**المحاضرة الثالثة : الشروط العلمية للاختبارات والقياسات**

الصدق –الثبات – الموضوعية – المعيارية

تمهيد :

 الاختبار هو موقف مقنن لقياس استجابة شخص أو سلوك فرد ، وقد يكون تمرينا أو سؤالا أو وضعية ، ولكن علينا أن نتفق أنه ليس كل تمرين أو موقف أو سؤال يصلح لكي يكون اختبار ..كما أن الاختبارات التي صممت في بيئة ما وتم التأكد منها لا يعني بالضرورة نجاحها في كل البيئات وبالخصوص تلك التي تبني الاستبيانات والمقاييس التي تعتمد على العبارات ، لأن العادات والتقاليد والذهنيات وتركيبات المجتمع ومستوى التحضر والتمدن يختلف من مجتمع لآخر ، وإذا لم تراعى هاته الأمور في الاختبار المعتمد فإننا الباحث لا يضمن التقويم الجيد ولا القياس الجيد للظاهرة ( كمثال على هذا وأثناء البحث عن اللحم المفضل لدى الأفراد فحينما يوزع السؤال في بيئة غربية فهو يذكر أغلب الحيوانات ومنها الخنزير ويتحصل خلال تفريغ النتائج على نسبة من الأشخاص قد اختاروا ذلك النوع من اللحم ، بينما عندما يوزع هذا السؤال في بيئة عربية إسلامية سيكون عدد الذين اختاروا هذا النوع من اللحوم هو 0 ، حينها يقول الباحث بأن هذا المجتمع لا يفضل تناول هذا اللحم ، ولكن هذا الحكم غير صحيح لأن البيئة التي يتواجد فيها المجتمع العربي أو المسلم يحرم عليه دينه تناوله وبالتالي فإن حكم الباحث بعيد كل البعد عن القياس الحقيقي .

 وقد يقيس أيضا باحث ما أهم النشاطات الرياضية المحببة لدى المراهق ومن ضمن الاقتراحات ركوب الأمواج أو الغطس أو الرماية بالسلاح بينما البيئة التي يتواجد بها أفراد عينة البحث لا تحوي مطلقا على أماكن لممارسة هاته الهوايات وبالتالي فإن عدم اختيارها ليس لعدم الرغبة بها وإنما لعدم تواجدها ) وبالتالي " فالاختبارات التي يتم بناؤها وتقنينها على عينات تمثل مجتمع المستفيدين تعد أصلح من غيرها التي تم بناؤها وتقنينها على عينات تمثل مجتمعا آخر مهما بلغت درجة التشابه بين المجتمعين "[[1]](#footnote-2) إذا لكي يكون الاختبار أداة علمية لقياس ظاهرة ما ، يتوجب على الباحث أن يوفر الشروط العلمية في هذا الاختبار وإلا أصبحت نتائجه غير موثوق بها وبالتالي لا تؤدي بالباحث إلى إصدار تقويم حقيقي عن الظاهرة محل البحث، وهنا يذكر الباحثين عدة شروط سنركز على أكثرها تداولا وتأثيرا على جودة نتائج القياس وهي :

1. الصدق
2. الثبات
3. الموضوعية
4. المعيارية

**1-الصدق : Validity**

**1-1تعريف الصدق:** أحد أكثر المفاهيم الأساسية أهمية في مجال الاختبارات ، حيث يُعتبر الصدق أهم شروط العلمية له. والاختبار الصادق هو الذي ينجح في قياس ما وُضع من أجله. حيث يُشير " رايتسون Wrightstone وجاستمان Justman وروبينز Robbins إلى أن صدق الاختبار يعني المدى الذي يحقق به الاختبار أو أي متغير آخر الغرض الذي وضع لأجله " ، كما يشير بارو Barrow ومك جي McGee إلى أن الصدق يعني المدى الذي يؤدي فيه الاختبار الغرض الذي وضع من أجله ، حيث يختلف الصدق وفقا للأغراض التي يراد قياسها [[2]](#footnote-3)

