الجمهورية الجزائرية الديمقراطية الشعبية
وزارة التعليم العالي والبحث العلمي
جامعة عبد الحميد بن باديس
كلية العلوم الإجتماعية
شعبة علم النفس

أطروحة مقدمة لنيل شهادة دكتوراه الطور الثالث في علم النفس
تخصص: القياس النفسي وتحليل المعطيات
تحت عنوان:

دراسة تقييمية لطرق تقدير صدق وثبات أدوات جمع المعطيات في أطروحات دكتوراه علم النفس وعلوم التربية بجامعة الجزائر2

من إعداد الطالب:
بن صافي عبد الرحمن.

تحت إشراف الأستاذ:
أ.د طاجين علي.

أعضاء لجنة المناقشة

هني حاج أحمد	أستاذ محاضر –أ–	جامعة مستغانم	رئيسا
طاجين علي	أستاذ التعليم العالي	جامعة مستغانم	مشرفا ومقررا
قماري محمد	أستاذ التعليم العالي	جامعة مستغانم	مناقشا
تيغزة أحمد	أستاذ التعليم العالي	جامعة وهران	مناقشا
بلقوميدي عباس	أستاذ محاضر –أ–	جامعة وهران	مناقشا

السنة الجامعية: 2017/2016

بسم الله الرحمن الرحيم

« وَإِنْ تَعُدُّوا نِعْمَةَ اللَّهِ لَا تُحْصُوهَا »

صدق الله العظيم

(سورة النحل، آية 18)

شكر وتقدير

أولا نشكر الله عزّ وجلّ القدير الذي أعاننا في إنجاز هذا العمل العلمي، ونتقدم بالشكر الجزيل إلى كل من ساعدنا من قريب أو بعيد ولو بابتسامة أو بجملة بسيطة؛ كما نتقدم بعميق شكرنا وخالص تقديرنا إلى أساتذتنا، وشكر خاص للبروفيسور "طاجين علي" الذي رغم انشغالاته الكثيرة أبى أن يكون صاحب الفضل في متابعة هذا العمل، بدون أن ننسى الأستاذ المسؤول عن مشروع دكتوراه "القياس النفسي وتحليل المعطيات" ومدير مخبر "تحليل المعطيات الكمية والكيفية للسلوكات النفسية والإجتماعية" البروفيسور "قماري محمد"، وكذا زملائي وزميلاتي في الدفعة.

إهداء

إلى من أطعماني ولم يأكلا حتى لا أذوق مرارة الجوع؛ إلى من لا أقدر على رد جميلها لو عملت لها طول الحياة، إلى من لا يكفي جهد الدهر لرد جميله ولو ليوم، إلى من لا يكفي قدر البحر مدادا للتعبير بالأحرف، ويعجز اللسان عن شكرهما بالكلمات، إلى أمي وروح أبي الطاهرة رحمه الله واسكنه فسيح جنانه، أهدي ثمرة جهدي وعملي، وألتمس منهما الرضا والعفو، إلى كل أفراد العائلة والأقارب، وإلى أصدقائي وزملائي في الجامعة والعمل، وإلى كل طالب علم نافع.

ملخص البحث

تهدف الدراسة الحالية إلى التعرف على مدى جودة طرق تقدير صدق وثبات أدوات جمع المعطيات في أطروحات دكتوراه علم النفس وعلوم التربية بجامعة الجزائر2.

لتحقيق أهداف هذه الدراسة، قام الباحث ببناء بطاقة لتقييم طرق تقدير الصدق والثبات، وتمثلت عينة الدراسة في 346 أداة جمع معطيات مستخدمة في 112 أطروحة دكتوراه في علم النفس وعلوم التربية بجامعة الجزائر2.

وقد أسفرت الدراسة على النتائج التالية:

1- تم التحقق من صدق أدوات جمع المعطيات بنسبة 72.00%، أما الثبات فبنسبة 65.90%.

2- أكثر طرق التحقق من الصدق شيوعا هي طريقة صدق المحكمين بنسبة 43.00%، أما بالنسبة للثبات فكانت طريقة معامل ألفا كرونباخ بنسبة 39.57%.

3- استخدامات طرق التحقق من الصدق كانت جيدة بنسبة 93.38%، أما الثبات فبنسبة 83.09%.

4- من الأخطاء التي وقع فيها الباحثون عند استخدامهم لطرق التحقق من الصدق: عدم كفاية طرق التحقق بنسبة 66.67%، عدم التحقق من الصدق على عينة ممثلة لمجتمع الدراسة بنسبة 53.61% وعدم التحقق من الصدق على عينة مناسبة الحجم بنسبة 07.83%؛ أما الأخطاء المتعلقة بالثبات فقد كانت: عدم التحقق من الثبات على عينة ممثلة لمجتمع الدراسة بنسبة 54.55%، عدم التحقق من الثبات على عينة مناسبة الحجم بنسبة 10.46%، كما لم يكن معامل الثبات مقبولا بنسبة 09.06%.

Résumé

La présente étude vise à déterminer la qualité des méthodes d'estimation de la validité et fidélité des instruments de collecte des données utilisés dans les thèses de doctorat en Psychologie et Sciences de l'Education à l'Université d'Alger 2.

Pour atteindre l'objectif de cette étude, le chercheur a élaboré une grille pour évaluer les méthodes d'estimation de la validité et la fidélité, l'échantillon est composé de 346 instruments de collecte des données utilisés dans 112 thèses de doctorat en Psychologie et Sciences de l'Education à l'Université d'Alger 2.

L'étude a révélé les résultats suivants:

1- Les instruments de collecte des données sont vérifiés au taux de 72,00% pour la validité et de 65,90% pour la fidélité.

2- La méthode d'estimation la plus fréquentée est la méthode d'inter-juges pour la validité (43.00%), et la méthode Alpha Cronbach pour la fidélité (39.57%).

3- Le taux de bonne utilisation des méthodes de l'estimation est 93.38% pour la validité et 83.09% pour la fidélité.

4- Les erreurs qui sont faites par les chercheurs quand ils vérifient la validité des instruments sont : l'insuffisance des méthodes utilisées (66.67%), l'échantillon utilisé n'est pas représentatif de la population (53.61%), la taille de l'échantillon est insuffisante (07.83%); et les erreurs concernant la fidélité sont : l'échantillon utilisé n'est pas représentatif de la population (54.55%), la taille de l'échantillon est insuffisante (07.83%) et les coefficients de corrélation sont inacceptables (09.06%).

Abstract

This study aims to determine the quality of the estimation methods of the validity and reliability of data collection instruments used in PhD theses in Psychology and Education Sciences at the University of Algiers 2.

To achieve the objective of this study, the researcher developed a grid to evaluate methods for estimating the validity and reliability, the sample consists of 346 data collection instruments used in 112 PhD theses in Psychology and Education Sciences at the University of Algiers 2.

The study revealed the following results:

1- The data collection instruments are verified at the rate of 72.00% for the validity and 65.90% for reliability.

2- The most common method to estimate the validity is the method of the judges at the rate of (43.00%), and the method of Cronbach Alpha for reliability (39.57%).

3- The rate of correct use of the methods for estimating the validity is 93.38% and 83.09% for the reliability.

4- The errors that are produced by researchers when they verified the validity of the instruments are: inadequate methods used for estimating (66.67%), the sample used is not representative of the population (53.61%) and the sample size is insufficient (07.83%); and the errors about reliability are: the sample used is not representative of the population (54.55%), the sample size is insufficient (07.83%) and the correlation coefficients are unacceptable (09.06%).

قائمة المحتويات

الموضـوع	الصفحة
شكر وتقدير	ب
إهداء	ت
ملخص البحث	ث
قائمة المحتويات	ح
قائمة الجداول	د
قائمة الملاحق	ذ
الفصل الأول: مدخل إلى البحث	
مقدمة	02
إشكالية البحث وتساؤلاته	04
أهداف البحث	06
أهمية البحث	06
حدود البحث	06
التعاريف الإجرائية لمصطلحات البحث	07
الفصل الثاني: الإطار النظري للبحث	
I: الصدق والثبات	**09-50**
أولا: الصدق	10
صدق المحتوى	13
الصدق المرتبط بمحك	16
صدق التكوين الفرضي	21
العوامل المؤثرة في الصدق	27
ثانيا: الثبات	28
طريقة إعادة الاختبار	30
طريقة الصور المتكافئة	33

معامل الاستقرار والتكافؤ	35
معامل الاتساق الداخلي	36
ثبات التقدير	43
طرق إحصائية أخرى	45
القيمة المقبولة لمعامل الثبات	46
العوامل المؤثرة في الثبات	47
العلاقة بين الصدق والثبات	50
II: الدراسات السابقة	**65-51**
الفصل الثالث: إجراءات البحث	
منهج البحث	67
عينة البحث	67
أداة البحث	72
الأساليب الإحصائية	77
الفصل الرابع: نتائج البحث	
عرض ومناقشة النتائج المتعلقة بالصدق	79
عرض ومناقشة النتائج المتعلقة بالثبات	96
إستنتاجات البحث	108
التوصيات والاقتراحات	112
المراجع	113
الملاحق	122

قائمة الجداول

الرقم	عنوان الجدول	الصفحة
01	عدد الأطروحات قبل الإستبعاد	68
02	توزيع الأطروحات حسب متغير التخصص	69
03	توزيع الأطروحات حسب متغير القسم	69
04	توزيع الأطروحات حسب سنة التخرج	70
05	توزيع الأدوات حسب متغير التخصص	70
06	توزيع الأدوات حسب متغير القسم	71
07	قائمة المحكمين لأداة الدارسة	73
08	تكرارات ونسب أدوات جمع المعطيات التي تم التحقق من صدقها	79
09	تكرارات ونسب طرق التحقق من الصدق	81
10	تكرارات ونسب الإستخدامات الجيدة لطرق التحقق من الصدق	83
11	تكرارات ونسب عدم كفاية طرق التحقق من الصدق	86
12	تكرارات ونسب عدم التحقق من الصدق على عينة ممثلة لمجتمع الدراسة	87
13	تكرارات ونسب عدم التحقق من الصدق على عينة مناسبة الحجم	88
14	نسب الأخطاء المتعلقة بالعوامل المؤثرة على الصدق	89
15	النتائج المتعلقة بجانب التحقق من صدق المحتوى	90
16	النتائج المتعلقة بجانب التحقق من الصدق المرتبط بمحك	91
17	النتائج المتعلقة بجانب التحقق من صدق التكوين الفرضي	93
18	تكرارات ونسب أدوات جمع المعطيات التي تم التحقق من ثباتها	96
19	تكرارات ونسب طرق التحقق من الثبات	98
20	تكرارات ونسب الإستخدامات الجيدة لطرق التحقق من الثبات	100
21	تكرارات ونسب عدم التحقق من الثبات على عينة ممثلة لمجتمع الدراسة	103
22	تكرارات ونسب عدم التحقق من الثبات على عينة مناسبة الحجم	104
23	تكرارات ونسب مستويات معامل الثبات	105
24	نسب الأخطاء المتعلقة بالعوامل المؤثرة على الثبات	107

قائمة الملاحق

الرقم	عنوان الملحق	الصفحة
01	أداة البحث	123
02	تعليمات ملء بيانات الشبكة ومعايير التقييم	127
03	نتائج تطبيق أداة البحث	140

الفصل الأول

مدخل إلى البحث

مقدمة:

شهدت حركة القياس النفسي تطورات كبيرة منذ بداية التجريب في علم النفس على يد فونت عام 1876 بألمانيا، فخلال الحرب العالمية الأولى انصب التركيز على بناء اختبارات ومقاييس نفسية اعتمدت على مبدأ الفروق الفردية، والتي كانت موجهة في البداية إلى مجال الإختيار والتوجيه المهني للجنود، لتتوسع بعد ذلك إلى مختلف مجالات الحياة الأخرى كالمؤسسات الصناعية والتربوية والعلاجية. ومن أهم مميزات هذا التطور في بناء أدوات القياس، الاهتمام بالخصائص السيكومترية التي من خلالها يمكن التحقق من مصداقية هذه الأدوات، حيث بدأ الاهتمام بتطوير الاختبارات والمقاييس من خلال إيجاد مؤشرات جديدة لها، ودراسة علاقة بعض المتغيرات أو تقديرها في الوصول إلى خصائص جيدة للاختبارات والمقاييس التي تحقق هدفها في قياس ما أعدت لقياسه.

إن الاهتمام بالخصائص السيكومترية جاء كرد على الإنتقادات الموجهة لعلم النفس عموما وللإختبارات النفسية خصوصا، وذلك لإضفاء قدر من الطابع العلمي لها، فمن خصوصيات القياس النفسي مقارنة بالقياس الطبيعي أنه أكثر تعرضا للخطأ (Ley, 1972: 102)، كما أنه غير مباشر أي أن الظواهر أو السمات توصف بتكوينها الفرضي وليس بوجودها المادي ويستدل عليها من خلال السلوك الدال عليها (عودة، 1998: 23-24)، ولهذا يصعب وضع قوانين تساعد على التنبؤ بالسلوك في مجال العلوم الإنسانية والإجتماعية -عكس العلوم الطبيعية- وبالتالي تعتمد على ما يسمى بالإحتمالات. كما أنه لا توجد وحدات قياس ثابتة متفق على استخدامها في القياس النفسي كما هو الحال في القياس الطبيعي (عيسوي، 1985: 166)، حتى في المجالات التي يصعب إجراء التجارب فيها نجد وحدات قياس متفق عليها، ففي الماكرو فيزياء تقاس المسافة في الفضاء الخارجي بالسنة الضوئية، وفي الميكرو فيزياء يقاس طول الالكترون في نواة الذرة بالنانومتر.

إن عملية بناء المقياس بمختلف مراحلها تعتبر خطوة منهجية مصيرية ومحددة للخطوات الموالية ولمسار البحث ككل، وعليه هناك خصائص ينبغي التحقق منها في أداة القياس، للاطمئنان على صلاحية النتائج المتحصل عليها وإمكانية اتخاذ قرارات هامة متعلقة بأفراد، جماعات أو منظمات، وهي ما يصطلح عليها بـ "الخصائص السيكومترية"، من بينها الصدق: أي أن المقياس يقيس فعلا ما وضع لقياسه، والثبات: أي غياب الأخطاء العشوائية، واتساق فقرات المقياس فيما بينها واتساق القياسات التي يتم الحصول عليها في مناسبات مختلفة وتحت نفس الظروف.

يعتبر الصدق والثبات من أهم الشروط التي ينبغى توفرها في أدوات القياس، فإذا لم يحدد الباحث في دراسته درجة صدق وثبات أدوات القياس التي يبنيها أو يستخدمها، فإن ذلك يفقد قيمة النتائج التي توصل لها في تلك الدراسة، حيث تشير أدبيات القياس النفسي إلى أساليب متعددة للتحقق من صدق المقياس، كمعامل الإتفاق بين المحكمين، معامل الإرتباط ببين المقياس والمحك، معامل الإرتباط بين درجة الفقرة ودرجة البعد أو الدرجة الكلية للمقياس، طريقة المقارنة الطرفية، المجموعات المتضادة وغيرها من الأساليب؛ أما بالنسبة للثبات فهناك عدة أساليب لتقدير قيم معامل الثبات مثل معامل الإرتباط بيرسون، معادلة سبيرمان براون ومعامل ألفا كرونباخ.

كما أن أي قصور في عملية بناء أداة جمع المعطيات أو في طرق التحقق من خصائصها السيكومترية، سيزيد من أخطاء القياس العشوائية، مما يعطي تفسيرا خاطئا وعدم مناسبة القرار للغرض المستخدمة فيه هذه الأداة، وبالتالي التقليل من قيمتها التعميمية والتطبيقية.

إشكالية البحث وتساؤلاته:

تعتبر الرسائل والأطروحات الجامعية جزءا هاما وأساسيا من مخرجات البحث العلمي، وذلك بحكم إعدادها وفق معايير المنهجية والقياس، مما يجعلها مؤهلة لئن تنتهي بتوصيات وآليات يمكن أن تسهم في إحداث التغيير أو معالجة المشكلات وبالتالي خدمة قضايا المجتمعات التنموية بأساليب علمية.

إلا أنه في بعض الأحيان لا يتم الأخذ بنتائج البحوث التربوية والنفسية في الجامعات ولايتم الاستفادة منها، وقد يرجع ذلك لعدم سلامة الإجراءات المنهجية، إما في مرحلة جمع المعطيات أو مرحلة التحليل والتفسير، حيث أن المرحلة الثانية تتعلق بنتائج المرحلة الأولى (جمع المعطيات) وتتحدد بها، وهذه الأخيرة تعتمد بدورها على أداة القياس المستخدمة، فكلما كانت هذه الأداة دقيقة وتتمتع بخصائص سيكومترية جيدة، كلما كانت نتائج الدراسة ذات مصداقية عالية ويمكن الاعتماد عليها في اتخاذ القرارات.

من خلال الإطلاع على البحوث في المجال النفسي والتربوي نلاحظ أن العديد من الباحثين يقعون في أخطاء منهجية عند التحقق من الصدق والثبات، حيث نجد أن البعض منهم يطبق أداة الدراسة الأولية على عينة غير ممثلة، أو يقتصر على طريقة واحدة فقط، كصدق المحكمين وثبات التجزئة النصفية، والبعض الأخر لا يراعي افتراضات وشروط التطبيق للطريقة المستخدمة؛ كما أن العديد من الباحثين لا يقومون بإعادة التحقق من صلاحية أدوات سابقة – عربية كانت أم أجنبية – ويطبقونها على عينة الدراسة مباشرة، حيث يكتفون بالخصائص السيكومترية الواردة للمقياس في البيئات الأصلية.

من هذا المنطلق يرى الباحث ضرورة التقويم المستمر والدائم للبحوث النفسية والتربوية، خصوصا فيما يتعلق بالطرق والأساليب المستخدمة في التحقق من صدق وثبات أدوات القياس.

وعليه يمكن تحديد مشكلة الدراسة في التساؤلات التالية:

أولا: الصدق:

1- هل تم التحقق من صدق أدوات جمع المعطيات في أطروحات دكتوراه علم النفس وعلوم التربية بجامعة الجزائر2؟

2- ما هي الطريقة الأكثر شيوعا في التحقق من صدق أدوات جمع المعطيات؟

3- ما نسبة الاستخدامات الجيدة لطرق التحقق من صدق أدوات جمع المعطيات؟

4- ما هي الأخطاء التي وقع فيها الباحثون عند التحقق من صدق أدوات جمع المعطيات؟

أ- هل كانت الطرق المستخدمة في التحقق من الصدق كافية؟

ب- هل تم التحقق من صدق الأداة على عينة ممثلة لمجتمع الدراسة؟

ت- هل تم التحقق من صدق الأداة على عينة مناسبة الحجم؟

5- هل تم مراعاة شروط التطبيق عند التحقق من صدق أدوات جمع المعطيات؟

ثانيا: الثبات:

1- هل تم التحقق من ثبات أدوات جمع المعطيات في أطروحات دكتوراه علم النفس وعلوم التربية بجامعة الجزائر2؟

2- ما هي الطريقة الأكثر شيوعا في التحقق من ثبات أدوات جمع المعطيات؟

3- ما نسبة الاستخدامات الجيدة لطرق التحقق من ثبات أدوات جمع المعطيات؟

4- ما هي الأخطاء التي وقع فيها الباحثون عند التحقق من ثبات أدوات جمع المعطيات؟

أ- هل تم التحقق من ثبات الأداة على عينة ممثلة لمجتمع الدراسة؟

ب- هل تم التحقق من ثبات الأداة على عينة مناسبة الحجم؟

ت- هل كانت مستويات معاملات الثبات مقبولة؟

أهداف البحث:

يهدف البحث الحالي إلى مايلي:

1- التعرف على مدى قيام طلاب الدراسات العليا بالتحقق من صدق وثبات أدوات جمع المعطيات في أطروحات دكتوراه علم النفس وعلوم التربية.

2- التعرف على مدى جودة طرق التحقق من صدق وثبات أدوات جمع المعطيات في أطروحات دكتوراه علم النفس وعلوم التربية.

أهمية البحث:

تظهر أهمية البحث الحالي من خلال أهمية الموضوع الذي نعالجه، وهو صدق وثبات أدوات جمع المعطيات، فالأداة التي لا تتوفر فيها هذين الخاصيتين، لا يمكن الوثوق فيما تمدنا به من معطيات، ومن ثم بالقرارات الناتجة عن طريقها، وبالتالي فهي لن تساعد في التحقق من فرضيات البحث أو الإجابة على تساؤلاته، ولا تحقيق الغرض الذي صممت من أجله.

من هنا يرى الباحث أن أهمية هذا البحث تتمثل في تقديم معطيات من الواقع إلى مشرفي الدراسات العليا في الجامعات الجزائرية حول موضوع صدق وثبات أدوات جمع المعطيات، لتحديد برامج التكوين والمقررات في الموضوع.

حدود البحث:

تقتصر الدراسة الحالية على وصف وتقييم الطرق المختلفة المستخدمة في التحقق من صدق وثبات أدوات جمع المعطيات في أطروحات دكتوراه علم النفس وعلوم التربية بجامعة الجزائر2.

التعاريف الإجرائية لمصطلحات البحث:

أ- الصدق:

يعرف ميسك Messick صدق الأداة على أنه "تقييم شامل يوفر من خلاله الدليل المادي والمبرر المنطقي اللازم لإثبات كفاية وملائمة ومعنى أي تأويل أو فعل يبنى على درجة الأداة". (الدوسري، 2000: 53).

ونقصد بالأداة الصادقة في بحثنا: الأداة التي تم التحقق من صدقها باستعمال طريقة واحدة على الأقل وفقا لمعايير الجمعية الأمريكية لعلم النفس.

ب- الثبات:

يعرف الثبات بأنه خلو أداة القياس نسبيا من الأخطاء العشوائية، أي قدرة الأداة على قياس الدرجة الحقيقية بأقل قدر ممكن من الخطأ العشوائي؛ ويشير ثبات الأداة إلى اتساقها الداخلي أي اتساق درجاتها في قياس ما يجب قياسه (علام، 2000؛ عبد الرحمن 1998).

ونقصد بالأداة الثابتة في بحثنا: الأداة التي تم التحقق من ثباتها باستعمال طريقة واحدة على الأقل وفقا لمعايير الجمعية الأمريكية لعلم النفس.

ج-أداة جمع المعطيات:

يقصد بها أداة القياس المستخدمة لجمع المعطيات الكمية أو الكيفية حول ظاهرة معينة، "وهي مجموعة من الأسئلة أو المواقف التي تمثل القدرة أو الخاصية المراد قياسها" (عبد الرحمن، 1998: 160).

أما إجرائيا فهي أحد الأدوات المستخدمة في الأطروحات: الملاحظة، المقابلة، الإختبار.

الفصل الثاني

الإطار النظري للبحث

I) الصدق والثبات

تمهيد: سنتطرق في هذا القسم إلى الصدق والثبات، حيث أكد المختصون في القياس النفسي أنه من المتعذر إعداد أدوات لقياس الظواهر النفسية من غير وجود أخطاء في القياس، سواء في النظرية الكلاسيكية للقياس أو النظرية الحديثة، "لذلك اتجه المتخصصون في القياس النفسي إلى تحديد بعض الخصائص السيكومترية، التي يمكن أن تحد من هذه الأخطاء" (Dick & Hagerty, 1971: 13).

"من أهم هذه الخصائص السيكومترية التي طورها المختصون في المقاييس النفسية، الصدق والثبات، إذ تعتمد عليهما دقة المعلومات التي توفرها هذه المقاييس" (ربيع، 1994: 42،91)، فبدونهما لا يمكن الوثوق في قدرة الأداة على قياس ما صممت لقياسه ولا مصداقية النتائج المتحصل عليها عند استخدامها لقياس السمات المختلفة.

"وهذا لا يعني إهمال الخصائص والسمات الأخرى التي يجب أن تتمتع بها أداة القياس، فهي بالإضافة إلى ذلك يجب أن تكون:

- شاملة وممثلة لجميع مكونات القدرة أو الخاصية المطلوب قياسها.
- أن تبنى وتحلل بطريقة موضوعية، مما يعني عدم تدخل العوامل الذاتية في بنائها وتحليلها وتصحيحها، وبالتالي فإن درجات الفرد ستبقى كما هي حتى باختلاف المصحح.
- أن تكون مفردات الأداة متصلة بالموضوع المراد قياسه اتصالا جيدا، مما يؤدي إلى إيجاد مدى واسع من انتشار وانحراف الدرجات حول المتوسط، وعندها تصبح الأداة مناسبة للعينة من حيث درجة الصعوبة والسهولة. (عبد الرحمن، 1998: 159).

أولا: الصدق

يعد الصدق من أهم الخصائص السيكومترية التي يتطلب توافرها في المقياس النفسي "لأنه يمثل إحدى الوسائل المهمة في الحكم على صلاحية المقياس" (Ebel, 1972 : 435).

ويرى كرونباخ (Cronbach, 1960) أنه "بقدر ما يكتمل تفسير درجة المقياس للسمة المعنية والثقة في هذا التفسير، بقدر ما يكتمل صدق المقياس، وهو بذلك يربط بين الدرجة على المقياس وقدرتها التفسيرية".

يعرف ميسك Messick صدق الأداة على أنه "تقييم شامل يوفر من خلاله الدليل المادي والمبرر المنطقي اللازم لإثبات كفاية وملائمة ومعنى أي تأويل أو فعل يبنى على درجة الأداة". (الدوسري، 2000: 53).

ويرى ثورنديك وهيجن Thorndike & Hagen (1989: 54) بأن "الصدق يشير إلى الدرجة التي يمكن فيها مقياس ما أن يقدم معلومات ذات صلة بالقرار الذي سيبنى عليه".

أي أن المعنى العام لمفهوم الصدق أن نقوم بجمع معلومات ودلائل من درجات أدوات القياس تفيدنا في معرفة إلى أي حد تقيس الأداة ما وضعت لقياسه، وقد ذكر سعد عبد الرحمن (1998) أن المقصود بقدرة الأداة على قياس ما وضعت لقياسه، أن تكون أداة ممثلة تمثيلا حقيقيا للسمة التي وضعت لقياسها، وأن تقيسها دون أن يختلط معها قياس سمات أخرى غير مطلوب قياسها، وأن تكون الأداة قادرة على التمييز بين طرفي القدرة التي تقيسها، أي التمييز بين الأداء القوي والأداء المتوسط والأداء الضعيف.

وبما أن الصدق يحسب من استجابات الأفراد عند تطبيقه، لذا يكون موقفياً إذ يتأثر بالعوامل المؤثرة في عملية القياس، "فإذا كان المقياس صادقا في موقف معين فإنه قد لا يكون صادقا في موقف آخر، فضلا عن أنه محدد بطبيعة السلوك الذي يمثل الخاصية أو السمة المراد قياسها" (عبد الرحمن، 1998).

"الصدق بلغة الاحصاء هو نسبة التباين الحقيقي المرتبط أو المنسوب للسمة المقاسة إلى التباين الكلي" (عودة، 1998: 338-339).

وللصدق مجموعة من الخصائص نعددها كالتالي:

1- الصدق خاصية تتعلق بنتائج الأداة وتفسيرنا لها وليس بالأداة في حد ذاتها، فالأداة قد تكون لها عدة مؤشرات للصدق بحسب تعدد أغراض إستخدامها، ونوع القرارات التي يمكن اتخاذها.

2- الصدق صفة نسبية متدرجة فلا توجد أداة عديمة الصدق تماما أو تامة الصدق.

3- الصدق صفة نوعية، أي أن الأداة التي تصدق في قياس سمة معينة قد تصدق في قياس سمة أخرى.

4- يعبر عن الصدق بدرجة وصفية، وقد تستخدم الأرقام (معامل الاتفاق أو الارتباط) لتوضيح الدرجة الوصفية. (السيد، 1997)، (غانم، 1997)، (علام، 2000)، (ملحم، 2005)، (مراد وسليمان، 2002).

ونتيجة للاستخدام الضيق والمحدود لمفهوم الصدق من قبل بعض الباحثين، ولأن الاختبارات والمقاييس تستخدم مجملها لاتخاذ قرارات وإصدار أحكام متنوعة كلا منها يتطلب نوعا معينا من الدراسات للتحقق من صدقه، "فقد أصدر علماء القياس في لجنة مشتركة من الجمعية الأمريكية لعلم النفس APA والجمعية الأمريكية للبحث التربوي AERA والمجلس القومي الأمريكي للقياس التربوي NCME كتيبا يضم أهم معايير الإختبارات والمقاييس التربوية والنفسية وتصنيف للقرارات والأحكام المتنوعة التي يهدف إليها القياس بصفة عامة، وهي:

- تحديد كيفية أداء الفرد في الوقت الحاضر في نطاق شامل لمواقف سلوكية تمثلها أسئلة الأداة تمثيلا جيدا.

- التنبؤ بالأداء المستقبلي للفرد أو تقدير مكانه في أحد المتغيرات ذات الأهمية.
- الإستدلال على درجة تملك الفرد سمة أو خاصية معينة من حيث تكوين فرضي يبدو أثرها في سلوك الفرد وأدائه.

حيث حددت ثلاثة مؤشرات (جوانب، أوجه، أبعاد) رئيسة لصدق المقاييس النفسية هي: صدق المحتوى، صدق البناء والصدق المرتبط بمحك (علام، 2000)، وذلك منذ عام 1966 (التصنيف الثلاثي)، بعد أن كان التصنيف الرباعي عام 1954.

حيث أن واحدا أو أكثر من الطرق الثلاثة من طرق الصدق (صدق المحتوى، صدق المحك، صدق التكوين الفرضي) تستخدم لتقدير صدق أدوات القياس حسب الاستخدام المرجو منها، وقد حدد هذه الاستخدامات في المجالات التالية وهي:

1. تمثيل أداة القياس لمجال محدد من محتوى معين.
2. اكتشاف وتحديد علاقة وظيفية بين أداة القياس وبعض المتغيرات المناسبة.
3. قياس السمات النفسية (بخش، 2007).

وحسب النظرية الحديثة فإن هذه الجوانب الثلاثة من الصدق ليست مختلفة بل هي متكاملة، وتعتبر أدلة على مدى صلاحية المقياس في المساعدة على اتخاذ القرارات وإطلاق الأحكام (تيغزة، 2009).

"بما أن الصدق مفهوم واحد، وما أصطلح عليه بأنواع الصدق ما هي إلا مؤشرات وطرائق لجمع الأدلة عنه، فإنه كلما كان المقياس يحمل أكثر من مؤشر للصدق زادت الثقة به في قياس ما أعد لقياسه (Jenkins, 1978: 93)، "ويتوقف على عاملين مهمين هما: الغرض من القياس، والفئة أو الجماعة التي سيطبق عليها المقياس" (الظاهر وآخرون: 1999: 132-133)، فيما يأتي توضيح لهذه الطرق:

1- صدق المحتوى:

يعني صدق المحتوى "مدى تمثيل عناصر الأداة وملاءمتها للمحتوى الذي تقيسه"، أي أنه لما كانت عناصر الأداة عبارة عن عينة لكافة المجالات التي يغطيها المحتوى فإن بينة الصدق تعتمد على مدى تمثيل العينة (الأداة) وكفايتها (الدوسري، 2000: 54).

" ويشير كروكر وألجينا Crocker & Algina 1986 إلى ضرورة الاهتمام بالاعتبارات التالية عند دراسة صدق المحتوى:

- إلى أي مدى كانت أوزان الأهداف تعكس أهميتها الفعلية؟
- ما مدى مطابقة فقرات أداة القياس لقائمة الأهداف؟
- ما جوانب الفقرة التي يجب فحصها؟
- كيف يمكن تلخيص نتائج وملاحظات الفاحص؟ " (النبهان، 2004: 275)

إن أدلة الصدق المرتبطة بالمحتوى مهمة بوجه خاص عندما نريد أن نصف كيفية أداء الفرد في مجال من المجالات التي مثلها المقياس، بمعنى أن الطريقة المرتبطة بالمحتوى تعتمد على مدى تمثيل فقرات المقياس تمثيلا سليما للمجال الذي نريد قياسه، "لذلك فان تحقيق صدق المقياس بهذه الطريقة يتطلب القيام بالخطوات الآتية:

- تحديد مجال المحتوى الذي نريد قياسه تحديدا واضحا ودقيقا بجميع أجزائه وعناصره.
- بناء مجموعة من الفقرات أو الأسئلة الممثلة لهذا المجال.
- تقديم المحتوى أو الفقرات أو الأسئلة إلى مجموعة من الخبراء ليقوموا بفحص الفقرات منطقيا وتقدير مدى صلاحيتها أو تمثيلها للمحتوى المراد قياسه" (أبو علام، 1987: 279-280).

فإذا كانت آراء الخبراء وأحكامهم من خلال العملية السابقة متفقة بنسبة كبيرة فهذا يعني أن المقياس يتمتع بصدق محتوى مرتفع، أما إذا تدنت المستويات التقديرية فإن صدق المقياس منخفض، وبصفة عامة فإنه عند بناء أداة لقياس قدرة معينة، فإنه يجب أن تكون ممثلة لتلك القدرة وللمحتوى.

يعد صدق المحتوى أكثر الأنواع التي تناسب الاختبارات التحصيلية، حيث يجرى تحليل للمحتوى الدراسي وتحديد للأهداف وبناء أسئلة تعكس عناصر المحتوى، وذلك عن طريق جدول المواصفات، ثم اختبار عينة من هذه الأسئلة والحكم على مدى تمثيلها للمحتوى عن طريق حساب درجة الاتفاق بين مجموعة من المحكمين، يشترط أن يكونوا متخصصين في المجال موضع القياس وألا يقل عددهم عن خمسة، ولا تقل درجة الاتفاق على كل سؤال عن 80% (الدوسري، 2000؛ مراد وسليمان، 2002).

يوجد نوعان من الصدق يتصلان بصدق المحتوى هما الصدق الظاهري والصدق المنطقي، حيث يهتم الصدق الظاهري بمضمون الفقرات والحكم على مدى علاقتها بمحتوى المادة، أما الصدق المنطقي فيتلخص في التأكد فيما إذا كانت فقرات المقياس تمثل مجال أداء يتعلق بسمة أو قدرة معينة.

فالصدق الظاهري (أو الصدق السطحي أو الشكلي)، يتمثل في الشكل العام للمقياس ومدى ملاءمته للغرض الذي وضع من أجله، ويتم التوصل إليه من خلال حكم مختص على مدى قياس الأداة للسمة المقاسة، حيث يعطى المقياس لأكثر من محكم، ويمكن تقييم درجة الصدق الظاهري للمقياس من خلال التوافق بين تقديرات المحكمين (يحسب عن طريق معامل إتفاق كاندل)؛ فإذا كان هناك توافق بين تقديراتهم وكانت بشكل عام متدنية فهذا مؤشر على ضعف الصدق الظاهري، فالمقياس الذي يقيس القدرة الحسابية مثلا لابد أن يغلب على فقراته الجوانب العددية والعمليات الحسابية بدلا من الصياغة اللغوية الخالية من الأعداد.

ترى أنستازي Anastasia 1976" بأن الصدق الظاهري ليس صدقا بالمعنى الحقيقي لأنه لا يشير إلى ما يقيسه المقياس فعلا وإنما ما يظهره المقياس في صورته السطحية أو الظاهرية"؛ ويرى سعد عبد الرحمن (1998) أنه ما هو إلا خطوة من خطوات صدق المحتوى، فهو ليس صدقا بالمعنى العلمي للكلمة، وإنما يدل على ما يبدو أن الأداة تقيسه من الظاهر وليس على ما تقيسه بالفعل، ويظهر مثل هذا الصدق في وضوح الأسئلة، ومدى علاقتها بالسمة التي تقيسها الأداة، وغالبا ما يقرر ذلك مجموعة من المحكمين المتخصصين في المجال الذي يفترض أنه تنتمي إليه أداة القياس.

