Institute Of Educational Studies
Educational Psychology Department

The Effect of the Sample Size and the Number of the Common Items on Calibration Items of Item Bank by Using the Partial Credit Model

A Thesis for Acquiring the Master Degree in Education
Submitted to

Educational Psychology Department

by

Ziad Ahmed El-Abdullah

Supervised by:

Prof. Dr. Ragaa Mahmoud Abou-Allam
Prof. of Foundations of Education
Department
Institute of Educational Studies
Cairo University

Dr. Amin Ali Sulaiman
Asst., Prof. Foundations of Education
Department
Institute of Educational Studies
Cairo University

1430H – 2009AD

معهد الدراسات التربوية
قسم علم النفس التربوى

أثر حجم العينة وعدد المفردات المشتركة على تدريج مفردات بنك الأسئلة باستخدام نموذج التقدير الجزئى

رسالة مقدمة لنيل درجة الماجستير في التربية
قسم علم النفس التربوى

إعداد الباحث

زياد أحمد العبد اللّه

إشراف

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

د. أمين على محمد سليمان
أستاذ مساعد بقسم علم النفس التربوى
معهد الدراسات التربوية
جامعة القاهرة

١٤٣٠هـ/٢٠٠٩ م

الفصل الأول

المشكلة وأهميتها

١- مقدمة

٢- مفهوم السمة الكامنة

٣- مشكلة الدراسة

٤- تساؤلات الدراسة

٥- أهمية الدراسة

٦- أهداف الدراسة

٧- متغيرات الدراسة

٨- التعريف الإجرائي لمصطلحات الدراسة

الفصل الأول

المشكلة وأهميتها

١. مقدمة :

حظي مجال القياس والتقويم التربوي والنفسي باهتمام بحثي متزايد من قبل علماء النفس والتربية، حيث كان الهدف الأساسي من ذلك هو الكشف عن الفروق الفردية بأنواعها سواء كانت فروقاً بين الأفراد أو بين الجماعات أو فروقاً في ذات الفرد، وقد تنوعت أساليب وأدوات القياس لتحقيق هذا الغرض، حيث كانت جودة التقويم تعتمد على جودة الأداة وجودة عملية القياس. وقد أثمرت هذه الجهود إلى الانتقال من النظرية التقليدية Classical Theory في تصميم الاختبارات، والتي استخدمت لفترة طويلة في التقويم التربوي والنفسي، إلى النظرية الحديثة المتمثلة بنظرية الاستجابة للمفردة Item Response Theory أو نظرية السمات الكامنة Latent Trait Theory .

وعلى الرغم من شيوع استخدام النظرية التقليدية فى تصميم الاختبارات وتحليل وتفسير نتائجها، إلا أن هناك بعض المشكلات التى تقلل من دقة هذا الاستخدام، ويمكن حصر تلك المشكلات فيما يلى:

١. تأثر خصائص مفردات الاختبار بقدرة الأفراد المفحوصين .

٢. تأثر الدرجة الكلية للفرد فى اختبار ما بمفرداته .

٣. عدم توافر وحدة قياس ثابتة .

٤. عدم إمكانية المقارنة بين تحصيل الأفراد فى حالة استخدام اختبارات تحصيلية مختلفة .

٥. تأثر ثبات الاختبار بالموقف الاختباري .

٦. صعوبة إثراء الاختبارات المعدة باستخدام النظرية التقليدية بمفردات جديدة .

٧. تساوي تباين أخطاء القياس لجميع أفراد العينة موضع الاختبار.

وقد استثارت تلك المشكلات العلماء المتخصصين فى القياس النفسي والتربوي للبحث عن الدقة والموضوعية فى القياس السلوكي حتى يقترب هذا القياس من العلوم الطبيعية، والتي تتميز بعدم تأثر نتائج القياس بالأداة المستخدمة طالما أنها أداة مناسبة لتقدير الظاهرة، كما يكون تدرج الأداة بوحدات قياس متســـاوية لا تعتمد ولا تتأثر بالعناصر التى تقدر عندها الظاهرة. ويذكر علام (٢٠٠٥) أن الموضوعية يقصد بها أن درجة الفرد فى الاختبار لا يجب أن تعتمد على عينة الأفراد، ففي الاختبارات النفسية المقننة تشتق الدرجات الخام من عينة التقنين ثم يوازن أداء

