الثبات في الاختبارات محكية المرجع: يعرف الثبات في الاختبار على انه اتساق أداء المفحوصين عبر التطبيقات المتكررة *** للاختبار نفسه أو الصيغ المتكافئة له عند تطبيق أي اختبار على الأفراد أنفسهم في ظروف مشابهة يفضل الحصول على النتائج نفسها ، ويطلق على هذا الاتساق في درجات الاختبار اسم الثبات. كانت وما زالت قضية عدم ثبات الاختبارات خاصة الاختبارات التربوية والنفسية محط اهتمام مطوري الاختبار ومستخدميه فالاختبارات النفسية جميعها تمتاز بأنها غير ثابتة وهذا يعود الى ان الدرجات التي نحصل عليها تحت الظروف الاختبارية عرضة لخطأ القياس حيث تصنف أخطاء القياس بأنها اما أخطاء عشوائية أو منتظمة كما يشير نموذج الدرجة الحقيقية التقليدي (النظرية الكلاسيكية في القياس) وأساس نموذج سبيرمان يقوم على أن أي بمعنى انه يمكن تصور (X= T+ E) درجة ملاحظة مكونة من عنصرين افتراضيين الدرجة الحقيقية وعنصر الخطأ العشوائي درجة المفحوص الملاحظة في الاختبار على أنها قيمة محتملة لقيم عديدة لمجموعة احتمالات غير معروفة وهذا ما يسمى الدرجة الحقيقية كما يمكن اعتبار درجة المفحوص الحقيقية هي متوسط الدرجات الملاحظة التي حصل عليها من تطبيق الاختبار نفسه لعدد لا نهائي من المرات . اما خطأ القياس فهو الفرق بين درجة المفحوص الملاحظة ودرجته الحقيقية ويعتبر الخطأ متغير بحيث يكون متوسط توزيع الخطأ .(T .و المقدار الثابت للمفحوص (X) عشوائي إذ يعرف بأنه الفرق بين المتغير العشوائي للمفحوص هو القيمة المتوقعة ***خصائص الدرجات الحقيقية ودرجات الخطأ : اهم افتراضات النظرية الكلاسيكية بما يتعلق بالثبات : ـ أن متوسط درجات الخطأ لأي مجتمع من المفحوصين = صفر . ـ أن الارتباط بين الدرجات الحقيقية ودرجات الخطأ للمفحوصين = صفر. ـ عندما يتقدم المفحوصين لاختبارين متوازيين فإن الارتباط بين درجات الخطأ للاختبارين= 0 .

*** دليل الثبات ومعامل الثبات يعبر معامل الارتباط عن درجة العلاقة بين الدرجات الحقيقية والملاحظة وهو ما يسمى معامل الثبات . و يمكن التعبير عن معامل الثبات بأنه نسبة الانحراف المعياري للدرجات الحقيقية إلى الانحراف المعياري لدرجات الخطأ . معامل الثبات يمكن تحديده رياضياً على أنه نسبة تباين الدرجة الحقيقية إلى تباين الدرجة الملاحظة أو أنه مربع معامل الثبات.

