

الفصل الثالث

الصدق

١- مفهوم الصدق

٢- الأنواع الأساسية للصدق

- صدق المحتوى (المضمون)

- الصدق المرتبط بالمحك

* الصدق التلازمى

* الصدق التنبؤى

- صدق التكوين الفرضى

٣- أنواع أخرى من الصدق

- الصدق الظاهرى

- الصدق الذاتى

- الصدق العاملى

٤- طرق حساب الصدق

- طريقة استطلاع آراء الخبراء (المحكمين)

- طريقة المحك

- طريقة المقارنة الطرفية

- طريقة المجموعات المتضادة

- طريقة مقارنة الدرجات قبل وبعد المعالجات الخاصة

- طريقة التحليل العاملى

Obaikandi.com

١ - مفهوم الصدق

يعد الصدق هو الآخر واحداً من أهم معايير جودة الاختبار، حيث يشير مصطلح Validity إلى الحقيقة Truth أو مدى الدقة Fidelity التي تقيس بها أداة القياس الشيء أو الظاهرة التي وضع لقياسها.

فالصدق لا يعنى ارتباط الاختبار بنفسه كما فى الثبات، ولكنه يعنى الارتباط بين الاختبار وبعض المحكات Criteria الخارجية Outside التي تتميز بأنها مستقلة عن الاختبار أو أداة القياس.

وقد عرف العديد من المتخصصين الصدق بطرق مختلفة، فقد ذكرت أنستازى Anastasi ١٩٩٠ م. - على سبيل المثال - أن صدق الاختبار يعنى «ما الذى يقيسه الاختبار؟ وكيفية صحة هذا القياس»، ويعرف ليندكويست Lindquist ١٩٥١ م صدق الاختبار على أنه «الدقة التي يقيس بها الاختبار ما وضع من أجله».

ومن ناحية أخرى يرى بعض علماء القياس أنه يفضل تعريف كل نوع من أنواع الصدق كل على حدة بدلاً من الكلام عن مصطلح «الصدق» كمفهوم مجرد. وبناء على ذلك أخذ تعريف الصدق يفقد أهميته، حيث بدأ الاهتمام بتعريف الصدق على أساس أنواعه (أنماطه) المختلفة. ومن الملاحظ أن كل من الجمعية الأمريكية لعلم النفس، والجمعية الأمريكية للبحوث التربوية، والمجلس القومى للقياس فى التربية قد تبناوا هذا الاتجاه منذ العام ١٩٧٤ م.

ومهما تعددت الأراء بالنسبة لمعنى ومفهوم الصدق، فإنه يمكن إيضاح مفهوم هذا المصطلح بشكل دقيق من خلال النقاط التالية:

١ - الصدق يشير إلى نتائج أداة القياس (الاختبار أو المقياس) التي تطبق على الأفراد وليس إلى الأداة ذاتها. معنى هذا أن الصدق يتعلق بنتائج

الاختبار أو المقياس وليس بالاختبار أو المقياس نفسه، فنحن نتحدث أحيانا عن صدق الاختبار أو المقياس، ولكن الأكثر ملاءمة هو أن نتحدث عن صدق نتائج الاختبار أو المقياس والتي تتعلق بشكل أكثر تحديداً بصدق التفسير الناجم عن نتائج الاختبار.

٢ - الصدق مسألة درجة، وهذا يعنى أن الصدق لا يتأسس على مبدأ الكل أو عدمه، بمعنى صدق أو لا صدق، وعليه يصبح من الملائم الإشارة إلى الصدق فى صورة تصنيفات تحدد درجته مثل : الصدق العالى أو الصدق المتوسط أو الصدق المنخفض.

٣ - الصدق يختص دائماً باستخدام معين، وعلى ذلك ينبغى عدم فهم الصدق على أنه صفة عامة لأداة القياس، فمثلا قد تحمل نتائج اختبار حركى معين درجة عالية من الصدق بالنسبة للياقة البدنية، ودرجة متوسطة من الصدق بالنسبة للتنبؤ بالنجاح فى نشاط رياضى ما، وقد لا تكون النتائج صادقة للتنبؤ بالتفكير الخططى مثلا، وعلى ذلك ينبغى عند وصف صدق الاختبار تحديد الاستخدام الخاص بالنتائج، إذ أن هذه النتائج لا تكون صادقة فى كل الأحوال بدرجة واحدة.

مما سبق يتضح أن صدق أداة القياس يشير إلى الدرجة التى تمتد إليها هذه الأداة فى قياس ما وضعت من أجله، فالاختبار أو المقياس الصادق هو الذى يقيس بدقة كافيته الظاهرة التى صمم لقياسها، بحيث لا يقيس شيئا بدلاً منها أو بالإضافة إليها.

فاستخدام اختبار ما لقياس ظاهرة غير التى صمم من أجلها إجراء خاطئ، فنحن لا نستطيع استخدام مقاييس الوزن عندما نريد قياس الطول، ولا نستطيع استخدام اختبار صمم لقياس التحمل العضلى لقياس القدرة العضلية، ولا نستطيع استخدام مقياساً للاتجاهات لقياس الذكاء.

وعلى الرغم من وضوح مفهوم الصدق، إلا أن بعض العاملين في المجال الرياضي كثيراً ما يقعون فى أخطاء تتعلق بهذا الموضوع، فقد يستخدمون بعض الاختبارات على أنها صادقة دون محاولة التأكد من ذلك، وهم يبررون ذلك بأنهم شاهدوا أو سمعوا أن اختباراً ما يمكن أن يستخدم لقياس صفة أو سمة أو قدرة معينة، مما قد ينتج عنه انتشار استخدام اختبارات غير صادقة أو ذات صدق منخفض.

وقد أجمع علماء القياس النفسى على أن الاختبار لا يكون صادقاً إلا توافرت فيه الشروط التالية:

١ - أن يكون الاختبار قادراً على قياس ما وضع لقياسه، بمعنى أن يكون الاختبار وثيق الصلة بما يقيسه.

٢ - أن يكون الاختبار قادراً على قياس ما وضع لقياسه فقط، بمعنى ألا يقيس شيئاً آخر مع ما وضع من أجله، أى يستطيع الاختبار أن يميز بين القدرة التى يقيسها والقدرات الأخرى التى يحتمل أن تختلط بها أو تتداخل معها كالقوة العضلية Muscular Strength وتحمل القوة (التحمل العضلى) Muscular endurance مثلاً .

٣ - أن يكون الاختبار قادراً على التمييز بين طرفى القدرة التى يقيسها، بمعنى أن يميز بين الأداء القوى والأداء المتوسط والأداء الضعيف، فإذا كانت الدرجات متقاربة دل ذلك على أن صدق الاختبار ضعيفاً.

٤ - أن يظهر الاختبار بوضوح الفروق الفردية بين الأفراد (وبخاصة فى حالة العينات العشوائية). ولكى تظهر هذه الفروق بشكل واضح ومميز فإنه يلزم استخدام وسائل قياس حساسة حتى تظهر هذه الفروق، فعلى سبيل المثال عند استخدام اختبار العدو ٣٠ متراً كقياس لسرعة الانتقال، فإنه يلزم تقدير الزمن باستخدام ساعات إيقاف تحسب الزمن بالثوانى (ث) و $\frac{1}{100}$ وجزء من مائة من الثانية، حتى تتضح الفروق بين المختبرين

(المفحوصين) بشكل تكون له دلالة مميزة، فقد أظهرت التجارب العملية أن حساب زمن المختبرين فى هذا الاختبار بالثوان وجزء من عشرة من الثانية فقط لا يظهر الفروق الفردية بين المختبرين على الرغم من تمتع الاختبار بالصدق.

٢ - الأنواع الأساسية للصدق

حددت الجمعية الأمريكية لعلم النفس American Psychological Association

١٩٧٤م ثلاثة أنواع أساسية من الصدق هي:

- صدق المحتوى (المضمون) Content Validity

- الصدق المرتبط بالمحك Criterion - Related Validity

- صدق التكوين الفرضى Construct Validity

وقد أقرت الجمعية الأمريكية لعلم النفس هذه الأنواع (الأنماط) الثلاثة على أساس الأغراض التي من أجلها تستخدم أدوات القياس measuring Instruments

، حيث وضعت هذه الأغراض فى ثلاث فئات عريضة هي:

١ - تمثيل محتوى خاص محدد من المجال الذى يراد اختبار المفحوص فيه، حيث يستهدف الاختبار تحديد كيف يؤدي المفحوص (فى الزمن الحاضر) عينة من المواقف Situations أو المحتويات Contents التى يدعى أن الاختبار يقيسها. فمثلا يستخدم اختبار Squat thrust (بيوربي) لأكبر عدد من المرات لتقدير المستوى الراهن للتحمل العضلى العام general muscular endurance لطلاب المدارس الثانوية. كما قد يستخدم اختبار لتهجئة فى اللغة الإنجليزية (كنوع من اختبارات التحصيل) لتقدير المستوى الراهن للتهجئة فى اللغة الإنجليزية لتلاميذ المرحلة الإعدادية.

٢ - التنبؤ بالسلوك (الأداء) المستقبلي للمفحوص بالنسبة لمتغير محدد، أو تقدير الحالة الراهنة للمفحوص بالرجوع إلى بعض الخصائص (المتغيرات) التي لا يتم قياسها مباشرة بأداة القياس . فبالنسبة لسلوك المستقبل فإنه يمكن - على سبيل المثال - إعطاء اختبار للاستعداد الدراسي Scholastic aptitude للتلاميذ في المرحلة المتوسطة من التعليم (المرحلة الإعدادية) للتنبؤ بالنجاح الأكاديمي academic success لهم في المرحلة الجامعية . أما بالنسبة للموقف الراهن Present standing للمفحوص ، فإنه يمكن تطبيق اختبار في علم الحساب (حسابي) arithmetic test على التلاميذ في المرحلة الإعدادية، حيث تمتد نتائج الاختبار في هذه الحالة لتظهر امكانية كل تلميذ في التعامل الدقيق مع الأرقام في الحياة العملية .

٣ - الاستدلال على قوة المفحوص في القدرة أو السمة أو الخاصية التي تظهرها أداة القياس ، كأن يقوم أحد الباحثين بالتعرف على درجات مجموعة من المفحوصين بالنسبة لبعض المفاهيم المجردة مثل: الذكاء والانبساط والانطواء والعصابية وما إذا كانت هذه الدرجات مرتفعة أم منخفضة .

وبناء على الأغراض الثلاثة السابقة فقد تم تحديد أنماط الصدق الثلاثة المذكورة، بحيث يطابق كل نمط منها غرضاً من تلك الأغراض الثلاثة، حيث يمكن تعريف كل نمط من هذه الأنماط على «أنه الدرجة التي تنجز بها أداة القياس الغرض المقترن بها وفقاً لطبيعة الاختبار المعطى للمفحوصين» .

ومن الطبيعي أن يحدث أن تصمم أداة ما من أدوات القياس لتلائم أكثر من غرض واحد من الأغراض الثلاثة المذكورة، حيث يقتضى الحال في مثل هذه الحالات حساب صدق مثل هذه الأداة بأكثر من طريقة من طرق الصدق . ومن ناحية تتجه بعض أنماط الصدق لتمثل أهمية كبيرة بالنسبة لأنواع محددة

من الاختبارات. فعلى سبيل المثال يلعب صدق المحتوى (المضمون) دوراً رئيسياً Key role بالنسبة لاختبار التحصيل Achievement testing ، فى حين يلعب الصدق المرتبط بالمحك هذا الدور بالنسبة لاختبار الاستعداد Aptitude testing .

وفيما يلى شرح موجز لكل نمط من أنماط الصدق الثلاثة:

صدق المحتوى (المضمون)

ويعرف باسم : الصدق المنهجي Curricular Validity والصدق التحليلي Relevance validity وصدق التمثيل Representativeness validity والصدق الداخلى (الذاتى) Intrinsic Validity . ويمثل صدق المحتوى أهمية خاصة بالنسبة للاختبارات التى تستهدف وصف أداءات الفرد على كل بعد domain من الأبعاد التى يفترض أنها تمثل الاختبار .



إذا لم تؤخذ القياسات بطريقة صحيحة، فإن عملية القياس تصبح غير صادقة وفى مجال القياسات الأثنوبوتورية تحدد النقاط التشريحية صدق المحتوى (المضمون).



اختبارات الأداء الحركى يمكن أن تستخدم لتعزيز الأداء
بالنسبة للعديد من المسابقات الرياضية

فمن المعروف أنه يوجد اختبارات تتكون من مجموعة من الوحدات items وقد وضعت كل وحدة من هذه الوحدات لقياس موضوعاً محدداً، وبحيث تشكل هذه الوحدات فى مجموعها الاختبار ككل. ولهذا السبب يهتم صدق المحتوى بتحليل المضمون relevane of the content أو الفقرات items فردياً وجمعياً، فكل فقرة أو مفردة من مفردات الاختبار يجب أن تمثل مقياساً صحيحاً ومناسباً بالنسبة للقدرة أو السمة أو المتغير الذى تسأل عنه الفقرة، وأن يتضمن الاختبار ككل فقط الفقرات التى تمثل السمات أو

المتغيرات المطلوب قياسها بواسطة الاختبار. وفي هذا الصدد تذكر أنستازى ١٩٦٨م «أن صدق المحتوى يستخدم بشكل رئيسى الفحص المنتظم لمحتوى الاختبار للتأكد من أن الاختبار يعطى عينة ممثلة من الأداءات (السلوك) المطلوب قياسها».

ويستخدم صدق المحتوى (المضمون) فى حالة بناء الاختبارات التى تستخدم لقياس التمكن من أداء المهارات الخاصة Specific skills فى مجال محدد، واختبارات الذكاء والتحصيل واستبيانات الشخصية واختبارات الاستعداد الدراسى وقوائم التوافق الذاتى الاجتماعى personal-social adjustment حيث يتطلب إعداد مثل هذه الاختبارات تحديد عدد كبير ومتنوع من الفقرات يختار منها أفضل الوحدات التى تشكل محتوى الاختبار.

ويشير صدق المحتوى (المضمون) بالنسبة للاختبارات المعرفية إلى الوحدات items التى يتضمنها الاختبار، ففى الاختبارات التحصيلية - على سبيل المثال - يستلزم بناء الاختبار تحليل المواد المراد قياسها، ثم اختيار عينات ممثلة لهذه المواد لوضعها فى الاختبار على أساس أنها ممثلة للسلوك الذى يظهر القدرة على التحصيل. معنى ذلك أن صدق المحتوى (المضمون) يتطلب التأكد من أن المفردات التى يتضمنها الاختبار (محتوى الاختبار) تعكس أو تمثل محتوى الظاهرة المستهدفة من القياس. فصدق المحتوى (المضمون) بالنسبة لاختبار يقيس المهارة فى كرة السلة يستلزم أن يتضمن الاختبار نظريا وحدات أساسية تمثل الأداء فى اللعبة مثل التصويب والمراوغة، وتنطيط الكرة، والتمرير، والتصويب، والوثب وغيرها.

وقد يظهر الاختبار أن له صدق محتوى (مضمون) عن طريق تمثيله للمجال المستهدف من القياس، إلا أن ذلك لا يؤكد نهائيا الثقة فى هذا الصدق، حيث يظهر ذلك جليا فى حالة قياس سمك طيات الجلد لتقدير نسبة الدهون فى الجسم. فمن المعروف أن البرجل المنزلق (الكالبير) يستخدم

كجهاز يتمتع بالصدق لقياس دهون الجسم، إلا أن الاستخدام الخاطئ لهذا الجهاز عن طريق أفراد غير مدربين، أو عدم اختيار المواقع الصحيحة المثلة لتراكم الدهون في الجسم، أو عدم الالتزام بالنقاط التشريحية الصحيحة للقياس قد يؤدي إلى عدم صدق النتائج.

كما سبق يتضح أن صدق المحتوى يقوم على أساس مدى تمثيل الاختبار أو المقياس لمكونات القدرة أو السمة أو الشيء أو الظاهرة التي يقيسها، حيث يكون محتوى الاختبار أو المقياس شاملاً لجميع عناصر الشيء أو القدرة التي يقيسها كشرط للصدق، معنى هذا: أن صدق المحتوى يستهدف معرفة مدى تمثيل الاختبار أو المقياس لعناصر (مكونات) السمة أو الصفة أو الظاهرة المطلوب قياسها، وعمّا إذا كان الاختبار أو المقياس يقيس جانباً محدداً من هذه الظاهرة أم يقيسها كلها.

أنواع المحكات المستخدمة للحكم على صدق المحتوى (المضمون)

تتمثل المحكات التي يمكن استخدامها لتقويم صدق المحتوى (المضمون) في المحكمين أو الخبراء الذين يسألون عن مدى صلاحية الاختبار من حيث تمثيله لمحتوى المجال المستهدف من القياس. ويعتبر التفكير الناقد والتحليل المنطقي وإعمال العقل من الأساليب المهمة التي يستخدمها الخبراء لتقويم صدق المحتوى (المضمون)، هذا بالإضافة إلى الخبرة السابقة في مجال التخصص.

وقد تكون هذه المحكات خبراء experts في الموضوع أو التخصص، أو محكمين Judges مدربين، أو زملاء في المجال Colleagues. وقد تكون هذه المحكات المراجع العلمية المتخصصة أو الكتب الدراسية أو الرسائل العلمية والتي تستخدم - بعضها أو كلها - كمصادر لتحديد المكونات الأساسية للظاهرة، ومن ثم الحكم على صدق محتوى الاختبار في ضوء تمثيله لهذه المكونات.

الصدق المرتبط بالمحك:

ويعرف هذا الصدق باسم: الصدق التجريبي experimental validity لكونه يستخدم محكا خارجيا criterion للحكم على مدى صلاحية صدق الاختبار الجديد، حيث يتم ذلك عن طريق (مقارنة Comaparing) درجات الاختبار بدرجات المحك، أو تقدير الارتباط بين درجات الاختبار ودرجات المحك.

وبناء على ذلك يمكن تعريف الصدق المرتبط بالمحك على أنه: العلاقة بين درجات الأداء على الاختبار ودرجات الأداء على بعض المقاييس الأخرى (المحكات) التي تقيس نفس الظاهرة (القدرة أو السمة أو غيرها). فالصدق المرتبط بالمحك يشير إلى مدى انتظام واتساق العلاقة بين درجات الاختبار بالنسبة لنتائج واحد أو أكثر من المحكات Criteria.

ويتم الحصول على الصدق المرتبط بالمحك عن طريق تطبيق الاختبار، ثم تطبيق المحك، ثم حساب الارتباط بين درجات الاختبار ودرجات المحك باستخدام معامل الارتباط التتابعى لبيرسون (PPM).

ويعرف الصدق المرتبط بالمحك فى بعض الأحيان باسم الصدق الإحصائى statistical validity والصدق الارتباطى correlational validity حيث يعزى السبب فى هذه التسمية إلى استخدام الارتباط لحساب العلاقة بين درجات الاختبار ودرجات المحك.

وينقسم الصدق المرتبط بالمحك إلى نوعين رئيسيين هما:

* الصدق التلازمى Concurrent validity.

* الصدق التنبؤى Predictive validity.

ويمكن التمييز بين هذين النوعين من الصدق على أساس الآتى:

أولاً : الهدف من عملية القياس:

في الصدق التلازمي يكون الهدف من عملية القياس هو تقدير الحالة الراهنة أو الحالية (الوضع الراهن)، أما في الصدق التنبؤي فيكون الهدف من عملية القياس هو التنبؤ بنتيجة معينة في المستقبل .

ثانياً : الفترة الزمنية التي يتم فيها جمع البيانات عن المحك:

في الصدق التلازمي يتم جمع بيانات المحك في نفس الوقت الذي يطبق فيه الاختبار، بمعنى تزامن تطبيق كل من الاختبار والمحك، أما في الصدق التنبؤي فيتم جمع بيانات المحك بعد فترة زمنية طويلة من تطبيق الاختبار .

الصدق التلازمي

هو نوع من أنواع الصدق التي ترتبط فيه الدرجات أو التقديرات أو النتائج التي تمثل الأداء الحالي بالنسبة لظاهرة ما بالدرجات التي يتم الحصول عليها نتيجة تطبيق الاختبار، حيث يدل معامل الارتباط المحسوب بين درجات المحك ودرجات الاختبار على الصدق التلازمي للاختبار .

فإذا كنا نريد أن نعرف كيف يستطيع أن يؤدي اللاعب مهارة ما في الوقت الراهن، فإنه يلزمنا استخدام اختبار له معامل صدق تلازمي مقبول، حيث يتم التحقق من ذلك عن طريق مقارنة درجات مجموعة من الأفراد على الاختبار المقترح بدرجات أدائهم على المحك الذي يعبر عن الأداء الأمثل بالنسبة للمهارة التي صمم الاختبار لقياسها، حيث يحدد المحك مدى تفوق هؤلاء الأفراد أو تخلفهم في أنواع الأداء الذي تتطلبه هذه المهارة، حيث يمكن عن طريق المقارنة بين درجات الاختبار ودرجات المحك تقدير الصدق التلازمي للاختبار .

ويشترط في الصدق التلازمي إعطاء الاختبار المقترح والمحك معا بالتعاقب، بمعنى أن يتزامن تطبيق الاختبار والمحك دون وجود فترات زمنية

طويلة بين التطبيقين، ويعتبر الصدق المرتبط بالمحك أكثر ملاءمة لحسب صدق العديد من الاختبارات من أهمها:

١ - اختبارات القدرات البدنية والمهارات الحركية فى الألعاب .

٢ - اختبارات الاستعدادات البدنية

٣ - الاختبارات التى تستهدف تشخيص الحالة الراهنة diagnosis of present status للمفحوصين .

٤ - اختبارات الذكاء .

ومن أمثلة المحكات التى تستخدم لحساب الصدق التلازمى فى المجال الرياضى ما يأتى:

١ - يستخدم القياس المعملى للحد الأقصى لاستهلاك الأوكسجين VO_{2max} كمحك لحساب صدق بعض اختبارات الجرى مثل: الجرى ١ ميل والجرى $1\frac{1}{4}$ ميل، الجرى ٩ دقائق، الجرى ١٢ دقيقة وغيرها من الاختبارات الميدانية، كما يستخدم نفس القياس لحساب صدق بعض الاختبارات التى تتطلب الأداء على أجهزة مثل الدراجة الأرجومترية والسير المتحرك .

