وزارة التعليم العالي و البحث العلمي.

جامعة العربي بن مهيدي أم البواقي.

كلية العلوم الاجتماعية و الإنسانية.

قسم العلوم الاجتماعية.

محاضرات مقياس:

القياس التربوي و بناء الاختبارات المدرسية.

موجهة لطلبة السنة الثانية علوم التربية.

السنة الجامعية: 2021/2020

الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne Démocratique et Populaire
وزارة التعليم العالي و البحث العلمي
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

جامعة العربي بن مهيدي – أم البواقي–
كلية العلوم لاجتماعية و الإنسانية
نيابة العمادة للدراسات ما بعد التدرج و البحث العلمي و العلاقات الخارجية

أم البواقي في:2021/12/07

مستخرج من محضر اجتماع المجلس العلمي للكلية

المنعقد يوم 2021/11/09

الموضوع:المصادقة على مطبوعة

بناء على محضر اللجنة العلمية لقسم العلوم الاجتماعية و على التقارير الإيجابية للخبراء: د.عداد وسام، د.سعادو أسماء ، د.معروف لمنور (جامعة خنشلة)، المعتمدين لتقييم المطبوعة الموسومة بـــ "محاضرات في القياس التربوي وبناء الاختبارات"، المقدمة من طرف الدكتورة: بوريو نضيرة- موجهة لطلبة السنة الثانية علوم التربية. تم اعتماد المطبوعة البيداغوجية على أن يتم وضع نسخة على الموقع الإلكتروني.

رئيس المجلس العلمي

رئيس المجلس العلمي
لكلية العلوم الاجتماعية و الانسانية
أ.د ضيف الله زدة

مقياس: القياس التربوي وبناء الاختبارات المدرسية (1):

السداسي: الثالث

أهداف التعليم:

أن يتعرف الطالب على القياس في التربية ونشأة القياس والتقويم في التربية ومجالات القياس وطرق قياس التحصيل الدراسي وحساب الصدق والثبات.

المعارف المسبقة المطلوبة:

- أن يكون للطالب دراية ببعض المبادئ والمعارف في علم النفس التربوي

محتوى المادة:

ـ مدخل إلى القياس في التربية

ـ نشأة القياس والتقويم في التربية

ـ مجالات القياس في التربية

ـ طرق قياس التحصيل الدراسي

ـ الصدق والثبات في القياس التربوي

ـ دور القياس في نجاح العملية التعليمية

* باقي العناصر في السداسي الرابع

طريقة التقييم: امتحان كتابي في نهاية السداسي بالنسبة للمحاضرات

تقييم متواصل خلال السداسي بالنسبة للأعمال التطبيق

السداسي الثالث: علوم التربية

وحدة التعليم	الحجم الساعي السداسي	الحجم الساعي الأسبوعي				المعامل	الأرصدة	نوع التقييم	
	14- 16 أسبوع	محاضرة	أعمال موجهة	أعمال تطبيقية	أعمال أخرى			متواصل	امتحان
وحدات التعليم الأساسية	180	6.00	6.00			10	20		
مذاهب ونظريات تربوية معاصرة	45	1.30	1.30	/	/	3	5	X	X
علم النفس التربوي (1)	45	1.30	1.30	/	/	3	5	X	X
طرق واستراتجيات التدريس الحديثة	45	1.30	1.30	/	/	2	5	X	X
إبستمولوجيا التربية	45	1.30	1.30	/	/	2	5	X	X
وحدات التعليم المنهجية	90	3	3	/	/	4	6	/	/
منهجية البحث التربوي (1)	45	1.30	1.30			2	3	X	X
القياس التربوي وبناء الاختبارات المدرسية (1)	45	1.30	1.30			2	3	X	X
وحدات التعليم الإسكتشافية	45	1,30	1.30	/	/	1	2	/	/
المادة الاختيارية: 1- علم النفس الاجتماعي المدرسي 2- علم نفس النمو 3- تاريخ الجزائر المعاصر	45	1.30	1.30	/	/	1	2	X	X
وحدات التعليم الأفقية	22.30		1.30			1	2		
لغة أجنبية	22.30		1.30	/	/	1	2	X	
مجموع السداسي 3	337.30	10.30	12.00	/	/	16	30	/	/

*– على كل طالب اختيار مادة واحدة فقط من الوحدة الاستكشافية.

قائمة المحتويات

المحتويات	الصفحة
أولا: مدخل للقياس و التقويم	
مدخل للقياس و التقويم	1
ثانيا: نبذة تاريخيّة عن القياس و التقويم	
نبذة تاريخيّة عن القياس و التقويم	3 -2
ثالثا: القياس	
1- مفهوم القياس	4
2- طبيعة القياس في العلوم السلوكيّة	5
3- مستويات القياس	6
أ) المقاييس الاسميّة	7-6
ب) المقاييس الرتبيّة	7
ج) المقاييس الفئويّة	8
د) المقاييس النسبيّة	9-8
رابعا: التقويم	
1- مفهوم التقويم	10
2- أهمية التقويم	11
3- العلاقة بين التقويم و القياس	12-11
4- خصائص التقويم الجيد	12
خامسا: مجالات القياس و التقويم التربوي	
1- مجالات تقويم التلاميذ	14-13
2- تقويم المنهج	15-14
3- تقويم المعلم	16-15
سادسا: قياس التحصيل الدراسي	
أدوات قياس التحصيل الدراسي	17
1- السجل القصصي	17
2- سلالم التقدير	18-17
3- قوائم الشطب	18
4- الاختبارات التحصيليّة	18
أ) تعريف الاختبارات التحصيليّة	19-18
ب) أنواع الاختبارات التحصيليّة	27-19
سابعا: ثبات الاختبار	
1- تعريف الثبات	29-28
2- أساليب تقدير الثبات	40-29
3- العوامل المؤثرة في الثبات	43-40
ثامنا: صدق الاختبار	
1- مفهوم الصدق	44-43
2- خصائص الصدق	46 -44
3- أنواع الصدق	50-46

4- العوامل المؤثرة في الصدق	52-50
تاسعا: دور القياس و التقويم في العمليّة التعليميّة	
1- اتخاذ القرارات	52
2- صياغة أهداف تعليميّة	53
3- زيادة الدافعيّة للتعلم	54-53
قائمة المراجع	56-55

أولا: مدخل للقياس و التقويم:

يعتبر القياس من أهم أدوات العلم الحديث، فبلا شك يحتل مكانة هامة في شتى ميادين الحياة وتبرز أهميته من خلال التجسيد اليومي لوحدات القياس في حياتنا اليومية.

إن موضوع القياس رغم كونه موضوعا كثير التداول في وقتنا الراهن إلا أن استعمال الإنسان له يعود إلى حقبة ما قبل التاريخ و لو بأشكاله البدائية، حيث أخذ بالتطور تدريجيا عبر العصور حتى صار على ما هو عليه الآن.

تُعنى الأنظمة التربوية في وقتنا الراهن بعملية القياس و التقويم لمعرفة مدى فعالية العملية التعليمية و السعي إلى تحسينها، إذ أن نتائج التقويم ضرورية بالنسبة للمتعلم من أجل تقويم أدائه ليتمكن من التعرف على مستواه ويبذل بدافع أكثر المزيد من المثابرة لتحقيق تحصيل أعلى. كما أنه ضروري بالنسبة للمعلم فيتأكد من مدى تحقق الأهداف التعليمية المسطرة.

يستهدف القياس التربوي التعرف على قدرات المتعلمين الخاصة و مواهبهم و استعداداتهم وميولهم الدراسية وذلك قصد تصنيفهم وتوجيههم إلى أنواع التعليم المختلفة التي تتوافق و ما لديهم من قدرات واستعدادات، و للقياس أسس ومبادئ لا بد من توافرها كما يستخدم في القياس أدوات واختبارات يجب أن تتوفر على صفات الاختبار الجيد المتمثلة في الصدق و الثبات.

ثانيا: نبذة تاريخية عن القياس و التقويم:

إن استخدام الإنسان للقياس و التقويم يعود إلى عصر ما قبل التاريخ ولو بأشكاله البسيطة البدائية، ويمكن القول أن تلك الممارسات قد ترجع منذ أن كان الإنسان يمارس صيد الحيوانات حيث كان يقدر حجم الصخرة التي يرميها للاصطياد، ويقدر الشخص الأكثر حجما و وزنا والأسرع حركة. فهذه الممارسات السلوكية اليومية تتضمن عمليات القياس و التقويم من خلال مقارنة الفرد بالآخرين الذين يعيشون معه.

كما سجل التاريخ أن بعض المجتمعات القديمة في بداية العصور التاريخية استخدمت وسائل في القياس حيث دأب الصينيون القدامى لاختبار الحكام و الإداريين لمختلف مقاطعات الصين كما أجريت امتحانات تحريرية كأساس للقبول و الترقية في الخدمة المدنية.

ودل تاريخ التربية أن الإغريق استخدموا القياس التحريري في تقدير نتائج التحصيل الدراسي (محمود، 2010، ص17) واعتبر سقراط أن التقويم اللفظي جزء من قياس نتائج التعلم. أما العرب في الجاهلية فقد قوموا نتائجهم الفكرية شعرا ونثرا من خلال عقد الندوات في الأسواق كسوق عكاظ، فيقوم خبير بإصدار أحكامه معتمدا على معايير متفق عليها. (الجلبي، 2005، ص5) وبحلول الإسلام تم مراعاة الفروق الفردية التي كانت واضحة وصريحة في القرآن الكريم كما قوم سلوك الناس وفق التعاليم الإسلامية.

مع حلول القرن التاسع عشر اخذ التقويم مجرى هاما في تطوره وهذا مع بداية الثورة الصناعية، ومر التقويم بمراحل صنفت كالتالي (علي و آخرون، 2009، ص36):

- فترة الإصلاح 1800 إلى 1900:

عام 1895 ظهرت نظرية التطور لدارون وكان لها الأثر الواضح على القياس النفسي عندما أكد على الفروق الفردية و التباين بين الأفراد وقد تأثر به فرنسيس جالتون الذي شق الطريق لحركة قياس الفروق الفردية على أسس سليمة، فأنشأ معملا في لندن لقياس قدرات الإنسان.

- **فترة ازدهار الاختبارات 1900 إلى 1930:**

ظهور أول مقياس للذكاء سنة 1905 وهو مقياس بينيه سيمون لتحديد المستوى العقلي للأطفال، وفي عام 1907 وضع المقياس الثاني الذي اشتمل على أعداد أخرى متزايدة من الأسئلة مقارنة بالمقياس الأول.

كما ظهرت اختبارات جمعية للذكاء اتجهت للحاجة العملية التي برزت في الحرب العالمية الأولى ثم عدلت وأجريت عليها العديد من المراجعات بعد الحرب ثم ظهرت اختبارات الاستعدادات والتحصيل المقنن و اختبارات قياس الشخصية. (المياحي،2011، ص24)

- **الفترة من 1930 إلى 1945:** ظهور تيلور الأب الروحي للتقويم وقد ركز اهتمامه على الأهداف المنشودة عند التقويم وأدى ذلك إلى ظهور ما يسمى بالمقاييس مرجعية المحك.

- **فترة الاستقرار 1945 إلى 1948:** استخدام نماذج تايلور وتطبيقها في المدارس الأمريكية وإدخال مقررات التقويم والقياس التربوي في إعداد المعلمين واعتبرت الاختبارات عنصرا رئيسيا في بناء النظم التعليمية.

- **فترة الازدهار و التوسع 1948 إلى 1972:** تم بناء عدد من البرامج التقويمية الهامة المتنوعة في أمريكا. استخدمت نماذج جديدة للتقويم الكيفي مثل نماذج النظم.

- **الفترة من 1973 حتى الآن:** ازداد الاهتمام بإصلاح التعليم في الثمانينات مما زاد من الاهتمام بالتقويم ووجود متخصصين قادرين على التطوير والتغيير المطلوب.

ثالثا: القياس:

1. مفهوم القياس:

كثيرا ما نتناول مصطلح القياس في أحاديثنا اليومية وتعاملاتنا مع الظروف الحياتية لأن مفهومه أصبح مألوفا وضروريا. ومع ذلك وردت تعاريف عديدة للقياس تباينت عند العلماء وقد عُرف القياس:

لغة: يقال قاس و إقتاس ويعني قاس الشيء بمثله أو على مثله، وهذا يعني تقدير الشيء من خلال مقارنته بغيره.

اصطلاحا: أورد قاموس ويبستر كلمة القياس بعدة معاني منها التعرف على مدى الشيء أو طاقته أو أبعاده من خلال معيار، كما جاء فيه بأن القياس يعني النشاط والجهد و العمل الذي نقوم به من أجل تحديد مدى وطاقة وأبعاد الشيء (الطربري،1997، ص4). وبالرجوع إلى القاموس الفرنسي نجد أن تعريف القياس معناه إعطاء قيم كمية أو عددية لبعض الصفات أو الخاصيات لشخص ما (مديرية التكوين، 2008، ص183).

عرف كامبل Campell القياس على أنه تمثيل للصفات أو الخصائص بأرقام وهو نفس المعنى الذي يراه جيلفورد Gelford بأن القياس وصف للبيانات أو المعطيات بالأرقام. (عبد الحميد، 2008، ص10). فكلاهما يرى أن القياس هو الارتباط بالقيم العددية.

في حين عرفه ستفنز Stevens: " القياس عملية التعبير الكمي عن الخصائص و الأحداث بناء على قواعد وقوانين محددة ". (النبهان، 2004، ص22) أضاف ستفنز أن القياس يتضمن الأرقام التي تكون مناظرة لفئة من الخصائص وهذا طبقا للقواعد محددة.

فالقياس عملية تكميم أو التعبير بلغة كمية عن صفات أو عوامل أو ظواهر لموضوعات نوعية أو معنوية أو سلوكية تتطلب إصدار حكم.

2. طبيعة القياس في العلوم السلوكية:

يتم القياس على الظواهر الطبيعية بطرق مباشرة فهي عمليات مقننة ومتفق عليها، وتتعامل مع ظواهر تعريفها مضبوط ووجودها ملموس كما يمكن تفسير نتائجها بدرجة عالية من الثقة والموضوعية. ومثال ذلك قياس الأطوال و المسافات والأوزان والزمن والحرارة فكلها تقاس بوحدات قياسية متفق على حدودها وتكون ثابتة دائما، فالقياس الطبيعي يمتاز بالتحديد الكمي وثبات وحدات القياس.

