، تحتوي كل منها على نفس المفردات الثمانية ولكن بمقياس استجابة مختلف (مقياس ليكرت، ومقياس التناظر البصري والمقياس ثنائي التفرع).

يتكون المقياس ثنائي التفرع من مربعين: [نعم] و [لا] للإشارة إلى الموافقة أو عدم الموافقة، على التوالي. تتألف أنواع ليكرت من ٥ نقاط في خمس مربعات: [غير صحيح تمامًا]، و [غير صحيح]، و [أحيانًا]، و]صحيح]، و]صحيح تمامًا]. يتألف مقياس التناظر البصري من خط أفقي يبلغ طوله ١٠ سم، والطرفان الأيسر والأيمن كما يلي: [ليس صحيحًا على الإطلاق] و [صحيح تمامًا]، على التوالي. تم اختيار ترتيب الأوراق الثلاث عشوائيًا عبر المشاركين، من أجل موازنة التأثيرات المحتملة لترتيب أشكال الاستجابة، مما أدى إلى ثلاث بطاريات اختبار مختلفة: (مقياس ليكرت، ومقياس التناظر البصري والمقياس ثنائي التفرع).

قامت الباحثة بترجمة النسخة الاجنبية من المقياس إلى اللغة العربية. تم حساب الثبات باستخدام اعادة الاختبار t-retest وجاء معامل الارتباط 0.691. كما تم حساب الثبات باستخدام معامل ألفا كرونباخ للاتساق الداخلي Cronbach's α، وكان 0.896. كما تم حساب صدق المحتوى لمظهر الإخلاص في بُعد" الضمير" من قبل مجموعة من ٦ خبراء، قيّموا مدى ملاءمة كل مفردة باستخدام مقياس Likert المكون من أربع نقاط (حيث يمثل ١ "غير ذي صلة" و٤ يمثل "ذو صلة كبيرة")، وقدموا اقتراحاتهم وتعليقاتهم. تم الحكم على المفردات الثمانية (٨) بأنها وثيقة الصلة نوعًا ما أو وثيقة الصلة بشكل كبير. تم حساب مؤشر صدق المحتوى على مستوى المفردة (I-CVI = 0.90).

تحليل البيانات

تم تحليل الثبات لكل من المقاييس الثلاثة بواسطة ألفا alpha وأوميجا omega. تتضمن صيغة ألفا تباينات المفردات والمقياس:

$$\alpha = \frac{k}{k-1}\left(1 - \frac{\sum_{i=1}^{k}\sigma_{Y_i}^2}{\sigma_X^2}\right)$$

تشير k إلى عدد المفردات، لذا فإن الكسر الأول يصحح لعدد المفردات في المقياس. يشير الرمز $\sigma_{Y_i}^2$ إلى الفروق بين المفردات ويشير σ_X^2 إلى تباين المقياس.

كما هو موضح في الصيغة التالية، يتم حساب أوميجا باستخدام عوامل التشبع والتباينات الفريدة للمتغيرات الواضحة في النموذج:

$$\omega = \frac{\left(\sum_{i=1}^{k}\lambda_i\right)^2}{\left(\sum_{i=1}^{k}\lambda_i\right)^2 + \sum_{i=1}^{k}(1 - h_i^2)}$$

يشير k إلى عدد المفردات (المتغيرات الواضحة) بينما يرمز λ_i إلى التشبعات العاملية على العامل العام ويشير h_i^2 إلى الفروق الفريدة لمفردات k (المتغيرات الواضحة).

أجريت تحليلات التباين للقياسات المتكررة (rmANOVAs) لاختبار الفروق في النتيجة الإجمالية وفقًا لمقاييس الاستجابة المختلفة بالإضافة إلى تأثيرات ترتيب

الاختبار. تمت مقارنة متوسطات الفرد على المقاييس باستخدام اختبارات t. تم حساب صدق المعيار، باستخدام معامل ارتباط بيرسون Pearson's r، بين مجموع درجات الاستبيانات والمعيار. أجريت اختبارات F لاختبار الفروق بين التباينات.

أجريت تحليلات الانحدار الخطي لتحليل القيمة التنبؤية للمتغيرات (وتأثيراتها التفاعلية) على مجموع درجات الاستبيانات والمعيار. تم البحث في العلاقات بين أنماط الاستجابة على مقياس ليكرت، ومقياس التناظر البصري والمقياس ثنائي التفرع من خلال الانحدار الخطي المعمم للاستجابات الترتيبية (انحدار الاحتمالات النسبي) والثنائية (الانحدار اللوجستي).

لتحليل الصدق التمييزي، تم تقسيم الاستبيان إلى مقياسين، أحدهما يشتمل على مفردات لا تتعلق بالعمل (المفردات ٢ و٣ و٤ و٥ و٦)، والآخر يشمل المفردات المتعلقة بالعمل (المفردات ١، ٧، ٨)، من أجل إعداد مصفوفة متعددة الأساليب. يحاكي نموذج المعادلة البنائية الخطية الناتج (SEM) يحاكي النموذج الذي استخدمه سايو، وو، ياو Hsiao, Wu, and Yao (٢٠١٤) لبناء مصفوفة متعددة الأساليب للبحث في الصدق البنائي.

أخيرًا، تم تحليل الهيكل العام لبطاريات الاختبار لتقدير نماذج المعادلة البنائية باستخدام تقديرات المربعات الموزونة المعدلة حسب المتوسط والتباين، والتي تسمح بتضمين المتغيرات ثنائية التفرع في النموذج. كما تم تعديل مستويات الدلالة من خلال تطبيق تصحيح بونفيرونى Bonferroni على مستوى ألفا الإجمالي 0.05. تم إجراء جميع الاختبارات ثنائية الذيل. اختلفت جميع المعلمات المقدرة اختلافًا دالًا عن الصفر ما لم يُنص على خلاف ذلك.

الإحصاء الوصفي وتقديرات الثبات

تمت مقارنة التأثيرات المحتملة لترتيب الاختبار (أي ما إذا كان الاستبيان في الموضع الأول أو الثاني أو الثالث ضمن بطاريات الاختبار) لكل شكل من أشكال الاستجابة على حدة. لم يلاحظ أي فرق دال بين متوسط الدرجات لأي من المقاييس ($p > .05$). يمكن رؤية الإحصاءات الوصفية لمقاييس كل شكل من أشكال الاستجابة في الجدول ٢.

الجدول (٢) الإحصائيات الوصفية لمقاييس الاستجابة

المقياس	العدد	المتوسط	الوسيط	الانحراف المعياري	اقل درجة	أعلى درجة	عدد الافراد ذوي اعلى الدرجات	عدد الافراد ذوي أقل الدرجات
مقياس التناظر البصري	300	184.00	604	2.49	184	260	3	صفر
مقياس ليكرت	300	23.61	24	3.69	9	32	8	صفر
المقياس ثنائي التفرع	300	6.43	7	1.29	صفر	8	207	3

