



كلية التربية  
قسم علم النفس التربوي

## تدرّيج و دمج اختباري تحصيل في مادة العلوم (مقالى - موضوعى ) باستخدام نموذج راش

### إعداد

أ.د/ مسعد ربيع عبد الله أبو العلا  
أستاذ علم النفس التربوي بكلية التربية  
جامعة بنها

أ.د/ محمد أحمد إبراهيم غنيم  
أستاذ علم النفس التربوي و رئيس مجلس قسم علم  
النفس التربوي بكلية التربية جامعة بنها

إيمان عبد الفتاح السيد  
مدرس مساعد بقسم علم النفس التربوي

سبتمبر ٢٠١٧

## مقدمة الدراسة:

نالت الاختبارات التحصيلية فى عملية بنائها و تطويرها اهتماماً كبيراً من التربويين و ذلك لما يمكن أن تقدمه من بيانات موضوعية يمكن الاعتماد عليها فى صنع القرارات السليمة التى تهدف الى تحسين عملية التعليم ، إلا أنها يشوبها قصور فى إعدادها ، فالكثير من أسئلة الاختبارات المدرسية يتصف بالغموض و الاستخدام العشوائى لأنماط الاسئلة الذى يرجع الى قناعات لدى معلمين بأفضلية نمط من الأسئلة على حساب الأنماط الأخرى ، كما تعانى من نقص فى تمثيل كافة مستويات الأهداف و أيضاً من عدم قدرة بعضها على التمييز بين الطلاب .

و على ذلك فإن جودة وسائل القياس و التقويم تساعد فى إتخاذ قرارات موضوعية بناءً على أسس علمية بإعتبار التقويم جزءاً عضوياً من نسيج النظام التعليمى ، فمن خلاله يتم الوقوف على مدى تحقيق أهداف النظام التعليمى ، و يقدم تغذية راجعة مستمرة تسهم فى تعديل و تطوير النظام و تُزيد من كفاءته و من ثم نوعية و جودة مخرجاته ، و قد يرجع ضعف النظام التعليمى إلى ضعف أدوات ووسائل التقويم المستخدمة ، و عليه فإن إصلاح التقويم و التعرف على المشكلات المرتبطة بأدواته يُعد مدخلاً جيداً لإصلاح التعليم و الارتقاء بمستوى جودة مخرجاته ، و يرتبط بعملية تقويم التحصيل الدراسى العديد من المشكلات و الصعوبات و التى تختلف باختلاف فلسفة القياس و الهدف منه سواء مشكلات ترتبط بأدوات التقويم - الاختبارات التحصيلية - أو مشكلات ترتبط بالمعايير التى يستند إليها فى تفسير درجة التلميذ التى حصل عليها فى اختبار ما حسب النظرية المستخدمة فى القياس (قياس كلاسي - قياس موضوعى) (محمد غنيم، ٢٠٠٣: ١٠).

و يؤكد عبد الفتاح القرشى ( ١٩٨٦ : ٣-٤ ) على ما سبق حيث يذكر أن تقويم الطلاب يعتمد على تحليل البيانات التى يتم الحصول عليها عن طريق الاختبارات و أدوات القياس المختلفة ، كما و يعتمد على كافة البيانات الأخرى التى تساعد على معرفة التغيرات التى تطرأ على نمو الطالب فى جميع جوانب نمو شخصيته ، و من هنا تظهر أهمية سلامة البيانات التى تم التقويم فى ضوءها ، و بالتالى التأكد من كفاءة الأدوات التى تستخدم فى جمع هذه البيانات .

و يذكر (على البثيتى ، ٢٠٠٢ : ٤) أن الاختبارات المدرسيه تعانى من مشكلات فنية تعوق الاستفادة من نتائجها فى عمليات التغذية الراجعة بهدف تطوير التعليم عن طريق تشخيص ما قد يعترى سبيل تحقيق الاهداف التربوية من عقبات و ما يواجهها من صعوبات ، و تتركز تلك المشاكل تحديداً فى عدم توافر الجوانب الفنية عند بنائها

و أثناء التطبيق و ما يترتب على ذلك من آثار سلبية على مصداقية النتائج المستمدة من تلك الاختبارات ، فالاختبارات التي لا يتوافر فيها الحدود الدنيا على الأقل من الصدق و الثبات قد تتحول الى غايات فى حد ذاتها بدلاً من كونها و سائل للكشف عن بعض معوقات تحقيق الأهداف التربوية و علاجها .

و يذكر (محمد غنيم ، ٢٠٠٣ ، ٩٠ - ١٠) أن من أبرز المشاكل التي ترتبط بالاختبارات التحصيلية هو ما يتعلق بما تقيسه أسئلة الاختبارات و نوعيتها و التي منها : أن أسئلة الاختبارات التحصيلية لا تعطى صورة حقيقية عما تقيسه من أهداف تعليمية ، بمعنى عدم صدق الاختبارات فى قياس الأهداف التعليمية كما تعكسها أهداف المحتويات الدراسية ، كما أن كثير من الاختبارات لا تراعى الوزن النسبى و الأهمية النسبية لكل من موضوعات المحتوي والأهداف التعليمية ، و تجدر الإشارة الى أن التذبذب فى نسب تمثيل الأهداف المعرفية بمستوياتها فى الاختبارات ربما يرجع لعدم الالتزام من قبل واضعى الاختبارات باستراتيجية واضحة مبنية على أهداف محددة تأخذ بعين الاعتبار مشكلات التقويم السابقة لتلافيها بهدف تطوير نظم الامتحانات و هذه أشكالية الاختبارات التقليدية التي هى أسرى النظرية الكلاسيكية .

و قد يرجع السبب فيما سبق الى إسناد مهمة إعداد الاختبارات الى معلمين ينقصهم الخبرة و التدريب الكافى ليتمكنوا من إعداد اختبارات سليمة .

كذلك قصور بعض الأساليب الاحصائية الخاصة ببناء الاختبار المستخدم مثل حساب معاملات الصعوبة و التمييز و الصدق و الثبات حيث تعتمد هذه الخصائص على عينة الافراد التي تطبق عليها الاختبار ، و على خصائص المفردات التي تشتمل عليها الاختبار أى خصائص الاختبار المستخدم & (Jandal,2001,53:54 Hambleton, 1991,2:3).

و يؤكد ذلك (صلاح علام ، ٢٠٠٦ ، ٦٨٠-٦٨١ ) حيث يشير إلى أن من المشكلات المتعلقة بالاختبارات التي يتم بناؤها فى إطار النظرية الكلاسيكية هى ، أن الخصائص السيكومترية للاختبارات تعتمد على خصائص عينة الافراد التي يجرى عليها الاختبار ، و على خصائص عينة المفردات التي يشتمل عليها الاختبار مما يؤدي الى نظام قياس نفسى و تربوى غير مستقر ، فتغير معاملات الصدق و الثبات و صعوبة المفردات بتغير قدرات الأفراد ، كما يُفترض أن الاختبار يقيس متغير أحادى البعد و ربما لا يكون ذلك صحيحاً فى كثير من الأحيان ، يفترض أن الدرجات التي يحصل عليها فرد على مفردات الاختبار يمكن جمعها كما لو كانت تمثل ميزاناً خطياً ، و الحقيقة أن هذا الميزان منحنيًا ، فالفرق الثابت بين درجتين من درجات الاختبار يختلف معناه بالنسبة



تم بناء عدد من الاختبارات التحصيلية و بنوك الأسئلة بالمركز القومي للامتحانات و التقويم التربوى و حيث أن نموذج راش يعتمد على فرضيات قوية يتحقق من خلالها أسس القياس الموضوعي و أن أداة القياس التي أُنتشئت باستخدام نموذج راش تُحقق أحادية البعد و استقلالية القياس بصرف النظر عن مجموعة الأفراد الذين أُجرى عليهم الاختبار أو صعوبات البنود التي يتضمنها الاختبار، مما يثبت فاعلية النموذج في تدريج بنود الاختبارات بحيث يتم وضع قدرات الأفراد وصعوبات البنود على ميزان تدريج واحد ، كما تشير إلى أن ارتفاع معامل الصدق و الثبات للاختبار بعد حذف البنود غير الملائمة وفقا لنموذج راش يوفر ذلك إحصاء لصدق و ثبات كل بند وكل فرد و أيضا قيمة الخطأ المعياري للقياس لتقدير صعوبة كل من بند وقدرة كل فرد بعكس الاختبارات الكلاسيكية التي توفر قيمة وحيدة للخطأ المعياري للاختبار الكلى .

و يشير ريكاس (Reckase,2009:275) أنه من المهم تطوير عدد كبير من مفردات الاختبار باستخدام بارامترات المفردة و ذلك فى نفس نظام الأحداثيات و للحصول على بارامترات المفردة فى نفس النظام المنهجي يتطلب ذلك ما يسمى بالربط linking وهناك العديد من الدراسات التي استخدمت نماذج نظرية الاستجابة للمفردة فى أعداد اختبارات تحصيلية تجمع مفرداتها بين الموضوعى و المقالى و منها:

**دراسة (Lancaster,etal.(1987)**، و التى هدفت الى المقارنة بين القدرة التمييزية لمفردات اختبار تحصيلي من نوع الاختيار من متعدد و مفردات اختبار تحصيلي ذات الاجابة القصيرة فى امتحانات منتصف و آخر العام بجامعة طيبة ، و تم التطبيق على عينة بلغت (١٧٧) طالباً خلال ثلاث سنوات وهى مدة الدراسة ، و تم تحليل المفردات باستخدام نماذج الاستجابة للمفردة ( الاحادى - الثنائى - الثلاثى) للحصول على تقدير لصعوبة المفردة و تقدير لتمييز المفردة و تم حساب ثبات كل اختبار باستخدام معادلة "كيودر ريتشاردسون" ، و توصلت نتائج البحث الى أن معامل ثبات الاختبار التحصيلي ذو الاختيار من متعدد يساوى (0.59) ، و معامل ثبات الاختبار التحصيلي ذو الاجابة القصيرة يساوى (0.86) ، و الأختبار ذو الاجابة القصيرة أقل صعوبة من الاختيار من متعدد و أفضل تمييزاً من الاختبار ذو الاختيار من متعدد و أن نسب كل من الصعوبة و التمييز لم تختلف كثيراً من حيث مصدر و نوع المفردة ، كما لا توجد فروق

دالة إحصائية بين مفردات الاختيار من متعدد ومفردات الاجابة القصيرة فى تقدير قدرات الأفراد .

#### أوجه الاستفادة:

١- معامل ثبات الاختبار التحصيلي ذو الاجابة القصيرة اكبر من الاختبار التحصيلي ذو الاختيار من متعدد .

٢- الأختبار ذو الاجابة القصيرة أقل صعوبة من الاختيار من متعدد و أفضل تمييزاً من الاختبار ذو الاختيار من متعدد .

٣- أن نسب كل من الصعوبة و التمييز لم تختلف كثيراً من حيث نوع المفردة .

٤- كما لا توجد فروق دالة إحصائية بين مفردات الاختيار من متعدد ومفردات الاجابة القصيرة فى تقدير قدرات الأفراد .

دراسة (Ercikan , et al.1996) ؛ التى هدفة إلى تدريج اختبارين فى (القراءة - اللغات - الرياضيات - العلوم ) علي عينة من ٨٠٠ طالب ، أحدهما اختبار أختيار من متعدد و تم تدريجة باستخدام النموذج اللوغاريتمى ثلاثى المعلم و الآخر يحتوي علي مفردات الاستجابة التكوينية و تم تدريجة باستخدام نموذج التقدير الجزئى ثنائى المعلم باستخدام برامج الحاسوب PARDUX و توصلت نتائج البحث عند مقارنة نتائج تدريج كل من الاختبارين كلا علي حدى أن قيمة ثبات الاختبار المقالي و قيم معلوماتة test information أقل منه فى الاختبار الموضوعى و ذلك نتيجة لارتفاع متوسط صعوبه الاختبار المقالي و انخفاض معاملات الارتباط فيه.

#### أوجه الاستفادة

١- قيمة ثبات الاختبار المقالي و قيم معلوماتة information test أقل منه في الاختبار الموضوعي .

٢- الاسلوب الاحصائي المتبع في تدريج المفردات : اختبار أختيار من متعدد و تم تدريجة باستخدام النموذج اللوغاريتمي ثلاثى المعلم و الآخر يحتوي علي مفردات الاستجابة التكوينية و تم تدريجة باستخدام نموذج التقدير الجزئي ثنائى المعلم باستخدام برامج الحاسوب PARDUX .

٣- حجم العينة المستخدمة (٨٠٠) طالب .

**دراسة (2000) Bastari** ؛ التى هدفة إلي بحث أثر طول الاختبار فى مادة اللغة الانجليزية فى مراحل تعليمية متعددة و حجم العينة و نمط المفردة - ثنائية الدرجة "صواب و خطأ" و متعددة الدرجة " اسئلة الاجابة القصيرة" - علي تقدير قدرة الفرد باستخدام النموذج اللوغاريتمي ثلاثى المعلم لمفردات الاختيار من متعدد و مفردات الاستجابة المتدرجه - الاستجابة التكوينية- ، و لدراسة أثر طول الاختبار أستخدمت الدراسة اختبارين الاول يتكون من (٣٠) مفردة و الثانى يتكون من (٥٠) مفردة و لمعرفة أثر حجم العينة استخدمت ثلاث حجوم مختلفة ( ٤٠٠ ، ١٠٠٠ ، ٣٠٠٠ ) ، و لدراسة أثر نوع المفردة التى يتضمنها الاختبار استخدمت مستويين من الاختبارات بحيث يتضمن الاختبار الاول على نسبة (٨٠%) مفردة من نوع الاختيار من متعدد و (٢٠%) مفردة من نوع متعددة الاستجابة ، أما الاختبار الثانى يتضمن على نسبة (٩٠%) مفردة من نوع الاختيار من متعدد و (١٠%) مفردة من نوع متعددة الاستجابة ، و قد استخدم برنامج MULTILOG لتقدير معالم المفردات ، و أسفرت النتائج عن أنه كلما زاد طول الاختبار و حجم العينة كلما زادت دقة معايير المفردة ، كما يوجد فروق ذات دلالة احصائية بين مفردات الاختيار من متعدد و مفردات الاستجابة التكوينية قى تقدير معلم (الصعوبة - التمييز) لصالح

مفردات الاستجابة التكوينية ، بينما لا توجد فروق بين مفردات الاختيار من متعدد و مفردات الاستجابة التكوينية فى تقدير قدرات الافراد باستخدام النموذج ثلاثى المعلم .

#### أوجه الاستفادة :

- ١- أشارت الدراسة إلى أنه بزيادة طول الاختبار و حجم العينة يزداد دقة بارامترات المفردة المكونه للاختبار .
- ٢- لا توجد فروق فى تقدير قدرات الافراد باختلاف نمط المفردة المكونه للاختبار مفردات الاختيار من متعدد و مفردات الاستجابة التكوينية .
- ٣- استخدمت الدراسة عدده اطوال للاختبار هي ( ٣٠ ، ٥٠ ) مفردة.
- ٤- استخدمت الدراسة عدة احجام للعينة هي ( ٤٠٠ ، ١٠٠٠ ، ٣٠٠٠ ) ممتحن .
- ٥- تم استخدام النموذج اللوغاريتمى ثلاثى المعلم لتحليل مفردات الاختيار من متعدد و مفردات الاستجابة المتدرجه "الاستجابة التكوينية" .
- ٦- تم استخدام برنامج MULTILOG لتقدير معالم المفردات .

دراسة **Famularo (2007)** و التى هدف إلى المقارنة بين مفردات الاختيار من متعدد ذات العبارة الاساسية و مفردات الاستجابة التكوينية باستخدام النموذج ثنائى المعلم و بحث إلى أى مدى يؤثر نوع المفردة على صعوبة مفردات اختبار الاختيار من متعدد ذات العبارة الأساسية و صعوبة مفردات الاستجابة التكوينية و كذلك بحث الى أى مدى يؤثر تقدير الدرجات على صعوبة مفردات الاختيار من متعدد و الاستجابة التكوينية ، و طبق الاختباران على عينة مكونة من (١٠٠٠) طالب من المدارس الإعدادية ، و أشارت نتائج البحث الى أن مفردات الاختيار من متعدد تعطي تقديرات لصعوبة المفردة أقل من مفردات الاستجابة التكوينية ، أى أن أسلوب تقديرات

الدرجات يساهم فى الاختلافات بين صعوبة مفردة الاختيار من متعدد و مفردة الاستجابة التكوينية ، و أنه توجد فروق ذات دلالة إحصائية فى تقدير الاخطاء المعيارية لقدرات الافراد لصالح مفردات الاختيار من متعدد عند مستوى (0.05) .

