أثر طريقة عرض مفردات الاختبار الإلكتروني وعددها في خصائصه السيكومترية والأداء عليه

إعداد
د/ محمد محمود محمد عبد الوهاب
قسم علم النفس التربوي كلية التربية – جامعة المنيا

المستخلص:

سعى البحث إلى التعرف على أثر طريقة عرض مفردات الاختبار الإلكتروني (الطريقة الكلية – الطريقة التتابعية مع حظر الرجوع – الطريقة التتابعية بدون حظر الرجوع) وعدد مفرداته (٣٠ – ٤٠ مفردة) في كل من الخصائص السيكومترية للاختبار، والأداء عليه، وذلك لدى عينة قوامها (٤٥٠) طالبة من طالبات كلية العلوم الإقتصادية والموارد الإسلامية بجامعة أم القرى، تم تقسيمها إلى (٦) مجموعات، وكان قوام كل مجموعة (٧٥) طالبة، بحيث تتعرض كل منها إلى اختبار تحصيلي إلكتروني مختلف في طريقة عرض مفرداته و / أو عددها، وبالمقارنة بين المجموعات تبين وجود فروق دالة إحصائيًا تعزى إلى اختلاف طريقة عرض المفردات – عند تثبيت عددها – في كل من معاملات ارتباط المفردات المصححة بالدرجة الكلية ومعاملات صعوبة المفردات ومعاملات تمييزها وأداء الطالبات على الاختبار، كما تبين عدم وجود فروق دالة إحصائيًا تعزى إلى عدد مفردات الاختبار – عند تثبيت طريقة عرضها – في معاملات ارتباط المفردات المصححة بالدرجة الكلية وأداء الطالبات على الاختبار، وقد ظهرت فروق دالة إحصائيًا تعزى لعدد مفردات الاختبار عند عرض المفردات بالطريقة التتابعية بدون حظر الرجوع فقط في كل من معاملات صعوبة المفردات ومعاملات تمييزها، كما أسفرت النتائج عن اختلاف بين الصور الاختبارية في معاملات ارتباط درجات الطالبات عليها بمعدلاتهن التراكمية، وفي معاملات ثباتها.

الكلمات المفتاحية: الاختبار الإلكتروني، طريقة عرض المفردات، عدد المفردات، الخصائص السيكومترية للاختبار الإلكتروني، الأداء على الاختبار الإلكتروني.

The effect of displaying method of electronic test items and their number on its psychometric properties and performance on it

BY

Dr/ Mohammed M. M. Abdul-wahab
Associate professor in dept. of Educational Psychology
Faculty of Education - Minia University - Egypt

Abstract

The research conducted to identify the effect of displaying method of the electronic test items (Total method – Sequential method with forbidden of return – Sequential method without forbidden of return) and number of its items (30-40 items) on each of the psychometric properties of the test, and performance on it, for a sample of female students in College of Economic Sciences and Islamic Resources at Umm Al-Qura University consists of (450) female students, they were divided into (6) groups, each group consists of (75) female students, so that each group is subjected to a different electronic achievement test in displaying method and / or number of items, and by comparison between groups It was found that there were statistically significant differences attributable to the different of displaying method - when fixing number of items - in each of corrected item -total correlation coefficients, items difficulty coefficients, discrimination coefficients and performance of students on the test. It was also found that there were no statistically significant differences attributable to number of test items - when displaying method was fixed - in corrected item -total correlation coefficients and the performance of the students on it. Statistically significant differences due to number of test items appeared only in each of item difficulty coefficients and discrimination coefficients when displaying item by item with prohibition of return.

Key words: The electronic test, Method of displaying items, Number of items, Psychometric properties of the electronic test, Performance on the electronic test

مقدمة:

شهد العالم في القرن الحادي والعشرين تقدمًا تقنيًا هائلًا، وتوظيفًا فاعلًا مستمرًا للتقنية الحديثة في مجال التعليم والتقويم؛ حيث توجهت الأنظمة والمؤسسات التعليمية في أغلب دول العالم إلى اعتماد أساليب وأدوات التقويم الإلكتروني لمخرجات التعلم، وقد توسعت معظم مؤسسات التعليم الجامعي في توظيف المستحدثات التكنولوجية في قياس وتقويم أداء طلبتها، وجعلت ذلك من أهم أولوياتها وعنصرًا أساسيًا في خطط واستراتيجيات تحسينها وتطويرها؛ بغية مواكبة التطور المتنامي في تكنولوجيا المعلومات.

وتعد الاختبارات الإلكترونية إحدى أدوات التقويم الإلكتروني التي انتشرت في العقدين الأخيرين بسبب تطور البرمجيات والتكنولوجيا (الزامل والحجيلان، ٢٠١٦) والتوجه لاستخدام منظومة بنوك الأسئلة في قياس الجوانب المعرفية كأحد نواتج التعلم أو القدرات العقلية المختلفة أو الكفاءات اللغوية (دلالعة ودلالعة، ٢٠١٩)، وكذلك التوجه نحو القياس التكيفي المفصل المحوسب Computerized adaptive measurement في القياسات النفسية (البربري، ٢٠٢٠؛ نور الدين، ٢٠٠٣)، وانتشار الجامعات الإلكترونية وما صاحبها من تعليم وتقييم عن بعد (عبيس، ٢٠٢٠).

ورغم انتشار الاختبارات الإلكترونية في كثير من دول العالم، إلا أن بعض المؤسسات التعليمية قد تأخرت في اعتمادها، إلى أن دعت الحاجة الملحة لها كإجراء احترازي للحد من انتشار جائحة كورونا (كوفيد ١٩) في عام ٢٠٢٠م (Wang, Kao & Chen, 2021)، والتي أدت إلى إغلاق أغلب الأنظمة التعليمية، واللجوء إلى الاختبارات الإلكترونية كوسيلة بديلة للاختبارات الورقية التقليدية.

وقد تركزت جهود علماء القياس النفسي لفترة طويلة من الزمن للكشف عن المتغيرات والعوامل التي تؤثر في الخصائص السيكومترية للاختبارات الورقية، وكذلك المتغيرات والعوامل التي تؤثر في أداء المختبرين عليها، إلا أن جهودهم في الكشف عن تأثير العوامل والمتغيرات المتعلقة بالاختبارات الإلكترونية سواء عند تصميمه أو عرض مفرداته أو ما يترتب عليه من تأثيرات أو نتائج ما تزال حديثة نسبيًا.

وتختلف طريقة عرض مفردات الاختبارات الإلكترونية على المختبرين، فقد فرق السيد وحسن (٢٠٢١) بين طريقتين للعرض، هما الطريقة الكلية والطريقة التتابعية، وقد أشارا إلى أن الطريقة الكلية يتم فيها عرض مفردات الاختبار بشكل كلي في شاشة واحدة، ويستطيع الطالب بذلك التنقل بحرية تامة بين المفردات أثناء الإجابة عنها، أما الطريقة التتابعية فهي التي يتم فيها عرض مفردات الاختبار بشكل تدريجي تتابعي يظهر للطالب فيها مفردة واحدة في الشاشة في كل مرة. وقد يمنع المستجيب في العرض التتابعي من العودة إلى المفردات التي أجاب عنها أو تركها مسبقًا إذا أراد تغيير إجاباته عنها عند المراجعة (خيار حظر الرجوع)، كما قد يقدم له خيار بترك المفردة بدون إجابة على أن يتم عرضها عليه مرة أخرى في نهاية الاختبار (محمود، ٢٠١٨).

ويطلق على العرض الكلي للمفردات عدة مسميات، مثل طريقة التمرير Scrolling (Klausch, et al., 2012)، أو طريقة المصفوفة Matrix (Liu & Cetant, 2018)، وطريقة الشبكة Grid (Callegaro, Shand-Lubbers & Dennis, 2009)، كما يطلق على طريقة العرض التتابعي للمفردات عدة مسميات، مثل طريقة المفردة الواحدة Single item (Klausch, et al., 2012) وطريقة مفردة تلو مفردة item by item (Liu & Cetant, 2018).

وقد اختلفت رؤى الباحثين حول أفضلية طرق عرض مفردات الاختبار الإلكتروني، فقد أوضح ليسون (Leeson, 2006) أن طريقة العرض التتابعي تجعل المختبر يتسرع في الإجابة ويزيد معدلات الخطأ؛ وهذا يؤثر سلبًا في الأداء؛ أما العرض الكلي للمفردات في شاشة واحدة يسمح للمختبر معاينة جميع المفردات، وتخطي ما يصعب عليه إجابتها وتأجيلها لنهاية الاختبار، وهذا يؤثر إيجابًا في أدائه، وقد أشار حميد (٢٠٢٠) أن العرض الكلي للمفردات يسهل على المختبر معاينة إجابته السابقة، وإعادة إجابة أي سؤال بشكل فوري دون تضييع وقت في البحث عنه، كما أن عرض المفردات بطريقة تتابعية يعد الأفضل عند وجود الطلبة معًا في نفس قاعة الاختبار.

بينما أشار خميس (٢٠٠٣) إلى أن العرض الكلي للمفردات على الطالب في شاشة واحدة يزيد من تشتت انتباهه، ويطيل زمن قراءة المفردات، ويقلل العرض التتابعي للمفردات

من قلق الطالب ويخفض من الضغوط النفسية لديه، وأكد عبد الرحيم ومرعي (٢٠٢٠) أن عرض مفردة واحدة في كل شاشة يقلل من تعرض المختبر لمشتتات الانتباه التي قد تحدث في حالة عرض المفردات كلها في شاشة واحدة، ورأى ويلسون وكول (Wilson & Cole, 1996) أن العرض الكلي للمفردات في شاشة واحدة تزيد من العبء المعرفي لدى المستجيب بسبب تعرضه لعدد كبير من المثيرات، وأشار كوبر ورفاقه (Couper, et al., 2013) إلى أن الطريقة الكلية تكون شاقة وتثني المستجيب عن استكمال الجهد المعرفي أثناء الاستجابة كلما زاد عدد المفردات في الشاشة، كما أوضحت مبارز وفخري (٢٠١٣) أن الطلبة يسهل عليهم إجراء الاختبار الإلكتروني عندما يكون هناك سؤال واحد في كل صفحة مع توضيح عدد المفردات التي تم حلها وعدد المفردات المؤجل الإجابة عنها وعدد المتبقية، مما يساعد الطلبة في إدارة وقت الاختبار، كما أكد الفيل (٢٠١٤) أن محدودية سعة الذاكرة العاملة تجعل من الصعب على المستجيب الانتباه لعدة مفردات في وقت واحد، وهذا يجعل العرض التتابعي للمفردات أكثر ملاءمة من العرض الكلي لها.

فيما يرى ليو وسيتانت (Liu & Cetant, 2018) أنه لا يزال من غير الواضح في أي موقف تكون إحدى طريقتي عرض المفردات أفضل من الأخرى، كما أكد الوكيل (٢٠٢١) أنه لا يوجد اتفاق بين الباحثين والخبراء على طريقة العرض الأفضل للمفردات (الكلي أم التتابعي)، وأن طريقة العرض الأنسب تعتمد على طبيعة البيئة التي يطبق فيها الاختبار ومدى جاهزية المؤسسة التعليمية تكنولوجيًا، ومدى تمكن الطلبة من مهارات التقنية. ومن ثم يتبين أن طريقة عرض مفردات الاختبار الإلكتروني تؤثر في متغيرات عقلية وانفعالية ومهارية، مثل الانتباه أو التذكر أو العبء المعرفي أو إدارة الوقت أو التغذية الراجعة أو التسرع في الإجابة أو القلق الإلكتروني ... الخ، وهذه المتغيرات قد تؤثر بشكل أو بآخر في أداء الطلبة على الاختبار الإلكتروني.

هذا، وقد ينتقل تأثير المتغيرات العقلية والمهارية والانفعالية الناجمة عن طريقة عرض مفردات الاختبار الإلكتروني إلى خصائص الاختبار نفسه، ففي ظل ارتباط الخصائص السيكومترية للاختبارات بأداء المختبرين؛ فإن أي تأثير قد يحدث في أداء المختبرين نتيجة اختلاف طريقة عرض مفردات الاختبار قد يؤدي إلى تغيير في صدق الاختبار وثباته

ومعاملات صعوبة مفرداته ومعاملات تمييزها. ومن زاوية أخرى أشار كاليجارو وشاندلوبز ودينيز (Callegaro, Shand-Lubbers & Dennis, 2009) أنه من المتوقع أن يترتب على العرض الكلي للمفردات (في صورة شبكة grid) انخفاض الوقت المستغرق في الإجابة عنها، ووجود عدد أكبر من المفردات المفقودة (بدون إجابة)، وعدد أكبر من المفردات اللاتمايزية بين المختبرين، وارتفاع أكبر في أخطاء القياس، وهذا كله قد يؤثر في الخصائص السيكومترية للاختبار؛ ومن ثم قد يؤثر في دقة الاختبار وصلاحيته للقياس.

ومن زاوية أخرى فإن طول الاختبار الإلكتروني أو عدد مفرداته يعد أحد المتغيرات التي نالت اهتمام الباحثين في مجال القياس النفسي لمعرفة العوامل التي يتم تحديده في ضوئها، سواء كانت عدد مخرجات التعلم المراد قياسها، أو الزمن المخصص للاختبار إن كان محددًا مسبقًا، أو طبيعة المرحلة التي يمر بها المختبر، أو نوع المفردات، وفي الاختبارات الإلكترونية يضاف لهذه العوامل المحددة لطول الاختبار عوامل أخرى تتعلق بالإمكانات التقنية المتاحة، ومدة التعرض المناسبة للشاشة الإلكترونية، والقلق الإلكتروني والمهارات التقنية ... الخ.

وفي ضوء عدد مفردات الاختبار التي تعرض على المختبرين تصنف الاختبارات الإلكترونية إلى اختبارات إلكترونية تقليدية يتم فيها تثبيت طبيعة مفرداتها وعددها لكل المختبرين، وهذا النوع من الاختبارات الإلكترونية هو ما سيتم تناوله في البحث الحالي، وهناك اختبارات إلكترونية تكيفية والتي تصنف بدورها حسب طولها إلى نوعين، هما اختبارات تكيفية ثابتة الطول، وفيها يتم عرض نفس العدد من المفردات على جميع المختبرين ولكن تختلف طبيعة المفردات المعروضة على المختبرين حسب قدراتهم وإجاباتهم عن المفردات الأولى التي تعرض عليهم، وكذلك اختبارات إلكترونية تكيفية متغيرة الطول، وفيها يتم عرض عدد متغير من المفردات ونوعية مختلفة من المفردات لكل مستجيب حسب قدراته وحسب قاعدة إنهاء الاختبار المتبعة (البربري، ٢٠٢٠؛ Conejos, et al., 2004; Meijer & Nering, 1999). كما تصنف الاختبارات الإلكترونية أيضًا إلى اختبارات ثابتة الطول واختبارات متعددة المراحل واختبارات تكيفية (Jodoin,2003; Jodoin, Zenisky & Hambleton, 2006).

ويعد تحديد عدد المفردات المناسب في الاختبار الإلكتروني أمرًا ضروريا للحصول على اختبار يتمتع بخصائص سيكومترية جيدة، مع ضرورة ألا يؤدي طول الاختبار إلى زيادة قلق المستجيبين التقليدي أو الإلكتروني، ومن ثم التأثير سلبيًا في أدائهم (صفوت وصبري، ٢٠٢٠)، كما أنه يجب أن يكون الاختبار طويلًا بما يكفي للحصول على اختبار هادف وثابت ويمكن المقارنة في ضوئه بين المختبرين، وفي نفس الوقت يجب أن يكون قصيرًا بما يكفي ليكون قابلًا للاستخدام والتطبيق (Schurig, Jungjohann & Gebhardt, 2021).

مشكلة البحث:

اتخذت الجامعات السعودية في العام الجامعي ٢٠٢٠ – ٢٠٢١م قرارًا بتطبيق الاختبارات الإلكترونية بدلًا من الاختبارات الورقية؛ بهدف الحد من انتشار جائحة كورونا (كوفيد ١٩)؛ لما ينتج عن ذلك من فرض التباعد الاجتماعي بين الطلبة أثناء العملية الاختبارية، ومنع تداول الأوراق الاختبارية التي قد تحمل الفيروس بين المعلم والطالب والمراقب وأعضاء لجان النظام والمراقبة والتي قد تزيد من فرص انتقال العدوى، ولكن اتخاذ القرار بالتحول المفاجئ من الاختبارات التقليدية إلى الاختبارات الإلكترونية دون وجود تدريب مسبق للمعلمين أو الطلبة، ودون وجود استعداد كافي لدى المؤسسات التعليمية، ودون وجود دراسات كافية عن تصميم هذه الاختبارات وعدد مفرداتها وأفضل طريقة لعرضها... الخ أدى إلى ظهور بعض أشكال التخبط في تصميمها أو الإجابة عنها.

ونظرًا لطبيعة الاختبارات الإلكترونية؛ فإن هناك بعض العوامل التي ترتبط بالاختبارات الإلكترونية دون الورقية، مثل: القلق الإلكتروني واعتياد استخدام الكمبيوتر (Berman, Haertel, & Pellegrino, 2020; Erdogan, 2008) وطريقة عرض الاختبار الإلكتروني وأسلوب التغذية الراجعة المصاحب له وحجم الشاشة ودرجة وضوحها ومدة عرض الفقرة على الشاشة (Bridgeman, Lennon & Jackenthal, 2001)، ومن ثم فإن الكشف عن الخصائص السيكومترية للاختبارات الإلكترونية ومعرفة العوامل التي يمكن أن تؤثر فيها يجب أن ينال قدرًا كافيًا من الدراسة، خاصة أن الدراسات السيكومترية لم تسفر عن نتائج مؤكدة بشأن تحقق التكافؤ بين الصورتين الإلكترونية والورقية للاختبار (العاجز، ٢٠١٩؛ Kingston, 2009; Wang, Kao & Chen, 2021)، كما أن

هناك حاجة ملحة لمعرفة العوامل التي قد تؤثر سلبًا أو ايجابًا في أداء المختبرين على الاختبارات الإلكترونية من أجل تفاديها أو الإبقاء عليها ومن ثم الحصول على أفضل تقدير ممكن لهذا الأداء؛ في ظل وجود ندرة في الدراسات التي اهتمت بتأثير المتغيرات المصاحبة للاختبار الإلكتروني على هذا الأداء سواء عند تصميمه أو عند عرضه على المختبرين أو بعد عرضه عليهم خاصة الدراسات العربية، وهذا ما أكده أيضًا الوكيل (٢٠٢١).

