

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ



كلية التربية
قسم علم النفس التربوي

استخدام نموذج التقدير الجزئي لتحليل مستويات بنية ناتج التعلم
للاختبارات العنقودية ذات الفقرات المترتبة هرميا لدي عينة من
طلاب كلية التربية بالمنيا

رسالة مقدمة من :

محمد إبراهيم محمد محمد

مدرس مساعد بقسم علم النفس التربوي

للحصول علي درجة دكتوراه الفلسفة في التربية

(تخصص قياس نفسي)

إشراف

أ.د. علي حسين بداري

أستاذ علم النفس التربوي

كلية التربية - جامعة المنيا

أ.د. أنور رياض عبد الرحيم

أستاذ علم النفس التربوي وعميد كلية التربية

كلية التربية - جامعة المنيا

أولاً - المقدمة :

شهد مجال القياس النفسي والتربوي في الآونة الأخيرة اهتماماً ملحوظاً على أيدي خبراء القياس النفسي والتربوي في كثير من دول العالم ، وذلك لبناء مفردات اختبارية تساعد على القياس الدقيق لمخرجات التعلم والحصول على تقديرات كمية تتسم بالموضوعية للسّمات التي تنطوي عليها استجابات الطلاب لتلك المفردات الاختبارية ، وقد تمثل هذا الاهتمام في ثلاثة اتجاهات رئيسة هي : **الاتجاه الأول** : تمثل في الاهتمام بالأسس التي يتم في ضوءها تصميم وبناء المفردات الاختبارية لقياس مخرجات التعلم ، وقد ساعد هذا الاتجاه على ظهور العديد من التقسيمات التربوية (**Taxonomies**) التي تفسر التعلم ومخرجاته وتوضح كيفية بناء المفردات الاختبارية لقياس مخرجات التعلم المعرفية، والوجدانية، والمهارية . ومن تلك التقسيمات التي تهتم بدراسة المخرجات المعرفية للتعلم : **تقسيم بلوم (Bloom's Taxonomy)** ، و**تقسيم بنية ناتج التعلم الملحوظ (The Structure of Observed Learning Outcome)** ، و**الاتجاه الثاني** : تمثل في الاهتمام بالشكل العام للاختبار النفسي والتربوي والعلاقات الداخلية بين المفردات ، وقد ساعد هذا الاتجاه على ظهور نظرية **الاختبارات العنقودية (Testlets)** لتوضيح كيفية بناء اختبارات عنقودية خطية أو هرمية، و**الاتجاه الثالث** : تمثل في الاهتمام بكيفية تحليل المفردات الاختبارية لتقدير العلاقة الاحتمالية بين أداء الطالب في الاختبار والسّمات أو القدرات الكامنة وراء هذا الأداء وتفسيرها ، وقد ساعد هذا الاتجاه على ظهور **نظرية الاستجابة للمفردة (Item Response Theory)** وما انبثق عنها من نماذج متعددة أكثرها شيوعاً نماذج **راش الاحتمالية (Rasch Models)** .

وقد برز هذا الاهتمام في مجال القياس والتقويم كرد فعل للنقد الذي تم توجيهه إلي تقسيم بلوم من ناحية ، والنقد الذي تم توجيهه إلي **نظرية القياس التقليدية (Classical Test Theory)** من ناحية أخرى ، ومن تلك الانتقادات التي وجهت إلي الافتراضات التي قام عليها تقسيم بلوم : أنه لا يتم تمثيل المعرفة هرمياً، ويصعب الفصل بين مستوياته الستة ، ويوجد تداخل بين تلك المستويات ، كما أنه لا يستطيع قياس المعالجات العميقة للتعلم ، فقد يعتمد الطالب على المعالجة السطحية في الإجابة عن أسئلة ذات مستويات عليا في تقسيم بلوم . (Hattie&Brown,2004,4)

كما تعد الانتقادات التي وجهت إلي فلسفة القياس النفسي والتربوي عامة ، والنظام المرجعي الجماعية خاصة ، الأسس التي يستند إليها هذا النظام في بناء الاختبارات والمقاييس النفسية والتربوية وتحليل مفرداتها دافعا نحو البحث عن نظرية جديدة في القياس ، إذ يعاب

