

الخطوات المنهجية لتصميم الاستبيان

أ. عايش صباح

باحثة في علم النفس الأسري

جامعة وهران، الجزائر

aichsabah@yahoo.fr



الخطوات المنهجية لتصميم الاستبيان

بقلم – الأستاذة عايش صباح

المقدمة:

تعتبر مرحلة جمع البيانات أصعب مراحل البحث العلمي، وأحد منابع صعوبة المرحلة هو اختيار (أو تصميم) أداة جمع بيانات مناسبة لهدف ومجتمع الدراسة، حيث يهتم الباحثون في مجالات العلوم النفسية والتربوية والاجتماعية بالحصول على معلومات وبيانات دقيقة حول الظواهر السلوكية، وفي سبيل الحصول على تلك البيانات أو المعلومات يستخدم الباحث عدداً من الأدوات السيكولوجية المتمثلة في الاستبيانات، وقوائم الملاحظة، وقوائم التقدير الذاتي، ومقاييس التقدير، وغيرها.

يعد الاستبيان من أدوات البحث العلمي الأكثر شيوعاً، حيث ازداد استعمال الاستبيان "Questionnaire" في البحوث العلمية على مدى العقود الماضية باعتباره من أدوات البحث الأكثر استخداماً، حتى أصبح الأداة الأولى التي تجمع بها المعلومات، والتي يمكن على ضوءها اختبار فروض الدراسة.

فالاستبيان شأنه شأن كل أدوات البحث العلمي، ذو طبيعة متماسكة، تترابط فيه المقدمات بالنتائج، يبدأ بتحديد الباحث لمشكلة الدراسة بناءً على تصورات نظرية، ووضع تصميم منهجي دقيق لكافة الخطوات التي يشتمل عليها البحث، بداية من وضع الفروض والتساؤلات، مروراً بجمع البيانات وتحليلها وانتهاءً بالوصول إلى النتائج التي تصب في نفس الإطار النظري الذي تم

الانطلاق منه، وكل مرحلة يجب أن تصمم بعناية كبيرة لأن النتيجة النهائية تعتمد على مدى ترابط جميع مراحلها.

يلخص "الضحيان والدليبي" هذه المراحل في ثلاثة أبعاد، وهي كشف الغموض الذي يكتنف المتغير التابع (مؤشرات المشكلة)، والتعرف على المتغيرات المستقلة الأكثر تأثيراً على المتغير التابع (المعرفة النظرية)، والتعرف على شدة ارتباط المتغيرات المستقلة بالمتغير التابع واتجاهها (تفسير وتحليل النتائج). (الضحيان والدليبي، 1998، 90).

هذه المراحل يجب أن تتم بهذا الترتيب والترابط، إذ تعتمد كل مرحلة على نتائج المرحلة التي تسبقها. لكن ما يبدو في كثير من البحوث هو عدم الاهتمام بهذا الترتيب والتسلسل في مراحل إعداد الاستبيانات، حيث يرى "حجر" (2000) أن الباحثين في كثير من الأحيان لا يتعاملون مع هذه المراحل بشكل واحد، لا من حيث نظرتهم إليها ولا من حيث تطبيقهم لها. (حجر، 2000، 121).

حيث يعتقد الكثيرون أن مكنم الصعوبة هو كيفية تحليل البيانات، وتراهم يركزون جل اهتمامهم حول هذا الموضوع، في حين يهملون المراحل الأولى التي يتأسس عليها البحث، أو بناء الاستبيان.

ويرجع ذلك إلى التقنيات الحديثة التي أصبحت تستخدم في تحليل البيانات من برامج إحصائية، اختصرت ذلك التعب المضي الذي كان يقضيه الباحث في الحساب اليدوي، والذي كان يأخذ منه أحيانا عدة أيام في حساب معادلة واحدة، مما جعل الأمر أكثر سهولة وجاذبية، ولا يستغرق من الباحث إلا القليل من الوقت.

حيث يسعى الطلبة في العلوم الاجتماعية لامتلاك واستخدام برامج التحليل الإحصائي على اختلافها، ويبدلون قصارى جهدهم لتعلم هذه البرنامج وكيفية استخدامها؛ ذلك أن هذه البرامج تتيح للباحث إمكانات جديدة وغير محدودة، وتكسب أعماله عمقا ودقة واكتمالا من الناحية الإحصائية.

لكن بقدر ما يعتبر هذا الأمر جيدا ومفيدا في تطوير الإمكانيات الفردية تماشيا مع التطور التكنولوجي الحاصل اليوم ومع عولمة البحث العلمي التي أصبحت ضرورية في كافة المجالات،

وضرورة التعامل مع الحاسوب والبرامج الإحصائية بكفاءة والتخلص من اليدوية في التحليل الإحصائي، بقدر ما يعتبر هذا غير كاف، ولا يغني عن التعمق في البحث حول كيفية بناء الاستبيانات، وكذا معرفة مختلف طرق حساب الخصائص السيكومترية لأدوات القياس، ليتمكن الباحث من اختيار الأسلوب المناسب لأداته وتفسير نتائج التحليل تفسيراً منطقياً. فالكمبيوتر وبرامج الإحصاء المتنوعة وسيلة مساعدة للإنسان، وليست بديلة عنه، كما أنها ليست أساسية في تكوين الباحث، فالمهارات تحتاج إلى عقل واع يتحكم فيها ويوجهها.

"إن الحاسب الآلي يقوم بهذه التحليلات مستخدماً أرقى ما توصل إليه علم الإحصاء من برامج، وعلى درجة كبيرة من الدقة، أما البيانات التي يتم إخضاعها لهذه العمليات فكثيراً ما تكون منتقاة لتلائم الأساليب التحليلية الحديثة المتوفرة حتى تأخذ ذلك "الشكل الجذاب". إن استخدام تلك الدقة الإحصائية العالية والتقنية الدقيقة في بيانات لم توفر بشكل سليم يؤدي إلى نتائج خاطئة بدرجة عالية." (حجر: 2000، 122)

إن الإحصاء هو وسيلة لخدمة البحث العلمي وليس غاية في حد ذاته، لذا لا بد من التيقن أن البرامج الإحصائية ليست كافية وحدها لحل مشاكل تحليل بيانات الاستبيان والوصول إلى النتائج المطلوبة، بل إلى جانب إتقان استخدام البرامج الإحصائية، لا بد من أن يكون الباحث على معرفة كافية ووافية بكيفية تصميم الاستبيان والمقاييس الاجتماعية أو النفسية.

لذلك عندما يستخدم الباحث استبياناً من أجل الحصول على معلومات تساعد في اتخاذ قرار ما، فإنه يواجه مشكلة أساسية تتعلق بكيفية تصميم الاستبيان. لذا كان الهدف من هذه الورقة البحثية التعرض لأهم خطوات ومراحل تصميم الاستبيان.

تعريف الاستبيان:

في حياتنا اليومية إذا أردنا معرفة شيء ما عن شخص معين، أو شخصية اجتماعية، فإن هذا يتطلب منا طرح أسئلة على الشخص المعني، وانتظار رده لمعرفة ما كنا نرغب في معرفته. لكن عندما يتم عرض هذه الأسئلة في تسلسل منظم في شكل مجموعة أسئلة هادفة، إلى واحد أو أكثر

من الناس تم اختيارهم مسبقا مع توقع أن نتلقى ردا كتابيا عن جميع الأسئلة التي طرحت عليهم، فإنه يصطلح على هذه المجموعة من الأسئلة "استبيان" في مجال البحوث وجمع البيانات.

كلمة استبيان مشتقة من الفعل استبان الأمر، بمعنى أوضحه وعرفه، والاستبيان بذلك هو التوضيح والتعريف لهذا الأمر، وهو ترجمة للكلمة الفرنسية "Questionnaire" (إبراهيم خضر: 2013، 221). يقصد بالاستبيان تلك الوسيلة التي تستعمل لجمع بيانات أولية وميدانية حول مشكلة أو ظاهرة البحث العلمي ("النوح: 2004، 100). ويعرفه كل من "غود و هات" Goode & Hatt (1952) بأنه مفهوم يشير إلى الأداة التي نحصل من خلالها على الأجوبة باستخدام نموذج الأسئلة، ويملاً من قبل المستجيب. كما يعرفه "بار، ديفيس، وجونسون" Barr, Davis and Johnso (1953) بأنه تجميع منهجي للأسئلة التي يتم إرسالها إلى عينة من الناس.

ويرى "كومار" Kumar (2005) أنه قائمة مكتوبة من الأسئلة، يتم تسجيل الأجوبة من قبل المشاركين الذين يقومون بقراءة الأسئلة والاستفسار وتسجيل الإجابات. وأخيرا يعرفه "روبنز" Robbins (2009) بأنه أداة يستخدمها الباحثون لقياس متغيرات مفيدة، تقيس ما نريد معرفته. وتكون للإجابة عن الأسئلة، أو لمعرفة البيانات القاعدية للمستجيبين مثل:

- الخصائص العامة، كالعمر والجنس والعرق.

- المعتقدات والمواقف الخاصة بهم.

- السلوكيات الخاصة بهم.

- المعارف الخاصة بهم (Mangal: 2013, 337).

باختصار يمكننا تعريف الاستبيان بأنه أحد الوسائل التي يعتمد عليها الباحث في تجميع البيانات والمعلومات من أفراد أو جماعات كبيرة الحجم، ذات كثافة سكانية عالية عن طريق عمل استمارة تضم مجموعة من الأسئلة أو العبارات، بغية الوصول إلى معلومات كيفية أو كمية، وقد تستخدم بمفردها أو قد تستخدم مع غيرها من أدوات البحث العلمي، وذلك للكشف عن الجوانب التي يحددها الباحث ويصوغها في استفسارات محددة.

خطوات تصميم الاستبيان:

إن بذل الجهد لإخراج استبيان قوي التصميم من ناحية أسئلته الموضوعية، وحجمه، وشكل إخراجها، حتما سيعطي بيانات يمكن أن يستفاد منها لاتخاذ قرار معين حول الموضوع المراد دراسته.

