الصدق والثبات في البحوث الاجتماعية

Validity and reliability in social research

د.بشتة حنان	د.بوعموشة نعيم
أستاذ محاضر أ	أستاذ مؤقت
جامعة محمد الصديق بن يحي جيجل	جامعة محمد الصديق بن يحي - جيجل
hananbechta@yahoo.com	naim.socio18@outlook.fr

ملخص:

يعتمد البحث دائما على القياس، وأيا كانت الأداة أو الأدوات التي يختارها الباحث عليه أن يراعي المعايير اللازمة لبنائها وصلاحيتها للقياس، إذ هناك خاصيتان مهمتان ينبغي أن تتوفر عليهما كل أداة قياس وهما الصدق والثبات. فالصدق يعتبر أهم وأكثر الخصائص السيكومترية أهمية، والذي يشير إلى المدى الذي تقيس فيه الأداة ما يراد قياسه. أما الثبات فيشير إلى مدى الدقة والاستقرار والاتساق في نتائج القياس، فهو المدى الذي تظل فيه أداة القياس ثابتة في قياس ما يراد قياسه، لو طبقت أداة القياس مرتين فأكثر على نفس الأفراد أو الأشياء لقياس الخاصية في مناسبات مختلفة.

الكلمات المفتاحية: الصدق، الثبات القياس، البحث العلمي، البحث الاجتماعي.

Abstract:

The research always depends on the measurement, and whatever instrument or tools the researcher chooses, he must observe the criteria necessary for its construction and its suitability for measurement, as there are two important characteristics that each measuring instrument must have, namely validity and reliabity. Truthfulness is considered the most important and most important psychometric properties, which indicates the extent to which the instrument measures what is intended to be measured. As for reliabity, it indicates the degree of accuracy, stability and reliabity in the measurement results. It is the extent to which the measuring instrument remains constant in measuring what is to be measured, if the measuring instrument is applied twice or more to the same individuals or things to measure the property on different occasions.

Keywords: Validity, reliability, analogy, scientific research, social research.

مقدمة:

تتميز البحوث التي تدرس سلوك الإنسان بالموضوعية من خلال اعتمادها على التقدير الكمي لعملية قياس ذلك السلوك، لأن غياب التقدير الكمي للسمات والخصائص والجوانب يؤدي إلى فقدان صفة الموضوعية، والصدق الخارجي (إمكانية تعميم النتائج) لأدوات القياس التي تقيس ذلك السلوك، فتبقى دراسة السلوك حينها في حدود الوصف البسيط والملاحظة التي تضفي عليها صفة الذاتية. ومع مطلع القرن التاسع عشر شهد مجال القياس والتجريب في ميدان العلوم الاجتماعية تطورا ملحوظا مما أدى إلى الاعتماد على أدوات القياس وتطورها، بالاعتماد على المقاييس والاختبارات التي من شأنها المساعدة على تكميم السلوك وإخضاعه للتقدير الكمي بهدف ليس الوصف فقط، بل تعدى الأمر إلى التحليل والاستدلال. ومن بين الإجراءات الواجب توفرها في هاته الأدوات كي نقول أنها صالحة للاستعمال وتكون هناك جودة في عملية القياس وثقة في النتائج المتوصل إليها، نجد هناك مؤشرين هما الصدق والثبات، واللذان يعتبران شرطين أساسيين للحكم على صلاحية الأدوات التي اعتمدت في عملية القياس. فإن لم يتم الحصول على البيانات بأدوات صادقة وموثوقة فستكون لدينا ثقة ضئيلة بالنتائج أو بالاستنتاجات القائمة على تلك النتائج. وعليه تهدف هذه الدراسة للتعرف على كل من الصدق والثبات وطريقة حسابهما أو التأكد منهما في البحوث الاجتماعية، لإعطاء مصداقية لأدوات البحث.

1. الصدق Validity:

1.1. مفهوم الصدق:

يعد الصدق من الشروط المهمة الواجب توفرها في أداة جمع البيانات، ويتعلق الصدق بالسؤال التالي: هل أنا أقيس ما نويت قياسه؟.

وعليه يقصد بالصدق "صلاحية الأسلوب أو الأداة لقياس ما هو مراد قياسه، أو بمعنى آخر صلاحية أداة البحث في تحقيق أهداف الدراسة، وبالتالي ارتفاع مستوى الثقة فيما توصل إليه الباحث من نتائج بحيث يمكن الانتقال منها إلى التعميم". (المشهداني، 2019، ص167.)

فالصدق يعني "صدق أسئلة الاختبار من حيث صياغتها ومحتواها وطريقة تطبيقها على المبحوثين لتحقيق الهدف من الاختبار". (عبد الرحمن، البدوي، 2007، ص345.)

والصدق أيضا "يعني صلاحية الأداة لقياس ما وضعت من أجل قياسه وصدقها في قياس السمة أو السمات التي يريد الباحث قياسها فعندما يكون الغرض قياس التحصيل في مادة معينة فإن صدق الأداة يعني أن الأداة صالحة لقياس التحصيل في تلك المادة وتوفير البيانات اللازمة عن

التحصيل، فالصدق يعني أن يكون المقياس صالحا لقياس الظاهرة أو السمة التي يراد قياسها". (عطية، 2009، ص108.)

ومن خلال التعاريف السابقة يتضح أن الصدق هو قدرة الأداة على قياس ما أعدت لقياسه فعلا. ويرى كثير من الباحثين أن صدق نتائج الاختبار تعتمد في جزء كبير منها على الصيغة التي يصاغ بها الاختبار وعلى الطريقة التي ينفد بها.

وقد حدد "جينجز" بعض الشروط التي ينبغي على أخصائي الاختبار السوسيومتري مراعاتها لضمان صدق بيانات الاختبار وهي: (عبد الرحمن، البدوي، 2007، ص346.)

"- صياغة كلمات وعبارات الأسئلة بدقة وسهولة بحيث تتاح للمبحوثين الفرصة لفهم الاختبار.

- إتاحة الوقت الكافي للإجابة على الاختبار.

- تحديد مسار السؤال للمبحوثين لكي يحدد المبحوث موقفه مقدما للاتجاه الذي يسير فيه الاختبار.

