وزارة التعليم العالي والبحث العلمي
جامعة الجزائر 2- أبو القاسم سعد الله-

كلية العلوم الاجتماعيّة
قسم علوم التربية

مطبوعة بيداغوجية

القياس التربوي

وبناء الاختبارات المدرسية

السنة الثانية جذع مشترك علوم التربية

إعداد الأستاذة: حومل زينب

السنة الجامعية: 2021-2022

فهرس المحتويات

أهداف المقياس.....1
المعارف المسبقة.....2
الكفاءات المستهدفة.....2
المحور الأول: مدخل عام في القياس النفسي والتربوي
مقدمة عامة 2
1. ماذا نقصد بالقياس 4
2. القياس النفسي وعلاقته بمفاهيم أخرى.....6
3. تطور حركة القياس النفسي والتربوي.....10
3-1. في مجال علم النفس.....11.
3-2. في المجال التربوي..... 18
4. مجالات القياس في علم النفس وعلوم التربية.....20
5. الهدف من القياس النفسي.....22
6. الأسس العلمية للقياس النفسي والتربية.....23
7. خصائص القياس النفسي والتربوي.....25
المحور الثاني: مستويات القياس وبعض المعالجات الإحصائية المناسبة
أولا: مستويات القياس..... 30
تمهيد..... 30
1-1. المستوى الاسمي.....32
1-2. المستوى الرتبي.....32
1-3. المستوى الفتري أو المسافة.....33
1-4. المستوى النسبي.....35

2- الفرق بين مستويات القياس 36
ثانيا : بعض المعالجات الإحصائية المستخدمة في أنواع المقاييس
2-1. المستوى الاسمي..... 39
2-2. المستوى الرتبي 41
2-3. المستوى الفتري أو المسافة 43
2-4. المستوى النسبي..... 44
المحور الثالث: أدوات القياس..... 45
تمهيد 45
1 – أدوات القياس والتقدير النفسي 45
1-1.أدوات القياس 46
الاختبار النفسي..... 46
المقياس.....48
الاستبيان أو الاستخبار 48
1-2. أدوات التقدير 52
سلالم التقدير 52
قوائم التقدير 60
المحور الرابع: الاختبارات
أولا : الاختبارات وتصنيفها 67
مفهوم الاختبار 67
تصنيف الاختبارات 68
ثانيا: الاختبارات التحصيلية..... 72
تمهيد 72
مفهوم الاختبار التحصيلي 72

وظائف الاختبارات التحصيلية 73
الأسس التي يقوم عليها الاختبارات التحصيلية 76
أنواع الاختبارات التحصيلية 77
الاختبارات التقليدية 77
الاختبارات التحصيلية الموضوعية 81
خطوات بناء الاختبارات التحصيلية 86
المحور الخامس: دراسة الخصائص السيكو مترية للاختبارات
مدخل لشروط الاختبار الجيد 93
1-الثبات 94
مفهومه 94
أنواعه وطرق تقديره 95
العوامل المؤثرة في الثبات 115
2-الصدق 118
مقدمة 118
مفهومه 119
أنواعه 121
العوامل المؤثرة على الصدق.....125
خلاصة 128

أهداف مقياس القياس التربوي وبناء الاختبارات المدرسية:

- التعرف على الباحثتين الرواد في مجال القياس النفسي والتربوي وأهم اسهاماتهم.
- الإلمام بالمفاهيم الأساسية المتعلقة بالقياس النفسي والتربوي (القياس، التقييم، التقويم، التقدير)
- التعرف على أهم الخصائص القياس النفسي والتربوي.
- القدرة على التمييز بين القياس في مجال العلوم الاجتماعية وعلوم الطبيعة.
- التعرف على الأسس العلمية التي بني عليها القياس النفسي والتربوي.
- التعرف على تصنيف ستيفنس لمستويات القياس وفهم المبادئ الأساسية التي يرتكز عليها هذا التصنيف.
- القدرة على التمييز بين مستويات القياس المختلفة وقواعد استخدام الأرقام في كل مستوى.
- التعرف على بعض الأساليب الإحصائية المستخدمة في المستويات المختلفة للقياس.
- التعرف على أهم أدوات القياس في المجال النفسي والتربوي.
- التعرف على أهم التصنيفات للمقاييس والاختبارات.
- التعرف على أهم المواصفات الاختبار الجيد (الموضوعية، الدقة، الصدق، والثبات)
- معرفة مختلف الأساليب التي تمكنه من تقدير الصدق.
- معرفة مختلف الأساليب التي تمكنه من تقدير الثبات.

المعارف المسبقة:

- مفاهيم أولية عامة في الإحصاء.
- معارف أولية حول البحث في مجال العلوم الاجتماعية.

الكفاءات المستهدفة:

- التمييز بين المفاهيم المتعلقة بالقياس مثل القياس، التقييم، التقويم والتقدير.
- التمييز بين مختلف أدوات القياس مثل الاختبار، المقياس، الاستبيان، سلالم التقدير.
- أن يكون قادرا على تقدير الثبات باستخدام الطرق المختلفة لحساب الثبات.

-

المحور الأول: مدخل عام في القياس النفسي والتربوي

مقدمة عامة:

لا شك أن القياس يعتبر من أهم أدوات العلم الذي يقوم على قياس الظواهر الطبيعية وتقديرها تقديرا كميا دقيقا. فالتقدم الذي عرفته علوم الحياة والطبيعة والهندسة والفلك وغيرها من العلوم في فهم والتحكم والتنبؤ في مجالات نشاطها يرجع بشكل أساسي لدقة التقديرات والقياسات التي تمنحها أدوات القياس التي تستخدمها من وزن وكيل ومساحات،.....الخ.

وعلى نفس الخطى، عرف القياس في مجال العلوم الاجتماعية نقلة علمية هامة، بفضل جهود باحثين وأخصائيين في مجالات متعددة من طب وفيزيولوجيا ورياضيات واحصاء وعلم النفس، الذين عرضوا السلوك الإنساني للملاحظة والتجريب، وطبقوا قواعد القياس الطبيعي على السلوك الإنساني في أبعاده المختلفة (نفسية، معرفية، حسية حركية، سلوكية،...الخ).

فامتد وتوسع مجال القياس في العلوم الاجتماعية إلى قياس الجوانب النفسية، المعرفية والسمات الشخصية للفرد كميا، كما أصبح يهتم أيضا بمجال التعلم و التعليم، بالمتعلم، بالمعلم و المنهاج....الخ.

وسنحاول من خلال هذه المطبوعة البيداغوجية لمقياس القياس التربوي وبناء الاختبارات المدرسية والتي تستهدف طلبة السنة الثانية جذع مشترك علوم التربية، تقديم المفاهيم الأساسية المتعلقة بالقياس التي يجب أن يكتسبها الطالب في هذا المستوى حتى يتسنى له في المستقبل اكتساب مهارات في للقيام بالبحوث في المجال العلوم الاجتماعية.

ومن أجل ذلك تم تقسيم محتويات المطبوعة إلى خمسة محاور أساسية تتمثل في مدخل عام في القياس النفسي والتربوي، نتطرق فيه إلى التعريف بأهم المفاهيم في القياس (من قياس، وتقويم، تقييم وتقدير)، لمحة تاريخية حول تطور القياس النفسي والتربوي، خصائص القياس النفسي، الأهداف من القياس، الأسس العلمية، ثم محور ثاني نعرض فيه تصنيف ستيفنس لمستويات القياس، فمحور ثالث نتعرض فيه لأدوات القياس، فمحور رابع نتطرق فيه للاختبارات وتصنيفاتها، ويتم التركيز فيه على الاختبارات التحصيلية، ونتناول في المحور الخامس والأخير مواصفات الاختبار الجيد وبالأخص دراسة الخصائص السيكو مترية.

المحور الأول: مدخل للقياس النفسي والتربوي

1-ماذا نقصد بالقياس؟:

القياس هو تلك العملية التي تهدف إلى جمع معلومات كمية حول ظاهرة ما، وهو تعيين أرقام للدلالة على الأشياء أو الأحداث أو الأفراد وهذا وفق قواعد معينة، كما يمكن اعتباره رصد إجرائي لسلوك ما بإعطائه قيمة عددية باستخدام اختبار ما.

وكلمة القياس مأخوذة من كلمة قاس بمعنى قدر وقاس الشيء بغيره أو على غيره أي قدره على مثاله. (مصطفى قمش وآخرون، 2008، ص.15)

وتقابل كلمة القياس باللغة الفرنسية كلمة « La Mesure » ، أما باللغة الإنجليزية فتقابلها كلمتين وهما « Measure » و « Measeurment » وتعني الكلمة الأولى النتيجة الكمية المحصل عليها من القياس، أما الكلمة الثانية فتعني التقدير الكمي للمتغيرات أي تحديد أرقام للشيء أو الظاهرة موضوع القياس (غريب العربي، بدون سنة، ص.28)

وتعتبر من الكلمات الشائعة حيث تستخدم بمعان متعددة، ومن المعاني التي أحصيت لكلمة قياس نذكر تلك التي تطرق لها (English & English, 1958)، سواء بوصفها فعلا أو اسما، وهي كآلاتي:

- النتيجة التي نتحصل عليها في عملية القياس.
- الوحدة أو المعيار الذي يستخدم في القياس (من غرامات وأمتار وساعات وغير ذلك من الوحدات...الخ، وفق المقاييس المختلفة المستخدمة.

- تقدير إحصائي لخصائص الأشياء (بشرى إسماعيل، 2004؛ صفوت فرج، 2007)

رغم الاستخدامات العديدة لمصطلح القياس، غير أنه عرف شبه اجماع بين الباحثين فيما يخص تعريفه، حيث عرفه براون على أنه " عملية **التحديد الكمي للمشاهدات** التي تجرى **على السمات** التي تخص الأفراد أو الموضوعات أو الحوادث" (عبد الله زيد الكيلاني و أخرون، 2009، ص.15). واتفق معه ستيفنس(S.Stevens, 1951) بحيث عرفه بأنه " عملية **التعبير الكمي عن الخصائص والأحداث** بناء على **قواعد وقوانين محددة**"(موسى النبهان،2004، ص. 22).

وفي نفس السياق، اعتبركامبل(Campell, 1952) القياس "عملية **تحويل الأحداث الوصفية** إلى أرقام بناء على **قواعد و قوانين** معينة"(بشير معمرية، 2012، ص.32؛ بشرى إسماعيل، 2004، ص.26). أما كيلنجر(Kerlinger)فيرى أنه " **نسب الأرقام إلى الأشياء أو الأحداث** وفق **قواعد معين**"(غريب العربي، بدون سنة، ص.28)،

وأكد ننللي(J.C.Nunnaly,1970) على التعاريف السابقة من خلال اعتبار القياس" **قواعد استعمال الأعداد**، لتدل على **الأشياء** بطريقة تشير إلى **كميات من خاصية**"، وهذا ليؤكد على القواعد التي تحكم استخدام الأرقام لوصف الظواهر النفسية.

وعليه أجمعت التعاريف السابقة على أن القياس هو استخدام الأرقام للدلالة على الأشياء أو الأحداث أو السمات المختلفة للأفراد وفق قواعد معينة، وعليه يتكون تعريف القياس من ثلاث مكونات أساسية تتمثل في: الخاصية المراد قياسها والرقم وهو الرمز الذي يصف كميا الخاصية، والقاعدة وهي العلاقة التي تربط الرقم بالخاصية.

وعلى خلاف ما سبق، اعتبر بين(K. Bean) القياس: " **مجموعة من المثيرات المرتبة** تقيس بطريقة **كمية بعض العمليات** الانفعالية أو العقلية أو النزوعية. وقد تكون المثيرات أسئلة مكتوبة أو شفهية أو في صورة سلسلة من الأعداد أو الأشكال أو النغمات "، وبذلك اقترن تعريف بين للقياس بالاختبار. أما كرونباخ(J. L. Cronbach,1970) فعرف القياس بأنه: " **الطريقة المنظمة لمقارنة سلوك** شخصين أو أكثر(بشير معمرية، 2012، ص.32)، أي الإجراءات العلمية المتبعة التي تسمح بالمقارنةبين سلوك الأشخاص، وبذلك فإن المقابلة لا تعتبر عملية قياس.

وبناء على ما سبق، يمكن القول بأن القياس مفهوم واسع لا يقتصر فقط على الاختبار، وهو الطريقة المنظمة التي نستخدم فيها الأرقام للدلالة على الأشياء أو الأحداث أو السمات والخصائص المرتبطة بالأشخاص للمقارنة بينها، حيث يستمد معناه العلمي من المقولة الشهيرة لثورندايك"**أن كل شيء يوجد بمقدار، وما وجد بمقدار يمكن قياسه.**" (بوسنة محمود،2007، ص.53؛ عبد الرحمن بن سليمان الطريري، 1997، ص.8)

2-**القياس النفسي و علاقته بمفاهيم أخرى :**

ارتبط مفهوم القياس (Measurement) في مجال علم النفس والتربية بمفهوم التقييم (Assessment)، التقويم ((Evaluation والتقدير (Rating) حيث الكثيرون ممن يتداولون مختلف هذه المصطلحات، يخلطون بينها ويستعملونها للدلالة على نفس المعنى، غير أن استخدامها في مجال علم النفس والتربية ميز بين هذه المفاهيم، و فيما يلي سنحاول بشكل موجز عرض أوجه الاختلاف بينها:

- **القياس(Measurement)** : هو الوصف الكمي والكيفي الموضوعي لأداء الفرد، والذي يبنى عليه اتخاذ قرار في مجال ما، فالدرجة التي يتحصل عليها الفرد عند قياس سمة ما ليس لها معنى في ذاتها بل تجد معناها عندما تعطى لها تقدير أو قيمة.
- التقدير (Jugement): يقابلها المصطلح الإنجليزي (Rating) وهو مفهوم أقدم من القياس من حيث الاستخدام، وهو مرادف للتخمين، أي تحديد شيء بالانطباع الشخصي أو بالحدس، ومن الناحية اللغوية نقول قدر شيء أي بين مقداره.

وفي الحياة اليومية يقوم الفرد بتقديرات متعددة بناء على تخمينات شخصية وكيفية مثل: أقل من، أكثر من، ضخم، هزيل، أكثر دقة،....الخ، والجدير بالذكر أن هذه التقديرات تبقى كيفية لا ترتقي إلى القياس الكمي.