يعتبر إعداد الاستبيان من أهم مراحل البحث العلمي، وليس مهمة بسيطة تعتمد على المهارة اللغوية وشيء من الحس العام والتفكير المنطقي، "فالقادم إلى ميدان البحث الاجتماعي لأول مرة قد يرى أن هذه العملية ليس لها من أسس تحكمها أو قوانين تستند عليها إجراءاتها" (حجر، 2000، 126).

من المهم إعداد الاستبيان على مراحل تبدأ بتعريف بمبررات اختيار البحث، وتحديد أهدافه التي سيتم اختبارها وينتهي باستخراج النتائج. كل مرحلة يجب أن تصمم بعناية كبيرة لأن النتيجة النهائية تعتمد على مدى ترابط جميع مراحلها. وفيما يلي أهم مراحل تصميم الاستبيان:

1- مبررات الاستبيان:

الاستبيان أحد الوسائل العديدة للحصول على البيانات، والباحث الذي يريد استخدام الاستبيان يجب أن يكون متأكدا من أنه لا توجد وسيلة أخرى أكثر صدقا وثباتا يمكن استخدامها في بحثه، ويعتمد هذا القرار على معرفته بنواحي قوة وضعف كل وسيلة (محمود علام، 2005، 405).

إن قلة الاستبيانات المناسبة لموضوع البحث، أو صعوبة الحصول عليها، أو مرور زمن طويل على بنائها يعطي الباحث مبررا لبناء استبيان جديد تتوافر فيه مواصفات الاستبيان الجيد، كما أن الحصول على هذه الاستبيانات باللغة الأجنبية يتطلب منه جهدا أكبر في ترجمتها، وتكييفها على البيئة التي سيطبقها عليها، وذلك لاختلاف الثقافة.

لا بد للباحث من ذكر مبررات استخدام الاستبيان، وبإمكان الباحث الاستعانة باستبيانات جاهزة إذا توفرت في مجال بحثه، بشرط تطويرها، وتجديدها، والتحسين فيها، بحيث تلائم

الاستبيان الجديد، أو موضوع البحث الذي سوف يتم الاستعانة به فيه. وإذا استطاع الباحث الحصول على استبيان سبق استخدامه، فإن ذلك يوفر له الوقت، والجهد، والمال.

2- أهداف الاستبيان:

بعد تحديد مبررات البحث تأتي الخطوة الثانية، وهي في العادة تدور حول أهداف البحث، أو أسئلة البحث. حيث يقوم الباحث بإعداد قائمة بالأهداف الخاصة التي سوف تحققها البيانات التي يتحصل عليها من الاستبيان، ويجب تحديد هذه الأهداف في ضوء أسئلة البحث ومشكلاته، مع توضيح كيف تستخدم كل جزئية من البيانات. (محمود علام، 2005، 405)

إذ على الرغم من أننا نبدأ في بناء الاستبيان من مفهوم علمي مقبول يمكن اختباره، إلا أن تحديد الهدف من الاستبيان يتطلب مزيداً من التفصيل (فرج، 2007، 131)، ويؤدي تحديد الهدف إلى تحديد نوع المعلومات التي يرغب الباحث في الحصول عليها. وإذا لم يستطع الباحث القيام بهذه الخطوة فهذا يعني أن مشكلة البحث ليست واضحة. (محمود علام، 2006، 406)

وكلما كان الباحث دقيقاً في تحديد الأهداف وتوضيحها كلما ساعده ذلك في كتابة الاستبيان بوضوح.

الاستبيان هو وسيلة يستخدمها الباحث للوصول إلى نتائج بحثه، وليس غاية في حد ذاته، فكلما كان اختيار هذه الوسيلة ملائماً لطبيعة البحث وأهدافه كانت نتائج البحث دقيقة.

3- التعريف الإجرائي للسمة أو الخاصية التي يقيسها الاستبيان:

ينبغي على الباحث بعد أن ينتهي من تحديد الهدف الذي يريد اختباره، والوظائف التي يهدف الاستبيان قياسها، أن يحدد بوضوح الإطار النظري الذي سوف يلتزمه في تصوره أو معالجته للخاصية أو القدرة أو السمة التي يستهدف قياسها.

بعبارة أخرى، عليه أن يعرف ما يود أن يقيسه تعريفاً إجرائياً دقيقاً، حيث تلعب التعريفات دوراً كبيراً في البحوث الأمبيريقية، لأنها المسئولة عن تحديد أبعاد المتغيرات البحثية، والتعريف الإجرائي يختلف عن التعريف المفاهيمي أو التكويني أو القاموسي. ويكمن الفرق بين هذين النوعين

من التعريفات في وجود عنصر القياس، ففي التعريف الإجرائي يوصف المتغير من خلال كيفية قياسه، أي وصف كيف يقاس.

إن التعريف الإجرائي يحدد معنى المتغير، وذلك عن طريق التوضيح الصريح لما يجب أن يقيسه المستقصي، والهدف من ذلك أن تكون الخطوط واضحة أمام الباحث منذ البداية كي يلتزم بها في خطواته المتبقية جميعا. ومن ناحية أخرى فإن هذا التحديد يتيح للباحثين الآخرين إمكانية متابعة ما أنجزه الباحث سواء بالتقييم أو بالتطوير. (حفي وآخرون، 1978، 29)

يعطي التعريف الإجرائي operational definition المفهوم معنى محسوساً محدداً، ويزودنا بالمعايير أو الخطوات المحسوسة اللازمة لقياس المفهوم موضوع الدراسة.

4- تحديد طبيعة وخصائص الأفراد الذين يصلح الاستبيان للتطبيق عليهم

بعد تحديد أهداف الاستبيان، تأتي خطوة اختيار المجتمع الذي ستطبق الدراسة عليه، والذي ستعكس النتائج عليه.

على الباحث أن يحدد تحديدا واضحا طبيعة الأفراد الذين سوف يطبق عليهم الاستبيان، ونعني بطبيعة الأفراد أبرز الخصائص التي تميزهم وتتحدد أبرز هذه الخصائص عادة في ثلاث مجالات: السن، المستوى التعليمي، نواحي العجز البدني والقصور العقلي. (حفي وآخرون، 1978، 29).

5- الاستعانة بالدراسات والاختبارات السابقة:

بعد أن تتضح الصورة النظرية للاستبيان في ذهن الباحث عليه أن يرجع إلى ما هو متوفر من الدراسات السابقة والاطلاع عليها وكذا الاستبيانات السابقة ذات الصلة بالموضوع، وذلك للاستفادة منها، ومعرفة نقاط الضعف فيها، ومناقشة المتخصصين حول السمة المراد قياسها.

إذ لابد للخاصية المقاسة أن تستند إلى أساس نظري يبرر مشروعيتها تناولها ويعرفها، وقد يعد الاستبيان للتأكد من مدى جدوى النظرية التي تفسر السمة أو الخاصية المقاسة، والنتيجة المستخلصة قد تفيد النظرية أو تعدلها.

6- اشتقاق فقرات أو أسئلة فرعية ذات صلة بأهداف أو أسئلة البحث:

إذا تم تحديد أهداف البحث، والعينة التي سيتم التعامل معها، وكذا تحديد التعريف الإجرائي للسمات التي سيقاسها الاستبيان، وتم التأكد من أنه لا توجد أداة جاهزة يمكن استخدامها، يبدأ الباحث في كتابة فقرات الاستبيان مع مراعاة جانبين مهمين هما:

1- أن تكون الأسئلة مرتبطة بأهداف الاستبيان.

2- تحديد أسلوب تحليل البيانات بعد جمعها، إذ يتوقف شكل الاستبيان وصورته على هذا الأمر. (محمود عام، 2005،، 406)

يتم اشتقاق هذه الفقرات من خلال العديد من المصادر، كطرح سؤال مفتوح أو أكثر على عدد من المجتمع الذي أعد الاستبيان من أجله، وتحليل الاستجابات باستخدام منهج تحليل المضمون، والخروج ببعض المؤشرات التي يمكن صياغة الفقرات منها، وكذا مناقشة المختصين في المجال بشكل عام، والعاملين فيه والمختصين بصفة خاصة لتحديد أبعاد ومكونات الصفة أو الخاصية المراد قياسها. كما يمكن الاستفادة من المقاييس السابقة، ولا مانع من الاستعانة ببعض بنودها أو مفرداتها خاصة إذا كانت معدة لمجتمعات مشابهة للمجتمع الذي يعد له الاختبار والاشارة إلى ذلك. والاستعانة بالدراسات السابقة في نفس الموضوع والتي يمكن أن تساعد معد الاستبيان ببعض الأفكار والمؤشرات التي يمكن أن تكون أساسا مناسباً للفقرات.

كما أن الملاحظة العيادية لمجموعة من أفراد في أثناء عملهم أو تفاعلهم مع بعضهم البعض، إذا كان هذا العمل أو التفاعل له علاقة بالموضوع، تمد معد الاستبيان بمعلومات قد لا يستطيع أن يحصل عليها من قراءة الدراسات السابقة. (أحمد عويس، 36).

و توجد أشكال عديدة لصياغة الفقرات يمكن تقسيمها إلى ما يلي:

1- فقرات تتناول اختيار إجابة من إجابتين أو إجابات متعددة:

وهي تعتمد على اختيار الإجابة الصحيحة بين بديلين أو عدة بدائل، ويتأثر هذا النوع بالتخمين تأثراً شديداً، ولذا تصحح الدرجة النهائية إحصائياً للتخلص من أثر هذا التخمين كلما قل عدد الاحتمالات المحددة لكل فقرة، ويقل كلما زاد هذا العدد، ويبلغ التخمين أقصاه عندما يصل إلى

اختيارين ويضعف عندما يصل إلى ستة احتمالات، ولتصحيح الدرجة من أثر التخمين يتم طرح عدد الإجابات الخاطئة من عدد الإجابات الصحيحة وتصاغ المعادلة في الصورة التالية:

$$\text{عدد الإجابات الصحيحة} = \frac{\text{عدد الإجابات الخاطئة}}{\text{عدد الاحتمالات} - 1}$$

يشترط في بناء هذا النوع من الفقرات أن تكون الإجابات المحتملة للفقرة تتناول إجابة واحدة صحيحة، وأن تحتوي تلك الإجابات على إجابة قريبة من الصحيحة كما يجب أن تخضع ترتيب الإجابات للتوزيع العشوائي. (محمد حسن، 27)

2- فقرات المزوجة أو المطابقة: يتكون من مجموعة من الفقرات التي يقابلها أو يوازنها مجموعة من البدائل أو الإجابات المحتملة لهذه الأسئلة. (الغامدي، 2008، 37)

3- الاستجابة الحرة: وفي هذه الحالة نكتفي بأن نقدم للمستجيب منها، ونتركه يستجيب كما يشاء. وتمتاز هذه الطريقة بالتخلص من أثر التخمين نهائياً. ولكنها تواجه بصعوبات أخرى وأهم تلك الصعوبات ذاتية التصحيح ومشكلة التوقيت، وهذا النوع من الفقرات أكثر شيوعاً في الاختبارات الإسقاطية. (قدرى حفى وأخرون، 1978، 33)

لا بد من الالتزام ببعض القواعد في صياغة الفقرات، منها قواعد تعلق بالشكل، إذ يجب أن تتسم فقرات الاستبيان بسهولة وبساطة التعبير، واختيار اللغة المناسبة للجمهور المستهدف بحيث تكون مفهومة وتستخدم فيها الألفاظ والتعبيرات المتداولة. مع مراعاة استخدام لغة الحديث في الحياة اليومية، وليس الفصحى أو التعبيرات البلاغية المجازية.

