جامعة القاهرة
معهد الدراسات التربوية
قسم علم النفس التربوي

أثر تصميمات المعادلة ومتوسط صعوبة الاختبارات وتوزيع القدرة على معادلة درجات الاختبارات متعددة الأبعاد باستخدام نظرية الاستجابة للمفردة

دراسة مقدمة من الطالبة

مايا إبراهيم بركات

للحصول على درجة دكتور فلسفة في التربية

– قسم علم النفس التربوي –

تخصص

القياس والتقويم التربوي

إشراف:

أ.د/ رجاء محمود أبو علام
أستاذ غير متفرغ بقسم علم النفس التربوي
معهد الدراسات والبحوث التربوية
جامعة القاهرة

د. أمين علي محمد سليمان
أستاذ مساعد متفرغ بقسم علم النفس التربوي
معهد الدراسات والبحوث التربوية
جامعة القاهرة

٢٠١٠

الفصل الأول

الإطار العام للمشكلة

المحتـوى	
أولاً	مقدمة
ثانياً	مشكلة الدراسة
ثالثاً	أهمية الدراسة
رابعاً	أهداف الدراسة
خامساً	مصطلحات الدراسة
سادسا	متغيرات الدراسة
سابعاً	حدود الدراسة

الفصل الاول
الإطار العام للمشكلة

أولاً: مقدمة:

تعد الاختبارات من الأدوات الأساسية المستخدمة داخل المؤسسات التعليمية في تقويم العملية التعليمية وتقويم مدى التحسن في تحصيل التلاميذ، بهدف اتخاذ القرارات الحاسمة المتعلقة بالطالب والمؤسسة التعليمية ككل.

لذلك فإن الاهتمام بجودة تلك الاختبارات يجب أن يكون الشغل الشاغل لعلماء التقويم والقياس التربوي، واللذين يجب أن يتيقظوا لما تعاني منه تلك الاختبارات من أوجه ضعف وقصور انعكس سلباً على دقة وموضوعية تلك الاختبارات ، خاصة وأن تفسير درجات تلك الاختبارات يعتمد على الدرجة الكلية للطالب وتلك الدرجة لا تعبر بموضوعية وصدق عن المستوى الحقيقي لقدرة الطالب فهي مرتبطة بمستوى صعوبة مفردات الاختبار، فاختلاف مستوى صعوبة مفردات اختبارين يقيسان محتوى واحد يؤدي إلى اختلاف مستوى القدرة عند نفس المتعلم ، كذلك يتأثر تقدير معالم المفردة (صعوبة المفردة وتمييزها...) بمستوى عينة التقنين (Embertson & Rice, 2000:32).

وتتطلب بعض مواقف الاختبار السرية التامة بالنسبة للإختبارات، مما يستدعي وجود صور عديدة للاختبار تقيس المحتوى ذاته، وتمكننا من مقارنة نتائج تلك الصور ، ووفقا للنظرية التقليدية الكلاسيكية في التقويم والقياس التربوي لا يمكن مقارنة صور تلك الاختبارات بسبب اختلاف مستوى صعوبة مفردات تلك الصور. وهذا ما جعل الباحثين يفكرون في طرق جديدة في تصميم الاختبارات تمكنهم من تجاوز تلك الثغرات بهدف تحقيق موضوعية قياس وتقويم الظاهرة التربوية والنفسية كما هو الحال في قياس الظواهر الفيزيائية، حتى يمكن تفسير درجة الطالب في أحد الإختبارات في مجال القياس التربوي والنفسي ومقارنتها بدرجات زملائه في الإختبار ذاته أو اختبارات أخرى تقيس نفس المحتوى، ولذلك لابد من إيجاد دوال وعلاقات رياضية تمكن من تحويل الدرجات الخام المراد تفسيرها إلى مقياس مشترك بصفر مشترك ووحدة قياس مشتركة.

من هنا انبثقت النظرية الحديثة في التقويم والقياس النفسي والتربوي والتي أطلق عليها نظرية الإستجابة للمفردة وهي عبارة عن دوال أو نماذج رياضية تحدد العلاقة بين استجابات الفرد على اختبار ما، وبين السمة أو القدرة الكامنة وراء تلك الاستجابات، ومن ثم تحديد مقدار تلك السمة والإستفادة منها في التنبؤ بسلوك هذا الفرد في مواقف اختبارية مماثلة، بهدف اتخاذ القرارات الخاصة بهذا الفرد، حيث أن تلك النظرية تفترض وجود سمات او قدرات مشتركة بين الأفراد جميعهم وهم يختلفون في مقدار تلك السمات أو القدرات والتي يستدل على مقدارها عند الفرد من استجابته على الاختبار الذي يقيس تلك السمة .

