فاعلية تصميم نظام اختبارات إلكترونية تكيفية على تقييم القدرة المعرفية وتعزيز التحصيل الدراسي

الاستـــلام: 28/فبراير/2025
التحكيم: 16/ مايو/2025
القبـــــول: 17/ مايو /2025

أحمد عبدالله بالحارث [1*]
ريم عبد عمر القامزي[1]

© 2025 University of Science and Technology, Aden, Yemen. This article can be distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

© 2025 جامعة العلوم والتكنولوجيا، المركز الرئيس عدن، اليمن. يمكن إعادة استخدام المادة المنشورة حسب رخصة مؤسسة المشاع الإبداعي شريطة الاستشهاد بالمؤلف والمجلة.

[1] قسم الحاسوب، كلية التربية، جامعة عدن، عدن، اليمن.
* عنوان المراسلة: balharethru@gmail.com

فاعلية تصميم نظام اختبارات إلكترونية تكيفية على تقييم القدرة المعرفية وتعزيز التحصيل الدراسي

أحمد عبدالله بالحارث
قسم الحاسوب، كلية التربية
جامعة عدن
عدن، اليمن
balharethru@gmail.com

ريم عبد عمر القامزي
قسم الحاسوب، كلية التربية
جامعة عدن
عدن، اليمن
reemabd88@gmail.com

الملخص:

استهدف البحث الحالي دراسة فاعلية تصميم نظام مقترح للاختبارات الإلكترونية التكيفية على تقييم القدرة المعرفية للطلبة ،ودورها في رفع التحصيل الدراسي لدى طلبة الجامعة ، من خلال المقارنة بين الاختبار الإلكتروني التكيفي، و الاختبار الإلكتروني (التقليدي) في تقدير القدرة باستخدام طريقة الأرجحية القصوى وقاعدة إنهاء الاختبار وفقًا لدرجة معيارية تم تحديدها سابقًا، بلغ عدد المشاركين في البحث (200) طالبًا من طلبة كليات جامعة العادل المستوى الأولى للعام الجامعي 2024/2023م. وقد توصلت نتائج الدراسة إلى وجد فروق ذات دلاله إحصائية عند مستوى $\alpha = 0.05$ بين متوسطي الخطأ المعياري لتقدير القدرة بين الاختبار التكيفي والاختبار الخطي باستخدام طريقة الأرجحية القصوى لتقدير القدرة وقاعدة إنهاء الاختبار وفقًا لدرجة معيارية ولصالح الاختبار التكيفي، ووجد فروق ذات دلاله إحصائية عند مستوى $\alpha = 0.05$ بين متوسطي درجات طلبة المجموعة التجريبية التي تم تقييمها باستخدام الاختبار التكيفي والمجموعة الضابطة التي تم تقييمها باستخدام الاختبار الخطي ولصالح الاختبار التكيفي، وتُعد النتيجة السابقة مؤشراً جيدًا على فاعلية الاختبار التكيفي في رفع التحصيل العلمي لدى الطلبة. واخيرًا خرجت الورقة بأهم التوصيات المتعلقة بالموضوع؛ حيث يسعى البحث إلى أن يقدم فكرة عامة لتطبيق التقنية الحديثة في التقييم، ودوروها في تطور أساليب التقويم باستخدام الاختبارات الإلكترونية التكيفية.

الكلمات المفتاحية : الاختبارات الإلكترونية التكيفية، الاختبارات الإلكترونية الخطية، التقييم، القدرة المعرفية، التحصيل الدراسي.

Abstract

The aim of the present study is to examine the effectiveness of designing a proposed adaptive electronic testing system in assessing students' cognitive ability and its role in increasing university students' academic achievement through a comparison between the adaptive electronic tests and the traditional electronic tests in estimating cognitive ability using the maximum likelihood method and the rule of completing the test according to a standard score previously determined. The participants of the study consisted of 200 first-year students from five colleges at Adel University for the academic year 2023/2024. The findings of the study revealed that statistically significant differences at the level of α=0.05 between the average standard error for estimating the ability between the adaptive test and the linear test using the maximum likelihood method. It is also observed that statistically significant differences at the level of $\alpha = 0.05$ between the mean scores of the experimental group, assessed using the adaptive test, and the control group, assessed using the linear test and in favor of the adaptive test, the previous findings is strong evidence of the effectiveness of adaptive test in improving students' academic achievement. Based on the findings, the study proposed some recommendations related to the topic, emphasizing the importance of integrating modern technology in assessment and highlighting the role of adaptive electronic test in advancing evaluation methods in higher education.

Keywords— *Adaptive electronic testing, linear electronic testing, assessment, cognitive ability, academic achievement.*

I. المقدمة

يشهد العصر الحالي تطورًا سريعًا في المعرفة كماً وكيفًا، مما يدفع المؤسسات التعليمية إلى البحث عن أدوات تمكنها من مواكبة هذا النمو المتتالي للمعرفة، لذلك تزايد الاهتمام بدعم بيئات التعلم ببرمجيات الحاسوب وتكنولوجيا المعلومات بشكل عام وبيئات التعلم التكيفية بشكل خاص، وذلك لما تقدمه من عدة مزايا، من أهمها سهولة الوصول إلى المعلومات واحتوائها على بيئة تفاعلية (بصرية وسمعية)، بالإضافة إلى أنها تضيف عوامل التشويق والإثارة إلى بيئة التعلم، مما يحسن الدافعية لدى المتعلمين.

ويرى الخزي والزكري [1] أن استخدام التقنيات الحديثة في التعليم يتوقف قبل عملية التقييم، وكأن التقييم لا يعتبر جزءاً من العملية التعليمية، بل إن التقنيات الحديثة يمكن أن تسهم بشكل فعال في تطوير التقييم، وليس من المنطقي أن يسعى التربويون إلى اكتشاف الاستراتيجيات الإلكترونية الحديثة في التعليم وهم يلتزمون بالطرق المعتادة في تقييم مخرجات ذلك التعليم.

لذا كان لا بد من زيادة الاهتمام ببناء الاختبارات حيث تعد من أهم المقاييس النفسية والتربوية التي تساعد في تقدير وتقييم القدرة المعرفية للطلبة والسعي إلى تطويرها لمساعدة متخذي القرار على التنظيم والتطوير والتحسين حتى يمكن اتخاذ القرارات التربوية بدقة أكبر، ولابد أن تتميز هذه الاختبارات والمقاييس العلمية بالصدق والاستقرار والموضوعية.

لذلك تعد الاختبارات بأنواعها من أهم أساليب التقويم في العملية التعليمية للتأكد من مدى تحقيق الأهداف التعليمية والتربوية المنشودة، ومن مدى التقدم الذي أحرزه المتعلمين.

وقد أكد ما سبق الرويس وويلز ووالد [2] في دراسة وجهت عدة انتقادات لعملية التقييم في صيغتها التقليدية المبنية على الاختبارات التقليدية، بما في ذلك مشكلاتها وأهمها عدم مراعاة خصائص الطلبة وما يصاحب ذلك من قلق وذاتية للمصحح؛ وهو ما ينعكس غالباً على عمليات التعليم والتعلم التي تتم بشكل متزايد إلكترونياً في جميع مراحل وقطاعات المنظومة التعليمية في ظل جائحة كورونا؛ مما فرض على عناصر المنظومة التعليمية واقع كيفية التعامل مع المناهج والأساليب الجديدة في التقييم وتفعيل استخدامات التكنولوجيا المتقدمة المستمدة من تطبيقات النظرية الحديثة في التقييم، وتمثل تكنولوجيا المعلومات والاتصالات العنصر الأساسي في تطبيق الاختبارات الإلكترونية أو الاختبارات الحاسوبية، وهي من الاتجاهات الحديثة في التقييم والتي من خلالها يمكن الحكم على مدى تحقيق الأهداف التعليمية، وفعالية أساليب واستراتيجيات التعليم والتعلم، وقدرات المتعلم وجاهزيته للتعليم، وكذلك فعالية مصادر التعلم.

أكد البربري ما تقدم، وأشار أيضاً إلى أن القاعدة الأساسية لهذه الاختبارات هي أنه ليس من المعقول قياس قدرة أو سمة مجموعة من الطلبة بنفس عدد كلمات المفردات ونفس مستوى السهولة والصعوبة، بل من الأجدى قياس أداء هؤلاء الطلبة بكلمات مفردات تختلف في مستوى صعوبتها حسب مستوى أداء الطالب في توليد الأسئلة من بنوك الأسئلة التي تحتوي على مئات كلمات المفردات أثناء الاختبار[3] .

لذلك ترى البياضة [4] إن ظهور الاختبارات التكيفية كان حلًا لمشكلة الاختبارات التقليدية والمتمثلة في تعريض المفحوصين للفقرات نفسها بغض النظر عن ملاءمة هذه الفقرات لقدراتهم حيث تشمل الاختبارات التقليدية بعض الفقرات السهلة التي يتعرض لها أفراد من ذوي القدرات المرتفعة، وهذا بحد ذاته يمثل مضيعة للوقت والجهد، ثم أن المفحوص قد يشعر بنوع من الملل من جراء اختباره بفقرات لا تتحد قدراته ؛ما يدفعه للإجابة عنها بشيء من عدم الاهتمام ،وكذلك عند اختبار الأفراد ذوي القدرات المنخفضة بفقرات صعبة، مما يشعرهم الإحباط لذا الإجابة عنها بشكل عشوائي.

يبدو من خلال استعراض الأدب النظري وجود إسناد نظري لموضوع الدراسة الحالية ولذلك يسعى هذا البحث إلى التغلب على مشكلات الاختبارات التقليدية من خلال استخدام الاختبارات التكيفية الإلكترونية، وقدرتها في تقييم القدرة المعرفية للطلبة، ورفع التحصيل العلمي للطلبة، وذلك كما أشار إلى أهميته كثير من الباحثين.

II. مشكلة البحث:

تتمثل المشكلة الرئيسية للبحث في كثرة مشكلات الاختبارات التي يتم إعدادها وتقديمها بصورة تقليدية، والتي من الممكن التغلب عليها باستخدام مزايا الاختبارات الإلكترونية التكيفية. وبسبب ندرة الدراسات والابحاث العربية التي كان لها اهتمام في تطبيق الاختبارات الإلكترونية التكيفية في قياس قدرات الطلاب في المقررات التعليمية، تبرز الحاجة إلى تصميم نظام اختبارات إلكترونية تكيفية وقياس فاعليته على تقييم القدرة المعرفية وتعزيز التحصيل الدراسي.

أسئلة البحث:

في ضوء مشكلة البحث المتمثلة في عدم وجود نظام اختبارات إلكترونية تكيفية، يمكن صياغة أسئلة الدراسة على النحو التالي:

1. السؤال الرئيسي:

- ما فاعلية تصميم نظام اختبارات إلكترونية تكيفية على تقييم القدرة المعرفية وتعزيز التحصيل الدراسي؟

2. الأسئلة الفرعية:

- ما معايير بناء وتصميم الاختبارات التكيفي الإلكترونية؟
- ما الفروق بين الاختبارات الخطية والاختبار التكيفي المحوسب في متوسط تقدير القدرة؟
- ما الفروق بين الاختبارات الخطية والاختبار التكيفي المحوسب في تعزيز التحصيل الدراسي؟

III. فرضيات الدراسة

1. لا توجد فروق ذات دلاله إحصائية عند مستوى $\alpha = 0.05$ بين متوسطي تقدير القدرة والخطأ المعياري لتقدير القدرة (كل على حده) للاختبار التكيفي والاختبار الخطي باستخدام طريقة الأرجحية القصوى لتقدير القدرة.
2. لا توجد فروق ذات دلاله إحصائية عند مستوى $\alpha = 0.05$ بين متوسطي درجات طلبة المجموعة التجريبية التي تم تقييمها باستخدام الاختبار التكيفي والمجموعة الضابطة التي تم تقييمها باستخدام الاختبار الخطي.
3. لا توجد فروق ذات دلاله إحصائية عند مستوى $\alpha = 0.05$ بين متوسطي درجات طلبة المجموعة التجريبية التي تم تقييمها باستخدام الاختبار التكيفي وبين المجموعة الضابطة التي تم تقييمها باستخدام الاختبار الخطي تعزى لمتغير الجنس (ذكور/أناث).
4. لا توجد فروق ذات دلاله إحصائية عند مستوى $\alpha = 0.05$ بين متوسطي درجات طلبة المجموعة التجريبية التي تم تقييمها باستخدام الاختبار التكيفي وبين المجموعة الضابطة التي تم تقييمها باستخدام الاختبار الخطي تعزى إلى متغير التخصص.

IV. أهداف البحث:

1. التعرف على الخلفية النظرية للاختبار الإلكتروني تكيفي.
2. الكشف عن اهم المعايير التي تساعد على بناء اختبار تكيفي الإلكتروني.
3. تصميم نظام مقترح لإعداد الاختبارات الإلكترونية التكيفية.
4. تقييم أثر استخدام الاختبار الإلكتروني التكيفي على التحصيل الدراسي.
5. الخروج ببعض التوصيات والاقتراحات المتعلقة بالاختبار الإلكتروني التكيفي.

V. أهمية البحث:

يمكن تحديد أهمية البحث في النقاط التالية:

1. يقدم البحث معلومات عن الاختبار الإلكتروني التكيفي الذي يتم بناءه وفق نظرية الاستجابة المفردة، موضحًا أهميتها كأحد أنواع الاختبارات الإلكترونية ودورها في معالجة القصور في جوانب الاختبار الإلكتروني التقليدية.
2. تنبع أهمية الدراسة من أهمية الاختبار الإلكتروني ودوره الهام في عملية تقييم العملية التعليمية.
3. توجيه نظر الباحثين إلى البحث واكتشاف طرق تقويم الإلكترونية تساعد على حل مشكلات الاختبارات الإلكترونية الخطية، وتتوافق مع قدرة الطالب والحد من عملية القلق والغش والتخمين.
4. توجيه اهتمام مصممي برمجيات الاختبارات الإلكترونية إلى تطوير طرق تقديم الاختبارات الإلكترونية مما يساعد في رفع تحصيل الطلاب وفق للنظريات الحديثة النفسية والتربوية.