- **التقييم (Assessement)**: في الكثير من الأحيان يستخدم كمرادف للتقويم، يعني لغة عملية إضافة قيمة للشيء، كما جاء في المعجم الوسيط بمعنى بيان قيمة الشيء، فالتقييم اصدار حكم عن قيمة الشيء (دون تعديل) وهو إعطاء قيمة لشيء وفقا لمستويات محددة. وقد ظهر المصطلح أول مرة في كتاب نشر لمكتب الخدمات الاستراتيجية بالولايات المتحدة تحت عنوان "تقييم الإنسان" في سنة 1948، الذي استخدم في الحرب العالمية الثانية من طرف علماء النفس الذين كانوا يشتغلون في المكتب والذين كلفوا بانتقاء الأفراد للقيام بمهام خاصة فيما وراء البحار (صلاح الدين محمود علام، 2000).

كما شاع استخدامه مؤخرا في مجال قياس الشخصية وفي وصف أداء الأفراد، وقد عرفه ساندبرج(Sundberg,1977) بانه مجموعة من العمليات التي تستخدم بواسطة أخصائيين للوصول إلى تصورات وانطباعات لاتخاذ قرار نحو سلوك معين وقد يأخذ هذا التقدير صيغة كمية أو كيفية لوصف الظاهرة موضع التقدير (فؤاد أبو حطب، سيد عثمان وأما صادق، 2008، ص. 18).

كما عرفه نيتسو (2000) بأنه" عملية جمع ووصف وتكميم البيانات عن مستوى أداء معين بقصد استخدامها في اتخاذ قرارات معينة". (موسى النبهان، 2004، ص ص.38-39)

وفي نفس السياق اعتبر هستوفل بيم وآخرون (Stuffelbeam & al,1971)" عملية الحصول على المعلومات المناسبة من أجل اتخاذ قرار من بين عدة خيارات. (عبد الرحمن بن سليمان الطريري، 1997، ص.5).

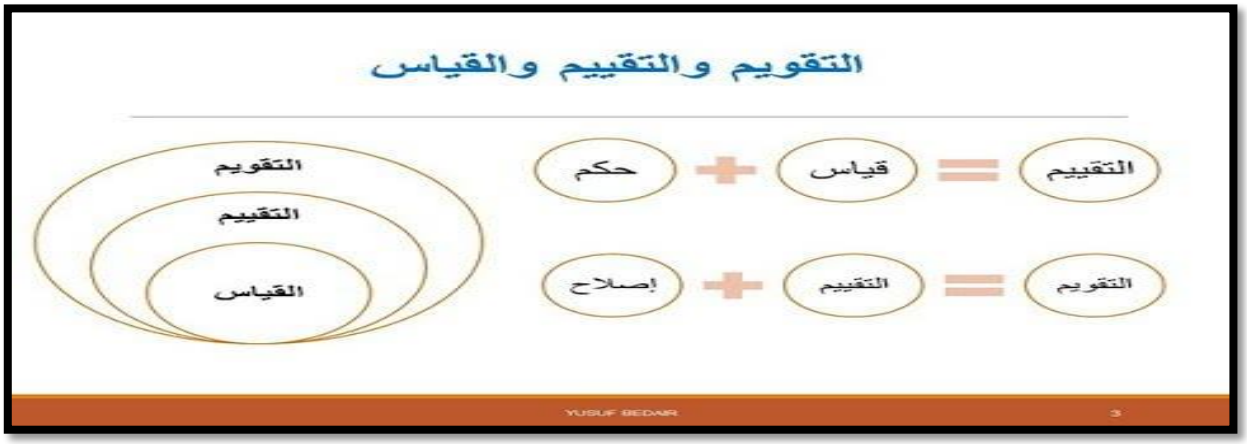
وبذلك فالتقييم عبارة عن اصدار حكم عن السلوك بالاعتماد على أدوات قياس متعددة (اختبارات الذكاء، التحصيل، الميول،.....الخ).

- **التقويم (Evaluation)**: يستخدم عادة في المجال التربوي، يقصد به لغويا بأنه بيان قيمة الشيء، ويعني كذلك تعديل أو تصحيح ما أعوج، فإذا قال شخص: إنه قوم المتاع فمعنى ذلك أنه ثمنه، وجعل له قيمة معلومة، وإذا قيل قوم الغصن فمعناه عدله وصححه أي جعله مستقيما (سبع محمد أبو لبدة، 1981، ص.61).

وبناء على هذا المعنى اللغوي لمفهوم التقويم يمكن تعريفه بانه: "تعديل مسار أي عملية تربوية، وتوجيهها الوجهة الصحيحة، وإصلاح نقاط القصور فيها.

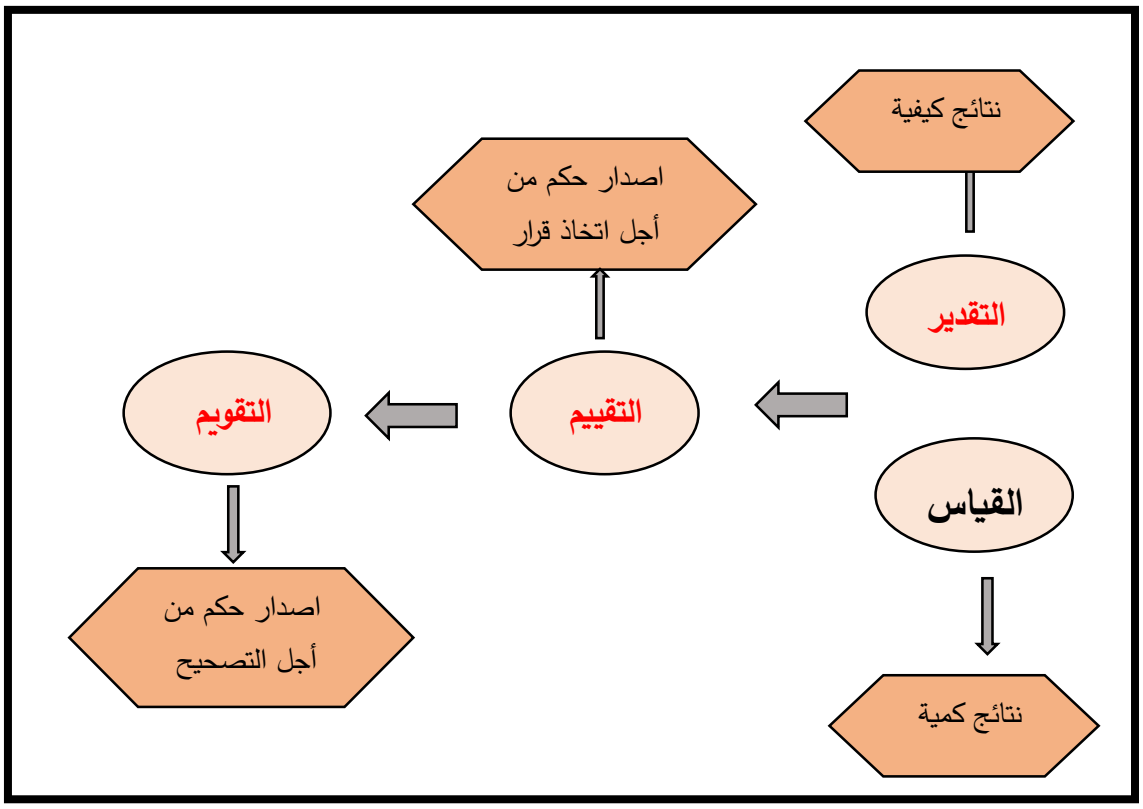
وعلى هذا الأساس فالعلاقة التي تجمع بين عملية القياس، التقييم والتقويم هي علاقة تكاملية حيث يعد القياس عملية ضرورية للتقييم والتقييم خطوة سابقة للتقويم، وفي هذا السياق يرى صبري ماهر إسماعيل أن:" كل عملية تقويم تنطوي على عمليتي التقييم والقياس، وأن كل عملية تقييم لا تعتمد بالضرورة على عملية قياس ولكن عمليتي القياس والتقييم لا تغنيان عن عملية التقويم"

والمخطط الموالي رقم (01) يوضح العلاقة بين التقويم والتقييم والقياس:



مخطط رقم (01): يمثل العلاقة بين القياس والتقييم والتقويم

ومزيدا من التوضيح في المخطط رقم (02) الذي يبين الفرق بين القياس والتقدير والتقييم والتقويم:



المخطط رقم 02: يمثل العلاقة بين القياس والتقييم والتقويم والتقدير

3- تطور حركة القياس:

تمهيد:

لقد كان الانسان منذ القديم يحاول أن يعرف مدى قدراته وامكانياته، فكان يقارنها بمن حوله من الكائنات فإن وجدها أقوى منه استسلم لها وقدم لها القرابين وربطها بطقوس معينة وإن تغلب عليها جعلها في خدمته.

مر القياس بمراحل وتطورات مختلفة، فقد بدأ باستخدام الفراسة، ثم اعتمد القياس على النواحي الحسية الحركية ثم تطور الى الاهتمام بالعقل والقدرات العقلية العليا ثم توسع في الأخير إلى مختلف الجوانب النفسية.

وسنحاول فيما يلي التطرق إلى حركة القياس من خلال عرض نشأتها وأهم المحطات والاسهامات في علم النفس أولا ثم ثانيا نعرض أهم مراحل نشأة القياس في مجال علوم التربية :

3-1. في مجال علم النفس:

المرحلة الأولى: الأحكام المبنية على الفراسة: ويقصد بها الاستدلال بالأمور الظاهرة على الأمور الباطنية والخفية، فكان يحكم على شخصية الفرد وقدراته العقلية من خلال مظهره الخارجي وملامح وجهه فتقارن بملامح الحيوانات التي تشبهه. كما كانت الفراسة تعتمد على شكل الجمجمة، فكانت تدرس السمات العقلية أو الملكات تبعا لحجم المخ وحجم الأمكنة المرتبطة بالملكات العقلية. (عبد الرحمن محمد عيسوي، 1999، ص ص.8-9)

غير أن الدراسات أثبتت خطأ أساليب الفراسة المختلفة في قياس الذكاء والقدرات العقلية، من هذه الأبحاث نذكر تلك التي قام بها كل من جالتون (Galton) وبيرسون(Berson) على فراسة الوجه (صلاح أحمد مراد & أمين علي سليمان،2005؛ أمين علي محمد سليمان، 2009).

المرحلة الثانية: بداية التجريب و القياس (قياس النواحي الحسية والحركية): بدأت هذه المرحلة في أواخر القرن الثامن عشر وبداية القرن التاسع عشر، وتميزت هذه المرحلة بقياس النواحي الحسية والحركية للمستويات الدنيا للنشاط العقلي لدلالة على الذكاء، مثل سرعة النبض قبل أداء وبعد أداء الاختبار، قوة الرفع باليد، تقدير الطول بالنظر، وقوة الصدر وسعته،....الخ.

غير أن الدراسات التي قام بها العديد من الباحثين أمثال كاتل وجيلبرت وكرابلين أثبتت اخفاق الاختبارات الحسية الحركية في قياسها الدقيق للذكاء، لاعتمادها على المستويات الدنيا للنشاط العقلي (أمين علي محمد سليمان، 2009)

وفيما يلي سنحاول التطرق لأهم الباحثين الذين ميزوا هذه الفترة:

من الغرابة عدم ظهور حركة القياس النفسي على أيدي علماء النفس، بل بدأت على أيدي علماء الفلك، ففي عام 1796 فصل ماسكلين(Maskelyne) مدير مرصد جرينيتش بإنجلترا مساعده كينيبروك(Kinnebrook) بعدما أخطا في رصد بعض الكواكب بزمن قدره ثانية واحدة، حيث اعتبر ذلك نوع من عدم الأمانة العلمية.

في عام 1816 لفتت هذه الحادثة انتباه العالم الألماني فريدريك فلهام بيسل(F.W Bessel) الذي كان يعمل فلكيا، فاهتم بجمع البيانات والمعلومات عن الأخطاء في تقدير الفلكيين، وخلص إلى وجود فروق طبيعية بين الأفراد في زمن الرجع، وصاغ على ضوء هذه النتائج عام 1820 ما يعرف باسم المعادلة الشخصية، والتي فحواها أن الأفراد يختلفون فيما بينهم من حيث زمن الرجع، أي الفترة الزمنية المحصورة بين ظهور المثير وحدوث الاستجابة (صلاح أحمد مراد & أمين علي سليمان،2005؛ بشير معمرية، 2012).

وقد أدى هذا الحادث إلى اهتمام الباحثين في النصف الأول من القرن التاسع عشر بقياس الفروق الفردية، إذ كان الظن السائد أنها أخطاء، لا بدا من التخلص منها.

تبنى هذا الاتجاه العالم الألماني فونت(Wundt) حيث درس الفروق الفردية وحاول الوصول إلى صياغة أوصاف معممة للسلوك أو قانون عام يصف السلوك الانساني.

فأنشأ أول معمل لعلم النفس سنة 1879، واهتم بالمثيرات الحسية مثل البصرية والسمعية ودراسة زمن الرجع. وكان لهذا الاتجاه أثر كبير في الاختبارات النفسية الأولى.

غير أنه هناك من يرى أن بدايات حركة القياس جاءت من ألمانيا ولكنب فضل أبحاث كل من ارنست هاينريخ فيبر (E.H Weber, 1878) في السيكوفيزيقا، والمعادلة التي تم صياغتها من طرف كل من ارنست فيبر& جوستاف تيودور فيخنر (G.T Fechner) والتي اعتبرت بمثابة تحول بارز في التعامل مع الظواهر النفسية من خلال الملاحظة المقننة والتحديد الموضوعي، بالإضافة إلى أعمال يوهانس مولر (J. Muller,1851) التي أثارت الاهتمام بدراسة الحواس وردود الأفعال (صفوت فرج، 2007؛ بشرى إسماعيل، 2004).

جاء بعد فونت تلميذه فرانسيس جالتون (F. Galton) العالم البيولوجي الانجليزي الشهير، فاهتم بدراسة الوراثة، واهتم بخصائص الأفراد الأقرباء وغير الأقرباء لتحديد درجة التشابه بين التوائم الاخوة والأخوات.

وفي سنة 1882 أنشأ معملا لعلم الانسان القياسي، وفيه أمكنه قياس حدة البصر والسمع، القوة العضلية، زمن الرجع ووظائف حسية حركية ومن ثم جمع أول وأضخم مجموعة من البيانات حول الفروق الفردية في العمليات النفسية البسيطة.

