مجلة جامعة الملك سعود، م24، العلوم التربوية والدراسات الإسلامية (4)، ص ص 1401-1426، الرياض (2012م/1433هـ)

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 2458، الرمز 11451

E – mail: ibursan@ksu.edu.sa

(قدم للنشر في 1431/10/17هـ؛ وقبل للنشر في 1432/6/1هـ)

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 التدريج وفقرات ذات إجابة منشأة متعددة التدريج في نفس الاختبار باستخدام بيانات مولدة من الحاسوب، حيث تم توليد إجابات (60) فقرة ثنائية التدريج و(10) فقرات متعددة التدريج لألفي مفحوص،جرى تدريجها معاً باستخدام برمجية (RUMM 2020)، واختيرت عينة للقياس التكيفي بلغت (100) مفحوص جرى استخدام بيانات استجاباتهم في القياس التكيفي الذي تكون من ثلاث مراحل تكونت المرحلة الاستطلاعية الأولى من خمس فقرات ذات إجابة منتقاة ثنائية التدريج، والمرحلة الثانية من خمس فقرات ذات إجابة منتقاة ثنائية التدريج، في حين تكونت المرحلة الثالثة والأخيرة من فقرتين ذاتي إجابة مُنشأة، 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 0.93، ومتوسط فرق في القدرة مقداره 0.087 لوجيت،إضافة إلى فرق في الخطأ المعياري مقداره في المتوسط 0.278 الأمر الذي يفيد بفاعلية عالية للقياس التكيفي المبني باستخدام فقرات ثنائية ومتعددة التدريج معاً خصوصاً عندما تتطلب النتاجات 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 من أن الإرهاصات الأولى له كفكرة قد بدأت عام

1905 في اختبارات بينية للذكاء.

ويعرف كل من مايلز وستوكنغ (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.

ويعرّفه هامبيلتون 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:

1- أحادية البعد (Unidimensionality)

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

2- 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 (θ)،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 χ^2 (Lord, 1980)، بينما يعتبر هاميلتون وسواميناثان أن هذا الافتراض يعتبر متحققاً ضمناً إذا ما تحقق الافتراض الأول (Hambleton & Swaminathan).

3- منحنى خصائص الفقرة (Item Characteristic Curve)

وهو منحنى يمثل العلاقة بين القدرة واحتمالية الإجابة الصحيحة على الفقرة، وهو منحنى متصاعد باضطراد وله مستوى تقاربي علوي (upper asymptotic – level) يقترب من 1، ومستوى تقاربي (Lower asymptotic level) أدنى يختلف باختلاف النموذج المستخدم في نظرية

الاستجابة للفقرة.

4- 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 كنجزبري وزارا (Kingsbury & Zara, 1989) ست خطوات لبناء الاختبار التكيفي هو:

1- بناء تجمع الفقرات (item pool): حيث تعد الفقرات ذات المعالم المقدرة جيداً أمراً ضرورياً للاختبار التكيفي، وترى 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.

2- 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 كالآتي:

أ – النموذج الأحادي المعلمة:

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

Pi (Θ): احتمال أن يجيب المفحوص ذو القدرة (Θ) على الفقرة (i) إجابة صحيحة.

D: ثابت قيمته 1,7.

bi: معامل الصعوبة للفقرة (i).

e: الأساس اللوغاريتمي الطبيعي (العدد النيبيري 2,7183).

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 Masters, 1982) والعلاقة التالية التي تعطي احتمالية وقوع المفحوص n ذي القدرة θ في المستوى X بدلاً من البقاء في المستوى (X – 1).

$$\Phi_{Xni} = \frac{\Pi_{Xin}}{\Pi_{(X-1)ni} + \Pi_{Xni}} = \frac{e^{(\theta_n - bi X)}}{1+e^{(\theta_n - bi X)}}$$

حيث:

Φ_{xni}: احتمال إنجاز المفحوص (n) الخطوة × في السؤال i ليأخذ العلامة × بدلاً من العلامة (1 – ×).

Π_{xin}: احتمال وقوع المفحوص (n) ذي القدرة θ في المستوى × من مستويات الأداء للفقرة i

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

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

3- 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 قدرة يختاره المفحوص وهو ما يسمى بـ (Self adaptive testing).

4- 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.

5- 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 التكيفية باختلاف الاستراتيجيات التي تتبعها، وقد قسم لورد (Lord, 1980) استراتيجيات القياس التكيفي إلى ثلاث استراتيجيات هي:

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

وتتكون هذه الاستراتيجية من مرحلتين أولاهما 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 التكيفي مجتمعتين (Hambleton & Swaminathan, 1985).

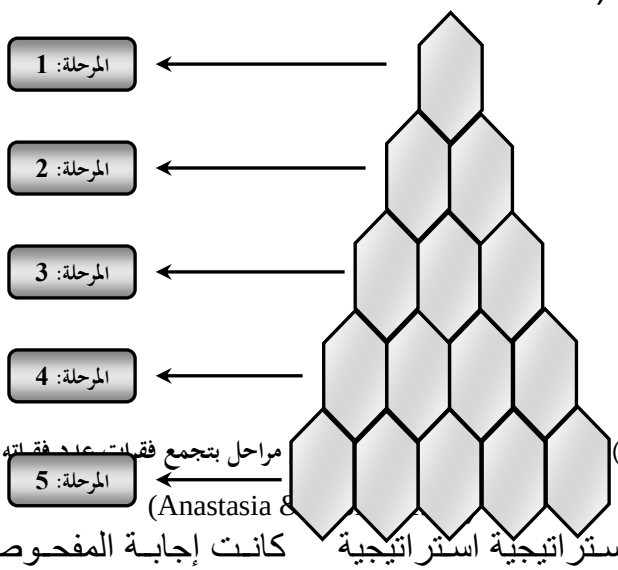
والإجراء العملي لاختيار فقرات الاختبار الاستطلاعي هو اختيار فقراته من بين فقرات الاختبارات الفرعية الخاصة بالمرحلة الثانية، فبعد تدريج تجمع الفقرات وتقسيمها إلى اختبارات فرعية ذات صعوبة متدرجة ومختلفة يتم انتقاء عدد قليل من الفقرات لتكوين الاختبار الاستطلاعي الذي يتم بواسطته تحديد مستوى قدرة ابتدائي 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 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 (1) الآتي يبين طريقة تنظيم الفقرات.



