

تأثير نماذج الاستجابة للمفردة على دقة تقديرات القدرة للأفراد باستخدام اختبار
المصفوفات المتتابعة لرافن

م.م علي عباس شنان الزاملي .
مديرة تربوية واسط

مشكلة البحث :

تتجه الدراسة الحالية نحو الكشف عن دقة تقديرات قدرات الافراد التي تعتمد بطبيعة الحال في تقديرها على متغيرات (الصعوبة والتميز) حيث ان نتائج دراستي كل من (Fan,2009). (Robert,2010) ان نموذج راش (الاحادي البارمتر) الاكثر دقة في تقدير كل من القدرة والصعوبة مقارنة بالنموذجين الاخرين الثنائي والثلاثي البارمتر في حيث اثبت (Van Der Maas,2011) P99 تفوق النموذج الثلاثي البارمتر على النماذج الاخرى في اتخاذ القرارات فيما يتعلق بالفروق الفردية ،كما اثبت دراستي (Courvill,2004 :87)،(Jairo,2003:54) ان النموذج الثنائي كان افضل النماذج الثلاثة واقلم من حيث معدلات الخطأ،بينما اثبت دراسة (العكالية، 2007: 66) و (والويلي، 2002 : 98) ان النموذج الثلاثي البارمتر هو النموذج الاكثر دقة في انتقاء البيانات مقارنة بالنموذجين الاخرين، مما يشير الى تعرض النتائج فيما يتعلق بدقة تقديرات معالم المفردات الاختبارية وكذلك معالم الافراد (متمثلة في تقديرات القدرة) باختلاف النماذج المستخدمة،وعلى الرغم من ان وجهة الدراسة الحالية لا تتجه نحو حسم الخلاف فيما يتعلق بدقة تقدير معالم المفردات الاختبارية (الصعوبة والتميز) وإنما نتائج الدراسات المشار اليها انفا اشارت لوجود اختلافات في تقدير للقدرة وفقا للنموذج المستخدم،اي نموذج من النماذج المرتبطة بنظرية (IRT)،مما اثار عدد من التساؤلات لدى الباحث تتعلق بمدى اختلاف هذه النماذج في تقديراتها لقدرات الافراد من خلال نموذج الاستجابة المستخدم في ضوء استخدام اسلوب مختلف عن الاساليب التي استخدمت بالدراسات السابقة (وهو اسلوب الخطأ المعياري لتقدير القدرة) في تقدير دقة قدرات الافراد،كما تعلق تلك التساؤلات بالنموذج الذي يمكن ان يحقق اعلى دقة تقديرات القدرة الافراد،كما

تعلقت تلك التساؤلات بالنموذج الذي يمكن ان يحقق اعلى دقة تقديرات القدرة للأفراد، وهذا ما تحاول ان تجيب عنه الدراسة .

ويتضح من العرض السابق لنموذج الاستجابة للمفردة -احادي او ثنائي أو ثلاثي البارمتر- الاختلاف الواضح للصيغ الرياضية الخاصة بتقدير البارمترات الخاصة بالمفردات الاختبارية وقدرات الافراد حيث يعتمد كل نموذج على متغيرات تختلف عن المتغيرات التي يعتمد عليها النموذجين الآخرين، حيث يعتمد النموذج الاحادي البارمتر على متغير صعوبة المفردة فقط، في حين يعتمد النموذج الثنائي البارمتر على متغيري الصعوبة والتميز، ويعتمد النموذج الثلاثي البارمتر على متغيرات الصعوبة والتميز والتخمين، ونظراً لان تقدير مستوى قدرة المستجيب للمفردة قد تتأثر كما تبين سابقا بمتغيرات التميز والتخمين فان ذلك قد يؤدي الى اختلاف في تقدير قدرات الافراد، التي يمكن الحصول عليها من التحليلات الخاصة بتلك النماذج مما يستوجب الكشف عن تأثير تلك الاختلافات على دقة تقديرات القدرة الناتجة من التحليلات الخاصة بالنماذج الثلاث وهذا ما تستهدفه الدراسة الحالية من خلال بعض المؤشرات التي يمكن من خلالها الكشف عن دقة تلك التقديرات، وبالتالي يمكن ان تتلخص مشكلة الدراسة في الاجابة عن الاسئلة التالية :

- 1- هل تختلف تقديرات القدرة لعينة متلقي الاختبار باختلاف نموذج التحليل المستخدم (النموذج احادي والثنائي والثلاثي البارمتر)؟
- 2- هل تختلف الاخطاء المعيارية لتقدير القدرة للعينة المستخدمة باختلاف نموذج التحليل المستخدم ؟
- 3- هل يختلف شكل توزيعات القدرة للأفراد باختلاف نموذج التحليل المستخدم ؟

أهمية البحث :

تبدو أهمية الدراسة (النظرية والتطبيقية) من خلال النقاط التالية :

1- الدراسة تهتم بتحليل اختبار المصفوفات المتتابعة (العادي) باستخدام بعض النماذج الرياضية اللوغارتمية المرتبطة بنظرية الاستجابة للمفردة وهذا له اهميته النظرية والعلمية في مجال القياس العقلي بصفة خاصة وفي مجال القياس السلوكي بصفة عامة .

2- تحاول الدراسة التعرف على اي من نماذج التحليل الثلاثة المستهدفة من الدراسة التي يمكن أن تعطي أفضل وأدق تقدير للقدرة المستهدف قياسها من الاختبار المستخدم بالدراسة وهذا له قيمة علمية في مجال التطبيقات التي تستخدم بها تلك النماذج وبصفة خاصة في مجالات القياس المختلفة والمتنوعة .

أهداف البحث :

يستهدف البحث الحالي معرفة الأتي:

1. الكشف عن مدى تأثير تقديرات القدرة بنموذج التحليل المستخدم .
2. الكشف عن دقة تقديرات القدرة المشتق من تحليل البيانات الخاصة باختبار المصفوفات المتتابعة بنموذج التحليل المستخدم .
3. الكشف عن تأثير توزيع بيانات اختبار المصفوفات المتتابعة بنموذج التحليل المستخدم .

حدود البحث :

- اقتصر البحث على بيانات اختبار المصفوفات المتتابع (العادي) لجون رافن .
- اقتصر البحث على تطبيق الاداة الخاصة به على الطالبات دون الطلاب .
- اقتصر البحث في تطبيق الاداة الخاصة به على طالبات المرحلة الثانوية في محافظة واسط من المدارس الحكومية فقط .
- اقتصر البحث على التحليلات بثلاث نماذج مرتبطة بنظرية الاستجابة للمفردة (النموذج الاحادي البارمتر -النموذج الثنائي البارمتر - النموذج الثلاثي البارمتر) .
- اقتصر البحث على تطبيق الاداة الخاصة به في العام الدراسي 2014 -2015 .

