

الفصل الثالث

خطوات بناء الاختبار أو المقياس

- خطة تصميم الاختبار
- تحديد الهدف أو الأهداف الرئيسة للاختبار
- تقرير محك أو معيار الدرجة
- ترجمة المفاهيم والأهداف إلى خصائص محددة
- تصميم فقرات مناسبة تعبر عن هذه الخصائص
- اختبار مستوى الصدق والصعوبة للفقرات
- فعالية المشتتات
- إعداد الاختبار للاستخدام
- تقنين الاختبار

Obeykhan.com

خطوات بناء الاختبار أو المقياس

يخضع تصميم وإعداد الاختبار الجيد لعدد من الاعتبارات والقواعد المنهجية، كما تتطلب فنية تصميم الاختبار دراية واسعة وعميقة بالأساليب المختلفة ومدى مناسبة كل منهما للأغراض المختلفة. وأساليب تصميم وانتخاب فقرات اختبار صادقة وثابتة ومميزة هي أساليب عملية وعلمية معاً. ويتضمن الجانب العلمي فيها المعرفة الدقيقة بكيفية وضع الفقرة ومهارة تصميمه مع حساب قدرته التمييزية ومستوى صعوبته وثباته ويأتي الجانب الفني في الربط التكاملي بين هذه الخصائص العلمية والخصائص اللفظية للفقرة، والاعتبارات العلمية التي سيستخدم فيها الاختبار. ولا يعد تصميم وإعداد اختبار جديد هدفاً في حد ذاته، فالباحث لا يلجأ عادة لهذا العمل إلا إذا دعت إليه الحاجة ملحة، وعندما يتضح له أن الاختبارات المتوفرة لا تفي بغرضه، ودون وجود الدواعي الضرورية لتصميم اختبار جديد، يظل استخدام الاختبارات المتوفرة في المجال أو تطويرها أفضل كثيراً للاستفادة من الدراسات والنتائج التي سبق التوصل إليها.

ويمكننا أن نلاحظ أن هناك اتجاهين بارزين في تصميم الاختبارات وهي:

الاتجاه الأول: علمي نظري يهدف إلى تصميم الاختبار وفق إطار نظري محدود، وللإجابة على فروض جيدة الصياغة يقدمها العالم، ويصمم وينتخب بعناية فائقة فقرات الاختبار وفقاً لمدى قربها من فروضه الأساسية.

الاتجاه الثاني: عملي أو فني وهو يرمي لمجرد تطوير اختبارات جديدة لتكون

أداة في يد الممارس أو الأخصائي النفسي بغض النظر عن اعتبارها أداة علمية للبحث.

والاختلاف بين هذين الاتجاهين له تأثيرات هامة على منهجية تصميم الاختبارات، فمصمحو الاختبارات ممن يأخذون بالمنحى العملي يطورون أساليب خاصة بهم تختلف بقدر ما عن تلك التي يستخدمها العالم الذي يصب القدر الأكبر من اهتمامه على صياغة فروضه النظرية. ورغم أن صياغة الفروض جوهرية للغاية

في تصميم الاختبار، إلا أن عمليات صياغة هذه الفروض تختلف تماماً في الحالتين بين العالم والفني. فبالنسبة للعالم تعد مشكلة صياغة الفروض أهم خطوات العمل كلها وأصعبها وهي ما يتطلب منه الأصالة أذ لا بد أن تكون فروضه متجهة إلى تقديم إضافة حقيقية مع اتساقها في الوقت نفسه مع الحقائق القائمة والملاحظات السابقة والتبريرات المنطقية التي يقدمها العالم لفروضه هي التي تميز عمله عن عمل الفني الذي يهدف إلى هدف محدود ويحكم إجراءاته المهارة في إصابة الهدف وتحقيق الأغراض المباشرة ورغم ذلك نستطيع أن نجد المزايا والعيوب في كلا المنحيين فالمنحى العملي الفني سريع النتائج، وتصميم الاختبارات فيه لا يتطلب وقتاً طويلاً، وقد يبدأ من معلومات خاطئة ولكنه يؤدي لنتائج عملية مفيدة ينتهي بمعلومات محدودة ذات قيمة نظرية ضئيلة. بينما المنحى النظري يتطلب وقتاً في إعداد الاختبارات، ووقتاً في تنظيم وتحليل النتائج وقد تكون نتائجه العملية محدودة ولكنه يبدأ من فروض صلبة وينتهي بمعارف هامة ويؤدي في أحيان كثيرة إلى الإسهام في تطوير النظرية.

خطة تصميم الاختبار:

تقسم خطة تصميم الاختبار إلى عدد من الخطوات المحددة وتختلف هذه الخطوات وتصنيفها من باحث لآخر.

آلا أنها تتضمن في كل الحالات في البداية التي تبدأ بها وتنتهي أليها كالآتي:

- ١- تحديد الهدف أو الأهداف الرئيسية من الاختبار.
- ٢- تقرير محك أو معيار الدرجة.
- ٣- ترجمة المفاهيم والأهداف إلى خصائص محددة.
- ٤- تصميم فقرات مناسبة تعبر عن هذه الخصائص
- ٥- اختبار مستوى الصعوبة والصدق للفقرات.
- ٦- فعالية المشتتات (الموهات)
- ٧- إعداد الاختبار للاستخدام.
- ٨- تقنين الاختبار.

١. تحديد الهدف أو الأهداف الرئيسة من الاختبار

أن تحديد الهدف من الاختبار من حيث ميادين استخدامه والمجتمع المراد تطبيقه عليه ، على جانب كبير من الأهمية. فإن كان الهدف من استخدام الاختبار في الحصول على بيانات دقيقة كان لابد من استخدام الطرق الإحصائية والفنية التي تحقق هذا الغرض ، أما إذا كان المطلوب جمع بيانات سريعة أو إذا كان المفروض أن يعهد بالاختبار إلى فئة من المعلمين غير المدربين تدريباً كافياً على استخدام الاختبارات المقننة روعي في تصميم الاختبار سهولة تفسير معاييرها حتى ولو أدت هذه السهولة إلى التضحية ببعض الشيء ببعض الدقة ذلك لأن تعقيد الاختبار من ناحية التعليمات أو التصحيح سيؤدي حتماً إلى أخطاء أكبر.

