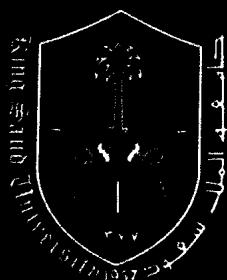


ردمك: ١٠١٨-٣٦٢٠



## العلوم التربوية والدراسات الإسلامية (٤)

أكتوبر (١٢٠٢م)  
ذو القعدة (١٤٣٣هـ)

## فاعلية القياس التكيفي باستخدام فقرات ذات إجابة مُنتقاة وفقرات ذات إجابة مُنشأة

إسماعيل سلامة البرصان

أستاذ مساعد، بقسم علم النفس، كلية التربية، جامعة الملك سعود

الرياض، المملكة العربية السعودية، ص.ب ٢٤٥٨، الرمز ١١٤٥١

E - mail: ibursan@ksu.edu.sa

(قدم للنشر في ١٧/١٠/١٤٣١هـ؛ وقبل للنشر في ١٤٣٢/٦/١هـ)

الكلمات المفتاحية: نظرية الاستجابة للفقرة، القياس التكيفي، النظرية الحديثة في الاختبارات، نماذج استجابة الفقرة. ملخص البحث. هدفت هذه الدراسة إلى الكشف عن فاعلية القياس التكيفي المبني من فقرات ذات إجابة منتقاة ثنائية التدرج وفقرات ذات إجابة منشأة متعددة التدرج في نفس الاختبار باستخدام بيانات مولدة من الحاسوب، حيث تم توليد إجابات (٦٠) فقرة ثنائية التدرج و(١٠) فقرات متعددة التدرج لألفي مفحوص، جرى تدريجها معاً باستخدام برمجية (RUMM 2020)، واختيرت عينة للقياس التكيفي بلغت (١٠٠) مفحوص جرى استخدام بيانات استجاباتهم في القياس التكيفي الذي تكون من ثلاث مراحل تكونت المرحلة الاستطلاعية الأولى من خمس فقرات ذات إجابة منتقاة ثنائية التدرج، والمرحلة الثانية من خمس فقرات ذات إجابة منتقاة ثنائية التدرج، في حين تكونت المرحلة الثالثة والأخيرة من فقرتين ذاتي إجابة منشأة، تم حساب القدرة لكل مفحوص في المراحل الثلاث، وتم مقارنتها بقدرة المفحوص المقدرة من الاختبار الخطي بفقراته السبعين، وأظهرت النتائج معامل ارتباط بين قدرة المقدرة من الاختبار التكيفي والقدرة المقدرة من الاختبار الخطي مقداره ٠.٩٣، ومتوسط فرق في القدرة مقداره ٠.٠٨٧ لوجيت، إضافة إلى فرق في الخطأ المعياري مقداره في المتوسط ٠.٢٧٨ الأمر الذي يفيد بفاعلية عالية للقياس التكيفي المبني باستخدام فقرات ثنائية ومتعددة التدرج معاً خصوصاً عندما تتطلب النتائج كلا النوعين من الفقرات.

### المقدمة والخلفية النظرية

يعد القياس التكيفي (adaptive testing) أحد أهم الأساليب الاختبارية المعتمدة في مجال قياس القدرات والسمات النفسية والتربوية، وأداته هي الاختبارات

التكيفية التي يطلق عليها أحياناً الاختبارات المفصلة (tailored tests)، ولم يصبح القياس التكيفي ممكناً للتطبيق إلا بعد ظهور نظرية الاستجابة للفقرة (item response theory) (Lord, 1980) على الرغم

الفقرة (item characteristic Curve) يكون واحداً لجميع أفراد المجموعة عند مستوى قدرة معين، بينما يؤدي انتهاك هذا الافتراض إلى أن يكون هناك أكثر من خط انحدار يختلف باختلاف أفراد المجموعة الجزئية عند كل مستوى قدرة (Hambleton & Swaminathan, 1985) ويتم التأكد من هذه الخاصية بأكثر من طريقة، من أشهرها التحليل العاملي لاستجابات المفحوصين على فقرات الاختبار.

## ٢- الاستقلال الموضعي (Local Independence)

ويعني هذا الافتراض أن إجابة المفحوص على فقرة ما لا يؤثر إيجاباً أو سلباً على إجابته على أي فقرة أخرى (Hambleton & Swaminathan, 1985) بينما يذكر وورم (warm, 1978) بأن هذا الافتراض يعني عدم ارتباط الفقرات مع بعضها البعض عند قدرة محددة ( $\theta$ )، أي أن الأفراد الذين يمتلكون نفس القدرة يجب أن يكون معامل ارتباط أدائهم على فقرة وأدائهم على فقرة أخرى مساوياً للصفر، ويمكن القول بأن هذا الافتراض يتضمن الاستقلال الإحصائي لاستجابة المفحوص على فقرات الاختبار كافة (Hambleton & Swaminathan, 1985) ويتم التأكد من هذا الافتراض باستخدام قوانين الاحتمالات الرياضية واختبار  $\chi^2$  (Lord, 1980)، بينما يعتبر هاميلتون وسواميناثان أن هذا الافتراض يعتبر متحققاً ضمناً إذا ما تحقق الافتراض الأول (Hambleton & Swaminathan).

من أن الإرهافات الأولى له كفكرة قد بدأت عام 1905 في اختبارات بينية للذكاء.

ويعرف كل من مايلز وستوكنج (Mills & Stocking, 1991) الاختبار التكيفي بأنه عملية تقييم يتم بواسطتها بناء اختبار أثناء قيام المفحوص بالإجابة عن الفقرة حيث يتم اختيار الفقرات اللاحقة من بنك الأسئلة اعتماداً على إجابات المفحوص على الفقرات السابقة.

ويعرفه هاميلتون وسواميناثان (Hambleton & Swaminathan, 1985) بأنه اختبار يوائم بين قدرة المفحوص وصعوبة الفقرات المقدمة له حيث لا يتقدم المفحوصون لنفس الفقرات بل يجري تعريض كل مفحوص لفقرات تناسب قدرته، الأمر الذي يجعل الحصول على درجات متسقة لعدة مفحوصين أمراً متعذراً في حال اعتماد النظرية التقليدية في الاختبارات أما في نظرية الاستجابة للفقرة فإن ذلك ممكن بسبب خاصية تحرر قدرات المفحوصين من خصائص الفقرات التي يتعرض لها المفحوص (item free)، وكذلك بسبب تحرر معالم الفقرات من خصائص المفحوص (Person free).

وحتى يتم استخدام تطبيق نظرية الاستجابة للفقرة لا بد من تحقق افتراضاتها وهي:

## ١- أحادية البعد (Unidimensionality)

ويعني هذا الافتراض أن فقرات الاختبار تقيس سمة واحدة بحيث إن خط الانحدار في منحني خصائص

### ٣- منحني خصائص الفقرة (Item Characteristic Curve)

وهو منحني يمثل العلاقة بين القدرة واحتمالية الإجابة الصحيحة على الفقرة، وهو منحني متصاعد باضطراد وله مستوى تقاربي علوي (upper asymptotic - level) يقترب من ١، ومستوى تقاربي أدنى (Lower asymptotic level) يختلف باختلاف النموذج المستخدم في نظرية الاستجابة للفقرة.

### ٤- التحرر من السرعة (Speediness)

ويعني هذا الافتراض أن عامل السرعة في الأداء لا يلعب دوراً في الإجابة على الفقرات، أي أن عدم إجابة الفرد بشكل صحيح عن فقرات الاختبار يعود إلى تدني قدرته وليس إلى تأثير السرعة في الإجابة أو عدم وصوله لهذه الفقرات نتيجة قصر الوقت (Hambleton, Swaminathan & Rogers, 1991).

وللاختبارات التكيفية باختلاف أنواعها خطوات تحدث عنها المختصون في هذا النوع من القياس، فقد أورد كنجزيري وزارا (Kingsbury & Zara, 1989) ست خطوات لبناء الاختبار التكيفي هو:

١- بناء تجمع الفقرات (item pool): حيث تعد الفقرات ذات المعالم المقدرة جيداً أمراً ضرورياً للاختبار التكيفي، وتري إمبرستون ورايس (Emberston & Raise, 2000) أنه حتى يتحقق القياس الجيد لا بد من أن يحتوي تجمع الفقرات على عدد كاف من الفقرات ذات صعوبة موزعة جيداً على متصل السمة.

### ٢- اعتماد نموذج لاستجابة الفقرة: وينبغي

هنا اختيار النموذج في سياق نظرية استجابة الفقرة التي تحتوى النموذج الأحادي المعلمة (نموذج راش) والنموذج الثنائي المعلمات، والنموذج الثلاثي المعلمات ومعادلتها الرياضية كالاتي:

أ - النموذج الأحادي المعلمة:

$$pi(\theta) = \frac{1}{1 + e^{D(\theta - bi)}}$$

Pi (Θ): احتمال أن يجيب المحفوض ذو القدرة

(Θ) على الفقرة (i) إجابة صحيحة.

D: ثابت قيمته ١.٧.

bi: معامل الصعوبة للفقرة (i).

e: الأساس اللوغاريتمي الطبيعي (العدد

النبيري ٢.٧١٨٣).

ب - النموذج ثنائي المعلمة:

$$pi(\theta) = \frac{1}{1 + e^{Da(\theta - bi)}}$$

حيث a: معلم التميز للفقرة i، وباقي الرموز

كما في معادلة النموذج الأحادي المعلمة.

ج - النموذج ثلاثي المعلمة:

$$pi(\theta) = Ci + (1 - Ci) \frac{e^{Da(\theta - bi)}}{1 + e^{Da(\theta - bi)}}$$

حيث Ci: معلم التخمين بالنسبة للفقرة i.

(Hambleton & Swaminathan, 1985).

د - يضاف إلى ذلك نموذج التقدير الجزئي الذي توصل إلى معادلته ماسترز 1982 Masters (والعلاقة التالية التي تعطي احتمالية وقوع المفحوص n ذي القدرة  $\theta$  في المستوى X بدلاً من البقاء في المستوى  $(x-1)$ .

$$\Phi_{xni} = \frac{\Pi_{xin}}{\Pi_{(x-1)ni} + \Pi_{xin}} = \frac{e^{(\theta_n - biX)}}{1 + e^{(\theta_n - biX)}}$$

حيث:

$\Phi_{xni}$ : احتمال إنجاز المفحوص (n) الخطوة x في السؤال i ليأخذ العلامة x بدلاً من العلامة  $(x-1)$ .

$\Pi_{xin}$ : احتمال وقوع المفحوص (n) ذي القدرة  $\theta$  في المستوى x من مستويات الأداء للفقرة i

$\Pi_{(x-1)ni}$ : احتمال وقوع نفس الفرد في المستوى السابق  $(x-1)$  في الفقرة نفسها i

وهنا تشير المعادلة إلى أن احتمال أن يصل المفحوص (n) ذو القدرة  $\theta$  إلى المستوى (x) يعتمد على صعوبة المستوى  $(b_{ix})$  وقدرة المفحوص  $\theta$  فقط.

٣ - اعتماد الطريقة المدخلية للاختبار، وهي الطريقة التي يتم فيها تحديد الفقرات التي تعطى للمفحوص في المرة الأولى، حيث يمكن البدء بفقرات ذات مستويات مختلفة في الصعوبة تغطي المناطق المختلفة للمحتوى وتراعي التوازن في التميز وهو ما يسمى بالاختبار الاستطلاعي (Routing Test)، ويمكن أن توكل اختيار الفقرات الأولى للمفحوص عن طريق تحديد مستوى قدرته يختاره المفحوص وهو ما يسمى بـ (Self adaptive testing).