صمم بنفسه وفي معمله اختبارات بسيطة، ويطبقها البعض حتى الآن مثل قضيب جالتون لتمييز البصري للأطوال وصفارة جالتون لتحديد أعلى مقام سمعي. (بشير معمرية، 2012، ص. 53) كما كان له الفضل في تطبيق مناهج الاستبيان والمقياس المتدرج وكذا استخدام التداعي الحر، كما ساهم في تطوير الطرق الاحصائية لتحليل البيانات عن الخصائص الانسانية وهذا بعدما استعان بكارل بيرسون (معامل الارتباط، الوسيط، طرق الترتيب...الخ). (سعد عبد الرحمن، 1998، ص. 22)

والى جانب الاهتمام بالحواس ظهرت محاولات لتدريب وتصنيف ضعاف العقول بعد أن كانوا يعاملون بقسوة وهذا نتيجة انتشار مؤسسات خاصة بضعاف العقول في أمريكا وأوروبا. والبداية كانت عن طريق الطبيب الفرنسي ايتارد (Itard) في أوائل القرن التاسع عشر عندما قام بتدريب الطفل المتوحش المعروف بأفيرون، وبعد ذلك ظهرت اسهامات الطبيب الفرنسي اسكيرول (Esquirol, 1837)من خلال كتابه المعنون الأمراض، حيث قام بوصف ضعاف العقول باستخدام أنواع من المقاييس الفيزيقية، كما ميز بين المرض العقلي والتأخر لعقلي.

وتوالت بعد ذلك حركة تدريب ضعاف العقول باستخدام الأسلوب الفيسيولوجي والتي قادها الطبيب الفرنسي الآخر سيجين (Seguin). (عبد الله الكيلاني وآخرون، 2009، ص.27).

المرحلة الثالثة: القياس العقلي: ظهرت هذه الحركة منذ القرن التاسع عشر، تميزت بقياس العمليات العقلية العليا، وما سمح بذلك هي الحاجة الماسة في فترة ما بين الحربين العالميتين لانتقاء وتوجيه الجنود والقادة إلى أنواع مختلفة من الأسلحة.

كان لجيمس ما كين كاتل(J.M Cattell) السيكولوجي الأمريكي جهود كبيرة في هذا المجال حيث انفرد عن باقي تلامذة فونت، باهتمامه بالفروق الفردية، ويعتبر أول من تحصل على شهادة الدكتوراه تحت إشراف فونت في جامعة ليبزغ في ألمانيا، حيث أنشأ معمل الاختبارات في جامعة بنسلفانيا عام 1888بعدما عاد إلى أمريكا، اتفق مع جالت ونعلى أنه يمكن قياس الوظائف العقلية من خلال اختبارات التمييز الحسي وزمن الرجع، فصمم العديد من الاختبارات الحسية الحركية والنفسية، كما اعتبر أول من طبق بطارية للاختبارات تتضمن اختبارات التداعي الحر والتداعي المقيد، وبعض العمليات الادراكية البسيطة والذاكرة وزمن الرجع على الطلاب المستجدين في كولومبيا بدءا من عام 1894، بالإضافة إلى أنه أول من استعمل مصطلح الاختبار العقلي سنة (1890) (صفوت فرج، 2007؛ عباس محمود عوض، 1998)

أما كرابلين فقد اهتم بالفحص الاكلينيكي لمرضى العقول، فاعتمد من اجل التشخيص على سلسلة من الاختبارات والعمليات الحسابية لقياس أثر التدريب والتذكر والقابلية للتعب والتشتت (صلاح أحمد مراد وآخرون، 2002، ص.19).

كما يعود الفضل كذلك في تطور القياس العقلي الى ألفرد بينيه (A. Binet) و تيودور سيمون (T. Simon)، ففي سنة 1904 كلف بينيه و سيمون من قبل وزارة المعارف الفرنسية للبحث في الطرق التي يمكن استخدامها لتعليم الأطفال المتخلفين عقليا والملتحقين بالمدارس النظامية فصمما مقياسهما الأول في سنة 1905 مكونا من 30 بندا مرتبة تصاعديا حسب الصعوبة. وكان يحدد

مستوى الصعوبة موضوعيا بتطبيق البنود على 50 طفلا سويا تتراوح أعمارهم ما بين 3 و11 سنة وعلى بعض ضعاف العقول وكانت البنود المصممة تغطي وظائف متعددة

مع التأكيد على الحكم والفهم والاستدلال والتي يعتبرها بينيه مكونات أساسية للذكاء وكان الاختبار في الغالب لفظيا.

عرف الاختبار رواجا بعدما راجعه لويس ترمان(L.Terman) سنة 1916 وعرف بمقياس ستانفورد وبينيه، وقام ترمان باستبدال مفهوم العمر العقلي بمفهوم نسبة الذكاء تبنيا للتعريف الذي اقترحه سترن (Stern, 1912)للعمر العقلي والذي يتمثل في النسبة بين العمر العقلي والزمني غير أنه أضاف له ضربه في 100.

وبعد ذلك تم استبدال نسبة الذكاء بمفهوم نسبة الذكاء الانحراف الذي اقترحه ويكسلر (Wechsler) في مقياسه الذي أعده لقياس الذكاء والذي قام بتسميته مقياس الذكاء ويكسلر –بلفيو نسبة للمستشفى الأمراض النفسية بلفيو الذي كان يعمل فيه في نيويورك (توما جورج الخوري، 2008؛ إبراهيم محمد محاسنة، 2013).

وفي بداية الحرب العالمية الأولى،ظهرتالحاجة إلى تظافر جهود علماء النفس لمساعدة أمريكا بعد دخولها الحرب سنة 1917، في انتقاء الجنود وتوجيهم وفق مؤهلاتهم وشخصيتهم وتكليفهم بمهام خاصة، واسندت مهمة الاشراف على الفريق والتنسيق للعالم النفسي يركس(Yerkes) ولهذا الغرض تم تصميم اختبارين عسكريين لقياس القدرات العقلية، يتم تطبيقه بشكل جماعي، يتمثل الاختبار الأول في اختبار آلفا(Alpha) وهو لفظي، يتكون من ثمانية أجزاء للذين يجيدون القراءة والكتابة،

أما الثاني فهو اختبار بيتا (Beta) والذي يعتبر اختبار أدائي من سبع أجزاء، وهو على خلاف اختبار ألفا يصلح لمن لا يجيدون القراءة والكتابة.(بشير معمرية، 2012؛ صفوت فرج،2007).

المرحلة الرابعة: تنوع الاختبارات لقياس النواحي النفسية المختلفة: رغما ظهور محاولات لقياس سمات الشخصية في أواخر القرن التاسع عشر قام بها كل من جالتون وبيرسون في إنجلترا، وروزانوف في أمريكا، يونج في سويسرا وكرابلن في ألمانيا، غير أن الانطلاقة الحقيقية لحركة قياس الشخصية كانت مع دخول أمريكا الحرب العالمية الأولى حيث ظهرت الحاجة إلى اسهامات الفنية لسيكولوجيين من اجل الكشف عن السمات الشخصية المرغوبة لتعزيزها، لغرض انتقاء الضباط الأكفاء وتشخيص حالات الإصابة بالاضطرابات النفسية.

لعل أهم السيكولوجيين الذين أسدوا مساعدة للجيش الأمريكي، نذكر روبرت وودورث (R.Woodworth,1917) الذي صمم نموذجا للاستبيان الشخصي أو بيان تقرير الحال، الذي يقدم على شكل مقابلة شخصية، بحيث يجيب المفحوص بنعم أو لا ، فيحكم على شخصية الفرد من حيث قدرته على تحمل الضغط العصبي و الأوامر و من ثم يمكن تحديد ملامحه الشخصية. كما ظهر في عام 1929 استبيانا للشخصية، أعده ثرستون وزوجته يحتوي على 223 بند، شاع استعماله في الحرب العالمية الثانية. وفي عام 1933 أعد بورنرويتر(Burnreuter) قائمة للشخصية، بينما صمم بال (H.m.Bell, 1934) قائمته لتوافق، وفي نفس السنة نشر ويلوبي قائمة مختصرة لاستبيان ثرستون.

كما انتشر عدد كبير من اختبارات الشخصية التي تعتمد على الأسلوب الاضفائي بفضل مفهوم الاسقاط عند فرويد، كاختبار روشاخ لهرمن (H.Roschach,1911)، اختبار تفهم الموضوع لهنري موراي و مورغان(H.A.Murray, K . Morgan1935)، واختبار رسم الرجل لفلورنس جودانف (Goode-nough1926) و اختبار رسم الشجرة لشارلز كوش و الاختبارات التي تعتمد على التقرير الذاتي مثل اختبار مينوسيتا (1939) و استبيان سترونج(E.K. Strong) للميول المهنية للرجال (1919-1920)(بشير معمرية،2011، ص ص. 63-70)

3-2. في المجال التربوي:

اما فيما يخص القياس في المجال التربوي يمكن ارجاع أصوله إلى اثنين وعشرين قرن قبل الميلاد إلى عملية الانتقاء التي كان يقوم بها أباطرة الصين لاختيار الأشخاص الأكفاء لتولي مناصب في الدولة. (صفوت فرج،2002)

أدخلت الاختبارات الشفهية إلى جامعة أكسفورد سنة 1636، أما الامتحانات التحريرية فقد أدرجت في جامعة كامبريج في سنة 1800 ضمن أساليبها في التقويم (أحمد مراد وعلي سليمان، 2002). أما في الولايات المتحدة الأمريكية فقد بدأت حملة الاختبارات في عام 1845 في المدارس الابتدائية والثانوية وأول اختبار فيها طبق في مدينة بوسطن في تلك السنة، كما ظهرت حركة نقدية للامتحانات وخاصة تلك المتعلقة بالامتحانات الشفهية، تلى ذلك ظهور إجراءات جديدة للقياس والتقويم، أهمها ضرورة استعمال الاختبارات الكتابية بالإضافة إلى الاختبارات الشفوية ولعل أول محاولة كانت جادة ظهرت على يد جورج فيشر الذي ضمن في كتابه "كتاب الميزان" مقياسا للكتابة لتبويب عدة أنواع

من الكتابات عند التلاميذ والذي يعتبر كذلك أول عالم زود عملية التعلم بالاختبارات المقننة في عدد من المواد(محاسنة، 2013).

تأسست لجنة في سنة 1894 أوكلت لها مهمة التحقق حول المردود التربوي بمدينة شيكاغو الأمريكية.

تميزت هذه الفترة بظهور فكرة التربية التجريبية حيث تم استخدام المفتشين الخارجيين في تقويم مدى التحسن في مستويات المدارس (غريب عربي، بدون سنة).

انتشرت ما بين 1900–1930 الاختبارات التحصيلية وبطاريات الاختبارات المقننة، ويعد روبرت ثورندايك(Thorndike) أحد أهم قيادات حركة القياس والتقويم في هذه الفترة فقد نشر اختبارا للخط عام 1910 وجعل للاختبارات فائدة عملية كبيرة من حيث أنها تعتبر عاملا أساسيا في عملية اتخاذ القرارات في مجال التربية، مثل تحديد مستويات النجاح والرسوب، كما قام بتدريس أول مقرر في القياس التربوي في جامعة كولومبيا عام 1902.

شهدت كذلك ظهور كتابات راش Rusch عام 1929 عن الاختبارات الموضوعية.

ظهرت في هذه الفترة كتابات حول صدق وثبات الاختبارات ومنها تلك التي تناولها كل من ثرستون1937 Thurstone، كيودرريتشارسون 1937 Kuder Richardson وجوتمان1945 Guttman.

عرفت سنة 1947 تأسيس جمعية خدمات الاختبار التربوي بالولايات المتحدة الأمريكية، وفي سنة 1948 انعقد لقاء لدراسة مشاكل التقويم برئاسة بلوم (مؤتمر الجمعية الأمريكية للسيكولوجيا) وقرر أن القاعدة التي يرتكز عليها التقويم هي الأهداف التربوية.

ظهرت سنة 1973اختبارات التفوق والاتقان، اختبارات القدرة والكفاية، اختبارات المهارات الأساسية والاختبارات المدنية والاختبارات التربوية المرجع وهي التسمية التي اقترحها كارفر سنة 1974(بن سي مسعود، 2008)

4- **مجالات القياس في علم النفس وعلوم التربية:** ما الذي نقيسه في مجال علم النفس وعلوم التربية ؟

- قياس العمليات المعرفية: يقوم المتخصص في علم النفس بقياس كمي دقيق لذكاء ولمختلف القدرات العقلية، كما يقوم بتمييز بينها. فهو يقيس الذكاء ويميز خصائصه، فيميز بين الذكاء اللفظي والأدائي والاجتماعي، ويميز بين الذاكرة الحسية والذاكرة القصيرة والطويلة المدى. كما يقيس المختص في الذكاء القدرة على الاستدلال، التجريد، التعميم والتمييز...الخ
- قياس سمات الشخصية: يقوم السيكولوجي بتقدير كمي للسمات المزاجية الشخصية للفرد، نذكر على سبيل المثال التعاون، السيطرة، التوتر، الانطواء والعصابية الاجتماعية.....الخ
- قياس المهارات الحركية: يقيس المختص في هذا المجال الدقة والسرعة في أداء الحركات الصادرة عن العضلات، مثلا نقيس مهارات الكتابة، تشغيل الآلات، العزف على الآلات الموسيقية، الجري...الخ.

- قياس الاضطرابات العقلية: يقيس المختص مظاهر سلوك الفرد الغير السوي كالقلق، الاكتئاب، الوسواس القهري والفصام...الخ .
- قياس الدوافع والاتجاهات والقيم والميول: يهتم القياس النفسي بالدوافع، فيقيس الدافع إلى العمل، الدافع للإنجاز والدافع للدراسة وغيرها. ويهتم أيضا بقياس الاتجاهات نحو القضايا المختلفة كالاتجاه نحو العمل، الدراسة، الانجاب...اخ. ويهتم بقياس مختلف ميول الأفراد مثل تلك المرتبطة بالمهنة والدراسة ...الخ، كما يولي الاهتمام بقيم التي يتمتع بها الفرد مثل القيم الدينية، الخلقية، الاقتصادية....الخ.
- قياس التفكير الابداعي وحل المشكلات: يتولى السيكولوجي أيضا قياس القدرات الابداعية لدى الأفراد ومدى وجود الفروق الفردية فيها وهذا بالتعبير على وجودها كميا، كالطلاقة، المرونة، الأصالة والحساسية للمشكلات وغيرها. كما يهتم بقياس الاستراتيجيات المختلفة المتبعة من طرف الأفراد لحل المشكلات. (بشير معمرية، 2012، ص. 38– 39).

