"المقارنة بين الاختبار الورقى والاختبار التكيفى المحوسب فى تقدير القدرة على الاستدلال اللفظى"

بحث مقدم من الباحث

مصطفى أحمد قاسم محمد

أخصائى تحليل احصائى

بالمركز القومى للامتحانات والتقويم التربوى

تحت إشراف

أ.د.أمينة محمد كاظم

أستاذ علم النفس التعليمي

بكلية البنات – جامعة عين شمس

١٤٣٦ هـ – ٢٠١٥ م

ملخص البحث :

يهدف هذا البحث إلى المقارنة بين تقدير قدرة الأفراد عند تطبيق مقياس الاستدلال اللفظى بكل من الصورة الورقية والصورة التكيفية المحوسبة.
وتمثلت عينة البحث الكلية في (٩٤) طالب وطالبة من طلاب المراحل الدراسية من الصف الثالث الابتدائى إلى الصف الثانى الثانوى بمحافظاتى القليوبية والجيزة وقد تكون مقياس الاستدلال اللفظى من (٩٦) مفردة. تم تقسيمهما فى النسخة الورقية للمقياس إلى (٢) صورة اختبارية وتم استخدام برنامج Fast Test Professional لتخزين مفردات المقياس بمؤشراتها الاحصائية أيضاً تم استخدامه فى تطبيق المقياس بالأسلوب التكيفى المحوسب .

وقد أسفرت النتائج عن تعادل القياس فى تقدير قدرة الأفراد عند استخدام كل من الصورة الورقية للمقياس والصورة التكيفية المحوسبة ، كما وجد اختلاف فى تقديرات القدرة باختلاف المرحلة الدراسية مما يظهر قدرة المقياس فى التمييز بين الأفراد فى المراحل المختلفة أيضا داخل كل مرحلة على حدة.

Abstract:

The current research **aimed at** comparing between the paper and pencil test and the computer adaptive test when applying the verbal reasoning scale.

The sample consisted of (٩٤) students from Primary Three to Secondary Two; from Qalyuibia and Giza governorates. The scale consisted of (96) items; divided into (2) forms on the paper version.

The Fast Test Professional Computer Program was used to storing the items with its statistical indicators and applying the scale by using the computerized adaptive method.

The results

- The equating measurement in the ability estimation of individuals on both versions of a scale (paper and pencil- computer adaptive).
- Due to the difference of the school stages, there is a difference between the ability estimations on the adaptive version of a scale; i.e, the more difference between the school stages, the more difference between the ability estimations.

المقارنة بين الاختبار الورقى والاختبار التكيفى المحوسب فى تقدير القدرة على الاستدلال اللفظى

١- مقدمة

يهدف علم النفس إلى فهم الظواهر السلوكية والتنبؤ بها والتحكم فيها ويؤدى القياس النفسى دوراً هاما فى تحقيق هذه الأهداف بدقة وموضوعية. ويعتبر التفكير الأداة الحقيقية التى يواجه بها الانسان متغيرات العصر ومن خلال التفكير يكتشف العلاقات بين هذه المتغيرات بماقد يؤدى إلى فهمها والتنبؤ بها وضبطها وعليه فإن اهتمام المجتمعات المتقدمة اصبح ينصب على تنمية مهارات التفكير لدى أفرادها لكى تفيدهم الإفادة المثلى، ومن أكثر طرق التفكير بروزا مهارات التفكير الاستدلالى حيث اتسع نطاق الاهتمام بدراستها وتنميتها.

تبدو القدرة على الاستدلال فى صورتين،الأولى تتمثل فى القدرة على الاستدلال الاستقرائى وتبدو فى الأداء العقلى الذى يتميز باستنتاج القاعدة العامة من جزئياتها وحالاتها الفردية، والثانية القدرة على الاستدلال الاستنباطى وتبدو فى الأداء العقلى الذى يتميز باستنباط الأجزاء من القاعدة العامة.**(أحمد المغربى ، ٢٠١٠ ، ص ١٦٧)**

ويستخدم الاستدلال فى البحث العلمى بطريقة واسعة حيث يعد من المهارات الأساسية المطلوبة فى الباحث العصرى، فهو يستخدم فى صياغة الفروض البحثية بطريقتين هما، الطريقة الاستنباطية والطريقة الاستقرائية، ويساعد ذلك فى بناء المعرفة حيث تزيد من قدرة الباحث على القيام ببحثه وتوجيهه الوجهة السليمة نحو بحث مشكلته.**(سمية على عبد الوارث أحمد، ٢٠١١، ص ٥٩)**

وقد تعددت الأدوات الخاصة بقياس الاستدلال ومن أهمها بطارية الاختبارات المعرفية العاملية لأكستروم وفرنش وهارمان وديرمين (1986) ،وبطارية المقاييس البريطانية للقدرات(1983) ،واختبار ألفا (1986) ، واختبار القدرة العقلية العامة لأوتيس و لينون(1967) ، واختبار تونى للذكاء لبراون وشيربينو ودولار (1983) ، واختبار تسجيل الخريجين للدراسات العليا Graduate Record Examination .

واعتمدت بعض تلك المقاييس فى بنائها على أسلوب القياس التقليدى جماعى المرجع الذى ساد ميدان القياس النفسى زمنا كبيراً بالرغم مماله من العديد من أوجه القصور مثل اختلاف تفسيرات النتائج المشتقة من مقاييس القدرات المعروفة بإختلاف الإختبار المستخدم وبإختلاف خصائص الجماعة المستخدمة في تحليل النتائج، وعدم وجود وحدات قياس علي المستويات المختلفة للمتغير ممايجعل من غير المستطاع قياس النمو النموذجي للقدرة عبر مراحل العمر المختلفة، بالإضافة إلى أنها تهدف إلي التوصل إلي درجة أو معلومات كلية شاملة أكثر من إهتمامها بتوفير معلومات تفصيلية عن نواحي القوة والضعف في مستوي القدرات المختلفة للفرد.**(أمينة كاظم ، ١٩٨٩ ، ص ص ٦-٧)**

وقد حاول علماء القياس النفسي تلافي تلك العيوب للتوصل للموضوعية فى القياس السلوكى أسوة بالقياس الفيزيائى وقد تطلب هذا:-

- بنود صادقة يمكنها تعريف المتغير موضوع القياس تعريفا اجرائيا .
- صدق التدرج لهذه البنود.
- أنماط استجابات صادقة-التوافق بين تدرج الأفراد على الاختبار ومميزات البنود.
- قياسات خطية يمكن استخدامها لدراسة النمو.

(أمينة كاظم ،١٩٩٦،ص ٣٠٧)

(ويتمثل القياس الموضوعى فى نظرية الإستجابة للمفردة) Item Response Theory أو ما كان يعرف سابقا بنظرية السمات الكامنة Latent Trait Theory والتى تتميز بالأتى :-

- لاتعتمد إحصاءات المفردة على عينة الأفراد ولاعلى خصائص الاختبار، وتكون إحصاءات المفردات وقدرة المفحوصين على نفس التدريج بوحدة قياس معرفة (اللوجيت). **Schumacker (, R. , 2005, P1)**
- لاتعتمد تقديرات الأفراد على خصائص الصورة الاختبارية المستخدمة.

- تساعد فى ترتيب الطلاب تبعا لقدراتهم، بالإضافة إلى أنها تعطى معلومات تشخيصية Diagnostic information عن المفحوصين.

(Hambelton ,R. and Swaminathan, H.&Rogers , H. ,1991 ,P155)

وقد حدث تطور هام فى مجال القياس النفسى حيث تم الاستفادة من تكنولوجيا الحاسب الألى متمثلا ذلك فى الاختبارات التكيفية المحوسبة التى تعتمد على نظرية الاستجابة للمفردة فى تدريج مفردات المقياس وتحقق إنتقاء المفردات المناسبة لقدرة الفرد بحيث تكون صورة اختبارية محبوكة على مستوى قدرة الفرد ومعادلة للمقياس الكلى للمتغير موضوع القياس كما يعتمد إجراء المقياس على الحاسب الآلى الذى له العديد من المميزات مثل:

- التقليل من طول الاختبار وذلك عند الوصول لمستوى مطلوب من دقة القياس.
- التوفير فى تكلفة طباعة الاختبار الورقى وتكلفة توفير أوراق للأجابة، أيضاً يساعد على التقليل من فرص إتلاف وضياع أوراق استجابات الطلاب
- اعطاء نتائج فورية للطالب، وبذلك يتلافى عيوب الاختبار الورقى فى تأخير إعلان النتائج.

(Douglas A. Strader,2012,p3)

ولما كان من الأهمية بمكان مواكبة التطور التكنولوجى فى العملية التربوية فقد تقدمت العديد من الدول فى مجال استخدام الحاسب الآلى فى القياس النفسى والتربوى، وأصبح لديها شركات متخصصة فى تكنولوجيا الكمبيوتر والإختبارات المعتمدة على الحاسب الآلى، لذلك تسعى الدراسة الحالية إلى توظيف بعض من التقنيات الحديثة فى القياس النفسى والتربوى ذلك لما لها من أهمية على المستوى المحلى والمستوى القومى.

٢- هدف الدراسة

تهدف الدراسة الحالية إلي:

توظيف تكنولوجيا الحاسب الآلى فى تطبيق الأسلوب التكيفى المحوسب عند إجراء المقياس والمقارنة بين تقدير قدرة الأفراد عند اختلاف طريقة التطبيق (الورقى /التكيفى المحوسب).