٤ - اعتماد طريقة لاختيار الفقرات (Lord, 1980) بناء على القدرة المتحصلة للمفحوص من الإجابة عن الفقرات السابقة، ويمكن أن تعتمد الفقرات اللاحقة بناءً على معلم الصعوبة أو دالة المعلومات.

٥ - اعتماد طريقة رصد العلامة، وهي إما طريقة الأرجحية العظمى (Maximum Likely hood) أو إحدى الطرق البايزية (Bayesian Methods). اعتماد قاعدة للتوقف عن إعطاء الفقرات للمفحوص.

وتختلف الاختبارات التكييفية باختلاف الاستراتيجيات التي تتبعها، وقد قسم لورد (Lord, 1980) استراتيجيات القياس التكييفي إلى ثلاث استراتيجيات هي:

أولاً: استراتيجية القياس التكييفي ثنائي المرحلة (Two Stage Strategy):

وتتكون هذه الاستراتيجية من مرحلتين أولاًهما الاختبار الاستطلاعي (routing test) حيث يجري تقدير قدرة أولية بواسطته، ثم يتم بعد ذلك توجيه المفحوص للمرحلة الثانية بحيث يتعرض المفحوص لاختبار فرعي يناسب قدرته المقدرة باستخدام الاختبار الاستطلاعي (Lord, 1980) وتحسب القدرة النهائية للمفحوص على أساس مرحلتي الاختبار التكييفي مجتمعين (Hambleton & Swaminathan, 1985). والإجراء العملي لاختيار فقرات الاختبار

يجري توجيه المفحوص إلى المرحلة التالية بناءً على القدرة المقدرة من المراحل السابقة، وقد تتضمن كل مرحلة فقرة واحدة أو مجموعة من الفقرات، ويتبع هذه الاستراتيجية استراتيجية القياس التكيفي الهرمي، حيث يتم ترتيب الفقرات في ترتيب هرمي مرتكز إلى صعوبة الفقرات ويتم هنا ابتداء الاختبار بفقرة ما وبعد الإجابة عنها يتم الانتقال إلى فقرتين إحداهما في حالة الإجابة الصحيحة والأخرى في حالة الإجابة الخطأ (Sands & Waters & Bride, 2001).

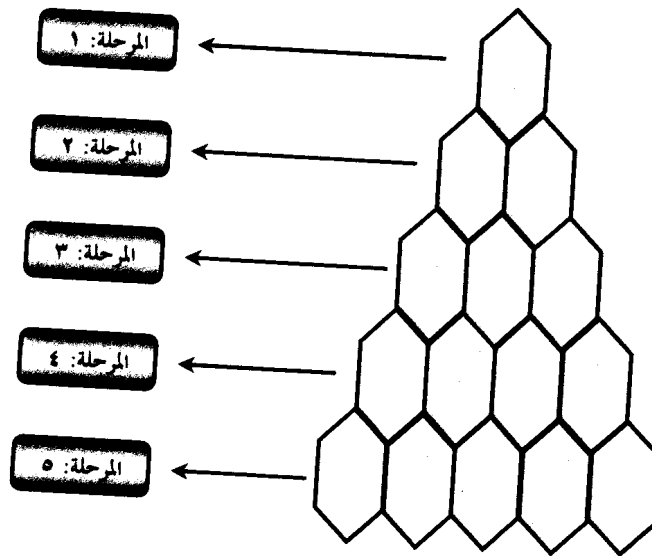
والشكل رقم (١) الآتي يبين طريقة تنظيم

الفقرات.

الاستطلاعي هو اختيار فقراته من بين فقرات الاختبارات الفرعية الخاصة بالمرحلة الثانية، فبعد تدريج تجمع الفقرات وتقسيمها إلى اختبارات فرعية ذات صعوبة متدرجة ومختلفة يتم انتقاء عدد قليل من الفقرات لتكوين الاختبار الاستطلاعي الذي يتم بواسطته تحديد مستوى قدرة ابتدائي للمفحوص، الأمر الذي يترتب عليه توجيه المفحوص لأحد الاختبارات الفرعية الخاصة بالمرحلة الثانية (Ghiselli, Campbell & Zedeck, 1981)، ويمكن أن تتداخل الاختبارات الفرعية فيما بينها لتكوين روابط لكن ذلك ليس ضرورياً (Wright & Stone, 1979).  
ثانياً: استراتيجية القياس التكيفي متعدد المراحل

(Multi - Stage Strategy)

وتتكون هذه الاستراتيجية من عدة مراحل



الشكل (١). اختبار تكيفي هرمي مكون من خمس مراحل بتجمع فقرات عدد فقراته ١٥ فقرة.

(Anastasia & Urbina, 1997)

وتتضمن هذه الاستراتيجية استراتيجية الاختبارات التكيفية الطبقية حيث يتم تصنيف جميع الفقرات إلى طبقات يجري التنقل بينها على أساس القدرة المقدرة بينما يتم اختيار الفقرة ضمن الطبقة اعتماداً على معامل التمييز (Kingsbury & Zara, 1989).

ثالثاً: استراتيجية القياس التكيفي المحوسب

(Computerized Adaptive Testing CAT)

حيث يتم في هذه الاستراتيجية تعريض المفحوص لبعض الفقرات باستخدام الحاسوب لتحديد مستوى قدرته المبدئي ثم تقدم له فقرات بعد ذلك تناسب التقدير المستمر لقدرته اعتماداً على سلسلة الاستجابات للفقرات السابقة (Murphy & Davidshofer, 1994) وعادة ما يقدم للمفحوص فقرة ذات صعوبة أعلى من سابقتها إذا ما كانت إجابته صحيحة للفقرة السابقة، وفقرة ذات صعوبة أدنى من سابقتها إذا ما كانت إجابة المفحوص خطأ للفقرة السابقة (Lincare, 2000) وينتهي الاختبار التكيفي إذا ما تم الوصول إلى تحقيق قاعدة التوقف، وقد أورد وورم (Warm, 1978) بعضاً من قواعد التوقف وهي:

- ١- التوقف عند استفاد الفقرات الموجودة في

تجمع فقرات بنك الأسئلة.

٢- التوقف عندما تصبح قيمة الخطأ المعياري

في التقدير أقل من ٠.٠٦٢٥

٣- التوقف عندما تصبح الفقرات الموجودة

غير كافية.

وبالنسبة للنقطة الثانية فإن إمبرستون ورايس (Emberston & Reise, 2000) يشيران إلى أن قيمة الخطأ المعياري التي يتوقف عندها الاختبار التكيفي يمكن أن يحددها الفاحص.

ويضيف لنكير (Lincare, 2000) في هذا المجال بعض القواعد الأخرى للتوقف مثل كون قياس القدرة بعيداً جداً عن المحك المطلوب، أو أن يُظهر المفحوص سلوكاً يؤدي لإنهاء الاختبار كأن يكون المفحوص سريعاً جداً في الإجابة أو بطيئاً جداً في الإجابة على فقرات الاختبار، وقد أضيفت فيما بعد قواعد أخرى مثل أقل دالة معلومات بالنسبة للاختبار التكيفي المتولد (Choi, Grady & Dodd, 2011) ويذكر كل من بويد ودود وكوي (Boyd, Dodd & Choi, 2010) أن أكثر قاعدة مستخدمة حالياً هي مقدار محدد سلفاً من قبل الفاحص للخطأ المعياري.

وقد أشار الكثير من الباحثين إلى فوائد الاختبارات التكيفية التي من أهمها اختزال عدد الفقرات المقدمة للمفحوصين حيث يذكر هاميلتون وسواميناثان أن الاختبار التكيفي يستخدم فقط ١٠٪ - ٥٠٪ من فقرات الاختبار التقليدي (Hambleton & Swaminathan, 1985) بينما ذكر وورم (Ward, 1984) أن الاختبار التكيفي يحتاج إلى فقرات تقل بـ (٥٠٪ - ٦٠٪) عن الاختبار التقليدي (الخطي) أما لندن وباشلي (Linden & Pashley, 2002) فيذكران

المحاكم والقضاء عند عرض نتائج الاختبارات التكيفية في بعض المجتمعات مثل المجتمع الأمريكي.

وقد أجرى الثوابية (٢٠٠٤) دراسة هدفت إلى استقصاء فاعلية القياس التكيفي ثنائي المرحلة في تقويم التحصيل في مبحث الأحياء لدى عينة من طلبة الصف الثاني الثانوي العلمي في الأردن حيث تم إعداد اختبار تحصيلي في وحدة الوراثة من كتاب الأحياء للصف الثاني الثانوي العلمي تكون من (٥٥) فقرة ثنائية

التدرج من نوع الاختيار من متعدد، وطبق هذا الاختبار على عينة تدرج بلغت (٧٢١) طالباً وطالبة واستخدم البرنامج الإحصائي (winsteps) في تدرج الفقرات، وقد اشتق بعد ذلك ستة اختبارات فرعية أحدها الاختبار الاستطلاعي الذي يتم توجيه المفحوص بناء عليه إلى أحد الاختبارات الفرعية

الخمس الأخرى التي تشكل المرحلة الثانية من الاختبار التكيفي ثم طبق القياس التكيفي على عينة خاصة بلغت (٨١) طالباً وطالبة وقد كان معامل الارتباط بين القدرة المقدرة من خلال الاختبار الاستطلاعي والعلامة المدرسية (٠.٧٥) وبينها وبين علامة الثانوية العامة (٠.٤٦)، أما معامل الارتباط بين الاختبار التكيفي ككل والعلامة المدرسية فكان (٠.٧٠)، ومع علامة الثانوية العامة (٠.٤٢)، بينما كان معامل الارتباط بين الاختبار الخطي ككل والعلامة المدرسية (٠.٧٦) وبينه وبين علامة الثانوية العامة (٠.٥٢).

وفي الدراسة التي أجراها العموش (٢٠٠٣)

أن الاختبار التكيفي يخفض عدد الفقرات التي تطبق بمقدار ٩٠.٨٪.

ويؤدي تقليل عدد الفقرات التي يتعرض إليها المفحوص إلى تحسن وضع سرية الأسئلة إذ إن انكشافها للمفحوصين يكون أقل (Wainer, 2000) يؤدي ذلك أيضاً إلى التقليل من احتمالية معاناة المفحوص من التعب والملل والإحباط (Hambleton, Swaminathan, Hulin, Drasgow & Parsons, 1983)، (1991) يضاف إلى ذلك الكفاءة والدقة في تقدير قدرة المفحوص من خلال عدد قليل من الفقرات التي تنسجم مع قدرة المفحوص (Crist, 1989)، (Wainer, 2000) ولم يغفل الباحثون محددات استخدام القياس التكيفي فقد ذكر كرس (Crist, 1989) المحددات الآتية:

- ١ - نقص الأجهزة والمحددات والبرمجيات.
- ٢ - ضرورة إتقان المفحوصين لمهارات استخدام الحاسوب.
- ٣ - تتطلب الاختبارات التكيفية دقة عالية في عملية معايرة الفقرات أو تدرجها (Item Calibration) ويضيف هامبلتون وزميله (Hambleton & Swaminathan, 1985) ما يلي:

- ١ - الكلفة المادية العالية من حيث الإعداد والتحضير والتجهيز والبرمجيات.
- ٢ - صعوبة الإعداد والتحضير بسبب استخدام عدد كبير من الفقرات لتكوين تجمع الفقرات.
- ٣ - بعض الإشكالات القانونية من قبل



بعنوان فاعلية القياس التكميلي في تقويم بعض القدرات المعرفية لدى طلبة السنة الأولى الجامعية والتي هدفت لاستقصاء فاعلية القياس التكميلي في تقويم القدرة اللفظية والقدرة التقليدية وفق الطريقتين التقليدية والحديثة أظهرت النتائج لاختبار القدرة الرياضية تفوق القياس التكميلي في مستويات القدرة المنخفضة وفي مستويات القدرة المرتفعة بينما في مستويات القدرة المتوسطة فقد كان القياس التقليدي أكثر فاعلية، أما في اختبار القدرة اللفظية فقد كان القياس التكميلي أكثر فاعلية في مستويات القدرة المتوسطة والمرتفعة، وبشكل عام أظهرت الدراسة أن القياس التكميلي كان أكثر فاعلية من الاختبار التقليدي اعتماداً على الكفاءة النسبية، واختزال الوقت والفقرات المستخدمة.

