تدريج مفردات اختبار تحصيلي في الدراسات الاجتماعية لمتعلمي التعليم المجتمعي باستخدام نموذج راش

إعداد

رضا شعبان عبدالسلام

باحث دكتوراة بقسم العلوم التربوية والنفسية برنامج علم نفس– كلية التربية النوعية– جامعة الزقازيق

المجلة العلمية المحكمة لدراسات وبحوث التربية النوعية

المجلد الثامن– العدد الأول– مسلسل العدد (15)– يناير 2022– الجزء الأول

رقم الإيداع بدار الكتب 24274 لسنة 2016

ISSN-Print: 2356-8690 ISSN-Online: 2356-8690

موقع المجلة عبر بنك المعرفة المصري https://jsezu.journals.ekb.eg

البريد الإلكتروني للمجلة E-mail JSROSE@foe.zu.edu.eg

تدريج مفردات اختبار تحصيلي في الدراسات الاجتماعية لمتعلمي التعليم المجتمعي باستخدام نموذج راش

إعداد

رضا شعبان عبدالسلام

باحث دكتوراة بقسم العلوم التربوية والنفسية برنامج علم نفس – كلية التربية النوعية – جامعة الزقازيق

مقدمة:

تعتبر دقة القياس هى الخطوة الأولى نحو التقدم العلمى وضبط ظواهره. وتعتمد دقة القياس على دقة أدواته. وقد كان التقدم العلمى فى مجال العلوم الفيزيائية انعكاساً لتقدم أدوات قياس ظواهره ودقتها منذ القرون الماضية، ولا تزال تتقدم بصورة مذهلة.

وقد تعددت المحاولات لبناء أدوات تتصف بالموثوقية فى مجال قياس التحصيل الدراسي للتوصل إلى مقاييس تتميز بالدقة والموضوعية لتقترب من تلك التى تستخدم فى قياس الظواهر الفيزيائية.

وإذا كانت الاختبارات من أهم أدوات القياس فى مجال التحصيل الدراسي ، فإن اختبارات التحصيل للمتعلمين أكثرها انتشاراً نظراً لسهولة إجراءاتها وتطبيقها، فهي تعتمد على موضوعات حياتية مألوفة لديهم.

ويري علماء القياس أن التحصيل الدراسي بالنسبة للمتعلمين يتم التعبير عنه باللغة وأشكال من أشكال التعبير المعرفى من خلال أعمال أدائية تعتمد فى جوهرها على مهارة التنظيم الإدراكى والمعرفي لدى الفرد وحسن تلقيه ومهارته فى تعبيره واستخدامه لعناصر المادة المتعلمة وإعادة تشكيلها فى تكوينات مركبة لذلك أصبح من الممكن النظر إلى الاختبارات التحصيلية من خلال دلالتها المهمة فى قياس التحصيل الدراسي.

وأدى تصميم البرامج الاختبارية على المستوى القومي والإقليمي في الولايات المتحدة الأمريكية إلى دفعة قوية لحركة تطوير الاختبارات التحصيلية، وخاصة الاختبارات المقننة على النحو الذي نشهده الآن وأصبحت الاختبارات التحصيلية التحريرية أكثر أدوات القياس والتقويم استخدماً بل تكاد تكون الوسيلة الأساسية التي تستخدم في توجيه المتعلمين نحو نوع التعليم وفق معدل درجاتهم معتمدة

علي نظرية الاستجابة للمفردة تفاديا للانتقادات التي وجهت إلي نظرية القياس التقليدية. (صلاح مراد وعلي سليمان: ٢٠٠٢: ٤٣) ولقد حاول بعض الباحثين العرب فى عدد لا بأس به من الدول العربية وفى هذا كان لمصر النصيب الأعظم تقنين بعض الاختبارات التحصيلية في بعض المواد الدراسية بالمراحل التعليمية المختلفة.

ولقد اعتمد بعض الباحثين فى تقنينهم للاختبارات التحصيلية على النظرية الكلاسيكية والتى يعيب عليها لويد Loyd (١٩٨٨: ١٣٦) فى قصور الخصائص السيكومترية لهذه المقاييس والتى تعتمد علي خصائص مفردات الاختبار فى تعريفها للمتغير موضوع القياس التى يشتمل عليها الاختبار والتى قد ترجع إلى تعدد المتغيرات التى تقيسها المفردات، ومن هنا تبدو الحاجة إلى التأكد من أن مفردات الاختبار تعرف وتقيس متغيراً واحداً فقط حتى يمكن الاعتماد على صدق الاختبار فى قياس هذا المتغير ويمكن تفسير معنى الدرجة الناتجة.

(هاشواى Hashway، ١٩٩٨: -٨٣)

ويحدد صلاح مراد وأمين سليمان (٢٠٠٠: ٤٢٢-٤٢٣) مشكلات القياس التقليدى (والذى أُعتمد عليه فى بناء الاختبارات التحصيلية ومنها:-

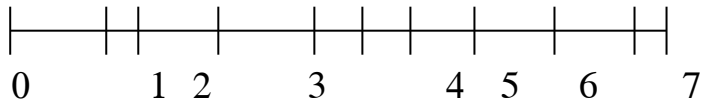
(١) تقيد الدرجة الكلية للاختبار بمفردات الاختبار:

تتلخص هذه المشكلة فى أن الدرجة التى يحصل عليها الفرد فى اختبار معين تعتمد على مفردات هذا الاختبار. فإذا حصل فرد ما على الدرجة ٣٥ فى اختبار ، فإننا لا نستطيع أن نجزم أنه سوف يحصل على نفس الدرجة إذا ما استبدل هذا الاختبار باختبار بديل يقيس نفس الموضوعات أى أن درجة الفرد على الاختبار تتغير بتغير الاختبار. وهذه المشكلة ليس لها وجود فى القياس الفيزيائى حيث إن الدلالة الكمية لطول أى قضيب من الحديد لا تتغير بتغير أداة القياس المستخدمة لقياس الطول.

(٢) انعدام خطية القياس:

ويقصد بخطية القياس وجود معدل ثابت لتدرج القياس على متصل المتغير موضوع القياس، أى أن الدرجات التى يحصل عليها الفرد فى مفردات اختبار يمكن جمعها كما لو كانت تمثل مقياساً خطياً (رايت وآخرون Wright et al، ١٩٨٠). ويتمثل المعدل الثابت بوحدة قياس ثابتة، وهو ما تتسم به المقاييس فى المجال الفيزيائى. حيث إن تقدير الفرق بين أى قياسيين متتالين على نفس التدريج يكون ثابتاً، ولا يتغير هذا الفرق بتغير الأداة المستخدمة طالما كانت أداة مناسبة ولها وحدة

قياس ثابتة. ومثال لذلك أن الفرق بين درجتى الحرارة ٢٥، ٣٠ هو نفس الفرق بين الدرجتين ٣٠، ٣٥. أما فى المجال السلوكى فإن درجات الاختبار لا تعطى قياسات خطية، فعلى الرغم من قدرة الدرجات على ترتيب مستويات الأفراد إلا أنها غير قادرة على تقدير المسافات بين هذه المستويات بصورة ثابتة. وذلك لأن المسافات بين درجات الاختبار غير متساوية. (أمينة كاظم،١٩٨٨: ١٣٠)



(٣) القياس فى أكثر من بعد:

تتميز أداة القياس الفيزيائى بالدقة أكثر من المقاييس النفسية بصورتها التقليدية، كما أن المقاييس الفيزيائية تقيس شيئاً واحداً مثل قياس الطول. إلا أن اختبار ما للتحصيل الدراسي قد يقيس متغيرات أخرى مثل ثراء البيئة أو الإمكانات المادية والحضارية وغيرها، وبالتالى فإن الدرجة التى يحصل عليها الفرد على هذه الاختبارات لا تعنى درجة معرفة الفرد فقط، بل هناك متغيرات أخرى تختلف عن حقيقة ما يهدف إليه الاختبار وما يقيسه وقد تسهم فى تلك الدرجة. ويعنى ذلك أن ذلك الاختبار لا يقيس متغيراً واحداً، وبالتالى تظهر صعوبة تقرير أحادية البعد فى كثير من الاختبارات التحصيلية. مما يستلزم التأكد من أن الاختبار يعرف متغيراً واحد.

(صلاح مراد وأمين سليمان، ٢٠٠٠: ٤٢٣)

(٤) تزايد درجات الاختبار مع مستوى المتغير المقاس:

تفترض الاتجاهات التقليدية للقياس أن درجات الاختبار دالة خطية متزايدة مع السمة المقاسة. بمعنى أنه كلما زادت درجة الفرد فى الاختبار دل ذلك على زيادة القدرة أو السمة لديه، وتبعاً لنظرية الاحتمالات قد يحصل بعض الأفراد مرتفعى القدرة على درجات منخفضة، كما يحدث العكس أحيانا لذوى القدرات المنخفضة.

وفى محاولة علماء القياس التغلب على الكثير من المشكلات الناجمة عن استخدام المدخل التقليدي فى بناء الاختبارات ومواجهة أوجه القصور فيها، اتجهوا إلى أساليب القياس الموضوعى. وقد اهتموا بما يسمى بنماذج السمات الكامنة والتى يمكنها فى إطار فروضها تحقيق موضوعية القياس السلوكي وتلافى الكثير من أوجه النقص فيه.

وتفترض نظرية السمات الكامنة Latent Trait Theory أنه يمكن تفسير الأداء الملاحظ للأفراد على الاختبار بسمات أو قدرات تميز الأفراد. وهذه السمات لا يمكن قياسها مباشرة (لهذا سميت بالسمات الكامنة). وتستخدم النماذج الرياضية لتصف هذه العلاقة بين الأداء الملاحظ للأفراد على الاختبار وتلك السمات الكامنة. (كلين Kline، ١٩٩٣: ٦٥) وذلك على غرار مفهوم طاقة الحركة والطاقة الكامنة.

ويفضل كثير من علماء القياس تسميتها بنظرية الاستجابة للمفردة لأنها تربط بين احتمال الاستجابة الصواب للفرد على مفردة اختباريه معينة وبين خصائص هذه المفردة وقدرة الفرد، ويعد نموذج "راش" Rasch Model الأحادي المعلم One Parameter أبسط نماذج الاستجابة للمفردة وأكثرها استخداماً فى بناء الاختبارات العقلية وغيرها وتحليل مفرداتها.

ويرى راش أن القياس يقوم فى جوهرة على التفاعل بين الخصائص المطلوب قياسها وبين أدوات القياس، ويمكن الحكم على القياس بالموضوعية إذا كانت نتائج القياس مستقلة عن الأداة المستخدمة فى التوصل إلى هذه النتائج، ويمكن التحقق من توفر متطلبات الموضوعية فى القياس عن طريق التحقق من تعريف المفردات لمتغير واحد، واستقلالية القياس وتتمثل في تحرر صعوبة المفردة من توزيع عينة الأفراد، وتحرر قدرة الفرد من تأثير صعوبة المفردات ،والتحقق من خطية القياس .

وتهدف الدراسة الحالية إلى:

تدريج مفردات اختبار تحصيلي لعينة من متعلمي التعليم المجتمعي باستخدام نموذج (راش) تبعاً لصعوبة مفرداته وتعقدها وتقدير قدرة الأفراد على الاختبار.

أهمية الدراسة :

قد تبدو أهمية الدراسة الحالية فيما يأتى:

- الأهمية النظرية:

- استخدام منحي جديد يتمثل فى استخدام نموذج (راش) فى تدريج اختبار تحصيلي لعينة من متعلمي التعليم المجتمعي تبعاً لصعوبة مفرداته وتعقدها بما يحقق أهداف القياس الموضوعى، ويتغلب على مشكلة اعتبار مفردات تصحيح الاختبار ذات صعوبة واحدة.

- إن استخدام نموذج (راش) فى تدريج اختبار تحصيلي يحقق صدق القياس وذلك بحذف الأفراد غير الملائمين لعملية التدريج وكذلك المفردات غير الملائمة، واستبقاء استجابات الأفراد الصادقة فى تدريجها على المتغير موضوع القياس، وكذلك استبقاء المفردات الصادقة فى تعريفها للمتغير موضوع القياس.
- يحقق استخدام نموذج (راش) فى تدريج اختبار تحصيلي ثبات القياس.

- يوفر لكل من صعوبة المفردات و قدرة الفرد وحدة قياس واحدة مطلقة وهي اللوجيت وهذه الوحدة يمكن تحويلها إلي وحدات أخرى جديدة لأغراض القياس المختلفة.
- من الممكن أن يساعد نموذج (راش) فى تطوير الاختبار التحصيلي عن طريق تعديل بعض المفردات التى بها عيوب فى البناء.

- الأهمية التطبيقية:

- يتيح استخدام نموذج (راش) فى تدريج مفردات الاختبار التحصيلي الفرصة لقياس النمو أو التغير الحادث فى القدرة لدي الفرد أو الأفراد على مدى زمنى معين أو فى مواقف تجريبية مختلفة.
- تتيح استقلالية القياس التى يوفرها نموذج (راش) فى تدريج مفردات الاختبار التحصيلي سحب صور اختبارية متعادلة القياس بمعنى لا يختلف تقدير قدرة الفرد باختلاف مجموعة المفردات المستخدمة لتكون صورة اختبارية يُعتمد عليها فى تقدير قدرة الأفراد دون التعرض لمشكلة تداخل موضوعات أخري.
- توفير اختبار مدرج وفق أحدث نظريات القياس لمتعلمي التعليم المجتمعي شأنهم في ذلك شأن المتعلمين في التعليم العام ، حيث تعاني تلك الفئة من المتعلمين الإهمال وغياب الاهتمام .