الشكل (1) مراحل بتجمع فقرات عند فقراته 15 فقرة.

وتتضمن هذه الاستراتيجية استراتيجية الاختبارات التكيفية الطبقية حيث يتم تصنيف تجميع 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:

1- التوقف عند استنفاد الفقرات الموجودة في تجمع فقرات بنك الأسئلة.

2- التوقف عندما تصبح قيمة الخطأ المعياري في التقدير أقل من 0.0625

3- 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 هامبيلتون وسواميناثان أن الاختبار التكيفي يستخدم فقط 10% – 50% من فقرات الاختبار التقليدي (Hambleton & Swaminathan, 1985)، بينما ذكر وورم (Ward, 1984) أن الاختبار التكيفي يحتاج إلى فقرات تقل بـ (50% – 60%) عن الاختبار التقليدي (الخطي) أما لندن وباشلي (Linden & Pashley, 2002) فيذكران أن الاختبار التكيفي يخفض عدد الفقرات التي تطبق بمقدار 90,8%.

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, 1991), (Hulin, Drasgow & Parsons, 1983) 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 يُغفل الباحثون محددات استخدام القياس التكيفي فقد ذكر كرست (Crist, 1989) المحددات الآتية:

1- نقص الأجهزة والمحددات والبرمجيات.

2- ضرورة إتقان المفحوصين لمهارات استخدام الحاسوب.

3- تتطلب الاختبارات التكيفية دقة عالية في عملية معايرة الفقرات أو تدريجها (Item Calibration)

ويضيف هامبيلتون وزميله (Hambleton & Swaminathan, 1985) ما يلي:

1- الكلفة المادية العالية من حيث الإعداد والتحضير والتجهيز والبرمجيات.

2- صعوبة الإعداد والتحضير بسبب استخدام عدد كبير من الفقرات لتكوين تجمع الفقرات.

3- بعض الإشكالات القانونية من قبل المحاكم والقضاء عند عرض نتائج الاختبارات التكيفية في بعض المجتمعات مثل المجتمع الأمريكي.

وقد أجرى الثوابية (2004) 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 (55) فقرة ثنائية التدريج من نوع الاختيار من متعدد، وطبق هذا الاختبار على عينة تدريج بلغت (721) طالباً وطالبة واستخدم البرنامج الإحصائي (winsteps) في تدريج الفقرات، وقد اشتق بعد ذلك ستة اختبارات فرعية أحدها الاختبار الاستطلاعي الذي يتم توجيه المفحوص بناء عليه إلى أحد الاختبارات

الفرعية الخمسة الأخرى التي تشكل المرحلة الثانية من الاختبار التكيفي ثم طبق القياس التكيفي على عينة خاصة بلغت (81) طالباً وطالبة وقد كان معامل الارتباط بين القدرة المقدرة من خلال الاختبار الاستطلاعي والعلامة المدرسية (0.75) وبينها وبين علامة الثانوية العامة (0.46)، أما معامل الارتباط بين الاختبار التكيفي ككل والعلامة المدرسية فكان (0.70)، ومع علامة الثانوية العامة (0.42)، بينما كان معامل الارتباط بين الاختبار الخطي ككل والعلامة المدرسية (0.76) وبينه وبين علامة الثانوية العامة (0.52).

وفي الدراسة التي أجراها العموش (2003) بعنوان فاعلية القياس التكيفي في تقويم بعض القدرات المعرفية لدى طلبة السنة الأولى الجامعية والتي هدفت لاستقصاء فاعلية القياس التكيفي في تقويم القدرة اللفظية والقدرة التقليدية وفق الطريقتين التقليدية والحديثة أظهرت النتائج لاختبار القدرة الرياضية تفوق القياس التكيفي 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 التكيفي أكثر فعالية في مستويات القدرة المتوسطة والمرتفعة، وبشكل عام أظهرت الدراسة أن القياس التكيفي كان أكثر فاعلية من الاختبار التقليدي اعتماداً على الكفاءة النسبية، واختزال الوقت والفقرات المستخدمة.

وفي دراسة أجراها البرصان (2006) هدفت إلى استقصاء أثر عدد مراحل القياس التكيفي وعدد أسئلة كل مرحلة في تقدير القدرة والخطأ المعياري في التقدير باستخدام فقرات ثنائية وفقرات متعددة التدريج، استخدم اختبارين تحصيليين في مادة الرياضيات تكون أحدهما من فقرات ثنائية التدريج، والآخر تكون من فقرات متعددة التدريج، طبقهما على عينتين من طلبة الصف الأول الثانوي في مدينة عمان، وتم تدريج كل من فقرات الاختبارين باستخدام برمجية (WINSTEPS)، طبق بعدها إستراتيجية القياس التكيفي متعدد المراحل،حيث حصل على خمسة تقديرات للقدرة لكل مفحوص في العينة الأولى التي طبق عليها القياس التكيفي باستخدام فقرات ثنائية التدريج، 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 التكيفي المكون من فقرات ثنائية التدريج حتى المرحلة الثالثة أما المرحلة الرابعة فلم تكن مؤثرة،أما في القياس التكيفي المعتمد على فقرات متعددة التدريج فكانت عملية الاقتراب من القدرة المرجعية والخطأ المعياري المرافق لها مجدية حتى الرحلة الثانية بينما لم تعد مجدية في المرحلتين الثالثة والرابعة، وقد أشارت النتائج كذلك إلى أن القياس التكيفي المبني على فقرات متعددة التدريج أدق من حيث تقدير القدرة والخطأ المعياري في التقدير من القياس التكيفي المعتمد على فقرات ثنائية التدريج.