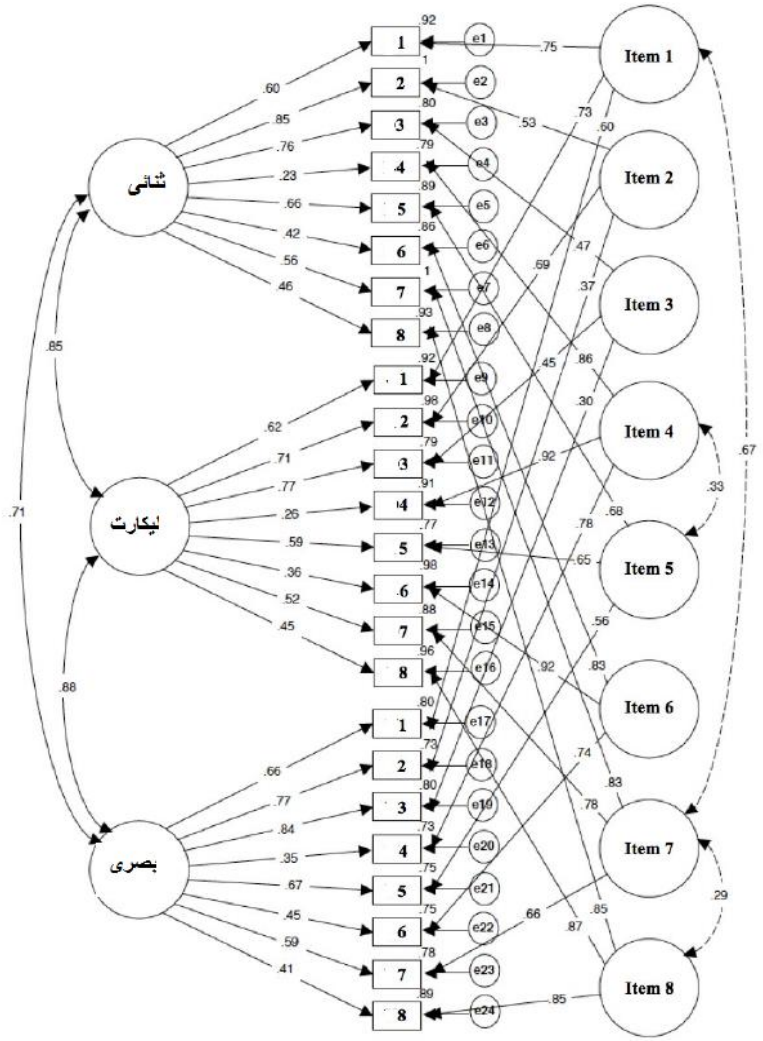
المقارنة بين متوسطات المقياس والتباين

قبل مقارنة متوسطات المقياس والتباين، تم توحيد المقاييس لنطاق محتمل من ٠-٨ (لأن كل مقياس يتكون من ثماني مفردات). كشفت تحليلات التباين للقياسات المتكررة rmANOVA عن فروق ذات دلالة إحصائية بين متوسطات المقاييس $F (2, 600) = 537.40, p < .05$. أظهرت اختبارات t أن جميع متوسط مجموع الدرجات اختلف بشكل دال، حيث أظهر مقياس التناظر البصري متوسط مجموع نقاط أعلى من مقياس ليكرت $t (865) = 4.08, p <. 01$؛ وأظهر المقياس ثنائي الاستجابة متوسط مجموع نقاط أقل من كل من مقياس التناظر البصرى ، $t (865) = 13.84, p < .01$ ، و مقياس ليكرت ، $t (865) = 16.95, p < .01.$ وهكذا، في المتوسط، وصف المستجيبون أنفسهم بأنهم أكثر إخلاصًا في مقياس التناظر البصري مقارنة بالمقياسين الآخرين، وأكثر إخلاصًا في مقياس مقياس ليكرت مقارنة بالمقياس ثنائي التفرع.

بنية المقياس والصدق التقاربى

تم مطابقة نموذج المعادلة البنائية الخطية SEM للبحث في بنية مظهر الإخلاص، كما تم قياسه بواسطة الاستبيان، وأشكال الاستجابة الثلاثة. أولاً، تم تطوير نموذج أساسي (الشكل ١، الخطوط الصلبة)، والذي يتكون من نماذج القياس لكل مقياس مع العوامل الكامنة وفقًا لذلك (تم قياس الإخلاص بمقياس استجابة ثنائي التفرع، وتم قياسه أيضًا باستخدام مقياس ليكرت الترتيبي، وباستخدام مقياس التناظر البصري. أسفر النموذج عن مطابقة ضعيفة للنموذج $\chi2(276) = 16598.14, p < .00$، بالإضافة إلى قيم مؤشر ملائمة غير مناسبة RMSEA = .075; 90% CI [.071, .079] ; WRMR = 2.093; CFI = .933

لذلك تم تعديل النموذج. كما هو مبين في الشكل ١ (الخطوط المتقطعة)، تمت إضافة ثلاث ارتباطات بين عوامل تفرد المفردة (معامل الارتباط بين المفردة ١، والمفردة ٧ = ٠,٦٧، معامل الارتباط بين المفردة ٤، والمفردة ٥ = ٠,٣٣، معامل الارتباط بين المفردة ٧، والمفردة ٨ = ٠,٢٩). تم إصلاح الفروق في عوامل تفرد المفردة (المفردة ١ - المفردة ٨) على ١. نظرًا لظهور تقديرات سلبية غير دالة للتباينات، تم تثبيت تباينات الخطأ للمتغيرين الظاهرين ٢ ثنائي التفرع، ٧ ثنائي التفرع إلى صفر وتم تثبيت مسارات كل منهما على ١.



الشكل ١. نموذج المعادلة البنائية المعدل. يشتمل النموذج على ثلاثة عوامل للإخلاص تُقاس بشكل الاستجابة المعني وعوامل ذات ثماني مفردات تشير إلى التباينات الفريدة لكل مفردة.

وفقًا لمطابقة النموذج، كان ينبغي رفض النموذج المعدل $\chi2(224) = 804.39, p$ < .00، ومع ذلك فإن قيم مؤشر المطابقة RMSEA = .055; 90% CI [.051, .059] ; WRMR = 1.52; CFI = .96. تطابق مع معايير القطع التي أوصى بها Hu and Bentler (1999)؛ RMSEA <.06؛ CFI> .95.

أشار هين واخرون Heene, Hilbert, Draxler, Ziegler, and Bühner (2011) إلى أن مؤشرات المطابقة تنخفض عندما تكون التشبعات العاملية أقل من تلك التي استخدمها هو ،بينتلر Hu and Bentler (1999). ومع ذلك، نظرًا لأن مؤشر الجذر التربيعي لمتوسط خطأ الاقتراب ومتوسط الجذر القياسي المتبقي قد تجاوزا بشكل طفيف قيم التشبعات العاملية وأشار مؤشر المطابقة المقارن إلى مطابقة مقبولة للنموذج، فقد أُعتبر النموذج مناسبًا بدرجة كافية لتحقيق أهداف الدراسة. من ناحية أخرى، تجاوز متوسط الجذر التربيعي المزون المتبقي WRMR بوضوح قيمة القطع (WRMR <1) التي اقترحها .Yu (2002) تم استخدام الارتباطات الناتجة بين عوامل أشكال الاستجابة (معامل ارتباط مقياس التناظر البصري، مقياس ليكرت الترتيبي= ٠,٨٨، معامل ارتباط مقياس التناظر البصري، المقياس ثنائي التفرع = ٠,٧١، معامل ارتباط مقياس ليكرت، المقياس ثنائي التفرع = ٠,٨٥) كتقدير للصدق التقاربى وثبت أنها جميعا مرتفع ولكن بعيدة عن بلوغ حد الكمال. كانت هناك فروق دالة بين كل الارتباطات (p <.05).

الصدق التمييزي

تم تقسيم الاستبيان إلى مقياسين منفصلين، بما في ذلك المفردات ٢ و٣ و٤ و٥ و٦ والمفردات ١ و٧ و٨ على التوالي. بعد ذلك، تم تقدير نموذج المعادلة البنائية، وهو مطابق للنموذج السابق ولكن بمقياسين بدلاً من مقياس واحد. تم تقدير ثلاث متغيرات كامنة (واحد لكل شكل استجابة) لكل مقياس (بالإضافة إلى متغيرات تفرد المفردة). يتم سرد الارتباطات الناتجة عن المتغيرات الكامنة في مصفوفة متعددة السمات - متعددة الأساليب، كما هو موضح في الجدول ٣. عدديًا، كان متوسط الارتباطات أعلى بالنسبة لنفس المقياس المقاس بأشكال استجابة مختلفة، وأقل لمقاييس مختلفة تم قياسها بنفس شكل الاستجابة، وأقل لمقاييس مختلفة تم قياسها باستخدام أشكال استجابة مختلفة.

