تكافؤ القياس للصورة الإلكترونية في مقابل التقليدية لمقياس روسنبرج لتقدير الذات

د. منى عبد الفضيل الألفي & د. محمود علي موسى

http://dx.doi.org/10.29009/ijres.2.3.10

تكافؤ القياس للصورة الإلكترونية في مقابل التقليدية لمقياس روسنبرج لتقدير الذات

د. منى عبد الفضيل الألفي، مدرس علم النفس التربوي، كلية التربية بالإسماعيلية، جامعة قناة السويس، مصر

Monaelalfy2013@yahoo.com

د. محمود علي موسى، مدرس علم النفس التربوي، كلية التربية بالإسماعيلية، جامعة قناة السويس، مصر

mahmod567@yahoo.com

قدمت للنشر في ١٨/ ١٢/ ٢٠١٨ م **قبلت للنشر في ١٧/ ٣/ ٢٠١٩ م**

ملخص: هدفت الدراسة إلى المفاضلة بين النسخة الورقية التقليدية والنسخة الإلكترونية لمقياس تقدير الذات لـــ Rosenberg (1965) باستخدام تصميم بين الموضوعات للتحقق من تكافؤ القياس. وقد طبقت الدراسة على عينتين تم اختيارهما بطريقة عشوائية بسيطة أحدهما للصورة التقليدية بإجمالي١٦٠ طالب وطالبة من طلاب البكالوريوس والدراسات العليا بكلية التربية بالإسماعيلية بمتوسط عمري ٢٣ عاماً وانحراف معياري ٣,٤٥ عاماً، بينما الصورة الإلكترونية بإجمالي١٩٧ متطوعاً ومتطوعة من مستخدمي شبكات التواصل الاجتماعي (تويتر، فيسبوك، سناب شات) بمتوسط عمري ٢٩,٧ عاماً وانحراف معياري ٢,١٣ عاماً. واستخدم الباحثان المنهج السببي المقارن تصميم بين المجموعات والذي يختبر التكافؤ للصور المتعددة والمقارنة بين أداء الطلاب عليها باستخدام اختبار ت للعينات المستقلة. وأظهرت النتائج تفوق الصورة الإلكترونية عن التقليدية في مؤشرات المطابقة وثبات ألفا لكرونباخ، كما أسفرت النتائج عن عدم تكافؤ الصورتين على مقياس روسنبرج لتقدير الذات. وقد أوصت الدراسة باستخدام الصور الإلكترونية للاختبارات بما تتمتع بهمن حتميه الاستجابة على جميع مفردات الاختبارات وبالتالي الحد من مشكلات التحليل الإحصائي. كما أسفرت نتائج تحليل التغاير عدم تأثير العمر على أداء المجموعتين.

الكلمات الدلالية: تكافؤ القياس، الاختبارات التكيفية، الاختبارات النفسية الإلكترونية.

http://dx.doi.org/10.29009/ijres.2.3.10

Measurement Equivalence of Rosenberg's Self-Esteem: Electronic versus Traditional Scale

Mahmoud Ali Mousa Mehanna, Lecturer Educational psychology, College of education, Suez canal university, Egypt

mahmod567@yahoo.com

Mona Abdelfadel Abdelfatah Elalfy, Lecturer Educational psychology, College of education, Suez Canal university, Egypt

Monaelalfy2013@yahoo.com

Received 18 Dec 2018 **Accepted 17 March 2019**

Abstract: The study aimed as comparing the factor structure of traditional and Electronic Self-esteem scale. Random sample has been selected and consisted of (n=160) for traditional, (n=197) for Electronic scale. Causal Comparative design and between groups method have been used. The findings revealed that the electronic scale has over weighted the traditional one on the fitting indices and on Cronbach's Alpha reliability coefficient. Furthermore, the results have shown that the sample of electronic version overpassed on the performance on Rosenberg's' Self-esteem scale. ANCOVA proved that was no significance of age variable of the samples on performance.

Keywords: traditional and Electronic scale- Rosenberg's' Self-esteem scale

http://dx.doi.org/10.29009/ijres.2.3.10

Summary

Introduction:

Alots of scales may be have an measurement errors, because of social desirability, emotional infidelity, missing data, and outliers. This study should be differentiate between the traditional and electronic scales. The study used self esteem as a personality trait which different from country to others, and from person to another.

The study discuss the statistical models of two forms of electronic and traditional self esteem scales. And test the structural validity of its others.

The study aimed as comparing the factor structure of traditional and Electronic Self-esteem scale with respect to:

- Between subjects designs: it concerns comparing the individuals response mean in the two forums (electronic and traditional) using differences tests such as independent samples t-test, ANOVA, ANCOVA. The non-significance refers to non-equivalence. Those designs depend on classical experimental experiments or meta analysis procedures which concerns with group equivalence, randomization procedure, of interaction of confounding or covariates variables.
- Within subject's designs: participants is the same in each other forums (electronic and traditional scales) there are an assumption assumed paired scores of two scales. We could use correlation or paired samples t-test techniques. We couldn't computes test retest reliability coefficient because of practice effects and person positive beliefs of personality. the alternative procedure here is to use three equivalence tests of electronic form and one traditional test as a control form.

http://dx.doi.org/10.29009/ijres.2.3.10

Problem and research motivation:

Literature reviews in psychology used classical procedures in differentiate between two forums (electronic and traditional) and comparing the reliability coefficients, convergence validity. Then other studies used confirmatory factor analysis, and factorial invariance for recognizing the scales performance for each groups. Finaly some studies used the between subjects or within subjects designs to test the structure fits, or factor stability using differences tests for comparing the two forums.

Hypothesis:

There were non-equivalence between electronic and traditional Rosenberg's self esteem scale using between subjects design.

Objectives:

Test the performance equivalence of electronic and traditional forums of Rosenberg's self-esteem scale using between subjects design.

Method:

Causal Comparative design and between groups method have been used.

Participants:

Random sample has been selected and consisted of (n=160) for traditional, (n=197) for Electronic scale. The participants was students in undergraduates and postgraduates degree in Ismailia college of Education. Traditional forum Participants were 23 years mean, and 3.45 years standard deviation. Then Electronic forum participants were users in social networking (Twitter, Facebook, Snapchat) in which age were 29.7 years mean with 2.13 years standard deviation.

http://dx.doi.org/10.29009/ijres.2.3.10

Results:

The findings revealed that the electronic scale has over weighted the traditional one on the fitting indices and on Cronbach's Alpha reliability coefficient. Furthermore, the results have shown that the sample of electronic version overpassed on the performance on Rosenbergs' Self-esteem scale.

http://dx.doi.org/10.29009/ijres.2.3.10

مقدمة:

تطورت طرق إعداد وإدارة المقاييس النفسية في الآونة الاخيرة وذلك بظهور صور إلكترونية توفر الوقت والجهد، كما كان لتطور عمليات القياس والتقويم المجال الأكبر في السيطرة على الاختبارات والمقاييس النفسية إذ تهتم بتقديم نسخ متوازية تسمح بتخطي المتعلم في مراحل التعلم طبقاً لقدرته التحصيلية. وتعتمد هذه الاختبارات على تقديم محتوى غير متسلسل من الأسئلة أو المفردة حسب الغرض المعدة من أجله. وتعد الاختبارات والمقاييس الإلكترونية طريقة موضوعية في تقويم أداء المتعلم نفسياً ومعرفياً وشخصياً. كما أنها تعطي تغذية راجعة تمكن المتعلم من الالمام بنواحي القصور البادية فيه.

وتتغلب الاختبارات الإلكترونية على مشكلة الزمان والمكان إذ يمكن للمتعلم أن يدخل على محتوى الاختبار ويجيب على مفرداته، بحيث لا يمكنه الغش إذ أن عدد المفردات محددة بوقت. كما أنها تتغلب على الاختبارات التقليدية في إدارتها إذ لا تحتاج إلى تجهيزات مادية من مكان ومراقبين ومتابعة المتعلم (Schmitz, Zimmermann, Gaunt, Stolze & Guttormsen-Schär, 2011). وتتميز الاختبارات الإلكترونية بعدم تخطي المتعلم للاختبار إلا بعدما يكون قد اجتاز كل مفرداته بالاستجابة، وتتصف بخلوها من أخطاء القياس الناتجة عن تفريغ البيانات وتفسيرها وتقلل من تأثير الفاحص وتحيزه عند تفريغ البيانات وتسجيل الدرجات وتفسيرها (سفيان، ٢٠٠٢).