تحديد المصطلحات :

أولاً: نظرية الاستجابة للمفردة

عريف مكينتيلير وميلر (McIntire & Miller, 2000)

أنها نظرية تربط بين الأداء على كل فقرة في الاختبار إلى قيمة محسوبة إحصائياً من قدرة المفحوص للخاصية أو بناء المقياس (McIntire & Miller, 2000:370).

ثانياً: اختبار المصفوفات المتتبع:

هو الاختبارات العقلية غير اللفظية، ويفترض أنه من الاختبارات المتحررة من أثر الثقافة أو الاختبارات غير المتحيزة ثقافياً، مما يتيح إمكانية المقارنات الحضارية في الذكاء. ويقاس هذا الاختبار مدى امتداد من القدرة العقلية، بحيث يصلح للاستخدام من مختلف المستويات العمرية من سن السادسة وحتى سن 65 سنة (Anastasi, 1976:302).

الفصل الثاني الإطار النظري والدراسات السابقة

نظرية الاستجابة للمفردة :

تعتبر نظرية الاستجابة للمفردة أحد الاتجاهات الحديثة للقياس ولقد حققت النماذج الرياضية اللوغارتمية المرتبطة بها طفرة هائلة ومتطورة في مجال القياس النفسي والسلوكي، ولقد أكدت العديد من الدراسات الأجنبية والعربية تفوق هذه النظرية في تحقيق الموضوعية المأمولة منها هي هذا المجال إذا ما قورنت بالنظرية الكلاسيكية، كما أن هذه النظرية تعتمد على افتراضات وشروط أقوى من الافتراضات والشروط ذات الصلة بالنظرية الكلاسيكية التي

هيمنت على مجريات وإجراءات القياس السلوكي والتحليلات الخاصة بأدوات القياس لفترة زمنية (Uttaro&Lehman,1994:41)(Rotou,2002:131).

ونتيجة لما حققته هذه النظرية وكذلك النماذج الرياضية المرتبطة بها في تحقيق الموضوعية في مجال القياس السلوك الإنساني، وفي بناء وإعداد اختبارات حققت تطوراً وتقدماً هائلاً في شتى ميادين القياس، فقد اتجهت العديد من الدراسات الأجنبية والعربية إلى بناء وتطوير اختبارات جديدة أو إعادة تطوير بعض الاختبارات العالمية المشهورة وشائعة الاستخدام في ضوء النظرية الحديثة للقياس والتحليلات الإحصائية المتعلقة بأحد النماذج الرياضية المرتبطة بها بهدف تحقيق المزيد من الموضوعية فيما تسفر عنه من قياسات (الزيد، 2007: 77)(Rouse,et al,1999:282).

ولقد اشتملت التحليلات الإحصائية الخاصة بالنماذج الرياضية المرتبطة بهذه النظرية الحديثة بعض المؤشرات الإحصائية التي تستخدم في التحقق من دقة وموضوعية نتائج القياس بالاختبار وبنوده المختلفة ومن بين هذه المؤشرات مؤشر الصعوبة ومؤشر التمييز ومؤشر التخمين والذي يمكن تقديرها لكل مفردة من مفردات الاختبار وكذلك الاختبار ككل (Hambleton&Swaminathan,1985:45).

وتختلف أساليب وصيغ التقديرات الخاصة بهذه المؤشرات باختلاف نموذج التحليل المستخدم حيث يعتمد كل نموذج من هذه النماذج على متغيرات تختلف عن المتغيرات التي تعتمد عليها النماذج الأخرى، ومما قد يؤدي إلى نتائج متباينة لتقديرات قيم تلك المؤشرات، ومما قد يؤدي في النهاية إلى اختلاف دقة تقدير القدرات التي يمكن استخلاصها من الاستجابات على مفردات وفقاً لاختلاف نموذج التحليل المستخدم (أحادي -ثنائي -ثلاثي) البارمتر، ونظراً لازدياد وتنامي استخدام التطبيقات الخاصة بنظرية الاستجابة للمفردة في الآونة الأخيرة على الصعيد العالمي وبصورة خاصة على الصعيد العربي في مختلف مجالات

القياس التي تهتم ببناء وإعداد أدوات جديدة متنوعة لجمع البيانات أو إعادة تطوير أدوات شائعة الاستخدام (منيرد، 200: 67) (Edward, 2010: 54).

ونظراً لاعتماد الإجراءات والتحليلات الخاصة بإعداد هذه الأدوات على المؤشرات الاحصائية (الصعوبة، التمييز، التخمين) والتي قد يختلف تقديرها باختلاف النماذج الرياضية المرتبطة (IRT) حيث يعتمد النموذج الاحادي البارمتر (راش) على مؤشر الصعوبة فقط، بينما يعتمد النموذج الثنائي البارمتر (لورد) على مؤشر الصعوبة والتمييز، بينما يعتمد النموذج الثلاثي البارمتر (بيرنوم) (لورد) على مؤشر الصعوبة والتمييز والتخمين، ونظراً لاختلاف المؤشرات الاحصائية التي يعتمد عليها كل نموذج من النماذج الثلاثة في تقديرات القدرة الخاصة بالإفراد، فقد دعت الحاجة الى ضرورة اجراء هذه الدراسة التي تستهدف عقد مقارنة بين تقديرات قدرات الافراد المشتقة من تحليل البيانات الخاصة بالمفردات الاختبارية (باستخدام اختبار المصفوفات المتتابعة لرافن العادي) والتي يمكن التوصل اليها باستخدام ثلاثة نماذج شائعة الاستخدام في هذا المجال وهي نماذج (احادية -ثنائية -ثلاثية) البارمتر، وذلك بهدف التعرف على اي النماذج الثلاثة التي يمكن أن تعطى اعلى وأفضل تقدير للقدرة، وبالتالي يمكن من خلال النتائج المتوقعة من هذه الدراسة تحديد اي هذه النماذج يمكن الاعتماد عليها مستقبلاً وفقاً لكفاءتها في تقديم افضل وأدق المعلومات عن قدرات الافراد في حال خضوعها لإجراءات وعمليات القياس

نظرية الاستجابة للمفردة Item Response

نظراً لنواحي القصور التي تعاني منها أساليب القياس التي يتم اعدادها وفقاً للنظرية التقليدية للقياس والتي بدت بصورة واضحة من خلال الانتقادات التي وجهت اليها، كانت هناك محاولات مستمرة من قبل المهتمين بمجال القياس النفسي والسلوكي أسفرت عن بعض الاتجاهات الحديثة للقياس، والتي تمثلت في ظهور ما يسمى بنظرية الاستجابة

للمفردة، والنماذج الرياضية المرتبطة بها، والتي حققت إلى حد بعيد الموضوعية المنشودة والتي تتصف بها التي تتصف بها أدوات القياس المستخدمة في المجال الفيزيقي (Linden, Hambleton, 1997:56)

تقديرات بارامترات المفردات

1. بارامترات التمييز: Discrimination Parameters

وهي مؤشرات المفردة التي تفيد بقدرتها على التمييز أو التفريق بين المستويات المختلفة من القدرة المقاسة لدى الممتحنين، وتقدر بارامترات التمييز في التحليلات الخاصة بنماذج التحليل المرتبطة بنظرية الاستجابة للمفردة بظل الزاوية التي يصنعها المماس عند نقطة منتصف المنحنى المميز على محور القدرة (Wang&Vispoel, 1998:116).