وقد هدفت دراسة أجراها عبيدات (2008) إلى فحص فاعلية الاختبار التكيفي المحوسب في دقة تقدير القدرة العقلية باستخدام اختبار مصفوفات رافن، ولتحقيق ذلك استخدم

اختبار مصفوفات رافن حيث تم تدريج الفقرات التي بلغت في صورتها النهائية (105) فقرات على عينة تدريج بلغت (2695) طالباً وطالبة، ثم طبق القياس التكيفي المحوسب على عينة ثانية بلغت (638) طالباً وطالبة باستخدام محكّين لإنهاء الاختبار هما عدد محدد من الفقرات (30) فقرة، وأدنى خطأ معياري (0.25)، وتم ذلك باستخدام طريقتين لتقدير القدرة هما طريقة الأرجحية العظمى (Maximum Likelihood estimation) (MLE) وطريقة التقدير البعدي الأعظم (MAP) (Maximum a posterior)، 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 (50%) من قاعدة عدد محدد من الفقرات لكن إذا تمت المقارنة بين القياس التكيفي والخطي وصل نسبة الاختزال لعدد الفقرات المطبقة إلى (70%) تقريباً، وكذلك كان الاختبار التكيفي أدق في تقدير القدرة من الاختبار التكيفي باستخدام قاعدة أدنى خطأ معياري، وقد كانت نسبة اختزال عدد الفقرات المطبقة بين الاختبار التكيفي والاختبار الخطي تساوي 7%.

أما في الدراسة التي أجراها روكس وديكريس (Roex & Degryse, 2004) بعنوان «اختبار تكيفي محوسب للمعرفة كأداة تقييم في الممارسة العامة» والتي هدفت إلى تقويم جدوى تحويل اختبار مكتوب إلى اختبار تكيفي محوسب ودراسة صدق المحتوى لهذا الاختبار، تكون الاختبار من (165) فقرة تطبق في فترة زمنية مقدراها (3) ساعات، وقد تم تدريج الفقرات باستخدام عينة بلغت (1000) طالب وطالبة، وقد استخدم برنامج (Fast test pro) لتطبيق الاختبار التكيفي المحوسب، واستخدمت طريقة الأرجحية العظمى لتقدير القدرة وقاعدة أدنى خطأ معياري للتوقف ($SE \leq 0.25$) وقد تم تطبيق الاختبار التكيفي على (50) طالباً، وأظهرت النتائج أن متوسط مدة الاختبار كان (47) دقيقة بانحراف معياري (19.38)، وكان عدد الفقرات التي جرى تعريض المفحوصين لها يتراوح بين (20) فقرة و (70) فقرة وبمتوسط (40.24) فقرة وبانحراف معياري (12.40)، أما صدق الاختبار فقد كان مجدياً وقابلاً للتطبيق، إلا الاختبار التكيفي 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 تكيفي محوسب للاكتئاب»، حيث تم استخدام (320) فقرة غطت اختبارات فرعية (النشاطات اليومية، المزاج، الشكاوى، الشخصية، الرضا عن الحياة، النرجسية، الصحة، القلق، فاعلية الذات) وباستخدام الاختبار التكيفي المحوسب تبين أن علامة الاختبار التكيفي المحوسب ارتبطت مع فقرات الاكتئاب بمعامل ارتباط (0.95)، ومع اختبار القلق بمعامل ارتباط (0.79) ومع اختبار الاكتئاب لمركز الدراسات الوبائية معامل ارتباط (0.76) وبشكل عام تم الاستنتاج أن الاختبار التكيفي المحوسب للاكتئاب يقيس أعراض الاكتئاب بدقة عالية وبعدد متدن من 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 (2200) مفحوص، حيث أظهرت النتائج أن الاختبار التكيفي الموسيقي يحتاج لفقرات تقل بنسبة (50% – 93%) 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 ويليك (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 محكّيْ دقة القياس والفعالية، ووجد أن الاختبارات التكيفية ثنائية المرحلة التي تستخدم مدى واسعا 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 (2003) والبرصان (2006)، ودراسة فليج وزملائه (Fliege, Becker, Walter, Bjorner, Klapp & Rose, 2005)، ودراسة هندركسون (Hendrickson, 2002)، ودراسة كيم وبليك (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 الفقرات من نوع الإجابة المنتقاة ذات التدريج الثنائي خصوصاً في الاختبارات التكيفية المحوسبة (Computerized adaptive CAT testing)، لكن ومن وجهة نظر كثير من اختصاصي القياس والتقويم بأن الفقرات من نوع الإجابة المنتقاة ذات التدريج الثنائي لا يمكنها قياس نتاجات التعلم المعقدة،هذا عدا عن أن الفقرات ذات الإجابة المُنشأة 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 المُنشأة هو الأكثر ملاءمة، وبناءً عليه فإن أي اختبار إذا ما أريد له أن يقيس نتاجات التعلم على اختلاف أنواعها فلا بد من أن يحتوي على فقرات من نوع الإجابة المنشأة (المصوغة) (Constructed Response Items) ذات التدريج متعدد الفئات، ومن المعلوم أن هناك صعوبات ترتبط بتطبيق الفقرات ذات الإجابة متعددة التدريج 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 سواءً كانت محوسبة أو باستخدام الورقة والقلم.

أهداف الدراسة:

تهدف هذه الدراسة إلى اختبار فاعلية القياس التكيفي المبني باستخدام فقرات إجابة منتقاة ثنائية التدريج (Dichotomous Items) وفقرات من نوع الإجابة المنشأة متعددة التدريج (Polytomous Items) معاً وذلك باستخدام طريقة قابلة للتطبيق وهي جعل المرحلة الأخيرة فقرات متعددة التدريج من نوع الإجابة المنشأة (المصوغة)، يجري إضافتها لإجابات المفحوص السابقة وحساب قدرته المتحققة من جميع الفقرات الثنائية التدريج والمتعددة التدريج 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 التدريج من نوع الفقرات ذات الإجابة المنشأة متعددة التدريج إلى القياس التكيفي وذلك لأهمية هذا النوع من الفقرات في قياس بعض نتاجات التعلم التي يصعب قياسها بفقرات ثنائية التدريج من نوع الفقرات ذات الإجابة المنتقاة، وبشكل هذا العمل داعماً لفكرة القياس التكيفي الذي أداته 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.