أما القياس في العلوم السلوكية عموما و العلوم التربوية والنفسية خصوصا فيختلف عن ما هو عليه في العلوم الطبيعية، كون طبيعة السلوك الإنساني و طبيعة الموضوعات التربوية والنفسية تتميز بدرجة من التعقيد وعدم الوضوح، الأمر الذي يجعل قياسها أكثر صعوبة.

إن الصفة أو الخاصية المقاسة في مجال التربية وعلم النفس ذات طبيعة خاصة يتعذر التعامل معها مباشرة، فالذكاء والتحصيل الدراسي و الاستعداد والدافعية وغيرها موضوعات غير ملموسة يتم التعامل معها من خلال التعامل مع الآثار الناجمة عنها وهو شأن أغلب الظواهر السلوكية، وعليه قياسها يكون استدلاليا غير مباشر. كما أن صعوبة قياس الخاصية التربوية النفسية يكمن في التداخل و التشابك فيما بينها فترتبط بظواهر وعوامل يصعب عزلها بشكل مستقل لقياسها، فقياس التحصيل الدراسي لتلميذ ما لا يمكن أن يكون بمعزل عن بعض الخصائص الأخرى له مثل أن يكون خائفا عند إجراء الاختبار، فهذا الخوف الذي يعانيه أثناء قياس تحصيله الدراسي ستكون له أثاره على أداءه في الاختبار ومن تم تكون نتيجة الاختبار غير دقيقة. لذلك يكون القياس التربوي والنفسي نسبيا وليس مطلقا، معرضا للخطأ أكثر مما في العلوم الطبيعية.

تقاس سمات السلوك الذي نستدل به على الخاصية النفسية والتربوية بافتراضات لوحدات سلوكية يعبر عنها في الغالب بفقرات الاختبارات لأي من مظاهر السلوك، التي كلما ازدادت تعقد القياس وارتفعت احتمالات الخطأ (الجلبي، 2005، ص19). مما يجعل القياس في مجال العلوم السلوكية يتطلب اختبارات دقيقة للقياس تشتمل على صفات الصدق و الثبات و القدرة على التمييز

3. مستويات القياس:

يعد التعرف على مستويات القياس من الأساسيات التي ينبغي الباحث في العلوم السلوكية التعرف عليها ليستخدم الطرق الإحصائية المناسبة لتحليل البيانات الخاصة ببحوثه استخداما مناسبا وذلك من أجل الوصول إلى نتائج صادقة و قرارات ذات فعالية.

لقد قدم ستيفنز STEVENS عام 1951 تصنيفا لمستويات القياس النفسي معتمدا على كيفية استخدام الأعداد في كل منها ولكل مستوى قواعده وضوابط استخدامه، تصنف مستويات القياس حسب ستيفنز إلى أربعة أنواع وهي في مجموعها ذات صفة تراكمية بمعنى أن مستويات القياس متضمنة فيما بينها فمثلا يتضمن المقياس الرتبي خصائص المقياس الاسمي وكذلك يحتوي المستوى الفئوي خصائص كل من المستوى الرتبي و الاسمي، ويتضمن المستوى النسبي كافة خصائص المستويات الأخرى. (النبهان، 2004، ص 27)

أ) المقاييس الاسمية:

يقع هذا النوع من المقاييس في أدنى مستويات القياس، حيث يستخدم العدد لتصنيف الأحداث أو الأشياء أو الأفراد في مجموعات وفقا لبعض الخصائص النوعية كأن نوزع الأفراد حسب جنسهم (ذكور، إناث)، أو حسب مناطقهم السكنية (شرق، غرب، جنوب، شمال). العدد المستخدم يشير فقط إلى اسم الفئة التي تم فيها التصنيف مثلا الذكور رقم (1) والإناث (2).

لا تعمل هذه المقاييس بأكثر من تصنيف الأشياء من أجل التمييز بينها اعتمادا على افتراض أن الأفراد يختلفون في صفة ما كالجنس، اللون، الخبرة، السن، المنطقة الجغرافية. ولتسهيل التعامل مع هذه المتغيرات بات تكميمها والتعبير عنها رقميا (النبهان،2004، ص28)

إن الأرقام في هذا المستوى أشبه بالأسماء فهي لا تتضمن معنى الأفضلية (الأكبر والأصغر) فالطالب الجامعي الذي رقمه 100 لا يعني أنه أفضل من الطالب الذي رقمه 85. كما أن هذه الأعداد المميزة لا يمكن أن تجري عليها أي من العمليات الحسابية الأربع (الجمع، الطرح، الضرب، القسمة) والإمكانية المتاحة هي عد عدد الحالات الموجودة في كل مجموعة أي تكرار كل فئة.

ب) المقاييس الرُتبية:

يعتبر هذا النوع من المقاييس تاليا من حيث المستوى للمقاييس الاسمية، حيث يقوم بالإضافة إلى تصنيف الأفراد و الأشياء في مجموعات متمايزة، فهو يرتب الأفراد في سلسلة تصاعديا أو تنازليا بناء على الخاصية المطلوب قياسها.

إن الأرقام التي تعطى للأشياء أو الأفراد وفق هذا المقياس لا تمثل كميات معينة، كما أن المسافات الفاصلة بين رقم و أخر لا يشترط أن تكون متساوية، فمثلا عند ترتيب خمسة طلاب حسب درجة مشاركتهم، يعطى رقم (1) لأكثر مشاركة و رقم (5) لأقل مشاركة، فإن الفرق بين الثاني و الثالث لا يساوي الفرق بين الثالث و الرابع. كذلك لا يشترط أن تكون مشاركة الطالب الأول مساوية لخمس أضعاف مشاركة الطالب الخامس. فالمقياس الرتبي لا يعطي صورة واضحة عن حجم الفروق الموجودة بين الأفراد المتجاورين في أية مجموعة. (النبهان ،2004، ص 29).

وعليه فالمقياس الرتبي لا يسمح لنا بتحديد الفروق بين رتبتين بدقة و الواقع إن الرتبة تعطي معنى كمي عام غير دقيق، حيث يسمح لنا فقط بترتيب الأفراد من الأدنى إلى الأعلى أو العكس دون أن تقدم لنا معلومات دقيقة عن كم و مقدار الخاصية موضوع القياس، أي أن المقاييس الرتبية تحدد لنا مرتبة الشيء أو مكانته في اختبار(مقياس) يقدم وضعا كيفيا مثل: الأول، الثاني، الأخير.

في المقاييس الرتبية أيضا لا يمكن استخدام العمليات الحسابية الأربع، ويمكن التعامل مع الرتب باستخدام المقاييس الإحصائية كمعامل الارتباط الرتب لأنه يعتمد في جوهره على ترتيب الأفراد في الصفة المقاسة.

ج) المقاييس الفئوية:

تسمى كذلك بمقاييس المسافة، ومقاييس الوحدات المتساوية ويتم فيها وصف الشيء وصفا كميا في ضوء قواعد متفق عليها. فما يميز هذا المقياس عن مقياس الرتب هو أنه يمكننا من معرفة مقدار المسافة بين شيئين أو شخصين. (تايلر ترجمة عبد الرحمان، 1982، ص22) يعد مقياس المسافة أكثر دقة من المقياسين السابقين، حيث تتساوى الفروق بين الأقسام المتتالية في السمة موضوع القياس، فهو يتميز بوحدات متساوية تسمح لنا بتحديد مدى بعد شيئين أو فردين عن بعضهما في الخاصية التي نقيسها من خلال استخدام وحدات أو مسافات متساوية. لذلك فالعمليات الحسابية المسموح بها في هذا المستوى من القياس هي عملية الجمع و الطرح دون الضرب و القسمة. فعندما تكون درجات أربعة تلاميذ هي 15،13، 17، 19 فإنه يمكن القول بأن الفرق بين (13) و (15) هو 2 ويساوي الفرق بين (17) و(19)، و لكن لا يجوز أن نقول أن نسبة الذكاء 140 هي ضعف نسبة الذكاء 70 ومنه فإن في هذا المستوى من القياس لا يمكننا أن نقول أن درجة فرد ما هي ضعف درجة فرد آخر، أو ثلاثة أمثالها.

إن الفرد الذي يحصل على الدرجة 0 لا يعني أنه لا ذكاء له لأن الصفر في هذه المقاييس اعتباطي. ومن الأمثلة على المتغيرات التي تقاس على المقياس الفئوي معدلات الذكاء و درجات الطلاب على الاختبارات التحصيلية. (النبهان، 2004، ص 30) أما الإجراء الأكثر شيوعا التي تلجأ إليه أغلب المقاييس النفسية هو حساب المتوسط الحسابي لدرجات الصفة المقاسة و الانحراف المعياري لهذه الدرجات. (أحمد عمر ،2010، ص 86)

د) المقاييس النسبية:

أعلى مستويات القياس، وتتميز بخصائص جميع المقاييس السابقة ، إضافة إلى وجود صفر مطلق أي حقيقي. إن صفة الصفر الحقيقي والوحدات المتساوية تجعل من مقياس النسبة من أرقى المقاييس. فتصلح معها كل العمليات الحسابية من جمع وطرح وضرب وقسمة و كل الطرق الإحصائية.

إن هذا النوع من القياس لا نتعامل معه في القياس النفسي لأننا لا نمتلك صفرا مطلقا للخصائص الإنسانية ويمكن استخدامه في قياس الخصائص العقلية إذا ما تم قياسها بوحدات فيزيائية كأن نقيس زمن الرجع أو التعلم بوحدات زمنية. (الجلبي ،2005، ص 21)

وفي الختام يمكن تلخيص مستويات القياس في هذا الجدول الذي يبين خصائص كل منها مع العمليات الممكنة في كل مستوى.

جدول يمثل مستويات القياس

مستوى القياس	خصائص القياس	أمثلة	العميات الممكنة
النسبي	– صفر حقيقي – تمثل الأرقام وحدات متساوية عن موقع الصفر الحقيق.	العمر، السرعة، الزمن.	جميع العمليات الحسابية وجميع الأساليب الإحصائية ومقاييس النزعة المركزية و مقاييس التشتت.
الفئوي (المسافة)	– مقدار المسافة بين قيم المتغير الذي يخضع للقياس معلوم. – الصفر افتراضي.	درجات الاختبارات التحصيلية ودرجات الذكاء	يمكن إجراء عمليات الجمع و الطرح. جميع مقاييس النزعة المركزية و التشتت.
الرتبي	تمثل الأرقام رتبا للأفراد حسب درجة امتلاكهم للسمة المقاسة .	ترتيب الطلبة حسب متغير : المواظبة ، المستوى الاجتماعي و الاقتصادي	لا يسمح بإجراء العمليات الأربع . يستعمل معامل إرتباط الرتب . يمكن استعمال الوسيط و الميئينيات.

التصنيفي (الاسمي)	الأرقام تستخدم للتصنيف فقط و لا تمثل أي صفة كمية	الرقم الجامعي ، أرقام الهواتف و المنازل ، الجنس ، اللغة ، الخبرة	لا يسمح بإجراء أي عمليات حسابية . يمكن استعمال المنوال .

رابعا: التقويم:

1. مفهوم التقويم :

التقويم في قواميس اللغة لفظ مشتق من الفعل "قوم " وقوم الشيء بمعنى قدره و وزنه وحكم على قيمته وعدله .

أما اصطلاحا فقد تباين تعريف التقويم حسب علماء التربية والمختصين في طرق التدريس والمناهج .

حيث يرى بلوم bloom أن التقويم هو إصدار حكم ما على الأفكار و الأعمال و الحلول وطرق التدريس وغيرها من الأمور التربوية .

بينما يرى جرولند GRONLUND " التقويم عملية تنظيمية لتحديد المدى الذي يحقق فيه التلاميذ الأهداف التربوية الموضوعة ".(محمد علي وآخرون،2009 ،ص32)
وترى الحريري أنه "عملية حكم على أهمية شيء ما من حيث الفعالية ويكون مقارنا أو نسبيا ، يتم في ضوء معايير متفق عليها أو محكات اختبارية . (الحريري ،2007 ،ص14)

أما يونس فيعرف التقويم بأنه العملية المنهجية التي تتضمن جمع المعلومات عن سمة معينة بالقياس الكمي أو غيره واستخدام المعلومات في إصدار الحكم على هذه السمة في ضوء أهداف محددة سلفا لمعرفة مدى كفايتها .(يونس ،2008،ص179)

إن مفهوم التقويم يقودنا إلى مقارنة النتائج المحصل عليها في الوضعيّة التعليميّة للوصول إلى اتخاذ قرارات (Dominique, 2005, p25).

وعليه فإن التقويم عملية جمع البيانات المتعلقة بالجانب التربوي المراد تقويمه ومعالجة تلك البيانات بطريقة مناسبة للحكم على مدى تحقيق الأهداف التي وضعت ومن ثم العمل على التحسين و العلاج.

2. أهمية التقويم:

ترجع أهمية التقويم كونه عنصرا أساسيا في العمليّة التعليميّة التعلميّة إلى ما يلي:

- معرفة مدى تحقيق الأهداف المرسومة للبرنامج المحدد.
- تزويد المتعلمين بالتغذية الراجعة.
- الكشف عن مدى ملاءمة المنهاج للمرحلة العمريّة و النمائيّة للمتعلمين.
- يساعد التقويم في الحكم على مدى فاعليّة أسلوب التعليم.
- يساعد التقويم المعلم حول مدى كفايته في التدريس و اختيار الوسائل و الطرق المناسبة.
- يوفر التقويم بيانات ودلالات للمخططين و المسؤولين عن التربية و واضعي البرامج التعليميّة.
- تزويد أولياء التلاميذ ببيانات حول مدى تقدم أبناهم في الدراسة.

3. العلاقة بين التقويم و القياس:

كثيرا ما يحدث خلط والتباس في مفهوم القياس و التقويم لأنهما عمليتان متلازمتان ، فالقياس وصف كمي لمقدار السمة التي يمتلكها الفرد، فإذا حصل المفحوص على درجة 15 في اختبار تحصيلي ما فإن هذا يعد قياسا لكن الحكم على تلك الدرجة وتفسير المستوى التي تدل عليه يعد تقويما .لأن التقويم عملية إصدار حكم انطلاقا من وصف كمي لمستوى الأداء وعليه :

- القياس جزء من التقويم وخطوة من خطواته, فالقياس أساس مهم لتسهيل عملية القيام بالتقويم
- القياس مفهوم ضيق لا يتعدى إعطاء فكرة معينة عن السمة المقاسة ،أما التقويم فهو المفهوم الأعم و لأشمل فيضم القياس و وسائل أخرى .
- يهتم القياس بالجانب الوصفي و التقدير الكمي للقيم العددية في حين يهتم التقويم بالجانب النوعي وإصدار الأحكام تبعا لمعايير محددة متفق عليها.

