

مشكلات في أعداد إجراءات البحوث التربوية والنفسية**أ.د. محمد أنور محمود****كلية التربية - ابن رشد/ جامعة بغداد****Problems in the procedures for the preparation of educational and psychological research****Prof. Mohammed Anwar Mahmood****College of Education - Ibn Rushd / University of Baghdad****Abstract**

The present study intends to highlight some of the problems faced by researchers in the area of educational and psychological studies regarding the preparation of research procedures.

The researcher deals with these problems as questions continuously raised about some of the important procedures when conducting researches, and as common errors used to be made by researchers either due to their ignorance of the correct procedure or depending on wrong procedures followed by some of the related previous studies.

The study also attempts to suggest scientific remedies for treating these problems and ultimately enabling researchers to handle their studies' procedures adequately to achieve a higher level of scientific adequacy

ملخص البحث

يهدف البحث الحالي الى الوقوف على بعض الإشكاليات التي تواجه الباحثين في مجال البحوث التربوية والنفسية عند أعدادهم لإجراءات البحث.

وتتاول الباحث هذه المشكلات بوصفها تساؤلات تثار باستمرار حول بعض الإجراءات الضرورية في البحث، أو أنها قد تكون أخطاءً شائعة درج بعض الباحثين على اعتمادها لعدم معرفتهم بالأجراء الصحيح من جانب، اعتمادهم على الإجراءات الخاطئة المتبعة في بعض الدراسات السابقة من جانب آخر.

كما يحاول البحث اقتراح الحلول العلمية التي تعالج هذه المشكلات وصولاً الى تمكين الباحثين من التعامل بدقة مع بعض المفاصل الإجرائية للبحوث بما يحقق مستوى أعلى من الدقة العلمية.

المحور الأول**التعريف بالبحث****مشكلة البحث:**

من خلال اطلاع الباحث على عدد كبير من البحوث والدراسات في مجال العلوم التربوية والنفسية، واشتراكه في لجان المناقشة لرسائل الماجستير وأطاريح الدكتوراه وفي تخصصات متباينة و تبين له أن بعض الباحثين يواجهون عدداً من المشكلات والأخطاء في التعامل مع بعض المفاصل الإجرائية في البحث، وهذه المشكلات أو الصعوبات أو الأخطاء الشائعة الاستخدام قد تكون متأتية بحسب اعتقاد الباحث من مصدرين أساسيين هما:

- النقص في الجانب المعرفي الخاص بالتعامل الصحيح مع هذه الإجراءات وتنفيذها بالأسلوب الأمثل.
- الاعتماد على بعض الجوانب الإجرائية المستخدمة في بعض البحوث والدراسات السابقة والتي تعد أحياناً تراكمياً من الأخطاء الإجرائية الشائعة الاستخدام.

وهذه المشكلات أو الأخطاء قد تكون سبباً مهماً في ابتعاد البحث عن الدقة العلمية المطلوبة وضعف الموثوقية بنتائجها، وذلك يتناقض مع ما يسعى إليه كل باحث علمي.

أهمية البحث:

- يمكن تلخيص الأهمية النظرية والتطبيقية للبحث الحالي في المحاور التالية:
- لا توجد دراسة سابقة (بحسب اطلاع الباحث) تناولت بعض الإشكاليات والأخطاء الشائعة في التعامل مع بعض المفاصل الإجرائية في مجال البحوث التربوية والنفسية.
- يوفر البحث الحالي قاعدة من المعلومات العلمية يمكن ان تسهم في زيادة الحصيلة المعرفية للباحثين الجدد.
- يمكن أن يساعد البحث الحالي التدريسيين المتخصصين في توجيه طلبة الدراسات العليا المشرفين عليهم في التعامل الصحيح مع بعض الجوانب الإجرائية في البحث.
- يمكن أن يوفر البحث الحالي إشارات بسيطة تعد بمثابة محفزات للباحثين نحو زيادة البحث والتقصي الدقيق من خلال الرجوع الى الأدبيات والمصادر العلمية الرصينة وصولاً الى تمكّن أفضل في التعامل مع الجوانب الإجرائية التي تناولها البحث.
- يمكن أن يعد البحث الحالي محفزاً للباحثين الآخرين للكشف عن مشكلات أو أخطاء إجرائية أخرى في مجال البحث العلمي.
- كما يمكن استثمار البحث الحالي بعده أداة لتقويم الإجراءات المتبعة في البحوث التربوية والنفسية.

هدف البحث:

يهدف البحث الحالي الى تحديد أهم المشكلات والأخطاء الشائعة التي يواجهها الباحثون عند أعدادهم لإجراءات البحوث التربوية والنفسية و محاولة رسم العلاجات الصحيحة لها.

المحور الثاني**خلفية نظرية:**

اعتمدت حركة القياس في علم النفس بشكل أساسي على مفهوم القياس النفسي ونظريته، إذ كان الهدف منه قياس الفروق الفردية بين الأفراد في القدرات العقلية أو السمات والخصائص الشخصية، والذي يتم فيه مقارنة درجة الفرد بدرجات مجموعته التي ينتمي إليها بالاعتماد على التوزيع الطبيعي للظواهر والخصائص النفسية، لذا سميت بالاختبارات والمقاييس معيارية المرجع Norm-Reference Test (Popham، 25، 1975)، إذ ان هدف القياس هنا إظهار الفروق الفردية في الخصائص والسمات النفسية المقيسة (Chiselli، al، 481، 1981).

وترجع نظرية القياس التقليدية Classical Test Theory الى نظرية الثبات التقليدية، لأن مهمتها الأساسية تقدير الثبات من خلال الدرجات الملاحظة للاختبار، أي أنها تحاول تقدير قوة العلاقة بين الدرجات الملاحظة والدرجات الحقيقية Theory Of True And Error Score (Allen & Yen، 25، 2002)، لذا فقد سميت هذه النظرية أيضا بنظرية الدرجة الحقيقية ودرجة الخطأ (Krus & Faller، 837، 1982).