على الاختبارات والمقاييس المقننة أنها تعتمد على موازنة أداء المتعلم بأداء أقرانه وفقا للمجموع الكلي لدرجاتهم الاختباريه وسواء أجريت هذه الموازنة على أساس الدرجات الخام أو الدرجات المعيارية بأنواعها المختلفة ، فإن هذه الدرجات تخضع لخصائص عينة الأفراد التي تستخدم في اشتقاق معايير الاختبار ، كما تخضع لخصائص عينة الأسئلة التي يشتمل عليها الاختبار واختلاف هذه الخصائص باختلاف عينة الأفراد المختبرين وعينة الأسئلة يؤثر بلا شك في صدق هذه الموازنات ، وبذلك يصعب تعميم نتائج الاختبارات أو الاستفادة منها ، فتدريج صعوبة أسئلة الاختبار يتباين بتباين قدرات الأفراد الذين يختبرون بهذه الأسئلة ، وقياس قدرات أو سمات الأفراد يتباين بتباين صعوبة الأسئلة ، كما أن ثبات درجات الاختبار يتأثر بتباين قدرات الأفراد الذين يجرى عليهم الاختبار ومستوي هذه القدرات .

(صلاح الدين علام، ٢٠٠١، ١٧ - ١٨)

ولتفادي تلك المشكلات فقد اقترح (Wainer & Kiely , 1987, 190) نظرية " الاختبارات العنقودية" لبناء اختبارات تحصيلية تساعد على تحقيق التوازن بين المحتوى الدراسي والأهداف التعليمية بحيث تعتمد على بناء حزم من المفردات تتعلق كل حزمة بمحتوى تعليمي محدد ، ويتم بناء تلك المفردات باستخدام تقسيم مستوى بنية ناتج التعلم ، وتحليلها باستخدام نماذج الاستجابة للمفردة بما يحقق دقة القياس وموضوعيته . ويهدف تقسيم بنية ناتج التعلم (Biggs & Collis , 1982, 31) إلي قياس جودة النواتج التعليمية من خلال قياس مستويات بنية ناتج التعلم ، حيث يرى أن الطلاب يتعلمون المادة الدراسية في خمسة مراحل مرتبة تصاعديا بشكل هرمي من المادي إلي المجرد، ومن البسيط إلي المعقد، تلك المراحل هي : مستوي ما قبل التركيب ، ومستوي التركيب الأحادي ، ومستوي التركيب المتعدد ، والمستوي العلاقاتي ، والمستوي المجرد . ويتميز تقسيم بنية ناتج التعلم بعدة مميزات تميزه عن تقسيم بلوم في تقويم الطلاب ، تلك المميزات لا تقتصر على بناء المفردات الاختبارية أو وضع التقديرات لها فقط ، حيث يتميز بأنه تقسيم هرمي يهدف إلي قياس جودة النواتج التعليمية ، ويوجه الانتباه نحو الكيفية التي يتعلم بها الطلاب ، والكيفية التي يستخدم المدرسون من خلالها إجراءات التدريس لتساعد الطلاب على استخدام عمليات عقلية أكثر تعقيدا في تعلم المهام ، وبذلك فهو لا يشبه تقسيم بلوم الذي يقتصر استخدامه على المدرسين أكثر من الطلاب ، فالطلاب لا يستطيعون تقييم ذاتهم في ضوء تقسيم بلوم، بينما في تقسيم بنية ناتج التعلم يمكن أن يدرس هذا التقسيم بسهولة للطلاب لمساعدتهم على تعلم كيفية التقدم بمستواهم للوصول إلي مستويات عليا ، حيث توجد علاقة بين الطريقة التي يتبعها المدرسون في التدريس والطريقة التي يستخدمها الطلاب في التعلم ، فقد يتقدم كل من المعلم والمتعلم من مستوي المعالجة السطحية إلي

مستوي المعالجة العميقة ، وهذا ينعكس على درجة المتعلم في تقسيم بنية ناتج التعلم بينما ليس من الضروري في تقسيم بلوم التقدم في سلوك التدريس والتعلم.

(Hattie &Purdie, 1998,17)

وتعد نظرية السمات الكامنة من أنسب النظريات التي تستخدم في تحليل فقرات الاختبارات النفسية والتربوية بعامة ، والاختبارات العنقودية بصفة خاصة حيث " تفترض أن هناك سمات أو خصائص معينة يشترك فيها جميع الأفراد، ولكنهم يختلفون في مقدارها وبالرغم من أن هذه السمات غير منظورة ، إلا أنه يمكن الاستدلال على مقدارها من السلوك الملاحظ للفرد المتمثل في استجاباته على مفردات الاختبار وهذا ما يبرر تسميتها بالسمات الكامنة . فالسمة التي تكمن وراء استجابة الفرد على مفردات اختبار لفظي،تختلف عن السمة التي تكمن وراء استجاباته على مفردات اختبار عددي أو مكاني، ولكن يمكن أن تكمن سمة واحدة وراء استجاباته على مفردات اختبارين متعلقين بنفس المحتوى".