** ثبات الدرجة الكلية(المركب): عرف المركب على أنه درجة كلية تعتمد على درجات اثنين أو أكثر من الاختبارات الفرعية و يتم تحديد ثبات المركب من خلال الخصائص الإحصائية للعناصر الفرعية، وسيتم طرح طريقتين يمكن بواسطتهما التعبير عن ثبات المركب بوساطة الخصائص الإحصائية لمكوناته: اولا : طريقة سبيرمان براون: وهذه الطريقة تسمح بتقدير ثبات مركب اختبارات متوازية عندما يكون ثبات إحداها معروفاً. ثانيا: طريقة كرونباخ ألفا: وهذه تسمح بتقدير ثبات المركب من معرفة تباين درجات المركب والتباين المشترك عبر عناصره أو مكوناته جميعها . ثبات المركب يمكن التعبير عنه كمعامل لثبات أحد مكوناته على افتراض أن مكوناتها جميعها متوازية ● ثبات المركب بمعامل ألفا يعرف معامل ألفا بانه الصيغة التي يتم اشتقاقها من ثبات المركب والتي يعبر عنها بصورة تباين درجات فقرات الاختبار والتباينات المشتركة للاختبارات التي تؤلف بمجموعها المركب . وبما انه لايمكن تحقيق التوازي للاختبارات الفرعية في المواقف الاختبارية الواقعية نلجأ لاستخدام مجموع التباينات المصاحبة وتباين المكونات لتقدير الحد الأدنى لثبات المركب يسمح الثبات بإعطاء وصف لنسبة التباين الحقيقي لدرجات الاختبار الملاحظة لمجموعة ما. وفي نظرية القياس التقليدية يمكنا وصف التباين المتوقع للدرجات الملاحظة لكل درجة حقيقية للمفحوص. ومن يشير الثبات إلى اتساق أداء (x2 O2 + O2 =E x2) خلالها نستطيع اشتقاق الخطأ المعياري للقياس باستخدام العلاقة التالية المفحوصين عبر التطبيقات المتكررة للاختبار نفسه أو الصيغ المتكافئة له، ومن المصادر الاساسية لعدم الثبات على الاختبار هو وجود أخطاء القياس العشوائية . وتزودنا النظرية الكلاسيكية في القياس بأساس نظري لتطوير دراسات تمكنا من التحقق من ثبات الاختبارات، وتعتمد هذه النظرية على مفهوم أن درجة المفحوص المشاهدة في الاختبار عبارة عن الدرجة الحقيقية والخطأ العشوائي في القياس، و باستخدام هذه النظرية يمكن تعريف معامل الثبات على أنه الارتباط بين القياسات المتوازية وهذا المعامل يساوي نسبة التباين الملاحظ الذي يعود إلى تباين الدرجات الحقيقية. يرتبط عدم ثبات درجات الاختبار بشكل رئيسي بأخطاء القياس، والتي تنقسم إلى نوعين: الأخطاء العشوائية والأخطاء المنتظمة. وتعرف الأخطاء العشوائية: بانها أخطاء غير منتظمة وعشوائية تؤدي إلى تباين غير متوقع في نتائج الاختبارات وتؤثر على تماسك وموثوقية النتائج . فمثلًا يمكن حدوث أخطاء نتيجة تصحيح أوراق بشكل غير منتظم أو أخطاء عشوائية في أداء المفحوص في كل اختبار مما يسبب فقدان الثبات. اما الأخطاء المنتظمة: فتتعلق بمشكلات غير عشوائية تؤثر على دقة الاختبار وتؤدي إلى اختلافات ثابتة نسبياً في النتائج بين التطبيقات وتقلل من دقة القياس بشكل عام . فإن سبب عدم الثبات الرئيسي هو تأثير أخطاء القياس التي تنقسم إلى عشوائية ومنتظمة، حيث تؤدي عدم السيطرة على هذين النوعين من الأخطاء إلى تفاوت النتائج في الاختبار. ولا تتغير بشكل عشوائي بين مرتي التطبيق. عليه، فإن

Summarized by © lakhasly.com