وفي دراسة أجراها البرصان (٢٠٠٦) هدفت إلى استقصاء أثر عدد مراحل القياس التكميلي وعدد أسئلة كل مرحلة في تقدير القدرة والخطأ المعياري في التقدير باستخدام فقرات ثنائية وفقرات متعددة التدرج، استخدم اختبارين تحصيليين في مادة الرياضيات تكون أحدهما من فقرات ثنائية التدرج، والآخر تكون من فقرات متعددة التدرج، طبقهما على عيّنتين من طلبة الصف الأول الثانوي في مدينة عمان، وتم تدرج كل من فقرات الاختبارين باستخدام برمجية (WINSTEPS)، طبق بعدها إستراتيجية القياس التكميلي متعدد المراحل، حيث حصل على خمسة تقديرات للقدرة لكل مفحوص في العينة الأولى التي

طبق عليها القياس التكميلي باستخدام فقرات ثنائية التدرج، وكذلك حصل على خمسة تقديرات للقدرة لكل مفحوص في العينة الثانية، وقد أشارت النتائج إلى أنه يستمر اقتراب تقدير القدرة والخطأ المعياري في التقدير من تقدير القدرة المرجعية والخطأ المعياري المرافق لها في القياس التكميلي المكون من فقرات ثنائية التدرج حتى المرحلة الثالثة أما المرحلة الرابعة فلم تكن مؤثرة، أما في القياس التكميلي المعتمد على فقرات متعددة التدرج فكانت عملية الاقتراب من القدرة المرجعية والخطأ المعياري المرافق لها مجدية حتى المرحلة الثانية بينما لم تعد مجدية في المرحلتين الثالثة والرابعة، وقد أشارت النتائج كذلك إلى أن القياس التكميلي المبني على فقرات متعددة التدرج أدق من حيث تقدير القدرة والخطأ المعياري في التقدير من القياس التكميلي المعتمد على فقرات ثنائية التدرج.

وقد هدفت دراسة أجراها عبيدات (٢٠٠٨) إلى فحص فاعلية الاختبار التكميلي المحوسب في دقة تقدير القدرة العقلية باستخدام اختبار مصفوفات رافن، ولتحقيق ذلك استخدم اختبار مصفوفات رافن حيث تم تدرج الفقرات التي بلغت في صورتها النهائية (١٠٥) فقرات على عينة تدرج بلغت (٢٦٩٥) طالباً وطالبة، ثم طبق القياس التكميلي المحوسب على عينة ثانية بلغت (٦٣٨) طالباً وطالبة باستخدام محكّن لإنهاء الاختبار هما عدد محدد من الفقرات (٣٠) فقرة، وأدنى خطأ معياري (٠.٢٥)، وتم ذلك

لتطبيق الاختبار التكميني المحوسب، واستخدمت طريقة الأرجحية العظمى لتقدير القدرة وقاعدة أدنى خطأ معياري للتوقف ( $SE \leq 0.25$ ) وقد تم تطبيق الاختبار التكميني على (٥٠) طالباً، وأظهرت النتائج أن متوسط مدة الاختبار كان (٤٧) دقيقة بانحراف معياري (١٩.٣٨)، وكان عدد الفقرات التي جرى تعريض المفحوصين لها يتراوح بين (٢٠) فقرة و (٧٠) فقرة وبمتوسط (٤٠.٢٤) فقرة وبانحراف معياري (١٢.٤٠)، أما صدق الاختبار فقد كان مجدياً وقابلاً للتطبيق، إلا أن الاختبار التكميني المحوسب فلم يكن بديلاً صادقاً لاختبار المعرفة المكتوب شديد الصعوبة أو عالي المعايير.

وقد أجرى فليج وآخرون (Fliege, Becker, Walter, Bjorner, Klapp & Rose, 2005) دراسة بعنوان «تطوير اختبار تكميني محوسب للاكتئاب»،

حيث تم استخدام (٣٢٠) فقرة غطت اختبارات فرعية (النشاطات اليومية، المزاج، الشكاوى، الشخصية، الرضا عن الحياة، النرجسية، الصحة، القلق، فاعلية الذات) وباستخدام الاختبار التكميني المحوسب تبين أن علامة الاختبار التكميني المحوسب ارتبطت مع فقرات الاكتئاب بمعامل ارتباط (٠.٩٥)، ومع اختبار القلق بمعامل ارتباط (٠.٧٩) ومع اختبار الاكتئاب لمركز الدراسات الوبائية معامل ارتباط (٠.٧٦) وبشكل عام تم الاستنتاج أن الاختبار التكميني المحوسب للاكتئاب يقيس أعراض الاكتئاب بدقة عالية وبعدد متدن من

باستخدام طريقتين لتقدير القدرة هما طريقة الأرجحية العظمى (MLE) Maximum Likelihood estimation) وطريقة التقدير البعدي الأعظم (MAP) (Maximum a posteriori)، وأظهرت النتائج أن قاعدة إنهاء الاختبار بعدد محدد من الفقرات توفر تقديرات أدق للقدرة ودالة معلومات أعلى من قاعدة أدنى خطأ معياري وذلك باستخدام طريقتي تقدير القدرة، ومن حيث اختزال الفقرات فقد كانت نسبة اختزال الفقرات باستخدام قاعدة أدنى خطأ معياري (٥٠٪) من قاعدة عدد محدد من الفقرات لكن إذا تمت المقارنة بين القياس التكميني والخطي وصل نسبة الاختزال لعدد الفقرات المطبقة إلى (٧٠٪) تقريباً، وكذلك كان الاختبار التكميني أدق في تقدير القدرة من الاختبار التكميني باستخدام قاعدة أدنى خطأ معياري، وقد كانت نسبة اختزال عدد الفقرات المطبقة بين الاختبار التكميني والاختبار الخطي تساوي ٧٪.

أما في الدراسة التي أجراها روكس وديكريس (Roex & Degryse, 2004) بعنوان «اختبار تكميني محوسب للمعرفة كأداة تقييم في الممارسة العامة» والتي هدفت إلى تقويم جدوى تحويل اختبار مكتوب إلى اختبار تكميني محوسب ودراسة صدق المحتوى لهذا الاختبار، تكون الاختبار من (١٦٥) فقرة تطبق في فترة زمنية مقدراها (٣) ساعات، وقد تم تدريج الفقرات باستخدام عينة بلغت (١٠٠٠) طالب وطالبة، وقد استخدم برنامج (Fast test pro)

الفقرات المطبقة على المستجيب.

وقد أجرى فيزبول ووانج ويلير (Vispoel, Wang & Blelier, 1997) دراسة بعنوان «الاختبار التكيفي المحوسب واختبار الفقرات الثابتة لمهارات الاستماع للموسيقى: مقارنة للكفاءة والصدق التلازمي» هدفت إلى مقارنة الكفاءة والدقة والصدق التلازمي بين الاختبارات التكيفية والاختبارات الخطية، واستخدم الباحثون بيانات مولدة (محاكاة) لعينة بلغت (٢٢٠٠) مفحوص، حيث أظهرت النتائج أن الاختبار التكيفي الموسيقي يحتاج لفقرات تقل بنسبة (٥٠٪ - ٩٣٪) ليطابق الثبات والصدق التلازمي للاختبار الخطي، بل إن الاختبار التكيفي أعطى مستويات أعلى من الثبات والصدق التلازمي من الاختبار الخطي عند إبقاء طول الاختبار التكيفي ثابتاً.

وقام هندركسون (Hendrickson, 2002) بدراسة هدفت إلى الكشف عن إمكانية استخدام الاختبار التكيفي ثنائي المرحلة بديلاً عن الاختبارات التحصيلية التقليدية، حيث تم استخدام اختبار في اللغة الإنجليزية لذلك على طلبة الصف السادس وأظهرت النتائج أن الخصائص السيكومترية المتعلقة بالثبات ودقة القياس تتحسن عند تدريج الاختبارات التقليدية لتصبح اختبارات تكيفية.

وقد قام كيم وبلير (Kim & Plake, 1993) بدراسة هدفت إلى المقارنة بين اختبار تكيفي محوسب واختبار تكيفي ثنائي المرحلة باستخدام المحاكاة، حيث

أظهرت النتائج أن الخصائص الإحصائية (صعوبة الفقرات للاختبار الاستطلاعي) لها أثر أساسي في دقة قياس القدرة وبشكل عام تفوق الاختبار التكيفي المحوسب على الاختبار ثنائي المرحلة المساوي له بعدد الفقرات وفق محكّ دقة القياس والفعالية، ووجد أن الاختبارات التكيفية ثنائية المرحلة التي تستخدم مدى واسعاً من الصعوبة وعدداً فردياً من الفقرات في المرحلة الثانية تعطي تقديرات أكثر دقة من الأشكال الأخرى للاختبارات ثنائية المرحلة، وأشار الباحثان إلى أن الاختبارات ثنائية المرحلة تعتبر بديلاً عملياً مناسباً عند عدم توفر أجهزة حاسوب.

ويتبين مما سبق أن الدراسات السابقة أكدت على فاعلية القياس التكيفي بشكل يفوق القياس الخطي مع الأخذ بعين الاعتبار دقة التقدير للقدرة كما في دراسة العموش (٢٠٠٣) والبرصان (٢٠٠٦)، ودراسة فليج وزملائه (Fliege, Becker, Walter, Bjorner, Klapp & Rose, 2005) ودراسة هندركسون (Hendrickson, 2002)، ودراسة كيم وبلير (Kim & Plake, 1993)، ومن حيث الصدق فقد أشارت الدراسات إلى تمتع الاختبار التكيفي بدرجة مناسبة من الصدق والثبات كما في دراسة كل من فيزبول وزملائه (Vispoel, Wang & Blelier, 1997) ودراسة ريكس وديكرس (Roex & Degryse, 2004).

مشكلة الدراسة:

عادة ما تعتمد الاختبارات التكيفية على

التكيفي المبني باستخدام فقرات إجابة منتقاة ثنائية التدرج (Dichotomous Items) وفقرات من نوع الإجابة المنشأة متعددة التدرج (Polytomous Items) معاً وذلك باستخدام طريقة قابلة للتطبيق وهي جعل المرحلة الأخيرة فقرات متعددة التدرج من نوع الإجابة المنشأة (المصوغة)، يجري إضافتها لإجابات المفحوص السابقة وحساب قدرته المتحققة من جميع الفقرات الثنائية التدرج والمتعددة التدرج فيما بعد، بعد أن ينهي الطالب إجابته عليها ورقياً أو حاسوبياً وتصحيحها من قبل مصححين ثم إدخالها مرمزة للبرنامج الحاسوبي الذي يقوم بحساب قدرة المفحوص معتمداً إحدى طرق التقدير في نظرية الاستجابة للفقرة. أهمية الدراسة:

تتمثل الأهمية النظرية للدراسة في إدخال فقرات متعددة التدرج من نوع الفقرات ذات الإجابة المنشأة متعددة التدرج إلى القياس التكيفي وذلك لأهمية هذا النوع من الفقرات في قياس بعض نتائج التعلم التي يصعب قياسها بفقرات ثنائية التدرج من نوع الفقرات ذات الإجابة المنتقاة، وبشكل هذا العمل داعماً لفكرة القياس التكيفي الذي أدواته الاختبارات التكيفية التي توائم بين قدرة المفحوص وصعوبة الفقرات التي يتعرض إليها، الأمر الذي يؤدي إلى اختصار الجهد والكلفة واختزال عدد الفقرات المستخدمة والتقليل من كشف الفقرات للمفحوصين، وبذلك تنتفي بعض المبررات التي يسوقها معارضو القياس التكيفي.