وظهور نظرية السمات الكامنة أو الاستجابة للمفردة واستخداماتها المتعددة أدى الى حل الكثير من المشكلات التي واجهت الاختبارات التي كانت تبنى وفقاً للنظرية الكلاسيكية في التقويم والقياس، حيث تفترض نظرية الاستجابة للمفردة عدداً من الشروط لابد من تحقيقها في تصميم الإختبارات. وأحد أهم مسلمات تلك النظرية افتراض الإستقلال الموضعي والذي يعني أن تقدير كمية السمة الكامنة وراء استجابات الفرد في أحد الإختبارات، مستقل عن تقدير معالم مفردات هذا الإختبار ، كما أن تقدير معالم مفردات الإختبار مستقل عن مستوى عينة التقدير، حيث أن تحقق هذا المسلم يعني أن الفرد سيحصل على نفس الدرجة عندما يجيب على اختبارين يقيسان مستويين مختلفين من المحتوى ذاته، فلا يصبح لمستوى صعوبة الاختبار أثر في انخفاض مستوى تقدير قدرة الفرد في اختبار ما أو تكون سهولة الإختبار سبباً في ارتفاع تقدير قدرة المفحوص في اختبار ما.

كما أن الإختبار المصمم وفقاً لنظرية الإستجابة للمفردة يجعل تقدير القدرة عند الفرد يتناسب مع مستوى القدرة الحقيقي لهذا الفرد ، حيث أن احتمال الإجابة الصحيحة للفرد على مفردات الإختبار يزداد بزيادة مقدار القدرة لهذا الفرد وليس تبعا لصعوبة أو سهولة مفردات الاختبار. ومن أهم نماذج نظرية الاستجابة للمفردة النموذج أحادي المعلم، والنموذج ثنائي المعلم، والنموذج ثلاثي المعلم وكذلك رباعي المعلم وغيرها من النماذج العديدة .

وعلى الرغم من الحلول المناسبة للكثير من المشكلات والتي قدمتها نظرية الإستجابة للمفردة وأهمها إستقلال تقدير معالم المفردة عن مجموعة المفحوصين، وكذلك إستقلال تقدير قدرة الأفراد عن مجموعة محددة من المفردات، إلا أن الحصول على المزايا التي تقدمها تلك النظرية بقي مرهوناً بتحقيق البيانات لمسلمات نماذج الاستجابة للمفردة وهي:

- أحادية البعد: يعني هذا المسلم أن تقيس كل مفردات الإختبار سمة واحدة فقط، أي تتطلب من الفرد إستخدام قدرة واحدة فقط في الإستجابة على مفردات الاختبار.
- الإستقلال الموضعي: وفيه يكون احتمال إستجابة الفرد على مفردة ما مستقلاً عن احتمال الإستجابة على مفردة أخرى في الإختبار – وذلك عند تثبيت قدرة الفرد – كما أنه لا توحي أية مفردة من مفردات الإختبار بإجابة مفردة أخرى في الإختبار.
- المنحنيات المميزة للمفردات: مجموعة من التوابع الرياضية التي تربط احتمال النجاح على مفردة ما بالقدرة التي تقيسها مفردات الإختبار، وهذا يعني وجود شكل محدد للعلاقة بين السمة أو القدرة وبين الإستجابة الملاحظة.

وعلى الرغم من محاولات معد الإختبار أن يحقق اختباره مسلمات نظرية الإستجابة للمفردة وفي مقدمتها مسلم أحادية البعد، إلا أنه وبعد تطبيق الإختبار يُفاجئ أن اختباره يقيس أكثر من سمة ويغطي أكثر من بعد، وتتعدد أسباب عدم تحقق أحادية البعد في البحوث النفسية والتربوية يعود بعضها إلى الصياغة غير الملائمة لمفردات الإختبار، أو قياس تلك المفردات لمكونات متعددة لقدرة كلية أو قياسها أنواعاً مختلفة من القدرات أو تعود إلى لجوء المفحوص إلى التخمين أو التسرع في الإستجابة، مما يستدعي من المفحوص استخدام أكثر من قدرة في الإستجابة على المفردات. وهذا ما جعل إمكانية الاستفادة من تطبيقات نظرية الاستجابة للمفردة ضئيلاً بسبب عدم إمكانية تحقق الافتراض الأول.