- تقديم الموقف الذي يدور حوله الاختبار بشيء من الاهتمام والحماس.

- ضرورة استيعاب الباحث الذي يستخدم الاختبارات السوسيومترية لجميع إجراءات الاختبارات لمواجهة المفاجآت بأسرع ما يمكن".

2.1. أنواع الصدق:

يتحدد صدق أداة البحث عادة من خلال العلاقة بين أداء المستجيب عليها وبين وظيفة تلك الأداة، ويمكن الحصول على عدد من المؤشرات التي تعزز صدق الأداة بعدة طرق، منها:

1.2.1. صدق المحتوى أو المضمون content validity:

"ويعني الدرجة التي يقيس بها الاختبار المحتوى المراد قياسه. ويتطلب صدق المحتوى شيئان هما: صدق الفقرات وصدق المعاينة. ويهتم صدق الفقرات في هل فقرات الاختبار تمثل المحتوى أم لا؟ أما صدق المعاينة فيهتم فيما إذا كانت عينة الاختبار شاملة للمحتوى أم لا؟". (الضامن، 2007، ص113.)

و"يعتمد هذا الأسلوب على مدى تمثيل بنود الاختبار تمثيلا جيدا للمجال المراد قياسه. ولذلك فإن الحصول على صدق الاختبار من خلال هذا الأسلوب يتوقف على تحديد المجال المراد قياسه تحديدا جيدا ثم بناء مجموعة من الأسئلة التي تغطي هذا المجال يعقب ذلك اختيار عينة ممثلة من هذه المجموعة". (المشهداني، 2019، ص168.)

وعليه فإن صدق المحتوى يتناول فقرات الأداة ومحتوياتها من حيث ترتيبها وعددها وتمثيلها للجوانب والأبعاد المراد دراستها تمثيلا جيدا، وفقا للوزن النسبي أو درجة الأهمية لكل جزء منها. بمعنى

أن بنود المقياس أو الاختبار أو فقراتها تعبر عن الظاهرة أو السمة أو الموضوع الذي يراد قياسه بدقة، وأن الأداة في ذاتها تنتمي إلى الموضوع الذي يراد فحصه وتصلح لقياسه.

"ويسمى هذا الصدق أيضا الصدق المنطقي ويحسب بفحص محتوى الاختبار وتحليل أسئلته لمعرفة مدى تمثيلها للسلوك الذي يقيسه الاختبار والتأكد من أن الأسئلة تغطي جميع جوانب السلوك، ويتطلب ذلك ما يلي:

- تحديد الأهداف والتأكد من أن الاختبار يضم أسئلة تغطي جميع هذه الأهداف.
- وصفا تفصيليا للمحتوى الذي يقيسه الاختبار.
- وواقع الأمر أن هذا النوع من الصدق هو الأكثر ملائمة للاختبارات التحصيلية". (نوفل، أبو عواد، 2010، ص271.)

"ويتم التحقق من صدق المحتوى عن طريق ملاحظة فقرات الأدوات وفحصها والتأكد من أن كل فقرة من تلك الفقرات معنية بقياس هدف من الأهداف التي يراد قياسها وتمثل جانبا من الجوانب التي يراد قياسها. وهذا يعني أن صدق المحتوى يتطلب توافق بنود الأداة أو فقرات الاختبار مع الموضوع الذي يراد قياسه، وأن تتوزع هذه البنود بين جوانب الموضوع الفرعية توزيعا مناسبا وتتوقف درجة صدق الأداة على مدى تمثيل فقراتها لما يراد قياسه. وهناك أكثر من طريقة للتأكد من صدق المحتوى لعل من أكثرها شيوعا عرض الأداة على مجموعة من المحكمين من ذوي الدراسة في مجال البحث وبنا المقاييس للحكم على مدى صلاحية الفقرات أو البنود التي تتضمنها الأداة لقياس ما وضعت من أجله. ويتم حساب معامل الصدق عن طريق نسبة الاتفاق بين المحكمين على فقرات المقياس بموجب المعادلة الآتية: $P = \frac{NP}{NP+NNP}$

على أن:

P: يعني معامل الاتفاق.

NP: يعني عدد مرات الاتفاق أو عدد الفقرات المتفق عليها.

NNP: يعني عدد مرات عدم الاتفاق أو البنود التي لم يتفق عليها". (عطية، 2009، ص109.)

2.2.1. صدق المحكمين trustees validity:

يعتبر صدق المحكمين أو استطلاع آراء المحكمين الخبراء من أكثر طرق الصدق شيوعا وسهولة وأشهرها استخداما لدى الباحثين. "ويتم الحصول على صدق المحكمين عن طريق عرض الاختبار على مجموعة من المحكمين المختصين في المجال وذلك للتأكد من سلامة صياغة البنود من ناحية ومدى مناسبتها للمجال المراد قياسه من ناحية أخرى". (المشهداني، 2019، ص168.)

وعليه فصدق المحكمين هو "أن يختار الباحث عددا من المحكمين المتخصصين في مجال الظاهرة أو المشكلة موضوع الدراسة، ويطلب منهم تصحيح الفقرات أو الحكم عليها بأنها مرتبطة بالبعد الذي يقيسه أم غير مرتبطة". (نجم، 2015، ص51.)

3.2.1. الصدق الظاهري face validity:

"يعتمد الحصول على صدق الاختبار في هذا الأسلوب على مظهر بنود الاختبار، فإذا ارتبطت هذه البنود بالسلوك أو السمة أو المجال المراد قياسه كان ذلك دليلا على صدق البنود. ويعتبر هذا الأسلوب في الصدق من أقل الأساليب قيمة ذلك أنه لا يمكنه تفسير النتائج في المجال المستخدم". (المشهداني، 2019، ص168.)