تحديد المشكلة:

تتحدد مشكلة الدراسة الحالية من خلال الأسئلة التالية و التى تهدف الدراسة الإجابة عنها:

(١) ما تقدير صعوبة مفردات الاختبار التحصيلي وتدرجها باستخدام نموذج (راش) ؟

(٢) ما تقدير قدرة الأفراد لكل درجة كلية محتملة على الاختبار؟

(٣) ما مدى صدق و ثبات المقياس فى صورته النهائية؟

(٤) ما شكل دالة الاختبار التى توضح العلاقة بين المفردات وتقديرات القدرة؟

مصطلحات الدراسة:

- **تدريج الاختبار :**

يقصد بتدريج أي اختبار أو مقياس من المقاييس (التى توضع لتقدير إحدى القدرات أو الصفات السلوكية) إنشاء ميزان scale لتدرج هذه القدرة أو السمة السلوكية يتحدد علي أساسه مستوي الأفراد علي هذه القدرة أو السمة مقدرا بوحدة معرفة، ويتمثل هذا فى تدرج مفردات المقياس تبعاً لصعوبتها (فى حالة نموذج (راش) أحادى البارامتر فقط) علي متصل القدرة أو السمة وذلك بوحدة قياس معرفة.

ويستخدم نموذج (راش) لهذا الغرض، حيث يشترط توفر ما يأتي:

(أ) أحادية البعد:

حيث تتدرج صعوبة مفردات الاختبار بحيث تعرف فيما بينها متغيراً واحداً فقط. ويتضمن هذا صدق تدرج هذه المفردات وصدق قياسها للمتغير. كما يتدرج الأفراد علي متصل هذا المتغير تبعاً لمستوى أدائهم بحيث يكون هؤلاء الأفراد ذوى قدرة أحادية البعد تحدد وحدها هذا المستوى للأداء على المفردات.

(ب) استقلالية القياس:

حيث لا يختلف القياس باختلاف عينة التدريج المستخدمة طالما أنها عينة مناسبة Sample Free ، كما لا يختلف القياس باختلاف مجموعة المفردات المستخدمة فى القياس طالما أنها مفردات مناسبة Item Free .

(ج) خطية القياس:

وتعني أن يكون هناك معدل ثابت لتدرج المقياس وذلك علي المدى الواسع من متصل المتغير (القدرة) موضوع القياس. ويتمثل هذا المعدل بواسطة وحدة قياس واحدة وهي (اللوجيت Logits) والتى يمكن تحويلها إلى وحدات أخري ولا يتغير عندئذ المعني الكمي للقياس.

التعليم المجتمعي: هو الأنشطة التعليمية التي تستهدف تحسن جودة التعليم و التي تنفذ من خلال شراكه فعالة وإيجابية من المجتمع ومؤسساته لتضمن استمرارية هذه الأنشطة ، وتضافر الجهود الأهلية مع الجهود الحكومية لتقديم تدخلات ومساهمات عينية وغير عينية لإحداث تحسين في جودة العملية التعليمية.

نموذج راش: هو أحد نماذج نظرية الاستجابة للمفردة، قام بوضعه جورج راش وهو نموذج إحصائي يعمل على تقدير احتمالية إجابة الفرد عن مفردة ما إجابة صحيحة، بدلالة قدرته، ومعامل صعوبة المفردة بغض النظر عن عددمفردات الاختبار وعدد أفراد العينة.

حدود الدراسة:

تتحدد الدراسة الحالية بما يلي:-

- اختبار تحصيلي في مادة الدراسات الاجتماعية للصف السادس الابتدائي للمتعلمين في التعليم النوعي.

- تتحدد المعايير التى تم التوصل إليها فى هذا البحث بالمدى العمري للعينة المستخدمة والذى يمتد من سن (١٢) إلى سن (١٣,٩).

الإطار النظري:

تتضمن نظرية الاستجابة على المفردة Item Response Theory على عدة نماذج رياضية لوغاريتمية منها أحادى البارامتر والثنائى والثلاثى ويرجع الفضل فى كل نموذج إلى أحد علماء القياس الحديث وكذلك يرجع الفضل لتطوير تلك النماذج وتطويعها للاستخدام الفعلى إلى فريق من الباحثين ممن تولوا تطويع هذه النماذج للاستخدام العملى.

ويعتبر عالم الرياضيات الدنماركى جورج راش G. Rasch أستاذ الرياضيات والإحصاء بجامعة كوبنهاجن بالدنمارك (١٩٤٥) أول من وضع أولى نماذج هذه النظرية وهو النموذج أحادى البارامتر أو المعلم One Parameter عام ١٩٤٥ وساعده فى تطويره وتطبيقه عالم القياس النفسى بنجامين رايت B. Wright بجامعة شيكاغو (١٩٦٠) وكان السبب فى النشاط البحثى الكبير فى جامعة شيكاغو فى هذا المجال والتى أصبحت اليوم أكثر الجامعات أبحاثاً فى هذا المجال كما توصل كل من بيرنبوم Birnbaum ولورد Lord فى الولايات المتحدة إلى نماذج أخرى وقام كذلك فريق من العلماء ينشر هذا الفكر (أمثال هامبلتون Hambleton وشوبين Chopin من الولايات المتحدة وديفد أندرش D. Andrich فى استراليا)

وتنال هذه النماذج فى وقتنا الحاضر اهتماماً بحثياً كبيراً يمكن ملاحظته بوضوح فى أدبيات القياس النفسى والتربوى المعاصر. وتحاول هذه البحوث التغلب على العديد من المشكلات المتعلقة بتقدير بارامتراتها وتقييم حسن مطابقة البيانات لهذه النماذج، وإمكانية استخدامها عندما تكون المتغيرات التى تقيسها الاختبارات متعددة الأبعاد، وغير ذلك من القضايا السيكومترية والتطبيقية.

وتهدف نماذج نظرية الاستجابة للمفردة إلى تحديد العلاقة بين أداء الفرد فى اختبار ما وبين السمة أو القدرة التى تكمن وراء هذا الأداء فى صورة دالة رياضية احتمالية تقوم على افتراض أساسى أى أنه كلما ازدادت قدرة الفرد عن صعوبة المفردة ازداد احتمال حدوث الاستجابة الصحيحة للمفردة والعكس صحيح.

(أمينة كاظم، ١٩٩٤ : ١١٦-١٢٥)

نموذج "راش" أحادى البارامتر Rasch One – Parameter Model

يُعد نموذج راش الأحادى المعلم One Parameter أبسط نماذج الاستجابة للمفردة وأكثرها استخداماً فى بناء الاختبارات العقلية وتحليل مفرداتها. ويرجع الفضل فى بناء هذا النموذج إلى جورج راش G. Rasch وطوعه للتطبيق العملى العالم الأمريكى بنجامين رايت B. Wright ، ويُستخدم نموذج راش قبل تطويره فى تحليل مفردات الاختبارات التى تُعطى فيها درجة خام واحدة عن كل استجابة صحيحة على المفردة ودرجة خام صفر عن كل استجابة خاطئة. وامتد بعد ذلك إلى استخدام الاستجابات المتدرجة Rating Scales فى تحليل المفردات. كما ساهم عالم القياس النفسى الاسترالى ديفد أندرش David Andrich (١٩٨٨ : ٧٩-٨٠) فى تطوير انتشار هذه النظرية خاصة فيما يخص استخدام الاستجابات المتدرجة والمنحنى المميز للمفردات ICC .

يقوم النموذج أحادى البارامتر على اعتبار أن جميع مفردات الاختبار تتساوى تقريباً من حيث قوة التميز، وهو بذلك يفترض أن التميز يساوى واحداً صحيحاً (a=1) فى جميع المفردات مع اعتبار الخطأ المعيارى، ومن ثم فإنه يخرجه من حسابه عند تقدير صعوبات المفردات وكذلك الأمر بخصوص التخمين حيث يعتبر النموذج أن التخمين فى حالته هذه أقل ما يكون وهو بذلك يضع المسئولية على الباحث فى أن يضع مفردات اختباره بحيث يراعى فيها على قدر المستطاع قلة التخمين كما يجب عليه أيضاً عند التطبيق أن يتأكد من ضبط الموقف الاختبارى الذى قد يسهم كذلك فى ضمان قلة التخمين أو الغش الذى قد يحاوله الممتحنون / الأفراد. وتظل مسئولية الباحث مستمرة بأن يقوم بعمل تنقيح للبيانات قبل بداية التحليل محاولاً التعرف على الطلاب الذين لم يؤدوا الاختبار بشكل حريص.

ويستند النموذج إلى افتراض أساسى يؤكد على أنه كلما ازدادت قدرة الفرد عن صعوبة المفردة ازداد احتمال حدوث الاستجابة الصحيحة للمفردة والعكس صحيح، وإذا تساوت قدرة الفرد

مع صعوبة المفردة فإن احتمال حدوث الاستجابة الصحيحة (٥٠%) أى يتساوى مع احتمال حدوث الاستجابة الخطأ.

(أمينة كاظم، ١٩٩٤ : ١١٦ – ١٢٥)

. تعريف المفردات لمتغير واحد:

يفترض نموذج (راش) أن يكون الاختبار مكوناً من مفردات ذات صعوبة أحادية البعد أي تتدرج من حيث الصعوبة معرفة متغيراً واحداً، كما تتدرج على المتغير قدرة الأفراد محددة مستوى أدائهم على هذا الاختبار، وهذا يعنى أن صعوبة المفردات وقدرة الأفراد تتدرج على متصل واحد يمثل متغيرا واحداً. [كما تستخدم المسطرة لقياس طول المنضدة.]

وهكذا يتميز قياس كل من قدرة الفرد و صعوبة المفردة وفقا للنموذج بأنه يمكن التعبير عنهما بمقياس واحد، أي يمكن رد التقديرات الخاصة بالفرد والتقديرات الخاصة بالمفردة إلى نقطة أصل واحد اختباريه (صفر تدرج واحد)، مما يجعل هذا النوع من القياس يقترب إلى حد كبير من الموضوعية المعهودة للقياس الفيزيائي؛ [بمعنى أن تتدرج أداة القياس بوحدة قياس مطلقة ثابتة تتوافق مع تدرج مستويات المتغير، موضوع القياس.]

(أمينة كاظم،١٩٩٤: ١١٦–١٢٥)

٢. استقلالية القياس:

وتبدو فى:–

(أ) تحرر صعوبة المفردة من توزيع أداء عينة الأفراد:

يتميز نموذج "راش" بخاصية إمكانية تقدير صعوبة كل مفردة من مفردات الاختبار بطريقة مستقلة عن عينة الأفراد المستخدمين فى تدريج مفردات الاختبار، أي أن تدرج صعوبة المفردة بين باقي مفردات الاختبار يظل ثابتاً، مهما اختلف الأفراد الذين يؤدون هذا الاختبار طالما أن هؤلاء الأفراد مناسبون لأدائه ويعنى هذا الاستقرار النسبي لتدرج المفردات. ويمكن تحديد اختيار أى مفردة كنقطة أصل لميزان تدرج صعوبة المفردات ، ثم يتم تدريج صعوبات جميع المفردات الأخرى التى يشتمل عليها الاختبار بدءاً من هذه النقطة على الميزان. أما تقدير صعوبة المفردة فهو داله لعدد الأفراد الذين أجابوا صوابا علي المفردة وقدرة هؤلاء الأفراد.

(أمينة كاظم،١٩٨٨: من ١٣٥ إلى ١٤٠)

(ب) تحرر قدرة الفرد من تأثير صعوبة المفردات:

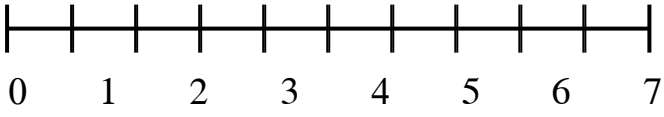
إن استخدام نموذج "راش" فى تحليل مجموعة من المفردات وتدريجها على ميزان تدرج واحد مشترك يتيح استخدام هذه المفردات فى تقدير قدرات الأفراد وفقاً لاستجاباتهم عليها وبالتالي تدريجها على المتغير المقاس الذى تعرفه هذه المفردات. وإذا كانت هناك أعداد كبيرة من هذه المفردات فإنه يمكن إعداد اختبار فرعى تسحب مفرداته منها، ثم تقدر قدرات الأفراد وفقا لهذا الاختبار الفرعي ولا تختلف قدرة الفرد سواء استخدمنا فى تقديرها كل المفردات المدرجة على المقياس الأصلي أو أى مجموعة فرعية مسحوبة منها، ومعنى هذا أن تقدير قدرة الفرد يستقل عن صعوبة المفردات المستخدمة فى القياس، ويمكن استخدام قدرة أي مجموعة حصل أفرادها على نفس الدرجة الكلية كنقطة أصل لكل ميزان تدرج قدرة الأفراد أو صعوبة المفردات. وتعتبر قدرة الفرد دالة لعدد المفردات التى أجاب عليها صواباً وصعوبة هذه المفردات.

(أمينة كاظم،١٩٨٨: من ١٣٥ إلى ١٤٠)

٣. خطية القياس Linearity:

وتعنى تساوى وحدات القياس على متصل القدرة موضع القياس بحيث يكون المعدل ثابتاً باستخدام وحدة قياس واحدة. وبذلك يكون الفرق بين أى قياسين متتالين على التدريج ثابتاً، ولا يتغير بتغير أداة القياس طالما أنها مناسبة. وهذا يختلف عن القياس الكلاسيكى حيث إن الفروق بين مفردات القياس غير متساوية (عدم خطية القياس). فإذا رغبنا فى مقارنة فردين أ، ب فإن عدم الخطية يعنى اختلاف الفرق بين قدرتهما باختلاف الاختبار المستخدم، كما فى الشكل (١-٢).

(صلاح مراد وأمين سليمان، ٢٠٠٠: ٤٢٥)



مما سبق يتبين لنا أن النظريات الحديثة فى القياس والتى من أهمها نموذج (راش) اللوغاريتمى يمكن أن تتلافى عيوب القياس التقليدى وإضفاء الموضوعية إلى القياس النفسي حيث يستطيع نموذج (راش) توفير ما يلى :

١. تدرج عناصر تصحيح مفردات الاختبار التحصيلي باستخدام نموذج (راش) بحيث تعطى لكل مفردة تقديراً يتناسب مع صعوبتها وتعقدها.

٢. إضفاء الموضوعية على لااختبارالتحصيلي والذى يتمثل فى:

أ. تحرر تقديرات صعوبات المفردات من تقديرات العينة التى استخدمت لتدريج المقياس. بمعنى لا يختلف تقدير صعوبة المفردات باختلاف خصائص عينة التدريج.