وقد ظهر لنظرية الإستجابة للمفردة العديد من التطبيقات العملية وأهمها استخدام تلك النظرية في معادلة درجات الإختبارات أي تحويل درجات الإختبارات إلى مقياس مشترك موحد يمكننا من مقارنة تلك الاختبارات التي تقيس السمة ذاتها.

وتبرز أهمية معادلة درجات الإختبارات في العديد من الأمور وأهمها : أن هناك العديد من المواقف التي تتطلب أكثر من صورة من الإختبار لضمان سرية الإختبار، حيث أن بعض البرامج تتطلب تطبيق عدة إختبارات على مجموعات مختلفة من الأفراد ومقارنة درجات الأفراد في المجموعات المختلفة ، وكذلك فإن متابعة تقدم طالب ما خلال مرحلة دراسية يتطلب تطبيق اختبارات مختلفة. وأيضا تقويم تقدم الطلاب خلال عام دراسي يتطلب استخدام اختبارات دورية . كل تلك المواقف وغيرها تحتاج إلى معادلة درجات الاختبارات حتى نتمكن من المقارنة بين

الأفراد المطبق عليهم صيغ مختلفة من الإختبار، وبدون عملية المعادلة تلك تصبح المقارنة بين الأفراد غير عادلة لأن بعض الإختبارات تحتوي مفردات سهلة وأخرى تحتوي مفردات صعبة .

ويتم في عملية معادلة درجات الإختبارات ضم مفردات اختبارين مختلفين أو أكثر يقيسان نفس المتغير في تدريج مشترك وفق تصميم معين لجمع البيانات الذي يعمل كرابط لضم هذين الإختبارين في تدريج مشترك، وتوجد ثلاثة تصميمات رئيسية لجمع البيانات التي تعمل كرابط بين الاختبارات المراد معادلتها وهي: تصميم المفردات المشتركة وهو يعتمد على مجموعة مفردات مشتركة متضمنة في كلا الاختبارين المراد معادلتهما، تعمل هذه المفردات المشتركة على ضم الإختبارين المراد معادلتهما في تدريج مشترك. والتصميم الثاني هو تصميم الأفراد المشتركين الذي يعتمد على مجموعة أفراد مشتركين بين مجموعتين مختلفتين من الأفراد يطبق عليهما الإختبارات المراد معادلة درجاتها ويشكلون رابطاً بين هذه الاختبارات. والتصميم الثالث هو تصميم المجموعات المتكافئة ويعتمد على وجود مجموعتين متكافئتين من الأفراد في مجموعتين مختلفتين من الأفراد تعمل هاتان المجموعتان المتكافئتان على ربط الاختبارات المراد معادلتها.

وتضيف عملية معادلة درجات الإختبارات في ضوء نظرية الإستجابة للمفردة شرطا آخر يضاف الى أحادية البعد وهو تماثل الإختبارات المراد معادلتها وفقاً للنظرية في القدرات التي تقيسها واتساقها فيما تقيسه، وهو من الأمور التي يصعب تحقيقها في أحيان كثيرة .

وفي ضوء صعوبة تحقق إفتراض أحادية البعد في الإختبارات التربوية والنفسية فقد قامت دراسات كثيرة باستخدام نماذج نظرية الإستجابة للمفردة أحادية البعد في معادلة درجات الإختبارات متعددة الأبعاد مثل دراسة دورانس وكينجستون Dorans & Kingston (1985) وأيضاً دراسة مودو Modu (1982) وكذلك دراسة منار طومان (٢٠٠٦) كل هذه الدراسات أشارت الى أن نتائج معادلة درجات الإختبارات متعددة الأبعاد باستخدام النماذج أحادية البعد جيدة كنظيراتها أحادية البعد، حيث أن بعض هذه الدراسات مثل دراسة Drasgow & Parsons (1982) وكذلك دراسة زينك Zeng (1989) التي اهتمت بتأثير تعدد الأبعاد على تقدير معالم القدرة والمفردة وعلى عملية المعادلة معاً وتوصلت إلى أن النماذج أحادية البعد تزود بوصف جيد لمعالم المفردة عندما تتطلب استجابة أكثر من (٨٠%) من مفردات الاختبار إستخدام الفرد القدرة الأساسية التي وضع الاختبار لقياسها.

