

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

محاضرات العام الجامعي 2025 - 2026
(السراسي الرابع)

مادة القياس التربوي وبناء الاختبارات المدرسية (2)

السنة الثانية علوم التربية

الدكتورة علوطي سهيلة

أولاً: 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:

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

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:

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

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 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).

ثانيا: 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 شرطا كافيا بل لا بد من توفر شروط أخرى من أهمها الثبات، غير أن الصدق أكثر أهمية من الثبات، فقد يكون الاختبار ثابتا 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 دليلا 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 أولا، قبل عمل مقارنة بين محتوى الاختبار وجدول المواصفات.

وتفسير 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. وبعد الصدق المرتبط بالمحك أكثر أساليب الصدق وضوحاً في الممارسات الواقعية، حيث يتحدد الصدق باعتباره ما يقيسه الاختبار وفق محك خارجي، والمحك هنا مستقل تماماً 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.

أولاً: 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 نسبياً من جمع المعلومات على الاختبار التي نبحث 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.
- وتفسير 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 نسبياً (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 شيوعاً واستخداماً 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 نبناها اعتماداً على درجات الاختبار، وبالتالي فالاختبار قادر على التنبؤ بالأفراد الذين سيحققون نجاحاً 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 تقديراً مرتفعاً في مقياس المحك والأخرى أخذت تقديراً منخفضاً في مقياس المحك. فإذا ثبت أنه هناك فرق ذو دلالة إحصائية بين متوسطي درجات هاتين المجموعتين في الاختبار كان ذلك دليلاً على صدق الاختبار.

وتفسير 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 نسبياً (الفترة التي ينهي فيها الطالب دراسته أو العامل تدريبه)، ثم جمع بيانات المحك.
3. ترتيب الأفراد طبقاً 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:

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

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

$\bar{X}_1$: 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 وتباين قدره 18,49، أما المجموعة الثانية فكان المتوسط الحسابي لدرجات أفرادها 40,61 وتباين قدره 7,84. المطلوب:

- ما نوع الصدق الذي أراد الباحث الحصول عليه؟
- ماهي الطريقة التي استخدمها؟
- هل هذا الاختبار 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.

➤ 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:

أولاً: الاتساق (الانسجام الداخلي) :

ويعني أن قوة الارتباط بين الفقرات المعدة لقياس السمة تعتبر مؤشراً إحصائياً 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.

ثانياً: الصدق التمييزي:

وتتلخص هذه الطريقة في إيجاد تصميم تجريبي فيه مجموعتان متميزتان منطقياً 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.
- مقارنة الفئات المتطرفة في المقياس ذاته.

3-العوامل المؤثرة في معامل صدق الاختبار:

من العوامل التي تؤثر في صدق الاختبار نجد:

- **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.
- **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 تسربهم أو غيابهم بسبب المرض أو انتقالهم إلى مكان آخر أو عدم تعاونهم.
- **طول الاختبار:** أية زيادة في عدد مفردات الاختبار تؤدي عادة إلى ارتفاع كل من صدقه وثباته.
- **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.

ثالثاً: التقويم التشخيصي

6- تعريفه.

7- أهميته.

8- أهدافه.

9- أنواعه.

10- مراحله.

11- وسائله.

1- تعريفه :

- التقويم التشخيصي هو عملية تقييم تجرى بهدف تحديد مستوى التلاميذ في المهارات والمعارف الأساسية، ومعرفة مكان القوة والضعف لديهم.
- التقويم التشخيصي هو عملية تقييم أساسية يمكن إجراؤها في أي وقت، ولكن عادة ما تجرى في بداية العام الدراسي، أو بداية وحدة دراسية جديدة بهدف تحديد احتياجات التلاميذ التعليمية، ويتم تنفيذه باستخدام مجموعة متنوعة من الأساليب مثل الاختبارات التحريرية والشفوية، والواجبات المنزلية والمشاريع.

2- أهميته :

- يعد التقويم التشخيصي أداة هامة في العملية التعليمية، حيث يساعد على تحديد نقاط القوة والضعف لدى التلاميذ، مما يمكن الأساتذة من تخصيص التدريس وتقديم الدعم اللازم للتلاميذ المتعثرين، كما تساعد نتائج التقويم التشخيصي المعلمين على تكييف الأهداف التعليمية التعليمية مع مستوى التلاميذ، واختيار الطرق والخطط الملائمة لتحقيق تلك الأهداف.
- ويمكن أن يساهم التقويم التشخيصي في تحسين مستوى التحصيل الدراسي، وتقليل الفجوات بين التلاميذ، وتحسين جودة العملية التعليمية بشكل عام.

3- أهدافه :

- تتمثل الأهداف الرئيسة للتقويم التشخيصي في الآتي:
- وضع خطة لتحديد مشكلات التعلم لدى التلاميذ ومعرفة أسبابها.
- وضع خطة لتعلم التلاميذ الذين تبرز لديهم نقاط ضعف في مادة أو أكثر من المواد الدراسية ليتم تعليمهم تعليماً علاجياً يحسن من مستوى تعلمهم في تلك المادة أو المواد، ويقلل بالتالي من نقاط الضعف التي يعانون منها في تلك المادة أو المواد.
- استثارة نقاط القوة لدى التلاميذ في تعلم جديد، خاصة إذا ما كشف المعلم عن نقاط قوة وفهم واضح لمادة تعليمية سابقة، فيقوم بتدريسهم مادة تعليمية جديدة مبنية على المادة التعليمية السابقة ومكملة لها.
- توجيه التلاميذ نحو المجالات التي يحتاجون فيها إلى تحسين مستواهم.

- تزويد مختلف الأطراف المعنية كالمجالس التعليمية وهيئة التفيتيش التربوي، بمعلومات دقيقة عن مستوى التحصيل الدراسي للتلاميذ.
- إطلاع الأسر على نقاط القوة والضعف لدى أبنائهم، من أجل مساعدتهم على تحسين تحصيلهم الدراسي.

4- أنواعه :

يوجد نوعان من التقويم التشخيصي، وهما تقويم المكتسبات السابقة وتقويم المستلزمات.

أ - تقويم المكتسبات السابقة:

يهدف إلى تحديد مستوى التلميذ في المهارات والمعارف الأساسية، من خلال قياس مدى تمكنه من الأهداف التعليمية للبرنامج الدراسي للسنة الماضية.

ب - تقويم المستلزمات:

يهدف إلى تحديد مدى استعداد التلميذ للتعلمات الجديدة، من خلال قياس مدى تمكنه من الأهداف التعليمية للبرنامج الدراسي للسنة الحالية.

وهكذا، فإن تقويم المكتسبات السابقة يستند إلى الأهداف التعليمية للبرنامج الدراسي للسنة الماضية، بينما يستند تقويم المستلزمات إلى الأهداف التعليمية للبرنامج الدراسي للسنة الحالية.

5- مراحله :

تعتمد إجراءات التقويم التشخيصي على أربعة مراحل رئيسية:

- تحديد مستوى التلاميذ في المهارات والمعارف الأساسية.
- تحديد التلاميذ المتعثرين وتحديد احتياجاتهم.
- تقديم الدعم اللازم للتلاميذ المتعثرين.
- تقييم أثر الدعم على نتائج التلاميذ.

6- وسائله :

يتم تنفيذ التقويم التشخيصي باستخدام مجموعة متنوعة من الوسائل كالاختبارات التحريرية والشفوية، والواجبات المنزلية، والمشاريع.

ولا شك بأن الاختبار التشخيصي يعد أداة هامة ورئيسة في هذا النوع من التقويم، وتهدف الاختبارات التشخيصية إلى تحديد نقاط الضعف والقوة لدى التلاميذ في موضوع دراسي معين أو مادة دراسية معينة.