أيضاً أحد العيوب الجوهرية التي يمكن أن يقع فيها مصمم الاستبيان هي صياغة بنود طويلة، فالمفحوص قد لا يتمكن من متابعة الصياغة الممتدة لفكرة معينة حتى نهايتها، فيفقد النقطة الرئيسية في الفقرة عند وصوله إلى نهايتها، وهذا يؤدي إلى إجابات عشوائية وغير متسقة.

كما يجب تجنب كتابة العبارات أو الأسئلة المزدوجة، وألا تحتوي الفقرات إلا على فكرة واحدة فقط. فالأسئلة المزدوجة التي تحتوي على أكثر من فكرة قد تترك المستجيب خصوصاً إذا

كان يريد الاستجابة على كل فكرة بطريقة مختلفة عن الأخرى. (محمود علام، 2005، 406) من ذلك أن بنداً مثل: أنا لا أحب ألا اتفق مع أصدقائي.

فالإجابة بلا-مثلاً- عن هذا البند قد تثير حيرة المفحوص، فيما إذا كانت تعني أنه يجب اللاتفاق أم يحبه (فرج، 2007، 142)

كما أن هناك بعض القواعد المتعلقة بالمضمون أو الأفكار التي تصاغ في محتوى البند ومنها: التركيز على ما يقاس، إذ يجب أن تتعلق الفقرات بجانب مهم من جوانب السلوك المقيس، وليس بالجوانب الهامشية غير المميّزة أو الأعراض سريعة التغير، أو الحالات المزاجية الوقوتية التي لا تكشف عن نمط سلوكي أو سلوك يتسم بالاستقرار النسبي على مدى زمني ممتد. وأن تكون كل فقرة مستقلة عن الأخرى، بحيث لا يترتب على إجابة محددة عن فقرة ما إجابة أخرى بالضرورة على فقرة أخرى. أو أن يترتب على عدم معرفة الإجابة الصحيحة عن فقرة معينة عدم معرفة الإجابة الصحيحة عن فقرة أو عدد آخر من الفقرات.

يؤدي مثل هذا الشكل من عدم الاستقلال إلى انخفاض في ثبات الاختبار، ناتج عن انخفاض طوله مادامت الإجابة عن عدد من الفقرات معا تكون بمثابة الإجابة على فقرة واحدة. (فرج، 2007، 143)

7- اختيار الفقرات:

بعد أن يستقر الباحث على شكل الفقرات، عليه أن يبدأ في اختيارها، والمفروض في هذه المرحلة أن يقوم بصياغة ضعف عدد الفقرات المطلوبة على الأقل. أي أن الباحث إذا كان يتصور أن الاستبيان النهائي مكون من 50 فقرة مثلاً، فعليه بصياغة فقرات تبلغ المائة تقريباً. ذلك أن الخطوات القادمة سوف تتطلب استبعاد عدد من هذه الفقرات، وبالتالي فإن مضاعفة عدد الفقرات يضمن عدم تأثر فقرات الاستبيان بعد استبعاد العبارات غير الصالحة. حيث يقوم بتحديد محاور الاستبيان الرئيسية، ثم كتابة الأسئلة لكل محور من هذه المحاور في مجموعة منفصلة عن المحاور الأخرى. وعند كتابة هذه الأسئلة يراعي الباحث النقاط التالية:

- اختصار أسئلة الاستبيانات .

- استخدام اللغة البسيطة أي اللغة السائدة والمناسبة لمستويات المبحوثين .

- أن لا تكون صيغة السؤال قابلة للتأويل .
- استخدام أشكال بسيطة للردود، مثل "نعم" أو "لا"، والخيارات المتعددة .
- تضمين خيار "ربما" أو "لا أعرف" في الأماكن الملائمة .
- تجنب طرح الأسئلة الشخصية .
- تجنب طرح الأسئلة التي توحى بإجابة معينة .
- طرح سؤال واحد فقط في الفقرة .
- ترك مساحة كافية للإجابة .
- طرح الأسئلة وفق ترتيب منطقي معين، وتدرجها من العام إلى الخاص بحيث تثير اهتمام الأفراد.

8- تعليمات الاستبيان:

قبل الانتقال إلى عمليات المعالجة الإحصائية لل فقرات المختارة يجب أن نعد تعليمات الاستبيان. وذلك بصياغة تلك التعليمات بصياغة واضحة، عن طريق الأخذ بالاعتبار طبيعة الأفراد الذين سوف يطبق عليهم الاستبيان، وبناء على ذلك تتحدد صياغة التعليمات إذا كانت ستلتزم بالفصحى أو العامية، كما أنها سوف تحدد كذلك إذا ما كنا سوف نكتفي بتعليمات مكتوبة يقرأها الجميع، أم تعليمات يقرأها الأخصائي على المفحوصين.

9- إخراج الاستبيان:

- في هذه الخطوة يقوم الباحث بتنسيق الاستبيان وإخراجه بشكل جيد بحيث تثير اهتمام المبحوثين. وهناك عدة نقاط يتم مراعاتها في عملية الإخراج:
- كتابة عنوان البحث في أعلى الاستبيان .
 - ترتيب الأسئلة في الصفحة بطريقة تسمح بالإجابة المناسبة .
 - أن يكون الاستبيان قصيراً قدر الإمكان .

- أن تكون تعليمات ملء الاستبيان واضحة وموجزة .
- أن يكون نوع الورق جيداً والكتابة على وجه واحد فقط .
- يجب تقسيم الأسئلة في محاور وتوضع لها عناوين واضحة .
- يجب في نهاية الاستبيان شكر المجيب على تعاونه.

10- ضبط الاستبيان قبل التطبيق الفعلي:

يقصد به عملية تحكيم الاستبيان، أي أن الاستبيان يقيس ما وضع لقياسه، ولمعرفة ذلك يتم عرض الاستبيان على مجموعة من الخبراء المتمرسين في القياس النفسي وبناء الاختبارات والمقاييس، وكذلك المتخصصون في موضوع البحث، وذلك لإقرار أو حذف أو تعديل أو إضافة فقرات الاستبيان حتى يتم الحصول على الصورة الأولية للاستبيان.

11- تطبيق الاستبيان على عينة الدراسة الاستطلاعية:

بعد ضبط الاستبيان يتم تطبيقه على عينة بسيطة من الأفراد، بحيث تكون هذه العينة خارج عينة البحث لكنها متفقة في خواصها مع أفراد البحث، وذلك لحساب الخصائص السيكومترية للاستبيان (معامل الصدق والثبات).

أولاً: حساب معامل الصدق:

تعد جوانب الصدق من أهم خصائص الاختبارات والمقاييس التربوية والنفسية، فصدق الاختبار يتعلق بالهدف الذي يبني الاختبار من أجله، وبالقرار الذي يتخذ استناداً إلى درجاته، فالاختبار يكون صادقاً عندما يقيس ما وضع لقياسه.

تعرض مفهوم الصدق للكثير من التعديل تبعاً لانتشار واتساع حركة الاختبارات والمقاييس في العلوم الإنسانية المختلفة، وأصبح هناك أنواعاً مختلفة يمكن تجميعها في ثلاثة أقسام رئيسية وهي:

الصدق المرتبط بالمحك Criterion-Related Validity، صدق المحتوى Content Validity، والصدق التكويني Construct Validity.

سنقوم باستعراض بعض الطرق التي يمكننا تعيين معامل الصدق من خلالها كما يلي:

أ- صدق المحتوى عن طريق طريقة استطلاع رأي الحكام:

يسمى هذا النوع من الصدق أحياناً بالصدق الظاهري والذي يعني مدى تطابق اسم الاختبار مع محتواه. تذكر أغلب البحوث والدراسات صدق المحكمين كنوع من أنواع صدق المحتوى Content Validity، ذلك أن صدق المحتوى يعتمد في بعض مراحله على مبدأ التحكيم، حيث يرى "محمود علام": "أن معظم أساليب تقدير صدق المحتوى تعتمد على الأحكام التقييمية لخبراء المواد التدريسية أو المهتمين بتنمية المهارات والكفايات التعليمية والمهنية والفنية، وتتعلق هذه الأحكام بتقدير مدى التناظر بين مفردات الاختبار والنطاق السلوكي الذي تمثله هذه المفردات، ويقوم المحكم بتقييم كل مفردة من مفردات الاختبار في ضوء هذه الأبعاد على ميزان التقدير الموضح في الاستمارة، ويفضل الاعتماد على أكثر من محكم للحصول على تقديرات متسقة بدرجة أفضل" (محمود علام: 2000، 194)

فيما تم الإشارة إليه كنوع منفصل من أنواع الصدق من قبل بعض الباحثين أمثال "عبد الرحمن" (2008) الذي أشار إلي طريقة استطلاع آراء الحكام كطريقة من طرق تعيين معامل صدق الاختبار.