#### أوجه الاستفادة

١- مفردات الاختيار من متعدد تعطي تقديرات لصعوبة المفردة أقل من مفردات الاستجابة التكوينية .

٢- أسلوب تقديرات الدرجات يساهم فى الاختلافات بين صعوبة مفردة الاختيار من متعدد و مفردة الاستجابة التكوينية .

٣- أنه توجد فروق ذات دلالة إحصائية فى تقدير الاخطاء المعيارية لقدرات الافراد لصالح مفردات الاختيار من متعدد عند مستوى (0.05) .

٤- العينة مكونة من (١٠٠٠) ممتحن .

دراسة زياد عبد الحسيب الخولى (٢٠١٣) ، و التى هدفت الى استخدام نموذج التقدير الجزئى متعدد التدريج فى تدريج اختبار يحتوى على مفردات ثنائية و متعددة التدريج باستخدام برنامج PARSCAL و ذلك من خلال تدريج اختبار افتراضى مكون من مفردات ثنائية و متعددة الاستجابة على مقياس مشترك بنقطة أصل مشتركة و وحدة قياس مشتركة ، قام الباحث بطريقة عشوائية بتوليد مصفوفة رقمية مكونة من ١٢٢٥ سطراً يمثل عدد الأفراد و ٥٣ عموداً يمثل عدد مفردات هذا الاختبار ثم باستخدام برنامج SPSS تم تخزين ملف البيانات المكون من المصفوفة السابقة بعد إضافة خمسة أعمدة منها أربعة خاصة بتعريف كل حالة و عمود يفصل بين أعمدة المفردات و أعمدة تعريف الحالة ، على أن تأخذ عناصر المصفوفة الخاصة بأول (٢٢) مفردة آحدى القيمتين (٠،١) بافتراض أنها مفردات ثنائية الاستجابة و أن (١٤) التاليين تأخذ آحدى القيم (٠،١،٢) بافتراض أنها

مفردات متعددة الاستجابة تتطلب خطوتين و (١٢) مفردة التاليين لما سبق تأخذ القيم (٠،١،٢،٣) بافتراض أن هذه المفردات متعددة الاستجابة و تتطلب ثلاث خطوات للحل و أن الخمس مفردات الاخيرة تأخذ إحدى القيم (٠،١،٢،٣،٤) بافتراض أن هذه المفردات متعددة الاستجابة و يتطلب حلها أربع خطوات ، ثم قام الباحث بالتأكد من تحقق فرضيات النموذج المتمثلة فى : أحادية البعد و ذلك بالاعتماد على التحليل العاملى و التحقق من وجود عامل عام و احد مسيطر فى أداء الأفراد على الاختبار و ذلك قيم الجذور الكامنه باستخدام طريقة المكونات الرئيسية و هو ما أظهرته النتائج ، و بالنسبة للتحقق من الاستقلال الموضعى تم التحقق منها نظرياً حيث أكد قابلية فصل و كفاية معالم نموذج التقدير الجزئى و بالتالى إمكانية تقدير معالم الصعوبة الفئوية بشكل مستقل عن قدرات أفراد عينة التدرج فضلاً عن إمكانية تقدير قدرة أفراد العينة بشكل مستقل عن معالم الصعوبة الفئوية ، أيضاً التحقق من فرضية تساوى قوة تمييز جميع المفردات و ذلك يمكن التأكد منه إجرائياً باستخدام برنامج التحليل المستخدم فى تدرج المفردات من خلال الرسوم البيانية .

#### أوجه الاستفادة

١- أكدت عملياً على إمكانية تدرج اختبار يتضمن أنماط مفردات متعددة من حيث طريقة التصحيح ( ثنائية الدرجة - متعددة الدرجة بمستويات تدرج مختلفة ) ، و هى تقيس قدرة أحادية البعد و تتوافق مع نموذج راش .

٢- أسخدمت نموذج التقدير الجزئى للتحقق من إمكانية تقدير معالم الصعوبة الفئوية بشكل مستقل عن قدرات أفراد عينة التدرج فضلاً عن إمكانية تقدير قدرة أفراد العينة بشكل مستقل عن معالم الصعوبة الفئوية، و هو النموذج التى ستعتمده الدراسة الحالية فى تحليل المفردات متعددة التدرج أداء البحث .

٣- تم استخدام نموذج التقدير الجزئي متعدد التدرج فى تدرج اختبار يحتوى على مفردات ثنائية و استخدام برنامج PARSCAL فى تدرج اختبار يحتوى على مفردات متعددة التدرج .

٤- حجم العينة كانت ١٢٢٥ ممتحن .

٥- عدد مفردات الاختبار ٥٣ مفردة منهم (٢٢) مفردة تأخذ إحدى القيمتين (٠،١) بافتراض أنها مفردات ثنائية الاستجابة و (١٤) التاليين تأخذ إحدى القيم (٠،١،٢) بافتراض أنها مفردات متعددة الاستجابة تتطلب خطوتين و (١٢) مفردة التاليين لما سبق تأخذ القيم (٠،١،٢،٣) بافتراض أن هذه المفردات متعددة الاستجابة و تتطلب ثلاث خطوات للحل و أن (٥) مفردات الاخيرة تأخذ إحدى القيم (٠،١،٢،٣،٤).

دراسة محمد دبوس (٢٠١٦) ، و التى هدفت إلى استخدام نظرية الاستجابة للفقرة فى بناء تجمع فقرات اختبار محكى المرجع فى الرياضيات وفق النموذج اللوجستى ثنائى المعلم لمعرفة قيم تقديرات معالم فقرات الاختبار ( الصعوبة و التمييز ) و القدرة للأفراد لكل فقرة من فقرات الثنائية التدرج باستخدام برنامج MULTILOG.7 و فقرات متعددة التدرج باستخدام نفس البرنامج معتمداً فيه على نموذج الاستجابة المتدرجة Grand Response Model (GRM) ، و تكونت أداة الدراسة من (٥٠) مفردة ثنائية التدرج لكل فقرة أربع بدائل و (١٠) مفردات متعددة التدرج من نوع الاسئلة محددة الاجابة بحيث كل فقرة مكونه من عدة خطوات متدرجة يعطى علامات جزئية على هذه الخطوات و روعى فى بناء الفقرات أن تقيس تحصيل الطلاب فى الرياضيات فى المعرفة المفاهيمية و المعرفة الإجرائية و حل المشكلات ، و تكونت عينة الدراسة من (٥٠٢) طالب و طالبة من طلاب الصف الثامن الاساسي أجابوا عن جميع الفقرات ، تم معالجة البيانات إحصائياً حسب نظرية الاستجابة للمفردة باستخدام البرمجيات الاجصائية

Excel, Minitab.15, Sps.15, Multilog.7 فى تحليل البيانات للعيننة الرئيسية و حساب تقدير معلم الصعوبة و الخطأ المعياري فى تقدير قيم الصعوبة للفقرات ثنائية التدرج و متعددة التدرج بواسطة نموذج الاستجابة المتدرجة ، توصلت النتائج إلى:

١. وجود تباين فى صعوبة الفقرات و هو ما نحتاجه فى بناء الاختبارات  
محكية المرجع •

٢. أن تقديرات معلم قدرة الأفراد على المفردات ثنائية التدرج تراوحت بين (-١.٥٧٥ : ٢.٨١٨) لوجيت و بمتوسط حسابي (٠.٢٤٨) لوجيت ، و تقديرات معلم صعوبة المفردات ثنائية التدرج تراوحت بين (-١.١١ : ١.٤١) لوجيت بمتوسط قدرة (-٠.٢٢٤) ، و هنا يلاحظ أن متوسط تقديرات القدرة للأفراد تزيد على متوسط تقديرات صعوبة الفقرات الثنائية التدرج مما يمكن القول أن الفقرات الثنائية التدرج بوجه عام كانت فى متناول قدرات أفراد العينة •

٣. أن تقديرات معلم قدرة الأفراد على المفردات متعددة التدرج تراوحت بين (-١.٧٥٧ : ٢.٢٩٦) لوجيت و بمتوسط حسابي (٠.٠٧١) لوجيت ، و تقديرات معلم صعوبة المفردات متعددة التدرج تراوحت بين (-٠.٠٩ : ٢.٦٢) لوجيت بمتوسط قدرة (٠.٩٥١) ، و هنا يلاحظ أن تقديرات متوسط صعوبة الفقرات تزيد على متوسط تقديرات القدرة للأفراد متعددة التدرج مما يمكن القول أن الفقرات متعددة التدرج بوجه عام كانت صعبة إلى حد ما •

أوجه الاستفادة

١- متوسط تقديرات القدرة للأفراد تزيد على متوسط تقديرات صعوبة الفقرات الثنائية التدرج مما يمكن القول أن الفقرات الثنائية التدرج بوجه عام كانت في متناول قدرات أفراد العينة .

٢- تقديرات متوسط صعوبة الفقرات تزيد على متوسط تقديرات القدرة للأفراد متعددة التدرج مما يمكن القول أن الفقرات متعددة التدرج بوجه عام كانت صعبة إلى حد ما .

٣- تم استخدام النموذج اللوجستي ثنائى المعلم لمعرفة قيم تقديرات معالم فقرات الاختبار ( الصعوبة و التمييز ) و القدرة للأفراد لكل فقرة من فقرات الثنائية التدرج باستخدام برنامج MULTILOG.7 و فقرات متعددة التدرج باستخدام نفس البرنامج معتمداً فيه على نموذج الاستجابة المترجة Grand (GRM) Response Model .

٤- تكونت أداة الدراسة من (٥٠) مفردة ثنائية التدرج لكل فقرة أربع بدائل و (١٠) مفردات متعددة التدرج من نوع الاسئلة محددة الاجابة بحيث كل فقرة مكونه من عدة خطوات مترجة يعطى علامات جزئية على هذه الخطوات .

٥- تكونت عينة الدراسة من (٥٠٢) طالب و طالبة .  
دراسة زياد أحمد العبد لله (٢٠٠٩) ، و التى تهدف إلى دراسة أثر حجم العينة على تدرج مفردات بنك الأسئلة باستخدام نموذج التقدير الجزئى ، و قد اعتمدت الدراسة على التصميم التجريبي و ذلك من خلال التعيين العشوائى لمجموعات العينه ، و تمثلت أداة الدراسة فى اختبار تحصيلي محكى المرجع فى وحدة " الحدود و المقادير الجبرية" من كتاب الجبر للصف الاول الاعدادى وتتوعت فيه المفردات بين اختيار من متعدد و الخطأ و الصواب و مفردات متعددة الاستجابة و التى تراوحت فئات استجابتها بين (٢) و (٥) فئات ، وقد تكون الاختبار من (٦٩) مفردة بحيث توزعت فيها

أنواع المفردات كما يلي ( ٢١ مفردة اختيار من متعدد - ١٤ مفردة صواب و خطأ - ١٣ مفردة متعددة الاستجابة ذات ثلاث فئات - ٦ مفردات متعددة الاستجابة ذات اربع فئات - ١٥ مفردة متعددة الاستجابة ذات خمس فئات )، كما بلغت عينة الدراسة (١١٣٦) تلميذ و تلميذة ، أوضحت النتائج الخاصة بالتساؤل عن أى حجم للعينة (١٠٠٠ - ٨٠٠ - ٤٠٠) الذى ينتج عنه أفضل تدريج ممكن لمفردات بنك الاسئلة أنه أفضل النسب هى المقابلة لحجم العينة (١٠٠٠) حيث كانت نسبة عدم التلائم أقل مايمكن و كذلك نسبة الفروق الجوهرية بالمقارنة مع النسب المقابلة للحجوم الأخرى .

### أوجه الاستفادة

١- تمت عملية التقدير لمعالم صعوبة المفردات باستخدام نموذج التقدير الجزئي عبر برنامج الحاسوب ConQuest و ذلك باستعمال كود ٧ من التعليمات الخاصة بالبرنامج .

٢- بلغت عينة الدراسة (١١٣٦) تلميذ و تلميذة .

٣- أحجام العينات المستخدمة هى (١٠٠٠ - ٨٠٠ - ٤٠٠) ممتحن .

٤- الحجم الأمثل للعينة و الذى ينتج عنه أفضل تدريج ممكن لمفردات بنك الاسئلة هو أكبر حجم .

٥- وقد تكون الاختبار من (٦٩) مفردة بحيث توزعت فيها أنواع المفردات كما يلي ( ٢١ مفردة اختيار من متعدد - ١٤ مفردة صواب و خطأ - ١٣ مفردة متعددة الاستجابة ذات ثلاث فئات - ٦ مفردات متعددة الاستجابة ذات اربع فئات - ١٥ مفردة متعددة الاستجابة ذات خمس فئات) .

و يتضح من نتائج الدراسات السابقة أنه :

- أختلفت الأبحاث و الدراسات على أفضلية النتائج الاحصائية الناتجة عن كلاً من الاختبار التحصيلي المقالي ( ذو الاجابة القصيرة ) و الاختبار التحصيلي الموضوعي ( ذو الاختيار من متعدد - التكميل ) و ذلك فى النتائج الخاصة بـ(معامل الثبات - صعوبة المفردات - التمييز - دالة معلومات معلومات الاختبار ) و ذلك فى نتائج دراسة (Lancaster , et. all,1987) فى مقابل نتائج كلاً من دراسة (Ercikan,et.all,1996) و دراسة محمد محمود البشير العلى (٢٠٠٤) و دراسة (Famularo ,2007) .
- بينما أوضحت دراسة (Lancaster , et. all,1987) و دراسة (Ercikan,et.all,1996) عدم وجود فروق دالة احصائياً بين نمطى المفردات فى تقدير قدرات الافراد بينما أظهرت دراسة (Famularo,2007) وجود فروق ذات دلالة احصائية فى تقدير الاخطاء المعيارية لقدرات الافراد لصالح مفردات الاختيار من متعدد - ثنائى الدرجة - .
- أشارت دراسة (Bastari,2000) ( أنه بزيادة طول الاختبار و حجم العينة يزداد دقة بارامترات المفردات المكونة للاختبار .
- تنوعت النماذج الاحصائية المستخدمة فى تدرج المفردات ثنائية الدرجة بين النموذج اللوغاريتمى الثلاثى المعلم كما فى دراسة (Ercickan,et.all,1996) و دراسة (Bastari,2000) ( و النموذج اللوغاريتمى ثنائى المعلم كما فى دراسة محمد دبوس (٢٠١٦) .
- تنوعت البرامج الاحصائية المستخدمة فى تدرج مفردات الاختبار متعدد الدرجة بين استخدام نموذج التقدير الجزئى كما فى دراسة (Ercikan,et.all,1996) و استخدام نموذج الاستجابة المترجحة كما

فى دراسة (محمد دبوس ،٢٠١٦) و نموذج سلم التقدير كما فى دراسة  
( ابتسام توفيق ،٢٠٠٩) .

• تتوعت البرامج الاحصائية المستخدمة وكانت ( MULTILOG.7,PARSCAL,  
WINSTEPS, BIGSTEPS,RUMM2020, MULTILOG , PARDUX).

و بذلك يمكن تحديد مشكلة الدراسة فى التساؤل التالي:

١- هل تختلف تقديرات معلم صعوبة المفردة و معلم قدرة الفرد و  
دالة معلومات الاختبار باختلاف نمط مفردة الاختبار(ثنائية  
الدرجة - متعددة الدرجة) التى يتضمنها الاختبار التحصيلي ؟  
٢- هل يمكن وضع اختباري التحصيل (موضوعي ، المقالي )  
فى صورة اختبارية مدمجة على تدرج واحد مشترك له صفر  
واحد؟

٣- هل يمكن اشتقاق صورة اختبارية مختصرة من الصورة  
الاختبارية المدمجة وتكون متعادلة معها ؟

أهداف الدراسة:

أهمية الدراسة:

مصطلحات الدراسة:

نموذج راش:

يذكر ( Rasch,1961 ) أن نموذج راش من أهم نماذج نظرية  
السمات الكامنة و يتميز بسهولة استخدامه و موضوعيته كما أن معالمه  
قابله للفصل أي فصل قدرة المفحوص عن صعوبة المفردة فى تحليل بيانات  
الاختبار و يترتب على هذا الفصل إمكانية مقارنة استجابة مفحوصين مختلفين  
على مدى معين من الجانب المُقاس (فقرة معينة أو مجموعة فقرات ) بغض  
النظر عن مستوى صعوبتها ، وكذلك إمكانية مقارنة صعوبة فقرتين مختلفتين  
أجاب عليهما مفحوص واحد وهذا يشير إلى استقلالية أداة القياس عن  
الأشياء المُقاسة واستقلالية قياس الأشياء عن أداة القياس (فى : زينبات دعنا  
، ٢٠٠٥ : ٤٣ ) .