ويؤدي هذا النوع من الاختبارات خدمة كبيرة لعملية التدريس التي يقوم بها المعلمون، وحتى يتم بناء اختبار تشخيصي جيد، يجب اتباع الخطوات التالية:

- تحديد موضوع دراسي معين يهدف المعلم لتعليمه لتلاميذه.
- تحديد محتوى هذا الموضوع الدراسي.
- التأكد من مدى مناسبة المحتوى لقدرات ومفاهيم التلاميذ.
- تصميم جدول مواصفات لهذا الموضوع الدراسي.
- تحديد الأهداف التي يأمل المعلم تحقيقها من الموضوع الدراسي الذي ينوي تعليمه.

- التأكد من مناسبة المحتوى لتحقيق الأهداف الموضوعية.
- تحديد المهارات اللازمة لقيام التلميذ بالعمل بشكل جيد.
- تحليل هذه المهارات إلى عناصرها الأولية تحليلًا دقيقًا.
- وضع أسئلة على هذه العناصر وعدم إغفال أي عنصر منها.
- ترتيب الأسئلة بحيث تدور حول موضوع معين بشكل متتابع.

الدكتورة س. علوي

رابعاً: التقويم التكويني (البنائي)

- 12- تعريفه .
- 13- أغراضه .
- 14- المبادئ التربوية التي يستند إليها .
- 15- وسائله .
- 16- دور المعلم فيه .
- 17- دور المتعلم فيه .

1- تعريفه :

التقويم التكويني عملية تقويمية منظمة تتم أثناء التدريس، وغرضها تزويد المعلم والمتعلم بتغذية راجعة من أجل تحسين العملية التعليمية التعليمية، ومعرفة مدى تقدم التلميذ. وبناء على ذلك، فإن الهدف الرئيس الذي يسعى إليه هذا النوع من التقويم هو تصحيح مسار العملية التعليمية التعليمية، ومعرفة المستوى الذي وصل إليه المتعلم في تعلمه. وهناك مظاهر عديدة لهذا النوع من التقويم يقوم بها المعلم مثل: مناقشة المتعلمين مع بداية الحصة الدراسية فيما تعلموه في وقت سابق، وتصحيح أخطاء التعلم، وإعادة تعليم التلاميذ الموضوعات التي لم يفهموها، وإعطاء المعلم للتلاميذ تمارين صفية، أو وظائف منزلية، وتصحيح هذه الوظائف وبيان الأخطاء التي قد يكون التلميذ قد وقع فيها.

2. أغراضه :

يمكن تحديد أغراض التقويم التكويني ضمن مجالين رئيسيين هما:

أ- الأغراض المباشرة للتقويم التكويني: وتتمثل في الآتي:

- التعرف على تعلم التلاميذ ومراقبة تقدمهم وتطويرهم خطوة خطوة.
- قيادة تعلم التلميذ وتوجيهه في الاتجاه الصحيح.
- تحديد الخلل في تعلم التلميذ.
- مساعدة المعلم على تحسين أسلوبه في التعليم أو إيجاد طريقة تعلم بديلة.
- إعادة النظر في المنهاج المدرسي وتعديله إذا كان عاملاً من عوامل عدم التعلم أو صعوبته.
- تزويد المعلم والمتعلم بتغذية راجعة عن التعليم والتعلم.
- وضع خطة للتعليم العلاجي وتصحيحه لتخليص التلميذ من نقاط الضعف.

ب. الأغراض غير المباشرة للتقويم التكويني: وتتمثل في الآتي:

- تقوية دافعية التلاميذ نحو التعلم وذلك نتيجة لمعرفته الفورية لنتائجه وأخطائه وكيفية تصحيحها.
- تثبيت التعلم.
- زيادة انتقال أثر التعلم.

3- المبادئ التربوية التي يستند إليها :

- التقويم مدمج في عملية التعليم والتعلم وليس منفصلا عنها.
- كل طالب عنصر فريد في الموقف التعليمي التعليمي له خصوصيته المميزة له، وهو ذو دور نشط وفعال في عملية التعليم والتعلم.
- التقويم له طابع علاجي وتعزيزي للطلاب.
- دور المعلم تيسير التعلم وتلبية حاجات الطلاب ومتطلبات المنهج المدرسي.
- إثارة اهتمام الطلاب وإثارة دافعتهم عوامل أساسية للتعلم.
- يتحقق التعليم والتعلم المتبادل بين المعلم والطلاب نتيجة للتفاعل الإيجابي بينهم.
- تنوع البيئة التعليمية التعليمية، وتنوع الوضعيات التعليمية متطلبات ومثيرات مهمة للتعليم والتعلم.

4- وسائله وأساليبه :

- يمكن للمعلم استخدام عدد من الأساليب والأدوات من أجل التأكد من تعلم التلاميذ للمادة العلمية، ومنها:
- طرح المعلم أسئلة شفوية أثناء الحصة الدراسية للتأكد من أن تلاميذه قد فهموا نقطة قد انتهى لتوه من شرحها.
- إجراء اختبار قصير أثناء الحصة.
- طلب المعلم من التلاميذ القيام بحل عدد من التمارين أثناء الحصة.
- استخدام قوائم التقدير في الدروس العملية.
- تصميم اختبارات لتغطية وحدة من وحدات المنهاج المدرسي الذي يقوم بتدريسه.
- ملاحظة أداء الطالب.
- المناقشة الصفية.
- الواجبات المنزلية ومتابعتها.

5- دور المعلم فيه :

- من الأدوار التي يتطلبها التقويم التكويني من المعلم ما يأتي:
- تقويم عمليات الإعداد في مراحلها كافة.
- تحديد الأخطاء التعليمية المتوقعة من الطلاب.
- تشجيع الطلاب على طرح الأسئلة والاستفسار والتساؤل حول تعليمهم وتعلمهم والمواضيع التي يتعلمونها.
- إشراك الطلاب في التخطيط للتعليم والتعلم وفي اتخاذ القرارات المتعلقة بهذا التعليم والتعلم.
- تفعيل عمل المجموعات التعليمية.
- ملاحظة أداء الطلاب ورصد تقدم الطلاب وكذلك أخطائهم.

- تقويم تعلم الطلاب بعد تنفيذ كل مهمة تعليمية تعليمية.
- استخدام التغذية الراجعة البناءة وتقديمها للطلاب.
- استخدام التغذية الراجعة في تطوير الممارسات التعليمية التعليمية المستقبلية.

6- دور المتعلم فيه :

- يتطلب التقويم التكويني من الطلاب دوراً إيجابياً ونشطاً، وتجعل منهم مشاركين مسؤولين عن تعلمهم. ويتمثل هذا الدور الإيجابي للطلاب في الآتي:
- مناقشة محتوى التعلم مع المعلم وإبداء وجهات نظره فيه.
 - العمل التعاوني لتحقيق التعلم وحل المشكلات.
 - المشاركة في تفعيل عمل المجموعات التعليمية.
 - المشاركة في المناقشات والتساؤل والاستفسار حول تعلمه.
 - مناقشة المعلم في التقويم.
 - المشاركة والتفاعل مع محتوى التعلم، وأهداف عملية التقويم ووسائله ومستوى الأداء المطلوب ومعايير النجاح.
 - مشاركة المعلم في هندسة البيئة الصفية واختيار الوسائط والمثيرات والتفاعلات اللازمة لبنائها، والعمل التعاوني لتحقيق أهداف التعلم وحل المشكلات.

خامسا: التقويم التجميعي (التحصيلي، النهائي، الختامي)

18- تعريفه.

19- أهميته.

20- أغراضه.

21- وسائله.