وقد ربط ليو وسيتانت (Liu & Cetant, 2018) بين طريقة عرض مفردات الاختبار الإلكتروني وعددها عندما أشارا إلى تفضيل الاعتماد على النمط الكلي لعرض مفردات الاختبار عندما يكون عددها قليل نسبيًا، أما نمط عرض المفردات التتابعي فيفضل استخدامه عندما يكون عدد المفردات كبيرًا نسبيًا، وقد أكد البربري (٢٠٢٠) أنه يمكن خفض قلق الطلبة من الاختبار الإلكتروني من خلال التحكم في طول الاختبار الإلكتروني وتصميمه وتطويره، كما أشارت صفوت وصبري (٢٠٢٠) إلى أن أسلوب عرض مفردات الاختبار الإلكتروني وعددها يعدان من العوامل التي تؤثر في قلق الاختبار الإلكتروني، ومن ثم فقد تؤثر في الأداء عليه، ومن ثم فإن طريقة عرض مفردات الاختبار الإلكتروني وعددها قد يؤثران في القلق الإلكتروني لدى الطلبة وهذا ما أكده أبو الشيخ (٢٠١٨) ، كما أنهما قد يؤثران في القدرة على الانتباه لديهم والعبء المعرفي وكفاءة الذاكرة العاملة (الفيل، ٢٠١٤)، وهذه العوامل كلها قد تؤثر في أداء المختبرين على هذه الاختبارات الإلكترونية؛ وبالتالي قد تؤثر في الخصائص السيكومترية لهذه الاختبارات، وهذا ما يسعى البحث الحالي إلى دراسته، ومن ثم فإن مشكلة البحث الحالي تثير السؤال الرئيس التالي: "ما أثر طريقة عرض الاختبار الإلكتروني وعدد مفرداته في الخصائص السيكومترية له وأداء المختبرين عليه؟"، ويتفرع من هذا التساؤل التساؤلات الفرعية الآتية:

١- ما الدلالة الإحصائية للفروق في معاملات ارتباط المفردات المصححة بالدرجة الكلية على الاختبار التي تعزى إلى:

١- أ) طريقة عرضها مع تثبيت عددها؟

١- ب) عددها مع تثبيت طريقة عرضها؟

١- ج) اختلاف طريقة عرضها وعددها؟

٢- ما ترتيب معاملات ارتباط درجات الطالبات على الاختبار الإلكتروني بالمعدل التراكمي باختلاف طريقة عرضها وعددها؟

٣- ما الدلالة الإحصائية للفروق في متوسطات معاملات صعوبة مفردات الاختبار الإلكتروني التي تعزى إلى:

٣- أ) طريقة عرضها مع تثبيت عددها؟

٣- ب) عددها مع تثبيت طريقة عرضها؟

٣- ج) اختلاف طريقة عرضها وعددها؟

٤- ما الدلالة الإحصائية للفروق في متوسطات معاملات تمييز مفردات الاختبار الإلكتروني التي تعزى إلى:

٤- أ) طريقة عرضها مع تثبيت عددها؟

٤- ب) مع تثبيت طريقة عرضها؟

٤- ج) اختلاف طريقة عرضها وعددها؟

٥- ما ترتيب معاملات ثبات الاختبار الإلكتروني للصور الاختبارية المختلفة وفقًا لطريقة عرض المفردات وعددها؟

٦- ما الدلالة الإحصائية للفروق في متوسطات درجات الطلبة على الاختبار الإلكتروني التي تعزى إلى:

٦- أ) طريقة عرضها مع تثبيت عددها؟

٦- ب) عددها مع تثبيت طريقة عرضها؟

٦- ج) اختلاف طريقة عرضها وعددها؟

أهداف البحث:

يسعى البحث الحالي إلى:

١- استكشاف مدى تأثير طريقة عرض المفردات وعددها في الخصائص السيكومترية للاختبار الإلكتروني (معاملات ارتباط المفردات المصححة بالدرجة الكلية – معاملات ارتباط درجات الاختبار الإلكتروني بالمعدل التراكمي – معاملات صعوبة المفردات – معاملات تمييز المفردات – معاملات ثبات الاختبار).

٢- استكشاف مدى تأثير طريقة عرض المفردات وعددها في أداء الطالبات عليه.

أهمية البحث:

تتمثل أهمية البحث الحالي فيما يلي:

- يعد البحث الحالي من البحوث العربية القليلة التي تناولت تأثير بعض المتغيرات على الخصائص السيكومترية للاختبار الإلكتروني وأداء المختبرين عليه.
- يتناول البحث موضوعًا مهمًا يتعلق بالاختبارات الإلكترونية التي كثر استخدامها أثناء جائحة كورونا كأداة للتقويم التربوي في المدارس والجامعات؛ من أجل الكشف عن تأثير طريقة عرض مفرداتها وعددها على أداء الطلبة والخصائص السيكومترية للاختبار.
- قد يساعد البحث الحالي القائمين على إعداد وتطبيق الاختبارات الإلكترونية في تحديد أفضل طرق عرض مفرداتها وعددها المناسب؛ من أجل تقدير أداء المختبرين بكفاءة؛ وقياس مخرجات التعلم بدقة عالية.
- قد يساعد البحث الحالي في اتخاذ القرار المناسب بشأن كيفية تطبيق الاختبارات الإلكترونية الطويلة التي تغطي محتوى كبيرًا، من خلال عدة جلسات اختبارية أو من خلال جلسة واحدة.
- قد تثير نتائج البحث الحالي وعي الباحثين في القياس والتقويم لإجراء مزيد من البحوث حول العوامل التي يمكن أن تؤثر في الخصائص السيكومترية للاختبار الإلكتروني أو في أداء المختبرين عليه.
- قد تؤدي نتائج البحث الحالي إلى تطوير إمكانات البرامج المستخدمة في إعداد وعرض الاختبارات الإلكترونية لتغيير طريقة عرض المفردات أو استيعاب عدد أكبر من المفردات

مصطلحات البحث:

- **<u>طريقة عرض مفردات الاختبار الإلكتروني</u>:** يعرفها الباحث على أنها: الطريقة التي يتم من خلالها تقديم مفردات الاختبار الإلكتروني للمختبرين، وذلك إما من خلال العرض الكلي للمفردات في شاشة واحدة فقط ويستخدم المختبر شريط التمرير لأسفل ولأعلى للوصول إلى كل المفردات في نفس الشاشة، أو العرض التتابعي للمفردات (كل مفردة في شاشة تواليًا) ويضغط المختبر على زر "التالي" للانتقال من المفردة إلى التي تليها، سواء كان هذا العرض التتابعي مع حظر الرجوع للمفردات السابقة أو مع السماح بالرجوع لها وتغيير الإجابة عنها.

- **الخصائص السيكومترية**: مجموعة مؤشرات وقيم تعبر عن إمكانية الثقة في نتائج أداة القياس، واستقرار النتائج واتساقها، والأسس التي يعتمد عليها في تفسيرها، وتتكون في الغالب من ثلاثة مؤشرات رئيسة، هي: الصدق، الثبات، والمعايير (Aiken,2000) ، وتتمثل في البحث الحالي في صدق الاختبار ومعاملات صعوبة المفردات ومعاملات تمييزها ومعامل ثبات الاختبار.
- **صدق الاختبار**: "أن يقيس الاختبار– أو المقياس– القدرة أو السمة أو الاتجاه أو الاستعداد الذي وضع لقياسه، أي يقيس فعلاً ما يقصد أن يقيسه" (عيسوي، ١٩٩٩: ٤٥)، ويستدل عليه في البحث الحالي من خلال الارتباط المصحح بين درجات الطالبات على المفردة والدرجة الكلية لهن على الاختبار، وكذلك من خلال الارتباط بين درجات الطالبات على الصور الاختبارية المختلفة ومعدلاتهن التراكمية.
- **معامل صعوبة المفردة**: هو حاصل قسمة عدد المختبرين الذين أجابوا بطريقة صحيحة عن المفردة إلى عدد أفراد العينة ككل (Aiken,2000)، وتعد المفردة مقبولة إذا تراوح معامل صعوبتها بين ٠.٢٥، ٠.٧٥.
- **معامل تمييز المفردة**: يشير إلى قدرة المفردة على التمييز بين المختبرين ذوي المستويات العليا (أعلى ٢٧%) والمختبرين ذوي المستويات الدنيا (أدنى ٢٧%) في المتغير موضع القياس (Aiken,2000)، ويحسب من خلال الفرق بين معامل الصعوبة في المجموعة العليا ومعامل الصعوبة في المجموعة الدنيا.
- **معامل ثبات الاختبار**: يشير إلى "مدى الخلو من الأخطاء غير المنتظمة التي تشوب القياس، أي مدى قياس المقدار الحقيقي للسمة التي يهدف لقياسها؛ فدرجات المقياس تكون ثابتة إذا كان يقيس سمة معينة قياساً متسقاً في الظروف المتباينة التي قد تؤدي إلى أخطاء القياس؛ فالثبات بهذا المعنى يعني الاتساق أو الدقة في القياس" (علام، ٢٠٠٤: ١٣١)، ويعبر عنه في البحث الحالي من خلال معامل كيودر ريتشاردسون للاختبار.
- **الأداء على الاختبار الإلكتروني**: يشير إلى مجموع درجات الطالبة على استجاباتها لمفردات الاختبار الإلكتروني، مع تحويل الدرجة على الاختبار إلى درجة سقفها (١٠٠)؛ لكي يمكن مقارنة أداء الطالبات على الصور الاختبارية المختلفة في عدد مفرداتها.

حدود البحث:

تمثلت حدود البحث الحالي فيما يلي:

- **الحدود الموضوعية:** حيث تناول البحث أثر طريقة عرض مفردات الاختبار الإلكتروني وعددها في كل من الخصائص السيكومترية للاختبار والأداء عليه مع تثبيت أثر حجم شاشة الجهاز المستخدم في الإستجابة.
- **الحدود البشرية:** عينة قوامها (٤٥٠) من طالبات كلية العلوم الإقتصادية والمالية الإسلامية بجامعة أم القرى
- **الحدود المكانية:** تم تطبيق البحث على الطالبات في فرع الزاهر بجامعة أم القرى في مكة المكرمة.
- **الحدود الزمانية:** تـم تطبيـق البحث في الفصل الدراسي الأول في العام الجامعي ٢٠٢٠ – ٢٠٢١م.
- **أدوات البحث:** تمثلت في صورتين مختلفتين في عدد المفردات لاختبار تحصيلي إلكتروني في مقرر الإحصاء الاستدلالي للإقتصاديين، وتقدم مفردات كل صورة من خلال ٣ طرق عرض مختلفة.

الإطار النظري:

مفهوم الاختبارات الإلكترونية والفرق بينها وبين الاختبارات الورقية التقليدية:

تعد الاختبارات الإلكترونية أكثر أدوات التقويم الإلكتروني استخدامًا؛ حيث تم استخدامها في قياس مخرجات التعلم المعرفية المستهدفة من العملية التعليمية والقدرات العقلية. وتعد نوعًا من الاختبارات المستخدمة لقياس نواتج التعلم من خلال شبكة الانترنت (عن بعد) أو الكمبيوتر داخل المؤسسة التعليمية (The International Test Commission, 2000)، كما أنها تلك الاختبارات التي تستخدم تكنولوجيا المعلومات والاتصالات في تصميم وإنتاج واستخدام وتقويم وإدارة الاختبارات (جليلة، ٢٠١٧)، وهي أيضًا إحدى تقنيات الحاسب الآلي التي يمكن توظيفها للتغلب على بعض الصعوبات والمعوقات التي تواجه الاختبارات الورقية، وتؤدي إلى زيادة التحصيل ومهارات التعلم الذاتي لدى الطالب (المرزوق، ٢٠٢٠).

وقد أوضح وانج وشين (Wang & Shin, 2010) وكذلك أبو الشيخ (٢٠١٨) أن الاختبارات الإلكترونية تطابق الاختبارات الورقية التقليدية إلا أنها تؤدى عن طريق الحاسب الآلي، وفيها يقوم المستجيب بقراءة المفردات من الشاشة والإجابة عنها من خلال أدوات الإدخال المتاحة، ولكن السيد وحسن (٢٠٢١) أشارا أن الاختبار لكي يعد إلكترونيا يجب أن يصمم ويطور ويستخدم ويدار ويصحح إلكترونيًا.

مزايا الاختبارات الإلكترونية:

تتعدد مزايا الاختبارات الإلكترونية في مقابل الاختبارات التقليدية؛ حيث تتميز الاختبارات الإلكترونية بأنها وسيلة فعالة لقياس مخرجات التعلم في حالات البعد المكاني بين المعلم والمتعلم (حميد، ٢٠٢٠)، ويمكن توظيف بعض الوسائط المتعددة كالصور والفيديوهات في صياغة مفرداتها (Krille, Wuttke & Seifried, 2017; Weir, Gwynllyw & Henderson, 2021)، ويمكن تقييد الاختبار بوقت محدد بدون تدخل العامل الإنساني (دلالعة ودلالعة، ٢٠١٩)، وتساعد بشكل كبير في تقديم تغذية راجعة فورية للطلبة وتسهيل مراجعة الإجابات ومعرفة الأخطاء (Bennett, 2002; Smolinsky, et al., 2020)، وتتيح فرصة التنوع في المفردات بين الموضوعية والمقالية وأسئلة التفاعل والمحاكاة وسهولة التنقل بينها (حميد، ٢٠٢٠؛ Wang, Kao & Chen, 2021)، كما يسهل تطبيقها وتصحيحها بقدر عال من الدقة والمصداقية والموضوعية (البربري، ٢٠٢٠؛ المرزوق، ٢٠٢٠)، وتوفر كثيرًا من الوقت المستهلك في عمليات التصحيح والتغذية الراجعة (Krille, Wuttke & Seifried, 2017; Wang, Kao & Chen, 2021) وتوفر الجهد المبذول من المعلم والمؤسسة التعليمية في تنظيم قاعات الاختبارات وطباعة وترتيب الأوراق الاختبارية وتصحيح المئات من الأوراق وتدقيقها ومراجعتها ورصدها (البربري، ٢٠٢٠؛ Parshall, et al., 2020)، فضلًا عن توفيرها تكاليف مكافآت التصحيح والكنترولات وأسعار الأوراق الاختبارية وتكاليف طباعتها (أبو الشيخ، ٢٠١٨؛ المقبل، ٢٠٢٠)، كما تخزن السجلات وبيانات الطلبة ودرجاتهم على مساحة قليلة من القرص المرن للجهاز أو من خلال التخزين السحابي الذي يضمن حفظ البيانات لفترات طويلة وعدم تعرضها للفقد أو التلف (المومني والنصراوين، ٢٠٢٠؛ Smolinsky, et al., 2020)، ويسهل إجراء

المعالجات الإحصائية لنتائجها ومقارنة النتائج عبر المجموعات المختلفة وفي الأعوام الدراسية المختلفة (عبد الحميد، ٢٠٠٥؛ المقبل، ٢٠٢٠)، وتساعد في تفريد عملية التعليم والتقويم (اليامي، ٢٠٢٠).

وكما أن الاختبارات الإلكترونية تقدم مزايا وفوائد جمة للمعلمين وللنظام التعليمي والأهداف التربوية فإنها تقدم أيضًا عدة مزايا للطلبة أنفسهم، فقد أكد حميد (٢٠٢٠) أن هذه الاختبارت تزيد من دافعية الطلبة للتعلم، كما أوضح المرزوق (٢٠٢٠) أن الطلبة يرون الاختبار الإلكتروني أكثر دقة وموضوعية وأنه لا يتأثر بشخصية المصحح، وأن هذا أدى إلى تكوينهم اتجاهات إيجابية نحوه، ورأى السيد وحسن (٢٠٢١) أن هذه الاختبارات ساهمت في التغلب على جوانب النقد الموجهة للاختبارات التقليدية، مثل عدم ملائمة نوع الاختبار لخصائص الطلبة والقلق المصاحب لها وذاتية التصحيح، كما أشار جليلة (٢٠١٧) أن هذه الاختبارات تزيد من تحفيز الطلبة ومشاركتهم بصورة أفضل وزيادة فهمهم واستيعابهم نظرًا لاحتوائها على وسائط متعددة من صوت وصورة وحركة، وقد أكد ثورلو (Thurlow, et al., 2010) أن هذه الاختبارات تلائم الطلبة ذوي صعوبات التعلم مثلما تلائم الطلبة العاديين.

عيوب الاختبارات الإلكترونية:

وعلى الرغم من كل المزايا المنسوبة للاختبارات الإلكترونية إلا أن هناك بعض التحديات التي تحد من اعتمادها في بعض الأنظمة التعليمية، والتي منها ضرورة امتلاك المعلم والطالب معارف ومهارات التعامل مع الحاسب الآلي أو الجوال الذكي (Berman, Haertel, & Pellegrino, 2020)، وعدم توفر إمكانات لدى بعض الأنظمة لإنشاء نظم الاختبارات الإلكترونية، وصعوبة تصحيح المفردات ذات الإجابات مفتوحة النهاية والإنشائية إلكترونيا (Alruwais, Wills & Wald, 2018)، ووجود مشكلات تتعلق بالاختراق تؤثر في أمن المفردات والإجابات (Kuikkaa, Kitolab & Laaksob, 2014)، ووجود بعض العوائق المادية تحول دون استخدام أنظمة التقويم الإلكتروني (اليامي، ٢٠٢٠)، وضرورة توفر مستوى عال من المراقبة لمنع الطلبة من الغش سواء من الآخرين أو من المصادر المتاحة على شبكة الانترنت (محمود، ٢٠١٨؛ المرزوق، ٢٠٢٠)، وحدوث بعض المشكلات التقنية مثل تعطل الأجهزة أو البرمجيات أو انقطاع شبكة الانترنت أثناء الاختبار

(الوكيل، ٢٠٢١؛ Krille, Wuttke & Seifried, 2017)، وصعوبة قياس بعض المهارات ومستويات التفكير العليا نتيجة اعتماد أغلب الاختبارات المحوسبة على مفردات الصواب والخطأ والاختيار من متعدد والمزاوجة ... الخ (عبد الحميد، ٢٠٠٥)، وعدم ملاءمة الاختبارات الإلكترونية لطبيعة بعض المقررات الدراسية (المقبل، ٢٠٢٠).

ولعل أكثر ما يتعلق بالطلبة من هذه التحديات ضرورة اكتسابهم مهارات التعامل مع التقنية الحديثة (المقبل، ٢٠٢٠)، ووجود نزعة لديهم لمقاومة التغيير (محمود، ٢٠١٨)، والقلق الإلكتروني الذي يصيبهم بسبب التقنية الحديثة، والذي يؤثر في أدائهم على الاختبار (أبو الشيخ، ٢٠١٨؛ الخزي، ٢٠١٣؛ Rana & Mahmood, 2010; Nicholson, 2009; Weiss, 2013).

الدراسات السابقة:

هدفت عدة دراسات إلى معرفة أثر طريقة عرض مفردات المقاييس الإلكترونية في بعض الخصائص السيكومترية للمقياس أو أداء المستجيبين عليه، وقد سعت بعض هذه الدراسات إلى الكشف عن عدد المفردات الأفضل التي يتم عرضها في الشاشة الواحدة أثناء تقديم المقياس الإلكتروني للمستجيبين، فقد قامت توبيويل وداس وسويست (Toepoel, Das & Soest, 2009) بدراسة أثر عدد المفردات المعروضة في الشاشة الواحدة على المستجيبين لاستبيان الإثارة الإلكتروني – المكون من (٤٠) مفردة – على عدد المفردات المفقودة والزمن اللازم للإجابة وتقييم المستجيبين لشكل الاستبيان، وذلك على عينة قوامها (٢٥٦٥) مفحوصًا تم تقسيمهم عشوائيًا لأربعة مجموعات رئيسة قوامهم (٣٥٢، ٧٢٧، ٧٦٨، ٧١٨) مستجيبًا، وتم عرض الاستبانة من خلال أربع صور مختلفة في عدد المفردات المعروضة في الشاشة الواحدة (١، ٤، ١٠، ٤٠) على الترتيب، وقد أسفرت النتائج عن زيادة عدد المفردات المفقودة والزمن اللازم للاستجابة عن الاستبيان بزيادة عدد المفردات المعروضة في الشاشة الواحدة، كما اتضح انخفاض تقييم المستجيبين لشكل الاستبيان مع زيادة عدد المفردات المعروضة في الشاشة الواحدة.

كما قام كاليجارو وشاندلوبز ودينيز (Callegaro, Shand-Lubbers & Dennis, 2009) بالكشف عن أثر طريقة عرض مفردات مقياس الحيوية والصحة العقلية في الاتساق

الداخلي للمفردات، وثبات ألفا كرونباخ وخطأ القياس وعدد المفردات المفقودة، وتكون المقياس من (٩) مفردات بواقع (٤) مفردات في بعد الحيوية، من بينهم مفردتين سالبتين، و(٥) مفردات في بعد الصحة العقلية، من بينهم أيضًا مفردتين سالبتين، فقاموا باختيار عينة عشوائية قوامها (١٤١٩) مشاركًا تزيد أعمارهم عن (١٨) عامًا، تم تقسيمهم عشوائيًا إلى خمس مجموعات قوامها (٢٨٤، ٢٨٧، ٢٨٣، ٢٦٨، ٢٩٧) مشاركًا، وتم عرض مفردات المقياس على المجموعات الخمس بطرق عرض مختلفة هي (طريقة العرض الكلي في شكل شبكة معيارية standard grid وطريقة العرض الكلي في صورة شبكة مظللة shaded grid وطريقة عرض مفردة واحدة في الشاشة والبدائل أفقية، وطريقة عرض مفردة واحدة في الشاشة والبدائل رأسية وطريقة عرض مفردة واحدة في الشاشة والبدائل رأسية مظللة)، وقد توصلوا إلى عدم وجود مفردات مفقودة في المجموعات نتيجة طرق العرض المختلفة، وعدم وجود فرق دال إحصائيًا في ثبات ألفا كرونباخ في بعدي الحيوية والصحة العقلية يعزى إلى اختلاف طريقة عرض المفردات، كما قاموا بحساب معامل ارتباط سبيرمان بين الدرجات على المفردات الموجبة والمفردات السالبة لحساب الاتساق الداخلي، ثم تحويل فيشر لقيمة معامل الارتباط إلى قيمة Z- Fisher، وقد تبين وجود فرق دال إحصائيًا في تحويل فيشر بين المجموعتين الأولى والثانية، ولم يظهر فرق دال إحصائيًا بين الخطأ المعياري للقياس بين المجموعات الخمس.