٣- أهمية الدراسة

تتحدد الأهمية النظرية والعملية للدراسة الحالية كما يلي:

- استخدام منحى جديد فى القياس النفسى والتربوى بتدريج مقاييس تكيفية محوسبة لقياس الاستدلال اللفظى، قد يفتح المجال لتشجيع الباحثين لبناء وتطوير العديد من المقاييس النفسية باستخدام هذا المنحى.

- ❖ إكمال منظومة التعليم عن بعد والتي تتم عن طريق الكمبيوتر بمقاييس نفسية تكيفية محوسبة ولها موضوعية في قياس الاستدلال.
- ❖ توفر المقاييس التكيفية المحوسبة الوقت والجهد والمال المتطلب في طباعة الاختبارات النفسية وخاصة في التطبيقات الجماعية وأيضا الدقة في رصد وتصحيح الدرجات، واستخدام عدد قليل من المفردات لتقدير القدرة.
- ❖ قد يعمل الأسلوب التكيفى المحوسب على الحد من الغش حيث يكون لكل فرد مجموعة المفردات الخاصة به والتى يرشحها البرنامج بماياسب قدرته.

٤- تحديد مشكلة الدراسة

تتحدد مشكلة الدراسة فى التساؤل التالي:

هل تختلف تقديرات قدرة أفراد عينة الدراسة على مقياس الاستدلال اللفظى باختلاف نوع التطبيق (ورقى/تكيفى محوسب)؟

٥- محددات الدراسة

تتحدد الدراسة الحالية بما يلي:-

- تقتصر **أدوات الدراسة** على مقياس الاستدلال اللفظى (التماثل اللفظى) المشتق من عدة مقاييس لقياس الاستدلال (سبق أن أعده الباحث).
- **عينة الدراسة** المتمثلة فى طلاب المراحل الدراسية من الصف الثالث الابتدائى إلى الصف الثانى الثانوى ببعض المدارس الحكومية والخاصة (لغات-عربى) بمحافظات القاهرة والجيزة والقليوبية.

٦- مصطلحات البحث

- **الاستدلال**

عملية تفكير تتضمن وضع الحقائق أو المعلومات بطريقة منظمة أو معالجتها بحيث تؤدى إلى استنتاج أو قرار أو حل لمشكلة، والاستدلال المنطقى هو عملية تفكير تستند إلى قواعد واستراتيجيات معينة تهدف إلى توليد معرفة جديدة عن طريق الاستنباط والاستقراء.

(فتحى عبد الرحمن جروان، ١٩٩٩، ص ٤٢١)

- **نظرية الإستجابة للمفردة**

هى اتجاه معاصر فى القياس النفسى يسعى إلى تحقيق الموضوعية وتوفير خطية واستقلالية القياس كما يسعى إلى تكامل أهداف القياس، وتفترض تلك النظرية وجود واحدة أو أكثر من السمات الأساسية غير الملاحظة التى تحدد استجابات الفرد الملاحظة لبنود اختبار ما، كما تهتم بالربط بين استجابة الفرد على مفردة اختبارية وبين خصائص هذه المفردة، ويندرج تحت هذه النظرية عدد من النماذج الرياضية الاحتمالية التى تحدد العلاقة بين الاستجابات الملاحظة على الاختبار والسمات أو القدرات غير الملاحظة التى يفترض أنها تحدد هذه الاستجابات.

- **بنك الأسئلة**

مجموعة كبيرة من المفردات تنضم فى تدريج واحد مشترك حتى تشكل شبكة تكون فيما بينها نسيجاً من الاختبارات المدرجة على ميزان واحد وبصفر واحد وتغطى مدى واسعاً من المتغير موضوع القياس.

- **الاختبار المحوسب الخطى**

تحويل النسخة الورقية المعتادة للاختبار إلى اختبار معتمد على الحاسب الآلي، وبالتالي يتسم الاختبار المعتمد على الحاسب الآلي بمنهج ثابت كما هو الحال في النسخ الورقية الأصلية للاختبار، فكل المتقدمين لأداء هذا الاختبار يجيبون عن الأسئلة بالترتيب نفسه الذي تقدم فيه الأسئلة في النسخة الورقية.

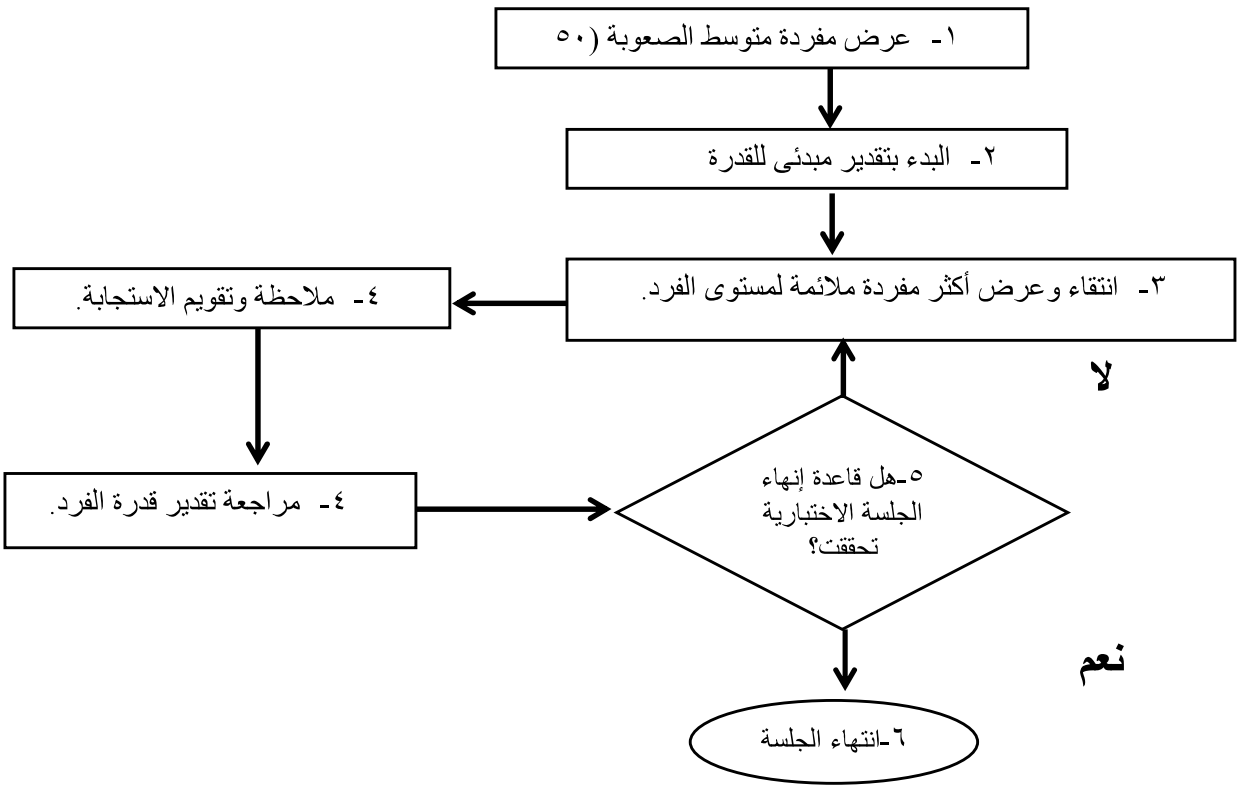
- **الاختبار التكيفى المحوسب**

هو الإختبار الذي يقوم فيه مؤدي الاختبار بالإجابة عن مجموعات مختلفة من الأسئلة، تطرح عليهم بناء على مستوياتهم، حيث تؤثر إجابتهم عن سؤال ما على نوعية الأسئلة التالية التي يختارها الحاسب لهم ففي حال الإجابة الصحيحة عن السؤال سيختار الحاسب سؤالاً أصعب قليلا من الأول، ويطرحه على المتقدم إلى الاختبار، وبالعكس إذا كانت الإجابة خاطئة سيختار الحاسب سؤالاً أقل صعوبة، ومن هنا جاءت تسمية "الاختبار االتكيفى"

(Peter Davidson ترجمة خالد بن عبد العزيز الدامغ, 2006 , P240)

٧- الاطار النظرى للدراسة

تعد نظرية الاستجابة للمفردة أحدى وسائل التقويم الحديثة التى تعمل على تقدير قدرة الفرد ولها العديد من التطبيقات العملية ومن أهمها الأسلوب التكيفى المحوسب لتطبيق الاختبارات والذى يقوم فلسفة تصميمه على انشاء صورة اختبارية خاصة لكل طالب تتعادل نتائجها مع نتائج بنك الأسئلة المسحوبة منه بحيث يتم توجيه هدف الاختبار لتحديد مستوى قدرة الطالب بأقل مجهود ووقت وبكفائة أعلى حيث يتم اختيار وتطبيق كل مفردة من مفردات الاختبار وتصحيحها بالكمبيوتر بناءاً على المفردة التى تعطى أكبر كم من المعلومات الملائم لمستوى قدرة المفحوص وذلك من خلال إجابته على المفردة السابقة، وبهذا فإن المفحوص يستقبل المفردات المتطابقة فقط مع قدرته ويتجنب المفردات السهلة جداً والصعبة جداً عن مستواه لأنها لاتعطى معلومات أكثر عن مستوى قدرة الطالب وهو مايوضحه الشكل التالى.



شكل (١) الخطوات المنطقية للاختبار التكيفى المحوسب.