1- تعريفه:

- يقصد بالتقويم التجميعي (الختامي) التقويم الذي يستند إلى نتائج الاختبارات التي يعطيها المعلم في نهاية الشهر أو في منتصف الفصل الدراسي أو نهايته، ثم رصد نتائجها في سجل العلامات من أجل تقويم تحصيل التلميذ بموجبها، تمهيدا لاتخاذ قرار بترفيعه أو ترسيبه أو تخرجه أو إعطائه شهادة تبين مقدار إنجازاته.
- يشير "بارلو" (Barlow) (1996) إلى أن التقويم التحصيلي هو الحكم على مجموعة أعمال نهائية.
- في حين يعرف "عيسى العباسي" (2006) التقويم التحصيلي بأنه إصدار الحكم النهائي على قيمة العملية التعليمية.
- ويؤكد المعهد الوطني للبحث في التربية "إنه تقويم يسمح للمعلم بوضع محصلة، وجرّد للكفاءات المكتسبة أو ترتيب للتعلمات، حيث تتخذ القرارات المتعلقة بتوجيه التلاميذ وهذا حسب مكتسباتهم، هذا التقويم ينتمي أكثر للمراقبة منها للتعديل من أجل منح علامة أو شهادة، "إنها تخدم الوجه المادي أو الإداري لفعل التقويم".
- أما "عبد السلام عبد الله الجقندي" (2008)، فيذكر بأن التقويم التحصيلي يقصد به قياس نواتج التعلم وتقديرها في ضوء الأهداف المنشودة، ويهدف إلى تحديد مستويات التلاميذ ومدى إنجازهم لمتطلبات النجاح تمهيدا لنقلهم إلى صف أعلى، أو منح شهادة أو اتخاذ قرار بنجاح التلميذ أو رسوبه".
- ويوضح "محمد عيسى الطيطي" (2008) أن هذا التقويم يشير إلى النوع الذي يستخدم في نهاية فترة زمنية أو نهاية مسار أو برنامج دراسي لأغراض وضع الدرجات، أو منح الشهادات أو مدى التقدم الذي تم تحقيقه في التعلم أو البحث عن فعالية منهج ما، أو مادة دراسية أو خطة تربوية معينة".
- التقويم التحصيلي يقرر المكتسبات التي توصل إليها المتعلم، فهو يجري إثر الانتهاء من العملية التعليمية التعليمية سواء في طور أو فصل أو سنة، ويجسد بنتائج معينة، فهو يقوم بوظيفة الإشهاد للعملية التعليمية - التعلمية، بمعنى حدوث تعلم، وبالتالي الانتقال إلى مستوى دراسي أعلى، أو عدم حدوثه، وبالتالي إعادة المستوى الدراسي أي رسوبه.

2- أهميته:

يعد هذا النوع من التقويم ذا أهمية كبيرة في الحكم على تقدير أداء المتعلم، فمن خلاله يتم تقويم المحصلة النهائية للتعلم تمهيدا لإعطاء تقديرات نهائية لأداء للمتعلمين لنقلهم لصفوف أعلى، ومراجعة فعالية طرق التدريس بشكل عام، وتقديم حصيلة عن تنمية الكفاءات النهائية التي حددها منهاج السنة أو المرحلة.

إضافة إلى أنه يهتم أيضا بالمسارات والاستراتيجيات المستخدمة لتحقيق الأهداف المرجوة، كما يلقي نظرة عاكسة على ما حققه التلميذ خلال الوقت المسخر للتعلم، ونظرة أخرى استكشافية على تقدم هذه التعلمات نفسها، ويعمل على مراقبة إدماج مكتسبات المعارف النهائية، وهذا بملاحظة سلوكيات نهائية في وضعيات الإنجاز.

كما أنه يساعد في التوجيه المهني والتربوي للتلاميذ وتوزيعهم على أنواع التعليم المناسبة لهم، أو استعماله كمعيار للتعين في بعض الوظائف أو الترقية، ومن أمثلة هذا النوع من التقويم الامتحانات التي تجرى في نهاية كل فصل دراسي أو سنة دراسية أو امتحانات نهاية المراحل للدراسة.

وهو يزود المدرس بمعلومات تساعد على اتخاذ قرارات ملائمة لتصحيح المسار التعليمي التعليمي ودعم التعلم، ويسهم في معرفة مدى تحقيق برنامج تعليمي معين لأهدافه المحددة، وذلك بعد الانتهاء من تنفيذه، فالتقويم الختامي إذن يركز على التقويم الإجمالي لجودة وتأثير البرنامج، ومدى تحقيقه للأهداف المرسومة له.

3- أغراضه :

- يمكن إجمال أغراض التقويم التجميعي في الآتي:
- معرفة مدى تحقيق الأهداف.
 - تسجيل علامة للتلميذ ليتم تقييمه بموجها.
 - إعطاء شهادات للتلاميذ.
 - اتخاذ قرارات إدارية كالترسيب والترفيه والفصل أو الطرد.
 - إعلام الآباء عن نتائج أبنائهم.
 - مقارنة نتائج شعب الصف الواحد.
 - ضبط العملية التعليمية وذلك بمقارنة نتائج الصفوف اللاحقة بنتائج الصفوف السابقة إن تماثلت في القدرة الدراسية.
 - اكتشاف الخلل في المنهاج المدرسي.
 - التنبؤ بأداء التلاميذ مستقبلا.
 - الحكم على العملية التعليمية بشكل نهائي صالحة أو غير صالحة.
 - الحكم على تحصيل المتعلم بشكل نهائي.
 - وضع درجات نهائية بتحديد معايير انتقالهم إلى صفوف أعلى.
 - تصنيف التلاميذ وتوزيعهم على صفوف وشعب مناسبة لمستويات تحصيلهم، سواء كان علميا أو أدبيا.
 - تقويم فاعلية التدريس.
 - التحقق من مدى تحكم المتعلم من الكفاءة المستهدفة أو الأهداف المسطرة والمكتسبات المنتظرة، ويمكن للإدارة إصدار أحكام حول نتائج دراسة المتعلم.
 - الحكم على درجة تحقق التعلم، وقياس الفوارق بين النتائج الفعلية والأهداف والكفاءات المسطرة.

3- وسائله :

تتميز اختبارات التقويم الختامي بأنها أطول وأعم وأشمل وأعم مقارنة بالاختبارات البنائية، كما أنها تدور حول العموميات وتتألف من أسئلة وهي عبارة عن عينة ممثلة للأهداف وأجزاء المادة، وهذا خلافا لأسئلة الاختبارات البنائية التي تضم النقاط الهامة جميعها لا عينة ممثلة عنها.

سادسا: نماذج التقويم

22- تعريف التقويم.

23- تعريف نموذج التقويم.

24- تصنيف نماذج التقويم.

25- أهم نماذج التقويم.

1. تعريف التقويم:

تعتبر عملية التقويم مقوما أساسيا من مقومات العملية التعليمية، وعنصراً أساسيا من عناصر المنهج، فهي الوسيلة التي يمكن بواسطتها التعرف على مدى نجاحنا في تحقيق الأهداف التعليمية التي نسعى إلى تحقيقها، وعلى كيفية الاستفادة من جوانب القوة في العملية التعليمية وتأكيدها، وكيفية معالجة جوانب القصور وتلافيا.

والتقويم هو عملية جمع بيانات كمية أو كيفية عن العملية التعليمية أو أحد جوانبها أو عن ظاهرة ما، ثم تصنيف هذه البيانات وتحليلها وتفسيرها بقصد استخدامها في إصدار حكم أو قرار يؤدي إلى تعديل الموقف نحو تحقيق الأهداف المرجوة، فالتقويم عملية تشخيصية علاجية.

والتقويم يشمل كل مكونات العملية التعليمية وعناصرها، ويساعد في الحكم عليها وتطويرها، وقد تشمل عملية التقويم التقويم نفسه كأحد المكونات المهمة للعملية التعليمية أو كأحد عناصر المنهج.

2- تعريف نموذج التقويم:

نموذج التقويم هو خطة عمل يبسطها معد النموذج ويرى أن اتباعها كفيل بإنتاج تقويم فعال، فهو إطار تصوري للتقويم يصف رؤية معدة لمفاهيم التقويم الأساسية، والكيفية التي يمكن أن توظف بها هذه المفاهيم للوصول إلى أحكام وتوصيات يمكن تبريرها.

ويستخدم نموذج التقويم بصورتين أساسيتين:

- **نموذج إرشادي:** وهو الأكثر شيوعا وهو عبارة عن مجموعة من القواعد والأوصاف والمحاذير والأطر المرشدة التي تحدد ماهية التقويم الجيد وكيفية إجرائه.
- **نموذج وصفي:** وهو مجموعة من الجمل والتعميمات التي تصف أنشطة التقويم أو تشرحها أو تنبأ بها.