VI. حدود الدراسة:

يتحدد البحث الحالي بالآتي:

- *الحدود البشرية:* عينة الدراسة (200) طالبًا وطالبة من طلبة جامعة العادل جميع الكليات المستوى الدراسي الأول
- *الحدود المكانية:* جامعة العادل (جميع كليات الجامعة).
- *الحدود الزمانية:* الفصل الدراسي الثاني للعام الدراسي 2024/2023م.
- *الحدود الموضوعية:* تم اختيار مقرر مهارات الحاسوب (2) وحدة الإنترنت.

VII. مصطلحات البحث:

في ضوء الاطلاع على التعريفات التي وردت في العديد من الأدبيات التربوية ذات العلاقة بمتغيرات البحث، ومراعاة طبيعة بيئة التعلم والعينة، وأدوات القياس بهذا البحث تم تعريف مصطلحات الدراسة إجرائيًا على النحو التالي:

- **الاختبار الالكتروني الخطي:**

ويعرف الاختبار الالكتروني الخطية إجرائيًا: هي إحدى الطرق التعليمية المنظمة التي تهدف لقياس مستوى أداء الطلاب بطريقة الإلكترونية سهلة مما يساعد على التغلب على صعوبات الاختبارات الورقية.

- **الاختبار الالكتروني التكيفي:**

الكلمات المفتاحية

- **ويعرف الاختبار الالكتروني التكيفي إجرائيًا:** هو الاختبارات الذي يقدم للطلبة المفردات التي تناسب مع قدراتهم ويختلف عدد المفردات ومراحلها وترتيبها من طالب إلى آخر حسب قدرة كل طالب.
- **التقييم الإلكتروني:**
- **ويعرف التقييم الإلكتروني إجرائيًا:** هو تطبيق الاختبارات باستخدام الكمبيوتر بدلًا من إجراء الاختبارات بشكل ورقي ويتم من خلالها توظيف البرامج وتقديم التغذية الراجعة الفورية، والتأكد من مدى اكتساب الطلبة للمعلومات.

القدرة المعرفية:

- **وتعرف القدرة المعرفية إجرائيًا:** هي مجموعة من العمليات الذهنية التي تمكن الطالب على التعامل مع المعلومات العلمية التي يحصل عليها عقلياً لتسهيل عملية التعلم.
- ***المعالجة الإحصائية المتبعة في البحث:***

تم استخدام الاختبارات الإحصائية التالية في تحليل البيانات من خلال برنامج الحزمة الإحصائية (spss25 Statistical Package for Social Sciences الإصدار الخامس والعشرون:

1. الإحصاءات الوصفية (الوسيط الحسابي –الانحراف المعياري-النسبة المئوية التكرار-الأهمية النسبية).
2. واستخدام اختبار t-test لعينتين.
3. للتحليل العاملي للتحقق من نظرية الاستجابة المفردة تم اخضاع بنك الأسئلة للتحليل العاملي وفق محك leckase.
4. تحليل التباين الأحادي (One Way ANOVA) لاختبار معنوية الفروق بين أفراد عينة الدراسة المتعلقة بالجنس وعدد مرات الاستخدام والتخصص.

VIII. الإطار النظري والدراسات السابقة:

الإطار النظري:

الاختبارات الإلكترونية الخطية:

لقد كانت الاختبارات الإلكترونية موجودة منذ زمن طويل، كما أشار كلاً من خان وجاويد [5] إلا أنها كانت تستخدم على نطاق محدود، بسبب بعض القضايا المتعلقة بموثوقية الاختبارات، إلى أن اجتاحت جائحة كورونا العالم، فحققت المؤسسات التعليمية قفزة نوعية على كافة مستويات التعليم الجامعي وما قبل الجامعي، وعلى المستويين الإقليمي والدولي في استخدام الاختبارات الإلكترونية، مع البحث عن طرق آمنة ومقبولة وموثقة وعادلة لتطبيق هذه الاختبارات.

تتسم الاختبارات الإلكترونية بمجموعة من الخصائص التي تتمثل فيمَا يلي:

- أقل تكلفة مقارنة بالاختبارات الورقية؛ حيث أنها توفر على المؤسسة تكاليف كل من الطباعة، والحفظ، والنقل.
- يمكن إعداد أنموذجات مختلفة من الاختبار الواحد في نفس الوقت.
- يسهل مراجعتها واكتشاف ما بها من أخطاء وتصحيحها [6].
- سهولة التصحيح تلقائيًا، ثَم يمكن للمعلمين معرفة درجة الواجبات المنزلية الأسبوعية للطلاب على الفور لذا تقديم ملاحظات بنائية لم تكن متاحة بواسطة التعليم اليدوي؛ ولا سيما مع الفصول ذات الأعداد الكبيرة.
- إن الاختبار يمكن الطلاب من رؤية نتيجة اختبارهم الذاتي لتقدمهم، بالإضافة إلى تلقي ردود فعل فورية على محاولاتهم. ومن السهل التأكد من حصول الطلاب على أسئلة مختلفة وقيم متغيرة في كل مرة[7] .

بالرغم من تلك المميزات للاختبارات الإلكترونية إلا أن قد أوضحت دراسة ويس[8] أن التطور النوعي في نظريات القياس من جانب والتطوير ونشر وتوزيع الاختبارات الإلكترونية من جانب آخر قد أظهر

العديد من جوانب القصور في الأنماط المختلفة من الاختبارات الإلكترونية التقليدية خاصة التي تعتمد على توليد العشوائية للمفردات من بنوك الأسئلة ذات قواعد بيانات هيكلية ثابتة دونما اعتبار لمستوى الأداء الفعلي للمتعلم في تحديد مستوى صعوبة أو سهولة المفردات

- آليات سحب المفردات منها وفقًا مستوى أداء المتعلم الأمر الذي قد أدى إلى ظهور قدر مرتفع من القلق الاختبار أثر بالسلب على أداء المتعلم وتكوين اتجاهات سلبية نحو التوسع في الاختبارات الإلكترونية عمومًا.

لذلك كان لظهور الاختبارات الإلكترونية التكيفية أهمية أكد عليها العديد من الدراسات منها دراسة [9] التي أوضحت الآتي:

إلى أنه في الاختبارات الورقة والقلم والاختبارات الإلكترونية الخطية، يتم تقديم نفس العناصر لجميع الأشخاص بغض النظر عن مستوى قدرتهم، وبغض النظر عما إذا كانت العناصر سهلة أو صعبة بالنسبة لهم، حيث لا توفر العناصر السهلة جدًا والصعبة جدًا معلومات عن مستوى أداء الشخص. وفي ضوء التناقض بين قدرة الشخص وصعوبة العنصر، سيؤدي هذا إلى انخفاض في كمية المعلومات، ومن ثم إلى انخفاض في دقة القياس.

الاختبارات الإلكترونية التكيفية:

تعد الاختبارات الإلكترونية التكيفية كما أوضح مارتن ولازنديك من الاتجاهات الحديثة لقياس القدرات العقلية وجوانب الإنجاز المختلفة، فالاختبارات الإلكترونية التكيفية هي اختبارات تتكيف مع قدرة المتعلم، بحيث لا يعرض على الطالب إلا الأسئلة أو المفردات المناسبة لمستواه، فلا يتعرض لمفردات سهلة جداً أو صعبة جداً، وبالتالي يتم الحصول على أقصى قدر ممكن من المعلومات عن قدرة الطالب بأقل عدد ممكن من كلمات المفردات، ومن ثم فإن عدد كلمات المفردات ومراحلها وترتيبها يختلف من طالب إلى آخر حسب قدرة كل طالب[10] .

أوضح كلًا من فون ديفر وماجس ين [11] وجود اختلاف بين الاختبارات المتكيفة مع الحاسوب (الإلكترونية التكيفية) بشكل واضح عن الاختبارات الخطية، وهي الاختبارات التقليدية الأكثر استخدامًا حتى الآن، وتُسمى أحيانًا اختبارات الورقة والقلم، وهي اختبارات يجيب فيها جميع الممتحنين على جميع الأسئلة في الاختبار بغض النظر عن صعوبة العناصر وبغض النظر عن مستوى الممتحن. تختلف الاختبارات المتكيفة مع الحاسوب أيضًا عن الاختبارات المعتمدة على الحاسوب (الإلكترونية الخطية) بشكل عام، حيث يمكن استخدام أجهزة الكمبيوتر لإدارة الاختبارات، ولكن أيضًا بطريقة خطية، كما هو الحال مع اختبارات الورقة والقلم.

أهمية الاختبارات الإلكترونية التكيفية في تنمية مهارات الطلبة:

أشار كلًا روبيرت وديفاد[12] إلى مجموعة الفوائد التي قد يحصل عليها الطالب بفضل تقييمه باستخدام هذا النوع من الاختبارات لخص بالنقاط الآتي:

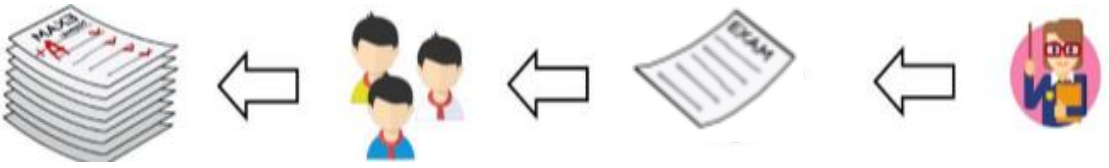
- تقييم الأداء بشكل دقيق وعادل: حيث يمكن تقييم أداء جميع الطلاب على نفس المستوى حتى إذا تفاعلوا مع مفردات مختلفة، ومن منظور نفسي فان هذا يزيد من الأنصاف في الدرجات وتقييم الأداء.
- زيادة الدافعية يتيح الاختبار التكيفي تحديًا مناسب لكل طالب يخضع للاختبار، وبذلك لا يشعر أصحاب الأداء المرتفع أو المنخفض بالملل في أثناء التقييم بالتفاعل مع المستويات المناسبة من التحدي.

مقارنة بين الاختبار التكيفي الإلكتروني والاختبار الخطي (التقليدي) الإلكتروني:

- وقد أشار كلًا من ماجس وين وديفر[11]بأن هناك اختلافات بشكل واضح بين الاختبارات الإلكترونية التكيفية والاختبارات الخطية Linear Tests أوضحه كالتالي:
- الاختبارات الخطية Linear Tests وهي الاختبارات التقليدية التي تعدّ الأكثر استخدامًا حتى الآن، ويطلق عليها أحيانًا اختبارات الورقة والقلم Paper-and-pencil Tests وهي تلك الاختبارات التي يجيب فيها جميع المفحوصين على جميع أسئلة الاختبار بغض النظر عن صعوبة تلك المفردات وكذلك بغض النظر عن مستوى المفحوص.
- وتختلف أيضًا الاختبارات الإلكترونية التكيفية عن الاختبارات المحوسبة Computer-based Tests عمومًا، حيث يمكن أن تستخدم الحواسيب في تطبيق الاختبارات ولكن بصورة خطية أيضًا كما هو الحال في اختبارات الورقة والقلم

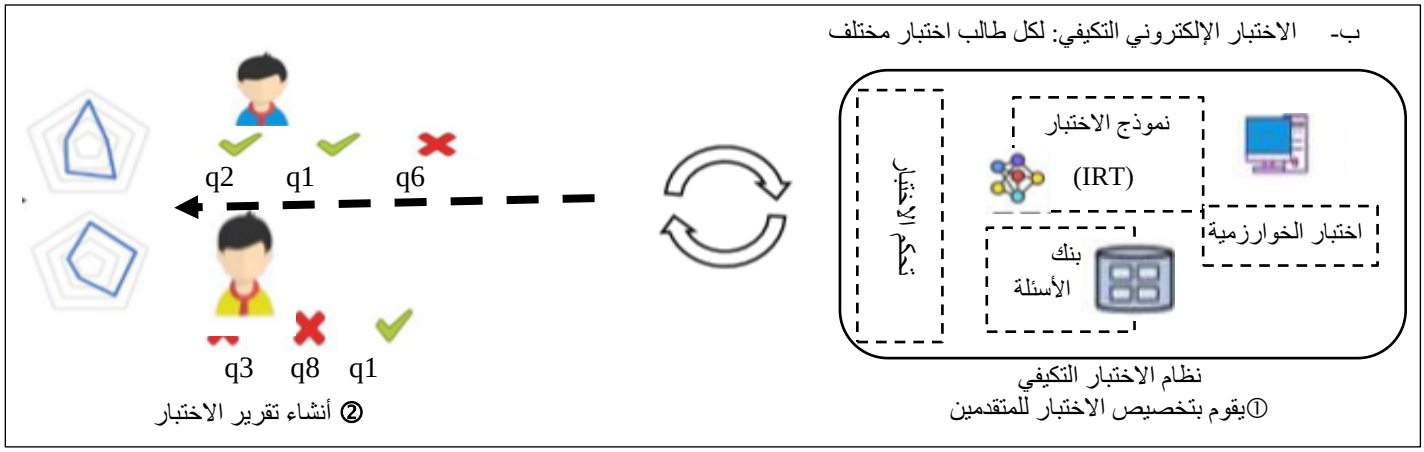
يمكن المقارنة بين الاختبارين من خلال الرسم التوضيحي الذي أشار إليه ليووآخرون [13] الذي يظهر في الأشكال التالية:

أ- الاختبار الإلكتروني الخطي (التقليدي): اختبار لكل طالب



①يقوم المعلم بتصميم الاختبار ②يقوم الطلبة بأداء نفس الاختبار ③يقوم النظام بإبلاغ الطلبة بالدرجة

الشكل 1: فكرة الاختبار الإلكتروني الخطي (التقليدي)



الشكل 2: فكرة الاختبار الإلكتروني التكيفي

وقد أضافة فرنك أن هذه الاختبارات تقيس قدرة الطلاب بكفاءة كونها تقدم سلسلة من الأسئلة التي تم تمييزها بدرجة صعوبة، ويتم اختيار الأسئلة لتتناسب مع مستوى قدرة الطالب، وعند نجاحه في السؤال المقدم له، يتم تقديم سؤال أكثر صعوبة بعد ذلك، وإذا أجاب عن السؤال بشكل غير صحيح يتم تقديم سؤال أقل تحديا، ويتوقف الاختبار عندما يتم تقدير قدرة المتقدم للاختبار بدقة المطلوبة. أي إلى أن يصل الطالب إلى قاعدة الإنهاء التي تم تحديدها من قبل المعلم[14] .