2. بارامترات الصعوبة: Difficulty Parameters

وتشير إلى مواقع المفردات على متصل وتدرج الصفة المقاسة من الاختبار وتحدد امتلاكها لمستويات محددة من هذه الصفة، وأكد الباحثون على أهمية أن تتوزع قيم بارامترات صعوبة المفردة بصورة واسعة (Jason, Edward&Walter, 2009:34)

3. بارامتر التخمين: Guessing Parameters

وهي المؤشرات التي تشير إلى توصل الممتحن إلى الإجابة الصحيحة عن المفردات التي لم تمكنه قدرته في الإجابة عنها عن طريق التخمين العشوائي (Ponsoda& et al, 1997:213).

نماذج الاستجابة للمفردة

1- النموذج احادي البارامتر

يعتمد هذا النموذج على فلسفة الاحتمالات والتي تتناسب مع طبيعة الظاهر السلوكية كونها ظاهرة احتمالية وليست تحديدية (Hambelton&Swaminathan,1985:46) ففروض نموذج راش: يستند نموذج راش إلى بعض الفروض أو القضايا الأساسية، وهي نفس الفروض التي تعتمد عليها أغلب النماذج الرياضية لنظرية الاستجابة للمفردة (IRT) ومن بين هذه الفروض:

1. يزداد احتمال الإجابة الصحيحة على مفردة اختبار بزيادة قدرة الفرد.

2. احتمال أن يجيب فرد إجابة صحيحة عن مفردة سهلة، أكبر من احتمال أن يجيب إجابة صحيحة عن مفردة صعبة (علام، 1986: 124).

3- النموذج اللوغاريتمي ذو المعلمين Two-Parameter Logistic Model

ينسب النموذج ثنائي المعلمة إلى لورد (1952)، والذي استخدم المنحنى الطبيعي التراكمي **Cumulative Normal Distribution** وتوصل بيرنوم (1968) إلى تعديل النسخة الرياضية لهذا النموذج مستخدماً المنحنى اللوجستي وذلك بإضافة معلم التميز لكل فقرة، ويتخذ المنحنى المميز للفقرة في هذا النموذج شكل دوال التوزيعات اللوغاريتمية ثنائية المعلمة التي على الصورة الآتية:

$$p_i(\theta) = \frac{e^{Da_i(\theta-b_i)}}{1 + e^{Da_i(\theta-b_i)}} \quad i = 1, 2, \dots, n$$

(Hambleton & Swaminathan 1987: 36).

ويختلف هذا النموذج عن سابقه في تقدير معلم جديد هو قوة تمييز المفردة (a_i) للمفردة (i) وبالتالي فإن النموذج يعتمد على تقدير معلمين أساسيين هما:-

1. صعوبة المفردة

2. قوة تمييز المفردة

حيث يبدو ذلك من الصيغة الرياضية الخاصة بالنموذج وهي:-

$$p_i(\theta) = \frac{e^{Da_i(\theta-b_i)}}{1 + e^{Da_i(\theta-b_i)}}$$

ويمكن كتابة $p_i(\theta)$ ، بصورة أخرى بعد ضرب حدي الكسر في $e^{-Dai(\theta-b_i)}$ وبالتالي يمكن

كتابة المعادلة السابقة على الصورة الآتية :

$$p_i(\theta) = [1 + e^{-Dai(\theta-b_i)}]^{-1}$$

والصورة الأخيرة للنموذج هي :

$$p_i(\theta) = \{1 + \exp[-Dai(\theta - b_i)]\}^{-1}$$

حيث a_i قوة تمييز المفردة (i)

b_i مستوى صعوبة المفردة i

θ مستوى قدرة الفرد

$p_i(\theta)$ هو احتمال الإجابة الصحيحة للفرد الذي قدرته θ على المفردة (i)

الثابت D ، هو مقياس العامل (Factor Scaling) ويظهر هذا العامل تقارباً شديداً في تقدير متغيري النموذج (قوة التمييز مستوى الصعوبة) مع متغيري نموذج التجميع الطبيعي *Normal Ogive* عندما يأخذ القيمة 1,7 أي $(D = 1.7)$ ، حيث تختلف تقديرات المتغيرات المتناظرة لكل من النموذجين بصورة مطلقة بأقل من 0,01 لكل قيم θ (Scherbaum,2003:39).

ويتميز هذا النموذج بخصائص رياضية تجعله أكثر استخداماً في التطبيقات السيكمترية من النموذج التجميعي الأعتدالي، ويفترض في النموذج اللوغاريتمي ثنائي المعلمه كما هو الحال في معظم نماذج السمات الكامنة عدم تأثر الإجابات بعامل التخمين. فالمعامل (bi) يرمز إلى معامل صعوبة الفقرة، وهو نقطة على الميزان القدرة **Ability Scale** يكون عندها مقدار احتمال الاجابة الصحيحة عن الفقرة = 0.50 ، وهذه تناظر الاحداثي الأفقي لنقطة انقلاب المنحنى، أي النقطة التي يكون عندها قدرة الفرد (θ) تساوي صعوبة الفقرة (bi) وهذا يحدث عندما تكون الفقرة مناسبة لمستوى قدرة الفرد. والمعلم (ai) يرمز الى معامل تميز الفقرة، وقيمه تتناسب مع ميل المنحنى عند النقطة $\theta = bi$ وهذان المعاملان معرفان على الفترة $[-\infty, +\infty]$ ، أي أن الميزان الذي تقاس عليه قدرة الأفراد هو نفس الميزان الخاص بتمييز الفقرات. ويتراوح مدى معلمة صعوبة الفقرة (bi) بين $[-2, +2]$ وحدة معيارية، على افتراض ان توزيع القدرة تم تدريجه على الفترة $[-3, +3]$ تقريباً. فالفقرات السهلة تقترب معلمة صعوبتها من -2 وحدة معيارية، أما الفقرات الصعبة جداً فتقترب معلمة صعوبتها من $+2$ وحدة معيارية، والفقرة متوسطة الصعوبة تكون معلمة صعوبتها صفراً. ويندر عملياً الحصول على قيم معلمة التميز (ai) اكبر من $+2$ وحدة معيارية، وتستبعد الفقرات التي معلمة تميزها سالبة من اختبارات القدرات. لذلك فإن المدى المعتاد لمعلمة تمييز الفقرة يقع في الفترة $[صفر, 2]$. ويزداد ميل المنحنى المميز للفقرة بزيادة قيمة (ai). فإذا كانت معلمة التميز (ai) = صفراً فإن المنحنى المميز للفقرة يكون مستقيماً موازياً للمحور الأفقي (المحور الذي يمثل القدرة). ويجب ملاحظة أن الفرق بين قدرة الفرد (θ) وصعوبة فقرة الاختبار (i) التي يجيب عنها (bi) أي $\theta - bi$ ، هو الذي يحدد احتمال توصله الى الإجابة الصحيحة للفقرة. لذلك فإنه كلما زاد الفرق زاد ميل المنحنى (النجار، 2010: 320-321).

