مجلة الروائز
ISSN: 2602-5914/ EISSN: 2716-9596
المجلد: 07 العدد:01 (السنة: 2023) ص ص: [21-40]

الخصائص السيكومترية لأدوات القياس في علم النفس
Psychometric characteristics of psychological measurement tools

خمنو دنيا *

جامعة مولود معمري تيزي وزو (الجزائر) dounia.khemnou @ummto.dz

تاريخ الإرسال: 2023/03/21 تاريخ القبول: 2020/05/24

ملخص:

تتميز الدراسة العلمية الموضوعية للسلوك الإنساني باعتمادها اعتمادا أساسيا على الأساليب الكمية في قياس السلوك بمختلف جوانبه ، فبدون هذه الأساليب تقتصر دراسة السلوك الإنساني على الوصف اللفظي ، و الملاحظة الذاتية ، و التأملات الفردية ، و قد أدت حركات التكميم في القرن التاسع عشر و التي تميزت باستخدام التجريب و القياس في ميدان علم النفس إلى ظهور الاختبارات والمقاييس التربوية و النفسية كأساليب بارا مترية تمكن من أجرأة المفاهيم المجردة والتعبير عنها مترولوجيا مما أضفى على هذه العلوم طابع الدقة والجرأة وسمة البارامترية، و لكن الجدير بالذكر هو ما مدى صلاحية هذه الأدوات للقياس ، و ما مدى صحة النتائج التي تستخلص منها من خلال عملية القياس ، هذا ما سنتعرض إليه من خلال عرض الخصائص السيكومترية المعبرة عن جودة أدوات القياس النفسي مما يمكن اعتماد النتائج المستخلصة منها .

الكلمات المفتاحية: الخصائص السيكومترية؛ القياس النفسي؛ الاختبارات والمقاييس النفسية؛ أدوات القياس ؛ الأساليب الكمية

Abstract:

The objective scientific study of human behavior is distinguished by its dependence mainly on quantitative methods in measuring behavior in its various aspects. Without these methods, the study of human behavior is limited to verbal description, self-observation, and individual reflections. The Quantification movements in the nineteenth century, which were distinguished by the use of experimentation and measurement in the field of psychology led to the emergence of educational and psychological tests and standards as parametric methods that enable the daring of abstract concepts and express them metrologically, which gave these sciences the character of accuracy, daring and the parametric feature, but it is worth noting the extent of the validity of these tools for measurement, and the extent of validity of the results that are drawned from it through the measurement process, This is what we will be exposed to by presenting the psychometric characteristics that express the quality of psychological measurement tools, so that the results drawn from them can be adopted.

Keywords: Psychometric characteristics, Psychological measurement, Psychological tests and measures, quantitative methods in measuring, measurement process

* خمنو دنيا

مقدمة:

يوجد العديد من المؤشرات التاريخية الدالة على أن الإنسان حاول منذ عدة قرون قياس و تقويم مختلف أنواع السلوك .إلا أن علم النفس القياسي يعتبر حديث العهد،فان بداياته الحقيقية تعود إلى حوالي المائة سنة فقط.

و تجدر الإشارة إلى أن أول من فكر في ضرورة تخصص لعلم النفس يهتم بالقياس و الكم هو فان و ولف van wolffفي سنة 1732 حيث ودعا أمكانية إنشاء فرع رياضي لعلم النفس أطلق عليه تسمية علم النفس القياسي و عرفه كما يلي"إن علم النفس القياسي الذي يجب إحداثه سيهتم بتطوير المعرفة الرياضية للنفس البشرية .إذ انه علينا تعليم كيف يمكن قياس مدى صحة و عدم صحة التقديرات و مدى يقينها"(عن بيرني و لويز Bernier et Pietrulewiez 1997).

إن المتأمل في المجتمع المعاصر يجد أن من أهم ما يميزه هو أولا الدقة في قياس الأشياء و ثانيا اتساع رقعة القياس حيث أصبحت تشمل جوانب كان يعتقد سابقا منذ عشريات قليلة أنها غير خاضعة للقياس، ومن هذه الجوانب الخصائص الذهنية والنفسية، والتقدم الذي أحرزته العلوم إنما كان بفضل الدقة في القياس التي تتميز بها.

الإشكالية:

إن التقدم السريع الذي عرفه علم النفس منذ بداية القرن التاسع عشر إنما كان بفضل تطور أدوات القياس، فمنذ اكتشاف ظاهرة الفروق الفردية في القدرات العقلية والخصائص النفسية ،حيث نجد كما يشير سعد جلال (1985) مربين تصدروا الساحة الفكرية في القرن الثامن عشر و اوائل التاسع عشر من أمثال روسو و بستالوتزي و هربارت و فرويل (المربين الطبيعيين) اهتموا بالطفل و بطبيعته و نادوا بملاحظته ووضع البرامج التي تتفق و هذه الطبيعة.و نجد في كتابتهم ما يدل على فهمهم لمبدأ الفروق الفردية ،غير أن اهتمامهم كان موجها إلى دراسة الطفل باعتباره ممثلا للأفراد و لم يكن اهتمامهم يقوم على أساس انه فرد يختلف عن غيره من الأفراد .كما تجدر الإشارة إلى إسهامات العالم البيولوجي البريطاني غالتون(1822_1911)Galtonفي القياس أهم معلم لتطور حركة القياس ببريطانيا .فقد كان غالتون مهتما بدراسة الوراثة لدى الإنسان و اهتم خاصة بدراسة الصفات المتشابهة و المختلفة بين الأقارب (أباء و أبناء و إخوة و أبناء العم ...)و عند مجموعة الأفراد لا تربط بعضهم بعض علاقة القرابة.و لقد انشأ في 1884 مخبرا لعلم الانتروبومتري حيث عمل على استقبال عدد كبير من المفحوصين أجرى عليهم ، بالإضافة إلى القياسات التي كانت تجرى في مخبر ليبزغ قياسات فيما يخص أبعاد الجسم،و كان يأمل إلى الوصول إلى تقدير المستوى العقلي للمختبر و ذلك من خلال قياس العمليات الحسية ،و قد اكتشف أن هناك نقصا في حواس ضعاف العقول مشير إلى أن القدرة على التمييز الحسي تكون أعلى عند القادرين عقليا،و من إسهامات غالتون تطويره لعمليات إحصائية بالتعاون مع تلميذه كارل بيرسون(Pearson)،و من أهم المقاييس الإحصائية التي تم التوصل إليها نذكر معامل الارتباط أو ما أصبح يعرف بمعامل ارتباط بيرسون و الوسيط و السلم الترتيبي و غير ذلك .و مما قاله بيرسون عن أستاذه غالتون هو أن هذا الأخير لا يقل أهمية عن فاندت حيث أن له الفضل في تأسيس منهج جديد في البحث النفسي و هو المنهج الإحصائي و قد تأثر عدة باحثين بريطانيين بمنهج غالتون ،فإلى جانب بيرسون نجد كل من سبيرمان (Spearman)وبيرت(Burt).قدم سبيرمان في سنة 1904بحثه المشهور حول الذكاء و طرق قياسه ،و طور معامل الارتباط الرتبي .و ترجع أهمية أعماله هذه إلى أنها أساس نظريته في الذكاء المعروفة بنظرية العاملين في الذكاء و التي قدمها سنة 1914.أما بيرت فقد طور

اختبارات لقياس العمليات العقلية العليا مثل التفكير و الانتباه،و قدم في عام 1913بحثا بعنوان اختبارات تجريبية للعمليات العقلية العليا،كما انه استعمل التحليل العاملي لتحليل القدرات الذهنية.

من هذا المنطلق أصبح علماء النفس يكرسون جهودهم وأعمالهم في تصميم المقاييس و الاختبارات النفسية التي تقيس مختلف الظواهر ذات الصلة بهذا العلم،وقد أدى هذا إلى ظهور المئات من الاختبارات التي أصبحت تستعمل لعدة أغراض وفي عدة مجالات كالاختيار و التوجيه و التدريب المهني و الإرشاد النفسي و التشخيص التربوي و العيادي و غير ذلك ، و في هذا السياق أصبحت الاختبارات و المقاييس النفسية أدوات لا يستغني عنها أي باحث في ميدان علم النفس ، بحيث أصبح السلوك الإنساني في كل أشكاله خاضعا للقياس و التقويم ، إذ أن معظم الأحكام و القرارات المتعلقة بالانتقاء والتشخيص و الإرشاد و التنبؤ بالسلوك الإنساني المستقبلي تعتمد على نتائج القياس و الاختبارات النفسية و التربوية ، و لهذا كانت هذه المقالة حول الخصائص السيكومترية (الميترولوجية)، أدوات القياس النفسي ، حيث تمكننا هذه الأخيرة من التأكد من مدى صدق و ثبات النتائج المتحصل عليها من هذه الأدوات.؟ ومامدى جودة أدوات القياس النفسي مما يمكن اعتماد النتائج المستخلصة منها؟وما مدى صحة النتائج التي تستخلص منها من خلال عملية القياس؟

1. القياس النفسي:

1.1 مفهوم القياس :

إن القياس كمفهوم امبريقي له عدة تعريفات، ولعل أبسطها أنه وصف البيانات المتعقلة بخصائص الأشياء باستخدام الأعداد أو الجوانب الكمية في وصف السمات أو خصائص الأفراد .و في هذا السياق أيضا يرى كل من " "Smithو "Adams " (1972) أن القياس بمعناه الواسع هو الجمع المنظم للمعلومات بترتيب معين ، و هذا يتضمن عملية الجمع و تنظيم المعلومات و ناتج هذه العملية " و يعرفه أيضا " نانالي " Nunnaly (1972) أنه يشتمل على قواعد تعيين أعداد الأشياء ،بحيث تمثل مقادير سمات هذه الأشياء ، ويتفق هذا التعريف مع سبق أن قدمه ستيفنز،Stevens(1951م) " أن القياس هو تعيين أعداد أو رموز رقمية للأشياء أو الأحداث وفقا لقواعد معينة ، فالقياس إذن يتطلب إجراءات أو عمليات معينة ، استنادا إلى قواعد محددة تستخدم في المقارنة بين الأشياء أو الأحداث وفقا لمعيار Standard ، أو ميزان Scale معرفا تعريفا دقيقا.

