نمطا تصميم الاختبار الإلكتروني التكيفي الثابت والمتغير الطول وأثرهما على خفض مستوى قلق الاختبار وتنمية الاتجاهات نحو الاختبارات الإلكترونية لدى طلاب كلية التربية

أ.م.د. رفيق سعيد إسماعيل البربرى
أستاذ تكنولوجيا التعليم المساعد
كلية التربية بشبين الكوم – جامعة المنوفية

مستخلص البحث

استهدف البحث الحالي دراسة أثر نمطي تصميم الاختبار الإلكتروني التكيفي الثابت والمتغير الطول على خفض مستوى قلق الاختبار وتنمية الاتجاهات نحو الاختبارات الإلكترونية لدى طلاب كلية التربية. وقد تم استخدام التصميم التجريبي القائم على امتداد المجموعة الضابطة ذو الاختبار القبلي البعدي Extended control group pre-posttest design. وتعرضت مجموعات البحث للمعالجات على النحو التالي: ١) المجموعة الضابطة تعرضت لاختبار إلكتروني تقليدي عدد مفرداته (٤٠) مفرده تم سحبها بشكل عشوائي من بنك أسئلة عدد مفرداته (٤٠٠) مفرده تم انشاءه على أحد نظم إدارة التعلم؛ ٢) تعرضت المجموعة التجريبية الأولى لاختبار تكيفي ذو التصميم ثابت الطول من خلال أحد المنصات المتخصصة في الاختبارات التكيفية . وقد ارتكزت بداية الاختبار على توليد وطرح مفردة متوسطة الصعوبة ثم سحب المفردات التالية من بنك الأسئلة بناء على استجابات المتعلم حتى نهاية المفردات المحدد عددها مسبقا لتحقيق شرط قاعدة انهاء الاختبار؛ ٣) تعرضت المجموعة التجريبية الثانية لاختبار تكيفي ذو التصميم متغير الطول من خلال أحد المنصات المتخصصة في الاختبارات التكيفية. وقد ارتكز الاختبار المتغير الطول على اختبار قبلي في تحديد المستوى المبدئي للمتعلم بعدد(١٢) مفرده متباينة في مستويات الصعوبة ومحتوى التعلم ومن ثم طرح المفردات التالية من حيث مستوى الصعوبة وفقا تقدير ديناميكي لاستجابات المتعلم ويتم انهاء الاختبار وفقًا طريقة الارجحية العظمى Maximum Like Hood Estimation في تقدير أداء المتعلم بشكل مستمر وفقا مقارنات بين التقدير المتوقع والاستجابة الفعلية، دونما اعتبار لعدد المفردات

المسحوبة من البنك لكل ممتحن ويتم انهاء الاختبار عند مستوى خطأ معياري (٠.٢٥) . وقد تمثلت أدوات البحث في مقياسين لمستوى قلق الاختبار والاتجاهات نحو الاختبار الإلكترونية من إعداد الباحث. أسفرت النتائج عن عدم وجود فرق ذو دلالة إحصائية بين المجموعة الضابطة التي استخدمت الاختبارات الإلكترونية الخطية مقارنة بالمجموعة التجريبية الأولى التي استخدمت الاختبارات التكيفية ثابتة الطول في خفض مستوى قلق الاختبار، بينما أظهرت نتائج المقارنة بين المجموعة الضابطة والمجموعة التجريبية الثانية التي استخدمت الاختبارات التكيفية متغيرة الطول وجود فرق لصالح المجموعة التجريبية الثانية في خفض مستوى القلق، كما أظهرت المقارنة بين المجموعة التجريبية الأولى والمجموعة التجريبية الثانية تفوق الأخيرة في خفض مستوى قلق الاختبار . كما أظهرت نتائج التطبيق البعدي لمقياس الاتجاهات على المجموعات الثلاث أن الاختبارات التكيفية ثابتة الطول أثرت بشكل أكبر في تكوين اتجاهات إيجابية لدى الطلاب مقارنة بالاختبارات الإلكترونية التقليدية، كذلك تفوقت الاختبارات التكيفية متغيرة الطول على الاختبارات الإلكترونية التقليدية الخطية في تكوين اتجاهات إيجابية لدى الطلاب نحو الاختبارات الإلكترونية . بينما تفوقت الاختبارات التكيفية ثابتة الطول على الاختبارات التكيفية متغيرة الطول في تكوين اتجاهات إيجابية لدى الطلاب نحو الاختبارات الإلكترونية .

الكلمات المفتاحية: الاختبارات التكيفية الإلكترونية ؛ متغيرات تصميم طول الاختبار التكيفي؛ قلق الاختبار؛ الاتجاهات نحو الاختبارات الإلكترونية ؛ مرحلة التعليم الجامعي؛ كليات التربية.

مقدمة

يعكس تطور أنظمة القياس والتقويم في أي نظام تعليمي قوة الأداء المؤسسي والبرامجي ودرجة ثقة المجتمع والمستفيدين Stakeholders فى تحقيق هذا النظام لمتطلبات سوق العمل وتحديات مواكبة تكنولوجيات المعلومات والاتصالات. مما دعي غالبية مؤسسات التعليم الجامعي الى اعتبار التوسع فى توظيف المستحدث التكنولوجي بشكل عام فى عمليات قياس وتقويم أداء الطلاب فى قمة أولوياتها وعنصرا أساسيًا فى خطط واستراتيجيات التحسين والتطوير المستقبلية. وتمثل الاختبارات الإلكترونية أحد أشهر أدوات التقويم التكويني Formative Evaluation والختامي Summative التي ساهمت فى التغلب على العديد من المظاهر السلبية فى الاختبارات الورقية التقليدية من حيث توفير الوقت والجهد ومستوى مناسب من الموضوعية والشفافية فى تقدير أداء المتعلم. إلا أن التطور النوعي فى نظريات القياس من جانب وتكنولوجيات تطوير ونشر وتوزيع الاختبارات الإلكترونية من جانب آخر؛ قد اظهر العديد من جوانب القصور فى الانماط المختلفة من الاختبارات الإلكترونية التقليدية خاصة التي تعتمد على التوليد العشوائي للمفردات

من بنوك أسئلة ذات قواعد بيانات هيكلية ثابته دونما اعتبار لمستوى الأداء الفعلي لمتعلم فى تحديد مستوى صعوبة أو سهولة المفردات التالية وكذا توقيت وعدد المفردات اللازمة لإنهاء الاختبار؛ مما يعكس وجود مشكلات فى تصميم بناء بنوك أسئلة الاختبارات الالكترونية التقليدية وآليات سحب المفردات منها وفقا مستوى أداء المتعلم، الامر الذي قد أدى الى ظهور قدرًا مرتفع من قلق الاختبار اثر بالسلب على أداء المتعلم وتكوين اتجاهات سالبة نحو التوسع فى الاختبارات الإلكترونية بشكل عام (Weiss, 2013, p.1).

وقد أكد عديد من الباحثين والخبراء فى تطوير نظم القياس والتقويم الإلكترونية على ضرورة الاهتمام برفع كفاءة تلك النظم من خلال تبنى نظريات حديثة فى القياس النفسي وتوظيف تكنولوجيات جديده فى بناء الاختبارات بأنواعها المختلفة مثل تكنولوجيا الذكاء الاصطناعي، وما يرتبط بها من مستحدثات مثل تعلم الآلة Machine Learning فى نمذجة أداء المتعلم Student Modelling للكشف عن الأداء الفعلي له بأقل عدد من المفردات، وأقصر مدى زمنى من خلال نظام اختباري تكيفي (CAT) Computer Adaptive Test يفّصل Tailored لتحديد مستوى قدرات الممتحن الفعلية، ويتعدى مستوى مراعاة الفروق بين مستويات أداء الأفراد إلى مراعاة الفروق بين مستويات الأداء لدى الفرد الواحد والاعتماد على التقدير الكمي والنوعي للاستجابات المتعلم على كل مفردة من مفردات الاختبار[1] (Conejo et al., 2004; Boyd, 2003; Wainer et al, 2000; Cruces and Arriaga, 2000)

كذلك أكد كلا من Magis and Raiche (2011) على ان التوسع في توظيف الاختبارات التكيفية في أنظمة القياس والتقويم في اختبارات القبول بالجامعات وتحديد المستوى وقياس القدرات ومعدلات التحصيل في المقررات الدراسة ضرورة تفرضها فاعلية تلك النوعية من الاختبارات في تقليل طول وزمن الاختبار ودقة قياس أداء الممتحنين، وكذا تقليل الضغوط عليهم اثناء الموقف الاختباري. خاصة وان العديد من الباحثين أكدوا على ان قلق الاختبار يمثل أحد اهم العوامل المؤثرة على أداء الممتحن والذي يتأثر بشكل كبير بطول الاختبار الإلكتروني (Hembree, 1998; Seipp, 1991). وقد اهتم العديد من الدراسات بمتغيرات تصميم وإنتاج ونشر الاختبارات الإلكترونية بصفة عامة مثل متغيرات تطبيق الاختبارات الإلكترونية القائمة على الويب، أو متغيرات ذات صلة بنمط الاستجابة على أسئلة الاختبار الإلكتروني (أسامة الهنداوي،

[1] استخدم الباحث فى التوثيق، وكتابة المراجع الإصدار السادس من نظام جمعية علم النفس الامريكية APA style، وفيه يكتب اسم العائلة للمؤلف، او المؤلفين، ثم السنة، ثم الصفحة أو الصفحات، بين قوسين، ويكتب المرجع فى قائمة المراجع. هذا بالنسبة للمراجع الأجنبية. أما بالنسبة للمراجع العربية فيكتب الاسم كما ورد فى قاعدة معلومات دار النشر المتاح فيها البحث، كما هو معروف فى البيئة العربية.

(Martin, 2009؛ ٢٠١٠،). وشهدت السنوات القليلة الماضية ظهور خدمات وتطبيقات ومنصات متخصصه ترتكز على تكنولوجيا متقدمة مثل الذكاء الاصطناعي والبيانات الضخمة Big Data فى بناء الاختبارات التكيفية تابعه لجامعات كبرى مثل جامعة Cambridge فى المملكة المتحدة Concerto ، وجامعة Kaplan بالولايات المتحدة الامريكية www.kaptest.com وكذلك ظهور شركات متخصصه فى تكنولوجيات البناء والنشر والتوزيع لتلك النوعية من الاختبارات مثل FastTest، ونظم خبيرة Expert system فى الاختبارات التكيفية مثل www.siette.org ، ومنظمات دولية متخصصه فى الاختبارات التكيفية مثل المنظمة الدولية للاختبارات الإلكترونية التكيفية www.iacat.org والتي جعلت مقرها الرسمي جامعة Minnesota للعلوم والتكنولوجيا التي تحتل ترتيباً متقدما في مؤسسات التعليم الجامعي في العامين السابقين.

وترتكز الاختبارات التكيفية الإلكترونية على أساس نظري مستمد من نظرية الاستجابة للمفردة Item Response Theory (IRT) التي تمثل أحد ركائز القياس النفسي الحديث، حيث ترى تلك النظرية أن تساوى طول الاختبار من حيث عدد ونوعية المفردات لكل الممتحنين من الناحية المنهجية غير مقبول في قياس الأداء، وأن تقدير أداء المتعلم يجب أن يكون قائمًا على بعدين أساسيين هما مدى امتلاك الفرد للقدرة او السمه موضع القياس، وخصائص المفردة السيكو مترية، ومن ثم اعتمادها في القياس على الأداء الفعلي للمتعلم أثناء الاختبار دونما اعتبار لفرضية اعتدالية خصائص العينة التي ينتمي اليها المتعلم كما في نظرية القياس التقليدية. حيث يتم تفسير العلاقة بين استجابات المفحوص وخصائص المفردة في رصد القدرة الفعلية للمتعلم بشكل مستقل عن العينة التي ينتمي اليها المفحوص (Wang, 2006). كما تتميز تلك النظرية بتوفير الوقت والجهد فى قياس قدرات المفحوص بأقل عدد من المفردات (Santamaría, Zorita and Serrano, 2006, p. 927).

ويرى الباحث ان الإطار المفاهيمي الداعم للتوسع فى توظيف الاختبارات الإلكترونية التكيفية يقوم على توجيه النقد لبعدين أساسيين شائعين فى تحقيق شرط العدالة فى الاختبار الإلكتروني وهما التساوي فى طول وزمن الاختبار لكل الممتحنين، بينما الاختبارات التكيفية قائمة على فكرة تكافؤ عدد المفردات مع قدرات الممتحن، ومن ثم فإن عدد المفردات ونوعيتها ومستوى صعوبتها يختلف من ممتحن لآخر حسب مستوى امتلاك المتعلم للقدرة او السمة موضع القياس. ويرتكز تصميم الاختبار التكيفي على أربع متغيرات أساسية تتمثل في أنماط تصميم قواعد بيانات بنك الأسئلة؛ ميكانزمات اختيار المفردات؛ خوارزميات تقدير الأداء الفعلي للمتعلم؛ ثم قواعد انهاء الاختبار Termination methodsوالتي يتوقف اختيارها على نمط تصميم

طول الاختبار من حيث كونه ثابت fixed length او متغير الطولvariable length حيث يرتكز تصميم الاختبار ذو الطول الثابت على آليات التدرج في مستوى الصعوبة المفردات وفقا استجابات المتعلم ، بينما الاختبار ذو الطول الديناميكي يرتكز على تقدير لقدرة المتعلم وفقا أدائه الحالي والسابق ويترب على هذا التقدير تحديد مستوى صعوبة المفردة التالية (Cisar, et al.,2010; Oppl, etal., 2017) . وقد ظهرت الحاجة الى دراسة متغيرات تصميم الاختبارات التكيفية بأنماطها المختلفة من خلال العديد من التحديات والمشكلات التي واجهت معدي تلك النوعية من الاختبارات من حيث كم ونوعية المفردات الواجب ان يحتوي عليها بنك الأسئلة اذا كان الاختبار ثابت الطول والذى يحتوى على عدد مفردات محدد سلفا، وكذا عدد المفردات الواجب سحبها من بنك الأسئلة لأنهاء الاختبار متغير الطول وما هو التتابع المناسب في طرح الأسئلة لتحديد المستوى المبدئي للمتعلم ، وكذلك إشكالية تحديد مستوى صعوبة المفردة الواجب طرحها في بداية الاختبار. مما دعي عديد من الدراسات ذات الصلة بتطوير الاختبارات التكيفية الى الاهتمام بمتغيرات تصميم تلك النوعية من الاختبارات مثل دراسة احمد عودة (٢٠١٣) 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 للمتعلم، كما وفرت في عدد المفردات المطروحة على المتعلم لتحديد مستوى قدراته بنسبة تصل إلى ٧٠% أقل في عدد المفردات مقارنة بالاختبارات الإلكترونية التقليدية ، كما اوصت بالتوسع في استخدمها في مرحلة التعليم الجامعي لما لها من دور في تقليل قلق الاختبار لدى المحنين. وكذلك ما أظهرته دراسة كل من Stone and Davey (٢٠١١) من مميزات ذات صلة بتقليص عدد الفقرات المطروحة على المفحوص في الاختبارات التكيفية ومن ثم الوقت المطلوب لإنهاء الاختبار والذي ينعكس بالإيجاب على تقليل تأثير عوامل التعب والملل ونقص الدافعية ومستوى القلق. كما قدمت دراسة يوسف محمد السوالمة (٢٠١١) أسلوباً مقترحاً مبسط لاختيار الفقرات الاختبارية قائم على نظرية القيم الوسيطة كنموذج مرجعي في بناء خوارزميات اختيار الفقرات وقد تمثلت مبررات تقديم هذا الأسلوب فى خفض تكاليف تطوير برمجيات الاختبارات التكيفية من خلال خوارزميات تعتمد على قيم متوسطات مفردات الاختبار كأساس في تحديد القدرة الفعلية للمتعلم. وتبرز الحاجة الى اجراء العديد من البحوث ذات الصلة بمتغير نمط تصميم طول الاختبار التكيفي مجال اهتمام البحث الحالي في كونه يؤثر وبشكل مباشر في بناء قواعد انهاء الاختبار كذلك يؤثر طول الاختبار من حيث كونه ثابتا او متغيراً في طريقة بداية الاختبار، وكذا تتابع المفردات بالإضافة الى آليات سحب المفردات من بنك الأسئلة. ونظرا للأهمية هذا المتغير أوصى

عدد من الباحثين بضرورة التحقق تجريبا من دور متغير تصميم طول الاختبار من حيث كونه ثابت او متغير في مستوى قلق الممتحنين مثل دراسة كلا من Tonidandel, Quiñones and (2002) Adams. كذلك أكدت نتائج التحقق التجريب عبر مدخل مقارن ان نمط طول الاختبار من حيث كونه (ثابتاً -متغيرًا) أكثر تأثير فى دقة قياس القدرة الفعلية للمتعلم مقارنة بمتغيرات تصميم قاعدة بيانات بنك الأسئلة من حيث حجم البنك وأسلوب سحب المفردات منه (Wang, ban, 2001; Weiss,2009).

كما ترجع مبررات الاهتمام بقلق الاختبار والاتجاهات نحو الاختبارات الإلكترونية كمتغيرات تابعة في الدراسة الحالية الى تأكيد نتائج العديد من البحوث على كون قلق الاختبار يمثل أكثر العوامل المؤثر على الأداء الفعلي للمتعلم في الموقف الاختباري (Rana, Mahmood, 2010; Nicholson, 2009). كذلك ظهور عديد من التوصيات بضروه التعرف على اتجاهات الطلاب في مرحلة التعليم الجامعي نحو التوسع في استخدام الاختبارات الإلكترونية (مفلح الجدعان، ٢٠١٧). بالإضافة الى ندرة الدراسات التي اهتمت بالتحقق تجريبيًا عبر مدخل مقارن بين الاختبارات التقليدية الإلكترونية ومتغير تصميم الاختبار التكيفي (متغير – ثابت) الطول على مستوى قلق الاختبار لدى المتعلم واتجاهاته نحو الاختبارات الإلكترونية بصفة عامة.

مشكلة البحث

نبع الجانب النظري من مشكلة البحث الحالي من توصيات عديد من البحوث والدراسات التي أكدت على الحاجة الى التوسع فى استخدام الاختبارات التكيفية فى منظومة القياس والتقويم بمرحلة التعليم الجامعي وضرورة التحقق تجريبياً من منظور مقارن بين التصميم الحالي للاختبار الإلكترونية التقليدية ومتغيرات تصميم الاختبارات التكيفية على مستوى قلق الاختبار كأحد اهم العوامل المؤثر فى أداء المتعلم (Magis and Raiche,2011; Tonidandel, Quinones and Adams, 2000). وكذا ضرورة دراسة أثر طول الاختبار التكيفي الإلكتروني ومدى تأثير معرفة الطالب مسبقا بعدد لمفردات على مستوى قلق الاختبار واتجاهاته نحو تلك النوعية من الاختبارات (Weiss, 2013). كما تمثلت اوجه القصور في تصميم الاختبارات التقليدية وفقا لتيار البحوث التي ينتمى اليها البحث الحالي فى كون غالبية الاختبارات التي تم استخدمها تنتمى من حيث التصميم الى الاختبارات ذات الطول والزمن الثابت لكل الممتحنين، مما يعنى أن الأساس النظري لتلك الاختبارات قائم على نظريات القياس التقليدية من جانب وارتكازها على الخطية في تصميم الاختبار و تطوير بنوك أسئلة تلك الاختبارات من جانب آخر، كذلك يرتكز تتابع طرح المفردات على التوليد العشوائي للمفردات وبدائلها من بنك الأسئلة دونما اعتبار لتقدير استجابات المفحوص على كل مفردة

في تحديد عدد المفردات التالية وأجمالي المفردات الواجب طرحها للأنهاء الاختبار مما يعنى من الناحية التكنولوجية عدم توفر خوارزميات اختيار مفردات الاختبار وفق مستوى قدرات المتعلم الفعلية (Vale, 2006; Timminga, van der Linden, & Schweizer, 1996).

ويتمثل الجانب العملي من مشكلة البحث الحالي فى وجود العديد من جوانب القصور فى تصميم وتطوير الاختبارات الإلكترونية فى مرحلة التعليم الجامعي المستخدمة حاليا والتي اظهرتها العديد من التقارير والشواهد والأدلة (مرفق ١)، كذلك اتضح للباحث من الاطلاع على الأدلة التعريفية للبرمجيات التي تم استخدمها عدم استنادها على مرجعيات من تقارير لخبراء ومتخصصين في مجال تكنولوجيا التعليم والقياس التربوي وهيئات استشارية تابعه لمنظمات دولية متخصصه في مجال القياس والتقويم الإلكترونية كجهات تحديد جودة واعتماد لنظم القياس والتقويم في مرحلة التعليم الجامعي واعتماد تبنى برمجيات تابعه لشركة او مؤسسات ما على مؤشر مدى انتشارها في قطاع تعليمي ما فى اتخاذ إجراءات تعميمها على باقي القطاعات على الرغم من عدم صدور تقارير عن مستوى كفاءة تلك البرمجيات من جهات رسمية محلية او دولية. مما يعنى من الناحية العملية وجود قصور في توفير مستحدثات تكنولوجية فى نظم الاختبارات المستخدمة حاليا فى مرحلة التعليم الجامعي لديها القدرة على تحديد المستوى الفعلي للمتعلم بأقل عدد من المفردات، كما ان المتوفر حاليا من برمجيات ذو مستوى تحكم مرتفع (Program Control) ومن ثم انخفاض مستوى تحكم المتعلم (Learner Control)فى مسار تتابع المفردات دونما اعتبار لاستجابات المتعلم كأساس موجه لطرح المفردات الاختبارية التالية. مما دعي الباحث الى الاهتمام بمتغيرات تصميم وتطوير الاختبارات التكيفية لكونها ترتكز على أساس نظري حديث تمثل فى نظرية الاستجابة للمفردة فى عمليات التصميم واستخدمها تكنولوجيات متقدمة مثل الذكاء الاصطناعي فى تطوير متغيرات التحكم فى طول الاختبار بشكل يمكن ان يساهم فى خفض قلق الاختبارات وتكوين اتجاهات موجبة نحو الاختبارات الإلكترونية بصفة عامة. وتتمثل مشكلة البحث في العبارة التقريرية التالية : "يوجد العديد من أوجه القصور في تصميم الاختبارات الالكترونية التقليدية ومن ثم الحاجة الى اجراء مزيد من البحوث التطويرية لأنماط تصميم الاختبارات الالكترونية التكيفية ودراسة أثرها على خفض مستوى قلق الاختبار والاتجاهات نحو الاختبارات الالكترونية لدى طلاب كلية التربية".

أهداف البحث

- الكشف عن أثر الاختبارات الإلكترونية التكيفية فى خفض مستوى قلق الاختبار لدى طلاب الفرقة الثالثة بكلية التربية فى مقرر تكنولوجيا التعليم.

- الكشف عن أثر الاختبارات الإلكترونية التكيفية فى تنمية الاتجاهات لدى الطلاب نحو التوسع فى استخدام الاختبارات الإلكترونية التكيفية فى مقررات مرحلة البكالوريوس.