الفرد الذى يطبق عليه الاختبار فيما بعد بمعايير مشتقة من هذه العينة، فإذا تغيرت العينة فقدت هذه المعايير دلالتها، أى يصبح الاختبار محكوماً بالعينة، كما أن الفرد يحصل على الدرجة نفسها فى أي اختبارين يقيسان نفس السمة، إلا أن درجة الفرد تختلف عادة باختلاف الاختبار الذى يطبق عليه، أى أن درجته تصبح محكومة بعينة المفردات التى يختبر بها. والتقويم الموضوعي لا يعتمد فى نظامه المرجعي على مقارنة نتائج أداء المتعلم بأداء الجماعة، أو بأداء الفرد ذاته فى وقت لاحق أو فى اختبار أو مقياس آخر، بل يعتمد على تقدير علاقة احتمالية بين الأداء الملاحظ للفرد فى الاختبار أو المقياس، والسمات أو القدرات التى تكمن وراء هذا الأداء وتفسره. وقد أسفرت جهود الباحثين فى مجال التقويم التربوى عما يسمى بالتقويم الموضوعي Objective Evaluation والذى يتمثل فى نظرية السمات الكامنة أو نظرية الاستجابة للمفردة.

٢. مفهوم السمة الكامنة:

تفترض الاختبارات النفسية والتربوية بعامة أن هناك سمات أو خصائص معينة يشترك فيها جميع الأفراد ولكنهم يختلفون فى مقدارها، وبالتالي ينبغي قياس هذه السمات بشكل كمي، وذلك على اعتبار أن الفروق الفردية هي فروق كمية أكثر منها وصفية (أبوعلام، شريف، ١٩٩٥). وبالرغم من أن هذه السمات غير منظورة ولا يمكن قياسها بطريقة مباشرة إلا أنه يمكن الاستدلال على مقدارها من السلوك الملاحظ للفرد والمتمثل فى استجاباته عن مفردات الاختبار، وهذا ما يبرر تسميتها بالسمات الكامنة، فالسمة التى تكمن وراء استجابة الفرد على مفردات اختبار لفظي، تختلف عن السمة التى تكمن وراء استجابته على مفردات اختبار مكاني أو عددي، ولكن يمكن أن تكمن سمة واحدة وراء استجابته على مفردات اختبارين مختلفين متعلقين بنفس المحتوى. لذلك فإن الهدف الأساسي لكل من النماذج التقليدية ونماذج السمات الكامنة هو تحديد العلاقة بين استجابات الأفراد على اختبار معين والسمة الكامنة وراء هذه الاستجابات، والأمر الأكثر أهمية فى القياس النفسي والتربوي بعامة هو تحديد مقدار السمات الكامنة وراء أداء الأفراد على الاختبارات المختلفة والاستفادة منها فى تفسير النتائج والتنبؤ بسلوكهم فى مواقف مماثلة واتخاذ قرارات معينة بشأنهم في ضوء هذا التقدير الكمى للسمات (علام، ١٩٨٧، ص٢١). ويمكن التوصل إلى متطلبات الموضوعية فى القياس عندما تتوافر فروض النموذج المستخدم كما حددتها أمينة كاظم (١٩٨٨، ص٤٢-٤٧) بالنقاط التالية:

١) أحادية القياس :

وتعنى أن مفردات الاختبار تكون متجانسة فيما بينها وتقيس فى أساسها نفس الصفة أو العملية العقلية، وهذا يعنى أن جميع المفردات المتدرجة الصعوبة تشترك معاً فى الإجراءات والعمليات السلوكية المتطلبة لحل أى منها ، وتختلف فيما بينها من حيث صعوبتها .

٢) استقلالية القياس:

وتبدو فى:

أ- تحرر القياس من خصائص العينة المستخدمة Sample free وهذا يعنى ثبات تقدير كل من قدرة الفرد وصعوبة المفردة واستقرارهما بالرغم من اختلاف عينة الأفراد المستخدمة فى تدريج المقياس طالما أنها عينة ملائمة.