ويشير مصطلح الصدق الظاهري إلى "الدرجة التي يقيس بها الاختبار ما يفترض قياسه وهو إجراء أولي لاختبار المقياس وصدق المحتوى هام في الاختبارات التحصيلية. والعلامات على الاختبار لا تعكس بدقة تحصيل الطالب إذا لم تقيس ما يفترض أن يتعلمه الطالب. ويبدو أن صدق المحتوى يشكل مشكل في كثير من الدراسات البحثية، فكثير من الدراسات يجري تصميمها لمقارنة فعالية طريقتين أو أكثر في التعليم لشيء واحد". (الضامن، 2007، ص.ص 114.113)

ومنه "تكون الأداة صادقة إذا كان مظهرها يشير إلى ذلك من حيث الشكل ومن حيث ارتباط فقراتها بالسلوك المقاس، فإذا كانت محتويات الأداة وفقراتها مطابقة للسمة التي تقيسها فإنها تكون أكثر صدقا. وهذا النوع من الصدق ليس صدقا حقيقيا إلا أنه ينال ثقة المستجيبين وتعاونهم مع الباحث". (نوفل، أبو عواد، 2010، ص271.)

وعليه تبدو الإشكالية الرئيسية للصدق الظاهري في عدم وجود إجراءات دقيقة وقابلة للتكرار لتقييم أداة القياس، مع صعوبة تكرار إجراءات التقييم هذه بدقة فإن الباحث يعتمد كليا على الأحكام الذاتية.

4.2.1. الصدق البنائي construct validity:

"نلجأ إلى الصدق البنائي عندما نريد أن نقيس مفاهيم معينة، والمقصود بالمفاهيم البنائية المفاهيم التي نصف رؤيتها أو سماعها أو لمسها مثل مفهوم الذكاء أو مفهوم تقدير الذات. فالشخص الذي يبني الاختبار يمكن أن يتوقع أن الطالب الذي لديه تقدير ذات عالي يمكن أن يدافع عن نفسه إذا تعرض لموقف ناقد، فالعلاقة بين مفهوم الشخص عن نفسه أو تقديره لذاته مع سلوكه هو ما ينظر إليه على أنه صدق بنائي أو صدق المفهوم". (الضامن، 2007، ص114.)

فالصدق البنائي يعني إلى أية درجة تؤكد نتائج تطبيق الأداة صحة الافتراضات المستخلصة من النظرية حول مفهوم السمة التي وضعت لقياسها. وهناك عدة طرق للتحقق من صدق البناء لاختبار ما منها: (عباس وآخرون، 2014، ص.ص 265.264)

"- مقياس التحليل لجوتمان Guttmann: فإذا كان بالإمكان ترتيب فقرات الاختبار خطيا أو هرميا فإنه يمكن التنبؤ بدرجة الفرد على الاختبار من معرفة هذا الترتيب، وفي الحالات التي لا يمكن التنبؤ فيها بدرجات الأفراد فإن ذلك قد يعود إلى أن التصنيف الهرمي غير دقيق.

- التحليل العاملي Factotial Analysis: حيث يتم استخراج دلالات عن صدق البناء بهذه الطريقة، من خلال مصفوفة الارتباطات بين الدرجات على فقرات الأداة في محاولة لإنقاص عدد العوامل أو المكونات التي تتجمع حولها فقرات الاختبار.

- المقارنات الطرفية: وتستخدم هذه الطريقة في حالة الرغبة في تعرف مدى قدرة الاختبار على التمييز بين المستويات المختلفة للسمة، أو التمييز بين المجموعات التي تمتلك درجات مرتفعة من السمة وتلك التي تمتلك درجات منخفضة منها، فقد تتضمن افتراضات السمة أو المفهوم وجود مجموعات مختلفة في درجة امتلاكها للسمة التي يقيسها الاختبار، فإذا كشفت نتائج الاختبار عن وجود فرق حقيقي بين المجموعات فإن هذا مؤشر لصدق البناء للاختبار".

5.2.1. الصدق التلازمي concurrent validity:

ويطلق عليه صدق المحك أيضا، ويعني "وجود ترابط بين الأداة والمقياس وأداة أخرى ثم التأكد من صدقها وثباتها أي بيان مدى اتفاق نتائج تطبيق الأداة الحالية مع نتائج محك آخر خارجي. فإذا ما قام الباحث بتطبيق الأداة أو المقياس الذي أعده على مجموعة أو عينة وطبق عليها في الوقت نفسه أداة أخرى تم التثبت من صدقها غير أنها أشمل وأكثر تعقيدا ووجد أن هناك اتفاقا بينهما من ناحية قياس سمات قائمة بالفعل عن طريق حساب معامل الثبات بين الأدائين، فهذا يعني أن الأداة التي بناها صادقة". (عطية، 2009، ص110.)

"ويتم الحكم على صدق الاختبار من خلال الارتباط بينه وبين أدوات أخرى صادقة على أن يتم تطبيق الأداتين في الفترة الزمنية نفسها أو في فترات متقاربة، وذلك بتطبيق الأداة المراد التحقق من صدقها وأداة أخرى تتمتع بدرجة عالية من الصدق وذات علاقة بالسمة المراد قياسها على مجموعة من المفحوصين، ثم يتم إيجاد معامل الارتباط بين نتائج المفحوصين على الأداتين، فإذا كان معامل الارتباط مرتفعا حكم بصدق الاختبار وإلا فالاختبار غير صادق. مثال ذلك عندما يطبق الباحث مقياسا لتقدير الذات على عينة من طلبة الصف الثامن الأساسي ويقارن درجاته على المقياس

بتقديرات مدرسيهم لسلوكهم التكيفي في الوقت نفسه، فإنه يحاول الحصول على دلالة للصدق التلازمي للمقياس". (عباس وآخرون، 2014، ص263.)

6.2.1. الصدق العاملي factotial validity:

يعتمد هذا الصدق على التحليل العاملي الذي يعتبر طريقة إحصائية لقياس العلاقة بين مجموعة من العوامل وفقا لما يلي: (عباس وآخرون، 2014، ص265.)

"- يطبق الباحث مجموعة أدوات تقيس السمة نفسها على عدد من المستجيبين.

- يحسب معامل الارتباط بين كل أداة والأدوات الأخرى.

- إذا وجد أن هناك معامل ارتباط عال بين أداتين منها فإن ذلك يعني أن هناك سمات مشتركة بين هاتين الأداتين، ويمكن وضعهما تحت عامل مشترك واحد يشملهما معا، ويتم إعطاؤه اسما".

