

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

محاضرات العام الجامعي 2025 - 2026
(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- دور القياس التربوي في نجاح العملية التعليمية - التعليمية.

1- تاريخ القياس النفسي والتربوي:

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

أ- 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 محكات 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 تعقيداً 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 أكدوا 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).

ب- 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".

ج- 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 1850م.

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.

أ- تعريف القياس:

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

ب- 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.

ج- 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.

د- 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:

- د- التقويم عملية شاملة، بينما القياس عملية جزئية تنصب على شيء واحد من الجوانب السابقة.
- ر- يهدف التقويم إلى التشخيص والعلاج، بينما يكتفي القياس بإعطاء بعض المعلومات الكمية المحددة عن الشيء موضوع القياس.
- ز- القياس عملية موضوعية وغير شخصية نسبياً، أما التقويم فهو تتصف غالباً بالذاتية.
- س- القياس سابق للتقويم وأساس له.
- ش- التقويم أكثر شمولاً واتساعاً من القياس.

3- خصائص القياس النفسي والتربوي:

يتميز القياس النفسي والتربوي بمجموعة من الخصائص نذكر منها:

- القياس النفسي والتربوي قياس غير مباشر.
- القياس النفسي والتربوي قياس نسبي وليس مطلقاً.
- القياس النفسي والتربوي قياس لعينة من السلوك.
- القياس النفسي والتربوي إجراء حيادي.
- القياس النفسي والتربوي يتضمن أخطاء معينة.
- القياس النفسي والتربوي أقل دقة من القياس المادي.
- القياس النفسي والتربوي مجرد وسيلة وليس غاية في حد ذاته.

4- أهداف القياس:

تتلخص أهداف القياس بصفة عامة في الآتي:

- الكشف عن الفروق بأنواعها المختلفة (بين الأفراد، في ذات الفرد، بين المهن المختلفة، بين الجماعات).
- الوصف الكمي لخصائص أو صفات الأشياء والأشخاص.

5- أنواع القياس:

يميز علماء القياس بين نوعين من القياس وهما:

- قياس العملية.

- قياس الناتج.

6- مراحل عملية القياس :

- تتضمن عملية القياس ثلاث مراحل أساسية وهي:
- مرحلة تعريف وتحديد الصفة أو الخاصية موضع القياس.
- مرحلة تحديد العمليات الإجرائية اللازمة للاستدلال على الصفة أو الخاصية موضع القياس.
- مرحلة التقدير الكمي للصفة أو الخاصية موضع القياس.

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:

أ- 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.

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، والمنوال وبعض المقاييس اللابرامترية مثل مربع كاي.

ب- المستوى الرتبي:

وهو يرتب الأشخاص أو الأشياء تصاعدياً أو تنازلياً على سلم متدرج، فيكون (س) مثلاً هو الأول و(ص) هو الثاني و(ع) هو الثالث وغيرها، ولكن قيمة هذه الأرقام هو في ترتيبها وليس في الكمية التي تحتوي عليها. فالوحدات في مقاييس الترتيب لا تكون متساوية في الحقيقة، ولو بدت من حيث الأرقام الدالة عليها وكأنها متساوية.

في هذا المستوى يعبر عن المتغير بترتيب؛ بحيث ترتب القياسات تصاعدياً أو تنازلياً، 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 الترتيب وهو مستوى أرقى من مستوى المقاييس الاسمية. وتعتمد مقاييس الترتيب على ترتيب مجموعة من الأفراد وفقاً لدرجاتهم في صفة من الصفات دون أن تعطى أي أهمية لتساوي الفروق بين الترتيب. القياس الترتيبي لا يسمح لنا بتحديد الفروق بين أي رتبتين بدقة.

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

ج- المستوى الفترتي:

يمكن أن نطلق عليه مقياس الوحدة المنتظمة أو مقياس المسافة، وتمتد مقاييس المسافة بفكرة ترتيب الترتيب 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. فمثلاً 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 المقاسة، 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.

• 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 وصفر مطلقاً كما في مقاييس الطول والوزن والحجم، وهو يستخدم على نطاق واسع في ميدان العلوم الطبيعية، بينما ينذر استعماله في العلوم التربوية والنفسية.

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

يلعب القياس التربوي دورا هاما في مجال التربية والتعليم وهذا الدور يجعل من العملية التعليمية-التعلمية 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: نظريتا القياس الكلاسيكية والحديثة

1. نظرية القياس الكلاسيكية.
2. نظرية القياس الحديثة (نظرية السمات الكامنة - نظرية الاستجابة للمفردة).
3. مقارنة بين نظريتي القياس الكلاسيكية والحديثة.

لكل نظرية من النظريات مجموعة من الفروض والمسلمات تقوم عليها من أجل تفسير الظواهر التي ترتبط بها، ولا بد أن تكون لهذه النظرية القدرة على التفسير والتحليل، حتى تكون نظرية صالحة للاستخدام والتطبيق.

1- نظرية القياس الكلاسيكية:

تقوم النظرية الكلاسيكية في القياس على أربع مسلمات رئيسية وهي:

- أداء الفرد يمكن قياسه وتقديره.
- أداء الفرد دال لخصائصه.
- الخاصية والأداء والعلاقة بينهما تختلف من فرد إلى آخر "الفروق الفردية".
- القياس الظاهري الكلي يتكون من قياس حقيقي وآخر يرجع إلى الخطأ.

نشأت نظرية القياس النفسي الكلاسيكية بصورة محددة من جهود "سبيرمان" (1927) في تقديم الأساس الرياضي لنظريته في الذكاء، والكثير من المصطلحات المعروفة مثل؛ الدرجة الحقيقية، والثبات، والصدق. تعد هذه النظرية من أهم وأقدم الطرق المستخدمة في القياس النفسي والتربوي، وقد انتشر استخدامها في بناء وتطوير مختلف أنواع الاختبارات النفسية؛ مثل اختبارات الذكاء، والاستعدادات، والميول، والاتجاهات، ومقاييس الشخصية، وغيرها من الاختبارات النفسية و الاختبارات التحصيلية، وتعتمد في تفسيرها في الثبات على الصور المتكافئة للمقاييس المستخدمة.

أ- الافتراضات الأساسية للنظرية الكلاسيكية:

- **الفرض الأول:** الدرجة الملاحظة (X) في الاختبار هي عبارة عن مجموع جزئين هما: الدرجة الحقيقية (T) ودرجة خطأ القياس (E)، ويمكن تفسير ذلك بالمعادلة التالية: الدرجة الملاحظة = الدرجة الحقيقية + درجة خطأ القياس.
$$X = T + E$$
- **الفرض الثاني:** لا يمكن قياس الدرجة الحقيقية، بل يمكن الاستدلال عليها أو تقديرها من خلال الدرجة الملاحظة، وذلك بحساب متوسط الدرجة الملاحظة؛ أي أن الدرجة الحقيقية للفرد تساوي متوسط الدرجات الملاحظة التي يحصل عليها الفرد عندما يطبق عليه عدد لا نهائي من نسخ الاختبارات المتكافئة.
- **الفرض الثالث:** تفترض هذه النظرية أن درجات الطلاب في السمة (القدرة) التي يقيسها الاختبار تتخذ شكل التوزيع الاعتمادي في المجتمع المستهدف، ويمكن أن يتحقق ذلك بانتقاء المفردات التي يكون قيم معامل صعوبتها حوالي (0,50).
- **الفرض الرابع:** لا يوجد ارتباط بين الدرجات الحقيقية وأخطاء القياس للمفحوصين على الاختبار.
- **الفرض الخامس:** لا يوجد ارتباط بين أخطاء القياس لاختبارين مختلفين على المفحوص نفسه.

- **الفرض السادس:** يشير إلى أن الدرجات على الاختبار الأول لا ترتبط بالدرجات الحقيقية على الاختبار الآخر.
- **الفرض السابع:** خطأ القياس هو خطأ عشوائي يحدد دقة القياس أو ما يسمى بثبات الاختبار، بينما الخطأ المنتظم هو خطأ يتعلق بصدق القياس أو بصدق الاختبار.
- تفترض هذه النظرية تساوي تباين أخطاء القياس لجميع المفحوصين الذين يطبق عليهم الاختبار أو أداة القياس.

ب- مميزات النظرية الكلاسيكية:

- للنظرية الكلاسيكية عدة مميزات، وقد ذكرها "سليمان" و"علام" (2010) وهي كالتالي:
- من السهل على غير المتخصصين في مجال القياس النفسي والتربوي من تحليل بيانات الاختبارات، أي المعالجة الإحصائية لمفرداتها، وحساب معاملات السهولة والصعوبة والتمييز لمفرداتها، واستيعاب المفاهيم المتعلقة بها.
- بساطة الافتراضات التي تقوم عليها النظرية التقليدية للقياس، والتي يمكن أن تلائم بيانات مختلف أنواع الاختبارات النفسية والتربوية، بالإضافة إلى أنها لا تحتاج إلى نماذج رياضية معقدة لتحليل بياناتها.
- يمكن تحليل بيانات الاختبارات المختلفة دون الحاجة لبرامج متخصصة لتحليل تلك البيانات، ولا تتطلب متخصصين لتحليل تلك البيانات، أو متخصصين في مجال الحاسب الآلي.