وقد هدف كلاوش ورفاقه (Klausch, et al., 2012) إلى معرفة أثر طريقة عرض مفردات المقياس (المصفوفة التقليدية ومصفوفة التمرير الأفقي) وعدد مفرداته في جودة البيانات المستمدة من تطبيق مفردات أحد المقاييس النفسية، فقام بتجربة عشوائية من خلال تصميم عاملي (٢×٣) يحتوي على متغيرين مستقلين أولهما طريقة عرض المفردات، حيث يتم عرض المفردات إما بالطريقة الكلية (المصفوفة التقليدية) أو بطريقة المفردة الواحدة من خلال مصفوفة التمرير الأفقي Horizontal Scrolling Matrix (HSM)، وثانيهما عدد مفردات المقياس، حيث يتكون المقياس من (٥) أو (١٠) أو (١٥) مفردة، وقد أشارت النتائج إلى أن طريقة عرض المفردة الواحدة من خلال التمرير الأفقي حسنت جودة البيانات في حالة المقاييس الأطول، وكان اتجاه المستجيبين نحو هذه الطريقة ايجابيًا مرتفعًا، وقد أوصت الدراسة باعتماد هذه الطريقة لمواجهة انخفاض دافعية الاستجابة عن المقاييس الطويلة.

كما سعى ريفيلا وتونينيلي واوشوا (Revilla, Toninelli & Ochoa, 2016) إلى دراسة أثر طريقة عرض مفردات الاستبانة (الكلي – مفردة مفردة) والجهاز المستخدم في الاستجابة على بعض الخصائص السيكومترية، حيث قاموا بتطبيق مقياس مرتين في شهري فبراير ومارس من عام (٢٠١٥) بأسبانيا، على عينة قوامها (١٦٠٨) مشاركًا في التطبيقين، وقد تم تقسيمهم عشوائيًا إلى ثلاث مجموعات هم: (٥٥٤) مستجيبًا عبر الكمبيوتر مع عرض كلي للمفردات، وعدد (٥١٨) مستجيبًا عبر الجوال الذكي مع عرض كلي للمفردات، وعدد (٥٣٦) مستجيبًا عبر الجوال الذكي مع عرض المفردات واحدة تلو الأخرى، وقد أسفرت النتائج عن مستويات متماثلة من الارتباطات البينية بين المفردات في استجابات المجموعات الثلاث، وكان الزمن اللازم للاستجابة أطول في حالة الاستجابة عبر الجوال الذكي مع العرض الكلي للمفردات، ولوحظ انخفاض في عدم التمايز بين المستجيبين عبر الكمبيوتر مع عرض المفردات كليًا، وانخفاض في عدم التمايز بين المستجيبين عبر الجوال الذكي مع عرض المفردات واحدة تلو الأخرى، وقد اتضح وجود تأثير لطريقة عرض المفردات ونوع الجهاز في الاستجابة عن مفردات الاستبانة.

وقد قارنت دراسة محمود (٢٠١٨) بين الخصاص السيكومترية للصورة الإلكترونية لاختبار تحصيلي معدة باستخدام برنامج البلاك بورد Blackboard وتم تقديمها حضوريا من خلال عرض مفردات الاختبار في شاشة واحدة، وصورة أخرى إلكترونية تم تقديمها عن بعد وكان عرض المفردات فيها تواليًا مع حظر الرجوع، وصورة ثالثة ورقية، وذلك من خلال تطبيق الصور الثلاثة على ثلاث عينات متكافئة بواقع صورة لكل عينة، وكان قوام العينات الثلاثة (٢٣، ٢٩، ٣١) طالبة بكلية التربية بجامعة الإمام عبد الرحمن بن فيصل، وكان عدد مفردات الاختبار (٢٠) مفردة من نوع الاختيار من متعدد، وقد توصلت إلى وجود فروق دالة إحصائيًا بين معاملات صعوبة مفردات الصور الثلاثة لصالح الصورة الإلكترونية المقدمة عن بعد، وعدم وجود فروق دالة إحصائيًا بين معاملات تمييزها، وكان أفضل معامل صدق تلازمي وأفضل معامل ثبات للصورة الإلكترونية المقدمة حضوريًا، وكان أقل معامل صدق تلازمي وأفضل معامل ثبات للصورة الإلكترونية المقدمة عن بعد.

وقد ربط ليو وسيتانت (Liu & Cetant, 2018) بين طريقة عرض مفردات االمقاييس الإلكترونية وعدد البدائل التي تحتويها المفردة الواحدة، فقاما بالمقارنة بين الطريقتين الكلية

والتتابعية عندما تم تطبيق مقاييس متعددة مختلفة في عدد خيارات الاستجابة بين (٢، ٣، ٤، ٥، ٧، ٩، ١١) خيارًا على (٥٦٤٤) مشاركًا، بينهم (٧٤ %) من الإناث، و (٦٨%) أعمارهم تزيد عن (٤٥) عامًا، وقد أسفرت النتائج عن أن خاصية عدم تمايز الإستجابة عن المفردات وكذلك الوقت اللازم للاستجابة لا يختلفان باختلاف طريقة عرض المفردات، أو باختلاف عدد خيارات الاستجابة، ولكن يزداد عدد المفردات المفقودة في حالة عرض المفردات بالطريقة الكلية، كما أظهرت النتائج تكافؤ القياس بين طريقتي العرض عندما يكون عدد الخيارات أقل من (٧)، بينما توجد اختلافات جوهرية بين طريقتي العرض الكلي والتتابعي للمفردات عندما تتكون خيارات الاستجابة من (٩، ١١) خيارًا.

كما هدفت دراسة مافليتوفا وكوبر وليبيديف (Mavletova, Couper & Lebedev, 2018) إلى الكشف عن تأثير كل من طريقة عرض المفردات (الكلي – التتابعي) ونوع الجهاز المستخدم في الاستجابة (الجوال الذكي – الحاسب الآلي) وعدد مفردات المقياس في بعض الخصائص السيكومترية، مثل خطأ القياس وصدق وثبات المقياس وتكافؤ القياس، وفي سبيل ذلك تم اختيار عينة قوامها (١٦٧٨) مشاركًا في التطبيق الأول، وعدد (١٠٧٩) مشاركًا في التطبيق الثاني، وقد أسفرت النتائج عن أن الصدق التلازمي في حالة العرض الكلي للمفردات كان أقل منه في حالة العرض التتابعي سواء عند عرض المفردات من خلال الجوال الذكي أو الحاسب الآلي، كما وجد خطأ قياس أعلى عند عرض المفردات بطريقة كلية بغض النظر عن الجهاز المستخدم، ولم يختلف معامل الثبات بطريقة إعادة التطبيق باختلاف طريقة عرض المفردات، كما لم يكن هناك اختلاف بين معدلات إنقطاع الاستجابة Breakoff rates في الحالتين، بينما أظهرت المقارنات متعددة المجموعات تحقق تكافؤ القياس بين طريقتي عرض المفردات، وقد ظهر تباينًا في تكافؤ القياس تعزى لطريقة عرض المفردات في حالة المقياس الأطول.

وقد سعى حميد (٢٠٢٠) إلى دراسة تأثير نمط عرض مفردات الاختبار الإلكتروني (مفردة واحدة في الشاشة – مفردات الاختبار كلها في الشاشة) في التحصيل والأداء المهاري والاتجاه نحو الاختبار الإلكتروني، وذلك لدى (١٦) طالبًا بمرحلة الدكتوراه بقسم تكنولوجيا التعليم في جامعة القصيم، وقد توصل إلى عدم وجود فرق دال إحصائيًا يعزى إلى اختلاف

طريقة عرض مفردات الاختبار بين متوسطي رتب درجات الطلاب في التحصيل أو الأداء المهاري أو الاتجاه نحو الاختبارات الإلكترونية، وقد أوصى بعرض مفردة واحدة في الشاشة عند وجود الطلاب معًا في نفس القاعة مع إتاحة خيار التراجع عن إجابة المفردة وتغيير الإجابة.

كما هدف السيد وحسن (٢٠٢١) إلى معرفة أثر نمط عرض مفردات الاختبار الإلكتروني (كلي – تتابعي) في الاحتفاظ بالتعلم ودافعية الإنجاز والضغوط النفسية لدى طلاب تكنولوجيا التعليم، معتمدان المنهج شبه التجريبي، وذلك لعينة قوامها (٤٠) طالبًا في قسم تكنولوجيا التعليم بجامعة عين شمس، وقد توصلا إلى تفوق الأفراد الذين تم تقديم المفردات لهم بالطريقة الكلية في الاحتفاظ بالتعلم ودافعية الإنجاز وفي القدرة على خفض الضغوط النفسية.

وقد قام عبد الوكيل (٢٠٢١) بدراسة استطلاعية قبل دراسته الأساسية؛ لمعرفة آراء الطلبة واتجاهاتهم نحو طريقة عرض مفردات الاختبارات الإلكترونية والصعوبات التي يواجهونها عند إجابتها، وذلك على (٣٠) طالبًا وطالبة من طلاب كلية المجتمع بالمستويين الثالث والرابع بقسم علوم الحاسب، وقد تبين له أن (٢٥) منهم يفضلون العرض التتابعي للمفردات وغير راضين عن طريقة العرض الكلي لمفردات الاختبار الإلكتروني، وغير راضين أيضًا عن حظر الرجوع في إجابة المفردات التي تمت الإجابة عنها، ومن ثم قام بدراسته الأساسية التي هدف منها إلى معرفة أثر التفاعل بين نمط التقديم التتابعي لمفردات الاختبارات الإلكترونية (حظر / سماح) الرجوع وعرض النتائج للطلاب (فوري / مرجأ) في التحصيل المعرفي وقلق الاختبار، وذلك على عينة قوامها (٦٠) طالبًا وطالبة بقسم علم الحاسب بكلية المجتمع في جامعة حائل مقسمين على أربع مجموعات تجريبية، من خلال المنهج شبه التجريبي، وأسفرت النتائج عن تفوق مجموعة (نمط نمط عرض المفردة السماح بالرجوع مع العرض الفوري للنتيجة) في التحصيل الدراسي، وانخفاض قلق الاختبار الإلكتروني لديهم عن باقي المجموعات، وبالتالي أوصت الدراسة بضرورة توفير خاصية السماح بالرجوع للمفردات مع العرض الفوري للنتيجة.

وقد هدف ديبيل ورفاقه (Debell, et al., 2021) إلى معرفة أثر طريقة عرض مفردات استبيان تقييم دراسات الحكومة والمجتمع الإلكتروني (الطريقة الكلية Grid – الطريقة التفريعية Branching – طريقة المفردة الواحدة Single item)، على الوقت اللازم للاستجابة وعدد المفردات المفقودة ومعدلات الانقطاع أثناء الاستجابة وصدق المحك للاستبيان، وذلك على عينة قوامها (١٢٠٩) مشاركًا في الولايات المتحدة الأمريكية، وقد أسفرت النتائج عن انخفاض الوقت اللازم للاستجابة وفق الطريقة الكلية عنه وفق طريقة المفردة الواحدة وجاء أطول وقت للاستجابة وفق الطريقة التفريعية، كما تبين أن عدد المفردات المفقودة أقل ما يمكن وفق طريقة المفردة الواحدة، ويزداد وفق الطريقة الكلية، ويكون أكبر عدد لها في حالة الطريقة التفريعية، ولم يظهر فرق دال إحصائيًا بين الطرق الثلاثة في معدلات انقطاع الاستجابة، كما كان صدق المحك للاستبيان أضعف في الطريقة التفريعية، ولم يكن هناك فرق دال إحصائيًا بين الطريقة الكلية وطريقة المفردة الواحدة، وتشير النتائج إلى عدم ملاءمة الطريقة التفريعية في جمع البيانات إلكترونيًا، وأفضلية طريقة المفردة الواحدة والطريقة الكلية (المعدة جيدًا) في جمع البيانات.

كما سعى وانج وكاو وشين (Wang, Kao & Chen, 2021) إلى معرفة تأثير طريقة عرض مفردات الاختبار الإلكتروني (كلي – تتابعي) وطريقة ترتيبها (من الصعب للسهل – من السهل للصعب) على أداء المختبرين، من خلال عينة قوامها (٣٨١) تلميذًا في الصف الخامس الابتدائي بتايوان، حيث تم إعداد اختبار تحصيلي مكون من (١٦) مفردة من نوع الاختيار من متعدد، وقد تم اعتماد التصميم التجريبي المختلط، حيث تم تقسيم العينة إلى مجموعتين مستقلتين تعرضت إحداهما للاختبار الإلكتروني من خلال العرض الكلي للمفردات والأخرى تعرضت للاختبار الإلكتروني من خلال العرض التتابعي للمفردات، وكل مجموعة منهما تعرضت للاختبار بطريقتين مختلفتين في ترتيب المفردات، وقد أسفرت النتائج عن عدم وجود فروق بين أداء المختبرين يعزى إلى طريقة عرض المفردات، بينما كشفت النتائج عن وجود تأثير لطريقة ترتيب المفردات على أداء الممتحنين، وقد أوضح الباحثين تفضيل الممتحنين طريقة العرض الكلي لمفردات الاختبار الإلكتروني المعروضة مفرداته بطريقة كلية، وأوصوا بذلك.

التعليق على الدراسات السابقة:

إن المتمعن لهذه الدراسات يتبين له ما يلي:

- إن أداة القياس الإلكترونية في أغلب الدراسات التي تمت الإشارة إليها عبارة عن استبانة أو استفتاء، ولكن الدراسات التي كانت فيها أداة القياس الإلكترونية عبارة عن اختبار تحصيلي هي دراسات (حميد، ٢٠٢٠٠؛ السيد وحسن، ٢٠٢١؛ عبد الوكيل، ٢٠٢١؛ محمود، ٢٠١٨؛ Wang, Kao & Chen, 2021).
- اهتمت جميع الدراسات السابقة بالمقارنة بين طريقة العرض الكلي وطريقة العرض التتابعي دون حظر الرجوع، ولم تهتم أي من الدراسات بطريق العرض التتابعي مع حظر الرجوع سوى دراسة محمود (٢٠١٨)، والتي لم يكن هدفها المقارنة بشكل مباشر بين طريقتي عرض مفردات الصورتين الإلكترونيين للاختبار (الكلية / التتابعية)، ولكنها سعت للمقارنة بين الصورتين الاختباريتين وفقًا لاختلاف بيئة تقديمهما حيث كانت إحداهما حضورية والأخرى عن بعد، ومن ثم لم يتضح بشكل مباشر تأثير طريقة عرض المفردات نفسها، وكذلك دراسة عبد الوكيل (٢٠٢١) التي تناولت المقارنة بين طريقتي العرض التتابعي للمفردات (مع وبدون) حظر الرجوع، دون أن تتعرض للطريقة الكلية.
- اختلفت نتائج الدراسات السابقة فيما يتعلق بأفضلية العرض الكلي أو التتابعي للمفردات، وليس هناك اتفاق من قبل الدراسات على أفضلية إحدى الطريقتين عن الأخرى.
- تبين وجود تأثير لحجم شاشة الجهاز المستخدم في الاستجابة أثناء مقارنة الطريقة الكلية بالطريقة التتابعية على الزمن اللازم للاستجابة كما في دراسة (Revilla, Toninelli & Ochoa, 2016) وعلى الصدق التلازمي للمقياس كما في دراسة (Mavletova, Couper & Lebedev, 2018).
- تعرضت مجموعات عينة البحث في أغلب الدراسات لمرحلتين من التطبيق بفاصل زمني، دون مراعاة عامل تذكر المفردات وألفة القياس الإلكتروني على نتائج التطبيق في المرة الثانية.

لذا سعى البحث الحالي إلى المقارنة بين ثلاثة طرق لعرض المفردات، هي: الطريقة الكلية والتتابعية مع حظر الرجوع والتتابعية بدون حظر الرجوع، مع تثبيت نوعية الأجهزة المستخدمة في الإجابة وحجم شاشتها من خلال اختبار الطالبات في معمل مجهز أجهزة حاسب آلي بالجامعة، ودون تعريض الطالبات لمرتين من التطبيق.

فروض البحث:

١- لا توجد فروق دالة إحصائيًا للفروق في معاملات ارتباط المفردات المصححة بالدرجة الكلية على الاختبار التي تعزى إلى:

١- أ) طريقة عرضها مع تثبيت عددها.

١- ب) عددها مع تثبيت طريقة عرضها.

١- ج) اختلاف طريقة عرضها وعددها.

٢- لا يختلف ترتيب معاملات ارتباط درجات الطالبات على الاختبار الإلكتروني بالمعدل التراكمي باختلاف طريقة عرضها وعددها.

٣- لا توجد فروق دالة إحصائيًا للفروق في متوسطات معاملات صعوبة مفردات الاختبار الإلكتروني التي تعزى إلى:

٣- أ)طريقة عرضها مع تثبيت عددها.

٣- ب) عددها مع تثبيت طريقة عرضها.

٣- ج) اختلاف طريقة عرضها وعددها.

٤- لا توجد فروق دالة إحصائيًا للفروق في متوسطات معاملات تمييز مفردات الاختبار الإلكتروني التي تعزى إلى:

٤- أ) طريقة عرضها مع تثبيت عددها.

٤- ب) عددها مع تثبيت طريقة عرضها.

٤- ج) اختلاف طريقة عرضها وعددها.

٥- لا يختلف ترتيب معاملات ثبات الاختبار الإلكتروني للصور الاختبارية المختلفة وفقًا لطريقة عرض المفردات وعددها.

٦- لا توجد فروق دالة إحصائيًا للفروق في متوسطات درجات الطلبة على الاختبار الإلكتروني التي تعزى إلى:

٦- أ) طريقة عرضها مع تثبيت عددها.

٦- ب) عددها مع تثبيت طريقة عرضها.

٦- ج) اختلاف طريقة عرضها وعددها.

منهج البحث:

تم اعتماد المنهج شبه التجريبي للإجابة عن تساؤلات البحث، حيث أحدث الباحث تغييرًا متعمدًا في المتغيرين المستقلين (طرق عرض مفردات الاختبار الإلكتروني وعددها)؛ لمعرفة مدى تأثير ذلك على المتغيرين التابعين (الخصائص السيكومترية للاختبار الإلكتروني وأداء الطالبات عليه)، مع ضبط تأثير أستاذ المقرر والأهداف التعليمية والمحتوى الدراسي وطريقة التدريس والمعدل التراكمي للطالبات وحجم شاشة الجهاز المستخدم في الإجابة.

عينة البحث:

تمثلت عينة البحث في (٤٥٠) من طالبات كلية العلوم الاقتصادية والمالية الإسلامية بجامعة أم القرى يمثلن الطالبات اللاتي يدرس لهن الباحث مقرر الإحصاء الاستدلالي للاقتصاديين بعد استبعاد الطالبات اللاتي تم التحقق من الخصائص السيكومترية للاختبار عليهن، وقد روعي بذلك تثبيت أثر بعض العوامل في التحصيل الدراسي للطالبات، مثل: أستاذ المادة والأهداف التعليمية للمقرر والمحتوى الدراسي وطريقة التدريس والمصادر التعليمية وتخصص الطالبات ... الخ. وقد قسمت العينة إلى ستة مجموعات متكافئة في المعدل التراكمي، قوام كل منهم (٧٥) طالبة.

أداة البحث:

تمثلت أداة البحث في اختبار تحصيلي في مقرر الإحصاء الاستدلالي للإقتصاديين، والذي تم بناؤه وفق الخطوات الآتية:

- تحديد الهدف من الاختبار: هدف الاختبار إلى قياس تحصيل الطالبات في موضوعات المقرر من خلال أربعة مستويات معرفية وفقًا لتصنيف بلوم، وهي التذكر والفهم والتطبيق والتحليل.