الفقرات من نوع الإجابة المنتقاة ذات التدرج الثنائي خصوصاً في الاختبارات التكيفية المحوسبة (Computerized adaptive testing CAT)، لكن ومن وجهة نظر كثير من اختصاصي القياس والتقويم بأن الفقرات من نوع الإجابة المنتقاة ذات التدرج الثنائي لا يمكنها قياس نتائج التعلم المعقدة، هذا عدا عن أن الفقرات ذات الإجابة المنشأة لا تعطي الفرصة لتخمين الإجابة الصحيحة إذا لم يكن يمتلك المعرفة الضرورية. ويشير كامبل (Campbell, 2000) في هذا المجال إلى أن المزج بين النوعين من الفقرات ذات الإجابة المنتقاة وذات الإجابة المنشأة هو الأكثر ملاءمة، وبناءً عليه فإن أي اختبار إذا ما أريد له أن يقيس نتائج التعلم على اختلاف أنواعها فلا بد من أن يحتوي على فقرات من نوع الإجابة المنشأة (المصوغة) (Constructed Response Items) ذات التدرج متعدد الفئات، ومن المعلوم أن هناك صعوبات ترتبط بتطبيق الفقرات ذات الإجابة متعددة التدرج خصوصاً في الاختبارات التكيفية المحوسبة (CAT)، لكن في هذه الدراسة سيتم التعرض لطريقة جديدة قابلة للتطبيق بإدخال فقرات من نوع الإجابة المنشأة (Constructed Response Items) متعددة التدرج في الاختبارات التكيفية سواء كانت محوسبة أو باستخدام الورقة والقلم.

#### أهداف الدراسة:

تهدف هذه الدراسة إلى اختبار فاعلية القياس

أما الأهمية العملية للدراسة فتتمثل في إيجاد طريقة قابلة للتطبيق لإدخال فقرات ذات إجابة منشأة متعددة التدرج في فقرات الاختبار التكيفي وذلك يجعل المرحلة الأخيرة للاختبار هي إجابة المفحوص عن فقرات ذات إجابة منشأة من نوع الفقرات ذات التدرج المتعدد وذلك اعتماداً على القدرة المقدرة من الفقرات ثنائية التدرج المقدمة سابقاً للمفحوص، سواء كان الاختبار محوسباً أو اختبار ورقة وقلم وقياس فاعلية الاختبار التكيفي المبني من فقرات ثنائية التدرج وفقرات متعددة التدرج معاً.

أسئلة الدراسة:

يتمثل السؤال الرئيس للدراسة بالسؤال:

ما فاعلية إدخال فقرات ذات إجابة منشأة متعددة التدرج كمرحلة أخيرة في القياس التكيفي متعدد المراحل على القدرة المقدرة للمفحوصين بشكل عام؟ مقاسة بما يأتي:

(أ) معامل الارتباط بين تقديرات القدرة المقدرة في المراحل الأولى والثانية، المتكونتين من فقرات ثنائية التدرج، والثالثة المتكونة من فقرات متعددة التدرج والقدرة الحقيقية المقدرة من الاختبار الخطي الذي يحتوي فقرات ثنائية ومتعددة التدرج.

(ب) متوسط القيم المطلقة للفروق بين تقديرات القدرة المقدرة في المراحل الثلاث والقدرة الحقيقية المقدرة من الاختبار الخطي الذي يحتوي فقرات ثنائية ومتعددة التدرج.

(ج) مقدار انخفاض الخطأ المعياري في التقدير أثناء التقدم في المراحل.

محددات الدراسة:

- ١ - استخدمت هذه الدراسة بيانات مولدة باستخدام الحاسوب.
  - ٢ - استخدمت هذه الدراسة القياس التكيفي متعدد المراحل.
- مصطلحات الدراسة:

القياس التكيفي (Adaptive Testing):

اختبار يتكون من فقرات متتابعة أو مجموعات متتابعة من الفقرات بحيث تتحدد الفقرة أو مجموعة الفقرات الآتية بناءً على نتيجة الفقرات السابقة وصولاً إلى مستوى تقاربي (Asymptotic level).

الأسئلة ذات الإجابة المنتقاة (ثنائية التدرج)

(dichotomous items):

أسئلة تتكون الإجابة فيها من متغير ثنائية التدرج حيث يعطي الدرجة صفراً للإجابة الخطأ والدرجة واحداً للإجابة الصحيحة وعادة ما تكون هذه الأسئلة ذات إجابة منتقاة.

الأسئلة ذات الإجابة المنشأة (متعددة التدرج)

(polytomous items):

أسئلة تتكون الإجابة فيها من متغير له تدرج ثلاثي أو رباعي أو أكثر بناءً على فئات الإجابة حيث تعطى المعرفة الجزئية علامات حيث تعطى الدرجة صفر للإجابة الخطأ، والدرجة (١) للإجابة الجزئية

٢- عينة القياس التكميلي: تكونت عينة القياس التكميلي من (١٠٠) مفحوص جرى اختيارهم من بين الألفي مفحوص الذين شكلوا عينة التدريج، وقد روعي في اختيارهم تمثيل جميع شرائح القدرة للمفحوصين.

٣- تدريج الفقرات: تم تدريج الفقرات الستين ذوات التدريج الثنائي والفقرات العشر ذوات التدريج المتعدد باستخدام برنامج RUMM 2020، الذي يوفر الإمكانية لتدريج فقرات ثنائية التدريج وفقرات متعددة التدريج في اختبار واحد، ويبين الجدول (١) قيم معلم الصعوبة للفقرات السبعين الناتجة باستخدام البرنامج (RUMM2020).

ذات خطوة واحدة صحيحة والدرجة (٢) للإجابة الجزئية التي فيها خطوتان صحيحتان وهكذا (وغالباً ما تكون هذه الأسئلة ذات إجابة منشأة وأحياناً تكون مقياس تقدير).

منهجية الدراسة:

١- عينة التدريج: تكونت عينة تدريج الفقرات من (٢٠٠٠) مفحوص جرى توليد إجاباتهم حاسوبياً باستخدام المحاكاة (simulation) وباستخدام برنامج (WENGEN) حيث تم توليد (٦٠) فقرة ثنائية التدريج لكل مفحوص بالإضافة إلى (١٠) فقرات متعددة التدريج لكل مفحوص كل فقرة تحتوي على (٥) مستويات للإجابة تراوحت بين صفر و(٤).

الجدول رقم (١). قيم معلم الصعوبة لفقرات الاختبار من نوعي ثنائي التدريج ومتعدد التدريج.

الفقرة	نوعها	معلم الصعوبة	الخطأ المعياري	إحصائي المطابقة
١٠٠٠١	ثنائية التدريج	- ١.١٩٦	٠.٠٥٤	٠.٥٨٨
١٠٠٠٢	ثنائية التدريج	- ٠.٢٣٥	٠.٠٤٩	٠.٢٦٩
١٠٠٠٣	ثنائية التدريج	١.٤٩٤	٠.٠٥٩	٠.٣٨١
١٠٠٠٤	ثنائية التدريج	- ٠.٢١٥	٠.٠٤٩	٠.٠٣
١٠٠٠٥	ثنائية التدريج	٠.٤٦٢	٠.٠٥	- ١.٢٧٧
١٠٠٠٦	ثنائية التدريج	٠.٧	٠.٠٥٢	٠.٨٠٣
١٠٠٠٧	ثنائية التدريج	- ٠.٢٣٧	٠.٠٤٩	- ٠.٠١١
١٠٠٠٨	ثنائية التدريج	- ٠.١٩١	٠.٠٤٩	- ٠.٧٩١
١٠٠٠٩	ثنائية التدريج	- ١.٠٧٩	٠.٠٥٣	- ٠.٣٣٢
١٠٠١٠	ثنائية التدريج	١.٧٩١	٠.٠٦٤	٠.٧٢٨
١٠٠١١	ثنائية التدريج	- ١.٦	٠.٠٥٨	١.٨٥٥
١٠٠١٢	ثنائية التدريج	٠.٥٨٤	٠.٠٥١	- ٠.٢٨٧
١٠٠١٣	ثنائية التدريج	٠.٨٣	٠.٠٥٢	- ٠.٢٠٨
١٠٠١٤	ثنائية التدريج	٠.١١	٠.٠٤٩	- ٠.٦٨٤
١٠٠١٥	ثنائية التدريج	٢.٢٨٦	٠.٠٧٣	٠.١٧٢

تابع الجدول رقم (١).

الفقرة	نوعها	معلم الصعوبة	الخطأ المعياري	إحصائي المطابقة
١٠٠١٦	ثنائية التدرج	٢.١٣٨ -	٠.٠٦٧	٠.٩٤٦
١٠٠١٧	ثنائية التدرج	٠.٥٣٣	٠.٠٥١	١.٧٠٥
١٠٠١٨	ثنائية التدرج	٠.٩٣٣ -	٠.٠٥٢	٠.٩٨٨
١٠٠١٩	ثنائية التدرج	١.٦٧٩	٠.٠٦٢	٠.٠٧٤ -
١٠٠٢٠	ثنائية التدرج	٠.٦٤ -	٠.٠٥	٠.٦١٣
١٠٠٢١	ثنائية التدرج	٠.٦٨٧ -	٠.٠٥١	٠.٥٦١ -
١٠٠٢٢	ثنائية التدرج	٠.٨٩٥	٠.٠٥٣	٠.٩٥٦
١٠٠٢٣	ثنائية التدرج	١.٢٥٩	٠.٠٥٦	٠.١٣٥ -
١٠٠٢٤	ثنائية التدرج	٢.٢٢ -	٠.٠٦٨	٠.٥٧٦
١٠٠٢٥	ثنائية التدرج	٢.٠٠٥	٠.٠٦٧	٠.٢٧٧ -
١٠٠٢٦	ثنائية التدرج	٠.١٠٩	٠.٠٤٩	٠.٠٣٥ -
١٠٠٢٧	ثنائية التدرج	١.٢٥٥ -	٠.٠٥٥	٠.٥٣٤ -
١٠٠٢٨	ثنائية التدرج	١.١٢٧	٠.٠٥٥	٠.٤٣١
١٠٠٢٩	ثنائية التدرج	١.٣٥٧	٠.٠٥٧	١.٠٢١
١٠٠٣٠	ثنائية التدرج	٢.٢٥٣ -	٠.٠٦٩	٠.٨٢٦ -
١٠٠٣١	ثنائية التدرج	٠.٦٦٧ -	٠.٠٥١	٠.٤٦٤ -
١٠٠٣٢	ثنائية التدرج	١.٤٥٧ -	٠.٠٥٧	٠.٥٨١
١٠٠٣٣	ثنائية التدرج	١.٠٢٩ -	٠.٠٥٣	٠.١٠٦
١٠٠٣٤	ثنائية التدرج	٠.٢٤٢ -	٠.٠٤٩	٠.١٩٨
١٠٠٣٥	ثنائية التدرج	٠.١٩٧	٠.٠٥	١.٥٥٢
١٠٠٣٦	ثنائية التدرج	٠.٢٧٦	٠.٠٥	٠.٤١٩
١٠٠٣٧	ثنائية التدرج	٠.١٦٤ -	٠.٠٤٩	٠.٦٥٨ -
١٠٠٣٨	ثنائية التدرج	٠.٦٤٦	٠.٠٥١	٠.١٤ -
١٠٠٣٩	ثنائية التدرج	٢.٣٣ -	٠.٠٧	٠.٢٣١ -
١٠٠٤٠	ثنائية التدرج	٠.٧٣٩	٠.٠٥٢	٠.٢٨٢ -
١٠٠٤١	ثنائية التدرج	٠.٨٠٤	٠.٠٥٢	١.١٦٨ -
١٠٠٤٢	ثنائية التدرج	٠.٤٢٥	٠.٠٥	٠.٦٤٦ -
١٠٠٤٣	ثنائية التدرج	٠.١٤٢ -	٠.٠٤٩	٠.٥٣٩
١٠٠٤٤	ثنائية التدرج	٠.٦٠٣	٠.٠٥١	٠.٨٨١ -
١٠٠٤٥	ثنائية التدرج	١.٣٢٤ -	٠.٠٥٥	٠.٥٩ -
١٠٠٤٦	ثنائية التدرج	٠.٥٤٥	٠.٠٥١	٠.٠٨٩
١٠٠٤٧	ثنائية التدرج	٠.٠٥٥	٠.٠٤٩	٠.٢٦٣ -
١٠٠٤٨	ثنائية التدرج	٠.٨٩٦	٠.٠٥٣	٠.٨٥٨
١٠٠٤٩	ثنائية التدرج	٠.٦٨٣	٠.٠٥٢	٠.١٥٥ -

تابع الجدول رقم (١).