وصف الاختبار التكيفي:

وصف مارتينيز بلومي وآخرون [15] الاختبارات الإلكترونية التكيفية ترتكز على:

أساس نظري مستمد من نظرية الاستجابة للمفردة (item Response Theory) IRT التي عرفها محاسنة[16] نظرية تدور حول المفردة والأداء على الاختبار وكيفية ارتباط الأداء بالقدرة التي تقاس بالفقرات واستجابة المفردة يمكن أن تكون منفصلة أو متصلة أو متفرعة.

كما تعتمد هذه النظرية على القيمة الاحتمالية لاستجابة المتعلم للمفردة الاختبارية وتكون دالة لكل القدرة التي يفترض أن يقيسها الاختبار لدى المتعلم، وخصائص المفردة التي يحاول الإجابة عنها، وإن ذلك يتطلب الحصول على معلومات من مصدرين :

1. أحدهما يتعلق بالمتعلم.
2. والآخر يتعلق بالمفردة الاختبارية؛ وعادة نحتاج إلى قيمة عددية واحدة تتعلق بالمتعلم، وهي بارامتر القدرة المقاسة لدى المتعلم، وقيمة عددية أو أكثر تتعلق بالمفردة الاختبارية، وتمثل هذه النظرية أحد ركائز القياس النفسي الحديث، حيث ترى أن تساوي طول الاختبار من حيث عدد ونوعية المفردات لكل الممتحنين من الناحية المنهجية غير مقبول في قياس الأداء، وأن تقدير أداء المتعلم **يجب أن يكون قائماً على بعدين أساسيين هما:**
1. مدى امتلاك الفرد للقدرة أو السمة موضع القياس.
2. خصائص المفردة السيكومترية ومن ثم اعتمادها في القياس على الأداء الفعلي للمتعلم أثناء الاختبار دون اعتبار الفرضية اعتدالية خصائص العينة التي ينتمي إليها المتعلم كما في نظرية القياس التقليدية. ويتم تفسير العلاقة بين استجابات المفحوص وخصائص المفردة في رصد القدرة الفعلية للمتعلم بشكل مستقل عن العينة التي ينتمي اليها المفحوص.

وأوضح الدامغ والهاجري أنه عند استخدام الاختبارات الإلكترونية التكيفية يمكننا التحكم في مستويات الصعوبة والسهولة بما يتناسب مع مستوى الممتحن، وهذا يؤثر ضمناً على اتجاه الاختبار، حيث قد يرتفع أو ينخفض بناءً على إجابات المتعلم على الاختبار، وبالتالي يمكن الحكم على مستواه دون الحاجة إلى عرض جميع الأسئلة [17].

ومن خلال ما سبق يتضح لنا أهمية هذا النوع من الاختبارات والذي يعد من الأساليب الحديثة التي تساعد على القياس والتقويم وفق نظرية الاستجابة للمفردة، وتختلف الاختبارات الإلكترونية التكيفية في طريقة تطبيقها من حيث:

- يزود الممتحن بالفقرات التي تناسب مستواه فقط وذلك بعرض بعض الفقرات لتحديد قدرته الأولية.
- بناءً على قدرته الأولية المقدرة من خلال أدائه على تلك الفقرة، يزود بفقرة أخرى لاحقة من بنك الأسئلة تتوافق مع التقييم المستمر لمستواه.
- يمكن تقدير القدرة بطريقتين: الارجحية القصوى وطريقة بييزي.
- تتميز هذه النظرية أيضًا بتوفير الوقت والجهد في قياس قدرات الممتحن بأقل عدد من الفقرات.

مبادئ الاختبار الالكتروني التكيفي:

يقوم الاختبار الإلكتروني التكيفي على مبدأين أساسيين هما:

1. بناء بنك أسئلة

عند تصميم الاختبار الإلكتروني التكيفي لا بد من بناء بنك الأسئلة والذي يعد خطوة أولى، كما أوضح بيدر وبيرجونز لأنه يعتمد على مجموعة كبيرة نسبياً من المفردات، إذ يتطلب التقييم الدقيق للمتعلمين ذوي الخصائص المختلفة أن يحتوي هذا البنك على عدد كبير من مفردات الاختبار، مختلفة في معاملات صعوبتها وموزعة بشكل جيد في بنك الأسئلة، يسهل اختيارها في أي اختبار بمواصفات محددة، ومن هذه المعلومات مثلاً: الإجابة الصحيحة، معامل الصعوبة، معامل التمييز، معامل التخمين، وأن عدد (30) مفردة بهذه المواصفات يمكن أن يكون كافياً لبناء اختبار إلكتروني تكييفي [18].

2. إدارة الاختبار الإلكتروني التكيفي:

وقد تم الاعتماد على نوعين من الأسئلة، النوع الأول هو الأسئلة الثنائية التي تسجل الإجابة بصواب أو خطأ وتحدد الدرجة بـ (0,1)، والنوع الأخر هو أسئلة الاختيار من متعدد والتي تكون الإجابة فيها باختيار إجابة واحدة صحيحة من بين أربعة بدائل، وتحدد الدرجة أيضًا بـ (0,1).

تصميم الاختبارات التكيفية المحوسبة:

أوضحت عدد من الدراسات الخطوات الأساسية التي يمر بها تصميم الاختبار التكيفي منها ما أشار إليه Younyoung &Cayce [19]:

فقد قسمها إلى مجموعة من الخطوات الأساسية التي يمر بها تصميم الاختبار التكيفي تتمثل في:

1. تصميم بنك الأسئلة الخاص بالاختبار والذي يتم من خلاله توليد الأسئلة وفق خصائص وقدرات المتعلم.
2. يليها تحديد نقطة البداية والتي تتمثل في اختيار أول سؤال من بنك الأسئلة.

في حالة قيام المتعلم بالإجابة بشكل صحيح يتمكن النظام من تحديد مستوى المتعلم ويترتب عليه تحديد المفردة التالية والتي تتدرج في مستوى الصعوبة ويكون ذلك وفق قدرة المتعلم حتى يصل المتعلم إلى العدد المطلوب من حيث الكم أو وفق القدرة المحددة حسب المتعلم وعليه يتم إنهاء الاختبار أو تكليف المتعلم بمهام أخرى وفق قدرته.

وأشارت دراسة بوساكوك وآخرون إلى ضرورة [20] التصميم الجيد للاختبارات الإلكترونية التكيفية.

كما حدد إبراهيم خطوات بناء الاختبارات الإلكترونية التكيفية فيما يلي [21]:

1. بناء بنك الأسئلة لمفردات الاختبار.
2. اختيار النموذج المناسب للدراسة من نماذج نظرية الاستجابة للمفردة.
3. تحديد استراتيجية انتقاء المفردة وعرض المفردات التالية.
4. اختيار أسلوب تقدير القدرة المناسب.
5. تحديد قاعدة إنهاء الاختبار.

أضاف ونج ولي [22] بإن الحاسوب يقوم باختيار المفردة وفقًا لما يلي:

1. قاعدة محددة، غالبًا ما تكون متوسطة الصعوبة، لتقديمها للممتحن في البداية.
2. ثم استخدامها في تقدير مستوى القدرة الأولي للفرد.
3. وبناء على المستوى الأولي يتم اختيار المفردة التالية لتقديمها للممتحن، وبعد ذلك يتم إعادة تقدير مستوى قدرة الفرد.
4. وتستمر هذه الخطوات حتى يتحقق محك يتم اختياره لإنهاء تطبيق الاختبار وتحديد القدرة النهائية للفرد ومستوى دَقَّة تقديرها.

IX. الدراسات السابقة:

1. دراسات أوصت باستخدام الاختبار الإلكترونية بصورة عامة:

تضمنت العديد من الدراسات السابقة التي أوصت بضرورة استخدام الاختبارات الإلكترونية وفعاليتها في العملية التعليمية منها:

a. فعاليتها في تطوير أساليب التقييم مثل:

- دراسة دماس [23]: التي أسفرت نتائجها على وجود موقف إيجابي عن الاختبارات الإلكترونية حيث أكدت الدراسة إن الاختبار الإلكترونية حققت رضا الطالب من حيث التصحيح الفوري والصلاحية والشفافية. وخلصت الدراسة إلى أن الاختبارات الإلكترونية ستصبح هي المفضلة لدى الطلبة وإنها سوف تكتسب فعاليتها في سياق أساليب التقييم..

b. دور استخدام الاختبارات الإلكترونية في تعزيز مصداقية النظام التعليمي مثل:

- **دراسة خميس [24]:** وقد أظهرت أن الاختبارات الإلكترونية تتمتع بدرجة عالية من الثقة والمصداقية بين الطلبة وأن هذا النوع من الاختبارات يحظى بقبول وتفضيل كبير.

2. دراسات أوصت باستخدام الاختبارات الإلكترونية التكيفي:

قامت بعض الدراسات والابحاث بتطبيق الاختبارات التكيفي الإلكترونية والتحقق من فاعليتها في العديد من المجالات منها:

a. فاعليتها في تقدير القدرة العقلية بالاعتماد على أسلوب تقدير القدرة بطريقة الارجحية القصوى مثل:

- **دراسة عودة وعبيدات [25]:** التي أسفرت نتائجها إلى زيادة دقة تقديرات القدرة من خلال الاختبارات التكيفية مقارنة بنظيرتها الخطية مع اختلاف أسلوب تقدير القدرة المستخدمين، وزيادة في كمية المعلومات لأسلوب بطريقة الارجحية القصوى.

- **دراسة نور الدين [26]:** التي أوصت بضرورة اعتماد الجامعات للاختبارات الإلكترونية التكيفية والتوسع في تحويل الاختبارات التحصيلية الورقية والخطية المحوسبة إلى اختبارات تكيفية محوسبة، والتحقق من فاعلية الاختبارات التكيفية المحوسبة في تقدير قدرة الأفراد في السياق الثقافي للبيئة العربية ومدى الاختلاف بين التقدير البيزي والأرجحية القصوى في تقدير قدرة الأفراد في الاختبارات التحصيلية.

- **دراسة فرحات [27]:** وجود فروق ذات دلالة إحصائية بين متوسطات الخطأ المعياري لتقدير القدرة للاختبار الخطي المحوسب والاختبار التكيفي المحوسب باستخدام أسلوب الاحتمال الأقصى أو بيزية لتقدير القدرة وقاعدة إنهاء الاختبار بعدد محدد من الفقرات أو بأقل خطأ معياري، وكانت النتائج لصالح الاختبار التكيفي المحوسب.

b. فاعليتها في قياس التحصيل المعرفي لطلبة مثل:

- **دراسة مارتين ولازيندي [10]** وجاءت نتائج البحث لصالح الاختبار التكيفي الإلكتروني والذي يعتبر أكثر دقة في قياس التحصيل.

- **دراسة أجراها أوزورت وأوزورت [28]:** وقد أكدت نتائج لاختبار إمكانية وفعالية استخدام أنظمة الاختبار التكيفي في المناهج التركية.

- **دراسة نور لدين [26]:** حيث أجرى بحثًا هدف إلى التعرف على مدى فاعلية القياس التكيفي المحوسب في قياس تحصيل طلبة الجامعة توصل أن الاختبار التكيفي المحوسب فعال على كل من الاختبار التقليدي الخطي أو الورقي في قياس تحصيل طلبة الجامعة.

- **دراسة جرجس دميانة [29]:** هدف الدراسة إلى قياس أثر الاختبارات التكويني التكيفي في تنمية التحصيل الدراسي المتعلق بالبرمجة لدى طلاب تكنولوجيا التعليم وقد أسفرت النتائج عن وجود فرق دال إحصائياً لصالح التطبيق البعدي.

c. فعاليتها في تخفيف مستوى القلق مثل:

- **دراسة جيمس[30] :** التي أجمع فيها الطلبة على دور الاختبارات المحوسبة في تخفيف مستوى القلق لديهم؛ إضافة لقلة التكاليف المادية عند عقد هذه الاختبارات .

ويمكن تلخيص ما توصلت إليه الدراسات السابقة بالنقاط التالية:

1. أكدت الكثير من الدراسات على إيجابيات اعتماد الاختبارات التكيفية بديلًا للاختبارات الخطية.
2. تأكد معظم الدراسات السابقة على عمل المزيد من الدراسات بشأن الاختبارات الإلكترونية التكيفية.
3. كما أكدت الدراسات السابقة على دور الاختبارات الإلكترونية التكيفية في رفع التحصيل الدراسي لدى الطلبة.
4. الفجوة البحثية تظهر الأدبيات محاولات لتطبيق فكر الاختبارات الإلكترونية التكيفية في البيئة العربية، وتعدّ تلك المحاولات قليلة إذا ما قورنت بنظيرتها في البيئة الأجنبية؛ لذلك أن ما يميز هذه الدراسة بأنه الأولى -في حدود علم الباحثين - التي تجرى في الجامعات اليمنية كون الاختبارات التكيفي حديثًا نسبيًا.

X. الطريقة والإجراءات:

منهج البحث:

ينتمي هذا البحث إلى فئة البحوث التطويرية التي تستخدم بعض مناهج الدراسات الوصفية في مراحل الدراسة والتحليل والتصميم والمنهج الشبة تجريبي في الكشف عن فاعلية المتغير المستقل المتمثل في الاختبارات الإلكترونية التكيفية على المتغير التابع تقييم القدرة المعرفية للطلبة في التحصيل الدراسي لدى طلبة جامعة العادل المستوى الأول، ويتمثل المنهج الشبة تجريبي في استخدام اختبارين والمقارنة بينهما للتحقق من نتائج الفرضيات.

A. متغيرات البحث:

أشتمل البحث على المتغيرات التالية:

https://doi.org/10.20428/jst.v30i6.2815

- **المتغير المستقل:** المتمثل في الاختبارات الإلكترونية التكيفية.
  ويشمل متغيرين فرعيين: (الجنس، التخصص)
- **المتغير التابع:** المتمثل في دور الاختبارات الإلكترونية التكيفية تقييم القدرة المعرفية وتعزيز التحصيل الدراسي لدى طلبة جامعة العادل.