(صلاح الدين علام ، ١٩٨٧ ، ٢٢)

ويعد نموذج التقدير الجزئي **Partial Credit** من أنسب نماذج الاستجابة للمفردة وأكثرها استخداما في بناء وتحليل الاختبارات العنقودية ، والفكرة الأساسية التي يقوم عليها هذا النموذج هي أن هناك عدة نجاحات جزئية في كل عنقود تمثل مستويات الطلاب المختلفة ، حيث تفترض أن الطالب يقوم بالإجابة عن العنقود في عدة خطوات تبدأ من البسيط إلى المعقد لذا يتم تقدير درجة الطالب في كل عنقود حسب عدد المفردات التي استطاع إنجازها للوصول إلى الإجابة الصحيحة . (Masters,1982,150,Lam&Foong,1996,4)

ولأن موضوعية وصدق نتائج الاختبارات تعتمد على دقة الأساليب المستخدمة في بنائه واختيار مفرداته وتفسير نتائجه ، ووصفها للقدرة التي يقيسها الاختبار لذلك فإن الأمر يتطلب ضرورة استخدام نظرية الاستجابة للمفردة التي أثبتت البحوث التجريبية أنها تحقق الدقة والموضوعية المنشودة في العلوم النفسية والتربوية .

ثانيا - مشكلة البحث :

تعاني الاختبارات التحصيلية المعتمدة على "نظرية القياس التقليدية" من مشكلة التحيز وعدم الموضوعية حيث تتأثر خصائص الاختبار بخصائص العينة وظروف التطبيق كما تتأثر تقديرات العينة بخصائص المفردات من حيث الصعوبة والتمييز والتخمين ، مما يؤدي إلى الحصول على تقديرات غير دقيقة لكل من الطلاب والمفردات ،بالإضافة إلى عدم دقة تلك الاختبارات نظرا لعدم الثقة في تقدير الثبات والصدق لها، حيث أن الخصائص السيكومترية للاختبار تتأثر بكل من طول الاختبار ونوعية المفردات وخصائصها وحجم العينة

وخصائصها وظروف التطبيق مما يؤدي إلى الشك في تقديرات الخصائص السيكومترية لتلك الاختبارات .

لذلك اتجهت "نظرية الاستجابة للمفردة" إلى وضع عدة افتراضات هي : أحادية البعد ، والاستقلال الموضوعي ، ومنحني المفردة المميز ، تهدف إلى تحقيق الموضوعية بحيث يمكن الحصول على تقديرات للمفردات مستقلة عن العينة ، و تقديرات لقدرات الأفراد مستقلة عن خصائص المفردات ، وتحقيق كفاءة عالية في الاختبار من خلال استقلالية خصائص الاختبار (الصدق والثبات) عن كل من العينة وخصائص المفردات .

(أمينة محمد كاظم، ١٩٨٨، ٤٢)

وتعد الاختبارات العنقودية من أنسب الطرق المستخدمة في بناء اختبارات تحصيلية تحقق التوازن بين المحتوى الدراسي والأهداف التعليمية من خلال بناء عدة عناوين يتكون كل عنقود من حزمة مفردات تتعلق بمحتوي محدد ، لذا تعددت الدراسات التي تناولت مدى ملاءمتها لنظرية الاستجابة للمفردة ومدى تحقيقها لفروضها الأساسية. وقد توصلت تلك الدراسات إلى نتائج متضاربة حول أربع نقاط رئيسية هي :

- ١- مطابقة البيانات المستمدة من الاختبارات العنقودية لنماذج الاستجابة للمفردة.
- ٢- تحقيق الاختبارات العنقودية لافتراض أحادية البعد .
- ٣- تحقيق الاختبارات العنقودية لافتراض الاستقلال الموضوعي .
- ٤- مدى كفاءة الثبات للاختبارات العنقودية .