3- تصنيف نماذج التقويم التربوي:

توجد العديد من محاولات تصنيف نماذج التقويم التربوي، ومن هذه المحاولات تصنيف "بوفام" وتصنيف "وورثن وساندرز":

أ- **تصنيف "بوفام":** تصنف نماذج التقويم إلى ثلاثة أنواع رئيسة وهي:

- **النماذج القائمة على تحقيق الأهداف:** ويندرج تحت هذا النوع من النماذج نماذج كثيرة من أهمها نموذج "تايلور"، وتنظر هذه النماذج إلى التقويم على أنه عملية التعرف على مدى تحقق الأهداف.
- **النماذج القائمة على إصدار الأحكام:** وتنقسم إلى نماذج تركز على المدخلات، ونماذج تركز على المخرجات.

- **النماذج المساعدة على اتخاذ القرارات:** وهي النماذج التي يتم التقييم فيها لاتخاذ القرار.

ب- تصنيف "وورثن" و"ساندرز": يعتبر هذا التصنيف أكثر التصنيفات شيوعاً، ويقسم النماذج إلى:

- **نماذج التقييم المعتمدة على الأهداف:** هي النماذج التي تنظر إلى التقييم على أنه عملية التعرف على مدى تحقق الأهداف، ومن أهمها نموذج "تايلور".
- **نماذج التقييم المتحررة من الأهداف:** وهي النماذج التي يتحرر فيها المقوم من القيود التي تفرضها الأهداف، ويرصد المقوم كل أثر للبرنامج سواء مقصود (مرتبط بالأهداف)، أو أثر جانبي (غير مرتبط بالأهداف)، ومن هذه النماذج نموذج "سكريفن".
- **نماذج التقييم الخادمة للإدارة:** وهي النماذج التي تهدف إلى تقديم معلومات عن البرنامج إلى صاحب القرار لاتخاذ قرار في البرنامج، ومن هذه النماذج نموذج "ستفلبيم" و"ألكن".
- **نماذج التقييم الطبيعي والتقييم الموجه نحو المشاركين:** وهي النماذج التي تصف البرنامج كما هو على الواقع دون أي محاولة لضبط المتغيرات، ويتم التقييم بصورة طبيعية ويبني على قضايا ومشاكل الأطراف المعنية بالتقييم، ومن هذه النماذج نموذج "ستيك".
- **نماذج التقييم الموجهة نحو المستهلك.**
- **نماذج التقييم بناء على أحكام أهل الاختصاص (المعتمد على الخبراء):** وهي النماذج التي يقوم فيها فريق من الخبراء بإصدار حكم حول جدارة أو جدوى موضوع التقييم.

4- أهم نماذج التقييم:

وتتنوع نماذج التقييم تبعاً للفلسفات والمفاهيم الخاصة بمطورها والغرض من التقييم، وقد ظهرت هذه النماذج استجابة لدعوات إصلاح التقييم وللمساعدة في تطوير الدراسات التقييمية، ولإيجاد لغة مشتركة تساعد في تنمية التواصل العلمي وتبادل الخبرات.

ونظراً لوجود العديد من نماذج التقييم فسيتم تناول أهمها:

أ- نموذج "رالف تايلور" (Ralph Tyler):

التقييم في هذا النموذج عبارة عن تقرير مدى التوافق بين الأهداف والنواتج؛ أي قياس مدى تحقق أهداف البرنامج التعليمي، ويتميز هذا النموذج برواجه الكبير وسرعة انتشاره وكثرة تطبيقاته، وذلك لبساطة إجراءات التقييم وفق هذا المدخل وسهولة استيعابها وتطبيقها من قبل المقومين، فمهمة المقوم تنحصر في تحديد أهداف البرنامج و صياغتها سلوكياً، ومقارنتها بالنواتج (الأداء).

ويمر التقييم وفقاً لهذا النموذج بسلسلة من الخطوات هي:

- تحديد الأهداف العامة (الغايات) وتصنيفها.
- ترجمة كل هدف عام إلى مجموعة من الأهداف الإجرائية.
- إعداد أو اختيار أدوات جمع البيانات.
- جمع البيانات عن أداء المتعلمين.
- مقارنة أداء المتعلمين مع الأهداف السلوكية.

وبناء على المقارنة بين الهدف والأداء تتضح أماكن النجاح والإخفاق في البرنامج، فالأهداف المحققة تعتبر نقاط قوة في البرنامج، أما الأهداف غير المحققة، فتعتبر نقاط ضعف فيه. فالأهداف هي الأساس التي يترتب عليها نجاح البرنامج أو إخفاقه، وصياغتها بصورة جيدة مهمة جدا في عملية التقويم.

ولكن يعاب على هذا النموذج تأخر نتائج التقويم إلى نهاية البرنامج مما يحد من الاستفادة منها في تحسين البرنامج، كذلك صعوبة صياغة الأهداف إجرائيا في بعض الأحيان، كما يؤخذ عليه أنه خطي في حين المنهج منظومة متكاملة.

ب- نموذج "هاموند" (Hammond):

يتعدى هذا النموذج نموذج "تايلور"، فلا يقف عند معرفة مدى تحقق الأهداف، بل يتعدى ذلك إلى محاولة فهم الأسباب التي تكون وراء نجاح أو إخفاق البرنامج في تحقيق أهدافه.

وتتلخص خطوات نموذج هاموند وفق الآتي:

- تحديد الجوانب المراد تقويمها في البرنامج.
- تحديد الأهداف العامة للبرنامج والعمل على تصنيفها تبعا لمجالاتها المختلفة.
- تحويل الأهداف العامة إلى أهداف إجرائية أو سلوكية محددة، تبعة لمستوياتها المتنوعة.
- تحديد المواقف التي يمكن أن نلاحظ فيها هذه الأهداف الإجرائية.
- تصميم أو انتقاء الأدوات الملائمة لقياس هذه الأهداف.
- جمع البيانات الخاصة بالأهداف باستخدام الأدوات المشار إليها في الخطوة السابقة.
- تحليل مستوى الأداء وتفسيره، وذلك من خلال البيانات التي تم جمعها حول الأداء الفعلي من جهة، وحول الأهداف التي تم تحديدها من جهة ثانية. وفي حال تحقيق تلك الأهداف، فإنه يتوجب تقديم بيانات موضوعية مناسبة، أما إذا لم يتم تحقيقها، فيجب توضيح أوجه القصور والضعف في البرنامج ذاته.

ومن مزايا هذا النموذج:

- يتصف بأنه نموذج دائري، وذلك حتى يتم ضمان وجود تغذية راجعة مستمرة تساعد في الغالب على إعادة تحديد الأهداف والعمل على صياغتها، ثم القيام بتعديل خطة التقويم في نهاية المطاف.
- تحديد الأهداف الإجرائية للمنهج، مما يسهل من مهمة التحقق منها، ثم القيام بعملية التقويم في ضوءها.
- استخدام التغذية الراجعة المستمرة لتحسين البرنامج.
- توضيح العلاقة بين مختلف المتغيرات الخاصة بالبرامج فيما أطلقت عليه "هاموند" بمكعب البرنامج.
- سهولة الفهم والمتابعة والتحقق، وبالتالي يوفر النموذج معلومات تهم المتخصصين التربويين.
- الدور الكبير للنموذج في حفز التربويين للمزيد من البحث والإجراءات التي ساهمت في إثراء الأدب التربوي بكثير من الأفكار الإبداعية.

وتقترح "هاموند" مكعبا يستفيد منه المقوم في وصف البرامج التربوية وتنظيم متغيرات التقويم، وهو ما يساعد في تحديد العوامل التي تكون وراء نجاح البرنامج أو إخفاقه وأطلقت على هذا المكعب اسم (بنية التقويم)، وكل بعد من أبعاده الثلاثة يتعلق بمتغير من متغيرات التقويم، والمتغيرات هي:

المتغير الأول: يتعلق بخصائص البرنامج التربوي، ويضم خمسة عناصر: التنظيم، المحتوى المراد تغطيته، النشاط التدريسي، التجهيزات، التكلفة.