ب. تحرر تقديرات قدرة الأفراد من تقديرات المفردات التى أجابوا عليها. وهذا يعنى عدم اختلاف تقدير قدرة الأفراد أو ترتيبها باختلاف مجموعة المفردات التى قام الفرد بالإجابة عنها.

٣. تحويل الدرجات فى الاختبار التحصيلي إلى مسافات متساوية أى قياس خطى وبذلك يتمتع بجميع خصائص القياس الخطى سواء بالنسبة لتدرج صعوبة المفردات أو تقدير قدرة الأفراد.

٤. التحقق من أحادية البعد فى الاختبار التحصيلي أى فى تعريفه لمتغير القياس (النضج العقلى).

يقف دور النموذج عند حد تدريج مفردات المقياس من حيث صعوبتها وكذا تقدير قدرات الأفراد ولا تتعداها إلى المعايير التى يمكن حسابها تبعاً لأهداف القياس.

مميزات نموذج "راش":

يرى هاملتون وسوامينثان Hambleton & Swaminthan ، ١٩٨٥ : ٤٦-٤٧) أن هذا النموذج يتمتع بعدة مميزات جعلت له قبولاً وجاذبية لدى مستخدميه وكانت السبب فى انتشاره وهى :-

١ – قلة عدد البارامترات التى يتعامل معها بالمقارنة مع باقى النماذج الخاصة بكل مفردة.

٢ – قلة الصعوبات التى يواجها الباحثون عند تقدير بارامترات النموذج عنها عن النماذج الأخرى (ولعل الدليل على ذلك أن عدد برامج الحاسب التى تستخدم النموذج أحادى البارامتر يفوق عدد البرامج التى تستخدم النموذج الثلاثى والثنائى بعشرة أضعاف).

٣ – يفترض أن كل المفردات تمتلك قوة تمييز متساوية (a = ١) حيث القيمة ١ هى ظل الزاوية التى يميل بها المنحنى المميز للمفردة على متصل القدرة وهى ٤٥° درجة.

٤ – التخمين يكون فى أضيق الحدود.

مفاهيم سيكومترية مرتبطة بنموذج راش:

- **دالة المعلومات: Information Function**

يمثل مفهوم دالة المعلومات فى نظرية الاستجابة للمفردة أحد المفاهيم الأساسية للنظرية الحديثة فى القياس، حيث إنها تساعد المهتمين بالقياس التربوى فى تحديد قدر المعلومات التى تقدمها مفردة ما عن فرد ما وذلك من خلال تحديد أقصى ارتفاع للمنحنى الممثل لدالة المعلومات لتلك المفردة على متصل القدرة، وبهذا يمكننا أن نحدد أى المفردات الاختبارية تقيس المتغير المراد قياسه (السمة) بدرجة أفضل عند مستويات محددة للقدرة.

(هامبلتون وآخرون ، Hambleton, et al., ، ١٩٩١ : ٧٩)

ويرجع الفضل فى المعنى الإحصائى لدالة المعلومات إلى فيشر R.A Fisher, الذى عرفها على أنها مقلوب قياس دقة تقدير البارامترات" وقياس الدقة يعنى التباين بين عناصر التقدير ويقصد به مربع الانحراف المعيارى (σ^2) وهذا ما يمكن أن نعبر عنه بالصيغة الرياضية التالية :-

$$I = \frac{1}{\sigma 2}$$

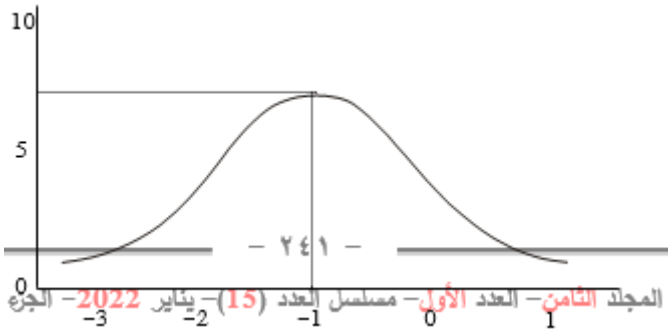
معادلة [٢-١٣]

حيث إن:-

I دالة المعلومات

σ^2 مقلوب مربع الانحراف المعيارى.

وهذا يعنى بالضرورة أنه كلما كان قدر المعلومات كبيراً كلما زادت دقة التقديرات، أى اقترب تقدير قدرة الأفراد من تقديراتها الحقيقية عند هذا المستوى من القدرة. ولكن إذا كان قدر المعلومات قليلاً فإن هذا يعنى ضعف دقة التقديرات وابتعادها عن التقديرات الحقيقية لها.



- ٢٤١ -
المجلد الثامن– العدد الأول– مسلسل العدد (15)– يناير 2022– الجزء

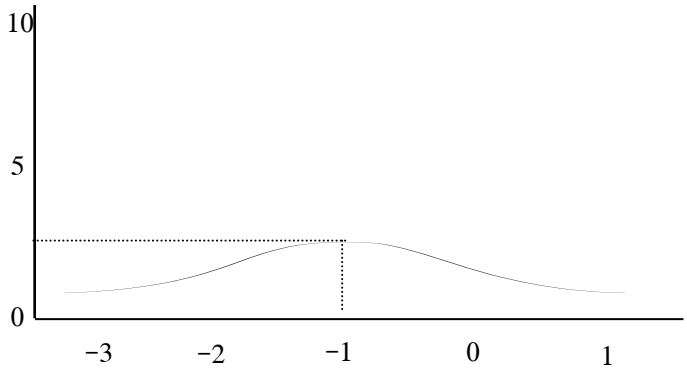
شكل (١-١) دالة المعلومات

يوضح الشكل (٢-١٠) أن قمة المعلومات التى يمكن أن نحصل عليها تكون عند القيمة (-١) للقدرة، ويقل قدر المعلومات كلما ابتعدنا عن هذه النقطة. ويجب هنا أن نذكر أن توزيع دالة المعلومات لا يعتمد على توزيع عينة الأفراد على متصل القدرة.

(بيكر Baker ، ٢٠٠١ : ١٠٧-١٠٨)

وقد استخدم أسلوب دالة المعلومات كل من هامبلتون Hambleton وسواميناثان Swaminathan وزال Zaal وغيرهم، حيث أكدوا أنها أفضل الطرق وأكثرها فعالية فى وصف كل من المفردات والاختبارات وانتقاء مفرداتها.

(أ) دالة معلومات المفردة Item Information Function:



شكل (١- ٢) دالة معلومات مفردة صعوبتها (-١)

أما الشكل (٢-١) فيوضح منحنى دالة معلومات مفردة صعوبتها (b = -1)، حيث يلاحظ أن أكبر قدر للمعلومات عند النقطة (-١) التى تمثل صعوبة المفردة وتقل المعلومات كلما ابتعدنا عن هذه النقطة على متصل القدرة. وتحسب دالة المعلومات فى النموذج أحادى البارامتر – نموذج (راش) – باستخدام المعادلة التالية:

$$Ii(\theta)=\frac{[Pi(\theta)]^2}{Pi(\theta).Qi(\theta)}$$

معادلة [٢ – ١٤]

حيث إن

$I_i(\theta)$ مقدار المعلومات المقدمة بواسطة المفردة (i) عند مستوى القدرة (θ)

$P_i(\theta)$ دالة الاستجابة للمفردة (i) عند مستوى القدرة (θ) كما تحسب فى ضوء النموذج أحادى البارامتر

$Q_i(\theta)$ وتعنى $1 - P_i(\theta)$

(بيكر Baker ، ٢٠٠١ : ١٠٨)

(ب) دالة معلومات الاختبار ككل Test Information Function:

يقوم حساب دالة المعلومات الخاصة بالاختبار ككل على افتراض مؤداه أن الاختبار هو مجموعة من المفردات، ومن ثم فإن دالة معلوماته على متغير السمة هى ببساطة مجموع دوال معلومات المفردات المكونة لهذا الاختبار على متغير السمة نفسه وتعرف بالصيغة الرياضية التالية :

$$I(\theta)=\sum_{i=1}^{N} Ii(\theta)$$

معادلة [٢–١٥]

حيث إن

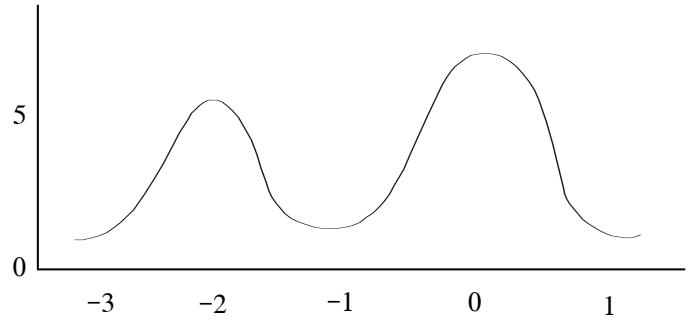
$I(\theta)$ دالة معلومات الاختبار

$\sum_{i=1}^{N} Ii(\theta)$ مجموع دوال المعلومات الخاصة بالمفردات الاختبارية

(بيكر Baker ، ٢٠٠١ : ١٠٩)

وبطبيعة الحال سوف يكون قدر المعلومات التى نحصل عليها من الاختبار أعلى من تلك التى نحصل عليها من مفردة واحدة. وبالتالى يمكن القول أن الاختبار يقيس القدرة بشكل أكثر دقة من مفردة واحدة. ولكن يجب الحذر لأن دالة المعلومات الخاصة بالاختبار ككل تزيد كلما زادت عدد مفرداته. ولكن تظل أعلى نقطة لدالة المعلومات الخاصة بالاختبار هى أفضل مستوى قدرة للأفراد الذين يمكن أن يطبق عليهم هذا الاختبار.

(بيكر Baker ، ٢٠٠١ : ١٠٩–١١١)



شكل (١-٣) منحنى معلومات اختبار لمفردات غير متساوية على طول الاختبار

ويتضح من الشكل (١-٣) أن مقدار المعلومات المقدمة من الاختبار ككل ليست متساوية على طول الاختبار، وليست بنفس القيمة عبر المستويات المختلفة للقدرة، ويرجع ذلك إلى أن مفردات الاختبار غير متساوية فيما بينها فى مقدار المعلومات المقدمة من كل منها، فمقدار المعلومات يختلف من مفردة إلى أخرى حسب مدى ملاءمة منحنيات المفردة لبيانات الاختبار. (بيكر Baker ، ٢٠٠١ : ١١٩)

أن كمية المعلومات المقدمة بواسطة الاختبار عند مستوى القدرة (θ) تكون مرتبطة عكسياً مع الدقة عند أى نقطة على متصل القدرة (θ) وهى الخطأ المعيارى لتقدير القدرة Standard Error of Estimation والذى يحصل عليه عندما يحصل على تقديرات الاحتمال الأقصى، وبصفة عامة فإن حجم الخطأ المعيارى يعتمد على:-

١ - مجموع مفردات الاختبار (أقل الأخطاء المعيارية ترتبط بالاختبارات الطويلة).

٢- مدى المطابقة بين صعوبة المفردة وقدرة الممتحنين (أقل الأخطاء المعيارية ترتبط بالاختبارات المكونة من مفردات مع معالم للصعوبة مساوية تقريباً لمعلم القدرة وهذا بخلاف الاختبارات السهلة نسبياً أو الصعبة نسبياً).

٣ - كفاءة مفردات الاختبار (أقل الأخطاء المعيارية ترتبط بالمفردات مرتفعة التمييز لأى إجابات صحيحة لا يمكن الوصول إليها بالتخمين).

(هامبلتون وآخرون Hambelton & et al ، ١٩٩١: ٩٣)

الدراسات السابقة:

١- دراسة هبة إبراهيم محمد على الناغى ٢٠١١) بعنوان أثر عدد مفردات الاختبار على تقدير قدرات الأفراد ودالة المعلومات لاختبار تحصيلي مرجعي المحك في

الرياضيات باستخدام نماذج نظرية الاستجابة للمفردة وهدفت الدراسة إلي الكشف عن أثر عدد مفردات الاختبار المكون من مفردات الاختيار من متعدد علي تقديرقدرات الأفراد ، وتكونت عينة الدراسة من ١٩٣٢ طالباً وطالبة من طلاب الصف الأول الثانوي بمدارس التعليم الثانوي العام موزعين على (٦٨٢) طالباً، و (١٢٥٠) طالبة وتم تطبيق اختبار في الجبر مكون من ٦٦ مفردة . وتشيرالنتائج إلي لا تختلف تقديرات قدرات الأفراد على الاختبار التحصيلي مرجعي المحك باختلاف عدد مفردات الاختبار باستخدام النموذج اللوغاريتمي ثلاثي المعلم اختلافاً دالاً إحصائياً.

٢- دراسة حمادنة وبن خالد)٢٠١٣ (هدفت هذه الدراصة إلي بناء مقياس اتجاهات نحو العنف الدراسي .لدى عينة من مستخدمي موا قع التواصل الاجتماعي بجامعة آل البيت، وفق النظرية الحديثة للقياس .وتم صياغة ٤٤ مفردة موزعة علي الأبعاد بشكل غير متساو .وتم تطبيق الأداة علي عينة مؤلفة من ٥٥٠ طالباً و طالبة بجامعة آل البيت ممن يستخدمون موا قع التواصل الاجتماعي . وعند تحليل استجابات أفراد العينة – وفق نظرية الاستجابة للمفردة IRT لم يتبق من مفردات المقياس سوى ٢٩ مفردة، موزعة في ثلاث مجالات رئيسة. وًتمتع المقياس بصورته النهائية ٢٩ مفردة بخصائص سسيكومترية مناسبة، حيث بلغت قيمة معامل الثبات للمقياس ٩٤. ،كما تمتع المقياس بدلالات متعددة للصدق.