بينما الصدق المنطقي (أو العيني) فيرتبط بتصنيف الفقرات وفقا لمجالات أداة القياس، ويتطلب هذا النوع تحديدا أدق للموضوعات التي يغطيها المقياس، "وكلما كانت هذه الموضوعات أكثر تحديدا فإنه يتوقع أن يكون الصدق العيني أعلى، "حيث يركز الصدق المنطقي على عدد الأسئلة، أما الصدق الظاهري فيركز على محتوى الأسئلة بصرف النظر عن عددها أو مدى تغطيتها للموضوعات أو الأهداف الرئيسية" (عودة، 1998).

بالرغم من أهمية صدق المحتوى فإنه يعتبر من أكثر مؤشرات الصدق عرضة لأخطاء التقدير لاعتماده على تقديرات المحكمين وتأثره بأحكامهم الذاتية؛ ويعلل عودة (1998) سبب الاعتماد على تقديرات المحكمين ضعف المؤشرات الإحصائية الموجودة، مثل قوة الارتباط بين أداتين متكافئتين تقيسان نفس المحتوى، وكذلك قوة الارتباط بين أسئلة الأداة، أو كما ذكر علام (2000) قوة الارتباط بين درجات أداة طبقت مرة قبل بدء عملية التعليم ومرة أخرى بعد الانتهاء منها.

معايير صدق المحتوى:

1- يجب تحليل المحتوى تحليلا دقيقا منظما، حتى يتم التأكد من أن جميع الجوانب الأساسية تغطيها أسئلة الأداة تغطية ملائمة وبالنسب الصحيحة حتى لا يكون هناك عدد كبير من الأسئلة في

جانب معين على حساب الجوانب الأخرى؛ ويتطلب ذلك وصف تفصيلي دقيق لمدى المهام والمثيرات والمواقف والمواد التي يستوعبها نطاق المحتوى، كما يتطلب أيضا توصيف الأهداف التعليمية أو التدريبية أو العلاجية، بشرط أن تصاغ في عبارات إجرائية سلوكية قابلة للملاحظة والقياس.

2- تحديد الإجراءات التي استخدمت في اختيار عينة الأسئلة (طرق تحديد صدق المحتوى) جدول المواصفات، أو طريقة التحكيم، أو أي إجراءات أخرى كالطرق الإحصائية.

3- يستخدم جدول المواصفات كطريقة من طرق صدق المحتوى، وذلك عن طريق الفحص الدقيق للكتب الدراسية من ناحية، ونواتج التعلم التي يجب اختبارها (في ضوء الأهداف) من ناحية أخرى، ثم تحديد الأهمية النسبية للموضوعات والأهداف، وعلى هذا الأساس يتم تحديد عدد الأسئلة التي ترتبط بكل هدف في كل موضوع، وذلك بإعداد جدول ثنائي (جدول مواصفات) توضح فيه الأهداف أفقيا والموضوعات رأسيا، ومنه يتضح عدد الأسئلة التي يشملها المقياس بالنسبة لكل موضوع من موضوعات المحتوى في علاقته بكل هدف من الأهداف التعليمية، وتحدد على شكل وزن نسبي لكل هدف وموضوع، والطريقة المتاحة لتحديد الوزن النسبي للأهداف هي استطلاع آراء الخبراء أو المحكمين، أما الوزن النسبي للمحتوى فيمكن تحديده موضوعيا في ضوء الوقت المخصص لتدريسه فعليا، ويفضل أن تكون جميع الحالات في صورة نسب مئوية.

2- الصدق المرتبط بمحك:

يشير الصدق المرتبط بمحك (أو الصدق التجريبي أو الإحصائي أو المعياري) إلى "مدى كفاءة المقياس في التنبؤ بأداء الفرد في أنشطة محددة، من خلال حساب معامل الارتباط بين درجات المقياس ومحك خارجي مستقل، هو السلوك أو النشاط نفسه الذي يقيسه المقياس" (عودة، 1998)، حيث يعتمد على مقارنة المقياس الحالي الذي طبق على الفاحص مع نتائج مقياس آخر لنفس العينة وبنفس

المواصفات، ومن هذه المقارنة يوضح مدى صدق المقياس الأول، ويسمى المقياس الآخر بالمحك، كما يدل الصدق المرتبط بمحك على قدرة المقياس على التنبؤ بسلوك المفحوص في موافق محددة أو تشخيص هذا السلوك.

يعرف المحك بأنه " مقياس أو اختبار مستقل خارجي لما تهدف الأداة إلى قياسه أو التنبؤ به أو تشخيصه، أو ميزان لتحديد مدى صلاحية الأداة" (أبو حطب وآخرون، 1997)، "وقد يكون المحك مقياسا آخرا يقيس السمة نفسها التي يقيسها المقياس الأول، أو تقديرات خارجية يصدرها المدرسون أو المسئولون عن العمل الذي يؤديه المفحوص، أو تقديرات الأصدقاء أو الزملاء، أو التقدير الذاتي الذي يعتمد على قيام الفرد بتقدير سماته الشخصية أو أدائه من خلال الحكم على تصرفاته كما يراها هو لا كما يراها الآخرون على وفق مقياس متدرج في درجة تقدير السمة المقاسة، قد يكون المحك أيضا مجموعة الأحكام التي أصدرها متخصصون في مدد طويلة ومتعاقبة بالنسبة لأنماط سلوكية يمكن الوثوق بها والاعتماد عليها" (عبد الرحمن، 1998: 185).

من أهم المشكلات التي تصادف الباحث عند تقدير الصدق المرتبط بمحك هو العثور على محك خارجي مناسب ذو صلة وثيقة بموضوع القياس، وخالي من التحيز، وثابت وصادق في درجاته (ثورندايك وهيجن Thorndike & Hagen، 1989).

بناء على المدة الزمنية التي تفصل بين تطبيق الأداة والمحك، هناك طريقتين من طرق الصدق المرتبط بمحك وهما: الصدق التلازمي والصدق التنبؤي.

أ- الصدق التلازمي:

يقصد بالصدق التلازمي (أو التزامني) مدى الارتباط (أو التلازم أو التطابق) بين نتائج المقياس الحالي ونتائج المحك، "من خلال تطبيقهما في وقت واحد أو في أوقات متقاربة جدا على مجموعة من الأفراد، ثم حساب معامل الارتباط بين درجتي المقياس الجديد والمحك" (سمارة وآخرون، 1989: 112).

أي الصدق التلازمي يهتم بجمع الأدلة التي تدل على حجم العلاقة بين درجات المفحوصين على مقياس ودرجاتهم على محك آخر، حيث لا يكون هناك فاصل زمني بين أداء المفحوصين على المقياس وأدائهم على المحك (مراد وسليمان، 2002).

يستخدم الصدق التلازمي لأغراض التشخيص وليس التنبؤ بنتائج المستقبل كما في الصدق التنبؤي، ويعد من أكثر أنواع الصدق استخداما في مقاييس الشخصية حينما يتوافر محك صادق يقيس الخاصية نفسها أو يؤشر عليها" (ربيع، 1994: 101)، كما نلجأ إليه كنوع من الحلول لمشكلات الصدق التنبؤي وما يتطلبه من وقت طويل (أبو حطب وآخرون، 1997)؛ وللقيام بذلك، فحسب الصيرفي (2003) هناك العديد من الطرق الإحصائية التي تستخدم حسب الهدف من عملية تحديد الصدق التلازمي وهي:

1. المجموعات المتضادة:

حيث يقسم أفراد العينة إلى مجموعتين أو أكثر حسب السمة التي تميز بينهم، وتحديد الفروق بينهم في السمات المميزة.

2. التقديرات:

تقوم هذه الطريقة على أساس إستخدام مقياس تقديري للسمة المراد قياسها والطلب من مجموعة من المحكمين إبداء وجهة نظرهم من حيث توفر هذه السمة في أفراد العينة.

3. معامل الارتباط مع غيره من المقاييس:

يحسب معامل الارتباط بين المقياس ومجموعة من المقاييس التي تقيس نفس السمة، ومن المهم أن يكون الارتباط إيجابيا ومرتفعا بينهم كدلالة على الصدق التلازمي للمقياس الجديد (بخش، 2007).

ب- الصدق التنبؤي:

يشير الصدق التنبؤي إلى مدى الدقة التي يمكن من خلالها تقديم تخمينات أو توقعات لمقدار وجود سمة معينة لدى الفرد تحدث في وقت آخر، من خلال مقياس يفترض أنه يقيس هذه الخاصية، "فهو مؤشر على مدى صلاحية المقياس الجديد في الإشارة إلى مستوى الإنجاز الذي يفترض أن يصل إليه الفرد في موضوع مماثل أو قريب، مما يساعد في التحقق من قدرته لاتخاذ قرارات مستقبلية" (إبراهيم، 1999).

"يستخدم هذا النوع من الصدق في مقاييس الانتقاء والتصنيف، لاسيما في اختبارات الاستعدادات" (الظاهر وآخرون: 1999: 134-135)، "ويكون في هذا النوع من الصدق جمع المعلومات على المحك بعد مدة طويلة نسبيا من جمع المعلومات على أداة القياس التي نريد تقدير صدقها" (عودة، 1998: 373).

أي أن الصدق التنبؤي يعني "قدرة المقياس وفاعليته في التنبؤ بنتيجة معينة في المستقبل، ويقوم على الموازنة بين درجات المفحوصين في المقياس ودرجاتهم على المحك" (الأنصاري، 2000: 97-98)، "وتتعلق أدلة الصدق التنبؤي بتقدير مدى صلاحية المقياس في التنبؤ بأداء الفرد المستقبلي الذي يقاس باختبار محك باستخدام درجات مقياس يطبق عليه في الوقت الحاضر، وبعبارة أخرى تتعلق هذه الأدلة بمدى اقتران تباين درجات مقياس منبئ بتباين مقياس محك" (علام، 2000: 195).

يقدر الصدق التنبؤي بعدة طرق هي: حساب معامل الارتباط بين درجات مقياس تنبؤي طبق على عينة من المفحوصين، ودرجات نفس المفحوصين على محك خارجي طبق عليهم بعد مرور فترة زمنية

طويلة، فإذا وجد ارتباطا قويا بين درجات المقياس التنبؤي والمحك، دل ذلك على قدرة الأداة على التنبؤ بالاختبار المستقبلي للمجموعة الإحصائية التي أخذت منها عينة المفحوصين (بخش، 2007).

ويمكن استخدام قيمة معامل الارتباط بين المقياس التنبؤي والمحك الذي يسمى بمعامل الصدق التنبؤي، وكذلك قيمة المتوسط والانحراف المعياري لدرجات كل منهما في التوصل إلى معادلة الإنحدار، التي تستخدم في التنبؤ بدرجة المفحوص في المحك بمعرفة درجته في المقياس التنبؤي، بشرط أن تكون العلاقة بين درجات المقياس التنبؤي والمحك خطية، ويمكن تقدير أخطاء التنبؤ بإستخدام معادلة الخطأ المعياري للتنبؤ؛ أما في حالة وجود أكثر من اختبار تنبؤي فإننا نستخدم نموذج الانحدار الخطي المتعدد (الطريري، 1998؛ علام، 2000؛ مهرنر ولهمن Mehrens & Lehmnn، 2003).

نشير أنه إذا اعتبرنا الدرجة الكلية للمقياس محكا راهنا، فيمكن اعتبار معامل الإرتباط بين فقرات المقياس ودرجته الكلية نوعا من صدق المحك، يطلق عليه صدق الإتساق الداخلي.

معايير الصدق المرتبط بمحك:

1- يحدد نوع المحك المستخدم سواء كان اختبارا، أو تقديرا، أو تحصيلا دراسيا، أو مجموعات متضادة، أو أداء فعليا.

2- يجب أن يكون المحك متعلقا بالوظيفة التي تقيسها الأداة، بحيث يكون بينهما ارتباط منطقي، كأن تقيس الأداة والمحك نفس السمة المقاسة، أو سمات نفسية مختلفة ولكن يفترض أن يكون بينهما ارتباط أو علاقة بناء على ما وضحته الأطر النظرية.

3- يجب أن تتوفر في المحك الجيد خاصية الثبات، فثبات المحك يؤثر إلى حد كبير في معامل الصدق أي الارتباط بين الأداة والمحك، لذلك من الضروري اختيار محك عالي الثبات عند حساب الصدق المرتبط بمحك.

4- يجب أن يكون المحك ثبت صدقه وفق ما تشير إليه البحوث والتقارير التي استند إليها البحث، ويتم الحكم على معامل الصدق الذي يشترط فيه أولا الدلالة الإحصائية طبقا للمستويات التالية:

أ- مستويات صدق مرتفع إذا بلغت قيمته 0,70 فأكثر

ب- مستويات صدق متوسط إذا تراوحت قيمته بين 0,30-0,69

ت- مستويات صدق منخفض إذا كانت قيمته أقل من 0,30

5- يجب تحديد الهدف من استخدام الأداة حتى نتمكن من تحديد الفترة الزمنية الفاصلة بين تطبيق الأداة وتطبيق المحك.

6- إذا كانت الأداة تقيس سمة نفسية واحدة فإنه يتم التحقق من الصدق بشكل كلي، أما إذا كانت الأداة تتكون من عدة مقاييس فرعية، بمعنى أنه تقيس أكثر من سمة حتى ولو ظهر قياس هذه السمات في أداة واحدة دون تقسيمها إلى مقاييس فرعية؛ كذلك عندما تكون الأداة متعددة الأبعاد أي تقيس سمة نفسية واحدة، ولكن تنقسم إلى عدة أبعاد، فإنه يتم التحقق من صدق كل مقياس فرعي أو كل بعد بشكل مستقل، ولا يكتفي بصدق الأداة بشكل عام.

3- صدق التكوين الفرضي:

يقصد بصدق التكوين الفرضي (أو صدق المفهوم أو البناء أو البنيوي أو النظري)، " الدرجة التي يقيس فيها المقياس بناءا نظريا أو سمة معينة دون غيرها أو مفهوما دون غيره، ويعتمد على التحقق التجريبي لدرجات المقياس من المفاهيم أو الافتراضات التي اعتمد عليها الباحث في بنائه" (Kroll, 1960 : 425)، وهناك عدة مؤشرات تدل على صدق التكوين الفرضي ينبغي تحديدها على شكل افتراضات أو مفاهيم نظرية قبل البدء ببناء المقياس، حيث "يؤشر مدى قياس المقياس النفسي لتكوين فرضي أو مفهوم نفسي معين من خلال التحقق التجريبي من مدى تطابق درجاته مع المفاهيم أو الافتراضات التي اعتمد عليها

الباحث في بناء المقياس" (Bechtoldt, 1959: 621) ، فإذا تطابقت نتائج القياس مع الافتراضات النظرية يمكن القول بأن المقياس صادق في بنائه، وفي حالة عدم تطابق النتائج التجريبية مع الافتراضات النظرية فهذا يعني أن المقياس غير صادق، أو إن هذه الإفتراضات خاطئة مما ينبغي إعادة النظر فيها.

تهتم هذه الطريقة بدراسة السمات النفسية والتي عرفها براون Brown (عودة، 1998) بأنها مجموعة من السلوكيات التي تميل إلى الحدوث معا، يطلق عليها مصطلح التكوين الفرضي، لأنه لا يمكن ملاحظتها بشكل مباشر، إنما يستدل عليها من خلال أنماط متعددة من السلوكيات المرتبطة بها.

فعند تطوير مقياس لقياس قدرة عقلية أو سمة نفسية معينة يجرى بناء أسئلة الأداة انطلاقا من إطارها النظري الذي يشمل على تعريف للسمة وللجوانب التي يمكن ملاحظتها للتعرف عليها، فنجمع الأدلة لإثبات علاقة نتائج الأداة بالإطار النظري للسمة والتي تهدف الأداة لقياسها (الدوسري، 2000).

وقد تختلف الافتراضات النظرية التي يعتمد عليها في بناء المقياس من سمة إلى أخرى بحسب طبيعة السمة أو الخاصية أو النظرية التي اعتمد عليها، "إذ هناك مؤشرات عديدة على صدق التكوين ينبغي تحديدها على شكل افتراضات أو مفاهيم نظرية يتبعها الباحث قبل البدء ببناء المقياس، ثم يطبق المقياس على عينة الصدق للتحقق من مدى تطابق درجاته مع هذه الافتراضات" (Adkins, 1974 : 125).

لقد ازداد اهتمام المتخصصين في القياس النفسي بصدق التكوين الفرضي وخاصة في مراحل بناء المقاييس، كونه متضمنا لكل أنواع أو مؤشرات الصدق الأخرى، فهو يتطابق معها في كثير من خطواته ومعانيه، وقد ظهر هذا التوجه جليا في مقولة " كرونباخ " Cronbach" : "أن الصدق كله واحد فقط، وأن الصدق هو صدق البناء" (Cronbach, 1960 : 126)

يشير كرونباخ وميل Cronbach & Meehl 1955 إلى " أن هناك خطوات للتحقق من صدق التكوين الفرضي تتمثل فيما يلي:

- وضع نظرية معينة يعتمد عليها المقياس، وتنص على فرضيات تفسير السلوك المتوقع.
- جمع البيانات للتحقق من صحة الفرضيات.
- تقييم النظرية في ضوء قدرتها الفعلية على تفسير البيانات تفسيرا جيدا" (علام: 2000، 219).

وقد بين (ساكس Sax وزيلر Zeller) أنه يمكن إيجاز عملية التوصل إلى أدلة تتعلق بصدق التكوين الفرضي في الخطوات التالية:

1. تبرير أهمية التكوين الفرضي من الناحية التربوية والنفسية على أن يكون التكوين الفرضي معرفا تعريفا إجرائيا ويدل على سمة قابلة للقياس.
2. الاستناد إلى نظرية تربوية أو نفسية أو اقتراح نموذج منطقي يوضح المفاهيم والعلاقات القائمة بينها.
3. التمييز بين التكوين الفرضي والتكوينات الفرضية الأخرى المماثلة لها.
4. التوصل إلى أدلة من مصادر متعددة باستخدام الطرق الارتباطية والتجريبية والمنطقية لتأكيد التكوين الافتراضي، وعندما تستخدم أدوات متعددة في هذا الشأن فإنه يمكننا الحصول على الصدق التقاربي للتكوين الفرضي.
5. التوصل إلى أدلة نتأكد منها أن التكوين الفرضي لا يرتبط بعوامل وقتية أو دخيلة لكي نحصل على الصدق التمايزي للتكوين الفرضي.
6. إجراء تعديلات مستمرة في التكوين الفرضي بما يتفق والأدلة والمعلومات الجديدة المتجمعة (علام، 2000: 228).

ويوجد العديد من الأساليب والطرق المستخدمة لجمع الأدلة التي تشير إلى صدق البناء التكويني، نذكر فيما يلي أكثرها استخدما (علام، 2000؛ الطريري، 1998؛ أبو حطب وآخرون، 1997):

أ- الطرق التي تعتمد على الارتباط:

1. الاتساق الداخلي: هدفه التأكد من أن الأداة تقيس مفهوما تكوينيا فرضيا واحدا، أي أن الأداة تتناول سمة نفسية واحدة. يتم ذلك بحساب معامل الارتباط بين كل سؤال من أسئلة الأداة والبعد الذي ينتمي إليه، أو بين الأبعاد الفرعية المكونة للأداة والدرجة الكلية للأداة، أو حساب معامل الارتباط بين كل سؤال من أسئلة الأداة والدرجة الكلية للأداة.

2. مصفوفة السمات (أو الطرق المتعددة): وتتمثل في الصدق التقاربي إذا كان الارتباط عالي بين أداة القياس وأدوات أخرى تقيس نفس السمة؛ والصدق التباعدي (أو التمايزي) إذا كان الارتباط ضعيف بين المقياس ومقاييس أخرى لا تقيس نفس السمة؛ وقد اعتبر كامبل وفيسك (Campbell & fiske, 1959) أن هذين النوعين معا يشكلان مؤشرا واحد على صدق البناء السيكولوجي للأداة.

3. التحليل العاملي: أي تحليل المفهوم إلى أبعاده بناء على مؤشرات: إما بنود أم أجزاء من إختبار أو إختبار، وذلك بهدف تلخيص المتغيرات المتعددة في عدد أقل تسمى (عوامل) بحيث يكون لكل عامل من هذه العوامل دالة تربطه ببعض (أو كل) هذه المتغيرات.

ب- الطرق التي تعتمد على الفروق:

1. المقارنة الطرفية: وهي تنقسم إلى قسمين: الأول مقارنة الأطراف في المقياس والمحك الخارجي، بمقارنة الثلث الأعلى في درجات المقياس بالثلث الأعلى في درجات المحك الخارجي، وكذلك الثلث الأدنى في المقياس بالثلث الأدنى في المحك، "ويكون المقياس صادقا إذا لم تكن هناك فروق ذات دلالة إحصائية بين متوسطات الدرجات العليا للمقياس والمحك، وكذلك الدرجات الدنيا. أما القسم الثاني فيتم مقارنة

الأطراف في المقياس فقط، بالاعتماد على درجات الثلث الأعلى والثلث الأدنى من المقياس (27% أعلى، 27% أدنى) وحساب الفرق بين المتوسطين، فإذا كانت هناك دلالة إحصائية بين المتوسطين دل ذلك على صدق المقياس" (عبد الرحمن، 1998: 191).

2. الفروق بين الجماعات: (المجموعات المتضادة) دراسة الفروق بين مجموعات مختلفة من الأفراد، بحيث تكون السمة مرتفعة لدى إحدى المجموعات ومنخفضة لدى المجموعات الأخرى، لأن التكوين سمة مجردة مفترضة تنعكس على أداء الأفراد للأدوات المختلفة، ولأن هذه السمات لا توجد بقدر متساو لدى كل الأفراد، فإنه من المنطقي افتراض أن الأفراد يختلفون في مدى ما لديهم من هذه السمات. ومثال ذلك أنه إذا أردنا قياس سمة الإقناع فإننا نختار مجموعة تتبلور فيها السمة كالمحللين السياسين، وجماعة تنخفض فيها هذه السمة كعمال المكتبات المكلفين بترتيب الكتب.

ج- الطرق التي تعتمد على التجريب:

تعتمد هذه الطرق على التدخل التجريبي لإحداث تغييرات في درجات المفحوصين في أداة ما، كوسيلة للتعرف على مدى تأثر الأداة بمتغيرات معينة مما يساعد على تأكيد بعض التفسيرات المتعلقة بنتائج المقياس أو رفضها؛ حيث تتم أكثر من معالجة تجريبية في مواقف مختلفة، فإذا تغيرت الدرجات والنتائج وفقا لذلك دل على أن الأداة تقيس السمة أو الخاصية المراد قياسها (مثال: سمة قلق الإمتحان).

الطرق التي تعتمد على التحليل المنطقي:

تعتمد هذه الطرق على الفحص الدقيق للأداة والأداء الذي تتطلبه، وإحداث تكامل بين نتائج هذا الفحص وبين النظرية التي تستند إليها الأداة وآراء المحكمين الذين لديهم خبرة ببناء أدوات القياس؛ كذلك من المفيد تحليل العمليات التي يستخدمها الأفراد في أدائهم أو في التوصل إلى إجاباتهم عن أسئلة الأداة (علام، 2000).

معايير صدق التكوين الفرضي:

1- يتم التعريف بالإطار النظري للسمة التي يفترض ارتباطها بنتائج الأداة.

2- يتم اشتقاق فرضيات حول نتائج الأداة من الإطار النظري للسمة المقاسة.

3- يتم التحقق من صحة الفرضيات المستمدة من الإطار النظري للسمة المقاسة منطقيا أو تجريبيا عن طريق اختيار الطرق ذات الصلة بالأداة والفرضيات، وكلما زاد عدد هذه الطرق والمؤشرات وحسن اختيارها كلما كان ذلك أفضل.

4- تحديد الطرق المستخدمة في التحقق من الصدق التكويني للأداة، سواء كانت: الإتساق الداخلي، التحليل العاملي، مصفوفة السمات (الصدق التقاربي أو التباعدي)، الصدق التمييزي (المقارنة الطرفية أو المجموعات المتضادة).

5- تعتبر الطرق المستخدمة في التحقق من الصدق التكويني للأداة مناسبة، في الحالات التالية:

أ- عندما تكون الأداة أحادية البعد واستخدم الباحث طريقة الاتساق الداخلي أو التحليل العاملي للكشف عن تجانسها.

ب- عندما تكون الأداة متعددة الأبعاد واستخدم التحليل العاملي للكشف عن أبعادها.

ج- عندما تكون السمة المقاسة تختلف اختلافا واضحا من مجموعة إلى أخرى وتم استخدام الفروق بين الجماعات للكشف عن هذه الفروق.

د- عندما تكون السمة المقاسة تختلف اختلافا واضحا من فئة عمرية إلى أخرى وتم استخدام الفروق بين الأعمار للكشف عن هذه الفروق. (علام، 2000).

العوامل المؤثرة في الصدق:

1- طول الأداة: كلما زاد عدد الأسئلة ارتفع معامل الصدق، لأن زيادة عدد الأسئلة يؤدي إلى شمول الأداة للمحتوى، ويقلل أيضا من أخطاء القياس.

2- طبيعة الأسئلة: غموض الأسئلة، عدم تدرجها في مستوى الصعوبة، الصياغة السيئة لها أو صعوبة لغتها كلها معوقات تقف أمام المفحوص أثناء إجابته على أسئلة الأداة، فقد يحصل المفحوص على درجات لا تعكس السمة التي تهدف الأداة إلى قياسها.

3- معامل الثبات: يتأثر صدق الأداة والمحك المستخدم في تقدير الصدق بمعامل ثباتها، 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.

4- تباين العينة: معامل صدق الأداة يتناسب طرديا مع تباين درجات المجموعة، بمعنى أنه كلما زاد تباين الدرجات أدى ذلك إلى زيادة قيمة معامل الصدق؛ لأن تباين الدرجات يعني اتساع مجال السمة المقاسة مما يدل على تغطية شاملة لمجال السمة المقاسة، ويؤدي هذا إلى ارتفاع معامل الارتباط مع المحك (إذا كانت درجاته متباينة أيضا) (عبد الرحمن، 1998؛ ملحم، 2005).

5- عوامل خارجية: مثل حالة الفرد النفسية والعقلية والجسمية أثناء تطبيق الأداة، ظروف التطبيق كالحرارة والوقت المخصص للإجابة؛ كلها عوامل قد تؤدي إلى انخفاض معامل الصدق.

ثانيا: الثبات

يعد الثبات من الخصائص السيكومترية الأساسية لأدوات القياس النفسية والتربوية، وعلى الرغم من أن الأداة الصادقة تعد ثابتة، فيما قد لا تكون الأداة الثابتة صادقة، إلا أن ذلك غير كاف، "وينبغي التحقق من ثبات الأداة على الرغم من مؤشرات صدقه لأنه لا توجد أداة تتسم بالصدق التام" (Brown, F.G, 1983 : 27).

هناك عدة معاني لمفهوم الثبات هي:

1- يعرف الثبات بأنه قدرة الأداة على قياس الدرجة الحقيقية، وبأقل قدر ممكن من الخطأ العشوائي، وهذا التعريف ناتج عن ارتباط الثبات بالنظرية التقليدية للقياس (الطريري، 1997؛ علام، 2000).

2- يشير ثبات الأداة إلى اتساق درجاتها في قياس ما يجب قياسه، وإعطاء نتائج مماثلة أو متقاربة لو كررت عملية القياس على الأفراد أنفسهم، وباختلاف العوامل والظروف الخارجية (عبد الرحمن، 1998: 163-164). بمعنى آخر "اتساق القياسات التي يتم الحصول عليها من نفس الأفراد عندما يتم إعادة اختبارهم باستعمال نفس الأداة في مناسبات مختلفة وتحت نفس الظروف" (Anastasi, 1997 :84).

3- يعد الثبات مؤشرا على درجة التجانس في نتائج الأداة، والذي يمكن أن يكون على نوعين هما: التجانس الداخلي الذي يشير إلى أن فقرات الأداة جميعها تقيس المفهوم نفسه، والتجانس الخارجي الذي يشير إلى استمرارية الأداة بإعطاء نتائج ثابتة بتكرار تطبيقه عبر الزمن (Holt,R.R, 1971 : 60-61).

4- مفهوم ثبات الأداة هو مفهوم إحصائي لا يمكن الاستدلال عليه من مجرد الفحص المنطقي للفقرات، كما هو الحال في بعض مؤشرات الصدق (الأنصاري، 2000: 114).

5- ثبات الأداة يعني خلوها نسبيا من الخطأ غير المنتظم (العشوائي) في القياس، "فالدرجة المستخلصة من أي مقياس لا تعبر تماما عن الأداء الحقيقي أو التباين الحقيقي للفرد، لتدخل بعض

المتغيرات الدخيلة المؤثرة في الظاهرة المقاسة في شكل زيادة في الدرجة أو نقصان، والتي تسمى بتباين الخطأ" (إبراهيم، 1999: 69).

6- المفهوم الإحصائي للثبات يشير إلى أنه كم من التباين الكلي في العلامات يمكن أن يكون تباينا حقيقيا، وهذه النسبة هي عبارة عن مربع معامل الارتباط بين الدرجة الحقيقية والدرجة الملاحظة" (عودة، 1998: 339-340).

7- معامل الثبات يحسب من درجات الأداة التي تتأثر بالموقف الذي تحدث فيه الأداة، لذا يكون الثبات موقفيا قد يختلف من موقف إلى آخر (Holt,R.R, 1971 : 62)، وأنه لا يمكن أن يكون تاما لأن الدرجات التي يحصل عليها الأفراد لا تمثل درجاتهم في كل السمة المراد قياسها وإنما جزءا منها.

8- الثبات صفة نسبية، وقيمة معامل الثبات تتراوح بين الصفر والواحد، وقبولها يعتمد على غرض المقياس وعلى دقة القرار المترتب، "حيث حدد "مهرنز ولهمان Mehrens & Lehman 1980" معامل الثبات المقبول للاختبارات المقننة الجمعية بأن لا يقل عن (0.65) " (Mehrens & Lehmann, 1986: 159) ، في حين يفضل بعض الباحثين أن يزيد معامل الثبات عن (0.70)، لأن معامل الاغتراب يكون فيه أقل من (0.50)، وبهذا تكون نسبة التباين المشترك أكثر من (0.50) نستطيع من خلالها تفسير وجود علاقة حقيقية بين المتغيرين (Lindquist, E.F, 1950: 57).

من خلال التعاريف السابقة يمكن تحديد التعريف التالي للثبات:

ثبات الأداة يعني قدرتها على قياس الدرجة الحقيقية بأقل قدر ممكن من الخطأ العشوائي، واتساق درجاتها في قياس ما يجب قياسه.

طرق تقدير الثبات:

تتعدد الطرق والأساليب المستخدمة لتقدير معامل ثبات أدوات القياس النفسية والتربوية، حيث تختص كل طريقة منها بتقدير نوعية معينة من تباين الخطأ، وتختلف كل طريقة عن الأخرى تبعا لاختلاف مصدر الأخطاء العشوائية –التي يود الباحث تقدير تأثيرها في الدرجات الملاحظة– التي تعتمد بدورها على طبيعة خصائص السمة التي تقيسها الأداة وأغراض استخدامها (علام، 2000)، وهذه الطرق هي:

1– طريقة إعادة الاختبار

تعتبر طريقة إعادة الاختبار أول الطرق التي استخدمها العلماء لتقدير الثبات، وتقوم هذه الطريقة على تطبيق أداة ما على مجموعة من الأفراد مرتين متتاليتين تحت نفس الظروف بينهما فاصل زمني مناسب، ثم حساب معامل ارتباط بيرسون بين درجات المفحوصين في التطبيقين للحصول على معامل ثبات الأداة (غانم، 1997)، أي تقدير إستقرار الأداة عبر الزمن، ولهذا يطلق على معامل الثبات المحسوب بهذه الطريقة بمعامل الاستقرار كما ذكر "إيبل وفرستن" (Ebel & Fristin, 1995)، ويسمى أحيانا بثبات الاستقرار.

وقد عرف فرج (2000: 300) معامل الاستقرار بأنه عبارة عن "تقدير لنسبة الاستقرار في درجة الفرد رغم التغيرات خلال مدى زمني مناسب".

فإذا كان معامل الاستقرار مرتفعا دل ذلك على أن أداء المفحوص في التطبيق الثاني لم يكن مختلفا عن أدائه في التطبيق الأول إلا بما يسمح به تأثير الأخطاء العشوائية، وإذا كان منخفضا دل على اختلاف الدرجات في المرتين، مما يعني أن الأداة غير ثابتة ولا يمكن الإعتماد عليها (الغريب، 1996).

تفترض هذه الطريقة أن السمة المقاسة ثابتة نسبيا، وأن التغيير في الدرجات بين التطبيقين راجع إلى الأخطاء العشوائية للقياس (علام، 2000).

وقد عدد فرج (2000) بعض مصادر الخطأ في هذه الطريقة كالتالي:

1. التذبذب العشوائي في أداء المفحوص الواحد في التطبيقين.

2. تعرض المفحوص لمتغيرات مختلفة خلال الفترة الزمنية بين التطبيقين.

3. التغيرات التي يمكن أن تحدث في موقف المقياس نفسه، كمتغير مكان التطبيق، أو قوة الإضاءة، وغيرها من الظروف الطبيعية المتغيرة.

4. بعض التغيرات التي يحتمل أن تتعرض لها السمة أو القدرة التي تقاس بهذه الأداة.

لذلك فإن معامل الثبات المحسوب بهذه الطريقة معرض لكثير من الأخطاء التي كانت السبب في توجيه العديد من الانتقادات، إذ أن تطبيق الأداة على المفحوصين في المرة الثانية يتأثر بالعوامل التالية:

1. اختلاف ظروف التطبيق في المرتين.

2. انتقال أثر التدريب والخبرة والألفة من التطبيق الأول إلى التطبيق الثاني.

3. الفترة الزمنية بين التطبيقين، حيث يؤثر هذا العامل بطريقتين:

أ- إذا كانت الفترة الزمنية بين التطبيقين طويلة، يتدخل أثر عامل النمو العقلي والجسمي والاجتماعي والوجداني الحادث لدى المفحوصين.

ب- إذا كانت الفترة الزمنية قصيرة، زاد احتمال تدخل عامل الذاكرة في التأثير على نتائج التطبيق الثاني، وخاصة في الأدوات القصيرة أو التي تتطلب وظائف عقلية عليا.

نتيجة للانتقادات السابقة سعى علماء القياس لإيجاد طريقة أخرى لتقدير معامل الثبات تتلافى عيوب الطريقة السابقة، فكانت طريقة الصور المتكافئة.

معايير استخدام طريقة إعادة الإختبار:

تتحدد معايير استخدام طريقة إعادة الاختبار على النحو التالي:

1- تستخدم في الأدوات التي تقيس سمة تتميز بالثبات النسبي خلال الفترة الزمنية الفاصلة بين التطبيقين مثل الإتجاهات والميول والقيم، وبعض القدرات العقلية الخاصة مثل الاستعدادات، والقدرة العقلية العامة (ذكاء البالغين)، وبعض سمات الشخصية (التكيف الشخصي والاجتماعي).

2- في حالة الأدوات التي تقيس سمات غير مستقرة مثل الحالة المزاجية، أو القلق، أو ذكاء الأطفال، والأدوات التي تهدف إلى قياس التذكر أو ترتبط ارتباطا مباشرا بهذه العملية العقلية، أو التي تتطلب أسئلتها استخدام العمليات العليا كالاستدلال والابتكار فإننا نستخدم طريقة إعادة الاختبار إما:

أ- بخفض الفترة الزمنية الفاصلة بين التطبيقين، بحيث لا تقل عن أسبوع ولا تزيد عن أسبوعين، إذا كانت العينة من الأطفال أقل من سبع سنوات أو الأشخاص الذين يعالجون بعقاقير تؤثر في ذاكرتهم، أو الأشخاص مرتفعي الذكاء.

ب- بزيادة الفترة الزمنية الفاصلة بين التطبيقين، بحيث تكون أكثر من أسبوعين، وأقل من ستة أشهر، إذا كانت العينة من البالغين أو كانت السمة المقاسة من النوع الذي يسهل تذكره، أو تعلمه أو الاستفادة منه في التدريب أو يتعرض أفراده فيه إلى الألفة.