-النموذج اللوغاريتمي ذو الثلاثة معالم

Three Parameter Logistic Model :

يمكن اعتبار هذا النموذج امتداد لنموذج "راش" النموذج اللوغاريتمي ذي المعلم الواحد، وكذلك اللوغاريتمي ذي المعلمين، إلا أن هذا النموذج يضيف في صيغته الرياضية معلماً ثالثاً وهو للمفردة الخط التقاربي الأسفل للمنحنى المميز للمفردة (Lower Asymptote (ICC)، ويرمز لهذا المعلم بالرمز C_i ، والصيغة الرياضية لهذا النموذج :-

$$p_i(\theta) = C_i + (1 - C_i) \frac{e^{Da_i(\theta - b_i)}}{1 + e^{Da_i(\theta - b_i)}}$$

حيث $p_i(\theta)$ ، $i = 1, 2, \dots, n$ احتمال الإجابة الصحيحة للفرد الذي له القدرة B للإجابة عن المفردة i

يمكن اختصار المعادلة السابقة على الصورة الآتية :

$$T = C_i + (1 - C_i) / \{1 + \exp\} - Da(\theta - b_i)$$

(Hambleton&Swaminthan,1985:38)

T : احتمال الإجابة الصحيحة إذ كان الاختبار متعدد البعد Multidimensional

(C_i) : معلمة تخمين الفقرة (i)

b_i : متغير صعوبة المفردة .

a_i : متغير تمييز المفردة .

exp : مقابل اللوغاريتمي الطبيعي ويساوي 2،27 .

D = 1.7 : عامل القياس (عامل التدرج) (Scaling Factor). ثابت يأخذ عادة القيمة 1.7

تقريباً.

θ : مستوى قدرة المفحوص.

a : المستوى المشترك من التمييز لكافة الفقرات. (عبابنة، 2006: 54).

وفي هذا النموذج إن احتمالية اختيار الإجابة الصحيحة هي دالة لمستوى السمة الكامنة لدى الفرد على صعوبة المفردة، وعلى قدرة المفردة على التمييز بين الأفراد ذوي المستويات المختلفة من السمة، وتدل على مقدار التخمين على هذه المفردة، وتشير القيم الأكبر لـ C_i إلى وجود نسبة كبيرة من الأفراد ذوي المستويات المنخفضة من θ ، ولكنهم اختاروا الإجابة الصحيحة عن طريق التخمين (Scherbaum, 2003: 39-40).

وإضافة هذا المعلم (الثالث) يكون بمثابة تعويضاً لعدم مطابقة المنحنيات المميزة للمفردات عند النهاية السفلى لمتصل القدرة نتيجة تأثير إجابات الممتحنين بعامل التخمين (علام، 1985: 116).

وقد يتطلب حسن مطابقة نتائج الاختبار للنموذج أن تكون القيم المفروضة للمعلم (C_i) أصغر من القيم التي سوف تنتج لو أن الممتحنين ذوي القدرة المنخفضة قد خمنوا الإجابة بطريقة عشوائية على المفردة، وقد لاحظ "لورد" lord إن هذه الظاهرة يكون توزيعها الاحتمالي يرجع إلى براعة وذكاء كاتب المفردة، في تطوير جاذبية الاختبارات غير الصحيحة، بحيث ينجذب الممتحنون ذوي القدرة المنخفضة لهذه الاختبارات، وإلا يحصلون على درجات عالية بواسطة التخمين العشوائي للإجابات الصحيحة، لهذا السبب فإن تقادي متغير التخمين الممثل بالمعلم C_i يبدوا من الأمور المرغوب فيها، ويعتبر هذا النموذج أكثر النماذج الاختبارية واسعة الانتشار لقياس القدرة متعددة البعد، ولكنه أكثر النماذج تعقيداً لأغلب الباحثين عند الاستخدام (Matthew, 1991: 567).

مؤشرات دقة القياس

ان مؤشرات دقة القياس يمكن القياس ان تنقسم يمكن ان تنقسم الى نوعين من المؤشرات

:

مؤشرات عامة حيث تعطي قيمة مختصرة عن دقة التقدير للاختبار ككل وتتمثل

- تتمثل في تقديرات الثبات ومؤشر التقارب والكفاءة (مؤشرات شرطية حيث يتم حسابها عند كل نقطة من القدرة).

- تتمثل فيالجزر التربيعي لمتوسطمربعات الخطأ.

- التحيز Bias.

- الخطأ المعياري لتقدير القدرة. (Divi,1989:146)

المؤشرات العامة لدقة تقدير القدرة

أ. الثبات Reliability

تؤكد نظرية الاستجابة للمفردة على دقة القياس المشروطة حيث تسمح بتحديد دقة القياس عند كل مستوى من القدرة باستخدام دالة فيشر للمعلومات وتشير دالة المعلومات الى دقة القياس الشرطي على مستوى القدرة الكامنة المقاسة، وبالتالي فإن دقة القياس يمكن تحديدها للمستويات المنخفضة والمرتفعة على حد سواء، باستخدام ثبات قيم دوال معلومات مفردات الاختبار

(Meij&Nerng,1999:191) مؤشر التقارب Convergence

يمثل مؤشر التقارب نسبة الحالات التي وفر لها اجراء اسلوب الاحتمال الاقصى تقدير قدرة جيد،وتكون درجة القدرة مقارنة اذا لم تساوي موجب أو سالب ما لانهاية. ولها خطأ معياري مساو لواحد او اقل، ويتم التعرف على التقارب بالرسم من خلال المحور الافقي الذي يمثل القدرة من (-2.5 الى +2.5)مثلا والمحور الراسي يمثل النسبة المئوية للتقارب من (0 الى

100)، ويكون الاختبار دقيقاً كلما ارتفعت النسبة المئوية للتقارب عند كل نقطة من القدرة (Vispoel, 1993:128-130).

ب. الدقة Fidelity

وتعرف على أنها العلاقة بين القدرة المقدرّة والقدرة الحقيقية، ويمكن أن تمثل بالرسم عن طريق المحور الأفقي الذي يمثل طول الاختبار والمحور الرأسي الذي يمثل قيم الدقة كدالة لطول الاختبار، فالعلاقة المرتفعة تعني دقة عالية ومربع الارتباط بين القدرة المقدرّة والقدرة الحقيقية يعطي مؤشر للثبات تماماً كما في العلاقة بين الدرجة الملاحظة والدرجة الحقيقية التي تمثل الاختبار في النظرية التقليدية (Vispoel, Wang & Bleiler, 1997:48-49).