**صدق الاختبار : هو الدرجة التي يقيس بها الاختبار الشيء المراد قياسه**

**1-2 مظاهر الصدق** : للاختبار الصادق ثلاث مظاهر تتمثل في

1. الثبات Reliability ويقصد به أن الاختبار الصادق يكون ثابتا في في معظم الأحيان إذا كانت ظروف القياس متوفرة ( إلا إذا تدخلت عوامل تحول دون ذلك )
2. التعلق Relevence ويقصد به مدى اقتراب درجات الاختبار من الدرجات الحقيقية الخاصة بالعينة كلها
3. الصدق النسبي Relative بمعنى أن الاختبار المقنن على عينة أو مجتمع يكون صادقا فيه ، وقد يكون أقل صدقا نسبيا إذا أجري على عينة أو مجتمع آخر

1-2- **انواع الصدق:** [[3]](#footnote-4)

1-2-1- الصدق الظاهريFace validity: ويقصد به ملائمة الاختبار للقياس ظاهريا ( أي يبدوا صادقا في صورته الظاهرية) ، و يُعتبر هذا النوع من الصدق أقل أنواعه أهمية أي أضعفها وأقلها استخداماً لأنه ويعتمد على منطقية محتويات الاختبار ومدى ارتباطها بالظاهرة المُقاسة وهو يُمثل الشكل العام للاختبار أو مظهره الخارجي، من حيث مفرداته ومدى وضوحها وموضوعيتها ووضوح تعليماتها للاختبار.

1-2-2 الصدق الذاتي : ويسمى أيضا Index of Reliability وهو صدق الدرجات التجريبية بالنسبة للدرجات الحقيقية التي خلصت من شوائب أخطاء الصدفة ، حيث تعتبر الدرجات الحقيقية ميزان أو محك صدق الاختبار حيث يتم حسابه بعد حساب الثبات بأسلوب الاختبار – إعادة الاختبار عن طريق المعادلة الصدق الذاتي

الصدق الذاتي =$\sqrt{الثبات}$

1-3-2- صدق المحتوى(المضمون) (المنطقي)Logical Validity : وهو أن يقيس الاختبار ما وُضع من أجله بدون أي زيادة غير ضرورية ويعتمد ذلك على الفحص الدقيق لمضمون الاختبار، ويُعرّفُهُ "حسنين" بأنّه: " مدى جودة تمثيل محتوى الاختبار لفئة من المواقف أو الموضوعات التي يقيسها". يعتمد صدق المحتوى للاختبار وبصورة أساسية على مدى إمكانية تمثيل الاختبار لمحتويات عناصره وكذلك المواقف والجوانب التي يقيسها تمثيلاً صادقاً ومتجانساً وذات معنوية عالية لتحقيق الهدف الذي وُضع من أجله الاختبار وللوصول إلى ذلك يجب على القائم بالاختبار مراعاة مايلي:

* تحديد السمة أو الظاهرة أو الخصوصية قيد البحث تحديدا منطقيا
* التعرف على أبعاد السمة أو الظاهرة أو الخصوصية المقاسة وأهمية كل جزء فيها .
* وُضع مفردات بما يتفق والابعاد أو الاجزاء التي استقر عليها الرأي في ضوء المرحلتين السابقتين.

1-3-3- الصدق المرتبط بالمحك(الصدق التجريبي): ويُقصد به مدى ارتباط درجات الاختبار المقنن وبين درجات اختبار معياري أو محكي خارجي مستقل يتناول السلوك نفسه أو النشاط الذي يتناوله الاختبار بالقياس ويتمتع بمستوى صدق عالي ، حيث يتم مقارنة درجات الاختبار بمتغير أو متغيرين خارجيين يُعتبران مناسبين لتوفير قياس للسمة موضوع الاختبار. و يُعتبر الصدق التجريبي من أفضل أنواع الصدق وأكثرها شيوعاً،.

مثال: إذا كنا بصدد قياس اللّياقة البدنية عن طريق اختبار جديد بُنى وقُنّن لهذا الغرض فمن الممكن إيجاد صدق هذا الاختبار عن طرق إيجاد معامل الارتباط بينه وبين اختبار آخر أُثبت صدقه لقياس اللّياقة البدنية.