مفهوم الاختبارات الإلكترونية:

عرف الخزي والزكري (٢٠١١) بأنه تلك الاختبارات التي تتم من خلال الحاسب الألي حيث يتم عرض مفردات الاختبار وعلى المتعلم الاستجابة إليها مباشرة. ويرى Ayhan (2018) أنه استخدام التكنولوجيا في شكل من أشكالها في عمليات التقييم النفسي والتعليمي. واتفق (Bird, Callisaya, Cannell, Gibbons, Smith & Ahuja, 2016; Dodou & De-Winter, 2014) أنها اختبارات تقدم تغذية راجعة عن أداء المتعلم وتتصف بعدة معايير منها: تعكس تقدم مسار التعلم باستمرار، وتسهيل عمليات التقييم الذاتي للمتعلم، واكتساب معلومات ذات كفاءة عالية من خلال التطبيق

http://dx.doi.org/10.29009/ijres.2.3.10

المستمر، وتشجع على التواصل المستمر بين المتعلم والأقران والمعلم حول مواد التعلم، وتنمي الدافعية والثقة بالذات، وتسد العجز أو الفجوة الناتجة عن تضخم درجات المتعلم تحصيلياً، وتمد المعلم بالمعلومات اللازمة لتعديل أساليب التعلم المفضلة للمتعلم.

أنماط الاختبارات الإلكترونية:

أشار Noyes & Garland (2008) إلى أن الصورة الإلكترونية للمقاييس والاختبارات تتميز بسرعة التتبع البصري للمثيرات المطروحة بها، وفرق بين نوعين من الصور الإلكترونية للمقاييس منها:

- صياغات المهام Proof-reading tasks: وهي مهام معرفية تقيس أداءات عقلية افتراضية كالتفكير والذاكرة والانتباه والادراك، ومعالجة المعلومات وهي جوانب يمكن الاستدلال عليها في ضوء استجاباته على مجموعة من المهام التي تعكس العملية العقلية.
- صياغات نصية Computer based text: وهي مقاييس تعتمد على التقدير الذاتي Self-report للمفحوص عن طريق الحكم والمراقبة الذاتية لأدائه.

وأكد Mangen, Walgermo & Bronnick (2013) أن الصور الورقية لاختبارات النواحي المعرفية (الفهم القرائي) في المراحل التعليمية الدنيا تؤدي إلى حدوث عبء معرفي. وتنقسم النصوص المقدمة في الاختبارات المعرفية إلى قسمين: أحدهما نصوص روائية سردية Narrative والتي تحوي قصة بمثيرات انفعالية أو درامية تتسلسل فيها الأحداث تتابعياً إذ تعتمد الطريقة السردية على دراسة العلاقة بين الحدث والزمان والمكان والشخصية وتخلق روحاً من الدعابة أثناء القراءة، والشكل الأخر للاختبارات النفسية تقدم في صورة نصوص توضيحية expository والتي تستخدم في بناء اختبارات الذكاء العملي أو الإبداعي والتي تتضمن بعض المثيرات بالقطع المقروءة ومن ثم يتم استنتاج اجابات المفردات استناداً لمهارات ما وراء المعرفة لاستنتاج أحداث وتوقعات وحسابات غير مدرجة بالقطعة

http://dx.doi.org/10.29009/ijres.2.3.10

المقروءة مسبقاً؛ ويفضل تقديم الأولى إلكترونيا، وتقديم الثانية ورقياً (Mangen et al., 2013; Wastlund, Reinikka, Norlander & Archer, 2005).

خصائص الاختبارات الإلكترونية:

أ- **خصائص تقنية**: هي مجموعة من الخصائص التي يوفرها الاختبار الالكتروني للفرد كالنضج الفني الذي يرتبط بالاستقرار والسرعة والأمان والخصوصية والعديد من الخصائص الذاتية والموضوعية في تصحيح عمليات الاختبار ومنها:

- توفير تغذية راجعة فورية بصورة تقلل عمليات العبء المعرفي التي يتعرض لها المتعلم جراء كثرة الأسئلة المقالية ووقت الفحص والاختبار الطويل، وتوفر الدعم لكل مفردة يختبر فيها، علاوة على الرضا النفسي للمتعلم جراء توفير صورة مرضيه عن أدائه النهائي(Schmitz et al., 2011).

- سرعة وقابلية الاستجابة Speed and Responsiveness: وتبدو في إمكانية حذف أو استبدال الإجابة في أي مرحلة من مراحل الاستجابة على مفردات المقياس (Mead & Blitz, 2003; Tyser et al., 2015). كما ينجم عن السرعة في حل المشكلات حصول المتعلم على درجات منخفضة نتيجة مرور الوقت دون انجاز (أبو السعد، ٢٠١١).

- الأمان privacy: ويقصد بها إلى أي درجة يعتقد المفحوص أنه في مأمن من أي تهديدات نفسية حيال استجاباته وبياناته الشخصية التي أدلى بها. ويفضل أن يكون الدخول إلى صفحات هذه المقاييس مستقل عن الحسابات الشخصية للمستخدم حتى لا تكون استجابات المفحوص مضللة (Tyser et al., 2015). وأثبت الصياد (١٩٩٠) أن عدم كتابة الفرد لاسمه بالصورة الورقية ينمي شعور الفرد بخصوصية استجابته وممارسة المفحوص حريته في التعبير

http://dx.doi.org/10.29009/ijres.2.3.10

عـن اسـتجاباته الحقيقيـة ممـا يـنعكس في زيـادة قيمـة الثبـات خصوصـاً في قيـاس المتغيرات اللامعرفية.

ب- **خصـائص إجرائيـة**: هـي مجموعـة مـن التسـهيلات التـي يوفرهـا الاختبـار الالكـتروني للفرد عقب وأثناء عمليات الاستجابة ومنها:

- سـهولة مراجعـة مفـردات الاختبـار: نتيجـة خفـض أخطـاء عمليـات الادخـال ومنع فقد الاستجابات Missing cases (Schmitz et al., 2011).
- تكلفـة الاسـتجابات Price knowledge: تـوفر تكـاليف المقـاييس التقليديـة، وتمكـن مـن إعطـاء قواعـد بيانـات للمفحوصـين، وإعطـاء احصـاءات وصـفية لطبيعـة الاسـتجابات المدخلـة. وحـذر Kim & Huynh (2007) مـن عـدم مناسـبة الصـور الإلكترونيـة لطـلاب المرحلـة الابتدائيـة إذ تعطـي نـواتج متحيـزة ومضـللة. كـما تعتمـد الاختبـارات الإلكترونيـة عـلى تطـوع أفـراد المجتمـع للاسـتجابة عـلى مفردات الاختبار.

ج- **خصـائص نفسـية**: هـي مجموعـة مـن التسـهيلات التـي يوفرهـا الاختبـار الالكـتروني للفرد عقب وأثناء عمليات الاستجابة ومنها:

- خفـض قلـق الاختبـار Test anxiety: أكـد Cassady, Budenz-Anders, Pavlechko & Mock (2001) في دراسـة أجراهـا عـلى طـلاب الجامعـة عـلى دور الاختبـارات الإلكترونيـة في خفـض قلـق الاختبـار. وأثبـت سـمعان (٢٠١٢) دور الـنمط التشـعبي للاختبـارات الإلكترونيـة في خفـض قلـق الرياضـيات لـدى طلاب المرحلة الإعدادية.
- المرغوبيـة الاجتماعيـة Social desirability: وتـرتبط الرغبـة الاجتماعيـة بـإبراز الفرد بمثاليـة واضـحة في اسـتجاباته عـلى مفـردات المقيـاس. وتتناسـب عكسـياً مـع

http://dx.doi.org/10.29009/ijres.2.3.10

خصوصية البيانات وعدم كشف هوية الفرد. ويميل الفرد في الاستجابة على مقاييس السمات الشخصية بالإيجاب أكثر من النفي في المفردات التي تمثل الاستجابة بالنفي تهديد ضمني أو انتقاد شخصي- وهذا ما دعا لوجود اختبارات ذات سياق اجتماعي تستوعب الفرد ذاتياً بعيداً عن هويته الشخصية (أبو السعد، ٢٠١١؛ Dodou & De-Winter, 2014). ويرى Schmitz et al. (2011) أن استجابات الأفراد على الصور الإلكترونية أكثر صدقاً. وتوصل Evan & Miller (1969) إلى أن مقياس الكذب للصور الإلكترونية أقل من نظيره للصور التقليدية.

د- **خصائص سيكومترية**: هي مجموعة من المميزات التي يمكن أن توفرها الصور الإلكترونية ومنها إعداد صور متكافئة من الاختبار الواحد. وتسهل الصور الإلكترونية استخدام البيانات المخزنة عن أداء المتعلم لحساب معايير المفردات ومعاملات للسهولة والصعوبة والتمييز والتخمين والتحيز.