5- **الهدف من القياس النفسي: لماذا** نقوم بعملية القياس في علم النفس؟

من أجل اثراء الموضوع، ارتأينا أن نبتدئ بمقولات لعلماء رواد في مجال علم النفس:

جيمس ماكين كاتل(J.M.Kattell) :

" لن يتمكن علم النفس من الوصول إلى يقين العلوم الطبيعية ودقتها. إلا إذا استند على أساس من التجربة والقياس".

كلفين (Calvin) :

" ان معرفة الانسان بالعلم، لتبدأ حين يمكنه أن يقيس ما يتحدث به، وأن يعبر عنه بالأرقام"

ادوين .ج.بورنج(Edwin.G.Boring) (بشير معمرية، 2012، ص. 38- 39).

" اننا لا نكاد ندرك موضوعا ما كموضوع علمي، إذا لم يكن القياس واحدا من أدواته"

مما لا شك فيه، أن المختص النفسي باعتماده إلى القياس فإنه يطمح في الوصول إلى الدقة العلمية التي تتصف بها العلوم الطبيعية، انه هدف ضمني ومسلم به.

فالباحث في علم النفس، عندما يجري عملية القياس فإنه يهدف كذلك إلى:

- **الوصف (المسح):** تصنيف الخصائص السلوكية و التعرف على جوانبها و المتغيرات التي تتداخل معها من أجل الوصول إلى قوانين تحكمها تمهيداً لتوظيف هذه الإمكانيات على الوجه الأكمل. (بشير معمرية، 2012، ص. 40)
- **التنبؤ** : نحن نقيس ونقيم الفرد و الجماعة في وظيفة معينة في وقت معين من أجل معرفة المستوى المتوقع الذي يمكن أن يتوصل إليه في نفس الوظيفة التي تم قياسها. وبناء على

فهم قدرات الأفراد ومقتضيات العمل ومتطلباته يمكن توجيه الأفراد للدراسة أو المهنة التي يكون احتمال نجاحهم فيها مرتفعاً. وبناء على الاختبارات المزاجية الشخصية يمكن التنبؤ بكيفية تصرف الأفراد في المواقف المختلفة.

- **التشخيص:** بناء على درجات التلاميذ في الاختبارات يمكن معرفة جوانب القوة والضعف لدى كل منهم.
- **العلاج:** بناء على التشخيص يقوم المختص بمعالجة وتعديل وتصحيح السلوك (سامي محمد ملحم،2017، ص.40).

6- **الأسس العلمية للقياس النفسي والتربوي:**

- يقوم القياس النفسي على ما نادى به " ثورنديك " في قوله " إذا وجد شيء فإنه يوجد بمقدار، وإذا كان موجوداً بمقدار فيمكن لنا قياسه ".
- القياس النفسي قياساً لعينة من السلوك، فعند تصميم اختبار تحصيلي في مادة الرياضيات مثلا، فإن الباحث المصمم للاختبار سيضع أسئلة حول بعض مما درسه التلميذ وليس جميع محتويات المادة، ولذلك سيختار عينة ممثلة لما حصله التلميذ في مادة الرياضيات ويختبره فيها في فترة زمنية محددة، هذه العينة من السلوك يجب أن تظهر في شكل أدائيات يمكن قياسها.
- يتوجب أن يكون الأداء بدرجات متفاوتة وبمقادير مختلفة لدى الأفراد أو بمعنى آخر وجود فروق فردية في الأداء. ويقصد بالفروق تلك: " الانحرافات الفردية عن المتوسط الجماعي

في الخصائص المعرفية والوجدانية والنزوعية". وتعتبر الفروق الفردية من أهم الأسس الذي يقوم عليها القياس، باعتباره يهدف إلى الكشف عن الفروق في نوع الخاصية أو درجتها، والتي تكون إما في الفرد ذاته أو بين الأفراد أو الجماعات أو المهن (بشير معمرية،2012، ص. 73).

- تعتمد الفكرة الأساسية للقياس على مقارنة ما نريد قياسه بمعيار دقيق نصطلح عليه، حيث القياس لا يكون له معنى في ذاته، فمثلا عندما يتحصل تلميذ ما على درجة 08، هذه الدرجة لا معنى لها إلا إذا ما قورنت بمعيار مثلا الدرجة القصوى التي يمكن تحصيلها مثلا 10. وبناء على ما سبق تختلف تفسير الدرجة وفق المعيار، فالدرجة 08 من 10 يختلف تفسيرها عن نفس الدرجة إذا ما كان معيار المقارنة 20 بدل من 10، كما يختلف تفسيرها إذا ما تبين وجود 80 % من التلاميذ من تحصلوا على درجة تفوق 08 ، وعليه القياس لا معنى له دون معيار.
- تعتمد عملية المقارنة على تحديد الصفة أو الخاصية (X) – القدرة – السمة لتحديد نوع المقياس الصالح لقياسها. فالقياس ليس اجراء اعتباطي، بل يقوم على تحديد السمة مراد قياسها في شكل أداء يمكن قياسه اجرائيا، ثم اختيار أنسب الأدوات التي تسمح بذلك.

مثال 01: السلوك العدواني الذي قد يكون لفظيا أو ماديا، يتم تحديده من طرف الباحث وفق تناوله النظري في البحث وأهداف بحثه، ثم يحدد مختلف السلوكيات التي تندرج ضمن السلوك العدواني، كالضرب بالعصا، الصفع، والركل، شد الشعر، الشتم والسب، ...الخ.

مثال 02: التحصيل في وحدة الإحصاء في السنة الأولى، فالأستاذ المكلف بالوحدة عند قياسه لما تم تحصيله في الوحدة، لا بدا إحصاء جميع محتويات ما قدمه في الوحدة ووفق الأهداف التي سطرها في بداية السنة والكفاءات المتوقع اكتسابها، فإنه يختار عينة من المفردات (أسئلة)

مثل: حول مقاييس النزعة المركزية والتشتت وحالات استخدامها، التمييز بين البيانات الكمية والكيفية وتصنيفها، مقاييس العلاقة المختلفة واستخداماتها تبعا لطبيعة البيانات، تفسير نتائج اختبارات الفروق،.....الخ.

7- خصائص القياس النفسي والتربوي:

يتميز القياس النفسي والتربوي بمجموعة من الخصائص العامة أهمها:

- القياس النفسي والتربوي عبارة عن تقدير كمي لبعد من أبعاد السلوك الانساني، عملا بمقولة ثرونديك بأن كل ما يوجد، يوجد بمقدار، وما يوجد بمقدار يمكن قياسه. (صالح أحمد مراد وآخرون، 2002)
- القياس النفسي والتربوي قياس غير مباشر فنحن لا نستطيع قياس الذكاء أو التحصيل أو أي صفة نفسية أخرى بطريقة مباشرة، مثلما نقيس طول الأفراد أو وزنهم، ولكننا نقيس ما يدل على وجود الصفة، مثلا: نقيس حدوث التعلم من خلال الزمن المستغرق لإنجاز عمل. ويشبه القياس النفسي في ذلك، قياس بعض الظواهر الطبيعية، مثل قياس الحرارة، فنحن لا

نقيس الحرارة إلا عن طريق أثرها على عمود من الزئبق أي نقيسها بطريقة غير مباشرة. (صالح أحمد مراد وآخرون، 2002)

- القياس في العلوم الانسانية قياس نسبى وليس مطلقاً، بمعنى أنه لا يكون للدرجة التي يتحصل عليها الفرد في الاختبار النفسي معنى في ذاتها. بل لابد من مقارنتها بمعيار يكسبها معنى تفهم وتفسر في إطاره (متوسط أداء الجماعة التي ينتمي إليها). والمعيار أساس للحكم مستمد من الخاصية ذاتها. (محمد سليمان وأبو علام، 2010) و (سامي محمد ملحم، 2017).
- لا توجد وحدة قياس واحدة معينة ثابتة القيمة متفق عليها تستخدم في قياس السمات المختلفة، كما هو عليه في قياس الأطوال التي تقاس باستخدام السنتيمتر كوحدة للقياس أو الأوزان. فعدم الاتفاق الباحثين على تعريف بعض المفاهيم موضع الدراسة لتعقد الظواهر النفسية يزيد من نسبية القياس النفسي من ناحية، ومن ناحية أخرى لا يساعد على مقارنة أداء فرد واحد على اختبارين مختلفين مقارنة دقيقة. لهذا ينصح ضرورة تقنين وحدات القياس واجراءاته. (موسى النبهان،2004)
- لا يمكن الحسم والتأكيد في مدى وإلى أي حد تقيس الأدوات المعتمدة من قبل الأخصائيين النفسيين للسمة المراد قياسها، عكس ما هو الحال عليه في العلوم الطبيعية (وهذا ما يهتم به القياس من خلال صدق أدوات القياس)

- إن أدوات القياس في علم النفس لا تتميز بالدقة والثبات التي تتصف بها أدوات القياس في علوم الطبيعة. وهذا معناه أننا لو قسنا ذكاء شخص ما، ثم قسنا ذكاءه مرة ثانية بعد أسبوعين، لما حصلنا على نفس الدرجة بل نحصل على درجة قريبة من الدرجة الأولى. إلا أن ذلك لا يحدث في قياس الظواهر الطبيعية، فلم يحدث في الظروف الطبيعية أن تجمد الماء فوق الصفر أو تحت درجة (100) مهما كررنا عملية القياس. (سامي محمد ملحم، 2017)
- توجد أخطاء في القياس النفسي شأنه في ذلك شأن القياس في أي ميدان من ميادين العلوم الطبيعية. وهذه الأخطاء قد ترجع إلى الفاحصين (أخطاء الملاحظة) أو أدوات القياس (عدم حساسية الأدوات في قياس الظاهرة) أو عدم ثبوت الظاهرة وتغيرها المستمر أو عدم الاتفاق حول ما يقاس بين الباحثين. (صالح أحمد مراد ومحمد سليمان، 2002؛ محمد سليمان & أبوعلام، 2010)

وسنتطرق لاحقا إلى أنواع هذه الأخطاء، حيث سنتطرق للأخطاء المنتظمة في الصدق، بينما سنتطرق للأخطاء العشوائية في الثبات.

- القياس النفسي عزل للخصائص والسمات. فالسمات لا توجد بمعزل بعضها عن بعض في الطبيعة بل توجد متشابكة ومتداخلة فيما بينها، فالذكاء يتداخل مع النضج الاجتماعي والتحصيل الدراسي والنضج العقلي وغير ذلك. وبالتالي فالقياس لابد من أن يعزل السمة

المدروسة عن غيرها من السمات، بحيث تكون التقديرات التي نصل إليها دقيقة في تعبيرها الكمي.

- إن الصفر الذي نتحصل عليه من خلال القياس في علم النفس ليس حقيقيا حيث لا يعبر على عدم وجود الصفة كما هو الحال في العلوم الطبيعية. فإذا تحصل الطالب على درجة الصفر فهذا لا يعني أنه لا يعرف شيئا مطلقا، بل لا يعرف شيئا بالنسبة لعينة الأسئلة المقترحة وفقط. (صالح أحمد مراد وآخرون، 2002)

وعلى ضوء ما سبق، سنلخص الخصائص التي يتميز بها القياس في المخطط رقم 03 أدناه:

مخطط رقم 03: يوضح الخصائص التي يتميز بها القياس النفسي والتربوي

أما الجدول الموالي سنحاول من خلاله توضيح أهم الاختلافات الموجودة بين القياس النفسي والطبيعي (الفيزيائي):

جدول رقم (01): يوضح أوجه الاختلاف بين القياس النفسي والقياس الطبيعي

القياس النفسي التربوي	القياس الفيزيائي
غير مباشر	مباشر
نسبي ومرتبط بغيره	مطلق ومستقل
أقل ضبطا ودقة	أكثر ضبطا ودقة
غير مقنن ومعير عالميا	مقنن ومعير عالميا
لا تتوافر وحدات محددة للقياس	تتوافر وحدات محددة للقياس
صدق الأدوات خاضع للشك	صدق الأدوات لا شك فيها

(سامي محمد ملحم، 2017، ص.35)

المحور الثاني: مستويات القياس وبعض المعالجات الاحصائية المناسبة

أولا: مستويات القياس:

تمهيد:

سبق أن أشرنا فيما سبق إلى أن القياس بمعناه الواسع يعني استخدام الأرقام في وصف الأحداث أو الأشياء بناء على قواعد معينة. واستخدام الأرقام تبعا لقواعد مختلفة سوف نتحصل على أنواع مختلفة من المقاييس، فمثلا إذا ما قلنا إن التلميذ استغرق ثلاثة دقائق لقراءة نص أو أنه تحصل على درجة 08 في القراءة أو أنه كان أفضل من ثلاث أفراد في القراءة أو قد صنف التلميذ تبعا لمهاراته في القراءة في فئة المتفوقين، فهذه الأمثلة توضح أربع أنماط مختلفة للتعبير بالأرقام عن الأحداث، إن هذه الأساليب المختلفة للتعبير عن الأرقام تسمى مستويات القياس.

لقد قدم ستيفنس(Stevens)عام 1951 تصنيفا للقياس في ميدان علم النفس والتربية، يرتكز على أربعة مستويات، مرتبة من البسيط إلى المعقد، تبعا لمدى استخدام وتطبيق وتداول العمليات الحسابية الأساسية عليها، من جمع وطرح وضرب وقسمة أي تبعا لدقة الصياغة الكمية /الرقمية للمتغيرات التي تدرسها (محمد سليمان وأبو علام،2010، ص.59).

فيما يلي سنتطرق إلى هذه المستويات مرتبة من أبسطها إلى أكثرها تعقيدا.

1-1.المستوى الإسمي:

يطلق عليه مقياس التصنيف، يعد هذا المستوى من أبسط وأدنى مستويات القياس من حيث ملاءمته للعمليات الحسابية الأساسية ومدى تطبيقها عليه، إذ يتم تسمية أو تصنيف الأشياء أو الأفراد إلى فئات، فالأرقام تشير إلى مسمى الفئة التي تم فيها التصنيف دون أن تحمل معنى كمي.

فمثلا قد نعطي رقم (1) للذكور و رقم (0) للإناث أو نعطي رقم (1) للتلاميذ الناجحين و رقم (2) للتلاميذ الراسبين أو قد نستخدم الأرقام للإشارة إلى لاعبي فريق كرة قدم أو مجموعة من المتسابقين في العدو، لتعريف بهوية الفرد، لتحديد العنوان .....الخ.