وقد اعتمدت العديد من الدراسات التي استخدمت نظرية الاستجابة للمفردة أحادية البعد مع بيانات متعددة الأبعاد ، النموذج أحادي البعد أحادي المعلم (نموذج راش) ، حيث أجمعت معظمها على أنه أفضل من النموذج ثنائي المعلم وثلاثي المعلم لأنه أبسط منهما واستخدامه أسهل لإحتوائه على معلم واحد فقط هو معلم الصعوبة .

وتختلف دقة معادلة درجات الإختبارات متعددة الأبعاد المدرجة وفقاً لنماذج نظرية الاستجابة للمفردة أحادية البعد، باختلاف التصميم المستخدم لجمع البيانات وهذا ما أكدته العديد من الدراسات ، كما تتوقف كفاءة معادلة درجات الاختبارات متعددة الأبعاد باستخدام النماذج أحادية البعد على بعض المتغيرات التي تلعب دوراً هاماً في نجاح أو فشل هذه العملية، ومن أهم تلك المتغيرات متغير "متوسط صعوبة الاختبارات" حيث أنه كلما اقتربت متوسطات صعوبة الإختبارات المراد معادلتها من بعضها البعض أدى ذلك إلى نتائج معادلة أفضل، حيث بينت العديد من الدراسات أنه كلما اقتربت متوسطات صعوبة الاختبارات المراد معادلتها من بعضها البعض أدى ذلك إلى نتائج معادلة أفضل، فقد أشار ين ويوجان yen & bogan إلى أنه على الرغم مما تقدمه معادلة درجات الإختبارات متعددة الأبعاد من نتائج جيدة، إلا أن اختلاف متوسط صعوبة الاختبارات المراد معادلتها يؤدي إلى نتائج سيئة لعملية المعادلة ، كما توصل بولت (Bolt,1999) إلى نتائج مماثلة حين بين أن نتائج معادلة درجات الاختبارات متعددة الأبعاد كانت جيدة نظراً لتشابه مستويات الصعوبة فيما بينها.
وفي دراسة دودي (Doody,1985) اختبرت تأثير اختلاف متوسط صعوبة الإختبارين اللذين تتم معادلتهما وتوصلت إلى أنه كلما اقترب متوسط الصعوبة للإختبارين من بعضهما البعض حصلنا على نتائج معادلة أفضل . بالمقابل فقد توصل جايلوكا Gialluca(1984,) إلى أن نتائج المعادلة لا تتأثر بسهولة أوصعوبة الإختبارين المعادلين .

ومن المتغيرات الأخرى المؤثرة في معادلة درجات الإختبارات متعددة الأبعاد متغير "توزيع القدرة" حيث يميل خطأ المعادلة للزيادة مع اختلاف توزيعات القدرة فقد توصل سنيكس (Snieckus, 1992) الى أن خطأ المعادلة يميل للزيادة مع اختلاف توزيعات القدرة . وهذا ما توصل إليه أيضاً سبينس (Spence,1996) بأن نتائج عملية المعادلة تتأثر في الحالات التي تكون فيها مجموعات المفحوصين غير متكافئة.

كما قام سكيج وليستس (Skaggs & Lissitz,1986) بدراسة لمقارنة معادلة نظرية الاستجابة للمفردة مع المعادلة التقليدية باستخدام مجموعتين مختلفتين من مستوى قدرة المفحوصين وكان لتوزيع القدرة تأثيرعلى معادلة درجات الإختبارين. وفي دراسة سلايند Slinde استخدم نموذج راش Rasch لمعادلة اختبارات الفهم من مستويات الصعوبة المختلفة جداً لثلاثة منْ مجموعات طلاب ذوي القدرة المختلفة جداً ، وتحت هذه الظروف المتطرّفة، كانت نتائج معادلة نموذج Rasch غير جيدة.