إذن رغم أننا قد نبدأ من مفهوم علمي مقبول يمكن اختياره بواسطة اختبار معين إلا أن تحديد الهدف من الاختبار يتطلب مزيداً من التفصيل فقد يكون الهدف من اختبار ما تحديد مستويات الأفراد وفقاً لمقدار امتلاكهم للخاصية وقد يكون الهدف من اختبار آخر التمييز بين الأفراد وفقاً لترتيبهم على الخاصية أو قد يتكون الهدف تحديد الصعوبات أو الاضطرابات التي يعاني منها الفرد وبالتالي تصبح مهمة الاختبار تحليل هذه الصعوبات بشكل تفصيلي وتصنيفي. ويختلف كل هدف من هذه الأهداف في طبيعته ، فالأساليب التي تهدف إلى تحديد مستوى الأفراد يكثر استخدامها في المجالات التحصيلية أو في قياس القدرات المعرفية. مما يتطلب تصنيف وتحليل العمليات المختلفة في المادة المحصلة وحيث يمكن استخدام هذا التصنيف كدليل وصفي لنتائج العملية التعليمية ومدى تحقق أهدافها.

٢ - تقرير محك أو معيار الدرجة: تفرض أهداف الاختبار نوع الإطار المرجعي الذي تفسر في ضوءه درجة الفرد وما إذا كان يعد جيداً الأداء أم لا ويمكن التمييز هنا بين نوعين من المحكات:

المحك المرجعي: ويستخدم هذا المحك أساساً في اختبارات التحصيل بقياس التحصيل يتضمن إشارة إلى حدود مقبولة للأداء على متصل يبدأ قطبه الأدنى من لا كفاءة على الإطلاق وينتهي قطبه الأعلى بأداء محكم تماماً وفق تعريفات وحدات

الاختبار مسبقاً التحديد وحتى يمكن تصميم اختبار جديد ذي محك مرجعي فمن الضروري العمل وفق خطوات محددة، من ذلك ضرورة وضع أشكال الأداء المطلوبة بصورة صريحة ومباشرة، وضرورة تحديد محكات قبول الأداء مسبقاً وفقاً لبرنامج تحصيل الموضوع المعين مثلاً، ويجب تحليل الأداء المطلوب لعناصره المختلفة حتى يمكن تقدير جوانب تحصيل المادة أو الخبرة موضوع الاختبار وأخيراً فإن النجاح على الاختبار يتقرر من خلال مقارنة الأداء لا بأداء أفراد آخرين ولكن بمقارنته بحجم المطالب المحددة في الاختبار وما أنجز منها. وسيتم لاحقاً تحديد الخطوات التفصيلية لبناء الاختبارات المحكية المرجع.

٢ - المحك المعياري: تختلف الاختبارات ذات المحك المعياري عن الاختبارات المحكية المرجع في أن الدرجة الخاصة بالفرد لا تقارن وفق حجم سابق التحديد للمطلب موضوع الاختبار، وإنما تقارن بالأداء الخاص ببقية الأفراد، فنحن هنا نهدف للتعرف على الوضع النسبي للفرد بين مجموعة الأفراد، وبهذا تكون الدرجة الخاصة بالفرد ذات قيمة من حيث أنها تفيدنا بمدى جودة هذا الأداء بالنسبة للمجموعة المختبرة ولكنها ليست ذات فائدة من حيث الإشارة إلى مدى جودة الأداء بالنسبة للمطالب الأساسية للاختبار أو المجال موضوع القياس. ويترتب على تقرير هذا المحك المعياري أن نتجه في تصميمنا للاختبارات إلى أن تكون الفقرات ذات قدرة على تفسير مدى الدرجات التي يمكن أن نحصل عليها من أية مجموعة من الأفراد.

معنى هذا أن الدرجة على المقياس في حد ذاته لا معنى لها على الإطلاق إلا إذا نسبت إلى إحدى المحكين: المحك المرجعي بأن تنسب إلى الأداء نفسه بوصفه مرجعاً للدرجة، أو المحك المعياري بأن تنسب إلى درجات بقية الأفراد بوصفهم معياراً لهذه الدرجة.

٢- ترجمة المفاهيم والأهداف إلى خصائص محددة:

المفاهيم العلمية كالذكاء والانبساط والتطرف والاجتماعية تجريدات لخصائص مفترضة في الأشياء ونستدل على هذه الخصائص من وقائع سلوكية

محددة سواء في شكل أفكار معبر عنها أو حلول لمشكلات أو استجابة لمنبهات محددة أو أساليب تعامل مع البيئة أو صفات شخصية تتسم بالاستقرار. ونقوم بترجمة هذه المفاهيم إلى خصائص محددة بصورة تسمح بصياغتها في وحدات معيارية للقياس ففي مفهوم مثل الطلاقة يمكن ترجمته الى أعمال محددة إذا وضع له التعريف الإجرائي المناسب مثل ((الطلاقة هي القدرة على إنتاج أكبر عدد من الأفكار ذات الدلالة استجابة لمنبه معين)).