يقوم نموذج راش على افتراض أن جميع مفردات الاختبار متساوية من حيث التمييز ولكنها تختلف عن بعضها في الصعوبة فقط بالإضافة إلي انعدام التخمين (صلاح علام، ٢٠٠٥ : ٧٠ ؛ Hambleton&swaminthan,1985,46-47).

#### تدرج المفردات:

يُقصد به انتظام مواقع المفردات على متصل صعوبتها حول صفر التدرج بحيث يُعبر موقع كل مفردة " بعدها عن صفر التدرج" عن مستوى صعوبتها (El-korashy, 1995 :756).

#### تقديرات الدرجات:

و يقصد بها نمط الدرجة التي يتم التصحيح على أساسها و هي إما ثنائية (١-٠) كما في المفردات الموضوعية أو متعددة (٢-١-٠) كما في المفردات المقالية.

#### نوع المفردة Item type :

##### • المفردة ثنائية الاستجابة Dichotomous item response

و يقصد بها كل مفردة تكون إجابتها إما صحيحة أو خطأ أي أنها تقيس مستويين من مستويات المتغير المراد قياسه و هما مستوى الاتقان التام و يمثل الاجابة الصحيحة و مستوى عدم الاتقان و يمثل الاجابة الخاطئة.

##### • المفردة متعددة الاستجابة Polytomous item response

و يقصد بها كل مفردة تكون الاجابة عليها مكونه من خطوتين على الاقل أي أنها تقيس أكثر من مستوى واحد من مستويات المتغير المراد قياسه ، بحيث تكون درجة الفرد بالنسبة للاجابة عن كل مفردة تساوي عدد الخطوات التي أجاب عليها بشكل صحيح و مستمر دون انقطاع ، أي اعتباراً من الخطوة الصحيحة الأولى إلى أول خطأ يرتكبه . (زياد العبد لله ، ٢٠٠٩ ، ٩).

#### (الدمج) دمج مفردات اختبارين في تدرج واحد

هي عملية تحويل التدرج المستقل لكل من الاختبارين أداتى الدراسة الحالية إلى تدرج واحد مشترك و يتطلب هذا التحويل القيام بعملية موازنة أو معادلة لتدرج البنود المكونة لكل من الاختبارين ، و تتم هذه العملية بأسلوبين يقوم

أولهما على استخدام بعض البنود المشتركة بين الاختبارين ، و يقوم  
الأسلوب الثانى على استخدام بعض الأفراد المشتركين فى أداء كل من  
الاختبارين (أمينة كاظم و آخرون ، ١٩٩٦ : ٢٦٧) .

### الاطار النظرى للدراسة

من منطلق ما تسعى إليه هذه الدراسة و هو التركيز على

#### فروض الدراسة:

نظراً لتباين نتائج الدراسات و البحوث السابقة فى هذا المجال ، فقد صاغت  
الباحثة فروض البحث على النحو التالى :

خطة و إجراءات الدراسة:

### أولاً: عينة الدراسة:

تم اشتقاق عينة الدراسة الحالية من بين تلاميذ الصف الثانى  
الأعدادى بمحافظة القليوبية فى العام الدراسى ٢٠١٦ / ٢٠١٧ م .

#### ١- عينة الدراسة الاستطلاعية:

تكونت عينة الدراسة الاستطلاعية من (100) من تلاميذ الصف  
الثانى الأعدادى بمتوسط عمرى (١٣.٣) سنة و انحراف معيارى (٠.٣٩)  
سنه بمدارس (أسامه بن زيد - بلال بن رباح - محمد متولى شعراوى ) بإدارة  
(بنها التعليمية) حيث تم تطبيق أدوات الدراسة الحالية فى النصف الاول من  
العام الدراسى بفواصل زمنى بين تطبيق الاختبارين من اسبوع و ذلك للوقوف  
على :

١-مدى وضوح تعليمات الأداة.

٢-مدى فهم و استيعاب أفراد العينة الاستطلاعية لمفردات الاختبار.

٤-الوقوف على الصعوبات التى تواجه الباحثة فى التطبيق المبدئى و العمل  
على إزالتها قبل التطبيق النهائى.

٥-الوقوف على الخصائص السيكومترية لأداة الدراسة (ثبات / صدق ) .

#### ٢- عينة الدراسة الأساسية:

تكونت عينة الدراسة الأساسية الأولية من ( ٥٦٩ ) تلميذ و تلميذة أختيروا عشوائياً لتطبيق عليهم الاختبارين (المقالي - الموضوعي) فى الوحدة الأولى بمقرر مادة العلوم للصف الثانى الاعدادى حيث حضر (٤٧٦) تلميذ و تلميذة جلستى الامتحانين ( المقالى - الموضوعي) ، بينما (٣٥) تلميذ و تلميذة حضروا جلسة الاختبار الموضوعي فقط لذا تم استبعادهم من العينة الأساسية و (٥٨) تلميذ و تلميذة حضروا جلسة الاختبار المقالي فقط لذا تم استبعادهم من العينة الاساسية نظرا لتغيبهم يوم التطبيق وذلك من مدارس (ناصر الأعدادية الحديثه بنين - ٢٥ يناير الأعدادية بنات - محمد متولى شعراوى الأعدادية بنين - بلال بن رباح الأعدادية بنين) بإدارة بنها التعليمية بفواصل زمنى بين تطبيق الاختبارين أسبوع ، متوسط عمرهم زمنى (١٣.٣) سنة و انحراف معيارى ( ٠.٣٤ ) سنة وهى التى استخدمت فى إجراء التحليل و التدريج و الإجابة على تساؤلات الدراسة و جدول (١) يوضح توزيع أفراد العينة .

#### جدول (١)

توزيع حجم العينة النهائى للتلاميذ الذين أنجزوا جلستى التطبيق

المجموع	إناث	ذكور	الجنس المدرسة
١٠٢	-	١٠٢	ناصر الأعدادية بنين
١١٨	١١٨	-	٢٥ يناير
١٠٠	-	١٠٠	محمد متولى شعراوى
١٥٦	٩٧	٥٩	بلال بن رباح
٤٧٦	٢١٥	٢٦١	المجموع

#### ثانياً: أدوات الدراسة

تم إعداد اختبارين محكى المرجع فى مادة العلوم للفصل الدراسى الأول للمرحلة الثانية من التعليم الأعدادى وفق الإجراءات الآتية التى استندت



الاختبار فى ضوء نموذج راش لذا تعين على الباحثة اعداد اختبار تحصيلي مرجعي المحك فى شكلين ( ثنائى الاستجابة - متدرج الاستجابة) فى مقرر الفصل الدراسى الأول من المرحلة الثانية من التعليم الأساسى فى مادة العلوم ، وذلك بإتباع خطوات منهجيه ، و يوضح (صلاح مراد ؛ أمين على سليمان ، ٢٠٠٥ : ١٤٢) أن نظرية السمات (الخصائص الكامنة) تقتضى إمكانية التنبؤ بالأداء التحصيلي للطالب أو تفسير أدائه على الاختبار التحصيلي فى ضوء خصائص مميزة لهذا الأداء تسمى السمات ، و تحاول نظرية السمات الكامنة تقدير درجات الأفراد فى هذه السمات من خلال أداء الأفراد على مجموعة من مفردات الاختبار مع ملاحظة أنه لا تختلف خطوات بناء الاختبار التحصيلي المرجع إلى المحك عن المرجع إلى المعيار إلا فى التحديد الدقيق لمستويات الأداء (محك الأتقان) و ذلك فى الخطوة الأولى الخاصة بتحديد الأهداف التعليمية .

اتبعت الباحثة الاجراءات التى أشار إليها(صلاح علام ، ١٩٨٧ : ٣٦-٣٧) عن كيفية بناء اختبار تحصيلي باستخدام أبسط نماذج السمات الكامنه و هو نموذج راش و فيما يلى خطوات إعداد الاختبار التحصيلي الموضوعي:

١. تحديد الاهداف العامة للوحدة موضوع الاختبار .
٢. تحديد الأهداف السلوكية.
٣. إعداد جدول المواصفات الناتج عن تحديد الموضوعات و أهميتها النسبية من خلال تحليل محتوى الوحدة موضوع الاختبار" فى ضوء بنيه العلم" و الاهمية النسبية للاهداف السلوكية .
٤. عرض تحليل المحتوى و جدول المواصفات علي الخبراء من المعلمين للتأكد من صحة تحليل المحتوى و مدى انتماء كل معلومة أشتمل عليها المحتوى لجانب التعلم الذى نسبت إليه ، كذلك التأكد من الاوزان النسبية ومدى تحقيق النسب المثلي لكل مستوى معرفي فى كل جانب من جوانب بنية العلم للوحدة الدراسية .

٥. عرض الاهداف السلوكية و ما يرتبط بها من اسئلة تقيس تلك الاهداف فى المستويات المعرفية المختلفة و فى جوانب بنية العلم للوحدة موضوع الاهتمام على الخبراء من و أساتذة مناهج و طرق تدريس العلوم بكلية التربية ، حيث تم التأكد من ملاءمة السؤال المصاغ لمستوى الهدف المراد قياسه و تعديل عدد من الاسئلة لا تقيس مستوى الهدف الخاص بها ومن ثم الوصول الى أسئلة الاختبار بحيث تقيس الاهداف فى المستويات المعرفية المراد قياسها .

٦. تجريب المفردات استطلاعياً للاطمئنان على جودة الصياغة ووضوح التعليمات و تحديد الزمن المناسب للاجابة .

٧. تقدير الخصائص السيكومترية لمفردات الاختبار .

أ- معاملات السهولة و الصعوبة لمفردات الاختبار .

ب- معاملات التمييز لمفردات الأختبار .

ت- معاملات الاتساق الداخلى للمفردات

ث- حساب صدق الاختبار .

ج- حساب ثبات الاختبار .

## ( أ ) :الاختبار التحصيلى الموضوعى

### وصف الاختبار التحصيلى فى صورته الاولى:

قامت الباحثة بطبع الاختبار الموضوعى والذي يتكون من (١٠٠) مفردة من نوع الاختيار من متعدد للتطبيق المبدئى بهدف التأكد من صلاحيته للتطبيق النهائى و تحديد زمن الاجابه الملائم .

### حساب صدق الاختبار الموضوعى :

الخصائص السيكومترية لحساب صدق الاختبار:

١. صدق المحكمين:

تم تقدير صدق المحكمين للوقوف على صدق الاختبار؛ وذلك بعرض الاختبار على مجموعة من السادة المحكمين<sup>1</sup> لأخذ آرائهم من حيث:

١. كفاية التعليمات المقدمة للتلاميذ للإجابة بطريقة صحيحة على الإختبار.

✓ أسفر ذلك عن وضوح التعليمات بالاتفاق مع آراء المحكمين .

٢. صلاحية المفردات علمياً، ولغوياً.

✓ أسفر ذلك عن إضافة للسؤال الاول ليكون صحيح علمياً ولذلك تم تعديله من

الصيغة (١-) تحتوى نواة الذرة على ..... أ- بروتونات و الكترونات ، ب-

الكترونات سالبة ، ج- نيترونات متعادلة ، د- بروتونات موجبة) ليصبح (١-)

تحتوى نواة الذرة على ..... أ- بروتونات و نيترونات ، ب- الكترونات سالبة

فقط ، ج- نيترونات متعادلة فقط ، د- بروتونات موجبة فقط) .

✓ وتم تعديل السؤال الثانى من الصيغة ( تدور الألكترونات حول نواة ذرة العنصر

فى مستويات طاقة عددها..... أ- ٢ ، ب- ٣ ، ج-٧ ، د-٥)

ليصبح (تدور الألكترونات حول نواة ذرة الصوديوم فى مستويات طاقة

عددها..... أ- اثنان ، ب- ثلاثة ، ج-سبعة ، د-خمسة) .

✓ كما تم تعديل السؤال الثامن من (أثناء التفاعل الكيميائى تفقد ذرة البوتاسيوم

الكترون التكافؤ لأنها ..... ألكترون . أ- أيون سالب ، ب-ذرة متعادلة ،

ج- عنصر شبه فلز ، د- أيون موجب ) إلى (أثناء التفاعل الكيميائى تفقد

ذرة البوتاسيوم الكترون التكافؤ متحوله إلى ..... أ- أيون سالب ، ب-ذرة

متعادلة ، ج- عنصر شبه فلز ، د- أيون موجب ) .

✓ وتعديل السؤال الثانى عشر من ( يستخدم لحفظ قرنية العين ٠٠٠٠٠٠٠٠-أ-

النيتروجين المسال ، ب- الهيدروجين ، ج-الماء ) إلى ( يقوم الأطباء بحفظ

<sup>1</sup> اسماء المحكمين فى ملحق (١) جدول (٣٤) .

قرنية العين فى مواد مبرده مثل ..... أ- النيتروجين المسال، ب- الهيدروجين ، ج-الماء ) .

✓ و أيضاً السؤال الحادى و العشرين من ( يتجمد الماء عند درجة..... درجة مئوية. أ- ٤ ، ب- ١٠٠ ، ج- ٩٠ ، د- صفر ) إلى (أعلى قيمة لكثافة الماء عند درجة ..... درجة مئوية. أ- ٤ ، ب- ١٠٠ ، ، ج- ٩٠ ، د- صفر ) .

٣. مناسبة المفردات لتلاميذ الصف الثانى الإعدادى.

✓ وفقاً لأراء المحكمين تم تعديل السؤال السادس والاربعين من (العدد الذرى لعنصر النحاس ٢٩ لذا فان عدد الالكترونات التى يحتويها مستوى طاقته الاخير هو ..... أ- ٢ ، ب- ٨ ، ج- ١ ، د- ١١) إلى (العدد الذرى لعنصر الصوديوم ١١ لذا فان عدد الالكترونات التى يحتويها مستوى طاقته الاخير هو ..... أ- ٢ ، ب- ٨ ، ج- ١ ، د- ١١) حيث أقتصر المقرر على توزيع الالكترونات فى مستويات الطاقة بحد أقصى ٢٠ إلكترون .

٤. مناسبة كل سؤال للمستوى الذى وضع لقياسه.

✓ تم تعديل السؤال الحادى و الثلاثين من الصيغة (يذوب ملح الطعام فى الماء لأن الماء ٠٠٠٠٠ أ- مذيب تساهمى ، ب- مذيب قطبى ضعيف ، ج- مذيب قطبى جيد د- مذيب أيونى) إلى (يذوب  $NaCl_2$  فى  $H_2O$  لأن الماء .... أ- مذيب تساهمى ، ب- مذيب قطبى ضعيف ، ج- مذيب قطبى جيد د- مذيب أيونى) ليتناسب مع مستوى التطبيق .

✓ كما تم تعديل السؤال السادس و الثلاثين من (يوصل الماء النقى للتيار الكهربى عند إضافة قطرات من ٠٠٠٠٠٠ أ- سكر الطعام ، ب- كلوريد الهيدروجين ، ج- البنزين ، د- حمض الكبرتيك ) إلى (عند تحليل  $H_2O$  كهربياً نضيف ..... إلى  $H_2O$  النقى أ- سكر الطعام ، ب- كلوريد الهيدروجين ، ج- البنزين ، د- حمض الكبرتيك ) ليتلاءم مع مستوى التطبيق .

## ٢. الصدق التمييزي:

للتحقق من القدرة التمييزية للاختبار؛ تم حساب الصدق التمييزي على عينة الدراسة الاستطلاعية البالغ عددها (١٠٠) تلميذ و تلميذة ؛ بحساب معامل التمييز لكل مفردة من مفردات الاختبار بعد ترتيب درجات التلاميذ تنازلياً حسب درجاتهم في الاختبار، وفصل ٢٧% من درجات أفراد العينة التي تقع في الجزء الأعلى (الإرباعى الأعلى)، وفصل ٢٧% من درجات أفراد العينة التي تقع في الجزء الأسفل (الإرباعى الأدنى) ثم استخدام معادلة جونسون لحساب معامل التمييز، وهى:  $\text{معامل التمييز} = \frac{\text{س-ص}}{\text{ن}}$  حيث أن: (س-ص) هى الفرق بين عدد الذين أجابوا على المفردة إجابة صحيحة فى كل من المجموعتين العليا والدنيا، (ن) هى عدد أفراد أى من المجموعتين (صلاح علام، ٢٠٠٠: ٢٨٤ - ٢٨٧)، (أمين محمد، رجاء أبو علام، ٢٠٠٩: ٣١٧).

و يعتبر معامل التمييز للمفردة مقبول إذا زاد عن (٠.٢)، ولذلك فإن الاختبار التحصيلي له القدرة على التمييز بين أفراد العينة. وبعد ذلك تم إعادة ترتيب مفردات الاختبار وفقاً لمعاملات الصعوبة لكل مفردة و جدول (٨) يوضح قيم معاملات التمييز لمفردات الاختبار التحصيلي الموضوعي.