3- يتناسب استخدام هذه الطريقة مع اختبارات السرعة.

4- تستخدم لتقدير ثبات التقارير الذاتية (أبو حطب وآخرون، 1997؛ علام 2000؛ فرج 2000؛ النبهان: 2004).

2- طريقة الصور المتكافئة (معامل التكافؤ)

تقوم الفكرة الأساسية لهذه الطريقة على إعداد صورتين متكافئتين لأداة القياس، وتطبيقهما معا أو بعد فترة زمنية قصيرة على نفس المفحوصين، ثم حساب معامل ارتباط بيرسون بين درجات المفحوصين في الصورتين (ملحم، 2005)، ويطلق على معامل الثبات المحسوب بهذه الطريقة معامل التكافؤ في حالة تطبيق الصورتين المتكافئتين في الوقت نفسه أو بفاصل زمني قصير.

ويقصد بالتكافؤ تساوي الصورتين من الناحية الإحصائية، وذلك بتساوي المتوسطات والإنحراف المعياري، وليس المقصود هنا تساوي القيم بل أن يكون الفرق غير دال إحصائيا، تشابه شكل التوزيع للدرجات على الصورتين، تساوي معامل الارتباط بين الأسئلة، وكذلك تساويها من ناحية بنية الأسئلة (مستوى الصعوبة، القدرة التمييزية، طول الأداة، طريقة صياغة الأسئلة، نوع السمة المقاسة، عدد مكونات السمة التي تقيسها الأداة، المحتوى الذي يقيسه السؤال)، وأخيرا توحيد إجراءات التطبيق في المرتين (التعليمات، طريقة الإجابة، التصحيح، زمن الإجابة) (أحمد، 1960؛ الغريب، 1996؛ عودة، 1998؛ عبد الرحمن، 1998).

إن مصدر تباين الخطأ في هذه الطريقة أكثر منه في طريقة الإعادة، فبالإضافة إلى مصادر الأخطاء في طريقة الإعادة، هناك مصدر آخر هو تباين الخطأ الناتج عن اختلاف محتوى الصورتين المتكافئتين.

وقد وجهت عدة انتقادات لهذه الطريقة، منها "صعوبة إعداد صورتين متكافئتين تماما، ومضاعفة الجهد والتكاليف في ذلك" (الأنصاري، 2000: 126)، "وأن عامل انتقال أثر التدريب والألفة بالمقياس يزيد كلما اقتربت الصورة من الأصل" (ربيع، 1994: 73). لذلك سعى علماء القياس لإيجاد طريقة أخرى لتقدير معامل الثبات لتلافي عيوب هذه الطريقة.

معايير استخدام طريقة الصور المتكافئة:

تتحدد معايير استخدام طريقة الصور المتكافئة على النحو التالي:

1- تساوي الصورتين من الناحية الإحصائية (المتوسطات، الانحراف المعياري، تساوي معاملات الارتباط بين الأسئلة)، وتساويها من ناحية بنية الأسئلة (مستوى صعوبة الأسئلة في حالة اختبارات الذكاء والقدرات العقلية والاختبارات التحصيلية والاستفتاءات، معامل شيوع استجابة الأسئلة في حالة اختبارات الشخصية، القدرة التمييزية، طول الأداة، طريقة صياغة الأسئلة، نوع السمة المقاسة، عدد مكونات السمة التي تقيسها الأداة، المحتوى الذي يقيسه السؤال) وتوحيد إجراءات التطبيق في المرتين (التعليمات، طريقة الإجابة، التصحيح، زمن الإجابة).

2- أن تكون الفترة الزمنية الفاصلة بين التطبيقين قصيرة، لكي لا تؤثر التغيرات الحاصلة للمفحوصين خلال هذا الفاصل الزمني في قيمة معامل التكافؤ.

3- تستخدم في حساب معامل الثبات لاختبارات الذاكرة والاختبارات التحصيلية المقننة لضعف تأثير عاملي الذاكرة والتدريب على قيمة معامل التكافؤ.

4- تستخدم في الدراسات التي تهدف لمعرفة أثر عامل تجريبي معين في الأداء الاختباري، خاصة إذا تطلب الأمر قياس قبلي وبعدي.

5- تستخدم في الاختبارات التي تهدف لقياس القدرة على حل المشكلات.

6- تستخدم عند تقويم الطلاب أو المناهج.

7- تستخدم في مواقف العلاج النفسي للتعرف على فاعلية العلاج.

8- يفضل استخدامها لتقدير ثبات الاختبارات التي تقيس ذكاء الأطفال.

9- لا يفضل استخدامها في اختبارات القدرات والميول والاتجاهات لصعوبة إعدادها.

10- لا يفضل استخدامها إذا كانت السمة التي تقيسها الأداة سمة عريضة ومركبة مثل القدرة على التفكير الناقد أو الإبداع (علام، 2000، النبهان، 2004: 240-241؛ Anastasi, 1997).

3- معامل الاستقرار والتكافؤ:

يمكننا جمع طريقتي إعادة الإختبار والصور المتكافئة بتطبيق إحدى الصورتين وبعد فاصل زمني طويل نسبيا تطبق الصورة الثانية؛ وعن طريق حساب معامل ارتباط بيرسون للدرجات في الصورتين، فإننا نقدر قيمة لمعامل ثبات الأداة تدل على كل من الاستقرار والتكافؤ (أبو حطب وآخرون، 1997: 150).

ولأن هذه الطريقة تجمع بين طريقة إعادة الاختبار وطريقة الصور المتكافئة، فإنها تؤدي إلى قيم تقديرية لمعامل الثبات أقل من غيرها من المعاملات المحسوبة بالطريقتين السابقتين، وتعد بمثابة الحد الأدنى لتقدير معامل الثبات، فهي تجمع بين مصادر تباين الخطأ (الأخطاء العشوائية) في الطريقتين (علام، 2000)، لذلك سعى علماء القياس لإيجاد طريقة أخرى لتقدير معامل الثبات لتلافي عيوب هذه الطريقة.

معايير استخدام طريقة معامل الاستقرار والتكافؤ:

تتحدد معايير استخدام طريقة معامل الاستقرار والتكافؤ على النحو التالي:

1- تساوي الصورتين من الناحية الإحصائية ومن ناحية بنية الأسئلة، وتوحيد إجراءات التطبيق في المرتين (مثلما أشرنا إليه في الطريقة السابقة).

2- أن تكون الفترة الزمنية الفاصلة بين التطبيقين طويلة.

3- تقدر قيمة معامل الاستقرار والتكافؤ عندما نرغب في التنبؤ بعيد المدى والاستدلال على محتوى دراسي معين (فرج 2000: 330؛ علام، 2000؛ 95: Anastasi, 1997).

4- معامل الاتساق الداخلي:

تتطلب الطرق الثلاث السابقة تطبيق أداة أو صيغ متكافئة منها مرتين بفواصل زمنية متباينة، غير أنه في بعض الأحيان يصعب بناء صيغتين متكافئتين، أو قد يصعب تطبيق الأداة مرتين، فعندئذ يفضل تقدير الثبات باستخدام أداة واحدة وتطبيقها مرة واحدة (علام، 2000).

ويمكن تقسيم طرق الاتساق الداخلي إلى:

1. طريقة التجزئة النصفية.
2. طريقة كودر-ريتشاردسون (20) (KR-20).
3. طريقة كودر-ريتشاردسون (21) (KR-21).
4. طريقة معامل ألفا كرونباخ.

أ- طريقة التجزئة النصفية:

يتطلب تقدير الثبات بهذه الطريقة الحصول على درجتين لكل فرد، وذلك بتقسيم الأداة إلى نصفين، بحيث نحسب درجة الفرد على كل نصف من الأداة، ثم يحسب معامل ارتباط بيرسون بين درجات الأفراد على النصفين، ويسمى معامل الثبات في هذه الطريقة بمعامل الاتساق الداخلي.

تتعدد مصادر تباين الخطأ في طريقة التجزئة النصفية، فهناك تباين الخطأ الناتج عن اختلاف التطبيق، والناتج عن التخمين، والناتج عن من يقوم بعملية تقدير الدرجة (موضوعية التصحيح)، وكذلك تباين الخطأ الناتج عن اختلاف معاينة محتوى كل من نصفي الأداة (Magnusson, 1967 : 113).

وقد وجهت لهذه الطريقة عدة انتقادات، منها أن القيمة التقديرية لمعامل الاتساق الداخلي المحسوب بطريقة التجزئة النصفية تختلف باختلاف الطرق المتعددة لتجزئة الأداة إلى جزئين متكافئين، فقد نستخدم النصف الأول مقابل النصف الأخير، إلا أن هذه التجزئة لا تعطينا نصفين متكافئين بسب الفروق في طبيعة الأسئلة ومستوى الصعوبة، بالإضافة إلى تدخل عوامل تختلف درجتها في بداية الاختبار

عن نهايته مثل الحماس والتعب والملل. ثاني طريقة للتجزئة هي استخدام الأسئلة ذات الأرقام الفردية مقابل الأسئلة ذات الأرقام الزوجية، وتعتبر طريقة موضوعية للحصول على قسمين متكافئين، وخاصة عندما تكون الأسئلة مرتبة حسب مستوى الصعوبة (ثورنديك وهيجن Thorndike & Hagen، 1989).

ومن الانتقادات أيضا أن القيمة التقديرية لمعامل الاتساق الداخلي المحسوب بطريقة التجزئة النصفية تمثل معامل ثبات نصف الأداة وليس كل الأداة، ولحل هذه المشكلة أوجد العلماء العديد من المعادلات لتصحيح معامل ثبات نصف الأداة ليدل على معامل ثبات كل الأداة، ومن هذه المعادلات:

- معادلة سبيرمان وبراون Spearman-Brown:

لقد اشتق سبيرمان وبراون معادلة رياضية عامة لتقدير معامل ثبات الأداة عندما تزيد عدد أسئلة الأداة عدد مرات (ن)، وفي حالة التجزئة النصفية فإن ن تساوي 2؛ حيث يحسب إحصائيا بقسمة ضعف معامل الثبات النصفي على نفسه مضافا إلى الواحد.

ويفترض لاستخدام هذه المعادلة تساوي تباين درجات كل من نصفي الأداة، وإلا فإن القيمة التقديرية للأداة ككل ستكون أكبر من القيمة التي يمكن الحصول عليها باستعمال المعادلات الاخرى لتقدير الاتساق الداخلي (علام، 2000).

ونتيجة لهذا العيب ظهرت معادلات أخرى تأخذ في الاعتبار اختلاف تباين درجات نصفي الأداة، من هذه المعادلات:

- معادلة رولون Rulon:

يعتبر رولون أن تباين درجات الفرق بين درجات نصفي الأداة هو تباين خطأ، حيث يرى أن "السبب في تباين درجات الأفراد على المقياس يرجع إلى تباين قدرات هؤلاء الأفراد واختلاف مستوياتهم الحقيقية فيما يتعلق بالسمة المقاسة، وإلى وجود أخطاء القياس العادية" (عبد الرحمن، 1998: 169).

وتعتمد هذه المعادلة على تباين الفرق بين درجات النصفين (Sd^2) والتباين الكلي (St^2)، حيث نقوم بقسمة تباين الفرق على التباين الكلي للأداة، ثم ننقص الحاصل من الواحد الصحيح؛ وبما أن الثبات هنا يحسب كاملا، فلا حاجة لاستخدام معادلة ثانية لتصحيح الطول، ويرجع ذلك إلى حقيقة أن الفرق بين درجتي الفرد على نصفي الأداة عبارة عن مجموع تباين الخطأ في النصفين معا، وليس في نصف واحد فقط، ولهذا الفروق بين النصفين خاصة بالأداة كلها.

- معادلة جتمان Gutman:

معادلة جتمان مثل معادلة رولون تأخذ في الاعتبار اختلاف تباين درجات نصفي الأداة، حيث يرى جتمان أنه "يمكن إهمال شرط تساوي التباين بين نصفي الأداة الذي تعتمد عليه معادلة سبيرمان وبراون" (ربيع، 1994، 74)، "وتعتمد هذه المعادلة على تباين نصفي الأداة والتباين الكلي لها" (الأنصاري، 2000: 125)، حيث نقوم بجمع تباين النصفين وقسمته على التباين الكلي للأداة، ثم ننقص الحاصل من الواحد الصحيح، ونضرب الناتج الكلي في إثنين، "ويؤدي استخدام معادلة جتمان أيضا إلى تقدير ثبات الأداة كاملا دون الحاجة لاستخدام معادلة أخرى لتصحيح الطول" (فرج، 2000).

القيم الناتجة باستخدام هذه المعادلة تتساوى مع الناتج من معادلة سبيرمان وبراون ومعادلة رولون إذا تساوي تباين درجات كل من نصفي الأداة (علام، 2000).

- معادلة هورست Horst:

تستخدم معادلة هورست في الحالات التي يصعب فيها تقسيم الأداة إلى نصفين متساويين (الطريري، 1997)، وتعتمد على نسبة الفقرات الفردية ونسبة الفقرات الزوجية إلى عدد الفقرات الكلي.

كما أن هناك عدة معادلات انطلقت من مبدأ معادلة سيرمان–براون، نذكر منها معادلة موزير Mosier، معادلة فلاناجان Flanagan، معادلة جلكسون Gulliksen ومعادلة رولسن Rohlsen.

معايير استخدام طريقة التجزئة النصفية:

تتحدد معايير استخدام طريقة التجزئة النصفية على النحو التالي:

1- تستخدم في الأدوات المتجانسة المحتوى التي تقيس سمة أحادية البعد، وإذا كانت الأداة غير متجانسة فإنه يمكن تجميع الأسئلة المتجانسة معا أو تقسيم الأداة إلى أدوات فرعية متجانسة في محتواها ثم تجزئة كل أداة فرعية إلى نصفين متكافئين؛ وإذا كانت الأداة تنقسم إلى أدوات فرعية متجانسة تقيس كل منها سمة معينة، نقوم بتقسيم كل أداة فرعية على حده (اختبارات التحصيل، القدرات العقلية).

2- إذا كانت الأداة تضم مجموعات الأسئلة التي تتعلق بمشكلة واحدة (مثل الأسئلة التي تتلو فقرة معينة في اختبار لفهم القراءة) فإن هذه الفقرة بأسئلتها جميعا يجب أن توضع في أحد نصفي الأداة.

3- لا تستخدم في الأدوات التي يكون الأداء عليها يعتمد بشكل تراكمي على الأداء في الخطوات السابقة، مثل اختبارات القدرة في حل المشكلات.

4- في حالة تجزئة الاختبار التحصيلي إلى نصفين حسب الأرقام (الأسئلة ذات الأرقام الزوجية مقابل الأسئلة ذات الأرقام الفردية)، يجب أن تكون الأسئلة مرتبة حسب مستوى الصعوبة.

5- يفضل استخدامها في الأدوات ذات الأسئلة الطويلة التي تصل إلى 60 فقرة أو أكثر، والأدوات القابلة للتجزئة النصفية (عدد الأسئلة زوجية).

6- لا تستخدم لتقدير ثبات التقارير الذاتية.

7- لا تستخدم لتقدير ثبات اختبارات السرعة، وإذا كان من الضروري ذلك، فإننا نقوم بتنصيف زمن الاختبار وليس الأسئلة (أبو حطب وآخرون، 1997؛ عودة، 1998؛ علام، 2000؛ فرج، 2000؛ Standards for Educational and Psychological Testing, 1999).

ب- طريقة كودر-ريتشاردسون (20):

تستخدم معادلة كودر وريتشاردسون Kuder & Richardson 20 (20) (KR-20) لتقدير معامل ثبات الأداة عن طريق حساب تباين الدرجات على كل سؤال من أسئلة الأداة ثم جمع هذه التباينات، ولقد عالجت هذه الطريقة أوجه قصور طريقة التجزئة النصفية، المتمثلة في صعوبة تجزئة الأداة إلى نصفين متكافئين (Aiken, 2000؛ الطريري، 1997).

يطلق على المعامل المحسوب معامل التجانس، وهو "قيمة تقديرية لمتوسط قيم معامل الارتباط بين كل من نصفي الأداة لجميع طرق التجزئة الممكنة دون أن نقوم بالتجزئة فعلا" (علام، 2000: 161)، وتعتمد هذه الطريقة على تجزئة الأداة إلى عدد كبير من الأجزاء، يتكون كل جزء من سؤال واحد فقط من أسئلة الأداة (فرج، 2000).

ويعاب على هذه الطريقة أن قيمتها تتضمن قدرا من التحيز يظهر في بعض الحالات، مثل إذا كانت الأداة قصيرة (عدد الأسئلة بين 9-18 سؤال)، أو إذا كانت الارتباطات بين الأسئلة شديدة الارتفاع (مابين 0,6-0,8)، كما أنها تتطلب حسابات كثيرة؛ مما أدى إلى ظهور معادلة أخرى تأخذ في الاعتبار سهولة تقدير معامل الثبات وهي معادلة كودر-ريتشاردسون (21) (فرج، 2000).

ج- طريقة كودر-ريتشاردسون (21):

تستخدم في حالة تقارب مستوى صعوبة الأسئلة، وقد اشتقت جبريا من معادلة (KR-20)، وتتطلب فقط حساب المتوسط والانحراف المعياري للدرجات الكلية.

إن معادلة (KR-21) تعطي دائما تقديرا أدنى عن معامل ثبات (KR-20)عندما تكون الأسئلة غير متقاربة في مستوى الصعوبة (Kuder & Richardson, 1937 :158).

لقد أدخل عدد من الباحثين عدة تعديلات على معادلة كودر –ريتشاردسون لمعالجة جوانب النقص فيهما، وتجاوز بعض الشروط التي لا يمكن أن تتحقق في غالب الأحيان، وإيجاد صيغة واحدة مهما كان عدد أوزان تدرج الأداة، ثنائية كانت أو متعددة. حيث قام تكر Tuker بإجراء بعض التعديلات على المعادلة بهدف تبسيط العمليات الحسابية التي تتطلبها (KR-20) وتجنب عدم الدقة التي تتضمنها (KR-21)، وذلك بإضافة متوسط نسبة تباين الإجابات الصحيحة على الأداة لمعادلة (KR-20) (فرج، 2000)؛ وأيضا معادلة دريسيل Dressel التي تعد تعديلا آخرا لمعادلتي كودر–ريتشاردسون في حالة الأدوات التي تتضمن نظاما للتصحيح يقوم على أوزان متدرجة للأسئلة؛ ولعل أشهر معادلة هي معادلة كرونباخ Cronbach (معامل ألفا) والتي سنتناولها بنوع من التفصيل.

معايير استخدام طريقة كودر–ريتشاردسون (20) و طريقة كودر–ريتشاردسون (21):

تتحدد معايير استخدام طريقتي كودر–ريتشاردسون على النحو التالي:

1- تستخدم في الأدوات المتجانسة المحتوى التي تقيس سمة أحادية البعد، وإذا كانت الأداة غير متجانسة فإنه يمكن تقسيم الأداة إلى أدوات فرعية متجانسة ثم حساب ثبات كل أداة فرعية على حده.

2- تستخدم لتقدير ثبات اختبارات القوة فقط.

3- تستخدم في الأدوات التي يكون نظام التصحيح فيها ثنائي أي إما واحد صحيح أو صفر.

4- يجب أن لا يكون عدد الأسئلة المتروكة كبيرا.

5- تستخدم معادلة كودر–ريتشاردسون (21) فقط إذا تساوت أسئلة الأداة في مستوى الصعوبة (حوالي 0.50) (أبو حطب وآخرون، 1997: 119 ؛ علام، 2000؛ فرج، 2000: 326؛ Standards for Educational and Psychological Testing, 1999 : 95).

د- طريقة معامل ألفا كرونباخ:

نظرا لكون معامل التجانس يقتصر استخدامه على المفردات ثنائية الدرجة، فقد اقترح كرونباخ Cronbach منذ عام 1951 معادلة تطورت بعد ذلك على يد كايزر و ميشيل Kaiser & Michael عام 1975 أطلق عليها اسم معامل ألفا كرونباخ، وفيها يحل مجموع تباين درجات جميع الأسئلة محل مجموع حاصل ضرب نسبة الأفراد الذين أجابوا إجابة صحيحة على كل سؤال ونسبة الذين لم يجيبوا على السؤال؛ وتعتبر معادلة (α Cronbach) تعميم لصيغة (K-R20) عندما تشتمل الأدوات على مفردات متعددة الدرجات، أين يستجيب الفرد لعبارات المقاييس على ميزان ثلاثي أو خماسي التدرج (أبو حطب، 1997: 119،122)، "وتعطي الحد الأدنى للقيمة التقديرية لمعامل ثبات درجات الأداة، وهي أنسب طريقة في البحوث المسحية كالاستبيانات أو مقاييس الاتجاهات ومقاييس الشخصية وفي حالة الاختبارات التحصيلية" (النبهان، 2004: 244)، "حيث تعتمد فكرة هذه المعادلة على تقسيم المقياس على عدد من الأجزاء يساوي عدد فقراته بحيث تشكل كل فقرة مقياسا فرعيا، وحساب الارتباطات بين درجات المقياس الكلي والمقاييس الفرعية" (عودة، 1998: 354)، حيث نقوم بحساب قيمة تباين كل مفردة ثم التباين الكلي لكل بعد ثم تباين الأداة ككل، وبعد ذلك إيجاد علاقة كل مفردة بالبعد ثم علاقة كل بعد بالدرجة الكلية وعلاقة كل مفردة بالدرجة الكلية، "ويمثل معامل ألفا (α) المتوسط الكلي للمعاملات الناتجة عن تجزئة الأداة إلى أجزاء بطرق مختلفة" (عبد الرحمن، 1998: 172-173).

ويعطي معامل ألفا الحد الأدنى للقيمة التقديرية لمعامل ثبات درجات الأداة، أي أن قيمة معامل الثبات عموما لا تقل عن قيمة معامل ألفا، فإذا كانت قيمة معامل ألفا مرتفعة، فإن هذا يدل بالفعل على ثبات درجات الأداة، أما إذا كانت منخفضة فربما يدل ذلك على أن الثبات يمكن أن تكون قيمته أكبر من ذلك باستخدام طرق أخرى.

وقد وجد كرونباخ أن هذا المعامل يعد مؤشرا للتكافؤ أي يعطي قيما تقديرية جيدة لمعامل التكافؤ، إلى جانب الاتساق الداخلي أو التجانس، "ومن الملاحظ أن معامل ألفا ومعامل التجانس لكودر- ريتشاردسون يتأثران بطول الأداة أي عدد مفرداتها ومدى تجانس هذه المفردات" (علام، 2000: 166).

معايير استخدام طريقة معامل ألفا كرونباخ:

تتحدد معايير استخدام طريقة معامل ألفا كرونباخ على النحو التالي:

1- تستخدم في الأدوات المتجانسة المحتوى التي تقيس سمة أحادية البعد، وإذا كانت الأداة غير متجانسة فإنه يمكن تقسيم الأداة إلى أدوات فرعية متجانسة ثم حساب ثبات كل أداة فرعية على حده.

2- تستخدم لتقدير ثبات اختبارات القوة فقط.

3- تستخدم في الأدوات التي يكون نظام التصحيح فيها يقوم على أوزان متدرجة للأسئلة مثل مقياس ليكرت بأنواعه، كما في استبيانات قياس الاتجاهات أو بعض مقاييس الشخصية، وكذلك في اختبارات التحصيل التي تشتمل على أسئلة مقال (أبو حطب وآخرون، 1997؛ عودة، 1998؛ علام 2000؛ فرج 2000؛Standards for Educational and Psychological Testing, 1999)

5. ثبات التقدير:

يحتاج كثير من الباحثين إلى مؤشرات لثبات التقدير وبخاصة عند قياس الشخصية التي تعتمد على أدوات الملاحظة أو المقابلة، وكذلك في قياس التحصيل الدراسي باستخدام اختبارات تشتمل على أسئلة المقال، حيث تعتمد جميعها على تقديرات المحكمين.

وللتعرف على مدى الثقة في هذه التقديرات، أي مقدار الخطأ الذي يعزى إلى فردية المحكمين، ويؤثر في الدرجات التي يقدرونها للمفحوصين، لابد من إيجاد معامل ثبات المحكمين (علام، 2000).

وأبسط الطرق هو إيجاد معامل الإرتباط بين تقديري محكمين مختلفين باستخدام معامل الارتباط الرتبي لسبيرمان، أو تقدير درجة الإتفاق بين المحكمين باستخدام معادلة كابا لِـ كوهين (عودة، 1998). كما أن هناك أيضا طرق أخرى مثل طريقة هولستي، سكوت، كريبندورف، وطريقة كوبر.

أما في حالة تعدد المحكمين فهناك عدة طرق، وهي كالتالي:

1- نحسب معامل الارتباط بين تقديرات كل المحكمين، ثم نحسب متوسط معاملات الارتباط الناتجة، وباستخدام معادلة سبيرمان وبراون نحصل على القيمة التقديرية لثبات متوسط تقديرات المحكمين.

2- استخدام طريقة جيلفورد Guilford، التي تعتمد على فكرة تحليل التباين الراجع إلى التباين بين المفحوصين الذين طبقت عليهم الأداة، وتباين الخطأ.

3- استخدام معادلة الغريب (1985)، والتي تعتمد على حساب متوسط الاتفاق بين المحكمين.

4- استخدام طريقة الصيرفي والصياد (1989)، التي تهدف إلى الكشف عن المحكمين الذين يتميزون بدرجة مقبولة من الاتساق مع أنفسهم كفرادى أو كمجموعات، بحساب متوسط الاتفاق بينهم.

معايير ثبات التقدير:

تتحدد معايير ثبات التقدير على النحو التالي:

1- تستخدم لتقدير اختبارات ومقاييس الشخصية.

2- تستخدم لتقديرات اختبارات التحصيل التي تتضمن أسئلة مقال.

4- تستخدم لتقدير أداة الملاحظة التي تشتمل على قوائم ملاحظة أو مقاييس التقدير (أبو حطب وآخرون، 1997؛ عودة، 1998؛ علام، 2000؛ فرج 2000؛ بخش، 2007).

6. طرق إحصائية أخرى:

أ) التحليل العاملي:

تقوم هذه الطريقة الإحصائية على تحديد أقل عدد من العوامل أو التكوينات الفرضية التي تفسر معاملات الإرتباط بين مفردات الأداة، وتصنيفها في مجموعات متجانسة تقيس كل واحدة منها سمة معينة، فإذا تم إيجاد عامل مشترك فإن هذا يعني أن مفردات الأداة متجانسة فيما تقيسه" (علام، 2000: 167).

معايير التحليل العاملي:

تتحدد معايير التحليل العاملي على النحو التالي:

1- نستخدم التحليل العاملي للكشف عن العوامل التي تتكون منها الأداة، عندما تكون الأداة عبارة عن اختبار متعدد الأبعاد أو يقيس أكثر من سمة.

2- يفضل أن يكون عدد أفراد العينة من خمسة إلى عشرة أضعاف عدد مفردات الأداة على الأقل، فكلما كان التباين بين أفراد العينة أكبر كانت العوامل الناتجة أكثر وضوحا ودلالة.

ب) تحليل التباين:

يمكن تقدير ثبات درجات الأداة باستخدام معادلة هويت Hoyt لتحليل التباين، لتعطي قيمة تدل على مدى تجانس أسئلة الأداة أو الاتساق الداخلي لها، وهذه الطريقة تفترض أن درجة الفرد على الأداة تقسم إلى أربعة عوامل: فهناك عوامل تتعلق بكل الفقرات عند كل الأفراد، وعوامل تتعلق بالفقرة، وعوامل تتعلق بالفرد، وعامل الخطأ؛ ويمكن إيجاد معامل الثبات في هذه الحالة بحساب خطأ التباين وقسمته على التباين بين الأفراد، ثم ننقص الناتج الكلي من الواحد الصحيح (علام، 2000).

ونشير إلى أن هناك طرق إحصائية متقدمة مثل التحليل العاملي التوكيدي (النمذجة)، ثبات المفهوم (المركب) والتباين المفسر (تيغزة، 2012).

القيمة المقبولة لمعامل الثبات:

معامل الثبات عبارة عن نسبة تباين لهذا يمكن أن يأخذ أي قيمة ضمن المدى (0-1)، إلا أننا لا نستطيع أن نحدد قيمة لمعامل الثبات ونعممها على كل أدوات القياس، فالأدوات تتفاوت في القيمة المقبولة حسب الغرض منها، حيث يعتمد ذلك على دقة القرار الذي سيترتب على النتائج، وهي كالتالي:

1- في حالة قياس مفاهيم نظرية افتراضية لأول مرة، فإن قيمة معامل الثبات المقبولة تتراوح بين 0,50-0,60.

2- في البحوث الأساسية قيمة معامل الثبات المقبولة تتراوح بين 0,70-0,80.

3- الأدوات المقننة تتطلب معاملات ثبات لا تقل عن 0,85 إذا كانت القرارات على مستوى أفراد، اما إذا كانت القرارات على مستوى جماعات فإن القيم المقبولة لمعامل الثبات تصل إلى 0,65.

4- الاختبارات التحصيلية المقننة يجب أن لا يقل معامل ثباتها عن 0,85 بينما يمكن أن يقل عن ذلك في اختبارات الشخصية (عودة، 1998؛ بخش، 2007).

ويتم الحكم على مستوى معامل الثبات للأداة وفق تصنيفات خمسة (تصنيف مرفي Murphy):

1- مستوى ثبات مرتفع جدا (أكثر من 0,95)

2- مستوى ثبات مرتفع (0,85-0,95)

3- مستوى ثبات متوسط (0,75-0,84)

4- مستوى ثبات منخفض (0,65-0,74)

5- مستوى ثبات غير مقبول (أقل من 0,65) (القضيب، 2001).

العوامل المؤثرة في الثبات:

إن قيم معامل الثبات – أيا كانت الطريقة المستخدمة في حسابها– ليست قيما مطلقة وإنما قيما تقديرية تؤثر فيها عوامل متعددة، ينبغي مراعاتها في تصميم أدوات القياس وعند انتقاء هذه الأدوات واستخدامها وتفسير نتائجها (علام، 2000)، حيث أن هناك عوامل خاصة بالأداة نفسها، وعوامل خاصة بالأفراد المفحوصين، وعوامل أخرى خارجية، وهي كالآتي:

أولا: العوامل التي تختص بالأداة:

1- طول الأداة:

توجد علاقة طردية بين طول الأداة وقيمة معامل الثبات، وذلك لأن أسئلة الأداة عبارة عن عينة تمثل السمة التي تقيسها الأداة وكلما كانت العينة كبيرة كانت الأداة أكثر دقة في قياس السمة (عبد الرحمن، 1998)، كما أن زيادة عدد الأسئلة يسمح للأخطاء العشوائية الموجبة والسالبة أن تلغي بعضها بعضا مما يجعل الدرجة الملاحظة للمفحوص تقترب من درجته الحقيقية (علام، 2000).

وهناك طريقة أخرى لتحديد العلاقة بين طول الأداة ودرجة ثباتها تبنى على حقيقة مهمة وهي: "إذا زاد طول الأداة ن مرة فإن التباين الحقيقي لدرجاتها يزيد ن^2 مرة ويزيد تباين الخطأ ن مرة"، وزيادة التباين الحقيقي يؤدي إلى معامل ثبات أكبر (عبد الرحمن، 1998: 180).

2- مستوى صعوبة الأسئلة:

توجد علاقة طردية بين مستوى صعوبة أسئلة الأداة وقيمة معامل الثبات، فإذا كانت الأسئلة سهلة جدا أو صعبة جدا لا تصلح الأداة لقياس الفروق الفردية بين الأفراد وكان معامل الثبات منخفضا، ويعود السبب إلى أن الأسئلة السهلة أو الصعبة تؤدي إلى انخفاض تباين الأسئلة وبالتالي انخفاض التباين الكلي للأداة، وبما أن تباين الخطأ ثابت فإن ناتج قسمة تباين الخطأ على التباين الكلي سوف يكون كبير، وعند

طرحه من الواحد الصحيح للحصول على معامل الثبات فإن معامل الثبات يكون منخفض، لذلك يفضل أن تشمل الأداة أسئلة تتراوح في مدى صعوبتها بين 25%-75% (بخش، 2007).

3- خصائص أسئلة الأداة:

تؤثر خصائص أسئلة الأداة في قيمة معامل الثبات للأداة، فالأسئلة الغامضة أو غير واضحة الصياغة أو الطويلة أو الإيحائية؛ كلها تعمل على خفض قيمة معامل الثبات.

ثانيا: العوامل التي تختص بالأفراد المفحوصين:

1- الحالة الصحية والنفسية:

تؤثر حالة الفرد الصحية والنفسية كشعوره بالتعب أو المرض أو التوتر على أدائه في الأداة، مما يؤدي إلى خفض قيمة معامل ثبات الأداة.

2- تجانس العينة:

يقصد بذلك تقارب مستوى أفراد العينة في الأداء، ويؤدي تجانس أفراد العينة إلى انخفاض القيمة التقديرية لمعامل الثبات، وذلك لأن قيمة معامل الثبات تعتمد على مدى الفروق بين أفراد العينة، فكلما قلت الفروق قل تباين الدرجات الحقيقية للمفحوصين، وبالتالي تقل قيمة معامل الثبات (علام، 2000).

3- مستويات قدرات المفحوصين:

تتاثر قيم معامل الثبات بمستوى قدرات المفحوصين، حيث ينبغي مراعاة انتقاء أدوات يعتمد تقدير ثبات درجاتها على عينة من الأفراد مستوى قدراتهم في نطاق مستوى المجموعة التي نهتم بدراستها، حتى لا تتأثر قيم معامل الثبات باختلاف مستوى قدرة المفحوصين (علام، 2000).

تطبيق أداة تكون سهلة بالنسبة لمجموعة من المفحوصين، لن تقيس الفروق الفردية بينهم، أما في حالة إذا كانت الأداة صعبة لمجموعة أخرى، فإنهم قد يلجأون إلى التخمين أثناء الإجابة على الأسئلة، مما قد يخفض قيمة معامل الثبات (ثورندايك وهيجن Thorndike & Hagen، 1989).

4- التخمين:

يلجأ بعض المفحوصين في حالة عدم تأكدهم من الإجابة الصحيحة في اختبارات الاختيار من متعدد أو الصح والخطأ إلى التخمين، الذي قد يختلف مكانه عند إجراء الأداة مرة أخرى، مما يقلل الارتباط بين درجات مرتي التطبيق، فتنخفض قيمة معامل الثبات، لذلك تستخدم معادلة تصحيح أثر التخمين للوصول إلى الدرجة الحقيقية للمفحوص.

ثالثا: عوامل خارجية:

1- زمن الاختبار:

تشير أبحاث لينكويست وكوك Lindquist & Cook إلى أن معامل الثبات يزداد تبعا لزيادة الزمن المستغرق للإجابة على فقرات الأداة يصل الحد المناسب، فيصل الثبات إلى نهايته العظمى ثم يقل تبعا لذلك كلما زاد الزمن عن ذلك الحد (ملحم، 2005).

2- موضوعية التصحيح:

عند تصحيح بعض أنواع الأدوات كالاختبارات المقالية أو الإسقاطية أو بعض مقاييس الشخصية فإنه تتدخل عوامل عديدة كالذاتية أو التحيز في تقدير الدرجات، مما يقلل من قيمة معامل الثبات. للتغلب على هذه المشكلة يجب أن يتضمن دليل الأداة وصفا دقيقا لشروط الإجابة المقبولة ومحكات هذه الاستجابة، بالإضافة إلى اللجوء إلى تقدير معامل ثبات التقدير الذي يشير إلى مدى تطابق تقديرات المصححين عند تصحيح الأداة، ومثل هذا الإجراء يؤدي إلى رفع معامل الثبات (فرج، 2000: 344).