B. المنهج التجريبي للبحث:

- استخدم هذا البحث التصميم الشبة تجريبي ذو المجموعتين المتكافئتين الذي يعتمد على مقارنة نتائج تقييم الطلبة عند استخدامهم الاختبار الإلكتروني الخطي وبعد استخدامهم الاختبار الإلكتروني التكيفي. ويبين الجدول الآتي التصميم التجريبي لهذا البحث:

جدول (1) التصميم التجريبي للبحث

المجموعة التجريبية	التطبيق القبلي	مادة المعالجة التجريبية	الأدوات
المجموعة التجريبية	اختبار الإلكتروني	اختبار الإلكتروني التكيفي	اختبار تحصيل
المجموعة الضابطة	خطي		

C. مجتمع البحث وعينته:

مجتمع البحث:

تكون مجتمع الدراسة من جميع الطلبة المسجلين والدارسين من طلبة جامعة العادل، وذلك للعام الجامعي 2024/2023م ويبلغ عددهم (988) طالبًا طالبة، بواقع (383) طالبًا و (605) طالبة خلال الفصل الدراسي الثاني.

عينة البحث:

بناءً على حجم مجتمع الدراسة من الذكور والإناث تم تقسيم عينة البحث إلى مجموعتين مجموعة ضابطة ومجموعة تجريبية تحوي كل مجموعة (100) طالبًا وطالبة من طلبة جامعة العادل المستوى الأول لجميع لكليات.

جدول (2) توزيع عينة البحث الأساسية

المجموعة	النوع	العينة	النسبة المئوية
المجموعة التجريبية الاختبار الإلكتروني التكيفي	ذكور	57	29%
	إناث	43	21%
المجموعة الضابطة الاختبار الإلكتروني خطي	ذكور	57	29%
	إناث	43	21%
المجموع	**200**		**100%**

D. إجراءات البحث:

أداة البحث:

تحقيقًا لأهدا ف الدراسة تم تطوير واستخدام أداة لتقييم طلبة جامعة العادل - جميع الكليات للمستوى الأول - عن طريقة استخدام الاختبار الإلكتروني التكيفي على التقييم المعرفي في التعليم، ودوره في رفع التحصيل الدراسي لدى طلبة جامعة العادل. وفي هذه الخطوة تم تحديد وبناء أداة البحث المتمثلة في: **الاختبار التحصيلي**.

E. ضبط أداة البحث: (الاختبار التحصيلي)

صدق الاختبار:

للتحقق من صدق الاختبار من تم وضع الاختبار في صورته الأولية حيث يتكون من (100) سؤال وعرضه على السادة المحكمين لإبداء الرأي من حيث : مناسبة الاختبار لقياس ما وضع لقياسه، مناسبة لأسئلة للمقرر المختار، مناسبة الأسئلة للمستويات المعرفية التي تم ذكرها مسبقًا، وضوح تعليمات الاختبار، وبعد تلقي التعليمات المحكمين ،تم إعادة صياغة بعض الأسئلة ،و تعديل بعض الأفعال للتناسب مع المستوى التي وضعت من أجله حيث تم حذف بعض الأسئلة من قبل السادة المحكمين ،وأصبح الاختبار مكون من (91) سؤال جاهز لتطبيقه على العينة الاستطلاعية.

ثبات الاختبار

من خلال تطبيق الاختبار على عينة تجريبيه عددها (65) طالب من طلبة كلية العلوم الأدرية تم اختيارهم بطريقة عشوائية من بين أقسام الكلية. حيث تتراوح معاملات الصعوبة للمفردات البنك بين (0.27-0.95) والذي يشير إلى مستوى مناسب.

تحديد زمن الاختبار:

تم تقدير الزمن من خلال جمع أزمنة إجابات الطلبة وقسمتها على العدد الكلي للعينة فكان الزمن اللازم تقريبًا (45) دقيقة، مع الإشارة إلى عدم احتساب الوقت المنقض في التوجيه وإعطاء التعليمات.

تصميم الاختبار والمقياس (أدوات البحث):

تحقيقًا لأهداف الدراسة تم تطوير واستخدام أداة لتقييم طلبة جامعة العادل- جميع الكليات للمستوى الأول- عن طريقة استخدام الاختبارات الإلكترونية التكيفية ودورها في تقييم القدرة المعرفية للطلبة في التعليم، وفي هذه الخطوة تم تحديد وبناء أداة البحث المتمثلة في: تصميم اختبار تحصيلي الإلكتروني. وفقًا لطبيعة البحث الحالي والتي تستهدف قياس فاعلية نمط الاختبار الإلكتروني التكيفي ومقارنة بنمط الاختبار الإلكتروني الخطي.

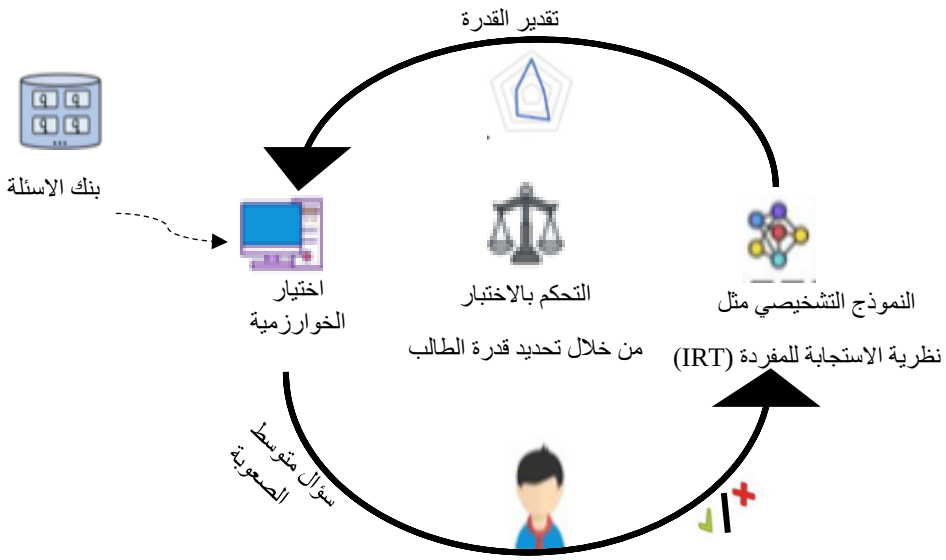
وبناء على ذلك تم تصميم النمطين:

a. تصميم الاختبار الإلكتروني التكيفي.
b. تصميم الاختبار الإلكتروني خطي.

أولًا: تصميم الاختبار الإلكتروني التكيفي.

1. مراحل تصميم الاختبار الإلكتروني التكيفي:

من خلال الاطلاع على الأدبيات ومنها ما أشار إليه كلًا من[19، 13] التي لخص مجموعة من الخطوات الأساسية التي يمر بها الاختبار الإلكتروني التكيفي. وقد تم وضع المعايير التي يقوم عليها تصميم الاختبار الإلكتروني التكيفي وفقًا لمخطط آليات السير البرنامج الذي أشار إليه ليو وآخرون كما هو موضح في الشكل (3):



رقم الشكل 3: مخطط بآليات سير البرنامج

ومن خلال ما تم عرضها أعلاه تم تحديد المراحل التي يمر بها تصميم الاختبار الإلكتروني التكيفي وهي كالتالي:

المرحلة الأولى: بناء مستودع (بنك)الاسئلة:

تم بناء بنك أسئلة يحوي (91) سؤالًا جاهز لتطبيقه على العينة بعد أن تم تقنين الأسئلة وتعديلها من قبل السادة المحكمين والتأكد من مناسبة الاختبار لقياس ما وضع لقياسه، مناسبة لأسئلة للمقرر المختار، مناسبة الأسئلة للمستويات المعرفية. وذلك كما أوصى به عدد من الباحثين منهم:

- علام [31] الذين أوصوا بكتابة أو جمع عدد كافٍ من المفردات الاختبارية تقيس مدى متسعًا من القدرة المعينة، وهذا يتطلب البدء بتحديد موصفات كل مفردة، ومحتواها، ونطاق المهارات التي تُمثلها، وكلما كانت هذه المواصفات واضحة، أمكن بناء عدد كافٍ ومتنوع من المفردات المتباينة في صعوبتها على مدى القدرة أو السمة المراد قياسها، كما ينبغي مراجعتها بدقة للتحقق من جودتها وعدم تحيزها لجماعة معينة دون غيرها.
- النظير الذي أكد أيضًا أن أهم مكون من مكونات الاختبار الإلكتروني التكيفي هو تحديد بنك الأسئلة حيث يمكن من خلاله تطبيق الاختبار التكيفي، ويجب أن يحتوي الاختبار على مجموعة كافية من الفقرات التي تتوزع صعوبتها حسب مستوى وخصائص الطلبة[32] .

حيث تم إجراء اختبار قبلي لتحديد درجة صعوبة كل سؤال من الأسئلة وتقسيم بنك الأسئلة إلى ثلاث مستويات (مستوى الأسئلة الصعبة – المتوسطة - السهلة) وتحديد مستوى الصعوبة لكل سؤال.

المرحلة الثانية: اختيار نموذج التحليل المناسب من نماذج نظرية الاستجابة للمفردة.

قبل أن نقوم بتطبيق الاختبار التكيفي لابد من اختيار أحد النماذج الرياضية المرتبطة بنظرية الاستجابة للمفردة.

نماذج الاستجابة للمفردة:

هناك عدد من النماذج في نظرية الاستجابة للمفردة LRT models لتحليل بيانات المفردات الاختبارية؛ غير أن هناك ثلاثة نماذج شائعة الاستخدام في البيانات ثنائية التقسيم التي تقدر بالدرجة (1) للإجابة الصواب، (0) للإجابة الخطأ، وهي النماذج (أحادية، وثنائية، وثلاثية) البارامتر، والتي أشار إليها عدد من الدراسات منها ما ذكرها التقي [33] حيث أظهرت العديد من الدراسات ومنها دراسة السباحة نتائج أوصت بضرورة استخدام نماذج الاستجابة للفقرات أو نموذج راش في بناء وتقنين مقاييس الاتجاه.

وقد تم استخدام نموذج راش (أحادي البارامتر) وهو النموذج الذي اتصل بالعينات صغيرة الحجم حيث يتطلب كحد أدنى (٢٠) مفردة و(٢٠٠) فرد. الذي يتميز ببساطته عن غيره من النماذج المتعلقة بنظرية استجابة الفقرة، لأنه يحتوي على معلمة واحدة فقط وهي الصعوبة، وبالتالي فهو يتجنب التعقيدات الإحصائية المتعلقة بعدة بارامترات التي تتعامل معها النماذج الأخرى، وبالتالي فهو يفسر الأداء وفق أقل عدد من المتغيرات، مما جعله أكثر استخداماً بين الباحثين.

المرحلة الثالثة: تحديد نقطة البداية:

وهو عبارة عن سؤال متوسط الصعوبة يتم اختياره عشوائيًا من بنك الأسئلة (من مستوى الأسئلة المتوسطة الصعوبة)، ومن خلال إجابة الطالب على السؤال يتدرج في الأسئلة التي تتناسب مع قدرته، وقد تم الاعتماد على أسلوب الأرجحية العظمى (أقصى معلومات):وهو بعد أن يتقدم المفحوص للفقرة أو الفقرات الأولية يتم تقدير القدرة له ثم يتم اختيار الفقرة ذات المعلومات الأكبر عند مستوى القدرة المقدر الحالي للمفحوص التي لم يتم تقديمها سابقًا وبعدها يتم تصحيحها وإيجاد تقدير جديد لقدرة المفحوص ثم يتم اختيار الفقرة التالية ذات المعلومات الأكبر عند مستوى القدرة الجديد وهكذا... حتى يتم إنهاء الاختبار. بناء إلى ما أشارت إليه عدد من الدراسات مثل دراسة كلًا من[26، 27،28،34،35] الذين أكدوا على فاعلية الاختبارات التكيفية المحوسبة في تقدير القدرة باستخدام طريقة الأرجحية العظمى.

المرحلة الرابعة: أساليب اختيار المفردات في الاختبارات الإلكترونية التكيفية:

خوارزمية اختيار المفردة التالية: تم اختيار خوارزمية لورد حيث تم تقدير قدرة الطالب المبدئية بالاستفادة من إجابتها على السؤال المتوسط الصعوبة، ثم تم جمع صعوبة السؤال مع أعلى صعوبة متوقعة.

حيث تم تصميم خوارزمية الاختيار للفقرات بطريقة بسيطة جدًا وواضحة، يتم تقدير قدرة أحد المتقدمين للاختبار ليكون له قياس قدرة معين واختياره باستخدام سؤال مستوى صعوبته يعادل مستوى قدرته، وإذا نجح في الاختيار يتم رفع مستوى تقدير المقدرة، وإذا خطأ في الإجابة عن هذا السؤال يتم خفض مستوى تقدير القدرة، وتقديم سؤال آخر إليه، وتكرر هذه العملية حتى يتم تقدير القدرة النهائي.

و تم استخدام إحدى استراتيجيات القياس التكيفي في اختيار المفردة التالية والتي أكد عليها كلًا [36] بأن المفردات التالية يتم اختيارها بما يتناسب مع قدرة كل ممتحن، حيث يتم اختيار درجة صعوبتها بحيث لا تكون صعبة جدًا أو سهلة جدا عليه، بمعنى أن احتمال إجابته عليها بشكل صحيح هو (50%)، ومع كل إجابة على المفردات التالية يصبح الحاسوب قادراً بشكل متزايد على تقييم معرفة الممتحنين وتقدير قدرته، لتحديد أفضل المفردات في التطبيق بعد ذلك، ولذلك فإن الاختبار التكيفي الإلكتروني هو نوع من الاختبارات التي تم تطويرها لزيادة كفاءة عملية تقدير معرفة الممتحنين، من خلال تعديل المفردات المقدمة لهم بناءً على إجاباتهم السابقة أثناء الاختبار.

وهذا ما أوضحت دراسة وايس بأن يتم تحديد العنصر التالي في المراحل الأولية من الاختبارات التكيفية المحوسبة من خلال قاعدة الخطوة. إذا تم الإجابة على العنصر الأول بشكل صحيح، يتم زيادة مقدار القدرة الأصلي السابق بمقدار، على سبيل المثال (0.50)، وإذا تم الإجابة عليه بشكل غير صحيح، يتم تقليل مقدار القدرة الأصلي السابق بنفس المقدار. بعد قياس كل عنصر وتسجيله، تُستخدم القدرة الجديدة لتحديد العنصر التالي، والذي يتم اختياره من بين جميع العناصر الموجودة في بنك الأسئلة التي لم يتم قياسها لهذا المختبر، من خلال تحديد العنصر الذي يوفر أكبر قدر من المعلومات للقدرة الحالية[37] .