1-4- العوامل المؤثرة في الصدق: [[4]](#footnote-5)

يتأثر صدق الاختبار بالعديد من العوامل أهمها : المحك الذي يستخدم لحساب صدق الاختبار ، السن ، الجنس ، النضج ، خبرة التعلم للأفراد الذين يحسب عليهم الصدق

**2- الثبات Reliability:[[5]](#footnote-6)**

2-1- تعريفه: يُمثل العامل الثاني في الأهمية بعد الصدق في عملية بناء وتقنين الاختبارات، ويعني أن يكون الاختبار على درجة عالية من الدقة والإتقان والاتساق والموضوعية فيما وُضع لقياسه. ويُعرّف بأنّه اتساق في النتائج ويُعتبر الاختبار ثابتاً إذا حصلنا منه على نفس النتائج لدى إعادة تطبيقه على نفس الأفراد وفي نفس الظروف.

**الثبات هو دقة الاختبار في ما يقيسه**

ويتضمن معنى الثبات في القياس النقاط التالية [[6]](#footnote-7)

-يعزى الثبات إلى النتائج التي نتحصل عليها من الاختبار وليس للاختبار نفسه، لأن الثبات يختلف طبقا لنوع العينة والموقف المستخدم فيه ، لذلك من الأحسن استخدام : ثبات درجات الاختبار وليس ثبات الاختبار.

-تقدير الثبات يشير دائما إلى نوع معين من الاتساق

-الثبات ضروري للاختبار ولكنه ليس بديلا للصدق

-الثبات يتسم بالصبغة الاحصائية على عكس الصدق الذي يتسم بالتحليل المنطقي دون أي دليل عملي .

**2-2- طرق حساب الثبات[[7]](#footnote-8):**

**2-2-1- طريقة إعادة الاختبارTest-retest reliability :** أشار إليها "فان دالين Van Dalen " حيث يُعطى الاختبار لنفس المفحوصين مرتين يفصل بينهما فارق زمني ثم يُحسب معامل الارتباط بين التطبيقين، في هذه الطريقة يتم إعداد أداة البحث على نفس أفراد العيّنة مرتين أو أكثر تحت ظروف متشابهة قدر الإمكان، ثُمّ يُستخدم معامل الارتباط بين نتائج التطبيق في المرتين ويُشير معامل الارتباط لثبات الأداء ويُسمى هذا المعامل بمعامل الاستقرار". وهي إحدى أهم الطرق لحساب الثبات خصوصا إذا كان الاختبار مقياسا أو موقفا أو أداءا حركيا يعبر عنه بالأرقام

**2-3-2- طريقة التجزئة النصفيةSplit –half relibility :** يُطبق الاختبار مرة واحدة، ولكن تُقسم بُنوده عشوائياً إلى نصفين، ويُحسب الارتباط بين درجات النصفين. وتتلخص إجراءات هذه الطريقة في تقسيم الاختبار إلى نصفين يُراعى فيهما تساوي عدد الوحدات.

هي كثيرة الاستخدام هاته الطريقة لم يتعلق الاختبار بالورقة والقلم ولا تصلح لاختبار الأداء في التربية البدنية والرياضية

2-3-3- **طريقة الصور المتكافئة Equivalence – form :** تُعد صورتين متكافئتين وتُطبقان على المفحوصين، ثم يحسب معامل الارتباط بين نتائج الصورتين. وتستخدم هذه الطريقة لإيجاد معامل الثبات عندما تتوافر صورتان أو أكثر متكافئتين من الاختبار، ومعنى التكافؤ هنا أن يكون قد تم بناء وتصميم كل صورة من الصورتين على حده وبطريقة مستقلة بشرط توافر عدد من المواصفات المحددة وهي:

* شروط التكافؤ الإحصائي (المتوسط، الانحراف المعياري، معامل الارتباط، معامل الصدق الداخلي...)
* تساوي عدد الأسئلة في الصورتين
* تماثل صياغة الأسئلة في الصورتين
* تماثل المحتوى في الصورتين
* تساوي مستوى الصعوبة في الصورتين
* تماثل متغيرات القياس في الصورتين (التعليمات، الزمن...).