2.1 مفهوم الاختبارات النفسية:

يمكن القول إن الاختبار ما هو إلا إجراء منظم لقياس عينة من السلوك، من خلال هذا التعريف يمكن الإشارة إلى قضايا متعددة :

_الاختبار إجراء منظم: يعني أن الاختبار طريقة تتألف من خطوات متتابعة، فكل خطوة تتطلب مجموعة من القوانين، أي أن الاختبار طريقة معيارية يؤخذ فيها كل عمل بناء على قوانين موصوفة .

_ السلوك: إذ أن فقرات الاختبار التحصيلي مثلا إنما هي الأعمال أو سلوكا يمكن أن يكون دليلا على تحصيل الطالب.

_ العينة: إذ أن لكل سمة من السمات عددا من السلوكيات التي تدل عليها، وبالتالي يمكن كتابة العديد من الفقرات التي تعكس كل السلوكيات الدالة على هذه السمة انطلاقا من العينة المستهدفة.

بعد ما قمنا بتعريف الاختبار بصفة عامة كان إلزاما علينا أن نتطرق إلى مفهوم الاختبار النفسي: يعرفه Chronback سنة 1984م أنه لا يوجد تعريف مقنع للاختبار النفسي، فكلمة اختبار عادة ما توحي في الذهن، أنه عبارة عن سلسلة من أسئلة المقننة التي تعرض على شخص معين، و يطلب منه الإجابة عنها كتابة أو شفهيا أما Tyler et Leona 1982م فيعرفانه أنه " موقف مقنن مصمم خصيصا للحصول على عينة من السلوك ". صالح مراد وفوزية هادي (2002 ص: 182)

2. طرق تصميم البحوث النفسية في علم النفس:

إن **القياس** النفسي لابد و أن يعتمد على بعض المبادئ الهامة التي فيها الموضوعية والدقة، يعني ألا يتأثر **الباحث** في وصفه للظاهرة التي يقيسها أو في تفسيرها بميوله الذاتية أو آرائه الشخصية أو تعصبانه أو تحيزا ته أو في عقائده و أفكاره و تجاربه الخاصة إنما يسجل الوقائع كما هي موجودة بالفعل لا كما يريدها أن تكون ، كذلك من مبادئ **القياس** الجيد أن تكون الاعتبارات و الأدوات المستخدمة صادقة بمعنى أنها تقيس فعلا السمة المراد قياسها و لا تقيس عرضا لسمات أخرى ، و يجب أيضا أن تكون ثابتة بمعني أن تعطي نتائج ثابتة كلما تم إعداد التجربة على نفس الأفراد و بدون تغيير في سير الظروف.

هذا ويجدر وجود معايير يتم السير وفقها للحصول على النتائج وهو ما يحتم ضرورة وجود خطوات مقننة للقياس النفسي على الأفراد المختلفين الذين يتبعون نفس الخطوات في سير البحث.

إن علماء النفس يهتمون بفهم الإنسان ككل، كما يهتمون بالتنبؤ بسلوكه ككل أيضا ويهتمون بالتحكم في السلوك، وإلى جانب هذا الاهتمام بالإنسان ككل هناك اهتمامات أخرى لعلماء النفس وهي الرغبة في فهم جوانب نوعية محددة جزئية من سلوك الإنسان، فعلماء النفس يحاولون أن يعرفوا أنواع السلوكات الجزئية التي تترابط معاً أو تلك التي تظهر معا أو تختفي معا، كما يهتمون بمعرفة أي نوع من السلوك عندما يوجد الفرد في موقف معين، ومن أمثلة هذه المشكلات النوعية المحددة التي يحاول علماء النفس إيجاد حلول لها ما يلي:

- هل يتعلم الفأر الجائع الخروج من المتاهة التي يوضع فيها أسرع من الفأر الشبعان؟.

- هل يستطيع الطالب الجامعي المستجد القلق استقبال المعلومات العلمية بنفس الدقة التي يستقبلها بها زميله المستريح؟.

- هل استذكار المادة ككل أسهل من استذكارها جزءا جزءا؟ و بعبارة أخرى هل يحفظ الطالب قائمة من المقاطع عديمة المعني أسرع إذا أخذ في حفظها كلها ككل دفعة واحدة عن إذا جزأها إلى أجزاء صغيرة و استذكرها جزءا جزءاً.

و هكذا بالنسبة لآلاف من المشكلات السلوكية التي يهتم بها علماء النفس و التي لابد من دراستها في ضوء الضبط التجريبي و الدقة الموضوعية. ومن أولي خطوات البحث العلمي تعريف المتغيرات أو العوامل أو السمات أو الظواهر التي يتناولها البحث، فالظاهرة التي تدرسها لابد من تعريفها تعريفاً إجرائياً موضوعياً دقيقاً، ولابد أيضاً من الاعتماد على المقاييس الكمية وليست العبارات الوصفية اللفظية ومعني ذلك الاعتماد على الوسائل الإحصائية.

كما نستطيع أن نتعرف على طبيعة مثير سمعي ما، فهناك بعض الأجهزة الإلكترونية التي تصدر صوتاً ما ذا كثافة أو شدة معينة أو ذا تكرار معين كما يريده **الباحث** و ذلك بمجرد إدارة قرص بسيط في الجهاز، و لكن الصعوبة

تكمن في قياس الإحساس الذي يتركه هذا المثير، إننا نريد أن نعرف العلاقة بين حدوث تغير في شدة المثير الذي يحدث في الإحساس هل يحدث تغيير في الإحساس بنفس المقدار أو الكم الذي يحدث بين التغير في المثير؟.

هل يتماشى التغير الذي يحدث في كثافة المثير مع المتغير الذي يتبعه في الإحساس؟ لقدر اخترع علماء النفس بعض المقاييس السيكو فيزيولوجية لقياس أبعاد الوعي أو الشعور، و لقد أخذنا زمن الرجع، هل حقيقة يعتبر هذا المتغير سهل **القياس**، هل نستطيع حقيقة أن نقيس المسافة أو الفترة الزمنية بين سماع الفرد صوتاً معيناً و قيامه بالضغط على زر معين قد يكون هذا في حد ذاته سهلا و لكن الصعوبة عندما يكتشف أن الشخص المعين ليس له معدلاً واحداً لزمن الرجع في الموقف الواحد، فإذا كررنا تجربة ما فإننا نحصل على درجات مختلفة لفرد معين في نفس الموقف أي أننا نحصل على توزيع الدرجات و لا نستطيع أن نحدد زمن الرجع الحقيقي لهذا الفرد. إن البحوث المعملية تحتاج إلى ما يلي:

- تعريف المتغيرات أو العوامل أو الظواهر المراد إجراء التجربة عليها.
- تصميم التجربة تصميماً دقيقاً.
- ضبط للعوامل والمتغيرات المتعلقة بالتجربة.
- قياس الاستجابات قياساً دقيقاً.
- تسجيل النتائج.
- العوامل المستقلة هي التي تقيس أثرها كالذكاء مثلاً.
- العوامل المعتمدة هي التي تقوم بملاحظتها كالتحصيل.

أي أننا لا نستطيع أن نتغلب على مشكلات المقارنة واستخلاص النتائج من البحوث النفسية إلا باستخدام **الأساليب الإحصائية**. (العيسوي، عبد الرحمن ،1997،ص 43).

3. الخصائص السيكومترية (المترولوجية) لأدوات القياس النفسي: او خصائص الاختبار و المقياس الجيد)

يعد تصميم الاختبارات النفسية من اجل التنبؤ بأمور مثل النجاح المدرسي من أوائل التطبيقات العملية للمبادئ النفسية.

إن جودة الأدوات المستخدمة في البحث هامة ألن النتائج التي يتوصل إليها الباحث تعتمد على المعلومات التي يتحصل عليها باستخدام هذه الأدوات و لهذا كان إلزاما التأكد من خصائصها السيكومترية الخاصة بها:

حدد علماء النفس أربع خصائص يبدو أنها أكثر الخصائص أهمية بالنسبة للاختبار الجيد وهذه الخصائص هي: الثبات، الصدق، التقنين، الموضوعية.