وأجرى Bird, Callisaya, Cannell, Gibbons, Smith & Ahuja (2016) ما وراء التحليل لمعاملات الثبات للاختبارات الخاصة بالصحة النفسية وقد بلغ ثباتها ٠,٩٠. وقارن Gwaltney, Shields & Shiffman (2008) بين تكافؤ الاختبارات العيادية المستخدمة في النواحي الاكلينيكية للصور الإلكترونية والتقليدية لعدد ٦٥ دراسة وبلغ ثبات الصورة الإلكترونية ٠,٩١ في مقابل ٠,٨٨. في حين قارن Duracinsky, Lalanne, Goujard, Herrmann, Cheung-Lung, Brosseau & Chassany (2014) بين ثبات الصورة الإلكترونية والتقليدية لمقياس جودة الحياة لدى المصابين بالأمراض الخطيرة وتوصلت الدراسة إلى عدم وجود فروق بين كلا استجابات المرضى على الصورتين، في حين كانت معاملات الثبات قريبة جداً وتراوحت بين ٠,٧١ و ٠,٩٤. بينما قارن Muehlhausen, Doll, Quadri, Fordham, O'Donohoe, Dogar & Wild (2015) بين

http://dx.doi.org/10.29009/ijres.2.3.10

تكافؤ قياسات نواتج التقرير المرضي للصور المداره إلكترونياً وتقليدياً لعدد ٦٢ دراسة في الفترة من ٢٠٠٧ إلى ٢٠١٣ وقد تفوق ثبات الصور الإلكترونية ٠,٩٩ عن نظيرتها التقليدية إذ بلغ الثبات الكلي ٠,٦٥. واستخدم Joubert & Kriek (2009) عينتين من خريجي الجامعات طبق على أحدها الصورة التقليدية لمقياس الشخصية وطبق على عينة أخرى المقياس إلكترونياً ثم قارن بين ثبات الأبعاد في كل نسخه وتوصلت الدراسة إلى معاملات ثبات تراوحت بين ٠,٦٢ إلى ٠,٨٢ للصورة التقليدية وقيم ثبات ٠,٦١ إلى ٠,٨٩ في الصور الإلكترونية. ومن أمثلة الخصائص السيكومترية للصور الإلكترونية ما يلي:

- التأثر بالبيئة الثقافية للعينة: يقع العديد في براثن قضية مصداقية المقاييس التي تتأثر بثقافة المجتمع، فالمقاييس التي بنيت في ضوء الثقافة السائدة في المجتمع ثبت ارتفاع ثباتها واستقرار أبعادها عن تلك التي بنيت مسبقاً وترجمت وطبقت في ثقافات مختلفة (عامر والألفي، ٢٠١٤).
- البيانات الغائبة Missing data: يعاب على الاختبارات التقليديةوجود البيانات الغائبة الناتجة عن شعور المفحوص بالخجل الاجتماعي أو عدم فهمه لمفردات الاختبار. وأكد Tyser, Beckmann, Weng, O'Farrell & Hung (2015) أن كثرة البيانات الغائبة في الصورة التقليديةراجعاً لعدم انتقال المفحوص من مفردة لآخري دون استكمال التي سبقتها.
- استبعاد المفردات Item Omitted: ويعتبر استبعاد المفردات رد فعل موازي للبيانات الغائبة والتي قد تعطي بيانات مضللة نتيجة إجراء التحليل دون معالجتها بأي من الاستراتيجيات المناسبة، كما أن صياغات بعض مفردات الاختبار تتضمن جوانباً متعددة للقياس مما يسبب نوع من الغموض في فهم المفردة أو راجعة لوجود مفردات تتناقض مع المفاهيم الاجرائية والنفسية للبعد المدرجة به المفردة.

http://dx.doi.org/10.29009/ijres.2.3.10

- المعلومات الشخصية للمفحوص personalization: حدد Tyser et al. (2015) إسهام المعلومات الشخصية المطروحة في التأثير على استجابات الفرد لمفردات في الصور التقليدية للاختبارات النفسية. واختلف معه الصياد (١٩٩٠) إذ يرى أن العبرة ليس بكون المستجيب يكتب بياناته خاصة اسمه أو لا وإنما العبرة بجودة أداة القياس خصوصاً أن معامل ثبات ألفا كان مرتفعاً، وأكد أن أداة القياس الجيدة تجعل المستجيب يستدعي الاستجابات الحقيقية بغض النظر عن استيفاء بياناته أو لا.

ويبدو وجود تعارضاً بين آراء الصياد (١٩٩٠) و Tyser et al. (2015) وهذا مدلوله أن الثبات يعد قدرة المفردات على إبراز التغاير في استجابات المفحوصين على نفس المفردات وأن حدوث بعض التغايرات في الاستجابات بكتابة الاسم من عدمه يزيد من معاملات الثبات لأنه يزيد تغايرات الاستجابة بين المفحوصين، كما يرى الباحثان أن معيار الحكم على جودة الاختبار هو صدقه وأن الثبات هو إجراء مكمل للصدق فكل اختبار صادق ثابت.

مميزات وعيوب الاختبارات الإلكترونية:

تتميز الاختبارات الإلكترونية بنمو الاتجاه الايجابي للمتعلم نحوها إذ أن تجربة الاختبار لديه تصبح أسهل عندما يعرف الطالب كيف يتصرف جيداً مما يزيد من ثقته بنفسه. يمكن استخدامها كصورة للتقويم المرحلي إذ تساعد المتعلم على التقييم لمستواه المعرفي كل فترة في ضوء فهمه (دوديين، ٢٠٠٥). وتتميز بخلوها من أخطاء القياس والتحيز (Ayhan, 2015). كما تتقارب معاملات الثبات للاختبارات الشخصية للصور الإلكترونية والتقليدية وهذا ما أكدت عليه الدراسات التجريبية ودراسات ما وراء التحليل (Duracinsky et al., 2014; Joubert & Kriek, 2009). خلو عمليات التقويم من الأخطاء العشوائية المرتبطة بالصور التقليدية. وتتميز هذه الاختبارات بالكفاءة Efficiancy إذ لا تتطلب جهداً تفسيرياً أو احصائياً (أبو السعد، ٢٠١١).

http://dx.doi.org/10.29009/ijres.2.3.10

ومن عيوبها اكتساب المتعلم للحكمة الاختبارية مما يسبب تضخم درجات الطالب Grade inflation في الاختبار مقارنة بزملائه ممن لديه نفس المستوى المعرفي (Tyser et al., 2015). كما يعاب عليها تخمين Guessing الاستجابات للاختبارات المعرفية عند يكون المتعلم في حالة من عدم التأكد أو عدم إلمامه بالاستجابة الصحيحة حيال موقف معين (Bird et al., 2016). ويعاب عليها عشوائية عرض مفرداتها بصورة تجعلها أكثر عرضة للتأويل (أبو السعد، ٢٠١١).

الإجراءات الفنية لتحسين مصداقية الاختبارات الإلكترونية:

اتفق (Bird et al., 2016; Hosseini, Abidin & Baghdarnia, 2014; Taylor, Kirsch, Jamieson & Eignor, 1999) على بعض الإجراءات الفنية لتحسين مصداقية الاختبارات الإلكترونية ومنها:

١. تعديلات طفيفة Minor modifications: وهي التي تجرى في ضوء المنطق على جميع نواحي المقياس فيما عدا التعليمات؛ كطريقة الاستجابة على المقياس وإجراءات عرض المفردات والتي تهتم بطرح مفردة واحدة أثناء الاستجابة في الصورة الإلكترونية عن مجموعة مفردات في الصورة التقليدية على ورقة واحدة.

٢. تعديلات متوسطة Moderate modifications: وهي تعديلات في محتوى ومعنى المفردات كطريقة عرض مفردات متعددة في إطار شاشة واحدة خصوصاً في مفردات الاختيار من متعدد، والتحكم في شكل ظهور وحجم الخط وتغيير ترتيب عرض المفردات، واستخدام شريط التمرير لعرض مفردات بعد أو مقياس واحد.

٣. تعديلات جوهريةSubstantial modifications: كاستبعاد مفردات تتطلب وقتاً أكثر للاستجابة واستحضار المعلومات للاستجابة من ذاكرة المتعلم طويلة الأمد.

http://dx.doi.org/10.29009/ijres.2.3.10

التصميمات الاحصائية لتكافؤ المقاييس:

- **تصميمات بين الموضوعات Between-Subjects Designs**: وهي تلك التي تهتم بالمقارنة بين متوسطات استجابات أفراد المجموعات داخل صورتي المقارنة الإلكترونية أو الورقية باستخدام اختبارات فارقة مثل: ت المستقلة أو تحليل التباين الأحادي أو تحليل التغاير البسيط وتشير الدلالة الإحصائية لعدم التكافؤ non-equivalence. وتعتمد هذه التصميمات على التجريبية التقليدية أو ما وراء التحليل، والتي تهتم في جل محتواها بتكافؤ المجموعات أو إجراء العشوائية وإلا فالتأثير يعتبر ناتجاً عن تفاعل متغيرات دخيلة. ويفضل استخدام النمذجة البنائية Structural equation modeling (SEM) والذي يفترض اختبار التكافؤ ومعاملات الارتباط الداخلية للمقاييس ويتصف حساسيته لتأثير المتغيرات الدخيلة. ويتميز هذا التصميم باختبار الثبات إما بطريقة الاختبار وإعادته أو باستخدام طريقة ألفا كرونباخ. ويعد مدخل المعادلة البنائية من الاساليب التي تهتم بتحديد النماذج تحت التحليل، وعدد القياسات المستخدمة، وقيم معاملات الارتباط الداخلية، وحجم العينة، وثبات المقاييس إذ تحتاج النمذجة البنائية لمقاييس ذات معاملات ثبات عالية إلى أحجام عينة كبيرة. وأثبتت Hosseini et al. (2014) أن الصور الإلكترونية للاختبار الفهم القرائي التي تتطلب الاستجابة على اختيار من متعدد تتفوق في خصائص الأداء عن نظيرتها الورقية وقد اعتمدت الدراسة على تصميم داخل المجموعات.