تساؤلات البحث

وعلى ضوء ما سبق يسعى البحث الحالي الى الإجابة عن الأسئلة التالية:

١- ما أثر نمطا تصميم الاختبار الإلكتروني التكيفي الثابت والمتغير الطول على خفض مستوى قلق الاختبار لدى طلاب كلية التربية فى مقرر تكنولوجيا التعليم؟

٢- ما أثر نمطا تصميم الاختبار الإلكتروني التكيفي الثابت والمتغير الطول على اتجاهات طلاب كلية التربية نحو الاختبارات الإلكترونية ؟

حدود البحث

اقتصر البحث على الحدود التالية:

- الحدود البشرية: عينة من طلاب الفرقة الثالثة بمرحلة البكالوريوس بكلية التربية جامعة المنوفية شعبتي (اللغة الإنجليزية – اللغة الفرنسية)
- الحدود الموضوعية:

١- مقرر تكنولوجيا التعليم

٢- أبعاد قلق الاختبار وأسبابه المختلفة (موضوعية – السياق وظروف التطبيق – الثقة بالنفس والعلاقة بالمحيطين)

٣- ابعاد الاتجاهات نحو الاختبارات الإلكترونية (التوسع فى استخدمها فى الاختبارات الرسمية – التوسع فى الدمج فى منظومة قياس وتقويم أداء الطلاب)

- الحدود المكانية: مدينة شبين الكوم كلية التربية – جامعة المنوفية.
- الحدود الزمنية: تم التطبيق الميداني فى الفصل الدراسي الأول من العام الجامعي ٢٠١٩/٢٠٢٠م.

عينة البحث

تم اختيار عينة البحث من طلاب الفرقة الثالثة بكلية التربية بشبين الكوم فى مرحلة البكالوريوس تخصص لغة انجليزية وفرنسية حيث تمثل المجتمع الأصلي للعينة والبالغ عددهم (٢٧٠) طالبا وطالبه. وقد بلغ عدد افراد العينة (٩٠) طالبا وطالبه من الدارسين لمقرر تكنولوجيا التعليم تم اختيارهم بشكل عشوائي من المجتمع الأصلي، وقد تم تقسيمهم الى ثلاث مجموعات تعرضت المجموعة الأولى الى نمط الاختبار الإلكتروني التقليدي كمجموعة ضابطة، بينما تعرضت المجموعة الثانية نمط تصميم الاختبار التكيفي ثابت الطول ، المجموعة الثالثة طبقت الاختبارات الإلكترونية متغيرة الطول .

منهج البحث

نظرًا لكون البحث الحالي ينتمي من حيث التصنيف الى البحوث التطويرية فى علم تكنولوجيا التعليم من ثم ارتكز على المناهج البحثية التالية:

١- المنهج الوصفي: فى تحديد سمات وخصائص ومظاهر قلق الاختبار وآليات التغلب عليه والحد منه من واقع البحوث السابقة، وكذا مقاييس الاتجاهات نحو بيئات التعلم الإلكتروني بصفة عامة والاختبارات الإلكترونية بصفة خاصة؛ وفيما يتصل بالمتغير المستقل تم وصف مفهوم الاختبار التكيفي وأهم الفروق بينه وبين الاختبارات الإلكترونية الأخرى ومعايير بناء الاختبارات التكيفية واهم سماتها ونماذج بناءها.

٢- المنهج التطويري فى تطوير النموذج المبدئي للاختبار التكيفي وفقا اليات وأدوات الذكاء الاصطناعي والنظم الخبيرة ذات الصلة بتلك بتطوير النوعية من الاختبارات ومراجعة جوانب القصور فى الاختبار كمنتج فى ضوء معايير تصميم وإنتاج تلك النوعية من الاختبارات

٣- المنهج التجريبي: استخدم الباحث المنهج التجريبي فى دراسة أثر التصميم المقترح للاختبار التكيفي والمقارنة بينه وبين الاختبارات الالكترونية التقليدية القائمة على أنظمة إدارة التعلم والتي يطلق عليها الاختبارات الالكترونية الخطية.

متغيرات البحث

أ- المتغيرات المستقلة

١- اختبار تكيفي ثابت الطول من حيث أجمالي عدد المفردات والصور المكافئة لها والتي يتم توليدها من بنك الأسئلة لطرحها على الممتحنين مع اختلاف نوعية ومستويات صعوبة تلك المفردات من ممتحن للأخر.

٢- اختبار تكيفي متغير الطول اختلاف عدد المفردات المطروحة على كل ممتحن حسب مستوى أدائه كأساس فى انهاء الاختبار.

ب- المتغيرات التابعة

١- المتغير التابع الأول: مستوى قلق الاختبار.

٢- المتغير التابع الثاني: اتجاهات الطلاب نحو الاختبارات الإلكترونية .

التصميم التجريبي

على ضوء متغيرات البحث تم استخدام التصميم التجريبي القائم امتداد المجموعة الضابطة- ذو الاختبار القبلي البعدي **Extended control group pre-posttest design.** ويوضح شكل (١) المتغيرات المستقلة للبحث الحالي والمتمثلة فى ثلاث متغيرات مستقلة تتمثل فى الاختبارات التقليدية الإلكترونية كمجموعة ضابطة بينما المجموعات التجريبية تمثلت فى مجموعة تجريبية أولى والتي تم تطبيق الاختبار التكيفي ثابت الإنهاء عليها، بينما المجموعة التجريبية الثانية تم تعريضها لاختبار إلكتروني تكيفي متغير الإنهاء .

المتغيرات المستقلة	مجموعات البحث		
	تجريبية اولى	تجريبية ثانية	ضابطة
اختبار إلكتروني تكيفي ثابت الإنهاء	x		
اختبار إلكتروني تكيفي متغير الإنهاء		x	
اختيار إلكتروني تقليدي			x

شكل (١) التصميم التجريبي للمعالجات حسب متغيرات البحث

فرضيات البحث

على ضوء مشكلة البحث وتساؤلاته وما تم طرحه فى الإطار النظري تم صياغة فرضيات البحث فى فرضيتين رئيسيتين الأولى صفريه والثانية موجه على النحو التالي:

١- لا يوجد فروق ذات دلالة إحصائية عند مستوى اقل من (٠.٠٥) بين متوسطات المجموعات التجريبية والمجموعة الضابطة فى التطبيق البعدي لمقياس مستوى قلق الاختبار.

ويتفرع من هذه الفرضية الصفرية الفرضيات الفرعية التالية:

١-أ- لا يوجد فرق ذو دلالة إحصائية عند مستوى اقل من (٠.٠٥) بين متوسط درجات المجموعة التجريبية الأولى والمجموعة الضابطة في التطبيق البعدي لمقياس قلق الاختبار.

١-ب- لا يوجد فرق ذو دلالة إحصائية عند مستوى اقل من (٠.٠٥) بين متوسط درجات كل من التجريبية الأولى والثانية فى التطبيق البعدي لمقياس قلق الاختبار.

١- ج – لا يوجد فرق ذو دلالة إحصائية عند مستوى اقل من (٠.٠٥) بين متوسطات درجات كل من التجريبية الثانية والضابطة فى التطبيق البعدي لمقياس قلق الاختبار.

٢- الفرض الموجه الثاني ينص على: توجد فروق ذات دلالة إحصائية بين متوسطات درجات المجموعات الضابطة والتجريبية فى التطبيق البعدي لمقياس الاتجاه نحو الاختبارات الإلكترونية لصالح المجموعات التجريبية

ويتفرع من هذه الفرضية الرئيس، الفرضيات الفرعية التالية:

٢-أ- يوجد فرق ذو دلالة إحصائية عند مستوى (٠.٠٥) بين متوسطات درجات المجموعة الضابطة والمجموعة التجريبية الأولى فى التطبيق البعدي لمقياس الاتجاهات نحو الاختبارات الإلكترونية لصالح المجموعة التجريبية الأولى .

٢-ب- يوجد فرق ذو دلالة إحصائية عند مستوى (٠.٠٥) بين متوسطات درجات المجموعة الضابطة والمجموعة التجريبية الثانية فى التطبيق البعدي لمقياس الاتجاهات نحو الاختبارات الإلكترونية لصالح المجموعة التجريبية الثانية.

٢-ج- يوجد فرق ذو دلالة إحصائية عند مستوى (٠.٠٥) بين متوسطات درجات المجموعة التجريبية الأولى والمجموعة التجريبية الثانية فى التطبيق البعدي لمقياس الاتجاهات نحو الاختبارات الإلكترونية لصالح المجموعة التجريبية الأولى .

أدوات البحث

- مقياس التقدير الذاتي لقلق الاختبار (من أعداد الباحث)
- مقياس الاتجاه نحو الاختبارات الإلكترونية (من أعداد الباحث)

أهمية البحث : تتمثل فى الاتي

- تقديم نموذج عملي يوضح اجراءات تطوير الاختبارات التكيفية فى أحد مقررات مرحلة التعليم الجامعي .
- مسايرة الاتجاهات الحديثة التي تدعم استخدام الأنظمة الخبيرة فى بناء الاختبارات التكيفية وتطوير أساليب وخوارزميات جديده فى نمذجة أداء الطالب وفقا آليات التعرف على اللغة الطبيعية للمتعلم وحالته المعرفية وحاجاته.
- الحد من ظهور متغيرات القلق والاضطراب فى الاختبارات التحصيلية من خلال نماذج تصميم للاختبارات تكيفية تراعى احتياجات المتعلم والتدرج فى مستويات صعوبة أسئلة الاختبارات الالكترونية.
- توجيه انظار الباحثين نحو تيار بحوث بناء أنظمة خبيرة فى نظم القياس والتقويم ونمذجة أداء المعلمين فى المواقف الاختبارية المختلفة .
- جذب انتباه القيادات الإدارية والأكاديمية نحو ضرورة تطوير البنية التحتية والمتطلبات اللازمة للتوسع فى تطبيقات الذكاء الاصطناعي فى التربية فى مجال القياس والتقويم.

مصطلحات البحث

- الاختبارات التكيفية:

يعرفها الباحث اجرائياً على انها نمط متقدم من أنماط الاختبارات الإلكترونية تعتمد على اليات استدلال قائمه على خوارزميات ونماذج رياضية فى تقدير المستوى المعرفي للمتعلم فى ضوء استجاباته السابقة والحالية على مفردات الاختبار كأساس فى تحديد شكل ومستوى صعوبة المفردة التالية.

١- الاختبارات التكيفية ثابتة الطول: يشير الى عدد محدد سلفا من المفردات تطرح على المفحوص للأنهاء الاختبار مع مراعاة ان تختلف تلك المفردات من متعلم لآخر

اعتمادا على انتاج صورة متكافئة متعددة من الاختبار من بنك الأسئلة به عدد كبير من المفردات.

٢- الاختبارات التكيفية متغيرة الطول: يكون عدد الأسئلة المطلوب طرحها للأنهاء الاختبار مرتبطة بمدى توفر السمة او القدرة محل القياس فى المتعلم ومن ثم عدد الأسئلة ومستويات صعوباتها يختلف من متعلم للأخر حسب نوعية استجابته على كل مفرده ويتم انهاء الاختبار وفقًا خوارزم طريقة الارجحية العظمى الذى تم تطويرها من قبل كل من Veerkamp and Berger (1997) حيث يتم إعطاء تقدير محتمل لقدرة المتعلم فى ضوء مستوى صعوبة المفردة كأساس فى سحب المفردة الأولى فى الاختبار وفى ضوء استجاباته الصحيحة او الخاطئة يتم تقدير مدى اقتران التقدير المحتمل مع القدرة الفعلية ثم حساب قيمة الخطأ المعياري لقدرة المتعلم بعد استجاباته على كل مفردة وفقا لمعامل تمييزها ومستوى صعوباتها ويتم حساب مدى الاقتران ومن ثم سحب المفردة التالية ويتم تكرار ذلك بعد كل مفردة الى ان تصل قيمة الخطأ المعياري الى (٠.٢٥) عندها يتوقف البرنامج فى سحب المفردات من البنك وفقا للمعادلة التالية :

$$\theta_{s+1} = \theta_s + \frac{\sum_{i=1}^{n} a_i [u_i - P_i(\theta_s)]}{\sum_{i=1}^{n} a_i^2 P_i(\theta_s) Q_i(\theta_s)} \dots\dots\dots\dots\dots\dots\dots (10)$$

$P_i(\theta_s)$: احتمال استجابة المفحوص على الفقرة i عند القدرة θ_s استجابة صحيحة

$Q_i(\theta_s)$: احتمال استجابة المفحوص على الفقرة i عند القدرة θ_s استجابة خاطئة.

θ_s : القدرة الأولية للمفحوص. a_i : معلم تمييز الفقرة .

ويتم تكرار هذه الحسابات غير مرّة للحصول على قيمة تقارب الصفر للحد الثاني من الطرف الأيمن من المعادلة، وبذلك تقترب قيمة θ_{s+1} من θ_s، وعندها تتوقف عملية

- قلق الاختبار:

حالة نفسية يمر بها الطالب اثناء الاختبار، تحدث نتيجة خوف الطالب من الفشل فى الاختبار، او الحصول على نتائج مرضية تحقق توقعات الاخرين فى مستوى تحصيله، بما يؤدى الى اثار سلبية فى العمليات العقلية والادراكية لدى المتعلم من تفكير وتذكر وانتباه وتركيز.

- الاتجاه نحو الاختبارات الإلكترونية :

الدرجة التي يحصل عليها الطالب فى مقياس الاتجاهات نحو الاختبارات الإلكترونية ، ويشير الى مستوى القبول، او الرفض للتوسع فى تطبيق الاختبارات الإلكترونية فى مؤسسات التعليم العالي.

الإطار النظري للبحث

الاختبارات التكيفية (نشأة وتطور المفهوم)

مصطلح الاختبارات الإلكترونية بصفة عامة كأحد أدوات التقويم يندرج تحت مسميات عدة مثل (التقييم الإلكتروني، التقييم القائم على الحاسوب، التقييم من خلال الشبكات). وقد أظهرت مراجعة عديد من اراء الباحثين الاهتمام بتعريف الاختبارات الإلكترونية بصفة عامة على انها أدوات قياس قائمه على الحاسوب وأنظمة المعلومات والاتصالات فى انتاج الأسئلة او البنود او المفردات الاختبارية بأنماطها المختلفة وكذا الصور المكافئة لها وطرحها على المتعلم بشكل عشوائي من بنوك او مستودعات للأسئلة (الغريب زاهر،٢٠٠٩؛ نبيل جاد، ٢٠١٤؛ Martin,2009).

وقد تم الإعلان عن الاختبارات التكيفية القائمة على الحاسب مجال اهتمام البحث الحالي فى اوائل التسعينات كأداة جديدة فى مجال القياس والتقويم، حيث قدمت نموذجا للاختبارات التي تراعى الفروق الفردية بين الممتحنين والتصحيح الفوري للإجابات والتغذية الراجعة متعددة المستويات والانماط، وارتكزت تلك الاختبارات على فرضية انه ليس من المعقول قياس قدرة او سمه لدى مجموعة من الافراد بنفس العدد من المفردات من حيث السهولة والصعوبة، ومن الاجدى ان نقيس أداء هؤلاء الافراد بمفردات تختلف فى مستوى صعوباتها حسب مستوى أداء المتعلم اناء الموقف الاختباري (Wainer,1990). ويؤكد ذلك ما ذكره Weiss (1990) أن الاختبارات التكيفية هي تلك الاختبارات التي تطبق فيها المفردات الاختبارية بشكل مختلف حسب الاختلاف بين الافراد في القدرة او السمة المقاسة. ويمثل رأى Eignor, et al (1993) اطارا مرجعيا لتعريفاً قياسياً للاختبارات التكيفية المحوسبة او الإلكترونية حيث وصفها بأنها اختبارات تنتقى فيها مفردات الاختبار الوأحد ة تلو الأخرى من مخزون كبير من الأسئلة، بحيث ان كل مفرده تم اختيارها مناسبه لمستوى قدرة الفرد الممتحن. وقد ذكر Boyd(2003) ان تلك النوعية من الاختبارات سميت بالاختبارات التكيفية المحوسبة لأنها تقدم للمفحوص فقرات تتناسب مع

قدراتـه وبنـاء علـى نمـط إجابـات للمفحـوص عـن الفقرات المقدمـة لـه يـتم تقدير قدراتـه. وقد أوضح Wainer, et al.(2000) ان السمة المميزة فى الاختبارات التكيفيـة فى كونهـا تستند على مستوى أداء المـتعلم فـى توليـد الأسـئلة مـن بنـوك أسـئلة تحتـوي علـى المئـات مـن المفـردات بمسـتويات سـهولة او صـعوبة حسـب أداء المـتعلم، ومـن ثـم ففرصـة اخفـاق الطالـب مرتفـع الأداء فـى الأسـئلة السهلة ضعيفة وكذا فرصـة الطالـب منخفض الأداء فــى الإجابــة الصــحيحة علــى الأســئلة الصــعبة منخفضه.

وقد عرف كل من Conejo et al. (2004) الاختبارات التكيفيـة على انهـا تمثل نظـام اختبـارات مدار بالحاسب الآلي دون تدخل للعنصر البشرى فى تنظيم تتابع المفردات الاختبارية واتخاذ قرار انهـاء الاختبـار وفقـا ملـف تعريـف المـتعلم Student Profile الذي يتم انشائه وتحديثه آليا اثنـاء تفاعل المتعلم مع مفردات الاختبار. وقد عرفها احمد عودة (٢٠١٣) بأنهــا نــوع مــن الاختبــارات يفُصــل Tailored لكل مفحوص على حدة، وذلك بعرض الفقرات التي تناسب مع مستوى قدراتـه، بمـا يمكن من تقدير ادق للقدرة بأقل عدد من الفقرات. وترتكز فكـرة التكيفيــة فــى الاختبــارات الإلكترونيــة علــى التعديل الـديناميكي فى مسـتويات صـعوبة مفـردات الاختبار وفقا اداء الطالـب السـابق والحـالي فيحصل البرنـامج علـى المعلومـات الكافيـة التـي تمكنـه مـن الحكم على مستوى المتقدم للاختبـار وتقدير درجـة معينـة لـه بنـاء علـى مستوى صـعوبة الأسـئلة التـي أجاب عنها، وليس بناء على عدد الأسئلة الصحيحة التي أجاب عنها.

ويـرى الباحـث علـى ضـوء التعريفـات السـابقة ان فكرة التوليد والترتيب العشوائي لمفردات الاختبـار الإلكترونـي التقليـدي وبدائلـه وفقـا سـياق ثابـت مـن حيث عدد المفردات وزمن الاختبار على الـرغم مـن توفيرها قدر مـن الموضـوعية ، الا انهـا فعليـا تفتقد الـى القدرة علـى الكشـف عـن مسـتوى أداء المـتعلم الفعلـي ممـا يعنـى عـدم مراعـاة مناسـبه مسـتوى صعوبة المفردات المطروحـة لقدرات الفرد؛ والـذى يعكس انخفاض مستوى جودة الاختبار كـأداة قيـاس فـى التحديـد الـدقيق لـلأداء المـتعلم بأقـل عـدد مـن المفـردات ممـا دعـي العديـد مـن المتخصصـين فـى القياس النفسي الـى ضـرورة تـوفر خاصـية التكيفيـة كأحـد اهـم عناصـر تصـميم الاختبـارات الإلكترونيـة وتبـرز الإشـكالية الكبـرى فـى تلـك النوعيـة مـن الاختبـارات فـى انهـاء الاختبـار وفقـا تقـدير دقيـق لمسـتوى أداء المـتعلم ممـا يعنـى عمليـا ان طـول الاختبـار مـن حيـث المفـردات الواجـب طرحهـا علـى المـتعلم مرهـون باسـتجابته مـن حيـث الصـواب والخطأ من جانب ومستوى صعوبة المفردات التـي تم طرحها من جانب اخر.

الفــروق بــين الاختيــارات الإلكترونيــة التقليدية والتكيفية

أظهـرت اراء عديـد مـن البـاحثين مجموعـة مـن الفــروق الجوهريــة بــين الاختبــارات التكيفيــة

والاختبارات الإلكترونية التقليدية مثل (Meijer) and nering, 1999؛(Conejo et al, 2004) . ويوضح جدول (٢) ملخص نتائج رصد تلك الآراء حول اهم الفروق بين الاختيارات الإلكترونية التقليدية والتكيفية.

جدول (1) الفروق بين الاختيارات الإلكترونية التقليدية والتكيفية

وجه المقارنة	الاختبارات الالكترونية التقليدية	الاختبارات التكيفية
تتابع طرح المفردات على الممتحن	خطية	شبكية ديناميكية
سحب المفردات من بنك الأسئلة	عشوائي، ترتيب تصاعدي او تنازلي ثابت.	حسب تقديرات استجابة المتعلم على كل مفردة
محك الحكم على قدرات المتعلم	الدرجة الكلية	خصائص المفردة والسمه المقاسة
طول الاختبار	ثابت	ثابت، متغير
نظرية القياس	القياس الكلاسيكي	نظرية الاستجابة للمفردة ونماذجها
نظم إدارة بنك الاختبار	إدارة قواعد البيانات الهيكلية	قواعد المعرفة وآليات استدلال قائمة على تكنولوجيا الذكاء الاصطناعي
الزمن الكلى للاختبار	ثابت	ثابت – متغير
مستوى الغش cheating level	مرتفع	منخفض

أنواع الاختبارات الإلكترونية التكيفية

أشار كل من Razzaq et al.(2005) الى ان أساليب التقييم بنظم التدريس الذكية التي تمثل الإطار المرجعي للاختبارات الإلكترونية التكيفية تنقسم الى نوعيين أساسيين النوع الأول قائم على المهام (Assessment Based Tasks) والثاني قائم على المفردات الاختبارية (Assessment Based Items) ومن ثم يمكن تصنيف الاختبارات التكيفية الى نوعيين رئيسيين موجه بالمهام وأخرى موجه بالمفردات. كما يمكن تصنيف الاختبارات التكيفية حسب اسلوب بداية الاختبار الى نمطين أساسيين، النمط الأول يعتمد على خوارزميات او تقنيات تحديد اول مفردة فى الاختبار قائم على قياس قبلي لمستوى قدرات المتعلم من خلال تكنيك

طرح مفردات متنوعة فى مستوى صعوبتها كاختبار قبلي وتمثل متوسطات استجابات الطالب الأداء الفعلي المعبر عن قدرته الحقيقة ومن ثم تحديد مستوى صعوبة اول مفردة فى بداية الاختبار، بينما النمط الثاني من الاختبارات التكيفية قائم على قاعدة القفز Step size فى مطابقة مستوى صعوبة المفردة مع مستوى القدرة الفعلي للمتعلم حيث يتم اختيار اول مفردة عند مستوى صعوبة متوسط بحيث اذا أجاب المفحوص او المتعلم اجابه صحيحه يتم طرح المفردة التالية عند مستوى الصعوبة الأعلى وفقا مصفوفة مستويات الصعوبة للاختبار، بينما اذا كانت اجابته خاطئة يتم اختيار مستوى الصعوبة الأدنى للمفردة التالية (Cisar, et al, 2010,p.135). كما أشار عديد من الباحثين الى تصنيف الاختبارات التكيفية حسب نمط انهاء الاختبار الى اختبارات تكيفية ثابتة الطول (Fixed Length) ومتغيرة الطول فى الإنهاء (variable length) من حيث عدد المفردات المطلوبة طرحها على المتعلم للأنهاء الاختبار. حيث يشير النمط التكيفي ثابت الإنهاء الى عدد محدد سلفا من المفردات تطرح على المفحوص للأنهاء الاختبار مع مراعاة ان تختلف تلك المفردات من متعلم للآخر اعتمادا على انتاج صورة متكافئة متعددة من الاختبار من بنك اسئلة به عدد كبير من المفردات، بينما فى نوعية الاختبارات التكيفية متغيرة النهاية يكون عدد الأسئلة المطلوب طرحها للأنهاء الاختبار مرتبطة بمدى توفر السمة او القدرة محل القياس فى المتعلم ومن ثم عدد الأسئلة ومستويات صعوباتها يختلف من متعلم للأخر حسب نوعية استجابته على كل مفرده (Stone and Davey, 2011).