ب- تحرر القياس من مجموعة المفردات المستخدمة Item free وهذا يعنى ثبات تقدير كل من قدرة الفرد وصعوبة المفردة واستقرارهما بالرغم من اختلاف مجموعة المفردات المستخدمة فى القياس، طالما أنها مفردات ملائمة، وأن هذه المجموعات المختلفة من المفردات تقع على ميزان تدرج واحد، أي أنها تعرف متغيراً واحداً .

٣) خطية القياس:

وتعنى أن يكون هناك معدل ثابت لتدرج القياس يتمثل فى وحدة قياس واحدة، عندئذ يكون الفرق بين أى قياسين متتاليين على هذا التدريج ثابتاً عند أى مستوى من مستويات المتغير. ويتيح توافر خطية القياس تقدير النمو أو التغير الحادث فى المتغير موضوع الدراسة كما يتيح أيضاً عمل المقارنات المختلفة التى يهتم بها الباحث.

وقد تنوعت نماذج نظرية الاستجابة للمفردة و لعل من أكثر هذه النماذج استخداماً هو نموذج راش أحادي المعلم وهذا ماجعل معظم الدراسات والأبحاث التي اهتمت بهذا النموذج تعتمد في تصميم المفردات الاختبارية على أنواع محددة من الاختبارات اقتصرت على مفردات من نوع الصواب أو الخطأ ومفردات من نوع الاختيار من متعدد وذلك لملاءمة هذه الأنواع من المفردات للنموذج المستخدم في الدراسة. إلا أن الباحث سيعتمد في دراسته هذه على نموذج التقدير الجزئي Partial Credit Model والذي يعد توسيعا لنموذج راش المتعلق بالمفردات ثنائية الدرجة بحيث يتناول المفردات التي تتطلب استجابات في قسمين مرتبين أو أكثر. لذلك فإن هذا النموذج يعد نموذجاً عاما للاستجابة للمفردة General Polytomous IRT وينتمي الى عائلة نماذج راش في القياس وقام بتطويره ماسترز في استراليا عام ١٩٨٢ وقد أعده في البداية لتحليل المفردات الاختبارية التي تتطلب خطوات متعددة مثل حل المسائل الحسابية، حيث يكون من الأفضل تعيين درجات جزئية Partial Credit عندما تستكمل خطوات متعددة

في عملية الحل (علام، ٢٠٠٥). وكذلك يتيح لنا هذا النموذج استخدام مفردات ثنائية ومتعددة الاستجابة والتي تعتبر نوعاً من غير المناسب الاستغناء عنه في تصميم الصور الاختبارية لمادة الرياضيات. وذلك لأن الاقتصار على استخدام مفردات من نوع الخطأ والصواب أو الاختيار من متعدد سيضع قيوداً على مصممي مفردات بنوك الأسئلة ربما تحد من شمولية تقييم قدرة الأفراد وبالتالي ينعكس أثر ذلك على موضوعية القياس ودقته. ومن هنا جاء الاهتمام بهذا النموذج الذي يتيح لنا تناول أنماط مختلفة من المفردات لتصميم بنوك الأسئلة سواء كانت مفردات ثنائية أو متعددة الاستجابة أو مفردات من نوع الاختيار من متعدد.

٣. مشكلة الدراسة:

تنوعت دراسات المهتمين بعلم القياس والتقويم حول الكثير من النقاط التي شكلت نوع من الالتباس فيما يخص نماذج النظرية المعاصرة في القياس. إلا أن جلّ هذه الدراسات انحصرت خياراتها فيما يتعلق بأداة القياس بنوع واحد من المفردات ألا وهو مفردات ثنائية الدرجة Dichotomously Scored Items والتي تصنّف ضمن نماذج الاستجابة للمفردة ثنائية الاستجابة Dichotomously IRT Models وذلك بهدف تحديد الخصائص السيكومترية للمفردة والعوامل المؤثرة عليها. وبالتالي كان لابد من الاعتماد على نوعين من المفردات المكونة للاختبار وهما مفردات من نوع الصواب أو الخطأ ومفردات من نوع الاختيار من متعدد. واقتصار الدراسات على استخدام هذه الأنواع من المفردات كان له أثر واضح على دقة تقدير كل من قدرة الفرد ومعالم صعوبة المفردات وذلك من خلال تجاهل قياس مستويات معرفية أعلى من مستوى التذكر والذي تركّز عليه المفردات ثنائية الاستجابة، هذا فضلاً عن كون هذه الأنواع من المفردات تضع الفرد أمام إحدى خيارين إما الإتقان التام في حال إجابته بشكل صحيح أو عدم الإتقان التام في حال فشله بالإجابة، وهذا سيكون له الأثر الكبير في تقدير القدرة لديه وخاصة فيما يتعلق بالأفراد ذوي القدرات المنخفضة والذين يعتمدون في غالبية إجاباتهم على التخمين وليس على مايملكون من قدرات بشكل فعلي. من هنا كان اهتمام الباحث بمحاولة إضافة نموذج آخر من المفردات وهي المفردات التي تعتمد على التقدير الجزئي Partial Credit في الإجابة والتي تتيح للفرد استخدام جميع المستويات المعرفية والمهارات التي يمتلكها في سبيل التوصل للإجابة الصحيحة وبالتالي التخلص من التقييد الذي كان يضعه أمام الإتقان أو عدم الإتقان. وهذا سيؤدي دون شك إلى زيادة في دقة تقدير السمة المراد قياسها لدى الفرد مع عدم إغفال الموضوعية والتي تشكل الأساس في النظرية المعاصرة. واستخدام هذا النوع من المفردات يتطلب نماذج مناسبة من نماذج الاستجابة للمفردة تختلف عن النماذج ثنائية الاستجابة والتي شاع استخدامها بشكل كبير وخاصة في منطقتنا العربية، ولعل السبب الرئيسي

وراء ذلك هو سهولة استخدامها من قبل غير المتخصصين فضلاً عن انتشار البرامج الحاسوبية اللازمة لتقدير معالمها. وتدعى النماذج الملائمة لتقدير معالم المفردات متعددة التدريج، والتي تعتمد على التقدير الجزئي للدرجة، بالنماذج متعددة الاستجابة Polytomous IRT Models والتي يندرج تحتها العديد من النماذج التي تعالج معالم هذا النوع من المفردات. وقد اختار الباحث نموذج التقدير الجزئي والذي سيتم من خلاله دراسة أثر أهم العوامل على تحديد الخصائص السيكومترية للمفردات وبالتالي على تدريج هذه المفردات على متصل قدرة واحد ومن ثم تقسيمها حسب معلم صعوبتها إلى صيغتي اختبار (سهلة و صعبة على الترتيب) وذلك بهدف معادلتها وصولاً لبناء بنك الأسئلة. وقد وجد الباحث من خلال اطلاعه على الكثير من الدراسات السابقة أن أهم المتغيرات التي تمت دراسة أثرها على دقة تدريج المفردات كان متغير حجم العينة ومتغير عدد المفردات المشتركة والمستخدمة في معادلة الصور الاختبارية. إلا أن غالبية هذه الدراسات تناولت أثر هذين المتغيرين ضمن إطار نماذج الاستجابة للمفردة ثنائية التدريج في حين أن القليل جداً منها اهتم بدراسة أثرهما على دقة التدريج ضمن إطار نماذج الاستجابة للمفردة متعددة التدريج (والتي تشكل محل اهتمام الباحث). ولقد حاول الباحث جاهداً الاطلاع على الدراسات التي تناولت النماذج متعددة التدريج (أياً كانت المتغيرات التي تدرسها) سواء في المنطقة العربية أو سواها إلا أنه عانى كثيراً في ذلك نظراً لندرة استخدامها، حيث في المنطقة العربية (وعلى حد علم الباحث) أجريت دراسة واحدة فقط (الخولي، ٢٠٠٦) اهتمت بهذا النوع من النماذج. وقلة تناول هذا النوع من النماذج بالبحث والدراسة على الرغم من الأهمية الكبيرة لها شجع الباحث للتطرق لها في محاولة لدراسة أثر حجم العينة وعدد المفردات المشتركة على دقة تدريج مفردات بنك الأسئلة من خلال استخدام نموذج التقدير الجزئي.