الفقرة	نوعها	معلم الصعوبة	الخطأ المعياري	إحصائي المطابقة
١٠٠٥٠	ثنائية التدرج	٠.٨١٨ -	٠.٠٥١	١.٧١٢
١٠٠٥١	ثنائية التدرج	١.١٦	٠.٠٥٥	٠.٩٣٢ -
١٠٠٥٢	ثنائية التدرج	١.٣٨١	٠.٠٥٨	٠.٦٥٨
١٠٠٥٣	ثنائية التدرج	١.٠٥٥ -	٠.٠٥٣	٠.٠٦١ -
١٠٠٥٤	ثنائية التدرج	١.٢٥٧	٠.٠٥٦	٠.٢١٥ -
١٠٠٥٥	ثنائية التدرج	٠.٤٧	٠.٠٥	٠.٠٩
١٠٠٥٦	ثنائية التدرج	٠.٣٠٢ -	٠.٠٤٩	٢.٨١١
١٠٠٥٧	ثنائية التدرج	٠.٧٣٣	٠.٠٥٢	٠.٠٩١ -
١٠٠٥٨	ثنائية التدرج	١.١٤٥	٠.٠٥٥	٠.١٨٣
١٠٠٥٩	ثنائية التدرج	٠.٤١٨	٠.٠٥	٠.٣٦٨ -
١٠٠٦٠	ثنائية التدرج	١.٤٩ -	٠.٠٥٧	٠.٠٧١ -
١٠٠٦١	متعددة التدرج	٠.٥٦٨ -	٠.٠٢٣	٠.٨٦٦ -
١٠٠٦٢	متعددة التدرج	٠.٤٧١ -	٠.٠٢٢	١.٠٥٨ -
١٠٠٦٣	متعددة التدرج	٠.٨٨٥ -	٠.٠٢٥	١.٤٢٢ -
١٠٠٦٤	متعددة التدرج	٠.٣٢١ -	٠.٠٢٤	١.٨٩٩ -
١٠٠٦٥	متعددة التدرج	٠.٠٠٢ -	٠.٠٢٤	٠.٨٧٢ -
١٠٠٦٦	متعددة التدرج	٠.٨٤٢	٠.٠٢٧	٠.٩٧٧ -
١٠٠٦٧	متعددة التدرج	٠.٠٦٦	٠.٠٢٢	٠.١٦٢
١٠٠٦٨	متعددة التدرج	٠.٩٩٨ -	٠.٠٢٦	١.٨٣٤ -
١٠٠٦٩	متعددة التدرج	٠.٥٢ -	٠.٠٢٥	٠.٧٦١ -
١٠٠٧٠	متعددة التدرج	٠.٩ -	٠.٠٢٦	٠.٩٨ -

الصعوبة لكل فقرة جرى اختيار (٥) فقرات ثنائية التدرج تمثل مستويات الصعوبة جميعها لتكون اختباراً استطلاعياً بشكل المرحلة الأولى من القياس التكيفي، وأرقام هذه الفقرات هي: (١٠، ١١، ١٨، ٤٧، ٤٨).

٢- جرى تقسيم الفقرات ثنائية التدرج الباقية وهي (٥٥) فقرة لتكون خمس مجموعات فرعية تشكل المرحلة الثانية من القياس التكيفي، والجدول رقم (٢) يبين أرقام الاختبارات الفرعية الخمسة.

وبلاحظ من الجدول رقم (١) أن قيم معلم الصعوبة تراوحت بين القيمة (٢.٣٣ -) لوجيت للفقرة رقم (٣٣) والقيمة (٢.٢٨٦) لوجيت للفقرة رقم (١٥)، كذلك كانت جميع الفقرات السبعين مطابقة لنموذج راش إذ إن إحصائي المطابقة لم يقع خارج الفترة (٢.٥ - ٢.٥) وذلك حسب شرط برنامج RUMM 2020. إجراءات القياس التكيفي:

١- بعد ترتيب الفقرات تصاعدياً حسب معلم



إسماعيل سلامة البرصان: فاعلية القياس التكميلي باستخدام فقرات...

الجدول رقم (٢). الفقرات التي شكلت المجموعات الفرعية.

المجموعة	أرقام الفقرات
الفرعية الأولى	٥٣، ٩، ١، ٢٧، ٤٥، ٣٢، ٦٠، ١٦، ٢٤، ٣٠، ٣٩
الثانية	٨، ٤، ٢، ٧، ٣٤، ٥٦، ٢٠، ٣١، ٢١، ٥٠، ٣٣
الثالثة	١٧، ٥٥، ٥، ٤٢، ٥٩، ٣٦، ٣٥، ١٤، ٢٦، ٤٣، ٣٧
الرابعة	٢٢، ١٣، ٤١، ٤٠، ٥٧، ٦، ٤٩، ٣٨، ٤٤، ١٢، ٤٦
الخامسة	١٥، ٢٥، ١٩، ٣، ٥٢، ٢٩، ٢٣، ٥٤، ٥١، ٥٨، ٢٨

المتعدد تتناسب مع قدرة المفحوص الجديدة وحساب القدرة النهائية باستخدام فقرات المراحل الثلاث معاً وبالتالي أصبح لكل مفحوص ثلاث قدرات  $\Theta_1$  متحصلة من الاستجابات على فقرات الاختبار الاستطلاعي،  $\Theta_2$  متحصلة من الاستجابات على فقرات اختبار المرحلة الثانية وفقرات المرحلة الأولى (الاستطلاعية) وجميعها ثنائية التدرج،  $\Theta_3$  متحصلة عن الاستجابات على فقرات اختبار المراحل الثلاث والتي كان آخرها من الفقرات ذات التدرج المتعدد.

#### النتائج

يبين الملحق رقم (١) تقديرات القدرة المتحصلة لأفراد عينة القياس التكميلي من تطبيق القياس التكميلي في المرحلة الأولى ( $\Theta_1$ ) وهي مرحلة الاختبار الاستطلاعي، و( $\Theta_2$ ) وهي القدرة المقدرة من المرحلة الثانية ثنائية التدرج مضافاً لها فقرات المرحلة الأولى ثنائية التدرج، وكذلك ( $\Theta_3$ ) وهي القدرة المقدرة من المرحلة الثالثة المكونة من فقرات متعددة التدرج مضافاً لها فقرات المرحلة الثانية ثنائية التدرج وفقرات المرحلة الأولى ثنائية التدرج إضافة للخطأ المعياري في التقدير لجميع المراحل وبجانبتها لكل مفحوص القدرة المقدرة من الاختبار الخطي بفقراته السبعين ( $\Theta_L$ ). وللإجابة عن السؤال الأول من أسئلة الدراسة يبين الجدول (٣) المستخلص من الملحق (١) المتوسطات الحسابية للقدرة المقدرة في المراحل الثلاث، ومعاملات الارتباط بيرسون بين كل من

٣- جرى تقسيم الفقرات متعددة التدرج وهي (١٠) فقرات لتكون (٥) اختبارات فرعية كل منها يتكون من فقرتين اثنتين.

٤- جرى تقديم فقرات الاختبار الاستطلاعي لكل المفحوصين في عينة القياس التكميلي وحساب قدرة مبدئية لكل منهم باستخدام برنامج (RUMM 2020) وذلك باستخدام البيانات السابقة التي تحتوي استجاباتهم على الاختبار الخطي، بحيث تم حذف الاستجابات جميعها باستثناء الاستجابات على فقرات الاختبار الاستطلاعي الخمس.

٥- بناءً على القدرة المحسوبة لكل مفحوص جرى تقديم خمس فقرات أخرى مختارة تتناسب مع قدرة المفحوص المبدئية من المجموعات الفرعية المكونة من فقرات ثنائية التدرج ذات إجابة منتقاة، وحساب قدرة المفحوص الجديدة باستخدام الفقرات الخمس الجديدة والفقرات السابقة التي كونت الاختبار الاستطلاعي.

٦- بناءً على القدرة المحسوبة من المرحلة الثانية جرى تعريض المفحوصين لفقرتين من ذوات التدرج

السبعين، والمتوسط الحسابي للخطأ المعياري في التقدير للمفحوصين، ومتوسط الفروق بين الخطأ المعياري في التقدير لكل مرحلة منسوبة إلى الخطأ المعياري في تقدير القدرة المتأتية من الاختبار الخطي بفقراته السبعين.

القدرة المقدرة في المراحل الثلاث والقدرة المقدرة من الاختبار الخطي بفقراته السبعين، ومتوسط القيم المطلقة للفروق في تقديرات القدرة في المراحل الثلاث مقارنة بالقدرة المقدرة من الاختبار الخطي بفقراته

الجدول رقم (٣). القيم الإحصائية المتعلقة بتقديرات القدرة في المراحل المختلفة في القياس التكميلي مقارنة بالقدرة الحقيقية المقدرة من الاختبار الخطي.

الاختبار الخطي ( $\Theta_L$ )	المرحلة الثالثة (فقرات ثنائية + متعددة التدرج) ( $\Theta_3$ )	المرحلة الثانية (فقرات ثنائية التدرج) ( $\Theta_2$ )	المرحلة الأولى (فقرات ثنائية التدرج) ( $\Theta_1$ )	
٠.٠٨٩ -	٠.١٧٦ -	٠.٢٤٩ -	٠.٢٥٦ -	المتوسط الحسابي للقدرة
١	٠.٩٣٠	٠.٨٨١	٠.٧٦٨	معامل الارتباط
٠	٠.٤١٣	٠.٥٨٢	٠.٧٦٧	متوسط القيم المطلقة للفروق
٠.٢٥٥	٠.٥٣٣	٠.٧٥٨	١.١٨٢	المتوسط الحسابي للخطأ المعياري
٠	٠.٢٧٨	٠.٥٠٤	٠.٩٢٨	متوسط الفروق في الخطأ المعياري

النوعين من الفقرات ثنائية ومتعددة التدرج، ويدعم ذلك أنه حينما أضيفت المرحلة الثالثة التي تحتوي على فقرات متعددة التدرج من نوع الإجابة المنشأة تحسن تقدير القدرة باتجاه القدرة الحقيقية حيث أصبح (٠.١٧٦) أي بفارق (٠.٠٨٧) لوجيت، وهذه النتائج تتفق مع دراسة البرصان (٢٠٠٦) من حيث المرحلة الأولى من القياس التكميلي بينما تختلف من حيث المرحلة الثانية للقياس التكميلي، وبالمثل مع دراسة الثوابية (٢٠٠٤) ودراسة هندركسون (٢٠٠٢)، وكما هو معلوم أن تقديرات القدرة تمتد في المدى السالب والمدى الموجب الأمر الذي يفرض أن يكون هناك مقارنة من نوع آخر وهي التي تستخدم الفروق المطلقة

ويظهر من الجدول رقم (٣) أن المتوسط الحسابي لقدرات المفحوصين في الاختبار الاستطلاعي ذي الفقرات الخمس كانت (٠.٢٥٦) لوجيت وهي قريبة نسبياً من المتوسط الحسابي للقدرة الحقيقية المتأتية من الاختبار الخامس بفقراته السبعين وهي (٠.٠٨٩) أي بفارق (٠.١٦٧) لوجيت، وفي المرحلة الثانية كان متوسط تقديرات القدرة يساوي (٠.٢٤٩) وهنا يظهر أن المرحلة الثانية لم تحسن تقدير القدرة للاقتراب من القدرة الحقيقية إلا بمقدار ضئيل وهو (٠.٠٠٧) وهو قريب من الصفر، والسبب هنا أن كلتا المرحلتين الأولى والثانية تتألفان من فقرات ثنائية التدرج من نوع الإجابة المنتقا، بينما تألف الاختبار الخطي من كلا

(القيم المطلقة لباقي الطرح) بين التقديرات للقدرة في مراحل القياس التكميلي والقدرة المقدرة من الاختبار الخطي، ذلك أن المتوسط الحسابي للتقديرات الموجبة والسالبة قد يعطي استنتاجاً مضللاً بسبب تعادل الموجب مع السالب.