جدول ٣. المصفوفة متعددة السمات ـ متعددة الأساليب

	المقياس ثنائي التفرع 1	مقياس ليكرت 1	مقياس التناظر البصري 1	المقياس ثنائي التفرع 2	مقياس ليكرت 2	مقياس التناظر البصري 2
المقياس ثنائي التفرع 1	(0.754)	*0.826*	*0.682*	**0.704**	0.487	0.557
مقياس ليكرت 1		(0.721)	*0.835*	0.487	**0.616**	0.620
مقياس التناظر البصري 1			(0.777)	0.490	0.613	**0.747**
المقياس ثنائي التفرع 2				(0.806)	*0.909*	*0.807*
مقياس ليكرت 2					(0.778)	*0.927*
مقياس التناظر البصري 2						(0.748)

ملحوظة: ١ = مقياس ١ (المفردات ٢، ٣، ٤، ٥، ٦)؛ ٢ = مقياس ٢ (المفردات ١، ٧، ٨). المعاملات في العناصر القطرية الرئيسية (بين قوسين) = أوميجا؛ المعاملات في المفردات خارج القطر = الارتباطات بين المتغيرات الكامنة؛ المعاملات بالنوع الغامق = المعاملات لمقاييس مختلفة تقاس بنفس شكل الاستجابة؛ المعاملات بالنوع المائل = المعاملات لنفس المقياس مقاسة بشكل استجابة مختلف.

تقديرات الاتساق الداخلي وصدق المعيار

يعرض الجدول (٤) ألفا وأوميجا لكل شكل استجابة على حدة. ازدادت ألفا Alpha بشكل رتيب مع زيادة عدد بدائل الاستجابة. من ناحية أخرى، ظلت أوميجا ثابتة بين المقياس ثنائي التفرع ومقياس ليكرت، مع قيمة أعلى قليلاً لمقياس ليكرت. أظهر مقياس التناظر البصري، كما هو الحال بالنسبة لألفا alpha، أعلى تقدير للثبات.

الجدول(٤) الخصائص السيكومترية لمقاييس الاستجابة

	المقياس ثنائي التفرع	مقياس ليكرت	مقياس التناظر البصري
ألفا	0.492 [0.439, 0.543]	0.691 [0.655, 0.719]	0.797 [0.843, 0.807]
أوميجا	0.808 [0.749, 0.862]	0.768 [0.732, 0.795]	0.817 [0.839, 0.783]
صدق المعيار	−0.146; p < .01	−0.148; p < .01	−0.149; p < .01

ملحوظة: صدق المعيار =معامل ارتباط بيرسون، P = احتمال ارتكاب خطأ من النوع الأول؛ ٩٥٪ فترات الثقة بين القوسين.

تأثير الجنس والعمر

لم يتم الكشف عن فروق ذات دلالة إحصائية لجميع أشكال الاستجابة (p> .05) من خلال اختبارات t- التي تقارن متوسط الإخلاص للمشاركين من الإناث والذكور. كما أظهرت تحليلات الانحدار الخطي أن الجنس والعمر لا يتفاعلان في التنبؤ بمجموع الدرجات. ومع ذلك، أظهر العمر علاقة موجبة ودالة مع متوسط درجات الإخلاص. كما هو موضح في الجدول (٥)، تنطبق العلاقة على جميع أشكال الاستجابة.

للبحث في تأثير الجنس، تم أيضًا تقدير جميع الخصائص السيكومترية المقدرة للمقاييس بشكل منفصل للمشاركين من الإناث والذكور. اختلفت التقديرات قليلاً عن المعاملات المقدرة للعينة بأكملها، كما هو موضح في الجدول (٦)، على الرغم من أن نمط النتائج ظل كما هو بشكل أساسي ولم تختلف الارتباطات بين المتغيرات الكامنة - التي تعمل كمؤشرات للصدق التقاربى - بشكل دال عن تلك التي تم الحصول عليها مع العينة بأكملها (p> .05). أيضًا، بالنسبة للإناث وكذلك للمشاركين الذكور، لم تختلف معاملات صدق المعيار بشكل دال، كما يتضح من الفروق في معاملات الارتباط المحولة لفيشر Fisher- transformed correlation coefficients (p> .05).

لتقدير تأثير العمر على الصدق التقاربى، تم تقدير نموذج المعادلة البنائية بما في ذلك العمر كمادة ظاهرة خارجية، مرتبطة بالمتغيرات الثلاثة الكامنة للمقاييس. أظهر العمر تشبعات دالة على جميع المتغيرات الثلاثة الكامنة (المقياس ثنائي التفرع=٠,٦٦٨، مقياس ليكرت=٠,٦٤٥، مقياس التناظر البصري=٠,٥٩٥(. أيضًا، أدى التضمين إلى انخفاض دال في الارتباطات بين المتغيرات الثلاثة الكامنة، كما يتضح من الفروق في معاملات الارتباط المحولة لفيشر (قيم-z): المقياس ثنائي التفرع- مقياس ليكرت=٠,٧٧٦ (z=-٤,٥٨٩، .01 > p)، المقياس ثنائي الاستجابة- مقياس التناظر البصري=٠,٥٦٩ (z=-٥,٠٠٩، .01 > p)، مقياس ليكرت- مقياس التناظر البصري=٠,٨٤١ (z=-٣,١٤٠، .01 > p).

الجدول (٥) انحدار مجموع درجات مقياس على العمر والجنس

	التقدير	الخطأ المعياري	ت	الدلالة
المقياس ثنائي التفرع				
التقاطع	5.393	0.3001	17.936	<.001
العمر	0.048	0.013	3.680	<.01
الجنس	0.301	0.574	0.524	غير دال
العمرX الجنس	−0.021	0.024	−0.876	غير دال
مقياس ليكرت				
التقاطع	20.869	0.855	24.403	<.001
العمر	0.132	0.037	3.556	<.01
الجنس	1.763	1.632	1.080	غير دال
العمرX الجنس	−0.112	0.069	−1.615	غير دال
مقياس التناظر البصري				
التقاطع	496.254	23.640	20.992	<.001
العمر	4.816	1.030	4.676	<.001
الجنس	60.662	45.128	1.344	غير دال
العمرX الجنس	−3.703	1.920	−1.928	غير دال

الجدول (٦) الخصائص السيكومترية للمشاركين من الإناث والذكور بشكل منفصل

	المقياس ثنائي التفرع	مقياس ليكرت	مقياس التناظر البصري
الاناث			
ألفا	0.499 [.559, .435]	0.712 [.747, .675]	0.805 [.829, .781]
أوميجا	0.833 [.893, .773]	0.796 [.828, .764]	0.827 [.862, .792]
صدق المعيار	−0.124; p < .01	−0.120; p < .01	−0.148; p < .01
الذكور			
ألفا	0.488 [.573, .394]	0.645 [.704, .581]	0.779 [.816, .739]
أوميجا	0.783 [.908, .658]	0.722 [.787, .657]	0.800 [.864, .735]
صدق المعيار	−0.172; p < .01	−0.182; p < .01	−0.133; p < .01

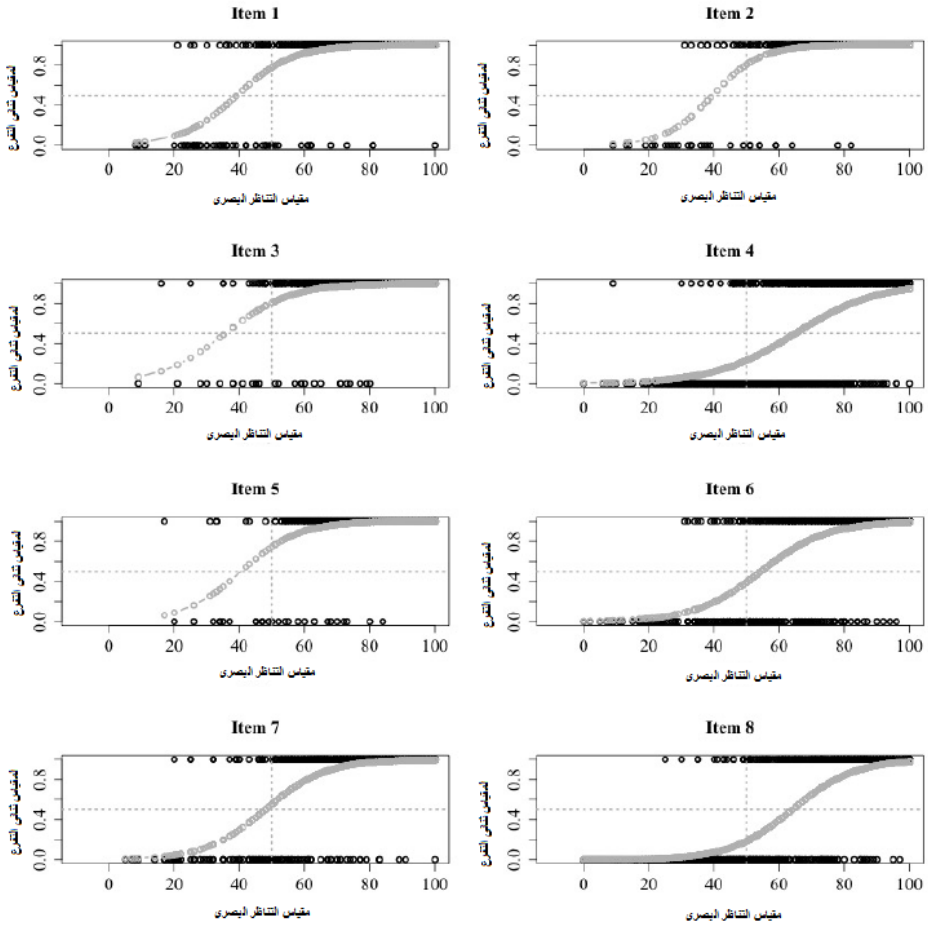
ملحوظة: صدق المعيار =معامل ارتباط بيرسون، P = احتمال ارتكاب خطأ من النوع الأول؛ ٩٥٪ فترات الثقة بين القوسين.