وقد استندت هذه النظرية الى أنموذج للقياس ينص على أن لكل فرد قدرا ما من السلوك غير الملاحظ (الكامن) وهي (الدرجة الحقيقية) التي لا يمكن ان تقاس بصورة مباشرة، لذا فإن الدرجة الملاحظة وفق هذه النظرية لابد ان تتضمن قدرا ما من الخطأ (Hambleton & Zaal، 60-70، 1991)، وعليه يمكن القول ان الدرجة الملاحظة التي نحصل عليها من مقياس ما تكون مركبة في حقيقتها من:

1- الدرجة الحقيقية True Score: ان الدرجة التي نحصل عليها من تطبيق الاختبار تسمى بالدرجة الخام Raw Score، وهذه الدرجة هي خليط من الدرجة الحقيقية وخطأ القياس Measurement Error، ويترجح درجة خطأ القياس من الدرجة الخام نحصل على الدرجة الحقيقية وفق المعادلة التالية: (الدرجة الحقيقية = الدرجة الخام - درجة الخطأ)، وتعبّر الدرجة الحقيقية كذلك عن متوسط درجات الفرد الواحد اذا طبق عليه الاختبار

عددا لانهايا من المرات، فالدرجة الحقيقية قد نصل اليها عن طريق تطبيق مقياس كامل الدقة، (Paul، 630: 1991) لذلك فأن القياس في هذه الحالة يكون من غير أخطاء، وتعتبر الدرجة على الاختبار عن الدرجة الحقيقية او القدرة الحقيقية للفرد (عبد و عثمان، 1995: 223).

2- درجة الخطأ Error Score: والمقصود بالخطأ هنا هو الخطأ الملازم لعملية القياس، والذي لا يمكن استبعاده او تجنبه، وإنما يمكن تقديره، ويطلق على درجة الخطأ بخطأ القياس Error Of Measurement (Lyman، 62-63: 1963)، وتقدر درجة الخطأ بالفرق بين درجة الاختبار الملاحظة والدرجة الحقيقية (Suen، 28: 1990)، وهذه الدرجة يمكن ان تكون موجبة او سالبة او صفرا، فاذا كانت موجبة تزيد الدرجة الملاحظة عن الدرجة الحقيقية للفرد، وبالعكس اذا كانت درجة الخطأ سالبة (علام، 2006: 91)، ودرجة الخطأ قد تأتي من ثلاث مصادر أساسية هي:

- أ- الخطأ الثابت: وهو نوع من الخطأ الذي يعود الى المقياس نفسه ويتكرر بصورة منتظمة، وله التأثير نفسه في كل درجة، واذا عرفت كميته فإنه لا يشكل مشكلة مهمة بالنسبة الى عملية القياس.
- ب- خطأ المقياس: وهو الخطأ الناتج عن استخدام الدرجة الملاحظة في القياس بدلا من الدرجة الحقيقية، وهو نوع من الخطأ يحتاج الى معالجة إحصائية خاصة للتحكم فيه.
- ج- خطأ الصدفة او العشوائية: وهذا النوع من الخطأ بحكم التسمية لا يمكن ضبطه او السيطرة عليه تماما لأنه عشوائي، وهذه الأخطاء العشوائية هي التي يلغي بعضها البعض الأخر عند تكرار القياس (Chislli، 196، al.et 1981: 87: 89).

ونظرا الى ان التغيير غير المنتظم في درجات الأفراد لأية سمة عندما تقاس عددا من المرات او عند إعادة تطبيق الاختبار نفسه او اختبار مماثل على عينة الأفراد أنفسهم هي ظاهرة مألوفة في المقاييس النفسية، ولأن مصادر هذا التباين في الدرجات لا يمكن تحديدها بشكل دقيق، فأن توافر أدلة على مدى استقرار نتائج القياس وأتساقها يعد مسألة أساسية بالنسبة لأي اختبار او مقياس حتى يمكن استخدامه في قياس السمات والقدرات التي وضع لقياسها (Graham & Lilly، 1984:29)

من هذا المنطلق حاول المتخصصون في القياس النفسي تركيز اهتمامهم منذ بدايات الربع الأخير من القرن العشرين الى تحسين أدوات القياس النفسي ووسائله، بدلا من التركيز على الجانب الكمي للمقاييس و الاختبارات النفسية، وذلك من خلال الدراسات التي استهدفت تطوير المقاييس وحساب الخصائص السيكومترية لها ولفقراتها، والتي تسهم في اعداد مقاييس تقيس الظواهر والخصائص التي ترمي الى قياسها بأقل ما يمكن من أخطاء (الجباري، 2000: 20).

فالدرجة التجريبية للقياس في مجال الشخصية لا تمثل الدرجة الحقيقية للسمة المقاسة من غير وجود أخطاء فيها، قد تزيد او تقلل الدرجة الحقيقية التي نسعى لقياسها، ويعد المقياس واحدا من أهم مصادر هذه الأخطاء وأكثرها تأثيرا في الدرجة التجريبية (Lord، 128: 1960)، لذلك أتجه المتخصصون في القياس النفسي الى تحديد بعض الخصائص السيكومترية للمقاييس النفسية التي يمكن ان تحد من هذه الأخطاء، او تبعد المقياس عن إعطاء نتائج دقيقة في قياس السمة (Dick & Hagert، 13: 1971).

ومن أهم هذه الخصائص السيكومترية التي طورها المختصون في القياس النفسي خصيصة الصدق والثبات، أذ تعتمد عليهما دقة المعلومات التي توفرها المقاييس النفسية (علام، 2007: 231)، وقد أضاف (جackson Jackson، 677، Neill & Jackson، 1970: 677)، ويمكن توضيح هذه الخصائص بالإيجاز التالي:

أولاً: صدق المقياس: Scale Validity

يعد الصدق أهم الخصائص السايكومترية التي يجب ان تتوافر في المقاييس النفسية Ebel، (435: 1972)، لأنه مؤشر على قدرة المقياس في قياس ما أعد لقياسه (ملحم، 2005: 270)، ومن خلاله يتحقق من مدى قدرة المقياس على تحقيق الغرض الذي أعد من أجله (عودة، 1998: 333-335).