ج. كفاءة الاختبار Efficiency:

تعني كفاءة الاختبار أنه الاختبار يعطي أفضل معلومات ممكنة عن المفحوص بأقل عدد ممكن من المفردات، فالمفردات الأقل تعني كفاءة أعلى، وكلما زاد طول الاختبار قلت الكفاءة بافتراض تثبيت الخطأ المعياري لتقدير القدرة (Rocklin, 1994:9).

ثانياً: مؤشرات الدقة المشروطة

أ. الجزر التربيعي لمتوسط مربعات الخطأ:

يعتبر الجزر التربيعي لمتوسط مربعات الخطأ هو مقياس الخطأ الكلي للتقدير، ويوضح بالرسم من خلال محورين المحور الأفقي الذي يمثل القدرة والمحور الرأسي من 0 إلى 1، ثم ترسم نقاط عند كل مستوى قدرة التي تمثل القيمة بموازاة المحور الأفقي من خلال المحور الرأسي، وكلما انخفضت القيمة واقتربت من الصفر دل هذا دقة أعلى (Weiss, 1982:483).

ب. الخطأ المعياري:

يشير (Urry,1977) أن الخطأ المعياري هو الانحراف المعياري لتقديرات القدرة حول خط الانحدار عند مستوى ثابت من القدرة الحقيقية (Urry,1977:45).

ج. التحيز Bias

يعرف بأنه الفرق بين متوسط الدرجة المقدرة والدرجة الحقيقية ويمكن ان يعبر عنه كما يلي

التحيز = مجموع (القدرة المقدرة - القدرة الحقيقية) / عدد المفحوصين عند مستوى القدرة الحقيقية المقدم وتكون قيمة التحيز مناسبة عندما تقترب من الصفر بدون زيادة في الخطأ المعياري، وإذا انخفض فيكون بخسا للتقدير وإذا ارتفع فيكون افراطا للتقدير (Weiss,1982:482). ومن خلال العرض السابق للأساليب التي تستخدم في الكشف عن دقة تقديرات القدرة الناتجة عن التحليلات، فسوف يستخدم الباحث الحالي اثنان من هذه الاساليب وهما الخطأ المعياري لتقدير القدرة، وكذلك ثبات دوال معلومات الاختبار وذلك للكشف عن دقة تقديرات القدرة التي يمكن الحصول عليها من تحليلات النماذج الثلاثة المستخدمة بالدراسة الحالية، نظرا لان هذين المؤشرين متوافرين ببرنامج التحليل المستخدم بالبحث الحالي وهو برنامج BILOG-MG3 (Zimowski,et.al,2003 :89).

دراسات سابقة

أ. دراسة منيرد (2000)

استهدفت هذه الدراسة استخدام نموذج راش في اعادة تدريج بنود اختبار رافن للمصفوفات المتتابعة العادية لرافن تبعا لمستوى صعوبتها، مع ما يستدعيه ذلك من حذف لبعض البنود غير الملائمة ان وجدت، وتحديد مدى الحاجة لإضافة بعض البنود جديدة، ثم عمل قدرات المعايير المختلفة التي تفسر مستويات قدرات الافراد للأعمار من 6 سنوات الى 13 سنة (منيرد، 2000: 219).

ب.دراسة الزاملي (2012)

استهدفت هذه الدراسة تدريج اختبار رافن للمصفوفات المتتابعة المتقدم وفقاً لنظرية السمات الكامنة وباستخدام انموذج راش، ولتحقيق هذا الهدف اتبع الباحث الخطوات العلمية في تطوير الاختبار بدأ بترجمة اجزاء مختارة من دليل الاختبار، وطبق الاختبار على عينة مكونة من (500) طالبا وطالبة للصفوف (الرابع والخامس والسادس بفرعية العلمي والأدبي في مدارس محافظة بغداد

واستخدم برنامج (Spss) وبرنامج (راسكال) لإجراء المعالجات الاحصائية اللازمة وفقاً لنموذج راش الاحادي المعلم (الزاملي، 2012: ح).

الفصل الثالث منهجية البحث وإجراءاته

مجتمع البحث :

يتألف مجتمع البحث الحالي من طالبات المرحلة المتوسطة للدراسة الصباحية لمحافظة واسط للعام الدراسي (2014-2015)، حيث بلغ المجموع الكلي للطالبات (15593) للصفوف الاول والثاني والثالث

عينة البحث : اعتمد البحث في اجراءات اختبار الفرضيات الخاصة به على عينة عشوائية طبقية من المرحلة المتوسطة بلغ تعدادها (512) طالبة بمتوسط (16.6) وانحراف معياري قدره (1.4) من جميع الصفوف الدراسية بالمرحلة المتوسطة أجري سحبها من خمسة من المدارس المتوسطة في محافظة واسط

أداة البحث : هو اختبار غير لفظي صممه (رافن) Raven في انجلترا يتكون من صورة واحدة فقط تتضمن (60) بنداً تصلح لمستويات عمرية مختلفة تبدأ من عمر الخامسة وحتى مستوى

الراشدين وبالإضافة إلى هذا يوجد اختلاف آخر في هذا الاختبار يدخل في فئة جديدة من فئات الاختبارات وهي الاختبارات المتحررة من اثر الثقافة أو الاختبارات غير متحيزة ثقافياً، بمعنى أن بنود الاختبار لا تتعلق بخصائص حضارية معينة بما يسمح بإمكان تطبيقها بصورتها نفسها دون تعديل في أي مجتمع، كما أنها تتيح إمكانية المقارنات الحضارية في الذكاء، ويتكون الاختبار من مجموعة من الأشكال ذات التصميم الهندسي بينها جزء ناقص ويطلب من المفحوص أن يتعرف على هذا الجزء الناقص ويحدده من بين عدد آخر من الأشكال للإجابة الصحيحة وتتزايد صعوبة البنود تدريجياً حتى نهاية الاختبار (Thorndike, 1962:238).

وقد وقع اختيار الباحث لهذا الاختبار نظراً لإمكانية استخدامه مع مدى واسع من الفئات العمرية المختلفة (من 6 سنوات إلى 65)، وقد اجري التحقق من الشروط والخصائص السيكومترية للاختبار من خلال العديد من الدراسات والأبحاث العربية والأجنبية التي استخدمت هذا الاختبار (معوض، 2007: 80)

الاجراءات

- 1- اجري تطبيق الاختبار على عينة الدراسة .
- 2- تم تحليل استجابات عينة الطالبات لجميع فقرات الاختبار باستخدام نماذج التحليل (احادي -ثنائي -ثلاثي البارمتر)، وذلك باستخدام برنامج BILOG-MG3 .
- 3- ايجاد تقديرات القدرة للإفراد لكل نموذج على حدة.
- 4- ايجاد الاخطاء المعيارية لتقديرات القدرة على حدة باستخدام التحليل الثلاثة.
- 5- التحقق من مدى اختلاف شكل توزيع قيم قدرات الافراد المشتقة من التحليلات الخاصة بنموذج التحليل الثلاث موضع المقارنة من خلال مقاييس التشتت.