الانحدار من مقياس التناظر البصري الى المقياس ثنائي التفرع

تربط دالة الانحدار اللوجستي قيمة الاستجابة المتوقعة (π) لمقياس ثنائي التفرع بالبارامتر الطبيعي (η) للمتنبئ الخطي لمقياس التناظر البصري عبر:

$$\pi = \frac{\exp(\eta)}{1 + \exp(\eta)},$$

مما أدى إلى المنحنيات اللوجستية كما في الشكل (٢).



الشكل (٢). مقياس الانحدار للمقياس ثنائي التفرع على مقياس التناظر البصري. النقاط السوداء = نقاط البيانات؛ النقاط الرمادية = القيم المتوقعة للمقياس ثنائي التفرع لكل نقطة بيانات لمقياس التناظري البصري.

يوضح الجدول (٧) نتائج الانحدار اللوجستي مع مفردات المقياس ثنائي التفرع كمتغير استجابة على مفردات مقياس التناظر البصري كمتغير مشترك. يوضح الشكل (٢) كذلك الارتباط بين الاستجابات للمقياسين. تختلف معاملات الانحدار بشكل كبير بين المفردات الفردية. يشار إلى إمكانية مقارنة مراكز المقياس بين الأشكال بنقطة انعطاف المنحنيات اللوجستية: فهي تشير إلى النقطة التي تصبح عندها الاحتمالات المقدرة لاختيار "نعم" أعلى من ٥٠٪ في المقياس ثنائي التفرع. كما هو مبين في الشكل (٢)، يتم إزاحة هذه النقطة من النقطة المركزية (٥٠ مم) من مقياس التناظر البصري لمعظم المفردات.

جدول (٧) الانحدار من مقياس التناظر البصري الى المقياس ثنائي التفرع

المفردة	التقاطع	الانحدار
1	−4.617	0.118
2	−5.466	0.138
3	−3.563	0.100
4	−5.207	0.080
5	−4.608	0.114
6	−5.383	0.099
7	−5.191	0.108
8	−6.725	0.105

ملحوظة. التقاطع = تقاطع المتنبئ الخطي للانحدار اللوجستي؛ المنحدر = ميل المتنبئ الخطي للانحدار اللوجستي

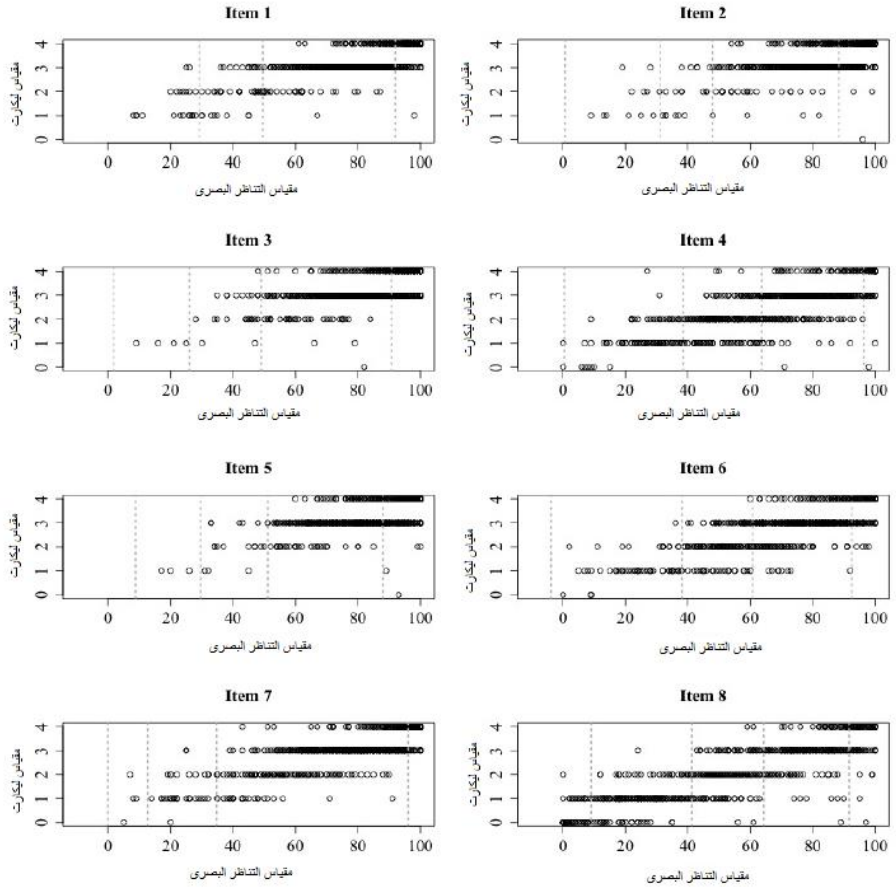
الانحدار من مقياس التناظر البصري الى مقياس ليكرت

يوضح الجدول (٨) معاملات الانحدار لنموذج الاحتمالات النسبية مع مفردات مقياس ليكرت كمتغير استجابة ومفردات مقياس التناظر البصري كمتغير مشترك. توضح العتبات المقدرة النقطة التي تصبح فيها احتمالات اختيار الفئة المعنية في مقياس ليكرت أعلى من الفئة السابقة. كما هو موضح بيانياً في الشكل (٣)، تختلف العتبات بشكل دال بين المفردات وهي بعيدة كل البعد عن مسافات متساوية. تم أيضًا مطابقة النماذج ذات العتبات المتساوية الثابتة وأسفرت عن مطابقة ضعيفة لكل مفردة، كما يُقاس باختبارات نسبة الاحتماليةlikelihood-ratio tests (p <.001).

جدول (٨) **الانحدار من مقياس التناظر البصري الى مقياس ليكرت**

المفردة	العتبة 1	العتبة 2	العتبة 3	العتبة 4	الانحدار
1	-	4.697	7.982	14.816	0.161
2	0.092	4.197	6.433	11.903	0.135
3	0.234	3.279	6.172	11.428	0.126
4	0.058	4.398	7.272	10.984	0.114
5	1.230	4.126	7.121	12.234	0.139
6	0.485−	4.920	7.811	11.897	0.129
7	0.088−	4.908	8.281	13.261	0.144
8	1.138	5.167	8.056	11.490	0.125

ملحوظة. العتبة = النقطة التي يصبح عندها احتمال اختيار الفئة التالية أعلى من احتمال اختيار الفئة السابقة؛ التقاطع = تقاطع المتنبئ الخطي للانحدار اللوجستي؛ المنحدر = ميل المتنبئ الخطي للانحدار اللوجستي.