إذ يتيح هذا التعريف الإجرائي وضع المنبهات لاستدعاء استجابات معينة هي التي يقاس بواسطة كمها واستيفائها لمحك الدلالة طلاقة المفحوص، ولم يكن متاحاً بالمثل قياس مفهوم الذكاء دون توفر التعريف الإجرائي الذي وضعه بينيه وترجمه إلى أعمال محددة يقوم بها الطفل. فالمطلوب في هذه الخطوة هو أن نحول الأهداف إلى سلسلة من الخطوات والأعمال لنتمكن من كتابة الفقرات المناسبة التي يمكن أن يقدم تقديراً صادقاً للجوانب التي تقوم بقياسها ويتم ذلك بتحديد عينه مقننة من هذا السلوك والذي يتعين أن يمثل بشكل جيد المفهوم الذي نصمم من أجله المقياس من خلال خطوات محددة وهي: -

١ - وصف السلوك المطلوب قياسه: فإذا كان المفهوم هنا هو ((القدرة

الحسابية)) فإن وصف السلوك يكون كالآتي:

يتضمن السلوك المطلوب قياسه القيام بحل عدد من المشكلات الحسابية البسيطة دون أخطاء ويمكن إجراء العمليات الحسابية جميعها إما داخلياً (عقلياً) أو باستخدام الورقة والقلم. ويقصد (بالعمليات الحسابية البسيطة) الجمع والطرح والضرب والقسمة والنسب المئوية والكسور ولأهمية السرعة في المطلوب، والإجابة الصحيحة هي المطلوبة ولذلك يمنح المفحوص وقتاً متسعاً للإجابة.

٢- تحليل السلوك المطلوب قياسه: ويقصد به تصنيف هذا السلوك للتعرف على مكانه بين أشكال السلوك الأخرى والإشارة إلى طبيعته والى عموميته وقابليته للتنبؤ. فإذا طبقنا مثل هذا التحليل على ((القدرة الحسابية)) فيمكن أن نصل إلى أن القيام بعمليات حسابية بسيطة يتأثر بكل من الاستعداد والتدريب أو العادات المستقرة لفترة طويلة والأفراد البارزين في هذا المجال لا يحتاجون لبذل مجهود كبير

لتجنب الأخطاء ، وحساباتهم دائماً صحيحة بصورة طبيعية ، والأفراد الذين يفتقدون هذه الخاصية يقعون بالخطأ بسهولة وإجاباتهم الصحيحة تعود في حقيقة الأمر إلى مراجعتهم الدقيقة وإعادة المراجعة ، والواقع أن الأشخاص أصحاب الدرجة المرتفعة يتمتعون بما يمكن اعتباره القدرة على المراجعة الآلية.

٣ - تقرير خصائص الفقرات: ويقصد بها التقدم نحو وصف نوع الفقرات التي يمكن أن تؤدي إلى صدق قياس وتقويم السلوك المعين فإذا تقدمنا من المثال الخاص بالقدرة الحسابية فسنجد أن قياس السلوك الحسابي يتطلب أن يتضمن الاختبار عدداً كبيراً من المشكلات البسيطة نسبياً والتي تتضمن عمليات جمع وطرح وضرب وقسمة وكسور عشرية ونسب.

ويتعين أن لا تتضمن الفقرات مشكلات صعبة أو بها خدعة يجب معرفتها للوصول إلى الحل ويجب أن تكون التعليمات الموجهة للمفحوص واضحة ومتضمنة وأن يعمل المفحوص حسب سرعته الخاصة ، ولكن دون إنفاق وقت كبير في المشكلة أو فقرة ما أكبر مما يجب.

٢ - تصميم فقرات مناسبة تعبر عن هذه الخصائص :

هناك أكثر من شكل وأسلوب يمكن أن تصاغ به فقرات الاختبار. وتتطلب كتابة فقرات مناسبة لاختبار جديد ، تحليل الفقرات تحليلاً كيفياً من حيث شكل الفقرات ومن حيث مضمونها إلى فقرات جيدة تقيس بشكل صادق السلوك المعين كما يؤدي هذا التحليل الكيفي إلى توضيح أهمية مراعاة شروط كتابة الفقرات الجيدة.

كما يتطلب الأمر تحليلاً كميّاً يهدف إلى تقدير مستوى صعوبة الفقرات ، وانتخاب الفقرات ذات الصعوبة المناسبة لأهداف الاختبار. ويوضح هذا النوع من التحليل علاقة صعوبة الفقرات بثباتها ، إذ أن صدق وثبات الاختبار يعتمد أساساً على خصائص الفقرات ويمكن التوصل إلى صدق وثبات مرتفعين للاختبار منذ البداية من خلال إجراءات تصميم الفقرات.

التحليل الكيفي:

تصنف الاختبارات من حيث بنيتها إلى اختبارات محددة البنية أو الاختبارات الموضوعية التي تتميز بأن المفحوص لا يقدم فيها إجابة للفقرة بل ينتخب إحدى الإجابات من بين عدد من الاختيارات التي يقدمها الاختبار، وتعتبر موضوعية حيث أن تصحيح الاختبار يتم بدون تدخل من جانب الفاحص. كما أن لهذه الاختبارات القائمة على أساس الاختيار أنواع نذكر منها:

١- اختبارات الصواب والخطأ

٢- اختبارات المطابقة

٣- اختبارات الاختيار من المتعدد

وهناك نوع آخر من الاختبارات والتي تتضمن فقرات غامضة أو لاتستدعي الاستجابات الموضوعية في التصحيح بل تحتمل تأويلات وتفسيرات مختلفة قد تختلف من باحث لآخر أو وفق مبدأ للتصحيح أو آخر، وهذه الاختبارات تسمى الاختبارات الإسقاطية. في فئة الثالثة نجد الاختبارات الإنتاجية والتي تتطلب من المفحوص ((تقديم)) استجابات قابلة لتصنيفات واضحة المعالم محددة بشكل مسبق مثلما نجد في الاختبارات الإبداعية أو بعض المقاييس الفردية للذكاء كاختبار الفهم أو المعلومات.

وقبل كتابة فقرات الاختبار على واضع الاختبارات أن يفكر في طول الاختبار وعدد فقراته، وإن أهداف الاختبار هي التي تحدد طوله فإن كان لأجل مقارنة سريعة فلا داعي للاختبار الطويل. أما إذا كان لأجل جمع بيانات دقيقة فمن الواجب أن يكون الاختبار طويلاً ومتعدد الأسئلة. وهناك عدد من القواعد العامة في كتابة الفقرات العامة في الوصول إلى مستوى صدق جيد للاختبار منها ما يلي:

١- احتواء الفقرة الواحدة على فكرة واحدة فقط والابتعاد عن الأسئلة المزدوجة

٢- الابتعاد عن استخدام الكلمات الغريبة أو الكلمات التي تحمل أكثر من

معنى واحد.