الأخطاء المنتظمة تخلق انحرافات ثابتة في نتائج الاختبارات، وتؤثر سلبًا على دقتها وموثوقيتها، وتفسيرها لأنها تعكس بشكل غير عادل أو غير دقيق الدرجة الحقيقية لمفحوصيهم. *** طرق حساب الثبات: يوجد عدة طرق لحساب الثبات مع الأخذ بعين الاعتبار أن معامل الثبات لايمكن تحديده بالضبط ولكن يمكن تقديره من العينة مع التركيز على ضرورة تحديد مصادر الخطأ في القياس لتفسير الدرجات بشكل مفيد . تقسم طرق حساب الثبات الى مجموعتين: ؟ طرق تتطلب تطبيق اختبارين ؟ طرق تتطلب تطبيق اختبار واحد. ـ طرق تتطلب تطبيق اختبارين 1- طريقة الاختبارات المتكافئة: تتطلب بناء صيغتي اختبار متوازيتين وتطبيق كلا الصيغتين على نفس المجموعة من المفحوصين خلال فترة زمنية قصيرة . ويُحسب معامل الارتباط بين مجموعتي الدرجات باستخدام معادلة معامل ارتباط بيرسون، ويُطلق عليه اسم معامل التكافؤ. 2 ـ طريقة إعادة الاختبار: يتم تطبيق الاختبار على مجموعة من المفحوصين، ثم يُعاد تطبيقه على نفس المجموعة بعد فترة زمنية، ويُطلق عليه اسم معامل الاستقرار. 3- طريقة اعادة الاختبار بوجود صيغ متبادلة: يتم تطبيق الصيغة الاولى للاختبار ثم تطبيق الصيغة الثانية بعد فترة من الزمن بحيث يكون ترتيب التطبيق لنصف المجموعة الاولى عكس الترتيب للنصف الثاني، ويطلق على معامل الارتباط معامل الاستقرار والتكافؤ. ـ طرق تتطلب تطبيق واحد للاختبار: 1- طرائق الاتساق الداخلي: ايجاد دوال لمعاملات الارتباط بين درجات الاجزاء الفرعية لنفس الاختبار بشكل منفصل ، ويتم تحديد مدى تطابق أداء المفحوصين عبر المجموعات الفرعية من الفقرات، بحيث يشير الاداء المتسق الى وجود تجانس بين فقرات الاختبار وبالتالي يمكن تعميم النتائج. 2- طريقة التجزئة النصفية: يتم تطبيق صيغة اختبارية واحدة على المفحوصين، ثم تُجزأ إلى مجموعتين فرعيتين متوازيتين قدر الإمكان، ويُحسب معامل الارتباط بين درجات المجموعتين بعد تصحيح كل نصف بشكل مستقل . ويوجد (4) طرق لتجزئة الاختبار: ـ اختيار الفقرات الفردية او الزوجية ـ ترتيب الفقرات حسب صعوبتها. ـ التعيين العشوائي للفقرات على كلا النصفين. ـ المزاوجة بين فقرات جزئي الاختبار. ويتم استخدام معادلة سبيرمان- براون لحساب معامل الارتباط، وفي حال كان هناك فرق واضح بين معامل الثبات بالطريقة النصفية ومعامل الثبات الكلي ويتم استخدام معادلة سبيرمان-براون التصحيحية للتقليل من الخلل في تقدير الثبات بين الدرجات الفرعية كما يمكن استخدام معادلة رولون او جوتمان لتقدير الثبات ؟؟؟+1 /؟؟؟2 = ؟؟؟ بين نصفي الاختبار والدرجة الكلية للاختبار الدرجة على ((A حيث D= A-B بديلا لمعادلة سبيرمان-براون والتي تعتمد على الفرق بين الدرجات بين نصفي الاختبار ويمكن استخدامه للفقرات ثنائية التصحيح أو (2X؟ /2D؟) -1 = ؟؟؟ .