2.3 الثبات:

يشير الثبات إلى الاتساق الذي يتم بمقتضاه الحصول على نفس النتيجة حيمما تستخدم صورا متماثلة أو متكافئة للاختبار الواحد. و نحن لا نتوقع الاتساق الكامل و إنما تعد الدرجة العالية من الثبات امرأ أساسيا، فمن المستحيل الحصول على تنبؤات دقيقة أو ذات معنى من نتائج اختبار لا يتصف بالثبات.

يقصد بثبات الاختبار مدى الدقة أو الاتساق أو استقرار نتائجه فيما لو طبق على عينة نمن الأفراد في مناسبتين مختلفتين، ويمكن النظر إلى ثبات الاختبار من ثلاث جوانب هي :

_ الحصول على نفس النتائج مهما أعدنا تطبيق الاختبار .

_ أن يقيس الاختبار ما اعد لقياسه .

_ احتمال و جود أخطاء في القياس يعني درجة الفرد في الاختبار تساوي مجموع الدرجة الحقيقية و الدرجة الخطأ . (عبد الكريم بوحفص(2011)،ص 221) .

1.2.3 العوامل المؤثرة في ثبات الاختبار :

يوجد العديد من العوامل التي تولد كمية من خطأ القياس و بالتالي تؤثر في ثبات الاختبار و من هذه العوامل ما يلي :

-الفترة الزمنية : الفترة الزمنية الموجودة بين تطبيق الاختبار و إعادة تطبيقه ، فكلما كانت طويلة كلما أثر على ثبات الاختبار هذا راجع إلى عامل النضج و التعلم.

- ظروف إجراء الاختبار

- استقلالية بنود الأداة بعضها عن بعض، بحيث أن الارتباط المرتفع بين بنود الأداة تقلل من معامل ثباته.

- .طول الاختبار : كلما كانت البنود أكثر كلما زاد ثبات الاختبار و هذا لزيادة تمثيل البنود للخاصية المقاسة.

-التخمين في الإجابة على البنود .

-مستوى صعوبة و سهولة البنود.

- موضوعية تصحيح المقياس.

- زمن الاختبار : حيث أن زمن الاختبار كلما كان طويل كلما انخفض معامل ثباته . (بشير معمرية(2007) ص: 90)

2.2.3 طرق اختبارثبات أدوات القياس النفسي:

توجد أربع طرق معروفة تستخدم لاختبار ثبات أدوات القياس هي :

أ- طريقة التطبيق وإعادة التطبيق :

يتمثل هذا الإجراء في تطبيق الأداة على نفس المجموعة في فترتين زمنيتين مختلفتين، ثم يتم حساب معامل الارتباط بين نتائج المجموعة على التطبيقين، وتشيرAnastasie Anneم(1982) أن معامل الثبات الجيد 0.90و0.80 مابين يتراوح.

ويسمى معامل الثبات بإعادة التطبيق باسم معامل الاستقرار، وهو يدل على استقرار الدرجات عبر الزمن ويتأثر معامل الثبات بعدد من العوامل أهمها أخطاء استقرار استجابات الأفراد، وتذكرهم لبنود الاختبار من التطبيق الأول.

وتصلح هذه الطريقة لحساب ثبات جميع الاختبارات، عدا اختبارات الذاكرة، شريطة أن يتراوح الفاصل الزمني بين أسبوعين وستة أسابيع، ومراعاة نفس الظروف.

ومن عيوبها صعوبة ضبط الظروف في التطبيقين وتأثر إجابات الأفراد، خاصة في الاختبارات التحصيلية، بانتقال أثر التدريب وعاملي النضج والتعلم، كما أنها مكلفة في الوقت والجهد.

ب- طريقة الصورتين المتكافئتين :

يشتمل هذا الإجراء على تطبيق الأداة محل البحث في نفس الوقت الذي تطبق فيه أداة أخرى يصممها الباحث بحيث تكون متكافئة مع الأداة محل البحث بحيث تكون تقيس نفس الخاصية ثم يقوم الباحث بحساب معامل الارتباط بين نتائج المجموعة الواحدة على الأداتين.

بالإضافة إلى ذلك فإن تكافؤ الصورتين يعنى تساوى معاملات الارتباط بين البنود (المعاملات البينية) في كلتيهما وكذلك تساوى المتوسط والانحراف المعياري لكلتا الصورتين ومن ثم ففي هذه الطريقة يفترض مصمم الاختبار أو الاستفتاء أنه يمكن تكوين صورتين متساويتين أو متكافئتين من الاختبار الواحد وهذا التكافؤ يشتمل على الجوانب التالية:

1_ عدد مكونات الوظيفة التي يقيسها الاختبار.

2_عدد الفقرات التي تخص كل منهما.

3_ مستوى صعوبة الفقرات.

4_طول الاختبار وطريقة إجرائه وتصحيحه وتوقيته.

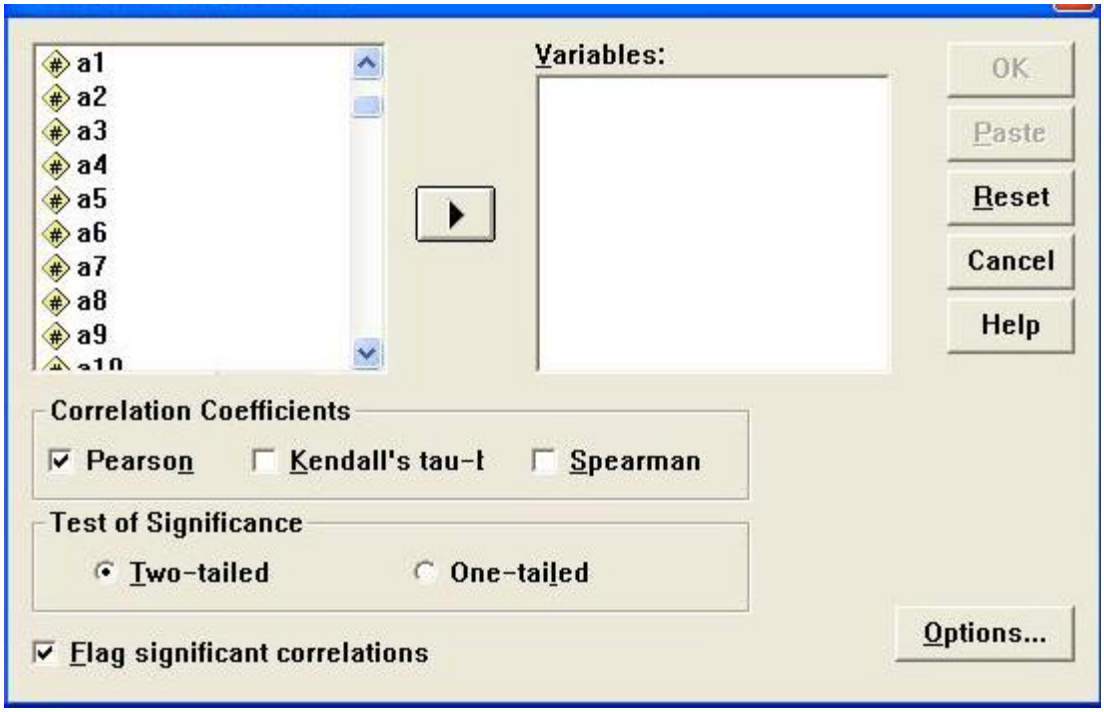
5_ تساوى متوسط وتباين درجات الأفراد على كل الصور.

ورغم إن طريقة الصور المتكافئة للاختبار تقدم لنا أساساً سليماً جداً لتقدير الدقة في اختبار نفسي إلا أن هذه الطريقة تثير عدداً من المشكلات العملية .فهي تطلب توفر صورتين متكافئتين للاختبار، وتوفر وقت يسمح باختبار كل فرد مرتين .في بعض الأحيان، ولا توجد صورة ثانية من الاختبار، أو قد لا يتوفر الوقت لإجراء تطبيق ثان .فتطبيق اختبار ثان منفصل كثيراً ما يمثل مطلباً ثقيلاً على الموارد المتاحة، مثل هذه الاعتبارات العملية المتعلقة بتوافر الإمكانيات جعلت الكثيرين من مؤلفي الاختبارات يتوجهون إلى إجراءات تعتمد على استخلاص تقدير للثبات من تطبيق صورة واحدة فقط من الاختبار .على أن مثل هذه الأساليب هي في أحسن الأحوال مجرد حلول وسط . والطريقة التي تعمد على استخراج الارتباط بين صورتين متكافئتين تطبقان في العادة بفاصل زمني يمتد إلى عدة أيام أو عدة أسابيع، تمثل الطريقة المفضلة في تقدير الثبات. (عبد الكريم بوحفص(2011)،ص 222) .

وتتميز هذه الطريقة بعدم التأثر بالتدريب أو الخبرة من صورة لأخرى، ويختفي أثر الألفة اختلاف بنود الصورتين، كما أنها تصلح لحساب معامل ثبات اختبارات الذاكرة والاختبارات التحصيلية.