العلاقة بين الصدق والثبات:

يرتبط مفهوم الصدق بمفهوم الثبات ارتباطا وثيقا، فكلاهما وجهان لشيء واحد هو مدى صلاحية ذلك المقياس في أن يقيس ما وضع لقياسه وفي إعطائه نتائج متماثلة، حيث أن الثبات يبحث في مدى اتساق مفردات الأداة، بينما الصدق يتعلق بالغرض الذي بنيت من أجله هذه الأداة، لذا ينبغي في أداة القياس أن تكون صادقة وثابتة في آن واحد، فالعلاقة بينهما يفترض أن تكون إرتباطية عالية.

إذا نظرنا إلى مفهوم الصدق الذاتي (قيمة مؤشر الثبات تعد الحد الأعلى لقيمة معامل الصدق) فإن الأداة الثابتة–أي إذا كان معامل ثباتها عاليا– هي أيضا أداة عالية الصدق من الناحية النظرية، ولكن قد تكون غير ذلك تماما من الناحية العملية التطبيقية، لأن تفسير الدرجة المستخرجة من أداة غير ثابتة سيكون بالتأكيد تفسيرا خاطئا؛ أما الأداة الصادقة فلابد وأن تكون أداة ثابتة من الناحية النظرية والتطبيقية، فلهذا فإن كل أداة صادقة فهي ثابتة.

كما أن معامل الاتساق الداخلي أو معامل التجانس للأداة يقدم دليلا آخر على صدق التكوين، فزيادة على كونه إحدى طرق حساب الثبات، فهو أيضا طريقة للتأكد من صدق الأداة التكويني، أين تعتبر الدرجة الكلية كمحك، وتحسب الإرتباطات بينها وبين المكونات الفرعية للإداة.

"ويرى كل من "كرونباخ Cronbach" و "جليزر "Glaser أن الصدق والثبات مفهومان مترابطان ويمكن أن يندرجان تحت اسم "مقايس قابلة للتعميم Generalizability Measures"، والفرق الرئيسي يكمن في الأبعاد التي نود التعميم لها" (علام، 2000).

II) الدراسات السابقة

تمهيد: يتضمن هذا القسم الإشارة للدراسات السابقة التي اطلع عليها الباحث وأفادته في بعض إجراءات بحثه، وبتتبعنا لأدبيات البحث العلمي وجدنا العديد من الدراسات التي تناولت تقييم البحوث النفسية والتربوية في المجال المنهجي، حيث وجدنا دراسات تقييمية للبحوث بصفة عامة، من ناحية الإشكالية، الفرضيات، الأساليب الإحصائية وغيرها؛ ودراسات تقييمية لأحد أدوات القياس، مثل الإختبارات التحصيلية أو الملاحظة؛ وهناك دراسات تناولت تقييم أدوات القياس بصفة عامة، سواء من الناحية المفاهيمية أو من ناحية أساليب التحقق وطرق تقدير الخصائص السيكومترية (الصدق والثبات). سنكتفي بتفصيل الدراسات التي تناولت تقييم طرق تقدير الصدق والثبات للأدوات المستخدمة في البحوث النفسية والتربوية، سواء الدراسات التي اهتمت بتقييم إحدى الأدوات أو بشكل عام.

تعتبر الإختبارات أكثر الأدوات التي حظيت بالمراجعة، لأنها تمثل أحد الأدوات الرئيسية التي تستخدم لجمع المعطيات لغرض القياس النفسي أو التقويم التربوي، وتستخدم على نطاق واسع سواء في مجال البحث العلمي، أو في المجال العملي: في المدارس، الجامعات، المستشفيات، المراكز التربوية...إلخ. حظي موضوع تقييم الإختبارات التحصيلية خاصة، باهتمام واسع وملحوظ على المستوى العالمي، ويظهر ذلك جليا من خلال العدد الكبير للدراسات التي تناولت تقييم الاختبارات التحصيلية على جميع المراحل التعليمية وفي جميع المواد الدراسية، خاصة في مادة الرياضيات. فهناك دراسات قومت أسئلة الاختبارات على ضوء تمثيل مفرداتها لمستويات المجال المعرفي، مثل دراسة العبد الوهاب (1973)، دراسة دانيال (1980)، دراسة العقلا (1986)، ودراسة العمر (1997)؛ ودراسات أخرى قومت أسئلة الاختبارات على ضوء تمثيلها للمحتوى الدراسي بالإضافة إلى مدى تمثيلها لمستويات المجال المعرفي حسب تصنيف

بلوم، مثل دراسة الغانم (1992)، دراسة المقبل (1994)، دراسة العتيبي (2001)؛ وهناك دراسات قومت أسئلة الاختبارات على ضوء معايير قام الباحث ببنائها، مثل دراسة إبراهيم (1988)، دراسة فاطمة الحمضي (1993)، دراسة وفاء باسروان (1995)، ودراسة مشاعل البكر (1998).

أما في البحوث النفسية، فتعتبر أداة الملاحظة من الأدوات التي يكثر استخدامها خاصة في الدراسات الإكلينيكية وأبحاث النمو؛ ومن الدراسات التي هدفت إلى التعرف على الطرق المستخدمة لتقدير ثبات أداة الملاحظة، دراسة قام بها أوتنباشير (Ottenbacher,1995) من خلال فحص ثلاثين دراسة منشورة في ثلاث مجلات لنمو الأطفال، بدءا من مجلد عام 1992 إلى 1995؛ وقد أوضحت نتائج الدراسة أن هناك أربع طرق لتقدير ثبات أدوات الملاحظة، وهي ثبات الملاحظين (الخبراء) استخدمت بنسبة 45%، وطريقة الإعادة استخدمت بنسبة 34%، وطريقة الاتساق الداخلي استخدمت بنسبة 11%، وطريقة الصور المتكافئة استخدمت بنسبة 10%، وكان معامل ارتباط بيرسون بين الزمن والنتائج (لأنها دراسة تتبعية) هي الطريقة المستخدمة لتقدير ثبات الملاحظين، وهي طريقة غير مناسبة لذلك؛ حيث أن معامل ارتباط بيرسون هو مقياس للثبات وليس لتقديم معلومات حول الإتفاق بين الملاحظين؛ وهذا يعني وقوع الباحثين في خطأ تفسير النتائج.

أما بالنسبة للدراسات التي تناولت موضوع صدق وثبات أدوات جمع المعطيات بشكل عام، فقد قام هميرسلي (Hamersley,1987) بدراسة بعنوان "بعض الملاحظات التي تتناول مصطلحي الصدق والثبات" استعرض فيها عينة من التعريفات الخاصة بمصطلحي الصدق والثبات الموجودة في مجموعة من كتب القياس، بهدف التعرف على المعاني المعطاة لهذين المصطلحين وكيفية ارتباطها بأهداف القياس؛ وقد أوضحت نتائج الدراسة أنه يوجد مجموعة من الأفكار المتنوعة حول المفهومين، كما أنهما يستخدمان استخداما غامضا وغير منسقا، وأغلبها تهتم بالطرق التي يتم عن طريقها تقدير الصدق والثبات ولم تحظ

القضايا الفكرية بنفس القدر من الاهتمام، كما أنه توجد أبعاد جوهرية وقدرا من التداخل بين التعريفات المختلفة الخاصة بكلا المفهومين.

ومن خلال الإجابة على مجموعة الأسئلة التي طرحها حول المفهومين، وضح التناقضات الموجودة في التعريفات الخاصة بكلا المفهومين وهي كما يلي:

1- ترتبط معظم مفاهيم الصدق والثبات بأهداف القياس بينما يرتبط البعض الآخر بطريقة القياس.

2- غالبا ما ينسب مصطلح الصدق إلى الأدوات وأحيانا إلى النتائج أو الفاحص، أما الثبات يعتبر غالبا سمة للنتائج وأحيانا ينسب إلى الأدوات أو الفاحص.

3- معظم تعريفات الصدق تشير إلى المدى الذي يتم فيه قياس الخصائص بواسطة الاختبارات المخصصة للقياس، أي الصدق يعرف في إطار العلاقة بين النتائج (الدرجات) والتباين في الخاصية التي يتم قياسها. أما الثبات يشير إلى الدرجات التي نحصل عليها عند تكرار قياس السمة باستخدام نفس الأداة، أي أنه يعرف في إطار العلاقات بين الدرجات التي يتم الحصول عليها باستخدام أداة واحدة أو أكثر.

وفي افتتاحية مجلة القياس التربوي والنفسي لعام 1994، نادى ثمبسون Thompson بضرورة استخدام لغة صحيحة ومناهج علمية وفنية مناسبة عند مناقشة مفهومي الصدق والثبات للمقاييس في الدراسات الخاصة بالصدق والثبات -بعد أن حدد الأخطاء التعبيرية والمنهجية-. لذلك قام ويتا ودانيال (Witta and Daniel, 1997,1998) بدراستين متتاليتين بعد نداء ثمبسون Thompson، وكذلك للتعرف على الطرق المستخدمة لتقدير صدق وثبات المقاييس؛ أما الدراسة الثانية فقد هدفت إلى تحديد مدى التأثير الذي أحدثه نداء ثمبسون Thompson على نسبة ظهور الأخطاء التعبيرية والمنهجية المتعلقة بالصدق والثبات، بالإضافة إلى التعرف على الطرق المستخدمة لتقدير الصدق والثبات في 150 دراسة

أخرى وفي نفس المجلات ولكن في مجلدات أعوام 1995، 1996، 1997 وهي الأعوام التي تلت نداء ثمبسون Thompson؛ وقد أوضحت نتائج الدراسة الأولى مايلي:

1- تحتوي أكثر من 50% من الدراسات على أكثر من خطأ تعبيري فيما يتعلق بالصدق والثبات، حيث أن (76) دراسة بنسبة 50,7% ذكرت عبارة الاختبار صادق، و(65) دراسة بنسبة 43,3% ذكرت عبارة الاختبار ثابت، و(38) دراسة بنسبة 50% ذكرت أن النتائج تثبت صدق الاختبار، و(06) دراسات بنسبة 4% ذكرت أن النتائج تثبت ثبات الاختبار، وعند التقييم حسب المجلة والعام، وجد أيضا أن أكثر من 50% من الدراسات قد أحتوت على أخطاء تعبيرية، وتندرج في أكثر من فئة واحدة.

2- خلو عينة الدراسة من الأخطاء المنهجية، ما عدا دراسة واحدة أشارت إلى أن التحليل الإحصائي يوضح أو يثبت الصدق والثبات.

3- توجد مجموعة من المقاييس تم التحقق من صدقها بنسبة 52,67%، وقد كانت طريقة ألفا كرونباخ أكثر الطرق شيوعا بنسبة 81%، ثم إعادة الاختبار بنسبة 23%، ثم ثبات التقدير، التجزئة النصفية، الصور المتكافئة بنسبة 6,3%-2,5%-1,3% على التوالي.

4- توجد مجموعة من المقاييس تم التحقق من صدقها بنسبة 61,3%، وكانت أكثر الطرق شيوعا في التحقق من الصدق هي طريقة صدق التكوين الفرضي بنسبة 55,4%، ثم طريقة الصدق التلازمي، الصدق التنبؤي، صدق المحتوى بنسبة 30,4%-26,1%-3,3% على التوالي، كما كانت أكثر طرق صدق التكوين الفرضي شيوعا هي طريقة التحليل العاملي الاستكشافي بنسبة 37%، ثم الصدق التمييزي، التحليل العاملي التوكيدي، السمات المتعددة، بنسبة 28,3%-20,7%-3,3% على التوالي.

وبعد دمج الدراستين (1997، 1998) في الدراسة الثانية وتصنيف في الدراسات حسب المتغيرات التالية: اسم المجلة، العام، الأخطاء التعبيرية، الأخطاء المنهجية، تم التوصل إلى النتائج التالية:

1- توجد علاقة ذات دلالة إحصائية بين عدد الدراسات ذات الخطأ الواحد أو أكثر والعام.

2- توجد علاقة ذات دلالة إحصائية بين العام وظهور الخطأ، فقبل النداء كانت 72% من الدراسات تحتوي على خطأ واحد أو أكثر، بينما قلت النسبة إلى 51% بعد نشر نداء ثمبسون Thompson.

3- توجد علاقة ذات دلالة إحصائية بين العام والخطأ بالنسبة لبيانات مجلة التقييم التربوي والنفسي.

4- لا توجد علاقة ذات دلالة إحصائية بين العام والخطأ بالنسبة لبيانات مجلة التقييم النفسي ومجلة التقييم التربوي والنفسي.

5- حدث انخفاض في جميع المجلات في عدد الدراسات ذات الخطأ الواحد أو أكثر بعد نشر النداء.

6- كانت مجلة القياس التربوي والنفسي التي نادت بهذه السياسة هي صاحبة أعلى معدل انخفاض في عدد الدراسات التي تحتوي على الأخطاء.

7- من 300 دراسة تم مراجعتها كانت طريقة ألفا كرونباخ هي أكثر الطرق المستخدمة لتقدير الثبات بنسبة 50,7%، ثم طريقة إعادة الاختبار بنسبة 16,7%، ثم ثبات التقدير، التجزئة النصفية، الصور المتكافئة بنسبة 6,3%-3%-1,35% على التوالي؛ وكانت طريقة صدق التكوين الفرضي أكثر الطرق المستخدمة في تقدير الصدق بنسبة 47%، 46%، منها التحليل العاملي الاستكشافي، و29% التحليل العاملي التوكيدي، و26,7% الصدق التمييزي، ثم طريقة الصدق التلازمي بنسبة 26,7%، ثم الصدق التنبؤي، صدق المحتوى بنسبة 26,1%-3,1% على التوالي؛ ويتضح من نتائج الدراسة الثانية حدوث تأثير إيجابي لنداء ثمبسون Thompson على دراسات الصدق والثبات في الأعوام التي تلت النداء.

صرحت العديد من الصحف النفسية المعاصرة بأن التقارير التي تناولت خصائص أدوات القياس في البحوث النفسية كثيرا ما يعيبها الخطأ أو النقصان، لذلك قام ماير ودافيس (Meier and Davis, 1990) ، بدراسة هدفت إلى فحص اتجاهات تدوين تقديرات الصدق والثبات لأدوات القياس في التقارير

التي تتناول خصائص أدوات القياس في بحوث الاستشارات النفسية المنشورة في مجلة مستشارك النفسي في مجلدات الأعوام 1967، 1977، 1987 والمشتملة على أدوات قياس، سواء كانت تلك الأدوات من إعداد الباحث صاحب الدراسة أو أعدها باحث آخر، ثم فحص تقارير عام 1987 بشكل أكثر دقة من أجل الحصول على رؤية أكثر تفصيلا عن الإجراءات المعاصرة فكانت النتائج كما يلي:

1- إن نسبة استخدام المقاييس التي أعدها باحثين آخرين كانت مرتفعة، فقد كانت بنسبة 97,8% في مجلد عام 1967 و75% في مجلد عام 1977 و81% في مجلد عام 1987، نسبة قليلة منها قام الباحثين بالتحقق من تقديرات الصدق والثبات لها قبل استخدامها، حيث كانت بنسبة 8% داخل مجلد عام 1967 و1987، وبنسبة 6% لمجلد عام 1977.

2- لم يدون الباحثين بمعدل متناقص مصدر معلوماتهم عن تقديرات الصدق والثبات الواردة في أبحاثهم، فكانت بنسبة 40% للمقاييس في مجلد عام 1967 و1% في مجلد عام 1977 و3% في عام 1987.

3- لم يدون الباحثين بمعدل متناقص تقديرات الثبات للمقاييس التي أعدوها بنسبة 93% للمقاييس في مجلد عام 1967 و97% في مجلد عام 1977 و77% في مجلد عام 1987.

4- لم يدون الباحثين بمعدل متزايد تقديرات الصدق بنسبة 98% للمقاييس التي أعدوها في مجلد عام 1967 و95% في مجلد عام 1977 و99% في مجلد عام 1987.

5- لم يدون بمعدل متناقض تقديرات الثبات للمقاييس التي أعدها باحثين آخرين بنسبة 95% للمقاييس في مجلد عام 1967 و85% في مجلد عام 1977 و60% في مجلد عام 1987.

6- لم يدون بمعدل متناقص تقديرات الصدق للمقاييس التي أعدها باحثين آخرين بنسبة 98% للمقاييس في مجلد 1967 و96% في مجلد عام 1977 و95% في مجلد عام 1987.

7- بالنسبة لمقاييس عام 1987 فقد كانت نسبة تدوين تقديرات الثبات تساوي 40% للمقاييس التي أعدها باحثين آخرين و33% للمقاييس التي أعدها الباحث صاحب الدراسة، وكانت نسبة تدوين تقديرات الصدق 5% للمقاييس التي أعدها باحثين آخرين و1% للمقاييس التي أعدها الباحث صاحب الدراسة، أي أن نسبة تدوين تقديرات الصدق منخفضة جدا مقارنة بنسبة تدوين تقديرات الثبات.

من الدراسات العربية التي اهتمت بتقويم طرق تقدير صدق وثبات المقاييس دراسة غادة بخش (2007)، التي هدفت إلى تقويم طرق تقدير ثبات وصدق أدوات القياس المستخدمة في بعض رسائل الماجستير بكلية التربية بجامعة أم القرى بالمملكة العربية السعودية، وقد أسفرت الدراسة عن النتائج التالية:

1- تم التحقق من ثبات أدوات القياس بنسبة 73,09%، كانت 81,08% منها من إعداد الباحث و63,29% منها من إعداد باحث آخر عربي و 58,21% منها من إعداد باحث آخر أجنبي؛ وتم التحقق من صدق أدوات القياس بنسبة 72.28%، كانت 84,68% منها من إعداد الباحث و 51,9% منها من إعداد باحث آخر عربي و 55,05% منها من إعداد باحث آخر أجنبي.

2- إن أكثر طرق التحقق من ثبات أدوات القياس شيوعا هي طريقة معامل ألفا كرونباخ بنسبة 52,8%، وأكثر طرق التحقق من الصدق شيوعا هي طريقة صدق المحكمين بنسبة 54,51%.

3- أن نسبة 87% من استخدامات طرق التحقق من ثبات أدوات القياس كانت جيدة، ونسبة 1,1% لم يحدد الباحث نوع الطريقة المستخدمة حتى يمكن الحكم بجودتها، وأن نسبة 46,24% من استخدامات طرق التحقق من الصدق كانت جيدة.

4- وقوع الباحثين في الأخطاء المتعلقة بالعوامل التي تؤثر على ثبات الأداة بشكل عام وبنسب متفاوتة، وهي: خطأ عدم التحقق من ثبات الأداة على عينة ممثلة لمجتمع الدراسة وقد ظهر بنسبة 2,97%، وخطأ عدم التحقق من ثبات الأداة على عينة مستقلة عن عينة الدراسة الأساسية بنسبة 55,02%، وخطأ عدم

تجانس العينة بنسبة 22,27%، وخطأ عدم التحقق من ثبات الأداة على الجماعات الفرعية بنسبة 80,72%، وخطأ عدم مناسبة طول الأداة بنسبة 13,01%، وخطأ عدم مناسبة حجم عينة الثبات بنسبة 72,49%، وخطأ عدم مناسبة قيمة معامل الثبات بـ 8,55% من المجموع الكلي للاستخدامات.

5- لا توجد علاقة بين متغير جنس الباحث ومتغير جودة استخدام طرق التحقق من الثبات، بينما توجد علاقة بين متغير قسم الباحث ومتغير جودة استخدام طرق التحقق من الثبات.

6- وقوع الباحثين في الأخطاء المتعلقة بالعوامل التي تؤثر على صدق الأداة بشكل عام وبنسب متفاوتة، وهي خطأ عدم التحقق من صدق الأداة على عينة ممثلة لمجتمع الدراسة وقد ظهر بنسبة 34,9%، وخطأ عدم التحقق من صدق الأداة على عينة مستقلة من عينة الدراسة الأساسية بنسبة 50,9%، وخطأ عدم التحقق من صدق الأداة على الجماعات الفرعية بنسبة 30,2%، وخطأ عدم كفاية طرق التحقق من صدق الأداة بنسبة 56%، وخطأ عدم سلامة إجراءات طرق التحقق من صدق الأداة بنسبة 49,6%، وخطأ عدم كفاية عدد المحكمين عند التحقق من صدق أدوات القياس قد ظهر بنسبة 28,5% من المجموع الكلي للاستخدامات.

7- لا توجد علاقة بين متغير جنس الباحث ومتغير جودة استخدام طرق التحقق من الصدق، بينما توجد علاقة بين متغير قسم الباحث ومتغير جودة استخدام طرق التحقق من الصدق.

من الدراسات التي سعت إلى تحقيق الهدف السابق دراسة نورة القضيب (2001)، التي هدفت إلى تقويم طرق التحقق من ثبات وصدق إحدى أدوات القياس وهي الاختبارات التربوية والنفسية وذلك في عينة من رسائل ماجستير لقسم علم النفس في بعض جامعات وكليات التربية بالمملكة العربية السعودية؛ وقد أوضحت نتائج الدراسة مايلي:

1- توجد مجموعة من الاختبارات لم يتم التحقق من ثباتها بنسبة 19,6%.

2- الطرق المستخدمة لتقدير الثبات هي: طريقة إعادة الاختبار 31,9%، طريقة التجزئة النصفية 29,1%، طريقة معامل ألفا كرونباخ 27,1%، طريقة كودر-ريتشاردسون(20) 3,2%، ثبات التقدير 1,9%، طريقة الصور المتكافئة 0,5%.

3- توجد بعض الممارسات الخاطئة عند التحقق من ثبات الاختبارات تمثلت في استخدام طرق ثبات غير مناسبة لطبيعة الاختبار، أو ارتضاء معاملات ثبات منخفضة للاختبار، أو عدم الدقة في اختيار وتحديد عينة الثبات.

4- توجد مجموعة من الاختبارات لم يتم التحقق من صدقها بنسبة 24,6%.

5- الطرق المستخدمة لتقدير الصدق هي: الاتساق الداخلي 24,8%، صدق المحكمين 20,3%، الصدق العاملي 9,5%، الصدق المرتبط بالمحك 9,3%، الصدق الذاتي 8,1%، صدق التكوين الفرضي وطريقة المقارنات الطرفية 6.7%، صدق المحتوى 4,5%، طريقة الفروق بين الجماعات 4.3%، الصدق الظاهري 2.4%، طريقة تحليل الأسئلة وطريقة ثبات التجانس 0,9%.

6- توجد بعض الممارسات الخاطئة في التحقق من صدق الاختبار، تمثلت في استخدام طرق صدق غير سليمة علميا (الصدق الظاهري، الصدق الذاتي، المقارنات الطرفية الخاطئة)، أو غير كافية، أو غير مناسبة لطبيعة الاختبار والغرض من استخدامه، أو عدم الدقة في اختيار وتحديد عينة الصدق.

7- توجد فروق دالة إحصائيا بالنسبة لمتغير الكلية في كل من جانبي الثبات والصدق.

8- لا توجد فروق دالة إحصائيا بالنسبة لمتغير نوع الاختبار في كل من جانبي الثبات والصدق.

9- توجد فروق دالة إحصائيا لجانب الثبات، بينما لا توجد فروق دالة إحصائيا لجانب الصدق بالنسبة لمتغير الفترة الزمنية.

من الدراسات العربية التي اهتمت أيضا بتقويم طرق تقدير صدق وثبات المقاييس دراستي الثبيتي 1998 (أ) و (ب). فقد قام الثبيتي 1998 (أ) بالدراسة الأولى تحت عنوان ثبات أدوات القياس في العلوم الإنسانية بين النظرية التقليدية للاختبار ونظرية التعميم، هدفت إلى وصف وتقويم طرق تقدير ثبات البيانات في عينة اشتملت على خمسين دراسة منشورة اختيرت من عشرة (10) مجلات عربية محكمة وتحديدا من عام 1995 إلى عام 1996 واشتملت على أداة قياس واحدة على الاقل.

وقد أوضحت نتائج الدراسة الأولى ما يلي:

1- توجد مجموعة من المقاييس تم التحقق من ثباتها بنسبة 71%.

2- أكثر الطرق شيوعا لتقدير الثبات هي طريقة معامل ألفا كرونباخ بنسبة 30%، ثم طريقة التجزئة النصفية بنسبة 22%، ثم طريقة إعادة الإختبار بنسبة 18%، ولم تستخدم طريقة الصورة المتكافئة.

3- توجد مجموعة من المقاييس لم يتم التحقق من ثباتها بنسبة 29%، حيث أن 55% منها كانت معدة باللغة العربية وسبق استخدامها في دراسات سابقة، و24% منها مترجمة سبق إعدادها واستخدامها في بيئات غير عربية، و26% منها من إعداد صاحب الدراسة وتستخدم لأول مرة في عملية جمع البيانات.

أما الدراسة الثانية 1998 (ب) فكانت بعنوان "صدق البناء النظري لأدوات جمع المعلومات في البحوث التربوية وانعكاس ذلك على تفسير النتائج"، وهدفت إلى وصف وتقويم طرق تقدير صدق البيانات في عينة اشتملت على خمسين دراسة منشورة اختيرت من عشرة دوريات محكمة من عام 1995 إلى عام 1996 واشتملت على أداة قياس واحدة على الأقل، وأوضحت نتائج الدراسة الثانية ما يلي:

1- توجد مجموعة من المقاييس تم التحقق من صدقها بنسبة 68%، ونسبة قليلة جدا منها بمختلف فئاتها (مقاييس مترجمة، مقاييس سبق استخدامها في دراسات سابقة، مقاييس من إعداد صاحب الدراسة وتستخدم لأول مرة) قد تم تقدير الصدق لها بالطرق الرئيسية الثلاثة (صدق المحتوى، الصدق المرتبط

بمحك، صدق التكوين الفرضي) حيث تمثل النسب التالية على التوالي 02%-14%-01% أي بنسبة 17% من مجموع المقاييس، و51% منها قد تم حساب معامل صدقها باستخدام طريقة لا تجيب عن السؤال الذي يطرحه صدق الأداة بأي حال من الأحوال وتتركز في طريقتين: الأولى عبارة عن استخدام طريقة الاتساق الداخلي، والثانية عبارة عن استخدام الجذر التربيعي لمعامل الثبات (الصدق الذاتي).

2- توجد مجموعة من المقاييس لم يتم التحقق من صدقها، حيث أن 14% منها كانت معدة باللغة العربية وسبق استخدامها في دراسات سابقة، و19% منها مقاييس مترجمة وسبق إعدادها واستخدامها في بيئات غير عربية، و67% منها مقاييس تم إعدادها من قبل صاحب الدراسة وتستخدم لأول مرة.

وقام حجر (2003) بدراسة نظرية حاول فيها مناقشة المعايير التي يتم على أساسها تحقيق شروط الموضوعية والصدق والثبات في البحث الكيفي. حيث أوضح الباحث أن الخصائص المميزة للظواهر الاجتماعية والنوعية الفريدة للعلاقة التي تربط ما بين الباحث وموضوع دراسته تتطلب تصورا بديلا للشروط التي يضعها منهج العلوم الطبيعية لتحقيق الصرامة وتوفير الشروط العلمية والصدق والثبات؛ وأن محاولة استجلاب المعايير الكمية للصدق والثبات ومحاولة تطبيقها على البحوث القائمة على المنهج الكيفي هي عملية غير مبررة عمليا، كما أنها لا تلائم طبيعة المعرفة التي يوفرها هذا المنهج.

فالباحث الكيفي يحاول تعميق فهمه للظاهرة الاجتماعية من خلال التحليل المتعمق والحساس لطبيعة وعي المتفاعلين في إطار تلك الظاهرة، وهو يقوم بقراءة الوثائق والمذكرات الميدانية بحثا عن إيحاءات ووسائل جديدة لفهم الموقف، وهي مهمة لا يصلح معها أساليب منهج العلوم الطبيعية. كما بين أنه لتحقيق الصدق يجب التأكد من عدم إنحياز الباحث، ويستخدم لذلك استراتيجية التأمل وهي بمثابة البحث الذاتي عن مواطن انحيازه المحتمل ومحاولة السيطرة عليها وتحديد الوسائل التي تمكنه من ذلك، وكذلك استراتيجية تعيين الحالات السالبة التي تستخدم لتقليل تأثير الميول الشخصية على البحث،

وتتضمن البحث المقصود عن الحالات التي لا تتفق مع تطلعاتهم وتفسيراتهم لما يقومون بدراسته، مما يؤدي إلى الوصول إلى نتائج أفضل والمساهمة في تدعيم مستوى صدق تلك النتائج.

أما الثبات فيمكن تحقيقه في البحث الاجتماعي الكمي عند الاعتماد على نفس الإطار النظري واستخدام نفس المنهج في جمع البيانات وتحليلها وتوفير أوضاعا متشابهة تمكن الباحثين من التوصل إلى نتائج متشابهة، مع إدراكهم لأسباب الاختلاف الذي قد يحدث كأمر طبيعي ناتج عن سرعة التغير في الظواهر الاجتماعية والمتغيرات المعقدة والمتجددة التي تؤثر فيها. حيث أنه من غير الممكن أن تتطابق نتائج دراسة ظاهرة إجتماعية مع نتائج دراسات سابقة بنفس الصورة التي تتحقق بها في البحث الكمي.

من الدراسات التي تناولت إحدى طرق الصدق دراسة قام بها القاطعي (1993) هدفت إلى معرفة علاقة المحكمين بالمفهوم الإحصائي لصدق الأسئلة، حيث طلب من عدد من المحكمين إبداء الرأي في مدى ملائمة ووضوح عدد الأسئلة الجيدة وغير الجيدة بناء على بعض المعايير الإحصائية والتي لم يكن لدى المحكمين معرفة مسبقة بها؛ وقد أظهرت نتائج الدراسة أن التحكيم قد يكون مفيدا في جانب الوضوح، أما في جانب الملائمة فقد كان دوره هامشيا.

حول تقويم صدق الإختبارات أجرت شيبارد (Shepard,1993) دراسة أكدت فيها على التطور السريع الذي مرت به نظرية الصدق خاصة منذ 1980، وقد أكدت الباحثة على أن هناك إجماعا بين علماء القياس على عدم جدوى استخدام الطرق التقليدية في عملية تقدير الصدق، وذلك لأن كلا من صدق المحتوى والصدق المرتبط بمحك لا يمكن الإعتماد على أي منهما بمعزل عن صدق التكوين الفرضي، وبررت أهمية هذه النظرية الحديثة في قدرتها على الربط بين جميع المفاهيم الظاهرة والعلاقات بين تلك المفاهيم وبين التحقق التجريبي من ذلك، وفي الدراسة دعوة إلى تطبيق النظرية العلمية لتوحيد مفهوم صدق الإختبار والذي يؤكد أهمية تفسير الدرجات واستعمالاتها في الميادين التطبيقية في ظل إطار مفاهيمي نظري

يصور العلاقة بين محتوى الأداة وما يمثله من استجابات. يتم ذلك عن طريق التحقق من البناء الداخلي والخارجي للأداة، حيث يتعلق البناء الداخلي بالنظرية المستخدمة وطريقة بناء الإختبار وارتباطه بالعمليات العقلية المقاسة، كما يرتبط البناء الخارجي للأداة بعلاقة الإختبار باختبارات أخرى، وقد اتفقت موس (Moss,1995) معها في ذلك ودعت إلى توسيع مجال مفهوم الصدق ليشمل الاهتمام بتبعات التقييم، ودعت أيضا إلى إجراء تنقيح لمعايير الصدق مرة كل عقد من الزمن لكي نعيد الاهتمام بالأساسيات الإرشادية، و القيام بالتغيير الذي سيكون له أثر إيجابي على دراسات الصدق.

وحول ثورة التغير في مجال القياس وخاصة فيما يتعلق بصدق الاختبارات فإن ميسك Messick يعتبر من الرواد في هذا المجال، حيث يعتقد بالنظرة الشمولية للصدق وإن خير ما يمثل هذه الوحدة هو صدق التكوين الفرضي، ففي دراسة أجراها (Messick,1994) حول صدق التقويم النفسي، أشار إلى أن مصادر عدم مصداقية الاختبار تتمثل في عدم تمثيل أسئلة الاختبار لمحتوى السمة من حيث التغطية أو من حيث العمليات العقلية التي توجه السلوك، وعدم وجود صلة بين التباين والسمة المراد قياسها، كأن يشمل الاختبار سمات أخرى لا علاقة لها بالسمة المقصود قياسها. كما أكد في دراسة أخرى (Messick,1995) بعنوان "معايير الصدق وصدق المعايير في تقويم الاختبار" أن الصدق ليس مجرد مبدأ من مبادئ القياس ولكنه يمثل قيمة اجتماعية ذات معنى وقوة تقويمية في اتخاذ القرارات، كما استعرض الباحث المعايير الستة التي يجب أن يقوم عليها صدق التكوين الفرضي وهي: صدق التكوين الفرضي للمحتوى، الجانب الأساسي للصدق، صدق الجانب البنائي للاختبار، صدق الجانب التعميمي، صدق الجانب الخارجي للبناء النظري، وصدق الجانب النظري للعواقب.

التعليق على الدراسات السابقة:

من خلال العرض السابق نلاحظ ما يلي:

1- ركزت الدراسات على تقويم الاختبارات أكثر من باقي أدوات القياس بالرغم من أهمية باقي الأدوات في جمع البيانات في الدراسات الإنسانية؛ لذلك ستتناول الدراسة الحالية تقويم أدوات القياس المستخدمة في بعض أطروحات الدكتوراه من حيث خاصيتي الصدق والثبات.

2- أن الدراسات التي اهتمت بتقويم الاختبارات، الغالبية العظمى منها ركزت على تقويم الاختبارات التحصيلية بناء على مدى التزامها ببعض المعايير (تصنيف بلوم للمجال المعرفي – نوعية الأسئلة – عمليات التعلم – الموضوعية – النواحي الفنية من حيث الصياغة والإخراج)، ما عدا دراسة وفاء باسروان (1995) التي اهتمت بتوفر شرط الصدق في الاختبارات، ودراسة نورة القضيب (2001) التي قومت طرق حساب الصدق والثبات في الاختبارات التربوية والنفسية فقط، ودراسة غادة بخش (2007) التي تناولت تقويم طرق تقدير صدق وثبات أدوات القياس؛ لذلك ستتناول الدراسة الحالية تقييم طرق تقدير صدق وثبات أدوات جمع المعطيات في أطروحات الدكتوراه.

3- تعدد القضايا التي تناولتها الدارسات حول الصدق والثبات، فهي إما حول المفهوم النظري للصدق والثبات، أو المشاكل التعبيرية أو المنهجية التي يقع فيها الباحثين عند إستخدامهما في البحوث، أو التعرف على مدى اهتمام الباحثين بتدوين أو تقدير القيم التقديرية للصدق أو الثبات، أو تقويم طرق التحقق من صدق وثبات الاختبارات التربوية والنفسية. أما في هذه الدراسة الحالية سوف يقوم الباحث بتقييم طرق تقدير صدق وثبات أدوات جمع المعطيات بناء على معايير حددها الباحث، والتعرف على الأخطاء التي يقع فيها الباحثين عند التحقق من الصدق والثبات.

4- يلاحظ أن أغلب الدراسات التي تمت في مجال الصدق ركزت على التطور السريع الذي تعرض له مفهوم الصدق خلال السنوات القليلة الماضية، ما عدا دراسة القاطعي (1993) التي هدفت إلى معرفة علاقة المحكمين بالمفهوم الإحصائي لصدق الأسئلة؛ ولذلك الدراسة الحالية سوف تهتم بتقييم طرق تقدير الصدق ومدى مناسبة الطريقة لطبيعة الأداة بناء على معايير الصدق التي حددها الباحث، وذلك وفق النظرية الكلاسيكية (التقليدية) للقياس.