٣- دراسة (لمياء رفقي محمود، سائد أحمد صباح ٢٠١٦) بعنوان بناء اختبار تحصيلي في الهندسة لطلبة الصف الخامس الأساسي باستخدام نموذج راش . وهدفت الدراسة إلى بناء اختبار تحصيلي في الهندسة لطلبة الصف الخامس الابتدائي باستخدام نموذج راش، وتكونت عينة الدراسة من (٢١٦) . وتم بناء اختبار تحصيلي مكون من (٣٠) مفردة من نوع الاختيار من متعدد. وأظهرت النتائج أن الاختبار ثابت، حيث بلغ معامل الثبات للمفردات (٩٨,.) ومعامل الثبات للأفراد (٧٨,.)وأن فقرات الاختبار غطت مدى الصعوبة بشكل جيد، فتراوحت قيمة الصعوبة للفقرات ما بين (-٢٩,٣ – ٦٩,٢ (لوجيت، وظهر أن معظم الفقرات تقع ضمن إحصائيات الملاءمة لنموذج راش.

٤- دراسة (معزوز جابر علاونه ٢٠١٧) عنوان الدراسة توظيف نموذج راش في بناء اختبار تحصيلي في منهاج التاريخ للصف الثاني الثانوي الأدبي بفلسطين . وهدفت الدراسة إلى توظيف نموذج راش في بناء اختبار تحصيلي في مقرر التاريخ، و قام الباحث ببناء اختبار تحصيلي لمنهاج التاريخ، تكون بصورته الأولية من (١٣٦) مفردة من نوع الاختيار من متعدد بأربعة بدائل. وتم تطبيقه على عينة مؤلفة من (٣٢٧) طالبا وطالبة من طلبة الصف الثاني الثانوي الأدبي في مدينة نابلس، ؛ وأشارت النتائج أن ثبات الأفراد الذين طبق عليهم الاختبار (٠,٨٠)في حين بلغ معامل ثبات فقرات الاختبار(٠,٧٣)

٥- دراسة حسن بكر الشريف ٢٠١٨ هدفت الدراسة للتعرف على الخصائص السيكومترية لاختبار مقرر مهارات التعلم والتفكير على عينة مكونة من١٦٢ من طلاب جامعة الملك فيصل باستخدام النظرية الكلاسيكية وكذلك تحديد الخصائص السيكومترية لنفس الاختبار باستخدام نموذج راش . أشارت النتائج إلي اختلاف المفردات المكونة للاختبار باختلاف النموذج الإحصائي المستخدم .

٦- دراسة (أشواق المرواني وشاهر سليمان٢٠١٩) بعنوان ,, أثر موقع البديل الصحيح في اختبار اختيار من متعدد على دقة تقدير معالم الفقرة وفق النموذج ثنائي المعلمة لنظرية الاستجابة للفقرة،، وهدفت الدراسة إلى الكشف عن أثر موقع البديل الصحيح في اختبار اختيار من متعدد ذي أربعة بدائل على دقة تقدير معلمة الصعوبة والتمييز للفقرات، وعلى تقديرات معاملات الثبات التجريبي ومعاملات ثبات الأفراد، وفق النموذج ثنائي المعلمة لنظرية الاستجابة للفقرة .وتكونت عينة الدراسة من ٥٠٠ طالبة من طالبات الصف الثالث المتوسط بمدينة ينبع بالمملكة العربية السعودية ، وتم بناء اختبار تحصيلي من نوع اختيار من متعدد ذي أربعة بدائل في مادة العلوم مكون في صورته النهائية من (٣١) مفردة، و من ثلاثة نماذج لها نفس المحتوي والمتغير هو موقع البديل الصحيح، حيث النموذج الأول يكون موقع البديل الصحيح على البديلين (أ، ب)، والنموذج الثاني يكون موقع البديل الصحيح على البديلين (ج، د)، والنموذج الثالث يتوزع البديل الصحيح بشكل عشوائي على جميع البدائل الأربعة. وأظهرت النتائج عدم وجود فروق دالة إحصائيا بين متوسطات دقة تقدير معلمة الصعوبة والتمييز للمفردات تعزي لموقع البديل الصحيح. وكذلك عدم وجود فروق ذات دلالة إحصائية بين

متوسطات تقديرات معاملات الثبات التجريبي تعزي لموقع البديل الصحيح، وربما يكون السبب هو أن الطالبات لا يمتلكن مهارة حكمة الاختبار. بينما توجد فروق ذات دلالة إحصائية بين متوسطات تقديرات معاملات ثبات الأفراد تعزي لموقع البديل الصحيح. وفي ضوء النتائج تم تقديم عدد من التوصيات من أبرزها توعية المعلمين والمعلمات إلى ضرورة العناية بفقرات اختبار الاختيار من متعدد عند إعدادها وكذلك ضرورة العناية بتوزيع الإجابات الصحيحة على جميع البدائل بشكل عشوائي.

***تعليق على الدراسات التى اهتمت بتناول الخصائص السيكومترية لاختبارات الاختيار من متعدد.**

١-تراوح عدد العينة المشاركة فى الدراسات السابقة بين ١٢٧ متعلما إلى ١٩٣٢ متعلما، وتنوعت دراستهم بين المرحلة الجامعية والمدرسة العليا والمدرسة الثانوية والمدرسة الابتدائية

٢-تراوحت أعداد المفردات بين ٣٠ مفردة و١٣٦ مفردة. واختلفت المواد الدراسية من المهارات المعرفية لطلاب الجامعة إلي الرياضيات (جبر وهندسة) وتاريخ.

٣- اهتمت دراسات علاء أيوب ٢٠٠١ وهبة إبراهيم محمد على الناغى ٢٠١١ و لمياء رفقي محمود، سائد أحمد صباح ٢٠١٦ و معزوز جابر علاونه ٢٠١٧ و أشواق المرواني وشاهر سليمان٢٠١٩ بالطرق الإحصائية المناسبة لمعالجة بيانات الاختبار.

٤- أجريت الدراسات السابقة علي متعلمين بالتعليم العام الابتدائي والثانوي والجامعي.

٥-اتفقت نتائج الدراسات السابقة علي صلاحية نموذج راش في تدريج مفردات الاختبارات المختلفة رغم حدوث بعض الأخطاء المنهجية التى ربما تكون قد وقعت فيها تلك الدراسات مثل صغر حجم العينات المستخدمة فى بعض الدراسات بالإضافة إلى أطوال الاختبارات المستخدمة كأدوات فى بعض الدراسات مما يؤدى إلى اختلاف النتائج بالإضافة إلى أن بعض الدراسات استخدمت اختبارات أعدها الباحثون والبعض الآخر استخدم اختبارات أو أجزاء من بطاريات معدة سلفاً بغرض الاستفادة من هذه الدراسات فى تحسين واقعنا التعليمى وحتى يمكن تحقيق الدقة والموضوعية فى التقويم التربوى.

إجراءات الدراسة: يتكون مجتمع البحث الحالي من طلبة الصف السادس الابتدائي في مدارس التعليم المجتمعي بمحافظتي الغربية وكفر الشيخ للعام الدراسي ٢٠١٩/٢٠٢٠والبالغ عددهم ٧٥٥ متعلماً ومتعلمة.

لتحقيق هدف الدراسة قام الباحث ببناء اختبار تحصيلي في مادة الدراسات الاجتماعية للصف السادس الابتدائي مكون من ٣٢مفردة من نوع الاختيار من متعدد.

ويمثل مقرر الصف السادس الابتدائي ٢٠١٨/٢٠١٩، الفصل الدراسي الأول الموضوعات الآتية:

البيئة الزراعية ٢- البيئة الصناعية ٣- مصر في ظل الحكم العثماني

الحملة الفرنسية علي مصر

وقد اختار الباحث منهم الموضوع الأول وهو: البيئة الزراعية

ويرجع سبب اختيار الباحث لهذه الوحدة إلى أنها تمثل محصلة نهائية لموضوع مترابط يعتبر مدخلاً لدراسة مادة الدراسات الاجتماعية وأساساً يتم البناء عليه بعد ذلك، بالإضافة إلى خواص وتطبيقات تشكل مدخلاً و أساساً مهماً لفهم واستيعاب العلاقات والارتباطات الجغرافية والنجاح في دراسة الدراسات الاجتماعية في المرحلة التالية من التعليم، وقد بلغ العدد النهائي للمفردات الاختبارية ٣٢ مفردة من نوع الاختيار من متعدد.

وقام الباحث بالاطلاع ومراجعة المعايير القومية للتعليم في جمهورية مصر العربية الخاصة بمادة الدراسات الاجتماعية للوقوف على أهم المعايير التي ينبغي أن تتحقق من خلال الوحدة محل الدراسة، وكذا مؤشرات تحقيق كل منها؛ ومن ثم إمكان قياسها إجرائيا.

وقد اتبع الباحث الخطوات التالية في بناء هذه المجموعة من المفردات:

تحليل محتوى الموضوعات المتضمنة في الوحدة (البيئة الزراعية)

٢- تحديد الأهداف الرئيسية الموجودة وتحليلها إلى مكوناتها وتقييمها.

٣- تقييم الأهداف السلوكية من قبل الباحث بعد فترة من صياغتها، وإعادة تقييمها من قبل محكمين متخصصين.

٤- بناء المفردات الاختبارية من نوع الاختيار من متعدد.

٥- تقييم المفردات الاختبارية من قبل الباحث بعد فترة من صياغتها، وإعادة تقييمها من قبل محكمين متخصصين.

٦- تقدير الخصائص السيكومترية لمجموعة المفردات الاختبارية.

وفيما يلي بعض التوضيح لكل خطوة من هذه الخطوات:

١- تحليل محتوى الموضوعات المتضمنة في الوحدتين الدراسيتين:

قام الباحث بتحليل محتوى الوحدة الدراسية (البيئة الزراعية) في ضوء تصنيف ثلاثي لجوانب التعلم في مادة الدراسات الاجتماعية يشتمل على العناصر الثلاثة الآتية: المفاهيم–العلاقات والتعميمات–المهارات) ،وذلك بغرض التعرف على التوزيع النسبي لجوانب التعلم الثلاثة في محتوى الوحدتين الدراسيتين المشار إليها.

جدول (١)

نتائج تحليل محتوى الوحدتين الدراسيتين في ضوء جوانب التعلم الثلاثة

(المفاهيم–العلاقات والتعميمات–المهارات)

المفاهيم		العلاقات والتعميمات		المهارات		المجموع	
العدد	النسبة	العدد	النسبة	العدد	النسبة	العدد	النسبة
٢٨	%٢٥,٩	٧١	%٦٥,٨	٩	%٨,٣	١٠٨	%١٠٠

ثم قام الباحث بعرض تحليل محتوى الوحدة الدراسية (البيئة الزراعية) الذي قام به على مجموعة من المحكمين. تتكون من: أساتذة جامعيين في مجال القياس والتقويم، والمناهج وطرق التدريس، ومستشار المواد الاجتماعية، وخبراء المادة، ومدرسون أوائل، حيث قام كل منهم بمراجعة تحليل المحتوى الذي قام به الباحث، وتحديد العناصر التي يتفق أو يختلف فيها مع الباحث في هذا التحليل.

٢– تحديد الأهداف الرئيسة الموجودة بالوحدة الدراسية وتحليلها إلى مكوناتها وتقييمها:

قام الباحث في هذه الخطوة بتحديد الأهداف الرئيسة المرجو تحقيقها لدى المتعلمين بعد دراستهم لوحدة (البيئة الزراعية)، مسترشداً في ذلك بنتائج تحليل محتوى الوحدة الدراسية، وآراء مجموعة من السادة المحكمين، وكذلك خبرة الباحث الشخصية في تدريس مادة الدراسات الاجتماعية، وإعداده لدليل تقويم المتعلم والتي تزيد عن عشرين عاما، وتوصل الباحث لبعض الأهداف الرئيسية التي ينبغي أن يتمكن منها المتعلم بعد دراسة هذه الوحدة الدراسية. وهى:

يعرف مفهوم البيئة.

يميز بين البيئات المصرية.

يتعرف خصائص البيئات المصرية .

يفسر اسباب تنوع البيئات المصرية .

يحدد مناطق البيئة الزراعية علي خريطة مصر.

يحلل الخصائص الطبيعية للبيئة الزراعية.

يفسر أسباب ندرة النباتات الطبيعية والحيوانات البرية في البيئة الزراعية.

يقدر أهمية الحفاظ علي الموارد الطبيعية في البيئة الزراعية.

يحدد خصائص سكان البيئة الزراعية.

يحلل خريطة تركز السكان في البيئة الزراعية في مصر.

يحدد العادات الايجابية والعادات السلبية لسكان البيئة الزراعية.

يفسر أسباب هجرة سكان البيئة الزراعية.

يذكر نتائج هجرة سكان البيئة الزراعية.

يحدد بعض المشكلات المرتبطة بسكان البيئة الزراعية.

يقترح حلولا مناسبة لحل المشكلة السكانية.

يحدد بعض الخدمات التي تقدمها الدولة لسكان البيئة الزراعية.

يميز بين حقوقه وواجباته تجاه البيئة التي يعيش فيها.

يحدد أهم الموارد المرتبطة بالبيئة الزراعية .

يحدد أهم الأنشطة الاقتصادية التي يمارسها سكان البيئة الزراعية.

يفسر أسباب اختلاف الأنشطة الاقتصادية في البيئة الزراعية.

يستنتج العلاقة بين الموارد والأنشطة الاقتصادية .

يحدد بعض المشكلات التي تعاني منها البيئة الزراعية.

يفسر دور الإنسان في ظهور المشكلات في البيئة الزراعية.

يحدد الآثار السلبية لبعض الأنشطة الاقتصادية علي البيئة.

يقترح وسائل وأساليب للحفاظ علي الموارد الطبيعية في البيئة الزراعية.

يحدد أهم المشروعات الاقتصادية ومجالات التنمية في البيئة الزراعية.

وقام الباحث بعرض هذه الأهداف الرئيسة على مجموعة من المحكمين لإبداء الرأي حول مدى أهمية هذه الأهداف لتلاميذ الصف السادس الابتدائي، ومدى كفايتها في تحقيق الهدف من تدريس الوحدة الدراسية المتعلقة بها، واقتراح ما يرونه مناسباً من تعديلات سواء بالحذف أو الإضافة.