### جدول ( ٨ )

معاملات التمييز لمفردات الاختبار التحصيلي لمجموعتى الطلاب العليا و الدنيا

$\frac{D-U}{27}$	D	U	ID	$\frac{D-U}{27}$	D	U	ID
0.78	٢	٢٣	51	0.85	٤	٢٧	1
0.81	٤	٢٦	52	0.74	٤	٢٤	2
0.85	٢	٢٥	53	0.48	١٣	٢٦	3
0.70	٣	٢٢	54	0.52	١٢	٢٦	4
0.56	٦	٢١	55	0.56	٩	٢٤	5
0.67	٧	٢٥	56	0.63	١٠	٢٧	6
0.70	٤	٢٣	57	0.63	١٠	٢٧	7

0.78	๐	๒๖	58	0.67	๓	๒๑	8
0.70	๔	๒๓	59	0.52	๘	๒๒	9
0.93	๒	๒๗	60	0.63	๘	๒๐	10
0.63	๖	๒๓	61	0.56	๘	๒๓	11
0.70	๑	๒๐	62	0.67	๔	๒๒	12
0.78	๖	๒๗	63	0.44	๑๔	๒๖	13
0.89	๓	๒๗	64	0.78	๒	๒๓	14
0.56	๑๑	๒๖	65	0.89	๒	๒๖	15
0.63	๖	๒๓	66	0.78	๓	๒๔	16
0.89	๓	๒๗	67	0.70	๐	๒๔	17
0.48	๑๓	๒๖	68	0.52	๑๓	๒๗	18
0.85	๔	๒๗	69	0.85	๓	๒๖	19
0.70	๐	๒๔	70	0.81	๓	๒๐	20
0.85	๓	๒๖	71	0.81	๑	๒๓	21
0.30	๑๐	๒๓	72	0.70	๘	๒๗	22
0.41	๑๔	๒๐	73	0.37	๑๓	๒๓	23
0.33	๑๔	๒๓	74	0.44	๑๐	๒๗	24
0.85	๓	๒๖	75	0.67	๗	๒๐	25
0.78	๑	๒๒	76	0.41	๑๖	๒๗	26
0.81	๐	๒๗	77	0.44	๑๔	๒๖	27
0.74	๐	๒๐	78	0.78	๓	๒๔	28
0.74	๔	๒๔	79	0.85	๓	๒๖	29
0.67	๔	๒๒	80	0.81	๑	๒๓	30
0.41	๑๓	๒๔	81	0.33	๑๖	๒๐	31
0.89	๒	๒๖	82	0.41	๑๑	๒๒	32
0.89	๓	๒๗	83	0.44	๑๓	๒๐	33
0.67	๒	๒๐	84	0.74	๐	๒๐	34
0.85	๒	๒๐	85	0.52	๑๐	๒๔	35
0.74	๒	๒๒	86	0.30	๑๒	๒๐	36

0.63	٥	٢٢	87	0.56	٥	٢٠	37
0.93	١	٢٦	88	0.37	١٣	٢٣	38
0.96	٠	٢٦	89	0.70	٣	٢٢	39
0.78	٥	٢٦	90	0.85	٠	٢٣	40
0.67	٢	٢٠	91	0.30	١٦	٢٤	41
0.70	٢	٢١	92	0.74	٥	٢٥	42
0.70	٤	٢٣	93	0.37	١٣	٢٣	43
0.81	٤	٢٦	94	0.85	٢	٢٥	44
0.78	٥	٢٦	95	0.70	٣	٢٢	45
0.37	١٤	٢٤	96	0.52	٧	٢١	46
0.78	٠	٢١	97	0.85	٢	٢٥	47
0.70	١	٢٠	98	0.63	٥	٢٢	48
0.78	٥	٢٦	99	0.81	٢	٢٤	49
0.56	٧	٢٢	100	0.59	٥	٢١	50

حيث أن Q تشير إلى رقم السؤال ، D تشير إلى عدد التلاميذ - من مجموعة التلاميذ منخفضي التحصيل- الذين اجابوا اجابه صحيحة على المفردة ، U تشير إلى عدد الطلاب-من مجموعة التلاميذ مرتفعي التحصيل- الذين اجابوا اجابه صحيحة على المفردة . كما يشير  $\frac{D-U}{27}$  إلى قيمة معامل التمييز و يوضح جدول(٨) قيم معاملات التمييز و التي تراوحت بين ( ٠.٣٠ للمفردة رقم ٣٦ إلى ٠.٩٦ للمفردة رقم ٨٩) و ذلك يشير إلى مستوى جيد لقدرة مفردات الاختبار علي التمييز ، حيث يذكر (على القرني ، ١٩٩٩ ، ٥٧) أنه أي مفردة (سؤال) ذات معامل تمييز بين ٢٠ إلى ٣٩% تعتبر ذات تمييز مقبول و أي مفردة ذات تمييز أعلى من ٣٩% تعتبر مفردة جيدة التمييز .

### ٣. الاتساق الداخلي :

تم حساب الاتساق الداخلي للاختبار باستخدام برنامج SPSS على عينة الدراسة الاستطلاعية (١٠٠) تلميذ و تلميذة من خلال حساب قيمة معاملات الارتباط بين درجة كل مفردة والدرجة الكلية للاختبار بعد حذف درجة المفردة من درجة الاختبار ككل و جدول (٩) يوضح ذلك .

### جدول (٩)

### معاملات الارتباط بين درجة المفردة والدرجة الكلية للاختبار التحصيلي

السؤال	معامل الارتباط						
1	0.701**	26	0.424**	51	0.594**	76	0.682**
2	0.586**	27	0.349**	52	0.649**	77	0.648**
3	0.343**	28	0.621**	53	0.666**	78	0.596**
4	0.456**	29	0.667**	54	0.570**	79	0.614**
5	0.472**	30	0.666**	55	0.464**	80	0.518**
6	0.477**	31	0.341**	56	0.553**	81	0.362**
7	0.528**	32	0.303**	57	0.537**	82	0.676**
8	0.455**	33	0.404**	58	0.580**	83	0.672**
9	0.380**	34	0.560**	59	0.512**	84	0.549**
10	0.405**	35	0.444**	60	0.721**	85	0.707**
11	0.304**	36	0.204*	61	0.516**	86	0.568**
12	0.509**	37	0.455**	62	0.563**	87	0.508**
13	0.420**	38	0.293**	63	0.636**	88	0.730**
14	0.551**	39	0.528**	64	0.698**	89	0.726**
15	0.632**	40	0.757**	65	0.466**	90	0.598**
16	0.591**	41	0.295**	66	0.527**	91	0.548**
17	0.515**	42	0.557**	67	0.717**	92	0.584**
18	0.389**	43	0.293**	68	0.472**	93	0.572**
19	0.645**	44	0.672**	69	0.703**	94	0.658**
20	0.588**	45	0.531**	70	0.585**	95	0.614**
21	0.641**	46	0.404**	71	0.652**	96	0.324**
22	0.586**	47	0.596**	72	0.288**	97	0.607**
23	0.297**	48	0.465**	73	0.333**	98	0.547**
24	0.378**	49	0.595**	74	0.294**	99	0.627**
25	0.495**	50	0.485**	75	0.655**	100	0.492**

يتضح من جدول (٩) أن قيم معاملات الارتباط جميعها دالة إحصائياً عند مستوى (٠.٠١)

مما يؤكد علي الاتساق الداخلي للاختبار التحصيلي الموضوعي .

#### ٤. معاملات السهولة و الصعوبة لمفردات الاختبار التحصيلي الموضوعي .

تم تقدير قيم معامل الصعوبة لكل مفردة من مفردات الاختبار التحصيلي الموضوعي من خلال قسمة عدد الذين أجابوا إجابة صحيحة علي المفردة على العدد الكلي للذين اجابوا على المفردة باستخدام برنامج SPSS و جدول (١٠) يوضح تلك القيم و التي تراوحت بين ٠.٢٤ للمفردة رقم (٤٠) إلى ٠.٨٨ للمفردة رقم (٢٦) و هي قيم مقبولة إحصائياً و تشير لصدق الأسئلة حيث يُعد السؤال (المفردة) مقبولا إذا تراوحت قيمة معامل الصعوبة له بين ( ٠.١٥ - ٠.٨٥ ) ، كون المفردة التي يقل معامل الصعوبة لها عن ٠.١٥ تكون شديدة

السهولة، والمفردة التي يزيد معامل الصعوبة لها عن ٠.٨٥ تكون شديدة الصعوبة ( صبحي أبو جلاله، ١٩٩٩ : ٢٢١) (صلاح الدين علام، ٢٠٠٠ : ٢٦٩).

### جدول (١٠)

#### قيم معاملات الصعوبة لأسئلة الأختبار الموضوعي

السؤال	قيم معاملات الصعوبة						
1	0.77	26	0.88	51	0.42	76	0.28
2	0.56	27	0.64	52	0.42	77	0.50
3	0.75	28	0.69	53	0.56	78	0.43
4	0.76	29	0.70	54	0.45	79	0.35
5	0.68	30	0.30	55	0.45	80	0.51
6	0.71	31	0.84	56	0.59	81	0.63
7	0.76	32	0.57	57	0.57	82	0.46
8	0.58	33	0.70	58	0.47	83	0.52
9	0.61	34	0.63	59	0.50	84	0.32
10	0.65	35	0.63	60	0.57	85	0.46
11	0.63	36	0.59	61	0.44	86	0.45
12	0.62	37	0.59	62	0.35	87	0.43
13	0.80	38	0.55	63	0.55	88	0.45
14	0.57	39	0.45	64	0.59	89	0.45
15	0.63	40	0.24	65	0.52	90	0.51
16	0.65	41	0.69	66	0.47	91	0.37
17	0.61	42	0.58	67	0.41	92	0.34
18	0.59	43	0.60	68	0.68	93	0.37
19	0.64	44	0.41	69	0.46	94	0.40
20	0.65	45	0.49	70	0.49	95	0.48
21	0.40	46	0.53	71	0.46	96	0.51
22	0.74	47	0.53	72	0.58	97	0.32
23	0.68	48	0.45	73	0.59	98	0.30
24	0.68	49	0.52	74	0.62	99	0.54
25	0.58	50	0.40	75	0.48	100	0.45

#### ثبات الاختبار التحصيلي الموضوعي " التعلق على الجدول " .

قامت الباحثة بحساب ثبات الاختبار باستخدام برنامج SPSS حيث كانت قيمة معامل

الثبات بعدة طرق هي

١. باستخدام ألفا كرونباك Cronbach's Alpha هي (٠.٩٧٦) و هي قيمة

مرتفعة يمكن الوثوق بها يتضح من الجدول السابق أن قيمة معامل الثبات

للاختبار هي قيمة مرتفعة، مما يدل على ثبات الإختبار وإمكانية الوثوق في نتائجه.

٢. باستخدام طريقة التجزئة النصفية حيث بلغت قيمة معامل الارتباط بين نصفي الاختبار (٠.٧٣٦) و باستخدام معادلة التصحيح لسبيرمان-Spearman Brown بلغت قيمة معامل الثبات (٠.٨٤٨) و هي قيمة مرتفعة يمكن الوثوق فيها.

### الاتساق الداخلي (ثبات المفردات)

قامت الباحثة بحساب معامل الارتباط بين (درجة كل مفردة و الدرجة الكلية للاختبار) لكل فرد باستخدام برنامج SPSS ويوضح الجدول (١١) قيم معاملات الارتباط<sup>٢</sup>

#### جدول (١١)

معاملات الارتباط بين مفردات الاختبار الموضوعي و الدرجة الكلية للاختبار

#### (ثبات المفردات)

السؤال	معامل الارتباط						
1	0.709**	26	0.435**	51	0.606**	76	0.691**
2	0.598**	27	0.365**	52	0.660**	77	0.659**
3	0.357**	28	0.632**	53	0.676**	78	0.608**
4	0.469**	29	0.676**	54	0.583**	79	0.626**
5	0.441**	30	0.676**	55	0.478**	80	0.532**
6	0.491**	31	0.353**	56	0.566**	81	0.378**
7	0.540**	32	0.320**	57	0.550**	82	0.686**
8	0.470**	33	0.419**	58	0.593**	83	0.682**
9	0.396**	34	0.573**	59	0.526**	84	0.561**
10	0.421**	35	0.459**	60	0.730**	85	0.716**
11	0.419**	36	0.222*	61	0.530**	86	0.581**
12	0.523**	37	0.470**	62	0.575**	87	0.522**
13	0.432**	38	0.311**	63	0.648**	88	0.738**
14	0.564**	39	0.542**	64	0.708**	89	0.735**
15	0.643**	40	0.764**	65	0.481**	90	0.610**
16	0.603**	41	0.311**	66	0.540**	91	0.561**
17	0.528**	42	0.570**	67	0.726**	92	0.596**

<sup>٢</sup> تعبر عن قيم جميعها دالة أحصائياً عن وجود اتساق داخلي بين درجات الطلاب على كل مفردة و المجموع الكلي لدرجات الطلاب على نفس المفردات .

0.585**	93	0.486**	68	0.310**	43	0.405**	18
0.668**	94	0.713**	69	0.683**	44	0.656**	19
0.626**	95	0.598**	70	0.545**	45	0.599**	20
0.341**	96	0.663**	71	0.420**	46	0.652**	21
0.618**	97	0.305**	72	0.608**	47	0.597**	22
0.559**	98	0.350**	73	0.480**	48	0.313**	23
0.639**	99	0.311**	74	0.608**	49	0.394**	24
0.506**	100	0.665**	75	0.499**	50	0.509**	25

يتضح من جدول (١١) أن قيم معاملات الارتباط دالة احصائياً<sup>٣</sup> عند مستوى (٠.٠١)

\*\* ، و عند مستوى (٠.٠٥) \*

### حساب زمن الإختبار:

تم تحديد الزمن اللازم لتطبيق الإختبار التحصيلي، عن طريق حساب المنوال ، فتم حساب منوال الأزمنة التي استغرقتها كل تلميذة من تلميذات العينة الاستطلاعية في الإجابة عن مفردات الإختبار، حيث بلغ عدد من أنهى الإختبار في زمن ٦٠ دقيقة هو (٤٠) ممتحن ، (٦٠) ممتحن في زمن ٩٠ دقيقة ، بناءً على ذلك فإن الزمن اللازم للإجابة عن مفردات الإختبار هو (٩٠) دقيقة .

## ( ب ) : الأختبار المقالى

### وصف الإختبار فى صورته الاولية:

يتكون الإختبار المقالى من (٤٦ مفردة) متعدد الاستجابة تقيس كل هدف من الأهداف السلوكية التى تمت صياغتها للاعداد الإختبار الموضوعى ، و قد تم مراعاة الأسس الفنية فى كتابة المفردات و تشير الباحثة إلى أنه تم استخدام الاهداف السلوكية - التى تم صياغتها للحصول على نسخة الأختبار الموضوعى- من أجل الحصول على نسخة الأختبار المقالى حيث طوعت الباحثة الاهداف بدمجها للحصول على اسئلة مقالية تقيس تمكن الطالب فى معلوماتين الأمر الذى يؤدى إلى الحصول على سؤال تتدرج فيه الاجابة بين الدرجة (٠ أو ١ أو ٢) ومما يحقق تكافؤ الأختبارين فى قياسهما لنفس الاهداف السلوكية لنفس العينة مع أختلاف نمط المفردة المصاغة و هو ما يحتاجه البحث للاجابة عن تساؤلات الدراسة .

### حساب صدق الإختبار المقالى :

<sup>٣</sup> قيمة معامل الارتباط الجدولية بدرجات حرية ٩٨ عند مستوى دلالة ٠.٠١ = ٠.٢٥٤ ، و عند مستوى دلالة ٠.٠٥ = ٠.١٩٥ .

تم حساب صدق الاختبار بالطرق الآتية:

#### ١. طريقة صدق المحكمين (صدق المحتوى):

أستخدم صدق المحكمين للوقوف على صدق الاختبار؛ وذلك بعرض الاختبار على مجموعة من السادة المحكمين لأخذ آرائهم من حيث:

• كفاية التعليمات المقدمة للتلاميذ للإجابة بطريقة صحيحة على الإختبار:

✓ أسفر ذلك عن وضوح التعليمات بالاتفاق مع آراء المحكمين .