مكونات الاختبارات الإلكترونية التكيفية.

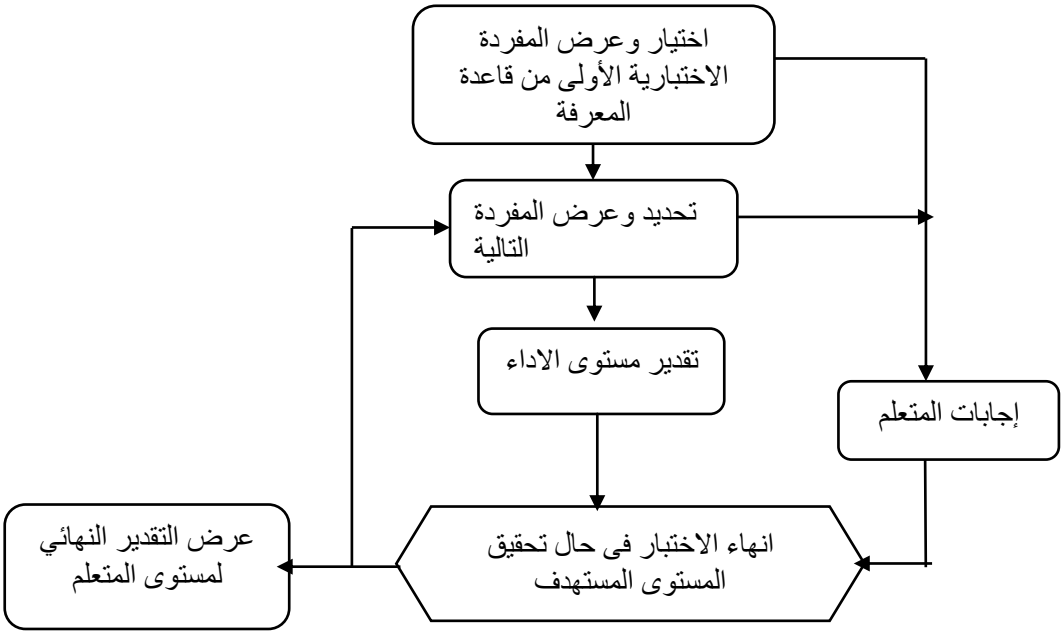
يرى بعض الباحثين ان الاختبار الإلكتروني بصفة عامة يتكون من ثلاث أجزاء رئيسية تعكس إجراءات تعامل المتعلم مع النظام الاختباري، والتي تتمثل فى دخول الاختبار، طرح مفردات الاختبار على المتعلم، عرض نتيجة الاختبار (ايناس مندور، ٢٠١٣). ويوجد تباين فى اراء الباحثين والمتخصصين حول العناصر الواجب توافرها فى تصميم الاختبار التكيفي بصفة خاصة مجال اهتمام البحث الحالي ولعل ذلك يرجع الى اختلاف الغرض من تصميم الاختبار التكيفي وكذا طبيعة الفئة الموجه لها هذا بالإضافة الى طبيعة المحتوى التعليمي والتكنولوجيا المستخدمة فى تطوير هذا الاختبار. ويرى عديد من الباحثين ان بيئة التعلم التكيفية بصفة عامة تتكون من أربعة نماذج أساسية تتمثل فى كل من: ١) نموذج المجال Domain Model؛ ٢) نموذج المتعلم Learner model، ٣) نموذج التكيف adaptation model؛ ٤) نموذج المجموعة group model (رشا حمدي هداية، ٢٠١٩؛ محمد عطية خميس، ٢٠١٨؛ نبيل جاد ومروة المحمدي، ٢٠١٧؛ Tadlaoui & Kommers, 2015؛ Bouamrane,2013).

ويمثل نموذج الطالب العنصر الأساسي فى تصميم الاختبار التكيفي لكونه الأساس فى تحديد مستوى الأداء الفعلي للمتعلم وكذا الأسلوب المناسب لعرض محتوى التعلم كما ونوعا، وفعليا يتم ذلك بطريقتين الأولى تعتمد على تعريف المتعلم لنفسه في بداية التعامل مع النظام التعليم من خلال سلسلة من الأسئلة والاستبيانات يستطيع من خلالها النظام بناء ملف تعريفي عن الطالب قبل بداية أحد اث التعلم، بينما تعتمد الطريقة الثانية على تتبع استجابات المتعلم وبناء تقديرات عن أدائه من خلال قواعد استدلال ومراجعة مستمرة للاداه الحالي والسابق من خلال تقنيات الذكاء الاصطناعي والتنقيب عن البيانات (Baylari& Montazer, 2009; iloglugil&Ineeoglu,2012).

خرائط مسار(خوارزميات) بناء الاختبار التكيفى

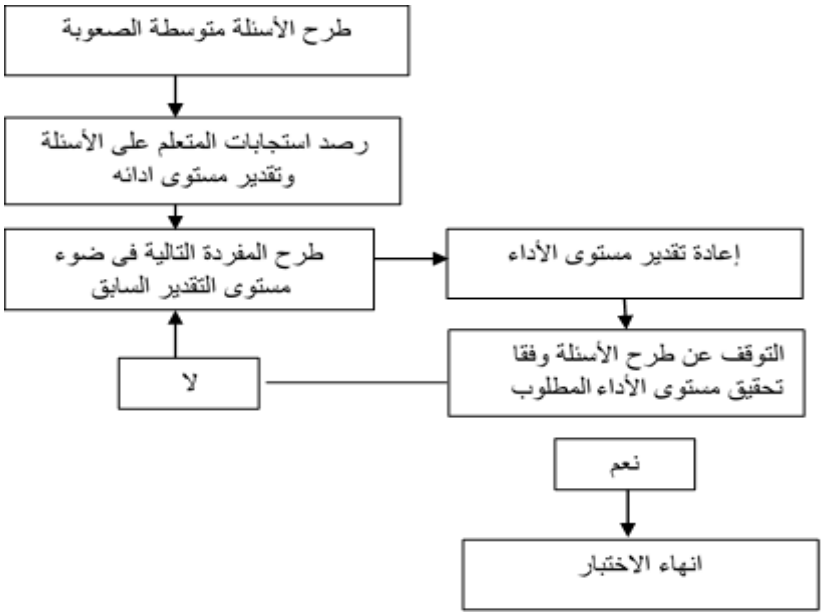
قدم كل من Olea & Ponsoda (١٩٩٦) مخطط كما هو موضح بالشكل (٢) لخريطة مسار بناء نموذج المتعلم فى الاختبارات التكيفية حيث تم اعداد خوارزميات التقدير المبدئي لمستوى أداء المتعلم وفقا الاجراءات التالية :

١. بناء قاعدة بيانات أسئلة الاختبار تتوفر بها خاصية الفحص المستمر لجميع المفردات لتحديد أفضل مفردة تالية فى متتابعة عرض الاسئلة فى ضوء استجابة المتعلم على المفردة الحالية.
٢. رصد وتحديد مستوى صعوبة المفردة الاختبارية الحالية المطروحة على المتعلم وكذا استجابته عليها.
٣. وفقا استجابات المتعلم يتم تحديث تقديرات أداء المتعلم بشكل دوري.
٤. يتم تكرار الخطوة الأولى والثالثة حتى يحقق المتعلم مستوى الأداء المستهدف.



شكل (٢) خريطة مسار بناء الاختبارات التكيفية (Olea & Ponsoda,1996).

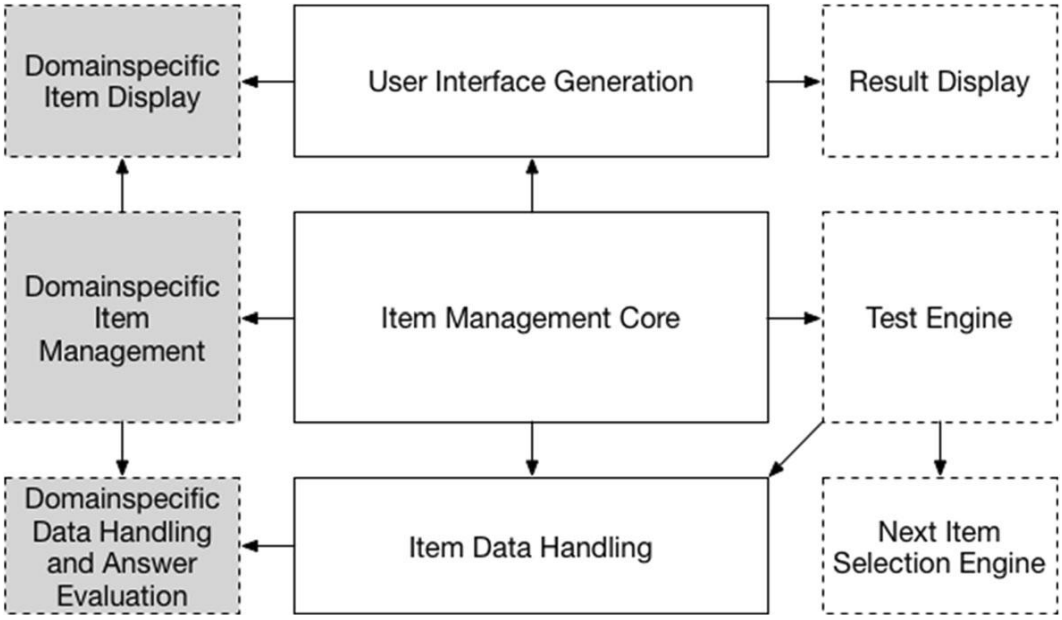
وعلى نفس النهج قدمت دراسة (2010) Wauters et al. مخطط بآليات عمل نموذج الطالب خاصة فيما يتصل بقواعد انهاء الاختبار التكيفي كما هو موضح بالشكل (٣)



شكل (٣) أليات عمل نموذج الطالب فى الاختبار الإلكتروني التكيفي

وترتكز قواعد انهاء الاختبار على الحصول على معلومات من نموذج الطالب ذات الصلة بتتبع الثبات في أداء المتعلم عند مستويات صعوبة محدده وعدم وجود بيانات تفيد بتغيير الاداء من واقع ملف تعريفه واداءه السابق كأساس في انهاء الاختبار للمتعلم (Leung & Li 2007;Baylari & Montazer 2009; Chen & Chen 2009; Veldkamp & Matteucci, 2013).

وقد صدر عن كلاً من Oppl, et al.(2017) كخبراء متخصصين في مجال تطوير منصات الاختبارات التكيفية نموذج لمخطط قياسي لنظام الاختبار التكيفي المحوسب كما هو موضح بالشكل (٤) يشتمل النظام الاختباري التكيفي على (٩) مكونات يتم الدمج بينها او استبعاد مكون حسب سيناريو بناء الاختبار التكيفي والمجال المعرفي لمحتوى التعلم، وتشير المربعات الموجودة في وسط المخطط الى المكونات التي لا يمكن الاستغناء عنها كمكون اصيل من مكونات الاختبار التكيفي والتي تتمثل في كل من مكون واجهة التفاعل؛ مركز إدارة المفردات الاختبارية؛ تنظيم بيانات المفردات من حيث خصائصها السيكو مترية.، وتشير المربعات ذات اللون الرمادي بمحيط خطى نقطي الى مجال او مكون (عرض المفردات – إدارة المفردات – كيفية التعامل مع بيانات المفردات السيكو مترية وتقييم استجابات الطلاب) . بينما تشير المربعات البيضاء ذات المحيط النقطي الى ثلاث مكونات تتمثل في (عرض النتائج – محرك تشغيل الاختبار – محرك اختيار المفردات التالية للمفردة المعروضة حاليا على المتعلم



شكل (٤) مكونات النظام الاختبار التكيفي (Oppl, et al., 2017)

ابعاد تصميم الاختبارات التكيفية

أشار عديد من الباحثين الى ثلاث ابعاد رئيسية تمثل الإطار المرجعي لضبط وتعديل خصائص بيئة التعلم وفقا للاحتياجات المتعلم وسياق التعلم وكذا تحديد المواصفات الفنية الواجب توفرها لتحقيق شروط التكيفية فى تصميم عناصر بيئة التعلم وقد تمثلت تلك الابعاد فى كل من الشكل (Form)، المصدر (Source)، الوسيط (Medium)

(Brusilovsky& Peylo 2003; Paramythis & Loidl-Reisinger2004)

وعلى ضوء ما سبق قدم كل من (2010) Wauters et al. وصفا تحليلاً للشروط والمواصفات الواجب توافرها في بيئة التعلم التكيفية وفقا للأبعاد السابقة تمثل التفاصيل الفنية التالية: -

البعد الأول والمتمثل فى شكل التكيفية formيمكن تقسيمهم الى ثلاث فئات:

- تمثيل العرض التكيفي Adaptive Form Representation

ويشير هذا البعد الى مراعاة مواصفات أسلوب عرض عناصر المحتوى التعليمي من صور ورسوم وفيديو ونصوص وروابط.

- تمثيل المحتوى التكيفي Adaptive content representation

ويشير هذا البعد الى ضرورة امداد المتعلم بالدعم المناسب فى كل خطوة من خطوات التعلم اثناء تعامله من مشكلات التعلم بحيث يكون قائما على اكتشاف الفجوات فى أداء المتعلم اثناء عملية التعلم.

- تتابع عناصر المحتوى التكيفي Adaptive curriculum sequencing

يرتكز تتابع عرض المحتوى التكيفي على طرح أسئلة متنوعة من حيث مستويات الصعوبة لتحديد مستوى كفاءة اكتساب المتعلم لمفاهيم مادة لتعلم وكذا أسلوب العرض المناسب لعرض عناصر المحتوى التالية فى ضوء تقدير مستوى أدائه الفعلي.

البعد الثاني: مصدر التكيفية يشير هذا البعد الى ثلاث فئات من المصادر المؤثر فى مستوى تكيفية بيئة التعلم

المصدر الأول ذو صلة بخصائص محتوى التعلم من حيث عناصر وموضوعات المحتوى ومستويات الصعوبة فى كل عنصر

المصدر الثاني خصائص المتعلم ويحتوي على أساليب تقدير المستوى المعرفي للمتعلم ومؤشرات دافعيته للتعلم ومستوى العبء المعرفي.

المصدر الثالث خصائص السياق وتحتوي على زمن ومكان التعلم والأجهزة

الإلكترونية المستخدمة فى الوصول الى مادة التعلم.

البعد الثالث الوسيط ترتبط مواصفات هذا البعد بنوع بيئة التعلم التكيفية التي تنقسم الى نوعين أساسيين هما بيئات التعلم التكيفية فائقة الوسائط ونظم التدريس الذكية حيث يعتمد النوع الأول على الدمج بين كمية ضخمة من عناصر محتوى التعلم عبر النصوص الفائقة كأساس فى توفير مستوى تكيفية مرتفع، بينما تعتمد نظم التدريس الذكية على دعم المتعلم اثناء عملية التعلم يتفق مع مستوى أدائه الفعلي (Brusilovsky, 1999).

معايير تصميم الاختبارات التكيفية

توصل الباحث من خلال استقراء اراء عديد من الباحثين الى مجموعة من المعايير التي تشير الى الحد الأدنى من المواصفات الواجب توافرها فى الاختبار التكيفي تمثلت فى الاتي:

- معيار الشخصنة The personalization

تمثل الشخصنة أحد اهم معايير التصميم الجيد لبيئة التعلم التكيفية من خلال تعديل وتغيير نمط عرض المحتوى وفقا خصائص المتعلم وتفضيلاته (تامر الملاح،٢٠١٧؛ اميرة عطا، ٢٠١٤). يشير هذا المعيار الى ضرورة ان يكون الاختبار التكيفي مُفصل وفق القدرات الفعلية للمتعلم من خلال نمذجة أداءه بشكل يحاكى احتمالات إجابة المفحوص على كل مفردة من مفردات الاختبار ووفقًا لمستوى قدراته من خلال تتبع استجابته الحالية من جانب وجمع معلومات عن استجاباته السابقة من جانب اخر. وتتمثل مؤشرات هذا المعيار فى وجود منظمات متقدمة لفرز وترتيب استجابات المعلمين الصحيحة والخطأ بشكل ديناميكي بحيث يمكن تقدير الأداء الفعلي للمتعلم.

- معيار التنوع فى مفردات بنك الأسئلة

يشير هذا المعيار الى أهمية التنوع وبعدد كبير فى أنماط المفردات الاختبارية من حيث مستويات الصعوبة والتمييز (Flaugher, 1990)

- معيار موائمة تتابع المفردات.

أشار عديد من الباحثين الى مداخل مختلفة واليات فى تحديد اول مفردة فى الاختبار التكيفي وما يتبعها من باقي المفردات من حيث السهولة او الصعوبة حسب الأداء الفعلي للمتعلم. يمثل هذا العنصر الأساس فى مناسبة الاختبار لمستوى قدرات المتعلم كما يترتب على مناسبه المفردة الاختبارية الأولى لمستوى أداء المتعلم العديد من القرارات مثل إجمالي عدد مفردات الاختبار ووقت انهاء الاختبار (طول الاختبار) بالإضافة الى تتابع المفردات التالية. وقد استخدم كلا من Cisar, et al.(2010) مدخل فقرات تحديد المستوى المبدئي للمتعلم حيث يعرض عليه بعض المفردات، وبناء على اداءه يتم اختيار المفردات التالية من بنك الأسئلة حيث يعتمد فى اختيار تلك المفردات على متغيرين

أساسيين هما استجابات المتعلم وخصائص المفردة المعروضة على المتعلم. يرتكز اختيار المفردة التالية على تقدير استجابة المفحوص على المفردة الحالية ومن ثم اليات تقدير أداء المتعلم تمثل الأساس فى جودة ودقة الاختبار التكيفي فى قياس مستوى القدرة التحصيلية لدى المفحوص.

- معيار الإنهاء المتمايز

يتم انهاء الاختبار التكيفي عند تمييز أداء المفحوص بشكل مستقل عن باقي افراد المجموعة التي ينتمي اليها، ومن الناحية العملية يتم انهاء الاختبار بعد زمن وعدد من المفردات يتوقف على الأداء الفعلي للمفحوص، وإجمالي عدد المفردات المطلوب سحبها من البنوك او الحد المطلوب من الأسئلة للانتهاء الاختبار يتوقف على العديد من المتغيرات تتمثل فى عدد الأسئلة المعروضة على المتعلم من حيث مستوي الصعوبة والتمييز وكذا تمثيل تلك الأسئلة للأهداف المراد تحقيقها.

مميزات الاختبارات التكيفية

قدم كل من Mislevy & Almond (1997) قائمة بأهم مميزات الاختبارات التكيفية تمثلت في الاتي:

- أسهمت الاختبارات التكيفية في انخفاض واضح في زمن الاختبار اللازمة لتقدير أداء المفحوصين.
- الدقة في تقدير المستوى المعرفي للمتعلم بأقل عدد من المفردات الاختبارية
- تحسين مستوى دافعية المفحوصين
- الاستفادة من بنوك الأسئلة التابعة لتلك الاختبارات في بناء أنماط أخرى من الاختبارات مثل اختبارات التسكين وتشخيص مستوى الصعوبة وكذا عدد أكبر من الصور المتكافئة من الاختبار المعد لقياس سمة او قدرة محددة
- تتسم أسئلة الاختبارات التكيفية باحتوائها وسائط متعددة (من أصوات وصور ورسوم) ترفع من مستوى فهم محتوى المفردة وإزالة اللبس الناجم عن ارتفاع مستوى التجريد في المفردات ذات الطابع النصي.
- تطوير خوارزميات اختيار المفردات وفقا مستوى أداء المتعلم ترفع من مستوى كفاءة الاختبار بشكل تراكمي بشكل يجعل منه نظام خبير في التقويم مع استمرارية تطوير تلك الخوارزميات.

ويؤكد ما سبق Cisar, et al.(2010) حيث ذكر ان تلك النوعية من الاختبارات أكثر مرونة من الاختبارات التقليدية المحوسبة من حيث عدد الصور المكافئة للاختبار. كذلك اظهرت دراسة كل من Stone and Davey (2011) مميزات تتمثل فى تقليل تأثير عوامل التعب والملل ونقص الدافعية. كما تتميز تلك النوعية من الاختبارات الإلكترونية بتوفيرها معلومات إحصائية عن أداء

المفحوص على كل مفردة من مفردات الاختبار وعدد اقل من المفردات للحصول على المستوى المطلوب من الدقة والثبات مقارنة بالاختبارات الإلكترونية التقليدية (Huo, 2009; Magis and Raiche,2011).

الأساس النظري لبناء الاختبارات التكيفية (نظرية الاستجابة للمفردة)

تمثل نظرية الاستجابة للمفردة Item Response Theory والنماذج التابعة لها الأساس الموجه لبناء الاختبارات التكيفية لكونها تمثل الاتجاه المعاصر فى القياس والتقويم. وقد أطلق المتخصصين فى القياس النفسي على هذه النظرية العديد من المسميات مثل نظرية السمة الكامنة Latent Trait Theory، نظرية الاختبارات المعاصرة Modern Test Theory، نظرية منحنى خاصية المفردة Item Characteristic Curve Theory

(Crocker & Algina, 1986,339 ;Janda, 1998, 134 ;Hambleton, 1982,2). ونظرية الاستجابة للمفردة من أولى النظريات التي قدمت بشكل رسمي الاختبارات التكيفية فى مجال القياس والتقويم (Fernandez, 2003). وترتكز نظرية الاستجابة للمفردة (IRT) على فرضية ان امتلاك الفرد لقدر معين من المتغير المراد قياسه يؤثر على كم ونوعية استجابته على مفردات الاختبار او المقياس المصمم لقياس هذا المتغير او السمة ومن ثم فأن احتمال إجابة الفرد على كل مفردة من مفردات الاختبار يمثل دالة في متغيرين، المتغير الأول السمة او المتغير او القدرة المراد قياسه والتي يستدل عليها من منحنى خصائص المفردة Item Characteristic Curve (ICC)، والمتغير الثاني خصائص المفردة والتي تتمثل في كل من معاملي الصعوبة والتمييز

(صلاح علام،٢٠٠١؛ George,2007؛ محمد عبد الوهاب، ٢٠١٠). ومن اهم سمات تلك النظرية تفسير العلاقة بين استجابات المفحوص وخصائص المفردة من خلال منحنى بياني يرصد أدائه الفعلي بشكل مستقل عن العينة التي ينتمي اليها المفحوص (Wang,2006)

<u>مميزات نظرية الاستجابة للمفردة فى بناء الاختبارات التكيفية.</u>

اسفرت مراجعات الادبيات ذات الصلة بمميزات الاختبارات الإلكترونية التكيفية عما يلي :

- نظرية الاستجابة للمفردة انتقلت فى تقدير أداء الفرد فى اختبار ما من مستوى التقييم القائم على الدرجة الكلية الى مستوى تقدير أداء الفرد على كل مفردة فى ضوء خصائصها السيكو مترية من حيث بارامترات مستوى الصعوبة والتمييز. مما يعنى ان تقدير الإجابات الخطأ والصواب على مفردات الاختبار يخضع بشكل أساسي الى خصائص تلك المفردات ووزنها النوعي فى الدرجة الكلية للاختبار (Baker,2001, p.6;Gomez,2008, p.866).