ومن عيوب هذه الطريقة صعوبة تصميم اختبارين متكافئين لقياس نفس السمة أو الخاصية. انتقال أثر التدريب وبخاصة في حالة اقتران عملية تطبيق الصورتين، وقصر الفترة الزمنية أو طولها بين تطبيق الصورتين المتكافئتين، والجهد المضاعف المبذول والوقت والتكلفة المبذولة في بناء الاختبار المكافئ.

(نود الإشارة هنا إلى أنه يمكن حساب معامل الارتباط بين الاختبار ونفسه " طريقة إعادة التطبيق " أو بين الصورتين المتكافئتين من الاختبار بالرجوع إلى قائمة Analysis ثم نختار Correlate ونتعامل مع درجات الاختبار ونفسه على أنهما متغيرين ونحسب معامل الارتباط بينهم ويكون هو معامل ثبات الاختبار) .



ج- طريقة التجزئة النصفية:

تتمثل هذه الطريقة في تطبيق الأداة مرة واحدة على مجموعة واحدة، ثم تقسيم الأداة إلى جزأين متساويين، وقد يكون من المناسب أن يشمل التقسيم البنود ذات الأرقام الفردية، في حين يشمل الجزء الثاني البنود ذات الأرقام الزوجية، و في مرحلة ثانية يقوم الباحث بحساب معامل الارتباط بين نتائج المتحصل عليها في الجزأين و ذلك باستخدام معاملات الارتباط التالية: Roulon ; Gulteman;Spearman etBrown

وهناك عدة معادلات تستخدم لتصحيح معامل ثبات نصفى الاختبار منها:

معادلة سبيرمان – براونSpearman – Brown

وفيها يتم التعويض بمعامل الارتباط بين نصفى الاختبار لنحصل على معامل ثبات الاختبار ككل .وهى شائعة الاستخدام وبخاصة في اختبارات التحصيل والقدرات تحت ظروف محددة. (سعد عبد الرحمان 2008ص 100).

معادلة رولونRulon

وتعتمد على حساب تباين درجات الاختبار ككل، ثم حساب تباين الفروق بين درجات الأفراد في النصف الأول ودرجاتهم في النصف الثاني ثم تطبيق المعادلة نحصل على معامل ثبات الاختبار ككل.

معادلة جتمانGuttmann

وفى هذه المعادلة يتم حساب تباين درجات النصف الأول، وتباين درجات النصف الثاني، وتباين درجات الاختبار ككل أي أنها تضع في الاعتبار احتمال اختلاف تباين درجات النصف الأول للاختبار عن تباين درجات النصف الثاني " وهذا لا يتحقق في المعادلتين السابقتين . "

واستخدام طريقة التجزئة النصفية في تعيين معامل ثبات الاختبار يثير عدة ملاحظات هي:

1_ قد يختلف النصف الأول عن النصف الثاني، وخاصة إذا أخذت البنود من (1_50 مثلاً) ثم من (51 _100)، وهذا يعنى أن إجابات الأفراد في النصف الثاني سوف تتأثر بعوامل الإجهاد والملل وضيق الوقت أكثر من إجابات الأفراد في النصف الأول . وهذا ما يعطى نتائج لا يمكن الوثوق بها بدرجة كبيرة.

2_في حالة تقسيم الاختبار إلى نصفين عن طريق أخذ الأسئلة الفردية، والأسئلة الزوجية فإنه من المحتمل أن يختلف تباين درجات النصف الأول عن تباين درجات النصف الثاني.

3_ من الممكن تجزئة الاختبار إلى نصفين بعدة طرق مختلفة، فقد نأخذ البنود ذات الأرقام الفردية في مقابل البنود ذات الأرقام الزوجية أو البنود من(1_50) تمثل النصف الأول في مقابل البنود(51_100) لتمثل النصف الثاني . وهذا يعنى أنه من المحتمل أن نحصل على معامل ارتباط بين نصفى الاختبار في الحالة الأولى يختلف عن المعامل في الحالة الثانية، وهذه الملاحظة صحيحة وخاصة إذا كانت جميع بنود الاختبار على درجة واحدة من الصعوبة، أو إذا كانت البنود واردة بدون ترتيب معين مثل قوائم الشخصية ، واختبارات السرعة . (سعد عبد الرحمان 2008ص 100).

د- طريقة الاتساق الداخلي :

هي طريقة تدل على الثبات والصدق معا و هي طريقة الهدف منها هو قياس مدى ثبات البنود بالنسبة لمحك داخلي هو معامل الثبات الكلي للاختبار و ذلك باستخدامKR20 , KR21 Cronback ،بحيث تستخدم وفقا للشروط التالية:

CronbackAlfa :يستخدم في حساب معامل الثبات الخاص بالسلالم و التقديرات KR20 ; KR21 .يستخدم في حساب معامل الثبات الخاص بالأدوات ذات البدائل _صحيح،خطأ_ بحيث تعطى ل: صحيح ،و خطأ.

وتعتمد هذه الطريقة على مدى ارتباط الوحدات أو البنود مع بعضها البعض داخل الاختبار، وكذلك ارتباط كل بند مع الاختبار ككل . ومن أكثر المعادلات استخداماً لقياس الاتساق الداخلي هي:

_ معادلة كودر وريتشاردسون رقم (20)Kuder – Richardson :

وتعتمد على توفر بيانات عن تباين كل مفردة من مفردات الاختبار . و في حالة عدم توفر هذه البيانات يمكن استخدام الصيغة رقم (21) بنفس المعادلة وهى تتميز بالسهولة والسرعة فى حسابها حيث إنها لا تحتاج إلى معرفة تباين البنود ، ولكن يعيبها أنها أقل دقة من الصيغ السابقة . وقد وضع كيودر و ريتشارد سون شروطاً لاستخدام هذه المعادلة وهى:

_ أن تكون درجة أسئلة الاختبار (صفر أو واحد) .

_ألا يكون عدد الأسئلة المتروكة كبير .

_ تقارب مستوى صعوبة الأسئلة .

تساوى معاملات الارتباط بين درجات الأسئلة . ومن الواضح أن هذه الشروط لا توجد في اختبارات المجال الوجداني ولذلك ننصح بعد استخدامها.

_ معادلة ألفا – كرونباخ Cronbach Alpha

ويعتبر معامل ألفا حالة خاصة من معادلة كيودر وريتشارد سون ٠ ويمثل معامل ألفا متوسط المعاملات الناتجة عن تجزئة الاختبار إلى بطرق مختلفة ، وبذلك فإنه يمثل معامل الارتباط بين أي جزئين من أجزاء الاختبار ٠ ويتم حساب تباين كل بند من بنود الاختبار ثم مجموع التباينات ، وكذلك تباين الدرجة الكلية للاختبار ، وتشترط أن تقيس بنود الاختبار سمة واحدة فقط ٠ وتستخدم هذه المعادلة فى المقاييس والاختبارات متعددة الاختيارات وليست الثنائية .

و يمكننا هنا أن نطرح تساؤل : هل الاتساق الداخلي طريقة من طرق حساب الصدق أم طريقة لحساب الثبات ؟ .

من العرض السابق يتضح لنا أن مفهوم الاتساق الداخلي أكثر ارتباطاً بمعنى الثبات حتى أنه من طرق حساب الثبات . فكل الطرق السابقة تعطينا معامل الثبات على أنه مؤشر للاتساق الداخلي أما استخدام البعض لهذه الطريقة في التأكد من صدق الاختبار فتعتمد على حساب الارتباط بين كل مفردة والمجموع الكلى لدرجات الاختبار ، أو الارتباطات بين المكونات الفرعية للاختبار ، أو الارتباط بين المفردة ومجموع درجات البعد في الاختبار وغير ذلك من الطرق ويتضح من خلال ما سبق أن الارتباطات الخاصة بالاتساق الداخلي سواء ما يبنى على أساس المفردات أو ما يبنى على أساس الاختبارات الفرعية هي أساساً قياسات لتجانس الاختبار وحيث أنها تساعد على تحديد مجال السلوك أو السمة موضع القياس والتي يعتبر الاختبار عينة ممثلة لها .فإن درجة التجانس هذه يكون لها بعض الدلالة لصدق الاختبار التكويني .