ج- عيوب النظرية الكلاسيكية:

- رغم شيوع نظرية الاختبارات الكلاسيكية واستخدامها من قبل العاملين في المقاييس والاختبارات، إلا أنها لم تخل من جوانب القصور في تحليل نتائج الاختبارات، ومن أهم جوانب القصور في هذه النظرية ما يلي:
- جميع الخصائص السيكمومترية التي تستند في بنائها على المدخل الكلاسيكي مثل؛ معاملات الصعوبة والتمييز والثبات تعتمد على خصائص عينة الأفراد التي يطبق عليها الاختبار، وعلى مدى صعوبة عينة الفقرات التي يشتمل عليها الاختبار.
- يفترض أن درجات الاختبار التي تمثل السمة أو القدرة المقاسة دالة خطية مطردة؛ أي أنه كلما زادت درجة الفرد في الاختبار دل ذلك على زيادة مقدار السمة أو القدرة لديه، غير أن بعض الأفراد من ذوي القدرات المرتفعة يحصلون أحيانا على درجات منخفضة في الاختبارات، وربما يحدث العكس بالنسبة لذوي القدرات المنخفضة، وذلك وفقا لمعنى الدرجة على الاختبار.
- يتغير تكوين الاختبارات بمرور الزمن؛ ويعني ذلك أن تكوين ومعنى فقرات الاختبار يتغير من زمن إلى آخر بالنسبة لمجموعات الأفراد التي أعد لها الاختبار، فالظروف البيئية تتغير، والظروف الاختبارية مقننة، وإن حذف أو تغيير أي فقرة من فقرات الاختبارات قد يؤدي إلى تغيير في درجات الأفراد تغيرا يصعب التنبؤ به، ويؤثر تأثيرا بالغا في تمثيل الفقرات لنطاقها السلوكي.
- إن نتائج الأفراد على اختبار ما تعتمد على خصائص عينة الفقرات التي يشتمل عليها الاختبار؛ أي أنه إذا سحبنا عينات من الفقرات المتباينة في صعوبتها من مجموعة كبيرة من الفقرات، بحيث تقيس القدرة أو السمة نفسها، فإن الدرجات المتوقعة للأفراد المفحوصين في الاختبارات التي تشتمل كل منها على إحدى هذه العينات سوف تختلف باختلاف صعوبة الفقرات على المجموعة نفسها من الفقرات أو مجموعة مكافئة لها.
- الاختصار على هدف واحد من أهداف القياس، وعدم الاهتمام بالأهداف الأخرى للقياس؛ مثل تقدير مستوى أداء الأفراد نسبة إلى الجماعات الأخرى التي لا ينتهي إليها هذا الفرد؛ كأن نقوم بمقارنة أداء هذا الفرد بأداء أفراد من الفئات العمرية المختلفة.
- النظرية الكلاسيكية في القياس لا تقدم تفسيراً واضحاً لنمط أداء الأفراد المفحوصين على فقرات الاختبار، بالرغم من أهمية هذا التفسير في التعرف على الأسباب الحقيقية وراء بعض الاستجابات الشاذة، ومن ثم اتخاذ القرار المناسب بشأن هذه الاستجابات.

- عدم تحقق الموضوعية للقياس؛ والتي تتمثل في اختلاف نتيجة القياس باختلاف كل من الاختبار المستخدم ومستوى العينة في حالة القياس معياري المرجع، أو مستوى المحك في حالة القياس محكي المرجع.

2- نظرية القياس الحديثة (نظرية السمات الكامنة - نظرية الاستجابة للمفردة) :

رغم ما قدمته النظرية الكلاسيكية إلا أنها عجزت عن تحقيق الموضوعية في القياس، مما جعل الباحثين يتجهون إلى مدخل آخر يعتمد على النظرية المعاصرة للاختبارات، والتي تعرف باسم نظرية السمات الكامنة، أو نظرية الاستجابة للمفردة. لذلك تم طرح عدة تساؤلات حول كيف يمكن جعل القياس النفسي والتربوي موضوعيا، وكيف تتحرر أدوات القياس من خصائص الأفراد التي نطبق عليهم هذه الأدوات، وكيف يتحرر قياس قدرات وسمات الأفراد من خصائص عينة الأسئلة التي يشتمل عليها الاختبار الذي يطبق عليهم.

يرى كل من "رايت" (1974) و "لورد" (1978) و "جروجيتز" (1983) و "هامبلتون" أن بناء نظام موضوعي لتحويل الملاحظات إلى قياسات يتطلب أن تتحرر تدريجات أدوات القياس من خصائص الأشياء التي تقيسها هذه الأدوات، وأن يتحرر قياس الخصائص أو الأشياء من خصائص أدوات القياس التي استخدمت في هذا القياس، وهذا يتطلب أن يشترك كل من قياس الفرد وتفسير الفقرات في وحده قياس ثابتة على المدى المتصل للمتغير، وأن يكون تدريج هذا المتصل خطيا ووحداته متساوية، وبذلك يسمح هذا النظام بتعميم القياسات على أي أدوات القياس غير المستخدمة في القياس، كما يسمح بالمقارنة بين خصائص الأشياء المراد قياسها بأدوات قياس مشابهة غير متطابقة، وأيضا يسمح بتوسيع أو تضيق مدى المحتوى الذي تقيسه أدوات القياس بما يتناسب مع المواقف والأغراض المختلفة للقياس، بحيث تناسب متطلبات قياس جديدة كما هو الحال في القياس الفيزيائي.

تلقي نظرية السمات الكامنة في الوقت الحالي اهتماما ملحوظا نظرا لتزايد المطالب التعليمية والتربوية لقياس فعالية البرامج التعليمية قياسا موضوعيا وصادقا، وتقدير تحصيل الطلاب في المقررات الدراسية بصورة موضوعية، والاستفادة منه في عمليات التقويم والتوجيه والانتقاء، وصنع القرار التربوي السليم، وبخاصة أن تلك النماذج جاءت نتيجة للانتقادات التي وجهت إلى نظرية القياس التقليدية التي لم تستطع أن تجعل القياس التربوي والنفسي موضوعيا كما هو الحال في القياس الفيزيائي.

ويطلق على هذه النظرية أحيانا بنظرية (المنحى المميز للفقرات) أو نظرية (الاستجابة للفقرات)، وتقوم على افتراض وجود واحد أو أكثر من المميزات الأساسية التي تحدد استجابات الفرد لفقرات الاختبار، وأن السمات لا يمكن ملاحظتها، وإنما يستدل على مقدارها من السلوك الملاحظ للفرد الذي يمثل في إجابته عن فقرات الاختبار. وقد انبثقت من هذه النظرية مجموعة من النماذج أو الدوال الرياضية التي تعرف بنماذج السمات الكامنة، وتستهدف جميعا تحديد العلاقة المتوقعة بين استجابة الفرد على الاختبار وبين السمات أو القدرات غير الملاحظة.

أ- الهدف الرئيس لنظرية السمات الكامنة:

تهدف نظرية السمات الكامنة والنماذج الرياضية المرتبطة بها إلى تقدير جميع إحصاءات المفردة والقدرة، وكلما كان هناك ملائمة بين النموذج المستخدم ومجموعة البيانات أدى ذلك إلى الحصول على تقديرات دقيقة لهذه الإحصاءات، حيث يجب وضع كل من الأفراد والمفردات على مقياس للقدرة، ويتم ذلك من خلال عمليات التقدير طالما توجد علاقة تقارب ممكنة بين الاحتمالات المتوقعة للممتحنين والاحتمالات الفعلية لأدائهم في كل مستوى من مستويات القدرة، مع الأخذ في الاعتبار أن تقديرات إحصاءات المفردة وتقديرات قدرة الممتحنين يجب أن تعدل وتراجع بصفة مستمرة حتى نحصل على أقصى اتساق ممكن بين التنبؤات على أساس تقديرات القدرة ومتغيرات المفردة وبين البيانات الواقعية الفعلية للاختبار، ويتم ذلك من خلال برامج الحاسب الآلي المعدة لذلك.

ب- متطلبات وشروط القياس في النظرية الحديثة:

- صدق الفقرات الاختبارية في تعريف المتغير موضوع القياس تعريفا إجرائيا.

- صدق تدريج الفقرات الاختبارية التي تقيس متغيراً معيناً، بحيث يمكن تمثيل هذا المتغير بواسطة خط مستقيم.
- ضرورة توافر قياسات خطية يمكن استخدامها لدراسة النمو، أو للمقارنة بين الأفراد أو المجموعات.
- التوافق بين تدريج الأفراد على الاختبار وخصائص هذه الفقرات، بحيث تؤدي إلى تقديرات لمستويات الأفراد لا تتقيد باختبار معين.

ج- الفرضية الأساسية للنظرية الحديثة:

تعتمد نظرية السمات الكامنة على فرضية أساسية مفادها أن القيمة الاحتمالية لاستعمال فرد لفقرة اختبارية تكون دالة لكل من السمة أو القدرة والخاصية التي يفترض أن الاختبار يقيسها، وخصائص الفقرة التي يحاول الإجابة عنها؛ أي أنها تفترض أنه هناك دالة احتمالية تربط بين معلمين أحدهما يتعلق بالفرد والآخر يتعلق بالمفردة التي يعتبر فيها، وأن احتمالية الاستجابة الصحيحة للفقرة تعد دالة لمستوى السمة الكامنة التي تظهر عند كل مستوى للسمة أو القدرة.

د- ركائز نظرية السمات الكامنة:

- أحادية البعد: وجود قدرة واحدة فقط يمكن أن تفسر أداء الفرد في اختبار ما، وتتميز مفردات الاختبارات التي يتم إعدادها وفقاً لهذه النماذج بأنها لا تختلف فيما بينها إلا من حيث مستوى الصعوبة.
- استقلالية القياس: عدم اعتماد تقدير صعوبة المفردة على صعوبة المفردات الأخرى المكونة للاختبار وعلى قدرة الأفراد الذين يجيبون عليها، ولا يعتمد تقدير قدرة الأفراد على قدرة أي مجموعة أخرى من الأفراد الذين يؤدون الاختبار نفسه، ولا على صعوبة المفردات التي يؤدونها.
- التحرر من السرعة: لا يتم إعداد اختبارات نماذج السمات الكامنة بحيث تكون موقوتة، حيث أن عامل السرعة لا يمثل دوراً في الإجابة على مفردات الاختبار، حيث يعود الفشل في الإجابة إلى مقدار القدرة المحددة، وليس إلى الفشل في الإجابة على كل مفردات الاختبار.
- المنحى المميز للفقرة: هو عبارة عن مخطط بياني لاحتمالية الإجابة الصحيحة على الفقرة، ويعد دالة للسمة التي يتضمنها الأداء على فقرات المقياس، وتفترض هذه النظرية أن المنحى المميز للفقرة دالة رياضية تربط بين احتمال نجاح الفرد في إجابته عن الفقرة، وبين السمة التي تقيسها الفقرات التي تشملها.
- عدم تباين معالم الفقرة: هو الوصف التقني لاستقلالية القياس؛ والتي تعتبر من الشروط الموضوعية في القياس النفسي بوجه عام وفي نظرية السمات الكامنة بوجه خاص، وهذا يعني أن معالم الفقرة لا تختلف عندما تحسب من مجتمعات مختلفة؛ أي أنها توفر مميزات القياس بوجه عام.