وقد اتفق المحكمون جميعاً على أهمية تمكن المتعلمين من هذه الأهداف الرئيسة، ورأوا أنها تعد أساساً مهماً لدراسة مقررات الدراسات الاجتماعية في المراحل التالية، وأنها في صورتها النهائية بعد إجراء التعديلات التي اقترحها بعضهم ،لا تتطلب المزيد من الإضافات.

٣-تقييم الأهداف السلوكية:

قام الباحث بمراجعة الأهداف السلوكية وعددها ٢٦ هدف بعد فترة من صياغتها، وقام بتعديل بعض الكلمات والألفاظ، ثم قام الباحث بعرض هذه الأهداف على مجموعة من المحكمين بغرض التحقق من وضوحها، وسلامة صياغتها اللغوية، وملاءمتها للهدف الأساسي لأداة البحث، وكذلك التحقق من كفايتها في تحقيق الأهداف الرئيسية للوحدة الدراسية (البيئة الزراعية). وقد أعد الباحث استمارة خاصة لتقييم هذه الأهداف اشتملت على الهدف الأساسي لأداة الدراسة، وكذلك نواتج التحليل السلوكي لأهداف الوحدة الدراسية، وطلب من كل محكم أن يجيب عن الأسئلة التالية:

١-هل الهدف واضح أم لا؟

٢-هل الهدف سليم الصياغة لغويا؟

٣-هل المستوى المعرفي للهدف مناسب؟

٤-هل الأهداف كافية لتحقيق المهارات الرئيسية للوحدة الدراسية؟

٥-هل الأهداف ملائمة للهدف الأساسي لأداة البحث؟

وقام الباحث بإجراء التعديلات التي اقترحها المحكمون على الأهداف، ثم قام بعرضها عليهم مرة أخرى، حيث وافقوا عليها وأقروها جميعاً.

وقد بلغ عدد الأهداف بعد عرضها على السادة المحكمين وإجراء التعديلات المقترحة عليها (٢٦) هدف حيث لم تتعد ملاحظات السادة المحكمين عليها بعض التعديلات المتعلقة بالصياغة.

* تقدير الاتساق الداخلي بين المحكمين:

نظراً لأن الباحث قد اعتمد في الخطوات السابقة عند إعداد أداة بحثه على مجموعة من السادة المحكمين، وحيث إن الباحث يضع آراء السادة المحكمين موضع التقدير، ويأخذ بها في تعديل وإثراء أداة بحثه، والتي سوف يعول عليها في صدق نتائج بحثه، لذا فقد كان من الضروري أن يتأكد الباحث من موضوعية وصدق وثبات آراء مجموعة المحكمين، فقام بحساب الاتساق

الداخلي بين المحكمين في كل من الخطوات السابقة من إعداد أداة الدراسة، والجدول التالي يوضح نتائج حساب الاتساق الداخلي بين المحكمين.

جدول (٢)

نتائج حساب الاتساق الداخلي بين المحكمين

المحكمون	معامل ارتباط بيرسون بعد حذف المحكم	معامل الثبات ألفا بعد حذف المحكم	معامل الثبات في وجود الجميع
محكم ١	٠,٩٩٩٠	٠,٩٩٨٨	
محكم ٢	٠,٩٩٨٩	٠,٩٩٨٨	
محكم ٣	٠,٩٩٥٩	٠,٩٩٩٠	٠,٩٩٩١
محكم ٤	٠,٩٩٥٦	٠,٩٩٩٠	
محكم ٥	٠,٩٩٩٨	٠,٩٩٨٨	
محكم ٦	٠,٩٩٩٢	٠,٩٩٨٨	
محكم ٧	٠,٩٩١٦	٠,٩٩٤	
محكم ٨	٠,٩٩٧٩	٠,٩٩٩٠	

ويتضح من الجدول السابق أن معامل الثبات للمحكمين مرتفع، وأن هناك اتساقا مرتفعا بين جميع المحكمين، وهذا يشير إلى جدية كل المحكمين فيما طلب إليهم من تحكيم، الأمر الذي يجعل الباحث ينفذ كل التعليمات والملاحظات التي اقترحوها بدرجة عالية من الثقة.

٤-صدق التدرج: ويعتمد على أن تدرج مفردات مجموعة من الصور الاختبارية تقيس نفس السمة على ميزان واحد مشترك باستخدام نموذج "راش" يعني أن هذه المفردات تحقق شرط أحادية القياس أي أنها تعرف متغيراً واحداً.

وتتحقق أحادية البعد بتحقق صدق المحتوى، وبملاءمة كل من الأفراد والمفردات لأسس القياس الموضوعي تبعاً لمحكات الملاءمة الخاصة ببرنامج WINSTEPS المستخدم في التحليل الإحصائي. حيث توضح هذه المحكات مدى قياس المفردة لما تقيسه باقي المفردات على متصل المتغير موضوع القياس، كما تبين مدى اتساق تدرج قدرة الفرد مع تدرج قدرات باقي الأفراد على

هذا المتصل. وتم حذف المفردات غير الملائمة وبلغ عددها سبع مفردات عند استخدام طريقة التعادل المتزامن، وتسع مفردات في حالة التعادل بتثبيت بارامترات المفردات المشتركة وفقا لتلك المحكات الإحصائية. وذلك لأسباب مختلفة كأن تعتمد المفردة على غيرها من المفردات، أو أن تكون المفردة ضعيفة الصياغة أو غير ذلك. كما تم حذف الأفراد غير الملائمين وذلك لأسباب مختلفة أيضاً كأن يلجأ الطالب إلى التخمين أو الغش أو يتسم أداؤه بالبطء الشديد أو السرعة أو يستجيب بعكس ما هو متوقع منه.

وبحذف الحالات غير الملائمة من الأفراد والمفردات يستبقى الأفراد الصادقون وتكون المفردات المتبقية هي المفردات الصادقة في تعريف متغير التحصيل الدراسي في مادة الدراسات الاجتماعية؛ مما يوفر الموضوعية في تقدير الأفراد على متصل التحصيل الدراسي، كما يوفرها نموذج راش من خلال شروطه. وبالإضافة إلى المفردات المستبقاة يستبقى أيضاً الأفراد الصادقون في استجاباتهم على تلك المفردات ويعنى هذا توفر شرطي الصدق والثبات لتقديرات كل من صعوبات مفردات الاختبار، وقدرات الأفراد أي يعنى تحقق صدق وثبات القياس.

٥- بناء المفردات الاختبارية:

من المعروف أن المفردات التي تقيس هدفاً سلوكيا معينا تعد بمثابة التعريف الإجرائي لهذا الهدف، مما يتطلب أن يعرف كل هدف سلوكي تعريفا إجرائياً دقيقاً، من خلال مجموعة متكاملة من المفردات الاختبارية التي تعتني ببنائها الجوانب المتعلقة بهذا الهدف. بحيث يمثل كل هدف من خلال مجموعة المفردات التي تقيسه، فإن كانت المفردات التي تقيس الهدف السلوكي قد اعتنى ببنائها وتعرف الهدف تعريفاً إجرائياً دقيقاً، فإن هذه المفردات تستوعب الفترة التي تمثل هذا الهدف، وتغطى كل جوانبه بحيث لا يشتمل على فجوات، حيث إن وجود فجوة فارغة لا يغطيها سؤال؛ يدل على وجود أحد الجوانب المتعلقة بالهدف المراد قياسه بلا قياس، مما يؤثر بالطبع في دقة وكفاءة تعريف الأهداف الذي ينعكس على دقة وكفاء الأداء.

وبناء على ذلك فقد قام الباحث الحالي بدراسة كل هدف سلوكي على حدة دراسة دقيقة، وتحديد كافة الأوجه التي يمكن أن يقاس من خلالها، وبحث كافة الأفكار التي يمكن أن تصمم من خلالها مفردات اختبارية يمكن باستخدامها قياس كل هدف سلوكي.

وعلى هذا فإن عدد البنود التي تقيس هدفاً سلوكياً معيناً لم يكن مساوياً للعدد الذي يقيس هدفاً آخر، فقد اعتمد عدد البنود التي تقيس هدفاً سلوكيا معيناً على كم الأفكار والجوانب المتعلقة بهذا

الهدف، والتي يمكن تطويعها ، وبناء بنود يعتمد عليها لقياس هذا الهدف، وكذلك على مدى صعوبة هذا الهدف ومدى أهميته.

وقد حاول الباحث الحالي أن يغطى جميع الجوانب المتعلقة بكل هدف، بحيث تشتمل المفردات التي تقيس الهدف السلوكي علي جميع الجوانب المتعلقة به.

وبناء على ما سبق، فقد قام الباحث ببناء مجموعة من المفردات من نوع الاختيار من متعدد ، لقياس كل هدف من الأهداف السلوكية التي تمت صياغتها ومراجعتها في الخطوة السابقة، واشتملت كل مفردة على خمس بدائل إحداها هي الإجابة الصحيحة للمفردة.

وقد بلغ عدد المفردات الاختبارية التي قام الباحث ببنائها في بادئ الأمر ٣٦ مفردة، وذلك قبل عرضها على السادة المحكمين، وقد ارتبط مستوى صعوبة كل مجموعة من المفردات الاختبارية بمستوى صعوبة الهدف السلوكي الذي تقيسه ومستواه المعرفي (تذكر–فهم–تطبيق).

٦– تقييم المفردات الاختبارية:

بعد بناء المفردات الاختبارية، أعد الباحث استمارة لتقييم هذه المفردات؛ اشتملت على الأهداف السلوكية، والمفردات الاختبارية التي يقيسها، حيث يقابل كل هدف سلوكي مجموعة المفردات التي تقيسه، وترك الباحث الأهداف السلوكية والمفردات الاختبارية مدة خمسة عشر يوماً، ثم أعاد قراءتها، وقام بتعديل بعض الكلمات، ثم قام الباحث بتقديمها لمجموعة من السادة المحكمين وعددهم اثني عشر محكماً، حيث أعطى لكل محكم نسخة من الاستمارة المعدة للتحكيم، وطلب من كل محكم أن يقرأ كل هدف سلوكي بعناية، ثم يقرأ كل مفردة من مجموعة المفردات الاختبارية التي تقابله ،ثم يجيب عن الأسئلة التالية بوضع علامة √ في المكان الذي يراه مناسباً:

س١: هل المفردة الاختبارية تقيس الهدف السلوكي المقابل لها بالفعل؟

س٢: هل صياغة المفردة الاختبارية واضحة وسليمة؟

س٣: هل الاختيارات المتعلقة بالمفردة مناسبة وتحقق شروط هذا النوع من المفردات؟

كما طلب من السادة المحكمين أيضاً كتابة أي ملاحظات أو اقتراحات يراها مناسبة،بشأن تعديل أي مفردة اختبارية تحتاج إلى التعديل، بغرض إثراء المفردات الاختبارية ورفع كفاءتها.

وتبين للباحث أن السادة المحكمين قد اتفقوا بنسبة (٩١%) على أن هذا العدد من المفردات الاختبارية يعد مناسبا للهدف الذي أعد من أجله.

كما اقترح السادة المحكمون حذف أربعة مفردات رأوا فيهما أفكار مكررة، أو متشابهة، وأنها لا تقيس الهدف السلوكي بشكل مباشر وصعبة جدا ولا تناسب تلاميذ الصف السادس الابتدائي، وبالتالي تحتاج إلى جهد كبير في الحل بما لا يجعلها تتناسب مع طبيعة بنود الاختيار من متعدد.

وقام الباحث بمراجعة المفردات الاختبارية، وحذف الأربعة مفردات التي وقعت في نطاق ما اقترحه السادة المحكمون حتى صار عدد المفردات الاختبارية ٣٢ مفردة، ثم أعاد استمارة التحكيم لكل محكم، وطلب منه الإجابة عن التساؤلات الثلاثة السابقة نفسها.

كما قام الباحث بإجراء التعديلات التي اقترحها بعض السادة المحكمون ،والتي تتلخص في إعادة صياغة بعض المفردات الغامضة، أو تعديلها بما يتناسب مع تلاميذ هذه المرحلة، وكذلك أيضا مراجعة بعض البدائل الخاصة ببعض المفردات.

واقترح السادة المحكمون تطبيق هذه المفردات في جلسة واحدة مدتها ساعة ونصف ، لتشابه في ذلك اختبار نهاية الفصل الدراسي.

* تعليمات التطبيق:

بعد تجميع المفردات الاختبارية في صورتها النهائية، قام الباحث بصياغة تعليمات التطبيق، متضمنة: الهدف من التطبيق، وعدد المفردات الاختبارية المطبقة، وتعليمات الإجابة عنها، وطريقة تسجيل العلامات الدالة على الإجابة في المربعات المخصصة لذلك في ورقة الإجابة، وكذلك أيضا مثالا محلولاً يوضح كيفية الإجابة عن أي مفردة.

* التجربة الاستطلاعية للمفردات الاختبارية:

تم طباعة وإعداد المفردات الاختبارية للتجريب، ووزعت الإجابات لكل بند أو مفردة توزيعاً عشوائيا، وذلك حتى لا يكتشف المختبر دلالة الإجابات في المفردات الاختبارية.

وتم تطبيق المفردات في صورتها المبدئية على عينة قوامها ٢٥٠ تلميذاً من تلاميذ الصف السادس الابتدائي في شهر نوفمبر ٢٠١٩م وذلك كما هو موضح في الجدول التالي:

جدول(٣)

العدد الكلى للعينة الاستطلاعية وعدد المتعلمين في كل مدرسة من مدارس العينة الاستطلاعية

المدرسة	الإدارة	عدد المتعلمين
الأحمدية الابتدائية	غرب طنطا	٧٥
صادق الرافعى الابتدائية	شرق طنطا	١٠٢
السادات الابتدائية	شرق طنطا	٧٣
المجموع		٢٥٠

وهؤلاء المتعلمين قد انتهوا من دراسة وحدة (البيئة الزراعية)

وقد تراوح العمر الزمني لأفراد العينة الاستطلاعية بين ١٢و ١٤ عام بمتوسط قدرة ١٣,١ وانحراف معياري قدرة ٠,٠٤ ، وقد اختار الباحث هذه المجموعة بالتحديد ليقوم بالتجريب الاستطلاعي للمفردات الاختبارية عليها ، لأنها تمثل العينة النهائية للدراسة الحالية تمثيلاً كاملاً.