(Wang J.,2009,P26)

ولتوضيح الأسلوب التكيفى المحوسب ينبغى الإشارة بأنه عند تطبيق الاختبار بواسطة الكمبيوتر فإنه يوجد نوعان من الاختبارات المحوسبة وهى:-

- **الاختبار المحوسب الخطى**

يتم تطبيق النسخة الورقية من الاختبار بواسطة الكمبيوتر بنفس ترتيب ونفس طريقة عرض المفردات ويكون الفرق الوحيد هو تغيير طريقة التطبيق من ورقية إلى محوسبة.

- **الاختبار التكيفى المحوسب**

يتم تطبيق الاختبار تكيفياً بواسطة الحاسب الآلى وذلك بتفصيل اختبار خاص لكل طالب ولاتقتصر طريقة التطبيق على الانتقال من الورقة إلى الحاسب ولكن يتم الانتقال من الأسلوب الخطى إلى الأسلوب التكيفى.

وفى السنوات الأخيرة ظهرت حاجة ملحة لاستخدام الأسلوب التكيفى المحوسب فى الاختبارات وذلك لمزاياه المتعددة والمتمثلة فى:-

- تحسين كفاءة وجودة الخصائص السيكوميترية للاختبار ويكون ذلك بتطبيق مجموعة من المفردات التى تتناسب مع قدرة المفحوص .
- الزيادة فى سرية الاختبارات خاصة عند إعادة استخدام المفردات فى اختبارات أخرى
- التصحيح والاعلان الفورى للنتائج
- يوفر التكلفة المادية بالتخلص من طباعة الاختبارات واوراق الأجابة
- يعمل على تحفيز المفحوصين

(Steven L. Wise,2014,p3)

- المرونة فى تطبيق الاختبار التكيفى المحوسب، حيث تعطى الاختبارات بناءا على طلب الممتحن "On demand". **(Thompson N. A.,no date ,P2)**
- صممت لكل فرد على حدة، فبالتالى لايحتاج المفحوص لإنتظار الأخرين حتى ينتقل من جزء لأخر فى الأختبار ممايحد من أحد مصادر قلق الأختبار بالإضافة إلى أنه يتيح العديد من الأشكال فى توزيع وقت الاختبار أو شكل المفردات، مثل أن تتضمن المفردات صوت، صورة، حركة.

(Linacre J.M., 2000, PP30-32)

وبالرغم من تلك المميزات إلا أنه توجد بعض العيوب التى يجب الأشارة إليها وهى:-

١- يحتاج تدريج مفردات المقياس إلى عينة كبيرة من الأفراد، ذلك لأنه يعتمد على أنواع معينة من نماذج نظرية الاستجابة للمفردة، قد يصل عدد العينة فى بعض النماذج إلى (١٠٠٠) مفحوص أو أكثر.

٢- يتطلب وجود خبراء ذو كفاءة عالية فى القياس التربوى، وذلك لمعايرة بنك الأسئلة وبناء اختبار تكيفى محوسب فعال، أيضا يتطلب توافر العديد من البرامج للتجهيز للاختبار التكيفى المحوسب.

(Thompson N. A.,no date,P2)

٣- عدم وجود اتفاق حول أفضل استراتيجيات انتقاء المفردات، والاستراتيجيات المختلفة تعطى نتائج متباينة وغير دقيقة.

٤- تأثر المشكلات المتعلقة بسياق عرض المفردات وترتيبها وتوازن محتواها تأثيرا سلبيا فى تقدير قدرة الأفراد.

٥- التكلفة العالية وخاصة فى بناء اختبارات محوسبة على نطاق واسع Large scale testing مثل اختبار القبول بالدراسات العليا GRE واختبار الاستعداد المدرسى SAT.

٦- فى بعض الأحيان يقيس الاختبار قدرات متعددة وبالتالى يجب استخدام نماذج نظرية الاستجابة للمفردة المتعددة الأبعاد ولكن يعاب على تلك النماذج تعقدها مقارنة بالنماذج أحادية البعد، أيضا يصعب تفسير بارمترات بعض النماذج خاصة نماذج المفردات متعددة الاستجابة. **(Thissen D.,Reeve B.B.&Bjorner J.B& others, 2007,P111)**

٧- عدم قدرة الطالب على تخطى أى مفردة أيضاً عدم القدرة على الرجوع للمفردات السابقة التى أجاب عليها مما يزيد من قلق الاختبار وتوتر الطالب، بالإضافة إلى تفاوت توقيتات الأجابة من طالب لأخر فمثلا الطالب القوى يعرض عليه مفردات صعبة تحتاج إلى وقت للأجابة عن المفردات السهلة.**(Colvin k.,2014,P18)**

وبوضع عيوب ومميزات الاختبار التكيفى المحوسب فى كفتى ميزان فمميزاته تتفوق على عيوبه وهو مايظهر عملياً فى معظم الاختبارات الدولية التى تستخدم الأسلوب التكيفى المحوسب مثل اختبار تسجيل الخريجين للدراسات العليا Graduate Record Examination واختبار التويفل الدولى TOEFL (IBT) وبناءاً على ماسبق ينبغى العمل على تطوير النسخ الورقية من الاختبارات إلى نسخ تكيفية محوسبة بحيث يكون ذلك موجهاً حسب الهدف من الاختبار.

ولتطوير النسخة الورقية من الاختبار إلى نسخة تكيفية محوسبة يوجد (٧) مكونات يجب وضعها فى الاعتبار وهى:-

- نموذج نظرية الاستجابة للمفردة المستخدم.
- بنك للمفردات تم تدريجه.
- طريقة تقدير القدرة.
- طريقة اختيار المفردة الأولى لبدء الاختبار.
- طريقة اختيار المفردات التالية فى الاختبار.
- قاعدة إنهاء الاختبار.

(Weiss D.,2012,P1)

ويوجد لكل من النقاط السابقة العديد من الاختيارات مثل طبيعة الأجابة على المفردة (اختيار من متعدد/مفتوح) وبارامترات المفردة التى تؤثر فى احتمال اختيار الأجابة الصحيحة، مستودعات مفردات متغيرة فى الحجم وتصنيفات المحتوى والخصائص السيكوميترية للمفردة.

وعند تطوير النسخة الورقية إلى نسخة تكيفية محوسبة فينبغى وضع بعض المحكات لتقويم جودة الاختبار التكيفى المحوسب ويكون ذلك من خلال ثلاثة أبعاد هى:

١-البعد التربوى

يرجع الحكم على المحتوى Content إلى عدد وجودة مفردات بنك الأسئلة، وطريقة عرض المفردات Presentation (وسائط متعددة وأشكال أخرى) فينبغى أن تكون سهلة وبسيطة وواضحة، أيضا التتابع Sequencing فى عرض المفردات على المفحوص ويكون ذلك قائما على إستجابته السابقة، بالإضافة إلى التغذية المرتدة Feed back التى تعطى للمفحوص بناءا على استجابته، حيث أنها من الممكن أن ترشده أوتساعده وتوضح له تقدمه ونقاط قوته ونقاط ضعفه.

٢- البعد الفنى

تفاعل المفحوص مع واجهة الاستخدام User Interface مثل الجرافيك والتعامل مع لوحة المفاتيح والماوس أو شاشة اللمس، قابلية الاختبار للإصلاح أثناء التقويم أى قابليتة للتحقق(الثبات) ، فيجب أن يتجنب الأخفاق وأن يكون متاحا للمفحوص فى أى وقت، أيضاً القابلية للصيانة والإصلاح

Maintainablity بمعنى الوقت والجهد اللازمين لإزالة الأخطاء أو التحديث أو الإصلاح، فاعلية وكفاءة الاختبار (الأداء) Performance ، الوظيفية Functionality أى الوظائف المتاحة فى الاختبار مثل (التذكير-التنبيه)، وجود الأدوات Tools(ألة حاسبة-مسطرة-منقلة-كاميرا-مسجل صوت)، التوصيل Connectivity بمعنى القدرة على التفاعل والاتصال مع نظم البرمجيات المختلفة وذلك عن طريق شبكات الكمبيوتر بأنواعها، السرية Secuirty بمعنى الحماية من الدخول غير المشروع على نظام الاختبار، أوتعديل معلومات الموجودة بالاختبار، التكيف (الموائمة) Adaption أى قدرته على تكييف وموائمة الاختبار مع مستوى المفحوص.

٣-البعد الاقتصادى

التكاليف Costs من تشغيل وصيانة وتطبيق وتطوير، العقود والتراخيص وذلك خاص بالمؤسسات الربحية فيجب عليها إمداد المفحوص بالمعلومات المتعلقة برسوم الاختبار، تأثير وفاعلية التكلفة فيجب أن يرى المفحوص فاعلية وجدوى التكلفة التى دفعها لاختبار تكيفى محوسب.

(Economides A.A., Roupas C., 2007, PP70-87)

وفى الآونة الأخيرة تزايدت الدراسات التى اهتمت بالمقارنة بين الاختبار التكيفى المحوسب والاختبار الورقى التقليدى وذلك للتأكد من صدق وثبات الاختبار التكيفى المحوسب فى تقدير قدرة الأفراد وتتم المقارنة من خلال عاملين ترى الدراسات أنه من الممكن أن يؤثران على اداء المفحوص وهما أسلوب التطبيق (خطى/تكيفى) وطريقة التطبيق (ورقى/محوسب)، والجدير بالذكر بأن بعض الدراسات قامت بتثبيت أحد العوامل وهو طريقة التطبيق وفحص تأثير العامل الثانى (أسلوب التطبيق) على أداء المفحوص وذلك بالمقارن بين النسخة الخطية المحوسبة والنسخة التكيفية المحوسبة لنفس الاختبار.

