

جامعة محمد الصديق بن يحيى، جيجل
كلية العلوم الإنسانية والاجتماعية
قسم علم النفس وعلوم التربية والأرطفونيا

محاضرات العام الجامعي 2022 - 2023

(السراسي الثالث)

مادة القياس التربوي
وبناء الاختبارات المدرسية (1)
السنة الثانية علوم التربية

د. علوطي سهيلة

أولاً: مدخل إلى القياس التربوي

- 1- تاريخ القياس النفسي والتربوي.
- 2- تعاريف هامة.
- 3- خصائص القياس التربوي.
- 4- أهداف القياس.
- 5- أنواع القياس.
- 6- مراحل عملية القياس.
- 7- مستويات القياس.
- 8- مجالات القياس التربوي.
- 9- أساليب قياس التحصيل الدراسي.
- 10- دور القياس التربوي في نجاح العملية التعليمية -التعلمية.

1- تاريخ القياس النفسي والتربوي:

يمكن تحديد مراحل تطور القياس النفسي بصفة عامة بارتباطه في البداية بمجال التخلف العقلي، ثم بالحربين العالميتين الأولى والثانية:

• القياس والتخلف العقلي:

كان مجال التخلف العقلي من أكثر المجالات التي برزت من خلالها مشكلة القياس والتمييز بين الأفراد، وكان ذلك مبكراً خلال النصف الأول من القرن التاسع عشر.

وكانت من أبرز الإنجازات في ذلك المجال أعمال "جان إسكيرو" وبعده بقليل العمل الرائد الذي قام به "سيجوين". فقد اهتم "إسكيرو" بالفرقة بين التخلف العقلي والمرض العقلي، وتحديد مستويات التخلف العقلي، وكان له الفضل في إدخال محركات موضوعية للتمييز بين درجات التخلف العقلي؛ كاستخدام المقاييس الفيزيائية أبرزها حجم الجمجمة ونسبتها إلى بقية حجم الجسم، ثم تبين له أن المحك المباشر للتمييز بين مستويات التخلف يكمن في استخدام اللغة.

أما الطبيب الفرنسي "سيجوين" فقد ابتدع لوحة أشكال باسمه تستخدم حتى الآن في أدوات القياس السيكولوجي التي يقاس بها مستوى الذكاء، وقد كان يرى أن الضعف العقلي يمكن علاجه فاهتم بتدريب وتنمية المتخلفين عقلياً.

وفي تلك الفترة استخدم العلماء اختبارات تقيس وظائف أكثر تعقيدا منهم "كرابلين" الذي طبق اختبارات تقيس العمليات الحسابية وسعة الذاكرة، وتكميل الجمل، وكذلك الفهم بالجانب الإكلينيكي لمرضى العقول. كذلك قامت بحوث أخرى اهتمت بمنحنيات التعلم، وتجارب عن التعب وقياسه، وانتقال أثر التدريب، والتعب الذاتي، وتشتت الانتباه.

وبذلك تعود أهمية "إسكيروول" و"سيجوين" في تاريخ القياس النفسي الذين أكدا على مشكلة التخلف العقلي وتمييز مستوياته والتي من خلالها طور "بينيه" مقاييس الذكاء حيث قدم في عام 1904 مجموعة من الاختبارات التي تقيس الذكاء، وقام بتعديلها سنة 1908، ثم التعديل الذي كان في سنة وفاته (1911).

• تطور القياس بين الحربين العالميتين الأولى والثانية:

كان لدخول الولايات المتحدة الأمريكية الحرب العالمية الأولى سببا في اتجاه العلماء نحو استخدام اختبارات جديدة وتطوير القديم الصالح منها، حيث طلب من عالم النفس الأمريكي "بيركس" تنظيم مساهمة السيكولوجيين الأمريكيين لتقديم يد المساعدة للجيش لفرز المجندين وفحص الأفراد وانتقاء القادة ورجال المهام الخاصة. وقد استفاد "بيركس" من اختبار الذكاء لـ "بينيه" حيث قام بتعديله وتجميعه مع اختبارات مشابهة. وترتب على تزايد الحاجة لاختبار الأعداد الكبيرة من أفراد الجيش تصميم اختبارات جمعية أيسر وأسهل في التطبيق أشهرها؛ اختبار "ألفا" وهو اختبار لفظي يصلح لمن يجيدون القراءة والكتابة، واختبار "بيتا" وهو اختبار أدائي صمم لقياس القدرات الفطرية التي لم تصقل من خلال التعليم، فهو يصلح لمن لا يجيدون القراءة والكتابة. وهكذا فقد حصل القياس النفسي على الدفعة الرئيسة خلال العامين الأخيرين من الحرب العالمية الأولى مما أدى إلى نموه وتطوره. بعد ذلك خلال فترة السلام بين الحربين كانت أهم معالم التطور في مقاييس الذكاء حيث ظهر التعديل الجديد لمقياس "بينيه" الذي نشره "تيرمان" وزميله "ميرل" سنة 1937، إضافة إلى ظهور مقياس "وكسلر بلفيو" لتدارك عيوب اختبار "بينيه".

• تطور القياس والتقويم في المجال التربوي:

عملية القياس والتقويم التربوي عملية يمكن إرجاعها إلى أقدم العصور البشرية أي إلى عصور ما قبل التاريخ، فالإنسان في العصر الحجري الحديث يقوم بإصدار نوع من الأحكام على زملائه وعشيرته، ولما تعقدت المجتمعات في بداية العصور التاريخية بدأ التقويم على يد معلم الحرفة في المجتمع البدائي حين كان يحاول تقويم المتعلمين على يديه بوضعهم في مواقف عملية تتطلب منهم القيام بأداء عمل معين، ثم يصدر حكمه ويقرر إلى أي مدى نجح "الصبي" في أداء هذا العمل، واستمرت هذه وسيلته في تقويم الأداء، حتى تعقد التراث الإنساني الذي يجب أن يكتسبه الخلف عن السلف، ونشأت المدرسة كمجتمع صناعي يمكن عن طريقه تعليم الصغار بعض ما خلفه السلف للخلف من تراث كبير.

هنا بدأ القياس والتقويم يتخذ صورة أخرى، حيث انتقل إلى مرحلة أخرى هامة في تطوره، ألا وهي المرحلة التي اتجه فيها القائمون على تربية النشء إلى تقويم وقياس نتائج التحصيل الدراسي، ففي المجتمع الصيني القديم عرفت وسائل التقويم التحريرية وعقدت امتحانات تحريرية كانت على درجة كبيرة من الصعوبة، وفي المجتمع اليوناني القديم استخدم نوع من التقويم والقياس التحريري في تقدير نتائج التحصيل الدراسي في هذا المجتمع، وكذلك في المدرسة الرومانية. وتجدر الإشارة إلى أن المجتمع العربي الإسلامي طور نظاما خاصا بالامتحانات وهذا من القرن الثاني هجري (القرن الثامن الميلادي) يدور حول الإجازة أي الشهادة التي يسلمها الشيخ (الأستاذ) إلى طالب العلم. كانت الامتحانات تتم بالنسبة لطلاب العلم في المرحلة العالية أمام لجنة أو لجان من كبار العلماء المعترف بهم في المؤسسات التعليمية (المساجد ودور العلم) وتدوم عدة أيام وهي مفتوحة للعامة. إن نظام الإجازة في طريقة مناقشته للممتحنين يشبه ما يجري حاليا فيما يخص مناقشة رسائل الدراسات العليا.

أما في القرون الوسطى فلم تعرف فيها طريقة تحريرية للتقويم والقياس، عدا مجرد عدد من الأسئلة تلقى شفاهية. واستمر الحال كذلك حتى القرن التاسع عشر لذلك كان تقويم وقياس التحصيل الدراسي في الولايات المتحدة الأمريكية حتى سنة 1850 م.

وفي النصف الثاني من القرن التاسع عشر استبدل الامتحان الشفهي بامتحان تحريري كأساس للالتحاق بالكليات والجامعات. ثم استخدمت هذه الطريقة في مدارس التعليم العام. ثم ظهرت بعد ذلك الامتحانات الموضوعية في أوائل القرن العشرين التي انتشرت بدرجة كبيرة في الوقت الحالي. وفي نقلة نوعية أخرى ظهرت أيضا اختبارات التحصيل المقننة؛ والتي تسمح لنا بمقارنات عديدة كالمقارنة بين المدارس في ولاية واحدة أو على مستوى الوطن.

2- تعاريف هامة:

يستخدم البعض من غير المتخصصين في القياس والتقويم التربوي كلا من مفهوم القياس، والتقويم، والاختبار على نحو تبادلي، وذلك نظرا لتداخل العلاقة الوظيفية بينها، غير أن المتخصصين يؤكدون على أن لكل منها وظيفة محددة في عمليات التقويم، لذا يبدو أنه من الضروري عرض هذه المفاهيم كل على حدة، وتوضيح الفرق بينها حتى يمكن استخدامها على نحو ملائم.

2-1- تعريف القياس:

- يعرف "ج Guilford (1973) القياس بأنه "نسب مقادير عددية (أرقام) إلى الأشياء أو الأحداث وفق قواعد منطقية.

- ويعرفه "كرونباخ" Cronbach (1975) بأنه الطريقة المنظمة لمقارنة سلوك شخصين أو أكثر.

ومنه يمكن القول أن القياس هو عملية تهدف إلى الحصول على قيمة كمية لخاصية معينة بواسطة أداة قياس ملائمة تم إعدادها لتقدير كم هذه الخاصية، الأمر الذي يتيح لنا مقارنة الأفراد في الخاصية أو الصفة موضع القياس.

2-2- تعريف التقويم:

- التقويم هو عملية إصدار حكم على قيمة الأشياء، أو الموضوعات، أو الأفكار بهدف اتخاذ قرارات أو إجراءات عملية بشأنها من حيث تبنيها مع تعديلها، أو تصحيحها وتخليصها من نقاط الضعف فيها، أو رفضها والعدول عنها.
- ويقصد به كذلك الأسلوب العلمي الذي يتم من خلاله تشخيص دقيق للعملية التعليمية وتعديل مسارها، وهو جزء مكمل للعملية التعليمية فهو أحد المؤشرات الهامة للتعرف على مدى كفاءة المناهج وطرائق التدريس في تحقيق الأهداف التربوية المنشودة.

ومما سبق يتضح أن مصطلح "التقويم" يستخدم بصورة أساسية في الميدان التربوي، وهو عملية تشخيصية علاجية تهتم بإصدار أحكام وقرارات بشأن النتائج التي يوفرها القياس في ضوء معايير محددة مسبقا، واتخاذ إجراءات عملية بشأنها.