وفقا لما أشار اليها الباحثين وبناء على ما سبق تم تحديد أليه انتقاء الفقرة التالية في البرنامج المقترح: حيث تم استخدام طريقة جمع صعوبة السؤال مع أكبر صعوبة إذا اجاب صح، والتي أكدت عليها بعض الدراسات ودعمت صحتها مثل دراسة كلًا لي وآخرون[38] حيث أوضحوا بأن هذه الطريقة تعطي تقديرات دقيقة بنفس القدر من الطرق التقليدية بالإضافة إلى ذلك، فإنها بسيطة وسهلة التنفيذ. هذا ما جعلها خيارًا جذابًا للاختبارات التكيفية. تعتمد على طرح سؤال متوسط الصعوبة ثم وفقًا لإجابة المتعلم يتم اختيار المفردة التالية وفقا للمعادلات التالية:

إذا كانت الإجابة صحيحة ينتقل الطالب إلى المفردة الأكثر صعوبة في نفس موضوع التعلم من خلال المعادلة التالية:

الإجابة صحيحة = (أعلى درجة صعوبة + درجة صعوبة السؤال السابقة) /2 (1)

وإذا كانت الإجابة خطأ ينتقل الطالب إلى السؤال الأقل صعوبة من خلال المعادلة التالية:

الإجابة خطأ = (أقل درجة صعوبة + درجة صعوبة السؤال السابقة) /2 (2)

وهكذا يستمر الطالب بين الانتقال بين الأسئلة حسب القاعدة المحددة حتى يصل إلى قاعدة الإنهاء، وقد استخدم البحث الاستراتيجية المحوسبة التي أشار إليها دعنا [39] في اختيار الفقرة التالية التي تقدم للطالب بعد تحديد مستوى القدرة لديهم حيث يقدم فقرات متتالية الصعوبة طبقا لصيغة رياضية تسمى دالة الخطوة التي تعني ان صعوبة الفقرة التي ترتيبها (ن+1) يساوي صعوبة الفقرة التي ترتيبها ن مضاف اليه أو مطروح منه قيمة معينة تسمى حجم الخطوة، واستخدم البحث قاعدة الإنهاء بأن تصل الدرجة المتوقعة للطالب ≤ 20

المرحلة الخامسة تحديد قاعدة إنهاء الاختبار:

بعد الاطلاع على عدد من الدراسات التي أوضحت عدد من الطرق يمكن من خلالها أنهاء الاختبار منها:

1. وفقًا لعدد محدد من المفردات.
2. وفقًا للخطأ المعياري.
3. عندما يكون قياس القدرة بعيدا جدا عن محك النجاح.
4. وفق درجة معيارية يتم تحديدها مسبقًا.

وقد أشار كلًا من مارتين وليزيندس [10]يمكن تمثل قاعدة الإنهاء المتمايز في عدد المفردات التي يحصل عليها الممتحنون، ويمكن أن تكون ذات الطول الثابت، والطول المتغير هما طريقتان لتحديد زمن وكيفية إنهاء الاختبار:

- يتطلب اختبار الطول الثابت أن يجيب جميع الممتحنين على نفس العدد من المفردات.
- أما اختبار الطول المتغير فيجيب الممتحنون على أعداد مختلفة من المفردات حتى يتم استيفاء مستوى تقدير القدرة في زمن محدد، أداء الممتحن بنفس درجة الاحكام.

وقد إشارة البياضة[4] إلى مجموعة من القواعد التي يمكن من خلالها إنهاء الاختبار التكيفي ومنها:

- التوقف عند استنفاذ الفقرات الموجودة في بنك الأسئلة وذلك في حالة البنوك الصغيرة.
- التوقف عند عدم كفاية الفقرات.
- في حالة الاختبارات الإلكترونية التكيفية عندما تكون إجابة المفحوص سريعة جدًا، أو بطيئة جدًا مما يدل على عدم جدية الممتحن.
- الوصول إلى الحد الأقصى لطول الاختبار المحدد سلفًا.
- الوصول الى مستوى معياري معين، يتم ايقاف الاختبار بمجرد وصول الطالب إلى الحد الأدنى للأداء المقبول.

وبناءً على ما سبق تم باختيار قاعدة لإنهاء وفق محك الطول المتغير كما أوصى بها ليروكس [40] والذي يعتمد على الوصول لمستوى معين من دِقة التقدير لقدرة المفحوص، ووفقًا لهذا المحك لا يتعرض جميع المفحوصين لنفس العدد من المفردات، ولكن بالقدر الذي يضمن تحقيق مستوى الدقة المحدد. حيث يتم ايقاف الاختبار بمجرد وصول الطالب إلى الحد الأدنى للأداء المقبول. وقد تم تحديدها بالوصول الطالب إلى مستوى معياري محدد، أي عندما تصل درجة الطالب الحقيقية ≤ 20كقاعدة لإنهاء الاختبار لتحديد قدرة المتعلم.

حساب الدرجة النهائية:

- ويتم حساب الدرجة النهائية للطالب من خلال الآتي:

- الدرجة المتوقعة: هي مجموع درجات الأسئلة المقدمة إلى الطالب على احتمال أن يجيب عليها جميعا بشكل صحيح.
- الدرجة الحقيقية: هي مجموع درجات الأسئلة التي أجاب عليها الطالب بشكل صحيح.

- حساب الدرجة الحقيقية في نموذج راش:

لا يوجد طريقة واحدة محددة لحساب الدرجة الحقيقية للطالب في نموذج راش، وقد تم استخدام طريقة "استخدام عدد الأسئلة التي أجاب عليها الطالب بشكل صحيح "حيث يتم فيها التالي:

- يتم حساب عدد الأسئلة التي أجاب عليها الطالب بشكل صحيح.
- يتم تحويل عدد الأسئلة الصحيحة إلى درجة على مقياس محدد.
- من المهم أن تكون طريقة حساب الدرجة الحقيقية للطالب عادلة لجميع الطلاب.
- يجب أن تأخذ طريقة الحساب بعين الاعتبار صعوبة الأسئلة التي أجاب عليها الطالب.
- حساب الدرجة الحقيقية للطالب = درجة صعوبة السؤال x الدرجة المخصصة للسؤال x احتمالية الإجابة (1,0) (3)
- حساب الدرجة المتوقعة للطالب= درجة صعوبة السؤال x الدرجة المخصصة للسؤال.

لإيجاد الدرجة النهائية استخدم البحث طريقة المجموع المرجح يتم حساب متوسط الدرجة المتوقعة والدرجة الحقيقية

- ولإيجاد الدرجة النهائية = (4)
نسبة الدرجة النهائية xالدرجة الكلية المحددة من قبل المعلم. (5)

https://doi.org/10.20428/jst.v30i6.2815

حيث تم برمجة الاختبار بحيث يختلف عرض الاسئلة الموضوعة من مختبر اخر.

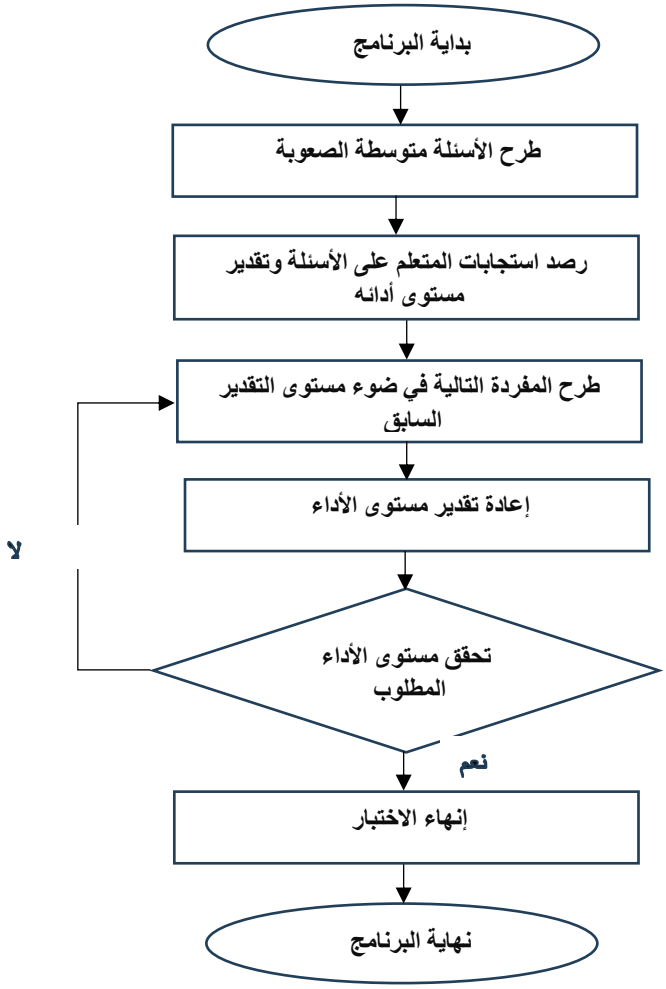
وأين كانت طريقة التوقف التي تم اختيارها فإن الاختبار التكيفي لا يقف إلا في الحالات التالية:

1. أن يصل الطالب إلى الحد الأدنى من عدد الأسئلة بحيث لا يشعر الطالب بأنه لم يتم قياسه بشكل دقيق إلا إذا أجاب على الأقل على 10 أو 20 سؤالاً. حيث روعي عند تصميم البرنامج هذه الخاصية.
2. حيث يمكن إنهاء الاختبار الإلكتروني لتكيفي وفقًا لعدد محدد من المفردات تم تحديدها ب (10) أسئلة من نوع الأسئلة الصعبة بدون الحاجة إلى عرض جميع فقرات الاختبار للطلبة ذو القدرة العالية وهذا أهم ما يميز هذا النوع من الاختبارات.
3. يغطي الاختبار جميع المواضيع، بمعنى أن الأسئلة المقدمة للطالب تمثل جميع مجالات الأسئلة التي يغطيها الاختبار، وخاصة في الاختبارات التي تشمل أكثر من مجال.

المرحلة السادسة تنفيذ الاختبار:

بناءً على ما سبق من خطوات لتصميم لاختبار تم تنفيذ الاختبار على عينة المتمثلة بطلبة جامعة العادل حيث تم تنصيب الاختبار في معمل الجامعة.

يمكن تلخيص الخطوات السابقة من خلال المخطط التالي الذي يوضح آليات عمل البرنامج حتى يصل بقواعد إنهاء الاختبار التكيفي كما هو موضح في الشكل (4).



رقم الشكل 4: مخطط يوضح

2. مميزات الاختبارات التكيفية:

من خلال تطبيق الاختبار التكيفي على عينة الدراسة – طلبة جامعة العادل – أظهر البحث عدة مميزات للاختبارات التكيفية مقارنة بالاختبارات الخطية أو التقليدية (اختبارات الورقة والقلم) وقد تطابقت تلك المميزات مع ما ارصدته الأدبيات، منها ما أشار إليه Linacre [41] ومن تلك المميزات التالي:

- سهولة إدارة المفردات الاختبارية.
- إتاحة درجة الفرد والتغذية الراجعة عن الأداء بشكل فوري.
- الحفاظ على مستوى دافعية المفحوص، حيث لا يتعرض لعدد كبير من المفردات.
- خفض معدلات القلق لديه كما أكد عليها عدد من الدراسات مثل [3، 28].

– بالإضافة إلى النقطة الأهم وهي دقّة تقدير مستوى القدرة للمفحوص أو مستوى السمة لديه.

– تمتع بأمان عال، وتجرى عند الطلب، ولا تحتاج إلى ورقة إجابة، وتلائم قدرة المختبر، ومرنة في اختيار المفردة، ومستوى تقنينها أعلى من الاختبار الخطي، ومدة المراقبة فيها قصيرة، وتعد فيها التقارير بصورة فورية ودقيقة وهذا ما أشار إليه[42،36] .

– كما تقوم بتوفير معلومات إحصائية عن أداء الممتحن في كل مفردة من مفردات الاختبار
حيث يمكن للطالب معرفة عدد الأسئلة التي أجاب عليها ودرجة صعوبة كل سؤال، والاجابة الصحيحة للسؤال في حالة إذا كان أجاب بشكل خاطئ على السؤال.

– تقدم عدد أقل من المفردات للحصول على المستوى المطلوب من الدقة والموثوقية مقارنة بالاختبارات الإلكترونية التقليدية. وقد أكد ذلك كلًّا من Huo, Magis and Raiche [44, 43].

3. مشكلات الاختبارات الإلكتروني التكيفي:

من خلال الاطلاع على بعض الدراسات والأبحاث السابقة التي أشارت إلى بعض المشكلات والعيوب للاختبارات الإلكتروني التكيفي منها دراسة علام [31] تم تلخيص تلك المشكلات كما يأتي:

1. عدم الاتفاق حول طريقة اختيار السؤال من بنك الأسئلة ما قد يؤدي إلى نتائج متباينة وقد تكون غير دقيقة.
2. تتطلب إجراءات تطبيقها إنشاء بنك أسئلة يشتمل على عدد كبير من المفردات تختزن في الحاسوب.
3. تحتاج المشكلات المتعلقة بتأثير سياق عرض المفردات وترتيبها، وتوازن محتواها إلى مزيد من الدراسات والبحوث التطويرية.
4. يتطلب أداء الاختبار التكيفي توافر بنية أساسية تدعم الاتصال بالشبكات المحلية وشبكة الإنترنت.
5. أن المختبر الذكي قد يجيب إجابة خاطئة عن الأسئلة الأولية بشكل متعمد، وسيفترض البرنامج بعد ذلك أن لديه قدرة منخفضة فيختار له سلسلة من الأسئلة السهلة.

4. معالجة مشكلات الاختبار الإلكتروني التكيفي:

تم التغلب على بعض المشكلات السابقة من خلال برنامج الاختبارات الإلكترونية التكيفية التي تم إنشاءه من خلال الآتي:

– تم التغلب على المشكلة الأولى من خلال: اختيار السؤال الأولي من بنك الأسئلة وهو سؤال متوسط الصعوبة حيث أكد على هذه الطريقة عدد من الدراسات: منها [25 , 38 ,27].