وعند تطبيق المفردات الاختبارية، راعى الباحث عدم الرد على أسئلة واستفسارات أفراد العينة الاستطلاعية، إلا في حدود تعليمات الاختبار التحصيلى، وذلك حتى لا يتجه المتعلمين نحو إجابة معينة للمفردة، وطلب من أفراد العينة اختيار إجابة واحدة فقط لكل مفردة من خمس إجابات (ا، ب، جـ، د، هـ)

وقام الباحث بتسجيل الزمن الذي استغرقه كل فرد من أفراد العينة في إجابة جميع المفردات الاختبارية كاملة ، وذلك بعد الانتهاء من الإجابة ، حتى يستطيع الباحث الإفادة من ذلك في تحديد الزمن اللازم لإجابة المفردات الاختبارية ، في ضوء متوسط الزمن لأفراد العينة الاستطلاعية، وتم تصحيح المفردات الاختبارية بواسطة الكمبيوتر.

وكان الهدف من التجريب الاستطلاعي للمفردات الاختبارية، هو: التعرف على المفردات الغامضة التي تحتاج إلى إعادة صياغة، أو التي تحتاج إلى تعديل. وكذلك للتعرف على البدائل غير المناسبة لتعديلها، أو إعادة صياغتها. وكذلك للتعرف على أي أخطاء فنية، أو لغوية في صياغة المفردات الاختبارية ؛ لضبطها ،وتعديلها، وكذلك أيضا لتحديد الخصائص السيكومترية لمجموعة المفردات ككل من صدق وثبات وقوة تشتيت للبدائل، وكذلك أيضا للتعرف على أي

مشكلات تتعلق بالتطبيق، والتي ربما تواجه الباحث أثناء التطبيق، ليمكن تداركها وحلها قبل التطبيق على العينة النهائية.

٧-تقدير الخصائص السيكومترية لمجموعة المفردات الاختبارية:

١- التحقق من صدق محتوى الاختبار: حيث إن صدق الاختبار مسألة على جانب كبير من الأهمية في تقرير صلاحية أو فائدة أي فرض علمي، أي أنه يحدد قيمة الاختبار ومغزاه، ويعد الاختبار صادقا إذا أمكن أن يصف أداء الفرد وصفا دقيقا، ولذا فإن الاعتماد على أسلوب صدق المحتوى Content validity أو الصدق الوصفي Descriptive validity كما يطلق عليه popham يعد مؤشرا جيداً لصدق الاختبار.

وهذا النوع من الصدق يقدم معلومات عن مدى تمثيل المفردات الاختبارية للأهداف التعليمية، في إطار الموضوعات الرئيسية المتعلقة بالوحدة الدراسية ،التي تم إعداد مجموعة المفردات الاختبارية من أجل قياس تحصيل المتعلمينفيها.

وتهدف هذه الطريقة للتحقق مما إذا كانت مواصفات النطاق السلوكي المراد قياسه محددة تحديداً تفصيلياً واضحاً ودقيقاً، وما إذا كان هناك اتفاق على أن مجموعة المفردات الاختبارية تعد بمثابة عينة كافية وممثلة للنطاق السلوكي الذي تقيسه أم لا.

ولذا فقد تحقق الباحث من صدق محتوى المفردات الاختبارية، وذلك بإعداد استمارة خاصة بذلك ، وقدم الباحث نسخة منها لكل محكم من السادة المحكمين، تتكون من عشرة خبراء في المجال، مرفقاً بها الأهداف السلوكية للوحدتين الدراسيتين، ويقابل كل هدف منها مجموعة المفردات الاختبارية التي تقيسه، ثم طلب من كل محكم أن يقرأ كل هدف سلوكي، وكذلك المفردات الاختبارية التي تقيسه، ثم يجيب عن الأسئلة التالية الواردة في الاستمارة وذلك بوضع علامة في العمود الذي يراه مناسباً، وهذه الأسئلة هي:

١- هل كل مفردة اختبارية تعد مقياساً مناسباً للهدف المقابل لها؟

٢- هل تعد مجموعة المفردات الاختبارية بمثابة عينة ممثلة لأهداف وحدة (البيئة الزراعية)؟

٣- هل مجموعة المفردات الاختبارية ككل متسقة مع الهدف الأساسي لأداة البحث؟

وقد استفاد الباحث من ملاحظات ومقترحات السادة المحكمين في إجراء بعض التعديلات اللازمة لبعض المفردات التي استلزمت ذلك، مما أدى إلى اتفاقهم على عناصر التحكيم بنسبة ١٠٠% ، وبذلك تعد مجموعة المفردات الاختبارية صادقة، وصالحة لقياس ما وضعت لقياسه.

ب- ثبات المفردات الاختبارية: قام الباحث بحساب ثبات المفردات الاختبارية بطريقة العلاقات الارتباطية، استناداً إلى درجات مجموعة تلاميذ العينة الاستطلاعية، فبعد تطبيق الاختبار على مجموعات المتعلمين الممثلين للعينة الاستطلاعية، قام الباحث بتصحيح الاختبار، وقام بإدخال البيانات الناتجة عن التصحيح في الحاسب الالى، في مصفوفة واحدة تتكون من ٣٢ مفردة و٢٥٠ تلميذاً هم أفراد العينة الاستطلاعية، وقام الباحث بحساب ثبات المفردات باستخدام برنامج (SPSS)، من خلال تحديد العلاقات الارتباطية بين درجة كل مفردة اختبارية ، والمجموع الكلى للمفردات الاختبارية في حالة حذف كل مفردة على حدة، وذلك للتعرف على أثر وجود المفردة ضمن المفردات أو حذفها على قيمة الثبات للمفردة ككل، وذلك بغرض التعرف على مدى اتساق كل مفردة مع مجموعة المفردات ككل.

وتبين أن قيمة معامل الثبات ألفا للمفردات الاختبارية ككل دون حذف أي منها هى (٠,٨٦٣٢) ، وتراوحت قيم معامل الثبات ألفا للمجموع الكلى للمفردات في حالة حذف كل مفردة على حدة ما بين (٨٥٧٣) و (٠,٨٦٥٨).

كما قام الباحث بحساب ثبات التدرج: حيث إن تدرج مفردات اختبار التحصيل الدراسي الكلية معاً على ميزان تدرج واحد مشترك وفقا لنموذج "راش" – بعد حذف الحالات غير الملائمة من الأفراد والمفردات – يعنى تحقق شروط النموذج ومنها شرط استقلال القياس، ويعنى ذلك ثبات تقديرات كل من الصعوبة والقدرة وعدم تأثرهما باختلاف الصورة الاختبارية المستخدمة والمسحوبة من الاختبار الكلي أو باختلاف عينة الأفراد المستخدمة في تدريجه (أمينة محمد كاظم،٢٠٠٠، ٢٣١).

وقد تم حساب معامل ثبات تقدير كل من صعوبة المفردات وقدرة الأفراد باستخدام معامل ثبات كيودريتشاردسون الذي يوفره برنامج WINSTEPS. وتشير قيم معاملات الثبات الموضحة بجدول ( ) إلى أن الصور الاختبارية الثلاث تتمتع بتقديرات تتميز بتقديرات ثبات مرتفعة لتقديرات صعوبات المفردات، كما توفر تقديرات ثبات مرتفعة لتقديرات قدرة الأفراد حيث بلغ معامل ثبات تقدير القدرة ( ) تقريباً.

حساب الخطأ المعياري لتقديرات كل من صعوبات المفردات وقدرات الأفراد: يوفر برنامج الحاسب الآلي WINSTEPS تقديرات للخطأ المعياري لتقدير قدرة كل فرد، وكذا صعوبة كل مفردة، والتي تعد مؤشرا لدقة القياس. ويتضح من جدول () أن قيم الخطأ المعياري لصعوبات

المفردات تعتبر منخفضة بوجه عام، حيث تتراوح هذه القيم بين (٠,١٣: ٠,١٥) لوجيت. كما يوضح جدول() انخفاض قيم الخطأ المعياري لتقديرات قدرات الأفراد على الصور الاختبارية الثلاث حيث تراوحت بين(٠,٤٥: ٠,٥٥) لوجيت.

جـ- قوة تشتيت البدائل: قام الباحث بحساب قوة تشتيت البدائل للمفردات الاختبارية باستخدام معادلة قوة تمييز البدائل (قوة تشتيت البدائل)

قوة تشتيت البدائل= **عدد اختيارات المشتت من طلاب الارباع الأعلى-عدد اختياراته لطلاب الارباع الادني**

حيث يرى الباحث أن تعديل البدائل في ضوء قوة تشتيتها ؛ سيرفع من مدى تمييزها، ومعامل ثباتها، ووجد الباحث أن تشتيت البدائل كان إيجابياً، لذا لم يغير الباحث أي بدائل في ضوء حساب قوة تشتيت البدائل.

*** زمن الإجابة على الاختبار التحصيلى:**

تم تحديد الزمن الكلى للإجابة على الاختبار التحصيلى عن طريقة عينة استطلاعية بلغت ٢٥٠ متعلماً موزعة على ثلاث مدارس هي مدرسة السادات الابتدائية ، ومدرسة صادق الرافعي الابتدائية، ومدرسة الأحمدية الابتدائية.

وتم حساب الزمن الكلى الذي استغرق تلاميذ العينة جميعاً وقسمت على عد المتعلمين:

الزمن الكلي الذي استغرقه جميع المتعلمين ١٣٧٥٢ دقيقة ÷ ٢٥٠ = ٥٥ دقيقة تقريباً

وأضاف الباحث خمس دقائق للمراجعة، فأصبح زمن الاختبار ستون دقيقة.

وكان الباحث يكتب زمن الاختبار على السبورة، حتى يراه جميع المتعلمين ، وعند منتصف الوقت كان الباحث يذكِر المتعلمين بالوقت، ثم يعيد تذكيرهم قبل انتهاء الوقت بعشر دقائق، ثم يعيد على مسامعهم (باقي خمس دقائق) ثم (باقي دقيقتان). وبعد ذلك ، وعند انتهاء الوقت يطلب الباحث من المتعلمين أن يتركوا القلم ، ويغلقوا كراسة الأسئلة، ويبدأ بتسلم كراسة الأسئلة متزامنة مع ورقة الإجابة.

*** ظروف إجراء الاختبار:**

حرص الباحث على تهيئة ظروف إجراء الاختبار، حتى يكون مطمئناً إلى أن درجة المتعلم على الاختبار تمثل بدرجة مقبولة تحصيله الحقيقي، وتمثل ذلك الحرص من قبل الباحث في:

١- اختيار الغرفة المجهزة لإجراء الاختبار بشكل جيد، من حيث: الإضاءة الجيدة، والتهوية الجيدة أيضاً.

٢- تم التطبيق على جميع أفراد العينة بواسطة الباحث، ولم يسمح بتدخل أي فرد في أي مدرسة.

٣- بالتعاون مع إدارات المدارس التي كان الباحث يقوم بتطبيق الاختبار بها،تم الاتفاق على أن تكون الغرف، أو الفصول ،التي يجرى بها الاختبار بعيدة عن أماكن الضوضاء، ومناطق تجمع المتعلمين.

٤- تم اختيار أوقات مناسبة لإجراء الاختبار في ضوء الظروف المدرسية، فكان الباحث يتجنب إجراء الاختبار بعد حصص التربية الرياضية، أو أثناء المباريات الرياضية، أو أثناء الحفلات المدرسية ، والرحلات، أو أثناء إجراء الاختبارات الشهرية أو النهائية.

٥- حرص الباحث على عدم مقاطعة المتعلمين أثناء الإجابة، حيث كان الباحث يزودهم بالتعليمات قبل بداية الإجابة عن الاختبار.

٦- حرص الباحث على عدم إعطاء أي توضيح لتلميذ بمفرده، وحرص الباحث على أن يكون أي توضيح لمجموع المتعلمين في الفصل.

٧- حرص الباحث على عدم السماح بالغش لجميع المتعلمين، أو التهاون في المراقبة.

٨- حرص الباحث على عقاب المتعلمين الذين يحاولون الغش بنقلهم من مكانهم إلى مكان آخر، فإذا عادوا لمثل هذا السلوك، فإن الباحث كان يستبعدهم من الفصل، ومن العينة نهائياً.

٩- حرص الباحث على عدم إعطاء الاختبار أهمية أكثر مما يستحق.

*** تصحيح الاختبار:**

اعتمد الباحث طريقة التصحيح الآلي للاختبارات، وذلك عن طريق إدخال استجابة المتعلم كالآتي: أ=١ ، ب=٢ ، جـ=٣ و د=٤ و هـ=٥ والمتروك=٩ ، وإذا أجاب المتعلم إجابتين يسجل رقم ٨.

ومن مزايا هذه الطريقة سهولة تغذية الحاسب الآلي (الكمبيوتر) بنواتج عملية التصحيح على مستوى المفردة الواحدة، وذلك لإجراء التحليلات الإحصائية المطلوبة، ومعرفة توزيع إجابات المتعلمين على البدائل.

قام الباحث بإدخال بيانات المتعلمين في الحاسب الآلي علي هيئة مصفوفة تتكون من ٣٢ عمود تمثل مجموعة المفردات الاختبارية التحصيلية ، و٧٥٠ صف تمثل مجموعة المختبرين .

قام الباحث بحساب صدق الاختبار التحصيلي.

قام الباحث بحساب ثبات الاختبار التحصيلي .

تحليل بيانات مفردات اختبار التحصيل الدراسي باستخدام برنامج WINSTEPS

حللت بيانات اختبار التحصيل الدراسي وفقاً لنموذج راش باستخدام برنامج الحاسب الآلي WINSTEPS بهدف تحديد واستبعاد البيانات غير الملائمة لأسس القياس الموضوعي وتدريج المفردات المختلفة.

وفيما يلي نتائج هذا التحليل بالتفصيل :

(١) حذف البيانات التامة والصفرية

يقوم برنامج WINSTEPS آليا باستبعاد البيانات التامة قبل التحليل والتي تتضمن:كل فرد حصل على الدرجة التامة (النهائية) أو الدرجة صفر على مجموعة المفردات التي أداها. وكذلك كل مفردة أجمع أفراد العينة على إجابتها سواء صواباً أو خطأ حيث تعتبر خارج نطاق القياس.