الدرجة على النصف الثاني (B) النصف الاول و تستخدمان للدرجات : ((KR21 و(KR20) الدرجات ذات المدى الواسع من الدرجات ـ صيغتي كوردر ـ ريتشاردسون المصممة ثنائيًا، وهما مكافئتان لمعامل ألفا في ظروف معينة ـ طريقة هويت: تعتمد على تحليل التباين بمعالجة الأفراد والفقرات ثبات تقديرات -4 Inter rater Reliability . على أنها مصادر للتباين، وتؤدي إلى نتائج مماثلة لتلك الناتجة عن معامل ألفا المقدرين يستخدم معامل إمكانية التعميم من خلال حساب معامل ثبات تقديرات المقدرين، ** العوامل المؤثرة في معاملات الثبات ـ تجانس المجموعة: يعتمد معامل الثبات على تباين الأفراد في درجاتهم الحقيقية والدرجات الخطأ، لذا فإن تجانس المجموعة يعد مهمًا في تطوير الاختبار وفي اختيار الاختبار ـ حدود الزمن: يؤثرفي مدى انتظام المفحوص في أدائه على صيغ الاختبار جميعها في الاختبارات المحددة بوقت. لذا فإن هذا المدى يصبح جزءًا من تباين الدرجة الحقيقية. ـ طول الاختبار: يؤثر طول الاختبار على تباين الدرجة الحقيقية والملاحظة، فالاختبار الأطول أكثر ثباتًا من الاختبار الأقصر ** تقدير الدرجات الحقيقية من خلال قيمة ((Y يمكن الحصول على تقدير للدرجة الحقيقية لمفحوص معين من خلال معادلة الانحدار العامة للتنبؤ بقيمة والتي تعتمد على تباين الدرجات الحقيقية على المتغيرين يمكن اختيار الطريقة المناسبة لقياس الثبات بناءً على .(X) معروفة الهدف من استخدام الاختبار من خلال النظر إلى طبيعة البيانات والنتائج المرغوبة، حيث يُفضَّل اختيار الطريقة التي تتوافق مع طبيعة الاستخدام وهدف الدراسة مع التركيز على تقليل تأثير أخطاء القياس وتحسين التكرارية والدقة. لذايُنصح بتحديد مصادر أخطاء القياس المهمة واختيار الطريقة التي تسمح بتقييم أثرها بشكل فعال، مع مراعاة الهدف المحدد من الاختبار. ** مصادر اخطاء القياس أهم مصادر الأخطاء التي تؤثر على ثبات درجات الاختبار تتعلق بالقياس والتطبيق والسياق الذي يجرى فيه الاختبار. من بين هذه المصادر: ـ أخطاء معاينة المحتوى، وهي الأخطاء الناتجة عن عدم تمثيل المحتوى بشكل دقيق أو متسق في فقرات الاختبار، مما يؤدي إلى قياس غير دقيق للأداء المرتبط بالمهارات أو المعرفة المستهدفة ـ أخطاء التطبيق والتصحيح، بالإضافة إلى الشكوك أو الأخطاء البشرية أثناء التصحيح . ـ أخطاء التخمين والتي قد تؤدي إلى استجابات غير دقيقة وغير ممثلة للمجال السلوكي، خاصة إذا كانت الاستجابات تعتمد على الحدس أو الحالة المزاجية للمفحوصين ـ أثر البيئة المحيطة أثناء

Summarized by © lakhasly.com

الأداء، حيث العوامل الخارجية مثل الضوضاء أو حالات الإضاءة قد تؤثر على أداء المفحوصين. ـ سوء تصميم الفقرات أو النصوص مما قد يقود إلى استجابات غير موضوعية أو غير ذات علاقة بالمحتوى وللحد من تأثير هذه المصادر على نتائج الاختبار وثباتها، يُنصح بما يلي: 1. تصميم فقرات الاختبار بشكل دقيق ومتسق ومراجعتها لضمان تمثيل المحتوى بشكل صحيح وخلوها من القصور التقني وتقليل العوامل المؤثرة سلباً على أداء المفحوصين 4. استخدام طرق قياس الثبات التي تتصف بالدقة لقليل أثر الاختلافات الطارئة 5. توثيق إجراءات الدراسة بشكل كافٍ مما يزيد من موثوقية درجات الاختبار وتقليل تأثير مصادر الخطأ المختلفة ** صدق الاختبار: يعتبر صدق الاختبار احد الخصائص السيكومترية الأساسية المستخدمة للتحقق من أن النتائج التي يعطيها الاختبار تعبر بشكل صحيح عن السمة المقاسة فعلاً. وتُركز الاختبارات على صدق المحتوى الذي يقيم مدى تمثيل فقرات الاختبار للنطاق السلوكي المستهدف من خلال استشارة خبراء لضمان أن الفقرات تعكس بشكل دقيق المحتوى المطلوب قياسه. تهدف هذه العملية إلى ضمان أن نتائج الاختبار يمكن الاعتماد عليها لاتخاذ قرارات صحيحة. يعرف الصدق على انه عملية جمع الأدلة لدعم نتائج درجات الاختبار او هو أن يقيس الاختبار ما وضع من اجله ويتضمن الصدق انواع عدة تتمثل في ثلاثة أنواع رئيسية هي: صدق المحتوى، والصدق المرتبط بمحك، وصدق البناء. يساعدك فهم هذه الأنواع على تقييم مدى دقة وموثوقية نتائج الاختبار وأهميتها في سياقات مختلفة. ● صدق المحتوى: يركز على مدى تمثيل عناصر الاختبار لنطاق الأداء أو المهارات التي يُقاسها، ويهتم بضمان أن فقرات الاختبار تغطي بشكل شامل ومناسب المحتوى المحدد. فهم هذا يساعدك على التأكد من أن الاختبار يقيس المدى الصحيح من المعارف أو المهارات المحددة. ● الصدق المرتبط بمحك: يقيم العلاقة بين درجات الاختبار وسلوك أو أداء خارجي يمكن قياسه بشكل مباشر. هذا النوع من الصدق يوضح مدى ارتباط النتائج بالحالة أو الأداء الفعلي في العالم الحقيقي، ● صدق البناء: يتعامل مع مدى توافق أدوات القياس مع النظرية أو البناء النظري نفسه. فهو يخبرك هل الاختبار يقيس بشكل صحيح المفهوم أو البناء الذي يُفترض أن يقيسه، ويساعدك على تقييم مدى موافقة نتائج الاختبار مع الافتراضات النظرية حول مكونات السلوك. 1- صدق المحتوى و تتضمن خطوات التحقق من صدق المحتوى : ـ توفير هيكل بنائي لعملية مطابقة الفقرات ـ وجمع وتلخيص البيانات الناتجة. كما يتطلب التخطيط لدراسة صدق المحتوى مجموعة من الاعتبارات العملية مثل موازنة الأهداف لتعكس أهميتها، وتكوين مهمة المزاوجة بين الفقرات والأهداف، وتحديد مظاهر الفقرة التي يجب فحصها. ولا يأخذ هذه النوع من الصدق أي قيم عددية للتعبير عنه. * تقييم وتلخيص النتائج في صدق المحتوى: ● تزويد الحكام بوصف واضح لخصائص الفقرة والنطاق ليأخذها بعين الاعتبار في عملية مزاوجة الفقرات مع النطاق، وتشمل الخصائص المهمة للموضوع، والعمليات العقلية، ومستوى تعقيد الأداء المطلوب، و نموذج الاستجابة المطلوب. ● تلخيص النتائج من خلال معاملات كمية مثل نسبة فقرات المزاوجة للأهداف، ونسبة فقرات المزاوجة للأهداف بتقديرات عالية الأهمية، والارتباط بين الأوزان النسبية للأهداف والفقرات التي تقيس هذه الأهداف. ● حساب معامل التوافق بين الفقرة والهدف ونسبة الأهداف التي لم تقيم بأي فقرة في الاختبار، مع الأخذ في الاعتبار خصائص هذه المعاملات عند تفسير النتائج. * اهم القضايا فيما يتعلق بتقييم صدق المحتوى: ● إحدى المشكلات الرئيسة هي عدم إمكانية المزاوجة بين الفقرات والأهداف بشكل جيد، حيث من الممكن أن لا تمثل الأهداف المكتوبة نطاق الأداء بشكل مناسب. ● يجب أن يكون الحكم على صدق المحتوى متعلقًا بمدى ملائمة فقرات الاختبار للنطاق المحدد، مع مراعاة خصائص المفحوصين الأساسية. ● يجب أخذ بيانات الأداء على الاختبار والفقرات بعين الاعتبار عند الحكم على صدق المحتوى، ويجب أن تكون الفقرات المطابقة للهدف نفسه مرتبطة على الأقل ارتباطًا متوسطًا فيما بينها. 