والجدير بالذكر أنه لا يمكن اجراء العمليات الحسابية الأساسية من جمع وطرح وضرب وقسمة، كما لا يمكننا المقارنة والتطرق للعلاقات أكبر أو أصغر أو يساوي بل يقتصر الأمر على تصنيف في فئات والتصنيف في نفس الفئة عندما يكون لدينا وحدتين لهما نفس الرقم. وكل ما يمكن اجراؤه هو العد أي عدد مرات وجود الصفة أو النسب المئوية. (سليمان الطريري، 1997، ص. 83)

ومن الأمثلة على المقاييس الاسمية ما يلي:

- ما يدل على العناوين أو التسمية : أرقام الشوارع، أرقام الهوية، أرقام السيارات، أرقام الهواتف، أرقام تسجيل الطلبة، .....الخ

- ما يدل على التصنيف: تصنيف الأفراد تبعا لجنسيتهم (جزائري، مغربي، ألماني،...الخ)، تصنيفهم تبعا لديانتهم (مسلم، مسيحي، يهودي،....)، تصنيف الطلبة لناجحين وراسبين، تصنيف المرضى تبعا لنوع المرض (سكري، سرطان، القلب،....)
- التصنيف غير البشري: تصنيف الصخور تبعا لصلابتها، تصنيف المواد الصلبة تبعا لتحللها في الماء،.....الخ. (امين علي محمود سليمان & رجاء محمود أبو علام، 2010)

1-2.المستوى الرتبي:

هو مستوى أرقى، يعتمد على ترتيب مجموعة من الأفراد وفقا لدرجتهم في الصفة موضوع الاهتمام، دون أن تعطى أي أهمية لتساوي الفروق بين الرتب، فالأول يتميز بأنه صاحب أكبر درجة لصفة المقاسة والثاني الذي يليه إلى غير ذلك، فالقياس لا يقدم لنا معلومات عن الكم أو مقدار الصفة المقاسة وإنما يحدد فكرة الزيادة والنقصان، كما يسمح لنا الترتيب التنازلي والتصاعدي على سلم متدرج.

يشيع استخدام هذا النوع من المقاييس في ميدان القياس التربوي والنفسي ولا سيما في حالة تعذر القياس الموضوعي للصفات كما في حالة قياس السمات الوجدانية/ الانفعالية التي تتعلق بالاتجاهات، والتفاضلات والآراء والتقديرات في مجال التربية البدنية، الموسيقى، الخط والرسم، التمثيل، الخطابة.

ومن أمثلة على المقاييس الرتبة نذكر على سبيل المثال: مقياس ليكرت الثلاثي، ومقياس ثرستون الخماسي في ترتيب الأفراد حسب ميولهم المهنية أو غير مهنية، حسب مهاراتهم لشغل وظيفة، ترتيب التلاميذ وفقا لنتائجهم في اختبار تحصيلي، ترتيب الأبناء حسب عمرهم الزمني.

في هذا المستوى لا يمكن كذلك القيام بالعمليات الحسابية الأساسية من جمع وطرح وضرب وقسمة (سعد عبد الرحمن، 1998، ص ص. 97-98؛ محمد سليمان وأبو علام، 2010) .

1-3.المستوى الفتري أو المسافة:

يمكن أن نطلق عليه مقياس الوحدة المنتظمة أو مقياس المسافة، يشيع استعماله في علم النفس والتربية، هذا النوع من المقاييس يقترب كثيرا إلى المعنى الكمي من النوعين السابقين (الإسمي والرتبي)، بحيث يجمع بين مفهومي العدد والكم على عكس المقاييس الإسمية والترتيبية التي لا يتوافران في آن واحد عليهما، فالمقاييس الإسمية هي أعداد بلا كميات، بينما المقاييس الترتيبية هي كميات بلا أعداد، وهذا ما يجعله يتميز بأكثر دقة على المستويين السابقين.

كما تتساوى الفروق بين الأقسام المتتالية في السمة موضوع القياس، ذلك يعني أنه يتميز بوحدات متساوية تسمح لنا بتحديد مدى بعد شيئين أو شخصين بعضهما البعض في الخاصية التي نقيسها من خلال استخدام وحدات متساوية للقياس، وبذلك يمكننا تحديد ما إذا كان شيء يساوي آخر أو أكبر أو أقل منه.

غير أن من عيوبه غياب الصفر الحقيقي أو المطلق الذي يدل على عدم وجود الصفة، ويتم تحديد الصفر بشكل نسبي متفق عليه، فمثلا في مقياس الحرارة (الترمومتر)، تم الاتفاق على أن الصفر هي الدرجة التي يتجمد عندها الماء، وأن درجة 100 هي الدرجة التي يغلي عندها الماء، ومن ثم يتم تقسيم المسافة بين الصفر وهذه المائة إلى مائة وحدة متساوية، كل منها تساوي درجة واحدة.

وبهذا فالصفر النسبي أو الوضعي يعتبر خاصية تميز مقياس الوحدات المتساوية.

وإلى جانب تساوي المسافات لمقياس الوحدات المتساوية، يرتكز هذا النوع من المقاييس على افتراض أن الخصائص أو الظواهر أو القدرات أو الأبعاد التي يطبق عليها هذا النوع من المقاييس تتوزع توزيعا اعتداليا بين أفراد العينة أو العينات التي يجرى عليها الاختبار (بشرى إسماعيل، 2004).

ومن الأمثلة على مقاييس المسافة نذكر ما يلي:

- في العلوم النفسية والتربوية: اختبارات الذكاء، اختبارات القدرات الطائفة، اختبارات الاستعدادات، الميول، الاتجاهات والاختبارات التحصيلية،....الخ.
- في العلوم الطبيعية: الترمومتر المئوي، الترمومتر الفهرنهيتي، مقياس الرطوبة، مقياس الضغط.

والجدير بالذكر أنه في هذا النوع من المقاييس يمكن اجراء العمليات الحسابية الأساسية من جمع وطرح عدا الضرب والقسمة التي لا نستطيع احداثها لأن الصفر يعتبر فرضي(محمد سليمان، محمود أبو علام، 2010، ص. 63)

4-1.المستوى النسبي:

يعد أعلى مستويات القياس دقة من الناحية الكمية، يتميز أن له وحدات متساوية وصفرا مطلقا كمقاييس الطول والوزن مثلا، ناذرا ما يستعمل في العلوم التربوية والنفسية. يسمح لنا بتقدير الاختلافات بين الأشياء والظواهر والمقارنة بينها بتحديد نسبة التشابه والاختلاف، يستعمل عادة في العلوم الطبيعية، لتوفره على خصائص مقياس المسافة اضافة إلى الصفر الحقيقي فيمكن استخدام عمليات الجمع والطرح والضرب والقسمة وكل الطرق الاحصائية الممكنة. (سعد عبد الرحمن، 1998، ص. 151)

تعد مقاييس النسبة من أعلى مستويات القياس دقة من الناحية الكمية، وقد جاءت تسمية هذا النوع من المقاييس بهذا الاسم لقابلية هذه المقاييس للتعبير عنها في صورة نسبية.

تتميز مقاييس النسبة بجميع خصائص المقاييس السابقة بالإضافة إلى وجود الصفر الحقيقي أي المطلق ومن مميزاتها:

- لها صفر مطلق يدل على عدم وجود الصفة.
- وحداته متساوية.

- يمكن قياس الصفة بطريقة مباشرة بواسطة وحدات معيارية من هذه الخاصية.
- النسبة بين أي درجتين لا تتأثر بوحدات القياس (مثال: قياس الوزن سواء بوحدات النظام الفرنسي (الغرام ومضاعفاته) أو بوحدات النظام الإنجليزي(الأوقية ومضاعفاتها) تكون النسبة بين الوزنين ثابتة.

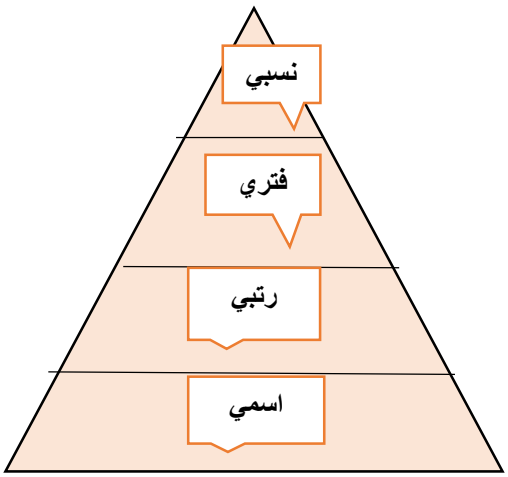
ومن الأمثلة على مقاييس النسبة ما يلي:

في العلوم النفسية والتربوية: قياس زمن الرجع، قياس التعلم بوحدات الزمن، وقياس أطوال وأوزان الأطفال بوحدات فيزيقية والجدير بالذكر ينذر استعمال مقياس النسبة في العلوم النفسية والتربوية وعادة ما يكون استعماله باللجوء إلى استخدام وحدات فيزيقية لقياس الخصائص العقلية أو الوجدانية.

في العلوم الطبيعية: قياس الأطوال (المتر ومضاعفاته)، مقاييس الوزن (الغرام ومضاعفاته)، وقياس الزمن وكذلك قياس الوحدات المشتقة كالسرعة،.... الخ(صلاح أحمد مراد& محمد سليمان، 2002، ص. 28).

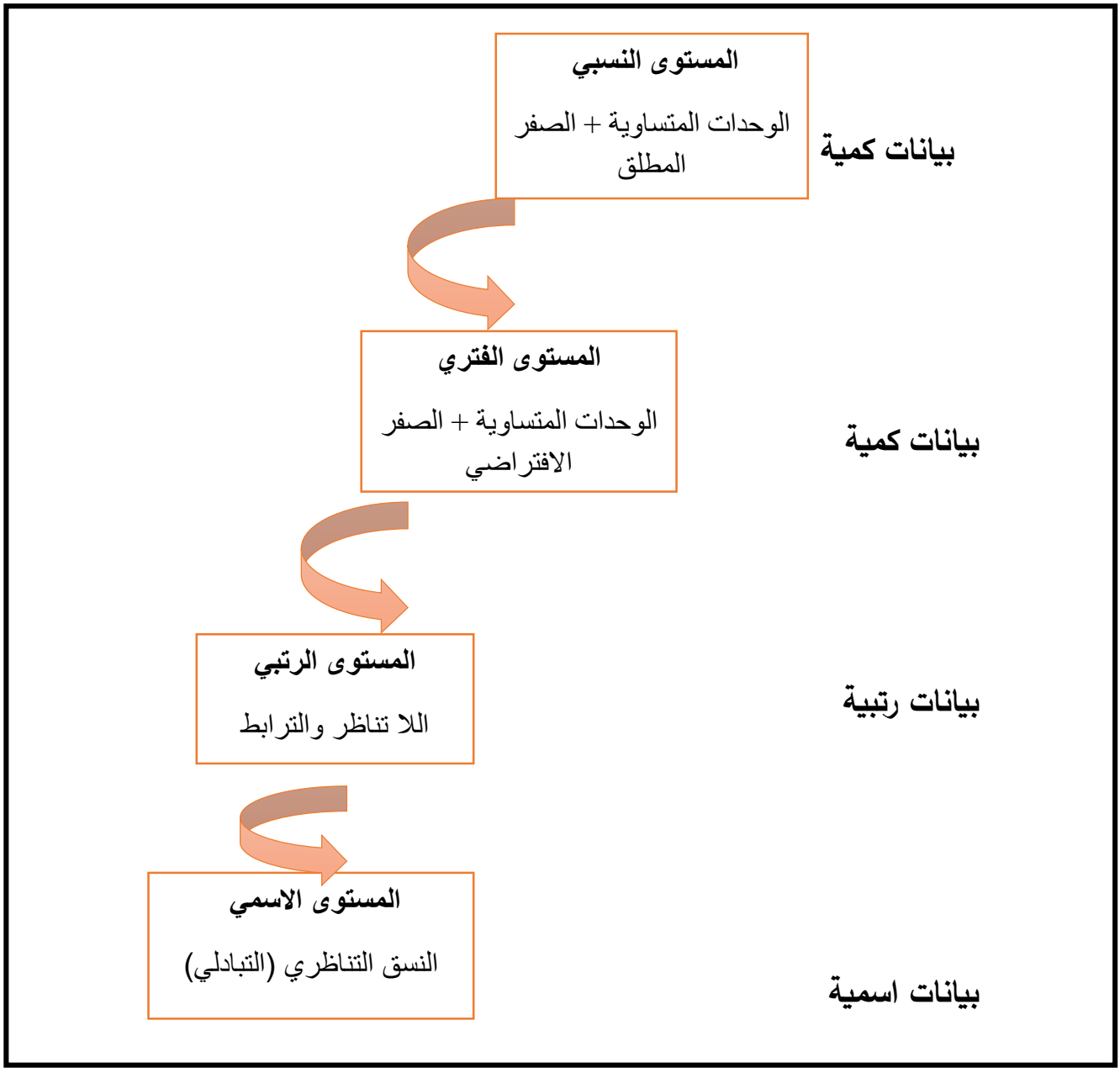
2-الفرق بين مستويات القياس:

- بناء على مدى توافر الخصائص القياسية في المقاييس، ترتب المقاييس بشكل هرمي من أدنى إلى أعلى، حيث نجد في القاعدة المقياس الاسمي والذي هو أدنى أنواع المقاييس من حيث استخدام العمليات الحسابية أما في أعلى الهرم نجد المستوى النسبي والذي يتوفر على جميع الخصائص القياسية، والمخطط رقم 04 يوضح ذلك:



مخطط رقم 04: يوضح التصنيف الهرمي لأنواع المقاييس النفسية والتربوية

- إن المقاييس الأربعة تراكمية من حيث مدى توافر الافتراضات الأساسية فيها: ونقصد بذلك أن كل مقياس أعلى ترتيبا يتضمن بالضرورة الافتراضات الأساسية للمستوى الأدنى، مثلا: إذا لم يتوفر في مقياس النسبة افتراض القابلية للتحويل إلى نسبة عن طريق القسمة، فإنه يصنف في أحد مقاييس الأدنى الذي يتمثل في المقياس الفتري لعدم وجود الصفر الحقيقي أو المطلق، أما إذا لم يتوفر في المقياس على افتراض الإضافة القائم على خاصيتي الوحدات المتساوية والصفر الافتراضي، صنف ضمن مقاييس الرتبة، وإذا لم تتوفر في المقياس على خصائص اللاتناظر والترابطية الذي تقوم عليه مقاييس الرتبة، أدرج ضمن المقاييس الاسمية.
- لكل مستوى من مستويات القياس عملياته الكمية الخاصة والتي ترتبط بأساليب إحصائية مناسبة لها.
- والجدير للإشارة أن الإحصاء اللابرامتري يمكن استخدامه في جميع مستويات القياس، بينما يتم تطبيق الإحصاء البرامتري في المستوى الفتري والنسبي.