– تم التغلب على المشكلة الثانية من خلال: أنشاء بنك أسئلة يحوي (91) سؤال تم تقنينه من قبل المحكمين. حيث أشار عدد من الدراسات بأنه يمكن أنشاء بنك أسئلة مكون (30) سؤلًا يعد كافيًا ومنها دراسة نور لدين [26]: حيث أجرى بحثًا هدف إلى التعرف على مدى فاعلية القياس التكيفي المحوسب في قياس تحصيل طلبة الجامعة وقام بإنشاء بنك أسئلة مكونًا من (48) سؤالًا، ودراسة Istiyono et al. [42] تألفت أدوات البحث من بنك للأسئلة يحتوي على (62) فقرة.

– وقد تم التغلب على المشكلة الثالثة من خلال: اختيار خوارزمية لانتقاء المفردة التالية تم الإشارة إليها سابقًا.

– وقد تم التغلب على المشكلة الرابعة من خلال: تم تصميم برنامج مجاني بدون الحاجة إلى استخدام الإنترنت لتطبيق الاختبارات التكيفية.

– وقد تم التغلب على المشكلة الخامسة من خلال: وتم التغلب على مشكلة المختبر الذكي قد يجيب إجابة خاطئة عن الأسئلة الأولية بشكل متعمد من خلال إيجاد الدرجة المتوقعة والدرجة الحقيقية.

ثانيًا: الاختبار الإلكتروني الخطي:

تم تصميم اختبار الإلكتروني خطي مكونًا من (15) سؤالًا من نمط الاختيار من متعدد – الصواب والخطأ ذي خصائص سيكومترية جيدة مشتق من بنك الأسئلة.

– تم تقدير أداء المتعلم عن طريقة عرض الأسئلة على المتعلم بشكل خطي.

– تحديد درجة كل سؤال على حسب مستوى الصعوبة الخاص به، حيث تم تقسيم الأسئلة على حسب مستوى الصعوبة الى (أسئلة صعبة-متوسطة-سهلة).

وعند تصميم الاختبارين تم الالتزام بالآتي:

- **هدف الاختبار:** هدف الاختبار تقييم طلبة جامعة العادل للمستوى الأول من خلال بناء نوعين من الاختبارات الإلكترونية (الاختبار الإلكتروني الخطي - الاختبار الإلكتروني التكيفي) والمقارنة بينهم وفقا للمحتوى التعليمي المقرر.
- **تحديد مفردات الاختبار:** من خلال جدول المواصفات التي إعداده سابقًا.
- **تصميم بنك الاسئلة:** تم تصميم بنك أسئلة يحوي (91) سؤال من نوع أسئلة الصواب والخطأ والاختيار من متعدد حيث روعي عند بناء الأسئلة صياغتها صياغة لغوية سليمة، وتم وضع البرنامج في Server، وتم ربط بقية الأجهزة به بحيث نستطيع مراقبة سير الاختبار وتصحيح الأسئلة ألياً، علمًا بأن النظام يتيح التحكم في أنواع الأسئلة ومستوى صعوبته وتحديد الاستجابات والتفاعلات، وتم تحديد خصائصها السيكومترية.
- وقد تم بناء قاعد بيانات عن طريق برنامج الاكسس يحتوي على رقم السؤال ونوعه ومستوى الصعوبة.

المعالجة الإحصائية:

تمر الدراسة الحالية بمجموعة مختلفة من التحليلات الإحصائية من أجل الإجابة على الأسئلة البحثية وهي:

أ- التحقق من افتراضات نظرية الاستجابة للمفردة:

– التحقق من افتراضية أحادية البعد:

للتحقق من نظرية الاستجابة المفردة تم اخضاع بنك الأسئلة للتحليل العاملي بطريقة المكونات الأساسية باستخدام برنامج spss25 إذ تعد المكونات الأساسية الناتجة عن التحليل العاملي مؤشرًا مقبولًا احاديًا عندما لا تقل نسبة المؤشر للعامل المفسر بالتراكمي عن (20%) للعوامل المستخلصة وفق مقياس leckase.
وذلك للتعرف على مدى تحقق أحادية البعد كونها أحد الافتراضات الأساسية التي تقوم عليها نظرية الاستجابة للمفردة.

جدول (3) قيم التباين المفسر للعوامل الثمانية مقارنة للتفسير التراكمي للعامل الأخير

العامل	نسبة التباين المفسر	نسبة التباين التراكمي المفسر
1	15.286	15.286
2	9.816	25.102
3	7.692	32.794
4	8.878	41.673
5	8.297	49.970
6	4.539	54.509
7	6.569	61.078
8	5.444	66.523

المتوسط	8.315125	43.36688

ثم نجد المتوسط لكلا العمودين ثم نطرح المتوسط التباين التراكمي من التباين المفسر يكون الناتج (35.05%) أي أن بنك الأسئلة يحقق محك leckase.

– **التحقق من افتراض الاستقلال الموضعي:**

ويقصد الاستقلال الموضعي أن احتمال الإجابة الصحيحة للسؤال يكون مستقل عن ناتج إجابته وعن أي سؤال أخرى في الاختبار عند ضبط كل من القيمة التقديرية لقدرته وصعوبة السؤال وعليه فأن الاستقلال الموضعي يعني عدم ترابط الأسئلة في الاختبار وهذا ما هدفت إليه النظرية الحديثة. حيث يمكن التحقق من افتراض الاستقلال الموضعي من خلال مؤشر (Q3) وهوعن معاملات الارتباط بين البواقي لكل زوج من أزواج مفردات الاختبار، ويتحقق الاستقلال الموضعي إذا كانت جميع معاملات الارتباط بين أزواج مفردات الاختبار أقل من نقطة القطه وهي القيمة المطلقة للعدد (0.2)، وكان متوسط قيم مؤشر(Q3) أقل من القيمة المتوقعة للمتوسط الحسابي للتوزيع العيني لهذا المؤشر.

– **التحقق من افتراض التحرر من السرعة:**

لتحيق هذا الفرض تم إعطاء الوقت الكافي للطلاب للإجابة على أسئلة الاختبار بحيث لا يؤثر عامل السرعة على أدائهم بحيث يمكن ارجاع إجاباتهم الخطاء إلى قدرتهم المحدودة وليس إلى عدم كفاية الوقت. حيث أنهى جميع الطلاب الإجابة على الاختبار ضمن الوقت المحدد للإجابة وهو ساعة.

ب- التحقق من فرضيات البحث:

يسعى البحث للإجابة على الفرضيات التالية:

- لا توجد فروق ذات دلاله إحصائية عند مستوى $\alpha = 0.05$بين متوسطي تقدير القدرة والخطأ المعياري لتقدير القدرة (كل على حده) للاختبار التكيفي والاختبار الخطي باستخدام طريقة الأرجحية القصوى لتقدير القدرة.
- لا توجد فروق ذات دلاله إحصائية عند مستوى $\alpha = 0.05$بين متوسطي درجات طلبة المجموعة التجريبية التي تم تقييمها باستخدام الاختبار التكيفي والمجموعة الضابطة التي تم تقييمها باستخدام الاختبار الخطي.
- لا توجد فروق ذات دلاله إحصائية عند مستوى $\alpha = 0.05$بين متوسطي درجات طلبة المجموعة التجريبية التي تم تقييمها باستخدام الاختبار التكيفي وبين المجموعة الضابطة التي تم تقييمها باستخدام الاختبار الخطي تعزى لمتغير الجنس (ذكور/أناث).
- لا توجد فروق ذات دلاله إحصائية عند مستوى $\alpha = 0.05$بين متوسطي درجات طلبة المجموعة التجريبية التي تم تقييمها باستخدام الاختبار التكيفي وبين المجموعة الضابطة التي تم تقييمها باستخدام الاختبار الخطي تعزى إلى متغير التخصص.

3.عرض النتائج والمناقشة:

عرض ومناقشة نتائج البحث وتفسيرها:

سؤال البحث الرئيسي

- **ما فاعلية تصميم نظام اختبارات إلكترونية تكيفية على تقييم القدرة المعرفية وتعزيز التحصيل الدراسي؟**

تمت الإجابة عن سؤال البحث الرئيسي عن طريق الإجابة عن أسئلة البحث الفرعية الثلاثة التالية:

1. **ما معايير بناء وتصميم الاختبارات التكيفي الإلكترونية؟**
2. **ما الفروق بين الاختبارات الخطية والاختبار التكيفي المحوسب في متوسط تقدير القدرة؟**
3. **ما الفروق بين الاختبارات الخطية والاختبار التكيفي المحوسب في تعزيز التحصيل الدراسي؟**

❖ **الإجابة على سؤال البحث الأول:**

بعد الانتهاء من كتابة الإطار النظري للبحث وأيضًا بعد الانتهاء من بناء الاختبار وضبط المعايير يكون قد تم الإجابة على السؤال الأول.

❖ **الإجابة على سؤال البحث الثاني:**

استخدم البحث طريق الأرجحية القصوى لتقدير القدرة وقاعدة أنهاء تم تحديدها بدرجة معيارية، وخطأ معياري وقدره (0.25)، وللإجابة على هذا السؤال تم اختبار صحة الفرضية القائلة:

- **الفرضية الأولى:**

" لا توجد فروق ذات دلاله إحصائية عند مستوى $\alpha = 0.05$ بين متوسطي تقدير القدرة والخطأ المعياري لتقدير القدرة (كل على حده) للاختبار التكيفي والاختبار الخطي باستخدام طريقة الأرجحية القصوى لتقدير القدرة "

من خلال اختبار T لعينتين مرتبطين كما يتضح من الجدول (4)

جدول (4) يبين اختبار T لعينتين مرتبطين

الاختبار	القياس	المتوسط الحسابي	الانحراف المعياري	قيمة t	مستوى الدلالة	النتيجة الإحصائية
تقدير القدرة	الخطي	5.29	1.54	0.06	0.980	غير دالة
	التكيفي	10.57	3.52			
الخطأ المعياري لتقدير القدرة	الخطي	0.08	0.01	9.98	0.000	دالة
	التكيفي	0.29	0.001			

حيث يتضح من الجدول (4) الآتي:

1- أن قيمة ت المحسوبة للفروق بين متوسطي تقدير القدرة بين للاختبار الخطي والاختبار التكيفي كانت (0.06) وكانت مستوى الدالة لهما (0.98) وهي أكبر من مستوى الدالة (0.05) بالتالي تكون الفرضية غير دالة احصائيا، بالتالي قبول الفرضية الصفرية:

- لا توجد فروق ذات دلاله إحصائية عند مستوى $\alpha = 0.05$بين متوسطي تقدير القدرة بين الاختبار التكيفي والاختبار الخطي باستخدام طريقة الأرجحية القصوى لتقدير القدرة. **أي ان مستوى القدرة بين الاختبارين لا فروق بينهما.**

2- أن قيمة ت المحسوبة للفروق بين مستوى الخطأ المعياري لتقدير القدرة بين الاختبار الخطي والاختبار التكيفي وهو (9.97) ومستوى الدالة (0.00) وهي أقل من مستوى الدالة (0.05) بالتالي تكون الدالة، بالتالي رفض الفرضية الصفرية وقبول البديلة:

- توجد فروق ذات دلاله إحصائية عند مستوى $\alpha = 0.05$بين متوسطي الخطأ المعياري لتقدير القدرة بين الاختبار التكيفي والاختبار الخطي باستخدام طريقة الأرجحية القصوى لتقدير القدرة لصالح الاختبار التكيفي. **وقد أتفق ما تم التوصل مع إليه مع عدد من الدراسات منها:**

- دراسة نور الدين [26] حيث أتتفق إلى عدم وجود فروق ذات دلالة إحصائية بين متوسطي تقدير القدرة التحصيلية للاختبار الخطي الورقي والاختبار التكيفي المحوسب باستخدام طريقة الأرجحية القصوى لتقدير القدرة وتحديد قاعدة إنهاء الاختبار بـ (20) مفردة، كما أن متوسط الخطأ المعياري لتقدير القدرة للاختبار الإلكتروني التكيفي كان أقل بصورة دالة عند مستوى دلالة (0,05) عن متوسط الخطأ المعياري لتقدير القدرة للاختبار الخطي الورقي.دالة عند مستوى دلالة (0,05) عن متوسط الخطأ المعياري لتقدير القدرة للاختبار الخطي الورقي.
- دراسة. Samsudin et al [35] الذي توصل إلى كفاءة الاختبار التكيفي المحوسب في قياس أداء طلاب الصف الثامن على الاختبارات الدولية في مادة العلوم (TIMSS)استخدام طريقة الأرجحية القصوى لتقدير القدرة وتحديد قاعدة إنهاء الاختبار بـ (20) مفردة.
- ودراسة Ozyurt and Ozyurt [28] الذي توصل إلى أن الاختبارات التكيفي المحوسبة والتي تم إعدادها في وحدة الاحتمالات من مادة الرياضيات لطلاب الصف الحادي عشر تتمتع بثبات مرتفع باستخدام طريقة الأرجحية القصوى لتقدير القدرة وتحديد قاعدة إنهاء الاختبار بـ (15) أو (20) مفردة.

وقد دلت النتائج السابقة على أن:

استخدام طريقة الأرجحية القصوى لتقدير القدرة في للاختبار الإلكتروني التكيفي يعطي نفس مستوى تقدير القدرة التي يعطيها الاختبار الخطي إلا أن:

- الاختبار الإلكتروني التكيفي أكثر دقة في تقدير القدرة لأنه يقدم أخطاء معيارية أقل من الاختبار الخطي.
- تعتمد طريقة الأرجحية القصوى على اختيار المفردات في المراحل الأولية الاختبار الإلكتروني التكيفي من خلال قاعدة الخطوة، حيث تحدد مقدار الانتقال على متصل القدرة لتحديد قدرة المختبر لإعطائه مفردة جديدة تتناسب مع القدرة الجديدة له حتى يتم الوصول إلى نمط الاستجابة الصحيحة والخاطئة.
- **الإجابة على سؤال البحث الثالث:**

استخدم البحث طريق الأرجحية القصوى لتقدير القدرة وقاعدة أنهاء تم تحديدها بدرجة معيارية، وخطأ معياري وقدره (0.25)، وللإجابة على هذا السؤال تم اختبار صحة الفرضيات القائلة:

- **الفرضية الثانية:**

" لا توجد فروق ذات دلاله إحصائية عند مستوى $\alpha = 0.05$ بين متوسطي درجات طلبة المجموعة التجريبية التي تم تقييمها باستخدام الاختبار التكيفي والمجموعة الضابطة التي تم تقييمها باستخدام الاختبار الخطي".