- تحديد محتوى الاختبار: تم قياس التحصيل الدراسي في موضوعات: نظرية الاحتمالات، التوزيعات الاحتمالية للمتغيرات العشوائية، توزيعات المعاينة ونظرية التقدير، اختبار الفرضيات الإحصائية.
- تحديد مخرجات التعلم المراد قياسها: هدف الاختبار إلى قياس (٢٠) مخرج تعليمي، وذلك في مستويات التذكر والفهم والتطبيق والتحليل بواقع خمسة مخرجات في كل مستوى معرفي.
- تحديد الأوزان النسبية للموضوعات المراد قياسها: في ضوء عدد الساعات التدريسية لكل موضوع من الموضوعات، تم تحديد الأوزان النسبية (٢٠%، ٣٠%، ٢٠%، ٣٠%) على الترتيب.
- تحديد الأوزان النسبية للمستويات المعرفية لمخرجات التعلم المراد قياسها: من خلال قسمة عدد المخرجات في كل مستوى على عدد المخرجات الكلي، وبذلك يكون وزن كل مستوى مساويًا (٢٥%).
- تحديد طول الاختبار: تم تحديد طولين مختلفين للاختبار التحصيلي لتحقيق الهدف من البحث، وكان طول الصورة الأولى (٣٠) مفردة، وطول الصورة الثانية (٤٠) مفردة.
- إعداد جدول مواصفات الاختبار: من أجل تحديد عدد المفردات التي تم من خلالها قياس كل مستوى من مستويات مخرجات التعلم في كل موضوعات المقرر وفقًا للأوزان السابقة، كما في جدول (١).

جدول (١)

عدد مفردات الاختبار موزعة في ضوء جدول مواصفات الاختبار

الموضوع	الصورة الأولى (٣٠) مفردة					الصورة الثانية (٤٠) مفردة				
	التذكر	الفهم	التطبيق	التحليل	المجموع	التذكر	الفهم	التطبيق	التحليل	المجموع
نظرية الاحتمالات	١	٢	٢	١	٦	٢	٢	٢	٢	٨
التوزيعات الاحتمالية	٢	٢	٣	٢	٩	٤	٤	٢	٢	١٢
توزيعات المعاينة	٢	٢	١	١	٦	٢	٢	٢	٢	٨
اختبار الفرضيات	٢	٢	٢	٣	٩	٢	٢	٤	٤	١٢
المجموع	٧	٨	٨	٧	٣٠	١٠	١٠	١٠	١٠	٤٠

- صياغة مفردات الاختبار: تمت صياغة (٤٠) مفردة على هيئة صواب وخطأ، ووضعها في الصورة الاختبارية الثانية، وتم اختيار (٣٠) مفردة منهم ووضعها في الصورة الأولى، مع الإلتزام بجدول المواصفات.
- تم عرض مفردات الاختبار ومخرجات التعلم المراد قياسها على (٨) من المحكمين من أعضاء هيئة التدريس بجامعة أم القرى من المتخصصين في تدريس مقررات الإحصاء؛ بهدف معرفة آرائهم في قياس المفردات لمخرجات التعلم ودقة الصياغة العلمية واللغوية لها، وقد ترتب على هذه الخطوة موافقة جميع المحكمين على جميع المفردات مع تعديل صياغة ثلاثة مفردات منها.
- تم تطبيق الصورتين الاختباريتين على عينتين استطلاعيتين قوام كل منهما (٥٠) طالبة من طالبات كلية العلوم الإقتصادية والمالية الإسلامية تم اختيارهن عشوائيًا من بين الطالبات اللاتي يدرس الباحث لهن المقرر من غير طالبات العينة الأساسية، بهدف التحقق من فهمهن للمفردات، والتحقق من الخصائص السيكومترية للصورتين الاختباريتين وحساب متوسط الزمن اللازم للإجابة عن كل صورة منهما.
- تم التحقق من صدق الصورتين الاختباريتين من خلال حساب معامل الارتباط بين درجة الطالبات على الصورة الاختبارية والمعدل التراكمي لهن، وكانت قيمة معاملات الارتباط في الحالتين (٠.٨٧، ٠.٨٥) على الترتيب، مما يشير إلى تمتع الصورتين الاختباريتين بصدق المحك، كما تم حساب معاملات ارتباط المفردات المصححة بالدرجة الكلية في الصورتين، وقد تراوحت بين (٠.٣٣ – ٠.٦٧) ، (٠.٣٧ – ٠.٥٨) في الصورتين على الترتيب، وكانت جميع معاملات الارتباط دالة إحصائيًا عند مستوى دلالة (٠.٠١).
- تم حساب معاملات صعوبة المفردات، وقد تراوحت بين (٠.٢٨ : ٠.٧١) في الصورة الاختبارية الأولى، وبين (٠.٣١ : ٠.٧٢) في الصورة الاختبارية الثانية، وكلها في المدى المقبول.
- كما تم حساب معاملات تمييز المفردات، وقد تراوحت بين (٠.٤٤ : ٠.٧٤)، في الصورة الاختبارية الأولى، وبين (٠.٤١ : ٠.٨٦) في الصورة الاختبارية الثانية، وهي معاملات تمييز جيدة.

- وتم حساب معامل ثبات الصورتين الاختباريتين باستخدام معادلة كيودر ريتشاردسون (٢١)، وقد بلغا (٠.٩٢ ، ٠.٨٩) على الترتيب.
- تم حساب متوسط الزمن المستغرق لإجابة الطالبات عن كل صورة من الصورتين الاختباريتين، وكان (٣٧، ٤٩) دقيقة تقريباً على الترتيب، وقد روعي في تحديد هذا الزمن البطء المحتمل حدوثه بسبب سرعة الانترنت وكذلك اختلاف خبرات الطالبات باستخدام التقنية الحديثة.
- تمت صياغة تعليمات الصورتين الاختباريتين لتوضيح كيفية تدوين الإجابة، ومنع استخدام الآلة الحاسبة والسماح بالإستعانة بجدول التوزيع الطبيعي القياسي والزمن المخصص للإجابة عن كل صورة اختبارية.

إجراءات البحث:

- تم تقسيم عينة البحث إلى ست مجموعات مستقلة (أ، ب، ج، د، هـ، و) وذلك من خلال الاختيار العشوائي الطبقي، حيث تم تقسيم الطالبات في ضوء معدلاتهم التراكمية إلى ثمانية مستويات (A^+, A, B^+, B, C^+, C, D^+, D)، وتم توزيع الطالبات في كل مستوى على المجموعات الست بطريقة عشوائية، بهدف تحقيق التكافؤ بين الطالبات في معدلاتهم التراكمية، وقد تم التحقق من تكافؤ المجموعات الست في المعدل التراكمي من خلال المقارنة بين متوسطات معدلاتهم الأكاديمية باستخدام تحليل التباين الأحادي One Way Anova، بعد التحقق من تجانس المجموعات الست واعتدالية التوزيع فيها، وكانت نتائج هذه المقارنة كما في الجدول (٢).

جدول (٢)

نتائج تحليل التباين الأحادي للمقارنة بين المجموعات المستقلة الست في معدلاتهم التراكمية

مصدر التباين	مجموع المربعات	درجة الحرية	متوسط المربعات	قيمة ف	مستوى الدلالة	الدلالة
المعدل التراكمي	٠.٠٤٨	٥	٠.٠١٠	٠.٠٨٩	٠.٩٩٤	غير دال
الخطأ	٤٧.٩٧٨	٤٤٤	٠.١٠٨			
الكلي	٤٨.٠٢٦	٤٤٩				

ويتضح من خلال جدول (٢) عدم وجود فرق دال إحصائيًا بين المعدلات التراكمية للطالبات في المجموعات الستة، مما يدل على تكافؤ الطالبات في معدلاتهم التراكمية.

- تم تقديم الصورة الاختبارية الأولى المكونة من (٣٠) مفردة على الطالبات في المجموعات الثلاثة الأولى (أ، ب، ج)، من خلال طرق عرض مختلفة، هي: الطريقة الكلية والطريقة التتابعية مع حظر الرجوع، والطريقة التتابعية بدون حظر الرجوع على الترتيب، وفي نفس الوقت تم تقديم الصورة الاختبارية الثانية المكونة من (٤٠) مفردة لطالبات المجموعات الثلاثة الأخيرة (د، هـ، و) من خلال نفس طرق العرض السابقة على الترتيب، وذلك من خلال الخيارات المتاحة في منصة البلاك بورد Blackboard، وقد اعتمد الباحث على هذا التصميم حتى تتعرض كل طالبة للاختبار مرة واحدة، تفاديًا لعامل تذكر الطالبات للمفردات والألفة بالقياس الإلكتروني.
- تم استخدام أجهزة حاسب آلي لها نفس حجم الشاشة لتقديم الاختبار لطالبات المجموعات الست؛ بهدف توحيد حجم شاشة الجهاز المستخدم من خلال عقد الاختبار في معمل به عدد كاف من أجهزة الحاسب الآلي.
- تم إجراء المعالجة الإحصائية للبيانات باستخدام برنامج (SPSS V. 25)، واستخراج نتائج البحث.

نتائج البحث وتفسيرها:

التحقق من قبول الفرض الأول:

ينص الفرض الأول على: لا توجد فروق دالة إحصائيًا في معاملات ارتباط المفردات المصححة بالدرجة الكلية على الاختبار التي تعزى إلى:

١- أ)طريقة عرضها مع تثبيت عددها

١- ب)عددها مع تثبيت طريقة عرضها.

١- ج) اختلاف طريقة عرضها وعددها

وللتحقق من هذا الفرض تم حساب معاملات ارتباط المفردات المصححة بالدرجة الكلية في الصور الاختبارية في المجموعات الستة (أ، ب، ج، د، ه، و)، وكانت الإحصاءات الوصفية لمعاملات الارتباط المصححة كما في جدول (٣)

جدول (٣)

الإحصاءات الوصفية لمعاملات الارتباط المصححة بين المفردات والدرجة الكلية في الصور الاختبارية

طريقة العرض	الكلية		التتابعية مع حظر الرجوع		التتابعية بدون حظر الرجوع	
عدد المفردات	٣٠ مفردة	٤٠ مفردة	٣٠ مفردة	٤٠ مفردة	٣٠ مفردة	٤٠ مفردة
المتوسط	٠.٦١٣	٠.٦٤٨	٠.٥٣٤	٠.٥٦٦	٠.٦٤٥	٠.٦٨١
الانحراف المعياري	٠.٠٩٥	٠.٠٨٦	٠.١١١	٠.٠٩٩	٠.٠٩١	٠.٠٨٣

ونظرًا لأن قيم معاملات الارتباط لا تخضع للتوزيع الطبيعي، فإنه تم تحويل قيم معاملات الارتباط إلى ما يقابلهما من قيم "Z" لفيشر (Z–Fisher)، وبالتالي يكون توزيع Z طبيعيا تقريبًا ثم تم ما يلي:

(أ) بالنسبة لتأثير طريقة عرض المفردات في معاملات ارتباطها المصححة بالدرجة الكلية مع تثبيت عددها:

تمت المقارنة بين متوسطات قيم Z في حالة اختلاف طريقة عرض مفردات الصورة الاختبارية الأولى المكونة من (٣٠) مفردة باستخدام تحليل التباين الأحادي، كما تمت هذه المقارنة أيضًا في حالة اختلاف طريقة عرض مفردات الصورة الاختبارية الثانية المكونة من (٤٠) مفردة باستخدام تحليل التباين الأحادي، وجاءت النتائج كما في جدول (٤).

جدول (٤)

نتائج تحليل التباين الأحادي للمقارنة بين متوسطات قيم "Z" لفيشر باختلاف طرق العرض

الاختبار	مصدر التباين	مجموع المربعات	درجة الحرية	متوسط المربعات	قيمة ف	مستوى الدلالة	مربع ايتا	حجم الأثر
الاختبار المكون من (٣٠) مفردة	طريقة العرض	٠.١٩٧	٢	٠.٠٩٨	٩.٩٦٦	٠.٠٠٠	٠.١٨٧	كبير
	الخطأ	٠.٨٥٩	٨٧	٠.٠١٠				
	الكلي	١.٠٥٦	٨٩					
الاختبار المكون من (٤٠) مفردة	طريقة العرض	٠.٢٧٩	٢	٠.١٤٠	١٧.٢٢٤	٠.٠٠٠	٠.٢٢٧	كبير
	الخطأ	٠.٩٤٨	١١٧	٠.٠٠٨				
	الكلي	١.٢٢٧	١١٩					

يتضح من جدول (٤) وجود فرق دال إحصائيًا بين متوسطات معاملات ارتباطات المفردات المصححة بالدرجة الكلية للاختبار (بعد تحويلها لقيم "Z" لفيشر)، وهذا الفرق يعزى إلى اختلاف طريقة عرض المفردات في حال تثبيت عدد مفرداته، وذلك عندما كان عدد مفردات الاختبار (٣٠) مفردة وكذلك عندما كان عددها (٤٠) مفردة، ولمعرفة اتجاه هذا الفرق في الحالتين تم إجراء مقارنات بعدية Post Hoc باستخدام اختبار شيفيه Scheffe للمقارنة بين كل مجموعتين تعرضتا لنفس عدد المفردات، وكانت النتائج كما في جدول (٥).

جدول (٥)

نتائج اختبار شيفيه بين متوسطي قيم "Z" لفيشر لكل طريقتين من طرق العرض في كل صورة اختبارية

عدد المفردات		٣٠ مفردة		٤٠ مفردة	
طريقة العرض		الكلية	التتابعية مع حظر الرجوع	الكلية	التتابعية مع حظر الرجوع
التتابعية مع حظر الرجوع	الفرق بين المتوسطين	٠.٠٧٩		٠.٠٨٢	
	مستوى الدلالة	*٠.٠١١		*٠.٠٠٠	
التتابعية بدون حظر الرجوع	الفرق بين المتوسطين	-٠.٠٣٢	-٠.١١١	-٠.٠٣٣	-٠.١١٥
	مستوى الدلالة	٠.٤٥٥	*٠.٠٠٠	٠.٢٦٥	*٠.٠٠٠

ويتضح من خلال جدول (٥) وجود فرق دال إحصائيًا بين متوسطي معاملات ارتباط المفردات بالدرجة الكلية (بعد تحويلها إلى قيم "Z" لفيشر) باختلاف طريقة عرضها بين الطريقة الكلية والطريقة التتابعية مع حظر الرجوع، في اتجاه الطريقة الكلية سواء عندما كان عدد مفردات الاختبار (٣٠) أو (٤٠) مفردة، كما يتضح وجود هذا الفرق الدال إحصائيًا عند المقارنة بين الطريقة التتابعية بدون حظر الرجوع والطريقة التتابعية مع حظر الرجوع، في اتجاه الطريقة التتابعية بدون حظر الرجوع سواء عندما كان عدد مفردات الاختبار (٣٠) أو

(٤٠) مفردة، ولم يكن هناك فرق دال إحصائيًا عند المقارنة بين الطريقتين الكلية والتتابعية بدون حظر الرجوع، سواء عندما كان عدد المفردات (٣٠) أو (٤٠) مفردة. وهذا يعني أن ارتباط المفردات بالدرجة الكلية عند عرض المفردات بالطريقتين الكلية والتتابعية بدون حظر الرجوع أعلى منه في حال عرض المفردات بالطريقة التتابعية مع حظر الرجوع، أي أن اتساق إجابات الطالبات وصدق الاختبار عند عرض المفردات بالطريقتين الكلية والتتابعية بدون حظر الرجوع أفضل منه عند عرضها بالطريقة التتابعية مع حظر الرجوع.

وقد ترجع هذه النتائج إلى أنه عند حظر رجوع الطالبة للإجابة عن المفردات التي لا تثق في صحة إجابتها أو التي تركتها بدون إجابة رغم ما قد تمتلكه من معرفة جزئية أو كلية بالإجابة، قد يؤثر سلبًا بشكل متفاوت في أداء الطالبات على المفردات، وقد يؤدي إلى ممارسة بعض الطالبات التخمين العشوائي للإجابة عن المفردات لأنهن لن يجدن فرصة أخرى للإجابة، ومن ثم فإن ذلك يضيف مصدرًا جديدًا من مصادر الخطأ العشوائي في القياس، وهذا يؤثر سلبًا في ارتباط الدرجة عن المفردات بالدرجة الكلية للطالبات، ولا يظهر هذا الأثر في حالة عرض المفردات بالطريقتين الكلية والتتابعية دون حظر الرجوع؛ لأن هاتين الطريقتين تسمحان للطالبات بمراجعة إجاباتهن عن المفردات التي يؤجلن الإجابة عنها لمزيد من التركيز أو لتذكر الإجابة طوال وقت الاختبار. كما أن حظر رجوع الطالبات للإجابة عن المفردات قد يؤدي إلى تفوق الطالبات القادرات على إدارة وقت الاختبار بكفاءة واللاتي لديهن مهارات الحكمة الاختبارية وضبط الانفعالات وإدارة القلق الإلكتروني، وبالتالي تزداد القدرات والمهارات المطلوبة للإجابة عن المفردات بجانب التحصيل الدراسي، وهذا يقلل من صدق الاختبار في قياس ما يسعى لقياسه.

وتتفق هذه النتيجة جزئيا مع نتيجة دراسة كاليجارو وشاندلوبز ودينيز (Callegaro, Shand-Lubbers & Dennis, 2009) التي توصلت إلى وجود فرق دال إحصائيًا بين تحويلات فيشر لمعاملات الارتباط باختلاف طريقة عرض المفردات، كما تختلف مع نتيجة دراسة ريفيلا وتونينيلي واوشوا (Revilla, Toninelli & Ochoa, 2016) التي توصلت إلى وجود مستويات متماثلة من الارتباطات بين المفردات رغم اختلاف طريقة عرضها.

(ب) بالنسبة لتأثير عدد المفردات في معاملات ارتباطها المصححة بالدرجة الكلية مع تثبيت طريقة عرضها:

تمت المقارنة بين متوسطي قيم "Z" لفيشر في حالة اختلاف عدد مفردات الاختبار عند عرض مفردات الاختبار بكل طريقة من الطرق الثلاثة على حده، باستخدام اختبار Z observed لمجموعتين مستقلتين وفقًا للمعادلة:

$$Z_{observed} = (z_1 - z_2) / (\text{square root of } [\ (1 / (N_1 - 3)) + (1 / (N_2 - 3))$$

حيث تشير Z_1, Z_2 إلى متوسطي قيم Z– Fisher للمفردات في الصورتين الاختباريتين الأولى والثانية المختلفتين في عدد المفردات مع تثبيت طريقة عرضها، وتشير N1, N2 إلى حجمي المجموعتين، ويتم استخدام توزيع t لإختبار الفرق عند مستوى دلالة (٠.٠٥)، حيث القيمة الحرجة هي (١.٩٦) وكانت النتائج كما في جدول (٦).

جدول (٦)

نتائج اختبار "Z observed" للمقارنة بين متوسطي قيم "Z" لفيشر باختلاف عدد المفردات

طريقة العرض	الصورة الاختبارية	Z	N	Z observed	مستوى الدلالة	الدلالة
الكلية	المكونة من ٣٠ مفردة	٠.٧١٤	٧٥	٠.٣٤٨	٠.٣٦٤	غير دالة
	المكونة من ٤٠ مفردة	٠.٧٧١	٧٥			
التتابعية مع حظر الرجوع	المكونة من ٣٠ مفردة	٠.٥٩٦	٧٥	٠.٢٧٥	٠.٣٩٢	غير دالة
	المكونة من ٤٠ مفردة	٠.٦٤٢	٧٥			
التتابعية بدون حظر الرجوع	المكونة من ٣٠ مفردة	٠.٧٦٧	٧٥	٠.٣٨٤	٠.٣٥٠	غير دالة
	المكونة من ٤٠ مفردة	٠.٨٣١	٧٥			

ويتضح من جدول (٦) أنه لا يوجد فرق دال إحصائيًا بين متوسطي معاملات ارتباط المفردات المصححة بالدرجة الكلية والمحولة إلى "Z" فيشر في الصورتين المختلفتين في عدد مفرداتهما، وذلك في كل طرق العرض الثلاثة، وقد يعود ذلك إلى أنه بزيادة طول الاختبار لا يتأثر اتساق إجابات الطالبات عن مفردات الاختبار؛ نظرًا لأن الصورتين الاختباريتين مبنيتين وفق جدول مواصفات واحد، له نفس الأوزان النسبية لموضوعات المحتوى ونفس الأوزان النسبية للمستويات المعرفية للأهداف، كما قد يعود إلى أن تأثير العوامل التي تؤثر في الإجابة عن مفردات الاختبار يكون ثابتًا رغم زيادة عدد المفردات.