في هذا النوع من الصدق يقدر المحكم المتخصص مدى علاقة كل فقرة من فقرات الاستبيان بالسمة أو القدرة المطلوب قياسها، وذلك بعد توضيح معنى السمة أو القدرة بصورة إجرائية. وتتلخص خطوات هذه الطريقة من طرق الصدق فيما يلي:

-يقوم الباحث بإعداد البنود أو العبارات التي يحتمل أن تقيس السمة المطلوبة، وبطبيعة الحال، كما ذكرنا سابقاً فإن على الباحث أن يضع عدداً من الفقرات يفوق بكثير العدد الذي يريد أن يتكون منه الاختبار المطلوب.

-عرض هذه البنود على مجموعة من المحكمين المتخصصين- وهم من الأساتذة المتخصصين في مجال البحث -ويستحسن أن يزيد عدد المحكمين على ثلاثة.(سعد عبد الرحمن، 2008، 201)

-تجهز التعليمات التي تسبق الفقرات، ويجب على الباحث أن يضع في اعتباره النقاط الآتية فيما يتعلق باستمارة المحكمين على وجه الخصوص:

- أن يبدي المحكم رأيه فيما إذا كانت أسئلة الاستبيان في كل مجموعة تغطي بطريقة صحيحة المحور الخاص بهذه المجموعة أو لا تغطيه، وأنها انعكاس صحيح أو غير صحيح لفروض الدراسة وتساؤلاتها.

- أن متغيرات الاستجابة للأسئلة؛ مثل: (نعم أو لا - أوافق، أو لا أوافق)، التي تقدم للمبحوثين يجب أن يصحبها متغيرات أخرى خاصة بالمحكم تحت عنوان (يقيس - لا يقيس - إلى حد ما)، بمعنى هل يقيس السؤال عناصر المحور المستخرج منه أم لا، ولهذا يجب أن تصحب استمارة التحكيم ورقة منفصلة توضح للمحكم أهداف الدراسة وفروضها وتساؤلاتها، حتى يبني قراره على أساس صحيح.

- يجب أن تخصص مساحة للمحكم في الاستمارة لإبداء آرائه ومقترحاته في استمارة البحث، ومدى تمثيلها لفروض وتساؤلات الدراسة إلى غير ذلك من ملاحظاته المفيدة للبحث والباحث.

إن صدق المحكمين هو مفهوم قديم نظرا لدخوله تحت مسمى "صدق المحتوى"، حيث أشار إلى ذلك العديد من الباحثين منهم: "جيون" "Guion" (1978)، "همبلتون وروجرز" "Hambleton & Rogers" (1991)، "لاوش" "Lawshe" (1975)، "تيتل" "Tittle" (1982)، "لين" "Lynn" (1986) واعتبروه نوعا من أنواع صدق المحتوى.

لكن "بالرغم من انطوائه تحت صدق المحتوى، إلا أن تطبيقاته ليست كثيرة، وهذه الطريقة لا يسهل استخدامها بالاعتماد على التكميم ومن هنا فقد استخدمت هذه الطريقة بطريقة عشوائية غير ممنهجة" (شحاتة: 2011، 166)

كما أن المحكمين يقومون فقط بالاطلاع على البنود ومدى ملاءمتها للتعريف الإجرائي الذي يضعه الباحث، بغض النظر عما إذا كان هذا التعريف مطابقا للسمة التي يقيسها أم لا، وفي غالب الأحيان يتم عرض الاستبيان على المحكمين الذين يعرفهم الباحث مسبقا بغرض تسهيل عملية التحكيم دون الأخذ في عين الاعتبار مدى ملاءمة تخصص المحكمين للموضوع الذي يتم تحكيمه، ناهيك عن اللامبالاة التي يبديها بعض المحكمين الذين لا يأخذون الوقت الكافي لإبداء آرائهم، حيث يكتفون بوضع علامات الموافقة دون التدقيق في العبارات ومدى ملاءمتها وانتمائها

للأبعاد، وأحيانا يقومون بالتحكيم في غياب الباحث، ذلك أن وجود الباحث مع المحكم يزيل الكثير من الغموض واللبس الذي يمكن أن يكتنف عملية التحكيم.

كما أن هؤلاء المحكمين سوف يعكسون خبراتهم الذاتية وتوجهاتهم البحثية في إصدار أحكامهم حول هذا الاستبيان، وينعكس هذا في جوانب كثيرة من الاستبيان، كتحديد الأبعاد والعبارات أو صياغة الأسئلة التي يتضمنها الاستبيان.

إن قياس الصدق عن طريق التحكيم أو آراء المحكمين غير كاف وحده ولا يعتبر صدقا بمعنى الكلمة، حيث أنه يقيس صدق الاستبيان ظاهريا أو سطحيا، ويعتمد على التقدير الذاتي للمحكمين، ويهتم فقط برأي المحكمين فيما يظهر أن الاستبيان سيقبضه وليس ما يقيسه بالفعل، مع أن الأهم من ذلك هو رأي العينة التي سيتم تطبيق الاستبيان عليها في الميدان العملي التطبيقي، والذي قد يظهر عكس ما اتفق المحكمون من بنود تقيس السمة فعليا.

(ب)- الصدق المرتبط بالمحك:

يدل الصدق المرتبط بالمحك على مدى كفاءة الاختبار في التنبؤ بأداء الأفراد في أنشطة محددة، ولهذا الغرض فإن الأداء على الاختبار تتم مراجعته أو ضبطه بالنسبة إلى محك، أي مقياس مباشر ومستقل يقيس ما صمم الاختبار نفسه للتنبؤ به، فبالنسبة لاختبار العصابية مثلا، يمكن ربط نتائجه بمقاييس التقدير أو أية بيانات متاحة عن سلوك الفرد في مختلف مواقف الحياة، ولا بد أن تتحقق في المحك المستخدم شروط عدة واحتياطات مهمة حتى لا تشوه النتائج (الأنصاري: 2000، 97)، ويوجد نوعين من صدق المحك، حيث تميز جمعية علم النفس الأمريكية بين الصدق التلازمي والصدق التنبؤي بناء على الفترة الزمنية الفاصلة بين تطبيق الاختبار والحصول على الدليل من المحك.

الصدق التلازمي يمكن حسابه من خلال معامل الارتباط بين درجات الأفراد على الاختبار ودرجاتهم في الأداء الفعلي في جوانب السلوك التي يقيسها الاختبار. بشرط أن تكون درجات أداء الأفراد الفعلية قد تم جمعها وقت إجراء الاختبار أو قبله. والصدق التلازمي هو أكثر أنواع الصدق ملاءمة للاختبارات التي تستخدم لأغراض التشخيص لا التنبؤ بنتائج المستقبل.

أما الصدق التنبؤي فإنه يقوم على أساس حساب القيمة التنبؤية للاختبار، أي معرفة مدى صحة التنبؤات التي ننبئها على درجات الاختبار. والتنبؤ هنا يقوم على أساس أن استجابات الفرد بالنسبة لسمة معينة- "القدرة اللفظية-التحصيل الدراسي - القدرة العددية- " دليل على أدائه في مجال معين بعد فترة من الوقت قد تكون سنة أو عدة سنوات.(أبو هاشم، 2006، 22)

ويمكن تلخيص أهم الخطوات التي يمكن إتباعها عند وجود محك خارجي:

1-يقوم الباحث باختيار المحك الصادق بناءً على الشروط والمعايير التي يجب أن تتوفر في المحك الصادق من كثرة الاستخدام والدراسات، ومن حيث مناسبتها للمرحلة العمرية التي صمم من أجلها الاختبار، وطبيعة المجموعة التي سوف يطبق عليها.

2- يتم تطبيق الاختبار المطلوب وإيجاد صدقه على العينة أولاً، ثم يتم بعد ذلك تطبيق الاختبار المحك- مع ملاحظة الفترة الزمنية- لتفادي عوامل الملل والإجهاد وغير ذلك.

3- يحسب معامل الارتباط بين درجات العينة على الاختبار المحك ودرجاتهم على الاختبار المطلوب تعيين معامل صدقه، ويدل هذا المعامل على صدق الاختبار.(سعد عبد الرحمن، 2008، 204)

ج- صدق التكوين الفرضي(صدق المفهوم):

يتناول هذا النوع من الصدق العلاقة بين نتائج الاختبارات والمقاييس وبين المفهوم النظري الذي يهدف الاختبار لقياسه، وبعبارة أخرى فإن صدق التكوين الفرضي يهدف لتحديد التكوينات الفرضية التي يعزى إليها تباين الأداء في الاختبارات، أي أن هذه التكوينات الفرضية هي التي يتركز عليها الاهتمام وليس درجات اختبار المحك أو سلوك الفرد.(محمود علام، 2000، 215)

هناك عدد من الأساليب والإجراءات التي يمكن من خلالها التأكد من صدق المفهوم منها:

صدق المقارنة الطرفية:

يقوم هذا النوع من الصدق على تقسيم درجات الاستبيان(المحك) إلى مستويين: ممتاز وضعيف، ثم مقارنة درجات السؤال (الفقرة) في المستوى الضعيف بدرجاته في المستوى الممتاز، وكلما زادت درجات الفقرة في المستوى الميزاني الممتاز عن درجاتها في المستوى الميزاني الضعيف كلما

زاد صدق المفردة، والعكس صحيح فكلما نقصت درجات المفردة في المستوى الميزاني الممتاز عن درجاتها في المستوى الميزاني الضعيف نقص صدق المفردة. (ابراهيم عثمان، 152).

من أبسط الطرق التي تستخدم لتحقيق هذه الفكرة مقارنة بمتوسطات درجات الأقوياء بمتوسطات درجات الضعفاء ثم حساب دلالة الفروق بين هذه المتوسطات، وعندما تصبح لتلك الفروق دلالة إحصائية واضحة نستطيع أن نقرر أن الاختبار يميز بين الأقوياء والضعفاء في الميزان، وبذلك نطمئن إلى صدقه، وعندما لا تصبح لتلك الفروق دلالة إحصائية واضحة فإننا لا نستطيع الاطمئنان إلى صدق مثل هذا الاختبار. إلا أن هذه الطريقة تدل على صدق الاختبار ولا تدل بطريقة عددية أكيدة على مقدار هذا الصدق.