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

أسئلة الدراسة:

يتمثل السؤال الرئيس للدراسة بالسؤال:

ما فاعلية إدخال فقرات ذات إجابة منشأة متعددة التدريج 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 التدريج، والثالثة المتكونة من فقرات متعددة التدريج والقدرة الحقيقية المقدرة من الاختبار الخطي الذي يحتوي فقرات ثنائية ومتعددة التدريج.

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

جـ) مقدار انخفاض الخطأ المعياري في التقدير أثناء التقدم في المراحل.

محددات الدراسة:

1- استخدمت هذه الدراسة بيانات مولدة باستخدام الحاسوب.

2- 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 التدريج) (dichotomous items):

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

الأسئلة ذات الإجابة المُنشأة (متعددة التدريج) (polytomous items):

أسئلة تتكون الإجابة فيها من متغير له تدريج ثلاثي أو رباعي أو أكثر بناءً على فئات الإجابة حيث تعطى المعرفة الجزئية علامات حيث تعطى الدرجة صفر للإجابة الخطأ، والدرجة (1) للإجابة الجزئية ذات خطوة واحدة صحيحة والدرجة (2) للإجابة الجزئية التي فيها خطوتان صحيحتان وهكذا (وغالباً ما تكون هذه الأسئلة ذات إجابة منشأة وأحيانا تكون مقياس تقدير).

منهجية الدراسة:

1- عينة التدريج: تكونت عينة تدريج الفقرات من (2000) مفحوص جرى توليد إجاباتهم حاسوبياً باستخدام المحاكاة (simulation) وباستخدام برنامج (WENGEN) حيث تم توليد (60) فقرة ثنائية التدريج لكل مفحوص بالإضافة إلى (10) فقرات متعددة التدريج لكل مفحوص كل فقرة تحتوي على (5) مستويات للإجابة تراوحت بين صفر و(4).

2- عينة القياس التكيفي: تكونت عينة القياس التكيفي من (100) مفحوص جرى اختيارهم من بين الألفي مفحوص الذين شكلوا عينة التدريج، وقد روعي في اختيارهم تمثيل جميع شرائح القدرة للمفحوصين.

3- 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 (1) قيم معلم الصعوبة للفقرات السبعين الناتجة باستخدام

البرنـــــــــــــــــــــامج (RUMM2020).

الجدول رقم (1). قيم معلم الصعوبة لفقرات الاختبار من نوعي ثنائي التدريج ومتعدد التدريج.

الفقرة	نوعها	معلم الصعوبة	الخطأ المعياري	إحصائي المطابقة
10001	ثنائية التدريج	1.196 –	0.054	0.588
10002	ثنائية التدريج	0.235 –	0.049	0.269
10003	ثنائية التدريج	1.494	0.059	0.381
10004	ثنائية التدريج	0.215 –	0.049	0.03
10005	ثنائية التدريج	0.462	0.05	1.277 –
10006	ثنائية التدريج	0.7	0.052	0.803
10007	ثنائية التدريج	0.237 –	0.049	0.011 –
10008	ثنائية التدريج	0.191 –	0.049	0.791 –
10009	ثنائية التدريج	1.079 –	0.053	0.332 –
10010	ثنائية التدريج	1.791	0.064	0.728
10011	ثنائية التدريج	1.6 –	0.058	1.855
10012	ثنائية التدريج	0.584	0.051	0.287 –
10013	ثنائية التدريج	0.83	0.052	0.208 –
10014	ثنائية التدريج	0.11	0.049	0.684 –
10015	ثنائية التدريج	2.286	0.073	0.172

تابع الجدول رقم (1).

الفقرة	نوعها	معلم الصعوبة	الخطأ المعياري	إحصائي المطابقة
10016	ثنائية التدريج	2.138 –	0.067	0.946
10017	ثنائية التدريج	0.533	0.051	1.705
10018	ثنائية التدريج	0.933 –	0.052	0.988
10019	ثنائية التدريج	1.679	0.062	0.074 –
10020	ثنائية التدريج	0.64 –	0.05	0.613
10021	ثنائية التدريج	0.687 –	0.051	0.561 –
10022	ثنائية التدريج	0.895	0.053	0.956
10023	ثنائية التدريج	1.259	0.056	0.135 –
10024	ثنائية التدريج	2.22 –	0.068	0.576
10025	ثنائية التدريج	2.005	0.067	0.277 –
10026	ثنائية التدريج	0.109	0.049	0.035 –
10027	ثنائية التدريج	1.255 –	0.055	0.534 –
10028	ثنائية التدريج	1.127	0.055	0.431
10029	ثنائية التدريج	1.357	0.057	1.021
10030	ثنائية التدريج	2.253 –	0.069	0.826 –
10031	ثنائية التدريج	0.667 –	0.051	0.464 –

10032	ثنائية التدريج	1.457 –	0.057	0.581
10033	ثنائية التدريج	1.029 –	0.053	0.106
10034	ثنائية التدريج	0.242 –	0.049	0.198
10035	ثنائية التدريج	0.197	0.05	1.552
10036	ثنائية التدريج	0.276	0.05	0.419
10037	ثنائية التدريج	0.164 –	0.049	0.658 –
10038	ثنائية التدريج	0.646	0.051	0.14 –
10039	ثنائية التدريج	2.33 –	0.07	0.231 –
10040	ثنائية التدريج	0.739	0.052	0.282 –
10041	ثنائية التدريج	0.804	0.052	1.168 –
10042	ثنائية التدريج	0.425	0.05	0.646 –
10043	ثنائية التدريج	0.142 –	0.049	0.539
10044	ثنائية التدريج	0.603	0.051	0.881 –
10045	ثنائية التدريج	1.324 –	0.055	0.59 –
10046	ثنائية التدريج	0.545	0.051	0.089
10047	ثنائية التدريج	0.055	0.049	0.263 –
10048	ثنائية التدريج	0.896	0.053	0.858
10049	ثنائية التدريج	0.683	0.052	0.155 –

تابع الجدول رقم (1).