– يعد القياس أكثر موضوعية كونه كمي رقمي لكنه لا يقدم لنا حكما على النتائج فبالتقويم نفسر ونقدر النتائج وفق معايير وضوابط محددة.

4. خصائص التقويم الجيد:

– الهدف، فعلى المقوم أن يحدد أهدافا للعمليّة التقويميّة يسعى لتحقيقها.

– الشموليّة، بأن يتناول الجوانب الرئيسية للوحدة الدراسيّة كما يشمل جميع المهارات التعليميّة (معرفيّة، وجدانيّة...).

– الاستمراريّة: فيجب أن تكون عمليّة التقويم مستمرّة تبدأ من نقطة تحديد الأهداف إلى معرفة خصائص الفئة المستهدفة، إلى التنفيذ و اختيار الطرق المناسبة و استخدام الوسائل التعليميّة الملبية للحاجات إلى آخر العمليّة التعليميّة (علي وآخرون، 2009، ص82).

– التعاون: فيجب أن يشترك في التقويم أطراف الفريق التربوي من معلم و متعلمين و إدارة وحتّى أولياء التلاميذ، فعمليّة التقويم عمليّة تفاعليّة تعاونيّة.

– اتصاف أدوات و وسائل التقويم بالصدق و الثبات و الموضوعية.

– إمكانيّة التنفيذ: أي أن تكون أساليب التقويم ممكنة التطبيق فلا تكون معقدة يصعب تنفيذها و لا حتى باهظة الكلفة يضطر المقوم إلى إهمالها.

– المرونة: حيث تكون قابلة للتعديل وفقا للظروف الطارئة أو إذا ما صادفتها عوائق.

– التنوع: يجب أن تتنوّع أساليب و أدوات التقويم حتى تسمح بالحصول على أكبر عدد ممكن من المعلومات.

خامسا: مجالات القياس و التقويم التربوي

لقد اتسعت مجالات استخدام القياس النفسي فشملت جميع ميادين الحياة واستخدم القياس في مجالات التربية لأغراض عديدة كالتالي :

1. مجالات تقويم التلاميذ :

تتعدد المجالات التي يشملها تقويم التلميذ ومنها :

أ- **قياس التحصيل الدراسي:** قد تظهر الاختبارات التي يعدها المعلمون ضعفا في ثباتها وأحيانا في صدقها و لكن بتنوعها يمكن أن تقيس تحصيل الطلبة ، فالامتحانات اليومية و الشهرية و الفصلية و السنوية يمكن أن تعطي صورة عن تحصيل التلاميذ بينما تستطيع الاختبارات التحصيلية المقننة المعدة من قبل جهات متخصصة قياس تحصيل التلاميذ بدرجة عالية من الموضوعية والدقة . (الجلبي،2005 ،ص 29)
فتسمح الاختبارات للمعلم من الكشف و التعرف على المهارات التي تحققت عند المتعلم و التي لم تتحقق ليعمل على علاج نواحي الضعف و تنميتها .

ب- **وظيفة إعطاء الدرجات لاتخاذ القرار** : تمثل الدرجات التي يحصل عليها المتعلم من جراء الاختبار التحصيلي بمثابة مخرجات تساعد في انتقال المتعلم و اتخاذ قرارات بشأن نجاحه وانتقاله من صف لأخر أو منحه شهادة أو ما إلى ذلك .

ت- **تصنيف التلاميذ** : تساعد الاختبارات والمقاييس المعلم في تصنيف تلاميذه من حيث ما يملكون من قدرات خاصة وذكاء كما يكتشف التلاميذ الذين يعانون من ضعف في القدرات العقلية .

ث- **تقويم استعداد التلميذ للتعلم:** يعني مدى قابلية التلميذ للتعلم، أو اكتساب مهارات جديدة تؤهله لاكتساب مهارات أخرى أعمق تعده للحياة، ويتأثر الاستعداد للتعلم بعاملين هما النضج و الخبرة (المبروك ،2016، ص28). وتوجد العديد من الاختبارات التي تقيس الاستعداد سواء لدراسة معينة أو تخصص و حتى لوظيفة ما.

ج- **تقويم القدرات العقلية (الذكاء)** : يعتبر الذكاء عامل مهم يؤهل المتعلم للالتحاق بالمدرسة و تعلم مختلف العلوم ،فقد أقر مجمع اللغة العربية استخدام الذكاء بمعنى القدرة على التحليل و التركيب و التمييز وعلى التكيف . لهذا تساعدنا اختبارات الذكاء في تشخيص الصعوبات التي يعاني منها المتعلم. توجد عدة مقاييس للذكاء وضعها العلماء أهمها مقايس دافيد وكسلر للراشدين(wais) وللأطفال (wisc) ولأطفال ما قبل المدرسة (wppsi) .

ح- **تقويم الشخصية:** يستند قياس الشخصية على أن الأفراد يختلفون فيما بينهم ، وأن دراسة الشخصية تتطلب إمكانية تقدير وقياس هذه الفروق .كما تشمل كل الجوانب المعرفية وغير المعرفية التي تتضمن الجوانب الميزاجية و الاجتماعية أيضا.(الجلبي ،2005 ،ص347) ،ويؤثر في شخصية الفرد عدة عوامل منها الوراثية ،كما يؤثر التكوين الجسمي للفرد على شخصيته وكذلك عامل البيئة الثقافية و الاجتماعية و الأسرية التي يعيش فيها، ودوره بالنسبة لغيره من أفراد الجماعة ، ومرحلة طفولته وغيرها من العوامل التي لها تأثير على شخصية الفرد وتعلمه (المبروك ،2016 ، ص29).فقياس شخصية المتعلم تسهل على المعلم اختيار الطريقة المثلى للتعامل معه ومن اختبارات الشخصية : اختبار تفهم الموضوع و اختبار بقع الحبر الروشاخ .

خ- **خدمة التوجيه التربوي:** حيث تقاس التلاميذ قدرات وميولهم واستعداداتهم الدراسية المختلفة وعلى أساس منها يمكن للإدارة المدرسية أن توجههم على أنواع التخصصات و الشعب التي تتناسب وتلك القدرات و الاستعدادات.

2. تقويم المنهج: يحتاج المنهج للتقويم ليواكب التطور العلمي و التكنولوجي الذي يشهده العصر فيلاءم احتياجات الفرد و المجتمع ،وعملية تقويم المنهج الدراسي قد تشترك فيه أكثر من جهة ، كل حسب مجاله والأهداف الموضوعة التي يسعى إلى تحقيقها فالمعلم يركز اهتمامه على الجانب المعرفي و الأنماط التدريسية التي يمكن أن يستخدمها لتحقيق ذلك وهناك المؤسسات التربوية التي توجه اهتمامها نحو التخطيط السليم و الإجراءات التي قد تتخذها باتجاه تحقيق أهدافها،(المياحي ،2011 ،ص59) كذلك الجهة الوصية من خلال الإصلاحات التي تشهدها هذه المناهج .

تشتمل عملية تقويم المنهج على عناصر هامة قد تشكل أساسا في عملية الإعداد و التخطيط التي يمكن أن يتوقف عليها نجاح المنهج أو فشله ومن أهمها:

- **الأهداف:** وتشمل الأهداف التربوية العامة أو المرحلية والسلوكية، فالحكم على مدى تحقيق الأهداف بمثابة الحكم على نجاح العملية التربوية.
- **المحتوى الدراسي:** ويتضمن الموضوعات و الكتاب المدرسي.
- **الأنشطة التعليمية وطرائق التدريس** المستخدمة لتطبيق المنهج.
- **الوسائل التعليمية** المستخدمة في تنفيذ المنهج بما فيها الوسائل القديمة و الحديثة.

يتم تقويم المنهج الدراسي عن طريق الملاحظة من خلال تطبيق المعلمين للمنهج وملاحظاتهم حوله ومن خلال الاستبيان الذي يشارك فيه المعلمون ويتناول تحليلا لجميع عناصر المنهج كذلك تحليل مضمون المنهج عن طريق متخصصين و التعرف على مدى تطابق المحتوى مع الأهداف المرسومة (المياحي ،2011، ص31). فنقيس درجة كفاءة الكتب و المقررات والأنشطة و الفعاليات المدرسية.

3. تقويم المعلم: يشير مفهوم تقويم أداء المعلم إلى جمع أدلة بشأن ما يقوم به المعلم من ممارسات مهنية سواء في حجرات الصف أو خارجها فقد رأى دانيل" أن تقويم أداء المعلم يتضمن مجموعة من الإجراءات التي تتخذ لجمع الأدلة و المعلومات عن جودة أداء المعلم و التي تحتاج إلى بيئة مدرسية صحية، يسودها مناخ من الثقة و الاحترام المتبادل و الموضوعية فضلا عن ضرورة وضوح الهدف من تقويم أداء المعلم و القائمين على تقويمه بحيث يشترك الجميع في رؤية موحدة بشأن تحسين أداء المعلم الذي ينعكس بدوره على تحسين أداء التلاميذ. (DUKE,1995, p65-66)

ويشير البيلاوي و آخرون أن الهدف من تقويم المعلم يتمثل في تحديد نقاط القوة في أداء المعلم لتدعيمها و التعرف على نقاط الضعف لاتخاذ التدابير العلاجية المناسبة. كذلك تحفيز المعلم على تدعيم ذاته ومراجعة أداءاته بهدف تطويرها و تحسينها. كما يساعد تقويم أداء المعلم في اتخاذ بعض القرارات المتعلقة بالترقية و المكافأة. (البيلاوي و آخرون، 2006، ص119). مما سبق نرى أن أهداف تقويم المعلم تكمن في تحسين نوعية التدريس المقدم للتلاميذ من خلال

توجيه التلاميذ نحو التعلم الفعال، ومنه تحسين ممارسات المعلم المهنية سواء من أجل التحسين الذاتي أو من أجل الترقية إلى وظائف أخرى.

للحصول على معلومات يمكن الاعتماد عليها بشأن تقويم أداء المعلم يتعين استخدام أكثر من أداة ولعل من أهمها: الملاحظة، التقويم الذاتي، تقويم الأقران، تقويم الأولياء، ملف الانجاز المهني، المنتجات الصفية، التقارير، نتائج التلاميذ. (الصغير، 2008، ص 14 -15)

- **مجال البحث العلمي**: تستخدم الاختبارات كأدوات في إجراء البحوث و الدراسات التربوية.

سادسا: قياس التحصيل الدراسي:

أدوات قياس التحصيل الدراسي:

تتمثل في مجموعة الأدوات التي يستعملها القائم على التعليم والتي تحدد مدى تحصيل المتعلم مقارنة بالأهداف التعليمية المسطرة كما تمكنه من اتخاذ القرارات بشأنه، فقد أصبح المتعلم يُقيَّم طبقا لمستوى أدائه باستخدام طرق أصيلة متنوّعة تتطلب من المتعلّم أعمالا كثيرة في أوقات مختلفة بدلا من عمل واحد في وقت واحد يقتصر على أداء التلميذ في ورقة الامتحان التقليدي.

1 – السجل القصصي: وهو عبارة عن وصف قصير من المعلم ليسجل ما يفعله الطالب و الحالة التي تمت عندها الملاحظة، أو هو وصف لسلوك يقوم به الطالب سواء كان السلوك لفظيا أو عمليا (متروك، 2015، ص417) و توثق الملاحظة بشكل كامل من حيث التاريخ و الموقف، كما تضاف بعض التفسيرات المحتملة سواء كانت إيجابيّة أو سلبيّة، و تُجمع هذه الملاحظات لتعصي مؤشرات صادقة في التعرف على اهتمامات الطالب و سلوكه، و توظَف لأغراض إرشاديّة، توجيهيّة و علاجيّة (علي و آخرون، 2009، ص108).

2 – سلالم التقدير: هو أداة بسيطة تظهر فيها إذا كانت مهارات المتعلم متدنية أو مرتفعة، حيث تخضع كل فقرة لتدريج من عدّة فئات أو مستويات، حيث يمثل أحد الطرفين انعدام أو وجود الصفة التي نقدّرها بشكل ضئيل و يمثل الطرف الآخر تمام أو كمال وجودها و ما بين الطرفين درجات متفاوتة من وجودها (متروك، 2015، ص418).

يقوم الفاحص أو الأستاذ بتحديد درجة توفر الصفة في سلوك المفحوص بأن يعلّم على قيمة التقدير المناسب مثلا: ممتاز– جيّد جدا– جيّد– مقبول– ضعيف.

عند التفكير في وضع سلّم تقدير يمكن الاعتماد على الاعتبارات التالية (النمر و آخرون، 2020، ص79–98):

أ– حدد أبعاد الصفة أو الناتج أو الأداء الذي يريد ملاحظته و تقديره.

ب– عرف هذه الأبعاد حيث يتضمن التعريف ناحية واحدة فقط.

ج- لا تصمم سلما لتقدير نواحي سلوكيّة لا يمكن ملاحظتها بصورة مباشرة.

د- حدد معاني الألفاظ الدالة على درجة وجود السمة أو الأداء.

هـ- بسط لغة القياس و اجعل ألفاظها واضحة و محددة.

3 - قوائم الشطب: تعرّف قوائم الرصد أو الشطب على أنها عبارة عن قائمة الأفعال/ السلوكات التي يرصدها المعلم أو الطالب أثناء تنفيذ مهمة أو مهارة تعليميّة و تعد من الأدوات المناسبة لقياس مدى تحقيق النتاجات التعليميّة و يستجاب على فقراتها باختيار إحدى الكلمتين من الأزواج التالية: (مُرضي- غير مُرضي) (نعم- لا) إلى غيرها، و تعد من أسهل أدوات التقويم إعدادا و تنفيذا و تصحيحا.

تتكون قائمة الرصد من مهارات و مفاهيم يعدّها المعلم و يقوم بشرحها مسبقا للطلبة و توضيح كيفيّة التعامل معها (العدوان و آخرون، 2016، ص213).

4 - الاختبارات التحصيلية:

أ) مفهومها: هي مقاييس للكشف عن أثر تعليم أو تدريب خاص، موضوعها واقع المواد التي درسها التلميذ (الفاخري، 2018 ،ص69) .