"ويمكن حساب الصدق العاملي عن طريق إيجاد معامل الارتباط بين فقرات الأداة الواحدة، كما يمكن حسابه عن طريق إيجاد معامل الارتباط بين كل فقرة وبين الأداة ككل. ولدا فإن الفقرة تكون صادقة إذا كان معامل الارتباط بينها وبين الأداة الكلية عاليا". (نوفل، أبو عواد، 2010، ص.ص 275.274)

7.2.1. الصدق التنبؤي predictive validity:

المقصود بالصدق التنبؤي هو "أن يكون هناك علاقة بين الاختبار والسلوك الحقيقي الذي نريد التنبؤ به، فإذا كانت نتائج الاختبار يمكن توقعها في التنبؤ بالسلوك أو الأداء، فإن الباحث يستطيع تقييم الصدق التنبؤي عن طريق الربط بين الأداء على الاختبار والسلوك". (الضامن، 2007، ص114.)

"ويقدر الباحثون الصدق التنبؤي من خلال التنبؤ بالنتائج التي يتوقعون الحصول عليها باستخدام قياس خارجي يعرف بالمعيار، وبمقارنة مخرجات أدوات القياس المستخدمة مع مخرجات أدوات قياس أخرى. بمعنى آخر الصدق التنبؤي هو عبارة عن درجة الارتباط، المعروفة باسم معامل الارتباط، بين نتائج أدوات القياس المستخدمة ومعيار خارجي". (فرانكفورت، ناشيماز، 2004، ص174.)

ومن أجل اختبار الصدق التنبؤي لابد من اهتمام الباحثين بقضيتين عامتين: الأولى تتعلق بضرورة استخدام أداة قياس بدلا من المعيار فقط، أما القضية الثانية فتتعلق بصدق المعيار.

فإذا ما أراد الباحث قياس الصدق التنبؤي لاختبار قام ببنائه لقياس مدى استعداد الطلبة في مادة دراسية معينة فإنه يقوم بالآتي: (عطية، 2009، ص.ص 111.110)

"- يطبق الاختبار على مجموعة من الطلبة ويرصد درجاتهم عليه.
- عندما تؤدي هذه المجموعة امتحانا شهريا أو نصف السنة أو آخرها ترصد درجاتهم.
- يحسب معامل الارتباط بين درجات الطلبة على الاختبار الذي أعده وبين درجاتهم التي حصلوا عليها في امتحان المادة فإن وجد أن هناك معامل ارتباط يزيد على 0.80 فهذا يعني أن الاختبار الذي بناه له قدرة على التنبؤ، وكثيرا ما يستخدم الصدق التنبؤي لاختبار الاستعداد في المجال التربوي".

وعليه يعتبر هذا النوع من الصدق من أهم أنواع الصدق حيث أنه يساعد في توفير الوقت والجهد والمال. "ويتم حسابه عن طريق إيجاد قوة العلاقة أو الارتباط بين الدرجات على الأداة المراد التحقق من صدقها ودرجات محك آخر تجمع عنه المعلومات بعد فترة زمنية طويلة نسبيا، فيطبق الباحث الاختبار مثلا ثم يتابع سلوك المستجيبين فيما بعد فإذا اتفق مستوى أدائهم على الاختبار مع سلوك آخر يتصل بما قاسه الاختبار فإن لهذا الاختبار قدرة تنبؤية، فالباحث الذي يريد أن يقيس القدرة اللغوية لدى الأطفال فإنه يطبق الاختبار ثم يتابع ملاحظة كلام الأطفال وقدراتهم اللغوية أثناء حديثهم وألعابهم، فإذا كانت نتائج الاختبار متفقة مع ملاحظات الباحث فإن الاختبار يتمتع بالصدق التنبؤي أي أنه يستطيع التنبؤ بسلوك الأطفال في المستقبل". (عباس وآخرون، 2014، ص263.)

إذ "يقاس الصدق التنبؤي عادة بمعامل الارتباط، مثل معامل ارتباط بيرسون والذي تتدرج قيمه من -1 إلى +1 وكلما كان الارتباط عاليا كلما كان الصدق أفضل. ويعتبر البعض أن الصدق التنبؤي يكون ضعيفا إذا كان معامل الارتباط حوالي (0.3)، ويرى البعض الآخر أن أقل معامل ارتباط يفترض أن يكون بحدود (0.4) أو أكثر". (الضامن، 2007، ص114.)

3.1. العوامل المؤثرة في صدق الأداة:

يتأثر صدق أداة القياس بعدد من العوامل يمكن تصنيفها ضمن الفئات التالية: (نوفل، أبو عواد، 2010، ص.ص 276.275)

أ). عوامل تتعلق بالأداة نفسها:

- لغة الأداة فإذا كانت فوق مستوى المفحوص فإنه سيعجز عن فهم السؤال، وبالتالي لا يجيب عنه مما يؤثر في أدائه على المقياس.
- غموض الفقرات مما يجعل المفحوص يفسرها تفسيرات متباينة ويجيب عنها إجابات خاطئة فيؤثر ذلك في مستوى أدائه على المقياس.
- سهولة الفقرات أو صعوبتها تجعل المفحوص يحصل على علامات لا يستحقها ولا تمثل مستوى قدرته الفعلية.

- صياغة الفقرات فقد تحمل بين طياتها أدلة ومؤشرات للإجابة.

ب). عوامل تتعلق بتطبيق الأداة وتصحيحها:

- العوامل البيئية تؤثر في أداء المفحوص فتقلل منه أو ترفعه، كالحرارة والبرودة والضوضاء والازدحام.
- وضوح الطباعة وإخراج الأداة وطريقة عرض الفقرات.
- التعليمات غير الواضحة التي لا تبين المطلوب من المفحوص.
- استخدام الأداة في غير ما ضعت له.
- عدم استخدام الأداة مع الفئة التي وضعت لها.

ج). عوامل تتعلق بشخصية المفحوص:

- اضطراب المفحوص وقلقه.
- التخمين أو الغش أو محاولة التأثير في الفاحص بطريقة ما".

كما تتمثل العوامل المؤثرة في الصدق أيضا في الآتي: (الدليمي، صالح، 2014، ص.ص 123.122)

"- طول الاختبار: يزداد صدق الاختبار تبعا لزيادة عدد الأسئلة، لأن زيادة عدد الأسئلة يؤدي إلى شمول الاختبار للمحتوى ويقلل أيضا من أخطاء القياس وبالتالي من الصدق.