2-3- تعريف الاختبار:

- يعرف "كرونباخ" Cronbach (1984) الاختبار بأنه إجراء منظم لملاحظة سلوك الفرد ووصفه بمعاونة قياس رقمي أو نظام تصنيفي.
- وترى "أناستازي" Anastasi (1976) بأن الاختبار النفسي هو عبارة عن قياس موضوعي ومقنن لعينة من السلوك.

ومنه يمكن القول أن الاختبار هو نمط معين من أدوات القياس ووسائله، وينطوي على مجموعة من الأسئلة أو العبارات أو المهام التعليمية تختار وتصاغ بطريقة منهجية معينة، بحيث توفر لدى إجابة المفحوص عنها قيمة رقمية لإحدى خصائصه المعرفية؛ كالتحصيل الدراسي أو الذكاء، أو غير المعرفية؛ كالخلفية الاجتماعية أو الاتجاهات أو الميول أو القيم.

2-4- الفرق بين القياس والتقويم:

على الرغم من أن البعض يستخدم المصطلحين (القياس والتقويم) على نحو تبادلي، وذلك بسبب العلاقة الوظيفية بينهما، إلا أنه في الحقيقة هناك اختلافات واضحة بين هذين المصطلحين تتلخص في:

- التقويم عملية شاملة؛ فتقويم المتعلم يمتد ليتناول جميع جوانب النمو والشخصية، وتقويم المنهج يمتد ليتناول جميع جوانب المنهج من أهداف وخطط وبرامج ومقررات وطرق تدريس وأنشطة ووسائل وامتحانات، بينما القياس عملية جزئية تنصب على شيء واحد من الجوانب السابقة.
- يهدف التقويم إلى التشخيص والعلاج، بينما يكتفي القياس بإعطاء بعض المعلومات الكمية المحددة عن الشيء موضوع القياس.
- القياس عملية موضوعية وغير شخصية نسبياً تهتم بتحديد مقدار السمة الذي يمتلكه الفرد أو مقدار تحصيله الدراسي، ومنه فإن درجة الفرد على اختبار ما تمثل القياس، أما التقويم فهو عملية إصدار أحكام أو قرارات بشأن النتائج التي يوفرها القياس في ضوء معايير محددة تحديداً مسبقاً، واتخاذ إجراءات عملية بشأنها، وهو عملية تتصف غالباً بالذاتية.
- القياس سابق للتقويم وأساس له.
- التقويم أكثر شمولاً واتساعاً من القياس.

3- خصائص القياس النفسي والتربوي؛

يتميز القياس النفسي والتربوي بمجموعة من الخصائص نذكر منها:

3-1- القياس النفسي والتربوي قياس غير مباشر؛

فنحن لا نقيس الصفات النفسية على نحو مباشر؛ بمعنى أننا لا نقيس هذه الصفات بالتناول المباشر أو الملاحظة المباشرة؛ بل نستدل عليها من مظاهرها السلوكية الدالة عليها.

فالخواص النفسية والتربوية؛ كالذكاء والتحصيل الدراسي والاتجاهات والدافعية لا يمكن ملاحظتها مباشرة، ولكن يمكن أن نستدل عليها من خلال السلوك الظاهري لها.

3-2- القياس النفسي والتربوي قياس نسبي وليس مطلقاً؛

يصل مستوى القياس في علم النفس وعلوم التربية في أحسن الحالات إلى مستوى المسافات المتساوية ومن النادر أن يصل إلى مستوى النسبة، والصفر في مستوى المسافات المتساوية هو صفر اعتباطي افتراضي وليس صفراً حقيقياً. بمعنى أنه لا يدل على عدم وجود السمة المقاسة بصورة تامة مثل حال القياس في العلوم الفيزيائية التي في الغالب يكون القياس بها على مستوى النسبة، فمثلاً إذا حصل مفحوص على درجة (صفر) في اختبار تحصيلي في مادة الرياضيات لمستوى معين، فإن هذه الدرجة لا تعني أن الطالب ليس لديه معارف على الإطلاق في مادة الرياضيات في هذا المستوى.

3-3- القياس النفسي والتربوي قياس لعينة من السلوك :

فاختبار تحصيلي في مادة ما لا يقيس كل ما حصله التلميذ في هذه المادة، بل عينة مما حصله فيها، فمعد الاختبار التحصيلي لا يقيس كل ما حصله التلميذ في مادته بل يختار عينة ممثلة لجميع الموضوعات التي تتضمنها المادة ويختبره فيها.

3-4- القياس النفسي والتربوي إجراء حيادي :

فالقياس لا يتضمن أحكاماً قيمة مثل جيد أو رديء على الصفات أو الخصائص موضع القياس.

3-5- القياس النفسي والتربوي يتضمن أخطاء معينة :

والتي تتمثل في أخطاء الملاحظة التي تعزى إلى القائم بعملية الملاحظة، وتلك التي ترجع إلى أدوات القياس ذاتها والتي تعزى إلى انخفاض ثباتها أو صدقها نتيجة عدم الاتفاق حول ما تقيسه هذه الأدوات.

3-6- القياس النفسي والتربوي أقل دقة من القياس المادي :

نظراً لأن القياس النفسي قياس غير مباشر للصفة المقيسة، إضافة إلى ما يتضمنه من أخطاء مقارنة بما يتضمنه القياس المادي.

3-7- القياس النفسي والتربوي مجرد وسيلة وليس غاية في حد ذاته :

فهو وسيلة لتطوير العمل وتحسين أداء الفرد؛ كمساعدة المتعلم في تحسين أدائه والمعلم والمرشد النفسي والمدير وغيرهم على تحسين أعمالهم وتطويرها، وبالتالي فهم السلوك الإنساني بصفة عامة.

4- أهداف القياس :

تتلخص أهداف القياس بصفة عامة في الآتي :

➤ الكشف عن الفروق بأنواعها المختلفة، والتي تشمل :

- الفروق بين الأفراد (مقارنة الفرد بغيره من أفراد الجماعة التي ينتمي إليها في ناحية من النواحي النفسية أو التربوية أو المهنية).
- الفروق في ذات الفرد (مقارنة الإمكانات المختلفة في الفرد نفسه).
- الفروق بين المهن المختلفة (تحديد القدرات والاستعدادات والسمات التي تتطلبها المهن المختلفة).
- الفروق بين الجماعات (دراسة سيكولوجية الجماعات المختلفة وخصائص النمو لدى هذه الجماعات ودراسة العوامل المسؤولة عن هذه الفروق).

➤ الوصف الكمي لخصائص أو صفات الأشياء والأشخاص:

فالقياس هو عملية تهدف إلى الحصول على قيمة رقمية أو تقدير كمي لمقدار ما يملكه فرد معين من صفة أو خاصية من الخصائص بمقياس معين وفقا لقواعد معينة، وذلك بهدف معرفة وفهم هذه الخاصية أو الصفة واتخاذ قرار بشأنها.

5-أنواع القياس:

يتميز علماء القياس بين نوعين من القياس وهما:

5-1- قياس العملية:

ويهتم بقياس كيفية أداء الفرد لعمل أو مهارة ما، ويتعلق بالطريقة أو الكيفية التي يؤدي بها الفرد مهارة أو عمل ما ويعالج بها موقفا ما. ويتم في المراحل الأولى من تعلم العمل أو المهارة، ويعد أكثر أهمية من النوع الثاني وذلك في المراحل الأولى من تعلم المهارات النفسحركية.

5-2- قياس الناتج:

ويهتم بقياس مقدار أو نوعية أو جودة العمل أو المهارة بصرف النظر عن كيفية أداء الفرد للعمل أو المهارة؛ فهو يهتم بقياس الناتج النهائي للأداء بغض النظر عن كيفية الأداء، ويتطلب توفر مقياس لقياس مقدار أو جودة الأداء.

6- مراحل عملية القياس:

تتضمن عملية القياس ثلاث مراحل أساسية وهي:

6-1- مرحلة تعريف وتحديد الصفة أو الخاصية موضع القياس:

يتطلب قياس صفة ما تعريف هذه الصفة وتحديد مكوناتها على نحو دقيق، ثم تحديد بعض الأنماط السلوكية الدالة عليها.

6-2- مرحلة تحديد العمليات الإجرائية اللازمة للاستدلال على الصفة أو الخاصية موضع القياس:

ويتم في هذه المرحلة اتخاذ مجموعة من العمليات الإجرائية اللازمة لاستثارة الصفة والاستدلال عليها؛ كتصميم أداة قياس ملائمة تتألف من مجموعة من الأسئلة أو الفقرات أو العبارات أو المهام التي تستثير الصفة، وتستدعي الأنماط السلوكية الدالة عليها بشكل مستمر.

6-3- مرحلة التقدير الكمي للصفة أو الخاصية موضع القياس:

ويتم في هذه المرحلة تحديد مدى ما يتوافر لدى الفرد من الصفة موضوع الاهتمام باستخدام وحدة قياس ملائمة، وتعتبر الدرجة وحدة القياس لأداء الفرد على معظم أدوات قياس الصفات النفسية والتربوية.

7- مستويات القياس:

القياس النفسي والتربوي يمكن أن يوجد على مستويات قياس مختلفة تختلف باختلاف الافتراضات الرياضية والمنطقية التي يؤسس عليها، أو التي يتم توفرها في المقاييس. والتمييز بين هذه المقاييس يتم في ضوء دقتها في قياس السمة أو القدرة موضع القياس.

قدم "ستيفنس" Stevens (1951) تصورا عاما للأنواع أو المستويات المختلفة للقياس، لكل نوع منها قواعده وحدوده الخاصة وطرق إحصائية معينة تناسبه. هذا التصور يقوم على مدى تطبيق العمليات الحسابية المألوفة (الجمع، الطرح، الضرب، القسمة)، وتنظم هذه المستويات في شكل هرمي يبدأ من البسيط إلى المعقد ومن الأدنى إلى الأعلى. ومستويات القياس ذات المستوى الأعلى تتضمن كل الافتراضات المنطقية والرياضية التي ينبغي توفرها في مستويات القياس ذات المستوى الأدنى، بالإضافة إلى الافتراضات الرياضية الخاصة بالمستوى الأعلى نفسه.

وفيما يلي نقدم هذه المستويات من أدناها إلى أكثرها تعقيدا:

1-7- المستوى الاسمي:

يعتبر هذا المستوى أبسط مستويات القياس، وذلك لمحدودية استخدام الأرقام والمعالجات الإحصائية التي يمكن أن نقوم بها، حيث أننا نستخدم الأرقام للدلالة فقط على تسمية الأفراد أو الأشياء، أو لتصنيفها إلى مجموعات (فئات). وهو يعتمد على تصنيف موضوعات القياس إلى فئات تبعا لاشتراكها في خاصية معينة؛ كمكان الإقامة أو لون العينين. ويمكن استخدام الأعداد والنسب المئوية للمفردات في كل فئة؛ كتحديد النسب المئوية لأعداد الطلبة في مدرسة معينة حسب المستوى الدراسي (الأول، الثاني، الثالث).