ولم يسفر التحليل الأولى لبيانات الاختبار التحصيلي الكلي عن استبعاد أي فرد أو مفردة تبعا لهذا المحك.

(٢) حذف الأفراد غير الملائمين لأسس القياس

بعد إجراء التحليل الأولى للبيانات باستخدام برنامج WINSTEPS تم تحديد وحذف الأفراد الذين تجاوزوا محكات الملاءمة الإحصائية (±٢) التي يوفرها برنامج WINSTEPS. وقد بلغ عدد الأفراد الذين حذفوا في ضوء هذه المحكات (١٤٥) تلميذ وتلميذة ، وبذلك لم تتجاوز نسبة هؤلاء الطلاب (١٩,٥ %) من عدد أفراد عينة التدريج (٧٤٤) تلميذ وتلميذة.

(٣) حذف المفردات غير الصادقة (غير الملائمة لأسس القياس الموضوعي) في تدرجها على متصل متغير التحصيل الدراسي

أعيد تحليل البيانات بعد حذف الأفراد غير الملائمين لأسس القياس الموضوعي ، بهدف تحديد وحذف المفردات التي تجاوزت محكات الملاءمة الإحصائية (±٢) التي يوفرها برنامج WINSTEPS. وقد بلغ عدد المفردات التي تجاوزت هذه القيمة من تدريج إحصاءات الملاءمة

(١) مفردات. ويلاحظ أن عدد المفردات المحذوفة لم يتجاوز (٣,١ %) فقط من العدد الكلى للمفردات (٣٢) مفردة.وبذلك يكون العدد الكلي للمفردات الملائمة لأسس القياس الموضوعي (٣١) مفردة.

ويتضمن جدول (٤) التالي ملخصاً لنتائج تحليل البيانات وفقا للخطوات التي اتبعت في تدريج مفردات اختبار التحصيل الدراسي.

جدول (٤)

ملخص نتائج تحليل بيانات اختبار التحصيل الدراسي

	عدد المفردات	عدد الأفراد	متوسط التقديرات		معامل ثبات التقديرات	
			للصعوبة	للقدرة	للصعوبة	للقدرة
التحليل الأول قبل الحذف	٣٢	٧٤٤	٠,٠٠	٠,٦٨	٠,٩٩	٠,٦٣
التحليل الثاني بعد حذف الأفراد	٣٢	٥٩٩	٠,٠٠	٠,٩٧	٠,٩٩	٠,٥٩
التحليل الثالث بعد حذف المفردات	٣١	٥٩٩	٠,٠٠	٠,٩٧	٠,٩٩	٠,٦٠

٤- التدريج النهائي لمفردات الاختبار التحصيلي

للحصول على التدريج النهائي لمفردات الاختبار التحصيلي أُتبعت الخطوات التالية:

- إعادة تحليل البيانات للمرة الثالثة بعد حذف المفردات غير الملائمة بهدف :

-تدريج مفردات الاختبار التحصيلي على تدريج واحد مشترك بصفر واحد مشترك تبعاً لصعوبتها باللوجيت وتكوين التدريج النهائي لمفردات الاختبار التحصيلي

-حساب تقديرات القدرة المقابلة لكل درجة خام كلية محتملة على الاختبار

- تحويل تقديرات كل من الصعوبة والقدرة من وحدة (اللوجيت) إلى وحدة (المنف) وذلك باستخدام معادلة التحويل الخطى – السابق الإشارة إليها في فصل الإطار النظري – وتناسب وحدة (المنف)مدى تقديرات القدرة المقابلة لكل درجة كلية محتملة على اختبار التحصيل .

ويتضمن جدول (٥) التدريج النهائي لمفردات اختبار التحصيل الدراسي في صورتها النهائية مرتبة تبعاً لصعوبتها بوحدتي اللوجيت والمنف بالإضافة إلى الخطأ المعياري مقدراً بوحدتي اللوجيت والمنف ، وذلك بعد حذف الأفراد والمفردات غير الملائمة لأسس القياس الموضوعي.

جدول (٥) التدريج النهائي لمفردات اختبار التحصيل الدراسي في صورتها النهائية

كود المفردة	الصعوبة باللوجت	الصعوبة بالمنف	الخطأ باللوجت	الخطأ بالمنف
V425	8.87	94	1.83	9.15
V419	3.29	66	0.14	0.70
V411	1.51	57	0.09	0.45
V407	1.34	56	0.09	0.45
V429	1.07	55	0.09	0.45
V431	0.98	54	0.09	0.45
V414	0.83	54	0.09	0.45
V409	0.8	54	0.09	0.45
V405	0.71	53	0.09	0.45
V427	0.66	53	0.09	0.45
V426	0.51	52	0.09	0.45
V415	0.48	52	0.09	0.45
V404	0.4	52	0.09	0.45
V412	0.23	51	0.09	0.45
V402	0.02	50	0.09	0.45
V428	-0.08	49	0.1	0.50
V401	-0.17	49	0.1	0.50
V422	-0.25	48	0.1	0.50
V408	-0.27	48	0.1	0.50

كود المفردة	الصعوبة باللوجت	الصعوبة بالمنف	الخطأ باللوجت	الخطأ بالمنف
V420	-0.46	47	0.1	0.50
V421	-0.51	47	0.11	0.55
V413	-0.52	47	0.11	0.55
V432	-0.66	46	0.11	0.55
V410	-0.68	46	0.11	0.55
V424	-0.94	45	0.12	0.60
V430	-1.09	44	0.12	0.60
V423	-1.14	44	0.13	0.65
V403	-1.22	43	0.13	0.65
V417	-1.22	43	0.13	0.65
V416	-1.62	41	0.15	0.75
V418	-2.01	39	0.17	0.85

ويتضمن جدول (٦) تقديرات القدرة المقابلة لكل درجة كلية محتملة على اختبار التحصيل الدراسي مقدرة بوحدتي اللوجيت والمنف.

جدول (٦) تقديرات القدرة المقابلة لكل درجة كلية محتملة على اختبار التحصيل الدراسي مقدرة بوحدتي اللوجيت والمنف.

الدرجة الخام	القدرة باللوجت	الخطأ باللوجت	القدرة بالمنف	الخطأ بالمنف
0	-5.06	1.84	24	9.20
1	-3.81	1.03	30	5.15
2	-3.05	0.75	34	3.75
3	-2.58	0.63	37	3.15

2.80	38	0.56	−2.23	4
2.60	40	0.52	−1.93	5
2.45	41	0.49	−1.68	6
2.30	42	0.46	−1.45	7
2.25	43	0.45	−1.25	8
2.15	44	0.43	−1.05	9
2.10	45	0.42	−0.87	10
2.10	46	0.42	−0.7	11
2.05	47	0.41	−0.53	12
2.05	48	0.41	−0.36	13
2.00	49	0.4	−0.2	14
2.00	49	0.4	−0.03	15
2.05	50	0.41	0.13	16
2.05	51	0.41	0.3	17
2.05	52	0.41	0.47	18
2.10	53	0.42	0.64	19
2.15	54	0.43	0.82	20
2.20	55	0.44	1.01	21
2.30	56	0.46	1.21	22
2.40	57	0.48	1.42	23
2.50	58	0.5	1.66	24
2.70	59	0.54	1.93	25
2.95	61	0.59	2.24	26
3.30	63	0.66	2.63	27
3.95	65	0.79	3.16	28

5.40	69	1.08	3.99	29
9.35	76	1.87	5.31	30

تعليق على نتائج تدريج مفردات اختبار التحصيل الدراسي في صورته النهائية

يتضح من جدول (٦) أن مفردات الاختبار وعددها (٣١) مفردة تغطى مدى الصعوبة بشكل مناسب ، كما يلاحظ أن هناك عدداً مناسباً من المفردات التي تغطي المستويات المختلفة من الصعوبات على مدى متصل الصعوبة. وجاءت قيم الأخطاء المعيارية صغيرة نسبياً مما يدل على دقة وثبات تقديرات صعوبات المفردات.

كما يتضح من جدول (٦) أن مدى تقديرات القدرة المقابلة لكل درجة كلية محتملة على اختبار التحصيل الدراسي يتراوح بين (٢٤: ٧٦) منف ، وتعتبر هذه الحدود هي أقصى حدود للقدرة يمكن تقديرها باستخدام اختبار التحصيل.

حساب صدق وثبات القياس للاختبار:

أولا : حساب الصدق: اتبع في حساب صدق القياس الطرق التالية:

- صدق المحتوى
- صدق التدرج (صدق الملاءمة) كما يوفره " نموذج راش ".

(١) صدق المحتوى

يشير صدق المحتوى إلى درجة التوافق بين مفردات الاختبار والأهداف أو محتوى جدول المواصفات وللتأكد من صدق المحتوى للصورة الأولية لاختبار التحصيل الدراسي اعتمد الباحث على:

- صدق البناء
- صدق المحكمين

ويرتبط صدق المحتوى بخطة بناء الاختبار؛ حيث تم بناء الاختبار وفق خطة محددة تشمل تحديد المجال السلوكي الذي يغطيه الاختبار وما يتضمنه من تحليل للمحتوى وتحديد لأوجه التعلم المتضمنة فيه، والأهداف السلوكية وما يترتب على ذلك من إعداد جدول المواصفات والذي في ضوئه يتم صياغة المفردات.

وقد سار الباحث في بنائه للصورة الأولية لاختبار التحصيل الدراسي في هذه الخطوات. وبهذا يكون قد تحقق النوع الأول من صدق المحتوى وهو صدق البناء كما يتمثل في إعداد

المفردات بحيث تقيس متغيراً واحداً وهو التحصيل الدراسي في مادة الدراسات الاجتماعية، وأن تكون الإجابة على كل مفردة من مفردات الاختبار مستقلة عن الإجابة عن أي من مفرداته الأخرى.

أما بالنسبة لصدق المحكمين فقد قام الباحث بعرض مجموعة المفردات الاختبارية على مجموعة من المحكمين وذلك لمراجعتها، والحكم على مدى تحقيقها للأهداف السلوكية التي وضعت لقياسها ومدى ملاءمتها للمستوى المعرفي الذي وضعت لقياسه.

وفى ضوء آراء المحكمين تم تعديل المفردات التي تحتاج إلى تعديل لتكون صالحة لقياس الهدف الذي تقيسه. هذا وبتحقق التطابق بين المفردات الاختبارية والأهداف التي تقيسها يتحقق صدق البناء للمفردات وبالتالي يتحقق صدق المحتوى.

(٢) صدق التدرج (صدق الملاءمة كما يوفره نموذج راش)

وكما سبق فإن تدرج مفردات مجموعة من الاختبارات – يفترض أنها تقيس نفس السمة – على ميزان تدرج واحد مشترك باستخدام نموذج "راش" يعنى أن هذه المفردات تحقق شرط أحادية القياس أي أنها تعرف متغيراً واحداً.

وما يتيحه نموذج "راش" من أحادية في القياس يحقق صدق تدرج المفردات في قياس المتغير موضوع القياس، وكذا صدق تدرج قدرات الأفراد على متصل المتغير والذي يقوم على صدق استجاباتهم على المفردات.

وتتحقق أحادية البعد بتحقق صدق المحتوى، وبملاءمة كل من الأفراد والمفردات لأسس القياس الموضوعي تبعاً لمحكات الملاءمة الخاصة ببرنامج WINSTEPS المستخدم في التحليل الإحصائي. حيث توضح هذه المحكات مدى قياس المفردة لما تقيسه باقي المفردات على متصل المتغير موضوع القياس، كما تبين مدى اتساق تدرج قدرة الفرد مع تدرج قدرات باقي الأفراد على هذا المتصل.

وكما سبق فقد تم حذف المفردات غير الملائمة من اختبار التحصيل الدراسي أربعة بدائل وفقا لتلك المحكات الإحصائية. وذلك لأسباب مختلفة كأن تعتمد المفردة على غيرها من المفردات أو أن تكون المفردة ضعيفة الصياغة أو غير ذلك. كما تم حذف الأفراد غير الملائمين وفقا لتلك المحكات الإحصائية وذلك لأسباب مختلفة أيضاً كأن يلجأ الطالب إلى التخمين أو الغش أو يتسم أداؤه بالبطء الشديد أو السرعة أو يستجيب بعكس ما هو متوقع منه.

وبحذف الحالات غير الملائمة من الأفراد والمفردات يستبقى الأفراد الصادقون وتكون المفردات المتبقية هي المفردات الصادقة في تعريف المتغير _ التحصيل الدراسي في مادة الدراسات_ مما يوفر الموضوعية في تقدير الأفراد على متصل التحصيل الدراسي، كما يوفرها نموذج راش من خلال شروطه. وبالإضافة إلى المفردات المستبقاة يستبقى أيضاً الأفراد الصادقون في استجاباتهم على تلك المفردات – ويعنى هذا توفر شرطي الصدق والثبات لتقديرات كل من صعوبات مفردات الاختبار، وقدرات الأفراد أي يعنى تحقق صدق وثبات القياس.