فالاختبار التحصيلي هو الاختبار الذي يقيس درجة التعلم والنجاح والانجاز وقد عرفه MFHREN و LEHMAN بأنه اختبار يقاس به مدى تحصيل الفرد لشيء من الأشياء أو اكتسابه لمعرفة محددة أو مهارات معينة وفي الغالب تتم هذه الأشياء نتيجة تعليم أو تدريب مخطط منظم .كما تعرف على أنه تلك الأدوات أو الوسائل التي يستعين بها المعلم لقياس مستوى تحصيل تلاميذه في مجال معرفي يتضمنه برنامج دراسي محدد أو مادة دراسية مقررة بهدف تحديد مدى تحقيق الأهداف التعليمية .(عبد الرحمان ، 1997 ، ص281) يمكن للاختبارات التحصيلية أن تطبق من طرف الأستاذ بشكل مستمر من أجل الوقوف على تحصيل المتعلم في مادة دراسية .كما قد تكون اختبارات تحصيلية مقننة للوقوف على عينة أداء المتعلم وتكون شاملة لمقرر دراسي مثلا وتطبق وفق تعليمات محددة وقواعد بناء خاصة . تعتبر الاختبارات التحصيلية أكثر الأدوات شيوعا في التقويم التربوي وتستخدم على مدار العام الدراسي

لغايات مختلفة كالتشخيص و التصنيف والتوجيه أو للانتقال من مستوى لأخر.
(محمود أحمد ،2010، ص97)

ب) أنواع الاختبارات التحصيلية:

للاختبارات التحصيلية المدرسية أهمية بالغة في العملية التعليمية وتكمن تلك الأهمية من خلال الدور الذي يمكن أن تلعبه في حياة الأستاذ و التلميذ وهي ثلاثة أنواع:

1 – الاختبارات الشفوية: تعتبر من أقدم الطرق التي استخدمت في تحديد استيعاب المتعلمين للدروس التي تلقوها، ويقال أن سقراط استعمل الاختبارات الشفوية منذ القرن الرابع قبل الميلاد للوقوف على مستوى مستمعيه لكي يبني تعليمه لهم على أساس خبرتهم الماضية.
(الجلبي،2005، ص2015)

تعرف الاختبارات الشفوية على أنها وسيلة اختبارية تستخدم في عملية التقويم التحصيلي للطلاب بهدف قياس قدرتهم على النطق و التعبير اللفظي و مدى فهمهم أو استيعابهم لموضوع الدرس عن طريق توجيه بعض الأسئلة الشفوية و تكون الإجابة المنتظرة شفوية أيضا .(المياحي ،2011 ،ص 92)

إذن تهدف الاختبارات الشفوية إلى التعرف على مدى إتقان المتعلم للمادة الدراسية بمعزل عن الكتابة.

من أهم مزاياها:

– تلعب دورا هاما في حث التلاميذ على استذكار موضوعاتهم وفسح المجال الواسع أمام معلميهم للتمكن من متابعتهم و الوقوف على مواطن قوتهم و ضعفهم التحصيلي. وهو ما يجعلها تناسب تقويم بعض الأهداف دون أخرى .

– الاختبارات الشفوية لا تسمح بالغش.

– توفر فرصة المواجهة بين الفاحص و المفحوص وجها لوجه .

– لا تحتاج الاختبارات الشفوية إلى وقت وجهد في إعدادها بعكس الاختبارات المكتوبة .

ورغم ما تحققه الاختبارات الشفوية من أهداف إلى أن جملة من الانتقادات توجه إليها.

من عيوبها :

– تتأثر الاختبارات الشفوية بحالة الأستاذ المزاجية وحالته النفسية عند إصدار الأحكام الذاتية حول إجابات التلاميذ على الأسئلة الموجهة إليهم . كما أن الحكم عن الإجابة يتأثر بعوامل ذاتية كالفكرة المسبقة التي يحملها الأستاذ عن التلميذ .

– عدم شمولية هذه الاختبارات ، فهي لا تغطي أكبر قدر ممكن من الموضوعات الدراسية التي يراد تحصيل التلاميذ فيها .

– التفاوت في مستوى السهولة و الصعوبة عند توجيه الأسئلة للتلاميذ حيث يلعب عامل الصدفة دورا هاما في هذا المجال . فقد يوجه إلى تلميذ سؤال سهل مقارنة بسؤال يوجه إلى تلميذ أخر.

– يتطلب إجراؤها وقتا طويلا مما لا يتيح المجال لإمكانية إخضاع الطلبة كافة للاختبار في الوقت المقرر .

– تأثر استجابة الطالب بموقف الامتحان وهو ما تلعبه العوامل التي تدخل ضمن إطار الحالة النفسية للطالب فقد يقدم الممتحن إجابة ناقصة أو خاطئة ليس لأنه يجهل المادة بل لأسباب أخرى كالخوف و الارتباك أمام الأستاذ في الامتحان نفسه .

– الأسئلة الممكن صياغتها محدودة خاصة وأنها تدور في نفس المواضيع، فقد تتكرر و يستفيد التلاميذ من إجابة الذين تقدموا قبلهم .

2 – الاختبارات العملية (الأدائية) : يهدف هذا النوع من الاختبارات إلى تقييم أداء أو مهارة عمل ، لتحقيق الأهداف التي تتعلق بالجانب الأدائي ذات الصفة العملية أكثر من الاهتمام بالجوانب الأخرى .و تنعكس فوائدها في ترجمة الجوانب النظرية إلى واقع عملي مدروس .

3– الاختبارات التحريرية : تتألف من الاختبارات المقالية و الاختبارات الموضوعية .

أ – الاختبارات المقالية: تعتبر من الاختبارات القديمة و تسمى بالتقليدية أو الإنشائية وتعتبر الأكثر شيوعا في محيطنا التربوي ،عادة تبدأ أسئلتها ببعض المصطلحات :ناقش ، تكلم عن ، وضح ، علل ، أعط رأيك و يكون المطلوب من الطالب إجابة تحريرية يعبر فيها بحرية عن أفكاره .و لهذه الطريقة إيجابيات و سلبيات .

الخصائص الايجابية للاختبارات المقالية :

– تعطي للطالب حرية التعبير عما لديه من أفكار و ألفاظ و ترتيبها بالشكل الذي براه مناسبا

– يفضلها الكثير من الأساتذة لسهولة إعدادها حيث لا تتطلب وقتا في الإعداد و لا جهدا .

– تسمح للطالب خاصة في مراحل التعليم المتقدمة بإظهار ما لديه من قدرة إبداعية و ابتكارية .

– احتمالية الغش فيها تقل مقارنة بالاختبارات الموضوعية .

الخصائص السلبية للاختبارات المقالية :

– ضعف شمولية الأسئلة فهي لا تغطي كل محاور المقرر الدراسي لكونها تقيس أجزاء محدودة منه . بمعنى أنه قد يأتي السؤال من جزء معين من المادة دون شمول الأجزاء الأخرى وهو ما يجعل احتمال استفادة بعض الطلبة من ذلك بالصدفة و تضرر بعضهم الأخر.

– يستغرق الطالب وقتا طويلا عند إجابته على هذا النوع من الأسئلة .

– يستغرق المعلم وقتا طويلا أثناء عملية التصحيح .

– كثيرا ما تكون غامضة ومبهمة حيث يصعب على الطالب فهمها أحيانا .

– ضعف ثبات و موضوعية التصحيح حيث يتأثر الأستاذ بعوامل عديدة أثناء التصحيح كالعوامل الذاتية و التي تخص حالته النفسية أو عوامل تخص ورقة الإجابة كالتنظيم و حسن الخط. الأمر الذي يبرز التباين في تقدير العلامات بين المصححين إذا أعيد تصحيح نفس الأوراق.

– يكون في كثير من الأحيان معامل صدقها ضعيفا لقلة عدد الأسئلة و عدم شموليتها على الموضوعات المقررة.

ب – الاختبارات الموضوعية: ظهرت هذه الاختبارات نتيجة الانتقادات وأوجه القصور المقدمة للاختبارات المقالية .

يعرف أبو علام (2005) الاختبارات الموضوعية على أنها الاتفاق في الأحكام ، فالسؤال الموضوعي هو السؤال الذي يمكن تقدير درجته تقديرا موضوعيا بحيث يمكن لعدة مقدرين أن يتوصلوا لنفس التقديرات بعيدا عن تدخل الأحكام الذاتية في بنائه وفي تحديد الإجابة الصحيحة (أبو علام ،2005 ،ص176).

يعرفها منسي محمود و زملاؤه (2007) بأنها تلك الاختبارات التي لا تتأثر بالحكم الشخصي لواضع الاختبار أو المصحح أو بغير ذلك من المؤثرات (منسي ، 2007 ،ص 373)

وعليه فالاختبار الموضوعي هو الاختبار الذي يحصل على قدر من الاتفاق بين المصححين حول طريقة تصحيحه و إعطاء الدرجات. ويمكن تحقيق الموضوعية عن طريق :

– وضع أسئلة بصورة دقيقة و خاصة بمعرفة معينة .

– تتطلب الأسئلة إجابات محددة .

– تحديد مفتاح التصحيح مسبقا لكي تتم عملية التصحيح بموجبه .

من إيجابياتها :

– شمولة الأسئلة للمنهج المقرر .

– معامل الثبات و الصدق و الموضوعية مرتفع إذا ما صيغت الأسئلة بصورة سليمة .

– تشكل الأسئلة عنصر تشويق و إثارة بالنسبة للطالب كما لا يستغرق وقتا طويلا في الإجابة عنها .

– سهولة تصحيحها

من سلبياتها :

– صعبة الإعداد و تتطلب وقتا وخبرة ودقة لتصميمها .

– تركز على قياس وحدات صغيرة من المعلومات غير المترابطة، بدلا من التركيز على المفاهيم الرئيسية و العلاقات البينية و الفهم و القدرة على التفسير. عمر وأخرون، 2010، ص (423

– تشجع الطالب على تخمين الإجابة الصحيحة.

– تحد من حرية الطالب في التعبير عن أفكاره و تنظيمها.

تتعدد أسئلة الاختبارات الموضوعية والتي يمكن أن نصنفها حسب نوعين:

النوع الأول: ينتج التلميذ فيه الإجابة في هيئة كلمة أو عبارة موجزة أو عدد أو رمز، فالإجابة المحتملة للسؤال تكتب كجزء من السؤال ومن أمثلة هذا النوع:

أ – أسئلة التكملة: يطلب فيه من التلميذ أن ينشئ الإجابة التي تكمل جملة حذفت منها كلمة أو عبارة .مثال الاختبار الثابت هو .....

<u>قواعد بنائها:</u>

– تجنب أخذ العبارات المقدمة في الدرس مباشرة لأنه يشجع التلاميذ على الحفظ الآلي.

– يجب أن يكون ملء الفراغ بكلمة أو جملة معينة و لا يحتمل الفراغ أكثر من إجابة.

– يستحسن ترك فراغ واحد في كل سؤال وهذا لتجنب الغموض الذي قد يحدث عند تعدد الفراغات.

– يستحسن جعل الفراغ قرب نهاية السؤال وتجنب وضعه في الأول وهذا يجعل التلميذ يقرأ المشكلة ليتوصل إلى الحل، وهو ما يجنبه الارتباك و الغموض خاصة إذا كانت الاختبارات موجه إلى تلاميذ المدرسة الابتدائية.

ب– أسئلة الإجابات الحرة: يطلب فيها من التلميذ تقديم إجابة حرة قصيرة عموما (بضع أسطر) حول السؤال المطروح. وفي هذا النوع من الأسئلة تكون عدة أجوبة محتملة. فيقل احتمال استخدام التخمين للوصول إلى الإجابة الصواب .

مثال: العلاقة التي تجمع بين التقويم و القياس.....

النوع الثاني: يقوم التلميذ باختيار الإجابة الصحيحة من بين المعلومات المقدمة له في السؤال وهو ما يعرف باختبارات الانتقاء أو الاستجابة المقيدة ومنها:

أ – أسئلة الصواب و الخطأ: يطلب من التلميذ أن يختار إجابة واحدة من إجابتين في كل سؤال (نعم /لا) أو (صواب /خطأ). حيث يتألف الاختبار من عبارات بعضها صحيح وبعضها خاطئ. يعاب على أسئلة الصواب والخطأ أنها تشجع التخمين و الغش.

مثال: الاختبار الجيد هو الاختبار الصعب جدا صواب خطأ

<u>قواعد بنائها:</u>

– تجنب العبارات المزدوجة التي تجمع بين الصواب و الخطأ.

– يفضل كتابة جميع المفردات أولا في صورة صحيحة ثم يتم تحويل البعض منها إلى مفردات خاطئة، ويكفل ذلك قدرا من التناسق في شكل وبنية المفردات كما يضمن أن يكون نصف الإجابات على الاختبار صواب و النصف الأخر خاطئ مما يقلل من أثر التخمين.
(عمر،2010 ،429)

– بجب أن تتسم العبارات بالدقة و الوضوح.

– الابتعاد عن النمطية في ترتيب الإجابات الصحيحة و غير الصحيحة.

– الابتعاد عن اقتباس العبارات حرفيا من المادة المعرفية المقدمة.

– تجنب ذكر بعض الكلمات التي توحي للتلميذ بالإجابة مثل نادرا، دائما، غالبا ...

ب – أسئلة الاختيار من متعدد: يتكون هذا النوع من الأسئلة من جزء ين رئيسيين الأول يسمى الأصل أو مقدمة السؤال وهي عبارة ناقصة خبرية تطرح على الطالب كمشكلة يتطلب منه حلها أو الإجابة عنها .أما الجزء الثاني يتشكل من قائمة تشمل إجابات تسمى البدائل وهي إجابات محتملة و يطلق عليها أيضا بالمشتتات وتتضمن إجابة صحيحة واحدة ضمن الإجابات الخاطئة

التي يفترض أن لا يقل عددها عن أربعة، وكلما ازداد عددها كانت الأفضل تجنبا للتخمين. (المياحي ،2010 ،ص118)

يطلب من التلميذ في سؤال الاختيار من متعدد أن يختار الجواب الصحيح من بين الإجابات الخاطئة الأخرى الموجودة في القائمة وذلك بوضع دائرة أو خط تحت الجواب الصحيح أو أي إشارة وهذا حسب التعليمات المطلوبة.