ومما سبق ترى الباحثة أنه وفي ظل انتهاك إفتراض أحادية البعد والذي ينتج عن عوامل كثيرة تفرضها الطبيعة المعقدة للسمات التربوية والنفسية والتي تجعل من الصعب جداً أن يقيس الإختبار سمة واحدة فقط لدى المفحوص ، ونظراً لمزايا معادلة درجات الإختبارات وفقاً لنظرية الإستجابة للمفردة أحادية البعد ، فإن استخدام نظرية الإستجابة للمفردة أحادية البعد مع بيانات متعددة الأبعاد في ظل عوامل متعددة لو تمت مراعاتها يمكن التوصل الى نتائج معادلة جيدة في ظل تعدد الأبعاد ، وأهم هذه العوامل تصميمات معادلة الإختبارات ومتوسط صعوبة الإختبارات وتوزيع قدرة المفحوصين وسوف تتناول الباحثة أثر تلك العوامل في معادلة بيانات متعددة الأبعاد باستخدام نماذج IRT أحادية البعد .

ثانياً: مشكلة الدراسة:

تنبع مشكلة الدراسة الحالية من أن دقة استخدام نماذج نظرية الإستجابة للمفردة أحادية البعد في معادلة درجات الإختبارات متعددة الأبعاد تتوقف على مجموعة كبيرة من العوامل والتي إذا لم يتم أخذها بعين الإعتبار فيمكن أن نصل لنتائج معادلة مضللة باستخدام تلك النماذج أحادية البعد وأهم تلك العوامل تصميمات جمع البيانات ، متوسط صعوبة الاختبارات ، وتوزيع القدرة وتلك العوامل الثلاثة ستتناولها البحثة بالدراسة ، وبذلك تتحدد مشكلة هذه الدراسة في الإجابة عن التساؤلات التالية:

١- هل تختلف دقة معادلة درجات الإختبارات متعددة الأبعاد باختلاف تصميم جمع البيانات (تصميم المفردات المشتركة، تصميم الأفراد المشتركين) باستخدام النموذج أحادي البعد أحادي المعلم؟

٢- ما أثر اختلاف "متوسط صعوبة الإختبارات "على دقة معادلة درجات الإختبارات متعددة الأبعاد باستخدام النموذج أحادي البعد أحادي المعلم ؟

٣- ما أثر اختلاف "توزيع القدرة" على دقة معادلة درجات الإختبارات متعددة الأبعاد باستخدام النموذج أحادي البعد أحادي المعلم؟

ثالثاً: أهمية الدراسة:

تكمن أهمية الدراسة الحالية في النقاط التالية:

- التعرف على ما تقدمه نظرية الإستجابة للمفردة كنظرية حديثة في القياس والتقويم في العلوم الإنسانية (النفسية – التربوية) من حلول للعديد من المشكلات التي عانت منها النظرية الكلاسيكية في التقويم على مدى سنوات طويلة.

- التعرف على إمكانية استخدام النماذج أحادية البعد مع بيانات متعددة الأبعاد، وبأن هذه النماذج تزود بوصف جيد للبيانات متعددة الأبعاد ، وذلك في حال عدم إمكانية تحقق الإفتراض الأول لنظرية الاستجابة للمفردة وهو افتراض أحادية البعد.

- التعرف على عملية معادلة درجات الإختبارات كأحد أهم تطبيقات نظرية الإستجابة للمفردة وأهميتها في وضع درجات الإختبارات على مقياس مشترك للإستفادة من ذلك في مقارنة درجات الأفراد، وكذلك التعرف على بعض العوامل المؤثرة في معادلة درجات الإختبارات متعددة الأبعاد باستخدام نظرية الاستجابة للمفردة أحادية البعد .

- التعرف على أفضل تصميم للمعادلة يمكن استخدامه في جمع البيانات متعددة الأبعاد.

رابعاً: أهداف الدراسة:

- دراسة أثر بعض المتغيرات في نتائج معادلة درجات الاختبارات متعددة الأبعاد عند استخدام نموذج الاستجابة للمفردة أحادي البعد أحادي المعلم ،وهذه المتغيرات هي:
  - تصميمات جمع البيانات .
  - متوسط صعوبة الإختبارات .
  - توزيع القدرة .
- معادلة درجات اختبارين ثنائيي البعد محكيي المرجع في مادة العلوم ووضعهما على مقياس مشترك بهدف الاستفادة من ذلك في العديد من الأمور منها مقارنة درجات الأفراد وغيرها.