3.2.3 طريقة تحليل التباين Analysis of Variance :

وهذه طريقة أخرى لتعيين معامل ثبات الاختبار عن طريق تحليل التباين والخاص بالمتوسطات المرتبطة والتباينات المختلفة حيث يكون معامل الثبات الاختبار: (علي عودة محمد 2012 ص202).

$$\text{معامل ثبات الاختبار} = \frac{\text{التباين بين الأفراد – تباين التفاعل}}{\text{التباين بين الأفراد}}$$

يتضح مما سبق وجود طرق مختلفة لحساب الثبات فإنه عندما تكون لدينا صورة واحدة من الاختبار يمكن أن نقدر الثبات بعدد من الطرق مثل إعادة الاختبار ، والتجزئة النصفية ، ومعامل ألفا ،أما الاختبارات التي يكون لها صورتان فيمكن أن نضيف طريقة الصور المتكافئة . ولكن أي من هذه الطرق هو الأفضل ؟ ومتى نستخدم طريقة دون الأخرى ؟

وعلينا هنا أن نحدد طبيعة وغرض الاختبار الذي نقدر له الثبات .فبالنسبة للاختبارات المعدة لتطبق مع الأفراد أكثر من مرة نتوقع أن يكون للاختبار ثبات عبر الزمن وفى هذه الحالة فإن الثبات المناسب هو ثبات إعادة الاختبار . وبالنسبة للاختبارات التي تهدف إلى أن يكون لها نقاء عاملي فإن معامل ألفا يصبح أساسياً لتقدير الثبات، وعلى العكس فإنه في حالة الاختبارات متعددة العوامل مثل مقاييس الذكاء فإننا لا نستفيد كثيراً من مقاييس الاتساق الداخلي وبذلك فإن معامل ألفا ومعادلة كيودر ريتشارد سون لن تعطينا تقديراً جيداً للثبات في هذه الحالة . ولهذا فإن معامل ألفا لا يعتبر مؤشراً جيداً للثبات لكل الاختبارات وإنما هو يناسب الاختبارات التي تشتمل على عامل واحد .كذلك فإن طريقة

تقدير الثبات باستخدام التجزئة النصفية تناسب أكثر الاختبارات التي تكون مفرداتها مرتبة بعناية تبعاً لمستوى الصعوبة.

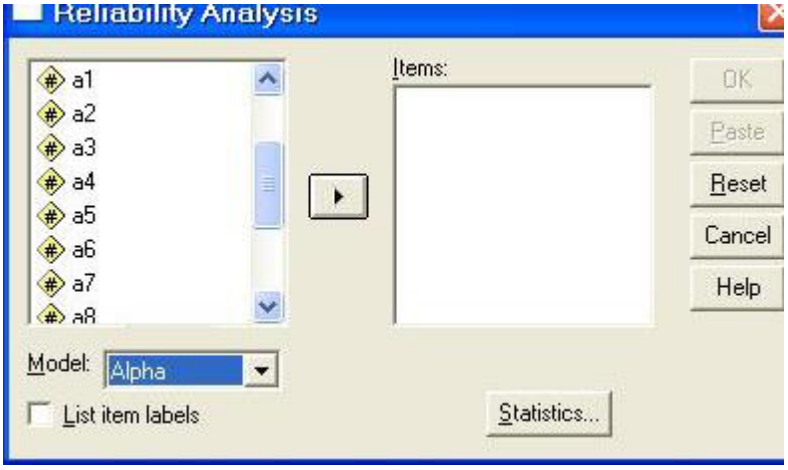
ويتبادر إلى الذهن سؤال هام هو : ما الحد الأدنى للثبات حتى يعتبر مقبولاً ؟ أو كم يجب أن يكون معامل الثبات مرتفعاً ؟

نحن نفضل أداة القياس الأكثر ثباتاً ، ولكن في الحقيقة لا توجد إجابة عامة عن هذا السؤال وإذا أردنا اتخاذ قراراً أو إجراءاً ما يتعلق بفرد معين ، فما علينا إلا أن نفعل ذلك معتمدين على أفضل ما يتيسر لنا من معلومات مهما كإثباتها ، بشرط واحد وهو أن يكون الثبات أكثر من الصفر ، "طبعاً هنا وفى جميع الأحوال . يكون الاعتبار الأول لصدق المقياس " إن تقويمنا لأي أداة جديدة يجب أن يكون دائماً بمقارنتها بأدوات أخرى تنافس معها ، فإذا كان لدينا اختبار للرياضيات في المرحلة الثانوية بمعامل ثبات (0.80) فإنه يصبح غير مفضل إذا توفرت لنا اختبارات أخرى يتراوح ثباتها (0.85_0.90). وقد أظهرت الدراسات إمكانية استخدام اختبارات ذات معاملات ثبات منخفضة إذا كنا نريد دراسة مجموعات وبخاصة إذا كانت مجموعات ذات أعداد كبيرة ، ولكننا نحتاج إلى اختبارات ذات ثبات مرتفع واضح إذا كنا نريد أن نستخلص نتيجة موثوق بها عن الأفراد ، وقد تدفعنا اعتبارات عملية مثل التكلفة المادية إلى استخدام اختبار أقل ثباتاً .

ويمكن القول أن معامل الثبات المناسب هو(0.70) فأكثر ويعد معامل الثبات مرتفعاً إذا بلغ (0.80)) فأكثر ومتوسطاً إذا تراوح بين (0.60_0.70)، ومنخفضاً إذا كان أقل من ذلك .وقد يذكر البعض أن معامل الثبات جيد لأنه دال ، والحقيقة أن مستوى الدلالة لمعامل الارتباط ليس محكاً لإقرار الثبات من عدمه ، لأن مستوى الدلالة يعتمد على حجم العينة ، فقد يكون معامل الارتباط(0.90) وغير دال إذا كانت العينة خمسة أفراد ، وقد يكون معامل الارتباط (0.20) ودالاً إذا كان حجم العينة(100) فرداً ، وعليه لا نستخدم مستوى الدلالة في قبول معامل الثبات ، إلا أن مستوى الدلالة هام جداً في حالة حساب معامل الصدق . (علي عودة محمد 2012 ص202).

والخلاصة أن الاختبار الذي ينخفض ثباته جداً(خطأ قياس كبير) سوف ينتج لنا معامل ثبات قريب من الصفر بينما الاختبار الثابت تماماً حيث لا يوجد خطأ قياس ينتج لنا معامل ثبات يساوى واحد صحيح وبذلك فإن المدى الممكن لمعامل الثبات يقع بين صفر ، والواحد الصحيح . وفى الواقع العملي فإن كل الاختبارات تنتج لنا معاملات ثبات تقع ما بين الصفر والواحد الصحيح ولكن كلما كانت قيمة هذا المعامل أقرب إلى الواحد الصحيح كان ذلك أفضل .

ولحساب معامل الثبات علينا أن ننشطReliability Analysis : يظهر الصندوق التالي:

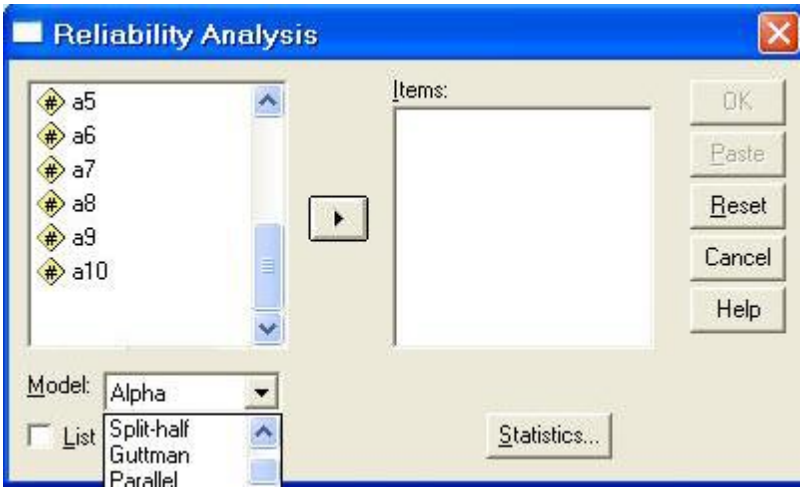


ويحتوى هذا الصندوق ما يلي:

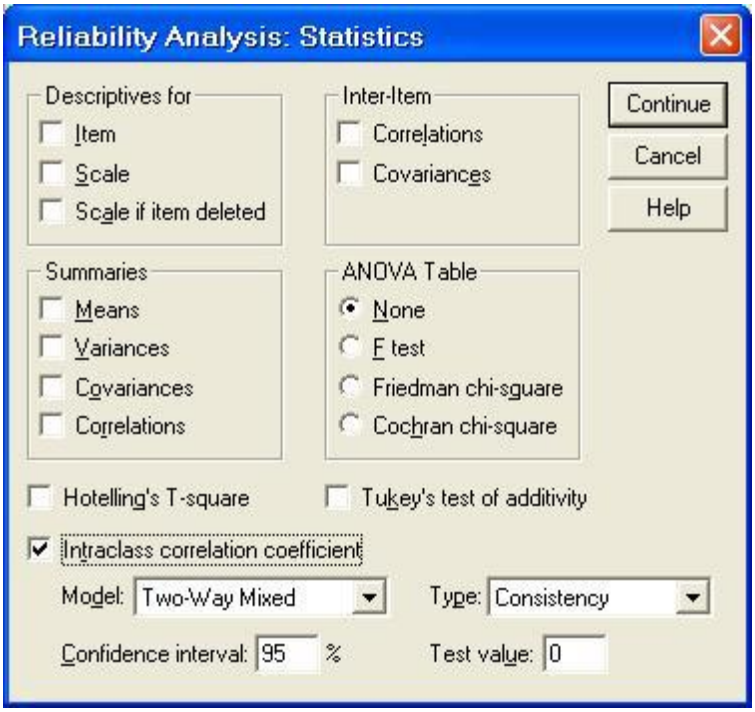
_الجزء الأيسر يحتوى جميع المتغيرات التي تم إدخالها إلى البرنامج .