6- وتم التحقق من تكافؤ توزيعات تقديرات القدرة المتناظرة المشتقة من التحليلات الخاصة بالنماذج الثلاث موضع المقارنة باستخدام القاعدة التالية: اذا كان الفرق بين احصاءات مقاييس التشتت المتناظرة اقل من او تساوي النسبة (0.05)) اعتبرت اشكال التوزيع متكافئة اما اذا ازدادت الفرق عن تلك النسبة اعتبرت توزيعات غير متكافئة .

منهاج الدراسة والأساليب الاحصائية

اعتمد البحث الحالي في اجراءاتها على المنهج الوصفي (التحليل -المقارن)مع استخدام الاساليب الاحصائية الاتية:

- اسلوب التحليل العاملي الاستكشافي للتحقق من توافر افتراض احادي البعد في بيانات اختبار المصفوفات المتتابع العادي المستخدم بالبحث .
- استخدم مؤشر الملاءمة الاحصائي مربع كاي المتوافر بتحليلات برنامج -BILOG MG3 للتحقق من توافر افتراض استقلالية المحل لمفردات اختبار المصفوفات العادي .
- اسلوب تحليل تباين القياس المتكرر المتوافر بتحليلات برنامج SPSS للتحقق من الفرض الاول والثاني .
- استخدام اختبار (Lsd) لعقد المقارنات الثنائية وليست المتعددة نظرا لان التحليلات تجرى على مجموعة واحدة فقط من الافراد،ولا تجرى على مجموعات متعددة.
- اختبار (كلمجروف-سميرنوف)المتوافر بتحليلات برنامج SPSS للتحقق من مدى اعتدالية توزيع قيم القياسات الخاصة بجميع المتغيرات موضع المقارنة.للتحقق من الفرض الثالث. وسوف يعتمد الباحث في تحليلاته الاحصائية باستخدام الاساليب سالفه الذكر على الحزمة الاحصائية SPSS الاصدار التاسع عشر .

الفصل الرابع

عرض النتائج ومناقشتها

سيتم عرض نتائج البحث الحالي ومناقشتها :

أولا : ((التحقق من توافر افتراضات نظرية الاستجابة للمفردة في بيانات اختبار الدراسة))

(أ) التحقق من توافر افتراض احادية البعد لبيانات اختبار المصفوفات المتتابعة وقد اعتمدت الباحث على استخدام التحليل العاملي الاستكشافي للتحقق من هذا الافتراض وأسفرت التحليلات الخاصة بهذا الاسلوب عن النتائج المبينة بالجدول التالي

الجدول (1) الجذور الكامنة للعوامل الثلاثة الناتجة عن التحليل العاملي لمفردات اختبار المصفوفات المتتابعة

الجذور الكامنة			العوامل
النسبة المئوية للتكرار	النسبة المئوية للتباين	الكلية	
التصاعدي			
26.146	26.146	10.288	1
33.108	6.962	4.177	2
39.161	6.052	3.631	3

يتضح من جدول التباين الكلية المفسر (السابق) ان قيمي الجذر الكامن للعامل الاول بلغت (10.288) كما بلغت قيمة الجذر الكامن للعامل الثاني (4.177)، وقد لوحظ ان نسبة التباين المفسر بالعامل الاول من نسبة الكلية للتباين المفسر بجملة العوامل التي استخرجت من التحليل بلغت (26) تقريبا كما نسبة قيمة الجذر الكامن للعامل الاول الى قيمة الجذر الكامن للعامل الثاني بلغت (2.5) تقريبا ، أي ان النسبة تزيد عن قيمة (2) احادية البعد

لبانات الاختبار، وهذه النتائج تشير الى توافر افتراض احادية البعد، حيث اذا فسر العامل الاول اكبر تباين، فان ذلك يعتبر مؤشر على احادية البعد

ب: ((التحقق من توافر استقلالية المحل لبيانات اختبار المصفوفات المتتابعة))

ويعني هذا الافتراض عدم اعتماد بعضها على البعض الاخر (أي ان الاجابة على أي مفردة اختبارية لا ينبغي ان تتأثر بالإجابة على المفردات الاخرى) وقد اسفرت نتائج التحليلات الخاصة باختبار هذا الافتراض والمتضمنة ببرنامج BILOG-MG3 عن عدم تجاوز أي مفردة من مفردات الاختبار لحدود مؤشر حسن المطابقة الاحصائية (ابراهيم، 2006: 87)

اما توافر افتراض السرعة والمقصود من هذا الافتراض أن الاسئلة المتروكة في نهاية الاختبار ينبغي ان تكون قد تركت بسبب ان قدرة المفحوص لم تمكنه من الاجابة عنها بشكل صائب وليس بسبب انه لم يكن هناك وقت كاف لمحاولة الممتحن ان يجيب عليها (الزيد، 2007: 67) وقد راعى الباحث في تطبيق اختبار البحث اعطاء وقتا للإجابة عن جميع فقرات الاختبار لحين انتهاء كل المفحوصين من محاولة الاجابة عن جميع الفقرات الاختبار، ولم يتبين للباحث وجود ردود افعال من الطالبات عن ضيق الوقت المخصص للإجابة عن فقرات الاختبار، وبالتالي فلم يكن لعامل السرعة اي تأثير في تحديد نوع الاستجابة على فقرات الاختبار مما يعني تحقق افتراض السرعة .

د: ((التحقق من حسن مطابقة بيانات الاختبار للنموذج))

لقد اسفرت نتائج المطابقة للإفراد والمفردات عن النتائج التالية

1- مطابقة المفردات لتوقعات النماذج اللوغاريتمية (الاحادية والثنائية و الثلاثية البارمتر) اعتمادا على دليل البرنامج وحيث بينت نتائج التحليل عدم وجود اية مفردة تجاوزت حدود المطابقة لأي نموذج من النماذج الثلاثة المستخدمة في التحليل .

