المجلة العربية للقياس والتقويم

تصميم أداة لقياس بعض المتغيرات السيكومترية المرتبطة بالاختبارات التكيفية المحوسبة من وجهة نظر المهتمين

إعـــــــــــداد

محمد "عبد الرازق الشهير صغير" مصطفى عبد الكريم

باحث ماجستير ــ قسم علم النفس التربوي

كلية التربية بقنا ــ جامعة جنوب الوادي

تصميم أداة لقياس بعض المتغيرات السيكومترية المرتبطة بالاختبارات التكيفية المحوسبة من وجهة نظر المهتمين

إعـــــــــداد

محمد "عبد الرازق الشهير صغير" مصطفى عبد الكريم

الملخـــــص:

هدف البحث إلى تصميم أداة لقياس بعض المتغيرات السيكومترية المرتبطة بالاختبارات التكيفية المحوسبة من وجهة نظر المهتمين، ولقد تم الكشف عن الخصائص السيكومترية للأداة، ولتحقيق هذا الهدف تم استخدام المنهج الوصفي، وتكونت عينة التحقق من الخصائص السيكومترية من (٢٠٠) معلم ومعلمة من معلمي ومعلمات المرحلتين الإعدادية والثانوية بمحافظة قنا، فلقد تم التحقق من الصدق باستخدام صدق المحكمين كخطوة من خطوات إعداد المقياس، ثم حساب معامل اتفاق كيندال للمحكمين الذي بلغت قيمته (٠.٨٦) عند مستوي دلالة (٠.٠٠٢)، وتم التحقق من الصدق التقاربي وذلك باستخدام اختبار مان ويتني، وتم التوصل إلى عدم وجود فروق بين متوسطي رتب الدرجات في كل بعد من الأبعاد الستة بين مجموعة معلمي ومعلمات التعليم الإعدادي ومجموعة معلمي ومعلمات التعليم الثانوي، وتم التحقق من الثبات باستخدام معامل ألفا كرونباخ حيث بلغت قيمته (٠.٧٣٠)، وبناءً عليه فأن الأداة تتسم بمؤشرات صدق وثبات مقبولة.

الكلمات المفتاحية: المتغيرات السيكومترية، الاختبار التكيفي المحوسب، المهتمين بالاختبارات.

Designing a tool to measure some psychometric variables associated with computerized adaptive tests from the perspective of those interested

Prepared by

Mohamed "Abdel Razek El Shaher Sagheer" Mustafa Abdel Karim

Abstract:

This paper aimed to design a scale to measure some of the psychometric variables associated with Computerized Adaptive Tests from the point of view of interested people by revealing the psychometric properties of this scale. To achieve this goal, Descriptive Approach was used on a sample consisted of (200) teachers from males and females who work in preparatory and secondary stages in Qena Governorate Its validity was verified by using the validity of the arbitrators as a step in preparing it, then calculating the Kendall agreement coefficient for the arbitrators, which had a value of (0.86) at the level of significance (0.002). Its convergent validity was verified by using Mann-Whitney Test., The overall results indicated that it the results showed that there were no differences between the average ranks of scores in each of the six dimensions between the group of preparatory school teachers and the group of secondary school ones. Its reliability was verified by using Cronbach's Alpha Coefficient, whose value was (0.730).and This scale was characterized by acceptable validity and reliability coefficients.

Keywords: psychometric variables, computerized adaptive testing, those interested in tests.

مقدمـــــــــــة:

تعد الاختبارات وسيلة لقياس الأداء والمهارات ومخرجات العملية التعليمية، وفي ظل التحول الرقمي الذي نعاصره كان من الضروري دراسة كيفية الانتقال من الاختبارات الورقية إلى الاختبارات المحوسبة؛ لأن دراسة هذه الأنواع من الاختبارات (التكيفية المحوسبة– الخطية المحوسبة– الورقية) تعد من الأمور المهمة التي تقوم عليها العملية التربوية.

وفي ظل توفر الحاسبات وشبكات الاتصال والإنترنت على المستوى الشخصي وداخل المؤسسات التعليمية، توجهت المؤسسات التعليمية نحو حوسبة التعلم والتقويم في التعليم الجامعي وقبل الجامعي (دعنا، ٢٠١٢)[١] ، ونتيجة لذلك فقد ظهر عدد من الاختبارات التي تعتمد على الحاسب الآلي، التي تعرف بالاختبارات المحوسبة، أو الاختبارات الإلكترونية.

فالاختبار التكيفي المحوسب (CAT)، هو حالة خاصة للاختبار القائم على الكمبيوتر، حيث يقوم كل ممتحن بإجراء اختبار مختلف مصمم وفقًا لمستوى قدرته كبديل لإعطاء كل ممتحن نفس الاختبار الثابت، وفي بداية الاختبار يتم تقديم أسئلة متوسطة الصعوبة من أجل التقييم المبدئي لمستوى كل فرد، وأثناء الاختبار يتم تسجيل كل إجابة على الفور وإذا أجاب الممتحن بشكل صحيح فإن الاختبار يقدر إحصائيًا قدرته على أنها أعلى ثم يقدم عنصراً يطابق هذه القدرة (Triantafillou et al.,2008).