٤. تساؤلات الدراسة:

تتحدد مشكلة الدراسة في محاولة الإجابة عن التساؤل الرئيسي التالي:

ما أثر حجم العينة وعدد المفردات المشتركة على تدريج مفردات بنك الأسئلة باستخدام نموذج التقدير الجزئي؟ ويتفرع عن هذا التساؤل الأسئلة التالية :

١) هل يختلف تدريج مفردات بنك الأسئلة في المقياس المرجعي عنه في المقاييس الناتجة عن معادلة مفردات اختباري البنك X و Y (سهل وصعب على الترتيب) في مقياس مشترك وذلك باختلاف حجم العينة المستخدمة ؟

٢) هل يختلف تدريج مفردات بنك الأسئلة في المقياس المرجعي عنه في المقاييس الناتجة عن معادلة مفردات اختباري البنك X و Y (سهل وصعب على الترتيب) في مقياس مشترك وذلك باختلاف عدد المفردات المشتركة ؟

٣) ماهو الحجم الأمثل للعينة والذي ينتج عنه أفضل تدريج ممكن **لمفردات بنك الأسئلة ؟**

٤) ماهي النسبة المثلى التي تشكلها المفردات المشتركة بين صورتي الاختبارين X و Y من العدد الكلي لمفردات **بنك الأسئلة** والتي ينتج عنها أفضل تدريج ممكن ؟

٥. أهمية الدراسة:

يمكن إجمال أهمية الدراسة الحالية من خلال تناولها للنقاط التالية :

١- شاع عند الباحثين تناول نظرية الاستجابة للمفردة من خلال نماذج ثنائية التدريج فقط، وهذا شكل قيوداً كثيرة لعملية قياس القدرة(q) لدى الأفراد، لذلك يرى الباحث أن الاعتماد على النماذج متعددة التدريج يعطي الكثير من المرونة وصولاً إلى تحديد دقيق للخصائص السيكومترية لعملية القياس ككل. ومن هنا جاء استخدام نموذج التقدير الجزئي والذي يمثل إحدى نماذج الاستجابة للمفردة متعددة التدريج.

٢- كثرت الدراسات حول بحث تأثير الكثير من المتغيرات المتعلقة بنماذج الاستجابة للمفردة ثنائية التدريج في حين أن ندرة الدراسات التي تناولت هذه المتغيرات وتأثيراتها المختلفة ضمن نماذج الاستجابة للمفردة متعددة التدريج كان واضحاً. ولعل صعوبة تقدير معالم الدوال الاحتمالية لهذه النماذج في ظل قلة برامج الحاسب والتي تقوم بتنفيذ العمليات الرياضية المعقدة من تفاضل وتكامل كان السبب الرئيسي وراء ذلك. إلا أن توافر هذه البرامج في الفترة الأخيرة شجع الباحث على تناول إحدى هذه النماذج وهو نموذج التقدير الجزئي من خلال دراسة متغيرين مهمين هما حجم العينة وعدد المفردات المشتركة وأثرهما على دقة تدريج مفردات بنك الأسئلة. وسيستخدم الباحث برنامج الحاسوب الـ ConQuest في إصداره الثاني لإنجاز العمليات الخاصة بتقدير معالم النموذج.

٣- تهتم الدراسة الحالية بدراسة أثر حجم العينة وعدد المفردات المشتركة على دقة التدريج وستعمل على بناء نواة لبنك أسئلة من خلال تعيير عدد مناسب من المفردات والتي من الممكن زيادتها إلى العدد الذي نريد وصولاً إلى كم كبير من المفردات، وهذا ما سيتيح لنا فرصة استخدام صور اختبارية مختلفة من حيث شكل المفردات ومستوى صعوبتها إلا أنها جميعاً تقيس سمة كامنة واحدة، والهدف من ذلك تحقيق أكبر قدر ممكن من العدل والمساواة بين الأفراد بالإضافة لتحقيق السرية اللازمة للعملية الاختبارية، والتي تشكل عاملاً كان يصعب السيطرة عليه من خلال استخدام النظرية التقليدية في القياس، بالإضافة لذلك توفر لنا مفردات البنك هذه إمكانية الحصول على الصور الاختبارية اللازمة بسهولة ويسر وهذا بدوره يوفر الكثير من الوقت والجهد.