2- مطابقة الافراد لتوقعات النماذج اللوغاريتمية، وقد تبين (24) طالبة جاء مؤشر المطابقة لقدراتهن خارج حدود المطابقة المسموح بها وتم استبعادهن من التحليلات النهائية .
ثانيا :عرض ومناقشة التحليلات الخاصة باختبار فروض الدراسة
(أ)مناقشة وعرض النتائج باختبار الفرض الاول من فروض الدراسة

أسفرت أجراء التحقق من الفرض ،والذي ينص على :لا تختلف قيم تقديرات القدرات الافراد عينة الدراسة باختلاف نموذج التحليل المستخدم والذي اعتمد على اسلوب تحليل القياس المتكرر باستخدام اختبار كلمجروف-سميرنوف للتحقق من استيفاء شرط اعتدالية توزيع بيانات المقارنة واسفر هذا الاجراء عن البيانات الموضحة في الجدول رقم (2)

قيم تقديرات	قيم تقديرات	قيم تقديرات	
القدرة	القدرة	القدرة	
النموذج	النموذج	النموذج الاحادي	
الثلاثي	الثنائي		
1.003	1.010	1.021	قيمة (Z) لاختبار
0.23	0.15	0.11	مستوى الدلالة

(ب) نتائج التحليلات الخاصة باستخدام أسلوب تحليل القياس المتكرر وفق النتائج المبينة ادناه

الانحرافات المعيارية	متوسطات قيم الاخطاء المعيارية لتقديرات القدرة	نموذج التحليل
1.0805019	0.007511	احادي البارمتر
1.1079355	0.003691	ثنائي البارمتر
1.253661	.036327	ثلاثي البارمتر

يتضح من الجدول (3) السابق ان هناك فروق ظاهرية بين متوسطات الاخطاء المعيارية لتقديرات القدرة الناتجة عن التحليل الخاص بكل نموذج من النماذج الثلاثة المستخدمة بالدراسة حيث بلغت اكبر قيمة لمتوسط الاخطاء المعيارية القيمة (0.007) تقريبا وكانت ناتجة من تحليلات النموذج احادي البارمتر، بينما بلغت أقل قيمة لمتوسط الاخطاء المعيارية القيمة (-0.36) تقريبا وكانت للنموذج الثلاثي البارمتر

الاختبار	قيمة	قيمة	درجات	مستوى
Pillai s Trace	0.019	4.549	2	0.011
Wilk s Lambda	0.981	4.549	2	0.011
Hotelling Trace	0.02	4.549	2	0.011
Roy s Larest Root	0.02	4.549	2	0.011

يتضح من الجدول (4) السابق الذي يتضمن نتائج الاختبارات المختلفة التي تبرز دلالة التباين بين متوسطات قيم تقديرات القدرة الناتجة عن التحليل بالنماذج الثلاثة.

جدول رقم (4) دلالة اختبار ماكولي للتحقق من شرط الدورية لقيم اخطاء تقديرات القدرة الناتجة عن تحليل استجابات عينة الدراسة

مستوى الدلالة	درجات الحرية	مربع كاي التقريبية	قيمة اختبار ماكولي
0.00	2	166.674	0.696

ويتضح من الجدول السابق أن قيمة اختبار ماكولي للدورية بلغت (0.696) وهي قيمة دالة حيث بلغت قيمة الدلالة المقترنة بها القيمة (0.00) مما يشير الى وجود اختلافا جوهريا بين تجانس القياسات الثلاثة الناتجة عن التحليل بالنماذج الثلاثة المستخدمة مما يشير الى عدم تحقق شرط الدورية وما يستوجب في مثل هذه الحالة استخدام اختبار "جرين هاوس-جيسر" عند اختبار دلالة التباين بين القياسات الثلاثة، وهو ما يتضح من خلال الجدول التالي

مصدر التباين	الاختبار	مجموع المربعات	درجات الحرية	متوسط المربعات	قيمة (ف)	مستوى الدلالة
بين القياسات	Sphericity	0.545	2	0.272	6.758	0.001
	Greenhouse-	0.454	1.534	0.355	6.758	0.001
	Huynh-feldt	0.454	1.538	0.354	6.758	0.003
	Lower-	0.545	1	0.545	6.758	0.01
داخل القياسات	Sphericity	37.164	922	0.04		
	Greenhouse	37.164	707.082	0.053		
	Huynh-feldt	37.164	708.974	0.052		
	Lowr-	37.164	461	0.081		

ويتضح من الجدول (5) السابق أن قيمة اختبار التباين (F) باستخدام اختبار جرين هاوس - جيسر بلغت (6.758) وهي قيمة دالة حيث بلغت قيمة الدلالة الخاصة بها (0.01) مما يشير الى وجود فروق جوهرية بين متوسطات تقديرات القدرة الناتجة عن التحليل باستخدام النماذج الثلاث موضع المقارنة وللتعرف على اتجاه هذه الفروق يجب استخدام جدول المقارنات الثنائية بطريقة LSD وهي موضحة في جدول رقم (6)

الفرق بين متوسط طرفي المقارنة (الاول- الثاني)	طرف المقارنة الثاني	طرف المقارنة الاول
0.004	النموذج الثنائي	النموذج الاحادي
0.44	النموذج الثلاثي	
0.04-	النموذج الاحادي	النموذج الثنائي
0.40	النموذج الثلاثي	

يتضح من الجدول السابق وجود فروق جوهرية بين متوسطي تقديرات القدرة الناتجة عن التحليل باستخدام النموذجين الاحادي والثلاثي البارمتر، في صالح النموذج الاحادي البارمتر، كما يتضح ايضا أن هناك فروق جوهرية بين متوسطي تقديرات القدرة الناتجة عن التحليل باستخدام النموذجين الثنائي والثلاثي البارمتر في صالح نموذج التحليل الثنائي البارمتر، كما يتضح عدم وجود فروق جوهرية بين نموذجي التحليل الاحادي والثنائي البارمتر .

مناقشة وعرض النتائج باختبار الفرض الثاني

أسفرت اجراءات التحقق من الفرض الثاني من فروض الدراسة والذي ينص على :لا تختلف الاخطاء المعيارية لتقديرات قدرات الافراد باختلاف نموذج التحليل المستخدم. والتي اعتمدت على استخدام أسلوب تحليل القياس المتكرر للاخطاء المعيارية لقياس تقديرات قدرات عينة الدراسة الناتجة عن تحليلات نماذج موضع المقارنة وذلك باستخدام اختبار كلمجروف-سميرنوف

للأحقق من اسأفاء شرط اعدالففة ءوزفع بفانات المقارنة اللازم لاسأءءام أسلوب أألفل القفاس المأكرر. مناقشة وعرض النأاف الخاصة بأأأبار الفرض الأال

لقد اسفرت الاجراءات الخاصة بالأأحقق من الفرض الأال من فروض الدراسة اعأمء على اسأءءام أسالفب الاحصاءات الوصففة وهف مقاففس النزعة المركزية والأأشأء، وكذلك المؤشراأ الاحصائفة لآوزفع البفاناء عن البفاناء الخاصة بأألفل اسأءاباء عفنة الدراسة باسأءءام النماأء الأال مؤضع المقارنة بالآءول (7)

الاحصاءات	النموء	النموء	النموء
	الاحاءف	الأأاف	الأأاف
	البارمأر	البارمأر	البارمأر
الانأراف المعفارف	12.60	1.1079355	10254
الأأافن	1.587	1.228	1.572
أقل آقأفر للقدرة	3.7314-	4.0000-	3.7600-
أقصى آقأفر للقدرة	2.5054	3.3682	2.5147
الارباعفاء	الارباعف	379600-	382200-
	الارباعف	0.1909	0.1605
	الارباعف الأال	0.7266	0.7245

جدول (8) الاحصاء الوصفي لتقديرات القدرة الناتجة من تحليل استجابات عينة الدراسة بنموذج التحليل الثلاث

ويتضح من خلال البيانات الموضحة بالجدول السابق، وجود تشابه وتقارب شديد جداً بين قيم الاحصاءات الوصفية (مقاييس النزعة المركزية والتشتت ومؤشرات توزيع البيانات) لكل من بيانات تقديرات القدرة لعينة الدراسة المشتقة نموذج التحليل احادي البارمتر والآخرى المشتقة من نموذج ثلاثي البارمتر .