هـ- عيوب النظرية الحديثة: (يعي سعد، 2024) <https://drasah.com/Description.aspx?id=8470>

- مع المزايا التي توفرها النظرية الحديثة في القياس، إلا أن لها بعض العيوب والقيود، والتي تشمل:
- تتطلب النظرية الحديثة في القياس أدوات وتحليلات متقدمة، مما يجعلها صعبة التطبيق بدون توفر برمجيات متطورة أو موارد تقنية متخصصة.
- تتطلب النظرية افتراضات معينة حول طبيعة البيانات وتوزيعها، مثل الفرضية بأن البيانات تتبع توزيعاً معيناً أو أن سلوك الأفراد متجانس. إذا لم تتحقق هذه الفرضيات، فقد تصبح النتائج غير دقيقة.
- في بعض الأحيان، تكون نتائج القياس المعتمدة على النظرية الحديثة معقدة، ويصعب تفسيرها مباشرة من قبل الأشخاص غير المتخصصين.
- تعتمد النظرية الحديثة على تجانس البيانات ووجود أنماط ثابتة، وقد تؤدي البيانات المتباينة إلى نتائج مضللة أو غير مستقرة.

- قد تكون هناك حاجة إلى برمجيات خاصة وتدريب لتطبيق النظرية الحديثة في القياس، مما يزيد من التكلفة والوقت اللازمين لإجراء البحث.
- في بعض الحالات، قد تكون النتائج صعبة التعميم على مجموعات أو سياقات أخرى، نظراً لأنها قد تكون خاصة بالعينات المدروسة أو المنهجية المتبعة.
- مع أنها تقدم إطاراً حديثاً، إلا أن النظرية قد لا تتكيف بشكل جيد مع التغيرات السريعة التي تحدث في بعض المجالات، مما يتطلب تحديثات مستمرة.

3- مقارنة بين نظريتي القياس الكلاسيكية والحديثة :

النظرية الحديثة (نظرية الاستجابة للمفردة - نظرية السمات الكامنة)	النظرية التقليدية (الكلاسيكية)
1-ارتفاع عدد الأفراد الذين يجيبون على مفردة ما إجابة صحيحة يعزى إلى قدرة السمة الكامنة لدى المفحوصين.	1-ارتفاع عدد الأفراد الذين يجيبون على مفردة ما إجابة صحيحة يعزى إلى سهولة هذه المفردة والعكس صحيح.
2-تفصل بين خصائص الاختبار والخصائص الكامنة للمفحوصين؛ بمعنى أن الدرجة المرتفعة في اختبار ما تعكس قدرة مرتفعة، وكذلك الدرجة المنخفضة في الاختبار على قدرة منخفضة.	2-لا تفصل بين خصائص الاختبار والخصائص الكامنة لدى المفحوصين.
3-ثبات الاختبار لا يرتبط بعدد مفردات الاختبار، ولكن بجودة مفرداته؛ بمعنى أنه قد يكون هناك اختبار قصير يتمتع بدرجة ثبات عالية بسبب جودة مفرداته.	3-ثبات الاختبار مرتبط بعدد مفردات الاختبار، فكلما زاد عدد المفردات زاد ثبات الاختبار، والعكس صحيح.
4-توزيع الدرجات لا تفترض التوزيع الاعتيادي.	4-توزيع الدرجات يأخذ شكل التوزيع الاعتيادي، مع العلم أن هذا التوزيع مرتبط بعينة الأفراد التي يطبق عليها الاختبار سواء من حيث حجم العينة أو نوعية الأفراد المفحوصين.
5-يعتمد تفسير نتائج الاختبار على ميزات مشتركة بين قدرة الأفراد وصعوبة المفردات، وبالتالي فإن تفسير الدرجة يرد إلى المفردة وليس إلى جماعه المرجع.	5-يعتمد تفسير نتائج الاختبار على موقع درجة الطالب بالنسبة لمتوسط أداء الجماعة، فالدرجة لا تشير إلى مستوى القدرة الواقعية للشخص وإنما تشير إلى الترتيب.
6-بالإمكان مقارنة نتائج اختبار لدى مجموعات مختلفة من المفحوصين، وأيضا لدى أشكال مختلفة من الاختبارات بسبب عدم تغيير معالم القدرة والمفردة.	6-نتائج الاختبار غير قابلة للمقارنة من عينة إلى عينة أخرى، ولا من صيغة اختبار إلى صيغة اختبار أخرى.
7-تسمح بالتنبؤ بكيفية أداء المفحوصين على الاختبار.	7-لا تسمح بالتنبؤ حول كيفية أداء المفحوصين على مفردات الاختبار.
8-الخطأ المعياري يكون لكل فرد بشكل منفرد ومستقل عن أفراد العينة ولكل مستوى معين من القدرة.	8-تفترض أن الخطأ المعياري ثابت بالنسبة للجماعة التي طبق عليها الاختبار؛ أي أن الخطأ المعياري ينطبق على كل الدرجات في المجتمع المعطى.

ثالثا: الإختبارات التحصيلية

1- لمحة تاريخية عن الاختبارات التحصيلية.

2- تعريف الاختبار التحصيلي.

3- خصائص الاختبار التحصيلي الجيد.

4- أهمية الاختبارات التحصيلية.

5- تصنيفات الاختبارات التحصيلية.

1- لمحة تاريخية عن الاختبارات التحصيلية:

كانت الاختبارات التحصيلية تعتمد اعتمادا أساسيا على الاختبارات الشفوية حتى نهاية النصف الأول من القرن التاسع عشر، وفي عام 1845 تزعم "هورانس مان" الذي كان يعد من القادة التربويين البارزين في ولاية "بوسطن" الأمريكية حركة تطوير التعليم العام، وأكد على ضرورة استخدام الامتحانات التحريرية التي تتسم بالموضوعية والاتساق بدلا من الاختبارات الشفوية في تقويم المتعلمين. ولكن استمر استخدام الاختبارات الشفوية في تقويم المتعلمين إلى أن بدأت في الاضمحلال وحلت محلها الاختبارات التحريرية.

وفي عام 1864 أعد المربي الانجليزي "جورج فيشر" الذي كان يعمل مدير مدرسة أول اختبار تحصيلي تحريري يتكون من عدة مقاييس متدرجة يمكن باستخدامها تقييم جودة الخط والنحو والتعبير والهجاء والرياضيات وغيرها من المواد الدراسية، وأطلق على هذه البطارية من الاختبارات "كتاب الموازين"، وقد وضح في هذا الكتاب كيفية تقدير درجات كل اختبار فرعي تشتمل عليه البطارية، والدرجات الكلية والمعايير التي يمكن استخدامها في مقارنة الطالب بنفسه في هذه المواد المختلفة، ورسم صفحة تربوية تبين جوانب قوته وضعفه.

وبعد مضي ما يقرب من ربع قرن أعد "رايس" عام 1895 أول اختبار تحصيلي موضوعي في الولايات المتحدة الأمريكية يقيس قدرة تلاميذ المدارس الابتدائية على الهجاء وطبقه على ما يقارب 16000 تلميذا في الصفوف من الرابع حتى الثامن. كما أعد اختبارا تحصيليا في الحساب واللغة الانجليزية. وتعد هذه الاختبارات لبنات لبناء الاختبارات التحصيلية المقننة، حيث شهدت أوائل القرن العشرين ظهور عدد من الاختبارات المقننة متمثلة في اختبار الحساب الذي أعده "ستون" عام 1908، واختبار جودة الخط الذي أعده "ثورندايك" عام 1909. وتزايد عدد هذه الاختبارات زيادة كبيرة قرب نهاية عام 1920 بما في ذلك بطاريات الاختبارات؛ مثل بطارية اختبارات "ستانفورد" التحصيلية للمرحلة الابتدائية عام 1923، وبطارية امتحانات "أيوا" للمحتوى الدراسي للمرحلة الثانوية عام 1925.

وتزايد عدد الاختبارات المقننة المنشورة التي استخدمت المفردات الموضوعية حتى وصل إلى المئات، كما ازداد الاهتمام ببناء اختبارات تقيس الفهم والاستنباط وتطبيق المعلومات، وغيرها من العمليات المعرفية العليا، والأهداف التعليمية.

2- تعريف الاختبار التحصيلي:

يقصد بالاختبار التحصيلي الأداة التي تستخدم في قياس المعرفة والفهم والمهارة في مادة دراسية أو تدريبية معينة أو مجموعة

من المواد.

ويعرف كذلك بأنه وضع المفحوص في موقف يطلب منه فيه أن يظهر معارفه، أو مهاراته، أو اتجاهاته، أو ميوله، أو جوانب منه تتصل بموضوع معين أو عدد من الموضوعات، فهو مجموعة من المواقف تمثل عينات من السلوك تعرض على المفحوصين، ويطلب منهم أن يقوموا بأداءات معينة، ويمكن اعتبارها دليلاً أو مؤشراً على تعلم الطالب.

3- خصائص الاختبار التحصيلي الجيد:

- ولكي تؤدي الاختبارات التحصيلية الجيدة وظائفها على أكمل وجه، لا بد من اتصافها بالآتي:
- الموضوعية: وتعني عدم تأثر نتائج التقويم بالعوامل الذاتية أو الشخصية للمصحح.
- الصدق: ويقصد به قياس الاختبار ما وضع لقياسه، ولا يقيس شيئاً آخر.
- الثبات: ويعني أن يعطي الاختبار النتائج نفسها إذا ما أعيد تطبيقه لمرات مختلفة وفي ظروف متشابهة.
- الشمول: وتعني تغطية الاختبار لكل جوانب المقرر الدراسي.