_الجزء الأيمن **Items** وننقل إليه المتغيرات التي نريد إجراء تحليلات لها وذلك من خلال زر التنشيط بين الجزأين.

النموذج**Model** ، وعلينا باختيار نموذج أو معادلة لحساب الثبات بما تتناسب ونوعية بيانات البحث ، حيث يوجد أكثر من معادلة وهى، التجزئة النصفية **Split – Half**،ألفا**Alpha** جتمان **Guttman**... الخ وتظهر في الصندوق التالي :



الإحصاء **Statistics** وعند تنشيطه يظهر الصندوق التالي :



وينقسم هذا الصندوق إلى ثلاثة أجزاء :

Descriptive for الوصف من خلال: **Inter – Item بين المفردات (البينية)**

Item المفردة **Correlations الارتباطات**

Scale المقياس **Covariances التغايرات**

Scale if item deleted المقياس بعد حذف درجة المفردة

ملخصات Summaries **ANOVA Table جدول تحليل التباين**

Means المتوسطات **Non لا**

Variances التباينات **F Test اختبارف**

Covariance التغايرات **Friedman Chi – Square ل فريدمان**

Correlations الارتباطات **Cochran Chi – Square ل كوجران**

Hotellings T Square ت 2 ل هوتلنج **Tukeys test of additivity اختبار توكى**

Interclass correlation Coefficient معاملات الارتباط الداخلية

Confidence Interval فترات الثقة **Test Value قيمة الاختبار**

مقاييس اختبار الفروض:(تصميم القياسات المتكررة)

أ- تحليل التباين أحادى الاتجاه للقياسات المتكررةANOVA one – Factor Experiment With Repeated Measurements من خلال خيار الثباتReliability :

يستخدم بعض الباحثين مجموعة واحدة من الأفراد في معالجات تجريبية مختلفة على التوالي وعقب الانتهاء من كل معالجة يجرى قياساً معيناً على كل فرد ، لذلك يسمى هذا التصميم **" تصميم القياسات المتكررة "** ، غير أن هناك بعض الضوابط التي ينبغي أن يراعيها الباحث إذا أراد استخدام هذا التصميم. إذ ينبغي مراعاة أن يكون تأثير جميع المعالجات التجريبية المستخدمة تأثيرات مؤقتة بحيث لا تؤثر إحدى هذه المعالجات في المعالجة التالية ، وأن يتحقق الباحث من عدم انتقال أثر معالجة إلى أخرى ، أو من قياس إلى آخر .

وهذا التصميم يعد من أكثر التصميمات ضبطًا لمتغيرات متعددة نظراً لاستخدام نفس مجموعة الأفراد بشرط مراعاة الأمور السابقة ، كما انه أكثر ملاءمة من الناحية العملية إذ أنه يحتاج إلى عدد أقل من الأفراد مما يقلل الوقت والجهد الذي يبذله الباحث .

وفى هذا التصميم نجد أن قيمة (ف) التي نحصل عليها في تحليل التباين للمجموعات المستقلة من قسمة التباين بين المجموعات على التباين داخل المجموعات تصبح في تحليل التباين للمجموعات المترابطة من قسمة التباين بين المجموعات على تباين التفاعل .

و على اعتبار الصفوف هي الأفراد **Rows are Individuals** و الأعمدة معالجات (قياسات تكرار تطبيق) **Columns are Treatments** ، ومع توفر الشروط التالية :

1_ وجود درجة لكل فرد في القياسات المختلفة.

2_ أن يكون توزيع الدرجات للظاهرة في المجتمع الأصلي اعتدالياً اختيرت منه عينة البحث عشوائياً وشكل توزيع الدرجات في كل معالجة طبيعي .

3_ تجانس تباين درجات المعالجات المختلفة (باستخدام أحد أساليب الكشف عن التجانس)

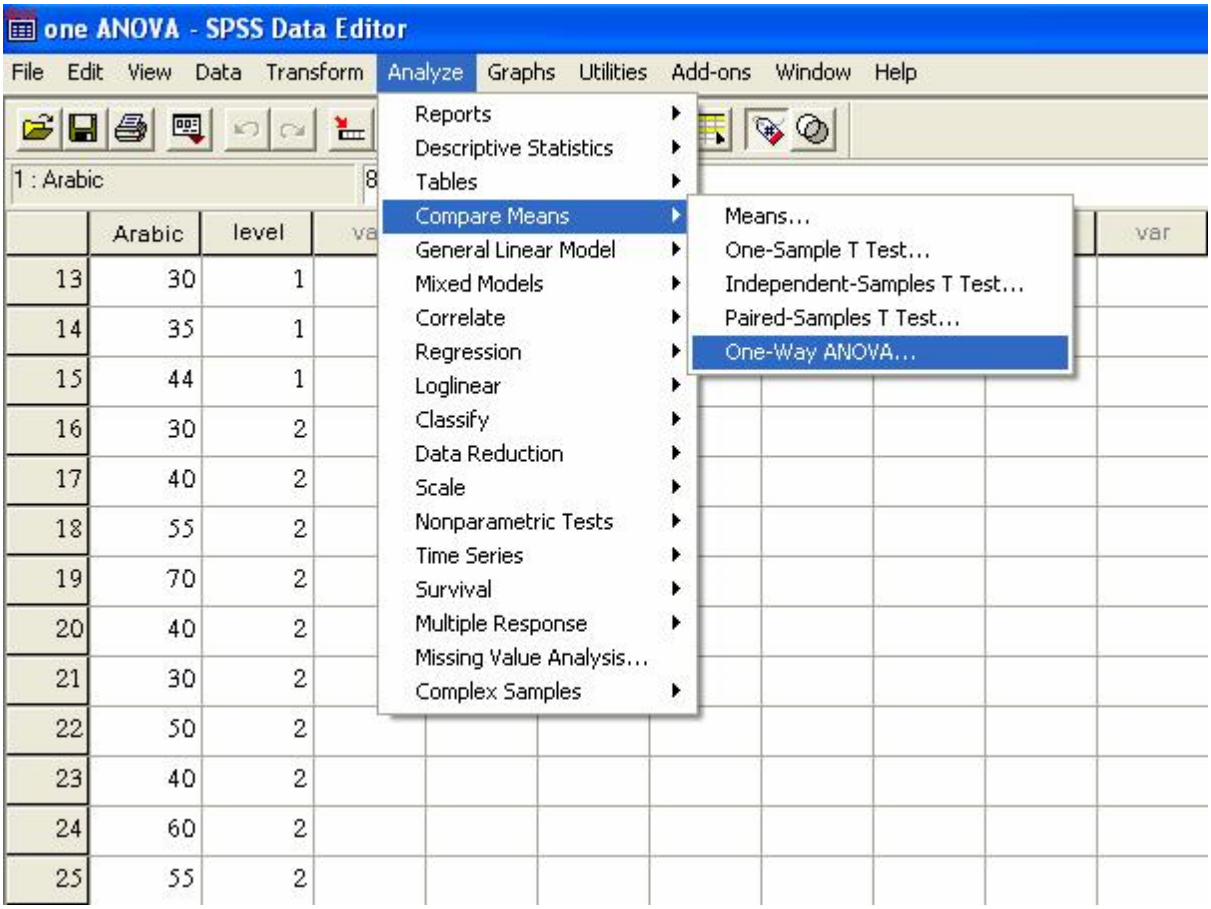
يعني أننا نريد إختبار الفرضيات التالية :

الفرضية الصفرية: لا يوجد فرق معنوي (حقيقي) بين متوسط المجتمعات.

الفرضية البديلة: يوجد فرق معنوي بين متوسط المجتمعات (أي أنه يوجد اختلاف بين متوسط مجتمعين على الأقل. وللقيام باختبار تحليل التباين الأحادي يجب تحقيق الشروط التالية قبل إجراؤه :كل فئة من المتغير النوعي حسب المتغير المتصل لها توزيع طبيعي .فئات المجتمع المختلفة متجانسة أي أن تباينها متساوي .المشاهدات أو القيم مستقلة . ونحن لسنا بصدد الآن دراسة هذه الشروط لذا سنعتبر أن الشروط السابقة متحققة .مثال:إذا كان لديك الجدول التالي الذي يمثل توزيع درجات الطلاب في مقرر اللغة العربية حسب السنة (المستوى) الدراسي.