٤- ستعمل الدراسة الحالية على تحديد حجم العينة الأمثل والذي ينتج عنه أفضل تدريج ممكن للمفردات من خلال مقارنة نتائج التدريج الناتجة عن استخدام حجوم متدرجة من العينات.

٥- ستعمل الدراسة الحالية على تحديد النسبة المثلى لعدد المفردات المشتركة المستخدمة في المعادلة من العدد الكلي لمفردات الاختبار وذلك بهدف تحقيق أفضل تدريج ممكن للمفردات من خلال مقارنة نتائج التدريج الناتجة عن استخدام أعداد مختلفة من المفردات المشتركة في كل حالة تدريج.

٦- تعتبر الدراسة الحالية من الدراسات القليلة عموماً وفي المنطقة العربية خصوصاً(على حد علم الباحث) والتي تناولت نموذج متعدد التدريج ممثلاً بنموذج التقدير الجزئي. حيث واجه الباحث صعوبة كبيرة في الحصول على دراسات سابقة عالجت نموذج التقدير الجزئي بجوانبه المختلفة. لذلك يأمل الباحث أن تشكل هذه الدراسة إضافة جديدة لمكتبتنا العربية وأن تحث المهتمين بهذا المجال على تناول هذا الجانب من النظرية المعاصرة في القياس والذي لم يأخذ حقه من البحث والدراسة لما له من أهمية واضحة.

٦. أهداف الدراسة :

هدفت الدراسة إلى تناول النقاط التالية :

١- استخدام إحدى نماذج الاستجابة للمفردة متعددة الاستجابة ممثلة بنموذج التقدير الجزئي وذلك بهدف تقدير كل من معالم صعوبات المفردات ومعالم قدرات الأفراد.

٢- تحديد حجم العينة الأمثل والذي من خلاله نحصل على أفضل النتائج في تدريج المفردات.

٣- تحديد نسبة المفردات المشتركة والمستخدمة في معادلة صيغتي الاختبار X و Y (السهل و الصعب على الترتيب) إلى العدد الكلي من مفردات الاختبار والتي ينتج عنها أفضل تدريج ممكن.

٤- بناء بنك أسئلة يضم مفردات تم تحديد خصائصها السيكومترية بشكل دقيق واستخدام هذا البنك للحصول على نماذج متنوعة من الاختبارات بهدف تحقيق العدالة والمساواة بين الأفراد مع الحفاظ على السرية المطلوبة للعملية الاختبارية. مع إمكانية توسيع هذا البنك عبر إضافة مفردات جديدة يتم تحديد خصائصها السيكومترية تحديداً دقيقاً قبل إضافتها، كما يمكن إجراء صيانة مستمرة لمفردات البنك من خلال عمليات الحذف والتعديل والإضافة وذلك حسبما يطرأ على المجال من تغيّر وتطوّر.

٧. متغيرات الدراسة:

أ) المتغيرات التابعة: حجم العينة وعدد المفردات المشتركة.

ب) المتغير المستقل: دقة التدريج لمفردات بنك الأسئلة.

٨. التعريف الإجرائي لمصطلحات الدراسة

١) نموذج التقدير الجزئي Partial Credit Model

هو أحد نماذج نظرية الاستجابة للمفردة متعددة الاستجابة، حيث يتم من خلاله تقدير معالم الصعوبة الفئوية لكل خطوة من خطوات الإجابة عن المفردة متعددة الاستجابة وكذلك تحديد معلم قدرة الفرد .

٢) الخصائص السيكومترية:

يقصد بالخصائص السيكومترية وفق النظرية الحديثة: معالم الصعوبة والتمييز والتخمين للمفردة، والصدق والثبات. وفي هذه الدراسة اقتصرت الخصائص السيكومترية على معالم الصعوبة، والصدق والثبات (علام، ٢٠٠١)

٣) بنك الأسئلة Item Bank :

يتضمن مجموعة كبيرة نسبياً من المفردات والتي تتميز بخصائص سيكومترية معينة تم تدريجها بوحدة قياس محددة بهدف تشكيل قياس له صفر واحد مشترك قادر على قياس أكبر قدر ممكن من المتغير المراد قياسه.