ويلجأ الباحثون إلى تقنين صدق مثل هذه الاختبارات لاستخدامها كوسائل قياس غير مباشرة لتقدير الحد الأقصى لاستهلاك الأوكسجين لكونها - أى الاختبارات - أسهل فى التطبيق، ولا تستغرق وقتاً طويلاً، كما أنها أقل تكلفة بالمقارنة بالطرق المعملية المعقدة التى تستخدم لتقدير الحد الأقصى لاستهلاك الأوكسجين .

٢ - وتستخدم طريقة الكثافة (الوزن تحت الماء) كمحك لتقدير الدهون فى الجسم، حيث تستخدم قياسات طيات الجلد فى بعض مناطق الجسم باستخدام جهاز الكاليسر caliper، وبعض القياسات الأنتروبومترية كوسائل لتقدير دهون الجسم body fat بدلا من طريقة الوزن تحت الماء

كمحك والتي يصعب تطبيقها من الناحية العملية بالإضافة إلى ما تتطلبه من إمكانات وأجهزة ووقت .

٣ - ويستخدم الأداء في المباريات game performance وتقديرات الخبراء experts كمحكات لحسب الصدق التلازمي للعديد من اختبارات المهارات الرياضية Sport Skill tests .

جدول (٣ - ١)

أمثلة للصدق المرتبط بالمحك في المجال الرياضي

الاختبار test	المحك criterion	معامل الصدق validity Coefficient
- اختبار كيمب وفينست للمهارة في التنس (١٩٦٨م)	نظام الدوري الكامل	٠,٨٤ للمبتدئين* ٠,٩٣ للمتقدمين
- اختبار هوايت للسرعة في الإرسال في التنس (١٩٦٨م)	نظام الدوري الكامل	٠,٨٩ للمبتدئين
- اختبار الجري - المشي $1\frac{1}{2}$ ميل (جيتشل وكيركندال وروبنز ١٩٩٠م).	الحد الأقصى لاستهلاك الأوكسجين vo ₂ max	٠,٩٢
- اختبار الجري - المشي ١ ميل (كلين وآخرون ١٩٨٧م)	الحد الأقصى لاستهلاك الأوكسجين vo ₂ max	٠,٨٨
- نسبة الدهون باستخدام جهاز الكالير.	وزن الجسم تحت الماء	٠,٨٠ للرجال من سن ١٧-٣٧ (ريكوروينك ١٩٦٩م) ٠,٧٨ للسيدات من سن ١٨ - ٢٣ (بولك وآخرون ١٩٧٥م) ٠,٨٧ للسيدات من سن ٣٣ - ٥٥ (بولك وآخرون ١٩٧٥م).

* تم حساب الصدق باستخدام معامل ارتباط فروق الرتب لسبيرمان .

كما سبق يتضح أن الصدق التلازمى يستخدم لتصميم بعض الاختبارات لتحل محل اختبارات أخرى لها صدق معروف، تلك الاختبارات تعرف بالملك criterion. حيث يشير معامل الصدق التلازمى إلى درجة العلاقة بين درجات الاختبار ودرجات الملك الذى يوثق فيه كقياس صادق للظاهرة.

مفهوم الملك : Criterion

الملك عبارة عن مقياس خارجى external مستقل independent يقيس بشكل أساسى نفس الظاهرة التى يفترض أن يقيسها الاختبار أو المقياس المزمع تقينته. ويمكن إيضاح مفهوم الملك بشكل أكثر تفصيلاً من خلال الآتى:

- أن الملك عبارة عن معيار أو ميزان قياسى يستخدم للحكم على صدق مقاييس أخرى جديدة.

- أن الثقة فى صدق الملك تكون كبيرة جداً.

- أنه يمكن النظر إلى متغير الملك Criterion variable على أنه متغير يمكن التنبؤ به من خلال متنبئ Predictor واحد (اختبار أو مقياس) أو من خلال الربط بين مجموعة من المتنبئات (اختبارات أو مقاييس).

- قد يكون الملك مجموعة من الدرجات أو الآراء أو الأحكام أو الأداءات على بعض الأجهزة أو الأدوات أو الاختبارات.

الملك المؤلف من مجموعة من العناصر: Composite criterion

هو ملك يتكون نتيجة الربط بين الدرجات المشتقة (المحولة) من عدد من الاختبارات، حيث تستخدم درجة هذا الملك لتقدير صدق الاختبار أو بطارية الاختبار. وعندما يتم هذا الإجراء فإنه يشير إلى ما يعرف «بالمحك المركب» والذى يتم الحصول عليه عن طريق تحويل درجات الاختبارات إلى الدرجات المعيارية Standard scores.

الدرجة على المحك (درجة المحك) : Criterion score

هى الدرجة التى يحصل عليها المفحوص عندما يؤدى على المحك (عندما يطبق عليه المحك)، حيث المطلوب هو التنبؤ بهذه الدرجة من خلال الدرجات على متغيرات أخرى تعرف بالمتنبآت (وسائل التنبؤ) أو الاختبارات.

المتنبئ (وسيلة التنبؤ) : The Predictor

يشير المتنبئ إلى الاختبار أو المقياس الذى يستخدم كوسيلة للتنبؤ بالمحك، حيث يلاحظ أنه لا يمكن تحديد المتنبئ (وسيلة التنبؤ) إذا لم يتم التعبير عن المحك تعبيراً كميّاً.

المحكات الضعيفة:

تتوقف دقة المتنبئ (الاختبار المستخدم فى التنبؤ predictor) على كفاية ودقة المحك. وقد قدم بروجدن وتيلور Brogden and Taylor ١٩٥٠م أربعة أوصاف للمحكات الضعيفة هى:

- قصور (ضعف) المحك : Criterion deficiency

ويظهر هذا المحك عند تحليل المواقف التى يحدث فيها سلوك العمل أو الأداء. ويحدث القصور (الضعف) فى المحك عندما يتم إهمال أو إغفال عناصر وثيقة الصلة بالمحك، وبخاصة فى حالة المحكات التى تتألف من مجموعة من العناصر Composite criterion.

- تلوث المحك: Criterion contamination

ويحدث خلال عملية التطبيق عندما يتعرض المحك لأخطاء القياس، وبعض العوامل الخارجية غير المتوقعة.

- تحيز المحك : Criterion scale unit bias

ويحدث فى حالة عدم تساوى وحدات المقياس الذى يستخدم لقياس المحك، بمعنى استخدام مقاييس تقدير ثنائية لقياس بعض العناصر، وأخرى

ثلاثية لقياس عناصر أخرى، وغيرها عشوائية، هذا بالإضافة إلى التحيز في تقدير الأوزان النسبية لمكونات أو عناصر المحك .

- تحريف المحك : Criterion distortion

ويظهر في حالة عدم الالتزام بالوزن النسبي للعناصر الداخلة في المحك، أو في حالة عدم الدقة في تقدير هذا الوزن، فإذا كانت القوة والدقة - على سبيل المثال - مكونات رئيسية بالنسبة للتصويب بيد واحدة في كرة اليد، ثم استخدمنا محك يستهدف فقط قياس الدقة، فمعنى ذلك أن هذا المحك قد أهمل عنصر (مكون) القوة وأعطاه وزناً نسبياً يساوى (صفرًا) وأعطى الدقة في التصويب وزناً نسبياً (يساوى ١٠٠).

الصدق التنبؤي

وقد صمم كصدق عملي (اختباري) *impirical validity* أو إحصائي *Statistical Validity* حيث يشير إلى ترابط الاختبار بالمحك للتنبؤ ببعض الأحداث الممكنة في المستقبل، بمعنى قدرة الاختبار على التنبؤ بأداء الفرد في موقف مستقبلي إذا كان هذا الموقف المستقبلي له علاقة بما يقيسه الاختبار. ويستهدف الصدق التنبؤي في المجال الرياضي الآتي:

- وصف الأداء الراهن على المحك *describing present criterion* performance .

- التنبؤ بالأداء المستقبلي على المحك *predicting future criterion* performance .

وصف الأداء الراهن على المحك:

عندما يستخدم الصدق التنبؤي لوصف الأداء الراهن على المحك، فإنه يشبه في هذه الحالة الصدق التلازمي أو يكون امتداداً له، حيث يمكن تقدير

الأداء على بعض المحركات المركبة عن طريق الدرجة التي يتم الحصول عليها على متغير التنبؤ (المتنبئ) وهو متغير يمتاز بأنه بسيط يسهل تطبيقه وحساب نتائجه. فإذا لم تكن هناك علاقة بين المتغيرين (الاختبار والمحرك) فإنه لا يمكن - فى مثل هذه الحالة - توظيف الاختبار للتنبؤ بالمحرك.

فعندما توصل جيتشل وآخرون. Gatchell et all. ١٩٧٧م - على سبيل المثال - إلى وجود علاقة ارتباط مقدارها (-٩١٥, ٠) بين اختبار الجرى - المشى واحد ونصف ميل واختبار السير المتحرك treadmill test والذى يقيس بطريقة مباشرة ودقيقه الحد الأقصى لاستهلاك الأوكسجين VO_{2max} ، فقد أوصى بناء على ذلك أنه يمكن استخدام اختبار الجرى - المشى واحد ونصف ميل كمتنبئ (وسيلة تنبئ) بالحد الأقصى لاستهلاك الأوكسجين.

وبناء على العلاقة السابقة بين اختبار الجرى - المشى $\frac{1}{4}$ ميل واختبار السير المتحرك، والتي بلغت (-٩١٥, ٠)، فإنه يمكننا استخلاص الآتى:

- أن لاختبار الجرى - المشى $\frac{1}{4}$ ميل صدق تلازمى بلغ (-٩١٥, ٠).
- أنه يمكن إحلال اختبار الجرى - المشى $\frac{1}{4}$ ميل بدلاً من اختبار السير المتحرك.

ويبقى السؤال المهم وهو:

هل يمكننا التنبؤ بقيم الحد الأقصى لاستهلاك الأوكسجين من خلال درجات الأداء على اختبار الجرى - المشى واحد ونصف ميل؟

والإجابة هى: نعم وذلك عندما نقوم بإعداد معادلة الانحدار regression equation التى يمكن استخدامها لتقدير الحد الأقصى لاستهلاك الأوكسجين VO_{2max} من خلال زمن الأداء على اختبار الجرى - المشى واحد ونصف ميل، وهى المعادلة التى أعدها جيتشل ومساعدوه عام ١٩٧٧م للإناث،

وصورتها الرياضية كالتالى :

$$y = 98,3 - 1,82x \text{ س}$$

حيث إن :

$$y = \text{القيمة المقدرة للحد الأقصى لاستهلاك الأوكسجين}$$

$$x = \text{الزمن الفعلى لأداء اختبار الجرى - المشى واحد ونصف ميل (بالدقائق).}$$

فلو افترضنا أن إحدى المتسابقات قطعت مسافة ال واحد ونصف ميل فى ١٢ دقيقة، فإنه يمكن تقدير الحد الأقصى لاستهلاك الأوكسجين لهذه المتسابقة بالتعويض فى معادلة الانحدار السابقة كالتالى :

$$y = 98,3 - (1,82 \times 12)$$

$$= 98,3 - 21,84$$

$$= 76,46$$

ومن الملاحظات الجديرة بالاهتمام فى هذا الخصوص هو أن الاختبار قد يكون له معامل ارتباط كبير مع المحك، ومع ذلك لا يكون متنبأً ممتازاً، فعندما نحصل على معامل ارتباط مرتفع بين الاختبار والمحك فإنه لا يعنى بالضرورة أن الاختبار صادق فى التنبؤ بالمحك. والقاعدة المهمة فى هذا الشأن هو ضرورة التحقق من أن التحسن فى درجات الاختبار يصاحبه تحسن فى درجات المحك، حيث يؤكد ذلك صدق الاختبار كوسيلة للتنبؤ بالمحك، لأنه لا يجوز الاعتماد فقط على معامل الارتباط بين الاختبار والمحك. فإذا اقترح بعض الباحثين - على سبيل المثال - أن الاتجاهات نحو النشاط البدنى يمكن أن تستخدم للتنبؤ بالنجاح فى دروس التربية الرياضية المدرسية، فإن ذلك يستلزم بالضرورة التأكد من أن التحسن فى اتجاهات الطلاب يتبعه تحسن فى درجاتهم فى التربية الرياضية، حيث لا يجب الاعتماد فقط على معامل الارتباط بين الاختبار (المتنبئ) والمحك للتحقق من الصدق التنبؤى.

ومن ناحية أخرى فقد يقترح بعض المدربين أن الوقوف على الرأس Headstand تعد وسيلة تنبؤ جيدة بالقدرة على العدو Sprinting ability، وقد جاء هذا الاقتراح نتيجة وجود معامل ارتباط (معامل صدق) بلغ (-.85، 0) بين الوقوف على الرأس (كمتنبئ) والقدرة على العدو (كمحك). ويرى علماء القياس أنه لا يجب الأخذ فقط بمعامل الارتباط لحسم هذه العلاقة، لأنه إذا كان الوقوف على الرأس وسيلة تنبؤ حقيقية بالنسبة للقدرة على العدو، فإن تدريب الأفراد على الوقوف على الرأس يجب أن يتبعه بالضرورة تحسن في قدرتهم على العدو وهو ما لم تؤكد الدراسات والبحوث العلمية المتخصصة وذلك على الرغم من وجود ارتباط مرتفع بين المتغيرين.

فقط لاحظ المتخصصون أن أى طريقة من طرق حساب الارتباط يمكن أن تعطى معاملات ارتباط مقبولة فيما بين المتغيرات، حيث تكشف هذه المعاملات عن درجة العلاقة بين الاختبارات كمقاييس مشتركة لعنصر عام Common element، إلا أنه يلاحظ أن معاملات الارتباط (معاملات الصدق) قد لا تعطينا معلومات عن درجة تأثير متغير على آخر، فمثلا يرتبط الوزن بالطول ارتباطا عاليا، لأن الأفراد طوال القامة يميلون لكى يكونوا أثقل وزنا، إلا أن الزيادة فى الوزن قد لا تحدث بالضرورة نتيجة الزيادة فى الطول، لذا نستخلص أن العلاقة بين المتنبئ (الاختبار) والمحك هى علاقة وظيفية Functional relationship أكثر منها علاقة سببية Causal.

مما سبق نستخلص الآتى:

- ١ - أن كلا من الصدق التلازمى والصدق التنبؤى يستخدمان المحك.
- ٢ - أن معادلة الانحدار هى أهم ما يميز بين الصدق التلازمى والوظيفة الأولى من وظائف الصدق التنبؤى، فالصدق التلازمى يكتفى بتقدير العلاقة بين الاختبار والمحك، بينما يستلزم الصدق التنبؤى (الوظيفة الأولى) إعداد معادلة الانحدار كأداة تستخدم درجات الأداء على الاختبار لتقدير الأداء (المتزامن) على المحك.

٣ - أنه لا يجوز الاعتماد فقط على معادلة الانحدار لتقدير الصدق التنبؤى للاختبار، حيث ينصح بعض علماء القياس بأنه بالإضافة إلى معادلة الانحدار، يجب دراسة العلاقة السببية بين الاختبار والمحك لتأكيد الثقة في معادلة الانحدار (التنبؤ).

ولزيادة الثقة في معادلة الانحدار ينصح علماء القياس بالآتي:

أولاً : يجب أن تكون عينة التقيين كبيرة نسبياً بحيث لا تقل عن ٢٠٠ فرد، وأن يتم اختيارها من مجتمعها الأصلي بالأسلوب العشوائى.

ثانياً : يجب أن يطبق الاختبار والمحك فى وقت واحد بالتبادل (التزامن فى التطبيق كما فى حالة الصدق التلازمى)، وأنه يفضل استخدام أكثر من اختبار للمفاضلة بينهم كمتنبآت بالمحك.

ثالثاً : أن يتم حساب العلاقة بين كل متنبئ على حدة والمحك، حيث يتم اختيار الاختبار الذى يظهر معامل ارتباط أكبر مع المحك على أنه الاختبار الأكثر صدقا، يلى ذلك إعداد معادلة الانحدار.

رابعا : ينصح علماء القياس بضرورة تقويم معادلة التنبؤ prediction equation، حيث يستخدم لإعداد المعادلة عينات عشوائية مختلفة مسحوبة من مجتمع إحصائى واحد، كما أنه يمكن استخدام عينة واحدة (ن = ٢٠٠) تقسم عشوائيا إلى مجموعتين متساويتين فى الحجم، حيث تستخدم كل مجموعة من المجموعتين فى اختبار صدق معادلة التنبؤ عن طريق مقارنة البيانات المستمدة من عينات مختلفة مختارة من نفس المجتمع الإحصائى، حيث يعرف هذا الإجراء باسم: التحقيق المتبادل للصدق Cross-Validitation وهو إجراء يستخدم لتقويم معادلة التنبؤ وقدرتها الوظيفية على التمييز.

التنبؤ بالأداء المستقبلي على المحك،

عندما يستهدف الصدق التنبؤى وصف المحك أو التنبؤ المستقبلي به، فإن ذلك يقتضى استخدام متنبئ أو أكثر للتنبؤ بالمحك عند مرحلة زمنية محددة، حيث يلاحظ أن درجات الاختبار (المتنبئ) ودرجات المحك لا يتم الحصول عليها فى نفس الفترة الزمنية، وإنما يطبق الاختبار (المتنبئ) أولاً، ثم يطبق المحك بعد مرور فترة زمنية طويلة من تطبيق الاختبار.

ويعرف الصدق التنبؤى بأنه قدرة الاختبار على التنبؤ ببعض المتغيرات بطريقة منتظمة، فالصدق التنبؤى يستهدف تقويم أداة القياس فى ضوء النتائج المستقبلية التى تتنبأ بها. معنى ذلك أن الصدق التنبؤى يستلزم تطبيق الاختبار أولاً، ثم نتابع سلوك المفحوص فيما بعد بالنسبة لبعض المتغيرات التى يتنبأ بها الاختبار. فإذا اتفق مستوى إنجاز المفحوص على متغيرات المحك ومستوى إنجازها على الاختبار، دل ذلك على أن الاختبار صادق، حيث تعرف هذه الطريقة باسم الطريقة التتبعية the fellow up method لحساب الصدق.

فإذا كانت اللياقة البدنية أساسية للنجاح الدراسى بكليات التربية الرياضية (كما ثبت بالخبرة والملاحظة)، فإن نتائج اختبار اللياقة البدنية الذى يطبق على الطلاب الراغبين فى الالتحاق بكليات التربية الرياضية يمكن أن يكون مؤشراً لانتقاء هؤلاء الطلاب، على أساس أن الطلاب المتفوقين فى اللياقة البدنية تكون أمامهم احتمالات أكبر من غير المتفوقين للنجاح الدراسى فى كليات التربية الرياضية فى المستقبل بعد أن يلتحقوا بها.

وإذا كان الصدق التنبؤى يدل على قدرة الاختبار فى التنبؤ بنتيجة معينة فى المستقبل، فإنه يستلزم بالتالى المقارنة بين درجات الأفراد على الاختبار وبين درجاتهم على محك يدل على أدائهم فى المستقبل، حيث يعبر الارتباط بين درجات الاختبار ودرجات هذا المحك على مدى قدرة الاختبار فى التنبؤ بنتائج المحك، حيث يتم ذلك وفقاً لمجموعة من العمليات والإجراءات الإحصائية.

ويذكر كيرتون Cureton ١٩٦٥م أن تقدير الصدق التنبؤى للاختبار يتم بحساب معامل الارتباط بين درجات الاختبار ودرجات المحك، ويشير مارشال وهاليس Marshal and Hales ١٩٧٢م أن معامل الصدق التنبؤى هو عبارة عن الارتباط التتابعى بطريق بيرسون بين درجات الأداء على الاختبار ودرجات الأداء على محك مناسب، حيث يدل الارتباط المسحوب على مؤشر معامل الصدق the index of validity coefficient .

ويستخدم الصدق التنبؤى لتقنين الاختبارات التنبؤية Predictive tests التى تستهدف بشكل خاص قياس القدرات والاستعدادات سواء كانت بدنية أو مهارية أو عقلية أو نفسية، كما يستخدم بالنسبة لاختبارات التحصيل الحركى أو المهارى لمحاولة التنبؤ بالقيام ببعض الأعمال أو الوظائف أو مظاهر السلوك فى المستقبل .

وقد أقر علماء القياس أن الصدق التنبؤى للاختبار يتحقق من خلال الارتباط بين درجات الاختبار ودرجات المحك، حيث توظف درجات الأداء على الاختبار للتنبؤ بالأداء على المحك بعد فترة زمنية طويلة نسبيا، وذلك من خلال استخدام ما يعرف بمعادلة التنبؤ Prediction equation .

ماهية معادلة التنبؤ:

هى معادلة انحدار a regresion equation تستخدم للتنبؤ بالأداء فى المستقبل بمتغير يعرف بمتغير المحك Criterion Variable ويرمز له بالرمز (y - ى) وذلك من خلال الأداء على واحد أو أكثر من الاختبارات التى تعرف بالمتنبئ predictor ويرمز له بالرمز (س - X) وغالبا ما يتم إعداد معادلة التنبؤ عن طريق البحوث الارتباطية .

مثال (٣ - ١):

أراد أحد الباحثين تقويم معلمى التربية الرياضية وفقا لفعاليتهم وكفاءتهم فى العمل ووفقا لعدد من متغيرات الشخصية، وقد وجد الباحث بعد تطبيق

مجموعة من الاختبارات والقوائم، وبعد تحليل العلاقات فيما بين هذه المجموعة من المقاييس، أن أفضل وسيلة للتنبؤ بالفعالية والكفاية في العمل (ي - y) هو متغير الثقة بالنفس (س - X)، حيث تمكن الباحث من صياغة معادلة التنبؤ التالية:

$$Y = 0.3 + 2.6X$$

فباستخدام هذه المعادلة يمكن التنبؤ بدرجة الكفاية المهنية للمعلم الذي يحصل على درجة تساوى ١٤ على مقياس الثقة بالنفس باستخدام معادلة التنبؤ السابقة كالتالى:

$$Y = 0.3 + 2.6(14)$$

$$= 0.3 + 36.4$$

$$= 36.7$$

ويرى علماء القياس أنه يمكن النظر إلى معامل الارتباط الذى يصف درجة العلاقة الخطية بين الاختبار والمحك، على أساس أن المحك (ي - y) متغير تابع dependent variable، وأن الاختبارات (المتنبآت) متغيرات مستقلة Independent Variables، وأنه بناء على ذلك فإنه يمكن تقدير الارتباط بين المحك وأفضل المتغيرات التى يمكن الربط فيما بينها كمتنبآت بالمحك.