الشكل (٣). انحدار مقياس ليكرت على مقياس التناظر البصري. النقاط السوداء = نقاط البيانات؛ الخطوط الرمادية = الحدود المقدرة بين الفئات على مقياس ليكرت

مناقشة النتائج

بحثت الدراسة الحالية في العلاقة بين المقياس ثنائي التفرع، ومقياس ليكرت من ٥ نقاط، ومقياس التناظر البصري. أظهرت مقاييس الاستجابة الثلاثة زيادة رتيبة في ألفا مع زيادة في احتمالات الاستجابة، بينما لم توجد فروق دالة في أوميجا. أيضًا، لم يظهر صدق المعيار أي فروق بين أشكال الاستجابة. بالنظر إلى الصدق التقاربى، كانت الارتباطات بين المقاييس عالية ولكنها بعيدة عن الوصول الى الكمال. في المصفوفة متعددة الأساليب ومتعددة السمات، أظهرت معاملات الارتباط التي تشير

إلى الصدق التمييزي نفس النمط لجميع أشكال الاستجابة. ثبت أن أنماط الاستجابة بين مقياس التناظر البصري والمقياسين مع الفئات الثابتة مترابطة - ومع ذلك، لوحظت فروق دالة بين المفردات الفردية للمقياس.

في حين كان للعمر تأثير دالة على الصدق التقاربى والدرجة الكلية للمقاييس، إلا أنه لم يكن للجنس تأثير دال. كذلك، لم يتفاعل الجنس والعمر في التنبؤ بنتيجة الإخلاص الكلي. وهذا يتوافق مع نتيجة دراسة Lehmann, Denissen, Allemand, and Penke (2013) الذين أشار الى أنه لم يكن التأثير الرئيسي للجنس ولا تفاعلات العمر × الجنس دالًا في حالة الضمير.

الثبات وصدق المعيار

تتفق النتيجة التي مفادها أن زيادة ألفا دالة رتيبة لمستوى التمايز في مقياس الاستجابة مع معظم الدراسات المتعلقة بالعلاقة بين أشكال الاستجابة والثبات (Cicchetti et al., 1985; Oaster, 1989; Weng, 2004). ومع ذلك، لم توجد فروق في أوميجا بين أشكال الاستجابة، وهذا يتفق مع ما توصلت اليه نتائج Maydeu-Olivares et al. (2009)، حيث أشارت النتائج إلى أنه مع زيادة عدد بدائل الاستجابة، (١) يزداد الاتساق الداخلي، (٢) لم يكن هناك تأثير على الصدق التقاربى، و (٣) انخفاض جودة المطابقة. تؤكد النتيجة على أهمية نموذج القياس (المفترض) والإطار السيكومترى. أوضح Gignac (2014) كيف تؤدي المقاييس متعددة الأبعاد بشكل خاص إلى التقليل من الثبات بواسطة ألفا alpha، والتي، مع الأخذ في الاعتبار الأبعاد الإضافية المسئولة عن ارتباطات المفردات ١ و٧ و٨ أو ارتباط المفردات ٤ و٥، يبدو أنه يمكن أن تنطبق على الدراسة الحالية.

يوضح فحص معادلات ألفا وأوميجا هذا التناقض: بينما يتم حساب ألفا بناءً على تباينات المفردات والمقاييس (التي اختلفت بشكل دال في الدراسة الحالية، بسبب المقاييس المختلفة وحتى بعد توحيد المجموع الكلى للدرجات)، تستخدم أوميجا التشبعات العاملية المعيارية والفروق الفريدة للمفردات في إطار نموذج العامل الكامن من أجل تقدير ثبات المقياس (Revelle & Zinbarg, 2009). نظرًا لأن التشبعات العاملية بين المقاييس الثلاثة (مقابل داخل المقاييس) يمكن مقارنتها في النموذج المستخدم في الدراسة الحالية، فمن المتوقع عدم تباين أوميجا.

يقلل ألفا من قيمة الثبات نظرًا لأنه يعتمد على تباينات المفردات والمقاييس (عن طريق الاتساق الداخلي) إذا كانت التشبعات العاملية للمفردات مختلفة: يشير الاختلاف في التشبعات العاملية إلى أن المفردات تقيس المتغير الكامن على مقياس مختلف، وبالتالي، فإن تباياناتها لا تشير إلى تباين المتغير الكامن بنفس الدرجة (Graham, 2006). نظرًا لعدم الوفاء الواضح أو عدم تحقق افتراض التشبعات العاملية المتساوية داخل كل مقياس في الدراسة الحالية، لا يمكن افتراض التكافؤ (الأساسي) وبالتالي ليست التقديرات الأقل لثبات ألفا مقارنة بأوميجا مفاجئة. بالإضافة إلى ذلك، يمكن أخذ الأبعاد المتعددة المحتملة للمقياس، كما هو موضح في الدراسة الحالية، في الاعتبار لتقدير الاتساق الداخلي (Dunn et al., 2014) من أجل الحصول على تقدير أكثر قوة للثبات - على الرغم من تناقص أوميجا في ظل حالات الأبعاد المتعددة (Socan, 2000).

إن النتيجة التي تفيد بأن صدق المعيار لم تظهر زيادة مع العدد المتزايد من احتمالات الاستجابة تتوافق مع نتائج Preston and Colman (2000). ومع ذلك، أظهر الإخلاص ارتباطًا ضعيفًا إلى حد ما مع المعيار مقارنة بالنتائج السابقة (Judge et al.,1997). ومع ذلك، فإن النتائج التي تم الحصول عليها تتعارض مع الافتراض القائل بوجود مكاسب محددة في صدق المعيار من خلال استخدام مقياس ليكرت ذي ٥ نقاط مقارنة بالمقياس ثنائي التفرع، على النحو الذي اقترحه Hancock and Klockars (1991).

بنية النموذج والصدق التقاربى

كان من المتوقع أن تعبر الاستبيانات عن نفس المتغير الكامن ("الإخلاص")، بغض النظر عن مقياس الاستجابة. ومع ذلك، فقد ثبت أن الارتباطات بين العوامل الكامنة الثلاثة عالية ولكنها بعيدة عن الوصول لدرجة الكمال (خاصة عند أخذ العمر في الاعتبار). تلقي الارتباطات التي تم ملاحظتها بظلال من الشك على الصدق التقاربى للمقياس (Bühner,2011): على الرغم من الجودة السيكومترية المماثلة، لا يبدو أن الاستبيانات تعبر نفس المعلومات إذا تم تغيير شكل الاستجابة. وفقًا لسكوارز (Schwarz ,1999)، تلعب أشكال الاستجابة، بالإضافة إلى توفير استجابة للمفردة، دورًا في جعل المفردة مفهومة للمستجيبين من خلال التلميح إلى نية مُعدّ الاستبيان.

تدعم النتائج الحالية وجهة نظر سكوارز بقوة: من المفترض أن المفردات المتطابقة قد تم تفسيرها بشكل مختلف عند تغيير شكل الاستجابة، على الأقل فيما يتعلق

بالاستجابة التي أثارتها. الأهم من ذلك، تشير هذه النتيجة إلى أنه بغض النظر عن كيفية قياس هذا الاستبيان لسمة ما، فإنه يقيس أبعادًا أخرى لا يمكن تفسيرها بخطأ قياس عشوائي - وبالتالي لا يتم قياس السمة بشكل خالص.

يجب التأكيد على هذه المشكلة بشكل أكبر من خلال حقيقة أن الارتباطات بين المقاييس تختلف اختلافًا دالا عن بعضها البعض، مما يشير إلى أن الأبعاد الإضافية يتم قياسها بواسطة أشكال الاستجابة. لم يذكر Maydeu- Olivares et al. (2009) أي فروق بين ارتباطات المتغيرات الكامنة التي تمثل أشكال الاستجابة. ومع ذلك، لم تتضمن نتائجهم مقياسًا للتناظر البصري قد يتفاعل بشكل إضافي مع الأبعاد الإضافية للمقاييس.

