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

١ - مفهوم الصدق

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

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

وقد عرف العديد من المتخصصين الصدق بطرق مختلفة، فقد ذكرت أنستازى Anastasi ١٩٩٠م. - على سبيل المثال - أن صدق الاختبار يعنى «ما الذى يقيسه الاختبار ؟ وكيفية صحة هذا القياس»، ويعرف ليندكويست Lindquist ١٩٥١م صدق الاختبار على أنه «الدقة التى 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 تبنوا هذا الاتجاه منذ العام ١٩٧٤م.

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) التى 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 والتى 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 أولا 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 كافيه الظاهرة التى صمم لقياسها، بحيث لايقيس 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 فى المجال الرياضى كثيرا 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 صادقا 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 شيئا 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 مثلا .

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 مترا كمقياس لسرعة الانتقال، فإنه يلزم تقدير الزمن باستخدام ساعات إيقاف تحسب الزمن بالثوانى (ث) و $\frac{١}{١٠٠}$ 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 وبجزء 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 هى:**

- 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 التى من أجلها تستخدم أدوات القياس measuring Instruments، حيث وضعت هذه الأغراض فى **ثلاث فئات عريضة هى:**

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

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

٣ - الاستدلال على قوة المفحوص فى القدرة أو السمة أو الخاصية التى تظهرها أداة القياس، كأن يقوم أحد الباحثين بالتعرف على درجات مجموعة من المفحوصين بالنسبة لبعض المفاهيم المجردة مثل: الذكاء والانبساط والانطواء والعصابية وعما 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 التى تنجز بها أداة القياس الغرض المقترن بها وفقاً لطبيعة الاختبار المعطى للمفحوصين».

ومن الطبيعى أن يحدث أن تصمم أداة ما من أدوات القياس لتلائم أكثر من غرض واحد من الأغراض الثلاثة المذكورة، حيث يقتضى الحال فى 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 المنهجى Curricular Validity والصدق التحليلى 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 الأنثروبوترية 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 وجمعيا، 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. وفى 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 المثل - 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 (الكاليبر) يستخدم

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

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

أنواع المحكات المستخدمة للحكم على صدق المحتوى (المضمون)

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

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

الصدق المرتبط بالمحك:

ويعرف هذا الصدق باسم: الصدق التجريبى 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) التى تقيس نفس الظاهرة (القدرة أو السمة أو غيرهما). 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:

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

ثانيا : الفترة الزمنية التى يتم فيها جمع البيانات عن المحك:

فى الصدق التلازمى يتم جمع بيانات المحك فى نفس الوقت الذى يطبق فيه الاختبار، بمعنى تزامن تطبيق كل من الاختبار والمحك، أما فى الصدق التنبؤى فيتم جمع بيانات المحك بعد فترة زمنية طويلة من تطبيق الاختبار.

الصدق التلازمى

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

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

ويشترط فى الصدق التلازمى 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 الأكسجين VO_{2max} كمحك لحساب صدق بعض اختبارات الجرى مثل: الجرى ١ ميل والجرى $1\frac{1}{2}$ 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 الأكسجين لكونها - أى الاختبارات - أسهل فى التطبيق، ولا تستغرق وقتا طويلا، كما أنها أقل تكلفة بالمقارنة بالطرق المعملية المعقدة التى تستخدم لتقدير الحد الأقصى لاستهلاك الأكسجين.

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

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

٣ - ويستخدم الأداء فى المباريات game performonce وتقديرات الخبراء experts كمحكات لحسب الصدق التلازمى للعديد من اختبارات المهارات الرياضية Sport Skill tests .

جدول (٣ - ١)

أمثلة للصدق المرتبط بالمحك فى المجال الرياضى

الاختبار test	المحك criterion	معامل الصدق validity Coefficient
- اختبار كيمب وفينست للمهارة فى التنس (١٩٦٨م)	نظام الدورى الكامل	٠,٨٤ للمبتدئين * ٠,٩٣ للمتقدمين
- اختبار هوايت للسرعة فى الإرسال فى التنس (١٩٦٨م)	نظام الدورى الكامل	٠,٨٩ للمبتدئين
- اختبار الجرى - المشى ١ $\frac{1}{2}$ ميل (جيتشل وكيركندال وروبنز ١٩٩٠م).	الحد الأقصى لاستهلاك الأكسجين vo_2max	٠,٩٢
- اختبار الجرى - المشى ١ ميل (كلين وآخرون ١٩٨٧م)	الحد الأقصى لاستهلاك الأكسجين vo_2max	٠,٨٨
- نسبة الدهون باستخدام جهاز الكاليبر.	وزن الجسم تحت الماء	٠,٨٠ للرجال من سن ١٧-٣٧ (ريكوروبينك 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.

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

مفهوم المحك: Criterion

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

- أن المحك عبارة عن معيار أو ميزان قياسى يستخدم للحكم على صدق مقاييس أخرى جديدة.
- أن الثقة فى صدق المحك تكون كبيرة جداً.
- أنه يمكن النظر إلى متغير المحك Criterion variable على أنه متغير يمكن التنبؤ به من خلال متنبِّئ Predictor واحد (اختبار أو مقياس) أو من خلال الربط بين مجموعة من المتنبآت (اختبارات أو مقاييس).
- قد يكون المحك مجموعة من الدرجات أو الآراء أو الأحكام أو الأداءات على بعض الأجهزة أو الأدوات أو الاختبارات.

المحك المؤلف من مجموعة من العناصر: Composite criterion

هو محك يتكون نتيجة الربط بين الدرجات المشتقة (المحولة) من عدد من الاختبارات، حيث تستخدم درجة هذا المحك 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 كمياً.

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 عشارية، هذا بالإضافة إلى التحيز فى تقدير الأوزان النسبية لمكونات أو عناصر المحك.

- **تحريف المحك**: Criterion distortion

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

الصدق التنبؤى

وقد صمم كصدق عملى (اختبارى) impirical validity أو إحصائى Statistical Validity حيث يشير إلى ترابط الاختبار بالمحك للتنبؤ ببعض الأحداث الممكنة فى المستقبل، بمعنى قدرة الاختبار على التنبؤ بأداء الفرد فى موقف مستقبلى إذا كان هذا الموقف المستقبلى له علاقة بما يقيسه الاختبار.

ويستهدف الصدق التنبؤى فى المجال الرياضى الآتى:

- وصف الأداء الراهن على المحك describing present criterion perfomance.
- التنبؤ بالأداء المستقبلى على المحك predicting future criterion performance .

وصف الأداء الراهن على المحك:

عندما يستخدم الصدق التنبؤى لوصف الأداء الراهن على المحك، فإنه يشبه فى هذه الحالة الصدق التلازمى أو يكون امتدادا له، حيث يمكن تقدير

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

فعندما توصل جيتشل وآخرون Getchell et all. ١٩٧٧م - على سبيل المثال - إلى وجود علاقة ارتباط مقدارها (٠,٩١٥-) بين اختبار الجرى - المشى واحد ونصف ميل واختبار السير المتحرك treadmill test والذى يقيس بطريقة مباشرة ودقيقه الحد الأقصى لاستهلاك الأكسجين VO_2max، فقد أوصى بناء على ذلك أنه يمكن استخدام اختبار الجرى - المشى واحد ونصف ميل كمتنبئ (وسيلة تنبئ) بالحد الأقصى لاستهلاك الأكسجين.

وبناء على العلاقة السابقة بين اختبار الجرى - المشى ١$\frac{١}{٢}$ ميل واختبار السير المتحرك، والتى بلغت (٠,٩١٥-)، **فإنه يمكننا استخلاص الآتى:**

- أن لاختبار الجرى - المشى ١$\frac{١}{٢}$ ميل صدق تلازمى بلغ (٠,٩١٥-).

- أنه يمكن إحلال اختبار الجرى - المشى ١$\frac{١}{٢}$ ميل بدلاً من اختبار السير المتحرك.

ويبقى السؤال المهم وهو:

هل يمكننا التنبؤ بقيم الحد الأقصى لاستهلاك الأكسجين 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 الأكسجين 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:

ى = ٩٨,٣ - ٤,١٨٢ س

حيث إن :

ى = القيمة المقدرة للحد الأقصى لاستهلاك الأكسجين

س = الزمن الفعلى لأداء اختبار الجرى - المشى واحد ونصف ميل (بالدقائق).

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

ى = ٩٨,٣ - (٤,١٨٢ × ١٢)

= ٩٨,٣ - ٥٠,١٨٤

= ٤٨,١

ومن الملاحظات الجديرة بالاهتمام فى هذا الخصوص هو أن الاختبار قد يكون له معامل ارتباط كبير مع المحك، ومع ذلك لا يكون متنبأ 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 ناحيـة أخرى فقد يقتـرح بعض المدربين أن الوقـــوف على الـرأس Headstand تعد وسيلة تنبؤ جيدة بالقدرة على العدو Sprinting ability، وقد جاء هذا الاقتراح نتيجة وجود مـعامل ارتباط (معامل صدق) بلغ (-٨٥, ٠) بين الوقوف على الـرأس (كمتنبئ) والقـدرة على العـــدو (كمحك). ويـرى علماء القيـاس أنه لا يجب الأخذ فقط بمعامل الارتباط لحـسم هذه العلاقة، لأنه إذا كان الوقـوف على الرأس وسيلـة تنبؤ حقـيقية بالـنسبة للقـدرة على العدو، فإن تدريب الأفراد على الوقوف على الرأس يجب أن يتبعه بالضرورة تحسن فى قدرتهم عـلى العدو وهو ما لم تؤكدة الدراسات والبحـوث العلمية المتخصصة وذلك على الرغم من وجود ارتباط مرتفع بين المتغيرين.