وتسمى المقاييس المعتمدة عليه بالمقاييس الاسمية، وفيها يستخدم العدد لتصنيف الأحداث أو الأشياء أو الأفراد في فئة أخرى، وفي هذه الحالة فإن العدد المستخدم يشير فقط إلى مسمى الفئة، أي يشير إلى اسم الفئة التي تم فيها التصنيف، فهو مجرد عنوان للفئة ومن أمثلة ذلك: عندما نعطي الذكور رقم (1) والإناث رقم (2) أيضا عندما نصنف الأفراد في فئات وفقا للسن، أو مستوى الخبرة في مجال معين بحيث نشير لكل فئة بعدد معين، وقد نعطي الرقم (2) للعاديين، والرقم (1) لذوي صعوبات التعلم، وقد يستخدم العدد كاسم، ومن أمثلة ذلك: أرقام السيارات في سباق للسيارات، أرقام القوارب في سباق للتجديف، أرقام فريق كرة القدم، أرقام الجلوس في الامتحانات أو أرقام الشوارع. (محمود أحمد عمر وآخرون، 2010، ص 83)

وطبيعة هذا المستوى أنه يسمح فقط باستخدام عملية العد، ولا يمكن إجراء أي من العمليات الحسابية الأربعة (الجمع، الطرح، الضرب، القسمة) في معالجتنا للأرقام، وذلك لأنها لا تتضمن أي تقدير ما عدا خاصية التصنيف. كما يمكن القيام بحساب التكرارات والنسب المئوية، والمونال وبعض المقاييس اللابرامترية مثل مربع كاي.

7-2- المستوى الرتبي :

وهو يرتب الأشخاص أو الأشياء تصاعديا أو تنازليا على سلم متدرج، فيكون (س) مثلا هو الأول و(ص) هو الثاني و(ع) هو الثالث وغيرها، ولكن قيمة هذه الأرقام هو في ترتيبها وليس في الكم الذي تحتوي عليه. فالوحدات في مقاييس الرتب لا تكون متساوية في الحقيقة، ولو بدت من حيث الأرقام الدالة عليها وكأنها متساوية.

في هذا المستوى يعبر عن المتغير برتب؛ بحيث ترتب القياسات تصاعديا أو تنازليا، وفيه تؤدي الأرقام وظيفة التمييز لكنها أكثر دقة منها في المستوى الاسمي؛ فهي تعطي فكرة عن موقع الفرد بالنسبة إلى بقية الأفراد.

وتسمى المقاييس المعتمدة عليه بمقاييس الرتبة وهو مستوى أرقى من مستوى المقاييس الاسمية. وتعتمد مقاييس الرتبة على ترتيب مجموعة من الأفراد وفقا لدرجاتهم في صفة من الصفات دون أن تعطى أي أهمية لتساوي الفروق بين الرتب. القياس الرتبي لا يسمح لنا بتحديد الفروق بين أي ربتين بدقة.

مع العلم أنه لا يمكن إجراء العمليات الحسابية الأربعة الأساسية المعروفة، ويمكن فقط حساب النسب المئوية والمنوال والوسيط. كما يمكن التعامل مع الرتب باستخدام معامل ارتباط الرتب والذي يعتمد في جوهره على ترتيب الأفراد في الصفة المقيسة، وتعتبر المئينيات من مقاييس الرتبة.

7-3- المستوى الفتري :

يمكن أن نطلق عليه مقياس الوحدة المنتظمة أو مقياس المسافة، وتمتد مقاييس المسافة بفكرة ترتيب الرتب لتشمل على مفهوم المسافات المتساوية بين الأحداث التي تم ترتيبها.

يعبر عن المتغير في هذا المستوى بقيم عددية، ويفترض أن المسافة بين القيمة والقيمة التي تليها متساوية. كما أن الصفر فيه غير حقيقي بل هو افتراضي؛ أي أنه لا يعبر عن غياب الظاهرة. فمثلا الطالب الذي يحصل على درجة الصفر في الإحصاء لا يعني هذا أن هذا الطالب ليست له معلومات عن وحدة الإحصاء المدرسة.

وما يميز هذا المستوى عن المستوى الترتيبي هو أنه يمكننا من معرفة مقدار المسافة بين شخصين أو شيئين في الخاصية المقيسة، ويعد أكثر دقة من المستويين السابقين، حيث تتساوى الفروق بين الأقسام المتتالية في السمة موضوع القياس. العمليات الحسابية المسموح بها في هذا المستوى من القياس هي عمليتا الجمع والطرح، دون الضرب والقسمة ويرجع ذلك إلى عدم وجود صفر مطلق في هذا المقياس.

7-4- المستوى النسبي :

يعتبر المستوى النسبي الأعلى موضوعية ودقة، وحسب ما تدل تسميته أنه يمكننا تقسيم عدد على آخر وأن نعبر عنهما في صيغة نسبة.

ويمتاز بأن له وحدات متساوية وصفرًا مطلقًا كما في مقياس الطول والوزن والحجم، وهو يستخدم على نطاق واسع في ميدان العلوم الطبيعية، بينما ينذر استعماله في العلوم التربوية والنفسية.

ونظرًا لتوافر جميع خصائص مقياس المسافة إضافة إلى الصفر المطلق في هذه المقاييس، يمكن استخدام كل العمليات الحسابية من جمع وطرح وضرب وقسمة مع هذه المقاييس، كما يمكن استخدام كل العمليات الرياضية العليا وكل الطرق الإحصائية الممكنة.

8- مجالات القياس والتقويم التربوي؛

يمكن إيجاز مجالات القياس والتقويم التربوي في المجال التربوي فيما يلي:

8.1- **التعلم:** ويتضمن تقدير درجات المتعلمين التحصيلية، وفي بقية نواتج التعلم وتحديد معدلات التعلم ومستويات الأداء المهاري للمتعلمين.

8.2- **التعليم:** ويشتمل على قياس وتقدير كفاءة التدريس ودرجة جودته وفعاليته في تحقيق الأهداف التعليمية المسطرة.

8.3- **المقررات:** ويشتمل على تقويم محتوى المقرر الدراسي وأساليب التعليم والتعلم، واختيار الطرق العلاجية المناسبة للتغلب على صعوبات تعلم موضوعات المقرر.

8.4- **البرامج:** ويتضمن قياس درجة فعالية برنامج تعليمي أو منهج دراسي ودرجة كفاءته، ومدى ملاءمته لتحقيق الأهداف التربوية في ضوء مستوى الإدارة المدرسية وفعالية التدريس.

8.5- **المؤسسات والنظم التعليمية:** ويشتمل على قياس مدى كفاءة البرامج المتعددة الجوانب، أو تقويم النظم التعليمية وتقويم المؤسسات التعليمية مثل الجامعات والمدارس.

9- أساليب قياس التحصيل الدراسي؛

من أساليب قياس التحصيل الدراسي نجد الاختبارات التحصيلية؛ والتي يمكن تقسيمها إلى ثلاثة أقسام رئيسة وهي:

9.1- **الاختبارات الشفوية:** ويقصد بها مجموعة الأسئلة التي تعطى للطالب دون استخدام الكتابة لقياس خبرة التلميذ في الموضوعات التي سبق أن تعلمها.

9.2- **الاختبارات التحريرية:** يقصد منها تقدير التحصيل الدراسي للتلاميذ باستخدام الكتابة؛ وتنقسم إلى اختبارات المقال والاختبارات الموضوعية.

9.3- **الاختبارات العملية:** وهي اختبارات ذات طابع عملي تتطلب تقدير الأداء والممارسة كقياس الأداء اللغوي في المختبر، والأداء الرياضي، والأداء في العلوم الطبيعية. وغالبا ما يستخدم لقياس مدى فهم الطلاب في الدراسة النظرية ومدى فاعليتها.

10- دور القياس التربوي في نجاح العملية التعليمية - التعليمية:

يلعب القياس التربوي دورا هاما في مجال التربية والتعليم وهذا الدور يجعل من العملية التعليمية-التعليمية ناجحة؛ إذ يستعمل للأغراض الآتية:

- معرفة مدى ما تحقق من الأهداف التعليمية.
- تزويد المعلم والمتعلم بتغذية راجعة عن مدى نجاح الأساليب المتبعة في المذاكرة والتحضير والمشاركة الصفية، وعن مدى كفاءة المواد الدراسية وأساليب التدريس والنشاطات الصفية في تحقيق الأهداف التعليمية.
- التشخيص والعلاج وتحسين مستوى التعلم، وتدعيم جوانب القوة ومعالجة جوانب الضعف لدى المتعلمين.
- الكشف عن المتخلفين في الدراسة، ومعرفة أسباب هذا التخلف.
- اختيار الطلبة الجدد، واختيار المواد الدراسية التي تناسب أعمارا معينة.
- توجيه التلاميذ إلى نوع الدراسة التي تناسب قدراتهم واستعداداتهم.
- إرشاد التلاميذ الذين يجدون مشاكل في دراستهم والذين يتعرضون لبعض الأزمات، واستعمال الاختبارات لمعرفة بعض الأعراض المرضية والأسباب التي أدت إلى هذه الأزمات.
- تقويم تحصيل التلاميذ ومدى فعالية برامج التعليم ومناهجه.
- تصنيف التلاميذ حسب قدراتهم إلى مستويات؛ بحيث يكون كل مستوى متجانسا من حيث القدرات والاستعدادات، الأمر الذي يساعد المؤسسة التعليمية على أداء مهمتها بنجاح.
- فحص مناهج التعليم المختلفة لتشخيص نقاط الضعف وعناصر القوة، باستخدام الأساليب الكفيلة لذلك.
- تقويم أداء المعلم للتأكد من فاعلية تنظيمه للبيئة التعليمية، بما يحقق التغيرات المرجوة في سلوك المتعلمين.

ثانياً: ثبات الاختبار

- 1- تعريف ثبات الاختبار.
- 2- طرق حساب ثبات الاختبار.
- 3- العوامل المؤثرة في ثبات الاختبار
- 4- تفسير معامل الثبات.
- 5- مستويات الثبات المقبولة.