وأشار Parshall et al. (2002) إلى أن الاختبار التكيفي المحوسب ينقسم إلى نوعين أساسيين، يتم تعريف هذين النوعين من خلال قواعد التوقف الخاصة بهم؛ إنها اختبارات ذات طول ثابت ومتغير، فالاختبارات التكيفية المحوسبة ذات الطول الثابت، لها نفس عدد العناصر لكل ممتحَن، لذلك يمكن اختبار الممتحنين المختلفين على مستويات مختلفة من الدقة، تماماً كما هو الحال مع الاختبار التقليدي غير المتكيف.

[١] يتم التوثيق وفق الجمعية الأمريكية لعلم النفس الإصدار السابع (APA7)

ويتطلب إنشاء الاختبار التكيفي المحوسب اختيار نموذج استجابة للمفردة مناسب (أحادي/ ثنائي/ ثلاثي المعلم)، وبناء مستودع المفردات (بنك الأسئلة)، وكذلك تحديد أسلوب انتقاء الفقرة المناسب، واختيار أسلوب تقدير القدرة المناسب وهناك أسلوبان رئيسان هما: (الأرجحية القصوى– البيزي) ، واختيار أسلوب قاعدة إنهاء الاختبار المناسبة، وهناك محكات مختلفة لإنهاء الاختبار، فقد ينتهي الاختبار بعد تطبيق عدد محدد من المفردات، أو استناداً إلى أدنى قيمة للمعلومات حيث يتوقف الاختبار عند الوصول إلى أدنى مستوى من معلومات المفردة المصاحبة لتقدير القدرة (نور الدين، ٢٠١٩).

ومن أوجه القصور التي تمثل مشكلة في الاختبارات التكيفية المحوسبة ما أشار إليه قاسم (٢٠٢١) أنها تحتاج إلى عينات كبيرة من الممتحَنين، كما أن بنك الأسئلة يحتاج إلى مئات المفردات، وتتطلب وجود خبراء ذوي كفاءة عالية في القياس التربوي، ويحتاج بناؤها إلى تكلفة مادية عالية، بالإضافة إلى عدم وجود اتفاق حول أفضل إستراتيجيات انتقاء الأسئلة، كما يصعب تفسير معالم بعض النماذج خاصة تلك النماذج الخاصة بالأسئلة متعددة الاستجابات.

وتؤثر تلك العيوب علي جودة الاختبار التكيفي المحوسب، ويؤيد هذا ما توصلت اليه دراسة (Kimura,2017)أن العديد من الممتحَنين الذين قدم لهم الاختبار التكيفي المحوسب شعروا بالإحباط، وهذا الشعور يمكن أن يقلل من التعلم والكفاءة الذاتية والتحفيز لديهم، ووفقًا لمسح أجرته تلك الدراسة، وجدت أن ما يقرب من ٩٠٪ من المتقدمين للاختبار زعموا أن الاختبار صعب، وشعر ٦٠٪ منهم أيضًا بالإحباط أو عدم الرضا عن التجربة، وزعموا أنهم عادة ما يشعرون بالرضا عندما تجاوزت نتائجهم ١٠٠/٧٦، وخيبة الأمل عندما كانوا أقل من ١٠٠/٤٥.

ولا يزال يواجه العديد من المشكلات العملية في عملية تنفيذ الاختبارات التكيفية المحوسبة، مثل الحفاظ على صلاحية الاختبار، فإنه يجعل اختيار الاسئلة مشكلة معقدة مع الكثير من القيود. كما أنها تحتاج إلى تكلفة عالية، وأيضاً مشكلة كيفية إنشاء عناصر عالية

الجودة تلقائيًا على نطاق واسع، ويعتمد حل هذه المشكلات على التطبيق العميق لتكنولوجيا الحاسب الآلي في تنفيذ عملية الاختبار التكيفي(Lu and Cong, 2016) .

وفي هذا الصدد توصلت دراسة (Lu and Cong, 2016) إلى أنه على الرغم من أن الاختبارات التكيفية قد حققت تقدمًا كبيرًا، إلا أنها كشفت تدريجياً عن بعض المشكلات وأوجه القصور في سنوات الممارسة وتحتاج إلى مزيد من الدراسة والتحسين

ويميز الاختبارات التكيفية أنها تختزل العدد الكلي لأسئلة الاختبار، حيث تحتاج إلى أقل من نصف عدد الأسئلة التي تطبق في الاختبارات الخطية المحوسبة أو الورقية، وهذا بدوره يعمل على تقليل وقت الاختبار؛ الذي يعمل علي تحفيز الممتحَن، ويميزها أيضاً المرونة في التطبيق، حيث يمكن تطبيقها عن طريق الإنترنت، أو عن طريق شبكات محلية، أو علي الجهاز الخاص بالممتحَن، وأيضا تتمتع بالمرونة في إدارة الاختبار، كما أنها تمنع الغش؛ لأن لكل ممتحن أسئلة خاصة تختلف عن أقرانه، و تتيح الاختبارات التكيفية عدداً من أشكال الأسئلة، مثل أن تتضمن الأسئلة صوتًا أو صورة أو حركة(قاسم،2021).

ويؤكد هذا دراسة (Yasuda et al.,2021) التي هدفت إلى قياس أثر الاختبار التكيفي المحوسب في اختصار وقت اختبار مفهوم القوة، ونتيجة لهذا فقد توصلت الدراسة إلى أنه يمكن تقليص عدد الأسئلة من ٤٠ سؤالًا إلى (١٥– ١٩) سؤالًا، وبالتالي تقليص وقت الاختبار إلى (٥٠٪ – ٦٣٪).