• صلاحية المفردات علمياً، ولغوياً:

✓ أسفر ذلك عن تعديل السؤال السادس (ناقش اثنين من الأسس المتبعة في تصنيف

العناصر بالجدول الدوري الحديث) إلى ( مألأساس العلمى لتصنيف العناصر بالجدول

الدورى الحديث) ، و السؤال السابع (ناقش بأسلوبك كيفية تحديد دورة العنصر و رقم

المجموعة ) إلى ( على ما يدل رقم الدورة و رقم المجموعة بالجدول الدوري الحديث) و

السؤال الثانى عشر ( أشرح ما تعرفه عن المركب القطبى موضعاً : التعريف ، أمثله

عليه) إلى ( عرف المركب القطبى مع ذكر أمثله ) ، و السؤال السابع عشر ( أذكر

النتائج المترتبة على تفاعل الماغنسيوم مع كلوريد الهيدروجين ، الاكسجين موضعاً ذلك

فى صورة معادلات كيميائية) إلى ( ما النتائج المترتبة على تفاعل الماغنسيوم مع كل

من : حمض الهيدروكلوريك ، الأكسجين) .

#### ٢. الصدق التمييزى:

للتحقق من القدرة التمييزية للاختبار؛ تم حساب الصدق التمييزى على عينة الدراسة

الاستطلاعية (١٠٠) تلميذ و تلميذة ؛ حيث قامت الباحثة بحساب الارباعى الأدنى و الارباعى

الاعلى للمجموع الكلى لدرجات التلاميذ

على الاختبار بواسطة برنامج SPSS حيث بلغت قيمة الارباعى الادنى (٢٠ درجة) و كان عدد

التلاميذ الذين حصلوا على مجموع درجات أقل من تلك الدرجة هم (٢٥ تلميذ و تلميذة) و هم

التلاميذ منخفضى التحصيل الدراسى ، و بلغت قيمة الارباعى الاعلى (٥٧.٥ درجة) و كان

عدد التلاميذ الذين حصلوا على مجموع درجات أكبر من تلك الدرجة هم (٢٦ تلميذ و تلميذة) و

هم التلاميذ مرتفعي التحصيل الدراسى و فى ضوء تلك القيم قامت الباحثة بحساب قدرة كل

مفردة علي التمييز بين التلاميذ منخفضى التحصيل و مرتفعي التحصيل و الجدول التالي يوضح ذلك.

### جدول (١٢)

معاملات التمييز لمفردات الاختبار المقالى

قيمة معامل التمييز	Q						
0.50	٤٠	0.69	٢٧	0.73	١٤	0.44	١
0.46	٤١	0.58	٢٨	0.80	١٥	0.52	٢
0.58	٤٢	0.84	٢٩	0.80	١٦	0.48	٣
0.50	٤٣	0.62	٣٠	0.76	١٧	0.52	٤
0.38	٤٤	0.88	٣١	0.77	١٨	0.48	٥
0.80	٤٥	0.88	٣٢	0.81	١٩	0.84	٦
0.72	٤٦	0.81	٣٣	0.77	٢٠	0.80	٧
		0.54	٣٤	0.58	٢١	0.69	٨
		0.46	٣٥	0.69	٢٢	0.60	٩
		0.77	٣٦	0.73	٢٣	0.76	١٠
		0.65	٣٧	0.65	٢٤	0.72	١١
		0.65	٣٨	0.50	٢٥	0.80	١٢
		0.65	٣٩	0.69	٢٦	0.72	١٣

ينتضح من جدول (١٢) أن قيم معاملات التمييز تراوحت بين ٠.٣٨ للمفردة رقم (٤٤) إلى ٠.٨٨ للمفردة رقم (٣٢ ، ٣١) و ذلك يشير إلى مستوى جيد لقدرة مفردات الاختبار علي التمييز ، حيث يعتبر معامل التمييز للمفردة مقبول إذا زاد عن (٠.٣) ، (على القرنى ، ١٩٩٩ ، ٥٧)، ولذلك فإن الاختبار التحصيلي له القدرة على التمييز بين التلاميذ مرتفعى و منخفضى التحصيل الدراسى .

### ٣. صدق الاتساق الداخلى:

وتم حساب الاتساق الداخلى للاختبار باستخدام برنامج SPSS من خلال حساب قيمة معامل الاتساق الداخلى وذلك بحساب قيمة معامل الارتباط بين درجة كل مفردة والدرجة الكلية للاختبار بعد حذف درجة المفردة من درجة الاختبار ككل .

جدول (١٣)

معاملات الاتساق الداخلي لمفردات الاختبار المقالي<sup>٤</sup>

قيمة الارتباط	مفردات الاختبار	قيمة الارتباط	مفردات الاختبار	قيمة الارتباط	مفردات الاختبار
**٠.٧٥٢	٣٣	**٠.٥٩٣	١٧	**٠.٤٢٧	١
**٠.٦٦٤	٣٤	**٠.٦١٨	١٨	**٠.٦٢٤	٢
**٠.٤٤٨	٣٥	**٠.٦٠٤	١٩	**٠.٥٣٢	٣
**٠.٦٨٧	٣٦	**٠.٥٨٠	٢٠	**٠.٥٠١	٤
**٠.٥٥١	٣٧	**٠.٦٣٢	٢١	**٠.٥٣١	٥
**٠.٥٩٤	٣٨	**٠.٦٧١	٢٢	**٠.٦٢٠	٦
**٠.٦٠١	٣٩	**٠.٦٤٩	٢٣	**٠.٦١١	٧
**٠.٥٥٨	٤٠	**٠.٥٩١	٢٤	**٠.٥٤٩	٨
**٠.٥٣٦	٤١	**٠.٥٧١	٢٥	**٠.٦٤٧	٩
**٠.٥٥٥	٤٢	**٠.٦٣٧	٢٦	**٠.٦٦٠	١٠
**٠.٤٩٧	٤٣	**٠.٦٢٧	٢٧	**٠.٥٩٤	١١
**٠.٦٤٦	٤٤	**٠.٥٧٨	٢٨	**٠.٦١١	١٢
**٠.٧١٥	٤٥	**٠.٨٠٨	٢٩	**٠.٤٥٢	١٣
**٠.٦١١	٤٦	**٠.٦٥٦	٣٠	**٠.٥٣٣	١٤
-	-	**٠.٨٠٨	٣١	**٠.٦٨٠	١٥
-	-	**٠.٨٠٥	٣٢	**٠.٥٩٥	١٦

يتضح من جدول (١٣) أن قيم معاملات الارتباط جميعها دالة عند مستوى (٠.٠١) مما يحقق صدق الاتساق الداخلي للاختبار المقالي .

<sup>٤</sup> قيمة معامل الارتباط الجدولية بدرجات حرية ٩٨ عند مستوى دلالة ٠.٠١ = ٠.٢٥٤ ، و عند مستوى دلالة ٠.٠٥ = ٠.١٩٥ .

#### ٤ . حساب قيم معاملات السهولة و الصعوبة لمفردات الاختبار المقالي.

تم حساب معاملات الصعوبة لكل مفردة من مفردات الاختبار عن طريق حساب المتوسط الحسابي للإجابة الصحيحة ، يذكر كلا من (على القرني ، ١٩٩٩ ، ٥٥) ، (إيمان الطائي ، ٢٠٠٩ ، ٥) ، (صلاح الدين علام، ٢٠٠٠ : ٢٦٩) أنه بشكل عام يعتمد معامل الصعوبة المطلوب على الغرض من الاختبار ، وفي الاختبارات التحصيلية العادية فإن أفضل معامل صعوبه للسؤال أو الفقرة هو ٥٠% أو (٠.٥) وما حولها.

ويمكن حساب معامل الصعوبة للأسئلة المقالية باستخدام المعادلة التالية :

مجموع الدرجات المحصلة على السؤال

معامل الصعوبة = —

عدد التلاميذ × درجة السؤال

لذا قامت الباحثة بتحديد الدرجة المخصصة لكل سؤال من أسئلة الأختبار المقالي و هي الدرجة (٢) ، ثم قامت بحساب المتوسط الحسابي لدرجات الأفراد لكل سؤال - حاصل قسمة (مجموع الدرجات المحصلة على السؤال ) علي ( حاصل ضرب عدد التلاميذ في الدرجة المخصصه للسؤال ) باستخدام برنامج SPSS ، وبقسمة المتوسط الحسابي لدرجات الأفراد للسؤال على الدرجة القصوى المحددة للسؤال يمكننا الحصول على معامل الصعوبة لذلك السؤال و جدول (١٤) يوضح قيم معاملات الصعوبة لأسئلة الأختبار المقالي و التي أمتدت من ٠.٢٢ إلى ٠.٨٥ ماعدا المفردات ( ٢٥ - ٢٨ - ٣٠ - ٣٤ - ٣٥ - ٤١ ) كانت قيم معاملات الصعوبه لهم علي الترتيب ( ٠.١٦ - ٠.١٥ - ٠.١٩ - ٠.١٧ - ٠.١٥ - ٠.١٨ ) بينما كانت قيم معاملات الصدق التمييزي هي علي الترتيب ( ٠.٥٠ - ٠.٥٨ - ٠.٦٢ - ٠.٤٦ ) كما كانت قيم معامل الاتساق الداخلي الدالة عن صدق الاتساق الداخلي للمفردات المكونه للاختبار هي علي الترتيب ( ٠.٥٧١ - ٠.٥٧٨ - ٠.٦٥٦ - ٠.٦٦٤ - ٠.٤٤٨ - ٠.٥٣٦ ) و كلاهما - قيم معامل التمييز و معامل الاتساق الداخلي - لتلك المفردات قيم داله احصائياً ، و يؤكد ( أنور الشرقاوى وآخرون ، ١٩٩٦ : ٧٣) أن الاختبار يعد صادقاً إذا أمكن باستخدامه وصف أداء الفرد بالنسبة للنطاق السلوكي الذي يقيسه الاختبار و الذي يكون محدداً تحديداً دقيقاً ، لذلك يعد أسلوب صدق المحتوى مؤشراً جيداً لصدق الاختبار .

كما يذكر (صبحي أبو جلاله، ١٩٩٩ : ٢٢١) أن السؤال (المفردة) تعد مقبوله إذا تراوحت قيمة معامل الصعوبة لها بين (٠.١٥ - ٠.٨٥) ، كون المفردة التي يقل معامل الصعوبة لها

عن ٠.١٥ تكون شديدة السهولة، والمفردة التي يزيد معامل الصعوبة لها عن ٠.٨٥ تكون شديدة الصعوبة .

لذا سوف تبقي الباحثة على تلك المفردات لما حققت من معاملات تمييز و معاملات اتساق داخلي مقبولة و لما أكده المحكمين من القائمين على تدريس المحتوى و الاكاديميين في مجال المناهج و طرق التدريس و لما كان هام للبحث الحالي تساوى الوزن المعلوماتي لكل من المعلومات التي يقيسها الاختبار المقالى والمعلومات التي يقيسها الاختبار الموضوعى حيث أسفرت النتائج الاحصائية عن صلاحيتها - يم عرضه فى البعد التالي- و الذى بدوره هام للاجابة على تساؤلات البحث الحالي.

#### جدول (١٤)

قيم معاملات الصعوبة لأسئلة الأختبار المقالى

السؤال	المتوسط الحسابى	معامل الصعوبة	السؤال	المتوسط الحسابى	معامل الصعوبة
١	١.٦٨	٠.٨٤	٢٤	٠.٦٦	٠.٣٣
٢	١.٣٦	٠.٦٨	٢٥	٠.٣٢	٠.١٦
٣	١.٥٣	٠.٧٦٥	٢٦	٠.٤٤	٠.٢٢
٤	١.٤٢	٠.٧١	٢٧	٠.٤٢	٠.٢١
٥	١.٥٣	٠.٧٦٥	٢٨	٠.٣٠	٠.١٥
٦	١.٣٧	٠.٦٨٥	٢٩	٠.٥٨	٠.٢٩
٧	١.٢٠	٠.٦٠	٣٠	٠.٣٧	٠.١٨٥
٨	٠.٩٦	٠.٤٨	٣١	٠.٦٦	٠.٣٣
٩	١.٣١	٠.٦٥٥	٣٢	٠.٦٠	٠.٣٠
١٠	١.٣٠	٠.٦٥	٣٣	٠.٦٠	٠.٣٠
١١	١.١٦	٠.٥٨	٣٤	٠.٣٣	٠.١٦٥
١٢	١.٣٦	٠.٦٨	٣٥	٠.٣٠	٠.١٥
١٣	١.٢٠	٠.٦٠	٣٦	٠.٤١	٠.٢٠٥
١٤	١.١٠	٠.٥٥	٣٧	٠.٥٦	٠.٢٨
١٥	١.١٧	٠.٥٨٥	٣٨	٠.٥٤	٠.٢٧
١٦	٠.٩٠	٠.٤٥	٣٩	٠.٥٢	٠.٢٦
١٧	٠.٩٣	٠.٤٦٥	٤٠	٠.٤٢	٠.٢١

٠.١٨	٠.٣٦	٤١	٠.٣٣	٠.٦٦	١٨
٠.٢١	٠.٤٢	٤٢	٠.٣٨٥	٠.٧٧	١٩
٠.٢١	٠.٤٢	٤٣	٠.٣٤	٠.٦٨	٢٠
٠.٣٧٥	٠.٧٥	٤٤	٠.٢٣	٠.٤٦	٢١
٠.٣٣	٠.٦٦	٤٥	٠.٢٩	٠.٥٨	٢٢
٠.٤٢٥	٠.٨٥	٤٦	٠.٣٠	٠.٦٠	٢٣

### ثبات الاختبار المقلبي

قامت الباحثة بحساب ثبات الاختبار باستخدام برنامج SPSS حيث كانت قيمة معامل الثبات بعدة طرق هي

١. باستخدام ألفا كرونباك هي (٠.٩٦٧) و هي قيمة مرتفعة يمكن الوثوق بها

٢. باستخدام طريقة التجزئة النصفية حيث بلغت قيمة معامل الارتباط بين نصفي الاختبار (٠.٥٥٧) و باستخدام معادلة التصحيح لسبيرمان بلغت قيمة معامل الثبات (٠.٧١٥) و هي قيمة مرتفعة يمكن الوثوق فيها.

٣. باستخدام الاتساق الداخلي لحساب ثبات المفردات حيث قامت الباحثة بحساب معامل الارتباط بين درجة كل سؤال و الدرجة الكلية للاختبار باستخدام حزمة البرنامج الاحصائي ببرنامج SPSS ويوضح الجدول (١٥) يوضح معاملات الارتباط الناتجة و التي تعبر عن قيم معاملات الثبات و جميعها دالة احصائيا عن مستوى دلالة ٠.٠٠١

### جدول (١٥)

معاملات الارتباط بين درجات مفردات الاختبار و الدرجة الكلية للاختبار<sup>٥</sup>  
(الاتساق الداخلي)

مفردات الاختبار	قيمة الارتباط	مفردات الاختبار	قيمة الارتباط	مفردات الاختبار	قيمة الارتباط
١	**٠.٤٥٥	١٧	**٠.٦١٧	٣٣	**٠.٧٦٧

<sup>٥</sup> قيمة معامل الارتباط الجدولية بدرجات حرية ٩٨ عند مستوى دلالة ٠.٠٠١  
قيمة معامل الارتباط الجدولية بدرجات حرية ٩٩ عند مستوى ٠.٠٠١ = ٠.٢٥٤

**٠.٦٧٩	٣٤	**٠.٦٤٠	١٨	**٠.٦٤٥	٢
**٠.٥٥٧	٣٥	**٠.٦٢٨	١٩	**٠.٥٥٨	٣
**٠.٧٠٢	٣٦	**٠.٦٠٤	٢٠	**٠.٥٢٨	٤
**٠.٥٧٥	٣٧	**٠.٦٥٢	٢١	**٠.٥٥٣	٥
**٠.٦١٧	٣٨	**٠.٦٩٠	٢٢	**٠.٦٤٤	٦
**٠.٦٢٣	٣٩	**٠.٦٦٨	٢٣	**٠.٦٣٦	٧
**٠.٥٨٠	٤٠	**٠.٦١٣	٢٤	**٠.٥٧٦	٨
**٠.٥٥٧	٤١	**٠.٥٩٠	٢٥	**٠.٦٦٧	٩
**٠.٥٧٦	٤٢	**٠.٦٥٦	٢٦	**٠.٦٧٨	١٠
**٠.٥٢١	٤٣	**٠.٦٤٦	٢٧	**٠.٦١٨	١١
**٠.٦٦٨	٤٤	**٠.٦٣٨	٢٨	**٠.٦٣٦	١٢
**٠.٧٣٣	٤٥	**٠.٨١٩	٢٩	**٠.٤٨٣	١٣
**٠.٦٣٦	٤٦	**٠.٦٧٣	٣٠	**٠.٥٥٨	١٤
-	-	**٠.٨٢٠	٣١	**٠.٧٠٠	١٥
-	-	**٠.٨١٧	٣٢	**٠.٦١٩	١٦

#### حساب زمن الإختبار:

تم تحديد الزمن اللازم لتطبيق الإختبار التحصيلي، عن طريق حساب المنوال، فتم حساب منوال الأزمنة التي استغرقتها كل تلميذ وتلميذة من تلاميذات العينة الاستطلاعية في الإجابة عن مفردات الإختبار، (Hays, 1980: 729)، حيث بلغ عدد من أنهى الإختبار في زمن ٦٠ دقيقة هو (٨) ممتحن، (٧٠) ممتحن في زمن ٩٠ دقيقة وهو زمن فتره دراسيه بالجدول المدرسي، (٢٢) ممتحن أنهى الإجابة في زمن ١١٠ دقيقة و هو زمن فترة دراسية بالاضافة الى فترة الراحة بالجدول المدرسي، وبناءً على ذلك فإن الزمن اللازم للإجابة عن مفردات الإختبار هو (٩٠) دقيقة.