1- تعريف ثبات الاختبار:

- يقصد بثبات الاختبار اتساق أداء الأفراد عبر الزمن إذا ما طبق عليهم الاختبار أكثر من مرة.
 - ويقصد به كذلك استقرار أداء الأفراد عبر صور متكافئة من الاختبار.
 - ويعرف أيضاً بأنه النسبة بين التباين الحقيقي إلى التباين المشاهد (التباين الكلي) لدرجات الاختبار.
- فالثبات يشير إلى استقرار النتائج التي تعطيها أداة القياس في مرات التطبيق المختلفة، وهو شرط ضروري للحكم على صلاحية هذه الأداة، غير أنه شرط غير كاف يستلزم خصائص أخرى كالصدق.

2- طرق حساب ثبات الاختبار:

توجد عدة طرق لحساب ثبات الاختبار نذكر منها:

أ- طريقة إعادة الاختبار:

وتقوم هذه الطريقة على تطبيق الاختبار على عينة من الأفراد (عينة التقنين)، ثم يعاد تطبيق الاختبار نفسه على العينة نفسها مرة أخرى بعد فترة زمنية لا تقل عن أسبوع (أو أسبوعين) ولا تزيد عن ستة أشهر، ثم يحسب معامل ثبات الاختبار من خلال حساب معامل الارتباط بين درجات الأفراد في التطبيق الأول ودرجاتهم في التطبيق الثاني، وذلك باستخدام إحدى صور معادلة معامل الارتباط البسيط عند "بيرسون".

ويستحسن استخدام المعادلة الآتية:

$$r = \frac{n (\sum xy) - (\sum x) (\sum y)}{\sqrt{[n \sum x^2 - (\sum x)^2] [n \sum y^2 - (\sum y)^2]}}$$

بحيث:

r : الارتباط بين درجات الأفراد في الاختبار في التطبيق الأول ودرجاتهم في التطبيق الثاني.

x : درجات الأفراد في الاختبار في التطبيق الأول.

y : درجات الأفراد في الاختبار في التطبيق الثاني.

n : عدد أفراد العينة.

يطلق على معامل الارتباط (r) في هذه الحالة بمعامل الاستقرار (تقدير لاتساق أداء الأفراد عبر الزمن).

مثال:

قام باحث ببناء اختبار تحصيلي جديد، ولغرض التأكد من ثباته قام بتطبيقه في 01 مارس على عينة مؤلفة من (05) طلاب، ثم أعاد تطبيقه على العينة نفسها في 01 أبريل. وقد حصل على البيانات الموضحة في الجدول أدناه.

المطلوب:

- وضح الطريقة التي استخدمها الباحث للتأكد من ثبات اختباره؟ وكيف يسمى معامل الثبات في هذه الطريقة؟
- احسب معامل ثبات هذا الاختبار.
- حدد نسبة التباين الحقيقي ونسبة تباين الخطأ في قيمة معامل الثبات الذي حصلت عليه. ماذا تستنتج؟

جدول يوضح درجات الطلبة في الاختبار في التطبيقين (01 مارس-01 أبريل)

رقم الطالب	الدرجات على الاختبار في 01 مارس	الدرجات على الاختبار في 01 أبريل
01	16	14
02	09	10
03	14	14
04	11	12
05	07	05

الحل:

- لحساب معامل الثبات استخدم الباحث طريقة إعادة الاختبار؛ والتي تقوم على تطبيق الاختبار على عينة من الأفراد (عينة التقنين)، ثم يعاد تطبيق الاختبار نفسه على العينة نفسها مرة أخرى بعد فترة زمنية لا تقل عن أسبوعين ولا تزيد عن ستة أشهر، ثم يحسب معامل الثبات من خلال حساب معامل الارتباط بين درجات الأفراد في التطبيقين الأول والثاني.

- يسمى معامل الثبات في هذه الطريقة بمعامل الاستقرار.

- حساب معامل ثبات الاختبار:

يحسب معامل الثبات من خلال حساب معامل الارتباط لـ "بيرسون" بين درجات الطلبة في التطبيق الأول ودرجاتهم في التطبيق الثاني باستخدام المعادلة الآتية:

$$r = \frac{n (\sum x y) - (\sum x) (\sum y)}{\sqrt{[n \sum x^2 - (\sum x)^2] [n \sum y^2 - (\sum y)^2]}}$$

رقم الطالب	الدرجات في التطبيق الأول (x)	الدرجات في التطبيق الثاني (y)	x ²	y ²	x y
01	16	14	256	196	224
02	09	10	81	100	90
03	14	14	196	196	196
04	11	12	121	144	132
05	07	05	49	25	35
المجموع	57	55	703	661	677

بالتعويض في المعادلة نجد:

$$r = \frac{(5 \times 677) - (57 \times 55)}{\sqrt{[5 \times 703 - (57)^2] [5 \times 661 - (55)^2]}} = \frac{3385 - 3135}{\sqrt{(3515 - 3249) (3305 - 3025)}}$$

$$= \frac{250}{\sqrt{266 \times 280}} = \frac{250}{\sqrt{74480}} = \frac{250}{272,91} = \underline{\underline{0,91}}$$

ومنه فإن قيمة معامل الثبات تساوي: **0,91**

- نسبة التباين الحقيقي هي: **91%**
- نسبة تباين الخطأ هي: **9%**

الاستنتاج: نستنتج أن معامل الثبات مرتفع لأن قيمة تباين الخطأ صغيرة مقارنة بقيمة التباين الحقيقي.

ب- طريقة الصور المتكافئة:

تقوم على إعداد صورتين متكافئتين من الاختبار، ثم تطبق هاتان الصورتان بحيث يكون بينهما فترة وجيزة على المجموعة نفسها من الأفراد، ثم يقدر معامل الثبات بحساب معامل الارتباط بين درجات الأفراد في الصورتين المتكافئتين. ويطلق على معامل الارتباط في هذه الحالة بمعامل التكافؤ.

تقوم هذه الطريقة على افتراض تكافؤ أو تساوي الصورتين في الآتي:

- عدد المكونات الوظيفية التي يقيسها الاختبار، ونسبة أو عدد المفردات التي تخص كل منها.
- تساوي مستوى صعوبة المفردات.
- تساوي طول الاختبار.
- تساوي متوسط وتباين درجات الأفراد.
- تماثل طريقة الإجابة عن الاختبار.
- تماثل طريقة التصحيح عن الاختبار.
- تساوي زمن تطبيق الاختبار.
- تكافؤ التعليمات.
- تكافؤ في الهدف والبنية.

ج- طريقة الاستقرار والتكافؤ:

تقوم هذه الطريقة على إعداد صورتين متكافئتين تماما من الاختبار، ثم تطبق هاتان الصورتان على مجموعة من الأفراد بحيث يكون بينهما فترة طويلة نسبيا (لا تقل عن أسبوعين ولا تزيد عن ستة أشهر)، ثم يقدر معامل الثبات بحساب معامل الارتباط بين درجات الأفراد في الصورتين المتكافئتين. ويطلق على معامل الارتباط في هذه الحالة بمعامل الاستقرار والتكافؤ.

د- طريقة التجزئة النصفية:

تقوم هذه الطريقة في تقدير معامل الثبات على فكرة تطبيق الاختبار مرة واحدة على عينة من الأفراد، ثم تقسيم مفردات الاختبار إلى نصفين متساويين، ومن وسائل ذلك:

- تقسيم المفردات إلى مفردات فردية ومفردات زوجية.
- تقسيم المفردات إلى جزأين متماثلين (النصف الأول والنصف الثاني).

ثم يحسب معامل الارتباط بين درجات الأفراد على نصفي الاختبار، وينتج معامل ثبات الأداء على نصف الاختبار

(معامل ثبات التجزئة النصفية). وهنا ينبغي استخدام معادلات تصحيح الطول، ومن أشهرها معادلة "سبيرمان- براون"

(Spearman-Brown)، وتقوم على افتراض أنه كلما زاد حجم عينة الأداء على الاختبار ارتفع معامل الثبات، وافترض تساوي تباين نصفي الاختبار.

ومعادلتها الرياضية على النحو الآتي:

$$r_{tt} = \frac{2 (r_{hh})}{1 + r_{hh}}$$

بحيث:

r_{tt} : معامل ثبات الاختبار بعد التصحيح.

r_{hh} : معامل ثبات نصف الاختبار.

ولقد اقترح "جوتمان" (Guttman) معادلة بديلة لحساب ثبات الاختبار في حالة التجزئة النصفية لا تقوم على افتراض تساوي تباين نصفي الاختبار، والصيغة الرياضية لهذه المعادلة هي على النحو الآتي:

$$r_{tt} = 2 \left[1 - \frac{S_1^2 + S_2^2}{S_t^2} \right]$$

بحيث:

$$S = 1/n \sqrt{n \sum x^2 - (\sum x)^2}$$

r_{tt} : معامل ثبات الاختبار ككل.

S_1^2 : تباين درجات الأفراد على النصف الأول من الاختبار (النصف الفردي).

S_2^2 : تباين درجات الأفراد على النصف الثاني من الاختبار (النصف الزوجي).

S_t^2 : تباين درجات الأفراد على الاختبار ككل.

ومن أهم مزايا هذه الطريقة أنها لا تتطلب حساب معامل الارتباط بين نصفي الاختبار، بالإضافة إلى أنها لا تتطلب تساوي تباين نصفي الاختبار.

مثال:

قام باحث ببناء اختبار تحصيلي جديد، ولغرض التأكد من ثباته قام بتطبيقه على عينة مؤلفة من (05) طلاب، ثم قام بحساب مجموع درجات كل طالب على المفردات الفردية والمفردات الزوجية. وقد حصل على البيانات الموضحة في الجدول أدناه.

المطلوب:

- وضح الطريقة التي استخدمها الباحث للتأكد من ثبات اختبار.
- احسب معامل ثبات هذا الاختبار.
- حدد نسبة التباين الحقيقي ونسبة تباين الخطأ في قيمة معامل الثبات الذي حصلت عليه. ماذا تستنتج؟

جدول يوضح درجات الطلبة على المفردات الفردية والزوجية

رقم الطالب	الدرجات على المفردات الفردية	الدرجات على المفردات الزوجية
01	04	05
02	07	08
03	09	08
04	06	06
05	05	05

الحل:

- الطريقة التي استخدمها الباحث للتحقق من ثبات اختبار هي طريقة التجزئة النصفية؛ والتي تقوم على تطبيق الاختبار مرة واحدة على أفراد عينة التقنين، ثم تقسيم مفردات الاختبار إلى نصفين متساويين (فردية وزوجية)، ثم حساب معامل الارتباط بين درجات الأفراد على نصفي الاختبار واستخدام معادلة التصحيح لـ "سبيرمان براون"، أو تطبيق معادلة "جوتمان" مباشرة والتي لا تتطلب حساب معامل الارتباط بين نصفي الاختبار ولا تشترط تساوي تباين نصفي الاختبار.