- **تصميمات داخل الموضوعات Within-Subjects Designs**: وفيها يستجيب أفراد نفس المجموعة على الاختبار الورقي والإلكتروني، ويفترض سيكومترياً أن تكون درجتي الاختبارين مرتبطة، وعادة ما تستخدم أسلوب الارتباط أو اختبار ت المرتبطة، ويتعذر حساب الثبات باستخدام الاختبار وإعادته نظراً للتلويث

http://dx.doi.org/10.29009/ijres.2.3.10

الإحصائي للاستجابات بسبب تأثير الممارسة أو بسبب رغبة المفحوص بتأمل الاتجاه الشخصي ـ للمفردات، ويفضل استخدام الاختبارات المتوازية إلكترونياً أو التجزئة النصفية لحساب ثبات الاختبارات. وقد استخدم Boo (1997) تصميم داخل المجموعات لنفس الاختبار وتوصلت دراسته إلى وجود فروق بين صورتي الاختبار لنفس المجموعة على ثلاثة صور متوازية إلكترونية واختبار تقليدي ورقي. في حين توصلت Taylor et al. (1999) إلى عدم وجود علاقة بين الصورة الإلكترونية والتقليدية لاختبار TOEFL.

خلفية ومشكلة الدراسة:

اعتمدت بعض الأدبيات النفسية على أساليب كلاسيكية في المفاضلة بين الصورتين الإلكترونية والورقية من خلال المقارنة بين معاملات الثبات أو الاعتماد على الصدق التقاربي لكل صورة على حدة باستخدام اختبار محك يقيس نفس السمة. واستخدمت بعض الدراسات التحليل العاملي التوكيدي للمقارنة بين مؤشرات المطابقة للصورتين، وثمة دراسات استخدمت التمايز العاملي Factorial Invariant للتعرف على وجود فروق في الأداء على الصورتين أم لا في أداء طلاب العينتين. وثمة دراسات اعتمدت على تصميم بين المجموعات وتصميم داخل المجموعات وذلك للتأكد من مطابقة بنية الصورتين، أو وجود استقرار للبنية العاملية للصورتين، واستخدام اختبارات فارقة أو فروض علاقة للتحقق من مدى تفوق أحد الصورتين على الاخرى، وفيما يلي بعض الدراسات السابقة المفسرة:

وميز Rice (1994) بين اختبارين للفهم القرائي أحدهما ورقية، والأخر صيغ في صورة مهام إلكترونية لعينة بلغت ١٢٠ متعلم بالمرحلة المتوسطة وقد أسفرت النتائج عن وجود فروق دالة إحصائياً في درجات اختبار الفهم القرائي لصالح الصورة الورقية. وفي دراسة Wise, Plake, Johnson & Roos (1992) أجريت على عينة مكونة من طلاب ٢٠٤ كلية فوضلت بين صورتين أحدهما إلكترونية والأخرى تقليدية في مقرر الجبر. وقد

http://dx.doi.org/10.29009/ijres.2.3.10

توصلت النتائج إلى تفوق أداء الاختبارات الإلكترونية، بالرغم من ارتفاع مستويات قلق الاختبار لدي المستجيبين. بينما قارن Parasuraman, Zeithaml& Malhotra (2005) بين الصدق البنائي لمقياس ESQUAL للصورة الورقية والصورة الإلكترونية وقد أسفرت النتائج عن مطابقة حسنة للصورة الإلكترونية للمقياس وارتفاع معاملات الثبات له والتي تراوحت بين ٠,٨٣٠ إلى ٠,٩٤.

وقدم Wastlund, Reinikka, Norlander & Archer (2005) اختباري للفهم القرائي أحدهما إلكترونياً والاخر ورقياً وصيغت مفردات الاختبار في صورة اختيار من متعدد لدى عينة من طلاب الثانوية العليا بالسويد وأسفرت النتائج عن وجود فروق ذات دلالة احصائية في أداء الفهم القرائي والعبء المعرفي بين الصورة الإلكترونية والورقية لصالح الصورة الإلكترونية. بينما أعد Green, Rafaeli, Bolger, Shrout & Reis (2006) مقياس نفسي متعدد الأبعاد وطبق في صورته الورقية والإلكترونية على عينة من الراشدين في البيئة السويسرية واختبر النموذج العاملي التوكيدي للنموذجين وقد اثبتت الدراسة مطابقة حسنة للنموذج التوكيدي للصورتين، كما أثبتت الدراسة عدم وجود فروق في الأداء على أبعاد الصورتين والدرجة الكلية لهما في ضوء اختبار ت المستقلة.

وقارن Kim & Huynh (2007) بين الصورة الإلكترونية والورقية لاختبار الجبر لعينة من طلاب الصف الخامس عشر بولاية ثاوزيرن الامريكية وقد توصلت الدراسة إلى مطابقة حسنة للنماذج العاملية التوكيدية للصورتين إلا أن الصورة الورقية أثبتت تفوق طفيف في مؤشري RMSEA و NNFI. وأجري Gwaltney, Shields & Shiffman (2008) ما وراء التحليل لعدد ٤٢ دراسة للصورة الإلكترونية والورقية لمقياس المثابرة والتي استخدمت تصميم داخل المجموعات وحسبت معامل الارتباط الكلي بين النسختين ٠,٩. واستخدم Joubert & Kriek (2009) النمذجة البنائية للتحقق من الصدق البنائي للصورة الإلكترونية والصورة الورقية لاختبار تقييم الشخصية لعمليات الانتقاء والتوجيه

http://dx.doi.org/10.29009/ijres.2.3.10

بالمؤسسات. وقد اسفرت النتائج عن تشابه كبير في الخصائص السيكومترية للمقياسين للنسختين.

وأعد Karkee, Kim & Fatica (2010) اختباراً للرضا عن عملية تدريس مقرر المواد الاجتماعية واستخدم التصميم داخل المجموعات بانتقاء عينة عشوائية من طلاب المدارس العليا بلغت ٢٥٠٠٠ طالب وتوصلت نتائج الدراسة إلى وجود تشابه عاملي بين الصورتين، بالإضافة إلى عدم وجود فروق دالة احصائياً بين تطبيق القياسين الالكتروني والورقي للمقياس. واستخدم Schmitz et al. (2011) تصميم داخل المجموعات للتحقق من وجود تكافؤ بين الصورة الإلكترونية والورقية لاختبار نفسي- بخمسة أبعاد مكون من ٣٦ مفردة معد في ضوء تدريج ليكرت الخماسي للاستجابات وقد عالج البيانات الغائبة بطرق الاستكمال الرياضي باستخدام برنامج spss وقد استخدم اختبار قابلية الاستخدام للتحقق من مدى مناسبة الصورتين للاستخدام وقد بلغ ثباته ٠,٩٢ وتوصلت النتائج على ١٠ طلاب إلى وجود فروق بين صورتي التطبيق مما يعني عدم تكافؤ الأداء على صورتي المقياس. وقد أبدى ثمانية من المستجيبين تفضيلهم للاستجابة عليالصور التقليدية.

واستخدم Mojarrad, Hemmati, Gohar & Sadeghi (2013) التصميم داخل المجموعات للمقارنة بين استخدام مجموعة من طلاب الجامعة الايرانية للصورة الورقية والإلكترونية لاختبار اتجاه المتعلم نحو الاستيعاب القرائي، وتوصلت النتائج إلى عدم وجود فروق بين متوسط درجات أداء الطلاب على صورتي الاختبار. وتحقق Mangen et al. (2013) من أثر النصوص الخطية المقروءة المقدمة تقليدياً وإلكترونياً على الفهم القرائي لدى عينتي من المدارس الابتدائية بالنرويج وتوصلت الدراسة إلى وجود فروق دالة إحصائياً في الفهم القرائي لدى الطلاب المقدم إليهم صور تقليدية ورقية مطبوعة عن ذويهم من المقدم لهم الصور الإلكترونية. وتوصلت دراسة Tyser et al. (2015) إلى وجود فروق دالة احصائياً بين الصورة التقليدية والصورة الإلكترونية في عدد المفردات المستبعدة لصالح

http://dx.doi.org/10.29009/ijres.2.3.10

الورقة والقلم، كما وجدت فروق دالة احصائياً بين الصورة التقليدية والصورة الإلكترونية في عدد البيانات الغائبة لصالح الورقة والقلم. وقارن Ayhan (2015) بين مجموعتين أحدهما اجتازت اختبار القبول الجامعية بالبيئة التركية لجامعة أنقرة ورقياً والأخرى اجتازت الاختبار إلكترونياً وقد اعتمدت الدراسة على تصميم بين المجموعات. وقد أسفرت النتائج عن ارتفاع معاملات الثبات للاختبار الالكتروني بالإضافة إلى ارتفاع معامل التخمين للصورة التقليدية للاختبار.