المتغير الثاني: يتعلق بخصائص الأفراد والجماعات المرتبطين بالبرنامج، ويضم ستة عناصر: التلميذ، المعلم، الإداري، المختص التربوي، الأسرة، المجتمع المحلي.

المتغير الثالث: يتعلق بأهداف البرنامج التربوي؛ والتي تكون إما معرفية أو انفعالية أو حركية.

ج- نموذج ستيك (Stake):

تناول "ستيك" التقييم من جهة تختلف جذريا عن الجهة التي تناولها "تايلور"، فقد انتقد نموذج "تايلور" لاعتماده على الأهداف فقط في عملية التقييم؛ إذ يرى أن الأهداف قد تتغير أثناء التنفيذ مما يسبب ضرورة اتصال المقوم بالمستفيد من البرنامج باستمرار لاكتشاف الأمور ذات الأهمية الخاصة بالبرنامج، وأخذها بالحسبان عند التقييم. كما انتقد "ستيك" التقييم لدى "تايلور" لحصره عملية التقييم في جوانب ضيقة ومحددة في البرنامج (تحقق الأهداف فقط) وطريقة جمع المعلومات التقييمية؛ والتي تقتصر على الاختبارات والأدوات المقننة والتجارب. فيرى "ستيك" أن الغرض من التقييم مساعدة المستفيد على تكوين رؤية شاملة عن موضوع التقييم، ويؤكد على ضرورة تجنب إصدار المقوم لأحكام نهائية حول موضوع التقييم؛ بل يجمع آراء الآخرين وأحكامهم، بالإضافة إلى رؤيته، ويقدمها إلى المستفيد في قالب مناسب. وقد ساهم "ستيك" بتطوير مفاهيم التقييم التربوي عندما قدم نموذجين للتقييم هما: نموذج الهيئة الكلية للتقييم، ونموذج التقييم المتجاوب، وفيما يلي تناول لهنذين النموذجين:

• نموذج الهيئة الكلية (الصورة الكاملة) للتقييم:

طرح "ستيك" رؤيته عن التقييم في مقالة نشرها عام 1967م انتقد فيها الأساليب المستخدمة في التقييم والتي لا تسهم في إعطاء صورة كاملة عن موضوع التقييم، فتم طرح هذا النموذج والذي يرى فيه أن أي عملية التقييم تحتوي على عمليتين أساسيتين هما الوصف والحكم (التحكيم)، وهذه العمليات تجري لثلاث مراحل من حياة البرنامج التعليمي: السوابق، العمليات، النواتج. وتوضح مراحل حياة البرنامج التعليمي بالتالي:

أما عملية الوصف، فيتم فيها جمع معلومات عن المراحل الثلاث: السوابق، العمليات، النواتج ليكون الوصف شاملا، ويحدد لكل مرحلة من المراحل الثلاث نوعين من البيانات هما مأمول (مقصود) وواقع (ملحوظ)، فالمأمول هو ما خطط له في كل مرحلة، أما الواقع فما تحقق فعلا.

ثم بعد الحصول على المعلومات الوصفية يتم التعامل معها بإحدى الطريقتين:

- إيجاد مدى التوافق بين الواقع والمأمول لكل من السوابق والعمليات والنواتج.
- إيجاد مدى اعتماد النواتج على العمليات والسوابق، واعتماد العمليات على السوابق.

بينما عملية التحكيم أو إصدار الأحكام على مراحل حياة البرنامج فتتم لجميع مراحل حياة البرنامج وفق معايير نسبية أو مطلقة فالمعايير النسبية تكون مستمدة من برامج أخرى مشابهة للبرنامج، أما المعايير المطلقة فهي التي حددت بناء على ما يجب أن يتوفر في البرنامج، وفي الحالتين لابد من الاتفاق على المعايير المستخدمة في إصدار الأحكام بين المقوم والمعنيين بالبرنامج حتى لا يكون هناك نظرة سلبية للمقوم أو محاولة حجب معلومات عنه.

كما يمكن للمقوم الحصول على تقدير عام لجدارة البرنامج من خلال دمج الأحكام المبنية على معايير مطلقة مع الأحكام المبنية على معايير نسبية.

ولا يكتمل التقييم إلا من خلال إعطاء مبرر لتقييم البرنامج، وإن كان المبرر غير صريح فيوضح الخلفية الفلسفية والأغراض الأساسية والحاجة للبرنامج.

وبالتالي فالخطوات التي يقوم بها المقوم في هذا النموذج هي:

- إعطاء مبرر للبرنامج.
- جمع المعلومات الوصفية، من خلال رصد الواقع والمأمول من السوابق والعمليات والنواتج.
- تحليل المعلومات الوصفية.
- تحديد المعايير المطلقة والنسبية.
- تقدير مدى قابلية البرنامج لتلك المعايير.
- إصدار الحكم على البرنامج من قبل المقوم بمفرده، أو بالتعاون مع المسؤولين في البرنامج وغيرهم من المتأثرين به.

• نموذج التقييم المتجاوب (الإيجابي، الحساس):

مع بداية السبعينيات بدأ "ستيك" يدعو إلى تقويم يعالج السلبيات المتأصلة في نماذج التقييم التقليدية؛ والمتمثلة في عدم استجابة تلك النماذج للقضايا والمشكلات التي تشغل بال المعنيين بالتقويم (المستفيدين من التقييم)، والتي تعتبر من وجهة نظره المحك النهائي لصديق التقييم.

ومن هذا المنطلق طرح "ستيك" نموذج التقييم المتجاوب، والذي يعتمد على الاستفسارات التي يطرحها المعنيون بالتقويم ومن القضايا والمشكلات التي تشغلهم، كما يعتمد في منهجيته لجمع المعلومات على نموذج الهيئة الكلية للتقويم. فالمقوم يحرص على فتح حوار للمعنيين بالتقويم ليحدد من خلاله الهدف من التقييم ونوعية المعلومات المطلوب جمعها، والتي تسهم في فك الغموض المحيط بالبرنامج، والإجابة على الاستفسارات التي يطرحها المعنيون بالتقويم وحل المشكلات التي تواجههم، ويستمر هذا التواصل بين المقوم والمعنيين بالتقويم طوال فترة التقييم لكي يستوعب التقييم أي اهتمامات ومطالب طارئة لم تكن قائمة عند البدء بالتقويم.

فالتقويم وفق هذا النموذج مرن ومتطور يسمح باستيعاب أي مستجدات، فالتخطيط للتقويم في بداية الأمر يكون بشكل عام، ثم يبدأ التخصيص وفقا لسير العمل، فالتقويم لا يكون محددا بكل دقة وتفصيل منذ البداية، بل يتطور أثناء إجراء التقييم. وفي سبيل الحصول على المعلومات فإن المقوم يحرص على جمع معلوماته بظروف طبيعية بعيدة عن التصنع من خلال الملاحظة المباشرة، والتحدث إلى الناس والاستماع إليهم، والاعتماد على عدة مصادر في جمع المعلومات، ويحرص المقوم على ألا يواجه المعنيون بالتقويم اهتمامهم بعيدا عن أوجه ضعفهم التي لا يريدون الإفصاح عنها عند تعبيرهم عن اهتماماتهم. وعند إصدار الحكم على جودة البرنامج لا يفرض المقوم قيمة معينة؛ بل تكون الأحكام النهائية مستمدة من القيم التي يحملها المعنيون بالتقويم.