يأتي المزيد من الدعم لهذا الافتراض من النتيجة التي مفادها أن متوسط الدرجة يختلف اختلافًا دالا بين جميع أشكال الاستجابة. تتفق هذه النتيجة مع نتائج مع نتائججويات واخرين Guyatt, Townsend, Berman, and Keller (1987)، الذين أعلنوا عن متوسط درجة أعلى بكثير لمقياس التناظر البصري مقارنة بمقياس ليكرت في استبيان حول الوظائف الجسدية. يمكن ملاحظة تأثير السقف الواضح أيضًا في المقياس ثنائي التفرع حيث وصل إجمالي عدد المشاركين إلى ٢٠٧ مشاركًا إلى أقصى درجة ثمانية مقارنة بثلاثة مشاركين فقط سجلوا صفرًا على هذا المقياس. في المقابل، وصل ثلاثة وأحد عشر مشاركًا فقط إلى الحد الأقصى للدرجات في مقياس التناظر البصري ومقياس ليكرت، على التوالي.

الصدق التمييزي

أظهرت المصفوفة متعددة الأساليب ومتعدد السمات multitrait-multimethod matrix، التي تم إنشاؤها في هذه الدراسة الصدق التمييزي: عدديًا، كان متوسط الارتباطات أعلى لنفس المقياس الذي تم قياسه بأشكال استجابة مختلفة، وأقل للمقاييس المختلفة باستخدام نفس شكل الاستجابة، وأدنى للمقاييس المختلفة المقاسة بأشكال مختلفة. هذه النتيجة تعني الصدق التمييزي للمقياسين، وبالتالي تتحدى بوضوح فرضية أحادية البعد للاستبيان. ومع ذلك، يؤكد النمط الذي تم الحصول عليه على فكرة أن أشكال الاستجابة المختلفة يمكن اعتبارها طرق قياس مختلفة وأن

الارتباطات بين السمات يمكن، بالتالي، أن تكون بمثابة مقياس للصدق التقاربى. الأهم من ذلك، أن هذا المدخل متعدد الأساليب لا يمكن تفسيره على أنه يبحث في الصدق التمييزي لمقياس الإخلاص dutifulness ككل، حيث تم تقسيم هذا المقياس إلى مقاييس فرعية منفصلة.

مقارنة الفترات الزمنية الفاصلة والمراكز للمقياس

أظهرت تحليلات الانحدار من المقياس ثنائي التفرع على مقياس التناظر البصري نقاط التواء متفاوتة بشدة للوظائف اللوجستية الناتجة بين المفردات الثمانية. تشير نقطة التواء المنحنى اللوجستي إلى النقطة التي يصبح عندها، في المقياس ثنائي التفرع، من المرجح أن يتم الموافقة مع العبارة المعروضة في الاستبيان بدلاً من عدم الموافقة (أي أن القيمة ١ تصبح أكثر احتمالية من القيمة صفر، أو "نعم" يصبح أكثر احتمالا من "لا"). يمكن النظر إلى هذه العتبة على أنها مركز مقياس الاستجابة ثنائي التفرع وينبغي، إذا اظهر المقياس ثنائي التفرع ومقياس التناظر البصري استجابات قابلة للمقارنة، يتوافق مع مركز مقياس التناظر البصري (أي القيمة ٥٠)، كما هو الحال بالنسبة للمفردة ٧. معظم المفردات (المفردات ١ و٢ و٣ و٥)، مع ذلك، تُظهر تحول دالة الانحدار إلى اليسار، مما يعني أنه من المرجح أن يختار المستجيبون "نعم" في المقياس ثنائي التفرع قبل أن يختاروا قيمة على مقياس التناظر البصري الأقرب إلى "اتفاق" بدلاً من "عدم الاتفاق" (أي أعلى من ٥٠). بالنسبة للمفردات التي تُظهر تحولًا إلى اليمين (المفردات ٤ و٦ و٨)، يتم توقع قيمة صفر في المقياس ثنائي التفرع لقيم مقياس التناظر البصري أعلى من ٥٠. على مستوى المفردة، لا يتوافق تصور الفاصل الزمني بين الاتفاق وعدم الاتفاق في الشكلين أو الصيغتين.

أظهر تحليل الانحدار من مقياس ليكرت على مقياس التناظر البصري أيضًا نمطًا سطحيًا إلى حد ما بين المفردات: تختلف العتبات المقدرة، التي تشير إلى النقطة التي تتجاوز فيها احتمالات اختيار فئة في مقياس ليكرت تلك الخاصة بالسابقة، بشدة بين المفردات فيما يتعلق بموقعهم والمسافة والنسبة. يشير هذا إلى أن الارتباط بين الفترات الفاصلة المدركة لمقياس التناظر البصري ومقياس ليكرت لا يمكن تعميمه ولكنه يعتمد على المفردة المقابلة. تعتبر الأحجام المختلفة للفترات الفاصلة بين العتبات، الواضحة في كل مفردة، مثيرة للاهتمام بشكل خاص: يبدو أن المستجيبين يحكمون على الفترات الفاصلة بين مستويات الاخلاص المختلفة بشكل مختلف اعتمادًا على شكل الاستجابة. تشير النتيجة إلى أن افتراض مقياس الفاصل الزمني لا ينطبق على مقياس واحد على الأقل من المقياسين، حيث يجب أن تكون العتبات

متساوية البعد إذا كان افتراض مقياس الفاصل ثابتًا على كل من مقياس ليكرت ومقياس التناظر البصري. ومع ذلك، على الأرجح، لا ينطبق هذا الافتراض على أي من المقياسين، كما أوضح ذلك باحثون أخرون (Clason & Dor- Clason, 1994; Goldstein & Hersen, 2000; Wewers & Lowe, 1990). ومع ذلك، فإن النتيجة القائلة بأن العلاقة الرتيبة بين عتبات مقياس ليكرت ومقياس التناظر البصري يمكن افتراضها تتفق مع ما توصل اليه Ferrando (2003)، حيث وجد توزيع الكثافة للمقياسين قابلاً للمقارنة.

إجمالا، توضح تحليلات الانحدار كيف تختلف الفواصل الزمنية والمراكز المدركة للمقياس الكامن المفترض بشدة بين أشكال الاستجابة. أيضًا، لا يمكن افتراض أن الارتباط بين المقاييس متساوٍ للمفردات.

تأثير العمر

كانت هناك علاقة موجبة ودالة بين العمر والاخلاص على مستوى الدرجة الكلية وكذلك على مستوى المتغيرات الكامنة. تتفق هذه النتيجة مع نتائج الأبحاث السابقة حول البعد الشخصي للضمير وخاصة جوانبها خلال العقد الأول من القرن الحادي والعشرين (McCrae, Martin, & Costa, 2005). لذلك ليس من المستغرب أن تكون العلاقات المتبادلة بين المتغيرات الكامنة التي تمثل الإخلاص قد انخفضت بعد أخذ العمر في الاعتبار.

الاستنتاجات والتضمينات التربوية

نظرًا لاعتماد الدراسة الحالية على تصميم القياسات المتكررة، فقد ألقت الضوء على العديد من القضايا المهمة: أولاً: لا يبدو أن الجودة السيكومترية في استبيان الشخصية تزداد بمقياس استجابة أكثر تمايزًا - ولا من حيث الثبات (عن طريق الاتساق الداخلي) ولا من حيث الصدق التقاربى أو صدق معيار. على الرغم من أن ألفا أظهرت زيادة كبيرة مع زيادة احتمالات الاستجابة، فإن أوميجا لم تظهر زيادة كبيرة مع زيادة احتمالات الاستجابة. يمكن تفسير جزء كبير من الارتباط بين المتغيرات الكامنة للمقاييس التي تم قياسها بواسطة أشكال الاستجابة المختلفة حسب العمر، والتي تنبأت أيضًا بشكل كبير بمتوسط الدرجة في الإخلاص.

ومع ذلك، من أجل الاستخدام العملي، من المهم مراعاة أنه كان من السهل فهم المقياس ثنائي التفرع ومقياس ليكرت للمستجيبين، على عكس مقياس التناظر البصري: كما لاحظ بالفعل العديد من الباحثين (Couper et al., 2006; Ferrando, 2003; Flynn et al., 2004) أظهر العديد من الأشخاص أن نقص الخبرة مع مقياس التناظر البصري أدى إلى بعض الارتباك وكان هناك حاجة إلى شرح إضافي في بعض الحالات. ثانيًا: يجب استخدام المقياس ثنائي التفرع فقط إذا لم يكن من المتوقع أن يحصل الأشخاص على مستويات عالية على مقياس معين، حيث لوحظت تأثيرات سقف واضحة.