وتختلف هذه النتيجة جزئيًا مع ما أشار إليه مافليتوفا وكوبر وليبيديف (Mavletova, Couper & Lebedev, 2018) من أن زيادة طول المقياس الإلكتروني يؤثر في تكافؤ القياس.

(ج)بالنسبة لتأثير لاختلاف طريقة عرضها وعددها في معاملات ارتباطها المصححة بالدرجة الكلية:

تمت المقارنة بين متوسطات قيم "Z" لفيشر في حالة اختلاف طريقة عرض المفردات وعددها باستخدام تحليل التباين الأحادي (أي في المجموعات الستة)، وكانت النتائج كما في جدول (٧).

جدول (٧)

نتائج تحليل التباين الأحادي للمقارنة بين متوسطات قيم "Z" فيشر باختلاف طرق عرض المفردات وعددها

مصدر التباين	مجموع المربعات	درجة الحرية	متوسط المربعات	قيمة ف	مستوى الدلالة	مربع ايتا	حجم الأثر
طريقة عرض المفردات وعددها	٠.٥٣٦	٥	٠.١٠٧	١٢.٠٩٩	٠.٠٠٠	٠.٢٢٩	كبير
الخطأ	١.٨٠٧	٢٠٤	٠.٠٠٩				
الكلي	٢.٣٤٣	٢٠٩					

ويتضح من جدول (٧) وجود فرق دال إحصائيًا بين معاملات ارتباط المفردات المصححة بالدرجة الكلية (بعد تحويلها إلى قيم "Z" فيشر) بين الصور الاختبارية المختلفة في عدد مفرداتها وطريقة عرضها والتي تم تطبيقها على مجموعات الطالبات الست، ولمعرفة اتجاه الفرق، تم إجراء مقارنة بعدية باستخدام اختبار شيفيه للمقارنة بين كل صورتين من الصور الاختبارية معًا، وكانت النتائج كما في جدول (٨).

جدول (٨)

نتائج اختبار شيفيه بين متوسطي قيم "Z" فيشر بين كل صورتين من الصور الاختبارية

المجموعة		أ	ب	ج	د	هـ
ب	الفرق بين المتوسطين	٠.٠٧٩				
	مستوى الدلالة	٠.٠٦٥				
ج	الفرق بين المتوسطين	-٠.٠٣٢	-٠.١١١			
	مستوى الدلالة	٠.٨٧٩	*٠.٠٠١			
د	الفرق بين المتوسطين	-٠.٠٣٤	-٠.١١٣	-٠.٠٠٢		
	مستوى الدلالة	٠.٧٩٩	*٠.٠٠٠	١		
هـ	الفرق بين المتوسطين	٠.٠٤٧	-٠.٠٣٢	٠.٠٧٩	٠.٠٨٢	
	مستوى الدلالة	٠.٥١٥	٠.٨٥٠	*٠.٠٣٦	*٠.٠١٢	
و	الفرق بين المتوسطين	-٠.٠٦٧	-٠.١٤٧	-٠.٠٣٥	-٠.٠٣٣	-٠.١١٥
	مستوى الدلالة	٠.١١٨	*٠.٠٠٠	٠.٧٨٥	٠.٧٨٢	*٠.٠٠٠

ويتضح من جدول (٨) وجود فرق دال إحصائيًا بين متوسطي معاملات ارتباط المفردات المصححة بالدرجة الكلية (بعد تحويلها إلى قيم "Z" فيشر) بين الصورة الاختبارية المكونة من (٣٠) مفردة ويتم عرض المفردات فيها تتابعيًا مع حظر الرجوع (المطبقة على المجموعة ب) وكل من الصورة الاختبارية المكونة من (٣٠) مفردة ويتم عرض المفردات فيها تتابعيًا دون حظر الرجوع (المطبقة على المجموعة ج)، والصورة المكونة من (٤٠) مفردة ويتم عرض المفردات فيها بطريقة كلية (المطبقة على المجموعة د)، والصورة المكونة من (٤٠) مفردة ويتم عرض المفردات فيها تتابعيًا دون حظر الرجوع (المطبقة على المجموعة و)، وهذه الفروق كلها في اتجاه الصور المطبقة على المجموعات (ج، د، و).

كما وجد فرق دال إحصائيًا بين متوسطي معاملات ارتباط المفردات المصححة بالدرجة الكلية (بعد تحويلها إلى قيم "Z" فيشر) بين الصورة الاختبارية المكونة من (٤٠) مفردة ويتم عرض المفردات فيها تتابعيًا مع حظر الرجوع (المطبقة على المجموعة ه)، وكل من الصورة الاختبارية المكونة من (٣٠) مفردة ويتم عرض المفردات فيها تتابعيًا بدون حظر

الرجوع (المطبقة على المجموعة ج)، وكذلك الصورة الاختبارية المكونة من (٤٠) مفردة ويتم عرض المفردات فيها بطريقة كلية (المطبقة على المجموعة د)، وكذلك الصورة الاختبارية المكونة من (٤٠) مفردة ويتم عرض المفردات فيها تتابعيًا بدون حظر الرجوع (المطبقة على المجموعة و)، وهذا الفرق في اتجاه متوسط معاملات الارتباط المصححة في الصور المطبقة على المجموعات (ج، د، و).

ويتضح من خلال هذه النتائج أفضلية جميع المقارنات الثنائية التي تمت بين معاملات ارتباط مفردات الصورة الاختبارية المصححة المحولة بالدرجة الكلية في أغلب الصور الاختبارية عن نظيراتها في الصورتين الاختباريتين (ب، هـ) اللتين تم عرض المفردات فيها بطريقة تتابعية مع حظر الرجوع، وعدم وجود فرق دال إحصائيًا بين معاملات الارتباط المصححة المحولة بين هاتين الصورتين (ب، هـ)، أي أن طريقة عرض المفردات مع حظر الرجوع تحقق أقل ارتباطات مصححة بين المفردات والدرجة الكلية مقارنة بطرق العرض الأخرى وبغض النظر عن عدد مفردات الاختبار، وهذا قد يرجع إلى أن حظر رجوع الطالبات للإجابة عن المفردات قد يؤدي بهن إلى التسرع في الإجابة عن المفردات إحساسًا منهن بأنه ليس لديهن فرصة أخرى، كما أنهن لا يأخذن الوقت الكافي للتفكير في الإجابة خوفًا من ضياع وقت الاختبار، وهذا يؤدي بهن إلى الإجابة عن بعض المفردات بطريقة لا تتسق مع إجاباتهن عن باقي المفردات مما يقلل متوسط معاملات ارتباط المفردات المصححة بالدرجة الكلية، وتتفق هذه النتيجة جزئيا مع نتيجة دراسة كاليجارو وشاندلوبز ودينيز (Callegaro, Shand-Lubbers & Dennis, 2009)، كما تختلف جزئيًا مع نتيجة دراسة ريفيلا وتونينيلي واوشوا (Revilla, Toninelli & Ochoa, 2016).

التحقق من الفرض الثاني:

ينص الفرض الثاني على: لا يختلف ترتيب معاملات ارتباط درجات الطالبات على الاختبار الإلكتروني بالمعدل التراكمي باختلاف طريقة عرضها وعددها، وللتحقق من هذا الفرض تم حساب معامل ارتباط درجات الطالبات في المجموعات الست بدرجاتهم في المعدل التراكمي، وكان ترتيب معاملات الارتباط كما في جدول (٩).

جدول (٩)

معاملات ارتباط درجات الطالبات في المجموعات الست بدرجاتهم في المعدل التراكمي

المجموعة	الصورة الاختبارية	معامل الارتباط
و	المكونة من (٤٠) مفردة، مع العرض التتابعي للمفردات بدون حظر بالرجوع	٠.٩١٥
د	المكونة من (٤٠) مفردة، مع العرض الكلي للمفردات	٠.٨٨٩
ج	المكونة من (٣٠) مفردة، مع العرض التتابعي للمفردات بدون حظر الرجوع	٠.٨٧٢
أ	المكونة من (٣٠) مفردة، مع العرض الكلي للمفردات	٠.٨٦٨
هـ	المكونة من (٤٠) مفردة، مع العرض التتابعي للمفردات مع حظر الرجوع	٠.٨٢١
ب	المكونة من (٣٠) مفردة، مع العرض التتابعي للمفردات مع حظر الرجوع	٠.٨٠٩

ويتضح من خلال جدول (٩) أنه على الرغم من تمتع جميع الصور الاختبارية بارتباطات عالية بالمعدل التراكمي للطالبات، إلا أن الصورة المكونة من (٤٠ مفردة) وتم عرض مفرداتها بطريقة تتابعية بدون حظر الرجوع حققت أعلى ارتباط، كما حققت الصورة المكونة من (٤٠ مفردة) وتم عرض مفرداتها بطريقة كلية المرتبة الثانية من حيث الارتباط، وهذا قد يعود إلى أنه مع زيادة عدد مفردات الاختبار تزداد قدرة الاختبار على تغطية المحتوى المراد قياسه بشكل أكبر مما يزيد من صدق الاختبار نسبيًا عن نظيراتها المكونة من (٣٠) مفردة وتم عرض مفرداتها بنفس الطريقتين، كما أنه عند عرض المفردات بطريقة تتابعية مع السماح بالرجوع أو بطريقة كلية فإنه توجد فرصة لدى الطالبات للمراجعة والتركيز بشكل أكبر في الإجابة عن المفردات وتذكر مزيد من المعلومات في المحتوى الدراسي، ومن ثم الاستفادة من المعرفة الجزئية أو الكلية بالمحتوى الذي تقيسه المفردات في الإجابة، وهذا يتسق مع مستوياتهم التحصيلية في المقررات الدراسية المختلفة الأخرى والتي يعبر عنها من خلال المعدل التراكمي، أما أقل ارتباط بالمعدل التراكمي فقد ظهر في الصورتين الاختباريتين اللتين تم عرض مفرداتهما بطريقة تتابعية مع حظر الرجوع، وقد يعود ذلك إلى أن هذه الطريقة في عرض المفردات قد لا تلائم كثيرًا من الطالبات اللائي يحتجن إلى تأجيل إجابة بعض المفردات الصعبة إلى نهاية الاختبار لكي يولون لها مزيدًا من

التركيز في الإجابة أو الوقت للتفكير في الإجابة عنها، وهذا ما لا يتاح لهن في هذه الطريقة من طرق العرض، بعكس ما يحدث في اختبارات أغلب المقررات الدراسية الأخرى، مما قلل من معامل ارتباط درجات الطالبات في هاتين الصورتين بالمعدل التراكمي عن باقي الصور الاختبارية.

وتتفق هذه النتيجة جزئيًا مع نتيجة دراسة محمود (٢٠١٨) التي أشارت إلى وجود اختلاف بين معاملات الصدق التلازمي بين الدرجات على الاختبار الإلكتروني والمعدل التراكمي، كما تتفق جزئيًا مع نتيجة دراسة ديبيل ورفاقه (Debell, et al., 2021) من وجود اختلاف في صدق المحك يعزى إلى طريقة العرض، وتختلف جزئيًا مع نتيجة دراسة مافليتوفا وكوبر وليبيديف (Mavletova, Couper & Lebedev, 2018).

التحقق من الفرض الثالث:

ينص الفرض الثالث على: لا توجد فروق دالة إحصائيًا في متوسطات معاملات صعوبة مفردات الاختبار الإلكتروني التي تعزى إلى:

٣- أ) طريقة عرضها مع تثبيت عددها.

٣- ب) عددها مع تثبيت طريقة عرضها.

٣- ج) اختلاف طريقة عرضها وعددها.

وللتحقق من هذا الفرض تم حساب معاملات صعوبة المفردات في الصور الاختبارية في المجموعات الستة (أ، ب، ج، د، ه، و)، وكانت الإحصاءات الوصفية لمعاملات الصعوبة كما في جدول (١٠)

جدول (١٠)

الإحصاءات الوصفية لمعاملات صعوبة المفردات في الصور الاختبارية

طريقة العرض	الكلية		التتابعية مع حظر الرجوع		التتابعية بدون حظر الرجوع	
عدد المفردات	٣٠ مفردة	٤٠ مفردة	٣٠ مفردة	٤٠ مفردة	٣٠ مفردة	٤٠ مفردة
المتوسط	٠.٤٣١	٠.٤٦٤	٠.٣٦٧	٠.٣٨٩	٠.٤٦٥	٠.٥٠٦
الانحراف المعياري	٠.٠٩٠	٠.٠٧٥	٠.١٠٨	٠.٠٩٨	٠.٠٨٤	٠.٠٦٤

(أ) بالنسبة لتأثير طريقة عرض المفردات في معاملات صعوبتها مع تثبيت عددها:

تمت المقارنة بين معاملات صعوبة المفردات في الثلاث مجموعات (أ، ب، ج) التي تعرضت للصورة الاختبارية الأولى المكونة من (٣٠) مفردة باستخدام تحليل التباين الأحادي، كما تم نفس الإجراء مع الثلاث مجموعات (د، هـ، و) التي تعرضت للصورة الاختبارية الثانية المكونة من (٤٠) مفردة، كما في جدول (١١).

جدول (١١)

نتائج تحليل التباين الأحادي للمقارنة بين متوسطات معاملات صعوبة المفردات باختلاف طرق العرض

الاختبار	مصدر التباين	مجموع المربعات	درجة الحرية	متوسط المربعات	قيمة ف	مستوى الدلالة	مربع ايتا	حجم الأثر
الاختبار المكون من (٣٠) مفردة	طريقة العرض	٠.١٥١	٢	٠.٠٧٥	٨.٤٢٦	٠.٠٠٠	٠.١٦٣	كبير
	الخطأ	٠.٧٧٨	٨٧	٠.٠٠٩				
	الكلي	٠.٩٢٩	٨٩					
الاختبار المكون من (٤٠) مفردة	طريقة العرض	٠.٢٧٨	٢	٠.١٣٩	٢١.٦١٥	٠.٠٠٠	٠.٢٧٠	كبير
	الخطأ	٠.٧٥٢	١١٧	٠.٠٠٦				
	الكلي	١.٠٢٩	١١٩					

ويتضح من خلال جدول (١١) وجود فرق دال إحصائيًا بين متوسطات معاملات صعوبة المفردات باختلاف طرق عرض المفردات، سواء في حالة الصورة الاختبارية المكونة من (٣٠ أو ٤٠) مفردة، أي أن صعوبة مفردات الاختبار تتوقف على طريقة عرض مفرداته، ولمعرفة اتجاه الفرق، تم إجراء مقارنة بعدية باستخدام اختبار شيفيه للمقارنة بين كل مجموعتين تعرضتا لنفس عدد المفردات، وكانت النتائج كما في جدول (١٢).

جدول (١٢)

نتائج اختبار شيفيه بين متوسطي معاملات صعوبة مفردات كل صورتين اختباريتين لهما نفس الطول

	عدد المفردات	٣٠ مفردة		٤٠ مفردة	
طريقة العرض		الكلية	التتابعية مع حظر الرجوع	الكلية	التتابعية مع حظر الرجوع
التتابعية مع حظر الرجوع	الفرق بين المتوسطين	٠.٠٦٥		٠.٠٧٤	
	مستوى الدلالة	*٠.٠٣٤		*٠.٠٠٠	
التتابعية بدون حظر الرجوع	الفرق بين المتوسطين	-٠.٠٣٤	-٠.٠٩٩	-٠.٠٤٢	-٠.١١٦
	مستوى الدلالة	٠.٣٨٣	*٠.٠٠٠	٠.٠٦٢	*٠.٠٠٠

ويتضح من جدول (١٢) أنه يوجد فرق دال إحصائيًا بين متوسطي معاملات صعوبة المفردات في الصورتين اللتين تم عرض مفرداتهما بالطريقة الكلية وبالطريقة التتابعية مع حظر الرجوع، وذلك في اتجاه متوسط معاملات صعوبة المفردات في حالة عرضها بطريقة كلية، وذلك سواءً كان عدد المفردات (٣٠ أو ٤٠) مفردة، وكذلك يتضح وجود فرق دال إحصائيًا بين متوسطي معاملات صعوبة المفردات في الصورتين اللتين تم عرض مفرداتهما بالطريقة التتابعية بدون حظر الرجوع وبالطريقة التتابعية مع حظر الرجوع، وذلك في اتجاه متوسط معاملات صعوبة المفردات في حالة عرضها بطريقة تتابعية بدون حظر الرجوع، وذلك سواءً كان عدد المفردات (٣٠ أو ٤٠) مفردة، أي بغض النظر عن طول الاختبار أو عدد مفرداته، وهذا يعني أن المفردات كانت أصعب على الطالبات عندما تم عرضها بالطريقة التتابعية مع حظر الرجوع عما كانت عليه عند عرض المفردات سواء بالطريقة الكلية أو التتابعية مع حظر الرجوع (معامل صعوبة المفردات يتناسب عكسيًا مع مستوى صعوبتها)، وقد يرجع ذلك إلى أن عرض المفردات بطريقة تتابعية مع حظر الرجوع لا تتيح للطالبات فرصة كافية في الإجابة عن المفردات التي قد تكون صعبة عليهن وتحتاج إلى مزيد من التركيز، كما أنها لا تسمح لهن بالاستفادة من الدلائل أو المنبهات أو المفاتيح التي قد تتواجد في المفردات الأخرى في الإجابة عن المفردة المعروضة أمامهن، وهذا يزيد من صعوبة الإجابة عن المفردة، وهذا الأثر لا يتواجد في حالة عرض المفردات بطريقة كلية أو بطريقة تتابعية مع السماح للطالبات بالرجوع للمفردات التي تركوا الإجابة عنها في نهاية

الاختبار، والاستفادة من الوقت المتبقي في نهاية الاختبار في زيادة التفكير والتركيز في الإجابة عنها، مما يزيد من فرص الإجابة الصحيحة عن المفردات، أي أن هناك فرصة أخرى لتغيير الإجابة عن المفردات وزيادة التفكير فيها وتذكر معلومات المحتوى، وهذا ما لا يتوفر في الطريقة التتابعية مع حظر الرجوع، وتتفق هذه النتيجة جزئيًا مع ما توصلت إليه نتيجة دراسة محمود (٢٠١٨) التي توصلت إلى وجود فرق دال في معاملات صعوبة المفردات تعزى لاختلاف طريقة عرضها.

(ب) بالنسبة لتأثير عدد المفردات في معاملات صعوبتها مع تثبيت طريقة عرضها:

تم استخدام اختبار "ت" للمقارنة بين متوسطي معاملات الصعوبة في الصورتين المختلفتين في عدد مفرداتهما في الحالات الثلاثة المختلفة وفقًا لطريقة العرض، وكانت النتائج كما في جدول (١٣).