- حساب معامل ثبات الاختبار:

نقوم بحساب معامل ثبات هذا الاختبار بطريقتين:

1- باستخدام معادلة "سبيرمان- براون" (التي تقوم على افتراض تساوي تباين نصفي الاختبار):

نقوم أولاً بحساب معامل الارتباط لـ "بيرسون" بين درجات الأفراد على المفردات الفردية ودرجاتهم على المفردات

الزوجية باستخدام المعادلة الآتية:

$$r_{hh} = \frac{n (\sum xy) - [(\sum x) (\sum y)]}{\sqrt{[n \sum x^2 - (\sum x)^2] [n \sum y^2 - (\sum y)^2]}}$$

رقم الطالب	الدرجات على المفردات الفردية (x)	الدرجات على المفردات الفردية (y)	x ²	y ²	x y
01	04	05	16	25	20
02	07	08	49	64	56
03	09	08	81	64	72
04	06	06	36	36	36
05	05	05	25	25	25
المجموع	31	32	207	214	209

بالتعويض في المعادلة نجد:

$$r_{hh} = \frac{(5 \times 209) - (31 \times 32)}{\sqrt{[5 \times 207 - (31)^2][5 \times 214 - (32)^2]}} = \frac{1045 - 992}{\sqrt{(1035 - 961)(1070 - 1024)}}$$

$$= \frac{53}{\sqrt{74 \times 46}} = \frac{53}{\sqrt{3404}} = \frac{53}{58,34} = \underline{0,90}$$

باستخدام معادلة التصحيح لـ "سبيرمان-براون" ومعادلتها:

$$r_{tt} = \frac{2 (r_{hh})}{1 + r_{hh}}$$

بالتعويض في المعادلة نجد:

$$r_{tt} = \frac{2 \times 0,90}{1 + 0,90} = \frac{1,8}{1,90} = \underline{0,94}$$

1- باستخدام معادلة "جوتمان" التي لا تتطلب تساوي تباين نصف الاختبار ولا حساب معامل الارتباط بين نصفين

الاختبار ومعادلتها:

$$r_{tt} = 2 \left[1 - \frac{S_1^2 + S_2^2}{S_t^2} \right]$$

$$S = 1/n \sqrt{n \sum x^2 - (\sum x)^2}$$

ومعادلة الانحراف المعياري:

رقم الطالب	الدرجات على المفردات الفردية (x ₁)	الدرجات على المفردات الفردية (x ₂)	x _t	x ₁ ²	X ₂ ²	x _t ²
01	04	05	09	16	25	81
02	07	08	15	49	64	225
03	09	08	17	81	64	289
04	06	06	12	36	36	144
05	05	05	10	25	25	100
المجموع	31	32	63	207	214	839

نقوم أولاً بحساب: S₁² ، S₂² ، S_t².

$$S_1 = \frac{1}{n} \sqrt{n \sum x_1^2 - (\sum x_1)^2} = \frac{1}{5} \sqrt{(5 \times 207) - (31)^2} = \frac{1}{5} \sqrt{1035 - 961}$$

$$= \frac{\sqrt{74}}{5}$$

ومنه:

$$S_1^2 = \left(\frac{\sqrt{74}}{5} \right)^2 = \frac{74}{25} = \underline{\underline{2,96}}$$

$$S_2 = \frac{1}{n} \sqrt{n \sum x_2^2 - (\sum x_2)^2} = \frac{1}{5} \sqrt{(5 \times 214) - (32)^2} = \frac{1}{5} \sqrt{1070 - 1024}$$

$$= \frac{\sqrt{46}}{5}$$

ومنه:

$$S_2^2 = \left(\frac{\sqrt{46}}{5} \right)^2 = \frac{46}{25} = \underline{\underline{1,84}}$$

$$S_t = \frac{1}{n} \sqrt{n \sum x_t^2 - (\sum x_t)^2} = \frac{1}{5} \sqrt{(5 \times 839) - (63)^2} = \frac{1}{5} \sqrt{4195 - 3969}$$

$$= \frac{\sqrt{226}}{5}$$

ومنه:

$$S_t^2 = \left(\frac{\sqrt{226}}{5} \right)^2 = \frac{226}{25} = \underline{9,04}$$

بالتعويض في معادلة "جوتمان" نجد:

$$r_{tt} = 2 \left[1 - \frac{2,96 + 1,84}{9,04} \right] = 2 \left(1 - \frac{4,80}{9,04} \right) = 2 (1 - 0,53) = 2 \times 0,47$$

$$= \boxed{0,94}$$

فإن قيمة معامل الثبات تساوي: $\boxed{0,94}$

- نسبة التباين الحقيقي هي: $\boxed{94\%}$
- نسبة تباين الخطأ هي: $\boxed{6\%}$

الاستنتاج: نستنتج أن معامل الثبات مرتفع لأن قيمة تباين الخطأ صغيرة مقارنة بقيمة التباين الحقيقي.

ملاحظة: قيمة معامل الثبات هي نفسها التي حصلنا عليها باستخدام معادلة "سبيرمان - براون".

هـ- معامل ثبات الاتساق الداخلي:

يقصد بالاتساق الداخلي مدى ارتباط مفردات الاختبار ببعضها البعض، أو مدى قياس المفردات للسمة نفسها، كما يعرف معامل ثبات الاتساق الداخلي بأنه متوسط مجموع تقديرات معاملات ثبات التجزئة النصفية الممكنة للاختبار نفسه، ومن المعادلات المستخدمة في هذه الطريقة معادلة "كيودر وريثشاردسون" ومعادلة "ألفا كرونباخ".

أولاً: معادلة "كيودر وريتشاردسون":

يمكن تقدير معامل الاتساق الداخلي باستخدام معادلة "كيودر وريتشاردسون" ولها صيغ رياضية عديدة غير أنه هناك صيغتان أكثر شيوعاً واستخداماً وهما الصيغتان (20) و (21) وسيتم تناول الصيغة (20) فقط لأنها الأكثر شيوعاً واستخداماً مقارنة بالأخرى كما أنها تتميز بسهولة عملياتها الحسابية، وصيغتها الرياضية على النحو الآتي:

$$KR_{20} = \frac{K}{K-1} \left(1 - \frac{\sum p q}{S_t^2} \right)$$

بحيث:

$$S = 1/n \sqrt{n \sum x^2 - (\sum x)^2}$$

KR_{20} : معامل ثبات الاتساق الداخلي (معامل التجانس).

K : عدد مفردات الاختبار.

P : نسبة من أجابوا إجابة صحيحة عن المفردة.

q : نسبة من أجابوا إجابة خاطئة عن المفردة.

S_t^2 : التباين الكلي المشاهد لدرجات الأفراد على الاختبار.

ملاحظة: تستخدم هذه الصيغة إذا كانت درجات المفردات ثنائية (0، 1) فقط.

ثانياً: معامل "ألفا كرونباخ":

اقترح "كرونباخ" معادلة لحساب الاتساق الداخلي للاختبار، وهي في حقيقة الأمر عبارة عن تعميم لمعادلة "كيودر وريتشاردسون"، ومن ثم فقد أطلق على هذه المعادلة بالمعادلة العامة للاتساق الداخلي، وتستخدم هذه المعادلة في حالة:

- تصحيح الاختبار ب (صح) أو (خطأ) أو على نحو ثنائي بحيث تمنح الإجابة الصحيحة واحداً صحيحاً وتمنح الإجابة الخاطئة صفراً.

- حساب الاتساق الداخلي للاختبار التي لا تستخدم التقسيم الثنائي (صح أو خطأ) بل تستخدم التقديرات المتصلة

كما في حالة مقاييس الاتجاهات التي تؤسس على طريقة "ليكرت" حيث تتراوح قيمة درجة الفرد على المفردة

في إطار مدى معين من القيم، ففي حالة المقياس الخماسي تتراوح درجة الفرد على المفردة بين (1) و (5)؛ بحيث

تعطى الدرجات (1، 2، 3، 4، 5) للاستجابات (معارض جداً، معارض، محايد، موافق، موافق جداً) وذلك على التوالي.

والصيغة الرياضية لهذه المعادلة هي على النحو الآتي:

$$\alpha = \frac{K}{K-1} \left(1 - \frac{\sum S_i^2}{S_t^2} \right)$$

بحيث:

$$S = 1/n \sqrt{n \sum x^2 - (\sum x)^2}$$

S_i^2 : تباين درجات الأفراد على المفردة الواحدة.

$\sum S_i^2$: مجموع تباين درجات الأفراد على المفردات.

S_t^2 : التباين الكلي المشاهد لدرجات الأفراد على الاختبار ككل.

K : عدد مفردات الاختبار.

مثال:

قام باحث بتطبيق اختبار يتألف من (10) مفردات، الإجابة عنها بصح أو خطأ. وذلك على عينة تتألف من (10) أفراد، وكانت درجاتهم على مفردات المقياس موضحة في الجدول أدناه.

المطلوب: احسب معامل الاتساق الداخلي لهذا الاختبار. ماذا تستنتج؟

جدول يوضح درجات الأفراد على الاختبار

درجات الطلاب على المفردات										أرقام الطلاب
10 ف	9 ف	8 ف	7 ف	6 ف	5 ف	4 ف	3 ف	2 ف	1 ف	
1	1	1	1	1	1	1	1	1	1	1
0	1	1	1	1	1	1	1	1	1	2
0	1	0	1	1	1	1	1	1	1	3
0	0	1	1	1	1	1	1	1	1	4
1	0	0	1	0	1	1	1	1	1	5
0	0	1	0	1	1	1	1	1	1	6
0	0	1	0	0	1	1	1	1	1	7
0	1	0	0	0	1	1	1	1	1	8
0	0	0	0	1	0	1	1	1	1	9
0	0	0	0	1	0	0	1	1	1	10

الحل:

بما أن الإجابة على مفردات هذا الاختبار ب(صح) أو (خطأ) أي على نحو ثنائي بحيث تمنح الإجابة الصحيحة واحدا صحيحا وتمنح الإجابة الخاطئة صفرا، فإن المعادلة المناسبة لحساب الاتساق الداخلي هي معادلة "كيودورويتشاردسون"، حيث تسير خطوات الحل كالآتي:

1. حساب الدرجة الخام (X_t) لكل طالب على حدة، ثم تربيعها (X_t^2).
2. حساب مجموع من أجابوا إجابة صحيحة عن كل مفردة.
3. حساب نسبة من أجابوا إجابة صحيحة عن كل مفردة (p)، ونسبة من أجابوا إجابة خاطئة (q) عن كل مفردة من مفردات الاختبار.
4. حساب تباين كل مفردة (pq).
5. حساب مجموع تباينات المفردات ($\sum pq$).
6. حساب التباين (S_t^2) لدرجات الطلاب على الاختبار.