ولذلك أكد Veldkamp and Matteucci (2013) أن الممتحنين في هذا العصر الرقمي يفضلون أن يخضعوا إلى اختبارات إلكترونية وبشكل خاص للاختبار التكيفي المحوسب بدلًا من الخضوع إلى الاختبارات الورقية التقليدية؛ فالاختبار التكيفي المحوسب يقلل من طول الاختبار بنسبة تصل إلى ٤٠٪ دون فقد في دقة القياس، كما أن الاختبار التكيفي المحوسب لا يشعر فيه الممتحنون بالملل أثناء تأديته، وقد أوصي نور الدين (٢٠١٩) بالتوسع في تحويل الاختبارات التحصيلية الورقية والخطية المحوسبة إلى اختبارات تكيفية محوسبة.

كما أوضح(2015) Rezaie and Golshan أن الاختبارات التقليدية قد تستغرق وقتاً طويلاً لتقييم عدد كبير من الطلاب، وقد يكون من الصعب تقديم نتائج دقيقة وسريعة للممتحنين؛ فيمكن أن تساعد الاختبارات التكيفية المحوسبة المؤسسات الأكاديمية في هذا الصدد، حيث من خلال استخدام الاختبارات التكيفية المحوسبة يمكن أن تتوفر نتائج التقييم بسرعة كبيرة، ويمكن تقليل الحاجة إلى الموارد البشرية، ويمكن أن يكون التقييم أكثر كفاءة.

وللاختبارات التكيفية المحوسبة متغيرات سيكومترية، ومن بين هذه المتغيرات الغش، وسهولة التصحيح، والدقة في التصحيح، وتغطية كافة الأهداف، وكذلك الإمكانيات المادية، وتفضيلات المهتمين، وتعد ظاهرة الغش في الامتحانات شكلًا من أشكال الخيانة؛ فهو مخالفة للقوانين والأخلاقيات المتعارف عليها، حيث يرى عطية (٢٠١٥) أنها تتمثل في حصول الممتحَن على الإجابة عن طريق نقل الجواب من وثيقة جاهزة أو زميل له، دون اعتبار لتعلم المادة أو شعور شخصي بأهميتها لحياته ومستقبله (ص٦٩).

وأما سهولة ودقة التصحيح يتأثران باختلاف نوع الاختبار، وذلك الاختلاف يسهم في توفير الراحة للممتحِن أثناء التصحيح وذلك فيما يتعلق بالسهولة، وفيما يتعلق بالدقة يسهم الاختلاف في تأثر دقة التصحيح بأداء الممتحِن، أو بالطريقة المحوسبة للتصحيح إذا كان نوع الاختبار محوسب، وتختلف النتائج حول تحديد النوع الأنسب من الاختبارات، والذي يسهم في جعل عملية التصحيح سهلة ودقيقة.

بينما فيما يخص التكاليف المادية المتعلقة بالاختبار التكيفي المحوسب، فإن الموقع المسؤول عن توفير البرامج المستخدمة، يقدم تلك البرامج بتكلفة مادية عالية جداً، مما يجعل الأمر معقدً عند الحصول على تلك البرامج للاستخدام مرات قليلة، بينما يعد الأمر غير مكلف مادياً إذا تم الحصول عليها من قبل المؤسسات التعليمية، التي تستخدمها على مدار سنوات عديدة، وتغطي عدد كبير من الممتحَنين، مما يجعل الاختبار التكيفي المحوسب ذا تكلفة مادية قليلة على المدي البعيد.

وأما فيما يتعلق بتغطية الأهداف فإن نوع الاختبار يوثر في تغطية تلك الأهداف السلوكية، حيث أن بعض الأنواع يسمح بعرض بعض أنواع الأسئلة دون غيره، مما يسهم في عدم تغطية كافة الاهداف المرجوة.

وعليه فإن هذا البحث يسعى لتصميم أداة لقياس مدى تأثر تلك المتغيرات السيكومترية المتعلقة بالاختبارات التكيفية المحوسبة من وجهة نظر المهتمين بتلك الاختبارات (معلمي ومعلمات المرحلة الابتدائية- معلمي ومعلمات المرحلة الإعدادية)، عن طريق فحص خصائصها السيكومترية.

مشكلة البحث:

مع بداية ظهور الاختبارات التكيفية المحوسبة، لم تكن الخيار الأفضل عند الممتحَنين؛ نظرا لأن أداءهم عليها لم يكن مقبولا مقارنة بأدائهم على الاختبارات الورقية، وذلك لأن الممتحَنين كانوا غير ملمين بمهارات الحاسب الآلي، وأيضا لعدم توفر التقنيات التكنولوجية التي ظهرت حديثاً.

ولكن توفر التقنيات التكنولوجية أسهم في قبول الاختبارات التكيفية المحوسبة لدى بعض الممتحَنين، حيث أوضح كل من Veldkamp and Matteucci (2013) أن الممتحَنين في هذا العصر الرقمي يفضلون أن يخضعوا إلى اختبارات إلكترونية، وبشكل خاص للاختبار التكيفي المحوسب بدلًا من الخضوع إلى الاختبارات الورقية التقليدية؛ فالاختبار التكيفي المحوسب يقلل من طول الاختبار بنسبة تصل إلى ٤٠٪ دون فقد في دقة القياس.