ولتحقق من نتيجة هذه الفرضية تم استخدام اختبار t لعينتين مرتبطتين بين متوسطي درجات طلبة المجموعة التجريبية التي تم تقييمها باستخدام الاختبار التكيفي والمجموعة الضابطة التي تم تقييمها الاختبار الخطي كما يتضح في جدول(5):

الجدول (5) اختبار t لعينتين مرتبطتين

الاختبار	المتوسط الحسابي	الانحراف المعياري	الخطأ المعياري	قيمة t	درجة الحرية	مستوى الدلالة
التكيفي	6.62	1.54	0.11	5.308		
الخطي	6.09	1.81	0.07		199	0.000

حيث يتضح من الجدول (5) الآتي:

أن المتوسط الحسابي لدرجات طلبة المجموعة التجريبية التي تم تقييمها باستخدام الاختبار التكيفي بلغت (6.62) والمتوسط الحسابي للمجموعة الضابطة التي تم تقييمها باستخدام الاختبار الخطي بلغت (6.09) وكانت الفروقات لصالح الاختبار التكيفي بفارق (0.53).

وأن قيمة ت المحسوبة لعينين مرتبطتين بلغت (5.308) وهي أكبر من الجدولية وفق درجة الحرية (199) كون عدد العينة (200) طالب وطالبة والتي بلغت (1.96) وكذلك مستوى الدلالة (0.000) وهي اقل من مستوى الدلالة (0.05) **بالتالي نرفض الفرضية الصفرية ونقبل الفرضية البديلة وهي:**

- توجد فروق ذات دلاله إحصائية عند مستوى $\alpha = 0.05$ بين متوسطي درجات طلبة المجموعة التجريبية التي تم تقييمها باستخدام الاختبار التكيفي والمجموعة الضابطة التي تم تقييمها باستخدام الاختبار الخطي ولصالح الاختبار التكيفي.

ويفسر ذلك بأن الاختبارات الإلكترونية التكيفية قد توافقت بشكل كبير مع الطلبة باختلاف الفروق الفردية لديهم، حيث أشتمل بنك الأسئلة على مجموعة متنوعة ومتعددة المستويات، والتي بدورها تقدم الاسئلة في ضوء الفروق الفردية لكل طالب، فيشعر ذوي القدرة الأقل من المتوسط نتيجة استجابته الصحيحة على الأسئلة المقدمة ومن ثم التفاعل بشكل أكبر مع الاستفادة من التغذية الراجعة المقدمة من خلال الاختبارات الالكترونية التكيفية، بينما الطلبة ذوي القدرة المتوسط يجدون تنوع الأسئلة ومناسبة مستوى تعقيدها لقدرتهم ، كما أن ذوي القدرة الأعلى من المتوسط يشعرون بالارتياح وعدم الملل نتيجة وجود تحدي في الأسئلة المقدمة عبر الاختبارات وبالتالي يزيد حماسهم نحو الاستمرار في التحدي والتعلم.

وقد اشارات عدد من الدراسات إلى هذه النتيجة منها:

- ما إشارات إليه دراسة [45] حيث أكدت على وجود فرق دال إحصائياً عند مستوى (0.05) بين متوسط مستوى درجات الطلاب في التطبيق القبلي والبعدي لصالح التطبيق البعدي (الاختبار التكيفي).
- ودراسة [25] التي استهدفت أثر الاختلاف بين الاختبار التكيفية والخطية في دقة تقدير قدرات المتعلم، وقد أظهرت النتائج أن الاختبارات التكيفية أدق من الاختبارات الإلكترونية الخطية في قياس القدرة الفعلية.
- أيضاً دراسة [46] والتي اسفرت نتائجها على وجود فرق دال إحصائياً بين درجات الطلاب لصالح التطبيق البعدي للاختبار التحصيلي المعرفي لمقرر الحاسب الآلي وأمن البيانات .
- دراسة أحمد عبد النبي والتي أشارت نتائجها إلى فاعلية نمط الاختبارات البنائية التكيفية فيما يخص الجانب المعرفي[32] .

- **الفرضية الثالثة:**

"لا توجد فروق ذات دلاله إحصائية عند مستوى $\alpha = 0.05$ بين متوسطي درجات طلبة المجموعة التجريبية التي تم تقييمها باستخدام الاختبار التكيفي وبين المجموعة الضابطة التي تم تقييمها باستخدام الاختبار الخطي تعزى لمتغير الجنس (ذكور/أناث)"

اختبار معنوية الفروق بين أفراد عينة الدراسة يعزو لمتغير الجنس، تم استخدام اختبار(t)لاختبار الفروق بين متوسطات لعينتين مستقلتين (ذكور، اناث):

جدول رقم (6) اختبار (t) لاختبار الفروق بين متوسطات لعينتين مستقلتين

الاختبار	الجنس	المتوسط الحسابي	العينة	الانحراف المعياري	قيمة t	درجة الحرية	مستوى الدلالة Sig	النتيجة الإحصائية
التكيفي	ذكور	6.23	52	1.345	-3.938-	198	0.000	دالة
	إناث	7.06	47	1.636				
الخطي	ذكور	5.74	52	1.618	-3.022-	198	0.003	دالة
	اناث	6.50	47	1.939				

يتضح من الجدول (6) الآتي:

- توجد فروق ذات دلالة احصائية عند مستوى دالة (0.05) بين درجات طلبة المجموعة التجريبية التي تم تقييمها باستخدام الاختبار التكيفي بين الذكور والإناث حيث بلغت t المحسوبة (3.938) وهي أكبر من الجدولية (1.96) وكذلك مستوى الدلالة (0.000) وهي أقل من مستوى الدلالة (0.05) ولذلك توجد فروق وكانت الفروقات المعنوية لصالح الإناث كونهما أكثر متوسط حسابي وكانت بفارق متوسط (0.83)، وهذا يعزو أن الطالبات الاناث كن أكثر اهتماماً وحرصاً من الذكور.
- توجد فروق ذات دلالة احصائية عند مستوى دالة (0.05) بين درجات طلبة المجموعة الضابطة التي تم تقييمها باستخدام الاختبار الخطي بين الذكور والإناث حيث بلغت ت المحسوبة (3.022) وهي أكبر من الجدولية (1.96) وكذلك مستوى الدلالة (0.003) وهي أقل من مستوى الدلالة (0.05) وكانت الفروقات المعنوية لصالح الإناث كونهما أكثر متوسط حسابي وكانت بفارق متوسط (0.76) وهذا يعزو أن الطالبات الاناث كن أكثر اهتماماً وحرصاً من الذكور

لذلك نرفض الصفرية ونقبل بالبديلة وهي:

- توجد فروق ذات دلاله إحصائية عند مستوى $\alpha = 0.05$ بين متوسطي درجات طلبة المجموعة التجريبية التي تم تقييمها باستخدام الاختبار التكيفي وبين المجموعة الضابطة التي تم تقييمها باستخدام الاختبار الخطي تعزى لمتغير الجنس (ذكور/أناث).
- حيث كانت نتيجة الاختبارين لصالح الإناث كونهما أكثر متوسط حسابي، وهذا يعزو أن الطالبات الاناث كن أكثر اهتماماً وحرصاً من الذكور في التحصيل العلمي ويؤكد ذلك ظهور نفس الفارق في الاختبارين.

- **الفرضية الرابعة:**

"لا توجد فروق ذات دلاله إحصائية عند مستوى $\alpha = 0.05$ بين متوسطي درجات طلبة المجموعة التجريبية التي تم تقييمها باستخدام الاختبار التكيفي وبين المجموعة الضابطة التي تم تقييمها باستخدام الاختبار الخطي تعزى إلى متغير التخصص ".

لاختبار معنوية الفروق بين أفراد اجابات عينة الدراسة يعزو لمتغير القسم تم استخدام تحليل التباين الأحادي (One Way ANOVA)، حيث تم الاعتماد على مستوى دلالة ($\alpha = 0.05$)، وتعد الفروق دالة إحصائيًا إذا كانت قيمة مستوى دلالة الاختبار (Sig.) أصغر من مستوى الدلالة الذي اعتمد ها البحث في دراستها، والعكس صحيح.

جدول (7) يوضح دلالة الفروق بين إجابات عينة الدراسة وفق متغير القسم

نوع الاختبار	مصدر التباين	مجموع المربعات	درجات الحرية	متوسط المربعات	قيمة ف	الدلالة	الدلالة الإحصائية
الخطي	بين المجموعات	53.940	3	17.980	5.877	0.001	دالة
	داخل المجموعات	599.651	196	3.059			
التكيفي	بين المجموعات	79.821	3	26.607	13.259	0.000	دالة
	داخل المجموعات	393.323	196	2.007			

يتضح من الجدول رقم (7) الآتي:

- توجد فروق ذات دلاله إحصائية عند مستوى $\alpha = 0.05$ بين متوسطي درجات طلبة المجموعة الضابطة التي تم تقييمها باستخدام الاختبار الخطي تعزو لمتغير التخصص كون مستوى الدلالة (0.001) وهي اقل من مستوى (0.05) وبالتالي تكون دالة وكانت الفروقات بحسب اختبار (LSD) لصالح قسم الحاسوب بمقدار (2.31).
- توجد فروق ذات دلاله إحصائية عند مستوى $\alpha = 0.05$ بين متوسطي درجات طلبة المجموعة التجريبية التي تم تقييمها باستخدام الاختبار التكيفي تعزو لمتغير التخصص كون مستوى الدلالة (0.000) وهي أقل من مستوى (0.05) وبالتالي تكون دالة وكانت الفروقات بحسب اختبار (LSD) لصالح قسم الحاسوب بمقدار (2.49) وقد يعزو ذلك ان طلبة التخصص الحاسوب لديهم معرفة ودراية أكثر من غيرهم في استخدام الحاسوب كون الاختبار التكيفي هو اختبار محوسب وايضًا لأن مقرر الاختبار كانت في مادة مهارات الحاسوب وهي من المواد التي لديها اتصال مباشر بهذا التخصص.

ولذلك نرفض الصفرية ونقبل بالبديلة وهي:

- توجد فروق ذات دلاله إحصائية عند مستوى $\alpha = 0.05$ بين متوسطي درجات طلبة المجموعة التجريبية التي تم تقييمها باستخدام الاختبار التكيفي وبين المجموعة الضابطة التي تم تقييمها باستخدام الاختبار الخطي تعزى إلى متغير التخصص.

4. الاستنتاجات والتوصيات والمقترحات:

تم عرض تحليل نتائج الدراسة وتفسيرها مستعين في ذلك بالمعالجة الإحصائية لاستجابات العينة للأدوات الدراسة وصولًا إلى الإجابة على أسئلة الدراسة وفرضياتها وسيتم عرض ملخص لعصارة ما توصل إليه البحث من النتائج في ضوء الفرضيات التي سعت لتحقيق أهداف الدراسة، وبناء على ذلك قدم البحث بعض التوصيات والمقترحات والتي

لها دور في تطوير أساليب التقييم الإلكتروني ورفع مستوى التحصيل لدى الطلبة.

4.1 ملخص الاستنتاجات:

توصلت الدراسة الحالية إلى النتائج التالية:

1. لا توجد فروق ذات دلاله إحصائية عند مستوى $\alpha = 0.05$ بين متوسطي تقدير القدرة بين الاختبار التكيفي والاختبار الخطي باستخدام طريقة الأرجحية القصوى لتقدير القدرة. أي ان مستوى القدرة بين الاختبارين لا فروق بينهما.
2. توجد فروق ذات دلاله إحصائية عند مستوى $\alpha = 0.05$ بين متوسطي الخطأ المعياري لتقدير القدرة بين الاختبار التكيفي والاختبار الخطي باستخدام طريقة الأرجحية القصوى لتقدير القدرة لصالح الاختبار التكيفي.
3. توجد فروق ذات دلالة إحصائية عند مستوى $\alpha = 0.05$ بين متوسطي درجات طلبة المجموعة التجريبية التي تم تقييمها باستخدام الاختبار التكيفي، والمجموعة الضابطة التي تم تقييمها باستخدام الاختبار الخطي ولصالح الاختبار التكيفي.
4. توجد فروق ذات دلاله إحصائية عند مستوى $\alpha = 0.05$ بين متوسطي درجات طلبة المجموعة التجريبية التي تم تقييمها باستخدام الاختبار التكيفي، وبين والمجموعة الضابطة التي تم تقييمها باستخدام الاختبار الخطي تعزى لمتغير الجنس (ذكور/أناث).

حيث كانت نتيجة الاختبارين لصالح الإناث كونهما أكثر متوسط حسابي، وهذا يعزو أن الطالبات الاناث كن أكثر اهتماماً وحرصاً من الذكور.

5. توجد فروق ذات دلاله إحصائية عند مستوى $\alpha = 0.05$ بين متوسطي درجات طلبة المجموعة التجريبية التي تم تقييمها باستخدام الاختبار التكيفي، وبين المجموعة الضابطة التي تم تقييمها باستخدام الاختبار الخطي تعزى إلى متغير التخصص. وقد يعزو ذلك لأن مقرر الاختبار كان في مساق مهارات الحاسوب وعلى الرغم أن هذا المساق كان إحدى المتطلبات الجامعية إلا أن طلبة قسم الحاسوب لديهم معرفة ودراية أكثر من غيرهم فيه، وعلى الرغم من سهوله واجهة الاختبار إلا أن أيضًا طلبة قسم الحاسوب كانوا أكثر قدره على استخدام الاختبار مقارنةً بباقي التخصصات.