مثال: يرجع الفضل في تصميم مقياس الذكاء إلى العالم

1 – فونت

2 – كهلر

3 – بينيه

4 – بافلوف

يستخدم هذا النوع من الأسئلة لقياس أهداف معرفية مختلفة كالتذكر و الفهم و التطبيق. كما أنها تلفت نظر التلميذ إلى ضرورة التمييز و معرفة الحقائق معرفة دقيقة. كما تعتبر صادقة وثابتة بدرجة كبيرة مقارنة بباقي الأنواع إذا ما أعدت بصورة سليمة .

من عيوب أسئلة الاختيار المتعدد أنها صعبة الإعداد فقد تصادف المعلم صعوبة في إيجاد العدد المناسب من المشتتات. كما أنها تحتاج وقتا أطولا للإجابة خاصة إذا كانت الإجابة تتطلب تمييزا دقيقا و فهما للأساسيات .

<u>قواعد بنائها:</u>

- يجب أن يصاغ السؤال بشكل مختصر ودقيق وواضح.

- الابتعاد عن تكرار الكلمة الموجودة بأصل السؤال في البدائل.

- يستحسن أن تكون البدائل متجانسة و منطقية أي من صنف أو عائلة واحدة وترتبط بمجال المشكلة موضوع البحث.

– أن يراعى توزيع الإجابة الصحيحة بصورة عشوائية و الابتعاد عن النمطية عند وضع الجواب الصحيح للبدائل بحيث لا تؤدي إلى الإيحاء وتعرف التلميذ بالإجابة الصحيحة.

– التقليل من استخدام الصياغات السلبية للسؤال (المنفية) لأنها تقلل من وضوحه.

– توخي الحذر بحيث لا يوحي سؤال ما بإجابة سؤال أخر.

ج – أسئلة المطابقة أو المزاوجة: يتكون سؤال المزاوجة من قائمتين متوازيتين تضم كل منها مجموعة من الكلمات أو العبارات يطلق على القائمة الأولى بالمثيرات أو المقدمات والمجموعة الثانية بالاستجابات ويمكن تنظيم القائمتين في عمودين متقابلين، ويطلب من التلميذ أن يربط أو يختار كل عبارة في قائمة المقدمات مع العبارة التي تتفق معها في قائمة الاستجابات.

تفيد أسئلة المطابقة في الربط بين الأشياء متماثلة العلاقة كاختبار معاني المفردات وتواريخ الأحداث و النظريات إلى أصحابها و الرموز وما تدل عليها ..... كما يمكن أن نستبدل قائمة الاستجابات بالصور و الخرائط و الرسوم البيانية.

يفضلها الأساتذة لسهولة صياغتها و تصحيحها وتعتبر مناسبة لمختلف المواد العلمية والأدبية. يعابُ على أسئلة المطابقة أن تقيس في كثير من الأحيان تذكر التلميذ للحقائق و الأحداث و الظواهر فقي لا تقيس القدرة على الفهم و التطبيق و النقد. وتقتصر على جوانب محدودة تتمثل في العلاقة بين عنصر آخر.

<u>قواعد بنائها:</u>

– أن تكون عناصر قائمة العمودين متجانسة أي من نفس الصنف و العائلة.
– يستحسن أن تكون عناصر إحدى القائمتين أكثر عددا من القائمة الأخرى ويفضل أن تكون عناصر قائمة الاستجابات أقل وأقصر من عناصر قائمة المقدمات بشكل معقول حتى يبقى البعض منها بدون مزاوجة لتقليل أثر التخمين.
– يحدد بوضوح الأساس الذي قامت عليه المزاوجة بين المقدمات و الاستجابات لأن السؤال لا يهدف إلى اكتشاف فهم الطالب لأساس المزاوجة، لكنه يهدف إلى معرفة ما إذا كان الطالب قادرا على إجراء المزاوجة بدقة. (عمر وأخرون، 2010، ص446)

– أن تكون المقدمات في القائمة الأولى و الاستجابات في القائمة الثانية مع الحرص على ترتيبها وفق تسلسلها الرقمي أو الزمني أو المنطقي (المياحي،2011 ،ص 131) .

د – أسئلة الترتيب: يطلب في هذه الأسئلة من التلميذ إعادة ترتيب أحداث أو تسلسل زمني أو منطقي. مثال: رتب مستويات بلوم للأهداف التعليمية

تطبيق – تركيب – معرفة – تحليل – تقويم – فهم

يعاب عليها أنه لا تقيس مستويات التحليل و النقد وتركز على المستويات الدنيا (المعرفة و الفهم)

<u>قواعد بنائها:</u>

– يجب أن تكون مقدمة السؤال واضحة لأنها الأساس الذي يعتمد عليه في الترتيب.

– بجب أن تكون بدائل الترتيب منطقية لكي يتوصل التلميذ للاختيار الصحيح.

سابعا: ثبات الاختبار:

1. تعريف الثبات:

يقصد بثبات الاختبار أن تكون أدوات القياس على درجة عالية من الدقة و الإتقان و الاتساق فيما تزودنا به من بيانات عن السلوك المفحوص، بمعنى اخر أنه متى كانت أداة القياس قادرة على قياس المقدار الحقيقي للخاصية المراد قياسها في ظروف مختلفة و متباينة كان المقياس ثابتا.

كما يقصد بالثبات أن الاختبار يعطي نفس النتائج كلما أعيد تطبيقه على نفس المجموعة من الأفراد أي يقيس الاختبار نفس المراد قياسه كلما أعدنا عملية القياس.

يستخدم مفهوم الثبات بالمعنى العام للدلالة على مدى اعتماد الفروق في درجات الاختبار على الفروق الحقيقيّة في السمة أو الخاصيّة المقيسة من جهة و أخطاء القياس العشوائيّة و الناجمة عن الصدفة من جهة أخرى. و من المعلوم أنّ الاختبار في التربية و علم النفس يستحيل عدّه مقياسا ثابتا و دقيقا كالمقاييس الطبيعيّة لأسباب عديدة بينها ما يتّصل بعدم استقرار السمات أو الصّفات المقيسة بمرور الوقت، و بينها ما يتصل بخطأ العيّنة أو المعاينة، بالإضافة إلى أخطاء التطبيق و التصحيح و غيرها. و يمكن إجمال مصادر الخطأ في القياس النفسي و التّربوي على النحو التالي:

1. أخطاء في المقياس نفسه: فالمقياس هو عيّنة من البنود و قد لا تكون هذه العيّنة ممثلة لجميع البنود و هذا ما يسمّى بخطأ العيّنة أو المعاينة.

2. الأخطاء الناجمة عن شروط التطبيق: فإذا لم تراع شروط التطبيق كما تنصّ عليها تعليمات الاختبار أو روعيت بصورة غير كافية فقد تؤدّي إلى نتائج غير صحيحة.

3. الأخطاء الناجمة عن التصحيح: و هذه الأخطاء تظهر بخاصّة في الأسئلة المقاليّة التي تفسح مجالا واسعا لظهور العوامل الذاتيّة للمصحح و احتمال نشوء الاختلاف بين مصحح و آخر في تقدير الدّرجات، أو حتى عند المصحح نفسه من وقت لآخر.

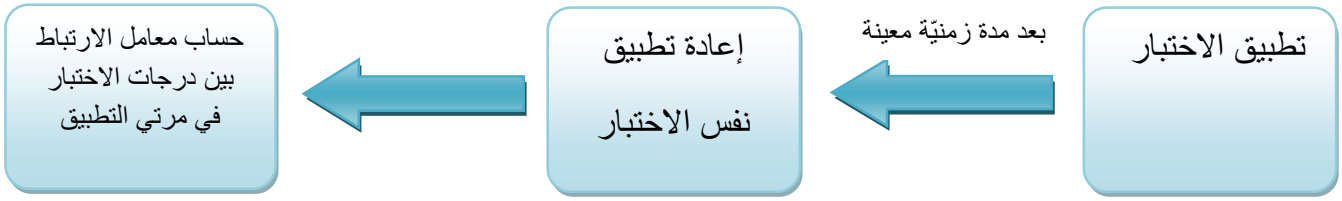
4. الأخطاء الناجمة عن المفحوص نفسه: وهذه الأخطاء كثيرة و متنوّعة منها ما يتصل بقلق الامتحان و منها ما يتصل بالتعليمات و الميل إل التخمين و الدافعيّة للأداء ودرجة التأثر بالعوامل الطبيعيّة كالحرارة و البرودة و الضجيج...و غيرها. هذا بالإضافة إلى عدم ثبات و استقرار السمة المقيسة عند المفحوص التي يمكن أن يتأثر بالعديد من العوامل كالتدريب و التذكر و النسيان و غيرها (ميخائيل، 2006، ص 182)

2. أساليب تقدير الثبات:

تتعدد أساليب حساب معامل الثبات ويختص كل أسلوب بمفهوم الثبات

المراد تقديره. لهذا على الباحث أن يحدد نوع معامل الثبات الذي يريد قياسه قبل استخدام الأسلوب الذي يقيسه ذلك النوع من الثبات.

أ- **ثبات الاستقرار** : ويمثل تطبيق الاختبار على مجموعة من المفحوصين وبعد فترة زمنية معقولة نعيد تطبيقه عليهم مرة أخرى تحت نفس الظروف التي طبق فيها الاختبار في المرة الأولى .يقدر معامل الثبات بحساب معامل الارتباط بين درجات أفراد المجموعة على الاختبار في التطبيقين .



مثال : يبين الجدول التالي درجات 8 أفراد طبق عليهم اختبار في المرة الأولى (X) وبعد مدة زمنية معينة أعيد تطبيق نفس الاختبار (Y) . احسب معامل الثبات.

N	X	Y	X^2	Y^2	XY
1	10	12	100	144	120
2	13	16	169	256	208
3	12	14	144	196	168
4	8	10	64	100	80
5	11	15	121	225	165
6	12	12	144	144	144
7	14	13	196	169	182
8	8	12	64	144	96
Σ	**88** Σ	**104**	**1002**	**1378**	**1163**

لحساب معامل الثبات في هذه الحالة نحسب معامل ارتباط بيرسون بين درجات التطبيقين

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

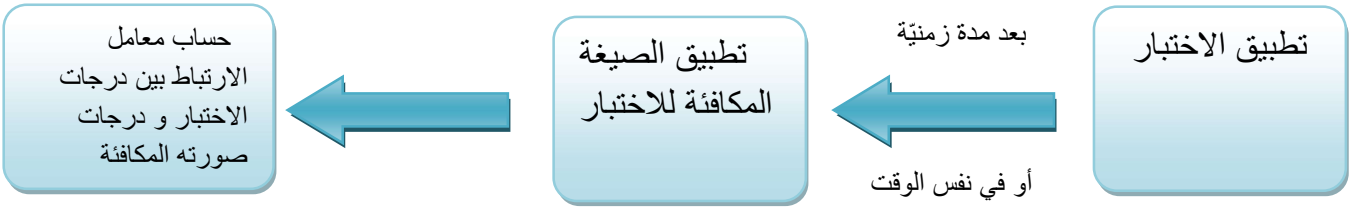
بالتعويض في المعادلة:

$$r = \frac{8\times 1163 - 88\times 104}{\sqrt{[8\times 100 \quad (88)^2][8\times 137 \quad (104)^2]}}$$

$$r = \frac{152}{\sqrt{272\times 208}} = \frac{152}{\sqrt{56576}}$$

r= 0.64

ب- ثبات التكافؤ: ويدل على مدى الارتباط بين درجات اختبار و اختبار أخر يكافئه، أي يقوم الباحث بتصميم صورتين من نفس الاختبار على أن يكونا متكافئتين ثم يطبقهما على أفراد العينة يمكن للباحث أن يقدم للمفحوص الاختبارين معا أو في وقت متقارب جدا.



يطبق معامل الارتباط بيرسون بين درجات الاختبار ودرجات الصورة المكافئة بنفس خطوات المثال السابق.

ج- ثبات الاتساق الداخلي:

هذا النوع من الثبات يجنب الباحث بعض الصعوبات التي تواجهه في إيجاد معامل الثبات بإعادة الاختبار أو بالصورة المكافئة.

يتجلى مفهوم الثبات في هذا النوع من الثبات في الاتساق الداخلي للاختبار ومدى تجانس فقرات الاختبار و مدى ترابط الاستجابات على الفقرة الواحدة مع درجة الاختبار الكلية.

- **التجزئة النصفية** : يعتمد هذا الأسلوب على تطبيق الاختبار مرة واحدة و تقسيم فقراته إلى قسمين متكافئين .

قد يعتمد الباحث **التقسيم النصفي** فيقسم اختباره إلى نصفيين متساويين يمثل النصف الأول

المجموعة الأولى من البنود فيما يمثل النصف الثاني المجموعة الثانية من البنود .لا يفضل استخدام هذه الطريقة في التقسيم لاعتبارات منها أن المفحوص قد لا يكون بنفس الدافعية و القدرة عند الإجابة على النصف الثاني .

كما يمكن أن يقسم الباحث اختباره إلى **مفردات فردية و أخرى زوجية** وهي أكثر الطرق استخداما في التجزئة النصفية.

من طرق حساب الثبات بأسلوب التجزئة النصفية:

- **معادلة سبرمان –براون spearman–bown:**

$$R_s = \frac{2r_{mm}}{1+r_{mm}}$$

R_s : معامل الثبات بطريقة سبرمان براون.

r_{mm} : معامل الارتباط لنصفي الاختبار.

مثال: اختبار مكون من 20 بند بعد تقسيمه إلى نصفين، حسب معامل الارتباط بين النصفين فوجد 0,40 . كم يكون معامل الثبات؟

بالتطبيق في معادلة سبرمان براون

$$R_s = \frac{2(0.40)}{1+0.40} = 0.62$$

- **معادلة جتمان:** تعتبر معادلة جتمان أكثر سهولة إذ يستخدم فيها التباين الخاص بكلا النصفين و كذا التباين الكلي للاختبار.

$$R_G=2(1-\frac{\sigma_1^2+\sigma_2^2}{\sigma_g^2})$$

R_G: معامل الثبات بطريقة جتمان.

σ_1^2 : تباين النصف الأول للاختبار.

σ_2^2 : تباين النصف الثاني للاختبار.

σ_g^2 : التباين الكلي لمجموع درجات نصفي الاختبار.

مثال : يبين الجدول التالي توزيعا لدرجات 10 طلا ب على اختبار يتألف من 20 بند ، بعد تقسيمه إلى درجات فردية و أخرى زوجية .أحسب معامل الثبات بطريقة جتمان .