2-4- **العوامل المؤثرة في الثبات:[[8]](#footnote-9)**

* خصائص أفراد العينة والفروق الفردية بينهم مثل الذكاء ، التوتر ، خبر التعلم والممارسة ، الحالة الصحة وكذا تعود أفراد العينة على الاختبارات
* خصائص الاختبار مثل طوله أو عدد محاولاته ، درجة صعوبته ، قدرته على التمييز ، طبيعة الأداء وكذا درجة التجانس في محاولات الاختبار
* خصائص عملية الاختبار كتنظيم البيئة أو الفصل الذي يتم فيه الاختبار ، الاحماء وتكافئ الفرص وكذا الاجراءات المستخدمة في عملية القياس ( عدد الدرجات ، تسجيل الأخطاء )
* خصائص القائمين بتطبيق الاختبار وذلك بكفائتهم ، ثقتهم واقتناعهم بأهمية الاختبار ، التركيز والتعود على استخدام أدوات الاختبار والقياس وعدم التحيز وأخيرا تناسب عددهم مع طبيعة الاختبار

**3- الموضوعية Objectivity :**

3-1- **تعريفها**: وتعني التحرر من التحيز والتعصب وعدم إدخال العوامل الشخصية للمختبر ، حيث تظهر عندما تقوم مجموعة مختلفة من المدرسين أو المحكمين بنفس الاختبار فتظهر نفس النتائج[[9]](#footnote-10)، فالموضوعية تعني أن تصف قُدُرات الفرد كما هي موجودة فعلاً، لا كما نريدها أن تكون. كما تعني عدم تأثر الاختبار بتغير المحكمين، أو أنّ الاختبار يُعطي نفس النتائج مهما كان القائم بالتحكيم.

**الموضوعية تعني قياس الاختبار للظاهرة بحيادية دون التأثر بالمحكم**

**3-2- العوامل المؤثرة في الموضوعية[[10]](#footnote-11):**

* درجة وضوح الاختبار : كلما كان الاختبار واضحا لجميع الممتحنين كلما كان أكثر موضوعية .
* معامل ثبات الاختبار : حيث ارتفاع معامل الثبات يؤدي دائما إلى إرتفاع معامل الموضوعية والعكس.
* مدى فهم المختبر لطبيعة الاختبار وطريقة التنفيذ والتسجيل تؤثر في معامل الموضوعية طبقا لمستوى هذا الفهم والالمام ، فالاستيعاب الكامل من قبل المختبر للاختبار يؤدي إلى ارتفاع معامل الموضوعية . وتتوضح هذه النقطة خصوصا عند توزيع الاستبيانات أو المقاييس ، حيث تؤثر درجة الفهم والالمام بموضوع القياس على إجابات المختبرين لذلك يجب أن يفهم المختبر أنه ليس محل تحقيق بهذا القياس وإنما سلوكه هو من يتعرض للقياس بعيدا كل البعد عن شخصه أو اسمه أو مكانته في المجتمع ، ولتبسيط هذا نعطي مثالا بالوقت الذي يخصصه الأشخاص للمطالعة والتعرف عما إذا كان ارتفاع سعر الكتاب يؤثر على المقروئية يجب أن يفهم المختبر أنه عليه أن يجيب بموضوعية وأننا لن نقوم بنشر إجابته للعام ونقوم بتشويه سمعته أو مكانته إذا كانت مقروئيته متدنية .

**3-3- شروط تحقيق الموضوعية**[[11]](#footnote-12):

* تعليمات دقيقة وواضحة لإجراءات القياس المختلفة
* تبسيط وتسهيل الإجراءات بحيث تكون قابلة للتطبيق العلمي .
* استخدام الاجهزة والأدوات كلما أمكن ذلك
* التقليل من استخدام الدرجات الناتجة عم العمليات الحسابية المعقدة والطويلة وتفضيل الاداء المباشر
* اختيار محكمين أذكياء ومدربين تدريبا جيدا
* الاطلاع المستمر على كل جديد بالنسبة لأسلوب القياس ، طرق تقنين الاختبارات وضبط المتغيرات ، أساليب التحليل الإحصائي المناسبة .

**4-المعايير Norms**

هي قيم تمثل أداء مجتمع خاص في اختبار معين [[12]](#footnote-13)

فمقارنة درجة الفرد بمعيار درجات مجموعة من الأفراد لا تدلنا عما يجب أن تكون عليه درجة هذا الفرد ، ولكنها تدلنا فقط كيف أن هذا الفرد أدى الاختبار عند مقارنته بآخرين من نفس مستواه ، كما أن وجود المعايير والمستويات يسمح للمختبر أن يتعرف على مركزه النسبي في المجموعة ، وهذا يعتبر إجراءا هاما وضروريا إذا كان الباحث يبحث عن نتيجة الاختبار من أجل تقويم لاعبيه وكشف مدى استعدادهم للمنافسة ، أو مستوى تحضير تلاميذه استعدادا للمسابقات الرسمية .