4- أهمية الاختبارات التحصيلية:

- تتجلى أهمية الاختبارات التحصيلية في كونها:
- تتيح للمعلمين إعادة النظر في الطرق والأساليب التي يستخدمونها في تدريسهم.
- تحفز التلاميذ على الدراسة.
- تبين للمعلمين النواحي التي يجب تأكيدها في التدريس.
- تساعد في تصنيف التلاميذ حسب المستوى.
- تكشف عن الصعوبات التي يواجهها بعض التلاميذ.
- تساعد على تثبيت التعلم.
- تقدم تغذية راجعة فورية حول سير العملية التعليمية.

رابعاً: أدوات القياس التربوي

- 1- الاختبارات.
- 2- قوائم الشطب.
- 3- سلاله التقدير.
- 4- البورتفوليو.
- 5- السجل القصصي.

1- الاختبارات: تم التطرق إليها في المحور السابق.

2- قوائم الشطب:

أداة مكونة من مجموعة فقرات ذات صلة بالمتغير أو السمة المقيسة، وكل فقرة من هذه الفقرات تعبر عن سلوك بسيط يخضع للكل أو للعدم؛ بمعنى أن التدرج هنا من مستويين فقط، وبالتالي فهي صورة من صور مقاييس التقدير. عرفها الفريق الوطني للتقويم بأنها قائمة الأفعال أو السلوكيات التي يرصدها المعلم أو المتعلم لدى قيامه بتنفيذ مهمة أو مهارة تعليمية واحدة أو أكثر، وذلك برصد الاستجابات على فقراتها باختيار أحد تقديريين من الأزواج التالية (صح أو خطأ)، (نعم أو لا)، (موافق أو غير موافق)، وتعد من الأدوات المناسبة لقياس مدى تحقق النتائج التعليمية. تعتبر قوائم الرصد وسيلة فعالة للحصول على معلومات بصيغة مختصرة، وتساعد الطالب والمعلم على تحديد مواطن القوة والضعف في الأداء بسرعة، وتساعد على تحديد الخطوات الضرورية لرفع كفاءة التعلم باستمرار.

خطوات تصميم قوائم الشطب أو الرصد:

1. تحليل المحتوى للحصول على نتائج التعلم.
2. اختيار المعايير المناسبة للتقويم.
3. تخصيص علامة مناسبة لكل فقرة حسب أهميتها.
4. مناقشة هذه القوائم مع الطلبة والاتفاق عليها.

3- سلاله التقدير:

سلم التقدير هو أداة بسيطة تظهر فيها إذا كانت مهارات المتعلم متدنية أو مرتفعة، حيث تخضع كل فقرة للتدرج من عدة فئات أو مستويات، حيث يمثل أحد طرفيه انعدام أو وجود الصفة التي نقدرها بشكل ضئيل، ويمثل الطرف الآخر تمام أو كمال وجودها، وما بين الطرفين درجات متفاوتة من وجودها.

يتم إعداد سلم التقدير حسب الخطوات الآتية:

1. تجزئة المهارة أو المهمة إلى مجموعة من المهام الأصغر، أو إلى مجموعة من السلوكيات المكونة للمهارة المطلوبة.
2. ترتيب السلوكيات المكونة للمهارة المقيسة حسب تسلسل حدوثها، أو بحسب تنفيذها من قبل المتعلمين.
3. اختيار التدرج المناسب على سلم التقدير لتقدير مدى إنجاز المهارة، وذلك وفقاً لطبيعة المهارة ومجموعة السلوكيات المتضمنة فيها والتي سيؤديها المتعلمون.

يمتاز سلم التقدير بما يلي:

1. يمكن استخدامه بصوره فعالة واقتصادية توفيراً لجهد المعلم ووقته.
2. يتصف بدرجة من الموضوعية والثبات أعلى مما يتوفر في أساليب التقويم القائمة على الملاحظة العادية.
3. يستخدم في تقويم أنواع مختلفة ومتعددة من أداء المتعلمين.
4. يحدد بشكل واضح مواطن القوة والضعف في أداء المتعلم، ومدى ما أحرزه من تقدم.

4- البورتفوليو:

يترجم مصطلح (Portfolio) في الأدبيات العربية بعدة تسميات منها: البورتفوليو، الحقيبة الوثائقية، ملف أعمال الطالب، ملف التعلم، ملف الإنجاز، حقيبة التعلم، ملف الأداء، السجل النامي. وهو الجمع الهادف الموثق لأعمال الطالب؛ والتي تعكس مدى جهده وتقدمه وتحصيله وإنجازاته في مجال أو مقرر ما. هذه الأعمال أو العينات تشمل أشياء كثيرة جداً من نماذج من الواجبات، والاختبارات وكتابات وانطباعات وآراء ونقد ذاتي، وقراءات وملخصات، وجمع قصاصات ومنتجات ومشروعات، وأبحاث قام بها الطالب بشكل تراكمي وعلى مدى فترة زمنية محددة.

أ- الفوائد التربوية للبورتفوليو:

يعد استخدام وتوظيف البورتفوليو كأداة من أدوات وأساليب تقويم تعلم التلاميذ من الاتجاهات الحديثة، وأحد أساليب التقويم والتعلم الأصيل، وهو:

- يقيس مجالات متعددة من جوانب تعلم التلاميذ.
- يوثق أفضل أعمال الطالب، وجهده ونموه الأكاديمي.
- يرسخ مفهوم (التقويم للتعلم)؛ وهو استخدام التقويم الصفي لتحسين التعلم.
- يرسخ مفهوم التقويم كوسيلة للتعلم، حيث أن عملية التقويم ليست منعزلة عن عملية التعليم.
- يركز على أبعاد متعددة للتقييم، بدلاً من بعد واحد كما هو الحال في أدوات التقويم التقليدية.
- يكشف عن قدرة الطالب على تطبيق المعرفة والأداء العملي والتفكير.
- يعتبر أحد أساليب التقويم الذاتي للطالب، والذي يتيح له تأمل أعماله وتقدمه الدراسي.
- ينمي لدى الطالب القدرة على تحمل المسؤولية واتخاذ القرار والإخلاص في العمل، والاهتمام بإتقانه.
- يعتبر أحد أساليب الكشف عن المواهب الإبداعية لدى الطالب.
- يعتبر وسيلة للتفاعل بين الطالب والمعلم، وبين الطالب وبقية الطلاب في حالة الاشتراك في ملف أو عرض ملفه للآخرين.
- يستخدم كأداة للتقويم المستمر والتقويم النهائي.
- ينمي استراتيجيات التعلم الذاتي، وتقدير الذات، وتنمية مهارات التفكير الواعي والناقد ومهارات البحث وحل المشكلات.
- يعتبر بمثابة مرجع موثق للطالب ونافذة لأعماله وجهوده.

ب- أبرز محتويات البورتفوليو:

يمكن أن يشتمل البورتفوليو على نطاق واسع من المحتويات بحسب موضوع الملف والأهداف المرجوة منه، ومما يمكن أن

يتضمنه:

- مقدمة لصاحب الملف يبين فيها هدف الملف والأغراض التي يريد أن يحققها من خلاله، والمبررات لما سوف يتضمنه الملف من محتويات وطريقة تنظيم الملف وترتيبه.
- صفحة جدول للمحتويات.
- عينات من أعمال وأنشطة الطالب المنتقاة.

- كتابات الطالب من مقالات وملخصات وتقارير.
- قراءات خارجية مختارة مع بيان سبب اختيارها وكيفية الاستفادة منها.
- صفحات من التأملات والانطباعات الذاتية للطالب.
- أنشطة تم إنجازها بصورة تعاونية تشاركية صفية أو لا صفية.
- أوراق الواجبات والاختبارات.
- مسودات معينة من أعمال الطالب ومشاريعه.
- قوائم بالمصادر التي اطلع عليها الطالب.
- مواد بصرية ووسائط إلكترونية من صور وملفات وفيديو وملتيميديا.
- معايير أو محكات تقويم البورتفوليو.
- محتويات اختيارية من قبل الطالب بشرط أن تكون متعلقة بموضوع وأهداف الملف (مقولات مختارة، مراجع ومصادر مشروحة).

ج- عوامل مهمة لإنجاح البورتفوليو:

- توعية الطالب بنوع الملف المطلوب وبأهدافه وأهميته، وما ينبغي أن يتضمن من محتويات.
- تدريب الطالب على بناء وتطوير الملف وفق تصور معين، وإشراكه في اختيار محتويات الملف.
- تحديد المدة الزمنية المطلوبة للبدء في عمل الملف وتنميته ووقت الانتهاء وتسليمه.
- متابعة المعلم الدورية لبناء ملف الطالب (بشكل يومي أو أسبوعي)، ومسايرته لتقديم ونمو ملف الطالب وتزويده بما يلزم من تغذية راجعة ونصائح واقتراحات.
- تشجيع الطالب على كتابة التفكير التأملية، وتوثيق انطباعاته حيال موضوعات وقضايا المادة الدراسية أو المقرر.
- تشجيع الطالب على كتابة جمل شارحة أو توضيحية عن العمل المختار، وتاريخ إنجازه وكيف تم إنجازه.
- إعلام الطالب بشكل مسبق بمحكات ومعايير تقويم الملف ومناقشته فيها والاتفاق عليها.
- المكافأة على مقدار الجهد المبذول، بالإضافة إلى إتقان ما تم تحصيله من معارف ومفاهيم ومهارات واتجاهات والتي يعكسها ملف الطالب.
- تحفيز الطالب على الإبداع والابتكار في تناول ملفه، واستشعاره بملكته ومسؤوليته تجاه تعلمه.
- إشراك البيت وتفعيل دور أولياء أمور الطلاب في معرفة وتشجيع وتقديم العون لأبنائهم، بما يسهل تطوير وإنجاز ملفاتهم.

د- أبرز المشكلات المتعلقة بتوظيف البورتفوليو:

- القصور المعرفي للبورتفوليو من قبل معظم المعلمين وعدم إدراكهم لفلسفته وأهميته التربوية وكيفية تقويمه.
- وجود اتجاهات سلبية نحو الطرق الحديثة؛ سواء في أساليب التدريس أو في أساليب التقويم من قبل بعض المعلمين وحتى الطلاب أنفسهم، لما تتطلبه هذه الأساليب الحديثة؛ ومنها (البورتفوليو) من جهد أكثر ومسؤوليات أكبر.
- عدم قناعة بعض أولياء الأمور بالبورتفوليو كأداة للتقييم.
- الاعتقاد الخاطئ بأن البورتفوليو عبارة فقط عن حشد وجمع للأوراق والأشياء.