وأسهم التوجه نحو قبول الاختبارات التكيفية المحوسبة في وجود العديد من الدراسات التي تقارن بينها وبين وغيرها من صور الاختبارات الاخرى، عن طريق قياس مدى كفاءة

تلك الصور من خلال تطبيق تلك الصور على الممتحَنين، وبناءً على استجابتهم يتم تحديد أي صورة أكثر كفاءة.

واستناداً على بعض قواعد البيانات مثل دار المنظومة وGoogle Scholar وResearchGate، تم التوصل إلى وجود ندرة في الدراسات التي تطرقت إلى فحص المتغيرات السيكومترية المتعلقة بالاختبارات التكيفية المحوسبة، فبعض النتائج توصلت إلى أن الاختبار التكيفي المحوسب أكثر كفاءة من غيره بناءً على أداء الممتحَنين عليه، حيث توصلت دراسة (Thomas et al., 2021)إلى أن الاختبار التكيفي المحوسب جعل المهمة أكثر كفاءة بنسبة ٣٦٪ وجعل تقدير الدرجات أكثر دقة بنسبة ٢٣٪ مقارنة بالاختبار غير التكيفي.

ودراسة (قاسم، ٢٠٢٠) التي توصلت إلى فاعلية الاختبار التكيفي المحوسب متعدد المراحل من خلال دقة تقديره لقياس الاستدلال البصري، وأظهرت دراسة (الكيال وآخرون،٢٠١٧) إلى أن الاختبار التكيفي المحوسب أكثر فاعلية من الورقي، وذهبت دراسة (دعنا، ٢٠١٢) إلى أن الاختبار التكيفي المحوسب يعمل بكفاءة.

بينما بعض النتائج توصلت إلي أن الاختبار التكيفي المحوسب أقل كفاءة من الصور الأخرى للاختبار، حيث توصلت دراسة (Kimura,2017)أن العديد من الممتحَنين الذين قدم لهم الاختبار التكيفي المحوسب شعروا بالإحباط، وهذا الشعور يمكن أن يقلل من التعلم والكفاءة الذاتية والتحفيز لديهم، ووفقًا لمسح أجرته تلك الدراسة، وجدت أن ما يقرب من ٩٠٪ من المتقدمين للاختبار زعموا أن الاختبار صعب، وشعر ٦٠٪ منهم أيضًا بالإحباط أو عدم الرضا عن التجربة، وزعموا أنهم عادة ما يشعرون بالرضا عندما تجاوزت نتائجهم ١٠٠/٧٦ ، و خيبة الأمل عندما كانوا أقل من ١٠٠/٤٥.

وأظهرت دراسة (AŞKAR et al., 2012) أن الدرجات أعلى في الاختبارات الورقية مقارنة بالاختبارات المحوسبة. وأوصى Rezaie and Golshan (2015) بإجراء المزيد من البحوث التي تتناول المقارنة بين الاختبار الإلكتروني بنوعيه (تكيفي – خطي) والاختبار

التقليدي (الورقة والقلم). كما توصلت دراسة (Lu and Cong, 2016) إلى أنه على الرغم من أن الاختبارات التكيفية قد حققت تقدماً كبيراً، إلا أنها كشفت تدريجياً عن بعض المشكلات وأوجه القصور في سنوات الممارسة وتحتاج إلى مزيد من الدراسة والتحسين.

ومن خلال عرض بعض الدراسات السابقة يتضح أن: جميع الدراسات قارنت بين أنواع الاختبارات عن طريق التطبيق الفعلي للاختبار على الممتحَنين، وندرة الدراسات التي تطرقت الي فحص المتغيرات السيكومترية المتعقلة بالاختبار التكيفي المحوسب، مما يجعل الميدان بحاجة إلى أدوات تقيس هذه المتغيرات، وعليه يمكن المقارنة بين أنواع الاختبارات من خلال التعرف على وجهة نظر المهتمين بها من معلمي ومعلمات المرحلتين الإعدادية والثانوية، وذلك بتصميم أداة لقياس بعض المتغيرات السيكومترية المرتبطة بالاختبارات التكيفية المحوسبة، وعليه يمكن صياغة المشكلة في:

١. هل يتمتع مقياس المتغيرات السيكومترية للاختبار التكيفي المحوسب بدرجة مقبولة من الصدق؟
٢. هل يتمتع مقياس المتغيرات السيكومترية للاختبار التكيفي المحوسب بدرجة مقبولة من الثبات

أهمية البحث:

يمكن تلخيص أهمية البحث في النقاط التالية:

١. تقديم البحث لخلفية نظرية عن الاختبارات التكيفية المحوسبة من حيث التعريف، والنشأة، وطريقة الإنشاء، وكذلك طريقة التطبيق.
٢. قد تفيد نتائج البحث القائمين على إعداد الاختبارات في تحديد مدى كفاءة الاختبار التكيفي المحوسب مقارنة بالأنواع الأخرى للاختبارات، كما تفيد أيضاً في التعرف على مدى إمكانية تطبيق مثل هذا النوع من الاختبارات في جمهورية مصر العربية.

٣. فتح المجال للمزيد من الباحثين لعمل بحوث حول الاختبارات التكيفية المحوسبة، كما يمكن للباحثين الاستفادة من أداة البحث في إجراء بحوث مماثلة.