٣- محاولة استخدام الصيغة الإيجابية للفقرات كلما أمكن ذلك.

- ٤- تجنب استخدام الإطلاق في العبارات مثل دائماً ابداً ، حتماً وفي كل مكان.
- ٥- البساطة والسهولة وعدم الغموض.
- ٦- الموضوعية وإمكانية تبويب الإجابات.
- ٧- تحاشي الأسئلة الإيحائية.
- ٨- تحاشي الأسئلة التطفلية والتي تسبب إحراجاً.

يضاف إلى ذلك عدد من القواعد الأساسية في تصميم الفقرات تتعلق بتصميم نظرية القياس والأسس الرياضية لمعنى الدرجة على المقياس. فمن الضروري أن يسعى مصمم الاختبار إلى محاولة الاقتراب منها إذ يبدو بعضها عسيراً في حالات كثيرة لهذا يحسن اعتبارها مؤشرات أساسية لتصميم الفقرات الجيدة للاختبار وتتلخص هذه القواعد في الآتي:

١- أن يتضمن الاختبار فقرات تشكل بصورة جيدة عينة ممثلة للمجال الذي تهدف لقياسه، أي أن تكون هذه العينة للسلوك الذي سبق أن قمنا بتحديدته من خلال تحديد الهدف من الاختبار.

٢- أن تتدرج فقرات الاختبار وفقاً لمحك معين. وتعني هذه الخاصية أن ترتب فقرات الاختبار من حيث الصعوبة بشكل متدرج وبحيث يكون مستوى صعوبتها موحداً بالنسبة لكل الأفراد. فإذا كان للفقرات مثل هذا الترتيب المتدرج فإن أي مفحوص سيتمكن من تقديم أجابه صحيحة لعدد من الفقرات عند حد معين لا يستطيع تجاوزه يقابل مستوى الصعوبة الملائم له، بحيث تعني الدرجة ١٧ مثلاً بالنسبة للفرد أنه أجاب إجابة صحيحة على أول ١٧ فقرة وأجاب إجابة خاطئة على بقية الاختبار. ويؤدي مثل هذا الترتيب للفقرات إلى إلغاء الغموض في تفسير الدرجة على الاختبار بحيث تعني الدرجة ١٧ نفس المعنى لدى كل من حصل عليها وهو أنهم جميعاً أجابوا إجابة صحيحة على الـ ١٧ فقرة الأولى.

٣- إن ترتيب فقرات الاختبار وفق متصل المتغير موضوع القياس بطريقة تجعلنا أمام مقياس مسافات ويلاحظ هنا أن الفقرات التي يتضمنها مثل هذا المقياس تتسم بخاصية التعبير عن فروق متساوية على القدرة بين الفقرات المتجاورة فيعبر الفرق بين الدرجتين ٥٤، ٥٣ عن فرق مماثل بين الدرجتين ٨٦، ٨٥ وهكذا.

٤ - أن يكون الصفر على أول فقرات المقياس تعبيراً جيداً عن كمية صفرية من القدرة التي تقاس فإذا أضفنا الخاصية السابقة لهذه الخاصية يصبح مقياسنا مقياس نسبة يمكن معالجة نتائجه بالعمليات الحسابية الأربعة.

٥ - أن تكون الفقرات بمثابة وحدات قياس ذات معنى واضح، لأن ذلك يؤدي إلى ربط وحدات القياس بالخطأ المعياري للمقياس، وهو ما يوفر ميزة تسمح لمستخدم الاختبار بترجيح أن الفروق بين الأفراد تعود أساساً إلى الخطأ في الأداء وليس لأي عامل آخر. تقودنا كل هذه الاعتبارات والقواعد العامة والخاصة إلى الانتقال للجانب الأساسي التالي في تصميم الفقرات وهو التحليل الكمي الذي يتعلق بدراسة وتحديد وترتيب مستوى صعوبة الفقرات في الاختبار.

٤ - اختبار مستوى الصدق والصعوبة لفقرات الاختبار

التحليل الكمي: هناك اعتباران أوليان يبرزان عند انتخاب فقرات اختبار ما: الأول هو: هل الفقرة صادقة في قياس السمة التي نرغب في قياسها بمعنى هل تمكنا هذه الفقرة من التمييز بين الأشخاص أصحاب القدر الضئيل منها؟ ويجب عادةً على هذا السؤال من خلال الأساليب الإحصائية الخاصة بحساب صدق فقرات الاختبار. أما الاعتبار الثاني فهو هل مستوى صعوبة الفقرة مناسبة للمجموعة من الأفراد الذين سيختبرون بهذا الاختبار؟

١- صدق الفقرات: ويعني مؤشر الصدق حسن قياس الفقرة أو تمييزها في أتساق مع بقية الاختبار أو حسن تنبؤه بالمحك الخارجي. ويمكن أن نلاحظ عدداً من المؤشرات الشائعة للصدق في الاستخدامات المختلفة من ذلك: نسبة المفحوصين الذين يجيبون على الفقرة أجابه مقبولة (تميزية) أو الارتباط بين الفقرة ومحك خارجي أو الارتباط بين الفقرة وبقية الاختبار وهو محك داخلي.