د- نموذج "سكريفن" (Scriven):

هاجم "سكريفن" أساليب التقييم التي كانت تركز على الأهداف فقط، حيث يرى أن الأهداف التي يقاس مدى تحققها قد تكون غامضة أو لا تكون ذات أهمية وغير مجدية أو تكون أهدافا سطحية، كما أن التركيز على تحقق الأهداف يعني التركيز على النتائج المقصودة، وإهمال النتائج غير المقصودة، والتي تسمى بالآثار الجانبية للبرنامج، كذلك التركيز يحصر عمل المقوم في زاوية ضيقة هي الأهداف، ويفرض عليه استخدام أساليب لجمع المعلومات تعتبر محدودة، وبالتالي فـ"سكريفن" ينادي بأهمية الآثار الجانبية للبرنامج، وأنها هي المحك الأصلي لتقرير جودة البرنامج، بغض النظر عن الأهداف وتحقيقها.

ويتم الحكم على جودة البرنامج من خلال قياس الآثار المترتبة على البرنامج الواقعية بشقيها المقصود وهو الناتج عن تحقق الأهداف وغير المقصود، مع الاهتمام بما يجري داخل البرنامج من تفاعلات وأنشطة.

وتتم عملية التقييم من خلال قيام المقوم بإعداد بيان بحاجات الجماعة محل الملاحظة في البرنامج، ثم يقيس نواتج البرنامج وآثاره الجانبية بطرق ذات طبيعة نوعية. وإذا ثبت أن للبرنامج آثار أو نتائج متماشية مع إحدى الحاجات التي حددت سابقا، فإن البرنامج يعتبر مفيدا، ويجب أن يكون المقوم ملاحظا حياديا موضوعيا.

مثلاً: لدينا برنامج دراسي في تعليم الرياضيات قد تحققت أهدافه، لكن وجد أن له آثار جانبية سلبية كتقليل اهتمام الطلاب بالرياضيات، فلا نستطيع الحكم على البرنامج بأنه مفيد تماماً، بينما في النماذج التي تهتم بقياس تحقق الأهداف يعتبر البرنامج السابق مفيداً جداً.

ويؤكد "سكريفن" على عدم الاقتصار على التقويم المعتمد على الأهداف فقط أو التقويم المتحرر من الأهداف؛ بل كل واحد منهما مكمل للآخر حتى يتم تناول الموضوع المراد تقويمه بصورة كاملة.

ويعتبر تقويم "سكريفن" موجه إلى خدمة المستهلك أو المنتفع من البرنامج التعليمي؛ إذ يرى أن تقويم البرامج المتحرر من الأهداف مشابه لتقويم سلعة أو منتج ما من خلال أخذ رأي المستخدمين والمستفيدين منه وليس المنتجين للمنتج.

هـ - نموذج "ستفليبم" (Stufflebeam):

ويرمز لهذا النموذج نموذج (CIPP) اختصار لـ (Context , Input , Process , Product) ؛ أي نموذج السياق والمدخلات والعملية والمخرجات لـ (ستفليبم). وكان مبرر قيام هذا النموذج أن البرامج المراد تقييمها لم يكن لها سوى أهداف عامة لا ينفع معها استخدام أسلوب "تايلور" السائد في تلك الفترة، مما أدى إلى قيام المسؤولين عن البرامج بالاجتهاد بتفسير تلك الأهداف مما ترتب عليه تباين في تطبيقات البرنامج الواحد من موقع لآخر، عندها اقتنع "ستفليبم" أنه هناك حاجة إلى تعريف أشمل للتقويم أبعد من مجرد قياس تحقق الأهداف، وطرح التعريف التالي للتقويم: (التقويم هو العملية التي تتكون من التخطيط للحصول على المعلومات المفيدة لصانع القرار، ومن ثم جمع هذه المعلومات وتقديمها بهدف الحكم على بدائل القرار).

كما يرى "ستفليبم" أن هناك أربعة أنواع من التقويم يخدم كل منها نوعاً من أنواع القرارات وهي :

النوع الأول: تقويم السياق (الإطار) العام:

ويخدم هذا النوع من التقويم قرارات التخطيط، وهو النوع الأساسي من أنواع التقويم الأربعة حسب النموذج وأكثرها شيوعاً، وهدف التقويم هو تحديد سبب أو مبرر لأهداف البرنامج القائم، والتي لا تقابل احتياجات صانع القرار عند دراسة البيئة والظروف والمشكلات التعليمية والإمكانات المتوفرة في النظام التعليمي.

النوع الثاني: تقويم المدخلات:

يخدم هذا التقويم قرارات التصميم، فالمعلومات التي يوفرها المقوم تساعد صانع القرار في اختيار أنسب المدخلات التي تخدم تحقيق الأهداف، فالمقوم يجمع المعلومات عن البدائل المختلفة المتوفرة، والتي تحتاج إلى توفير من حيث متطلبات تنفيذها من قوة بشرية ومصادر وتجهيزات وميزانية وجدول زمني، ثم ينقدها ويفحصها ويرتبها ثم يقدمها لصانع القرار ليختار البديل المناسب.

النوع الثالث: تقويم العملية:

يخدم هذا التقويم قرارات التنفيذ؛ وهي القرارات التي يتخذها صانع القرار في مرحلة تطبيق البرنامج على أرض الواقع، وهذا النوع من التقويم يمد صانع القرار بتغذية راجعة دورية عن سلامة التطبيق، ويهدف هذا التقويم إلى إبراز مواقع الخلل في تصميم الخطة أو تنفيذها، ومتابعة مكونات الخطة الأخرى، والتعرف على المشكلات الموجودة في العمل وعلاجها وتهيئة الظروف المناسبة لسير العمل على ما خطط له.

النوع الرابع: تقويم الناتج:

يخدم هذا التقويم قرارات التدوير وهي القرارات التي يتخذها صانع القرار لتحديد مستقبل البرنامج، إما بالاستمرار أو بالإيقاف أو بالتعديل، والهدف الأساسي لهذا التقويم هو التحقق من درجة مقابلة أهداف البرنامج لاحتياجات المجموعة المخدومة بالبرنامج.

وفي هذا التقييم يجمع المقوم معلومات عن درجة التوافق بين الأهداف والأداء، ويكشف عن الآثار غير المقصودة للبرنامج المرغوبة وغير المرغوبة، ثم تقدم إلى صانع القرار ليتخذ القرار المناسب.

و- نموذج "ألكن" (Alkin):

طور "ألكن" هذا النموذج أثناء إشرافه على مركز دراسات التقييم التابع لجامعة كاليفورنيا فرع (لوس أنجلوس)، ويسمى هذا النموذج (نموذج مركز دراسات التقييم) (The CSE Model) وهي اختصار لـ (Center for the Study of Evaluation). والتقييم وفق هذا النموذج عبارة عن تحديد أنواع القرارات التي سيتم اتخاذها إزاء كل مرحلة من مراحل تطوير البرنامج، ثم يتم جمع المعلومات وتحلل وتقدم لصانع القرار ليصدر فيها قرار.

وتوجد خمسة أنواع من التقييم يخدم كل منها نوعاً من أنواع القرارات وهي كالتالي:

1. تقدير الاحتياجات: وفي هذا النوع من التقييم يجمع المقوم معلومات عن الوضع الراهن للنظام، ويحدد الاحتياجات من خلال التفاوت الحاصل بين الأهداف والأداء، ثم يتم إيجاد قائمة بالأهداف المرغوبة التي تلي تلك الاحتياجات، ويشبه هذا التقييم إلى حد كبير تقييم السياق العام في نموذج (CIPP).
2. التخطيط للبرنامج: وفي هذا التقييم يجمع المقوم المعلومات التي تساعد صانع القرار على اختيار البدائل المناسبة التي تفي بمقابلة الاحتياجات التي تم تحديدها في مرحلة تقدير الاحتياجات، ويشبه هذا التقييم إلى حد كبير تقييم المدخلات في نموذج (CIPP).
3. تنفيذ البرنامج: وفي هذا التقييم يجمع المقوم معلومات عن مدى تنفيذ البرنامج وفقاً للخطة المقترحة.
4. تحسين البرنامج: وفي هذا التقييم يجمع المقوم معلومات عن النجاح النسبي لمختلف مكونات البرنامج وعن مدى تحقق الأهداف المقصودة، ويشبه هذا التقييم إلى حد كبير تقييم العملية في نموذج (CIPP).
5. إجازة البرنامج (تقويم النواتج): وفي هذا التقييم يجمع المقوم معلومات تساعد صانع القرار على اتخاذ قرار بشأن القيمة النهائية للبرنامج ككل، وهو القرار الذي يحدد مصير البرنامج من حيث إما الإبقاء على البرنامج بوضعه الحالي أو تنقيحه أو إيقافه.