٤) المعادلة Equating :

هي عملية إحصائية يتم من خلالها تحديد ما تساويه الدرجة التي يحصل عليها الفرد في أحد الاختبارات (والتي تقيس مستوى معين من متغير ما) من درجة ممكن أن يحصل عليها في اختبار يقيس مستوى مختلف من نفس المتغير وفقاً لمعايير ثابتة.

٥) المفردة ثنائية الاستجابة Binary Response Item

هي كل مفردة تكون إجابتها إما صحيحة كلياً أو خاطئة كلياً أي إنها تقيس مستويين من مستويات المتغير المراد قياسه هما مستوى الاتقان التام ويمثل الاجابة الصحيحة ومستوى عدم الاتقان التام ويمثل الاجابة الخاطئة .

٦) المفردة متعددة الاستجابة Multi-Response Item :

هي كل مفردة تكون الإجابة عليها مكونة من خطوتين على الأقل أي إنها تقيس أكثر من مستوى واحد من مستويات المتغير المراد قياسه، بحيث تكون درجة الفرد بالنسبة للإجابة عن كل مفردة تساوي عدد الخطوات التي أجاب عليها بشكل صحيح ومستمر دون انقطاع، أي اعتباراً من الخطوة الصحيحة الأولى إلى أول خطأ يرتكبه .

٧) المفردات المشتركة Common Items :

وهي المفردات التي تشكل الرابط بين الاختبارين المراد معادلة درجاتهما.

٨) **حجم العينة Sample size** :

تمثل عدد الأفراد المتلقين للإختبار الكلي وللإختبارات الفرعية التي تتطلبها الدراسة. وسيتم اختيار العينة بطريقة المعاينة العنقودية العشوائية

٩) **تدريج المفردات Items Calibration** :

هو انتظام مواقع المفردات حول صفر تدريج متغير مستوى الصعوبة بحيث يعبر موقع كل مفردة (بعدها عن نقطة صفر التدريج) عن مستوى صعوبتها فقط (WRight & Stone ,1979, pp. 28-30)

١٠) **التدريج المرجعي Referential Scaling** :

هو التدريج الذي يعبر عن مستويات الصعوبة المختلفة من متغير ما وينتج من خلال تقدير معالم صعوبة مفردات بنك الأسئلة ككل ومن ثم تدريجها بوحدة قياس مشتركة وذلك دون المرور بمرحلة المعادلة، حيث تعتبر مفردات البنك جميعاً صورة اختبارية واحدة.

١١) **اللوجيت Logits**:

هي وحدة قياس كل من قدرة الفرد وصعوبة المفردة، وتعرف بأنها اللوغاريتم الطبيعي لاحتمال نجاح الفرد على البنود التي تعبر نقطة صفر التدريج عن صعوبتها، عندما يساوي هذه الاحتمال ثابتا هو الأساس الطبيعي (e)، ويكون عندئذ احتمال نجاحه ٠,٧٣(كاظم، ١٩٨٨، ص٥٧)

الفصل الثاني

الإطار النظري

المقدمة

أولاً: نظرية الاستجابة للمفردة (IRT).

ثانياً: بنك الأسئلة

ثالثاً: معادلة الاختبارات

١. مقدمة

٢. أهمية طرق المعادلة القائمة على نماذج نظرية الاستجابة للمفردة

٣. تعريف عملية معادلة الاختبارات

٤. تصميمات معادلة الاختبارات

٥. شروط إجراء معادلة الاختبارات

٦. طرق معادلة الاختبارات

٧. طريقة المتوسط والانحراف المعياري: Mean and Sigma Method

رابعاً: أثر بعض العوامل على دقة نتائج معادلة الاختبارات

١) تأثير حجم العينة على دقة المعادلة

٢) تأثير عدد المفردات المشتركة على دقة نتائج المعادلة

٣) معادلة درجات الاختبارات غير المتوازية للحصول على درجات متكافئة

٤) أخطاء معادلة الاختبار

خامساً: ارشادات عامة لعملية المعادلة