- حققت نظرية الاستجابة للمفردة فعالية أفضل من نظرية القياس التقليدي فى تصميم وبناء بنوك الأسئلة لكون الخصائص السيكو مترية لكل مفردة قائمه على تحليلات دقيقة لا تتأثر بخصائص العينة التي تم تطبيق المفردات عليها (عماد غصاب عبابنة، ٢٠٠٨). مما يعنى استقلالية تقدير أداء الفرد عن الجماعة المرجعية او العينة التي ينتمي اليها، كما يشير أيضا الى أهمية وجود عدد كبير من المفردات فى بنوك الأسئلة للحد من تأثير العينة على خصائص المفردات.
- يعتمد تطوير الاختبارات القائمة على نظرية الاستجابة للمفردة على نماذج رياضية تحاكى احتمالات إجابة الفرد على كل مفردة فى كل مستوى من مستويات المتغير المراد قياسه من خلال الاختبار مما يعنى ان تقدير استجابات الافراد يتم وفقا تحليل منحنى أدائهم الفعلي على كل مفردة .مما يسمح بإمكانية المقارنة بين الافراد المطبق عليهم اختبارات مختلفة لنفس المتغير الذى يقيسه الاختبار وكذا تطبيق النتائج الخاصة بتحليل المفردات على مجموعات أخرى غير المجموعة او العينة التي طبق عليها الاختبار وذلك لتوفر استقلالية السمة عن العينة (Crocker&Algina,1986, p.339).
- يرجع التوسع فى استخدام مفاهيم نظرية الاستجابة للمفردة فى تصميم الاختبارات الإلكترونية التكيفية الى كونها ترتكز على قاعدة تلقى المفحوصين عدد من المفردات الاختبارية تتناسب مع قدراتهم وامكاناتهم (Hambleton, 1982,p.24;McGlohen,2004,p.36).
- توفير الوقت والجهد فى قياس قدرات المفحوص بأقل عدد من المفردات (Sanz-Santamaría, Zorita&Serrano,2006, p927).
- ساهمت نظرية الاستجابة للمفردة فى تحديد مستوى تحيز المفردة (Item Bios) من خلال تتبع استجابات افراد من مجموعات مختلفة لنفس المفردة (رضا سمير عوض ،٢٠٠٨، ٥؛ Hungi, 2005, p.143-144)
- بناء الصور المتكافئة من الاختبار وفقا فرضيات نظرية الاستجابة للمفردة أفضل من نظرية القياس التقليدية فقد أصبحت نظرية الاستجابة للمفردة الاكثر تطبيقا على نطاق واسع من قبل مصممي ومطوري الاختبارات وكذا تحليل الاختبارات محكية ومعيارية المرجع (Fan,1998,p363;Önder, 2007,p210)

العلاقة بين متغيرات تصميم الاختبارات الإلكترونية ومستوى قلق الاختبار والاتجاهات نحو الاختبار الإلكترونية

يمثل قلق الاختبار أحد العوامل المؤثرة على أداء الطلاب فى الاختبارات التحصيلية بصفة عامة ومن ثم توجد علاقة بين ارتفاع او انخفاض مستوى قلق الاختبار ومعدلات تحصيل الطلاب فى اختبار ما. وقد

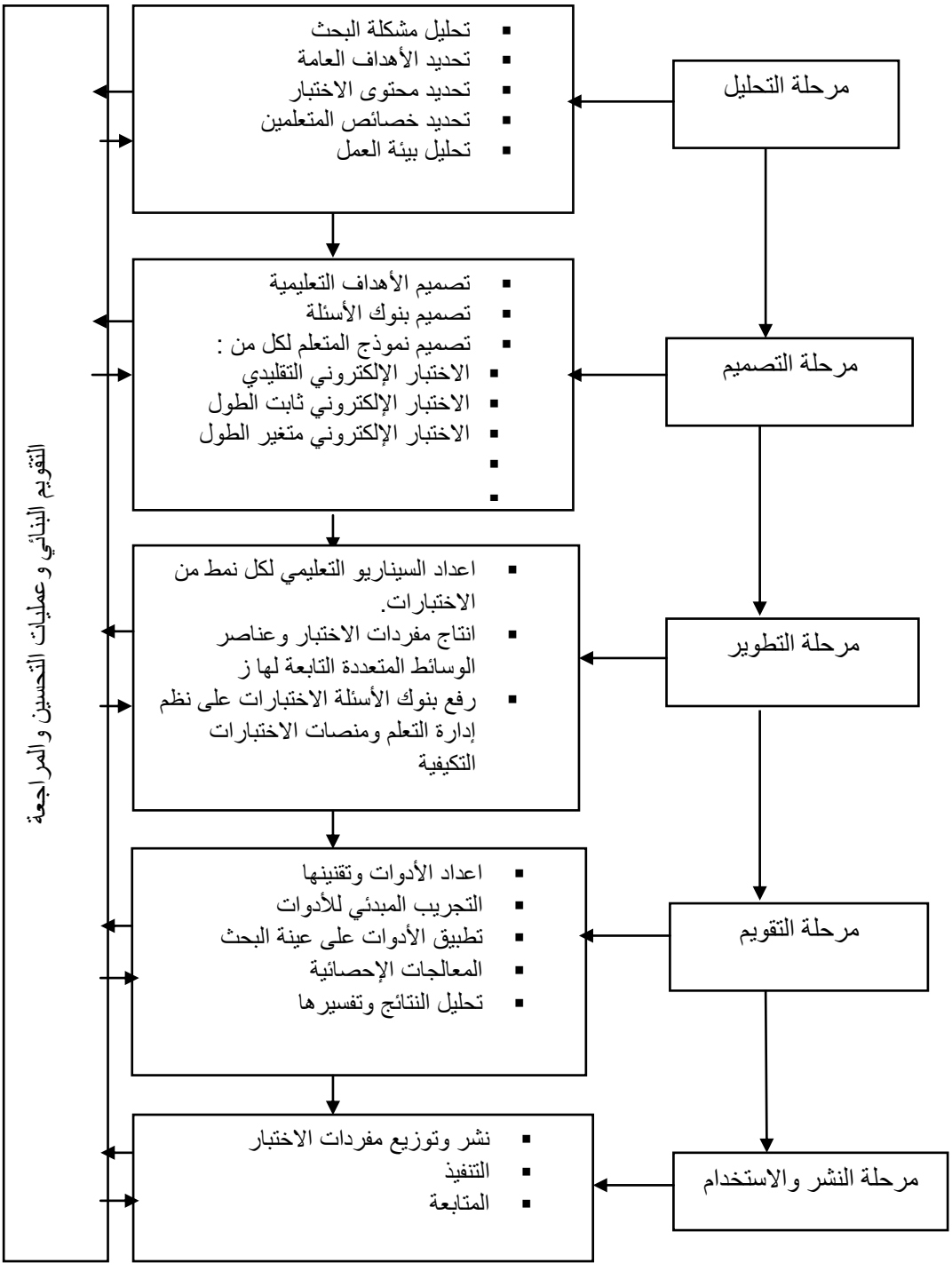
تناولت العديد من البحوث فى مجال علم النفس قلق الاختبار كأحد المتغيرات النفسية المؤثرة على مستوى تحصيل الطلاب فى المراحل الدراسية المختلفة. ويمثل كتاب كلا من Sarason, sarason(1990) الإطار المرجعي في وصف متغيرات قلق الاختبار ومدى تأثيرها على تحصيل الطلاب، على سبيل المثال متغيرات أهمية الاختبار للمتعلم وكذا مستوى صعوبة الأسئلة وقد أظهرت نتائج العديد من البحوث التجريبية للباحثين السابق ذكرهما ان قلق الاختبار يؤثر بالسلب على مستوى تحصيل الطلاب في الاختبارات التي تعتمد على المهارات العقلية بينما في اختبارات الأداء توفير قدر من التوتر في بداية الاختبار يؤثر بالإيجاب على رفع مستوى الأداء. وقد اكدت اراء عديد من الباحثين على وجود تأثير بشكل واضح لترتيب الأسئلة بشكل تصاعدي او تنازلي او بشكل عشوائي من حيث مستوى الصعوبة على مستوى قلق الاختبار التي تعتمد على وسيط مثل الحاسب الآلي في نشر وتوزيع المفردات الاختيارية على الممتحنين (Dordi Nejd, et al., 2011). وقد ذكر Sax, Gotmak(2013) ان تأثير عنصر تصميم تتابع أسئلة الاختبار من حيث كونه تصاعدي او تنازلي او عشوائي على مستوى قلق الممتحن مرهون بزمن الاختبار بمعنى كلما كان زمن الاختبار أطول قل تأثير عامل قلق الاختبار بغض النظر عن نوعية تتابع الأسئلة حيث تلاشت الفروق بين متوسطات درجات طلاب التي تعرضت لمفردات تصاعدية الترتيب في مقابل المجموعات التي تعرضت لمفردات تم توليدها عشوائيا او تنازليا عندما تم زيادة زمن انهاء الاختبار لجميع المجموعات التي تعرضت لنفس العدد من المفردات مع اختلاف نمط تصميم ترتيب المفردات. وقد أشار Weiss(2013) الى ان البحوث في هذا المجال يجب ان تهتم بمتغيرات ذات صلة بميول واتجاهات الطلاب نحو الاختبارات الإلكترونية خاصة مع تعدد الأجهزة ومنصات تقديم الاختبارات الإلكترونية بصفة عامة والتكيفية بصفة خاصة، ومن اكثر المتغيرات الواجب دراستها ذات التأثير النفسي الدور الذى يلعبه حجم بنك الأسئلة في الاختبارات التكيفية ومنهجيته في قياس أداء المتعلم على مستوى قلق الممتحنين، وكذا وضعية تطبيق الاختبار من حيث كونه في وضعية (off/on line) وكذا طول الاختبار ومدى تأثير معرفة الطالب مسبقا بعدد لمفردات على مستوى قلق الاختبار واتجاهاته نحو تلك النوعية من الاختبارات . وقد أظهرت عديد من الدراسات وجود تأثير لمتغيرات تصميم الاختبارات الإلكترونية على اتجاهات الطلاب فنجد على سبيل المثال أظهرت دراسة أسامة الهنداوي، نهاري الغبيشى (٢٠١٢) وجود تأثير للأنماط الاستجابة في الاختبارات الإلكترونية بصفة عامة على أداء الطلاب واتجاهاتهم نحو الاختبارات الإلكترونية . ويتفق ذلك مع ما توصلت اليه دراسة كلا من سليمان احمد حرب (٢٠١٨)؛ مفلح ال جديع(٢٠١٧) التي أظهرت وجود اتجاهات إيجابية

لدى أعضاء هيئة التدريس وطلاب كلية التربية نحو الاختبارات الإلكترونية .ويتضح مما سبق أهمية دراسة مدى الاختبارات مدى تأثير متغيرات تصميم الاختبارات الإلكترونية بصفة عامة والتكيفية بصفة خاصة على بعد الابعاد النفسية للمتعلم مثل مستوى قلق الاختبار والاتجاهات نحو الاختبارات الإلكترونية بصفة عامة.

إجراءات البحث

١- اجراءات بناء الاختبارات الإلكترونية

اطلع الباحث على مراحل بناء الاختبارات الإلكترونية أشار اليها عديد من الباحثين مثل (محمد محمود،٢٠١٧)؛ (ايناس محمد، ٢٠١٣)؛ (زاهر الغريب، ٢٠٠٩)؛ (محمد عطية خميس،٢٠٠٧). ويمثل نموذج محمد عطية خميس (٢٠٠٧) بمراحله الأساسية كما هو موضح بالمخطط (٥) الأنسب لتوجهات البحث الحالي ومن ثم تم توظيفه فى بناء الاختبارات التكيفية والتقليدية كمواد معالجة مع مراعاة خصائص ومعايير بناء تلك النوعية من الاختبارات فى الخطوات الفرعية من كل مرحلة من المراحل الأساسية للنموذج على النحو التالي:



مخطط (٥) نموذج محمد عطية خميس (٢٠٠٧)

أولا: مرحلة التحليل وشملت الإجراءات التالية:

١- تحليل مشكلة البحث وتحديد ابعادها:

تمثل مشكلة البحث فى الحاجة الى تطوير اختبارات تكيفية تراعى القدرات الفعلية للمعلمين والمقارنة بين أنماط الإنهاء المختلفة فى تلك النوعية من الاختبارات التكيفية والاختبارات الإلكترونية التقليدية، والتعرف على مدى اسهامها فى خفض مستوى قلق الاختبار لدى المتعلمين؛ وتكوين اتجاهات إيجابية لديهم نحو التوسع فى الاختبارات الإلكترونية بصفة عامة ودمج الاختبارات التكيفية فى مصفوفة تقييم الطلاب بالمقررات الدراسية فى مرحلة التعليم الجامعي.

٢- تحليل الأهداف العامة للاختبارات الإلكترونية بأنواعها المختلفة.

تم تحديد الأهداف العامة للاختبارات الإلكترونية مجال الدراسة من واقع توصيف مقرر تكنولوجيا التعليم للفرقة الثالثة المعتمد من القسم العلمي التابع له المقرر بكلية التربية جامعة المنوفية والموضح بملحق(٢) وقد تمثلت الأهداف العامة للمقرر فى التالي :

- التعرف على المصطلحات والمفاهيم الأساسية فى علم تكنولوجيا التعليم
- التمييز بين المراحل المختلفة من تطور العلم
- التعرف على مكونات منظومة تكنولوجيا التعليم
- التعرف على اهم التكنولوجيات المستخدمة في العملية التعليمية
- التميز بين أنظمة التعلم الإلكتروني المختلفة.
- اكتساب بعض مهارات توظيف التكنولوجيات المختلفة في مجال التخصص

٣- تحديد محتوى مفردات الاختبارات.

تم تحديد محتوى المفردات من واقع المحتوى العلمي لمقرر تكنولوجيا التعليم والذي يتناول المفاهيم الأساسية لعلم تكنولوجيا التعليم ومراحل تطوره واهم عناصره واسهاماته فى تطوير العملية التعليمية واهم المستحدثات التكنولوجية التي يتم توظيفها حاليا من تطبيقات وخدمات كذا اهم الاستراتيجيات التعليمية المستخدمة فى بيئات التعلم الافتراضية واهم المهارات الواجب توافرها فى المعلم لتوظيف المستحدث التكنولوجي. لموضوعات المحتوى.

٤- تحليل خصائص المتعلمين: تمثلت خصائص الفئة المستهدفة فى طلاب الفرقة الثالثة بكلية التربية من الدارسين لمقرر تكنولوجيا التعليم، لديهم مهارات التعامل مع التطبيقات والخدمات المتاحه من خلال الشبكات السلكية واللاسلكية من خلال الاجهزة الحاسوبية المكتبية والهواتف الذكية، وقد بلغ عدد طلاب وطالبات الفئة المستهدفة ٢٧٠ طالبا وطالبه شعب اللغة الإنجليزية عام واللغة الفرنسية تمثل المجتمع الأصلي لعينة البحث من الدراسين لمقرر

تكنولوجيا التعليم فى الفصل الدراسي الأول من العام الجامعي ٢٠١٩-٢٠٢٠م .

٥- تحليل بيئة العمل

قام الباحث بمراجعة الفنين والمسئولين فى وحدة الدعم والخدمات التكنولوجية بالكلية لتحديد واقع الخدمات والتطبيقات والأجهزة الحاسوبية فى الكلية عبر خدمات شبكة WIMAX، كما يمكن للطلاب الدخول عبر الهواتف الذكية من خلال

ثانيا: مرحلة التصميم وقد استلزمت الإجراءات التالية:

١- صياغة الأهداف التعليمية

تم صياغة الأهداف التعليمية للاختبارات وتوزيعها حسب موضوعات التعلم ونواتج التعلم لمقرر تكنولوجيا التعليم والموضحة فى توصيف المقرر بملحق (٢)

٢- تصميم بنوك الأسئلة

تمثل بنوك الأسئلة نموذج المجال Content Model في البيئة الاختبارية التكيفية ويحتوي بنك الأسئلة في البحث الحالي على ٤٠٠ مفردة اختبارية تم تحديد خصائصها السيكو مترية خلال العام الدراسي السابق للتجربة الحالية من واقع نتائج تطبيق الاختبارات التي تم تصحيحها اليا لمقرر تكنولوجيا التعليم من خلال برنامج Remark الإصدارات (٨، ١٠) حيث يقدم البرنامج تحليلات إحصائية من واقع استجابات الطلاب على كل مفردة من مفردات الاختبار وقد تراوحت معاملات الصعوبة لمفردات البنك ما بين (٠.٤-٠.٨) والذى يشير الى مستوى مناسب موضح بملحق (٣) معاملات صعوبة المفردات .وقد تم بناء قاعد بيانات البنك بحيث يحتوى على رقم المفردة ونوعها من بين أنماط الأسئلة الموضوعية ونص او جذر المفردة ثم البدائل التابعة لها ومستوى الصعوبة والتمييز الخاص بها، واكواد موضوع التعلم المرتبط بها ، واكواد مخرج التعلم والمستوى المعرفي للمفردة . موضح بنك الاسئلة بملحق(٤) ٢- تصميم نموذج المتعلم

وفقا لطبيعة البحث الحالي والتي تستهدف المقارنة بين نمطين الاختبارات التكيفية تمثلت إجراءات تصميم نموذج المتعلم فى كلا النمطين كالتالي :

أ- تصميم نموذج المتعلم فى الاختبارات التكيفية ثابتة الإنهاء (الطول الثابت)

ارتكز هذا النموذج فى تقدير أداء المتعلم على الخوارزميات التي أشار اليها كلا من (Olea & Ponsoda,1996)؛ (Wauters 2010 et al., فى الإطار النظري للبحث الحالي والذي يعتمد على طرح مفردة متوسطة الصعوبة فى بداية الاختبار ووفقا إجابة المتعلم يتم اختيار مفردات اعلى او أدني فى الصعوبة حتى انتهاء مفردات الاختبار المحددة سلفا بعدد ٤٠ مفردة مع اختلاف عرض ترتيب البدائل التابعة لكل سؤال موضوعي من ممتحن لأخر من خلال

تفعيل خاصية shuffle في عرض الأسئلة على عينة البحث.

ب- نموذج المتعلم فى الاختبارات التكيفية متغيرة الإنهاء (الطول المتغير)

يرتكز هذا النموذج على اسلوب الاختبار القبلي القصير الذي يعتمد على طرح (٥) مفردات متنوعة الصعوبة قبل بداية الاختبار ثم يقوم البرنامج بحساب متوسط استجابات المتعلم لتحديد مستوى القدرة الفعلي ومن ثم تحديد اول مفردة فى الاختبار من حيث مستوى الصعوبة حتى يصل المتعلم الى مرحلة الثبات فى القدرة والمقدر بعدد (٥) استجابات صحيحه او خاطئة متصلة فى متسلسلة عرض مفردات الاختبار كأساس فى انهاء سحب الأسئلة ذات الصلة بموضوع التعلم الحالي والانتقال الى أسئلة الموضوع التالي.

رابعا: مرحلة التطوير واشتملت على الإجراءات التالية:

أولا: اعداد السيناريو التعليمي

١- سيناريو بناء الاختبار الإلكتروني التقليدي

يرتكز السيناريو الخاص بالاختبار الإلكتروني التقليدي على التوليد العشوائي للمفردات من بنك الأسئلة دونما اعتبار للاستجابات المتعلم ويحتوى هذا البنك على عدد ٤٠٠ مفرده وهو نفس العدد من المفردات فى الاختبارات التكيفية، ويوضح (ملحق ٥) الشاشة الافتتاحية وتعليمات الاختبار وتم عرض كل سؤال في صفحة مستقلة به ازرار الانتقال بين الأسئلة التالية او السابقة وتسليم الاستجابة وكذا عناصر الوسائط المتعددة في كل مفردة من صوت وصوره ونص او فيديو وقد اشتمل الاختبار على عدد أجمالي ٤٠ مفردة مرتبة عشوائيا من حيث نمط المفردة ومستوى صعوباتها.