أخيرًا، يمكن أن تنحدر استجابات المقياس ثنائي التفرع ومقياس ليكرت على مقياس التناظر البصري ولكنها تميل إلى التباين بشكل كبير بين المفردات والأشكال الفردية. لذلك من المهم مراعاة أن الفترات الزمنية الفاصلة بين المستويات المختلفة وكذلك مراكز المقاييس يبدو أنه يتم الحكم عليها بشكل مختلف اعتمادًا على شكل الاستجابة. يمكن افتراض العلاقة بين الأشكال الثلاثة، على الأقل على مستوى المفردة، على أنها رتيبة بطريقة ترتيبية، ولكن ليست خطيًة. لذلك يجب أن يخضع اختيار شكل الاستجابة إلى دراسة مكثفة فيما يتعلق بالسياق والوقت المتاح وخبرة المستجيبين مع المقياس.

المراجع

حجاج غانم أحمد (٢٠١١). تأثير عدد بدائل ليكرت على الخصائص السيكومترية للمقياس النفسي وافتراضات التصميم العاملي ثنائي الاتجاه. **مجلة الإرشاد النفسي**، جامعة عين شمس - مركز الإرشاد النفسي، ع ٢٩، ص ص ٥٦ – ١٣١.

سامي محمد ملحم (٢٠٠٩). **القياس والتقويم في علم النفس**، دار المسيرة للنشر والتوزيع والطباعة عمان

صلاح الدين علام (٢٠٠٦). **القياس والتقويم التربوي والنفسي أساسياته وتطبيقاته وتوجهاته المستقبلية**، دار الفكر العربي، القاهرة.

Alkadi, L., Emad, T., Tarig, Mohamud, M. and Farook, M. (2022) A Likert Scale Versus a Visual Analogue Scale and the Participant Response: A cross Sectional Study. *Journal of International Dental and Medical Research* 15(2):255-262.

Berth, H., Goldschmidt, S., Ostendorf, F., & Angleitner, A. (2006). NEO-PI-R. NEO-Persönlichkeitsinventar nach Costa und McCrae. Revidierte Fassung [NEO-PI-R personality inventory, adapted from Costa & McCrae. Revised version]. *Diagnostica*, *52*, 95-99.

Birkett, N. J. (1986). Selecting the number of response categories for a Likert-type scale. In *Proceedings of the American Statistical Association, Section on Survey Research Methods* (pp. 488-492). Washing- ton, DC: American Statistical Association.

Bond, T.G., & Fox, C.M. (2001). *Applying the Rasch model: Fundamental measurement in the human sciences*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Bowling, A. (2005). Mode of questionnaire administration can have serious effects on data quality. *Journal of Public Health, 27*(3), 281–291.

Briggs, M., & Closs, J. S. (1999). A descriptive study of the use of visual analogue scales and verbal rating scales for the assessment of postoperative pain in orthopedic patients. *Journal of Pain and Symptom Management, 18,* 438-446. doi:10.1016/S0885-3924(99)00092-5

Brunier, G., & Graydon, J. (1996). A comparison of two methods of measuring fatigue in patients on chronic haemodialysis: Visual analogue versus Likert scale. *International Journal of nursing Studies*, *33*, 338-348.

Bühner, M. (2011). *Einführung in die Test-und Fragebogenkonstruktion* [Introduction to test and questionnaire construction]. München, Germany: Pearson-Education.

Byrom, B., Elash, C.A., Eremenco, S. *et al.* (2022). Measurement Comparability of Electronic and Paper Administration of Visual Analogue Scales: A Review of Published Studies. *Ther Innov Regul Sci* **56,** 394–404. https://doi.org/10.1007/s43441-022-00376-2

Callegaro, M., Manfreda, K. L., & Vehovar, V. (2015). *Web survey methodology*. Los Angeles, CA: Sage.

Cicchetti, D. V., Shoinralter, D., & Tyrer, P. J. (1985). The effect of number of rating scale categories on levels of interrater reliability: A Monte Carlo investigation. *Applied Psychological Measurement, 9*, 31-36. doi:10.1177/014662168500900103

Clason, D. L., & Dormody, T. J. (1994). Analyzing data measured by individual Likert-type items. *Journal of Agricultural Education, 35*, 31-35.

Couper, M. P., Tourangeau, F., Conrad, F. Crawford,S.D. (2004). What they see is what we get. Response options for web surveys. *Social Science Computer Review*, 22(1): 111–127.

Couper, M. P., Tourangeau, R., Conrad, F. G., & Singer, E. (2006). Evaluating the effectiveness of visual analog scales: A web experiment. *Social Science Computer Review*, *24*, 227-245. doi:10.1177/0894 439305281503

Dawes, J. (2008). Do data characteristics change according to the number of scale points used? An experi- ment using 5-point, 7-point and 10-point scales. *International Journal of Market Research*, *51*, 61- 77.

Dourado GB, Volpato GH, de Almeida-Pedrin RR, Pedron Oltramari PV, Freire Fernandes TM, de Castro Ferreira Conti AC. (2021) Likert scale vs visual analog scale for assessing facial pleasantness. Am J Orthod Dentofacial Orthop. 2021 Dec;160(6):844-852. doi: 10.1016/j.ajodo.2020.05.024.

Dunn, T. J., Baguley, T., & Brunsden, V. (2014). From alpha to omega: A practical solution to the perva- sive problem of internal consistency estimation. *British Journal of Psychology*, *105*, 399-412. doi: 10.1111/bjop.12046

Dutheil F, Pereira B, Bouillon-MinoisJ- B, et al. (2022). Validation of Visual Analogue Scales of job demand and job control at the workplace: a cross-sectional study. BMJ Open ,12: e046403. doi:10.1136/ bmjopen-2020-046403

Faulbaum, F., Prüfer, P., & Rexroth, M. (2009). *Was ist eine gute Frage? Die systematische Evaluation der Fragenqualität* [What is a good question? The systematic evaluation of question quality]. Wiesbaden, Germany: Springer.

Ferrando, P. J. (2003). A kernel density analysis of continuous typical-response scales. *Educational and Psychological Measurement*, *63*, 809–824.

Flynn, D., van Schaik, P., & van Wersch, A. (2004). A comparison of multi-item Likert and visual analogue scales for the assessment of transactionally defined coping function. *European Journal of Psychological Assessment*, *20*(1), 49-58.

Funke, F & Reips, U.-D. (2012). Why semantic differentials in Web-based research should be made from visual analogue scales and not from 5-point scales. Field Methods. 24: 310–327. doi:10.1177/1525822X12444061

Gignac, G. E. (2014). On the inappropriateness of using items to calculate total scale score reliability via coefficient alpha for multidimensional scales. *European Journal of Psychological Assessment*, *30*, 130-193. doi:10.1027/1015-5759/a000181

Goggin, S., & Stoker, L. (2014). Optimal scale length and single-item attitude measures: Evidence from simulations and a two-wave experiment. *American Political Science Association Annual Meeting Paper*, Washington, DC. Retrieved from http://www.ssrn.com/link/APSA-2014.html

Goldstein, G., & Hersen, M. (2000). *Handbook of psychological assessment*. Oxford, UK: Elsevier.

Graham, J. M. (2006). Congeneric and (essentially) tau-equivalent estimates of score reliability what they are and how to use them. *Educational and Psychological Measurement*, *66*, 930-944. doi:10.1177/ 0013164406288165

Grant, S., Aitchison, T., Henderson, E., Christie, J., Zare, S., McMurray, J., & Dargie, H. (1999). A com- parison of the reproducibility and the sensitivity to change of visual analogue scales, Borg scales, and Likert scales in normal subjects during submaximal exercise. *Chest Journal, 116*, 1208-1217. doi:10.1378/chest.116.5.1208

Guyatt, G. H., Townsend, M., Berman, L. B., & Keller, J. L. (1987). A comparison of Likert and visual an- alogue scales for measuring change in function. *Journal of Chronic Diseases, 40*, 1129-1133. doi: 10.1016/0021-9681(87)90080-4

Hancock, G. R., & Klockars, A. J. (1991). The effect of scale manipulations on validity: Targetting frequency rating scales for anticipated performance levels. *Applied Ergonomics*, *22*, 147-154. doi:10. 1016/0003-6870(91)90153-9

Hasson, D., & Arnetz, B.B. (2005). Validation and findings comparing VAS vs. Likert scales for psychosocial measurements. *International Electronic Journal of Health Education, 8*, 178-192.