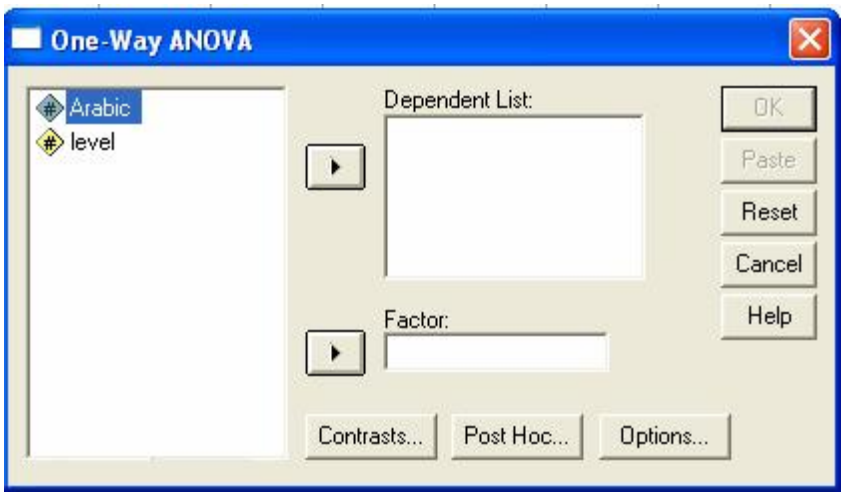
ونريد مقارنة متوسطات درجات الطلبة في مادة اللغة العربية Arabic حسب متغير المستوى الدراسي Level عند مستوى الدلالة α05.0 = ؟

الحل :الفرضية الصفرية: لا يوجد فروق في متوسطات درجات الطلاب في مقرر اللغة العربية حسب متغير المستوى الدراسي عند مستوى الدلالة α05.0 = ؟ وللإجابة على الفرضية نقوم بالخطوات التالية: 1)نعرف متغيرين Arabic) و Levelبحيث يكون كود متغير Level هو 1 = المستوى الاول، 2=المستوى الثاني، 3 =المستوى الثالث، 4 =المستوى الرابع . 2)ندخل البيانات بحيث تكون درجات الطلاب في متغير Arabic) في عمود مستقل ومن ثم ندخل بيانات المميز في المتغير Level)بحيث يكون كل درجة ومصدرها في نفس السطر كما في الشكل التالي:

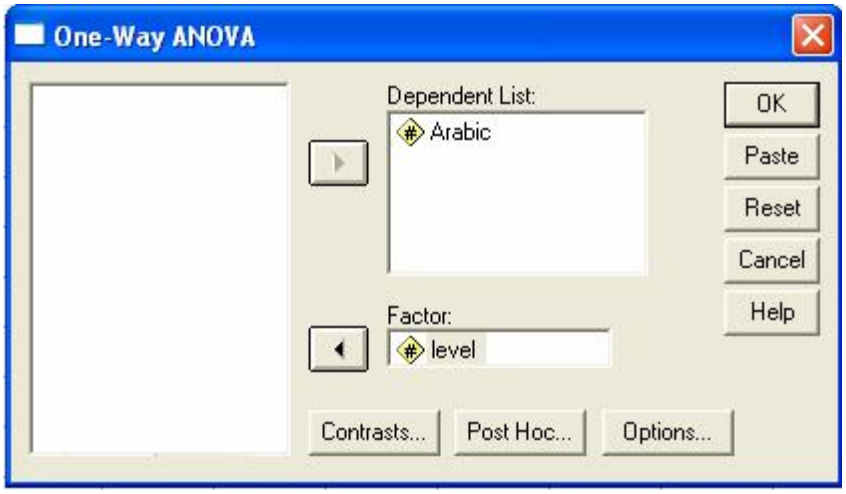


3.من خلال قائمة التحليل اختار القائمة الفرعية) مقارنة المتوسطات*Analyze* (ومن ثم اختار) تحليل التباين الاحادي*Compare Mean*، *One Way ANOVA* فسيظهر المربع الحواري التالي:

4.يظهر المربع الحواري التالي



5.ننقل متغير الدرجات إلى خانة المتغير التابع(*Arabic*) و المتغير *Factor* إلى خانة(*Level*)



6.الآن أصبحنا جاهزين لتنفيذ الاختبار فنضغط على الناتج التالي :،Ok

ناتج اختبار تحليل التباين الأحادي

Oneway

ANOVA

Arabic

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	3772.336	3	1257.445	3.959	.012
Within Groups	18105.861	57	317.647		
Total	21878.197	60			

النتائج:

يتكون ناتج جدول تحليل التباين الأحادي من جدول واحد فقط وهو جدول تحليل التباين الأحادي وهو مشابه تمامًا للجدول الذي درسناه في الجانب النظري من المقرر.

فيبين الجدول أعلاه ما يلي:

-**مجموع المربعات** :يبين الجدول أعلاه أن مجموع مربعات طرق المعالجة هي3772.336و أن مجموع مربعات الخطأ 18105.861وأن مجموع المربعات الكلي هو21878.197

-**درجات الحرية** :يبين الجدول أن درجة حرية طرق المعالجة هي 3 (عدد العينات ناقصًا)و ان درجة حرية الخطأ هي 57 و مجموع درجات الحرية هي 60

-**معدل المربعات** :يبين الجدول أعلاه أن معدل مربعات طرق المعالجة و الخطأ على التوالي هو 1257.445 و317.647

-**قيمة اختبار تحليل التباين الأحادي** هي3.959

-وأخيرً يبين الجدول أن قيمة مستوى دلالة الاختبار هي2 *Sig.= 0.01* هي أقل من مستوى دلالة الفرضية الصفرية

0.05 = αو بالتالي نرفض الفرضية الصفرية، ونستنتج أن هناك فروق ذات دلالة إحصائية بين متوسطات درجات الطلاب حسب المستوى الدراسي أي نرفض الفرضية الصفرية.

ب- كا 2 لفريدمانFriedman Chi- Square :

تواجه الباحث في كثير من البحوث التربوية والنفسية والاجتماعية حالات تكون فيها لديه عدة مواقف تجريبية أو عدد من الاختبارات ، ويتطلب أن تتحدد أهمية أو رتبة كل منها بالنسبة لكل فرد من أفراد العينة ٠ كأن تكون هناك أربع مهن مختلفة ويطلب من كل فرد في العينة أن يبدى رأيه بأفضلية كل مهنة ، أو بمعنى آخر أن يعطى الترتيب الأول للمهنة التي يفضلها أكثر من غيرها ، والترتيب الثاني للمهنة التالية ... وهكذا . حتى يعطى الترتيب الرابع للمهنة التي يفضلها أقل من غيرها . أو كمثال آخر يطلب من مجموعة من الطلاب أن يوضحوا المادة الدراسية التي يحبونها أكثر من غيرها من المواد الثلاث أو الأربع الأخرى ، ثم تعطى رتبة لكل مادة دراسية من هذه المواد الأربع ، في مثل هذه الحالات ولأجل المقارنة بين المواقف التجريبية المتنوعة أو الاختيارات المتعددة يستخدم كا 2 لفريدمان .

ج- اختبار كا 2 لكوكران لعدة عينات مترابطةCochran Chi – Square

إذا استخدم الباحث في دراسته عينات مرتبطة إما عن طريق المزاوجة أو بتكرار القياس عدة مرات على مجموعة واحدة من الأفراد ، وكانت البيانات من المستوى الأسمى ، فإنه يمكنه المقارنة بين هذه العينات المرتبطة باستخدام اختبار كوكران ، ويتطلب هذا الاختبار أن يكون المتغير الاسمي ثنائي التصنيف مثل الاستجابة " نعم – لا " أو " موافق – غير موافق " ، وأن تكون العينات منتقاة عشوائيا .

ويعد اختبار كوكران من الاختبارات القليلة التي يمكن استخدامها في تقييم الفروق بين عدة عينات مرتبطة اعتماداً على كا 2 حيث أن توزيع القيم الناتجة يقترب من توزيع كا 2، كما أن هذا الاختبار يتميز بأن قوته تتراوح بين(80%_85%) فيما يتعلق برفض الفرض الصفري، ولكن لا يجوز استخدامه إذا كان عدد أفراد المجموعات المختلفة أقل من(5). (العيسوي، عبد الرحمن 1997 ص190).

مثال :معرفة ما إذا هناك فوارق في الكفاءة في التسيير وهل يمكن إرجاعها لجنس المسير؟ ويمكن طرح المشكل على النحو الآتي :هل يختلف مستوى الكفاءة في التسيير تبعا للجنس؟ (سعد عبد الرحمن، 1998 ، ص ص78_89).

4. الصدق:

يقصد بالصدق أن الاختبار يقيس ما يجب قياسه ، إن الاختبار الصادق يتنبؤ بما يجب التنبؤ به ، و قد يكون للاختبار أنواع عديدة من الصدق ، فمثلا : قد يتصف الاختبار بدرجة عالية من الصدق للتنبؤ بالنجاح المدرسي في دروس الأدب بينما تنخفض درجة صدقه في التنبؤ بالنجاح في دروس الرياضيات .

فمن الأهمية الكبيرة أن يتحدد بدقة ما يقيسه الاختبار ، هو الاستخدام الغير مناسب للاختبارات قد يؤدي إلى تنبؤات غير دقيقة و من ثم إلى التوجيه و الإرشاد الخاطئ .

يعد الصدق من أهم المفاهيم التي يجب وضعها بعين الاعتبار عند إعداد أو اختيار أدوات الدراسة ، و يعرف هذا الأخير أنه عملية جمع أدلة و شواهد تدعم الاستدلالات و الاستنتاجات التي توصل إليها الباحث من خلال عملية جمع البيانات ، و هذا يعني أن المفهوم الامبريقي للصدق يعني أنه استدلال أو استنتاج خاص باستخدامات الأداة و ليس الأداة ذاتها و يمكن حصر الأدلة الخاصة بالصدق فيما يلي:

_ أدلة تتعلق بصدق المحتوى : و هي أدلة متعلقة بطبيعة المحتوى المتضمن الموجود في الأداة ومدى ملاءمته للخاصية المقاسة .