فقط لاحظ المتخـصصون أن أى طريقة من طرق حسـاب الارتباط يمكن أن تعطى مـعاملات ارتباط مـقبولة فـيما بين المتـغيرات، حيـث تكشف هذه المعاملات عن درجـة العلاقة بين الاختبارات كـمقاييس مشتركـة لعنصر عام Common element، إلا أنه يلاحظ أن معاملات الارتباط (معاملات الصدق) قد لا تعطينا معلـومات عن درجة تأثير متغير على آخـر، فمثلا يرتبط الوزن بالطول ارتباطا عاليا، لأن الأفراد طوال القامة يميلون لكى يكونوا أثقل وزنا، إلا أن الزيادة فى الوزن قد لا تحـدث بالضرورة نتيجـة الزيادة فى الطول، لذا نستـخلص أن العلاقـة بين المتنبئ (الاخـتبار) والمحك هى عـلاقـــة وظيفـية Functional relationship أكثر منها علاقة سببية Causal.

مما سبق نستخلص الآتى:

١ - أن كلا من الصدق التلازمى والصدق التنبؤى يستخدمان المحك.

٢ - أن معادلـة الانحدار هى أهم ما يـميز بين الصـدق التلازمى والوظيـفة الأولى من وظائف الصـدق التنبؤى، فالصـدق التلازمى يكتفى بـتقدير العلاقـة بين الاختبار والمحك، بـينما يستلـزم الصدق التنبؤى (الوظـيفة الأولى) إعـداد معـادلة الانحـدار كأداة تسـتخـدم درجات الأداء عـلى 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 فى معادلة الانحدار (التنبؤ).

ولزيادة الثقة فى معادلة الانحدار ينصح علماء القياس بالآتى:

أولا : يجب أن تكون عينة التقنين كبيرة نسبيا بحيث لا تقل عن ٢٠٠ فرد، وأن يتم اختيارها من مجتمعها الأصلى بالأسلوب العشوائى.

ثانيا : 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.

ثالثا : 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 المستقبلى على المحك:

عندما يستهدف الصدق التنبؤى وصف المحك أو التنبؤ المستقبلى 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) أولا، 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 أولا، ثم نتابع سلوك المفحوص فيما بعد بالنسبة لبعض المتغيرات التى يتنبأ بها الاختبار. فإذا اتفق مستوى إنجاز المفحوص على متغيرات المحك ومستوى إنجازه 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 المحسوب 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 (ى - y) وذلك من خلال الأداء على واحد أو أكثر من الاختبارات والتى 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 لفعالياتهم 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 فى العمل (ى - y) هو متغير الثقة بالنفس (س - x) ، حيث تمكن الباحث من صياغة معادلة التنبؤ التالية:

$$ى = ٣,٥٠ + ٦,٢ س$$

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

$$ى = ٣,٥٠ + ٦,٢ (١٤)$$

$$= ٣,٥٠ + ٣٦,٤$$

$$= ٨٦,٧$$

ويرى علماء القياس أنه يمكن النظر إلى معامل الارتباط الذى يصف درجة العلاقة الخطية بين الاختبار والمحك، على أساس أن المحك (ى - 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 (ى - y) والترابط الخطى لاثنين أو أكثر من المتغيرات المستقلة (س١، س٢ ، ... 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:

$$ى = ب + ب_{١}س_{١} + ب_{٢}س_{٢} + ...$$

حيث إن :

ى = المحك (المتغير التابع)

س = متغيرات مستقلة (س١ ، س٢ . . إلخ)

ب = 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:

س١ (الثقة بالنفس)

س٢ (القدرة على التفاعل مع الآراء والاتجاهات المختلفة)

س٣ (قوة الإرادة)

وكانت معادلة التنبؤ المتعدد باستخدام هذه المتغيرات الثلاثة كالتالى:

$$ى = ٢٨,٢ + ٣,١س_{١} + ٠,٧س_{٢} + ١,٦س_{٣}$$

فإذا كانت درجات أحد المعلمين بالنسبة س١ = ٨، س٢ = ١٣، س٣ = ١١، فإنه يمكن التنبؤ بالكفاية المهنية لهذا المعلم عن طريق تطبيق المعادلة السابقة كالتالى:

$$ى = ٢٨,٢ + ٣,١ (٨) + ٠,٧ (١٣) + ١,٦ (١١)$$

$$= ٢٨,٢ + ٢٤,٨ + ٩,١ + ١٧,٦$$

$$= ٧٩,٧$$

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 المتنبآت (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 - Predicotors) 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 يتهربون من استخدام هذا النوع من أنواع الصدق، فى حين يضطر بعض الباحثين إلى استخدام هذا

الصدق فـى الحالات التى يجـدون فيها أن كـلا من صدق المحتـوى والصدق المرتبط بالمحك غير مناسبين لتقنين أدواتهم.

ويعرَّف صدق التكويـن الفرضى بمسميات أخرى منها الـصـدق العاملى Factorial Validity وصدق السـمة trait validity، وتحدد أنسـتازى ١٩٩٠م، مفهـوم صدق التكوين الـفرضى على أنه « المدى الـذى يمكن أن يمتـد إليه الاختبـار كمقيـاس لمفهوم افـتراضى 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، ففى مجـال التربية الرياضية - علـى سبيل المثال - يمكن أن نطـلق مصطلح التكوينات Constucts على المهارات أو القدرات أو السمـات التى نفترض أنها تشكل فى مـجموعهـا اختباراً واحـداً يقيس ظاهرة معـينة أو خاصيـة مميزة، فعندمـا نقوم بتصمـيم اختبار ما لـقياس ظاهرة معـينة أو خاصية مميـزة 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 التـكوينات الفرضية للقدرة المهارية فى كرة السلة تتكون من المحاورة بالكرة والتصويب والتمرير، فإذا كان الاختبار المعد لقياس هذه القدرة يشتمل على وحدات لقياس كل تكوين من التكوينات الفرضية السابقـة (مهارات المحاورة والصويب والتـمرير) فإن هذا يدل على أن الاخـتبار صادق ككل فى قياس القدرة المهارية فى كرة السلة، وذلك على أساس أن هناك وحـدة اختـبار لكل مهـارة من المهـارات الثلاث السابقـة، وأن الـربط بين هـذه الوحدات يشكل اختباراً صادقاً لقياس القدرة المهارية فى كرة السلة.

وفى مجـال قياس بـعض القدرات والسـمات النفـسية المـرتبطة بالنـشاط الرياضى يشـير صدق التكويـن الفرضى إلى مدى قـياس الاختبار لـتكوينات فرضيـة محددة، من هذه التـكوينات الفرضيـة - على سبيل المثـال - 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 ضمنا 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 للظهور وتقبلها لوصف سلوك المفحوص والتنبؤ به . فإذا قلنا - على سبيل المثال -

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

مما سبق يتضح أن «مصطلح التكوينات» عبارة عن ضرب من المفاهيم a sort of concepts التى يتم اقتراحها بشكل نظامى يتضمن التعريف بها بوضوح وبحيث تكون مؤسسة على بيانات ومعلومات تعتمد على الملاحظة أو التجربة. وفى هذا الصدد تشير نانلى Nunnally ١٩٧٠م إلى أن «التكوينات» عبارة عن فروض توضح لنا أن تنوعات فى السلوك سوف يرتبط كل منها بالآخر فى الدراسات التى تتناول الفروق الفردية Individual differences فى الإنسان، وفى ضوء ذلك يمكن تعريف صدق التكوين الفرضى 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 الفرضى 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 التى أعدت لقياس الذكاء، فإنه يمكن من خلال دراستنا ومعرفتنا عن الذكاء وضع التنبؤات التالية:

- الدرجات الخام على الاختبار سوف تزيد بزيادة العمر الزمنى (لقد أكدت الدراسات والبحوث العلمية أن الذكاء يزداد مع زيادة العمر حتى سن ١٦ سنة تقريباً».
- درجات الاختبار سوف تتنبأ بالنجاح فى التحصيل المدرسى school achievement (وهى علاقة تنبؤية أظهرتها نتائج الدراسات والبحوث التربوية والنفسية على مدى سنوات طويلة).
- درجات الاختبار سوف ترتبط إيجابيا مع التقديرات التى يضعها معلم الفصل لذكاء التلاميذ.
- درجات الاختبار سوف تميز بين مجموعات التلاميذ المتفوقين وغير المتوفقين دراسياً فى المدارس.
- درجات الاختبار سوف لا تتأثر بدرجة جوهرية بالتدريس الذى يحدث فى 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 الزمنى - فإننا نستخلص أن الاختبار إما أن يكون غير صادق لقياس الذكاء، أو أن هناك خطأ ما فى التكوينات النظرية.

طرق الحصول على صدق التكوين الفرضى:

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

لا توجد طريقة واحدة مناسبة لبنـاء أو إعداد صدق التكوين الفرضى كما هو متبع بالنسبة لصدق المحتـوى (المضمون) أو الصدق المرتبط بالمحك. وقد قدم كرونـباخ ومـيل Cronbach and Meehl ١٩٥٥م مـجمـوعة من الإجـراءات لحساب صدق التكوين الفرضى هى:

١ - تحليل العمليات العقلية لإعداد فقرات الاختبار:

وتتضمن هذه الطريقة تحليل العمليات العقلية المطلوب قياسها والتى يجب أن تتضمنها وحـدات الاختبار، حيث يتضمن هذا الـتحليل 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 مـن الأفراد. فدرجـات مجمـوعات الأفـراد يمكن أن تخـتلف بالنسبـة لاختبار ما وفقاً لـلعمر الزمنى للمـجموعات age groups ووفقاً للنوع (بنين أو بنات)، ووفقاً للحالة التدريبية (مجموعات مدربة وأخرى غيـر مدربة)، ووفقـاً لدرجة التكـيف الاجتمـاعى (مجمـوعات تتـمتع بدرجات عالـية من التكيف وأخرى تتـمتع بدرجات أقل)، فعلـى سبيل المثال تزداد معظم القدرات (البدنية والعقلية) مع تقدم العمر الزمنى، وأن البنيـن يحصلون على درجـات أعلى من الدرجات الـتى تحصـل عليـها البنـات فى اختـبارات مـحددة مـثـل اخـتـبـار الاستـيعـاب الميكانـيكى Mechanical Comprehension وهو اختبار يقيـس الفهم الآلى، فى حين يحصل البنات على درجات أكبر من الدرجـات التى يحصل عليها البنين فى بـعض الاختـبـارات مثـل اختـبار الاسـتـيعـاب القـرائى Reading Comprehension.