ونادراً ما يلجأ مصمم الاختبار إلى حساب الارتباط بين الفقرة والمحك الخارجي وقليلاً ما يحسب الارتباط بين الفقرة وبقية الاختبار والأسلوب الأكثر شيوعاً هو حساب القدرة التمييزية للاختبار من خلال نسبة المفحوصين الذين يتجاوزونه ودون

إجراء مثل هذا التحليل لصدق الفقرات. ومهما كانت خبرة مصمم الاختبار وموضوعيته فسيظل الاختبار في حاجة إلى فحصه لتقرير هذه الخصائص التمييزية للفقرات. تستخدم المجموعات المتطرفة أو المتعارضة عادة في حساب صدق الفقرات وذلك بأن ننتخب مثلاً الفقرات التي يجيب عليها أعلى ١٠٪ من المجموعة (أ) وأقل ١٠٪ من المجموعة (ب) المتطرفة معها في السمة المقاسة. وبقدر تعارض المجموعتين بقدر حدة التمييز الذي نتوصل إليه. إلا أن استخدام المجموعات الشديدة التطرف بهذه الصورة يؤدي إلى خفض الثبات نتيجة للعدد الصغير من الحالات المستخدمة حيث تكون المجموعتان أ، ب عبارة عن الحالات المتطرفة في عينة واحدة مستخدمة لحساب صدق الفقرات. ويقترح كيلي Kelly استخدام أعلى وأدنى ٢٧٪ من التوزيع باعتبارهما المجموعتين المتطرفتين بشرط إعتدالية التوزيع. بينما يرى كيرتون cureton أن النسبة مقبولة إذا كانت تتراوح بين ٢٧٪ - ٣٣٪ وإذا كان المنحنى أكثر استواءاً من المنحنى الإعتدالي التقليدي. وتضيف أنستازي إلى أن خطأ العينة يصبح كبيراً في حالة العينات الصغيرة. ولهذا يصبح من الأفضل أن لا نحدد هذه النسب بشكل حاسم وتعتبر النسب التي تتراوح بين ٢٥٪، ٣٣٪ وافية بالغرض بصورة مقبولة، ومؤشر التمييز المقبول بصفة عامة باعتباره مؤشراً لصدق الفقرات هو ما توضحه المعادلة الآتية لجونسون Johnson:

يتم حساب معامل التمييز بهذه الطريقة بطرح عدد الأفراد الذين اختاروا الخيار الصحيح للسؤال في المجموعة الدنيا من عدد الأفراد الذين اختاروا الخيار الصحيح للسؤال من المجموعة العليا ثم قسمة الناتج على عدد الأفراد في إحدى المجموعتين فقط.

$$\frac{أع - أد}{ن} = م ع د$$

م ع د = مؤشر الفرق بين أعلى وأدنى المجموعتين

أ ع = عدد الإجابات الصحيحة على الفقرة في المجموعة العليا

أ = عدد الإجابات الصحيحة على الفقرة في المجموعة الدنيا

ن = عدد أفراد العينة في إحدى الفئتين

كما يمكن إيجاد تباين كل فقرة حسب المعادلة الآتية:

تباين الفقرة = نسبة الأفراد الذين أجابوا إجابة صحيحة عن الفقرة \times نسبة

الأفراد الذين أجابوا إجابة خاطئة عن الفقرة

ولكن يعاب على هذه الطريقة أن التوزيع العيني لقيمتها غير معروف وبالتالي لا

نستطيع ان نجيب عن الأسئلة التالية:

- هل قيمة معامل التمييز الناتجة اكبر من الصفر ؟

- ما حجم الفرق بين قيمة معامل التمييز والصفر ؟

- هل الفرق دال إحصائياً ؟

ورغم ذلك استمرت هذه الطريقة تستخدم كواحدة من الطرق العملية لحساب

معامل التمييز وتصلح لاختبارات المدرس لسهولة استخدامها وحسابها وتفسيرها.

ويمكن حساب معامل التمييز باستخدام معامل الارتباط المنصف للسلسلة ويتم

استخدام هذه الطريقة عندما يكون لدينا متغيران أحدهما ثنائي الدرجة ، درجة

كل فقرة: (صفر أو واحد) والآخر فئوي متصل (الدرجة الكلية).

حيث يعتبر معامل الارتباط المنصف للسلسلة أكثر مناسبة للاستخدام بين

سؤال موضوعي والدرجة الكلية على الاختبار أو على محك خارجي مثل تقديرات

المعلمين أو درجات الطلاب في اختبارات أخرى.

ويمكن حساب معامل التمييز باستخدام معامل الارتباط المنصف للسلسلة

الحقيقي.

وتستخدم هذه الطريقة عند تقسيم المتغير المتصل (الدرجات الكلية للاختبار)

إلى قسمين متساويين يشتمل أحدهما على الدرجات العليا والآخر على الدرجات

الدنيا.

حيث ينتج لدينا متغيران أحدهما ثنائي وهو (درجة كل فقرة) والآخر فئوي

متصل (الدرجات الكلية) ولكن تم تقسيمه إلى متغير ثنائي.

$$\text{رب س} = \frac{\text{م س} - \text{م ك}}{\text{ع س}} \times \left[\frac{\text{ص}}{\text{أ}} \right]$$

رب س = معامل الارتباط المنصف للسلسلة الحقيقي
 م س = متوسط الدرجة الكلية في الاختبار للمجموعة التي إجاباتها صحيحة
 على الفقرة. م ك = متوسط الدرجة الكلية في الاختبار (أو المحك) لكامل المجموعة
 ع س = الانحراف المعياري للدرجة الكلية في الاختبار. ص = صعوبة الفقرة
 أ = ارتفاع منحنى التوزيع الاعتمالي عند الدرجة (ز) التي تقسم التوزيع عند
 النسبة المعيارية للصعوبة (ص)
 متى تستخدم الطرق الثلاث السابقة:

- أ- في حالة الأسئلة متوسطة الصعوبة لا يوجد فرق بين الطرق الثلاث وبالتالي
 فإن طريقة مقارنة الأطراف تكون هي الأفضل وذلك لسهولة حسابها.
 ب - إذا كان الغرض اختبار مدى بعد قيمة معامل التمييز عن الصفر فتستخدم
 إحدى الطريقتين الأخيرتين.
 ج- أما إذا كان الغرض اختيار أسئلة ذات مستوى صعوبة متطرفة (سهلة او صعبة)
 ولكن أيضا تميز بين الأطراف فتستخدم طريقة معامل الارتباط المنصف للسلسلة.
 د - إذا كان من بنى الاختبار في شك أن العينات في المستقبل سوف تختلف في
 القدرة عن عينة التقنين فيستخدم معامل ارتباط ثنائي التسلسل حتى يحصل على
 معامل تمييز عالي.
 هـ - إذا كان من بنى الاختبار على ثقة بأن العينات في المستقبل سوف لا تختلف
 في القدرة عن عينة التقنين فيستخدم معامل الارتباط المنصف للسلسلة الحقيقي.
 و- إذا كان السؤال والمحك ثنائيا فإنه يمكن استخدام معامل الارتباط المنصف
 للسلسلة الحقيقي. وكلما ارتفع معامل تمييز الفقرة كلما كان إسهامه أفضل في