N	درجة النصف الفردي **X**	درجة النصف الزوجي **Y**	الدرجة الكلية لاختبار **G**
1	10	10	20
2	8	9	17
3	5	4	9
4	4	5	9
5	4	7	11
6	9	8	17
7	7	4	11
8	4	4	8
9	8	8	16

10	3	3	6
التباين σ^2	18	35	102, 3

بالتعويض في معادلة جتمان :

$$R=2\ (1\ -\ \frac{18+}{102.3}\ )\ =\ 0.96$$

- **معادلة رولون Rulon**: تعتبر معادلة رولون طريقة مختصرة لحساب معامل الثبات الكلي للاختبار ولا تحتاج إلى حساب معامل الارتباط بين جزئي الاختبار كما في معادلة سبرمان براون.

$$R_R=1-\frac{\sigma_d^2}{\sigma_g^2}$$

R_R: معامل الثبات بطريقة رولون.

σ_d^2: تباين الفرق بين نصفي الاختبار.

σ_g^2 : تباين الدرجة الكلية للاختبار.

مثال1 : طبق اختبار رولون على مجموعة من الأفراد و قد وجد تباين الفرق بين نصفي الاختبار يساوي 2، أما تباين الدرجة الكليّة فيساوي 8.

$$R_R=1-\frac{2}{8}$$

$$R_R=0.75$$

مثال: طبق باحث اختبارا تحصيليّا على 10 تلاميذ و بعد تصحيحه قسّمه إلى مفردات فرديّة و أخرى زوجيّة. رصدت النتائج في الجدول الموالي:

N	بنود فردية x	بنود زوجية y	d	d^2	G	g^2
1	2	4	−2	4	6	36
2	2	1	1	1	3	9
3	2	2	0	0	4	16
4	3	2	1	1	5	25
5	2	3	−1	1	5	25
6	3	3	0	0	6	36
7	4	3	1	1	7	49
8	4	4	0	0	8	64
9	4	5	−1	1	9	81
10	5	4	1	1	9	81
$\sum$	31	31	0	10	62	422

– أحسب معامل الثبات بطريقة رولون.

بتطبيق قانون التباين الموالي:

$$\sigma^2 = \frac{1}{N^2}\left(N\sum x^2 - \left(\sum x\right)^2\right)$$

نجد:

$$\sigma_d^2 = 1$$

$$\sigma_g^2 = 3.76$$

بتطبيق معادلة رولون نجد:

$$R_R = 1 - \frac{1}{3.76}$$

$$R_R = 0.73$$

- **عدم تجزئة الاختبار**: إن طريقة التجزئة النصفية تصلح لحساب معامل الاتساق الداخلي (تجانس نصفي الاختبار) لكنها لا تقيس تجانس الأسئلة ذاتها، لهذا يسمى هذا النوع من الثبات بمعامل الاتساق داخل أسئلة أو بنود الاختبار. من أشهر الطرق لحساب معامل اتساق الأسئلة مايلي:
  - معادلة كيودر- رتشاردسون KUDER-RICHARDSON:

تطبق هذه المعادلة إذا صححت بنود الاختبار بشكل ثنائي (صحيحة 1، خاطئة 0)

$$R_{k_{21}} = \frac{n}{n-1}\left(1 - \frac{\bar{x}(n-\bar{x})}{n\sigma^2}\right)$$

$R_{k_{21}}$: معامل ثبات الاختبار بمعادلة كيودر

$\bar{x}$: المتوسط الحسابي لدرجات الاختبار.

n : عدد بنود الاختبار

σ^2 : تباين درجات الاختبار

مثال 1: طبق اختبار مكون من 10 بنود، حُسب متوسط درجات الاختبار فوجد 8 وتباين الدرجات5. ما هو معامل الثبات؟

بالتطبيق في معادلة كيودر

$$R_{k_{21}} = \frac{10}{10-1}\left(1 - \frac{8(10-8)}{10(5)}\right) = 0.75$$

مثال 2: يبين الجدول التالي استجابات 10 طلاب على اختبار يتألف من 10 بنود ، مع العلم أن هذا الاختبار صحح وفق الثنائية (0 ،1)

ما هو معامل ثبات هذا الاختبار؟

رقم المفحوص	البند 1	البند 2	البند 3	البند 4	البند 5	البند 6	البند 7	البند 8	البند 9	البند 10
1	1	1	1	1	1	1	1	0	1	1
2	1	1	1	1	1	1	1	0	1	1
3	1	1	1	1	1	1	1	0	1	0
4	1	0	0	1	1	1	1	0	1	1
5	1	1	1	0	1	1	0	1	0	0
6	0	0	0	1	1	1	0	1	0	1
7	0	1	1	1	0	0	0	1	1	0
8	1	1	1	1	0	0	0	0	0	0
9	0	0	0	0	1	0	0	1	1	1
10	1	0	0	1	1	0	0	0	0	0

نطبق معادلة كيودر:

	الدرجة الكلية G	مربع الدرجة الكلية G^2
1	9	81
2	9	81
3	8	64
4	7	49
5	6	36
6	5	25
7	5	25
8	4	16
9	4	16
10	3	9
المجموع	60	402

$$\sigma^2=\frac{1}{N^2}(N\sum x^2-(\sum x)^2)$$

$\sigma^2= 4.20$

$$\bar{x}=\frac{60}{10}=6$$

بالتعويض في معادلة كيودر نجد:

$$R_{k_{21}}=\frac{10}{10-}\left(1-\frac{6(10-\ )}{10(4.20)}\right)= 0.47$$

- **معادلة ألفا كرونباخ:** هي معادلة اقترحها كرونباخ عندما لا يتم تصحيح الفقرات بشكل ثنائي ويشيع استخدام هذه الطريقة في تقدير ثبات مقاييس الاتجاهات و استطلاع الرأي العام و في مقاييس الشخصية.

$$r_\alpha=\frac{n}{n-1}\left[1-\frac{\sum\sigma_i^2}{\sigma_G^2}\right]$$

n: عدد بنود الاختبار.

$\sum\sigma_i^2$: مجموع تباينات البنود.

σ_G^2 : تباين الدرجة الكلية للاختبار.

r_α : معامل الثبات بطريقة ألفا كرونباخ.

مثال 1: طبق باحث اختبارا على 10 أفراد، تضمن هذا الاختبار 8 عبارات على مقياس ثلاثي التدرج، قدر مجموع تباينات البنود ب 3,50 و تباين درجات الاختبار ككل 7. ما مقدار معامل الثبات ألفا؟

بالتطبيق في المعادلة:

$$r_\alpha = \frac{8}{8-1}\left[1 - \frac{3.5}{7}\right] = 0.57$$

مثال 2: يبين الجدول التالي نتائج 10 أفراد طبق عليهم مقياس للاتجاهات يتضمن 6 بنود على مقياس ثلاثي التدرج و صحح وفق (موافق 3، محايد 2، غير موافق 1). أحسب معامل الثبات؟

البنود / الأفراد	البند 1	البند 2	البند 3	البند 4	البند 5	البند 6
1	2	1	3	2	3	1
2	1	2	3	3	2	3
3	2	1	2	3	1	**2**
4	2	3	3	1	1	1
5	1	2	3	2	3	2
6	2	2	1	2	2	2
7	3	2	3	3	2	3
8	2	2	2	2	1	2
9	1	1	1	2	1	2
10	2	1	2	2	2	2

قبل تطبيق معادلة ألفا كرونباخ نقوم بإعداد الجدول التالي لحساب تباينات البنود والتباين الكلي للاختبار.

البنود / الأفراد	البند 1		البند2		البند 3		البند 4		البند5		البند6		المجموع	
	x_1	x_1^2	x_2	x_2^2	x_3	x_3^2	x_4	x_4^2	x_5	x_5^2	x_6	x_6^2	G	G^2
1	2	4	1	1	3	9	2	4	3	9	1	1	12	144
2	1	1	2	4	3	9	3	9	2	4	3	9	14	196
3	2	4	1	1	2	4	3	9	1	1	**2**	4	11	*121*
4	2	4	3	9	3	9	1	1	1	1	1	1	11	121
5	1	1	2	4	3	9	2	4	3	9	2	4	13	169
6	2	4	2	4	1	1	2	4	2	4	2	4	11	121
7	3	9	2	4	3	9	3	9	2	4	3	9	16	256
8	2	4	2	4	2	4	2	4	1	1	2	4	11	121
9	1	1	1	1	1	1	2	4	1	1	2	4	8	64

10	2	4	1	1	2	4	2	4	2	4	2	4	11	121
المجموع	18	36	17	33	23	59	22	52	18	38	20	44	118	1434
σ^2	0,36		0,41		0,61		0,36		0,56		0,40		4,16	

$\sum \sigma_i^2 = 2.7$

-يظهر الجدول تباين كل بند وتباين الدرجة الكلية للاختبار، تم حسابهم من خلال اعتماد معادلة التباين التالية:

$$\sigma^2 = \frac{1}{N^2}(N\sum x^2 - (\sum x)^2)$$

N : عدد الأفراد

x : *متغير(في هذه الحالة يمثل بنود الاختبار)*

بعدها نطبق معادلة ألفا كرونباخ وبالتعويض نجد :

$$r_\alpha = \frac{6}{6-1}\left[1 - \frac{2.7}{4.16}\right] = 0.43$$

3. العوامل المؤثرة في الثبات:

هناك عوامل عديدة يمكن أن تؤثر على الثبات، وبالتالي على الباحث أن يتحكم بها قدر الإمكان ويخفّض من تدخلها، و من هذه العوامل:

1.طول الاختبار (عدد الأسئلة في الاختبار):

يؤثر طول الاختبار في درجة ثباته، إذ كلما تزايد عدد الأسئلة كلما ارتفع ثبات الاختبار، لأن العدد الأكبر للأسئلة يؤدي إلى الحصول على عينة أكبر من السلوك و بالتالي يكون من المتوقع أن يمثل مكونات السلوك أو السمة أو القدرة المقاسة، كما يعكس الفروق الحقيقيّة في هذا السلوك. (بمعنى آخر أن الأسئلة التي زدناها و هذا بعد تحليلها تقيس جوانب نفسيّة جديدة). في

حين أن الأسئلة القليلة في الاختبار تقلل من احتمال استقرار مكونات السلوك أو الصفة المراد قياسها، و بالتالي انخفاض معامل الثبات.

لهذا إذا أراد باحث الوصول إلى معامل الثبات المرغوب لاختباره ينبغي عليه إضافة عدد آخر من الأسئلة إلى الاختبار، ويمكن استخدام معادلة سبرمان– براون في تحديد عدد الفقرات التي ينبغي إضافتها. حيث أن معادلة سبرمان– براون تساعد على التنبؤ بالزيادة المنتظرة في معامل الثبات إذا أضفنا إلى الاختبار عددا من البنود التي تقيس الخاصية موضوع الاختبار (الجلبي،2004، ص 121). وبعبارة أخرى تمكن هذه المعادلة من تحديد العلاقة بين الزيادة في طول الاختبار والزيادة في معامل الاختبار وهذه المعادلة كالتالي:

$$r_{\mathrm{ap}} = \frac{n r_{av}}{1+(n-1) r_{av}}$$

r_{ap} : الثبات بعد الزيادة

r_{av} : الثبات قبل الزيادة

n : عدد مرات زيادة الأسئلة.

مثال: حصل باحث على معامل ثبات 0,60 فأراد أن يرفع من عدد بنوده 4 مرات العدد السابق. فكم يحصل على الثبات بعد الزيادة ؟

بالتطبيق في المعادلة السابقة

$$r_{\mathrm{ap}} = \frac{4(0.60)}{1+(4-1)0.60}$$

$$r_{\mathrm{ap}} = 0.85$$

2. تباين درجات الاختبار:

يزداد معامل ثبات الاختبار إذا ازدادت درجة تباين مجموعة المفحوصين شريطة ضبط العوامل الأخرى المؤثرة في ثبات الاختبار بمعنى انه كلما كان أداء أفراد مجموعة المفحوصين متباينا أكثر كلما زاد مقدار معامل ثبات الاختبار.

متى تكون درجات الاختبار متقاربة؟ عندما يكون الاختبار غير قادر على التمييز و في هذه الحالة تكون المفردات سهلة جدا أو صعبة جدا، فالمفردات السهلة جدا أو الصعبة جدا لا تبرز الفروق الفرديّة و هي لا تظهر التباين، الأمر الذي يجعل معامل الثبات منخفضا.

3. موضوعيّة المصحح:

بعد ضبط العوامل المختلفة المؤثرة في معامل الثبات، فإن الاختبار الذي يُصحّح بموضوعيّة عالية يؤدي إلى زيادة تقدير معامل الثبات، فيجب ألا تختلف الدرجات من مصحح إلى آخر، ومن المعلوم أن الاختبارات الموضوعيّة عموما تحقق قدرا عاليا من الثبات بخلاف الاختبارات المقاليّة التي تفسح المجال لتدخل الذاتية للمصحح.

4. زمن الاختبار:

لكل اختبار زمن معيّن و مناسب يجب أن يقدّم للأفراد، بحيث لا يقل ذلك الزمن عن المحدد و لا يزيد عنه. فإذا ازداد الزمن عن المحدد عن حده ينخفض الثبات و كذلك إذا انخفض ينخفض الثبات . لأن زيادة الزمن عن المطلوب تفتح الباب أمام الممتحنين لكي يراجعوا الإجابات التي قدموها، كما يتعرض هؤلاء الممتحنون للقلق مما يؤدي بهم إلى إلغاء أو ترك أو تغيير الإجابات أو حتى الوقوع في التخمين.

و عندما يقل الوقت فإن الممتحنين يتركون الأسئلة خاصة الأخيرة لأن الوقت لا يكفيهم، فتتوتر أذهانهم و يجيبون عشوائيا. فالثبات يزداد في الزمن المحدد للاختبار فقط.

5. عوامل تؤثر بمقادير مختلفة في ثبات الاختبار:

كضبط موقف التطبيق و دافعيّة المفحوص و المؤثرات الفيزيقيّة في موقف الاختبار و الحالة الصحيّة و الانفعاليّة للمفحوصين، هذه الأخيرة تؤثر بصورة واضحة على الثبات. فالحالة الجيّدة للفرد تؤدي إلى ثبات مرتفع، لهذا على الباحث أن يقوم بضبط دقيق لكل هذه العوامل حتى يتمكّن من الوصول إلى معاملات ثبات مرتفعة.