الفقرة	نوعها	معلم الصعوبة	الخطأ المعياري	إحصائي المطابقة
10050	ثنائية التدريج	0.818 –	0.051	1.712
10051	ثنائية التدريج	1.16	0.055	0.932 –
10052	ثنائية التدريج	1.381	0.058	0.658
10053	ثنائية التدريج	1.055 –	0.053	0.061 –
10054	ثنائية التدريج	1.257	0.056	0.215 –
10055	ثنائية التدريج	0.47	0.05	0.09
10056	ثنائية التدريج	0.302 –	0.049	2.811
10057	ثنائية التدريج	0.733	0.052	0.091 –
10058	ثنائية التدريج	1.145	0.055	0.183
10059	ثنائية التدريج	0.418	0.05	0.368 –
10060	ثنائية التدريج	1.49 –	0.057	0.071 –
10061	متعددة التدرج	0.568 –	0.023	0.866 –
10062	متعددة التدرج	0.471 –	0.022	1.058 –
10063	متعددة التدرج	0.885 –	0.025	1.422 –
10064	متعددة التدرج	0.321 –	0.024	1.899 –
10065	متعددة التدرج	0.002 –	0.024	0.872 –
10066	متعددة التدرج	0.842	0.027	0.977 –

10067	متعددة التدرج	0.066	0.022	0.162
10068	متعددة التدرج	0.998 –	0.026	1.834 –
10069	متعددة التدرج	0.52 –	0.025	0.761 –
10070	متعددة التدرج	0.9 –	0.026	0.98 –

ويلاحظ من الجدول رقم (1) أن قيم معلم الصعوبة تراوحت بين القيمة (– 2.33) لوجيت للفقرة رقم (33) والقيمة (2.286) لوجيت للفقرة رقم (15)، كذلك كانت جميع الفقرات السبعين مطابقة لنموذج راش إذ إن إحصائي المطابقة لم يقع خارج الفترة (– 2.5 – 2.5) وذلك حسب شرط برنامج RUMM 2020.

إجراءات القياس التكيفي:

1- بعد ترتيب الفقرات تصاعدياً حسب معلم الصعوبة لكل فقرة جرى اختيار (5) فقرات ثنائية التدريج تمثل مستويات الصعوبة جميعها لتكون اختباراً استطلاعياً بشكل المرحلة الأولى من القياس التكيفي، وأرقام هذه الفقرات هي: (10، 11، 18، 47، 48).

2- جرى تقسيم الفقرات ثنائية التدريج الباقية وهي (55) فقرة لتكون خمس مجموعات فرعية تشكل المرحلة الثانية من القياس التكيفي، والجدول رقم (2) يبين أرقام الاختبارات الفرعية الخمسة.

الجدول رقم (2). الفقرات التي شكلت المجموعات الفرعية.

المجموعة الفرعية	أرقام الفقرات
الأولى	39، 30، 24، 16، 60، 32، 45، 27، 1، 9، 53
الثانية	33، 50، 21، 31، 20، 56، 34، 7، 2، 4، 8
الثالثة	37، 43، 26، 14، 35، 36، 59، 42، 5، 55، 17
الرابعة	46، 12، 44، 38، 49، 6، 57، 40، 41، 13، 22
الخامسة	28، 58، 51، 54، 23، 29، 52، 3، 19، 25،

	15

3- جرى تقسيم الفقرات متعددة التدريج وهي (10) فقرات لتكون (5) اختبارات فرعية كل منها يتكون من فقرتين اثنتين.

4- جرى تقديم فقرات الاختبار الاستطلاعي لكل المفحوصين في عينة القياس التكيفي 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.

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

6-بناء على القدرة المحسوبة من المرحلة الثانية جرى تعريض المفحوصين لفقرتين من ذوات التدريج المتعدد تتناسبان مع قدرة المفحوص الجديدة وحساب القدرة النهائية باستخدام فقرات المراحل الثلاث معاً وبالتالي أصبح لكل مفحوص ثلاث قدرات Θ_1 متحصلة من الاستجابات على فقرات الاختبار الاستطلاعي، Θ_2 متحصلة من الاستجابات على فقرات اختبار المرحلة الثانية وفقرات

المرحلة الأولى (الاستطلاعية) وجميعها ثنائية التدريج، Θ_3 متحصلة عن الاستجابات على فقرات اختبار المراحل الثلاث والتي كان آخرها من الفقرات ذات التدريج المتعدد.

النتائج

يبين الملحق رقم (1) تقديرات القدرة المتحصلة لأفراد عينة القياس التكيفي من تطبيق القياس التكيفي في المرحلة الأولى (Θ_1) وهي مرحلة الاختبار الاستطلاعي، و(Θ_2) وهي القدرة المقدرة من المرحلة الثانية ثنائية التدريج مضافاً لها فقرات المرحلة الأولى ثنائية التدريج، وكذلك (Θ_3) وهي القدرة المقدرة من المرحلة الثالثة المكونة من فقرات متعددة التدريج مضافاً لها فقرات المرحلة الثانية ثنائية التدريج وفقرات المرحلة الأولى ثنائية التدريج إضافة للخطأ المعياري في التقدير لجميع المراحل وبجانبها لكل مفحوص القدرة المقدرة من الاختبار الخطي بفقراته السبعين (Θ_L).

وللإجابة عن السؤال الأول من أسئلة الدراسة يبين الجدول (3) المستخلص من الملحق (1) المتوسطات الحسابية للقدرات المقدرة في المراحل الثلاث، ومعاملات الارتباط بيرسون بين كل من القدرات 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 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.

الجدول رقم (3). القيم الإحصائية المتعلقة بتقديرات القدرة في المراحل المختلفة في القياس التكيفي مقارنة بالقدرة الحقيقية المقدرة من الاختبار الخطي.