- ثبات الاختبار: يتأثر صدق الاختبار بمعامل ثباته فمعامل الثبات المرتفع يزيد من احتمال الصدق لكنه لا يضمن ارتفاع الصدق، ومعامل الثبات المنخفض يدل على عدم الصدق. لكن الصدق يضمن الثبات، فالصدق الجيد يضمن ارتفاع معامل الثبات.

- ثبات المحك: إذا كان المحك جيدا (ثابتا وصادقا) فان ذلك يزيد من صدق الاختبار.

- تباين الدرجات: زيادة تباين درجات الاختبار تؤدي إلى ارتفاع الثبات والصدق، لأن تباين الدرجات يعني اتساع مجال السمة المقاسة مما يدل على تغطية شاملة لمجال السمة المقاسة".

2. الثبات Reliability

1.2. مفهوم الثبات:

يحظى الثبات باهتمام كبير من قبل الباحثين في العلوم الاجتماعية لأن أداة القياس المستخدمة نادرا ما تكون صادقة بشكل كامل، وفي كثير من الحالات لا يتوفر لدى الباحث أي دليل حول الصدق. لذلك يلجأ الباحثون إلى تقييم أداة القياس من ناحية خصائص أخرى وافتراض صدقها، ومن بين الطرق المستخدمة في ذلك هي درجة ثباتها.

"إن كلمة الثبات قد تعني الاستقرار، بمعنى أنه لو كررت عمليات قياس الفرد الواحد لأظهرت درجته شيئا من الاستقرار، كما أن الثبات قد يعني الموضوعية، بمعنى أن الفرد يحصل على نفس

الدرجة مهما اختلف الباحث الذي يطبق الاختبار أو الذي يصححه وفي هذه الحالة يكون الاختبار الثابت اختبار يقدر الفرد تقديرا لا يختلف في حسابه اثنان". (المشهداني، 2019، ص169.)

فالثبات يعني أن تكون النتائج التي تظهرها الأداة ثابتة، بمعنى تشير إلى النتائج نفسها لو أعيد تطبيقها على العينة نفسها في نفس الظروف بعد مدة زمنية ملائمة. فإذا لم تتغير النتائج بعد إعادة تطبيق الأداة، ولا تختلف استجابة المبحوثين فهذا يعني أن الأداة ثابتة.

والمقياس الثابت هو "المقياس الذي يعطي النتائج نفسها إذا قاس الشيء نفسه مرات متتالية تحت الظروف نفسها، أي عدم تناقض المقياس مع نفه، ولا تصل المقاييس النفسية إلى دقة مقياس الظواهر المادية المختلفة كالطول والوزن والزمن". (الدليمي، صالح، 2014، ص119.)

فالأداة الثابتة "هي الأداة التي تعطي نتائج متقاربة أو نفس النتائج إذا طبقت أكثر من مرة في ظروف متماثلة. ومن النادر أن يوجد مقياس صادق ولا يكون ثابتا، فالمقياس الصادق هو مقياس ثابت لكن العكس ليس صحيحا". (نوفل، أبو عواد، 2010، ص276.)

فإذا قام مثلا الباحث ببناء اختبار وطبقه على عينة معينة ورصد درجاتهم عليه، ثم قام بعد أسبوعين بتطبيق الاختبار نفسه على العينة نفسها في الظروف نفسها، ورصد درجاتهم عليه وحسب معامل الارتباط بين درجات العينة في التطبيق الأول وبين درجاتها في التطبيق الثاني، فوجد أن معامل الارتباط ذو قيمة عاليه، فهذا يعني أن الاختبار ثابت. أي أن ثباته يعني إعطاء النتائج نفسها لو أعيد تطبيقه على العينة نفسها. ويعبر عن الثبات بقيمة معامل الارتباط، فعندما تكون قيمة معامل الثبات عالية تكون درجة الثبات عالية.

2.2. الطرق الإحصائية لحساب الثبات:

هناك عدة طق تستخدم للتحقق من ثبات أداة القياس، منها:

1.2.2. طريقة إعادة الاختبار test retest method:

"تتسم هذه الطريقة بالبساطة والسهولة حيث تقوم على فكرة إعادة نفس الاختبار على نفس عينة البحث مرة أخرى بعد فترة زمنية يحددها الباحث. ويحسب معامل ارتباط درجات المرة الأولى، ومعامل ارتباط درجات المرة الثانية والمقارنة بينهما للحصول على معامل ثبات الاختبار". (عبد الرحمن، البدوي، 2007، ص344.)

بمعنى أن هذه الطريقة تقوم على تطبيق الباحث الاختبار أو المقياس الذي أعده على عينة من أفراد المجتمع المستهدف بالبحث ويرصد استجاباتهم عليه، ثم إعادة تطبيق نفس الاختبار أو المقياس على نفس العينة وفي الظروف نفسها بعد مضي فترة زمنية معينة –غالبا لا تقل عن الأسبوعين-

ويرصد استجاباتهم، وهكذا يحصل كل فرد على درجة في التطبيق الأول في الاختبار وعلى درجة أخرى في التطبيق الثاني للاختبار. ثم يحسب معامل الارتباط بين درجاتهم في المرتين، وكلما ارتفع معامل الارتباط دل ذلك على ثبات الأداة. "فبمعالجة هذه البيانات عن طريق معامل الارتباط الملائم نحصل على مقدار الثبات الذي يفضل أن لا يقل عن (0.80) لغرض قبول الاختبار واعتماده لأغراض جمع المعلومات". (عطية، 2009، ص207.)

وتتمثل عموما عيوب هذه الطريقة في: (نوفل، أبو عواد، 2010، ص.ص 277.276)

"- أن المستجيبين قد يتعلمون من التطبيق في المرة الأولى مما يفيدهم في المرة الثانية كما أن درجاتهم في المرة الثانية تتأثر بألفتهم لفقرات الأداة وانخفاض عامل التوتر وانتقال أثر التدريب.

- عندما تطول الفترة بين مرتي التطبيق فمن المتوقع أن يكون المستجيبون في المرة الثانية أكثر نموا ونضجا وخبرة ومعرفة.