5- بالنسبة للعينة فإن معظم الدراسات استخدمت عينة عبارة عن دراسات منشورة، سواء في دوريات عربية كما في دراستي الثبيتي (1998)، أو دوريات أجنبية كما في دراسة كل من (Ottenbacher, 1995 و (Meir and Davis,1990) و (Witta and Daniel,1997,1998)، ما عدا دراسة (Hamersley,1987) التي كانت عينتها مجموعة من كتب القياس، ودراستي كل من نورة القضيب (2001) التي كانت عينتها رسائل الماجيستير في قسم علم النفس، ودراسة غادة بخش (2007) عينتها عبارة عن رسائل الماجيستير في قسم علم النفس وقسم الإدارة والتخطيط التربوي وقسم المناهج وطرق التدريس؛ أما الدراسة الحالية فسوف تتضمن العينة أطروحات الدكتوراه في قسم علم النفس وعلوم التربية.

6- لم تهتم الدراسات السابقة بدراسة العوامل المؤثرة في قيمة معامل الصدق والثبات ومدى التزام الباحثين بضبطها، ما عدا دراسة نورة القضيب (2001) ودراسة غادة بخش (2007)، فقد تناولتا البعض منها؛ لذلك سوف تسعى الدراسة الحالية إلى دراسة هذا الجانب بالتفصيل.

الفصل الثالث

إجراءات البحث

تمهيد:

يتضمن هذا الفصل إجراءات البحث المنهجية، من حيث اختيار منهج البحث المناسب ووصف عينة البحث، ومن ثم اختيار أداة البحث وطريقة بنائها، وأخيرا الإشارة إلى الأساليب الإحصائية المستخدمة، وفيما يأتي توضيح لهذه الإجراءات:

أولا: منهج البحث:

قام الباحث باتباع المنهج الوصفي التقويمي لمناسبته غرض الدراسة وطبيعة بياناتها، والذي يعتمد على الجداول الإحصائية والرسوم البيانية لوصف واقع الأشياء أو الظواهر التي تتناولها الدراسة، فغرض البحث الحالي هو تقييم طرق التحقق من صدق وثبات أدوات جمع المعطيات المستخدمة في أطروحات دكتوراه علم النفس وعلوم التربية بجامعة الجزائر2 وذلك باستخدام معايير محددة.

ثانيا: عينة البحث:

تتمثل عينة الدراسة في أدوات جمع المعطيات المستخدمة في أطروحات دكتوراه علم النفس وعلوم التربية بجامعة الجزائر2، حيث تقتصر هذه الدراسة على الأطروحات المتاحة للباحث بصورة كاملة على الموقع الإلكتروني للبوابة الوطنية للإشعار عن الأطروحات (*www.pnst.cerist.dz*)، والتي استخدمت أداة واحدة على الأقل.

حيث يتوفر في الموقع الإلكتروني للبوابة الاطروحات المناقشة في جامعة الجزائر2 فقط، باعتبار أن الجامعات الآخرى لم تقم برقمنة الأطروحات ورفعها على الموقع؛ كما أننا لم نتمكن من الحصول على نسخ ورقية في جامعات أخرى (صعوبات إدارية).

تم تحميل الأطروحات يوم الإثنين 09 جوان 2014: من الساعة 16 حتى الساعة 20 مساء، وكانت كالآتي:

جدول رقم (01) يوضح عدد الأطروحات قبل الإستبعاد

التخصص	الأطروحات المتوفرة	الأطروحات للتحميل
علم النفس	81	68
علم النفس الاجتماعي	18	16
علم النفس العمل والمنظمات	15	14
علم النفس وعلوم التربية	48	42
علوم التربية	25	25
أرطفونيا	11	09
المجموع	198	174

يوضح الجدول السابق المجتمع الأصلي للأطروحات (حسب التسمية الأصلية في الموقع)، حيث تتوفر في الموقع 198 أطروحة، من بينها 174 قابلة للتحميل (أخطاء تقنية في الموقع).

بعد تحميل هذه الأطروحات قام الباحث بتصنيفها حسب معياري: القسم والتخصص، وذلك لاختلاف برامج التكوين فيها خاصة في مقاييس المنهجية والإحصاء والقياس النفسي؛ حيث استثنينا فقط أطروحات قسم علم النفس (تخصص: إجتماعي، عمل وتنظيم، عيادي) وعلوم التربية، واستبعدنا أطروحات قسم الأرطفونيا لقلة عددها وطبيعة أدواتها المستخدمة (البروتوكل، والأجهزة الإلكترونية مثل IRM...) باعتبارها لا تقيس سمات نفسية ولم نجد ذكر لمصطلح صدق أو ثبات فيها؛ واستبعدنا تخصصات أخرى هي: الارشاد النفسي، علم النفس التربوي، علم النفس العصبي، علم النفس اللغوي والمعرفي لقلة عدد الأطروحات فيها والذي لا يتجاوز 03 في كل تخصص، مما يؤثر على تجانس العينة وبالتالي على نتائج الدراسة.

والجدولين التاليين يوضحان عينة الأطروحات حسب متغيري القسم والتخصص، وذلك بعد تعديل تسمية التخصصات، حيث تم اعتماد التسميات الموجودة على الصفحة الواجهة للأطروحة، وليس التسمية الأصلية الموجودة في الموقع.

جدول رقم (02) يوضح توزيع الأطروحات حسب متغير التخصص

التخصص	عدد الأطروحات	النسبة
علم النفس الاجتماعي	19	% 17
علم النفس العمل والمنظمات	21	% 18
علم النفس العيادي	30	% 27
علوم التربية	42	% 38
المجموع	112	% 100

يمثل الجدول رقم (02) عدد الأطروحات ونسبتها حسب متغير التخصص، حيث نلاحظ أن عدد الأطروحات التي ستتم عليها الدراسة هو 112 أطروحة، وأن أكبر نسبة كانت لـ علوم التربية بـ 38 % أي 42 أطروحة، ثم تليها التخصصات الأخرى كالتالي: علم النفس العيادي، علم النفس العمل والتنظيم، علم النفس العمل الإجتماعي (30، 21، 19 على الترتيب)؛ مما يبين لنا مدى تقارب نسب وأعداد الأطروحات في التخصصات.

جدول رقم (03) يوضح توزيع الأطروحات حسب متغير القسم

التخصص	عدد الأطروحات	النسبة
علم النفس	70	% 62
علوم التربية	42	% 38
المجموع	112	% 100

الجدول رقم (03) يمثل توزيع أطروحات الدراسة حسب متغير القسم، فأكبر نسبة هي لقسم علم النفس (62 %) بـ 70 أطروحة، ثم النسبة الأخرى 38% لقسم علوم التربية بـ 42 أطروحة، وبالتالي ليس هناك فرق كبير في النسبة.

ويمكن تفسير هذه الإختلافات في النسب بأن عملية رقمنة الأطروحات الجامعية غير كاملة، وقد تكون أيضا بسبب اختلاف عدد طلبة الدكتوراه، حيث أن أغلب مشاريع الدكتوراه المقبولة في الجامعات هي من قسم علم النفس ولاسيما تخصص العيادي.

جدول رقم (04) يوضح توزيع الأطروحات حسب سنة التخرج

سنة التخرج	التكرار	النسبة
2006	10	% 08.9
2007	16	% 14.3
2008	23	% 20.5
2009	14	% 12.5
2010	07	% 06.3
2011	14	% 12.5
2012	15	% 13.4
2013	13	% 11.6
المجموع	112	% 100

الجدول السابق يوضح عدد ونسبة الأطروحات حسب سنة المناقشة (التخرج)، والتي تتراوح بين 06.3 % (سنة 2010) و 20.5 % (سنة 2008).

حيث يمكن أن للمتغير الزمني علاقة مع جودة استخدام طرق التحقق من صدق وثبات أدوات القياس، ففي سنة 2006 لم تكن تتوفر برامج الحاسوب الآلي كالـ SPSS مثلما هو في سنة 2013. كما أن برامج التكوين في النظام الجديد LMD تختلف عن برامج النظام الكلاسيكي، خاصة فيما يتعلق بمقاييس منهجية البحث والإحصاء والقياس النفسي.

جدول رقم (05) يوضح توزيع الأدوات حسب متغير التخصص

التخصص / الأداة	الملاحظة	المقابلة	الإختبار	المجموع	
				التكرار	النسبة%
علم النفس الاجتماعي	3	3	49	55	15.9
علم النفس العمل والمنظمات	2	5	66	73	21.1
علم النفس العيادي	6	16	98	120	34.7
علوم التربية	10	3	85	98	28.3
المجموع	21	27	298	346	100.0

يمثل الجدول رقم (05) عدد الأدوات المستخدمة في أطروحات الدكتوراه حسب متغير التخصص، حيث نلاحظ أن أكبر عدد مستخدم من الأدوات هو في تخصص علم النفس العيادي، خاصة فيما يخص أداة المقابلة وأداة الإختبار. وبقراءة الجدول رقم (05) في ضوء الجدول رقم (02) نلاحظ أن نسبة

الأدوات المستخدمة في تخصص علم النفس العيادي أكبر منه في علوم التربية (34.7% مقابل 28.3%) بالرغم من أن عدد الأطروحات في علوم التربية أكبر منه في علم النفس العيادي (38% مقابل 27%)، وقد يرجع هذا لتعقد الظاهرة النفسية في الدراسة العيادية وتداخل متغيراتها وكثرة عددها، وبالتالي حاجة المختص لتوفر أدوات كافية كما ونوعا بهدف تشخيص الإضطرابات ومن ثم علاجها.

جدول رقم (06) يوضح توزيع الأدوات حسب متغير القسم

القسم / الأداة	الملاحظة	المقابلة	الإختبار	المجموع	
				التكرار	النسبة%
علم النفس	11	24	213	248	71.7
علوم التربية	10	3	85	98	28.3
المجموع	21	27	298	346	100.0

يمثل الجدول السابق عدد الأدوات المستخدمة في أطروحات الدكتوراه حسب متغير القسم، حيث نلاحظ أن العدد الإجمالي للأدوات يقدر ب 346 أداة –وهي تمثل العينة الفعلية للدراسة الحالية–، وأن ما يفوق نسبة الثلثين من هذه الأدوات استخدمت في أطروحات علم النفس (71.7%). بينما نجد أن النسبة المستخدمة في أطروحات علوم التربية لا تتجاوز الثلث (28.3%). وبالنظر إلى الجدول رقم (03) نجد أن هذه النتائج منطقية باعتبار عدد الأطروحات في قسم علم النفس أكبر منه في قسم علوم التربية (62% مقابل 38%).

كما نلاحظ أن أكثر أداة مستخدمة بغض النظر عن القسم هي أداة الإختبار، وذلك لسهولة إعدادها وتنفيذها، ثم تليها أداتي المقابلة والملاحظة، واللتين استخدمتا بكثرة في أطروحات علم النفس العيادي لطبيعة البحث (بحث كيفي) وطبيعة العينة (حالات) في هذا التخصص. أما بالنظر إلى متغير القسم، فنجد أن نسبتي استخدام أداة الملاحظة متساويتين تقريبا، بينما أداتي المقابة والملاحظة فاستخدمتا في أطروحات علم النفس بصفة أكبر، خاصة في تخصص علم النفس العيادي.

ثالثا: أداة البحث:

لتحقيق أهداف هذه الدراسة، قام الباحث ببناء شبكة لتقييم طرق تقدير الصدق والثبات المستخدمة في أطروحات الدكتوراه، مستفيدا من مراجعته لأدبيات القياس النفسي، خاصة فيما يتعلق بالخصائص السيكومترية لأدوات القياس، وكذلك من خلال إطلاعه على الدراسات السابقة التي تناولت تقويم الأساليب المنهجية للبحوث الإجتماعية عموما وأدوات جمع المعطيات خصوصا، كما سبق ذكرها في الفصل الخاص بالإطار النظري للبحث.

وتتكون أداة الدراسة من 34 فقرة موزعة على ثلاثة أقسام:

القسم الأول: يحتوي على البيانات العامة للرسالة، وهي: الرقم التسلسلي للأطروحة، رقم الأداة في الأطروحة، القسم، التخصص، إعداد الأداة، نوع الأداة، والسمة المقاسة (من الفقرة 1-1 إلى 1-7).

القسم الثاني: يضم بيانات متعلقة بالطرق المستخدمة للتحقق من صدق الأداة، وبدورها تحتوي على جزئين:

أ. أسئلة عامة وهي: التحقق من الصدق، الطرق المستخدمة في ذلك، كفاية هذه الطرق، تمثيل عينة الصدق لمجتمع الدراسة، وكفاية حجم العينة (من الفقرة 2-1 إلى 2-5).

ب. أسئلة خاصة بكل نوع من أنواع الصدق الثلاث:

1. صدق المحتوى: تحليل المحتوى بالوصف التفصيلي أو توصيف الأهداف، عمل جدول مواصفات، حساب نسبة الإتفاق، ومناسبة الطريقة للغرض من الأداة (من الفقرة 2-1-1 إلى 2-1-4).

2. الصدق المرتبط بمحك: نوعه تلازمي أم تنبؤي، المحك المستخدم، تعلق المحك بالسمة، صدق المحك، ثبات المحك، مناسبة الطريقة للغرض، ومستوى معامل الصدق (من الفقرة 2-2-1 إلى 2-2-7).

3. صدق التكوين الفرضي: التعريف بالإطار النظري للسمة، اشتقاق الفرضيات، طرق التحقق، مناسبة الطرق للغرض، وكفاية الطرق المستخدمة (من الفقرة 2-3-1 إلى 2-3-5).

القسم الثالث: يتعلق بطرق التحقق من الثبات ويحتوي على عدة بيانات وهي: التحقق من الثبات، الطرق المستخدمة، مناسبة الطرق للغرض من الأداة، تمثيل عينة الثبات لمجتمع الدراسة، كفاية حجم العينة، ومستوى معامل الثبات (من الفقرة 3-1 إلى 3-6). (أنظر الملحق رقم 01).

صدق الأداة:

للتأكد من صدق الأداة تم استخدام طريقة صدق المحكمين حيث قام الباحث بعرض شبكة التقييم على 10 محكمين متخصصين في مجال القياس النفسي وتحليل المعطيات (06 أستاذ التعليم العالي، 04 أستاذ محاضر-أ-؛ بكل من جامعة: مستغانم، وهران، البليدة، الجزائر، باتنة، سطيف)، كما عرضت أيضا على أعضاء مخبر تحليل المعطيات بمستغانم في شكل مناقشة جماعية (مدير المخبر و10 طلبة دكتوراه تخصص القياس النفسي)، وقد تم تعديل بعض بنود الشبكة ومعايير التحليل في ضوء آراء المحكمين، ويتعلق الأمر بالبنود 2-2، 2-3-3، 2-3-4، 3-2، 3-3، 3-6. (أنظر الملحق رقم 01).

جدول رقم (07) يوضح قائمة المحكمين

رقم المحكم	الدرجة العلمية	الجامعة الأصلية
01	أستاذ التعليم العالي	مستغانم
02	أستاذ التعليم العالي	مستغانم
03	أستاذ التعليم العالي	وهران
04	أستاذ التعليم العالي	البليدة
05	أستاذ التعليم العالي	باتنة
06	أستاذ التعليم العالي	سطيف
07	أستاذ محاضر-أ-	الجزائر
08	أستاذ محاضر-أ-	مستغانم
09	أستاذ محاضر-أ-	مستغانم
10	أستاذ محاضر-أ-	وهران

كما استخدم الباحث طريقة صدق المحتوى لمعرفة مدى تمثيل بنود الأداة للأهداف ولجوانب المجال المقاس، وللتأكد من مرونة الشبكة في التطبيق واستيفائها لكل جوانب التقييم، حيث قام بإجراء تطبيقي للأداة عن طريق فحص 12 أطروحة من مختلف الأقسام والتخصصات ضمت 30 أداة قياس.

أولا: وصف الأهداف:

تهدف هذه الشبكة إلى:

1- تقويم طرق التحقق من صدق أدوات جمع المعطيات المستخدمة في أطروحات الدكتوراه، ويتفرع من هذا الهدف عدة أهداف فرعية هي:

1. أن يتحقق الباحث من صدق الأداة.
2. أن يحدد البحث الطرق المستخدمة في التحقق من صدق الأداة.
3. أن يستخدم الباحث طرق كافية في التحقق من صدق الأداة.
4. أن يراعي الباحث مواصفات عينة الصدق.
5. أن يراعي الباحث شروط التحقق من صدق المحتوى.
6. أن يراعي الباحث شروط التحقق من الصدق المرتبط بمحك.
7. أن يراعي الباحث شروط التحقق من صدق التكوين الفرضي.

2- تقويم طرق التحقق من ثبات أدوات جمع المعطيات المستخدمة في أطروحات الدكتوراه، ويتفرع من هذا الهدف عدة أهداف فرعية هي:

1. أن يتحقق الباحث من ثبات الأداة.
2. أن يحدد البحث الطرق المستخدمة في التحقق من ثبات الأداة.
3. أن يستخدم الباحث طرق مناسبة في التحقق من ثبات الأداة.

4. أن يراعي الباحث مواصفات عينة الثبات.

5. أن يعتمد الباحث على معامل ثبات مقبول.

حيث لاحظنا في جانب الصدق أن كل هدف من الأهداف الثلاثة الأولى اشتمل على بند واحد، أما الهدف الرابع فقد اشتمل على بندين، وكذلك الهدف الخامس اشتمل على أربعة بنود، والهدف السادس فقد اشتمل على سبعة بنود، وأخيرا الهدف السابع اشتمل على خمسة بنود.

أما جانب الثبات فيلاحظ أن كل هدف اشتمل على بند واحد، عدا الهدف الرابع فقد اشتمل على بندين.

ثانيا: تحديد المجال المقاس:

يتكون المجال المقاس من الصدق والثبات، فيما يلي تحديد لكل منهما:

1- الصدق: ويتكون من النقاط التالية:

1. حالات التحقق من صدق الأداة.
2. الطرق المختلفة في التحقق من صدق الأداة.
3. الطرق الكافية في التحقق من صدق الأداة.
4. مناسبة طريقة التحقق من الصدق مع الهدف من استخدام الأداة.
5. عينة الصدق.
6. جانب التحقق من صدق المحتوى.
7. جانب التحقق من الصدق المرتبط بمحك.
8. جانب التحقق من صدق التكوين الفرضي.

2- الثبات: ويتكون هذا الجانب من عدة نقاط هي:

1. حالات التحقق من ثبات الأداة.
2. الطرق المختلفة في التحقق من ثبات الأداة.
3. الإفتراضات الأساسية لطرق التحقق من الثبات.
4. مناسبة طريقة التحقق من الثبات مع الهدف من استخدام الأداة.
5. عينة ثبات الأداة.
6. معامل ثبات الأداة.

أما بالنسبة لجوانب المجال المقاس فيلاحظ في جانب الصدق أن كل جانب من الجوانب الأربعة الأولى اشتمل على بند واحد، أما الهدف الخامس فقد اشتمل على بندين، وكذلك الجانب السادس اشتمل على أربعة بنود، والجانب السابع فقد اشتمل على سبعة بنود، وأخيرا الجانب الثامن اشتمل على خمسة بنود.

أما جانب الثبات فيلاحظ أن كل جانب اشتمل على بند واحد، عدا الجانب الخامس فقد اشتمل على بندين.

وبناء على هذا تم صياغة الشبكة في صورتها النهائية، حيث اشتمل جانب الصدق على 21 بندا (05 بنود عامة، صدق المحتوى 04 بنود، صدق المرتبط بمحك 07 بنود، صدق التكوين الفرضي 05 بنود)، أما جانب الثبات فاشتمل على 06 بنود (أنظر الملحق رقم 01).

ثبات الأداة:

قام الباحث بتقدير ثبات شبكة التقييم عن طريق حساب معامل الإتفاق بين الباحث ونفسه (إعادة التطبيق)، وبين الباحث وباحث آخر (متخصص في القياس النفسي) بفاصل زمني قدره أسبوعين، باستخدام معادلة هولستي Holsti (1969):

$$\text{معامل الاتفاق} = \frac{2 \text{ (س ص)}}{\text{س1} + \text{ص1}}$$

حيث أن:

س1: عدد الفقرات التي نتجت عن التطبيق الأول.

س2: عدد الفقرات التي نتجت عن التطبيق الثاني.

س ص: عدد الفقرات التي تطابقت في التطبيقين.

حيث تكونت عينة الثبات من 36 أطروحة اشتملت على 98 أداة موزعة كالتالي: من حيث القسم (18 علم النفس، 18 علوم تربية)، من حيث التخصص (6 علم النفس الاجتماعي، 6 عمل وتنظيم، 6 عيادي)، من حيث نوع الأداة (اختبار 92، دليل مقابلة 5، شبكة ملاحظة 1).

ولقد بلغ معامل الاتفاق العام (الوسيط) بين الباحث ونفسه 0,98، أما بين الباحث وزميله فقد بلغ 0,90، مما يدل على ثبات عالي لشبكة التقييم الخاصة بالدراسة الحالية.

رابعا: الأساليب الإحصائية:

للإجابة على تساؤلات البحث تم استخدام مجموعة من الأساليب الإحصائية الوصفية، حيث إعتمدنا على التكرارات والنسب المئوية.

الفصل الرابع

نتائج البحث

تمهيد:

يتضمن هذا الفصل عرض ومناقشة النتائج المترتبة على الإجابة عن كل سؤال من أسئلة البحث، والتي سيتم عرضها في جانبين، جانب الصدق وجانب الثبات:

أولا: عرض ومناقشة النتائج المتعلقة بالصدق:

السؤال الأول:

هل تم التحقق من صدق أدوات جمع المعطيات في أطروحات دكتوراه علم النفس وعلوم التربية بجامعة الجزائر2؟

للإجابة على هذا السؤال قام الباحث بإحصاء الأدوات التي تم التحقق من صدقها في أطروحات الدكتوراه، سواء كانت من إعداد الباحث صاحب الدراسة أو باحث آخر، والجدول الآتي يوضح ذلك:

جدول رقم (08) يوضح تكرارات ونسب أدوات جمع المعطيات التي تم التحقق من صدقها

التحقق من صدق الأداة	إعداد الأداة								المجموع	
	الباحث نفسه		باحث عربي		باحث أجنبي		لم يحدد			
	التكرار	ن %	التكرار	ن %	التكرار	ن %	التكرار	ن %	التكرار	ن %
تم التحقق منه	113	45.38	29	11.65	93	37.35	14	05.62	249	72.00
لم يتم التحقق منه	35	36.08	16	16.50	45	46.39	01	01.03	97	28.00
المجموع	148	/	45	13.00	138	/	15	/	346	100.0

من خلال الجدول السابق يتضح أنه لم يتم التحقق من صدق أدوات القياس بنسبة 28.00% من مجموع الأدوات المستخدمة في أطروحات دكتوراه علم النفس وعلوم التربية بجامعة الجزائر2، والبالغ عددها الإجمالي 346 أداة؛ وبالنظر إلى متغير إعداد الأداة نجد أن 46.39% من الأدوات التي لم يتم التحقق من صدقها كانت من إعداد باحث آخر "أجنبي"، ونسبة 36.08% منها كانت من إعداد الباحث نفسه،

أما نسبة الأدوات التي كانت من إعداد باحث آخر "عربي" فقدرت بـ 16.50%، بينما الأدوات التي لم يتم التحقق من صدقها ولم يحدد الباحث الذي أعدها فكانت نسبتها 01.30%.

وفي هذه النتائج دليل على إهمال الباحثين لاعتبارات الدقة في نتائج دراساتهم، فنسبة 28.00% من الأدوات التي لم يتم التحقق من صدقها تعد نسبة ملفتة للنظر (حوالي الثلث)، خاصة عند ملاحظة أن 36.08% منها كانت من إعداد الباحث نفسه، أي أنها أداة جديدة لم تستخدم من قبل، ورغم ذلك أهمل الباحث تقدير صدقها.

وتتفق هذه النتائج مع ما توصلت إليه دراسة نورة القضيب (2001) حيث وجدت أن 24.60% من الاختبارات المستخدمة في عينة دراستها لم يتحقق من صدقها.

ومن جهة أخرى نجد أن نسبة 46.39% من الأدوات التي لم يتم التحقق من صدقها كانت من إعداد باحث آخر "أجنبي"، رغم أنها طبقت بلغة مختلفة وعلى بيئة مختلفة من جميع النواحي، ويمكن تفسير هذه النتيجة بأن ارتفاع هذه النسبة راجع لأن أغلب هذه الأدوات تم استخدامها على بيئة عربية وأظهرت صدق مناسب فاستخدمها الباحث على هذا الأساس وتجاهل اعتبارات مثل اختلاف الثقافة والعينة.

وتتفق هذه النتيجة مع نتيجة دراسة ماير ودافيس (1995) فقد وجدا أن نسبة الأدوات التي كانت من إعداد باحث آخر مرتفعة 72.00% وأن نسبة قليلة منها قام الباحثين بالتحقق من صدقها بنسبة 08.00% للمقالات في عينة دراستهما.

أما النتيجة المتعلقة بالأدوات التي لم يتم التحقق من صدقها وكانت من إعداد باحث آخر "عربي" والتي قدرت بـ 16.50%، فقد اكتفى الباحثين بذكر الخصائص السابقة للأدوات، مبررين ذلك بتشابه خصائص البيئة الجزائرية والعربية؛ ولكن عند الرجوع إلى معايير الاختبارات النفسية والتربوية (APA, 1999) نجد أن هذا الإجراء خاطئ لأنه "ليس هناك اختبار يعتبر صادقا لجميع الأغراض والمواقف، فكل

استخدام أو تفسير مقترح للاختبار يتطلب عملية تحقق من الصدق، وتحديد -بلغة واضحة- المجتمع الذي يلائمه الاختبار والبناء الذي يقيسه الاختبار، والأسلوب والسياق اللذين سيتم توظيف الدرجات فيهما".

وتتفق هذه النتائج مع دراسة الثبيتي (1998-ب) حيث وجد أن 32% من الأدوات لم يتم التحقق من صدقها، حيث أن 14% منها كانت معدة باللغة العربية وسبق استخدامها و 19% منها مترجمة استخدمت في بيئات عربية و67% من إعداد صاحب الدراسة وتستخدم لأول مرة.

كما وجدت غادة بخش (2007) أن 27.72% من أدوات القياس في الرسائل لم يتم التحقق من صدقها، كانت 33.33% منها من إعداد الباحث و 37.26% منها من إعداد باحث آخر "عربي" و29.41% منها من إعداد باحث آخر "أجنبي".

السؤال الثاني:

ما هي الطريقة الأكثر شيوعا في التحقق من صدق أدوات جمع المعطيات في أطروحات دكتوراه علم النفس وعلوم التربية بجامعة الجزائر2؟

للإجابة على هذا السؤال قام الباحث بإحصاء طرق الصدق المستخدمة في عينة البحث، وهي ستة (06) طرق، موضحة في الجدول التالي:

جدول رقم (09) يوضح تكرارات ونسب طرق التحقق من الصدق

طريقة التحقق من الصدق	التكرار	النسبة المئوية
صدق المحتوى	14	03.56
الصدق المرتبط بمحك	11	02.80
صدق التكوين الفرضي	126	32.06
صدق المحكمين	169	43.00
الصدق الذاتي	48	12.21
الصدق الظاهري	24	06.11
طريقة أخرى	01	00.26
المجموع	393	100.0

من خلال الجدول السابق يتضح أن أكثر طرق التحقق من صدق أدوات القياس شيوعا هي طريقة صدق المحكمين بنسبة 43.00% من مجموع استخدامات طرق التحقق من صدق أدوات جمع المعطيات في أطروحات دكتوراه علم النفس وعلوم التربية بجامعة الجزائر2 (393 طريقة)، ثم طريقة صدق التكوين الفرضي بنسبة 32.06% من الاستخدامات، ثم الصدق الذاتي، الصدق الظاهري، صدق المحتوى، الصدق المرتبط المحك بنسبة 12.21%، 06.11%، 03.56%، 02.80% على التوالي. كما استخدمت طريقة أخرى واحدة وهي الصدق الثقافي (استخدم التحكيم) بنسبة 00.26%.

ويمكن تفسير هذه النتائج أنه كلما تقدمت الطريقة وزادت تعقيدا كلما قل استخدامها، حيث أن أبسط طريقة من حيث المتطلبات والاجراءات (الوقت، الجهد، المعالجات الإحصائية) من وجهة نظر الباحثين هي صدق المحكمين (43.00%)، ثم صدق التكوين الفرضي بنسبة 32.06% من الاستخدامات في أدوات جمع المعطيات في أطروحات دكتوراه علم النفس وعلوم التربية بجامعة الجزائر2.

لاحظ الباحث أثناء تحليل الأطروحات أن أغلب الباحثين استخدموا نفس الطرق في كل أدواتهم المستخدمة في أطروحتهم، كما لاحظ أن الطلاب يقلدون بعضهم البعض، بحيث يكونوا ينتمون إلى نفس القسم أو يؤطرهم نفس المشرف، حيث أنه مهما اختلفت أدواتهم وأهدافهم فإنهم يستخدمون نفس الطريقة، حتى أصبح الخطأ مركب وتأخذ الطريقة صفة الشيوع، ليس لمناسبتها ولكن لعدم إدراك الباحثين أن استخدام كل طريقة يجب أن يكون له ما يبرره من حيث الغرض من الأداة والسمة التي تقيسها.

وتتفق هذه النتيجة مع ما أكدت عليه نتائج دراسة ميسيك (1994) من أهمية استخدام طريقة صدق التكوين الفرضي وأنه من الطرق التي تجعل الأسئلة ممثلة لمحتوى السمة مما يزيد من مصداقية النتائج.

كما تتفق مع دراستي ويتا ودانيال (1997، 1998) حيث كانت طريقة صدق التكوين الفرضي أكثر الطرق استخداما بنسبة 55.40%-47%.

كما أن أكثر الطرق تعقيدا هو الصدق المرتبط بمحك وحقق أقل نسبة 02.80%، حيث أنه يتطلب إجراءات تجريبية تحوي العديد من التقنيات التي تحتاج دراية وخبرة علمية للتحقق منه بكل دقة، مثل إيجاد المحك المناسب الذي يتعلق بالسمة التي تقيسها الأداة، مما يشكل عبئا على الباحث من حيث الجهد والوقت المبذولين.

تتفق النتيجة تقريبا مع ما توصلت إليه نورة القضيب (2001) حيث كانت نسبة استخدام صدق التكوين الفرضي 24.80% وبلغت نسبة صدق المحكمين 20.30%، وطريقة الصدق المرتبط بمحك 09.30%، صدق المحتوى 04.50%، الصدق الذاتي 07.10%، والصدق الظاهري 02.40%.

وتتفق نتائج السؤال مع نتائج دراسة غادة بخش (2007) حيث توصلت إلى أن الطريقة الأكثر شيوعا في التحقق من الصدق في عينة دراستها من الرسائل كانت طريقة صدق المحكمين بنسبة 54.51% ثم طريقة صدق التكوين الفرضي بنسبة 28.95%.

السؤال الثالث:

ما نسبة الاستخدامات الجيدة لطرق التحقق من صدق أدوات جمع المعطيات في أطروحات دكتوراه علم النفس وعلوم التربية بجامعة الجزائر2؟

للإجابة على هذا السؤال قام الباحث بإعداد الجدول التالي:

جدول رقم (10) يوضح تكرارات ونسب الإستخدامات الجيدة لطرق التحقق من الصدق

حالة الاستخدام	طرق التحقق من الصدق						المجموع	
	المحتوى		المرتبط بمحك		التكوين الفرضي			
	التكرار	ن %	التكرار	ن %	التكرار	ن %	التكرار	ن %
جيد	09	64.29	11	100.0	121	96.03	141	93.38
غير جيد	05	35.71	00	00.00	05	03.97	10	06.62
المجموع	14	100.0	11	100.0	126	100.0	151	100.0

من خلال الجدول السابق يتضح أن نسبة 93.38% من استخدامات طرق التحقق من صدق أدوات جمع المعطيات في أطروحات دكتوراه علم النفس وعلوم التربية بجامعة الجزائر2 كانت جيدة، وأن نسبة 06.62% منها غير جيدة.

حيث كانت طريقة الصدق المرتبط بمحك جيدة بنسبة 100%، وطريقة التكوين الفرضي جيدة بنسبة 96.00%، أما طريقة صدق المحتوى فكانت جيدة الاستخدام بنسبة 64.29%.

وتجدر الإشارة إلى أننا اقتصرنا على الطرق الثلاثة (صدق المحتوى، الصدق المرتبط بمحك، صدق التكوين الفرضي)، أما الطرق الأخرى (المحكمين، الذاتي، الظاهري) فلم نعتمدها باعتبارها طرق غير علمية بناءا على معايير الاختبارات النفسية والتربوية (APA, 1999). حيث لاحظ الباحث عند مراجعته للأطروحات في عينة الدراسة الحالية أن هناك خلط بين طريقة صدق المحكمين والصدق الظاهري، وكذلك بين طريقة صدق المحكمين وطريقة ثبات التقدير من حيث المسمى فقط، وتتفق هذه النتيجة مع نتيجة دراسة هامرسلي (1987) التي توصلت لوجود مجموعة من الأفكار المتنوعة حول مفهومي الصدق والثبات وأنهما يستخدما استخداما غامضا وغير منسق، كما تتفق مع توصلت إليه دراسة ويتا ودانيال (1997، 1998) حيث وجدت انتشارا ملحوظا لإساءة استعمال اللغة عند تسمية طرق وإجراءات التحقق من الصدق والثبات في المقالات التي تم فحصها.

كما أن الإجراء في طريقة الصدق الذاتي بحساب الجذر التربيعي لمعامل الثبات، لا ينتمي للصدق بمعناه الدقيق، فحسب فرج (2000) أن استخدام الصدق الذاتي لا يتعارض فقط مع المبادئ السيكومترية بل يعد تشويها لمعنى الصدق، ويسمى أحيانا بالصدق الهروبي لأن الباحث يلجأ إليه عندما يعجز أو يتكاسل عن التحقق من صدق أدواته بالطرق السليمة، وهو غير مقبول كبديل لطرق الصدق الأخرى.

قد يرجع انتشار الخطأ في استخدام طرق التحقق من الصدق إلى أن بعض الباحثين يعتقدون –لقلة الخبرة وقلة الإعداد الأكاديمي– أن كل أداة يجب تقدير صدقها بما يتيسر للباحث من الطرق سواء من حيث الجهد أو الوقت، وتبين هذه النتيجة أهمية ما أوصت به نتائج دراسة شيبارد (1993)، بضرورة توحيد مفهوم صدق الأداة والتي تؤكد أهمية تفسير الدرجات واستعمالاتها في الميادين التطبيقية في ظل إطار مفاهيمي نظري يصور العلاقة بين المحتوى والأداء وما يمثله من استجابات، وهذا ما أكد عليه الباحثين في الاتجاهات الحديثة لنظرية الصدق أمثال تيغزة (2009).

تتفق هذه النتيجة مع نتائج دراسة الثبيتي (1998-ب) حيث وجد أن 51.00% من عينة دراسته تم التحقق من صدقها باستخدام طرق لا تجيب عن السؤال الذي يطرحه صدق الأداة بأي حال من الأحوال، من بينها استخدام الجذر التربيعي لمعامل الثبات.

كما تتفق جزئيا مع النتائج التي توصلت إليها دراسة غادة بخش (2001) حيث كانت طريقة الصدق المرتبط بمحك جيدة بنسبة 66.67%، وطريقة التكوين الفرضي جيدة بنسبة 98.70%، أما طريقة صدق المحتوى فكانت جيدة الاستخدام بنسبة 80.00%.

السؤال الرابع:

ما هي الأخطاء التي وقع فيها الباحثون عند التحقق من صدق أدوات جمع المعطيات في أطروحات دكتوراه علم النفس وعلوم التربية بجامعة الجزائر2؟

يتفرع من هذا السؤال عدة أسئلة، يتعلق كل سؤال بعامل من العوامل التي تؤثر على صدق الأداة بشكل عام، بغض النظر عن الطريقة المستخدمة في تقدير الصدق؛ وفيما يلي عرض نتيجة كل سؤال من تلك الأسئلة الفرعية ثم استخلاص الإجابة على السؤال العام.