زيادة ثبات الاختبار ورفع قيمة تباينه وعلى هذا الأساس يمكن تقويم معاملات تمييز الفقرات بناءً على المعيار الموضح في الجدول أدناه والذي وضعه ايبيل ١٩٦٣

التقويم	مستوى التمييز
فقرة جيد جداً	من ٠,٤٠ فأعلى
فقرة جيدة بدرجة معقولة ولكن يمكن تحسينها	من ٠,٣٠ الى ٠,٣٩
فقرة هامشية تحتاج الى تحسين	من ٠,١٩ الى ٠,٣٩
فقرة ضعيفة تحذف او تعدل	اقل من ٠,١٩

٢- صعوبة الفقرات: أن فعالية الاختبار تتأثر بشكل مباشر في عدد من المواقف بتشتت قيم صعوبة الفقرات. وبصيغة عامة تحدد صعوبة الفقرات بحساب نسبة الأفراد الذين يجيبون على الفقرة أجابه صحيحة، فإذا كانت الإجابة على الفقرة أما صواب أو خطأ فالأسلوب المباشر لتحديد صعوبته هو حساب متوسط الدرجة عليه أو نسبة الإجابة بصواب للمجموع الكلي للأفراد. وهي تمثل الاحتمالية العملية لقدرة الجمهور المعنى بالفقرة أو المتميز بالقدرة على الإجابة عليه أجابه صحيحة. ولأن إمكانية توحيد مستوى الصعوبة بشكل حاسم في كل فقرات الاختبار تؤدي إلى آثار جانبية على كل من صدق الاختبار وتجانسه الداخلي، فيتعين الاهتمام في هذه الحالة بترتيب الفقرات من حيث مستوى صعوبتها من الأسهل إلى الأصعب على أن يكون مدى الصعوبة بين أسهل وأصعب فقرة محدوداً بقدر الإمكان ويدور حول نسبة ٥٠٪ صعوبة ويؤدي ترتيب الفقرات من حيث الصعوبة إلى عدد من المزايا من ذلك:

تتمية ثقة المفحوص بإجابته على الفقرات السهلة أولاً قبل أن يلتقي بالفقرات الصعبة قرب الحد المقابل لأقصى قدرته، كما يؤدي هذا الترتيب إلى خفض إمكانية تضییع جزء من الوقت في فقرات تفوق قدرة الفرد وإهماله لفقرات سهلة وفي متناوله. ويقتصر تحليل مستوى الصعوبة على اختبارات الاستعدادات والقدرات والتحصيل وهو أكبر أهمية بالنسبة لاختبارات الاستعدادات والقدرات أما اختبارات التحصيل فيحدث أحياناً أن يكون تقدير خبير في الموضوع أكثر أهمية من حساب مستوى الصعوبة إحصائياً. كما أن اختبارات السرعة بكل أشكالها تخرج عن إطار مشكلة الصعوبة إذ تتضمن أساساً فقرات سهلة ويفضل اختيار الفقرات ذات

الصعوبة المتوسطة وبمدى ضيق حول متوسط الصعوبة (٥٠٪) لانه يعطى أعلى قيمة تباين للفقرة، وهو ما يؤدي الى ارتفاع قيمة تباين الاختبار ككل وبالتالي ارتفاع قيمة ثباته، أما في اختبارات القوة فيفضل تدرج فقرات الصعوبة من الأسهل للأصعب، وبالتالي احتواء الاختبار على مدى واسع من مستويات الصعوبة. ويمكن الوصول إلى مقياس أفضل لمستوى الصعوبة بواسطة تحويل نسب الصعوبة إلى درجات معيارية باعتبار النسب من الأفراد الذين يجيبون أجابه صحيحة على الفقرة مقابلة للنسب تحت المنحنى الأعتدالي.

$$\text{معامل الصعوبة} = \frac{\text{ص ع} + \text{ص د}}{\text{ن}^2}$$

ص ع = عدد الأفراد في الفئة العليا الذين أجابوا إجابة صحيحة

ص د = عدد الأفراد في الفئة الدنيا الذين أجابوا إجابة صحيحة

ن = عدد الأفراد في إحدى الفئتين

العلاقة بين مستوى الصعوبة ودرجة التمييز للفقرات

تعتمد قيم معاملات تمييز الفقرات المتحصل عليها باستخدام طريقتي معامل الارتباط المنصف للسلسلة الحقيقي، ومقارنة الأطراف العليا والدنيا على مستوى صعوبة الفقرات، كلما ابتعدت قيمة معامل صعوبة الفقرة عن (٠,٥٠) قلت قيمة معامل التمييز للفقرة، وذلك لان الطريقتين السابقتين تقسم الدرجة الكلية الى قسمين في الطريقة الأولى والى ثلاثة أقسام في الطريقة الثانية ثم يتم المقارنة بين القسمين الممثلين للفئة العليا والدنيا، مما يجعلها تتأثر بمعامل الصعوبة.

أما معاملات تمييز الفقرات المتحصل عليها باستخدام طريقة الارتباط المنصف للسلسلة فلا تتأثر بقيمة معامل الصعوبة حيث يتم الحصول على قيمتين مستقلتين لكل من معاملي الصعوبة والتمييز، وذلك لأن هذه الطريقة تعتبر الدرجات الكلية متغيراً متصلاً ولا تقسمه إلى قسمين مما يجعلها لا تتأثر بصعوبة الفقرة، وبالتالي تصبح هذه الطريقة المفضلة عند احتواء الاختبار على تدرج واسع من مستويات الصعوبة.