خامساً: مصطلحات الدراسة:

– نماذج الإستجابة للمفردة: دوال رياضية تتنبأ بأداء الفرد على مفردة من خلال القدرات الكامنة وراء أدائه ومن خلال خواص المفردة . وسوف تستخدم الدراسة الحالية النموذج التالي:

- النموذج أحادي البعد أحادي المعلم: يستخدم هذا النموذج في صيغته الرياضية معلم أو بارامتر صعوبة المفردة .

– الإختبارات متعددة الأبعاد: هي الاختبارات التي تتطلب من الفرد أن يستخدم أكثر من قدرة واحدة أو سمة واحدة في الاستجابة لمفردات الإختبار.

– تصميمات المعادلة: إجراءات تستخدم لجمع البيانات تعمل على ربط الإختبارات المراد معادلتها ووضعها على تدريج مشترك .

– معادلة درجات الاختبارات: عملية إحصائية يتم فيها وضع درجات الإختبارات المختلفة والتي تقيس نفس السمة على مقياس واحد مشترك بحيث يحصل المفحوص على نفس التقدير للسمة التي يقيسها الإختبار بغض النظر عن اختلاف صورة الإختبار.

– متوسط صعوبة الإختبارين: متوسط صعوبة المفردات في كلا الاختبارين اللذين تتم معادلتهما، ومدى اقترابها من بعضها .

– توزيع القدرة: توزيع قدرة المفحوصين فيما يتصل بالسمة المقاسة ومدى الإنحراف لقدرة المفحوصين عن التوزيع الطبيعي.

سادساً: متغيرات الدراسة:

المتغيرات المستقلة : توجد ثلاثة متغيرات مستقلة في الدراسة الحالية وهي تصميمات المعادلة ، متوسط صعوبة الإختبارين وتوزيع القدرة .

المتغير التابع : معادلة درجات الإختبارات متعددة الأبعاد .

سابعاً: حدود الدراسة :

تتمثل حدود الدراسة الحالية في عدد من النقاط تجملها الباحثة فيما يلي :

١– إستخدام النموذج أحادي البعد أحادي المعلم أحادي البعد (نموذج راش) من ضمن النماذج الشائعة لنظرية الإستجابة للمفردة

٢– استخدام تصميمن فقط من تصميمات جمع البيانات وهما (تصميم المفردات المشتركة ، تصميم الأفراد المشتركين) في معادلة درجات الإختبارات متعددة الأبعاد باستخدام نموذج راش .

٣- إقتصرت الدراسة الحالية على (٩٨٢) طالباً وطالبةً من طلبة الصف الأول الإعدادي في محافظة الجيزة .

٤- اقتصرت الدراسة الحالية على ثلاثة عوامل فقط مؤثرة في معادلة درجات الإختبارات متعددة الأبعاد باستخدام نموذج أحادي البعد .

٥- اقتصرت الدراسة الحالية على وحدة " المادة وتركيبها " من وحدات مادة العلوم المقررة على تلاميذ الصف الأول الأعدادي في جمهورية مصر العربية .

٦- اقتصرت الدراسة الحالية على اختبار ثنائي البعد في مادة العلوم ، حيث أن الإختبار المكون من بعدين هو الأكثر شيوعاً في الدراسات السابقة التي تناولت معادلة الإختبارات متعددة الأبعاد .

الفصل الثاني

الإطار النظر

المحتـوى

أولاً **مقدمة**

ثانياً **النظرية الكلاسيكية في التقويم والقياس**

ثالثاً **النظرية الحديثة في التقويم والقياس :**

١- إفتراضات نظرية الاستجابة للمفردة أحادية البعد.

٢- نماذج نظرية الاستجابة للمفردة أحادية البعد.

رابعاً **معادلة درجات الإختبارات :**

١- تعريف معادلة الاختبارات

٢- شروط معادلة الاختبارات

٣- أنواع معادلة الاختبارات

٤- تصميمات معادلة الاختبارات

٥- طرق معادلة الاختبارات

٦- استخدام نظرية الاستجابة للمفردة أحادية البعد في معادلة الاختبارات متعددة الأبعاد

خامساً **العوامل المؤثرة في معادلة درجات الاختبارات متعددة الأبعاد :**

١-تصميمات جمع البيانات

٢-متوسط صعوبة الاختبارات

٣-توزيع القدرة

٤- سيطرة الأبعاد

٥- عدد المفردات متعددة الأبعاد

٦- قوة الارتباط بين الأبعاد

٧- عدد الأبعاد

الفصل الثاني

الإطار النظري

أولاً: مقدمة:

يشهد العصر الحالي العديد من التطورات في مختلف ميادين الحياة التربوية والتعليمية ، وكان التقويم والقياس التربوي والنفسي من أهم الميادين التي تطورت تطوراً ملموساً في العقود الثلاثة الماضية ، وكان لهذا التطور تأثيراته على أساليب التعليم ومنهجياته ، كما كان الإهتمام الأساسي لخبراء التقويم التربوي منصباً على تطوير الأساليب الكلاسيكية في التقويم والقياس والتي تمثلت في التقويم مرجعي المعيار Norm-Referenced Evaluation وما يعتري هذا الأسلوب من قصور في تحقيق الأهداف التربوية المتمثلة بشكل أساسي في تحديد المهارات الفعلية التي يتقنها الطالب ، حيث يقوم هذا الأسلوب على تصنيف الطلاب والمقارنة فيما بينهم بصرف النظر عن كفاياتهم التعليمية ومستوى أدائهم الفعلي.

ومع ظهور نقاط الضعف في التقويم مرجعي المعيار ، ظهرت الحاجة إلى نوع آخر من التقويم الذي اهتم بالكفايات التعليمية للطالب ومقارنة أداءه بمحك أو مستوى أداء معين بدلاً من مقارنته بمعيار الجماعة التي ينتمي إليها، وأطلق على هذا النوع التقويم مرجعي المحك Criterion-Referenced Evaluation ، وعلى الرغم مما حققه هذا النوع من التقويم من تحسن ملحوظ في درجات الطلاب مقارنة بالتقويم معياري المرجع ، إلا أنه بدأ يثير الكثير من الجدل حوله نتيجة الخوف من تركيزه على الحد الأدنى للمهارات وإهماله للمعرفة العليا ، وفي ضوء الإنتقادات الكثيرة التي وجهت لكلا الأسلوبين التقليديين في التقويم (التقويم مرجعي المعيار والمحك) قام علماء التقويم التربوي بمحاولات لإبتكار طرق جديدة في القياس والتقويم تماثل القياس الفيزيائي ، من هنا ظهرت نظرية جديدة في التقويم والقياس أُطلق عليها نظرية السمات الكامنة أونظرية الإستجابة للمفردة Item Response Theory .

وقد ظهرت نظرية الإستجابة للمفردة كرد فعل لأوجه القصور في الأساليب التقليدية للتقويم وتعتمد تلك النظرية على أداء الفرد على مفردة الإختبار في تقدير قدرة الفرد التي يقيسها الإختبار، وتقوم تلك النظرية على عدد من الافتراضات التي ينبغي أن تتحقق لكي تحقق الفائدة

منها ، وظهر لهذه النظرية العديد من التطبيقات أهمها معادلة درجات الإختبارات والتي سوف تذكرها الباحثة بالتفصيل في سياق عرض الإطار النظري لتلك الدراسة ، كما أمكن استخدام تلك النظرية في معادلة درجات الاختبارات متعددة الأبعاد ، وهذا ما سيكون المحور الأساسي للدراسة ، بالإضافة إلى دراسة أهم المتغيرات التي يمكن أن تؤثر في عملية معادلة درجات الإختبارات متعددة الأبعاد باستخدام نظرية الإستجابة للمفردة أحادية البعد. وفيما يلي عرض تفصيلي للتراث النظري للدراسة :

ثانياً: النظرية الكلاسيكية في التقويم والقياس Classical Theory :

يعتبر التقويم والقياس أحد فروع علم النفس التربوي ويقصد بالتقويم Evaluation إصدار حكم على قيمة الأشياء أوالأشخاص أوالموضوعات ، وفي التربية قوم المعلم أداء تلاميذه أي أعطاه قيمة ووزناً بهدف معرفة إلى أي حد استطاع التلميذ الإفادة من عملية التعلم ، والمدى الذي أدت هذه الإفادة إلى إحداث تغييرات في سلوكهم واكتسابهم للمهارات التي تساعدهم على مواجهة الحياة الإجتماعية .