أ- السؤال الفرعي الأول:

هل كانت الطرق المستخدمة في التحقق من الصدق كافية؟

للإجابة على هذا السؤال قام الباحث بإعداد الجدول التالي:

جدول رقم (11) يوضح تكرارات ونسب عدم كفاية طرق التحقق من الصدق

كفاية طرق التحقق من الصدق	التكرار	النسبة المئوية
نعم	83	33.33
لا	166	66.67
المجموع	249	100.0

من خلال الجدول السابق يتضح أن خطأ عدم كفاية طرق التحقق من صدق الأدوات من الاخطاء التي وقع فيها الباحثون عند استخدامهم لطرق التحقق من الصدق في أدوات جمع المعطيات في أطروحات دكتوراه علم النفس وعلوم التربية بجامعة الجزائر2، حيث ظهر هذا الخطأ في 166 استخدام بنسبة 66.67% أي ثلثي مجموع الأدوات (249 أداة)، وتدل النسبة على أن ظهور هذا الخطأ كان متوسطا.

ويمكن تفسير ارتفاع نسبة خطأ عدم كفاية طرق التحقق من الصدق في أطروحات عينة الدراسة الحالية بأنه راجع لاستخدام أغلب الباحثين طريقة واحدة (غير علمية) للتحقق من أدواتهم والتي هي طريقة صدق المحكمين بالدرجة الأولى، ثم طريقة الصدق الذاتي وطريقة الصدق الظاهري.

وتتفق نتيجة هذا السؤال الفرعي مع ما توصلت إليه نورة القضيب (2001) حيث كانت أساليب التحقق من الصدق غير كافية بنسبة 59.70%، كما جاءت نتائج الدراسة الحالية مقاربة لنتائج دراسة غادة بخش (2007) حيث وجدت أن نسبة الوقوع في خطأ عدم كفاية طرق التحقق من صدق أدوات القياس 56.00%.

ب- السؤال الفرعي الثاني:

هل تم التحقق من صدق الأداة على عينة ممثلة لمجتمع الدراسة؟

للإجابة على هذا السؤال قام الباحث بإعداد الجدول التالي:

جدول رقم (12) يوضح تكرارات ونسب عدم التحقق من الصدق على عينة ممثلة لمجتمع الدراسة

التحقق من الصدق على عينة ممثلة	التكرار	النسبة المئوية
نعم	77	46.39
لا	89	53.61
المجموع	166	100.0

من خلال الجدول السابق نجد أن خطأ عدم التحقق من صدق الأداة على عينة ممثلة لمجتمع الدراسة من الأخطاء التي وقع فيها الباحثون عند استخدامهم لطرق التحقق من صدق أدواتهم، ويلاحظ أن عدد الأدوات الكلي 166 أداة، حيث تم استثناء 83 أداة وهي التي استخدمت طريقة صدق المحكمين أو صدق المحتوى فقط، لأنها لا تحتاج إلى عينات للتحقق من صدقها، وقد ظهر هذا الخطأ في 89 استخدام بنسبة 53.61% من مجموع الاستخدامات، مما يدل على أن ظهور هذا الخطأ كان متوسطا.

ويعود سبب الوقوع في هذا الخطأ لإهمال بعض الباحثين أو جهلهم بالعوامل المؤثرة على طرق التحقق من الصدق، حيث نجد أن بعض الباحثين يكتفون بالتحقق من الصدق على طبقة واحدة فقط، مثل الذكور دون الإناث أو فئة عمرية دون أخرى، ويترك بقية الطبقات ويستخرج مؤشر الصدق لها ويرتضيه كمعامل صدق لبقية الفئات رغم اختلاف خصائصها. كما أن بعض الباحثين لم يتطرق أصلا لخصائص العينة التي استخدمها في التحقق من صدق أداته، مما يدل على الإهمال الكبير لهذا الجانب.

وتختلف نتيجة هذا السؤال الفرعي مع ما توصلت إليه نورة القضيب (2001) حيث أنها وجدت أن عينة الصدق غير ممثلة لمجتمع الدراسة الأساسية بنسبة منخفضة 10.30%، ولكن تتفق إلى حد ما مع دراسة غادة بخش (2007) حيث ظهر هذا الخطأ بنسبة 34.90% من المجموع الكلي للاستخدامات.

ج- السؤال الفرعي الثالث:

هل تم التحقق من صدق الأداة على عينة مناسبة الحجم؟

للإجابة على هذا السؤال قام الباحث بإعداد الجدول التالي:

جدول رقم (13) يوضح تكرارات ونسب عدم التحقق من الصدق على عينة مناسبة الحجم

التحقق من الصدق على عينة مناسبة الحجم	التكرار	النسبة المئوية
نعم	102	61.45
لا	13	07.83
لم يحدد	51	30.72
المجموع	166	100.0

من خلال الجدول السابق نجد أن خطأ عدم التحقق من صدق الأداة على عينة مناسبة الحجم من الأخطاء التي وقع فيها الباحثون عند استخدامهم لطرق التحقق من صدق أدوات جمع المعطيات في أطروحات دكتوراه علم النفس وعلوم التربية بجامعة الجزائر2، ويلاحظ أن عدد الأدوات الكلي 166 أداة، حيث تم استثناء 83 أداة (استخدم فيها طريقة صدق المحكمين أو صدق المحتوى فقط)، وقد ظهر هذا الخطأ في 13 استخدام بنسبة 07.83% من مجموع الاستخدمات، مما يدل على أن ظهور هذا الخطأ كان منخفضا، كما أن 30.72% من الاستخدامات لم يذكر فيها الباحثون أي معلومات عن هذا الجانب.

ويعود سبب الوقوع في هذا الخطأ لإهمال بعض الباحثين أو جهلهم بالعوامل المؤثرة على طرق التحقق من الصدق، حيث نجد أن بعض الباحثين يكتفون باختيار عينة صغيرة الحجم، مما يدل على عدم وعيهم بأهمية حجم العينة، فكلما كان كبيرا بما فيه الكفاية كلما أمكن الحصول على نتائج إحصائية دقيقة تقلل من أخطاء اختيار العينة وبالتالي تمكن من تعميم هذه النتائج.

تتفق نتيجة هذا السؤال الفرعي مع ما توصلت إليه نورة القضيب (2001) حيث وجدت أن عينة الصدق غير مناسبة الحجم بنسبة 05.10%، في حين لم يحدد الباحث حجم العينة بنسبة 15.50%.

بعد أن ناقش الأسئلة الفرعية الثلاثة (كفاية الطرق، تمثيل وحجم العينة) التي من خلال الإجابة عليها تمت الإجابة على السؤال الرابع الذي يبحث في الأخطاء المتعلقة بالعوامل التي تؤثر على صدق الأداة بشكل عام التي وقع فيها الباحثون عند استخدام طرق التحقق من صدق أدوات جمع المعطيات في أطروحات دكتوراه علم النفس وعلوم التربية بجامعة الجزائر2، قام الباحث بإعداد الجدول التالي:

جدول رقم (14) يوضح نسب الأخطاء المتعلقة بالعوامل المؤثرة على الصدق

الأخطاء المتعلقة بالعوامل المؤثرة على صدق الأداة	النسبة المئوية	نسبة الإنتشار
عدم كفاية طرق التحقق من الصدق	66.67	52.04
عدم التحقق من الصدق على عينة ممثلة لمجتمع الدراسة	53.61	41.85
عدم مناسبة حجم عينة الصدق	07.83	06.11
المجموع	/	100.0

من خلال الجدول السابق يتضح أن هناك 03 أخطاء متعلقة بالعوامل المؤثرة على الصدق، حيث ظهرت بنسب متفاوتة، وكان خطأ عدم كفاية طرق التحقق من الصدق هو الأكثر ظهورا من بين هذه الأخطاء بنسبة 52.04%، ثم خطأ عدم التحقق من الصدق على عينة ممثلة لمجتمع الدراسة بنسبة 41.85%، ويأتي في الأخير خطأ عدم مناسبة حجم عينة الصدق بنسبة 06.11%.

وقد يرجع تكرار ظهور الأخطاء في جميع طرق الصدق تقريبا عند التحقق من صدق أدوات الدراسة إلى عدم وعي الباحثين بأهمية أدوات البحث وبالأسس التي تبنى عليها وبالعوامل التي تؤثر عليها، والباحث المبتدئ من طلاب الدراسات العليا قد لا يدرك الفرق بين الأدوات المختلفة من حيث طبيعتها والغرض من استخدامها، وبالتالي بمفاهيم ومبادئ القياس المتعلقة بكل نوع والتي من أهمها مفهوم الصدق والطريقة الملائمة لإيجاد مؤشراته بالنسبة لكل نوع من أنواع الأدوات في ضوء أهداف القياس في كل دراسة على حدة.

السؤال الخامس:

هل تم مراعاة شروط التطبيق عند التحقق من صدق أدوات جمع المعطيات في أطروحات دكتوراه علم النفس وعلوم التربية بجامعة الجزائر2؟

وللإجابة على السؤال تم عرض النتائج لكل طريقة من الطرق الثلاث للتحقق من صدق الأدوات على حدة، وهي صدق المحتوى، الصدق المرتبط بمحك وصدق التكوين الفرضي.

أ- صدق المحتوى:

للإجابة على السؤال المتعلق بطريقة صدق المحتوى قام الباحث بإعداد الجدول التالي:

جدول رقم (15) يوضح النتائج المتعلقة بجانب التحقق من صدق المحتوى

البند	الإستجابات	التكرار	النسبة المئوية
هل تم تحليل المحتوى بالوصف التفصيلي أو توصيف الأهداف؟	نعم	04	28.57
	لا	10	71.43
	المجموع	**14**	**100.0**
هل تم عمل جدول مواصفات؟	نعم	03	21.43
	لا	11	78.57
	المجموع	**14**	**100.0**
هل تم حساب نسبة الإتفاق بين المحكمين؟	نعم	05	35.71
	لا	09	64.29
	المجموع	**14**	**100.0**

من خلال الجدول السابق يتضح أن الباحثين وقعوا في 03 أخطاء متعلقة بإجراءات تطبيق طريقة صدق المحتوى، حيث ظهرت بنسب متفاوتة، فنجد أن نسبة 78.57% من الاستخدامات لم يتم عمل جدول مواصفات، ونسبة 71.43% لم يتم تحليل المحتوى بالوصف التفصيلي أو توصيف الأهداف، بينما نسبة 64.29% لم يتم حساب نسبة الإتفاق بين المحكمين.

وقد يرجع ذلك إلى عدم تحكم الباحثين بشروط وافتراضات طريقة صدق المحتوى أو بأهميتها.

ب- الصدق المرتبط بمحك:

للإجابة على السؤال المتعلق بطريقة الصدق المرتبط بمحك قام الباحث بإعداد الجدول التالي:

جدول رقم (16) يوضح النتائج المتعلقة بجانب التحقق من الصدق المرتبط بمحك

البند	الإستجابات	التكرار	النسبة المئوية
صدق المحك	تلازمي	11	100.0
	تنبؤي	00	00.00
	المجموع	**11**	**100.0**
نوع المحك	مقياس	05	45.45
	تقدير	01	09.10
	تحصيل دراسي	05	45.45
	المجموع	**11**	**100.0**
هل المحك متعلق بالسمة المقاسة؟	نعم	11	100.0
	لا	00	00.00
	المجموع	**11**	**100.0**
هل المحك صادق؟	نعم	00	00.00
	لا	00	00.00
	لم يحدد	11	100.0
	المجموع	**11**	**100.0**
هل المحك ثابت؟	نعم	00	00.00
	لا	00	00.00
	لم يحدد	11	100.0
	المجموع	**11**	**100.0**
ماهو المستوى المعتمد لمعامل الصدق؟	مرتفع (أكثر من 0.70)	06	54.55
	متوسط (0.30-0.69)	05	45.45
	منخفض (أقل من 0.30)	00	00.00
	المجموع	**11**	**100.0**

من خلال الجدول السابق يتضح أن كل طرق الصدق المتعلق بمحك المستخدمة كانت من نوع الصدق التلازمي ولم يستخدم الصدق التنبؤي على الإطلاق، ويرجع ذلك إلى طبيعة البحوث والغرض منها، والتي تناسبها طريقة الصدق التلازمي.

كما تم استخدام محكات متنوعة في التحقق من الصدق المرتبط بمحك وهي: مقياس آخر، تحصيل دراسي، تقديرات، حيث كانت نسبها على التوالي: 45.45%، 45.45%، 09.10%؛ ولاحظ الباحث أن أغلب المحكات التي استخدمت كانت جاهزة، وبعض الباحثين استخدموا الأدوات التي قاموا ببنائها كمحكات فيما بينها؛ فقد لاحظنا أن الكثير من الباحثين استخدم الطريقة الاقل تكلفة من حيث الجهد والوقت، ولو كانت على حساب الدقة العلمية.

ونلاحظ أيضا من خلال الجدول أن كل الباحثين في عينة الدراسة ركزوا على تعلق المحك المستخدم بالسمة المقاسة، وبالمقابل لم يذكروا أي معلومات حول صدق المحك أو ثباته؛ وهذا دليل على إهمال الباحثين لهذه الشروط المتعلقة بالتحقق من الصدق المرتبط بمحك، وعدم إدراكهم بأهميتها.

أما بالنسبة للمستويات المعتمدة لمعامل الصدق المتعلق بالمحك فقد كانت كالتالي: مرتفعة (أكثر من 0.70) بنسبة 54.55%، ومتوسطة (0.30-0.69) بنسبة 45.45%.

وتتفق نتائج هذا السؤال مع ما توصلت إليه دراسة نورة القضيب (2001) حيث وجدت أن المحكات المستخدمة كانت عبارة عن اختبارات بنسبة 69.10%، وتحصيل دراسي بنسبة 21.40%، وتقديرات بنسبة 09.50%؛ وكانت نسبة تعلق المحك بالسمة التي يقيسها الاختبار 100.0%. أما فيما يخص صدق وثبات المحك المستخدم، فكانت نسبة 17.90%، 20.50% من الاستخدامات لم يتم فيها ذكر معلومات حول الصدق والثبات على التوالي؛ ولكن تختلف مع نتيجة نورة القضيب (2001) حيث وجدت أن المستويات المعتمدة لمعامل صدق الاختبار في عينة دراستها من الرسائل كانت كالتي: مستويات مرتفعة بنسبة 23.10%، مستويات متوسطة بنسبة 43.60%، بينما كانت المستويات المنخفضة بنسبة 33.30%؛ ويرجع ذلك لاختلاف حجم العينة بين الدراستين، حيث بلغ عدد طرق التحقق من الصدق المرتبط بمحك في الدراسة الحالية 11 طريقة مقابل 39 طريقة في دراسة نورة القضيب (2001).

ج- صدق التكوين الفرضي:

للإجابة على السؤال المتعلق بطريقة صدق التكوين الفرضي قام الباحث بإعداد الجدول التالي:

جدول رقم (17) يوضح النتائج المتعلقة بجانب التحقق من صدق التكوين الفرضي

البند	الإستجابات	التكرار	النسبة المئوية
هل تم التعريف بالإطار النظري للسمة المقاسة؟	نعم	10	07.94
	لا	116	92.06
	المجموع	**126**	**100.0**
هل اشتقت الفرضيات من الإطار النظري للسمة المقاسة؟	نعم	10	07.94
	لا	116	92.06
	المجموع	**126**	**100.0**
ماهي طرق التحقق من صدق التكوين الفرضي للأداة؟	الإتساق الداخلي	72	49.32
	المقارنة الطرفية	56	38.36
	التحليل العاملي	07	04.79
	طريقة أخرى	10	06.85
	لم يحدد الطريقة	01	00.68
	المجموع	**146**	**100.0**
هل طرق التحقق من صدق التكوين الفرضي للأداة مناسبة؟	نعم	140	96.55
	لا	05	03.45
	المجموع	**145**	**100.0**
هل طرق التحقق من صدق التكوين الفرضي للأداة كافية؟	نعم	124	99.20
	لا	01	00.80
	المجموع	**125**	**100.0**

نلاحظ من خلال الجدول السابق أن هناك قصورا ملحوظا من قبل الباحثين باتباع الخطوات المنهجية للتحقق من صدق التكوين الفرضي، حيث لم يشر إلى البناء النظري للسمة المقاسة بنسبة 92.06%. فبالرغم من أن الباحث يفرد له فصلا كاملا من فصول البحث، ولكن يتحتم عليه عندما يستخدم هذا النوع من الصدق أن يذكر البناء النظري الذي اعتمد عليه ولو بالإشارة بشكل موجز إلى نظرية معينة أو مجموعة من النظريات التي حاول الباحث أن يوفق بينها.

الخطوة التالية التي يجب أن يتبعها الباحث عند التحقق من صدق التكوين الفرضي بعد أن يذكر النظرية التي اعتمد عليها، أن يحاول اشتقاق فرضيات مصاغة من إطاره النظري ثم يتحقق منها إحصائيا أو تجريبيا (Gronglund, 1976)، هذه الخطوة تم اغفالها كذلك من الباحثين بنسبة 92.06%.

استخدم الباحثون مجموعة من الطرق (الأنواع) للتحقق من التكوين الفرضي وهي حسب نظرية الصدق الحديثة بمثابة بينات أو دلالات أو مؤشرات (تيغزة، 2009)، وكانت الطريقة الأكثر استخداما هي طريقة الاتساق الداخلي بنسبة 49.32% ثم المقارنة الطرفية بنسبة 38.06%. كما تم استخدام طريقة أخرى تمثلت في الصدق التقاربي بنسبة 06.85%، واستخدمت طريقة التحليل العاملي بنسبة 04.79%، بينما لم تحدد الطريقة في استخدام واحد فقط أي بنسبة 00.68%.

ولقد كانت هذه الطرق غير مناسبة بنسبة 03.45% فقط، ويرجع ذلك في استخدام طريقة الاتساق الداخلي بصورة خاطئة، حيث قام بعض الباحثين بحساب معاملات الارتباط بين درجة البند والدرجة الكلية لأدوات متعددة المقاييس، في حين كان يجب في هذه الحالة استخدام الدرجة الكلية للمقاييس الفرعية وليس الأداة كلها.

كما نلاحظ أيضا أن طرق التحقق من صدق التكوين الفرضي لم تكن كافية بنسبة 00.80% من 125 استخدام بعد استثناء استخدام واحد لم يتم فيه تحديد الطريقة وتوضيح إجراءات التحقق، حيث وجد الباحث إجراء واحد تم فيه استخدام أحد اساليب ثبات التجانس للحكم على صدق بناء الأداة.

وتتفق هذه النتائج مع ما توصل إليه الثبيتي (1998) في دراسته بأن الملفت للنظر في النتائج التي توصل إليها أن صدق التكوين الفرضي (البناء النظري) الذي يعتبر العمود الفقري لعملية التحقق من صدق أدوات المعطيات لم يلق أي اهتمام من قبل الباحثين، مما قد يكون له الأثر الواضح على تحليل النتائج وتفسيرها وعمليات التطبيق.

كما تتفق أيضا مع نتائج دراسة نورة القضيب (2001) حيث توصلت إلى أن الباحثين لم يشيروا إلى البناء النظري بنسبة 96.40%، ولم يقوموا باشتقاق فرضيات مصاغة من الإطار النظري بنسبة 81.50%. كما وجدت أنه استخدمت طرق (بينات) مختلفة للتحقق من صدق التكوين الفرضي للأدوات المستخدمة في عينة دراستها وهي: الاتساق الداخلي، التحليل العاملي، الفروق بين الأعمار، تحليل البنود وبينات أخرى، الفروق بين الجماعات، وكانت نسبها على التوالي: 31.40%، 28.60%، 17.10%، 08.60%، 08.60%، 05.70%.

ويمكن تفسير النتائج التي تم التوصل إليها فيما يتعلق ببعض الممارسات الخاطئة لبعض الباحثين عند التحقق من صدق أدواتهم، في عدم التحكم بمفهوم القياس النفسي وبناء وتطوير الاختبارات، ينتج عن ذلك عدم الفهم الصحيح لطبيعة الأداة المستخدمة والغرض من استخدامها، وعدم نجاعة استعمال المفاهيم المتعلقة بالقياس النفسي وبناء الاختبار وأهمها الصدق؛ يتضح ذلك في ميل بعض الباحثين الى استخدام الطرق الشائعة ومحاكاة الباحثين الآخرين (خاصة إذا كانت الأطروحات يشرف عليها نفس الأستاذ) حتى ولو كانت غير سليمة أو غير دقيقة أو لم تستخدم استخداما صحيحا. كما يفسر الباحث أوجه القصور في عملية التحقق من صدق المحك في الصعوبة التي يواجهها الباحثون في الحصول على محك جيد يستوفي أهم الشروط التي يجب توافرها فيه، وهي مشكلة متأزمة دائما ما ينادي المختصون في القياس النفسي بحلها عن طريق استخدام محكات متعددة أو تصحيح معامل الصدق في ضوء ثبات المحك وثبات الاختبار، وغالبا ما يكتفي بتصحيح معامل الصدق في ضوء ثبات المحك فقط. أما فيما يتعلق بعدم اتباع الشروط المنهجية في التحقق من صدق التكوين الفرضي، فيفسرها الباحث بأن هذه الممارسات ما هي إلا انعكاس لأدبيات القياس النفسي والتي تنظر إلى الصدق وفق منظور تعددي، وتنظر إلى صدق التكوين الفرضي (البناء، المفهوم) على أنه نوع منها يتم التحقق منه عن طريق بينات محددة (تيغزة، 2009).

ثانيا: عرض ومناقشة النتائج المتعلقة بالثبات:

السؤال الأول:

هل تم التحقق من ثبات أدوات جمع المعطيات في أطروحات دكتوراه علم النفس وعلوم التربية بجامعة الجزائر2؟

للإجابة على السؤال السابق قام الباحث بإحصاء الأدوات التي تم التحقق من ثباتها في أطروحات الدكتوراه، سواء كانت من إعداد الباحث صاحب الدراسة أو باحث آخر، والجدول الآتي يوضح ذلك:

جدول رقم (18) يوضح تكرارات ونسب أدوات جمع المعطيات التي تم التحقق من ثباتها

التحقق من ثبات الأداة	إعداد الأداة								المجموع	
	الباحث نفسه		باحث عربي		باحث أجنبي		لم يحدد			
	التكرار	ن %	التكرار	ن %	التكرار	ن %	التكرار	ن %	التكرار	ن %
تم التحقق منه	88	38.60	33	14.47	94	41.23	13	05.70	228	65.90
لم يتم التحقق منه	60	50.85	12	10.17	44	37.29	02	01.69	118	34.10
المجموع	148	/	45	/	138	/	15	/	346	100.0

من خلال الجدول السابق يتضح أن نسبة 34.10% من الأدوات المستخدمة في أطروحات دكتوراه علم النفس وعلوم التربية بجامعة الجزائر2 لم يتم التحقق من ثباتها على الإطلاق؛ وبالنظر إلى متغير إعداد الأداة نجد أن 50.85% من الأدوات التي لم يتم التحقق من ثباتها كانت من إعداد الباحث نفسه، ونسبة 37.29% منها كانت من إعداد باحث آخر "أجنبي"، أما نسبة الأدوات التي كانت من إعداد باحث آخر "عربي" فقدرت بـ 10.17%، بينما الأدوات التي لم يتم التحقق من ثباتها ولم يحدد الباحث الذي أعدها فكانت نسبتها 01.69%.

ويتضح مما سبق أن هناك تجاوزا كبيرا في عملية تقدير ثبات أدوات جمع المعطيات المستخدمة في أطروحات دكتوراه علم النفس وعلوم التربية بجامعة الجزائر2، فنتيجة أن نسبة 50.85% من الأدوات التي لم يتم تقدير ثباتها كانت من إعداد الباحث نفسه، دليل على أن الباحثين لا يولون أهمية كبيرة لتقدير ثبات أدوات القياس؛ وكذلك بعض الباحثين لم يتحققوا من ثبات الأداة لأنه قد سبق استخدامها في دراسات سابقة، وله قيمة معامل ثبات عالي، ظنا منهم أن هذه المؤشرات تكفي لاستخدام الأداة عدة مرات دون تقدير ثباتها؛ إلا أن هذا الإجراء خاطئ نظرا لأن الثبات خاصية للمعطيات وليس خاصية للأداة؛ ولأن المعطيات تتأثر بطبيعة تجانس أو عدم تجانس العينة التي تطبق عليها الأداة، وتتأثر بالظروف المحيطة التي تطبق فيها، كان لزاما على الباحثين تقدير ثبات معطياتهم في كل مرة يطبقون فيها الأداة.

تتفق هذه النتائج مع ما توصلت إليه دراستي ويتا ودانيال (1997، 1998) حيث وجدا أن 47.33% من المقاييس لم يتم التحقق من ثباتها وهي نسبة مرتفعة مقارنة بنتيجة الدراسة الحالية.

وتتفق نتائج السؤال مع نتائج دراسة الثبيتي (1998-أ) فقد وجد أن 29.00% من أدوات القياس لم يتم التحقق من ثباتها، حيث أن نسبة 55.00% منها كانت معدة باللغة العربية وسبق استخدامها، 24.00% منها مترجمة استخدمت في بيئات عربية و26.00% منها من إعداد صاحب الدراسة وتستخدم لأول مرة.

كما وجدت نورة القضيب (2001) في دراستها أن 19.60% من الاختبارات المستخدمة في عينة دراستها من الرسائل لم يتحقق من ثباتها بنسبة 19.60%.

وتتفق النتائج أيضا مع نتائج دراسة غادة بخش (2007) حيث توصلت إلى أنه لم يتم التحقق من ثبات أدوات القياس بنسبة 26.90%، كانت 42.42% منها من إعداد الباحث، و29.30% من إعداد باحث آخر "عربي"، وأن 28.28% منها من إعداد باحث آخر "أجنبي".

السؤال الثاني:

ما هي الطريقة الأكثر شيوعا في التحقق من ثبات أدوات جمع المعطيات في أطروحات دكتوراه علم النفس وعلوم التربية بجامعة الجزائر2؟

للإجابة على هذا السؤال قام الباحث بإحصاء طرق الثبات المستخدمة في عينة الدراسة، وهي ستة (06) طرق، موضحة في الجدول التالي:

جدول رقم (19) يوضح تكرارات ونسب طرق التحقق من الثبات

طريقة التحقق من الثبات	التكرار	النسبة المئوية
إعادة التطبيق	76	27.34
التجزئة النصفية	79	28.41
كيودر ريتشاردسون	02	00.72
ألفا كرونباخ	110	39.57
ثبات التقدير	08	02.88
طريقة أخرى	03	01.08
المجموع	278	100.0

من خلال الجدول السابق يتضح أن أكثر طرق التحقق من ثبات أدوات القياس شيوعا هي طريقة معامل ألفا كرونباخ بنسبة 39.57% من مجموع استخدامات طرق التحقق من ثبات أدوات جمع المعطيات في أطروحات دكتوراه علم النفس وعلوم التربية بجامعة الجزائر2 (278 طريقة)، ثم طريقة التجزئة النصفية بنسبة 28.41% وطريقة إعادة التطبيق بنسبة 27.34% من الاستخدامات، ثم طريقة ثبات تقدير المحكمين بنسبة 02.88%. كما استخدمت طريقة أخرى وهي تحليل التباين بنسبة 01.08%، أما طريقة كيودر ريتشاردسون فبنسبة 00.72%، حيث تم استخدام الصيغة (KR-20) لمرة واحدة، والمرة الثانية لم يتم تحديد الصيغة (21/20)؛ ويتضح أيضا أنه لم تستخدم طريقة الصور المتكافئة على الإطلاق.

ويفسر الباحث في الدراسة الحالية أنه كلما تقدمت الطريقة وزادت تعقيدا كلما قل استخدامها، حيث أن أبسط الطرق استخداما هي طريقة معامل ألفا كرونباخ الذي يستخدم في مرة تطبيق واحدة، ويتميز بسهولة الحساب خاصة مع توفر برنامج الحزم الاحصائية SPSS، لذلك حقق أكبر نسبة استخدام عند تقدير ثبات أدوات جمع المعطيات في العينة الحالية، كما أن أغلب الأدوات التي استخدمت كانت متعددة البدائل، وهذا يعتبر سببا آخرا في زيادة استخدام طريقة ألفا كرونباخ لتقدير الثبات.

وأكثرها تعقيدا هي طريقة الصور المتكافئة والتي لم تستخدم على الإطلاق في عينة الدراسة الحالية، حيث أنها تتطلب إعداد صورتين متكافئتين من الأداة، مما يشكل عبئا على الباحث من حيث الجهد والوقت المبذولين.

ومن خلال ما سبق يتضح أن استخدام الباحثين لطرق التحقق من ثبات أدوات جمع المعطيات في أطروحات دكتوراه علم النفس وعلوم التربية بجامعة الجزائر2 يركز على أبسط الأنواع من حيث الإجراءات والقيود والجهد المبذول، وقد يعود ذلك لقلة الوقت المخصص للدراسة أو لقلة خبرة الباحث بتطبيق الطرق المختلفة للتحقق من ثبات أدوات القياس مما يدفعه لأن يختار أبسطها.

تتفق نتيجة هذا السؤال مع دراسة الثبيتي (1998-أ) حيث وجد أن أكثر الطرق شيوعا هي طريقة معامل ألفا كرونباخ بنسبة 30.00%، ثم التجزئة النصفية 22.00%، ثم إعادة الاختبار 18.00%، كما أنه لم تستخدم طريقة الصورة المتكافئة.

وتتفق أيضا مع توصلت إليه دراستي ويتا ودانيال (1997، 1998) حيث كانت طريقة معامل ألفا كرونباخ أكثر الطرق استخداما بنسبة 81.00%-50.70% على التوالي.

وتختلف جزئيا مع نتيجة دراستي ويتا دانيال (1997، 1998) ودراسة نورة القضيب (2001) حيث استخدمت طريقة الصور المتكافئة بنسبة 01.35%-00.50% على التوالي.

كما تتفق نتائج هذا السؤال مع ما توصلت إليه دراسة غادة بخش (2007) حيث وجدت أن أكثر الطرق استخداما هي ألفا كرونباخ بنسبة 52.80% ثم التجزئة النصفية بنسبة 27.50% من الاستخدامات، ثم طريقة إعادة الاختبار، ثبات التقدير، كودر-ريتشاردسون(20)، معامل الاستقرار والتكافؤ بنسبة 11.50%، 04.50%، 02.20%، 00.40% على التوالي، كما لم تستخدم أيضا طريقة الصور المتكافئة وطريقة كودر-ريتشاردسون(21).

السؤال الثالث:

ما نسبة الاستخدامات الجيدة لطرق التحقق من ثبات أدوات جمع المعطيات في أطروحات دكتوراه علم النفس وعلوم التربية بجامعة الجزائر2؟

للإجابة على هذا السؤال قام الباحث بإعداد الجدول التالي:

جدول رقم (20) يوضح تكرارات ونسب الإستخدامات الجيدة لطرق التحقق من الثبات

حالة الاستخدام	طرق التحقق من الثبات												المجموع	
	إعادة التطبيق		التجزئة النصفية		كودر ريتشاردسون		ألفا كرونباخ		ثبات التقدير		تحليل التباين			
	ت	ن %	ت	ن %	ت	ن %	ت	ن %	ت	ن %	ت	ن %	ت	ن %
جيد	68	89.47	66	83.54	02	100.0	85	77.27	08	100.0	02	66.67	231	83.09
غير جيد	08	10.53	13	16.46	00	00.00	25	22.73	00	00.00	01	33.33	47	16.91
المجموع	76	100.0	79	100.0	02	100.0	110	100.0	08	100.0	03	100.0	278	100.0

من خلال الجدول السابق يتضح أن نسبة 83.09% من استخدامات طرق التحقق من ثبات أدوات جمع المعطيات في أطروحات دكتوراه علم النفس وعلوم التربية بجامعة الجزائر2 كانت جيدة، وأن نسبة 16.91% منها غير جيدة.

وبتفصيل أكثر كانت طريقة كيودر ريتشاردسون جيدة على الإطلاق حيث كانت الأسئلة متجانسة ونظام التصحيح ثنائي، وكذلك طريقة ثبات التقدير كانت جيدة على الأطلاق حيث تم حساب درجة

الاتفاق بين تقديرات المحكمين. ثم طريقة إعادة التطبيق كانت غير جيدة بنسبة 10.53% حيث كانت بعض الأدوات التي استخدمت فيها هذه الطريقة تقيس سمات غير مستقرة ولم يراعي الباحث فيها معايير التطبيق، وكانت طريقة التجزئة النصفية غير جيدة بنسبة 16.46% حيث كانت بعض الأدوات إما غير متجانسة الأسئلة أو إنها اختبارات سرعة، كما أن بعض الباحثين لم يوضحوا طريقة التصحيح والبعض الآخر لم يقوموا بالتصحيح أصلا.

أما طريقة ألفا كرونباخ فقد كانت غير جيدة بنسبة 22.73% حيث اكتفى الباحثين بتقدير المعامل لكل الأداة (أداة عبارة عن مقاييس فرعية)، كما استخدم بعض الباحثين أيضا هذه الطريقة مع الأدوات ثنائية الاستجابة (نعم/لا)؛ وبالنسبة للطريقة الأخرى والتي تمثلت في طريقة تحليل التباين فقد كانت غير جيدة بنسبة 33.33% حيث استخدمت لتقدير معامل ثبات كل الأداة.

قد يرجع انتشار الخطأ في استخدام طرق التحقق من الثبات إلى أن بعض الباحثين لا يعرفون الفرق بين معنى معامل الثبات في كل طريقة من الطرق المختلفة ويعتقدون –لقلة الخبرة وقلة الإعداد الأكاديمي– أن كل أداة يجب تقدير ثباتها بما يتيسر للباحث من الطرق سواء من حيث الجهد أو الوقت.

كما لاحظ الباحث أثناء تحليل الأطروحات أن أغلب الباحثين استخدموا نفس الطرق في كل أدواتهم المستخدمة في أطروحاتهم، كما لاحظ أن بعض الطلاب يقلدون بعضهم البعض، بحيث يكونوا ينتمون إلى نفس القسم أو يؤطرهم نفس المشرف، حيث أنه مهما اختلفت أدواتهم وأهدافهم فإنهم يستخدمون نفس الطريقة في التحقق من الثبات، حتى أصبح الخطأ مركب وتأخذ الطريقة صفة الشيوع، ليس لمناسبتها ولكن لعدم إدراك الباحثين أن استخدام كل طريقة يجب أن يكون له ما يبرره من حيث الغرض من الأداة والسمة التي تقيسها، وبالتالي درجة الدقة والثقة في النتائج المترتبة على تطبيق الأدوات مثل معامل الثبات.

وتتفق نتائج هذا السؤال في بعض جوانبها مع ما توصلت إليه نورة القضيب (2001) حيث كانت الطرق المستخدمة في التحقق من ثبات الأدوات في عينتها غير جيدة بنسبة 17.70%، وطريقة ألفا كرونباخ غير جيدة بنسبة 21.80%، وطريقة إعادة الاختبار غير جيدة بنسبة 11.70%.

ولكنها تختلف في أن نورة القضيب (2001) وجدت في دراستها أن طريقة التجزئة النصفية غير جيدة بنسبة 32.10% أي أكبر من نسبة الدراسة الحالية (16.46%)، وطريقة كيودر ريتشاردسون غير جيدة بنسبة 33.30% بينما في الدراسة الحالية كان استخدام الطريقة جيد على الإطلاق.

كما تتفق النتائج أيضا مع ما توصلت إليه دراسة غادة بخش (2007) حيث كانت نسبة الاستخدامات الغير جيدة 11.90%، وكانت كل من طريقة كيودر ريتشاردسون وطريقة ثبات التقدير جيدة على الإطلاق.

وتختلف جزئيا مع نتائج غادة بخش (2007) حيث وجدت أن طريقة إعادة الاختبار كانت جيدة 100%، وكانت طريقة التجزئة النصفية غير جيدة بنسبة 35.14%، وطريقة ألفا كرونباخ غير جيدة بنسبة 04.23%، بينما كانت في الدراسة الحالية غير جيدة بنسبة 10.53%، 16.46%، 22.73% على التوالي.