وسيأتي تطبيق الخطوات الخمسة الأولى في الجدول الآتي:

أرقام الطلاب	درجات الطلاب على المفردات											الدرجة على الاختبار X_t	مربع الدرجة على الاختبار X_t^2
	ف1	ف2	ف3	ف4	ف5	ف6	ف7	ف8	ف9	ف10			
1	1	1	1	1	1	1	1	1	1	1	10	100	
2	1	1	1	1	1	1	1	1	1	0	9	81	
3	1	1	1	1	1	1	1	0	1	0	8	64	
4	1	1	1	1	1	1	1	1	0	0	8	64	
5	1	1	1	1	1	0	1	0	0	1	7	49	
6	1	1	1	1	1	1	0	1	0	0	7	49	
7	1	1	1	1	1	0	0	1	0	0	6	36	
8	1	1	1	1	1	0	0	0	1	0	6	36	
9	1	1	1	1	0	1	0	0	0	0	5	25	
10	1	1	1	0	0	1	0	0	0	0	4	16	
Σ	10	10	10	9	8	7	5	5	4	2	70	520	
P	1	1	1	0,9	0,8	0,7	0,5	0,5	0,4	0,2			
Q	0	0	0	0,1	0,2	0,3	0,5	0,5	0,6	0,8			
P q	0	0	0	0,09	0,16	0,21	0,25	0,25	0,24	0,16			
Pq Σ	0+ 0+ 0+ 0,09+ 0,16+ 0,21+ 0,25+ 0,25+ 0,24+ 0,16= <u>1,36</u>												

- حساب التباين الكلي درجات الطلبة على الاختبار (S_t^2):

لدينا:

$$S_t = \frac{1}{n} \sqrt{n \sum x_t^2 - (\sum x_t)^2} = \frac{1}{10} \sqrt{(10 \times 520) - (70)^2}$$

$$= \frac{1}{10} \sqrt{5200 - 4900} = \frac{\sqrt{300}}{10} = \frac{10\sqrt{3}}{10} = \underline{\sqrt{3}}$$

$$S_t^2 = (\sqrt{3})^2 = \underline{3}$$

بالتعويض في معادلة "كيودورويتشاردسون" نجد:

$$KR_{20} = \frac{K}{K-1} \left(1 - \frac{\sum p q}{S_t^2} \right) = \frac{10}{10-1} \left(1 - \frac{1,36}{3} \right)$$

$$= \frac{10}{9} (1 - 0,45) = 1,11 \times 0,55 = \underline{0,61}$$

الاستنتاج:

بلغت قيمة معامل الاتساق الداخلي (0.61) وهي قيمة متوسطة وتشير إلى أن (61٪) من تباين درجات الأفراد على الاختبار هو تباين حقيقي و(39 ٪) هي تباين الخطأ، وهو معامل ثبات غير مقبول في ضوء محركات الثبات المقبولة.

3-العوامل التي تؤثر في ثبات الاختبار:

توجد عدة عوامل تؤثر في ثبات الاختبار منها:

أ. طول الاختبار:

زيادة طول الاختبار تؤدي إلى ارتفاع معامل ثباته وذلك لأن طول الاختبار معناها زيادة عدد مفرداته، الأمر الذي يترتب عليه الحصول على عينة أكبر من السلوك تسمح بأن يمثل فيها بشكل مستقر العدد الأكبر من مكونات السلوك أو الصفة المقيسة.

ب. صعوبة الاختبار:

ارتفاع أو انخفاض مستوى صعوبة الاختبار بصفة عامة تؤدي إلى التقليل من مدى الفروق الفردية، وبالتالي يقل ثبات الاختبار. ويشير ذلك إلى أنه لكي نحصل على اختبار يتمتع بمعامل ثبات عال يجب أن يبلغ معامل صعوبة الاختبار (0,50) تقريباً.

ج. استقلال مفردات الاختبار:

ينبغي أن تكون المفردات مستقلة، حتى وإن كانت تقيس سمة واحدة أو قدرة واحدة.

د. تشابه محتوى مفردات الاختبار:

يؤدي زيادة تشابه محتوى مفردات الاختبار إلى زيادة قيمة معامل ثباته.

هـ. التخمين:

يعد التخمين من العوامل التي تؤثر في ثبات الاختبار على نحو سلبي.

و- زمن الاختبار:

يؤثر زمن الاختبار على ثباته على نحو سلبي عندما يكون الاختبار شديد الصعوبة والزمن اللازم للإجابة عن مفرداته غير كاف، الأمر الذي يدفع المفحوصين إلى التخمين، وهذا يترتب عنه انخفاض قيمة معامل ثبات الاختبار. كما نجد أن الاختبارات الموقوتة (اختبارات السرعة) تكون معاملات ثباتها مرتفعاً أكثر من الاختبارات التي تعطي متسعاً من الوقت (يكون فيها مدى الفروق ضيقاً).

ز- تجانس عينة التقنين:

يقل ثبات الاختبار بزيادة تجانس المجموعة المطبق عليها، ويرتفع الثبات بزيادة تباين درجات أفرادها عليها.

4- تفسير معامل ثبات الاختبار:

يشير معامل الثبات على نحو مباشر إلى نسبة التباين المشاهد الذي يعزى إلى نسبة التباين الحقيقي في درجات الأفراد على الاختبار، ونسبة تعزى إلى أخطاء القياس.

مثال: إذا بلغت قيمة معامل ثبات الاختبار (0,91) فهذا معناه أن (91٪) من التباين المشاهد تعزى إلى الفروق في الدرجات الحقيقية وأن (9٪) من التباين المشاهد تعزى إلى أخطاء القياس.

5- مستويات الثبات المقبولة:

قيمة معامل ثبات الاختبار تتوقف على عدد من العوامل منها نوع الاختبار ذاته. وعلى الرغم من أن معامل ثبات قدره (0,90) يعتبر معامل ثبات مقبول لأي اختبار إلا أن معاملات ثبات اختبارات الاستعدادات العقلية والاختبارات التحصيلية المقننة يجب ألا تقل عن (0,90)، بينما تتراوح معاملات الثبات المقبولة لاختبارات الشخصية واستبيانات الميول بين (0,70) و (0,80)، وقد تصل معاملات ثبات بعض مقاييس الشخصية إلى (0,90).

ثالثاً: صدق الاختبار

- 1- تعريف صدق الاختبار.
- 2- أنواع صدق الاختبار.
- 3- العوامل المؤثرة في معامل صدق الاختبار
- 4- مستويات الصدق المقبولة.
- 5- العلاقة بين الصدق والثبات.

1- تعريف صدق الاختبار:

يعتبر صدق الاختبار خاصية سيكومترية تكشف عن مدى تأدية الاختبار للغرض الذي أعد من أجله، أو مدى قياسه لما أعد لقياسه.

فالصدق يحدد قيمة الاختبار وصلاحيته في قياس ما وضع لقياسه، فهو شرط ضروري لصلاحية الاختبار غير أنه ليس شرطاً كافياً بل لا بد من توفر شروط أخرى من أهمها الثبات، غير أن الصدق أكثر أهمية من الثبات، فقد يكون الاختبار ثابتاً ولكنه غير صادق، أما الاختبار الصادق فهو بالضرورة ثابت، لأن الصدق يتضمن الثبات.

2-أنواع صدق الاختبار.

هناك أنواع عديدة لصدق الاختبار سنتطرق إلى ثلاثة منها وهي:

- صدق المحتوى.
- الصدق المرتبط بالمحك.
- صدق التكوين الفرضي.

➤ صدق المحتوى:

ويسمى كذلك بالصدق المنطقي، أو الصدق بحكم التعريف، أو صدق عينة الاختبار؛ ويقصد به مدى تمثيل الاختبار للجوانب التي وضع لقياسها، أو مدى تمثيل الاختبار لجوانب المجال موضع القياس. وأفضل المجالات مناسبة لقبول صدق المحتوى هو مجال الاختبارات التحصيلية التي نهدف من استخدامها لتقدير حجم المعلومات التي تم تحصيلها عن مجال معين.

إجراءات تقدير صدق المحتوى :

الإجراء الأساسي في دراسة صدق محتوى الاختبار التحصيلي يتم من خلال تحديد درجة التطابق بين فقرات الاختبار وجدول المواصفات، والمشكلة في هذا الإجراء هو أن ارتفاع درجة التطابق بين فقرات الاختبار وجدول المواصفات لا يعد دليلاً على صدق محتوى الاختبار، ذلك لأن خبراء المادة قد لا يوافقون على الموضوعات والأهداف والمواصفات التي يهدف الاختبار إلى قياسها.

ومن ثم يجب أن تقوم مجموعة من الخبراء المختصين بتقدير درجة جودة جدول المواصفات أولاً، قبل عمل مقارنة بين محتوى الاختبار وجدول المواصفات.

وتسير إجراءات تقدير صدق محتوى اختبار تحصيلي ما سبق إعداده على النحو الآتي:

1. تحديد أهداف المقرر الدراسي التي يهدف الاختبار التحصيلي إلى قياسها.
2. تحديد الجوانب المختلفة للمجال موضع القياس، وإعطاء أوزان مناسبة لكل منها في إطار أهداف المقرر تتفق مع أهمية كل منها.
3. عمل تحليل منطقي لمحتوى الاختبار وفقراته، لتحديد الجوانب المختلفة الممثلة فيه ونسبة كل منها إلى الاختبار ككل.
4. عمل مقارنة بين فقرات الاختبار والمجال الذي يقيسه لتحديد مدى تمثيل فقرات الاختبار للمجال موضع القياس، وذلك بعرض الصورة المبدئية للاختبار وجدول المواصفات على مجموعة من المحكمين وهم خبراء المادة الدراسية.

ولا تؤسس درجة التطابق بين فقرات الاختبار والمجال الذي يقيسه على مؤشر إحصائي أو على تقديرات كمية، بل تؤسس على أحكام منطقية.

وفي الأخير نشير إلى أن صدق المحتوى بشكل عام يعتمد على تقديرات المحكمين، ولهذا فهو أكثر أنواع الصدق عرضة لأخطاء التقدير. وفي الوقت الذي يكون فيه عرضة لأخطاء التقدير إلا أنه أهم أنواع الصدق في الاختبارات بشكل عام واختبارات التحصيل بشكل خاص. أما عن سبب الاعتماد على طريقة التحكيم في تقدير صدق المحتوى فهو غياب المؤشرات الإحصائية، ولذلك فإن الاتجاه العام في تحديد صدق المحتوى هو الاعتماد على زيادة عدد المحكمين بقدر الإمكان للكشف عن مدى الاتفاق في تقديراتهم، بحيث لا يقل العدد عن خمسة محكمين مختصين في المادة الدراسية، ولديهم معرفة جيدة بأساليب القياس والتقويم وإعداد الاختبارات التحصيلية حتى تتصف أحكامهم بالدقة والموضوعية. ولكي يتصف الاختبار بدرجة مرتفعة من صدق المحتوى فيجب ألا تقل درجة الاتفاق بين المحكمين عن (80٪)، فإذا قلت درجة الاتفاق بينهم عن (60٪) فيجب مراجعة فقرات الاختبار وفحصها مرة أخرى.