أهداف البحث:

يهدف البحث الحالي إلى التحقق من الكفاءة السيكومترية لمقياس المتغيرات السيكومترية للاختبار التكيفي المحوسب من خلال:

١. التحقق من صدق المقياس.
٢. التحقق من ثبات المقياس.

مصطلحات البحث:

١. **المتغيرات السيكومترية المرتبطة بالاختبار التكيفي المحوسب:** وهي (الغش في الامتحان، سهولة التصحيح، دقة التصحيح، التكاليف المادية المطلوبة، تغطية كافة الأهداف، تفضيلات المهتمين)، فهي الأمور التي يرتبط بإعدادها الاختبار التكيفي المحوسب، والاختبار التكيفي المحوسب هو حالة خاصة للاختبارات الإلكترونية، يتم فيه عرض أسئلة مختلفة على كل طالب، ويتم اختيار الأسئلة لكل طالب ومدى صعوبتها بناءً على إجابته على الاسئلة السابقة، ومن المحتمل أن يختلف عدد الاسئلة المحددة وترتيب عرض كل سؤال من طالب إلى آخر.
٢. **المهتمين:** هم معلمو ومعلمات المرحلتين الإعدادية والثانوية المهتمون بالاختبارات من حيث إعدادها وتطبيقها.

حدود البحث:

الحد البشري: يتمثل في عينة من معلمي المرحلتين الإعدادية والثانوية بمحافظة قنا.

الحد المكاني: ثلاث مدارس بكل مركز من مراكز محافظة قنا بإجمالي ٢٧ مدرسة للمرحلتين(الإعدادية–الثانوية).

الحد الزمني: الفصل الدراسي الثاني للعام الدراسي ٢٠٢٢/٢٠٢٣.

منهج البحث وإجراءاته:

أولاً: منهج البحث: يعتمد البحث الحالي على المنهج الوصفي.

ثانياً: عينة البحث: بلغ عدد عينة البحث (٢٠٠) معلم ومعلمة من معلمي المرحلتين الإعدادية والثانوية بمدارس محافظة قنا.

ثالثاً: أداة البحث:

وصف المقياس: تكون المقياس في صورته الأولية من (٣٦) مفردة، تم تقسيمها على ستة أبعاد، في كل بعد ست مفردات، وتكون الاستجابة على مفردات المقياس وفق تدريج ليكرت الثلاثي على أن تكون البدائل (يتوافق بشدة)، (يتوافق)، (لا يتوافق)، وتقابل تلك الاستجابات الدرجات الآتية على الترتيب (٣)، (٢)، (١)، ويسبق المقياس شرح مبسط لبعض أنواع الاختبارات وعلى رأسها الاختبار التكيفي المحوسب، وأيضاً تم إضافة صفحة للتعليمات توضح للمستجيب طريقة الإجابة على مفردات المقياس.

رابعاً: المعالجات الإحصائية:

للتحقق من أهداف البحث، تم استخدم عدة أساليب إحصائية بواسطة برنامج (SPSS v25)، وبرنامج (JASP V0.18.1)، وتلك الأساليب هي:

- معامل اتفاق كيندال.
- اختبار ليفين للتجانس.
- كولمجروف سميرنوف– شبيرويلك للاعتدالية.

- مان ويتني.
- معامل ألفا كرونباخ.

نتائج البحث:

الإجابة عن السؤال الأول: هل يتمتع مقياس المتغيرات السيكومترية للاختبار التكيفي المحوسب بدرجة مقبولة من الصدق؟

وللإجابة عن السؤال الأول؛ تم استخدام صدق المحكمين كمياً وكيفياً، وتم استخدام الصدق التقاربي، تم عرض المقياس في صورته الأولية على (١٦)[١] أستاذ جامعي، وذلك للتحقق من مدى صحة صياغة المفردات لغوياً، ومدي ملاءمتها للعينة، واتفق المحكمون على كل مفردات المقياس كيفياً وذلك بحصول جميع العبارات على نسبة مئوية تتخطي ال ٧٠٪، ويوضح هذا تفصيلاً في جدول (١)، مع تقديم بعض التعديلات اللغوية التي تم العمل بها.

جدول (١)
النسبة المئوية لاتفاق المحكمين

م	رقم المفردة	عدد المحكمين	موافق	غير موافق	النسبة
١	١-٢-٣-٦-٨-٩-١٤-١٥-١٨-١٩-٢١-٢٢-٢٣-٢٤-٣٢-٣٣-٣٦	١٦	١٦	٠	١٠٠٪
٢	٧-١٠-١١-١٢-١٣-١٦-٢٠-٢٥-٢٧-٣١-٣٤		١٥	١	٩٣,٧٥٪
٣	٤-٢٦-٢٨-٢٩-٣٠-٣٥		١٤	٢	٨٧,٥٪

[١] أسماء السادة المحكمين: أ.د/ محمد حبشي حسين، أ.د/ حجاج غانم أحمد، أ.د/ ربيع عبده أحمد رشوان، أ.د/ عبد الناصر السيد عامر، أ.د/ أحمد علي حسن المعمري، أ.د/ محمد محمد فتح الله، أ.د/ عبد المنعم احمد حسين، أ.د/ فريال أحمد أبو عواد، أ.د/ حلمي الفيل، أ.د/ أسامة أحمد عطا، أ.د/ رباب صلاح الدين إسماعيل، أ.د/ ديانا فهمي حمّاد، أ.د/ علي صلاح عبد المحسن، أ.د/ ليلي عابد طوخي، د. أشرف سعد جادالله، د/ دعاء هريدي.