وتشير رابطة السيكولوجيين الأمريكية Amirecan Psycgological Association الى ان الصدق عبارة عن تجميع الأدلة التي يستدل بها على ان المقياس يقيس الغرض الذي أعد لقياسه (A. P. A، 9: 1985)، وبما أن صدق المقياس يحسب من استجابات الأفراد عند تطبيقه، لذا يكون صدق المقياس موقفياً اذ يتأثر بالعوامل المؤثرة في القياس، وعليه فأن صدق المقياس نوعي او محدد، بمعنى ان المقياس يقيس جانباً مما وضع لقياسه، كما انه محدد بطبيعة عينة السلوك التي تمثل الخبيصة او السمة المراد قياسها (ربيع، 1994: 92)، وعلى هذا فأن الصدق في حقيقته هو صدق درجات المقياس وليس صدق المقياس (عودة، 1998: 333-344).

وقد حددت رابطة السيكولوجيين الأمريكية الصدق بثلاث مؤشرات هي:

1- صدق المحتوى: Content Validity

ويقصد به الدرجة التي يقيس فيها المقياس ما صمم لقياسه في محتوى موضوع معين من خلال التحليل المنطقي لمحتوى المقياس، والتحقق من تمثيله للمحتوى المراد قياسه، لذا يطلق عليه احياناً بالصدق المنطقي Logical Validity، او الصدق بالتعريف Validity Of Definition (Gronlund، 67: 1981).

ويشير (ألدسون) Alderson الى ان الجانب الأساسي لهذا الصدق هو ان تكون عينة الفقرات مناسبة وممثلة لنطاق السلوك المراد قياسه (Alderson، 136: 1981)، ويسمى هذا الصدق في مجال قياس الشخصية بالصدق التمثيلي او الصدق العيني (ثورندايك وهيجن، 1989: 56).

2- الصدق المرتبط بمحك: Criterion – Related Validity

وهو أحد مؤشرات الصدق الذي يعين تجريبياً بناء على مؤشرات إحصائية أهمها الارتباط بين المقياس الجديد ومحك خارجي (Anastasi، 140: 1976)، وقد يكون المحك الخارجي مقياساً آخر، أو تقديرات شخصية مثل التقديرات التي يصدرها المدرسون او المشرفون او رؤساء المهنة... الخ، او تقديرات الأصدقاء والزملاء خاصة اذا كانت التقديرات نتاج متوسط تقديرات أكثر من فرد واحد (عبد الخالق، 1993: 158)، او التقدير الذاتي الذي يعتمد على قيام الفرد بتقدير سماته الشخصية او أدائه من خلال الحكم على تصرفاته كما يراها هو لا كما يراها الآخرون في ضوء مقياس متدرج، وقد يكون المحك مجموعة الأحكام التي أصدرها متخصصون في فترات طويلة ومتعاقبة بالنسبة لأنماط سلوكية يمكن الوثوق بها والاعتماد عليها (عبد الرحمن، 1998: 185).

ويقسم الصدق المرتبط بمحك الى نوعين:

أ- الصدق التنبؤي: Predictive Validity**ب- الصدق التلازمي: Concurrent Validity**

ويعود التمييز بينهما الى طريقة استخدام المحك، ففي الصدق التلازمي يطبق المقياس والمحك معاً، أما في الصدق التنبؤي فيطبق المقياس ثم الانتظار لمدة قد تطول او تقصر تبعاً لطبيعة السمة وهدف المقياس، وبعدها يطبق المحك، وفي كلا النوعين يحسب معامل الارتباط بين درجات المقياس ودرجات المحك الذي يمثل معامل الصدق التنبؤي للمقياس (Gronland، 83-84: 1976) (سمارة وآخرون، 1989: 112).

وترى (أنستازي) Anastasi ان الصدق التلازمي هو أنسب الأساليب وأكثرها ملائمة لاختبارات الشخصية من خلال حساب ارتباط درجة الفقرة بمحك داخلي او خارجي، وحيثما لا يتوفر محك خارجي مناسب فإن الدرجة الكلية للمجيب تمثل أفضل محك داخلي في حساب هذه العلاقة (Anastasi، 1976: 151).

3- صدق البناء: Construct Validity

يوصف صدق البناء بأنه أكثر أنواع الصدق تمثيلاً لمفهوم الصدق، والذي يسمى احياناً صدق المفهوم، او صدق التكوين الفرضي، ويقصد به مدى قياس المقياس النفسي لتكوين فرضي او مفهوم نفسي معين (ربيع، 1994: 98)، ويتم ذلك من خلال التحقق التجريبي من مدى مطابقة درجات المقياس للمفاهيم او الافتراضات التي أستند اليها الباحث في بناء الاختبار (أسعد، 1981: 331)،

ثانياً: ثبات المقياس:

يشير ثبات المقياس الى دقة وأتساق درجاته في قياس ما يجب قياسه، وإعطاء نفس النتائج او نتائج متقاربة لو كررت عملية القياس على الأفراد أنفسهم (الأنصاري، 2000: 114)، فمتى ما كانت درجات أداة القياس خالية من الأخطاء العشوائية، وكانت قادرة على قياس المقدار الحقيقي للسمة او الخاصية المراد قياسها قياساً متسقاً وفي ظروف مختلفة ومتباينة و كان المقياس عندئذ ثابتاً (علام و 2000: 167).

وبما ان الثبات يؤشر تجانس مفردات الاختبار التي يفترض ان تقيس ما يجب قياسه، لذا يمكن ان يكون تجانسا خارجياً External Consistency والذي يتحقق حينما يستمر الاختبار بإعطاء نتائج ثابتة بتكرار تطبيقه عبر الزمن، أو تجانسا داخلياً Internal Consistency والذي يتحقق من خلال كون فقرات الاختبار جميعها تقيس المفهوم نفسه (Franse Ila، 1981: 47).