المخطط رقم 05: يوضح الخصائص الأساسية لكل مستوى من مستويات القياس

ثانيا: بعض المعالجات الإحصائية المستخدمة والملائمة لأنواع القياس:

مقدمة :من البديهي أننا لا توقف فقط على قياس ظاهرة نفسية من أجل وصفها وفقط، فإننا لا نكتفي مثلا بمعرفة النتيجة التي تحصل عليها الطلبة في وحدة ما، بل نبحث عن أسباب النجاح أو الفشل واتخاذ القرارات اللازمة تبعا لذلك، وهذا لا يتأتى إلا بالاعتماد على الإحصاء.

لهذا ارتأينا في هذا الشأن إلى التطرق بشكل مختصر إلى أهم المعالجات الإحصائية المعتمدة في البحوث النفسية تبعا لملاءمتها لأنواع القياس.

المستوى الاسمي: إن المعالجات الإحصائية المناسبة تقوم على فكرة العد البسيط (سعد عبد الرحمن، 1998: 77)

مقاييس النزعة المركزية:

- المنوال: في هذا المستوى يمكن استعمال فقط المنوال والتي تتمثل في القيمة الأكثر تكرارا في التوزيع الدرجات (سليمان الطريري، 1997، ص. 98).

مقاييس الدلالة الاحصائية:

- اختبار كا2x^2: هي أداة احصائية تمكننا من معرفة مدى التشابه بين توزيعين أحدهما ملاحظ و الآخر نظري (متوقع).

مثال: معرفة ما إذا هناك فوارق في الكفاءة في التسيير وهل يمكن ارجاعها لجنس المسير؟ ويمكن طرح المشكل على النحو الآتي: هل يختلف مستوى الكفاءة في التسيير تبعا للجنس؟ (سعد عبد الرحمن، 1998، ص ص. 78-89)

- اختبار ماك نمار لدلالة التغير: يستخدم في حالة المقارنة بين استجابات الفرد في المجموعة الواحدة نتيجة احداث معالجة معينة أفرزت تغير ما بعد فترة زمنية. (سعد عبد الرحمن، 1998، ص. 93)

مثال: معرفة مدى تأثير حملة إعلامية من أجل التعبئة لانتخاب حزب ما سوف تقارن استجابة نفس الأفراد ما قبل وبعد الحملة الإعلامية.

مقاييس الارتباط: ان مقياس التصنيف يعتبر الخطوة الأولى لقياس الارتباط بين الظواهر مع بعضها البعض (سعد عبد الرحمن، 1998، ص. 77)

- معامل كرامر: يستعمل في حالة بيانات نوعية ووجود متغيرين أحدهما يكون له أكثر من تقسيمين، وهو صورة معدلة لمعامل فاي.

مثال: هل هناك علاقة بين المستوى التعليمي للفرد (ابتدائي، متوسط، ثانوي) و الهواية المختارة (الرياضة، المطالعة و المراسلة)؟. (محمد بوعلاق، 2012، ص. 95-97)

- معامل الترافق: هو عبارة عن أداة يستعمل لقياس العلاقة الارتباطية بين متغيرين نوعيين، يطبق عندما يتم تصنيف المتغيرين أو تصنيف أحدهما إلى صنفين أو أكثر.

مثال: هل هناك علاقة بين نوعية الدراسة (أكاديمية، تطبيقية، عملية وتجارية) والطبقة الاجتماعية (1، 2، 3، 4)؟. (محمد بوعلاق، 2012، ص ص. 109-111) و (سعد عبد الرحمن، 1998، ص ص. 90-91)

- معامل فاي: يستعمل لدراسة العلاقة بين متغيرين نوعيين و ثنائيين التقسيم، يكثر استعماله عند دراسة العلاقة بين مختلف بنود اختبار معين وعندما تكون الاجابة على البنود ب نعم أو لا. مثال: هل هناك علاقة بين الجنس والتغيب؟ (محمد بوعلاق،2012، ص ص. 93-94)

المستوى الرتبي: تقوم المعالجات الإحصائية على مبدأ ترتيب الوحدات على معيار أو أكثر

مقاييس النزعة المركزية:

- الوسيط: هي القيمة التي تقسم التوزيع إلى مجموعتين متساويتين في عدد القيم بحيث يكون عدد القيم التي تسبق القيمة الفاصلة أو القيمة الوسطية تساوي عدد القيم التي تليها (سليمان الطريري، 1997، ص. 97).
- المنوال: هي القيمة الأكثر تكرارا

مقاييس التشتت: يعرف (Gay, 1990)التشتت لأي مجموعة من القيم، بالتباعد بين مفرداتها أو التفاوت أو الاختلاف بينها. ويمكننا أن نتخذ مقدار التشتت قليلا كان أم كبيرا، كدليل على تجمع القيم وقربها من بعضها أو تفرقها وتباعدها عن بعضها البعض. وهذا يكون مقياس لمقدار تجانس المجموعات الإحصائية او تجانسها (سامي ملحم، 2012، ص.99)

- الانحراف الربيعي: هو أفضل من المدى المطلق لأنه لا يتأثر بالقيم المتطرفة حيث يعتمد حسابه على تعيين الربيعين ر1 وهو النقطة التي يقع عندها ربع أفراد المجموعة أي 25٪ ور3 النقطة التي يقع عندها لتي يقع عندها ربع أفراد المجموعة ويقل عنها 75٪ منهم. (كامل محمد مغربي، 2017، ص. 184)

مقاييس المكانة النسبية:

- العشاريات: هو المقياس الذي بموجبه يتم تقسيم التوزيع إلى 10 أجزاء متساوية.
- الرتبة المئينية: هو المقياس الذي بموجبه يتم تقسيم التوزيع إلى 100 جزء متساو ويمكن تعبير عنه بمكانة الفرد النسبية في التوزيع. (بن سليمان الطريري، 1997، ص. 120).

مقاييس العلاقة:

- معامل ارتباط سبيرمان لرتب: هو أداة احصائية تستعمل لقياس العلاقة بين متغيرين أو أحدهما يتضمن بيانات رتبية، يعتمد على حساب الفرق بين الرتب. مثال: هل هناك علاقة بين تقييمي أستاذين في علم النفس التربوي؟ (سعد عبد الرحمن، 1998، ص. 112)
- معامل كندال للارتباط أو التوافق: يستخدم عندما يتم ترتيب المجموعة الواحدة بناء على ثلاث معايير. مثال: هل هناك علاقة ارتباطية بين نتائج طلبة علم النفس في القياس النفسي، علم النفس المرضي وعلم النفس الفيزيولوجي. (محمد بوعلاق،2012، ص ص. 115-117)

المستوى الفتري أو المسافة:

مقاييس النزعة المركزية:

- المتوسط الحسابي: يسمح بتحديد مركز القيم لمقارنة نتائج كل فرد ومدى قربه أو ابتعاده عن متوسط أداء الجماعة الذي ينتمي اليها. (كامل محمد مغربي،2017، ص. 171-177)
- الوسيط والمنوال المستخدمين في المستويات الدنيا

مقاييس التشتت:

- الانحراف المعياري: يبين مدى تشتت درجات أفراد مجموعة حول المتوسط. ويعرف على أنه الجدر التربيعي لمربع مجموع الفروق بين قيم المفردات ومتوسطها الحسابي مقسوما على عدد المفردات. (كامل محمد مغربي، 2017، ص. 189)

معامل الاختلاف: يعبر عن الانحراف المعياري بالنسبة المئوية للمتوسط، يستعمل للمقارنة بين المجموعات. وهو الجدر التربيعي لتباين، يفضل استخدامه بدلا من التباين لأن وحدة القياس فيه مساوية لوحدة القياس في البيانات الأصلية (محمد بوعلاق، 2012، ص. 67)

مقاييس المكانة النسبية:

- الدرجة المعيارية: تدل على المسافة التي يبعدها الفرد عن المتوسط المجموعة التي ينتمي إليه، تستعمل لمقارنة الدرجات في المجموعة الواحدة وبين مجموعات مختلفة. (سامي محمد ملحم، 2015: 179)

2-3-4. اختبارات الدلالة الإحصائية:

- الخطأ المعياري للأدوات الإحصائية مثل المتوسط، الوسيط، الانحراف المعياري، الانحراف الإرباع، النسبة المئوية ومعامل الارتباط (سعد عبد الرحمن، 1998: 126-130)
- النسبة التائية أو اختبار (ت): يستعمل هذا الاختبار لدراسة دلالة الفروق بين المتوسطات المرتبطة وغير مرتبطة للعينات المتساوية وغير متساوية (محمد بوعلاق، 2012: 141)

2-3-5. الارتباط:

- معامل ارتباط بيرسون: يستعمل لدراسة العلاقة بين متغيرين كميين س وع أحدهما مستقل والآخر تابع.
- معامل الارتباط المتعدد: يستعمل بكثرة في البحوث النفسية والتربوية والاجتماعية لتعقد وتداخل المتغيرات، يدرس العلاقة بين متغير تابع وعلى الأقل متغيرين مستقلين أو متغير مستقل وعلى الأقل متغيرين تابعين. (محمد بوعلاق، 2012: 79-87)
- معامل ارتباط الثنائي: يستعمل لدراسة العلاقة بين متغير كمي وآخر نوعي يكون فيه التقسيم غير حقيقي.
- معامل الارتباط الثنائي الأصيل: يستعمل لدراسة العلاقة بين متغير كمي وآخر نوعي يكون فيه التقسيم حقيقي.

(محمد بوعلاق، 2012: 98-104).

المحور الثالث: أدوات القياس

تمهيد:

عرفت أدوات القياس النفسي والتربوي في العقود الأخيرة انتشارا واسعا وتطورا هاما من حيث التنوع والاستخدامات المتعددة ولا سيما في مجال التقويم التربوي، حيث تعدى استخدامها قياس مخرجات العملية التعليمية، بل أصبحت تستخدم في مختلف مراحل التعلم: تشخيص المكتسبات أو المعارف السابقة، تشخيص صعوبات التعلم للوقوف عليها ومعالجتها، التقييم النهائي ومعرفة مدى تحقق الأهداف المسطرة مسبقا، بالإضافة إلى الاستعانة بها لتقييم مهارات المعلم، المناهج، محتويات التعليم، الكتب المدرسية، طرائق التدريس، كفاءات المعلمين، المهارات الإدارية ،...الخ.

وفي هذا السياق سنحاول التطرق إلى أهم الأدوات المستخدمة في القياس والتقدير النفسي والتربوي باعتبارها من الأدوات الضرورية التي قد يلجا لها المختص التربوي لاتخاذ قرارات في مجال التقييم والتقويم التربويين.

1-أدوات القياس والتقدير النفسي والتربوي:

تعتمد العلوم كثيرا على القياس في تطورها، ولقد قطع علم النفس أشواطا معتبرة في ميدان القياس والتقدير رغم تعقد وصعوبة قياس وتقدير أغلبية الظواهر السيكولوجية لكونها لا تخضع للملاحظة المباشرة مثل الظواهر الفيزيائية. وللتغلب على هذه الصعوبة عمد الباحثون السيكولوجيون على تطوير تقنيات بناء على المؤشرات الخاصة أو الدالة على السلوك موضوع القياس أو التقدير وذلك

من أجل الكشف عن المقدار أو مقدار الاستعداد الذي يتمتع به الفرد فيما يخص هذا السلوك، مثلا الذكاء، الذاكرة، الميول، القيم، أبعاد الشخصية.....الخ.(بوسنة، 2007، ص.127)

1-1. أدوات القياس:

إن أدوات القياس متنوعة، ومن أهمها نذكر الاختبار النفسي، الاستبيان والمقياس. سنحاول التطرق لها ونبين الاختلافات فيما بينها.

الاختبار النفسي (Test) :

يعرفه كل من لي و كرونباخ(Cronbach & Lee, 1984) على "أنه طريقة منظمة لمقارنة سلوك شخصين أو أكثر «، وفي نفس الاتجاه ذهبت أنستازي(Anastasi, 1982) حيث ترى أن الاختبار عبارة عن" مقياس موضوعي مقنن لعينة معينة من السلوك"(بشير معمرية ، 2012، ص.104).

كما تعرفه كل من الجمعية الأمريكية للبحث التربوي (AERA) والجمعية الأمريكية لعلم النفس (APA) والمجلس القومي الأمريكي للقياس التربوي (NCME)على أنه:" اجراء يتم بواسطته الحصول على عينة من سلوك الفرد وتقييمها وتقدير درجاتها باستخدام إجراءات مقننة (رينولدسو ليفينقستون ترجمة محمود علام، 2013، ص. 38).

وبذلك فالاختبار النفسي قياس موضوعي ومقنن لظاهرة سلوكية باستخدام عينة من السلوك وترجمتها كميا وتفسيرها تبعا لمعايير محددة.

ويقصد بالموضوعية بعدم تدخل ذاتية الفاحص في تقدير النتائج المتحصل عليها من الاختبار، أما فيما يخص عينة السلوك، فالقياس النفسي يحذو حذو القياس الفيزيائي الذي يختبر عينة أو أكثر للدم مثلا لقياس ضغط الدم، فالقياس النفسي يقوم على انتقاء عينة ممثلة للسمة المراد قياسها وتعميم نتائجها.

في حين يقصد بالتقنين توحيد إجراءات التطبيق، التصحيح والتفسير ودراسة الخصائص السيكو مترية من ثبات وصدق على عينة ممثلة وواسعة لمجتمع الدراسة.

ويجدر الاشارة بان الاختبارات النفسية صممت بالدرجة الأولى لقياس القدرات العقلية، بينما صممت بعض الاختبارات النفسية لغرض قياس القدرات الحسية الحركية وأخرى لقياس خصائص الشخصية، ونذكر من الاختبارات على سبيل المثال اختبار بينه وسيمون للذكاء، اختبار ويكسلر واختبار تفهم الموضوع TAT واختبار الروشاخ و...الخ. (د بوسنة ، 2007)

أما في المجال التربوي، بالإضافة إلى الاختبارات النفسية التي قد يلجأ لها المختص التربوي من أجل مساعدته على اتخاذ قرارات صائبة في مجال التقويم، يرتكز الأخصائي التربوي أساسا لمعرفة نتاجات التعلم على استخدام الاختبارات التحصيلية وهي أنواع سنتحدث لها بشكل من التفصيل في المحور الخاص بالاختبارات التحصيلية.