- في حالة قصر الفترة الزمنية بين مرتي التطبيق فإن المستجيبين يتذكرون بعض أجزاء الأداة، ولدا يفضل أن لا تقل الفترة بين التطبيقين عن أسبوع مع محاولة ضبط العوامل الداخلية المؤثرة في التجربة".

2.2.2. طريقة التجزئة النصفية split-half method:

"في هذه الطريقة يقوم الباحث بتقسيم الاختبار (المقياس) بعد تطبيقه إلى نصفين متكافئين، وحساب معامل ارتباط كل نصف على حدى ثم الحصول بعد ذلك على معامل ثبات الاختبار ككل من خلال المعادلات الإحصائية المناسبة لذلك". (عبد الرحمن، البدوي، 2007، ص344.) حيث "يتم تصحيحه بمعادلة إحصائية مثل معادلة سبيرمان براون spearman brown أو جثمان guttmann أو رولون rolon، ويكون المقياس ثابثا إذا كان معامل الارتباط عاليا". (نوفل، أبو عواد، 2010، ص.ص 278.277)

وبعبارة أخرى "التجزئة النصفية تعني تقسيم بنود الاختبار على نصفين الأول يشتمل على البنود أو الأسئلة ذات التسلسلات أو الأرقام الفردية والثاني يشتمل على البنود ذات التسلسلات الزوجية لدا فإن هذه الطريقة صالحة لقياس الاتساق الداخلي للاختبار. ويحسب الثبات بموجب هذه الطريقة بإتباع الخطوات الآتية:

- يختار عينة من أفراد المجتمع المستهدف بالقياس.
- يقدم جزئي الاختبار إلى أفراد العينة للإجابة على بنودها (تطبيق الجزأين).

- يرصد درجات أفراد العينة على بنود كل من النصف الأول والثاني، فيكون لديه درجتان لكل طالب درجة على بنود النصف الأول وأخرى على بنود النصف الثاني.
- يحسب معامل الارتباط بين درجات الطلبة على النصف الأول ودرجاتهم على النصف الثاني". (عطية، 2009، ص208.)

فعلى سبيل المثال إذا كان الاختبار مكون من عشر أسئلة فيمكن اعتبار النصف الأول من الاختبار الأسئلة من رقم (1) إلى (5)، والنصف الثاني الأسئلة من رقم (6) إلى (10). وهناك أسلوب آخر لقسمة الاختبار بحيث يتكون النصف الأول من الأسئلة ذات الأرقام الفردية (1، 3، 5، 7، 9) والنصف الثاني الأسئلة الزوجية (2، 4، 6، 8، 10).

وتجدر الإشارة إلى أنه يمكن استعمال أربع معادلات في طريقة التجزئة النصفية هي: (الدليمي، صالح، ص.ص 121.120)

"- معامل ارتباط بيرسون: يقيس هذا المعامل في حالة التجزئة النصفية ثبات نصف الاختبار فقط، لهذا لقياس ثبات كل الاختبار نطبق معادلة سبيرمان براون.

- معامل ارتباط رولون.

- معامل ارتباط جتمان: يستعمل عندما تتساوى أو لا تتساوى الانحرافات المعيارية لجزئي الاختبار.

- معامل ارتباط جالكسون للاختبارات الموقوتة: ويرمز له (ر. جك) ويستعمل عندما لا يجيب المفحوص على كل أسئلة الاختبار. ومعادلته: ر. جك = ر. سب-

حيث: ر. جك: معامل الثبات بعد تصحيح السرعة، ر. سب: معامل ارتباط سبيرمان براون.

م ت: متوسط الأسئلة المتروكة، ع^{2}خ: تباين الخطأ (مجموع الاجابات الخاطئة والمتروكة)".

أما عن عيوب هذه الطريقة فتتمثل في: (عباس وآخرون، 2014، ص268.)

"- تقسيم المقياس إلى قسمين قد يؤدي إلى عدم تجانس القسمين الجديدين أو تكافؤهما.

- يقل صدق كل من المقياسين الجديدين لأن الفقرات لا تقيس جميع الأبعاد المتعلقة بالسمة.

- الناتج في هذه الطريقة هو معامل الثبات النصفي والذي يحتاج إلى تصحيح باستخدام معادلات إحصائية مثل معادلة سبيرمان براون: معامل الثبات المصحح = (معامل الثبات النصفي X 2) / (معامل الثبات النصفي + 1)".

3.2.2. طريقة الصور المتكافئة equivalent forms method:

"تعتمد هذه الطريقة على تكوين صورتين متكافئتين ومتماثلتين تماما للاختبار، وتطبيقهما في نفس الوقت على نفس الأفراد، ثم حساب معامل ارتباط الصورة الأولى من الاختبار

ومقارنتها بمعامل ارتباط الصورة الثانية، ويدل معامل ثبات الصورتين المتماثلتين على معامل ثبات الاختبار". (عبد الرحمن، البدوي، 2007، ص345.)

وبصورة أضح يجري حساب الثبات في هذه الطريقة بإتباع الإجراءات الآتية: (عطية، 2009، ص.ص 208.207)

- يعد الباحث صيغتين متكافئتين من الاختبار.
- يختار عينة من أفراد المجتمع المستهدف يفضل أن لا يقل عددها عن 30 فردا.
- يطبق الصيغة الأولى على أفراد العينة وبعد الانتهاء منها يطبق الصيغة الثانية عليهم، في جلسة أخرى أو في الجلسة نفسها.
- يصحح إجابات أفراد العينة على كل صيغة من الصيغتين، فتكون لديه درجتان أو علامتان لكل فرد.
- يحسب معامل الارتباط بين درجات أفراد العينة على الصيغة الأولى والثانية فإن وجده عاليا يمكن القول إن الاختبار ثابت ويصلح الاعتماد عليه، ومن عيوب هذه الطريقة صعوبة إعداد صيغتين متكافئتين".

وعليه في طريقة الصور المتكافئة يقوم الباحث بإعداد صيغتين متكافئتين أو متماثلتين من الاختبار أو المقياس من حيث جوهر البنود أو الأسئلة وعددها وقياسهما الظاهرة أو السمة نفسها، ووحدة بنيتهما وصعوبتهما وتعليماتهما، وأسلوب تصحيحهما وتفسير نتائجهما، ويختلفان في صياغة الأسئلة والبنود فقط. ثم يطبق الصورة الأولى والصورة الثانية، ثم يحسب معامل الارتباط بين درجات المستجيبين على الصورتين، فيكون هو معامل ثبات الأداة.