ويعبر معامل الثبات عن نسبة التباين الحقيقي الى التباين الكلي في درجة المجيبين، وكلما زاد التباين الحقيقي ارتفع ثبات المقياس وقل الخطأ المعياري (Lemke & Weirsma، 1976: 71).

وبما ان معامل الثبات يحسب من درجات المقياس التي تتأثر بالموقف الذي يحدث فيه القياس، لذا يكون الثبات موقفياً قد يختلف من موقف الى آخر و تبعاً لتأثر نتائج القياس بحالة الفرد النفسية والجسمية، بالإضافة الى العوامل والعوارض الطارئة التي لا يمكن ضبطها تجريبياً في المرات التي تجري فيها عملية القياس (عودة، 1998: 386-389) (ربيع، 2009: 75).

فدرجة أي فرد في أي اختبار تنقسم الى جزئين، جزء جوهري ثابت لا يتأثر بالعوامل الخارجية، ويسمى بالدرجة الحقيقية، وجزء يتأثر بالعوامل الخارجية ويختلف تبعاً لاختلاف هذه العوامل، ويسمى بالدرجة الخاطئة، ومعامل ارتباط تلك الأجزاء الخاطئة يساوي صفراً، ومن ثم عندما يحسب معامل ارتباط الاختبار بنفسه، فأننا نحصل على قيمة عددية تدل على الجزء الثابت من هذا الاختبار، أي ان الثبات يقيس الجزء الحقيقي من التباين العام للاختبار، وهذا الجزء الحقيقي هو الذي يعطينا القيمة العددية لارتباط الاختبار بنفسه (السيد، 1971: 415)، فعندما يطبق الاختبار على مجموعة من الأفراد ثم نعيد تطبيق الاختبار نفسه او صورة مكافئة له بعد مدة مناسبة من الزمن على المجموعة نفسها، فإن معامل الارتباط بين التطبيقين هو ما يعرف بمعامل الثبات (عبد السلام، 1987: 163).

ويمكن التحقق من ثبات المقاييس والاختبارات النفسية بعدة طرائق، ويتوقف اختيار الطريقة الملائمة على نظرية القياس النفسي التي أعتمدها الباحث في بناء اداة القياس، وعلى خصائص السمة المقیسة، وأهمية القرار الذي ينشأ من استخدام المقياس، وهذه الطرائق هي:

1- طريقة إعادة الاختبار: Test Re – Test

ويطلق على معامل الثبات المحسوب بهذه الطريقة بمعامل الاستقرار او التجانس الخارجي، وتقوم الفكرة الأساسية في هذه الطريقة على إجراء الاختبار على مجموعة من الأفراد ثم إعادة تطبيقه مرة ثانية على نفس المجموعة في ظروف مماثلة وبعد مضي فترة زمنية مناسبة و استخراج معامل الارتباط بين درجات الإجراءين (ربيع، 2009: 85).

2- طريقة الصور المتكافئة: Equivalent Forms

تقوم الفكرة الأساسية لهذه الطريقة على أعداد صورة مكافئة تماما للمقياس الذي يحسب ثباته، من حيث خصائص الفقرات وطبيعتها وسهولتها وصعوبتها، ويطبق المقياس وصورته المكافئة على الأفراد أنفسهم ويحسب معامل الارتباط بين درجتهما (Eble، 1972: 412)، ويطلق على معامل الثبات بهذه الطريقة بمعامل التكافؤ في حالة تطبيق الصورتين المتكافئتين في الوقت نفسه، وبمعامل الاستقرار والتكافؤ معا إذا تم تطبيق الصورتين المتكافئتين بفواصل زمني (أبو حطب وآخرون، 1987:122).

3- طريقة التجزئة النصفية: Split – Half

يحسب معامل الثبات بهذه الطريقة من خلال تقسيم فقرات المقياس الى جزئين بعد تطبيقه على عينة الثبات، ومن أكثر الطرق شيوعا تقسيمها الى الفقرات الفردية مقابل الفقرات الزوجية، ثم حساب معامل الارتباط بين درجتهما، وتصحيح معامل الارتباط باستخدام معادلة (سبيرمان – براون) للحصول على معامل ثبات الاختبار الكلي المحسوب (ملحم، 2005: 263).

4- طريقة تحليل التباين: Varianc Analysis

تعد طريقة تحليل التباين من الطرق الشائعة في حساب ثبات الاختبارات التربوية والنفسية، إذ يشير معامل الثبات المحسوب بهذه الطريقة الى الاتساق داخل فقرات الاختبار، او الى التجانس الداخلي لفقراته (Weiner & Stewart، 1984: 61).

المحور الثالث**إجراءات البحث**

يتضمن هذا المحور استعراضاً لأهم المشكلات والتساؤلات التي تواجه الباحثين عند أعدادهم لإجراءات البحث، ومحاولة اقتراح الحلول العلمية لها وكما يأتي:

- التقيد بمستوى واحد من مستويات الدلالة الإحصائية:

يتساءل كثير من الباحثين (هل ان من الضروري التقيد والالتزام بمستوى واحد من مستويات الدلالة الإحصائية في كل إجراءات البحث ونتائجه؟ وهل يتوجب أن يثبت مستوى محدد مسبقاً في فرضيات البحث؟ أم ان للباحث الحق في التنقل بين هذه المستويات ضمن البحث الواحد؟).

للإجابة عن هذه التساؤلات لابد من التذكير بأن مستويات الدلالة الإحصائية تمثل أساساً المستوى المسموح به من هامش الخطأ لقبول النتائج في مجال البحوث الإنسانية عموماً، ومنها البحوث التربوية والنفسية، وعليه فإن من حق الباحث اختبار مدى الدلالة الإحصائية للنتائج التي توصل اليها عند مستوى أعلى إذا ما تبين له بأنها كانت دالة عند الحد الأدنى وهو مستوى (05،0) لقبول هذه النتائج.