فباستخدام المتوسط الحسابي للفروق المطلقة يظهر أن تقديرات المرحلة الأولى الاستطلاعية تبعد في المتوسط عن القدرة الحقيقية بمقدار (٠.٧٦٧) لوجيت، وفي المرحلة الثاني أصبحت تبعد بمقدار (٠.٥٨٢) لوجيت، وفي المرحلة الأخيرة التي احتوت فقرات متعددة التدرج كانت القدرة في المتوسط تبعد بمقدار (٠.٤١٣) لوجيت وهي تعد قريبة من القدرة الحقيقية، لاسيما لو أننا أعدنا تطبيق الاختبار الخطي بفقراته السبعين فسنحصل على فروقات ربما تساوي الفروق الحالية مع القياس التكميلي الذي استخدم فقط (١٢) فقرة من أصل (٧٠) فقرة مع الانتباه لإشكالية احتواء القياس التكميلي على كلا النوعين من الفقرات.

ومن حيث المؤشر الثالث على فاعلية القياس التكميلي المبني باستخدام فقرات ثنائية ومتعددة التدرج وهو معامل الارتباط فيظهر أن معامل الارتباط في المرحلة الاستطلاعية الأولى مع القدرة الحقيقية كان (٠.٧٦٧) بتباين مفسر مقداره (٠.٥٨٨) أما في المرحلة الثانية فقد أصبح معامل الارتباط (٠.٧٦٦) ويلاحظ هنا مدى التحسن في التقدير باتجاه القدرة الحقيقية، لكن في المرحلة الأخيرة التي احتوت فقرات متعددة

التدرج بالإضافة للفقرات الثنائية السابقة أصبح معامل الارتباط (٠.٩٣٠) ويتباين مفسر مقداره (٠.٨٦٥) الأمر الذي يبين مدى اقتراب تقدير القدرة باستخدام الاختبار التكميلي الذي احتوى فقط (١٢) فقرة من تقدير القدرة باستخدام الاختبار الخطي بفقراته السبعين، علماً بأن معاملات الارتباط المذكورة هي دالة عند مستوى  $(\alpha = 0.01)$ .

وبالنسبة للخطأ المعياري في التقدير يُظهر الجدول أن متوسط الخطأ المعياري في التقدير كان في المرحلة الأولى الاستطلاعية (١.١٨٢) ثم أصبح في المرحلة الثانية (٠.٧٥٨) ويلاحظ هنا مقدار الانخفاض في الخطأ المعياري في التقدير الذي اعتمد حتى الآن على فقرات ثنائية التدرج، لكن في المرحلة الثالثة التي احتوت على فقرات متعددة التدرج أصبح (٠.٥٣٣) وهو يبعد بمقدار (٠.٢٧٧) عن متوسط الخطأ المعياري في تقديرات القدرة في الاختبار الخطي والذي يبلغ (٠.٢٥٥) بينما كان يبعد بمقدار (٠.٥٠٤) في المرحلة الثانية، وبمقدار (٠.٩٢٨) ويلاحظ أن الخطأ المعياري بدأ بمقدار كبير في المرحلة الأولى الاستطلاعية إلا أنه وصل لمقدار معقول في المرحلة الثالثة حيث كانت نسبة متوسط الخطأ المعياري في المرحلة الثالثة إلى متوسط الخطأ المعياري في الاختبار الخطي تساوي تقريباً (١ : ٢) بينما كانت نسبة عدد فقرات الاختبار الخطي إلى عدد فقرات الاختبار التكميلي تقريباً (١ : ٦) وهذا يبين مقدار الخفض الكبير في الكلفة والجهد للفاحص

ذلك الذي يحتوي كلا النوعين من الفقرات ذات الإجابة المنتقاة، وذات الإجابة المنشأة في مختلف أنواع الاختبارات.

والمفحوص في الاختبار التكيفي مقارنة بالاختبار الخطي في مقابل التضحية بمقدار قليل جداً من الدقة في القياس، وهذا يقود إلى التوصية باعتماد القياس التكيفي بشكل عام مكان القياس الخطي، وخصوصاً

ملحق (١)

جدول يبين تقديرات القدرة المتحصلة لأفراد عينة القياس التكيفي من تطبيق القياس التكيفي في المراحل الثلاث والقدرة المقدرة من الاختبار الخطي والأخطاء المعيارية لكل تقدير -

المفحوص	مرحلة (١) فقرات ذات إجابة منتقاة		مرحلة (٢) فقرات ذات إجابة منتقاة + فقرات المرحلة السابقة		مرحلة (٣) فقرات ذات إجابة منشأة + فقرات المرحلتين السابقتين		الاختبار الخطي (٦٠) فقرات ذات إجابة منتقاة + ٩٠ فقرات ذات إجابة منشأة	
	$SEE_1$	$(\Theta_1)$	$SEE_2$	$(\Theta_2)$	$SEE_3$	$(\Theta_3)$	$SEE_L$	$(\Theta_L)$
١	١.٠٥٢	٠.٤٦٦-	٠.٦٩١	٠.٤٥٤-	٠.٤٦٢	٠.١١٨-	٠.٢١٣	٠.٠٠٨
٢	١.٠٥٢	٢.٦٠١-	٠.٧٣٤	١.٠٧٩-	٠.٤٧٦	٠.٦٤٨-	٠.٢٢٢	٠.٣٨٣
٣	١.٠٥٢	٠.٤٦٦-	٠.٩٦٣	٠.٤١١	٠.٥٣٣	٠.٨٦٥	٠.٢٦	١.٢٨٧
٤	١.٠٥٥	٠.٥٣	٠.٦٨٢	٠.٢٨٨	٠.٥١٧	٠.٤٩٣	٠.٢٢٧	٠.٥٣٣
٥	١.٠٥٥	٢.٦٠١-	٠.٩٧٥	٣.٣١٥-	٠.٥٥٣	٢.٢٠٧-	٠.٢٤٤	١.٥-
٦	١.١٦٩	١.٥١٩-	٠.٨٠٦	١.٦١٧-	٠.٥٠١	١.٣١٣-	٠.٢١٧	٠.٨٨-
٧	١.٠٥٥	٠.٥٣	٠.٦٩٨	٠.١٥-	٠.٤٧٥	٠.٠٢١	٠.٢٢٣	٠.٤٣٢
٨	١.١٦٩	١.٥١٩-	٠.٩٧٥	٢.٣١١-	٠.٥٦٩	٢.٠٣٣-	٠.٢٣٤	١.٣٣-
٩	١.١٧٨	١.٦٠١	٠.٦٨٩	٠.٥١٥	٠.٤٩٨	٠.٠٦٣-	٠.٢١٨	٠.٢٣٩
١٠	١.٠٥٥	٢.٦٠١-	٠.٩٧٥	٣.٣١٥-	٠.٧٠٨	٢.٩٤٨-	٠.٢٥٨	١.٦٨٦-
١١	١.٥٣٩	٢.٧١١	٠.٨٠٣	٢.٣٨٦	٠.٥٥٥	١.٦٩٣	٠.٢٤١	٠.٩١٣
١٢	١.٠٥٥	٠.٥٣	٠.٧٤	٠.٦٢٨-	٠.٤٩٨	٠.٩٩٧-	٠.٢١	٠.٤٣٣-
١٣	١.١٧٨	١.٦٠١	٠.٧٩٥	١.٩٢٢	٠.٥٤٥	١.٤٣١	٠.٢٨٣	١.٦٥١
١٤	١.٠٥٢	٠.٤٦٦-	٠.٦٩١	٠.٤٥٤-	٠.٤٧٣	٠.٨٨٩-	٠.٢١٦	٠.٨٣٤-
١٥	١.٠٥٢	٠.٤٦٦-	٠.٧٢٧	٠.٩١٤-	٠.٤٧١	٠.٢١٩-	٠.٢١١	٠.٢١٤-
١٦	١.١٦٩	١.٥١٩-	٠.٧٣٢	١.٠٨١-	٠.٤٦١	٠.٦٨٦-	٠.٢٣٧	١.٣٨٥-
١٧	١.١٦٩	١.٥١٩-	٠.٧٣٢	١.٠٨١-	٠.٤٦١	٠.٦٨٦-	٠.٢٢٥	١.١٢١-
١٨	١.١٦٩	١.٥١٩-	٠.٨٠٦	١.٦١٧-	٠.٥٣٤	١.٥٦٢-	٠.٢١٢	٠.٦٥٣-
١٩	١.١٦٩	١.٥١٩-	٠.٦٩٩	٠.٦١-	٠.٤٥٤	٠.٥٦١-	٠.٢٢١	١.٠٢٣-
٢٠	١.٠٥٢	٠.٤٦٦-	٠.٤٥٤-	٠.٦٩١	٠.٤٥٤	٠.٣١١-	٠.٢١٤	٠.٠٥٤-
٢١	١.١٧٨	١.٦٠١	٠.٩٥٢	١.٤٤٧-	٠.٤٩٥	٠.١٧٤	٠.٢١٧	٠.١٩٢
٢٢	١.٠٥٢	٠.٤٦٦-	٠.٨٠٤	١.٤٤٧-	٠.٥٨٤	١.٧٨-	٠.٢١٣	٠.٦٩٨-
٢٣	١.٠٥٥	٠.٥٣	٠.٧١٩	٠.٢٨٦	٠.٤٩٨	٠.٥٢٧	٠.٢٣٧	٠.٧٩٩
٢٤	١.١٦٩	١.٥١٩-	٠.٧٠٥	٠.٢٨٦	٠.٥٣٥	٠.٥١١	٠.٢١٥	٠.٠٩٩
٢٥	١.٠٥٢	٠.٤٦٦-	٠.٦٨	٠.٠٢٣-	٠.٤٧٨	٠.٠٩	٠.٢١١	٠.٥٦٤-
٢٦	١.٠٥٥	٠.٥٣	٠.١٥-	٠.٦٩٨	٠.٥١٧	٠.٤٩٣	٠.٢١٩	٠.٢٨٦
٢٧	١.٠٥٥	٠.٥٣	٠.١٥-	٠.٦٩٨	٠.٤٧٥	٠.٠٢١	٠.٢٢٥	٠.٤٨٢

تابع ملحق رقم (١).