الارتباط المتعدد : Multiple Correlation

هو معامل ارتباط يصف درجة العلاقة الخطية التى يتم بها التنبؤ بالمحك عن طريق الاختبارات (المتنبآت).

الانحدار المتعدد : Multiple Regression

ويشير إلى العلاقة بين المحك كمتغير تابع (ي - y) والترابط الخطى لاثنين أو أكثر من المتغيرات المستقلة (س_١، س_٢، ... إلخ)، حيث يتم وصف هذه العلاقة عن طريقة معادلة الانحدار المتعدد أو ما يعرف بمعادلة

التنبؤ المتعدد Multiphe Prediction equation ، والصورة الرياضية لهذه المعادلة كالتالى :

$$Y = b_0 + b_1S_1 + b_2S_2 + \dots$$

حيث إن :

Y = المحك (المتغير التابع)

S = متغيرات مستقلة (S₁، S₂، .. إلخ)

b = الوزن الذى يعطى للمتغيرات المستقلة، وهى معاملات الانحدار المتعدد.

فإذا تمكن الباحث (فى المثال الخاص بتقويم معلمى التربية الرياضية) من تحديد ثلاثة متغيرات مستقلة يمكن استخدامها كأفضل المتنبآت بالكفاية المهنية لمعلمى التربية الرياضية، وكانت هذه المتغيرات هى :

S₁ (الثقة بالنفس)

S₂ (القدرة على التفاعل مع الآراء والاتجاهات المختلفة)

S₃ (قوة الإرادة)

وكانت معادلة التنبؤ المتعدد باستخدام هذه المتغيرات الثلاثة كالتالى :

$$Y = 28,2 + 1,3S_1 + 0,7S_2 + 1,6S_3$$

فإذا كانت درجات أحد المعلمين بالنسبة S₁ = 8، S₂ = 13،

S₃ = 11، فإنه يمكن التنبؤ بالكفاية المهنية لهذا المعلم عن طريق تطبيق

المعادلة السابقة كالتالى :

$$Y = 28,2 + 1,3(8) + 0,7(13) + 1,6(11)$$

$$= 28,2 + 10,4 + 9,1 + 17,6$$

$$= 65,3$$

وقد بدأ علماء النفس الرياضى فى السنوات الأخيرة تقويم البروفيل النفسى Psychological Profiles للاعبين المتميزين وغير المتميزين كوسيلة مهمة فى التنبؤ بالتفوق فى المجال الرياضى، حيث توصلوا إلى أنه يمكن استخدام العديد من العوامل النفسية والبيولوجية لدراسة الخصائص المرتبطة بالتفوق فى رياضة المستويات العليا high-level توطئة لإعداد البروفيل الخاص بكل نشاط، واستخدام مثل هذا البروفيل فى التنبؤ والانتقاء.

أمثلة لبعض التنبآت (وسائل التنبؤ) فى المجال الرياضى:

١ - تستخدم الحالة الغذائية الراهنة Present diet، والممارسة الرياضية exercise behaviors، وضغط الدم blood pressure، والتاريخ المرضى للأسرة كوسائل تنبؤ (متنبآت - Predicitors) للكشف عن أمراض القلب heart disease المتوقع أن يصاب بها الفرد فى المستقبل.

٢ - تستخدم نتائج الدرجات فى المرحلة الجامعية الأولى (البكالوريوس أو الليسانس)، ونتائج الامتحانات التأهيلية (الأهلية) Qualifying examination كوسائل تنبؤ (تنبآت) بالنجاح فى الدراسات العليا فى التخصصات المختلفة.

٣ - تستخدم القدرات البدنية Physical abilities والقدرات المعرفية Cognitive abilities كوسائل تنبؤ (متنبآت) للأداء الناجح فى العمل.

صدق التكوين الفرضى

أول من قدم مصطلح «صدق التكوين الفرضى» هى الجمعية الأمريكية لعلم النفس عام ١٩٥٤م، وكان ذلك ضمن التوصيات الفنية والعلمية التى أقرت بها الجمعية بخصوص تقدير صدق بعض أدوات القياس. ويتميز صدق التكوين الفرضى بأنه أكثر تعقيداً وصعوبة من كل من صدق المحتوى والصدق المرتبط بالمحك، ومن ثم يلاحظ أن بعض الباحثين يتهبون من استخدام هذا النوع من أنواع الصدق، فى حين يضطر بعض الباحثين إلى استخدام هذا

الصدق فى الحالات التى يجدون فيها أن كلا من صدق المحتوى والصدق المرتبط بالمدك غير مناسبين لتقنين أدواتهم.

ويعرف صدق التكوين الفرضى بمسميات أخرى منها الصدق العاملى Factorial Validity وصدق السمة trait validity، وتحدد أنستازى ١٩٩٠م، مفهوم صدق التكوين الفرضى على أنه « المدى الذى يمكن أن يمتد إليه الاختبار كمقياس لمفهوم افتراضى hypothetical Construct أو شىء نظرى theoretical Construct أو سمة trait من السمات».

ما هو المقصود بالتكوين الفرضى؟

التكوين الفرضى Construct عبارة عن شىء ينشأ على سبيل الافتراض، فى مجال التربية الرياضية - على سبيل المثال - يمكن أن نطلق مصطلح التكوينات Constucts على المهارات أو القدرات أو السمات التى نفترض أنها تشكل فى مجموعها اختباراً واحداً يقيس ظاهرة معينة أو خاصية مميزة، فعندما نقوم بتصميم اختبار ما لقياس ظاهرة معينة أو خاصية مميزة فإننا نفترض أن هناك بعض التكوينات Constructs (قدرات أو مهارات أو سمات) سوف تقيس الخاصية أو الظاهرة ككل.

فعلى سبيل المثال قد يرى الباحث أن التكوينات الفرضية للقدرة المهارية فى كرة السلة تتكون من المحاوره بالكرة والتصويب والتمرير، فإذا كان الاختبار المعد لقياس هذه القدرة يشتمل على وحدات لقياس كل تكوين من التكوينات الفرضية السابقة (مهارات المحاوره والتصويب والتمرير) فإن هذا يدل على أن الاختبار صادق ككل فى قياس القدرة المهارية فى كرة السلة، وذلك على أساس أن هناك وحدة اختبار لكل مهارة من المهارات الثلاث السابقة، وأن الربط بين هذه الوحدات يشكل اختباراً صادقاً لقياس القدرة المهارية فى كرة السلة.

وفى مجال قياس بعض القدرات والسمات النفسية المرتبطة بالنشاط الرياضى يشير صدق التكوين الفرضى إلى مدى قياس الاختبار لتكوينات فرضية محددة، من هذه التكوينات الفرضية - على سبيل المثال - الذكاء

الخططى والاتجاهات نحو النشاط البدنى والعدوان والقلق وغيرها. هذه التكوينات الفرضية عبارة عن مفاهيم نظرية تعبر عن سلوك معين أو مظاهر لخاصية مميزة، حيث يتطلب هذا النوع من أنواع الصدق فهما دقيقا لهذا السلوك أو المظاهر المميزة المطلوب قياسها، وأن مجموع الوحدات أو العبارات التى يتضمنها المقياس أو الاختبار سوف تقيس هذا التكوين الفرضى (القدرة أو السمة النفسية) ككل .

ويشير علماء القياس النفسى على أن صدق التكوين الفرضى هو أعلى صيغ وأشكال الصدق لكونه يربط بين كل من الصدق المنطقى والصدق الإحصائى، ويقررون أن هذا النمط من أنماط الصدق يستخدم تحديداً عندما نريد تفسير الأداء على الاختبار وفقاً لبعض المفاهيم والمصطلحات النفسية .

ويرى هؤلاء العلماء أن صدق التكوين الفرضى عبارة عن «صفة نفسية - Psychological quality» نفترض أنها تمتد لكى تشرح لنا بعض مظاهر السلوك .

فالقدرة على التفكير Reasoning ability والذكاء Intelligence، والاتجاهات attitudes والتفكير الناقد Critical thinking، والاستعداد الرياضى mathematical aptitude، والاستيعاب القرائى Reading Comprehension جميعها تكوينات فرضية Constructs تتعلق بالسلوك المعرفى للمفحوص .

فالقدرة على التفكير كتكوين فرضى، يتم تفسير درجات الاختبار الذى يقيسها على أساس أنه يدلنا ضمناً على الصفة التى يمكن أن نسميها «القدرة على التفكير»، وأن هذه القدرة يمكن تقديرها وفقاً لدرجات الأداء على بعض الاختبارات .

ومن الملاحظات المهمة هو أن كل تكوين Construct من التكوينات السابقة لديه نظرية يستتر خلفها، هذه النظرية يمكن استدعاؤها لظهور وتقبلها لوصف سلوك المفحوص والتنبؤ به . فإذا قلنا - على سبيل المثال -

أن المفحوص مرتفع الذكاء فإننا نستطيع أن نعرف على الفور ماهية مظاهر السلوك التي نتوقعها من هذا المفحوص في المواقف الخاصة.

مما سبق يتضح أن «مصطلح التكوينات» عبارة عن ضرب من المفاهيم a sort of concepts التي يتم اقتراحها بشكل نظامي يتضمن التعريف بها بوضوح وبحيث تكون مؤسسة على بيانات ومعلومات تعتمد على الملاحظة أو التجربة. وفي هذا الصدد تشير نانلى Nunnally ١٩٧٠م إلى أن «التكوينات» عبارة عن فروض توضح لنا أن تنوعات في السلوك سوف يرتبط كل منها بالآخر في الدراسات التي تتناول الفروق الفردية Individual differences في الإنسان، وفي ضوء ذلك يمكن تعريف صدق التكوين الفرضي على أنه:

«المدى الذى تظهر به الدرجات على الاختبار متسقة مع تكوينات فرضية محددة»

أو هو:

«المدى الذى يستطيع فيه الأداء على الاختبار أن يمثل تكوينات نفسية محددة».

وتشتمل عمليات تحديد صدق التكوين الفرضي على الخطوات التالية بالترتيب:

١ - تحديد التكوينات Constructs التى يفترض أن يقيسها الأداء على الاختبار.

٢ - اشتقاق الفروض الخاصة بالأداء على الاختبار بالنسبة لكل تكوين، ويتم ذلك من خلال المفهوم النظرى لكل تكوين وعلى أساس النظرية التى يستتر خلفها الأداء.

٣ - التأكد من صحة الفروض عن طريق التفكير المنطقى والخبرة السابقة والتجربة العملية، ولنفترض - على سبيل المثال - أننا أردنا أن نفحص

أحد الاختبارات الجديدة التي أعدت لقياس الذكاء، فإنه يمكن من خلال دراستنا ومعرفتنا عن الذكاء وضع التنبؤات التالية:

- الدرجات الخام على الاختبار سوف تزيد بزيادة العمر الزمني (لقد أكدت الدراسات والبحوث العلمية أن الذكاء يزداد مع زيادة العمر حتى سن ١٦ سنة تقريباً).

- درجات الاختبار سوف تتنبأ بالنجاح في التحصيل المدرسي school achievement (وهي علاقة تنبؤية أظهرتها نتائج الدراسات والبحوث التربوية والنفسية على مدى سنوات طويلة).

- درجات الاختبار سوف ترتبط إيجابياً مع التقديرات التي يضعها معلم الفصل لذكاء التلاميذ.

- درجات الاختبار سوف تميز بين مجموعات التلاميذ المتفوقين وغير المتفوقين دراسياً في المدارس.

- درجات الاختبار سوف لا تتأثر بدرجة جوهرية بالتدريس الذي يحدث في المدرسة.

٤ - يتم اختبار كل تنبؤ من التنبؤات السابقة واحداً بعد الآخر، فإذا جاءت نتائج جميع التنبؤات مقبولة، تبين لنا أن الاختبار الجديد اختباراً مقبولاً لقياس الذكاء. أما إذا جاءت إحدى التنبؤات بنتائج سلبية (عكس المعترف به) - كأن لا تزداد درجات الاختبار مع تقدم العمر الزمني - فإننا نستخلص أن الاختبار إما أن يكون غير صادق لقياس الذكاء، أو أن هناك خطأ ما في التكوينات النظرية.

طرق الحصول على صدق التكوين الفرضي:

يعتمد صدق التكوين الفرضي على الاستدلالات المنطقية Logical inferences التي يتم استخلاصها من أنماط متنوعة ومتعددة من البيانات. وقد لوحظ أنه

لا توجد طريقة واحدة مناسبة لبناء أو إعداد صدق التكوين الفرضى كما هو متبع بالنسبة لصدق المحتوى (المضمون) أو الصدق المرتبط بالمشك. وقد قدم كرونباخ وميل Cronbach and Meehl ١٩٥٥م مجموعة من الإجراءات لحساب صدق التكوين الفرضى هي:

١ - تحليل العمليات العقلية لإعداد فقرات الاختبار:

وتتضمن هذه الطريقة تحليل العمليات العقلية المطلوب قياسها والتي يجب أن تتضمنها وحدات الاختبار، حيث يتضمن هذا التحليل الكشف عن العوامل المهمة الواجب مراعاتها فى عملية القياس.

٢ - مقارنة درجات مجموعات معروفة:

فى بعض الحالات يمكن التنبؤ بأن الدرجات سوف تختلف من مجموعة لأخرى من الأفراد. فدرجات مجموعات الأفراد يمكن أن تختلف بالنسبة لاختبار ما وفقاً للعمر الزمنى للمجموعات age groups ووفقاً للنوع (بنين أو بنات)، ووفقاً للحالة التدريبية (مجموعات مدربة وأخرى غير مدربة)، ووفقاً لدرجة التكيف الاجتماعى (مجموعات تتمتع بدرجات عالية من التكيف وأخرى تتمتع بدرجات أقل)، فعلى سبيل المثال تزداد معظم القدرات (البدنية والعقلية) مع تقدم العمر الزمنى، وأن البنين يحصلون على درجات أعلى من الدرجات التى تحصل عليها البنات فى اختبارات محددة مثل اختبار الاستيعاب الميكانيكى Mechanical Comprehension وهو اختبار يقيس الفهم الآلى، فى حين يحصل البنات على درجات أكبر من الدرجات التى يحصل عليها البنين فى بعض الاختبارات مثل اختبار الاستيعاب القرائى Reading Comprehension.

وأيضاً ولنفس الأسباب، فإنه يتوقع أن تميز درجات بعض الاختبارات بين المجموعات وفقاً للاختلاف فى درجات بعض المتغيرات مثل التدريب

training، والتكيف adjustment وغيرهما. فالتنبؤ بدرجات اختبار محدد يمكن استخدامه لتقويم صدق هذا الاختبار على أساس قدرته على التمييز بين مجموعات من المعروف أنها تختلف فيما بينها فى السمة النفسية أو العقلية التى يقيسها الاختبار.

٣ - مقارنة الدرجات قبل وبعد إجراء بعض المعالجات الخاصة (انظر طرق حساب الصدق).

٤ - الارتباطات باختبارات أخرى (انظر طرق حساب الصدق).

الإجراءات الارتباطية لحساب صدق التكوين الفرضى:

لعل من أهم الصعوبات التى تواجه الباحثين عند بناء صدق التكوين الفرضى للاختبار هو عدم وجود اختبار واحد مفرد يمكن الثقة فيه كمحك قياسى لتقنين صدق الاختبار، وكما ذكرنا من قبل فإنه لا توجد طريقة واحدة تكفى لبناء صدق التكوين الفرضى للاختبار. وقد قدم كرونباخ ١٩٧١م شرحاً للإجراءات الارتباطية Correlational Procedures التى تستخدم لوصف صدق التكوين الفرضى للاختبار موضحاً أنه يلزم لتحديد هذا الصدق القيام بالآتى بالترتيب.

أولاً: حساب الصدق الداخلى للاختبار internal validity: وتتم هذه الخطوة على أساس افتراض أن الاختبار الذى يقيس السمة له خصائص داخلية محددة مثل الاتساق الداخلى Internal Consistency، وتعدد أبعاد Multidimensionality الاختبار.

ثانياً: حساب الصدق الخارجى للاختبار external validity: ويستخدم لتأكيد صدق الاختبار، لكون الصدق الداخلى لا يكفى للثقة فى صدق التكوين الفرضى للاختبار.

وقد استخدم كينيون Kenyon ١٩٦٨م هذا الإجراء عند إعدادة لقائمة الاتجاهات نحو النشاط البدنى (ATPA) على أساس افتراض أن السمة تتكون

من أكثر من بعد واحد (متعددة الأبعاد)، ثم قام بصياغة بنية (تركيب) افتراضية hypothesized structure للاتجاهات نحو النشاط البدني، ثم قام باختبارها باستخدام التحليل العاملي، حيث كشفت النتائج عن ستة أبعاد dimensions ممثلة للسمة المقيسة. وقد ضمنها قائمة صممت لقياس كل بعد على حدة، حيث ينظر إلى هذا الإجراء على أنه يمثل الصدق الداخلي للقائمة.

وبالإضافة إلى التحليل العاملي - كأسلوب لحساب الصدق الداخلي لها - فقد قام كينيون بالإضافة إلى ذلك بحساب العلاقة بين التكوينات Constructs وبعض المحركات الخارجية بغية التحقق من الصدق الخارجي للقائمة. فقد قام كينيوي بصياغة الفروض المتعلقة بخصائص الأفراد الذين لهم درجات منخفضة على القائمة بالمقارنة بالأفراد الذين لهم درجات مرتفعة.

ومن ناحية أخرى فقد قام كينيون بحساب العلاقة بين درجات البعد المتعلق «بالصحة واللياقة the health and fitness» بمستوى كفاءة الجهازين الدوري والتنفسي كما يقيسها اختبار بالك the Balke treadmill test، وقد افترض كينيون أن العلاقة بين درجات هذا البعد سوف تكون مرتفعة مع نتائج اختبار بالك، وقد جاءت النتائج لتثبت صحة هذه الفرضية وتؤكد الصدق الخارجي لهذا البعد من أبعاد القائمة.

مفهوم الصدق التقاربي وصدق التمييز:

يوضح كامبل وفيسك Campbell and Fiske ١٩٥٩م، أن تحقيق كل من الصدق التقاربي Convergent validation وصدق التمييز Discriminant Validation، يعد من الإجراءات المهمة لبناء صدق التكوين الفرضي لأي اختبار من الاختبارات، وأن من العوامل الأساسية بالنسبة لصدق التكوين الفرضي هو أن يرتبط الاختبار مع غيره من المقاييس التي يجب أن يرتبط بها،

وألا يرتبط بمقاييس أخرى لا يجب أن يرتبط بها. فعلى سبيل المثال يجب أن يرتبط اختبار للقابلية العددية numerical aptitude test باختبار للاستدلال الحسابي arithmetical reasoning test أو اختبار للجبر algebra test، في حين يجب ألا يرتبط نفس الاختبار مع اختبار للتاريخ أو اختبار للجغرافيا، فعندما يكون للاختبار ارتباطات بالسلوك المتوقع أنه يقيسه فإن هذه العملية تعرف بـ «تحقيق الصدق التقاربي»، وأنه عندما يكون للاختبار ارتباطات ضعيفة مع مقاييس لا يجب أن يرتبط بها لكونها تختلف عما يتوقع أنه يقيسه، فإنه يطلق على هذه العملية اسم «تحقيق صدق التمييز» حيث يعد الارتباط المنخفض أحد أهم مظاهر صدق التمييز.

كما سبق يتضح أن الصدق التقاربي يعكس درجة العلاقة بين المقياس ومقاييس أخرى تقيس نفس التكوين the same Construct، وأن صدق التمييز يعكس درجة العلاقة بين المقياس ومقاييس أخرى لا تقيس نفس التكوين.

كما يتضح أن الصدق التقاربي للاختبار يتحقق عندما تكون هناك علاقات ارتباط جوهريّة بين الاختبار ومقاييس أخرى تقيس نفس التكوين، كما أن صدق التمييز يتحقق من خلال علاقات الارتباط المنخفضة بين الاختبار ومقاييس أخرى عندما تختلف السمة أو التكوين. وبناء على ذلك يمكننا القول بأن:

«الصدق التقاربي يعنى تقارب الاختبار مع مقاييس أخرى تقيس نفس السمة أو التكوين وأن صدق التمييز يعنى اختلاف الاختبار مع نفس المقاييس عندما تختلف السمة أو التكوين.

ومن أمثلة التجارب التي استخدمت لحساب الصدق التقاربي في المجال الرياضي تلك الإجراءات التي اتبعها راينر مارتنز Reiner Martens عند إعداده لاختبار قلق المنافسة الرياضية (Sport Competition Anxiety Test (SCAT).

فقد قام مارتنز بمقارنة نتائج اختبار (SCAT) بنتائج مقياس للقلق العام general anxiety ، حيث قام بفحص الصدق التقاربي لاختباره على أساس الفرضين التاليين :

١ - أن معاملات الارتباط بين اختبار قلق المنافسة الرياضية وأى من اختبارات القلق العام سوف لا تكون مرتفعه جداً، وذلك لكون مقياس قلق المنافسة الرياضية قد أعد لدراسة الفروق الفردية فى سمة القلق المتعلق بالمنافسات الرياضية، ومن ثم فإنه لا يصلح أن يستخدم كمقياس للقلق العام.

٢ - أن معاملات الارتباط بين اختبار قلق المنافسة الرياضية وأى من اختبارات القلق العام سوف لا تكون صفرية، وأنه يتوقع أن تظهر هذه المعاملات بقيم متوسطة (من $+٠,٣$ إلى $+٠,٥$) وذلك لوجود عوامل مشتركة بين اختبار قلق المنافسة الرياضية وأى اختبارات أخرى للقلق العام.

وعندما قام مارتنز بحساب معاملات الارتباط بين اختباراه واختبارات أخرى للقلق العام على عينات من الذكور والإناث، تراوحت معاملات الارتباط المحسوبة ما بين ($٠,٣٦$ و $٠,٤٦$) وهى معاملات ارتباط تؤكد صحة الفرضين الأول والثانى، وتؤكد بالتالى الصدق التقاربي لاختبار قلق المنافسة الرياضية الذى أعده مارتنز.

وللتحقق من صدق التمييز لاختبار قلق المنافسة الرياضية، قام مارتنز بدراسة العلاقة بين اختباراه لقلق المنافسة الرياضية ودافعية الانجاز Achievement motivation توطئة لاختبار الفرضية التالية:

«أن اختبار قلق المنافسة الرياضية سوف لا يظهر أية علاقات ارتباط مع دافعية الانجاز، انطلاقاً من أن اختبار القلق يقيس فقط قوة الدافع فى تجنب الفشل Failure ، وأنه لا يقيس قوة الدافع فى النجاح فى الإنجاز، وأنه بناء على ذلك فإنه يتوقع أن الأفراد الذين يحصلون على درجات منخفضة فى دافعية الإنجاز سوف يحصلون على درجات مرتفعة فى الخوف من الفشل

Fear of failure ويحصلون على درجات مرتفعة في الدافع نحو النجاح في الإنجاز the motive to achieve success .