أما القياس Measurement فهو أضيق في معناه من التقويم ويعني غالباً جمع معلومات كمية عن موضوع القياس ، ويُعرّف القياس في الإحصاء تعريفاً إجرائيا أنه تقدير الأشياء والمستويات تقديراً كمياً وفق إطار معين من المقاييس المدرجة ، وذلك اعتماداً على الفكرة السائدة انه كل ما يوجد يوجد بمقدار وكل مقدار يمكن قياسه ، وتتوقف دقة نتائج القياس على دقة المقياس المستخدم في القياس ، وتؤثر السمة المراد قياسها في نوع المقياس المستخدم ، فهناك أشياء تقاس بشكل مباشر كما في قياس الطول ، ولكن غالبية السمات تقاس بطريقة غير مباشرة كما في الإستعدادات العقلية والسمات المزاجية والشخصية ، ولذلك اختلفت المقاييس التي تقيسها عن المقاييس السابقة ليس فقط في طبيعتها بل في درجة دقتها ، فالقياس المباشر أدق وأسهل من القياس غير المباشر ، ويعتبر القياس إحدى وسائل التقويم الهامة ، ولا يوجد تقويم دون قياس في أي صورة من صوره فالتقويم أشمل من القياس. (رمزية الغريب ، ١٩٩٦ : ص ٧-١١)

ويظن بعض المربين أن مفهوم القياس التربوي مرادفاً لمفهوم التقويم على اعتبار أن كليهما يتطلب جمع بيانات تتعلق بالأفراد أوالجماعات ، غير أن القياس يقتصر على تحديد درجة امتلاك الفرد لإحدى السمات . (صلاح الدين علام ،٢٠٠٣: ص ١٢)

طُورت النظرية التقليدية في التقويم والقياس من قبل شارل سبيرمان Sperman عام ١٩٠٤ ، وأحياناً يطلق عليها نظرية الدرجة الحقيقية ودرجة الخطأ ، ومن ثم ازدهرت بواسطة عدد من السيكومتريين المعاصرين مثل فيلدت وبرينان Feldt & Brennan عام (1989) وكلاين kline عام (1986) ، ورغم ظهور نماذج حديثة في القياس والتقويم فما زالت النظرية الكلاسيكية سائدة في هذا المجال ، وتبدأ تلك النظرية من منطلق أن درجة الإختبار تنتج من تأثير عاملين هما :

١- العوامل التي تسهم في الاتساق : وهي تتألف من الخصائص المستقرة للفرد والتي يحاول الباحث قياسها.
٢- العوامل التي تسهم في عدم الاتساق وتشتمل على خصائص الفرد والإختبار أوالموقف التي ليس لها دور في السمة التي نقيسها ولكنها تؤثر في درجة الإختبار . والعامل الأول مرغوب لأنه يمثل الكمية الحقيقية من الخاصية موضوع القياس ، بينما يمثل العامل الثاني التشويش الذي لا يمكن تجنبه لعوامل الخطأ . (إسماعيل محمد الفقي ، ٢٠٠٥: ص٣٢)

ويرى كاراس(Caras, 2001: p 7-9) أن النظرية الكلاسيكية للإختبار كما يشير إسمها تمثل طريقة تقليدية في التقويم وكانت قاعدة للإختبار لمدة قرن تقريباً ، وتفترض تلك النظرية أن درجة المفحوص على الإختبار تساوي الدرجة الحقيقية للمفحوص مضافاً اليها مكون الخطأ ، وهذا يمكن أن يعبر عنه رياضياً كالتالي : $X = T + E$.

حيث أن X هي الدرجة الملاحظة ، و T هي الدرجة الحقيقية ، و E هو الخطأ المرتبط بالقياس .

والدرجة الحقيقية T هي درجة غير ملاحظة ومعرّفة كمتوسط نظري لنتائج الاختبار المستقل المتكرر للفرد ذاته على الاختبار ذاته ، ولو أن المفحوص كان لابد ان يأخذ نفس الاختبار عدداً غير محدداً من المرات ، فإن توزيع الدرجات يبدو طبيعياً تقريباً بمتوسط الدرجة الحقيقية للمفحوص ، وإنحراف معياري يساوي الخطأ المعياري للقياس ، وخطأ القياس يوضع بالإعتبار أنه عشوائي ، غير نظامي وغير مرتبط بالدرجة الحقيقية ، والهدف الأساسي لكل مطور إختبار تقليل خطأ القياس والحصول على أفضل تقدير للدرجة الحقيقية للمفحوص . (Rotou, 2001:p 10)