	المرحلة الأولى (فقرات ثنائية التدريج) (Θ_1)	المرحلة الثانية (فقرات ثنائية التدريج) (Θ_2)	المرحلة الثالثة (فقرات ثنائية + متعددة التدريج) (Θ_3)	الاختبار الخطي (Θ_L)
المتوسط الحسابي للقدرة	0.256 –	0.249 –	0.176 –	0.089 –
معامل الارتباط	0.768	0.881	0.930	1
متوسط القيم المطلقة للفروق	0.767	0.582	0.413	0
المتوسط الحسابي للخطأ المعياري	1.182	0.758	0.533	0.255
متوسط الفروق في الخطأ المعياري	0.928	0.504	0.278	0

ويظهر من الجدول رقم (3) أن المتوسط الحسابي لقدرات المفحوصين في الاختبار الاستطلاعي ذي الفقرات الخمس كانت (– 0.256) لوجيت وهي قريبة نسبيا من المتوسط الحسابي للقدرة الحقيقية المتأتية من الاختبار الخامس بفقراته السبعين وهي (– 0.089) أي بفارق (0.167) لوجيت، وفي المرحلة الثانية كان متوسط تقديرات القدرة يساوي (–

0.249) وهنا يظهر أن المرحلة الثانية لم تحسن تقدير القدرة للاقتراب من القدرة الحقيقية إلا بمقدار ضئيل وهو (0.007) وهو قريب من الصفر، والسبب هنا أن كلتا المرحلتين الأولى والثانية تتألفان من فقرات ثنائية التدريج من نوع الإجابة المنتقاة،بينما تألف الاختبار الخطي من كلا النوعين من الفقرات ثنائية ومتعددة التدريج، ويدعم ذلك أنه حينما أضيفت المرحلة الثالثة التي تحتوي على فقرات متعددة التدريج من نوع الإجابة المنشأة تحسن تقدير القدرة باتجاه القدرة الحقيقية حيث أصبح (– 0.176) أي بفارق (0.087) لوجيت، وهذه النتائج تتفق مع دراسة البرصان (2006) من حيث المرحلة الأولى من القياس التكيفي بينما تختلف من حيث المرحلة الثانية للقياس التكيفي، وبالمثل مع دراسة الثوابية (2004) ودراسة هندركسون (2002)، 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 (القيم المطلقة لباقي الطرح) بين التقديرات للقدرة في مراحل القياس التكيفي والقدرة المقدرة من الاختبار الخطي،ذلك أن المتوسط الحسابي للتقديرات الموجبة والسالبة قد يعطي استنتاجا مضللاً بسبب تعادل الموجب مع السالب.

فباستخدام المتوسط الحسابي للفروق المطلقة يظهر أن تقديرات المرحلة الأولى الاستطلاعية تبعد في المتوسط عن القدرة الحقيقية بمقدار (0.767) لوجيت، وفي المرحلة الثاني أصبحت تبعد بمقدار (0.582) لوجيت، وفي المرحلة الأخيرة التي احتوت فقرات متعددة التدريج كانت القدرة في المتوسط تبعد بمقدار (0.413) 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 التكيفي الذي استخدم فقط (12) فقرة من أصل (70) فقرة مع الانتباه لإشكالية احتواء القياس التكيفي على كلا النوعين من الفقرات.

ومن حيث المؤشر الثالث على فاعلية القياس التكيفي المبني باستخدام فقرات ثنائية ومتعددة التدريج وهو معامل الارتباط فيظهر أن معامل الارتباط في المرحلة الاستطلاعية الأولى مع القدرة الحقيقية كان (0.767) بتباين مفسر مقداره (0.588) أما في المرحلة الثانية فقد أصبح معامل الارتباط (0.766) ويلاحظ هنا مدى التحسن في التقدير باتجاه القدرة الحقيقية، لكن في المرحلة الأخيرة التي احتوت فقرات متعددة التدريج بالإضافة للفقرات الثنائية السابقة أصبح معامل الارتباط (0.930) وبتباين مفسر مقداره (0.865) الأمر الذي يبين مدى اقتراب تقدير القدرة باستخدام الاختبار التكيفي الذي احتوى فقط (12) فقرة من تقدير القدرة باستخدام الاختبار الخطي بفقراته السبعين،علماً بأن معاملات الارتباط المذكورة هي دالة عند مستوى (α=0.01).

وبالنسبة للخطأ المعياري في التقدير يُظهر الجدول أن متوسط الخطأ المعياري في التقدير كان في المرحلة الأولى الاستطلاعية (1.182) ثم أصبح في المرحلة الثانية (0.758) ويلاحظ هنا مقدار الانخفاض في الخطأ المعياري في التقدير الذي اعتمد حتى الآن على فقرات ثنائية التدريج، لكن في

المرحلة الثالثة التي احتوت على فقرات متعددة التدريج أصبح (0.533) وهو يبعد بمقدار (0.277) عن متوسط الخطأ المعياري في تقديرات القدرة في الاختبار الخطي والذي يبلغ (0.255) بينما كان يبعد بمقدار (0.504) في المرحلة الثانية، وبمقدار (0.928) 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 (1: 2) بينما كانت نسبة عدد فقرات الاختبار الخطي إلى عدد فقرات الاختبار التكيفي تقريباً (1: 6) وهذا يبين مقدار الخفض الكبير في الكلفة والجهد للفاحص 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 ذلك الذي يحتوي كلا النوعين من الفقرات ذات الإجابة المنتقاة، وذات الإجابة المُنشأة في مختلف أنواع الاختبارات.