وبالتأمل في الدراسات السابقة لوحظ اختلاف بين النتائج التي تناولت المفاضلة بين النسخة الورقية التقليدية والنسخة الإلكترونية حيث هم متشابهين من حيث الثبات وكذا ارتفاع ثبات الصورة الإلكترونية عن الصورة الورقية ولذا تحاول هذه الدراسة محاولة الحسم فيما الاختلاف بين كل منهما وذلك من خلال استخدام بيانات مقياس تقدير الذات عبر الفيسبوك والذى طبقت صورتيه الورقية والإلكترونية. وفي ضوء الطرح السابق للأدبيات النفسية يمكن صياغة مشكلة الدراسة في السؤال الآتي: هل توجد تكافؤ أداء الطلاب على الصورة الإلكترونية والصورة الورقية لمقياس روسنبرج تقدير الذات؟

فرض الدراسة:

لا يوجد تكافؤ بين الصورة الإلكترونية والتقليدية لمقياس روسنبرج لتقدير الذات باستخدام تصميم بين المجموعات.

هدف الدراسة:

التحقق من تكافؤ في الأداء على الصورة الإلكترونية والتقليدية لمقياس روسنبرج لتقدير الذات باستخدام تصميم بين المجموعات.

http://dx.doi.org/10.29009/ijres.2.3.10

أهمية الدراسة:

تكمن أهمية الدراسة في عرض الإجراءات التقنية والسيكومترية والإحصائية والفنية اللازمة لإعداد الصور الإلكترونية للمقاييس، كما طرحت الدراسة إجراءات تحسين مصداقية المقاييس التقليدية وإدخال اجراءات بديلة تحسن من صدق وثبات المقاييس. وسد الفجوة الناتجة عن إدمان الانترنت والحاسب الألي والعزوف والاتجاهات السلبية نحو استخدام الاختبارات الإلكترونية في عمليات القياس والتقويم داخل الجامعات المصرية.

حدود الدراسة:

تقع حدود هذه الدراسة زمنيا بالنسبة للعينه الورقية والإلكترونية في الفصل الدراسي الثاني ٢٠١٦/ ٢٠١٧. ويجب الإشارة إلى أن العينة الورقية تم اختيارها عشوائيا بينما العينة الإلكترونية فقد كانت من متطوعين من مستخدمي شبكات التواصل الاجتماعي. ويمكن تعميم نتائج الدراسة على عينات مختلفة من طلاب الجامعات المصرية بكليات التربية المصرية بغض النظر عن التخصصات التربوية المختلفة التي يؤدي فيها المتعلم الاختبار.

الطريقة والاجراءات

أولاً: منهجية الدراسة: اعتمدت الدراسة على استخدام المنهج السببي المقارن، تصميم بين المجموعات والذي يهتم باختبار مجموعتين مختلفتين واحدة لكل نوع من الاختبارات.

ثانياً: عينة الدراسة: اختيرت عينه الصورة الورقية للمقياس بصورة عشوائية بسيطة؛ بلغت ١٦٠ طالب وطالبة من طلاب البكالوريوس والدراسات العليا بكلية التربية بالإسماعيلية بمتوسط عمري ٢٣ عاماً وانحراف معياري ٣,٤٥ عاماً. بينما الصورة الإلكترونية للمقياس فقد تكونت العينة من ١٩٧ متطوع من مستخدمي شبكات التواصل الاجتماعي

http://dx.doi.org/10.29009/ijres.2.3.10

(تـويتر، فيسـبوك، سـناب شـات) بمتوسـط عمـري ٢٩,٧ عامـاً وانحـراف معيـاري ٢,١٣ عاماً.

ثالثاً: مقياس روسنبرج لتقدير الذات:

طبـق مقيـاس تقـدير الـذات لــ (1965) Rosenberg وهـو مكـون مـن عشر مفـردات، صـيغت سـتة مفـردات صـياغة موجبـة (١ و ٣ و ٤ و ٧ و ٨ و ١٠) مثـل "أنـا راضي عـن نفسي بصفة عامة"، "وأنـا لـدي اتجـاه ايجـابي نحـو نفسي" وصـيغت أربعـة مفـردات صـياغة سـلبية (٢ و ٥ و ٦ و ٩) مثـل "اعتقـد بـأنني لـيس لـدي شيء افتخـر بـه". وصُـحح المقيـاس في ضـوء مقياس ليكـرت الخـماسي ولـيس الربـاعي كـما وضـعه مصـمم المقيـاس وعليـه فالدرجـة القصـوىٰ هـي ٥٠ والدرجـة الـدنيا ١٠ ، وقـدر ثبـات المقيـاس باسـتخدام الفـاكر وبنـاخ وبلـغ ٠,٧٠ تقريباً وتراوح معامل الارتباط المصحح من ٠,١٧٧ إلى ٠,٥٤ .

الصدق البنـائي: قـارن عـامر (٢٠١٦) بـين العديـد مـن الابنيـة العامليـة لمقيـاس روسـنبرجلتقدير الذات لطـلاب المرحلـة الثانويـة وتفـوق نمـوذج العامـل العـام في مطابقتـه. وتحقـق الباحثـان مـن البنيـة العامليـة التوكيديـة لمقيـاس تقـدير الـذات للصـورة التقليديـة والصـورة الإلكترونيـة باسـتخدام التحليـل العـاملي التوكيـدي باسـتخدام طريقـة أقصىٰ احـتمال Maximum likelihood (ML)، وفيما يلي قيم مؤشرات المطابقة للنماذج المقاسة:

جدول (١): مطابقة النماذج العاملية لمقياس تقدير الذات.

	X^2	X^2/df	RMSEA	NNFI	GFI	AGFI
الصورة التقليدية	٤٠٤٣,٤ P= 0.000	١١٥,٥	٠,٤٠٥	-٠,٠٤٨	٠,٤٥٦	٠,١٤٦
الصورة الإلكترونية	٦٩,٣١ P= 0.001	٢	٠,٠٧٤	٠,٩٢٦	٠,٩٢١	٠,٨٨١

http://dx.doi.org/10.29009/ijres.2.3.10

وأسفرت النتائج عن مؤشرات مطابقة حسنة للصورة الإلكترونية لمقياس روسنبرج لتقدير الذات بينما أسفرت النتائج عن مطابقة غير مقبولة لنموذج الصورة الورقية كما اتفقت نتائج التحليل التوكيدي مع نتائج عامر (٢٠١٥) في تدعيم نموذج العامل العام مع وجود ارتباطات بين بواقي المفردات الموجبة وبين بواقي المفردات السالبة (المفردة ٦ والمفردة ٨) وهذا تدعيم الطبيعة أحادية البعد مع وجود تأثيرات الطريقة. وفيما يلي تشبعات المفردات:

جدول (٢): تشبعات مفردات مقياس روسنبرج لتقدير الذات.

م	المفردة	التشبع	الخطأ المعياري	قيمة ت	معامل R^2
١	عموما أنا راضي عن نفسي	٠,٧٤٣	٠,٠٧١	١٠,٥٠	٠,٥٥
٢	دائما أنا غير مبسوط	٠,٦٤٨	٠,٠٧٤	٨,٧٤	٠,٤٢
٣	اعتقد أني امتلك صفات حميدة	٠,٥٩٦	٠,٠٧٦	٧,٨٨	٠,٣٦
٤	أنا قادر علي انجاز المطلوب مني مثل الآخرين	٠,٥٨٠	٠,٠٧٦	٧,٦٢	٠,٣٤
٥	اعتقد أني لا امتلك شيء افتخر به	٠,٧٥٠	٠,٠٧١	١٠,٦٣	٠,٥٦
٦	غالبا اعتقد بأنني غير فعال	٠,٥٨٨	٠,٠٧٦	٧,٧٤	٠,٣٥
٧	اعتقد بأنني شخص ذو قيمة في المجتمع	٠,٣٦٩	٠,٠٨١	٤,٥٥	٠,١٤
٨	أتمنى الحصول على احترام أكثر من الآخرين	٠,٤٨١	٠,٠٧٨	٦,١١	٠,٢٣
٩	أنا غير مبالي بالشعور بالفشل	٠,٧٥٣	٠,٠٧٠	١٠,٧٠	٠,٥٧
١٠	أنا لدي اتجاه ايجابي نحو نفسي	٠,٧٩١	٠,٠٦٩	١١,٤٩	٠,٦٣

أسفرت النتائج عن تشبعات دالة احصائياً عند مستوى دلالة ٠,٠٥ لتجاوز قيمة اختبار ت لها عن ١,٩٦ وتراوحت قيم التشبعات بين ٠,٣٦٩ إلى ٠,٧٩١.

http://dx.doi.org/10.29009/ijres.2.3.10

ثبات مقياس روسنبرج لتقدير الذات:

تراوحت قيم معاملات الثبات باستخدام طريقة ألفا بين ٠,٧٠ إلى ٠,٩٢ في الدراسات النفسية المختلفة والدراسات عبر الثقافية. وحسب الباحثان ثبات الصورة الإلكترونية باستخدام معامل ألفا وبلغت ٠,٨٧٠ بينما بلغ معامل ألفا للصورة الورقية ٠,٧٧٩.