ومن بين عيوب هذه الطريقة: (نوفل، أبو عواد، 2010، ص277.)

"- صعوبة التأكد من تكافؤ صورتي الأداة.
- صعوبة توفير الظروف المتشابهة تماما.
- الجهد والوقت الذي يبذل في الإعداد والتطبيق".

3.2. العوامل المؤثرة في ثبات الأداة:

يتأثر ثبات الأداة أو الاختبار بعدد من العوامل، منها ما له علاقة بالأداة نفسها، ومنها ما له علاقة بالمجموعة التي تخضع للأداة. ومن أبرز هذه العوامل ما يلي: (عباس وآخرون، 2014، ص272.)

"- طول الاختبار: يزداد ثبات الاختبار بزيادة طول الاختبار لأن ذلك يعني قدرته على تمثيل عينة واسعة من السلوك الذي يقيسه ويقل الثبات إذا كان الاختبار قصيرا، ولدا يمكن رفع درجة الثبات بزيادة عدد الأسئلة في الاختبار شريطة ألا يكون ذلك مدعاة الملل المستجيبين.

- زمن الاختبار: يزداد الثبات بزيادة الوقت الذي يستغرقه المفحوص في أداء الاختبار، ويقل الثبات بانخفاض مدة الاختبار.

- تجانس المستجيبين: يزداد الثبات كلما قل تجانس المستجيبين لأنهم يحصلون على درجات متفاوتة، بينما يقل الثبات كلما ارتفع التجانس بين المستجيبين لتقارب درجاتهم وتغير ترتيبهم عند إعادة تطبيق الاختبار.

- مستوى صعوبة الاختبار: يقل ثبات الاختبار كلما ازدادت سهولة الاختبار لأن ذلك يفقده القدرة على التمييز، كما يقل الثبات إذا ازدادت صعوبة الاختبار لأن ذلك سيدفع المستجيبين إلى التخمين وفي حالتي السهولة أو الصعوبة يأخذ المستجيبون درجات متقاربة ويصبح من السهل عند إعادة الاختبار أن يتغير ترتيب درجاتهم ولدا يقل الثبات".

ويمكن أيضا الإشارة لعوامل أخرى تؤثر في ثبات الاختبار، نوجزها في النقاط التالية: (الدليمي، صالح، 2014، ص123.)

"- عدد البنود: حيث تزداد قيمة معامل الثبات بزيادة عدد الأسئلة (بنود) الاختبار.

- صياغة البنود: الأسئلة الموضوعية تزيد من معامل الثبات كما أن الأسئلة الغامضة والطويلة مثل أسئلة المقال تقلل من معامل الثبات.

- تباين الدرجات: أن الأسئلة منخفضة التباين (السهلة أو الصعبة) تؤدي إلى خفض معامل الثبات بينما الأسئلة المرتفعة التباين (متوسط السهولة) تؤدي إلى زيادة معامل الثبات.

- زمن أداء الاختبار: زيادة الزمن تؤدي بالفرد إلى الحصول على أعلى درجة متسقة مع قدرته، إلا أن زيادة الزمن بدرجة أكبر من اللازم قد تؤدي إلى الارتباك في الإجابة ومن ثم خفض معامل الثبات.

- التخمين: يقل ثبات الاختبار بارتفاع نسبة التخمين (أو الغش) ولذلك فإن أسئلة الصواب أو الخطأ تقلل معامل الثبات عن أسئلة الاختيار من متعدد.

- الحالة الصحية والنفسية للفرد: تؤثر على معامل الثبات، فإذا كان متعبا أو مريضا أو متوترا فإن الثبات يقل".

ومن هذا المنطلق فإن مصادر عدم الثبات متعددة، وهذه المصادر هي: (الضامن، 2007، ص122.)

"- الفقرات الخاطئة faulty items: إذ تكون الفقرات أحيانا غامضة وخادعة،أو أن يتم عرضها بطريقة غير مشوقة، فإذا كانت الفقرة غامضة يضطر المستجيب لعملية التخمين، وكذلك الحال إذا كانت مكررة أو صيغت بطريقة غير صحيحة.

- صعوبة بالغة في الفقرات excessively difficult items: فإذا كانت الفقرات على درجة كبيرة من الصعوبة فإن هذا يجبر المستجيب على التخمين أيضا.

- سهولة بالغة في الفقرات excessively easy items: فإذا كانت جميع الأسئلة سهلة فكأن الاستبيان فقرة واحدة يجيب عليها المستجيبون بنفس الأسلوب.

- عدم ملائمة عدد الفقرات inadiquate number of items: إذ كلما كان عدد الفقرات قليلا كلما كانت أخطاء الصدفة كبيرة وهذا بدوره يزيد من عدم الثبات.

- عدم قياس الفقرات للخصائص المشتركة dissimilarity of the items: فإذا كانت الفقرات تقيس خصائص مشتركة، فإن الثبات سيكون عاليا، أما إذا كانت الفقرات تقيس خصائص مختلفة فإن الثبات سيكون منخفضا.

- خصائص المستجيبين characteristics of the respondents: يقل الثبات نتيجة الخصائص المؤقتة للمستجيبين والتي تتسبب في اختلاف الاستجابات لديهم، وتشمل هذه الخصائص عدم قدرة المستجيبين على التركيز في فترة زمنية معينة، وتقلب المزاج، وعدم الثبات في استدعاء المعلومات. وعلى سبيل المثال إذا اكتسب المستجيب اتجاه معين، فنحن نتوقعه أن يعكس ذلك الاتجاه على الاستبيان المقدم له.

- الإدارة السيئة للأداة faulty administration of the instrument.

- إجراءات التصحيح الخاطئة faulty scoring procedures: إذ من الممكن أن يكون التصحيح خاطئا، ومن الممكن أن تكون ورقة الأسئلة غامضة أو أن يكون حساب الدرجات خاطئا أيضا".

وعليه فإن أفضل طريقة لزيادة الثبات لأدوات القياس التي نستخدمها هي معرفة مصادر عدم الثبات، ولكي نزيد الثبات نتبع الخطوات التالية: (الضامن، 2007، ص120.)