جدول (١٣)

نتائج اختبار "ت" للمقارنة بين متوسطي معاملات الصعوبة في الصورتين المختلفتين في عدد مفرداتهما

طريقة العرض	الصورة الاختبارية	المتوسط	الانحراف المعياري	قيمة ت	مستوى الدلالة	الدلالة
الكلية	المكونة من ٣٠ مفردة	٠.٤٣١	٠.٠٩٠	١.٦٣٣	٠.١٠٧	غير دالة
	المكونة من ٤٠ مفردة	٠.٤٦٤	٠.٠٧٥			
التتابعية مع حظر الرجوع	المكونة من ٣٠ مفردة	٠.٣٦٧	٠.١٠٨	٠.٩٣٢	٠.٣٥٥	غير دالة
	المكونة من ٤٠ مفردة	٠.٣٨٩	٠.٠٩٨			
التتابعية بدون حظر الرجوع	المكونة من ٣٠ مفردة	٠.٤٦٥	٠.٠٨٤	٢.٣١٨	٠.٠٢٣	دالة
	المكونة من ٤٠ مفردة	٠.٥٠٦	٠.٠٦٤			

ويتضح من خلال جدول (١٣) عدم وجود فرق دال إحصائيًا بين متوسطي معاملات صعوبة المفردات في الصورتين الاختباريتين المكونتين من (٣٠ و ٤٠) مفردة، سواء عند عرض المفردات بالطريقة الكلية أو عند عرضها بالطريقة التتابعية مع حظر الرجوع، بينما وجد فرق دال إحصائيًا بين متوسطي معاملات صعوبة المفردات في الصورتين الاختباريتين المكونتين من (٣٠ و ٤٠) مفردة عند عرض المفردات بالطريقة التتابعية بدون حظر

الرجوع، وقد يعود ذلك إلى أنه مع زيادة عدد المفردات في ظل طريقة العرض التتابعي بدون حظر الرجوع تزداد فرصة الاستفادة من إجابة بعض المفردات في تذكر الإجابة عن مفردات أخرى، ويزداد الوقت المتاح أمام الطالبة للإجابة عن المفردات التي تركتها، فيزداد تركيزها في الإجابة ويسهل عليها الإجابة بطريقة صحيحة، كما أن زيادة عدد المفردات قد يذكر الطالبات بالإجابة عن بعض المفردات الصعبة في الاختبار والتي يمكنهن الرجوع إليها وتغيير الإجابة عنها، وهذا يقلل من مستوى صعوبة المفردات ويزيد من معامل صعوبتها في الصورة المكونة من (٤٠) مفردة عن المكونة من (٣٠) مفردة، وتتفق هذه النتيجة جزئيًا مع نتيجة دراسة محمود (٢٠١٨).

(ج)بالنسبة لتأثير اختلاف طريقة عرض المفردات وعددها في معاملات صعوبتها:

تمت المقارنة بين متوسطات قيم معاملات الصعوبة في حالة اختلاف طريقة عرض المفردات وعددها باستخدام تحليل التباين الأحادي، وكانت النتائج كما في جدول (١٤).

جدول (١٤)

نتائج تحليل التباين الأحادي بين متوسطات معاملات الصعوبة باختلاف طرق عرض المفردات وعددها

مصدر التباين	مجموع المربعات	درجة الحرية	متوسط المربعات	قيمة ف	مستوى الدلالة	مربع ايتا	حجم الأثر
طريقة عرض المفردات وعددها	٠.٤٨١	٥	٠.٠٩٦	١٢.٨٣٥	٠.٠٠٠	٠.٢٣٩	كبير
الخطأ	١.٥٣٠	٢٠٤	٠.٠٠٧				
الكلي	٢.٠١١	٢٠٩					

ويتضح من جدول (١٤) وجود فرق دال إحصائيًا بين متوسطات معاملات صعوبة المفردات في الصور الاختبارية الست المختلفة في عددها وطريقة عرضها، ولمعرفة اتجاه هذا الفرق، تم إجراء مقارنة بعدية باستخدام اختبار شيفيه للمقارنة بين كل صورتين معًا، وكانت النتائج كما في جدول (١٥).

جدول (١٥)

نتائج اختبار شيفيه بين متوسطي معاملات صعوبة المفردات بين كل صورتين من الصور الاختبارية

		أ	ب	ج	د	هـ
ب	الفرق بين المتوسطين	٠.٠٦٥				
	مستوى الدلالة	٠.١٤٣				
ج	الفرق بين المتوسطين	-٠.٠٣٤	-٠.٠٩٩			
	مستوى الدلالة	٠.٨٠٤	*٠.٠٠٢			
د	الفرق بين المتوسطين	-٠.٠٣٢١	-٠.٠٩٧	٠.٠٠٢		
	مستوى الدلالة	٠.٧٩٦	*٠.٠٠١	١		
هـ	الفرق بين المتوسطين	٠.٠٤٢	-٠.٠٢٣	٠.٠٧٦	٠.٠٧٤	
	مستوى الدلالة	٠.٥٥٨	٠.٩٤٣	*٠.٠٢٦	*٠.٠١٥	
و	الفرق بين المتوسطين	-٠.٠٧٥	-٠.١٣٩	-٠.٠٤١	-٠.٠٤٣	-٠.١١٦
	مستوى الدلالة	*٠.٠٢٨	*٠.٠٠٠	٠.٥٧٦	٠.٤٣٥	*٠.٠٠٠

ويتضح من جدول (١٥) وجود فرق دال إحصائيًا في سبع مقارنات ثنائية، حيث تزداد مستويات صعوبة (أي تقل معاملات صعوبة) مفردات الصورة الاختبارية المكونة من (٣٠) مفردة والمعروضة مفرداتها بالطريقة التتابعية مع حظر الرجوع المطبقة على المجموعة (ب) عن الصورة المطبقة على المجموعة (ج) المكونة من (٣٠) مفردة بطريقة تتابعية بدون حظر الرجوع، وكذلك عن الصورة المطبقة على المجموعة (د) المكونة من (٤٠) مفردة بطريقة كلية، وكذلك عن الصورة المطبقة على المجموعة (و) المكونة من (٤٠) مفردة بطريقة تتابعية بدون حظر الرجوع.

كما يتضح ازدياد مستويات صعوبة (أي قلة معاملات صعوبة) مفردات الصورة الاختبارية المكونة من (٤٠) مفردة والمعروضة مفرداتها بالطريقة التتابعية مع حظر الرجوع المطبقة على المجموعة (هـ) عن الصورة المطبقة على المجموعة (ج) المكونة من (٣٠) مفردة بطريقة تتابعية بدون حظر الرجوع، وعن الصورة المطبقة على المجموعة (د) المكونة من (٤٠) مفردة بطريقة كلية، وكذلك عن الصورة المطبقة على المجموعة (و) المكونة من (٤٠) مفردة بطريقة تتابعية بدون حظر الرجوع.

وبالتالي يتضح سهولة إجابة الطالبات عن المفردات التي يتم عرضها سواء بطريقة كلية أو بطريقة تتابعية بدون حظر الرجوع عن المفردات التي يتم عرضها تتابعيًا مع حظر الرجوع سواء كان عدد المفردات (٣٠ أو ٤٠) مفردة، وهذا يؤكد ما سبقت الإشارة إليه من أن حظر رجوع الطالبات للمفردات التي أجابوا عنها يؤدي إلى عدم قدرة الطالبات على التركيز بالشكل الكافي في الإجابة عن المفردات الصعبة والتي اعتدن على تركها لنهاية الاختبار لكي يحصلن على فرصة أكبر ووقت أطول في التفكير فيها؛ مما يؤثر في صحة إجابتهن عن هذه المفردات، فيظهر عدد أكبر من الأخطاء في الإجابة عن مثل هذه المفردات فتقل معاملات صعوبتها، وذلك مقارنة بما يحدث للطالبات في حالة العرض الكلي للمفردات أو العرض التتابعي بدون حظر الرجوع؛ ففي هذين العرضين تستطيع الطالبات ترك المفردات الصعبة إلى نهاية الاختبار من أجل مزيد من التركيز والتدقيق واستعادة المعلومات.

كما يتضح ازدياد مستويات صعوبة مفردات الصورة الاختبارية المكونة من (٣٠) مفردة والمعروضة مفرداتها بالطريقة الكلية المطبقة على المجموعة (أ) (أي قلة معاملات صعوبتها) عن الصورة المطبقة على المجموعة (و) المكونة من (٤٠) مفردة بطريقة تتابعية بدون حظر الرجوع، وقد يرجع ذلك إلى أنه عند عرض المفردات بطريقة كلية قد يزداد العبء المعرفي عند الطالبات نتيجة ظهور جميع المفردات في شاشة واحدة؛ مما قد يضعف تركيزهن في الإجابة، عن العبء المعرفي الذي تتعرض له الطالبات في حالة العرض التتابعي للمفردات بدون حظر الرجوع، كما قد ينسيهن الإجابة عن بعض المفردات المتروكة بدون إجابة أو ينسيهن تغيير الإجابة عنها عند المراجعة.

التحقق من الفرض الرابع:

ينص الفرض الرابع على: "لا توجد فروق دالة إحصائيًا في متوسطات معاملات تمييز مفردات الاختبار الإلكتروني التي تعزى إلى:

٤- أ) طريقة عرضها مع تثبيت عددها.

٤- ب) عددها مع تثبيت طريقة عرضها.

٤- ج) اختلاف طريقة عرضها وعددها.

وللتحقق من هذا الفرض تم حساب معاملات تمييز المفردات في الصور الاختبارية في المجموعات الستة (أ، ب، ج، د، ه، و)، وجاءت إحصاءاتها الوصفية كما في جدول (١٦).

جدول (١٦)

الإحصاءات الوصفية لمعاملات تمييز المفردات في الصور الاختبارية

طريقة العرض	الكلية		التتابعية مع حظر الرجوع		التتابعية بدون حظر الرجوع	
عدد المفردات	٣٠ مفردة	٤٠ مفردة	٣٠ مفردة	٤٠ مفردة	٣٠ مفردة	٤٠ مفردة
المتوسط	٠.٤٣١	٠.٤٦٤	٠.٣٦٧	٠.٣٨٩	٠.٥٥٩	٠.٥٩٢
الانحراف المعياري	٠.٠٩٠	٠.٠٧٥	٠.١٠٨	٠.٠٩٨	٠.٠٧٥	٠.٠٥٥

(أ) بالنسبة لتأثير طريقة عرض المفردات في معاملات تمييزها مع تثبيت عددها:

تمت المقارنة بين معاملات تمييز المفردات في المجموعات الثلاث (أ، ب، ج) التي تعرضت للصورة الاختبارية الأولى المكونة من (٣٠) مفردة باستخدام تحليل التباين الأحادي، كما تم نفس الإجراء مع الثلاث مجموعات (د، هـ، و) التي تعرضت للصورة الاختبارية الثانية المكونة من (٤٠) مفردة، كما في جدول (١٧).

جدول (١٧)

نتائج تحليل التباين الأحادي للمقارنة بين متوسطات معاملات تمييز المفردات باختلاف طرق العرض

الاختبار	مصدر التباين	مجموع المربعات	درجة الحرية	متوسط المربعات	قيمة ف	مستوى الدلالة	مربع ايتا	حجم الأثر
الاختبار المكون من (٣٠) مفردة	طريقة العرض	٠.١٥٠	٢	٠.٠٧٥	١٠.٠٨٤	٠.٠٠٠	٠.١٨٨	كبير
	الخطأ	٠.٦٤٦	٨٧	٠.٠٠٧				
	الكلي	٠.٧٩٦	٨٩					
الاختبار المكون من (٤٠) مفردة	طريقة العرض	٠.٢١٢	٢	٠.١٠٦	١٩.٦٠	٠.٠٠٠	٠.٢٥١	كبير
	الخطأ	٠.٦٣٢	١١٧	٠.٠٠٥				
	الكلي	٠.٨٤٤	١١٩					

ويتضح من خلال جدول (١٧) وجود فرق دال إحصائيًا بين متوسطات معاملات تمييز المفردات باختلاف طرق عرضها، سواء في حالة الصورة الاختبارية المكونة من (٣٠ أو ٤٠) مفردة، ولمعرفة اتجاه الفرق، تم إجراء مقارنة بعدية باستخدام اختبار شيفيه بين كل مجموعتين تعرضتا لنفس عدد المفردات، كما في جدول (١٨).

جدول (١٨)

نتائج اختبار شيفيه بين متوسطي معاملات تمييز مفردات كل صورتين اختباريتين لهما نفس الطول

	عدد المفردات	٣٠ مفردة		٤٠ مفردة	
طريقة العرض		الكلية	التتابعية مع حظر الرجوع	الكلية	التتابعية مع حظر الرجوع
التتابعية مع حظر الرجوع	الفرق بين المتوسطين	٠.٠٦٣		٠.٠٦٨	
	مستوى الدلالة	*٠.٠٢٢		*٠.٠٠٠	
التتابعية بدون حظر الرجوع	الفرق بين المتوسطين	-٠.٠٣٦	-٠.٠٩٨	-٠.٠٣٣	-٠.١٠١
	مستوى الدلالة	٠.٢٨٢	*٠.٠٠٠	٠.١٣٤	*٠.٠٠٠

ويتضح من خلال جدول (١٨) أفضلية معاملات تمييز المفردات عند عرضها بالطريقة الكلية أو عند عرضها بالطريقة التتابعية بدون حظر الرجوع عن معاملات تمييز المفردات في حال عرضها بالطريقة التتابعية مع حظر الرجوع، سواء عندما كان عدد المفردات (٣٠ أو ٤٠) مفردة، وقد يرجع ذلك إلى أن الطالبات منخفضات المستوى في حالة عرض المفردات مع حظر الرجوع قد يلجأن إلى التخمين العشوائي للإجابة عن المفردة التي يصعب عليهن الإجابة عنها بدلًا من تركها بدون إجابة، وهذا قد يؤدي إلى إجابتهن بطريقة صحيحة عن المفردات الصعبة دون معرفة بالمحتوى المراد قياسه؛ كما أن الطالبات مرتفعات المستوى قد يتسرعن في الإجابة عن المفردات نتيجة قلقهن من ضياع الوقت في التفكير وعدم وجود فرصة أخرى للتفكير في الإجابة في نهاية الاختبار؛ مما قد يزيد من خطأهن في الإجابة عن المفردات، وهذا كله يؤدي إلى انخفاض قدرة المفردة على التمييز بين الطالبات مرتفعات ومنخفضات التحصيل الدراسي، وتختلف هذه النتيجة مع ما توصلت إليه نتيجة دراسة محمود (٢٠١٨) التي توصلت إلى عدم وجود فرق دال إحصائيًا في معاملات تمييز المفردات وفق طريقة عرضها.

(ب) بالنسبة لتأثير عدد المفردات في معاملات تمييزها مع تثبيت طريقة عرضها:

تم استخدام اختبار "ت" للمقارنة بين متوسطي معاملات التمييز في الصورتين المختلفتين في عدد مفرداتهما في الحالات الثلاثة المختلفة وفقًا لطريقة العرض، كما في جدول (١٩).

جدول (١٩)

نتائج اختبار "ت" للمقارنة بين متوسطي معاملات التمييز في الصورتين المختلفتين في عدد مفرداتهما

طريقة العرض	الصورة الاختبارية	المتوسط	الانحراف المعياري	قيمة ت	مستوى الدلالة	الدلالة
الكلية	المكونة من ٣٠ مفردة	٠.٤٣١	٠.٠٩٠	١.٩٧٦	٠.٠٥٢	غير دالة
	المكونة من ٤٠ مفردة	٠.٤٦٤	٠.٠٧٥			
التتابعية مع حظر الرجوع	المكونة من ٣٠ مفردة	٠.٣٦٧	٠.١٠٨	١.٣٠٧	٠.١٩٦	غير دالة
	المكونة من ٤٠ مفردة	٠.٣٨٩	٠.٠٩٨			
التتابعية بدون حظر الرجوع	المكونة من ٣٠ مفردة	٠.٥٥٩	٠.٠٧٥	٢.٠٩٥	٠.٠٤٠	دالة
	المكونة من ٤٠ مفردة	٠.٥٩٢	٠.٠٥٥			

ويتضح من جدول (١٩) وجود فرق دال إحصائيًا بين متوسطي معاملات تمييز المفردات في الصورتين الاختباريتين المكونتين من (٣٠ و ٤٠) مفردة في حالة عرض المفردات بطريقة تتابعية بدون حظر الرجوع، بينما لم يتضح وجود فرق دال إحصائيًا بين متوسطي معاملات تمييز المفردات في الصورتين الاختباريتين المكونتين من (٣٠ و ٤٠) مفردة، في حالة عرض المفردات بطريقة كلية أو تتابعية مع حظر الرجوع.

أي أنه مع زيادة طول الاختبار تزداد قدرة المفردات على التمييز عند عرضها تتابعيًا بدون حظر الرجوع، وقد يرجع ذلك إلى أن عرض المفردات بهذه الطريقة يتيح للطالبات في المستويات العليا الفرصة لترك المفردات التي يصعب عليهن الإجابة عنها إلى نهاية الاختبار وإمعان التفكير فيها، وهذا يزيد من فرصة الطالبات مرتفعات المستوى في الوصول إلى الإجابة الصحيحة عنها؛ في الوقت الذي قد تجيب فيه الطالبات منخفضات المستوى عن هذه المفردات إجابة خطأ؛ مما يزيد من قدرة المفردات على التمييز، وتزداد هذه القدرة على التمييز مع زيادة عدد المفردات ووجود مزيد من المفردات الصعبة المميزة بين الطالبات مرتفعات ومنخفضات المستوى؛ ومع زيادة عدد المفردات تزداد قدرة الاختبار بشكل عام على تغطية المحتوى بشكل أكبر مما يسهم في زيادة التمييز بين طالبات المستويين المرتفع والمنخفض.

(ج) بالنسبة لتأثير اختلاف طريقة عرض المفردات وعددها في معاملات تمييزها:

تمت المقارنة بين متوسطات قيم معاملات التمييز في حالة اختلاف طريقة عرض المفردات وعددها باستخدام تحليل التباين الأحادي، كما في جدول (٢٠).

جدول (٢٠)

نتائج تحليل التباين الأحادي بين متوسطات معاملات التمييز باختلاف طرق عرض المفردات وعددها

مصدر التباين	مجموع المربعات	درجة الحرية	متوسط المربعات	قيمة ف	مستوى الدلالة	مربع ايتا	حجم الأثر
طريقة عرض المفردات وعددها	٠.٤١٧	٥	٠.٠٨٣	١٣.٢٩٥	٠.٠٠٠	٠.٢٤٩	كبير
الخطأ	١.٢٧٨	٢٠٤	٠.٠٠٦				
الكلي	١.٦٩٥	٢٠٩					

ويتضح من جدول (٢٠) وجود فرق دال إحصائيًا بين متوسطات معاملات تمييز المفردات في الصور الاختبارية المختلفة في عددها وطريقة عرضها، ولمعرفة اتجاه هذا الفرق، تم إجراء مقارنة بعدية باستخدام اختبار شيفيه للمقارنة بين كل صورتين معًا، وكانت النتائج كما في جدول (٢١).

جدول (٢١)

نتائج اختبار شيفيه بين متوسطي معاملات تمييز المفردات بين كل صورتين من الصور الاختبارية

		أ	ب	ج	د	هـ
ب	الفرق بين المتوسطين	٠.٠٦٣				
	مستوى الدلالة	٠.٠٩٦				
ج	الفرق بين المتوسطين	-٠.٠٣٦	-٠.٠٩٩			
	مستوى الدلالة	٠.٦٩٣	*٠.٠٠٠			
د	الفرق بين المتوسطين	-٠.٠٣٥	-٠.٠٩٨	٠.٠٠١		
	مستوى الدلالة	٠.٦٤٤	*٠.٠٠٠	١		
هـ	الفرق بين المتوسطين	٠.٠٣٣	-٠.٠٣٠	٠.٠٦٨	٠.٠٦٨	
	مستوى الدلالة	٠.٧١٢	٠.٧٧٣	*٠.٠٢٩	*٠.٠١٤	
و	الفرق بين المتوسطين	-٠.٠٦٨	-٠.١٣١	-٠.٠٣٣	-٠.٠٣٣	-٠.١٠١
	مستوى الدلالة	*٠.٠٢٩	*٠.٠٠٠	٠.٧١٣	٠.٦٢١	*٠.٠٠٠

ويتضح من جدول (٢١) وجود فرق دال إحصائيًا في سبع مقارنات ثنائية، حيث تنخفض معاملات تمييز مفردات الصورة الاختبارية المكونة من (٣٠) مفردة والمعروضة مفرداتها بالطريقة التتابعية مع حظر الرجوع المطبقة على المجموعة (ب) بفرق دال إحصائيًا عن الصورة المطبقة على المجموعة (ج) المكونة من (٣٠) مفردة بطريقة تتابعية بدون حظر الرجوع، وكذلك تنخفض بفرق دال إحصائيًا عن الصورة المطبقة على المجموعة (د) المكونة من (٤٠) مفردة بطريقة كلية، كما تنخفض بفرق دال إحصائيًا عن الصورة المطبقة على المجموعة (و) المكونة من (٤٠) مفردة بطريقة تتابعية بدون حظر الرجوع.