٥. عالية المشتتات (الموهات)

المشتتات عبارة عن الخيارات الأخرى غير الصحيحة للسؤال ذي الاختيار من متعدد ، ويفترض في المشتتات التالي:

أ- أن تكون جذابة وبالذات للأفراد الذين لا يعرفون الاختيار الصحيح، أي يتم اختيار أي مشتت من قبل مفحوص أو أكثر بنسبة لا تقل عن ٥٪ من المفحوصين.

ب - المشتتات عبارة عن إجابة خاطئة، لذلك فالمشتت الجيد هو الذي يكون عدد المفحوصين الذين يختارونه من الفئة العليا اقل من عددهم في الفئة الدنيا أي انه (يميز باتجاه معاكس لتمييز الفقرة). وكلما كانت قيمة المموه بالسالب كان هذا دليل على ان المشتت جيد.

$$ن ع - ن د$$

$$\text{معامل فعالية المشتت} = \frac{\text{-----}}{ن}$$

ن

$$ن ع = \text{عدد الأفراد في الفئة العليا الذين اختاروا المشتت}$$

$$ن د = \text{عدد الأفراد في الفئة الدنيا الذين اختاروا المشتت}$$

$$ن = \text{عدد الأفراد في إحدى الفئتين}$$

٦- إعداد الاختبار للاستخدام:

بعد أن ينتهي الباحث من جميع المراحل الأساسية لتصميم الاختبار يصبح من الضروري أن يضع الاختبار في صورته النهائية التي سيقدم بها إلى المفحوصين، من ذلك ما إذا كانت فقراته ستقدم في شكل بطاقات مستقلة تتضمن كل بطاقة فقرة واحدة أو ستوضع الفقرات في كتيب مستقل على أن يرفق بها صحيفة منفصلة للإجابة. كما يجب وضع تعليمات مناسبة للاختبار قبل تجربة الاختبار.

وفيما يلي بعض الأمور الواجب أتباعها عند وضع التعليمات للاختبار وتجربته

١- يجب شرح فكرة الاختبار شرحاً دقيقاً حتى لا يكون عدم فهم الأفراد

لفكرة الاختبار سبباً في الإخفاق فيه.

٢ - يجب أن تبرز التعليمات الهامة وأن تكتب بخط واضح ومن المستحسن أن تكون حروفها أكبر من بقية التعليمات العادية.

٣ - من الأفضل أن تحتوي التعليمات على مثال أو نوع من التمرين للمختبر قبل كل اختبار.

٤ - يحسن أن تكون تعليمات الأجزاء المختلفة للاختبار متسقة أو موحدة إن أمكن. فقد ثبت بالتجربة أن توحيد التعليمات يؤدي إلى دقة أكبر وإلى ألفة المختبر بالتعليمات وتمسكه بها، بينما اختلافها قد يؤدي إلى بعض الأخطاء أو إلى التضحية بالدقة.

٥ - بالإضافة إلى أن التعليمات يجب أن تحتوي على طريقة الإجابة ومكانها وأسلوبها ويجب أن تحتوي أيضاً على الزمن المطلوب مراعاته عند الإجابة إذا كان الاختبار يتطلب ذلك

٦ - من الأفضل إجراء تجربة استطلاعية على عينة صغيرة لغرض التعرف على مدى الاختبار ومدى فهم العينة لهذه التعليمات ويقصد بمدى ملائمة الاختبار أي مدى وضوح العبارات والمدة اللازمة للإجابة على الاختبار ومن ثم تعديل فقرات الاختبار في ضوء التجربة الاستطلاعية. ومن الأمور المهمة التي ينبغي أن يراعيها واضع الاختبار هي وضع مفتاح للتصحيح إذ أن مفاتيح التصحيح تختلف باختلاف طريقة الإجابة أي قد تكون الإجابة على كراس الاختبار نفسه، أو إذا كانت الإجابة على ورقة أجابه خاصة ودون كتابة شيء على كراس الاختبار. وطريقة الإجابة على الورقة أفضل بكثير وأكثر اقتصادية من الإجابة على كراس الاختبار نفسه إلا في الحالات الاستثنائية مثل اختبارات الدقة والتي تدخل فيها أرقام كثيرة أو رموز أو رسوم بيانية أو في حالة كون التلاميذ صغار العمر وخبرتهم محدودة في الاختبارات الموضوعية. وإن مفاتيح التصحيح تكون بأنواع متعددة منها:

١ - المفتاح ذو المروحة وهو الأكثر شيوعاً وهو أن يطبق المفتاح على حسب أقسامه الرأسية فيصبح على شكل مروحة ويلاحظ أن بالمفتاح قوائم بدرجات خام حولت إلى درجات معيارية.