➤ الصدق المرتبط بالمحك:

ويسمى كذلك بالصدق التجريبي، أو الصدق الواقعي؛ ويقصد به مدى الارتباط بين درجات الأفراد على الاختبار ودرجاتهم على محك خارجي مستقل ثبت صدقه وثباته، ويقاس الخاصية السلوكية نفسها التي يتناولها الاختبار بالقياس.

ويعد الصدق المرتبط بالمحك أكثر أساليب الصدق وضوحا في الممارسات الواقعية، حيث يتحدد الصدق باعتباره ما يقيسه الاختبار وفق محك خارجي، والمحك هنا مستقل تماما عن الاختبار، فإذا تزامن معه فإننا نطلق عليه اسم الصدق التلازمي، أما إذا كان فارق زمني بين الدرجة على الاختبار والدرجة على المحك فنكون إزاء ما نطلق عليه اسم الصدق التنبئي.

أولا: الصدق التنبئي:

ويقصد به قدرة الاختبار على التنبؤ بالأداء اللاحق للفرد على محك خارجي مستقل للأداء، والهدف منه هو التنبؤ على المدى البعيد بأداء الفرد في عمل لاحق يرتبط بما يقيسه الاختبار.

يتم في هذا النوع من الصدق جمع المعلومات على المحك، بعد فترة زمنية طويلة نسبيا من جمع المعلومات على الاختبار التي نبحت عن صدقه.

ويعتبر هذا النوع من الصدق أهم من غيره عندما يكون الغرض من الاختبار هو التنبؤ بسلوك مرتقب مثل، التنبؤ بنجاح الطالب في التحصيل الجامعي، أو التنبؤ بنجاح موظف في عمله.

أساليب قياس الصدق التنبئي:

• **طريقة معامل الارتباط:** وتقوم هذه الطريقة على حساب معامل الارتباط بين درجات الأفراد على الاختبار

ودرجاتهم على مقياس موضوعي خارجي للأداء اللاحق الذي يرتبط بما يقيسه الاختبار، ويطلق على معامل

الارتباط الناتج بمعامل الصدق التنبئي، وتعتبر هذه الطريقة أفضل وأدق الطرق في قياس الصدق التنبئي.

وتسير إجراءات حساب الصدق التنبئي باستخدام طريقة معامل الارتباط على النحو الآتي:

1. تطبيق الاختبار على عينة من الأفراد ممثلة للمجتمع الأصلي (عينة التقنين).
2. متابعة العينة فترة زمنية طويلة نسبيا (الفترة التي ينهي فيها الطالب دراسته أو العامل تدريبه)، ثم جمع بيانات المحك.

3. حساب معامل الارتباط بين بيانات الاختبارات وبيانات المحك، وذلك باستخدام معامل الارتباط الملائم لنوع البيانات المستخدمة، ومن أكثر معاملات الارتباط شيوعا واستخداما هو معامل الارتباط البسيط عند "بيرسون".

يدل حجم معامل الارتباط الناتج على القيمة التنبئية للاختبار، فكلما ارتفع معامل الصدق كلما دل ذلك على صحة التنبؤات التي نبنيها اعتمادا على درجات الاختبار، وبالتالي فالاختبار قادر على التنبؤ بالأفراد الذين سيحققون نجاحا في المستقبل.

• **طريقة المجموعات المتضادة (طريقة المتوسطات) :** وتقوم هذه الطريقة على حساب دلالة الفرق بين متوسطات درجات مجموعتين متطرفتين من الأفراد في الاختبار، إحداهما أخذت تقديرا مرتفعا في مقياس المحك والأخرى أخذت تقديرا منخفضا في مقياس المحك. فإذا ثبت أنه هناك فرق ذو دلالة إحصائية بين متوسطي درجات هاتين المجموعتين في الاختبار كان ذلك دليلا على صدق الاختبار.

وتسير إجراءات حساب الصدق التنبئي باستخدام طريقة المتوسطات على النحو الآتي:

1. تطبيق الاختبار على عينة من الأفراد ممثلة للمجتمع الأصلي (عينة التقنين).
2. متابعة العينة فترة زمنية طويلة نسبيا (الفترة التي ينهي فيها الطالب دراسته أو العامل تدريبه)، ثم جمع بيانات المحك.
3. ترتيب الأفراد طبقا لدرجاتهم على المحك.
4. تكوين مجموعتين متطرفتين على مقياس المحك، وذلك باختيار أعلى (27%) من العينة وأدنى (27%) من العينة.
5. حساب متوسط درجات مجموعة الأداء الأعلى في الاختبار ($\bar{X}_1$) ومتوسط درجات مجموعة الأداء الأدنى على الاختبار ($\bar{X}_2$).

6. حساب دلالة الفرق بين متوسطي درجات المجموعتين باستخدام اختبار (t) للمجموعات غير المتجانسة.

ومعادلته الرياضية هي على النحو الآتي:

$$t = \frac{|\bar{X}_1 - \bar{X}_2|}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

بحيث:

$$S = 1/n \sqrt{n \sum x^2 - (\sum x)^2}$$

$\bar{X}_1$: متوسط درجات مجموعة الأداء الأعلى على الاختبار.

$\bar{X}_2$: متوسط درجات مجموعة الأداء الأدنى على الاختبار.

S_1^2 : تباين درجات مجموعة الأداء الأعلى على الاختبار.

S_2^2 : تباين درجات مجموعة الأداء الأدنى على الاختبار.

n_1 : عدد أفراد مجموعة الأداء الأعلى على الاختبار.

n_2 : عدد أفراد مجموعة الأداء الأدنى على الاختبار.

إذا كانت قيمة (t) المحسوبة أعلى من قيمة (t) الجدولية لدرجة الحرية دليل على أنه هناك فرق جوهري بين متوسطي درجات مجموعتي المقارنة عند مستوى الدلالة المحدد، فالاختبار يقيس ما يقيسه المحك، أي أن الاختبار صادق ويقاس ما يدعي قياسه، حيث أنه يميز بين مجموعة الأداء الأعلى على المحك ومجموعة الأداء الأدنى على المحك.

أما إذا كانت قيمة (t) المحسوبة أقل من قيمة (t) الجدولية لدرجة الحرية دليل على عدم وجود فرق جوهري بين متوسطي درجات مجموعتي المقارنة عند مستوى الدلالة المحدد، ويشير ذلك إلى عدم قدرة الاختبار على التمييز بين مجموعة الأداء الأعلى على المحك ومجموعة الأداء الأدنى على المحك، وأنه لا يقيس ما يدعي قياسه.

ملاحظة: يتم تقدير درجة الحرية كالآتي:

$$df = (n_1 + n_2) - 2 \quad \text{عندما يكون } (n_1 \neq n_2)$$

$$df = 2(n - 1) \quad \text{عندما يكون } (n_1 = n_2)$$

مثال:

قام باحث ببناء اختبار جديد، وأراد التأكد من صدقه، فقام بتطبيقه على عينة من الطلبة وتابعهم حتى إنهاء دراستهم. بعدها طبق عليهم مقياس "المحك"، ثم رتبهم طبقاً للدرجات التي حصلوا عليها. ثم كون مجموعتين ضمت كل واحدة منهما 14 طالباً (27% أعلى، 27% أدنى)، ثم قام بحساب المتوسط الحسابي للمجموعة الأولى فوجده 55,32 بانحراف معياري قدره 18,49، أما المجموعة الثانية فكان المتوسط الحسابي لدرجات أفرادها 40,61 وانحراف معياري قدره 7,84.

المطلوب:

• ما نوع الصدق الذي أراد الباحث الحصول عليه؟

• ماهي الطريقة التي استخدمها؟

- هل هذا الاختبار صادق؟

الحل:

- الصدق الذي أراد الباحث الحصول عليه هو: الصدق المرتبط بالمحك وتحديد الصدق التنبئي.
- الطريقة التي استخدمها الباحث هنا هي طريقة المجموعات المتضادة (طريقة المتوسطات).
- نتأكد من صدق هذا الاختبار بحساب دلالة الفرق بين متوسطي درجات المجموعتين باستخدام اختبار (t) للمجموعات غير المتجانسة باستخدام المعادلة الآتية:

$$t = \frac{|\bar{X}_1 - \bar{X}_2|}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

بالتعويض في المعادلة نجد:

$$t = \frac{|55,32 - 40,61|}{\sqrt{\frac{18,49}{14} + \frac{7,84}{14}}} = \frac{14,71}{\sqrt{1,32 + 0,56}} = \frac{14,71}{\sqrt{1,88}} = \frac{14,71}{1,37} = \underline{10,74}$$

$$df = 2(n - 1) = 2(14 - 1) = 2 \times 13 = \underline{26}$$

نلاحظ أن قيمة (t) المحسوبة (10,74) أكبر من قيمة (t) الجدولية (2,77) عند مستوى الدلالة (0,01)، وهذا يدل على أن الفرق بين متوسط مجموعة الأداء الأعلى ومتوسط مجموعة الأداء الأدنى دال إحصائياً عند مستوى الدلالة (0,01)، فالاختبار يميز بين المجموعتين، وبالتالي فالاختبار صادق.

ثانياً: الصدق التلازمي:

ويقصد به الارتباط بين درجات الأفراد على الاختبار وأدائهم الراهن على محك خارجي مستقل يقيس ما يدعي الاختبار قياسه، وهو الأكثر ملاءمة لاختبارات الشخصية التشخيصية، وقد يستخدم في حساب صدق اختبارات الاستعدادات والقدرات، والهدف منه هو تحديد إمكانية أن يحل الاختبار الجديد بإجراءاته البسيطة وغير المكلفة محل المحك الذي

ثبت صدقه وثباته، والذي قد يكون مكلفا ومعقدا ويحتاج إلى جهود كبيرة، فالهدف هنا هو أن يكون الاختبار بديلا أبسط وأسرع، وأقل تكلفة في الوقت والجهد من المحك.