وأيضا ولنفس الأسباب، فإنه يتوقع أن تميز درجات بعض الاختبارات بين المجموعـات وفقا للاختلاف فى درجـات بعض المتغيرات مـثل 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 كرونباح 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 الاختبار Multidimenionality.

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 إعداده 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 hyothesized structure للاتجاهات نحو النشاط البدنى، ثم قام باختبارها باستخدام التحليل العاملى، 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 الداخلى للقائمة.

وبالإضافة إلى التحليل العاملى - كأسلوب لحساب الصدق الداخلى لها - فقد قام كينيون بالإضافة إلى ذلك بحساب العلاقة بين التكوينات Constructs وبعض المحكات الخارجية بغية التحقق من الصدق الخارجى للقائمة. فقد قام كينيوى 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 الدورى والتنفسى كما يقيسها اختبار بالك the Balke treadmill test، وقد افترض كينيون أن العلاقة بين درجات هذا البعد سوف تكون مرتفعه مع نتائج اختبار بالك، وقد جاءت النتائج لتثبت صحة هذه الفرضية وتؤكد الصدق الخارجى لهذا البعد من أبعاد القائمة.

مفهوم الصدق التقاربى وصدق التمييز:

يوضح كامبل وفيسك Campbell and Fiske ١٩٥٩م ، أن تحقيق كل من الصدق التقاربى Convergent validation وصدق التمييز Discriminant Validation ، يعد من الإجراءات المهمة لبناء صدق التكوين الفرضى لأى اختبار من الاختبارات، وأن من العوامل الأساسية بالنسبة لصدق التكوين الفرضى هو أن يرتبط الاختبار مع غيره من المقاييس التى 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 الحسابى arithmetical reasoning test أو اختبار للجبر algebra test، فى 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 التقاربى»، وأنه عندما يكون للاختبار ارتباطات ضعيفة مع مقاييس لايجب 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.

مما سبق يتضح أن الصدق التقاربى 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 التقاربى للاختبار يتحقق عندما تكون هناك علاقات ارتباط جوهرية 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 التقاربى 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 التى استخدمت لحساب الصدق التقاربى فى المجال الرياضى تلك الإجراءات التى اتبعها راينر مارتنز Reiner Martens عند إعداده لاختبار قلق المنافسة الرياضية Sport Competition Anxiety Test (SCAT).

فقد قام مارتنز بمقارنة نتائج اختباره (SCAT) بنتائج مقاييس للقلق العام general anxiety ، حيث قام بفحص الصدق التقاربى 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 ±٣, ٠ إلى ± ٥, ٠) وذلك لوجود عوامل مشتركة بين اختبار قلق المنافسة الرياضية وأى اختبارات أخرى للقلق العام.

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

وللتحقق من صدق التميز لاختبار قلق المنافسة الرياضية، قام مارتنز بدراسة العلاقة بين اختباره 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 علـى درجات مرتفعة فى الدافع نحـو النجاح فى الإنجاز the motive to achieve success.

وقد توصـل مارتنز إلى صحـة الفرضيـة السابقة، حـيث كشفت نـتائج الدراسات التى قام بهـا عن معاملات ارتباط بين دافعية الإنجـاز 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 التقاربى وصدق التمييز:

«قدم كامـبل وفيسك ١٩٥٩م» تصـميماً تجريـبياً خاصاً لـدراسة كل من الصدق التـقاربى وصدق التمـييز، هذا التصمـيم يشير إلى مـا يعرف باسم: المصفوفة متعددة الطرق ومتعددة السمات.

Multitrait - Multimethod Matrix (MTMM)

ووفقا لما هو موضح من الاسم، فإن استخدام هذه المصفوفة يتطلب سمتين traits أو أكثر، يتـم قياسها بطريـقتين methods أو أكثر، حيـث يتم تحليل معاملات الارتبـاط البينية intercorrelations داخل المصفوفة لتـحقيق كل من الصدق التقاربى وصدق التمييز للاختبار وفقا لمبدئين مهمين هما:

١ - أن قياس نفـس السمة باستخـدام طرق مختلفة سـوف يؤدى إلى وجود معامـلات ارتباط مرتفـعـة فيمـا بين المقاييس المـستخدمـة، مما يؤكـد الصدق التقاربى.

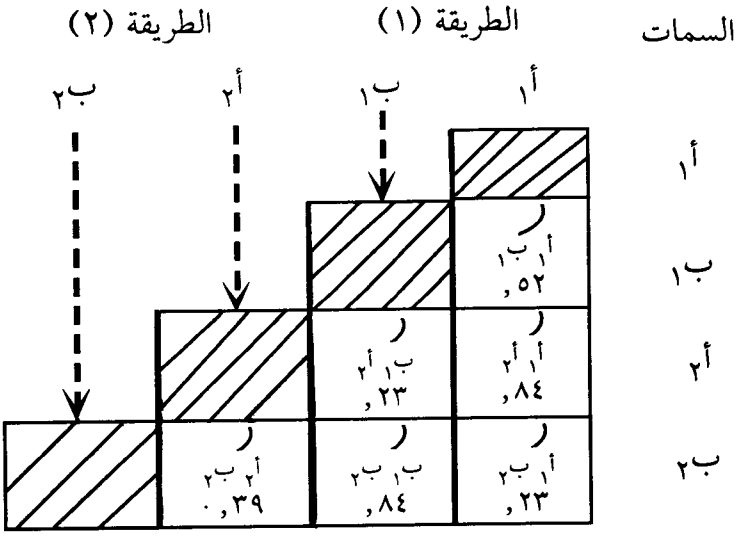
٢ - أن استخدام نفس الطرق لقياس سـمات مختلفة سوف يؤدى إلى وجود معـاملات ارتباط مـنخفضـة فيمـا بين هذه المقـاييس، مما يؤكـد صدق التمييز، حيث يستخلص من هذين المبدأين الآتى:

«أن اتفاق السـمة واختلاف الطـريقة يؤدى إلى معـاملات ارتباط مرتـفعة» وأن «اختلاف السمة واتفاق الطريقة يؤدى إلى معاملات ارتباط منخفضة».

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

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

(أ) العدوان Agression، (ب) العصابية Neuroticism ، وأنه قد تم قياس هاتين السمتين بطريقتين أعطيت الأرقام (١ , ٢)، وكانت هاتين الطريقتين (١) قائمة لقياس الشخصية ، (٢) اختبار إسقاطى Progective test ووفقا لهذه المعطيات فإنه يمكن وضع التصميم التالى:



شكل (٣ - ١)

المصفوفة متعددة الطرق ومتعددة السمات

ويلاحظ من الشكل (٣ - ١) الآتى:

(أ١) تدل على درجة العدوان على قائمة الشخصية

(أ٢) تدل على درجة العدوان على الاختبار الإسقاطى

(ب١) تدل على درجة العصابية على قائمة الشخصية

(ب٢) تدل على درجة العصابية على الاختبار الإسقاطى

وعندما حسبت معاملات الارتباط بين هذه السمات جاءت كالتالى:

$\overset{\frown}{أ_١ب_١}$ = ٥٢, ٠ (العدوان على قائمة الشخصية مع العصابية على نفس القائمة).

$\overset{\frown}{أ_١أ_٢}$ = ٨٤, ٠ (العدوان على قائمة الشخصية مع العدوان على الاختبار الإسقاطى).

$\overset{\frown}{أ_١ب_٢}$ = ٢٣, ٠ (العدوان على قائمة الشخصية مع العصابية على الاختبار الإسقاطى).

$\overset{\frown}{ب_١أ_٢}$ = ٢٣, ٠ (العصابية على قائمة الشخصية مع العدوان على الاختبار الإسقاطى).

$\overset{\frown}{ب_١ب_٢}$ = ٨٤, ٠ (العصابية على قائمة الشخصية مع العصابية على الاختبار الإسقاطى).

$\overset{\frown}{أ_٢ب_٢}$ = ٣٩, ٠ (العدوان على الاختبار الإسقاطى مع العصابية 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:

$\overset{\frown}{أ_١أ_٢}$ = ٨٤, ٠ (ويمثل درجة العدوان على قائمة الشخصية مع درجته على الاختبار الاسقاطى).

$\overset{\frown}{ب_١ب_٢}$ = ٠,٨٤ (ويمثل درجة العصابية على قائمة الشخصية مع درجتها على الاختبار الاسقاطى).

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

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

$\overset{\frown}{أ_١ب_١}$ = ٠,٥٢ (ويمثل درجة العدوان على قائمة الشخصية مع درجة العصابية على نفس القائمة).

$\overset{\frown}{أ_٢ب_٢}$ = ٠,٣٩ (ويمثل درجة العدوان على الاختبار الإسقاطى مع درجة العصابية على نفس الاختبار).

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

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

($\overset{\frown}{أ_١أ_٢}$ = ٠,٨٤ ، $\overset{\frown}{ب_١ب_٢}$ = ٠,٨٤ وهى أكبر من

$\overset{\frown}{أ_١ب_١}$ = ٠,٥٢ ، $\overset{\frown}{أ_٢ب_٢}$ = ٠,٣٩)

٣ - أنواع أخرى من الصدق

بالإضافة إلى الأنواع الأساسية للصدق والتى (سبق الإشارة إليها) توجد ثلاثة أنواع أخرى هى:

الصدق الظاهرى Face Validity

يحدث فى كثير من الأحيان أن يخلط بعض الباحثين بين صدق المحتوى (المضمون) والصدق الظاهرى، فالصدق الظاهرى ليس صدقا حقيقيا بالمعنى العلمى لكلمة «الصدق»، ولكنه يعنى ببساطة أن الاختبار أو المقياس يبدو صادقاً فى صورته الظاهرية.

ويتأسس الصدق الظاهرى 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 الذى ينتمى إليه الاختبار، حيث يطلب منهم قراءة الاختبار، ثم الحكم على الصدق من خلال الآتى:

١ - فحوى بنوده ومدى علاقتها بما تقيسه منطقيا.

٢ - طبيعة الفقرة - ووضوحها.

٣ - تعليمات الاختبار.

٤ - الزمن المخصص للاختبار . .إلخ.