الاختبار الحظي (٦٠ فقرة ذات إجابة متقاة + ٩٠ فقرات ذات إجابة منشأة)		مرحلة (٣) فقرات ذات إجابة منشأة + فقرات المرحلتين السابقتين		مرحلة (٢) فقرات ذات إجابة متقاة + فقرات المرحلة السابقة		مرحلة (١) فقرات ذات إجابة متقاة		المفحوص
SEE <sub>L</sub>	(Θ <sub>L</sub> )	SEE <sub>3</sub>	(Θ <sub>3</sub> )	SEE <sub>2</sub>	(Θ <sub>2</sub> )	SEE <sub>1</sub>	(Θ <sub>1</sub> )	
٠.٢٢٣	١.٠٧٢-	٠.٤٧٨	١.٤٩١-	٠.٧٥١	٢.٠٧-	١.٥٢٥	٢.٦٠١-	٢٨
٠.٢١	٠.٣٤٥-	٠.٤٧١	٠.٢١٩-	٠.٧٢٧	٠.٩١٤-	١.٠٥٢	٠.٤٦٦-	٢٩
٠.٢٢	٠.٩٧٥-	٠.٤٦١	٠.٦٢١-	٠.٧٢٧	٠.٩١٤-	١.٠٥٢	٠.٤٦٦-	٣٠
٠.٢٥٣	١.١٥٦	٠.٥٧٢	١.٣١٦	٠.٧٢٣	١.١٧٤	١.٠٥٥	٠.٥٣	٣١
٠.٢١٦	٠.١٤٥	٠.٥٠٥	٠.٦٤٨-	٠.٦٨٩	٠.٥١٥	١.١٧٨	١.٦٠١	٣٢
٠.٢٣٢	٠.٦٩	٠.٤٧٨	٠.٠٩	٠.٦٨	٠.٠٢٣-	١.٠٥٢	٠.٤٦٦-	٣٣
٠.٢١٨	٠.٩٢٧-	٠.٤٧٧	١.١-	٠.٩٧٥	٢.٣١١-	١.١٦٩	١.٥١٩-	٣٤
٠.٢١٤	٠.٧٤٣-	٠.٤٦٦	٠.٨٨٥-	٠.٧٣٢	١.٠٨١-	١.١٦٩	١.٥١٩-	٣٥
٠.٢٤٤	٠.٩٧١	٠.٥١٣	٠.٧٦٨	٠.٦٨٩	٠.٧١٩	١.٠٥٥	٠.٥٣	٣٦
٠.٢١	٠.٤٣٣-	٠.٤٧٥	٠.٠٢١	٠.٦٩٨	٠.١٥-	١.٠٥٥	٠.٥٣	٣٧
٠.٢٢٧	١.١٧٢-	٠.٤٧٧	١.٠٧٦-	٠.٧٥١	٢.٠٧-	١.٥٢٥	٢.٦٠١-	٣٨
٠.٢١٧	٠.٨٨-	٠.٤٧٨	١.٠٦٧-	٠.٧٣٤	١.٠٧٩-	١.٥٢٥	٢.٦٠١-	٣٩
٠.٢٨٩	١.٧٣٣	٠.٥٧٩	١.٤٦	٠.٦٩٢	٠.٩٥٢	١.١٧٨	١.٦٠١	٤٠
٠.٢١٣	٠.٠٠٨	٠.٤٦٢	٠.١٨-	٠.٦٩٩	٠.٦١-	١.١٦٩	١.٥١٩-	٤١
٠.٢٤٧	١.٠٣١	٠.٥٠٥	٠.٦٤٨-	٠.٦٩٢	٠.٩٥٢	١.١٧٨	١.٦٠١	٤٢
٠.٢١١	٠.٥٢-	٠.٤٩٧	١.١٠٧-	٠.٦٩١	٠.٤٥٤-	١.٠٥٢	٠.٤٦٦-	٤٣
٠.٢٣٩	٠.٨٥٥	٠.٥٣٦	١.٠٢٧	٠.٦٨٩	٠.٧١٩	١.٠٥٥	٠.٥٣	٤٤
٠.٢٣	٠.٦٣٧	٠.٤٦٢	٠.١٨-	٠.٦٩٩	٠.٦١-	١.١٦٩	١.٥١٩-	٤٥
٠.٢١١	٠.١٢٦-	٠.٤٦٢	٠.٤٨٨-	٠.٧٣٢	١.٠٨١-	١.١٦٩	١.٥١٩-	٤٦
٠.٢١٢	٠.٠٣٧-	٠.٤٥٩	٠.٦٨٩-	٠.٦٨	٠.٠٢٣-	١.٠٥٢	٠.٤٦٦-	٤٧
٠.٢٩	٢.٠٥٤-	٠.٦١٢	٢.٥٣-	٠.٨١٥	٢.٦١٩-	١.٥٢٥	٢.٦٠١-	٤٨
٠.٢٤٩	١.٥٦-	٠.٤٦٥	٠.٨٢-	٠.٧٢٧	٠.٩١٤-	١.٠٥٢	٠.٤٦٦-	٤٩
٠.٢٥	١.٠٩٣	٠.٥٣٦	٠.٥٨١	٠.٦٨	٠.٠٢٣-	١.٠٥٢	٠.٤٦٦-	٥٠
٠.٢١٢	٠.٠٨١-	٠.٥٣٦	١.٠٢٧	٠.٧٢٣	١.١٧٤	١.٠٥٥	٠.٥٣	٥١
٠.٢١	٠.٣٤٥-	٠.٤٨٦	٠.٣١٣-	٠.٦٩٣	٠.٤١١	١.٠٥٢	٠.٤٦٦-	٥٢
٠.٢٥٣	١.٦٢٢-	٠.٥٠١	١.٣١٣-	٠.٨٠٦	١.٦١٧-	١.١٦٩	١.٥١٩-	٥٣
٠.٢٦٤	١.٣٥٥	٠.٦٢١	٢.٠٦	٠.٧٩٥	١.٩٢٢	١.١٧٨	١.٦٠١	٥٤
٠.٢١٢	٠.٦٠٨-	٠.٤٦١	٠.٦٢١-	٠.٧٢٧	٠.٩١٤-	١.٠٥٢	٠.٤٦٦-	٥٥
٠.٢٢٥	٠.٤٨٢	٠.٤٩٣	٠.٢٤٥	٠.٦٨٢	٠.٢٨٨	١.٠٥٥	٠.٥٣	٥٦
٠.٢١١	٠.١٧-	٠.٤٩٥	٠.١٧٤	٠.٦٩٢	٠.٩٥٢	١.١٧٨	١.٦٠١	٥٧

تابع ملحق رقم (١).

الاختبار الخطي (٦٠ فقرات ذات إجابة متقاة + ٩٠ فقرات ذات إجابة منشأة)		مرحلة (٣) فقرات ذات إجابة منشأة + فقرات المرحلتين السابقتين		مرحلة (٢) فقرات ذات إجابة متقاة + فقرات المرحلة السابقة		مرحلة (١) فقرات ذات إجابة متقاة		المفحوص
SEE <sub>L</sub>	(Θ <sub>L</sub> )	SEE <sub>3</sub>	(Θ <sub>3</sub> )	SEE <sub>2</sub>	(Θ <sub>2</sub> )	SEE <sub>1</sub>	(Θ <sub>1</sub> )	
٠.٢٤٧	١.٠٣١	٠.٥٤٤	١.١٤١	٠.٧٠٩	١.٣٩	١.٥٣٩	٢.٧١١	٥٨
٠.٢١	٠.٣٠٢-	٠.٥١١	٠.٦٣٨-	٠.٦٨٩	٠.٧١٩	١.٠٥٥	٠.٥٣	٥٩
٠.٢١٩	٠.٢٨٦	٠.٤٩٣	٠.٢٤٥	٠.٦٩٨	٠.١٥-	١.٠٥٥	٠.٥٣	٦٠
٠.٢٨٩	١.٧٣٣	٠.٥٣٣	٠.٨٦٥	٠.٦٩٣	٠.٤١١	١.٠٥٢	٠.٤٦٦-	٦١
٠.٣٣٤	٢.٤٣٦-	٠.٥٥٣	٢.٢٠٧-	٠.٨١٥	٢.٦١٩-	١.٥٢٥	٢.٦٠١-	٦٢
٠.٣٤٩	٢.٥٥١-	٠.٦١٢	٢.٥٣-	٠.٨١٥	٢.٦١٩-	١.٥٢٥	٢.٦٠١-	٦٣
٠.٢١	٠.٢٥٨-	٠.٤٦٢	٠.١١٨-	٠.٦٩١	٠.٤٥٤-	١.٠٥٢	٠.٤٦٦-	٦٤
٠.٢٧٥	١.٨٩٦-	٠.٥٦٩	٢.٠٣٣-	٠.٩٧٥	٢.٣١١-	١.١٦٩	١.٥١٩-	٦٥
٠.٣٧٥	٢.٥٧١	٠.٧٠٢	٢.٤٥٩	٠.٩٥٨	٢.٥٩٧	١.١٧٨	١.٦٠١	٦٦
٠.٢٢٩	١.٢٢٣-	٠.٤٦٥	٠.٨٢-	٠.٧٢٧	٠.٩١٤-	١.٠٥٢	٠.٤٦٦-	٦٧
٠.٣١٢	٢	٠.٥٢	٠.٨٩٥	٠.٦٩٢	٠.٩٥٢	١.١٧٨	١.٦٠١	٦٨
٠.٢٢٨	٠.٥٨٤	٠.٥٣٥	٠.٥١١	٠.٦٩١	٠.١٦٥-	١.١٦٩	١.٥١٩-	٦٩
٠.٢٤	١.٤٤٢-	٠.٥٨٤	١.٧٨-	٠.٧٢٧	٠.٩١٤-	١.٠٥٢	٠.٤٦٦-	٧٠
٠.٣٣٢	٢.٢٠٤	٠.٥٧٩	١.٤٦	٠.٦٩٢	٠.٩٥٢	١.١٧٨	١.٦٠١	٧١
٠.٢١	٠.٤٧٦-	٠.٥٣٥	٠.٥١١	٠.٧٠٥	٠.٢٨٦	١.١٦٩	١.٥١٩-	٧٢
٠.٢٣٤	٠.٧٤٤	٠.٤٩٨	٠.٥٢٧	٠.٦٨٩	٠.٧١٩	١.٠٥٥	٠.٥٣	٧٣
٠.٢٨٢	١.٩٧٣-	٠.٥٥٣	٢.٢٠٧-	٠.٨١٥	٢.٦١٩-	١.٥٢٥	٢.٦٠١-	٧٤
٠.٢٩٦	١.٨١٨	٠.٥٤٥	١.٤٣١	٠.٧٩٥	١.٩٢٢	١.١٧٨	١.٦٠١	٧٥
٠.٢٧٨	١.٥٧٣	٠.٥٤٥	١.٤٣١	٠.٧٩٥	١.٩٢٢	١.١٧٨	١.٦٠١	٧٦
٠.٣٨٩	٢.٨٢١-	٠.٨٣	٢.٨٩-	٠.٩٧٥	٢.٣١١-	١.١٦٩	١.٥١٩-	٧٧
٠.٢١	٠.٣٨٩-	٠.٥١٣	٠.٧٦٨	٠.٦٨٩	٠.٧١٩	١.٠٥٥	٠.٥٣	٧٨
٠.٢١	٠.٣٠٢-	٠.٤٦٤	٠.١٨٥-	٠.٦٨٢	٠.٢٨٨	١.٠٥٥	٠.٥٣	٧٩
٠.٣٢١	٢.٠٩٩	٠.٨٨٨	٣.٣٣٣	٠.٩٦٤	٣.٠٦٢	١.٥٣٩	٢.٧١١	٨٠
٠.٣٥٨	٢.٤٣٩	٠.٦٣١	٢.٣٣٦	٠.٩٦٤	٣.٠٦٢	١.٥٣٩	٢.٧١١	٨١
٠.٢٦٣	١.٧٥٣-	٠.٥٣٤	١.٥٦٢-	٠.٨٠٦	١.٦١٧-	١.١٦٩	١.٥١٩-	٨٢
٠.٢٥٧	١.٢٢	٠.٥٧٩	١.٤٦	٠.٦٩٢	٠.٩٥٢	١.١٧٨	١.٦٠١	٨٣
٠.٢٣٧	١.٣٨٥-	٠.٥٣١	١.٤٩-	٠.٨٠٤	١.٤٤٧-	١.٠٥٢	٠.٤٦٦-	٨٤
٠.٣٩٥	٢.٧١٨	٠.٥٧٩	١.٤٦	٠.٦٩٢	٠.٩٥٢	١.١٧٨	١.٦٠١	٨٥
٠.٢٣٢	١.٢٧٦-	٠.٥٣٤	١.٥٦٢-	٠.٧٣٢	١.٠٨١-	١.١٦٩	١.٥١٩-	٨٦
٠.٣٠٣	١.٩٠٧	٠.٥٧٢	١.٣١٦	٠.٧٢٣	١.١٧٤	١.٠٥٥	٠.٥٣	٨٧

تابع ملحق رقم (١).