رابعاً: إجراءات الدراسة:

١. برمجة المقياس في صورته الإلكترونية باستخدام برنامج Google forums بحيث يتم عرض جميع مفردات المقياس جميعها في منصة العرض الكترونياً في صفحة والبيانات الديمغرافية.

٢. اختيرت عينتان بصورة عشوائية من طلاب المرحلة الجامعية بمستوييها (البكالوريوس والدراسات العليا) وزعت عليها الصورة الورقية للمقياس وبلغت إجمالي العينة الأولى ١٦٠ طالب وطالبة، بينما بلغت العينة الإلكترونية ١٩٧ متطوع.

٣. تم التحقق من الصدق البنائي والثبات لمقياس روسنبرج لتقدير الذات لكلاً من النسخة الإلكترونية والنسخة الورقية. واستخدم تصميم بين المجموعات وذلك باختيار كل صورة على عينة

٤. حساب الفروق بين مجموعتين التصميم بين الموضوعات باستخدام اختبار ت للعينات المستقلة لأداء المجموعتين على أداة الدراسة كلاً على نسخته.

٥. إعادة التحقق من النتائج لتصميم بين المجموعات باستخدام اختبار تحليل التغاير للتحقق من انعدام تأثير متغير العمر على نتائج الأداء على مقياس روسنبرج لتقدير الذات.

http://dx.doi.org/10.29009/ijres.2.3.10

نتائج الدراسة

نتائج الفرض ومناقشته:

ينص فرض الدراسة على لا يوجد تكافؤ بين الصورة الإلكترونية والتقليدية لمقياس روسنبرج لتقدير الذات باستخدام تصميم بين المجموعات. ويمكن ترجمة الفرض احصائياً إلى الفرض الاحصائي التالي: توجد فروق ذات دلالة احصائية بين الصورة الإلكترونية والصورة الورقية لمقياس روسنبرج لتقدير الذات لصالح الصورة الإلكترونية. وباستخدام تصميم بين المجموعات وذلك بإجراء اختبار ت للعينات المستقلة لدراسة تكافؤ الأداء الطلاب الصورة الإلكترونية والورقية لمقياس روسنبرج لتقدير الذات، وفيما يلي نتائج اختبار الفرض احصائياً:

جدول (٥): تكافؤ الصورة الإلكترونية والورقية لمقياس تقدير الذات.

الصورة	ن	المتوسط	الانحراف المعياري	د.ح	قيمة ت	الدلالة	حجم تأثير مربع ايتا
الإلكترونية	١٩٧	٣٣,٦٦	٧,٥٧	٣٥٥	٤,١٣	٠,٠٠٠ دال	٠,٠٤٦ صغير
الورقية	١٦٠	٣٠,٣٨	٧,٣٥				

وأسفرت النتائج عن وجود فروق ذات دلالة احصائية بين الصورة الإلكترونية والورقية في متوسط درجات مقياس روسنبرجتقدير الذات لصالح النسخة الإلكترونية. ويلاحظ من نتائج اختبار ت للعينات المستقلة دلالة الفروق بين الصورتين مما يعني عدم تكافؤ الأداء على كلا النسختين وهذا قد يفسر بعدم تشوه الاستجابة اجتماعياً لربما أن الصور أصبحت سمة ممثلة للعصر والحديث عنها يعتبر مجدياً ونظراً لحداثة متغيرات علم النفس الالكتروني وما يجد فيه الفرد من تقدير الاخرين لذاته وبالتالي تتضخم فيه الذات المثالية لديه؛ ومن ثم فالمفحوص شغوفاً بالتعرف صورة أدائه عند الاخرين فتتولد لديه قدر من الاعتمادية في استخدام التقنية.

http://dx.doi.org/10.29009/ijres.2.3.10

واتفقت الدراسة الحالية مع نتائج التحليل التوكيدي الذي أثبت تفوق التحليل العاملي التوكيدي للصورة الإلكترونية عن الصورة الورقية، وهذا قد يرجع إلى تفوق الصورة الإلكترونية في بعض المعايير التقنية التي عدم قبول النسخ التي تحوي بيانات غائبة من ناحية، بالإضافة إلى عدم سماح النسخة الإلكترونية لكل حاسب شخصي- من الاستجابة على المقياس الا مرة واحدة فقط.

وبالتأمل في قيمة حجم التأثير يتضح تدني قيمتها وهذا قد يرجع إلى أن تقدير الذات هي أحد السمات الشخصية المميزة للفرد وبالتالي فيميل الفرد إلى اختيار الاستجابة المرغوبة اجتماعياً بدلاً من اختيار تلك التي تعبر عن شخصيته وهذا يتفق مع (أبو السعد، ٢٠١١؛ Dodou & de Winter, 2014). كما أن هذا يؤكده معاملات الثبات المتقاربة للدراسة الحالية بين كلا الصورتين وفي دراسة Joubert & Kriek (2009) والتي استخدمت مقياس العوامل الخمسة الكبرى للشخصية. ويوصي الباحثان بضرورة صياغة مؤشراً للكذب داخل مقاييس الدراسة لاستبعاد حالات اختيار الاستجابات المرغوبة وذلك انطلاقاً من نتائج Evan & Miller (1969).

اختلف النتائج للدراسة الحالية مع دراسة Rice (1994) في توصلت لتفوق النسخة الورقية وهذا يرجع إلى اعتماده على مجموعة من المهام المعرفية الإلكترونية بينما اعتمدت الدراسة الحالية على التقدير الذاتي للمفحوص لمفردات مقياس تقدير الذات، بالإضافة إلى اعتماد Rice (1994) على طلاب المرحلة المتوسطة والتي تقابل المرحلة الاعدادية بالبيئة المصرية.

واثبتت الدراسة تفوق الصورة الإلكترونية على الورقية باستخدام اختبارات فارقة وهذا اتفق مع (Wastlund et al., 2005; Wise et al., 1992) وتعارضت النتائج مع Rice (1994) بالرغم من اتفاق الدراستين في قياس سمة معرفية إلا أن التفوق الحادث راجع

http://dx.doi.org/10.29009/ijres.2.3.10

لاختلاف بيئة الدراسة وارتفاع مستوى نضج الطلاب حيث اتفقت الدراسة مع نتائج الدراسات التي اعتمدت على عينة من طلاب المرحلة الثانوية أو الجامعة.

كما أن تفضيل الصور الورقية لطلاب المراحل الأولى والوسطي من التعليم وذلك لعدم نضج الطلاب وتأهيلهم للعمل على الحواسب الإلكترونية. وتفوقت النسخ الإلكترونية عن الورقية في تقليل التشوه الاجتماعي للاستجابات وذلك بمراعاة اخفاء هوية المتعلم، وعدم استخدامه لحسابه الشخصي لمواقع التواصل أو بريده الالكتروني.

المناقشة والتعليق

معظم الدراسات النفسية اعتمدت على تصميم بين المجموعات باستخدام مجموعتين تستجيب أحداها على الصورة الورقية والإلكترونية، واهتمت بعض الدراسات بالمفاضلة بين الاداء، أو النواحي السيكومترية كالصدق أو الثبات، أو تفضيل بعض الصور؛ بينما اختلفت الدراسة الحالية في أنها درست تكافؤ الاداء والخصائص السيكومترية للنسختين.

اتفقت الدراسة الحالية مع دراسة Bennett (2003) في استخدام تصميم بين المجموعات والمقارنة بين البنية العاملية للصورتين الإلكترونية والورقية وتفوق الصورة الإلكترونية بالرغم من اختلاف الثقافة وأهداف الدراسة والمرحلة الدراسية المطبق عليها (ابتدائي وثانوي). واتفقت الدراسة مع Schmitz et al. (2011) في تفوق الصورة الإلكترونية في الأداء على طلاب الجامعة بالرغم من اختف خلفية الطلاب الثقافية والجامعية فركزت الدراسة الحالية على التربويين وكان حجم العينة أكبر واختلاف الدراستين في تصميم التكافؤ فاعتمدت الدراسة الحالية على تصميم بين المجموعات.

اعتمدت الدراسات السابقة (Flowers & Kim, 2011; Karkee et al., 2010; Kim & Huynh, 2007) في البنية العاملية على مقاييس ذات العامل العام، واعتمد بعضها

http://dx.doi.org/10.29009/ijres.2.3.10

(Green et al., 2006; Joubert & Kriek, 2009; Parasuraman et al., 2005) على مقاييس متعددة الابعاد وقد بلغ متوسط الثبات للدراسات السابقة التي اعتمدت على العامل العام القيمة ٠,٨٣٠ في حين بلغ متوسط الثبات للدراسات التي اعتمدت على تعدد الأبعاد الابعاد القيمة ٠,٩٠٠، بينما بلغ ثبات العامل العام للصورة الإلكترونية ٠,٨٧٠ وهذه القيمة تجاوزت متوسط ثبات دراسات العامل العام وهذا قد يرجع إلى تمتع مقياس روسنبرج لتقدير الذات بثبات مرتفع عبر ثقافات مختلفة.