٤	٥-١٧	١٢ ٤	٧٥٪

معامل اتفاق كيندال(Kendall's w) :

تم ايضاً عرض المقياس على (١٦) أستاذ جامعي، لمعرفة مدى انتماء كل مفردة للبعد، عن طريق إعطاء علامة صح للتقدير الذي تم اختياره، والتقديرات متمثلة في (ينتمي بدرجة كبيرة – ينتمي بدرجة متوسطة – ينتمي بدرجة منخفضة)، وتقابل تلك التقديرات الدرجات الآتية على الترتيب (٣)، (٢)، (١)، وتم تحليل البيانات باستخدام أسلوب معامل اتفاق كيندال، حيث أشارت النتائج إلى وجود اتفاق بين أراء المحكمين، حيث كانت قيمة معامل الاتفاق (٠.٨٦) وهذه القيمة تحقق الدلالة العملية، ويتفق هذا مع ما أشار إليه Rovai et al. (2013) أن قيمة حجم تأثير معامل اتفاق كيندال تتراوح من ٠ الي ١، حيث أن من ٠ الي ٠,٢٠ تأثير ضعيف جدا، و من ٠,٢٠ الي ٠,٤٠ تأثير ضعيف، ومن ٠,٤٠الي ٠,٦٠ تأثير معتدل، ومن ٠,٦٠ الي ٠,٨٠ تأثير قوي، ومن ٠,٨٠الي ١,٠٠ تأثير قوي جدا. كما أشارات النتائج أيضا الي ان مستوى الدلالة الإحصائية أقل من ٠.٠٥، مما يحقق الدلالة الإحصائية، كما هو موضح بالجدول (٢).

جدول (٢)

قيمة معامل اتفاق كيندال

م	عدد المحكمين	قيمة معامل كيندال	كا^{٢}	مستوي الدلالة
١	١٦	٠.٨٦	٤٨٣.٥١٩	٠.٠٠٢

الصدق التقاربي:

الموقف العام استخدام تحليل التباين متعدد المتغيرات التابعة أحادي الاتجاه (MANOVA)؛ وذلك لوجود متغير مستقل واحد وهو المعلم، ومتغيرات تابعة متعددة وهي أبعاد المقياس الستة، ولكن قبل الشروع وفي التطبيق يجب التحقق من الافتراضات:

أ- التعددية الخطية:

وذلك عن طريق حساب الارتباطات بين المتغيرات التابعة الستة، وكانت الارتباطات في أغلبها أقل من ٠.٥ كما هو موضع في جدول (٣)، وبرغم عدم وجود تعددية خطية إلا أن ضعف الارتباطات سينقلنا من تحليل التباين متعدد المتغيرات التابعة أحادي الاتجاه (MANOVA)، إلى تحليل التباين مستقل المتغيرات التابعة (ANOVA).

جدول (٣)

قيمة معاملات ارتباط بيرسون

		الغش	سهولة التصحيح	دقة التصحيح	التكاليف المادية	تغطية الأهداف	تفضيلات المهتمين
الغش	Pearson Correlation	1	-.001	-.073	-.173*	.181*	.107
	Sig.		.988	.305	.015	.010	.131
سهولة التصحيح	Pearson Correlation	-.001	1	.132	.223**	.301**	.213**
	Sig	.988		.062	.002	.000	.002
دقة التصحيح	Pearson Correlation	-.073	.132	1	.309**	.103	.323**
	Sig	.305	.062		.000	.148	.000
التكاليف المادية	Pearson Correlation	-.173*	.223**	.309**	1	-.150*	.017
	Sig	.015	.002	.000		.034	.814
تغطية الاهداف	Pearson Correlation	.181*	.301**	.103	-.150*	1	.582**

	Sig.	.010	.000	.148	.034		.000
تفضيلات المهتمين	Pearson Correlation	.107	.213**	.323**	.017	.582**	1
	Sig.	.131	.002	.000	.814	.000	

الاعتدالية والتجانس:

أ- **التجانس:** تم اختبار التجانس باستخدام ليفين للمجموعتين (معلمو المرحلة الإعدادية والثانوية) مقابل كل متغير من المتغيرات التابعة، واتضح أن جميع البيانات متجانسة وذلك لأن قيمة الدلالة الإحصائية أكبر من ٠.٠٥ كما هو موضوح بالجدول (٤)

جدول (٤)
قيمة اختبار ليفين للتجانس

م	البعد	ليفين	مستوي الدلالة
١	الغش في الامتحان	1.524	٠.٢١٨
٢	سهولة التصحيح	1.424	٠.234
٣	دقة التصحيح	2.083	٠.150
٤	التكاليف المادية	٥.١١٨	٠.٠٢٥
٥	تغطية الاهداف	٠.٢٥٩	٠.٦١١
٦	تفضيلات المهتمين	٠.١٨٧	٠.٦٦٦

ب-**الاعتدالية:** تم اختبار الاعتدالية باستخدام اختبار كولمجروف سميرنوف-شبيرويلك، ولم تتحقق الاعتدالية وذلك لأن جميع مستويات الدلالة أقل من ٠.٠٥، كما هو موضح بجدول (٥) وبناءً عليه يتم الانتقال من تحليل التباين مستقل المتغيرات التابعة (ANOVA)، إلى مان ويتني.