وقد توصل مارتنز إلى صحة الفرضية السابقة، حيث كشفت نتائج الدراسات التي قام بها عن معاملات ارتباط بين دافعية الإنجاز وقلق المنافسة الرياضية تراوحت من (٠,٠٢ إلى ٠,١٥) لكل من الذكور والإناث مما يؤكد صدق التمييز لاختبار قلق المنافسة الرياضية.

كيفية تحقيق الصدق التقاربي وصدق التمييز:

«قدم كامبل وفيسك ١٩٥٩م» تصميماً تجريبياً خاصاً لدراسة كل من الصدق التقاربي وصدق التمييز، هذا التصميم يشير إلى ما يعرف باسم: المصفوفة متعددة الطرق ومتعددة السمات.

Multitrait - Multimethod Matrix (MTMM)

ووفقاً لما هو موضح من الاسم، فإن استخدام هذه المصفوفة يتطلب سمتين traits أو أكثر، يتم قياسها بطريقتين methods أو أكثر، حيث يتم تحليل معاملات الارتباط البينية intercorrelations داخل المصفوفة لتحقيق كل من الصدق التقاربي وصدق التمييز للاختبار وفقاً لمبدئين مهمين هما:

١ - أن قياس نفس السمة باستخدام طرق مختلفة سوف يؤدي إلى وجود معاملات ارتباط مرتفعة فيما بين المقاييس المستخدمة، مما يؤكد الصدق التقاربي.

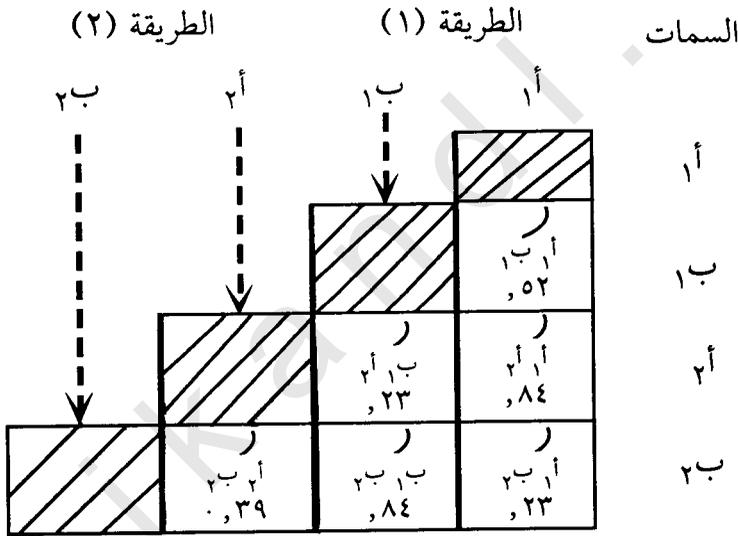
٢ - أن استخدام نفس الطرق لقياس سمات مختلفة سوف يؤدي إلى وجود معاملات ارتباط منخفضة فيما بين هذه المقاييس، مما يؤكد صدق التمييز، حيث يستخلص من هذين المبدئين الآتي:

«أن اتفاق السمة واختلاف الطريقة يؤدي إلى معاملات ارتباط مرتفعة» وأن «اختلاف السمة واتفاق الطريقة يؤدي إلى معاملات ارتباط منخفضة».

والشكل (٣ - ١) يعرض نموذجاً لتصميم تجريبي مأخوذاً عن «كامبل وفيسك ١٩٥٩م» لتحقيق الصدق التقاربي وصدق التمييز، حيث يمكن استخدام هذا التصميم وفق المثال التالي:

لنفرض أننا أعطينا الرمزين (أ، ب) لسمتين نفسييتين هما:

(أ) العدوان Agression، (ب) العصاوية Neuroticism، وأنه قد تم قياس هاتين السمتين بطريقتين أعطيت الأرقام (١، ٢)، وكانت هاتين الطريقتين (١) قائمة لقياس الشخصية، (٢) اختبار إسقاطي Projective test ووفقاً لهذه المعطيات فإنه يمكن وضع التصميم التالي:



شكل (٣ - ١)

المصفوفة متعددة الطرق ومتعددة السمات

ويلاحظ من الشكل (٣ - ١) الآتي:

- (أ) تدل على درجة العدوان على قائمة الشخصية
- (أ) تدل على درجة العدوان على الاختبار الإسقاطي
- (ب) تدل على درجة العصاوية على قائمة الشخصية
- (ب) تدل على درجة العصاوية على الاختبار الإسقاطي

وعندما حسبت معاملات الارتباط بين هذه السمات جاءت كالتالى :

أ_١ب_١ = ٠,٥٢ . (العدوان على قائمة الشخصية مع العصاوية على نفس القائمة).

أ_١أ_٢ = ٠,٨٤ . (العدوان على قائمة الشخصية مع العدوان على الاختبار الإسقاطى).

أ_١ب_٢ = ٠,٢٣ . (العدوان على قائمة الشخصية مع العصاوية على الاختبار الإسقاطى).

ب_١أ_٢ = ٠,٢٣ . (العصاوية على قائمة الشخصية مع العدوان على الاختبار الإسقاطى).

ب_١ب_٢ = ٠,٨٤ . (العصاوية على قائمة الشخصية مع العصاوية على الاختبار الإسقاطى).

أ_٢ب_٢ = ٠,٣٩ . (العدوان على الاختبار الإسقاطى مع العصاوية على الاختبار الإسقاطى).

من معاملات الارتباط السابقة يمكن استخلاص الآتى:

أولاً: أن معاملات الارتباط الأكبر فى الخلايا تدل على الصدق التقاربى للاختبار، حيث تتمثل هذه المعاملات فى الآتى :

أ_١أ_٢ = ٠,٨٤ . (ويمثل درجة العدوان على قائمة الشخصية مع درجته على الاختبار الإسقاطى).

ب_١ب_٢ = ٠,٨٤ (ويمثل درجة العصائية على قائمة الشخصية مع درجتها على الاختبار الإسقاطي).

وهذه النتائج تؤكد الصدق التقاربي للاختبار انطلاقاً من أن اتفاق السمة واختلاف الطريقة يعطى معاملات ارتباط مرتفعة بين الاختبارات.

ثانياً : أن معاملات الارتباط الأقل في الخلايا تدل على صدق التمييز للاختبار، حيث تتمثل هذه المعاملات في الآتى :

أ_١ب_١ = ٠,٥٢ (ويمثل درجة العدوان على قائمة الشخصية مع درجة العصائية على نفس القائمة).

أ_٢ب_٢ = ٠,٣٩ (ويمثل درجة العدوان على الاختبار الإسقاطي مع درجة العصائية على نفس الاختبار).

وهذه النتائج تؤكد صدق التمييز انطلاقاً من أن اختلاف السمة واتفاق الطريقة يعطى معاملات ارتباط منخفضة بين الاختبارات.

ثالثاً : أنه يلزم أن تكون معاملات الارتباط التى تدل على الصدق التقاربي أكبر من معاملات الارتباط التى تدل على صدق التمييز، حيث يلاحظ أن :

أ_١أ_٢ = ٠,٨٤ ، ب_١ب_٢ = ٠,٨٤ وهى أكبر من

أ_١ب_١ = ٠,٥٢ ، أ_٢ب_٢ = ٠,٣٩

٣ - أنواع أخرى من الصدق

بالإضافة إلى الأنواع الأساسية للصدق والتي (سبق الإشارة إليها) توجد ثلاثة أنواع أخرى هي:

الصدق الظاهري Face Validity

يحدث في كثير من الأحيان أن يخلط بعض الباحثين بين صدق المحتوى (المضمون) والصدق الظاهري، فالصدق الظاهري ليس صدقاً حقيقياً بالمعنى العلمي لكلمة «الصدق»، ولكنه يعنى ببساطة أن الاختبار أو المقياس يبدو صادقاً في صورته الظاهرية.

ويتأسس الصدق الظاهري على نقطتين مهمتين هما:

١ - مدى مناسبة الاختبار لما يقيسه .

٢ - مدى مناسبة الاختبار لمن يطبق عليهم .

ويتم تقويم هذا الصدق من قبل المتخصصين في المجال الذي ينتمى إليه الاختبار، حيث يطلب منهم قراءة الاختبار، ثم الحكم على الصدق من خلال الآتى:

١ - فحوى بنوده ومدى علاقتها بما تقيسه منطقياً .

٢ - طبيعة الفقرة - ووضوحها .

٣ - تعليمات الاختبار .

٤ - الزمن المخصص للاختبار . . إلخ .

كما سبق يتضح أن الصدق الظاهري يشير فقط إلى المظهر الخارجى للاختبار Appearance، حيث يتم فحص وحدات Items الاختبار من قبل بعض المتخصصين، لإقرار مدى صلاحية هذا الاختبار .

ومن الملاحظات الجديرة بالاهتمام فى هذا الخصوص أن الصدق الظاهري إجراء إضافي يمكن استخدامه بالنسبة لأنواع الصدق الأساسية (الأنواع الثلاثة) كمدخل مهم لتأكيد صلاحية بعض وسائل القياس، ومع ذلك فإنه لا يجب أن يكون بديلاً عن الأنواع الثلاثة الأساسية السابقة .

ويستهدف الصدق الظاهري إعداد تقرير علمي يعطى للباحث الثقة والطمأنينة في مدى صلاحية الاختبار قبل قيامه بإجراءات التقنين العملية (التجريبية) للاختبار، فعندما يقترح الخبراء استبعاد بعض فقرات الاختبار التي تبدو غير ذات صلة بالموضوع الذي تقيسه Irrelevant فإن ذلك يوفر على الباحث جهداً كبيراً ويزيد من الثقة في صدق الاختبار.

ويلعب الصدق الظاهري دوراً مهماً بالنسبة للمفحوصين، فإذا كان الاختبار يبدو في نظر المفحوصين أنه غير ذي صلة بالظاهرة المطلوب قياسها، فإنهم قد يقومون بأداء الاختبار بصورة غير جدية أو بصورة يغلب عليها الاستهتار لكونهم يعتقدون أن نتائجه غير مجدية.

الصدق الذاتي Intrinsic Validity

ويقصد به الصدق الداخلي للاختبار، وهو عبارة عن مؤشر الثبات Index of reliability أو الارتباط بين الدرجات الملاحظة على الاختبار والدرجات الحقيقية (الافتراضية). فالصدق الذاتي للاختبار عبارة عن الدرجات التجريبية للاختبار منسوبة للدرجات الحقيقية الخالية من أخطاء القياس.

وبما أن الثبات يقوم في جوهره على الدرجات الحقيقية للاختبار إذا أعيد تطبيقه على نفس مجموعة الأفراد أى عدد من المرات، لذا نجد أن الصلة بين الثبات والصدق صلة وثيقة.

ويقاس الصدق الذاتي بحساب الجذر التربيعي لمعامل ثبات الاختبار، حيث يمكن حسابه كالتالي:

$$\text{معامل الصدق الذاتي} = \sqrt{\text{معامل الثبات}}$$

فإذا كان معامل ثبات أحد الاختبارات يساوي ٠,٦٥، حينئذ يصبح معامل الصدق الذاتي لهذا الاختبار هو:

$$= ٠,٦٥$$

إذاً معامل الصدق الذاتي = ٠,٨١

ونود أن نلفت النظر إلى أن الصدق الذاتى المحسوب من معامل الثبات يحدد النهاية العظمى لمعاملات الصدق التجريبية للاختبارات المستخدمة، فإذا كان معامل ثبات اختبار تمرير كرة السلة على الحائط wall-pass لمدة ٢٥ ثانية يساوى ٠,٨٤ ، فإن الصدق الذاتى لهذا الاختبار يكون:

$$0,92 = \sqrt{0,84} =$$

وهذا يعنى أن معامل الصدق التجريبى لاختبار تمرير كرة السلة على الحائط لمدة ٢٥ ثانية يجب أن يساوى أو يقل عن ٠,٩٢ ، وأنه فى أغلب الأحيان لا يصل إلى ٠,٩٢ إلا من الناحية النظرية فقط .

الصدق العاملى Factorial Validity

وهو نوع من الصدق يعتمد على منهج التحليل العاملى Factorial analysis ، كأسلوب إحصائى يستهدف التحليل متعدد المتغيرات multivariate analysis بغرض وصف العلاقات فيما بين مجموعة كبيرة من المتغيرات وسوف نقدم مزيداً من التفاصيل عن هذا النوع من أنواع الصدق عند الحديث عن التحليل العاملى فى إصدار لاحق بإذن الله تعالى .

حساسية الاختبار Sensitivity

تعد حساسية الاختبار من الشروط المهمة التى يجب مراعاتها فى أدوات القياس ، فقد تكون أداة القياس صادقة فى قياس مما وضعت لقياسه ولكنها لا تكون حساسة .

ويقصد بحساسية المقياس أو الاختبار مدى مناسبه لما يقيس ، بالرغم من كونه يتمتع فى الأصل بالصدق ، فالمسطرة المدرجة تتمتع بالصدق كأداة لقياس المسافة ، ولكنها تصبح غير حساسة لقياس المسافة بين القاهرة والإسكندرية مثلا ، واختبارات الاستعدادات البدنية التى تستخدم لانتقاء الموهبين رياضيا من الأطفال قد لا تكون حساسة لقياس الاستعدادات البدنية للأطفال العاديين

إنطلاقاً من أن مثل هذه الاختبارات لا تتمتع بالقدرة على إظهار الفروق الفردية في الاستعدادات البدنية للأطفال غير الرياضيين (العاديين).

فعندما تعجز أداة القياس عن إظهار الفروق فيما نقيسه، فإنه يمكن أن نصف هذه الأداة بأنها غير حساسة لما نقيسه. فعندما يحاول شخص ما أن يقيس مقدار ما فقده من وزن جسمه نتيجة انتظامه في برنامج لإنقاص الوزن، فإن وحدات الوزن التي تستخدم الكيلوجرامات أو الأرتال (الميزان الطبلية) قد لا تكون حساسة لإظهار الفروق في وزن الجسم قبل وبعد الانتظام في البرنامج، وذلك لكون هذا الفرق قد يكون محدوداً بجرامات قليلة لا يظهرها ميزان الطبلية لعدم حساسيته لهذا الفرق.

وتمثل حساسية أدوات القياس أحد الشروط المهمة التي يجب مراعاتها عند اختيار هذه الأدوات، وذلك لما لهذا الشرط من أهمية بالنسبة لإظهار مقدار التقدم والتحسين في بعض المجالات المهمة المتعلقة بالتعلم وتذكر (استرجاع) المعلومات، واللياقة الوظيفية والبدنية، والقياسات المورفولوجية، وإنقاص الوزن وغيرها.

ومن ناحية أخرى فقد تلعب حساسية أدوات القياس دوراً مهماً بالنسبة لاستشارة دافعية وحماس جماهير الناس عندما يشعرون بنجاحهم وتحسين مستوياتهم، فعندما نستخدم لقياس الأداء مقاييس عامة gross measures غير مناسبة لإظهار الفروق داخل الفرد أو بين الأفراد، فإنه يتوقع أن تفشل مثل هذه المقاييس في إظهار تلك الفروق، أو ربما تظهرها بشكل محدود جداً. فعندما يطبق اختبار تمرير كرة السلة على دائرة واحدة مرسومة على حائط (بحيث تعطى الدرجة (١) عند إصابة الهدف، والدرجة (صفر) إذا لم تصب الكرة الهدف).

فإن هذا الاختبار يكون غير حساس لإظهار التقدم الذي يحققه لاعب كرة السلة من ناحية، والفروق بين هذا اللاعب وغيره من اللاعبين من ناحية

أخرى، وذلك في حالة إعطاء الدرجة (١) عندما تصيب الكرة الهدف،
والدرجة (صفر) إذا خرجت الكرة بعيداً عن الهدف.

ويكون نفس الاختبار السابق حساساً لقياس الدقة والسرعة في التمرير في
كرة السلة إذا استخدمت مجموعة من الدوائر المركزية مختلفة الأقطار
والدرجات، بحيث تمنح الدرجة الأقل كلما تحركت الكرة بعيداً عن المركز.

وقد تشير حساسية الاختبار إلى استجابته أو قابليته للتغير Responsivity
to change، بمعنى قدرته على إظهار الفروق في الدرجات كلما تحسن
مستوى المفحوص، وأيضاً قدرته على إظهار الفروق بين الأفراد من ذوى
القدرات المختلفة، حيث يعنى ذلك أن الاختبار قد صمم بطريقة تسمح له بأن
يظهر التقدم والتحسين الذى يحدث للمفحوص فى مدى واسع من الدرجات،
فقياس زمن العدو ٣٠ متراً للاعبى المنتجب القومى فى العدو باستخدام
ساعات إيقاف تسجل الزمن بالثوانى فقط، يعد قياساً غير حساس لإظهار
التقدم الذى يحدث لهؤلاء اللاعبين، وأيضاً غير حساس لإظهار الفروق بينهم
فى هذا الاختبار، ومن ثم يلزم فى مثل هذه الحالة استخدام ساعات إيقاف
تقوم بتسجيل الزمن بالثوانى و $\frac{1}{10}$ من الثانية.

خريطة الصدق Validity Chart

يوضح الشكل (٣ - ٢) رسماً تخطيطياً لأنماط (أنواع) الصدق،
ومجالات القياس المستهدفة، والأساليب الإحصائية التى يمكن استخدامها
لحساب الصدق، والمحكات المختارة لتقدير كل نمط من أنماط الصدق.

٤ - طرق حساب الصدق

يستخدم لحساب الصدق عدة طرق، حيث يلاحظ أن كل هذه الطرق ليست صالحة بالنسبة لكل أنواع الاختبارات. وفيما يلي عرض لهذه الطرق مع تقديم شرح موجز لكل منها:

طريقة استطلاع آراء الخبراء (المحكمين) Expert's Judgement

يتم اختيار الخبراء وفقاً لشروط يحددها الباحث والمشرّف ويمكن أن تتضمن المؤهل الدراسي، التخصص الدقيق، عدد سنوات الخبرة في التخصص... إلى غير ذلك من الشروط، وفي جميع الحالات يجب أن يراعى بالنسبة للخبراء الآتي:

- أن يكونوا من الدارسين ذوي الخبرة في مجال التخصص الذي يعد له الاختبار أو المقياس.
- ألا يقل عددهم عن ٣٠ خبيراً وكلما زاد هذا العدد كان ذلك أفضل.
- توضع قائمة بأسماء ووظائف هؤلاء الخبراء بمرفقات البحث أو الدراسة.

متى تستخدم هذه الطريقة؟

تستخدم هذه الطريقة لحساب صدق المحتوى (المضمون) والصدق الظاهري معاً، حيث يطلب من الخبير (المحكم - المتخصص) إبداء رأيه في مكونات الاختبار أو المقياس، والتي يرى أنها تشكل من وجهة نظره أهمية خاصة بالنسبة للظاهرة المقيسة، حيث يعتمد في ذلك على التفكير المنطقي أو ما يعرف بالتفكير الناقد Critical thinking والخبرة الذاتية.

٢ خطوات تطبيق هذه الطريقة:

أولاً: بالنسبة للاختبارات البدنية والحركية (الاختبارات العملية):

- تحليل القدرة أو المهارة، إلى مكوناتها الأساسية - principle components التي تشكل في مجموعها محتوى تلك القدرة أو المهارة.
- عرض هذه المكونات على الخبراء لاستطلاع آراءهم حول صدق هذه المكونات بالنسبة للقدرة أو المهارة الحركية المطلوب قياسها.
- تحليل آراء الخبراء بالنسبة لكل مكون من المكونات التي كشف عنها التحليل.

مثال (٣ - ١):

أراد أحد الباحثين تحديد المكونات الأساسية للقدرة الحركية motor ability في المجال الرياضي، توطئة لوضع بطارية اختبار لقياس هذه القدرة المركبة. ولتحقيق ذلك فقد قام الباحث بتحليل هذه القدرة معتمداً في ذلك على الدراسات السابقة وخبرته الذاتية في مجال التخصص، حيث توصل إلى تحديد ١٧ مكوناً جدول (٣ - ٢).

وللتحقق من صدق كل مكون من المكونات الـ ١٧ قام الباحث باختيار مجموعة من الخبراء تتكون من ٣٠ خبيراً من المتخصصين في القياس والتقويم في المجال الرياضي، وقد تم استطلاع آراءهم في هذه المكونات وفق الخطوات التالية:

- إعداد استمارة استطلاع رأى تتضمن المكونات الـ ١٧ موضحاً بها معنى القدرة الحركية وكل مكون من المكونات التي تقيس تلك القدرة.
- عرض الاستمارة على مجموعة الخبراء، حيث طلب من كل واحد منهم الحكم على علاقة كل مكون بالقدرة الحركية وذلك بوضع دائرة حول

الرقم الذى يتفق مع وجهة نظره وذلك وفق مقياس تقدير عشارى
Rating scale كالتالى :

١٠	٩	٨	٧	٦	٥	٤	٣	٢	١
----	---	---	---	---	---	---	---	---	---

حيث تعنى هذه المجموعة من الدرجات الآتى :

- الدرجة (١) وتعنى أن المكون لا يقيس هذه القدرة
 - الدرجة (٥) وتعنى أن المكون يقيس هذه القدرة بدرجة متوسطة.
 - الدرجة (١٠) وتعنى أن المكون يقيس هذه القدرة بدرجة كبيرة جداً.
 - الدرجة الأقل من (٥) تعنى علاقة ضعيفة بين المكون والقدرة
 - الدرجة الأكثر من (٥) تعنى علاقة قوية بين المكون والقدرة.
- تجميع الاستمارة لتصنيف آراء الخبراء بالنسبة لكل مكون على حدة وفق التدرجات من (١-١٠) ثم تحسب النسبة المئوية لكل تدرج.