ويمكن الإشارة هنا إلى أن المقارنة الطرفية تتم بأسلوبين:

أ- مقارنة الأطراف في الاختبار والمحك الخارجي:

وفيه يتم مقارنة الثلث الأعلى في درجات الاختبار بالثلث الأعلى في درجات المحك الخارجي، والثلث الأدنى في درجات الاختبار بالثلث الأدنى في درجات المحك الخارجي، وتستخدم لهذه المقارنة اختبار "ت". فإذا لم تكن هناك دلالة إحصائية للفرق بين المتوسطين في حالة مقارنة الثلث الأعلى في درجات المحك بالثلث الأعلى في درجات الاختبار، وإذا لم تكن هناك دلالة إحصائية للفرق بين المتوسطين في حالة مقارنة الثلث الأدنى في درجات المحك بالثلث الأدنى في درجات الاختبار. في هذه الحالة يمكن أن نقول إن الاختبار صادق - بطبيعة الحال نحن نفترض صدق المحك الخارجي الذي يتم اختياره من أجل تعيين صدق الاختبار - كما نفترض أيضا تكافؤ المحك الخارجي مع الاختبار من حيث البناء.

ب- مقارنة الأطراف في الاختبار فقط:

وهذا أسلوب آخر يعتمد على مقارنة درجات الثلث الأعلى بدرجات الثلث الأدنى في الاختبار، وتتم هذه المقارنة عن طريق حساب الدلالة الإحصائية للفرق بين المتوسطين. فإذا كانت هناك دلالة إحصائية واضحة للفرق بين متوسط الثلث الأعلى ومتوسط الثلث الأدنى يمكن القول بأن الاختبار صادق.

الحقيقة أن هذه الطريقة عموماً طريقة سهلة وأقل دقة من طريقة التحليل العاملي أو المحك الخارجي، ولكنها تعطى مؤشراً سريعاً عن صدق الاختبار. (أبو هاشم، 2006، 26)

استخدام معاملات الارتباط في حساب صدق الاستبيان (الاتساق الداخلي)

تستخدم هذه الطرق للتحقق من الارتباطات الداخلية للاختبار، سواء بين الفقرات ببعضها البعض أم بين المفردات والأبعاد التي يتكون منها المقياس، أم بين الأبعاد ببعضها البعض وهذه الطرق هي:

(أ)- حساب معاملات الارتباط بين مفردات الاستبيان ببعضها البعض، حيث يتم حساب الارتباطات البينية بين مفردات الاختبار، وينتج عن هذا الإجراء التوصل إلى مصفوفة معاملات الارتباط. وإذا كانت معاملات الارتباط ذات قيم موجبة عالية بين فقرات الاستبيان هذا يدل على أنها جميعاً تقيس نفس الهدف، وإذا كانت القيم منخفضة وتقرب من الصفر فهذا يدل على عدم وجود علاقة بين الفقرات وبالتالي تحتاج إلى إعادة صياغتها أو إلغائها.

(ب)- حساب معاملات الارتباط بين أبعاد المقياس ببعضها البعض، ويتم التحقق من ذلك بحساب معاملات الارتباط بين الأبعاد، فإذا كانت معاملات الارتباط مرتفعة بينهم هذا يعني أنها تقيس نفس الهدف مما يدل على صدق المقياس لما يقيسه.

(ج)- حساب معاملات الارتباط بين درجة كل فقرة والدرجة الكلية للمقياس، وفي هذه الطريقة تحسب معاملات الارتباط بين درجة كل مفردة والدرجة الكلية للمقياس، وتعتبر كل فقرة صادقة عندما يكون الارتباط بالدرجة الكلية للمقياس دال إحصائياً، ويستبعد الباحث الفقرات الضعيفة التي لا تظهر ارتباطاً دالاً بالدرجة الكلية للمقياس، فارتباط درجة الفقرة بالدرجة الكلية تعني أن الفقرة تقيس نفس الهدف الذي يقيسه الاستبيان، وبالتالي تدل على أن الاستبيان على درجة من الإتساق الداخلي.

(د)- حساب معاملات الارتباط بين درجة الفقرات ودرجة كل بعد من أبعاد الاستبيان.

(هـ)- حساب معاملات الارتباط بين درجة كل بعد والدرجة الكلية للمقياس. (غنيم، 2004،

(د) - طريقة التحليل العاملي:

التحليل العاملي هو أسلوب إحصائي يستهدف تفسير معاملات الارتباط الموجبة التي لها دلالة إحصائية بين مختلف المتغيرات، ويعد منهجا إحصائيا لتحليل بيانات متعددة ارتبطت فيما بينها بدرجات مختلفة من الارتباط في صورة تصنيفات مستقلة قائمة على أسس نوعية للتصنيف.

حيث يعمل على تلخيص العديد من المتغيرات لعدد أقل يعرف بالعوامل (Factors) وكل مجموعة من المتغيرات تربط بعامل واحد فقط بواسطة دالة، ويبدأ بحساب معاملات الارتباطات بين عدد من المتغيرات، وعندما يتم الحصول على مصفوفة من الارتباطات بين هذه المتغيرات لدى عينة البحث التي تم إجراء القياس عليها، يتم تحليل هذه المصفوفة الارتباطية تحليلا عاملياً لنصل إلى أقل عدد ممكن من المحاور أو العوامل والتي تمكنا من التعبير عن أكبر قدر من التباين بين هذه المتغيرات.

ثانيا: حساب معامل الثبات

حينما نحكم على اختبار بأنه ثابت، فإن ذلك يعني أن هذا الاختبار على درجة عالية من الدقة والإتقان والاتساق وأنه موثوق فيه، ويعتمد عليه. وهذا هو المعنى اللغوي للكلمة الأجنبية Reliability وهو قريب من المعنى الاصطلاحي.

فثبات الاختبار يشير إلى المطابقة بين نتائج الاختبار في المرات المتعددة التي يطبق فيها هذا الاختبار على نفس الأفراد. (أحمد عويس، 51). كما يقصد به مدى خلو المقياس من الأخطاء غير المنتظمة التي تشوب القياس، أي مدى قياس الاختبار للمقدار الحقيقي للسمة التي يهدف لقياسها، فدرجات الاختبار تكون ثابتة إذا كان الاختبار يقيس سمة معينة قياسا متسقا في الظروف المتباينة التي قد تؤدي إلى أخطاء القياس. والثبات بهذا المعنى يعني الاتساق أو الدقة في القياس.

الأخطاء غير المنتظمة التي تؤثر في درجات الاختبارات تكون أخطاء عشوائية يصعب التنبؤ بها من موقف إلى آخر، ولذلك تعمل على خفض ثبات الدرجات. وهذه الأخطاء العشوائية ترجع إلى عوامل بعضها يتعلق بالاختبار، مثل عدم وضوح مفرداته، وغموض تعليماته، وعدم تحديد

محكات تصحيح مفرداته، والبعض الآخر يتعلق بالظروف البيئية مثل الإضاءة والتهوية والضوضاء، وملائمة غرفة الاختبار. (محمود علام، 2000، 131)

وهناك مفاهيم ثلاثة أساسية لثبات درجات الاختبار هي:

1- أن يعطي الاختبار نفس النتائج تقريبا إذا أعيد على نفس المجموعة من الأفراد: وهذا يعني أن الاختبار، أو بمعنى أدق درجات الاختبار لا تتأثر بتغير العوامل والظروف الخارجية، حيث أن إعادة تطبيق الاختبار والحصول على نفس النتائج يعني دلالة الاختبار على الأداء الفعلي أو الحقيقي للفرد مهما تغيرت الظروف.

2- بناءً على المفهوم السابق فإن ثبات الاختبار يعني أيضا دلالة الاختبار على الأداء الفعلي أو الأداء الحقيقي للفرد- هذا الأداء يعبر عنه بالدرجة الحقيقية (د ح) التي يحصل عليها الفرد في اختبار ما. (وهذه غير معلومة).

والأداء الحقيقي هو جزء من الأداء العام أو الكلي الذي يعبر عنه بالدرجة الكلية (دك)، وهي الدرجة الملاحظة أو المسجلة على الاختبار والتي حصل عليها الفرد. أما الجزء الآخر فهو الأداء الذي يعود إلى أخطاء الصدفة أو الظروف الخارجية البعيدة عن موضوع الاختبار، ويعبر عنه بدرجة الخطأ. (دخ) وهي غير معروفة أيضا. وهكذا فإن:

$$\text{دك} = \text{دخ} + \text{دح}$$

أي أن الدرجة الكلية = الدرجة الحقيقية + درجة الخطأ

ولهذا يمكن القول إن معنى دلالة ثبات الاختبار على الأداء الحقيقي إنما هو الدلالة على التباين الحقيقي والارتباط به. حيث أن معامل ثبات درجات الاختبار يساوي النسبة بين التباين الحقيقي إلى التباين العام أي أن:

$$\text{معامل الثبات} = \frac{\text{التباين الحقيقي}}{\text{التباين العام}} = \frac{\text{دح}^2}{\text{دك}^2}$$

ولنوضح ذلك بمثال:

عندما تذهب إلى السوق لتشتري صندوقاً من البرتقال من بائع معين، فإن وزن الصندوق ليس هو وزن ما تأكله من البرتقال فقط، ولكنه أيضاً يشمل قشر البرتقال والورق الذي يغلف البرتقال والمادة المصنوع منها الصندوق.

وهذا ما يقابل التباين الكلي أو التباين العام (الوزن الكلي للصندوق)، أما وزن قشر البرتقال والورق المغلف للبرتقال والمادة المصنوع منها الصندوق - وهذا ما سوف نتخلص منه-، وهو يختلف من صندوق لآخر- فهو يقابل تباين الخطأ، أما وزن ما سوف نأكله من برتقال فهو يقابل التباين الحقيقي.

وعليه فإنه كلما زادت نسبة وزن ما سوف نأكله من برتقال إلى نسبة وزن الصندوق ككل كنت مقتنعاً تماماً بما دفعته من ثمن في هذا الصندوق والعكس صحيح. وبالمثل فإن درجات الاختبار التي ترتفع فيها نسبة المكون الحقيقي للتباين العام تكون أكثر ثباتاً من تلك الدرجات التي تقل فيها هذه النسبة.

وللتلخيص فإننا نقول: إن درجات الاختبار تعتبر ثابتة إذا ارتفعت نسبة المكون الحقيقي في التباين العام لهذه الدرجات أي أ: $\frac{ع^2}{ع^2}$ تكون أعلى ما يمكن، بينما $\frac{ع^2}{ع^2}$ تكون أقل ما يمكن.