Hewson, C., Vogel, C., & Laurent, D. (2015). *Internet research methods*. London: Sage.

Jaeschke, R., Singer, J., & Guyatt, G. H. (1990). A comparison of seven-point and visual analogue scales: Data from a randomized trial. *Controlled Clinical Trials*, *11*, 43-51. doi:10.1016/0197-2456(89)90005-6

Jenkins, C. R. and Dillman, D. A. (1995). Towards a theory of self-administered questionnaire design. In L. Lyberg, P. Biemer, M. Collins, et al. eds., *Survey Measurement and Process Quality. New York: Wiley, 165–196.*

Jones, W.& Loe, S. (2014) Optimal Number of Questionnaire Response Categories More May Not Be Better. SAGE Open 3(2) DOI:10.1177/2158244013489691.

Judge, T. A., Martocchio, J. J., & Thoresen, C. J. (1997). Five-factor model of personality and employee absence. *Journal of Applied Psychology, 82*, 745-755. doi:10.1037/0021-9010.82.5.745

Hancock, G. R., & Klockars, A. J. (1991). The effect of scale manipulations on validity: Targetting fre- quency rating scales for anticipated performance levels. *Applied Ergonomics*, *22*, 147-154. doi:10. 1016/0003-6870(91)90153-9

Heene, M., Hilbert, S., Draxler, C., Ziegler, M., & Bühner, M. (2011). Masking misfit in confirmatory fac- tor analysis by increasing unique variances: A cautionary note on the usefulness of cutoff values of fit indices. *Psychological methods*, *16*, 319-336.

Hsiao, Y.-Y., Wu, C.-H., & Yao, G. (2014). Convergent and discriminant validity of the WHOQOL-BREF using a multitrait-multimethod approach. *Social Indicators Research*, *116*, 971-988. doi:10.1007/ s11205-013-0313-z

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conven- tional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*, 1-55. doi:10.1080/10705519909540118

Lehmann, R., Denissen, J. J., Allemand, M., & Penke, L. (2013). Age and gender differences in motiva- tional manifestations of the Big Five from age 16 to 60. *Developmental Psychology*, *49*, 365-383. doi:10.1037/a0028277

Lozano, L. M., García-Cueto, E., & Muñiz, J. (2008). Effect of the number of response categories on the reliability and validity of rating scales. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, *4*, 73-79. doi:10.1027/1614-2241.4.2.73

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 intra-individual effects. *Behavior Research Methods*, *41*, 295-308. doi:10.3758/BRM.41.2.295

McCrae, R. R., Martin, T. A., & Costa, P. T. (2005). Age trends and age norms for the NEO Personality Inventory-3 in adolescents and adults. *Assessment*, *12*, 363-373. doi:10.1177/1073191105279724

Oaster, T. R. F. (1989). Number of alternatives per choice point and stability of Likert-type scales. *Percep- tual and Motor Skills*, *68*, 549-550. doi:10.2466/pms.1989.68.2.549

Preston, C. C., & Colman, A. M. (2000). Optimal number of response categories in rating scales: Reliabil- ity, validity, discriminating power, and respondent preferences. *Acta Psychologica*, *104*, 1-15. doi: 10.1016/S0001-6918(99)00050-5

Reips, U.-D. (2002). Standards for Internet-based experimenting. *Experimental Psychology*, 49, 243–256.

Reips, U.-D., & Birnbaum, M. H. (2011). Behavioral research and data collection via the Internet. In K.-L. Vu & R. W. Proctor (Eds.), *The handbook of human factors in web design* (2nd ed., pp. 563–585). Mahwah, New Jersey: Erlbaum.

Reips, U.-D., & Funke, F. (2008). Interval-level measurement with visual analogue scales in Internet-based research: VAS generator. *Behavior Research Methods*, 40, 699–704. doi:10.3758/BRM.40.3.699

Revelle, W., & Zinbarg, R. E. (2009). Coefficient's alpha, beta, omega, and the glb: Comments on Sijtsma. *Psychometrika*, *74*, 145-154. doi:10.1007/s11336-008-9102-z

Schwarz, N. (1999). Self-reports: How the questions shape the answers. *American Psychologist*, *54*, 93- 105. doi:10.1037/0003-066X.54.2.93

Socan, G. (2000). Assessment of reliability when test items are not essentially tau-equivalent. *Development in Survey Methodology*, *15*, 23-35.

Shulman, H. C., & Boster, F. J. (2014). Effect of test-taking venue and response format on political knowledge tests. *Communication Methods and Measures*, *8*, 177-189.

Simms, L. J., Zelazny, K., Williams, T. F., & Bernstein, L. (2019). Does the Number of Response Options Matter? Psychometric Perspectives Using Personality Questionnaire Data. *Psychological Assessment*. Advance online publication. http://dx.doi.org/10.1037/pas0000648

Smith, T. (1995). Little things matter: a sampler of how differences in questionnaire format can affect survey responses. In American Statistical Association ed., Proceedings of the American Statistical Association, Survey Research Methods Section. New York: Springer, 187–201.

Toepoel, V., Das, J. W. M., and Van Soest, A. H. O. (2009). Design of web questionnaires: the effect of layout in rating scales. *Journal of Official Statistics*, 25(4): 509–528.

Toepoel, V. and Dillman, D. A. (2011). Words, numbers, and visual heuristics in web surveys. Social Science Computer Review, 29(2): 193–207.

Tucker-Seeley, R. (2008). The Effects of Using Likert vs. Visual Analogue Scale Response Options on the Outcome of a Web-based Survey of 4th Through 12th Grade Students, Boston College, http://hdl.handle.net/2345/2624.

Weng, L.-J. (2004). Impact of the number of response categories and anchor labels on coefficient alpha and test-retest reliability. *Educational and Psychological Measurement*, *64*, 956-972. doi:10.1177/0013 164404268674

Wewers, M. E., & Lowe, N. K. (1990). A critical review of visual analogue scales in the measurement of clinical phenomena. *Research in Nursing & Health*, *13*, 227-236. doi:10.1002/nur.4770130405

Wuensch, K. (2005). What is a Likert Scale? and How Do You Pronounce 'Likert? East Carolina University.

Yu, C. Y. (2002). *Evaluating cutoff criteria of model fit indices for latent variable models with binary and continuous outcomes* (Doctoral

dissertation, University of California Los Angeles). Retrieved from http://www.statemodel.com/download/Yudissertation.pdf

Abstract

The influence of different forms of response on the psychometric properties of the Internet-based personality scale using Likert, visual analog and dichotomous scales.

The study aimed to investigate the effect of different response formats on the psychometric properties of the Internet-based personality scale using Likert scale, visual analogue scale and dichotomous scale. The present study, therefore, uses a repeated measures design (allowing for the analysis of convergent validity and regression analysis between scales) including a DS, a 5-point LTS, and a VAS to compare the response formats regarding their psychometric properties and scale intervals. In order to gain information about the use of a VAS in a personality questionnaire, the eight items measuring the facet "dutifulness" of the dimension "conscientiousness" (Berth, Goldschmidt, Ostendorf, & Angleitner, 2006) are used. Criterion-related validity and coefficients of reliability are compared between the different response formats. Of all five facets of conscientiousness, dutifulness, in particular, has shown the numerically highest correlation with absent days from work. Also, the perceived intervals and central points of the VAS are compared to the dichotomous and the 5-point scale to obtain information about the comparability of the response scales' category thresholds and centers. Concerning estimates of reliability, it was shown that Cronbach's alpha increased with the number of response alternatives of the scales, while McDonald's omega and validity coefficients remained steady. It is argued that omega provides the more adequate estimate of internal consistency. Generalized linear regression analyses indicated that the participants judged the intervals between varying levels of dutifulness on the VAS differently from the other two scales. It is concluded that the choice of response format should not be exclusively based on desired psychometric properties but rather on practical considerations.

Keywords: response formats, psychometric properties, personality scale, Likert scale, visual analogue scale, dichotomous scale.