السؤال الرابع:

ما هي الأخطاء التي وقع فيها الباحثون عند التحقق من ثبات أدوات جمع المعطيات في أطروحات دكتوراه علم النفس وعلوم التربية بجامعة الجزائر2؟

هذا السؤال يتفرع منه عدد من الأسئلة، يتعلق كل سؤال بعامل من العوامل التي تؤثر على ثبات الأداة بشكل عام، بغض النظر عن الطريقة المستخدمة في تقدير ثبات الأداة؛ وفيما يلي عرض نتيجة كل سؤال من تلك الأسئلة الفرعية ثم استخلاص الإجابة على السؤال العام.

أ- السؤال الفرعي الأول:

هل تم التحقق من ثبات الأداة على عينة ممثلة لمجتمع الدراسة؟

للإجابة على هذا السؤال قام الباحث بإعداد الجدول التالي:

جدول رقم (21) يوضح تكرارات ونسب عدم التحقق من الثبات على عينة ممثلة لمجتمع الدراسة

التحقق من الثبات على عينة ممثلة	التكرار	النسبة المئوية
نعم	100	45.45
لا	120	54.55
المجموع	220	100.0

من خلال الجدول السابق نجد أن خطأ عدم التحقق من ثبات الأداة على عينة ممثلة لمجتمع الدراسة من الأخطاء التي وقع فيها الباحثون عند استخدامهم لطرق التحقق من ثبات أدواتهم، ويلاحظ أن عدد الأدوات الكلي 220 أداة، حيث تم استثناء 08 أدوات وهي التي استخدمت طريقة تقدير ثبات المحكمين فقط، لأنها لا تحتاج إلى عينات للتحقق من ثباتها، وقد ظهر هذا الخطأ في 120 استخدام بنسبة 54.55% من مجموع الاستخدمات، مما يدل على أن ظهور هذا الخطأ كان متوسطا.

ويعود سبب الوقوع في هذا الخطأ لإهمال بعض الباحثين أو جهلهم بالعوامل المؤثرة على طرق التحقق من الثبات، حيث نجد أن بعض الباحثين يكتفون بالتحقق من الثبات على طبقة واحدة فقط، ويترك بقية الطبقات ويستخرج معامل الثبات لها ويرتضيه كمعامل ثبات لبقية الفئات رغم اختلاف خصائصها. كما أن بعض الباحثين لم يتطرق أصلا لخصائص العينة التي استخدمها في التحقق من ثبات أداته، مما يدل على الإهمال الكبير لهذا الجانب.

وتختلف نتيجة هذا السؤال مع ما توصلت إليه نورة القضيب (2001) حيث أن نسبة 10.90% من الرسائل في عينة دراستها وقعت في خطأ عدم التحقق من ثبات الأدوات على عينة ممثلة لمجتمع الدراسة، وهي نسبة منخفضة مقارنة مع نسبة الدراسة الحالية.

كما تختلف أيضا مع نتيجة غادة بخش (2007) حيث وجدت أن نسبة 02.97% من مجموع استخدامات التحقق من الثبات في الرسائل لم تتم على عينات ممثلة لمجتمعات الدراسة، وهي نسبة خطأ منخفضة جدا مقارنة بما توصلت إليه الدراسة الحالية.

ب- السؤال الفرعي الثاني:

هل تم التحقق من ثبات الأداة على عينة مناسبة الحجم؟

للإجابة على هذا السؤال قام الباحث بإعداد الجدول التالي:

جدول رقم (22) يوضح تكرارات ونسب عدم التحقق من الثبات على عينة مناسبة الحجم

التحقق من الثبات على عينة مناسبة الحجم	التكرار	النسبة المئوية
نعم	148	67.27
لا	23	10.46
لم يحدد	49	22.27
المجموع	220	100.0

من خلال الجدول السابق نلاحظ أن خطأ عدم التحقق من ثبات الأداة على عينة مناسبة الحجم من الأخطاء التي وقع فيها الباحثون عند استخدامهم لطرق التحقق من ثبات أدوات جمع المعطيات في أطروحات دكتوراه علم النفس وعلوم التربية بجامعة الجزائر2، ويلاحظ أن عدد الأدوات الكلي 220 أداة، حيث تم استثناء 08 أدوات (استخدم فيها طريقة ثبات التقدير فقط)، وقد ظهر هذا الخطأ في 23 استخدام بنسبة 10.46% من مجموع الاستخدمات، مما يدل على أن ظهور هذا الخطأ كان منخفضا، كما أن 22.27% من الاستخدامات لم يذكر الباحثون أي معلومات عن هذا الجانب.

ويعود سبب الوقوع في هذا الخطأ لإهمال بعض الباحثين أو جهلهم بالعوامل المؤثرة على طرق التحقق من الثبات، حيث نجد أن بعض الباحثين يكتفون باختيار عينة صغيرة الحجم، مما يدل على عدم

وعيهم بأهمية حجم العينة، فكلما كان كبيرا بما فيه الكفاية كلما أمكن الحصول على نتائج إحصائية دقيقة تقلل من أخطاء اختيار العينة وبالتالي تمكن من تعميم هذه النتائج.

وتتفق نتيجة هذا السؤال الفرعي مع ما توصلت إليه نورة القضيب (2001)، حيث كانت نسبة الرسائل في عينة دراستها التي وقعت في هذا الخطأ 05.20%، ونسبة 18.80% من الاستخدامات لم يذكر فيها الباحثون أي معلومات عن هذا الجانب، وهذه النسب مقاربة لما توصلت إليه الدراسة الحالية.

لكن تختلف النتيجة الحالية مع نتيجة دراسة غادة بخش (2007)، حيث ظهر خطأ عدم التحقق من الثبات على عينة مناسبة الحجم بنسبة 50.55% من مجموع استخدامات التحقق من الثبات في الرسائل، وهي نسبة خطأ مرتفعة جدا مقارنة بما توصلت إليه الدراسة الحالية.

ج- السؤال الفرعي الثالث:

هل كانت مستويات معاملات الثبات مقبولة؟

للإجابة على هذا السؤال قام الباحث بإعداد الجدول التالي:

جدول رقم (23) يوضح تكرارات ونسب مستويات معامل الثبات

مستوى معامل الثبات	التكرار	النسبة المئوية
مرتفع جدا (أكثر من 0.95)	17	06.41
مرتفع (0.85-0.95)	89	33.58
متوسط (0.75-0.84)	84	31.70
منخفض (0.65-0.74)	51	19.25
منخفض جدا (أقل من 0.65)	24	09.06
المجموع	265	100.0

من خلال الجدول السابق نلاحظ أن هناك تباين لمستويات معاملات الثبات الناتجة عن تطبيق طرق التحقق من الثبات، كما نلاحظ أن عدد المعاملات الكلي بلغ 265 معامل (لكل طريقة)، حيث تم استثناء 08 طرق استخدم فيها طريقة ثبات التقدير تم فيها تحديد نسبة الإتفاق فقط، و05 طرق أخرى

لم يتم تحديد معامل الثبات فيها، وقد بلغت نسب مستويات معامل الثبات على التوالي: 33.58%، 31.70%، 19.25%، 09.06%، 06.41% في كل من المستوى: المرتفع، المتوسط، المنخفض، المنخفض جدا، المرتفع جدا على الترتيب.

بالرغم من القيم المنخفضة لمعاملات الثبات في بعض الاستخدامات فقد ارتضاها الباحثون كمعاملات ثبات لأدواتهم، وبالتالي اعتمدوا على نتائج تلك الأدوات؛ حيث من الصعوبة بمكان تحديد معامل ثبات مقبول وتعميم هذه القيمة على كل الأدوات، وتتحدد القيمة المقبولة لمعامل الثبات بنوع الأداة، فيلاحظ أن اختبارات القدرة والاستعداد ترتفع معاملات الثبات فيها عنها في اختبارات الشخصية.

ويجتهد بعض العلماء بوضع قيم مقبولة لمعاملات الثبات بحسب نوع الاختبار، حيث يرى جاي (1993) أن اختبارات الاستعداد العقلي يجب أن يكون لها ثبات مرتفع (أكبرمن 0.90)، إلا أنه يمكن قبول معاملات ثبات في الثمانينات؛ ومقاييس الشخصية عادة لا تبلغ هذا الثبات العالي، حيث يمكن قبول معامل في السبعينات. كما أشار أهمان وجلوك (1980) إلى أن معامل ثبات الاختبارات التحصيلية المقننة يجب أن لا يقل عن 0.85، بينما يمكن أن يقل عن ذلك في اختبارات الشخصية. كما كانت هناك اجتهادات أيضا من قبل بعض العلماء في تحديد القيم المقبولة لمعامل الثبات في ضوء الغرض من الاختبار سواء كان الغرض من استخدام الاختبار اتخاذ قرارات على مستوى أفراد أو جماعات، حيث يطالب بثبات مرتفع عندما يكون الغرض من الاختبار اتخاذ قرار على مستوى أفراد، ومعامل ثبات أقل عندما يكون القرار على مستوى جماعات، إلا أنه يفضل أن يكون المعامل مرتفعا في كل الحالات.

وتتفق نتيجة هذا السؤال مع ما توصلت إليه نورة القضيب (2001) حيث كانت نسب مستويات معامل الثبات على التوالي: 51.70%، 25.30%، 10.10%، 06.70%، 06.20% في المستوى: المرتفع، المتوسط، المنخفض، المرتفع جدا، المنخفض جدا؛ وهي نسب مقاربة من نسب الدراسة الحالية.

كما تتفق أيضا مع نتيجة غادة بخش (2007) حيث وجدت أن نسبة معامل الثبات غير مقبولة بنسبة 08.55% عند التحقق من ثبات الأدوات المستخدمة في عينة دراستها.

بعد أن ناقش الباحث الأسئلة الفرعية الثلاثة التي من خلال الإجابة عليها تمت الإجابة على السؤال الرابع الذي يبحث في الأخطاء المتعلقة بالعوامل التي تؤثر على ثبات الأداة بشكل عام التي وقع فيها الباحثون عند استخدام طرق التحقق من ثبات أدوات جمع المعطيات في أطروحات دكتوراه علم النفس وعلوم التربية بجامعة الجزائر2، قام الباحث بإعداد الجدول التالي:

جدول رقم (24) يوضح نسب الأخطاء المتعلقة بالعوامل المؤثرة على الثبات

الأخطاء المتعلقة بالعوامل المؤثرة على ثبات الأداة	النسبة المئوية	نسبة الإنتشار
عدم التحقق من الثبات على عينة ممثلة لمجتمع الدراسة	54.55	73.65
عدم مناسبة حجم عينة الثبات	10.46	14.12
عدم مناسبة قيمة معامل الثبات	09.06	12.23
المجموع	/	100.0

من خلال الجدول السابق يتضح أن هناك 03 أخطاء متعلقة بالعوامل المؤثرة على الثبات، حيث ظهرت بنسب متفاوتة، وكان خطأ عدم التحقق من الثبات على عينة ممثلة لمجتمع الدراسة هو الأكثر ظهورا من بين هذه الأخطاء بنسبة 73.65%، ثم خطأ عدم مناسبة حجم عينة الثبات بنسبة 14.12%، ويأتي في الأخير خطأ عدم مناسبة قيمة معامل الثبات بنسبة 12.23%.

وقد يرجع تكرار ظهور الأخطاء في جميع طرق الثبات تقريبا عند التحقق من ثبات أدوات الدراسة إلى عدم وعي الباحثين بالأسس التي تبنى عليها أدوات البحث وبالعوامل التي تؤثر عليها، والباحث المبتدئ قد لا يدرك الفرق بين الأدوات المختلفة من حيث طبيعتها والغرض من استخدامها.

استنتاجات البحث:

من خلال مناقشة نتائج البحث في ضوء تساؤلاته، نستنتج أنه:

أولا: جانب الصدق:

1- تم التحقق من صدق أدوات جمع المعطيات بنسبة 72.00% من مجموع الأدوات المستخدمة في أطروحات دكتوراه علم النفس وعلوم التربية بجامعة الجزائر2، والبالغ عددها الإجمالي 346 أداة؛ ولم يتم التحقق منه بنسبة 28.00%؛ وبالنظر إلى متغير إعداد الأداة نجد أن 46.39% من الأدوات التي لم يتم التحقق من صدقها كانت من إعداد باحث آخر "أجنبي"، ونسبة 36.08% منها كانت من إعداد الباحث نفسه، أما نسبة الأدوات التي كانت من إعداد باحث آخر "عربي" فقدرت بـ 16.50%، بينما الأدوات التي لم يتم التحقق من صدقها ولم يحدد الباحث الذي أعدها فكانت نسبتها 01.30%.

2- أن أكثر طرق التحقق من صدق أدوات جمع المعطيات شيوعا هي طريقة صدق المحكمين بنسبة 43.00% من مجموع استخدامات طرق التحقق من الصدق (393 طريقة)، ثم طريقة صدق التكوين الفرضي بنسبة 32.06% من الاستخدامات، ثم الصدق الذاتي، الصدق الظاهري، صدق المحتوى، الصدق المرتبط بمحك بنسبة 12.21%، 06.11%، 03.56%، 02.80% على التوالي؛ كما استخدمت طريقة أخرى واحدة بنسبة 00.26%.

3- أن نسبة 93.38% من استخدامات طرق التحقق من صدق أدوات جمع المعطيات كانت جيدة، وأن نسبة 06.62% منها غير جيدة، حيث كانت طريقة الصدق المرتبط بمحك جيدة بنسبة 100%، وطريقة التكوين الفرضي جيدة بنسبة 96.00%، أما طريقة صدق المحتوى فكانت جيدة الاستخدام بنسبة 64.29%.

4- أن خطأ عدم كفاية طرق التحقق من صدق الأداة من الاخطاء التي وقع فيها الباحثون عند استخدامهم لطرق التحقق من الصدق، حيث ظهر هذا الخطأ في 166 استخدام بنسبة 66.67%.

5- أن خطأ عدم التحقق من صدق الأداة على عينة ممثلة لمجتمع الدراسة من الأخطاء التي وقع فيها الباحثون عند استخدامهم لطرق التحقق من صدق أدوات جمع المعطيات، وقد ظهر هذا الخطأ في 89 استخدام بنسبة 53.61% من مجموع الاستخدمات الكلي.

6- أن خطأ عدم التحقق من صدق الأداة على عينة مناسبة الحجم من الأخطاء التي وقع فيها الباحثون عند استخدامهم لطرق التحقق من صدق أدوات جمع المعطيات، وقد ظهر هذا الخطأ في 13 استخدام بنسبة 07.83% من مجموع الاستخدمات، كما أن 30.72% من الاستخدامات لم يذكر فيها الباحثون أي معلومات عن هذا الجانب.

7- أن الباحثين وقعوا في 03 أخطاء متعلقة بإجراءات تطبيق طريقة صدق المحتوى، حيث أن نسبة 78.57% من الاستخدامات لم يتم عمل جدول مواصفات، و71.43% لم يتم تحليل المحتوى بالوصف التفصيلي أو توصيف الأهداف، بينما 64.29% لم يتم حساب نسبة الإتفاق بين المحكمين.

8- أن كل طرق الصدق المتعلق بمحك المستخدمة كانت من نوع الصدق التلازمي ولم يستخدم الصدق التنبؤي على الإطلاق، كما تم استخدام المحكات المختلفة في التحقق من الصدق المرتبط بمحك وهي: مقياس آخر، تحصيل دراسي، تقديرات، حيث كانت نسبها على التوالي: 45.45%، 45.45%، 09.10%، كما أن كل الباحثين في عينة الدراسة ركزوا على تعلق المحك المستخدم بالسمة المقاسة، لكنهم بالمقابل لم يذكروا أي معلومات حول صدق المحك أو ثباته؛ أما بالنسبة للمستويات المعتمدة لمعامل الصدق المتعلق بمحك فقد كانت مرتفعة (أكثر من 0.70) بنسبة 54.55%، ومتوسطة (0.30-0.69) بنسبة 45.45%.

9- أن هناك قصورا ملحوظا من قبل الباحثين باتباع الخطوات المنهجية للتحقق من صدق التكوين الفرضي، حيث لم يشر إلى البناء النظري للسمة المقاسة بنسبة 92.06%، كما لم يتم اشتقاق فرضيات مصاغة من الإطار النظري بنسبة 92.06%، واستخدم الباحثون مجموعة من الطرق للتحقق من التكوين الفرضي، حيث كانت الطريقة الأكثر استخداما هي طريقة الاتساق الداخلي بنسبة 49.32% ثم المقارنة الطرفية بنسبة 38.06%، كما تم استخدام طريقة أخرى تمثلت في الصدق التقاربي بنسبة 06.85%، واستخدمت طريقة التحليل العاملي بنسبة 04.79%، بينما لم تحدد الطريقة في استخدام واحد فقط أي بنسبة 00.68%؛ ولقد كانت طرق التحقق من صدق التكوين الفرضي غير مناسبة بنسبة 03.45%، كما أن هذه الطرق لم تكن كافية بنسبة 00.80%.

ثانيا: جانب الثبات:

1- تم التحقق من ثبات أدوات جمع المعطيات بنسبة 65.90% من مجموع الأدوات المستخدمة في أطروحات دكتوراه علم النفس وعلوم التربية بجامعة الجزائر2، ولم يتم التحقق منه بنسبة 34.10%؛ وبالنظر إلى متغير إعداد الأداة نجد أن 50.85% من الأدوات التي لم يتم التحقق من ثباتها كانت من إعداد الباحث نفسه، ونسبة 37.29% منها كانت من إعداد باحث آخر "أجنبي"، أما نسبة الأدوات التي كانت من إعداد باحث آخر "عربي" فقدرت بـ 10.17%، بينما الأدوات التي لم يتم التحقق من ثباتها ولم يحدد الباحث الذي أعدها فكانت نسبتها 01.69%.

2- أن أكثر طرق التحقق من ثبات أدوات جمع المعطيات شيوعا هي طريقة معامل ألفا كرونباخ بنسبة 39.57% من مجموع استخدامات طرق التحقق من الثبات (278 طريقة)، ثم طريقة التجزئة النصفية بنسبة 28.41% وطريقة إعادة التطبيق بنسبة 27.34% من الاستخدامات، ثم طريقة ثبات تقدير المحكمين بنسبة 02.88%. كما استخدمت طريقة أخرى وهي تحليل التباين بنسبة 01.08%، أما

طريقة كيودر ريتشاردسون فبنسبة 00.72%، حيث تم استخدام الصيغة (KR-20) لمرة واحدة، والمرة الثانية لم يتم تحديد الصيغة (21/20)، كما أنه لم تستخدم طريقة الصور المتكافئة على الإطلاق.

3- أن نسبة 83.09% من استخدامات طرق التحقق من ثبات أدوات جمع المعطيات كانت جيدة، وأن نسبة 16.91% منها غير جيدة، حيث كانت كل من طريقة كيودر ريتشاردسون وطريقة ثبات التقدير جيدة على الأطلاق، أما طريقة إعادة التطبيق كانت غير جيدة بنسبة 10.53% ، وكانت طريقة التجزئة النصفية غير جيدة بنسبة 16.46%، بينما طريقة ألفا كرونباخ فقد كانت غير جيدة بنسبة 22.73%، وبالنسبة للطريقة الأخرى والتي تمثلت في طريقة تحليل التباين فقد كانت غير جيدة بنسبة 33.33%.

4- أن خطأ عدم التحقق من ثبات الأداة على عينة ممثلة لمجتمع الدراسة من الأخطاء التي وقع فيها الباحثون عند استخدامهم لطرق التحقق من ثبات أدوات جمع المعطيات، وقد ظهر هذا الخطأ في 120 استخدام بنسبة 54.55% من مجموع الاستخدمات.

5- أن خطأ عدم التحقق من ثبات الأداة على عينة مناسبة الحجم من الأخطاء التي وقع فيها الباحثون عند استخدامهم لطرق التحقق من ثبات أدوات جمع المعطيات، وقد ظهر هذا الخطأ في 23 استخدام بنسبة 10.46% من مجموع الاستخدمات، بينما 22.27% من الاستخدامات لم يذكر الباحثون أي معلومات عن هذا الجانب.

6- أن هناك تباين لمستويات معاملات الثبات الناتجة عن تطبيق طرق التحقق من الثبات، كما نلاحظ أن عدد المعاملات الكلي بلغ 265 معامل (لكل طريقة)، حيث بلغت نسب مستويات معامل الثبات على التوالي: 33.58%، 31.70%، 19.25%، 09.06%، 06.41% في كل من المستوى: المرتفع، المتوسط، المنخفض، المنخفض جدا، المرتفع جدا على الترتيب.

التوصيات والإقتراحات:

بناءا على نتائج الدراسة الحالية فإن الباحث يوصي بما يلي:

1. تركيز أساتذة قسم علم النفس وعلوم التربية الذين يقومون بتدريس مادة القياس النفسي على خطوة جمع المعطيات، باعتبارها خطوة هامة من خطوات البحث العلمي، وخاصة موضوع الصدق والثبات وكيفية التحقق منهما.
2. تبصير طلاب الدراسات العليا بالأخطاء التي قد يقع فيها الباحثين فيما يتعلق بطرق التحقق من صدق وثبات أدوات جمع المعطيات، من خلال تقويم الرسائل والأطروحات الجامعية المتوفرة في القسم، وذلك للتعرف على أنماط الأخطاء الشائعة بين الباحثين في هذا المجال بصورة عملية.
3. عقد ورشات ودورات تدريبية لطلاب الدراسات العليا حول كيفية التحقق من صدق وثبات أدوات جمع المعطيات، مع التركيز على توضيح شروط تطبيق كل طريقة من الطرق المستخدمة.
4. إشراك طلاب الدراسات العليا في نشاطات المخابر الجامعية كمشاريع البحث CNEPRU وذلك للتدرب العملي على البحوث العلمية وخاصة بما يتعلق ببناء أدوات جمع المعطيات.
5. إجراء دراسات تقييمية دورية شاملة لمراحل بناء وتطوير الاختبارات النفسية والتربوية المستخدمة في الرسائل الجامعية والأبحاث المنشورة في المجلات العلمية.
6. إعادة النظر في الطرق النمطية الكلاسيكية، والإطلاع على التطورات الحديثة في القياس النفسي والطرق المتقدمة مثل التحليل العاملي التوكيدي (النمذجة)، ثبات المفهوم (المركب)، التباين المفسر ونظرية الإستجابة للمفردة IRT.

المراجع

أولا: المراجع العربية

1- إبراهيم بن مبارك الدوسري: **الإطار المرجعي للتقويم التربوي**، ط2، الرياض: مكتب التربية العربي، (2000).

2- أحمد الخطيب وآخرون: **دليل البحث والتقويم التربوي**، عمان: دار المستقبل، (1985).

3- أحمد سليمان عودة: **القياس والتقويم في العملية التدريسية**، الأردن: دار الأمل للنشر والتوزيع، (1998).

4- أحمد محمد عبد الخالق: **قياس الشخصية**، ط1، الكويت: مطبوعات الجامعة، (1996).

5- امحمد بوزيان تيغزة: الاتجاهات الحديثة لنظرية الصدق ومتضمناتها التطويرية للقياس وتصميم الاختبارات، **بحوث ندوة قسم علم النفس وكلية التربية لجامعة الملك سعود حول علم النفس وقضايا التنمية الفردية والمجتمعية**، الرياض، (2009).

6- امحمد بوزيان تيغزة: تجويد البحوث العلمية والرسائل الجامعية – نحو منهجية لتطوير آليات التنظير والأطر النظرية السائدة، **المؤتمر العربي الدولي الثاني لضمان جودة التعليم العالي**، البحرين، (2012).

7- بدر محمد الأنصاري: **قياس الشخصية**، الكويت: دار الكتاب الحديث، (2000).

8- بشير معمرية: **القياس النفسي وتصميم أدواته – للطلاب والباحثين في علم النفس والتربية**، ط2، الجزائر: منشورات الحبر، (2007).

9- رجاء محمود أبو علام: **مدخل إلى مناهج البحث التربوي**، الكويت: مكتبة الفلاح، (1987).

10-رجاء محمود أبو علام: **مناهج البحث الكمي والنوعي والمختلط**، ط1، الأردن: دار المسيرة، (2013).

11-رمزية الغريب: **التقويم والقياس النفسي والتربوي**، القاهرة: مكتبة الأنجلو، (1996).

12-روبرت ثورندايك وإليزبيث هيجن: **القياس والتقويم في علم النفس والتربية** (ترجمة عبد الله زيد الكيلاني و عبد الرحمن عدس)، ط4، عمان: مركز الكتب الأردني، (1986).

13-زكريا محمد الظاهر وآخرون: **مبادئ القياس والتقويم في التربية**، عمان: دار الثقافة للنشر والتوزيع، (1999).

14-سامي محمد ملحم: **القياس والتقويم في التربية وعلم النفس**، الأردن: دار المسيرة، (2005).

15-سبع محمد أبولبدة: **مبادئ القياس النفسي والتقييم التربوي**، عمان: دار الفكر، (2008).

16-سعد جلال: **القياس النفسي، المقاييس والاختبارات**، القاهرة: دار الفكر العربي، (1985).

17-سعد عبد الرحمن: **القياس النفسي – النظرية والتطبيق**، ط3، القاهرة: دار الفكر العربي، (1998).

18-صفوت فرج: **القياس النفسي**، ط4، القاهرة: مكتبة الأنجلو المصرية، (2000).

19-صفوت فرج: **القياس النفسي**، ط6، القاهرة: مكتبة الأنجلو المصرية، (2007).

20-صلاح أحمد مراد وأمين علي سليمان: **الاختبارات والمقاييس في العلوم النفسية والتربوية**، القاهرة: دار الكتاب الحديث، (2002).

21-صلاح الدين محمود علام: **القياس والتقويم التربوي والنفسي أساسياته وتطبيقاته وتوجهاته المعاصرة**، ط1، القاهرة: دار الفكر العربي، (2000).

22-عباس محمود عوض: **القياس النفسي بين النظرية والتطبيق**، الإسكندرية: دار المعرفة الجامعية، (1998).

23-عبد الرحمن بن سليمان الطريري: **القياس النفسي والتربوي، نظرياته، أسسه، تطبيقاته**، الرياض: مكتبة الرشد، (1997).

24-عبد الرحمن عيسوي: **القياس والتجريب في علم النفس والتربية**، الإسكندرية: دار المعرفة الجامعية، (1985).

25-عبد العزيز بوسالم: **القياس في علم النفس والتربية – الأسس النظرية والمبادئ التطبيقية**، ط1، الجزائر: دار قرطبة للنشر والتوزيع، (2014).

26-عبد القادر كراجة: **القياس والتقويم في علم النفس**، ط2، عمان: دار اليازودي العلمية للنشر والتوزيع، (2001).

27-عزيز سمارة وآخرون: **مبادئ القياس والتقويم في التربية**، الإسكندرية: دار الفكر للنشر والتوزيع، (1989).

28- علي حامد الثبيتي: أخطاء شائعة بين تصاميم البحوث التربوية والنفسية وعلاقة ذلك بالصدق الإحصائي للنتائج وتعميمها، **رسالة الخليج العربي**، السنة 1، ع(44)، 51-89، (1992).

29-علي حامد الثبيتي: صدق البناء النظري لأدوات جمع المعلومات في البحوث التربوية والنفسية وانعكاس ذلك على تفسير النتائج، **رسالة الخليج العربية**، السنة 19، ع(96)، 17-62، (1998).

30-غادة بنت عبد القادر يار محمد بخش، **تقويم طرق تقدير ثبات وصدق أدوات القياس المستخدمة في بعض رسائل الماجستير بكلية التربية بجامعة أم القرى**، رسالة ماجستير غير منشورة، جامعة أم القرى، كلية التربية، مكة المكرمة، (2007).

31-فؤاد أبو حطب وآخرون: **التقويم النفسي**، ط4، القاهرة: مكتبة الأنجلو المصرية، (1997).

32-فؤاد البهي السيد: **علم النفس الإحصائي وقياس العقل البشري**، ط3، القاهرة: دار الفكر العربي، (1979).

33-فيصل عباس: **الإختبارات النفسية، تقنياتها و إجراءاتها**، بيروت: دار الفكر العربي، (1996).

34-ليونا تايلر: **الإختبارات والمقاييس** (ترجمة:سعد عبد الرحمن)، القاهرة: دار الشروق، (1998).

35-محمد شحاتة ربيع: **قياس الشخصية**، الإسكندرية: دار المعرفة الجامعية، (1994).

36-محمد عبد السلام أحمد: **القياس النفسي والتربوي**، القاهرة: مكتبة النهضة المصرية، (1960).

37-محمود محمد غانم: **القياس والتقويم**، حائل: دار الأندلس، (1997).

38-مروان عبد المجيد إبراهيم: **الأسس العلمية والطرق الإحصائية للاختبارات والقياس في التربية الرياضية**، عمان: دار الفكر للطباعة والنشر، (1999).

39-موسى النبهان: **أساسيات القياس في العلوم السلوكية**، ط1، الأردن: دار الشروق للنشر والتوزيع، (2004).

40-نورة عبد الرحمن القضيب: **أساليب التحقق من ثبات وصدق الاختبارات المستخدمة في رسائل الماجستير في اقسام علم النفس في بعض جامعات المملكة العربية السعودية – دراسة تقويمية مقارنة**، رسالة ماجستير غير منشورة، جامعة الملك سعود بن عبد العزيز، كلية التربية، الرياض، (2001).

41-وليم مهرنز وإرفن ولهمن: **القياس والتقويم في التربية وعلم النفس** (ترجمة هيثم الزبيدي وماهر أبو هلالة)، العين: دار الكتاب الجامعي، (2003).

ثانيا: المراجع الأجنبية

1- Adkins, D.C. (1974). **Test Construction**. Johio, Abell & company.

2- Aiken, L. (2000). **Psychological Testing and Assessment**. Boston: Boston.

3- American Psychological Association. (1985). **Standards For Education and Psychological Tests**, Washing Gotoo, d.c. Acethor.

4- American Psychological Association. (1999). **Standards For Education and Psychological Tests**, Washing Gotoo, d.c. Acethor.

5- Anastasi, A & Urbina , S. (1997). **Psychological testing**, 7th ed; New York; prentice hall.

6- Anderson,J.C.(1981). **Issues in Language Testing**, "ELT" Documents lll. The British Council.

7- Bechtoldt, H, P. (1959). Construct validity: **A critique**, **American Psychologist, 14,5.**

8- Brewer, J.K. (1969). **Checklist for Test and Manual Evaluation**. Master Dissertation, university of Tennessee.

9- Brown, F. G. (1983). **Principles of education and psychological testing**, 4th ed, New York, Holt. Rineh & Winston.

10- Crocker. L., & Algina, J. (1986). **Introduction to classical and modern test theory**. New York: Holt, Rinehart, and Winston, In.

11- Cronbach, L. J. (1960). (1951). Coefficient alpha and the internal structure of tests, **Psychometrika, 16, 3**: 297-334.

12- Cronbach, L. J. (1960). **Essentials of psychological testing**. New York, Harper & Row.

13- Dick, C. Hagert. Y. (1971). **Topics in Measurement**. New York, McGraw. Hill.

14- Ebel, R. L. (1972). **Essentials of Educational measurement**, New Jersey; prentice Hall.

15- Edwards, A.L. (1964). **Techniques of Attitude Scale Construction,** Appletion _Century_Crofts, New York.

16- Grawitz, M. (1996). **Méthodes des sciences sociales**, Paris, Dalloz, coll. Précis Dalloz, 10e éd.

17- Guilford, J. P. (1954). **Psychometric methods**. N. Y. McGraw, Hill.

18- Gutman, L. (1945). Basis for analyzing test-retest reliability, **Psychometrika, 10, 4**: 255-282.

19- Hammersley, M. (1987). Some Notes on the terms Validity and Reliability. **British Educational Research Journal, 13, 1**: 73: 81.

20- Helnstadter, G. C. (1966). **Principles of psychological measure**. London, Methuen and co, ltd.

21- Holt, R.R. (1971). **Assessing personality**, New York, Harcount Brance jovanorich.

22- Jenkins, J.G. (1978). Validity for what? **Journal of consulting psychology.**

23- Kroll, A. (1960). Item Validity as a Factor in test Validity. **Journal of Education Psychology**, 31, 2.

24- Kuder, G. and Richardson, M. W. (1937). The theory of the estimation of test reliability. **Psychometrika**. **2 (3),** 151-165.

25- Ley, P. (1972). **Quantitative Aspects of Psychological Assessment: An Introduction**; London; Gerald Duck Worth & co.

26- Lindquist, E.F. (1950). **Statistical Analysis in Educational Research**, Boston, Houghton Mifflin.

27- Magnusson, David. (1967), **Test Theory**. Palo Alto, Addison-Wesley publishing.

28- Maloney, P.m & Ward, P.M (1980). **Psychological Assessment A conceptual Approach**. New York; Oxford University Press.

29- Mehrens. W. A. & Lehmann. (1986). **Measurement and evaluation in education psychological. Testing principles and application**. New York; Hall international inc.

30- Meir, S. Davis, S. (1990). Trends in reporting psychometric properties of scales used in counseling psychological research. **Journal of counseling Psychology, 37, 1:** 113-115.

31- Messick, S. (1994b**)**. **Validity of psychological Assessment: Validation of Inferences from persons' Response and Performances as scientific Inquiry into Score meaning**. Princeton, N. J: Educational Testing Services. (ERIC No. ED 380466).

32- Messic, S. (1995b). Standards of Validity and the Validity of Standards in Performance Assessment. **Educational Measurement: Issues and Practice**, winter.

33- Moss, P. A. (1995). Themes and Variation in Validity theory. **Educational Measurement: Issues and Practice, 14, 2**: 5-12.

34- Ottenbacher, K. (1995). An examination of reliability in developmental research. **Journal of Developmental and Behavioral Pediatrics, 19, 3:** 177-182.

35- Shepard, L.A. (1993). Evaluating Test Validity. Review of Research in Education. Edited by Linda Darling Hammond. **Colombia University**, 19, 405-449.

36- Spearman, C. (1910). Coefficient of correlation calculated from faulty data. **British Journal of Psychology, 3:** 271-295.

37- Sundberg, N. D. (1977). **Assessment of persons**. Englewood cliffs, N. J. prentice- Hall.

38- Weiner, E. A. & Stewart, B. J. (1984). **Assessing individuals Psychological and Educational tests and measurement**. New York. Litlle. brown & co.

39- Witta, E. Lea and Daniel, Larry G. (1997). Implication for Teaching Graduate Students Correct Terminology for Discussing Validity and Reliability Based on a Content Analysis of Three social Science Measurement Journals. **Paper Presented at the annual meeting of the American Educational Research Association.** Chicago. IL. March 24-28. (ERIC No. ED 408853).

40- Witta, E. Lea and Daniel, Larry G. (1998). The Reliability and Validity of Test Scores: Are Editorial Policy Changes Reflected in Journal Articles? **Paper presented at the annual meeting of the American Educational Research Association.** San Diego? CA April 13-17. (ERIC No. ED 422366).