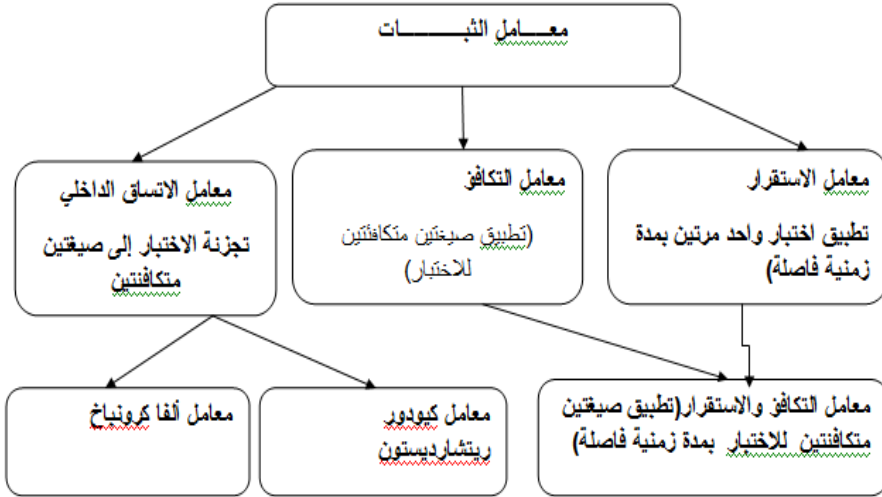
3- أن تكون هناك علاقة قانونية بين وحدات الاختبار أو بنوده، فإن ذلك يدل على التناسق في البناء الداخلي للاختبار، وهذا يعني أن معامل ثبات الاختبار سوف يتوقف على العلاقة أو الارتباط بين كل وحدة ووحدة أخرى (الارتباطات البينية)، كما يتوقف أيضاً على ارتباط كل وحدة بالاختبار ككل. ويتضح من هذا أن تماسك الاختبار أو تناسق بنائه يدل على ثبات درجاته. (عبد الرحمن، 1998، 164، 166)

طرق حساب معامل الثبات:

عندما نقوم بتطبيق أحد الاختبارات النفسية على عينة من مجتمع ما، فإن الدرجات التي نحصل عليها لكل فرد من أفراد هذه العينة على هذا الاختبار (تعرف بالدرجات المشاهدة). و تشمل بداخلها على الدرجات الحقيقية وعلى جزء من الدرجات المشاهدة الراجعة إلى خطأ

القياس، و ليس بوسع الفاحص أن يفصل بين الدرجة الحقيقية ودرجة الخطأ، لكن من الممكن أن نحصل على معاملات الثبات عن طريق مجموعة من الأساليب المختلفة نعرضها فيما يلي:

الشكل التخطيطي التالي يوضح هذه الأنواع من معاملات الثبات:



شكل (1) يوضح الأنواع المختلفة من معاملات الثبات من تصميم الباحث صلاح الدين محمود علام

1- طريقة إعادة تطبيق الاختبار

من العوامل الهامة التي تؤثر في ثبات درجات الاختبار "عامل الزمن"، لذلك فإن هذه الطريقة تعتمد على فحص مدى ثبات الاختبار في ضوء متغير الزمن، وذلك بتطبيق الاختبار، ثم إعادة تطبيق نفس الاختبار على نفس العينة من الأفراد تحت نفس الظروف بعد فترة زمنية. ثم حساب معامل الارتباط بين الدرجات التي تحصل عليها في مرتي التطبيق، ويسمى معامل الارتباط الناتج "معامل الاستقرار بين مرتي التطبيق" "coefficient of stability"، حيث يعبر عن مدى الاتساق بين درجات نفس الاختبار إذا ما أعيد تطبيقه على نفس الأفراد.

يمكن استخدام الطريقة العامة لحساب معامل الارتباط (باستخدام الدرجات الخام مباشرة)، والمشكلات التي تواجه هذه الطريقة في حساب الثبات هي:

- أ- طول الفترة الواقعة بين مرتي التطبيق حيث يتوقف طول الفترة الزمنية على:
- العمر الزمني للمفحوصين: فتكون قصيرة نسبيا مع الصغار عن الكبار بسبب تدخل عامل النضج العقلي والجسمي والاجتماعي في التأثير على إجابات المفحوصين في التطبيق الثاني.
 - طبيعة الاختبار: عند تطبيق الاختبارات التي تتطلب حل المشكلات فإن المفحوص سوف يحل المشكلات بنفس الطريقة التي سبق أن توصل إليها في المرة الأولى.
 - ينصح أن لا تقل الفترة الزمنية بين مرتي التطبيق عن أسبوعين ولا تزيد عن ستة أشهر على أن تحدد في ضوء عوامل عمر المفحوصين وطبيعة الاختبار.
 - ضمان أن تكون ظروف تطبيق الاختبار في المرة الثانية هي نفسها ظروف التطبيق الأول من حيث المكان "حالة المناخ"، "الضوضاء"، "التهوية"،... وغيرها من العوامل التي يمكن أن تسهم في التباين الخطأ لدرجات الاختبار عند إعادة التطبيق.
 - عندما يؤدي المفحوص مهام الاختبار بأقصى قوة بدنية له- في اختبارات المهارات الحركية- فإن تكرار أداء نفس المهام عند إعادة التطبيق سوف يتأثر بحالة التعب البدني للمفحوص، وخاصة إذا كانت الفترة الزمنية بين التطبيقين قصيرة نسبيا، مما يؤدي إلى تغير درجات المفحوص في الاختبار الثاني، وبالتالي انخفاض في قيمة معامل الثبات. (غنيم، 2004، 71، 73)

2- طريقة الصور المتكافئة

هنا يتم إنشاء صورتين متكافئتين من الاختبار الواحد، من حيث تمثيل جانب السلوك المطلوب قياسه. أي أن التكافؤ يجب أن يشمل الجوانب الآتية:

- 1- عدد أسئلة الاختبار.
- 2- عدد مكونات الوظيفة التي يقيسها الاختبار.
- 3- نسبة البنود التي تخص كلا منها.
- 4- مستوى صعوبة البنود.
- 5- طريقة صياغة البنود.

6- طول الاختبار وطريقة إجرائه وتصحيحه وتوقيتته.

7- تساوي متوسط تباين (ع) درجات الأفراد على كل من الصورتين.

8- تعليمات التطبيق والتصحيح. (الأنصاري، 2000، 123)

تتلخص طريقة الصيغ المتكافئة في اختبار الأفراد أنفسهم بإحدى الصيغ في المرة الأولى، ثم تستخدم صيغة مكافئة لها في المرة الثانية، ويستخرج معامل الارتباط بين الدرجات في المرتين، وهو يمثل- عندئذ- معامل ثبات الاختبار.

فيما يختص بالفواصل الزمنية بين الصيغتين فقد يكون أحد نوعين هما:

1- تطبيق الصيغتين في الجلسة ذاتها تطبيقا متعاقبا في الحال.

2- تطبيق إحدى الصيغتين في جلسة، ثم الصيغة الثانية في جلسة أخرى مع فاصل زمني بينهما.

يشير النوع الأول إلى الثبات عبر الصيغتين فقط، أما النوع الثاني فيعد مقياسا أو دليلا على كل من الاستقرار عبر الزمن واتساق الاستجابات فيما يختص بعينات مختلفة من البنود (أو صيغ الاختبار)، ومن ثم فإن هذا المعامل يجمع بين نوعين من الثبات كلاهما مهم لمعظم أغراض القياس، ولذلك يمدنا ثبات الصيغ المتعاقبة بمقياس مفيد لتقييم عديد من الاختبارات. (عبد الخالق، 1996، 46).

3- طريقة التجزئة النصفية:

في هذه الطريقة يتم تجزئة المقياس إلى نصفين، ويعطى كل فرد درجة في كل نصف، أي أننا بعد تطبيق المقياس نقسمه إلى صورتين متكافئتين. وأفضل أساس للتقسيم في هذه الطريقة هي أن يحتوي القسم الأول على المفردات الفردية (1، 3، 5، ...)، والقسم الثاني يحتوي على المفردات الزوجية (2، 4، 6، ...).، حتى نقلل ما أمكن من العوامل المؤثرة في أداء الأفراد، مثل الوقت، والجهد، والتعب، والملل، وغيرها، وميزة هذه الطريقة هو توحيد ظروف الإجراء توحيدا تاما، ونظرا لأن معامل الارتباط في هذه الحالة يكون بين نصفي المقياس. فقد ظهرت عدة معادلات تعدل

معامل الارتباط بين النصفين، بحيث تأخذ في اعتبارها مضاعفة طول الاختبار، ومن أهم هذه المعادلات معادلة سبيرمان:

$$\frac{2r_s}{r_s + 1} = 11r$$

حيث:

11r :: معامل ثبات الاختبار كله

rs :: معامل الارتباط بين نصفي الاختبار

فإذا كان معامل الارتباط بين نصفي الاختبار 0.60 فإن:

$$0.75 = \frac{1.20}{1.60} = \frac{(0.60)2}{0.60+1} = 11r$$

ويلاحظ أن طريقة التجزئة النصفية تؤكد حقيقة مهمة فيما يتعلق بالثبات وهي أن معامل ثبات الاختبار يرتبط بطوله فيزيد معامل الثبات عندما يزداد طول الاختبار.

ويندر أن يكون معامل ثبات الاختبارات التي تقل عن عشر مفردات مرتفعاً، ولذلك يعتبر 10 مفردات هو الحد الأدنى للاختبار الثابت. (محمود علام، 2008،، 473)

4- طريقة الاتساق الداخلي Internal Consistency:

تعتمد 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 نذكر:

1-4- معادلة كودر ريتشاردسون: Kuder - Richardson:

تهدف طريقة كودر ريتشاردسون للتوصل إلى قيمة تقديرية لمعامل ثبات الاختبارات غير الموقوتة أي اختبارات القوة، والتي تكون درجات مفرداتها ثنائية أي إما واحد صحيح أو صفر (محمود علام، 2000، 160)، وتعتمد هذه المعادلة على توفر بيانات عن تباين كل مفردة من

مفردات الاختبار . وفي حالة عدم توفر هذه البيانات يمكن استخدام الصيغة رقم (21) بنفس المعادلة وهي تتميز بالسهولة والسرعة في حسابها حيث أنها لا تحتاج إلى معرفة تباين البنود، ولكن يعيها أنها أقل دقة من الصيغ السابقة . وقد وضع كيودر ريتشاردسون شروطاً لاستخدام هذه المعادلة وهي:

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

- ألا يكون عدد الأسئلة المتروكة كبير .