هـ- كيفية تصحيح وتقدير الدرجة الممنوحة للبورتفوليو:

- يكون تقييم بورتفوليو الطلاب من خلال استخدام وتصميم ما يعرف بميزان التصحيح (التقدير) الوصفي، والخطوات المتبعة في تصميم ميزان تصحيح مناسب للبورتفوليو هي كالآتي:
1. تحديد المحكات والجوانب والمهارات التي ستكون هي المراد تقويمها وفقاً لأهداف وموضوع ونوع التقويم.

2. تحديد مستويات الأداء المتوقع لكل جانب (محك).
3. شرح وتوصيف المعايير المحققة لكل مستوى من المستويات في كل جانب أو محك.
4. تصميم ميزان وصفي ومتدرج للأداء على شكل جدول، وتوضع فيه المحكات أو الجوانب المراد قياسها بشكل رأسي، ويوضع مقابلها مستويات الأداء بشكل أفقي ومتدرج ومتتابع، وكل مستوى يستحق درجة محددة، مع إعطاء وصف مفصل لكل خانة من خانات المحكات.

5- السجل القصصي:

السجل القصصي عبارة عن وصف قصير من المعلم ليسجل ما يفعله الطالب والحالة التي تمت عندها الملاحظة، أو هو وصف لسلوك يقوم به الطالب سواء كان هذا السلوك لفظيا أو عمليا. وكثيرا ما يلاحظ المعلم في غرفة الصف أو خارجها مثل هذه السلوكات، ويفضل أن يسجلها على بطاقات خاصة يصف فيها السلوك وزمن القيام به ومن قام به، بالإضافة الى تعليق المعلم أو الملاحظ على هذا السلوك إذا كان يستحق التعليق فقط.

فقد تكون هذه الملاحظات ذات علاقة بكتابه المدرسي المقرر، أو بأسلوب التدريس، أو بالطالب نفسه.

وبالتالي فإن تجميعه في ملف خاص لكل طالب يمكن أن يكون له فائدة عملية في مجالات متعددة، فهي تخدم برنامج الإرشاد والتوجيه في المدرسة، وتخدم برنامج مجالس أولياء الأمور والمعلمين وتطوير المناهج وأساليب التدريس.

أ- خطوات تصميم السجل القصصي:

1. ملاحظة سلوك المتعلم.
2. تسجيل الاحداث بطريقة وصفية.
3. تحديد التاريخ والمكان.
4. التعرف على النمط السلوكي الذي يتكرر حدوثه.
5. وضع فروض على سلوك المتعلم.
6. اختبار الفروض في ضوء الأنماط المتكررة، ليرى المعلم ما إذا كانت الملاحظات تحقق الفروض.
7. توثيق اسم الملاحظ.
8. إضافة بعض التفسيرات للسلوك سواء كان إيجابيا أو سلبيا.

ب- خصائص السجل القصصي:

- يعطي مؤشرات صادقة في التعرف على مهارات واهتمامات المتعلم وسلوكه وشخصيته بشكل عام.
- يوظف لأغراض تنبئية، أو إرشادية توجيهية، أو علاجية.
- يتطلب وقتا طويلا للكتابة والمتابعة والتفسير.

ج- دور المعلم في تطوير واستخدام السجل القصصي:

- إعداد طريقة للرصد عند إكمال السجلات.
- تحديد الملاحظات المهمة أو ذات الدلالة للمتعم.
- توثيق الملاحظة بأسرع ما يمكن لعد الملاحظة.
- تفسير المعلومات المسجلة، للمساعدة في تخطيط الخطوات اللاحقة للمتعم.

خامسا: أنواع مهمة للإختبارات التحصيلية

1- الاختبارات معيارية المرجع.

2- الاختبارات محكية المرجع.

3- الاختبارات الأدائية (العملية).

4- الاختبارات المقننة.

5- الاختبارات المقالية.

6- الاختبارات الموضوعية.

1- الاختبارات معيارية المرجع:

وفها يتم إصدار الحكم على أداء الفرد عن طريق مقارنة أدائه بأداء الأفراد الآخرين على المقياس نفسه أو الاختبار المستعمل، أي أنّ هذا النوع من التقويم يهتم بمقارنة أداء الفرد بأداء أقرانه من المجموعة التي ينتمي إليها للحصول على معنى للدرجة التي تحصل عليها.

مثال: الدرجة التي تحصل عليها المفحوص دليل على مدى ارتفاع أو انخفاض فعاليته مع درجة مجموعته التي من خلالها نستطيع الوقوف على علامة كل فرد من أفراد هذه المجموعة. إذا حصل مفحوص ما في اختبار ما على 30 من 50 فإنّ الدرجة 30 لن تكون لها قيمة إلا عند الرجوع إلى أداء مجموعته في نفس الاختبار. ومن هنا نصل إلى الوضع النسبي لجودة المفحوص بهذا المقياس. تتميز الاختبارات مرجعية المعيار بـ:

- ✓ الكشف عن الفروق الفردية حيث تكون الدرجات قريبا من المنحنى الاعتدالي، حيث يفسر الأداء في الاختبار مرجعي المعيار موازنة بأداء الأفراد الآخرين في الاختبار نفسه، أي أنّ الأداء نسبي.
- ✓ استخدام اجراءات مقننة في تطبيق ورصد درجات المفحوصين في نفس الفئة (العمر، الجنس...)
- ✓ يتم من خلال هذا الاختبار تفسير سلوك المفحوص من حيث خصائصه وصفاته، حيث أنّ هذا الاختبار يستخدم لقياس مجال الاهتمام.

2- الاختبارات محكية المرجع:

يهدف الاختبار محكي المرجع إلى تقدير أداء الفرد بالنسبة إلى محك أو مستوى أداء مطلق دون اللجوء إلى مقارنة أدائه بأداء الأفراد الآخرين.

في الاختبار مرجعي المحك يتطلب تحديد مستويات مسبقة للأداء، وذلك لما يحتوي من معلومات محددة ومفصلة عن المفحوصين بالنسبة إلى مجال أو موضوع معين، ويستخدم في الغالب لوصف تقدمهم فيه.

تتميز الاختبارات محكية المرجع بـ:

- ✓ الكشف عن مستوى المفحوص بشكل مطلق دون التقيد بموازنته بأداء الأفراد الآخرين.
- ✓ معرفة مدى إتقان الفرد عن موضوع أو مهارة ما، كما يصف مدى تقدمه في مجال ما تحديدا.
- ✓ التطلع إلى معرفة مدى تمكن المفحوص من حيث الوصول إلى المستوى المرجو لديه فيما يفعل، بغض النظر عمّا يحصل عليه من درجات.

- ✓ الاهتمام بتصنيف المفحوصين إلى مستويين فقط هما: مستوى الذين استطاعوا الوصول إلى مستوى الأداء المطلوب من التمكن والذين لم يتمكنوا من الوصول إليه.
- ✓ السعي نحو تحديد مجال سلوكي تحديدا واضحا.
- ✓ تحديد العلاقة بين أداء الفرد ومستواه الواقعي في المجال السلوكي.
- ✓ يجب أن يكون الاختبار محكي المرجع متعلقا بالوظيفة التي يقيسها.
- ✓ درجات أو رتب الأفراد على المحك يجب ألا تتأثر بالشخص الذي يقوم بإعطاء الدرجة أو التقدير.

3- الاختبارات الأدائية (العملية) :

- يعرف اختبار الأداء بأنه اختبار يتطلب عادة استجابة يدوية أو استجابة حركية عموما يقوم بها الفرد.
- الاختبارات الأدائية غالبا ما تكون فردية، إذ يصعب توفير مجموعة من الأجهزة والمواد تكفي للإجراء الجمعي، كما يصعب ضبط الموقف، ويكلف الكثير، ويحصل الغش بالمحاكاة.
- تطبق اختبارات الأداء أحيانا على مجموعات صغيرة من الأشخاص، وفي مثل هذه الحالة يوضع حاجز على مقاعد الاختبار كي لا يغش الطلبة.

أنواع الاختبارات العملية:

- اختبار الورقة والقلم: في عدد من الحالات قد يساعد اختبار الورقة والقلم العملي على قياس نواتج للتعليم لها أهميتها التربوية.
- مثال: رسم خريطة جغرافية، رسم أجزاء لنبات معين، الكتابة بخط الرقعة.
- اختبارات التعرف: تهدف هذه الاختبارات إلى قياس قدرة المتعلم على التعرف على الخصائص الأساسية لأداء معين، أو نتيجة أداء معين أو التعرف على بعض الأشياء.
- مثال: العينات البيولوجية، عزف قطعة موسيقية على إحدى الآلات ويطلب من المتعلم بيان الأخطاء أو النغمات في عزف القطعة، وفي أنواع أخرى من الاختبارات قد يتطلب من المتعلم ان يتعرف على أجزاء معينة من الأجهزة ووظائفها، أو اختيار الأداة المناسبة أو الجهاز المناسب لعمل محدد.
- اختبارات المحاكاة: تصمم هذه الاختبارات العملية أحيانا لمحاكاة موقف من المواقف الحقيقية، وذلك بغرض عزل هذا الموقف بطريقة تمكن المتعلم من القياس بنفس الحركات التي يتطلبها الموقف الحقيقي، ولكن تحت ظروف مزيفة، أو غير حقيقية.
- مثال: في التربية البدنية (القيام بحركات السباحة خارج الماء)، وفي القانون (تمثيل محاكمة، وقيام المتعلم بدور المحامي أو القاضي).
- اختبارات عينات العمل (أداء المهمات الفعلية): تتكون هذه الاختبارات من موقف يمثل موقفا حقيقيا لمجال العمل ويطلب من المتعلم أداء المهمات الفعلية لهذا العمل.
- مثال: اختبار قياده السيارات؛ حيث يقوم المتعلم بأهم الاعمال التي تبين إلمامه بالقيادة في ظروف مختلفة.