ملحق (1)

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 (1) فقرات ذات إجابة منتقاة		مرحلة (2) فقرات ذات إجابة منتقاة + فقرات المرحلة السابقة		مرحلة (3) فقرات ذات إجابة منشأة + فقرات المرحلتين السابقتين		الاختبار الخطي (60 فقرة ذات إجابة منتقاة + 10 فقرات ذات إجابة منشأة)	
	(Θ_1)	SEE_1	(Θ_2)	SEE_2	(Θ_3)	SEE_3	(Θ_L)	SEE_L
1	0.466 –	1.052	0.454 –	0.691	0.118 –	0.462	0.008	0.213
2	2.601 –	1.525	1.079 –	0.734	0.648 –	0.476	0.383	0.222
3	0466. –	1.052	0.411	0.963	0.865	0.533	1.287	0.26
4	0.53	1.055	0.288	0.682	0.493	0.517	0.533	0.227
5	2.601 –	1.525	3.315 –	0.975	2.207 –	0.553	1.5 –	0.244
6	1.519 –	1.169	1.617 –	0.806	1.313 –	0.501	0.88 –	0.217
7	0.53	1.055	0.15 –	0.698	0.021	0.475	0.432	0.223
8	1.519 –	1.169	2.311 –	0.975	2.033 –	0.569	1.33 –	0.234
9	1.601	1.178	0.515	0.689	0.063 –	0.498	0.239	0.218
10	2.601 –	1.525	3.315 –	0.975	2.948 –	0.708	1.686 –	0.258
11	2.711	1.539	2.386	0.803	1.693	0.555	0.913	0.241
12	0.53	1.055	0.628 –	0.74	0.997 –	0.498	0.433 –	0.21
13	1.601	1.178	1.922	0.795	1.431	0.545	1.651	0.283
14	0.466 –	1.052	0.454 –	0.691	0.889 –	0.473	0.834 –	0.216
15	0.466 –	1.052	0.914 –	0.727	0.219 –	0.471	0.214 –	0.211
16	1.519 –	1.169	1.081 –	0.732	0.686 –	0.461	1.385 –	0.237
17	1.519 –	1.169	1.081 –	0.732	0.686 –	0.461	1.121 –	0.225
18	1.519 –	1.169	1.617 –	0.806	1.562 –	0.534	0.653 –	0.212
19	1.519 –	1.169	0.61 –	0.699	0.561 –	0.454	1.023 –	0.221
20	0.466 –	1.052	0.454 –	0.691	0.311 –	0.454	0.054 –	0.214
21	1.601	1.178	0.952	0.692	0.174	0.495	0.192	0.217
22	0.466 –	1.052	1.447 –	0.804	1.78 –	0.584	0.698 –	0.213
23	0.53	1.055	0.719	0.689	0.527	0.498	0.799	0.237
24	1.519 –	1.169	0.286	0.705	0.511	0.535	0.099	0.215
25	0.466 –	1.052	0.023 –	0.68	0.09	0.478	0.564 –	0.211
26	0.53	1.055	0.15 –	0.698	0.493	0.517	0.286	0.219
27	0.53	1.055	0.15 –	0.698	0.021	0.475	0.482	0.225

تابع ملحق رقم (1).

المفحوص	مرحلة (1) فقرات ذات إجابة منتقاة		مرحلة (2) فقرات ذات إجابة منتقاة + فقرات المرحلة السابقة		مرحلة (3) فقرات ذات إجابة منشأة + فقرات المرحلتين السابقتين		الاختبار الخطي (60 فقرة ذات إجابة منتقاة + 10 فقرات ذات إجابة منشأة)	
	(Θ_1)	SEE_1	(Θ_2)	SEE_2	(Θ_3)	SEE_3	(Θ_L)	SEE_L
28	2.601 –	1.525	2.07 –	0.751	1.491 –	0.478	1.072 –	0.223

0.21	0.345 –	0.471	0.219 –	0.727	0.914 –	1.052	0.466 –	29
0.22	0.975 –	0.461	0.621 –	0.727	0.914 –	1.052	0.466 –	30
0.253	1.156	0.572	1.316	0.723	1.174	1.055	0.53	31
0.216	0.145	0.505	0.648 –	0.689	0.515	1.178	1.601	32
0.232	0.69	0.478	0.09	0.68	0.023 –	1.052	0.466 –	33
0.218	0.927 –	0.477	1.1 –	0.975	2.311 –	1.169	1.519 –	34
0.214	0.743 –	0.466	0.885 –	0.732	1.081 –	1.169	1.519 –	35
0.244	0.971	0.513	0.768	0.689	0.719	1.055	0.53	36
0.21	0.433 –	0.475	0.021	0.698	0.15 –	1.055	0.53	37
0.227	1.172 –	0.477	1.076 –	0.751	2.07 –	1.525	2.601 –	38
0.217	0.88 –	0.478	1.067 –	0.734	1.079 –	1.525	2.601 –	39
0.289	1.733	0.579	1.46	0.692	0.952	1.178	1.601	40
0.213	0.008	0.462	0.18 –	0.699	0.61 –	1.169	1.519 –	41
0.247	1.031	0.505	0.648 –	0.692	0.952	1.178	1.601	42
0.211	0.52 –	0.497	1.107 –	0.691	0.454 –	1.052	0.466 –	43
0.239	0.855	0.536	1.027	0.689	0.719	1.055	0.53	44
0.23	0.637	0.462	0.18 –	0.699	0.61 –	1.169	1.519 –	45
0.211	0.126 –	0.462	0.488 –	0.732	1.081 –	1.169	1.519 –	46
0.212	0.037 –	0.459	0.689 –	0.68	0.023 –	1.052	0.466 –	47
0.29	2.054 –	0.612	2.53 –	0.815	2.619 –	1.525	2.601 –	48
0.249	1.56 –	0.465	0.82 –	0.727	0.914 –	1.052	0.466 –	49
0.25	1.093	0.536	0.581	0.68	0.023 –	1.052	0.466 –	50
0.212	0.081 –	0.536	1.027	0.723	1.174	1.055	0.53	51
0.21	0.345 –	0.486	0.313 –	0.693	0.411	1.052	0.466 –	52
0.253	1.622 –	0.501	1.313 –	0.806	1.617 –	1.169	1.519 –	53
0.264	1.355	0.621	2.06	0.795	1.922	1.178	1.601	54
0.212	0.608 –	0.461	0.621 –	0.727	0.914 –	1.052	0.466 –	55
0.225	0.482	0.493	0.245	0.682	0.288	1.055	0.53	56
0.211	0.17 –	0.495	0.174	0.692	0.952	1.178	1.601	57

تابع ملحق رقم (1).