المفحوص		مرحلة (١) فقرات ذات إجابة متقاة		مرحلة (٢) فقرات ذات إجابة متقاة + فقرات المرحلة السابقة		مرحلة (٣) فقرات ذات إجابة منشأة + فقرات المرحلتين السابقتين		الاختبار الخطي (٦٠) فقرة ذات إجابة متقاة + ١٠ فقرات ذات إجابة منشأة	
		$SEE_1$	$(\Theta_1)$	$SEE_2$	$(\Theta_2)$	$SEE_3$	$(\Theta_3)$	$SEE_L$	$(\Theta_L)$
٨٨		١.١٧٨	١.٦٠١	٠.٧٩٥	١.٩٢٢	٠.٥٧٤	١.٧٢٨	٠.٢٧٣	١.٤٩٨
٨٩		١.١٦٩	١.٥١٩-	٠.٩٧٥	٢.٣١١-	٠.٥٦٩	٢.٠٣٣-	٠.٣٦٨	٢.٦٧٩-
٩٠		١.٠٥٢	٠.٤٦٦-	٠.٤١١	٠.٤١١	٠.٤٨٦	٠.٣١٣-	٠.٢١٥	٠.٧٨٨-
٩١		١.٥٢٥	٢.٦٠١-	٠.٨١٥	٢.٦١٩-	٠.٥٥٣	٢.٢٠٧-	٠.٣٠٩	٢.٢٣١-
٩٢		١.٠٥٢	٠.٤٦٦-	٠.٨٠٤	١.٤٤٧-	٠.٤٩٩	١.٢٤٤-	٠.٢٦٩	١.٨٢٣-
٩٣		١.١٧٨	١.٦٠١	٠.٧٢٢	١.٤٠٦	٠.٥٢	٠.٨٩٥	٠.٢٦٨	١.٤٢٥
٩٤		١.١٧٨	١.٦٠١	٠.٩٥٨	٢.٥٩٧	٠.٧٠٢	٢.٤٥٩	٠.٤٩٧	٣.٢٩٢
٩٥		١.٥٢٥	٢.٦٠١-	٠.٧٥١	٢.٠٧-	٠.٤٩١	١.٧٠٦-	٠.٢٩	٢.٠٥٤-
٩٦		١.١٧٨	١.٦٠١	٠.٧٢٢	١.٤٠٦	٠.٥٢	٠.٨٩٥	٠.٢٣٢	٠.٦٩
٩٧		١.٥٢٥	٢.٦٠١-	٠.٩٧٥	٣.٣١٥-	٠.٥٥٣	٢.٢٠٧-	٠.٣٢١	٢.٣٢٩-
٩٨		١.١٦٩	١.٥١٩-	٠.٩٧٥	٢.٣١١-	٠.٨٣	٢.٨٩-	٠.٤١٦	٢.٩٨٢-
٩٩		١.٠٥٢	٠.٤٦٦-	٠.٦٨	٠.٠٢٣-	٠.٤٦٢	٠.١١٨-	٠.٢١٧	٠.١٩٢
١٠٠		١.٠٥٥	٠.٥٣	٠.٨	١.٦٩٦	٠.٧٠٣	٢.٣٤	٠.٤٢١	٢.٨٨٢



## المراجع

### أولاً: المراجع العربية:

البرصان، إسماعيل. أثر عدد مراحل القياس التكيفي وعدد أسئلة كل مرحلة في تقدير القدرة والخطأ المعياري في التقدير باستخدام أسئلة ثنائية ومتعددة التدرج. رسالة دكتوراه غير منشورة، جامعة عمان العربية للدراسات العليا، ٢٠٠٦م.

الثوابية، أحمد. فاعلية القياس التكيفي في تقويم التحصيل في مبحث الأحياء للصف الثاني الثانوي العلمي في الأردن. رسالة دكتوراه غير منشورة، جامعة عمان العربية للدراسات العليا، ٢٠٠٤م.

عبيدات، عمر. فاعلية الاختبار التكيفي المحوسب في تقدير القدرة العقلية باستخدام مصفوفات رافن. رسالة دكتوراه غير منشورة، جامعة اليرموك، ٢٠٠٨م.

العموش، جميل. فاعلية القياس التكيفي في تقويم بعض القدرات المعرفية لدى طلبة السنة الأولى الجامعية. رسالة دكتوراه غير منشورة، جامعة عمان العربية للدراسات العليا، ٢٠٠٣م.

### ثانياً: المراجع الأجنبية:

Anastasi, A. Urbina, S. Psychological Testing. 7<sup>th</sup> Ed, New York: Prentice Hall, 1997.  
Brengstrom, B.A., Lunz, M.E., Gershon, R.C.

Altering the Level of Difficulty in Computer Adaptive Testing: Applied Measurement in Education, 5(2), (1992) 137 – 149.

Boyd, A. M., Dodd, B. G., & Choi, S. W. Polytomous models in computerized adaptive testing. In M. L. Nering & R. Ostini (Eds), Handbook of polytomous item response theory models. New York, NY: Routledge, 2010.

Campbell, J. Cognitive process elicited by multiple – choice and constructed response questions on an assessment of reading comprehension. Dissertation Abstract International. 95001 – 128, (2000). <http://erica.net/eda/ed315425.htm>.

Choi, W. Seung, Grady, W. Matthew, Dodd, G. Barbra. A New Stopping Rule for Computerized Adaptive Testing. Educational and Psychological Measurement. 2001171:37, (2011) <http://epm.sagepub.com/content/71/1/37>

Crist, S. Computerized Adaptive tests: Eric Digest No. 107, Eric database, 1989. <http://erica.net/eda/ed315425.htm>.

Embreston, S. E. Riese, S. P. Item Response Theory for Psychologists. Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers, 2000.

Fliege, H., Backer, J., Walter, O., Bjorner, J., Klapp, B., & Rose, M. Development of a computer – adaptive test for depression (D – CAT). Quality of life research, 14, (2005), 2277 – 2291.

Ghiselli, E., Campbell, J., & Zedeck, S. Measurement Theory for the Behavioral Science. San Francisco: Freeman company, 1981.

Hambleton, R., Swaminathan, H. Item Response Theory. Principles and Application. Boston: Kluwer – Nighoff Publishing, 1985.

Hambleton, R., Swaminathan, H., and Rogers, H. Fundamental of Item Response Theory. New York: Sage Publication the International Professional

- tool in general practice: a pilot study. Medical teacher, 26(2), (2004), 178 – 183.
- Sands, W., Waters, B., and Mc Bride, J.** computerized Adaptive testing: From Inquiry to Operation. American Psychological Association Washington, 2001.
- Vispoel, W., Wang, T., & Blieler, T.** Computerized adaptive and fixed – item testing of music listening skill: A comparison of efficiency, precision and concurrent validity. Journal of Educational measurement, 34, (1997), 34 – 63.
- Ward, W.** Using micro computers to administer measurement. Issues and practices, 3, (1984), 16 – 20.
- Wainer, H., Dorans, N. J., Flaunter, R., Jreen, B. F., Mislevy, R. J., Steinberg, L., and Thissen, D.** Computerized Adaptive Testing: A primer. Second Edition Hillsdale, NJ: Lawrence Erlbaum, 2000.
- Warm, A.** A primer of Item Response Theory: US. Coast Guard Institute Oklahoma, (1978), 73/69.
- Wright, B. D., Stone, M.H.** Best Test Design: Rash Measurement, Chicago. MESA Press, 1979.
- Publishers, 1991.
- Hendrickson, A.** Scaling of Two – Stage Adaptive Test Configurations For Achievement Testing. Unpublished Doctoral Dissertation. University of Iowa, 2002.
- Hulin, C., Drasgow, F., & Parsons, C.** Item response theory: Application to Psychological Measurement. Dow Jones Irwin Home Wood, 1983.
- Kim, H., Plake, B.** Monte Carlo Simulation Comparison of Two – stage Testing and Computerized Adaptive Testing, ERIC: ED 357041, 1993.
- Kingsbury, G.G., & Zara, A.R.** Procedures for Selecting Items for Computerized Adaptive Testing Applied Measurement in Education, 9, (1989), 287 – 304.
- Lincare, J.** computer – Adaptive Testing: A methodology whose Time Has Come. Seol: Komesa Press, 2000.
- Linden, W.J & Pashley, P.J.** Item Selection and Ability Estimation in Adaptive Testing. In Linden, 2002.
- Linden, W.J., & Glass, C.A (eds).** Computerized Adaptive Testing: Theory and practice. Kluwer Academic Publishers, 2000.
- Lord, F.** Application of Item Response Theory to Practical Testing Problems. New Jersey: Lawrence Erlbaum Associates Publishers, 1980.
- Masters. N.G.** A Rasch Model for Partial Credit Scoring. Psychometrika. Vol. 47. No.(2), (1982).
- Mills, C., & Stocking, M.** Practical Issues in Large Scale Computerized Adaptive Testing. Applied Measurement in Education, 9(4), (1996), 287 – 304.
- Murphy, K., & Davidshofer, C.O.** Psychological Testing: Principles and Applications, 3<sup>rd</sup> ed., New Jersey: Prentice Hall. New Jersey: Lawrence Erlbaum Associates, Inc, 1994.
- Roex, A. & Degryse, J.** A Computerized adaptive knowledge test as an assessment

## The Effectiveness of the Adaptive Testing Using Selected Response Items and Constructed Response Items

Ismail Al-Bursan

Assistant Professor, Department of Psychology,  
College of Education, King Saud University  
Riyadh, Kingdom of Saudi Arabia, p.o box: 2458, Postal Code: 11451  
E-mail: ibursan@ksu.edu.sa  
(Received 17/10/1431H; accepted for publication 1/6/1432H.)

**Key Words:** Item response theory, adaptive testing, modern test theory, item response models.

**Abstract:** This study aimed at investigate the effectiveness of the adaptive testing which consists of both dichotomous selected response items and polytomous items by using generated data from computer, 60 dichotomous responses and 10 polytomous responses were generated for 2000 examinees. These 70 items calibrated by RUMM 2020. the sample of adaptive testing was 100 examinee from the calibration sample, their data used in the adaptive testing which consisted of three stages, the first stage was the routing test consisted of 5 dichotomous selected response items, the second stage also consisted of 5 dichotomous selected response items, while the third consisted of 2 polytomous constructed response. The ability of each examinee was calculated for the three stages. The results indicated that the correlation coefficient between the ability which estimated from adaptive testing and the ability estimated from the linear test was 0.93, and the average of differences between the two abilities equal 0.087 logit. the average of the differences between the related two standards error of estimation equal 0.287. this indicated that the adaptive testing using both dichotomous and polytomous is highly effective, especially when the outcomes requested the two types of items.