أساليب تقدير الصدق التلازمي:

طريقة معامل الارتباط هي الأكثر شيوعا واستخداما، وتسير إجراءات حساب الصدق التلازمي باستخدام طريقة معامل الارتباط على النحو الآتي:

1. تطبيق الاختبار على عينة من الأفراد ممثلة للمجتمع الأصلي (عينة التقنين).
 2. جمع بيانات المحك قبل أو خلال الفترة نفسها التي يتم فيها جمع بيانات الاختبار، بحيث يتزامن جمع بيانات الاختبار مع بيانات المحك.
 3. حساب معامل الارتباط بين بيانات الاختبارات وبيانات المحك، وذلك باستخدام معامل الارتباط الملائم لنوع البيانات المستخدمة، ومن أكثر معاملات الارتباط شيوعا واستخداما هو معامل الارتباط البسيط عند "بيرسون".
- وكلما ارتفع معامل الارتباط معناه ارتفاع معامل الصدق التلازمي، أي أن الاختبار صادق وقادر على التمييز بين الأفراد على مقياس المحك.

الفرق بين الصدق التنبئي والصدق التلازمي:

- الهدف من الصدق التنبئي هو التنبؤ على المدى البعيد بأداء الفرد في عمل لاحق أو دراسة لاحقة، بينما الهدف من الصدق التلازمي هو قياس حالة قائمة بالفعل ومعرفة إمكانية إحلال الاختبار الجديد مكان المحك.
- في حالة الصدق التنبئي يتم جمع بيانات المحك بعد إجراء الاختبار بفترة زمنية معينة، بينما في حالة الصدق التلازمي يتم جمع بيانات كل من الاختبار والمحك في الوقت نفسه تقريبا.

➤ صدق التكوين الفرضي:

ويسمى كذلك بالصدق البنائي، ويقصد به مدى قياس الاختبار للتكوين الفرضي أو السمة التي يهدف إلى قياسها، ويتطلب هذا النوع من الصدق توفر قدر كبير من المعلومات التي تلقي الضوء على طبيعة السمة موضوع القياس. وللكشف عن صدق البناء للاختبار فقد أشار "جرونلاند" (Gronlund) إلى ثلاث خطوات هي:

1. التعريف بالإطار النظري للسمة التي يفترض ارتباطها بنتائج الاختبار.
2. اشتقاق فرضيات حول نتائج الاختبار من خلال الإطار النظري.
3. التحقق من صحة الفرضيات منطقيا أو تجريبيا.

مثال: الذكاء هو بناء سيكولوجي نفترض وجوده، وله أكثر من تعريف لهذا تعددت الاختبارات التي حاولت قياسه، إلا أنه هناك بعض المعلومات النظرية التي تصلح كمؤشرات عن طبيعة الذكاء، فعلى باني الاختبار جمع أكبر عدد ممكن من المؤشرات التي يمكن أن تساهم في تأكيد الصدق البنائي للاختبار.

فيما يلي بعض المؤشرات التي تعتبر بمثابة فرضيات يمكن اشتقاقها من خلال الإطار النظري للذكاء:

- درجات الأفراد على اختبار الذكاء تزيد بزيادة العمر.
- درجات الأفراد على اختبار الذكاء تتنبأ بنجاحهم الأكاديمي.
- درجات الأفراد على اختبار الذكاء ترتبط طرديا مع درجاتهم على اختبارات أخرى أعدت للغرض نفسه.
- درجات الأفراد على اختبار الذكاء تميز بين الموهوبين والمعوقين عقليا.

في حالة عدم تحقق إحدى هذه الفرضيات يعني ثلاثة احتمالات:

- وجود خلل تجريبي في فحص الفرضية.
- وجود خلل في الإطار النظري.
- الاختبار لا يقيس السمة التي أعد لقياسها، وهو الاحتمال الذي نشك فيه.

هناك العديد من الأساليب، والمؤشرات الإحصائية وغير الإحصائية (المنطقية) المستخدمة في الكشف عن الصدق

البنائي ومنها:

أولا: الاتساق (الانسجام الداخلي) :

ويعني أن قوة الارتباط بين الفقرات المعدة لقياس السمة تعتبر مؤشرا إحصائيا لصدق البناء.

الخاصية الأساسية لهذه الطريقة هي أن المعيار ما هو إلا الدرجة الكلية للأداة، حيث تحسب الارتباطات بين البنود منفردة والدرجة الكلية للأداة، وكلما كان الارتباط موجبا ودالا دل ذلك على أن البند يغطي بعدا من أبعاد مجال السلوك أو الاتجاه أو المهارة محل الدراسة. كما تستخدم طريقة الاتساق الداخلي في اختبار العلاقة بين المقاييس الفرعية للأداة والدرجة الكلية لها، ويترتب عن هذه الطريقة بعض الإجراءات التجريبية، بحيث يمكن للباحث أن يرفع صدق الأداة بالتخلي عن البنود التي تظهر ارتباطا ضعيفا أو سالبا مع الدرجة الكلية للأداة.

ثانيا: الصدق التمييزي :

وتتلخص هذه الطريقة في إيجاد تصميم تجريبي فيه مجموعتان متميزتان منطقيا بالنسبة للصفة المقيسة، فإذا كشفت

نتائج الاختبار عن وجود فرق حقيقي بين المجموعتين، فإن هذا مؤشر من مؤشرات صدق البناء السيكولوجي للاختبار.

ولتحقيق ذلك يمكن اعتماد إجراءين وهما:

✓ **طريقة الفروق المتقابلة:** وتعتمد على مقارنة نتائج مجموعتين مختلفتين على المقياس المعد، كأن نقارن نتائج طلبة معهد تخصص معين بنتائج مجموعة غير منتقاة من طلبة الجامعة، وهذه الطريقة شائعة الاستخدام لقياس صدق اختبارات الشخصية ومقاييس الاتجاهات والميول.

✓ **طريقة الفروق الطرفية:** وتعتمد على قدرة المقياس في التمييز بين طرفي القدرة أو السمة التي يقيسها المقياس المعد، ويكون ذلك بحساب دلالة الفرق بين الطرفين من أفراد المجموعة وهما الفئة العليا التي تحصلت على درجات عالية، والفئة الدنيا التي تحصلت على درجات صغرى، ويمكن إجراء هذه المقارنة بإحدى هاتين الطريقتين:

- **مقارنة الفئات المتطرفة في المقياس مع محك خارجي:** تتم في هذه الطريقة مقارنة الربع الأعلى في درجات الاختبار بالربع الأعلى في المحك الخارجي، ومقارنة الربع الأدنى في درجات الاختبار بالربع الأدنى في المحك الخارجي. وتجرى هذه المقارنات عن طريق الكشف عن الدلالة الإحصائية للفرق بين المتوسطات بتطبيق اختبار (t)، فإذا لم تظهر فروق دالة إحصائية في المقارنتين بين الفئتين أمكن القول بأن المقياس صادق بدلالة المحك.

- **مقارنة الفئات المتطرفة في المقياس ذاته:** تتم في هذه الطريقة مقارنة الربع الأعلى في درجات الاختبار بالربع الأدنى في الدرجات في الاختبار نفسه. وتجرى هذه المقارنات عن طريق الكشف عن الدلالة الإحصائية للفرق بين المتوسطات بتطبيق اختبار (t)، فإذا ظهرت فروق دالة إحصائية في المقارنتين بين الفئتين أمكن القول بأن المقياس صادق.

3-العوامل المؤثرة في معامل صدق الاختبار:

من العوامل التي تؤثر في صدق الاختبار نجد:

3-1- تجانس العينة المستخدمة في التقنين:

فكلما زاد التجانس بين أفراد العينة في الصفة المقيسة، كلما أدى إلى انخفاض ملموس في قيمة معامل الارتباط بين درجات الأفراد على الاختبار ودرجاتهم على المحك، وهذا يعني انخفاضاً في قيمة معامل صدق الاختبار.

3-2- الفترة الزمنية الفاصلة بين تطبيق الاختبار على أفراد عينة التقنين وجمع المعلومات على مقياس المحك:

عندما يزداد طول هذه الفترة يقل معامل الصدق التنبئي للاختبار، وذلك لتأثير الصدفة وعوامل أخرى؛ مثل نقص عدد أفراد عينة التقنين بسبب تسريحهم أو غيابهم بسبب المرض أو انتقالهم إلى مكان آخر أو عدم تعاونهم.

3-3- طول الاختبار:

أية زيادة في عدد مفردات الاختبار تؤدي عادة إلى ارتفاع كل من صدقه وثباته.

3-4- ثبات الاختبار:

كلما ارتفعت قيمة معامل الثبات ارتفعت قيمة معامل الصدق، وأية عوامل تؤدي إلى انخفاض معامل الثبات تؤدي إلى انخفاض معامل الصدق.

4- مستويات الصدق المقبولة:

- قيمة معامل صدق الاختبار تتوقف على الظروف التي تم تحديد معامل الصدق في إطارها فمثلاً:
- الارتباط بين درجات الذكاء وتقديرات المعلمين تتراوح بين (0,30) و (0,50)، وهي قيمة منخفضة لتدخل عوامل أخرى في تقديرات المعلمين غير الذكاء كاتجاهاتهم نحو طلابهم، وميولاتهم، وإعدادهم المسبق، وقدرتهم على التدريس.
- الارتباط بين درجات اختبارات التحصيل المقننة والمعدل التراكمي تتراوح بين (0,60) و (0,70).
- يشير "نول" و "سكانيل" (Noll & Scannel) إلى أن الارتباط بين درجات اختبارات الذكاء المقننة ودرجات اختبارات التحصيل المقننة تتراوح بين (0,60) و (0,80).

5- العلاقة بين الصدق والثبات:

- يمكن التعبير عن العلاقة بين الصدق والثبات بالمعادلة الآتية:
- معامل الصدق = الجذر التربيعي لمعامل الثبات، وهي قيمة الصدق الذاتي. وتشير هذه المعادلة إلى أن قيمة معامل الصدق لا يمكن أن تزيد عن الجذر التربيعي لقيمة معامل الثبات، فالعلاقة بينهما طردية؛ فكلما ارتفعت قيمة معامل الثبات ارتفعت قيمة معامل الصدق، وأية عوامل تعمل على انخفاض معامل الثبات تؤدي بدورها إلى انخفاض معامل الصدق، وهذا يعني أن معامل الارتباط بين الاختبار والمحك (الصدق) يزداد كلما زاد معامل الارتباط بين الاختبار وذاته (الثبات).
- فالصدق والثبات وجهان لعملة واحدة، ولا يمكن الاستغناء عن أحدهما، فرغم أن الصدق يتضمن الثبات، فيجدر بالباحث أن يتأكد من ثبات اختباريه بجانب تأكده من صدقه، لأنه بذلك يعزز حكمه على صلاحية اختباريه.