كما أن هناك العديد من المحكات لمقارنة الاختبار الورقى بالاختبار التكيفى المحوسب وهى كالتالى:-

- **نسبة الأجابة الصحيحة**

تستخدم للمقارنة بين نوعى التطبيق (الورقى/التكيفى المحوسب) فى ضوء متوسط الصعوبة ومن المتوقع بأن يكون الاختبار التكيفى المحوسب متوسطه أكبر من الاختبار الورقى.

- **التحيز**

يعنى الفرق بين تقدير القدرة والقدرة الحقيقية، ويستخدم لتقييم دقة الاختبار فى تقدير االقدرة موضع الاهتمام فمعنى ان يكون التحيز قيمته منخفضة فإن ذلك يعنى دقة أكبر فى تقدير القدرة ومن المتوقع أن يكون الاختبار التكيفى المحوسب له مستوى من التحيز أقل من الاختبار الورقى وخاصة عند المستويات المتطرفة من القدرة.

- **خطأ القياس**

يتم حسابه عن طريق الخطا المعيارى للقياس ويستخدم فى تقييم قرار لتقديرات القدرات لكل اختباروكلماكان الخطا المعيارى أقل كلماكان تقدير القدرة متحرراً من الخطأ، ومن المتوقع أن يكون الاختبار التكيفى المحوسب أكثر تحرراً من الخطأ فى تقدير القدرة من الاختبار الورقى.

- **الكفاءة النسبية**

تستخدم للمقارنة بين معلومات الاختبار لكل من نوعى تطبيق الاختبار، ويتم حسابهاعن طريق قسمة معلومات الاختبار التكيفى المحوسب على معلومات الاختبار الورقى، ورياضياً فإن معلومات الاختبار هى مقلوب الجذر التربيعى للخطأ المعيارى.

- **استخدام مستودع المفردات**

وذلك بحساب نسبة عدد مرات تطبيق كل مفردة على المفحوصين (عدد مرات عرض المفردة على الأفراد)Item Exposure

(Widiatmo H.,2004,P74-76)

٨-الدراسات السابقة :

صنف الباحث الدراسات السابقة إلى محورين أساسين هما:
المحور الأول: دراسات تناولت المقارنة بين الاختبار التكيفى المحوسب والاختبار الورقى ببيانات واقعية.
المحور الثانى : دراسات تناولت المقارنة بين الاختبار التكيفى المحوسب والاختبار الورقى ببيانات واقعية وبيانات محاكاة للواقع .

وتمثلت أهم دراسات المحور الأول فى :-
- **دراسة أمين نور الدين (٢٠٠٢)**

بعنوان: فاعلية استخدام الاختبار الموائم باستخدام الاختبار المحوسب الآلى فى تقدير قدرة الأفراد وتحديد الخصائص السيكوميترية للمقياس

وهدفت الي :-
اختبار مدى فاعلية الاختبار التواؤمى المحوسب من حيث مؤشرات دقة القياس ودراسة خصائصه السيكومترية.

وتكونت عينه الدراسة من :-
(٥٠٠٠) طالب طالبة من الفرقة الثانية والثالثة والرابعة بكلية التربية جامعة عين شمس وذلك لتقدير بارمترات مفردات الاختبار، ثم تم تطبيق الاختبار بصورته التواؤمية المحوسبة على (٥٠٠) طالب وطالبة أخرين من مختلف التخصصات.

وتمثلت أدوات الدراسة في :-
اختبار رافن للمصفوفات المتدرجة بصوره المختلفة الملونة والعادية، والمتقدمة وطبق(١٣٧) مفردة من تلك الاختبارات ،وقد تم تقسيم المفردات بصورة مبدئية إلى اختبارين يطبق كل منهما بصورة منفصلة.

وتمثلت أهم نتائج الدراسة في :-
أن الاختبار التواؤمى المحوسب أكثر فعالية ويعطى دقة أكبر فى القياس ويوفر فى عدد مفردات الاختبار، كما أنه يعطى تقديرا للقدرة لايختلف متوسطه عن متوسط تقدير القدرة للاختبار المحوسب العادى.

-دراسة مارى بومريش Mary Pommerich(٢٠٠٧)
بعنوان: تأثير استخدام بارمترات المفردة المشتقة من التطبيق الورقى فى التطبيق التكيفى المحوسب
The Effect Of Using Item Parameters Calibarated from Paper Administrations in Computer Adaptive Test Administrations

وهدفت الي :-
معرفة الفروق فى تقديرات بارمترات المفردة عند اختلاف تطبيق الاختبار (ورقى – تكيفى محوسب) لصور اختبارية فى الفهم القرائى فى مادة العلوم فى مجال القراءة والاستدلال العلمى.

وتكونت عينه الدراسة من :-
(٦١٩٥) طالبا بالمرحلة التعليمية (١١-١٢) وقسمت العينة إلى (٣١٧١) طالب طبق عليهم اختبار القراءة بصورتيه (الورقية – التكيفية المحوسبة)، وعدد (٣٠٢٤) طالب طبق عليهم اختبار الاستدلال العلمى بصورتيه (الورقية – التكيفية المحوسبة).

وتمثلت أدوات الدراسة في :-
اختبار القراءة ويتكون من (أربع) قطع للفهم القرائى وكل قطعة بها مفردات اختيار من متعدد على موضوع القطعة، اختبار الاستدلال العلمى ويتكون من (سبع) قطع وبكل قطعة من (٥-٧) أسئلة

اختيار من متعدد على موضوع القطعة، تم تدريج الاختبارين بإستخدام النموذج الثلاثى البارامتر لنظرية الاستجابة للمفردة.

وتمثلت أهم نتائج الدراسة في :-

تقدير بارمترات المفردة يختلف اختلافاً بسيطاً باختلاف نوع التطبيق(ورقى – تكيفى محوسب) على الرغم من أن النتائج وجدت أن الثبات من الممكن أن يقل فى بعض الحالات وذلك فى حالة تقدير البارامترات عند التطبيق الورقى للاختبار.

وتمثلت أهم دراسات المحور الثانى فى :-

- **دراسة فيكينو وماكبرايد (1997)**

تم تطبيق نسخة من اختبار ورقى للطلاب واختبار تكيفى محوسب لنفس الطلاب وذلك بترتيب عشوائى ومن خلال النتائج اتضح أن الاختبار التكيفى المحوسب أظهر نفس تقديرات قدرة الطلاب على الاختبار الورقى وذلك بالرغم من قصر طول الاختبار.

وعملت تلك الدراسات على استخدام المحاكاة كخطوة أولى وتحليل النتائج المستمدة من الاختبارات الواقعية كخطوة تالية وذلك لأهمية دراسات المحاكاة فى تقييم مكونات تقنية هامة

تعليق الباحث على نتائج الدراسات السابقة وإمكانية الاستفادة منها في البحث الحالي :

- أوضحت الدراسات عدم وجود فروق جوهرية بين التطبيق الورقى والمحوسب لنفس الاختبار.
- فاعلية الاختبار التكيفى المحوسب فى زيادة دقة القياس ولكنه يحتاج إلى عينة كبيرة من الأفراد بالإضافة إلى التكلفة المادية والمجهود المبذول.

٩- خطة الدراسة وإجراءاتها

٩-١- المنهج المستخدم

استخدمت هذه الدراسة السيكومترية المنهج التجريبي حيث يتمثل المتغير المستقل فى أسلوب تطبيق المقياس (ورقى/تكيفى محوسب) ويتمثل المتغير التابع فى تقدير قدرة الأفراد.

٩-٢ عينة الدراسة

تتضمن العينة (٩٥) طالب من مراحل دراسية مختلفة وهى كالتالى:-

جدول (٢) يوضح توزيع أفراد عينة المقارنة بين الاختبار الورقى والاختبار التكيفى المحوسب تبعا للمراحل الدراسية

م	المرحلة الدراسية	عدد الطلاب
١	الرابع الابتدائى	١٦
٢	الخامس الابتدائى	١٦
٣	السادس الايتدائى	١١
٤	الأول الاعدادى	٢٠
٥	الثانى الاعدادى	١٠
٦	الثالث الاعدادى	١٤
٧	الأول الثانوى	٤
٨	الثانى الثانوى	٣
المجموع		٩٤

٩-٣ -أدوات الدراسة

استخدم الباحث مقياس الاستدلال اللفظى (إعداد الباحث) ويتكون من (٩٦) مفردة متدرجة بإستخدام نظرية الاستجابة للمفردة (النموذج الثلاثى البارامتر) تغطي المستويات المختلفة على

مدى متصل تقديرات المفردات وتتمثل الخصائص السيكوميترية لمفردات المقياس فى الجدول التالى:-

جدول (٣) الخصائص السيكومترية لمقياس الاستدلال اللفظى

بارامتر المفردات	القيمة الصغرى	القيمة العظمى	المتوسط	الانحراف المعيارى	متوسط الخطأ المعيارى
التمييز	0,44	2,36	1,20	0,39	0,22
الصعوبة	1,38-	2,28	0,57	0,81	0,14
التخمين	0,08	0,43	0,18	0,06	0,04

يتضح من الجدول أن قيم الخطأ المعياري لتقديرات المفردات تعتبر منخفضة بوجه عام حيث ويزيد ذلك من ثبات المقياس اللفظى.