4- الاختبارات المقننة :

ويقصد بالاختبارات المقننة الاختبارات التي يقوم بها خبراء في القياس، وهي تتيح الفرصة لاستخدام طرائق وأدوات الحصول على عينات من السلوك باستخدام إجراءات منتظمة التي تعني أن محتوى الاختبار نفسه يطبق بالتعليمات والتوقيت نفسه المحدد للأداء، كما أن طريقة احتساب النتائج تتضمن إجراءات منتظمة وثابتة وبصورة موضوعية، فضلاً عن توافر المؤشرات الأساسية

للاختبار الجيد مثل الصدق والثبات، كما أن هذه الاختبارات تكون في الغالب قد جرى تطبيقها على مجموعة معيارية (أو مجموعات معيارية) حتى يمكن تفسير أداء الفرد في ضوء هذه المعايير .

ويمكن تصنيف الاختبارات المقننة طبقاً لكثير من وجهات النظر المختلفة، إلا إن التصنيف الشائع هو التصنيف وفقاً لما يقيسه الاختبار، والمتمثل في:

أ- اختبارات القدرات (القدرة العامة والقدرات الحركية والقدرات الخاصة).

ب- اختبارات التحصيل (اختبارات التنبؤ والاختبارات المرتبطة بنشاط معين).

ت- اختبارات الميول والشخصية والاتجاهات.

واختبارات القدرات والتحصيل غالباً ما تتضمن اختبارات لقياس الأداء الأقصى، أما اختبارات الشخصية، فإنها تتضمن اختبارات الأداء المميز.

كما يرى بعض العلماء أن اختبارات القدرات واختبارات التحصيل هي اختبارات معرفية، أما اختبارات الميول والشخصية والاتجاهات فهي اختبارات غير معرفية، لأن هذه الاختبارات لا تتميز بوجود إجابات صحيحة أو خاطئة على بنود أو عبارات الاختبار.

5- الاختبارات المقالية:

تعتبر من الأنماط التقليدية الشائعة منذ زمن بعيد، وتستخدم من أجل تقييم تحصيل الطالب في الموضوعات الدراسية التي يقوم بتعلمها. ويمثل الاختبار المقالي تقويماً للقدرة على التعبير اللغوي، والإبداع الفكري وتنظيمه وتكامله، وإبداء الرأي وتقديم الحجة المناسبة، وهو الأسلوب الأمثل الذي يعبر فيه الطالب عن وجهة نظره والتعبير عنها بالطريقة التي يريد. وهو اختبار كتابي يطلب ممن يؤديه كتابة مقال أو موضوع إنشائي يتحدد حسب ما يطلب في السؤال، والذي قد يبدأ بكلمة: ناقش، تحدث عن، ابحث في وغيرها. ويستطيع الطالب إطلاق العنان لقلمه والاسترسال في الكتابة، مع مراعاة الصحة في التعبير والدقة في استخدام المفاهيم والأفكار والمصطلحات والقواعد العلمية، والقدرة على العرض والشرح والتحليل والاستنباط، وربط المعلومات بعضها ببعض.

مزايها:

- تعطي للتلميذ الحرية اللازمة لإبراز معارفه وقدرته على تنظيمها بشكل متجانس، وتقديم التوضيحات المهمة؛ ولها قدرة في قياس عميق لتعلمات معقدة.
- تسمح بإظهار مهارات أخرى وجوانب مهمة من شخصية المتعلم.
- تمكن من الاستخدام الجيد للأساليب اللغوية، من حيث الكلمات المنتقاة والقواعد التي يخضع لها التركيب اللغوي للجملة، لتؤدي بمجموعها إلى وضوح الفكرة والعرض المنظم لها.
- مجال التخمين ضئيل في هذا النوع من الاختبارات.
- غير مكلفة.
- تعتبر نافعة بشكل واضح مع مراحل التعليم الأعلى.
- يتم إعدادها بسرعة وسهولة، مما يوفر الكثير من وقت وجهد المعلم.
- تساعد على التمييز بين التلاميذ الذين تقوم دراستهم على الفهم، والتلاميذ الذين تقوم دراستهم على الحفظ دون فهم واستيعاب.
- تهتم بقياس قدرة الطالب الذاتية على التفكير والإبداع.

عيوبها:

- لا تسمح لنا بقياس تعلمات متعددة في الوقت نفسه، بل تغطي جانباً محدوداً من المقرر الدراسي.

- تلعب فيها الصدفة أو الحظ دورا كبيرا.
- تتيح للتلميذ التحايل والمراوغة عندما لا يعرف شيئا عن الموضوع، فيحصل على بعض الدرجات بدلا من الحصول على الصفر.
- من الصعب التحكم في مستوى صعوبتها.
- تتأثر بعدة عوامل كخط التلميذ وأسلوبه وحسن ترتيبه لورقة إجابته، كما تتأثر بالهالة.
- تتصف بعض أسئلتها بالغموض.
- ينفق فيها التلميذ معظم وقته في الكتابة أكثر من التفكير المنظم.
- يتقلص فيها مدى الدرجات.
- تشجع على الحفظ.
- تصحيحها ذاتي وبطيء وصعب وغير ثابت.
- يصعب إخضاع نتائجها لطرق البحث والإحصاء، بسبب صعوبة وضع معايير واضحة لأداء التلاميذ على أسئلتها.

6- الاختبارات الموضوعية:

وتعتبر أكثر الاختبارات التحصيلية شيوعا واستخداما لدى المعلمين، باعتبارها إحدى وسائل التقويم المتبعة. وهي أساليب أكثر تطورا في قياس التحصيل الدراسي، ولقد نشأت فكرة الاختبار الموضوعي عندما نشأت فكرة القياس في الاختبارات العقلية والنفسية. وعندما ظهرت مفاهيم تقدير كفاءة الأدوات من حساب للثبات والصدق والمعايير وغيرها، ولقد وضعت لتلافي نقص الاختبارات المقالية. وغالبا ما تتضمن أربعة نماذج من الأسئلة وهي: الصح والخطأ، الاختيار من متعدد، التكملة، المزاوجة. وتنسب لها صفة الموضوعية لأن بنودها تمثل بدرجة كبيرة الموضوع المراد قياسه، وهذه البنود تكون دقيقة وتتطلب إجابات دقيقة ومحددة، ومن جهة أخرى اتصافها بموضوعية التصحيح؛ لأن درجة التلميذ لا تتأثر بذاتية المصحح.

مزاياها:

- عملية تصحيحها تكون بسيطة وسريعة.
- صالحة لقياس تذكر الحقائق، وبعض أنواعها صالحة لقياس الفهم ومهارات التفكير المتوسطة وبعض المردودات التعليمية المتقدمة.
- تغطي عينة واسعة لمحتوى المساق، وهذا لإمكانية شمولها على عدد كبير من الأسئلة.
- تحدد نوع الاستجابة المطلوبة بدقة، وتمنع المتاهات في الإجابة أو التفتن بالمهارة الكتابية.
- تشجع المتعلم على المعرفة الشاملة للحقائق، وتنمي القدرة على التمييز الدقيق بينها.
- تلعب دورا مهما في التقويم التكويني، حيث توفر فرصة سريعة لجمع معلومات حول تعلم الطالب، وحول الأخطاء الشائعة، ونقاط القوة والضعف لديه.
- تكشف عن قدرة الطالب على سرعة التفكير بعيدا عن الحفظ والاستظهار.
- الدرجات فيها لا تتأثر بذاتية المفحوص.

عيوبها:

- بناء هذا النوع من الاختبارات صعب، ويتطلب جهدا كبيرا ووقتا طويلا.
- غير صالحة لقياس القدرة على تنظيم الأفكار أو انتقائها، ولا على قياس بعض أنواع مهارات حل المشكلات والاكتشاف.

- إجابة المتعلم فيها تكون خاضعة للتخمين أحيانا.
- تتطلب مجرد التعرف على الإجابة الصحيحة، بدلا من استدعائها أو بنائها.
- تؤكد على مهارات القراءة فقط، دون مهارة الكتابة رغم أهميتها.
- صعوبة وضع مفردات تقيس عمليات عقلية عليا؛ كالتحليل والتركيب.
- لا تقيس جوانب الأداء الإبداعي للطالب.
- تتعامل مع الطلبة كآلات يبحثون عن الإجابة فقط.
- تشجع على التفكير المتسرع.
- تقلل من فرص الطلبة المتأملين في الحصول على الدرجات التي تعكس مستواهم التحصيلي.

سادسا: خطوات بناء الإختبار التحصيلي

- 1-تحديد الأهداف (النواتج) التي يهدف المقرر الذي يقوم المعلم بتدريسه إلى تحقيقها.
- 2-تحديد محتوى الاختبار.
- 3-تحديد نوع مفردات الاختبار.
- 4-تحديد عدد مفردات الاختبار.
- 5-كتابة مفردات الاختبار ومراجعتها.
- 6-كتابة تعليمات الاختبار (التعليمات الخاصة بالمفحوصين).
- 7-إعداد الصورة التجريبية للاختبار (الأولية).
- 8-إعداد مفتاح تصحيح الاختبار (نموذج الإجابة).
- 9-تطبيق الصورة التجريبية للاختبار (الأولية).
- 10-تصحيح الاختبار ورصد الدرجات على نحو يسمح بمعالجتها إحصائيا.
- 11-القيام بتحليل المفردات.
- 12-إجراء التحليلات الإحصائية اللازمة لتقدير معاملات ثبات وصدق الاختبار تقديرا مبدئيا.
- 13-القيام بالتعديلات اللازمة في الصورة الأولية للاختبار.
- 14-إعادة ترتيب المفردات.
- 15-تطبيق الصورة التجريبية المعدلة للاختبار على مجموعة أخرى درست المقرر نفسه.
- 16-تصحيح الصورة التجريبية المعدلة للاختبار ورصد الدرجات على نحو يسمح بمعالجتها إحصائيا.
- 17-القيام بتحليل المفردات.
- 18-إجراء التحليلات الإحصائية اللازمة لتقدير معاملات ثبات وصدق الاختبار.