ثانياً : ثبات القياس

اعتمد حساب ثبات القياس على الطرق التالية:

- ثبات التدرج

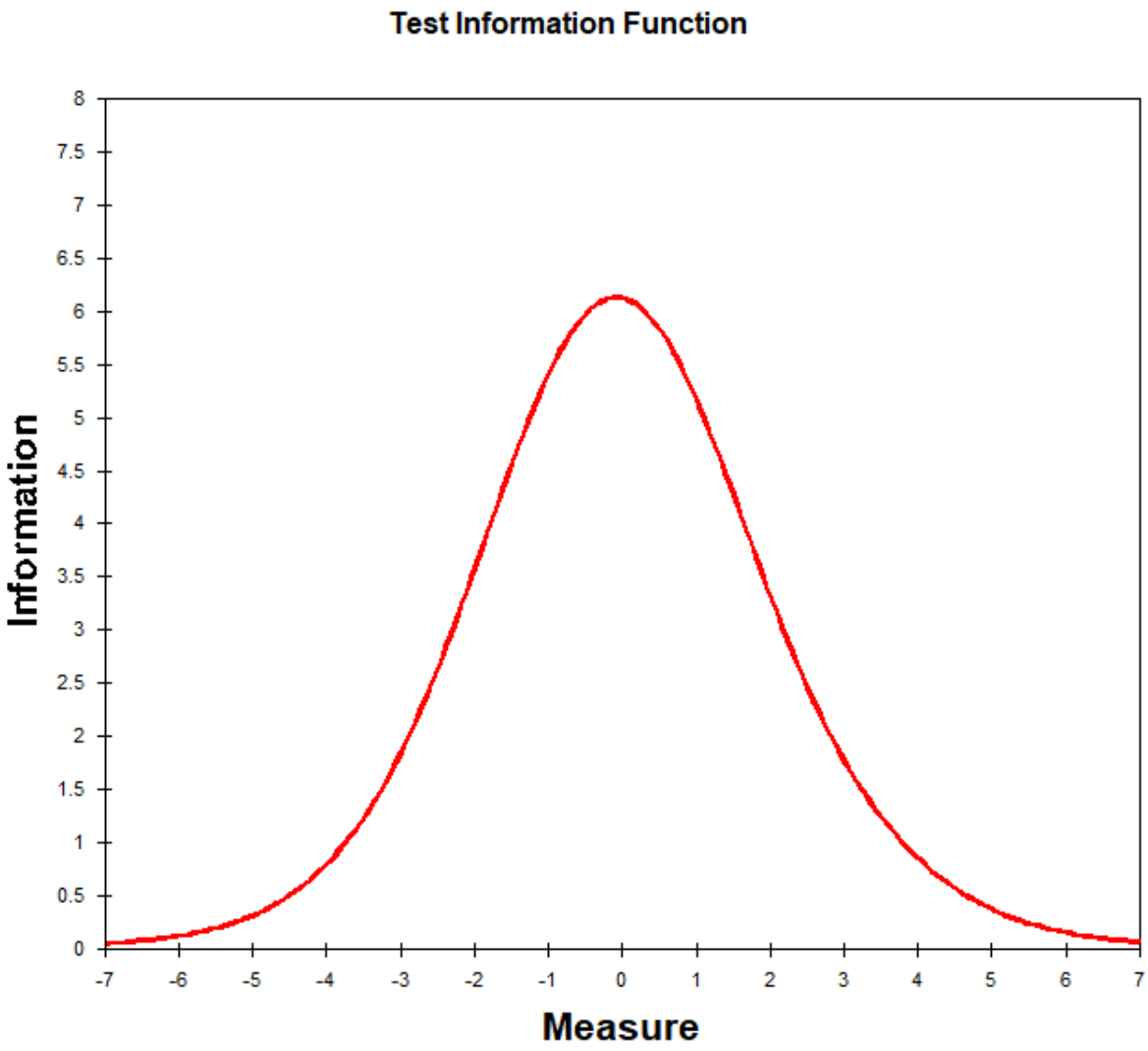
إن تدرج مفردات اختبار التحصيل الدراسي معاً على ميزان تدرج واحد مشترك وفقا لنموذج "راش" – بعد حذف الحالات غير الملائمة من الأفراد والمفردات – يعنى تحقق شروط النموذج ومنها شرط استقلالية القياس ، ويعنى ذلك ثبات تقديرات كل من الصعوبة والقدرة وعدم تأثرهما باختلاف الصورة الاختبارية المستخدمة والمسحوبة من الاختبار الكلي أو باختلاف عينة الأفراد المستخدمة في تدريجه(أمينة محمد كاظم،٢٠٠٠، ٢٣١).

وقد حسب معامل ثبات تقدير كل من صعوبة المفردات وقدرة الأفراد باستخدام معامل ثبات كيودريتشاردسون الذي يوفره برنامج WINSTEPS. وتشير قيم معاملات الثبات الموضحة بجدول ( ) إلى أن اختبار التحصيل الدراسي يتميز بتقديرات ثابتة للصعوبة والقدرة.

- حساب الخطأ المعياري لتقديرات كل من صعوبات المفردات وقدرات الأفراد

يعد تقدير الخطأ المعياري لتقدير قدرة كل فرد ، وكذا صعوبة كل مفردة مؤشراً لدقة القياس. ويوفر برنامج الحاسب الآلي WINSTEPS هذه التقديرات للخطأ المعياري ويتضح من جدول( ) أن قيم الخطأ المعياري لصعوبات المفردات تعتبر منخفضة بوجه عام.

ويوضح الشكل البياني التالي دالة المعلومات لاختبار التحصيل الدراسي



شكل (٣): دالة معلومات اختبار التحصيل الدراسي أربعة بدائل

ويتضح من شكل (٣) السابق أن أقصى كم من المعلومات يمكن الحصول عليه من اختبار التحصيل الدراسي يقابل مستوى القدرة ( ) لوجيت. كما يشير المنحنى إلى تمتع الاختبار بدرجة عالية من الثبات خاصة في المناطق المتوسطة للقدرة

تركز الدراسة الحالية على استخدام أحد نماذج نظرية الاستجابة للمفردة – وهو نموذج "راش" أحادى البارامتر – فى تطوير أداة من أدوات القياس التحصيلي شائعة الاستخدام وذلك عن طريق إنشاء تدريج مفردات الاختبار، وحذف المفردات غير الملائمة للقياس فى ضوء المحكات الإحصائية الملائمة الخاصة بالنموذج.

وقد أمكن لهذه الدراسة الاستفادة من خطية القياس التى يتميز بها نموذج "راش"، حيث تتوفر وحدة قياس واحدة لكل من صعوبة المفردة وقدرة الفرد وهى وحدة اللوجيت والتى أمكن تحويلها فى الدراسة الحالية إلى وحدة المنف.

كما توصلت الدراسة الحالية إلى المعايير التى تفسر على أساسها مستوى قدرة الأفراد على الاختبار الكلى، والاختبار المختصر فى إطار أسلوب القياس جماعى–المرجع .

وفى ضوء ما توصلت إلية الدراسة الحالية، فإن الباحث يخلص إلى مجموعة من الدراسات المقترحة وهى كما يلى:

دراسات خاصة بنموذج "راش" ونظرية الاستجابة للمفردة:

- على أساس ما يحققه نموذج "راش" من خطية التدريج وأحادية القياس واستقلاليته، وما تتميز به أداة القياس التى طورت باستخدام النموذج من تحقيق لموضوعية القياس، توصى الدراسة باستخدام نماذج نظرية الاستجابة للمفردة فى تطوير وبناء الاختبارات التحصيلية المعروفة والشائعة الاستخدام للتغلب على أوجه النقد التى وجهت لتلك المقاييس.
- كانت من المشكلات التى واجهت الباحث عند بداية الدراسة العينات المناسبة للتحليل وعلاقتها بعدد المفردات وعلاقة ذلك بعدد الأفراد المشتركين اللازم. لذلك توصى الدراسة الحالية فى التوسع فى الدراسات الخاصة بشروط التحليل باستخدام الأفراد المشتركين.

دراسات خاصة بالاختبار التحصيلي:

- توصى الدراسة باستخدام النموذج الثنائى المعلم Two Parameters Model فى تدريج الاختبار التحصيلي كطريقة للتعرف على قوة تمييز مفردات الاختبار.
- توصى الدراسة الحالية بإجراء دراسات تكون بمثابة امتداد لهذه الدراسة بهدف التوسع فى المدى العمرى للعينات المستخدمة ..
- توصى الدراسة الحالية بإجراء دراسات تكون بمثابة امتداد لهذه الدراسة بهدف التعرف على قدرة الاختبار فى صورته الجديدة ووضع المعايير الخاصة بهم على الاختبار.

- توصى الدراسة الحالية بإجراء دراسة بمثابة امتداد لهذه الدراسة تقوم بتعديل المفردات المحذوفة وإضافة مفردات جديدة للاختبار وإعادة تدريج الاختبار.

أولا: المراجع العربية:

١- أمين محمد صبرى نور الدين (١٩٩٥). بعض الخصائص السيكومترية لمقياس ستانفورد-بنيه المعدل لدى عينة من الأطفال فى مرحلة ما قبل المدرسة، رسالة ماجستير، القاهرة: كلية التربية، جامعة عين شمس.

٢- أمينة محمد كاظم (١٩٨٨). دراسة نظرية نقدية حول القياس الموضوعى للسلوك (نموذج راش)، سلسلة الكتب المتخصصة، الكويت: مؤسسة الكويت للتقدم العلمى.

٣- أمينة محمد كاظم (١٩٨٩). مقاييس القدرات، تعريب وتجريب للمقاييس البريطانية للقدرات ، الدليل، المقياس الأول: سرعة تجهيز المعلومات، الكويت: دار العلم.

٤- أمينة محمد كاظم (١٩٩٦). نماذج السمات الكامنة فى: أنور محمد الشرقاوى، سليمان الخضرى الشيخ، أمينة محمد كاظم، نادية محمد عبد السلام، اتجاهات معاصرة فى القياس والتقويم النفسى والتربوى، القاهرة: مكتبة الأنجلو المصرية:٢١٨-٥٤٣.

٥- خالد حسن بكر الشريف: فاعلية نموذج راش في تدريج مفردات اختبار مقرر مهارات التعلم والتفكير على عينة من طلاب كلية التربية جامعة الملك فيصل مجلة العلوم الإنسانية والاجتماعية - جامعة الإمام محمد بن سعود الإسلامية

٦- رشدى فام (يونيو١٩٩٧). حجم التأثير الوجه المكمل للدلالة الإحصائية، المجلة المصرية للدارسات النفسية، العدد ١٦، المجلد السابع، ، القاهرة: مكتبة الأنجلو المصرية:٥٧ -٧٥.

٧- صفوت فرج (١٩٨٠). القياس النفسى، القاهرة: دار الفكر العربى.

صلاح أحمد مراد، أمين على سليمان (٢٠٠٠). الاختبارات والمقاييس فى العلوم النفسية والتربوية- خطواتها إعدادها وخصائصها، دار الكتاب الحديث.

٨- صلاح الدين محمود علام (١٩٩٥). الاختبارات التشخيصية مرجعية المحك فى المجالات التربوية والنفسية والتدريبية، ط١، القاهرة: دار الفكر العربى.

٩- صلاح الدين محمود علام (٢٠٠٠). القياس والتقويم التربوى والنفسى أساسياته وتطبيقاته وتوجهاته المعاصرة، ط١، القاهرة: دار الفكر العربى.

١٠- علاء الدين عبد الحميد أيوب محمد (٢٠٠١). دراسة سيكومترية للمهارات المعرفية باستخدام نظرية الاستجابة للمفردة IRT ، رسالة ماجستير غير منشورة: كلية التربية أسوان، جامعة جنوب الوادى.

١١- فؤاد أبو حطب (١٩٧٧). بحوث في تقنين الاختبارات النفسية، المجلد الأول ، القاهرة: الأنجلو المصرية.

ثانياً: المراجع الإنجليزية:

1 1- Andrich, D. (1998). Rasch Models for Measurement, Beverly Hills, London: Sage Publications.

2 2- Andrich, D. & others (2001). The Manual of Program "RUMM2010" For Analysing Assessment and Attitude Questionnaire Data, part1:Getting Started, Austria: RUMM Laboratory Pty Ltd.

3 3- Andrich, D. & others (2001). The Manual of Program "RUMM2010" For Analysing Assessment and Attitude Questionnaire Data, part2:Extending RUMM2010, Austria: RUMM Laboratory Pty Ltd.

4 4- Bachman, L.F. (1997). *Fundamentals Considerations in Language Testing*, 4th edition, London: Oxford University Press.

5 5- Baker, F. (2001). The Basics of Item Response Theory, *Eric Clearinghouse on Assessment & Evaluation*, U.S.A.:

University of Maryland, College Park, MD.

6 6- Embretson, S.E. & Reise, S. P. (2000). *Items Response Theory for Psychologists*, London, Lawrence Erllaum Associates.

7 7- Ground, N.E.& Linn, R.L. (1994). *Measurement and Evaluation Teaching*, 6th Edition, London: Collier Macmillan Publishers.

8- Hambleton, R.K. & Swaminthan, H. & Rogers, H. J. (1991). *Fundamentals of Item Response Theory*, Teler Road, Newbury Park, California, U.S.A.: Sage Publications Inc.

9- Hambleton, R.K. & Swaminthan, H. (1995). *Item Response Theory, Principles and Application*, U.S.A: Boston, , Norwell, Assinippi, U.S.A.: Kluwer Academic Publishers.

10- Kaplan, R.M. & Saccuzzo, D.P. (2001). *Psychological Testing Principles, applications and Issues*, 5th Edition, U.S.A.: Wadsworth Thomson Learning.(A Modified Edition)

11- Kline, P. (1993). *The Handbook of Psychological Testing*, London: Rout ledge.

12- Martlew, Margaret & Connolly, Kevin J. (1996 December). Human Figure Drawings by Schooled and Unschooled Children in Papua New Guinea, *Children Development*, 67(6),p.p.2743-62.

13- Niko, J. Anthony (2001). *Educational Assessment of Students*, 3rd Edition, New Jersey, U.S.A: Prentice-Hall Inc..

14- Rust, J.& Golombok (1995). Modern Psychometrics, A Review of National Norming Studies and Ethnic and Socioeconomic Variation within Unites States, *Journal of Education Measurement*, 26 (1): p.p.1-16

15- Suen, H.K. (1990). *Principles of Test Theories*, New Jersey, U.S.A.: Laurence Erlbaum Associates Publishers.

16- Wright, B. & Linacre, J. M. (1985). *Microscale Manual*, Version 2, Mediax Interactive Technology.

17- Wright, B. and Mead, R. Kidmap (1980). Analysis

and Reporting Individual Performance, *Research Memorandum*, No. 29, Chicago, U.S.A: MESA Press.

18- Wright, B. D. & Linacre, J. M. (1998). *A User's Guide to Winsteps/ Bigsteps*, Version 2.88, ,Chicago, U.S.A.: MESA Press.