#### **رابعاً: منهج الدراسة:**

استخدمت الباحثة فى سبيل الإجابة على تساؤلات الدراسة الحالية المنهج الوصفى الاحصائى ، و ذلك باستخدام برنامج Winsteps لمعالجة البيانات .

### خامساً: إجراءات الدراسة:

بعد الانتهاء من تطبيق أداتى الدراسة على عينة التدريج التى بلغت ( ٤٧٦ ) تلميذ و تلميذة ، و التصحيح وفقاً لمفتاح التصحيح المعدة يدوياً لكل من الاختبار التحصيلى الموضوعى و المقالي ، ثم استخدام البيانات لتحليلها كلا على حده ثم معا فى ملف بيانات واحد حيث أن كلا الاختبارين قام نفس التلاميذ بالاستجابة عليهما كاختبار مختلط (موضوعى و مقالي ) وفق نموذج (راش) باستخدام البرنامج الإحصائى "WINSTEPS" بهدف استبعاد الأفراد و المفردات غير الملائمة إحصائياً لأسس القياس الموضوعى و أيضاً تدريج المفردات ، و تقدير قدرات الأفراد المقابلة لكل درجة خام احتمالية و إيجاد قيمة دالة المعلومات لكل اختبار .

و فيما يلى نتائج تحليل و تدريج الاختبار الموضوعى ( عدد مفردات الاختبار =

١٠٠ مفردة ، عدد العينة = ٤٧٦ تلميذ و تلميذة) بالتفصيل:

#### ✓ حذف البيانات التامة والصفيرية :

من خصائص برنامج WINSTEPS القيام ألياً باستبعاد البيانات الخاصة بكل الأفراد الذين حصلوا على درجات نهائية (تامة) ، أو درجة (صفر) على مجموعة المفردات التى أداها الفرد ، كذلك يستبعد البيانات المتعلقة بالمفردات التى أجمع أفراد العينة على إجابتها سواء بالصواب أو الخطأ ، حيث تعتبر خارج نطاق القياس ، و ذلك بعد التحليل الأول ، و لم يسفر التحليل الأول عن حذف أى فرد أو مفردة تبعاً لهذا المحك .

#### ✓ حذف الأفراد غير الملائمين إحصائياً للنموذج:

بعد الخطوة السابقة ، تم التحليل الثانى باستخدام نفس البرنامج ، لحذف الأفراد بعد تحديدهم من التحليل الأول و الذين تجاوزوا محكات الملاءمة الإحصائية التقاربية و التباعدية ( $2\pm$ ) التى يوفرها البرنامج ، و يعنى حذف الأفراد غير الملائمين لأسس القياس الموضوعى، فتجاوز الحد الفاصل ( $2+$ ) يعنى تجاوز الحد المقبول لاختلاف نمط استجابات الفرد عما يتوقع منه ، فيفشل فى حل المفردات الأقل فى صعوبتها عن تقدير قدرته ، و يجيب صواباً على المفردات الأكثر

٩ حيث قامت الباحثة بالتدريب على استخدام برنامج Winsteps :

(١) فى دورة " تحليل و تدريج بنود الاسئلة باستخدام نظرية الاستجابة للمفردة " ورشة تدريبية بكلية البنات جامعة عين شمس ، وحدة التقويم و الامتحانات فى الفترة من ٢٣/٤/٢٠١١م الى ٢٨/٤/٢٠١١م .

(٢) فى دورة " تحليل و تدريج بنود الاسئلة باستخدام نظرية الاستجابة للمفردة و توظيفها فى تطبيق الاختبار التكيفى المحوسب" ورشة تدريبية بكلية البنات جامعة عين شمس ، وحدة التقويم و الامتحانات فى الفترة من ٢٨/٩/٢٠١٥م الى ١/١٠/٢٠١٥م .

فى صعوبتها عن تقدير قدرته، و ذلك بسبب التسرع أو التخمين ، أو عدم الجدية فى الإجابة ، أما تجاوز الحد الفاصل (-2) يعنى أن نمط استجابات الفرد متنسق بدرجة غير واقعية و يرجع ذلك إلى كونه بطيئاً أو شديد الحرص ، و قد يلجأ إلى النقل أو الغش .

وقد أسفرت هذه الخطوة عن حذف (٩٠) تلميذ و تلميذة أكوادهم هى،  
(١٢٨،١٢٥،١٠٣،١٠٠،٩٦،٩٢،٩٠،٨٨،٧٧،٦٧،٦٥،٥٣،٥٠،٤٧،٤٥،٣٧،٣٤،٢٩،٢٢،١٩)  
(١٩٨،١٩٧،١٩٥،١٩٠،١٨٢،١٧٨،١٧٥،١٧١،١٦٨،١٦٥،١٥٥،١٥٤،١٤٨،١٤١،١٣٧،  
(٢٥٧،٢٥٦،٢٥٥،٢٥٣،٢٤٢،٢٣٩،٢٣٧،٢٣٦،٢٢٩،٢٢٨،٢٢٥،٢٢٢،٢١٨،٢٠٥،٢٠٤  
(٣٤٥،٣٣٠،٣١٤،٣٠٣،٣٠٠،٢٩٤،٢٩٣،٢٨٣،٢٨١،٢٧٧،٢٧٣،٢٦٧،٢٦٣،٢٦٢،٢٦١  
(٤١٠،٣٩٨،٣٩٧،٣٩٢،٣٩٠،٣٨٨،٣٧٨،٣٧٥،٣٦٧،٣٦٦،٣٦٢،٣٦٠،٣٥٤،٣٥٢،٣٤٧  
(٤٤١،٤٤٧،٤٥٢،٤٥٣،٤٥٧،٤٥٨،٤٦٢،٤٦٧،٤٧٢،٤٧٥) تلميذ و تلميذة فى ضوء مؤشر  
الملاءمة التقاربية و التباعدية ، و بذلك تكون نسبة التلاميذ المحذوفين (١٨.٩%) من عدد أفراد  
عينة التدرج ، ليكون العدد النهائي للعينة ٣٨٦ تلميذ و تلميذة .

#### ✓ حذف المفردات غير الملائمة للنموذج :

أجرى التحليل الثالث بعد حذف الأفراد غير الملائمين لأسس القياس الموضوعى ، بهدف تحديد و حذف المفردات التى تتجاوز محكات الملاءمة الإحصائية ( $\pm 2.5$ ) و التى يوفرها البرنامج ، و يعنى هذا حذف المفردات غير الملائمة لأسس القياس الموضوعى، فإذا تجاوزت المفردات الحد الفاصل ( $+2.5$ ) فهذا يعنى عيب فى المفردة أو عدم صدقها فى قياس ما تقيسه باقى المفردات، أما إذا تجاوزت المفردة الحد الفاصل ( $-2.5$ ) فهذا يعنى عدم استقلال المفردة عن غيرها من المفردات أى تعتمد الإجابة عليها على الإجابة على غيرها من المفردات ، أو أنها تقيس متغير آخر شديد الارتباط بالمتغير موضوع القياس .

وقد أسفرت هذه الخطوة عن حذف (عدد ٣٠) مفردة أكوادهم  
هى(٢،٣،١١،١٧،١٨،٢١،٣١،٣٨،٤٠،٤٦،٤٩،٥٣،٥٨،٦١،٦٣،٦٤،٦٥،٦٨،٦٩)  
(٧٢،٧٧،٨٣،٨٤،٨٥،٨٦،٨٨،٨٩،٩٠،٩٦،١٠٠) ، ليصبح عدد مفردات الاختبار  
الموضوعى (٧٠) مفردة و جدول (١٦) يوضح ملخص نتائج التحليل الأول و الثانى  
و الثالث .

#### جدول ( ١٦ )

## نتائج خطوات تحليل الاختبار الموضوعى باستخدام برنامج "WINSTEPS"

التحليل الثالث	التحليل الثانى	التحليل الأول	البيانات
386	386	467	عدد الأفراد
70	100	100	عدد المفردات
0.00	0.00	0.00	متوسط تقديرات الصعوبة
1.25	1.15	0.99	الانحراف المعياري لتقديرات الصعوبة
0.99	0.98	0.99	معامل ثبات تقدير الصعوبات
0.88	0.71	0.52	متوسط تقدير القدرات
1.36	1.25	1.22	الانحراف المعياري لتقديرات القدرة
0.93	0.95	0.95	معامل ثبات تقدير القدرات

### حساب صدق و ثبات الاختبار الموضوعى :

#### ١- حساب صدق الاختبار الموضوعى:

أعتمدت الباحثة فى التحقق من صدق القياس فيما يتيح نموذج راش من صدق المفردات فى صدق تعريف المتغير موضوع القياس ، و صدق التدرج ( الملائمة للنموذج ) ، و فيما يلى تفصيل لذلك :

#### ١-١ صدق التدرج (صدق الملاءمة) .

إذا كانت جميع مفردات الاختبار تعبر عن نفس المتغير ، لأنها متسقة فيما بينها ، و تقيس فى أساسها نفس المتغير ، و كان نمط استجابات كل فرد يتسق مع استجابات معظم الأفراد أو مع ما يتوقع منه ، فان هذا يوضح مدى صدق تدرج المفردة و الفرد على المتغير موضوع القياس (Linacre&Wright,1985:4-20) .

و بما أنه قد تم حذف الاستجابات غير الملائمة للقياس - وفق لمحكات الملاءمة - و

ذلك لأسباب مختلفة كالتخمين و الغش أو السرعة فى الاستجابة على الاختبار .

كذلك تم حذف المفردات غير الملائمة للقياس وفق المحكات التي يوفرها نموذج (راش) ، و تبقى فقط الأفراد الملائمون للقياس ، و المفردات الصادقة فى تعريف المتغير ، فإن ذلك يتيح صدق تدرج المفردات فى قياس المتغير موضوع القياس و كذلك صدق تقدير قدرات الأفراد .

#### ٢-١ صدق تعريف المتغير :

يتطلب التحقق من عدم وجود فجوات بين مفردات المقياس ، ومن خلال خريطة المتغير فى الشكل (١٢) فى ملحق الاشكال (٢) تم تحديد المفردات التى تقع بينها فجوات و من ثم تم التحقق من طبيعة الفرق بين تقديرى صعوبة كل مفردتين متتاليتين بينهما مسافة و التى تعنى الفرق فى الصعوبة و الذى يعبر عنه جدول (١٧) وفيه يشير الحكم (صفر) أى عدم وجود فجوة عندما يكون مجموع الخطأ المعيارى أكبر من فرق الصعوبة بين أى مفردتين متتاليتين ، و الحكم (١) أى يوجد فجوة بين المفردتين عندما يكون مجموع الخطأ المعيارى أصغر من الفرق فى الصعوبة ، يتضح من جدول (١٧) أن هذا الفرق يزيد بمقدار ضئيل عن مجموع الخطأ المعيارى لهما ، أى عدم وجود فجوات واضحة على متصل صعوبات المفردات و خاصة فى وسط المتصل ، أى أن مفردات الاختبار منتظمة فى تدرجها على مدى ميزان القياس ، أى أنها تعرّف مفردات الاختبار لجميع مستويات المتغير الذى يمثل هذا المتصل ، مما يشير إلى صدق تعريف المفردات للمتغير موضوع القياس .

#### جدول (١٧)

العلاقة مسافة الفراغ (فجوة) بين كل مفردتين متتاليتين على خريطة المتغير للاختبار الموضوعى بالمنف

الوصف الاحصائى المستخدم لتحديد العلاقة						
المسلسل	المفردتين المتتاليتين	الخطأ المعيارى	تقديرات الصعوبة	فرق الصعوبة	مجموع الخطأ	الحكم
1	30	51.9	78.55	20.85	102.5	0
	92	50.6	57.7			
2	19	50.85	41.75	3.35	101.9	0
	20	51.05	38.4			
3	20	51.05	38.4	7.95	103	0
	1	51.95	30.45			
4	36	50.6	50.9	1.45	101.2	0
	34	50.6	49.45			
5	39	50.65	48.35	1.3	101.3	0
	42	50.65	47.05			
6	26	50.8	43.05	1.15	101.65	0
	7	50.85	41.9			
7	7	50.85	41.9	3.3	101.9	0

			38.6	51.05	22	
0	101.2	1.7	55.15	50.6	94	8
			53.45	50.6	91	
0	101.2	1.6	50.7	50.6	44	9
			49.1	50.6	41	
0	101.5	-1.7	43.3	50.8	6	10
			45	50.7	5	
0	101.5	1.95	45	50.7	5	11
			43.05	50.8	26	
0	101.2	2.5	54.6	50.6	95	12
			52.1	50.6	73	
0	101.2	2.1	51.2	50.6	45	13
			49.1	50.6	43	
0	101.35	4.3	49.1	50.6	43	14
			44.8	50.75	9	
0	101.55	1.6	44.8	50.75	9	15
			43.2	50.8	29	
0	101.2	2.9	55.1	50.6	98	16
			52.2	50.6	74	
0	101.2	1.85	51.1	50.6	66	17
			49.25	50.6	56	
0	101.4	5.95	49.25	50.6	56	18
			43.3	50.8	6	
0	101.2	2.55	52.05	50.6	70	19
			49.5	50.6	57	

## ٢- حساب ثبات الاختبار :

إعتمد التحقق من ثبات الاختبار التحصيلي الموضوعي على ما يلي :

### ٢-١ ما يوفره نموذج راش من استقلالية القياس :

إن إستقلالية القياس و تحرره الذي يوفره نموذج " راش " أتاح الفرصة لثبات القياس ، حيث لا يختلف القياس ( سواء أكان ذلك لقدرة الفرد أم لصعوبة المفردة ) باختلاف عينة التدرج (أمينة كاظم و آخرون ، ١٩٩٤ : ١٣١ ) .

ولقد تحققت استقلالية القياس بتحقق ملاءمة كل من المفردات و الأفراد للنموذج ، و ذلك وفق لمحكات الملاءمة الإحصائية المختلفة ، مما يحقق ثبات قياس كل من الأفراد و المفردات.

### ٢-٢ ما يوفره برنامج Winsteps من تقديرات للخطأ المعياري :

يوفر برنامج Winsteps تقديراً للخطأ المعياري لتقدير القدرة العقلية لدى كل فرد ، و كذلك لكل مفردة من مفردات الاختبار ، و يعتبر هذا بدوره مؤشراً دقيقاً لمدى ثبات المقياس ، و

بالرجوع إلى جدول تدرج الاختبار في صورته النهائية - جدول (١) في ملحق (١) لوحظ أن قيم الخطأ المعياري لتقديرات صعوبة المفردات تعتبر صغيرة نسبياً حيث تراوحت ما بين ( 0.12 ) و ( 0.39 ) لوجيت مما يدل على ارتفاع ثبات مفردات الاختبار في صورته النهائية .

أما قيم الخطأ المعياري لتقديرات قدرة الأفراد فقد تراوحت ما بين ( 0.27 ) و ( 1.23 ) لوجيت مما يحقق ثبات تقدير قدرات الأفراد ، كما يتضح من جدول (١٢) في ملحق (١) .

### ٢-٣ معامل كيودر ريتشارسون 20 :

يقوم برنامج Winsteps بحساب معامل الثبات لكل من قدرة الأفراد و صعوبة المفردات و هو معامل المكافئ لمعامل ثبات كيودر - ريتشارسون 20 حيث كانت قيمته :

- معامل ثبات تقدير قدرة الأفراد = 0.93 .

- معامل ثبات تقدير صعوبة المفردات = 0.99 .

ويدل ارتفاع قيمة معامل ثبات صعوبة المفردات (0.99) على ارتفاع ثبات القياس لتقدير صعوبات مفردات الاختبار في صورته النهائي ، أما قيمة معامل ثبات الأفراد فجاءت (0.93) و هذا يعتبر ثبات تقدير قدرات الأفراد و يؤكد ثبات المقياس في صورته الجديدة .