ومن الأخطاء الشائعة في هذا المجال هو التحديد المسبق لمستوى الدلالة الإحصائية المعتمد في البحث من خلال تثبيت هذا المستوى (غالباً ما يكون مستوى 05،0) في فرضيات البحث ويبقى الباحث نفسه ملزماً بهذا المستوى في إجراءات البحث ونتائجه، في الوقت الذي يمكن فيه ان تكون هذه النتائج دالة إحصائياً عند مستويات دلالة أعلى وهو ما يؤشر دقة وموثوقية أفضل بالنتائج.

- التحليل المنطقي لفقرات المقياس أو الأداة (الصدق الظاهري) في مقابل التحليل الإحصائي لها:

تشير معظم أدبيات القياس والتقويم الى أن التحليل المنطقي لفقرات المقياس أو الأداة يعد إجراءً ضرورياً كخطوة أولى في مرحلة أعداد المقياس أو أداة البحث لأنه يؤثر مدى تمثيل الفقرة ظاهرياً للسمة أو الخاصية المقاسة، الا أنه في ذات الوقت يعد إجراءً غير مكتمل الدقة لأنه قد يكون مضللاً لاعتماده على الآراء الذاتية للمحكمين في تقديرهم لمدى ارتباط الفقرة كما تبدو ظاهرياً في صياغتها بالسمة أو الخاصية التي أعدت لقياسها. (الكبيسي، 2001: 171)

وعلى الرغم من أن معظم الأدبيات في مجال بناء الاختبارات تشير الى أن هذا الإجراء لا يمتاز بالدقة والموضوعية الكافية للحكم على صلاحية الفقرات في قياس ما أعدت لأجل قياسه، الا أن معظم الباحثين يناقضون هذا التنظير ويتقاطعون معه عند استخدامهم للصدق الظاهري إجرائياً، فالباحث يلجأ الى حذف الفقرات التي لا تحصل على الموافقة الظاهرية من قبل المحكمين بعد عرضها عليهم من قائمة الفقرات أو من أداة البحث بصورة نهائية.

والمشكلة هنا أن هذه الفقرات المحذوفة يمكن ان تؤثر في دقة الأداة أو المقياس في تمثيلها للصفة أو الخاصية المقاسة، خصوصاً عندما يكون عدد الفقرات المحذوفة من الأداة أو المقياس كبيراً.

والأفضل هنا وتماشياً مع الاستخدام الأمثل لمؤشر الصدق الظاهري أن يلجأ الباحث الى تحديد الفقرات التي لم تحصل على نسبة اتفاق عالية من قبل المحكمين، ولكن لا يتم حذف هذه الفقرات بل يتم إخضاعها للتحليل الإحصائي وعند ذلك يتخذ القرار بحذفها عندما تتطابق نتائج التحليل الإحصائي مع نتائج التحليل المنطقي في تحديد عدم صلاحية الفقرة في قياس ما أعدت لأجل قياسه.

وعلى الباحث هنا ترجيح كفة التحليل الإحصائي للفقرات على التحليل المنطقي لها، حيث يشير (أبيل) Ebel الى أن التحليل المنطقي للفقرات قد لا يكشف أحياناً عن صلاحيتها أو صدقها بشكل دقيق، في حين أن التحليل الإحصائي للدرجات التجريبية يكشف عن دقة الفقرات في قياس ما وضعت لأجل قياسه (Ebel, 1972:406).

وعليه فإن من الأفضل أن يحدد الباحث محكاً مسبقاً يستند إليه في اتخاذ القرار بحذف الفقرة من الاختبار أو الإبقاء عليها، وهذا المحك يحدد في ضوء عدد المؤشرات المستخدمة في البحث، كأن يكون توافر اثنين أو أكثر من المؤشرات التالية (الصدق الظاهري للفقرة، معامل تمييزها، معامل ارتباط درجة الفقرة بالدرجة الكلية للمقياس).

وفي طور الحديث عن مجال الصدق الظاهري لابد من الإشارة الى ضرورة استبدال مفردة (الخبراء) بمفردة (المحكمين) وهم التدريسيين أو المتخصصين الذين عرضت عليهم قائمة الفقرات لمراجعتها ظاهرياً، فالخبير هو الذي يكون متخصصاً تماماً في ذات الموضوع الذي يتناوله الباحث، وهذا أمر نادر الحدوث، في حين أن المحكم يكون متخصصاً في المجال العام الذي يصنف ضمنه البحث.

- أنواع الصدق ومؤشراته:

لقد حددت جمعية السيكولوجيين الأمريكية American Psychological Association ثلاثة أنواع رئيسية من الصدق هي: (صدق المحتوى، الصدق المرتبط بمحك، صدق البناء) (A.P.A, 1985: 9)، في حين أن هنالك بعض الإجراءات التي تعد بمثابة مؤشرات تدل على تحقق أحد هذه الأنواع الرئيسية من الصدق، ويمكن أيجاز هذه المؤشرات بالآتي:

أ- الصدق الظاهري: يعد مؤشراً على صدق المحتوى.

ب- الخارطة الاختبارية (جدول المواصفات): تعد مؤشراً جيداً ودقيقاً لصدق المحتوى عند بناء الاختبارات التحصيلية.

ت- القوة التمييزية للفقرات: تعد مؤشراً لصدق البناء.

ث- معامل ارتباط درجة الفقرة بالدرجة الكلية للمقياس أو الأداة: تعد مؤشراً لصدق البناء.

ج- مصفوفة الارتباطات الداخلية: تعد مؤشراً على صدق البناء.

ومن الأخطاء الشائعة في هذا المجال أن الباحث يتحقق أحياناً من مؤشر من مؤشرات هذه المؤشرات ألا أنه لا ينسبها الى نوع الصدق الذي ينتمي إليه هذا المؤشر.