- تقارب مستوى صعوبة الأسئلة .

- تساوى معاملات الارتباط بين درجات الأسئلة (أبو هاشم، 2006، 08)

2-4- معادلة ألفا كرونباخ Cronbach Alpha:

يعتبر معامل ألفا حالة خاصة من معادلة كيودر ريتشاردسون، اقترحه كرونباخ 1951، ويمثل معامل ألفا متوسط المعاملات الناتجة عن تجزئة الاختبار بطرق مختلفة، وبذلك فإنه يمثل معامل الارتباط بين أي جزأين من أجزاء الاختبار.

$$\text{ومعامل } \alpha = \frac{E - E^2}{E^2} \times \frac{N}{1 - N}$$

حيث E^2 هو مجموع تباين البنود أو الأسئلة، بمعنى أن يحسب تباين كل بند من بنود الاختبار (من درجات الأفراد في هذا البند)، ثم يوجد مجموع هذه التباينات لنحصل على E^2 ، N = عدد البنود، E^2 تباين الاختبار ككل، وتشتت أن تقيس بنود الاختبار سمة واحدة فقط، وتستخدم هذه المعادلة في المقاييس والاختبارات متعددة الاختيارات وليست الثنائية. (سعد عبد الرحمن، 1998، 172)

يتم تفسير معامل الناتج في ضوء معامل الارتباط الناتج، وأعلى قيمة يمكن أن يصل إليها معامل الثبات هي (1)، وهي قيمة لا نصل إليها في أغلب الأحوال وخاصة على المستوى الإنساني.

على هذا الأساس إذا حصلنا على معامل ثبات قيمته (0.95) فإن هذا يعني أن 0.95 من تباين درجات الاختبار هو حقيقي، وأن ما تبقى (0.05) هو من نوع التباين الخطأ. (غنيم، 2004، 70).

ويتأثر تقدير الثبات بمجموعة من العوامل منها:

- طول الاختبار:

حيث تزداد قيمة معامل الثبات بزيادة عدد الأسئلة أو بنود الاختبار، ويرجع ذلك إلى أن زيادة عدد الأسئلة يؤدي إلى شمول أكثر للمحتوى، ومن ثم صدق محتوى مرتفع.

هنا يمكن القول أن العلاقة بين عدد بنود الاختبار ومعامل ثباته علاقة طردية، بمعنى أنه إذا زاد عدد البنود ارتفع معامل ثبات الاختبار، ومن ثم عندما يزداد طول الاختبار فإن أخطاء القياس الناتجة عن الصدفة يختزل بعضها بعضاً، وتعتمد الدرجات أكثر فأكثر على خصائص الشخص المختبر، وتكون النتيجة الحصول على تقدير أكثر دقة لهذه الخصائص.

بالطبع توجد اعتبارات عملية تحدد مقدار الزيادة الممكنة في طول الاختبار، منها أن الوقت الذي يتوفر لتطبيق الاختبار محدود، ومنها ما يتعلق بعوامل التعب والملل عند المفحوصين. وتتحدد في بعض الأحيان بعدم قدرتنا على كتابة فقرات جديدة بنفس جودة الفقرات الأصلية. ولكن، في حدود هذه الاعتبارات، يمكننا أن نرفع ثبات الاختبار، بالقدر الذي نريده، بزيادة طول الاختبار (أبو هاشم، 2006، 11).

تباين درجات الأفراد:

عند حساب معامل الارتباط بين متغيرين فإن هذا المعامل يتأثر بمدى كل متغير منهما، ومن ثم فإن ضيق المدى أو اتساعه يؤثر على معامل الارتباط، أو بمعنى آخر معامل ثبات الاختبار. يقل معامل الثبات بزيادة تجانس المجموعة المطبق عليها وكلما ارتفع تباين الأفراد كلما ارتفع معامل الثبات، لذلك يجب أن يصاحب معامل الثبات وصف دقيق لنوع المجموعة التي حسب من درجات أفرادها.

هذا يعني أن معامل الثبات يقل عندما يقل التباين في مجموعة ما، وعليه نقول إن العلاقة بين التباين ومعامل الثبات هي علاقة طردية، مع ملاحظة أننا نتحدث عن التباين الحقيقي

كسبب لزيادة التباين العام. أما إذا افترضنا أن التغير في التباين العام أدى إلى التغير في تباين الخطأ، وليس إلى التباين الحقيقي، فإن العلاقة بين تباين الدرجات ومعامل الثبات تصبح غير ذلك تماماً حيث تصبح علاقة عكسية. (أبو هاشم، 2006، 12)

مستوى صعوبة/سهولة فقرات الاختبار:

تشكل مستوى صعوبة/سهولة مفردات الاختبار عامل مهم في التأثير على ثباته، إذ في حالة الاختبار الذي يتضمن أسئلة شديدة الصعوبة تنخفض قيمة معامل الثبات، ويرجع ذلك إلى لجوء الطالب للتخمين حتى يتمكن من الإجابة على الأسئلة، أو قد يلجأ إلى الغش، وكلها أساليب تؤدي إلى اختلاف درجات الطالب عبر مرات تطبيق الاختبار، وبالتالي تؤدي إلى عدم الاتساق فيما نحصل عليه من نتائج. كما تؤدي الصعوبة الشديدة للاختبار إلى خفض التباين بين درجات الطلاب ويصبح المدى بينها صغيراً حيث أن الطالب المتفوق والطالب الضعيف سوف يحصلون على درجات متقاربة، مما يؤدي إلى انخفاض قيمة التباين الحقيقي وبالتالي انخفاض في قيمة معامل ثبات الاختبار.

كما أن الاختبارات التي تتضمن أسئلة سهلة جداً يستطيع جميع الطلاب الإجابة عليها بسهولة بصرف النظر عن مستواهم المعرفي، مما يؤدي إلى ضيق مدى الفروق الفردية وانخفاض قيمة التباين الحقيقي، وبالتالي انخفاض قيمة معامل الثبات الناتج. (غنيم، 2004، 82، 83).

من خلال هذه الطرق المختلفة لحساب صدق وثبات الاستبيان يكون الباحث قد تأكد من صدق وثبات أداة الدراسة بعد تطبيقها على عينة الدراسة الاستطلاعية وأصبحت جاهزة لتطبيقها على عينة الدراسة الأساسية.

بناءً على ما تقدم ذكره من طرق لتقدير الصدق والثبات، يمكن القول أن تفسير النتائج والقرارات المبنية على حسابهما تستمد قوتها من البيانات التي جمعها الباحث، وليس من الأداة في حد ذاتها، فالصدق والثبات في الأساس موجه للسلوك المقاس وليس لأداة القياس وما تحويه من بنود، لأن دور الاستبيان في الأساس هو سحب عينات من السلوك الذي يريد الباحث قياسه ومن ثم التحقق من صدقه وثباته.

إذا على الباحث أن لا يتلقى بيانات الدراسة بشكل سلبي ثم يقوم بتحليلها وتعبئتها في أطر نظرية، وإنما يتحصل عليها بشكل واع ومقصود وبناء ويقوم بتكوينها عن طريق العناية بانتقاء وتصميم وتنفيذ الطرق التي يتم بها الحصول على البيانات.

"فالارقام والبيانات تتميز بقابليتها لان تقول كل ما نريد أن نجعلها تقوله". (حجر: 2000، 123)

خلاصة وتوصيات:

يعتبر الاستبيان وسيلة من وسائل جمع البيانات والمادة العلمية في الدراسات الاجتماعية المستعملة على نطاق واسع وشائع، ورغم الانتقادات الكثيرة الموجهة لهذه الأداة إلا أن أهميتها تكمن في سمة الاقتصاد في الجهد والوقت إذا ما قورنت بغيرها من وسائل جمع البيانات.

لذا نجد أغلب البحوث والدراسات العربية تعتمد على الاستبيان كوسيلة لجمع البيانات، بالرغم مما تنطوي عليه هذه الوسيلة من عيوب يمكن أن تؤثر على نتائج الدراسة، وأول عائق يواجهه الباحث هو عدم الاستجابة على الأسئلة من قبل المبحوثين والتي قد تكون كلية أو جزئية، إما بدافع الخوف أو الخجل أو بسبب الجهل واللامبالاة، مما يؤدي إلى زيادة في حجم البيانات المفقودة التي تؤثر على النتائج.

في الغالب يتم التعامل مع الأسئلة غير المُجابهة، أو القيم المفقودة، بالإهمال والتجاهل، مما قد يؤدي إلى تقديرات ذات كفاءة أقل، حيث يقوم الباحثون بإلغاء هذه الاستثمارات وهذا بدوره يؤدي إلى ضياع الكثير من الجهد، بالرغم من وجود بعض الطرق الإحصائية التي تتعامل مع البيانات المفقودة.

كما أنه من المشاكل العويصة في البحوث التي تعتمد على أداة الاستبيان في جمع البيانات هو ميل بعض المستجيبين إلي تقديم معلومات غير دقيقة أو معلومات جزئية، أو قد يخشون من التعبير الصريح عن آرائهم ومواقفهم نتيجة لاعتبارات معينة، ما يؤثر على الإجابة المعطاة لبعض النماذج من الأسئلة. كما أن طبيعة الشخص المبحوث وظروفه النفسية تؤثر في استجابته ورغبته في الإجابة على الأسئلة أو في صدقه في إعطاء الإجابات.

كما تعتمد عملية بناء الاستبيان على القيام بمجموعة من الخطوات المتسلسلة التي تؤدي في النهاية إلى تجنب كثير من الأخطاء، وتتيح إمكانية إعداد استبيان جيد يعتمد عليه. وتتوقف سلامة بناء الاستبيان على ضمان صدق الأداة وثباتها، فإذا لم يتم تحديد درجة صدق وثبات الاستبيان الذي يتم بناؤه فإن ذلك سيقبل من قيمة النتائج التي يتم التوصل إليها في الدراسة.