٢- سيناريو بناء الاختبار الإلكتروني التكيفي ثابت الإنهاء

ارتكز سيناريو بناء الاختبار التكيفي القائم على الإنهاء ثابت الطول على السيناريو المرئي لنظام التقييم الذي قدمته دراسة (Lazarinis, Retalis, 2007). كما هو موضح بالمخطط (6) حيث يتم طرح مفرده متوسطة الصعوبة على المتعلم فى بداية الاختبار من الموضوع الأول ، إذا كانت إجابة الطالب صحيحه ينتقل الطالب الى مفردة الأكثر صعوبة في نفس موضوع التعلم، ام إذا كانت إجابة الطالب خطأ ينتقل الى المفردة الأقل منها في مستوى الصعوبة حسب متوالية تنظيم المفردات في موضوع التعلم ،ويتم تكرار ما سبق حتى يتم الانتهاء من جميع المفردات المحدد عددها مسبقا ب(٤٠) مفردة . تتمثل قاعدة القفز بين موضوعات التعلم على تحقيق شرط الانتهاء من اعلى المفردات صعوبة في نفس موضوع التعلم للانتقال الى موضوع التعلم الذي يليه. ويتم حساب الدرجة الكلية لكل طالب وفقا للمعادلة التالية التي تعتمد على جمع الدرجة

المخصصة لكل مفردة والتي تمثل حاصل ضرب مستوى الصعوبة للسؤال في عدد المفردات في نفس موضوع التعلم

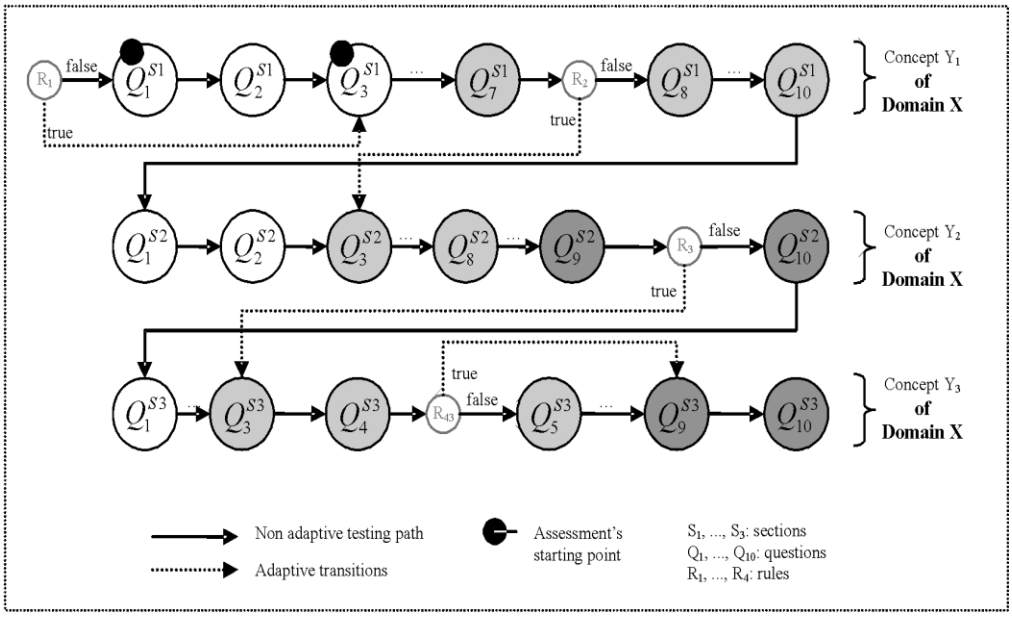
$$MS = w1*MQ1 + w2*MQ2 + \ldots + wn*MQn$$

حساب الدرجة الكلية

n: number of section questions عدد المفردات في موضوع التعلم

wi: difficulty weight of each question مستوى صعوبة السؤال

Qi: mark in each question الدرجة المخصصة للسؤال



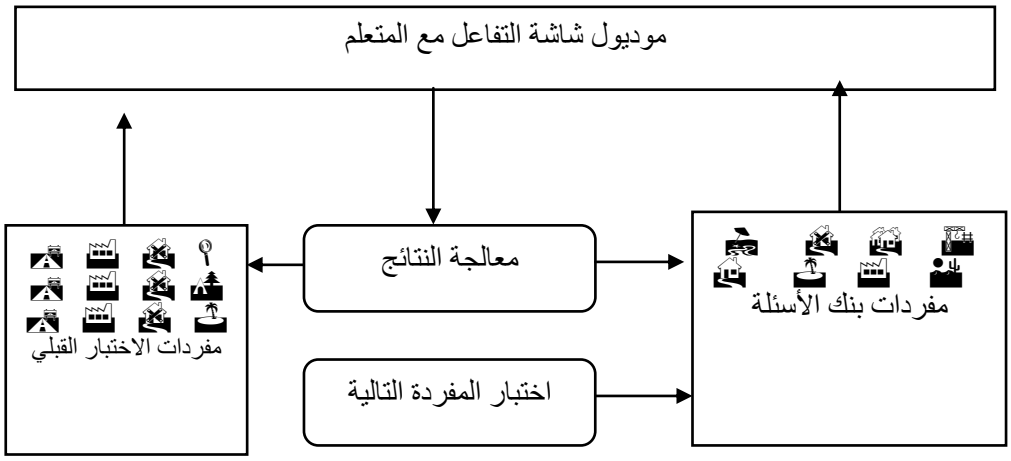
مخطط (٦) السيناريو المرئي لخطوات الانتقال بين مفردات الاختبار

Visualization of assessment scenario

Adapted by Lazarinis and Retalis, (2007)

٣- سيناريو بناء الاختبار الإلكتروني التكيفي متغير الإنهاء

يعتمد سيناريو الانتقال بين المفردات فى الاختبار متغير الإنهاء على المسار الموضح فى المخطط (٧) حيث تم طرح مجموعة من المفردات متباينة فى مستويات الصعوبة ما بين (٠.٤-٠.٩) ومعامل تمييزاي (٠.٣-) فى جميع موضوعات التعلم كاختبار قبلي والتي تتمثل فى عدد (١٢) مفردة بواقع (٣) مفردات من كل موضوع تعلم ولا يتكرر ظهورها فى الاختبار الفعلي وبناء على استجابات الطالب يتم تقدير مستوى أداء المتعلم فى كل موضوع من تلك الموضوعات ومتسلسلة طرح المفردات التالية من بنك الاسئلة والتي يختلف عددها من ممتحن للأخر حسب استجابته.



مخطط (٧) سيناريو الانتقال بين مفردات الاختبار التكيفي متغير الإنهاء

ثانيا: تطوير الاختبار الإلكتروني التقليدي

اعتمد الباحث على مجموعة office365في اعداد عناصر الوسائط المتعددة من أصوات وصور ورسوم ونصوص كمتطلب اعداد محتوى المفردات الاختبارية كما استخدم الباحث نظام إدارة التعلم https://www.schoology.com في بناء الاختبار الإلكتروني التقليدي حيث قام الباحث بإنشاء مساحة عمل لمقرر تكنولوجيا التعليم على هذا النظام

ثانيا: تطوير الاختبارات الإلكترونية التكيفية ثابتة ومتغيرة الإنهاء .

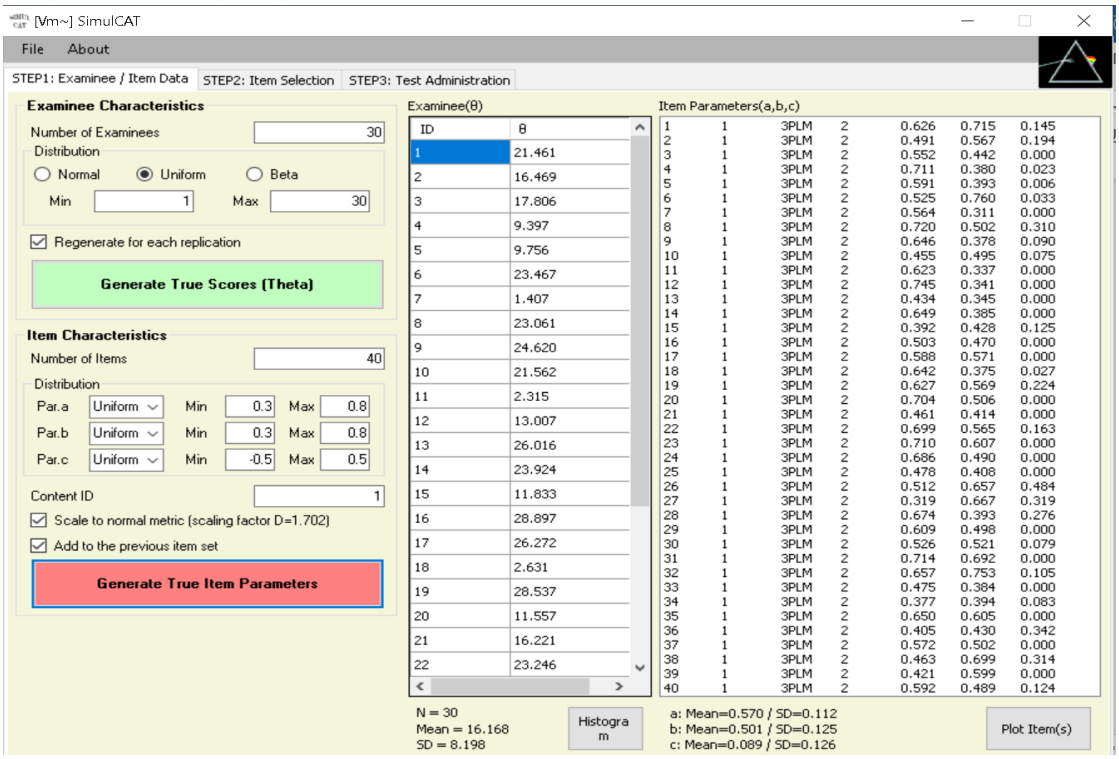
ارتكز الباحث على منصتين اساسيتين في تطوير وإنتاج الاختبارات التكيفية الثابتة الإنهاء والمتغيرة الإنهاء وفقا الإجراءات التالية:

١- اعتمد الباحث على برنامج SimuICAT المتخصص في تحديد بارامترات الاختبار التكيفي الثابت والمتغير الطول ذات الصلة بخصائص الممتحنين، خصائص المفردات ومواصفاتها السيكو مترية وكذا الخوارزميات الرياضية اللازمة لتقدير المستوى الفعلي للمتعلم وفقا نظرية الاستجابة للمفردة، وقد تم تصدير ملف بارامترات الاختبار الى منصة نشر وتوزيع الاختبار التكيفي الثابت والمتغير الطول للتحكم في اليات سحب المفردات الاختبار من بنك الأسئلة (مرفق (٨) دليل استخدام برنامج SimuICAT في تحديد بارامترات الاختبارات التكيفية الثابتة والمتغيرة الطول)

ويوضح الشكل (٨) إجراءات معايرة مفردات بنك الأسئلة Item Bank Calibration وتحديد قيم القدرات المبدئية المحتملة لكل مفردة من مفردات الاختبار التكيفي متغير الطول وفقا الخطوات التالية :

- تحديد البيانات ذات الصلة بالممتحنين ومفردات الاختبار
  - من حيث عدد الممتحنين في كل مجموعة من مجموعات الاختبار

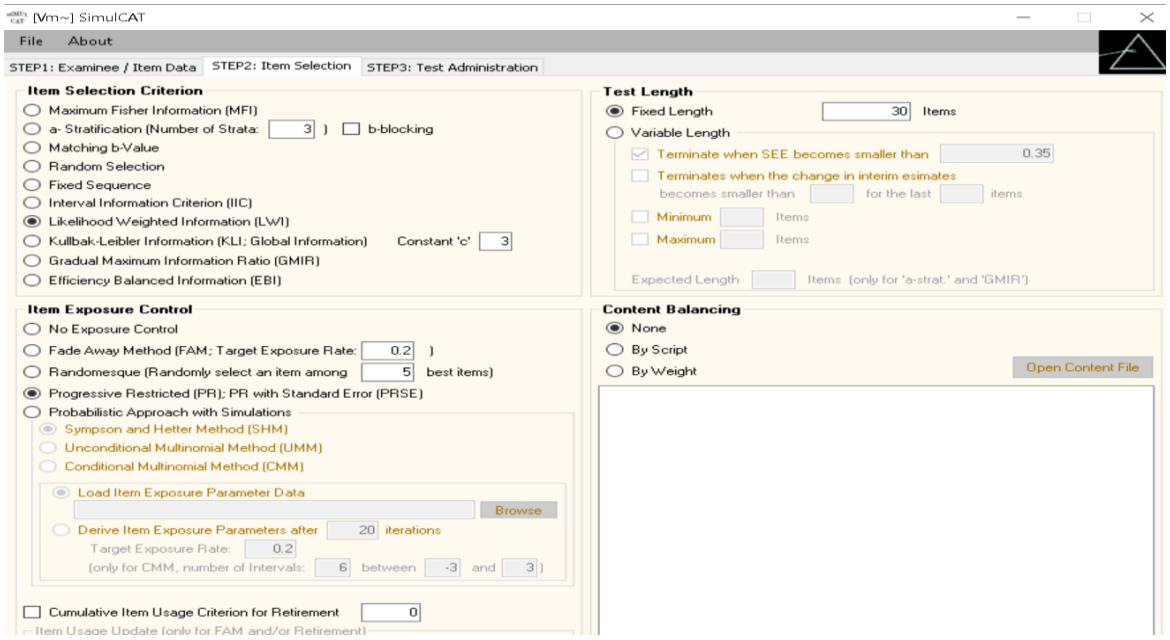
التكيفي سواء ثابت الإنهاء او متغير الإنهاء .



شكل (٨) تحديد خصائص الممتحنين والخصائص السيكو مترية للمفردات حسب مستويات صعوباتها

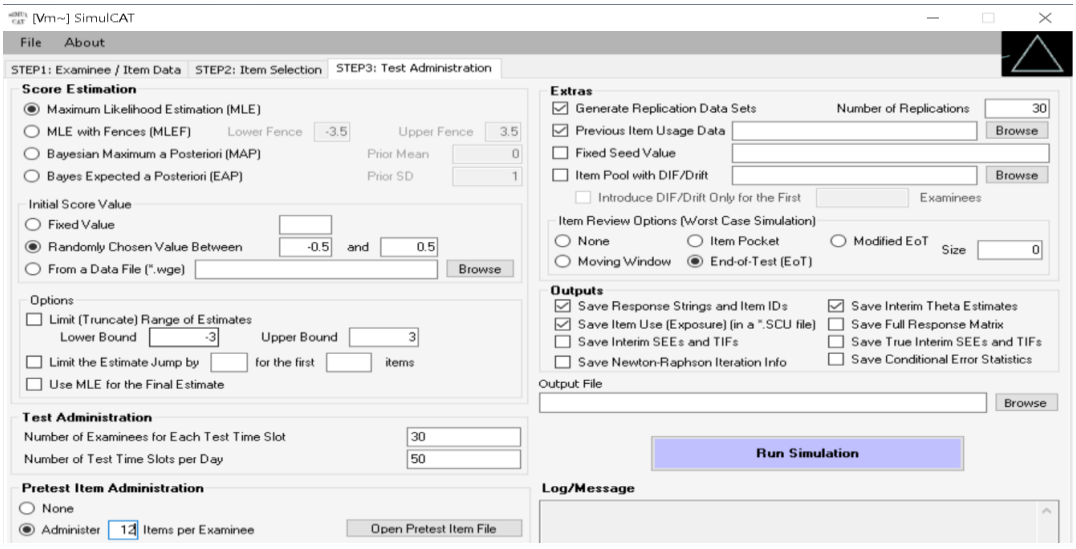
٢ـ محكات سحب المفردة من بنك الأسئلة

يوضح شكل (٩) محكات اختيار المفردات من بنك الأسئلة وفقا الخوارزميات الموضحة في الشكل وقد استخدم الباحث دالة تقدير أداء المتعلم like hood weight (information) حيث يتم سحب المفردة الأولى وتتابع طرحها وفق قيم متوسطات لمعامل الصعوبة والتمييز ، وكذا يتوقف سحب المفردات في الاختبارات التكيفية عند قيمة ٠.٢٥لمعامل الخطأ المعياري كشرط انهاء الاختبار.



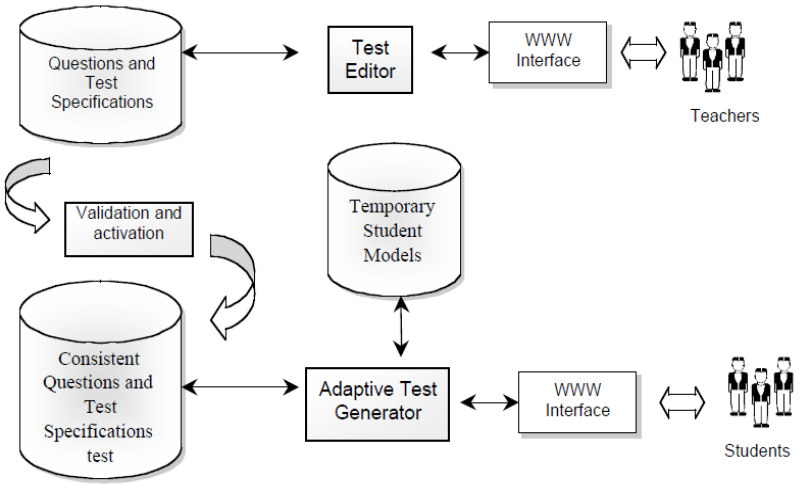
شكل (٩) محكات سحب المفردات وفقا التقدير المتوقع لقدرات المتعلم القائم على خوارزميات الارجحية العظمى

٣- انشاء ملف إدارة الاختبار التكيفي كما هو موضح بالشكل(١٠) ويشمل تحديد نظام حساب الدرجة الكلية للطالب وعدد مفردات الاختبار القبلي ومعدل سحب المفردات وفقا تقدير أداء المتعلم في المفردة الحالية وفقا أدائه في المفردة السابقة وزمن اتاحة الاختبار يوميا للممتحنين.



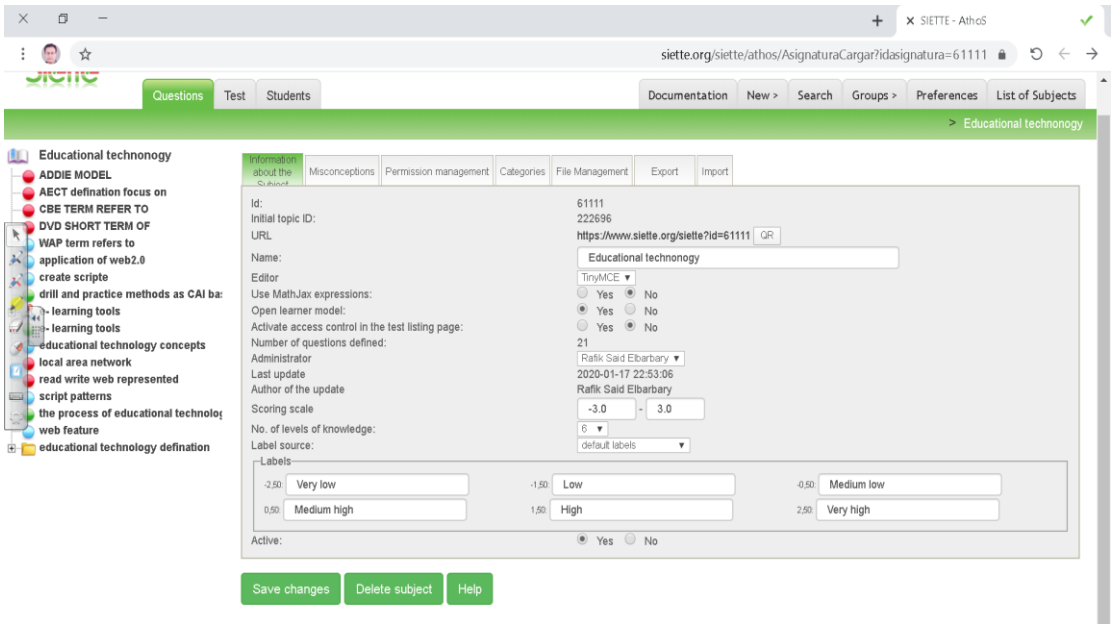
شكل (١٠) إدارة نظام الاختبار التكيفي

٢- اعتمد الباحث على انشاء حساب ذو غرض بحثي وليس تجارى على منصة https://www.siette.org. في انتاج الاختبارات التكيفية ثابتة ومتغيرة الإنهاء وهي أحد المنصات المتخصصة التي اعتمدت على الدمج بين لغات الذكاء الاصطناعي مثل Python27 ولغات تطوير خدمات الويب مثل PHP,XHTML في تطوير الاختبارات التكيفية من قبل مجموعة من الخبراء والمتخصصين في برمجيات الذكاء الاصطناعي والقياس النفسي كما يرتكز على نظرية الاستجابة للمفردة في بناء المفردات ويسمح هذا النظام بالتعامل مع أنظمة إدارة التعلم من حيث استيراد او تصدير ملفات بنوك أسئلة وملفات خصائص الطلاب وكذا التعامل مع خلال الهواتف الذكية، كما يتميز هذا النظام الخبير بإمكانية إضافة او حذفة موضوعات يتم تناولها في الاختبار وما يرتبط بها من مفردات وكذا المفاهيم الخاطئة الشائع ظهوره في كل موضوع من موضوعات التعلم التي يتناولها الاختبار وكذا عرض النتائج ومخططات تتبع أداء المتعلم ويوضح شكل (١١) المخطط العام لنظام siette الذى يحتوى على المعلم التي تتكون من محرر لاختبار التكيفي وبنك أسئلة يشتمل على مفردات تم تحديد موصفاتها وفقا لنظرية الاستجابة للمفردة ، وواجهة التفاعل مع الطلاب التي تعتمد على استجابات الطلاب في بناء بيانات نموذج الطالب



شكل (١١) المخطط العام للنظام الخبير siette (Conejo et al., 2004)

ويوضح الشكل (١٢) الوجه الأساسية للنظام الخبير والتي تحتوي على ثلاث تبويب ذات الصلة بإجراءات اعداد المفردات في كل موضوع من موضوعات التعلم ، الاختبار، الطلاب



خامسا: مرحلة التقويم

تم تجريب الاختبارات الإلكترونية التقليدية والتكيفية على عينة استطلاعية من (٢٠) طالبا وطالبه من طلاب الفرقة الثالثة ومن الدارسين لمقرر تكنولوجيا التعليم وفقا لتوصيف المقرر ذاته الذي تم الاعتماد عليه في بناء بنوك الأسئلة وقد أظهرت نتائج التجريب الاستطلاعي ضرورة اجراء بعد التعديلات ذات الصلة بضبط متغيرات نظام الاختبار التكيفي من حيث ضبط القدرة المحتملة لكل مفردة من مفردات الاختبار وكيفية رفع جداول قواعد البيانات ذات الصلة بإداء الطلاب بشكل دوري على مصفوفة انتاج المفردات من بنك الأسئلة.

سادسا : مرحلة النشر والتوزيع

١- تم نشر محتوى الاختبار الإلكتروني التقليدي على نظام schoology وتوزيع كود دخول الطلاب رقم **VJ7ZX-M5HHF** على الطلاب عبر الحساب الخاص بهم على هذا النظام وقبول انضمامهم لنشاط اختبار اعمال السنه من خلال الإجراءات التالية:

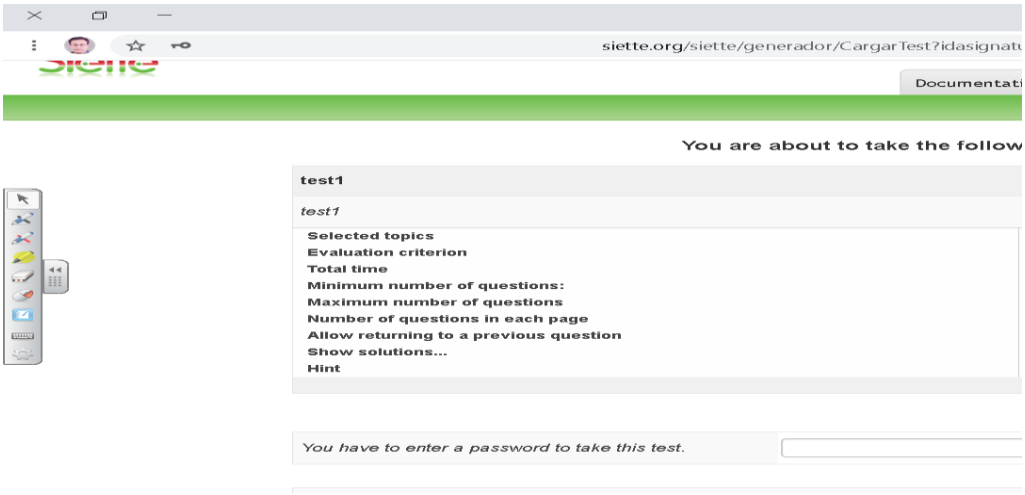
- ارسال اشعارات للطلاب عينة البحث الموجه لهم الاختبار
- انشاء بنك اسئلة على نظام ادارة التعلم السابق ذكره يحتوي على (٤٠٠) مفرده.
- سحب مفردات الاختبار الإلكتروني التقليدي والبالغ عددها (٤٠) مفرده من هذا البنك من قبل أستاذ المقرر حيث تم ترتيب الأسئلة بشكل مسبق من حيث التدرج من السهولة الى الصعوبة وكذا الترتيب العشوائي لمفردات وبدائل

الأسئلة من حيث تتابع عرضها على المتعلم.

- توفير صلاحيات تمكن الطلاب من مشاهدة نتائج الاختبار والدرجة الكلية ومراجعة استجابتهم من حيث الصواب والخطأ.