### ▪ تقدير قدرة الأفراد لكل درجة كلية محتملة على الاختبار الموضوعي لعينة متوسطة الحجم :

استخدمت الباحثة برنامج Winsteps مرة أخرى و ذلك لاستخراج الجدول رقم (٢٣) في ملحق (١) و الذى يوضح العلاقة بين كل درجة كلية محتملة على الاختبار في صورته النهائية و القدرة المقابلة لها باللوجيت ، ثم قامت الباحثة بتحويل الدرجة باللوجيت إلى المنف حتى تتخلص من القيم السالبة ، يتضح من جدول (٢٣) ملحق (١) ما يلي:

✓ أن العلاقة بين الدرجة الكلية الخام المحتملة على الاختبار و القدرة المقابلة لها

هى علاقة طردية موجبة حيث تزداد القدرة كلما ازدادت الدرجة الكلية الخام

المقابلة لها .

✓ يمتد مدى الدرجات المحتملة من (٠) إلى (٧٠) درجة خام ، كما تمتد القدرة

من (-٦.٢٧) إلى (٧.٠١) لوجيت أى من (١٨.٦٥) إلى (٨٥.٠٥) منف و

هو مدى واسع .

✓ يلاحظ أن قيم الخطأ المعياري يقل بدرجة كبيرة كلما ابتعد عن الأطراف و هي نتيجة أن أصعب المفردات و أسهلها عادة تكون أبعد المفردات عن مستوى قدرة أفراد العينة ، مما قد يزيد من أخطاء القياس .

يقدم البرنامج المستخدم تقديراً تقريبياً للقدرة المقابلة لكل من الدرجة الخام (صفر) و الدرجة العظمى (٧٠) و يلاحظ ارتفاع الخطأ المعياري لهاتين القدرتين .

### (ب) الاختبار المقالى التحصيلى

و فيما يلى نتائج تحليل و تدرج الاختبار المقالى ( عدد مفردات الاختبار = ٤٦ مفردة ، عدد العينة = ٤٧٦ تلميذ و تلميذة) بالتفصيل:

✓ حذف البيانات التامة والصفرية :

من خصائص برنامج WINSTEPS القيام آلياً باستبعاد البيانات الخاصة بكل الأفراد الذين حصلوا على درجات نهائية (تامة) ، أو درجة (صفر) على مجموعة المفردات التى أداها الفرد ، كذلك يستبعد البيانات المتعلقة بالمفردات التى أجمع أفراد العينة على إجابتها سواء بالصواب أو الخطأ ، حيث تعتبر خارج نطاق القياس ، و ذلك بعد التحليل الأول ، و لم يسفر التحليل الأول عن حذف أى فرد أو مفردة تبعاً لهذا المحك .

✓ حذف الأفراد غير الملائمين إحصائياً للنموذج:

بعد الخطوة السابقة ، تم التحليل الثانى باستخدام نفس البرنامج ، لحذف الأفراد بعد تحديدهم من التحليل الأول و الذين تجاوزوا محكات الملاءمة الإحصائية التقاربية و التباعدية ( $2\pm$ ) التى يوفرها البرنامج ، و يعنى حذف الأفراد غير الملائمين لأسس القياس الموضوعى، فتجاوز الحد الفاصل ( $2+$ ) يعنى تجاوز الحد المقبول لاختلاف نمط استجابات الفرد عما يتوقع منه ، فيفشل فى حل المفردات الأقل فى صعوبتها عن تقدير قدرته ، و يجيب صواباً على المفردات الأكثر فى صعوبتها عن تقدير قدرته، و ذلك بسبب التسرع أو التخمين ، أو عدم الجدية فى الإجابة ، أما تجاوز الحد الفاصل ( $2-$ ) يعنى أن نمط استجابات الفرد متنسق بدرجة غير واقعية و يرجع ذلك إلى كونه بطيئاً أو شديد الحرص ، و قد يلجأ إلى النقل أو الغش .

وقد أسفرت هذه الخطوة عن حذف (٣٩) تلميذ و تلميذة أكوادهم هى،

(٢١٤،٢١٣،١٩٠،١٨٤،١٧٢،١٤٨،١٤٧،١١٩،٨٣،٨٠،٧٦،٦٧،٥٧،٤٦،٤٠،٣٦،٣٥،١٩)

٣٧٦،٣٦٩،٣٥٨،٣٣٦،٢٩١،٢٨٦،٢٧٥،٢٦٧،٢٦٢،٢٥٩،٢٥٨،٢٥٤،٢٤٦،٢٣١،٢٢٠  
 ٤٠٢،٤٠٦،٤١١،٤١٨،٤٤٣،٤٧٤) فى ضوء مؤشرى الملاءمة التقاربية و التباعدية ، و بذلك  
 تكون نسبة الطلاب المحذوفين (٨.١٩%) من عدد أفراد عينة التدرج .

#### ✓ حذف المفردات غير الملائمة للنموذج :

أجرى التحليل الثالث بعد حذف الأفراد غير الملائمين لأسس القياس الموضوعى ،  
 بهدف تحديد و حذف المفردات التى تتجاوز محكات الملاءمة الإحصائية ( $\pm 2.5$ ) و التى  
 يوفرها البرنامج ، و يعنى هذا حذف المفردات غير الملائمة لأسس القياس الموضوعى، فإذا  
 تجاوزت المفردات الحد الفاصل ( $+2.5$ ) فهذا يعنى عيب فى المفردة أو عدم صدقها فى قياس  
 ما تقيسه باقى المفردات، أما إذا تجاوزت المفردة الحد الفاصل ( $-2.5$ ) فهذا يعنى عدم استقلال  
 المفردة عن غيرها من المفردات أى تعتمد الإجابة عليها على الاجابة على غيرها من المفردات  
 ، أو أنها تقيس متغير آخر شديد الارتباط بالمتغير موضوع القياس .

وقد أسفرت هذه الخطوة عن حذف (عدد ٩) مفردة أكوادهم

هى (٢،٣،٦،٧،٨،١٩،٢٣،٢٧،٣٤)

#### جدول (٢٦)

نتائج خطوات تحليل الاختبار المقالي باستخدام برنامج "WINSTEPS"

التحليل الثالث	التحليل الثانى	التحليل الأول	البيانات
437	437	476	عدد الأفراد
37	46	46	عدد المفردات
0.00	0.00	0.00	متوسط تقديرات الصعوبة
0.89	0.81	0.78	الانحراف المعياري لتقديرات الصعوبة
0.99	0.99	0.99	معامل ثبات تقدير الصعوبات
-0.46	0.43	- 0.37	متوسط تقدير القدرات
1.25	1.18	1.16	الانحراف المعياري لتقديرات القدرة
0.93	0.94	0.94	معامل ثبات تقدير القدرات

حساب صدق و ثبات الاختبار المقالي :

## ١- حساب صدق الاختبار المقالي:

أعتمدت الباحثة فى التحقق من صدق القياس فيما يتيح نموذج راش من صدق المفردات فى صدق تعريف المتغير موضوع القياس ، و صدق التدرج ( الملائمة للنموذج ) ، و فيما يلى تفصيل لذلك :

### ١-١ صدق التدرج (صدق الملاءمة) .

إذا كانت جميع مفردات الاختبار تعبر عن نفس المتغير ، لأنها متنسقة فيما بينها ، و تقيس فى أساسها نفس المتغير ، و كان نمط استجابات كل فرد يتسق مع استجابات معظم الأفراد أو مع ما يتوقع منه ، فان هذا يوضح مدى صدق تدرج المفردة و الفرد على المتغير موضوع القياس (Linacre & Wright,1985:4-20) .

و بما أنه قد تم حذف الاستجابات غير الملائمة للقياس - وفق لمحكات الملاءمة - و ذلك لأسباب مختلفة كالتخمين و الغش أو السرعة فى الاستجابة على الاختبار .

كذلك تم حذف المفردات غير الملائمة للقياس وفق المحكات التى يوفرها نموذج (راش) ، و تبقى فقط الأفراد الملائمون للقياس ، و المفردات الصادقة فى تعريف المتغير ، فإن ذلك يتيح صدق تدرج المفردات فى قياس المتغير موضوع القياس و كذلك صدق تقدير قدرات الأفراد .

### ١-٢ صدق تعريف المتغير :

يتطلب التحقق من عدم وجود فجوات بين مفردات المقياس ، ومن خلال خريطة المتغير فى الشكل (١٧) فى ملحق الاشكال (٢) ، تم تحديد المفردات التى تقع بينها فجوات و من ثم تم التحقق من طبيعة الفرق بين تقديرى صعوبة كل مفردتين متتاليتين بينهما مسافة و التى تعنى الفرق فى الصعوبة و الذى يعبر عنه جدول (٢٧) وفيه يشير الحكم (صفر) أى عدم وجود فجوة عندما يكون مجموع الخطأ المعيارى أكبر من فرق الصعوبة بين أى مفردتين متتاليتين ، و الحكم (١) أى يوجد فجوة بين المفردتين عندما يكون مجموع الخطأ المعيارى أصغر من الفرق فى الصعوبة ، يتضح من جدول (٢٧) أن هذا الفرق يزيد بمقدار ضئيل عن مجموع الخطأ المعيارى لهما ، أى عدم وجود فجوات واضحة على متصل صعوبات المفردات و خاصة فى وسط المتصل ، بذلك فإن مفردات الاختبار منتظمة فى تدرجها على مدى ميزان القياس ، أى أنها

تعرف مفردات الاختبار لجميع مستويات المتغير الذي يمثل هذا المتصل ، مما يشير إلى صدق تعريف المفردات للمتغير موضوع القياس .

### جدول (٢٧)

العلاقة مسافة الفراغ (فجوة) بين كل مفردتين متتاليتين على خريطة المتغير للاختبارالمقالى بالمنف

الوصف الاحصائي المستخدم لتحديد العلاقة						
المسلسل	المفردتين المتتاليتين	الخطأ المعياري	تقديرات الصعوبة	فرق الصعوبة	مجموع الخطأ	الحكم
1	35	50.4	53.95	1.4	100.75	0
	21	50.35	52.55			
2	10	50.35	45.75	1.9	100.75	0
	5	50.4	43.85			
3	12	50.4	43.15	10.85	101.15	0
	1	50.75	32.3			
4	41	50.4	54.4	2.8	100.75	0
	38	50.35	51.6			
5	24	50.35	48.3	5.3	100.75	0
	4	50.4	43			
6	45	50.4	53.95	2.5	100.75	0
	39	50.35	51.45			
7	43	50.4	54.9	3.45	100.75	0
	46	50.35	51.45			
8	36	50.35	50.6	1.65	100.7	0
	33	50.35	48.95			
9	44	50.35	51.2	1.7	100.7	0
	37	50.35	49.5			

### ٢- حساب ثبات الاختبار :

إعتمد التحقق من ثبات الاختبار التحصيلي الموضوعي على ما يلي :

#### ٢-١ ما يوفره نموذج راش من استقلالية القياس :

إن إستقلالية القياس و تحرره الذي يوفره نموذج " راش " أتاح الفرصة لثبات القياس ، حيث لا يختلف القياس ( سواء أكان ذلك لقدرة الفرد أم لصعوبة المفردة ) باختلاف عينة التدرج (أمينة كاظم و آخرون ، ١٩٩٤ : ١٣١ ) .

ولقد تحققت استقلالية القياس بتحقق ملائمة كل من المفردات و الأفراد للنموذج ، و ذلك وفق لمحكات الملاءمة الإحصائية المختلفة ، مما يحقق ثبات قياس كل من الأفراد و المفردات .

#### ٢-٢ ما يوفره برنامج Winsteps من تقديرات للخطأ المعياري :

يوفر برنامج Winsteps تقديراً للخطأ المعياري لتقدير القدرة العقلية لدى كل فرد ، و كذلك لكل مفردة من مفردات الاختبار ، و يعتبر هذا بدوره مؤشراً دقيقاً لمدى ثبات المقياس ، و بالرجوع إلى جدول تدريج الاختبار في صورته النهائية - جدول (٢) في ملحق (١) لوحظ أن قيم الخطأ المعياري لتقديرات صعوبة المفردات تعتبر صغيرة نسبياً حيث تراوحت ما بين ( 0.07 ) و ( 0.15 ) لوجيت مما يدل على ارتفاع ثبات مفردات الاختبار في صورته النهائية .

أما قيم الخطأ المعياري لتقديرات قدرة الأفراد فقد تراوحت ما بين ( 0.21 ) و ( 1.8 ) لوجيت مما يحقق ثبات تقدير قدرات الأفراد ، كما يتضح من جدول (١٣) في ملحق (١) .

#### ٢-٢ معامل كيودر ريتشارسون 20 :

يقوم برنامج Winsteps بحساب معامل الثبات لكل من قدرة الأفراد و صعوبة المفردات و هو معامل المكافئ لمعامل ثبات كيودر - ريتشارسون 20 حيث كانت قيمته :

- معامل ثبات تقدير قدرة الأفراد = 0.93 .

- معامل ثبات تقدير صعوبة المفردات = 0.99 .

ويدل ارتفاع قيمة معامل ثبات صعوبة المفردات (0.99) على ارتفاع ثبات القياس لتقدير صعوبات مفردات الاختبار في صورته النهائية ، أما قيمة معامل ثبات الأفراد فجاءت (0.93) و هذا يعتبر ثبات تقدير قدرات الأفراد و يؤكد ثبات المقياس في صورته الجديدة .

#### ▪ تقدير قدرة الأفراد لكل درجة كلية محتملة على الاختبار :

استخدمت الباحثة برنامج Winsteps مرة أخرى و ذلك لاستخراج الجدول رقم (٢٤) في ملحق (١) و الذى يوضح العلاقة بين كل درجة كلية محتملة على الاختبار في صورته النهائية و القدرة المقابلة لها باللوجيت ، ثم قامت الباحثة بتحويل الدرجة باللوجيت إلى المنف حتى تتخلص من القيم السالبة ، يتضح من جدول (٢٤) ملحق (١) ما يلى:

✓ أن العلاقة بين الدرجة الكلية الخام المحتملة على الاختبار و القدرة المقابلة لها

هى علاقة طردية موجبة حيث تزداد القدرة كلما ازدادت الدرجة الكلية الخام

• المقابلة لها

✓ يمتد مدى الدرجات المحتملة من (١) إلى (٧٤) درجة خام ، كما تمتد القدرة من (-٥.٣ : ٤.٧٥) لوجيت أى (٢٣.٥ : ٧٣.٧٥) منف و هو مدى واسع .

✓ يلاحظ أن قيم الخطأ المعياري يقل بدرجة كبيرة كلما ابتعد عن الأطراف و هي نتيجة أن أصعب المفردات و أسهلها عادة تكون أبعد المفردات عن مستوى قدرة أفراد العينة ، مما قد يزيد من أخطاء القياس .

✓ يقدم البرنامج المستخدم تقديراً تقريبياً للقدرة المقابلة لكل من الدرجة الخام (صفر) و الدرجة العظمى (٧٤) و يلاحظ ارتفاع الخطأ المعياري لهاتين القدرتين .

و فيما يلي نتائج تحليل و تدريج الاختبار المختلط ( عدد مفردات الاختبار = ١٤٦ مفردة ، عدد العينة = ٤٧٦ تلميذ و تلميذة) بالتفصيل:

✓ حذف البيانات التامة والصفرية :

من خصائص برنامج WINSTEPS القيام آلياً باستبعاد البيانات الخاصة بكل الأفراد الذين حصلوا على درجات نهائية (تامة) ، أو درجة (صفر) على مجموعة المفردات التي أداها الفرد ، كذلك يستبعد البيانات المتعلقة بالمفردات التي أجمع أفراد العينة على إجابتها سواء بالصواب أو الخطأ ، حيث تعتبر خارج نطاق القياس ، و ذلك بعد التحليل الأول ، و لم يسفر التحليل الأول عن حذف أى فرد أو مفردة تبعاً لهذا المحك .

✓ حذف الأفراد غير الملائمين إحصائياً للنموذج:

بعد الخطوة السابقة ، تم التحليل الثانى باستخدام نفس البرنامج ، لحذف الأفراد بعد تحديدهم من التحليل الأول و الذين تجاوزوا محكات الملاءمة الإحصائية التقريبية و التباعدية ( $2\pm$ ) التي يوفرها البرنامج ، و يعنى حذف الأفراد غير الملائمين لأسس القياس الموضوعى، فتجاوز الحد الفاصل ( $2+$ ) يعنى تجاوز الحد المقبول لاختلاف نمط استجابات الفرد عما يتوقع منه ، فيفشل فى حل المفردات الأقل فى صعوبتها عن تقدير قدرته ، و يجيب صواباً على المفردات الأكثر فى صعوبتها عن تقدير قدرته، و ذلك بسبب التسرع أو التخمين ، أو عدم الجدية فى الإجابة ،

أما تجاوز الحد الفاصل (-2) يعنى أن نمط استجابات الفرد متسق بدرجة غير واقعية و يرجع ذلك إلى كونه بطيئاً أو شديد الحرص ، و قد يلجأ إلى النقل أو الغش .