صدق وثبات مقياس الاستدلال اللفظى

اعتمد تحديد الصدق على مايلى:

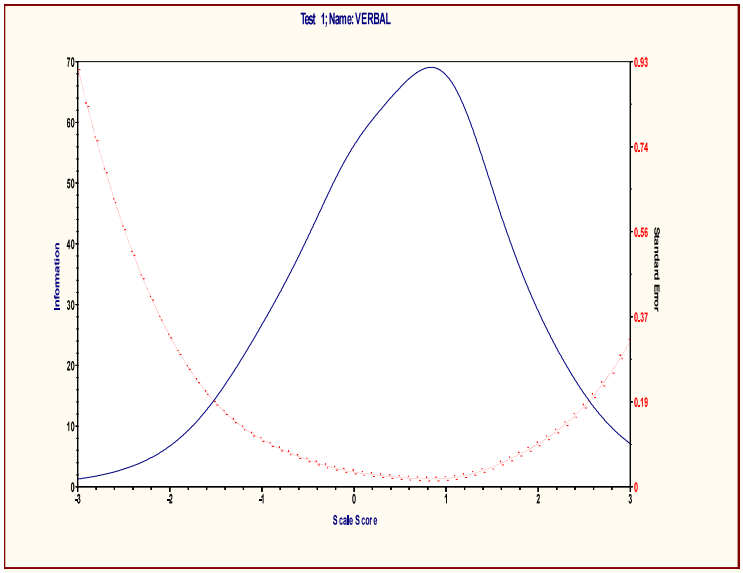
- صدق المحتوى (صدق البناء)
- صدق التدرج (صدق الملاءمة)
- صدق تعريف المتغير (عدم وجود فجوات)
- صدق التمييز (حساسية المقياس للتمييز بين مستويات الأفراد)

أما تحديد الثبات فقد اعتمد على:

- ثبات التدرج
- حساب الخطأ المعياري لتقديرات كل من معالم المفردات والأفراد

وتستخدم دالة معلومات الاختبار Test Information Function والمنحنى الدال عليها كمؤشر لثبات تقديرات القدرة على اختبار ما،حيث أنها تشير إلى كمية المعلومات التى يقدمها الاختبار عند المستويات المختلفة للقدرة ، وهى عبارة مقلوب مربع الخطأ المعياري والذي يشير إلى معامل الخطأ المعياري في تقدير قدرات العينة، ومن ثم فإن كمية المعلومات المقدمة بواسطة الاختبار عند مستوى القدرة (θ) تكون مرتبطة عكسياً مع الدقة عند هذه النقطة على متصل القدرة (θ). **(هامبلتون وآخرون Hambleton & et al ،١٩٩١: ٩٣)**

وتعتبر دوال معلومات المفردات ودالة معلومات الاختبار أحد الاستراتيجيات المستخدمة لانتقاء المفردات فى الاختبار التكيفى المحوسب حيث يتم ترشيح المفردة المناسبة لقدرة الطالب بناءاً على اختيار المفردة التى تعطى أكبر كم من المعلومات عند مستوى قدرة الطالب، وقد اعتمد الباحث على دالة معلومات الاختبار في تحديد ثبات تقديرات القدرة على الصور الاختبارية التي تكون منها مقياس الاستدلال اللفظى **حيث بلغ معامل ثبات الأفراد (٩١%) ومعامل ثبات المفردات (٩٨%)** وفيما يلي شكل دالة المعلومات لمقياس الاستدلال اللفظى.



شكل(٢) دالة معلومات مقياس الاستدلال اللفظى (النسخة الورقية)

يتضح من الشكل السابق أن أكبر كم من المعلومات حوالى (٦٩) وكان عند مستوى قدرة (١) لوجيت أيضاً يتضح إنخفاض الخطأ المعيارى ويعد ذلك مؤشراً لدقة القياس .

٩-٤ خطوات الدراسة

- **تجهيز المقياس للتطبيق التكيفى المحوسب بإستخدام برنامج fast test professional وتتضمن ذلك:-**

- تجهيز برنامج الاختبار التكيفى المحوسب وتخزين مفردات مقياس الاستدلال اللفظى ومؤشراتها الاحصائية بعد التدريج النهائى واستبعاد المفردات غير الملائمة للقياس.
- تحديد **محك بداية الاختبار** وهى متوسط القدرة باللوجيت(صفر).
- تحديد **محك اختيار المفردات التالية** وهو المفردة التى تعطى أكبر كم من المعلومات عند مستوى قدرة المفحوص Maximum Information Function.
- تحديد **محك إنهاء الاختبار** وهو أن يكون انهاء الاختبار عندما يكون الخطأ المعيارى يساوى(٠.٢٠).

- **تطبيق المقياس ورقياً وتكيفياً محوسباً**

قام الباحث بتطبيق مفردات المقياس ككل(البنك) ورقياً وتطبيق المقياس بالأسلوب التكيفى المحوسب ثم تقدير قدرة الطلاب فى كلا من نوعى التطبيق والمقارنة بينهما، ولتحقيق موضوعية التطبيق أجرى الباحث التطبيق بصورة تبادلية (تكيفى محوسب / ورقى – ورقى /تكيفى محوسب).

١٠- نتائج الدراسة ومناقشتها :-

فيما يلى عرض لنتائج الدراسة وذلك للأجابة على سؤال الدراسة.

وتتضمن الوصف الاحصائى للبيانات الخاصة بكل من الأسلوبين الورقى والتكيفى المحوسب وكذا نتائج المقارنة بينهما

أولاً: الوصف الاحصائى لكل من الأسلوبين (ورقى/تكيفى محوسب)

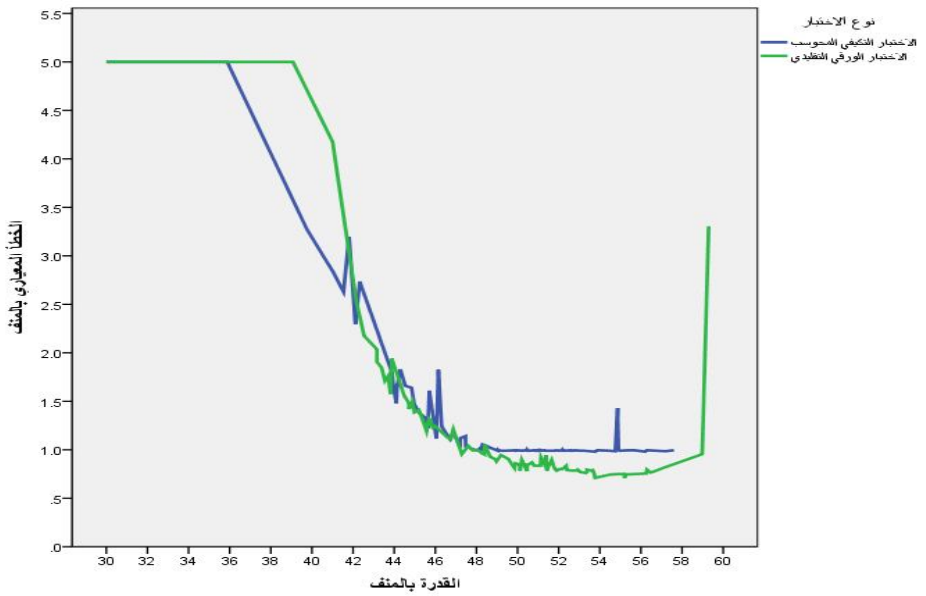
فيما يلى متوسط تقديرات القدرة على المقياس ورقيا وتكيفيا بالحاسب أيضا متوسط الخطأ المعيارى والزمن الكلى وعدد المفردات لكل مرحلة دراسية ويوضح ملحق (١) بملاحق الدراسة تقدير قدرات الأفراد بالمنف على مقياس الاستدلال اللفظى بالأسلوب الورقى والأسلوب التكيفى المحوسب حيث يتضح من جدول (٤)

أنه كلما كانت المرحلة الدراسية أكبر كلما قل عدد المفردات المعروضة للطالب وهذا ليس شرطا ولكنه يختلف من طالب لأخر فى المراحل المختلفة وفى كل مرحلة وذلك بناءا على قدرته، أما بالنسبة لمتوسط الخطأ المعيارى لتقدير القدرة فى الاختبار التكيفى المحوسب فالباحث ثبت قيمة الخطأ المعيارى كمحك لإنهاء الأختبار التكيفى المحوسب بقيمة (0,20) لوجيت وتقابل (1,00) منف لأنها تقابل معامل ثبات عالى وهى أحدى الاستراتيجيات المستخدمة لأنها الاختبار، اما بالنسبة لزمن تطبيق الاختبار التكيفى المحوسب فيتراوح من (١٧) إلى(٣٩) دقيقة، **ويتضح ذلك من الجدول التالى :**

جدول (٤) متوسط قدرة افراد عينة المقارنة بين الاختبار الورقى والاختبار التكيفى المحوسب تبعا للمرحلة الدراسية

م	المرحلة الدراسية	الاختبار الورقى		الاختبار التكيفى المحوسب			
		متوسط القدرة بالمنف	متوسط الخطأ المعيارى بالمنف	متوسط القدرة بالمنف	متوسط الخطأ المعيارى بالمنف	متوسط العدد الكلى للمفردات	متوسط الزمن بالدقيقة
١	الرابع الابتدائى	44,54	1,90	46,07	1,71	79	39
٢	الخامس الابتدائى	46,29	1,99	43,22	2,22	70	32
٣	السادس الابتدائى	43,42	2,36	46,35	1,70	67	24
٤	الأول الاعدادى	50,23	0,99	49,94	1,24	52	20
٥	الثانى الاعدادى	51,02	0,89	51,04	1,05	43	17
٦	الثالث الاعدادى	52,83	0,82	52,27	1,00	38	20
٧	الأول الثانوى	54,49	0,79	54,00	1,00	32	27
٨	الثانى الثانوى	56,30	0,83	55,70	1,00	30	20