ثامنا: صدق الاختبار

يعتبر الصدق من الخصائص المهمة التي يجب الاهتمام بها في بناء الاختبارات كما يعد أكثر المفاهيم الأساسيّة أهميّة في مجال القياس النفسي إذ لم يكن أهمها على الإطلاق.

1. مفهومه:

هنالك تعاريف عديدة للصدق منها:

– تعريف ليندكويست Lindquist: درجة الصحة التي يقيس بها الاختبار ما نريد قياسه. (مخائيل، 2006، ص141)

– تعريف ادجارتون Edgerton: الصدق يشير إلى المدى الذي تكون به أداة القياس مفيدة لهدف معيّن. (مخائيل ،2006، ص142)

– تعريف ساكس Sakse: هو الدرجة التي تكون فيها المقاييس ذات فائدة في اتخاذ القرارات المتعلقة بهدف أو غرض من الأغراض. (الطربري، 1997، ص218)

– تعريف آلن و وين Alin et Win: يعرّفان الصدق بأنه قدرة الاختبار على أن يقيس ما وضع لقياسه. (الطربري ،1997 ،ص219)

كل التعاريف السابقة تتفق في معنى الصدق وهو مدى نجاح الاختبار في قياس السمة المراد قياسها.

تعريف جيلكسون Gulikson: الصدق هو ارتباط الاختبار ببعض المحكات محددا بذلك إن الارتباط بمحك خارجي في شكل معامل هو مؤشر الصدق.

تعريف جلفورد: الصدق هو تحديد لمعامل الارتباط بين الاختبار و بعض محكات الأداء في مواقف الحياة (لا يختلف هذا التعريف عن تعريف جيلكسون و جزء من تعريفات الصدق).

_ تعريف كاتل Katel: الصدق في معناه الواسع هو قدرة الاختبار على التنبؤ ببعض الوظائف أو أشكال السلوك المحددة و المستقلّة عن الاختبار. (الجلبي ،2005، ص85).

إن تعاريف جليكسون وجلفورد وكاتل تنص على أنه ينبغي أن يوجد محك، ثم يرتبط الاختبار الجديد به، و هذا كي نحقق الصدق.

تبدو كل التعاريف التي قدمت للصدق متكاملة ومتشابهة فيما بينها فبعضها إجرائية وبعضها نظرية. و عليه فمفهوم الصدق يمكن أن نلخصه في النقاط التالية:

1 – أن يكون الاختبار قادرا على قياس ما وضع لقياسه : أي أن يكون الاختبار ممثلا تمثيلا للقدرة التي صمم لقياسها .

2 – أن يكون الاختبار قادرا على قياس ما وضع لقياسه فقط، أي أن الاختبار يقيس القدرة التي صمم لقياسها دون أن يشرك معها قياس قدرة أخرى غير مطلوب أداء قياسها.

3 – أن يكون الاختبار قادرا على التمييز بين طرفي الخاصية التي يقيسها، أي أن للاختبار القدرة التمييزية التي تسمح له بالتمييز بين الأداء المرتفع و الأداء المتوسط أو الأداء المنخفض للأفراد. (الجلبي، 2005، ص86)

عموما فالصدق هو الدرجة التي يمكن أن يقدم فيها الاختبار معلومات ذات صلة بالقرار الذي سيبنى عليها. أو مقدرته على قياس ما وضع من أجله أو السمة المراد قياسها، و إحصائيّا يعرف على أنه نسبة التباين الحقيقي المرتبط و المنسوب للسمة المقاسة.

2. خصائص الصدق:

1 – الصدق صفة نسبية لأداة القياس ليست مطلقة، لهذا لا ينصح بالحكم على الاختبار أنه صادق أو غير صادق فلا وجود لاختبار صادق بصورة كلية أو مطلقة، ولكن تحدد درجة صدقه كأن نقول صدق مرتفع أو صدق منخفض.

2 – الصدق يرتبط بالفئة أو العينة التي صمم لها: فإذا أُعِدَّ الاختبار لفئة معيّنة يجب أن يُطبّق عليها و ليس على غيرها لأن كل فئة تتسم بخصائص مختلفة عن الفئة الأخرى، فإذا صُمِّم الاختبار ليطبّق على الأسوياء فهو لا يصلح ليطبّق على فئة المعوّقين مثلا أو جماعة أخرى.

3– يرتبط الصدق مباشرة بالغرض أو الاستعمال الخاص الذي صمم الاختبار من أجله.
فالاختبار الذي أعد بغرض التنبؤ بأداء المفحوصين في فترة مقبلة قد لا يصلح لغرض التشخيص و الكشف عن نقاط الضعف و القوة في أداء المفحوصين، كما أن الاختبار الذي أعد لقياس الدافعيّة مثلا يجب أن يقيسها قياسا دقيقا شاملا و لا يقيس شيئا غير ذلك.

4– الصدق صفة تتعلّق بنتائج البحث، فلا يصلح أن نقول أن الاختبار صادق بل الأصح أن نقول أن نتائج الاختبار صادقة، فالصدق لا يرتبط بأداة القياس ذاتها بل بطريقة تفسير الدرجات المستخرجة من تلك الأداة.

فصدق أداة القياس يأخذ تنوّعا من المعاني بناء على الغرض الذي يبنى لأجله في كلّ مرّة، إذ عند الحديث عن صدق أداة القياس يتجه الاهتمام حتما نحو صلاحيّة تلك الأداة في تحقيق غرض معيّن و ليس نحو الأداة نفسها.

و بنفس السياق، يرى الباحثون في مجال القياس النفسي و التربوي أن الصدق هو مجموعة الأدلة التي نسترشد بها للتحقق من وجود الصدق و من درجته (النبهان، 2004، ص273).

5 – للصدق علاقة بالثبات: يكمن مفهوم الصدق في مدى فائدة أداة القياس في اتخاذ القرار المتعلق بالغرض و يتحقق هذا في اتساق المفردات و ثباتها، و منه فالمقياس الذي يعطي نتائج صادقة لا بد أن يعطي نتائج تتمتع بصفة الثبات، و عليه فإذا زاد معامل الصدق فإن معامل الثبات يزداد حتما.
و بالمقابل، قد يكون الاختبار ثابتا و لكنّه غير صادق و هذا لأنّ اتساق مفردات اختبار ما لا يعني أنها تقيس الغرض الذي أعدّ له الاختبار، كما أنّ تباين الدرجات المحصّل عليها من خلال الاختبار لا يعود إلى فروق حقيقيّة بين الأفراد في السّمة المقاسة بل يعود إلى فروق ناجمة عن الخطأ. لهذا، الاختبار الثابت قد لا يكون صادقا، فمعامل الثبات أشمل من معامل الصّدق.
و عليه فإن ما يعرف بالصدق الذاتي (الصدق الذاتي= الجذر التربيعي للثبات) لا يعتبر صدقا

بالمفهوم الحقيقي للصّدق، لأنه يتجاهل العلاقة التي تجمع الصدق بالثبات وهو أن الاختبار الصادق ثابت وليس كل اختبار ثابت صادق.

3. أنواع الصدق:

ارتبطت الطرق المختلفة للتحقق من صدق المقياس بأنواع عديدة ليست مختلفة بل هي متكاملة، و تعتبر أدلة على مدى صلاحيّة المقياس في المساعدة على اتخاذ القرارات و اتخاذ الأحكام، و تتمثل في:

أ) صدق المحتوى:

يسمى أيضا بصدق المضمون والصدق المنطقي وفيه يقتضي التأكد من تمثيل جميع المواقف التي تبدو فيها القدرة المراد قياسها .(العيسوي ، 2005 ،ص40) وهو الصدق الذي يتم عن طريق إجراء تحليلي منطقي لفقرات و بنود المقياس لتحديد مدى تمثيلها لموضوع القياس و المواقف التي نقيسها، و بعبارة أخرى أن يقوم الفاحص بفحص مضمون الاختبار فحصا دقيقا منظما و تحديد ما إذا كان يشتمل على عينة ممثلة لميدان السلوك الذي يقيسه.

لا يمكن استخدام صدق المضمون في اختبارات التحصيل و الكفاية و قد لا يصلح مع بعض الاختبارات النفسيّة و الاجتماعيّة.

من طرق حساب صدق المحتوى طريقة تقديرات الخبراء والمحكمين: و فيها يقوم الباحث بإعداد بنود اختباره و التي يجب أن تكون أكثر من عدد البنود التي سيَحويها اختباره في صورته النهائيّة، ثم يرفق هذه البنود بتعليمات تطبيق الاختبار و بتعريف إجرائي للخاصيّة التي هو بصدد دراستها، بعدها يقدّم نسخ ما أعد إلى خبراء و مختصّين في القياس أو في الخاصيّة المدروسة أو في المادّة الدراسيّة التي صمم الاختبار لقياسها إذا كان الاختبار تحصيليّا.

يطلب الباحث منهم تقدير مدى علاقة كل مفردة من مفردات الاختبار بالخاصيّة المراد قياسها.

يستحسن ألا يقل عدد الخبراء عن ثلاثين (30) خبيرا.

بعدما يقوم الباحث باسترجاع النسخ المسلّمة للخبراء، يقوم بتفريغ تكرارات تقديرات كلّ بند ثم يحوّلها إلى نسب مئويّة حيث يقسّم عدد المحكّمين الذين اختاروا أحد البنود على أنّه يقيس الخاصيّة المدروسة على العدد الكلّي للمحكّمين و ضرب الناتج في مئة، ثم يختار لاختباره البنود

التي حصلت على 80 % فأكثر و يرفض الباقي.

كما قد يلجأ الباحث لتطبيق معادلة لوشي

$$V_e = \frac{N_m - \frac{N}{2}}{\frac{N}{2}}$$

V_e : صدق المحكمين للمفردة.

N_m : عدد المحكمين الذين اعتبروا أن المفردة تقيس البعد المراد قياسه.

N : العدد الإجمالي للمحكمين.

ب) الصدق الظاهري (السطحي):

ويعني البحث عما يبدو أن المقياس يقيسه، وهو المظهر العام للاختبار أو الصورة الخارجيّة من حيث نوع المفردات و كيفية صياغتها و مدى وضوحها ، كما يجب أن تبدو مفردات الاختبار قوية الصلة بما يفترض أن الاختبار يقيسه .فالصدق الظاهري ليس صدقا حقيقيا بالمعنى التقني كما هو الحال في صدق المحتوى وصدق المحك وصدق البناء . ومع هذا يجب أن يكون الاختبار صادق ظاهريا ليرفع من دافعية المفحوص نحو الإجابة .

ج) الصدق المرتبط بمحك :(الصدق التجريبي)

يعتمد هذا النوع من الصدق على مقارنة نتائج الاختبار الذي يود الباحث التأكد من صدقه مع نتائج اختبار آخر طبق على نفس العينة و بنفس الظروف، حيث نسمي الاختبار الآخر بالمحك والذي يشترط فيه اتصافه بالموضوعية و الصدق و الثبات. و **هو نوعان :**

ج.1- الصدق التلازمي:

يتم فيه الحصول على بيانات المحك في نفس الوقت تقريبا الذي يتم فيه الحصول على درجات الاختبار، أي أنهما متلازمان، وكلما كان الارتباط قويا بين الاختبار والمحك قويا دل ذلك على معامل صدق تلازمي و العكس .

ج.2- الصدق التنبؤي:

يقصد به قدرة الاختبار على التنبؤ بنتيجة معينة في المستقبل كأن يطبق باحث اختبار على مجموعة من التلاميذ خلال العام الدراسي و الدرجات التي يحصل عليها أفراد العينة يحتفظ بها إلى أن يتم اختبارهم في وقت لاحق كمحك ثم يقارن بين درجات الاختبار ودرجات المحك فإذا كان هناك ارتباط قوي دل ذلك على أن اختباره صادق تنبؤيا .

- يحسب الصدق التلازمي و التنبؤي من خلال حساب معامل الارتباط بين درجات الاختبار ودرجات المحك .

د) صدق البناء (المفهوم): إن الاختبار المصمم لقياس بنية ما يجب أن يقدر مدى وجود هذه الصفة أو السمة الكامنة معتمدا مفردات الاختبار والتي تعد بمثابة مؤشرات على وجود هذه البنية .

يرى علماء القياس النفسي أن صدق البنية هو مفهوم موحد لكل أنواع الصدق ، هناك عدة طرق يمكن أن يلجأ إليها الباحث للوصول إلى هذا النوع من الصدق منها :

- إيجاد معاملات الارتباط مع اختبارات أخرى : ويتم ذلك عن طريق حساب معامل الارتباط بين الاختبار الذي نريد الوصول إلى صدق بنائه وبين الاختبارات التي تتشابه معه في الوظائف .
- إيجاد العلاقة بين درجات الفقرات والاختبار : تحسب العلاقة الارتباطية بين كل فقرة من فقرات الاختبار والدرجة الكلية للاختبار .
- طريقة المقارنة الطرفيّة: و فيها يقوم الباحث باستخراج نسبة 27% من أفراد العينة من كل طرف من طرفي التوزيع كمجموعتين متناقضتين (بعد أن يقوم بترتيب درجات الاختبار ترتيبا تصاعديّا أو تنازليّا) حيث تمثل المجموعة الأولى القسم العلوي للدرجات(درجات الأقوياء) بينما تمثل المجموعة الثانية القسم السفلي (درجات الضعفاء)، بعدها تتم المقارنة بينهما لإيجاد الدلالة الإحصائيّة للفرق بين المتوسطين باستخدام أسلوب إحصائي ملائم (النسبة الحرجة) فإذا كان الفرق دالا احصائيّا دلّ ذلك على أنّ الاختبار له القدرة على التمييز بين الأقوياء و الضعفاء و من ثمّ فهو صادق،

أمّا إذا كان الفرق غير دال إحصائيّا نقول أن الاختبار غير قادر على التمييز و من ثم فهو غير صادق.

كما تستخدم هذه الطريقة أيضا في حساب صدق المحتوى.

- **استخدام التحليل العاملي :** ويهدف إلى معرفة مدى قياس الاختبار للظاهرة التي صمم الاختبار لأجلها وتعتمد عل حساب معاملات الارتباط بين كل عبارة و العبارات الأخرى وتتجمع نتيجة المعاملات الارتباطية مصفوفة ارتباطات تنقسم التجمعات ، يجمع بين كل مجموعة عامل أو أكثر ونتيجة هذه العملية فإن الاختبار يختزل إلى عدد صغير من العوامل أو السمات المشتركة و التي يطلق عليها المكونات الأساسية للظاهرة التي يقيسها الاختبار .