كما انخفضت معاملات تمييز مفردات الصورة الاختبارية المكونة من (٤٠) مفردة والمعروضة مفرداتها بالطريقة التتابعية مع حظر الرجوع المطبقة على المجموعة (هـ) بفرق دال إحصائيًا عن الصورة المطبقة على المجموعة (ج) المكونة من (٣٠) مفردة بطريقة تتابعية بدون حظر الرجوع، وكذلك تنخفض بفرق دال إحصائيًا عن الصورة المطبقة على المجموعة (د) المكونة من (٤٠) مفردة بطريقة كلية، كما تنخفض بفرق دال إحصائيًا عن الصورة المطبقة على المجموعة (و) المكونة من (٤٠) مفردة بطريقة تتابعية بدون حظر الرجوع.

وبالتالي يتضح قدرة المفردات التي يتم عرضها سواء بطريقة كلية أو بطريقة تتابعية بدون حظر الرجوع في التمييز بين المستويات المرتفعة والمنخفضة للطالبات عن المفردات التي يتم عرضها تتابعيًا مع حظر الرجوع سواء كان عدد المفردات (٣٠ أو ٤٠) مفردة، وقد يرجع ذلك إلى أن حظر رجوع بعض الطالبات مرتفعات المستوى إلى المفردات التي يصعب عليهن الإجابة عنها أو تتطلب منهن وقتًا أطول للتفكير فيها، قد يؤدي إلى وقوعهن في أخطاء متعددة نتيجة التسرع في الإجابة والخوف من ضياع الوقت؛ فلا يساعدهن ذلك على أخذ الوقت الكافي للتفكير في الإجابة عنها بدقة؛ مما قد لا يظهر الفروق الحقيقية بينهن وبين الطالبات منخفضات المستوى؛ مما قد يقلل من معاملات تمييز المفردات عند عرضها بطريقة تتابعية مع حظر الرجوع.

كما انخفضت معاملات تمييز مفردات الصورة الاختبارية المكونة من (٣٠) مفردة والمعروضة مفرداتها بالطريقة الكلية المطبقة على المجموعة (أ) بفرق دال إحصائيًا عن

الصورة المطبقة على المجموعة (و) المكونة من (٤٠) مفردة بطريقة تتابعية بدون حظر الرجوع، وقد يعزى ذلك إلى أنه في حالة عرض المفردات بطريقة كلية يحدث نوعًا من التشتت الذهني لدى بعض الطالبات نتيجة عرض أكثر من مثير في الشاشة الواحدة، وهذا قد يؤدي إلى نسيان الإجابة عن بعض المفردات أو تشتت الانتباه عند الإجابة، وذلك لدى الطالبات بشكل عام؛ مما يضعف من قدرة مفردات الاختبار في المتوسط العام على إظهار الفروق بين الطالبات مرتفعات ومنخفضات المستوى، وفي ظل نقص عدد المفردات في الصورة المعروضة مفرداتها بالطريقة الكلية عن عددها في الصورة المعروضة مفرداتها بالطريقة التتابعية بدون حظر الرجوع؛ فإن فرصة الطالبات مرتفعات المستوى في التميز عن منخفضات المستوى تكون أقل في ظل ما قد ينتاب الفئتين من تشتت انتباه بسبب المثيرات المعروضة أمامهن عند عرض المفردات بطريقة كلية.

التحقق من الفرض الخامس:

ينص الفرض الخامس على: لا يختلف ترتيب معاملات ثبات الاختبار الإلكتروني للصور الاختبارية المختلفة وفقًا لطريقة عرضها وعددها، وللتحقق من هذا الفرض تم حساب ثبات كل صورة من الصور الاختبارية المختلفة في عدد مفرداتها وطريقة عرضها باستخدام معادلة كيودر ريتشاردسون (٢١)، كما في جدول (٢٢)

جدول (٢٢)

ترتيب معاملات ثبات الصور الاختبارية المختلفة في عدد مفرداتها وطريقة عرضها

المجموعة	الصورة الاختبارية	معامل الثبات
و	المكونة من (٤٠) مفردة، مع العرض التتابعي للمفردات بدون حظر بالرجوع	٠.٨٨٤
ج	المكونة من (٣٠) مفردة، مع العرض التتابعي للمفردات بدون حظر الرجوع	٠.٨٧١
د	المكونة من (٤٠) مفردة، مع العرض الكلي للمفردات	٠.٨٦٩
أ	المكونة من (٣٠) مفردة، مع العرض الكلي للمفردات	٠.٨٤٢
هـ	المكونة من (٤٠) مفردة، مع العرض التتابعي للمفردات مع حظر الرجوع	٠.٨١٣
ب	المكونة من (٣٠) مفردة، مع العرض التتابعي للمفردات مع حظر الرجوع	٠.٧٦٩

ويتضح من جدول (٢٢) أن قيم معاملات الثبات تدل على تمتع جميع الصور الاختبارية بدرجة ثبات مقبولة، كما يتضح اختلاف قيم معاملات ثبات الصور الاختبارية باختلاف عدد مفرداتها وطريقة عرضها، حيث كان أعلى قيمة للثبات لصالح الصورتين الاختباريتين المكونتين من (٤٠، ٣٠) مفردة، والمعروضتان مفرداتهما تتابعيًا بدون حظر بالرجوع، وتلتهما الصورتان الاختباريتان المكونتان من (٤٠، ٣٠) مفردة، والمعروضتان مفرداتهما بطريقة كلية، وفي الترتيبين الأخيرين جاءت الصورتان الاختباريتان المكونتان من (٤٠، ٣٠) مفردة، والمعروضتان مفرداتهما تتابعيًا بدون حظر بالرجوع، وجاءت الصورتان الاختباريتان المكونتان من (٤٠، ٣٠) مفردة، والمعروضتان مفرداتهما تتابعيًا مع حظر بالرجوع في المرتبتان الأخيرتان.

ويتضح من ذلك أن ثبات الاختبار الإلكتروني لا يزداد بزيادة عدد مفرداته بشكل مطلق، وإنما يكون ذلك مشروطًا بطريقة عرض هذه المفردات، ويمكن تفسير ذلك بأن زيادة طول الاختبار تؤدي إلى زيادة قدرته على الكشف عن التباين الحقيقي بين الطالبات، حيث تزداد الفروق بين درجات الطالبات على مجموع المفردات، ولكن هذا الأمر مشروطًا بضرورة استخدام طريقة عرض للمفردات يمكن من خلالها أيضًا الكشف عن الفروق بين استجابات الطالبات على الاختبار بشكل أفضل، وقد كانت طريقة عرض المفردات تتابعيا بدون حظر الرجوع أفضل الطرق للحصول على معامل ثبات أعلى؛ نظرًا لأن هذه الطريقة تيسر على الطالبات الأداء على الاختبار بشكل عام بدون الأخطاء العشوائية المتعددة في القياس، تتمثل في التسرع في الإجابة عن المفردات الصعبة وعدم تأجيل الإجابة عنها لنهاية الاختبار حتى يتم التفكير فيها بشكل أفضل وارتياحية أكبر، والخوف من انتهاء وقت الاختبار بدون استكمال الإجابة عن باقي المفردات، وعدم القدرة على توزيع وقت الاختبار بشكل مناسب؛ وما قد ينتج عن هذا كله من ممارسة التخمين العشوائي في الإجابة، أما عند عرض المفردات تتابعيًا بدون حظر الرجوع، فإن الإجابة عنها لا تتأثر بخطأ القياس العشوائي الناتج عن تشتت الانتباه في الإجابة عن مفردات الاختبار المعروضة بطريقة كلية في شاشة واحدة، والناتج عن كثرة المثيرات وزيادتها عن العبء المعرفي للطالبة أثناء الإجابة، ومن ثم فإن أقل طرق عرض المفردات تعرضًا لأخطاء القياس العشوائية هي الطريقة التتابعية بدون حظر الرجوع، تليها الطريقة الكلية ثم الطريقة التتابعية مع حظر الرجوع في الترتيب

الأخير، وتختلف هذه النتيجة مع النتيجة التي توصلت إليها دراسة كاليجارو وشاندلوبز ودينيز (Callegaro, Shand-Lubbers & Dennis, 2009)، ومع النتيجة التي توصلت إليها نتيجة دراسة مافليتوفا وكوبر وليبيديف (Mavletova, Couper & Lebedev, 2018)، بينما تتفق جزئيًا مع نتيجة دراسة محمود (٢٠١٨).

التحقق من الفرض السادس:

ينص الفرض السادس على: لا توجد فروق دالة إحصائيًا في متوسطات درجات الطلبة على الاختبار الإلكتروني التي تعزى إلى:

٦- أ) طريقة عرضها مع تثبيت عددها.

٦- ب) عددها مع تثبيت طريقة عرضها.

٦- ج) اختلاف طريقة عرضها وعددها.

وللتحقق من هذا الفرض تم تحويل سقف درجات الطالبات على الصور الاختبارية المختلفة في عدد مفرداتها وطريقة عرضها إلى الدرجة (١٠٠)؛ لكي يسهل المقارنة بينها في ظل اختلاف سقف الدرجة الكلية بين الصورتين الاختباريتين المختلفتان في العدد، وكانت الإحصاءات الوصفية لهذه الدرجات المحولة، وكانت النتائج كما في جدول (٢٣).

جدول (٢٣)

الإحصاءات الوصفية للدرجات المحولة للطالبات على الصور الاختبارية المختلفة

طريقة العرض	الكلية		التتابعية مع حظر الرجوع		التتابعية بدون حظر الرجوع	
عدد المفردات	٣٠ مفردة	٤٠ مفردة	٣٠ مفردة	٤٠ مفردة	٣٠ مفردة	٤٠ مفردة
المتوسط	٦٩.٧٤	٦٩.٣٤	٦١.٥٤	٦١.٥٣	٧٢.٦٩	٧٢.٦٧
الانحراف المعياري	١١.٦٥	١١.٢٦	١١.٠٥	١١.٠٧	١١.٤٤	١١.٤٦

(أ) بالنسبة لتأثير طريقة عرض المفردات في أداء الطالبات مع تثبيت عددها:

تمت المقارنة بين الدرجات المحولة في الثلاث مجموعات (أ، ب، ج) التي تعرضت للصورة الاختبارية الأولى المكونة من (٣٠) مفردة باستخدام تحليل التباين الأحادي، كما تم نفس الإجراء مع الثلاث مجموعات (د، هـ، و) التي تعرضت للصورة الاختبارية الثانية المكونة من (٤٠) مفردة، كما في جدول (٢٤).

جدول (٢٤)

تحليل التباين الأحادي للمقارنة بين متوسطات الدرجات للطالبات باختلاف طرق العرض

الاختبار	مصدر التباين	مجموع المربعات	درجة الحرية	متوسط المربعات	قيمة ف	مستوى الدلالة	مربع ايتا	حجم الأثر
الاختبار المكون من (٣٠) مفردة	طريقة العرض	٤٩٩٦.٩٥٥	٢	٢٤٩٨.٤٧٧	١٩.٢٨٩	٠.٠٠٠	٠.١٤٨	كبير
	الخطأ	٢٨٧٥٤.٦٧٩	٢٢٢	١٢٩.٥٢٦				
	الكلي	٣٣٧٥١.٦٣٤	٢٢٤					
الاختبار المكون من (٤٠) مفردة	طريقة العرض	٤٨٩٩.٤٢٠	٢	٢٤٤٩.٧١٠	١٩.٢٩٩	٠.٠٠٠	٠.١٤٨	كبير
	الخطأ	٢٨١٧٩	٢٢٢	١٢٦.٩٣٢				
	الكلي	٣٣٠٧٨.٤٢	٢٢٤					

ويتضح من خلال جدول (٢٤) وجود فرق دال إحصائيًا بين متوسطات معاملات صعوبة المفردات باختلاف طرق عرض المفردات، سواء في حالة الصورة الاختبارية المكونة من (٣٠ أو ٤٠) مفردة، ولمعرفة اتجاه الفرق تم إجراء مقارنة بعدية باستخدام اختبار شيفيه للمقارنة بين كل مجموعتين تعرضتا لنفس عدد المفردات، وكانت النتائج كما في جدول (٢٥).

جدول (٢٥)

نتائج اختبار شيفيه بين متوسطي الدرجات على كل صورتين اختباريتين لهما نفس الطول

	عدد المفردات	٣٠ مفردة		٤٠ مفردة	
طريقة العرض		الكلية	التتابعية مع حظر الرجوع	الكلية	التتابعية مع حظر الرجوع
التتابعية مع حظر الرجوع	الفرق بين المتوسطين	٨.١٩٠		٧.٨٠٨	
	مستوى الدلالة	*٠.٠٠٠		*٠.٠٠٠	
التتابعية بدون حظر الرجوع	الفرق بين المتوسطين	-٢.٩٥	-١١.١٤	-٣.٣٢٥	-١١.١٣٣
	مستوى الدلالة	٠.٢٨٦	*٠.٠٠٠	٠.١٩٨	*٠.٠٠٠

ويتضح من خلال جدول (٢٥) وجود ارتفاع بفارق دال إحصائيًا لدرجات الطالبات عند عرضها بالطريقة الكلية أو عند عرضها بالطريقة التتابعية بدون حظر الرجوع عن درجاتهن في حال عرضها بالطريقة التتابعية مع حظر الرجوع في الحالتين سواء عندما كان عدد المفردات (٣٠ أو ٤٠) مفردة؛ وقد يرجع ذلك إلى أن الطالبات اللائي تعرض عليهن المفردات تتابعيًا مع حظر الرجوع قد يزداد لهن القلق أثناء الإجابة عن المفردات، نظرًا لأنهن يشعرن بعدم وجود فرصة أخرى للتفكير في الإجابة، وقد يؤدي ذلك إلى تسرعهن في الإجابة خوفًا من انتهاء الوقت، كما أن حظر الرجوع للمفردات لا يساعدهن في الاستفادة من بعض المفاتيح أو المنبهات التي يجدونها في أثناء الإجابة عن مفردات أخرى وتذكرهن بالإجابة الصحيحة عن مفردات سابقة، وهذا كله لا يحدث عند عرض المفردات بطريقة كلية في شاشة واحدة، أو عند عرضها تتابعيا بدون حظر الرجوع، وتتفق هذه النتيجة مع نتيجة دراسة ريفيلا وتونينيلي واوشوا (Revilla, Toninelli & Ochoa, 2016)، ومع نتيجة دراسة عبد الوكيل (٢٠٢١)، ولكنها تختلف مع نتائج دراسات كل من (حميد، ٢٠٢١؛ Wang, Kao & Chen, 2021).

(ب) بالنسبة لتأثير عدد المفردات في أداء الطالبات مع تثبيت طريقة عرضها:

تم استخدام اختبار "ت" لمقارنة متوسطي درجات الطالبات المحولة على كل صورتين باختلاف عدد مفرداتهما في حالات العرض الثلاث، كما في جدول (٢٦).

جدول (٢٦)

نتائج اختبار "ت" للمقارنة بين متوسطي الدرجات المحولة في الصورتين المختلفتين في عدد مفرداتهما

طريقة العرض	الصورة الاختبارية	المتوسط	الانحراف المعياري	قيمة ت	مستوى الدلالة	الدلالة
الكلية	المكونة من ٣٠ مفردة	٦٩.٣٤٢	١١.٦٤٨	٠.٢١١	٠.٨٣٣	غير دالة
	المكونة من ٤٠ مفردة	٦٩.٧٣٧	١١.٢٦٤			
التتابعية مع حظر الرجوع	المكونة من ٣٠ مفردة	٦١.٥٤٧	١١.٠٤٨	٠.٠٠٧	٠.٩٩٤	غير دالة
	المكونة من ٤٠ مفردة	٦١.٥٣٣	١١.٠٧٣			
التتابعية بدون حظر الرجوع	المكونة من ٣٠ مفردة	٧٢.٦٨٧	١١.٤٣٨	٠.٠١١	٠.٩٩١	غير دالة
	المكونة من ٤٠ مفردة	٧٢.٦٦٧	١١.٤٥٩			

ويتضح من جدول (٢٦) عدم وجود فرق دال إحصائيًا بين درجات الطالبات على الصورتين المختلفتين في عدد مفرداتهما (٣٠ ، ٤٠ مفردة) سواء عند عرضها بطريقة كلية أو بطريقة تتابعية مع حظر الرجوع أو بطريقة تتابعية بدون حظر الرجوع، أي أن طول الاختبار لم يؤثر في أداء الطالبات على الاختبار، وقد يعود ذلك إلى أن تم وضع مفردات كلا الصورتين الاختباريتين وفق جدول مواصفات واحد، يحتوي على نفس النسب الوزنية لموضوعات المحتوى والمستويات المعرفية للأهداف، ومن ثم فإن الزيادة في طول الصورة الاختبارية الثانية تمت بشكل متوازن مع باقي المفردات التي استخدمت في الصورة الاختبارية الأولى، وبالتالي كانت متسقة في قياسها للتحصيل الدراسي وفق نفس المستويات المعرفية للأهداف ووفق نفس الموضوعات، ومن ثم فلم تكن هناك اختلافات بين الطالبات في أدائهن على الصورتين المختلفتين في الطول في ظل تثبيت طريقة عرض المفردات وأن الطالبات تم توزيعهن في المجموعات المختلفة بطريقة عشوائية تحقق تكافؤهن في المعدل التراكمي، وتتعارض هذه النتيجة جزئيًا مع ما أشار إليه كوبر ورفاقه (**Couper, et al., 2013**) من أن الطريقة الكلية تحول دون استكمال المستجيبين الجهد المعرفي المبذول أثناء الاستجابة مع زيادة عدد المفردات المعروضة في الشاشة.

(ج) بالنسبة لتأثير اختلاف طريقة عرض المفردات وعددها في أداء الطالبات:

تمت مقارنة متوسطات درجات الطالبات المحولة في حالة اختلاف طريقة عرض المفردات وعددها باستخدام تحليل التباين الأحادي، كما في جدول (٢٧).

جدول (٢٧)

نتائج تحليل التباين الأحادي بين متوسطات درجات الطالبات باختلاف طرق عرض المفردات وعددها

مصدر التباين	مجموع المربعات	درجة الحرية	متوسط المربعات	قيمة ف	مستوى الدلالة	مربع ايتا	حجم الأثر
طريقة عرض المفردات وعددها	٩٨٩٨.٦٦٩	٥	١٩٧٩.٧٣٤	١٥.٤٣٩	٠.٠٠٠	٠.١٤٨	كبير
الخطأ	٥٦٩٣٣.٦٧٩	٤٤٤	١٢٨.٢٢٩				
الكلي	٦٦٨٣٢.٣٤٧	٤٤٩					

ويتضح من جدول (٢٧) وجود فرق دال إحصائيًا بين متوسطات درجات الطالبات المحولة في الصور الاختبارية المختلفة في عددها وطريقة عرضها، ولمعرفة اتجاه هذا الفرق، تم إجراء مقارنة بعدية باستخدام اختبار شيفيه للمقارنة بين كل صورتين معًا، كما في جدول (٢٨).