جدول (٥)
قيم اختباري كولمجروف سميرنوف- شبيرويلك

م	المرحلة	كولمجروف سميرنوف	مستوي الدلالة	شبيرويلك	مستوي الدلالة
١	الإعدادية	٠.178	٠.000	٠.878	٠.000
٢	الثانوية	٠.153	٠.000	٠.911	٠.000
٣	الإعدادية	٠.185	٠.000	٠.947	٠.001
٤	الثانوية	٠.212	٠.000	٠.929	٠.000
٥	الإعدادية	٠.132	٠.000	٠.965	٠.010
٦	الثانوية	٠.106	٠.007	٠.960	٠.004
٧	الإعدادية	٠.152	٠.000	٠.923	٠.000
٨	الثانوية	٠.191	٠.000	٠.922	٠.000
٩	الإعدادية	٠.165	٠.000	٠.949	٠.001
١٠	الثانوية	٠.131	٠.000	٠.946	٠.000
١١	الإعدادية	٠.137	٠.000	٠.937	٠.000
١٢	الثانوية	٠.138	٠.000	٠.932	٠.000

اختبار مان ويتني:

وبتطبيق اختبار مان ويتني تبين عدم وجود فروق بين متوسطي رتب الدرجات في كل بعد من الأبعاد الستة، بين مجموعة معلمين التعليم الإعدادي ومجموعة معلمين التعليم الثانوي، كما هو موضح في الجدول التالي:

جدول (٦)

قيم اختبار مان ويتني

م		البعد الأول	البعد الثاني	البعد الثالث	البعد الرابع	البعد الخامس	البعد السادس
١	مان ويتني	4623.500	4474.000	4860.000	4357.500	4990.000	4824.000
٢	مستوي الدلالة	.346	.183	.730	.111	.980	.665

وعدم الدلالة تحقق الصدق التقاربي؛ لأنها تعنى عدم وجود فروق بين وجهات نظر معلمي المرحلة الإعدادية ومعلمي المرحلة الثانوية في المتغيرات السيكومترية، أي أنهما متقاربان في الرؤى، ويفسر الباحث ذلك بأن عدم وجود فروق في وجهات النظر بين المجموعتين قد يرجع إلى أن برنامجهما في مؤسسات الإعداد واحد، وهي كلية التربية، حيث

برنامج التعليم العام، وتحقق الصدق التقاربي عن طريق المقارنة بين وجهات نظر المجموعتين يؤكد ان الأداة تتسم بمؤشرات صدق مقبولة، ويؤكد هذا ما جاءت به دراسة (قاسم، ٢٠٢٠) التي عملت على تصميم اختبار تكيفي محوسب متعدد المراحل للتحقق من أثره على دقة قياس الاستدلال البصري، وأظهرت النتائج فاعلية الاختبار التكيفي المحوسب متعدد المراحل من خلال دقة تقديره لقياس الاستدلال البصري.

وأيضاً ما أشار إليه Thompson and Weiss (2019) أن الاختبار التكيفي المحوسب يختلف عن غيره من صور الاختبارات الأخرى، حيث إن له خصائص فريدة ومرحل إعداد وتطبيق مختلفة، وبالتالي مشكلاته الخاصة، ولذلك تطوير الاختبار التكيفي المحوسب يحتاج إلى خبره كبيرة في مجال القياس النفسي، لذا تترك عملية تطويره إلى المتخصصين العاملين بالمجال، وأكدت دراسة (حجازي والشريف، ٢٠١٧) على وجود فروق ذات دلالة إحصائية بين معاملي الصدق المحكي للاختبار، لصالح الصورة الورقية.

ولكن هناك دراسات لا تتفق مع هذه النتيجة، مثل دراسة (Lu and Cong, 2016) التي أشارت إلى أن الاختبار التكيفي المحوسب يحتاج إلى تكلفة عالية، وأيضاً مشكلة كيفية إنشاء عناصر عالية الجودة تلقائياً على نطاق واسع، ويعتمد حل هذه المشكلات على التطبيق العميق لتكنولوجيا الحاسب الآلي في تنفيذ عملية الاختبار التكيفي، وأنه على الرغم من أن الاختبارات التكيفية قد حققت تقدماً كبيراً، إلا أنها كشفت تدريجياً عن بعض المشكلات وأوجه القصور في سنوات الممارسة وتحتاج إلى مزيد من الدراسة والتحسين.

ويرجع الباحث هذا الاختلاف إلى أن الاختبار التكيفي المحوسب يحتاج إلى الإلمام التام بمهارات الحاسب الآلي، ويحتاج إلى الدراية بمجال القياس والتقويم، ويؤثر هذا على المتغيرات السيكومترية المتعلقة بالاختبار، مما يجعل تطبيقه غير صالح في بعض الأحيان.

الإجابة عن السؤال الثاني: هل يتمتع مقياس المتغيرات السيكومترية للاختبار التكيفي المحوسب بدرجة مقبولة من الثبات

وللإجابة عن السؤال الثاني؛ تم حساب قيمة معامل ثبات "ألفا كرونباخ" الخاصة بالمقياس ككل، ولقد بلغت قيمته (٠.٧٣٠)، ويدل هذا على ان المقياس يتمتع بدرجة مقبولة من الثبات.

وبناءً على ما سبق؛ يتضح ان أن الأداة المُعدة لقياس بعض المتغيرات السيكومترية المرتبطة بالاختبارات التكيفية المحوسبة من وجهة نظر المهتمين، تتمتع بخصائص سيكومترية مقبولة.

التوصيات:

١-إجراء المزيد من البحوث المعتمدة على الاختبارات التكيفية المحوسبة

٢-تشجيع القائمين على إعداد الاختبارات في المؤسسات التعليمة على استخدام الاختبارات التكيفية المحوسبة.

٣-تدريب كل من المعلم والمتعلم على كيفية استخدام التقنيات التكنولوجية الحديثة؛ وذلك لجعل الاختبارات التي تعتمد على التقنيات الحديثة كالاختبار التكيفي المحوسبة ممكنة التطبيق في جمهورية مصر العربية.

المراجع:

الكيال، مختار أحمد ونور الدين، أمين صبري والظنحاني، ميسون محمد. (٢٠١٧). الخصائص السيكومترية لاختبار تواؤمي محوسب لتشخيص صعوبات تعلم الرياضيات. *مجلة الإرشاد النفسي*، ٥٠(٥٠- الجزء الثاني)، ٢١٣-٢٥٤.

حجازي، تغريد، والشريف، نايف. (٢٠١٧). الخصائص السيكومترية لصورتي اختبار في الرياضيات للمرحلة المتوسطة: دراسة مقارنة. *مجلة الدراسات التربوية والنفسية*، ١١(١)، ١٩٠-١٧٨.

دعنا، زينات يوسف. (٢٠١٢). بناء برنامج اختباري محبوك محوسب وفق الاستراتيجية ثنائية المرحلة مع نموذج تطبيقي لاختبار محبوك ثنائي المرحلة في مقرر مدخل إلي علم النفس. *المجلة العربية للعلوم الاجتماعية*، ١ (١)، ٢٣٠-٢٠٥.

عطية، محي الدين محمد (٢٠١٥). *الإدارة المدرسية الناجحة*. الوراق للنشر والتوزيع.

قاسم، مصطفي أحمد (٢٠٢١). *الاختبارات التكيفية المحوسبة: النظرية والتطبيق*. دار الفكر العربي: القاهرة.

قاسم، مصطفي أحمد. (٢٠٢٠). *تصميم الاختبار التكيفي المحوسب متعدد المراحل وأثره علي دقة قياس الاستدلال البصري باستخدام نظرية الاستجابة للمفردة "دراسة محاكاة"* [رسالة دكتوراه غير منشورة] . كلية البنات للآداب والعلوم التربوية بجامعة عين شمس.

نور الدين، أمين صبري. (٢٠١٩). فاعلية القياس التكيفي المحوسب في قياس تحصيل طلاب الجامعة. *المجلة السعودية للعلوم النفسية*، (٦٤)، 29-47.

Aşkar, P., Altun, A., Cangöz, B., Çevik, V., Kaya, G., & Türksoy, H. (2012). A comparison of paper-and-pencil and computerized forms of line orientation and enhanced cued recall tests. *Psychological reports*, *110*(2), 383-396. DOI: https://doi.org/10.7275/wqzt-9427

Kimura, t. (2017). The impacts of computer adaptive testing from a variety of perspectives. *Journal of Educational Evaluation for Health professions*,*14*,1-5. https://doi.org/10.3352/jeehp.2017.14.12

Lu, p., and cong, x. (2016). The Research on Computerized Adaptive Testing. *Journal of physics,710*(1),1-10. DOI 10.1088/1742-6596/710/1/012029

Parshall, c.g.,Spray,j.a., Kalohn, j.c., & Davey, t.(2002). *Practical Considerations in Computer-Based Testing*. Springer Science +Business Media.

Rezaie, M., & Golshan, M. (2015). Computer adaptive test (CAT): Advantages and limitations. *International Journal of Educational Investigations, 2*(5), 128-137.

Rovai, A. P., Baker, J. D., & Ponton, M. K. (2013). *Social science research design and statistics: A practitioner's guide to research methods and IBM SPSS*. Watertree Press LLC.

Thomas, m.l., Brown, g.g., Patt, v.m., & Duffy, j.r. (2021). Latent Variable Modeling and Adaptive Testing for Experimental Cognitive Psychopathology Research. *Educational and PsychologicalMeasurement,81*(1),155-181. DOI: 10.1177/0013164420919898

Thompson, Nathan A. and Weiss, David A.(2019) AFramework for the Development of Computerized AdaptiveTests," *Practical Assessment, Research, and Evaluation,16*(1), 1-11.

Triantafillou, E., Georgiadou, E., & Economides, A. (2008). The Design and evaluation of a Computerized adaptive test on mobile devices. *Computers & Education*, *50*(4),1319-1330.

Veldkamp, B.p., & Matteucci, M. (2013). Bayesian Computerized Adaptive Testing. *Enasio: aval.Pol.Publ. Edu,21*(78),57-82.

Yasuda, j.i., Mae, n., Hull, m.m., &Taniguchi, m.a. (2021). Optimizing the length of computerized adaptive testing for the Force Concept Inventory. *Physical Review Physics Education Research,17*(1),1-15. DOI: 10.1103/PhysRevPhysEducRes.17.010115