جدول (٣ - ٢)

التقديرات										المكونات
١٠	٩	٨	٧	٦	٥	٤	٣	٢	١	١ - القوة العضلية
١٠	٩	٨	٧	٦	٥	٤	٣	٢	١	٢ - القدرة العضلية
١٠	٩	٨	٧	٦	٥	٤	٣	٢	١	٣ - السرعة الانتقالية
١٠	٩	٨	٧	٦	٥	٤	٣	٢	١	٤ - سرعة رد الفعل
١٠	٩	٨	٧	٦	٥	٤	٣	٢	١	٥ - الرشاقة
١٠	٩	٨	٧	٦	٥	٤	٣	٢	١	٦ - المرونة
١٠	٩	٨	٧	٦	٥	٤	٣	٢	١	٧ - التوازن
١٠	٩	٨	٧	٦	٥	٤	٣	٢	١	٨ - الدقة
١٠	٩	٨	٧	٦	٥	٤	٣	٢	١	٩ - التوافق البدني
١٠	٩	٨	٧	٦	٥	٤	٣	٢	١	١٠ - التحمل العضلي
١٠	٩	٨	٧	٦	٥	٤	٣	٢	١	١١ - التحمل الدوري التنفسي
١٠	٩	٨	٧	٦	٥	٤	٣	٢	١	١٢ - الايقاع
١٠	٩	٨	٧	٦	٥	٤	٣	٢	١	١٣ - التوقيت
١٠	٩	٨	٧	٦	٥	٤	٣	٢	١	١٤ - البنيان الجسماني
١٠	٩	٨	٧	٦	٥	٤	٣	٢	١	١٥ - الطول
١٠	٩	٨	٧	٦	٥	٤	٣	٢	١	١٦ - الوزن
١٠	٩	٨	٧	٦	٥	٤	٣	٢	١	١٧ - القابلية للتعلم الحركي

* تحسب درجة صدق كل مكون من المكونات الـ ١٧ باستخدام المعادلة التالية:

$$\text{ص} = \text{ل} + \frac{٠,٥ - \text{مجن}}{\text{ن}} \quad \text{معادلة (٣ - ١)}$$

حيث إن:

ص = درجة صدق المكون (العنصر)

ل = الحد الأدنى للدرجة الوسيطة

٥,٠ = مقدار ثابت

مجن = مجموع النسب التي تقع قبل الدرجة الوسيطة

ن = النسبة الوسيطة

فإذا جاءت استطلاعات رأى الخبراء بالنسبة لمكون «القوة العضلية» فى

مثالنا الحالى كالتالى:

(مكون القوة العضلية)

	١٠	٩	٨	٧	٦	٥	٤	٣	٢	١	التقديرات
(٣٠)	٥	٥	٥	٤	٣	٣	٢	٢	١	٠	عدد المحكمين
(١,٠٠)	٠,١٧	٠,١٧	٠,١٧	٠,١٣	٠,١	٠,١	٠,٠٧	٠,٠٧	٠,٠٣	٠	نسبة المحكمين

(عدد المحكمين = ٣٠)

فإنه يمكن حساب المعطيات الخاصة بالمعادلة (٣ - ١) كالتالى:

- حساب نسبة المحكمين الخاصة بكل تقدير وتتم بقسمة عدد المحكمين

أسفل كل تدرج مقسوماً على العدد الكلى للمحكمين كالتالى:

$\frac{٠}{٣} = ٠$ ، $\frac{١}{٣} = ٠,٣$ ، $\frac{٢}{٣} = ٠,٦٧$ ، وهكذا بالنسبة لبقية التقديرات.

- نقوم بحساب الوسيط للتقديرات (١, ٢, ٣, ٤, ٥, ٦, ٧, ٨, ٩, ١٠)

كالتالى:

$$\text{ترتيب الوسيط} = \frac{١ + ن}{٢}$$

$$\frac{1 + 1.0}{2} =$$

$$0,50 =$$

$$\frac{6 + 5}{2} = \text{إذا قيمة الوسيط}$$

$$0,5 =$$

- حساب الحد الأدنى للدرجة الوسيطة كالتالى:

$$0,5 - 0,5 = 0,0$$

$$0 =$$

- نقوم بحساب مجز ن كالتالى:

$$0,27 = 0,1 + 0,07 + 0,07 + 0,03 + 0,0$$

$$0,27 = \text{إذا مجز ن}$$

- نقوم بحساب النسبة الوسيطة ن فنجد أن تساوى 0,1

- نقوم بالتعويض فى المعادلة (3 - 1) كالتالى:

$$\frac{0,27 - 0,5}{0,1} + 5 = \text{ص}$$

$$\frac{0,23}{0,1} + 5 =$$

$$2,3 + 5 =$$

$$7,3 =$$

إذا درجة صدق مكون (عنصر) القوة العضلية يساوى 7,3 .

وإذا جاءت استطلاعات رأى الخبراء بالنسبة لمكون «القدرة العضلية» فى

نفس المثال كالتالى:

	١٠	٩	٨	٧	٦	٥	٤	٣	٢	١	التقديرات
(٣٠)	٥	٥	٥	٣	٢	٣	٤	٢	١	٠	عدد المحكمين
(١,٠٠)	٠,١٧	٠,١٧	٠,١٧	٠,١	٠,٠٧	٠,١	٠,١٣	٠,٠٧	٠,٠٣	٠	نسبة المحكمين

(عدد المحكمين = ٣٠)

من هذه النتيجة نجد أن:

الوسيط = ٥,٥ (نفس النتيجة السابقة)

ل = ٥ (نفس النتيجة السابقة)

ومجدن = $٠,٣ + ٠,٠٧ + ٠,١٣ + ٠,١$

= ٠,٣٣

ن = ٠,١

وبالتعويض فى المعادلة (٣ - ١) يتج أن:

$$ص = \frac{٠,٣٣ - ٠,٥}{٠,١} + ٥ =$$

$$= \frac{٠,١٧}{٠,١} + ٥ =$$

$$= ١,٧ + ٥ =$$

$$٦,٧ =$$

إذا درجة صدق مكون «القدرة العضلية» يساوى ٦,٧ .

وهكذا يتم تطبيق المعادلة (٣ - ١) بالنسبة لبقية المكونات (العناصر) الواردة

بالتداول (٣ - ٢).

● يتم ترتيب قيم (ص) تنازليا بحيث نبدأ بأعلى درجة وننتهى بأقل درجة، ثم نأخذ نسبة ٣٣٪ من عدد المكونات لبناء بطارية الاختبار الخاصة بقياس القدرة الحركية، حيث تصل هذه النسبة إلى:

عدد المكونات = ١٧

$$\frac{17 \times 33}{100} = ٥,٦ =$$

٥,٦ =

٦ = مكونات (عناصر)

وسوف نجد أن هذه المحاولة قد أسفرت عن ستة مكونات هي:

١- القوة العضلية Muscular strength

٢- القدرة العضلية Muscular power

٣- التوافق البدني Body coordination

٤- السرعة الانتقالية Running speed

٥- الرشاقة (سرعة تغيير الاتجاه) Agility

٦- التحمل الدورى التنفسى Cardiovascular-respiratory endurance

حيث إن هذه المكونات هي الأكثر صدقا بالنسبة لقياس القدرة الحركية.

ثانياً: بالنسبة للاختبارات غير العملية (اختبارات الورقة والقلم):

وتستخدم هذه الطريقة في حالة تقنين اختبارات التحصيل والذكاء وقوائم

الشخصية وبالنسبة للعديد من المقاييس في علم النفس الرياضى Sport and

exercise psychology، حيث يتم تطبيقها وفق الخطوات التالية:

● إعداد البنود items (الفقرات - العبارات - الجمل) بحيث يكون عددها يفوق العدد المتوقع أن يتكون منه المقياس أو الاختبار، وذلك لكون عمليات التقنين المختلفة (الصدق والثبات) سوف تستبعد العديد من هذه الفقرات حيث يستلزم ذلك مراعاة الشروط الخاصة بإعداد الفقرات (راجع هذه الشروط).

● إعداد تعليمات المقياس أو الاختبار مع ملاحظة أن هذه التعليمات تختلف وفقاً لطبيعة المقياس أو الاختبار. فعند إعداد مقياس للاتجاهات نحو التربية الرياضية، يمكن أن تكون التعليمات التي تسبق العبارات على النحو التالي:

- فيما يلي بعض العبارات المرتبطة بالتربية الرياضية والمطلوب منك أن تحدد مدى موافقتك أو عدم موافقتك على كل عبارة بالصورة التالية:

أ - إذا كنت توافق على العبارة بدرجة كبيرة ضع علامة (x) أسفل كلمة (أوافق بدرجة كبيرة) وأمام رقم العبارة.

ب - إذا كنت توافق على العبارة ضع علامة (x) أسفل كلمة (أوافق) وأمام رقم العبارة.

ج - إذا كانت لا توافق على العبارة ضع علامة (x) أسفل كلمة (أعارض) وأمام رقم العبارة.

د - إذا كنت لا توافق بدرجة كبيرة على العبارة ضع علامة (x) أسفل كلمة (أعارض بدرجة كبيرة) وأمام رقم العبارة.

- وقد وضعت أوزان العبارات بغض النظر عن اتجاه العبارة كالتالي:

(أوافق بدرجة كبيرة) (أوافق) (أعارض) (أعارض بدرجة كبيرة)

١

٢

٣

٤

- يرجى منك قراءة كل عبارة جيداً، ثم حدد مدى موافقتك أو عدم موافقتك عليها.

ملحوظة :

لا يؤثر اتجاه العبارة على مدى موافقة أو عدم موافقة الخبير، ففي حالة إعداد «مقياس الاتجاهات نحو التربية الرياضية» نجد أن العبارتين (١, ٢) كالتالى :

١ - تهتم التربية الرياضية ببناء العضلات فقط . (اتجاه العبارة سلبى)
(أوافق بدرجة كبيرة) (أوافق) (أعارض) (أعارض بدرجة كبيرة)
×

٢ - حصص التربية الرياضية تتيح الفرصة لإقامة الصداقات بين التلاميذ (اتجاه إيجابى)
(أوافق بدرجة كبيرة) (أوافق) (أعارض) (أعارض بدرجة كبيرة)
×

يلاحظ أن العبارة الأولى تقيس الاتجاهات نحو التربية الرياضية بالرغم من أنها سلبية الاتجاه، ولذلك حظيت بالموافقة بدرجة كبيرة، وأن العبارة الثانية تقيس أيضا نفس الاتجاه وهى موجبة الاتجاه وقد حظيت بالموافقة بدرجة كبيرة وهكذا بالنسبة لبقية العبارات .

● يطرح المقياس على مجموعة من الخبراء المتخصصين فى علم النفس الرياضى لاستطلاع آرائهم حول صدق العبارات التى يتضمنها المقياس، وبحيث لا يقل عدد هؤلاء الخبراء عن ٣٠ خبيراً .

● حساب درجة صدق كل عبارة من عبارات المقياس وفق الأسلوب الذى اتبع فى المثال السابق (٣ - ١) .

ملحوظة مهمة:

يمكن استطلاع آراء الخبراء حول صدق عبارات مقياس الاتجاهات المقترح وفق مقياس تقدير عشاري كما في المثال (٣ - ١) كالتالي:

العبارات مقياس التقدير

١ - تهتم التربية الرياضية ببناء العضلات فقط

١٠، ٩، ٨، ٧، ٦، ٥، ٤، ٣، ٢، ١

٢ - حصص التربية الرياضية تتيح الفرصة لإقامة الصداقات بين التلاميذ.

١٠، ٩، ٨، ٧، ٦، ٥، ٤، ٣، ٢، ١

٣ -

٤ -

يلى ذلك تصفيه عبارة المقياس وفق الإجراءات التي تم اتباعها في المثال السابق (٣ - ١) أيضاً.

طريقة المحك Criterion Method

يقصد بالمحك المعيار أو الميزان الذي نحكم به على صدق أداة القياس المطلوب تقويم صدقها وتنقسم المحكات إلى نوعين رئيسيين هما:

المحك الداخلي Internal Criterion

ويستخدم للحكم على مدى الاتساق فيما بين وحدات (بنود) المقياس أو الاختبار. وتقوم فكرة هذا المحك على أن جميع وحدات (بنود) المقياس أو الاختبار تقيس نفس الشيء، وأنها إذا كانت كذلك فإنها سوف تظهر فيما بينها معاملات ارتباط غير صفرية مما يؤكد اتساق هذه الوحدات أو البنود. ويستخدم هذا النوع من المحكات لتقدير ما يعرف بالاتساق الداخلي للمقياس أو الاختبار، حيث يكثّر استخدامه عند تقدير صدق المحتوى (المضمون)

بالنسبة للعديد من المقاييس النفسية الاستبيانات الشخصية، كما يستخدم لتقدير صدق المحتوى (المضمون) بالنسبة لبعض اختبارات التحصيل والاستعدادات والذكاء وغيرها. حيث يلجأ الباحثون إلى استخدام هذا المحك في حالة عدم إمكانية استخدام النوع الثانى من المحكات وهو: المحك الخارجى، حيث يقرر معظم علماء القياس النفسى أنه يمكن الاعتماد بطريقة الاتساق الداخلى Internal Consistency كوسيلة لتقدير الصدق الداخلى للمقياس أو الاختبار (راجع الاتساق الداخلى كطريقة لحساب الثبات).

المحك الخارجى External Criterion

وهو عبارة عن مقياس خارجى external مستقل independent يستخدم بشكل رئيسى لقياس نفس الظاهرة التى يفترض أن المقياس أو الاختبار الذى يتم إعداده يقيسها. ويفترض فى هذا المحك أنه على درجة عالية جداً من الصدق تستوجب الثقة التامة فيه، وذلك لكون الباحث يعتمد عليه كمعيار (ميزان) للحكم على مدى صدق المقياس أو الاختبار الذى يقوم بإعداده.

وتكمن الصعوبات الرئيسية بالنسبة للمحك الخارجى فى النقاط المهمة التالية:

- نقص المحكات Criteria المناسبة لتقدير الصدق وبخاصة فى مجالات قياس الشخصية والتحصيل الدراسى والاستعدادات المهنية.
- قد يكون للمحكات الخارجية معاملات صدق منخفضة.
- قد تكون المحكات الخارجية قد أعدت فى الأصل على عينات من أفراد لهم خصائص وسمات وقدرات تختلف عن عينة الأفراد التى يقن لها المقياس أو الاختبار الذى يعده الباحث، حيث تمثل الفروق العرقية Races والفروق فى المستوى الثقافى والحضارى محددات أساسية بالنسبة للمحكات الخارجية.

خصائص المحكات الخارجية:

للمحكات الخارجية عدة خصائص نذكر منها الآتى:

- أنها تعبر عن الشيء المطلوب قياسه أفضل تعبير.
- أنها تعبر عن المستوى الأفضل من الأداء الناجح.
- أنها تقيس بطريقة مباشرة الظاهرة المطلوب قياسها أفضل من الأداة المقترحة (المقياس أو الاختبار).
- أن درجات المحك الخارجى دائما أكبر من درجات الأداة المقترحة.
- أن كل من المحك الخارجى وأداة القياس المقترحة يظهران كمتغيرين مستقلين، إلا أنه يفترض أن كلا منهما يقيس شيئاً واحداً.

أنواع المحكات الخارجية فى المجال الرياضى

يستخدم لتقنين المقاييس والاختبارات فى مجال التربية البدنية والرياضة العديد من المحكات الخارجية منها:

١ - نتائج المسابقات الرياضية:

يمكن المستخدم الأداء الحقيقى للظاهرة المقيسة كمحك خارجى لحساب الصدق بالنسبة للعديد من اختبارات المهارات الرياضية Sports skills وبخاصة فى الألعاب الفردية Individual sport كالتنس والريشة الطائرة والجودو وغيرها، وذلك بأن تجرى بين الأفراد (المفحوصين) مسابقات حقيقية جادة فى النشاط الرياضى المستهدف عن طريق نظام الدورى tournament أو أى طريقة أخرى، ثم نقارن بين ترتيب هؤلاء الأفراد وفقاً لنتائجهم فى الدورى وبين ترتيبهم وفقاً لنتائجهم فى الاختبار المقترح، يلى ذلك حساب معامل ارتباط سبيرمان لفروق الرتب Rho حيث يدل معامل الارتباط المحسوب على معامل صدق الاختبار المقترح (الصدق التلازمى).

ومن ناحية أخرى فقد يمكن منح الأفراد درجات (نقاط) فى ضوء نتائجهم فى المسابقات (كما يحدث فى بعض الألعاب)، حيث يمكن استخدام معامل ارتباط بيرسون Pearson للدرجات الخام لتقدير العلاقة بين الدرجة الخام التى حصل عليها كل لاعب طبقاً لنتائجه فى المسابقات وبين الدرجة التى حصل عليها فى الاختبار، وذلك بدلاً من نظام الترتيب، حيث يدل معامل الارتباط الحسوب على معامل صدق الاختبار(*) .

٢ - معاملات الارتباط باختبارات أخرى:

يمكن استخدام بعض المقاييس أو الاختبارات التى ثبت صدقها كـمحكات خارجية لتقدير صدق أدوات القياس المزمع إعدادها، حيث يتم ذلك عن طريق حساب معامل الارتباط (بيرسون) بين درجات الأداء على المحك ودرجات الأداء على أداة القياس المقترحة لتقدير صدق هذه الأداة .

٣ - تقديرات الخبراء :

قد تستخدم تقديرات الخبراء experts للأداء الفعلى للأفراد (المتسابقين - اللاعبين) كمحك خارجى لتقدير صدق بعض أدوات القياس، وتستخدم هذه الطريقة عندما يكون المحك المطلوب خاص بأداء الفرق الجماعية (ألعاب الفريق) team sport فى مهارة واحدة أو مهارتين، حيث يتم حساب معامل الارتباط بين تقديرات الخبراء (كمحك خارجى) والدرجات على أداة القياس المقترحة لتقدير صدق هذه الأداة (المقياس أو الاختبار).

* راجع طرق حساب كل من معامل ارتباط سيرمان ومعامل ارتباط بيرسون فى كتاب: الإحصاء الوصفى فى علوم التربية البدنية والرياضية، دار الفكر العربى، ٢٠٠٢م (للمؤلف).

وتتطلب هذه الطريقة الآتى:

- اختيار الخبراء بدقة متناهية فى ضوء معايير واضحة ومحددة، مع مراعات أن يكونوا من ذوى الخبرة العملية والنظرية فى مجال الظاهرة (اللعبة) المطلوب قياسها، حيث يلاحظ أن المهم نوعية هؤلاء الخبراء وليس عددهم.

- وضع مقاييس تقدير Rating scales تتضمن عناصر الأداء المطلوب تقويمها بالنسبة للمحك، على أن تكون هذه العناصر محددة تحديداً دقيقاً وواضحاً لدى كل خبير.

- توضيح شروط الأداء الناجح لكل مهارة من المهارات المطلوب تقويمها وكيفية تقدير الدرجة على مقاييس التقدير.

- إعطاء الخبراء فرصة ملاحظة الظاهرة المطلوب تقويمها (المهارة) أكثر من مرة فى مواقف مختلفة، وذلك بغرض إكسابهم الألفة والخبرة بهذه المهارة مما يزيد من موضوعية Objectivity تقدير الدرجة.

- يستخدم متوسط درجات جميع الخبراء لكل فرد (مفحوص) على حدة كدرجة نهائية تمثل الأداء على المحك فى مقابل درجة نفس الفرد فى الأداء على الأداة المقترحة كزوجين من الدرجات لحساب معامل الارتباط الذى يدل على صدق هذه الأداة.

كما سبق يتضح أن المحك الخارجى عبارة عن:

- ميزان ثبت بالتجربة أنه يتمتع بدرجة عالية جدا من الصدق.
- أو تم التأكد من صدقه نتيجة كثرة استخدامه فى الدراسات والبحوث العلمية مثل اختبارات: الشد لأعلى أو الدفع على المتوازيين فى اختبارات تحمل القوة، واختبارات بينيه أو وكسلر بالنسبة لبحوث الذكاء.

- قد يكون نتائج الأداء فى مسابقات رياضية حقيقية .
- أو مجموعة من التقديرات (الاحكام) أصدرها خبراء متخصصون .

خطوات حساب صدق الاختبار بطريقة المحك الخارجى:

- ١ - اختيار المحك
- ٢ - اختيار عينة التقنين بحيث لا تقل عن ٣٠ فرداً وبحيث يتم تحديد خصائص هذه العينة تحديداً دقيقاً .
- ٣ - تطبيق الاختبار (أداة القياس) على عينة التقنين (أولاً) .
- ٤ - تطبيق المحك على عينة التقنين (ثانياً) .
- (يلاحظ ضرورة التزامن بين تطبيق كل من الاختبار والمحك - بمعنى ألا يفصل بين التطبيقين فترة زمنية طويلة) .
- ٥ - حساب معامل الارتباط بين درجات الاختبار ودرجات المحك حيث تكشف قيمة معامل الارتباط المحسوب عن معامل الصدق .

ماهية معامل الصدق Validity Coefficient

يستخدم الارتباط Correlation للتعبير عن العلاقة relationship بين درجات الاختبار ودرجات المحك (الخارجى)، حيث يطلق على هذا الارتباط اسم معامل الصدق، هذا المعامل يكشف لنا عن المدى الذى يقيس به الاختبار ما يقيسه المحك . وبناء على ذلك يمكن تعريف صدق الاختبار على أنه ارتباط الاختبار ببعض المحكات الخارجية .

وفى هذه الخصوص نود أن نشير إلى أنه يمكن حساب معامل الارتباط بطرق مختلفة هي:

- ١ - معامل ارتباط بيرسون Pearson وهو الأكثر شيوعاً .

- ٢ - الارتباط الثنائي Biserial Correlation .
- ٣ - الارتباط الرباعي Tetrachoric Correlation .
- ٤ - معامل فاي Phi - Coefficient .
- ٥ - الارتباط المتعدد Multiple Correlation .
- ٦ - ارتباط الرتب (سبيرمان) Rank- Order Correlation .

حيث تستخدم هذه المعاملات في الحالات التالية:

- يستخدم الارتباط الثنائي عندما تقسم درجات الاختبار أو المحك في فئتين مثل (ناجح - راسب) أو (صح - خطأ) بمعنى أن تكون الإجابة على أى من الاختبار أو المحك ثنائية البعد.
- يستخدم الارتباط الرباعي ومعامل فاي عندما تكون درجات كل من الاختبار والمحك مقسمة في فئتين، بمعنى أن الإجابة على كل منهما تكون ثنائية البعد.
- يستخدم الارتباط المتعدد عندما يتكون المحك من أكثر من مقياسين، بمعنى أن المحك يتكون من درجة مركبة Composite لمقياسين أو أكثر حيث تدل العلاقة بين مقياس واحد وهذه الدرجة على معامل الصدق (وسوف نتعرض لهذه الطريقة لاحقاً).
- يستخدم ارتباط الرتب (معامل سبيرمان) عندما تكون درجات كل من الاختبار والمحك في هيئة رتب، بمعنى أن هذا المعامل يستخدم في حالة البيانات المتقطعة discrete data أو حالة البيانات المتصلة التي يمكن وضعها في شكل رتب.
- يستخدم معامل ارتباط بيرسون (وهو الأكثر شيوعاً) في حالة البيانات المتصلة Continuous data للعينات التي تساوى أو تزيد عن ٣٠ مفردة.