مما سبق يتضح أن الصدق الظاهرى 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 الظاهرى إجراء إضافى 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 الظاهرى إعداد تقرير علمى يعطى للباحث الثقة والطمأنينة فى مدى صلاحية الاختبار قبل قيامه بإجراءات التقنين العملية (التجريبية) للاختبار، فعندما يقترح الخبراء استبعاد بعض فقرات الاختبار التى تبدو غير ذات صلة بالموضوع الذى تقيسه Irrelevant فإن ذلك يوفر على الباحث جهداً كبيراً ويزيد من الثقة فى صدق الاختبار.

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

الصدق الذاتى Intrinsic Validity

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

وبما أن الثبات يقوم فى 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 الذاتى بحساب الجذر التربيعى لمعامل ثبات الاختبار، حيث يمكن حسابه كالتالى:

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

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

$$= \sqrt{٠,٦٥}$$

إذاً معامل الصدق الذاتى = ٨١ , ٠

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:

$$\text{الصدق الذاتى} = \sqrt{\text{معامل الثبات}} = \sqrt{٠,٨٤} = ٠,٩٢$$

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 مناسبته 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

إنطلاقـا من أن مثل هذه الاخـتبارات لا تتـمتع بالقـدرة على إظهـار الفروق الفردية فى الاستعدادات البدنية للأطفال غير الرياضيين (العاديين).

فعندما تعجـز أداة القياس عن إظهار الفروق فيما نقـيسه، فإنه يمكن أن نصف هذه الأداة بأنهـا غير حساسة لما نقـيسه. فعندما يحـاول شخص ما أن يقيس مقدار ما فقده من وزن جسمه نتيجة انتظامه فى برنامج لإنقاص الوزن، فإن وحدات الوزن التى تسـتخدم الكيلوجرامات أو الأرطال (الميـزان الطبلية) قد لا تكون حـساسة لإظهار الفـروق فى وزن الجسم قبل وبعـد الانتظام فى البرنامج، وذلك لكون هذا الفرق قد يكون محدوداً بجرامات قليلة لا يظهرها ميزان الطبلية لعدم حساسيته لهذا الفرق.

وتمثل حساسـية أدوات القياس أحد الشـروط المهمة التى يجب مـراعاتها عند اختـيار هذه الأدوات، وذلك لما لهـذا الشرط من أهمـية بالنسبـة لإظهار مقـدار التقدم والتـحسن فى بعـض المجالات المهمـة المتعلقـة بالتعلـم وتذكر (استرجاع) المعلومات، واللياقة الوظيفيـة والبدنية، والقياسات المورفولوجية، وإنقاص الوزن وغيرها.

ومن ناحية أخرى فقـد تلعب حساسية أدوات القياس دوراً مهـماً بالنسبة لاستثـارة دافعية وحـماس جماهيـر الناس عندما يشـعرون بنجاحـهم وتحسن مستوياتهم، فعندما نستـخدم لقياس الأداء مقاييس عامة gross measures غير مناسبة لإظهـار الفروق داخل الفرد أو بين الأفراد، فإنه يتـوقع أن تفشل مثل هذه المقاييـس فى إظهار تلك الفروق، أو ربمـا تظهرها بشكل مـحدود جداً. فعندمـا يطبق اختبـار تمرير كرة السلة علـى دائرة واحدة مرسومة عـلى حائط (بحيث تعطى الدرجة (١) عند إصابة الهدف، والدرجة (صفر) إذا لم تصب الكرة الهدف).

فإن هذا الاختـبار يكون غير حسـاس لإظهار التقدم الذى يحـققه لاعب كرة السلة من ناحية، والفروق بين هذا اللاعب وغيره من اللاعبين من ناحية

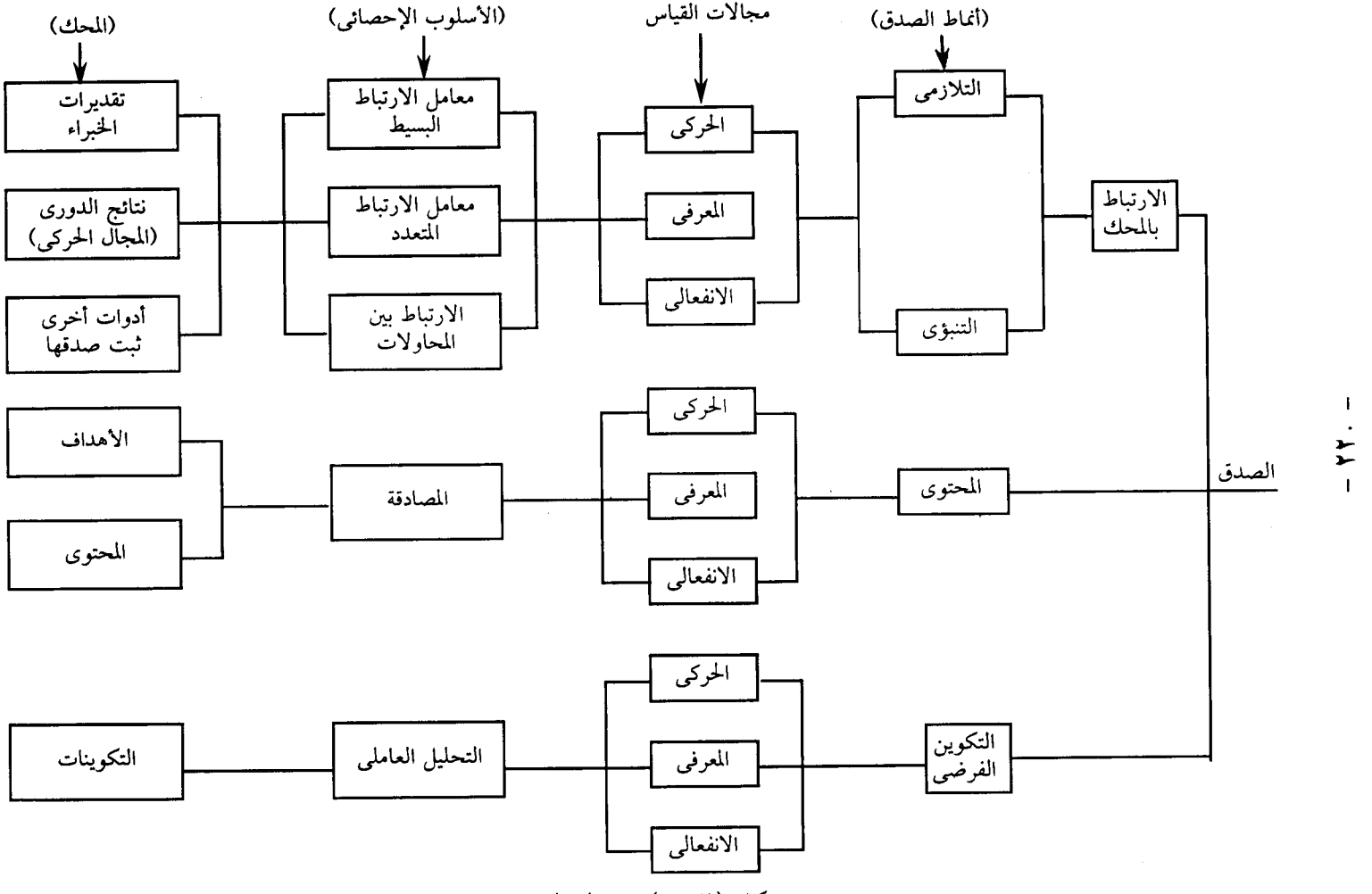
أخرى، وذلك فى حالة إعطاء الدرجة (١) عندما تصيب الكرة الهدف، والدرجة (صفر) إذا خرجت الكرة بعيدا عن الهدف.

ويكون نفس الاختبار السابق حساساً لقياس الدقة والسرعة فى التمرير فى 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 فى الدرجات كلما تحسن مستوى المفحوص، وأيضا قدرته على إظهار الفروق بين الأفراد من ذوى القدرات المختلفة، حيث يعنى ذلك أن الاختبار قد صمم بطريقة تسمح له بأن يظهر التقدم والتحسن الذى يحدث للمفحوص فى مدى واسع من الدرجات، فقياس زمن العدو ٣٠ مترا للاعبى المنتجب القومى فى العدو باستخدام ساعات إيقاف تسجل الزمن بالثوانى فقط، يعد قياساً غير حساس لإظهار التقدم الذى يحدث لهؤلاء اللاعبين، وأيضا غير حساس لإظهار الفروق بينهم فى هذا الاختبار، ومن ثم يلزم فى مثل هذه الحالة استخدام ساعات إيقاف تقوم بتسجيل الزمن بالثوانى و $\frac{١}{١٠٠}$ 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.



شكل (٣ -٢) خريطة الصدق

٤ - طرق حساب الصدق

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

طريقة استطلاع آراء الخبراء (المحكمين) Expert's Judgement

يتم اختيار الخبراء وفقا لشروط يحددها الباحث والمشرف ويمكن أن تتضمن المؤهل الدراسى، التخصصى الدقيق، عدد سنوات الخبرة فى التخصص... إلى غير ذلك من الشروط، وفى جميع الحالات يجب أن يراعى بالنسبة للخبراء الآتى:

- أن يكونوا من الدارسين ذوى الخبرة فى مجال التخصص الذى 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 الظاهرى معا، حيث يطلب من الخبير (المحكم - المتخصص) إبداء رأيه فى مكونات الاختبار أو المقياس، والتى يرى أنها تشكل من وجهة نظره أهمية خاصة بالنسبة للظاهرة المقيسة، حيث يعتمد فى ذلك على التفكير المنطقى أو ما يعرف بالتفكير الناقد Critical thinking والخبرة الذاتية.

٢خطوات تطبيق هذه الطريقة:

أولاً : بالنسبة للاختبارات البدنية والحركية (الاختبارات العملية) :

- تحليل القدرة أو المهارة، إلى مكوناتها الأساسية principle components التى تشكل فى 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 التى كشف عنها التحليل.

مثال (٣ - ١):

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

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

- إعداد استمارة استطلاع رأى تتضمن المكونات الـ ١٧ موضحاً بها معنى القدرة الحركية وكل مكون من المكونات التى 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 يـتفق مع وجهـة نظره وذلك وفق مقـياس تقدير عـشارى 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 التدريجات من (١-١٠) ثم تحسب النسبة المئوية لكل تدريج.

جدول (٣ - ٢)

المكونات	التقديـــــرات									
١ - القوة العضلية	١	٢	٣	٤	٥	٦	٧	٨	٩	١٠
٢ - القدرة العضلية	١	٢	٣	٤	٥	٦	٧	٨	٩	١٠
٣ - السرعة الانتقالية	١	٢	٣	٤	٥	٦	٧	٨	٩	١٠
٤ - سرعة رد الفعل	١	٢	٣	٤	٥	٦	٧	٨	٩	١٠
٥ - الرشاقة	١	٢	٣	٤	٥	٦	٧	٨	٩	١٠
٦ - المرونة	١	٢	٣	٤	٥	٦	٧	٨	٩	١٠
٧ - التوازن	١	٢	٣	٤	٥	٦	٧	٨	٩	١٠
٨ - الدقة	١	٢	٣	٤	٥	٦	٧	٨	٩	١٠
٩ - التوافق البدنى	١	٢	٣	٤	٥	٦	٧	٨	٩	١٠
١٠ - التحمل العضلى	١	٢	٣	٤	٥	٦	٧	٨	٩	١٠
١١ - التحمل الدورى التنفسى	١	٢	٣	٤	٥	٦	٧	٨	٩	١٠
١٢ - الايقاع	١	٢	٣	٤	٥	٦	٧	٨	٩	١٠
١٣ - التوقيت	١	٢	٣	٤	٥	٦	٧	٨	٩	١٠
١٤ - البنيان الجسمانى	١	٢	٣	٤	٥	٦	٧	٨	٩	١٠
١٥ - الطول	١	٢	٣	٤	٥	٦	٧	٨	٩	١٠
١٦ - الوزن	١	٢	٣	٤	٥	٦	٧	٨	٩	١٠
١٧ - القابلية للتعلم الحركى	١	٢	٣	٤	٥	٦	٧	٨	٩	١٠

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

$$ص = ل + \frac{٠,٥ - مجـ ن}{نَ}$$ معادلة (٣ - ١)

حيث إن:

ص = درجة صدق المكون (العنصر)

ل = الحد الأدنى للدرجة الوسيطية

٠,٥ = مقدار ثابت

مجـ ن = مجموع النسب التى تقع قبل الدرجة الوسيطية

نَ = النسبة الوسيطية

فإذا جاءت استطلاعات رأى الخبراء بالنسبة لمكون «القوة العضلية» فى مثالنا الحالى كالتالى:

(مكون القوة العضلية)

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

(عدد المحكمين = ٣٠)

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

- حساب نسبة المحكمين الخاصة بكل تقدير وتتم بقسمة عدد المحكمين أسفل كل تدريج مقسوماً على العدد الكلى للمحكمين كالتالى:

$\frac{٠}{٣٠}$ = ٠ ، $\frac{١}{٣٠}$ = ٠,٠٣ ، $\frac{٢}{٣٠}$ = ٠,٠٧ وهكذا بالنسبة لبقية التقديرات.

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

ترتيب الوسيط = $\frac{ن + ١}{٢}$

$$= \frac{١٠ + ١}{٢}$$

$$= ٥,٥٠$$

إذاً قيمة الوسيط $= \frac{٥ + ٦}{٢}$

$$= ٥,٥$$

- حساب الحد الأدنى للدرجة الوسيطية كالتالى:

ل = ٥,٥ - ٠,٥

= ٥

- نقوم بحساب مجـ ن كالتالى:

٠ + ٠,٠٣ + ٠,٠٧ + ٠,٠٧ + ٠,١٠ = ٠,٢٧

إذاً مجـ ن = ٠,٢٧

- نقوم بحساب النسبة الوسيطية نَ فنجد أن تساوى ٠,١

- نقوم بالتعويض فى المعادلة (٣ - ١) كالتالى:

$$ص = ٥ + \frac{٠,٥ - ٠,٢٧}{٠,١}$$

$$= ٥ + \frac{٠,٢٣}{٠,١}$$

= ٥ + ٢,٣

= ٧,٣

إذاً درجة صدق مكون (عنصر) القوة العضلية يساوى ٧,٣ .

وإذا جاءت استطلاعات رأى الخـبراء بالنسبة لمكون «القدرة الـعضلية» فى نفس المثال كالتالى:

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

(عدد المحكمين = ٣٠)

من هذه النتيجة نجد أن:

الوسيط = ٥,٥ (نفس النتيجة السابقة)

ل = ٥ (نفس النتيجة السابقة)

ومجـ ن = ٠,٠٣ + ٠,٠٧ + ٠,١٣ + ٠,١

= ٠,٣٣

نَ = ٠,١

وبالتعويض فى المعادلة (٣ - ١) ينتج أن:

$$ص = ٥ + \frac{٠,٥ - ٠,٣٣}{٠,١}$$

$$= ٥ + \frac{٠,١٧}{٠,١}$$

= ٥ + ١,٧

= ٦,٧

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

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

إذاً نسبة ٣٣٪ = $\frac{١٧ \times ٣٣}{١٠٠}$

= ٥,٦

= ٦ مكونات (عناصر)

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

١- القوة العضلية Muscular strength

٢- القدرة العضلية Muscular power

٣- التوافق البدنى Body coordination

٤- السرعة الانتقالية Running speed

٥- الرشاقة (سرعة تغيير الاتجاه) Agility

٦- التحمل الدورى التنفسى Cardiovascular-respiratory endurance

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

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

وتستخدم هذه الطريقة فى حالة تقنين اختبارات التحصيل والذكاء وقوائم الشخصية وبالنسبة للعديد من المقاييس فى 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 التى تسبق العبارات على النحو التالى:

- فيما يلى 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 (×) أسفل كلمة (أوافق بدرجة كبيرة) وأمام رقم العبارة.

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

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

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

- وقد وضعت أوزان العبارت بغض النظر عن اتجاه العبارة كالتالى:

(أوافق بدرجة كبيرة)	(أوافق)	(أعارض)	(أعارض بدرجة كبيرة)
٤	٣	٢	١

- 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، ففى 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 عشارى كما فى المثال (٣ - ١) كالتالى:

العبارات **مقياس التقدير**

١ - تهتم التربية الرياضية ببناء العضلات فقط. . . .

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

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

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

٣ - . . . . . .

٤ - . . . . . .

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

طريقة المحك Criterion Method

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

المحك الداخلى 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 الشىء، 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 الداخلى للمقياس أو الاختبار، حيث يكثر 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 فى حالة عدم إمكانية استخدام النوع الثانى من المحكات وهو: المحك الخارجى، حيث يقرر معظم علماء القياس النفسى أنه يمكن الاعتداد 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 يقنن 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 وفقا 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)، حينئذ يمكن استخدام معامل ارتباط بيرسون Pearson للدرجات الخام لتقدير العلاقة بين الدرجة الخام التى حصل عليها كل لاعب طبقا 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 كمحكات 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 سبيرمان 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.

مما 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 (ثانيا).

(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 هى:

١ - معامل ارتباط بيرسون Pearson وهو الأكثر شيوعاً.

٢ - الارتباط الثنائى Biserial Correlation .

٣ - الارتباط الرباعى Tetrachoric Correlation .

٤ - معامل فاى Phi - Coefficient .

٥ - الارتباط المتعدد Multiple Correlation .

٦ - ارتباط الرتب (سبيرمان) Rank- Order Correlation .

حيث تستخدم هذه المعاملات فى الحالات التالية:

- يستخدم الارتباط الثنائى عندما تقسم درجات الاختبار أو المحك فى فئتين مثل (ناجح - راسب) أو (صح - خطأ) بمعنى أن تكون الإجابة على أى من الاختبار أو المحك ثنائية البعد.
- يستخدم الارتباط الرباعى ومعامل فاى عندما تكون درجات كل من الاختبار والمحك مقسمة فى 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 فى هيئة رتب، بمعنى أن هذا المعامل يستخدم فى حالة البيانات المتقطعة discrete data أو حالة البيانات المتصلة التى يمكن وضعها فى شكل رتب.
- يستخدم معامل ارتباط بيرسون (وهو الأكثر شيوعاً) فى حالة البيانات المتصلة Continuous data للعينات التى 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 وفقا 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 الذى يساوى ٠,٨٠ فأكثر يمكن قبوله.

* يلاحظ أن مربع معامل الصدق يمثل كمية التباينات الشائعة بين الاختبار والمحك، فمعامل الصدق الذى 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 منخفضا 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 التجريبى تستخدم المعادلة التالية:

$$\text{س ص}_2 = \frac{\text{س ص}_1}{\sqrt{\text{س س} \times \text{ص ص}}}$$ معادلة (٣ - ٢)

حيث إن :

س ص٢ = معامل صدق الاختبار بعد التعديل.

س ص١ =معامل صدق الاختبار قبل التصحيح (معامل الصدق التجريبى).

س س = معامل ثبات الاختبار.

ص ص = معامل ثبات المحك.

مثال : (٣ -٢):

استخدم أحد الباحثين اختبار الشد لأعلى pull - ups كمحك خارجى لحساب الصدق التجريبى 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 التجريبى = ٠,٧٠٤

معامل ثبات الاختبار = ٠,٩٦٠

معامل ثبات المحك = ٠,٩٥٠

وبالتعويض فى المعادلة (٣ -٢) ينتج أن:

$$\overline{\overline{س ص}}_{٢} = \frac{٠,٧٠٤}{\sqrt{٠,٩٦٠ \times ٠,٩٥٠}}$$

$$= \frac{٠,٧٠٤}{\sqrt{٠,٩١٢٩٥}}$$

$$= \frac{٠,٧٠٤}{٠,٩٥٥٤٨}$$

$$= ٠,٧٣٧$$

إذاً معامل الصدق التجريبى لاختبار الانبطاح 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 والتى تستخدم المعادلة التالية :

$$\overline{\overline{س ص}}_{٢} = \frac{\overline{س ص}_{١}}{\sqrt{\overline{ص ص}}}$$

معادلة (٣ - ٣)

حيث إن :

$\overline{\overline{س ص}}_{٢}$ = معامل صدق الاختبار بعد التصحيح.

$\overline{\overline{س ص}}_{١}$ = معامل صدق الاختبار قبل التصحيح (معامل الصدق التجريبى).

$\overline{ص ص}$ = معامل ثبات المحك.

مثـــال (٣-٣):

قام أحد البـاحثين بإعداد اخـتبار لقياس مـهارة الدقة فى التمـرير فى كرة القدم، وقد بلغ معامل ثبات هذه الاختبار ٨٨, ٠ ، ولتقدير الصدق التجريبى لهذا الاختـبار استخدام الباحث مـجموعة من الخبراء لتـقدير درجة الأداء فى هذه المهارة لمجـموعة من اللاعـبين أثناء الأداء الفعـلى فى المباريات الرسـمية كمحك خـارجى، وعندما قام البـاحث بحساب الصـدق التجريبى للاخـتبار المقتـرح وجد أنه يسـاوى ٦٥, ٠ وهو معـامل صدق منخـفض، فهل يـأخذ الباحث بهذا المعامل أم يرفضه إذا كان معامل ثبات المحك هو ٦٧, ٠ ؟

الحـــل:

بالتعويض فى المعادلة (٣ - ٣) ينتج أن:

$$\text{س ص}_{٢} = \frac{٠,٦٥}{\sqrt{٠,٦٧}}$$

$$= \frac{٠,٦٥}{٠,٨١٩}$$

$$= ٠,٧٩٤$$

إذاً معـامل الصدق التـجريبى للاخـتبار المـقترح يسـاوى ٧٩٤, ٠ وهـو معامل مقبول.

ملحوظة مهمة:

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

- أن القيـمة $\sqrt{٠,٩٥٠ \times ٠,٩٦١}$ والتى تـساوى ٩٥٥٤٨, ٠ (٩٥٥, ٠ تقريبا) تعـنى أن معامل الارتباط بين الاختبـار والمحك فى المثال (٣ - ٢) لن يتعدى بأى حال من الأحوال القيمة ٩٥٥, ٠ .

- وأن القيـمة $\sqrt{٠,٦٧}$ والتى تسـاوى ٨١٩, ٠ تعنـى أن معـامل الارتبـاط بين الاختبار والمحك فى المثال (٣ - ٣) لن يتعدى بأى حال من الأحوال قيمة ٨١٩, ٠ .

طريقة المقارنة الطرفية:

وتستخدم هذه الطريقة لتقدير الصدق المرتبط بالمحك (الصدق التلازمى)، وهى تعتمد على قدرة الاختبار على التمييز بين طرفى السمة أو القدرة التى يقيسها، حيث يستخدم اختبار ت-t لدراسة دلالة الفروق بين المتوسطات الحسابية بدلاً من استخدام الارتباط. حيث تتم المقارنة فى هذه الطريقة بأسلوبين هما:

الأسلوب الأول (خاص بكل من الاختبار والمحك الخارجى) :

أ - ويتم بمقارنة الثلث الأعلى فى درجات الاختبار بالثلث الأعلى فى درجات المحك.

ب - مقارنة الثلث الأدنى فى درجات الاختبار بالثلث الأدنى فى درجات المحك. فإذا كان الفرق بين المتوسط الحسابى للثلث الأعلى للاختبار والمتوسط الحسابى للثلث الأعلى للمحك، وكان الفرق بين المتوسط الحسابى للثلث الأدنى للاختبار والمتوسط الحسابى للثلث الأدنى للمحك غير دال إحصائيا كان ذلك مؤشرا على أن الاختبار صادق، بمعنى أن الاختبار يكافؤ المحك الخارجى من حيث البناء.

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

$$\bar{س}_{١} = \bar{س}_{٢}$$

أو $\bar{س}_{١} - \bar{س}_{٢} = ٠$

الأسلوب الثانى (خاص بالاختبار فقط):

ويتم بمقارنة درجات الثلث الأعلى للمحك بدرجات الثلث الأدنى للاختبار، وبحيث يستخدم اختبار ت - t لحساب دلالة الفرق بين

المتوسطين، فإذا كان هذا الفرق دالا إحصائيا كان الاختبار صادقاً. 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 توزيعا مقبولا 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 ١٩٥٩م 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 - ١٣سنة) 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 العضليةMuscular 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 الميكانيكىMechanical 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 ينتهكون 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 فى هذه السمة أو القدرة، ومجموعة أخرى ممن يتميزون بمستوى منخفض (ضعيف) فى 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 الإحصائى المناسب (اختبار ت - t أو اختبار ف - f).

استخدام هذه الطريقة :

تستخدم طريقة المجموعات المتضادة لتقدير صدق التكوين الفرضى للعديد من الاختبارات والمقاييس فى 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

Comparsion 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 (تقنينه)، فنحن

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

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

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

ويتوقف استخدام أساليب وطرق المعالجات التجريبية فى 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 الفرضى للاختبار أو المقياس.

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

وبخاصة في مجال علم النفس الرياضي، ولتوضيح ذلك نسوق المثال التالي(*):

«لحساب صدق التكوين الفرضي لمقياس الاتجاهات نحو النشاط البدني ATPA الذي أعده جيرالد كينيون Gerald kenyon ١٩٦٨م، والذي يتكون من ٥٤ عبارة موزعة على ستة أبعاد للاتجاهات نحو النشاط البدني هي:

- النشاط البدني كخبرة اجتماعية (٨ عبارات)
- النشاط البدني للصحة واللياقة (١١ عبارة)
- النشاط البدني كخبرة توتر ومخاطرة (٩ عبارات)
- النشاط البدني كخبرة جمالية (٩ عبارات)
- النشاط البدني لخفض التوتر (٩ عبارات)
- النشاط البدني كخبرة للتفوق الرياضي (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 خماسي 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 وفقا لمقياس تقدير خماسى هو:

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

وقد قام بعد هذه الخطوة بمقارنة درجات التفضيل العالى لكل بعد فى 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 التى تقيس هذا البعد، وأن المفحوصين الذين يعبرون عن تفضيل منخفض نحو بعدما يتخذون اتجاهاً أكثر سلبية نحو العبارات التى تقيس هذا البعد.

وقد قام كينيون بحساب المتوسطات الحسابية والانحرافات المعيارية لدرجات الاستجابات لكل من مجموعة التفضيل العالى فى 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 إحصائيا لصالح مجموعات التفضيل العالى.

طريقة التحليل العاملى Factorial analysis

التحليل العاملى أسلوب رياضى mathematics يستخدم كمنهج إحصائى 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 العاملى حساب معاملات الارتباط البينية intercorrelations coefficients ثم وضعها فى مصفوفة معاملات ارتباط Correlation matrix، حينئذ نقوم بتحليل هذه المصفوفة عامليا بإحدى الطرق الرياضية للتحليل العاملى 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 العاملى من إيجاد هذه المكونات وصف الترابط association بين المتغيرات قيد التحليل ما أمكن ذلك.

وقد نحصل من التحليل العاملى 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 العاملى 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 ٨٨,٠ ، فإننا نستطيع أن نقرر أن هذا الاختبار يعد صادقا فى قياسه لهذا العامل وأن معامل صدقه العاملى يساوى ٨٨,٠

وإذا كان اختبار الجرى - المشى لمسافة واحد ونصف ميل يتشبع بعامل التحمل الدورى التنفسى (اللياقة الهوائية) بمقدار ٩٤,٠ ، فإننا نستطيع أن نقرر أن هذا الاختبار يعد صادقا فى قياسه لهذا العامل وأن معامل صدقه العاملى يساوى ٩٤,٠٠

ملحوظات مهمة:

- يستخدم التحليل العاملى المحكات الخارجية، هذه المحكات 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 صدقًا 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 obstract 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 (لمتغيرين) 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 تطرح 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 المئين 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 المهارية 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 المهارية 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 (٠-١٩) 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 لاعبى كرة القدم

فئات درجات الاختبار	تكرارات الاختبار	تكرارات المحك					المجموع
		ضعيف	متوسط	جيد	جيد جدا	ممتاز	
٠ - ١٩	١٢	٧	٥	٠	٠	٠	١٢
٢٠ - ٣٩	١٨	٦	٧	٥	٠	٠	١٨
٤٠ - ٥٩	٢٥	٩	٧	٧	٢	٠	٢٥
٦٠ - ٧٩	٨٠	١٠	٢٠	٣٠	١١	٩	٨٠
٨٠ - ٩٩	٣٥	٠	٤	١٠	١٥	٦	٣٥
١٠٠ - ١١٩	٢٠	٠	٠	٥	١٠	٥	٢٠
١٢٠ - ١٣٩	١٠	٠	٠	٠	٧	٣	١٠
	ن = ٢٠٠						ن = ٢٠٠

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

٠ ٠ ٠ ٠ ٠ 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:

$\frac{١٠٠ \times ٧}{١٢}$ = ٥٨٪، $\frac{١٠٠ \times ٥}{١٢}$ = ٤٢٪ (بالنسبة الفئة الأولى)

$\frac{١٠٠ \times ٦}{١٨}$ = ٣٣٪، $\frac{١٠٠ \times ٧}{١٨}$ = ٣٩٪، $\frac{١٠٠ \times ٥}{١٨}$ = ٢٨٪

(بالنسبة للفئة الثانية)

. . . . . وهكذا بالنسبة لجميع الخلايا

نقوم بوضع جميع النسب المئوية فى جدول التوقع expectancy table التالى:

جدول التوقع: ويظهر العلاقة بين درجات اختبار المهارة فى كرة القدم ودرجات الأداء على محك له خمسة مستويات (ن =٢٠٠).

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

يلاحظ من جدول التوقع الآتى:

- أن فئة اللاعبين الحاصلين على درجات الاختبار بين (٠-١٩) احتمال حصولهم على مستوى (جيد جدا) فى المحك (القدرة المهارية فى كرة القدم) هو صفر٪ .
- بينما نجد أن هذا الاحتمال يصل هذا الاحتمال إل ٧٠٪ بالنسبة للاعبين الحاصلين على درجات فى الاختبار تقع بين (١٢٠ - ١٣٩).

- 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 أيضا كلما زادت درجات الأداء على الاختبار، حيث كانت هذه النسب كالتالى:

٠٪، 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 طولا يصبح أكثر ثباتا وصدقا، فنحن نعرف أن صدق أى اختبار يتأثر بثبات هذا الاختبار، ونعرف أيضا أنه نتيجة استخدام معادلة سبيرمان - 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 الكاليبر (المساك) 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 وفقا لخمس حالات هى:

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

$$\overline{س}_{(ن\ ص)} = \frac{ن \times \overline{س}_{ص}}{\sqrt{ن + ن\ (ن - ١)\ \overline{١٠١}}}$$ معادلة (٣ - ٤)

حيث إن :

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

$\overline{س}_{ص}$ = الارتباط بين المحك والاختبار قبل تطويله (الصدق التجريبى للاختبار).

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

$\overline{١٠١}$ = معامل ثبات الاختبار.

مثال (٣ - ٤):

لنفترض أن معامل صدق أحد الاختبارات كان ٠,٥٠ ، وكان معامل ثباته هو ٠,٤٠ فكم يكون صدق هذا الاختبار إذا زاد طوله ٤ مرات؟

الحل:

بالتعويض فى المعادلة (٣ - ٤) ينتج أن:

$$س_{(ن\ ص)} = \frac{٤ \times ٠,٥٠}{\sqrt{٤ + ٤\ (٤ - ١)\ ٠,٤٠}}$$

$$= \frac{٢}{\sqrt{٤ + ٤\ (٣)\ (٠,٤٠)}}$$

$$= \frac{٢}{\sqrt{٨ \times ٣ \times ٠,٤٠}}$$

$$= \frac{٢}{\sqrt{٩,٦}}$$

$$= \frac{٢}{\sqrt{٣,٠٩٨}}$$

$$= ٠,٦٤٥٥٧$$

$$= ٠,٦٥$$

يلاحظ أن الاختبار الذى زاد طوله ٤ مرات زاد معامل صدقه من ٠,٥٠ إلى ٠,٦٥ .

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

ويستخدم لذلك المعادلة التالية:

$$\text{س}_{(\text{ن م})} = \frac{\text{س ص}}{\sqrt{\left(\frac{\text{ن}}{\text{ن} + (\text{ن} - 1)\text{ر}_{11}}\right)\left(\frac{\text{م}}{\text{م} + (\text{م} - 1)\text{ر}_{11}}\right)\text{ر}_{22}}}$$

معادلة (٣ -٥)

حيث إن :

$\text{س}_{(\text{ن م})}$ = الارتباط بين الاختبار بعد زيادته (ن) مرة، وبين المحك بعد زيادته (م) مرة.

$\text{س}_{\text{ص}}$ = معامل الارتباط بين الاختبار والمحك قبل الزيادة (الصدق التجريبى للاختبار).

ن = عدد مرات زيادة طول الاختبار أو إعادة تكراره.

م = عدد مرات زيادة طول المحك.

ر_{11} = معامل ثبات الاختبار.

ر_{22} = معمل ثبات المحك.

مثال (٣ - ٥):

لنفترض أن اختبار ما بلغ معامل ثباته ٠,٥٦ ، وكان معامل الارتباط بين هذا الاختبار والمحك الخارجى هو ٠,٦٥ فكم يبلغ معامل الصدق الجديد للاختبار إذا زاد طول كل من الاختبار والمحك ثلاث مرات، وكان معامل ثبات المحك هو ٠,٥٨ .

الحل:

من المثال السابق يتضح أن:

س ص = ٠,٦٥

ن = ٣

م = ٣

ر_{11} = ٠,٥٦

$$\sqrt{٢٠٢} = ٠,٥٨$$

وبالتعويض فى المعادلة (٣ -٥) ينتج أن:

$$\text{س (ن م)} = \frac{٠,٦٥}{\sqrt{\left(\frac{١ - ٠,٥٦}{٣} + ٠,٥٦\right)\left(\frac{١ - ٠,٥٨}{٣} + ٠,٥٨\right)}}$$

$$= \frac{٠,٦٥}{\sqrt{\left(\frac{٠,٤٤}{٣} + ٠,٥٦\right)\left(\frac{٠,٤٢}{٣} + ٠,٥٨\right)}}$$

$$= ٠,٩٢$$

مما سبق يتضح أن زيادة طول كل من الاختبار والمحك ثلاث مرات بالنسبة لطولهما الأصلى قد أدى إلى زيادة معامل الصدق التجريبى للاختبار من ٠,٦٥ إلى ٠,٩٢.

مثال (٣ - ٦):

فى المثال السابق (٣ - ٥) المطلوب حساب معامل الصدق الجديد للاختبار إذا زاد طول الاختبار ثلاث مرات، وزاد طول المحك مرتين.

الحل:

$$\text{س ص} = ٠,٦٥$$

$$\text{ن} = ٣$$

$$\text{م} = ٢$$

$$\sqrt{١٠١} = ٠,٥٦$$

$$\sqrt{٢٠٢} = ٠,٥٨$$

بالتعويض فى المعادلة (٣ - ٥) ينتج أن:

$$س(ن\ م) = \frac{٠,٦٥}{\sqrt{\left(\frac{١ - ٠,٥٦}{٣} + ٠,٥٦\right)\left(\frac{١ - ٠,٥٨}{٢} + ٠,٥٨\right)}}$$

$$= \frac{٠,٦٥}{\sqrt{\left(\frac{٠,٤٤}{٣} + ٠,٥٦\right)\left(\frac{٠,٤٢}{٢} + ٠,٥٨\right)}}$$

$$= \frac{٠,٦٥}{\sqrt{(٠,١٥ + ٠,٥٦)(٠,٢١ + ٠,٥٨)}}$$

$$= \frac{٠,٦٥}{\sqrt{٠,٧١ \times ٠,٧٩}}$$

$$= \frac{٠,٦٥}{\sqrt{٠,٥٦٠٩}}$$

$$= \frac{٠,٦٥}{٠,٧٤٩}$$

$$= ٠,٨٧$$

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:

$$ن = \frac{س^{٢}(ن\ ص) \times (١ - ١٠١)}{(س\ ص)^{٢} - س^{٢}(ن\ ص) \times ١٠١}$$ معادلة (٣ - ٦)

حيث إن :

ن = عدد المرات التى يجب أن يزيدها الاختبار.

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

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

١٠١ = معامل ثبات الاختبار.

مثال (٣ - ٧):

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

الحل:

بتطبيق المعادلة (٣ - ٦) ينتج أن:

$$\text{ن} = \frac{(٠,٦٥)^{٢}\,(١-٠,٤٠)}{(٠,٥٠)^{٢} - (٠,٦٥)^{٢}\,(٠,٤٠)}$$

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

$$= \frac{٠,٢٥٢}{٠,٢٥ - ٠,١٦٨}$$

$$= \frac{٠,٢٥٢}{٠,٠٨٢}$$

$$= ٣,٠٧$$

$$= ٣ \text{ مرات}$$

مما سبق يلاحظ أنه لكى نحصل على المستوى المطلوب لمعامل صدق الاختبار الذى يساوى ٠,٦٥ ، فإنه يلزم تطويل الاختبار المستخدم ثلاث مرات.

الحالة الرابعة : عندما يزيد طول الاختبار إلى مالا نهاية

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

$$\text{س ص \&} = \sqrt{\frac{\text{س ص}}{\text{١٠١}}} \qquad \text{معادلة (٣ - ٧)}$$

حيث إن :

س ص & = معامل صدق الاختبار بعد الزيادة

س ص = معامل صدق الاختبار قبل الزيادة

١٠١ = معامل ثبات الاختبار

مثال (٣ - ٨):

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

الحل :

بتطبيق المعادلة (٣ - ٧) ينتج أن :

$$\text{س ص \&} = \sqrt{\frac{٠,٦٥}{٠,٨٢}}$$

$$= \frac{٠,٦٥}{٠,٩٠٦}$$

$$= ٠,٧٢$$

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

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

$$\&\& = \frac{س ص}{\sqrt{١٠١ \times ٢٠٢}}$$ معادلة (٣ -٨)

حيث إن :

$\&\&$ = معامل صدق الاختبار بعد الزيادة.

$س ص$ = معامل صدق الاختبار قبل الزيادة.

١٠١ = معامل ثبات الاختبار.

٢٠٢ = معامل ثبات المحك.

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

الحل:

بالتعويض فى المعادلة (٣ - ٨) ينتج أن:

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

$$= \frac{٠,٦٥}{\sqrt{٠,٧٣٨}}$$

$$= \frac{٠,٦٥}{٠,٨٥٩}$$

$$= ٠,٧٦$$

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

٢ - عدم وضوح التعليمات Unclear Direction

عندما تكـون تعليمات الاختـبار المعطاة للمفـحوصين غامـضة بمعنى أن تكـون غيـر واضـحة وغـيـر محـددة، فـإن ذلك يؤدى إلـى اختـلاف المفحـوصين فى تفسيـر هذه التعليمـات مما يؤثر على صدق الاخـتبار، فتعليمات الاختبار غير الواضحـة لا تمكن المفحوص من كيفية الاستجابة على وحدات الاخـتبار أو المقـياس، مما قد يشـجع على الاعتمـاد على التخـمين Guessing، مما يترتـب عليه بالتالـى الحصول على مـعاملات صدق منخفضة.

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

Reading vocabulary and sentence too difficult

أظهرت التجارب العلمية فى مجـال القياس النفسى أن الكلمات والجمل المعقدة التراكيب بالنسبة لمسـتوى المفحوص الذى يعطى له الاختبار سوف تؤثر على صدق الاختبار، وبناء على ذلك يـنصح 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 تـكون فقرات الاختبـار سهلة جداً أو صـعبة جداً، فإن الاخـتبار يفقد قـدرتـه على التميـيز فيما بين المفـحوصين، ومن ثم يكون لـه مـعامـل 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 اختبارت 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 فى الاختبارات المقيدة بزمن مثل اختبارات : التحصيل والذكاء - مما يؤثر على صدق الاختبار.

ثانيا : العوامل التى تتعلق بتطبيق وحساب درجات الاختبار :

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

ثالثا : العوامل التى 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 والذى يجعل المفحوص (وبخاصة صغير السن) غير قادر على الاستجابة على الأسئلة بصورة طبيعية، كما أن الاضطراب الانفعالى emotional disturbcnce يؤثر هو الآخر على استجابة المفحوصين، كما تمثل الدافعية motivation أحد أهم العوامل التى تؤثر على استجابات المفحوصين وبخاصة بالنسبة لاختبارات الأداء الأقصى maximum performance كاختبارات القدرات البدنية والعقلية واختبارات التحصيل وغيرها.

وفى هذا الخصوص يبين كرنباخ Cronbach,L.J ١٩٧٠م أن اتجاه الاستجابة response set على العبارات يمثل أحد العوامل التى تؤثر على صدق الاختبار، فعندما يكون ميل المفحوصين نحو نمط محدد من الاختبارات متناغما Consitent tendency، فإنه يلزم فى هذه الحالة أن يكون عدد العبارات (الجمل) الصحيحة true مساوياً لعدد العبارات الخطأ false، ففى الاختبارات التى تتطلب الاستجابة بصح(✓) أو

خطأ (×) يميل بعض المفحوصين إلى الإجابة بصح (✓) عندما لا يعرفون الإجابة الصحيحة للفقرة، بينما يميل البعض الآخر الى الإجابة بخطأ (×). 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 الضرورى أن تكون عبارات الاختبار نصفها صحيحا والنصف الآخر خطأ.

رابعاً : العوامل المتعلقة بمجموعة التقنين والمحك

من المعروف أن صدق الاختبار يختص بمجموعة محددة من الأفراد من حيث الجنس، والعمر الزمنى، والحالة التدريبية، والتعليمية، والصحية، والعرق Race ، والمستوى الاقتصادى الاجتماعى إلى غير ذلك من المحددات التى تصف لنا مجموعة الأفراد التى أعد لها الاختبار فى الأصل.

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

فطبيعة صدق الاختبار تتوقف على خصائص Characteristics ، مجموعة (عينة) التقنين التى أعد لها الاختبار. وقد أشرنا من قبل إلى أن مدى السمة أو القدرة Range of ability المقيسة يعد من العوامل التى تؤثر على صدق الاختبار، فعندما تكون مجموعة التقنين غير متجانسة تجانساً شديداً بالنسبة للسمة أو القدرة المقيسة، فإن هذا بدوره يؤدى 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 يؤدى إلى توزيع الدرجات فى انتشار ضيق مما يؤدى إلى انخفاض صدق الاختبارات المرتبطة بالمحك (لاحظ أن الصدق المرتبط بالمحك يعتمد فى حسابه - فى معظم الحالات - على معامل ارتباط بيرسون الذى 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 التى تستخدم اللغة المكتوبة (اختبارات الورقة والقلم)، فالاختبار الذى أعد لثقافة محددة ربما يصبح غير صادق بالنسبة لثقافة أخرى، وذلك بسبب الاختلافات فى القيم والعادات والتقاليد والمستوى الاقتصادى الاجتماعى 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 فى صدق الاختبار الجديد إلا بعد التأكد من ذلك عمليا.

ويتفق معظم علماء القياس أن الاختبارات عبر - ثقافية - cros cultural tests التى يتم إعدادها فى 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 فى المجال الرياضى.

---

* راجع: كتاب «الإحصاء الوصفى فى علوم التربية البدنية والرياضة، دار الفكر العربى، 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 الذى لا يرتبط بنفسه لا يتوقع له أن يرتبط مع محكات خارجية مستقلة. ومن ناحية أخرى فإن الاختبار الذى يظهر ثباتاً منخفضاً فإنه لا يتوقع أن يكون له صدقاً مرتفعاً، لهذا يعتمد الصدق على الثبات وبخاصة بالنسبة للاختبارات المتجانسة (الموحدة) Homogeneous tests . أما فى حالة الاختبارات غير المتجانسة، فإن صدق الاختبار قد يظهر مرتفعاً دون أن تكون هناك حاجة لكى يظهر الاختبار ثباتا مرتفعا كما فى حالة ثبات الاتساق الداخلى للاختبار أو المقياس، حيث يعزى ذلك إلى أن كل جزء من أجزاء الاختبار يقيس وظيفة مستقلة مما يترتب عليه ظهور ثبات الاتساق الداخلى منخفضا لكون وحدات الاختبار لا ترتبط بعضها مع بعض، بينما تظهر هذه الوحدات ارتباطات مرتفعة مع بعض المحكات الخارجية المستقلة، فى مثل هذه الحالة نستطيع القول بأن هذا الاختبار الذى أعطى ثباتا منخفضا عن طريق الاتساق الداخلى له صدق مرتفع مع بعض المحكات الخارجية.

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

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

يقومون بإعداده ، مع ملاحظة أن أهـداف وأساليب بناء كل من ثبات وصدق الاختبـار تبدو متضاربة، فـقد تتعارض متطـلبات الحصول على ثبـات مرتفع للاختبار مـع متطلبات الحصول على صـدق مرتفع لنفس الاختبـار، فالثبات المرتفع يتطـلب فقرات متسـاوية فى درجة الصعـوبة، كما يتطلب مـعاملات ارتباط بينية مرتفعـة (فيما بين الفقرات) Intercorrelations ، فى حين يتطلب الصدق المرتفـع لنفس الاختبار فـقرات مختلفـة فى درجات صعوبتهـا، كما يتطلب معاملات ارتباط بينية (فيما بين الفقرات) منخفضة.

وللتغلب على هذه المشكلة يقترح كل من توكر Tucker ١٩٤٦م وجيلفورد Guilford ١٩٥٦م أنه لكى يمكن الحـصول على صدق مرتفع بـالإضافة إلى ثبات مرتفع للاختـبارات، فإنه يجب أن تتراوح الارتباطات البينـية للفقـــرة inter-item correlations (معاملات ارتـباط كل فقرة مع غيرهـا من الفقرات) من (١,٠ إلى ٦,٠) لضمان الحـصول على معاملات ثبات وصـدق مرضية للاختبار أو المقياس.

معجم إنجليزى - عربى للمصطلحات

A

Abstract Constructs تكوينات مجردة

Adolescence مرحلة المراهقة

American psychological association الجمعية الأمريكية لعلم النفس

Ambiguity الغموض

Academic success النجاح الأكاديمى

Arithmetic test اختبار فى الحساب

Achievement testing اختبار التحصيل

Aptitude tests اختبارات الاستعدادات

Appearance المظهر الخارجى للاختبار

B

Bivaraite distribution التوزيع التكرارى الثنائى

Biserial Correlation الارتباط الثنائى

C

Clusters تجمعات

Critical thinking التفكير الناقد

Constructs تكوينات

Concurrent validity الصدق التلازمى

Content validity صدق المحتوى (المضمون)

Cros - Cultural tests اختبارات عبر - ثقافية

Criterion Relate validity الصدق المرتبط المحك

Construct validity صدق التكوين الفرضى

Contents المحتويات (محتويات)

Curricular validity الصدق المنهجى

Criterion method طريقة المحك

Correlation matri مصفوفة ارتباط

Continuous data بيانات متصلة

Composite score الدرجة المركبة

Contrasted group المجموعات المتضادة

Childhood مرحلة الطفولة

Commulation شيوع (اشتراكيات)

D

Discrete data بيانات متقطعة

Discriminat validation صدق التمييز

Domain محال

E

Experimental validlity الصدق التجريبى

External خارجى

External Criterion محك خارجى

Experts الخبراء (المحكمون)

Expectancy tables جداول التوقع

Emotional disturbunce الاضطراب الانفعالى

F

Fidelity الدقـة

Factorial validity الصدق العـاملى

Face Validity الصدق الـظاهري

Factorial analysis التـحليل العـاملى

Factor Loading تشبع الاختبارات على العامل

H

Hypothetical Construct تكوين افتراضى

Homogeneous Lengthening
التطويل المتجانس فى الاختبار

I

Intrinsic Validity الصدق الداخلى(الذاتى)

Independent مستـقبل

Impiyical validity
الصدق العملى (الإحصائى)

Inded of validity coefficeent
مؤشر معامل الصدق

Irrelevant ليس له صلة بالموضوع

Index of reliability مؤشر الثبات

Internal criterion محك داخلى

Intercorrelations coefficients
معاملات الارتباط البينية

Interrelationships
العلاقات التبادلية

L

Linear combination تراكيب خطية

Loadings تشبعات الاختبارات على العوامل

M

Multivariate analysis
التحليل متعدد المتغيرات

Multipple correlation الارتباط المتعدد

Mechanical comprehension
الاستيعاب الميكانيكى

Mechanical ability
القدرة الميكانيكية

Maximum validity
الحد الأقصى لصدق الاختبار

U

Uncorrelated غير مترابطة

Unclear directions عدم وضوح التعليمات

F

Predictive efficiency
الكفاءة التبنؤية للاختبار

Pure factors matrix مصفوفة العوامل الثقة

Principle Components المكونات الأساسية

Phi-Coefficient معامل فاى

Present standing الوضع الراهن

Predictive tests اختبارات تنبؤية

تشبعات الاختبارات على العوامل Saturations

T

الصح - الخطأ True - False

صدق السمة Trait Validity

تكوين نظرى Theoretical construct

كفاءة الاختبار Test efficiency

نظام الدورى Tournament

الارتباط الرباعى Tetrachoric Correlation

مجموعات مدربة Trained groups

V

الصدق Validity

معامل الصدق Validity Coefficient

R

اتجاه الاستجابة Response set

علاقة Relationship

صدق التمثيل Representativeness validity

الصدق التحليلى Relevance validity

معامل ارتباط سبيرمان لفروق الرتب Rho

مقاييس تقدير Rating scales

الاستيعاب القرائى Reading comprehension

S

مواقف situations

الصدق الإحصائى Statistical validity

القدرة المكانية Spatial ability

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.