الملاحق

ملحق رقم (01):

أداة البحث

أولا: البيانات العامة

1-1	الرقم التسلسلي للأطروحة: ...		**2-1**	رقم الأداة في الأطروحة:				
3-1	القسم	1	علم النفس	2	علوم التربية			
4-1	التخصص في علم النفس	1	اجتماعي	2	عمل وتنظيم	3	عيادي	
5-1	إعداد الأداة	1	الباحث نفسه	2	باحث عربي	3	باحث أجنبي	
6-1	نوع الأداة	1	اختبار	2	مقابلة	3	ملاحظة	
7-1	السمة المقاسة:							

ثانيا: جانب الصدق

أ- أسئلة عامة:

الرقم	**السؤال**	**نعم**	**لا**	**ملاحظات**
1-2	هل تم التحقق من صدق الأداة؟			
2-2	ما هي الطرق المستخدمة في التحقق من صدق الأداة؟ 1. ..... 2. ..... 3. ..... 4. ..... 5. ..... 6.			
3-2	هل طرق التحقق من صدق الأداة كافية؟			
4-2	هل عينة الصدق ممثلة لمجتمع الدراسة؟			
5-2	هل حجم عينة الصدق كاف؟			

ب–أسئلة خاصة:

الرقم	الأسئلة الخاصة بصدق المحتوى	نعم	لا	ملاحظات
2-1-1	هل تم تحليل المحتوى بالوصف التفصيلي أو توصيف الأهداف؟			
2-1-2	هل تم عمل جدول مواصفات؟			
2-1-3	هل تم حساب نسبة الإتفاق بين المحكمين؟			
2-1-4	هل الطريقة مناسبة للغرض من الأداة؟			
الرقم	**الأسئلة الخاصة بالصدق المرتبط بمحك**	**نعم**	**لا**	**ملاحظات**
2-2-1	هل صدق المحك: 1- تلازمي 2- تنبؤي			
2-2-2	نوع المحك:			
2-2-3	هل المحك متعلق بالسمة المقاسة؟			
2-2-4	هل المحك صادق؟			
2-2-5	هل المحك ثابت؟			
2-2-6	هل الطريقة مناسبة للغرض من الأداة؟			
2-2-7	ماهو المستوى المعتمد لمعامل الصدق؟ 1-مرتفع 2-متوسط 3-منخفض			
الرقم	**الأسئلة الخاصة بصدق التكوين الفرضي**	**نعم**	**لا**	**ملاحظات**
2-3-1	هل تم التعريف بالإطار النظري للسمة المقاسة؟			
2-3-2	هل اشتقت الفرضيات من الإطار النظري للسمة المقاسة؟			
2-3-3	ماهي طرق التحقق من صدق التكوين الفرضي للأداة؟ 1. ..... 2. ..... 3.			
2-3-4	هل طرق التحقق من صدق التكوين الفرضي للأداة مناسبة؟ .الطريقة رقم 1 .الطريقة رقم 2 .الطريقة رقم 3			
2-3-5	هل طرق التحقق من صدق التكوين الفرضي للأداة كافية؟			

ثالثا: جانب الثبات

الرقم	السؤال	نعم	لا	ملاحظات
3-1	هل تم التحقق من ثبات الأداة؟			
3-2	ماهي الطرق المستخدمة في التحقق من ثبات الأداة؟ 1. ..... 2. ..... 3.			
3-3	هل الطرق المستخدمة في التحقق من ثبات الأداة مناسبة؟ .الطريقة رقم 1 .الطريقة رقم 2 .الطريقة رقم 3			
3-4	هل عينة الثبات ممثلة لمجتمع الدراسة؟			
3-5	هل حجم عينة الثبات كاف؟			
3-6	ماهو المستوى المعتمد لمعامل الثبات؟ .الطريقة رقم 1 .الطريقة رقم 2 .الطريقة رقم 3			

ملحق رقم (02):

تعليمات ملء بيانات الشبكة ومعايير التقييم

أولا: تعليمات ملء بيانات شبكة التقييم:

تخصص لكل أداة على حده شبكة تقييم مستقلة، حيث أن الاجابة عن الأسئلة الواردة بالشبكة، ستكون مخصصة لطرق التحقق من الصدق والثبات الواردة بأداة واحدة فقط، وتتكون الشبكة من ثلاثة (03) أقسام، وذلك على النحو التالي:

القسم الأول: يتضمن بيانات عامة عن الأطروحة التي استخدمت فيها الأداة، وعن الأداة موضوع التقييم، ويراعى في ملء البيانات ما يلي:

1- كتابة الرقم التسلسلي للأطروحة في المربع المخصص له.

2- كتابة رقم الأداة المستخدمة في كل أطروحة على حده.

3- وضع علامة (X) في المربع أمام القسم الذي تنتمي إليه الأطروحة التي استخدمت فيها الأداة، إما علم النفس أو علوم التربية.

4- وضع علامة (X) في المربع أمام التخصص الذي تنتمي إليه الأطروحة (من قسم علم النفس) التي استخدمت فيها الأداة، إما علم النفس الإجتماعي، عمل وتنظيم أو علم النفس العيادي.

5- وضع علامة (X) في المربع أمام من أعد الأداة، إما الباحث نفسه، باحث عربي أم باحث أجنبي.

6- وضع علامة (X) في المربع أمام نوع الأداة، سواء كانت اختبار، مقابلة أم ملاحظة.

7- كتابة السمة التي تقيسها الأداة: ذكاء، شخصية، اتجاهات، ميول، قيم، قدرة/استعداد، درجة التحصيل أو سمة أخرى.

القسم الثاني: يتضمن هذا القسم جانبين، الأول يتعلق بأسئلة عامة في التحقق من الصدق، أما الثاني يتعلق بأسئلة خاصة بكل طريقة من طرق التحقق من الصدق الثلاثة وهي: صدق المحتوى، الصدق المرتبط بمحك، وصدق التكوين الفرضي؛ ويندرج تحت كل جانب عدد من الأسئلة، ولكل سؤال عدة إجابات، وقد وضع بجانب كل إجابة من هذه الإجابات مجموعة من الأرقام ويتبع ما يلي للإجابة عن هذه الأسئلة:

1- تتم الإجابة عن الأسئلة المتعلقة بالجانب الاول الذي هو عبارة عن أسئلة عامة تسري على جميع طرق التحقق من الصدق، بحيث تتم قراءة كل سؤال من هذا الجانب ثم وضع علامة (X) على الإجابة التي تنطبق عليه.

2- تحدد الطريقة الأولى المستخدمة في التحقق من صدق الأداة، ثم تتم قراءة كل سؤال من هذه الطريقة، ثم بعد ذلك يوضع علامة (X) على الإجابة التي تنطبق عليه؛ وإذا تم استخدام طرق أخرى للتحقق من صدق نفس الأداة، يتم اتباع نفس الخطوات كما في الطريقة الأولى.

القسم الثالث: يتضمن هذا القسم الأسئلة المتعلقة بالثبات، بحيث يندرج تحت هذا الجانب عدد من الأسئلة العامة، ولكل سؤال عدة إجابات، وقد وضع بجانب كل إجابة من هذه الإجابات مجموعة من الأرقام، ويتبع ما يلي للإجابة عن هذه الأسئلة:

تتم الإجابة عن الأسئلة المتعلقة بهذا الجانب الذي هو عبارة عن أسئلة عامة تسري على جميع طرق الثبات، بحيث تتم قراءة كل سؤال من هذا الجانب ثم وضع علامة (X) على الإجابة التي تنطبق عليه.

ثانيا: معايير التقييم:

قبل القيام بتحليل مضمون طرق التحقق من الصدق والثبات، يجب قراءة التعليمات التالية التي تعتبر بمثابة معايير أساسية في عملية التقييم، والإجابة عن الأسئلة التي تتضمنها شبكة التقييم، وهذه المعايير تتضمن جانبين هما الصدق والثبات، فيما يلي تفصيل للمعايير المتعلقة بكل منهما:

أولا: جانب الصدق:

وينقسم هذا الجانب إلى معايير عامة ومعايير خاصة، فيما يلي عرض لكل منهما:

أ– معايير عامة:

هذه المعايير عامة بمعنى أنها تسري على كل طرق التحقق من الصدق المستخدمة وهي:

2-1: تعتبر الأداة المستخدمة في الأطروحة تم التحقق من صدقها (بغض النظر عن الطريقة المستخدمة في التحقق من الصدق)، إذا تحقق الباحث من صدق أداته في دراسته الحالية، وذلك إذا كانت الأداة من إعداد الباحث نفسه، أو كانت الأداة تستخدم لأول مرة في البيئة الجزائرية، ويستوي الحال في ذلك أن تكون الأداة مترجمة تستخدم لأول مرة، أو سبق استخدامها في بيئات عربية أخرى، أما إذا كانت الأداة المستخدمة قد سبق إعدادها على البيئة الجزائرية فيكتفي بصدق الأداة الأصلي، أو صدق الأداة في دراسات سابقة، إلا إذا اختلف مجتمع البحث، أو مضى على إعدادها أكثر من عشر سنوات، أو اختلف الهدف من استخدام الأداة كأن تكون الأداة قد أعدت لأغراض التنبؤ ويستخدمها الباحث في دراسته الحالية لأغراض التشخيص، في هذه الحالات فإن على الباحث أن يتحقق من صدق أداته في دراسته.

2-2: على الباحث تسجيل طرق الصدق التي استخدمها للتحقق من صدق الأداة سواء كانت:

1- صدق المحتوى: عن طريق التحكيم أو جدول المواصفات.

2- الصدق المرتبط بمحك: الارتباط بين المقياس ومحك موضوعي موثوق منه: اختبار آخر، تقديرات، تحصيل دراسي، مجموعات متضادة، أداء فعلي.

3- صدق التكوين الفرضي (صدق البناء أو المفهوم): باستخدام أحد المؤشرات التالية: الإتساق الداخلي (الارتباط بين درجة الفقرة ودرجة البعد أو الدرجة الكلية للمقياس ككل، أو الارتباط بين درجة البعد والدرجة الكلية للمقياس)، التحليل العاملي، الفروق بين الجماعات (إيجاد الفروق بين درجات مجموعة المرتفعين ومجموعة المنخفضين في السمة المقاسة)، الفروق بين الأعمار (إيجاد الفروق بين درجات مجموعة الأكبر عمرا ومجموعة الأصغر عمرا في السمة المقاسة)، مصفوفة السمات، الطرق التجريبية، تحليل الأسئلة (أو تحليل البنود، عن طريق حساب: معامل التمييز، معامل الصعوبة، معامل الثبات للبند)، المقارنة الطرفية (تحليل المجموعتين المتضادتين في المحك على المفردة أو المقياس الفرعي)، ثبات التجانس (معامل ألفا كرونباخ أو معادلتي كيودر-ريتشاردسون 21/20).

4- صدق المحكمين: إذا تم استخدام إجراءات التحكيم فقط.

5- الصدق الذاتي: إذا تم حساب الجذر التربعي لمعامل الثبات.

6- الصدق الظاهري (السطحي): أي تحقق سطحي دون اللجوء إلى تحكيم أو إعداد جدول مواصفات.

7- عند استخدام طريقة أخرى غير الطرق السابقة، فإنه يتم تحديدها.

2-3: تعتبر الطريقة المستخدمة في التحقق من صدق الأداة غير كافية في الحالات التالية:

أ- عندما يقتصر التحقق من صدق الأداة على الصدق الظاهري أو الصدق الذاتي.

ب- عندما يقتصر التحقق من صدق الأداة على صدق المحكمين لما يكون المجال المقاس غير محدد.

ت- عندما يقتصر التحقق من صدق الأداة على صدق المحتوى لما يكون المجال المقاس غير محدد.

ث- عندما يقتصر التحقق من صدق الأداة على طريقة ثبات التجانس.

ج- عندما تكون الأداة عبارة عن مقياس أحادي البعد ولم يستخدم طرق للكشف عن تجانسه بطريقة الاتساق الداخلي أو التحليل العاملي.

ح- عندما تكون الأداة عبارة عن مقياس متعدد الأبعاد أو يقيس أكثر من سمة ولم يستخدم التحليل العاملي للكشف عن العوامل التي يتكون منها.

2-4: يتم التحقق من صدق الأداة، بتطبيقها على عينة ممثلة تمثيلا جيدا لنفس مجتمع الدراسة الأساسية.

2-5: يتم التحقق من صدق الأداة على عينة مناسبة من حيث العدد، حيث تعتبر عينة الصدق مناسبة إذا بلغت 30 فردا فأكثر، وتعتبر غير مناسبة إذا قلت عن هذا العدد.

ب- معايير خاصة بكل طريقة من طرق التحقق من الصدق:

1) المعايير الخاصة بصدق المحتوى:

يختص هذا الجانب بالإجابة عن الأسئلة الخاصة بإصدار أحكام حول طريقة التحقق من الصدق باستخدام صدق المحتوى، وفيما يلي المعايير المتعلقة بالإجابة عن كل سؤال من هذه الأسئلة:

2-1-1: يتم تحليل المحتوى تحليلا دقيقا حتى يتم التأكد من أن جميع الجوانب الأساسية تغطيها أسئلة الأداة تغطية ملائمة وبالنسب الصحيحة، ويتم ذلك إذا قام الباحث بما يلي:

أ-وصف تفصيلي دقيق لمدى المهام والمثيرات والمواقف والمواد التي يستوعبها نطاق المحتوى.

ب- توصيف الأهداف التعليمية أو التدريبية أو العلاجية، شرط أن تصاغ في عبارات إجرائية سلوكية قابلة للملاحظة والقياس.

2-1-2: يتم استخدام جدول المواصفات كطريقة من طرق تحليل المحتوى، وذلك بتحديد الأهمية النسبية للأهداف ثم مناظرة الأسئلة بالأهداف، وعلى هذا الأساس يتم تحديد عدد المفردات التي ترتبط بكل هدف في كل موضوع، وذلك بإعداد جدول ثنائي توضع فيه الأهداف التعليمية أفقيا والموضوعات رأسيا

والوزن النسبي لكل موضوع، وهذه الأوزان يجب تحديدها مسبقا باستطلاع آراء الخبراء، أما الوزن النسبي للمحتوى فيحدد موضوعيا في ضوء الوقت المخصص لتدريسه فعليا، ومن الأفضل أن تكون نسب مئوية.

2-1-3: يتم استخدام التحليل الكيفي للبنود للحكم على مدى فعاليتها من حيث صياغتها وصلاحيتها، وفي هذا الصدد يلجأ الباحث إلى مجموعة من المحكمين لإبداء الرأي في مختلف الجوانب الكيفية للسؤال، وبعد ذلك يتم حساب نسبة الاتفاق بين المحكمين.

2-1-4: تستخدم طريقة صدق المحتوى، عندما يكون المجال المقاس محددا أو معرفا، بحيث يعتمد على مقرر دراسي أو برنامج تدريبي أو على مجموعة موحدة من الخبرات السابقة التي يمكن منها صياغة محتوى الأداة، وعليه فالاختبارات التحصيلية من أفضل المجالات مناسبة لقبول صدق المحتوى بخلاف اختبارات الاستعدادات ومقاييس الشخصية، وحتى اختبارات التحصيل التي تقيس العمليات المعرفية العليا، فصدق المحتوى لا يعد كافيا لمثل هذه الأنواع من الاختبارات، بل لا يعد أساسيا؛ ذلك أن هذا النوع من الاختبارات لا تتشابه ضمنيا مع ميدان السلوك الذي تعد عينة منه.

2) المعايير الخاصة بالصدق المرتبط بمحك:

يختص هذا الجانب بالإجابة عن الأسئلة الخاصة بإصدار أحكام حول طريقة التحقق من الصدق باستخدام الصدق المرتبط بالمحك، وفيما يلي المعايير المتعلقة بالإجابة عن كل سؤال من هذه الأسئلة:

2-2-1: يحدد نوع صدق المحك، إما تلازمي أو تنبؤي.

2-2-2: يحدد نوع المحك المستخدم، سواء كان اختبارا، تقديرا، تحصيلا دراسيا، مجموعات متضادة، أو أداء فعليا؛ وعندما تستخدم أنواع أخرى من المحكات يتم تحديدها.

2-2-3: يجب أن يكون المحك متعلقا بالوظيفة التي تقيسها الأداة، بحيث يكون بينهما ارتباط منطقي.

2-2-4: يجب أن يكون المحك ثبت صدقه وفق ما تشير إليه البحوث والتقارير.

2-2-5: يجب أن تتوفر في المحك الجيد خاصية الثبات، فثبات المحك يؤثر إلى حد كبير في معامل الصدق، أي الإرتباط بين المقياس والمحك، لذلك من الضروري اختيار محك عالي الثبات عند حساب الصدق المرتبط بمحك.

2-2-6: يستخدم الصدق التلازمي عندما يكون الغرض من استخدام الأداة هو التشخيص، حيث يستخدم مع الاختبارات التشخيصية التي يلجأ إليها في حل المشكلات الراهنة أكثر من مشكلات المستقبل، وتجمع به بيانات الاختبار والمحك في نفس الفترة الزمنية، أو بفاصل زمني قصير جدا.

أما الصدق التنبؤي فيستخدم عندما يكون الغرض من استخدام الأداة هو التنبؤ طويل المدى، حيث يستخدم مع اختبارات الاستعدادات والاختبارات التي تستخدم في انتقاء الأفراد وتصنيفهم وتوجيههم تربويا أو مهنيا أو عسكريا، وكذلك لأغراض التنبؤ الإكلينيكي، ويكون هناك فاصل زمني طويل بين جمع بيانات الاختبار والمحك (قد تصل إلى عدة سنوات).

2-2-7: يتم الحكم على معامل الصدق طبقا للمستويات التالية:

-مستوى صدق مرتفع إذا بلغت قيمته من 0,70 فأكثر.

-مستوى صدق متوسط إذا تراوحت قيمته بين 0,30-0,69.

-مستوى صدق منخفض إذا كانت قيمته أقل من 0,30.

3) المعايير الخاصة بصدق التكوين الفرضي:

يختص هذا الجانب بالإجابة عن الأسئلة الخاصة بإصدار أحكام حول طريقة التحقق من الصدق باستخدام صدق التكوين الفرضي (البناء، المفهوم)، وفيما يلي المعايير المتعلقة بالإجابة عن كل سؤال من هذه الأسئلة:

2-3-1: يتم التعريف بالإطار النظري للسمة التي يفترض ارتباطها بالنتائج.

2-3-2: يتم اشتقاق فرضيات من الإطار النظري للسمة المقاسة.

2-3-3: تحدد الطرق المستخدمة في التحقق من الصدق التكويني للأداة، سواء كانت: الإتساق الداخلي، التحليل العاملي، الفروق بين الجماعات، الفروق بين الأعمار، مصفوفة السمات-الطرق المتعددة، الطرق (الإجراءات) التجريبية، تحليل الأسئلة (معامل التمييز، معامل الصعوبة للبنود)، المقارنة الطرفية، ثبات التجانس. وإذا تم استخدام طريقة أخرى غير الطرق السابقة يتم تحديدها.

2-3-4: تعتبر الطرق المستخدمة في التحقق من الصدق التكويني للأداة مناسبة، في الحالات التالية:

أ- عندما يكون المقياس أحادي البعد واستخدم الباحث طريقة الاتساق الداخلي أو التحليل العاملي للكشف عن تجانسه.

ب- عندما يكون المقياس متعدد الأبعاد واستخدم التحليل العاملي للكشف عن أبعاده.

ت- عندما تكون السمة المقاسة تختلف اختلافا واضحا من مجموعة إلى أخرى وتم استخدام الفروق بين الجماعات للكشف عن هذه الفروق.

ث- عندما تكون السمة المقاسة تختلف اختلافا واضحا من فئة عمرية إلى أخرى وتم استخدام الفروق بين الأعمار للكشف عن هذه الفروق.

2-3-5: تعتبر الطرق المستخدمة في التحقق من الصدق التكويني للأداة كافية، في الحالة التالية:

- عندما لا يقتصر على أساليب ثبات التجانس للحكم على صدق بناء المقياس.

ثانيا: جانب الثبات:

ويتضمن هذا الجانب المعايير المتعلقة بجميع الطرق المستخدمة في التحقق من الثبات:

3-1: تعتبر الأداة المستخدمة في الأطروحة تم التحقق من ثباتها (بغض النظر عن طريقة التحقق المستخدمة)، إذا تحقق الباحث من ثبات أداته في دارسته الحالية، وذلك إذا كانت الأداة من إعداد

الباحث نفسه، أو كانت الأداة تستخدم لأول مرة في البيئة الجزائرية، ويستوي الحال في ذلك أن تكون الأداة مترجمة تستخدم لأول مرة، أو سبق استخدامها في بيئات عربية أخرى، أما إذا كانت الأداة المستخدمة قد سبق إعدادها على البيئة الجزائرية فيكتفي بثبات الأداة الأصلي، أو ثبات الأداة في دراسات سابقة، إلا إذا اختلف مجتمع البحث، أو مضى على إعدادها أكثر من عشر سنوات، أو اختلف الهدف من استخدام الأداة كأن تكون الأداة قد أعدت لأغراض التنبؤ ويستخدمها الباحث في دراسته الحالية لأغراض التشخيص، في هذه الحالات فإن على الباحث أن يتحقق من ثبات أداته في دراسته الحالية.

3-2: على الباحث أن يذكر الطريقة المستخدمة في التحقق من ثبات الأداة، سواء كانت: إعادة الإختبار، الصور المتكافئة، التجزئة النصفية، طرق التجانس (معادلة كرونباخ، معادلة كيودر-ريتشاردسون20، معادلة كيودر-ريتشاردسون21)، أو ثبات التقدير، أو أي طريقة أخرى.

3-3: يستخدم في التحقق من ثبات الأداة الطريقة التي تتناسب مع طبيعة الأداة والغرض من استخدامها بحيث:

- تستخدم طريقة إعادة الإختبار مع الأدوات التي تقيس سمات تتمتع بالاستقرار عبر الوقت مثل الإتجاهات، الميول، القيم، الاستعدادات، ذكاء البالغين وبعض سمات الشخصية (التكيف الشخصي والاجتماعي). بينما في حالة الاختبارات التي تقيس سمات نفسية غير مستقرة مثل الحالة المزاجية أو القلق أو ذكاء الأطفال أو التذكر، أو الاختبارات التي تتطلب أسئلتها استخدام العمليات العليا كالاستدلال والابتكار، فإننا نقوم باستخدام طريقة الإعادة إما: بخفض الفترة الزمنية الفاصلة بين التطبيقين، بحيث لا تقل عن أسبوع ولا تزيد عن أسبوعين، إذا كانت العينة من الأطفال أقل من سبع سنوات أو الأشخاص الذين يعالجون بعقاقير تؤثر في ذاكرتهم، أو الأشخاص مرتفعي الذكاء؛ أو بزيادة الفترة الزمنية الفاصلة بين

التطبيقين، بحيث تكون أكثر من أسبوعين وأقل من ستة أشهر، إذا كانت العينة من البالغين أو كانت السمة المقاسة من النوع الذي يسهل تذكره أو تعلمه أو يتعرض أفراده فيه إلى الألفة.

كما يتناسب استخدام هذه الطريقة مع اختبارات السرعة، وكذلك تستخدم لتقدير ثبات التقارير الذاتية.

– تستخدم طريقة الصور المتكافئة في حال تواجد صورتين متكافئتين، بحيث يبرر الباحث تكافؤ هذين الصورتين من حيث: تساوي الصورتين من الناحية الإحصائية (المتوسطات، الانحراف المعياري، تساوي معاملات الارتباط بين الأسئلة)، وتساويها من ناحية بنية الأسئلة (مستوى صعوبة الأسئلة في حالة اختبارات الذكاء والقدرات العقلية والاختبارات التحصيلية والاستفتاءات، معامل شيوع استجابة الأسئلة في حالة اختبارات الشخصية، القدرة التمييزية، طول الأداة، طريقة صياغة الأسئلة، نوع السمة المقاسة، عدد مكونات السمة التي تقيسها الأداة، المحتوى الذي يقيسه السؤال) وتوحيد إجراءات التطبيق في المرتين (التعليمات، طريقة الإجابة، التصحيح، زمن الإجابة). كما يجب أن تكون الفترة الزمنية بين التطبيقين قصيرة، لكي لا تؤثر التغيرات الحاصلة للمفحوصين خلال هذا الفاصل الزمني في قيمة معامل التكافؤ.

تناسب هذه الطريقة اختبارات الذاكرة والاختبارات التحصيلية المقننة لضعف تأثير عاملي الذاكرة والتدريب على قيمة معامل التكافؤ، وتستخدم في الدراسات التي تهدف لمعرفة أثر عامل تجريبي معين في الأداء، خاصة إذا تطلب الأمر قياس قبلي وبعدي؛ كما تستخدم في الاختبارات التي تهدف لقياس القدرة على حل المشكلات، وعند تقويم الطلاب أو المناهج، وفي مواقف العلاج النفسي للتعرف على فاعلية العلاج.

يفضل استخدامها لتقدير ثبات الاختبارات التي تقيس ذكاء الأطفال، بينما لا يفضل استخدامها في اختبارات القدرات والميول والاتجاهات لصعوبة إعدادها، وكذلك في الأدوات التي تقيس سمة عريضة ومركبة مثل القدرة على التفكير الناقد أو الإبداع.

– تستخدم طريقة معامل الاستقرار والتكافؤ عندما نرغب في التنبؤ بعيد المدى والاستدلال على محتوى دراسي معين؛ وزيادة على طريقة الصور المتكافئة يجب أن تكون الفترة الزمنية الفاصلة بين التطبيقين طويلة.

– تستخدم طريقة التجزئة النصفية مع الأدوات التي تتسم بتجانس المحتوى، بحيث يكون جميعه يقيس نفس السمة النفسية، وإذا كانت الأداة تقيس أكثر من سمة فالأفضل أن تقسم كل أداة فرعية على حده، ومن ثم حساب ثباتها بشكل منفصل. كما تستخدم طريقة التجزئة النصفية عندما لا يكون الأداء على الأداة يعتمد بشكل تراكمي على الأداء في الخطوات السابقة مثل اختبارات القدرة على حل المشكلات؛ وإذا كانت الأداة تضم مجموعة الأسئلة التي تتعلق بمشكلة واحدة فيجب أن توضع المجموعة في أحد نصفي الأداة؛ أما في حالة تجزئة الاختبار التحصيلي حسب الأرقام (فردية، زوجية) فيجب أن تكون الأسئلة مرتبة حسب مستوى الصعوبة. كما لا تتناسب هذه الطريقة مع اختبارات السرعة، والتقارير الذاتية.

– تستخدم طرق التجانس المختلفة (معادلة كرونباخ، معادلتي كيودر–ريتشاردسون20 و21)، عندما يكون المقياس متجانسا، أي يقيس خاصية أو قدرة واحدة؛ أما لو تعددت السمات والقدرات التي تقيسها الأداة فلن تكون هذه الطرق مناسبة لتقدير ثبات المقياس بكامله، لكن تفيد في حساب ثبات الأجزاء بشكل منفصل، كما يقتصر استخدام طرق التجانس على اختبارات القوة. تستخدم معادلة كرونباخ في الأدوات التي يكون نظام التصحيح فيها يقوم على أوزان متدرجة للأسئلة مثل مقياس ليكرت بأنواعه (كاستبيانات الاتجاهات، بعض مقاييس الشخصية واختبارات التحصيل المقالية)، بينما تستخدم معادلة كيودر–ريتشاردسون في الأدوات التي يكون نظام التصحيح فيها على شكل (1،0)، وتضيف الصيغة (21) منها شرطا آخر هو تقارب البنود في مستوى الصعوبة.

– تستخدم طرق ثبات التقدير مع الأدوات ذات الاستجابات المفتوحة والتي غالبا ما تتأثر بالذاتية، مثل: الاختبارات الاسقاطية، الاختبارات الخاصة بالقدرات الابداعية، اختبارات المقالة أو الاختبارات

الإنشائية، مقاييس التقدير المتدرجة وأدوات الملاحظة؛ ولا يكتفي بمعامل ثبات التقدير وحده كمعامل ثبات للمقياس، بل يتعين حساب ثبات المقياس بأحد الطرق المناسبة إلى جانب حساب ثبات التقدير.

3-4: يتم التحق من ثبات الأداة، بتطبيقها على عينة ممثلة تمثيلا جيدا لمجتمع الدراسة الأساسية.

3-5: يتم التحقق من ثبات الأداة على عينة مناسبة من حيث العدد، حيث تعتبر عينة الثبات مناسبة، إذا بلغت (30) فردا فأكثر، وتعتبر غير مناسبة إذا قلت عن هذا العدد.

3-6: يتم الحكم على مستوى معامل الثبات المعتمد للأداة وفق تصنيفات خمسة (تصنيف مرفي):

1- مستوى ثبات مرتفع جدا (أكثر من 0,95)

2- مستوى ثبات مرتفع (0,85-0,95)

3- مستوى ثبات متوسط (0,75-0,84)

4- مستوى ثبات منخفض (0,65-0,74)

5- مستوى ثبات منخفض جدا (أقل من 0,65)

ملحق رقم (03):

نتائج تطبيق أداة البحث

رقم البند	البند	الاستجابات	التكرار	النسبة المئوية
1-1	الأطروحات	-	112	-
2-1	الأدوات	-	346	-
3-1	القسم	علم النفس	248	71.70
		علوم التربية	98	28.30
		المجموع	346	100.0
4-1	التخصص في علم النفس	اجتماعي	55	22.18
		عمل وتنظيم	73	29.43
		عيادي	120	48.39
		المجموع	248	100.0
5-1	إعداد الأداة	الباحث نفسه	148	42.80
		باحث آخر عربي	45	13.00
		باحث آخر أجنبي	138	39.90
		لم يحدد	15	04.30
		المجموع	346	100.0
6-1	نوع الأداة	اختبار	298	86.10
		مقابلة	27	07.80
		ملاحظة	21	06.10
		المجموع	346	100.0
1-2	هل تم التحقق من صدق الأداة؟	نعم	249	72.00
		لا	97	28.00
		المجموع	346	100.0
2-2	ما هي الطرق المستخدمة في التحقق من صدق الأداة؟	صدق المحكمين	169	43.00
		صدق التكوين الفرضي	126	32.06
		الصدق الذاتي	48	12.21
		الصدق الظاهري	24	06.11
		صدق المحتوى	14	03.56
		صدق المحك	11	02.80
		طريقة أخرى	01	00.26
		المجموع	393	100.0

3-2	هل طرق التحقق من صدق الأداة كافية؟	نعم	83	33.33
		لا	166	66.67
		المجموع	249	100.0
4-2	هل عينة الصدق ممثلة لمجتمع الدراسة؟ (باستثناء صدق المحكمين وصدق المحتوى)	نعم	77	46.39
		لا	00	00.00
		لم يحدد	89	53.61
		المجموع	166	100.0
5-2	هل حجم عينة الصدق كاف؟ (باستثناء صدق المحكمين وصدق المحتوى)	نعم	102	61.45
		لا	13	07.83
		لم يحدد	51	30.72
		المجموع	166	100.0
1-1-2	هل تم تحليل المحتوى بالوصف التفصيلي أو توصيف الأهداف؟	نعم	04	28.57
		لا	10	71.43
		المجموع	14	100.0
2-1-2	هل تم عمل جدول مواصفات؟	نعم	03	21.43
		لا	11	78.57
		المجموع	14	100.0
3-1-2	هل تم حساب نسبة الإتفاق بين المحكمين؟	نعم	05	35.71
		لا	09	64.29
		المجموع	14	100.0
4-1-2	هل الطريقة مناسبة للغرض من الأداة؟	نعم	09	64.29
		لا	05	35.71
		المجموع	14	100.0
1-2-2	صدق المحك	تلازمي	11	100.0
		تنبؤي	00	00.00
		المجموع	11	100.0
2-2-2	نوع المحك	مقياس	05	45.45
		تقدير	01	09.10
		تحصيل دراسي	05	45.45
		المجموع	11	100.0
3-2-2	هل المحك متعلق بالسمة المقاسة؟	نعم	11	100.0
		لا	00	00.00
		المجموع	11	100.0

4-2-2	هل المحك صادق؟	نعم	00	00.00
		لا	00	00.00
		لم يحدد	11	100.0
		المجموع	11	100.0
5-2-2	هل المحك ثابت؟	نعم	00	00.00
		لا	00	00.00
		لم يحدد	11	100.0
		المجموع	11	100.0
6-2-2	هل الطريقة مناسبة للغرض من الأداة؟	نعم	11	100.0
		لا	00	00.00
		المجموع	11	100.0
7-2-2	ماهو المستوى المعتمد لمعامل الصدق؟	مرتفع (أكثر من 0.70)	06	54.55
		متوسط (0.30-0.69)	05	45.45
		منخفض (أقل من 0.30)	00	00.00
		المجموع	11	100.0
1-3-2	هل تم التعريف بالإطار النظري للسمة المقاسة؟	نعم	10	07.94
		لا	116	92.06
		المجموع	126	100.0
2-3-2	هل اشتقت الفرضيات من الإطار النظري للسمة المقاسة؟	نعم	10	07.94
		لا	116	92.06
		المجموع	126	100.0
3-3-2	ماهي طرق التحقق من صدق التكوين الفرضي للأداة؟	الإتساق الداخلي	72	49.32
		المقارنة الطرفية	56	38.36
		التحليل العاملي	07	04.79
		طريقة أخرى	10	06.85
		لم يحدد الطريقة	01	00.68
		المجموع	146	100.0
4-3-2	هل طريقة الإتساق الداخلي مناسبة؟	نعم	67	93.06
		لا	05	06.94
		المجموع	72	100.0
4-3-2	هل طريقة المقارنة الطرفية مناسبة؟	نعم	56	100.0
		لا	00	00.00
		المجموع	56	100.0

4-3-2	هل طريقة التحليل العاملي مناسبة؟	نعم	07	100.0
		لا	00	00.0
		المجموع	07	100.0
4-3-2	هل الطرق الأخرى مناسبة؟	نعم	10	100.0
		لا	00	00.00
		المجموع	10	100.0
4-3-2	هل طرق التحقق من صدق التكوين الفرضي للأداة مناسبة؟	نعم	140	96.55
		لا	05	03.45
		المجموع	145	100.0
5-3-2	هل طرق التحقق من صدق التكوين الفرضي للأداة كافية؟	نعم	124	99.20
		لا	01	00.80
		المجموع	125	100.0
1-3	هل تم التحقق من ثبات الأداة؟	نعم	228	65.90
		لا	118	34.10
		المجموع	346	100.0
2-3	ماهي الطرق المستخدمة في التحقق من ثبات الأداة؟	إعادة التطبيق	76	27.34
		التجزئة النصفية	79	28.41
		ألفا كرونباخ	110	39.57
		كيودر ريتشارسون	02	00.72
		ثبات التقدير	08	02.88
		الصور المتكافئة	00	00.00
		طريقة أخرى	03	01.08
		المجموع	278	100.0
1-3-3	هل طريقة إعادة التطبيق مناسبة؟	نعم	68	89.47
		لا	08	10.53
		المجموع	76	100.0
2-3-3	هل طريقة التجزئة النصفية مناسبة؟	نعم	66	83.54
		لا	13	16.46
		المجموع	79	100.0
3-3-3	هل طريقة ألفا كرونباخ مناسبة؟	نعم	85	77.27
		لا	25	22.73
		المجموع	110	100.0

4-3-3	هل طريقة كيودر ريتشارسون 20 مناسبة؟	نعم	01	100.0
		لا	00	00.00
		المجموع	01	100.0
5-3-3	هل طريقة كيودر ريتشارسون مناسبة؟	نعم	01	100.0
		لا	00	00.00
		المجموع	01	100.0
6-3-3	هل طريقة ثبات التقدير مناسبة؟	نعم	08	100.0
		لا	00	00.00
		المجموع	08	100.0
7-3-3	هل الطرق الأخرى مناسبة؟	نعم	02	66.67
		لا	01	33.33
		المجموع	03	100.0
8-3-3	هل الطرق المستخدمة في التحقق من ثبات الأداة مناسبة؟	نعم	231	83.09
		لا	47	16.91
		المجموع	278	100.0
4-3	هل عينة الثبات ممثلة لمجتمع الدراسة؟ (باستثناء ثبات التقدير)	نعم	100	45.45
		لا	02	00.91
		لم يحدد	118	53.64
		المجموع	220	100.0
5-3	هل حجم عينة الثبات كاف؟ (باستثناء ثبات التقدير)	نعم	148	67.27
		لا	23	10.46
		لم يحدد	49	22.27
		المجموع	220	100.0
6-3	ماهو المستوى المعتمد لمعامل الثبات؟	مرتفع جدا (أكثر من 0.95)	17	06.41
		مرتفع (0.85-0.95)	89	33.58
		متوسط (0.75-0.84)	84	31.70
		منخفض (0.65-0.74)	51	19.25
		منخفض جدا (أقل من 0.65)	24	09.06
		المجموع	265	100.0