تقويم معامل الصدق:

يمكن تقويم معاملات الارتباط التى تدل على معاملات الصدق على أساس الدلالة الإحصائية لنتائج الأسلوب الإحصائى المستخدم، ومع ذلك يضع علماء القياس بعض الشروط المهمة لتقويم معاملات الصدق، نذكر من هذه الشروط ما يلى:

- يرى «بارو وماكجى، ١٩٧٩م» أنه يلزم تقويم معاملات الصدق المرتبط بالمحك فى مجال القياس النفس - حركى وفقاً للمستويات التالية:

معامل الصدق	تقديرات الصدق
٠,٨٥ - ٠,٩٩	ممتاز
٠,٨٠ - ٠,٨٤	جيد جداً
٠,٧٠ - ٠,٧٩	مقبول
٠,٦٠ - ٠,٦٩	مرفوض

- ويرى «كيركيندال وآخرون، ١٩٨٧م» أن معامل الصدق الذى يتم تقديره على ١٠٠ مفحوص أو أكثر، يمكن تقويمه على أساس المستويات التالية:

معامل الصدق	تقديرات الصدق
٠,٨٠ - ١,٠٠	ممتاز
٠,٧٠ - ٠,٧٩	مرتفع
٠,٥٠ - ٠,٦٩	متوسط
أقل من ٠,٥٠	مرفوض

- وتشير «سافريت، ١٩٨٦» أنه يمكن تقويم معامل الصدق على أساس الآتى:

* يكون معامل الارتباط الذى يدل على معامل الصدق مرتفعاً إذا كان يساوى ٠,٩٠ فأكثر.

* معامل الصدق الذي يساوى ٠,٨٠ ، فأكثر يمكن قبوله .

* يلاحظ أن مربع معامل الصدق يمثل كمية التباينات الشائعة بين الاختبار والمحك، فمعامل الصدق الذي يساوى ٠,٥٠ له تباينات تساوى ٢٥٪ من التباينات بين درجات الاختبار ودرجات المحك، وأنه كلما زادت التباينات بين درجات الاختبار ودرجات المحك، كلما دل ذلك على أن الاختبار والمحك يقيسان نفس الشيء والعكس صحيح .

* يمكن قبول الصدق التنبؤى إذا كان معامل الصدق يتراوح (من ٠,٥٠ إلى ٠,٦٠) وذلك بسبب توقع انخفاض التباينات بين درجات الاختبار ودرجات المحك لاختلاف طبيعة كل منهما .

تصحيح معامل الصدق المرتبط بالمحك (الخارجى):

يتأثر معامل صدق الاختبار المحسوب باستخدام بعض المحكات الخارجية بمعامل ثبات كل من الاختبار والمحك الخارجى، حيث يتطلب ذلك تصحيح معامل الصدق المحسوب بأى من الطريقتين التاليتين:

الطريقة الأولى: وتتضمن استخدام كل من الاختبار والمحك .

الطريقة الثانية: وتتضمن استخدام المحك فقط .

الطريقة الأولى :

وتستخدم هذه الطريقة عندما يكون ثبات كل من الاختبار والمحك منخفضاً لأسباب تتعلق بطريقة بناء (إعداد) كل منهما، حيث يتوقع فى مثل هذه الحالة أن يكون معامل الارتباط بين الاختبار والمحك هو الآخر منخفضاً، ولتصحيح معامل الصدق التجريبي تستخدم المعادلة التالية:

معادلة (٣ - ٢)

$$\frac{\sqrt{s_1}}{\sqrt{s_2 \times s_3}} = \sqrt{s_2}$$

حيث إن :

$\sqrt{s_2}$ = معامل صدق الاختبار بعد التعديل .

$\sqrt{s_1}$ = معامل صدق الاختبار قبل التصحيح (معامل الصدق التجريبي) .

$\sqrt{s_3}$ = معامل ثبات الاختبار .

$\sqrt{s_2}$ = معامل ثبات المحك .

مثال : (٣-٢) :

استخدم أحد الباحثين اختبار الشد لأعلى pull - ups كمحك خارجي لحساب الصدق التجريبي لاختبار الانبطاح المائل ثنى الذراعين Push-ups ، وقد حصل الباحث على معامل صدق بلغ ٠,٧٠٤ ، فإذا كان معامل ثبات المحك هو ٠,٩٥٠ ، وكان معامل ثبات اختبار الانبطاح المائل ثنى الذراعين هو ٠,٩٦١ ، فما هو معامل صدق الاختبار بعد التصحيح :

الحل :

معامل الصدق التجريبي = ٠,٧٠٤ ،

معامل ثبات الاختبار = ٠,٩٦٠ ،

معامل ثبات المحك = ٠,٩٥٠ ،

وبالتعويض فى المعادلة (٣- ٢) ينتج أن:

$$\frac{0,704}{\sqrt{0,950 \times 0,960}} = \sqrt{\text{ص}_2}$$

$$\frac{0,704}{\sqrt{0,91295}} =$$

$$\frac{0,704}{0,95548} =$$

$$0,737 =$$

إذاً معامل الصدق التجريبي لاختبار الانبساط المائل ثنى الذراعين بعد التعديل أصبح يساوى ٠,٧٣٧ بدلا من ٠,٧٠٤

الطريقة الثانية:

وتستخدم هذه الطريقة عندما يكون معامل ثبات المحك يقل كثيرا عن معامل ثبات الاختبار نفسه، فى مثل هذه الحالة يصبح من الضرورى إجراء عملية التصحيح بالنسبة للمحك فقط حيث تعرف هذه الطريقة باسم: التصحيح فى اتجاه واحد - one-way - Correction of attenuation - والتي تستخدم المعادلة التالية :

$$\frac{\sqrt{\text{ص}_1}}{\sqrt{\text{ص}}} = \sqrt{\text{ص}_2}$$

معادلة (٣ - ٣)

حيث إن :

$$\sqrt{\text{ص}_2} = \text{معامل صدق الاختبار بعد التصحيح.}$$

$$\sqrt{\text{ص}_1} = \text{معامل صدق الاختبار قبل التصحيح (معامل الصدق التجريبي).}$$

$$\sqrt{\text{ص}} = \text{معامل ثبات المحك.}$$

مثال (٣-٣):

قام أحد الباحثين بإعداد اختبار لقياس مهارة الدقة فى التمرير فى كرة القدم، وقد بلغ معامل ثبات هذه الاختبار ٠,٨٨ ، ولتقدير الصدق التجريبي لهذا الاختبار استخدام الباحث مجموعة من الخبراء لتقدير درجة الأداء فى هذه المهارة لمجموعة من اللاعبين أثناء الأداء الفعلى فى المباريات الرسمية كمدك خارجى ، وعندما قام الباحث بحساب الصدق التجريبي للاختبار المقترح وجد أنه يساوى ٠,٦٥ ، وهو معامل صدق منخفض، فهل يأخذ الباحث بهذا المعامل أم يرفضه إذا كان معامل ثبات المدك هو ٠,٦٧ ؟

الحل:

بالتعويض فى المعادلة (٣ - ٣) ينتج أن:

$$\frac{0,65}{0,67} \sqrt{2} = 0,794$$

$$\frac{0,65}{0,819} =$$

$$0,794 =$$

إذاً معامل الصدق التجريبي للاختبار المقترح يساوى ٠,٧٩٤ وهو معامل مقبول.

ملحوظة مهمة:

من المثالين (٣ - ٣)، (٢ - ٣) يلاحظ الآتى:

- أن القيمة $\sqrt{0,961} \times 0,950$ والتى تساوى ٠,٩٥٥٤٨ (٠,٩٥٥) (تقريباً) تعنى أن معامل الارتباط بين الاختبار والمدك فى المثال (٢ - ٣) لن يتعدى بأى حال من الأحوال القيمة ٠,٩٥٥ .

- وأن القيمة $\sqrt{0,67} \times 0,819$ والتى تساوى ٠,٨١٩ تعنى أن معامل الارتباط بين الاختبار والمدك فى المثال (٣ - ٣) لن يتعدى بأى حال من الأحوال قيمة ٠,٨١٩ .

طريقة المقارنة الطرفية:

وتستخدم هذه الطريقة لتقدير الصدق المرتبط بالتحك (الصدق التلازمي)، وهي تعتمد على قدرة الاختبار على التمييز بين طرفي السمة أو القدرة التي يقيسها، حيث يستخدم اختبار ت-t لدراسة دلالة الفروق بين المتوسطات الحسابية بدلاً من استخدام الارتباط. حيث تتم المقارنة في هذه الطريقة بأسلوبين هما:

الأسلوب الأول (خاص بكل من الاختبار والتحك الخارجي):

أ - ويتم بمقارنة الثلث الأعلى في درجات الاختبار بالثلث الأعلى في درجات التحك.

ب - مقارنة الثلث الأدنى في درجات الاختبار بالثلث الأدنى في درجات التحك. فإذا كان الفرق بين المتوسط الحسابي للثلث الأعلى للاختبار والمتوسط الحسابي للثلث الأعلى للتحك، وكان الفرق بين المتوسط الحسابي للثلث الأدنى للاختبار والمتوسط الحسابي للثلث الأدنى للتحك غير دال إحصائياً كان ذلك مؤشراً على أن الاختبار صادق، بمعنى أن الاختبار يكافؤ التحك الخارجي من حيث البناء.

ويستخدم اختبار ت - t لحساب دلالة هذا الفرق حيث يلزم أن تكون قيمة ت - t المحسوبة غير دالة إحصائياً لقبول صدق الاختبار، حيث يمكن التعبير عن ذلك كالتالي:

$$\bar{s}_1 = \bar{s}_2$$

$$\text{أو } \bar{s}_1 - \bar{s}_2 = 0$$

الأسلوب الثاني (خاص بالاختبار فقط):

ويتم بمقارنة درجات الثلث الأعلى للتحك بدرجات الثلث الأدنى للاختبار، وبمقارنة درجات الثلث الأدنى للاختبار بدرجات الثلث الأدنى للتحك، وبمقارنة درجات الثلث الأعلى للاختبار بدرجات الثلث الأعلى للتحك، وبمقارنة درجات الثلث الأوسط للاختبار بدرجات الثلث الأوسط للتحك، حيث يمكن التعبير عن ذلك كالتالي:

المتوسطين، فإذا كان هذا الفرق دالا إحصائيا كان الاختبار صادقا. ويعاب على هذا الأسلوب أنه أقل دقة من الأسلوب الأول لكونه لا يستخدم المحك الخارجى.

وإذا كان هذا الأسلوب يستخدم بكثرة لحساب الصدق لسهولته، فإنه يلزم عند استخدامه مراعاة الشروط التالية:

- أن تكون عينة التقنين كبيرة العدد نسبيا.
- أن تكون عينة التقنين غير متجانسة تجانساً شديداً بالنسبة للسمة أو القدرة التى يقيسها الاختبار.
- أن تختار عينة التقنين بالأسلوب العشوائى.
- أن يفترض أن بيانات السمة أو القدرة المقيسة تتوزع توزيعاً مقبولاً بالنسبة لعينة التقنين، بمعنى أن تكون خالية تماماً من عيوب التوزيعات غير الاعتدالية.

طريقة المجموعات المتضادة Contrasted Groups

تستهدف هذه الطريقة تقدير صدق الاختبار على أساس قدرته على التمييز discrimination بين أصحاب الدرجات المرتفعة وأصحاب الدرجات المنخفضة فى السمة أو القدرة التى يقيسها الاختبار. وقد قدم كامبل وفيسك Compbell and fiske 1959م هذه الطريقة التى عرفت باسم «تحقيق صدق التمييز discriminat validation» باعتبارها أحد أهم طرق حساب صدق التكوين الفرضى.

ويمكن تعريف القدرة على التمييز على أنها: قدرة الاختبار المقترح على التفريق بين الأفراد الذين يتمتعون بدرجة مرتفعة من السمة أو القدرة من ناحية، وبين الأفراد الذين يتمتعون بدرجة منخفضة من نفس السمة أو القدرة من ناحية أخرى.

وتكمن المشكلة الحقيقية بالنسبة لهذه الطريقة فى كيفية تحديد المجموعات المتباينة فى مستوى السمة أو القدرة التى يقيسها الاختبار، إلا أن جرونلند Gronlund ١٩٨١م يوضح أن الخبرة الذاتية لمن يقوم بإعداد الاختبار بالإضافة إلى نتائج الدراسات والبحوث المتخصصة يمكن الاستفادة منها فى تحديد المجموعات المتضادة فى مستوى السمة أو القدرة التى يقيسها الاختبار بصورة يمكن الثقة فيها كمحك لتقويم صدق الاختبار، فمثلا يمكن استخدام متغير السن Age لتحديد المجموعات المتضادة بالنسبة للقدرة العقلية والقدرات البدنية على أساس أن هذه القدرات تزداد مع زيادة السن وتستمر كذلك خلال مرحلة الطفولة Childhood (من سن ٦ - ١٣ سنة) وحتى مرحلة المراهقة Adolescence. كما يمكن استخدام الجنس كمتغير لتحديد المجموعات المتضادة (الأولاد فى مقابل البنات) بالنسبة للعديد من القدرات البدنية كالقوة العضلية Muscular strength والقدرة العضلية Muscular power والسرعة فى العدو sprint وغيرها حيث تؤكد البحوث والدراسات العلمية أن الأولاد يحققون درجات مرتفعة تفوق الدرجات التى يحققها البنات على مثل هذه الاختبارات.

ومن ناحية أخرى تؤكد نتائج البحوث العلمية أن البنين يتفوقون على البنات فى الاستيعاب الميكانيكى Mechanical Comprehension وهى قدرة مركبة من القدرة الميكانيكية Mechanical ability والقدرة المكانية Spatial ability، فى حين يتفوق البنات على الأولاد فى الاستيعاب القرائى (فهم المادة المقروءة) Reading Comprehension، وعليه فإنه يمكن استخدام كل من الاستيعاب الميكانيكى والاستيعاب القرائى لإعداد المجموعات المتضادة للتحقق من قدرة بعض الاختبارات على التمييز.

وبالمثل يمكن استخدام المجموعات المدربة Trained فى مقابل المجموعات غير المدربة Untrained للتحقق من صدق التمييز بالنسبة للعديد من

الاختبارات البدنية واختبارات المهارات الحركية فى الأنشطة والألعاب الرياضية المختلفة، تأسيساً على أن المجموعات المدربة تتميز بمستوى أفضل من التحصيل Achievement «الاكتساب» بالمقارنة بالمجموعات غير المدربة .

وقد تستخدم نتائج الفحوص الطبية فى تعيين المجموعات المتضادة، فقد استخدم ماك كيردى ولارسون Mac Curday and Larson نتائج الفحوص الطبية لتحديد ثلاث مجموعات متباينة المستوى (متضادة) للتحقق من قدرة اختبارهما للكفاءة الوظيفية organic efficiency test على التمييز، وكانت هذه المجموعات كالتالى :

المجموعة الأولى: وتضم الأفراد الذين يتمتعون بحالة بدنية ممتازة .

المجموعة الثانية: وتضم الأفراد الذين يتمتعون بحالة بدنية متوسطة .

المجموعة الثالثة: وتضم الأفراد الذين يتمتعون بحالة بدنية تتميز بالضعف والوهن .

ثم قاما بتقويم صدق الاختبار على أساس قدرته على التمييز بين المجموعات الثلاثة .

ومن ناحية أخرى فقد تستخدم الملاحظة فى تعيين المجموعات المتضادة، فقد قام واشبورن Washburne - على سبيل المثال - بتقدير صدق التمييز للقائمة الخاصة بالتكيف الاجتماعى social adjustment inventory لتلاميذ المدارس الثانوية، وذلك من خلال تعيين أربع مجموعات متباينة فى التكيف الاجتماعى عن طريق الملاحظة حيث كانت هذه المجموعات كالتالى :

- مجموعة التلاميذ الذين يظهرون تكيفا اجتماعيا عاليا .

- مجموعة التلاميذ الذين يظهرون تكيفا اجتماعيا متوسطا .

- مجموعة التلاميذ الذين يظهرون عدم التوافق مع النظام المدرسى .

- مجموعة التلاميذ الذين يتتهكون القوانين والنظام المدرسى ويسببون إزعاجاً لإدارة المدرسة .

وقد قام بتقويم صدق القائمة على أساس قدرتها على التمييز بين درجات المجموعات الأربع .

إجراءات تطبيق هذه الطريقة:

للتعرف على قدرة الاختبار على التمييز ، تتبع الإجراءات التالية :

* تحديد السمة أو القدرة المراد قياسها تحديداً واضحاً ودقيقاً .

* تحديد مجموعة من الأفراد ممن يتميزون بمستوى عال في هذه السمة أو القدرة، ومجموعة أخرى ممن يتميزون بمستوى منخفض (ضعيف) في نفس السمة أو القدرة . (قد يتم تحديد أكثر من مجموعتين) .

* يطبق الاختبار المقترح لقياس السمة أو القدرة على المجموعتين (أو المجموعات) ثم نقارن نتائج درجات الاختبار بالنسبة للمجموعتين (المجموعات) للتعرف على دلالة الفروق بين المجموعات باستخدام الأسلوب الإحصائي المناسب (اختبار ت - t أو اختبار ف - f) .

استخدام هذه الطريقة :

تستخدم طريقة المجموعات المتضادة لتقدير صدق التكوين الفرضي للعديد من الاختبارات والمقاييس في مجال التربية الرياضية وعلم النفس الرياضى وبخاصة بالنسبة لاختبارات الذكاء Intelligence ، والاتجاهات Attitudes ، والتفكير الناقد Critical thinking ، والاستيعاب القرائى Reading Comprehension والمهارات فى الأنشطة والألعاب الرياضية، وبالنسبة للعديد من الخصائص والسمات النفسية .

طريقة مقارنة الدرجات قبل وبعد المعالجات الخاصة

Comparison of scores before and after some particular treatments

قد تستخدم بعض المعالجات الخاصة كإجراءات تجريبية تسهم فى التحقق من صدق التكوين الفرضى للاختبار أو المقياس المزمع إعداده (تقنيته)، فنحن

نعلم أن الذكاء مثلا لا يتأثر بالتدريب، وعليه فإننا نتوقع ألا يحدث تغيير في درجات اختبار ما للذكاء عندما يطبق قبلها وبعديا على مجموعة تتدرب على برنامج لتطوير اللياقة البدنية، وذلك لكون الذكاء قدرة فطرية تمتاز بالثبات النسبي وأنها غير قابلة للتغير نتيجة التدريب البدني، بينما نتوقع أن درجات اختبار للتحصيل سوف تزداد بعد انتهاء العام الدراسي عنه عند بداية العام، وذلك لكون التحصيل الدراسي خاصية قابلة للتحسن والزيادة نتيجة التدريب والتعليم الذي توفره المدرسة.

ومن ناحية أخرى فقد لوحظ أن درجات بعض الاختبارات يمكن أن تتغير وفقا لبعض المعالجات (الإجراءات) التجريبية المحددة، ولتوضيح ذلك نسوق المثال الآتي:

«عند تطبيق مقياس القلق كحالة على مجموعة من الرياضيين مرتين، المرة الأولى قبل الاشتراك في مباراة حساسة ومهمة، والمرة الثانية في موقف استرخائي معين. في هذه الحالة يتوقع أن تكون درجات القلق كحالة في القياس الأول تزيد عن درجاته في القياس الثاني (الموقف الاسترخائي)، وذلك على افتراض أن درجات المقياس في موقف المنافسة الرياضية يستدعي درجات أعلى من القلق عنه في حالة الاسترخاء، فإذا كان الفرق بين القياسين دالا إحصائيا دل ذلك على صدق التكوين الفرضي للمقياس».

ويتوقف استخدام أساليب وطرق المعالجات التجريبية في هذا الخصوص على طبيعة السمة أو القدرة المقيسة، بحيث يمكننا وضع التنبؤات الخاصة باختبار ما على أساس أن درجات الاختبار يمكن أن تتغير (أو تظل ثابتة) وفقا لبعض الشروط، فإذا جاءت التنبؤات صحيحة دل ذلك على صدق التكوين الفرضي للاختبار أو المقياس.

ومن ناحية أخرى يمكن استخدام التقديرات الذاتية للمفحوصين وفقا لبعض المعالجات الخاصة للتحقق من صدق التكوين الفرضي لبعض المقاييس

وبخاصة فى مجال علم النفس الرياضى، ولتوضيح ذلك نسوق المثال التالى (*):

«لحساب صدق التكوين الفرضى لمقياس الاتجاهات نحو النشاط البدنى ATPA الذى أعده جيرالد كينيون Gerald kenyon ١٩٦٨م، والذى يتكون من ٥٤ عبارة موزعة على ستة أبعاد للاتجاهات نحو النشاط البدنى هى:

- النشاط البدنى كخبرة اجتماعية (٨ عبارات)

- النشاط البدنى للصحة واللياقة (١١ عبارة)

- النشاط البدنى كخبرة توتر ومخاطرة (٩ عبارات)

- النشاط البدنى كخبرة جمالية (٩ عبارات)

- النشاط البدنى لخفض التوتر (٩ عبارات)

- النشاط البدنى كخبرة للتفوق الرياضى (٨ عبارات)

ولحساب صدق التكوين الفرض للأبعاد الستة للمقياس، قام كينيون بعرض الأبعاد على مجموعة من المفحوصين لسؤالهم عن مدى تفضيلهم لكل بعد من الأبعاد الستة وفقا لمقياس تقدير حماسى هو:

- أفضل بدرجة كبيرة جدا (٥ درجات)

- أفضل بدرجة كبيرة (٤ درجات)

- أفضل بدرجة متوسطة (٣ درجات)

- أفضل بدرجة ضئيلة (درجتان)

- لا أفضله مطلقا (درجة واحدة)

* انظر: محمد حسن علاوى: موسوعة الاختبارات النفسية للرياضيين، مركز الكتاب للنشر، ١٩٩٨م.