 كما يصادف الباحث أو المربي أو المدرب أحيانا بعض المتدربين الذي لديهم ثقة زائدة بأنفسهم أو النوع الأخر الذي يحتقر أدائه وموهبته فيضطر إلى استخدام المعايير لجعل النوع الأول أكثر حرصا على المزيد من العمل والنوع الآخر على الثقة في النفس وذلك عندما ينظر النوعين إلى نتيجة أدائهم مقارنة بأفراد آخرين في نوادي وفرق أخرى

 كما أن المعايير تقسم إلى أربعة من حيث المستويات[[13]](#footnote-14) :

* المعايير الوطنية : تتضمن الكثير من معايير الاختبارات المنشورة ، وغالبا ما يتم نشرها على أساس ( السن ، الجنس ، الصف الدراسي ) وتتضمن التحصيل الدراسي ، الصفات البدنية
* المعايير الخاصة بمجموعة خاصة : وتكون خاصة باختصاص رياضي معين أو عينة خاصة ، مثل نتائج اختبار سارجنت للقفز العمودي من الثبات لدى لاعبي كرة السلة أقل من 18 سنة ، أو مقياس الثقة بالنفس لدى لاعبي كرة القدم أقل من 14 سنة لنادي اتحاد الجزائر ، وهكذا .
* المعايير المحلية : ويقصد بها المعايير الخاصة بمدينة أو منطقة سكنية أو جماعة أو مدرسة أو نادي أو شركة محددة ، وهي محدودة مقارنة بالمعايير المدروسة سابقا ( الوطنية والخاصة ) لأنها تستخدم للمقارنة داخل المجموعة نفسها .( كأن نقول عدد الغيابات المسجلة بين العمال أو الموظفين خلال شهر سبتمبر أو عدد التأخرات ، عدد العطل المرضية ) كما يمكن أن تكوت المقارنات داخل الفريق أيضا ( عدد الاصابات التي تلقاها اللاعب خلال موسم رياضي وهكذا .
* المعايير المدرسية ( معايير الصفوف ) وتستخدم عند مقارنة متوسط أداء قسم دراسي معين على اختبار معين بالنسبة للأداء الكلي للمدرسة على نفس الاختبار ، كأن نقول أن معدل القسم 13.21 هو الأعلى في المؤسسة التربوية من بين أربع أقسام في السنة الأولى وقسمين في السنة الثانية وخمسة أقسام في السنة الثالثة . أو نقول كذلك أن متوسط السرعة في سباق 100 هو 13.02 ثا لقسم سنة ثانية الفوج 1 هو الأعلى في أقسام السنة الدراسية في تلك المؤسسة .
1. محمد صبحي حسانين : طرق بناء وتقنين الاختبارات والمقاييس في التربية البدنية .ط3.دار الفكر العربي . القاهرة .1987.ص56 [↑](#footnote-ref-2)
2. محمد صبحي حسانين .2001 . مرجع سابق545 .ص 138 [↑](#footnote-ref-3)
3. نفس المرجع ص 141-144 [↑](#footnote-ref-4)
4. محمد حسن علاوي ومحمد نصر الدين رضوان . مرجع سابق.ص 276 [↑](#footnote-ref-5)
5. محمد صبحي حسانين .2001. مرجع سابق ص 145 [↑](#footnote-ref-6)
6. محمد حسن علاوي ومحمد نصر الدين رضوان . مرجع سابق.ص 278 [↑](#footnote-ref-7)
7. محمد حسن علاوي ومحمد نصر الدين رضوان . مرجع سابق.ص 148 [↑](#footnote-ref-8)
8. نفس المرجع ص 297 [↑](#footnote-ref-9)
9. محمد صبحي حسانين 2001. مرجع سابق . ص152 [↑](#footnote-ref-10)
10. أحمد محمد خاطر وعلي فهمي البيك : القياس في المجال الرياضي . دار الكتاب الحديث ط4 . مصر .1996 ص 25 [↑](#footnote-ref-11)
11. محمد حسن علاوي ومحمد نصر الدين رضوان . مرجع سابق.ص 300 [↑](#footnote-ref-12)
12. نفس المرجع ص 300 [↑](#footnote-ref-13)
13. محمد حسن علاوي ومحمد نصر الدين رضوان . مرجع سابق.ص 202 [↑](#footnote-ref-14)