سابعاً: نظم وضع الدرجات وكتابة التقارير

26- نظام التقدير باستخدام الدرجة الخام أو النسبة المئوية.

27- نظام التقدير بالحروف أو الأرقام.

28- نظام التقدير ناجح - راسب.

29- أساليب مكملة لنظم التقدير.

30- إرشادات لتطوير أنظمة وضع الدرجات وكتابة التقارير.

توجد نظم مختلفة لوضع الدرجات وكتابة التقارير، فهناك نظم تقليدية، وأخرى نظم بديلة طرحت لتحل محل النظم التقليدية، وسيتم عرض لأهم هذه النظم:

1- نظام التقدير باستخدام الدرجة الخام أو النسبة المئوية:

إن نظام التقدير باستخدام الدرجة الخام أو النسبة المئوية لا يزال يستخدم على نطاق واسع في المدارس، وفيه يتم وضع الدرجة التي تمثل تحصيل الطالب في كل مادة دراسية على حدة، بالإضافة إلى العلامة الإجمالية لجميع المواد الدراسية، والتي يعبر عنها برقم واحد. ويتخذ هذا النظام صورتين على النحو التالي:

- الدرجة الخام التي حصل عليها الطالب مقرونة بالنهايتين العظمى والصغرى.
- النسبة المئوية، وهي تعطى بناء على مقياس حده الأدنى صفراً وحده الأعلى مئة.

وحق يمكن إعطاء معنى للدرجات الخام أو النسب المئوية، فيجب تفسيرها من خلال إطار مرجعي، والإطار المرجعي قد يربط الدرجة بترتيبها ضمن توزيع علامات المجموعة الصفية، أو يربط العلامة بنسبة معينة من الأهداف أو الإجابات الصحيحة، والتفسير الأول هو تفسير يستند إلى إطار معياري المرجع، في حين أن التفسير الثاني يستند إلى إطار محكي المرجع.

2- نظام التقدير بالحروف أو الأرقام:

إن نظام التقدير بالحروف أو الأرقام يعبر عن مستوى تحصيل الطالب بدرجة واحدة يعبر عنها بحرف (أ، ب، ج، د، هـ)، أو برقم (1، 2، 3، 4)، ويوضح الجدول الآتي استخدام الحروف الأبجدية للتعبير عن مستوى تحصيل الطالب بحيث يشير الحرف (أ) إلى أعلى تحصيل، في حين يشير الحرف (هـ) إلى أدنى تحصيل.

جدول يوضح الوصف النوعي للتقديرات الخماسية المقابلة للحروف

التقدير بالحروف	الوصف النوعي للتقدير
أ	ممتاز
ب	جيد جداً
ج	جيد
د	مقبول
هـ	ضعيف

ويمكن أن تظهر هذه الحروف وما يقابلها من تقديرات بصورة موسعة، والجدول الآتي يعرض توسيعا للتقديرات السابقة، ويبين وصفا نوعيا للتقديرات المقابلة للحروف.

جدول يوضح الوصف النوعي للتقديرات الموسعة المقابلة للحروف

التقدير بالحروف	الوصف النوعي للتقدير
أ	ممتاز
ب ⁺	جيد جدا
ب	جيد
ج ⁺	فوق المتوسط
ج	متوسط
د ⁺	مقبول
د	ضعيف
هـ	ضعيف جدا

- وعلى الرغم من أن هذا النظام يعطى تلخيصا جيدا لمستوى تحصيل الطالب وييسر فهم الدرجة الممثلة بالحروف بسهولة، سواء من قبل الطلاب أنفسهم أو من قبل أولياء أمورهم، إلا أن به عدة عيوب نجملها فيما يلي:
- 1- لا يسمح بإجراء تحليلات إحصائية إلا بعد تحويل العلامات الممثلة بالحروف إلى أرقام.
 - 2- لا يعطي صورة دقيقة عن مستوى تحصيل الطالب، فقد تتضمن العلامة تلخيصا لعوامل أخرى بجانب التحصيل كالجهد، والسلوك الشخصي والاتجاه.
 - 3- الدرجات في صورة حروف أدت إلى اهتمام زائد عن الحد بالدرجات من قبل الطلاب وأولياء الأمور، باعتبارها أهدافا لا بد من تحقيقها.
 - 4- يختلف معنى الدرجات بالحروف من معلم لآخر، ومن مقرر لآخر، ومن مجموعة صفية لأخرى، ومن مؤسسة تعليمية لأخرى، وهذا كله يزيد من صعوبة تفسيرها.

تستخدم بعض المؤسسات التعليمية نظام التقدير بالأرقام على أساس أنه يوفر للمعلم إمكانية إجراء التحليلات الإحصائية على العلامات بصورة مباشرة، ودون اللجوء للتحويل. ويبين الجدول الآتي وصفا للتقديرات بالأرقام المقابلة للحروف.

جدول يبين وصفا للتقديرات بالأرقام المقابلة للحروف

التقدير بالحروف	الوصف النوعي للتقدير
أ	4
ب ⁺	3,5
ب	3
ج ⁺	2,5
ج	2
د ⁺	1,5
د	1
هـ	صفر

3- نظام التقدير ناجح - راسب:

رغم أن استخدام النظامين السابقين شائعا في أغلبية المدارس والجامعات، سواء في البلاد العربية أو الأجنبية، إلا أنه في السنوات الأخيرة بذلت محاولات عديدة من قبل خبراء وعلماء القياس التربوي في الولايات المتحدة الأمريكية لتحسين تلك الأنظمة، وخاصة نظام التقدير الخماسي بالحروف، وذلك من خلال تغيير عدد الحروف أو الرموز المستخدمة مقابل التقديرات الخماسية، فيتم على سبيل المثال إنقاص عدد الحروف إلى ثلاثة يقابلها ثلاثة تقديرات مثل (ممتاز، مرض، غير مرض) على الترتيب، وفي محاولات أخرى تم إنقاصها إلى رمزين يقابلهما تقديرين هما: (مرض - غير مرض) أو (ناجح - راسب).

وفي هذا النظام يستطيع الطالب أن يدرس بعض المقررات، وخاصة الاختيارية منها، ويتم تقدير الأداء فيها على أساس (ناجح - راسب)، ولكن هذا التقدير لا يدخل في حساب معدل علامات الطالب، ومن هنا فإن هذا النظام يشجع الطلاب على الدخول في مجالات جديدة للدراسة دون خوف من أن تنقص العلامة في هذا المجال أو المقرر من المعدل العام لدرجاتهم، ولكن عندما يكون نظام التقدير مكونا من فئتين فإنه يزودنا بمعلومات أقل عن مستوى الطالب في المقررات الدراسية بالمقارنة بنظام التقدير بالحروف، كما أن هذا النظام لا يشجع الطلاب على بذل الجهد في الدراسة والاستذكار في هذه المقررات، وإنما بذل أقل جهد ممكن. فإذا سجل الطالب في مقرر أو مقررين يقدران على أساس (ناجح - راسب)، فإنه بسهولة سيقرر تخصيص معظم وقته لدراسة واستذكار المقررات الدراسية الأخرى ذات الدرجات المتدرجة المستوى والتي يكون لها تأثير على معدله العام.

ويمكن التقليل من عيوب هذا النظام بتقييد أو تحديد عدد المقررات الدراسية التي يسمح بالتسجيل فيها وفق التقدير (ناجح - راسب).