ثم قام كينيون بعد هذه الخطوة بعرض المقياس (الأبعاد والعبارات) على نفس مجموعة المفحوصين للإجابة على عباراته الـ ٥٤ وفقاً لمقياس تقدير خماسي هو:

- أوافق بدرجة كبيرة (٥ درجات)
- أوافق (٤ درجات)
- لم أكون رأياً (٣ درجات)
- أعارض (درجتان)
- أعارض بدرجة كبيرة (درجة واحدة)

وقد قام بعد هذه الخطوة بمقارنة درجات التفضيل العالي لكل بعد في مقابل درجات التفضيل المنخفض لنفس البعد وذلك على افتراض أن المفحوصين الذين يعبرون عن تفضيل كبير جداً نحو بعدما يتخذون اتجاهاً أكثر إيجابية نحو العبارات التي تقيس هذا البعد، وأن المفحوصين الذين يعبرون عن تفضيل منخفض نحو بعدما يتخذون اتجاهاً أكثر سلبية نحو العبارات التي تقيس هذا البعد.

وقد قام كينيون بحساب المتوسطات الحسابية والانحرافات المعيارية لدرجات الاستجابات لكل من مجموعة التفضيل العالي في مقابل مجموعة التفضيل المنخفض لكل بعد من الأبعاد الستة كل على حدة، ثم قام بحساب دلالة الفروق بين درجات المجموعتين حيث كانت جميع الفروق دالة إحصائياً لصالح مجموعات التفضيل العالي.

طريقة التحليل العائلي Factorial analysis

التحليل العائلي أسلوب رياضي mathematics يستخدم كمنهج إحصائي لدراسة العلاقات التبادلية interrelationships فيما بين مجموعة من المتغيرات

variables لاختصار (تخفيض) reduction هذه المتغيرات فعندما يكون لدينا استجابات (درجات) لعدد كبير من الاختبارات (القياسات)، فإنه يصبح من الأفضل تخفيض هذا العدد الكبير من البيانات (الدرجات) إلى تجمعات Clusters (تجمعين أو أكثر) من البيانات حتى يسهل التعامل معها وفهمها وتفسير دلالتها.

ويتطلب التحليل العاىلى حساب معاملات الارتباط البينية intercorrelations coefficients ثم وضعها فى مصفوفة معاملات ارتباط Correlation matrix، حيث نقوم بتحليل هذه المصفوفة عامليا بإحدى الطرق الرياضية للتحليل العاىلى لإيجاد التراكيب الخطية (المشتركة) Linear Combination للمتغيرات الداخلة فى التحليل كما تم وصفها عن طريق العلاقات التبادلية (البينية) لمجموعة المتغيرات.

ويطلق على التراكيب الخطية للمتغيرات اسم: المكونات الأساسية principale components أو العوامل factors، حيث يستهدف التحليل العاىلى من إيجاد هذه المكونات وصف الترابط association بين المتغيرات قيد التحليل ما أمكن ذلك.

وقد نحصل من التحليل العاىلى على العديد من المكونات الأساسية فى حالة تحليل عدد كبير من المتغيرات (الاختبارات)، ومع ذلك فكل مكون من المكونات الأساسية يتم استخلاصه وفقا لقواعد رياضية (حسابية) mathematical تجعلها مستقلة independent أى غير مترابطة uncorrelated مع أى من المكونات الأساسية الأخرى.

* سوف نتعرض بالتفصيل لمنهج التحليل العاىلى فى كتاب آخر تحت الطبع.

وينتهى التحليل العاملى إلى مصفوفة العوامل النقية pure factors matrix وتشبعات saturations or Loadings كل اختبار من الاختبارات الداخلية فى التحليل بالعوامل (المكونات) المستخلصة، وكذلك قيم شيوع (اشتراكيات) Commulation الاختبارات بالنسبة لهذه العوامل.

ويكشف التحليل العاملى عن علاقات الاختبارات بكل عامل من العوامل المستخلصة من التحليل العاملى، مما يمكننا من تحديد صدق الاختبار صدقا عامليا Factorial validity، حيث يشير مقدار تشبع الاختبار Factor Loading على العامل إلى صدق الاختبار بالنسبة لقياس هذا العامل وهكذا بالنسبة لبقية العوامل.

فإذا كان اختبار العدو ١٠٠ متر يتشبع بعامل السرعة الانتقالية بمقدار ٠,٨٨، فإننا نستطيع أن نقرر أن هذا الاختبار يعد صادقا فى قياسه لهذا العامل وأن معامل صدقه العاملى يساوى ٠,٨٨.

وإذا كان اختبار الجرى - المشى لمسافة واحد ونصف ميل يتشبع بعامل التحمل الدورى التنفسى (اللياقة الهوائية) بمقدار ٠,٩٤، فإننا نستطيع أن نقرر أن هذا الاختبار يعد صادقا فى قياسه لهذا العامل وأن معامل صدقه العاملى يساوى ٠,٩٤.

ملحوظات مهمة:

- يستخدم التحليل العاملى المحركات الخارجية، هذه المحركات قد تكون اختبارات ثبت صدقها بالإضافة إلى الاختبار الذى يراد التحقق من صدقه.
- يستهدف التحليل العاملى بالدرجة الأولى تحديد التكوينات الفرضية التى يقررها الباحث وعزل الاختبارات الأكثر صدقا بالنسبة للتكوينات التى يتم عزلها isolated نتيجة عمليات التحليل.

- تبين كل من سافريت ١٩٧٥م، وياتس Yates ١٩٧٧م أن العوامل المستخلصة من التحليل العاملى عبارة عن تكوينات مجردة abstract constructs تصاغ على سبيل الافتراض القائم على التفكير المنطقى، وعليه فإن طريقة التحليل العاملى تصبح من الطرق المناسبة لتقدير صدق التكوين الفرضى للعديد من الاختبارات والمقاييس.

- يتفق معظم علماء القياس فى مجال التربية البدنية والرياضة على أنه يمكن استخدام التحليل العاملى لتحديد المكونات الأساسية للياقة البدنية physical fitness والمهارة فى الألعاب والأنشطة الرياضية sport skill وأنه ليس الطريقة الأمثل لتقنين صدق الاختبارات والمقاييس، حيث ينظر إليه - فى حالة عدم استخدام محكات خارجية - على أنه وسيلة من وسائل وصف الصدق الداخلى لوسائل القياس التى تظهر أعلى تشبعا بالعوامل المستخلصة.

طريقة جداول التوقع Expectancy tables

تعتبر جداول التوقع إحدى الطرق التى تظهر العلاقة بين درجات الاختبار ودرجات أو مستويات المحك. وعادة ما تتضمن جداول التوقع التكرارات والنسب المئوية المناظرة لدرجات الأداء على المحك لكل مفحوص له درجات على الاختبار المطلوب تقدير صدقه. معنى هذا أن جداول التوقع تعتمد على حساب التكرار المزدوج لدرجات الاختبار ودرجات أو مستويات الأداء على المحك، لهذا يستخدم التوزيع التكرارى الثنائى (المتغيرين) bivaraite distribution لإعداد مثل هذه الجداول وذلك عن طريق توزيع درجات الاختبار فى مقابل درجات المحك، حيث يتم عن طريق هذا التوزيع حساب معامل صدق الاختبار.

وبناء على مسمى جداول التوقع، فإنها تستخدم لحساب الكفاءة التنبؤية Predictive efficiency للاختبار الذى يتم تقدير صدقه، ومن ثم طرح جداول التوقع الإجابة على العديد من الأسئلة، من هذه الأسئلة على سبيل المثال:

- ما هى احتمالات أن يحصل التلميذ الذى معامل ذكاؤه يساوى ١٠٠ على المرتبة الأولى فى فصله فى اختبار نهاية العام؟
- ما هى احتمالات أن يحصل شخص ما على المثين الـ ٩٠ فى اختبار للسرعة والدقة فى تمرير كرة السلة على الحائط؟

ولتوضيح كيفية استخدام جداول التوقع نسوق المثال التالى:

«لنفترض أن أحد الباحثين أراد حساب صدق اختبار يقيس القدرة مهارية فى كرة القدم، فقام بتطبيق الاختبار على ٢٠٠ من اللاعبين المبتدئين فى كرة القدم، واختار مجموعة من الخبراء المتخصصين فى كرة القدم وطلب منهم ملاحظة أداء هؤلاء اللاعبين وتقدير مستويات كل منهم فى القدرة مهارية فى كرة القدم». ولإعداد جدول التوقع تتبع الخطوات التالية:

* يتم توزيع اللاعبين بناء على تقديرات الخبراء (المحك) إلى خمسة مستويات هى: (ضعيف، متوسط، جيد، جيد جدا، ممتاز).

* وضع درجات الأداء على الاختبار فى شكل فئات مغلقة طول الفئة ٢٠ درجة (نقطة).

* حساب تكرارات اللاعبين المقابلة لكل فئة من فئات درجات الاختبار ووضع النتائج فى جدول توزيع تكرارى مزدوج حيث يلاحظ من هذا الجدول الآتى:

١ - أن عدد اللاعبين الذين حققوا درجات على الاختبار (١٩-٠) هو ١٢ لاعبا، وعندما تم توزيعهم وفقا لتقديرات الخبراء (المحك) وجد أن ١٢ منهم كان مستواهم ٧ ضعيف، ٥ متوسط، صفر جيد، صفر جيد جدا، صفر ممتاز.

جدول التوزيع التكرارى المزدوج لدرجات ٢٠٠ من لاعبي كرة القدم

المجموع	تكرارات المحك					تكرارات الاختبار	فئات درجات الاختبار
	ممتاز	جيد جدا	جيد	متوسط	ضعيف		
١٢	٠	٠	٠	٥	٧	١٢	١٩ - ٠
١٨	٠	٠	٥	٧	٦	١٨	٣٩ - ٢٠
٢٥	٠	٢	٧	٧	٩	٢٥	٥٩ - ٤٠
٨٠	٩	١١	٣٠	٢٠	١٠	٨٠	٧٩ - ٦٠
٣٥	٦	١٥	١٠	٤	٠	٣٥	٩٩ - ٨٠
٢٠	٥	١٠	٥	٠	٠	٢٠	١١٩ - ١٠٠
١٠	٣	٧	٠	٠	٠	١٠	١٣٩ - ١٢٠

ن = ٢٠٠

ن = ٢٠٠

٢ - وأن عدد اللاعبين الذين حققوا درجات على الاختبار (٣٩ - ٢٠) هو ١٨ لاعبا، منهم ٦ مستواهم ضعيف، ٧ متوسط، ٥ جيد، صفر جيد جدا، صفر ممتاز، وفقا لتقديرات نفس الخبراء.
..... وهكذا بالنسبة لبقية سطور الجدول.

ولاستكمال إعداد جدول التوقع فى المثال الحالى نقوم بتحويل التكرارات داخل الخلايا إلى نسبة مئوية percentage وذلك بضرب تكرار الخلية فى ١٠٠ وقسمة الناتج على مجموع تكرارات نفس الفئة كالتالى:

$$(بالنسبة الفئة الأولى) \quad \%.٤٢ = \frac{١٠٠ \times ٥}{١٢} \quad , \quad \%.٥٨ = \frac{١٠٠ \times ٧}{١٢}$$

$$\%.٢٨ = \frac{١٠٠ \times ٥}{١٨} \quad , \quad \%.٣٩ = \frac{١٠٠ \times ٧}{١٨} \quad , \quad \%.٣٣ = \frac{١٠٠ \times ٦}{١٨}$$

(بالنسبة للفئة الثانية)

..... وهكذا بالنسبة لجميع الخلايا

نقوم بوضع جميع النسب المئوية في جدول التوقع expectancy table التالي:
جدول التوقع: ويظهر العلاقة بين درجات اختبار المهارة في كرة القدم ودرجات الأداء على محك له خمسة مستويات (ن=٢٠٠).

المجموع	النسب المئوية لكل مستوي من مستويات المحك					فئات درجات الاختبار
	ممتاز	جيد جدا	جيد	متوسط	ضعيف	
٪١٠٠	٠	٠	٠	٤٢	٥٨	١٩ - ٠
٪١٠٠	٠	٠	٢٨	٣٩	٣٣	٣٩ - ٢٠
٪١٠٠	٠	٨	٢٨	٢٨	٣٦	٥٩ - ٤٠
٪١٠٠	١١	١٤	٣٨	٢٥	١٢	٧٩ - ٦٠
٪١٠٠	١٧	٤٣	٢٩	١١	٠	٩٩ - ٨٠
٪١٠٠	٢٥	٥٠	٢٥	٠	٠	١١٩ - ١٠٠
٪١٠٠	٣٠	٧٠	٠	٠	٠	١٣٩ - ١٢٠

يلاحظ من جدول التوقع الآتي:

- أن فئة اللاعبين الحاصلين على درجات الاختبار بين (١٩-٠) احتمال حصولهم على مستوى (جيد جدا) في المحك (القدرة المهارية في كرة القدم) هو صفر٪ .

- بينما نجد أن هذا الاحتمال يصل هذا الاحتمال إل ٪٧٠ بالنسبة للاعبين الحاصلين على درجات في الاختبار تقع بين (١٣٩ - ١٢٠).

- ومن ناحية أخرى يتضح من جدول التوقع السابق أن النسب المئوية للحاصلين على مستوى (جيد جداً) على المحك تزداد كلما زادت درجات الأداء على الاختبار، حيث كانت هذه النسب كالتالى:

٠٪، ٠٪، ٨٪، ١٤٪، ٤٣٪، ٥٠٪، ٧٠٪.

- كما يتضح من نفس الجدول أن النسب المئوية للحاصلين على مستوى (ممتاز) على المحك تزداد أيضاً كلما زادت درجات الأداء على الاختبار، حيث كانت هذه النسب كالتالى:

٠٪، ٠٪، ١١٪، ١٧٪، ٢٥٪، ٣٠٪.

(وهذه النتائج مجتمعة توحى بصدق الاختبار المقترح).

٥ - العوامل التى تؤثر على صدق الاختبار

يتأثر صدق الاختبار بالعديد من العوامل، هذه العوامل يمكن وضعها فى أربع مجموعات رئيسية هى:

- عوامل تتعلق بالاختبار نفسه.
 - عوامل تتعلق بتطبيق وحساب درجات الاختبار.
 - عوامل تتعلق باستجابات المفحوصين.
 - عوامل تتعلق بمجموعة التقنين والمحك.
- وفيما يلى شرح موجز لكل مجموعة من هذه العوامل:

أولاً : العوامل التى تتعلق بالاختبار نفسه ، وتشمل الآتى :

١ - طول الاختبار : Length of the test

يؤدى التطويل المتجانس Homogeneous Lengthening للاختبار إلى زيادة كل من ثبات وصدق هذا الاختبار، فعندما يصبح الاختبار أكثر طولاً يصبح أكثر ثباتاً وصدقا، فنحن نعرف أن صدق أى اختبار يتأثر بثبات هذا الاختبار، ونعرف أيضاً أنه نتيجة استخدام معادلة سيرمان - براون فإن ثبات الاختبار

يزداد كلما زاد طول الاختبار أو زاد عدد محاولاته . فاختبار تنطيط الكرة dribbling test واختبار التصويب shooting test فى كرة السلة - على سبيل المثال - يصبحان مؤشرين أكثر صدقا فى التنبؤ بالقدرة على الإنجاز فى كرة السلة كلما أعطيت للمفحوصين عدداً أكبر من المحاولات، كما أظهرت نتائج الدراسات والبحوث العلمية أن قياس سمك ثنايا الجلد Folds of skin باستخدام جهاز الكالير (المسك) skinfold caliper أربعة أو خمسة قياسات متتالية للمنطقة الواحدة يزيد من ثبات وصدق النتائج أفضل من أخذ القياس لمرة واحدة أو مرتين .

ويمكن حساب الصدق الجديد للاختبار وفقاً لخمس حالات هى :

الحالة الأولى : عندما يزيد طول الاختبار ويبقى المحك الخارجى كما هو :
ويستخدم لذلك المعادلة التالية :

$$\text{معادلة (٣ - ٤)} \quad \frac{ن \times س \text{ ص}}{\sqrt{ن + ن(١ - ن)}} = س(ن \text{ ص})$$

حيث إن :

$س(ن \text{ ص})$ = الارتباط بين المحك والاختبار الذى تم تطويله (ن) من المرات .

$س \text{ ص}$ = الارتباط بين المحك والاختبار قبل تطويله (الصدق التجريبي للاختبار) .

ن = عندما يزيد طول الاختبار (ن) مرة (عدد مرات زيادة طول الاختبار أو إعادة تكراره) .

١.١ = معامل ثبات الاختبار .

مثال (٣ - ٤):

لنفترض أن معامل صدق أحد الاختبارات كان ٥, ٠، وكان معامل ثباته هو ٠, ٤٠. فكم يكون صدق هذا الاختبار إذا زاد طوله ٤ مرات؟

الحل:

بالتعويض في المعادلة (٣ - ٤) ينتج أن:

$$\frac{0,50 \times 4}{0,40(1-4)4 + 4} \sqrt{\quad} = \text{س(ن ص)}$$

$$\frac{2}{(0,40)(3)4 + 4} \sqrt{\quad} =$$

$$\frac{2}{0,40 \times 3 \times 8} \sqrt{\quad} =$$

$$\frac{2}{9,6} \sqrt{\quad} =$$

$$\frac{2}{3,098} \sqrt{\quad} =$$

$$0,64557 =$$

$$0,65 =$$

يلاحظ أن الاختبار الذي زاد طوله ٤ مرات زاد معامل صدقه من ٥, ٠،

إلى ٥, ٦٥.

الحالة الثانية: عندما يزيد طول الاختبار ويزيد طول المحك الخارجى:

ويستخدم لذلك المعادلة التالية:

$$\text{معادلة (3-5)} \quad \frac{\overline{S^2}}{\sqrt{\left(\frac{1}{m} + \frac{1}{n} - 1\right) \left(\frac{1}{m} + \frac{1}{n} - 1\right)}} = \overline{S^2} \text{ (ن م)}$$

حيث إن :

$\overline{S^2} \text{ (ن م)}$ = الارتباط بين الاختبار بعد زيادته (ن) مرة، وبين المحك بعد زيادته (م) مرة.

$\overline{S^2}$ = معامل الارتباط بين الاختبار والمحك قبل الزيادة (الصدق التجريبي للاختبار).

ن = عدد مرات زيادة طول الاختبار أو إعادة تكراره.

م = عدد مرات زيادة طول المحك.

1.1 = معامل ثبات الاختبار.

2.2 = معامل ثبات المحك.

مثال (3-5):

لنفترض أن اختبار ما بلغ معامل ثباته 0.56، وكان معامل الارتباط بين هذا الاختبار والمحك الخارجى هو 0.65، فكم يبلغ معامل الصدق الجديد للاختبار إذا زاد طول كل من الاختبار والمحك ثلاث مرات، وكان معامل ثبات المحك هو 0.58.

الحل:

من المثال السابق يتضح أن:

$$\overline{S^2} = 0.65$$

$$3 = n$$

$$3 = m$$

$$1.1 = \overline{S^2}$$

$$0,58 = 2.2$$

وبالتعويض في المعادلة (3-5) ينتج أن:

$$\begin{aligned} 0,65 &= \sqrt{(0,58 + \frac{0,58-1}{3})(0,56 + \frac{0,56-1}{3})} \quad \text{س (ن م)} \\ 0,65 &= \sqrt{(0,58 + \frac{0,42}{3})(0,56 + \frac{0,44}{3})} \\ 0,92 &= \end{aligned}$$

مما سبق يتضح أن زيادة طول كل من الاختبار والمحك ثلاث مرات بالنسبة لطولهما الأصلي قد أدى إلى زيادة معامل الصدق التجريبي للاختبار من 0,65 إلى 0,92.

مثال (3-6):

في المثال السابق (3-5) المطلوب حساب معامل الصدق الجديد للاختبار إذا زاد طول الاختبار ثلاث مرات، وزاد طول المحك مرتين.

الحل:

$$\begin{aligned} 0,65 &= \sqrt{\text{س ص}} \\ 3 &= \text{ن} \\ 2 &= \text{م} \\ 0,56 &= 1.1 \\ 0,58 &= 2.2 \end{aligned}$$

بالتعويض في المعادلة (٣ - ٥) ينتج أن:

$$\begin{aligned}
 & \frac{0,65}{\left(0,58 + \frac{0,58-1}{2}\right) \left(0,56 + \frac{0,56-1}{3}\right)} \sqrt{\quad} = \text{س(ن م)} \\
 & \frac{0,65}{\left(0,58 + \frac{0,42}{2}\right) \left(0,56 + \frac{0,44}{3}\right)} \sqrt{\quad} = \\
 & \frac{0,65}{(0,58 + 0,21)(0,56 + 0,15)} \sqrt{\quad} = \\
 & \frac{0,65}{0,79 \times 0,71} \sqrt{\quad} = \\
 & \frac{0,65}{0,56 \cdot 9} \sqrt{\quad} = \\
 & \frac{0,65}{0,749} = \\
 & 0,87 =
 \end{aligned}$$

إذا كان معامل الصدق الجديد هو ٠,٨٧ ، حيث يلاحظ انخفاض هذا المعامل من ٠,٩٢ إلى ٠,٨٧ عندما كان طول المحك (م) مرتين بدلا من ثلاث مرات.

الحالة الثالثة : تحديد عدد مرات تطويل الاختبار للحصول على معامل صدق محدد:

قد يحدث أن يطلب من الباحث أن يكون للاختبار مستوى محدد يجب ألا يقل عنه معامل صدق الاختبار، حيث يستخدم لذلك المعادلة التالية:

$$\text{معادلة (٣ - ٦)} \quad \frac{\text{س(ن ص)}^2 \times (1.1 - 1)}{\text{س(ن ص)}^2 - \text{س(ص)}^2} = \text{ن}$$

حيث إن :

ن = عدد المرات التي يجب أن يزيد بها الاختبار.

س(ن ص) = مستوى معامل الصدق المطلوب تحقيقه.

س ص = معامل الارتباط بين الاختبار والمحك (الصدق التجريبي).

١.١ = معامل ثبات الاختبار.

مثال (٣-٧):

لنفترض أن معامل ثبات أحد الاختبارات كان ٠,٤٠ ، ومعامل صدقه ٠,٥٠ ، فما هي عدد مرات تطويل الاختبار لكي يصل معامل صدقه إلى ٠,٦٥ .