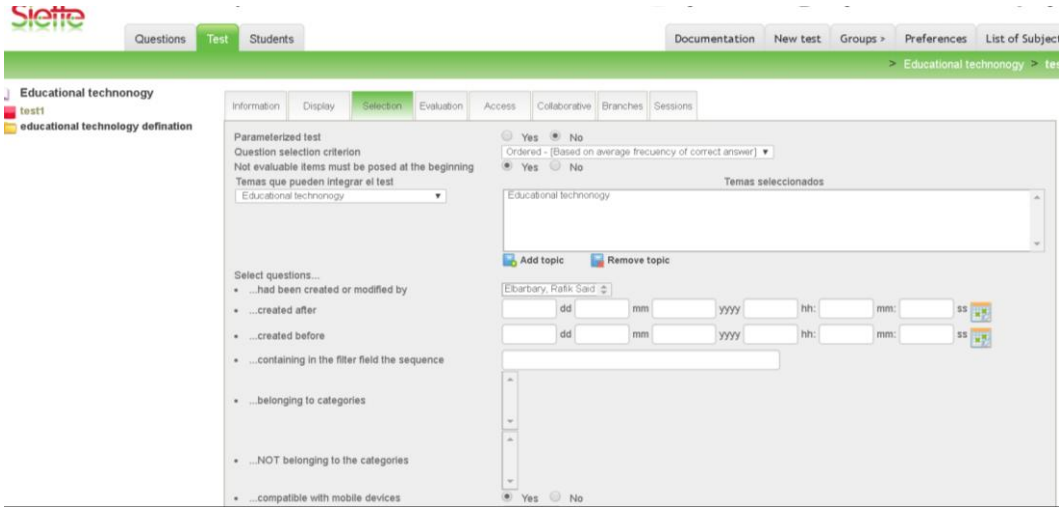
٢- تم نشر الاختبارات التكيفية ثابتة ومتغيرة الطول عبر https://www.siette.org.

ويوضح الشكل (١٣) كود الدخول على رابط الاختبار



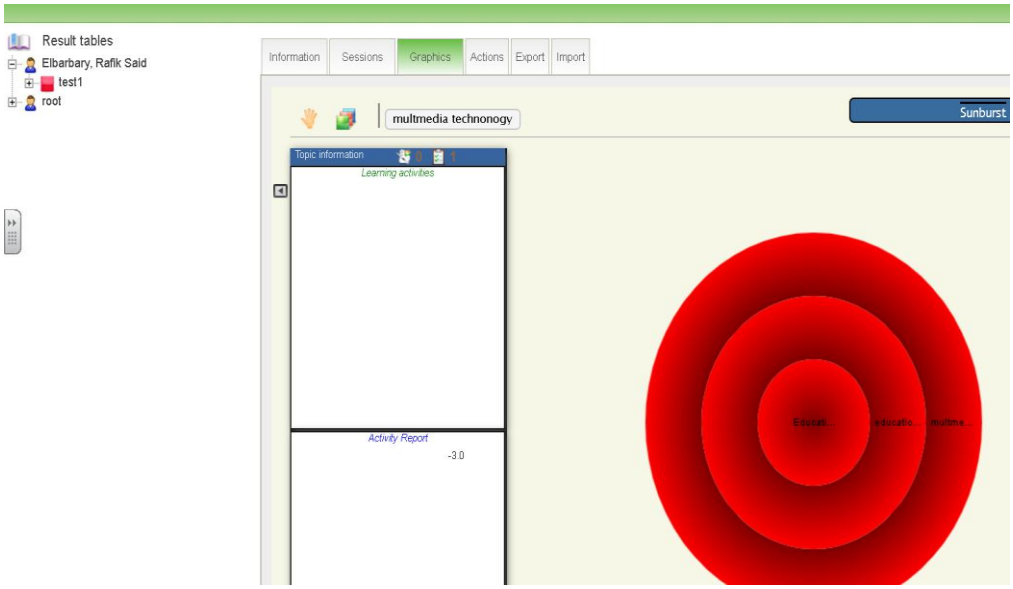
شكل (١٣) كود دخول الاختبار التكيفى على منصة SIETTE

كما يوضح شكل (١٤) واجهة ضبط بارامترات الاختبار الثابت والمتغير الطول على مجموعتي البحث وتوقيت نشره على أجهزة الطلاب



شكل (١٤) واجهة ضبط بارامترات الاختبار التكيفي وتوقيت نشره على أجهزة الطلاب المختلفة

ويوضح شكل (١٥) واجهة شاشة عرض نتائج الطلاب ونتائج المعالجات الاحصائية



شكل (١٤) واجهة شاشة عرض النتائج على الطالب في كل موضوع من موضوعات الاختبار

٤- اجراءات اعداد أدوات البحث

أولا: إجراءات اعداد مقياس اتجاهات الطلاب نحو الاختبارات الإلكترونية

ارتكز الباحث فى اعداد مقياس الاتجاهات نحو الاختبارات الإلكترونية على مراجعة الادبيات ذات الصلة بأعداد مقاييس الاتجاهات نحو التعليم الإلكتروني بصفة عامة والاختبارات الإلكترونية بصفة خاصة مثل: مقياس سليمان حرب (٢٠١٨)، عبد العزيز الجنزوري (٢٠١٧)، مقياس محمد عطا الله (٢٠١٦)، رباب البلاصى (٢٠١٦)، مروة الباز (٢٠١٣)، نضال الطعانى (٢٠١١)، وقد تم توزيع بنود المقياس على الابعاد التالية:

- الاتجاهات نحو التوسع فى استخدام الاختبارات الإلكترونية فى منظومة القياس والتقويم.
- الاتجاهات نحو اشكال دمج الاختبارات الإلكترونية فى منظومة القياس والتقويم

المقياس فى صورته النهائية من ثلاثون فقرة (٣٠)، ويوضح جدول (٢) مواصفات فقرات المقياس.

جدول (٢) مواصفات فقرات مقياس الاتجاهات نحو الاختبارات الإلكترونية

الفقرات / البعد	الفقرات الموجبة	المجموع	الفقرات السالبة	المجموع	المجموع الكلى	النسبة المئوية
الاتجاهات نحو التوسع فى تطبيق الاختبارات الإلكترونية	١،٢،٣،٤، ٥،٧،٨، ١٤،١٥	٩	٦، ٩، ١٠، ١١، ١٢، ١٣	٦	١٥	%٥٠
الاتجاهات نحو الدمج فى منظومة الامتحانات	١٦،١٧، ١٨، ١٩، ٢١، ٢٢، ٢٣، ٢٥، ٢٦، ٢٧، ٢٨،٢٩	١٢	٢٠، ٢٤،٣٠	٣	١٥	%٥٠
المجموع		٢١		٩	٣٠	%١٠٠

تقدير الاستجابات على فقرات المقياس

استخدم الباحث مقياس ليكرات خماسي الابعاد (موافق بشدة، موافق، محايد، غير موافق، غير موافق بشدة)، بحيث يعطى التدرج التنازلي للعبارات الموجبة والتصاعدي للعبارات السالبة.

التحقق من موضوعية المقياس:

أولا: التحقق من صدق المقياس

اعتمد الباحث على مدخلين فى تقدير مستوى صدق المقياس فى تحقيق الأهداف التي وضع من اجلها كالتالي:

- صدق المحكمين:

عرض الباحث المقياس فى صورته الأولية على مجموعة من المحكمين فى مجال تكنولوجيا التعليم وعلم النفس ملحق (٦). وقد قام الباحث بحساب نسب الاتفاق بين المحكمين على فقرات المقياس وفق معادلة كوبر حيث تم اعتبار الفقرة الصالحة فى حال الاتفاق فى الآراء بنسب تتعدى ٧٥% (Chang, Cooper,1970).

- صدق الاتساق الداخلي

قام الباحث بتطبيق المقياس على عينة استطلاعية عددها (٢٠) طالبا مرتين بفاصل زمني اسبوعان ثم استخدم الباحث معامل

بيرسون (Pearson) فى حساب معامل الارتباط بين كل فقره من فقرات البعد والدرجة الكلية له لتحديد دلالة الاتساق الداخلي وموضح بالجداول (٣)، (٤) نتائج حساب معاملات الارتباط لكل بعد من ابعاد المقياس.

جدول (٣) معاملات ارتباط فقرات بعد الاتجاهات نحو التوسع فى تطبيقات الاختبارات الإلكترونية

رقم الفقرة	معاملات الارتباط	الدلالة	رقم الفقرة	معاملات الارتباط	الدلالة
١	٠.٦٥	٠.٠١	٢	٠.٦٢	٠.٠١
٣	٠.٨٦	٠.٠١	٤	٠.٨١	٠.٠١
٥	٠.٦٥	٠.٠١	٦	٠.٧٦	٠.١
٧	٠.٧١	٠.٠١	٨	٠.٦٥	٠.٠١
٩	٠.٧١	٠.٠١	١٠	٠.٣٧	٠.٠١
١١	٠.٥٨	٠.٠١	١٢	٠.٥٤	٠.٠١
١٣	٠.٧٦	٠.٠١	١٤	٠.٥٨	٠.٠١
١٥	٠.٤٨	٠.٠١			

جدول (٤) معاملات ارتباط فقرات بعد الاتجاهات نحو اشكال دمج تطبيقات الاختبارات الإلكترونية فى منظومة القياس والتقويم

رقم الفقرة	معاملات الارتباط	الدلالة	رقم الفقرة	معاملات الارتباط	الدلالة
١٧	٠.٦١	٠.٠١	١٦	٠.٧٤	٠.٠١
١٩	٠.٥١	٠.٠١	١٨	٠.٧١	٠.٠١
٢١	٠.٥٤	٠.٠١	٢٠	٠.٥٨	٠.٠١
٢٣	٠.٧٩	٠.٠١	٢٢	٠.٨١	٠.٠١
٢٥	٠.٦٦	٠.٠١	٢٤	٠.٨٤	٠.٠١
٢٧	٠.٨٢	٠.٠١	٢٦	٠.٥٤	٠.٠١
٢٩	٠.٧١	٠.٠١	٢٨	٠.٦٤	٠.٠١
			٣٠	٠.٦١	٠.٠١

- التحقق من ثبات المقياس

تم التحقق من ثبات المقياس من خلال معامل ألفا كرو نباخ، حيث تم حسابها لفقرات المقياس ككل وبلغت ٠.٧٤ وهي قيمه ذات دلالة على ارتفاع الثبات ويوضح ملحق (٦) المقياس فى صورته النهائية والرابط الموضح ادناه يمثل صورته النهائية الإلكترونية التي تطبيقها على المجموعات عينة البحث

https://docs.google.com/forms/d/1W_thZ2qGB4ncESITQ1x7pgKT1ZQSuiP4rRGX9c4u8Xc/edit.

ثانيا : خطوات بناء مقياس قلق الاختبار

قام الباحث بمراجعة العديد من مقاييس قلق الاختبار مثل: مقياس Richard Driscoll (٢٠٠٤) العضو بالجمعية الامريكية لقلق الاختبار، مقياس STAB Suinn(1969) الذي يمثل الإطار المرجعي في بناء تلك النوعية من المقاييس.

ويحتوي مقياس القلق فى البحث الحالي على عدد (٢٠) بنداً موزعه على الابعاد التالية:

- قلق موضوعي مرتبط بطبيعة الاختبار.
- قلق مرتبط بظروف تطبيق الاختبار.
- قلق مرتبط بعلاقة المتعلم بإقرانه ومستوى الثقة بالنفس.

تقدير الاستجابات على فقرات المقياس.

استخدم الباحث مقياس ليكرات خماسي الابعاد (دائما، عادة، احيانا، نادرا، لا يحدث)، حيث اعطى الباحث قيم الدرجات (٥، ٤، ٣، ٢، ١) حسب ترتيب الابعاد السابقة. أي ان اعلى مستوى للقلق من حيث القيمة دائما (٥) درجات، عادة (٤)، أحيانا (٣)، واقل قيمة (١) لا يحدث.

حساب صدق المقياس

عرض الباحث المقياس فى صورته الأولية على مجموعة من المحكمين فى مجال علم النفس ملحق (٤). وقد قام الباحث بحساب نسب الاتفاق بين المحكمين على فقرات المقياس وفق معادلة كوبر حيث تم اعتبار الفقرة صالحة فى حال الاتفاق فى الآراء بنسب تتعدى ٧٥% (Chang, Cooper,1970).

حساب ثبات المقياس

تم حساب ثبات المقياس ككل من خلال تطبيق المقياس ثم إعادة تطبيقه مرة ثانية بفاصل زمنى أسبوعين على عينة من (١٥) طالب من طلاب الفرقة الثالثة لغة انجليزية ثم استخدام قيم معاملات الارتباط وفقا معادلة بيرسون بين درجات التطبيق الأول والثانى لتقدير قيم الثبات فى كل بعد من ابعاد المقياس حيث بلغت قيمة معامل الثبات لبعد القلق الموضوعي (٠.٧١)، وكذلك بلغت قيمه معامل الارتباط لبعد القلق من ظروف تطبيق الاختبار (٠.٨٤)، بينما بلغت قيمة معامل الارتباط الخاص ببعد مستوى الثقة بالنفس والعلاقة بين الاقران الى (٠.٧٦) وتشير القيم السابقة الى مستوى ثبات

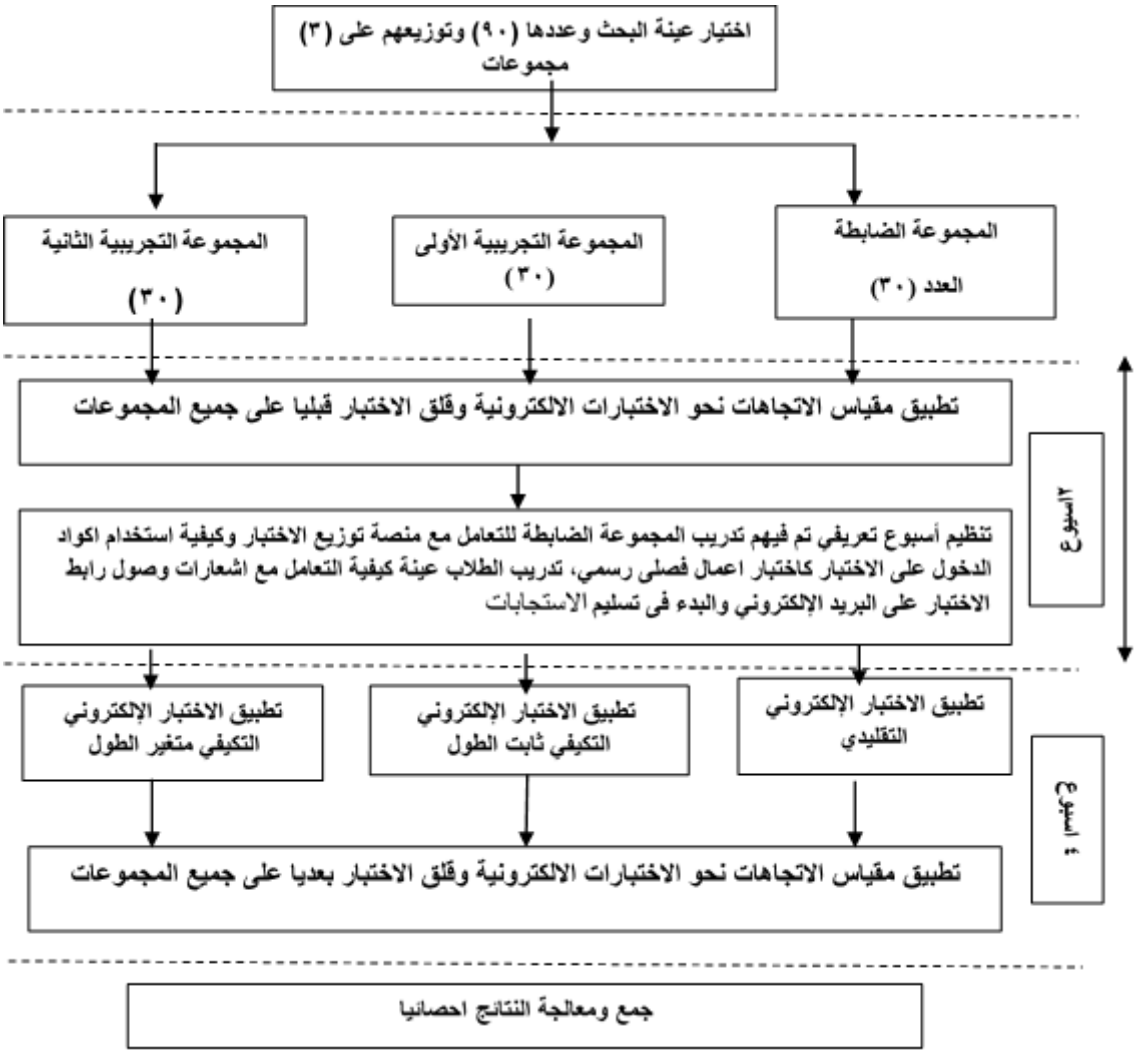
مناسب ومن ثم صلاحية استخدمه فى البحث الحالي.

ويوضح ملحق (٧) الصورة النهائية لمقياس القلق والرابط الموضح ادناه يمثل صورته النهائية الإلكترونية .

https://docs.google.com/forms/d/1FOdfTbRsD4-tRIe3gf7mqrJVow8VRyN20suntBudH.Bo/edit

إجراءات تنفيذ التجربة

يوضح الشكل (١٤) إجراءات تنفيذ التجربة من حيث المدى الزمنى لكل مرحلة وتعامل المجموعات التجريبية والضابطة مع مواد المعالجة التجريبية وفقا التتابع التالي:



شكل (١٤) إجراءات تنفيذ التجربة

أ- تقسيم عينة البحث الى ثلاث مجموعات (مجموعتين تجريبيتين ومجموعة وأحد ة ضابطة) بواقع ٣٠طالبا وطالبه فى كل مجموعة من تلك المجموعات.

ب- تطبيق أدوات البحث قبليا على جميع المجموعات والمتمثلة فى كل من مقياس قلق الاختبار والاتجاهات نحو الاختبارات الإلكترونية . وللتأكد من تكافؤ المجموعات قام الباحث بالمعالجات الإحصائية التالية:

١- للتأكد من تكافؤ المجموعات فى التطبيق القبلي لكل من مقياس الاتجاهات نحو الاختبارات الإلكترونية ومستوى قلق الاختبار. تم استخدام اختبار " تحليل التباين أحادي الاتجاه" لتحديد قيمة "ف" ومستوى دلالتها كمؤشر على حالة التكافؤ بين المجموعات ويوضح جدول(5)، (6) نتائج اختبار تحليل التباين وقيم "ف" فى كلا المقياسين.

جدول (٥) قيمة "ف" ودلالتها الإحصائية للفروق بين متوسطات الدرجات فى التطبيق القبلي لمقياس الاتجاهات نحو الاختبارات الإلكترونية

مصدر التباين	مجموعات مربعات الدرجات	درجات الحرية	متوسط المربعات الدرجات	قيمة "ف"	الدلالة
بين المجموعات	١١٦.١٥	٢	٥٨.١	١.٤١	٠.٢٤٩
داخل المجموعات	٣٥٧٥.٨	٨٧	٤١.١		
المجموع الكلى	٣٦٩١.٩٥	٨٩			

جدول (٦) قيمة "ف" ودلالتها الإحصائية للفروق بين متوسطات درجات التطبيق القبلي لمقياس قلق الاختبار على مجموعات البحث

مصدر التباين	مجموعات مربعات الدرجات	درجات الحرية	متوسط المربعات	قيمة "ف"	الدلالة
بين المجموعات	١٢١٧.٢	٢	٦٠٨.٦	٢.٤	٠.٤
داخل المجموعات	١٥٧٥٥.٣	٨٧	١٨١.١		
المجموع الكلى	١٦٩٧٢.٥	٨٩			

يتضح من الجداول (٥)، (٦) ان قيم "ف" المحسوبة على الترتيب (١.٤١- ٢.٤) لمتوسطات درجات التطبيق القبلي لكل من مقياس الاتجاهات نحو الاختبارات الإلكترونية

وقلق الاختبار، وبالرجوع الى قيمة "ف" الجدولية الموضح بالرابط :

http://socr.ucla.edu/Applets.dir/F_Table.html . .

عند قيم درجات الحرية الموضحة بالجداول السابقة وجد الباحث ان قيمه "ف" (3.8046)

مما يعنى ان قيمة ف المحسوبة < من "ف" الجدولية والذي يعنى عدم وجود فروق بين متوسطات الدرجات ومن ثم تحقيق شروط التكافؤ فى كلا المقياسيين .

ج- تدريب طلاب عينة المجموعة الضابطة على كيفية انشاء حسابات على نظام إدارة التعلم https://www.schoology.com وكيفية استخدام اكواد الوصول للدخول على الاختبار الذى تم التعامل معه على انه اختبار اعمال رسمي لمقرر تكنولوجيا التعليم لضمان تعامل العينة مع الاختبار بشكل يحاكى ظروف الاختبار الحقيقي، وقد تم ارسال اشعارات رسمية للطلاب بموعد الاختبار والدرجة الكلية وتوزيع الدرجات على كل مفردة من مفردات الاختبار وموعد تسليم الإجابات وقد تم تحديد زمن انهاء الاختبار الإلكتروني التقليدي ب(٥٠) دقيقة في ضوء نتائج تطبيق الاختبار الإلكتروني التقليدي على عينة مكونه من (٢٠) طالبا ثم حساب متوسط انهاء طلاب تلك العينة للاختبار الإلكتروني التقليدي.

د- تم تدريب الطلاب عينة البحث على انشاء حساب عبر منصةhttps://www.siette.org كما تم ارسال اكواد وصول طلاب المجموعتين التجريبيتين الى الاختبارات التكيفية، عبر البريد الإلكتروني لكل طالب موضح شاشات ارسال محتوى الاختبارات واشعارات الوصول موضح بملحق

(٥) ويشير الباحث ان زمن الاختبار في المجموعتين التجريبيتين لم يتم تحديده لكون انهاء الاختبار يتم من قبل البرنامج بعد نفاذ عدد المفردات المقرر طرحها فى الاختبارات التكيفية ثابتة الطول وكذا وفقا حسابات ذات صلة بتقدير استجابات المتعلم وفقا مستوى صعوبة وعدد المفردات التي أجاب عليها سواء بالصواب او الخطأ كشرط من شروط انهاء الاختبار .

د- تطبيق كل من مقياس قلق الاختبار والاتجاهات نحو الاختبارات الإلكترونية بعديا

هـ - جمع وتصنيف بيانات التطبيق من حيث تعامل الطلاب مع الاختبارات التقليدية والتكيفية واستجابات الطلاب على كلا المقياسين إلكتروني وتصدير ملفات ردود الطلاب على عبارات المقاييس الى برنامج التحليل الإحصائي SPSS version24.

و- اعتمدت المعالجة الإحصائية على تحليل نتائج التطبيق البعدي لكل من مقياس

الاتجاهات نحو الاختبارات الإلكترونية وقلق الاختبار باستخدام اختبار "ت" للمجموعات المستقلة فى المقارنة بين متوسطات درجات المجموعات الثلاث حسب الفرضيات الصفرية والموجه للبحث.