عينات البحث:

عموما هنالك ثلاث أنواع من العينات التي تستخدم في معظم البحوث التربوية والنفسية، وتتباين هذه العينات في حجمها، والهدف من انتقائها، والفائدة المتأتية منها، ويمكن ترتيب هذه العينات كالآتي:

أ - العينة الاستطلاعية: وتسمى أحياناً (بعينة وضوح التعليمات والفقرات)، وهي عينة صغيرة من الأفراد تنتقى من مجتمع البحث بالطريقة العشوائية البسيطة، ويتم تطبيق أداة البحث عليها للتحقق من وضوح تعليمات الأداة او المقياس وفقراته بالنسبة للمستجيبين، والكشف عن جوانب الغموض وسوء الفهم الذي قد يحصل لدى المستجيبين، كما يتم من خلال هذه العينة التعرف على متوسط الوقت التقريبي المستغرق للإجابة.

ب - عينة التحليل الإحصائي: وهي عينة من الافراد تنتقى بشكل دقيق من مجتمع البحث، والغاية منها هو حساب الخصائص السيكومترية للمقياس وفقراته، وهناك اتجاهين في تحديد حجم هذه العينة يمكن أيجازهما بالآتي:

- الرأي الأول يشير الى أن حجم عينة التحليل الإحصائي ينبغي أن لا يقل عن (400) فرد ينتقون بدقة من مجتمع البحث، ومن أبرز أصحاب هذا الاتجاه (أنستازي) Anastasi و (ثورندايك) Thorndike (1971: 132).
- الرأي الثاني يجيز على وفق ما أفتخره (نانلي) Nunnally أن يكون حجم عينة التحليل الإحصائي بما لا يقل عن خمسة أفراد (كحد أدنى) مقابل كل فقرة اختبارية (Nunnally، 200: 1978)، وعليه فأن الحد الأدنى المسموح به لحجم العينة على وفق هذا الاتجاه يحسب كالآتي: (عدد أفراد عينة التحليل الإحصائي = عدد فقرات المقياس × 5).
- عينة البحث الأساسية: وتسمى أحياناً ب (عينة استخلاص النتائج)، وهي العينة التي يتم انتقائها من مجتمع البحث بدقة ليطبق عليها المقياس أو الأداة المعتمدة في البحث تحقيقاً للأهداف المعتمدة فيه. وتجدر الإشارة هنا الى أن بالإمكان اعتماد عينة التحليل الإحصائي بعدها عينة البحث الأساسية بشرط أن عملية التحليل الإحصائي لفقرات المقياس أو الأداة لم تتمخض عن استبعاد أي من هذه الفقرات.

التمثيل المتناسب للعينة:

- حينما تكون عينة البحث هي من نوع العينة متعددة المراحل، او من نوع العينة الطبقيّة متعددة المراحل، فأن من الأفضل أن تنتقى هذه العينة من المرحلة الأخيرة أو من الطبقات بصورة متناسبة وكالآتي:
- يتم أولاً تحديد حجم العينة التي يرغب الباحث بانتقائها.
 - تحسب نسبة هذه العينة بالنسبة للمجتمع الذي احتيرت منه وفق المعادلة التالية:
(حجم العينة ÷ حجم المجتمع × 100).
 - تنتقى من كل طبقة من طبقات المجتمع (جنس، صف، تخصص.... الخ) عدداً من الأفراد يقابل النسبة التي حسبت في الخطوة السابقة، ولتحديد ذلك تستخدم المعادلة التالية:
(حجم الطبقة ÷ 100 × النسبة المئوية المحسوبة).

متوسط الوقت المستغرق في الإجابة:

هذا الأجراء يحسب من خلال التطبيق الاستطلاعي للأداة على العينة الاستطلاعية، ومن الأخطاء الشائعة في حساب متوسط الوقت هو ما يقوم به بعض الباحثين من خلال الأجراء التالي: (الزمن الذي أستغرقه أول فرد أنهى الإجابة

+ الزمن الذي أستغرقه آخر فرد أنهى الإجابة ÷ 2)، واعتبار ناتج هذه المعادلة هو متوسط وقت الإجابة، وهذا الأجراء يتناقض مع مفهوم الوسط الحسابي.

لذا فإن متوسط الوقت التقريبي للإجابة يحسب من خلال إيجاد مجموع التوقيتات المستغرقة للإجابة ولجميع أفراد العينة الاستطلاعية مقسوماً على عددهم.

الدلالة الإحصائية لمعامل الارتباط:

من الإجراءات الشائعة الاستخدام التي يتبعها الباحثون للتحقق من التجانس الداخلي لفقرات المقياس أو الأداة والذي يعد مؤشراً على صدق البناء هو حساب قيمة معامل الارتباط بين درجة كل فقرة من فقرات المقياس وبين الدرجة الكلية له، حيث تشير (أنستازي) Anastasi الى أن ارتباط درجة الفقرة بمحك داخلي او خارجي يعد مؤشراً على صدقها، وحيثما لا يتوافر محك خارجي مناسب فإن الدرجة الكلية للمجيب تمثل أفضل محك داخلي في حساب هذه العلاقة (Anastasi، 1988: 211).

وللتحقق من هذا المؤشر يستخدم اثنين من معاملات الارتباط الشائعة الاستخدام وكالاتي:

- معامل ارتباط بيرسون: عندما تكون الإجابة عن فقرات المقياس متدرجة (دائماً، أحياناً، أبداً.... الخ)، وتعطى عند التصحيح تقديرات متدرجة (1، 2، 3، ... 1).
- معامل ارتباط بوينت بايسيرال: عندما تكون الإجابة عن الفقرات ثنائية منقطعة، وتعطى عند التصحيح الدرجات (0، 1).