(Hendrickson,2001,4-5,Wang,Bradlow and Wainer,2002,1-2)

ويرجع ذلك التضارب في نتائج الدراسات السابقة لاختلافها في الإجابة عن السؤالين

التاليين:

- أ- ما الطريقة المناسبة لتصحيح الاختبارات العنقودية ؟
 - ب- ما النموذج المناسب من نماذج الاستجابة للمفردة لتحليل الاختبارات العنقودية؟
- حيث جاءت الإجابة عن السؤالين في اتجاهين مختلفين ، الاتجاه الأول : أشار إلى أن الاختبار يتكون من مجموعة من المفردات وتمثل المفردة وحدة الاختبار وبذلك يتم تصحيح الاختبار من خلال تقدير كل مفردة على أنها وحدة مستقلة بذاتها وتحليله باستخدام نماذج الاستجابة للمفردة ثنائية التدرج التي منها (نموذج راش Rasch Model)، في حين أشار الاتجاه الثاني : إلى أن الاختبار يتكون من مجموعة من العناوين ويمثل العنقود وحدة الاختبار وبذلك يتم تصحيح الاختبار من خلال تقدير كل عنقود على أنه وحدة مستقلة بذاتها وتحليله باستخدام نماذج الاستجابة للمفردة متعددة التدرج التي منها نموذج التقدير الجزئي (Wang&Wilson,2005,126)

ونتيجة لاختلاف الدراسات السابقة حول كل من طريقة تصحيح الاختبارات العنقودية والنموذج المناسب لتحليلها فقد جاءت النتائج متضاربة حول **مطابقة** تلك الاختبارات لنماذج نظرية الاستجابة للمفردة ومدى تحقق الموضوعية والثبات لتلك الاختبارات حيث أشارت دراسات (Rosenbaum,1988,Thissen,Streinberg and Mooney,1989, Yen,1992,1933 , Wainer & Thissen ,1996, Wainer&Lukhele,1997, إلى عدم مطابقة تلك الاختبارات لنماذج نظرية الاستجابة للمفردة وعدم تحقق الموضوعية للاختبارات العنقودية وذلك نتيجة عدم تحقق فروض نظرية الاستجابة للمفردة والتي تتمثل في (أحادية البعد والاستقلال الموضعي) وبالتالي انخفاض معدل الثبات، في حين توصلت دراسات كـ_____ لـ_____:

(Wilson&Masters,1993,Wilson&Adams,1995,Lam&Foong,1996, Davis&Dodd,2003,Lian&Idris, 2006). إلى وجود مطابقة للبيانات المستمدة من الاختبارات العنقودية لتوقعات نموذج التقدير الجزئي وتحقيقها للموضوعية والثبات حيث توصلت إلى تحقق فروض نظرية الاستجابة للمفردة الاختبارية وبالتالي ارتفاع معدل الثبات . ومن ثم يهدف البحث الحالي إلى محاولة التحقق من فاعلية نموذج التقدير الجزئي في تحليل الاختبارات العنقودية مقارنة بنموذج راش نظراً لتضارب نتائج الدراسات السابقة حول كل من طريقة تدريج الاختبارات العنقودية ومدى تحقيقها للموضوعية ، وذلك من خلال التساؤلات التالية :

- ١- هل تختلف قيم المطابقة المستمدة من بيانات اختبار عنقودي مرجعي المحك باختلاف النموذج المستخدم (نموذج راش ، نموذج التقدير الجزئي) ؟
- ٢- هل يوجد اختلاف في تحقيق افتراض أحادية البعد لاختبار عنقودي مرجعي المحك باختلاف النموذج المستخدم (نموذج راش ، ونموذج التقدير الجزئي) ؟
- ٣- هل يوجد اختلاف في تحقيق افتراض الاستقلال الموضعي لاختبار عنقودي مرجعي المحك باختلاف النموذج المستخدم (نموذج راش ، ونموذج التقدير الجزئي) ؟
- ٤- هل يوجد اختلاف في تحقيق افتراض منحنى خصائص المفردة عنقودي مرجعي المحك باختلاف النموذج المستخدم (نموذج راش ، ونموذج التقدير الجزئي) ؟
- ٥- هل تختلف قيم معاملات ثبات الأفراد والمفردات لاختبار عنقودي مرجعي المحك باختلاف النموذج المستخدم (نموذج راش ، نموذج التقدير الجزئي).؟
- ٦- هل يختلف تدريج اختبار عنقودي مرجعي المحك باختلاف النموذج المستخدم (نموذج راش ، نموذج التقدير الجزئي).؟

- ٧- هل يختلف ثبات التدرّيج لاختبار عنقودي مرجعي المحك باختلاف مستويات قدرات الأفراد في الحصول على التدرّيج باستخدام نموذج راش ونموذج التقدير الجزئي؟
- ٨- هل تختلف كمية المعلومات التي يقدمها اختبار عنقودي مرجعي المحك عند مستويات القدرة المختلفة باختلاف النموذج المستخدم (نموذج راش ونموذج التقدير الجزئي)؟