وفيما يلى شكل يوضح العلاقة بين القدرة والخطأ المعيارى بالمنف لكلا من الاختبار التكيفى المحوسب والاختبار الورقى

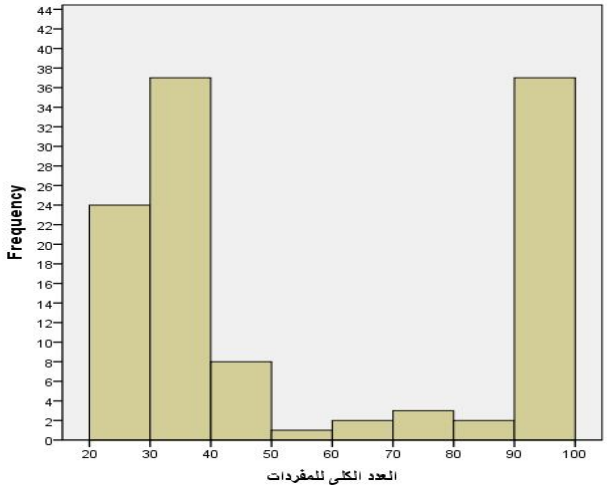


شكل (٣) العلاقة بين القدرة والخطأ المعيارى بالمنف لكل من الاختبار التكيفى المحوسب والاختبار الورقى

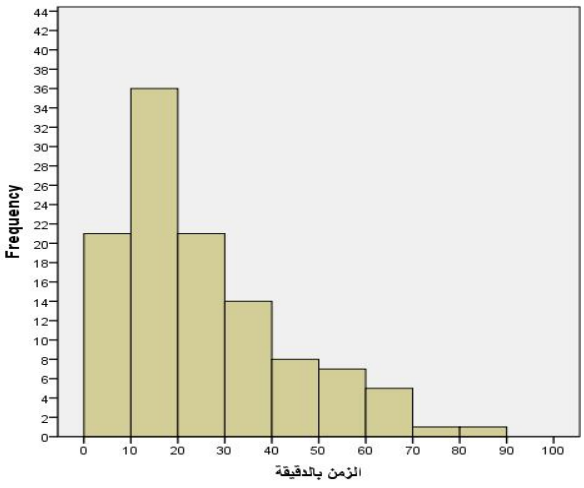
يتضح من الشكل السابق انخفاض الخطأ المعيارى لتقدير القدرة خاصة عند المستويات المنخفضة من القدرة ويوجد ثبات إلى حدما فى قيمة الخطأ المعيارى للقدرة عند

المستويات العليا وذلك فى الاختبار التكيفى المحوسب وقد يرجع ذلك إلى طبيعة التطبيق فالباحث قد حدد محك إنهاء الاختبار بقيمة ثابتة معينة للخطأ المعيارى، ويدل ذلك الشكل على دقة الاختبار التكيفى المحوسب فى القياس عن الاختبار الورقى التقليدى.

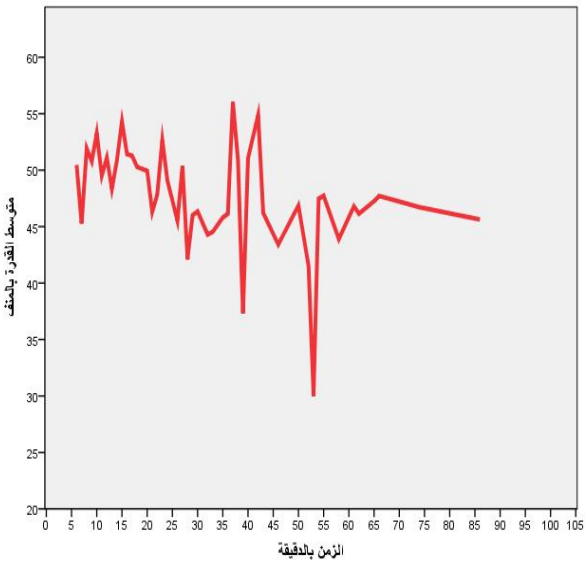
بسبب أن الأسلوب التكيفى المحوسب يُفصَل الاختبار حسب قدرة الطالب فبالتالى تختلف عدد المفردات وزمن تطبيق الاختبار من طالب لأخر لذلك حاول الباحث ايجاد علاقة بين كل من الزمن وعدد المفردات ومتوسط قدرة الطالب بالمنف وهو ماتوضحه الأشكال التالية:-



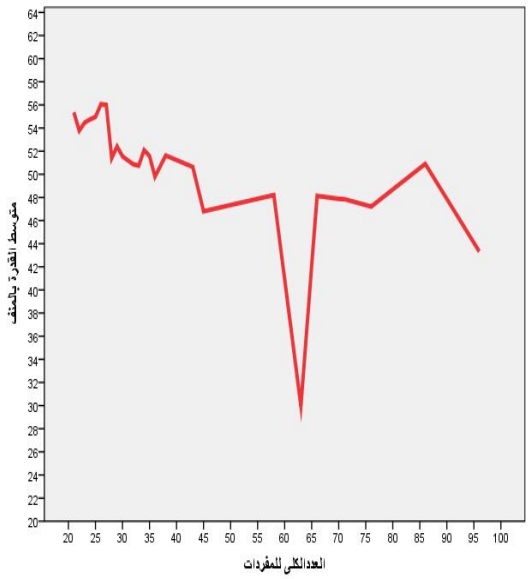
شكل (٤)
تكرار الأفراد مقابل العدد الكلى لمفردات الاختبار التكيفى المحوسب



شكل (٥)
تكرار الأفراد مقابل الزمن الكلى للاختبار التكيفى المحوسب



شكل (٦)
العلاقة بين متوسط القدرة بالمنف والعدد الكلى للمفردات المجاب عليها فى الاختبار التكيفى المحوسب



شكل (٧)
العلاقة بين متوسط القدرة بالمنف وزمن الاختبار التكيفى المحوسب

يتضح من الأشكال السابقة أن العدد الكلى لمفردات الصورة الاختبارية التكيفية المحوسبة يتراوح بين (٢٠) إلى (٩٦) مفردة وتراوحت عدد المفردات لنسبة كبيرة من أفراد العينة بين (٣٠) إلى (٤٠) مفردة كما هو موضح بشكل (٤) ، أما بالنسبة للزمن الكلى المستغرق فى الأجابة على الصورة الاختبارية التكيفية المحوسبة فيختلف من طالب لأخر وقد تراوح بين (١٠) إلى (٩٠) دقيقة وكانت نسبة كبيرة من أفراد العينة قد استغرقوا من (١٠) إلى (٢٠) دقيقة كماهو موضح بشكل (5)، أيضاً لاحظ الباحث أنه كلما زاد متوسط القدرة كلما انخفض عدد مفردات الصورة الاختبارية أيضاً ينخفض الزمن المستغرق للأجابة وهو ماتوضحه الأشكال (٦) ، (٧).

ثانياً: نتائج المقارنة بين تقديرات قدرة الأفراد على كل من الأسلوب الورقى والأسلوب التكيفى المحوسب

استخدم الباحث حوالى (١٠%) من العينة وذلك للتحقق من مدى تعادل القدرة على الأسلوبين وفيما يلى جدول يوضح ذلك:-

جدول (٤) تقديرات القدرة والخطأ المعيارى لبعض أفراد العينة والفرق بين تقدير القدرة ومجموع الخطأ المعيارى

م	رقم الطالب	الاختبار الورقى		الاختبار التكيفى المحوسب		الفرق فى تقدير القدرة	مجموع الخطأ المعيارى
		القدرة بالمنف	الخطأ المعيارى بالمنف	القدرة بالمنف	الخطأ المعيارى بالمنف		
١	٢٦	٤٧.٦	١.٠٤	٥٠.٩	٠.٩٩	٣.٣	٢.٠٣
٢	١٣٤	٤٢.٥٥٣	٢.١٧٥	٣٥.٨٨	٥	٦.٦٥٥	٧.١٧٥
٣	٥١	٤١.٠٠٥	٤.١٧٥	٤٣.٩	١.٧٩	٢,٨٩٥	٥.٩٦٥
٤	٧	٤٨.٩٧٥	٠.٨٨	٤٩.٠٤٥	١	٠.٠٧	١.٨٨
٥	٢	٤٥.٧٦	١.٣٠٥	٤٤.١٤	١.٦٦٥	١.٦٢	٢,٩٧
٦	١٣	٥٢.٩٢٥	٠.٧٩٥	٥٢.٥٧	٠.٩٩	٠.٣٥٥	١.٧٨٥
٧	١٢	٥٥.٢٤٥	٠.٧٤٥	٥٥.٣٦	٠.٩٩٥	٠.١١٥	١.٧٤
٨	٣١	٥٢.٦٤٥	٠.٧٨٥	٥١.٠٥	١	١.٥٩٥	١.٧٨٥
٩	٤٧	٥٦.٥٥٥	٠.٧٧	٥٥.٦٣	٠.٩٩٥	٠.٩٢٥	١.٧٦٥

يتضح من الجدول السابق تعادل تقديرات قدرة الأفراد على الاختبار التكيفى المحوسب مع الاختبار الورقى فى جميع المراحل الدراسية المختلفة.