اعتمدت الدراسات السابقة على الاداء النهائي على المقاييس باستخدام اختبارات فارقة واهملت معاملات التحيز والسمات النفسية التي تؤثر على الاستجابة على مفردات المقياس كالحالة الانفعالية للمفحوص وعامل التاريخ وتكرار التطبيق في تصميم داخل المجموعات وتشوه الاستجابة والقابلية للاستخدام وتفضيل الصور الإلكترونية إلا دراسة (Schmitz et al., 2011) فقد استخدمت صور متوازية الكترونياً ومقارنتها بصورة ورقية داخل المجموعات وذلك للتخلص من أثر عامل التكرار.

اختلفت النتائج مع دراسات (Mangen et al., 2013; Taylor et al., 1999) عدم وجود فروق بين النسختين وهذا يرجع إلى اعتماد هذه الدراسات على اختبارات الأداء المعرفي (الفهم القرائي، اختبارات تعلم اللغة الانجليزية كلغة اجنبية TOEFL)، بينما اتفقت الدراسة مع (Schmitz et al., 2011; Wastlund et al., 2005; Wise et al., 1992)في فروق في الاداء وهذا قد يرجع إلى قلة عدد المفردات المطروحة بداخل المقياس، ويرى Noyes & Garland (2003) أن التعارض الحادث في النتائج يرجع إلى اختلاف المنهجية والتصميمات التجريبية المقارنة وبخاصة استخدام تصميم داخل المجموعات.

اعتمدت الدراسة على مقياس روسنبرج لتقدير الذات إذ يتمتع بدرجة عالية من الاتساق الداخلي بين مقاييس تقدير الذات الأخرى، علاوة على أنه أثبت ثباته في العديد من

http://dx.doi.org/10.29009/ijres.2.3.10

الدراسات عبر الثقافية ودراسات ما وراء التحليل والذي تراوح ثبات مفرداته بين ٠,٧ إلى ٠,٩ كما أنه يتمتع بأحادية البعد.

واعتمدت بعض الدراسات (Mojarrad et al., 2013; Schmitz et al., 2011) إجراءات متعددة لمعالجة البيانات الغائبة مثل استبدالها بأحد طرق الاستكمال الرياضي ببرنامج Spss 23. إلا أن الدراسة الحالية جعلت جميع مفردات الصورة الإلكترونية حتمية بحيث لا تقبل استجابات المفحوص إلا بعض استجابته على جميع مفرداته أما الصورة التقليدية فقد استبعد الباحثان المقاييس التي وجدت بها بيانات غائبة تماما عن التحليل.

وبالتأمل في قيم المتوسط العمري لكل من العينتين يتضح ان أحدهما من مرحلة المراهقة المتأخرة والأخرى من مرحلة الشباب وهنا درس الباحثان تأثير متغير العمر الزمني على نتائج التصميم باستخدام اختبار تحليل التغاير وقد كان متغير العمر غير دال احصائياً (F=0.30, P= 0.59)، في حين أن كلا المجموعتين تأكد وجود فروق دالة لهما في الأداء على نسخة الاختبار التي طبقت على كل مجموعة فيهما (F=10.51, P= 0.001). ومن ثم فإنه يمكن القول بان التغير الحادث في أداء كلا المجموعتين كان مغايراً وهذا يرجع لفروق التصميم وليس للمرحلة العمرية أي أن النتائج تتمتع بالصدق الداخلي.

وتعاني الدراسة الحالية من بعض المحددات منها أنه لا يمكن تعميم النتائج على عينات تختلف في الخصائص العمرية أو الثقافية لعدم اختبار الصدق العبوري لنتائج الدراسة الحالية. ويوصي الباحثان بإعادة إجراء الدراسة مرة أخرى للمقارنة بين تصميم بين المجموعات وتصميم داخل المجموعات للوقوف على أفضل هذه التصميمات ومدى قابليته للاستخدام ومدى تفضيل المتعلم لأحد هذه الصور. كما يوصي الباحثان استخدام الصور الإلكترونية لما تتمتع به من حتمية الاستجابة على جميع مفردات الاختبارات والمقاييس ومن ثم التخلص من مشكلة البيانات الخطية والبيانات المتطرفة التي ينجم عنها

http://dx.doi.org/10.29009/ijres.2.3.10

مشكلات التحليل الاحصائي مثل الاعتمادية الخطية ومحدد المصفوفة السالب خصوصاً في التحليل العاملي للمقاييس.

ويمكن استخلاص أن النتائج الإلكترونية للمقاييس النفسية هي أدق في رصد العمليات النفسية للتقرير الذاتي للمتعلم أو المفحوص. كما يمكن الاستفادة من هذا في الجوانب التي يتسبب قياسها في بعض الحرج للمفحوص في التعبير صراحة في استجاباته خوفاً من تعليق الفاحص على استجاباته أو لتكوين صورة سلبية عنه. ومن ثم فإنه قد يزيف استجاباته في الصور التقليدية، وهذا ما بدئ في نتائج الدراسة حيث أن تقدير الذات هي سمة شخصية بدئ قياسها الكترونياً أفضل من نظيره التقليدي.

ويوصي الباحثان ضرورة استخدام الصور الإلكترونية في النواحي المعرفية التي قد تعكس أداءً موضوعياً بعيداً عن تحيز المقدّر، وتبعد المتعلم عن العوامل البيئية التي تسبب شتات الانتباه أثناء استجابته على الاختبارات المعرفية، كما أنها تمنع المتعلم من الغش الكائن في اللجان الاختبارية إذ أنها محددة بوقت، وتجعل المتعلم أكثر رضى في تقبل نتاج تعلمه المعرفي. كما أنها سوف توفر على المعلم عمليات الرصد وأخطاء القياس البادية في معالجة البيانات ورصد وغيرها. كما يوصي الباحثان أيضاً في حالة توفر وقت لدئ الإدارة المدرسية بضرورة عمل صورتين أحدهما تقليدية كما هو مألوف بمدارسنا العربية وأخرئ إلكترونية للتعرف على مدئ تساهل أو تشدد المعلم في تقييم أداء متعلمه، الأمر الذي يعطي التعلم جودة ومعنئ.

http://dx.doi.org/10.29009/ijres.2.3.10

المراجع

١. أبو السعد، أحمد (٢٠١١). دليل المقاييس والاختبارات النفسية والتربوية، الطبعة الثانية. عمان: مركز ديبونو لتعليم التفكير.

٢. الخزي، فهد والزكري، محمد. (٢٠١١). تكافؤ الاختبارات الإلكترونية مع الاختبارات الورقية في قياس التحصيل الدراسي: دراسة تجريبية على طلبة كلية التربية بجامعة الكويت. مجلة دراسات الخليج والجزيرة العربية، ٣٧، ١٤٣، ١٦٧-١٩٨.

٣. دودين، حمزة. (٢٠٠٥). تدريس استراتيجيات تقديم الاختبارات. مجلة التربية، قطر، ٣٤، ١٥٢، ١٠٢-١١٧.

٤. سفيان، نبيل صالح (٢٠٠٢): تصميم برنامج على الحاسوب لمقياس التوافق النفسي الاجتماعي وأثره على بعض المتغيرات. منشور في مجلة دراسات عربية في علم النفس، العدد الثاني، ابريل ٢٠٠٢، مصر، كلية الآداب، قسم علم النفس، عرض في المؤتمر التربوي الأول – التعليم وتحديات القرن الواحد والعشرين – جامعة مؤتة = الأردن من ٢٧-٣١ أكتوبر http://arabpsynet.com/archives/op/OP.Sofiane.SoftPersScale.htm.

٥. الصياد، عبد العاطي أحمد. (١٩٩٠). المعامل ألفا كرونباخ ونوعية استجابة المستجيب لعبارة "اكتب اسمك إذا رغبت". المؤتمر السنوي السادس لعلم النفس في مصر المنعقد بالجمعية المصرية للدراسات النفسية في الفترة من ٢٢- ٢٤ يناير.

٦. عامر، عبد الناصر السيد. (٢٠١٤أ). تقييم استخدام تطبيقات نمذجة المعادلة البنائية في البحث النفسي. مجلة دراسات عربية في علم النفس (رانم)، المجلد ١٣، العدد ٢.

٧. عامر، عبد الناصر السيد. (٢٠١٤ب). نمذجة المعادلة البنائية الاستكشافية في مقابل التحليل العاملي التوكيدي للبنية الداخلية لأهداف الانجاز. المجلة المصرية للدراسات النفسية، ٢٤, ٤٠٤-٤٣٠.

http://dx.doi.org/10.29009/ijres.2.3.10

٨. عامر، عبد الناصر السيد. (٢٠١٥). فحص تأثيرات الطريقة في مقياس تقدير الذات لروسنبرج: نماذج عالمية متنافسة. المجلة المصرية للدراسات النفسية، المجلد ٢٥، العدد ٨٨، ١ – ٣١.