وبلجاً كثير من الباحثين الى الحكم على قيمة معامل الارتباط المحسوب من خلال مقارنته بجدول القيم الحرجة لمعاملات الارتباط وهذه القيم الجدولية خاصة بمعاملات ارتباط بيرسون دون معاملات ارتباط بوينت بايسيرال. وسعيًا لمستوى أعلى من الدقة في الحكم على قيمة معامل الارتباط المحسوب فإن من الأفضل استخدام الاختبار التائي لدلالة معامل الارتباط باستخدام القانون الآتي:

وتعتمد الدلالة الإحصائية للقيمة التائية المحسوبة من خلال مقارنتها بالقيم التائية الجدولية بدرجة حرية (2 - n) معياراً للحكم على صدق الفقرة أم لا.

أسلوب صياغة الفقرات:

حينما يتطلب تحقيق أهداف البحث أعداد مقياس لقياس بعض السمات والخصائص النفسية أو الانفعالية فإن معظم الباحثين يشيرون في طور الحديث عن أعداد الفقرات الى أسلوب الصياغة المتبع في بنائها.

ومن الأخطاء الشائعة في هذا المجال هو الخلط الحاصل بين أسلوبين من أساليب بناء الفقرات الاختبارية، فبعض الباحثين يشيرون الى أن الأسلوب المتبع في صياغة الفقرات هو أسلوب (ليكرت)، على الرغم من ان أسلوب صياغة الفقرات هنا هو أسلوب (العبارات التقريرية) وهو واحد من الأساليب الشائعة الاستخدام في صياغة فقرات مقاييس الشخصية، وفي هذا الأسلوب تكون بدائل الإجابة التي تتبع كل فقرة متدرجة في شدة قياسها للسمة او الخاصية، فقد تكون هذه البدائل ثلاثية (دائماً، أحياناً، نادراً) أو رباعية (دائماً، أحياناً، نادراً، أبداً)، أو خماسية... الخ.

أما أسلوب (ليكرت) فهو من أكثر الأساليب شيوعاً في صياغة فقرات مقاييس الاتجاهات (حصراً)، وفي هذا الأسلوب يكون عدد بدائل الإجابة هو عدداً فردياً، بمعنى أن هذه البدائل أما ان تكون ثلاثية أو خماسية أو سباعية، وفلسفة هذا التدرج للبدائل يمكن توضيحه بالمخطط الآتي:

(الموافقة ----- الحياد ----- الرفض)

وعلى هذا الأساس فإن استجابة الفرد على كل فقرة في أي مقياس للاتجاه نحو موضوع ما يمكن أن تتدرج من الموافقة التامة الى الرفض التام ولا بد من وجود درجة الحياد التي تمثل منتصف المسافة بين القبول والرفض.

تحديد درجة القطع عند أعداد البرامج التدريبية أو الإرشادية:

عندما يكون هدف البحث هو بناء برنامج تدريبي أو إرشادي فإن الباحث يلجأ إلى أنتقاء عينة من الأفراد يطبق عليهم البرنامج التدريبي أو التعليمي أو الإرشادي للتعرف على مدى فاعليته. وانتقاء العينة الملائمة في مثل هذه البحوث البحوث يعد أمراً مهماً جداً، فمعظم الباحثين في هذا المجال يلجأون إلى تطبيق مقياس أو اختبار تشخيصي يتم من خلاله تحديد عينة الأفراد الذين سيطبق عليهم البرنامج من خلال مقارنة متوسط درجاتهم مع المتوسط النظري للمقياس بعده درجة القطع التي تفصل بين من يمتلكون السمة أو لا يمتلكونها، واختيار الفئة الأعلى أو الأدنى من المتوسط النظري (بحسب هدف البرنامج). والانتقاد الموجه لهذا الإجراء هو أن الأفراد الذين يتم اختيارهم بهذه الطريقة من الممكن أن يكون متوسط درجاتهم قريباً من المتوسط النظري للمقياس عند الاختبار القبلي، كما ويمكن أن يرتفع هذا المتوسط عن المتوسط النظري بدرجة بسيطة في الاختبار البعدي عند انتهاء تطبيق البرنامج، ولاكن لا يمكن الجزم بأن جلسات البرنامج كانت هي صاحبة التأثير، فقد تلعب عوامل النضج والتعلم والخبرة البسيطة المستمدة من بعض جلسات البرنامج في أحداث هذه الزيادة البسيطة.

والإجراء الأفضل هنا هو أن يطبق المقياس على عينة من الأفراد ثم يحسب الانحراف المعياري للدرجات، ثم يتم أنتقاء العينة التي ستشمل جلسات البرنامج من خلال الركون إلى درجة القطع التالية:
(درجة القطع = المتوسط النظري للمقياس + الانحراف المعياري المحسوب للعينة)
ولزيادة الدقة يمكن للباحث أن يضاعف قيمة الانحراف المعياري المضاف للمتوسط النظري لتكون درجة القطع كالآتي:

$$\text{درجة القطع} = \text{المتوسط النظري} + (\text{الانحراف المعياري للعينة} \times 2)$$

معادلة (ألفا - كرونباخ) لحساب ثبات المقياس أو الاختبار:

تعد طريقة تحليل التباين واحدة من الطرائق المستخدمة في التحقق من ثبات المقاييس والاختبارات النفسية والتربوية و هنالك أربع معادلات فرعية تدرج ضمن هذه الطريقة وهي:

- معادلة كيودر - ريتشاردسون 20.
- معادلة كيودر - ريتشاردسون 21.
- معادلة ألفا - كرونباخ.
- معادلة هويت.

وتختلف هذه المعادلات الأربعة من حيث الشروط الخاصة باستخدام كلاً منها، وتعد معادلة (ألفا - كرونباخ) أسهل هذه المعادلات استخداماً وأكثرها شيوعاً، لأنها لا تحتاج إلى متطلبات أو شروط للاستخدام كما هو الحال في المعادلات الثلاث الأخرى، فهي تصلح لحساب ثبات جميع أنواع الاختبارات والمقاييس سواءً أكانت ذات إجابة متقطعة أو متدرجة، فبالإمكان استخدام هذه المعادلة لحساب ثبات الاختبارات التحصيلية ومقاييس الشخصية ومقاييس الاتجاهات... الخ.