4- أساليب مكملة لنظم التقدير:

هناك أساليب مكملة لأنظمة التقدير التقليدية المتبعة في المدارس، وخاصة في المدارس الابتدائية منها:

1- تقارير أو رسائل المدرسة للآباء:

الرسائل أو التقارير المصاحبة لكشف العلامات والموجهة للآباء أو أولياء الأمور تصف مستوى التحصيل الدراسي للطلاب كما تحدد حاجاته، وتتناول جوانب القوة ونقاط الضعف لديه، ثم تقترح خطة محددة لمعالجة النقاط الضعيفة وتحسين أدائه وتنميته. وتتميز الرسالة أو التقرير بأنها تشتمل على تفاصيل ومعلومات كثيرة تعطي صورة واضحة عن مدى تقدم التلميذ في المجالات المعرفية والشخصية.

وبالتالي فهي تعتبر تكملة جيدة للدرجات أو لأنماط التقارير الأخرى التي ترسلها المدرسة للآباء أو أولياء أمور الطلبة. ومع ذلك فلها مشاكلها أيضا، حيث تتطلب قدرا كبيرا من الوقت والمهارة، ويصعب تسجيل محتواها في سجل تراكمي مجمع، كما قد يساء فهمها أحيانا من قبل أولياء الأمور.

2- قائمة الأهداف التعليمية:

وهي تتكون عادة من مجموعة من الأوصاف السلوكية التي توضح مستوى أداء الطالب، ومدى التقدم الذي أنجزه نحو تحقيق الأهداف التعليمية في كل موضوع أو مقرر من المقررات الدراسية، ويتم تقدير مستوى الطالب في كل هدف من هذه الأهداف بواسطة تقدير رقمي متدرج (1، 2، 3، 4)، أو باستخدام الحروف (أ، ب، ج، د، هـ)، أو بإضافة نوعية (ممتاز، جيد جدا، جيد، مقبول، ضعيف)، أو (ممتاز - ضعيف)، أو (مرض - غير مرض). وتتميز مثل هذه القوائم بأنها تزود المعلمين وأولياء الأمور بتحليل جيد لنواحي القوة ونقاط الضعف في تحصيل الطالب، مما ييسر القيام بخطوات تعليمية محددة ومنظمة من قبل المعلمين، وبالتعاون مع الآباء أو أولياء الأمور المساعدة الطلاب وتحسين أو رفع مستوى تقدمهم وتحصيلهم الدراسي.

ويبقى أن نشير إلى أن هذه الأساليب لا تصلح منفردة لتقدير مستوى تحصيل المتعلم، إلا أنها تكون فعالة أكثر وتؤدي دورا هاما في فهم كشف درجات الطالب وتقديره المدرسي ومعرفة جوانب القوة ونقاط الضعف عنده، فضلا عن إعطاء ولي الأمر والأب معلومات أكثر تفصيلا عن الجوانب المختلفة المعرفية والانفعالية المتصلة بنمو ولده.

5- إرشادات لتطوير أنظمة وضع الدرجات وكتابة التقارير:

- حتى يكون لدينا نظام واضح لوضع الدرجات وكتابة التقارير داخل المؤسسات التعليمية المختلفة فإنه يجب أن يكون:
 - نظام تقديم المعلومات بصورة مبسطة وموجزة تناسب وتلائم جميع مستخدمي هذه التقارير، رغم تفاوت خلفياتهم الثقافية والأكاديمية.
 - نظام يحقق معظم الأغراض التي تستهدفها هذه التقارير.
 - نظام شامل ومفصل ويتضمن معلومات يفهمها المتعلمون والمرشدون والإداريون، وكذلك الآباء.
 - نظام تقبله كل مدرسة، لأنه يتفق مع ظروفها وامكانياتها ومسئوليات طلابها وحاجات مجتمعها.
- وحتى نضع نظاما يتميز بهذه الخصائص أو بمعظمها، فإننا نعرض لبعض الأسس أو المبادئ التي يمكن الاسترشاد بها لتحقيق ذلك، وهذه الأسس هي:

أ- يجب أن يستند نظام وضع العلامات وكتابة التقارير إلى أهداف تعليمية واضحة ومحددة:

الأهداف التعليمية هي التي يجب استخدامها كأساس عند وضع الدرجات وكتابة التقارير، وعند وضع نظام لتقدير الدرجات، لابد من تحديد أفضل طريقة لوضع نظام لتقدير الدرجات يكشف لنا عن مستوى تقدم الطالب نحو تحقيق هذه الأهداف.

ب- يجب أن يستند نظام وضع الدرجات وكتابة التقارير إلى تقويم واف وملائم:

عند كتابة التقرير يفترض أن يتم تقويم الأداء من قبل المعلمين بطريقة موضوعية، ولا يعتمد على أحكام تستند إلى تجميع واسترجاع أحداث عابرة، ولهذا فعند التخطيط لنظام وضع الدرجات وكتابة التقارير، يكون من الضروري أن وضع في الاعتبار أنواع بيانات التقويم اللازمة، وأن تكون الفقرات المتضمنة في التقرير هي تلك الفقرات التي يمكن للمعلمين أن يحصلوا على معلومات ثابتة وصادقة عنها.

ج- يجب أن يسترشد بالوظائف التي يخدمها نظام وضع الدرجات وكتابة التقارير عند وضعه:

يجب أن يكون نوع المعلومات التي يحتاجها مستخدمو التقارير موجودة فيه. وهذا يتطلب إجراء دراسة للوظائف التي يخدمها هذا النظام والمتعلقة بالطلاب، والآباء، والمعلمين، والمرشدين النفسيين، والقائمين على الإدارة. ورغم أنه يصعب تلبية جميع حاجات هذه الجماعات، إلا أنه يمكن تحقيق حل وسط مرض إذا عرفت هذه الحاجات وتم تحديدها. ومن المرغوب فيه تزويد هذه التقديرات بالحروف في كل مادة بتقرير منفصل عن أهداف المقرر الدراسي، والجهد المبذول، والخصائص الشخصية والاجتماعية، وعادات العمل.

د- يجب أن يوضع نظام وضع الدرجات وكتابة التقارير بشيء من التفصيل مما يجعله تشخيصيا، وبشيء من الإيجاز مما يجعله عمليا:

- وإذا أردنا توجيه التعلم والنمو عند الطلاب فإننا نحتاج إلى صورة شاملة ومعلومات عن نواحي القوة ونقاط الضعف لدى هؤلاء الطلاب، ولكن الرغبة في الحصول على معلومات مفصلة يجب أن تتوازن مع المطالب العملية التالية:
- يجب أن يعطى زمن معقول يكفي لإعداد التقارير واستخدامها.

- يجب أن تكون التقارير قابلة للفهم بوضوح من قبل الطلاب والآباء والمسؤولين في الإدارة المدرسية.
- يجب أن يكون من السهل تلخيص التقارير وحفظ الملخص في السجلات المدرسية.

والتوفيق بين الشمول والنواحي العملية يقتضي تزويد أو تكملة نظام التقدير الذي يستخدم الحروف، بتقارير مفصلة عن الجوانب الأخرى لنمو الطالب.

هـ - يجب أن يطور وينمو نظام وضع العلامات وكتابة التقارير بشكل يتعاون فيه الآباء والطلاب، والمسؤولون عن الإدارة في المدرسة:

فالتقارير المدرسية يمكن أن تحقق أعظم فائدة لها إذا أسهم في وضعها كل من يستخدمها، ويمكن أن يتم هذا بتشكيل لجنة أو هيئة تتألف من ممثلين للطلاب والآباء والمعلمين والمرشدين والمسؤولين عن الإدارة المدرسية، بحيث يقوم ممثلو هذه الجماعات بتزويد اللجنة بالأفكار والمقترحات، أو نقل هذه الأفكار والمقترحات إلى جماعاتهم التي يمثلونها، حتى تتم مناقشتها أو الموافقة عليها. وهذه المشاركة التعاونية تؤدي لوضع نظام مناسب لوضع الدرجات وكتابة التقارير المدرسية، كما تؤدي لفهم هذه التقارير فهما كاملا من قبل المهتمين بها ومستخدميها.