وقد أسفرت هذه الخطوة عن حذف (119) تلميذ و تلميذة أكوادهم هي ( ٦٧،٦٦،٦٥،٦٤،٥٤،٥٣،٥٢،٥٠،٤٥،٣٧،٣٤،٢٦،٢٥،٢٤،٢٣،٢٢،٢٠،١٩،١٧،١٤،٧،٦،١٧١،١٦٩،١٦٢،١٥٣،١٤٨،١٢٩،١٢٥،١٢٢،١١٧،١١٦،١٠٠،٩٦،٩٠،٨٧،٨٣،٧٤،٧٠،٢٣٨،٢٣٢،٢٢٩،٢٢٨،٢٢٢،٢١٧،٢١٣،١٩٩،١٩٨،١٩٥،١٩٠،١٨٩،١٨٨،١٨٧،١٧٨،٢٩٤،٢٩٣،٢٨٩،٢٨١،٢٧٩،٢٧٧،٢٧٥،٢٧٣،٢٦٧،٢٦٢،٢٦١،٢٦٠،٢٥٧،٢٥١،٢٥٠،٣٥٨،٣٥٧،٣٥٤،٣٥٢،٣٥١،٣٥٠،٣٤٥،٣٣٦،٣٣٠،٣٢٩،٣٢١،٣٢٠،٣٠٨،٣٠٧،٣٠٢،٤١٦،٤٠٢،٤٠٠،٣٩٨،٣٩٧،٣٩٠،٣٨٨،٣٨٧،٣٨٠،٣٧٥،٣٦٨،٣٦٧،٣٦٦،٣٦٤،٣٦٢،٤٦١،٤٥٧،٤٥١،٤٤٩،٤٤٧،٤٤٤،٤٤٠،٤٣٩،٤٣٧،٤٣٦،٤٣٣،٤٣٢،٤٣٠،٤٢٨،٤٢٦،٤٦٥،٤٦٧،٤٦٨،٤٧٠،٤٧٢) فى ضوء مؤشرى الملاءمة التقاربية و التباعدية ، و بذلك تكون

نسبة التلاميذ المحذوفين (٢٥%) من عدد أفراد عينة التدرج .

#### ✓ حذف المفردات غير الملائمة للنموذج :

أجرى التحليل الثالث بعد حذف الأفراد غير الملائمين لأسس القياس الموضوعى ، بهدف تحديد و حذف المفردات التى تتجاوز محكات الملاءمة الإحصائية ( $\pm 2.5$ ) و التى يوفرها البرنامج ، و يعنى هذا حذف المفردات غير الملائمة لأسس القياس الموضوعى، فإذا تجاوزت المفردات الحد الفاصل ( $+2.5$ ) فهذا يعنى عيب فى المفردة أو عدم صدقها فى قياس ما تقيسه باقى المفردات، أما إذا تجاوزت المفردة الحد الفاصل ( $-2.5$ ) فهذا يعنى عدم استقلال المفردة عن غيرها من المفردات أى تعتمد الإجابة عليها على الاجابة على غيرها من المفردات ، أو أنها تقيس متغير آخر شديد الارتباط بالمتغير موضوع القياس .

وقد أسفرت هذه الخطوة عن حذف (عدد ٢٧) مفردة أكوادهم هي (٢،١١،٣٠،٣٨،٤٠،٤١،٤٥،٥٨،٦٠،٦١،٦٢،٧٧،٨٣،٨٤،٨٦،٨٨،٨٩،١٠٢،

١٠٣،١٠٥،١٠٦،١٠٧،١٠٨،١٠٩،١١٩،١٢٣،١٤٥)

#### جدول ( ٣٦ )

نتائج خطوات تحليل الاختبار المختلط باستخدام برنامج "WINSTEPS"

التحليل الثالث	التحليل الثاني	التحليل الأول	البيان
357	357	476	عدد الأفراد
119	146	146	عدد المفردات
0.00	0.00	0.00	متوسط تقديرات الصعوبة
0.99	0.98	0.91	الانحراف المعياري لتقديرات الصعوبة
0.99	0.99	0.99	معامل ثبات تقدير الصعوبات
0.37	0.27	0.21	متوسط تقدير القدرات
0.97	0.91	0.86	الانحراف المعياري لتقديرات القدرة
0.96	0.97	0.96	معامل ثبات تقدير القدرات

#### حساب صدق و ثبات الاختبار المختلط :

##### ١- حساب صدق الاختبار المختلط:

أعتمدت الباحثة في التحقق من صدق القياس فيما يتيح نموذج راش من صدق المفردات في صدق تعريف المتغير موضوع القياس ، و صدق التدرج ( الملائمة للنموذج ) ، و فيما يلي تفصيل لذلك :

##### ١-١ صدق التدرج (صدق الملاءمة) .

إذا كانت جميع مفردات الاختبار تعبر عن نفس المتغير ، لأنها متنسقة فيما بينها ، و تقيس في أساسها نفس المتغير ، و كان نمط استجابات كل فرد يتسق مع استجابات معظم الأفراد أو مع ما يتوقع منه ، فان هذا يوضح مدى صدق تدرج المفردة و الفرد على المتغير موضوع القياس (Wright & Linacre,1985:4-20) .

و بما أنه قد تم حذف الاستجابات غير الملائمة للقياس - وفق لمحكات الملاءمة - و ذلك لأسباب مختلفة كالتخمين و الغش أو السرعة في الاستجابة على الاختبار .

كذلك تم حذف المفردات غير الملائمة للقياس وفق المحكات التي يوفرها نموذج (راش) ، و تبقى فقط الأفراد الملائمون للقياس ، و المفردات الصادقة في تعريف المتغير ، فإن ذلك

يتيح صدق تدرج المفردات فى قياس المتغير موضوع القياس و كذلك صدق تقدير قدرات الأفراد .

#### ٢-١ صدق تعريف المتغير :

يتطلب التحقق من عدم وجود فجوات بين مفردات المقياس ، ومن خلال خريطة المتغير فى الشكل (٢٢) فى ملحق الاشكال (٢) ، تم تحديد المفردات التى تقع بينها فجوات و من ثم تم التحقق من طبيعة الفرق بين تقديرى صعوبة كل مفردتين متتاليتين بينهما مسافة و التى تعنى الفرق فى الصعوبة و الذى يعبر عنه جدول (٣٧) وفيه يشير الحكم (صفر) أى عدم وجود فجوة عندما يكون مجموع الخطأ المعيارى أكبر من فرق الصعوبة بين أى مفردتين متتاليتين ، و الحكم (١) أى يوجد فجوة بين المفردتين عندما يكون مجموع الخطأ المعيارى أصغر من الفرق فى الصعوبة ، يتضح من جدول (٣٧) أن هذا الفرق يزيد بمقدار ضئيل عن مجموع الخطأ المعيارى لهما ، أى عدم وجود فجوات واضحة على متصل صعوبات المفردات و خاصة فى وسط المتصل و بذلك تكون المفردات منتظمة فى تدرجها على مدى ميزان القياس ، أى أنها تعرف مفردات الاختبار لجميع مستويات المتغير الذى يمثل هذا المتصل ، مما يشير إلى صدق تعريف المفردات للمتغير موضوع القياس .

#### جدول (٣٧)

العلاقة مسافة الفراغ (فجوة) بين كل مفردتين متتاليتين على خريطة المتغير

للاختبار المختلط بالمنف

الوصف الاحصائى المستخدم لتحديد العلاقة						
نوع المفردة	المفردتين المتتاليتين	الخطأ المعيارى <sup>٧</sup>	تقديرات الصعوبة	فرق الصعوبة	مجموع الخطأ	الحكم
mcq	19	50.85	40.8	1.2	101.55	0
cr	1	50.7	39.6			
mcq	22	50.95	38.4	1.25	102	0
mcq	20	51.05	37.15			
mcq	20	51.05	37.15	3.45	102.45	0
mcq	1	51.4	33.7			
crq	35	50.4	57.15	1.75	100.8	0
crq	27	50.4	55.4			
mcq	48	50.6	50.4	0.2	101.2	0
mcq	18	50.6	50.2			
mcq	10	50.6	47.4	1.55	101.25	0
mcq	42	50.65	45.85			
mcq	3	50.7	43.65	1.55	101.45	0
mcq	26	50.75	42.1			
crq	40	50.4	57.4	2.55	101	0

<sup>٧</sup> الدرجة بالمنف = ٥٠+٥٠ \* الدرجة باللوجيت .

			54.85	50.6	71	mcq
0	100.7	0.9	53.9	50.35	29	cr
			53	50.35	15	cr
0	100.95	0.6	52.25	50.35	33	cr
			51.65	50.6	75	mcq
0	101.2	0.9	50.65	50.6	55	mcq
			49.75	50.6	27	mcq
0	101.35	2.75	46.9	50.65	39	mcq
			44.15	50.7	25	mcq
0	101.55	1.85	43.35	50.75	31	mcq
			41.5	50.8	4	mcq
0	101	2.1	57.3	50.4	43	crq
			55.2	50.6	76	mcq
0	100.7	0.65	53.75	50.35	32	crq
			53.1	50.35	16	crq
0	100.95	1.2	53.1	50.35	16	crq
			51.9	50.6	100	mcq
0	101.2	3.4	51.9	50.6	100	mcq
			48.5	50.6	68	mcq
0	101.2	0	50.4	50.6	70	mcq
			50.4	50.6	47	mcq
0	101.2	1.35	50.4	50.6	47	mcq
			49.05	50.6	49	mcq
0	101.35	2.6	46.85	50.65	64	mcq
			44.25	50.7	9	mcq
0	101.5	2.1	43.45	50.7	37	mcq
			41.35	50.8	7	mcq
0	101.25	1.8	57.05	50.65	65	mcq
			55.25	50.6	92	mcq
0	100.7	0.75	53.65	50.35	36	crq
			52.9	50.35	37	crq
0	101	6	52.9	50.35	37	crq
			46.9	50.65	39	mcq
0	101.25	-4.1	46.9	50.65	39	mcq
			51	50.6	72	mcq
0	101.2	0.95	51	50.6	72	mcq
			50.05	50.6	50	mcq
0	101.2	2.6	50.05	50.6	50	mcq
			47.45	50.6	69	mcq
0	101.3	3.7	47.45	50.6	69	mcq
			43.75	50.7	5	mcq
0	101.2	0.9	53.75	50.6	52	mcq
			52.85	50.6	21	mcq
0	101.2	0.65	52.85	50.6	21	mcq
			52.2	50.6	82	mcq
0	101.2	1.75	52.2	50.6	82	mcq
			50.45	50.6	73	mcq

0	101.2	0.25	50.45	50.6	73	mcq
			50.2	50.6	51	mcq
0	101.2	2.45	50.2	50.6	51	mcq
			47.75	50.6	63	mcq
0	101.2	-1.35	53.85	50.6	79	mcq
			55.2	50.6	76	mcq
0	101.2	3.1	55.2	50.6	76	mcq
			52.1	50.6	85	mcq
0	101.2	1.4	52.1	50.6	85	mcq
			50.7	50.6	74	mcq
0	101.2	0.6	50.7	50.6	74	mcq
			50.1	50.6	66	mcq

## ٢- حساب ثبات الاختبار :

إعتمد التحقق من ثبات الاختبار التحصيلي المختلط على ما يلي :

### ٢-١ ما يوفره نموذج راش من استقلالية القياس :

إن إستقلالية القياس و تحرره الذى يوفره نموذج " راش " أتاح الفرصة لثبات القياس ، حيث لا يختلف القياس ( سواء أكان ذلك لقدرة الفرد أم لصعوبة المفردة ) باختلاف عينة التدرج (أمينة كاظم و آخرون ، ١٩٩٤ : ١٣١ ) .

ولقد تحققت استقلالية القياس بتحقق ملائمة كل من المفردات و الأفراد للنموذج ، و ذلك وفق لمحكات الملاءمة الإحصائية المختلفة ، مما يحقق ثبات قياس كل من الأفراد و المفردات.

### ٢-٢ ما يوفره برنامج Winsteps من تقديرات للخطأ المعياري :

يوفر برنامج Winsteps تقديراً للخطأ المعياري لتقدير القدرة العقلية لدى كل فرد ، و كذلك لكل مفردة من مفردات الاختبار ، و يعتبر هذا بدوره مؤشراً دقيقاً لمدى ثبات المقياس ، و بالرجوع إلى جدول تدرج الاختبار فى صورته النهائية - جدول (٣) فى ملحق (١) لوحظ أن قيم الخطأ المعياري لتقديرات صعوبة المفردات تعتبر صغيرة نسبياً حيث تراوحت ما بين (٠.٠٧ : ٠.٢٨) لوجيت مما يدل على ارتفاع ثبات مفردات الاختبار فى صورته النهائية .

أما قيم الخطأ المعياري لتقديرات قدرة الأفراد فقد تراوحت ما بين (0.16) و (0.40) لوجيت مما يحقق ثبات تقدير قدرات الأفراد ، كما يتضح من جدول (١٤) فى ملحق (١) .

### ٢-٢ معامل كيوذر ريتشارسون 20 :

- يقوم برنامج Winsteps بحساب معامل الثبات لكل من قدرة الأفراد و صعوبة المفردات و هو معامل المكافئ لمعامل ثبات كيودر - ريتشارسون 20 حيث كانت قيمته :
- معامل ثبات تقدير قدرة الأفراد = 0.96
  - معامل ثبات تقدير صعوبة المفردات = 0.99
- ويدل ارتفاع قيمة معامل ثبات صعوبة المفردات (0.99) على ارتفاع ثبات القياس لتقدير صعوبات مفردات الاختبار في صورته النهائيه ، أما قيمة معامل ثبات الأفراد فجاءت (0.96) و هذا يعتبر ثبات تقدير قدرات الأفراد و يؤكد ثبات المقياس في صورته الجديدة .

#### ▪ تقدير قدرة الأفراد لكل درجة كلية محتملة على الاختبار :

استخدمت الباحثة برنامج Wensteps مرة أخرى و ذلك لاستخراج الجدول رقم (٢٥) فى ملحق (١) و الذى يوضح العلاقة بين كل درجة كلية محتملة على الاختبار فى صورته النهائية و القدرة المقابلة لها باللو جيت ، ثم قامت الباحثة بتحويل الدرجة باللو جيت إلى المنف حتى تتخلص من القيم السالبة ، يتضح من جدول (٢٥) ملحق (١) ما يلى:

✓ أن العلاقة بين الدرجة الكلية الخام المحتملة على الاختبار و القدرة المقابلة لها هى علاقة طردية موجبة حيث تزداد القدرة كلما ازدادت الدرجة الكلية الخام المقابلة لها .

✓ يمتد مدى الدرجات المحتملة من (٠) إلى (١٥٥) درجة خام ، كما تمتد القدرة من ( - ٦.٥٤ : ٦.١٤ ) لو جيت و هو مدى واسع .

✓ يلاحظ أن قيم الخطأ المعيارى يقل بدرجة كبيرة كلما ابتعد عن الأطراف و هى نتيجة أن أصعب المفردات و أسهلها عادة تكون أبعد المفردات عن مستوى قدرة أفراد العينة ، مما قد يزيد من أخطاء القياس .

✓ يقدم البرنامج المستخدم تقديراً تقريبياً للقدرة المقابلة لكل من الدرجة الخام (صفر) و الدرجة العظمى (١٥٥) و يلاحظ ارتفاع الخطأ المعيارى لهاتين القدرتين .

نتائج الدراسة و مناقشتها: يُعرض فيما يلى نتائج الدراسة و مناقشتها:

١- نتائج الفرض الأول للدراسة: ينص الفرض الأول على أنه " لا تختلف تقديرات معلم صعوبة المفردة و معلم قدرة الفرد و دالة معلومات الاختبار باختلاف نمط مفردة الاختبار(ثنائية الدرجة - متعددة الدرجة) التي يتضمنها الاختبار التحصيلي " ، و للتحقق من صحة هذا الفرض تم

٢- نتائج الفرض الثاني للدراسة: ينص الفرض الثاني على أنه " يمكن وضع اختباري التحصيل (موضوعي ، المقالى ) فى صورة اختبارية مدمجة على تدرج واحد مشترك له صفر واحد " ، و للتحقق من صحة هذا الفرض تم

٣- ٢- نتائج الفرض الثالث للدراسة: ينص الفرض الثالث على أنه " يمكن اشتقاق صورة اختبارية مختصرة من الصورة الاختبارية المدمجة وتكون متعادلة معها " ، و للتحقق من صحة هذا الفرض تم

#### التوصيات المقترحة:

يمكن الخروج بالتوصيات التالية بناء على نتائج الدراسة الحالية:

#### مراجع الدراسة