المقياس (Mesure) :

تعرف أنستازي (1976) المقياس بأنه "أداة موضوعية مقننة لتحديد عينة من السلوك " وهو نفس التعريف الذي قدمته للاختبار النفسي، وهذا ما لا يتفق تماما مع وجهة نظر ليونا تايلر 1982، حيث ترى أنه هناك تداخلا في المعنى بين الاختبار والمقياس، غير أنهما ليسا مترادفان.

فالمقاييس تهتم بقياس الجوانب الفيزيائية او الطبيعية مثل مقاييس الفيسيولوجية (نبضات القلب، المخطط الكهربائي للدماغ، القدرة على الابصار، القدرة على التركيز ...الخ) ومقاييس الظروف الفيزيقية للعمل (الضوضاء، الحرارة، درجة الإضاءة .....الخ) وكذا بالجوانب النفسية الفيزيقية كمقاييس السيكو فيزيائية (العتبات وزمن الرجع، القدرة على الابصار، القدرة على التركيز، تردد الصوت لمعرفة حدود السمع الدنيا والعليا، مستويات الانتباه،.....الخ). (د. بوسنة، 2007؛ د. معمرية، 2012)

الاستبيان أو الاستخبار (Questionnaire)

الاخبار في اللغة العربية هو السؤال عن الخبر، ويقترب هذا المعنى من اللغوي مما نقصده في علم النفس من اجراء الاستخبار، وهو معرفة "اخبار" أو "معلومات معينة عن الشخص.

ويعرف أحمد محمد عبد الخالق الاستخبار بأنه: " طريقة من طرق قياس السمات أو الأبعاد الأساسية للشخصية، وهو نوع من المقابلة المقننة، ويتكون من مجموعة من الأسئلة أو العبارات التقريرية المطبوعة غالبا، يجيب عنها المفحوص بنفسه (كتابة غالبا وشفاهية أحيانا) بناء على

احتمالات أو فئات للإجابة محددة سلفا، مثل نعم لا، أو موافق غير موافق في موقف قياس فردي أو جماعي "

وعليه تهتم أسئلة الاستخبار بالجوانب الوجدانية والانفعالية أو بالسلوك الإنساني في المواقف الاجتماعية، بالميول والاتجاهات والقيم...الخ.

يجيب المفحوص على أسئلة الاستخبار وفق معرفته والوعي بمشاعره، انفعالاته وسلوكياته.

يكون الاستخبار أحادي بحيث يقيس سمة واحدة، وقد يشمل على أبعاد متعددة في نفس الوقت مثل استخبارات لأبعاد الشخصية وسماتها (الانبساطية، العصابية، المثابرة والسيطرة،....الخ).

ويندرج الاستخبار ضمن اختبارات الورقة والقلم، لذلك هناك من يعتبره اختبار لهذا الغرض، غير أنه يختلف عنه، فالاختبار أدق ويهتم بالاستعدادات والذكاء والقدرات، بينما الاستخبار يهتم بجوانب مختلفة من الشخصية والدوافع والحاجات والاتجاهات والقيم، كما يمكن استخدام الاستخبار بطريقة غير مباشرة (الاتصال غير مباشر بين الفاحص والمفحوص) على عكس الاختبار الذي ينبغي تطبيقه على المفحوص مباشرة. (د. معمرية، 2012)

وقد صدرت بين الحربين العالميتين استخبارات كثيرة تقيس سمات عامة أو محددة كالانبساط– الانطواء، والسيطرة –الخضوع، والأمان–عدم الأمان، ...الخ.

ومن الاستخبارات المعروفة نذكر: قائمة منيسوتا متعددة الأوجه للشخصية (1943)، (1989) وقائمة كاتل ذات 16 عاملا في الشخصية (1949)، قائمة ايزنك للشخصية (1961)، قائمة كاليفورنيا النفسية (1957)،.....الخ.

مزايا الاستخبارات: هناك عدة مزايا تترتب على استخدامها نذكر من أهمها:

- إجراءات واضحة وموضوعية لجمع البيانات
- المرونة في التطبيق، إذ تطبق فرديا وفي جماعة في نفس الوقت مما يوفر الجهد والوقت.
- منخفضة التكاليف من حيث الوقت والجهد والمال مقارنة بالطرق الأخرى.
- سهولة التصحيح وموضوعية نظام التصحيح.
- إمكانية تحديد معايير لها أو تقنينها على مجموعات كبيرة مما يسمح بالمقارنة بين الأشخاص.
- يمكن ان تعالج الدرجات المستخرجة منها إحصائيا.
- معاملات ثباتها ليست منخفضة مقارنة بالاختبارات الإسقاطية
- تستخدم في مجالات متعددة منها الارشاد والتوجيه، المجال الاكلينيكي، الصناعي،....الخ (أحمد عبد الخالق، 1993)

عيوب الاستخبارات: واجهت الاستخبارات انتقادات عديدة ومنها نذكر:

- تأثير تغير صياغة السؤال: لوحظ أن نسبة الإجابة بـ "نعم" أو "لا" تتغير وفق اتجاه الصياغة اللفظية للعبارة رغم ان المحتوى نفسه.

مثال حول تغيير اتجاه العبارة او البند: (كثيرا ما أصاب بالصداع أو قلما أصاب بالصداع)

- مشكلة صياغة الإجابة: كثيرا ما يعترض المفحوصين مشكلة تحديد الإجابة بـ نعم أو لا، كما يختلف المفحوصين في تفسير الفئات الخماسية للإجابة (غالبا، كثير ما، أحيانا، نادرا).
- اختلاف اتجاهات المفحوصين نحو الاستخبار: يستجيب المفحوص للاستخبار تبعا لدوافعه، فقد يجيب تبعا لنوع الشخصية المناسبة للموقف وليس وفق لسماته الفعلية وما يشعر به في الحقيقة.
- تنوع العوامل التي تؤثر في استجابة المفحوص: اتجاه المفحوص نحو موقف القياس، عدم معرفة الشخص بذاته، عامل الإيحاء، ذكاء المفحوص وفهمه لبنود الاستخبار، كل هذه العوامل تتدخل في نوع الاستجابة التي يدلي بها المفحوص.
- تأثير الحالة المزاجية الراهنة والخبرات الحديثة: حساسية الاستخبارات لحالات مثل التعب، الضيق والسرور وللخبرات التي يمر بها الفرد حديثا.
- عدم الدقة في التقنين للاستخبارات: كثيرا من الاستخبارات قننت على فئة الجامعيين، والتي تعتبر فئة مختارة لا تمثل المجتمع.
- اختلاف ظروف التطبيق والتقنين: يصعب على الفاحص المقارنة بين الظروف التي يجيب فيها المفحوص والظروف السائدة خلال عملية التقنين.

1-2. أدوات التقدير:

هي أقدم من أدوات القياس باعتبار أن التقدير يقوم على التخمين كما سبق أن تطرقنا له في المفاهيم ذات صلة بالقياس، استخدمت لأول مرة في القرن التاسع عشر من قبل البحرية البريطانية لوصف حالات الطقس، وبعد ذلك استخدمها فرانسيس جالتون لتقييم حيوية الصور، أما في بداية القرن العشرين فقد استعان بها كارل بيرسون في تقدير الذكاء، كما تم استخدامها من طرف علماء النفس على نطاق واسع في بداية الحرب العالمية الأولى لتقدير كفاءة الضباط، ومن ذلك الوقت راج استخدامها في مجالات متعددة، سواء في علم النفس أو المجال التربوي.

ويجدر الإشارة أن أدوات التقدير تستخدم في تسجيل الملاحظات حول سلوك الفرد، قدراته، مهاراته، طرق تعلمه،....الخ ووجدت لتقدير جوانب متعددة من المواضيع التي يصعب أو يتعذر قياسها باستخدام أدوات قياس دقيقة أو لعدم وجود أدوات قياس أكثر موضوعية ودقة من أدوات التقدير (الرسم، التعبير الكتابي، اكتساب المهارات المختلفة، تقدير الميول والاتجاهات، تكيف أطفال الروضة ....الخ) (توما جورج الخوري، 2008).

وأدوات التقدير نوعان وهما: قوائم التقدير وسلالم التقدير سيتم تناولهما على التوالي:

سلالم التقدير(Rating scales) :

تسمح لنا بجمع بيانات حول مقدار وجود السمة، مثلا وجودها بشكل عال، عال نوعا ما، متوسط، تحت المتوسط ومنخفض.

ويجدر الإشارة أن عملية التقدير قد تكون من الأخصائي النفسي ونتكلم في هذه الحالة عن التقدير الموضوعي، وقد تتم عملية التقدير من طرف المفحوص نفسه وهنا نتكلم عن تقدير ذاتي، حيث يقوم المفحوص بتأشير حول وجود أو عدم وجود السمة موضع التقدير أو وضع علامات تقدير درجة وجود السمة.

فباستثناء التقدير الذاتي للمفحوص لسمة موضع القياس فإن سلالم التقدير تقوم على ملاحظة الفاحص لسلوك المفحوص، حيث يسجل ملاحظاته عن طريق سلم التقدير المصمم، مباشرة أو بعد اجراء الملاحظة. (بوسنة، 2007؛ الانصاري،2009)

توجد عدة أنواع من سلالم التقدير منها نذكر سلم التقدير العددي (سلم عددي مدرج من 1 إلى 5 أو 7...الخ)، سلم التقدير البياني، تمايز المعاني، سلم التقدير المعتمد على السلوك، سلم التقدير ذو الاختيار المقيد، سنحاول التطرق لها بشكل موجز فيما يلي:

سلم التقدير الرقمي Numerical Rating Scale:

في هذه الطريقة يتم التقدير بتحديد رقم واحد من بين عدة أرقام تشير لوصف محدد للخصائص، والمثال الموالي يوضح ذلك:

الجدول رقم (02): نموذج حول سلم التقدير الرقمي

الصفة	لا تصفه أبدا				تصفه تماما
الخجل	1	2	3	4	5
الغضب	1	2	3	4	5
القلق	1	2	3	4	5
الاكتئاب	1	2	3	4	5
الخوف	1	2	3	4	5
الذنب	1	2	3	4	5
الحرج	1	2	3	4	5

(الأنصاري، 2009، ص.404)

ومن أمثلة على سلالم التقدير في المجال التربوي:

سلم لتقدير عادات العمل لدى التلميذ

الاسم: .....الصف: ..... المدرسة: .....المربي:

التعليمة: ضع مربعا حول ما ينطبق على التلميذ

الجدول رقم (03): نموذج حول سلم التقدير الرقمي في المجال التربوي

السمة		ضعيف	متوسط	جيد	ممتاز
الانتباه	1	2	3	4	5
الاجتهاد	1	2	3	4	5
المثابرة	1	2	3	4	5
الدقة	1	2	3	4	5
الاعتماد على النفس	1	2	3	4	5
الهدوء أثناء العمل	1	2	3	4	5

(توما جورج الخوري، 2008، ص.50)

سلم التقدير البياني Graphic Rating Scale:

هذا النوع أكثر سلالم التقدير شيوعا، وفيه تمثل السمة المراد تقديرها على خط بياني مقسم إلى 5 أو 7 درجات (أو وحدات متساوية) غالبا، يطلب من القائم على التقدير تعيين العلامة التي تتطابق مع سلوك المفحوص.

الانطواء

منطوى جدا منطوى متوسط منبسط منبسط جدا

(الأنصاري، 2009، ص. 405)

وفي المجال التربوي قد يستخدم المعلم سلما بيانيا لتقدير مدى اشتراك التلميذ في المناقشة، والمثال أدناه يوضح ذلك:

1-هل يطلب الاشتراك في المناقشة برفع اليد ؟

2-هل يحترم حق الآخرين في الكلام؟

يتبين من المثال أن مقدار الصفة محدد بألفاظ (أبد، أحيانا،...)،حيث يختلف مدلولها من فاحص إلى آخر يسمى هذا النوع من سلالم التقدير بالبياني اللفظي.

لذا يستحسن استخدام أوصاف تحدد مستويات أو مقدار السمة بدلا من الألفاظ، بحيث التقدير يكون أكثر موضوعية ودقة، ويسمى في هذه الحالة بالسلم البياني الوصفي، كما هو موضح في هذا المثال:

1-ما مدى مساهمة التلميذ في المناقشة؟

(توما جورج الخوري، 2008، ص ص.54-55)

تمايز المعاني Semantic Differential:

هو نوع من مقاييس التقدير الرقمية، يستخدم بشكل شائع لدراسة الشخصية وفي مجال علم النفس الاجتماعي.

يطلب من المفحوص تقدير سلسلة من المفاهيم على مقياس ثنائي القطب كما هو مبين في الجدول:

الجدول رقم (04): مقياس التقدير الرقمي لمفهوم الحاسب

الحاسب (الكمبيوتر)							
قلق	1	2	3	4	5	6	7 هادئ
متشائم	1	2	3	4	5	6	7 متفائل
قبيح	1	2	3	4	5	6	7 جميل
سلبي	1	2	3	4	5	6	7 ايجابي
صغير	1	2	3	4	5	6	7 كبير
ضعيف	1	2	3	4	5	6	7 قوي
انطوائي	1	2	3	4	5	6	7 انبساطي
عاجز	1	2	3	4	5	6	7 جبار
متخلف	1	2	3	4	5	6	7 ذكي عقليا

(الأنصاري،2009، ص.406)

سلم التقدير ذو الاختيار المقيد Choice Rating Scale Forced:

يقدم في هذا النوع وصفان أو أكثر ويطلب منه أن يحدد الوصف الذي يميز الفرد المفحوص تميز جيدا.

- يتحمل المسؤولية بسهولة.
- لا يعرف كيف يتفاوض ومتى
- لديه اقتراحات بناءة كثيرة
- لا يستمع إلى اقتراحات الأخرين.