- ٢ - المفتاح الشريطي ويشبه المفتاح ذو المروحة إلا أنه يختلف عنه في كون التصحيح الخاص لكل صحيفة من صحف الاختبار قد كتبت على حدة كما أن المفتاح قد كتب على ورق مقوى وليس على ورق عادي.
- ٣ - المفتاح المثقوب في مثل هذا المفتاح قد عملت ثقوب لبيان الكلمات أو الجمل الممثلة للإجابة في مفردات التكملة وهو يصلح لاختبارات التكملة.
- ٤ - المفتاح الشفاف وهو مصنوع من شفاف وشفافية تساعد على سرعة التصحيح وعلى تمييز المفردات التي قد تصحح مرتين من جهة أخرى ويستخدم بكثرة في تصحيح الاختبارات الموضوعية.
- ٥ - المفتاح الكربون تصاحب هذا المفتاح ورقة الإجابة حيث تحدد أماكن الإجابات الصحيحة على ورقة مستقلة تلتصق أطرافها في ظهر ورقة الإجابة حيث تكون مستقرة بالنسبة للمختبر ويطلق ظهر ورقة الإجابة بطلاء أسود بحيث يترك أثراً لأية كتابة أو علامة تسجل على ورقة الإجابة. بالرغم من سرعة هذا المفتاح في التصحيح إلا أنه كثير التكاليف.
- ٦ - المفتاح المطبوع فوق ورقة الإجابة وهو أن يلصق مفتاح التصحيح فوق ورقة الإجابة أي يطبع هذا المفتاح على ورقة الإجابة بعد أن ينتهي المختبرين من الإجابة.
- ٧ - المفتاح الآلي يمكن أن يستخدم المفتاح المثقوب في التصحيح بالآلات إذا أعدت أوراق الإجابة للتصحيح بالآلة وتكون وظيفة المفتاح الآلي هو تقسيم نقط الاتصال إلى دائرتين إحداهما للإجابة الصحيحة والآخر للإجابات الخاطئة كذلك قد يستخدم أحياناً مفتاح لاستبعاد المفردات حتى تطمس اجابات المفردات غير المرغوب فيها.
- وبعد أن تمت كتابة المفردات وأعدت التعليمات يتم تجربة الاختبار تجربة مبدئية على عينة ممثلة للمجتمع المراد عمل الاختبار له وتسجل الملاحظات المختلفة فيما يتعلق بصلاحيه التعليمات وفهم المختبرين لها وكذلك وقت الاختبار وتجمع كافة الملاحظات عن المفردات المختلفة وذلك تمهيداً لتعديل ما يحتاج منها إلى تعديل.

وأثر الانتهاء من التجارب الاستطلاعية الأولية يعمد المصمم إلى اختيار عينة أكبر من الأفراد المستهدفين في البحث بعد إجراء التعديلات المناسبة للاختبار ثم يجري مصمم الاختبار تجربة أساسية للاختبار Try-Out يطبق فيها الاختبار على عينة من الأفراد قوامها حوالي (٤٠٠) فرد تحمل خصائص المجتمع المراد أعداد الاختبار له ومن خلال هذا التطبيق يتم بموجبه اختيار الفقرات الصالحة للصيغة النهائية للاختبار.

وعند ذلك يجب مراعاة نواحي متعددة عند اختيار الفقرات الخاصة بالصورة النهائية للاختبار وهي:

١- صعوبة المفردات ومدى تدرجها. ٢- قدرة المفردات على التمييز. ٣- صدق المفردات. وعن طريق إجراء التحليلات الإحصائية يتم تطبيق هذه الشروط وقد ذكرنا ذلك عند الحديث عن تصميم فقرات الاختبار في الفقرة الخاصة باختبار مستوى الصدق والصعوبة للفقرات.

٧- تقنين الاختبار

يمكن القيام بعملية تقنين الاختبار باتباع الخطوات الآتية:

أ- تحديد المجتمع الذي سيقنن عليه الاختبار تحديداً إجرائياً دقيقاً: حيث ان هذه الخطوة هي الأساس لضمان صحة الخطوات اللاحقة في عملية التقنين، وتتضمن هذه الخطوة تحديد أهم خصائص وسمات المجتمع الديمغرافية، من حيث الخصائص الجغرافية والسكانية والاقتصادية، وتوزيع الفئات العمرية فيه، ونوعية التعليم، والتركيبية الاجتماعية، حيث يتم من خلال هذه المعلومات تحديد العينة التي تمثل المجتمع تمثيلاً جيداً، كما ان هذه المعلومات تمثل خصائص المجتمع والتي على ضوءها يتم تعميم نتائج الاختبار على المجتمعات الأخرى.

ب- اختيار العينة الممثلة للمجتمع وتحديد أسلوب اختيارها: وتعتمد هذه الخطوة على الخطوة السابقة، حيث إن تحديد حجم العينة وأسلوب اختيارها يعتمد بشكل كبير على المعلومات المتوفرة عن مجتمع الدراسة كما تعتمد على

الإمكانات المادية والبشرية المتاحة، وبصفة عامة كلما كان حجم العينة كبيراً كلما كان أفضل وأقرب إلى التمثيل الجيد للمجتمع.

ج- التخطيط الجيد والمسبق لتطبيق الاختبار: وتتمثل هذه الخطوة في وضع خطة شاملة لتطبيق الاختبار تتضمن تحديد الإجراءات والخطوات التي سوف تتبع، وتجهيز جميع أدوات ومستلزمات الاختبار، مع وضع قوائم بأسماء الأماكن التي سوف يتم تطبيق الاختبار فيها (أسماء المدارس مثلاً) مع وضع برنامج زمني للتنفيذ.

د- تطبيق الاختبار: وتستلزم هذه الخطوة توحيد ظروف إجراء وتطبيق الاختبار لجميع أفراد العينة وذلك لتحقيق مبدأ تكافؤ الفرص أمام الجميع لضمان أن الفروق التي رصدها الاختبار تعود للفروق في أداء الأفراد فقط.

هـ - تحليل فقرات الاختبار: وذلك للتعرف على مدى فعالية فقرات الاختبار، ومدى إسهامها في الحصول على خصائص سيكومترية (الصدق والثبات) تتفق مع خصائص الاختبار الجيد، وتشمل هذه الخطوة التعرف على صعوبة كل فقرة من الفقرات، وقدرتها التمييزية، وتباينها، كما تشمل التعرف على مدى فعالية المشتتات لكل فقرة.

و- إيجاد الخصائص السيكومترية للاختبار (الصدق والثبات): وذلك للتأكد من أن الاختبار فعلاً يقيس ما وضع لقياسه فقط، وأن الفرق بين أداء الأفراد في الاختبار ناتج عن الأداء الحقيقي للأفراد وليس بسبب أخطاء القياس العشوائية.

ز- إيجاد معايير الأداء: وذلك من خلال أداء مجموعة التقنين وذلك بغرض توفير إطار مرجعي يمكن من خلاله الحكم على أداء الفرد في الاختبار. والشكل الآتي يوضح الخطوات الأساسية لعملية تقنين الاختبارات النفسية:

