المدخل إلى القياس

في التربية البدنية والرياضة

الدكتور

محمد نصر الدين رضوان

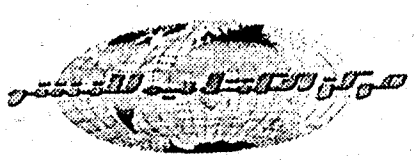
أستاذ القياس والتقويم

ورئيس قسم علم النفس الرياضي (الأسبق)

بكلية التربية الرياضية للبنين - جامعة حلوان

الطبعة الأولى

١٤٢٦ هـ - ٢٠٠٦م



مقدمـــة

يسعدنى أن أقدم إلى جمهور القراء من الطلاب والباحثين والمهتمين كتاب «المدخل إلى القياس فى التربية البدنية والرياضة» وهو كتاب يستهدف التعريف بالتخصص الرئيسية بالنسبة لقياس وتقويم السلوك البشرى فى هذه التخصص، حيث يتضح أن هذا الفرع من فروع التخصص لم يعد قاصراً أبداً على الموضوعات التقليدية التى توارثناها منذ أن أدرجت «مادة الاختبارات والمقاييس» لأول مرة ضمن البرنامج الدراسى بالمعاهد العليا للتربية الرياضية «كليات التربية الرياضية الآن» عام ١٩٦١م.

وقد أردت بهذا الكتاب أن أضمنه خبراتى فى التدريس والبحث خلال ثلث قرن قضيتها محاضراً للقياس والتقويم والإحصاء والبحث العلمى ومناقشاً ومشرفاً على العديد من البحوث والرسائل العلمية بكليات التربية والتربية الرياضية فى جمهورية مصر العربية والمملكة العربية السعودية، حيث كشفت هذه الخبرة عن ثلاث نقاط مهمة هى :

أولاً : أن القياس فى التربية البدنية والرياضة يعد علماً حديثاً نسبياً، وهو يرجع فى أصوله ونشأته إلى «القياس النفسى».

ثانياً : أن المتتبع لحركة القياس المعاصرة فى مجال العلوم الإنسانية سوف يلاحظ أنه يلزم إعادة النظر بالنسبة لبعض القضايا العلمية والاهتمامات والأساليب الفنية الخاصة بالقياس فى المجال الرياضى.

ثالثاً : أن تقدم أى علم من العلوم إنما يقاس بقدرة هذا العلم على تطويع واستخدام رياضياته، وأن القياس هو «رياضيات» التربية البدنية والرياضة، فهو المدخل الحقيقى للنهوض بهذا التخصص من حيث التنبؤ والتحكم فى البيئة الخارجية، وضبط متغيراتها والسيطرة عليها.

وينتظم هذا الكتاب فى ستة فصول رئيسية تعرض للموضوعات الأساسية التى تهم القارئ المتخصص وتفتقر إليها المكتبة الرياضية العربية، حيث روعى فى هذه الموضوعات أن مجالات الاهتمام التى تناولتها تمثل أحدث ما توصلت إليه نظريات القياس وتقنياته، وقد حرصنا على أن ينتهى كل فصل من فصول الكتاب بمعجم إنجليزى - عربى للتعريف بالمصطلحات المستخدمة فى كل فصل، وقائمة بالمراجع العلمية الخاصة بكل فصل على حدة ، هذا بالإضافة إلى تضمين الكتاب لجداول وأشكال ونماذج إيضاحية لبعض الموضوعات التى اقتضى الأمر إيضاح محتواها.

ونود أن نشير إلى أن موضوعات هذا الكتاب تستلزم الاهتمام لكونها تمثل الأساس بالنسبة للعديد من عمليات وإجراءات القياس المتعلقة بمختلف المظاهر السلوكية فى الرياضة البدنية.

وأخيراً .. فإنه يسعدنى .. أن أتقدم بهذا الكتاب للدارسين فى كليات التربية الرياضية فى الجامعات المصرية والعربية، وإلى المهتمين بهذا الفرع من فروع التخصص (القياس والتقويم) راجياً أن يحقق هذا الكتاب الهدف من إعداده، وأن يلبى رغبة الباحثين فى التعرف على مجالات هذا التخصص ومساراته وتوجهاته المستقبلية.

والله سبحانه وتعالى من وراء القصد ،،،

أ.د. محمد نصر الدين رضوان

الجيزة: الهرم

أول مايو ٢٠٠٤

المحتويـــات

الموضوع — الصفحة

مقدمـــة ٣
المحتويات ٥

الفصل الأول

طبيعة القياس فى التربية البدنية والرياضة ١٧

١ - القياس والاختبار والتقويم ١٩
- مفهوم القياس ٢٠
- مفهوم الاختبار ٢١
- الفرق بين المقياس والاختبار ٢٢
- مفهوم التقويم ٢٣
- العلاقة بين القياس والتقويم ٢٤
- الفرق بين القياس والتقويم ٢٦
٢ - مجالات قياس السلوك الإنسانى ٢٨
المجال المعرفى ٢٨
- المستويات الرئيسية للمجال المعرفى ٢٨
- الأغراض المعرفية فى المجال الرياضى ٣١
المجال الانفعالى ٣٢
- المستويات الرئيسية للمجال الانفعالى ٣٣
- الأهداف الانفعالية فى الرياضة المدرسية ٣٤
المجال النفس - حركى ٣٧
- المستويات الرئيسية للمجال النفس - حركى ٣٨
- العلاقة بين المجالات الثلاثة والبنية المعرفية للتربية البدنية والرياضة ٤٥
٣- وظائف القياس فى المجال الرياضى ٤٧
- تحديد الأهداف ٤٧

الموضوع	الصفحة

- الحث الدافعى ٤٨
- تحديد التحصيل ٥٠
- التصنيف ٥١
- التنبؤ بالأداء المستقبلى ٥٤
- التوجيه والإرشاد ٥٥
- مراقبة التقدم ٥٦
- تعيين المستوى ٥٧
- تحسين التعليم (التدريب) ٥٩
- تقدير الدرجات ٥٩
- البحث العلمى ٦٠

٤ - تصنيف الاختبارات والمقاييس فى المجال الرياضى ٦١

التعريف ببعض المصطلحات ٦١

- مفهوم الاستجابة ٦١
- مفهوم السلوك ٦٢
- مفهوم الأداء (الإنجاز) ٦٤
- الأداء الأقصى ٦٥
- الأداء المميز ٦٥

أسس تصنيف الاختبارات والمقاييس ٦٥

- التصنيف وفقاً لطبيعة الأداء (الإنجاز) ٦٥
- التصنيف وفقاً لأسلوب التطبيق ٦٨
- التصنيف وفقاً لأنماط الفقرات التى يتضمنها الاختبار .. ٧١
- التصنيف وفقاً لطرق الاستجابة ٧٤
- التصنيف وفقاً لأسس بناء الاختبار ٧٦
- التصنيف وفقاً لطرق تفسير النتائج ٧٩

- **معجم إنجليزى - عربى للمصطلحات** ٩٠
- **مراجع الفصل الأول** ٩٢

الموضوع | الصفحة

الفصل الثانى

الثبـــات ٩٣

١ - مقدمـــة ٩٥
٢ - نظرية الثبات ٩٦
أ- مفهوم الثبات ٩٨
- أهمية الثبات ١٠١
٣ - مفهوم الثبات النسبى وطرق حسابه ١٠٢
- طريقـة إعادة الاختبـار ١٠٤
- طريقة الصـور المتكافئة ١٠٨
- طريقة التجزئة النصفية ١١٠
• طرق تصحيح معامل ثبات نصف الاختبار ١١٢
- معادلة سبيرمان - براون ١١٢
- معادلـــة رولـــون ١١٤
- معادلـــة فلانجـان ١١٨
• مميزات طريقة التجزئة النصفية ١٢٧
• المآخذ على هذه الطريقة ١٢٨
- طريقة الاتساق الداخلى ١٣١
* معادلات كيودر - ريتشاردسون ١٣١
* معامـل ألـفا ١٣٨
- طريقة تحليل التباين ١٣٩
- ما هى القيم المقبولة لمعامل الثبات ؟ ١٤٥
٤ - مفهوم الثبات المطلق وطرق حسابه ١٤٦
- ما الذى يعنيه الخطأ المعيارى ؟ ١٤٧
- العلاقة بين الانحراف المعيارى والخطأ المعيارى ١٤٩
-مؤشر (دليل) الثبات ١٥٠

الموضوع	الصفحة

٥ - **الثبات الخارجى والثبات الداخلى** ١٥٠
- الثبات الخارجى ١٥٠
- الثبات الداخلى ١٥١
- خريطة الثبات ١٥١
٦ - **العوامل التى تؤثر على ثبات درجات الاختبار** ١٥٣
- العوامل الخارجية ١٥٣
- العوامل الداخلية ١٥٤
٧ - **التصحيح للتنقية** ١٦٣
- كيف نحسن من ثبات درجات الاختبار ؟ ١٦٥
- تأثير العمر الزمنى على الثبات ١٦٥
٨ - **علاقة الموضوعية بالثبات** ١٦٦٦
- مفهوم الموضوعية ١٦٧
- تعريف الموضوعية ١٦٨
- كيفية حساب الموضوعية ١٦٩
• **معجم إنجليزى - عربى للمصطلحات** ١٧١
• مراجع الفصل الثانى ١٧٣

الفصل الثالث

الصدق

١٧٥
١ - مفهوم الصدق ١٧٧
٢ - الأنواع الأساسية للصدق ١٨٠
- صدق المحتوى (المضمون) ١٨٢
- **الصدق المرتبط بالمحك** ١٨٦
• **الصدق التلازمى** ١٨٧
- مفهوم المحك ١٩٠
- المحك المؤلف من مجموعة من العناصر ١٩٠

الصفحة

الموضوع

- الدرجة على المحك (درجة المحك) ١٩١
- المتنبئ (وسيلة التنبؤ) ١٩١
- المحكات الضعيفة ١٩١
• **الصدق التنبؤى** ١٩٢
- التنبؤ بالأداء المستقبلى على المحك ١٩٧
- ماهية معادلة التنبؤ ١٩٨
- الارتباط المتعدد ١٩٩
- الانحدار المتعدد ١٩٩
• **صدق التكوين الفرضى** ٢٠١
- طرق الحصول على صدق التكوين الفرضى ٢٠٥
- الإجراءات الارتباطية لحساب صدق التكوين الفرضى ... ٢٠٧
- مفهوم الصدق التقاربى وصدق التمييز ٢٠٨
- كيفية تحقيق الصدق التقاربى وصدق التمييز ٢١١
٣ - أنواع أخرى من الصدق ٢١٥
- الصدق الظاهرى ٢١٥
- الصـــدق الـذاتى ٢١٦
- الصـــدق العاملى ٢١٧
- حساسية الاختبار ٢١٧
- خريطـة الصـدق ٢١٩
٤ - طرق حساب الصدق ٢٢١
- طريقة استطلاع آراء الخبراء (المحكمين) ٢٢١
- طريقة المحـك ٢٣١
* المحك الداخلى ٢٣١
* المحك الخارجى ٢٣٢
• خصائص المحكات الخارجية ٢٣٣
• أنواع المحكات الخارجية فى المجال الرياضى ٢٣٣

الموضوع	الصفحة

• خطوات حساب صدق الاختبار بطريقة المحك الخارجى ٢٣٦
• ماهية معامل الصدق ٢٣٦
• تقويم معامل الصدق ٢٣٨
• تصحيح معامل الصدق المرتبط بالمحك (الخارجى) ٢٣٩
- الطريقة الأولى (طريقة الاختبار والمحك) ٢٣٩
- الطريقة الثانية (طريقة المحك فقط) ٢٤١
- طريقة المقارنة الطرفية ٢٤٣
- طريقة المجموعات المتضادة ٢٤٤
- طريقة مقارنة الدرجات قبل وبعد المعالجات الخاصة ٢٤٧
- طريقة التحليل العاملى ٢٥٠
- طريقة جداول التوقع ٢٥٣
٥ - العوامل التى تؤثر على صدق الاختبار ٢٥٧
- العوامل التى تتعلق بالاختبار نفسه ٢٥٧
- العوامل التى تتعلق بتطبيق وحساب درجات الاختبار ٢٦٨
- العوامل التى تتعلق باستجابات المفحوصين ٢٦٨
- العوامل التى تتعلق بمجموعة التقنين والمحك ٢٦٩
٦ - العلاقة بين صدق الاختبار وثباته ٢٧١
• معجم إنجليزى - عربى للمصطلحات ٢٧٣
• مراجع الفصل الثالث ٢٧٦

الفصل الرابع
القياس المعرفى

٢٧٧
١ - طبيعة القياس المعرفى ٢٧٩
- مفهوم القياس المعرفى ٢٧٩
- متطلبات بناء الاختبارات المعرفية ٢٨٠
٢ - خطوات بناء الاختبار المعرفى ٢٨٠

الموضوع	الصفحة
أولاً : التخطيط للاختبار	٢٨٠
- ماهية جدول المواصفات	٢٨٣
- جدول المواصفات من اتجاهين	٢٨٧
ثانياً : كتابة فقرات الاختبار	٢٨٩
- تعريف الفقرة	٢٨٩
- متطلبات إعداد وصياغة الفقرات	٢٨٩
- فقرات اختبار المقال	٢٩١
- توجيهات عامة لكتابة فقرات المقال	٢٩٢
- كيفية تقدير درجات فقرات المقال	٢٩٣
- فقرات الاختبار الموضوعى	٢٩٤
ثالثاً : تطبيق الاختبار	٣١٢
- إجراءات تتخذ قبل تطبيق الاختبار	٣١٣
- إجراءات تتخذ أثناء تطبيق الاختبار	٣١٤
- إجراءات تتخذ بعد تطبيق الاختبار	٣١٥
رابعاً : تحليل الفقرة	٣١٥
- الأهداف الرئيسية لتحليل الفقرة	٣١٥
- خطوات تحليل الفقرة	٣١٦
• حساب مؤشر الصعوبة	٣١٧
- دلالات مؤشر الصعوبة	٣٢٠
- صعوبة الفقرة والاختبار المرجعى المحك (CR)	٣٢٢
- صعوبة الفقرة والاختبار المرجعى المعيار (NR)	٣٢٣
- علاقة مؤشر الصعوبة بثبات الاختيار	٣٢٣
• حساب مؤشر التمييز	٣٢٦
- طريقة النسب المئوية	٣٢٧
- الطرق الارتباطية	٣٣٠
• التخمين وعوامل الصدفة	٣٣١

الموضوع	الصفحة

• تصحيح أثر التخمين ٣٣١
• متى نلجأ إلى تصحيح أثر التخمين ؟ ٣٣٦
خامساً : مراجعة وتنقيح الفقرة ٣٣٩
• معجم إنجليزى - عربى للمصطلحات ٣٤١
• مراجع الفصل الرابع ٣٤٣

الفصل الخامس

القيـاس الانفعالى ٣٤٥

١ - طبيعة القياس الانفعالى ٣٤٧
- أهمية القياس الانفعالى ٣٤٧
- فلسفة وأهداف القياس الانفعالى ٣٤٨
٢ - أساليب القياس الانفعالى ٣٥٠
أولاً : أسلوب التقرير الذاتى ٣٥٠
ثانياً : أسلوب التقارير التى يقوم بها آخرون ٣٥١
ثالثاً : أسلوب القياس الاجتماعى ٣٥٢
٣ - أدوات القياس الانفعالى ٣٥٥
- قوائم الميول (الاهتمامات) ٣٥٥
- قوائم الاتجاهات ٣٥٦
- قوائم الروح الرياضية ٣٥٧
- مقاييس القيادة ٣٥٧
- قوائم النمو الاجتماعى ٣٥٨
- تقديرات السلوك ٣٥٨
- قوائم (مقاييس) الشخصية ٣٥٩
٤ - الطرق الوصفية للقياس الانفعالى ٣٥٩
أولاً : الملاحظة ٣٦٠
- الملاحظة غير النظامية ٣٦٢

الموضوع	الصفحة
- الملاحظة النظامية	٣٦٢
- الملاحظة والاختبارات الموقفية	٣٦٤
- تدريب وإعداد القائم بالملاحظة	٣٦٤
ثانياً : المقابلة الشخصية	٣٦٦
- متى تستخدم المقابلة الشخصية	٣٦٦
- المقابلة الشخصية كاختبار	٣٦٧
- علاقة المقابلة الشخصية بالملاحظة	٣٦٨
- متطلبات نجاح المقابلة الشخصية	٣٦٩
- تسجيل المقابلة الشخصية	٣٧١
- أنماط المقابلات الشخصية	٣٧١
- طبيعة الأسئلة فى المقابلات الشخصية	٣٧٥
- أنواع الأسئلة فى المقابلة الشخصية	٣٧٩
- الإطار الزمنى للأسئلة	٣٨٢
- تسلسل الأسئلة	٣٨٣
٥ - الطرق الكمية للقياس الانفعالى	٣٨٣
أولاً : مقاييس ليكرت	٣٨٤
ثانياً : مقاييس تمايز المعانى	٣٨٨
ثالثاً : مقاييس التقدير	٣٩٣
١- مقاييس التقدير الرقمية	٣٩٧
٢- المقاييس التخطيطية (البيانية)	٤٠٣
٣- تقدير النسبة المئوية	٤٠٤
رابعاً : الاستبيانات (الاستخبارات)	٤٠٤
- مميزات الاستبيانات	٤٠٦
- عيوب الاستبيانات	٤٠٦
• **معجم إنجليزى - عربى للمصطلحات**	٤٠٩
• **مراجع الفصل الخامس**	٤١٢

الموضوع	الصفحة

الفصل السادس

القياس النفس - حركى	٤١٣
١ - طبيعة القياس النفس - حركى	٤١٥
- الاختبارات النفس - حركية	٤١٧
- كيفية اختيار الاختبارات النفس - حركية	٤١٨
٢ - إجراءات القياس النفس - حركى	٤١٩
أولاً : تدابير وإجراءات ما قبل التطبيق	٤٢٠
ثانيًا : تدابير وإجراءات أثناء التطبيق	٤٣١
ثالثاً : تدابير وإجراءات ما بعد التطبيق	٤٣٦
٣ - أنماط الاختبارات النفس - حركية	٤٣٨
أولاً : الاختبارات الموضوعية	٤٣٨
- الاختبارات الموقوتة	٤٣٩
- اختبارات المسافة	٤٤٠
- اختبارات الدقة	٤٤١
- اختبارات الأداء المتكرر	٤٤٢
ثانياً : التقديرات الذاتية	٤٤٢
١ - مقاييس التقدير النسبية	٤٤٣
أ - نظام الرتب	٤٤٤
ب- المسافات المتساوية ظاهرياً	٤٤٦
جـ- المقارنة الزوجية	٤٤٧
٢ - مقاييس التقدير المطلقة	٤٥٠
أ - المقاييس العددية	٤٥١
ب- المقاييس الوصفية	٤٥٤
جـ- المقاييس التخطيطية	٤٥٥
د- قوائم المراجعة (التدقيق)	٤٥٦

الموضوع	الصفحة
٤ - خطوات بناء الاختبارات النفس - حركية	٤٦١
- مراجعة معايير الاختبار الجيد	٤٦١
- تحليل الأداء	٤٦٦
- مراجعة الأدب السابق	٤٦٧
- اختيار أو بناء وحدات الاختبار	٤٦٨
- إعداد تعليمات الاختبار	٤٦٩
- مراجعة الخبراء	٤٦٩
- التطبيق الأولى للاختبار	٤٧٠
- تقدير الصدق والثبات والموضوعية	٤٧١
- إعداد المعايير	٤٧٣
- إعداد دليل الاختبار	٤٧٣
• **معجم إنجليزى - عربى للمصطلحات**	٤٧٥
• **مراجع الفصل السادس**	٤٧٧

الفصل الأول

طبيعة القياس فى التربية البدنية والرياضة

١ - القياس والاختبار والتقويم

٢ - مجالات قياس السلوك الإنسانى

٣ - وظائف القياس فى المجال الرياضى

٤ - تصنيف الاختبارات والمقاييس فى المجال الرياضى

- معجم إنجليزى - عربى للمصطلحات
- مراجع الفصل الأول

١ - القياس والاختبار والتقويم

قد يختلط مفهوم كلمة القياس مع مفهوم بعض الكلمات الأخرى المشتقة منها، ومن ثم فقد وجدنا أنه لكى نزيل اللبس الذى يمكن أن يقع فيه القارئ، فإنه يلزم توضيح المدلول الاصطلاحى لكل كلمة من هذه الكلمات ، وذلك على النحو التالى :

- مصطلح Measurement (اسم) ويقصد به عملية الحصول على المعلومات الكمية عن السمة أو القدرة أو الظاهرة المقيسة، فالقياس يعنى العملية التى يتم بها أخذ قياس الشىء، وعليه فقد يعنى القياس طرق جمع البيانات العددية، كما قد يعنى أيضاً نظام القياس a System of measurement .
- أما كلمة Measure فتعنى مقياس (اسم) وجمعها مقاييس Measurements ويقصد بها (الأداء أو الأدوات) التى تستخدم فى جمع البيانات عن الظاهرة المقيسة. وفى بعض الأحيان يستخدم مصطلح Measurements بمعنى أخذ عدد من القياسات مثل قياسات : اللياقة البدنية والذكاء وسمات الشخصية لفرد واحد أو لمجموعة من الأفراد.
- ومن ناحية أخرى فقد تعنى كلمة measure (فعل القياس)، ومن مشتقاتها Measuring (يقيس) و Measures (مقيس).

فعلى سبيل المثال نستطيع أن نقول :

« . . . أن القياس Measurement فى مجال القدرات البدنية يمكن أن يستخدم العديد من المقاييس Measurements لقياس Measuring القوة العضلية، ونقول: إن الباحث قد انتهى من قياس Measured طول التلميذ باستخدام وحدات قياس Unit of measurement تتبع نظام القياس System of measurement المترى، وأن الطول المقيس Measured بلغ ١٧٠ سم، وأن الباحث قد أخذ عدداً من القياسات Measurements الجسمية للتلميذ. . ».

مفهوم القياس Measurement :

مهما تعددت الكلمات وتداخلت المعانى المتعلقة بمصطلح «القياس» ، فإنه يمكن إيضاح مفهوم هذا المصطلح على نحو أكثر تفصيلاً، وذلك على النحو التالى :

١ – القياس Measurement ظاهرة واسعة الانتشار فى مجال العلوم الإنسانية، وهو يستهدف التقدير الكمى للسمة أو القدرة أو الظاهرة المقيسة .

٢ – القياس يجيب عن السؤال : كم How - Much ؟ مما يتطلب التحديد الكمى لما نقيسه، هذا التحديد الكمى يتم على أساس استخدام وحدات عد لها صفة الثبات النسبى، مثل قياس طول القامة بالسنتيمترات، أو قياس وزن الجسم بالكيلو جرامات، أو قياس الذكاء عن طريق نسبة الذكاء .

٣ – القياس يعنى تحديد أرقام (أعداد) لموضوعات أو أحداث معينة طبقاً لقواعد واضحة ومحددة تحديداً دقيقاً، مما يتيح له خاصية التعامل مع المقادير الكمية والتفكير الحسابى .

٤ – تشير نتائج القياس – دائمًا – إلى أرقام عددية، إذ إن نتائج القياس تصبح غير ذات مدلول ما لم تعبر عن نفسها رقمياً .

مما سبق نستخلص أن :

«مصطلح القياس Measurement يشير إلى تلك الإجراءات التى يتم بواسطتها تعيين أو تخصيص قيم عددية Numerical Values لشىء ما وفقاً لمجموعة من القواعد المحددة تحديداً دقيقاً، بحيث تشتمل هذه القواعد على طرق وشروط تطبيق أدوات القياس المستخدمة» .

كما نستخلص أن :

مصطلح أدوات القياس Measuring insruments يعنى :

« ما نقيس به الشىء من أدوات أو مقاييس بحيث يتم التعبير عن الناتج كمياً، فاختبارات اللياقة البدنية، واختبارات المهارات الحركية، والاختبارات النفسية، والاختبارات المعرفية جميعها أدوات قياس (مقاييس) تستخدم فى المجال الرياضى» .

مفهوم الاختبار Test :

١ – كلمة اختبار فى اللغة تحمل معنى الامتحان Examination ، وكلمة اختبره Tested تعنى «امتحنه» أو «جربه» ، وقد يشار إلى لفظة اختبار test فى اللغة الإنجليزية بكلمة Trial أو Proof وجميعها تعنى «التجربة – الامتحان – البرهان» أى تطبيق الأداة أو الاختبار على عينة من الأفراد .

٢ – وتشير بعض القواميس إلى أن الكلمة Test قد تكون مرادفة للكلمة Standard or Criterion بمعنى (معيار أو محك) .

٣ – والاختبار عبارة عن مجموعة من الأسئلة Questions أو المشكلات Problemes صممت لتقدير المعرفة أو الذكاء أو غيرهما من القدرات والخصائص ، فمفهوم الاختبار يعنى ضمنًا طلب الإجابة على مجموعة من الأسئلة المعدة سلفاً ، بحيث نحصل بناءً على إجابات المفحوص عن تلك الأسئلة على نتائج تكون فى شكل قيم عددية Numerical values عن سمات وقدرات المفحوص الذى أجاب عن تلك الأسئلة .

٤ – ويعرَّف الاختبار فى المجالين النفسى والتربوى على أنه : إجراءات منظمة تستهدف القياس الكمى أو الكيفى لمظهر واحد أو أكثر لسمة أو قدرة من القدرات عن طريق عينة من السلوك اللفظى أو غير اللفظى Verbal or non-verbal behaviours .

٥ – يستهدف الاختبار فى مجال الدراسات النفسية غرضين رئيسين هما : الأول : محاولة مقارنة نفس الفرد فى مظهرين أو أكثر بالنسبة للسمة أو القدرة المقيسة ، والثانى : مقارنة فردين أو أكثر فى مظهرين أو أكثر من مظاهر السمة أو القدرة المقيسة .

٦ – يوضح بين Bean ١٩٥٣ م أن الاختبار عبارة عن «مجموعة متعاقبة من المثيرات Stimuli نظمت لقياس بعض العمليات العقلية أو سمات الشخصية كمياً أو كيفياً» ، حيث يلاحظ أن هذا التعريف يكشف عن خاصيتين مهمتين للاختبار

النفسى والتربوى هما: الخاصية الأولى : أن الاختبار عبارة عن تنظيم متعاقب لمثيرات، هذه المثيرات تنظم فى شكل سلسلة متتابعة من العبارات أو الجمل وفقاً لبعض المبادئ التى يتطلبها بناء الاختبار، والخاصية الثانية: أنه يمكن استخدام كل من القياسات الكمية والقياسات الكيفية بالنسبة للاختبارات النفسية والتربوية.

مما سبق يتضح الآتى:

- أن الاختبار Test مصطلح يشير إلى مجموعة من البنود (الأسئلة أو المهام أو غيرهما) مرتبة بشكل يتيح تصحيح الإجابات أو الأداءات، حيث تستخدم الدرجات فى تقدير الفروق الفردية للمفحوصين.
- وأن الاختبار عبارة عن أداة قياس Measurement device تستخدم للحكم على جانب (مظهر) أو جوانب محددة بالنسبة للمفحوص مثل: الإنجاز أو الذكاء أو الشخصية أو غيرها.

الفرق بين المقياس والاختبار:

يرى بعض علماء القياس فى المجال التربوى الرياضى أنه طالما أن الاختبار يعنى بالنسبة للمفحوص «موقف الامتحان»، لذا نجد أن التفاعل بين المفحوص وموقف الاختبار يعد أحد أهم متطلبات الاستجابة (الأداء) القوية بالنسبة للاختبار.

فالإنجاز على الاختبار يستهدف الأداء بأقصى ما يستطيع المفحوص أن يقوم به، حيث يعبر هذا الأداء عن مستويات التحصيل أو القدرات أو الاستعدادات أو الكفاءة فى مجال محدد من المجالات، ويكون المفحوص هو المسئول الأول عن مستوى إنجازه وعما إذا كان هذا المستوى ممتازاً أم ضعيفاً، صحيح أم خطأ، مقبول أو غير مقبول.

وبناء على ما سبق يقرر علماء القياس فى التربية البدنية والرياضة الآتى:

١ - أنه يمكن اعتبار كل الاختبارات tests مقاييس Measurement لكون كل منهما وسائل (أدوات) لجمع البيانات.

٢ – أنه لا يجوز اعتبار كل المقاييس اختبارات، فالمقاييس الجسيمة كمقاييس الطول والوزن، ومقاييس الميول الرياضية، والاتجاهات نحو النشاط البدنى، ومقاييس الشخصية، لا يجوز وصفها بأنها اختبارات لكونها لا تتطلب من المفحوص التفاعل أثناء التطبيق، ولكونها أيضاً لا تحمل بالنسبة للمفحوص معنى «الامتحان».

٣ – أن القياس يعد أكثر اتساعاً من الاختبار، فنحن نستطيع أن نقيس بعض الصفات أو الخصائص باستخدام الاختبارات أو بدونها، فقد يستخدم لقياس خصائص وسمات معينة بعض الأساليب كالملاحظة Observation أو المقابلات الشخصية interviews وغيرها من الوسائل التى يمكن أن تعطينا معلومات فى شكل بيانات كمية Quantitative Form عن الظاهرة المقيسة.

مفهوم التقويم:

يقصد بالتقويم Evaluation إصدار الأحكام القيمية، واتخاذ القرارات والإجراءات العملية بشأن موضوع أو برنامج ما أو حتى فرد ما، وقد تتأسس الأحكام القيمية على مدى تحقق الأهداف على النحو الذى تحددت به للبرنامج أو المشروع، وقد تتأسس الأحكام على مدى قيمة أو جدوى أو فعالية برنامج أو طريقة أو مشروع ما، وذلك بهدف اتخاذ قرارات عملية بشأن الاستمرار فى المشروع أو تعديله أو تطويره أو التخلى عنه.

وإذا كان التقويم عملية تستهدف الحكم على قيمة الأشياء أو الموضوعات أو المواقف، فإنه لا بد وأن يعتمد فى ذلك على المحكات والمعايير التالية:

١ – المحكات (من خارج الظاهرة) Criteria:

وهى عبارة عن وسائل (قد تكون درجات أو آراء أو غيرهما) نرتضيها كأسس لمقارنة نتائج القياس بها، لإصدار أحكام قيمية إما كيفية أو كمية على هذه النتائج، ولعل من أهم ما يميز المحكات هى أنها تكون دائماً من خارج الظاهرة.

٢ - المستويات (المطلقة) Standards :

وهى عبارة عن وسائل «قد تكون درجات أو آراء أو غيرهما» تستخدم لتفسير درجة المفحوص بردها إليها، وتعد المستويات كأسس للتقويم من داخل الظاهرة المقيسة، إلا أنها تمثل ما يجب أن تكون عليه هذه الظاهرة، وتعرف المستويات باسم المحكات (CR) لكونها لا تتطلب مقارنة المفحوص بالمجموعة التى ينتمى إليها.

٣ - المعايير (النسبية) Norms :

وهى عبارة عن أسس أعدت فى ضوء الخصائص الواقعية لأداء المفحوصين بالنسبة للظاهرة المقيسة، وتستخدم كأسس داخلية (من داخل الظاهرة) لتفسير درجة المفحوص بمقارنتها بدرجات زملائه من نفس المجموعة التى ينتمى إليها، وتظهر المعايير - فى معظم الحالات - فى شكل جداول Tables تتضمن الدرجات الخام ومقابلاتها من الدرجات المحولة للعينة المرجعية التى أعدت لها المعايير فى الأصل.

وتنقسم المعايير من وجهة نظر علماء القياس النفسى إلى نوعين هما:

أ - المعايير السيكومترية (الجمعية المرجع):

وهى معايير تستخدم فى مقارنة درجات المفحوص فى اختبار ما بغيرها من الدرجات التى حصل عليها أفراد آخرون فى نفس الاختبار، هؤلاء الأفراد قد يكونون نفس المجموعات التى ينتمى إليها المفحوص، أو أفراد من مجموعات أخرى لها نفس خصائص المفحوص.

ب- المعايير الأديومترية (فردية المرجع):

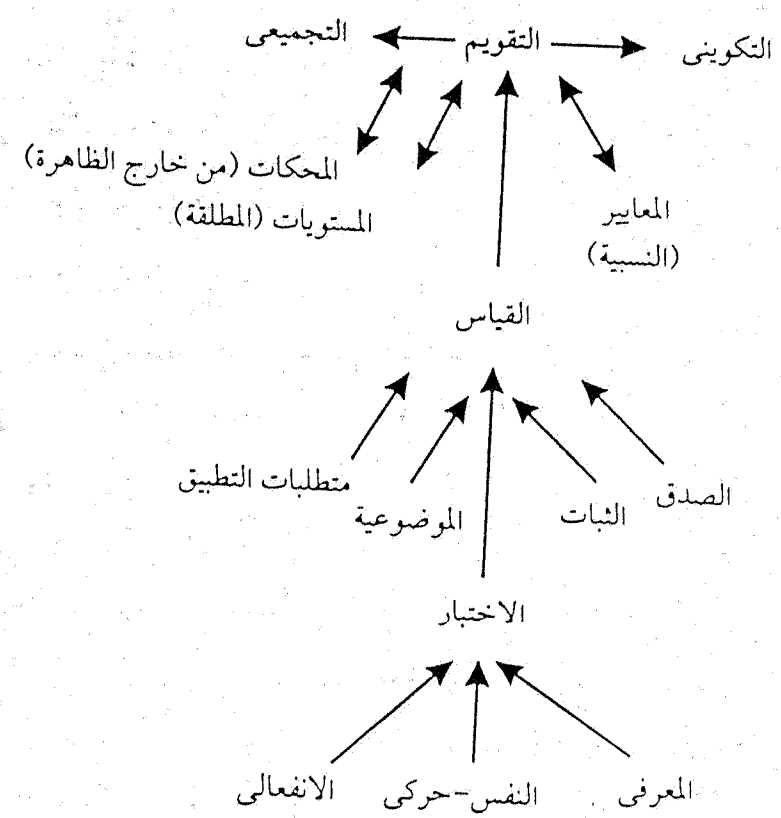
وهى معايير Norms تستخدم لتفسير درجة المفحوص على أساس مقارنته بنفسه من وقت لآخر، وغالباً ما تستهدف هذه المعايير مقارنة الدرجة التى يحصل عليها المفحوص فى ضوء مستواه فى الماضى.

العلاقة بين القياس والتقويم:

يتجاوز مصطلح التقويم فى مفهومه واستخداماته مصطلح القياس، فمصطلح التقويم يعتبر أعم وأشمل من مصطلح القياس، فالبيانات التى يتم الحصول عليها من

عمليات القياس تعد مدخلات inputs بالنسبة لعمليات التقويم، لهذا السبب يمكن اعتبار عملية القياس من متطلبات التقويم، أو مرحلة من أهم مراحله، ولأن عملية التقويم تتأثر بدقة البيانات المتجمعة من عمليات القياس، لذا يصبح من الضرورى الاهتمام بسلامة ودقة الأدوات المستخدمة فى القياس حتى تكون عملية التقويم عملية دقيقة.

ويمكن التعبير عن العلاقة بين القياس والتقويم وفق الشكل (١ - ١) حيث يلاحظ من الشكل الآتى:



شكل (١ - ١) : العلاقة بين الاختبار والقياس والتقويم

- أن التقويم أعم وأشمل من كل من القياس والاختبار.
- أن كل الاختبارات مقاييس وليست كل المقاييس اختبارات.
- أن الصدق والثبات والموضوعية ومتطلبات التطبيق من شروط الجودة المطلوب توافرها فى أدوات ووسائل القياس المختلفة.
- أن التقويم (بنمطية الكبيرين : التكوينى والتجميعى) يتطلب استخدام كل من المعايير أو المستويات أو المحكات كأسس للحكم على الدرجات المتجمعة من عملية القياس بغرض إصدار الأحكام القيمية عليها.

الفروق بين القياس والتقويم :

يمكن إدراك الفروق بين القياس والتقويم عن طريق عقد مقارنة بينهما وفق ما هو موضح بالجدول (١ - ١).

جدول (١ - ١) : بعض الفروق المهمة بين القياس والتقويم

القياس	التقويم
١ - يهتم بوصف السلوك. ٢ - يقتصر على التقدير الكمى للسلوك. ٣- يستخدم الأرقام فى التعبير عن الظاهرة. ٤ - يستهدف الحصول على نتائج دقيقة. ٥ - يعتمد على جمع المعلومات فقط. ٦ - حيادى لا يتضمن أية أحكام قيمية. ٧ - له وظيفة محدودة هى: الحصول على النتائج.	١ - يهتم بالحكم على قيمة السلوك. ٢- يشتمل على التقدير الكمى والكيفى للسلوك. ٣ - يقارن الأرقام بمعايير محددة لكى تصبح ذات معنى. ٤ - يستهدف تفسير هذه النتائج. ٥- يعتمد على المقارنات وإصدار الأحكام. ٦ - صريح: فالحكم هو وظيفته الأساسية. ٧ - له وظائف متعددة تتمثل فى التشخيص والعلاج، والتصحيح، وتحديد الأهداف، واختيار الوسائل وغيرها.

من الجدول (١ - ١) يلاحظ الآتى :

- أن القياس يحدد قيماً عددية للظاهرة المقيسة وفقاً لقواعد معينة، فى حين يصدر التقويم حكماً على هذه القيم وفقاً لمحكات ومعايير محددة.
- أن النتائج هى محور اهتمام القياس والتقويم، غير أن كلاً منهما يتناولها فى حدود وظيفته الأساسية، فالقياس يُعنى بوصف النتائج وإعطاء تقديرات كمية للسلوك، بينما يُعنى التقويم بالحكم على قيمة هذه النتائج وأنه أى التقويم يعطى اهتماماً خاصاً بالمحكات والمعايير.
- يمثل القياس حجر الزاوية بالنسبة لعملية التقويم، فالاقتصار على نتائج القياس وحدها لا يكفى، لأن الحصول على نتائج دقيقة وموضوعية من غير تقدير لقيمتها لا يعنى شيئاً بالنسبة للمهتمين والمسئولين.
- يستفاد من نتائج التقويم فى مساعدة المتعلمين والمتدربين على التقدم بمستوياتهم، وكمدخل لتحسين خطط وبرامج التعليم والتدريب، وفى تصحيح المسار عن طريق الحكم على مدى صلاحية العمل.

٢ - مجالات قياس السلوك الإنسانى

يستهدف القياس بالنسبة للأداء (السلوك) الإنسانى Human Performance ثلاثة مجالات Domains رئيسية عريضة هى :

- المجال المعرفى Cognitive domain
- المجالى الانفعالى Affective domain
- المجال النفس - حركى Psychomotor domain

وفيما يلى شرح موجز لكل مجال من هذه المجالات :

المجال المعرفى:

يهتم المجال المعرفى أساساً بالإنجاز (الأداء) العقلى Mental performance الذى يتعلق باسترجاع (تذكر) Recall المعلومات والعمل على تطوير القدرات والمهارات العقلية المعرفية Intellectual Abilities .

المستويات الرئيسية للمجال المعرفى :

قدم بلوم وآخرون .Bloom et al ١٩٥٦م ستة مستويات للمعرفة متدرجة من حيث درجة صعوبتها ودرجة أهميتها ما بين عمليات الاستدعاء البسيط مثل التذكر إلى العمليات المركبة مثل تركيب وتحليل وتقويم المعلومات والمعارف والأفكار، وهذه المستويات الستة كالآتى :

١ - تذكر المعلومات (المعرفة) Knowledge :

ويعنى استرجاع الموضوعات التى سبق تعلمها، ويضم هذا المستوى تذكر (استرجاع) مدى عريض بالنسبة للمادة الدراسية بدءاً بالحقائق الخاصة Specific facts بالمادة وحتى النظريات والتراكيب البنوية لها . ويتطلب التذكر (الاسترجاع) استحضار المعلومات المناسبة فى العقل (الذهن) Mind عن الحقائق، وطرق ووسائل التعامل مع تلك الحقائق، بالإضافة إلى العموميات والتجريدات فى ميدان معين .

وتمثل المعرفة أدنى مستويات نواتج Learning Outcomes بالنسبة للمجال المعرفى من حيث صعوبتها أو تعقيدها ، مع ملاحظة أن إتقان تذكر المعلومات ضرورى لإتقان المستويات المعرفية الأخرى بما تتضمنه من عمليات مركبة، فقد اعتبره بلوم وزملاؤه أول مستويات المجال المعرفى لكون النمو المعرفى يبدأ من البسيط إلى المركب، ومن المحسوس إلى المجرد.

٢ - الفهم Comprehension :

ويقصد به القدرة على إدراك معانى المادة، ويمكن أن يظهر فى ترجمة المادة (المعلومات) من صورة إلى أخرى، وفى تفسيرها عن طريق الشرح أو التلخيص، وفى إمكانية تقدير الاتجاهات المستقبلية (عن طريق التنبؤ بالآثار والأشياء المترتبة على شىء أو فعل معين) وفى أن يعرف الفرد الشىء المراد التفاعل معه، وأن يعرف الرسالة التى يبنى عليها عملية الاتصال والتى على أساسها يستطيع استخدام بعض الأفكار المتضمنة فى الرسالة. وهذه النواتج تمثل خطوة أبعد من مجرد تذكر المادة أو تذكر المعلومات.

٣ - التطبيق Application:

وهو يشير إلى القدرة على استخدام ما تعلمه الفرد من أفكار عامة ومبادئ فى مواقف جديدة محددة ومتماسكة. معنى ذلك أن التطبيق يتضمن استخدام القواعد rules والطرق methods والمفاهيم Concepts والمبادئ Principles ، والقوانين Laws والنظريات theories ، حيث تتطلب نواتج التعلم فى هذا التصنيف مستوى من الفهم أكبر من التصنيف السابق، فالتطبيق يتطلب أساساً التجريدات abstractions فى مواقف محددة، هذه التجريدات تقوم على التحليل والمقارنة، ثم تنتهى بإعطاء تعميمات Generalized فى شكل أفكار أو قواعد عامة.

٤ - التحليل Analysis:

وهو يشير إلى القدرة على تفكيك (تحليل - تجزئة) Break down المادة إلى مكوناتها الأساسية بما يساعد على فهم تنظيمها البنائى، حيث يتطلب ذلك

تحديد المكـونات (الأجزاء) وتعريفـها، وتحليل العـلاقات فيمـا بين المكونات أو العناصـر أو الأجزاء، وتحليل الأسـس التنظيمـية Organizational principles التى تربـط المكونات أو العنـاصر بعضـها ببعض. وتمـثل نواتج التعـلم فى هذا التصنيف مستوى عقليًا معرفيًا أعلى من مستويات الفهم والتطبيق لكونها تتطلب فهمًا لـكل من المحتوى Content والشكل البنائى Structural Form للمادة أو المحتوى المعرفى.

٥- التركيب Synthesis:

وهو يشير إلى القـدرة على وضع الأجزاء أو العناصر معاً لـتكوين كل فى شكل جديد، وهو ما يتطلب التعامل مع الأجزاء أو العناصـر للتوفيق فيما بينها لتعطى تركيبـاً أو نمطًا لم يكن واضحـاً أو مفهوماً من قـبل. ومن ناحية أخـرى يقصد بالتركيب العمـل على معالجة الأجزاء (العناصر - المكونات) وتنـظيمها والربط فيما بينهـا على نحو يؤلف نمطًا لم يكن واضحاً من قبـل. ويتكون التركيب من إنتاج اتصال مـتفرد Unique، وبناء خطة عـمليات، واشتـقاق مجمـوعة من العلاقـات المجردة، ونواتج التعـلم لهذا المستـوى تؤكد السلوك الابـتكارى مع التأكيد على تكوين أشكال أو أنماط جديدة.

٦- التقويـم Evaluation:

مصطلـح يهتم بالقـدرة على الحكم على قـيمة المادة أو الشىء، هـذا الشىء قد يكون عبارة عن قصة أو شعر أو فن أو تقريـر بحث، وتقوم أحكامه على معايير Criteria محددة، تلك المعايير قد تكون معايير داخلية Internal خاصة بالتنظيم أو معايير خارجية External خاصة بالغرض أو الهدف. وعلى المتعلم أن يحدد نوع المحك المستخدم. وتمثل نواتج التعلم لهذا المستوى أعلى مستويات التعلم فى المجال المعرفى وهى تتضمن عناصر من المستـويات الخمسة السابقة بالإضافة إلى أحكام قيمية Value Judgments واعية تستند إلى معايير محددة.

مما سبق يتضح أن المجال المعرفى يشتمل على ستة مستويات (فئات) رئيسية يمكن وضعها فى فئتين كبيرتين هما:

أ - **المعرفة** Knowledge وتشمل المستوى الأول وتدور حول تذكر واسترجاع المعلومات.

ب- **القدرات والمهارات العقلية المعرفية** Intellectual abilities and Skills وتشمل المستويات الخمسة المتبقية وهى: الفهم والتطبيق والتحليل والتركيب والتقويم.

ويوضح بلوم أن هذه الفئات المتنوعة من المستويات، وما ينتج عن تحليل كل فئة منها من عدد من النواتج، هى التى تشكل موضوع القياس التربوى edumetrics، حيث يلاحظ أن الكثير من هذه النواتج يتطلب استخدام العمليات العقلية العليا Cognitive processes باستثناء تذكر (استرجاع) المعلومات retention of information لكونها تمثل أدنى مستويات المجال المعرفى من حيث صعوبتها أو تعقيدها.

وبناء على ما سبق أصبحت الأهداف المعرفية بفئاتها ومستوياتها، وما ترتب عن تحليلها من تفاصيل، تشكل مواد (موضوعات) الاختبارات التربوية Educational tests التى تهتم بقياس التحصيل التربوى (المدرسى) Educational achievement فى الموضوعات التى تقدمها المدرسة، وقياس التحصيل المهنى Occupational achievement فى موضوعات وأعمال خاصة تتطلبها بعض المهن والوظائف.

ونظراً لأن التحصيل يرتبط بكل ما هو مكتسب بالتعليم والتعلم Learning and Instruction فى مجال الموضوعات الدراسية والمهنية المختلفة، لذا نجد أن الاختبارات فى المجال المعرفى قد صممت لقياس موضوعات تتعلق بالمعرفة والفهم والتطبيق والتحليل والتركيب والتقويم فى مجال دراسى أو مهنى معين.

الأغراض المعرفية فى المجال الرياضى:

تهتم الأغراض المعرفية فى المجال الرياضى بالتحصيل والفهم والتطبيق والتحليل والتركيب والتقويم بالنسبة للمعلومات المرتبطة بموضوعات مثل بنية ووظيفة الجسم، الصحة، الإسعافات الأولية، تاريخ التربية البدنية والرياضة، فلسفة

التربية البدنية، النمو والتطور، التنظيم والإدارة الرياضية، التعلم الحركى، فسيولوچيا الرياضة، الأسس الميكانيكية للحركة، اللياقة البدنية، القوام، المنهج والطرق والوسائل، تواعد واستراتيچيات وخطط اللعب وغيرها.

وقد حددت أنجيلا لومبكن Lumpkin,A. ١٩٩٤م مستويين كبيرين للأغراض المعرفية فى المجال الرياضى هما:

أولاً: أغراض تذكر المعلومات Knowledge وتشمل الموضوعات التالية:

١ - وظائف الجسم.

٢ - الصحة.

٣ - عمليات النمو والتطور.

٤ - التعلم الحركى.

ثانياً: أغراض الفهم Understanding وتشمل الموضوعات التالية:

١ - قوانين وقواعد الألعاب والأنشطة الرياضية.

٢ - المهارات.

٣ - الاستراتيچيات وخطط اللعب.

٤ - القواعد والبرتوكولات.

المجال الانفعالى:

تبين المعاجم المتخصصة أن مصطلح Affective يعنى الانفعال emotion أو الشعور Feeling أو الاعتقاد بشىء ما believing أكثر من كونه مصطلحا يرتبط بالتفكير Thinking أو بالإدراك الحسى Perceiving أو الفعل doing.

وفى كثير من الأحيان يستخدم مصطلح المجال الانفعالى كمرادف لمصطلح السلوك الانفعالى Affective Behavior، وهو يتضمن أنماط السلوك المرتبطة بالمشاعر والعواطف والتى تظهر فى الاتجاهات Attitudes والاهتمامات (الميول)

Interests والقيم Values والتذوق appreciations والنزعات الانفعالية emotional biases وغيرها من الموضوعات، حيث أن أحاسيس الفرد وانفعالاته بالنسبة لهذه الموضوعات تظهر من خلال السلوك الانفعالى لهذا الفرد.

ومن الملاحظ أن القياس والتقويم فى المجالين التربوى والنفسى يركزان على الجانبين المعرفى والانفعالى، حيث تستهدف التربية والتعليم العمل على تطوير وتحسين الجانب الانفعالى للتلميذ (الطالب) من حيث: الوعى، والانتباه، والدافعية، ومفهوم الذات، والرضا، والطموح، والتسامح، والميول، والقيم، وغيرها من الحصائل Outcomes المتعلقة بهذا المجال.

وعلى عكس المجال المعرفى فإن وحدات قياس البعد الانفعالى ليس لها إجابات صحيحة أو خاطئة، وإنما تحسب الدرجات على أساس مفتاح تصحيح يحدد اتجاهات الفقرة (الوحدة أو العبارة) التى تعكس توجه الاختيار الذى يقرره أو يرغبه المفحوص.

المستويات الرئيسية للمجال الانفعالى:

أوضح كراشول وبلوم ومازيا Krathwohl , Bloom, and Masia ١٩٦٤م أن المجال الإنفعالى يتضمن أنماط السلوك المرتبطة بالمشاعر والعواطف والتى تظهر فى الاتجاهات والاهتمامات والتقدير والقيم وغيرها، وقد قدموا تصنيفاً لأنماط السلوك فى هذا المجال يقع فى خمسة مستويات كبيرة مرتبة هرمياً من البسيط إلى المركب وهى:

١ - الاستقبال (الحضور) Receiving - Attending :

ويتضمن: الوعى Awareness ، والرغبة فى الاستقبال Willingness to receive والانتباه الانتقائى (الاختيار) Selected attention.

٢ - السلوك (الاستجابى) Responding :

ويتضمن: الإذعان فى الاستجابة acquiescence ، والرغبة فى الاستجابة Willingness ، والرضا فى الاستجابة Satisfaction in response.

٣ - السلوك القيمى Valuing :

ويتضمن : تقبل القيمة Acceptance ، وتفضيل القيمة Preference For a value ، والالتزام بالقيمة Commitment .

٤ - السلوك التنظيمى Organizing :

ويتضمن : تحديد الإطار المفاهيمى للقيمة Conceptualization of a value ، وتنظيم النسق القيمى Organization of a value .

٥ - الوصف Characterizing :

ويتضمن : تعميم generalized مجموعة الاتجاهات والقيم المعطاة فى نسق متسق فى سلوك الفرد ، وكذا فلسفته فى الحياة One's philosophy of life .

الأهداف الانفعالية فى الرياضة المدرسية Affective objectives :

تعد أغراض النمو النفسى والاجتماعى من الأغراض المهمة فى مجال التربية الرياضية المدرسية لكونها تستهدف العديد من الحصائل كالتعاون والقيادة والأمانة والكياسة وحسن المعاملة والطاعة واحترام القواعد والقوانين والسلطة وغيرها .

ومع أن العديد من معلمى التربية الرياضية يدركون أهمية الأهداف الانفعالية ، إلا أنهم لا يهتمون كثيراً بقياس حصائل تلك الأهداف ، ومن ثم لا يهتمون بالعمل على تحقيقها ، انطلاقاً من كونهم يعتقدون أن الأهداف المعرفية والأهداف النفس - حركية هى الأولى بالاهتمام وأن العمل على تحقيق الأهداف المعرفية والنفس حركية يمكن أن يتضمن تحقيق الأهداف الانفعالية .

وفى هذا الصدد يحدد كل من فيلبس وهورناك Phillips and Hornak ١٩٧٩م ومود Mood ١٩٨٢م أسباب عدم اهتمام معلمى التربية الرياضية بقياس السلوك الانفعالى على النحو التالى :

١ - أن كل من السلوك المعرفى والسلوك المهارى يمكن أن يتغير خلال فترات زمنية قصيرة بالمقارنة بالسلوك الانفعالى ، فالاتجاهات والاهتمامات والقيم تتطلب فى العادة فترات زمنية أطول لكى تتغير .

٢ – صعوبة تحديد وتقويم العديد من الموضوعات المتعلقة بالمجال الانفعالى كالتكيف والميول (الاهتمامات) والاتجاهات والتذوق والقيم وغيرها. بالإضافة إلى أن إجراءات تقويم السلوك الانفعالى غير مستقرة (ثابتة) ومن ثم يرى المعلمون أنه قد يصعب الاعتداد بنتائج هذا التقويم.

٣ – إدراك بعض المعلمين لصعوبة تدريس بعض الموضوعات المتعلقة بالمجال الانفعالى، كأن يقوم المعلم بتدريس كيفية اكتساب التلميذ اتجاهات إيجابية نحو دروس وأنشطة التربية الرياضية مثلا.

٤ – نظراً لكون القياس والتقويم فى المجال الانفعالى يتطلب استخدام أساليب التقرير الذاتى Self - reporting or rating على القوائم أو (الاستخبارات – الاستبيانات) وهى تتطلب استجابة المفحوص عما يعتقد أنه حقيقى، لذا فقد يحاول معظم التلاميذ والطلاب أن يُجملوا استجابتهم لإرضاء المعلم وإدارة المدرسة.

٥ – أن معظم معلمى التربية الرياضية غير مدربين تدريباً مناسباً على قياس وتقويم السلوك الانفعالى.

٦ – قد يطرح بعض معلمى التربية الرياضية التساؤل التالى: لماذا يفترض أن تطوير السلوك الانفعالى للتلاميذ والطلاب فى المدرسة هو مسئوليتهم؟ ولماذا لا يكون مسئولية المعلمين الآخرين بالمدرسة كمعلمى اللغات أو الرياضيات أو غيرهم من المعلمين.

٧ – أن الوقت المخصص لدروس التربية الرياضية فى معظم نظم التعليم لا يسمح بقياس كل من المجالات: المعرفية والنفس حركية والانفعالية، ومن ثم تعطى الألوية للمجالين الأول والثانى.

وتشير لومبكين ١٩٩٤م إلى أن الأغراض الانفعالية فى المجال الرياضى تستهدف تطوير الاتجاهات Attitudes والتذوق (التقدير)Appreciations والقيم

Values ، وتبين أن هذا المجال يتكون من كل من البعدين الاجتماعى والانفعالى Social and emotional وقد حدت أغراض كل بعد من البعدين كالتالى :

أولاً : البعد الاجتماعى Social dimension ويشمل :

١ - الثقة بالنفس Self - Confidence

٢ - الانتماء Affiliation

٣ - الأحكام القيمية Value Judgments

٤ - نمو الشخصية Character Development

٥ - مهارات الاتصال Communication Skills

٦ - الكياسة (اللباقة) Courtesy

٧ - اللعب النظيف Fair Play

ثانياً : البعد الانفعالى Emotional ويشمل :

١ - التحكم الذاتى (السيطرة على الذات) Self - Control

٢ - التعبير عن الذات Self - Expression

٣ - خفض التوتر Tension reduction

٤ - تعلم كيفية الفوز والهزيمة Learning how to win and how to lose

٥ - الانضباط الذاتى Self - Discipline

وتبين (لومبكين) أن تعلم الثقة بالنفس والكياسة (اللباقة) واللعب النظيف والروح الرياضية والأحكام القيمية يتم من خلال التعاون بين الطلاب (اللاعبين) فى الأنشطة الرياضية الجماعية . وأن مهارات الاتصال والانتماء تتحقق أيضاً من خلال كل من الأنشطة الرياضية الجماعية والفردية . فى حين تتحقق جميع أغراض البعد الانفعالى عندما يشترك الطالب (اللاعب) فى أنشطة رياضية تتطلب أن يتحدى فيها نفسه أو آخرين .

المجال النفس - حركى :

فى عام ١٩٥٦م نشر بلوم Bloom, B.S تصنيفه الشهير للمجال المعرفى والذى شرح فيه الوظائف العقلية المرتبطة بالأهداف التربوية educotional objectives. وبعد مرور فترة من الزمن قدم كراثهول وبلوم ومازيا Krathwohle, Bloom, and Masia ١٩٦٤م التصنيف الخاص بالمجال الانفعالى، وحتى عام ١٩٨٥م صدر ما لا يقل عن اثنى عشر تصنيفاً للمجال النفس - حركى، كما ظهر العديد من المجالات الفرعية Subdomains لهذا المجال أسوة بكل من المجالين المعرفى والانفعالى.

فالمجال الانفعالى - على سبيل المثال - يتألف من الشخصية، والميول، والقيم، والاتجاهات، والتكيف الاجتماعى وغيرها من المجالات الفرعية، التى تبدو مستقلة كل منها عن الآخر.

وفى مجال الأداء الحركى عند الإنسان Human Motor Performance نجد أن معلمى التربية الرياضية يركزون على مجالات فرعية كاللياقة البدنية، والمهارات الحركية الرئيسية Fundamental Motor Skills، والمهارات الخاصة بالأنشطة الرياضية Specific Sport Skills.

ومع مطلع السبعينيات من القرن الماضى أقر كل من مارتيوزا Martuza ١٩٧٧م، وسافريت Safrit ١٩٨١م، وهوبكنز وستانلى Hopkins and Stanely ١٩٨١م، وبرينان Brennen ١٩٨٣م، وأولريش Ulrich ١٩٨٤م استخدام مصطلح «مجال - Domain» أو «Universe» لكل مظاهر السلوك المتعلقة بالأداء الحركى Motor performance.

ويرى كيركندال وجروبر وجونسون Kirkendall, Gruber, and Johnson ١٩٨٧م أنه يمكن استخدام مصطلح « المجال الحركى Motor domain» كمرادف لمصطلح «المجال النفس - حركى Psychomotor domain» حيث يتطلب هذا المجال القيام بواجبات (مهام) حركية تتضمن المهارات الحركية الدقيقة Fine motor Skills كالنقر على الآلة الكاتبة، أو العزف على البيانو، أو القيام بواجبات (مهام) تتضمن

المهارات الحركية الكبيرة Gross Motor Skills وهى مهارات تركز على الواجبات والمهام التى تتطلب استخدام حركات الأطراف والجسم خلال الفراغ (المكان).

ويضيف كيركندال وجروبر وجونسون أن المجال النفس - حركى (المجال الحركى) يتضمن العديد من الاختبارات التى تستهدف قياس اللياقة البدنية Physical Fitness، واللياقة الحركية Motor Fitness ، والقدرة الحركية Physical ability ، والإدراك الحركى Motor - Perceptual ، والمهارات فى الألعاب الرياضية Sports Skills.

مما سبق يتضح أن المجال النفس - حركى يهتم بالسلوك الحركى motor Behavior (Movement) وهو سلوك يشير إلى الحركات العلنية (الظاهرة) للعضلات الهيكلية Skeletal Muscles الإرادية، فكل حركة للجسم يمكن ملاحظتها تعد سلوكًا حركيًا، حيث ينظر إلى الحركة Movement على أنها تحدث نتيجة للانقباضات العضلية التى تحدث نتيجة لعمل الجهاز الحركى الذى يتضمن عمل مجموعة من أجهزة الجسم كالجهاز العضلى والجهازين العظمى والمفصلى الذين يعملون كروافع Levers والجهاز العصبى وغيرها من الأجهزة.

ويشتمل المجال النفس - حركى على تنوع عريض من أنماط السلوك الحركى، فإذا كان المجال المعرفى يشتمل على أنماط السلوك (يعرف - Knowing) والمجال الانفعالى يشتمل على أنماط السلوك (يشعر-Feeling) ، فإن المجال النفس - حركى يشتمل على أنماط السلوك (يفعل - Doing) ، وفى هذا الخصوص يوضح سنجر Singer ١٩٧٢م، أن من أهم ما يميز المجال النفس - حركى هو أن الاستجابات فيه استجابات بدنية Physical responses ، وهو مجال يركز على الحركات البدنية وكيفية التحكم فيها وتوجيهها.

المستويات الرئيسية للمجال النفس - حركى :

تصنيف سمبسون Simpson ١٩٧٢م:

يتضمن هذا التصنيف سبعة مستويات كبيرة هى :

١ - الإدراك : Perception

٢ - التهيؤ (الاستعداد) : Set or Readiness

٣ - الاستجابة الموجهة : Guided Response

٤ - آلية الفعل (الأداء) : Mechanism

٥ - الاستجابة الظاهرة المركبة : Complex Overt response

٦ - المواءمة (التعديل) : Adaptaion

٧ - الإبداعية : Origination

ويحدد هارو Harrow ١٩٧٢م ستة مستويات للمجال النفس - حركى هى:

١ - الحركات المنعكسة Reflex movements وتتضمن منعكسات موضعية Segmental reflexes ، والمنعكسات بين الفلقية Intersegmental reflexes ، والمنعكسات فوق الفلقية Suprasegmental reflexes .

٢ - الحركات الأساسية الأصلية Basic - Fundamental Movements وتتضمن الحركات الانتقالية Locomotor Movements ، والحركات غير الانتقالية Nonlocomotor Movements ، وحركات المعالجة اليدوية Mani Pulative Movements .

٣ - القدرات الإدراكية Perceptual abilities وتتضمن: التمييز الحركى Kinesthetic discrimination ، والتمييز البصرى Visual discrimination ، والتمييز السمعى auditory discrimination ، والتمييز اللمسى tactile discrimination ، والقدرات التوافقية Coordinated abilities .

٤ - القدرات البدنية Physical abilities وتتضمن: التحمل endurance والقوة العضلية Strength ، والمرونة Flexibility ، والرشاقة agility .

٥ - الحركات الماهرة : Skilled movements وتتضمن: المهارة التكيفية البسيطة Simple adaptive Skill ، والمهارة التكيفية المركبة Compound adaptive Skill ، والمهارة التكيفية المعقدة Complex adaptive .

٦ – حركات الاتصال Nondiscursive movements وتتضمن: الحركات التعبيرية expressive movements ، والحركات التفسيرية Interpretive Movements .

ويحدد جويت ومولان Jewett and Mullan ١٩٧٧م ثلاثة مستويات للمجال النفس – حركى هى:

١ – حركة عامة Generic movement وتشمل: الإدراك Perceiving ، والتنميط patterning .

٢ – حركة مرتبة Ordinative movement وتشمل: التعديل adapting ، والصقل refining .

٣ – حركة ابتكارية Creative movement وتشمل: التنويع Varying ، والإبداع المرتجل Improvising ، والتركيب (الدمج) Composing .

ويقدم جانسما Jansma ١٩٨٨م وصفاً لأربعة مكونات أساسية للمجال النفس – حركى هى:

١ – المكون البدنى Physical Component

٢ – المكون الحركى Motor Component

٣ – مكون اللياقة Fitness Component

٤ – مكون اللعب Play Component

ويصف «جانسما» المكون البدنى على أنه الحالة البنائية (التركيبية) للفرد من الناحية التشريحية anatomical ، ويدل المكون الحركى على نوعية أنماط الحركة لدى الفرد من حيث الجودة quality of Movement Patterns ، ويشير مكون اللياقة إلى مقدار الحركة quantity of movement بمعنى الاستمرارية فى بذل الجهد، ويشير مكون اللعب إلى ذروة التطور الذى يحدث للفرد خلال المجال النفس–حركى ويضيف «جانسما» أنه لكى يصبح الفرد لاعباً متميزاً فإنه يحتاج إلى كل من الكفاءة (المقدرة) البدنية والحركية واللياقة مجتمعة .

وتحدد أنجيلا لومبكن Angela Lumpkin ١٩٩٤م مستويين رئيسين للمجال النفس - حركى هما :

أولاً : المهارات الحركية الأساسية Basic movement skills :

وتتضمن الآتى :

١ - مفاهيم الحركة Movement Concepts وتتضمن : الوعى البدنى Body awareness ؛ والوعى المكانى Spatial awareness الذى يتضمن : المكان (الحيز) Space ، والاتجاه Direction ، والمستوى Level ، والسبل (الطرق) Pathways ، وخصائص الحركة qualities of movement من حيث التوقيت والقوة والانسيابية .

٢ - الحركات الانتقالية Locomotor Movements مثل : حركات المشى Walking والجرى Running والوثب Jumping والإنزلاق Sliding وغيرها .

٣ - الحركات غير الانتقالية Nonlocomotor Movements مثل : حركات الدفع Pushing والإطالة (الشد) Stretching والرفع Lifting والتأرجح Swinging واللوى (اللف) Twisting وغيرها .

٤ - مهارات المعالجة اليدوية Manipulative skills وتتضمن : الرمى Throwing والمسك Catching والضرب Striking والركل Kicking .

ومن الملاحظ أن هذه المهارات يمكن أن تؤدى باستخدام الكرات balls والحبال Ropes والأطواق Hoopes والمضارب rackets (كمضرب التنس وتنس الطاولة) والعصى bats (كعصا الهوكى والبيسبول والكركيت) وغيرها من الأدوات والوسائل .

٥ - المهارات الحركية الإدراكية Percepual - Motor Skills وتظهر عند الحا إلى التوافق بين العين واليد لضرب كرة التنس بالمضرب أو عند الحاجة إلى ز رد الفعل (زمن الرجع) Reaction Time للتصرف بسرعة وفقاً لحرك

المنافس. وتظهر المهارات الحركية الإدراكية كقدرات مستقلة تلعب دوراً مهماً فى إتقان التعليم وتطوير مستوى الفرد الرياضى فى مهارات الألعاب games Skills كالمسك والرمى والضرب وغيرها.

٦ - **المهارات الرئيسية فى اللعبة** Fundamental game Skills وهى تشير إلى المهارات التى يتعلمها الفرد فى الأنشطة الرياضية المختلفة وهى تعتمد على كل ما سبق الإشارة إليه فيما يتعلق بالمستوى (أولاً) والخاص بالمهارات الحركية الأساسية.

ثانياً : اللياقة البدنية Physical Fitness وهى تمثل المستوى الآخر من نواتج المجال النفس - حركى، حيث يأتى فى المقام الأول كلا من:

١ - **التحمل الدورى التنفسى** Cardiorespiratory Endurance وهو يشير إلى القدرة على الاستمرار فى العمل البدنى الذى يتطلب استخدام أكسچين الهواء الجوى للاستمرار فى بذل المجهود والذى يعد المكون الأكثر أهمية بالنسبة للياقة البدنية.

٢ - **القوة العضلية والتحمل** Muscular Strength and Endurance وتشير إلى القدرة على الاستمرار فى بذل القوة القصوى عند مقاومة تتطلب استخدام القوة العضلية والتحمل العضلى (تحمل القوة).

٣ - **المرونة** Flexibility وتشير إلى مدى الحركة على مفاصل الجسم المختلفة، وهى كصفة يمكن اكتسابها عن طريق تدريب العضلات على الإطالة Stretching.

٤ - **تركيب الجسم** Body Composition وهو يشير إلى مجموعة الأجزاء أو العناصر التى تشكل الكل عندما تترابط مع بعضها البعض، ويهتم تركيب الجسم بتحديد الأجزاء والعناصر التى يتكون منها الكل (الجسم)، والطريقة التى تترابط بها تلك الأجزاء والعناصر لتشكل الكل، وكذا التنظيم الذى يتكون

منه هذا الكل، ويرتبط تركيب الجسم باللياقة البدنية لكونه يستهدف تقدير الكمية النسبية للدهون فى الجسم.

٥ - **الرشاقة** (سرعة تغيير الاتجاه) Agility ، والتوازن Balance ، والتوافق العضلى العصبى Neuromuscular Coordination ، والسرعة Speed ، والقدرة العضلية Muscular power وجميعها قدرات تلعب أدواراً ووظائف مهمة بالنسبة للياقة البدنية، فالرشاقة مثلاً تمكن الفرد من تغيير اتجاهه بسرعة خلال الأداء فى النشاط، والتوازن يمكن الفرد من تهيئة الجسم لأداء الحركات والمهارات بثبات ودقة، والتوافق العضلى العصبى يتضمن : زمن الرجع (زمن رد الفعل)، وتكامل الإدراك الحسى مع وظيفة الحركة، بينما تمكن السرعة الفرد من التنفيذ السريع للحركة، وتمكنه القدرة من التنفيذ السريع القوى للحركة لكونها تربط بين السرعة والقوة العضلية فى آن واحد.

وتشير لوريتا ستالنجيز Stallings , L.M ١٩٨٢م إلى أن السلوك النفس-حركى يمكن أن يتضمن خمسة مجالات رئيسية هى:

١ - **التربية الرياضية والمهارات الترويحية**، وتشمل: الرياضة Sports والرقص Dance والتمرينات Exercise.

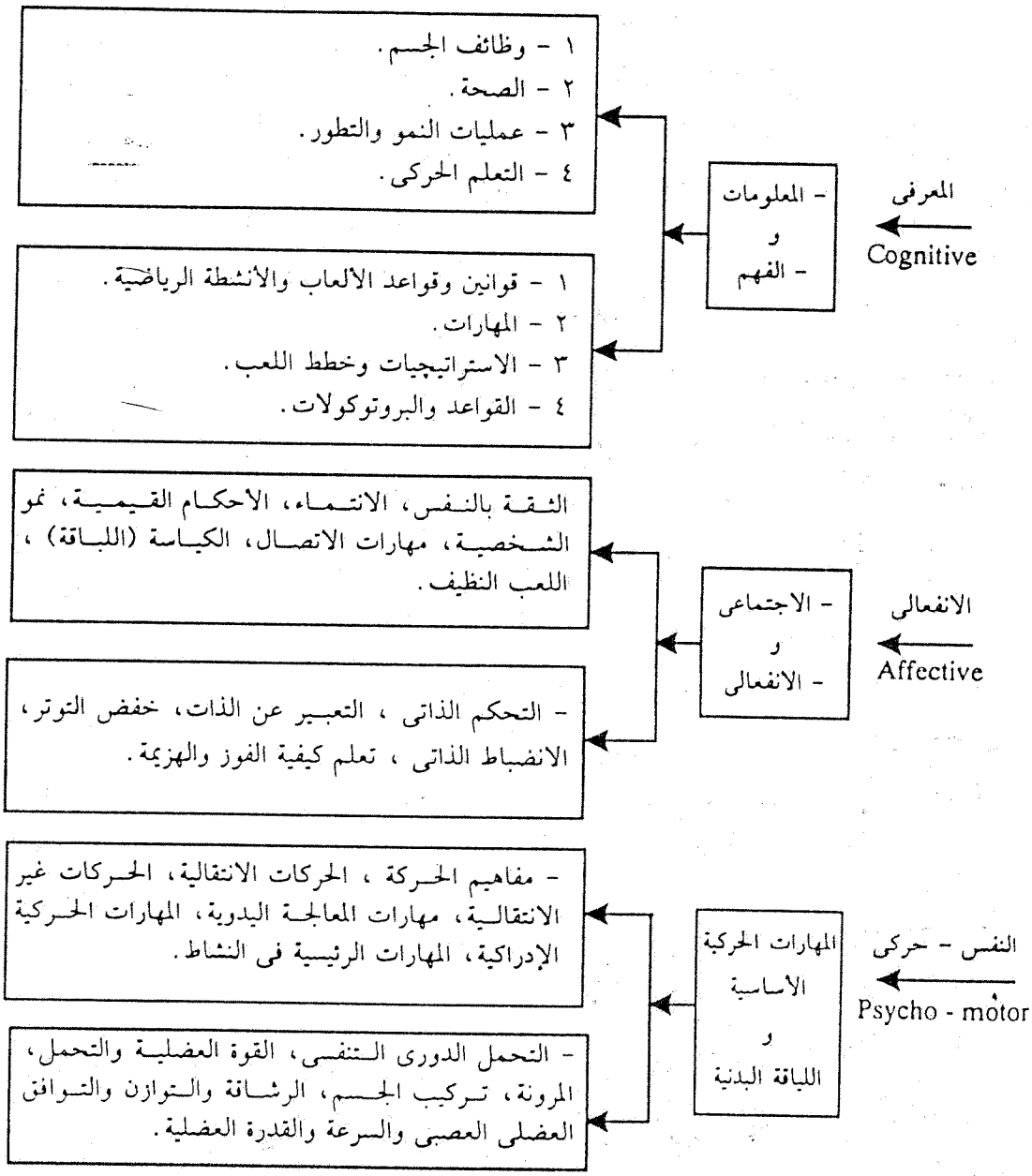
٢ - **مهارات الاتصال** Communication Skills وتشمل: الكتابة باليد والنقر بالأصابع (على الآلة الكاتبة مثلا) والاختزال Shorthand.

٣ - مهارات اللغة Language Skills وتشمل: الكلام، والتعبير بالوجه Facial والتعبيرات والإيماءات الخاصة Gestures.

٤ - **المهارات المهنية** Vocational Skills وتشمل: الحرف اليدوية Crafts واستخدام الأدوات أو الآلات (كقيادة السيارة أو الطائرة مثلا) وتشغيل الأجهزة equipments operation وغيرها.

٥ - مهارات الفنون الجميلة Fine Art Skills وتشمل: اللعب بالأدوات Instrument Playing ، وأعمال الخزف والتصوير والرسم Drawing والغناء Singing وغيرها.

ويلخص الشكل (١ - ٢) الأهداف المعرفية والانفعالية والنفس- حركية للتربية البدنية والرياضة.



شكل (١ - ٢)

أهداف التربية البدنية والرياضة

(عن Lumpkin ١٩٩٤م ص : ١٥)

يتضح مما سبق أن المجال المعرفى يهتم بالأداء العقلى Mental Performance الذى يستهدف تقويم تذكر قواعد وقوانين اللعب والتاريخ والمبادئ الأساسية بالنسبة للأنشطة والألعاب الرياضية، وأيضاً تقويم القدرة على فهم وتحليل وتطبيق الاستراتيجيات وقواعد اللعب المختلفة، كما نستخلص أن القياس المعرفى يتطلب تحديد الأهداف السلوكية Behavioral Objectives فى هيئة مصطلحات سلوكية خاصة حتى يتمكن المفحوص من إنجازها. حيث يلاحظ أن هذه الأهداف السلوكية يجب أن تتضمن مصطلح (العمل أو الفعل - Doing) حتى يتمكن المفحوص من إنجازها من خلال استخدام أفعال Verbs تدل «على الأداء أو العمل - action» ، ومن أمثلة ذلك الآتى: (يصف، يفسر، يعرّف، يحدد، يفرّق، . . . إلخ).

كما يتضح أن القياس الانفعالى يهتم بالاتجاهات والسمات النفسية Psychological Traits وغيرها من الموضوعات المرتبطة بالأداء البشرى فى مجال الرياضة البدنية، وأن المفحوص يستخدم بعض الأفعال التى تدل على الأداء أو العمل فى المجال الانفعالى مثل (يسأل ، يرد أو يجيب) يُنصت، يظهر أو يكشف، يصمم أو يعزم، . . . إلخ).

ويهتم المجال النفس - حركى بالأداء البدنى Physical Perfoemance فى الرياضة، ويتطلب القياس النفس - حركى تحديد الأهداف السلوكية حتى يتمكن المفحوص من إنجازها باستخدام أفعال تدل على الأداء أو العمل مثل: (يتحرك، يقفز، يجرى، يثب، يسبح ، يمسك، يرمى، . . . إلخ).

وعموماً فقد اتفق معظم العلماء، على أن الاستجابة المعرفية تستخدم فعل (يعرف Knowing) وأن الاستجابة الانفعالية تستخدم فعل (يعى أو يشعر aware or feeling)وأن الاستجابة النفس - حركية تستخدم الفعل (يفعل - doing).

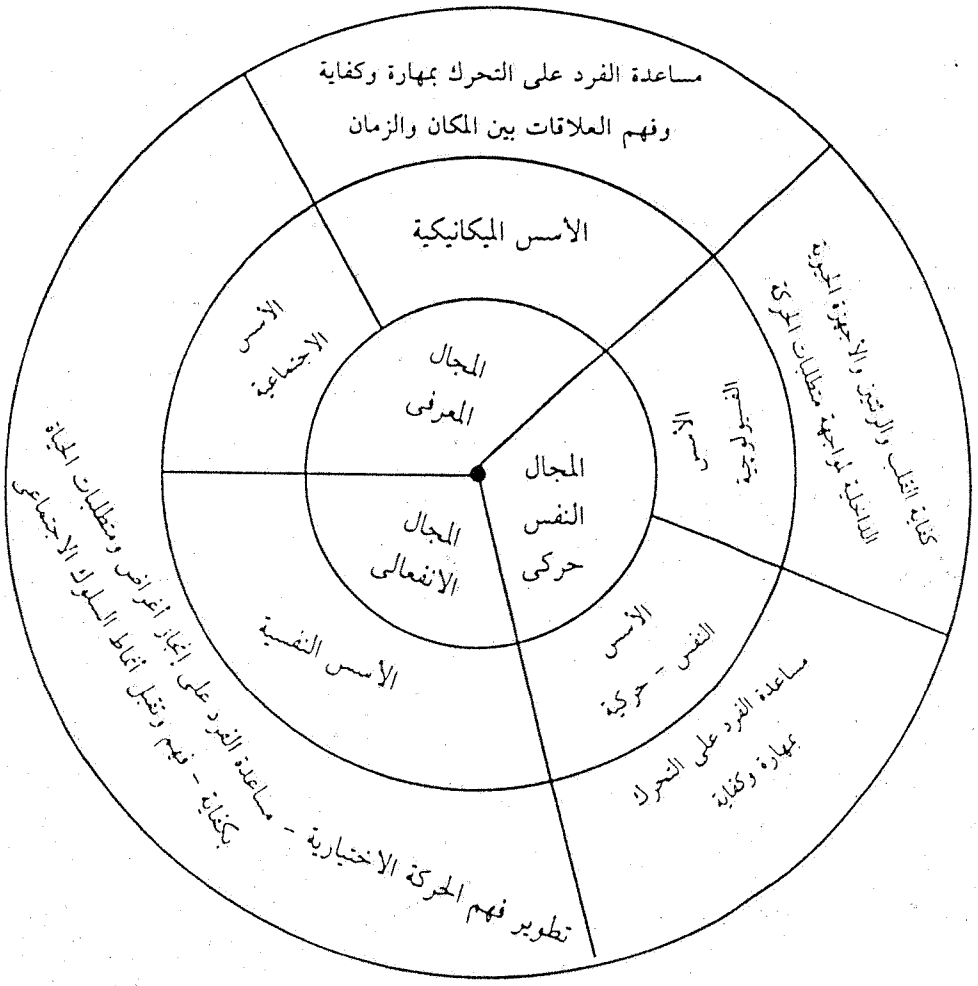
العلاقة بين المجالات الثلاثة والبنية المعرفية للتربية البدنية والرياضة:

يوضح الشكل (١ -٣) العلاقة بين المجالات الثلاثة: المعرفى، والانفعالى، والنفس - حركى، حيث يلاحظ من هذا الشكل الآتى:

١ - تتضمن الدائرة الداخلية المجالات الثلاثة ، المعرفي، والانفعالي، والنفس حركي كمجالات للتعلم Domains of Learning .

٢- تتضمن الدائرة الوسطى البنية المعرفية Body of Knowledge للتربية البدنية والرياضة متمثلة فى خمس مجموعات من الأسس Foundations هى : الأسس الميكانيكية ، والأسس الاجتماعية ، والأسس النفسية، والأسس النفس - حركية، والأسس الفسيولوجية .

٣- الدائرة الثالثة وتتضمن أهداف التربية البدنية والرياضة .



شكل (١ - ٣)
العلاقة بين البنية المعرفية وأهداف التربية البدنية والرياضة

٤- أن البنية المعرفية للمجال النفس - حركى تتضمن كلاً من الأسس النفس-حركية والأسس الفسيولوجية للحركة متمثلة فى التوافق العضلى العصبى، المهارات الحركية، التحمل العضلى، القوة العضلية، كفاية القلب والجهازين الدورى والتنفسى.

٥ - وأن البنية المعرفية للمجال الانفعالى تتضمن الأسس النفسية والاجتماعية متمثلة فى القيم والشعور والانفعالات والاتجاهات والميول وغيرها.

٦- وأن البنية المعرفية للمجال المعرفى تتضمن الأسس الاجتماعية والأسس الميكانيكية متمثلة فى استرجاع الحقائق، الفهم، معرفة المبادئ الميكانيكية، تذكر القوانين والقواعد والاستراتيجيات وغيرها.

٣ - وظائف القياس فى المجال الرياضى

للقياس فى المجال الرياضى مجموعة مهمة ومتباينة من الوظائف (الاستخدمات)، من هذه الوظائف (الاستخدامات) ما يلى:

تحديد الأهداف Determining The Objectives

تستخدم نتائج الاختبارات والمقاييس فى المجال الرياضى فى إقرار الأهداف المستهدفة من عمليات التعليم والتدريب حيث يتم ذلك عن طريق تحديد الحاجات needs الحقيقية للتلميذ أو اللاعب من خلال ما تكشف عنه نتائج القياس المختلفة.

فالقياس قبل البدء فى عملية التعليم (التدريب) وأثناء الاستمرار فيها يعطى للمعلم (المدرب) فرصة الوقوف على المستوى البدنى والمهارى والمعرفى للتلميذ (اللاعب)، ومن ثم يتاح للمعلم (المدرب) إمكانية تحديد الأهداف التعليمية والتعديل فى هذه الأهداف وفق ما تظهره عمليات القياس من نتائج قبل وأثناء تنفيذ برنامج التعليم (التدريب).

الحث الدافعى Motivating :

يمكن الاستفادة من نتائج الاختبارات والمقاييس فى المجال الرياضى كوسائل لاستثارة دافعية الأفراد نحو الممارسة والتعليم والتدريب ومحاولة الوصول إلى أعلى المستويات الرياضية، وكذا محاولة التفوق على الزملاء والمنافسين، وزيادة الاهتمام بعمليات التعليم والتدريب، وبخاصة عندما يتعرفون على مستوياتهم الفعلية ويقارنون مستوياتهم بمستويات أقرانهم.

فقد أظهرت نتائج الدراسات والبحوث العلمية المختلفة أن للاختبارات تأثير إيجابى كبير فى حث الأفراد المتدربين، فتوقع المتدرب أنه سيتم امتحانه أو اختباره فى موضوع ما قد ينبهه Stimulate لكى يعد نفسه لهذا الامتحان وذلك بالاجتهاد فى الدراسة والتدريب على الموضوعات التى سوف يختبر فيها، حيث تؤدى مثل هذه الدراسة والتدريب السابق على موقف الامتحان إلى تطوير وتحسين أداء هذا المتدرب.

فعندما يعلن المدرب الرياضى - على سبيل المثال - أنه سوف يقوم بتقويم مستوى الدهون فى الجسم لفريق المصارعة الذى يقوم بتدريبه، فإنه يتوقع نتيجة هذا الإجراء أن يزداد اهتمام اللاعبين الذين يتمتعون بوزن زائد بالعمل على تخفيض أوزان أجسامهم.

وعندما يطبق اختبار لقياس المهارة فى نشاط ما من الأنشطة الرياضية، فإنه يمكن الاستفادة من نتائج هذا الاختبار فى تعريف التلاميذ (اللاعبين) بقدرتهم المهارية فى هذا النشاط مما يزيد من حماسهم ودوافعهم نحو محاولة تطوير مهاراتهم فى هذا النشاط.

ويلاحظ أنه لكى يستفاد من الدافعية Motivation فى مجال القياس، فإنه يجب على المعلم (المدرب) أن يمد المفحوصين بالتغذية المرتدة (الرجعية) Feedback لكى يحاولوا المحافظة على الجوانب المتميزة التى تكشف عنها نتائج القياس.

ولكى يستفاد من معلومات التغذية المرتدة فإنه يجب إمداد المفحوصين بهذه المعلومات بصورة مباشرة وفورية، لأن مبدأ الإمداد الفورى بالتغذية المرتدة يعد من أهم متطلبات الحث الدافعى، حيث يمكن الاستفادة منها فيما يلى:

- تنشيط المفحوصين وجعلهم أكثر ميلاً للمثابرة والاستمرار فى بذل الجهد البدنى أثناء المنافسة والتدريب على الأنشطة البدنية التى تتطلب بذل مجهود عالى الشدة.
- مساعدة المفحوصين على التقويم الذاتى لأنفسهم، عندما يعرفون مستوياتهم، وعما إذا كانت جيدة أم رديئة بطيئة أم سريعة.
- تزيد من كفاءة العمل العضلى والأداء من مواقف التعلم والتدريب والمنافسة.
- تساعد على تثبيت الاستجابات الناجحة ومحاولة إصلاح أو تجنب الاستجابات الخاطئة.
- زيادة الميل والاتجاه نحو الوصول إلى المستويات الأعلى.

ومن ناحية أخرى فقد لوحظ أن كل مفحوص يميل إلى مقارنة مستوى أدائه بمستوى أداء غيره من المفحوصين، كما أنه يميل أيضاً إلى مقارنة أدائه السابق بأدائه اللاحق لمعرفة ما إذا كان الأداء الثانى أفضل من الأول أم لا؟ . ولعل من أهم الشروط التى يجب على المعلم (المدرب) أن يتبعها لضمان فعالية عمليات القياس كوسائل مؤثرة فى دافعية الإنجاز لدى المفحوصين ما يأتى:

- توضح الهدف من الاختبار.
- تطبيق الاختبارات (وبخاصة بالنسبة للاختبارات الميدانية) فى وجود جمهور من المشاهدين.
- الإمداد الفورى للمفحوصين بمعلومات التغذية المرتدة التى تتناول نتائج إنجازهم على الاختبار.

- إلقاء بعض التعليقات الخاصة التى تعطى المفحوص فكرة واضحة عما أنجزه بنجاح، وعما يحتاج إليه من تقوية بالنسبة لبعض الجوانب.

تحديد التحصيل Determining the achievement :

من أهم دواعى القياس فى المجال الرياضى هو التعرف على مستوى التحصيل الذى يصل إليه المفحوص، ولأن التحصيل يرتبط بكل ما هو مكتسب بالتعلم والتدريب والمران والممارسة، لذا نجد أن اختبارات التحصيل تتناول موضوعات تتعلق بالمعارف والحقائق والمعلومات والمهارات المكتسبة بالتعلم، وإن كانت هناك بعض الاختبارات التى تتناول بشكل محدد الاتجاهات والميول والقيم وغيرها من الموضوعات النفسية انطلاقاً من كونها موضوعات قد تكتسب جزئياً بالتعلم.

ويعتبر التحصيل مؤشراً مهمًا للنمو والتقدم والتحسن الذى يحققه التلميذ (اللاعب) فى مجال ما من المجالات، لذلك يستهدف قياس التحصيل التعرف على مدى التقدم الذى يشير إلى مدى الزيادة فى التحصيل بالنسبة لمهارة حركية ما، ومدى النجاح الذى تحققه عملية التعليم أو التدريب، ومدى فاعلية البرنامج الدراسى أو التدريبى.

ومن ناحية أخرى يستخدم قياس التحصيل للتحقق من مدى إنجاز أهداف البرنامج، وإنجاز الأهداف الشخصية لكل مفحوص، فقد وجد أن التلاميذ (اللاعبين) يرغبون فى معرفة مستوى التحصيل الذى وصلوا إليه ومقدار التقدم الذى حققوه بالنسبة للموضوع الدراسى أو التدريبى الذى يتلقونه.

وقد تستخدم طريقتى القياس القبلى Pre - Test والبعدى Post - Test لتقدير مستوى التحصيل قبل البرنامج الدراسى (التدريبى) وبعده، كأن يقوم المعلم (المدرب) بقياس مهارة حركية ما قبل البرنامج، ثم يعيد قياس نفس المهارة بعد انتهاء الفترة الزمنية المقررة للبرنامج لكى يقف كل من المعلم والمدرب والتلميذ واللاعب على مستوى التحصيل الذى حققوه كنتيجة للبرنامج.

وقد تستخدم مقاييس التقدير Rating Scales بدلاً من الاختبارات المقننة لقياس

التحصيل فى بعض المهارات الحركية دون الحاجة إلى استخدام طريقتى القياس القبلى والبعدى.

فقد يقوم المعلم (المدرب) بتقدير درجة التلميذ (اللاعب) فى مهارة الدحرجة الأمامية Forward roll فى الجمباز الأرضى مثلاً فى نهاية العام الدراسى (الموسم التدريبى) باستخدام مقياس للتقدير، حيث تشير الدرجة التى يحصل عليها المفحوص على مستوى تحصيله لهذه المهارة، وبمقارنة درجة هذا المفحوص بدرجات غيره من المفحوصين فى نفس الفصل (الفريق) يمكن تقويم مستوى تحصيل هذا المفحوص بالمقارنة بمستويات تحصيل أقرانه من نفس الفصل (الفريق) لتحديد موقعه النسبى وعما إذا كان هذا الموقع ممتازاً أو متوسطاً أو ضعيفاً.

ومن الملاحظ أن تقويم مستوى تحصيل التلميذ (اللاعب) يستخدم المعيار المرجعى – Norm reference ، ففى طريقتى القياس القبلى والبعدى يكون هذا المعيار هو: مقارنة المفحوص بنفسه، وفى طريقة القياس الواحد كما فى نهاية العام أو الموسم التدريبى، يكون المعيار هو: مقارنة التلميذ (اللاعب) بغيره من التلاميذ (اللاعبين).

مما سبق يتضح أن تحديد تحصيل التلميذ (اللاعب) يستهدف الآتى:

١ – تقدير مدى تحصيل التلميذ (اللاعب) بمقارنته بنفسه.

٢ – التقدير النهائى لمستوى تحصيل التلميذ (اللاعب) بالمقارنة بأقرانه.

٣ – تقدير مدى مناسبة المنهج والبرنامج الدراسى (التدريبى).

٤ – تقدير مدى فعالية طرق التدريس (التدريب) المستخدمة.

٥ – تقدير مدى فعالية كل من المعلم والمدرب.

التصنيف Classification :

تستخدم نتائج القياس فى المجال الرياضى لتصنيف التلاميذ واللاعبين لوضعهم فى فئات أو مجموعات متجانسة (متشابهة) وفقاً للعديد من المتغيرات كالجنس، والطول، والوزن، ونمط الجسم، والحالة الصحية، والقدرات البدنية، والمهارات

الحركيـة الرياضية، والـقدرة العقليـة، والسمات الانفعـالية والمزاجيـة، والقدرات المعرفية، وغيرها من المتغيرات.

ويقصـد بالتصنيف هنا، مـحاولة وضع كل مجـموعة متـشابهة - فى متـغير التصنيف - فى فئة أو مجموعة واحدة، هذه المجموعة تمتاز بأنها متجانسة، بمعنى أن الفروق الـفردية فيمـا بين أفراد هذه المجـموعة تكون مـحدودة أى ليست واسـعة. ويستفـاد من المجموعات المتجـانسة فى مجال التعـليم والتدريب الرياضى لتـحقيق مجموعة مهمة من الوظائف منها:

١ - أنها تجعل البرنامج (التعليمى أو التدريبى) أكثر فعالية.

٢ - أنها تسهل على المعلم والمدرب القيام بعمليتى التعليم والتدريب.

٣ - أنها تـزيد من تماسك الجـماعة وتجـعلها أكـثر تفاعـلاً أثناء عمـليات التـعليم والتدريب.

٤ - أنها تجعل الأفراد يشعرون بالرضا والراحة الـنفسية عندما يشتركون فى الأداء مع غيرهم فى نفس مستواهم.

٥ - أنها تكفل بدرجة كبيرة الوقاية من الإصابة وبخاصة فى أنشطة الاحتكاك المباشر كالمصارعة والملاكمة والجودو والكارتيـه وغيرها، وأيضاً فى الأنشطة الرياضية التى تتطلب بذل مجهود بدنى عالى الشدة يستغرق فترات زمنية طويلة نسبياً.

٦ - أنها تحد من الإحبـاط وتزيد من الدافعية عندما يشتـرك الأفراد مع منافسين من نفس مستواهم فى اللعبة أو النشاط الذى يـتنافسون فيه، هذا بالإضافة إلى أنها تحد من فاقد العملية التعليمية.

ويستخـدم معلمو التربـية الرياضية والمدربون عنـد تصنيف أو تجميع التـلاميذ (اللاعبين) Classification or Grouping فى مجـموعات متـجانسة شـكلين من أشكال التصنيف هما:

التصنيف العام General Classification :

وهو يعنى وضع التلميذ أو اللاعب وفقاً لمستوى قدراته واستعداداته العامة فى الصف الدراسى أو المستوى التدريبى الذى يناسبه، وذلك من أجل تحقيق الأهداف العامة للتعليم أو التدريب، ويتم ذلك فى بداية العام أو الفصل الدراسى أو موسم التدريب.

التصنيف الخاص Specific Classification :

وهو يعنى وضع التلميذ (اللاعب) فى المجموعة التى تناسبه وفقاً لمستوى قدراته ومهاراته الخاصة فى بعض الأنشطة والألعاب الرياضية، وذلك من أجل تحقيق الأهداف التعليمية الخاصة بالوحدة التعليمية المقررة عند البدء فى تعليم وحدات جديدة فى المهارات الحركية فى الأنشطة الرياضية المختلفة.

وكما ذكرنا فإنه يمكن تصنيف التلاميذ (اللاعبين) وفقاً لعدد من متغيرات التصنيف كالوزن والطول واللياقة البدنية والمهارات الحركية الرياضية والذكاء وغيرها، حيث تستخدم نتائج القياس لوضع المفحوصين فى فئة أو طبقة Category or Class من بين عدد من الفئات (الطبقات) المتمايزة كالتالى:

- بالنسبة لوزن الجسم فيمكن وضع المفحوص فى فئة واحدة من الفئات التالية (أصحاب الوزن المنخفض - أصحاب الوزن المتوسط - أصحاب الوزن الزائد).
- وبالنسبة للطول فيمكن وضع المفحوص فى فئة واحدة من الفئات التالية (قصير - متوسط - طويل).
- وبالنسبة للذكاء فيمكن وضع المفحوص فى فئة واحدة من الفئات التالية (منخفض الذكاء - متوسط الذكاء - موهوب).
- وبالنسبة للياقة البدنية والمهارات الحركية الرياضية فيمكن وضع المفحوص فى فئة واحدة من الفئات التالية (المستوى المنخفض - المستوى المتوسط - المستوى المرتفع).

التنبؤ بالأداء المستقبلى Predicting Future Performance :

يمكن استخدام نتائج الاختبارات والمقاييس فى التنبؤ بالنجاح المستقبلى للتلاميذ (اللاعبين) فى العديد من الأعمال والواجبات (الأداءات) ، ويقصد بالتنبؤ هنا توقع نتائج محددة فى المستقبل فى ضوء نتائج عمليات القياس التى تتم. فنحن نقيس القدرة الحركية (MA) للتلميذ أو اللاعب فى وقت معين، وفى ضوء معرفة مستواه الراهن فى هذه القدرة، فإنه يمكن تقدير المستوى المتوقع له فى بعض أنواع السلوك التى ترتبط بهذه القدرة كالنجاح الدراسى بكليات التربية الرياضية، أو النجاح فى الأداء فى نشاط معين من الأنشطة الرياضية. فعلى سبيل المثال يمكن استخدم كل من الحد الأقصى لاستهلاك الأكسچين VO_2 Max ومستوى القلق Anxiety level كقياسات يمكن توظيفها كمتنبئات Predictors بالأداء المستقبلى فى الرياضة، كما تستخدم الكليات العسكرية فى معظم دول العالم نتائج الأداء البدنى Physical performance والفحوص الطبية Medical examination عند الالتحاق بهذه الكليات للتنبؤ بالنجاح المتوقع للطلاب كعسكريين متميزين فى المستقبل. وبناء على ذلك فإنه يمكن للمدرب الرياضى استخدام نتائج القياس فى انتقاء Selection الناشئين لبعض الأنشطة الرياضية، فإذا تقدم للإلتحاق بفريق كرة القدم بأحد الأندية الرياضية عدد ٥٠٠ ناشئ، فإن المدرب يستطيع عن طريق قياس بعض السمات والصفات والخصائص النفسية والصحية والجسمية المرتبطة بالتفوق فى كرة القدم أن يختار العدد المطلوب لتكوين الفريق على أساس أفضل المستويات فى تلك السمات والخصائص.

وعموماً فإنه يمكن استخدام نتائج أدوات القياس المقننة لتحقيق بعض الوظائف المرتبطة بالتنبؤ فى المجال الرياضى، من هذه الوظائف ما يلى:

١ - **التنبؤ بالأداء المستقبلى** Future Performance فى نشاط رياضى محدد على أساس نتائج الأداء الحالى على بعض اختبارات القدرات أو الاستعدادات البدنية المرتبطة بهذا النشاط.

٢ - **توجيه الفرد** الرياضى إلى الانتقال من نشاط رياضى إلى آخر، على أساس أن ما يمتلكه هذا الفرد من قدرات وسمات ترتبط بالتفوق فى نشاط آخر، غير النشاط الذى يتدرب عليه.

٣ - **التنبؤ بالنجاح** فى بعض الأنشطة الرياضية المتشابهة، فالتلميذ الذى يظهر مقدرة على التحكم فى الكرة أثناء التحرك فى دروس التربية الرياضية بالمدرسة يمكن التنبؤ له بالتفوق فى بعض الأنشطة التى ترتبط بهذه المقدرة ككرة السلة واليد مثلاً.

التوجيه والإرشاد Cuidance and Counselling :

يقصد بالتوجيه والإرشاد فى المجال الرياضى تلك المجموعة من الخدمات التى يقدمها المعلم (المدرب) أو إدارة المدرسة (النادى) للتلميذ (اللاعب) بهدف مساعدته على الآتى:

- اختيار الأنشطة الرياضية التى تتناسب مع قدراته وميوله.
- إيجاد الحلول المناسبة للمشكلات الشخصية التى تواجه بعض التلاميذ (اللاعبين) وبخاصة فيما يتعلق بالمشكلات المرتبطة بالنشاط الرياضى كالمشكلات البدنية، والحركية، والمهارية، والنفسية، والاجتماعية.

وتستهدف عمليات التوجيه والإرشاد فى المجال الرياضى تعريف التلميذ (اللاعب) بالمشكلات المختلفة المرتبطة بهذا المجال، والعمل على تقديم الخدمات والمساعدت المناسبة للتغلب على هذه المشكلات.

ولكى يمكن القيام بذلك، فإنه يلزم توافر معلومات وبيانات كافية ودقيقة عما يمتلكه التلميذ (اللاعب) من سمات وقدرات وميول واتجاهات، بالإضافة إلى معلومات عن السمات والخصائص المميزة Attribute لكل تلميذ (لاعب) على حدة.

وتعتبر الاختبارات والمقاييس من أهم الوسائل التى تستخدم للحصول على المعلومات التى يمكن الإفادة منها بالنسبة لعمليات التوجيه والإرشاد فى المجال الرياضى، وذلك لكونها تساعد كل من التلميذ والمعلم على الآتى:

- استكشاف التلميذ لذاته Self - exploration ، وفهمه لنفسه .
- الكشف عن جوانب القوة وجوانب الضعف فيما يمتلكه التلميذ من سمات وقدرات واستعدادات ومهارات وصفات وغيرها ، مما يتيح الفرصة للمعلم لكى يقوم بإرشاد التلميذ للعمل على التغلب على جوانب الضعف ، والتأكيد فى الوقت نفسه على استثمار جوانب القوة (التميز) بصورة تمكنه من الوصول إلى أعلى المستويات . وحتى تحقق عمليات التوجيه والإرشاد النتائج المرجوة منها ، فإنه يلزم أن يتعاون كل من التلميذ (اللاعب) والمعلم (المدرب) فى التعرف على المشكلة ، واقتناع كل منهما بأهمية هذه المشكلة ، ثم الاشتراك معاً فى وضع الحلول المناسبة لها ، وكذا تحديد طبيعة خدمات التوجيه والإرشاد اللازمة والعمل على أن يتقبل التلميذ (اللاعب) تلك الخدمات بوعى واقتناع .

مراقبة التقدم Monitoring Progress :

عندما يتم القيام بعملية القياس بطريقة دورية ومنتظمة ، فإنه يمكن الاستفادة من النتائج التى تكشف عنها هذه العمليات فى تطوير مهارات الطالب (اللاعب) Developing Student Skills على النحو التالى :

١ - يمكن استخدام النتائج الدورية لعملية القياس كمرشد وموجه للطالب (اللاعب) يبين مقدار التقدم الذى يحققه نحو أهداف العملية التعليمية (التدريبية) . فعمليات القياس الدورية أثناء القيام بالتعليم (التدريب) يمكن الاستفادة منها فى تعديل طريقة التدريس (التدريب) ، وكذا التعرف على صعوبات التعلم والأخطاء الشائعة ، مما يترتب عليه اتخاذ إجراءات التعديل الإصلاحى Corrective action المناسبة بصورة فورية وعاجلة ، مما يؤدى بدوره إلى تحسين وتطوير مهارات الطالب (اللاعب) .

٢ - ويمكن أيضاً الاستفادة من النتائج الدورية لعملية القياس فى تقويم المجهودات المبذولة من قبل المعلم (المدرب) لتصحيح مسار عملية التعليم (التدريب) ، فالقياسات الدورية تكشف بشكل واضح عن مدى فعالية إجراءات التعديل الإصلاحى التى قدمها المعلم (المدرب) ومدى نجاح أو فشل هذه التعديلات .

٣ - تعرف الاختبارات التى تستخدم لمراقبة تقدم الطالب (اللاعب) باسم : الاختبارات التشخيصية Diagnostic Tests لكونها تستهدف الكشف عن جوانب الضعف والقصور بالنسبة لموضوعات تتعلق بالتحصيل والإنجاز فى القدرات والمهارات الحركية الرياضية .

٤ - ويمكن الاستفادة من القياسات الدورية فى الكشف عن الأفراد الذين يعانون من صعوبات فى التعلم، فعندما يظهر بعض الطلاب (اللاعبين) مستويات أداء منخفضة فى مرات القياس المختلفة التى تؤخذ لهم خلال السنة أو الفصل الدراسى أو الموسم التدريبى، فإنه يلزم فى مثل هذه الحالة توجيه هؤلاء الطلاب (اللاعبين) نحو العمل على تحسين مستوياتهم، أو تغيير مسارهم بالنسبة للنشاط الذى يتدربون عليه .

٥ - ويمكن كذلك الاستفادة من القياسات الدورية للطلاب (اللاعبين) فى زيادة الاحتفاظ بالتعلم increase retention of learning وذلك من خلال انتقال أثر القياس من مرة لأخرى ومن موضوع لآخر، وتمكين الطلاب transfer (اللاعبين) من تطبيق ما تعلموه من مهارات ومعلومات فى مواقف متجددة ، هذا بالإضافة إلى تركيز انتباه الطلاب (اللاعبين) على الحصائل الأكثر أهمية بالنسبة للموضوعات التى يتدربون عليها .

والخلاصة أن القياسات الدورية للطلاب (اللاعبين) تعد من أهم وظائف القياس فى المجال الرياضى، لكونها تعمل على مراقبة تقدم مستوى الطالب (اللاعب) لضمان التطوير والتحسين المستمرين لهذا المستوى .

تعيين المستوى Placement :

يقصد بتعيين المستوى الوقوف على مستويات التلاميذ (اللاعبين) قبل أن يلتحقوا بالبرنامج التعليمى (التدريبى) . فمن المعروف أنه عندما يبدأ التلاميذ (اللاعبين) فى تلقى مهارة جديدة، فإنهم يبدءون بمستويات متباينة فى درجة الكفاءة والخبرة، وعليه فإنه يمكن الاستفادة من بعض وسائل القياس - عند بداية التعليم أو

التدريب - فى التعرف على مستويات التلاميذ (اللاعبين) فى المهارة الجديدة، وكذا تحديد المتميزين منهم وغير المتميزين فى أداء هذه المهارة.

وتستخدم الاختبارات والمقاييس لتعيين مستوى التلميذ (اللاعب) بتطبيقها قبل البدء فى التعليم (التدريب) Pretest وأثناء الاستمرار فيه، حيث يستفاد من هذا الإجراء فى الآتى:

١ - الوقوف على مستوى الاستعدادات والقدرات اللازمة كمتطلب أساسى Prerequisite للنجاح فى المهارة الجديدة والاستمرار فيها.

٢ - تقويم مستوى تحصيل التلميذ (اللاعب) فى المهارة الجديدة.

٣ - وضع التلميذ (اللاعب) فى الصف أو المستوى الدراسى (التدريبى) الذى يتناسب مع مستواه فى الموضوع الجديد للتعلم (المهارة الجديدة).

٤ - الوقوف على الميول والاتجاهات الشخصية للتلميذ (اللاعب) نحو موضوع التعلم (المهارة الجديدة).

٥ - تحديد نقطة البداية المناسبة للبرنامج التعليمى (التدريبى) فى ضوء المعلومات المتوفرة عن المتطلبات القبلية اللازمة لموضوع التعلم (المهارة الجديدة).

٦ - اختيار طريقة التدريس (التدريب) المناسبة.

٧ - تقويم الحالة الصحية للتلميذ (اللاعب) فى ضوء نتائج الفحوص الطبية التى تتم قبل الالتحاق بالبرنامج، والتى تستهدف الكشف عن الأمراض والإعاقات الظاهر منها وغير الظاهر، واتخاذ قرارات التصفية المناسبة التى تتعلق بالإعفاءات من بعض أوجه النشاط الرياضى وبخاصة فى المدارس.

ويستخدم لتعيين المستوى فى المجال الرياضى العديد من وسائل القياس، من هذه الوسائل ما يلى:

١ - أساليب الملاحظة.

٢ - سجلات المدرسة (النادى).

٣ - اختبارات الاستعدادات والميول والاتجاهات وبعض اختبارات سمات الشخصية، والتى تدور بشكل رئيسى حول الاستعداد المختص specialized aptitude بالتحصيل فى مجال واحد محدد «حيث يطلق على مثل هذا النوع من الاختبارات اسم: اختبارات التهيؤ Readiness ».

تحسين التعليم (التدريب) Improving instruction :

عندما تكشف نتائج القياس عن وجود فروق differences كبيرة (حادة) فيما بين التلاميذ (اللاعبين) فى درجات تحصيل المهارة التى يتدربون عليها، وذلك على الرغم من أن المعلم (المدرب) يقوم بتدريس هذه المهارة لجميع التلاميذ (اللاعبين) بطريقة موحدة، فإن ذلك يدل على عدم ملائمة طريقة التدريس (التدريب) بالنسبة لجميع التلاميذ (اللاعبين). فعندما تظهر نتائج القياسات الدورية أن الفروق الفردية فيما بين المتدربين على مهارة حركية تزداد اتساعاً، فإن ذلك يعنى عدم مواءمة برنامج التعليم (التدريب) adapation of instruction المقدم، حينئذ يصبح من الضرورى على المعلم (المدرب) اتخاذ الإجراءات المناسبة لتحسين عملية التعليم (التدريب)، وذلك بأن يقوم بتعديل طريقة التدريس (التدريب) وإعادة تقويم الأدوات والأجهزة والوسائل المساعدة، ومحاولة التعرف على العوامل التى تؤثر فى عمليتى التعليم والتعلم Learning and instruction والتغلب عليها، بما يحقق مواءمة البرنامج من أجل تحسين التعليم (التدريب).

تقدير الدرجات Grading:

تستخدم نتائج القياس فى التربية الرياضية المدرسية لإعطاء الدرجات grades للتلاميذ حيث تعد هذه المهمة من أهم وأصعب مسئوليات معلمى التربية الرياضية. وعموماً فإنه يمكن الاستفادة من تقدير الدرجات فى التربية الرياضية المدرسية فى الآتى:

١ - تنشيط التلاميذ وزيادة حماسهم.

٢ - تزويد أولياء الأمور والقائمين على العملية التعليمية بالمعلومات اللازمة عن

حالات التلاميذ.

٣ - تقويم تحصيل التلميذ وقدرته على التعلم.

٤ - الكشف عن مدى تقدم التلميذ فى بعض المجالات الخاصة.

ومن الملاحظ أن وضع الدرجات فى التربية الرياضية المدرسية يعتبر أحد أهم مظاهر التقويم التجميعى (الختامى) Summative evaluation الذى يستخدم القياس التجميعى Summative testing لتقدير التحصيل فى نهاية المقرر الدراسى أو الوحدة التعليمية.

البحث العلمى Research:

البحث العلمى عبارة عن استقصاء Investigation يستهدف تنظيم الحقائق والمعلومات التى يتم الحصول عليها وفقاً لخطة هادفة، هذه الحقائق والمعلومات قد تسهم بشكل مباشر أو غير مباشر فى تطوير برامج الخدمات المهنية التى تقدمها بعض الهيئات أو المؤسسات، أو حل المشكلات الحيوية فى بعض المجالات الحيوية.

وقد تدور البحوث العلمية حول التنظير من حيث طبيعة المدركات والمفاهيم التى تتناولها، وقد تهتم بدراسة النظرية والفروض العلمية وعلاقتها ببعض الظواهر الاجتماعية أو النفسية أو غيرها.

ويأخذ البحث العلمى أشكالاً متعددة، فقد يكون تنظير لقضية أو مشكلة من المشكلات كما فى الدراسات الفلسفية، وقد يكون اختباراً للفروض العلمية كما فى الدراسات التجريبية، وقد يكون استرجاعاً للماضى كما فى الدراسات التاريخية، أو تحليلاً للحاضر كما فى الدراسات المسحية الوصفية descriptive وما إلى ذلك من الأشكال.

ولعل أهم ما يميز البحث العلمى، هو اختبار الفروض العلمية، حيث يتطلب ذلك التنقيب المنظم عن المعلومات فى موضوع معين، وذلك بغرض اكتشاف الحقائق العلمية، أو التحقق من صحة بعض القوانين والنظريات فى بعض المجالات.

ويعتبر ميدان التربية البدنية والرياضة من الميادين التى تتضمن العديد من المشكلات التى تتطلب القيام بالبحوث العلمية المتخصصة. وتعد الاختبارات والمقاييس أدوات رئيسية ووسائل مهمة لجمع المادة العلمية عند القيام بمثل هذه البحوث. فالكشف عن حقائق جديدة ، ودراسة تأثير متغير تجريبى على متغير تابع ، وتحقيق التكافؤ بين المجموعات ، واتخاذ بعض التصميمات التجريبية المناسبة ، كل ذلك يتطلب استخدام العديد من وسائل القياس. وفى هذا الصدد يتفق معظم العلماء على أنه يمكن اعتبار القياس «بمثابة القلب بالنسبة لأى من البحوث النفسية أو التربوية Measurement may be regared as the heart of any psychological or Educational research».

وتكمن أهمية الاختبارات والمقاييس بالنسبة للبحوث العلمية ليس فى كونها الوسيلة الأساسية لجمع البيانات Data فحسب ، وإنما فى مدى سلامة ودقة البيانات التى يتم تجميعها عن الظواهر المقيسة والتى تعتمد بدورها على مدى صلاحية الاختبارات والمقاييس المستخدمة ، وأيضاً على مدى الدقة فى عمليات القياس.

ولا يتوقف استخدام الاختبارات والمقاييس فى المجال الرياضى على البحث العلمى فقط ، وإنما يمتد ليشمل العديد من البحوث التى تستهدف إعادة تقنين أو بناء أدوات قياس تصلح لتغطية النقص فى بعض المجالات التى تفتقر إلى مثل هذه الأدوات.

٤ - تصنيف الاختبارات والمقاييس فى المجال الرياضى

التعريف ببعض المصطلحات :

قبل أن نستعرض التصنيفات المختلفة للاختبارات والمقاييس فى المجال الرياضى ، فإنه يلزم توضيح المقصود بالمصطلحات التالية :

مفهوم الاستجابة Response :

يقصد بالاستجابة رد فعل الإنسان لمثير ما ، أى أن الاستجابة عبارة عن عملية عضلية أو غددية تشير إلى سلوك أو فعل صادر عن الإنسان كرد فعل لمثير أو منبه ما ،

وقد تكون الاستجابة إجابة على سؤال أو استفتاء ما، وقد يكون السلوك صريح معلن، أو خفى مضمر.

مفهوم السلوك Behavior:

يطلق السلوك على الاستجابة التى يقوم بها الإنسان كنتيجة للموقف الذى يواجهه (المثير): أى أن الاستجابة تشير إلى كل ما يصدر عن الإنسان من أنشطة (أنماط سلوك) يمكن ملاحظتها وقياسها . معنى ذلك أن الكلام سلوك، والحركة سلوك، والاستجابات الداخلية (كالتخيل) سلوك.

ويتضمن السلوك كنشاط كلى مركب ثلاثة مجالات رئيسية هى:

- **المجال الإدراكى المعرفى Perceptual and Cognitional:**

وهو يختص بإدراك المظاهر والأحداث المختلفة التى تدور حول الفرد، والتى يحدث فيها تفاعل يشتمل على العمليات العقلية كالإدارك والتميز والتفكير والتصور والتخيل والتذكر والتعبير اللغوى، إلى غير ذلك من الجوانب المعرفية والعقلية.

- **المجال الانفعالى الوجدانى Emotional and Sentimental:**

وهو مجال يختص بالحالة الانفعالية المصاحبة للسلوك، كالميل والإقبال على موضوع ما، حيث يمثل هذا الميل نشاط ومحركات للسلوك، كما يتضمن المجال الانفعالى الشعور بالارتياح أو عدم الارتياح تجاه موضوع معين، مما يؤثر فيما يعرف بالتثبيت الانفعالى Affective Fixation أو تدعيم أو كف الاستجابة inhibition الخاصة بهذا السلوك.

- **المجال الحركى الإجرائى Motional and Operational:**

وهو مجال يختص بالاستجابات الحركية التى تتم عندما يواجه الفرد مواقف تتطلب منه الاستجابة عن طريق الفعل أو الحركة كالمشى أو الجرى أو غيرهما، وجميعها استجابات تحدث نتيجة متغيرات أو تنبيهات قد تكون لفظية يعقبها استجابات حركية إجرائية.

ومن ناحية أخرى يمكن أن يأخذ السلوك الصور التالية :

السلوك الظاهر Apparent behavior :

ويتمثل فى السلوك الحركى فى الرياضة - وفى غيره من مظاهر السلوك كالأكل والشرب وقيادة السيارة وغيرها .

السلوك الباطن Unapparent behavior:

ويتمثل فى التفكير والتخيل والتذكر، وجميعها أنماط من السلوك لا يمكن ملاحظتها، ويمكن أن يكون السلوك الباطن عقلياً كالتفكير أو انفعالياً كالميول والاتجاهات والانفعالات والعواطف وغيرها .

السلوك المكتسب Acquired behavior :

وهو سلوك مكتسب عن طريق التعلم من البيئة الاجتماعية كالرغبة فى التفوق، والرغبة فى الشعور بالأمن والتقبل الاجتماعى وغيرها .

السلوك السوى Normal behavior :

وهو السلوك الذى يتفق مع عادات وتقاليد وأعراف المجتمع، حيث تختلف معايير السواء والإنحراف من مجتمع لآخر ، وإن كانت النظرة إلى السلوك السوى تتأسس على ما يتفق وتقاليد المجتمع أو يتفق مع غالبية سلوك أفراد المجتمع .

السلوك المنحرف (الشاذ) Abnormal behavior :

وهو السلوك الذى يغاير ما اتفقت عليه الجماعة .

مما سبق يتضح أن السلوك الإنسانى يشير إلى الأنشطة العقلية Mental activities والأنشطة النفسية psychological وبنية الشخصية بالإضافة إلى الأنشطة الحركية motor activities حيث تدخل هذه الأنشطة فى تشكيل غالبية مظاهر السلوك وعمليات الاستجابة فى المجال الرياضى، كما يتضح أن بنية السلوك وحدة لا تتجزأ برغم تميز المظاهر السلوكية فى ناحية على أخرى، وأن من أهم ما يميز السلوك أنه عملية دينامية قابلة للتعديل والتغيير والتكييف والتطور نتيجة تفاعل الفرد مع بيئته الخارجية .

- مفهوم الأداء (الإنجاز) Performance :

يستخدم مصطلح الأداء (الإنجاز) للتعبير عن كل المظاهر التى يمكن رؤيتها أو ملاحظتها فى المجال المعرفى Cognitive والمجال الانفعالى Affective ، والمجال النفس - حركى Psychomotor . ويعرّف جرونلند Gronlund ١٩٨١م - على سبيل المثال - الأداء (الإنجاز) بأنه استجابة قابلة للقياس أو الملاحظة فى المجال المعرفى أو المجال الانفعالى أو المجال النفس - حركى.

ويقصر توكمان Tuckman ١٩٧٥م مصطلح الأداء (الإنجاز) على مظاهر السلوك المتعلقة بالمجالين المعرفى والنفس - حركى فقط، ويستخدم مصطلح السلوك Behavior للدلالة على المظاهر المتعلقة بالمجال الانفعالى Affective.

ومن ناحية أخرى يرى فريمان Freeman ١٩٦٥م أن مصطلح الأداء يمكن أن يطلق على أى شىء يفعله الفرد بشكل علنى وصريح، سواء استخدم فى ذلك اللغة أو لم يستخدمها، ويضيف "فريمان" أنه عند استخدام مصطلح الأداء Performance فى مجال القياس النفسى، فإن هذا المصطلح يصبح مقصوراً على المعنى الخاص بالاختبارات العملية فقط.

وإذا كان كل من فريمان ١٩٦٥م وتوكمان ١٩٧٥م يقصران مصطلح الأداء (الإنجاز) على الاستجابات الخاصة بالمجالين المعرفى والنفس - حركى ، فإن جرونلند يعتبر الاستجابات الخاصة بالمجال الانفعالى (القياس النفسى) تندرج تحت مفهوم الأداء (الإنجاز) . وقد رأينا أن نأخذ بوجهة نظر جرونلند فى هذا الخصوص.

والخلاصة أن الأداء (الإنجاز) مصطلح يشير إلى الفعل "act" أو الأسلوب style الذى يتم به تأدية العمل أو الشغل أو المهمة، وأن هذا المصطلح يعنى «الاستجابة response» فى مجالات قياس الذكاء، والجانب المعرفى، والجانب الانفعالى، والجانب النفس - حركى.

الأداء الأقصى Maximum Performance :

وهو نمط من الأداء (الإنجاز) يتطلب من المفحوص الآتى :

- تحقيق أقصى أنتاج Production (درجة أو كمية أو رقم) ممكن خلال الزمن المقرر للأداء .
- أو الوصول إلى أعلى مستوى يقرره الاختبار (خلال الزمن المقرر له) .
- أو الوصـــول للـنهايـــة العظمـى للدرجـــة Maximum degree التـى يحددها الاختبار .

الأداء المميز Typical Performance :

وهو نمط من الأداء (الإنجاز) يتطلب من المفحوص الآتى:

- الاستجابة المعتادة (العادية) الى تعبـر عما يعتقده المفحوص أنه أفضل حل أو أفضل رأى ، أو الحل الذى يود أن يقوم به .
- الاستجابـة التى تقيـس ما يقـوم به المفحوص بالفعل ، وليس ما يستطيع أن يقوم به .

أسس تصنيف الاختبارات والمقاييس:

● **التصنيف وفقاً لطبيعة الأداء (الإنجاز) According To Performance :**

وفقاً لطبيعة الآداء فإنه يمكن تصنيف الاختبارات والمقاييس فى مجال التربية البدنية والرياضة إلى الأنماط التالية :

أولاً : مقاييس لا تتطلب الأداء :

وتشمل مجموعة المقاييس التالية :

١ - المقاييس الأنثروبومترية Anthropometric measurements .

٢- مقـاييس النمـو والحـالـة الغـذائيـة Growth and nutritional status measurements .

٣ - مقاييس أنماط الأجسام Body Type Measurements .

وتمتاز هذه المقاييس بالآتى :

- أنها لا تتطلب الاستجابة من المفحوص، بمعنى أنها لا تتطلب الأداء.
- أنها غير مقيدة بزمن محدد تقرره عملية القياس.
- تستهدف التعرف على الحالة الراهنة للصفات والخصائص المقيسة بالنسبة للمفحوص.
- أن معظم هذه المقاييس يستخدم أدوات وأجهزة خاصة، وقليل منها يستخدم الورقة والقلم.

ثانياً: مقاييس تتطلب الأداء الأقصى:

وتشمل مجموعة المقاييس التالية :

١ - اختبارات الذكاء Intelligence tests (القدرة العقلية العامة).

٢ - اختبارات القدرات العقلية المتخصصة Abilities tests

٣ - اختبارات التحصيل Achievement tests (المدرسى والمهنى).

٤ - اختبارات الاستعدادات Aptitudes tests (البدنية والعقلية).

٥ - اختبارات القدرات البدنية والحركية Physical and Motor abilities.

٦ - اختبارات اللياقة البدنية Physical Fitness tests.

٧ - اختبارات اللياقة الحركية Motor Fitness tests.

٨ - اختبارات القدرة الحركية Motor ability tests.

وتمتاز هذه المقاييس (الاختبارات) بالآتى:

- أنها فى - معظم الأحيان - مقيدة بزمن محدد يستغرقه الاختبار، أو العمل ضد زمن مقرر سلفاً للأداء على الاختبار.

- أنها تدور حول مجالين رئيسيين هما: المجال المعرفى، والمجال النفس حركى.
- أن بعض هذه المقاييس يستخدم أدوات وأجهزة خاصة، والبعض الآخر منها يستخدم الورقة والقلم.
- أنها تقيس ما يستطيع الفرد أن يفعله What individual can do?
- أنها تهتم بمقدار الأداء (الاستجابة)، وليس بكيفية الأداء (الاستجابة).
- أنها تتضمن استجابات صحيحة وأخرى خاطئة، أو استجابات قوية وأخرى ضعيفة.

ثالثاً: مقاييس تتطلب الأداء المميز:

١ - الملاحظة Observation

٢ - المقابلات الشخصية Interviews

٣ - الاستبيانات Questionnaires

٤ - سجلات السير الذاتية Anecdotal records

٥ - أساليب ووسائل التقرير الذاتى Self - report

٦ - مقاييس (قوائم) الشخصية Personality measurements

٧ - مقاييس الميول Interests measurements

٨ - مقاييس الاتجاهات والقيم Attitudes and values measurements

٩ - أساليب قياس العديد من الموضوعات النفسية فى مجال علم نفس الرياضة والتمرين البدنى Sport and exercise psychology.

وتمتاز هذه الأساليب والطرق (المقاييس) بالآتى:

- أنها أساليب تستهدف إظهار سمات الشخصية كالسمات المزاجية والميول والاتجاهات والقيم وغيرها من الموضوعات النفسية.

- أنها تتطلب من المفحوص نوع من التقرير الذاتى إزاء موقف معين.
- أنها لا تتضمن استجابات صحيحة وأخرى خاطئة.
- تكون كل استجابة صحيحة طالما أنها تعبر عن السلوك الانفعالى الحقيقى للمفحوص.
- أنها ليست اختبارات قوة Power tests أو اختبارات سرعة Speed tests لكونها غير مقيدة بزمن محدد سلفاً للأداء.
- أنها تستهدف إظهار استجابة المفحوص (السلوك الانفعالى) الذى يعبر عن أدائه نحو موضوع نفسى محدد.
- أن طبيعة الاستجابة التى يظهرها المفحوصين تتوقف على طبيعة الخصائص المميزة للمظاهر السلوكية المزاجية للشخص المفحوص.
- أن معظم هذه المقاييس يستخدم الورقة والقلم.
- تعرف مقاييس الأداء المميز باسم: المقاييس الموقفية Situational التى تستهدف معرفة ما يفعله المفحوص أو يفكر فيه عندما يواجه موقفاً من المواقف التى يمكن أن يقابلها فى حياته اليومية.

● **التصنيف وفقاً لأسلوب التطبيق** According to mode of adminstration:

تصنيف الاختبارات والمقاييس فى مجال التربية البدنية والرياضة وفقاً لأسلوب وطريقة التطبيق إلى الآتى:

أولاً: اختبارات السرعة Speed Tests:

وهى اختبارات تتطلب من الفحوص أن يجيب على مجموعة متتالية من الأسئلة، أو يؤدى مجموعة متتالية من الأعمال أو الواجبات المتماثلة فى درجة صعوبتها فى وقت محدد من الزمن. وتصمم اختبارات السرعة على أساس أن مستوى الأسئلة (الفقرات) أو الأعمال تكون متساوية فى درجة صعوبتها. وهى

تتطلب من المفحوص أن يجيب على كل الأسئلة إجابات صحيحة وبسرعة، وهذا معناه أنه فى حالة تطبيق اختبارات السرعة فإنه يطلب من المفحوص أن يجيب أو يحل كل الأسئلة أو يقوم بكل الأعمال والواجبات فى حدود الزمن المقرر للاختبار، حيث يكون الزمن محدداً لدرجة صعوبة الاختبار.

وفى هذا النمط من الاختبارات تكون درجة المفحوص هى عدد الأسئلة أو الأعمال التى يجيب عنها أو ينجزها بطريقة صحيحة خلال الوقت المخصص للاختبار.

وتعتبر اختبارات الاستعدادات Aptitude tests من أفضل الأمثلة لهذا النمط (النوع) من الاختبارات.

ملاحظات مهمة:

١ - هذا النمط من الاختبارات غير شائع الاستخدام فى مجال القياس فى التربية البدنية والرياضة.

٢ - تتمثل الصعوبات الرئيسية فى هذا النمط من الاختبارات فى كيفية تحديد وضبط الآتى: الزمن المخصص للاختبار، مستوى الصعوبة، تماثل درجة الصعوبة فيما بين الأسئلة أو الأعمال المختلفة التى يتضمنها الاختبار.

ثانياً: اختبارات القوة Power Tests:

وهى اختبارات تتضمن أسئلة أو أعمال أكثر صعوبة من نظيراتها فى اختبارات السرعة، حيث يلاحظ أن قوة الأداء تكون عكس سرعة الأداء. فإذا كانت اختبارات السرعة تتقيد بوقت محدد للإجابة لكونها اختبارات موقوتة Tiemed tests، فإن اختبارات القوة تعطى للمفحوص وقتاً كافياً للإجابة على كل وحدات الاختبار.

ومن خصائص اختبارات القوة الآتى:

- تعطى متسعاً من الوقت يسمح للمفحوص بالإجابة على كل الأسئلة:

- تتضمن أسئلة أو أعمال أو واجبات تكون متدرجة فى صعوبتها .

- تكون درجة المفحوص هى عدد الإجابات الصحيحة التى أنجزها .

- من أمثلة هذا النوع من الاختبارات : اختبارات الذكاء واختبارات التحصيل .

ثالثاً : الاختبارات الفردية Individual Tests :

هناك بعض الاختبارات تطبق على مفحوص واحد فى الوقت نفسه ، بمعنى أنه لا يمكن تطبيق الاختبار إلا على مفحوص واحد فقط فى ذات الوقت . ولعل من أهم الأسباب التى تستدعى التطبيق الفردى لبعض الاختبارات والمقاييس هو صعوبة توفير الأجهزة والأدوات والمعامل التى تكفى للتطبيق الجمعى لهذه الاختبارات والمقاييس ، هذا بالإضافة إلى أن طبيعة بعض المقاييس والاختبارات يستلزم تطبيقها تطبيقاً فردياً ، وتندرج معظم الاختبارات فى المجال الرياضى تحت هذا النوع من الاختبارات ، ومن أمثلة ذلك ما يأتى :

١ - اختبارات القدرات البدنية والحركية فى المجال الرياضى .

٢ - اختبارات اللياقة البدنية ، واللياقة الحركية ، والقدرة الحركية .

٣ - اختبارات المهارات الحركية فى الأنشطة الرياضية .

٤ - بعض اختبارات الذكاء (المعملية) .

٥ - الملاحظة والمقابلة الشخصية ودراسة الحالة .

رابعاً : الاختبارات الجمعية Group Tests :

وهى اختبارات تعطى لفرد واحد أو أكثر فى نفس الوقت . ويغلب استخدام هذا النمط من الاختبارات فى الحالات التى تتطلب من المفحوص استخدام اللغة المقروءة والمكتوبة كما فى اختبارات الورقة والقلم التى تتطلب الاستيعاب اللفظى والاستجابات الرمزية Symbolic Responses ، ومن أمثلة هذا النوع من الاختبارات ما يأتى :

١ - اختبارات الذكاء (التى تستخدم الورقة والقلم).

٢ - اختبارات الاستعدادات العامة.

٣ - اختبارات التحصيل (الدراسى والمهنى).

٤ - قوائم (مقاييس) الشخصية واختبارات المقال.

ملاحظات مهمة:

- يستغرق تطبيق الاختبارات الفردية وقتاً أطول من الاختبارات الجمعية.
- يتطلب تطبيق الاختبارات الفردية من القائمين على التنفيذ بذل مجهود يفوق بكثير المجهود المبذول فى تطبيق الاختبارات الجمعية.
- تتعرض الاختبارات الفردية لبعض مصادر الخطأ المتعلقة بالقائمين على تنفيذ الاختبار والتى تتمثل فى الآتى: التعب الذى يمكن أن يؤثر على متابعة الأداء وتسجيل الدرجات، التباين فى تعليمات وشروط الاختبار فى كل مرة تعطى لكل مفحوص على حدة. بينما يلاحظ أن هذه التعليمات تعطى فى الاختبارات الجمعية مرة واحدة لكل المفحوصين مجتمعين.

● **التصنيف وفقاً لأنماط الفقرات التى يتضمنها الاختبار**

According to types of items included

تصنف الاختبارات والمقاييس وفقاً لطبيعة الفقرات التى يتضمنها الاختبار إلى ثلاث فئات رئيسية هى:

أولاً: الاختبارات الشفهية Oral tests:

وتستخدم لقياس التحصيل فى مجالات التعليم، وكوسائل لاختبار المعارف والفهم للطلاب فى مراحل الدراسات العليا حيث يتم ربطها بالاختبارات الكتابية Written Tests فى معظم الأحيان.

وتمتاز الاختبارات الشفهية بأنها تستخدم كوسائل للتشخيص Diagnosis، فالأسئلة الدقيقة المعدة بعناية، يمكن أن تظهر جوانب القوة وجوانب الضعف فى موضوعات دراسية أو علمية محددة، كما قد تفيد مثل هذه الأسئلة فى التعرف على الأخطاء المتكررة التى يقع فيها المفحوص بالنسبة لبعض الموضوعات، والعوامل المسببة لها.

كما أن الاختبارات الشفهية تمثل موقف عملياً يستطيع المفحوص من خلاله أن يوظف معلوماته فى مواقف أخرى لم يتعود عليها، ويحتاج أن يعبر عنها عن طريق الحوار المباشر.

ويمكن تحديد خصائص الاختبارات الشفهية فى الآتى:

١ - أنها تفتقد فى كثير من الأحيان إلى التخطيط المتكامل، بمعنى أن طبيعة الأسئلة التى توجه إلى المفحوص قد تتحدد بناء على طبيعة استجاباته عن أول سؤال يوجه إليه، ومن ثم يمكن أن تتجه الأسئلة الموجهة إلى المفحوص اتجاهاً عفوياً.

٢ - قد تتأثر نتائجها بالتحيز.

٣ - قد تفتقد إلى الشمول، بمعنى أنها تتناول - فى كثير من الأحيان - عينات من الموضوعات والمعارف والمعلومات.

٤ - يتأثر المفحوص - كما يتأثر ممثلو المسرح - بجمهور الحضور من الزملاء ولجنة الامتحان، مما قد يؤثر على قدرته على الكلام والتعبير عن نفسه نتيجة ما يعتريه من توتر وقلق.

٥- عندما يستخدم فى الاختبارات الشفهية لغة أجنبية، فأنه يلاحظ أن إجابة المفحوص قد تتأثر بالقدرة على الاستماع، وبدرجة إجادته لهذه اللغة مما قد يؤثر على مستواه فى الموضوع الذى يتم اختباره فيه.

٦- تستغرق وقتاً طويلاً بالنسبة للقائمين على تطبيقها وبالنسبة للمفحوصين لأنها تطبق فى معظم الحالات تطبيقاً فردياً.

٧- أنها تفتقد لأية سجلات كتابية تبين إجابات المفحوص عن الأسئلة التى وجهت إليه ، بمعنى توثيق إجابات المفحوص.

ثانياً: اختبارات المقال Essay tests :

وتعرف باسم الاختبارات حرة الإجابة Free - answer tests ، وهى اختبارات تتكون فى العادة من فقرات أو أسئلة تستهدف التعرف على قدرة المفحوص على التنظيم الشامل لمعلومات وأفكار تتناول موضوعات متعددة، وتتطلب استرجاع المعلومات واختيار المهم منها.

وتمتاز هذه الاختبارات بأنها تعطى للمفحوص الحرية فى الكتابة وتنظيم الإجابة، ومن ثم فهى اختبارات تستخدم الورقة والقلم Paper and pencil ، بدلاً من الاختبارات الشفهية ، وذلك بهدف التغلب على المشكلات الناشئة عن الإجابة أمام لجنة الاختبار (الامتحان) وجهاً لوجه ، فبدلاً من أن نسأل المفحوص ليجيب عن السؤال شفهياً Oral ، فإننا نسأله ليجيب عليه كتابة Written .

مميزات اختبارات المقال :

١ - تتضمن أسئلة موحدة لكل المفحوصين.

٢ - يمكن تطبيقها على عدد كبير من المفحوصين فى وقت واحد، لذا فهى لا تستغرق وقتاً طويلاً بالمقارنة بالاختبارات الشفهية.

٣ - تتطلب عناية وجهد كبيرين فى إعدادها.

٤ - تعطى فرصاً جيدة للمفحوصين لكى ينظموا أفكارهم قبل البدء فى الاجابة، كما أنها تمكنهم من تنسيق وتنظيم إجاباتهم.

ثالثاً: الاختبارات الموضوعية Objective tests:

وهى اختبارات تحتوى على وحدات أو أسئلة تتضمن المشكلة بالإضافة إلى الإجابة التى تؤول إليها، فالمشكلة تكون معروفة من خلال الفقرة، إلا أن اختيار الإجابة الصحيحة يتوقف على مجموعة من البدائل تتطلب اتخاذ القرار الصحيح من المفحوص.

وتمتاز الاختبارات الموضوعية بأنه لا مجال فيها لتدخل ميول المفحوص أو ذاته فى الإجابة، وبأنها أيضاً تتجنب عدم موضوعية مصحح الاختبار (مقدر الدرجات)، وتعتبر طرق الإجابة التى تتطلب الصواب والخطأ True - Fals ، أو المزاوجة Matching ، أو الاختيار من متعدد Multiple - Choice من طرق تقدير الدرجات التى تخلو من الأحكام الذاتية Subjective Judgments لمقدرى الدرجات.

ويمكن اعتبار التكميل Completion أحد طرق الإجابة فى الاختبارات الموضوعية، ولكنه نظراً لتأثره بعض الشىء بالأحكام الذاتية عند تقدير الدرجات، لذا يمكن اعتباره أسلوب شبه موضوعى Semiobjective لكونه لا يخلو تماماً من الأحكام الذاتية لمقدرى الدرجات.

- **التصنيف وفقاً لطرق الاستجابة** According to methods of response :

تصنف الاختبارات والمقاييس فى مجال التربية البدنية والرياضة وفقاً لطرق الاستجابة (الإجابة) إلى ثلاث فئات عريضة هى:

- اختبارات تتطلب الاستجابة اللفظية.
- اختبارات تتطلب الاستجابة غير اللفظية.
- اختبارات تتطلب الاستجابة الحركية.

أولاً: الاختبارات التى تتطلب الاستجابة اللفظية Verbal response tests:

وهى اختبارات تتطلب من المفحوص أن يعبر عن استجاباته عن طريق اللغة المنطوقة Spoken Language (الشفهية) أو المكتوبة Written ، ويتحدد نوع الاستجابة فى هذا النمط من الاختبارات فى ضوء التعليمات الى يتضمنها الاختبار. وتستخدم الاستجابات اللفظية (سواء المنطوقة أو المكتوبة) فى الاختبارات الشفهية وفى جميع الاختبارات التى تستخدم الورقة والقلم سواء فى مجال قياس الذكاء أو القدرات أو سمات الشخصية أو فى غيرها من المجالات.

ثانياً: الاختبارات التى تتطلب الاستجابة غير اللفظية

Nonverbal response tests

وتعرف هذه الاختبارات باسم: الاختبارات غير اللفظية، وذلك لكونها لا تستخدم اللغة المقروءة أو المكتوبة بالنسبة لاستجابات المفحوصين، فقد يواجه العاملون فى مجال القياس التربوى والنفسى فئات من الأفراد لا يمكن إعطاؤهم الاختبارات اللفظية لأسباب تتعلق بالتعليم كما فى حالة الأميين، أو بالسن كما فى حالة الأطفال الذين لا يستطيعون القراءة والكتابة، أو لأسباب تتعلق ببعض الإعاقات التى قد يكون لها تأثير على الفهم والإدراك اللغوى كما فى حالات التخلف العقلى للصغار والكبار وحالات الصمم وغيرها. ففى مثل هذه الحالات تطبق اختبارات لا تستخدم أى لغة كتابة أو كلاماً، سواء أكان ذلك بالنسبة للاستجابة أو بالنسبة للتعليمات.

ويستخدم هذا النمط من الاختبارات الصور، والرموز، والرسوم، حيث يسجل للمفحوص استجابات بطريقة موضوعية عن طريق التعامل مع تلك الصور أو الرموز أو المكعبات أو ما إلى ذلك من وسائل. ومن الملاحظ بالنسبة لهذا النمط من الاختبارات أنه قد تعطى تعليمات الاختبار إما بالإشارة والحركات وعرض الصور أو الرسوم، وإما شفهياً بالكلام، حيث يترك للمفحوص أن يعبر عن استجابته بطريقة غير لغوية.

ويكثر استخدام هذا النمط من الاختبارات فى مجال الدراسات التربوية والنفسية والاجتماعية فى حالات قياس الذكاء وبعض القدرات العقلية المتخصصة، وأيضاً فى حالة قياس الاستعدادات العامة للأسوياء والمعاقين.

ثالثاً: الاختبارات التى تتطلب الاستجابات الحركية

Motor response tests

وهى نمط من الاختبارات يتطلب الاستجابة الحركية Motor response وليس الاستجابة اللفظية Verbal response، ويطلق على هذا النمط من الاختبارات فى

كثير من الأحيان اسم: الاختبارات الأدائية (العملية) Performance tests لكونها تتطلب العمل البدنى Physical work فى معالجة الأشياء أو القيام بأعمال كبيرة تمثل بعض أوجه النشاط الحركى فى الرياضة البدنية باستخدام بعض الأجهزة والأدوات الخاصة.

وبناء على ذلك يظهر أن الاختبار العملى (الأدائى) يكون عكس الاختبار اللغوى، ففى الاختبار اللغوى تكون الاستجابة ملفوظة أو مكتوبة، بينما يطلب من المفحوص فى الاختبارات العملية (الأدائية) أن يؤدى عملاً، وليس أن يقول شيئاً، حيث يقل استخدام (اللغة)، وقد يمنع استخدامها تماماً.

ولأن الاختبار العملى (الأدائى) يتطلب الاستجابة عن طريق الحركة Movement، لهذا يذكر مصطلح «الاختبار العملى» على أنه مرادف لمصطلح «اختبار غير لفظى». وهناك وجهة نظر تصنف كل الاختبارات التى تتطلب استخدام أدوات أو أجهزة كاختبارات عملية. وعليه فإنه يمكن اعتبار معظم اختبارات الأداء الحركى فى الرياضة اختبارات عملية، لكون معظم اختبارات القدرات البدنية والمهارات الحركية فى الأنشطة الرياضية تستخدم الأدوات.

- **التصنيف وفقاً لأسس بناء الاختبار**

According to principles of test construction

تصنف الاختبارات وفقاً لأسس بناء (تركيب) الاختبار إلى نمطين رئيسين هما:

- الاختبارات المقننة.
- الاختبارات التى يعدها المعلم (المدرب الرياضى).

أولاً: الاختبارات المقننة Standardized tests:

يقصد بالاختبارات المقننة تلك الاختبارات التى يتم إعدادها بمعرفة باحث أو فريق من الباحثين المتخصصين فى أى من مجالات القياس التربوى أو النفسى أو غيرهما. وتتميز الاختبارات المقننة بأنها مصممة بعناية فائقة، وأنه قد تم تجريبها مرات كثيرة للتحقق من مدى صلاحيتها.

ولعل من أهم الشروط الواجب توافرها فى الاختبارات المقننة ما يأتى:

١ - أن يكون للاختبار تعليمات تعطى لجميع المفحوصين بطريقة موحدة، وعلى أن تشمل هذه التعليمات: الهدف من الاختبار، الشروط الخاصة بالأداء، الأخطاء الشائعة، الزمن المخصص للأداء، كيفية الاجابة (الأداء) على الاختبار.

٢ - أن يكون للاختبار مفتاح تصحيح (اختبارات الورقة والقلم) لتعيين الإجابات الصحيحة على أسئلة أو وحدات للاختبار.

٣ - أن يكون للاختبارات معاملات ثبات وصدق معلنة وصريحة بالنسبة لمجموعات الأفراد الذين أعد لهم الاختبار فى الأصل (مجموعات أو عينات التقنين المرجعية).

٤ - أن يكون للاختبار معايير Norms (جداول مستويات) تظهر بوضوح درجات ومستويات أداء عينة التقنين الأصلية على الاختبار. ومن الملاحظات الجديرة بالاهتمام بالنسبة لهذا الشرط هو أن كرونباخ Cronbach ١٩٧٠م يرى أن شرط المعايير قد يكون غير ملزم بالنسبة للاختبارات المقننة، إلا أن مجموعة كبيرة من علماء القياس النفسى يخالفونه فى هذا الرأى.

مما سبق يتضح أن الاختبارات المقننة تتميز بالآتى:

١ - أنها تتمتع بمعاملات صدق وثبات مقبولة بالنسبة لعينة التقنين التى أعدت لها فى الأصل.

٢ - أن لها معايير تعكس مستويات أداء عينة (مجموعة) التقنين الأصلية.

٣ - أن لها كراسة تعليمات توضح كافة المعلومات اللازمة بالنسبة لتطبيق الاختبار.

٤ - أن لها مفتاح تصحيح يوضح اتجاهات العبارات (الأسئلة) الموجبة والعبارات السالبة لتعيين الدرجات الخاصة بكل إجابة.

٥ - تعرف هذه الاختبارات - فى معظم الأحيان - باسم الاختبارات المنشورة Published tests، وذلك لكونها تحظى بالنشر من قبل الدوريات والمراجع العلمية المتخصصة والتى تتمتع بسمعة محلية وعالمية محترمة.

ومن أمثلة الاختبارات المقننة ما يأتى:

- مقياس ستانفورد - بينية للذكاء Stanford - Binet Scale
- مقاييس وكسلر للذكاء Wechsler Scales
- بطارية اختبار الاستعداد العام General Aptitude Test Battery
- اختبار القدرة الرياضية العامة General Athetic Ability
- اختبار ماك كيردى للوسع البدنى McCurdy physical Capacity Test
- اختبار هارفارد للخطوة Harvard Step Test
- اختبار تشمبرين للكرة الطائرة Chamberlain Volleyball Test

ثانياً: الاختبارات التى يعدها المعلم (المدرب الرياضى)

Teacher - Made Tests

وهى نمط من الاختبارات يقوم المعلم (المدرب الرياضى) بإعدادها كوسائل قياس تقديرية وغير مقننة أو غير رسمية Informal tests ، بمعنى أنها لم تخضع لشروط التقنين Standardization الخاصة بالاختبارات المقننة. ويعرف هذا النمط من الاختبارات باسم : اختبارات غرفة الصف Classroom tests لأن معلمى الصف يقومون بإعدادها لقياس تحصيل التلاميذ فى المستويات الدراسية المختلفة.

وتكمن القيمة الحقيقية فى هذا النمط من الاختبارات فى أنها لا تستغرق وقتاً طويلاً لكونها تطبق تطبيقاً جمعياً، وأيضاً لكونها تستخدم لتقويم التحصيل الدراسى لتلاميذ كل صف دراسى على حدة، كما تمتاز بأنها تعد من قبل المعلم على أساس الخصائص المميزة للتلاميذ فى كل فصل وفى كل صف دراسى كل على حدة، حيث يكون المعلم أكثر من غيره معرفة بتلك الخصائص، لهذا نجده أفضل شخص يمكن أن يقوم بإعداد مثل هذا النمط من الاختبارات لتكون على درجة مناسبة من الصعوبة حتى تصلح لقياس تحصيل التلاميذ فى الموضوعات الدراسية التى قدمها المعلم.

وتعتبر الاختبارات الشفهية، واختبارات الورقة والقلم، والمقابلات الشخصية نماذج لمثل هذا النوع من الاختبارات.

- **التصنيف وفقاً لطرق تفسير النتائج**

According to interpreting the results

تصنف الاختبارات والمقاييس على أساس طرق تفسير الدرجة إلى نمطين رئيسيين من الاختبارات هما :

- اختبارات لها معيار مرجعى Norm - Referenced Tests (NR)
- اختبارات لها محك مرجعى Criterion - Referenced Tests (CR)

وفيما يلى شرح موجز لكل نمط من هذين النمطين :

أولاً : اختبارات لها معيار مرجعى (NR) :

وهى اختبارات تستخدم عند محاولة تفسير أداء كل مفحوص بالمقارنة بأداء غيره من المفحوصين من نفس مجموعته، وهى تعرف باسم الاختبارات معيارية التفسير لكونها تعتمد على مقارنة أداء الفرد بمعيار Norm يتمثل فى أداء المجموعة التى ينتمى إليها أو أى مجموعة أخرى مشابهة لمجموعته.

فالاختبارات التى لها معيار مرجعى (NR) تمكننا من مقارنة أداء أى فرد أو مفحوص بمعيار أداء مجموعة مشابهة. فالدرجات التى يتم الحصول عليها من هذا النمط من الاختبارات يتم تفسيرها وفقاً لمستويات نسبية Relative Standards من الدرجات. ويكون للاختبارات معيارية المرجع (NR) جداول معيارية Norm tables تشتمل على معدلات Averages الأداء لعينة أو مجموعة نمطية typical group أو مجموعة التقنين Standardization group.

ماهية المجموعة المرجعية (مجموعة التقنين أو الإسناد)

المجموعات المرجعية أو مجموعات الإسناد أو مجموعات التقنين هى المجموعات التى أعد لها الاختبار فى الأصل، واستخدمت درجاتها فى الأداء على

الاختبار لإعداد جداول المعايير Tables of Norms ، حيث تمثل هذه الدرجات الأداء الصحيح والجاد لمجموعة التقنين.

ويتم إعداد معايير الاختبارات فى المجال الرياضى عادة على أساس النوع (الجنس)، والعمر الزمنى، والطول، والوزن، أو المستوى الدراسى. ويفضل عند بناء معايير الاختبارات معيارية المرجع (NR) مراعاة الآتى:

١ - أن يكون حجم مجموعة (عينة) التقنين مناسباً من حيث الحجم.

٢ - أن تكون مجموعة التقنين ممثلة للمجتمع الأصلى بكل خصائصه وفئاته تمثيلاً جيداً.

٣ - ألا يكون قد مضى على المعايير فترة زمنية طويلة (أكثر من ١٠ سنوات) لأن مثل هذه المعايير تستخدم كمحكات لتقويم أداء المفحوصين عن طريق مقارنة درجاتهم بهذه المعايير التى أعدت على مجتمع مشابه تماماً لمجتمعهم.

مميزات الاختبارات معيارية المرجع (NR):

١ - يكثر استخدامها فى مجال النشاط الرياضى، لأن معظم الاختبارات المتاحة للقياس فى هذا المجال اختبارات معيارية المرجع.

٢ - صممت الاختبارات معيارية المرجع على أساس الكشف عن الفروق الفردية بين الأفراد، ومن ثم فهى تستخدم فى قياس التحصيل لأغراض التقويم التجميعى، كما يستفاد منها فى أغراض التصنيف والانتقاء.

٣ - يستخدم هذا النمط من الاختبارات للتمييز بين الأفراد والجماعات الرياضية، ولذلك يطلق عليها فى بعض الأحيان اسم: الاختبارات التمييزية Discrimination tests لأن نتائجها تسمح بترتيب الأفراد تنازلياً أو تصاعدياً حسب درجات المفحوصين بالنسبة للقدرة أو السمة المقيسة.

٤ - تقيس الاختبارات معيارية المرجع الحالة الراهنة Status quo للمفحوصين، ومن ثم لا يمكن اعتبار نتائجها تمثل المستوى الأمثل الذى يفترض أن تكون عليه حالة هؤلاء المفحوصين.



قیاس له معیار مرجعی (NR)



قیاس له محك مرجعی (CR)

ثانياً : اختبارات لها محك مرجعى (CR):

تطبق الاختبارات والمقاييس لجمع البيانات عن الظواهر المقيسة، وذلك بغرض التقويم، فعندما يتم الانتهاء من جمع البيانات وإجراء المعالجات الإحصائية المناسبة لها، يقوم الباحث باتخاذ القرارات التقويمية Evaluative Decisions فيما يتعلق بالنتائج التى تم التوصل إليها وذلك بأى من الطريقتين التاليتين :

- المعيار المرجعى Norm - Referenced
- المحك المرجعى Criterion - Referenced

وقد أشرنا إلى أن المعيار المرجعى (NR) عبارة عن محك يستخدم لتفسير الأداء عن طريق مقارنة درجات المفحوص بدرجات غيره من المفحوصين من نفس مستواه أو مجموعته.

وأما فيما يتعلق بالاختبارات التى لها محك مرجعى، فهى نمط من الاختبارات تستخدم المحك المرجعى لإصدار القرارات التقويمية بالنسبة للمفحوصين، وذلك على أساس معرفة ما الذى يستطيع المفحوص أن يفعله؟ وليس على أساس مقارنة أداء هذا المفحوص بأداء غيره من المفحوصين. فالاختبارات التى لها محك مرجعى تستخدم لتقوم أداء الأفراد على أساس معيار ثابت (قياسى) وليس على أساس مقارنة أداء الأفراد بعضهم ببعض أو مقارنة أدائهم بأداء مجموعة مشابهة.

وقد يستخدم البعض مصطلح «اختبارات مطلقة المرجع» كمرادف لمصطلح «اختبارات لها محك مرجعى» حيث يستهدف استخدام الاختبارات مطلقة المرجع (CR) التعرف على الأداء المتوقع من الأفراد بالنسبة لمستوى خاص ومحدد من التحصيل Specific Level of Achievement ، حيث يتوقف اجتياز الاختبار بنجاح على درجة تمكن الفرد من المتطلبات الأساسية للأداء المقبول Acceptable Performance والتى تمثل الحد الأدنى Minimum level من الأداء وليس على أساس مقارنة أداء الفرد بأداء الآخرين.

والاختبار مطلق المرجع (CR) يتضمن استخدام الأغراض السلوكية behavioral objectives التى يتم تحديدها مسبقاً كمستوى متوقع لأداء الأفراد، ومن أمثلة ذلك فى المجال الرياضى ما يأتى:

- لتحقيق النجاح فى اختبارات الجرى كمؤشر للكفاءة الوظيفية للجهازين الدورى والتنفسى لطلاب الجامعات هو أن يتمكن الطالب من الجرى - المشى مسافة ٢ ميل فى ١٤ دقيقة أو أقل.
- أن يجيب الطالب إجابات صحيحة على ٨٠٪ أو أكثر من الأسئلة التى يتضمنها اختبار للمعرفة الرياضية.
- للنجاح فى مقرر دراسى فى فسيولوچية الرياضة يجب أن يجيب الدارس إجابات صحيحة على ٣٠ سؤالاً على الأقل فى اختبار معرفى يتكون من ٥٠ سؤلاً.

مما سبق يلاحظ أن أداء الطالب لم يقارن بأداء غيره من الطلاب، وإنما تمت مقارنته بمستويات Levels محددة سلفاً إذا اجتازها الطالب يتم تقويمه على أنه (ناجح-Pass) ، وإذا فشل فى اجتيازها يقوّم على أنه (راسب - Fail) ، وبناء على ذلك قد يرى بعض العلماء أن القياس مطلق المحك (CR) يعد محدوداً لكونه مقيداً بمستوى (ناجح أو راسب) ، حيث لا يساعد هذا الأسلوب من أساليب التقويم على إظهار مستوى قدرة المفحوص من حيث هى جيدة أو متوسطة - أو ضعيفة.

مميزات الاختبارات مطلقة المرجع (CR) :

١ - هى اختبارات أعدت فى الأصل لكى تستخدم فى اتخاذ قرارات عن مستويات Levels تمكن mastery الطالب أو اللاعب أو الفرد الرياضى بالنسبة لموضوع محدد. لذلك نجدها تعرف باسم اختبارات التمكن Mastery أو اختبارات الكفاءة Proficiency tests .

٢ - نظراً لكون الاختبارات مطلقة المرجع تستخدم لتحديد التمكن أو الكفاءة بالنسبة لموضوع ما، لذا نجدها تتطلب من المفحوص الالتزام الكامل بكل ما تتضمنه

من تفاصيل وإجراءات وتعليمات، كما أن إعدادها فى مجال النشاط الرياضى يحتاج إلى متخصصين على مستوى عال من الكفاءة، كما يتطلب إقرار صلاحيتها ضرورة موافقة المتخصصين من ذوى الخبرة المتميزة، كما تستلزم المراجعة والتنقيح من وقت لآخر.

٣ - للاختبارات مطلقة المرجع أهمية خاصة فى المجال الرياضى لكونها تبين للفرد مدى تمكنه من موضوع ما يهمه، كما أنها تمثل بالنسبة لمعلم التربية الرياضية والمدرب الرياضى وسيلة مهمة لانتقاء اللاعبين على أساس اختيار كل من يحقق المستويات المحددة سلفاً للأداء، أو على أساس الوصول إلى نسبة مئوية محددة من الإنجازات يستطيع الفرد الرياضى أن يحققها بنجاح وفقاً لمستوى محدد من درجة الصعوبة.

٤ - تتطلب الاختبارات مطلقة المرجع وضع الدرجة الفاصلة Cutting Score لتحديد مستوى التمكن أو الكفاءة، والدرجة الفاصلة عبارة عن نقطة أو رتبة يتم على أساسها قبول ما فوقها ورفض ما دونها من نتائج أو أشياء، وهى تستخدم لتقسيم الأفراد (المفحوصين) إلى فئتين مثل ناجح أو راسب، أو مقبول ومرفوض، وتمثل الدرجة الفاصلة الهدف الذى يستوجب الوصول إليه كمستوى يمثل الحد الأدنى من التمكن أو الكفاءة فى الاختبارات مطلقة المرجع (CR).

٥ - تستخدم نتائج الاختبارات مطلقة المرجع فى وضع الدرجات فى التربية الرياضية، فعلى سبيل المثال يمكن إعطاء الطالب الذى يتمكن من جرى مسافة ٢ميل فى ١٣ دقيقة أو أقل التقدير (أ - A) ، وإعطاء الطالب الذى يتمكن من الجرى مسافة ٢ ميل فى ١ ث : ١٣ ق إلى ٣٠ث : ١٣ ق التقدير (ب-B) وإعطاء الطالب الذى يتمكن من جرى مسافة ٢ ميل فى ٣١ ث : ١٣ ق إلى ث ٠٠ : ١٤ ق التقدير (جـ -C) وهكذا بالنسبة لبقية التقديرات.

٦ - تمثل الاختبارات مطلقة المرجع مستويات خاصة من الأداءات يُستهدف تحقيقها (الوصول إليها) حيث تستخدم هذه المستويات كمحكات للتقويم.

٧ - المستويات الخاصة Specific Levels التى تتضمنها الاختبارات مطلقة المرجع تعد مستويات مطلقة Absolute Standards ، هذه المستويات لا تهتم بدرجة الأداء ، وإنما تهتم بالمنافسة ضد المستويات الموضوعة لمحاولة الوصول إليها .

المقارنة بين المعيار المرجعى والمحك المرجعى:

حتى نهاية العقد الثانى من القرن العشرين لم يكن قد ظهر بعد مصطلح المعيار المرجعى (NR) أو مصطلح المحك المرجعى (CR) وذلك على الرغم من تبنى استخدام الاختبارات المقننة Standardized Tests على نطاق واسع خلال فترة الحرب العالمية الأولى، وإن كان من الملاحظ أن معظم الاختبارات كانت تستخدم فى تلك الفترة وعن غير قصد المعيار المرجعى Norm - Referenced .

ولقد بدأت المحاولات الأولى لاستخدام المحك المرجعى - Criterion Referenced على يد كارلتون واشبورت ومعاونوه Carlton Washburne etal عام ١٩٢٢م ، والتى تطورت على يد هنرى مورسون Henry Morrison فى عام ١٩٢٦م وكان ذلك نتيجة الاهتمام بمحاولة تقويم مدى إنجاز الأهداف النهائية (الختامية) Terminal Objectives للخطط والبرامج التعليمية. ولقد تأصل استخدام المحك المرجعى خلال حقبة الخمسينيات من القرن الماضى مع مجىء التعليم المبرمج Programmed instruction ، كما كانت للجهود التى قدمها عالم النفس الشهير برهس سكنر Skinner ١٩٥٤م تأثير واضح على تطور مفهوم واستخدامات المحك المرجعى، وذلك عندما قدم سكنر مصطلح الهدف السلوكى behavioral objective كمصطلح جديد فى المجال التربوى.

وتكمن الفروق الأولية بين المعيار المرجعى والمحك المرجعى فى الآتى:

١ - تستخدم الاختبارات التى لها معيار مرجعى لتحديد مستويات السلوك فى مجالات المعرفة والقدرات، بينما تستخدم الاختبارات التى لها محك مرجعى لتقدير مدى إنجاز الأهداف الخاصة Specific Objectives فى أى مجال من المجالين.

٢ - نظرًا لأن الاختبارات التى لـها محك مرجعى (CR) لا تهتم بالإنجاز النسبى أو التحصيل النسبى Relative performance للأفراد، لذا نجدها لا تهتم بدرجة صعوبة أو سهولة الفـقرة، ولا بقدرة الفقرة على التميز كمـعيار لبناء مثل هذه الاختبارات، وإنما نجدها تـهتم باختيار الفقرات على أساس كـيف تظهر الفقرة بشكل جيد المهمات (الواجبات) المطلوب قياسها، فإذا كانت المهمات المطلوب قياسها سهلة كانت فقـرات (وحدات) الاختبار هى الأخرى سهلة، وإذا كانت المهمات (الواجبات) صعبة كانت فقرات الاختبار صعبة.

٣ - يستخدم كل من المـعيار المرجعى والمحك المرجعى لتوفيـر المعلومات التى يمكن الاستفادة منها بالنسبة لأربعة أنماط من التقويم هى:

- يستخدم كل من التقويم التكوينى (البنائى) Formative والتقويم التشخيصى Diagnostic evaluation أدوات قياس لها محك مرجعى (CR).
- يستخدم التقويم التجميعى (الإجمالى أو الختامى) Summative Evaluation أدوات قياس لها معيار مرجعى (NR).
- يستـخدم التقويم بـغرض تعيين المسـتوى Placement Evaluation كل من المحك المرجـعى (CR) لتـقـويم المهـارات اللازمـة كمـتطـلب أسـاسى Prerequisite Skills للـنجاح فـى دراسة وحـدة أو مقـرر دراسى مـعين، ويستخدم المعـيار المرجعى (NR) لتحديد مـستويات الأداء بالنسبـة لتعيين المستوى المتقدم.

ومن الملاحظات الجديرة بالاهتمام فى هذا الخصوص هو التغير الذى حدث فى العقود الأخيرة من القرن العشرين بالنسبة لتقويم نتائج اختبارات اللياقة البدنية، التى كانت تستخـدم المعيار المرجعى (NR) فى تقويم لـياقة الأفراد على أساس متـغيرين مهمين هما: الجـنس (ذكر - أنثى)، والعمر الزمنى، حيث يـلاحظ أن الكثير من هذه الاختبارات يستخدم - الآن - المحك المرجعى جدول (١ - ٢).

ويعرض الجدول (١ -٢) المحك المرجعى (CR) والمعيار المرجعى (NR) لنتيجة اختبار الجلوس من الرقود Sit - Ups test لتلميذ فى سن ١٠ سنوات، فالتلميذ الذى يؤدى ٣٠ مرة يقابل المحك المرجعى (CR) على اختبار معهد أبحاث اللياقة الهوائية (١٩٨٧م)، ويفشل فى نفس الوقت فى أن يقابل المحك المرجعى (CR) لنفس السن فى اختبار برنامج الأحسن بدنيا (١٩٨٨م) ، ويوضح نفس الجدول (١-٢) أن عدد مرات الأداء ٣٠ تكافؤ المئينى الـ ٣٥ بالنسبة لمعايير الاختبار الأمريكى للصحة والتربية البدنية والترويح والرقص (AAHPERD) (١٩٨٠م) ، فى حين تقابل الدرجة ٣٤ المئينى الـ ٥٠ بالنسبة لمعايير نفس الاختبار الأمريكى AAHPERD (١٩٨٠م).

جدول (١ - ٢) المحك المرجعى (CR) والمعيار المرجعى (NR)
لتلميذ سنه ١٠ سنوات عندما أنجز ٣٠ مرة
فى اختبار الجلوس من الرقود
(١ ق)

بطارية الاختبار	الدرجة المرجعية للاختبار	المئينيات على اختبار AAHPERD للياقة البدنية ***
- اختبار معهد أبحاث اللياقة * (Cooper , 1992)	٣٠ مرة	المئينى الـ ٣٥
- اختبار الأحسن بدنياً ** (AAHPERD , 1988)	٣٤ مرة	المئينى الـ ٥٠

* Cooper Institute For Aerobic Research , 1992.

** AAHPERD , Physical Best , 1988.

*** AAHPERD , Youth Fitness Test, 1980

ويبين الجدول (١ - ٢) أن التلميذ أنجز فى اختبار الجلوس من الرقود (١ق) عدد ٣٠ مرة، وأنه عند تقويم هذه النتيجة وفقاً للمحك المرجعى (CR) فى اختبار معهد أبحاث اللياقة فسوف نجد أن التلميذ قد حقق هذا المستوى، ومع ذلك فقد فشل فى تحقيق المستوى الذى حدده اختبار الأحسن بدنياً والذى يساوى ٣٤ مرة.

وفيما يتعلق بالمستويات النسبية (المعيار المرجعى - NR) فنجد أن درجة التلميذ ٣٠ مرة تكافئ المئينى الـ ٣٥ على جداول المعايير التى أعدت عام ١٩٨٠م للاختبار الأمريكى للياقة البدنية AHHPERD ، فى حين تقابل الدرجة ٣٤ المئينى الـ ٥٠ بالنسبة لمعايير نفس الاختبار . ومن المحكات المرجعية (CR) الآتى:

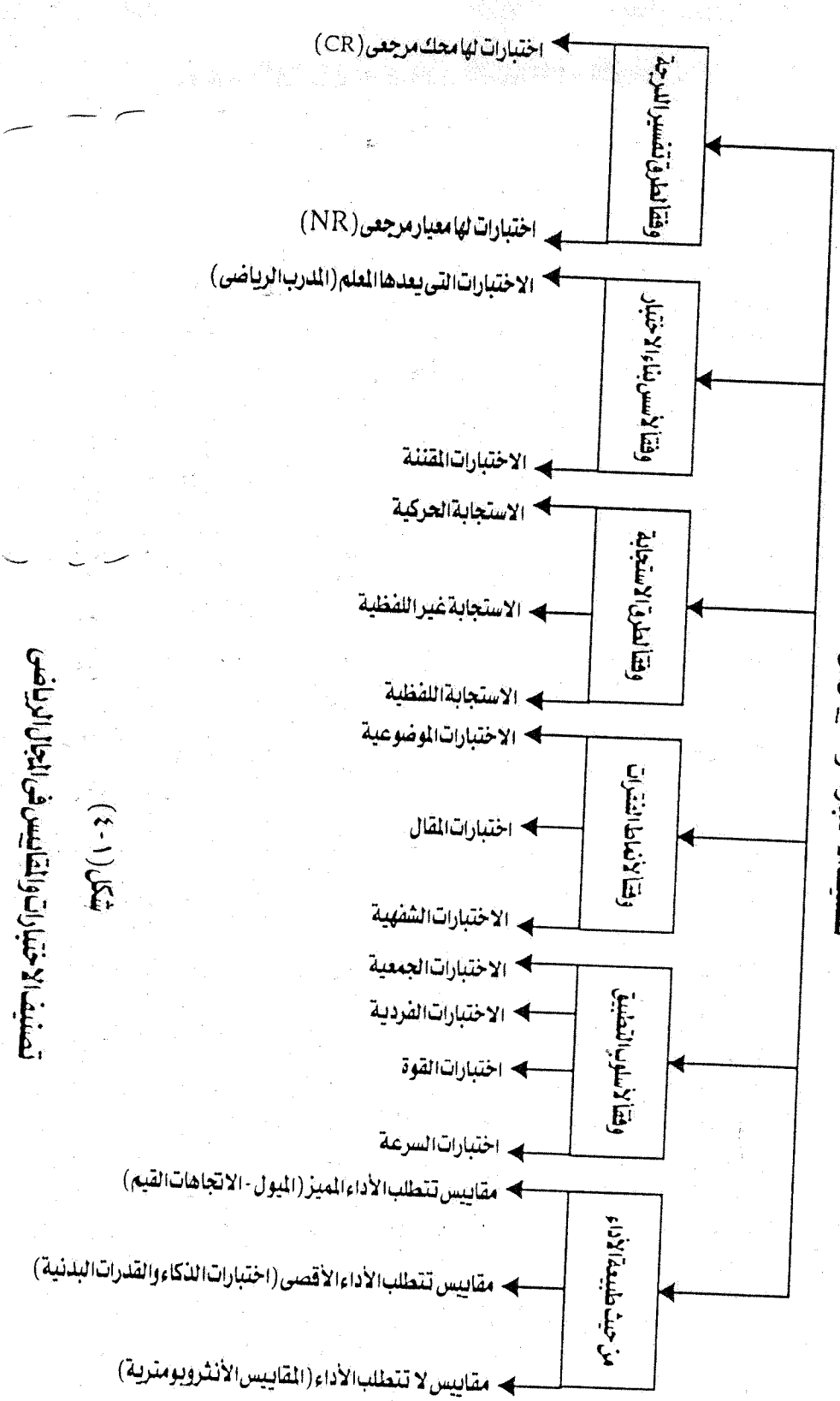
١ - المحك المرجعى (CR) لضغط الدم الطبيعى فى الإنسان هو الذى يساوى $\frac{١٢٠}{٨٠}$ مم/ زئبق.

٢ - المحك المرجعى (CR) لضغط الدم الزائد هو ما يساوى أو يزيد عن $\frac{١٤٠}{٩٠}$ مم/ زئبق.

٣ - المحك المرجعى (CR) للكوليسترول الكلى فى الدم هو ما يقل عن ٢٤٠ مليجرام.

ملحوظة مهمة:

يلخص الشكل (١ - ٤) التصنيفات المختلفة للاختبارات والمقاييس فى المجال الرياضى.



شكل (١-٤)
تصنيف الاختبارات والمقاييس في المجال الرياضي

معجم إنجليزى - عربى للمصطلحات

A

Acceptance	تقبل القيمة
Affective domain	المجال الانفعالى
Application	التطبيق
Analysis	تحليل
Abstractions	تجريدات
Attitudes	اتجاهات
Appreciation	تذوق
Affiliation	الانتماء
Adaptation	المواءمة (التعديل)
Attendance	الحضور

C

Character development	نمو الشخصية
Cognitive processes	العمليات العقلية العليا
Cognitive domain	المجال المعرفى
Comprehension	الفهم
Complex adaptive skill	المهارة التكيفية المعقدة
Complex overt respose	الاستجابة الظاهرة المركبة
Commitment for a value	الالتزام بالقيمة
Characterizing	الوصف
Conceptualization of a value	الإطار المفاهيمى للقيمة
Courtesy	الكياسة (اللباقة)
Communication Skills	مهارات الاتصال
Creative movement	حركة ابتكارية
Compound adaptive Skill	المهارة التكيفية المركبة

E

Edumetrics	القياس التربوى
Evaluation	تقويم
Emoting	انفعال
Emotional biases	نزعات انفالية
Expressive movements	الحركات التعبيرية

F

Fair Play	اللعب النظيف
Fundamental game Skills	المهارات الرئيسة فى اللعبة (المسابقة)

G

Games Skills	مهارات اللعب
Guided Response	استجابة موجهة
Generic Movement	حركة عامة

H

Human Performance	الأداء البشرى

I

Improvising	الإبداع المرتجل
Intellectual abilities	القدرات العقلية
Instruction	تعليم
Interests	ميول
Interpretive movements:	حركات تفسيرية

K

Knowledge	تذكر المعلومات

L

Locomotor movements	الحركات الانتقالية

M

Manipulative Skills	مهارات المعالجة اليدوية
Measuring instruments	أدوات القياس
Measurement device	أداة القياس
Mental Performance	الانجاز (الأداء) العقلى
Mind	الذهن (العقل)
Movement Concepts	مفاهيم الحركة
Mechanism	آلية الفعل (الأداء)

N

Numerical Values	قيم عددية
Nonlocomotor movements	حركات غير انتقالية

O

Organization of a value	تنظيم النسق القيمى
Organizational principles	الأسس التنظيمية
Occupational achievement	التحصيل المهنى
Organizing	السلوك التنظيمى
Ordinative	حركات مرتبة

P

Psychomotor domain	المجال النفس - حركى
Principles	مبادئ
Preference for a value	تفضيل القيمة
Physical responses	استجابات بدنية
Perception	إدراك
Patterning	التنميط
Perceptual abilities	قدرات إدراكية

Q

Qualitlies of movement	خصائص الحركة

R

Retention of Information	استرجاع (تذكر) المعلومات
Responding	السلوك الاستجابى
Reflex movements	الحركات المنعكسة
Refining	يصقل ينقى

S

Set or readiness	تهيؤ (استعداد)
Skilled movements	الحركات الماهرة
Synthesis	التركيب
Structural form	الشكل البنائى
Satisfaction in response	الرضا فى الاستجابة
Selected attention	الانتباه الانتقائى
Self - reporting	التقرير الذاتى
Self - Control	التحكم الذاتى (السيطرة على الذات)
Self - expression	التعبير عن الذات
Self - discipline	الانضباط الذاتى
Self - Confidence	الثقة بالنفس
Spatial awareness	الوعى المكانى

U

Understanding	فهم
Unique	متفرد

V

Valuing	السلوك القيمى
Value Judgments	احكام قيمية
Varying	التنويع

W

Willingness to receive	الرغبة فى الاستقبال

مراجع الفصل الأول

١- أمين أنور الخولى، وجمال الدين الشافعى (٢٠٠٠م): مناهج التربية البدنية المعاصرة. دار الفكر العربى، القاهرة.

٢- محمد حسن علاوى، ومحمد نصر الدين رضوان (٢٠٠٠م): القياس فى التربية الرياضية وعلم النفس الرياضى، دار الفكر العربى، القاهرة.

٣- محمد نصر الدين رضوان، وكمال عبد الحميد إسماعيل (١٩٩٤م): مقدمة التقويم فى التربية الرياضية. دار الفكر العربى، القاهرة.

٤- على ماهر خطاب (٢٠٠١م): القياس والتقويم فى العلوم النفسية والتربوية والاجتماعية. الأنجلو المصرية، القاهرة.

5- Lumpkin, A. (1994): Physical education and Sport - A Contemporary introduction. Mosby, New York.

6- Marsh, J. J. (1984): Measuring affective objectives in Physical education. Physical educator, 41 (2) : 77-81.

7- Merrill, M. D. (1972): Taxonomies, classifications and theory. in R. N. Singer (ed.), the psychomotor domain: movement behavior, philadelphia: Lea and Febiger.

8- Nixon, J., and Jewett, A. (1980): An introduction to physical education. W. B Saunders, Philadelphia.

9- Page, G. T., and Thomas, J. B. (1979): International Dictionary of Education. Kogan Page, London.

10- Sax, G. (1974): Principles of educational measurement and evaluation. Belment, Calif: wadsworth publishing Co.,In C..

11- Stallings, L. M. (1982): Motor Learning - from theory to practice, St. Louis: The C. V. Mosby Company.

الفصل الثانى

الثبات

١ - مقدمة

٢ - نظرية الثبات

٣ - مفهوم الثبات النسبى وطرق حسابه

- طريقة إعادة الاختبار
- طريقة الصور المتكافئة
- طريقة التجزئة النصفية
- طريقة الاتساق الداخلى
- طريقة تحليل التباين

٤ - مفهوم الثبات المطلق وطريق حسابه

٥ - العوامل التى تؤثر على ثبات درجات الاختبار

أولا : العوامل الخارجية

ثانياً : العوامل الداخلية

٦- التصحيح للتنقية

٧ - علاقة الموضوعية بالثبات

- معجم إنجليزى - عربى للمصطلحات
- مراجع الفصل الثانى

١ - مقدمـــة

يواجه القـائمين بالقيـاس فى مجال التربيـة البدنية والرياضة الكـثير من المشكلات التى تُعـرِّض عملية القياس لـلعديد من الصعوبات التـى قد يترتب عليها حدوث أخطاء فى عملية القياس. فقـد تستهدف عمليات القياس تقدير خصائص أو صفات بسيطة واضحة ومحددة مثل قياس بعض أبعاد الجسم، أو تستهـدف قياس بعض الصفات أو الـسمات أو القدرات المعقـدة مثل الذكاء، والقلق، والعدوان، واللياقة البدنية، أو القدرة الحركية وغيرها.

ومن الملاحظ أن القياس فى التـربيـة البدنية والرياضيـة قليـلا ما يتعامل مع الصفـات أو الخصائص البـسيطة Simple Qualities، فهو يتـعامل - فى معظم الحـالات - مـع العديـد مـن الصفـات والخصائـص والسمـات المعقـدة Complex traits، وهى لسوء الحظ صفات وخصائص وسمات لا ترى بالعين المجردة ولايـمكن حسـها أو لمسـها باليـدين، وإنما هى صـفات وخصـائص وسمـات يمكن التـعبـير عنهـا من خلال مظـاهر السلوك (الأداء) الـتى تدل عليـها، فى ذات الوقت فإنـه لا يمكن استـخدام المسطرة المدرجـة أو شريط القياس لقياس مثل تلك الصفات والخصائص.

وقد اهتم علماء القـياس بتقويم الأدوات التى يستخدمـونها سواء أكانت هذه الأدوات المسطرة المدرجة أو شريط القيـاس أو الميزان الطبى أو غيرها من أدوات ووسائل القياس، وفى هـذا الصدد يوضح نانلى Nunnally ١٩٧٨م أن أدوات القياس سواء استخدمت لقياس صفات بسيطة أو معقدة يجب أن تكون قادرة على إعـطاء نتائج دقيقـة فى كل مرة تستـخدم فيها مـا دامت الشروط موحدة فـى جميع مرات الـتطبيق. فإذا كـان شريط القيـاس المستخـدم قابلا للاستطالة (المط) أثناء التطبيق، فإنـه سوف يعطينا إما قراءات أكثر أو أقل من القراءات الحقيقية للظاهرة المقيسة.

من هذا المنطلق يحاول علماء القياس فهم سلوك الإنسان Human behavior على أساس تقويم ثبات الكثير من الاختبارات والمقاييس حتى لا يقعوا فى الأخطاء التى يمكن أن يتعرض لها «النجار» الذى يحاول بناء منزل من الخشب باستخدام شريط قياس من المطاط لن يعطيه أبداً أطوالا موحدة لنفس قطع الخشب المطلوب أن تكون متساوية فى الطول.

٢ - نظرية الثبات

يمكن القول بأن نظرية الثبات Reliability theory تعنى نظرية خطأ القياس the theory of measurement error، وهى نظرية تم تطويرها خلال المحاولات التى تمت فى مجال القياس النفسى psychometrics، والتى كانت تستهدف دائماً مراجعة وتقويم الثبات حتى تكون المقاييس measurements المتاحة للاستخدام يعتد بها من حيث الثقة فى النتائج التى يتم الحصول عليها.

وأخطاء القياس measurement errors ليست قاصرة على القياس التربوى والنفسى، ولكنها تحدث بالنسبة للقياس فى مجال علوم التربية البدنية والرياضة، وفى مجال العلوم الاجتماعية والبيولوجية وغيرها، فمن المعروف أن مقاييس measures الناتج القومى National product فى مجال الاقتصاد، ومقاييس ضغط الدم blood pressure فى مجال الطب تتعرض لأخطاء القياس كما يحدث بالنسبة للاختبارات التربوية والنفسية برغم الاهتمام الشديد بإعداد مثل هذه الاختبارات. وبسبب ظاهرة أخطاء القياس سوف يظل الثبات يستحوذ على اهتمام علماء القياس فى المجالات المختلفة.

دراسات سبيرمان المبكرة:

لقد حقق علم النفس تقدماً مشهوداً بالنسبة لتطوير ثبات أدواته ووسائله المختلفة وكان الفضل فى ذلك يرجع إلى المحاولات المبكرة التى قام بها عالم

النفس الإنجليزى الشهير تشارلز سبيرمان C. Spearman والذى استفاد فيها بما أشار إليه سلفه دى موفر De movire عن خطأ المعاينة Sampling error وهو خطأ ينجم عن اختيار عينة لا تمثل مجتمعها الأصلى. كما استفاد سبيرمان كذلك من الجهود التى قدمها كارل بيرسون Karl pearson فى إعداده لمعامل الارتباط التتابعى product moment correlation والذى نشره عام ١٨٩٦م. فقد وظف سبيرمان هذين الانجازين معا فى سياق عملية القياس ليضع بذلك معظم الأسس التى نعرفها الآن بالنسبة لنظرية الثبات والتى نشرها لأول مرة فى عام ١٩٠٤م تحت عنوان:

"The proof and measurement of association between two things"

ولأن مجلة علم النفس الإنجليزية British Journal of psychology لم تبدأ فى الصدور إلا عام ١٩٠٧م، لذا فقد قام سبيرمان بنشر أعماله التى توصل إليها فى مجلة علم النفس الأمريكية American Journal of psychology عام ١٩٠٤م.

وعندما نشر سبيرمان مقالته فى مجلة علم النفس الأمريكية فى يناير عام ١٩٠٤م لفتت انتباه رائد القياس النفسى فى أمريكا ادوارد ل. ثورندیك Edward L. Thorndike الذى كان يعد الطبعة الأولى من كتابه الشهير «مقدمة فى نظرية المقاييس العقلية والاجتماعية "Introduction to the theory of mental and social measurements".

وقد كان كتاب ثورندیك من الأعمال الأولى الرائدة التى ظهرت فى بداية القرن العشرين والتى تميزت بمستوى علمى رفيع لفت أنظار كل معاصريه من علماء القياس، فقد كان من أهم نتائج هذا الكتاب هو بدء الاهتمام بموضوع الثبات فى مجال القياس ليس فى أمريكا فحسب ولكن فى أوروبا أيضاً.

وبالإضافة إلى إنجازات كل من سبيرمان وثورندیك بالنسبة لنظرية الثبات تأتى إسهامات كل من كيودر، وريتشاردسون Kuder and Richardson ١٩٣٧م. وذلك بما قدماه من بعض الطرق الجديدة لتقدير معاملات الثبا

Reliabitity coefficients، والتى انتهت بما قدمه كرونباخ Cronbach ١٩٧٢م ومعاونوه من تطوير طرق تقويم العديد من مصادر الخطأ Sources of error فى البحوث السلوكية.

مفهوم الثبات:

ثبات القياس Reliability of measurement يعنى أن الدرجات التى يتم الحصول عليها دقيقة accurate وخالية من الخطأ error، وهذا يعنى أنه فى حالة تطبيق نفس أداة القياس (الاختبار أو المقياس) على نفس الفرد أو الشىء أى عدد من المرات بنفس الطريقة والشروط، فإننا سوف نحصل على نفس القيمة فى كل مرة، حيث تدل هذه القيمة على أن الشىء الذى تم قياسه لم يتغير فى غضون فترات القياس المختلفة.

ومن المتفق عليه أن الثبات يشير الى مدى الدقة والإتقان أو الاتساق الذى يقيس به الاختبار الظاهرة التى وضع من أجلها، لذلك يمثل الثبات أحد أهم الشروط الواجب توافرها فى أدوات القياس أيا كانت هذه الأدوات، وذلك لكونه - أى الثبات - يشير إلى مدى الوثوق فى الدرجات التى يتم الحصول عليها من تطبيق الأداة، بمعنى عدم تأثير الدرجات التى يتم الحصول عليها من تطبيق الأداة بأخطاء الصدفة، حيث يطلق على الثقة فى نتائج الأداة اسم: ثبات درجات الاختبار أو المقياس، فنحن لا نستطيع أن نثق فى اختبار للعدو طبق على تلميذ فى المرحلة الإعدادية فكان زمنه ٤ ثوان، وبعد أسبوع طبق نفس الاختبار على نفس التلميذ فكان زمنه ٧ ثوان. فكأن الثبات بهذا المعنى يدلل على الضمانات الخاصة بالحصول على نفس النتائج - تقريبا - إذا أعيد تطبيق الاختبار أو المقياس على نفس مجموعة الأفراد مع توحيد كافة الشروط والمتغيرات فى مرات التطبيق المختلفة، وهو ما يؤكد ضعف تأثير عوامل الصدفة أو العوامل الدخيلة أو العشوائية على نتائج الاختبار أو أداة القياس.

ومهما تعددت مفاهيم الثبات، فإنه يمكن إيضاح معناه فى النقاط التالية:

١- يختص الثبات بالنتائج التى يتم الحصول عليها من الاختبار أو المقياس، ولا يختص بالاختبار أو المقياس نفسه، وبناء على ذلك يصبح من الأكثر دقة أن نشير إلى ثبات درجات الاختبار أو المقياس بدلا من الإشارة إلى ثبات الاختبار أو أداة القياس.

٢- يبين الثبات مدى تأثير عوامل الصدفة (العشوائية) على نتائج الاختبار أو المقياس، وعليه فكلما كان الثبات مرتفعًا كان ذلك تأكيدًا على ضعف تأثير عوامل الصدفة، وكانت الدرجات التى يتم الحصول عليها دالة على المستوى الحقيقى للظاهرة المقيسة.

٣- الثبات يشير دائما إلى مستوى معين من الاتساق فى الدرجات، هذا المستوى من الاتساق مشروط بفترات زمنية، وعينات من الأداءات أو الأسئلة، ومفحوصين لهم خصائص محددة. فالثبات يختلف تبعا للفترات الزمنية بين مرات التطبيق المختلفة، ووفقا لخصائص الاختبار أو المقياس، ووفقا لخصائص عينة الأفراد.

٤- الثبات يعرَّف منطقيا على أنه نسبة التباين الحقيقى True variance إلى التباين العام total variance. والمقصود بالتباين هنا هو تباين الدرجات (القياسات) وليس تباين أداة القياس Measuring instrument، ويرجع السبب فى ذلك إلى أن القياس فى العلوم النفسية والتربوية والاجتماعية يتوقف على خصائص المجتمع الإحصائى الذى نقيسه باستخدام أداة القياس، وبناء على ذلك يظهر الثبات كمصطلح يشير إلى ثبات درجات الاختبار أو المقياس، وليس إلى ثبات الاختبار أو أداة القياس، فالاختبار الواحد يظهر قيما مختلفة من الثبات عندما يطبق على مجتمعات متباينة الخصائص مما يدل على أن الثبات يختص بالدرجة وليس بالأداة.

مما سبق يتضح أن ثبات أداة القياس يشير إلى الاتساق والدقة بالنسبة للنتائج التى يتم الحصول عليها من الأداة عندما يتم تكرار تطبيقها على نفس

٢ - لا يجوز أن يكـون الثبات بديلا عن الـصدق، فصدق الاختـبار أو أداة القياس يأتى فى المرتبة الأولى، لكون الصدق يعد أول معايير (محكات) الجودة الواجب توافـرها فى أداة القياس حتـى يمكن إقرار صلاحيـتها، ولكن تظل أداة القياس تتمتع بالصـدق طالما كانت تستخدم فى حدود ما وضعـت من أجله، وبدون أى حـذف أو إضافة أو تـغييـر فى محـتوى بنودها، وأيضا دون تغـيير فى خصائص عينة الأفراد الـتى أعدت عليهم فى الأصل. إلا أن الأمـر يختلف بالـنسبة للثـبات فعلـى الرغم من أن الأداة التى تتمتع بالصدق تتمتع بثـبات النتائج، إلا أن تقدير الثبات يظل إجراء تستـلزمه التجارب البـحثية فى كل مـرة تطبق فيهـا أداة القياس، فالثبات بالنسبة لنفس الاختبار أو المقياس يتعرض لأخطاء تتعلق بالتطبيق وعوامل الصدفة (الـعوامل العشوائية)، هذه الأخطاء والعـوامل تختلف مـن باحـث لآخـر، ومن مـوقف لآخـر، ومن وقـت لآخـر، ومن مفحوص لآخر.

٣ - يتسم الثبات بالصبغة الإحصائية على عكس من الصدق، فقد يتم تقدير الصدق عن طريق التـحليل المنطقى، إلا أن التحليل المنطـقى لأى اختبار لا يعطى أى دليـل عملى عن الثبـات، ومن ثم أصبحت طـرق حساب الثبات تتـطلب فهم العديد من العمـليات الإحصائية التى أضـفت عليه أهمية خاصة.

٣ - مفهوم الثبات النسبى وطرق حسابه

يصنف الثبات وفقاً لطرق حسابه إلى نوعين رئيسيين هما:

- الثبات النسبى Relative reliability
- الثبات المطلق Absolute reliability

وفيما يلى شرح للثبات النسبى وطرق حسابه:

الثبات النسبى:

وهو ثبات يتأسس على مدى استقرار الأوضاع النسبية للأفراد داخل المجموعة أو العينة فى مرات التطبيق المختلفة، ويستخدم لحساب هذا النوع من الثبات أسلوبين من أساليب التحليل الإحصائى هما:

١- حساب الارتباط بين الفئات (مجموعات الدرجات) interclass Correlation ويطبق هذا الأسلوب بالنسبة لطرق: التجزئة النصفية، والصور المتكافئة وإعادة الاختبار.

٢- تحليل التباين ANOVA ويستخدم لحساب الارتباط داخل الفئات (مجموعات الدرجات) intraclass Correlation وهو أسلوب يستخدم فى الحالات التى يتكرر فيها الاختبار عدد من المحاولات، حيث تستخدم معادلتى كيودر - ريتشاردسون KR_{20} , kR_{21} ومعامل ألفا كرونباخ alpha.

ويستخدم لتقدير ثبات درجات الاختبار (المقياس) خمس طرق رئيسية هى:

١ - طريقة إعادة الاختبار Test-Retest Method.

٢ - طريقة الصور المتكافئة Equivalent - Forms Method.

٣ - طريقة التجزئة النصفية Split - Half Method.

٤ - طريقة الاتساق الداخلى Internal - Consistancy Method.

٥ - طريقة تحليل التباين Analysis of Variance Method.

ومن الملاحظ أن هذه الطرق تستخدم الدرجات الخام التى يتم الحصول عليها نتيجة تطبيق أدوات القياس المختلفة، وأن استخدام أيا من هذه الطرق يتوقف على ثلاثة شروط رئيسية هى:

أ - طبيعة الاختبار.

ب - إحصاء العينة وإحصاء المجتمع الأصلى.

جـ - الغرض الذى سوف يستخدم من أجله إحصاء العينة.

وفيما يلى شرح موجز لكل طريقة من هذه الطرق:

طريقة إعادة الاختبار Test - Retest Method

تستخدم طريقة إعادة الاختبار لتقويم الخطأ المرتبط بتطبيق الاختبار مرتين مختلفتين بالنسبة لوقت التطبيق، بمعنى أن الثبات فى هذه الطريقة يستخدم المعاينة الزمنية time sampling لتسجيل درجات الأداء على الاختبار أو المقياس فى فترتين زمنيتين.

وتعد هذه الطريقة من أبسط الطرق وأسهلها فى تقدير الثبات، فهى تقوم على أساس تطبيق الاختبار أو المقياس على مجموعة من الأفراد، ثم يعاد التطبيق مرة أخرى على نفس المجموعة، ثم يحسب معامل الارتباط بين التطبيقين لنحصل بذلك على معامل ثبات الاختبار أو المقياس، حيث يدل معامل الثبات المحسوب على معامل استقرار الاختبار stability coefficient.

ويمكن حساب معامل الارتباط بين التطبيقين بطريقة بيرسون أو سبيرمان ويتم حساب الدلالة الإحصائية له باستخدام الجداول الإحصائية الخاصة بمعاملات الارتباط، وإن كان يفضل استخدام الاغتراب كمؤشر لقوة معامل الارتباط الدال على معامل الاستقرار (الثبات).

ويوصى علماء القياس باستخدام طريقة إعادة الاختبار فى الحالات التالية:

١- فى حالة قياس السمات والخصائص التى يتوقع لها الاستقرار stable traits، بمعنى أنها لا تتغير تغيرات سريعة وحادة عبر الزمن، ومن أمثلة هذه السمات: الذكاء العام والقدرات البدنية والحركية.

٢- فى حالة الاختبارات غير المتجانسة heterogeneous وهى نمط من الاختبارات تستخدم لقياس متغيرات بدنية أو حركية أو نفسية مختلفة من

حيث كونها تتضمن أبعاداً أو أجزاءً مختلفة، ففى الاختبارات غير المتجانسة نجد أن أجزاء مختلفة من الاختبار تقيس أبعادًا (أجزاء) مختلفة من السمات أو الخصائص، لذلك لا يتوقع فى هذه الحالة الحصول على مؤشر مرتفع لثبات الاتساق الداخلى لدرجات الاختبار أو المقياس ، وبناء على ذلك يوصى علماء القياس باستخدام طريقة إعادة الاختبار Retest لحساب ثبات الاختبارات غير المتجانسة فى المحتوى، حيث تشير هذه الطريقة إلى ثبات الاختبار عن طريق إعادة تطبيقه the Retest reliability أو معامل إعادة الاختبار Retesting coefficient وهو ما يعرف باستقرار stability درجات الاختبار.

وتصنف اختبارات السرعة speed tests (اختبارات الورقة والقلم) كاختبارات غير متجانسة لكونها تتطلب الإجابة على مجموعة من الأسئلة خلال زمن محدد، حيث لا يستطيع كل المفحوصين الإجابة عن كل الأسئلة، وعليه يمكن النظر إلى الأسئلة التى تم الإجابة عليها على أنها غير متجانسة بالنسبة للأسئلة التى لم يتم الإجابة عليها تأسيساً على أن الزمن المخصص للاختبار يكون محدداً لدرجة صعوبته، بحيث لا يكون فى إمكان أيا من المفحوصين أن يجيب على كل الوحدات أو الفقرات أو الأسئلة فى الزمن المخصص للاختبار على الرغم من فرضية أن جميع الأسئلة متماثلة فى درجة صعوبتها.

ويوصى علماء القياس بعدم استخدام طريقة إعادة الاختبار فى الحالات التالية:

١- بالنسبة للاختبارات التى تقيس سمات أو صفات أو خصائص تكتسب بالتعلم وتتأثر بالمران والممارسة والتدريب مثل اختبارات المهارات skills الحركية فى الألعاب والأنشطة الرياضية.

٢- بالنسبة للاختبارات التى تقيس ظواهر تتأثر بالنضج الناشئ عن العمر الزمنى وبخاصة فى مراحل سن ٤ و ٥ سنوات مثل اختبارات الذكاء واختبارات النمو الجسمى والحالة الغذائية وبعض اختبارات الإدراك الحس - حركى والتوافقات.

٣- بالنسبة للاختبارات التى تقيس ظواهر تتأثر بالأحداث التى تقع بين التطبيقيين الأول والثانى كالقراءة ومشاهدة التليفزيون والتفاعل مع الأقران من أفراد الأسرة أو زملاء الدراسة أو العمل أو غيرها، مثل استبيانات استطلاعات الرأى بالنسبة لبعض القضايا والمشكلات الاقتصادية والاجتماعية والسياسية وغيرها.

٤- بالنسبة للاختبارات التى تقيس ظواهر تتميز بالتذبذب أو التأرجح Fluctuate بمرور الوقت، فقد أقر عدد من علماء النفس أن بعض مظاهر النزوع السلوكى behavioral disposions قد لا تظهر بشكل متسق طوال الوقت، فالشخص العدوانى قد لا يظهر نفس درجة العدوان طوال الوقت، وعلى هذا الأساس يقرر أتكنسن atkinson ١٩٨١م أنه لا يفضل حساب ثبات مقاييس العدوان والدافعية باستخدام طريقة إعادة الاختبار، وبالمثل فإنه لا يفضل على الإطلاق استخدام نفس الطريقة لحساب ثبات بعض القياسات الفسيولوجية كمعدل القلب مثلا.

٥- بالنسبة للاختبارات التى ينتقل فيها تأثير الأداء فى التطبيق الأول الى الأداء فى التطبيق الثانى carry - over effect ومن أمثلة ذلك اختبارات التحصيل والاختبارات المعرفية حيث يتوقع أن يتذكر المفحوصون إجاباتهم فى التطبيق الأول، وأن يعرفوا الإجابات الصحيحة لبعض الأسئلة خلال الفترة بين التطبيقين.

الفترة الزمنية بين التطبيقين الأول والثانى:

يشير جرونلند Gronlund ١٩٨١م إلى أن طريقة إعادة الاختبار تعنى إعطاء نفس الاختبار مرتين لنفس مجموعة الأفراد، على أن تكون الفترة الزمنية بين المرتين تتراوح بين عدد من الدقائق حتى عدد من السنوات، وعليه سوف تظل الفترة الزمنية المناسبة بين التطبيقين الأول والثانى، مثار جدل مستمر بين علماء القياس، فإذا كانت هذه الفترة قصيرة لمدة - يوم أو يومين - فإن اتساق النتائج سوف يزداد نتيجة تذكر التلاميذ لبعض إجاباتهم فى التطبيق الأول فينقلونها إلى التطبيق الثانى، هذا من ناحية، ومن ناحية أخرى فقد يحدث أن يؤثر التعب الناشئ عن التطبيق الأول - وبخاصة فى الاختبارات البدنية - تأثيراً سلبياً على الأداء فى التطبيق الثانى فينخفض اتساق النتائج (معامل الثبات).

وإذا كانت الفترة الزمنية بين التطبيقين طويلة - شهرًا أو أكثر - فسوف يحدث الآتى:

- توقع حدوث تغيرات فى شروط وإجراءات التطبيق التى اتبعت فى المرة الأولى.
- توقع حدوث تغييرات فى سلوك المفحوصين خلال هذه الفترة الزمنية، وعليه فسوف يتوقع الحصول على معامل ثبات منخفض نتيجة طول الفترة الزمنية بين التطبيقين الأول والثانى.

وعموما فقد رأى علماء القياس أن الفترة الزمنية المناسبة بين التطبيقين تتوقف بشكل كبير على استخدامات النتائج التى سوف يتم الحصول عليها من الاختبار، فإذا كان الاختبار سوف يستخدم فى التنبؤ، حينئذ يشترط أن تتمتع درجات الاختبار بالاستقرار على مدى شهور قليلة حتى بضع سنوات، فاختبارات الاستعدادات البدنية تتمتع بثبات نسبى خلال نفس العام، وتصبح غير مستقرة إذا زادت الفترة عن العام.

وللخروج من هذه المشكلة الجدالية فقد أقر معظم المتخصصين فى القياس أن فترة أسبوعين بين التطبيقين الأول والثانى فترة مناسبة جدا لحساب الثبات بطريقة إعادة الاختبار لكونها ليست بالفترة القصيرة جداً أو الطويلة جداً.

طريقة الصور المتكافئة Equivalent - Forms Method:

يتطلب تقدير الثبات بطريقة الصور المتكافئة Parallel or equivalent forms استخدام صورتين متكافئتين من الاختبار، حيث يقصد بالتكافؤ هنا الآتى:

١- تساوى عدد الأسئلة فى الصورتين.

٢- تكافؤ درجة سهولة وصعوبة الأسئلة المتقابلة فى الصورتين (السؤال الأول فى الصورتين، والسؤال الثانى... إلخ).

٣- تساوى قيم معاملات الارتباط بين البنود فى كل من الصورتين (كل عبارة مع البعد الذى تنتمى إليه).

٤- تساوى قيم المتوسط الحسابى (س) والانحراف المعيارى (ع) لكلتا الصورتين ويتم استخدام هذه الطريقة **وفق الخطوات التالية:**

١- إعداد صورتين متكافئتين للاختبار إعدادا جيداً. وقد تكون لبعض الاختبارات صورتين (أو صيغتين) الصورة (أ) والصورة (ب).

٢- نقوم بتطبيق الصورتين المتكافئتين على مجموعة التقنين، وقد يتم تطبيق الصورتين معا فى نفس اليوم، وقد يتم تطبيق الصورة (أ) فى يوم والصورة (ب) بعد فترة زمنية مناسبة (يوم أو يومين أو ١٥ عشر يوما بحد أقصى).

٣- نقوم بحساب معامل الارتباط بين درجات مجموعة التقنين على الصورتين، حيث يدل معامل الارتباط المحسوب على معامل التكافؤ coefficient of equivalence والذى يعرف بعدة أسماء هى:

- معامل ثبات الصور المتبادلة Alternate - Forms Reliability

- معامل ثبات الصورة المتوازية Parallel - Forms Reliability

- معامل ثبات الصور المتساوية Comparable - forms Reliabitity

وعادة ما يستخدم معامل ارتباط بيرسون لحساب هذا المعامل، حيث يدلنا هذا المعامل على الآتى:

- يعد مقياساً للتكافؤ.
- يبين الدرجة التى تقيس بها كل من صورتى الاختبار نفس المظاهر أو الموضوعات المطلوب قياسها.
- إذا كان معامل الارتباط (الثبات) مرتفعاً دل ذلك على صلاحية الصورتين وحسن إعدادهما، وإذا جاء هذا المعامل منخفضاً دل ذلك على عدم صلاحية إحدى الصورتين أو كلاهما.

ومن الملاحظ أن طريقة الصور المتكافئة لا تدلنا على ثبات الخصائص المقيسة، وإنما تبين لنا إلى أى مدى يمثل الاختبار عينة من الخصائص والسمات ومظاهر السلوك المطلوب قياسها. فعند قياس التحصيل - على سبيل المثال - تظهر الآلاف من الأسئلة التى تستلزم أن يتضمنها الاختبار، ومع ذلك، ولأسباب تتعلق بالوقت والإمكانات المتاحة، يحدث أن يقتصر إعداد مثل هذه النوعية من الاختبارات على أعداد محدودة من الأسئلة التى يمكن استخدامها، هذه المجموعة من الأسئلة تمثل أفضل عينة من الأسئلة الممكنة لتمثيل المجال الذى تقيسه.

ويمكن النظر إلى طريقة الصور المتكافئة على أنها تطبيق للاختبار مرتين متتاليتين، حيث تختص كل مرة بإحدى صور الاختبار، ووفقا لهذا فإنه يمكن النظر إلى معامل الثبات المحسوب بطريقة الصورة المتكافئة على أنه مؤشر لاستقرار stability الاختبار وتكافؤ صورتيه equivalence فى ذات الوقت.

متى تستخدم طريقة الصور المتكافئة؟

تستخدم هذه الطريقة بشكل محدود فى مجال الاختبارات البدنية والحركية، لكونها خاصة باختبارات الورقة والقلم paper - and pencil tests التى تتطلب من المفحوص أن يعبر عن استجابته باستخدام الكتابة، وهى اختبارات تختلف تماماً عن الاختبارات العملية (الأدائية) performance tests التى تتطلب الاستجابة الحركية motor response وليس الاستجابة اللفظية (اللغوية) verbal response.

وعموما فإنه يمكن استخدام هذه الطريقة بالنسبة لاختبارات الورقة والقلم المقننة Standardized tests لكون معظم هذه الاختبارات له صورتان أو أكثر، من هذه الاختبارات على سبيل المثال: اختبارات التحصيل، والاختبارات المعرفية، وبعض القوائم والصور الخاصة بقياس الشخصية.

طريقة التجزئة النصفية Split - Half Method:

يمكن تقدير ثبات درجات الاختبار أو المقياس عن طريق التطبيق لمرة واحدة فقط، فى هذه الطريقة يطبق الاختبار أو المقياس على مجموعة واحدة من المفحوصين ثم يجزء الاختبار أو المقياس إلى نصفين متكافئين كالتالى:

١ - النصف الأول من العبارات فى مقابل النصف الثانى من العبارات:

الـ ٥٠ سؤالا الأولى فى مقابل الـ ٥٠ سؤالا الثانية

(فى حالة اختبار يتكون من ١٠٠ سؤال أو عبارة)

أو:

٢ - الأسئلة (العبارات) ذات الأرقام الفردية فى مقابل الأسئلة (العبارات) ذات الأرقام الزوجية.

٣ - وفق ما سبق يصبح لكل مفحوص مجموعتان من الدرجات الخام هما:

- مجموع درجات النصف الأول من العبارات (س)

- مجموع درجات النصف الثانى من العبارات (ص)

أو :

- مجموع درجات العبارات الفردية (س)

- مجموعة درجات العبارات الزوجية (ص)

٤ - يحسب معامل الارتباط بين (س، ص) لكل المفحوصين باستخدام معامل ارتباط بيرسون فنحصل على معامل ثبات نصف الاختبار وهو:

$\frac{١}{٢} \cdot \frac{١}{٢}$ = معامل ثبات نصف الاختبار

ملحوظة مهمة:

تتطلب طريقة التجزئة النصفية تقسيم الاختبار إلى نصفين متساويين أو قريبين من التساوى، فإذا كان لدينا اختبار يتكون من (١٠) عبارات أو أسئلة، فإنه يمكن تقسيم هذا الاختبار إلى نصفين متساويين كالتالى:

أ- النصف الأول من العبارات = ٥ عبارات وهى العبارات أرقام (١، ٢، ٣، ٤، ٥) فى مقابل النصف الثانى من العبارات = ٥ عبارات وهى العبارات أرقام (٦، ٧، ٨، ٩، ١٠).

أو :

ب- العبارات الفردية = ٥ عبارات وهى العبارات أرقام (١، ٣، ٥، ٧، ٩) فى مقابل العبارات الزوجية = ٥ عبارات وهى العبارات أرقام (٢، ٤، ٦، ٨، ١٠).

أما إذا كان عدد العبارات ١١ عبارة، فإنه يمكن تقسيم هذا الاختبار إلى نصفين قريبين من التساوى nearly equal halves كالتالى:

أ- النصف الأول من العبارات = ٥ عبارات وهى العبارات أرقام (١، ٢، ٣، ٤، ٥)، فى مقابل النصف الثانى من العبارات = ٦ عبارات وهى العبارات أرقام (٦، ٧، ٨، ٩، ١٠، ١١).

أو:

ب- العبارات الفردية = ٦ عبارات وهى العبارات أرقام (١، ٣، ٥، ٧، ٩، ١١) فى مقابل العبارات الزوجية = ٥ عبارات وهى العبارات أرقام (٢، ٤، ٦، ٨، ١٠).

طرق تصحيح معامل ثبات نصف الاختبار:

يستخدم لتصحيح معامل ثبات نصف الاختبار للحصول على معامل ثبات الاختبار ككل أى من المعادلات التالية:

معادلة سبيرمان - براون: Spearman - Brown formula الصورة المختصرة وهى:

$$\text{معامل الثبات} = \frac{\text{٢} \times \text{معامل ثبات النصف}}{\text{١} + \text{معامل ثبات النصف}} \qquad \text{معادلة (٢-١)}$$

حيث إن:

r_{11} = معامل ثبات كل الاختبار.

$r_{\frac{1}{2}\frac{1}{2}}$ = معامل الارتباط بين نصفى الاختبار (معامل ثبات نصف الاختبار).

٢ = مقدار ثابت.

١ = مقدار ثابت.

مثال: (٢ - ١)

احسب معامل ثبات الاختبار ككل إذا كان معامل الارتباط بين نصفى الاختبار يساوى ٠,٧٠

بتطبيق المعادلة (٢ - ١) ينتج أن:

$$r_{11} = \frac{٢ \times ٠,٧٠}{١ + ٠,٧٠}$$

$$= \frac{١,٤}{١,٧٠}$$

$$= ٠,٨٢$$

وتستخدم هذه المعادلة فى اختبارات التحصيل والقدرات، حيث يشترط بالنسبة لاستخدامها أن تكون تباينات variances نصفى الاختبار متساوية.

معادلة رولون The Rulon formula **وهى:**

$$\overset{\frown}{ر}_{ث ث} = ١ - \frac{ع_{د}^{٢}}{ع_{ت}^{٢}}$$

معادلة (٢ - ٢)

حيث إن:

$\overset{\frown}{ر}_{ث ث}$ = معامل ثبات كل الاختبار.

$ع_{د}^{٢}$ = تباين الفرق بين درجات نصفى الاختبار بالنسبة لكل مفحوص.

$ع_{ت}^{٢}$ = تباين الدرجة الكلية (ويقصد بها درجة كل مفحوص على نصفى الاختبار).

مثال: (٢ - ٢)

طبق اختبار يتكون من ١٢ عبارة على ١٠ مفحوصين فكان مجموع درجات العبارات الفردية، ومجموع العبارات الزوجية لكل مفحوص كالتالى:

المفحوص:	أ	ب	جـ	د	هـ	و	ز	ح	ط	ي
مجموع العبارات الفردية:	٦	٦	٢	٣	٤	٢	٤	٦	٥	٥
مجموع العبارات الزوجية:	٥	٥	٢	٢	٣	٠	٤	٦	٦	٤

والمطلوب:

حساب معامل ثبات كل الاختبار باستخدام معادلة «رولون».

الحـــل: نقـــوم بعمل جدول إحـصائى متـضمنًا العـمليات الإحصـائية الأساسية كالتالى:

المفحوص	قيم العبارات الفردية $س_ي$	قيم العبارات الزوجية $س_ز$	الفرق $ع_د$ ($س_د - س_ز$)	مربع الفرق $ع_د^2$	قيم جميع العبارات س	مربع $س^2$	مربع $س_د^2$	مربع $س_ز^2$
أ	٦	٥	+١	١	١١	١٢١	٣٦	٢٥
ب	٦	٥	+١	١	١١	١٢١	٣٦	٢٥
جـ	٢	٢	٠	٠	٤	١٦	٤	٤
د	٣	٢	+١	١	٥	٢٥	٩	٤
هـ	٤	٣	+١	١	٧	٤٩	١٦	٩
و	٢	٠	+٢	٤	٢	٤	٤	٠
ز	٤	٤	٠	٠	٨	٦٤	١٦	١٦
ح	٦	٦	٠	٠	١٢	١٤٤	٣٦	٣٦
ط	٥	٦	-١	١	١١	١٢١	٢٥	٣٦
ى	٥	٤	+١	١	٩	٨١	٢٥	١٦
	مجـ $س_ي$ = ٤٣	مجـ $س_ز$ = ٣٧	مجـ $ع_د$ = ٦	مجـ $ع_د^2$ = ١٠	مجـ س = ٨٠	مجـ $س^2$ = ٧٤٦	مجـ $س_د^2$ = ٢٠٧	مجـ $س_ز^2$ = ١٧١

يلاحظ من العمليات الإحصائية التى يتضمنها الجدول السابق أن:

مجموع قيم العبارات الفردية مجـ $س_د$ = ٤٣ درجة

ومجموع قيم العبارات الزوجية مجـ $س_ز$ = ٣٧ درجة

ومجموع الفروق مجـ $ع_د$ = ٦

ومجموع مربع الفروق مجـ $ع_د^2$ = ١٠

ومجموع قيم جميع العبارات مجـ س = ٨٠

ومجموع مربعات جميع القيم مجـ $س^2$ = ٧٤٦

ومجموع مربعات المحاولات الفردية مجـ $س_د^2$ = ٢٠٧

ومجموع مربعات المحاولات الزوجية مجـ $س_ز^2$ = ١٧١

ولحساب تباين الفرق $ع_د^2$ تطبق المعادلة التالية:

$$ع_د = \frac{1}{ن} \sqrt{ن \text{ مجـ } ع_د^2 - (\text{مجـ } ع_د)^2}$$ معادلة (٢ - ٣)

وبالتعويض فى هذه المعادلة ينتج أن:

$$ع_د = \frac{1}{١٠} \sqrt{[١٠ \times ١٠ - (٦)^2]}$$

$$= \frac{1}{١٠} \sqrt{(١٠٠ - ٣٦)}$$

$$= ٠,٨$$

إذا $ع_د^2$ $= (٠,٨)^2 = ٠,٦٤$

(حيث إن التباين = مربع الانحراف المعيارى)

وبالمثل يتم حساب التباين الكلى للاختبار $ع^{2}_{ن}$ بتطبيق المعادلة التالية:

$$ع_{ت} = \frac{1}{ن}\sqrt{ن\ مجـ\ س^{2} - (مجـ\ س)^{2}}$$ معادلة (٢ -٤)

وبالتعويض فى هذه المعادلة ينتج أن:

$$ع_{ت} = \sqrt{(١٠ \times ٧٤٦ - (٨٠)^{2}}$$

$$= ٣,٢٦$$

$$إذا\ ع^{2}_{ن} = (٣,٢٦)^{2} = ١٠,٦٣$$

(حيث إن التباين = مربع الانحراف المعيارى)

أسفرت العمليات الإحصائية السابقة عن الآتى:

$$ع^{2}_{ن} = ٠,٦٤$$

$$ع^{2}_{ن} = ١٠,٦٣$$

حينئذ يمكن حساب معامل ثبات كل الاختبار بتطبيق معادلة رولون كالتالى:

$$\frac{ن}{ن-١} = ١ - \frac{٠,٦٤}{١٠,٦٣}$$

$$= ١ - ٠,٠٦$$

$$= ٠,٩٤$$

معادلة فلانجان Flangan Formula وهى:

$$\text{ر}_{\text{ن ن}} = ٢ \left(١ - \frac{\text{ع}_{١}^{٢} + \text{ع}_{٢}^{٢}}{\text{ع}_{\text{ن}}^{٢}}\right) \qquad \text{معادلة (٢ - ٥)}$$

حيث إن :

$\text{ر}_{\text{ن ن}}$ = معامل ثبات كل الاختبار

$\text{ع}_{٢}^{٢}$ = تباين درجات النصف الأول للاختبار

$\text{ع}_{٢}^{٢}$ = تباين درجات النصف الثانى للاختبار

$\text{ع}_{\text{ن}}^{٢}$ = التباين الكلى للاختبار

ويمكن تطبيق معادلة فلانجان على البيانات المعطاة فى المثال السابق رقم (٢ - ٢) لحساب معامل ثبات كل الاختبار وفق الخطوات التالية:

أولا : حساب تباين درجات النصف الأول للاختبار ($\text{ع}_{١}^{٢}$) بتطبيق المعادلة (٢ - ٤) وهى:

$$\text{ع}_{١} = \sqrt{\frac{\text{ن مجـ س}_{\text{ف}}^{٢} - (\text{مجـ س}_{\text{ف}})^{٢}}{\text{ن}^{٢}}}$$

حيث إن :

ن = ١٠

مجـ $\text{س}_{\text{ف}}^{٢}$ = ٢٠٧ (للعبارات الفردية)

مجـ $\text{س}_{\text{ف}}$ = ٤٣ (مجموع قيم العبارات الفردية)

وبالتعويض فى المعادلة رقم (٢ - ٤) ينتج أن:

$$ع_{١} = \frac{١}{١٠} \sqrt{١٠ \times ٢٠٧ - (٤٣)^{٢}}$$

$$= \frac{١}{١٠} \sqrt{٢٠٧٠ - ١٨٤٩}$$

$$= ١,٤٩$$

إذاً $ع_{١}^{٢} = (١,٤٩)^{٢} = ٢,٢٢$

ثانيا : حساب تباين درجات النصف الثانى للاختبار ($ع_{٢}^{٢}$) بتطبيق نفس المعادلة (٢ - ٤) حيث إن :

ن = ١٠

مجـ س٢ = ١٧١ (للمحاولات الزوجية)

مجـ س = ٣٧ (مجموع قيم العبارات الزوجية)

وبالتعويض فى المعادلة (٢ - ٤) ينتج أن:

$$ع_{٢}^{٢} = \frac{١}{١٠} \sqrt{١٠ \times ١٧١ - (٣٧)^{٢}}$$

$$= \frac{١}{١٠} \sqrt{١٧١٠ - ١٣٦٩}$$

$$= ١,٨٥$$

إذا $ع_{٢}^{٢} = (١,٨٥)^{٢} = ٣,٤٢$

ثالثا : لحساب التباين الكلى للاختبار ($ع_{ت}^{٢}$) بتطبيق نفس المعادلة (٢ - ٤) حيث يظهر أن :

$ع_{ت}^{٢} = ١٠,٦٣$ (من تطبيق معادلة رولون)

رابعا : يتم حساب ثبات كل الاختبار بتطبيق معادلة فلانجان (٢-٥) كالتالى :

$$\text{ر}_{١٠١} = ٢ \left(١ - \frac{٣,٤٢ + ٢,٢٢}{١٠,٦٣}\right)$$

$$= ٢ \left(١ - \frac{٥,٦٤}{١٠,٦٣}\right)$$

$$= ٠,٩٣٩$$

إذاً $\text{ر}_{١٠١} = ٠,٩٤$

(وهى نفس النتيجة التى تم الحصول عليها من معادلة رولون)

مما سبق يلاحظ الآتى :

أ - يتطلب استخدام معادلة رولون توافر نوعين من البيانات هما :

١ - تباين الفرق بين درجات المفحوصين على نصفى الاختبار .

٢ - التباين الكلى للاختبار .

ب- ويتطلب استخدام معادلة فلانجان ثلاثة أنواع من البيانات هى :

١ - تباين درجات النصف الأول للاختبار .

٢ - تباين درجات النصف الثانى للاختبار .

٣ - تباين درجات كل الاختبار .

ومن الملاحظات الجديرة بالاهتمام أنه يمكن استخدام طريقة التجزئة النصفية لحساب ثبات درجات بعض الاختبارات الحركية والقياسات الجسمية التى يمكن تطبيقها لعدد من المحاولات أو المرات (القياسات) المتتالية دون أن يتأثر الأداء بالتعب أو بانتقال أثر التدريب .

ومن أمثلة هذه الاختبارات:

١ - **اختبارات القدرة العضلية** Muscular power (القوة السريعة أو المتفجرة) مثل اختبارات: الوثب العمودى vertical Jump ، والوثب العريض من الثبات standing broad jump، ودفع الكرة الطبية ٣كجم Medicine Ball Put وغيرها، حيث تسمح شروط الاختبار بإعطاء المختبر (المفحوص) ثلاث أو أربع محاولات متتالية فى نفس الوقت.

٢ - **اختبارات التوافق والدقة** وبعض الاستجابات الحركية الانتقائية كما فى حالة اختبار نيلسون للاستجابة الحركية الانتقائية Nelson choice Response movement حيث يعطى المختبر ١٠ محاولات متتالية بين كل محاولة والأخرى ٢٠ ثانية، واختبار نيلسون للسرعة الحركية Nelson speed of movement حيث يعطى المختبر ٢٠ محاولة متتالية، وغيرها من الاختبارات(١).

٣ - **اختبارات المهارات فى الألعاب الرياضية**، حيث تسمح الشروط الخاصة بمعظم هذه الاختبارات بتنفيذ الاختبار لعدد من المحاولات المتتالية مرة واحدة(٢).

٤ - **القياسات الجسمية (المورفولوجية)** كما فى حالة قياس أطوال أو محيطات أو اتساعات بعض أجزاء الجسم حيث يفضل إجراء القياس الواحد لعدد ٣ أو ٤ مرات متتالية(٣).

---

(١) انظر هذه الاختبارات وغيرها فى المرجع الآتى:
محمد حسن علاوى ومحمد نصرالدين رضوان: اختبارات الأداء الحركى، دار الفكر العربى، ٢٠٠١م.

(٢) انظر هذه الاختبارات وغيرها فى المراجع التالية:
محمد حسن علاوى ومحمد نصر الدين رضوان: الاختبارات المهارية والنفسية فى المجال الرياضى، دار الفكر العربى، ١٩٨٧م.

(٣) محمد نصر الدين رضوان: المرجع فى القياس الجسمية، دار الفكر العربى، ١٩٩٧م.

مثال : (٢ - ٣)

طبق اختبار دفع الكرة الطبية (٣ كجم) على خمسة لاعبين، وكان كل لاعب يقوم بتنفيذ الاختبار عدد ٦ محاولات متتالية، فكانت النتائج مسجلة بالمتر كالتالى:

	المحاولات					
اللاعب	١	٢	٣	٤	٥	٦
أ	١٠	١١	١١	١٢	١١	١٢
ب	١٤	١٣	١٤	١٤	١٥	١٤
جـ	١٢	١٢	١٢	١٣	١٢	١٣
د	٩	٨	٨	٩	٩	٨
هـ	٧	٩	٨	٨	٩	٩

والمطلوب :

أولا : حساب معامل ثبات الاختبار ككل باستخدام معادلة سبيرمان - براون.

ثانيا : حساب معامل ثبات الاختبار ككل باستخدام معادلة كل من: رولون، وفلانجان.

الحل :

أولا : يتم حساب ثبات الاختبار ككل باستخدام معادلة سبيرمان - براون كالتالى:

١ - تقسم محاولات الاختبار إلى نصفين، حيث يمثل النصف الأول مجموع نتائج المحاولات الفردية، ويمثل النصف الثانى مجموع نتائج المحاولات الزوجية لكل لاعب على حدة كالتالى:

اللاعب	مجموع المحاولات الفردية	مجموع المحاولات الزوجية
أ	٣٢	٣٥
ب	٤٣	٤١
جـ	٣٦	٣٨
د	٢٦	٢٥
هـ	٢٤	٢٦

٢ - مما سبق يتبين أن درجات اللاعبين فى المحاولات الست للاختبار أصبحت تتكون من مجموعتين، المجموعة الأولى وتضم درجات المحاولات الفردية (س)، والمجموعة الثانية وتضم درجات المحاولات الزوجية (ص)، وبتطبيق معامل ارتباط بيرسون للدرجات الخام، فإنه يمكن حساب معامل الارتباط بين نصفى الاختبار، حيث نجد أن هذا المعامل يساوى ٨٩ , ٠ أى أن:

$$\text{(معامل ثبات نصف الاختبار)} \quad = ٨٩ , ٠ \quad \sqrt{\frac{١}{٢} \cdot \frac{١}{٢}}$$

وبتطبيق معادلة سبيرمان - براون (معادلة : ٢ - ١) لتصحيح معامل ثبات نصف الاختبار والحصول على معامل ثبات الاختبار ككل ينتج أن:

$$\text{معامل ثبات الاختبار ككل} = \frac{٢ \times ٨٩ , ٠}{١ + ٨٩ , ٠}$$

$$= \frac{١ , ٨٧}{١ , ٨٩}$$

$$= ٩٤ , ٠$$

إذا معامل الثبات لجميع المحاولات (٦ محاولات) هو ٩٤ , ٠

ثانيا : لحساب معامل ثبات الاختبار ككل باستخدام معادلة رولون، ومعادلة فلانجان، تستخدم الخطوات التالية:

١ - نقوم بوضع نتائج المحاولات الفردية والمحاولات الزوجية فى جدول إحصائى كالتالى:

المفحوص	قيم المحاولات الفردية $س_د$	قيم المحاولات الزوجية $س_ز$	الفرق $ع_د$ $(س_د - س_ز)$	مربع الفرق $ع_د$	قيم جميع المحاولات س	مربع القيم $س^2$	مربع $س^2_د$	مربع $س^2_ز$
أ	٣٢	٣٥	-٣	٩	٦٧	٤٤٨٩	١٠٢٤	١٢٢٥
ب	٤٣	٤١	+٣	٤	٨٤	٧٠٥٦	١٨٤٩	١٦٨١
جـ	٣٦	٣٨	-٢	٤	٧٤	٥٤٧٦	١٢٩٦	١٤٤٤
د	٢٦	٢٥	+١	١	٥١	٢٦٠١	٦٧٦	٦٢٥
هـ	٢٤	٢٦	-٢	٤	٥٠	٢٥٠٠	٥٧٦	٦٧٦
	مجـ $س_د$ = ١٦١	مجـ $س_ز$ = ١٦٥	مجـ $ع_د$ = -٤	$ع^2_د$ = ٢٢	مجـ س = ٣٢٦	مجـ $س^2$ = ٢٢١٢٢	مجـ $س^2_د$ = ٥٤٢١	مجـ $س^2_ز$ = ٥٦٥١

٢ - نقوم بإجراء العمليات الإحصائية وفق ما هو مبين بالجدول السابق حيث يلاحظ أن:

مجـ $س_د$ = ١٦١

مجـ $س_ز$ = ١٦٥

مجـ $ع_د$ = -٤

مجـ $ع^2_د$ = ٢٢

مجـ س = ٣٢٦

مجـ $س^2$ = ٢٢١٢٢

مجـ $س^2_د$ = ٥٤٢١

مجـ $س^2_ز$ = ٥٦٥١

٣ – يتم حساب تباين الفرق ($ع^{٢}_{د}$) بتطبيق المعادلة (٢ – ٣) كالتالى :

$$ع_{د} = \frac{١}{٥} \sqrt{٥ \times ٢٢ - (-٤)^{٢}}$$

$$= ٠,٢ \sqrt{١١٠ - ١٦}$$

$$= ٠,٢ \times ٩٤$$

$$= ٠,٢ \times ٩,٧٠$$

$$= ١,٩٣٩$$

إذًا $ع^{٢}_{د} = (١,٩٣٩)^{٢} = ٣,٧٦$

٤ – وبالمثل يتم حساب التباين الكلى للاختبار $ع^{٢}_{ت}$ بتطبيق المعادلة (٢-٤) كالتالى:

$$ع_{ت} = \frac{١}{٥} \sqrt{٥ \times ٢٢١٢٢ - (٣٢٦)^{٢}}$$

$$= ٠,٢ \sqrt{١١٠٦١٠ - ١٠٦٢٧٦}$$

$$= ٠,٢ \sqrt{٤٣٣٤}$$

$$= ٠,٢ \times ٦٥,٨٣٣$$

$$= ١٣,١٦٧$$

إذًا $ع^{٢}_{ت} = (١٣,١٦٧)^{٢} = ١٧٣,٣٦$

٥ - بالتعويض فى معادلة رولون ينتج أن

$$\text{ر}_{١٠١} = ١ - \frac{٣,٧٦}{١٧٣,٣٦}$$

$$= ١ - ٠,٠٢$$

$$= ٠,٩٨$$

وبالمثل يمكن تطبيق معادلة فلانجان وفق الخطوات التالية:

١ - حساب تباين درجات النصف الأول من للاختبار ($ع_{١}^{٢}$) بتطبيق المعادلة (٢ - ٤) كالتالى :

$$ع_{١} = \frac{١}{٥} \sqrt{٥ \times ٥٤٢١ - (١٦١)^{٢}}$$

$$= ٢,٠ \sqrt{٢٧١٠٥ - ٢٥٩٢١}$$

$$= ٢,٠ \sqrt{١١٨٤}$$

$$= ٢,٠ \times ٣٤,٤١ = ٦,٨٨$$

$$\text{إذاً } ع_{١}^{٢} = (٦,٨٨)^{٢} = ٤٧,٣٣$$

٢- حساب تباين درجات النصف الثانى للاختبار ($ع_{٢}^{٢}$) بتطبيق نفس المعادلة (٢-٤) كالتالى :

$$= \frac{١}{٥} \sqrt{٥ \times ٥٥٦٥١ - (١٦٥)^{٢}}$$

$$= ٢,٠ \sqrt{٢٨٥٥ - ٢٧٢٥}$$

$$= ٢,٠ \sqrt{١٠٣٠}$$

$= ٠,٢ \times ٣٢,٠٩$

$= ٦,٤٢$

إذا $ع_{١}^{٢} = (٦,٤٢)^{٢} = ٤١,٢٢$

٣ - حساب التباين الكلى للاختبار ($ع_{ن}^{٢}$) وقد بلغ ١٧٣,٣٦ (الخطوة رقم: ٣ من إجراءات تطبيق معادلة رولون).

٤ - وبتطبيق معادلة فلانجان معادلة (٢-٥) على هذه المجموعة من المعطيات ينتج أن :

$$\text{ر}_{١٠١} = ٢ (١ - (\frac{٤١,٢٢ + ٤٧,٣٣}{١٧٣,٣٦}))$$

$$= ٢ (١ - (\frac{٨٨,٥٥}{١٧٣,٣٦}))$$

$$= ٢ (١ - ٠,٥١)$$

$$= ٢ \times ٠,٤٨٩$$

$= ٠,٩٨$ (وهى نفس النتيجة التى تم الحصول عليها من معادلة رولون)

مميزات طريقة التجزئة النصفية:

تمتاز طريقة التجزئة النصفية بالآتى:

١- تستخدم عندما يتعذر استخدام أيا من الطريقتين السابقتين (إعادة الاختبار أو الصور المتكافئة).

٢- تعد مناسبة لبعض الاختبارات التى تتأثر بشكل حساس بالتعليم وانتقال أثر التدريب أو التعب الذى يحدث نتيجة التطبيق الأول ويمكن أن ينتقل تأثيره إلى التطبيق الثانى.

٣ - تطبيق الاختبار مرة واحدة بدلاً من مرتين أو إعداد صورة متكافئة للاختبار، يعمل على توفير المجهود والوقت والتغلب على ظاهرة تسرب بعض المفحوصين (المختبرين) فى التطبيق الثانى.

٤ - تطبيق الاختبار لمرة واحدة فقط يجنبنا جدلية الفترة الزمنية الفاصلة بين التطبيقين الأول والثانى والتى ما تزال محل اهتمام علماء القياس.

٥ - تمكننا من التغلب على المشكلات المتعلقة باختلاف شروط التطبيق والتى يحتمل أن تحدث فى المرة الثانية للتطبيق، وأيضا التغلب على التباين المحتمل حدوثه بالنسبة للتعليمات المعطاة للمختبرين عندما يطبق عليهم الاختبار مرتين.

٦- تصنف طريقة التجزئة النصفية كأحد مقاييس الاتساق الداخلى Measure of internal consistency لكونها تتعامل مع جميع مفردات الدرجات مقسمة على قسمين أو نصفين.

المآخذ على هذه الطريقة:

١ - قد يتعرض المفحوصين لعوامل الإجهاد والملل والضيق عندما تطبق اختبارات أو قوائم تحتوى على عدد كبير من الأسئلة مثلا (١٢٠ سؤال أو عبارة)، فإذا أخذنا البنود أرقام من (١ - ٦٠) ممثلة للنصف الأول من الاختبار، والبنود من (٦١ - ١٢٠) ممثلة للنصف الثانى من الاختبار، فى هذه الحالة نجد أن إجابات المفحوصين على النصف الثانى من الاختبار سوف لا تعطى نتائج يوثق بها، وذلك لسببين هما:

أ- أن تركيز الاهتمام فى الإجابة على عبارات النصف الأول قد يستهلك جهد وطاقة المفحوص.

ب- أن ضيق الوقت أو التعب والملل أو غيرهما قد تجعل الاهتمام بالإجابة على عبارات النصف الثانى أقل منه بالنسبة للنصف الأول من الاختبار.

٢- أنه فى حالة الأخذ بأسلوب التجزئة النصفية للاختبار على أساس العبارات الفردية odd فى نصف، والعبارات الزوجية even فى النصف الآخر، فى هذه الحالة فإنه يحتمل أن نحصل على تباينات غير متساوية لنصفى الاختبار لكون العبارات فى النصفين ليست على درجة واحدة من الصعوبة.

٣- قد تتضمن بعض الاختبارات أو القوائم الأسئلة (البنود أو العبارات) بدون ترتيب محدد كقوائم الشخصية واختبارات السرعة مثلا، مما يمثل تحديا لهذه الطريقة، لذا ينصح بعض المتخصصين القيام بترتيب الأسئلة فى مثل هذه الحالة حسب درجة صعوبتها ثم إعطائها أرقاماً مسلسلة بعد ذلك.

٤- يقتصر بعض الباحثين عند حساب الثبات بطريقة التجزئة النصفية على تقدير معامل الارتباط بين نصفى الاختبار فقط، دون تطبيق معادلة سبيرمان - براون لتصحيح هذا المعامل، وهذا الخطأ الشائع يجب التنبه له لكون معامل الارتباط بين نصفى الاختبار يشير فقط إلى ثبات نصف الاختبار.

والجدول (٢-٢) يبين تأثير استخدام معادلة سبيرمان - براون على معاملات الارتباط بين نصفى الاختبار.

جدول (٢-٢)

يوضح تأثير استخدام معادلة سبيرمان - براون لتصحيح معاملات الارتباط بين نصفى الاختبار

معاملات الارتباط بين نصفى الاختبار	معاملات الارتباط بعد التصحيح	مقدار التغير
٠,٠٥	٠,٠٩	٠,٠٤
٠,١٥	٠,٢٦	٠,١١
٠,٢٥	٠,٤٠	٠,١٥
٠,٣٥	٠,٥٢	٠,١٧
٠,٤٥	٠,٦٢	٠,١٧
٠,٥٥	٠,٧١	٠,١٦
٠,٦٥	٠,٧٩	٠,١٤
٠,٧٥	٠,٨٦	٠,١١
٠,٨٥	٠,٩٢	٠,٠٧
٠,٩٥	٠,٩٧	٠,٠٢

طريقة الاتساق الداخلى Internal Consistency Method

توجد بالإضافة إلى طريقة التجزئة النصفية مجموعة أخرى من الطرق لتقدير الاتساق الداخلى للاختبار. ففى عام ١٩٣٧م قدم كيودر وريتشاردسون Kuder and Richardson إنجازهما المتميز لتقويم الثبات من خلال تطبيق الاختبار مرة واحدة دون اللجوء إلى تقسيم الاختبار إلى نصفين بأية طريقة تحكمية، وأيضا للتغلب على مشكلة عدم تساوى التباينات بين نصفى الاختبار والتى تعد من أهم متطلبات طريقة التجزئة النصفية.

وتُعنى طريقة الاتساق الداخلى بالآتى:

١- مدى ارتباط البنود أو الفقرات أو الوحدات مع بعضها البعض داخل الاختبار أو المقياس.

٢- مدى ارتباط كل بند أو فقرة أو وحدة مع الاختبار ككل.

٣- التحقق من مدى الاتساق فيما بين البنود أو الفقرات أو الوحدات لكون هذا الاتساق يتأثر بمصدرين للخطأ هما:

أ) أخطاء محتوى البنود (الفقرات أو الوحدات).

ب) أخطاء عدم تجانس البنود (الفقرات أو الوحدات).

ويشكل التجانس homogeneus موضوع الاهتمام الرئيسى بالنسبة للاتساق الداخلى، حيث تستهدف طريقة الاتساق الداخلى البحث عن الدرجة التى يمكن الوثوق فيها كمؤشر لثبات الاتساق الداخلى للاختبار أو المقياس Internal-consistency reliability، حيث يعرف هذا بالثبات الداخلى لأداة القياس.

معادلات كيودر - ريتشاردسون :

وضع كيودر وريتشاردسون معادلتين ٢٠ و ٢١ واللتين حازتا على شهرة فائقة وأصبحتا تعرف بـ K-R20 وهى المعادلة الرئيسية، والمعادلة الثانية

وتعرف بـ K-R21 وهى مشتقة - مع التعديل - من المعادلة الرئيسية K-R20.

المعادلة K-R20:

وصورتها الرياضية كالتالى:

$$\text{ر}_{\text{١٠١}} = \frac{\text{ن}}{\text{ن} - ١} \left(\frac{\text{ع}^{٢} - \text{مجـ ب ك}}{\text{ع}^{٢}} \right)$$ معادلة (٢ ـ ٦)

حيث إن:

$\text{ر}_{\text{١٠١}}$ = معامل الثبات المقدر من المعادلة K-R20

$\text{ع}^{٢}$ = التبايـن الكلى لدرجات الاختبار

ب = نسبة الذيـن أجابـوا إجابات صحيحـة على كل وحـدة من وحدات الاختبار.

ك = نسبة الذين أجابوا إجابات خـاطئة على كل وحدة من وحدات الاختبار حيث ك بالنسبة لكل وحدة تساوى (١- ب).

ن = عدد بنود أو وحدات الاختبار.

مجـ ب ك = مـجموع حـاصل ضرب نسبـة الإجابات الصحـيحة × نسـبة الإجابات الخاطئة.

ويشترط لاستخدام هذه المعادلة الآتى:

١ - أن تكون كل بنـود أو وحدات أو فقرات الاختـبار متجانسـة، بمعنى أن تقيس كل فقرة أو عبارة نفس البعد أو المحور الذى تنتمى إليه.

٢- أن تكون معاملات الارتبـاط البينية للفقرات المندرجـة تحت البعد الواحد غير صفرية.

٣- أن تكون الدرجة المعطاة لكل فقرة أو عـبارة إما (١ أو صفر) حيث تمنح الدرجة (١) للإجابة الصحيحـة، والدرجة (صفر) للإجابة الخطأ. وهذا

يعنى أن معادلة كيودر - ريتشاردسون تستخدم فى حالة الاختبارات ثنائية البعد، والتى تدور الاستجابة فيها حول احتمالين اثنين فقط إما صح (✓) وإما خطأ (×).

٤- ألا تتباين معاملات صعوبة الفقرات تباينا كبيراً، فى حين يشترط فى تطبيق المعادلة K-R21 أن يكون لكل الفقرات نفس درجة الصعوبة، لأنه فى حاله تباين درجات صعوبة الفقرات تباينًا كبيراً فإن الثبات المحسوب بالمعادلة K-R21 يكون منخفضا.

مثال: (٢-٤)

طبق اختبار يتكون من ٦ فقرات على ٥٠ تلميذاً، وكانت الإجابة على الفقرات وفقاً لمقياس تقدير ثنائى (صح أو خطأ). وقد جاءت الإجابات على الاختبار كالتالى:

الوحدات	: الأولى	الثانية	الثالثة	الرابعة	الخامسة	السادسة
الإجابة الصحيحة	: ١٢	٤١	١٨	٢٩	٣٠	٤٧

والمطلوب:

حساب معامل ثبات هذا الاختبار باستخدام معادل K-R20 .

الحل:

١ - نقوم بإيجاد قيمة (ب) وذلك بقسمة عدد الإجابات الصحيحة لكل فقرة على العدد الكلى للمفحوصين الذين طبق عليهم الاختبار كالتالى:

$\frac{١٢}{٥٠} = ٠,٢٤$ ، $\frac{٤١}{٥٠} = ٠,٨٢$ ، $\frac{١٨}{٥٠} = ٠,٣٦$ وهكذا حتى الفقرة السادسة

٢ - نقوم بإيجاد قيمة (ك) لكل فقرة بطرح قيمة ب (فى الخطوة السابقة) من ١ (واحد)، فنحصل بذلك نسب الإجابات غير الصحيحة لكل فقرة كالتالى:

١ - ٠,٢٤ = ٠,٧٦

١ - ٠,٨٢ = ٠,١٨

١ - ٠,٣٦ = ٠,٦٤

. . . . . .

وهكذا حتى الفقرة السادسة

٣ - نقوم بإيجاد (ب × ك) لكل فقرة بضرب الناتج فى الخطوتين (١، ٢) كالتالى:

٠,٢٤ × ٠,٧٦ = ٠,١٨

٠,٨٢ × ٠,١٨ = ٠,١٥

٠,٣٦ × ٠,٦٤ = ٠,٢٣

. . . . . .

وهكذا حتى الفقرة السادسة

٤ - نقوم بحساب مجـ ب ك بجمع القيم المحسوبة من الخطوة ٣.

٥ - توضع النتائج السابقة فى جدول إحصائى كالتالى:

الفقرات	الإجابات الصحيحة (س)	(ب) (من الخطوة١)	ك (من الخطوة٢)	ب × ك (من الخطوة٣)
١	١٢	٠,٢٤	٠,٧٦	٠,١٨
٢	٤١	٠,٨٢	٠,١٨	٠,١٥
٣	١٨	٠,٣٦	٠,٦٤	٠,٢٣
٤	٢٩	٠,٥٨	٠,٤٢	٠,٢٤
٥	٣٠	٠,٦٠	٠,٤٠	٠,٢٤
٦	٤٧	٠,٩٤	٠,٠٦	٠,٠٦

مجـ ب × ك = ١,١٠

٦ - نقوم بحساب التباين (ع^{٢}) لدرجات الاختبار فنجد أنه يساوى ٢,٨.

٧ - مما سبق يتضح أن :

ع^{٢} = ٢,٨

مجـ ب ك = ١,١٠

ن = ٦

وبتطبيق المعادلة K-R20 معادلة : (٢ - ٦) ينتج أن:

$$\frac{٦}{٥} \left(\frac{٢,٨ - ١,١}{٢,٨} \right) = ١٠١$$

$$= ١,٢ \left(\frac{١,٧}{٢,٨} \right)$$

$$= ١,٢ \times ٠,٦٠٧$$

= ٠,٧٣

إذا معامل الثبات المقدر للاختبار باستخدام معادلة K-R20 = ٠,٧٣

مثـال : (٢ - ٥)

طبق اختبار معرفى يتكون من ١٥ فقرة يجاب عليها بصح (✓) أو خطأ (×) على عدد ١٠ مفحوصين فكانت النتائج كالتالى:

مجـ س = ٨٠

سَ = ٨

مجـ س^{٢} = ٧٩٦

مجـ ب ك = ٢,٦٢

والمطلوب :

حساب معامل ثبات هذا الاختبار باستخدام معادلة K-R20.

الحـل :

١ - نقوم أولا بحساب التباين الكلى لدرجات الاختبار بتطبيق المعادلة (٢-٤) كالتالى :

$$ع_{ت} = \frac{١}{١٠} \sqrt{١٠(٧٩٦) - (٨٠)^{٢}}$$

$$= \frac{١}{١٠} \sqrt{٧٩٦٠ - ٦٤٠٠}$$

$$= \frac{١}{١٠} \sqrt{١٥٦٠}$$

$$= \frac{١}{١٠} \times ٣٩,٤٩٧$$

$$= ٣,٩٥$$

إذا $ع_{ت}^{٢} = (٣,٩٥)^{٢} = ١٥,٦٠$

٢ - نقوم بالتعويض فى المعادلة K- R20 كالتالى :

$$ر_{١٠١} = \frac{١٥}{١٥ - ١} \left(\frac{١٥,٦٠ - ٢,٦٢}{١٥,٦٠} \right)$$

$$= \frac{١٥}{١٤} \left(\frac{١٢,٩٨}{١٥,٦٠} \right)$$

$$= ١,٠٧١ \times ٠,٧٣٢$$

$$= ٠,٨٩١$$

إذا معامل الثبات المقدر للاختبار باستخدام معادلة K-R20 = ٠,٨٩

المعادلة K-R21 :

هذه المعادلة لا تتطلب تحليل الفقرات (العبارات أو الأسئلة) كما فى المعادلة السابقة، ولكنها تتطلب المتوسط الحسابى ($\bar{X}$ - سَ) لكل الاختبار وكذا الانحراف المعيارى (ع - SD) ، والصورة الرياضية لهذه المعادلة كالتالى :

$$\text{ر}_{\text{ن ن}} = \frac{\text{ن ع}^2 - \bar{\text{س}}\ (\text{ن} - \bar{\text{س}})}{\text{ع}^2\ (\text{ن} - 1)}$$ معادلة (٢-٧)

حيث إن :

$\bar{\text{س}}$ = متوسط درجات الاختبار

ن = عدد وحدات (بنود) الاختبار

ع^2 = التباين الكلى لدرجات الاختبار

وبتطبيق هذه المعادلة على البيانات المعطاة فى المثال السابق (٢-٥) ينتج أن:

$$\text{ر}_{\text{ن ن}} = \frac{١٥ \times ١٥,٦ - ٨\ (١٥ - ٨)}{٢,٨ \times (١٥ - ١)}$$

$$= \frac{٢٣٤ - ٥٦}{١٥,٦٠ \times ١٤}$$

$$= \frac{١٧٨}{٢١٨,٤}$$

= ٠,٨١٥

إذاً معامل الثبات المقدر للاختبار باستخدام معادلة K-R21 = ٠,٨٢

مثال (٢-٦) :

لنفرض أننا طبقنا اختبار معرفى لكرة القدم يتكون من ١٠٠ عبارة (سؤال) على مجموعة من المدربين فكانت بياناتهم كالتالى:

$\bar{\text{س}}$ = ٦٥ درجة (متوسط درجات الاختبار)

ع = ١٠ درجات (الانحراف المعيارى)

والمطلوب :

حساب معامل ثبات هذا الاختبار باستخدام معادلة K-R21 .

الحل:

نقوم بالتعويض فى المعادلة (٢-٧) كالتالى:

$$\text{١٠١} = \frac{١٠٠\,(١٠)^{٢} - ٦٥\,(١٠٠ - ٦٥)}{(١٠)^{٢}\,(١٠٠ - ١)}$$

$$= \frac{١٠٠ \times ١٠٠ - ٦٥ \times ٣٥}{١٠٠ \times ٩٩}$$

$$= \frac{١٠٠٠٠ - ٢٢٧٥}{٩٩٠٠}$$

$$= \frac{٧٧٢٥}{٩٩٠٠}$$

$$= ٠,٧٨$$

معامل ألفا ∝:

فى عام ١٩٥١م قدم كرونباخ Cronbach, L.J معادلة عامة يمكن تطبيقها فى الحالات التى تكون فيها الاستجابة على الفقرة (العبارة أو السؤال) متعددة الاختيار، أى فى الحالات التى يكون للفقرة (العبارة أو السؤال) الواحدة عدد كبير من الدرجات المحتملة، حيث تكون احتمالات الإجابة ليست (صفرا) أو (واحداً) كما فى حالة الفقرات التى تتطلب الاستجابة بصح أو خطأ.

فمن المعروف أنه يوجد العديد من الاختبارات والمقاييس التى لا تتطلب الاستجابة عليها بصح أو خطأ كما فى مقاييس الشخصية والاتجاهات، ففى مقياس كينيون Kenyon للاتجاهات نحو النشاط البدنى - على سبيل المثال - يقوم كل مفحوص بالإجابة عن كل عبارة بما يتناسب مع اتجاهه نحوها طبقا لمقياس مدرج من ٥ تدريجات هى: أوافق بدرجة كبيرة، أوافق، لم أكون رأيًا بعد، أعارض، أعارض بدرجة كبيرة. حيث لا تكون هناك إجابة صحيحة وأخرى خاطئة، فالإجابة التى يختارها المفحوص تعد مؤشراً عن اتجاهه بين الموافقة وعدم الموافقة بالنسبة لموضوع ما، وبحيث تمنح

الدرجـات: ٥، ٤، ٣، ٢، ١ أو ١، ٢، ٣، ٤، ٥ للتدريجـات الخمسة بالترتيب حسب اتجاه العبارة.

ولكى يمكن تطبيـق معادلة كيودر - ريتشاردسون علـى مثل هذه المقاييس أعد كرونبـاخ معادلته المعروفة بمـعامل ألفا (α) coefficient alpha لتقدير الاتساق الداخـلى للاختبارات والمـقاييس متعـددة الاختيار، أى عنـدما تكون احتمالات الإجابة ليست صفراً أى ليست ثنائية البعد.

والصورة الرياضية لمعامل ألفا (α) كالتالى:

$$= \left(\frac{ن}{ن-١}\right)\left(١ - \frac{مجـ ع^{٢}ك}{ع^{٢}}\right)$$ معادلة (٢ - ٨)

حيث إن:

* α = معامل ألفا (معامل الثبات المقدر).

ن = عدد الفقرات.

ع٢ = تباين الدرجات على كل الاختبار.

مجـ ع٢ك = مجـموع تباين كل فقـرة من فقرات الاختـبار (من درجات المفحوصين على الفقرة).

طريقة تحليل التباين: ANOVA method

يطبق تحليل الـتباين analysis of variance (ANOVA) فى بعض الأحيان كمقيـاس لحساب ما يعرف بالثبـات النسبى relative reliability حيث يرجع الفضل فى استخدامه كأسلوب لتقدير ثبات الاختبارات والمقاييس إلى كل من جاكـسون Jackson ١٩٣٩م، وهويت Hoyt ١٩٤١م، والكسندر Alexander ١٩٤٧م. وكانت مجهودات هؤلاء العلماء قد بنيت على مسلمة رئيسية هى :

- أن الدرجة الكليـة للمفحوص على الاختـبار يمكن تقسيمـها إلى أربعة مكونات مـستقـلة هى: (أ) المكون الشـائع بين كل المـفحوصـين وكل فقرات الاختبار (ب) المكون المرتبط بفـقرات الاختبار فقط (جـ) المكو

المرتبط بالمفحوصين فقط د) مكون الخطأ وهو يرتبط بالمكونات الثلاثة السابقة. وحتى يمكننا فهم ما سبق فإنه يلزم توضيح الآتى:

١ - التباين الكلى (s^2t) : total variance

وهو يشير إلى مجموع مربعات انحرافات جميع القيم عن متوسطها الحسابى. ومن الملاحظ أن التباين الكلى لمجموعة من القياسات يتألف من مصدرين من مصادر التباين هما : التباين الحقيقى true variane وتباين الخطأ error variance، حيث يتم التعبير عن هذه العلاقة كالتالى:

الدرجة الملاحظة (التجريبية) = الدرجة الحقيقية + درجة الخطأ

$$س_م = س_ح + س_خ$$

حيث إن:

$س_م$ = الدرجة التى يتم الحصول عليها من التجربة نتيجة عملية القياس.

$س_ح$ = الدرجة الحقيقية.

$س_خ$ = درجة الخطأ (مكون الخطأ).

الدرجة الملاحظة (التجريبية) observed or experimental score

هى الدرجة التى يحصل عليها الفرد نتيجة الأداء على الاختبار، وتمتاز بأنها درجة يمكن ملاحظتها وتسجيلها (أى أنها درجة معلومة) تعبر عن الأداء العام أو الكلى للفرد.

الدرجة الحقيقة true score

وهى درجة تعبر عن الأداء الحقيقى للفرد، وتمثل جزءًا من الأداء العام أو الكلى على الاختبار (وهى درجة غير معلومة).

درجة الخطأ error score

هى درجة تعزى لأخطاء الصدفة أو للظروف الخارجية البعيدة عن موضوع الاختبار (وهى أيضا درجة غير معلومة).

أى أن:

التباين الكلى = التباين الحقيقى + تباين الخطأ

معادلة (٢ -٩)

$$ع^{٢}_{م} = ع^{٢}_{ح} + ع^{٢}_{خ}$$

ومن هذا يمكننا أن نقول إن :

$$\text{معامل الثبات} = \frac{\text{التباين الحقيقى}}{\text{التباين الكلى}}$$

معادلة (٢ - ١٠)

$$= \frac{ع^{٢}_{ح}}{ع^{٢}_{م}}$$

أو :

$$\text{معامل الثبات} = \frac{\text{التباين الحقيقى}}{\text{التباين الحقيقى + تبادين الخطأ}}$$

معادلة (٢ - ١١)

$$= \frac{ع^{٢}_{ح}}{ع^{٢}_{ح} + ع^{٢}_{خ}}$$

أى أن معامل ثبات درجات الاختبار تساوى النسبة بين التباين الحقيقى إلى التباين العام.

مثال : (٢-٧)

لنفرض أن لدينا مجموعة تتكون من ١٠ قياسات، وكان المتوسط الحسابى للدرجات الحقيقية هو ٢٥، وكان التباين الكلى للدرجات $ع^{٢}_{م}$ = ١٢٠، وتباين الدرجات الحقيقية $ع^{٢}_{ح}$ = ١٠٥ ، وتباين الخطأ $ع^{٢}_{خ}$ = ١٥,٢ .

(جدول : ٢ - ٣) التالى :

جدول (٢ - ٣)

الدرجات الملاحظة	الدرجات الحقيقية	مكونات الخطأ
٣	٥	-٢
١٧	١٥	+٢
١٦	٢٠	-٤
٢٣	٢٥	-٢
٢٧	٢٥	+٢
٢٥	٢٥	٠
٣٥	٢٥	+١٠
٢٦	٣٠	-٤
٣٣	٣٥	-٢
٤٥	٤٥	٠
مجـ س = ٢٥٠	٢٥٠	٠
سَ = ٢٥	٢٥	٠
مجـ س٢ = ١٢٠٢	١٠٥٠	١٥٢
$ع^{٢}$ = ١٢٠,٢	١٠٥	١٥,٢

من الجدول السابق يلاحظ الآتى:

بالنسبة للدرجات الملاحظة (التجريبية):

مجموع القيم مجـ س = ٢٥٠

المتوسط الحسابى سَ = ٢٥

مجموع مربعات انحرافات القيم عن وسطها الحسابى مجـ س٢ = ١٢٠٢

تباين الدرجات الملاحظة (التباين الكلى) $ع^{٢}_{م}$ = ١٢٠,٢

وبالنسبة للدرجات الحقيقية :

مجموع القيم مجـ س = ٢٥٠

المتوسط الحسابى سَ = ٢٥

مجموع مربعات انحرافات القيم عن وسطها الحسابى مجـ س٢ = ١٠٥٠

تباين الدرجات الحقيقية $ع^{٢}_{ح}$ = ١٠٥

وبالنسبة لتباين الخطأ :

مجموع القيم مجـ سَ = ٠

س = ٠

مجـ س٢ = ١٥٢

$ع^{٢}_{خ}$ = ١٥,٢

وبتطبيق المعادلة (٢ - ٩) يتضح أن التباين الكلى يساوى تباين الدرجات الحقيقية + تباين الخطأ :

١٢٠,٢ = ١٠٥ + ١٥,٢

١٢٠,٢ = ١٢٠,٢

ولقد استخدمت المعادلة (٢-٩) للتأكد من أن التباين الكلى $ع^{٢}_{م}$ الذى تم

حسابه من مجموع مربعات انحرافات القيم عن وسطها الحسابى يساوى تباين الدرجات الحقيقية ع^{٢}ح + تباين الخطأ ع^{٢}خ .

ولحساب نسبة التباين الكلى إلى التباين الحقيقى نقوم بتطبيق المعادلة (٢-١٢) التالية :

$$\frac{ع^{٢}م}{ع^{٢}م} = \frac{ع^{٢}ح}{ع^{٢}م} + \frac{ع^{٢}خ}{ع^{٢}م}$$ معادلة (٢ - ١٢)

وبالتعويض فى المعادلة السابقة (٢-١٢) يتضح الآتى :

$$\frac{١٢٠,٢}{١٢٠,٢} = \frac{١٠٥}{١٢٠,٢} + \frac{١٥,٢}{١٢٠,٢}$$

١ = ٠,٨٧٤ + ٠,١٢٦

١ = ١

ويمكن حساب ثبات مجموعة الدرجات الموضحة بالجدول (٢ - ٣) بحساب نسبة تباين الدرجات الحقيقية ع^{٢}ح مقسوماً على التباين الكلى ع^{٢}م معادلة (٢ - ١٠) كالتالى :

الثبات = $\frac{١٠٥}{١٢٠,٢}$ = ٠,٨٧

كما يمكن حساب الثبات بتطبيق المعادلة التالية (٢ - ١٣) كالتالى :

الثبات = ١ - $\frac{١٥,٢}{١٢٠,٢}$ معادلة (٢ - ١٣)

= ١ - $\frac{ع^{٢}خ}{ع^{٢}م}$

= ١ - ٠,١٢٦

= ٠,٨٧

ما هى القيم المقبولة لمعاملات الثبات؟

يتساءل معظم الباحثين عن قيم معاملات الثبات التى تجعلهم يثقون فى الأدوات التى سوف يستخدمونها فى جمع المادة العلمية فى الدراسات والبحوث التى يقومون بها، وتتوقف الإجابة على هذا التساؤل على الغرض من استخدام أداة القياس، فكثير من الباحثين يحكمون على مستوى معامل الثبات المحسوب من خلال الدلالة الإحصائية لمعامل الارتباط الذى يدل معامل الثبات، فإذا كان معامل الارتباط هذا دالاً إحصائيا كان معامل الثبات هو الآخر دالاً إحصائياً (أى مقبولاً) والعكس صحيح، إلا أن علماء القياس يرفضون هذا الأسلوب، ويؤكدون على أن القيمة المقبولة لمعامل الثبات المحسوب تتوقف أساساً على الغرض من استخدام أداة القياس (المقياس أو الاختبار)، ويتفقون فيما بينهم على ضرورة أن تكون هذه القيم كالتالى :

١- بالنسبة للاختبارات التشخيصية diagnosis واختبارات الذكاء Intelligence واختبارات الاستعدادات aptitudes واختبارات التحصيل achievement يجب ألا تقل قيم معاملات الثبات عن ٩٠ , ٠ .

٢- فى الاختبارات الإكلينيكية clinical يجب أن تكون معاملات الثبات مرتفعة جداً (تقترب من ٠٠ , ١).

٣- فى الاختبارات التى تستخدم فى اتخاذ قرارات مهمة عن مستقبل بعض الأفراد أو المؤسسات فإنه يلزم أن تكون معاملات الثبات ٩٥ , ٠ فأكثر.

٤- فى الاختبارات التى تستخدم لمقارنة المتوسطات الحسابية لمجموعتين أو أكثر من الأفراد فإنه يمكن قبول معامل الثبات الذى يساوى ٦٠ , ٠ كحد أدنى.

٤ - مفهوم الثبات المطلق وطرق حسابه

هو ثبات يعكس اتساق درجات الفرد على أساس الدرجات الملاحظة (التجريبية)، حيث يستخدم الخطأ المعيارى للقياس standard error of measurement كمؤشر لوصف ثبات الاختبار وإقرار مدى كفاية وصلاحية هذا الثبات. فالخطأ المعيارى طريقة إحصائية تستخدم لإيضاح مفهوم الثبات وملاحظة التغيرات (الاختلافات) التى تحدث فى درجات الاختبار عند تطبيقه لعدد من المرات على نفس الأفراد مع توحيد الشروط فى جميع المرات.

ويستخدم الخطأ المعيارى (SE) فى القياس لتحديد مدى الاختلاف الحادث فى درجات المقياس أو الاختبار عند تطبيقه لعدد من المرات على نفس الأفراد، فإذا طبق الاختبار ثم أعيد تطبيقه أى عدد من المرات على نفس الأفراد، مع افتراض عدم حدوث تغير فى الصفة أو القدرة أو السمة المقيسة، فإن الانحراف المعيارى(ع - SD) لتوزيع الدرجات الفعلية فى كل مرة من مرات القياس يشير إلى الخطأ المعيارى (SE) الذى يحدث فى عملية القياس أو فى درجات الاختبار الذى طبق على الأفراد.

فمعامل الثبات يدل على اتساق درجات الاختبار عندما يتم تطبيقه أكثر من مرة على نفس الأفراد، حيث يكون هذا المعامل منخفضًا فى حالة وجود اختلافات كبيرة فى الدرجات، ويكون مرتفعًا عندما يكون الاختلاف فى الدرجات محدوداً.

ويستهدف الثبات المطلق تطبيق أداة القياس عدد من المرات على نفس الفرد (المفحوص) حتى يمكن تقدير كمية التغير variation التى يمكن توقعها بالنسبة لدرجات الاختبار، هذا التقدير estimate يسمى بالخطأ المعيارى للقياس (SE).

ومن الملاحظ أن دليل الاختبارات الكتابية المقننة (الاختبارات المنشورة) يتضمن - عادة - الخطأ المعيارى للقياس كجزء مهم من شروط الاختبار، وفى مجال الاختبارات النفس - حركية يلعب الخطأ المعيارى للقياس دوراً مهمًا بالنسبة لتقدير درجة المفحوص على هذه الاختبارات.

ما الذى يعنيه الخطأ المعيارى للقياس؟

لنفترض أنه تم تطبيق اختبار الرمية الحرة فى كرة السلة free throw ٢٠ مرة على ١٠٠ لاعب، وقد حصل أحد اللاعبين على الدرجة ١٢ من المحاولات العشرين، وكان الخطأ المعيارى للقياس (SE) يساوى ٢، فما الذى يعنيه هذا الرقم بالرجوع إلى درجة هذا اللاعب والتى تساوى ١٢؟

إن هذه الدرجة تظهر كمية الخطأ amount of error التى يمكن وضعها فى الاعتبار عند تحليل الدرجة التى حصل عليها اللاعب والتى تساوى ١٢. بمعنى آخر، تدلنا الدرجة ٢ على الحدود limits التى تمتد إليها الدرجة الحقيقية للقدرة على الرمية الحرة لهذه اللاعب، فإذا كانت درجات اختبار الرمية الحرة ثابتة تماماً perfect reliabitity ، فإنه فى حالة تكرار تطبيق الاختبار على هذا اللاعب أى عدد من المرات تحت نفس الشروط، فإن حوالى ٦٨٪ من الدرجات الحقيقية سوف تقع فى مدى وحدة خطأ معيارى واحدة بالنسبة للدرجة الملاحظة (التجريبية) observed score (١٢ ± ٢) ، إن حوالى ٩٥٪ من الدرجات سوف تقع فى مدى وحدتين خطأ معيارى (١٢ ± ٤)، وأن حوالى ٩٩٪ من الدرجات سوف تقع فى مدى ثلاث وحدات خطأ معيارى (١٢ ± ٦).

وقد أخذ بهذه القاعدة وفقًا للمساحات تحت منحنى التوزيع الاحتمالى، وعلى أساس أن الدرجة الملاحظة (التجريبية) تم تحويلها إلى الدرجة ذ - z ، وأن الخطأ المعيارى (SE) يكافئ درجة واحدة من الدرجات ذ - z.

ووفقاً لهذه القاعدة فإنه يمكن حساب مستوى الثقة فى درجات الرمية الحرة للاعب فى المثال السابق على النحو التالى :

عدد وحدات الخطأ المعيارى (SE)	مستوى الثقة (L.C)	وحدات الخطأ للاختبار	مدى الدرجات (R.S)
١	٦٨٪	٢	١٠ - ١٤
٢	٩٥٪	٤	٨ - ١٦
٣	٩٩٪	٦	٦ - ١٨

وعلى هذا الأساس نستطيع أن نقرر أن الدرجة الحقيقية لهذا اللاعب فى اختبار الرمية الحرة فى كرة السلة تنحصر ما بين (٦ إلى ١٨ درجة) خلال ثلاث وحدات انحراف معيارى ومستوى ثقة ٩٩٪، فى حين أن درجة الأداء الفعلية لهذا اللاعب بالنسبة لنفس الاختبار تميل إلى أن تنحصر ما بين (١٠ إلى ١٢ درجة) .

ويمكن حساب الخطأ المعيارى للقياس (SE) باستخدام المعادلة التالية :

$$ع_{ن} = ع \sqrt{١ - ر_{١٠١}}$$

حيث إن :

$ع_{ن}$ = الخطأ المعيارى لعملية القياس (درجات الاختبار).

ع = الانحراف المعيارى لدرجات الاختبار.

$ر_{١٠١}$ = معامل ثبات درجات الاختبار.

وتستخدم هذه المعادلة لتقدير الخطأ المعيارى (SE) لعملية القياس التى تنتج من تكرار تطبيق الاختبار (القياس)، حيث يدل الخطأ المعيارى على مدى تطابق الدرجات الفعلية التجريبية Actual scores بالنسبة للدرجات الحقيقية True scores عند تطبيق الاختبار عدد من المرات.

فإذا افترضنا أن الانحراف المعيارى(ع - SD) فى المثال السابق الخاص باختبار الرمية الحرة فى كرة السلة كان يساوى ٤، وكان معامل الثبات يساوى ٠,٧٥ فإنه يمكن حساب الخطأ المعيارى للقياس بالتعويض فى معادلة الخطأ المعيارى كالتالى :

$$ع_{ن} = ٤ \sqrt{١ - ٠,٧٥}$$

$$= ٤ \sqrt{٠,٢٥}$$

$$= ٤ \times ٠,٥$$

$$= ٢$$

إذا الخطأ المعيارى للقياس يساوى فى هذه الحالة ٢ .

وإذا افترضنا أن الخطأ المعيارى للقياس كان يساوى ٣٣٣, ٠ بدلاً من ٢، فإن مستوى الثقة فى درجات اختبار الرمية الحرة فى المثال السابق يصبح كالتالى:

(١٢ ± ٣٣٣, ٠) = ١١, ٦٦٧ إلى ١٢, ٣٣٣ عند مستوى ثقة ٦٨٪.

(١٢ ± ٦٦٦, ٠) = ١١, ٣٣٤ إلى ١٢, ٦٦٦ عند مستوى ثقة ٩٥٪.

(١٢ ± ٩٩٩, ٠) = ١١, ٠٠ إلى ١٣, ٠٠ عند مستوى ثقة ٩٩٪.

حيث يلاحظ أن الدرجة تنحصر ما بين ١١و ١٣ درجة بدلاً من ٦و ١٨ درجة.

العلاقة بين الانحراف المعيارى والخطأ المعيارى :

يزداد معامل ثبات الاختبار كلما نقص الانحراف المعيارى (ع – SD) والخطأ المعيارى (ع$_{ن}$ – SE)، لهذا السبب كان لمعاملات الثبات المرتفعة أخطأ صغيرة، فى حين تتميز معاملات الثبات المنخفضة بوجود أخطاء كبيرة. فإذا كان معامل ثبات أحد الاختبارات يساوى ٨٤, ٠ ، وكان الانحراف المعيارى لدرجات الاختبار يساوى ١٠ ، فإن الخطأ المعيارى للقياس يمكن حسابه كالتالى :

$$ع_{ن} = ١٠ \sqrt{١ - ٨٤, ٠}$$

$$= ١٠ \sqrt{١٦, ٠}$$

$$= ١٠ \times ٤, ٠$$

$$= ٤, ٠٠$$

وإذا كان الانحراف المعيارى يساوى ٥ بدلاً من ١٠ ، فإن الخطأ المعيارى لعملية القياس (درجات الاختبار) يكون كالتالى:

$$ع_{ت} = ٥ \sqrt{١ - ٨٤,}$$

$$= ٥ \times ٤,٠$$

$$= ٢,٠٠$$

ويلاحظ أن الخطأ المعيارى للقياس انخفض إلى النصف ، عندما انخفض الانحراف المعيارى للدرجات، وذلك على أساس افتراض أن معامل الثبات ظل ثابتا.

مؤشر (دليل) الثبات : index of reliabitity

هو مفهوم نظرى يشير إلى العلاقة بين الدرجات الملاحظة (التجريبية) والدرجات الحقيقية التى خلصت من أخطأ القياس. ويمكن حساب مؤشر (دليل) الثبات باستخدام المعادلة التالية:

$$\text{مؤشر الثبات} = \sqrt{ر_{١٠١}}$$

حيث إن:

$ر_{١٠١}$ = معامل ثبات الاختبار (الدرجات الملاحظة أو التجريبية).

٥ - الثبات الخارجى والثبات الداخلى

هناك وجهة نظر ترى أنه يمكن تصنيف الثبات إلى نمطين عريضين هما :

الثبات الخارجى : interclass reliabitity

وهو يستخدم عندما يتم أخذ القياسات لمرتين فقط، حيث يتم حساب المعاملات الخارجية interclass coeffients باستخدام معامل الارتباط التتابعى (ppm) بطريقة بيرسون، ويتضمن هذا النمط الطرق الآتية :

١ - ثبات إعادة الاختبار (استقرار درجات الاختبار)

test-retest reliability

٢ - ثبات الصور المتكافئة (المتبادلة - المتوازية - المتساوية)
equivalent reliability

٣ - ثبات نصف الاختبار split - halves reliability

حيث يستخدم لتصحيح ثبات نصف الاختبارات المعادلات التالية :

- معادلة سبيرمان - براون.
- معادلة رولون.
- معادلة فلانجان.

الثبات الداخلى : Intraclass reliability

وهو نمط من الثبات يختلف عن سابقة فهو يستخدم عندما يتم إجراء القياسات لثلاثة محاولات متتالية أو أكثر، حيث يستخدم تحليل التباين ANOVA كأسلوب إحصائى لتقدير الثبات. ويتضمن الثبات الداخلى استخدام بعض الأساليب الإحصائية المهمة هى :

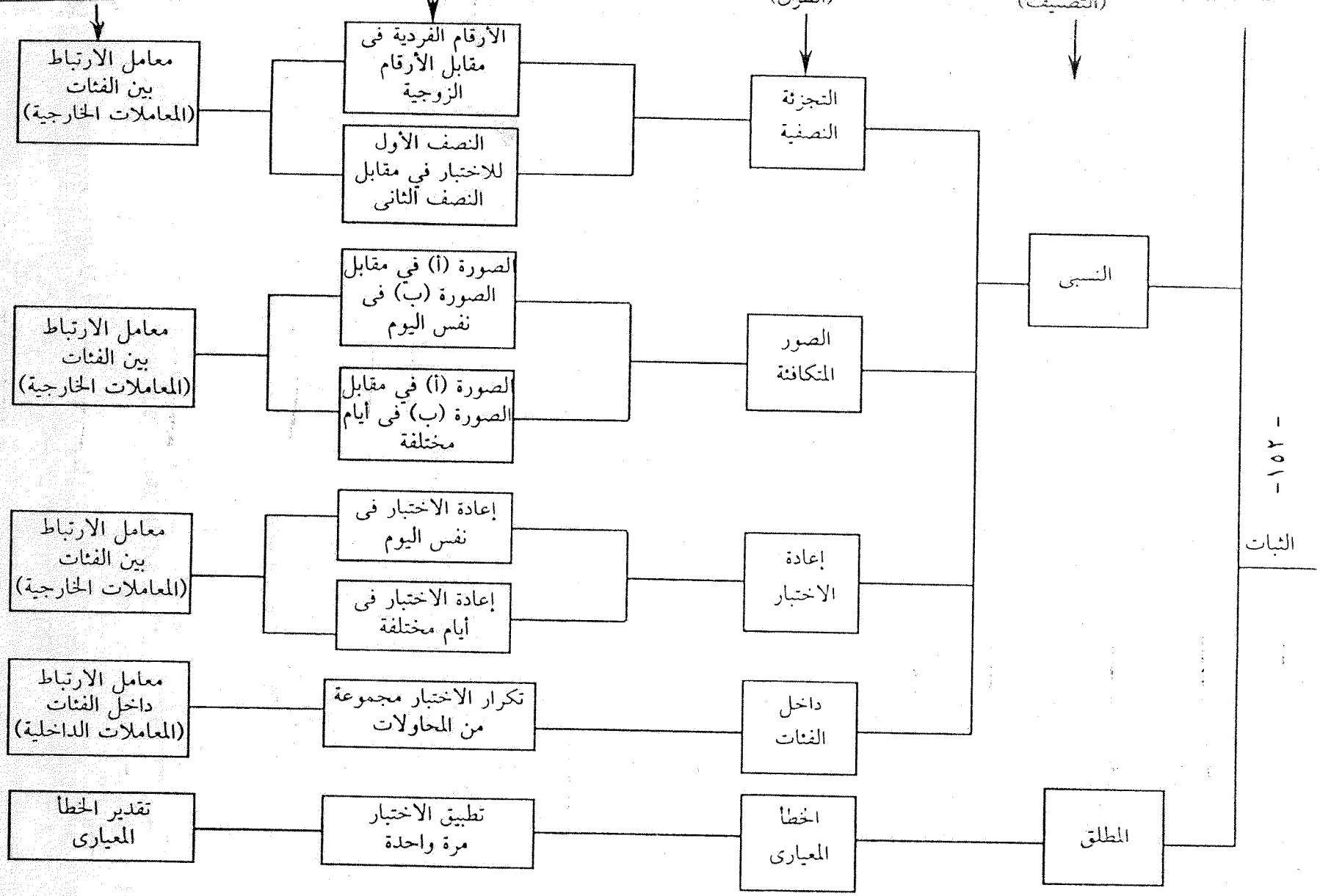
١ - معادلات كيودر - ريتشاردسون (KR21) (KR20).

٢ - معامل ألفا α Coeffient alpha.

٣ - تحليل التباين (ANOVA).

خريطة الثبات : Reliability chart

يعرض شكل (٢-١) ملخصًا لتصنيف الثبات، والطرق المستخدمة لحساب كل تصنيف وفقًا لطبيعة عملية القياس من حيث عدد مرات التطبيق، والفترات الزمنية بين مرات التطبيق المختلفة، هذا بالإضافة إلى الأساليب الإحصائية المناسبة لتقدير معاملات الثبات بالنسبة لكل طريقة على حدة.



شكل (٢ - ١) خريطة الثبات

٦- العوامل التى تؤثر على ثبات درجات الاختبار

يتأثر ثبات درجات الاختبار بعدد كبير من العوامل، هذه العوامل يمكن وضعها فى فئتين عريضتين هما:

- عوامل خارجية Ixtrinsic factors
- عوامل داخلية Intrinsic factors

العوامل الخارجية:

وهى عوامل تحدث بعيداً عن الاختبار نفسه وتؤدى إلى أن يصبح الاختبار ثابتًا أو غير ثابت، ولعل من أهم هذه العوامل ما يأتى:

١) **تشتت مجموعة الأفراد** : Group Variability

فعندما تكون مجموعة المفحوصين متجانسة homogeneous فى القدرة المزمع قياسها، فمن المرجح الحصول على معاملات ثبات منخفضة، أما إذا كانت مجموعة المفحوصين تتفاوت فيما بينها تفاوتاً كبيراً بالنسبة للقدرة المقيسة، فمن المرجح الحصول على معاملات ثبات مرتفعة.

٢) **التخمين**: Guessing

ويقصد به تخمين المفحوصين بالنسبة للعبارات (الأسئلة) الخاصة باختبارات الورقة والقلم، والذى يعد من أهم مصادر عدم الثبات، ففى حالة الاستجابات ثنائية البعد (صح أو خطأ) تكون فرصة الإجابات الصحيحة على العبارات على أساس التخمين هى ٥٠٪، فى حين ينخفض تأثير التخمين بالنسبة للعبارات التى تتطلب الاستجابة عليها الاختيار من متعدد (Mcq).

وبشكل عام فقد يؤدى التخمين إلى زيادة الدرجة الكلية للاختبار فيزيد بذلك معامل الثبات، أو يزيد من أخطاء القياس مما يؤدى إلى انخفاض معامل الثبات.

٣) الاشتراطات البيئية : Environmental Conditions

وتتمثل هذه الاشتراطات فى مدى ضبط وتنظيم بيئة الاختبار مثل: الإضاءة، الضوضاء، درجات الحرارة، واتجاهات الرياح، ونسبة الرطوبة، والزى (وبخاصة فى الاختبارات البدنية) إلى غير ذلك من العوامل التى يكون لها تأثير على ثبات درجات الاختبار.

٤) الأحداث الطارئة : Momentary Fluctuations

تؤثر بعض الأحداث الطارئة على ثبات درجات الاختبار إما بالزيادة أو بالنقصان، فالأحداث الطارئة التى تحدث للمفحوص أثناء الأداء على الاختبار قد تؤدى إلى الاضطراب والتشويش وعدم التركيز، فالأصوات العنيفة المفاجئة، وتداعى بعض المواقف والأحداث غير السارة، وضياع القلم أو كسره قبل أو أثناء الإجابة، جميعها عوامل قد تؤدى إلى قلق وتوتر المفحوص فيقع فى أخطاء تؤثر فى ثبات درجات الاختبار.

العوامل الداخلية:

وهى عوامل تختص بالاختبار نفسه وتؤثر على ثباته، ومن أهم هذه العوامل ما يأتى:

١) طول الاختبار : length of the test

يقصد بطول الاختبار عدد الوحدات (البنود - الفقرات - الأسئلة) التى يتضمنها، فكلما زاد طول الاختبار زاد بالتالى ثبات هذا الاختبار، فالاختبارات الأطول تميل إلى إظهار معاملات ثبات أعلى من الاختبارات الأقصر. وتتأسس هذه النتيجة على حقيقة علمية مؤداها أن الاختبارات الطويلة تتأثر بدرجة أقل بعوامل الصدفة والتخمين من الاختبارات القصيرة.

ولتوضيح ذلك، لنفترض أننا أردنا قياس القدرة على التهجئة Spelling ability فقمنا بسؤال التلاميذ لتهجئة كلمة واحدة فقط، فى هذه الحالة سوف نحصل على نتائج غير ثابتة، حيث يرجع ذلك إلى أن التلاميذ الذين استطاعوا الإجابة على الكلمة نجحوا، والتلاميذ الذين لم يستطيعوا الإجابة على نفس الكلمة رسبوا فى الاختبار. فإذا كانت الكلمة «صعبة» فإن معظم التلاميذ سوف يرسبون، وإذا كانت «سهلة» فإن معظم التلاميذ سوف ينجحون، وعليه فإن كلمة واحدة (عبارة أو فقرة أو سؤال) سوف يظهر عدم ثبات درجات الاختبار عندما يستخدم هذا الاختبار لتقدير قدرة التلميذ على التهجئة.

فإذا ضمنا اختبار التهجئة السابق قائمة من الكلمات المتساوية فى درجة صعوبتها، حينئذ يمكننا تقدير قدرة التلميذ على التهجئة بطريقة جيدة، وذلك لكون الدرجات التى يتم الحصول عليها من عدد كبير من الكلمات تظهر بشكل أفضل الفروق الحقيقية فى القدرة على التهجئة ومن ثم تصبح أكثر استقراراً فعن طريق زيادة حجم عينة الأسئلة (الفقرات أو البنود) التى يتضمنها الاختبار يزيد بالتالى الاتساق الداخلى لهذا الاختبار.

مما سبق يتضح أن طول الاختبار يؤثر على ثباته، حيث يتحدد هذا التأثير **من خلال نقطتين مهمتين هما:**

أ - أن العلاقة بين طول الاختبار (عدد الوحدات) وثباته علاقة طردية.

ب - أنه كلما زاد عدد الوحدات (الأسئلة أو البنود) ارتفع معامل ثبات الاختبار. ويستخدم لتقدير معامل ثبات الاختبار بعد زيادة طوله معادلة سبيرمان - براون التالية :

$$\text{ر}_{\text{ن ن}} = \frac{\text{ن} \times \text{ر}}{1 + (\text{ن} - 1)\,\text{ر}} \qquad \text{معادلة (٢ - ١٤)}$$

حيث إن :

$\rho_{1\cdot1}$ = معامل ثبات الاختبار بعد زيادة طوله (عدد وحداته)

ρ = معامل ثبات الاختبار قبل زيادة طوله (عدد وحداته)

ن = عدد مرات زيادة الاختبار (نسبة عدد الوحدات بعد الزيادة إليها قبل الزيادة)

مثال : (٢ - ٨)

طبق اختبار للذكاء يتكون من ١٠٠ سؤال (عبارة أو فقرة)، وكان معامل ثبات هذا الاختبار ٠,٨٠ ، فإذا زاد هذا الاختبار ٤ مرات بالنسبة لطوله، فكم يكون معامل ثبات الاختبار بعد زيادة طوله؟

الحل :

$$\rho_{1\cdot1} = \frac{٤ \times ٠,٨٠}{١ + (٤ - ١) ٠,٨٠} \quad \text{معادلة (٢ - ١٤)}$$

$$= \frac{٣,٢}{١+٣\times٠,٨٠}$$

$$= \frac{٣,٢}{١ + ٢,٤}$$

$$= \frac{٣,٢}{٣,٤}$$

$$= ٠,٩٤$$

يلاحظ من المثال السابق أن زيادة طول الاختبار ٤ مرات بالنسبة لطوله الأصلي زاد من معامل ثبات هذا الاختبار - عندما طبق على نفس المفحوصين - من ٠,٨٠ إلى ٠,٩٤ .

مثـــال : (٢ - ٩)

طبق اختبار للذكاء يتكون من ١٠٠ سؤال وكان معامل ثباته ٠,٨٠ ، فكم يكون معامل ثبات هذا الاختبار إذا أصبح يتكون من ٤٠٠ سؤال ؟

الحـــل:

١ - نقوم بحساب (ن) والتى تمثل النسبة بين عدد الوحدات بعد الزيادة إلى عدد الوحدات قبل الزيادة كالتالى :

$$\text{ن} = \frac{٤٠٠}{١٠٠} = ٤$$

٢ - نقوم بالتعويض فى معادلة سبيرمان - براون (معادلة : ١٤) كالتالى:

$$\frac{٤ \times ٠,٨٠}{١ + (٤ - ١) ٠,٨٠} = ١٠١ \qquad \text{معادلة (٢ - ١٤)}$$

$$= \frac{٣,٢}{٣,٤}$$

$$= ٠,٩٤ \qquad \text{وهى نفس النتيجة فى المثال السابق.}$$

مثال : (٢ - ١٠)

طبق اختبار معرفى للثقافة الصحية يتكون من ٥٠ سؤلاً (فقرة أو بند) وكان معامل ثبات هذا الاختبار ٠,٧٥ فكم يكون معامل ثباته إذا أصبح عدد وحداته ١٥٠ سؤالاً ؟

١ - نقوم أولاً بحساب (ن) كالتالى:

$$\text{ن} = \frac{١٥٠}{٥٠} = ٣$$

٢ - نقوم بتطبيق (معادلة : ١٤) كالتالى:

$$\frac{\text{ل}}{\text{ل}_{1.1}} = \frac{٣ \times ٠,٧٥}{١ + (١-٣)\ ٠,٧٥} \quad \text{معادلة (٢ - ١٤)}$$

$$= \frac{٢,٢٥}{١+٢\times٠,٧٥}$$

$$= \frac{٢,٢٥}{١ + ١,٥} = \frac{٢,٢٥}{٢,٥٠}$$

$$= ٠,٩٠$$

مثال : (٢ - ١١)

طبق اختبار للثقافة الصحية معامل ثباته ٠,٧٥ عندما كان عدد وحداته ٥٠ سؤلاً (فقرة أو بند) ، فكم يصبح معامل ثبات هذا الاختبار إذا أضيف إلى عدد وحداته ١٠٠ سؤال؟

الحل:

١ - نقوم أولاً بحساب عدد الوحدات بعد الزيادة كالتالى:

عدد الوحدات بعد الزيادة = ٥٠ + ١٠٠ = ١٥٠

٢ - نقوم بحساب (ن) كالتالى:

$$\text{ن} = \frac{١٥٠}{٥٠} = ٣$$

٣ - نقوم بالتعويض فى المعادلة (٢ - ١٤) كالتالى :

$$\frac{\text{٣} \times \text{٠,٧٥}}{\text{١} + (\text{٣}-\text{١})\ \text{٠,٧٥}} = \sqrt{\text{١٠١}}$$

$$= \text{٠,٩٠} \qquad \text{(وهى نفس النتيجة السابقة)}$$

مثال : (٢ - ١٢)

طبق اختبار للثقافة الصحية يتكون من ٥٠ سؤلاً ومعامل ثباته ٠,٧٥ ، والمطلوب أن يكون معامل ثبات هذا الاختبار ٠,٩٠ ، فكم يكون عدد وحداته؟

الحل :

بالتعويض فى المعادلة الأساسية (٢ - ١٤) ينتج أن :

$$\sqrt{\text{١٠١}} = \frac{\text{ن} \times \sqrt{\quad}}{\sqrt{(\text{ن}-\text{١})} + \text{١}}$$

$$\text{٠,٩٠} = \frac{\text{ن} \times \text{٠,٧٥}}{\text{١} + (\text{ن}-\text{١})\ \text{٠,٧٥}}$$

$$\text{٠,٧٥ ن} \times \text{٠,٩٠} = \text{٠,٧٥} [\text{١} + (\text{ن} - \text{١})] \text{٠,٩٠}$$

$$\text{٠,٧٥ ن} = \text{٠,٩٠} + (\text{٠,٩٠ ن} - \text{٠,٩٠})\text{٠,٧٥}$$

$$\text{٠,٧٥ ن} = \text{٠,٩٠} + (\text{٠,٦٨ ن} - \text{٠,٦٨})$$

$$\text{٠,٧٥ ن} = \text{٠,٢٢} + \text{٠,٦٨ ن}$$

$$\text{٠,٧٥ ن} - \text{٠,٦٨ ن} = \text{٠,٢٢}$$

$$\text{٠,٠٧ ن} = \text{٠,٢٢}$$

ن = $\frac{٢٢}{٧}$ = ٣,١٤ = ٣ مرات تقريبا حيث تستبعد الكسور الأقل من ٥, ٠

مما سبق يتضح أنه لكى نرفع معامل ثبات اختبار الثقافة الصحية من ٧٥, ٠ إلى ٩٠, ٠ ، فإنه يجب أن نزيد عدد أسئلته من ٥٠ سؤلاً إلى ١٥٠ سؤلاً، حيث (ن = ٣) أى ٣ × ٥٠ = ١٥٠

ولتسهيل العمليات الحسابية فى المثال (٢ - ١٢) يمكن تطبيق المعادلة التالية لحساب قيمة (ن) بطريقة مباشرة، وهذه المعادلة هى :

$$\text{ن} = \frac{\text{معامل الثبات المطلوب} \times (١ - \text{المعامل الحالى})}{\text{المعامل الحالى} \times (١ - \text{المعامل المطلوب})}$$ معادلة (٢ - ١٥)

وبتطبيق هذه المعادلة على المعطيات فى المثال (٢ - ١٢) ينتج أن:

$$= \frac{٩٠, \times (١ - ٧٥, ٠)}{٧٥, \times (١ - ٩٠, ٠)}$$

$$= \frac{٩٠, \times ٢٥, ٠}{٧٥, \times ١, ٠}$$

$$= \frac{٢٢٥, ٠}{٧٥, ٠}$$

إذاً ن = ٣ (وهى نفس النتيجة السابقة تقريبا)

٢) درجة صعوبة الفقرات : Difficulty value of items

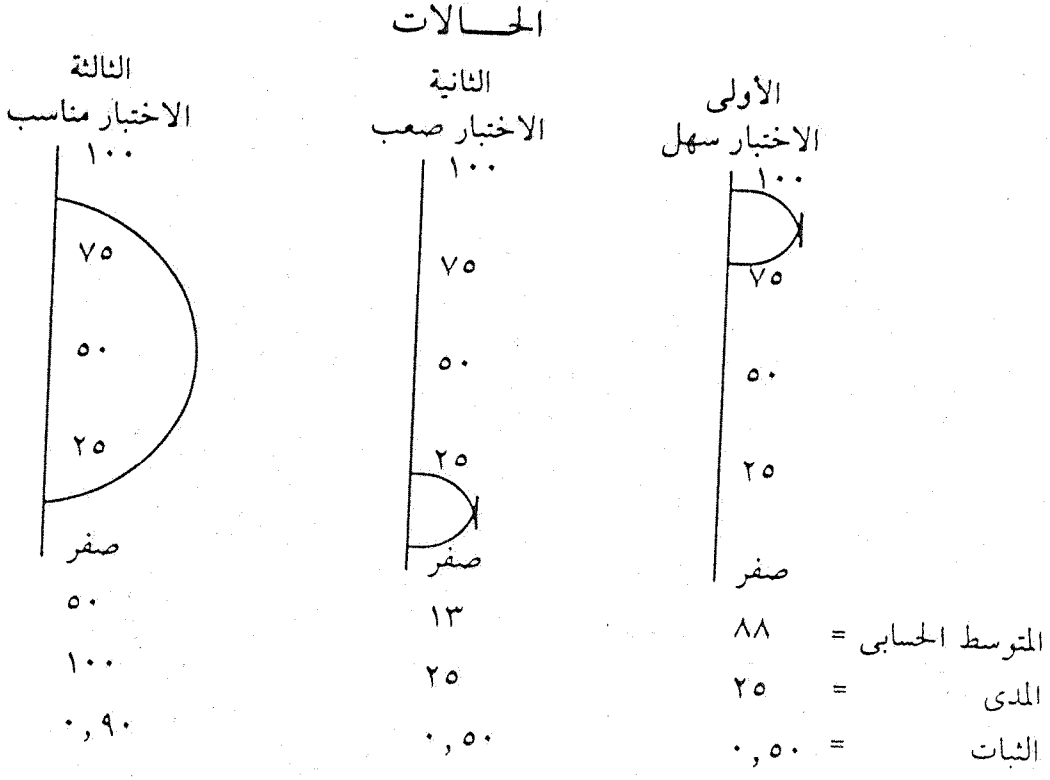
من المعروف أن درجة المفحوص على الاختبار أو المقياس هى مجموع أو محصلة درجاته على فقراته، فالفقرة هى الوحدة التى تكسب المقياس خصائصه، حيث تتوقف الخصائص الإحصائية للمقياس على المعالم الإحصائية لفقراته، فقد تبين أن عدد ومستوى صعوبة الفقرات يؤثر على ثبات وصدق المقياس أو الاختبار.

فالفقرات التى لـها مؤشر صعوبة يساوى ٥, ٠ يكون لها مـعامل ثبات أعلى، من الفقرات التى لهـا مؤشرات صعوبة أكبر من ٥, ٠ ، كما أن الـفقرات السهلة جداً والفقرات الـصعبة جداً تؤدى إلى معـاملات ثبات منخفضة، ويـبين الشكل (٢- ٢) مقارنة افتراضية لتوزيعات درجات اخـتبار للتحصيل الدراسى يتكون من ١٠٠ فقرة ومعاملات ثبات درجات هذا الاخـتبار عندما طبق على ثلاث عينات مختلفة من التلاميذ، حيث يتضح الآتى :

الحالة الأولى : عندمـا بلغ معامـل الصعوبـة ٢٥, ٠ ، نجد أن معامـل ثبات الاختبار بلغ ٥٠, ٠ .

الحالة الثـانية : عندمـا بلغ مـعامل الصعوبة ٧٥, ٠ نجد أن أن معـامل الثبات بلغ ٥٠, ٠ .

الحالة الثالثة : عندما بلغ معامل الصعوبة ٥, ٠ نجد أن معامل الثبات بلغ ٩٠, ٠ .



شكل (٢ - ٢)

٣) مدى الدرجات : Range of scores

عندما يكون مدى الدرجات الكلية للاختبار ضيقا (محدوداً) يقل الانتشار (التشتت) Variability فينخفض بذلك معامل الثبات، ففى الشكل السابق (٢ - ٢) يظهر أن مدى الدرجات يساوى فى الحالة الأولى ٢٥ (١٠٠ - ٧٥=٢٥)، ويساوى فى الحالة الثالثة حاولى ١٠٠ (١٠٠ - ٠ =١٠٠)، وعليه يظهر أنه كلما كان انتشار درجات الاختبار كبيرا كان معامل ثبات هذه الدرجات هو الآخر كبيراً، والعكس صحيح.

ومن ناحية يمكننا القول أنه كلما كان الانحراف المعيارى (ع -SD) للدرجات الكلية للاختبار مرتفعا كان معامل الثبات هو الآخر مرتفعا، وكلما كان الانحراف المعيارى منخفضا كان معامل الثبات هو الآخر منخفضا.

٤) تجانس الفقرات : Homogeneity of items

يعد تجانس الفقرات أحد أهم العوامل التى تؤثر على الثبات، حيث يقصد بالتجانس هنا الآتى:

- معامل ثبات الفقرة (معاملات ارتباط الفقرة مع غيرها)
- السمات والخصائص والقدرات التى تقيسها كل فقرة (وظائف الفقرة).

فعندما تقيس الفقرات وظائف مختلفة وتكون معاملات الارتباط البينية فيما بين الفقرات Intercorrelations صفرية (أو قريبة من الصفر) فإن معامل الثبات فى هذه الحالة يظهر منخفضًا جداً (صفر أو قريب منه) وذلك بسبب عدم تجانس الفقرات Hetrogeneous فى الوظائف التى تقيسها، أما إذا كانت الفقرات تقيس نفس الوظائف (نفس السمة أو القدرة) حينئذ تظهر معاملات الارتباط البينية (فيما بين الفقرات) مرتفعة، وبالتالى يظهر ثبات الاختبار هو الآخر مرتفعًا.

٥) القيمة التميزية للفقرة : Discriminative value

ويقصد بها الارتباط بين كل فقرة من فقرات الاختبار والدرجة الكلية له، وتعرف هذه العملية بتحليل الفقرة item analysis والتى تبين الارتباط بين درجة الأداء على الفقرة والدرجة الكلية للاختبار. فعندما يكون الارتباط بين درجة الفقرة والدرجة الكلية للاختبار منخفضاً، فإن ذلك يعنى أن الفقرة تقيس شيء آخر مختلف عما تقيسة الفقرات الأخرى للاختبار، كما قد يعنى هذا الارتباط أيضا أن الفقرة كانت سهلة جداً أو صعبة جداً حيث لم تقدم لنا فروقاً ظاهرة فيما بين استجابات المفحوصين عليها، وعليه يظهر معامل ثبات الاختبار منخفضاً حيث يفضل فى مثل هذه الحالة حذف مثل هذه الفقرة، أما إذا كان الارتباط بين الفقرة والدرجة الكلية للاختبار مرتفعاً فإن معامل ثبات الاختبار يصبح هو الآخر مرتفعاً.

٧ - التصحيح للتنقية

تستخدم معاملات الثبات للاختبارات والمقاييس المختلفة لتصحيح الارتباط بين أى مقياسين انطلاقاً من أن الارتباط بين أى مقياسين لا يخلو من الخطأ الذى حدث أثناء عمليات القياس.

وتستخدم عملية التصحيح للتنقية Correction for Attenuation كطريقة لتصحيح معاملات الارتباط بين الاختبارات من أخطاء القياس Measurement errors ، حيث تستلزم هذه الطريقة توافر الآتى:

١ - معاملات ثبات الاختبارين (كل اختبار على حدة).

٢ - معامل الارتباط بين الاختبارين.

ويستخدم لذلك المعادلة التالية:

$$r = \frac{r_{2.1}}{\sqrt{r_{1.1} \times r_{2.2}}}$$ معادلة (٢ - ١٦)

حيث إن :

r = الارتباط الحقيقى بين الاختبارين (١ ,٢)

$r_{2.1}$ = الارتباط الملاحظ (التجريبى) بين الاختبارين (١ ,٢)

$r_{1.1}$ = معامل ثبات الاختبار الأول (١)

$r_{2.2}$ = معامل ثبات الاختبار الثانى (٢)

مثال : (٢ - ١٣)

حسب الارتباط بين اختبار للتوافق بين العين وحركة اليد واختبار للمهارة الرياضية العامة general athletic skill فكان هذا الارتباط ٠,٣٤ ، فإذا كان معامل ثبات الاختبار الأول ٠,٧٥ ، ومعامل ثبات الاختبار الثانى ٠,٨٢ ، فما هو الارتباط الحقيقى (الخالى من أخطاء القياس) بين الاختبارين.

الحل :

نقوم بتطبيق المعادلة (٢ - ١٦) كالتالى :

$$r = \frac{٠,٣٤}{\sqrt{٠,٧٥ \times ٠,٨٢}}$$

$$= \frac{٠,٣٤}{\sqrt{٠,٦١٥}} = \frac{٠,٣٤}{٠,٧٨} = ٠,٤٤$$

يتبين من المثال السابق أن الارتباط بين الاختبارين قد زاد من ٠,٣٤ إلى ٠,٤٤ عندما استخدمت طريقة التصحيح.

كيف نحسن من ثبات درجات الاختبار؟

يمكن تحسين ثبات درجات الاختبار عن طريق ضبط جميع العوامل الخارجية والداخلية التى توثر على الثبات والتى سبق الإشارة إليها، ولضمان الحصول على معاملات ثبات جيدة فإنه يجب مراعاة الشروط التالية:

١ - يجب أن تكون عينة المفحوصين غير متجانسة بالنسبة للقدرة أو السمة أو الخاصية التى يقيسها الاختبار.

٢ - يجب أن تكون وحدات (الأسئلة أو الفقرات) متجانسة.

٣ - يفضل أن يكون الاختبار طويلا.

٤ - يفضل أن تكون وحدات الاختبار متوسطة الصعوبة، بمعنى أن تتراوح مؤشرات الصعوبة من ٠,٤٠ إلى ٠,٦٠

٥ - أن تتمتع الفقرات بالقدرة على التمييز.

تأثير العمر الزمنى على الثبات:

أظهرت التجارب العلمية المختلفة أن ثبات درجات الاختبار يزداد كلما تقدم السن (النضج)، وأنه توجد فروقاً بين البنين والبنات لصالح البنين، فقد أظهرت نتائج التجارب التى أجريت لحساب ثبات اختبار ستانفورد بينية Stanford- Binet scale لنسبة الذكاء IQ عن طريق إعادة اختبار معاملات الثبات التالية:

فئات السن	معاملات الثبات بنين	بنات
٦ - ١٠ سنوات	٠,٧٧	٠,٧٤
٧ - ١٠ سنوات	٠,٧٩	٠,٧٦
٨ - ١٠ سنوات	٠,٨٩	٠,٨٣
٩ - ١٤ سنة (للجنسين)	٠,٨٢	

كما أظهرت التجارب أن طريقة إعادة الاختبار هى أكثر طرق حساب الثبات التى تتأثر بالسن (النضج)، وأنه كلما تقدم السن مع توحيد الفترة الزمنية بين التطبيقين الأول والثانى زاد الثبات، فقد سجل فريمان Freeman ١٩٦٥م. لاختبار ستانفورد بينية معاملات الثبات التالية:

السن عند بدء الاختبار	معامل الثبات عن طريق إعادة الاختبار
٤ سنوات	٠,٧٣
٩ سنوات	٠,٨٧
١١ سنة	٠,٩٢

ومن ناحية أخرى أظهرت التجارب العلمية أن طول الفترة الزمنية بين التطبيقين الأول والثانى يؤدى إلى انخفاض معاملات الثبات بالنسبة لاختبارات الذكاء والاستعدادات العقلية، فقد حصل «فريمان» ١٩٦٥م على معاملات الثبات التالية:

الفترة الزمنية	معامل الثبات
- إعادة الاختبار بعد يوم واحد	٠,٩٠ - ٠,٩٥
- إعادة الاختبار بعد سنة	٠,٨٥
- إعادة الاختبار بعد سنتين ونصف	٠,٨٠
- إعادة الاختبار بعد تسع سنوات	٠,٧٨

٨ - علاقة الموضوعية بالثبات

يرى ستراند وويلسون Strand and wilson ١٩٩٣م أن الموضوعية Objectivity صورة (شكل أو صيغة) Form من صور الثبات، وهى تعرف بالثبات المقدر Rater Objectivity وهو ثبات يحدث عندما يقوم شخصان أو أكثر بتطبيق نفس الاختبار على نفس الأفراد ثم يحصلون على نفس النتائج.

ويمكننا القول بأن موضوعية الاختبار تتأثر تأثرا كبيرا عندما تختلف درجات المحكمين Judges عندما يقومون بتقدير درجة موضوع واحد، حيث تصبح الدرجات التى يتم الحصول عليها من تطبيق الاختبار لا يعتد بها كدرجات تتمتع بالثبات والصدق. فالفروق فى النتائج تحدث عندما يستخدم المحكمون أساليب مختلفة للحكم أو التقويم أو تقدير الدرجة، وأيضا عندما لا يلتزمون بإجراءات وتعليمات تطبيق الاختبار.

مفهوم الموضوعية:

تشير الموضوعية إلى مدى إمكانية الحصول على درجة صحيحة عندما يقوم فردين أو أكثر بتطبيق الاختبار على نفس الأفراد فى مناسبتين (موقفين) مختلفين. والموضوعية تشبه الثبات ولكنها تستلزم قيام اثنين أو أكثر بتطبيق نفس الاختبار وتقدير درجاته.

ويتفق علماء القياس على أن الموضوعية Objectivity عكس الذاتية Subjectivity، فمن المعروف أن المعلم فى المدرسة يعتمد على أسلوب التقدير الذاتى كأسلوب سهل ومعتاد - فى معظم نظم التعليم - لتقدير درجات الاختبارات (الامتحانات) وبخاصة التى تستخدم الورقة والقلم.

وقد لوحظ أنه عندما يقوم اثنين أو أكثر من المعلمين بتصحيح نفس ورقة الإجابة، فإنه يظهر تباين فى تقديرهم للدرجة، حيث يعزى هذا التباين إلى الفروق فيما بين هؤلاء المعلمين بالنسبة لتقديرهم لدرجة الامتحان، لهذا السبب تتجه البحوث العلمية إلى استخدام أساليب الاختبارات الموضوعية بالنسبة لاختبارات الورقة والقلم بدلاً من الأساليب التقديرية.

مما سبق يتضح أن مصطلح الموضوعية يعنى: موضوعية درجة أداة القياس عندما يقوم شخصان (محكمان) أو أكثر بتقدير درجة أى اختبار يطبق على مفحوص واحد فى أوقات مختلفة، حيث تظهر هذه الموضوعية فى مدى اتساق الدرجات التى يتم الحصول عليها، وأنه عندما تختلف الدرجات التى

يضعها المحكمان اختلافات كبيرة، فإن عملية القياس تصبح أقل موضوعية. وكقاعدة عامة: فإن الاختبار الأكثر موضوعية هو الاختبار الأكثر ثباتاً وصدقاً كوسيلة أو أداة للقياس.

وتعد اختبارات الصواب - الخطأ، والاختيار من متعدد ، نماذج للاختبارات الكتابية (الورقة والقلم) التى تتمتع بدرجة عالية من الموضوعية، بينما يتضح أن اختبارات المقال واختبارات الإجابات القصيرة - Short answer نماذج لاختبارات أقل موضوعية لتأثرها بدرجة أكبر بتحيز القائم على تقدير الدرجة (المحكم).

وفى المجال الرياضى يوجد العديد من وسائل القياس التى تتأثر برأى مقدر الدرجة (المحكم) حتى فى حالات استخدام مقاييس التقدير Rating scales لتقويم الأداء فى بعض الأنشطة كالغطس Diving ، والجمباز Gymnastic وبخاصة الجمباز الأرضى، وفى حالات تقويم القوام Posture ، والأداء فى مواقف اللعب الحقيقية أثناء المنافسات الرياضية وغيرها، حيث يصعب الاتفاق فيما بين العديد من مقدرى الدرجة أو القائمين على تنفيذ الاختبار (المحكمين)، وربما يرجع ذلك إلى عدم فهم المحكمين لطرق وأساليب تقدير الدرجة، مما يجعلهم يلجأون إلى التقديرات الذاتية لتقويم الأداء والتى عادة ما تتأثر بميول واتجاهات هؤلاء المحكمين، فكلما كان هناك قصور فى فهم تعليمات الاختبار، وشروط تطبيقه، كلما أدى ذلك إلى انخفاض الثبات نتيجة عدم اتفاق القائمين على التنفيذ (المحكمين) بالنسبة لتقدير الدرجة على الاختبار. ومن ثم فإنه يلزم - للتغلب على هذه المشكلة - الاهتمام بإعداد تعلميات الاختبارات بشكل محدد وواضح (مقنن) بغية الحصول على معاملات موضوعية وثبات وصدق مرتفعة.

تعريف الموضوعية:

تعرف الموضوعية على أنها «درجة الاتفاق فيما بين مقدِّرى الدرجة».

مفهوم الاختبار الموضوعى:

الاختبار الموضوعى objective test هو أى اختبار تمَّ إعداده بطريقة تضمن أقل درجة من التحيز bias أو الأحكام الذاتية لمقدر الدرجة على الاختبار (المحكم). وكلما كانت إجراءات تقدير الدرجة واضحة ومحددة تحديداً دقيقا، زادت موضوعية الدرجة التى يحصل عليها المفحوص.

كيفية حساب الموضوعية:

يتم تقدير الموضوعية - بشكل عام - وفق الأساليب المتبعة فى تقدير الثبات، حيث تستهدف عملية القياس توفير مجموعتين من الدرجات لكل مفحوص، ثم يستخدم أسلوب الارتباط المناسب لتقدير العلاقة بين هاتين المجموعتين من الدرجات.

ويسفر قياس العلاقة بين مجموعتى الدرجات عن معامل ارتباط يعرف باسم معامل الموضوعية objectivity coefficient وهو معامل ليس له رمز رياضى يميزه ولكنه يكون بديلا لمعامل الارتباط الذى استخدم لحسابه.

ولحساب موضوعية الاختبار (الثبات المقدر) يقوم اثنان من المحكمين بتقدير درجة الاختبار، أو الحكم على الأداء لعدد من المفحوصين (الطلاب أو اللاعبين) فى نفس الوقت، بحيث يكون كل محكم بعيداً عن الآخر حتى لا يتأثر برأيه، ثم يلى ذلك حساب معامل الارتباط بين تقديرات المحكمين، حيث يدل معامل الارتباط المحسوب على درجة الاتفاق فيما بين تقديرات المحكمين الاثنين.

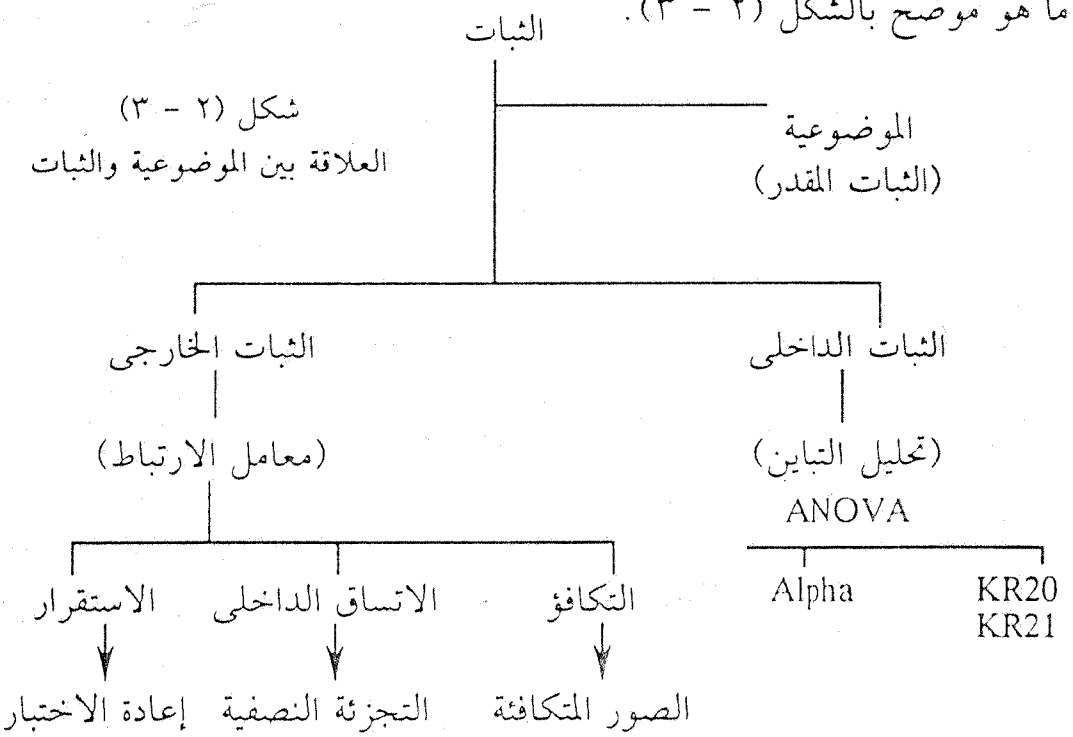
وهناك طريقة أخرى لحساب الموضوعية، إلا أنها طريقة محدودة الاستخدام، وهى طريقة تعتمد على تطبيق الاختبار على نفس مجموعة الأفراد فى موقفين مختلفين، حيث يقوم بتقدير الدرجة فى كل موقف محكم مختلف، يلى ذلك حساب الارتباط بين درجات كل محكم، حيث يلاحظ

أن هذا الأسلوب يقترب من طرق حساب الثبات، ومن ثم نجد أنه أسلوب محدود الاستخدام حيث يغنى عنه تقدير الثبات.

وقد يستخدم تحليل التباين ANOVA لحساب الموضوعية كما هو متبع بالنسبة للثبات، وذلك فى الحالات التى يستخدم فيها تقديرات ثلاثة محكمين أو أكثر. ويتم تقويم موضوعية الاختبارات والمقاييس فى المجال الرياضى وفقا للمعايير التى ذكرها كير كندال وآخرون ١٩٨٧ وهى :

معامل الارتباط (الموضوعية)	مستوى الموضوعية
٠,٩٥ إلى ١,٠٠	ممتاز
٠,٨٥ إلى ٠,٩٤	مرتفع
٠,٧٠ إلى ٠,٨٤	متوسط
٠,٠٠ إلى ٠,٦٩	غير مقبول

ويحدد بومجارتنر وجاكسون ١٩٧٥ م العلاقة بين الموضوعية والثبات وفق ما هو موضح بالشكل (٢ - ٣).



شكل (٢ - ٣)
العلاقة بين الموضوعية والثبات

معجم إنجليزى - عربى للمصطلحات

A

Analysis of variance methocl

طريقة تحليل التباين

Alternate - Forms reliability

معامل ثبات الصور المتبادلة

B

Behavioral disposions

النزوع السلوكى

C

Correlation For attenuation

التصحيح للتنفية

Complex traits

السمات المعقدة

Carry - over effect

انتقال تأثير الاداء

Coefficient of equivalence

معامل التكافؤ

Comparable - forms reliability

معامل ثبات الصور المتساوية

Clinical tests

الاختبارت الاكلينيكية

D

fficulty value of items

درجة صعوبة الفقرات

Discriminative value

القيمة التميزية للفقرة

E

Equivalent - forms method

طريقة الصور المتكافئة

Error score

درجة الخطأ

Environmental conditions

الاشتراطات البيئية

G

Guessing

التخمين

Group Variability

تشتت مجموعة الافراد

H

Human behavior

سلوك الإنسان

Heterogeneous

عدم التجانس

Hemogeneus

التجانس

Homogeneity of items

تجانس الفقرات

I

Internal Consistancy method

طريق الاتساق الداخلى

Internal - Consistency reliabitity

ثبات الاتساق الداخلى

ixtrinsic Factors

عوامل خارجية

Intrinsic factors عوامل داخلية

Item analysis تحليل الفقرة

Intercorrelation الارتباطات البينية

M

Measurement errors أخطاء القياس

Measuring instrument أداة القياس

Motor response الاستجابة الحركية

Momentory fluctuations الاحداث الطارئة

O

Observed score الدرجة الملاحظة (التجريبية)

P

Psychometrics القياس النفسى

Product - moment correlation معامل الارتباط التتابعى

Parallel forms الصور المتكافئة - المتوازية

Parallel - forms reliability معامل ثبات الصور المتوازية

Performance tests الاختبارات العملية (الادائية)

Paper-and pencil tests اختبارات الورقة والقلم

R

Range of scores مدى الدرجات

Reliability theory نظرية الثبات

Reliabitity Coefficients معاملات الثبات

Reliability of measurements ثبات القياس

Retesting coefficient معامل اعادة الاختبار

Retest reliability ثبات الاختبار عن طريق اعادته

Relative reliability بات النسبى

S

Simple qualities الخصائص البسيطة

Sampling error خطأ المعاينة

Source of error مصادر الخطأ

Split - Half method طريقة التجزئة النصفية

Stability Coefficient معامل الاستقرار

Speed tests اختبارات السرعة

Spelling ability القدرة على التهجئة

T

True variance التباين الحقيقى

total variance التباين العام

Time Sampling المعاينة الزمنية

True score الدرجة الحقيقية

V

Verbal response الاستجابة اللفظية (اللغوية)

Variability الانتشار (التشتت)

مراجع الفصل الثانى

١ - سعد عبدالرحمن (١٩٩٨) : القياس النفسى - النظرية والتطبيق. دار الفكر العربى، القاهرة.

٢ - محمد حسن علاوى، ومحمد نصر الدين رضوان (٢٠٠١) : اختبارات الأداء الحركى . دار الفكر العربى، القاهرة.

٣ - محمد حسن علاوى، ومحمد نصر الدين رضوان (٢٠٠٠) : القياس فى التربية الرياضية وعلم النفس الرياضى. دار الفكر العربى، القاهرة

٤ - محمد حسن علاوى، ومحمد نصر الدين رضوان (١٩٨٧) : الاختبارات المهارية والنفسية فى المجال الرياضى . دار الفكر العربى، القاهرة.

٥ - محمد نصر الدين رضوان (١٩٩٧) : المرجع فى القياسات الجسمية، دار الفكر العربى، القاهرة.

6 - Baumgartner, T.A. and Jakson, A.S. (1982) : Measurement for evaluation in physical education. 2 n d ed., Dubuque , Iowa : Wm.C.Brown.

7 - Jackson, A., Jackson, A.S., and Bell,J. (1980) : A Comparison of alpha and the intraclass reliability Coefficients. R.Q., 51,568-571.

8 - Safrit, M.J., and wood, T.M. (1989) : Measurement Concepts in physical education and execise science . champaign, Ill. : Human Kinetics.

9 - Sax, G. (1980) : principles of educational and psychological measurement and evaluation. Belmont, C. A : Wadsworth.

10 - Verducci, F.M. (1980) : Measurement Concepts in physical education. st. Louis : C.V. Mosby.

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
- طريقة جداول التوقع

٥ - العوامل التى تؤثر على صدق الاختبار

٦ - العلاقة بين صدق الاختبار وثباته

- معجم إنجليزى - عربى للمصطلحات
- مراجع الفصل الثالث

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 الشىء 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 حا العينات العشوائية). ولكى تظهر هذه الفروق بشكل واضح ومميز فإ يلزم استخدام وسائل قياس حساسة حتى تظهر هذه الفروق، فعلى سب المثال عند استخدام اختبار العدو ٣٠ مترا كمقياس لسرعة الانتقال، ف يلزم تقدير الزمن باستخدام ساعات إيقاف تحسب الزمن بالثوانى (و $\frac{١}{١٠٠}$ وجزء من مائة من الثانية، حتى تتضح الفروق بين المختب

(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 .

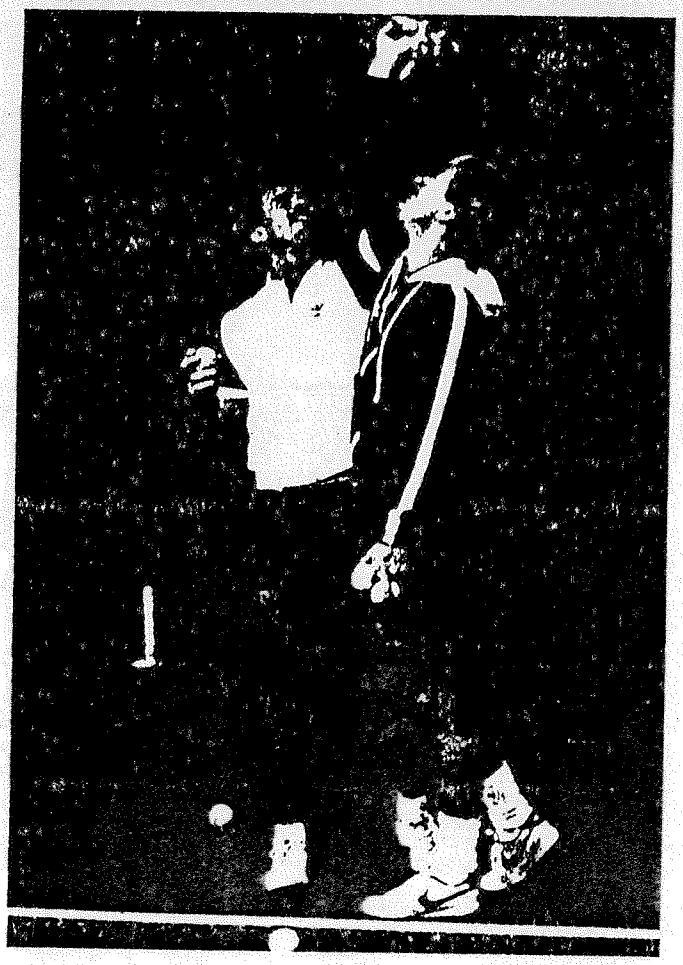
وفيما يلى شرح موجز لكل نمط من أنماط الصدق الثلاثة :

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 ١/٢ 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 $1\frac{1}{2}$ ميل واختبار السير المتحرك، والتى بلغت (٠,٩١٥-)، **فإنه يمكننا استخلاص الآتى:**

- أن لاختبار الجرى - المشى $1\frac{1}{2}$ ميل صدق تلازمى بلغ (٠,٩١٥-).
- أنه يمكن إحلال اختبار الجرى - المشى $1\frac{1}{2}$ 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 ناحية أخرى فقد يقترح بعض المدربين أن الوقوف على الرأس Headstand تعد وسيلة تنبؤ جيدة بالقدرة على العدو Sprinting ability، وقد جاء هذا الاقتراح نتيجة وجود معامل ارتباط (معامل صدق) بلغ (-٨٥, ٠) بين الوقوف على الرأس (كمتنبئ) والقدرة على العدو (كمحك). ويرى علماء القياس أنه لا يجب الأخذ فقط بمعامل الارتباط لحسم هذه العلاقة، لأنه إذا كان الوقوف على الرأس وسيلة تنبؤ حقيقية بالنسبة للقدرة على العدو، فإن تدريب الأفراد على الوقوف على الرأس يجب أن يتبعه بالضرورة تحسن فى قدرتهم على العدو وهو ما لم تؤكدة الدراسات والبحوث العلمية المتخصصة وذلك على الرغم من وجود ارتباط مرتفع بين المتغيرين.

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

مما سبق نستخلص الآتى:

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

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

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

$$ى = ٣,٥٠ + ٢,٦ س$$

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

$$ى = ٣,٥٠ + ٢,٦ (١٤)$$

$$= ٣,٥٠ + ٣٦,٤$$

$$= ٨٦,٧$$

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

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

ما هو المقصود بالتكوين الفرضى؟

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

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

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

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

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

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

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

فالقدرة على التفكير كتكوين فرضى، يتم تفسير درجات الاختبار الذى يقيسها على أساس أنه يدلنا 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 لبناء أو إعداد صدق التكوين الفرضى كما هو متبع بالنسبة لصدق المحتوى (المضمون) أو الصدق المرتبط بالمحك. وقد قدم كرونباخ وميل Cronbach and Meehl ١٩٥٥م مجموعة من الإجراءات لحساب صدق التكوين الفرضى هى:

١ - تحليل العمليات العقلية لإعداد فقرات الاختبار:

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

٢ - مقارنة درجات مجموعات معروفة:

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

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

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

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

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

الإجراءات الارتباطية لحساب صدق التكوين الفرضى:

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

أولاً: حساب الصدق الداخلى للاختبار internal validity: وتتم هذه الخطوة على أساس افتراض أن الاختبار الذى يقيس السمة له خصائص داخلية محددة مثل الاتساق الداخلى Internal Consistency، وتعدد أبعاد 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 التمييز 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 على درجات مرتفعة فى الدافع نحو النجاح فى الإنجاز the motive to achieve success .

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

كيفية تحقيق الصدق التقاربى وصدق التمييز:

«قدم كامبل وفيسك ١٩٥٩م» تصميماً تجريبياً خاصاً لدراسة كل من الصدق التقاربى وصدق التمييز، هذا التصميم يشير إلى ما 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، يتم قياسها بطريقتين methods أو أكثر، حيث يتم تحليل معاملات الارتباط البينية intercorrelations داخل المصفوفة لتحقيق كل من الصدق التقاربى وصدق التمييز للاختبار وفقا لمبدئين مهمين هما:

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

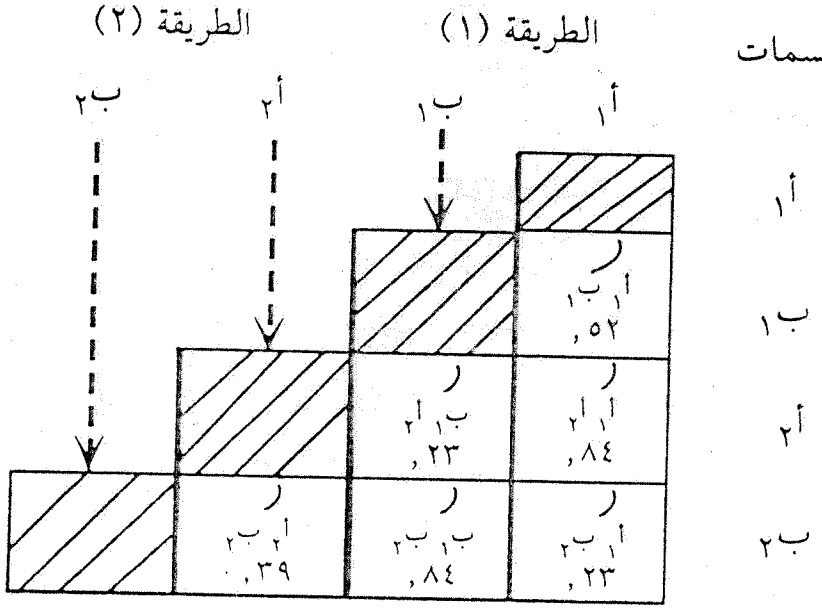
٢ - أن استخدام نفس الطرق لقياس سمات مختلفة سوف يؤدى إلى وجود معاملات ارتباط منخفضة فيما بين هذه المقاييس، مما يؤكد صدق التمييز، حيث يستخلص من هذين المبدأين الآتى:

«أن اتفاق السمة واختلاف الطريقة يؤدى إلى معاملات ارتباط مرتفعة» 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:

أ١ب١ = ٠,٥٢ (العدوان على قائمة الشخصية مع العصابية على نفس القائمة).

أ١أ٢ = ٠,٨٤ (العدوان على قائمة الشخصية مع العدوان على الاختبار الإسقاطى).

أ١ب٢ = ٠,٢٣ (العدوان على قائمة الشخصية مع العصابية على الاختبار الإسقاطى).

ب١أ٢ = ٠,٢٣ (العصابية على قائمة الشخصية مع العدوان على الاختبار الإسقاطى).

ب١ب٢ = ٠,٨٤ (العصابية على قائمة الشخصية مع العصابية على الاختبار الإسقاطى).

أ٢ب٢ = ٠,٣٩ (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:

أ١أ٢ = ٠,٨٤ (ويمثل درجة العدوان على قائمة الشخصية مع درجته على الاختبار الاسقاطى).

ب$_{١}$ب$_{٢}$ = 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:

أ$_{١}$ب$_{١}$ = ٠,٥٢ (ويمثل درجة العدوان على قائمة الشخصية مع درجة العصابية على نفس القائمة).

أ$_{٢}$ب$_{٢}$ = 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:

(أ$_{١}$أ$_{٢}$ = ٠,٨٤ ، ب$_{١}$ب$_{٢}$ = ٠,٨٤ وهى أكبر من

أ$_{١}$ب$_{١}$ = ٠,٥٢ ، أ$_{٢}$ب$_{٢}$ = 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:

$$= \sqrt{٦٥ , ٠}$$

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 ٨٤, ٠ ، فإن الصدق الذاتى لهذا الاختبار يكون:

$$= \sqrt{٨٤, ٠} = ٩٢, ٠$$

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

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

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

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

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

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

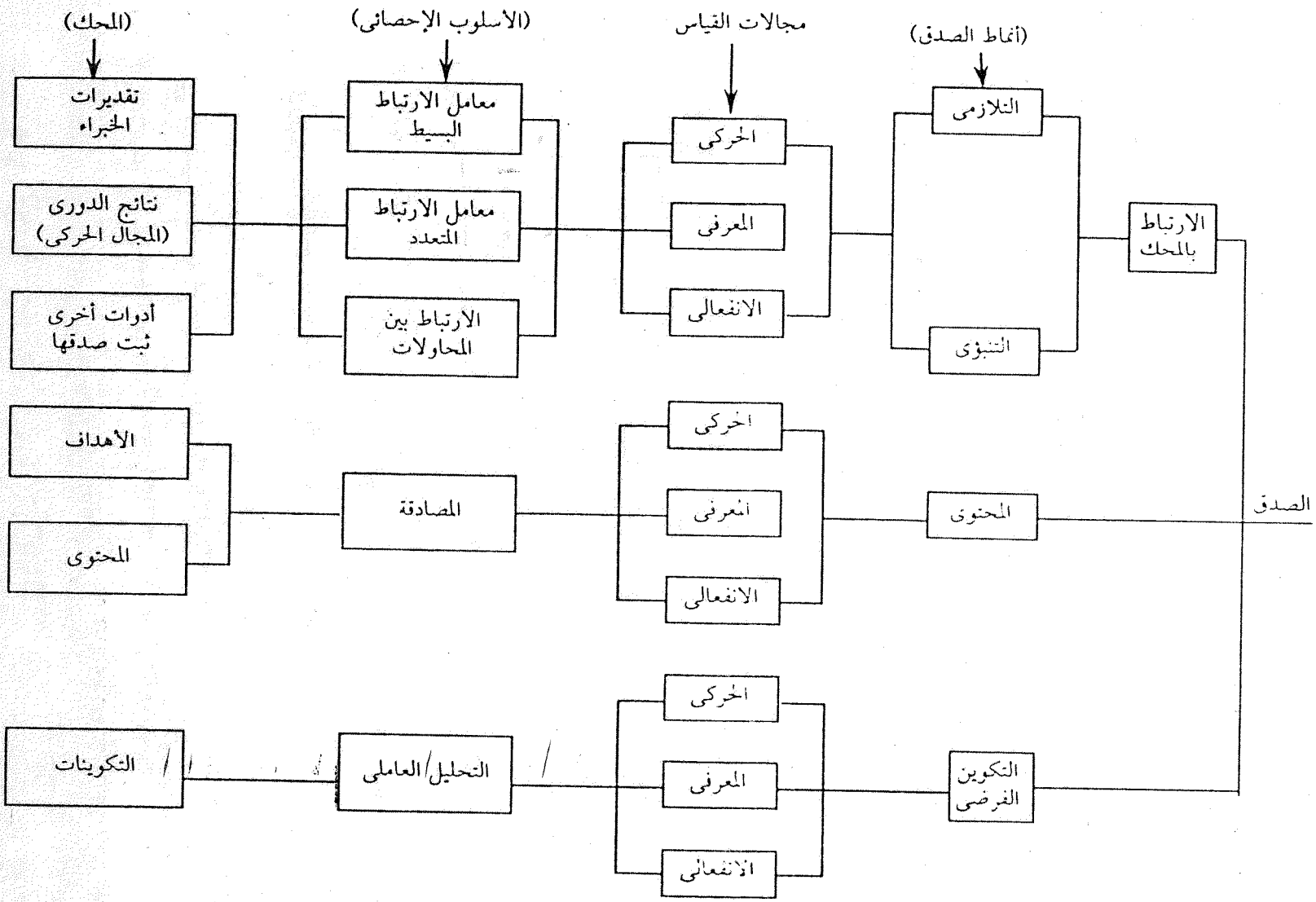
أخرى، وذلك فى حالة إعطاء الدرجة (١) عندما تصيب الكرة الهدف، والدرجة (صفر) إذا خرجت الكرة بعيدا عن الهدف.

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 (٢-٢) 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 المهـــارة، إلى مكوناتها الأساسية principle com-ponents التى تشكل فى مجموعها محتوى تلك القدرة أو المهارة.
- عرض هذه المكونات على الخبـراء لاستطلاع آراءهم حول صدق هذه المكونات بالنسبة للقدرة أو المهارة الحركية المطلوب قياسها.
- تحليل آراء الخـبـراء بالنسبـة لـكل مكـون مـن المكونات الـتى كشف عنها التحليل.

مثال (٣ - ١):

أراد أحد الباحـثين تحديد المكونات الأسـاسيـة للـقدرة الحـركيـة motor ability فى المجال الرياضى، توطئة لوضع بطارية اخـتبار لقياس هذه القدرة المركبة. ولتحقـيق ذلك فقد قام الباحث بتحليل هذه القـدرة معتمداً فى ذلك على الدراسات الســابقة وخبرته الذاتيـة فى مجال التخصـص، حيث توصل إلى تحديد ١٧ مكونا جدول (٣ - ٢).

وللتحقق مـن صدق كل مكون من المكونات الـ ١٧ قام الباحـث باختيار مجموعة من الخبراء تتكون من ٣٠ خبيرا من المتخصصين فى القياس والتقويم فى المجـال الريـاضى، وقـــد تم اسـتطلاع آراءهـم فى هـذه المكونـات 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 بالقـدرة الحركية وذلك بوضع دائرة حول

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

١٠	٩	٨	٧	٦	٥	٤	٣	٢	١

حيث تعنى هذه المجموعة من الدرجات الآتى :

- الدرجة (١) وتعنى أن المكون لا يقيس هذه القدرة
- الدرجة (٥) وتعنى أن المكون يقيس هذه القدرة بدرجة متوسطة.
- الدرجة (١٠) وتعنى أن المكون يقيس هذه القدرة بدرجة كبيرة جداً.
- الدرجة الأقل من (٥) تعنى علاقة ضعيفة بين المكون والقدرة
- الدرجة الأكثر من (٥) تعنى علاقة قوية بين المكون والقدرة.

● تجميع الاستمارة لتصنيف آراء الخبراء بالنسبة لكل مكون على حدة وفق 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 :

$$ص = ل + \frac{٠,٥ - مجـ ن}{\sqrt{ن}}$$ 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 + ٠,١ = ٠,٢٧

إذاً مجـ ن = ٠,٢٧

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

- 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 الخبراء بالنسبة لمكون «القدرة العضلية» 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
نسبة المحكمين	٠	٠,٠٣	٠,٠٧	٠,١٣	٠,١	٠,٠٧	٠,١	٠,١٧	٠,١٧	٠,١٧

(٣٠)

(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 + ١,٠

= ٠,٣٣

نَ = ٠,١

وبالتعويض فى المعادلة (٣ - ١) ينتج أن:

$$\text{ص} = ٥ + \frac{٠,٥ - ٠,٣٣}{٠,١}$$

$$= ٥ + \frac{٠,١٧}{٠,١}$$

= 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 الشىء 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 المحسوب 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 التقين (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 الشىء 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 ص_2 = معامل صدق الاختبار بعد التعديل.

س ص_1 =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:

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

$$= \frac{٠,٧٠٤}{\sqrt{٠,٩١٢٩٥}}$$

$$= \frac{٠,٧٠٤}{٠,٩٥٥٤٨}$$

$$= ٠,٧٣٧$$

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 :

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

معادلة (٣ - ٣)

حيث إن :

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

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

ص ص = معامل ثبات المحك.

مثـــال (٣-٣):

قام أحد البـاحثين بإعداد اختـبار 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 $\sqrt{٠,٩٦١ \times ٠,٩٥٠}$ 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 القـيمـة $\sqrt{٠,٦٧}$ 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 التميز 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 فى مجال علم النفس الرياضى، ولتوضيح ذلك نسوق المثال التالى(*):

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

- النشاط البدنى كخبرة اجتماعية (٨ عبارات)
- النشاط البدنى للصحة واللياقة (١١ عبارة)
- النشاط البدنى كخبرة توتر ومخاطرة (٩ عبارات)
- النشاط البدنى كخبرة جمالية (٩ عبارات)
- النشاط البدنى لخفض التوتر (٩ عبارات)
- النشاط البدنى كخبرة للتفوق الرياضى (٨ عبارات)

ولحساب صدق التكوين الفرض للأبعاد الستة للمقياس، قام كينيون بعرض الأبعاد على مجموعة من المفحوصين لسؤالهم عن مدى تفضيلهم لكل بعد من الأبعاد الستة وفقا لمقياس تقدير خماسى 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	
١٩ - ٠	١٢	٧	٥	٠	٠	٠	١٢
٣٩ - ٢٠	١٨	٦	٧	٥	٠	٠	١٨
٥٩ - ٤٠	٢٥	٩	٧	٧	٢	٠	٢٥
٧٩ - ٦٠	٨٠	١٠	٢٠	٣٠	١١	٩	٨٠
٩٩ - ٨٠	٣٥	٠	٤	١٠	١٥	٦	٣٥
١١٩ - ١٠٠	٢٠	٠	٠	٥	١٠	٥	٢٠
١٣٩ - ١٢٠	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.

. . . . . 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:

$$\text{س}_{(\text{ن ص})} = \frac{\text{ن} \times \text{س}_{\text{ص}}}{\sqrt{\text{ن} + \text{ن}(\text{ن} - 1)\,\text{س}_{\text{ص}}^{2}}}$$

معادلة (٣ - ٤)

حيث إن :

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

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

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

$\text{س}_{\text{ا ا}}$ = معامل ثبات الاختبار.

مثال (٣ - ٤):

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

الحل:

بالتعويض فى المعادلة (٣ - ٤) ينتج أن:

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

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

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

$$= \frac{٢}{\sqrt{٩,٦}}$$

$$= \frac{٢}{\sqrt{٣,٠٩٨}}$$

$$= ٠,٦٤٥٧$$

$$= ٠,٦٥$$

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:

$$\text{س}_{(\text{ن م})} = \frac{\text{س ص}}{\sqrt{\left(\frac{\text{ن}}{\text{ن} - 1 + \text{ر}_{101}}\right)\left(\frac{\text{م}}{\text{م} - 1 + \text{ر}_{101}}\right)\left(\text{ر}_{101} + \text{ر}_{202}\right)}}$$ معادلة (٣ -٥)

حيث إن :

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

س 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_{101} = معامل ثبات الاختبار.

ر_{202} = معمل ثبات المحك.

مثال (٣ - ٥):

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

الحل:

من المثال السابق يتضح أن:

س ص = ٦٥ , ٠

ن = ٣

م = ٣

ر_{101} = ٥٦ , ٠

$$\sqrt{٢٠٢} = ٠,٥٨$$

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:

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

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

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

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

$$= \frac{٠,٦٥}{\sqrt{٠,٥٦٠٩}}$$

$$= \frac{٠,٦٥}{٠,٧٤٩}$$

$$= ٠,٨٧$$

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:

$$\text{ن} = \frac{\text{س}^{٢}(\text{ن ص}) \times (١ - ١٠١)}{(\text{س ص})^{٢} - \text{س}^{٢}(\text{ن ص}) \times ١٠١}$$ 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 :

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

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 :

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

$$= \frac{٠,٦٥}{٠,٩٠٦}$$

$$= ٠,٧٢$$

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 :

$$\text{&\&} = \frac{\text{س ص}}{\sqrt{\text{١٠١} \times \text{٢٠٢}}}$$

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:

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

$$= \frac{٠,٦٥}{\sqrt{٠,٧٣٨}}$$

$$= \frac{٠,٦٥}{٠,٨٥٩}$$

$$= ٠,٧٦$$

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

٢ - عدم وضوح التعليمات Unclear Direction

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

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

Reading vocabulary and sentence too difficult

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

- الاستيعاب القرائى للمفحوص Reading comprehension
- الذكاء Intelligence

٤ - عدم ملاءمة مستوى صعوبة فقرات الاختبار

Inappropriate level difficulty of the test items

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

٥ – الغموض The Ambiguity

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

٦ – الاختبار قصير أكثر مما ينبغى Test too short

من المعروف أن 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 disturbence 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 (x) يميل بعض المفحوصين إلى الإجابة بصح (✓) عندما لا يعرفون الإجابة الصحيحة للفقرة، بينما يميل البعض الآخر الى الإجابة بخطأ (x). فعندما يتضمن الاختبار عدداً كبيراً من العبارات (الجمل) الصحيحة، فإن مثل هذا الاختبار يعطى فرصة أكبر للفئة الأولى من المفحوصين على حساب الفئة الثانية، ومن ثم يصبح من 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 أهداف وأساليب بناء كل من ثبات وصدق الاختبار تبدو متضاربة، فقد تتعارض متطلبات الحصول على ثبات مرتفع للاختبار مع متطلبات الحصول على صدق مرتفع لنفس الاختبار، فالثبات المرتفع يتطلب فقرات متساوية فى درجة الصعوبة، كما يتطلب معاملات ارتباط بينية مرتفعة (فيما بين الفقرات) Intercorrelations ، فى حين يتطلب الصدق المرتفع لنفس الاختبار فقرات مختلفة فى درجات صعوبتها، كما يتطلب معاملات ارتباط بينية (فيما بين الفقرات) منخفضة.

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

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 الدقة

Factorial validity الصدق العاملى

Face Validity الصدق الظاهري

Factorial analysis التحليل العاملى

Factor Loading تشبع الاختبارات على العامل

H

Hypothetical Construct تكوين افتراضى

Homogeneous Lengthening

التطويل المتجانس فى الاختبار

I

Intrinsic Validity الصدق الداخلى(الذاتى)

Independent مستقبل

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

P

Predictive efficiency

الكفاءة التنبؤية للاختبار

Pure factors matrix مصفوفة العوامل النقة

Principle Components المكونات الأساسية

Phi-Coefficient معامل فاى

Present standing الوضع الراهن

Predictive tests اختبارات تنبؤية

تشبعات الاختبارات على العوامل Saturations

R

Response set	اتجاه الاستجابة
Relationship	علاقة
Representativeness validity	صدق التمثيل
Relevance validity	الصدق التحليلى
Rho	معامل ارتباط سبيرمان لفروق الرتب
Rating scales	مقاييس تقدير
Reading comprehension	الاستيعاب القرائى

S

situations	مواقف
Statistical validity	الصدق الإحصائى
Spatial ability	القدرة المكانية

T

True - False	الصح - الخطأ
Trait Validity	صدق السمة
Theoretical construct	تكوين نظرى
Test efficiency	كفاءة الاختبار
Tournament	نظام الدورى
Tetrachoric Correlation	الارتباط الرباعى
Trained groups	مجموعات مدربة

V

Validity	الصدق
Validity Coefficient	معامل الصدق

مراجع الفصل الثالث

١ - سعد عبد الرحمن (١٩٩٨م) : القياس النفسى - النظرية والتطبيق. دار الفكر العربى، القاهرة.

٢ - صفوت فرج (٢٠٠٠م) القياس النفسى. مكتبة الأنجلو المصرية. القاهرة.

٣ - صفوت فرج (١٩٨٠م): التحليل العاملى فى العلوم السلوكية. دار الفكر العربى، القاهرة.

٤ - محمد حسن علاوى، ومحمد نصر الدين رضوان (٢٠٠٠م): القياس فى التربية الرياضية وعلم النفس الرياضى. دار الفكر العربى، القاهرة.

5 - Anastasi, A. (1990) : Psychological testing, New york : Macmillan.

6 - Cronbach, L.J. (1970) : Test Validation . In R.L. Thorndike (Ed.) , Educationul measurement. washington, DC: American Council on Education.

7 - Kaplan, R.M., and Saccuzzo, D.P. (1988) : Psychological Testing, Principles, Applecations, and Issues. Brooks Cole Publishing Company, California.

8 - Miller, D.K. (1994) : Measurement - By the physical educators : why an How. WCB Brown and Benchmark ppublishers.

9 - Morris, L.L., Gibben, C.T.F. , and Lindheim, E. (1987) : How to measure performance and use tests. saGe publications, Inc., London New dehil.

10 - Singh, A.K. (1986) : Tests, Measurements and Research methods in Behavioural Sciences. Tata McGraw- Hill Publishing Company Limited , New Delhi.

الفصل الرابع

القيـاس المعـرفى

١- طبيعة القياس المعرفى

٢ - خطوات بناء الاختبارات المعرفية

أولاً : التخطيط للاختبار

ثانيا : كتابة فقرات الاختبار

ثالثا : تطبيق الاختبار

رابعا : تحليل الفقرة

خامساً : مراجعة وتنقيح الفقرة

- معجم إنجليزى - عربى للمصطلحات
- مراجع الفصل الرابع

١ - طبيعة القياس المعرفى

يستهدف القياس بالنسبة للبعد المعرفى فى المجال الرياضى التعرف على التغيرات والتطورات التى تحدث فى المعلومات والفهم للعديد من الموضوعات المتعلقة بتاريخ التربية البدنية، وتاريخ الأنشطة الرياضية، وقواعد التحكيم، والمعلومات الأساسية المتعلقة بالجسم البشرى والتى تؤثر على كفاءته فى النشاط الرياضى، إلى غير ذلك من الموضوعات.

ويشير الفهم فى المجال الرياضى إلى القدرة على استخدام المعلومات والحقائق العلمية المختلفة فى مجال التطبيق العملى، فمثلا يمكننا أن نتوقع من اللاعب الذى يعرف الأوضاع والتحركات الصحيحة فى طريقة «رجل لرجل» الدفاعية فى كرة السلة، أن يقوم بتطبيق هذه الطريقة بشكل جيد أثناء المباريات فى كرة السلة.

مفهوم القياس المعرفى:

يشير القياس المعرفى إلى الاختبارات التى تستخدم لقياس عمليات التفكير thought processes ، حيث يلاحظ أن هذه الاختبارات تستخدم اللغة المكتوبة، بمعنى أنها اختبارات كتابية Written tests، وأنها تستخدم فى المجال الرياضى لقياس مستوى تحصيل الأغراض المعرفية.

وتأخذ الإجابة فى الاختبارات الكتابية نمطين رئيسيين هما :

النمط الأول : وفيه يطلب من المفحوص اختيار الإجابة (الاستجابة) المناسبة من بين عدد من الإجابات، ويتضمن هذا النمط وحدات الاختبار التالية: الصواب - الخطأ، الاختيار من متعدد، المزاوجة.

النمط الثانى : وهو يستلزم إمداد المفحوص بالمعلومات الضرورية التى تمكنه من الإجابة (الاستجابة)، ويتضمن هذا النمط الإجابات القصيرة كالتكميل وكتابة المقال.

متطلبات بناء الاختبارات المعرفية:

لكى يستطيع الباحث أو (القائم بإعداد الاختبار المعرفى) النجاح فى مهمته فإنه يلزم أن يتوافر له الآتى:

١ - أن يكون ملما بالأساليب الصحيحة والمناسبة لبناء الاختبارات.

٢ - أن يكون على دراية بمجال المادة (موضوع التخصص) الذى يعد له الاختبار.

٣ - أن تكون لدية المهارة فى التعبير المكتوب (التعبير كتابة) written expression عن الموضوعات التى يعدلها الاختبار.

٤ - أن يكون على وعى بمستويات فهم المفحوصين اللذين يعد لهم الاختبار.

٥ - الصبر والاستعداد للعمل لفترات زمنية طويلة.

٢ - خطوات بناء الاختبار المعرفى

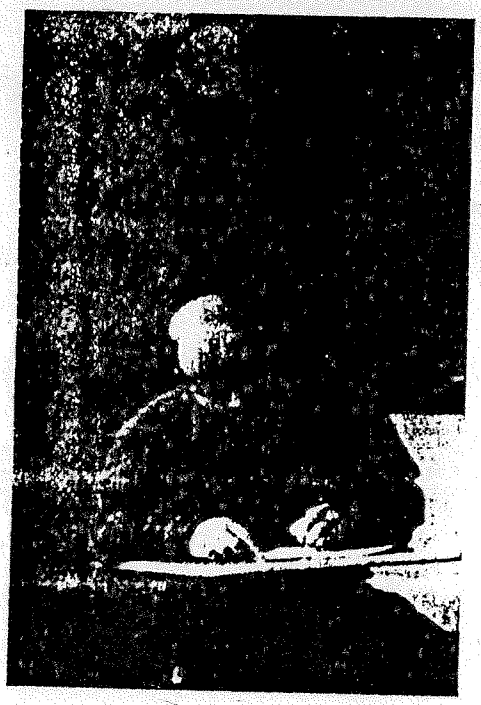
لبناء الاختبار المعرفى تتبع الخطوات الخمس التالية:

- التخطيط للاختبار Planning of the test
- كتابة فقرات الاختبار writing the items of the test
- تطبيق الاختبار test administration
- تحليل الفقرة item analysis
- مراجعة وتنقيح الفقرة item revision

وفيما يلى شرح موجز لكل خطوة من هذه الخطوات:

أولاً: التخطيط للاختبار:

إن أول خطوة يجب أن توضع فى الاعتبار عند بناء أى اختبار هى التخطيط للاختبار، حيث تتضمن هذه الخطوة الآتى:



القياس المعرفى



يمكن استخدام الاختبار لقياس معلومات الطالب عن الأساليب الفنية للعب

١ - أن يحدد القائم بوضع الاختبار مجال عمل (تطبيق) الاختبار والأغراض الخاصة به.

٢ - تحديد المصطلحات التى يمكن أن يتضمنها الاختبار تحديداً دقيقاً وواضحاً مثال ذلك (الذكاء الخطى - اللياقة البدنية، التحصيل الدراسى.. وغيرها).

٣ - أن يحدد المسئول عن إعداد الاختبار مسبقاً طبيعة محتوى الاختبار (الفقرات - العبارات) وطبيعة الاستجابة عليها، مثال ذلك (المقال - صح أو خطأ - التكميل - الاختيار من متعدد .. إلخ).

٤ - تحديد أسلوب العينة اللازمة لبناء الاختبار موضحا نوع العينة، وحجمها، وأسلوب اختيارها، والمجتمع الأصلى للعينة.

٥ - وضع التفاصيل والترتيبات الخاصة بالتطبيق الأولى والتطبيق النهائى للاختبار.

٦ - تقدير الوقت المتوقع أن يستغرقه الاختبار (طول الاختبار - والوقت اللازم لإكمال الاختبار).

٧ - تحديد الأساليب الإحصائية المتوقع التعامل معها لبناء الاختبار.

٨ - تحديد حجم العينة اللازم لإعداد (بناء) معايير الاختبار.

٩ - إعداد الدليل الخاص بالاختبار manual والذى يتضمن التعليمات وشروط تطبيق الاختبار.

وهناك وجهة نظر ترى أنه يجب أن يؤخذ فى الاعتبار صدق المحتوى (المضمون) Content validity بالنسبة لفقرات الاختبار test items كخطوة أولى عند التخطيط لبناء الاختبار، وأن يمثل الاختبار الأغراض التعليمية instructional objectives، التى تتضمنها الوحدة التدريسية أو التدريبية. ولاستكمال صدق المحتوى للفقرات يقترح علماء القياس استخدام ما يعرف بجدول مواصفات الاختبار table of test specifications، حيث يستفاد من جدول المواصفات فى الآتى:

- تحديد أنواع وعدد فقرات (أسئلة) الاختبار.
- تحديد أنواع المهام (الأعمال) tasks بالنسبة لعمليات التفكير المختلفة وأيضا تحديد عدد الفقرات الخاصة بكل عملية من تلك العمليات.
- تحديد العناصر أو المجالات الفرعية للموضوع الرئيسى الذى يعد له الاختبار وعدد الوحدات الخاصة بكل عنصر أو مجال فرعى. ويتضمن الجدول (٤ - ١) مثالاً لجدول المواصفات لاختبار للكرة الطائرة.

جدول (٤ - ١) : المواصفات الخاصة باختبار معرفى للكرة الطائرة

يتكون من ٥٠ فقرة

(المهام أو الأعمال - وعدد الأسئلة - والنسب المئوية)				
المجالات الفرعية	المعرفة	الفهم	التحليل	التطبيق
- التاريخ	٢ (٤٪)			
- قواعد اللعب	٥ (١٠٪)	٥(١٠٪)		٥(١٠٪)
- النواحى الفنية	٢ (٤٪)	٥(١٠٪)	٨(١٦٪)	
- استراتيجية الهجوم	٤ (٨٪)			٥(١٠٪)
- استراتيجية الدفاع	٤(٨٪)			٥(١٠٪)
المجموع	١٧(٣٤٪)	١٠(٢٠٪)	٨(١٦٪)	١٥(٣٠٪)
				= ٥٠ (١٠٠٪)

ماهية جدول المواصفات:

جدول المواصفات عبارة عن مخطط أو رسم أو تصميم يبين الأهمية النسبية لوحدات الاختبار فيما يتعلق بكل مجال من المجالات الفرعية وفقا للفئات الرئيسية للمجال المعرفى وكذا النسب المئوية لعدد الفقرات فى كل مجال من هذه المجالات الفرعية.

ويتكون جدول المواصفات من اتجاهين two-ways ، الأول ويمثل المحور الرأسى ويتضمن العناصر أو المجالات الفرعية للموضوع المطلوب قياسه، والثانى ويمثل المحور الأفقى ويتضمن الفئات الرئيسية للمجال المعرفى (الأغراض التربوية) ففى الجدول (٤ - ١) على سبيل المثال يمثل المحور الرأسى المجالات الفرعية للموضوعات التى سوف يتناولها القياس المعرفى بالنسبة للكرة الطائرة وهى خمسة مجالات فرعية، بينما يمثل المحور الأفقى الفئات الرئيسية لمستويات القياس المعرفى والتى تظهر فى أربعة مستويات هى: المعرفة والفهم والتحليل والتطبيق.

ويستخدم جدول المواصفات كوسيلة مهمة لتأكيد صدق المحتوى (المضمون) للاختبار المقترح حيث يشير هذا الصدق إلى المدى الذى تقيس به وحدات الاختبار عينات مناسبة من الموضوع الدراسى أو القدرات التى صمم الاختبار لقياسها.

فلو طلب منا - على سبيل المثال - القيام بإعداد جدول المواصفات لاختبار معرفى للريشة الطائرة (البادمنتون)، فى هذه الحالة يمكننا اتباع الخطوات التالية:

الخطوة ١: تحديد المجالات الفرعية (العناصر) التى تستوجب أن يشملها الاختبار: ويتم ذلك عن طريق الآتى:

- الرجوع إلى المراجع العلمية المتخصصة.
- تحليل الدراسات والبحوث السابقة المرتبطة.
- استطلاع آراء المتخصصين فى المجال.

الخطوة ٢ : عرض المجالات الفرعية (العناصر) على مجموعة من الخبراء فى مجال التخصص: ويتم كالتالى:

- إعداد استمارة لاستطلاع رأى الخبراء.

- وضع معايير اختيار الخبراء.

- تحديد الخبراء بحيث يكون عددهم مناسبا (توضع قائمة بأسمائهم فى المرفقات).

- عرض الاستمارة على الخبراء وذلك بغرض الآتى:

أ) التأكد من صحة المجالات (العناصر) الفرعية.

ب) اقتراح حذف أو إضافة مجالات (عناصر) أخرى.

جـ) تقدير الأهمية النسبية لكل مجال (عنصر) من المجالات الفرعية.

د) تقدير النسبة لكل مستوى من مستويات المجال المعرفى.

فلو افترضنا أن الباحث توصل نتيجة الخطوة الأولى إلى ثمانية مجالات (عناصر) فرعية هى: التاريخ، القيم، الأجهزة والأدوات، آداب وتقاليد اللعبة، الأمان، القواعد والقوانين المنظمة، الاستراتيجيات، الأساليب الفنية للعب. وعندما قام الباحث بعرض هذه المجالات الفرعية على مجموعة من الخبراء، جاءت الأهمية النسبية للمجالات (العناصر) الفرعية كالتالى :

١ - تاريخ اللعبة	٥٪
٢ - القيم	٥٪
٣ - الأجهزة والأدوات	١٠٪
٤ - آداب وتقاليد اللعبة	١٠٪
٥ - الأمان والسلامة	١٠٪
٦ - القواعد والقوانين	٢٠٪
٧ - استراتيجات اللعب	١٥٪
٨ - الأساليب الفنية للعب	٢٥٪
	١٠٠٪

كما جاءت الأهمية النسبية لكل مستوى من مستويات المجال المعرفى (ستة مستويات) كالتالى:

١ - تذكر المعلومات	٣٠٪
٢ - الفهم	١٠٪
٣ - التطبيق	٣٠٪
٤ - التحليل	٢٠٪
٥ - التركيب	٠٪
٦ - التقويم	١٠٪
	١٠٠٪

الخطوة ٣: يتم وضع النتائج السابقة وفقا لأهميتها النسبية فى جدول المواصفات بحيث تمثل المجالات الفرعية (الثمانية للعبة المحور الرأسى للجدول ، وتمثل مستويات المجال المعرفى (٦ مستويات) المحور الأفقى للجدول. وبناء على هذا يمكننا إعداد جدول للمواصفات (فى اتجاهين) يتكون من ٩ أعمدة، ١٠ صفوف جدول (٤ - ٢) كالتالى:

جدول (٤ - ٢): جدول المواصفات (من اتجاهين) لاختبار معرفى للريشة الطائرة يتكون من ٢٠٠ بند أو عبارة

المجالات (العناصر) الفرعية	الأهمية النسبية	تذكر المعلومات	الفهم	التطبيق	التحليل	التركيب	التقويم	المجموع
		٣٠٪	١٠٪	٣٠٪	٢٠٪	٠٪	١٠٪	١٠٠٪
التاريخ	٥٪	١,٥٪ ٣	٠,٥٪ ١	١,٥٪ ٣	١٪ ٢	٠٪ ٠	٠,٥٪ ١	١٠
القيم	٥٪	١,٥٪ ٣	٠,٥٪ ١	١,٥٪ ٣	١٪ ٢	٠٪ ٠	٠,٥٪ ١	١٠
الأجهزة والأدوات	١٠٪	٣٪ ٦	١٪ ٢	٣٪ ٦	٢٪ ٤	٠٪ ٠	١٪ ٢	٢٠
الآداب والتقاليد	١٠٪	٣٪ ٦	١٪ ٢	٣٪ ٦	٢٪ ٤	٠٪ ٠	١٪ ٢	٢٠
الأمان والسلامة	١٠٪	٣٪ ٦	١٪ ٢	٣٪ ٦	٢٪ ٤	٠٪ ٠	١٪ ٢	٢٠
القواعد والقوانين	٢٠٪	٦٪ ١٢	٢٪ ٤	٦٪ ١٢	٤٪ ٨	٠٪ ٠	٢٪ ٤	٤٠
الاستراتيجيات	١٥٪	٤,٥٪ ٩	١,٥٪ ٣	٤,٥٪ ٩	٣٪ ٦	٠٪ ٠	١,٥٪ ٣	٣٠
الأساليب الفنية	٢٥٪	٧,٥٪ ١٥	٢,٥٪ ٥	٧,٥٪ ١٥	٥٪ ١٠	٠٪ ٠	٢,٥٪ ٥	٥٠
المجموع	١٠٠٪	٦٠	٢٠	٦٠	٤٠	٠	٢٠	٢٠٠

الخطوة ٤ : تحديد الأوزان النسبية relative weights لكل خلية من خلايا الجدول والتى تمثل نقطة التقاء كل مجال فرعى (المحور الرأسى - ٨ مجالات) بكل مستوى من المستويات المعرفية (المحور الأفقى - ٦ مستويات)، فعلى سبيل المثال يمكن تحديد الوزن النسبى للخلية الأولى والخاصة (بتذكر المعلومات التاريخية للريشة الطائرة) كالتالى :

الخلية الأولى= الوزن النسبى للتاريخ × الوزن النسبى لتذكر المعلومات.

= ٠,٠٥ × ٠,٣٠ = ٠,٠١٥

= ٠,٠١٥ ÷ ١٠٠ = ١,٥ أى ١,٥٪

ويمكن تحديد الوزن النسبى للخلية الثانية كالتالى:

= ٠,٠٥ × ٠,١٠ = ٠,٠٠٥

= ٠,٠٠٥ ÷ ١٠٠ = ٠,٥ أى ٠,٥٪

وهكذا بالنسبة لبقية خلايا الجدول والبالغ عددها ٤٨ خلية (٦×٨ =٤٨).

ويوضع الناتج فى داخل الخلايا وفق ما هو موضح بالجدول (٤ - ٢).

الخطوة ٥ : تحديد عدد العبارات (الأسئلة أو الفقرات) الخاصة بكل خلية:

ويسبق هذه الخطوة تحديد العدد الكلى للعبارات (الأسئلة) للاختبار ككل، فلو افترضنا أن الباحث قرر أن يكون عدد عبارات الاختبار المعرفى للريشة الطائرة هو ٢٠٠ عبارة، فى هذه الحالة يتم تحديد عبارات كل خلية من خلايا جدول المواصفات (٤٨ خلية) كالتالى:

$$\text{عدد عبارات الخلية الأولى} = \frac{\text{الوزن النسبى للخلية} \times \text{العدد الكلى للعبارات}}{١٠٠}$$

$$= \frac{٢٠٠ \times ١,٥}{١٠٠}$$

= ٣ عبارات

$$\text{وبالنسبة للخلية الثانية} = \frac{٢٠٠ \times ٠,٥}{١٠٠}$$

= ١ عبارة واحدة

وهكذا بالنسبة لبقية الخلايا.

ثانيا : كتابة فقرات الاختبار:

يمثل إعداد الفقرات items الخطوة الثانية من خطوات إعداد اختبارات الورقة والقلم paper and pencil tests ، ووفقا لتعريف بين Bean ١٩٥٣م يتبين أن الفقرة عبارة عن سؤال مفرد أو واجب أو عمل لا يمكن تحليله إلى ما هو أبسط.

تعريف الفقرة:

تعرَّف الفقرة item على أنها عمل أو واجب أو سؤال مفرد لا يمكن تحليله أو تجزئته إلى وحدات أصغر أو أبسط. وتعد الفقرة وحدة الاختبار الأساسية، فكل مقاييس أو اختبار يتكون من عدد من الفقرات (العبارات أو الأسئلة أو البنود).

متطلبات إعداد وصياغة الفقرات:

عندما ينتهى الباحث من التخطيط للاختبار، يبدأ فى إعداد وكتابة الفقرات معتمداً فى ذلك على المصادر التالية:

١ - المراجع العلمية فى مجال التخصص.

٢ - الكتب الدراسية المقررة والمذكرات التى يعدها المتخصصون.

٣ - خبراته السابقة فى المجال الذى يعد له الاختبار.

ولكى ينجح الباحث فى هذه المهمة فإنه يلزم توافر الآتى:

- أن يكون لدى الباحث معرفة وتمكن من الموضوع الذى يكتب عنه، وأن يكون ملما بكل الحقائق facts والمبادئ principles والمفاهيم والمدركات الخاطئة misconceptions المتعلقة بهذا الموضوع.
- أن يكون ملما بطبيعة الأشخاص الذين يعد لهم الاختبار، كأن يكون على علم بمستوى ذكائهم وقدراتهم وميولهم ومستوياتهم

الاقتصادية والاجتماعية، لأن ذلك يمكنه من الاختيار الصحيح للفقرة والصياغة الجيدة لها من حيث الألفاظ المستخدمة ودرجة الصعوبة.

- أن يكون ملماً بالعبارات والألفاظ المناسبة وغير المناسبة، الواضحة وغير الواضحة، المحددة والعامة إلى غير ذلك من العبارات والألفاظ.
- أن تكون لديه حصيلة لغوية (عامة وخاصة) واسعة، وبحيث تشمل هذه الحصيلة: أصل الكلمة ومرادفاتها، وأن يكون الباحث ملماً بمختلف المعانى الخاصة بالكلمة word التى يختارها، وأن يكون قادراً على صياغة الفقرة بأبسط لغة ممكنة.
- أن يكون لدى الباحث القدرة على الاطلاع والبحث، وكذا القدرة على التصور والتخيل والابتكار حتى يتمكن من اختيار فقرات جيدة.
- بعد أن يقوم الباحث بتحديد وصياغة الفقرات، فإنه يفضل أن يقوم بعرضها على مجموعة من الخبراء لإبداء آراءهم فى تلك العبارات.

وقبل أن يبدأ الباحث فى كتابة الفقرة item writing يلزمه الإجابة على السؤالين التاليين:

السؤال الأول: ما هو نوع الاختبار المطلوب إعداده؟ وهل هذا الاختبار سيستخدم فقرات الصواب والخطأ، أو الاختيار من متعدد، أو المقال، أو التكميل أم كلها مجتمعة؟ وتكمن أهمية تحديد نوع الاختبار، فى كون جميع الاختبارات لا تستخدم نفس البنية أو الشكل، فاختبارات الشخصية والذكاء - على سبيل المثال - تتطلب كل أنواع الاستجابات.

السؤال الثانى: ما هو نمط الاستجابة على الاختبار؟ ففى الاختبارات التى تتطلب الإجابة بصح أو خطأ تستخدم أسئلة الصواب والخطأ، والاختيار من متعدد، والمزاوجة، والمقال.

وفى جميع الحالات يجب على المسئول عن إعداد الاختبار أن يقرر مبكراً عند التخطيط لبناء الاختبار أى الفقرات سوف يختار، ففى حالة اختيار فقرات المقال essay items يكون الاختبار «اختبار مقال - essay test»، وفى حالة اختيار فقرات: الصواب - الخطأ، والاختيار من متعدد، والمزاوجة يكون الاختبار «اختبار موضوعى - objective test».

فقرات اختبار المقال Essay test items

لعل من أهم ما يميز فقرات اختبار المقال أنها تعتمد على أسلوب التقويم الذاتى subjective evaluation، حيث تصمم الفقرات لقياس قدرة الأفراد على استخدام العمليات العقلية العليا higher mental processes فى التعبير عن أنفسهم عن طريق الكتابة بالنسبة لموضوعات تتطلب: تعيين هوية الشىء identifying ، وتفسير أو توضيح الشىء interpreting ، والتوحيد والدمج integrating ، والتنظيم organizing ، والتركيب synthesizing.

ولعل من أهم ما تتميز به فقرات اختبار المقال مايلى:

- سهولة الإعداد، بحيث يمكن وضع العديد من الفقرات فى دقائق معدودة.
- تستخدم لقياس المفاهيم المعقدة complex concepts، والقدرة على التفكير thinking ability ، ومهارات حل المشكلات problem - solving skills.
- تشجع المفحوصين: على كيف يتعلمون بفاعلية ونظام، وكيف يعبرون عن أفكاهم ideas.
- التخمين فى هذا النوع من الاختبارات يكون أقل ما يمكن.

ومن عيوب فقرات اختبار المقال مايلى:

- تتطلب وقتا طويلا عند تقويمها لتقدير الدرجات.
- تقدير الدرجة يتطلب اتخاذ بعض القرارات فيما يتعلق بتسجيل أجزاء الدرجة، مما يؤدى إلى تناقص ثبات هذه الاختبارات.
- نظراً لكونها تستغرق وقتا طويلا فى الإجابة، لذا فقد لا يتمكن المفحوص من الإجابة على كل الفقرات مما يجعلها محدودة كوسيلة لقياس المعرفة.

توجيهات عامة لكتابة فقرات المقال :

لبناء فقرات اختبار المقال بسهولة ويسر، يجب اتباع التوجيهات التالية:

١ - أن تعبر الفقرة عن المعرفة الرئيسية بالنسبة للموضوع الدراسى الذى درسه الطالب (الموضوع المطلوب قياسه)، حيث يلزم أن تستدعى الفقرة المادة الرئيسية فى الكتاب الدراسى الذى درسه الطالب.

٢ - أن يكون للفقرة إجابة واحدة صحيحة، لأنه فى حالة وجود أكثر من إجابة صحيحة للفقرة، فإنه يصبح من الصعب تقويم مستوى معرفة الفرد (المفحوص).

٣ - أن تتضمن الفقرة مجال واتجاهات الإجابة المطلوبة، لأن الفقرات (الأسئلة أو البنود) الغامضة غير الواضحة vague تؤدى إلى تباين واسع فى استجابات المفحوصين مما يجعل عملية التقويم أكثر صعوبة.

٤ - أن تكون جميع الفقرات (الأسئلة) موحدة بالنسبة لكل المفحوصين، بمعنى أن يطلب من كل المفحوصين الإجابة على نفس الفقرات (الأسئلة)، لأنه فى حالة الإجابة على فقرات مختلفة، فإنه يصعب مقارنة درجات المفحوصين بعضهم ببعض، لأنه عندما يختار المفحوص

الفقرات، فإن ذلك يعطيه ميزة تحقيق «أفضل إجابة the best - answer» مما قد يترتب عليه أن يصبح مدى الدرجات على الاختبار محدوداً وبالتالى تناقص ثبات درجات الاختبار.

٥ - إظهار الزمن (الوقت) التقريبى المخصص للإجابة على كل فقرة (سؤال)، وكذا الدرجات المخصصة لكل فقرة، لأن العديد من المفحوصين يحتاجون إلى دليل يرشدهم إلى الزمن الذى يستغرقونه للإجابة على كل سؤال.

٦ - وفى جميع الحالات يفضل استخدام عدد مناسب من فقرات المقال القصيرة short essay items فى مقابل عدد قليل من الفقرات الطويلة، لأن الفقرات القصيرة تكون أقل فى الغموض ambiguous بالمقارنة بالفقرات الطويلة مما يسهل على المفحوص فهمها.

كيفية تقدير درجات فقرات المقال:

لتقدير درجات فقرات اختبارات المقال، يلزم اتباع التوجيهات التالية:

- الاتفاق على طريقة موحدة لتقدير درجات الاختبار. لأن بعض المعلمين يركزون على النقاط الرئيسية التى يجب أن يتضمنها الاختبار، فى حين يركز البعض الآخر على ترتيب الفقرات وفقا لنوعية الاستجابة quality of response .

- يفضل تقويم نفس الفقرة same item بالنسبة لكل المفحوصين قبل الانتقال للفقرة التالية. كما يفضل أن يقوم مقدر الدرجات بمراجعة اتساق تقديره لدرجات الفقرة عن طريق الرجوع إلى أول مجموعة من الأوراق قام بتصحيحها.

- يفضل عدم ذكر اسم المفحوص على ورق الإجابة، لأن تقويم درجات المفحوص على الاختبار بدون أن يكون اسمه معروفا للقائم بعملية التقويم إجراء يكفل الحد من التحيز biases.

أمثلة لفقرات المقال:

اقرأ كل فقرة بعناية ثم أجب على كل واحدة فيها إجابة كاملة قدر الإمكان. الوقت التقريبى اللازم للإجابة على كل فقرة والدرجة المخصصة لها موضح بين الأقواس.

١ - عرِّف الارتباط، ثم وضح أوجه الاختلاف بين معامل إرتباط سبيرمان لفروق الرتب ومعامل الارتباط التتابعى لبيرسون (١٠ دقائق - ٢٠ درجة).

٢ - عرف الدرجة المعيارية، ثم اكتب المعادلات التى تستخدم لحساب كل من الدرجة التائية ت -T ، والدرجة ذ - Z ، والمئينيات Percentiles (١٠ دقائق - ٢٠ درجة).

٣ - أراد أحد المعلمين أن يتعرف على فعالية طريقته الجديدة فى التدريس. اشرح كيف يستطيع هذا المعلم التحقق من فعالية هذه الطريقة، على أن يتضمن هذا الشرح الآتى:

- إجراءات القياس التى يجب أن يستخدمها المعلم.
- متى تتم عملية القياس.
- التحليل الإحصائى الذى يجب أن يستخدمه المعلم (١٥ دقيقة - ٣٠ درجة).

فقرات الاختبار الموضوعى Objective test items

يعتقد بعد المتخصصين أن الاختبارات الموضوعية يمكن أن تستخدم لقياس المستويات المعرفية البسيطة، وأنها لا تصلح لقياس عمليات التفكير بما تتضمنه من مستويات الفهم والتحليل والتطبيق، وأن اختبارات المقال هى المقاييس المناسبة لعمليات التفكير. إلا أن مجموعة كبيرة من علماء القياس يقررون أن

الفقرات الموضوعية objective items يمكن أن تستخدم لقياس أنواع مختلفة من عمليات التفكير، إلا أن إعداد مثل هذا النمط من الفقرات يستغرق وقتاً وجهداً كبيرين.

أنماط فقرات الاختبارات الموضوعية:

تتضمن الاختبارات المعرفية الموضوعية مجموعة مختلفة من أنماط الفقرات هى : الصواب - الخطأ، الاختيار من متعدد، التكميل، المزاوجة، وفيما يلى شرح موجز لكل نمط من هذه الأنماط.

١) فقرات الصواب - الخطأ The true - false items

وهى فقرات تقدم للمفحوص فى شكل جملة أو عبارة أو مشكلة أو مسألة ويطلب منه أن يحدد صحة الفقرة بوضع علامة (صح - ✓) أو (خطأ - ×) أو وضع خط أسفل كلمة صح أو خطأ.

وتستخدم فقرات الصواب - الخطأ فى الاختبارات المعرفية عند قياس تذكر المعلومات عن طريق اختيار الإجابة الصحيحة لحقائق أو مصطلحات أو مبادئ محددة فى مجال محدد.

معنى هذا أن فقرات الصواب - الخطأ عبارة عن صياغات (جمل) تقريرية صريحة تعطى للمفحوص لكى يقرر ما إذا كانت الجملة صحيحة أو غير صحيحة. ويلاحظ أن هذا النمط من الفقرات يستخدم على نطاق واسع فى الاختبارات التى يعدها معلم الفصل teacher-made tests لسرعة كتابة الفقرات وسهولة تقدير الدرجة.

مميزات فقرات الصواب - الخطأ:

- تستخدم لتغطية مدى كبير وواسع من المادة الدراسية عن طريق إجراء القياس لمرة واحدة فقط. فالوقت المطلوب لكى يستجيب المفحوص على فقرات الصواب - الخطأ أقل من الوقت المطلوب للاستجابة على فقرات

الاختيار من متعدد أو التكميل، لذا يمكن تضمين اختبارات الصواب - الخطأ عدد كبير من الفقرات يفوق أى نمط آخر من الاختبارات.

- سهولة التصحيح وتقدير الدرجات.
- سهولة إعداد الفقرات وتطبيقها وبخاصة عند استخدامها لقياس تذكر المعلومات، إلا أنها قد تستغرق وقتا أطول عندما تعد لقياس عمليات التفكير المركبة.

عيوب فقرات الصواب - الخطأ:

- تتأثر بالتخمين العشوائى random guessing لكونها تتضمن إجابتين محتملتين فقط، حيث ينتج عن هذا التخمين العشوائى أن يصبح لدى المفحوص فرصة ٥٠٪ للحصول على إجابات صحيحة نتيجة التخمين.
- نظراً لوجود فرصة للحصول على ٥٠٪ من الإجابات الصحيحة نتيجة التخمين، لذا يتجه ثبات درجات اختبارات الصواب - الخطأ نحو الانخفاض.
- غالبا ما تتوقف (تعتمد) الإجابة الصحيحة على كلمة واحدة فقط مما يزيد من احتمالات الوقوع فى الخطأ.

توجيهات عامة لكتابة فقرات الصواب - الخطأ:

يمكن استخدام فقرات الصواب - الخطأ بفاعلية إذا اتبع عند إعدادها الآتى:

- تجنب استخدام العبارات التى توحى بإجابات معينة، وتجنب استخدام التعميمات فى الزمان أو المكان مثل: دائما - أبدا - فى كل وقت - أحيانا - عادة إلى غير ذلك من الكلمات.
- يجب أن يكون عدد الفقرات الصواب مساويا لعدد الفقرات الخطأ، أو أن يكون عدد الفقرات الخطأ أكبر من عدد الفقرات الصواب، لأن للفقرات الخطأ قدرة على التمييز أكبر من الفقرات الصواب.

- تجنب استخدام لغة الكتاب المدرسى لكونها لغة جافة قد تتميز بالصرامة.
- تجنب استعمال الصيغ السلبية كأن تقول «أنا لا أنام فى الليل كثيراً..» ولكن الأفضل فى الصياغة هو « أنا أنام فى الليل قليلا...»، وذلك لكون الصياغات السلبية ربما تربك المفحوص، وقد لا تسهم بفاعلية عند قياس القدرات. وينصح فى حالة الحاجة إلى استخدام العبارات السلبية أن يوضع خط أسفل المصطح السلبى كالتالى «أنا لا أنام فى الليل كثيرا..».
- تجنب استخدام الكلمات الغامضة والتى قد تحمل أكثر من معنى مثل كلمة «ثقافة» لكون مثل هذه الكلمات تثير الشك والحيرة لدى المفحوص، فالأساس فى الصياغات (الجمل) بالنسبة لمثل هذا النوع من الفقرات هو الآتى:

«ألا تحمل الجملة أى شك فى أنها صحيحة تماماً أو خطأ تماماً».

- العمل على أن يكون لجميع الفقرات نفس الطول تقريبا، حيث يلاحظ فى بعض الاختبارات المعرفية أن الجمل (الفقرات) التى لها إجابات صحيحة تكون أطول من الجمل التى لها إجابات خاطئة.
- أن تسأل الفقرة (السؤال أو الجملة) عن مفهوم أو موضوع أو فكرة واحدة مفردة فقط. فالفقرات التى تشتمل على أكثر من مفهوم أو موضوع أو حقيقة علمية غالبا ما تؤدى إلى إرباك المفحوصين.

أمثلة لفقرات الصواب - الخطأ:

مثال (٤ ـ ١): اقرأ بعناية كل جملة من الجمل التالية، ثم ضع علامة صح (✓) أمام الجملة التى تعتقد أنها صحيحة، وعلامة خطأ (×) أمام الجملة التى تعتقد أنها خطأ (لكل إجابة صحيحة ٣ درجات).

- ١) إذا زاد الانحراف المعيارى (SD) لدرجات المجموعة، يزيد بالتالى التشتت.
- ٢) الدرجة المعيارية ت - T التى تساوى ٦٠ تقابل وحدة انحراف معيارى واحدة فوق المتوسط الحسابى.
- ٣) معامل الارتباط +٦٠,٠ يساوى مرتين معامل الارتباط +٣٠,٠.
- ٤) إذا كان منحنى توزيع البيانات متجها نحو اليمين كان الالتواء سالباً.

مثال (٤-٢): الجمل التالية تصف بعض المواقف فى رياضة التنس. والمطلوب: هو أن تقرأ الجمل جيداً، ثم تضع علامة صح(✓) أمام الجملة التى تعتقد أنها صحيحة، وعلامة خطأ أمام الجملة التى تعتقد أنها خطأ (×) (لكل إجابة صحيحة ٣ درجات):

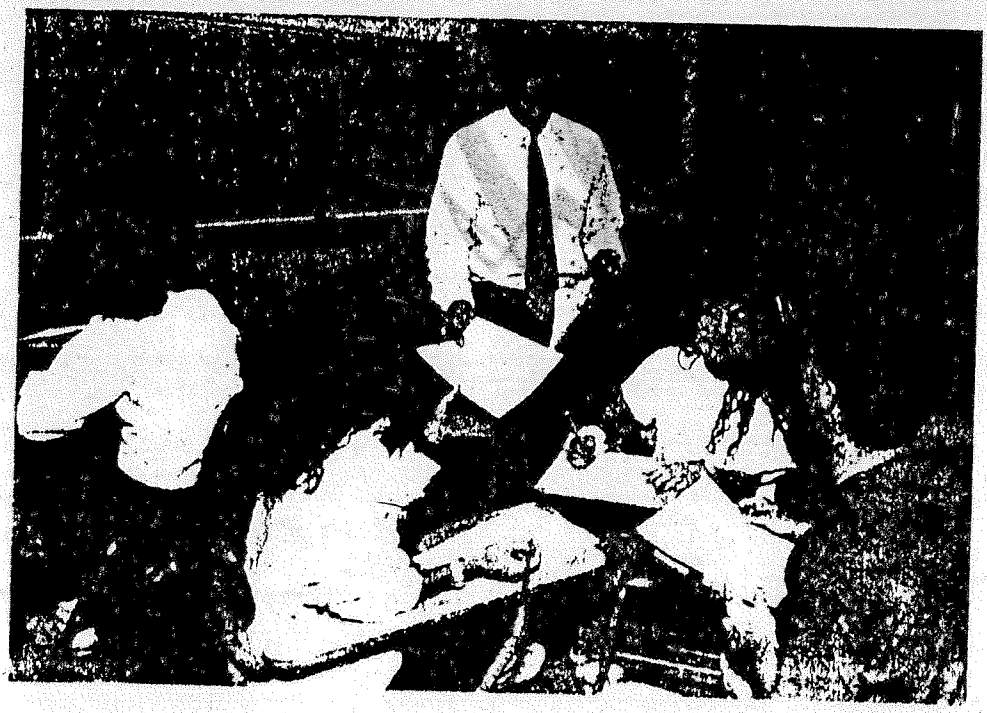
«اللاعب (أ) قام بأداء إرسال جيد، فقام اللاعب (ب) برد الإرسال بنجاح، فأى اللاعبين يمكن أن يفقد (نقطة واحدة) إذا:

- ١) ارتدت الكرة مرتين فى ملعبه.
- ٢) اللاعب ضرب الكرة فلمست الشبكة قبل الهبوط فى ملعب خصمه.
- ٣) بعد ضرب الكرة سقط المضرب من يد اللاعب.
- ٤) اللاعب ضرب الكرة قبل أن تعبر الشبكة فى اتجاه ملعبه.

٢) فقرات الاختيار من متعدد

The Multiple - Choice items (MCq)

وفيها يعطى المفحوص السؤال يتبعه عدة إجابات إحداها فقط صحيحة، وعليه أن يختار الإجابة التى يعتقد هو أنها صحيحة. فالإجابات فى هذا النمط من الاختبارات يكون فيها إجابة واحدة صحيحة وعدد من الإجابات



القياس المعرفى يستهدف قياس تحصيل الأهداف المعرفية



اختبارات الاختيار من متعدد يمكن تطبيقها على أعداد كبيرة من المفحوصين

غير صحيحة incorrect responses والتى تعرف بالإجابات «المضللة dis tractor» وهى إجابات تعد بحيث تغرى المفحوص باختيارها فى اختبارات الاختيار من متعدد أو اختبارات المزاوجة.

وتستخدم فقرات الاختيار من متعدد لقياس نواتج التعلم بالنسبة لمستويات متبانية من المعرفة بدءا بالمستويات البسيطة وحتى المستويات المركبة، كما تستخدم عند قياس التحصيل بالنسبة للعديد من الموضوعات الدراسية.

وتتطلب فقرات الاختيار من متعدد - فى بعض الأحيان - أن يختار المفحوص «أفضل إجابة - the best answer » من بين مجموعة من الإجابات الصحيحة التى تتطلب من المفحوص اختيار أفضلها، حيث يعنى ذلك أنه يطلب من المفحوص أن يختار من بين مجموعة من الاستجابات الصحيحة أفضل الاستجابات المناسبة للإجابة على السؤال الذى تتضمنه الفقرة.

ومن الملاحظ أن فقرات الاختيار من متعدد MCq تتضمن مكونين رئيسيين هما:

١ - الأرومة (البند الرئيسى) Stem : وهى عبارة عن سؤال مباشر أو جملة غير تامة (غير كاملة) تعرض لمشكلة تفاصيلها غير كافية ولا يوجد بها أى غموض فيما يتعلق بالشىء الذى تسأل عنه.

٢ - الاستجابات Responses: وهى تتراوح من ٣ إلى ٥ استجابات واحدة منها صحيحة وبقية الاستجابات الأخرى غير صحيحة، حيث يلاحظ أن الاستجابات غير الصحيحة تشير إلى إجابات خاطئة تغرى المفحوص باختيارها، حيث تعرف فى مجال القياس المعرفى باسم «المضلل أو المشتت - Distractor».

مميزات فقرات الاختيار من متعدد:

- تستخدم لقياس القدرة على الفهم.
- تصلح لقياس أنماط مختلفة من المواد الدراسية.
- لا تستغرق وقتا طويلا فى التطبيق وتصحيح النتائج.
- فرص التخمين بالنسبة للإجابات الصحيحة أقل بكثير من فقرات الصواب - الخطأ.
- تستخدم كثيراً فى الاختبارات المعرفية فى مجالات: الرياضيات والهجاية ومفردات اللغة وغيرها.

عيوب فقرات الاختيار من متعدد:

- صعوبة إعدادها بالمقارنة بالفقرات الأخرى، حيث يستغرق إعدادها وقتا أطول وذلك لكونها تتطلب من (٣ - ٥) استجابات لكل فقرة (سؤال أو جملة).
- تشجع - فى بعض الأحيان - على استظهار (الحفظ عن ظهر قلب) memorization الحقائق أكثر من فهم understanding تلك الحقائق.
- تمتاز اختبارات الاختيار من متعدد بأنه يمكن تضمينها عدد أقل من الفقرات بالمقارنة باختبارات الصواب - الخطأ، وذلك عند قياس نفس الموضوع أو المادة الدراسية، لذلك فهى تستغرق وقتا أقل.

توجيهات عامة لكتابة فقرات الاختيار من متعدد:

مع أن إعداد فقرات الاختيار من متعدد من الصعوبة بمكان، إلا أنه يمكن الاسترشاد بالتوجيهات التالية عند إعدادها:

أولاً: بالنسبة للأرومة (البند الرئيسى) Stem:

١- يجب أن يكون البند الرئيسى مختصرا وسهل القراءة والفهم، وأن يحتوى على المسألة أو القضية الرئيسية للفقرة، كما يشترط عدم

تكرار أية كلمات يتضمنها البند الرئيسى فى الاستجابات، كما ينصح بتجنب استخدام أية مواد (موضوعات) غير ذات صلة irrelevant بالأرومة (البند الرئيسى).

كما يفضل أن تصاغ الأرومة فى جملة أو سؤال غير كامل حتى يتمكن المفحوص من تكميله عن طريق الاستجابة.

٢- يجب تجنب التعميمات مثل: دائما - مطلقا - أبدا - كلية وغيرها.

٣- يجب تحديد الأرومة (البند الرئيسى) بكلمات وتعبيرات موجبة Positive words ، وفى حالة الكلمات السالبة فإنه يفضل وضع خط أسفل هذه الكلمات أو كتابتها ببنط كبير يميزها.

٤- يفضل أن تبدأ الأرومة (البند الرئيسى) بالحرف «W» أى بكلمات مثل: أى - أية which، لماذا why ، أين where ، ما - ماذا what، متى when، من who ، لأن مثل هذه الكلمات تقدم للمفحوص الموضوع الرئيسى أو النقاط الرئيسية للفقرة أو السؤال.

٥- لا يجب تضمين الأرومة (البند الرئيسى) أية كلمات تستهدف سؤال المفحوص عن رأيه opinion، وإنما المطلوب هو أن يحدد الإجابة الصحيحة.

ثانيا: بالنسبة للاستجابات:

١ - يجب أن تكون جميع الاستجابات معقولة وقابلة ظاهريا للتصديق plausible ، وأن توجد بين هذه الاستجابات استجابة واحدة فقط صحيحة، كما يجب عدم تضمين الاختبارات أية استجابات هزلية أو ضعيفة.

٢ - يجب ألا يقل عدد الاستجابات عن ثلاث استجابات وألا يزيد عن خمس استجابات بالنسبة لكل فقرة، وفى جميع الحالات يفضل ألا يقل عدد الاستجابات عن أربع استجابات لأن هذا العدد يجعل عامل التخمين guesing factor عند حدود ٢٠٪ .

٣- يجب أن تكون جميع الاستجابات متسقة فى المحتوى والصياغة وقواعد النحو ولها نفس الطول تقريبا. ويفضل تجنب تضمين الاستجابة الصحيحة معلومات أكثر من الاستجابة الخاطئة.

٤- إذا أعطيت الفقرات الأرقام (١, ٢, ٣, ٤، . . . إلخ)، فإنه يلزم استخدام الحروف (أ، ب، جـ، د، هـ. . .) كأسلوب لترقيم الاستجابات، كما يفضل أن توضع الاستجابات مرتبة رأسيا لأن هذا الأسلوب يساعد على قراءتها بسهولة أكثر من وضعه فى نظام الترتيب الأفقى.

٥- تجنب إعطاء الاستجابات الصحيحة نمط ثابت من الترقيم، كأن تعطى جميع الاستجابات الصحيحة فى الاختبار الحرف (جـ) ولكنه يفضل أن يتم توزيع جميع الاستجابات الصحيحة فى الاختبار على الحروف (أ، ب، جـ، د، هـ) بالتساوى تقريبا.

أمثلة لفقرات الاختيار من متعدد:

ضع الحرف الذى يدل على الإجابة الصحيحة على الخط الأسود الموجود على يمين رقم الفقرة (كل فقرة لها درجتان).

- ١) فى اختبار معرفى، كان للدرجة ٥٥ رتبة مئينية ٤٠٪ ما دلالة ذلك؟

أ - أن ٤٠ تليمذا أجابوا على ٥٥٪ من أسئلة الاختبار إجابات صحيحة.

ب- أن ٤٠ ٪ من التلاميذ لديهم الدرجة ٥٥ أو أقل.

جـ- أن ٥٥٪ من التلاميذ لديهم الدرجة ٤٠ أو أقل.

د - أن ٤٠٪ من التلاميذ لديهم الدرجة ٥٥ أو أكثر.

- ٢) إذا كان معامل الارتباط بين متغيرين هو - ٨٠, ٠ فمعنى ذلك أن:

أ - العلاقة بين المتغيرين منخفضة.

ب - المتغيرين غير مترابطين.

جـ - النقص فى قيم أحد المتغيرات يتبعه نقص فى قيم المتغير الآخر.

د - الزيادة فى قيم أحد المتغيرات يتبعها نقص فى قيم المتغير الآخر.

- ٣) أى العضلات تعمل على مد الساق وقبض الفخذ ؟

أ - العضلة التوأمية (خلف الساق).

ب - العضلة الفخذية ذات الرأسين.

جـ - العضلة الإلية الكبيرة.

د - العضلة الفخذية ذات الأربع رؤوس.

-٤) ما هو بالتقريب الحد الأقصى لمعدل القلب لفرد عمره ٢٠ سنة؟

أ - ١٨٠ نبضة.

ب- ١٩٠ نبضة.

جـ- ٢٠٠ نبضة.

د - ٢١٠ نبضة.

٣) فقرات الإجابة القصيرة والتكميل

Short - answer and completion items

يكمن الفرق بين فقرات الإجابة القصيرة والتكميل فى نقطتين أساسيتين هما: الطول والشكل (البنية) format. فالفقرات فى الإجابة القصيرة تتطلب من المفحوص أن يجاوب على السؤال فى كلمة word أو جملة sentence أو بهما معاً. بينما فى فقرات التكميل تحذف كلمة أو عدد من الكلمات من الجملة، ويطلب من المفحوص استكمال الكلمات الناقصة. وتصلح فقرات الإجابة القصيرة وفقرات التكميل لقياس المعرفة الحقيقية factual knowledge والفهم والقدرة على تحديد وتعريف المفاهيم والمدركات.

ففى فقرات الإجابة القصيرة يطلب من المفحوص تحديد أو تسمية شىء ما فى كلمة أو جملة أو رقم، ومن أمثلة ذلك ما يأتى:

- من هو أول رئيس للجمهورية فى مصر؟

- $\frac{١}{٨} + \frac{١}{٤} + \frac{١}{١٦} =$

- ما هى الغدة التى تنتج هرمون ACTH فى جسم الإنسان؟

مميزات فقرات الإجابة القصيرة والتكميل:

- أقل تأثراً بالتخمين من الصواب - والخطأ أو الاختيار من متعدد.
- تصلح لتقويم تذكر المعلومات أكثر من أى نمط آخر من الفقرات.
- يسهل إعدادها.

عيوب فقرات الإجابات القصيرة والتكميل:

- يستغرق حساب الدرجات فيها وقتا أطول من فقرات الاختيار من متعدد، لكونها تتطلب تصحيح مدى دقة وسلامة اللغة المستخدمة فى الإجابة.
- قد تكون بعض الإجابات صحيحة ككل أو صحيحة جزئيا مما يضع مقدر الدرجة فى حيرة بالنسبة لتقدير درجات كل فقرة، وبناء على ذلك يرى علماء القياس أن الشخص القادر على تقدير الدرجات على اختبارات الإجابات القصيرة والتكميل هو الشخص الذى يقوم ببناء الاختبار.
- تشجع على التعلم الصمّى rote Learning وهو تعلم يقوم على التكرار بقصد الاستظهار (الحفظ) دون الاهتمام بالفهم وغيره من المستويات المعرفية، ومع أن هذا النمط من الفقرات يتضمن العديد من الفرص الملائمة للاسترجاع والحفظ عن ظهر قلب memorization فإننا نجدها

مناسبة لإعداد الاختبارات بالنسبة لبعض المجالات مثل الإسعافات الأولية وبالنسبة لعمليات إنقاذ الغرقى والإنعاش فى حالات الأزمات القلبية.

توجيهات عامة لكتابة فقرات الإجابات القصيرة والتكميل:

تتميز فقرات اختبارات الإجابات القصيرة والتكميل بأنها أسهل فى إعدادها من غيرها من فقرات الاختبارات الموضوعية الأخرى، ومع ذلك فإنه يمكن الاسترشاد بالتوجيهات التالية عند إعداد مثل هذا النوع من الفقرات:

١- يجب العمل على أن تكون الإجابة على الفقرة بكلمة واحدة مفردة أو بعدد من الكلمات وأنه توجد إجابة واحدة صحيحة فقط.

٢- التأكد من أن جميع المفحوصين يعرفون ما هو نمط الاستجابة المطلوب، وكذا المؤشرات الخاصة بدقة الاستجابة وبخاصة فى الحالات التى تستخدم فيها أجزاء الدرجة أو الكسور العشرية.

٣- يجب على من يقوم بإعداد مثل هذا النمط من الاختبارات أن يفكر أولا فى الإجابة، ثم يحاول بعد ذلك كتابة الفقرة التى تمثل الإجابة التى أعدها، فعن طريق هذا الإجراء يمكن تجنب إعداد فقرات لها أكثر من إجابة صحيحة.

٤- فى فقرات التكميل، يفضل تخصيص منطقة فراغ خالية من الكتابة قرب نهاية الجملة، حيث يفيد هذا الإجراء فى جعل الجملة أكثر وضوحاً وتجنب وضع أكثر من إجابة للجملة.

٥- يجب عدم استخدام أكثر من منطقتين (فراغين) فى الجملة الواحدة، لأن وجود أكثر من فراغين فى الجملة يزيد من غموضها مما يؤدى إلى اضطراب المفحوص.

٦- عدم نقل الفقرات مباشرة من الكتاب المدرسى، لأن الجملة التى تؤخذ من فقرة فى كتاب قد لا تعبر بشكل مناسب عن المحتوى المراد التعبير عنه.

٧- أن تكون جميع الفراغات متساوية الطول بالنسبة لجميع الاستجابات، لأن الاختلاف فى أطوال الفراغات وفقا لأطوال الإجابات المتوقعة، قد يعطى المفحوص إيحاءات بأهمية الإجابة، ولتسهيل تقدير الدرجة والحد من الأخطاء، يجب وضع خطوط سوداء قصيرة فى الفراغات الخاصة بكل جملة لكى يسجل عليها المفحوص استجاباته.

أمثلة لفقرات الإجابة القصيرة:

١- ما هى المقاييس الثلاثة للنزعة المركزية؟

٢- أظهرت دراسة لقوة الرجل وقوة القبضة ارتباط مقدارة (-٠,٩٠) فسر هذا الارتباط؟

٣- فى توزيع للدرجات له متوسط حسابى ٢٢ وانحراف معيارى ٢، ما هى الدرجة المئوية للدرجة ١٦ والدرجة ٢٨؟

أمثلة لوحدات التكميل:

أكمل الفقرات التالية بكتابة الإجابة الصحيحة على الخط الأسود الموجود داخل كل فقرة (كل فقرة لها درجتان)

- ١) يستخدم ـــــــ كواحد من أهم مقاييس التشتت

- ٢) الدرجة التائية ت - T لها متوسط حسابى هو ـــــــ وانحراف معيارى هو ـــــــ .

- ٣) المقياس Z-Scale له متوسط حسابى هو ـــــــ وانحراف معيارى هو ـــــــ .

٤) **فقرات المزاوجة (المواءمة)** Matching items

تعتمد اختبارات المزاوجة (المواءمة أو التوفيق) على المواءمة بين عدة فقرات (بنود) فى قائمتين، حيث يطلب من المفحوص أن يربط بين كل فقرة

فى القائمة الأولى مع ما يناسبها من القائمة الثانية، ويحدد السؤال طبيعة المواءمة، وهل هى ترادفاً أو تضاداً أو علاقة سببية أو علاقة زمنية أو علاقة مكانية أو غيرها.

معنى ذلك أن اختبار المزاوجة matching test يتكون من عمودين (أيمن، وأيسر) بكل عمود مجموعة من الفقرات (الكلمات أو الجمل) التى تستخدم كمثيرات، حيث يلاحظ أن الفقرات فى العمود الأيمن متزاوجة paired مع الفقرات فى العمود الأيسر. وتمثل الفقرات فى الجهة اليمنى المقدمات المنطقية premises للسؤال، بينما تمثل الفقرات فى الجهة اليسرى الاستجابات على السؤال.

وتشبه فقرات المزاوجة (المواءمة) فقرات الاختيار من متعدد، فى كونهما يعطيان المفحوص حق الاختيار فيما بين مجموعة من البدائل alternatives كما أنها تشبه فقرات الإجابة القصيرة لكونها تتطلب استخدام معلومات حقيقية ومختصرة ومحددة، هذه المعلومات قد تكون أسماء أو تواريخ أو أرقاماً أو ألقاباً.

مثال (٤ - ٣):

أمامك قائمتان، القائمة اليمنى تتضمن بعض الأساليب الإحصائية، بينما تتضمن القائمة اليسرى أسماء العلماء، الذين ابتكروا هذه الأساليب الإحصائية.

والمطلوب: المزاوجة بين اسم العالم والأسلوب الإحصائى الذى ابتكره عن طريق توصيل خط بينهما.

(ب) معامل الارتباط التتابعى (ppm)	(أ) فيشر
(جـ) معامل ارتباط فروق الرتب (Rho)	(ب) بيرسون
(د) معامل الاتفاق المطابقة (W)	(جـ) سبيرمان
(أ) النسبة الفائية (F-R)	(د) كندال

يلاحظ من المثال السابق وجود قائمتين، الأولى تتضمن مجموعة من المقدمات، والثانية تتضمن أسماء مجموعة من العلماء، حيث يلاحظ أن القائمتين متساويتين فى العدد.

وعادة ما يطلب من المفحوصين المزاوجة بين الفقرات فى القائمتين، هذه الفقرات قد تكون أحداث وتواريخ، أو أحداث وأسماء ، أو كتب وأسماء مؤلفين، أو مصطلحات وتعاريف إلى غير ذلك. لهذا تعد فقرات المزاوجة (المواءمة) من الوسائل التى تقيس قدرة المفحوص على ربط شىء بآخر.

ومن مميزات فقرات المزاوجة مايأتى:

- يمكن الاستفادة منها فى قياس كمية كبيرة من المعلومات الحقيقية المرتبطة بالمادة الدراسية فى وقت قصير.
- يسهل إعدادها وتقدير درجات الاستجابة عليها بالمقارنة بفقرات الاختيار من متعدد.
- تثير دافعية المفحوصين نحو المعرفة المتكاملة عن طريق التفكير فى العلاقة فيما بين الفقرات.
- يظهر التخمين بالنسبة للإجابات الصحيحة أقل بالمقارنة بالأنماط الأخرى من الفقرات.

ومن عيوب فقرات المزاوجة ما يأتى:

- تستخدم فقط لقياس المعرفة الحقيقية المتعلقة بالحقائق والأسماء والأرقام الجافة.
- محدودة الاستخدام بالنسبة للموضوعات الدراسية، لكونها لا تصلح بالنسبة لكل الموضوعات.
- المغالاة فى التركيز على الحفظ عن ظهر قلب للعلاقات البسيطة والاستجابات الصحيحة لحقائق علمية.

توجيهات عامة لكتابة فقرات المزاوجة (المواءمة):

لبناء فقرات المزاوجة بكفاية ووضوح تتبع الإرشادات التالية:

١- يجب أن يكون عدد الإجابات البديلة (الاستجابات أو الاختيارات) أكبر من عدد المقدمات (فى العمود الأول) وذلك لضمان الحد من احتمالية تأثير المزاوجة وبخاصة بالنسبة للزوجين الأخيرين من الوحدات، فلو كان عدد الوحدات فى القائمة الأولى أربع وحدات، وعدد الاختيارات فى القائمة الثانية أربع اختيارات، واستطاع المفحوص أن يعرف الإجابات الصحيحة لأى ثلاثة أزواج من الفقرات، حينئذ يستطيع تلقائيا أن يجيب الإجابة الصحيحة على الزوج الأخير من الفقرات، أما فى حالة ما إذا كان عدد الفقرات فى القائمة الثانية (قائمة الاستجابات أو الاختيارات) خمس أو ست فقرات مثلا، فى هذه الحالة تقل احتمالية التصحيح الذاتى للفقرة.

٢- يجب أن تكون أسس المزاوجة والاختيارات الخاصة بكل فقرة واضحة فى التعليمات المعطاة، بمعنى أن توضح التعليمات المعطاة للمفحوص عما إذا كان لكل فقرة إجابة واحدة صحيحة، أو أكثر وأيضا كيفية عمل المزاوجة بين الفقرات.

مثال (٤ - ٤):

أمامك قائمتان، الأولى تتضمن بعض الدول، والثانية تتضمن الألعاب الرياضية التى قامت بابتكارها هذه الدول. والمطلوب المزاوجة بين أسماء الأنشطة وأسماء الدول التى ابتكرتها وذلك عن طريق اختيار الحرف المناسب من العمود الثانى ووضعه كتابة على الخط الموجود أمام الإجابة المناسبة فى العمود الأول (كل إجابة صحيحة لها ٣ درجات).

(العمود الأول)	(العمود الثانى)
أ، ب ١) الصين	أ - المصارعة
جـ ٢) الأغريق القدماء	ب - الملاكمة
و ٣) الهند	جـ - الميدان والمضمار
د ٤) ألمانيا	د - الجمباز
هـ ، ز ٥) انجلترا	هـ - كرة القدم
	و - الريشة الطائرة
	ز - التنس

يلاحظ من هذا المثال الآتى:

- عدم تساوى الفقرات فى العمودين (٥ فقرات فى العمود الأول) فى مقابل (٧ اختيارات فى العمود الثانى).
- تتطلب طريقة المزاوجة فى هذا المثال كتابة الحرف المعبر عن الاختيار المناسب على الخط الموجود أمام الفقرة.
- يلاحظ أنه يوجد أكثر من اختيار واحد للفقرة، فالإجابات الصحيحة لهذا السؤال تبين أن الفقرة (١) لها إجابتان صحيحتان هما (أ، ب)، وأن الفقرة (٥) لها أيضاً إجابتان صحيحتان هما (هـ، ز)، وأن لكل من الفقرات أرقام (٢، ٣، ٤) إجابة واحدة صحيحة.

٣ - يجب وضع كل الفقرات والاختيارات (الاستجابات) فى صفحة واحدة وفى عمودين متقابلين، ولا يجب بأى حال من الأحوال توزيعها على صفحتين، لأن تنظيمها فى عمودين فى صفحة واحدة يساعد المفحوص على الإجابة فى أقصر وقت، كما يحد من وقوعه فى الخطأ نتيجة التقليب فى الصفحات.

٤ - يجب استخدام الشكل المناسب لتقديم فقرات المزاوجة، حيث يفضل أن توضع الفقرات فى عمود على اليمين، وتوضع الاختيارات (الاستجابات) فى عمود على اليسار، ولتسهيل عملية الاستجابة، وتقدير الدرجة، يعمل خط أسود أمام الفقرة ليكتب عليه الحرف الذى يعبر عن الاختيار المناسب.

٥ - يتم ترتيب الاختيارات (الاستجابات) بإعطائها الحروف الأبجدية (أ، ب، جـ، .. إلخ). أو أى ترتيب منطقى آخر.

٦ - يفضل أن يكون هناك أكثر من اختيار (استجابة) لكل فقرة، فوجود اثنين أو ثلاثة أو أكثر من الاختيارات بالنسبة للفقرة الواحدة يحد من تأثير التخمين guessing.

٧ - يفضل أن يكون عدد الفقرات فى القائمتين مناسباً للمرحلة العمرية التى يعطى لها الاختبار، فيكون فى المرحلة الابتدائية من ٥ إلى ٦ فقرات، ويزيد هذا العدد مع تقدم السن ليصبح من ١٠ إلى ١٥ فقرة، وفى جميع الحالات يجب ألا يكون عدد الفقرات والاختيارات طويل جداً، وألا يكون قصير جداً، لأنه فى حالة ما إذا كان عدد الفقرات والاختيارات (الاستجابات) طويلاً جداً، فإنه يحتمل أن تفقد تجانسها بالإضافة إلى أنها سوف تستغرف وقتا طويلا فى التطبيق، وإذا كان عدد الفقرات قصيراً جداً، حينئذ يزيد تأثير عامل التخمين وينخفض صدق مزاوجة (مواءمة) الفقرة the validity of the matching item .

ثالثا : تطبيق الاختبار:

بعد الانتهاء من إعداد الصورة الأولية للاختبار المقترح، يبدأ الباحث فى تطبيق الاختبار وفق مجموعة من التدابير والإجراءات يمكن تحديدها على النحو التالى:

إجراءات تتخذ قبل تطبيق الاختبار : وتشمل الآتى:

١ - الإعداد لتنفيذ الاختبار : ويتضمن الآتى :

- تجهيز عدد كاف من بطاقات الإجابة (نسخ مكتوبة من الاختبار).
- توفير كافة الأدوات اللازمة لتطبيق الاختبار مثل: الأقلام، وساعة الإيقاف، وأوراق المسودات وغيرها من الأدوات.

ويفضل إعداد هذه الأدوات قبل تطبيق الاختبار بأسبوع على الأقل، لإتاحة الفرصة أمام المسئولين لمراجعة متطلبات التطبيق، والتغلب على أية مشكلات يتوقع حدوثها. كما يفضل تخزين مثل هذه الأدوات فى مكان أمين وقريب من مكان تطبيق الاختبار.

٢ - اختيار الموقع المناسب لتطبيق الاختبار:

ويفضل أن تكون غرفة من غرف الصف الدراسى المعتادة فى حالة تطبيق الاختبار على عينات من التلاميذ، وعموما يشترط فى الغرفة أو القاعة أو الصالة التى يتم فيها تطبيق الاختبار الآتى :

- أن تكون مناسبة لعدد المفحوصين بحيث يجلسون وبينهم مسافات كبيرة.
- أن تكون مناسبة من حيث المساحة والتهوية والإضاءة والحرارة حتى تحقق للمفحوصين الراحة البدنية والنفسية.
- يفضل أن يكون مكان تطبيق الاختبار (غرفة الصف أو غيرها) بعيداً عن الشوارع أو أية مصادر للضوضاء يمكن أن تشتت انتباه وتركيز المفحوصين.
- يتم جلوس المفحوصين بطريقة تسمح لهم رؤية وسماع المسئول عن إدارة الاختبار (الباحث)، وأن تسمح للباحث بالتحرك فيما بين المفحوصين بسهولة عند الرد على تساؤلاتهم واستفساراتهم.

٣ - توفير الاحتياطات المناسبة لتطبيق الاختبار، مثل:

- عمل لوحات إرشادية بموعد وأماكن تنفيذ الاختبار.

- عمل لوحات إرشادية تفيد بعدم دخول الطلاب (اللاعبين) أو الزوار إلى المنطقة التى يتم فيها تنفيذ الاختبار.

- التنبيه بعدم استخدام جرس المدرسة أو مكبرات الصوت أو إقامة المسابقات الرياضية أثناء تطبيق الاختبار.

إجراءات تتخذ أثناء تطبيق الاختبار: وتشمل الآتى:

١- قراءة تعليمات الاختبار قبل البدء فى الإجابة على الأسئلة.

٢- بدء الإجابة على الاختبار وحساب الزمن الذى يستغرقه التطبيق بدقة، حيث يتم تشغيل ساعة الإيقاف عند إعطاء اشارة البدء، وإيقاف الساعة عند انتهاء الزمن المقرر للاختبار كالتالى:

	الساعة	الدقيقة	ثانية
- زمن بدء الاختبار	٢	١٠	٠
- الزمن المسموح به		٢٠	٠
- ينتهى زمن الاختبار	٢	٣٠	٠

٣- يجلس المسئول عن إدارة الاختبار (الباحث) فى المكان المعتاد الذى يجلس فيه معلم الفصل، ويفضل عدم جلوسه بجوار أى مفحوص، أو التحرك أثناء الإجابة على الاختبار.

٤- يقوم المسئول عن إدارة الاختبار بالرَّد على أية أسئلة أو استفسارات يقدمها المفحوصين. ولتجنب تشتت انتباه المفحوصين، يتم الرد على كل مفحوص على حدة وهو جالس فى مكانه، وبحيث يكون الرد بصوت منخفض لا يصل إلى بقية زملائه.

٥- حث المفحوصين على الأداء لأقصى ما تمكنهم قدراتهم لتحقيق أفضل نتائج ممكنة.

٦- تسجيل كل الأحداث والملاحظات المهمة التى تحدث خلال التطبيق مثل التوتر الزائد للمفحوصين، السرحان والنظر إلى النافذة لفترات زمنية طويلة، أو الإجابة بطريقة عشوائية على الأسئلة، إلى غير ذلك من الملاحظات التى يمكن الاستفادة منها عند إعادة تطبيق الاختبار.

إجراءات تتخذ بعد تطبيق الاختبار، وتشمل:

١- تصحيح الاختبار وتقدير درجات كل مفحوص على حدة.

٢- رصد الدرجات فى بطاقة تسجيل جمعية توطئة لإجراء المعالجات الإحصائية المناسبة، والتى تتضمن التحليل الإحصائى للفقرة، وهو ما سوف نعرض له فى الخطوة التالية (الخطوة الرابعة) من خطوات بناء الاختبار المعرفى.

رابعاً: تحليل الفقرة:

بعد كتابة فقرات الاختبار ومراجعتها بعناية وإعدادها مكتوبة وتطبيقها لتوفير البيانات الأولية اللازمة للتحقق من نوعية الفقرات quality of items وذلك من خلال ما يعرف بتحليل الفقرة item analysis.

ما هو المقصود بتحليل الفقرات؟

تحليل الفقرات هو الدراسة التى تعتمد على التحليل المنطقى الإحصائى والتجريبى لوحدات الاختبار لمعرفة خصائصها وحذف أو تعديل أو إبدال أو إضافة أو إعادة ترتيب هذه الفقرات حتى يتسنى الوصول إلى اختبار ثابت وصادق ومناسب من حيث الطول ومستوى الصعوبة.

الأهداف الرئيسية لتحليل الفقرة:

١- الكشف عما إذا كانت الفقرات صعبة جداً، أو سهلة جداً، أو متوسطة الصعوبة، حيث يكشف التحليل الإحصائى عن مستوى صعوبة الفقرة difficulty level أو ما يعرف بمؤشر الصعوبة index of difficulty.

٢- الكشف أيضاً عما إذا كان للفقرة قدرة على التمييز بين الأفراد المتميزين والأفراد غير المتميزين، حيث يكشف التحليل الإحصائى عن القوة التمييزية للفقرة discriminative power أو ما يعرف بمؤشر (دليل) التمييز the index of discimination وهو مؤشر يدل على صدق الفقرة item validity.

٣- إظهار لماذا فقرة معينة لا تؤدى وظيفتها بكفاية، وكيف يمكن تعديل هذه الفقرة لتصبح أكثر فعالية.

وتتوقف الإجراءات الإحصائية لتحليل الفقرة على أنماط الفقرات المستخدمة فى الاختبار، وعلى عدد المفحوصين، والغرض من التحليل، فالثقة تزداد فى الفقرة عندما يتم التحليل الإحصائى على مائة مفحوص أو أكثر، وتنخفض عندما يقل العد عن مائة.

ويهتم تحليل الفقرة فى الاختبارات المعرفية بثلاثة مصطلحات إحصائية هى بالترتيب:

- صعوبة الفقرة item difficulty
- القدرة على التمييز discrimination power
- نوعية الاستجابة response quality

خطوات تحليل الفقرة:

يتم تحليل الفقرة وفقا للخطوات التالية:

الخطوة ١: ترتيب الإجابات فى بطاقة التسجيل الجماعية من الأعلى إلى الأدنى.

الخطوة ٢: يتم تصنيف الإجابات فى بطاقة التسجيل الجماعية إلى ثلاثة مجموعات فرعية كالتالى:

- المجموعة الأعلى: وتمثل حوالى ٢٧٪ من عدد الإجابات.

- المجموعة الأدنى: وتمثل حوالى ٢٧٪ من عدد الإجابات.

- المجموعة المتوسطة: وتمثل حوالى ٤٦٪ من عدد الإجابات.

وتضم المجموعة الأعلى المفحوصين الذين حققوا أعلى درجات على الاختبار، وتضم المجموعة الأدنى المفحوصين الذين حققوا أقل درجات على الاختبار، وتستخدم المجموعة الأعلى والمجموعة الأدنى كمجموعات متضادة تعكس الحد الأقصى للفروق فى الأداء على الاختبار.

ملحوظة مهمة:

يتوقف حجم كل مجموعة من المجموعتين (الأعلى والأدنى) على العدد الملائم بالنسبة لعملية التحليل الإحصائى، كما يتوقف على العدد الكلى للإجابات. وقد ارتضى علماء القياس أن تتراوح هذه النسبة من ٢٥٪ وحتى ٣٣٪ بالنسبة لكل مجموعة.

الخطوة ٣: حساب عدد الذين أجابوا إجابات صحيحة على كل فقرة من فقرات الاختبار بالنسبة للمجموعة الأعلى upper group.

الخطوة ٤: حساب عدد الذين أجابوا إجابات صحيحة على كل فقرة من فقرات الاختبار بالنسبة للمجموعة الأدنى Lower group.

حساب مؤشر الصعوبة index of difficulty

يتم حساب مؤشر الصعوبة لكل فقرة من فقرات الاختبار باستخدام نتائج الخطوتين (٣، ٤) وتطبيق المعادلة (٤-١) التالية:

$$ب = \frac{س_{١} + س_{٢}}{ن_{١} + ن_{٢}} \times ١٠٠$$ معادلة (٤-١)

حيث إن:

ب = مؤشر الصعوبة.

$ر_1$ = عدد الإجابات الصحيحة فى المجموعة الأعلى.

$ر_2$ = عدد الإجابات الصحيحة فى المجموعة الأدنى.

$ن_1$ = عدد المفحوصين فى المجموعة الأعلى.

$ن_2$ = عدد المفحوصين فى المجموعة الأدنى.

مثال: (٤ - ٥):

طبق اختبار معرفى على ١٨٥ لاعب، والمطلوب حساب مؤشر الصعوبة للفقرات أ، ب، جـ، د. إذا كانت الاستجابات الصحيحة بالنسبة لكل فقرة موزعة بالنسبة للمجموعة الأعلى والمجموعة الأدنى كالتالى:

المجموعات	أ	ب	جـ	د	المجموع
- المجموعة الأعلى ٢٧٪	٢٨	٢	٦	١٤	٥٠
- المجموعة الأدنى ٢٧٪	٢٤	٨	١	١٧	٥٠

والمطلوب: حساب مؤشر صعوبة كل فقرة من الفقرات الأربع.

بتطبيق المعادلة (٤-١) على البيانات المذكورة سلفاً نحصل على مؤشر الصعوبة كالتالى:

$$\text{مؤشر صعوبة الفقرة (أ)} = \frac{٢٨ + ٢٤}{٥٠ + ٥٠} \times ١٠٠$$

$$= \frac{٥٢}{١٠٠} \times ١٠٠$$

$$= ٥٢٪$$

ومؤشر صعوبة الفقرة (ب) = $\frac{٨ + ٢}{٥٠ + ٥٠} \times ١٠٠$

$= \frac{١٠}{١٠٠} \times ١٠٠$

$= ١٠٪$

ومؤشر صعوبة الفقرة (جـ) = $\frac{١ + ٦}{٥٠ + ٥٠} \times ١٠٠$

$= \frac{٧}{١٠٠} \times ١٠٠$

$= ٧٪$

ومؤشر صعوبة الفقرة (د) = $\frac{١٧ + ١٤}{٥٠ + ٥٠} \times ١٠٠$

$= \frac{٣١}{١٠٠} \times ١٠٠$

$= ٣١٪$

وهناك طريقة أخرى لحساب مؤشر صعوبة الفقرة بدون تقسيم المفحوصين إلى مجموعتين (أعلى وأدنى)، حيث يتم تقدير عدد الإجابات الصحيحة على الفقرة بالنسبة لكل المفحوصين، ثم تطبق المعادلة (٤ - ٢) وهى:

$$ب = \frac{ل}{ن} \times ١٠٠ \quad \text{معادلة (٤-٢)}$$

حيث إن:

ب = مؤشر الصعوبة.

ل = عدد من أجابوا إجابات صحيحة على الفقرة.

ن = العدد الكلى للمفحوصين الذين طبق عليهم الاختبار.

مثال (٤-٦):

طبق اختبار معرفى على ٦٠٠ طالب، وقد بلغ عدد الإجابات الصحيحة على إحدى فقرات الاختبار ٣٢٠، فما هو مؤشر صعوبة هذه الفقرة ؟

بالتعويض فى المعادلة (٤-٢) ينتج أن:

$$\text{مؤشر صعوبة الفقرة} = \frac{٣٢٠}{٦٠٠} \times ١٠٠$$

$$= ٥٣٪$$

دلالات مؤشر الصعوبة:

مما سبق يتضح أن مؤشر صعوبة الفقرة يشير إلى النسبة المئوية للأفراد أو المفحوصين الذين أجابوا إجابات صحيحة على الفقرة، هذه النسبة تعرف باسم مؤشر الصعوبة أو معامل الصعوبة، ويمكن تقويم هذه النسبة كالتالى:

١- إذا كانت هذه النسبة (مؤشر الصعوبة) ٩٠٪ دل ذلك على أن الفقرة سهلة جداً، ومن ثم تصبح هذه الفقرة غير قادرة على التمييز.

٢- إذا كانت هذه النسبة ١٠٪ فإن ذلك يدل على أن الفقرة صعبة جداً.

٣- إذا كانت هذه النسبة ١٠٠٪ أو صفراً ٪ فهذا يعنى عدم صلاحية الفقرة لكونها لا تستطيع التمييز بأى حال من الأحوال بين المفحوصين.

٤- يتم تقويم مؤشر صعوبة الفقرة وفقا للمستويات التى حددها ابل Ebel ١٩٧٩ كالتالى:

مؤشر الصعوبة	التقويم
٨٠٪ فأكثر	تستبعد الفقرة
من ٧١٪ الى ٧٩٪	يمكن قبول الفقرة إذا كان لها مؤشر تمييز مقبول
من ٣٠٪ الى ٧٠٪	الفقرة جيدة
من ٢٠٪ الى ٢٩٪	يمكن قبول الفقرة إذا كان لها مؤشر تمييز مقبول
أقل من ٢٠٪	تستبعد الفقرة

٥- تصنف الفقرات التى لها مؤشر صعوبة من ٤٠٪ الى ٦٠٪ كفقرات متوسطة الصعوبة. ويفضل وضع الفقرات التى لها مؤشر صعوبة يزيد عن ٦٠٪ فى أول الاختبار، والفقرات التى لها مؤشر صعوبة أقل من ٤٠٪ فى نهاية الاختبار، حتى نضمن أن يكون لتوزيع الفقرات خصائص التوزيع المعتدل normal distribution بالنسبة لمستوى صعوبة الفقرات.

مما سبق يتضح الآتى:

- يجب ألاَّ تكون الفقرة سهلة جدا بحيث ينجح فيها جميع المفحوصين، وألا تكون صعبة جداً بحيث يرسب فيها كل المفحوصين، لأن الفقرات المتطرفة فى مستويات صعوبتها أو سهولتها، لا تمدنا بالمعلومات الكافية عن مستوى تحصيل المفحوصين، وقد ارتضى علماء القياس أن يتراوح مستوى صعوبة الفقرة من ٣٠ إلى ٧٠٪.
- تحقق الفقرة أقصى قدرة على التمييز بين المفحوصين إذا كان مستوى صعوبتها يسمح بنجاح ٥٠٪ من المفحوصين الذين أجابوا عليها.
- كلما زادت النسبة المئوية للنجاح زادت سهولة الفقرة، وكلما زادت النسبة المئوية للرسوب، زادت صعوبة الفقرة.
- لتقدير مؤشر صعوبة الفقرة يمكن استخدام أى من المعادلتين (٤ : ١أو٢)، حيث يلاحظ أن المعادلة (٤-١) أسهل فى التطبيق لكونها تتعامل مع حوالى ٥٤٪ من عدد المفحوصين، وإن كان يعاب عليها أنها تفترض أن المجموعة الوسطى والتى تمثل ٤٦٪ من عدد المفحوصين تدخل فى إطار التوزيع المعتدل للبيانات وتكون المعادلة (٤-١) أفضل فى التطبيق فى حالات الأعداد الكبيرة، ففى حالة وجود ١١١٢ مفحوص مثلاً، يكون حجم المجموعة الأعلى (٢٧٪) يساوى ٣٠٠ مفحوص، وحجم المجموعة الأدنى (٢٧٪) يساوى ٣٠٠ مفحوص

أيضاً، فى هذه الحالة يتم التعامل مع درجات ٦٠٠ مفحوص فقط بدلاً من ١١١٢ مفحوص، بينما تكون المعادلة (٤-٢) أنسب فى حالات المجموعات صغيرة العدد، فإذا كان عدد المفحوصين ٥٠ مفحوصاً - على سبيل المثال - فإنه لا يفضل تقسيم هذا العدد إلى ثلاث مجموعات، المجموعة الأعلى (٢٧٪) وتضم حوالى ١٤ مفحوصاً، والمجموعة الأدنى (٢٧٪) وتضم ١٤ مفحوصاً أيضاً.

صعوبة الفقرة والاختبار المرجعى المحك (CR):

تتحدد صعوبة الفقرة فى اختبارات التمكن مرجعية المحك (CR) وفقا للطبيعة الخاصة للإنجاز المستهدف من عملية القياس، فإذا كانت المهمات المطلوب انجازها سهلة، كانت فقرات الاختبار هى الأخرى سهلة، وإذا كانت المهمات (الواجبات) متوسطة الصعوبة، كان من الضرورى أن تكون فقرات الاختبار هى الأخرى متوسطة الصعوبة، ولا يجوز بأى حال من الأحوال محاولة إجراء أى تعديل فى مستوى صعوبة الفقرة فى الاختبارات مرجعية المحك، أو حذف الفقرات السهلة من الاختبار لتوسيع مدى درجات الاختبار.

ولا يتأسس مستوى صعوبة الفقرة فى الاختبارات مرجعية المحك على قدرة الفقرة على التمييز بين المستوى المرتفع والمستوى المنخفض للمفحوصين - كما فى حالة الاختبارات معيارية المرجع (NR) - وإنما تتحدد صعوبة فقرة الاختبار فى ضوء طبيعة المهمة أو الواجب المطلوب القيام به، وبناء على ذلك فإنه يمكن تطبيق أى من المعادلتين السابقتين (٤ : ١ أو ٢) للتحقق من مؤشر صعوبة الفقرة فى الاختبارات مرجعية المحك (CR)، ومع ذلك فإن النتائج التى يتم الحصول عليها من تطبيق أى من المعادلتين قد لا تستخدم فى اختيار فقرات الاختبار أو فى معالجة صعوبة الفقرة، فمعظم الفقرات فى اختبارات التمكن مرجعية المحك (CR) يكون لها مؤشرات صعوبة مرتفعة (بمعنى أن النسبة المئوية المطلوبة للنجاح على الاختبار تكون مرتفعة).

صعوبة الفقرة والاختبار المرجعى المعيار (NR):

لأن الاختبارات مرجعية المعيار صممت لإعطاء رتب للمفحوصين وفقا لترتيبهم فى الإنجاز، وهى أيضا اختبارات صممت لإظهار الفروق الفردية بين المفحوصين فى مدى واسع من الدرجات، ولكى يتحقق لها ذلك فإنها تستبعد الوحدات السهلة التى يستطيع المفحوصين الإجابة عليها إجابات صحيحة لكونها لا تظهر الفروق فيما بين المفحوصين. فالفروق بين المفحوصين فى الاختبارات مرجعية المعيار (NR) فى حالة اختبار يتكون من ١٠٠ فقرة - على سبيل المثال - تظهر بأقصى قدر لها إذا كان مدى الدرجات يتراوح من ٥ إلى ٩٥ بمتوسط ٥٠ درجة.

وإذا كانت الصعوبة بالنسبة للاختبارات مرجعية المحك (CR) قد أعدت على أساس الحد الأدنى لمستوى الكفاءة the minimum proficiency Level، لذا فإنها تعد على أساس توقع نجاح ما لا يقل عن ٨٠ إلى ٨٥٪ من المفحوصين عندما تطبق عليهم هذه الاختبارات، وفى المقابل تعد الاختبارات مرجعية المعيار (NR) على أساس أن الفقرة متوسطة الصعوبة.

علاقة مؤشر الصعوبة بثبات الاختبار:

يتفق علماء القياس على أن مستويات صعوبة الفقرات يؤثر على ثبات الاختبار، وأن مستويات الصعوبة الأمثل التى تعطى أعلى ثبات للاختبار تتوقف على مجموعة من العوامل من أهمها عدد فقرات الاختبار، ونوع الفقرات التى يتكون منها الاختبار، حيث يوصى علماء القياس بالآتى:

- يزداد ثبات الاختبار عندما يكون عدد فقراته ٣٠ فقرة أو أقل، ويميل إلى الانخفاض كلما زاد عدد الفقرات عن ٣٠ فقرة.
- فى فقرات الصواب - الخطأ أو (موافق - غير موافق) يجب أن يكون مؤشر صعوبة الفقرة فوق ٧٠٪.

- وأن هذا المؤشر بالنسبة لفقرات الاختيار فيما بين (٢ من البدائل) يجب أن يكون ٨٥٪، وفيما بين (٣ من البدائل) ٧٧٪، وفيما بين (٤ من البدائل) ٧٤٪، وفيما بين (٥ من البدائل) ٦٩٪.

ويمكن حساب ثبات درجة الفقرة (البند) بتطبيق المعادلة (٤-٣) التالية:

$$\text{ر} = \frac{\text{ن}}{\text{ن} - ١} \times \left(\text{ك} - \frac{١}{\text{ن}} \right) \qquad \text{معادلة (٤-٣)}$$

حيث إن:

ر = معامل ثبات الفقرة (البند).

ن = عدد الاختيارات (احتمالات الإجابة).

ك = أعلى تكرار نسبى فى الاختيارات.

مثال: (٤-٧)

أجاب ٢٥٠ مفحوص على إحدى فقرات الاختبار، وكانت الإجابة على هذه الفقرة تتطلب اختيار إجابة من خمس إجابات هى: أ، ب، جـ، د، هـ. وقد جاءت الإجابات (الاستجابات) كالتالى:

الاختيارات	أ	ب	جـ	د	هـ	المجموع
- تكرارات الاستجابة	٢٥	٤٥	٤٠	١٠٠	٤٠	٢٥٠
- التكرار النسبى	,١٠	,١٨	,١٦	,٤٠	,١٦	١,٠٠

والمطلوب : حساب درجة ثبات هذه الفقرة.

ولتحقيق ذلك يتبع الآتى:

١- حساب تكرارات الـ ٢٥٠ مفحوص على كل اختيار من الاختيارات (الاستجابات) الخمسة فنجد أنها تساوى بالترتيب: ٢٥، ٤٥، ٤٠، ١٠٠، ٤٠ (على سبيل المثال).

٢- حساب التكرار النسبى لكل اختيار (استجابة) من الاستجابات الخمس، وذلك بقسمة تكرار كل اختيار على العدد الكلى للمفحوصين، فنجد أن هذه التكرارات النسبية كانت كالتالى:

(١,٠ + ١٨,٠ + ١٦,٠ + ٤,٠ + ١٦,٠ = ١,٠).

٣- تحديد أعلى تكرار نسبى، فنجد أنه يساوى ٤,٠.

٤- نقوم بالتعويض فى المعادلة (٤-٣) كالتالى:

$$= \frac{٥}{١ - ٥} \times \left(٤,٠ - \frac{١}{٥} \right)$$

$$= \frac{٥}{٤} \times (٤,٠ - ٢,٠)$$

$$= ٢٥,٠$$

إذاً معامل ثبات الفقرة يساوى ٢٥,٠

ومن ناحية أخرى يمكن استخدام مؤشر الصعوبة لحساب الانحراف المعيارى (ع - SD) للفقرة باستخدام المعادلة (٤-٤) التالية:

$$ع = \sqrt{ب \times (١٠٠ - ب)}$$ معادلة (٤-٤)

حيث إن:

ع = الانحراف المعيارى للفقرة (البند).

ب = مؤشر صعوبة الفقرة (البند).

١٠٠ = مقدار ثابت.

ففى المثال (٤-٥) كانت مؤشرات الصعوبة للفقرات أ، ب، جـ، د كالتالى بالترتيب: ٥٢٪، ١٠٪، ٧٪، ٣١٪.

وبتطبيق المعادلة (٤-٤)، يمكن حساب ثبات كل من هذه الفقرات كالتالى:

بالنسبة للفقرة (أ)

$$ع = \sqrt{٠,٥٢ \times (٠,٥٢ - ١,٠٠)}$$

$$= \sqrt{٠,٥٢ \times ٠,٤٨}$$

$$= \sqrt{,٢٤٩٦}$$

إذاً ع = ٠,٥٠ تقريباً

بالنسبة للفقرة (ب)

$$ع = \sqrt{٠,١٠ \times (٠,١ - ١,٠)}$$

$$= \sqrt{٠,١ \times ٠,٩}$$

$$= \sqrt{,٠٩}$$

إذاً ع = ٠,٣

كما يمكن حساب تباين الفقرة حيث (التباين = ع٢) وبالتالى يكون تباين الفقرة (أ) = $(٠,٥)^{٢}$ = ٠,٢٥ ، وتباين الفقرة (ب) = $(٠,٣)^{٢}$ = ٠,٠٩ ، ومن الملاحظ أن أعلى قيمة لتباين الفقرة هى ٠,٢٥ ، وهى القيمة التى يساوى فيها مؤشر الثبات حوالى ٥٠٪.

حساب مؤشر التمييز (I D) : Index of discrimination

يشير مؤشر (دليل) التمييز إلى الدرجة التى تميز بها الفقرة (البند) بين مستويات المفحوصين، بمعنى قدرة الفقرة على أن تميز بين الأفراد الذين يحصلون على درجات مرتفعة، والأفراد الذين يحصلون على درجات منخفضة فى الصفة أو السمة التى تقيسها كل فقرات المقياس أو الاختبار. فعندما يحصل عدد كبير من المفحوصين فى المجموعة الأعلى upper group على إجابات صحيحة على الفقرة، نقول: إن القدرة التمييزية للفقرة إيجابية،

بمعنى أن للفقرة مؤشر تمييز إيجابى positive discrimination، أى أن اتجاه التمييز للفقرة هو نفس اتجاه الدرجة الكلية للاختبار.

فإذا كان لفقرة الاختبار القدرة على التمييز بصورة جيدة، فإن ذلك يعنى أن المفحوصين الذين يحققون درجات عالية (مرتفعة) على الاختبار سوف يجيبون على هذه الفقرة إجابة صحيحة، عكس المفحوصين الذين لهم درجات منخفضة على الاختبار، وتعد الفقرة غير القادرة على التمييز فقرة ضعيفة البناء poorly Construction .

بينما يشير التمييز السلبى للفقرة negative discrimination إلى أن أكثر المفحوصين فى المجموعة الأدنى Low group يحصلون على إجابات صحيحة للفقرة.

ويعرف مؤشر (دليل) التمييز (ID) بمؤشر صدق الفقرة item validity index والذى يعد أحد الإجراءات المهمة بالنسبة لتحليل الفقرة. حيث يمكننا من تصنيف كل فقرات الاختبار إلى ثلاثة أنواع من الفقرات هى:

١ - فقرات لها مؤشر تمييز إيجابى.

٢ - فقرات لها مؤشر تمييز سلبى.

٣ - فقرات ليس لها قدرة تمييزية.

ويمكن حساب مؤشر تمييز الفقرة بإحدى الطريقتين التاليتين:

١ - طريقة النسب المئوية Percentages

وتتم هذه الطريقة بتقسيم المفحوصين إلى مجموعتين متساويتين فى العدد وفقا للدرجة الكلية للاختبار (درجات جميع الفقرات)، بحيث تضم المجموعة الأولى (ن١) المتفوقين، وتضم المجموعة الثانية (ن٢) غير المتفوقين.

وتستخدم المجموعتين (ن١ ، ن٢) فى المقارنة كمجموعتين متطرفتين. ويمكن استخدام أى من النسب المئوية التالية لتحديد حجم كل مجموعة.

٣٣٪ كقيم طرفية لكل مجموعة - كيرتون Cureton ١٩٥٧م.

٢٧٪ كقيم طرفية لكل مجموعة - كيلى Kelly ١٩٣٩م.

٢٥٪ كقيم طرفية لكل مجموعة - ايبل Ibel ١٩٧٩م.

٩ ٪ كقيم طرفية لكل مجموعة - ايبل Ibel ١٩٧٩م.

٢ - حساب عدد المفحوصين الذين أجابوا إجابات صحيحة على الفقرة المستهدفة بالنسبة للمجموعة المتميزة (ن١) حيث نرمز لذلك بالرمز ($\checkmark_1$).

٣ - بالمثل يتم حساب عدد من أجابوا إجابات صحيحة على نفس الفقرة بالنسبة للمجموعة التى حصلت على أقل درجات على الاختبار (ن٢) ، حيث نرمز لذلك بالرمز ($\checkmark_2$).

٤ - يتم طرح عدد الاستجابات الصحيحة للمجموعة الضعيفة ($\checkmark_2$) من عدد الاستجابات الصحيحة للمجموعة المتميزة ($\checkmark_1$) ، ثم يقسم الناتج على العدد الكلى للمجموعة المتميزة (ن١) باستخدام معادلة مارشال وهاليس marshall and hales ١٩٧٢م، وصورتها الرياضية كالتالى:

$$د = \frac{\checkmark_1 - \checkmark_2}{ن_1}$$ معادلة (٤ - ٥)

حيث إن:

د = مؤشر التمييز للفقرة (صدق الفقرة).

$\checkmark_1$ = عدد الإجابات الصحيحة فى المجموعة المتميزة.

$\checkmark_2$ = عدد الإجابات الصحيحة فى المجموعة الضعيفة.

ن١ = عدد المفحوصين فى المجموعة المتميزة.

مثـــال :

فى المثال : (٤-٥) كانت بيانات الإجابة على الفقرة (أ) كالتالى:

$$\text{ل}_{١} = ٢٨$$

$$\text{ل}_{٢} = ٢٤$$

$$\text{ن}_{١} = ٥٠$$

إذاً يمكن حساب مؤشر التمييز لهذه الفقرة بتطبيق المعادلة (٤-٥) كالتالى:

$$\text{د} = \frac{٢٨ - ٢٤}{٥٠}$$

$$= \frac{٤}{٥٠}$$

$$= ٠,٠٨$$ إذاً مؤشر التمييز لهذه الفقرة يساوى (١ , ٠) تقريبا.

دلالات مؤشر التمييز (ID):

١- يتراوح مؤشر التمييز من (+١,٠٠ إلى - ١,٠٠) ولكنه نادراً ما يصل إلى أى من القيمتين.

٢- عندما يكون مؤشر التمييز سلبيا، فإن ذلك يعنى أن عدد المفحوصين الذين أجابوا إجابات صحيحة على الفقرة فى المجموعة الأدنى ($ن_{٢}$) أكثر ممن أجابوا على نفس الفقرة فى المجموعة الأعلى ($ن_{١}$)، ومن ثم يحذف مثل هذا النوع من الفقرات تماماً.

٣- بالنسبة للاختبارات مرجعية المعيار (NR) يقبل مؤشر تمييز الفقرة عندما يساوى (٠,٤) فأكثر. وقد حدد ايبل Ibel ١٩٧٩ مؤشرات التمييز للفقرات فى الاختبارات مرجعية المعيار كالتالى:

مؤشر التمييز	التقويم
٠,٤٠ فما فوق	لها قدرة على التمييز جيدة

من ٠,٣٠ إلى ٠,٣٩	لها قدرة على التمييز مناسبة، وربما تحتاج إلى تعديل وبخاصة فيما يتعلق بصعوبة الفقرة
من ٠,٢٠ إلى ٠,٢٩	تحتاج إلى مراجعة
أقل من ٠,٢٠	تستبعد الفقرة

ملحوظة مهمة:

لاتطبق المستويات السابقة الخاصة بمؤشرات التمييز على الاختبارات مرجعية المحك (CR).

٢- الطرق الارتباطية Correlational techniques

قدم ليندكويست Lindquist ١٩٥١م الدرجة الكلية للاختبار كمحك داخلى لحساب صدق الفقرات بالرجوع إلى هذا المحك الداخلى، حيث أطلق ليندكويست على هذه الطريقة اسم (مؤشر تمييز الفقرة عن طريق الاتساق الداخلى Internal Consistency). وتعتمد هذه الطريقة على تحليل العلاقة (الارتباط) بين درجة المفحوص على الفقرة ودرجته الكلية على الاختبار ككل والتى تستخدم كمحك (أو ميزان داخلى) لتقويم صدق كل فقرة من فقرات الاختبار.

فالعلاقة بين الفقرة والدرجة الكلية للاختبار تظهر كيف تقيس الفقرة بشكل جيد الوظائف التى يقيسها الاختبار نفسه، حيث يبين مؤشر الصدق المدى الذى تمتد إليه فقرة محددة فى التمييز فيما بين المفحوصين الذين يظهرون فروقاً حادة فى الوظائف التى يقيسها الاختبار ككل.

ويرى بعض خبراء القياس أنه يمكن النظر إلى ارتباط الدرجة الكلية للفقرة على أنه مؤشر جيد للتمييز، حتى أنه يمكن اختيار الفقرات على أساس الارتباطات الكلية للفقرة، إذا كانت نسبة الارتباطات الموجبة لا تقل عن ٧٥٪ وتزيد قيمها عن ٠,١٥.

ويستخدم لقياس الارتباط بالنسبة لهذا الموضوع خمسة مقاييس شائعة هى: الارتباط التتابعى (PPM) والارتباط الثنائى biserial (r) والارتباط الثنائى الحقيقى Point biserial (r) ، والارتباط الرباعى tetrachoric ومعامل فاى Phi Coefficient - . ويعتبر فلانجان Flanagan أول من استخدم معاملات الارتباط الثنائى لتقدير نسبة المفحوصين الناجحين على الفقرة بالنسبة لـ ٢٧٪ العليا كمجموعة مرجعية Criterion group ونسبة المفحوصين الناجحين على نفس الفقرة فى الـ ٢٧٪ الدنيا، وذلك على أساس ألا يقل حجم كل مجموعة عن ١٠٠ مفحوص وأن يكون العدد الكلى للمفحوصين ٣٧٠ مفحوصا.

التخمين وعوامل الصدفة:

من العوامل التى تؤثر سلبيا على صعوبة الفقرة - عند استخدام النسبة المئوية للإجابات الصحيحة - هو لجوء بعض المفحوصين إلى التخمين guessing فيما بين الاختيارات المعروضة عليهم، بمعنى أن يختار المفحوص جواباً من بين عدة أجوبة دون دراية بسبب الاختيار ثم يحدث أن تكون هذه الإجابة صحيحة عن طريق الصدفة by chance. فالتخمين يعنى أن نسبة معينة من المفحوصين الذين لا يعرفون الإجابة الصحيحة أجابوا بمحض الصدفة إجابات صحيحة على الفقرة.

تصحيح أثر التخمين:

للتغلب على هذه المشكلة يتم ما يعرف بتصحيح التخمين أو أثر التخمين (تصحيح الصدفة) Correction for guessing والتى تستهدف استبعاد النسبة المئوية للتخمين من النسبة المئوية للناجحين فى الفقرة للوصول إلى النسبة المئوية الحقيقية للإجابات الصحيحة للفقرة. ويتم تصحيح أثر التخمين كالتالى:

أولاً : بالنسبة لاختبارات الصواب - الخطأ

تستخدم المعادلة التالية:

س = ل - خ معادلة (٤-٦)

حيث إن :

س = الدرجة بعد التصحيح.

ل = عدد الإجابات الصحيحة.

خ = عدد الإجابات الخاطئة.

وتتأسس هذه المعادلة على مسلمة مؤداها أن الإجابات الخاطئة جاءت نتيجة التخمين.

مثال (٤-٨):

طبق اختبار الصواب - الخطأ يتكون من ٧٥ فقرة على أحد التلاميذ، وقد تمكن التلميذ من الإجابة على ٦٠ فقرة إجابات صحيحة، والإجابة على ١٥ فقرة إجابات خاطئة. والمطلوب حساب درجة هذا التلميذ بعد تصحيح أثر التخمين.

لتحقيق المطلوب يتبع الآتى:

- حساب الدرجة بعد التصحيح بتطبيق المعادلة (٤-٦) كالتالى:

س = ٦٠ - ١٥

= ٤٥ درجة

أى أن درجة التلميذ أصبحت ٤٥ بدلا من ٦٠

ثانيا : بالنسبة لاختبارات الاختيار من متعدد (MCq)

تستخدم المعادلة التالية :

$$س = ل - \frac{خ}{ن - ١}$$ معادلة (٤-٧)

حيث إن :

س = الدرجة بعد التصحيح .

ل = عدد الإجابات الصحيحة .

خ = عدد الإجابات الخاطئة .

ن = عدد البدائل (الاحتمالات أو الاختيارات) لكل فقرة .

ملحوظة مهمة :

المعادلة (٤-٦) هى نفسها المعادلة (٤-٧) وذلك فى حالة ما إذا كانت (ن=٢) كالتالى :

$$س = ل - \frac{خ}{ن - ١} \quad \text{معادلة (٤-٧)}$$

$$= ل - \frac{خ}{١}$$

$$س = ل - خ \quad \text{معادلة (٤-٦)}$$

وإذا كان التلميذ قد أجاب إجابات صحيحة على ٦٠ فقرة (المثال : ٤-٨) فى اختبار الاختيار من متعدد (mcq) احتمالات الإجابة فيه (٣, ٤, ٥) فى هذه الحالة يمكن حساب درجة التلميذ بعد التصحيح بتطبيق المعادلة (٤-٧) كالتالى :

فى حالة (٣ اختيارات) $= ٦٠ - \frac{١٥}{٣ - ١}$

$$= ٦٠ - \frac{١٥}{٢}$$

$$= ٦٠ - ٧,٥$$

$$= ٥٢,٥$$

٥٢,٥

فى حالة (٤ اختيارات) $= ٦٠ - \frac{١٥}{٤ - ١}$

$$= ٦٠ - \frac{١٥}{٣}$$

$$= ٥٥$$ ٥٥

فى حالة (٥ اختيارات) $= ٦٠ - \frac{١٥}{١ - ٥}$

$$= ٦٠ - \frac{١٥}{٤}$$

$$= ٦٠ - ٣,٧٥$$

$$= ٥٦,٢٥$$ ٥٦,٢٥

مما سبق يلاحظ الآتى:

- فى حالة الفقرات التى تعتمد على إجابة واحدة من إجابتين، فإن إجابات المفحوص تتأثر بعوامل التخمين أكثر منه بالنسبة لأى نمط آخر من الاختيارات، فقد بلغ عدد الإجابات الصحيحة ٤٥ إجابة بعد التصحيح فى مقابل ٦٠ إجابة قبل التصحيح.
- أنه فى حالة فقرات الاختيار من متعدد يقل تأثير التخمين فى مقابل الفقرات ثنائية الاختيار (صح - خطأ). وأنه كلما زاد عدد الاختيارات (الاحتمالات) قل تأثير التخمين. ففى حالة (٥ اختيارات) كان تأثير التخمين = ٣,٧٥ درجة وفى حالة (٤ اختيارات) كان تأثير التخمين = ٥ درجات، وفى حالة (٣ اختيارات) كان تأثيره = ٧,٥ درجة، وفى حالة (اختيارين فقط) كان تأثيره = ١٥ درجة.

مثال (٤ - ٩):

طبق اختبار معرفى الإجابة عليه بطريقة الاختيار من متعدد (mcq) لعدد أربعة اختيارات هى (أ، ب، جـ ، د) وبعد ترتيب استجابات المفحوصين وفقا للدرجة الكلية للاختبار من الأعلى إلى الأدنى، أخذت المجموعة الأعلى

(ن١) والمجموعة الأدنى (ن٢) كمجموعتين طرفيتين، وكانت استجابات كل مجموعة وفقا للبدائل (الاختيارات) المطروحة كالتالى:

المجموعات/ الفقرات	أ	ب	جـ	د	المجموع
المجموعة الأعلى(ن١)	١	٤	١٧	٥	٢٧
المجموعة الأدنى (ن٢)	٣	٧	٨	٦	٢٧

والمطلوب: تقويم هذه المجموعة من الفقرات وفقا لنتائج الاستجابة عليها.

لتحقيق ذلك نقوم بحساب مؤشر صعوبة كل فقرة بتطبيق المعادلة (٤-١) كالتالى:

بالنسبة للفقرة (أ) = $\frac{١+١}{٢٧+٢٧} \times ١٠٠ = \frac{٢}{٥٤} \times ١٠٠ = ٤٪$

وبالنسبة للفقرة (ب) = $\frac{٧+٤}{٥٤} \times ١٠٠ = \frac{١١}{٥٤} \times ١٠٠ = ٢٠٪$

وبالنسبة للفقرة (جـ) = $\frac{٨+١٧}{٥٤} \times ١٠٠ = \frac{٢٥}{٥٤} \times ١٠٠ = ٤٦٪$

وبالنسبة للفقرة (د) = $\frac{٦+٥}{٥٤} \times ١٠٠ = \frac{١١}{٥٤} \times ١٠٠ = ٢٠٪$

ووفقا للمعايير التى حددها إيبل ١٩٧٩م لمؤشر الصعوبة، فإنه يمكن استبعاد الفقرات (أ،ب،د) والابقاء فقط على الفقرة (جـ) لأن لها معامل صعوبة ينحصر فيما بين (٣٠ إلى ٧٠٪).

وللابقاء على هذه الفقرة أو استبعادها، يتم تصحيح مؤشر الصعوبة من أثر التخمين وذلك باستخدام المعادلة (٤-٨) كالتالى:

مؤشر الصعوبة بعد التصحيح = ب - $\frac{(١٠٠ - ب)}{ن - ١}$ معادلة (٤-٨)

حيث إن:

ب = مؤشر الصعوبة قبل التصحيح.

ن = عدد الاختيارات (احتمالات الإجابة).

١٠٠ = مقدار ثابت.

وبالتعويض فى المعادلة (٤-٨) ينتج أن:

مؤشر صعوبة الفقرة (جـ) بعد التصحيح

$$= ٤٦ - \frac{١٠٠ - ٤٦}{١ - ٤}$$

$$= ٤٦ - \frac{٥٤}{٣}$$

$$= ٤٦ - ١٨$$

$$= ٢٨$$

إذاً مؤشر صعوبة الفقرة (جـ) يساوى ٢٨ ٪.

ووفق المستويات التى حددها ايبل ١٩٧٩م لمؤشرات الصعوبة، فإنه يلاحظ أن حذف هذه الفقرة أو الابقاء عليها يتحدد على أساس قدرة الفقرة على التمييز وذلك بتطبيق المعادلة (٤-٥) كالتالى:

$$\text{مؤشر التمييز للفقرة (جـ)} = \frac{١٧ - ٨}{٢٧}$$

$$= \frac{٩}{٢٧}$$

= ٠,٣٣ (وهو معامل تمييز يجيز قبول الفقرة)

متى نلجأ إلى تصحيح أثر التخمين؟

نلجأ إلى تصحيح أثر التخمين فى الحالات التالية:

١ - فى الاختبارات المعرفية التى تتطلب السرعة فى الأداء، والتى تعرف باسم: اختبارات السرعة speed tests ، لكونها اختبارات تعتبر السرعة فى الأداء عاملا مهما، لذلك يتوقع أن يحدث التخمين عند الاستجابة على مثل هذه الاختبارات.

٢ - فى الاختبارات المعرفية التى تتميز بدرجة عالية من الصعوبة وتكون محدودة بزمن.

٣ - قد لا تكون هناك حاجة لتصحيح أثر التخمين عندما ينتهى جميع المفحوصين من الاختبار قبل انتهاء الزمن المخصص له، لأن هذا يعنى أن أثر التخمين فى الإجابات كان ضعيفا للغاية لعدم وجود أية ضغوط مرتبطة بالزمن قد تؤدى إلى التخمين.

٤ - يلزم تصحيح أثر التخمين فى الاختبارات التى تصمم بحيث لا يستطيع المفحوص الإجابة على جميع وحداتها خلال الزمن المقرر للاختبار.

نوعية الاستجابة Response Quality

يطلق على اختيار الإجابات على الفقرة فى حالة اختبارات الاختيار من متعدد (mcq) اسم: الاستجابات responses، حيث تشير الاستجابات غير الصحيحة - فى هذا النمط من الاختبارات - إلى ما يعرف بالمضللات (المشتتات) distractors والتى تتمثل فى إجابات خاطئة تغرى المفحوص باختيارها فى اختبارات الاختيار من متعدد أو اختبارات المزاوجة (المواءمة).

ففى اختبارات الاختيار من متعدد (mcq) يمكن أن تختار كل استجابة عن طريق بعض المفحوصين، فإذا لم تختار الاستجابة بواسطة أى مفحوص، فإن ذلك يعنى أن هذه الاستجابة ليس لها إسهامات تذكر بالنسبة للاختبار، ومن ثم فإنه يلزم حذف هذه الاستجابة من استجابات الاختيار من متعدد أو المزاوجه.

وقد اتفق علماء القياس على أن يتم الإبقاء على الاستجابات إذا اختارها ما لا يقل عن (٢ إلى ٣٪) من عدد المفحوصين الذين أعطى لهم الاختبار.

ويتفق علماء القياس على أن المضللات (المشتتات) تكون فعالة إذا كان عدد المفحوصين الذين اختاروها فى المجموعة الأدنى أكثر من عدد المفحوصين الذين اختاروها فى المجموعة الأعلى، لأن ذلك يعنى أن لهذه المضللات قدرة على التمييز بين المجموعة الأدنى والمجموعة الأعلى، وأن التمييز فى الاتجاه الإيجابى لكونه فى صالح المجموعة الأعلى صاحبة الاستجابات الأفضل على الاختبار.

مثال : (٤-١٠):

البيانات التالية تبين استجابات المجموعة الأعلى (ن١ = ٢٧) ، والمجموعة الأدنى (ن٢ = ٢٧) على فقرة فى اختبار للاختيار من متعدد (mcq) عدد الاختيارات فيه أربع احتمالات هى: أ ، ب، (جـ) ، [د] ، حيث يلاحظ الاستجابة (جـ) هى الاستجابة الصحيحة، وأن الاستجابة د هى الاستجابة المضللة.

المجموعة / الاختيارات	أ	ب	(جـ)	[د]
المجموعة الأعلى (ن١=٢٧)	٤	٠	١٢	١١
المجموعة الأدنى (ن٢=٢٧)	٩	٠	٨	١٠

ويمكن تقويم الاستجابات على هذه الفقرة كالتالى:

- تستبعد الاستجابة (ب) لعدم فعاليتها، كاستجابة بديلة لكونها لم يتم اختيارها من قِبل أى مفحوص فى المجموعتين.
- الاستجابة (جـ): هى الاستجابة الصحيحة لقدرتها على التمييز الإيجابى بين المجموعتين الأعلى والأدنى، لأن عدد المفحوصين الذين استخدموها فى المجموعة الأعلى (١٢ مفحوصا) وهو أكثر من العدد الذى اختارها فى المجموعة الأدنى (٨ مفحوصين).
- الاستجابة (د) لا تصلح كاستجابة مضللة لأن عدد الذين اختاروها من المجموعة الأعلى (١١ مفحوص) وهو أكبر من عدد الذين اختاروها فى المجموعة الأدنى (١٠ مفحوصين)
- الاستجابة (أ) تظهر كاستجابة مضللة، لأن المفحوصين الذين اختاروها فى المجموعة الأدنى (٩ مفحوصين) وهو أكبر من عدد المفحوصين الذين اختاروها فى المجموعة الأعلى (٤ مفحوصين).

مما سبق يتضح أن الفقرة فى اختبارات الاختيار من متعدد أو اختبارات المزاوجة يكون لها مجموعة من الاستجابات، يطلق على الاستجابة الصحيحة اسم: الاختيار الصحيح، وتعرف الإستجابات غير الصحيحة باسم: المضللات أو المشتتات، هذه المضللات (المشتتات) عبارة عن اختيارات (بدائل) غير صحيحة incorrect alternatives تستمد اسمها من وظيفتها لكى تضلل المفحوصين المتشككين فى الإجابة الصحيحة للفقرة.

خامساً: مراجعة وتنقيح الفقرة:

بعد الانتهاء من تحليل الفقرة - وهى الخطوة الرابعة من خطوات بناء الاختبارات المعرفية - يبدأ الباحث بعمل المراجعات والتعديلات الضرورية لتنقيح الاختبار. حيث تشتمل هذه المراجعات على الآتى:

- تغيير بعض الاستجابات.
- حذف بعض الفقرات (البنود) وإعادة صياغة البعض الآخر منها.
- نقل الفقرات من اختيار لآخر (كأن تنقل الفقرة من صيغة الاختيار من متعدد إلى صيغة الصواب - الخطأ).
- استبعاد العبارات (البنود) التى لا تحقق مؤشر التمييز المقبول وفقا للمعايير الخاصة بذلك.

معجم إنجليزى - عربى للمصطلحات

A

Ambiguous	غامض
Alternatives	بدائل

B

Biases	تحيز

C

Correction for guessing	تصحيح أثر التخمين
Completion	التكميل
Completion items	فقرات التكميل
Complex concepts	مفاهيم معقدة
Criterion group	مجموعة مرجعية

D

distrator	مشتت - مضلل
difficulty level	مستوى الصعوبة
discriminative power	القوة التمييزية للفقرة

E

Essay	كتابة المقال
Essay test	اختبار مقال
Essay items	فقرات المقال
Essay test items	فقرات اختبار المقال

F

Facts	حقائق
Factual Knowledge	المعرفة الحقيقية

G

Guessing factor	عامل التخمين

H

Higher mental processes	العمليات العقلية العليا

I

Interpreting	توضيح - تفسير
Integrating	التوحيد والدمج
Item analysis	تحليل الفقرة
Item revision	مراجعة وتنقيح الفقرة
Item writing	كتابة الفقرة
Item validity	صدق الفقرة
Irrelevant	غير ذات صلة
Item difficulty	صعوبة الفقرة
Index of the difficulty	مؤشر الصعوبة

M

Misconceptions	مدركات خاطئة
Mcq	الاختيار من متعدد
Memorization	الحفظ عن ظهر قلب

Matching test	اختبار المزاوجة
Matching items	فقرات المزاوجة

N

NR	اختبارات مرجعية المعيار
Negative discrimination	

O

objective test	اختبار موضوعى
organizing	تنظيم
objective test item	فقرات الاختبار الموضوعى
opinion	رأى

P

paper and pencil tests	اختبارات الورقة والقلم
Principles	مبادئ
Positive words	تعبيرات موجبة
plausible	قابل ظاهريا للتصديق
positive discrimination	التمييز الإيجابى للفقرة
Poorly constructions	فقرة ضعيفة البناء

Q

Quality of items	نوعية الفقرات

R

Relative weights	الأوزان النسبية
Random guessing	التخمين العشوائى
Responses	استجابات
Rote learning	التعلم الصمى (الحفظ)
Response quality	نوعية الاستجابة

S

Subjective evaluation	تقويم ذاتى
Synthesizing	تركيب - تاليف
Short essay items	فقرات المقال القصيرة
Stem	الأرومة (البند الرئيسى)

T

thinking ability	القدرة على التفكير
true - false items	فقرات الصوات - الخطأ

V

Vague	غير واضح

W

Written expression	التعبير كتابة
word	كلمة

١- طبيعة القياس الانفعالى

أشرنا (فى الفصل الأول) إلى أن المجال الانفعالى يتضمن خمس مستويات كبيرة مرتبة فى شكل تنظيم هرمى من البسيط إلى المركب كالتالى: الاستقبال، والسلوك الاستجابى، والسلوك القيمى، والسلوك التنظيمى، والوصف. وذكرنا أن السلوك الانفعالى affective behavior مصطلح يطلق على المجال الانفعالى affictive domain الذى يشتمل على الميول والاتجاهات والقيم والتذوق وغيرها من الموضوعات الانفعالية التى تنعكس خلال أحاسيس الفرد وانفعالاته.

ويشير علم النفس المعاصر إلى أن السلوك الانفعالى يعتمد على التفاعل بين عدد من العوامل، بعض هذه العوامل يعتمد على التكيف الناشئ عن التعلم، والبعض الآخر يعتمد على الوظائف الكيميائية وبنية الجسم ومستوى الذكاء العام للفرد. فالطفل صغير الحجم قد يصبح عدوانيا للتعويض عن حجم جسمه، وربما ينسحب من دروس وأنشطة التربية الرياضية المدرسية ويعوض هذا الانسحاب بالتوجه نحو أنشطة تتطلب المهارات العقلية أو الفنية.

أهمية القياس الانفعالى:

على الرغم من قلة الاهتمام بقياس السلوك الانفعالى فى مجال التربية الرياضية المدرسية، إلا أنه فى حالة تطبيق بعض أدوات القياس المناسبة، فإنه يمكن توظيف النتائج التى يتم الحصول عليها كالتالى:

- تحديد مكانة التلميذ أو الطالب بالنسبة لجماعة الفصل أو الفريق.
- معرفة اتجاهات وميول التلميذ أو الطالب نحو العديد من الأنشطة الرياضية بحيث يمكن الاستفادة من ذلك فى إعداد المنهاج والبرنامج.
- تمكين المعلم من كتابة التقارير المدرسية عن التلميذ أو الطالب وبخاصة فى حالات الطالب المشاغب والطالب المقصر دراسيا.

- الحث الدافعى عن طريق التخطيط الجيد للأغراض الانفعالية من جانب كل من المعلم والطالب، بحيث يمكن الاستفادة من ذلك فى حث الطلاب نحو العمل على إنجاز تلك الأهداف.

- مساعدة التلاميذ والطلاب على معرفة أنفسهم، فالقياس الانفعالى يجعل التلاميذ والطلاب على وعى باتجاهاتهم واهتماماتهم وقيمهم وقبول الجماعة لهم group acceptance وتقدير الذات self-esteem، فعند الوعى بتلك النتائج يتجه الطالب للتعاون مع المعلم للعمل من أجل تحقيق التقدم الذاتى Self - improvement .

- تستخدم نتائج قياس السلوك الانفعالى فى انتقاء الطلاب لبعض الأدوار المهمة بالنسبة للفرق الرياضية حيث يستخدم لذلك النتائج التى يتم الحصول عليها من قياسات الاتجاهات والقيادة والتنافسية Competitiveness وبعض سمات الشخصية وغيرها.

فلسفة وأهداف القياس الانفعالى:

يهتم المجال الانفعالى بالقيم والميول والاتجاهات وبالسلوك الاجتماعى للشخص. ويذكر ستونر stoner ١٩٨٢م أن للقياس الانفعالى فى المجال الرياضى أهمية خاصة لكون النمو النفسى والاجتماعى للفرد يعد واحداً من الأهداف العامة للتربية الرياضية، حيث تشكل القيم والاتجاهات والتذوق وتحقيق الذات مكونات الشعور feelings والانفعالات (العواطف) emotions وهى مكونات تمثل متحصلات (نواتج) مهمة بالنسبة للخبرات الإنسانية فى المجال الرياضى.

ويهتم العاملون فى المجال الرياضى بالنمو الانفعالى للاعبين والممارسين، ومن ثم فهم يضعون لذلك أهدافا واقعية يمكن تحقيقها. وفى هذا الصدد يشير «ستونجر» ١٩٨٢م إلى أن الأهداف الانفعالية يجب أن توضع بالنسبة للنشاط الرياضى والتربية البدنية بطريقة مناسبة، فليس من المعقول أن

مراجع الفصل الرابع

١ - أمين أنور الخولى، وجمال الدين الشافعى (٢٠٠٠م) : مناهج التربية البدنية المعاصرة. دار الفكر العربى، القاهرة.

٢ - أمين أنور الخولى، ومحمود عبدالفتاح عنان (١٩٩٩م): المعرفة الرياضية (الإطار المفاهيمى - اختبارات المعرفة الرياضية - أسس بنائها ونماذج كاملة منها) دار الفكر العربى، القاهرة.

٣ - أنور محمد الشرقاوى (٢٠٠٣م) : علم النفس المعرفى المعاصر. مكتبة الأنجلو المصرية، القاهرة.

٤ - ليلى السيد فرحات (٢٠٠١م) : القياس المعرفى الرياضى. مركز الكتاب للنشر، القاهرة.

٥ - نورمان جرونلند (١٩٧٦م): الأهداف التعليمية - تحديدها السلوكى وتطبيقاته. ترجمة الدكتور أحمد خيرى كاظم، دار النهضة العربية، القاهرة.

6- Bloom, B.S. (Ed.) (1956): Taxonomy of educational objectives: Cognitive domain. Mckay : New Nork.

7 - Gronlund, N.E. (1981): Measurement and evaluation in teaching. Macmillan publishing Co., Inc. New york.

8 - Johnson, B.L., and Nelson, J.K (1986): Practical measurements for evaluation in physical education. Minneapolis: Burgess publishing company.

9 - Miller,D.K. (1994): Measurement by the physical educator : why and How, WCB Brown and Benchmark Publishers.

10 - Morrow, J.R., Jackson, A.W. Disch, J.G., and Mood, D.P. (1995): Measurement and evaluation in human performance. Human Kinetics.

الفصل الخامس

القياس الانفعالى

١- طبيعة القياس الانفعالى

٢ - أساليب القياس الانفعالى

٣ - أدوات القياس الانفعالى

٤ - الطرق الوصفية للقياس الانفعالى

أولا : الملاحظة

ثانيا : المقابلة الشخصية

٥ - الطرق الكمية للقياس الانفعالى

أولا : مقاييس ليكرت

ثانيا : مقاييس تمايز المعانى

ثالثا : مقاييس (التدريج) التقدير

رابعا : الاستبيانات (الاستخبارات)

- معجم إنجليزى - عربى للمصطلحات
- مراجع الفصل الخامس

تخصص وحدة دراسية لكل هدف من الأهداف الانفعالية، وإنما يمكن تحقيق الأهداف الانفعالية فى المجال الرياضى كالتالى:

- أن تعلم الالتزام بالقواعد المنظمة للعديد من الألعاب والمسابقات الرياضية يساعد على تطوير وتنمية المعرفة والاتجاهات المرتبطة بتقبل القرارات الملزمة.
- أن الأنشطة الجماعية مثل ألعاب الفرق والألعاب الجماعية تعمل على تطوير وتنمية التعاون، والقيادة، والفرد كعضو فى جماعة followership.
- أن الأنشطة الحركية تتطلب من الفرد أن يكون عارفا بجسمه وبقدراته وبحدود تلك القدرات، مما يزود هذا الفرد بأسس لتطوير الذات self concepts المرتبطة بصورة الجسم (الصورة الذهنية للجسم) Body image.
- أن التربية الرياضية المدرسية تعمل على تنمية اتجاهات إيجابية نحو النشاط البدنى كطريقة وأسلوب للحياة.
- أن الأنشطة الرياضية والليلقة يمكن أن تعلم الطلاب قيمة المثابرة perseverance تحت ظروف المحاولة.

وهناك وجهة نظر ترى أن القيام بقياس وتقويم النمو الانفعالى بالنسبة للتربية الرياضية المدرسية مهمة لا لزوم للقيام بها، فمثلا يوصى مود mood ١٩٨٢م باستبعاد أهداف النمو الانفعالى من الأهداف الأولية للتربية الرياضية المدرسية، وذلك للأسباب التالية :

١ - أن معلمى التربية الرياضية بالمدارس لديهم ما يكفيهم من الواجبات والمهام الأساسية المتعلقة بالتدريس والتعليم وتدريب الفرق الرياضية والإشراف على النشاطين الداخلى والخارجى مما لا يسمح لهم بالقيام بقياس وتقويم السلوك الانفعالى للطلاب.

٢ - يرى معلمو التربية الرياضية أن التنمية الانفعالية للطلاب ليست مهتمهم وحدهم فقط، ولكنها أيضا مهمة جميع المعلمين بالمدرسة.

٣ - يرى معلمو التربية الرياضية أن أدوات القياس المتعلقة بالمجال الانفعالى غير متاحة بالقدر الكافى، وأن ما هو متاح منها مقاييس تفتقر إلى الثبات.

٢- أساليب القياس الانفعالى

يتم تجميع البيانات الخاصة بالمجال الانفعالى بإحدى ثلاثة أساليب هى:

- التقرير (التقدير) الذاتى الذى يقوم به المفحوص نفسه.
- تقدير سلوك المفحوص عن طريق الملاحظة التى يقوم بها الآخرون.
- التقديرات التى يضعها أفراد المجموعة لفرد واحد من المجموعة.

وفيما يلى شرح موجز لكل أسلوب من هذه الأساليب.

أولا : أسلوب التقرير الذاتى Self - report

وهو يتضمن كل الإجراءات التى تستهدف سؤال الشخص (المفحوص) لكى يقدم أو تقدم تقريراً report عن الذات، هذا التقرير يمكن أن يزودنا بالمعلومات التى يقررها الشخص عن ذاته بإحدى الطريقتين التاليتين:

١ - طريقة التعبير بالكلام Oral method

وهى طريقة تتطلب من المفحوص أن يعبر عن استجاباته بالكلام (شفهيا) حيث يقوم غيره من الأفراد المسئولين بتسجيل نتائج هذه الاستجابات، وذلك كما فى حالات المقابلات الشخصية والاستخبارات (الاستبيانات).

٢ - طريقة التعبير بالكتابة Written method

وهى طريقة يقوم فيها المفحوص بالتعبير كتابة وبطريقة مباشرة عن استجاباته من خلال بعض وسائل القياس مثل الاستبيانات (الاستخبارات)، وتوائم الميول والاتجاهات والشخصية وغيرها.

ويتأسس استخدام أسلوب التقرير الذاتى على فرضية أن المفحوصين المستهدفين من القياس لديهم الوعى بالذات Self - awareness لكى يقروا بمعتقداتهم وأحاسيسهم، وأن لديهم القدرة على التعبير عن ذلك بوضوح.

ثانيا : أسلوب التقارير التى يقوم بها آخرون Repports of others

وهى تقارير يقدمها آخرون عن المفحوص، هذه التقارير تتضمن معلومات عن أحاسيس وانفعالات ومعتقدات المفحوص، حيث يتم الحصول عليها عن طريق المقابلات الشخصية والاستبيانات والملاحظات وغيرها. وقد يقوم بهذه التقارير المعلم، أو المدرب، أو أحد الأبوين، أو المشرف، أو الباحث أو أى شخص آخر يوثق فى قدرته على القيام بهذه المهمة.

وقد يقوم بكتابة التقارير أشخاص على علاقة بالمفحوص، أو أشخاصاً ليست لديهم أى علاقة بالمفحوص، وفى جميع الحالات تظل كفاءة وتدريب القائم بإعداد التقرير من أهم متطلبات هذه الطريقة. ومن ناحية أخرى يمكن النظر إلى هذه الطريقة على أنها طريقة مرغوبة لتقويم سلوك الآخرين، لأن ما يكتبه الباحث عما يشاهده بالنسبة لسلوك المفحوص يكون - فى كثير من الحالات - أكثر مصداقية مما يكتبه المفحوص عن نفسه.

وتستخدم التقارير التى يكتبها الآخرون لاستحضار (استرجاع) المعلومات التى يتم تسجيلها عن طريق الملاحظة المنظمة وغير المنظمة. وتأخذ هذه التقارير كل من الصور الكلامية (الشفهية) والصورة الكتابية، حيث تتمثل الصور الكلامية فيما يعبر عنه المفحوص من إجابات عن أسئلة توجه إليه، وتتمثل الصورة الكتابية فى التقارير التى يكتبها الباحث عن استجابات المفحوص عن الأسئلة التى وجهت إليه. وتتضمن هذه الطريقة المقابلات الشخصية، الملاحظات، والاستبيانات، وتقارير الأحداث السلوكية المهمة فى حياة المفحوص anecdotal reports، وقوائم المراجعة checklists.

ثالثا : أسلوب القياس الاجتماعى Sociometric

وهو أسلوب يتطلب من أفراد جماعة ما أن يقدموا تقريراً عن اتجاهاتهم - كل فرد منهم نحو الآخر - انطلاقاً من أن كل فرد من أفراد المجموعة أمامه فرصة كافية لكى يلاحظ عينة ممثلة من سلوك زملائه الآخرين فى مجموعته. ومن أساليب القياس الاجتماعى أسلوب التقدير بواسطة الرفاق peer rating، ومخططات العلاقات الاجتماعية sociograms، ومقاييس المسافة الاجتماعية social - distance scales ، وطريقة خمن - من Guess who .

طرق القياس الانفعالى كما يحددها أمان وجلوك:

يحدد أمان وجلوك Ahmann and Glock ١٩٨١م أربعة طرق (أساليب) لتقويم السلوك الانفعالى هى:

أولاً : طرق تتأسس على الملاحظة Methods based on observation.

وتشمل:

- سجلات السيرة الذاتية Anecdotal records
- مقاييس التقدير rating scales
- قوائم المراجعة Checklists
- المقابلات الشخصية Interviews

ثانياً : أساليب القياس الاجتماعى Sociometric techniques

وتشمل:

- طريقة خمن - من Guess who
- مخططات العلاقات الاجماعية Sociograms
- مقاييس المسافة الاجتماعية Social-distance scales



السلوك القيادى للمدرب الرياضى

تفاعل أفراد الفريق فى الألعاب الجماعية

ثالثاً : قوائم التقرير الذاتى Self - report inventories

وتشمل :

- قوائم الشخصية Personal inventories
- قوائم الميول Interests Inventories
- مقاييس الاتجاهات والاستبيانات

Attitude scales and Questionnaires

رابعاً : الأساليب الإسقاطية Progective techniques

وتشمل:

- طرق عرض الصورة وبقع الحبر

Inkblot and picture presentations

- طرق إكمال الجمل Sentence completions
- طرق عرض المثير اللفظى Presentations of verbal stimuli

ويحدد مورو وآخرون Morrow et.al. ١٩٩٥م أسلوبين رئيسيين للقياس الانفعالى هما:

أولاً : الأسلوب الخاص بالطرق الوصفية Qualitative methods

ويشمل:

- الملاحظات Observations
- المقابلات الشخصية Interviews

ثانياً : الأسلوب الخاص بالطرق الكمية Quantitative methods

ويشمل:

- مقاييس ليكرت Likert scales
- مقاييس تمايز المعانى Semantic differential scales

ومن الملاحظات الجديرة بالاهتمام فى هذا الخصوص هو أن أسلوب التقرير الذاتى يعطى نتائج أكثر ثباتا من الأساليب الأخرى، كما أظهرت نتائج بعض البحوث أن المعلمين والمدربين يصنفون على أنهم وسائل ضعيفة للتنبؤ بالاستجابات الانفعالية لكل من الطلاب واللاعبين، كما وجد كاول Cowell ١٩٨٥م أن العلاقة بين تقديرات المعلم وتقديرات الطالب لغيره من الطلاب بلغت ٥, ٠ بالنسبة للتكيف الاجتماعى، مما يعنى أن الاختلاف في التعبير عن السلوك الانفعالى للمفحوص كان كبيرا عندما قام بتقويم هذا التكيف كل من المعلم والطالب، وقد حصل نولاند وجروبر Noland and Gruber ١٩٧٨م على معامل ارتباط منخفض جداً بين تقديرات مجموعتين مختلفتين من المعلمين بالنسبة لمشكلة محددة على قائمة مراجعة Checklist.

٣ - أدوات القياس الانفعالى

يتضمن القياس الانفعالى فى المجال الرياضى العديد من الأدوات، بعض هذه الأدوات أعد للاستخدام فى الرياضة المدرسية، والبعض الآخر أعد للاستخدام فى مجال رياضة البطولة (الرياضية القمية)، والبعض الثالث منها أعد للاستخدام فى الدراسات والبحوث العلمية التى تستهدف التعرف على تأثير النشاط البدنى على الصحة الانفعالية emotional health ، هذا بالإضافة إلى بعض الأدوات التى أعدت لقياس التحسن الذى يحدث لمرضى القلق والإحباط عندما يتعرضون لبعض وسائل العلاج بالتمرينات.

ويستخدم لقياس السلوك الانفعالى فى مجال التربية البدنية والرياضة العديد من الأدوات والوسائل هى:

١ - قوائم الميول (الاهتمامات) Interests inventories

وهى عبارة عن أدوات قياس تستخدم أسلوب التقرير الذاتى للإجابة، بمعنى أن المفحوص يعبر بنفسه عما يحبه أو يكرهه بالنسبة لأشياء أو أفراد أو مواقف أو برامج أو أنشطة محددة.

وتستخدم قوائم الميول لاختيار الأنشطة الرياضية المفضلة لدى التلاميذ والطلاب والأعضاء عند إعداد البرامج الرياضية فى المدارس والجامعات والأندية الرياضية. وعادة ما تستخدم قوائم الميول فى المجال الرياضى عند بناء المناهج وإعداد محتوى البرامج وبخاصة البرامج الترويحية، وبرامج النشاط الرياضى الداخلى intramural فى المدارس.

وقد أظهرت دراسات الميول الرياضية أن الميل وعدم الميل نحو نشاط رياضى ما يتحدد فى ضوء مجموعة من العوامل هى:

- الخبرة السابقة.
- الضغوط الناشئة عن الأقران والزملاء.
- تأثير المعلم أو الأبوين.
- العوامل الثقافية والحضارية.

وقد تعكس الميول اتجاهات الفرد الرياضى نحو أشكال محددة من النشاط، فعندما يفضل أحد الأفراد كرة القدم على المصارعة، فإن هذا التفضيل ربما يعكس اتجاهاته نحو كل من كرة القدم والمصارعة.

٢ - **قوائم الاتجاهات** Attitudes inventories

هى أيضاً أدوات قياس تستخدم أسلوب التقرير الذاتى للإجابة، فالمفحوص يحدد توجهاته أو شعوره نحو أشياء خاصة، أو أحداث، أو أشخاص، أو أنشطة، أو تنظيمات أو قوانين اجتماعية محددة أو غيرها. فالاتجاه نحو الشىء يعنى إحساس الفرد نحو هذا الشىء، حيث يعتقد أن هذا الإحساس يمثل قوة دافعة تؤثر مستقبلا على اختيار هذا الشىء من عدمه.

فمن خلال استجابات الفرد على قوائم الاتجاهات يتم الكشف عما يشعر به الفرد (المفحوص) نحو أحداث أو أشخاص أو أنشطة أو أفكار أو سياسات

أو أيديولوجيات أو مؤسسات ونظم اجتماعية أو غيرها، فاتجاه الفرد نحو نشاط ما سوف يكون له تأثيره فى المستقبل على اشتراك الفرد فى هذا النشاط، فإذا كان اتجاه الفرد نحو نشاط ما سلبيا فإنه يتوقع ألا يمارس هذا الشخص هذا النشاط.

ويستهدف الكشف عن الاتجاهات فى المجال الرياضى تحديد الأفراد ذوى النظرة السلبية نحو النشاط البدنى physical activity بشكل عام، ونحو بعض الأنشطة الرياضية بشكل خاص، وذلك بغرض العمل على تغيير اتجاهات هؤلاء الأفراد عن طريق وضع البرامج المناسبة لذلك.

٣ - قوائم الروح الرياضية Sportsmanship iventories

وهى قوائم تستخدم أما أسلوب التقرير الذاتى أو التقارير التى يكتبها آخرون عن مدى التزام الفرد الرياضى بالقواعد والقوانين، ومظاهر الفوز والهزيمة. وغالبا ما ينظر إلى الروح الرياضية على أنها «الخلق Character» الذى يظهره الفرد الرياضى فى مواقف اللعب والمنافسات.

وقد يطبق المعلم أو المدرب الرياضى مقاييس التقدير rating scales فى بداية العام الدراسى أو الموسم التدريبى، وذلك بغرض التعرف على مظاهر السلوك المنافية للروح الرياضية، حتى يستطيع ملاحظة أصحاب هذا السلوك خلال العام الدراسى (الموسم التدريبى) ومحاولة التعرف على تأثير البرنامج الدراسى أو التدريبى عليهم.

٤ - مقاييس القيادة Leadership scales

عادة ما يسأل أعضاء الجماعات الرياضية sport groups (الفصل الدراسى أو الفريق الرياضى) لكى يحددوا فرداً منهم يصلح كقائد متميز للجماعة.

ويستخدم فى المجال الرياضى عدد من المقاييس والقوائم التى تطبق لقياس السلوك القيادى وبخاصة بالنسبة للمدربين. هذه المقاييس والقوائم عبارة عن تقارير تتم بمعرفة أعضاء الجماعة، حيث يطلب منهم تحديد أى فرد

فى الجماعة يفضلونه كقائد للفصل أو الفريق. ويعاب على هذا الأسلوب أنه لا يعطى لكل أفراد الجماعة (الفريق) فرصا متكافئة لممارسة القيادة، وللتغلب على هذه المشكلة فإنه يفضل تحديد واجبات خاصة بالقيادة لكل فرد من أفراد الفصل أو الفريق لكى يتدربو عليها خلال السنة الدراسية أو الموسم التدريبى.

٥ - قوائم النمو الاجتماعى Social development inventories

وهى عبارة عن وسائل (أدوات) صممت لقياس مظاهر النمو الاجتماعى. وقد تأخذ هذه الوسائل (القوائم) بعض أساليب التقرير الذاتى، أو التقديرات التى يضعها المعلم أو المدرب، أو التقديرات التى تضعها جماعة الرفاق peer group .

وتستهدف قوائم النمو الاجتماعى تقدير درجة تكيف الفرد للبيئة الاجتماعية social environment بالمدرسة أو النادى، كما توجد بعض الدلالات التى تعطينا مؤشرات عن مدى تقبل جماعة الرفاق (طلاب الفصل أو أعضاء الفريق) للفرد، أو ما يمكن أن نطلق عليه «شهرة أو شعبية الفرد - popularity» أو مكانة الفرد لدى الرفاق peer statuas .

٦ - تقديرات السلوك Behavior ratings

وتتم بمعرفة المعلم أو المدرب أو الباحث بعد ملاحظته سلوك المفحوص لفترة زمنية. ويستخدم هذا الأسلوب لقياس بعض الظواهر النفسية مثل : السلوك الانسحابى withdrawal، واضطراب العلاقات مع الرفاق، وعدم النضج immaturity، والولاء (الإخلاص) Loyalty، وضبط النفس self-control، والقيادة، والاجتماعية (القدرة الاجتماعية) sociability، والتعاون وغيرها.

٧ - قوائم (مقاييس) الشخصية Personality inventories

وتعرف باسم استبيانات (استخبارات) الشخصية، وهى عبارة عن أدوات (وسائل) تستخدم أسلوب التقرير الذاتى عن طريق الورقة والقلم، حيث يجيب المفحوص عن مجموعة من الأسئلة تصف أنماط السلوك المميز typical behavioral patterns.

وتتضمن قوائم الشخصية عبارات (أسئلة) كثيرة ترتبط بالخصائص الشخصية والسلوك، وبحيث يجيب الفرد (المفحوص) على هذه العبارات على أساس ما إذا كانت تنطبق أو لا تنطبق عليه. وتعتبر السمات traits كالتحكم الانفعالى emotional control ، والثقة بالنفس self - confidenc، والدافعية motivation، والعدوانية aggressiveness، والتصميم determination، والصلابة mental toughness وغيرها هو ما تستهدفه قوائم الشخصية.

وتتكون درجة المفحوص من عدد الأسئلة التى أجاب عليها فى الاتجاه الذى ينبغى أن تظهره السمة السلوكية التى يتم قياسها. وهناك بعض قوائم الشخصية التى تقيس سمة واحدة مثل القلق، وقوائم أخرى تقيس عدداً من السمات.

٤ - الطرق الوصفية للقياس الانفعالى

يقصد بالطرق الوصفية (الكيفية) qualitative للقياس الانفعالى تلك الطرق التى تعتمد على أساليب وإجراءات تحليلية وصفية لفهم وتفسير الظاهرة المستهدفة من عملية القياس حيث تقدم هذه الطرق معلومات فيما يتعلق بالمفحوصين فى إطار عدد كبير من المتغيرات والظروف فى واقعها الفعلى.

وتعد الملاحظات الميدانية والمقابلات الشخصية من الطرق المهمة للحصول على البيانات الوصفية (النوعية) qualitative data. وفيما يلى شرح لكل طريقة (وسيلة) من هذه الطرق.

أولاً : طرق الملاحظة Observation

تعتبر الملاحظة وسيلة مهمة من وسائل جمع البيانات فى مجال القياس النفسى. ويقصد بالملاحظة أن يقوم آخر (المعلم أو المدرب أو غيرهما) بملاحظة الشيء، هذا الشيء قد يكون التلميذ أو اللاعب أو الطريقة أو البرنامج أو غير ذلك من الأشياء المستهدفة بالملاحظة. ولكى يتحقق للملاحظة المصداقية والثراء فى المعلومات التى يمكن أن توفرها، فإنه يلزم أن تتم الملاحظة أثناء وقوع الأحداث الفعلية، كأن نلاحظ اللاعب أثناء المباريات، أو البرنامج أثناء التطبيق، أو المعلم أثناء التدريس، حيث يطلق على هذا الأسلوب من أساليب الملاحظة اسم: الملاحظة فى الموقع on- site observation.

ولعل من أهم ما يميز طريقة الملاحظة ما يأتى:

- أنها تتيح الفرصة للتعامل مع الأشياء أو الأحداث عن قرب، فهى تتيح للملاحظ إمكانية التحدث مع التلاميذ أو اللاعبين أو المعلمين أو غيرهم من موضوعات الملاحظة.
- أنها تضع القائمين بها فى أقرب مكان لسير العمل (البرنامج أو المشروع) مما يمكنهم من استكشاف الحقائق من مصادرها الحقيقية، وتسجيل ما يمكن رؤيته خلال سير العمل بدقة.
- أن الملاحظة تستهدف أن تقدم للقائم بعملية التقويم تقريراً مكتوباً عن الشيء المستهدف من الملاحظة، وعليه يجب أن تكون البيانات المتجمعة عن طريق الملاحظة بيانات عميقة وبإسهاب، وأن تتناول هذه البيانات وصف الأشياء وصفاً دقيقاً وشاملاً، حتى يتمكن القارئ (القائم بالتقويم) من فهم ما يحدث وكيف يحدث، استناداً إلى المقولة التى تزعم أن «القائم بالملاحظة هو أذن وعين القارئ أو القائم بعملية التقويم».

- أن صدق الملاحظة - كإحدى طرق (وسائل) جمع البيانات - يتوقف على مهارة القائم بها، لأنه وفقا لما أقره جوبا ولينكلن Guba and Lincoln ١٩٨١م من أن القائم بالملاحظة هو أداة القياس observer is the instrument، لذا نجد أن القائم بالملاحظة يمكن أن يؤثر على النتائج وفقا لمتغيرات التعب، والخبرة السابقة، والتدريب، والمهارة والقدرة، والخلفية الثقافية والحضارية، وغيرها.

وتصنف الملاحظة كوسيلة من وسائل جمع البيانات النوعية qualitative data . وقد استخدمت الملاحظة فى معظم الدراسات المبكرة التى أجريت فى مجال الدراسات النفسية فى بداية القرن الماضى، وما تزال من الوسائل المهمة فى جمع البيانات فى الوقت الحاضر وذلك بعد الاستعانة بأجهزة الفيديو لتسجيل سلوك المفحوص ثم عرض هذا السلوك مرة أخرى للقيام بعمليات التحليل اللازمة. وقد لاحظ الباحثون أنه عندما يشعر المفحوص بأنه يراقب ويتم تصويره بالفيديو، فإنه يغير من سلوكه الحقيقى. وللتغلب على هذه المشكلة كان لزاما على القائم بالملاحظة أن يتجاهل المفحوص المستهدف من الملاحظة، وألا يهتم به حتى لا يشعر أنه المستهدف من الملاحظة، كما يفضل ملاحظة نفس المفحوص وتسجيل سلوكه لأكثر من مرة فى أيام متباعدة حتى يمكن الوقوف على سلوكه بشكل طبيعى.

أنواع الملاحظة :

تنقسم الملاحظة إلى نوعين هما :

- الملاحظة غير النظامية informal observation
- الملاحظة النظامية formal observation

الملاحظة غير النظامية:

وهى أبسط أنواع الملاحظة وأكثرها شيوعا، وهى عبارة عن الانطباعات الشخصية لأى شخص خارج الإطار الرسمى للمؤسسة أو الهيئة أو المشروع أو البرنامج قدر له أن يشهد الأحداث خلال سيرها عن قرب، بحيث يلاحظ بمحض الصدفة التسهيلات والإمكانات المتاحة للمشروع أو البرنامج، والتفاعل بين التلاميذ والمعلم أو اللاعبين والمدرب، أو بين أعضاء الفريق أو غير ذلك.

وتمتاز الملاحظة غير النظامية بأنها تعطى للقائم بها إحساساً حقيقيا بسياق وبيئة العمل أو البرنامج، ولكنه يعاب عليها أنها غير دقيقة حيث يطلق عليها اسم: الملاحظة غير المنتظمة unsystematic observation أو غير الرسمية أو الحرة، وذلك للأسباب التالية:

- أن القائمين بها قد يكونون غير معدين لذلك، ومن ثم فهم لا يعرفون ما ينبغى البحث عنه.
- أنهم قد يسيئون تفسير ما يلاحظونه.
- أنهم قد يقومون بالملاحظة فى الوقت غير المناسب، وأنهم قد لا يستغرقون فى الملاحظة وقتا كافيا.

الملاحظة النظامية:

هى الملاحظة المرتبطة بالقضايا ذات الأهمية بالنسبة لموضوع البحث، حيث يخطط لها تخطيطا جيداً، ويتم تدريب الملاحظين على كيفية وصف وتسجيل ما ينبغى ملاحظته وكيفية تسجيله.

وتتضمن الملاحظة الرسمية الخطوات التالية:

١ - إعداد قائمة بالسمات أو الصفات أو الخصائص التى ينبغى أن يكون عليها الشىء المطلوب ملاحظته.

٢ - إعداد قائمة بمظاهر السلوك أو الأحداث التى لا ينبغى حدوثها.

٣ - تحديد عدد المرات المطلوبة لإجراء الملاحظة.

٤ - تحديد المدة الزمنية التى تستغرقها كل مرة من مرات الملاحظة .

٥ - إعداد استمارات التسجيل.

٦ - اختيار الملاحظين.

٧ - تدريب الملاحظين.

٨ - القيام بالملاحظة.

٩ - تجميع البيانات واعدادها للتحليل والتفسير والعرض.

وتستخدم الملاحظة فى المجال الرياضى فى ملاحظة الطالب (اللاعب) فى المواقف الطبيعية، وبناء على ذلك يمكن الاستفادة منها كأسلوب لجمع البيانات على النحو التالى:

- يمكن الاستفادة من المعلومات التى يتم جمعها عن طريق الملاحظة كمعلومات إضافية بالنسبة للمعلومات التى يتم جمعها بوسائل أخرى.

- أن الملاحظة المنتظمة للطالب (اللاعب) قد تمد المعلمين والمدربين بمعلومات من غير الممكن الحصول عليها بطريقة أخرى. فالملاحظة تهيئ الفرصة لكى يرى الملاحظ العديد من مؤشرات السلوك التى تكشف عن الميول والاتجاهات، هذه الفرصة قد لا تتاح لأى وسيلة أخرى تطبق فى مواقف اصطناعية (غير طبيعية) artificial situations.

- أن الملاحظة تعطى الفرصة لمعاينة sampling السلوك الفعلى للطالب أو اللاعب. ووفقا للملاحظة النظامية، فإنه لا يتم سؤال الطالب أو اللاعب عما يشعر به - كما فى حالة الإجابة على الاستبيانات، أو فى حالة تطبيق قوائم التقرير الذاتى - وإنما يتم التوجه مباشرة إلى الملاحظة التى تكون أكثر صدقا من غيرها من الوسائل بالنسبة لبعض الظواهر.

- عندما يقوم المعلم أو المدرب بالملاحظة، فإن ذلك يتيح الفرصة لإقامة علاقات طيبة بينه وبين الطلاب (اللاعبين)، كما تمكنه من مساعدة الطلاب (اللاعبين) والاستماع إليهم ومحاولة فهمهم بعمق.

ولعل من أهم ما يؤخذ على أسلوب الملاحظة هو أن المعلومات التى يتم جمعها قد تتعرض للتحيز، بسبب أن المفحوصين قد يغيرون من سلوكهم عندما يعلمون أنهم خاضعون للملاحظة.

ومن ناحية أخرى تتميز طرق الملاحظة عن قوائم التقرير الذاتى، فى أن الملاحظة قد تتخذ إما موقفا منظما structured أو موقفا غير منظم unstructured situation . فالموقف المنظم هو الموقف الذى يتم التحكم فيه وضبطه والعكس بالنسبة للموقف غير المنظم. ويرى علماء القياس النفسى أنه يفضل بالنسبة لسمات الشخصية ملاحظة المفحوص بعناية فى الموقفين المنظم وغير المنظم حتى يمكن الاعتداد بنتائج الملاحظة كأساس لتقويم سمات الشخصية.

الملاحظة والاختبارات الموقفية:

الاختبارات الموقفية situational tests عبارة عن اختبارات تعطى للمفحوص وهو فى موقف حقيقى لرصد تصرفاته وردود أفعاله ، ولهذا يطلق عليها اسم: الاختبارات السلوكية behavioral tests . وبناء على ذلك تصنف الملاحظة كوسيلة قياس موقفية فى مجال القياس الانفعالى وذلك لكونها تهتم مباشرة بسلوك المفحوص فى مواقف فعلية.

تدريب وإعداد القائم بالملاحظة:

تستهدف الملاحظة تقديم وصف كامل عن كافة التفاصيل عن سلوك المفحوص فيما يتعلق بالموضوع المطلوب ملاحظته. ولما كانت الملاحظة تتميز بأنها مصدر مباشر للبيانات لكونها تتطلب من القائم بها الذهاب إلى ميدان العمل field work الحقيقى، لذا تتوقف دقة البيانات المستمدة من الملاحظة على الآتى:

- مهارة ومقدرة القائم بالملاحظة.
- تدريب القائم بالملاحظة.
- قدرته على الاستماع والتركيز.
- الصبر وسعة الصدر.
- القدرة على تحديد الظواهر ووصفها بدقة.
- القدرة على تحمل الأعباء الذهنية والبدنية.

ومن المعروف أن الناس لا تثق فى البيانات المتجمعة من الملاحظة التى يقوم بها أفراد غير مدربين لم يتم إعدادهم للقيام بها، ومن المتفق عليه أن طرق الملاحظة تتطلب نظاماً للتدريب والإعداد يستهدف خمس وظائف رئيسية هى:

- الإحساس والبصيرة Sense and Sight
- الذوق Taste
- الاستماع Hearing
- الشم Smell
- اللمس Touch

وذلك بغرض الآتى:

- تعلم كيفية كتابة التقارير.
- وكيفية تسجيل الملاحظات الميدانية.
- وكيفية كتابة التفاصيل دون تداخل بين الموضوعات المختلفة.

وحيث يتضمن التدريب الآتى:

- الإعداد العقلى المعرفى.

- الإعداد البدنى.

- الإعداد النفسى.

وفى هذا الصدد يقول العالم الشهير لويس باستور Louis pasteur (١٨٢٢ - ١٨٩٥): إن الملاحظة الميدانية تتطلب إعداد العقل (الذهن) عن طريق تعلم كيفية التركيز أثناء الملاحظة، وأن التدريب يجعل القائم بالملاحظة قادراً على جعل بيانات الملاحظة أكثر دقة وصدقاً وثباتاً.

ثانيا : المقابلة الشخصية Interview

المقابلة الشخصية عبارة عن لقاء بين شخصين أو أكثر وجهاً لوجه، حيث يقوم فيها المفحوص (الفرد الرياضى) بالإجابة على الأسئلة التى يطرحها عليه الباحث أو القائم بعملية المقابلة Interviewer . وقد تكون الأسئلة الموجهة للمفحوص محددة مسبقا، ومع ذلك فللقائم بالمقابلة الحرية فى توجيه ما يراه من أسئلة تقتضى استجابات ذات أهمية خاصة. وعادة ما يتم تسجيل ملاحظات مختصرة عن استجابات المفحوص أثناء المقابلة، فى حين يتم إعداد ملخص كامل لهذه الاستجابات فيما بعد الانتهاء من المقابلة.

متى تستخدم المقابلات الشخصية؟

تستخدم المقابلات الشخصية للحصول على بيانات ومعلومات عن الناس الذين لا يستطيعون القراءة والكتابة، والناس الذين لا يتحدثون لغة أخرى غير اللغة الأصلية لهم، والناس الذين يواجهون صعوبات فى فهم وكتابة الأسئلة التحريرية.

وتتسم المقابلات الشخصية بالمرونة، كما أنها تعطى للقائم بالمقابلة الفرصة لكى يتتبع الأفكار والاتجاهات الرئيسية للمفحوص، وإن كان يعاب عليها أنها تستنفد الكثير من الوقت، وأنه قد يحدث أن يؤثر القائم بالمقابلة على استجابات المفحوصين وجميعها من الأمور غير المستحبة فى المقابلة.

المقابلة الشخصية كاختبار:

تصنف المقابلة الشخصية كأداة لجمع البيانات والمعلومات عن طريق الاستجابة للكلمة المنطوقة word - of mouth ، بمعنى أن القائم بالمقابلة يتكلم والمفحوص يسمع ويرد بالكلمة المنطوقة. وتتم المقابلة - عادة - فى غياب الأقران مما يعطى الفرصة لملاحظة ردود أفعال المفحوصين فى سرية تامة بعيداً عن الناس.

ووفقاً لمجموعة من الاعتبارات يمكن النظر إلى المقابلة الشخصية على أنها اختبار، يشبه أى اختبار نفسى، فالمقابلة الشخصية طريقة لجمع البيانات أو المعلومات عن الفرد (المفحوص)، حيث تستخدم هذه المعلومات فى وصف هذا الفرد، أو التنبؤ بسلوكه فى المستقبل، أو بكليهما معاً.

ويمكن للقائم بالمقابلة تقويم صدق وثبات المقابلة كما يحدث بالنسبة لأدوات القياس الأخرى، فالمقابلة الشخصية - مثل أى اختبار - تتضمن التفاعل بين شخصين أو أكثر. وبعض المقابلات تأخذ طريقها فى التطبيق مثل الاختبارات الفردية، بمعنى أن القائم بالمقابلة يتعامل فقط مع مفحوص واحد فى ذات الوقت.

وفى بعض المقابلات الشخصية - وبخاصة المقابلات الشخصية للأسرة - يقوم فرد واحد فقط (القائم بالمقابلة أو الباحث) بإجراء المقابلة الشخصية مع اثنين أو أكثر من أفراد العائلة فى نفس الوقت، تماماً كما يحدث بالنسبة للتطبيق الجمعى لأى اختبار.

وكأى اختبار، يكون للمقابلة الشخصية هدف محدد، ومجالات اهتمام محددة أيضا، كما تقع على المسئول عن إجراء المقابلة مسئوليات تزيد فى أهميتها وخطورتها عن تلك المسئوليات التى تقع على القائم بتنفيذ وإدارة الاختبار.

ويمكن النظر إلى المقابلة الشخصية كاختبار مهم جدا، وبخاصة فى بعض الحالات التى تكون فيها الاختبارات المستخدمة تفتقر إلى الدقة، ففى حالة اختبار تفهم الموضوع thematic apperception test (TAT) – على سبيل المثال – تستخدم المقابلة الشخصية كأحد الوسائل المهمة للحصول على البيانات المناسبة والمكملة للبيانات التى يتم الحصول عليها من الاختبار، فما تزال المقابلات الشخصية من أهم مصادر الحصول على المعلومات فى حالات استخدام الأساليب الإسقاطية فى القياس النفسى.

ومن ناحية أخرى تعتبر المقابلة الشخصية الطريقة الرئيسية لجمع البيانات فى مجال الطب النفسى العلاجى clinical psychiatry، وفى بعض المجالات الأخرى كمجال الصحة العامة وطب المجتمع، ومجال الصناعة، والإرشاد والتوجيه النفسى فى مجال التعليم، وعلم النفس الإصلاحى correctional psychology .

وتستخدم المقابلة الشخصية أيضاً كأداة لجمع البيانات فى مجال الدراسات الاجتماعية وبخاصة الدراسات المتعلقة بظواهر كالزواج والطلاق وغيرها، وفى مجالات دراسات الإدارة والتسويق والاستهلاك وغيرها.

علاقة المقابلة الشخصية بالملاحظة:

أساس المقابلة الشخصية هو «المقابلة وجهاً لوجه face- to face interview» ومن ثم يمكن القول بأن المقابلة الشخصية تعتمد على الملاحظة observation فى جمع البيانات والمعلومات، وأن المقابلة الشخصية يمكن أن تتضمن الملاحظة، فالباحث الماهر فى المقابلة سوف يكون ماهرا أيضا فى الملاحظة، بمعنى أنه سوف يكون قادراً على قراءة الرسائل غير اللفظية nonverbal messages الصادرة عن المفحوص، وسوف يكون حساساً لكيفية إجراء المقابلة، وكيف أنها تتأثر بما يقوله، وكيف يتعامل بحذر مع من يجرى معهم المقابلة من حيث التفاعل والعلاقة معهم.

ومن ناحية أخرى، تعد مهارات المقابلة أساسية للقائم بالملاحظة (الملاحظ)، لكون القائم بالملاحظة يقوم بالزيارات الميدانية للتعرف على الواقع، ومن ثم فهو يحتاج إلى مقابلة الناس والتحدث معهم وأن يشرح لهم ويستمع منهم.

متطلبات نجاح المقابلة الشخصية:

يقتضى القيام بالمقابلة الشخصية بكفاية ونجاح أن يتخذ القائم بها بعض الإجراءات والتدابير المهمة وهى:

١ - أن يكون للمقابلة هدف محدد ومقبول.

٢ - العمل على إقامة صلة الوئام والثقة مع المفحوص، حيث يتوقع عدم تعاون المفحوص إذا فقد الثقة فى القائم بالمقابلة (الباحث)، فالطالب أو اللاعب الذى تعرض لعقوبة وقعت عليه من المعلم أو المدرب سوف لا يتقبل إجراء المقابلة مع أى منهما.

٣ - أن يشرح القائم بالمقابلة للهيئة المعنية (المدرسة أو النادى أو غيرهما) الغرض من المقابلة، واسم الهيئة التى يمثلها القائم بالمقابلة، وأهمية المعلومات التى سيتم الحصول عليها بالنسبة للبحث العلمى.

٤ - يجب إعطاء الفرصة للمفحوصين للتعبير عن مشاعرهم، وعلى القائم بالمقابلة عندما يوجه الأسئلة للمفحوص ألا يوحى إليه بإجابات معينة، وعليه أيضا أن يضمّن المقابلة أسئلة تستهدف التأكد من المعلومات التى قدمها المفحوص، لأن ذلك يعمل على تخفيف الصورة التى قد تكون لدى المفحوص والتى تتمثل فى «أن شخصاً يقوم باستجوابه».

٥ - يجب على القائم بالمقابلة أن يتجنب توجيه الأسئلة التى يصعب فهمها، وفى جميع الحالات يجب عليه تجنب استخدام النبرة التهديدية عند توجيه الأسئلة إلى المفحوصين.

٦ - على القائم بالمقابلة (الباحث)، أن يطمئن المفحوص وأن ينصحه - قبل أن تبدأ المقابلة - بالتعاون، وأن يؤكد له سرية المعلومات التى سيتم الحصول عليها، وأنها لن تستخدم إلا فى أغراض البحث العلمى.

٧ - على القائم بالمقابلة أن يسجل كتابة وبشكل مختصر النقاط الرئيسية للمقابلة وذلك بعد الانتهاء من المقابلة مباشرة وقبل أن يغادر المفحوص (المفحوصون) مكان المقابلة.

وبشكل عام، يتوقف نجاح المقابلة الشخصية على نجاح القائم بها فى إقامة الألفة بينه وبين المفحوص، لكى يسمح له ذلك بوصف أحاسيسه والتعبير عنها بصدق، ومن المهم كذلك ألا يصدر القائم بالمقابلة أية أحكام على استجابات المفحوص، وفى جميع الحالات يجب أن يكون القائم بالمقابلة مستمعاً جيداً.

الشروط الواجب توافرها فى القائم بالمقابلة:

يشترط فى القائم بالمقابلة الشخصية الآتى:

- أن يكون مدربا تدريباً جيداً على كيفية إجراء المقابلة.
- أن يضع الأسئلة فى ترتيب منطقى، وعند طرحها لا يقفز من موضوع لآخر، فعند الانتهاء من الأسئلة التى تتناول موضوعاً واحداً، يتم الانتقال إلى الأسئلة التى تتناول موضوعاً آخر.
- أن يقوم بعمل مقدمة للمقابلة قبل البدء فى توجيه الأسئلة.
- أن يكون مدرباً على تقبل الإجابات المفاجئة والعدوانية من بعض المفحوصين.
- أن يعمل جاهداً على أن يزيل من المفحوصين الإحساس بأنه مندوباً أو ممثلاً للسلطة.
- أن يكون محايداً، وأن تكون لدية القدرة على التعامل مع المفحوصين الذين يتسمون بالعناد.

تسجيل المقابلة الشخصية:

ينبغى - قدر الإمكان - تسجيل المقابلات الشخصية وبخاصة المقابلات غير المنظمة على شرائط كاسيت كوسيلة للاحتفاظ بالمعلومات وحمايتها من الضياع، واسترجاعها فيما بعد لتقديمها للتحليل الإحصائى، فالمقابلات المسجلة تمكن القائم بالمقابلة (الباحث) من تلخيص المعلومات وذلك عن طريق اقتباس مقتطفات من حديث المفحوص. ويقوم القائم بالمقابلة بتدوين الملاحظات أثناء المقابلة لربطها مع المعلومات المسجلة على الشرائط للحصول على نتائج على مستوى عال من الموضوعية.

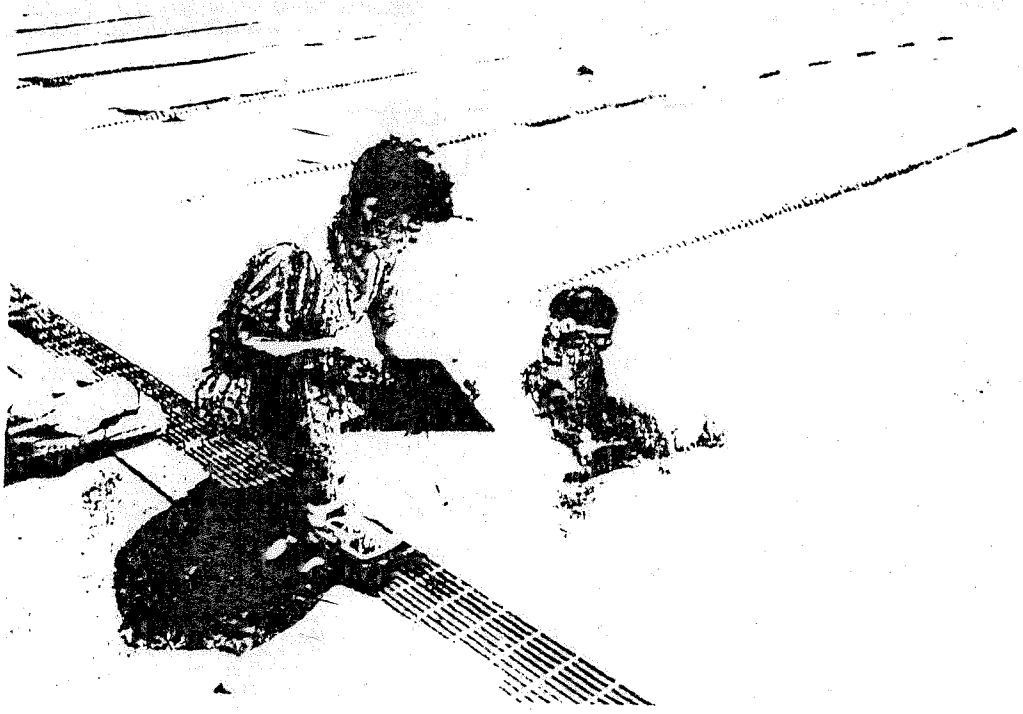
وقد يعاب على طريقة التسجيل أن بعض المفحوصين لا يفضلون تسجيل المقابلات معهم، بالإضافة إلى ما قد تستغرقه هذه الطريقة من وقت ومن الكثير من عدد الشرائط المسجلة، وأيضاً ما قد يستغرقه نقل الشريط وتدوين ما به من معلومات من وقت وجهد. فقد أظهرت الدراسات أن هذا الإجراء قد يستغرق ضعفين أو ثلاثة أضعاف الزمن الذى تستغرقه المقابلة ذاتها. وقد يكون البديل لذلك هو أن يقوم القائم بالمقابلة بتدوين الملاحظات خلال المقابلات غير المنظمة، كما ينبغى إعداد ملخصاً عاماً لكل استجابة، بالإضافة إلى التسجيل الحرفى لبعض العبارات الأكثر شيوعاً.

أنماط المقابلات الشخصية:

يمكن تحديد ثلاثة أنماط للمقابلات الشخصية هى:

١) المقابلة محددة الإطار (المنظمة) structured interview

وتعرف باسم المقابلة النظامية (المقننة أو المنهجية أو الرسمية) formal interview، وهى مقابلة تقدم فيها الأسئلة إلى المفحوص وفقا لترتيب وتسلسل محدد، وبحيث تصاغ هذه الأسئلة على أساس نوع المعلومات المطلوب الحصول عليها من المفحوص، ويتم تسجيل الإجابات وفقا لصيغ (أشكال أو صور) مقننة standardized forms.



فى رياضة السباحة تستخدم قائمة المراجعة كوسيلة لتقويم العمليات ونواتج التعلم



يمكن استخدام الكاسيت كوسيلة مساعدة لاسترجاع البيانات فى المقابلة الشخصية

وتعرف المقابلة المنظمة أيضاً باسم المقابلة النمطية patterned interview لكونها تتطلب تحديد الأسئلة تحديداً جيداً قبل أن تبدأ المقابلة، ولكونها أيضا تتبع إجراءات منظمة systematic proced فى جمع المعلومات عن المفحوص، ولهذا السبب تتميز المقابلات المنظمة بأنها تعطى بيانات أكثر صدقا من المقابلات غير المنظمة.

ويتوقف الإعداد الجيد للمقابلة المنظمة على التحديد المسبق للأسئلة، وتسلسل توجيهها للمفحوص، وطرق تقدير الدرجات، وتدريب القائمين على المقابلة. وقد يحتاج القائم بالمقابلة المنظمة عمل بروفة للأسئلة التى يتم اختبارها للمقابلة، ويمكن للقائم بالمقابلة (الباحث) أن يقوم بهذه البروفة مرة أو مرتين مع أناس آخرين من غير المفحوصين المستهدفين من المقابلة وقد يكون هؤلاء الناس الزوجة أو أحد الأبناء أو أحد الزملاء أو غيرهم.

ويعتبر هذا الإجراء بمثابة اختبار لكل من الأداة (المقابلة) والقائم بها (بإجراء المقابلة) interviewer، كما يعتبر من الوسائل المهمة لتجنب الأسئلة الصعبة، والأسئلة غير المفهومة، وأيضا تجنب عدم الاتساق فى منطق ترتيب وتتابع الأسئلة.

ويرى جيلفورد Guilford ١٩٦٦م أنه قد يعاب على المقابلة المنظمة الآتى:

- أنها تستغرق وقتا طويلا وجهداً كبيراً فى إعداد الأسئلة وتنفيذ المقابلة، ومن ثم يقترح جيلفورد استخدام المقابلات المنظمة فى الحالات التى يصعب فيها تطبيق المقابلات غير المنظمة.
- أن صدق وثبات وموضوعية المقابلات الشخصية عموماً أقل فى العادة من الوسائل التى تستخدم الورقة والقلم وبالتحديد قوائم التقدير inventories وبعض الطرق الأخرى الشائعة كالاختبارات النفسية المقننة أو تحليل بيانات السيرة الذاتية (السجل الشخصى) bio- data analysis.

٢) المقابلة غير محددة الإطار (غير المنظمة) Unstructured interview

وتعرف باسم المقابلة غير النظامية (غير المقننة)، وهى نمط لا يتضمن وجود مجموعة من الأسئلة ولا ترتيب محدد لتوجيه الأسئلة، وإنما يترك للقائم بالمقابلة أن يسأل بعض الأسئلة بأى طريقة يراها مناسبة وفقا لعدد من الاعتبارات أو النقاط المتعلقة بما يجب أن تتضمنه المقابلة.

وتعد المقابلة غير المنظمة أكثر استخداماً من المقابلة المنظمة، وذلك لكونها تتميز بأنها طريقة مرنة لجمع البيانات، ولكونها طريقة يكون فيها القائم بالمقابلة أكثر اجتهاداً وتعمقاً فى فهم سلوك المفحوص، فعن طريق حريته فى توجيه الأسئلة يستطيع أن يشكلها بالأسلوب الذى يضمن له الحصول على الاستجابات المستهدفة من المقابلة.

ويؤخذ على المقابلة غير المنظمة ما يأتى:

- أنها تتأثر بالقائم بالمقابلة أكثر منها بالنسبة للمقابلة المنظمة، وعليه تصبح المقابلة غير المنظمة أقل ثباتا من المقابلة المنظمة.
- أنها تتطلب أن يكون القائم بها أكثر مهارة من القائم بالمقابلة المنظمة، وبخاصة فيما يتعلق بالذكاء والحس الاجتماعى والمعرفة العميقة والشاملة بالنسبة لموضوعات المقابلة، وهو ما قد يصعب تحقيقه بالنسبة لبعض الدراسات والبحوث النفسية.
- أنه يصعب تحويل البيانات المتجمعة من المقابلات غير المنظمة إلى بيانات كمية للقيام بتحليلها، وذلك بسبب بعض الصعوبات المتعلقة بحرية توجيه الأسئلة للمفحوص وما يمكن أن ينتج عن ذلك من اختلاف فى طبيعة الأسئلة للمفحوصين وما يتبع ذلك من اختلاف فى طبيعة النتائج التى يتم الحصول عليها من استجابات المفحوصين.

٣) المقابلة شبه المحددة Sem-structured interveiw

وهى نمط يجمع ما بين التحديد والمرونة فى توجيه الأسئلة، فكل مفحوص يستجيب لمجموعة عامة من الأسئلة، فى حين تستخدم أسئلة للمتابعة و"جس النبض" وهى أسئلة تعتمد على طبيعة استجابة المفحوص.

ويعد هذا النمط من المقابلات الشخصية من أكثر الأنماط استخداما فى مجال بحوث علم النفس الرياضى، وبخاصة عندما تتم المقابلة مع الأبوين أو مع المدرب الرياضى، وبشكل عام، تستخدم المقابلات الشخصية مع الأبوين كوسيلة مفيدة جداً لتقويم السلوك الانفعالى لأبنائهم سواء أكانوا طلاباً أو لاعبين، وذلك لكون الأبوين يمثلان مصدراً حقيقياً للمعلومات المتعلقة بالمشكلات السلوكية لأبنائهم. ويمكن إجراء المقابلات مع الأبوين بدعوتهم لزيارة المدرسة أو النادى، أو الذهاب إليهم فى منازلهم، وإن كان علماء القياس النفسى يرون أن زيارة الأبوين فى المنزل تعطى القائم بالمقابلة (الباحث) صورة حقيقية عن الحياة الأسرية للطالب أو اللاعب (المفحوص).

طبيعة الأسئلة فى المقابلات الشخصية:

ينظر للسؤال question فى المقابلة الشخصية كمثير يستهدف إحداث أو توليد الاستجابة من الشخص الذى يتم مقابلته (المفحوص)، فالكلمات والألفاظ التى تتضمنها الأسئلة تعد من أكثر العناصر أهمية بالنسبة لتحديد كيف يستجيب المفحوص بالنسبة لتلك الأسئلة. ويتفق علماء القياس النفسى على أن توجيه الأسئلة فى المقابلات الشخصية هو بالدرجة الأولى «فن an art» ، وأن الأسئلة الجيدة هى التى تمكن الباحثين من القيام بالتقويم الكمى للنتائج، وأنه يجب أن يراعى فى تلك الأسئلة ما يأتى:

١ - أن تكون الأسئلة غير محددة (مفتوحة الطرف):

Open - ended questions

وهى أسئلة تتيح الفرصة للشخص الذى تجرى معه المقابلة (المفحوص) لكى يستجيب مستخدماً مصطلحاته وألفاظه وتعبيراته الشخصية وحيث تكون استجابته غير مقيدة بزمن.

وتعرف المقابلة التى تستخدم هذا النوع من الأسئلة باسم: المقابلة مفتوحة النهاية open-ended interview أو المقابلة حرة الاستجابة. وقد يتصور بعض الباحثين أن طريقة استخدام الأسئلة مفتوحة الطرف يعنى التخلى عن الاستجابة المحددة structured response، ولكن المقصود هو أن الأسئلة مفتوحة الطرف تعطى الفرصة للمفحوص ليجيب على الأسئلة إجابات غير نمطية وغير معدة مسبقا مثلما يحدث فى حالة الأسئلة المغلقة closed questions .

ومن أمثلة الأسئلة غير المحددة (مفتوحة الطرف)، هو أن نسأل أحد الأعضاء عن برنامج النشاط الرياضى الذى يقدمه النادى الأسئلة التالية:

- ما هو شعورك بالنسبة للبرنامج الرياضى الذى يقدمه النادى؟
- ما هو رأيك فى البرنامج الرياضى الذى يقدمه النادى؟

فمثل هذه الأسئلة تتيح الفرصة للمفحوصين لكى يحددوا اتجاهات الإجابة واستخدام الكلمات التى يريدونها.

٢ - أن تكون الأسئلة واضحة: Asking clear questions

من مسئولية القائم بالمقابلة أن يوضح للمفحوص ما الذى يسأل عنه. كما أن توجيه الأسئلة سهلة الفهم يمثل جزءا مهما بالنسبة لبناء الألفة والانسجام rappport بين المفحوص والقائم بالمقابلة. كما أن الأسئلة غير الواضحة يمكن أن تجعل المفحوص يشعر بعدم الراحة، وأنه عديم المعرفة ignorant، بالإضافة إلى البلبة والارتباك. فالأسئلة التى تركز على موضوع محدد تساعد على جعل الاستجابات واضحة. ولعل من أهم العوامل التى يجب مراعاتها

بالنسبة للمقابلات الشخصية هى لغة الناس الذين سيتم مقابلتهم، حيث يلزم أن يقوم الباحث (القائم بالمقابلة) بالتعرف على مستوى اللغة التى سوف يستخدمها فى الحديث مع المفحوصين وبخاصة المفردات والمفاهيم والمصطلحات التى سوف يستخدمها بالنسبة لموضوع معين.

ومن المتطلبات المهمة لإنجاح المقابلات الشخصية هو أن تكون اللغة المستخدمة سهلة الفهم واعتبار ذلك إطاراً مرجعياً مهماً بالنسبة لنجاح المقابلة. وفى أثناء المقابلة يجب التنبه إلى اللغة التى يستخدمها المفحوص للتعبير عن رأيه لوصف المشتركين فى البرنامج أو وصف الأنشطة، لأن مثل هذه الإجراءات يمكن القائم بالمقابلة من استخدام نفس مفردات اللغة التى يستخدمها المفحوص وبخاصة فى أثناء فترات الراحة، فالأسئلة التى تستخدم لغة المفحوص هى نمط الأسئلة الأكثر وضوحاً بالنسبة لمثل هذا المفحوص.

٣- أن يتم توجيه الأسئلة-سؤلاً بعد الآخر Asking singular questions

قد يحدث أن يقوم القائم بالمقابلة بإلقاء عدد من الأسئلة مرة واحدة لكى يجيب عليها المفحوص، وقد يعزى ذلك إلى عدم وضوح ما الذى يريد أن يسأل عنه القائمون بالمقابلات.

وفى حالات المقابلات العميقة depth interviews يجب توجيه الأسئلة وفق تسلسل محدد وبحيث يجيب عليها المفحوص سؤلاً سؤلاً، وفق المثال التالى: - القائم بالمقابلة يسأل:

«الآن أحب أن أسألك عما تراه من جوانب القوة والضعف بالنسبة لبرنامج النشاط الرياضى بالنادى، بعد ذلك عليك أن تقترح التحسينات التى يمكن إدخالها على البرنامج»

أولا: ما هى جوانب القوة فى البرنامج؟

(المفحوص يجيب)

- القائم بالمقابلة: حسنا، وماذا عن جوانب الضعف ؟

(المفحوص يجيب)

- القائم بالمقابلة: ما هى مقترحاتك لتطوير البرنامج الحالى؟

(المفحوص يجيب)

فى هذا المثال يتبين مدى وضوح ودقة الأسئلة، كما يتضح أنها تقدم للمفحوص سؤلاً بعد الآخر وفق تسلسل منطقى، وأن المثال لا يتضمن إعطاء سؤالين فى وقت واحد، وأن الأسئلة مفتوحة الطرف وغير مقيدة بزمن للإجابة.

٤ - أن تكون هناك مجسات وأسئلة للمتابعة

Probes and follow - up questions

تستخدم «المجسات probes» لتعميق الإستجابة على السؤال، وزيادة وثراء البيانات التى يتم الحصول عليها، كما تستخدم كمؤثرات cues تقدم للمفحوص لتحسين مستوى الاستجابة على السؤال.

ويمكن النظر إلى المجسات probes على أنها أسلوب للمقابلة يستخدم للدخول بعمق فى إستجابات المفحوص، وتستخدم المجسات فى الحديث، وتقدم بأسلوب وبصوت طبيعى، وتستخدم لمتابعة الإستجابات الأولية intitial responses، وتتكون مجسات الحديث من مجموعة طبيعية من الأسئلة الموجهة للتفاصيل detail-oriented questions ، وهى أسئلة تبدأ دائما بكلمات رئيسية هى: من who وأين where وماذا what، ومتى when، وكيف how، لأن مثل هذه الكلمات التى تبدأ بها الأسئلة يمكن أن تستخدم لإعطاء تفاصيل كاملة عن الموضوع الذى نسأل عنه، كما أنها تستخدم لتغيير مسار الاستجابة أو لجعل المفحوص يتكلم أكثر عن الموضوع الذى يسأل عنه.

ويمكن النظر إلى "المجسات" على أنها فن ومهارة التحقق من دقة الاستجابة، وهى تأتى من خلال معرفة ما الذى ننتظره من المقابلة، ويعد

الاستماع بعناية لما يقوله المفحوص، والحساسية لردود أفعاله أثناء المقابلة من الأساليب التى تستهدف التحقق من دقة الاستجابة.

أنواع الأسئلة فى المقابلة الشخصية:

يتضمن إعداد المقابلات الشخصية إقرار الآتى:

- ماهية الأسئلة التى سيتم توجيهها إلى المفحوص.
- كيفية تتابع الأسئلة.
- التفاصيل المستهدف الحصول عليها من كل سؤال.
- الوقت اللازم لإجراء المقابلة.
- كيفية إلقاء (توجيه) الأسئلة.

ومن المعروف أن الأسئلة التى تتضمنها المقابلة عبارة عن وسائل قياس تستهدف قياس ظاهرة ما، وأن نوعية استجابات المفحوصين فى المقابلات الشخصية تتأثر بتلك الأسئلة.

ويوجد ستة أنوع (الأنماط) رئيسية من الأسئلة يمكن أن يسألها الناس، بمعنى أنه يمكن سؤال أى من هذه الأسئلة بالنسبة لأى موضوع من الموضوعات، وهذه الأنواع (الأنماط) كالتالى:

١ - **أسئلة السلوك** Behavior Questions

وهى عبارة عن أسئلة توجه لمعرفة ما فعله، أو يفعله الشخص، وتستهدف تلك الأسئلة استدعاء ووصف خبرات وسلوك وأحداث وأنشطة يرغب القائم بالمقابلة فى استرجاعها مرة أخرى.

٢ - أسئلة الاعتقاد (الرأى) Belief (opinion) Questions

وهى أسئلة توجه بغرض استظهار الفهم والمعلومات وعمليات التفسير والشرح عند الشخص، حيث تدل الإجابة عن تلك الأسئلة على فكر الناس عن بعض الظواهر والأحداث والقضايا، كما تدل الإجابة عليها على

اهتمامات ورغبات وطموحات وقيم الناس. وتحمل تلك الأسئلة الشخص المفحوص على التفكير العقلانى عند اتخاذ القرار.

٣ - أسئلة الشعور (الإحساس) Feeling Questions

وهى أسئلة توجه لفهم الاستجابات الانفعالية للناس فى إطار خبراتهم وأفكارهم انطلاقا من افتراض التلقائية spontaneity فيما يتعلق بمنشأ (مصدر) الاستجابات الانفعالية. فالشعور (الإحساس) يحدث فى داخل الفرد، وتحدث الاستجابات الانفعالية وفقا للأحداث التى تقع حولنا أو تقع لنا.

وعندما توجه أسئلة الشعور (الإحساس) فإنه يجب على القائم بالمقابلة (الباحث) interviewer أن ينظر إلى استجابات المفحوص ومحاولة وصف هذه الاستجابات من حيث كونها تعبر عن الشعور: بالسعادة، القلق، الكراهية، الخوف، الثقة، الأمان وغيرها من مظاهر الاستجابات الانفعالية.

ومن الملاحظات الجديرة بالاهتمام أن بعض القائمين بالمقابلة يخلطون بين الإحساس (الشعور) feeling والرأى opinion لكونهم لا يعرفون كيف يميزون بينهما. ولعل من أهم الأساليب التى يمكن استخدامها للتمييز بين الاثنين يكمن فى طريقة توجيه السؤال، حيث يتم ذلك كالتالى:

- عندما نريد أن نفهم الاستجابات الانفعالية للمفحوص، فإنه يصبح من الضرورى أن نسأله عن : شعوره، كالتالى:

« ما هو شعورك بالنسبة لـ . . . . ؟»

«كيف تشعر بـ . . . . ؟»

- وعندما نريد أن نفهم فكر المفحوص عن شىء ما ، فإنه يصبح من الضرورى أن يطلب السؤال بوضوح : رأى المفحوص opinion، أو معتقداته beliefs أو حكمه judgment بالنسبة لموضوع ما ، وليس إحساسه أو شعوره، كالتالى:

« ما هو رأيك بالنسبة لـ . . . ؟»

« ما الذى تعتقده بالنسبة لـ . . . . ؟»

٤ - أسئلة المعرفة Knowledge Questions

وهى أسئلة توجه لاكتشاف ما الذى يمتلكه المفحوص من معلومات حقيقية، انطلاقا من افتراض أنه توجد أشياء محددة يجب أن تكون معروفة، تلك الأشياء ليست آراء opinions، وليست أحاسيس أو شعوراً، ولكنها أشياء يعرفها الشخص وتتضمن حقائق facts عن الشيء أو الحالة التى يسأل عنها.

وقد ينظر البعض إلى المعرفة على أنها مجموعة من المعتقدات beliefs أكثر من كونها مجموعة من الحقائق، ومع ذلك يظل الفيصل فى أسئلة المعرفة يتحدد - أثناء المقابلة - على أساس صحة ما يعرفه المفحوص، لا على أساس ما يعتقد أنه صحيح.

٥ - الأسئلة المتعلقة بالحواس Sensory Questions

وهى أسئلة تستهدف ما يراه أو يسمعه أو يلمسه أو يشمه أو يتذوقه المفحوص، وهى أسئلة تستهدف السماح للقائم بالمقابلة من الاقتراب من حواس المفحوص، كأن يوجه إليه الأسئلة التالية:

«عندما كنت تسير فى فناء المدرسة صباح اليوم، ما الذى شاهدته؟» أو «صف لى ما الذى رأيته عندما كنت تتدرب فى الملعب أمس؟»

٦ - الأسئلة المتعلقة بالجغرافيا البشرية (السكانية)

Demographic Questions

وهى أسئلة تهتم بتحديد خصائص الأفراد الذين يتم مقابلتهم، حيث تساعد الإجابة عن الأسئلة على تحديد موقع المفحوص بالنسبة لغيره من الناس. فالعمر الزمنى والتعليم والوظيفة والدخل والإقامة وغيرها تمثل موضوعات مهمة للأسئلة المتعلقة بالجغرافيا البشرية (السكانية).

مما سبق يتضح أنه يوجد أنواع (أنماط) من الأسئلة عن السلوك، والرأى، والشعور، والمعرفة، والإحساس، والجغرافيا البشرية، وأن تلك الأنواع من الأسئلة يمكن استخدامها فى المقابلات الشخصية، وأنه يمكن استخدام أى نوع من هذه الأنواع، وأن تدريب القائم بالمقابلة على هذه الأسئلة، وتدريبه على الاحتفاظ بها فى ذهنه يمكن أن يساعده على التخطيط الشامل للمقابلة، وعند ترتيب تسلسل الأسئلة التى سيوجهها للمفحوص.

الإطار الزمنى للأسئلة:

يتم توجيه أى من الأسئلة التى سبق الإشارة إليها فى المقابلات الشخصية وفقا لإطار زمنى هو:

- فى الزمن الحاضر Present tense
- فى الزمن الماضى Past tense
- فى المستقبل Future tense

فعلى سبيل المثال يمكن سؤال المفحوص السؤال التالى:

- ما الذى تتدرب عليه الآن ؟ كمثال للزمن الحاضر.
- ما الذى كنت تتدرب عليه أمس؟ كمثال للزمن الماضى.
- ما الذى تخطط لعمله فى العام القادم؟ كمثال للمستقبل.

وعندما يتم الربط بين الإطار الزمنى للأسئلة والأنواع الستة المختلفة لها، فإنه يمكن بناء مصفوفة matrix تتضمن ثمانية عشر نوعاً (نمطاً) مختلفاً من الأسئلة، وفق ما هو موضح بالجدول (٥ - ١) :

جدول (٥ - ١) : مصفوفة اختيارات الاسئلة

أنماط	فى الماضى	فى الحاضر	فى المستقبل
١- أسئلة السلوك			
٢- أسئلة الاعتقاد (الرأى)			
٣- أسئلة الشعور (الإحساس)			
٤- أسئلة المعرفة			
٥- الأسئلة المتعلقة بالحواس			
٦- الأسئلة المتعلقة بالجغرافيا البشرية			

ويلاحظ أنه يمكن توجيه كل أنماط الأسئلة الثمانية عشر بالنسبة لأى موقف أو موضوع أو حدث أو نشاط خاص، وبخاصة عن تكرار إعطاء الأسئلة خلال المقابلة الشخصية لمجموعة مختلفة من المفحوصين، ولكن مع ضرورة الالتزام بتسلسل إعطاء الأسئلة بالنسبة لجميع المفحوصين، ومن ناحية أخرى يتوقف اختيار نمط الأسئلة على ما يراه الباحث (القائم بالمقابلة) مناسباً لجمع المادة العلمية فى بحثه أو دراسته.

تسلسل الأسئلة:

لا توجد قاعدة ثابتة لتنظيم تسلسل الأسئلة فى المقابلة الشخصية، فنادراً ما تلتزم المقابلات الشخصية غير النظامية بتسلسل ثابت للأسئلة، أما بالنسبة للمقابلات النظامية فإنه يجب إعطاء الأسئلة وفق تسلسل موحد بالنسبة لجميع المفحوصين.

٥ - الطرق الكمية للقياس الانفعالى

يقصد بالطرق الكمية للقياس الانفعالى تلك الطرق التى تستخدم مقاييس معيارية مقننة standardized measures لتحويل الآراء والخبرات المختلفة وفقا لفئات من الاستجابة مقدرة سلفا.

وتتميز الطرق الكمية فى مجال القياس الانفعالى بأنها تتضمن العديد من التصميمات التجريبية والأساليب الإحصائية التى تستهدف التحكم فى المتغيرات وتحليل النتائج. فالطرق الكمية تستخدم - فى العادة - وسائل قياس تتمتع بالموضوعية والثبات والصدق، مثال ذلك القياسات المعملية والاستبيانات (الاستخبارات) والاختبارات النفسية الميدانية وغيرها.

وتستخدم المقاييس المختلفة بعض الطرق للحصول على البيانات الكمية بالنسبة للمجال الانفعالى، ومن أهم هذه المقاييس ما يلى:

- مقاييس ليكرت Likert scales
- مقاييس تمايز معانى الألفاظ semantic differential
- مقاييس (التدريج) التقدير rating scales
- الاستبيانات (الاستخبارات) questionnaires

أولا : مقاييس ليكرت Likert scales

تعتبر صيغة ليكرت من أكثر الصيغ استخداماً فى مجال قياس الاتجاهات وسمات الشخصية، حيث يطلب من المفحوص أن يعبر عن درجة موافقته agreement عن سؤال اتجاهى attitudinal question خاص. وقد أعد ليكرت هذا الأسلوب كطريقة لقياس الاتجاهات عام ١٩٣٢م، على أساس أن أى مقياس يستخدم طريقة ليكرت يمكن أن يتضمن سلسلة من الفقرات أو الجمل أوالعبارات مثل «أنا أخاف من الأماكن المرتفعة» وبدلاً من أن يعطى المفحوص إجابته بنعم أو لا، يرجع المفحوص إلى خمس بدائل (اختيارات) مقدمة له وهى :

- غير موافق بشدة Strongly Disagree (SD)
- غير موافق Disagree (D)

- محايد (لم أكون رأياً) (U) Ueutral (Undecided)

- أوافق (موافق) Agree (A)

- أوافق بشدة Strongly Agree (SA)

وذلك لكى يختار منها ما يعبر عن درجة موافقته agree أو عدم موافقته disagree على العبارة المقدمة إليه. فطريقة ليكرت تتضمن تقديم مجموعة من الفقرات (العبارات) للمفحوص لكى يعبر عن مدى موافقته أو عدم موافقته على الفقرة أو العبارة وفقا لمقياس خماسى النقاط five -point scale، تخصص فيه النقاط كالتالى:

١ نقطة لـ أوافق بشدة.

٢ نقطة لـ أوافق.

٣ نقاط لـ محايد (لم أكون رأياً).

٤ نقاط لـ غير موافق.

٥ نقاط لـ غير موافق بشدة.

أو كالتالى:

٥ نقاط لـ أوافق بشدة.

٤ نقاط لـ أوافق.

٣ نقاط لـ محايد.

٢ نقاط لـ غير موافق.

١ نقطة لـ غير موافق بشدة.

حيث يتوقف ذلك على الكيفية التى تعبر بها الكلمات عن الفقرة، حيث يظهر ذلك من خلال المثال التالى:

لقياس اتجاهات التلاميذ نحو النظام المدرسى، طبق مقياساً للاتجاهات، وكان من ضمن عبارات المقياس الجملتين التاليتين:

١- المدرسة مضيعة للوقت (غيرموافق بشدة) (غير موافق) (محايد) (أوافق) (أوافق بشدة)

١ ٢ ٣ ٤ ٥

٢- المدرسة تعلم النشء مهارات نافعة للحياة

٥ ٤ ٣ ٢ ١

من الملاحظ أن أعلى تقدير (٥) قد منح للاتجاه المفضل favorable attitude نحو المدرسة، فقد أعطى فى العبارة الأولى للاستجابة (غير موافق بشدة)، بينما أعطى فى العبارة الثانية للاستجابة « أوافق بشدة».

وتقدم مقاييس ليكرت للمفحوص لكى يصف إحساسه أو شعوره نحو الشيئ لعمل دائرة حول الاستجابة (الاختيار) الذى يناسبه. وللتعبير عن الاستجابات على العبارات (الجمل) تعطى الدرجة ٥ للاستجابة الأكثر تفضيلاً (قبولاً)، وتعطى الدرجة ١ للاستجابة الأقل تفضيلا (قبولا)، وتعطى الدرجات ٢, ٣, ٤، للاستجابات الوسيطية (المتداخلة) intervening responses حيث يتم تقدير إحساس أو شعور أو اتجاهات كل مجموعة المفحوصين على العبارة الواحدة عن طريق حساب المتوسط الحسابى لدرجات المجموعة.

ومن الملاحظات المهمة بالنسبة لمقياس ليكرت أنه يظهر دائما على هيئة تدريج خماسى الاستجابة، فلم يحدث منذ أن ظهر هذا المقياس أنه تضمن أقل أو أكثر من خمس استجابات. ويتطلب مقياس ليكرت العناية الشديدة عند إعطاء الدرجة (تصحيح الاختبار) scoring، وبخاصة فيما يتعلق بالعبارات السالبة، حيث يلزم أن يتم طرح الدرجة على الجمل السالبة من العدد الكلى لفئات الاستجابة (خمس فئات) مضافاً إليها واحد (٥ +١)،

وذلك لتجنب الحصـول على الدرجة صفر بالنسبـة لأى سؤال أو جملة. فإذا كانت الدرجة السالبة هى ٥، ومجموعة فئات الاستجابة تساوى ٥، إذاً:

٥ - ٥ = صفر

أما فى حالة إضافة واحد، فتصبح درجة هذه الجملة كالتالى:

(٥ + ١) - ٥ = ٦ - ٥ = ١ بدلا من صفر

وبعد تـصحيح الدرجـة السالبة، تحـسب الدرجة الكـلية للمقـياس، أو الدرجة الكلية لكل مجال فرعى من مجالات المقياس، أو الدرجة الكلية لكل جملة أو عبارة.

وقد تتضـمـن صيغـة ليكـرت أكثـر من خمس فئـات (اخـتيـارات) Categories أو فئتين (اختيارين)، حيـث يفضـل استخـدام فئتين (اختيارين) فقـط بالنسبـة للأطفـال صغار الـسن مثل: نعـم - لا ، أحب - لا أحب، حاضر - غـائب، وغيرها من اخـتيارت الاستـجابة response choices ثنائية البـعد two- dimensional. ومن المـلاحظات الجـديرة بـالاهتـمام فـى هذا الخصوص هو ما ذكـره جيلفورد Guilford ١٩٥٤م من أن معامـلات الثبات تزداد كلمـا زادت اختيارات الاسـتجابة، وأن أكبـر ثبات وجد عندمـا كانت اختيارات الاستجابة سبع اختيارات (١ - ٧ نقاط).

ثانيا : مقاييس تمايز المعانى Semantic differential

هو أسلوب أعده فى الأصل كل من أسجود، وسوسى، وتانينبيوم Osgood, Suci, and Tannenbaum ١٩٥٧م. حيث أصبح هذا الأسلوب معروفا - فيما بعد - كطريقة سريعة تستخدم لقياس الاتجاهات.

ويتكون مقياس تمايز المعانى من سلسلة من الأوصاف وعكسها توضع فى قائمة فى شكل عمودين متضادين، كل عمود على جانب من جانبى الصفحة، وبحيث تكون المسافة بين الجانبين مقسمة إلى سبعة أوضاع تمثل أوضاع الاتجاهات attitude positions السبعة، ويوضع على قمة الصفحة اسم موضوع الاتجاهات، الذى يتم التعبير عنه فى شكل كلمة أو جملة كالتالى:

اللياقة البدنية

ردىء ـــ ـــ ـــ ـــ ـــ ـــ ـــ جيد

صح ـــ ـــ ـــ ـــ ـــ ـــ ـــ خطأ

مناسب ـــ ـــ ـــ ـــ ـــ ـــ ـــ غير مناسب

ويعد تمايز المعانى أداة جيدة لقياس انفعالات الناس الإيجابية والسلبية وشعورهم نحو موضوع الاتجاه attitude object ، حيث تعكس درجة المفحوص - بشكل عام - انطباعاته عن موضوع الاتجاه. ويمتاز مقياس تمايز المعانى بأنه يحد من استخدام اللغة فى تقرير اتجاه المفحوص نحو موضوع معين، كما يستفاد منه فى المواقف التى يظهر فيها الناس ردود فعل انفعالية emotional reactions نحو موضوع أو مسألة معينة، حيث لا يكون هناك اهتمام بآراء الناس وأفكارهم بالنسبة لموضوع الاتجاه.

وبناء على ما سبق أصبح مقياس تمايز المعانى من الوسائل المناسبة للتطبيق على الأطفال صغار السن (٩ سنوات فأكثر) لكونه يحد من استخدام اللغة، وفى عام ١٩٦٥م أعد ديفيستا Divesta مجموعة من الأوصاف ثنائية البعد يمكن استخدامها بالنسبة للأطفال صغار السن وهى:

الاتجاه الإيجابى	الاتجاه السلبى	الاتجاه الإيجابى	الاتجاه السلبى
جيد	ردىء	سريع	بطىء
صديق	عدو	الأول	الأخير
جميل	قبيح	شجاع	جبان
صح	خطأ	جديد	قديم
سعيد	حزين	قوى	ضعيف

ويمكن استخدام طريقة تمايز المعانى لقياس الاتجاهات فى المجال الرياضى، وبخاصة فيما يتعلق بالاتجاهات نحو المفاهيم والمعانى المجردة concepts مثل التمرينات أو اللياقة البدنية أو غيرهما والتى تقاس وفقا لثلاثة أبعاد من المعانى أسماها أسجود «مجال معانى المفاهيم semantic space هذه الأبعاد أو العوامل تكون كالتالى:

١) بالنسبة لعامل التقويم evaluation، يمكن أن يتضمن الألفاظ التالية:

جيد	good	(فى مقابل) ردىء	bad
سار	pleasant	(فى مقابل) غير سار	unpleasant
ناجح	successful	(فى مقابل) غير ناجح	unsuccessful

٢) وبالنسبة لعامل القدرة (الفاعلية) potency ، فيمكن أن تتضمن الألفاظ التالية:

قوى	strong	(فى مقابل) ضعيف	weak
ذكرى	musculine	(فى مقابل) أنثوى	feminine
ثقيل	heavy	(فى مقابل) خفيف	Light

٣) وبالنسبة لعامل النشاط activity، فيمكن أن يتضمن الألفاظ التالية:

مسترخ relaxed (فى مقابل) متوتر tense

سريع fast (فى مقابل) بطئ slow

ايجابى active (فى مقابل) سلبى passive

نشيط busy (فى مقابل) كسول Lazy

ويلاحظ أن الأبعاد أو العوامل السابقة ثنائية البعد (القطب) bipolar، بمعنى أن لكل بعد أو عامل زوجين من الأوصاف pairs of adjectives يضعهما المفحوص فى الاعتبار. ففى حالة الاستجابة للفظين (سار - غير سار) - مثلا - فإنه يتم عمل تدريج من ٧ نقاط يقوم المفحوص بعمل دائرة حول الرقم الذى يتفق مع رأيه أو إحساسه أو شعوره، مع ملاحظة أن الدرجة الأقل توضع فى الجهة اليمنى عندما تمثل الاتجاه السلبى للصفة أو السمة أو القدرة ، وأن الدرجة الأكبر توضع فى الجهة اليسرى عندما تمثل الاتجاه الإيجابى للصفة أو السمة أو القدرة كالتالى:

	١	٢	٣	٤	٥	٦	٧	
غير سار	-	-	-	-	-	-	-	سار
ردىء	-	-	-	-	-	-	-	جيد

ويتطلب مقياس تمايز المعانى من المفحوص أن يستجيب لنشاط ثنائى البعد (القطب) فالأوصاف ثنائية البعد تمثل المعانى المتضادة للكلمات مثال ذلك: جديد فى مقابل قديم، وجيد فى مقابل ردىء، وسريع فى مقابل بطىء إلى غير ذلك من المعانى المتضادة. حيث يسأل المفحوص لكى يضع علامة (×) على خط واحد من سبعة خطوط تعكس إحساسه أو شعوره بالنسبة للمفهوم الذى يسأل عنه، وتتراوح درجة كل فقرة أو عبارة من (١-٧) درجات، فإذا وضعت الأوصاف الإيجابية على الناحية اليمنى، فإن التدريج يبدأ:

من ٧ إلى ١ كالتالى:

٧	٦	٥	٤	٣	٢	١
سار -	-	-	-	-	-	- غير سار

ويقاس تمايز المفاهيم والمعانى وفقا لثلاثة عوامل (أبعاد) كبيرة هى:

١ - التقويم Evaluation

وهو أكثر العوامل (الأبعاد) شيوعاً، ويشير إلى درجة جودة goodness الصفة أو السمة التى يعبر عنها المفهوم أو المعنى concept. ومن أمثلة الأوصاف ثنائية البعد التى تستخدم بالنسبة للتقويم: جيد - ردئ، جميل - قبيح، سار - غير سار وغيرها.

٢ - المقدرة (الفاعلية والتأثير) Potency

ويشمل القوة التى يمتد إليها المفهوم أو المعنى، ومن الأوصاف ثنائية البعد (القطب) التى تقيس هذا العامل: قوى - ضعيف، ثقيل - خفيف، صلب - رخو وغيرها.

٣ - النشاط Activity

ويتضمن الفعل أو العمل action ويقاس بأوصاف ثنائية البعد تصف الفعل أو العمل مثل: سريع - بطىء، إيجابى - سلبى، سعيد - حزين، وغيرها.

وتستخدم الأوصاف المتضادة (ثنائية البعد) لقياس العديد من المفاهيم والمعانى المختلفة، حيث يلاحظ أن أى مقياس من مقاييس تمايز المعانى يجب أن يتضمن ثلاثة عبارات أو فقرات على الأقل تكون ممثلة لكل عامل من العوامل الثلاثة التى سبق الإشارة إليها. والمثال التالى يبين مقياساً لتمايز المعانى يستخدم لتقويم المعلم (٩ عبارات):

المعلم

١ - ناجح	- - - - - - - ×̲	غير ناجح	(تقويم)
٢ - سلبى	×̲ - - - - - - -	إيجابى	(نشاط)
٣ - وسيم	- - - - - - - ×̲	قبيح	(مقدرة)
٤ - غير أمين	- - - - - - - ×̲	أمين	(تقويم)
٥ - قوى	×̲ - - - - - - -	ضعيف	(مقدرة)
٦ - متزن	×̲ - - - - - - -	عصبى	(نشاط)
٧ - كسول	- - - - - - ×̲ -	نشيط	(نشاط)
٨ - جيد	- - - - - ×̲ - -	ردىء	(تقويم)
٩ - صلب	- - - - ×̲ - - -	متساهل	(مقدرة)

حيث يلاحظ بالنسبة المقياس الآتى:

- سوف يطبق المقياس على طلاب الفصل للتعبير عن إحساسهم نحو المعلم وذلك بوضع علامة (×) على خط واحد من الخطوط السبعة المحصورة بين كل زوجين من الأوصاف.
- تشتمل كل قائمة (عمود) على كل من الأوصاف الإيجابية والسلبية للحد من عامل التخمين، بمعنى عدم وضع كل العبارات الإيجابية أو السلبية فى عمود واحد.
- توجد ثلاث عبارات على الأقل ممثلة لكل عامل من العوامل الثلاثة، حيث تمثل العبارات أرقام (١, ٤, ٨) عامل التقويم، وتمثل العبارات أرقام (٣, ٥, ٩) عامل المقدرة، وتمثل عبارات أرقام (٢, ٦, ٧) عامل النشاط.
- أنه تم ترتيب العبارات بشكل عشوائى، بمعنى مزج العبارات الخاصة بالعوامل الثلاثة معا دون الالتزام بإعطائها ترتيب محدد.

- تحسب درجة الاستجابة على أساس الآتى: تمنح الدرجة (١) للاستجابة الأكثر سلبية (كما فى العبارات أرقام: ١, ٢, ٣، على سبيل المثال ، وتمنح الدرجة (٧) للاستجابة الأكثر إيجابية (كما فى العبارات أرقام: ٤, ٥, ٦) على سبيل المثال أيضا، وتمنح الدرجات من (٢-٦) للاستجابات الوسيطية (كما فى العبارات أرقام : ٧, ٨, ٩)، حيث يلاحظ أن درجة اتجاهات المفحوص نحو المعلم - فى هذا المثال - هى متوسط درجاته فى العبارات التسع، والتى تم حسابها كالتالى:

= ١ + ١ + ١ + ٧ + ٧ + ٧ + ٦ + ٣ + ٤

$$= \frac{٣٧}{٩}$$

= ٤,١١

إذاً متوسط استجابات هذا الطالب بالنسبة للاتجاه نحو المعلم بلغت ٤,١١ درجة .

ثالثا : مقاييس التقدير Rating scales

يعرَّف مقياس التقدير على أنه أسلوب (طريقة) من خلاله يقوم الملاحظ بتصنيف الأشياء أو الأحداث أو الظواهر أو الأشخاص فى فئات (طبقات) تظهر فى شكل سلسلة متصلة من الأرقام، وذلك بهدف معرفة أى أنواع الانطباعات التى يقدمها المفحوصين على التقديرات ratees.

وقد تأخذ مقاييس التقدير نقطتين أو ثلاث، أو خمس، أو سبع، أو تسع نقاط، توضع على خط واحد يصف الفئات فى نهايتى طرفيه، ويصف الفئات الوسطى فى هئية تدريج متصل.

وتستخدم مقاييس التقدير (أو مقاييس التدريج) كوسيلة لجمع المعلومات أثناء القيام بالملاحظة، وذلك عندما يقوم الملاحظ المدرب تدريبا جيداً بتقدير الأشياء أو الأحداث أو الأشخاص على المقياس المعطى له وفقا لانطباعاته عما

يلاحظه. وقد تكون التقديرات لأحداث ماضية retrospective أو لأحداث متزامنة cuncurrent ، وعندما تكون التقديرات لأحداث ماضية فإنه يلزم تلخيص كل الانطباعات المتجمعة بمعرفة الملاحظ (القائم بالملاحظة) خلال فترة زمنية محددة. أما إذا كانت التقديرات لأحداث متزامنة فإنه يلزم تلخيص الانطباعات التى يتم جمعها كما حدثت أثناء الملاحظة أو المقابلة.

وتتميز مقاييس التقدير (التدريج) بأنها سهلة البناء والاستخدام، إلا أن الصعوبة الرئيسية التى تتكرر كثيرا عند استخدامها هو أن السمات traits التى يتم تقديرها قد لا يتم تحديدها تحديداً دقيقاً مما ينتج عنه وجود معانى مختلفة للسمة وبالتالى تقديرات مختلفة لهذه السمة.

وللتغلب على هذه المشكلة يلزم عند إعداد مقاييس التقدير الأخذ بثلاثة تدابير مهمة هى:

١ - أن يتم تعريف السمة المطلوب تقديرها تعريفاً واضحاً ، وشرحها بإعطاء أمثلة خاصة لها.

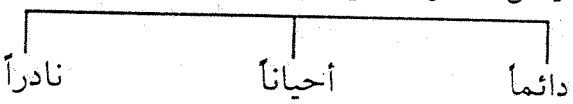
٢ - أن تكون التدريجات المختلفة على المقياس scale - سواء أكانت النقاط أم الفئات - محددة بوضوح، حيث يلاحظ أن التدريجات الخماسية أو السباعية الفئات أو النقاط هى الأكثر استخداما فى تقدير سمات الشخصية والاتجاهات.

٣ - أظهرت الدراسات والبحوث النفسية أنه يمكن تقدير بعض السمات النفسية بدرجة عالية من الثبات، من هذه السمات على سبيل المثال: القيادة، الأمانة، التعاون وغيرها، بينما توجد صعوبة فى تقدير بعض السمات الأخرى بدرجة ثبات مقبولة، من هذه السمات: قوة الأنا ego-strength، الرضا عن العمل job satisfaction، الثبات الانفعالى emotional stability وغيرها، ومن ثم يفضل الابتعاد عن استخدام مقاييس التقدير لتقويم مثل هذه السمات.

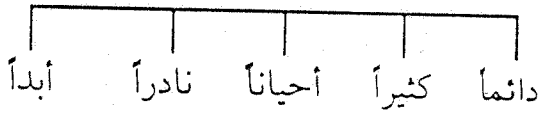
ويلاحظ أنه يوجد لأى مقياس تقدير مكونين components اثنين، الأول ويختص بالأسماء أو الصفات المتعلقة بالفقرة أو السؤال ويعرف بالمتغير المثير stimulus variable، والثانى ويختص باختيارات الاستجابة response options والتى يتم التعبير عنها فى شكل أرقام أو طبقات أو فئات تصنيف.

وقد تأخذ مقايس التقدير الأشكال التالية:

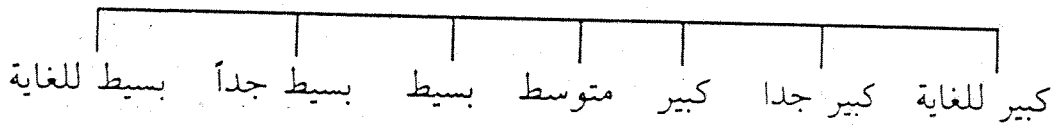
(١) مقياس تقدير ثلاثى النقاط (الاختيارات)



(٢) مقياس تقدير خماسى النقاط (الاختيارات)



(٣) مقياس تقدير سباعى النقاط (الاختيارات)



وتختلف مقاييس التقدير عن مقاييس ليكرت، فى أن مقاييس التقدير تستخدم تعبيرات أو مصطلحات وصفية descriptive terms لكل فئة (اختيار) من فئات الاستجابة، وذلك بدلا من استخدام مجموعة مقننة من درجات الموافقة وعدم الموافقة بالنسبة لكل جملة أو فقرة كما فى مقياس ليكرت.

معنى هذا أن مقاييس التقدير عندما تستخدم التعبيرات أو المصطلحات الوصفية، فإنها تضمن أن كل المفحوصين سوف يجيبون على فئات خاصة من السلوك، هذه الفئات تختلف من مجال لآخر كالتالى:

١ - فى حالة استخدام مقياس تقدير يستهدف معرفة إحساس (شعور) مجموعة من الطلبة نحو معلمهم وكانت الأسئلة كالتالى:

«يهتم بالطلاب. . .»



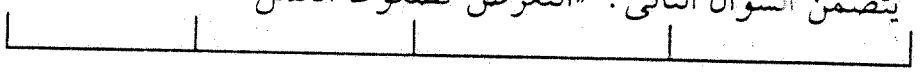
لايهتم متمركز حول ذاته يهتم نادراً يهتم فى معظم الأحيان يهتم دائما

« حماس المعلم. . . . .»



غير متحمس متحمس نادراً متحمس أحياناً متحمس غالبا متحمس بقوة

٢ - وفى حالة مقياس تقدير يختص بتعرض العمال لضغوط العمل، وكان يتضمن السؤال التالى: «التعرض لضغوط العمل»



لا يتعرض لأى ضغوط يتعرض نادراً يتعرض أحيانا يتعرض غالبا يتعرض دائما

أنماط مقاييس التقدير:

يوجد أنماط مختلفة من مقاييس التقدير (التدريج)، ووفقا لما ذكرة جيلفورد Guilford ١٩٥٤ من أنه يمكن وضع جميع مقاييس التقدير فى خمس فئات رئيسية هى:

- مقاييس التقدير الرقمية Numerical rating scales
- المقاييس المعيارية Standard scales
- المقاييس التخطيطية (البيانية) Graphic scales
- مقاييس (تقديرات) النسبة المئوية Percentage scales
- المقاييس قهرية الاختيار Forced choice scales

وفيما يلى شرح موجز لأهم هذه الأنماط:

١ - مقاييس التقدير الرقمية Numerical rating scales

هى عبارة عن مقاييس تعتمد على الترقيم فى تقدير درجة الظاهرة أو السمة التى يقوم بوصفها الملاحظ، حيث يتم تعيين درجات لكل فئة من فئات التصنيف، كأن يوضع تدريج يتكون من خمس نقاط أو سبع نقاط أمام الصفة أو العنصر أو إحدى مكونات السمة، ثم يطلب من القائم بالملاحظة إعطاء درجة للمفحوص بعمل دائرة حول الرقم الذى يتطابق مع مضمون الصفة أو السمة.

وتدل الأرقام فى هذا النمط من أنماط مقاييس التقدير على الدرجة التى تمتد إليها الصفة أو السمة موضوع السؤال، حيث يلاحظ أن سلسلة الأرقام التى يتضمنها المقياس (١، ٢، ٣، ٤، ٥ . . . .) تقابلها صيغ لفظية تصف السمة أو الخاصية عند كل رقم من الأرقام التى يتضمنها المقياس وذلك على أساس أن الدرجة الأكبر تعبر عن القيمة (المقدار) الأكبر لوجود السمة أو الصفة، وأن الدرجة الأصغر تمثل العكس. فإذا فرضنا أنه طلب من أحد اللاعبين تقويم سمات الشخصية للمدرب، وذلك بإعطائه قائمة للشخصية تتضمن مجموعة من العبارات (الفقرات أو الجمل) ليقوم اللاعب بالإجابة عليها برسم دائرة حول الرقم الذى يتفق مع رأيه، فإذا كانت القائمة تتضمن السؤال التالى:

١ - المدرب يتمتع بالقدرة على ضبط النفس. . . ١ ٢ ٣ (٤) ٥ ٦ ٧

ولكى ينجح اللاعب فى هذه المهمة، كان عليه أن يعرف الآتى:

- أن الدرجة (١) تعبر عن أقل مقدار للصفة أو العنصر أو السمة المطلوب تقويمها.
- وأن الدرجة (٧) تعبر عن أعلى مقدار لهذه الصفة أو السمة.
- وأن الدرجة (٤) تعبر عن وجود الصفة أو السمة بمقدار متوسط.
- وأن الدرجة التى تقل عن (٤) تكون أقل من المتوسط، وأن الدرجة التى تزيد عن (٤) تكون أكبر من المتوسط.

- وأنه يجب على اللاعب (وهو القائم بالملاحظة) مراجعة سلوك المدرب، ثم مطابقة هذا السلوك بالرقم المناسب على التدريج، وذلك بوضع دائرة حول هذا الرقم.

وقد تأخذ مقاييس التقدير الرقمية شكل المقاييس العشارية النقاط، حيث يعرف هذا النمط من المقاييس باسم: مقياس التقدير عشارى النقاط Ten-point rating scale ، وهو نمط من المقاييس يستخدم عند القيام بتقويم الأشياء أو البرامج أو الخطط أو الأفراد أو غيرهم، حيث يطلب من القائم بالتقويم (الملاحظ) تقدير موضوع التقويم وفقا لمقياس من ١٠ نقاط.

ويستخدم مقياس التقدير عشارى النقاط بشكل شائع فى البحوث النفسية فى المجال الرياضى، فمثلا قد يطلب من مدرب لكرة السلة بأحد الأندية تقدير مستوى أداء الفريق الأول بالنادى على مقياس عشارى النقاط.

ومن ناحية أخرى فقد يطلب من مدير إحدى المدارس تقويم شخصية معلم التربية الرياضية بمدرسته، وذلك عن طريق ملاحظته للمعلم، ووضع تقديرات له وفق ما هو مبين بالمثال التالى:

سمات الشخصية	التقديرات
١ - يتمتع بالثبات الانفعالى	١ ٢ ٣ ٤ ٥ ٦ ٧ ٨ ٩ ١٠
٢ - يتقبل النقد ...	١ ٢ ٣ ٤ ٥ ٦ ٧ ٨ ٩ ١٠
٣ - واثق من نفسه ...	١ ٢ ٣ ٤ ٥ ٦ ٧ ٨ ٩ ١٠
٤ - يتمتع بالقدرة على ضبط النفس.	١ ٢ ٣ ٤ ٥ ٦ ٧ ٨ ٩ ١٠
٥ - متعقل ومتزن فى تصرفاته ...	١ ٢ ٣ ٤ ٥ ٦ ٧ ٨ ٩ ١٠
٦ - متوافق ومتسق فى تعاملاته مع الطلاب.	١ ٢ ٣ ٤ ٥ ٦ ٧ ٨ ٩ ١٠

ولعل من أهم دوافع استخدام مقياس التقدير عشارى النقاط هو أنه يعطى تقديرات لها مؤشرات صدق وثبات تفوق الطرق الأخرى، وفى هذا

الخصوص ينبه كل من ستيفنز Stevens ١٩٦٦م وباردوسى Parducci ١٩٨٢م إلى أن المقاييس عشارية النقاط تتأثر بخصائص مجموعة الأشياء أو الأفراد الذين يتم تقويمهم، فإذا قام أحد المدربين - على سبيل المثال - بتقدير مهارات ٢٠ لاعباً دولياً متميزاً جداً فى كرة السلة، فإنه قد يواجه مشكلة التمييز فيما بين هؤلاء اللاعبين، هذا من ناحية، ومن ناحية أخرى فقد وجد أن اللاعب الذى يحصل على الدرجة (٦) بالنسبة للاعبين المتميزين، يمكن أن يحصل على الدرجة (٩) عندما يتم تقويم مستواه إذا كان ضمن مجموعة أخرى من لاعبى كرة السلة من ذوى المهارات الضعيفة.

وللتغلب على هذه المشكلة يقترح كابلان وارنست Kaplan and Ernst ١٩٨٠م أنه يلزم أن نشرح لمدير المدرسة والمدرب الرياضى ما الذى تعنيه الدرجة (١٠) والدرجة (١)، فمثلا بالنسبة لمدرب كرة السلة، فإنه يمكن أن نعرض عليه فيلماً مصوراً يبين له أداء اللاعب الذى يستحق (١٠) درجات، وأداء اللاعب الذى يستحق (١) درجة واحدة، حيث يقوم المدرب بتقويم أداء اللاعبين وفقا لمقياس تقدير ينحصر بين هاتين القيمتين.

وقد يتساءل البعض لماذا لا نستخدم مقاييس تقدير أكبر من المقياس عشارية النقاط (كأن تتكون من ١٥ نقطة أو ٢٠ نقطة مثلا)؟ وللرد على مثل هذه التساؤلات لا بد أن نرجع إلى ما أقره سيموندز Symonds ١٩٢٤م واتفق عليه علماء القياس النفسى منذ العام ١٩٢٤م وهو أن أفضل مقاييس التقدير الرقمية هى المقاييس التى تدور حول (٧) نقاط، بينما يرى تشامبنى ومارشال Chompney and Marshall ١٩٣٩م أن أفضل مقاييس التقدير الرقمية هى المقاييس ثلاثية النقاط.

ومع ذلك فقد اتفق العلماء على أن عدد النقاط يتوقف على الدرجة المطلوبة للتمييز فيما بين المفحوصين، ففى مجالات البحث التى تقتضى العناية والاهتمام، فإنه يلزم زيادة عدد الفئات (النقاط) فى المقياس،

والعكس بالنسبة للمجالات الأقل أهمية، فإنه يمكن استخدام مقاييس تقدير ثلاثية النقاط مثلا، حيث يقرر أندرسون Anderson ١٩٧٦م أن المقياس عشارى النقاط يتمتع بقدرة كافية على التمييز بين الأشياء وفقاً لتنوع واسع من المثيرات.

وقد تأخذ مقاييس التقدير الرقمية صيغة الرقم والوصف كالتالى:

الرقم:	١	٢	٣	٤	٥	٦	٧
الوصف:	غير موافق بدرجة كبيرة جداً	غير موافق بدرجة كبيرة	غير موافق	محايد	أوافق	أوافق بدرجة كبيرة	أوافق بدرجة كبيرة جداً

حيث يلاحظ على هذه الصيغة أن عدد الاختيارات (غير موافق بدرجة كبيرة جداً حتى أوافق بدرجة كبيرة جداً) قد أعدت لتكشف الانطباعات والاتجاهات المختلفة نحو الشىء أو الحدث أو الشخص المعين.

وقد يحدث أن تأخذ مقاييس التقدير الرقمية صيغة الوصف بدون أرقام كالتالى:

- أوافق بدرجة كبيرة	(أو)	- دائما	(أو) - كثير جداً
- أوافق		- غالبا	- كثيراً
- محايد		- أحيانا	- أحياناً
- غير موافق		- نادراً	- نادراً
- غير موافق بدرجة كبيرة		- أبداً	- لا يحدث أبداً

ويلاحظ على هذه الصيغة أن الباحث (الملاحظ) هو الذى يحدد الأرقام بالنسبة لكل صيغة من الصيغ بعد إجراء عملية الملاحظة.

ويرى علماء القياس النفسى أن وضع النقاط (١, ٢, ٣، . . . .) أمام الفئات المختلفة (الأوصاف) للمقياس تعكس فى ذهن القائم بالملاحظة أن الأوصاف على المقياس موضوعة فى شكل فئات (تصنيفات) متساوية، مما

يزيد من موضوعية تقدير الدرجة وتسهيل عملية التحليل الإحصائى للنتائج. ومن ثم فإنهم يفضلون استخدام هذه الصيغة على جميع الصيغ الأخرى.

وقد تأخذ مقاييس التقدير الرقمية الصيغة التالية:

الرقم	المعنى
+٣	أوافق بدرجة كبيرة جداً
+٢	أوافق بدرجة كبيرة
+١	أوافق
صفر	محايد
-١	غير موافق
-٢	غير موافق بدرجة كبيرة
-٣	غير موافق بدرجة كبيرة جداً

حيث يلاحظ على هذه الصيغة الآتى:

- أعطيت النقطة (صفر) للصفة (محايد).
- وأعطيت النقاط (+٣، +٢، +١) للأوصاف (الاختيارات) المختلفة فى الاتجاه الموجب (النهاية الموجبة).
- وأعطيت النقاط (-٣، -٢، -١) للأوصاف (الاختيارات) فى الاتجاه السالب (النهاية السالبة للمقياس).

وقد لاحظنا أن جيلفورد Guilford ١٩٥٤م لا يوصى باستخدام هذه الطريقة والتى تقسم المقياس من المنتصف إلى نصفين لسببين مهمين هما:

١ - أنه لا يوافق على منح درجة الصفر المطلق لأى اختيار من الاختيارات.

٢- أن استخدام الأعداد السلبية فى التقديرات قد يوقع القائم بالملاحظة (بعملية التقويم) فى أخطاء، هذا بالإضافة إلى ما يتطلبه استخدام الأرقام السلبية من خلفية رياضية قد لا تتوافر للعديد من القائمين بعملية الملاحظة.

الاتجاه الموجب للعبارة Positive direction

هو الاتجاه الذى يظهر الصفة أو السمة بالشكل الذى يحب معظم الناس أن يصفوا أنفسهم بها، بل ويحرصون دائماً على أن تكون مثل هذه الصفة أو السمة ضمن ما يمتلكونه من صفات أو سمات. ومن أمثلة ذلك: شجاع، جيد، جميل، صحيح، مضحك، سريع، قوى، جديد، ...إلخ. ويمكن أن يعبر الاتجاه الموجب للعبارة عن مصطلحات مثل: الاستجابة الموجبة أو الإحساس الإيجابى أو الرأى الإيجابى.

الاتجاه السالب للعبارة Negative direction

هو الاتجاه الذى يظهر الصفة أو السمة بشكل يرفض معظم الناس أن يصفوا أنفسهم بها، بل ويحرصون دائما أن ينكروا مثل هذه الصفة أو السمة، ومن أمثلة ذلك: جبان، ردىء، خطأ، محزن، بطىء، ضعيف، قديم.

الاتجاه الموجب	الاتجاه السالب
شجاع	جبان
جيد	ردئ
جميل	قبيح
صحيح	خاطئ
مضحك	محزن
سريع	بطىء
قوى	ضعيف

الاتجاه المحايد Neutral direction

هو الاتجاه الذى يظهر الصفة أو السمة بشكل لا يهتم به الناس من حيث أن يصفوا أو لا يصفوا أنفسهم بمثل هذه الصفات أو السمات، وتكون آرائهم واتجاهاتهم نحوها محايدة.

٢ - المقاييس التخطيطية (البيانية) Graphic scales

هى عبارة عن مجموعة من التعبيرات أو الصيغ تصف السمة أو الخاصية بحيث توضع على خط أفقى فى شكل متتابع، حيث يقوم الملاحظ بوضع علامة على التعبير أو الصيغة التى يرى أنها تتفق مع الشىء الذى يقوم بتقويمه.

مثال :

طلب من أحد المعلمين تحديد درجة فعالية أحد التلاميذ فى الدروس التى يقدمها بالنسبة لمادته، وذلك بوضع علامة (×) أسفل أى اختيار من الاختيارات الموجودة على امتداد الخط الأفقى كالتالى :

١ - إلى أى مدى يشترك التلميذ فى المناقشة؟

لا يشترك أبداً	يشترك نادراً	يشترك أحياناً	يشترك كثيراً	يشترك دائماً

٢ - إلى أى مدى ترتبط تعليقات التلميذ بموضوع المناقشة؟

لاترتبط أبداً	ترتبط نادراً	ترتبط أحياناً	ترتبط كثيراً	ترتبط دائماً

يلاحظ من المثال السابق استخدام نفس مجموعة الأوصاف بالنسبة لاختيارات الاستجابة لكل فقرة أو سؤال، حيث يعرف هذا الأسلوب باسم «مقياس البدائل الثابتة Constant - alternatives scale»، وهو مقياس تتباين

فيه فئات الاستجابة من وصف إلى وصف آخر يعرف باسم «البديل المتغير - a changing - alternatives».

وقد أعد الخط الأفقى horizontal Line فى مقياس التقدير البيانى graph-ic rating بغرض وضع التقديرات عليه فى شكل نقاط، يتم التعبير عنها بكلمات مفردة تحدد الفئة أو الاختيار أو الاستجابة، حيث لا توجد ضرورة لاستخدام الأعداد. ولعل من أهم الكلمات المستخدمة فى وصف فئات أو اختيارات الاستجابة ما يأتى: نادراً، أحياناً، كثيراً، إلى غير ذلك من الكلمات والمصطلحات.

٣ - تقدير النسبة المئوية Percentage rating

ويستخدم هذا الأسلوب عندما يريد الباحث إجراء تقدير سريع مع تحقيق أقصى اتساق للنتائج، حيث يتطلب هذا الأسلوب من القائم بالملاحظة وضع المفحوصين فى مجموعات مختلفة وفقا لنسب مئوية خاصة أو وفقا لمئينات percentiles أو ارباعيات quartiles مختلفة كالتالى:

- أعلى ٥٪.
- ثانى أعلى ٥٪.
- أعلى ٢٥٪ بدون الـ ١٠٪ الأعلى.
- الـ ٥٠٪ الأعلى بدون الـ ٢٥٪.
- الـ ٥٠٪ الأدنى.

ويلاحظ أن تقديرات النسب المئوية تستخدم كثيرا من قبل المعلمين عندما يطلب منهم وضع طلابهم فى فئات وفقاً للتحصيل الدراسى ككل.

رابعا: الاستبيانات (الاستخبارات) Questionnaires

الاستبيان (الاستفتاء - الاستقصاء - الاستخبار) عبارة عن وسيلة من وسائل جمع البيانات التى لها صلة وثيقة بمشاعر الأفراد ودوافعهم وعقائدهم

واتجاهاتهم نحو موضوع معين، حيث يستخدم فى الحالات التى يصعب فيها الحصول على البيانات عن طريق الملاحظة كما فى المواقف الماضية والمستقبلية.

والاستبيان عبارة عن تكوين أو شكل يتضمن مجموعة من الأسئلة يتم عرضها على المفحوصين لكى يجيبوا عليها بأسلوب التقرير الذاتى self-report اللفظى وقد ترسل الاستبيانات إلى المفحوصين إما بطريق البريد، أو تنشر على صفحات الجرائد والمجلات، أو عن طريق الإذاعة والتليفزيون ليجيب عليها المفحوصين ثم يقوموا بإرسالها إلى الهيئة المشرفة على البحث، أو تسلم باليد عن طريق باحث مدرب يقوم بعرض الأسئلة على المفحوصين عن طريق المقابلة الشخصية.

ويستخدم الاستبيان البريدى فى المجالات التى يصعب فيها جمع المفحوصين فى مكان واحد، حيث يتولى المفحوص بنفسه الإجابة على أسئلة الاستبيان وتسجيلها بنفسه دون مساعدة من جانب الباحث. فى حين يستخدم الاستبيان غير البريدى فى الحالات التى يمكن فيها جمع المفحوصين فى مكان واحد، حيث يتم توزيع استمارات الاستبيان عليهم عن طريق الباحث نفسه ثم يقوم بجمعها بعد الإجابة عليها.

وتبدأ الأسئلة عادة ببيانات عن الاسم، والجنس، وتاريخ الميلاد، وعدد سنوات الخبرة، والحالة الاجتماعية، ومحل الإقامة، وغيرها. يلى ذلك عرض مجموعة من الأسئلة تستهدف الحصول من المفحوص على معلومات خاصة عن آرائه وميوله وقيمه واهتماماته بالنسبة لموضوعات محددة، وبحيث يقوم المفحوص بتسجيل استجاباته وفقا لمجموعة مختلفة من التقديرات منها: الاستجابة بكلمة واحدة، أو بوضع علامة على الوصف المناسب على مقياس التقدير.

وتستخدم الاستبيانات كثيراً فى مجال التربية الرياضية والترويح فى البحوث المسحية التى تستهدف التعرف على الميول interests والحاجات

needs لإعداد وبناء البرامج التعليمية والترويحية المناسبة. وعادة ما يستخدم الاستبيان فى الحالات التى يحتاج فيها الباحث (القائم بعملية التقويم) إجابات على عدد مختلف من الأسئلة، لذا فهى تصمم بحيث يعبر كل سؤال عن موضوع محدد، وبحيث يمكن جمع درجات الإجابات على كل الأسئلة فى درجة واحدة مفردة.

مميزات الاستبيانات:

١ - تسمح للمفحوص بعدم ذكر اسمه، مما يعطيه الفرصة لكى يعبر عن آرائه ومعتقداته بحرية تامة دون خوف أو حرج.

٢ - تعطى للمفحوص وقتاً كافياً لكى يقرأ الأسئلة بإمعان ويفكر فى الإجابة عليها قبل تسجيل الإجابات.

٣ - يمكن إعطاء الاستبيان لمجموعة كبيرة من الأفراد فى وقت واحد.

٤ - تمثل بالنسبة لكل المفحوصين مواقف قياس على درجة عالية جداً من التناسق والتوحد والانتظام، وذلك لكون كل مفحوص يجيب على نفس الأسئلة، بعكس المقابلة التى قد تختلف فيها الأسئلة من مفحوص لآخر.

٥ - تمد الباحثين ببيانات يمكن تجميعها وتحليلها وعرضها بسهولة أكثر من البيانات التى يتم تجميعها من الاستجابات الشفهية oral responses.

٦ - يخضع الاستبيان لشروط التقنين أكثر من الطرق والوسائل الأخرى، وذلك نتيجة ما يتضمنه من تدقيق بالنسبة للألفاظ، وترتيب الأسئلة، وأسلوب تسجيل الاستجابات، وهو ما يكفل للمفحوصين مواقف متجانسه تزيد من صدقه وثباته كوسيلة من وسائل جمع البيانات.

عيوب الاستبيانات:

١ - تفتقد إلى المرونة الموجودة فى المقابلات الشخصية، ففى المقابلات الشخصية يمكن اكتشاف الأسئلة الصعبة أو غير المفهومة مما يترتب عليه

توضيح المقصود من السؤال أو إعادة صياغته بطريقة أخرى، وهذه الميزة لا تتوافر فى الاستبيانات.

٢ - يميل الناس - عموما - إلى التعبير عن آرائهم وأفكارهم ووجهات نظرهم بطريقة شفهية أكثر منها كتابة.

٣ - تعتمد الاستبيانات على استخدام اللغة المقروءة والمكتوبة، لذا نجدها لا تصلح إلا إذا كان المفحوصون يتمتعون بمستوى معقول من القدرة على القراءة والكتابة.

٤ - قد يؤدى الاستبيان إلى الملل وإهمال الإجابة على بعض الأسئلة إذا كان عددها كبيراً.

٥ - يتوقف صدق وثبات المعلومات المتجمعة من الاستبيانات على وضوح الكلمات ودقة الصياغة والبعد عن المصطلحات الفنية والتعبيرات المعقدة، لذا نجد أن الاستبيانات لا تصلح فى حالة الأسئلة الصعبة أو التى يرتبط بعضها ببعض، أو التى تتطلب قدراً كبيراً من الشرح، وذلك لكون المفحوص يجيب على الأسئلة بدون توجيه من الباحث.

٦ - تفتقد الاستبيانات إلى إمكانية مناقشة المفحوصين والاستفسار عما يقدمونه من آراء شخصية وبخاصة فى الحالات التى لا يذكر فيها المفحوص اسمه.

معجم إنجليزى - عربى للمصطلحات

A

Anecdotal reports تقارير الأحداث السلوكية

Anecdotal records سجلات السير الذاتية

Aggressiveness العدوانية

Attitudinal question سؤال اتجاهى

B

Body image الصورة الذهنية للجسم

Behavior ratings تقديرات السلوك

Bio-dataanalysis تحليل بيانات السجل الشخصى

Belief questions أسئلة الاعتقاد (الرأى)

C

Competitiveness التنافسية

Checklists قوائم مراجعة

Character الخلق

Clinical psychiatry الطب النفسى العلاجى

Correctional psychology علم النفسى الإصلاحى

Closed questions الأسئلة المغلقة

Cuncurrent متزامن

D

Determination التصميم

Depth interview المقابلة العميقة

Demographic questions الأسئلة المتعلقة بالجغرافيا البشرية

E

Emotional Control التحكيم الانفعالى

Emotional stability الثبات الانفعالى

Ego- strength قوة الأنا

F

Followership عضو فى جماعة

Face - to face interview المقابلة وجها لوجه

Formal interview المقابلة النظامية الرسمية

Feeling questions أسئلة الشعور (الإحساس)

G

Guess - who طريقة خمن من

Graphic scales المقاييس التخطيطية (البيانية)

H

Hearing الاستماع

Observation	ملاحظة
On- site observation	الملاحظة فى الموقع
Oriented questions	أسئلة موجهة
Opinion	رأى

P

Perseverance	المثابرة
Peer rating	التقدير بواسطة الرفاق
Peer group	جماعة الرفاق
Popularity	شعبية الفرد
Peer statuas	مكانة الفرد لدى الرفاق
Probes	مجسات
Potency	الفاعلية والتأثير

Q

Qualitative methods	الطرق الوصفية
Quantitative methods	الطرق الكمية

R

Response options	اختيارات الاستجابة
Rapport	الالفة والانسجام
Rating scales	مقاييس التقدير

S

Self-confidence	الثقة بالنفس
Self-Control	ضبط النفسى
Self-awareeness	الوعى بالذات

I

Immaturity	عدم النضج
Interviews	المقابلات الشخصية
Inkblot and picture	طريقة الصورة وبقع الحبر
Intitial responses	استجابات أولية
Intervening responses	استجابات متداخلة

K

Knowledge questions	أسئلة المعرفة

L

Loyalty	الولاء - الإخلاص
Likert scales	مقاييس ليكرت

M

Mental toughness	الصلابة

N

Nonverbal messages	رسائل غير لفظية
Numerical rating	مقاييس التقدير الرقمية
Neutral direction	اتجاه محايد
Negative direction	اتجاه سالب (سلبى)

O

Oral method	التعبير بالكلام

Self- Improvement	التقدم الذاتى
Self-esteem	تقدير الذات
Self- report	التقرير الذاتى
Sociometric	القياس الاجتماعى
Sociogram	مخطط العلاقة الاجتماعية
Social - distance scales	مقاييس المسافة الاجتماعية
Semantic differential	تمايز المعانى
Sociabiity	الاجتماعية
Scoring	إعطاء الدرجة (تصحيح الاختبار)

T

Taste	الذوق
Touch	اللمس
Thematic apparception test (TAT)	اختيار فهم الموضوع
Two dimensional	ثنائية البعد

W

Written method	التعبير بالكتابة
Withdrawal	السلوك الإنسحابى
Word - of mouth	الكلمة المنطوقة

مراجع الفصل الخامس

١ - أحمد حسين اللقانى، ومحمد أمين المفتى (١٩٨٢م): قائمة ملاحظة لتقويم طلاب التربية العملية ، القاهرة، مكتبة الأنجلو المصرية.

٢ - أنور محمد الشرقاوى (٢٠٠٠)م: الدافعية والإنجاز الأكاديمى والمهنى وتقويمه (الجزء الأول)، القاهرة، مكتبة الأنجلو المصرية.

٣ - محمد العربى شمعون (١٩٩٩م) : علمى النفس الرياضى والقياس النفسى، القاهرة ، مركز الكتاب للنشر.

٤ - محمد حسن علاوى (١٩٩٨م): مدخل فى علم النفس الرياضى، القاهرة، مركز الكتاب للنشر.

5 - Henerson, M.E., Morris, L.L. and Gibbon, C.T.F. (1987) : How to measure attitudes, sage publications, London / New Delhi.

6- King, J.A., Morris, L.L. and Gibbon, C.T.F. (1987): How to assess program implementation, sage publications, London/ new Delhi.

7 - Morris , L.L., Gibbon, C.T.F., and Freeman, M.E. (1987) : How to Communicate evaluation Findings, sage publica tions, London / New Delhi.

8 - Morrow, J.R. Jackson, J.A.w. Disch, J.G. and Mood D.P. (1995): Measurement and evaluation in human performance, Human Kinetics.

9 - Patton, M.Q. (1987) : How to use Qualitative method in Evaluation, Sage publications, London / new Delhi.

10 - Singh, A.K. (1986) : Tests Measurements and Research methods in Behavioral Sciences, Tata Mc Graw- Hill publishing comp., New Delhi

الفصل السادس

القياس النفس - حركى

١- طبيعة القياس النفس - حركى

٢ - إجراءات القياس النفسى - حركى

٣ - أنماط الاختبارات النفس - حركية

٤ - خطوات بناء الاختبارات النفس - حركية

- مراجعة معايير الاختبار الجيد
- تحليل الأداء
- مراجعة الأدب السابق
- اختيار أو بناء وحدات الاختبار
- إعداد تعليمات الاختبار
- مراجعة الخبراء
- التطبيق الأولى للاختبار
- تقدير الصدق والثبات والموضوعية
- إعداد المعايير
- إعداد دليل الاختبار

• معجم إنجليزى - عربى للمصطلحات

• مراجع الفصل الثانى

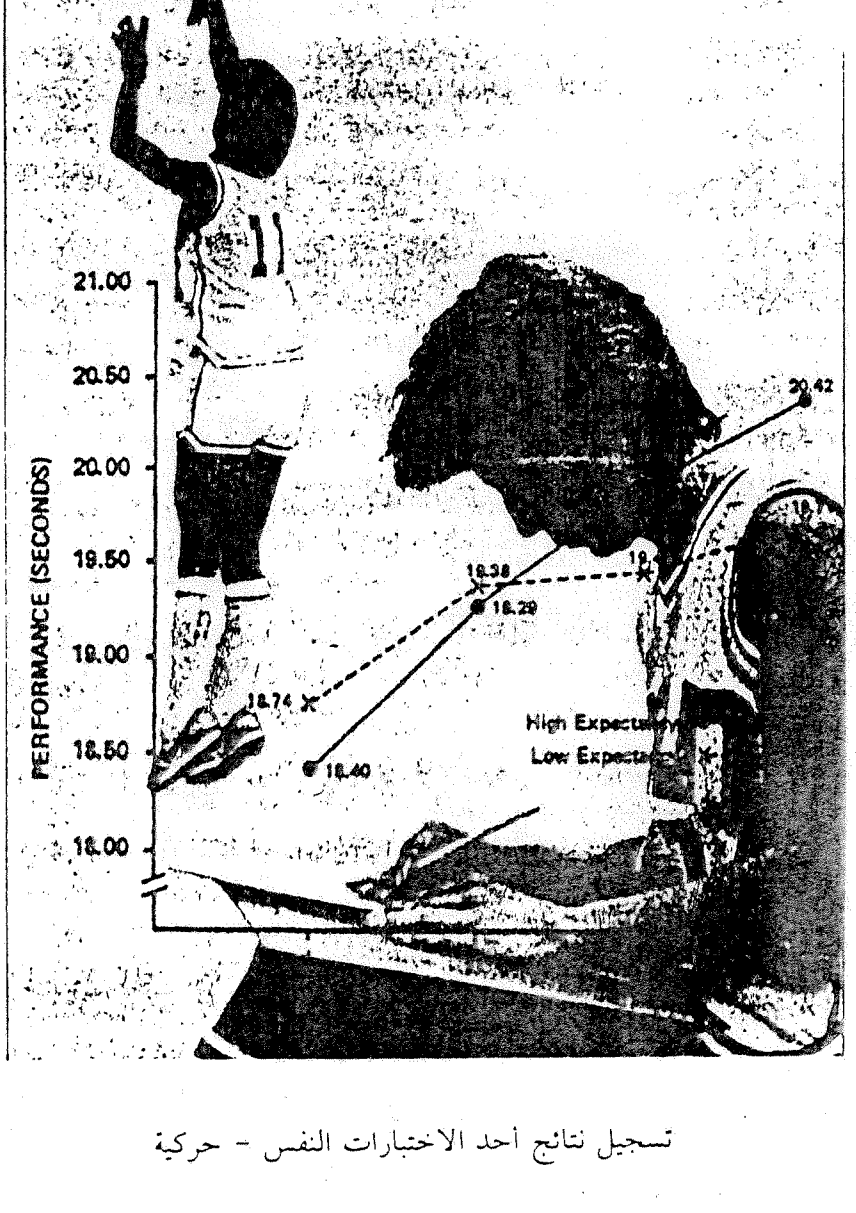
١ - طبيعة القياس النفس - حركى

ذكرنا فى الفصل الأول أن القياس فى التربية البدنية والرياضة يستهدف ثلاثة مجالات رئيسية عريضة domains هى:

- **المجال المعرفى:** ويهتم بالأداء العقلى mental performance
- **المجال الانفعالى:** ويهتم بالاتجاهات والإدراك وسمات الشخصية وغيرها.
- **المجال النفس - حركى:** ويهتم بالأداء البدنى physical performance

وذكرنا أيضا أن جانسيما Jansma ١٩٨٨م قد حدد أربعة مكونات أساسية للمجال النفس - حركى هى: المكون البدنى، والمكون الحركى، ومكون اللياقة، ومكون اللعب.

ومن ناحية أخرى تشكل المهارات فى الألعاب (الرياضات) sport skills والقدرات الحركية motor abilities الموضوعات الأساسية للقياس النفس - الحركى. وفى هذا الخصوص يشير فليشمان Fleishman ١٩٦٤م إلى أن الأساس الحديث للقياس النفس - حركى يستلزم تحديد الفرق بين المهارات skills والقدرات abilities ، ووفقا لما قدمه فليشمان من توضيح فى هذا الشأن يتبين أن المهارة skill عبارة عن سمة (قدرة) تكتسب بالتعلم، هذا الاكتساب يتوقف على ما يمتلكه الشخص من قدرات abilities. ويضيف فليشمان أن القدرات تعزى لعوامل وجينات وراثية أكثر من المهارات، بمعنى أن القدرات فطرية والمهارات مكتسبة. ويرى فليشمان أن المهارات أكثر خصوصية more specific فى الرياضة، بينما القدرات أكثر عمومية more general.



تسجيل نتائج أحد الاختبارات النفس - حركية

وبالنسبة لمصطلح العمومية فى مقابل الخصوصية generality Vs Specificity يوضح باتينللى Battinelli ١٩٨٤م أن اكتساب القدرات الحركية Motor abilities والمهارات الحركية motor skills من خلال عمليات التعلم الحركى motor Learning تعتمد على العوامل العامة كما تعتمد على العوامل الخاصة. ويضيف (باتينللى) بأن المكونات العامة general components للقدرة الحركية motor ability هى: القوة العضلية، والتحمل العضلى، والقدرة العضلية، والسرعة، والتوازن، والمرونة، والرشاقة، والتحمل الدورى التنفسى، وأن هذه المكونات تعمل كعوامل بدنية داعمة للتعلم الحركى، بينما تظهر المهارة كنموذج للعمليات الفسيولوجية العصبية neural physiological التى حدثت خلال التعلم.

الاختبارات النفس - حركية Psychomotor tests

يتفق علماء القياس فى المجال الرياضى على أن الاختبارات النفس - حركية يمكن أن تتضمن نمطين رئيسين من الاختبارات هما:

- اختبارات المهارات فى الألعاب : tests of sport skills
- اختبارات القدرات الحركية: tests of motor abilities

ويتفقون كذلك على أن هذين النمطين من الاختبارات يمكن أن يندرجا تحت مظلة ما يمكن أن نسميه: اختبارات الأداء الحركى tests of motor performance. وبناء على ذلك نود أن نشير إلى أن مصطلح اختبارات الأداء الحركى يعنى الإشارة إلى الاختبارات النفس - حركية، حيث تستهدف هذه الاختبارات قياس الآتى:

- المهارات فى الألعاب sport skills ويشمل الألعاب الفردية والجماعية.
- القدرات الحركية motor abilitres وتشمل: القوة العضلية، التحمل العضلى، القدرة العضلية، السرعة، التوازن، المرونة، الرشاقة، التحمل الدورى التنفسى.

كيفية اختيار الاختبارات النفس - حركية ؟

يستلزم اختيار أو إعداد الاختبارات التى تقيس المهارات فى الألعاب أو القدرات البدنية اتباع إجراءات محددة، هذه الإجراءات تصبح أكثر وضوحاً عندما يتم الإجابة على السؤال التالى:

هل المطلوب : هو اختيار اختبارات مقننة standardized tests أم اختبارات من إعداد المعلم (المدرب) teacher - made tests؟

ذلك أن تحديد ما إذا كانت الاختبارات مقننة أو من إعداد المعلم (المدرب) يتوقف على الاستخدام الخاص بالاختبار، وعلى طبيعة عملية القياس وأهميتها، فعندما تستهدف عملية القياس المقارنة بين مجموعات من الأفراد، فإنه يفضل فى هذه الحالة استخدام أنماط الاختبارات المقننة، أما إذا كانت المعلومات المستهدفة من عملية القياس سوف تستخدم لأغراض بسيطة وسريعة تخدم بصورة مباشرة وفورية تقويم عمليات التعليم والتدريب، فإنه يمكن فى هذه الحالة استخدام اختبارات من إعداد المعلم (المدرب)، حيث يتم إعداد مثل هذه الاختبارات بطريقتين هما:

الطريقة الأولى: وهى أن يقوم المعلم (المدرب) بتعديل الاختبار المقنن ليصبح أكثر بساطة وسهولة ولا يستغرق وقتاً طويلا فى التطبيق.

الطريقة الثانية: وهى أن يقوم المعلم (المدرب) بإعداد اختبارات يرى أنها مناسبة معتمداً فى ذلك على خبراته السابقة فى مجال تخصصه.

وقد قدم الاتحاد الأمريكى للصحة والتربية البدنية والترويح والرقص (AAHPERD) ١٩٨٩م مجموعة من التوجيهات التى يجب وضعها فى الاعتبار عند اختيار الاختبارات النفس - حركية (اختبارات المهارات فى الألعاب واختبارات القدرات الحركية)، حيث تشير هذه التوجيهات إلى أنه يلزم بالنسبة لهذه الاختبارات الآتى:

١ - أن تتمتع هذه الاختبارات بالحد الأدنى المقبول بالنسبة للثبات والصدق.

٢ - أن تكون بسيطة ويسهل تطبيقها.

٣ - أن يكون لها تعليمات واضحة يسهل فهمها.

٤ - ألا تتطلب أجهزة أو أدوات مكلفة الثمن.

٥ - أن تشجع على الأداء الجيد (الصحيح) فى الأنشطة والألعاب المختلفة.

٦ - أن تكون مناسبة من حيث درجة الصعوبة (ليست صعبة بحيث تسبب إحباطاً للمفحوصين، أو سهلة بحيث تفقدهم الحماس والميل إلى التحدى).

٧ - أن تكون جذابة ومشوقة ولها معنى بالنسبة للمؤدى (المفحوص).

٨ - أن تكون خالية من الغموض وألا تتأثر بأية متغيرات دخيلة أو طارئة.

٩ - أن تتمتع بنظام دقيق لتقدير الدرجات.

١٠ - أن يتضمن الاختبار عدداً كافياً من المحاولات للحصول على نتائج موثوق فيها، لأن الاختبارات التى تتطلب الدقة تستلزم الأداء لأكثر من محاولة.

٢ - إجراءات القياس النفس - حركى

يتطلب القياس النفس - حركى اتخاذ بعض التدابير والإجراءات المهمة وذلك بغرض الاقتصاد فى الوقت والمجهود من ناحية، والحصول على نتائج دقيقة من ناحية أخرى. ويمكن تقسيم هذه التدابير والإجراءات إلى الآتى:

- تدابير وإجراءات قبل تطبيق الاختبار pretest.
- تدابير وإجراءات أثناء التطبيق testing.
- تدابير وإجراءات بعد التطبيق post test.

أولاً : تدابير واجراءات ما قبل التطبيق

وهى تتعلق بالتخطيط الأولى لتطبيق الاختبار، وتعد من الأساسيات الملزمة بالنسبة للقائم على إدارة الاختبار، ولعل من أهم هذه التدابير والواجبات ما يلى:

١- الإعداد المسبق للاختبار:

وهو إجراء يستهدف إعداد الاختبار قبل تطبيقه فى التجربة الأساسية للدراسة أو البحث، حيث يتضمن هذا الإعداد وصف الاختبار كتابة من حيث الآتى: الغرض من الاختبار، مستوى السن والجنس، والمعاملات العلمية للاختبار، عدد وحدات الاختبار، طريقة الأداء، عدد المحاولات، كيفية تقدير الدرجة.

ومن ناحية أخرى يستهدف الإعداد المسبق للاختبار زيادة الألفة بين الاختبار والمسئول الأول عن إدارته وذلك من خلال الاستيعاب الكامل لكل التفاصيل والنواحى الفنية للاختبار.

٢ - تحضير الأجهزة والأدوات:

وهى خطوة تأتى بعد تحديد الاختبارات التى وقع عليها الاختيار، حيث يتم تحضير الأجهزة والأدوات اللازمة للتطبيق فى ضوء عدد الاختبارات (الوحدات) التى يتم اختيارها وما يتطلبه تطبيق هذه الاختبارات من أجهزة وأدوات. من هذه الأدوات - على سبيل المثال - طبيعة الأرض (أسفلت - رمل - نجيل - باركية)، الحوائط walls، أشرطه القياس، ساعات إيقاف، الرايات (العلامات)، الكرات الطبية، الأثقال، الحبال، أقلام وورق، ألواح توضع عليها الأوراق للكتابة clipboords وغيرها. ويفضل عمل قائمة مراجعة checklist بكل المتطلبات اللازمة لتطبيق الاختبارات.

ومن ناحية أخرى يلزم تحديد المساحات الفضاء المطلوبة، فعلى سبيل المثال تحتاج وحدة اختبار «رمى الكرة الناعمة softball لأطول مسافة ممكنة»

مساحات كبيرة من الأرض قد لا تقل عن ملعب قانونى لكرة القدم، وقد يحتاج اختبار الجرى - المشى ٨٠٠ متر مضماراً لألعاب القوى، ومن ثم فإنه فى حالة عدم توافر مساحات الأرض المناسبة للتطبيق، فإنه يمكن فى هذه الحالة اتخاذ **واحد من أمرين هما:**

الأول: تخطيط مساحة الأرض المتاحة بأسلوب يمكننا من تطبيق الاختبار، كأن يتم تنفيذ اختبار الجرى - المشى - ٨٠٠ متر بعمل خطين المسافة بينهما ٥٠ مترا بحيث يتم الجرى الارتدادى بين الخطين ١٦ مرة (٥٠×١٦ = ٨٠٠ متر).

الثانى: فى حالة صعوبة تنفيذ الحل الأول يستبدل الاختبار باختبار آخر مناسب.

وفى جميع الحالات يجب أن يتوافر فى الأدوات والأجهزة ومساحات الأرض الفضاء عوامل الأمن والسلامة، كأن تكون الأرض خالية من الحفر التى يصعب رؤيتها، أو العوائق التى يمكن أن تتسبب فى الإصابة عند الجرى لكونها غير مرئية، أو لكونها ليست على بعد مسافة مناسبة من مجال تطبيق الاختبار مما ينتج عنه الاصطدام بها وبخاصة فى حالة اختبارات العدو واختبارات الجرى الارتدادى وغيرها.

ويستخلص مما سبق أن تحضير الأدوات والأجهزة والإمكانات اللازمة لتطبيق الاختبارات يستهدف الإجابة على الأسئلة التالية:

- ما هى الأدوات والأجهزة والإمكانات اللازمة للتطبيق؟
- هل الأدوات والأجهزة اللازمة للتطبيق متاحة فى البيئة المحلية ؟
- هل الأدوات والأجهزة اللازمة للتطبيق فى تناول يد الباحث ؟

٣ - تحديد مكان تطبيق الاختبارات:

يتم اختيار مكان أو موقع تطبيق الاختبارات (مدرسة أو ناد) على أساس ثلاثة معايير رئيسية هى:

- الإمكانات المتوفرة بالمكان وبخاصة مساحات الأرض الفضاء والصالات المغطاة والأدوات والأجهزة التى يتطلبها التطبيق.
- مدى توافر الشروط الصحية بالمكان مثل: الإضاءة والتهوية ودورات المياه والحمامات والأدشاش، وموقع المكان بالنسبة لمصادر التلوث والضوضاء.
- أن يكون المكان فى موقع مناسب يسهل الوصول إليه بالنسبة لجميع المفحوصين. وفى جميع الحالات يجب تحديد الموقع وتجهيزه من حيث تخطيط الأرض وتحديد المحطات، وتوفير الأدوات والأجهزة قبل عملية التطبيق وليس أثناءها، وإن كانت هناك وجهة نظر ترى أنه يفضل القيام بهذه الإجراءات أثناء عملية القياس، وذلك لضمان استمرار وجود الأدوات والأجهزة فى أماكنها وعدم العبث بها.

ومن الشروط الواجب توافرها بالنسبة لأماكن تطبيق الاختبارات ما يأتى:

- أن تكون الأرض مستوية ونظيفة وخالية من أية عوائق أو حفر أو مواقع تسبب خطراً على المفحوصين.
- أن تكون أرضيات الملاعب ذات أسطح لها معامل احتكاك يسمح للمفحوص بأن يؤدى بأفضل ما يمكن وبخاصة فى حالة الاختبارات التى تتطلب الحركة بالرجلين والجسم مثل اختبارات العدو، والجرى الارتدادى، والوثب العريض وغيرها.

٤ - تخطيط محطات الاختبار:

- يمثل تخطيط محطات الاختبار testing stations واحداً من الواجبات التى يجب أن تؤخذ فى الاعتبار قبل تنفيذ بطاريات الاختبارات التى تتضمن عدداً من الوحدات items، وذلك لسببين مهمين هما:
- أنها تعد واحدة من أهم الإجراءات التنظيمية التى تستهدف توفير الوقت والمجهود.

- أنها تحد من ظاهرة التعب مما قد يمكن المفحوصين من الأداء بأفضل ما تمكنهم قدراتهم.

ويتضمن تخطيط محطات الاختبار تحديد عدد المحطات، وتوزيع الاختبارات على هذه المحطات، ووضع ترتيب لتطبيق هذه الاختبارات على أساس المجهود المبذول والمجموعات العضلية المشتركة فى الأداء. فقد أظهرت التجارب العلمية أن من أهم دواعى استخدام نظام المحطات هو: أن يتم توزيع تطبيق وحدات البطارية على عدد من الأيام المتباعدة وذلك بغرض الحد من تأثير عامل التعب إذا ما تم تطبيق جميع وحدات البطارية (٧ وحدات مثلا) فى يوم واحد.

ويجب أن يتضمن تخطيط المحطات تحديد عددها ومكان (موقع) كل منها، وتحضير الأدوات والأجهزة اللازمة لكل محطة قبل البدء فى تنفيذ عملية القياس. وقد يتضمن هذا الإجراء تزويد المفحوصين بمعلومات مبسطة عن موقع تطبيق عملية القياس، ومواعيد التطبيق، والزى الذى يجب أن يرتديه المفحوص.

٥ - ترتيب تنفيذ وحدات الاختبار:

أشرنا إلى أن عدد المحطات يتحدد دائماً فى ضوء عدد الاختبارات التى يتضمنها برنامج القياس، حيث يتم تحديد الاختبارات التى تؤدى فى كل محطة من هذه المحطات. فإذا كان برنامج القياس يتضمن - على سبيل المثال - ١٨ اختباراً تقيس ستة مكونات حركية رئيسية هى: القوة العضلية، القدرة العضلية، السرعة، الرشاقة، التوافق، التحمل الدورى التنفسى.

فى هذه الحالة يمكن تحديد ستة محطات رئيسية هى:

- المحطة الخاصة بتنفيذ اختبارات القوة العضلية.
- المحطة الخاصة بتنفيذ اختبارات القدرة العضلية.
- المحطة الخاصة بتنفيذ اختبارات السرعة.

- المحطة الخاصة بتنفيذ اختبارات الرشاقة.

- المحطة الخاصة بتنفيذ اختبارات التوافق.

- المحطة الخاصة بتنفيذ اختبارات التحمل الدورى التنفسى.

وقد تم تحديد هذا العدد من المحطات على أساس أن تختص كل محطة بتنفيذ عدد من الاختبارات المتشابهة.

يأتى بعد ذلك تحديد عدد الأيام اللازمة لتنفيذ الاختبارات، فإذا تم توزيع الاختبارات الـ ١٨ على ثلاث أو أربع أو خمس مجموعات، فإنه يمكن تنفيذ كل مجموعة فى يوم واحد. ويتم تحديد عدد الاختبارات التى يمكن تطبيقها فى اليوم الواحد **على أساس الآتى:**

-المجهود الذى يتطلبه الأداء.

- المرحلة السنية (أطفال - راشدين - كبار).

- النوع (بنين - بنات).

- الحالة البدنية والتدريبية للمفحوصين.

ويرى علماء القياس النفسى- حركى أنه يفضل أن يكون عدد الاختبارات التى يتم تطبيقها - على أى مجموعة من المفحوصين - فى اليوم الواحد هو ٦ اختبارات. ووفقا لهذا المبدأ فإنه يمكن توزيع الاختبارات الـ ١٨ - فى مثالنا الحالى - على ثلاث مجموعات تضم كل مجموعة ستة اختبارات.

كيفية ترتيب أداء الاختبارت داخل كل مجموعة:

يراعى ألا تضم مجموعة الاختبارات الواحدة أكثر من اختبار واحد يقيس نفس المكون (الصفة أو العنصر)، كما يراعى أن يتم ترتيب تنفيذ الاختبارات داخل المجموعة الواحدة على أساس الآتى:

- التدرج فى المجهود المبذول.

- طبيعة العمل العضلى.

- تأثير التعب على أداء الاختبارات اللاحقة.

ويستهدف هذا الإجراء إتاحة الفرصة للمفحوص لكى يؤدى أفضل أداء نتيجة التدريج فى المجهود المبذول. فقد لوحظ أن معظم الاختبارات النفس - حركية تبدأ بقياس التوافق (المهارة)، يليه السرعة، ثم القدرة العضلية، فالرشاقة ، فالقوة العضلية، وأخيراً التحمل الدورى التنفسى.

وبناء على ذلك فإنه يمكن تنفيذ الاختبارات الستة فى مثالنا الحالى كالتالى بالترتيب: اختبار التوافق، اختبار السرعة، اختبار القدرة العضلية، اختبار الرشاقة، اختبار القوة العضلية، وأخيراً اختبار التحمل الدورى التنفسى.

ولقد بدأ الترتيب السابق بقياس التوافق أولاً، يلية قياس السرعة، وذلك بسبب أن الأداء فى هذين المكونين يتأثر بالتعب أسرع منه بالنسبة لبقية المكونات الأخرى.

مما سبق يتضح أن تحديد عدد المحطات، وتوزيع الاختبارات على مجموعات (إذا كان عددها يزيد عن ٦ اختبارات) ووضع ترتيب لتطبيق هذه الاختبارات، يعد من الإجراءات الإدارية المهمة التى تسبق تنفيذ أى مشروع للقياس النفس - حركى، وذلك لأن مثل هذه الإجراءات يسهل انتقال المفحوصين من محطة لأخرى، ويزيد من إحساسهم بأهمية عملية القياس، ويحد من الهرج والإزعاج الذى قد يحدث أثناء عملية القياس.

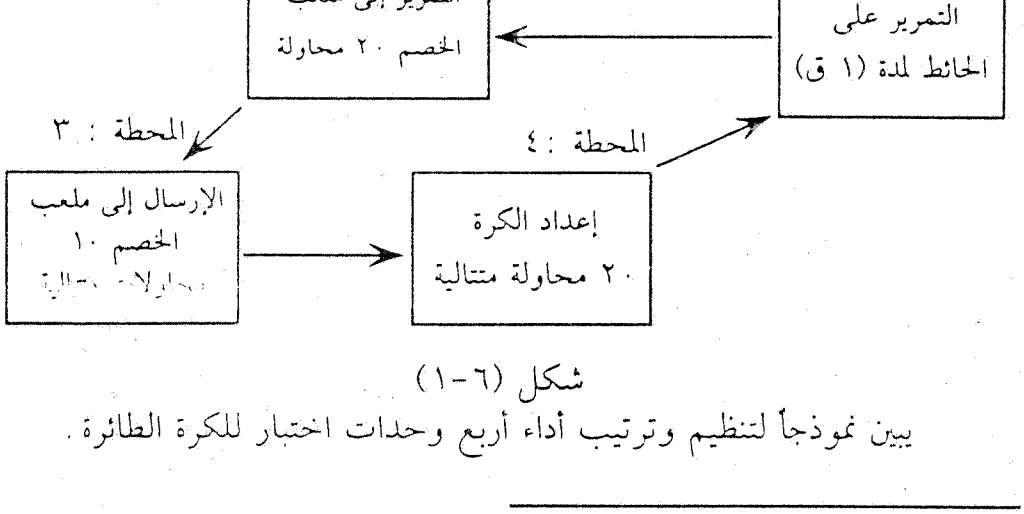
وعندما يتحرك المفحوصون من محطة لأخرى، فإنه يلزم عدم السماح لهم بالتداخل مع مجموعات أخرى، كما يجب إخطار المفحوصين كيف ومتى يتحركون من محطة لأخرى. ولتحقيق ذلك يلزم إعطاء المفحوصين التعليمات بصورة شفوية بالإضافة إلى عمل تخطيط لميدان تطبيق الاختبارات موضحاً به ترتيب المحطات، والاختبارات التى تتضمنها كل محطة.

وفى هذا الصدد يوضح فيلبس وهورناك Phillips and Hornak ١٩٧٩م أنه يلزم ترتيب المحطات والاختبارات بأسلوب يحد من التعب، فيجب أن يتحرك

المفحوص من المحطة الأقل مجهوداً إلى الأكثر، وأن يتحرك من المحطة التى تستخدم مجموعة عضلية معينة إلى المحطة التى تستخدم مجموعة عضلية أخرى.

ويعرض الشكل رقم (٦-١) نموذجاً لتنظيم وتريب أداء أربع وحدات اختبار للأداء المهارى فى الكرة الطائرة. حيث يلاحظ أن هذه الوحدات هى: التمرير على حائط الصد Volleying لمدة دقيقة، الارسال serving إلى ملعب الخصم ١٠ مرات متتالية، التمرير passing إلى ملعب الخصم ٢٠ محاولة متتالية ١٠ محاولات على كل جانب، إعداد الكرة set-up ٢٠ محاولة متتالية.(*)

وقد وضع الاتحاد الأمريكى للصحة والتربية البدنية والترويح ١٩٦٩ ترتيباً لتطبيق الوحدات السابقة، بحيث يبدأ التطبيق بالمحطات ذات الرقم الأقل إلى المحطات ذات الرقم الأكبر مع الدوران فى اتجاه عكس عقارب الساعة، بحيث يبدأ الأداء بالمحطة: ١، يليها المحطة: ٢، ثم المحطة: ٣، وأخيراً المحطة: ٤.



شكل (٦-١)
يبين نموذجاً لتنظيم وترتيب أداء أربع وحدات اختبار للكرة الطائرة.

(*) انظر طرق تطبيق هذه الوحدات بالتفصيل فى كتاب: الاختبارات المهارية والنفسية فى المجال الرياضى - دار الفكر العربى، القاهرة، ١٩٨٧ (ص: ٢٤٠ - ٣٥٢).

٦ - إعداد بطاقات تسجيل البيانات:

تؤثر كيفية تسجيل البيانات على دقة نتائج الاختبارات النفس - حركية، من ثم يصبح من الضرورى العناية بإعداد بطاقات التسجيل المناسبة score cards or score sheets للبيانات بحيث يتم إعداد هذه البطاقات قبل تنفيذ مشروع القياس. وقد يستخدم المسئول الأول عن عملية القياس بطاقات تسجيل معدة سلفاً، وقد يقوم بنفسه بإعداد بطاقات جديدة تتناسب مع متطلبات عملية وخصائص العينة التى يتم تطبيق الاختبارات عليها.

وبشكل عام فإنه يشترط فى بطاقات التسجيل الآتى:

- أن يسهل قراءة وفهم ما تحتويه من بيانات ومعلومات.
- أن تصمم بطريقة تجعلها وسيلة مأمونة للاحتفاظ بالبيانات.
- أن تصمم بطريقة تتضمن ترتيب تطبيق وحدات الاختبار وفقا للتنظيم الفعلى للتطبيق.
- أن تصمم بحيث تشتمل على بعض الفراغات أو الهوامش التى تتيح الفرصة للقائم على تنفيذ الاختبار بتدوين أية معلومات أو ملاحظات تعن له بخصوص المفحوص.
- لتوفير الوقت وزيادة الدقة يتم تخصيص بطاقة لكل مفحوص يكتب عليها اسمه، ويوضع عليها صورته، قبل البدء فى تنفيذ برنامج القياس بفترة كافية.

وتشتمل بطاقات تسجيل البيانات على نمطين من البطاقات هما:

النمط الأول: وهو عبارة عن بطاقة تختص بتسجيل كل البيانات الخاصة بمفحوص واحد فقط (بطاقة تسجيل فردية - individual score sheet)، وتنتقل هذه البطاقة مع المفحوص من محطة إلى أخرى. وقد يعاب على نقل البطاقات مع المفحوصين أنها تتعرض للتلف أو الضياع أو يختلط بعضها مع البعض.

النمط الثانى: وهو عبارة عن بطاقة تضم كل وحدات البطارية وكل أسماء المفحوصين، حيث توزع على كل مسئول بطاقة واحدة من هذه البطاقات. وعندما يدور المفحوصون على المحطات، يقوم المسئول بالنداء على المفحوص، ويتم تسجيل الدرجة له بعد الانتهاء من الأداء.

وتعرف هذه البطاقة بالبطاقة الرئيسية (الأصلية أو الأم) master sheet ، حيث يعاب عليها أن عملية تسجيل البيانات قد تستغرق وقتا طويلا بعد الانتهاء من التطبيق Postest، وإن كان العديد من علماء القياس النفس - حركى ينصحون باستخدام هذا النمط من بطاقات التسجيل شكل رقم (٦-٢).

٧ - اختيار المساعدين وتدريبهم::

قد يحدث نتيجة كثرة عدد الاختبارات وكبر حجم العينة (عدد المفحوصين) أن يلجأ الباحث الرئيسى (المسئول الأول عن برنامج القياس) إلى الاستعانة بعدد من المساعدين assitans الذين يقومون بمعاونته فى تطبيق الاختبارات على المفحوصين.

ويتوقف عدد هؤلاء المساعدين على عدد من العوامل من أهمها ما يأتى:

- طول الاختبار (بمعنى عدد الوحدات التى يتضمنها الاختبار).
- عدد المفحوصين الذى سيتم اختبارهم (حجم العينة).
- عدد محطات الاختبار المستخدمة.
- الزمن الذى تستغرقه كل وحدة من وحدات الاختبار.

ويفضل أن يكون المساعدون ممن لهم خبرة سابقة فى مجال القياس، كما يفضل تدريبهم على كيفية تطبيق الاختبارات التى سوف يقومون بالإشراف على تطبيقها، وبحيث يقوم بهذا التدريب المسئول عن إدارة الاختبارات (الباحث الرئيسى)، وعلى أن يتضمن هذا التدريب الآتى:

شكل (٦-٢) : نموذج لبطاقة تسجيل رئيسية أو أصلية (بطاقة جمعية)

اسم المدرسة (أو النادى) : . . . . . .
الصـف الـدراسـى: . . . . . .
١- الوثب العريض ٢- العدو ٣٠م ٣- التصويب على هدف

الاسم (ثلاثى)	النوع	المحاولات			الدرجة	المحاولات		الدرجة	المحاولات					الدرجة
		١	٢	٣		١	٢		١	٢	٣	٤	٥	
١ -														
٢ -														
٣ -														
٤ -														
٥ -														
٦ -														
إلخ														

- تزويد المساعدين بتعليمات مكتوبة عن كيفية تطبيق الاختبارات.
- تدريب المساعدين على الاختبار بالشرح أولاً، ثم بالتطبيق العملى على عينة من خارج عينة الدراسة الأساسية ثانيا.
- أن يتم تكرار هذا التدريب لأكثر من مرة حتى يتم التأكد من أنهم قد أتقنوا فهم الاختبارات التى كلفوا بالإشراف على تطبيقها.
- يتم توزيع المساعدين على محطات الاختبار، بحيث تكون كل مجموعة من المساعدين مسئولة عن تنفيذ اختبار واحد يوميا، وذلك بشكل ثابت طوال تطبيق البرنامج.

- وفى جميع الحالات فإنه يفضل تزويد جميع المساعدين المشتركين فى برنامج القياس بدليل مكتوب عن الاختبارات موضحا به تعليمات وشروط تنفيذ كل منها، وتوزيع هذه الاختبارات على المحطات، وترتيب أدائها، وأماكن تنفيذها، وذلك من أجل ضمان الحصول على مؤشرات صدق وثبات عالية، وكذا توفير الدقة والموضوعية والسرعة فى تطبيق الاختبارات وتسجيل النتائج.

وقد يكون المساعدون طلابا، أو معلمين، أو معيدين ومدرسين مساعدين بالجامعات، أو من زملاء الباحث الرئيسى، وأيا كانت نوعية المساعدين، فإنه يجب أن يتوافر فيهم شرطان رئيسيان هما:

- أن يكونوا مدربين.
- وغير متحيزين.

وقد تحتاج معظم الاختبارات النفس-حركية لأكثر من فرد واحد للإشراف على تنفيذها، وبناء على ذلك يصبح من الضرورى تحديد المسئوليات الخاصة بتنفيذ كل اختبار، وتوزيع هذه المسئوليات على المساعدين، فمثلا قد تحتاج بعض الاختبارات إلى: ميقاتى لإعلان الزمن، ومسجل للنداء على المفحوصين وتسجيل النتائج، وملاحظ للأداء وتقدير الدرجات.. إلخ، وعليه فإنه يتوقع أن يختلف عدد المساعدين بالنسبة لكل اختبار من الاختبارات وفقا لطبيعة الاختبار. وفى كل الحالات يجب أن يكون هناك مسئولاً رئيسياً عن برنامج القياس، هذا المسئول غالبا ما يكون الباحث الأول (الرئيسى)، أو رئيس فريق البحث.

٨ - توحيد تعليمات الاختبار:

يقصد بتوحيد تعليمات الاختبار هو أن يتم تنميط التعليمات الخاصة بكل اختبار بحيث تكون موحدة ومتسقة بالنسبة لجميع القائمين على التنفيذ، وتكمن أهمية هذا الإجراء فى أنه يزيد من ثبات وموضوعية الدرجات.

وتتضمن معظم الاختبارات النفس - حركية مجموعتين من التعليمات هما:

المجموعة الأولى: وتختص بالتعليمات التى تعطى للقائمين على تنفيذ الاختبار (المساعدين)، وهى تعليمات تشرح الاختبار، وكيفية تطبيقه، وكيفية تقدير الدرجة، وكذا التدريب على تقديم نموذج عملى عن كيفية أداء الاختبار.

المجموعة الثانية: وتختص بالتعليمات التى تعطى للمفحوصين، وهى تعليمات تتعلق بالأداء على الاختبار، وتختص بالأساليب الفنية للأداء، والأخطاء الشائعة، وكيفية تحقيق أفضل النتائج.

ويشترط بالنسبة لكل مجموعة من التعليمات التى سبق الإشارة إليها أن تكون معايَرة بدقة (مقننة)، وأن تقدم مكتوبة لكل فرد مسئول عن تنفيذ الاختبارات.

وبشكل عام يشترط فى التعليمات أن تكون موجزة وأن تكتب بطريقة مختصرة ووافية، وأن تركز على الأداء الصحيح أكثر من الأداء غير الصحيح. وفى جميع الحالات - يجب توحيد التعليمات التى تعطى لكل المفحوصين، بحيث لا يعطى أى مفحوص معلومات إضافية أو مختلفة عن التعليمات التى تعطى لمفحوص آخر، كما يفضل توحيد الطريقة التى تقدم بها التعليمات للمفحوصين.

ثانيا: تدابير واجراءات أثناء التطبيق

وهى تدابير وإجراءات تتخذ عند الوصول إلى مكان تطبيق الاختبارات (صالة أو ملعب)، وتتضمن كافة التجهيزات الضرورية لتنفيذ عملية القياس. وتشمل هذه التدابير والإجراءات الآتى:

١) تنظيم المفحوصين:

يقوم المسئول الأول عن عملية القياس (الباحث الرئيسى) بتحديد عدد المفحوصين المسموح باختبارهم فى اليوم الواحد، حيث يتم تحديد هذا العدد فى ضوء الآتى:

- الإمكانات ومساحات الأرض الفضاء المتاحة.
- عدد المساعدين المشتركين فى تنفيذ عملية القياس.
- عدد الاختبارات (الوحدات) المقرر تطبيقها.

وقد أظهرت التجارب أن تنظيم المجموعة الواحدة غير مناسب بالنسبة لجميع مواقف الاختبار، فإذا كانت الإمكانات متوفرة وعدد المساعدين كافيا، فإنه يمكن فى هذه الحالة استخدام أكثر من مجموعة. أما إذا كانت الإمكانات المتاحة محدودة فإن استخدام المجموعة الواحدة يصبح أمراً ملزماً. وهناك بعض الأنماط الخاصة بتنظيم المجموعات فى القياس النفس - حركى، منها ما يلى:

١ - تجميع كبير العدد:

ويستخدم فى حالة تطبيق اختبارات مثل: الجلوس من الرقود situps، والانبطاح المائل ثنى الذراعين push-ups، الجرى ميل the mile run حيث يقوم نصف المفحوصين بأداء الاختبار، ويقوم النصف الآخر بتقدير الدرجات وإبلاغها للمسئول عن تسجيل النتائج، ففى اختبار الجلوس من الرقود مثلا يقوم نصف المفحوصين بالأداء بينما يقوم النصف الآخر بتثبيت أقدام زملائهم وحساب عدد مرات الأداء.

٢- تجميع صغير العدد:

وهو تنظيم يستهدف وضع المفحوصين فى مجموعات متساوية صغيرة العدد، حيث يتوقف عدد المجموعات على عدد محطات الاختبار، وبحيث تدور كل المجموعات على محطات الاختبار فى نفس الوقت وفقا للترتيب المعد لذلك مسبقا.

٣ - التجميع الحر (غير مقيد بعدد):

وهو تنظيم يستخدم عندما يكون ترتيب أداء الاختبارات غير ملزم، حيث يصبح المفحوص غير مقيد بالبقاء فى أى مجموعة، حيث يستطيع أن يتوجه

لأى محطة من المحطات لأداء الاختبار، ثم يتوجه بعد ذلك لأى محطة أخرى دون التقيد بترتيب محدد للمحطات.

وقد يواجه المسئولين عن إدارة الاختبارات النفس-حركية بعض المشكلات التى تؤثر على المجموعات صغيرة العدد كتغيب بعض المفحوصين، مما يلزم هؤلاء المسئولين بالإسراع فى إعادة تنظيم المفحوصين فى مجموعات جديدة.

مثال تطبيقى :

كلف أحد الباحثين بتطبيق عدد ثمانية عشر اختبارا حركيا على عدد ٢٠٠ طالب من طلاب الجامعات، ولتنظيم القيام بهذه المهمة قام الباحث بالإجراءات التالية:

أولا : توزيع الاختبارات الثمانية عشر على ثلاث مجموعات، كما قام بتحديد ترتيب أداء الاختبارات داخل كل مجموعة (جدول٦-١).

ثانيا : توزيع المفحوصين على عدد مناسب من المجموعات، فإذا كان حجم المجموعة ٢٥ مفحوصاً، يصبح بذلك عدد المجموعات هو ٨ مجموعات تأخذ الأرقام (١, ٢, ٣, ٤, ٥, ٦, ٧, ٨).

ثالثا : يمكن تطبيق الاختبارات الـ ١٨ وفق الترتيب الآتى:

اليوم / المجموعات	(١، ٢، ٣، ٤)	(٥، ٦، ٧، ٨)
السبت ٥ / ١٠	✓	–
الأحد ٦ / ١٠	–	✓
الاثنين ٧/ ١٠	✓	–
الثلاثاء ٨/ ١٠	–	✓
الأربعاء ٩/ ١٠	✓	–
الخميس ١٠/ ١٠	–	✓

جدول رقم (٦-١): توزيع الاختبارات الـ ١٨ على ثلاث مجموعات وترتيب كل منها داخل كل مجموعة

مجموعة الاختبار	اسم الاختبار وترتيب تنفيذه	موعد التنفيذ	أرقام المحطات	تطبيق الاختبارات على المجموعات الأربع الأول (١، ٢، ٣، ٤)
المجموعة الأولى	١ - رمى الكرة الناعمة لأطول مسافة ٢ - العدو (٣٠ مترا) من البدء المنطلق ٣ - دفع الكرة الطبية (٣ كجم) لمسافة ٤ - الجرى المتعرج ٥ - الشد لأعلى ٦ - الجرى - المشى ٨٠٠ متر	اليوم الأول	الأولى الثانية الثالثة الرابعة الخامسة السادسة	(١)، (٢)، (٣)، (٤) (١)، (٢)، (٣) ..... (١)، (٢) ..... (١) ..... ↓ ..... ↓ .....
المجموعة الثانية	١ - التعلق على العقلة لدفع الجسم للأمام ٢ - العدو (٥٠ مترا) من البدء العالى ٣ - الجرى (٤×٩٠ متر) ٤ - الدفع على المقوارين ٥ - بيو ربى (١٠ ث) ٦ - الجرى - المشى (١٠٠٠ متر)	اليوم الثانى	الأولى الثانية الثالثة الرابعة الخامسة السادسة	↓
المجموعة الثالثة	١ - تمرير الكرة على حائط (٢٥ ث) ٢ - العدو (٦ ث) من البدء العالى ٣ - الوثب العريض من الثبات ٤ - جرى الزحزاح (١٠ متر) ٥ - الانبطاح المائل ثنى الذراعين ٦ - الجرى - المشى (١ ميل)	اليوم الثالث	الأولي الثانية الثالثة الرابعة الخامسة السادسة	↓
	١٨ اختباراً	ثلاثة أيام	ست محطات	(٤ مجموعات × ٢٥ = ١٠٠ مفحوص)

يلاحظ من التوزيع السابق ما يلى:

- المجموعات (١, ٢, ٣, ٤) وتضم ١٠٠ مفحوص استغرقت ثلاثة أيام للانتهاء من أداء الاختبارات الـ ١٨ هى أيام: السبت، الاثنين، والأربعاء، حيث يلاحظ أنه يوجد يوم للراحة بين كل مجموعة وأخرى من مجموعات الاختبار الثلاث.
- المجموعات (٥, ٦, ٧, ٨) وتضم ١٠٠ مفحوص استغرقت هى الأخرى ثلاثة أيام هى أيام: الأحد، الثلاثاء، الخميس، حيث يلاحظ أنه يوجد يوم للراحة بين كل مجموعة وأخرى من مجموعات الاختبار.
- يستغرق تطبيق جميع الاختبارات (١٨ اختبارا) على جميع المفحوصين (٢٠٠ مفحوص) أسبوعاً كاملا يبدأ من (يوم السبت ٥ / ١٠ وينتهى يوم الخميس ١٠/ ١٠) حيث يتم اختبار ١٠٠ مفحوص فى اليوم الواحد.
- يؤدى المفحوص الواحد ٦ اختبارات مختلفة فى اليوم، بين كل اختبار والآخر فترة راحة كافية.
- تبدأ مجموعات المفحوصين بالمرور على المحطات الست وفق الترتيب المتفق عليه مسبقا والمعلن على المفحوصين، وعلى أن يكون دوران المجموعات فى عكس اتجاه عقارب الساعة. فبالنسبة لليوم الأول : تبدأ المجموعة (١) بالمحطة الأولى بالنسبة لمجموعة الاختبارات الأولى، وعندما تنتقل إلى المحطة الثانية، تبدأ المجموعة (٢) بالانتقال إلى المحطة الأولى، وهكذا بالنسبة لبقية المجموعات ومحطات الاختبار.
- يقوم المساعدون كل فى مجال تخصصه بشرح الاختبار للمفحوصين قبل الأداء مباشرة، مع تقديم نموذجاً عملياً لهم موضحاً من خلاله طريقة وشروط الأداء والأخطاء الشائعة وكيفية تقدير الدرجة، حيث يلاحظ توحيد التعليمات بالنسبة لجميع المفحوصين، وهناك وجهة نظر ترى أنه يمكن إعطاء المفحوصين تعليمات الاختبارات مكتوبة قبل تطبيقها بفترة كافية.

- السماح للمفحوصين بالإحماء بالنسبة للاختبارات التى تشترط ذلك، وكذا إعطاؤهم الفرصة لكى يؤدوا بأفضل ما تمكنهم قدراتهم عن طريق توفير وسائل الحث الدافعى المناسبة.

٢) مراجعة التجهيزات :

تتطلب مراجعة التجهيزات قيام المسئول عن برنامج القياس بعمل فحص يومى للأدوات والأجهزة والملاعب، حيث يستهدف هذا الفحص والمراجعة الآتى:

- ملائمة الأجهزة والأدوات والملاعب واستمرار صلاحيتها للاستخدام طوال فترة تطبيق الاختبارات.
- عدم استبدال أية أدوات أو أجهزة بأخرى لا تتوافر فيها الشروط العلمية مثل الكرات الطبية، الأثقال وغيرها.
- التأكد من سلامة ودقة بعض أدوات القياس كالساعات وأجهزة قياس الطول والوزن وغيرها.

ثالثا : تدابير واجراءات ما بعد التطبيق:

وهى إجراءات تتم بعد الانتهاء من عمليات القياس، وتشتمل على تسجيل درجات الاختبارات توطئة لتحليلها إحصائيا. وتتوقف طريقة التسجيل على الكيفية التى يتم بها جمع البيانات. كما تتوقف عملية التحليل على الغرض من عملية القياس ذاتها. ومن الملاحظ أن عملية تحليل البيانات قد تستخدم العديد من الأساليب والطرق الإحصائية التى تستهدف تصنيف واختصار وعرض وتحليل البيانات توطئة للقيام بعملية التقويم التى تتضمن الآتى:

١- مقارنة أداء المفحوص بأداء المجموعة التى ينتمى إليها، حيث يستخدم أسلوب التحليل معيارى المرجع النسبى norm-referenced analysis، وهو

أسلوب يستخدم لتفسير الأداء عن طريق مقارنة درجات المفحوص بدرجات غيره من المفحوصين من نفس مجموعته، ولعل من أهم ما يميز هذا الأسلوب أنه يتضمن مستويات نسبية relative standards من الدرجات.

٢- مقارنة أداء المفحوص بمستويات مطلقة absolute standards تتضمن درجات قياسية لتفسير الأداء، حيث يستخدم هذا الأسلوب التحليل معيارى المرجع القياسى criterion- Referenced analysis.

ومن الملاحظات الجديرة بالاهتمام فى هذا الخصوص هو أن تقويم نتائج الاختبارات النفس - حركية تستخدم المعايير المرجعية النسبية Norms لعقد المقارنات فيما بين الأفراد داخل المجموعة الواحدة عن طريق تحديد مستويات نسبية مختارة تستخدم كمحكات للمقارنة، وأن معظم المعايير المتاحة للاستخدام فى المجال الرياضى - فى الوقت الراهن - قد أعدت فى عقود الخمسينيات والستينيات من القرن الماضى، بالإضافة إلى أن الكثير من هذه المعايير قد أعدت على عينات من تلاميذ وطلاب المدارس مما يستلزم توخى الحذر عند استخدام هذه المعايير كمحكات للتقويم.

٣- عرض النتائج على المفحوصين، وهو إجراء يستهدف إمداد المفحوصين بطريقة فورية بالمعلومات الراجعة feedback التى توضح لهم مستويات إنجازهم وجوانب القوة والضعف فى هذا الإنجاز، وبالتالى القيام بعملية التصحيح اللازمة وإعادة توجيه السلوك فى الاتجاه المطلوب. ويعد الإمداد الفورى بالمعلومات من أهم وسائل الحث الدافعى للمفحوصين عند تطبيق برامج القياس النفس - حركى.

وقد يتضمن عرض نتائج الاختبارات النفس - حركية تزويد المفحوصين بالمعايير التى تعكس مستويات الأداء المختلفة كالمتوسطات (المعدلات) والمئينيات وغيرها، بحيث يستطيع المفحوص مقارنة درجاته بهذه المعايير لمعرفة ما إذا كان ممتازاً أو متوسطاً أو ضعيفاً. ويوصى بعض علماء القياس أن ثقة

المفحوص فى أن درجاته التى يتم الحصول عليها ستكون سرية تعد من المواثيق الأخلاقية المهمة بالنسبة لبرامج القياس فى مجال العلوم الإنسانية، وأن على المفحوص الذى يرغب فى معرفة درجاته أن يطلبها مباشرة من المسئول الأول عن إدارة الاختبارات.

٣ - أنماط الاختبارات النفس - حركية

يستخدم فى القياس النفس - حركى نمطان رئيسيان من المقاييس هما:

- الاختبارات الموضوعية Objective tests
- التقديرات الذاتية Subjective ratings

وفيما يلى شرح موجز لكل نمط من هذه الأنماط:

أولاً: الاختبارات الموضوعية

وهى مقاييس تتميز بأنها أقل عرضة للأخطاء errors لكونها تخلو من الأحكام الذاتية Subjective judgments عندما يتم تقدير درجات المفحوص أو تصحيح إجاباته على الاختبار. وينتشر استخدام الاختبارات الموضوعية فى مجال القياس النفس - حركى تحت اسم الاختبارات المقننة standardization tests، حيث تعد الموضوعية objectivity فى تقدير الدرجات من أهم مميزات هذه الاختبارات، كما تتميز هذه الاختبارات بأنها تستخدم وحدات قياس معتمدة ومعايرة بدقة حتى يمكن استخدامها بمعرفة أفراد من بيئات مختلفة دون أن يتأثر تقدير الدرجات فيها سلبا أو إيجابا باتجاهات الشخص الذى يقوم بتطبيقها.

ويمكن تصنيف الاختبارات الموضوعية فى مجال القياس النفس - حركى إلى أربعة فئات رئيسية هى:

١ - الاختبارات الموقوتة Timed tests

٢ - اختبارات المسافة Distance tests

٣ - اختبارات الدقة Accuracy tests

٤ - اختبارات الأداء المتكرر Repetitive - Peformance tests

الاختبارات الموقوتة:

هى نمط من الاختبارات تستخدم الوقت لتقويم الأداء مثل اختبار العدو ٥٠ ياردة، حيث يستخدم الزمن الذى يسجله المفحوص كمؤشر لتقويم السرعة فى العدو. هذا من ناحية، ومن ناحية أخرى تعنى الاختبارات الموقوته بأنه يوجد زمن محدد لأداء الاختبار، وأنه يسمح للمفحوص أن يؤدى الاختبار خلال الزمن المقرر له فقط، فاختبار العدو لمدة ٤ ثوان من البدء العالى four second dash test يستهدف قطع أكبر مسافة ممكنه خلال زمن ٤ ثوان فقط، واختبار بيوربى (١٠ ث) Burpee test لقياس الرشاقة يستهدف تحقيق أكبر عدد من المرات (الإنتاج) خلال الزمن المقرر للاختبار وهو ١٠ ثوان، واختبار الجلوس من الرقود (١ ق) sit - ups test يستهدف الجلوس من الرقود أكبر عدد من المرات خلال الزمن المقرر للاختبار وهو دقيقة واحدة.

وقد يحدث أن تتضمن الاختبارات الموقوتة تكرار أداء نفس المهارة the same skill لفترة زمنية محددة كما فى اختبار التصويب على السلة لمدة نصف دقيقة basket shooting حيث يستهدف الاختبار تسجيل أكبر عدد ممكن من الأهداف فى زمن ٣٠ ثانية، واختبار تمرير الكرة على حائط لمدة (١٥ ث) wall- volley لتسجيل أكبر عدد من التمريرات الصحيحة خلال الزمن المقرر للاختبار وهو (١٥ ث).

وترى سافريت safrit ١٩٩٠م أنه فى حالة الاختبارات التى تتضمن تمرير أو ضرب الهدف فى الهواء كما فى الاختبارين الأخيرين (التصويت على السلة وتمرير الكرة على الحائط)، فإن الزمن المخصص لأداء الاختبار قد يصبح غير مناسب كمقياس للأداء، حيث ترى أن الدقة فى الأداء هى العامل

الأكثر تأثيراً فى النتائج، وتضيف «سافريت» أن الزمن يمكن أن يكون مقياساً مناسباً بالنسبة للاختبارات التى تتم على الأرض وليس فى الهواء.

ومن ناحية أخرى يرى فيلبس وهورناك Phillips and Hornak ١٩٧٩م أن الاختبارات الموقوتة وبخاصة الاختبارات التى تتطلب أداء نفس المهارة قد تدفع المفحوص إلى أن يضحى بالدقة فى الأداء Accuracy فى مقابل السرعة فى تحقيق أكبر عدد ممكن من المرات (الأهداف) خلال الزمن المسموح به للاختبار، فقد يجد المفحوص صعوبة فى استلام الكرة عند أداء اختبار التمرير على الحائط Wall-volley نتيجة تركيزه الشديد على السرعة فى الأداء من ناحية، وانخفاض مستواه فى الدقة فى توجيه الكرة من ناحية أخرى، مما قد يؤثر على صدق الاختبار.

اختبارات المسافة:

تستخدم اختبارات المسافة شريط القياس لقياس مسافة الرمى كما فى اختبار رمى الكرة الناعمة Softball throw لأطول مسافة ممكنة، أو لقياس مسافة الوثب كما فى اختبار الوثب العمودى Vertical Jump، أو الوثب الطويل Long Jump ، أو عند دفع جسم أو هدف لأطول مسافة ممكنة كما فى اختبار دفع الكرة الطبية Medicine ball but أو فى غيرها.

فقد اتفق على أن المسافة Distance التى يقطعها الفرد أو الأداة Object أثناء العمل البدنى تعد من الوسائل المهمة فى القياس النفس - حركى ، ويتمثل ذلك فى المسافات التى يقطعها المفحوص فى الوثب أو القفز، وفى المسافات التى تقطعها الكرات فى الرمى، أو الدفع، أو الركل، أو الضرب، أو التمرير، وتتمثل كذلك فى المسافات التى يقطعها المفحوص خلال فترة زمنية محددة.

فالمسافات تستخدم كمقاييس تدل على الأداء فى العديد من الاختبارات مثل: الوثب العمودى، والوثب العريض وغيرها، كما تستخدم المسافة

كمقياس للأداء بالنسبة لبعض الأنشطة الرياضية مثل السباحة، والعدو والجرى، والوثب الطويل، والقفز بالزانة، والرماية بالسهام وغيرها.

ويرى بومجارتنر وجاكسون Baumgartner and Jakson ١٩٨٢م أن اختبارات المسافة تعطى دائما معاملات ثبات مستقرة نظراً لأنه يمكن قياس المسافة التى تقطعها الأداة بدقة، وأن من أهم ما يميز هذا النوع من الاختبارات هو أنها تتطلب كل من القوة والدقة.

اختبارات الدقة:

وهى نمط من الاختبارات تستخدم أهدافاً targets خاصة لتقويم دقة الرميات Throws كما فى كرة السلة، والضربات strikes كما فى الإرسال فى التنس، وفى الركلات kicks كما فى كرة القدم، وفى غيرها.

وتتطلب اختبارات الدقة استخدام أهداف ترسم فى شكل دوائر أو مربعات أو مستطيلات متداخلة، تخصص درجة لكل منها بحيث تكون الدرجة الأكبر للهدف الأصغر الأقل فى المساحة، وفى معظم الحالات تحدد هذه الأهداف بألوان مميزة على حوائط للصد أو على الأرض.

وتتطلب اختبارات الدقة القيام بعدد كبير نسبيا من المحاولات يفوق عدد المحاولات فى غيرها من الاختبارات وذلك حتى تعطى نتائج أكثر اتساقا، كما تتطلب تحديد الأهداف بألوان واضحة ومختلفة حتى يمكن التمييز بينها بسهولة مما يزيد من موضوعية تقدير الدرجة. ويتأثر الأداء فى هذا النوع من الاختبارات بقوة الإبصار، ومن ثم كان من الضرورى مراعاة ذلك عند تطبيق هذا النمط من الاختبارات.

ويرى بومجارتنر وجاكسون ١٩٨٢م أنه كلما كان نظام تقدير الدرجات فى اختبارات الدقة واسعا وعدد المحاولات (التكرارات) مناسبا، كلما أدى ذلك إلى ارتفاع صدق وثبات الاختبارات، فنظام تقدير الدرجات من (١-١٠ درجات) أفضل من نظام (١-٥ درجات) وأفضل من (٠-٢ درجة).

اختبارات الأداء المتكرر:

تتطلب بعض الاختبارات حساب مرات الأداء الصحيحة التى ينجح فيها المفحوص خلال فترة زمنية محددة، أو عندما يؤدى لعدد محدد من التكرارات أو المحاولات. وفى العادة يمنح المفحوص نقطة واحدة one point فى كل مرة (محاولة) من مرات الأداء الصحيحة، حيث يمثل مجموع النقاط فى جميع المرات (المحاولات) درجة المفحوص.

ففى اختبار داير Dyer للتنس مثلا تحتسب للاعب درجة واحدة لكل كرة صحيحة تلمس حائط الصد فوق خط الشبكة خلال زمن ٣٠ ثانية (*).

وفى اختبار جونسون Johnson لقياس الدقة فى التمرير فى كرة السلة يعطى المفحوص ١٠ محاولات متتالية بدون زمن لتمرير الكرات نحو أهداف متداخلة محددة على حائط حيث يخصص لكل هدف درجة معينة، بحيث يمثل مجموع الدرجات للمحاولات العشرة الدرجة الكلية للاختبار(*).

ثانياً: التقديرات الذاتية

تستخدم التقديرات الذاتية لقياس الأداء النفس - حركى فى مجال الأنشطة الرياضية الفردية كالمصارعة والجودو والسلاح والغطس والجمباز والرقص وغيرها، كما تستخدم لقياس الأداء فى بعض الأنشطة الجماعية ككرة القدم وغيرها، وفى بعض الأحيان تستخدم التقديرات الذاتية كوسائل للحصول على معلومات إضافية عن الأداء فى بعض الأنشطة التى تستخدم فيها الاختبارات الموضوعية، وذلك بهدف تقويم وحدات الأداء المهارى فى اللعبة بعد تحليلها، والحصول على معلومات إضافية عن بعض النواحى الفنية للنشاط أو اللعبة.

(*) انظر كتاب «الاختبارات المهارية والنفسية فى المجال الرياضى» دار الفكر العربى، القاهرة، ١٩٨٧م (ص: ٢٣٧ - ٢٤٢)

(*) انظر المرجع السابق (ص: ٣٥٧ - ٣٦٤).

وتشكل مهارة توجه المهمة Process-oriented skill واحدة من أشكال (صور) form المهارة التى يتم تقويمها، حيث تركز مهارة توجه المهمة على عملية إنجاز المهمة التى سيقوم المفحوص (الفرد الرياضى) بأدائها، وذلك عند الاشتراك فى مواقف الإنجاز، وهذا يعنى محاولة المفحوص أداة المهمة المكلف بها بأفضل ما يمكن.

وتعد المهارات فى الغطس والجمباز أمثلة لما سبق، حيث يتم تقويم اللاعبين على أساس مدى فهمهم لطبيعة المهمة (الأداء) ، أو محاولته الاشتراك فى بعض المهام المعينة التى تظهر قدرته على الإنجاز.

أنماط مقاييس التقدير:

حدد فيردوس verducci ١٩٨٠م نمطين من مقاييس التقدير rating scales هما:

- مقاييس التقدير النسبية relative rating scales
- مقاييس التقدير المطلقة absolute rating scales

١) مقاييس التقدير النسبية:

وفيها يتم مقارنة المفحوص بغيره من المفحوصين فى مظاهر محددة من السلوك، وتعد المقاييس النسبية مناسبة جداً للاستخدام فى حالات المجموعات صغيرة العدد، ولكنها قد تكون مرهقة عندما تكون المجموعات كبيرة العدد.

وتتخذ مقاييس التقدير النسبية العديد من الأشكال (الصيغ) forms ، لعل من أهم هذه الصيغ ما يأتى:

أ - نظام الرتب rank order

ب - المسافات المتساوية ظاهريا equal-appearing intervals

جـ - المقارنة الزوجية paired comparison

أ - نظام الرتب:

هو عبارة عن ترتيب تدريجى تنازلى أو تصاعدى للمفحوصين حسب مستوياتهم، ولعل أكثر الأساليب استخداماً هو الترتيب من الأعلى إلى الأدنى، بمعنى إعطاء المفحوصين رتبا عددية تبدأ من المستويات الأفضل ثم تتدرج إلى المستويات الأقل (الضعيفة)، وذلك عن طريق ملاحظة الأداء الفعلى للمفحوصين سواء أكان الأداء فى مهارة واحدة أو فى اللعبة (المسابقة) ككل.

ويتم وضع نظام الرتب عند تقويم الأداء الذى يدل على المهارة أو القدرة وذلك عن طريق الحكم على هذا الأداء عن طريق الملاحظة، حيث يتم ترتيب المفحوصين ranked بإعطائهم رتباً عددية تبدأ من المستوى الأعلى وتتدرج إلى المستوى الأدنى. ويتم تحديد رتبة المفحوص على أساس مقارنة سلوكه (مستواه فى الأداء) بغيره من المفحوصين من نفس مجموعته.

فإذا افترضنا أنه يوجد لدينا خمسة سباحين (أ،ب، جـ، د، هـ)، ثم طلب منا ترتيب هؤلاء السباحين وفق قدرتهم على القفز الأمامى front dive فى السباحة فإذا جاءت نتيجة الترتيب كالتالى:

السباح	الرتبة	
جـ	١	← أفضل مستوى فى المجموعة
أ	٢	
د	٣	← المستوى المتوسط فى المجموعة
ب	٤	
هـ	٥	← المستوى الضعيف فى المجموعة

فإنه يلاحظ من الترتيب السابق أن أفضل لاعب فى المجموعة هو اللاعب (جـ)،وأن أقل مستوى فى المجموعة هو اللاعب (هـ).

ويمكـن النظر إلى نظـام الرتب rank-order على أنـه نظام يتطـلب من الملاحظ (المحكم) أن يـميز differentiate فيمـا بين الأفراد بأسلـوب تحكمى (قسرى) متجاهلا فـى ذلك الدرجات الفعلية scores التى يمكن أن تمنح لكل فرد، ذلـك أن الملاحظ (المحكـم) يعطى أعـداداً numebers تشيـر إلى رتب ranks تبين المواقع النسبية للأفراد.

ففى نـظام الرتب يدل حـجم الرقم (الرتبـة) المعطى للـمفحـوص على الوضع النسبى له بالمقارنة ببقية زملائـه فى التجربة، مما يعنى أن المسافات فيما بين الرتب المختلفة تـكون غير متساوية، ذلك أن القياس التـرتيبى يقوم على فكرة رئيسيـة هى (أن ملاحظة أو مشاهدة واحدة تكون أكـبر من، أو مساوية إلى، أو أصغـر من ملاحظة أو مشـاهدة أخرى). وبناء على ذلـك يتضح أن نظام الرتب لا يعكس الدرجات الخام الحقيـقية، هذا بالإضافة إلى أنه نظام لا يظهر الفروق الحقـيقية بين درجات الأفراد، فلو افترضنـا أن الدرجات الفعلية للسباحين الخمسة فى القدرة على القفز الأمامى فى السباحة كانت كالتالى:

السباح	الرتبة	الدرجة
جـ	١	٧٥
أ	٢	٦٩
د	٣	٦٧
ب	٤	٦٠
هـ	٥	٥٥

ووفقا لدرجات السـباحين فى القدرة على القفـز الأمامى، فإنه يلاحظ أن اللاعب الأول على المجمـوعة حصل على درجة فعـلية هى ٧٥ درجة، وأن اللاعب الثانى حـصل على ٦٩ درجة، وأن اللاعب الثالـث حصل على ٦٧ درجة ، من هذا نلاحظ الآتى:

أن الفروق الحقيقية بين درجات اللاعبين الأول والثانى هى:

٧٥ - ٦٩ = ٦ درجات

وأن الفروق الحقيقية بين درجات اللاعبين الثانى والثالث هى:

٦٩ - ٦٧ = ٢ درجتان

وبناء على ما سبق يتضح أنه بالرغم من أن الفروق بين رتبتى اللاعبين الأول والثانى (٢-١=١) تساوى الفرق بين رتبتى اللاعبين الثانى والثالث (٣-٢=١)، فإن الفرق بين الدرجات الحقيقية للاعبين الأول والثانى (٦ درجات) لا تساوى الفرق بين الدرجات الحقيقية للاعبين الثانى والثالث (٢ درجتان). مما سبق نستخلص أن نظام الرتب يستهدف تقويم كل الأفراد داخل المجموعة الواحدة بإعطائهم رتباً للتمييز (للتفريق) differentiation فيما بينهم، أكثر من كونه يستهدف حساب الدرجات degrees التى تمنح لكل منهم.

ب - المسافات المتساوية ظاهرياً:

يستخدم هذا النظام فى حالة المجموعة كبيرة العدد، فبدلا من إعطاء رتبة لكل مفحوص، يقسم المفحوصين إلى مجموعات متكافئة السعة، حيث يوضع الأفراد (اللاعبين - المفحوصين) المتشابهين فى الأداء فى نفس المجموعة.

فعن طريق هذا النظام يستطيع الملاحظ (الخبير - المحكم) أن يصنف - على سبيل المثال - ٢٢ لاعبا وفقا لمستويات أدائهم فى الجمباز الأرضى كالتالى:

أولا: يتم تحديد اللاعبين عن طريق إعطائهم الحروف الأبجدية (أ، ب، جـ، د، هـ و. . . إلخ).

ثانيا: يتم تحديد عدد المجموعات وفقا لمقياس تقدير rating scale سباعى يتراوح من ١ إلى ٧ نقاط points كالتالى:

١ ٢ ٣ ٤ ٥ ٦ ٧

حيث تعنى هذه النقاط الآتى:

الدرجة (١) : المستوى الأفضل

الدرجة (٤) : المستوى المتوسط

الدرجة (٧) : المستوى الضعيف جدا

ثالثاً: يتم توزيع المجموعات وفقا لمقياس التقدير السباعى، ثم يوزع اللاعبين على المجموعات وفقا لمستوياتهم كالتالى:

			ش			
		ص	ط	ك		
غ	س	ل	ح	أ		
ى	هـ	أ	و	ن	ز	
جـ	ب	م	ف	د	ى	ت
المجموعة ١	المجموعة ٢	المجموعة ٣	المجموعة ٤	المجموعة ٥	المجموعة ٦	المجموعة ٧

جـ- المقارنة الزوجية:

وتعرف بالمقارنة الثنائية، بمعنى مقارنة كل مفحوص فى المجموعة مع نظير له من نفس المجموعة، فهى طريقة تتضمن وضع مجموعة الأفراد المطلوب اختبارهم (فحصهم) فى شكل أزواج pairs وفقا لمجموعة من المتغيرات تقدم للمحكم (المقوم) ليقوم بإصدار أحكام مقارنة على كل زوج pair.

وتعد طريقة المقارنة الزوجية أكثر إجراءات التقديرات النسبية relative ratings دقة، حيث يقوم المحكم (الخبير او المقوم) بمقارنة كل مفحوص (طالب أو لاعب) بكل مفحوص آخر لتحديد أى زوج من المفحوصين ممتازا فى السمات أو الصفات أو الخصائص التى يتم ملاحظتها.

ويجب على المحكم أن يقوِّم ما إذا كان كل مفحوص أفضل من كل المفحوصين فى المجموعة التى ينتمى إليها أم لا؟، ويتم ترتيب ranking المفحوصين على أساس عدد المرات التى يتم فيها فحص كل منهم فحصاً دقيقا بالمقارنة بزملائه.

مثال :

طلب من أحد الخبراء فى كرة السلة تقدير مستويات أربعة لاعبين (أ،ب،جـ، د) عن طريق المقارنة الزوجية فى القدرة على مسك الكرة. ولتحقيق ذلك قام الخبير بالآتى:

أولا: وضع اللاعبين (أ،ب،جـ،د) فى جدول يعرف باسم: جدول تقدير المقارنات الزوجية paired-comparisons rating جدول رقم (٦-٢):

جدول (٦-٢) : تقدير المقارنات الزوجية

	أ	ب	جـ	د		الرتب
أ	▨	+	−	+	٢	٢
ب	−	▨	−	+	١	٣
جـ	+	+	▨	+	٣	١
د	−	−	−	▨	٠	٤

مجموع الإشارات الموجبة للمقارنات

ثانياً: يبدأ الخبير بمقارنة (أ) مع (ب) على أساس القدرة على مسك الكرة، فإذا كان اللاعب (أ) أفضل من اللاعب (ب) مثلا، فيتم إعطاؤه الإشارة (+) بحيث توضع فى الصف الأفقى المقابل للاعب (أ)، وفى المربع أسفل اللاعب (ب). فوجود الإشارة (+) أسفل اللاعب (ب) وفى الصف الأفقى للاعب (أ) تعنى أن مستوى (أ) أفضل من مستوى (ب)، ومن ثم يجب وضع الإشارة (-) أسفل اللاعب (أ) وفى الصف الأفقى المقابل للاعب (ب). وهكذا تستمر بقية المقارنات حتى يتم مَلء جميع مربعات الجدول (١٢ مربعا).

ثالثاً: يقوم الخبير بعد ذلك بجميع الإشارات الموجبة فى كل صف ووضع الناتج فى العمود قبل الأخير، حيث يشير هذا الناتج إلى مجموعة تقديرات كل لاعب، فاللاعب (أ) له نقطتان، واللاعب (ب) له نقطة واحدة، واللاعب (جـ) له ثلاث نقاط، واللاعب (د) لم يحصل على أية نقاط.

ووفقاً لمجموع النقاط فى العمود قبل الأخير، يتم ترتيب اللاعبين بأن يمنح اللاعب الذى يحصل على أكبر النقاط الرتبة (١)، وهكذا بالنسبة لبقية النقاط والرتب، ثم يوضع الناتج فى العمود الأخير، حيث يلاحظ من هذا العمود أن اللاعب (جـ) كان ترتيبة الأول، واللاعب (د) كان ترتيبه الأخير.

ويلاحظ أن هذا الأسلوب غير عملى عندما يكون عدد الأفراد كبيرا، فمثلا فى حالة وجود ٥٠ لاعبا، فإن الخبير يحتاج إلى القيام بعدد (١٢٢٥) من المقارنات، انطلاقاً من القانون التالى:

$$\text{عدد المقارنات} = \frac{\text{ن (ن-١)}}{٢}$$

$$= \frac{٥٠ \times ٤٩}{٢}$$

= ١٢٢٥ مقارنة

وهو بالطبع عدد كبير جداً يصعب القيام به من الناحية العملية.

مقاييس التقدير المطلقة:

إذا كانت مقاييس التقدير النسبية تستهدف مقارنة أداء الفرد بغيره من أقرانه من نفس مجموعته، فإن مقاييس التقدير المطلقة تستهدف مقارنة أداء الفرد بمعايير حتمية مقدرة سلفا، بمعنى تقويم أداء الفرد (المفحوص) بمقياس ثابت fixed scale له وحدات قياس معايرة بدقة مثل وحدات قياس الطول، أو الوزن ، أو الزمن، أو المسافة، أو غيرها.

وتمتاز مقاييس التقدير المطلقة بأنه يمكن مقارنة مجموعة من الأفراد بمجموعة أخرى، لأن تقويم الأفراد فى المجموعتين يتم باستخدم نفس المعايير أو وحدات القياس، ومن ثم يمكن تطبيق هذه الطريقة على مجموعات كبيرة العدد، نظراً لكونها تستغرق فى التطبيق وقتا أقل من مقاييس التقدير النسبية.

وقد تتطلب مقاييس التقدير المطلقة أن يكون المحكم (القائم بعملية التقويم) مدربا على استخدام المقياس، وأن يكون ملما ومتفهما لدلالة الوحدات المستخدمة لقياس القدرات abilities والسمات traites والخصائص الجسمية والبدنية وغيرها حتى يستطيع تقويم ما يلاحظه بكفاية وموضوعية.

فقد يرغب مدربٌ للسباحة استخدام مقاييس التقدير المطلقة لتقويم أداء أحد اللاعبين، حينئذ يمكنه ملاحظة أداء السباح عندما يؤدى مهارة محددة ثم يقارن نتائج ملاحظاته بمعايير معدة سلفا للحكم عما إذا كان أداء السباح (ضعيفا أو متوسطاً. جيداً، أو ممتازاً) حيث يلاحظ أنه يتم مقارنة أداء السباح بمعايير معدة مسبقا، وليس مقارنته بأداء أقرانه من نفس مجموعته.

وتتخذ مقاييس التقدير المطلقة الصورة (الأشكال) الخاصة التالية:

أ - المقاييس العددية Numerical scales

ب - المقاييس الوصفية Descriptive scales

جـ - المقاييس التخطيطية Graphic scales

د - قوائم المراجعة (الدقيق) Checklists

أ - المقاييس العددية:

المقاييس العددية تعد واحدة من أبسط صور (أشكال) مقاييس التقدير المطلقة، حيث يقوم الملاحظ (المقوم) بعمل دائرة حول العدد (الرقم) الذى يدل على الدرجة التى تعبر عن السلوك الذى لاحظه، معنى هذا أن المقاييس العددية تستخدم الأرقام (الأعداد) numbers لوصف المستويات المختلفة للأداء أو لوصف الأشياء أو الأشخاص أو الأحداث أو الظواهر أو غيرها.

وتشتمل المقاييس العددية على سلسلة من الأعداد تظل ثابتة بالنسبة لكل التقديرات حيث يكون مسموحًا باستخدام أى عدد من النقاط، إلا أنه يفضل استخدام مجموعات الأعداد الفردية، بحيث تأخذ واحداً من التدريجات التالية:

. ١ ٢ (مقياس ثلاثى به درجة الصفر المطلق)

١ ٢ ٣ (مقياس ثلاثى النقاط أو التدريج)

١ ٢ ٣ ٤ ٥ (مقياس خماسى)

١ ٢ ٣ ٤ ٥ ٦ ٧ (مقياس سباعى)

١ ٢ ٣ ٤ ٥ ٦ ٧ ٨ ٩ (مقياس تساعى)

ويلاحظ على التدريجات السابقة ما يلى:

- أنها تتضمن مجموعة فردية من الأعداد (٣، ٥، ٧ أو ٩) موضوعة بشكل متتال من الأصغر إلى الأكبر.

- أنه لا يفضل استخدام أية تدريجات بها الصفر المطلق.

- أنه يمكن استخدام أيا من التدريجات (الخماسية أو السباعية أو التساعية) النقاط، وأنه يفضل ألا يزيد تدريج المقياس لأكثر من ٩ نقاط.

ويلزم اخطـار الملاحظ (المقوِّم) أن الـعدد الأقل فى التـدريج يعنى انـعدام وجود الصفة أو الخاصية أو وجودها بـأقل مقدار ، وأن أكبر عدد يعنى وجود الصفة بأعلى مقدار لها، وأن العدد الأوسط يشـير إلى التقدير المتوسط للصفة أو الخاصية أو السمة. فإذا كان مدى التـدريج الموصى به للمقياس يتراوح من ثلاثة إلى تسعة أعداد، فإنه يمكن وصف هذه المقاييس كالتالى:

* بالنسبة لمقياس التقدير الثلاثى:

١	٢	٣
(الأداء ضعيف)	(متوسط)	(جيد)

* بالنسبة لمقياس التقدير الخماسى:

١	٢	٣	٤	٥
(الأداء ضعيف)	(معقول)	(متوسط)	(جيد)	(ممتاز)

* بالنسبة المقياس التقدير السباعى:

١	٢	٣	٤	٥	٦	٧
(ضعيف جداً)	(أقل من المتوسط)		(متوسط)	(فوق المتوسط)		(قوى جداً)

* بالنسبة لمقياس التقدير التساعى:

١	٢	٣	٤	٥	٦	٧	٨	٩
ضعيف جداً	أقل من المتوسط		(متوسط)		(فوق المتوسط)			(قوى جدا)

ومن ناحـية أخرى فقـد تظهر المقـاييس العـددية فى شكل وصف لفـئات (مستويات) الأداء دون وجود أية أعداد. حيث يقـوم الباحث الرئيسى بتحديد أعداد لكل فئة أو مستوى من مستويات الأداء كالتالى:

٥ درجات (موافق بشدة)

٤ درجات (موافق)

٣ درجات (وسط)

٢ درجتان (غير موافق)

١ درجة واحدة (غير موافق بشدة)

وهناك بعض الاتجاهات التى ترى أنه يمكن إعطاء فئات الاستجابة السابقة الدرجات التالية:

(موافق بشدة)	(موافق)	(وسط)	(غير موافق)	(غير موافق بشدة)
+٢	+١	٠	-١	-٢

وإن كان جيلفورد Guilford ١٩٥٤م يرفض هذا الاتجاه، لأنه لا يقر استخدام الصفر المطلق من ناحية، ولصعوبة التعامل مع الأعداد السالبة من ناحية أخرى.

مثال:

يرجى منك تحديد الدرجة التى تعكس مدى امتلاك اللاعب: أحمد سعيد للمهارة الأساسية فى التنس، وذلك بعمل دائرة حول الرقم الذى يدل على هذه المهارات، مع ملاحظة أن:

الدرجة (٥) تشير إلى المستوى الممتاز

الدرجة (٤) تشير إلى المستوى فوق المتوسط

الدرجة (٣) تشير إلى المستوى المتوسط

الدرجة (٢) تشير إلى المستوى الأقل من المتوسط

والدرجة (١) تشير إلى المستوى الضعيف.

(١) ما مدى فعالية اللاعب بالنسبة لأداء الضربات الأرضية الأمامية فى التنس؟

١	٢	٣	٤	٥

(٢) مـا مـدى فعاليـة اللاعـب بالنسبـة لأداء الضربـات الأرضية الخلفيـة فى التنس؟

١ ٢ ٣ ٤ ٥

(٣) .....

ب- المقاييس الوصفية:

هى عبارة عن مقاييس تقدير تستخدم التعبير بالكلمات لوصف مستويات الأداء بدلاً من الأعداد numbers، وعادة ما يتـم صياغة العبـارات أو الجمل بشكل موجز وواضح يتم وضعهـا بشكل منفصل أسفل المربعات boxes التى تصف سلوك المفحوص (اللاعب أو الطالب) عند نقاط مختلفة على المقياس. وبحيث يـقوم الملاحظ (المقوِّم) بوضـع علامة mark صح (✓) فى المربع الذى يعلو الصياغات التى تصـف أداء المفحوص أفضل وصف، والشكل رقم (٦-٣) يبين مثالا لمقياس تقدير وصفى.

ومن مميزات المقاييـس الوصفية أنه فى حالة وصف الصـفة أو السمة فى شكل مصطلحات قـابلة للملاحظة observable terms ، فإن ذلك يؤدى إلى زيادة موضوعية وثـبات الدرجات، بصورة أفضل من العبارات الـتى يتم فيها وصف الأداء بكلمات مثل (ضعيف، ومقـبول، وجيد، وممتاز) لكون الأخيرة قد تحمل دلالات تختلف من ملاحظ إلى آخر مما قد يؤدى إلى الحصول على معاملات موضوعية وثبات منخفضة.

جـ- المقاييس التخطيطية:

عبارة عن مقاييس تقدير rating scales تقدم وصفا للصفات أو الخصائص فى هيئة خط أفقى Horizontal Line يتراوح فيه مدى التقديرات من الأدنى إلى الأعلى، بحيث يطلب من القائم بالملاحظة وضع العلامة التى تشير إلى تقديره للصفة أو الخاصية التى قام بملاحظتها. ويتضمن الخط الأفقى تحديد أوضاع مجموعة من فئات التصنيف الخاصة التى تمتد على طول الخط شكل رقم (٦-٤).

ويلاحظ من الشكل رقم (٦-٤) أنه يستخدم مجموعة من فئات التصنيف تختلف من صفة لأخرى، لذا يطلق عليه فى بعض الأحيان اسم المقياس المتغير البدائل.

التعليمات:

يرجى منك تحديد الدرجة التى تشير إلى امتلاك اللاعب للمهارة التالية فى التنس عن طريق وضع علامة صح (✓) فى المربع الذى يعلو الكلمات التى تصف أداء اللاعب بدقة:

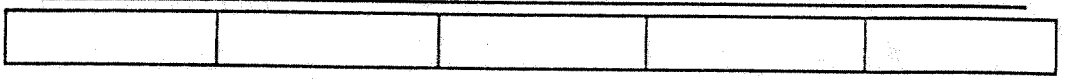
(١) ما مدى فعالية أداء اللاعب بالنسبة لأداء الضربة الأرضية الأمامية فى التنس؟

مع ملاحظة أنه يمكن وصف الضربة الأرضية الأمامية فى التنس كالتالى:

* اللاعب تحرك تحركات خاطئة ٩ مرات أو أكثر * أداء اللاعب يفتقد إلى السيطرة على الكرة لأكثر من ٥٠٪ من الوقت.	* له من ٦-٨ تحركات خاطئة أثناء الأداء. * عدم السيطرة على الكرة من ٣٠٪ إلى ٥٠٪ من الوقت.	* له من ٣ - ٥ تحركات خاطئة أثناء الأداء * عدم السيطرة على الكرة من ١٥٪ إلى ٢٩٪ من الوقت.	* له ١ أو ٢ تحركات خاطئة أثناء الأداء. * عدم السيطرة على الكرة من ٥٪ إلى ١٤٪ من الوقت.	* يؤدى بدون أخطاء، وبشكل جيد من حيث التوقيت، والدقة. * عدم السيطرة على الكرة حوالى ٤٪ أو أقل من الوقت.

(٢) ما مدى فعالية اللاعب بالنسبة لأداء الضربات الأرضية الخلفية فى التنس ؟

مع ملاحظة أنه يمكن وصف الضربة الأرضية الخلفية فى التنس كالتالى:



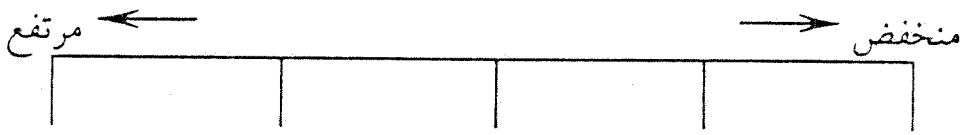
. . . . . إلى آخره من الجمل والأسئلة.

شكل (٦ - ٣) : يبين نموذجاً لأحد المقاييس الوصفية

التعليمات:

يرجى منك تحديد مدى مساهمة التلميذ عبد الرحمن فى المناقشات العلمية التى تدور فى الفصل، وذلك بوضع العلامة (×) على أى مكان عبر الخط الأفقى أسفل أى تقدير تراه يعبر عن رأيك:

(١) إلى أى مدى يشترك التلميذ فى المناقشة؟



لايشترك أبداً يشترك نادراً يشترك أحياناً يشترك كثيراً يشترك دائماً

(٢). . . . . إلخ

شكل (٦ -٤) : مقياس تقدير تخطيطى

د - قوائم المراجعة (التدقيق):

تستخدم قوائم المراجعة عندما يكون المطلوب هو تقويم كل من العمليات process والحصائل (النتائج) outcomes. وتختلف قوائم المراجعة عن مقاييس التقدير الأخرى على أساس نمط الحكم أو القرار type of judgment المطلوب من القائم بالملاحظة، فمقاييس التقدير تتضمن التعبير عن الخاصية أو

السمة على أساس مدى واسع من التقديرات، بينما تنحصر قوائم المراجعة فى إصدار أحكام أو قرارات بسيطة تتمثل فى كلمتين هما: نعم أو لا، بمعنى هل الصفة أو الخاصية أو السمة موجودة أم لا ؟ ، فهى مقاييس ثنائية التقدير dichotmous rating ، بمعنى أنها مقاييس تقدير من نقطتين 2-point rating scale أو بعدين.

وتعتبر قوائم المراجعة من الوسائل المفيدة فى تقدير الأداء وبخاصة عند تقدير الأداء فى البدايات الأولى من التعليم أو التدريب، حيث يتم ملاحظة اللاعبين أو الطلاب أكثر من اختبارهم. كما تستخدم على نطاق واسع فى قياس القوام posture وتحليل المهارات الحركية. فلو افترضنا أن أحد معلمى التربية الرياضية قرر أن يصنف أحد الفصول إلى مجموعتين لدراسة وحدة تعليمية فى كرة القدم، فى هذه الحالة نجد أن هذا المعلم يقوم بوضع التلاميذ المتميزين فى مهارة كرة القدم فى مجموعة واحدة، والتلاميذ غير المتميزين فى مجموعة أخرى، حيث نجده يستخدم قائمة المراجعة للقيام بهذا التصنيف.

ومن ناحية أخرى تستخدم قوائم المراجعة عند ملاحظة السلوك الحركى، حيث يتم تحليل هذا السلوك إلى سلسلة متتالية من المكونات القابلة للملاحظة، حتى يستطيع الملاحظ (المقوِّم) القيام بفحص ما إذا كان المكون مرضيا أو غير مرضى.

وقد تستخدم قائمة المراجعة لتقويم مهارات التمكن master skills فى العديد من المجالات بمعنى أنها تستخدم كاختبار تمكن mastery test تستهدف قياس المعلومات أو المهارات التى اكتسبها الطالب أو اللاعب والتى تظهر مدى تمكنه وكفايته بالنسبة لمجال من المجالات أو بالنسبة لموضوع من الموضوعات شكل (٦-٥).

قائمة مراجعة المهارات الأساسية فى السباحة

(مستوى المبتدئين فى السباحة)

التعليمات :

أمامك مجموعة من المهارات الأساسية الضرورية الواجب وضعها فى الاعتبار حتى يكون الشخص مؤهلاً للسباحة عند مستوى المبتدئين. والمطلوب منك هو وضع دائرة حول نعم أو لا لتوضيح ما إذا كانت المهارة (الصفة - السمة - الخاصية) موجودة فى الشخص أو غير موجودة.

نعم لا ١ - يحبس النفس تحت الماء لمدة ١٠ ثوان.

نعم لا ٢ - يسبح بطريقة الزحف على الصدر لمسافة ٢٠ قدما.

نعم لا ٣ - يسبح بطريقة الزحف على البطن لمسافة ٢٠ ياردة.

نعم لا ٤ - يسبح على الظهر لمسافة ١٠ ياردات.

نعم لا ٥ - يسبح على البطن باستخدام حركات الرجلين فقط لمسافة ٢٠ قدما.

نعم لا ٦- يسبح على الظهر باستخدام حركات الرجلين فقط لمسافة ٢٠ قدما.

نعم لا ٧ - الطفو على الظهر.

نعم لا ٨ - الطفو على البطن.

نعم لا ٩ - الوثب (بالرجلين) فى الماء.

نعم لا ١٠ - القفز (بطريقة الغطس) فى الماء.

شكل (٦-٥) يبين قائمة مراجعة لتقويم مهارات التمكن للمبتدئين فى السباحة

ويلاحظ بالنسبة لقائمة المراجعة شكل (٦-٥) أنه قد تم مراجعة كل مهارة (صفة - خاصية - سمة) من المهارات على أساس مقياس تقدير ثنائى يستهدف الإجابة على سؤال محدد هو :

- هل المهارة موجودة أم غير موجودة؟

ويلاحظ أيضا أن مثل هذا النوع من القوائم يستخدم المعيار المرجعى القياسى criterion-referenced لتقويم النتائج على أساس النقطة الفاصلة بين (نعم ولا).

وقد يحدث أن تتضمن قائمة المراجعة مجموعة من الوظائف أو الواجبات المطلوب ملاحظتها وتقويمها على أساس مقياس تقدير ثنائى هو : موجودة بدرجة مرضية stastisfactory ، أو بدرجة غير مرضية unsatisfactory.

ففى قائمة المراجعة شكل (٦-٦) يلاحظ أنها تتضمن وصف القوام posture أثناء المشى، بحيث يطلب من الملاحظ (المقوم) ملاحظة كل مفحوص أثناء المشى لاختبار حالته القوامية ووضع أى من الإشارتين التاليتين :

(+) إذا كانت الحالة مرضية (أو طبيعية).

(-) إذا كانت الحالة غير مرضية (يوجد انحراف أى غير طبيعية).

على أن تكون الدرجة النهائية للمفحوص هى مجموع الإشارات (+).

واذا كانت قائمة المراجعة شكل (٦-٥) تشتمل فقط على كل مظاهر السلوك الإيجابية (الأفعال أو الأعمال أو الواجبات أو المهارات) المطلوبة للأداء الأفضل، فإنه يلاحظ أن قائمة المراجعة شكل (٦-٦) تشتمل فقط على مظاهر السلوك السلبية التى تعبر عن أخطاء أو عيوب فى الأداء، حيث يطلب من الملاحظ (المقوم) فى الحالتين ملاحظة ماذا كان هذا السلوك يحدث أم لا؟ وبناء على ذلك فإنه يصبح من الضرورى على القائم بالملاحظة التأكد ما إذا كان نمط السلوك المستهدف من الملاحظة يمكن وصفه من حيث الكيف (النوع) وصفاً واضحاً ودقيقاً أم لا ؟، وما إذا كان هذا النمط من السلوك موجوداً أم غائباً بالنسبة لكل مفحوص.

قائمة مراجعة الحالة القوامية لطالبات الجامعات

التعليمات :

* المطلوب ملاحظة كل طالبة أثناء المشى، ثم تقدير الحالة القوامية لها بوضع إشارة (+) أو إشارة (-) فوق الخط أمام كل انحراف تتضمنه عبارات القائمة، مع ملاحظة أن الإشارة (+) تعنى أن الحالة طبيعية، وأن الإشارة (-) تعنى أن الانحراف (التشوه) موجود.

* الدرجة النهائية للطالبة = عدد الإشارات الموجبة (+)

- ١ - سيطرة تحميل وزن الجسم على قدم واحدة.
- ٢ - تأرجح الساقين على الجانبين بدلا من الاتجاه للأمام.
- ٣ - الزيادة فى حركة مفصلى الكعب.
- ٤ - المشى على العقبين.
- ٥ - تقوس الساقين.
- ٦ - اصطكاك الركبتين.
- ٧ - المغالاة فى مرجحة الذراعين.
- ٨ - الميل بالرأس للأمام.
- ٩ - المبالغة فى الارتداد (الحركة) لأعلى.
- ١٠ - تجر قدميها أثناء المشى بدلاً من نقلهما.

شكل (٦-٦) يبين قائمة مراجعة لتقويم حالة القوام أثناء المشى.

٤ - خطوات بناء الاختبارات النفس - حركية

يلجأ العاملون فى المجال الرياضى إلى بناء الاختبارات النفس - حركية فى **الحالات التالية:**

١- عندما تكون الاختبارات المنشورة فى المراجع والدوريات العلمية المتخصصة غير مناسبة للبيئة المحلية، لما تتطلبه من أدوات أو أجهزة أو مساحات أرض غير متاحة.

٢- عندما تتطلب الاختبارات المنشورة استخدام بعض الأدوات والأجهزة الخاصة المعقدة والمكلفة الثمن.

٣- عندما تكون المعلومات المنشورة عن الاختبار غير مستوفاة لكل جوانب وتعليمات الاختبار.

٤- عندما لا تتضمن المعلومات المنشورة ما يشير إحصائيا إلى صدق وثبات الاختبار، وإلى أنواع المحكات criteria التى استخدمت لحساب الصدق، وإلى طبيعة وحجم العينة المرجعية (عينة التقنين) standardization sample.

٥- عندما تكون هناك حاجة إلى تطوير وسائل القياس المتاحة والعمل على ابتكار وسائل وأدوات جديدة للاستفادة منها فى البيئة المحلية.

ويتطلب بناء اختبارات نفس-حركية جديدة، أو تعديل بعض الاختبارات المتاحة، أو انتقاء الاختبارات المناسبة للتطبيق، اتباع خريطة التدفق شكل (٦-٧) والتى تتضمن عشرة خطوات رئيسية لاستكمال هذه العملية، وهذه **الخطوات كالتالى:**

الخطوة الأولى: مراجعة معايير الاختبار الجيد:

وتشمل هذه المراجعة المعايير criteria الواجب مراعاتها بالنسبة للاختبارات النفس - حركية، وهذه المعايير هى:

١- **الصدق** Validity : ويعد أحد أهم الشروط الواجب توافرها فى أدوات القياس، وينظر إلى الصدق على أنه الدرجة التى يقيس بها الاختبار ما وضع من أجله. ويقبل الصدق على أساس معاملات الارتباط التى تشير إلى هذا الصدق.

ويشير كل من ماتيوز Mathews ١٩٦٨م وبارو وآخرون Barrow et.al. ١٩٨٩م وسافريت ١٩٩٠م إلى أن معاملات الارتباط التى تدل على الصدق تتوقف على مدى مناسبة المحك criterion من ناحية، وعلى الغرض من استخدام الاختبار من ناحية أخرى.

ويتوقف قبول معاملات الصدق على نمط (نوع) الصدق المستخدم، فمعاملات صدق التكوين الفرض construct validity يجب ألا تقل عن ٠,٨٠ بينما يمكن قبول معاملات الصدق التنبؤى إذا كانت ٠,٥٠، ويقترح بارو ١٩٧٩م المستويات الموضحة بالجدول (٦-٣) لتقويم معاملات الارتباط التى تدل على الصدق (معامل ارتباط الصدق - validity correlation coefficient).

جدول (٦-٣): مستويات معاملات الارتباط التى تدل على صدق الاختبار.

المعاملات	الصدق
من ٠,٨٥ - ٠,٩٥	ممتاز
من ٠,٨٠ - ٠,٨٤	جيد جداً
من ٠,٧٠ - ٠,٧٩	مقبول
من ٠,٦٠ - ٠,٦٩	موضع شك

٢- **الثبات:** Reliability: وهو يشير إلى مدى الدقة والاتساق الذى يقيس به الاختبار الظاهرة التى وضع من أجلها. ويشير بوسكو وجستافن Bosco and Gustafson ١٩٨٣م إلى أن الثبات المرتفع يعنى انخفاض أخطاء القياس measurement errors وهى أخطاء تتطلب التحكم فى مجموعة من العوامل والمتغيرات مثل: أساليب تطبيق الاختبار، شروط التطبيق، حالة المفحوصين، الأدوات والأجهزة المستخدمة، الشروط البيئية والمناخية وغيرها.

ويقترح باروماك جى ١٩٧٩م المستويات الموضحة بالجدول (٦-٤) لتقويم معاملات الارتباط التى تدل على الثبات (معامل ارتباط الثبات reliability correlation coefficient).

٣- **الموضوعية:** Objectivity: وهى صورة أو شكل من أشكال الثبات لكونها مسألة تتعلق بمقدر الدرجة. فعندما يقوم اثنان أو أكثر بتطبيق نفس الاختبار على نفس المفحوص، ثم يحصلون على نفس الدرجة تقريبا فإن ذلك يعنى ارتفاع موضوعية القياس.

جدول (٦-٤): مستويات معاملات الارتباط التى تدل على ثبات درجات الاختبار.

المعاملات	الثبات
من ٠,٩٥ - ٠,٩٩	ممتاز
من ٠,٩٠ - ٠,٩٤	جيد جدا
من ٠,٨٠ - ٠,٨٩	مقبول
من ٠,٧٠ - ٠,٧٩	ضعيف
من ٠,٦٠ - ٠,٦٩	موضع شك

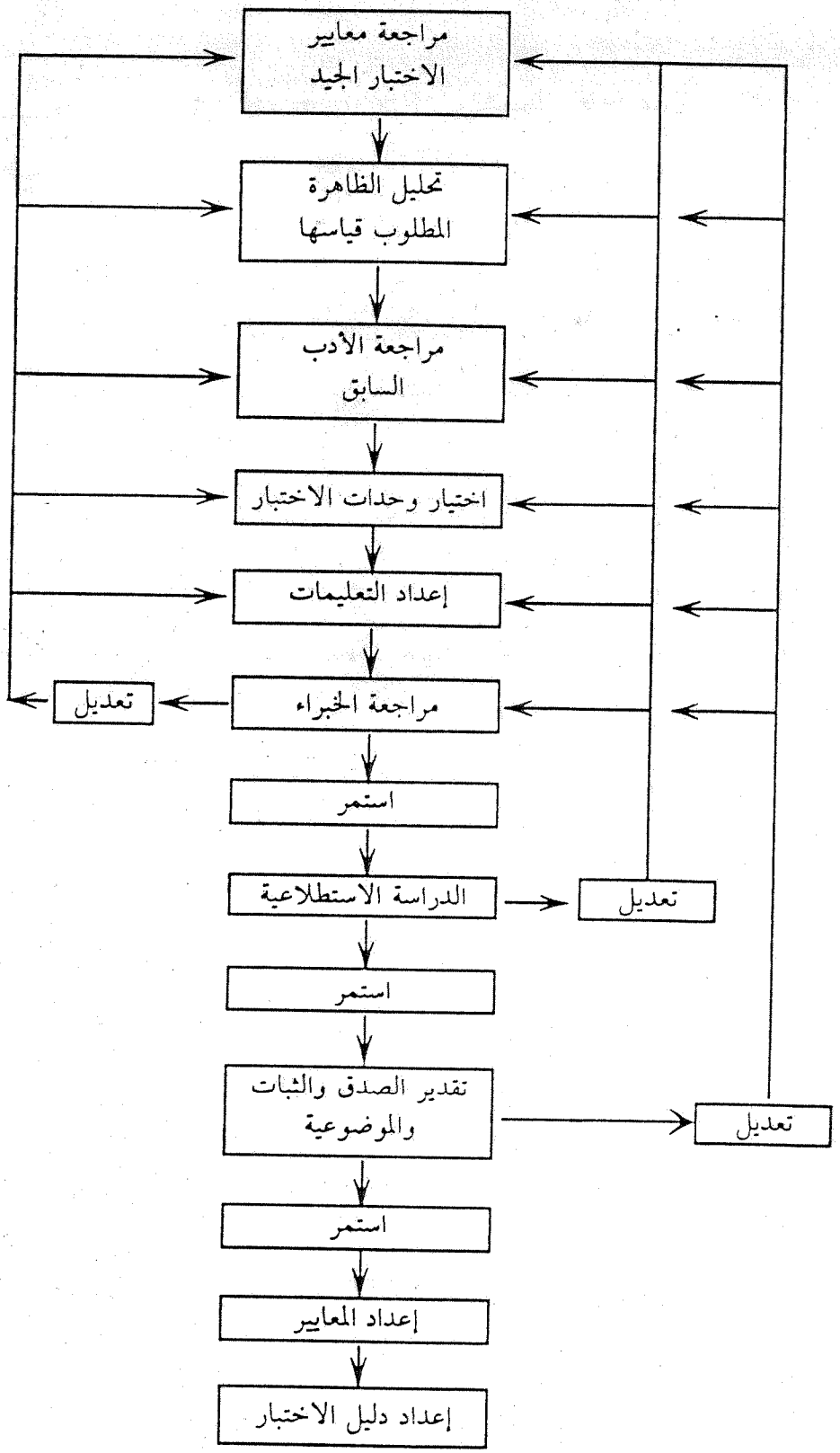
وفى هذا الصدد يشير جونسون ونيلسون Johnson and Nelson ١٩٨٦م إلى أن المعامل الذى يتم الحصول عليه عندما نقارن الدرجات التى يضعها أكثر من محكم يعرف بمعامل الموضوعية objectivity coefficient.

ويتوقف معامل الموضوعية بدرجة كبيرة على كيفية إعطاء التعليمات للمفحوصين وعلى مدى شمول ووضوح هذه التعليمات. فعندما تختلف درجات المحكمين اختلافات كبيرة عندما يقدرون الدرجات لنفس الشىء، فإن ذلك يعنى غياب الموضوعية، وأنه لا يوثق فى ثبات أو صدق أى من هذه الدرجات.

فالفروق الحادة فى تقدير الدرجة تحدث عندما يستخدم المحكمون أساليب وطرق وإجراءات مختلفة لتقدير الدرجة على الاختبار.

٤- **المعايير:** Norms : وهى عبارة عن درجات تعكس بوضوح أداء عينات التقنين (العينات المرجعية) على الاختبار. وتعرض المعايير فى جداول تعرف بجداول المعايير norms tables، وهى جداول تبين الدرجات الخام raw scores والدرجات المشتقة derived scores فى شكل أعمدة متوازية.

وتعد المعايير على أساس الطول، والوزن، والعمر الزمنى، والصف الدراسى، أو على أساس الربط بين بعض أو كل هذه المتغيرات. ولا يجب النظر إلى المعايير على أنها مستويات قياسية، أو على أنها تمثل الأداء الأمثل الذى يجب أن يصل إليه المفحوصين، أو على أنها تمثل ما يجب أن يصل إليه التلاميذ أو اللاعبين فى التحصيل، وإنما يمكن النظر إليها على أنها وسيلة من وسائل المقارنة أو محكات للتقويم.



شكل (٦-٧): خريطة التدفق لبناء الاختبار

٥- **متطلبات التطبيق:** Requirements: وتشمل الأجهزة والأدوات ومساحات الأرض الفضاء، وعدد المساعدين، والوقت الذى يستغرقه التطبيق، وسهولة التطبيق وغيرها من المتطلبات التى تلزم لتنفيذ الاختبار. وتستخدم متطلبات التطبيق كشرط من شروط المفاضلة بين اختبار وآخر. ففى حالة توافر الشروط السابقة (الصدق - الثبات - الموضوعية - المعايير) فى أكثر من اختبار، فإنه يفضل اختيار الاختبار المناسب للتطبيق على أساس أنه:

- لا يتطلب استخدام أجهزة أو أدوات مكلفة غالية الثمن.
- لا يستغرق وقتا طويلا فى التطبيق.
- لا يحتاج إلى تقنيات وخبرات فنية عالية لتشغيل الأجهزة أو استخدام الأدوات.
- فى جميع الحالات تفضل الاختبارات التى لا تتطلب استخدام أية أجهزة أو أدوات أو مساحات فضاء كبيرة من الأرض.

الخطوة الثانية: تحليل الأداء

يسبق بناء أو إعداد أى اختبار جديد تحديد طبيعة الأداء الذى سوف يقيسه هذا الاختبار، فالمهارات الحركية فى الأنشطة والألعاب الرياضية، واللياقة البدنية، والقدرات البدنية، والقدرات الحركية وغيرها، جميعها مفاهيم ومصطلحات تستلزم عند قياسها تحديدها تحديدا واضحا ودقيقا بحيث لا يكون هناك أى تداخل أو لبس فيما بين أى من هذه المفاهيم.

ويقصد بتحليل الأداء analysis the performance تحديد جميع المكونات components التى يتطلبها الأداء الناجح، ففى المجال الرياضى يستهدف تحليل الأداء تحديد المهارات أو القدرات المطوب قياسها (فالأداء الناجح فى كرة السلة - على سبيل المثال - يتضمن مهارات مسك الكرة وتنطيط الكرة، والمحاورة، والتمرير، والتصويب وغيرها).

وبعد أن يتم تحديد المكونات، يبدأ الباحث فى اختيار المكونات التى يريد قياسها، حيث تمثل المكونات التى يتم تحديدها المكونات التى يتضمنها الاختبار (البطارية) المزمع إعداده أو بناؤه.

الخطوة الثالثة: مراجعة الأدب السابق

وتشمل هذه الخطوة مراجعة الدراسات السابقة review Literature وكل ما كتب عن الموضوع فى المراجع العلمية المتخصصة. وهذه الخطوة ضرورية للتعرف على بطاريات الاختبار السابقة test batteries التى تستخدم لقياس المهارات الخاصة فى الألعاب أو لقياس الأداء الحركى فى مجالاته المختلفة.

ويستهدف الباحث عند مراجعته للدراسات والبحوث السابقة المرتبطة بالموضوع أو الظاهرة (موضوع الاختبار) **مجموعة من الأغراض هى:**

١- التعرف على أساليب وكيفية بناء اختبارات الأداء فى المجال النفس - حركى.

٢- قد يحصل الباحث على بعض الأفكار أو الوسائل التى تساعده على تنفيذ مشروعه.

٣- قد يختار الباحث بعض وحدات الاختبار المنشورة published test items لتكون ضمن الاختبار الذى يقوم بإعداده.

٤- قد يجد الباحث فى الاختبارات السابقة (المنشورة)، اختبارات تصلح لكى يستخدمها كمحكات خارجية external criteria لتقنين صدق الاختبار الجديد.

٥- حصر جميع المكونات الأساسية أو العوامل factors أو المهارات الخاصة specific skills التى يجب أن يتضمنها الاختبار الجديد.

وينصح علماء القياس النفس - حركى أنه فى حالة عدم الاستفادة من الكتب المتخصصة، فإنه يمكن للباحث أن يرجع إلى الدوريات العلمية المتخصصة فى هذا الشأن، **ومن أهمها ما يأتى:**

- مجلة الأبحاث الربع السنوية للتمرين البدنى والرياضة
Research Quarterly for Exercise and Sport

- معلم التربية الرياضية The physical Educator

- صحيفة التربية الرياضية والترويح والرقص
Journal of physical Eduction Recreation and Dance

وجميعها يصدر تحت إشراف الاتحاد الأمريكى للصحة والتربية الرياضية والترويح والرقص AAHPERD.

الخطوة الرابعة: اختيار أو بناء وحدات الاختبار

وتتضمن هذه الخطوة قيام الباحث باختيار selection وحدات الاختبار المناسبة له، أو بناء وحدات جديدة إذا كانت هناك ضرورة لذلك. ومن المهم بالنسبة لوحدات الاختبار هو أن تمثل هذه الوحدات - فى مجموعها - جميع المهارات أو الصفات النوعية specific qualities التى يستهدفها الاختبار، هذا بالإضافة إلى الآتى:

- أنه يمكن تطبيق مثل هذه الوحدات بسهولة.
- أن تكون الوحدات مرتبطة بالأداء الفعلى المطلوب قياسه.
- أن تتشابه مواقف الأداء فى الوحدات مع مواقف الأداء المستهدفة من القياس وبخاصة فى حالة اختبارات المهارات فى الألعاب.
- أن تشجع الوحدات على أشكال الأداء الجيد.
- أن يكون للوحدة معنى بالنسبة للمفحوصين وأن تتميز بالتشويق وفى جميع الحالات يفضل اختيار أكثر من وحدة اختبار واحدة لقياس المكون الواحد، مع ملاحظة أن تغطى الوحدات المختارة أو التى يتم إعدادها فى مجموعها جميع المظاهر السلوكية أو الأدائية الخاصة بالمكون الواحد، وألا تكون بعيدة عن هذا المضمون.

ويرى بعض علماء القياس أنه ينبغى اختيار وحدات الاختبار على أساس الشروط الخاصة بمتطلبات التطبيق بغض النظر عن مدى صلاحية هذه الوحدات، فالوقت اللازم للتنفيذ، والإمكانات المطلوبة، وتكاليف الأجهزة والأدوات، وصعوبة تقدير الدرجة وتسجيلها، وعدد المساعدين وغيرها، تعد من الشروط المهمة عند اختيار الوحدات اختياراً أولياً.

الخطوة الخامسة: إعداد تعليمات الاختبار

وتتضمن إعداد التعليمات directions وطريقة الأداء، وعدد المحاولات، والأخطاء التى تؤثر على الأداء، وطريقة حساب الدرجة، وترتيب أداء الوحدات وغيرها.

ويجب أن تكون التعليمات واضحة ودقيقة حتى لا يتأثر ثبات وموضوعية الدرجات، بالإضافة إلى أن دقة ووضوح التعليمات من العوامل المهمة جداً لوقاية المفحوصين من الاضطراب والحد من تأثير التخمين.

ويجب إعداد التعليمات كتابة حتى يمكن الالتزام بها بالنسبة لجميع القائمين على التنفيذ، ولكون هذا الإجراء يسهل على المحكم (المساعدين) وعلى المفحوص فهم الاختبار فهما كاملا مما يكون له تأثيره المباشر على نجاح عمليات القياس.

الخطوة السادسة: مراجعة الخبراء

وتستهدف هذه الخطوة عرض وحدات الاختبار التى استقر رأى الباحث عليها ومعها التعليمات على مجموعة من المتخصصين والخبراء من زملاء وأساتذة الباحث، ومن غير العاملين معه فى مشروع البحث، حيث يطلب من هؤلاء الخبراء تقديم المساعدات التى يمكن أن تساهم فى نجاح المشروع، وتقديم ملاحظاتهم مكتوبة، واقتراح أية تعديلات يرونها بالنسبة لصدق وثبات الوحدات وإجراءات تطبيقها، وفى هذا الخصوص نود أن ننبه الباحثين إلى الملاحظات المهمة التالية:

١- ألا يكون الباحث الرئيسى حساساً لأية آراء أو تعديلات تقدم إليه شفاهة أو كتابة من زملائه.

٢- ألا يقبل كل المقترحات التى تقدم إليه دون مراجعة وفحص وتدقيق.

٣- أن يعد قائمة بأسماء ووظائف وعناوين هؤلاء الخبراء، وعلى أن تكون هذه القائمة من الوثائق الرئيسية فى مشروع البحث، فقد لوحظ أن هذا الإجراء يزيد من إحساس الخبراء بالمسئولية مما يدفعهم إلى العناية والاهتمام بتسجيل ملاحظاتهم.

٤- يستوجب تقدير الخبراء واحترامهم والعناية بالمادة المكتوبة المقدمة إليهم شكلا وموضوعاً، واختيار الأوقات المناسبة للاتصال بهم، وفى جميع الحالات يفضل إجراء مقابلة شخصية - إن أمكن - للرد على استفساراتهم وكذا الاستماع إلى ما يبدونه من آراء وملاحظات قد يتعذر تسجيلها كتابة.

الخطوة السابعة: التطبيق الأولى للاختبار

عندما يتم اختيار وحدات الاختبار فى ضوء آراء الرفاق (الخبراء) Experts، يقوم الباحث الرئيسى بتطبيق الاختبار على مجموعة صغيرة من المفحوصين (الطلاب أو اللاعبين) وذلك بغرض تحقيق الآتى:

١ - مدى مناسبة وحدات الاختبار.

٢ - تحديد الزمن الكلى الذى يستغرقه تطبيق الوحدات.

٣ - مدى وضوح التعليمات.

٤ - الكشف عن أية مشكلات غير متوقعة قد تحدث أثناء التطبيق.

ويفضل أن يتم التطبيق الأولى للاختبار (الوحدات) على مجموعة من المفحوصين من نفس المجتمع الذى يعد له الاختبار، أو بمعنى آخر أن يكون لدى مجموعة المفحوصين نفس خصائص المجتمع الأصلى الذى سوف يطبق

عليه الاختبار، كما يفضل أن يقوم المساعدون بتنفيذ الاختبار (الوحدات)، فى نفس الوقت الذى يتولى فيه الباحث (المسئول الأول عن المشروع) ملاحظة وتسجيل كافة المشكلات والصعوبات التى يراها أو تبلغ إليه من المساعدين.

وقد ينظر إلى التطبيق الأولى للاختبار على أنه تجربة استطلاعية pilot study، وأنه بدون هذه التجربة قد لا تظهر بعض المشكلات التى تظل خافية فيكون تأثيرها شديداً عندما يطبق الاختبار على مجموعات كبيرة من المفحوصين. وبعد الانتهاء من التجربة الاستطلاعية، يقوم الباحث الرئيسى بعمل التغيرات الضرورية، فكلما استخدم الباحث وحدات اختبار أكثر فى التجربة الاستطلاعية، كلما مكنه ذلك من استبعاد وحدات الاختبار غير المناسبة.

الخطوة الثامنة: تقدير الصدق والثبات والموضوعية

وتستهدف هذه الخطوة تعيين الخصائص الإحصائية للاختبار، ووفقا للمعلومات السابقة عن معاملات الصدق جدول (٦-٣) ومعاملات الثبات جدول (٦-٤) يستطيع الباحث أن يقدر هذه المعاملات عن طريق تطبيق الاختبار (الوحدات) على عدد كبير من المفحوصين، حيث يشترط فى هؤلاء المفحوصين الآتى:

١- أن تكون العينة ممثلة للمجتمع الأصلى - الذى ستطبق عليه الوحدات فيما بعد - تمثيلا دقيقا.

٢- أن تختار العينة بالطريقة العشوائية وأن تكون مناسبة من حيث العدد.

ويرى علماء القياس أنه يفضل استخدام الصدق التلازمى concurrent validity بالنسبة للاختبارات التى لها معيار مرجعى norm-referenced tests حيث يمكن حساب هذا الصدق عن طريق استخدام واحدة من الطرق التالية:

طريقة الدورى:

وتستخدم هذه الطريقة لتقدير صدق اختبارات المهارات فى الألعاب، حيث يقام دورى كامل أو نصف دورى فى اللعبة أو النشاط بين المفحوصين (اللاعبين) ، يلى ذلك ترتيب هؤلاء المفحوصين وفقا لنتائجهم فى الدورى، ثم نقارن بين ترتيبهم فى الدورى وترتيبهم على درجات الاختبار المقترح، ثم نقوم بحساب الصدق التلازمى للاختبار باستخدام معامل ارتباط فروق الرتب لسبيرمان أو معامل ارتباط بيرسون للدرجات الخام. فإذا كان الاختبار صادقا، فإن المفحوص الذى يؤدى جيدا على الاختبار سوف يؤدى جيداً فى الدورى، وأن المفحوص الذى يؤدى بمستوى ضعيف على الاختبار سوف يكون مستواه ضعيفا فى الدورى.

معاملات الارتباط باختبارات أخرى ثبت صدقها:

وتعتمد هذه الطريقة على حساب معامل الارتباط بين درجات الاختبار ودرجات اختبار آخر ثبت صدقه بالنسبة لقياس السمة أو الصفة أو المهارة المطلوب قياسها، فإذا كان معامل الارتباط المحسوب مقبولا دل ذلك على أن الاختبار الجديد (المقترح) صادقا. وتكمن الصعوبة الحقيقية فى هذه الطريقة فى أن الاختبارات التى تستخدم كمحك خارجى قد لا تكون متوفرة، أو تكون مكلفة ماديا، أو تستغرق وقتا طويلا فى التطبيق، أو نتيجة لبعض الأسباب الأخرى التى قد تحول دون ذلك، وفيما عدا ذلك فإنه فى حالة توافر مثل هذه الاختبارات، فإن هذه الطريقة تعد من أفضل طرق حساب الصدق.

تقديرات الخبراء:

ويمكن حساب الصدق التلازمى للاختبار المقترح عن طريق حساب معامل الارتباط بين تقديرات الخبراء experts للأداء الفعلى للمفحوصين فى النشاط ودرجات المفحوصين فى الأداء على الاختبار المقترح.

وفيما يتعلق بالثبات فيمكن تقديره بأى طريقة من طرق حساب الثبات التى تم الإشارة إليها بالتفصيل فى الفصل الثانى من هذا الكتاب. ويشير ميللر Miller ١٩٩٤م إلى أن طريقة إعادة الاختبار the test - retest قد تكون مناسبة لتقدير معاملات الثبات والموضوعية فى نفس الوقت.

الخطوة التاسعة: إعداد المعايير

يتم إعداد المعايير norms فى حالة الاختبارات التى لها معيار مرجعى نسبى norm - referenced tests ، ويتم إعداد المستويات standards فى حالة الاختبارات التى لها محك مرجعى قياسى criterion- referenced tests. والنسبة للاختبارات التى لها معيار مرجعى نسبى فإنه يمكن بناء جداول المعايير tables of norms عن طريق حساب الدرجات (ذ - Z-scores) أو الدرجات (ت - T- scores) أو (المئينيات - percentiles) وتعد المعايير على أساس العمر الزمنى والجنس (بنين - بنات)، وعادة ما يستخدام لإعداد مثل هذه المعايير عينات كبيرة الحجم.

الخطوة العاشرة: إعداد دليل الاختبار

وتمثل الخطوة الأخيرة من خطوات بناء الاختبار، ويتضمن دليل الاختبار test manual الشروط والإجراءات النهائية للاختبار فى ضوء التطبيق النهائى لها ، وكذا التعليمات ووصف الاختبارات، وكيفية تخطيط الملاعب وإعداد الأدوات، وكيفية حساب الدرجات، وكذا المعايير، وأية أشكال أو صور توضيحية بالإضافة إلى قائمة بالمراجع الرئيسية التى يمكن الرجوع إليها للاستزادة.

معجم إنجليزى - عربى للمصطلحات

A

Assistans مساعدون

Absolute standards مستويات مطلقة

Accuracy tests اختبارات الدقة

Abilities قدرات

Analysis the performance تحليل الأداء

C

Competence كفاءة - مقدرة

Concurrent validity الصدق التلازمى

Components مكونات

Construct validity صدق التكوين الغرضى

Criteria محكات

Criterion-referenced المعيار المرجعى القياس

Changing-alternatives scale
مقياس متغير البدائل

Checklist قائمة مراجعة (تدقيق)

Criterion - referenced analysis
التحليل معيارى المرجع القياسى

D

Directions تعليمات الاختبار

Dichotomous rating مقاييس ثنائية التقدير

Domains مجالات

Distance tests اختبارات المسافة

Descriptive scales المقاييس الوصفية

E

Equal- appearing intervals
المسافات المتساوية ظاهريا

External Criterion محك خارجى

F

عوامل

Factors

Forms أشكال - صيغ - صور

Feedback تغذية راجعة - مرتدة

Four- seconds dash test
اختبار العدو ٤ ثوان من البدء العالى

Fixed scale مقياس ثابت

G

Graphic scales المقاييس التخطيطية

General Components مكونات عامة

Generality vs. specificity
العمومية فى مقابل الخصوصية

I

items وحدات

Individual score sheet بطاقة تسجيل فردية

M

Master shee البطاقة الرئيسية

Mental Performance الأداء العقلى

Motor abilities القدرات الحركية

Motor Learning التعلم الحركى

Motor skills مهارات حركية

Mastery tests اختبارات التمكن

.astery skills مهارات التمكن

N

Norm tables جدول المعايير

Norm- referenced analysis

التحليل معيارى المرجع النسبى

Norms معايير نسبية

Numerical scales المقاييس العددية

O

objectivity Coefficient معامل الموضوعية

objective tests اختيارات موضوعية

P

Pilot study دراسة استطلاعية

Psychomotor tests اختبارات نفس - حركية

retest اختبار قبلى

osttest اختبار بعدى

rocess - oriented skill مهارة توجه المهمة

aired comparison المقارنة الزوجية

ublished tests اختبارات منشورة

R

ative standards مستويات نسبية

titive - performance tests

اختبارات الأداء المتكرر

Relative rating scales

مقاييس التقدير النسبية

nk - order نظام الرتب

ng ترتيب

scale مقياس تقدير

Reliability Correlation Coefficient

معامل ارتباط الثبات

S

specitic qualities صفات نوعية

standardization sample عينة التقنين

Sports skills المهارات فى الألعاب

standardized tests اختبارات مقننة

subjective ratings التقديرات الذاتية

subjective judgments الأحكام الذاتية

side- step test اختبار الخطوة الجانبية

sit - ups tes اختبار الجلوس من الرقود

softball throw رمى الكرة الناعمة

T

Timed tests اختبارات موقوتة

Testing stations محطات اختبار

Tests of sport skills

اختبارات المهارات فى الألعاب

Teacher- made tests

اختبارات من إعداد المعلم

Tests of motor performance

اختبارات الأداء الحركى

Test batteries بطاريات الاختبار

Targets أهداف

Traits سمات

V

Vertical Jump الوثب العمودى

معامل ارتباط الصدق

Validity correlation coefficient

مراجع الفصل السادس

١- محمد حسن علاوى، ومحمد نصرالدين رضوان (١٩٧٨): الاختبارات المهارية والنفسية في المجال الرياضى. دار الفكر العربى، القاهرة.

٢- محمد حسن علاوى، ومحمد نصر الدين رضوان (٢٠٠٠): القياس فى التربية الرياضية وعلم النفس الرياضى. ط٦، دار الفكر العربى، القاهرة.

3 - AAHPERD (1988) : Physical Best. Reston, VA : American Alliance for Health, physical Education, Recreation and Dance.

4 - Hastad, D. H, and Lacy, A.C. (1988) : Measurement and evaluation in contemporary of physical education. scottsdale, AZ : Gorsuch Scarisbrick.

5 - Safrit, M.J. (1990) : Jntroduction to measurement in physicl education and exercise science. St. Louis : Times Mirrorl Mosby.

6 - Standards for educational and psychological testing (1985) : Washington, DC : Americal psychologial association.

7 - Strand, B.N. , and wilson , R. (1993) : Assessing sport skills. Champaign, lL : Humon kinetics.

8 - Verducci, F.M (1980) : Measurement Concepts in physical Educati The C. V. Mosby Comppany.

المؤلف فى سطور

الدكتور/ محمد نصر الدين رضوان

- بكالوريوس فى التربية الرياضية مع مرتبة الشرف ١٩٦٣م.
- ماجستير فى التربية الرياضية ١٩٧٢م.
- دكتوراه الفلسفة فى التربية الرياضية جامعة حلوان ١٩٧٧م.
- دراسات فى الإحصاء التربوى من معهد الدراسات والبحوث الإحصائية بجامعة القاهرة (٧٢ - ١٩٧٣م).
- مؤسس وأول رئيس لقسم التربية الرياضية والصحية بجامعة الملك عبد العزيز بالسعودية (٨٠ - ١٩٨٤م).
- أول رئيس لقسم علم النفس الرياضى بكلية التربية الرياضية للبنين - جامعة حلوان - ١٩٩٧م.
- له العديد من الإصدارات العلمية والترجمات فى مجالات القياس والتقويم والبحث العلمى والإحصاء.
- أشرف على العديد من البحوث والرسائل العلمية للحصول على درجتى الماجستير والدكتوراه فى مختلف الجامعات المصرية.
- شارك (كمحكم) فى عدد من المؤتمرات العلمية من مصر والسعودية والسودان والأردن والعراق.
- أمين شعبة التعليم بالنقابة العامة للمهن الرياضية (٩٧ - ٢٠٠١م).
- عضو لجنة التعليم بالحزب الوطنى الديمقراطى (أمانة القاهرة).
- عضو لجنة التخطيط الفنى باللجنة الأولمبية المصرية (٢٠٠١م).
- عضو اللجنة العلمية الدائمة للترقى لوظائف الأساتذة والأساتذة المساعدين (٩٨ - ٢٠٠١م).
- عضو المجالس القومية المتخصصة (٢٠٠٤م).
- رئيس مجلس إدارة الاتحاد المصرى للقوس والسهم.

هذا الكتاب

يقدم أحدث المعارف والمعلومات التى تعالج الموضوعات الرئيسية فى
ياس والتقويم فى المجال الرياضى.

فهو يتناول طبيعة القياس فى هذا المجال موضحاً مجالاته ووظائفه
واته، كما يقدم شرحاً وافياً وحديثاً لشروط الثبات والصدق وكيفية
ديرها وتقويمها.

ويقدم الكتاب شرحاً وافياً للقياس المعرفى موضحاً ماهيته وخطوات بناء
ختبارات المعرفية، كما يعرض لموضوع القياس الانفعالى وأساليبه، وأدواته،
لطرق الوصفية والكمية له.

ويتناول الكتاب القياس النفس - حركى موضحاً طبيعته وإجراءاته
ماط الاختبارات النفس - حركية وخطوات بنائها.

ويأتى فى نهاية كل فصل معجم إنجليزى - عربى للمصطلحات، وقائمة
لمراجع الأصلية اللازمة لخدمة القراء وبخاصة طلاب الدراسات العليا الذين
رغبون فى تناول موضوعات الكتاب بمزيد من الدراسة والبحث.