_ أدلة تتعلق بصدق المحك : و هي أدلة تعنى بالمقارنة بين النتائج المتحصل عليها عن طريق الأداة و أداة أخرى مشابهة لها من حيث النتائج.

_ أدلة تتعلق بصدق التكوين : و هو يشير إلى طبيعة التكوين النفسي أو الخاصية المقاسة المراد قياسها عن طريق الأداة ، أي مدى شرح و تفسير هذا التكوين للفروق الفردية بين الأفراد في السلوك أو الأداء.

1.4 أنواع الصدق:

أ _ صدق المحتوى : يطلق على هذا النوع من الصدق أحيانا بالصدق المنطقي ، أو الصدق بحكم التعريف ، أي تعريف المفهوم أو المتغير محل القياس ، و يطلق عليه أيضا صدق العينة أي عينة السلوك المقاس ، و هناك هدفان ينبغي الوصول إليهما لتحقيق صدق المحتوى هما:

الهدف الأول: أن تكون الخاصية المقاسة محددة بدقة في مجموعة من البنود .

الهدف الثاني: أن تمثل البنود المجالات الفرعية للخاصية المقاسة وأبعاد الخاصية ، و كذلك التوازن بين هذه المجالات، و أن تقيس البنود ما أعدت لقياسه .

ب_ الصدق المرتبط بمحك : يطلق على هذا النوع من الصدق اسم الصدق الواقعي أو الصدق العملي أو الصدق التجريبي، ويقصد به مجموع الإجراءات التي نتمكن من خلالها من حساب الارتباط بين درجات الاختبار و درجات محك خارجي مستقل .

ج_ صدق التكوين (الصدق الفرضي)يعد صدق التكوين الفرضي الدليل الأكثر شمولا من بين أدلة الصدق الثلاثة ، و لا يوجد دليل واحد يمكنه أن يغطي جميع جوانب صدق التكوين ، بل توجد ثلاث خطوات للحصول على أدلة لصدق التكوين هي :

تعريف المتغير المراد قياسه تعريفا واضحا : يجب أن تكون الفروض التي تم وضعها على أساس النظرية الخاصة بالمتغير المراد قياسه ، فروض توضح إمكانية التمييز بين من لديهم الكثير من المتغير المراد قياسه و ممن لديهم القليل من المتغير (أي بين المرتفعين والمنخفضين في المتغير المقاس)

د _ الصدق التمييزي : تقوم هذه الطريقة على أحد مفاهيم الصدق ، و هو قدرة الأداة على التمييز بين طرفي الخاصية التي يقيسها ، و غالبا ما تستخدم اختبارات الفروق للتأكد منه Welcoxon test-t ,whitney-Mann,

2.4 المقارنة بين الصدق و الثبات :

من الضروري أن نشير إلى أن الاختبار قد يتصف بدرجة عالية من الثبات و لكنه قد يكون غير صادق ، و يمكن أن يحدث هذا عندما يقيس الاختبار بعض الخصائص بصورة متسقة، إلا انه لا يقيس ما صمم لقياسه. و جدير بالذكر أن الاختبار الصادق عادة ما يتصف بالثبات (فإذا كان الاختبار يقيس ما يقصد قياسه فانه يجب أن يتصف بدرجة عالية من الاتساق و بخلاف ذلك فان الاختبار لن يستطيع التنبؤ بصورة دقيقة). (العيسوي، عبد الرحمن 1997 ص194).

3.4 العلاقة بين ثبات الاختبار وصدقه :

كلما كان الاختبار ثابتا كلما دل ذلك على صدقه و الدليل على ذلك الصدق الذاتي ، و الذي يعرف على أنه جذر معامل الثبات ، وغالبا ليس كل اختبار صادق هو اختبار ثابت .سعد عبد الرحمان (2008) ص: 213

1.3.4 التقنين :

يشير التقنين إلى الكيفية التي يطبق بها الاختبار حيث يجب أن تتبع كل جوانب إجراء الاختبار التي تتضمن تطبيق الاختبار و إعطاء الدرجات وتقويم النتائج بنفس الأسلوب في كل مرة يعطى فيها الاختبار و إذا تقنن طريقة إجراء الاختبار فان الفروق في أداء المفحوصين على الاختبار قد يكون نتيجة للتباين في أساليب إجرائه ولا تكون مؤثرا دقيقا للاختلافات بين الحالات.

2.3.4 المعايير:

يتمثل احد أساليب تقنين الاختبار في بناء المعايير ، وهي عبارة عن دراجات تم الحصول عليها من مجموعة من الناس طبق عليهم الاختبار . و بمجرد تحديد المعايير فانه يمكن مقارنة أداء الأفراد الآخرين الذين يطبق عليهم نفس الاختبار بهذه المعايير.

أ_ المعايير في مقابل المحاكاة :

غالبا ما تساعد المعايير في تفسير نتائج الاختبارات النفسية وإلا أنها في بعض الأحيان قد تكون خادعة مضللة و من المستحسن في مثل هذه الحالات استخدام المحك او المعيار المطلق لتقويم أداء المفحوصين.

ب_ المعايير و الصيغ النمطية :

نظرا لان المعايير تبنى على أساس نتائج الاختبار و غالبا ما تصاغ في صورة إحصائية فقد يتخذها القارئ الساذج كمؤثر جوهري حقيقي و من ثم يمكن أن يساء تفسيرها . فمثلا قد ينظر الشخص إلى المعيار و يتبع هذه السلسلة من الاستدلال:

1_ إذا كانت هذه هي الدرجة المتوسطة أو النمطية لمجموعة من الأفراد.

2 _ شخص ما عضو في تلك المجموع

3_ لابد إذن أن تنطبق هذه الدرجة أيضا على ذلك الشخص . وهنا نجد انه على الرغم من أن نتائج الاختبار، قد تنطبق على هذا الشخص إلا أن المعيار المستخدم لا يقدم أي ضمان على أنها كذلك.

الموضوعية

يقصد بالموضوعية ألا تلعب تحيزات أو أهواء القائم بالملاحظات أو القياس دورا في الملاحظة أو القياس (فكل المقاييس تصاغ في صورة محايدة أو منزهة عن الهوى) و الملاحظون الذين يسمحون لتحيز أنهم بالتأثير في حكمهم أو قياسهم يوصفون بالذاتية و اللا موضوعية. (العيسوي، عبد الرحمن 1997 ص195).

5. خاتمة :

إن الخصائص السيكومترية (الميترولوجية)لأدوات القياس النفسي هي بمثابة الأدلة على صدق النتائج المتوصل إليها و بالتالي توفر له الإمكانية لتعميم النتائج وفق بعض الشروط المتفق عليها ، و بالتالي فإذا اعتمد الباحث على أدوات قياس سواء كانت مستعارة من أدوات تم بناءها و تم التأكد من خصائصها السيكومترية أو أدوات تم بناءها من طرف الباحث ، و لم يتم التأكد من خصائصها السيكومترية ، فهذا يؤدي بالباحث إلى التجوال في حلقة مفرغة يميزها الغموض فيما إذا كانت نتائجه المتوصل إليها صادقة أم لا ، و لهذا كانت مرحلة التأكد من الخصائص السيكومترية لأدوات من الإجراءات المنهجية الأكثر أهمية في البحوث النفسية و هذا ما غاب للأسف في بعض المذكرات و الأطروحات الجامعية ، و لهذا يمكن وضع توصية للباحثين في الجامعة الجزائرية في ميدان علم النفس ، و هي يجب التأكد من الخصائص السيكومترية لأدوات البحث قبل وضع النتائج الخاصة بالبحث.

6. المراجـــع :

- أحمد، محمد عبد السلام. (1960). القياس النفسي والتربوي: التعريف بالقياس ومفاهيمه وأدواته، بناء المقاييس ومميزاتها والقياس التربوي. القاهرة: مكتبة النهضة المصرية.

- صالح الدين محمود عالم (2006). الاختبارات والمقاييس التربوية والنفسية.ط1. الأردن: دار الفكر.

- علي عودة محمد(2012). مناهج البحث في التربية وعلم النفس.ط1. العراق: مكتبة عدنان للطبع والنشر والتوزيع.

- بشير معمريه(2007). القياس النفسي وتصميم أدواته. ط2. الجزائر: منشورات الحبر.

- صالح مراد وفوزية هادي(2002). طرائق البحث العلمي _تصميماتها وإجراءاتها. الكويت: دار الكتب الحديثة.

- مقدم عبد الحفيظ (1993). الإحصاء والقياس النفسي والتربوي مع نماذج من المقاييس والاختبارات. الجزائر: ديوان المطبوعات الجامعية.

- عبد الكريم بوحفص(2011). أسس ومناهج البحث في علم النفس. الجزائر: ديوان المطبوعات الجامعية.

- سعد عبد الرحمان (2008). القياس النفسي النظرية والتطبيق. ط5. القاهرة: هبة النيل العربية للنشر والتوزيع.

- العيسوي، عبد الرحمن (1997). أصول البحث السيكولوجي. لبنان: دار الراتب الجامعية.