الحل:

بتطبيق المعادلة (٣-٦) ينتج أن:

$$ن = \frac{(٠,٦٥)^2(٠,٤٠-١)}{(٠,٤٠)^2(٠,٦٥) - (٠,٥٠)^2}$$

$$= \frac{٠,٦٠ \times ٠,٤٢}{(٠,٤٠)(٠,٤٢) - (٠,٢٥)}$$

$$= \frac{٠,٢٥٢}{٠,١٦٨ - ٠,٢٥}$$

$$= \frac{٠,٢٥٢}{٠,٠٨٢}$$

$$= ٣,٠٧$$

$$= ٣ مرات$$

كما سبق يلاحظ أنه لكي نحصل على المستوى المطلوب لمعامل صدق الاختبار الذي يساوى ٠,٦٥ ، فإنه يلزم تطويل الاختبار المستخدم ثلاث مرات.

الحالة الرابعة : عندما يزيد طول الاختبار إلى ما لا نهاية

عندما يزيد طول الاختبار إلى ما لا نهاية ، فإن الصدق المحسوب في هذه الحالة يعبر عن الحد الأقصى لصدق الاختبار the maximum validity ، حيث يحسب هذا الصدق باستخدام المعادلة التالية :

$$\text{معادلة (٣ - ٧)} \quad \frac{\text{س ص}}{1.1} = \text{س ص} \&$$

حيث إن :

$$\text{س ص} \& = \text{معامل صدق الاختبار بعد الزيادة}$$

$$\text{س ص} = \text{معامل صدق الاختبار قبل الزيادة}$$

$$1.1 = \text{معامل ثبات الاختبار}$$

مثال (٣ - ٨) :

إذا كان معامل ثبات أحد الاختبارات هو ٠,٨٢ ، ومعامل صدقه التجريبي هو ٠,٦٥ ، فكم يصل الحد الأقصى لهذا الصدق إذا زاد طول الاختبار إلى ما لا نهاية :

الحل :

بتطبيق المعادلة (٣ - ٧) ينتج أن :

$$\begin{aligned} \frac{0,65}{0,82} \sqrt{\text{س ص}} &= \text{س ص} \& \\ \frac{0,65}{0,906} &= \\ 0,72 &= \end{aligned}$$

إذاً الحد الأقصى لمعامل الصدق التجريبي لهذا الاختبار هو ٠,٧٢ .

الحالة الخامسة: عندما يزيد طول كل من الاختبار والمحك إلى ما لا نهاية:
ويستخدم ذلك المعادلة التالية :

$$\text{معادلة (٣-٨)} \quad \frac{\overline{س ص}}{\sqrt{٢.٢ \times ١.١}} = \&\&$$

حيث إن :

$$\text{معامل صدق الاختبار بعد الزيادة.} = \&\&$$

$$\text{معامل صدق الاختبار قبل الزيادة.} = \overline{س ص}$$

$$\text{معامل ثبات الاختبار.} = ١.١$$

$$\text{معامل ثبات المحك.} = ٢.٢$$

فإذا كان معامل ثبات المحك في المثال السابق رقم (٣-٨) هو ٠,٩٠ فكم
يكون معامل صدق الاختبار بعد الزيادة.

الحل :

بالتعويض في المعادلة (٣ - ٨) ينتج أن:

$$\frac{٠,٦٥}{\sqrt{٠,٩٠ \times ٠,٨٢}} = \&\&$$

$$\frac{٠,٦٥}{\sqrt{٠,٧٣٨}} =$$

$$\frac{٠,٦٥}{٠,٨٥٩} =$$

$$٠,٧٦ =$$

يلاحظ أن معامل الصدق زاد بزيادة طول كل من الاختبار والمحك إلى ما لا نهاية، فأصبح ٠,٧٦، بدلا من ٠,٦٥ .

٢ - عدم وضوح التعليمات Unclear Direction

عندما تكون تعليمات الاختبار المعطاة للمفحوصين غامضة بمعنى أن تكون غير واضحة وغير محددة، فإن ذلك يؤدي إلى اختلاف المفحوصين في تفسير هذه التعليمات مما يؤثر على صدق الاختبار، فتعليمات الاختبار غير الواضحة لا تمكن المفحوص من كيفية الاستجابة على وحدات الاختبار أو المقياس، مما قد يشجع على الاعتماد على التخمين Guessing، مما يترتب عليه بالتالى الحصول على معاملات صدق منخفضة.

٣ - صعوبة قراءة الكلمات والجمل (في اختبارات الورقة والقلم)

Reading vocabulary and sentence too difficult

أظهرت التجارب العلمية فى مجال القياس النفسى أن الكلمات والجمل المعقدة التراكيب بالنسبة لمستوى المفحوص الذى يعطى له الاختبار سوف تؤثر على صدق الاختبار، وبناء على ذلك ينصح علماء القياس النفسى أن تصاغ الكلمات والجمل فى اختبارات الورقة والقلم فى ضوء معيارين أساسيين هما:

- الاستيعاب القرائى للمفحوص Reading comprehension

- الذكاء Intelligence

٤ - عدم ملاءمة مستوى صعوبة فقرات الاختبار

Inappropriate level difficulty of the test items

عندما تكون فقرات الاختبار سهلة جداً أو صعبة جداً، فإن الاختبار يفقد قدرته على التمييز فيما بين المفحوصين، ومن ثم يكون له معامل صدق منخفض.

٥- الغموض The Ambiguity

عندما يتم صياغة الجمل (الأسئلة) التى يتضمنها الاختبار بدون عناية وتدقيق فى المعانى والألفاظ ودلالاتها، فإنها تحمل المفحوص إلى حالة من الاضطراب وعدم القدرة على تفسير الجملة. وقد أظهرت نتائج الدراسات والبحوث العلمية أن المفحوصين ذوى المستويات المتميزة يتأثرون بهذه المشكلة أكثر من المفحوصين ذوى المستويات غير المتميزة، مما يفقد الجملة (الفقرة) قدرتها على التمييز، وربما يجعل قدرتها على التمييز تكون فى الاتجاه السلب (أى فى صالح المفحوصين غير المتميزين).

٦ - الاختبار قصير أكثر مما ينبغى Test too short

من المعروف أن اختبارات الورقة والقلم تتضمن عينة من كثير من الأسئلة التى يجب أن توجه إلى المفحوص، فليس من المعقول أن يتضمن الاختبار كل الأسئلة التى يجب أن توجه إلى المفحوص لأسباب تتعلق بالإمكانات والوقت المتاح للتطبيق وغيرهما. وبناء على ذلك يلجأ علماء القياس إلى إعداد اختبارات الورقة والقلم بحيث تتضمن عينة من الأسئلة تكون ممثلة للسمة أو القدرة المطلوب قياسها أفضل تمثيل. وقد لوحظ أن صدق الاختبار يتأثر سلبا كلما كانت عينة الأسئلة التى يتضمنها الاختبار قليلة العدد.

٧ - عدم ملاءمة تنظيم الفقرات Improper arrangement of items

للحصول على معاملات صدق مرتفعة لاختبارات الورقة والقلم يجب تنظيم فقرات الاختبار (الجمل) بحيث توضع الفقرات السهلة أولا يليها بعد ذلك الفقرات الصعبة، لأنه فى حالة وضع الفقرات الصعبة فى بداية الاختبار سوف يجعل المفحوص يستغرق وقتا أطول مما ينبغى للإجابة على هذه الفقرات، مما يترتب عليه عدم وجود وقت كاف للإجابة على

الفقرات السهلة - وبخاصة في الاختبارات المقيدة بزمان مثل اختبارات :
التحصيل والذكاء - مما يؤثر على صدق الاختبار.

ثانيا : العوامل التي تتعلق بتطبيق وحساب درجات الاختبار :

يمثل أسلوب تطبيق وحساب درجات الاختبار أحد أهم العوامل التي تؤثر على صدق النتائج التي يتم الحصول عليها. فعدم الالتزام بشروط التطبيق، والوقت المخصص للإجابة (الأداء)، وكيفية تقدير الدرجات الخام أو المحولة باستخدام مفتاح التصحيح كلها عوامل تؤثر في صدق الاختبار.

ثالثا : العوامل التي تتعلق باستجابات المفحوصين :

تتأثر استجابات المفحوصين أثناء موقف الأداء على الاختبار بالعديد من العوامل، من هذه - العوامل على سبيل المثال لا الحصر - الخوف من موقف الاختبار والذي يجعل المفحوص (وبخاصة صغير السن) غير قادر على الاستجابة على الأسئلة بصورة طبيعية، كما أن الاضطراب الانفعالي emotional disturbance يؤثر هو الآخر على استجابة المفحوصين، كما تمثل الدافعية motivation أحد أهم العوامل التي تؤثر على استجابات المفحوصين وبخاصة بالنسبة لاختبارات الأداء الأقصى maximum performance كاختبارات القدرات البدنية والعقلية واختبارات التحصيل وغيرها.

وفي هذا الخصوص يبين كرنباخ Cronbach, L.J. 1970م أن اتجاه الاستجابة response set على العبارات يمثل أحد العوامل التي تؤثر على صدق الاختبار، فعندما يكون ميل المفحوصين نحو نمط محدد من الاختبارات متناغما Consistent tendency، فإنه يلزم في هذه الحالة أن يكون عدد العبارات (الجملة) الصحيحة true مساوياً لعدد العبارات الخاطئة false، ففي الاختبارات التي تتطلب الاستجابة بصح (✓) أو

خطأ (x) يميل بعض المفحوصين إلى الإجابة بصح (✓) عندما لا يعرفون الإجابة الصحيحة للفقرة، بينما يميل البعض الآخر إلى الإجابة بخطأ (x). فعندما يتضمن الاختبار عدداً كبيراً من العبارات (الجمل) الصحيحة، فإن مثل هذا الاختبار يعطى فرصة أكبر للفئة الأولى من المفحوصين على حساب الفئة الثانية، ومن ثم يصبح من الضروري أن تكون عبارات الاختبار نصفها صحيحاً والنصف الآخر خطأ.

رابعاً : العوامل المتعلقة بمجموعة التقنين والمحك

من المعروف أن صدق الاختبار يختص بمجموعة محددة من الأفراد من حيث الجنس، والعمر الزمني، والحالة التدريبية، والتعليمية، والصحية، والعرق Race، والمستوى الاقتصادي الاجتماعي إلى غير ذلك من المحددات التي تصف لنا مجموعة الأفراد التي أعد لها الاختبار في الأصل.

ولما كان صدق الاختبار يتأثر بالعديد من العوامل المرتبطة بخصائص مجموعة التقنين التي طبق عليها كل من الاختبار والمحك مثل: الجنس، والسن، ومستوى القدرة، والحالة التعليمية، والخلفية الثقافية، وغيرها. لذا يصبح من الضروري استخدام الاختبار وفقاً لشروط التقنين الذي تتضمنها كراسة التعليمات الخاصة بالاختبار test manual، لأن عدم الالتزام بأي شرط من هذه الشروط قد يضعف من صدق الاختبار، أو يجعل الاختبار غير صادق بالمرّة.

فطبيعة صدق الاختبار تتوقف على خصائص Characteristics، مجموعة (عينة) التقنين التي أعد لها الاختبار. وقد أشرنا من قبل إلى أن مدى السمة أو القدرة Range of ability المقيسة يعد من العوامل التي تؤثر على صدق الاختبار، فعندما تكون مجموعة التقنين غير متجانسة تجانساً شديداً بالنسبة للسمة أو القدرة المقيسة، فإن هذا بدوره يؤدي إلى الحصول

على درجات لها انتشار واسع spread scores بالنسبة لكل من الاختبار والمحك مما يزيد من صدق الاختبارات وبخاصة الصدق المرتبط بالمحك .

وعلى عكس ما سبق، فعندما تكون مجموعة التقنين متجانسة تجانساً شديداً بالنسبة للسمة أو القدرة المقيسة، فإن ذلك يؤدي إلى توزيع الدرجات في انتشار ضيق مما يؤدي إلى انخفاض صدق الاختبارات المرتبطة بالمحك (لاحظ أن الصدق المرتبط بالمحك يعتمد في حسابه - في معظم الحالات - على معامل ارتباط بيرسون الذي يعتمد على الانتشار الواسع لتوزيع درجات أى متغيرين) (*).

ومن ناحية أخرى تؤثر الفروق الثقافية Cultural differences فيما بين المجتمعات المختلفة على صدق الاختبار وبخاصة الاختبارات التي تستخدم اللغة المكتوبة (اختبارات الورقة والقلم)، فالاختبار الذي أعد لثقافة محددة ربما يصبح غير صادق بالنسبة لثقافة أخرى، وذلك بسبب الاختلافات في القيم والعادات والتقاليد والمستوى الاقتصادي الاجتماعي بين الثقافتين، هذا من ناحية، ومن ناحية أخرى فإن نقل الاختبار من لغته الأصلية (اللغة الأم)، إلى اللغة الجديدة (عن طريق الترجمة) قد يفقد الاختبار بعض خصائصه مما يستلزم عدم الثقة في صدق الاختبار الجديد إلا بعد التأكد من ذلك عملياً.

ويتفق معظم علماء القياس أن الاختبارات عبر - ثقافية - cross cultural tests التي يتم إعدادها في ثقافات مختلفة تعد اختبارات شاملة، ومن ثم لا يتأثر صدقها بالفروق الثقافية، من هذه الاختبارات - على سبيل المثال - اختبارات الذكاء وقوائم الشخصية واختبارات القدرات البدنية والمهارات الحركية في المجال الرياضي.

* راجع: كتاب «الإحصاء الوصفي في علوم التربية البدنية والرياضة، دار الفكر العربي، القاهرة، ٢٠٠٢م (للمؤلف).

٦ - العلاقة بين صدق الاختبار وثباته

الثبات والصدق وجهين لعملة واحدة، أو بعدين لشيء واحد، هذا الشيء هو كفاءة (فعالية) الاختبار efficiency فيما يقيسه، فإذا كان الصدق هو ارتباط الاختبار ببعض المحكات الخارجية المستقلة، وكان ثبات الاختبار هو ارتباط الاختبار بنفسه، فإن الاختبار الذي لا يرتبط بنفسه لا يتوقع له أن يرتبط مع محكات خارجية مستقلة. ومن ناحية أخرى فإن الاختبار الذي يظهر ثباتاً منخفضاً فإنه لا يتوقع أن يكون له صدقاً مرتفعاً، لهذا يعتمد الصدق على الثبات وبخاصة بالنسبة للاختبارات المتجانسة (الموحدة) Homogeneous tests. أما في حالة الاختبارات غير المتجانسة، فإن صدق الاختبار قد يظهر مرتفعاً دون أن تكون هناك حاجة لكي يظهر الاختبار ثباتاً مرتفعاً كما في حالة ثبات الاتساق الداخلي للاختبار أو المقياس، حيث يعزى ذلك إلى أن كل جزء من أجزاء الاختبار يقيس وظيفة مستقلة مما يترتب عليه ظهور ثبات الاتساق الداخلي منخفضاً لكون وحدات الاختبار لا ترتبط بعضها مع بعض، بينما تظهر هذه الوحدات ارتباطات مرتفعة مع بعض المحكات الخارجية المستقلة، في مثل هذه الحالة نستطيع القول بأن هذا الاختبار الذي أعطى ثباتاً منخفضاً عن طريق الاتساق الداخلي له صدق مرتفع مع بعض المحكات الخارجية.

كما سبق يتضح أن الثبات يعد أحد المتطلبات الواجب توافرها في أداة القياس، ولكنه لا يعد بالضرورة، أحد الشروط اللازمة للصدق، فقد يكون صدق الاختبار مرتفعاً عن ثباته، ولكنه ليس أكثر ارتفاعاً من مؤشر الثبات index of reliability لكون مؤشر الثبات يضع الحد الأقصى للارتباط الذي يمكن أن يكون بين الاختبار والمحك.

ومن الملاحظ في هذا الشأن أن المهتمين ببناء الاختبارات والمقاييس يستهدفون دائماً الحصول على معاملات ثبات وصدق مرتفعة لكل اختبار

يقومون بإعداده ، مع ملاحظة أن أهداف وأساليب بناء كل من ثبات وصدق الاختبار تبدو متضاربة، فقد تتعارض متطلبات الحصول على ثبات مرتفع للاختبار مع متطلبات الحصول على صدق مرتفع لنفس الاختبار، فالثبات المرتفع يتطلب فقرات متساوية في درجة الصعوبة، كما يتطلب معاملات ارتباط بينية مرتفعة (فيما بين الفقرات) Intercorrelations ، في حين يتطلب الصدق المرتفع لنفس الاختبار فقرات مختلفة في درجات صعوبتها، كما يتطلب معاملات ارتباط بينية (فيما بين الفقرات) منخفضة .

وللتغلب على هذه المشكلة يقترح كل من توكر Tucker ١٩٤٦م وجيلفورد Guilford ١٩٥٦م أنه لكي يمكن الحصول على صدق مرتفع بالإضافة إلى ثبات مرتفع للاختبارات، فإنه يجب أن تتراوح الارتباطات البينية للفقرة inter-item correlations (معاملات ارتباط كل فقرة مع غيرها من الفقرات) من (٠,١ إلى ٠,٦) لضمان الحصول على معاملات ثبات وصدق مرضية للاختبار أو المقياس .

معجم إنجليزي - عربي للمصطلحات

A		Criterion Relate validity	
Abstract Constructs			الصدق المرتبط المحك
	تكوينات مجردة	Construct validity	صدق التكوين الفرضي
Adolescence	مرحلة المراهقة	Contents	المحتويات (محتويات)
American psychological association		Curricular validity	الصدق المنهجي
	الجمعية الأمريكية لعلم النفس	Criterion method	طريقة المحك
Ambiguity	الغموض	Correlation matri	مصنوفة ارتباط
Academic success	النجاح الأكاديمي	Continuous data	بيانات متصلة
Arithmetic test	اختبار في الحساب	Composite score	الدرجة المركبة
Achievement testing	اختبار التحصيل	Contrasted group	المجموعات المتضادة
Aptitude tests	اختبارات الاستعدادات	Childhood	مرحلة الطفولة
Appearance	المظهر الخارجي للاختبار	Commulation	شيوخ (اشتراكيات)
B		D	
Bivariate distribution		Discrete data	بيانات متقطعة
	التوزيع التكراري الثنائي	Discriminat validation	صدق التمييز
Biserial Correlation	الارتباط الثنائي	Domain	محال
C		E	
Clusters	تجمعات	Experimental validlity	الصدق التجريبي
Critical thinking	التفكير الناقد	External	خارجي
Constructs	تكوينات	External Criterion	محك خارجي
Concurrent validity	الصدق التلازمي	Experts	الخبراء (المحكمون)
Content validity	صدق المحتوى (المضمون)	Expectancy tables	جداول التوقع
Cros - Cultural tests		Emotional disturbunce	
	اختبارات عبر - ثقافية		الاضطراب الانفعالي

F		L	
Fidelity	الدقة	Linear combination	تراكيب خطية
Factorial validity	الصدق العاملي	Loadings	تشبعات الاختبارات على العوامل
Face Validity	الصدق الظاهري		
Factorial analysis	التحليل العاملي	M	
Factor Loading	تشبع الاختبارات على العامل	Multivariate analysis	التحليل متعدد المتغيرات
		Multiple correlation	الارتباط المتعدد
H		Mechanical comprehension	الاستيعاب الميكانيكي
Hypothetical Construct	تكوين افتراضى	Mechanical ability	القدرة الميكانيكية
Homogeneous Lengthening	التطويل المتجانس فى الاختبار	Maximum validity	الحد الأقصى لصدق الاختبار
I			
Intrinsic Validity	الصدق الداخلى (الذاتى)		
Independent	مستقبل		
Impiycal validity	الصدق العملى (الإحصائى)	U	
Inded of validity coefficeent	مؤشر معامل الصدق	Uncorrelated	غير مترابطة
Irrelevant	ليس له صلة بالموضوع	Unclear directions	عدم وضوح التعليمات
Index of reliability	مؤشر الثبات	F	
Internal criterion	معك داخلى	Predictive efficiency	الكفاءة التنبؤية للاختبار
Intercorrelations coefficients	معاملات الارتباط البينية	Pure factors matrix	مصفوفة العوامل الثقة
Interrelationships	العلاقات التبادلية	Principle Components	المكونات الأساسية
		Phi-Coefficient	معامل فاي
		Present standing	الوضع الراهن
		Predictive tests	اختبارات تنبؤية

	R		Saturations	تشبعات الاختبارات على العوامل
Response set		اتجاه الاستجابة	T	
Relationship		علاقة	True - False	الصحيح - الخطأ
Representativeness validity		صدق التمثيل	Trait Validity	صدق السمة
Relevance validity		الصدق التحليلي	Theoretical construct	تكوين نظري
Rho		معامل ارتباط سبيرمان لفروق الرتب	Test efficiency	كفاءة الاختبار
Rating scales		مقاييس تقدير	Tournament	نظام الدوري
Reading comprehension		الاستيعاب القرائي	Tetrachoric Correlation	الارتباط الرباعي
			Trained groups	مجموعات مدربة
	S			
situations		مواقف	V	
Statistical validity		الصدق الإحصائي	Validity	الصدق
Spatial ability		القدرة المكانية	Validity Coefficient	معامل الصدق

مراجع الفصل الثالث

- ١ - سعد عبد الرحمن (١٩٩٨م) : القياس النفسى - النظرية والتطبيق . دار الفكر العربى ، القاهرة .
- ٢ - صفوت فرج (٢٠٠٠م) القياس النفسى . مكتبة الأنجلو المصرية . القاهرة .
- ٣ - صفوت فرج (١٩٨٠م) : التحليل العاملى فى العلوم السلوكية . دار الفكر العربى ، القاهرة .
- ٤ - محمد حسن علاوى ، ومحمد نصر الدين رضوان (٢٠٠٠م) : القياس فى التربية الرياضية وعلم النفس الرياضى . دار الفكر العربى ، القاهرة .
- 5 - Anastasi, A. (1990) : Psychological testing, New york : Macmillan.
- 6 - Cronbach, L.J. (1970) : Test Validation . In R.L. Thorndike (Ed.) , Educationul measurement. washington, DC: American Council on Education.
- 7 - Kaplan, R.M., and Saccuzzo, D.P. (1988) : Psychological Testing, Principles, Applecations, and Issues. Brooks Cole Publishing Company, California.
- 8 - Miller, D.K. (1994) : Measurement - By the physical educators : why an How. WCB Brown and Benchmark ppublishers.
- 9 - Morris, L.L., Gibben, C.T.F. , and Lindheim, E. (1987) : How to measure performance and use tests. saGe publications, Inc., London New dehil.
- 10 - Singh, A.K. (1986) : Tests, Measurements and Research methods in Behavioural Sciences. Tata McGraw- Hill Publishing Company Limited , New Delhi.