ثالثا- فروض البحث :

يحاول البحث الحالي التحقق من الفروض التالية :

- ١- لا تختلف قيم المطابقة المستمدة من بيانات اختبار عنقودي مرجعي المحك باختلاف النموذج المستخدم (نموذج راش ، نموذج التقدير الجزئي).
- ٢- لا يوجد اختلاف في تحقيق افتراض أحادية البعد لاختبار عنقودي مرجعي المحك باختلاف النموذج المستخدم (نموذج راش ، ونموذج التقدير الجزئي)
- ٣- لا يوجد اختلاف في تحقيق افتراض الاستقلال الموضوعي لاختبار عنقودي مرجعي المحك باختلاف النموذج المستخدم (نموذج راش ، ونموذج التقدير الجزئي)
- ٤- لا يوجد اختلاف في تحقيق افتراض منحنى خصائص المفردة لاختبار عنقودي مرجعي المحك باختلاف النموذج المستخدم (نموذج راش، ونموذج التقدير الجزئي)
- ٥- لا تختلف قيم معاملات ثبات الأفراد والمفردات لاختبار عنقودي مرجعي المحك باختلاف النموذج المستخدم (نموذج راش ، نموذج التقدير الجزئي).
- ٦- لا يختلف تدرّيج اختبار عنقودي مرجعي المحك باختلاف النموذج المستخدم (نموذج راش ، نموذج التقدير الجزئي).
- ٧- لا يختلف ثبات التدرّيج لاختبار عنقودي مرجعي المحك باختلاف مستويات قدرات الأفراد في الحصول على التدرّيج باستخدام نموذج راش ونموذج التقدير الجزئي.
- ٨- لا تختلف كمية المعلومات التي يقدمها اختبار عنقودي مرجعي المحك عند مستويات القدرة المختلفة باختلاف النموذج المستخدم (نموذج راش ونموذج التقدير الجزئي).

رابعا - حدود البحث :

يلتزم البحث الحالي بالحدود التالية :

- أ- أدوات البحث : اختبار عنقودي مرجعي المحك لقياس مستويات بنية ناتج التعلم في مقرر القياس النفسي والتقويم التربوي (إعداد الباحث) .
- ب- عينة البحث : وهي مكونة من (٧٧١) طالبا وطالبة من طلبة الدبلوم العام بكليّة التربية - جامعة المنيا للعام الدراسي (٢٠٠٨ / ٢٠٠٩)

خامسا - نتائج البحث :

تؤكد نتائج البحث الحالي فاعلية استخدام نموذج التقدير الجزئي في بناء وتحليل وتدرج الاختبارات العنقودية مرجعية المحك مقارنة بنموذج راش من خلال تحقق المؤشرات التالية :

١- تحقق حسن مطابقة للبيانات المستمدة من الاختبار مع نموذج التقدير الجزئي أفضل مقارنة بنموذج راش .

٢- تحقق افتراضات نظرية الاستجابة للمفردة (أحادية البعد - الاستقلال الموضوعي - توازي منحنيات خصائص المفردة) بفاعلية مقارنة بنموذج راش .

٣- ارتفاع معدل الثبات لكل من الأفراد والعناقيد باستخدام نموذج التقدير الجزئي عنه عند استخدام نموذج راش.

٤- ثبات التدرج برغم اختلاف مستوى الطلاب بكفاءة ودقة أعلى مقارنة بنموذج راش .

٥- ارتفاع دالة معلومات الاختبار عند استخدام نموذج التقدير الجزئي مقارنة بنموذج راش .

٦- الدقة في قياس دالة معلومات العناقيد ودالة الاختبار .

ومما سبق يتضح فاعلية استخدام نموذج التقدير الجزئي مع الاختبارات العنقودية في النظام التعليمي عامة والقياس النفسي والتقويم التربوي خاصة حيث إن استخدام نموذج التقدير الجزئي قد تلاشي العديد من المشكلات التي ظهرت عند بناء وتدرج الاختبارات باستخدام نموذج راش ثنائي التدرج .

د.محمد ابراهيم محمد محمد ، مدرس علم النفس التربوي ، تخصص القياس النفسي ،

كلية التربية ، جامعة المنيا ، مصر

[edutests @gmail.com](mailto:edutests@gmail.com)

موبيل ٠٢٠١٠٩٣١٦٣١١٦

صفحتي الشخصية على الفيسبوك : <http://facebook.com/mhmm.ibrahim>