الاختبار الخطي **(60 فقرة ذات إجابة منتقاة** **+** **10 فقرات ذات إجابة منشأة)**		**مرحلة (3)** **فقرات ذات إجابة منشأة** **+** **فقرات المرحلتين السابقتين**		**مرحلة (2)** **فقرات ذات إجابة منتقاة** **+** **فقرات المرحلة السابقة**		**مرحلة (1)** **فقرات ذات إجابة منتقاة**		**المفحوص**
SEE_L	(Θ_L)	SEE_3	(Θ_3)	SEE_2	(Θ_2)	SEE_1	(Θ_1)	
0.247	1.031	0.544	1.141	0.709	1.39	1.539	2.711	58
0.21	0.302 –	0.511	0.638 –	0.689	0.719	1.055	0.53	59
0.219	0.286	0.493	0.245	0.698	0.15 –	1.055	0.53	60
0.289	1.733	0.533	0.865	0.693	0.411	1.052	0.466 –	61
0.334	2.436 –	0.553	2.207 –	0.815	2.619 –	1.525	2.601 –	62
0.349	2.551 –	0.612	2.53 –	0.815	2.619 –	1.525	2.601 –	63
0.21	0.258 –	0.462	0.118 –	0.691	0.454 –	1.052	0.466 –	64
0.275	1.896 –	0.569	2.033 –	0.975	2.311 –	1.169	1.519 –	65
0.375	2.571	0.702	2.459	0.958	2.597	1.178	1.601	66
0.229	1.223 –	0.465	0.82 –	0.727	0.914 –	1.052	0.466 –	67
0.312	2	0.52	0.895	0.692	0.952	1.178	1.601	68

0.228	0.584	0.535	0.511	0.691	0.165 –	1.169	1.519 –	69
0.24	1.442 –	0.584	1.78 –	0.727	0.914 –	1.052	0.466 –	70
0.332	2.204	0.579	1.46	0.692	0.952	1.178	1.601	71
0.21	0.476 –	0.535	0.511	0.705	0.286	1.169	1.519 –	72
0.234	0.744	0.498	0.527	0.689	0.719	1.055	0.53	73
0.282	1.973 –	0.553	2.207 –	0.815	2.619 –	1.525	2.601 –	74
0.296	1.818	0.545	1.431	0.795	1.922	1.178	1.601	75
0.278	1.573	0.545	1.431	0.795	1.922	1.178	1.601	76
0.389	2.821 –	0.83	2.89 –	0.975	2.311 –	1.169	1.519 –	77
0.21	0.389 –	0.513	0.768	0.689	0.719	1.055	0.53	78
0.21	0.302 –	0.464	0.185 –	0.682	0.288	1.055	0.53	79
0.321	2.099	0.888	3.333	0.964	3.062	1.539	2.711	80
0.358	2.439	0.631	2.336	0.964	3.062	1.539	2.711	81
0.263	1.753 –	0.534	1.562 –	0.806	1.617 –	1.169	1.519 –	82
0.257	1.22	0.579	1.46	0.692	0.952	1.178	1.601	83
0.237	1.385 –	0.531	1.49 –	0.804	1.447 –	1.052	0.466 –	84
0.395	2.718	0.579	1.46	0.692	0.952	1.178	1.601	85
0.232	1.276 –	0.534	1.562 –	0.732	1.081 –	1.169	1.519 –	86
0.303	1.907	0.572	1.316	0.723	1.174	1.055	0.53	87

تابع ملحق رقم (1).

الاختبار الخطي (60 فقرة ذات إجابة منتقاة + 10 فقرات ذات إجابة منشأة)		**مرحلة (3) فقرات ذات إجابة منشأة + فقرات المرحلتين السابقتين**		**مرحلة (2) فقرات ذات إجابة منتقاة + فقرات المرحلة السابقة**		**مرحلة (1) فقرات ذات إجابة منتقاة**		**المفحوص**
SEE_L	(Θ_L)	SEE_3	(Θ_3)	SEE_2	(Θ_2)	SEE_1	(Θ_1)	
0.273	1.498	0.574	1.728	0.795	1.922	1.178	1.601	88
0.368	2.679 –	0.569	2.033 –	0.975	2.311 –	1.169	1.519 –	89
0.215	0.788 –	0.486	0.313 –	0.693	0.411	1.052	0.466 –	90
0.309	2.231 –	0.553	2.207 –	0.815	2.619 –	1.525	2.601 –	91
0.269	1.823 –	0.499	1.244 –	0.804	1.447 –	1.052	0.466 –	92
0.268	1.425	0.52	0.895	0.722	1.406	1.178	1.601	93
0.497	3.292	0.702	2.459	0.958	2.597	1.178	1.601	94
0.29	2.054 –	0.491	1.706 –	0.751	2.07 –	1.525	2.601 –	95
0.232	0.69	0.52	0.895	0.722	1.406	1.178	1.601	96
0.321	2.329 –	0.553	2.207 –	0.975	3.315 –	1.525	2.601 –	97
0.416	2.982 –	0.83	2.89 –	0.975	2.311 –	1.169	1.519 –	98
0.217	0.192	0.462	0.118 –	0.68	0.023 –	1.052	0.466 –	99
0.421	2.882	0.703	2.34	0.8	1.696	1.055	0.53	100

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 التدريج.* رسالة دكتوراه غير منشورة، جامعة عمان العربية للدراسات العليا، 2006م.

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، 2004م

عبيدات، عمر. *فاعلية الاختبار التكيفي المحوسب في تقدير القدرة العقلية باستخدام مصفوفات رافن.* رسالة دكتوراه غير منشورة، جامعة اليرموك، 2008م.

العموش، جميل. فاعلية القياس التكيفي في تقويم بعض القدرات المعرفية لدى طلبة السنة الأولى الجامعية. رسالة دكتوراه غير منشورة، جامعة عمان العربية للدراسات العليا، 2003م.

ثانياً: المراجع الأجنبية:

Anastasi, A. Urbina, S. Psychological Testing. 7th 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 Publishers, 1991.

Hendrickson, A. Scaling of Two – Stage Adaptive Test Configurations For Achievement Testing. Unpublished

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.

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, 3rd ed., New Jersey: Prentice Hall. New Jersey: Lawrence Erlbaum Associates, Inc, 1994.

Roex, A. & Degryse, J. A Computerized adaptive knowledge test as an assessment 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

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.