XI. التوصيات:

توصلت نتائج الدراسة إلى فاعلية لاختبارات الإلكترونية التكيفية في تقييم القدرة المعرفية للطلبة في التعليم، ودورها في رفع التحصيل لدى طلبة الجامعة، وفي ضوء هذه النتائج يقترح البحث التوصيات التالية:

1. ضرورة تطور أساليب التقويم باستخدام الاختبارات الإلكترونية في التعلم، متمثلة في الاختبارات التكيفية الإلكترونية، ويقترح البحث الحالي أن يتم البدء أولًا بتحويل الاختبارات التي تقدم لطلاب السنة الأول كمتطلبات للجامعة، مثل: مواد الثقافة الإسلامية، واللغة العربية العامة، اللغة الإنجليزية مهارات الحاسب إلى اختبارات الإلكترونية تكيفية ثم التوسع في ذلك وصولًا إلى إجراء الاختبارات التكيفية الإلكترونية لجميع المساقات الجامعية في جميع الجامعات، وذلك للاستفادة من مزايا الاختبار التكيفي.
2. الاستفادة من التصميم المستخدم في الاختبارات الإلكترونية التكيفية للتغلب على مشكلة القلق المصاحب للاختبارات والغش لدى الطلبة.
3. نشر ثقافة الاختبارات الإلكترونية التكيفية بين معلمي التعليم قبل الجامعي.
4. دعوة الخبراء في مجالي القياس والتقويم والتعلم الإلكتروني لإقامة ورش عمل في مجال الاختبارات الإلكترونية التكيفية وآليات الاستفادة منها بالعملية التعليمية.
5. توفير بيئة صفية مناسبة لعقد الاختبارات المحوسبة وجعله ملائمة لكل من الطلبة والمدرسين
6. تضمين مهارات تصميم وبناء الاختبارات الإلكترونية التكيفية ضمن برامج تنمية قدرات أعضاء هيئة التدريس.
7. التوسع في إجراء الدراسات حول الاختبارات الإلكترونية التكيفية في كافة المراحل التعليمية وفي مختلف الثقافات.

XII. المقترحات:

وختامًا حتى يمكن تفعل تطبيقات الاختبار التكيفي المحوسب لتطور أساليب التقويم للاختبارات الإلكترونية في التعلم الجامعي، اقترح البحث ما يلي:

1. عمل دراسات مشابهة على عدد من الكليات الأخرى والخروج باقتراحات ونتائج متفق عليها من أجل تحسين جودة تقديم الاختبارات المحوسبة لما فيه من مصلحة للطلبة والمدرسين أنفسهم.
2. أن تعمل الجامعات على إنشاء مراكز متخصصة في القياس والتقويم، تتوفر بها البنية التحتية، والبرامج الخاصة بتطوير وتحليل الاختبارات وتضم كفاءات متخصصة للإشراف على تطوير منظومة التقويم في كل جامعة.
3. عمل دراسات خاصة بالخصائص الفنية للاختبارات المحوسبة وفق النظرية الجديدة في القياس؛ لكي يتم تقديم اختبارات مناسبة للطلبة تتمتع بخصائص سيكومترية مناسبة.
4. إجراء مزيد من الدراسات حول مدى تقبل الطلاب مستحدثات تكنولوجية بما فيها الإلكترونية التكيفية.
5. إجراء دراسات لحل مشكلة التخمين في الأسئلة المحوسبة؛ وذلك أثناء القيام بغربلة الفقرات المقدمة للطلبة في كل مادة دراسية؛ حتى يزيد من مستوى فاعلية الفقرات التي تقدم للطلبة في الاختبارات المحوسبة.
6. إجراء دراسات للمقارنة بين نتائج الطلاب على الاختبارات الورقية والاختبارات الإلكترونية، والإلكترونية التكيفي.

XIII. المراجع العربية والأجنبية:

[1] ف. ع. الخزي و م. أ. الزكري، "تكافؤ الاختبارات الإلكترونية مع الاختبارات الورقية في التحصيل الدراسي: دراسة تجريبية على طلبة كلية التربية بجامعة الكويت"، مجلة دراسات الخليج والجزيرة العربية، مج. 37، ص. 167–198، 2011.

[2]N. Alruwais, G. Wills, and M. Wald, "Advantages and challenges of using e-assessment," International Journal of Information and Education Technology, vol. 8, no. 1, pp. 34-37, 2018.

[3] ر. س. البربري، "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، 2020.

[4] آ. م. معزي البياضة، "بناء اختبار التكيفية للقدرة الرياضية للصف السابع الأساسي وفق الاستراتيجية الهرمية باستخدام نظرية الاستجابة للفقرة"، رسالة ماجستير، عمادة الدراسات العليا، جامعة مؤته، الأردن، 2011.

[5]Khan and M. Jawaid, "Technology Enhanced Assessment (TEA) in COVID 19 Pandemic," Pakistan Journal of Medical Sciences, vol. 36, no. COVID19-S4, pp. S108–S110. doi:

10.12669/pjms.36.COVID19-S4.2795, accessed on: 21 January 2022.
[6] ع. إسماعيل، "قلق الاختبارات الالكترونية وعلاقته بالأداء في نظر عينة من طالبات كلية الأميرة عالية بجامعة البلقاء التطبيقية - الأردن"، المجلة التربوية، كلية التربية، جامعة سوهاج، مج. 52، ص. 799–823، 2018.
[7] J. Rossiter, "Evaluation of software tools for formative assessment of control topics," IFAC Papers OnLine, vol. 52, no. 9, pp. 292–297, 2019.
[8] D. J. Weiss, "Item banking, test development, and test delivery," in The APA Handbook on Testing and Assessment, K. F. Geisinger, Ed. Washington, DC: American Psychological Association, 2013.
[9] Y. Zheng and H. Chang, "On-the-fly assembled multistage adaptive testing," Applied Psychological Measurement, vol. 39, no. 2, pp. 104-118, 2015.
[10] A.J. Martin and G. Lazendic, "Computer- adaptive testing: Implications for students' achievement, motivation, engagement, and subjective test experience," Journal of Educational Psychology, vol. 110, no. 1, pp. 27, 2018.
[11] D. Magis, D. Yan, and A. A. von Davier, Computerized Adaptive and Multistage Testing With R: using packages catr and mstr, Cham, Switzerland: Springer International Publishing, 2017.
[12] D. Robert, G. David, and B. Benjamin, "Computerized Adaptive Tests for Rapid and Accurate Assessment of Psychopathology Dimensions in Youth," Journal of the American Academy of Child & Adolescent Psychiatry, vol. 9, pp. 122–151, 2020.
[13] Q. Liu, Y. Zhuang, H. Bi, Z. Huang, W. Huang, J. Li, J. Yu, Z. Liu, Z. Hu, and Y. Hong, "Survey of Computerized Adaptive Testing: A Machine Learning Perspective," arXiv preprint arXiv:2404.00712, 2024. [Online]. Available: arxiv.org.
[14] A. Franco, "The Adaptive Quiz Activity," Retrieved from: https://moodle.org/plugins/mod_adaptivequiz, accessed on: 21 January 2022.
[15] F. Martinez-Plumed, R. B. Prudêncio, A. Martínez-Usó, and J. Hernández-Orallo, "Item response theory in AI: Analysing machine learning classifiers at the instance level," Artificial Intelligence, vol. 271, pp. 18-42, 2019.
[16] إ. محاسنة، "القياس النفسي في ظل النظرية التقليدية والنظرية الحديثة"، عمان: دار جرير للنشر والتوزيع، 2013.
[17] خ. الدامغ و هـ. الهاجري، "تصميم الاختبارات الإلكترونية لمتعلمي اللغة العربية -لغة ثانية-," مجلة كلية الدراسات الإسلامية والعربية للبنات بالإسكندرية, vol. 2, no. 35, pp. 823-875, 2019.
[18] Bider and E. Perjons, "Design science in action: developing a modeling technique for eliciting requirements on business process management (BPM) tools," Software & Systems Modeling, vol. 14, no. 3, pp. 1159–1188, 2015.
[19] C. Youn Young and M. Cayce, "Development of Adaptive Formative Assessment System Using Computerized Adaptive Testing and Dynamic Bayesian Networks," Applied Sciences, vol. 10, pp. 2-1, 2020.
[20] M. Boussakuk et al., "Design of Computerized Adaptive Testing Module into our Dynamic Adaptive Hypermedia System," International Journal of Emerging Technologies in Learning (iJET), vol. 16, no. 18, pp. 113-128, 2021. doi: 10.3991/ijet.v16i18.23841.
[21] أ. محمد إبراهيم، "الاختبارات المحبوكة المحوسبة: تطوير لبنوك الأسئلة," مجلة ابداعات تربوية, رابطة التربويين العرب, no. 3, ISSN: 2536-9407, أكتوبر 2017.
[22] C. Wang and H. Lu, "Mediating Effects of Individuals' Ability Levels on the Relationship of Reflective-Impulsive Cognitive Style and Item Response Time in CAT," Educational Technology & Society, vol. 21, no. 4, pp. 89–99, 2018.
[23] Dammas, "Investigate Students' Attitudes toward Computer Based Test (CBT) at Chemistry Course," Archives of Business Research, vol. 4, no. 6, pp. 58-71, 2016.
[24] م. ع. خميس، "التحول الرقمي: قضايا ومشكلات التعلم"، المؤتمر العلمي السادس عشر للجمعية العربية لتكنولوجيات التربية، كلية الدراسات التربوية، جامعة القاهرة، 2019.
[25] العودة و ع. عبيدات، "فاعلية الاختبار التكيفي المحوسب في تقدير القدرة، باستخدام مصفوفة رافن"، دراسات العلوم التربوية، مج. 40، ع. 3، ص. 1602–162، 2013.
[26] نور الدين، "فاعلية القياس التكيفي المحوسب في قياس تحصيل طلاب الجامعة"، المجلة السعودية للعلوم النفسية، جامعة الملك سعود، الجمعية السعودية للعلوم التربوية والنفسية، ع. 64، ص. 29–47، 2019.
[27] ر. فرحات، "الفروق بين الاختبارات المحوسبة الموائمة والخطية في تقدير القدرة وفقاً لنظرية الاستجابة للمفردة لطلبة كلية التربية جامعة الأزهر"، التربية، أكتوبر 2021.
[28] H. Özyurt and Ö. Özyurt, "Ability level estimation of students on probability unit via computerized adaptive testing," Eurasian Journal of Educational Research, vol. 58, pp. 27-44. doi: 10.14689/ejer.2015.58.7.
[29] د. جرجس، "أثر الاختبارات البنائية التكيفية في تنمية التحصيل الدراسي المرتبط بالبرمجة لدى طلاب تكنولوجيا التعليم"، كلية التربية النوعية، جامعة بنها، 2021.
[30] R. James, "Tertiary student attitudes to invigilated, online summative examinations,"

International Journal of Educational Technology, vol. 13, no. 19, pp. 2-13, 2016.
[31] ص. د. علام، "نماذج الاستجابة للمفردة الاختبارية أحادية البعد ومتعددة الأبعاد وتطبيقاتها في القياس النفسي والتربوي"، دار الفكر العربي، 2005.
[32] أ.عبد النبي عبد الملك النظير، "التفاعل بين نمط أسئلة الاختبار إلكتروني التكيفي (ثنائية/متعددة) الاستجابة ومركز التحكم (داخلي/ خارجي) عبر المنصات التعليمية الرقمية وأثره في تنمية مهارة تطوير البرمجيات التعليمية واتخاذ القرار والكفاءة الذاتية لطلاب تكنولوجيا المعلومات"، كلية التربية النوعية، جامعة عين شمس، الجمعية المصرية تكنولوجيا المعلومات، 2020.
[33] م. التقي، "النظرية الحديثة في القياس"، عمان: دار المبسرة للنشر والتوزيع، 2009.
[34] أ. علي الشهري، "تأثير قواعد التعليق ذات الطول الثابت على إدارة الاختبارات التكيفية"، المجلة العلمية لكلية التربية بجامعة الوادي الجديد، 2023.
[35] M. Samsudin, T. SomChut, and M. Ismail, "Evaluating computerized adaptive testing efficiency in measurement student's performance in science timss," Indonesian Journal of Science Education, vol. 8, no. 4, pp. 547-560, 2019.
[36] D. Cisar, D. Radosav, B. Markoski, R. Pinter, and P. Cisar, "Computer adaptive testing of student knowledge," Acta Polytechnic Hungarica, vol. 7, no. 4, pp. 139-152, 2010.
[37] D. Weiss, "Computerized adaptive testing for effective and efficient measurement in counseling and education," Measurement and Evaluation in Counseling and Development, vol. 37, no. 2, pp. 70-84, 2004.
[38] M. K. O. Lee, C. M. K. Cheung, and Z. Chen, "Acceptance of Internet Based learning medium: The role of extrinsic and intrinsic motivation," Information & Management, vol. 42, 2005.
[39] ز. يوسف دعنا، "بناء اختبار محبوك هرمي في الرياضيات للصف الثامن الأساسي وفق نموذج راش في نظرية السمة الكامنة"، دراسات العلوم، مج. 1، ع. 32، 2005.
[40] A. J. Leroux and B. G. Dodd, "A comparison of stopping rules for computerized adaptive screening measures using the rating scale model," Journal of Applied Measurement, vol. 15, pp. 213–226, 2014.
[41] J. M. Linacre, "Computer-Adaptive Testing: A Methodology Whose Time Has Come," Rasch Measurement Transactions, vol. 69, 2000. [Online]. Available: http://www.rasch.org/memo69.pdf
[42] E. Istiyono, W. Brams, R. Setiawan, and I. Megawati, "Developing computerized adaptive testing to measure physics higher order thinking skills of senior high school students and its feasibility of use," European Journal of Educational Research, vol. 9, no. 1, pp. 91-101, 2020.
[43] Y. Huo, "Variable-length Computerized Adaptive Testing: adaptation of the a-stratified strategy in item selection with content balancing," Ph.D. dissertation, Graduate College, University of Illinois at Urbana-Champaign, 2009.
[44] D. Magis and G. Raiche, "CatR: An R package for computerized adaptive testing," Applied Psychological Measurement, vol. 33, no. 7, pp. 576-577, 2011.
[45] أ. م. عبد الرحمن، تصميم الاختبارات, عمان: دار أسامة للنشر والتوزيع, 2011.
[46] د. ج. القس جرجس، "أثر الاختبارات البنائية التكيفية في تنمية التحصيل الدارسي المرتبط بالبرمجة لدى طلاب تكنولوجيا التعليم," كلية التربية النوعية, جامعة بنها, 2021.