على الرغم من ذلك فإن بعض الباحثين يلجؤون إلى استخدام معادلتَي (كيودر - ريتشاردسون 20 و 21) ويغفلون المتطلبات الصعبة أحياناً الواجب توافرها لاستخدام هاتين المعادلتين، كما أن استخدام معادلة (هويت) يتطلب أعداداً جداولاً لتحليل التباين.

من ناحية أخرى فإن من الأخطاء الشائعة هو ما تشير إليه بعض البحوث من أنه قد تم استخدام طريقة ألفا - كرونباخ في حساب الثبات، والأفضل هو الإشارة إلى اعتماد طريقة تحليل التباين باستخدام معادلة ألفا - كرونباخ.

المصادر

- ابو حطب، فؤاد وآخرون (1987)، **التقويم النفسي**، ط3، مكتبة الانجاء، القاهرة.
- اسعد، ميخائيل ابراهيم (1981)، **القياس النفسي**، مطبعة الجمهورية، دمشق.
- ثورندايك، روبرت واليزابيث، هيجن (1989)، **القياس والتقويم في علم النفس والتربية**، ترجمة عبدالله زيد الكيلاني وعبد الرحمن عدس، مركز الكتب الاردني، عمان.
- الجباري، محمد محي الدين صادق (2000)، دراسة مقارنة في الخصائص السيكمترية بين طريقتي ثرستون وليكرت في بناء مقاييس الاتجاهات، اطروحة دكتوراه (غير منشورة)، كلية التربية - جامعة بغداد
- ربيع، محمد شحاتة (2009)، **قياس الشخصية**، دار المسيرة للنشر والتوزيع، عمان.
- (1994)، **قياس الشخصية**، دار المعرفة، القاهرة.
- سمارة، عزيز وآخرون، (1989)، **مبادئ القياس والتقويم في التربية**، دار الفكر للنشر والتوزيع، الاسكندرية.
- السيد، فؤاد البهي، (1976)، **الذكاء**، دار الفكر العربي، القاهرة.
- عبد الخالق، احمد محمد (1993)، **أستخبارات الشخصية**، ط2، الاسكندرية، دار المعرفة الجامعية.
- عبد الرحمن، سعد (1983)، **القياس النفسي**، ط1، مكتبة الفلاح، الكويت.
- عبد السلام، نادية محمد (1987)، دراسة في حساب ثبات وصدق الاختبار هدفي المرجح، **مجلة علم النفس**، العدد 2، ص: 155 - 165.
- علام، صلاح الدين محمود (2000)، **القياس والتقويم التربوي والنفسي**، أساسياته، تطبيقاته، توجهاته المعاصرة، دار الفكر العربي، القاهرة.
- (2007)، **القياس والتقويم التربوي في العملية التدريسية**، دار المسيرة للنشر والتوزيع، عمان.
- عودة، احمد سليمان (1998)، **القياس والتقويم في العملية التدريسية**، دار الامل للنشر والتوزيع، الاردن.
- الكبيسي، كامل ثامر (2001)، العلاقة بين التحليل المنطقي والتحليل الإحصائي لفقرات المقاييس النفسية، مجلة الأستاذ، العدد (25)، جامعة بغداد، كلية التربية - ابن رشد.
- ملحم، سامي محمد (2005)، **القياس والتقويم في التربية وعلم النفس**، دار المسيرة للنشر والتوزيع، عمان.
- Allen, M.S. & yen, W.M (2002) , **Introduction to measurement - , Theory**, long Grove, Il Waveland press.
- Alderson, J. C,(1981), **Issues in language testing**, Documents 3, the British council. - American psychological Association (1985),**Standards for Educational and psychological tests**, Washington D.C: Author
- Anastasi,A. (1976), **Psychological testing**, New York, Macmillan publishing. -Anastasi, A. (1988), **Psychological Testing**, 6th ed, New York, MacMillan.
- Dick, W.& Hagert, Y. (1971), **Topics in Measrement**, New York, McGraw -Hill.
- Ebel,R.L.(1972), **Essentials of Educational Mesurement**,New Jersey prentice Hall Inc.
- Ghiselli,E.et al (1981) **Measurement theory for the behavioral sciences**,san francies& company.
- Graham,J.R & Lilly, R.S(1984), **Psychological Testing**, Englewood Cliffs, Prentice, Hall.
- Gronlund, NE(1981), **measurement and Evauation in teaching**, NewYork,Mac-Millan.

- Krus, D.J & Fuller,E.A.(1982), Illustration Of The Theory Of True And Error Scores,Computer Based Method Of Idealized Data Sets, **Education And Psychological**, vol: 42.
- Hambleton,R.K.&Zaal J.N.(1991),**Advances in Educational testing**, Boston,Kluwer Academic publishers.
- Lemke,E & Wiersms, W (1976), **Principles Of Psychological Measurement**, Chicago.
- Lord, F.M.(1960), **Psychological scaling**, New York, John Wiley.
- Neill, m.A. & Jackson, D.N. (1970), "An Evaluation of Item selection strategies in personality scale construction" **Educational and Psychological measurement**, Vol (30) no.3 ,p.p.647-661.
- Nunnally, J. C. (1978), **Psychometric Theory**, New York: McGraw Hill Company.
- Popham, W.J. (1975), **Educational Evaluation**,New Jersey:prentice Hall Inc.
- Paul,J.(1991), **Interpretation of test Reliability**, journal of Educational and Psychological measurement, Vol (51), P.P. 625-635.
- Suen, H.K.(1990), **Principles of test theories**, new jersey, Lawrence Erlbaum Associates, Inc,Hill saddle.
- Thorndike, R.L (1971), **Educational measurement**, 2nd ed,Washington American council of education.