وتتسق هذه النتيجة ايضاً مع النتائج التي انتهت اليها بعض الدراسات مثل دراسة (Raykov, T., et al., 2010:33) التي انتهت الى ان النموذج الثنائي البارمتر اعطى اخطاء معيارية للتقدير اقل من نظائرها بالنموذج الاحادي البارمتر ، كما ان هذه النتيجة التي تفيد بان النموذج الثنائي البارمتر اسفر تحليلاته عن اقل اخطاء معيارية لقياس تقديرات القدرة التي تشير الى النموذج ثنائي البارمتر اكثر النماذج دقة في تقدير القدرة وفي ضوء النتائج السابقة يمكن رفض الفرض الثالث

التوصيات والمقترحات :

في ضوء نتائج البحث يوصي الباحث بما يأتي :

1. اجراء دراسة للكشف عن دلالة الفروق بين النماذج الثلاثة في دقة تقديرات القدرة للمفردات الاختبارية .
 2. اجراء دراسة للكشف عن مدى اختلاف قيم الدوال معلومات الاختبار باختلاف نموذج التحليل المستخدم (احادي ،ثنائي، ثلاثي) البارمتر .
 3. اجراء دراسة تأثير اختلاف حجم عينة الافراد على دقة تقدير المشتق من التحليلات الخاصة بنموذج التحليل المستخدم (احادي ،ثنائي، ثلاثي) البارمتر
- المقترحات :

1. يجب الاعتماد على النموذج الثنائي البارمتر في التحليلات الخاصة بالمفردات الاختبارية نظرا لدقة تقديرات القدرة المشتقة من هذا النموذج بالمقارنة بالنموذجين الآخرين .
2. يجب الاعتماد على النموذج الثنائي البارمتر في التحليلات الخاصة بإنشاء بنوك الأسئلة نظرا لدقة القدرة المشتقة من هذا النموذج بالمقارنة بالنموذجين الآخرين .

* المصادر العربية

1. العكايلة، عبد الناصر (2007) : دراسة مقارنة بين النظرية الكلاسيكية والنظرية الحديثة في القياس في كشف وتقدير الخطأ المعياري في اختبارات القدرات المعرفية لطلبة المرحلة الأساسية في الأردن ، اطروحة دكتوراه غير منشورة، جامعة القاهرة، مصر
2. الوليلي، أسماعيل حسن (2002) : دراسة مقارنة لبعض نماذج الاستجابة للمفردة في انتقاء مفردات الاختبارات مرجعية المحك ، اطروحة دكتوراه غير منشورة ، كلية التربية، جامعة القاهرة .
3. الزيد، جواهر محمد (2007) : فاعلية نظرية الاستجابة للمفردة في بناء الاختبارات (اعداد بنك أسئلة في مقرر علم النفس التربوي) اطروحة دكتوراه غير منشورة ، جامعة الامام محمد بن سعود الاسلامية ، كلية العلوم الاجتماعية.
4. الزالملي ، علي عباس ،(2012): استخدام نظرية السمات الكامنة وفقا لانموذج راش في تطوير اختبار رافن للمصفوفات المتابعة المتقدم على طلبة المرحلة الاعدادية في مدينة بغداد، رسالة ماجستير، كلية التربية ابن رشد، جامعة بغداد.
5. النجار، نبيل جمعة صالح،(2010): القياس والتقويم منظور تطبيقي مع تطبيقات برمجية SPSS، ط1، عمان، دار الحامد للنشر والتوزيع
6. منيرد، منى ربيع طنطاوي (2000) : دراسة سيكومترية حول اختبار المصفوفات المتتابعة لرافن باستخدام نموذج راش، رسالة ماجستير ،كلية البنات للاداب والعلوم والتربية ، جامعة عين الشمس
7. علام ،صلاح الدين ، (1986) : تطورات معاصرة في القياس النفسي والتربوي، ط1، الكويت ، مطابع القبس التجارية
8. معوض، خليل ميخائيل(2007): القدرات العقلية . ط4، مصر الاسكندرية للكتاب .

* المصادر الاجنبية

9. Anastasi, A (1976).Psycholoical Testing, (4th,ed)New Yourk:The Macmillan,Co.Inc
9. Corville,T.g, (2004): an empirical comparison of item response theory and classical test theory item –person statistics unpublished dissertation ,office of gradute studies of texas a&w university.
10. Edward, H. (2010): Interpretation of the three-parameter Testlet Response Model and Information Function .Applied Psychological Measurement,,34(7):467-482
11. Fan,a (2009): Item Response Theory Model:an Empirical Comparison of Thir Item Statistics Educational and Psychological measurement,44(2):256-270
12. Hambleton&Swaminathan, H. (1985):Item Response Theory Principle and Applicationns,The Netherland,Kluwer Nijhoff Publishing
13. Hambleton& Swaminathan, H. (1987):Item Response Theory Principle and Applicationns,The Netherland,Kluwer Nijhoff Publishing,United States of America, 3rd
14. McIntire&Miller S,A&(2000):Foundations of Psychological Principles and Application,4th Ed,New jesity, printice-Hall Measurement VO,13,NO.4,PP38
15. Linden W& Hambleton,R (1997):Hanbook of Moderen Item Response Theory Spring,New York
16. Wang,T. & Vispoel, W.(1998).Properties of ability estimation methods in computerized adaptive testing ,Journal of Educational Measurement,35,109-135
17. Jason I. Meyers, G Edward Miller,Walter D. Way,(2009)Item Position and Item Difficulty Change In An IRT-Based Common Item Equating Design.Applied Measurement in Education,22(3):38-60,
18. Ponsoda V. wise S., Olea,J & Revuelta , J.(1997).An investigation of self- adaptive testing in a Spanish high school population.Educational and Psychological Measurement,57,210-221
19. Reckase, MD.(1999).The past and future of Multidimensional Item Response Theory .Applied Psychological Measurement;21(3):25-36
20. Urry.V(1977) Tailored Testing A successful application of latent trait theory,Journal of Educational Measurement,14,181-196
21. Reckase,M.D(1999).The past and Future of Multidimensional Item Response Theory Applid psychological Measurement .21(3):25-39
22. Zimowski,M,.Muraki,E,.Mislevy,R.J.& Bock,R.D(2003).BILOG-MG 3:Item Analysis and Test Scoring With Binary logistic Models .Chicago: Scientific Software
23. Sternbeg, R. J. Zhang, (2001): Perspectives Thinking, Learning and Cognitive Styles. Mahwah, nj: Erlbaum Associates.