٩. عامر، عبدالناصر السيد والألفي، منى عبد الفضيل. (٢٠١٤). الصدق البنائي لقائمة العوامل الخمسة الكبرى للشخصية لكولدبيرج واختبار عوامل عالية الرتبة. مجلة كلية التربية بالمنصورة.

http://dx.doi.org/10.29009/ijres.2.3.10

References

- Ayhan, A. S. (2015). Comparability of scores from cat and paper and pencil implementations of student selection examination to higher education. PhD dissertation, Bilkent University.
- Bennett, R. E. (2003, October). Online assessment and the comparability of score meaning. In International Association for Educational Assessment Annual conference, Manchester, October 2003.
- Bird, M. L., Callisaya, M. L., Cannell, J., Gibbons, T., Smith, S. T., & Ahuja, K. D. (2016). Accuracy, validity, and reliability of an electronic visual analog scale for pain on a touch screen tablet in healthy older adults: A clinical trial. Interactive journal of medical research, 5(1).
- Boo, J. (1997). Computerized versus paper-and-pencil assessment of educational development: Score comparability and examinee preferences (Doctoral dissertation, University of Iowa).
- Cassady, J. C., Budenz-Anders, J., Pavlechko, G., & Mock, W. (2001). The Effects of Internet-Based Formative and Summative Assessment on Test Anxiety, Perceptions of Threat, and Achievement.
- Dodou, D., &De-Winter, J. C. (2014). Social desirability is the same in offline, online, and paper surveys: A meta-analysis. Computers in Human Behavior, 36, 487-495.
- Duracinsky, M., Lalanne, C., Goujard, C., Herrmann, S., Cheung-Lung, C., Brosseau, J. P., &Chassany, O. (2014). Electronic versus paper-based assessment of health-related quality of life specific to HIV disease: reliability study of the PROQOL-HIV questionnaire. Journal of medical Internet research, 16(4).
- Evan, W. M., & Miller, J. R. (1969). Differential effects on response bias of computer vs. conventional administration of a social science questionnaire: an exploratory methodological experiment. Behavioral Science,14, 216–227.

http://dx.doi.org/10.29009/ijres.2.3.10

- Flowers, C., Kim, D. H., Lewis, P., & Davis, V. C. (2011). A comparison of computer-based testing and pencil-and-paper testing for students with a read-aloud accommodation. Journal of Special Education Technology, 26, 1, 1-12.
- Green, A. S., Rafaeli, E., Bolger, N., Shrout, P. E., & Reis, H. T. (2006). Paper or plastic? Data equivalence in paper and electronic diaries. Psychological methods, 11, 1, 87.
- Gwaltney, C. J., Shields, A. L., &Shiffman, S. (2008). Equivalence of electronic and paper-and-pencil administration of patient-reported outcome measures: a meta-analytic review. Value in Health, 11(2), 322-333.
- Gwaltney, C. J., Shields, A. L., Shiffman, S. (2008). Equivalenceof electronic and paper-and-pencil administration of patientreportedoutcome measures: a meta-analytic review. Value Health, 11, 322–333.
- Hosseini, M., Abidin, M. J. Z., &Baghdarnia, M. (2014). Comparability of test results of Computer based Tests (CBT) and Paper and Pencil Tests (PPT) among English language learners in Iran. Procedia-Social and Behavioral Sciences, 98, 659-667.
- Joubert, T.,&Kriek, H. J. (2009). Psychometric comparison of paper-and-pencil and online personality assessments in a selection setting. SA Journal of Industrial Psychology, 35, 1.
- Joubert, T., &Kriek, H. J. (2009). Psychometric comparison of paper-and-pencil and online personality assessments in a selection setting. SA Journal of Industrial Psychology, 35(1), 78-88.
- Joubert, T., &Kriek, H. J. (2009). Psychometric comparison of paper-and-pencil and online personality assessments in a selection setting. SA Journal of Industrial Psychology, 35(1), 78-88.
- Karkee, T., Kim, D. I., &Fatica, K. (2010, April). Comparability study of online and paper and pencil tests using modified internally and externally

http://dx.doi.org/10.29009/ijres.2.3.10

matched criteria. In annual meeting of the American Educational Research Association, Denver, CO.

- Kim, D. H., & Huynh, H. (2007). Comparability of computer and paper-and-pencil versions of algebra and biology assessments. The Journal of Technology, Learning and Assessment, 6, 4.
- Liu, C., Ang, R.,& Lwin, P. (2013). Cognitive, personality, and social factors associated with adolescents' online personal information disclosure. Journal of Adolescence, 36, 629- 638.
- Mangen, A., Walgermo, B. R., &Brønnick, K. (2013). Reading linear texts on paper versus computer screen: Effects on reading comprehension. International Journal of Educational Research, 58, 61-68.
- Mead, A. D., & Blitz, D. L. (2003). Comparability of paper and computerized non-cognitive measures: A review and integration. Paper presented at the eighteenth annual conference of the Society for Industrial andOrganizational Psychology, April 2003, Orlando, Florida.
- Millsap, C. M. (2000). Comparison of computer testing versus traditional paper and pencil testing.
- Mojarrad, H., Hemmati, F., Gohar, M. J., & Sadeghi, A. (2013). Computer-basedassessment (CBA) VS. Paper/pencil-basedassessment (PPBA): aninvestigationin totheperformanceandattitudesofiranian EFL learners' readingcomprehension. International Journal of Language Learning and Applied Linguistics World, 4, 4, 418-428.
- Moore, T. M., Reise, S. P., Gur, R. E., Hakonarson, H., & Gur, R. C. (2015). Psychometric properties of the Penn Computerized Neurocognitive Battery. Neuropsychology, 29, 2, 235.
- Muehlhausen, W., Doll, H., Quadri, N., Fordham, B., O'Donohoe, P., Dogar, N., & Wild, D. J. (2015). Equivalence of electronic and paper administration of patient-reported outcome measures: a systematic review

http://dx.doi.org/10.29009/ijres.2.3.10

and meta-analysis of studies conducted between 2007 and 2013. Health and quality of life outcomes, 13(1), 167.

- Naglieri, J. A., Drasgow, F., Schmit, M., Handler, L., Prifitera, A., Margolis, A., & Velasquez, R. (2004). Psychological testing on the Internet: new problems, old issues. American Psychologist, 59, 3, 150.
- Noyes, J. M., & Garland, K. J. (2008). Computer-vs. paper-based tasks: Are they equivalent?.Ergonomics, 51, 9, 1352-1375.
- Parasuraman, A., Zeithaml, V. A., & Malhotra, A. (2005). ESQUAL: A Multiple-Item Scale for AssessingElectronic Service Quality. Journal of Service Research, 7, 3, 213-233.
- Rice, G. E. (1994). Examining constructs in reading comprehension using two presentation modes:Paper vs. computer. Journal of Educational Computing Research, 11, 153–178.
- Richman, W. L., Kiesler, S., Weisband, S., &Drasgow, F. (1999). A meta-analytic study of social desirability distortion in computer-administered questionnaires, traditional questionnaires, and interviews. Journal of Applied Psychology, 84, 5, 754.
- Rosenberg, M. (1965). Society and adolescents self image. Princeton, NJ: Prinction University.
- Schmitz, F., Zimmermann, P., Gaunt, K., Stolze, M., &GuttormsenSchär, S. (2011). Electronic rating of objective structured clinical examinations: mobile digital forms beat paper and pencil checklists in a comparative study. Information Quality in e-Health, 501-512.
- Seo, D. G., & De Jong, G. (2015). Comparability of Online-and Paper-Based Tests in a Statewide Assessment Program Using Propensity Score Matching. Journal of Educational Computing Research, 52, 1, 88-113.

http://dx.doi.org/10.29009/ijres.2.3.10

- Sun, S. Y., Shieh, C. J., & Huang, K. P. (2013). A research on comprehension differences between print and screen reading. South African Journal of Economic and Management Sciences, 16, 5, 87.
- Taylor, C., Kirsch, I., Jamieson, J., &Eignor, D. (1999). Examining the relationship between computer familiarity and performance on computer-based language tasks. Language Learning, 49(2), 219-274.
- Tyser, A. R., Beckmann, J., Weng, C., O'Farrell, A.,& Hung, M. (2015, March). A Randomized Trial of the Disabilities of the Arm, Shoulder, and Hand Administration: Tablet Computer Versus Paper and Pencil. The Journal of Hand Surgery, 40, 3, 554-559.
- Wastlund, E., Reinikka, H., Norlander, T., & Archer, T. (2005). Effects of VDT and paper presentation on consumption and production of information: Psychologicaland physiological factors. Computers in Human Behavior, 21, 377–394.
- Wise, S. L., Plake, B. S., Johnson, P. L., & Roos, L. L. (1992). A comparison of self-adaptedand computerized adaptive tests. Journal of Educational Measurement, 29, 4, 329-339.

http://dx.doi.org/10.29009/ijres.2.3.10