إن التحليل العاملي يعتبر نوع من أنواع الصدق وهو أسلوب رياضي يمثل عدد كبير من المعالجات الرياضية في تحليل الارتباطات، و الجدير بالذكر إن التحليل العاملي كمنهج رياضي في التحليل يعد أسلوبا ليس من السهل تناوله وفهمه بل أسلوب معقد ساعد التقدم في برامج الحاسوب في سهولة استخدامه و نشره .

جدول يوضح أنواع الصدق

نوع الصدق	المعنى	طريقة استخراجه
صدق المحتوى	قياس مدى تمثيل الاختبار لنواحي الجانب المقيس لبحثه ، تحلل مواد الاختبار وعناصره تحليلا منطقيا لتحديد الوظائف و الجوانب الممثلة فيه ونسبة كل منهما إلى الاختبار بأكمله .	– طريقة استطلاع أراء المحكمين و الخبراء . – مقارنة محتوى الاختبار بمحتوى المنهاج و الأهداف .
صدق المفهوم	تحليل لمعنى درجات الاختبار في ضوء مفهوم نظرية ما .	– يقاس باستخدام طريقة الصدق العاملي . – طريقة المقارنة الطرفية

		– يقاس بحساب معامل الارتباط بيرسون بين درجات الاختبار واختبار اخر ثبت صدقه . – حساب معامل الارتبط بين الدرجة الكلية للاختبار و درجة كل مفردة من مفردات الاختبار . – حساب معامل الارتباط بين درجة كل فقرة من فقرات الاختبار و الدرجة الكلية للاختبار .
صدق المحك التلازمي	مدى ارتباط الدرجة على الاختبار بموازين الأداء الراهنة أو برتبة المفحوص أو ترتيبه أو مركزه الحالي (محك)	حساب معامل الارتباط بيرسون بين درجات الاختبار ودرجات المحك التلازمي
صدق المحك التنبؤي	مدى ارتباط الاختبار بمحك للأداء أو النجاح في وقت لاحق لإجراء الاختبار .	حساب معامل الارتباط بيرسون بين درجات الاختبار ودرجات المحك التنبؤي
الصدق الظاهري	مدى وضوح التعليمات ومدى مناسبة المفردات للسلوك المراد قياسه	استطلاع أراء الحكام و الخبراء

4. العوامل المؤثرة في الصدق:

ثمة عوامل عديدة يمكن أن تؤثر في صدق الاختبار و تضعفه ومنها:

- **عوامل متصلة بالمفحوص:**

– قد لا يبدي المفحوص دافعيته اتجاه الاختبار فلا يبذل جهد على النحو المطلوب، وخاصة الاختبارات التي تجرى بهدف البحث فهو يرى أنها لا تقدم له أي فائدة.

– يتعرض المفحوص لقلق أثناء أداء الاختبار خاصة الاختبارات التحصيلية مما تؤثر على إجاباته ومن ثم نتيجة الاختبار و صدقه.

- **عوامل متصلة بالاختبار نفسه:**

– **التعليمات غير واضحة:** فالتعليمات التي لا تشير بوضوح إلى ما يجب أن يفعله المفحوص، كيف يجيب وكيف يسجل إجاباته تضعف الصدق.

– **صعوبة الأسئلة أو سهولتها:** الأسئلة السهلة جدا أو الصعبة جدا لا تميز بين الأقوياء و الضعفاء تحصيليا فتُضعف الصدق. (ميخائيل، 2006، ص175)

– **البنود الموحية بالإجابة:** فالمفحوص يكتشف الإجابة الصحيحة اعتمادا على التلميحات والإيحاءات التي تتضمنها البنود وبالتالي لا تكشف الفروق في السمة أو الخاصية موضع القياس (مخائيل ،2006، ص176)

– **عدد البنود لا تغطي السمة المقاسة:** فالاختبار هو عينة من بنود أو أسئلة، فإذا كان قصيرا لا يمثل الخاصية بكل فروعها وجوانبها ومنه ضعف الصدق.

– **ترتيب البنود بصورة غير ملائمة:** يرتب الاختبار عادة وفق مبدأ التدرج في الصعوبة فإذا وضعت البنود الصعبة في البداية يقضي المفحوص الكثير من الوقت للإجابة عنها وقد لا يتسع له الوقت للإجابة عن البنود السهلة و الأخيرة، كما أن هذا الترتيب يضعف دافعيتهم مما ينعكس سلبا على أدائهم (مخائيل ،2006، ص176).

- **عوامل متصلة بإدارة الاختبار:**

- **الشروط البيئية المحيطة:** الحرارة و البرودة والضوضاء و غيرها كلها عوامل تؤثر في أداء المفحوص.

- **الوقت المخصص للإجابة:** فالوقت المخصص للاختبار يجب أن يكون مناسبا فإذا كان قصيرا لا يكفي المفحوص للإجابة يوقعه في الضغط و القلق، وإذا كان طويلا أكثر يؤدي أيضا بالمفحوص للارتباك و الشك في الإجابات وتغييرها وإتاحة الفرصة للغش والفوضى كما قد يقع المفحوص في التخمين.

- **تطبيق الاختبار على مجموعة لم يوضع لها الاختبار:** كأن يطبّق الاختبار على فئة المتفوقين و هو قد صمم للفئة المتوسطة أو العاديّة. (الجلبي، 2005، ص106)

تاسعا: دور القياس و التقويم في العملية التعليمية :

إن القياس و التقويم عمليتان متلازمتان و متكاملتان وتعدان من العناصر الأساسية للعملية التعليمية مثلهما مثل الأهداف و مدخلات التلاميذ و التخطيط للنشاط التعليمي وتنفيذه، حيث يعتبر التقويم التربوي من الفعاليات الأساسية في نشاط النظم التربوية عامة و المؤسسات التعليمية خاصة لضمان سيرها في الاتجاه السليم و المخطط له وضمان تطورها واستمرارها. وقد زاد الاهتمام بالتقويم التربوي لزيادة الحاجة لرفع الكفاءة الاقتصادية للنظام التعليمي وتطويره. فيعتمد نجاح العملية التعليمية على مدى نجاح عمليتا القياس و التقويم ومن أهم أدوارهما في العملية التعليمية ما يلي:

1 – **اتخاذ القرارات:** تعتبر الاختبارات أداة لقياس مستوى تحصيل المتعلمين، مما تمكن الإدارة المدرسية من اتخاذ القرار في تصنيف و توزيع التلاميذ و انتقالهم من مستوى دراسي إلى أخر كما تساعد الأستاذ في اتخاذ قرار بمدى استعداد تلاميذه للانتقال إلى موضوع جديد أو وحدة دراسية جديدة.

2 – صياغة أهداف تعليمية :

يرى ملحم سامي (2009) بأن الهدف التعليمي عبارة عن وصف لنمط من أنماط السلوك يُنتَظر حدوثه في شخصية المتعلم نتيجة لمروره بخبرة تعليمية أو موقف تعليمي معين، كما يُعرف أيضا على أنّه وصف التغيير المرغوب فيه في مستوى من مستويات خبرة أو سلوك المتعلم معرفيّا أو مهاريّا أو وجدانيا و هذا عند إنهاء خبرة تربوية معينة بنجاح بحيث يكون هذا التغيير قابلا للملاحظة و التقويم، فالهدف التعليمي هو الهدف السلوكي الذي يوضح التغيير المرغوب و المتوقع حدوثه في سلوك المتعلم بعد مروره بخبرة تعليمية معينة و يكون قابل للتقويم.

لقد شدد التربويون مند زمن بعيد على أهمية الأهداف التربوية ومدى تحقيقها. فمع تحديدها وتوجيه عملية التعلم والتعليم نحو تحقيقها، فإن النواتج تصبح قابلة للقياس، فمحاولة قياس أشياء غير محددة الأهداف هو بمثابة عمل ضائع بالنسبة لكل من المعلم والمتعلم. فوضوح الأهداف يتيح امكانية التحكم في عمل المتعلم وتقييمه بمعنى أنه كلما كانت الأهداف التعليمية محددة، كان من السهل قياس نواتج التعليم التي نلمسها من خلال الاختبارات.

وبناء على ما تقدم يمكن أن نقول أنّ تحديد الأهداف التعليمية ضرورة ملحة في عملية اختيار وسائل التقويم المناسبة، فالأهداف المحددة هي التي تحدد لنا ما نقوم به وفي الوقت نفسه ترشدنا إلى اختيار أنسب الوسائل والأدوات التي تصلح لتقويم ما نهدف إليه، كما أن الاختبارات تتأثر بالأهداف التعليمية، فكلما تحددت الأهداف على شكل نواتج تعليمية تتمثل في سلوك المتعلم تقوم الاختبارات بقياسها والحكم عليها بصورة دقيقة وصادقة. فتتحدد أهمية الاختبارات باعتبارها الوسيلة التي تمكن القائمين على التعليم من التعرف على مدى تحقق الأهداف التعليمية.

3 – زيادة الدافعية للتعلم:

حسب أحمد عمر وآخرون فإن التقويم يقوم بالوظائف الثلاث الرئيسية للدوافع في التعلم وهي :

أ– وظيفة التنشيط: أي زيادة المستوى العام للنشاط و الجهد المبذول. ويمكن التدليل على ما تحدثه الامتحانات السنوية العامة في ازدياد النشاط و الجهد العقلي في فترات العمل الشديد قبيل الامتحانات، وكشف كوك أنه عند مناقشة النتائج واستخراج الأخطاء الفردية يتحسن تحصيل

التلاميذ وأكدت أغلب الدراسات أن معرفة نتائج الامتحانات تفيد التعلم بمقدار النجاح الذي تتضمنه النتائج.

ب – وظيفة التوجيه: وفيها يتم توجيه سلوك المتعلم ونشاطه إلى المسالك المرغوب فيها فمثلا في الاختبارات الموضوعية التي تعتمد على الحقائق تتطلب من التلميذ تأكيد التفاصيل و الأسماء و التواريخ و نتائج التجارب .

ج – وظيفة الانتقاء: تعنى هذه الوظيفة بتحديد الاستجابات التي سوف بتم تثبيتها وبقاؤها عند التلاميذ ، و الاستجابات التي سوف تحذف أو تربط هذه الوظيفة بالجانب التشخيصي للتقويم . وتؤكد البحوث التجريبية أنه كلما كانت معرفة التلاميذ لجوانب الصواب و الخطأ في استجاباتهم وأسبابها معرفة مباشرة في موقف الاختبار ن فإن ذلك يؤدي إلى تثبيت الاستجابات الصحيحة أو السلوك المرغوب فيه وحذف الأخطاء واستبعادها. (أحمد عمر،2010 ،ص23)

قائمة المراجع:

1- أحمد حسين الصغير (2008)، **معايير تقويم أداء المعلم:** "نموذج مقترح دراسة ميدانية في مجتمع الإمارات " مجلة جامعة الشارقة للعلوم الإنسانية و الاجتماعية المجلد 5 العدد2.

2- البيلاوي حسن حسين وأخرون (2006)، **الجودة الشاملة في التعليم بين مؤشرات التمييز ومعايير الاعتماد،** عمان، دار الميسرة للنشر.

3- الجلبي سوسن شاكر (2005) ، **أساسيات بناء الاختبارات و المقاييس النفسية و التربوية** مؤسسة علاء الدين للطباعة و التوزيع ، الطبعة الأولى، دمشق ، سوريا .

4- الحريري رائدة عمر (2007)، **التقويم التربوي** ،الطبعة الأولى ،عمان الأردن .

5- الطربري عبد الرحمان بن سليمان (1997)،**القياس النفسي و التربوي** ،مكتبة الرشد للنشر ،الطبعة الأولى ،المملكة العربية السعودية .

6- العدوان زيد سليمان و أحمد عيسى داود (2016)، **استراتيجيات التدريس الحديثة،** مركز ديبونو لتعليم التفكير، دبي.

7- العيسوي عبد الرحمان (2005)، **فن القياس النفسي** ، دار الفكر العربي ، بيروت ، لبنان.

8- الفاخري سالم عبد الله سعيد، (2018) ، **التحصيل الدراسي** ، مركز الكتاب الأكاديمي، عمان .

9- المياحي جعفر عبد كاظم (2011)، **القياس النفسي و التقويم التربوي،** دار كنوز للمعرفة، الطبعة الأولى، عمان، الأردن.

10- النمر عصام و تيسير الكوفحي (2020)، **مناهج و أساليب التدريس في التربية و التربية الخاصة،** دار اليازوري العلميّة للنشر و التوزيع، عمان، الأردن.

11- عباس محمود عوض (2006)، **القياس النفسي بين النظرية و التطبيق**، دار المعرفة الجامعية، مصر.

12- علي عبد الحميد محمد ومنى إبراهيم قرشى (2009)، **الاتجاهات الحديثة في القياس النفسي والتقويم التربوي** ، مؤسسة طيبة للنشر و التوزيع ،الطبعة الأولى القاهرة ،مصر.

13- فرج المبروك عمر عامر (2016)، **التقويم و القياس التربوي الحديث بين الواقع و المأمول**. حميثرا للنشر والترجمة، مصر.

14- متروك نهى خليل (2015)، **أدوات القياس المفضلة و علاقتها بالتحصيل الدراسي لدى طلبة الصفوف الأولى من وجهة نظر المعلمين في محافظة معان**، مجلة كليّة التربية جامعة الأزهر العدد 164 الجزء الثالث، مصر.

15- محمود أحمد عمر وآخرون (2010)، **القياس النفسي و التربوي** ، الطبعة الأولى ، دار الميسرة عمان ، الأردن .

16- مديرية التكوين (2008)، **تربية وعلم النفس**، وزارة التربية الوطنية، الجزائر.

17- موسى النبهان (2004)، **أساسيات القياس في العلوم السلوكية**، دار الشروق للنشر، الطبعة الأولى، عمان، الأردن.

18- يونس محمد عبد السلام (2008)، **القياس النفسي**، دار الحامد للنشر، الطبعة الأولى، عمان، الأردن.

19- Dominique Galiana (2005), **Mémento de l'évaluation Analyser et améliorer sa pratique de l'évaluation guide méthodologique**, Educagri éditions, France.

20- DUKE, L. DANIEL (1995). **Teacher Evaluation policy: from accountability to professional development**. U.S.A. State University of New York press.