اختلف الباحثون في تحديد خطوات بناء الاختبار التحصيلي رغم اتفاقهم في نقطة البداية والنهاية وسوف يتم التعرض إلى الخطوات التي اقترحها "علي ماهر خطاب" لاتصافها بالدقة والتفصيل، وهي كالآتي:

1- تحديد الأهداف (النواتج) التي يهدف المقرر الذي يقوم المعلم بتدريسه إلى تحقيقها:

يتم تحديد هذه الأهداف عندما يخطط المعلم للمقرر وليس عندما يبدأ في تحديد محتوى وشكل الاختبار، ويجب أن تصاغ هذه الأهداف في صورة عبارات سلوكية تحدد نوع السلوك الذي يتوقع أن يؤديه المتعلم بعد انتهائه من دراسة وحدة دراسية معينة أو موضوع معين.

ويجب أن يستعين المعلم في تحديد هذه الأهداف بالبحوث والكتب الخاصة بمادته، ويرأي وخبرة زملائه، وبالأهداف التي تحددها الوزارة.

2- تحديد محتوى الاختبار:

يتم في هذه الخطوة تحديد الموضوعات التي يغطيها الاختبار في ضوء الأهداف التي يسعى الاختبار إلى تحقيقها، ومن وسائل ذلك عمل جدول مواصفات ويسمى كذلك بخطة الاختبار وهو جدول ثنائي البعد يتضمن (الموضوعات) و(الأهداف التعليمية للمقرر الدراسي) ويحتوي على الأهمية النسبية (الوزن النسبي للموضوعات والأهداف). استخدام جدول المواصفات يزيد من احتمالية تمثيل الاختبار للجوانب الهامة للمقرر الدراسي، وهذا يرفع من صدق محتوى الاختبار، كما أنه يعد كموجه لمعد الاختبار في اختيار الأفكار التي يجب أن يتضمنها الاختبار. إعداد جدول المواصفات يعد من أفضل يعد من أفضل الأنشطة التي يجب أن يقوم بها المعلم ويهتم بها.

خطوات إعداد جدول المواصفات:

- تحديد الأهداف التعليمية للمقرر الدراسي (نواتج التعلم) والأوزان النسبية لكل منها والتي تعكس الاهتمام الذي تحظى به عملية التدريس، وبعد تحديدها تكتب أعلى أعمدة جدول المواصفات أي توضع أفقياً أعلى الجدول.
- تحديد موضوعات المقرر الدراسي ونسبة تمثيل كل منها، ولكي يتسنى للمعلم أو معد الاختبار تحديد الأوزان النسبية لموضوعات المقرر الدراسي يمكنه الاستعانة بالموجهات الآتية:

- الزمن المخصص لتدريس كل موضوع من موضوعات المقرر الدراسي.
 - درجة الاهتمام الذي يحظى به كل موضوع في عملية التدريس.
 - خبرته الشخصية في تدريسه للمقرر الدراسي وتحليل محتواه.
 - استشارة الخبراء في تحديد الأوزان النسبية للموضوعات والأهداف.
- وبعد تحديد الموضوعات وأوزانها النسبية تكتب أفقياً على صفوف الجدول. وينشأ عن تقاطع الأعمدة (الأهداف) والصفوف (الموضوعات) عدد من الخلايا (الخانات) والتي تحدد نسبة الأسئلة التي يتضمنها الاختبار بالنسبة لكل موضوع من موضوعات المحتوى في علاقته بكل هدف من الأهداف.

ولتحديد الأهمية النسبية لكل خلية من خلايا جدول المواصفات يتم اتباع الخطوات الآتية:

- تحديد وضع الخلية.
- تحديد الصف (الموضوع) الذي يتقاطع مع الخلية.
- تحديد النسبة المئوية الكلية للصف.
- تحديد العمود (الهدف) الذي يتقاطع مع الخلية.
- تحديد النسبة المئوية الكلية للعمود.
- ضرب قيمة النسبة المئوية الكلية للصف بقيمة النسبة المئوية للعمود، ويمثل الناتج نسبة الأسئلة التي تمثل الخلية والتي يجب أن يتضمنها الاختبار.

مثال:

نفترض أننا بصدد إعداد جدول مواصفات في مادة الإحصاء. احتوى المقرر الدراسي على أربعة (4) موضوعات لتحقيق مجموعة من الأهداف التعليمية موزعة على أربعة (4) مستويات، ويوضح الجدول الآتي الأهمية النسبية لكل موضوع ولكل هدف.

المطلوب: أكمل إعداد هذا الجدول موضحاً الأهمية النسبية لكل خانة.

جدول يوضح الأهمية النسبية لكل موضوع ولكل هدف.

المجموع	الهدف الرابع	الهدف الثالث	الهدف الثاني	الهدف الأول	الأهداف الموضوعات
	تحقيق المبادئ	فهم المبادئ	معرفة الحقائق النوعية	معرفة المصطلحات	
23%					مقاييس النزعة المركزية
24%					مقاييس التشتت
27%					مقاييس الارتباط
26%					الفروق بين المتوسطات
100%	30%	40%	16%	14%	المجموع

3- تحديد نوع مفردات الاختبار:

يجب على المعلم تحديد نوع المفردات التي سوف يتألف منها الاختبار، وأن يكون ملماً بأسس صياغة المفردات وعيوبها ومزاياها. ✓ وهناك عدة أنواع للأسئلة أو المفردات التي تتألف منها الاختبارات منها: اختيار إجابة من إجابتين، الاختيار من متعدد، المزوجة، التكملة (الإجابات القصيرة- الاستدعاء)، أسئلة الترتيب.

4- تحديد عدد مفردات الاختبار:

يجب أن يكون عدد المفردات يعكس الأوزان النسبية لخلايا جدول المواصفات التي تم تحديدها في الخطوة الثانية، ومناسبا للزمن المخصص لتطبيق الاختبار.

مثال:

بالعودة إلى المثال السابق والجدول الذي يوضح جدول المواصفات لمقياس الإحصاء، نفترض أن هذا لاختبار مؤلف من 100 مفردة. المطلوب: اعتمادا على الأهمية النسبية لكل خانة، احسب عدد المفردات في كل خانة.

5- كتابة مفردات الاختبار ومراجعتها:

يجب كتابة عدد كاف من المفردات أكثر مما هو متوقع أن تتضمنه الصورة النهائية للاختبار، وذلك لضمان الحصول على العدد المطلوب من المفردات الجيدة بعد تجربة الاختبار. بعد الانتهاء من كتابة المفردات يراجع معد الاختبار المفردات لغويا للتأكد من سلامة اللغة وسهولة التعبير وبساطة المفردات لإزالة الغموض وتقوية البدائل.

6- كتابة تعليمات الاختبار (التعليمات الخاصة بالمفحوصين):

تهدف هذه التعليمات إلى توضيح ما يجب على المفحوصين أدائه للإجابة عن مفردات الاختبار، ويجب أن تطبع هذه التعليمات في بداية الاختبار، ويجب أن تكون هذه التعليمات بسيطة، صريحة، واضحة، محددة.

7- إعداد الصورة التجريبية للاختبار (الأولية):

8- إعداد مفتاح تصحيح الاختبار (نموذج الإجابة):

9- تطبيق الصورة التجريبية للاختبار (الأولية):

يتم تطبيقها على مجموعة من الأفراد (فصل دراسي أو مجموعة فصول دراسية) التي يهدف المعلم إلى قياس تحصيلها في المقرر الدراسي الذي قام بتدريسه لهم.

10- تصحيح الاختبار ورصد الدرجات على نحو يسمح بمعالجتها إحصائياً.

11- القيام بتحليل المفردات :

والهدف من هذا الإجراء هو دراسة وتحديد خصائص مفردات الاختبار وذلك من خلال حساب (معاملات الصعوبة، معاملات التمييز، صدق المفردات، ثبات المفردات).

12- إجراء التحليلات الإحصائية اللازمة لتقدير معاملات ثبات وصدق الاختبار تقديراً مبدئياً.

13- القيام بالتعديلات اللازمة في الصورة الأولية للاختبار.

14- إعادة ترتيب المفردات.

15- تطبيق الصورة التجريبية المعدلة للاختبار على مجموعة أخرى درست المقرر نفسه.

16- تصحيح الصورة التجريبية المعدلة للاختبار ورصد الدرجات على نحو يسمح بمعالجتها إحصائياً.

17- القيام بتحليل المفردات.

18- إجراء التحليلات الإحصائية اللازمة لتقدير معاملات ثبات وصدق الاختبار.

سابعاً: تحليل مفردات الاختبار التحصيلي

1- تعريف تحليل المفردات.

2- الهدف من تحليل المفردات.

3- معامل التمييز.

4- معامل الصعوبة.

1-تعريف تحليل المفردات :

تحليل المفردات هو إجراء يستخدم للحكم على جودة مفردات الاختبار. ويقصد به كذلك عملية فحص استجابات المفحوصين لكل مفردة من مفردات الاختبار.

2-الهدف من تحليل المفردات :

وفي هذا الصدد يذكر "بورج" و"جال" (Borg & Gall) أن الهدف الأساسي من تحليل المفردات هو اختيار المفردات الصادقة والثابتة ذات معاملات الصعوبة الملائمة، وتحديد المفردات التي تحتاج إلى مراجعة أو تعديل (تحديد المفردات الجيدة والردئية). ويعتمد تحليل المفردات على طبيعة الاختبار، ففي حالة الاختبارات التحصيلية مرجعية المعيار يكون الاهتمام في تحليل المفردات منصبا على مستوى صعوبة كل مفردة وقدرتها على التمييز بين الأفراد ذوي الأداء المرتفع وذوي الأداء المنخفض.

3- معامل التمييز:

معامل التمييز هو مقياس مدى قياس المفردة لما يقيسه الاختبار ككل، كما يعتبر مقياسا لمدى قدرة المفردة على قياس الفروق الفردية بين الأفراد. وتقدر قيمته بالفرق بين نسبة من أجابوا إجابة صحيحة على المفردة في مجموعة الأداء العالي على الاختبار ونسبة من أجابوا إجابة صحيحة على المفردة في مجموعة الأداء المنخفض على الاختبار، وذلك طبقا للتقسيم الذي يحدده واضع الاختبار.