نتائج البحث

أولا: اختبار صحة الفرضية الصفرية الأولى للبحث وقد تطلب ذلك اختبار صحة الفرضيات الفرعية التابعة لها وفقا المعالجات التالية:

١- نتائج اختبار صحة الفرضية الصفرية الأولى .

يوضح جدول (٧) قيم المتوسطات (م) والانحرافات المعيارية (ع) ودلالة الفروق لقيمة "ت" المحسوبة لكل من المجموعة الضابطة والتجريبية الأولى فى التطبيق البعدي لمقياس مستوى قلق الاختبار.

جدول (٧) قيمة "ت" ودلالتها الإحصائية للفروق بين متوسطات درجات التطبيق البعدي لمقياس مستوى قلق الاختبار لدى طلاب المجموعة الضابطة والتجريبية الأولى

المجموعات	ن	م	ع	قيمة ت	درجات الحرية	مستوى الدلالة
الضابطة	٣٠	٥٩.٦	١٣.٥	٠.٦	٥٨	0.53
التجريبية الأولى	٣٠	٥٧.٣	١٥.٢			

يتضح من الجدول (٨) ان قيمة "ت" (٠.٦) للفرق بين متوسطات درجات التطبيق البعدي لمقياس قلق الاختبار عند مستوى دلالة (<٠.٠٥) غير دال احصائيا. كما اتضح للباحث ان بعد مراجعة قيمة "ت" الجدولية عند درجات الحرية الموضح بالجدول (٨) والموضح بالرابط http://socr.ucla.edu/Applets.dir/T_Table.html. ان قيمة "ت" الجدولية (١.٧) مما يعنى ان قيمة "ت" الجدولية أكبر من قيمة " ت" المحسوبة ما يؤكد عدم وجود فرق دال احصائيا بين متوسطات درجات المجموعتين الضابطة والتجريبية فى مستوى خفض قلق الاختبار ومن ثم قبول الفرضية الصفرية الفرعية الأولى .

٢- نتائج اختبار صحة الفرضية الصفرية الثانية

يوضح جدول (٨) نتائج التحليل الإحصائي للاختبار "ت" للمجموعات المستقلة لتحديد دلالة الفرق بين متوسطات درجات التطبيق البعدي لمقياس قلق الاختبار على المجموعة الضابطة والتجريبية الثانية.

جدول (٨) قيمة "ت" ودلالتها الإحصائية للفروق بين متوسطات درجات التطبيق البعدي لمقياس مستوى قلق الاختبار لدى طلاب المجموعة الضابطة والتجريبية الثانية

المجموعات	ن	م	ع	ت	درجات الحرية	مستوى الدلالة
الضابطة	٣٠	٥٩.٦	١٣.٥	2.5	٥٨	0.01
التجريبية الثانية	٣٠	٤٩.٦	١٦.٦			

يتضح من جدول (٨) ان قيمة "ت" (٢.٥) ومستوى دلالتها في الفرق بين متوسطات درجات التطبيق البعدي لمقياس مستوى قلق الاختبار على المجموعة الضابطة والتجريبية الثانية ذا دلالة إحصائية عند مستوى (< ٠.٠٥) لصالح المجموعة الضابطة. كما يتضح من مقارنة قيمة "ت" الجدولية بقيمة "ت" المحسوبة ان الأخيرة أكبر من قيمة "ت" الجدولية مما يعنى وجود دلالة للفرق بين متوسطات درجات التطبيق البعدي لمقياس القلق. أي ان مستوى قلق الاختبار فى المجموعة الضابطة التي استخدمت الاختبار الإلكتروني التقليدي اعلى من المجموعة التجريبية الثانية التي استخدمت الاختبار التكيفي متغير الإنهاء . مما يعنى رفض الفرضية الصفرية الفرعية الثانية التابعة للفرض الصفري الأول.

٣- نتائج اختبار صحة الفرضية الفرعية الثالثة.

يوضح جدول (٩) نتائج اختبار "ت" لتحديد دلالة الفرق بين متوسطات التطبيق البعدي على المجموعة التجريبية الأولى والثانية.

جدول (٩) قيمة "ت" ودلالتها الإحصائية للفروق بين متوسطات درجات التطبيق البعدي لمقياس مستوى قلق الاختبار لدى طلاب المجموعة التجريبية الأولى والثانية.

المجموعات	ن	م	ع	قيمة ت	درجات الحرية	مستوى الدلالة
التجريبية الأولى	٣٠	٥٧.٣	١٥.٢	١.٨٦	٥٨	0.05
التجريبية الثانية	٣٠	٤٩.٦	١٦.٦			

يتضح من جدول (٩) ان قيمة "ت" (١.٨٦) وهي قيمة دالة عند مستوى دلالة (=٠.٠٥) وليس اقل من (<٠.٠٥) ولتحديد دلالة الفروق تم المقارنة بين قيمة "ت" الجدولية والتي تساوى (١.٧) أي ان قيمة "ت" المحسوبة (>) من قيمة "ت" الجدولية مما يعنى وجود دلالة للفرق بين المجموعة التجريبية الأولى والثانية فى التطبيق البعدي لمقياس القلق لصالح المجموعة التجريبية

الأولى مما يعنى ارتفاع مستوى قلق الاختبار فى الاختبارات التكيفية ثابته الطول عنه فى متغيرة الطول ومن ثم قبول الفرض الصفري الفرعي الثالث.

مما سبق يتضح ان نتائج اختبار الفرضية الرئيسة الصفرية الأولى والتي تنص" لا توجد فروق ذات دلالة إحصائية عند مستوى اقل من (٠.٠٥) بين متوسطات المجموعات التجريبية والمجموعة الضابطة فى التطبيق البعدي لمقياس مستوى قلق الاختبار". اسفرت عن قبول الفرضية التي تشير الى عدم وجود أثر ذو دلالة إحصائية فى مستوى قلق الطلاب الذين استخدموا الاختبارات الإلكترونية التقليدية عند مقارنها بالطلاب الذين استخدموا الاختبارات الإلكترونية التكيفية ثابتة الطول . بينما تم رفض الفرضية التي تشير الى عدم وجود قلق اختبار لدى الطلاب الذين استخدموا الاختبارات الإلكترونية التقليدية مقارنة بالطلاب الذين استخدموا الاختبارات التكيفية متغيرات الطول حيث أظهرت النتائج ارتفاع مستوى قلق الاختبار لدى الطلاب الذين استخدموا الاختبارات التقليدية مقارنة بالطلاب الذين استخدموا الاختبارات التكيفية متغيرة الطول . كذلك أظهرت نتائج الاختبار الإحصائي ارتفاع مستوى قلق الاختبار لدى المجموعة التي استخدمت الاختبار التكيفي ثابت الطول فى مقابل المجموعة التي استخدمت الاختبار التكيفي متغير الإنهاء .

ثانيا: اختبار صحة الفرضية الموجه الثاني وتتطلب ذلك اختبار صحة الفرضيات الفرعية التابعة لها كالتالي:

ا- اختبار صحة الفرض الفرعي الموجه الأولى

تم حساب قيمة "ت " للمجموعات المستقلة من خلال الفروق بين متوسطات(م) درجات التطبيق البعدي لمقياس الاتجاهات نحو الاختبارات الإلكترونية وقيم الانحرافات المعيارية(ع) لكلا المجموعتين الضابطة والتجريبية الأولى كما هو موضح بجدول (١٠).

جدول (١٠) قيمة "ت" ودلالتها الإحصائية للفروق بين متوسطات درجات التطبيق البعدي لمقياس الاتجاهات نحو الاختيارات الإلكترونية لدى طلاب المجموعة الضابطة والتجريبية الأولى

المجموعات	ن	م	ع	ت	درجات الحرية	مستوى الدلالة
الضابطة	٣٠	١٢٩.٩	٧.٤	٥.٨	٥٨	٠.٠١
التجريبية الأولى	٣٠	١٣٩.٦	٥.٢			

يتضح من جدول (١٠) ان قيمة "ت" (٥.٨) لدلالة الفرق بين متوسطات التطبيق البعدي لمقياس الاتجاهات ذو دلالة إحصائية عند مستوى (< ٠.٠٥) لصالح المجموعة التجريبية الأولى . كما اتضح للباحث بعد مراجعة قيمة "ت" الجدولية عند درجات الحرية الموضح بالجدول

(١١) والموضح بالرابط http://socr.ucla.edu/Applets.dir/T_Table.html. وجد الباحث ان قيمة "ت" الجدولية (١.٧) عند درجات حرية (٥٨) ومستوى دلالة (٠.٠٥) مما يعنى ان قيمة (ت) المحسوبة أكبر من قيمة "ت" الجدولية والذي يشير الى وجود فروق لصالح متوسط الدرجات الأكبر والمتمثل فى درجات المجموعة التجريبية الأولى . مما يعنى قبول الفرض الفرعي الأول من الفرضية الرئيسية الثاني والتي تنص على انه " يوجد فرق ذو دلالة إحصائية عند مستوى (٠.٠٥) بين متوسطات درجات المجموعة التجريبية الأولى والمجموعة الضابطة فى التطبيق البعدي لمقياس الاتجاهات نحو الاختبارات الإلكترونية لصالح المجموعة التجريبية الأولى ".

٢- اختبار صحة الفرضية الفرعية الثانية من الفرض الرئيس الموجه الثاني

يوضح جدول (١١) قيمة "ت" للفروق بين المتوسطات والانحرافات المعيارية لدرجات التطبيق البعدي لمقياس الاتجاهات على المجموعة الضابطة والمجموعة التجريبية الثانية.

جدول (١١) قيمة "ت" ودلالتها الإحصائية للفروق بين متوسطات درجات التطبيق البعدي لمقياس الاتجاهات نحو الاختيارات الإلكترونية لدى طلاب المجموعة الضابطة والتجريبية الثانية

المجموعات	ن	م	ع	ت	درجات الحرية	مستوى الدلالة
الضابطة	٣٠	١٢٩.٩	٧.٤	٧.٤	٥٨	٠.٠٥
التجريبية الثانية	٣٠	١٣١.١	٥.٥			

يتضح من جدول (١١) ان قيمة (٧.٤) مما يعنى ان دلالة الفرق بين متوسطات التطبيق البعدي لمقياس الاتجاهات ذو دلالة إحصائية عند مستوى اقل (< ٠.٠٥) لصالح المجموعة التجريبية الثانية.

كما اتضح للباحث بعد مراجعة قيمة "ت" الجدولية عند درجات الحرية الموضح بالجدول (١١) والموضح بالرابط http://socr.ucla.edu/Applets.dir/T_Table.html. وجد الباحث ان قيمة "ت" الجدولية (١.٧) عند درجات حرية (٥٨) ومستوى دلالة (٠.٠٥) مما يعنى ان قيمة (ت) المحسوبة أكبر من قيمة "ت" الجدولية والذي يشير الى وجود فرق لصالح متوسط درجات التجريبية الثانية. مما يعنى قبول الفرض الفرعي الثاني من الفرضية الرئيسية الثاني والتي تنص على انه " يوجد فرق ذو دلالة إحصائية عند مستوى (٠.٠٥) بين

متوسطات درجات المجموعة التجريبية الأولى والمجموعة التجريبية فى التطبيق البعدي لمقياس الاتجاهات نحو الاختبارات الإلكترونية لصالح المجموعة التجريبية الثانية".

٣- اختبار صحة الفرضية الفرعية الثالثة من الفرض الرئيس الموجه الثاني.

يوضح جدول (١٢) قيمة "ت" (٦.١) للفروق بين متوسطات درجات المجموعة التجريبية الأولى والثانية فى التطبيق البعدي لمقياس الاتجاهات نحو الاختبارات الإلكترونية .

جدول (١٢) قيمة "ت" ودلالتها الإحصائية للفروق بين متوسطات درجات التطبيق البعدي لمقياس الاتجاهات نحو الاختيارات الإلكترونية لدى طلاب المجموعة الأولى والتجريبية الثانية

المجموعات	ن	م	ع	ت	درجات الحرية	مستوى الدلالة
التجريبية الأولى	٣٠	١٣٩.٦	٥.٢	٦.١	٥٨	٠.٠١
التجريبية الثانية	٣٠	١٣١.١	٥.٥			

يتضح من جدول (١٢) ان قيمة (٦.١) ان دلالة الفرق بين متوسطات التطبيق البعدي لمقياس الاتجاهات ذو دلالة إحصائية عند مستوى اقل (< ٠.٠٥) لصالح المجموعة التجريبية الثانية. كما اتضح للباحث بعد مراجعة قيمة "ت" الجدولية عند درجات الحرية بقيمة (٥٨) والموضح بالرابط http://socr.ucla.edu/Applets.dir/T_Table.html. وجد الباحث ان قيمة "ت" الجدولية (١.٧) عند درجات حرية (٥٨) ومستوى دلالة (٠.٠٥) مما يعنى ان قيمة (ت) المحسوبة أكبر من قيمة "ت" الجدولية والذي يشير الى وجود فرق لصالح متوسط الدرجات التجريبية الأولى . مما يعنى رفض الفرض الفرعي الثاني من الفرضية الرئيسية الثاني والتي تنص على انه " يوجد فرق ذو دلالة إحصائية عند مستوى (٠.٠٥) بين متوسطات درجات المجموعة التجريبية الأولى والمجموعة التجريبية فى التطبيق البعدي لمقياس الاتجاهات نحو الاختبارات الإلكترونية لصالح المجموعة التجريبية الثانية".

يتضح مما سبق من نتائج اختبار صحة الفرضية الرئيسية الموجه الثانية اتجاه الفرق بين متوسطات درجات التطبيق البعدي لمقياس الاتجاهات على المجموعة الضابطة والتجريبية الأولى عند مستوى دلالة (٠.٠٥) لصالح متوسط درجات المجموعة التجريبية الأولى مقارنة بالمجموعة الضابطة مما يعنى ان الاختبارات الإلكترونية التكيفية ثابتة الإنهاء اثرة بشكل أكبر

فى اتجاهات الطلاب نحو الاختبارات الإلكترونية مقارنة بالاختبارات الإلكترونية التقليدية. كذلك اثرت الاختبارات الإلكترونية التكيفية متغيرة الإنهاء بشكل أكبر فى اتجاهات الطلاب مقارنة بالاختبارات الإلكترونية التقليدية. كما اظهر نتائج المقارنة بين الاختبارات الإلكترونية التكيفية ثابتة الإنهاء فى مقابل متغيرة الإنهاء ان الاختبارات التكيفية ثابتة الإنهاء الأكثر تأثيرا فى اتجاهات الطلاب.

مناقشة وتفسير النتائج

أولا: مناقشة وتفسير نتائج التحقق من فرضيات أثر الاختلاف بين أنماط التصميم للاختبارات التكيفية على مستوى قلق الاختبار.

- مناقشة وتفسير النتائج ذات الصلة بصحة الفرضية الفرعية الأولى .

يتضح من المعالجة الإحصائية للتأكد من صحة الفرض الصفري الأول فى ضوء التحقق من نتائج التطبيق البعدي لمقياس قلق الاختبار والتي أظهرت عدم وجود فروق ذو دلالة إحصائية بين متوسطات درجات كل من المجموعة الضابطة التي استخدمت الاختبارات الإلكترونية التقليدية والمجموعة التجريبية الأولى التي استخدمت الاختبارات التكيفية فى مستوى القلق مما يعنى انه تصميم الاختبار ثابت الطول من حيث عدد المفردات الكلية المطروحة على المتعلم فى كل من المجموعة الضابطة والتجريبية أدى الى عدم ظهور فرق دال احصائيا بين كلا المجموعتين على الرغم من اختلاف نمط توليد المفردات ومفردات بداية وانهاء الاختبار فى كلا المجموعتين مما يعنى ان التوليد العشوائي للمفردات الاختبارية وعدم مراعاة الاستجابات الفعلية للمتعلم كما فى اختبارات التكيفية لم يحدث فارق فى مستوى من القلق ذو دلالة إحصائية مقارنة بالمجموعة التجريبية التي استخدمت الاختبارات التكيفية ثابتة الطول ويمكن تفسير ذلك فى كون المدى الزمنى الاجمالي للاختبار والذى تم تحديده فى الاختبار التقليدي (٥٠) دقيقه كان مناسب بحيث لم يظهر مستوى مرتفع من القلق، وفى المقابل الاختبار التكيفي الثابت الطول من حيث عدد المفردات والمحدد ب(٤٠) مفرده والمرتبط انهاء الاختبار به بطرح هذا العدد على كل متعلم من افراد المجموعة التجريبية على الرغم من وجود زمن مقرر لكل مفردة فى الاختبار التكيفي ثابت الطول دونما تحديد زمن أجمالي لإنهاء لكونه مختلف من طالب للأخر حسب خصائص المفردات التي أجاب عليها ونوعية استجابته من حيث الصواب والخطأ مما يعنى ان ثبات الطول من حيث عدد المفردات المطروحة على جميع افراد المجموعة فى الاختبار التكيفي ثابت الطول لم يحدث مستوى من القلق مقارنة بالاختبار الإلكتروني التقليدي على الرغم من تغير زمن

انهاء الاختبار فى الاختبار التكيفي من ممتحن للأخر حسب خصائص المفردة المطروحة من حيث مستوى الصعوبة والزمن اللازم للإجابة عليها وتقييم استجابة المتعلم عليها . ويؤكد النتيجة السابقة ما اشار اليه كلا من Sax, Gmatik(2013) على ان ترتيب الأسئلة من حيث الصعوبة او السهولة او بشكل عشوائي على الرغم من ثبات عدد المفردات في كل المتتابعات السابقة يحد من الاثار الناجمة عن القلق في حال زيادة المدى الزمنى المقرر للإجابة على الاختبار ككل او كل مفردة مما يعنى ان متغير زمن الإجابة على الاختبار اجمالا او المفردة الاختبارية تفصيلا يؤثر بشكل مباشر في مستوى القلق على الرغم من ثبات طول الاختبار دونما اعتبار للاستجابات المتعلم .كما تؤكد النتيجة السابقة في ان جودة الاختبارات التكيفية في انتاج صور متكافئة من المفردات من بنوك الأسئلة حققت مستوى منخفض من القلق لدى الممتحن ومن ثم حققت قبولا واسعا من قبل مصممي ومطوري الاختبارات وكذا تحليل الاختبارات محكية ومعيارية المرجع (Fan,1998,363;Önder, 2007,210).

كما يرى الباحث ان التوسع فى توظيف تكنولوجيا الذكاء الاصطناعي والبيانات الضخمة يمكن ان يقدم تفسيرات ذات صلة بمتغير الزمن فى تصميم ثبات الطول للاختبار التكيفي التي تعتمد على الصور المتكافئة فى انتاج مفردات متساوية العدد لكل الممتحنين من خلال تحليلات عميقة للعلاقات بين قدرات المتعلم الفعلية وخصائص المفردة فى ضوء مدى زمنى محدد، بمعنى تقنى تحديد الزمن المناسب للإجابة على المفردة فى تصميمات الاختبارات التكيفية ثابتة الطول من خلال تحليل كافة الاستجابات السابقة للمتعلم على مقياس متغير الزمن لتوفير الزمن المناسب للإجابة بالقدر الذى يقلل من ظهور قلق سببه زمن اتاحه دون دراسة مسبقه لخصائص المتعلم فى سرعة الاستجابة مع المفردات المشابه للمفردة المطروحة حاليا من حيث الخصائص السيكو مترية.

- مناقشة وتفسير النتائج ذات الصلة بصحة الفرضية الفرعية الثانية.

يتضح من المعالجة الإحصائية رفض الفرض الصفري الفرعية الثاني فى ضوء التحقق من نتائج التطبيق البعدي لمقياس قلق الاختبار التي أظهرت وجود فرق دال احصائيا بين متوسطات درجات المجموعة الضابطة التي استخدمت الاختبارات التقليدية والمجموعة التجريبية الثانية التي استخدمت الاختبارات التكيفية متغيرة الطول لصالح المجموعة الضابطة مما يعنى ان مستوى القلق فى المجموعة الضابطة اعلى منه فى التجريبية التي استخدمت الاختبارات التكيفية ذات

التصميم متغير الطول ويرجع ذلك الى كون غياب وجود زمن كلى للاختبار فى المجموعة التجريبية الثانية وفر قدر من الحماسة والدافعية لدى المتعلم للأنهاء الاختبار فى اقل زمن ممكن لكون تتابع المفردات من حيث مستوى الصعوبة يعتمد على الاستجابة الحالية ومن ثم أجمالي عدد المفردات المطروحة من حيث العدد لتحقيق درجه اجمالية final score مرتفعة ويتفق ذلك مع ما ذكره Wauters et al.(2010) فى ان الاختبارات التكيفية أسهمت فى رفع مستوى دافعية لدى المتعلم لكون انهاء الاختبار مرتبطة بمعدلات الأداء المرتفعة وليس زمن محدد . كما يرى الباحث ان ارتباط الاختبار التكيفي بقدرات المتعلم يعظم من ثقة الطالب فى نفسة فى حال احساسه بأن استجابته صحيحه وإمكانية انهاء الاختبار فى اقل وقت وبأقل عدد من المفردات المطروحة الامر الذى يحد من ظهور مستوى مرتفع من القلق الذى يتناسب طرديا مع زيادة عدد المفردات الاجمالية للاختبار.ويؤكد وجهة نظر الباحث ما توصل اليه Sarason, sarason (1990) تجريبا بشأن اختبارات الأداء التكيفية حيث وفر غياب وجود زمن محدد للاختبار وفر قدر إيجابي من التوتر ساعد فى رفع مستوى رغبة الممتحنين تحقيق مستوى أداء مرتفع بأقل عدد من المحاولات من خلال مراعاة قواعد الوصول الى المستوى المطلوب فى اقل وقت ممكن. وكذا ما ذكره Davey And Stone(2011)حول المميزات ذات صلة بتقليص عدد الفقرات المطروحة على المفحوص ومن ثم الوقت المطلوب للإنهاء الاختبار والذي ينعكس بالإيجاب على تقليل تأثير عوامل التعب والملل ونقص الدافعية

- مناقشة وتفسير النتائج ذات الصلة بصحة الفرضية الفرعية الثالثة.

يتضح من المعالجة الإحصائية رفض الفرض الصفري الفرعية الثالث فى ضوء التحقق من نتائج التطبيق البعدي لمقياس قلق الاختبار التي أظهرت وجود فروق دال احصائيا بين متوسطات درجات المجموعة التجريبية الأولى التي استخدمت الاختبارات التقليدية والمجموعة التجريبية الثانية التي استخدمت الاختبارات التكيفية متغيرة الطول لصالح المجموعة التجريبية الأولى والذى يرجع الى ان عامل عدد وجود زمن او عدد محدد من المفردات للأنهاء الاختبار وفر قدرا من التوتر الإيجابي ومن ثم ارتفاع مستوى الدافعية فى تحقيق درجات الى من خلال الوصول الى مفردات وزنها النوعي او قيمة الدرجة المخصصة لها مرتفعة والتي تتيح انهاء الاختبار فى زمن اقل منع فى الاختبار التكيفية التي تعتمد على طور متكافئة وعدد ثابت من المفردات يتم طرحها

على جميع الممتحنين مع اختلاف تتابعها ومستويات صعوباتها من ممتحن للأخر ويتفق ذلك مع اراء عديد من البحثين فى كون الاختبارات التكيفية متغيرة الطول اكثر مرونة فى تطبيق قاعدة القفز الى مفردة اكثر صعوبة او مفردة فى موضوع تعلم اخر بناء على اليات استدلال في تقدير قدرات المتعلم قائمه على تكنولوجيا الذكاء الاصطناعي من خلال تحديد مستوى القدرة الفعلي للمتعلم ثم المقارنة بين مستوى صعوبة المفردات التي إجابة عليها المتعلم إجابة صحيحه ومستوى الصعوبة التي يجب ان يصل اليه المتعلم كمحك في انهاء الاختبار او طرح عدد اخر من المفردات لها نفس مستوى الصعوبة في حال ثبات استجابة المفحوص من حيث الصح/ الخطأ كأساس في انهاء الاختبار في مستوى الأداء الفعلي للمتعلم (Leung & Li 2007;Baylari & Montazer 2009; Chen & Chen 2009; Veldkamp & Matteucci, 2013).