جدول (٢٨)

نتائج اختبار شيفيه بين متوسطي درجات الطالبات المحولة بين كل صورتين من الصور الاختبارية

		أ	ب	ج	د	هـ
ب	الفرق بين المتوسطين	٨.١٩٠				
	مستوى الدلالة	٠.٠٠٢				
ج	الفرق بين المتوسطين	٢.٩٥٠-	١١.١٤٠-			
	مستوى الدلالة	٠.٧٧٠	*٠.٠٠٠			
د	الفرق بين المتوسطين	٠.٣٩٠	٧.٧٩٥-	٣.٣٤٥		
	مستوى الدلالة	١	*٠.٠٠٤	٦٥٨		
هـ	الفرق بين المتوسطين	٨.٢٠٣	٠.٠١٣	١١.١٥٣	٧.٨٠٨	
	مستوى الدلالة	٠.٠٠٢	١	*٠.٠٠٠	*٠.٠٠٤	
و	الفرق بين المتوسطين	٢.٩٣-	١١.١٢٠-	٠.٠٢٠-	٣.٣٢٥-	١١.١٣٣-
	مستوى الدلالة	٠.٧٧٥	*٠.٠٠٠	١	٠.٦٦٤	*٠.٠٠٠

ويتضح من جدول (٢٨) انخفاض درجات الطالبات على الصورتين الاختباريتين المعروضة مفرداتهما بطريقة تتابعية مع حظر الرجوع بفرق دال إحصائيًا عن درجات الطالبات على باقي الصور الاختبارية المعروضة مفرداتها كليًا أو تتابعيًا بدون حظر الرجوع، وقد يعود ذلك إلى أنه بغض النظر عن طول الاختبار فإنه عند حظر الطالبات من الرجوع إلى المفردات التي أجبن عنها ولا يستطعن العودة إليها مرة أخرى في نهاية الاختبار قد يزيد من قلقهن وتوترهن أثناء الإجابة عن المفردات وقد يؤدي إلى تسرعهن في الإجابة عن المفردات، وهذا يؤثر سلبا في درجاتهن في الاختبار، وعلى النقيض فإن الطالبات اللائي عرضت عليهن المفردات بطريقة تتابعية بدون حظر الرجوع أو كليًا كان لديهن فرصة أخرى لإعادة التفكير في إجابة المفردات التي صعب عليهن الإجابة عنها ومن ثم إمعان التركيز في الإجابة، كما يمكنهم الاستفادة من المفاتيح أو المنبهات أو المعلومات التي

يتذكرونها عند إجابة بعض المفردات في الإجابة عن بعض المفردات الأخرى، كما يمكنهن تنظيم وقت الاختبار بشكل أفضل، كما أن العرض الكلي لمفردات الاختبار الإلكتروني يسمح للطالبات بالإجابة عن المفردات وفق الترتيب الذي يتوافق معهن، سواء حسب صعوبة المفردات أو حسب تسلسل عرضها في المحتوى الدراسي أو حسب الدرجات المخصصة لها، واتفقت هذه النتائج رغم اختلاف طول الاختبار، مما يدل على أن تأثير طريقة عرض المفردات أكبر من تأثير طول الاختبار في أداء الطالبات على الاختبار.

توصيات البحث: من خلال نتائج البحث الحالي يوصي الباحث بما يلي:

- تجنب استخدام خيار حظر الرجوع عند تصميم الاختبارات الإلكترونية لما قد يكون له من تأثير سلبي على الخصائص السيكومترية للاختبار وأداء الطلبة عليه.
- إمكانية زيادة عدد مفردات الاختبار الإلكتروني بما يحقق جدول المواصفات ولا يؤدي إلى إرهاق الطلبة في الإجابة عند استخدام الطريقة التتابعية بدون حظر الرجوع في عرض المفردات، وتخفيف عدد المفردات عند عرض المفردات بالطريقتين الكلية أو التتابعية مع حظر الرجوع.
- مطالبة المتخصصين في تقنية المعلومات بتطوير المنصات الإلكترونية التي يتم عرض الاختبارات الإلكترونية فيها بحيث يسمح للمختبر أن يغير طريقة عرض المفردات بين الطريقتين الكلية والتتابعية بدون حظر الرجوع، بما يناسب قدراته وإمكاناته وتفضيلاته الشخصية.

البحوث المقترحة: يمكن اقتراح بعض البحوث في ضوء البحث الحالي، كما يلي:

- أثر استخدام خيار حظر الرجوع في الخصائص السيكومترية للاختبار الإلكتروني والأداء عليه لدى المتروين والمندفعين معرفيًا.
- أثر طريقة عرض المفردات وعددها في دقة تقدير بارامترات الأفراد والمفردات في ضوء نظرية الاستجابة للفقرة.
- تحليل بعدي لنتائج البحوث التي تناولت العوامل المؤثرة في الخصائص السيكومترية للاختبار الإلكتروني.
- أثر طريقة عرض بدائل المفردات (أفقية – رأسية) في الطريقتين الكلية والتتابعية على الخصائص السيكومترية للاختبار الإلكتروني والأداء عليه.

قائمة المراجع

اولاً المراجع العربية :

أبو الشيخ، عطية إسماعيل (٢٠١٨). قلق الاختبارات الإلكترونية وعلاقته بالأداء في نظر عينة من طالبات كلية الأميرة عالية بجامعة البلقاء التطبيقية – الأردن، *المجلة التربوية لكلية التربية بجامعة سوهاج*، ٥٢ (٥٢)، ٧٩٩ – ٨٢٣.

البربري، رفيق سعيد (٢٠٢٠). نمطا تصميم الاختبار الإلكتروني التكيفي الثابت والمتغير الطول وأثرهما على خفض مستوى قلق الاختبار وتنمية الاتجاهات نحو الاختبارات الإلكترونية لدى طلاب كلية التربية. . *تكنولوجيا التعليم: سلسلة دراسات وبحوث محكمة (الجمعـية المصرية لتكنولوجيا التعليم)*، ٣٠ (١)، ٢٣ – ٨٧.

جليلة، محمود محسوب (٢٠١٧). أثر اختلاف متغيرات تصميم الاختبار الإلكتروني على الدافعية للإنجاز الدراسي لدى تلاميذ المرحلة المتوسط. *دراسات عربية في التربية وعلم النفس*، ٨٣، ٥٢٥ – ٥٦٠.

حميد، عبد الرحمن أحمد (٢٠٢٠). أثر التفاعل في الاختبارات الإلكترونية بين ترتيب الأسئلة متدرجا / عشوائيًا ونمط عرضها سؤالًا واحدا في الشاشة / الاختبار كله في الشاشة في التحصيل والأداء المهاري والاتجاه نحوها لدى طلاب الدراسات العليا بجامعة القصيم. *مجلة البحث العلمي في التربية (جامعة عين شمس)*، ٥ (٢١)، ٦٠٣ – ٦٩٥.

الخزي، فهد (٢٠١٣). أثر قلق الاختبارات الإلكتروني على الأداء في الاختبارات الإلكترونية. *المجلة الدولية للأبحاث التربوية (جامعة الإمارات العربية المتحدة)*، ٣٣، ١ – ٢٩.

خميس، محمد عطية (٢٠٠٣). *منتوجات تكنولوجيا التعليم*. القاهرة: دار الحكمة.

دلالعة، أسامة؛ دلالعة، طارق (٢٠١٩). معايير الجودة في تصميم الاختبارات الإلكترونية ودرجة توفرها في اختبارات المستوى في الجامعات الأردنية. *مجلة المنارة للبحوث والدراسات*، ٢٥ (١)، ٤٥ – ٩٦.

الزامل، رنا بنت محمد؛ الحجيلان، محمد بن إبراهيم (٢٠١٦). العوامل المؤثرة في اتجاه معلمات الحاسب الآلي حول تطبيق الاختبارات الإلكترونية في مادة الحاسب الآلي وتقنية المعلومات بمدينة الرياض. *المجلة التربوية الدولية المتخصصة*، ٥ (١٠)، ١٩٣ – ٢٠٧.

السيد، محمد حمدي؛ حسن، أمل حسان (٢٠٢١). التفاعل بين نمط عرض الاختبارات الإلكترونية "كلي / تتابعي" ومستوى قلق الاختبار وأثره في تنمية الاحتفاظ بالتعلم ودافعية الإنجاز وخفض الضغوط النفسية لدى طلاب تكنولوجيا التعليم. *تكنولوجيا التعليم: سلسلة دراسات وبحوث محكمة (الجمعية المصرية لتكنولوجيا التعليم)*، ٣١ (٢)، ٣ – ٩٠.

صفوت، إيناس محمد ؛ صبري، نصر محمود (٢٠٢٠). التحكم الانتباهي كمتغير وسيط بين التفكير السلبي وقلق الاختبار باستخدام التابلت لدى طلبة الصف الأول الثانوي العام. *دراسات تربوية ونفسية*، ٣٥ (١٠٨)، ٣٩ – ١١٤.

العاجز، عبد الحليم إبراهيم (٢٠١٩). مقارنة الخصائص السيكومترية لاختبار المصفوفات المتتابعة العادية لرافن ورقيا وإلكترونيا لدى طلاب كلية التربية. *مجلة كلية التربية بالمنصورة*، ١٠٨ (٥)، ١٣٢٧ – ١٣٦٥.

عبد الحميد، محمد (٢٠٠٥). *منظومة التعلم عبر الشبكات*. القاهرة: عالم الكتب.

عبد الرحيم، المعتز بالله زين الدين؛ مرعي، السيد محمد (٢٠٢٠). فاعلية برنامج يرتكز على نظرية تجهيز المعلومات في تنمية مهارات تصميم الاختبارات الإلكترونية في العلوم ومستوى الطموح العلمي لدى طلاب شعبة الكيمياء والفيزياء بكلية التربية جامعة الأزهر. *مجلة البحث العلمي في التربية بعين شمس*، ٢١ (١٠)، ٥٧٠ – ٦٠٥.

عبيس، محمد فرحان (٢٠٢٠). تقييم برنامج ريسبوندس Respondus للمراقبة الذكية في الاختبارات الإلكترونية من وجهة نظر أساتذة كلية الإمام الكاظم "ع" – أقسام بابل. *مجلة الفنون والأدب وعلوم الإنسانيات والاجتماع (كلية الإمارات للعلوم التربوية)*، ٦٠، ٣٥٣ – ٣٣٩.

علام، صلاح الدين محمود (٢٠٠٤). *القياس والتقويم التربوي والنفسي*. القاهرة: دار الفكر العربي للنشر والتوزيع.

عيسوي، عبدالرحمن محمد (١٩٩٩). *القياس والتجريب في علم النفس والتربية*. الإسكندرية: دار المعرفة الجامعية.

الفيل، حلمي محمد (٢٠١٤). مبادئ تصميم التعلم الإلكتروني المشتقة من نظرية العبء المعرفي. *مجلة بحوث في العلوم والفنون النوعية، ١* (١)، ٤ – ٤٢.

مبارز، منال عبد العال؛ فخري، أحمد محمود (٢٠١٣). *التعليم الإلكتروني: مفهومه – بيئاته – مقرراته – إدارته – تقويمه – تطبيقاته المتقدمة*. الرياض: دار الزهراء للنشر والتوزيع.

محمود، سومية شكري(٢٠١٨). مقارنة الخصائص السيكومترية بين الاختبارات التحصيلية الإلكترونية والورقية. *مجلة كلية التربية بأسيوط، ٣٤*(٦)، ١١٦– ١٤٦.

المرزوق، محمد حسن (٢٠٢٠). *الاختبارات الإلكترونية: خصائصها ومزاياها*. ورقة عمل في المؤتمر الدولي (الافتراضي) لمستقبل التعليم الرقمي في الوطن العربي، إثراء المعرفة للمؤتمرات والأبحاث بمدينة الطائف خلال الفترة من ٣٠ أكتوبر – ٢ نوفمبر، ص ص: ٢٧٠ – ٢٧٨.

المقبل، مها ناصر (٢٠٢٠). اتجاهات معلمات العربية لغة ثانية نحو الاختبارات الإلكترونية ومعوقات التطبيق بمعهد اللغويات العربية بجامعة الملك سعود. *مجلة كلية التربية (جامعة كفر الشيخ)*، ٢٠ (٤)، ٣٣١ – ٣٥٢.

المومني، جهاد علي؛ النصراوين، معين سليمان (٢٠٢٠). أثر استخدام الاختبار الإلكتروني بالباركود عبر الهاتف الذكي في تحصيل طلبة الدراسات العليا في كلية العلوم التربوية بجامعة عمان العربية واتجاهاتهم نحوه. *مجلة البلقاء للبحوث والدراسات (جامعة عمان الأهلية)*، ٢٣ (عدد خاص)، ٥٧ – ٦٧.

نور الدين، أمين محمد صبري (٢٠٠٣). فاعلية استخدام الاختبار الموائم المحوسب. *مجلة العلوم الإجتماعية، ٣١* (٣)، ٧٧٩ – ٧٨٣.

اليامي، أميرة بنت صغير (٢٠٢٠). *مدى فعالية التقويم الإلكتروني للطالبات من وجهة نظر المعلمات في المرحلة المتوسطة والثانوية بمدينة نجران*. ورقة عمل في المؤتمر الدولي (الافتراضي) لمستقبل التعليم الرقمي في الوطن العربي، إثراء المعرفة للمؤتمرات والأبحاث بمدينة الطائف خلال الفترة من ٣٠ أكتوبر – ٢ نوفمبر، ص ص: ٣٠٨ – ٣٢١.

الوكيل، محمد أبو الليل (٢٠٢١). أثر التفاعل في الاختبارات الإلكترونية بين نمط تقديم الأسئلة (حظر / سماح) بالرجوع وعرض النتائج للطلاب (فوري / مرجأ) على تنمية التحصيل المعرفي ومستوى قلق الاختبار لدى طلبة كلية المجتمع. مجلة البحوث في مجالات التربية النوعية، ٧ (٣٤)، ١٧٧١ – ١٨٦١.

Aiken,L.(2000). *Psychological Testing and Assessment* (10th Ed.). Allyn & Bacon, USA.

Alruwais, N., Wills, G & Wald, M. (2018). Advantages and Challenges of Using e-Assessment. International Journal of Information and Education. *Technology*,*8*,1, 34-37.

Bennett, R. E. (2002). Using Electronic Assessment To Measure Student Performance: Online Testing. *State Education Standard, 3* (3), 23-29.

Berman, A. I., Haertel, E. H., & Pellegrino, J. W. (2020). *Comparability of Large-Scale Educational Assessments: Issues and Recommendations*. Washington, DC: National Academy of Education

Bridgeman, B., Lennon, M. L., & Jackenthal, A. (2001). *Effects of screen size, screen resolution, and display rate on computer-based test performance. (ETS RR-01-23). Princeton NJ: Educational Testing Service.* http://www.ets.org/Media/Research/pdf/RR-01-23-Bridgeman.pdf

Callegaro, M., Shand-Lubbers, J., Dennis, J. M. (2009). *Presentation of a single item versus a grid: Effects on the vitality and*

Mental Health Scales of the SF-36v2 health survey. Presented at the AAPOR, Knowledge Networks.

http://www.amstat.org/sections/srms/Proceedings/y2009/Files/400045.pdf

Conejo, R., Guzmán, E., Millán, E., Trella, M., Pérez-De-La-Cruz, J. L., & Ríos, A. (2004). SIETTE: A web-based tool for adaptive testing. *International Journal of Artificial Intelligence in Education, 14*(1), 29-61.

Couper, M. P., Tourangeau, R., Conrad, F. G., & Zhang, C. (2013). The Design of Grids in Web Surveys. *Social science computer review*, *31*(3), 322–345. https://doi.org/10.1177/0894439312469865

Debell, M.; Wilson, C.; Jackman,S. & Figueroa,L. (2021). Optimal Response Formats for Online Surveys: Branch, Grid, or Single Item?. *Journal of Survey Statistics and Methodology*, *9* (1), 1 – 24, https://doi.org/10.1093/jssam/smz039

Erdogan, Y. (2008). Paper-based and computer-based concept mapping: The effects on computer achievement, computer anxiety and computer attitude. *British Journal of Educational Technology*, *40*(5), 821-836.

Jodoin, M. G. (2003). *Psychometric properties of several computer-based test designs with ideal and constrained item pools*. [Doctoral Dissertations] 1896 - Department of Psychology, University of Massachusetts Amherst, ScholarWorks@UMass Amherst. February 2014. 3299. https://scholarworks.umass.edu/dissertations_1/3299.

Jodoin, M. G. ; Zenisky, A. L. & Hambleton, R. K. (2006). Comparison of the Psychometric Properties of Several Computer-Based Test Designs for Credentialing Exams with Multiple Purposes. *Applied Measurement in Education 19* (3), 203-220.

Kingston, N. M. (2009). Comparability of Computer- and Paper-Administered Multiple-Choice Tests for K-12 Populations: A Synthesis. *Applied Measurement in Education, 22* (1), 22-37.

Klausch, T., de Leeuw, E. D., Hox, J. J., Roberts, A., de Jongh, A. (2012). *Matrix vs. single question formats in web surveys: Results from a large scale experiment*. Paper presented at the General Online Research conference(GOR), Mannheim, Germany, 5–7 March

Krille C., Wuttke E., Seifried J. (2017) Mode Effects in Correcting Students' Errors: A Comparison of Computer-Based and Paper-Pencil Tests. In: Wuttke E., Seifried J. (eds) *Professional Error Competence of Preservice Teachers*. SpringerBriefs in Education. Springer, Cham. https://doi.org/10.1007/978-3-319-52649-2_4

Kuikkaa, M., Kitolab, M., & Laaksob, M. J. (2014). Challenges when introducing electronic exam. *Research in Learning Technology*, 22, 22817. http://doi.org/10.3402/rlt.v22.22817

Leeson, H.V. (2006). The mode effect: A literature review of human and technological issues in computerized testing. *International Journal of Testing*, *6*, 1–24

Liu, M., & Cerant, A. (2018). Item-by-item Versus Matrix Questions: A Web Survey Experiment. *Social Science Computer Review, 36*(3), 690-706.

Mavletova, A., Couper, M. P., & Lebedev, D. (2018). Grid and Item-by-Item Formats in PC and Mobile Web Surveys. *Social Science Computer Review*, *36*(6), 647–668. https://doi.org/10.1177/0894439317735307

Meijer, R. R., & Nering, M. L. (1999). Computerized Adaptive Testing: Overview and Introduction. *Applied Psychological Measurement, 23* (3), 187– 194. https://doi.org/10.1177/01466219922031310

Nicholson, A. M. (2009). *Effects of test anxiety on student achievement (ACT) for college bound students*. Dissertation Abstract International. DAI-A-70/07, AAT 3366126

Parshall, C.G.; Spray, J.A.; Kalohn, J.C.; Davey, T. (2020). *Practical Considerations in Computer-Based Testing*; Springer: New York, NY, USA, 2020; Available online: https://link.springer.com/book/10.1007%2F978-1-4613-0083-0

Rana, R., & Mahmood, N. (2010). The relationship between test anxiety and academic achievement. *Bulletin of Education and Research, 32*(2), 63-74.

Revilla, M., Toninelli, D., Ochoa, C. (2016) An experiment comparing grids and item-by-item formats in web surveys completed through PCs and smartphones. *Telematics and Informatics*, 34, 30–42.

Schurig, M. ; Jungjohann, J. & Gebhardt, M. (2021). Minimization of a Short Computer-Based Test in Reading. *Frontiers in Education*, 6, 220-233, https://www.frontiersin.org/article/10.3389/feduc.2021.684595

Smolinsky, L., Marx, B. D., Olafsson, G., & Ma, Y. A. (2020). Computer-Based and Paper-and-Pencil Tests: A Study in Calculus for STEM Majors. *Journal of Educational Computing Research, 58*(7), 1256–1278.

The International Test Commission (2006). International Guidelines on Computer-Based and Internet Delivered Testing. *International Journal of Testing, 6* (2), 143-171.

Thurlow, M., Lazarus, S. S., Albus, D., & Hodgson, J. (2010). Computer-based testing: Practices and considerations (Synthesis Report 78). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes

Toepoel, V., Das, M., Soest, A. V. (2009). Design of web questionnaires: The effects of the number of items per screen. Field Methods, 21, 200–213. Retrieved from http://doi.org/10.1177/1525822X08330261

Wang, H. & Shin, C. D. (2010). Comparability of computerized adaptive and paper-pencil tests. *Measurement and Research Service Bulletin*, 13, 1 – 7.

Wang, T.H.; Kao, C.H. & Chen, H. C. (2021). Factors Associated with the Equivalence of the Scores of Computer-Based Test and Paper-and-Pencil Test: Presentation Type, Item Difficulty and Administration Order. *Sustainability, 13*(17) , 48 – 95 https://doi.org/10.3390/su13179548

Weir, I., Gwynllyw, R., & Henderson, K. (2021). A case study in the e-assessment of statistics for nonspecialists. *Journal of University Teaching & Learning Practice*, *18*(2). https://ro.uow.edu.au/jutlp/vol18/ iss2/05

Weiss, D. J. (2013). Item banking, test development, and test delivery. In K. F. Geisinger (Ed.), *The APA Handbook on Testing and Assessment*. Washington, WA: American Psychological Association.

Wilson, B. & Cole, P. (1996). Cognitive teaching models. In D. h. Jonassen (Ed.), *Handbook of research for educational communications and technology*. New York: Macmillan, 601 – 621.