تعليق عام على نظام الاختبارات التكيفية المحوسبة

- تتميز الاختبارات التكيفية المحوسبة بأنها جعلت العملية الاختبارية ديناميكية وتقنية بدلا من اختبارات الورقة والقلم الاستاتيكية التى تسهم فى كثير من الأحيان فى إثارة عوامل القلق وخفض الدافعية.
- تطور العديد من الدول المتقدمة من خوارزميات برامج الحاسب الآلى لتنفيذ نظم الاختبارات التكيفية المحوسبة، حيث انتشر استخدام تلك الاختبارات لتشمل استخدامها فى اختبارات مزاولة مهنتى الطب والتمريض.
- نظر لأن معظم الاختبارات التكيفية المحوسبة اختبارات قوة Power Tests، إلا أن السرعة تعد مكونا هاما من مكونات القدرة، لذلك تبدو الحاجة لتطوير نماذج الاختبارات التكيفية المحوسبة التى تسمح للفرد بتكييف استجابته الكامنة أو زمن الاستجابة.
- توجد جهود بحثية لاستخدام نماذج الاستجابة المتعددة الأقسام Polytomus IRT Models مثل نموذج الاستجابة المتدرجة ونموذج الدرجات الجزئية.

دراسات وبحوث مقترحة :

من خلال نتائج الدراسة الحالية، يمكن للباحث أن يقترح بعض الدراسات والبحوث التي قد تُسهم في تطوير مجال الدراسة الحالية، وذلك كما يلي:-

- اجراء المزيد من الدراسات التى تقوم على المقارنة بين كلا من الاختبار التقليدى والاختبار التكيفى (ورقياً) ثم المقارنة بين نفي النوعين ولكن باستخدام الحاسب الآلى وذلك لنفس مجموعة المفردات.
- محاولة اجراء الدراسات للتعرف على العوامل المؤثرة على أداء الطالب عند أدائه للاختبار التكيفى المحوسب.
- محاولة انشاء اختبارات تكيفية محوسبة تتضمن مفردات استجابة متعددة ومفردات مفتوحة
- يمكن استخدامها فى تشخيص الأشخاص ذوى صعوبات التعلم.
- تصميم الاختبارات بحيث تتيح القدرة للصم والبكم وذوى الاحتياجات الخاصة فى الأجابة على الاختبارات.
- من الممكن تصميم بعض الاختبارات العملية وعمل مواقف محاكاة للأشياء الواقعية التى يستحيل اختبار الطالب بالطريقة التقليدية فمثلا من الممكن تصميم مواقف اختبارية بتكنولوجيا (3D) وتكنولوجيا (7D).
- محاولة دراسة التغيرات الفسيولوجية فى المخ عند أداء الطالب على مفردات الاختبار.
- تفعيل استخدام الأسلوب التكيفى المحوسب فى إعداد اختبارات قدرات للقبول فى الجامعات .

قائمة المراجع

أحمد المغربى (٢٠١٠). **مقاييس واختبارات الذكاء فى ميزان نظرية الذكاء الكلى**.القاهرة : دار الفجر.

أمين محمد نور الدين (٢٠٠٢) . **فاعلية استخدام الاختبار الموائم باستخدام الاختبار المحوسب الآلى فى تقدير قدرة الأفراد وتحديد الخصائص السيكوميترية للمقياس**، رسالة دكتوراة ، غير منشورة ، كلية التربية، جامعة عين شمس.

أمينة محمد كاظم (١٩٨٩) . **مقاييس القدرات ، تعريب وتجريب للمقاييس البريطانية للقدرات ، الدليل ، المقياس الأول سرعة تجهيز المعلومات** ، القاهرة ، المركز القومى للأمتحانات والتقويم التربوى.

أمينة محمد كاظم (١٩٩٦). **نماذج السمات الكامنة في: أنور محمد الشرقاوي، سليمان الخضري الشيخ، أمينة محمد كاظم، نادية محمد عبد السلام، اتجاهات معاصرة في القياس والتقويم النفسي والتربوي**. القاهرة: مكتبة الأنجلو المصرية.

بيتير ديفيدسون ، ترجمة عبد العزيز الدامغ (٢٠٠٦) . **الاختبارات المعتمدة على الحاسب الألى** . جامعة الملك سعود ، الرياض.

سمية على عبد الوارث أحمد (٢٠١١). **البحث التربوى والنفسى: دليل تصميم البحوث**: القاهرة: مكتبة الأنجلو المصرية.

فتحى عبد الرحمن جروان (١٩٩٩) . **تعليم التفكير : مفاهيم وتطبيقات** . العين : دار الكتاب الجامعى.

COLVIN K. (2014). **EFFECT OF AUTOMATIC ITEM GENERATION ON ABILITY ESTIMATES IN A MULTISTAGE TEST.** Doctoral Dissertation, University of Massachusetts Amherst.

Douglas A. Strader(2012). **Comparability of Computer Delivered versus Traditional Paper and Pencil Testing,**Phd dissertation,Wilmington University**.**

Economides A.A., Roupas C. (2007). **Evaluation of Computer Adaptive Testing Systems.** International Journal of Web, Web-Based Learning and Teaching Technologies.(2) 1.

Hambelton, R. , Swaminathan, H. & Rogers, H. J. (1991) . **Fundamentals Of Item Response Theory** , California :SAGE Publications.

Linacre, J. M.(2000).**Computer-Adaptive Testing:A Methodology Whose Time Has Come**. Development of Computerized Middle School Achievement Test in Korean,Seoul,South Korea :Komesa Press,MESA Memorandum No.69.
Pommerich M. . (2007). **The Effect Of Using Item Parameters Calibarated from Paper Administrations in Computer Adaptive Test Administrations**. Journal of Technology, Learning, and Assessment , 5 (7) .
Schumacker,R.E.(2005).**ItemResponseTheory,AppliedMeasurementAssociates,**WhitePapers,from http://www.appliedmeasurementassociates.com/articles/ITEM%20RESPONSE%20THEORY.pdf.
Steven L. Wise. .(2014). **Adaptive Testing and Unmotivated Examinees.** Journal of Computerized Adaptive Testing, 2 (1) .
Thissen D.,Reeve B.B.&Bjorner J.B& others.(2007). **Methodological issues for building item banks and computerized adaptive scales**.Qual Life Res(2007)16:109-119,Springer Science&Business Media B.V.
Thompson N.A. ( ),**Adaptive Testing:Is It Right for me?** ,White papers Assessment Systems Coroporation.
Wang J. (2009). **Using real-data simulations to compare computer adaptive testing and static short-form administrations of an upper extremity item bank**. Unpublished Doctoral Dissertation, University of Florida.
Weiss,D.(2012). **Termination Criteria in Computerized Adaptive Tests: Do Variable-Length CATs Provide Efficient and Effective Measurement?.** Journal of Computerized Adaptive Testing, 1 (1) .
Widiatmo,H.(2004). **A SIMULATION AND EVALUATION OF COMPUTERIZED ADAPTIVE TESTING DESIGNS FOR THE VERBAL BATTERY OF THE COGNITIVE ABILITIES TEST (CogAT)**. Doctoral Dissertation, Graduate School of Iowa University.

ملاحق الدراسة

ملحق (١)

تقدير قدرات الأفراد بالمنف على مقياس الاستدلال اللفظى بالأسلوب الورقى والأسلوب التكيفى المحوسب

م	رقم الطالب	الاختبار الورقى		الاختبار التكيفى المحوسب		الفرق فى القدرة	مجموع الخطأ المعيارى بالمنف	الحكم
		القدرة بالمنف	الخطأ المعيارى بالمنف	القدرة بالمنف	الخطأ المعيارى بالمنف			
1	1	51.14	0.92	47.475	1.035	3.665	1.955	0
2	2	45.76	1.31	44.14	1.665	1.62	2.97	1
3	5	56.29	0.80	57.185	0.985	0.895	1.78	1
4	7	48.98	0.88	49.045	1	0.07	1.88	1
5	9	52.86	0.79	50.865	0.995	1.99	1.78	0

1	7.525	7.325	5	34.84	2.53	42.17	10	6
1	1.74	0.46	0.98	53.78	0.76	53.32	11	7
1	1.74	0.115	0.995	55.36	0.75	55.25	12	8
1	1.815	1.505	0.995	49.875	0.82	51.38	23	9
1	1.785	0.355	0.99	52.57	0.80	52.93	13	10
0	1.95	2.695	0.995	56.285	0.96	58.98	18	11
0	2.09	4.595	1.145	46.8	0.95	51.40	20	12
1	1.79	1.065	0.99	53.235	0.80	52.17	22	13
1	2.705	0.925	1.22	45.63	1.49	44.71	24	14
1	2.795	0.52	1.375	45.24	1.42	44.72	25	15
0	2.55	4.62	0.99	49.1	1.56	44.48	157	16
0	6.94	13.89	5	30	1.94	43.89	154	17
0	6.76	13.71	5	30	1.76	43.71	36	18
0	6.845	13.39	5	30	1.85	43.39	39	19
1	3.375	0.605	1.665	44.155	1.71	43.55	41	20
1	1.76	1.14	0.995	57.6	0.77	56.46	42	21
1	1.765	0.93	0.995	53.975	0.77	53.05	43	22
1	1.86	1.17	1	51.065	0.86	49.90	44	23
0	1.795	1.87	0.995	53.87	0.80	52.00	45	24
1	1.7	0.73	0.99	54.47	0.71	55.20	46	25
0	2.695	3.57	1.64	44.84	1.06	48.41	66	26
1	1.785	1.44	0.99	50.54	0.80	51.98	73	27
1	5.965	2.895	1.79	43.9	4.175	41.005	51	28
0	3.465	11.71	1.43	54.865	2.035	43.16	63	29
1	1.93	0.78	1	47.875	0.93	48.655	158	30