ثانيا: مناقشة وتفسير نتائج التحقق من فرضيات أثر الاختلاف بين أنماط التصميم للاختبارات التكيفية على اتجاهات الطلاب نحو الاختبارات الإلكترونية .

- مناقشة وتفسير النتائج ذات الصلة بصحة الفرضية الفرعية الأولى .

يتضح من نتائج التحقق من الفرضية الأولى التابعة للفرض الرئيس الموجه الثاني وجود فروق ذو دلالة إحصائية بين متوسطات درجات التطبيق البعدي لمقياس الاتجاهات نحو الاختبارات الإلكترونية على المجموعة الضابطة التي تعرضت لاختبار إلكتروني تقليدي والمجموعة التجريبية الثانية التي تعرضت للاختبار تكيفي ثابت الطول لصالح المجموعة التجريبية الأولى والذى يرجع الى توفير تلك النوعية من الاختبارات درجة دقة وموضوعية انعكست فى ثقة الطلاب فى الاختبارات ومن تكون اتجاهات إيجابية لديهم نحو التوسع فى استخدامها ودمجها فى منظومة التقويم الإلكتروني ويتفق ذلك مع ما اظهرته نتائج عديد من البحوث ذات الصلة بأثر متغيرات تصميم الاختبارات الإلكترونية بصفة عامة على اتجاهات طلاب كلية التربية نحو الاختبارات الإلكترونية مثل دراسة كل من دراسة أسامة الهنداوي، نهارى الغبيشى (٢٠١٢)، سليمان احمد حرب (٢٠١٨)؛ مفلح ال جديع(٢٠١٧).كما يرجع تفوق الاختبارات التكيفية ثابتة الطول في تكوين اتجاهات إيجابية لدى الطلاب مقارنة بالاختبارات الإلكترونية التقليدية في توفير تلك النوعية من الاختبارات تقدير دقيق لمستوى الأداء الفعلي لدى المتعلم في متتابعة عرض المفردات بشكل يساعد في تكوين

شعور لدى المتعلم ان الاختبار مفصل لقياس قدراته وفقا ما ذكر كل من احمد عودة (٢٠١٣)؛ Wauters et ؛ Wang(2006) al.(2010). كما تؤكد النتائج دور الإطار النظري الذي ترتكز عليه الاختبارات التكيفية والمتمثل في نظرية الاستجابة للمفردة التي تعتمد على نماذج رياضية تحاكى احتمالات إجابة الفرد على كل مفردة فى كل مستوى من مستويات المتغير المراد قياسه من خلال الاختبار مما يعنى ان تقدير استجابات الافراد يتم وفقا تحليل أدائهم الفعلي على كل مفردة (Crocker&Algina,1986, p.339)٠

- مناقشة وتفسير النتائج ذات الصلة بصحة الفرضية الفرعية الثاني.

يتضح من دلالة الفروق بين متوسطات درجات التطبيق البعدي لمقياس الاتجاهات نحو الاختبارات الإلكترونية على المجموعة الضابطة التي تعرضت للاختبارات التقليدية والمجموعة التجريبية الثانية التي استخدمت الاختبارات التكيفية متغيرة تفوق الأخيرة على المجموعة الضابطة في تكوين اتجاهات لدى الطلاب نحو الاختبارات الإلكترونية ويرجع ذلك الى تحقيق عنصر غياب زمن وعدد محدد من المفردات كما فى الاختبارات التقليدية قدر من الحماس والدافعية. ويتفق ذلك مع ما أشار اليه Davey and Stone(٢٠١١) في ان مميزات تقليص عدد الفقرات المطروحة على المفحوص للوصول الى تقدير دقيق لمستوى اداءه ومن ثم الوقت المطلوب للإنهاء الاختبار ينعكس بالإيجاب على اتجاهات الطلاب نحو تلك النوعية من الاختبارات.

- مناقشة وتفسير النتائج ذات الصلة بصحة الفرضية الفرعية الثالث.

يتضح من نتائج التحقق من الفرضية الثالث التابعة للفرض الرئيس الموجه الثاني وجود فرق ذو دلالة إحصائية بين متوسطات درجات التطبيق البعدي لمقياس الاتجاهات نحو الاختبارات الإلكترونية على المجموعة التجريبية الأولى التي تعرضت لنمط تصميم اختبار تكيفي ثابت الطول والمجموعة التجريبية الثانية التي تعرضت لنمط الاختبار التكيفي متغير الطول لصالح المجموعة التجريبية الأولى والذى يرجع الى أن ثبات الطول فى الاختبارات التكيفية من حيث عدد المفردات المطروحة على الممتحنين يساعد فى إدارة المتعلم لوقت الاختبار بشكل اكثر دقة وعمل تفضيلات الازاحة او الاستجابة لمفردات محدده ، كذلك يرجع السبب فى اتجاهات الطلاب نحو الاختبار التكيفي ثابت الطول لكونه اكثر وضوحا للمتعلم من حيث معلومية العدد الكلى للاختبار .

التوصيات :

على ضوء نتائج البحث الحالي يوصى الباحث بما يلي:

- التوسع فى استخدام منصات وتطبيقات وبرمجيات الاختبارات التكيفية فى المقررات الدراسة المختلفة بكليات التربية.
- استخدام الاختبارات التكيفية فى اختبارات القبول والتسكين وتجديد رخص مزاولة المهنة فى برامج كليات مرحلة التعليم الجامعي.
- انشاء مراكز متخصصة فى تطوير اختبارات تكيفية لقياس المهارات العملية فى التخصصات العلمية المختلفة.
- تدريب أعضاء هيئة التدريس فى مهارات استخدام وتطوير اختبارات تكيفية فى مجال تخصصهم.
- دعم مشروعات تطوير اختبارات تكيفية قائمه على خوارزميات تعلم الالة.

البحوث المقترحة

يقدم الباحث مجموعة من المقترحات تمثلت فى الاتي :

- دراسة أثر اختلاف نمط تصميم تتابع المفردات فى الاختبارات التكيفية على الأداء الفوري والمؤجل لدى طلاب كليات التربية.
- دراسة أثر نمط تصميم التغذية المرتدة فى الاختبارات التكيفية التشخيصية على مستوى تقدير الذات والقدرة على حل المشكلات لدى طلاب المرحلة الثانوية.
- دراسة أثر التفاعل بين الأسلوب المعرفي (مندفع – متروى) ونمط تصميم بداية الاختبار التكيفي على مستوى التحصيل والدافعية للتعلم.

Title: Impact of fixed and variable length computer adaptive test designs in reducing test anxiety and developing attitudes towards online exams: Faculty of education students' case

Abstract: The present study investigates the impact of fixed and variable computer adaptive test designs in reducing test anxiety and developing attitudes towards online exams for faculty of education students. The experimental design has been set as extended control group pre-posttest design. Research groups included: 1) The control group that was subjected to a traditional computer test of (40) items. These have been randomly withdrawn from a question bank of 400 items created on a learning management system 2) The first experimental group had an adaptive test of fixed-length design through one of the specialized adaptive testing platforms. The beginning of the test was based on providing a medium-difficulty item then withdrawing the following items from the question bank based on the learner's responses until pre-determined items come to end; according to the rule of test-ending; 3) The second experimental group was subjected to an adaptive test of variable length design through one of the specialized platforms in adaptive tests. It was based on a pre-test for determining the initial level of the learner with (12) items of varying levels of difficulty and learning content. The following items have been provided, in terms of difficulty level, according to a dynamic estimation of the learner's responses. The test is ended according to the maximum likelihood estimation method in estimating the learner's performance without regard to the number of items withdrawn from the bank for each examinee. The test is completed at a standard error level of (0.25). Research tools consisted of two measures of the level of test anxiety and trends towards the electronic test prepared by the researcher. The results showed that there was no

statistically significant difference between the control group, which used linear designs computer tests, and the first experimental group, which used fixed-length adaptive tests, for reducing the level of test anxiety. Comparison between the control group and the second experimental group that used variable-length adaptive tests has shown a difference in favor of the second experimental group in reducing the level of anxiety. Comparison between the first experimental group and the second experimental group has shown the latter's superiority in reducing the test anxiety level. The results of the post-application of the scale of trends on the three groups have also shown that fixed-length adaptive tests have a greater impact in forming positive trends among students compared to traditional computer tests. Also, variable-length adaptive tests have excelled written computer tests in forming positive trends among students towards online exams tests. Meanwhile, fixed-length adaptive tests have outperformed variable-length adaptive tests in forming positive trends for students toward online exams.

Keyword: adaptive test length design patterns; test anxiety; attitudes toward online exam; higher education students.

المراجع

احمد سعيد العطار (٢٠١٧). فاعلية نظام إلكتروني تكيفي قائم على أسلوب التعلم والتفضيلات التعليمية على تنمية مهارات البرمجة لدى طلاب تكنولوجيا التعليم، مجلة البحث العلمي فى التربية،٦(١٨)، ص ٣٤٩-٤٠٨.

احمد عودة، عمر عبيدات (٢٠١٣). فاعلية الاختبار التكيفي المحوسب فى تقدير القدرة باستخدام مصفوفة رافن. دراسات – العلوم التربوية، مج٤٠، ع٢،ص ص ١٦٠٢-١٦٢ .

أسامة الهنداوي (مشرف)، نهاري الغبشى (باحث) (٢٠١٢). أثر بعض متغيرات تصميم الاختبارات الإلكترونية على أداء طلاب الصف الثالث الثانوي واتجاهاتهم نحوها. رسالة ماجستير، كلية التربية، جامعة طيبة، السعودية.

أسامة سعيد هنداوي (٢٠١٠). أثر ثلاث تصميمات لأنماط الاستجابات على الاختبارات الإلكترونية على معدل الأداء الفوري والمؤجل لطلاب الجامعة فى الاختبار. مجلة العلوم التربوية، كلية الدراسات العليا، جامعة القاهرة،١٨(٣)، ص ١٠٢-١٤٢.

اميرة عطا (٢٠١٤). التكيف فى بيئات التعلم: تحدى جديد من اجل الأفضل. مجلة التعليم الإلكتروني – متاح على الرابط

http://emag.mans.edu.eg/index.php?page=news&task=show&id=94&sessionID=17

ايناس محمد الحسيني مندور (٢٠١٣). أثر برنامج تدريبي لطلاب الدراسات العليا بكلية التربية فى تصميم الاختبارات الإلكترونية وفقا لمعايير الجودة المقترحة. دراسات تربوية اجتماعية، مصر، ١٩(٢)،٣٩١-٤٦٠.

تامر المغاورى الملاح (٢٠١٧). التعلم الكيفي. بيئات التعلم التكيفية. القاهرة: دار السحاب للنشر والتوزيع.

رباب عبد المقصود البلاصى (٢٠١٦). اتجاهات أعضاء هيئة التدريس بجامعة حائل نحو استخدام نظم إدارة التعلم. دراسات عربية فى التربية وعلم النفس، ع٦٩.

رشا حمدى هدية (٢٠١٩). تصميم بيئة تعلم تكيفية وفقا للذكاءات المتعددة وأثرها فى تنمية مهارات أنتاج الاختبارات الإلكتروني لدى طلاب كلية التربية. الجمعية العربية لتكنولوجيا التربية، ع٣٨، ص ص ٤٧٣-٥٤٠.

رضا سمير عوض (٢٠٠٨). بعض المتغيرات المؤثرات على دقة كشف الدالة المميزة للمفردة في ضوء نظرية الاستجابة للمفردة. رسالة دكتوراه بمعهد الدراسات التربوية بجامعة القاهرة.

زاهر إسماعيل الغريب (٢٠٠٩). المقررات الإلكترونية – تصميمها- انتاجها- إنتاجها- تطبيقها- تقويمها. القاهرة: عالم الكتب.

سليمان احمد حرب (٢٠١٨). المعوقات والاتجاهات نحو استخدام الاختبارات الإلكترونية ووضع تصور مقترح لتطبيق الاختبارات الإلكترونية بجامعة الأقصى بعزة. مجلة جامعة الأقصى للعلوم التربوية والنفسية، مج١(ع١)، ص ١٥٥-١٩٧.

صلاح علام (٢٠٠٥). نماذج الاستجابة للمفردة الاختبارية الأحادية ومتعددة الابعاد وتطبيقاتها فى القياس النفسي والتربوي، القاهرة: دار الفكر العربي.

محمد إبراهيم عطا الله (٢٠١٦). اتجاهات أعضاء هيئة التدريس بجامعة المنصورة نحو التقويم الإلكتروني ومعوقات تطبيقه، مجلة كلية التربية دراسات تربوية ونفسية، جامعة الزقازيق، ع٩٠.

محمد عبد الوهاب (٢٠١٠). استخدام نماذج تدرج المفردة الاختبارية بعض الاختبارات المعرفية، رسالة دكتوراه، كلية التربية جامعة المنيا، متاحه على الرابط https://www.researchgate.net/publication/277011471.

محمد عطية خميس (٢٠٠٧). الكمبيوتر التعليمي وتكنولوجيا الوسائط المتعددة. القاهرة: دار السحاب للنشر والتوزيع.

محمد عطية خميس (٢٠١٨). بيئات التعلم الإلكتروني (الجزء الأول). ط١. القاهرة: دار السحاب للنشر والتوزيع.

مروة محمد الباز (٢٠١٣). فاعلية برنامج تدريبي قائم على الويب٢.٠فى تنمية مهارات التدريس الإلكتروني والاتجاه نحو لدى معلمي العلوم اثناء الخدمة. مجلة التربية العلمية، ١٦(٢)، ١١٣-١٦٠.

مفلح ال جديع (٢٠١٧). اتجاهات أعضاء هيئة التدريس نحو اجراء الاختبارات الإلكترونية ومعوقات تطبيقها فى جامعة تبوك. المجلة الدولية التربوية المتخصصة، مج (٦)، ع٢، ص ص ٧٧-٨٧.

نبيل جاد عزمي (٢٠١٤). تكنولوجيا التعليم الإلكتروني. القاهرة: دار الفكر العربي.

نضال بهجت الطعان (٢٠١١). اتجاهات أعضاء هيئة التدريس بجامعة البلقاء التطبيقية نحو استخدام التعليم الإلكتروني فى التدريس الجامعي، مجلة كلية التربية بجامعة الجوف.

يوسف محمد السوالمة(مشرف)، فؤاد محمد مسلم(باحث) (٢٠١١). اسلوب مقترح للاختيار الفقرات وتقدير القدرة فى الاختبارات التكيفية. رسالة دكتوراه، كلية التربية، جامعة اليرموك، الأردن.

Administration. Dissertation Presented to The Faculty of The Graduate

Baker, F. (2001): The Basics of Item Response Theory (2nd ed.). The ERIC Clearinghouse on Assessment and Evaluation, USA.

Baylari A. & Montazer G.A. (2009) Design a personalized e-learning system based on item response theory and artificial neural network approach. Expert Systems with Applications 36, 8013–8021.

Baylari, A., & Montazer, G. A. (2009). Design a personalized e-learning system based on item response theory and artificial neural network approach. *Expert Systems with Applications*, *36*(4), 8013-8021.

Boyd, A. M. (2003). Strategies for controlling testlet exposure rates in computerized adaptive testing systems.

Brusilovsky P. & Peylo C. (2003) Adaptive and intelligent Web-based educational systems. International Journal of Artificial Intelligence in Education 13, 156–169.

Brusilovsky P. (1999) Adaptive and intelligent technologies for Web-based education. Künstliche Intelligenz 4,25-29.

Center for the study of Evaluation Graduate School of Education, University of California, Los Angeles, USA.

Chang, J. S., & Cooper, G. (1970). A practical difference scheme for Fokker-Planck equations. Journal of Computational Physics, 6(1), 1-16.

Chen C.M.&Chen M.C. (2009) Mobile formative assessment tool based on data mining techniques for supporting Web based learning. Computers & Education 52, 256–273.

Ciloglugil, B., & Inceoglu, M. M. (2012, June). User modeling for adaptive e-learning systems. In *International Conference on Computational Science and Its Applications* (pp. 550-561). Springer, Berlin, Heidelberg.

Cisar, D., Radosav, D., Markoski, B., Pinter, R. and Cisar, P. (2010). Computer adaptive testing of Student knowledge. Acta Polytechnic Hungarica, 7(4), 139-152.

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.

Crocker, L&Algina, J (1986): Introduction to Classical and Modern Test Theory. New York: Holt, Rinehart and Winston, USA.

Cruces, A. L. L., & De Arriaga, F. (2000). Reactive agent design for intelligent tutoring systems. Cybernetics & Systems, 31(1), 1-47.

DordiNejad, F. G., Hakimi, H., Ashouri, M., Dehghani, M., Zeinali, Z., Daghighi, M. S., & Bahrami, N. (2011). On the relationship between test anxiety and academic performance. Procedia-Social and Behavioral Sciences, 15, 3774-3778.

Driscoll, R. (2007). Westside Test Anxiety Scale Validation. Online Submission.

Eignor, D. R., Stocking, M. L., Way, W. D., & Steffen, M. (1993). Case Studies in Computer Adaptive Test Design Through Simulation 1, 2. Ets

environment. Lecture Notes in Computer Science, 12–20.

Fernandez, G. (2003). Cognitive scaffolding for a web-based adaptive learning

Hambleton. (1982): Item Response Theory: The Three-Parameter Logistic Model.

Hembree, R. (1988). Correlates, causes, effects, and treatment of test anxiety. Reviews of educational research, 58, 47-77.

Huo, Y. (2009). Variable-length Computerized Adaptive Testing: adaptation of the a-stratified strategy in item selection with content balancing. Dissertation of Doctor of Philosophy in psychology in the graduate college of the University of Illinois at Urbana-Champaign.

Janda, L. (1998). Psychological Testing: Theory and Applications. Allyn and Bacon, USA.Kacmar, K.M, Farmer, W.L, Zivnuska, S.&Witt, L.A. (2006): Applying Multidimensional Item Response Theory Analysis to a Measure of Meta-Perspective.

Kacmar, K.M, Farmer, W. L, Zivnuska, S.&Wattala. (2006). Applying multidimensional Item Response Theory Analysis to a Measure of Meta-Perspective.

Lazarinis, F., & Retalis, S. (2007). Analyze me: open learner model in an adaptive web testing system. International Journal of Artificial Intelligence in Education, 17(3), 255-271.

Leung E.W.C. & Li Q. (2007) An experimental study of a personalized learning environment through open-source software tools. IEEE Transactions on Education 50, 331–337.

Magis, D. and Raiche, G. (2011). CatR: An R package for computerized adaptive testing, Applied Psychological Measurement, 33(7): 576-577.

McGlohen, M. (2004). The Application of a Cognitive Diagnosis Model Via an Analysis of a Large-Scale Assessment and A Computerized Adaptive Testing

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

Olea, J., & Ponsoda, V. (1996). Tests adaptativos informatizados. In J. Muñiz (Ed.) Psicometría (pp. 731-783). Madrid: Universitas

Önder, I. (2007). An Investigation of Goodness of Model Data Fit Model Veri Uyumunun Araştirilmasi. Hacettepe Üniversitesi Eğitim Fakültesi Dergisi, (32),210.

Oppl, S., Reisinger, F., Eckmaier, A., & Helm, C. (2017). A flexible online platform for computerized adaptive testing. International Journal of Educational Technology in Higher Education, 14(1), 2.

Paramythis A. & Loidl-Reisinger S. (2004) Adaptive learning environment and e-learning standards. Electronic Journal of eLearning 2, 181–194.

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

Sanz-Santamaría, S., Zorita, J.&Serrano. (2006). Mixing Standards, IRT and Pedagogy for Quality e-Assessment. Current Developments in Technology-Assisted Education, University of the Basque Country (UPV-EHU), Pº Manuel de Lardizabal 1, 20018 San Sebastián, Spain.

Sarason, I. G., & Sarason, B. R. (1990). Test anxiety. In Handbook of social and evaluation anxiety (pp. 475-495). Springer, Boston, MA.

School of The University of Texas At Austin In Partial Fulfillment Of

Stone, E. and Davey, T. (2011). Computer-Adaptive Testing for Students with Disabilities: A Review of the Literature. ETS, Princeton, New Jersey, Retrieved at online: http:// www.ets. org/ research/ contact. html.

Suinn, R. M. (1969). The STABS, a measure of test anxiety for behavior therapy: Normative data. Behaviors Research and Therapy, 7(3), 335-339.

the Requirements for The Degree of Doctor of Philosophy.

Tonidandel, S., Quiñones, M. A., & Adams, A. A. (2002). Computer-adaptive testing: The impact of test characteristics on perceived performance and test takers' reactions. Journal of Applied Psychology, 87(2), 320.

Vale, C. D. (2006). Computerized item banking. In S. M. Downing & T. M. Haladyna (Eds.), Handbook of test development. Mahwah NJ: Erlbaum.

Veerkamp, W. J. J., & Berger, M. P. F. (1997). Some new item selection criteria for adaptive testing, Journal of Educational and Behavioral Statistics, 22(2), 203–226.

Veldkamp, B. P., & Matteucci, M. (2013). Bayesian computerized adaptive testing. Ensaio: Avaliação e Políticas Públicas emEducação, 21(78), 57–82. http://doi.org/10.1590/S0104-40362013005000001.

Wainer, H., & Lewis, C. (1990). Toward a psychometrics for testlets. Journal of Educational Measurement, 27(1), 1-14.

Wainer, H., Dorans, N. J., Flaugher, R., Green, B. F., & Mislevy, R. J. (2000). Computerized adaptive testing: A primer. Routledge.

Wang, F.-H. (2006). Application of componential IRT model for diagnostic test in a standard conformant eLearning system. In Sixth international conference on advanced learning technologies (ICALT'06).

Wang, T., & Ban, J. C. (2001). Effects of Scale Transformation and Test- Termination Rule on the Precision of Ability Estimation in Computerized Adaptive Testing. Journal of Educational Measurement, 38(3), 267-292.

Wauters, K., Desmet, P., & Van Den Noortgate, W. (2010). Adaptive item- based learning environments based on the item response theory: possibilities and challenges. Journal of Computer Assisted Learning, 26(6), 549-562.

Weiss, D. J. (2013). Item banking, test development, and test delivery, in Press, 2011 The APA Handbook on Testing and Assessment Kurt F. Geisinger, Editor-in-Chief Washington DC: American Psychological Association.

Weiss, H. M. (1990). Learning theory and industrial and organizational psychology.

Wiberg, M (2004). Classical Test Theory vs. Item Response Theory: An Evaluation of The Theory Test in The Swedish Driving-license Test. University UMEA. EM No 50.