2- الصدق المرتبط بمحك يسعى مستخدم الاختبار في معظم الحالات إلى التوصل الى النتائج من درجات الاختبار لفحص السلوك على محك أداء معين لا يمكن قياسه مباشرة من خلال الاختبار وهذا يسمى بصدق المحك. ويتضمن التصميم العام لدراسة الصدق المرتبط بمحك تحديد سلوك المحك المناسب وطريقة لقياسه، وتطبيق الاختبار، وتحديد قوة العلاقة بين درجات الاختبار والأداء على المحك. ويمكن التمييز بين نوعين من الصدق المرتبط بمحك: ▯ الصدق التنبؤي (درجة تنبؤ الاختبار بالقياس المحكي المستقبلي) ▯ الصدق التلازمي (العلاقة بين درجات الاختبار وقياس محكي تم تطبيق الاختبارين في الوقت نفسه) يواجه تطبيق الصدق المرتبط بمحك عدة مشكلات منها: تحديد المحك المناسب، وعدم كفاية حجم العينة، وعدم ثبات درجات المتنبئ أو المحك. اولا: تحديد المحك المناسب: ● يمكن اعتبار مقاييس المحك محكات فورية أو متوسطة أو قطعية، بينما المحك القطعي يتميز بأهميته الأساسية ولكنه صعب التحديد الإجرائي والقياس. ● يجب أن يكون المحك قابلاً للقياس ويتسم بالثبات ضمن حدود الوقت وكلفة الدراسة، وأن يكون له علاقة بالمحك القطعي الذي يهتم به مستخدم الاختبار. ثانيا: عدم كفاية حجم العينة ● يجب أن

Summarized by © lakhasly.com

يكون حجم العينة كافيًا، حيث إن خطأ المعاينة في العينات الصغيرة يكون كبيراً نسبياً، وقد تنخفض القدرة الإحصائية الاستدلالية المستخدمة في تحليل بيانات دراسة الصدق. ● يجب ألا تؤثر درجات الأفراد الذين قد يؤثروا على درجات المحك ضمن درجات المفحوصين على المتنبيء لتجنب تلوث المحك. ● يجب تمييز أن تغيير مدى الدرجات على المحك أو المتنبيء قد يؤدي إلى تضعيف معامل الصدق الملاحظ بسبب تقييد المدى. ثالثا: والتلويث وعدم ثبات درجات المتنبئ أو المحك ● يجب تجنب تقليص أو اختصار التباين قدر الإمكان بدلاً من الاعتماد على طرائق التصحيح الإحصائي في التغلب على مشكلة تقييد المدى. ● يجب أن يكون خطأ القياس أقل ما يمكن، وكذلك فحص ثبات كل من المتنبيء والمحك. ● يجب أن يكون الاختبار المتنبىء أكثر تجانساً في المحتوى بزيادة الاتساق الداخلي، وأن يكون ارتباط الفقرات عالياً مع المحك ولكن ليس فيما بينها. * تفسير وتحليل نتائج الصدق المرتبط بمحك ● يمكن تحليل البيانات المتجمعة من دراسة الصدق المرتبط بمحك بعدة طرائق، مثل حساب معامل ارتباط بيرسون بين درجات الاختبار وقياس المحك (معامل الصدق). ● يمكن مقارنة متوسط درجة المتنبئ للناجحين بمتوسط درجات الذين لم ينجحوا إذا كان المتغير المحك فئويًا.

Summarized by © lakhasly.com