ملحق (١)

تقدير قدرات الأفراد بالمنف على مقياس الاستدلال اللفظى بالأسلوب الورقى والأسلوب التكيفى المحوسب

م	رقم الطالب	الاختبار الورقى		الاختبار التكيفى المحوسب		الفرق فى القدرة	مجموع الخطأ المعيارى بالمنف	الحكم
		القدرة بالمنف	الخطأ المعيارى بالمنف	القدرة بالمنف	الخطأ المعيارى بالمنف			
1	1	51.14	0.92	47.475	1.035	3.665	1.955	0
2	2	45.76	1.31	44.14	1.665	1.62	2.97	1
3	5	56.29	0.80	57.185	0.985	0.895	1.78	1
4	7	48.98	0.88	49.045	1	0.07	1.88	1
5	9	52.86	0.79	50.865	0.995	1.99	1.78	0
6	10	42.17	2.53	34.84	5	7.325	7.525	1
7	11	53.32	0.76	53.78	0.98	0.46	1.74	1
8	12	55.25	0.75	55.36	0.995	0.115	1.74	1

1	1.815	1.505	0.995	49.875	0.82	51.38	23	9
1	1.785	0.355	0.99	52.57	0.80	52.93	13	10
0	1.95	2.695	0.995	56.285	0.96	58.98	18	11
0	2.09	4.595	1.145	46.8	0.95	51.40	20	12
1	1.79	1.065	0.99	53.235	0.80	52.17	22	13
1	2.705	0.925	1.22	45.63	1.49	44.71	24	14
1	2.795	0.52	1.375	45.24	1.42	44.72	25	15
0	2.55	4.62	0.99	49.1	1.56	44.48	157	16
0	6.94	13.89	5	30	1.94	43.89	154	17
0	6.76	13.71	5	30	1.76	43.71	36	18
0	6.845	13.39	5	30	1.85	43.39	39	19
1	3.375	0.605	1.665	44.155	1.71	43.55	41	20
1	1.76	1.14	0.995	57.6	0.77	56.46	42	21
1	1.765	0.93	0.995	53.975	0.77	53.05	43	22
1	1.86	1.17	1	51.065	0.86	49.90	44	23
0	1.795	1.87	0.995	53.87	0.80	52.00	45	24
1	1.7	0.73	0.99	54.47	0.71	55.20	46	25
0	2.695	3.57	1.64	44.84	1.06	48.41	66	26

ملحق (١)

تقدير قدرات الأفراد بالمنف على مقياس الاستدلال اللفظى بالأسلوب الورقى والأسلوب التكيفى المحوسب

م	رقم الطالب	الاختبار الورقى		الاختبار التكيفى المحوسب		الفرق فى القدرة	مجموع الخطأ المعيارى بالمنف	ك
		القدرة بالمنف	الخطأ المعيارى بالمنف	القدرة بالمنف	الخطأ المعيارى بالمنف			
27	73	51.98	0.80	50.54	0.99	1.44	1.785	1
28	51	41.005	4.175	43.9	1.79	2.895	5.965	1
29	63	43.16	2.035	54.865	1.43	11.71	3.465	0
30	158	48.655	0.93	47.875	1	0.78	1.93	1
31	34	49.19	0.945	48.29	1.055	0.9	2	1
32	37	50.125	0.785	49.915	0.99	0.21	1.775	1
33	38	31.09	5	44.3	1.83	13.21	6.83	0
34	47	56.555	0.77	55.63	0.995	0.925	1.765	1
35	48	53.77	0.71	53.105	0.99	0.665	1.7	1
36	49	53.375	0.795	56.205	0.98	2.83	1.775	0
37	55	50.7	0.87	48.13	1	2.57	1.87	0
38	56	50.8	0.835	51.28	0.995	0.48	1.83	1
39	61	50.465	0.845	48.205	0.995	2.26	1.84	0
40	67	52.365	0.835	49.98	0.995	2.385	1.83	0
41	68	50.115	0.845	52.21	0.99	2.095	1.835	0

1	1.775	0.59	0.99	52.01	0.785	51.42	69	42
0	1.82	2.795	0.99	52.665	0.83	49.87	64	43
1	2.86	1.9	1.63	43.97	1.23	45.87	52	44
1	1.99	0.48	1.015	47.72	0.975	48.2	121	45
1	1.89	0.47	0.99	49.395	0.9	48.925	122	46
1	1.87	0.28	1	50.46	0.87	50.18	123	47
0	2.655	3.275	1.66	44.535	0.995	47.81	124	48
1	2.24	0.14	1.11	46.715	1.13	46.575	127	49
0	1.945	1.965	0.995	50.36	0.95	48.395	128	50
1	1.745	0.285	0.995	54.905	0.75	55.19	129	51

ملحق (١)

تقدير قدرات الأفراد بالمنف على مقياس الاستدلال اللفظى بالأسلوب الورقى والأسلوب التكيفى المحوسب

الحكم	مجموع الخطأ المعيارى بالمنف	الفرق فى القدرة	الاختبار التكيفى المحوسب		الاختبار الورقى		رقم الطالب	م
			الخطأ المعيارى بالمنف	القدرة بالمنف	الخطأ المعيارى بالمنف	القدرة بالمنف		
1	2.385	1.48	1.14	47.47	1.245	45.99	131	52
1	7.295	5.04	2.295	42.11	5	37.07	132	53
1	1.74	0.005	0.99	54.96	0.75	54.955	133	54
1	3.055	0.29	1.48	44.105	1.575	43.815	135	55
0	6.14	18.27	2.835	41.025	3.305	59.295	137	56
1	2.88	0.23	1.47	44.975	1.41	45.205	138	57
1	2.645	0.15	1.325	45.53	1.32	45.68	139	58
0	1.775	2.355	0.985	54.75	0.79	52.395	140	59
1	2.735	1.43	1.24	46.31	1.495	44.88	141	60
1	2.305	0.475	1.115	46.05	1.19	45.575	145	61
1	1.77	0.61	0.99	51.075	0.78	50.465	146	62
0	2.375	6.85	0.99	51.83	1.385	44.98	147	63
0	6.61	8.35	1.61	45.71	5	37.36	148	64
1	8.27	2.285	3.27	39.765	5	37.48	150	65
1	1.98	0.08	1.025	47.205	0.955	47.285	151	66
1	7.735	3.25	2.735	42.325	5	39.075	152	67
1	2.115	0.855	1.12	47.22	0.995	48.075	155	68
1	1.905	1.7	1	47.855	0.905	49.555	27	69
1	1.805	0.765	0.995	52.555	0.81	51.79	28	70
0	1.775	2.38	0.985	56.035	0.79	53.655	29	71
1	1.785	1.595	1	51.05	0.785	52.645	31	72
1	1.77	1.195	1	52.18	0.77	53.375	33	73
0	1.735	2.215	0.99	52.31	0.745	54.525	70	74
1	1.81	0.525	0.995	50.95	0.815	51.475	71	75
1	1.825	0.625	0.99	50.49	0.835	51.115	72	76
1	1.835	0.27	0.99	50.235	0.845	50.505	74	77

ملحق (١)

تقدير قدرات الأفراد بالمنف على مقياس الاستدلال اللفظى بالأسلوب الورقى والأسلوب التكيفى المحوسب

م	رقم الطالب	الاختبار الورقى		الاختبار التكيفى المحوسب		الفرق فى القدرة	مجموع الخطأ المعيارى بالمنف	الحكم
		القدرة بالمنف	الخطأ المعيارى بالمنف	القدرة بالمنف	الخطأ المعيارى بالمنف			
٨٧	75	53.59	0.785	51.89	0.995	1.7	1.78	1
٧٩	77	52.975	0.785	51.325	0.99	1.65	1.775	1
٠٨	76	51.895	0.785	52.165	0.995	0.27	1.78	1
١٨	78	51.66	0.895	49.03	0.99	2.63	1.885	0
٢٨	115	30	5	41.805	3.195	11.81	8.195	0
٣٨	118	30	5	39.825	3.25	9.825	8.25	0
٤٨	120	43.165	1.905	41.535	2.63	1.63	4.535	1
٥٨	26	47.6	1.04	50.9	0.99	3.3	2.03	0
٦٨	57	53.715	0.76	52.76	0.995	0.955	1.755	1
٧٨	16	56.25	0.755	56.345	0.995	0.095	1.75	1
٨٨	17	57.12	0.815	55.67	0.995	1.45	1.81	1
٨٩	62	46.875	1.21	46.15	1.825	0.725	3.035	1
٠٩	134	42.535	2.175	35.88	5	6.655	7.175	1
١٩	58	49.85	0.81	50.805	0.995	0.955	1.805	1
٢٩	59	49.705	0.855	50.45	0.995	0.745	1.85	1
٣٩	60	50.2	0.895	48.17	1.01	2.03	1.905	0
٤٩	50	46.75	1.11	46.87	1.185	0.12	2.295	1
٥٩	65	46.825	1.1	49.12	1	2.295	2,1	0

الحكم: أى أن الفرق بين تقدير القدرتين أقل من أو يساوى مجموع الخطأ المعيارى ممايدل على تعادل تقدير القدرتين

(٠) تعنى عدم وجود تعادل بين تقدير القدرتين

(١) تعنى وجود تعادل بين تقدير القدرتين