إلا أن هناك بعض الأخطاء التي يقع فيها الكثيرون عند تقدير صدق و ثبات الاستبيان منها الاكتفاء بصدق المحكمين فقط، باعتباره من أكثر طرق الصدق شيوعاً وسهولة، وأشهرها استخداماً لدى الباحثين، لكنها ليست دقيقة، لأن بعض المحكمين قد لا يكون مخلصاً أميناً في تحكيم الاستبيان، لذا لا بد من تدعيمه بأنواع أخرى من أنواع قياس الصدق التي قمنا بذكرها سالفاً.

كما قد يقوم بعض الباحثين بإعادة تحكيم مقاييس مقننة على نفس البيئة التي سيتم التطبيق عليها، والأفضل أن يكتفي بذكر الخصائص السيكومترية للاستبيان، ذلك أن التحكيم يكون فقط للاستبيانات التي تتأثر بالثقافة ويتم تطبيقها على بيئة مختلفة عن البيئة التي طبقت فيها أول مرة.

يقوم بعض الباحثين أيضاً بحساب الخصائص السيكومترية للاستبيان على العينة الاستطلاعية بعد تطبيقها على عينة الدراسة الأساسية. حيث يقومون باقتصاص جزء من العينة النهائية التي طبق عليها الاستبيان، وحساب الخصائص السيكومترية، بالرغم من أن هناك خطوات علمية يجب إتباعها، وقد تم ذكرها آنفاً.

يلجأ البعض إلى تطبيق الاستبيانات على عينة كبيرة من أجل الحصول على ثبات عال، مع أن الصدق والثبات لا يتأثران بحجم العينة فقط، بل هنالك عوامل مجتمعة تؤثر على الصدق والثبات، كذلك عدم اختيار عينة عمدية بل يجب أن تكون العينة عشوائية.

في طريقة المقارنة الطرفية يتم تقسيم العينة إلى مرتفعين ومنخفضين وحساب الفرق بينهما، حيث أن تقسيم عينتين إلى مرتفعين ومنخفضين حتما سيعطينا فرقا، وبالتالي اعتبار الاختبار صادقا، لذا يفضل حساب المقارنة الطرفية باستخدام محك آخر، كما وضحنا في حساب صدق المقارنة الطرفية.

زيادة عدد البنود رغبة في الحصول على ثبات مرتفع، بالرغم من أن هناك حدا معيناً لا يجب تعديده، فالشخص المجيب قد يصاب بالملل. لذا يجب أن تغطي الأسئلة جميع محاور الاستبيان وخصائصه وسماته، كما يجب أن يكون هناك توازن في توزيع الأسئلة على الأبعاد.

زيادة عدد خيارات الإجابة: يمكن زيادة عدد خيارات الإجابة من قياس ثبات الاستبيان بدرجة أكبر، فبدلاً من الاعتماد على بديلين فقط ب(نعم أو لا)، يمكن الاعتماد من خلال مقياس ذي أربع درجات، أو خمس درجات، أو سبع.

ومن خلال العرض الذي تم تقديمه يمكن التقدم بعدد من التوصيات للباحثين منها:

- إتباع الخطوات المنهجية في بناء الاستبيان، فعدم تصميم الاستبيان تصميمًا دقيقاً يمكن أن يؤدي إلى عدم دقة إجابات المفحوص.
- لا بد من اختيار عينة ممثلة ومناسبة لبنود وعبارات الاستبيان حتى تكون البيانات المتحصل عليها ذات فائدة.
- التأكد من ثبات الاستبيان وصدقه من أجل ضمان استقرار النتائج لاستخدامها في قرارات دقيقة وموثوق منها.
- يجب أن يدرك الباحث أن معامل الصدق والثبات يعود إلى البيانات التي يجمعها، وبالتالي فالصدق والثبات مرتبطين بالبيانات نفسها وليس بأداة القياس.
- من الضروري أن يطلع الباحث على أساليب تقدير الصدق والثبات المختلفة كي يتم استعمالها استعمالاً سليماً، بدلاً من الاهتمام بالبرامج الإحصائية دون أي خلفية سيكومترية.
- خلاصة لما سبق يمكن القول أن الاستبيان يعد من بين أهم أدوات جمع البيانات في البحوث الاجتماعية، وبالرغم من كثرة استخدامه إلا أن كثرة عيوبه تفرض على الباحث أن يستخدمه بنوع من الحذر والتنبه حتى لا يقلل من مصداقية بياناته التي يجمعها عن طريق الاستبيان، ويبقى السؤال المطروح هو: كيف يمكننا الوصول باستخدام الاستبيان إلى مستويات أعلى من الدقة والمصداقية في البحوث الاجتماعية؟، وإذا كان الاستبيان غير قادر على تقديم معلومات دقيقة وموثوقة، هل بالإمكان اليوم تقديم أدوات جديدة تمكننا من الاستغناء عن هذه الأدوات المستخدمة في بحثنا اليوم؟

المراجع والهوامش

- زياد بن علي بن محمود الجرجاوي (2010): القواعد المنهجية التربوية لبناء الاستبيان، الطبعة الثانية، مطبعة أبناء الجراح، القدس.
- أحمد إبراهيم خضر (2013): إعداد البحوث والرسائل العلمية من الفكرة حتى الخاتمة، قسم الخدمة الاجتماعية وتنمية المجتمع، جامعة الأزهر.
- السيد محمد أبو هاشم حسن (2006): الخصائص السيكومترية لأدوات القياس في البحوث النفسية والتربوية باستخدام SPSS، كلية التربية، جامعة الملك سعود.
- الضحيان، سعود بن ضحيان، الدليمي، عبد الله حمد (1998): المنهجية والرسائل الجامعية العربية دراسة حالة مجلة العلوم الاجتماعية - الكويت، مج 26، ع 4، ص 109 - 87.
- بدر محمد الأنصاري (2000): قياس الشخصية، دون طبعة، دار الكتاب الحديث، الكويت.
- حجر، خالد أحمد مصطفى (2000): بناء الاستبيان في البحث الاجتماعي: بعض الاعتبارات المنهجية والفنية، مجلة آداب - كلية الآداب - جامعة الخرطوم - السودان، ع 18، ص 119 - 153.
- رجاء محمود أبو علام (2005): مناهج البحث في العلوم النفسية والتربوية، الطبعة الخامسة، دار النشر للجامعات، القاهرة.
- سعد عبد الرحمن (1998): القياس النفسي (النظرية والتطبيق)، الطبعة الثالثة، دار الفكر العربي، القاهرة.
- شحاتة، سامية سمير (2011): مدى فعالية صدق المحكمين بالمقارنة بأنواع الصدق الأخرى: دراسة تحليلية، علم النفس، مصر.
- صفوت فرج (2007): القياس النفسي، الطبعة السادسة، مطبعة محمد عبد الكريم حسان، القاهرة.
- صلاح أحمد مراد، أمين علي سليم (2005): الاختبارات والمقاييس في العلوم النفسية والتربوية، دار الكتاب الحديث، الطبعة الثانية.
- صلاح الدين محمود علام (2000): القياس والتقويم التربوي والنفسى، الطبعة الأولى دار الفكر العربي، القاهرة.
- عباس محمود عوض (1998): القياس النفسي بين النظري والتطبيق، دار المعرفة الجامعية الإسكندرية.
- عبد الله محمد الشريف (1996): مناهج البحث العلمي، دليل الطالب في كتابه الأبحاث والرسائل العلمية، دون طبعة، مكتبة الإشعاع للطباعة والنشر والتوزيع، الإسكندرية.

- عفاف أحمد عويس(2000): المقاييس والاختبارات النفسية، دون طبعة، القاهرة.
- قدرى حفنى، سليمان خليل (1978):القياس النفسى، دار فينوس للطباعة والنشر، القاهرة..
- محمد عبد السلام غنيم (2004):مبادئ القياس والتقويم النفسى والتربوي، جامعة حلوان.
- محمود السيد أبو النيل(1987): الإحصاء النفسى والاجتماعى والتربوي، دار النهضة العربية، بيروت.
- مساعد بن عبد الله النوح (2004): مبادئ البحث التربوي، الطبعة الأولى، كلية المعلمين، الرياض.
- ناجى محمد حسن درويش:التقويم والقياس النفسى، جامعة 6 أكتوبر، كلية علوم اجتماعية، قسم علم النفس.
- أحمد عبد الرحمن إبراهيم عثمان: تقويم التحصيل الدراسى للطلاب الجامعى فى ضوء الاتجاهات العالمية المعاصرة، جامعة الزقازيق، مصر.
- أحمد محمد عبد الخالق(1996): قياس الشخصية، الطبعة الأولى، جامعة الكويت، الكويت.
- سعد عبد الرحمن (2008): القياس النفسى النظرية والتطبيق، الطبعة الخامسة، هبة النيل العربية للنشر والتوزيع، الجيزة.
- عبد الله بن أحمد آل شويل الغامدى(2008): أثر عدد البدائل فى الخصائص السيكمترية للاختبار التحصيلى فى الرياضيات، متطلب تكميلي للحصول على درجة الماجستير فى علم النفس تخصص (قياس وتقويم). جامعة أم القرى.
- مروان عبد المجيد إبراهيم (2000): أسس البحث العلمى لإعداد الرسائل الجامعية، الطبعة الأولى، مؤسسة الوراق، عمان.

-Hambleton, R. K. & Rogers, H. J. (1991). Advances in criterion-references measurement.(In R. K. Hambleton & J. N. Zaal (Eds.), Advances in educational and psychological testing: Theory and applications (pp. 3—43). Boston: Kluwer Academic.

-Lawshe, C. H. (1975). The quantitative approach to content validity .Personnel Psychology, 28, 563—575.

-Lynn, M. R. (1986). Determination and quantification of content validity .Nursing Research, 35, 382—385

-MANGAL S. K. , SHUBHRA MANGAL(2013): Research methodology in behavioural sciences, PHI Learning Pvt. Ltd.

-Tittle, C. K. (1982). Use of judgmental methods in item bias studies.(In R. A. Berk (Ed.), Handbook of methods for detecting bias (pp. 31—63), Baltimore: Johns Hopkins University Press.