يمكن أن يكون التقسيم إلى المجموعة (50 % العليا) والمجموعة (50 % الدنيا) اعتمادا على درجات الاختبار الكلية إذا كانت مجموعة المفحوصين صغيرة.

أما عندما يكون حجم العينة كبيرا، فإن أفضل التقسيمات هو (27 % العليا و 27 % الدنيا) من المجموعة الكلية، كما يستخدم البعض التقسيم (25 % العليا) و(25 % الدنيا) من المجموعة الكلية.

تتراوح قيمة معامل التمييز بين (-1) و (+1)، ويتم حساب معامل التمييز باستخدام المعادلة الآتية:

$$D = \frac{TU - TL}{n}$$

بحيث:

D: معامل التمييز.

TU: مجموع الإجابات الصحيحة في مجموعة الأداء العالي.

TL: مجموع الإجابات الصحيحة في مجموعة الأداء المنخفض.

n: عدد الأفراد في إحدى المجموعتين.

يذكر "أهمان" و"جلوك" (Ahman & Glock) أنه:

- إذا كانت قيمة معامل التمييز أكبر أو تساوي (0,40) فهو معامل تمييز جيد جدا.
- إذا كانت قيمة معامل التمييز محصورة بين (0,20) و (0,40) فهو معامل تمييز مرض.
- إذا كانت قيمة معامل التمييز محصورة بين الصفر (0) و (0,20) فهو معامل ضعيف أو محل تساؤل.

خطوات حساب القوة التمييزية للمفردة (معامل التمييز):

تسير خطوات حساب القوة التمييزية للمفردة كالآتي:

- ترتيب أوراق الإجابة حسب الدرجة الكلية على الاختبار ترتيبا تنازليا.
- تحديد مجموعتي المقارنة الطرفية (مجموعة الأداء العالي ومجموعة الأداء المنخفض) طبقا لمحك التقسيم المتبني.
- حساب وتسجيل عدد الأفراد الذين أجابوا إجابة صحيحة على كل مفردة في كل مجموعة.
- حساب معامل التمييز لكل مفردة.

مثال:

يوضح الجدول الآتي عدد الإجابات الصحيحة في عدد من المفردات في مجموعتي الأداء العالي والأداء المنخفض.

المطلوب: احسب معامل تمييز هذه المفردات. علق على هذه القيم.

المفردة	مجموعة الأداء العالي (n=50)	مجموعة الأداء المنخفض (n=50)
	عدد الإجابات الصحيحة	عدد الإجابات الصحيحة
1	40	20
2	15	18
3	15	05
4	20	04
5	30	40

فعالية البدائل:

عند دراسة معامل التمييز يجب دراسة المجموع الكلي للاختيارات التي تلقاها كل بديل لأن:

- البديل الذي لا يختاره أي أحد من الأفراد دليل على عدم فاعليته أو وضوح خطئه فهذا المشتت لا فائدة منه.
 - البديل الذي يختاره معظم الأفراد وخاصة الفئة العليا ولم يختاروا الإجابة الصحيحة، فإن هذا ربما يكون هو الإجابة الصحيحة بالفعل وليس التي حددها معد الاختبار، وهنا يجب مراجعة الإجابة الصحيحة والتأكد من صحتها وعدم صحة المشتت.
 - يجب أن تكون للبدائل الجاذبية نفسها لدى المستجيب الذي لم يكن قد سبق له أن درس الموضوع الذي يجري فيه الاختبار دراسة وافية.
- لتحليل المفردات في حالة الاختيار من متعدد، نستخدم النماذج الآتية كأمثلة لدراسة إلى أي مدى يختار المفحوصون البدائل كإجابات صحيحة.

مثال (1):

		الاختيارات				
		أ	ب	ج	د	هـ
n ₁ = 100 n ₂ = 100	مجموعة الأداء العالي	10	12	50	7	21
	مجموعة الأداء المنخفض	13	17	31	11	27

$D = \frac{50-31}{100} = 0,19$ (معامل تمييز موجب ولكنه منخفض)

البدائل الخاطئة يجب أن تجذب مجموعة الأداء المنخفض أكثر مما تجذب مجموعة الأداء العالي، ولكن يلاحظ تقارب كبير بين استجابات مجموعة الأداء المنخفض ومجموعة الأداء العالي. لذلك يفضل تعديل البدائل الخاطئة أو استبدالها.

مثال (2):

		الاختيارات				
		أ	ب	ج	د	هـ
n ₁ = 100 n ₂ = 100	مجموعة الأداء العالي	0	0	10	75	15
	مجموعة الأداء المنخفض	30	0	20	45	5

$D = \frac{75-45}{100} = 0,30$ (معامل تمييز موجب ومقبول)

رغم أن المفردة تميز تمييزاً موجباً ولكن يجب:

- استبدال أو تعديل البديل (ب).
- التأكد من أن البديل (هـ) تمثل إجابة خاطئة، وينصح باستبدال هذا البديل أو تعديله.

مثال (3):

		الاختيارات				
		أ	ب	ج	د	هـ
n ₁ = 100 n ₂ = 100	مجموعة الأداء العالي	5	84	0	11	0
	مجموعة الأداء المنخفض	14	70	4	12	0

$D = \frac{84-70}{100} = 0,14$ (معامل تمييز موجب لكنه منخفض)

معامل التمييز منخفض، لذلك يجب:

- استبدال أو تعديل البديلين (ج) و(هـ)، لأنهما لا يضيفان شيئاً لجودة المفردة.
- تعديل البديل (د) لأنه لا يميز بين مجموعتي الأداء العالي والمنخفض.

4-معامل الصعوبة:

يقصد بمعامل صعوبة المفردة نسبة الأفراد الذين أجابوا إجابة صحيحة على المفردة، ويعد أحد المعاملات الإحصائية الهامة في تحليل مفردات الاختبارات الموضوعية. غير أن البعض يشير إلى أن التسمية الصحيحة لهذا المعامل هو معامل السهولة على اعتبار أن ارتفاع هذا المعامل دليل على سهولة المفردة وليس على صعوبتها. ويقدر بـ:

$$P = \frac{TU+TL}{N}$$

بحيث:

P : معامل الصعوبة.

TU : عدد الإجابات الصحيحة في المجموعة العليا.

TL : عدد الإجابات الصحيحة في المجموعة الدنيا.

N : مجموع الأفراد في المجموعتين.

ولغرض الحكم على مستوى صعوبة الفقرة الاختبارية يمكن الاستفادة من المحكات المحددة في الجدول الآتي:

معامل صعوبة المفردة	تقويم المفردة
أقل من 0,20	صعبة جدا.
من 0,20 إلى 0,39	صعبة.
من 0,40 إلى 0,59	متوسطة الصعوبة.
من 0,60 إلى 0,79	سهلة.
0,80 فأكثر	سهلة جدا.

خطوات حساب معامل صعوبة المفردة:

تسير خطوات حساب معامل صعوبة المفردة كالآتي:

- بعد تطبيق الاختبار على الطلبة، نصحح إجاباتهم على كل مفردة، وذلك بمنح كل إجابة صحيحة درجة واحدة وكل إجابة خاطئة صفرا.
- حساب الدرجة الكلية لكل طالب على الاختبار، وذلك بجمع الدرجات التي حصل عليها في جميع مفردات الاختبار.
- ترتيب أوراق الإجابة حسب الدرجة الكلية على الاختبار ترتيبا تنازليا.
- تحديد مجموعتي المقارنة الطرفية (مجموعة الأداء العالي ومجموعة الأداء المنخفض) طبقا لمحك التقسيم المتبنى.
- حساب وتسجيل عدد الأفراد الذين أجابوا إجابة صحيحة على كل مفردة في كل مجموعة.
- حساب معامل الصعوبة لكل مفردة.

مثال:

يوضح الجدول الآتي عدد الإجابات الصحيحة في عدد من المفردات في مجموعتي الأداء العالي والأداء المنخفض.

المطلوب: احسب معامل صعوبة هذه المفردات. علق على هذه القيم.

المفردة	مجموعة الأداء العالي (n=100)	مجموعة الأداء المنخفض (n=100)
	عدد الإجابات الصحيحة	عدد الإجابات الصحيحة
1	70	60
2	90	70
3	20	05
4	60	40
5	40	20

ثامنا: 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 أيضا 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:

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

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:

- وضع الطريقة التي استخدمها الباحث للتأكد من ثبات اختبار؟ 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

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:

- وضع 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

هـ- معامل ثبات الاتساق الداخلي:

ومن المعادلات المستخدمة في هذه الطريقة معادلة "كيودورويتشاردسون" ومعادلة "ألفا كرونباخ".

أولاً: معادلة "كيودورويتشاردسون":

يمكن تقدير معامل الاتساق الداخلي باستخدام معادلة "كيودورويتشاردسون" ولها صيغ رياضية عديدة غير أنه هناك صيغتان أكثر شيوعاً واستخداماً وهما الصيغتان (20) و(21) وسيتم تناول الصيغة (20) فقط، 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_i^2 : التباين الكلي للمشاهد لدرجات الأفراد على الاختبار.

ثانيا: 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. والصيغة الرياضية لهذه المعادلة هي على النحو الآتي:

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

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^2 : التباين الكلي للمشاهد 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
ف10	ف9	ف8	ف7	ف6	ف5	ف4	ف3	ف2	ف1	
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

3-العوامل التي تؤثر في ثبات الاختبار:

توجد عدة عوامل تؤثر في ثبات الاختبار منها:

- **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.
- **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) تقريبا.
- **استقلال مفردات الاختبار:** ينبغي أن تكون المفردات مستقلة، حتى وإن كانت تقيس سمة واحدة أو قدرة واحدة.
- **تشابه محتوى مفردات الاختبار:** يؤدي زيادة تشابه محتوى مفردات الاختبار إلى زيادة قيمة معامل ثباته.
- **التخمين:** يعد التخمين من العوامل التي تؤثر في ثبات الاختبار على نحو سلبي.
- **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).
- **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).