"- استخدام فقرات لا يكتنفها الغموض وأفضل طريقة للتخلص من هذا الغموض هو إخضاع الأداة للفحص.

- اعمل على تقنين الإجراءات الإدارية للأداة.

- اعمل على تقنين إجراءات التصحيح.

- احرص على أن يكون عدد الفقرات كافية لتخفيف عامل الصدفة.

- لا تعكي أداة القياس للمستجيبين إذا كان مزاجهم متعكرا أو إذا كانوا في حالة قلق.

- تأكد أن الأداة تقيس الجوانب المعرفية لدى المستجيب وليس التخمين".

ومن المهم هنا أن نشير إلى أن هنالك علاقة واضحة بين الصدق والثبات، وكأن أحدهما متضمن في الآخر، ومن ذلك مثلا: (عمر، 2009، ص110.)

"- إن الاختبار الصادق لابد أن يكون ثابتا أيضا، لأنه لا يعقل أن يكون الاختبار متفقا مع وظيفة ما ولا يكون متفقا مع نفسه.

- إن معامل الصدق لا يمكن أن يزيد على معامل الثبات لأنه لا يعقل أن يكون الاختبار متفقا مع وظيفة ما بمقدار ما يكون متفقا مع نفسه.

- إن الاختبار قد يكون ثابتا ولا يكون صادقا لأنه يقيس وظيفة أخرى عدا الوظيفة المخصص لقياسها أو وظيفة أخرى إلى جانب الوظيفة المخصص لها".

خاتمة:

يعد الصدق والثبات من التقنيات التي يستعملها الباحث الجاد لإعطاء مصداقية لأدوات بحثه. إذ لابد أن تقيس أداة الدراسة ما يريد الباحث أن يقيسه، كما يجب أن تكون الإجابة عليها هي ذاتها نفس الإجابة لو تكرر إجراء القياس، ولهذا كان الصدق والثبات من أهم الشروط المنهجية في تصميم أدوات البحث. لدا يجب إعطاء الاهتمام بصدق أدوات جمع البيانات؛ أي ما إذا كانت تقيس حقا ما يفترض قياسه، وينبغي أن يكون للأداة صدق ظاهري إذ ينبغي أن يبدو صادقا من أجل هدفه المقصود، فالأفراد يميلون أكثر إلى الإجابة على أسئلة يدركون صلتها ومعناها أكثر من أسئلة لا يدركون هدفها. ولعل من بين الأنواع الأوضح للصدق هو الصدق المرتبط بالمضمون أو المحتوى، والذي يتم بمساعدة الباحثين المختصين، الذين يقومون بفحص البنود للحكم عما إذا كانت مناسبة لقياس ما يراد قياسه. وتجدر الإشارة هنا إلى أن هناك متغيران مهمان يؤثران على صدق الأداة، أولا ما مدى أهمية الموضوع للمبحوث؟ فيمكننا افتراض إجابات صادقة أكثر من أشخاص مهتمين بالموضوع أو مطلعين عليه، ثانيا هل تضمن الأداة إغفال شخصية المبحوث؟ فمن المعقول الافتراض أنه سيتم الحصول على بيانات صدق أكبر إذا كان بوسع المبحوثين البقاء مجهولين خصوصا إذا تم طرح أسئلة حساسة أو شخصية.

كما يجب أيضا أن تتمتع بيانات أداة الدراسة بالثبات لتحقيق الفائدة منها، فإذا لم تكن إجابات المبحوثين ثابتة فإن صدق البحث يكون موضع شك. ويمكن التحقق من الثبات الداخلي عن طريق بناء ما يزيد عن حاجة الأداة –بنود- حول الموضوع ذاته، والتي قد يعاد كتابتها وتكرارها في أداة الدراسة (الاستبيان أو المقابلة)، وكلما كانت الإجابات متسقة كان الثبات عاليا. كما يمكن أيضا تكرار الأداة (الاستبيان أو المقابلة) مع الأشخاص أنفسهم بعد فترة من الزمن، أو تطبيق شكلين مختلفين من الأداة نفسها على ذات الأشخاص، إلا أن ما يعاب على هذه الطريقة أنها باهظة في الغالب وتستغرق وقتا كبيرا في ذلك، وهي غير عملية نوعا ما إذ ليس من السهولة إيجاد أفراد يرغبون في تكرار

تطبيق الأداة، بالإضافة إلى أن بعض الإجابات على الأسئلة التي تعالج مظاهر سلوكية أقل استقرارا قد تتغير بشكل منطقي مع مرور الزمن.

قائمة المراجع:

- الدليمي عصام حسن وصالح علي عبد الرحيم (2014). **البحث العلمي أسسه ومناهجه**. ط1. عمان: دار الرضوان.
- الضامن منذر (2007). **أساسيات البحث العلمي**. ط1. عمان: دار المسيرة.
- عباس محمد خليل وآخرون (2014). **مدخل إلى مناهج البحث في التربية وعلم النفس**. ط5. عمان: دار المسيرة.
- عبد الرحمن عبد الله محمد والبدوي محمد علي (2007). **مناهج وطرق البحث الاجتماعي**. ط2. الإسكندرية: مطبعة البحيرة.
- عطية محسن علي (2009). **البحث العلمي في التربية مناهجه أدو اته وسائله الإحصائية**. عمان: دار المناهج.
- عمر سيف الإسلام سعد (2009). **الموجز في منهج البحث العلمي في التربية والعلوم الانسانية**. ط1. دمشق: دار الفكر.
- فرانكفورت شاقا وناشيمياز دافيد ناشيماز (2004). **طرائق البحث في العلوم الاجتماعية**. ترجمة ليلى الطويل. ط1. دمشق: بترا للنشر والتوزيع.
- المشهداني سعد سلمان (2019). **منهجية البحث العلمي**. ط1. عمان: دار أسامة.
- نجم طه عبد العاطي (2015). **مناهج البحث الإعلامي**. ط1. الإسكندرية: دار كلمة.
- نوفل محمد بكر وأبو عواد فريال محمد (2010). **التفكيروالبحث العلمي**. ط1. عمان: دار المسيرة.