مزايا سلالم التقدير:

- تسمح سلالم التقدير بالوصف الكمي للسمات والخصائص السلوكية
- تتيح سلالم التقدير إمكانية اشتراك عدد من الحكام في الحكم على المفحوص باستخراج متوسط أحكامهم وذلك للتقليل من التحيز
- إمكانية أن يتناقش الحكام بخصوص المفحوص وهذا ما يرفع ثبات تقديرات الحكام.
- تتميز بالمرونة حيث يمكن استخدامها لتقدير مدى واسع من الخصائص السلوكية والشخصية. (أحمد عبد الخالق، 1996، ص.164)

عيوب سلالم التقدير:

من أهم عيوب سلالم التقدير نقلا عن (أحمد عبد الخالق، 1996):

- غموض السمات المراد تقديرها، فقد يعرف الأخصائي القائم بالتقدير صفة القيادة مثلا بالقدرة على استخدام السلطة، بالقرارات الصارمة والسيطرة العامة أو قد يقصد بها تشجيع المرؤوسين على العمل التعاوني واتخاذ قرارات مشتركة، وعليه يختلف تقدير سمة القيادة باختلاف مفهومها.
- غموض وحدات القياس ودرجاتها، وعدم وضوح الفروق بين كل درجة والتي تليها، كما أن الوصف الكيفي المستخدم مثل: كثيرا غالبا وأحيانا ونادرا يختلف تفسيره بين قائمي على التقدير النفسي، كما يتباين تفسيره لدى الفرد ذاته وفق السياقات المختلفة.
- عدم اتاحة الوقت الكافي للقائم على تقدير لملاحظة المفحوص، ولذلك سيحاول الفاحص ملأ الثغرات بطريقة ما، فقد يقوم بالمبالغة في تعميم الملاحظة خاصة إذا كانت إيجابية (أحمد عبد الخالق، 1996، ص ص. 165-169).

قوائم التقدير(Rating cheecklists) : إن قوائم التقدير تسمح لنا بالتقصي حول وجود السمة التي هي موضع التقدير، وتعتبر من الصيغ الشائعة التي تستخدم لقيا س الشخصية.

بحيث يقدم للمفحوص مجموعة من الصفات، ويطلب منه أن يحدد ما إذا كانت كل صفة تميزه أو لا، وذلك بالتأشير على الصفة التي تنطبق عليه وترك الصفات التي لا تنطبق عليه.

تستخدم قوائم الصفات إما لوصف شخص ذاته (التقدير الذاتي)، أو وصف شخص آخر يعرفه تمام المعرفة.

بالإضافة إلى استخدام قوائم الصفات لدراسة الشخصية وتحديد المشكلات التي يواجها الفرد من مختلف النواحي: الانفعالية والصحية والشخصية والأسرية والمالية والاجتماعية، لها استخدامات كلينيكيه وبحثية ودراسة السلوك في العمل.

أمثلة على بعض نماذج قوائم الصفات:

- قائمة الصفات لـ (Gough,1952) والتي تم اعادة نشرها عام 1980 بعد تنقيحها بالاشتراك مع (Heilbrun) وتشتمل على 300 صفة مرتبة أبجديا، بحيث يطلب من المفحوص ان يضع علامة على الصفات التي يرى أنها تصفه، وتقدر الدرجات على أساس 37 مقياسا فرعيا. وقام عبد الرحيم بخيت (1987) بتعريب القائمة واعدادها للبيئة العربية.
- قائمة الصفات المزاجية: صمم هذه القائمة (Nowlis,1965)، حيث تشتمل على 140 صفة مثل مستوى النشاط، التركيز، الضبط....الخ ، يجيب عليها المفحوص بوضع علامة صحيح على أحد الاحتمالات الأربعة (تنطبق تماما، إلى درجة بسيطة، غير محدد، لا تنطبق أبدا) التي تبين إلى أي مدى تصف كل صفة مشاعره.
- قائمة الصفات الوجدانية: تم وضع هذه القائمة من طرف (Zuckerman& Lubin) لوصف حالات القلق، ثم صدرت نسخة أخرى عام 1965 لترصد حالات القلق والاكتئاب والعداء (صيغة طويلة تشمل على 132 بند وصيغة مختصرة تتكون من 48 بند)، كما تم مراجعتها عام 1985، بحيث أصبحت تقيس خمسة صفات وجدانية أساسية تتمثل في: القلق، الاكتئاب، العداء، الوجدان الإيجابي، والبحث عن الإثارة (تضم 132 صفة)

المجال التربوي:

ولا يقتصر استخدام قوائم التقدير على المجال النفسي، حيث يتم الاستعانة بها كذلك في المجال التربوي لتقييم التعلم، الكشف عن مشكلات المتعلمين، تقويم المناهج، تقويم مهارات المتعلمين والمعلمين على السواء،.....الخ.

فقد تستخدمها مربية الروضة لتقييم نظافة الطفل، فتحدد مسبقا مختلف المؤشرات التي تدل على النظافة مثل: تسريحة الشعر، نظافة الملابس، نظافة الجسم (الأيدي، الأسنان، الأنف، الرقبة، الأظافر)،....الخ، وفق لحاجتها وتقدمها لأم الطفل لمراقبته يوميا، كما قد تستخدم القائمة لتدريب الأطفال على العادات الصحية في الروضة(توما جورج خوري، 2008).

كما قد يستخدمها المعلم لتقييم ما اكتسبه التلميذ من معارف ومهارات مثل مهارات الانشاء، وفيما يلي مقترح لقائمة تقدير الانشاء وضعها توما جورج خوري(2008)، يمكن للمعلم أن يضيف ما يراه ضروري وفق ما تم تلقينه في الدروس:

قائمة تقدير الانشاء لتوما جورج خوري(2008)

- أخطاء املائية
- أخطاء في الإعراب
- أخطاء نحوية
- لا يعرف التنقيط
- مفردات عامية
- لا يعرف كيف يبدا الفقرات وكيف ينهيها
- تفكير غير منظم
- لا أصالة في التفكير
- لا هوامش مناسبة ولا تاريخ
- خط سيئ وغير واضح

- استعمال الكلمات في غير معناها -حقائق وأفكار خاطئة

(توما جورج خوري، 2008، ص.40)

وإلى جانب تقييم ما تعلمه المتعلم من معارف ومهارات وسلوكيات مرغوب فيها، يمكن استخدام قوائم التقدير في تقييم المنهاج، والمثال الآتي يوضح ذلك:

نموذج تقييم الكتاب المدرسي

إن هدف هذا النموذج تحسين الكتاب المدرسي. لذا عبر عن رأيك بأن تكتب "نعم" أو كلمة "لا" إلى يمين كل عبارة من العبارات التالية:

1- مقدمة هذا الكتاب تساعد الطالب في دراسته.

2-أسلوب الكتاب مشوق للقراءة.

3-أسلوب الكتاب يحفز الطالب على التفكير والبحث.

4-أسلوب الكتاب يساعد الطالب على الاعتماد على نفسه في فهم المشكلات وحلها.

5- مادة الكتاب تناسب مستوى الطالب.

6-الكتاب يخلو من التكرار بين وحداته.

7-عدد الحصص المقررة يكفي لاستيعاب مادة الكتاب.

8-تتابع الأفكار والمفاهيم في الكتاب جاء متسلسلا واضحا.

9-لغة الكتاب سليمة واضحة.

10- الكتاب يراعي الدقة في مادته.

11- الكتاب يراعي الحداثة في مادته.

12- قائمة المصطلحات الواردة في الكتاب تساعد الطالب في دراسته.

13- الأسئلة الواردة في نهاية كل وحدة تساعد على تركيز فهم الطالب.

14- أورد الكتاب عدة مواقف تتحدى تفكير الطالب وتحفزه للمشاركة في النشاطات.

15- يوجد تداخل وتكرار غير ضروريين بين هذا الكتاب والكتب الأخرى.

16- الكتاب يخلو من الأخطاء المطبعية (أذكر الأخطاء إذا وجدت)

17- طباعة الكتاب واضحة ومريحة للبصر.

18- الرسوم والوسائل التعليمية الواردة في الكتاب واضحة (إن كانت لك اقتراحات حول هذه الرسوم أذكرها في ورقة منفصلة).

19- (ملاحظات أخرى (أكتبها في ورقة منفصلة).

20- هذا النموذج لتقويم الكتاب المدرسي مناسب (إن كان جوابك لا، اكتب الأسباب خلف النموذج) (توما جورج الخوري، 2008، ص ص. 41-42).

كما تستخدم قوائم التقدير في الكشف عن المشكلات التي يعاني منها المتعلمين ومن أهم هذه القوائم نذكر:

- قائمة موني (R.L.Mooney& al) للمشاكل وهي قائمة معدة لدراسة مشاكل الطلبة في مختلف المراحل التعليمية من اعدادية وثانوية وجامعية، تتكون من 330 بند ا موزعة على

ميادين الصحة، المالية، الحياة الاجتماعية، العلاقات الإنسانية، الدين، الدراسة، الجنس، المنهاج وطريقة التدريس، المستقبل التعليمي والمهني والبيت والأسرة.

تستخدم هذه القائمة من طرف التلميذ وليس الملاحظ، حيث يقوم بقراءة البنود وإذا تطابق محتواها عليه يضع سطرا تحت البند.

ومن الأمثلة على البنود ما يلي:

1-أشعر بتعب في معظم الأوقات

2- وزني أقل من الوزن الطبيعي

3- وزني أكثر مما ينبغي

4-لا أمارس الرياضة بشكل كاف

5-لا أحصل على نوم كاف

6-نقودي لا تكفي لشراء الملابس

7-مساعدة البيت لي قليلة

8-يحصل أصدقائي على نقود أكثر

9-لا أستطيع تدبير أموري المالية

10- بحاجة لعمل جزئي (توما جورج خوري، 2008، ص.39)

على ضوء ما سبق، يمكن القول إن قوائم التقدير في المجال التربوي تستخدم على نطاق واسع في مجالات متعددة منها تقييم منتوج التعلم (المعارف والمهارات للمتعلمين)، تحقيق الأهداف

التعليمية، طرق التعلم، تقويم مهارات المعلمين، طرق التعليم، المناهج التعليمية، بالإضافة إلى الإشراف التربوي والإداري.

مزايا قوائم الصفات:

- تعتبر طريقة فعالة وبسيطة لوصف الشخص وتقويمه
- طريقة مباشرة نسبيا للحصول على أوصاف للشخصية وتوصيف مشكلات الفرد.
- أسهل في تصميمها مقارنة بمقاييس التقدير والاستخبارات، وهذا لا يعني أنها أقل صدقا منها.
- تكون مفيدة بشكل خاص عندما يكون المفحوص متعاونا مستبصرا بذاته.
- استخداماتها متعددة، يمكن استخدامها في المجالات الاكلينيكية والبحثية ودراسة السلوك في مواقف العمل وغير ذلك من المواقف، وأكثر استخداماتها في دراسة الشخصية وتحديد المشكلات الشخصية والاضطرابات الانفعالية.

عيوب قوائم الصفات:

- تعتمد غالبا في الإجابة عنها على التصنيف الكيفي: حيث يطلب من المفحوص بأن يجيب بإحدى الطريقتين: إما أن الصفة تميزه فيضع علامة عليها، أو انها لا تميزه فيتركها. وهذا ما لا يسمح بتقدير كمي لكل صفة على حدا، بل يتحصل الفرد على درجة كلية. على الرغم من أن تطبيقها سهلا ولا يتطلب وقت غير أن تقدير درجاتها امر مجهد ويستغرق وقتا طويلا (الأنصاري، ص ص. 438-490).

المحور الرابع: الاختبارات

تمهيد:

تعتبر الاختبارات التحصيلية من أهم الأدوات القياس التي يعتمد عليها المختص التربوي والمعلم في تقويم مخرجات التعليم.

ولكن قبل الخوض في التعريف بأنواعها، ارتأينا أولا تذكير مفهوم الاختبار بشكل عام والتطرق للتصنيفات المختلفة المتداولة، وذلك اعتبارا أن المختص التربوي قد يستعين باختبارات متعددة لمساعدته على اتخاذ قرارات في المجال التربوي، فعلى المتعلم أن يكون ملما بكل ما يتعلق بالاختبارات في مجالاتها المختلفة.

أولا: الاختبارات وتصنيفاتها

مفهوم الاختبار:

عرف فؤاد أبو حطب الاختبار النفسي بانه: " طريقة منظمة للمقارنة بين الأفراد أو داخل الفرد الواحد في السلوك أو في عينة منه في ضوء معيار أو مستوى أو محك" (فؤاد أبو حطب، السيد عثمان وأمال صادق، 2008، ص.32)

كما يعتبر حسب ما ورد في قاموس (English & English, 1958)" مجموعة من الظروف المقننة أو المضبوطة، تقدم بنظام معين للحصول على عينة ممثلة للسلوك، في ظروف أو متطلبات بيئية معينة، أو في مواجهة تحديات تتطلب بذل أقصى جهد او طاقة، غالبا ما تأخذ هذه الظروف او التحديات شكل الأسئلة اللفظية (بشير معمرية، 2012، ص ص. 103-104).

ومن التعاريف السابقة فإن الاختبار كل اجراء وفق قواعد منظمة ومضبوطة يهدف للمقارنة بين الأشخاص للكشف عن الفروق الفردية أو المقارنة داخل نفس الفرد الواحد للوصول إلى قوانين تحكم السلوك الإنساني.

تصنيف الاختبارات:

يمكن تقسيم الاختبارات على أسس كثيرة منها حسب المحتوى، الغرض، تفسير الدرجات، طريقة التطبيق،....الخ، وسنعرض فيما يلي بعض التصنيفات :

- **اختبارات الأداء الأقصى واختبارات الأداء المميز:** اقترح هذا التصنيف كرونباخ (1990)، حيث ميز بين اختبارات الأداء الأقصى والتي تشمل على اختبارات القدرة والتحصيل، وهي تقيس الحدود العليا لمعارف الطالب وقدراته واختبارات الأداء المميز التي تهتم بقياس السلوك المعتاد للأفراد وخصائصهم.

يمكن تقسيم اختبارات الأداء الأقصى إلى أقسام فرعية:

1. **اختبارات تحصيلية أو اختبارات الاستعدادات:** حيث تهدف الاختبارات التحصيلية تقييم معارف ومهارات الفرد في نطاق محتوى معين ومعرفة مدى تم تحصيله من طرف المتعلم، وعلى عكس ذلك صممت اختبارات الاستعدادات لقياس المهارات والقدرات المعرفية والمعارف التي تجمعت لدى الفرد نتيجة خبراته مدى حياته، وتستخدم عادة للتنبؤ بالأداء المستقبلي سواء كان أكاديميا أو في مجال العمل.

2. **اختبارات موضوعية وذاتية:** وتشير الموضوعية إلى مدى اتفاق القائمين بتقدير درجات اختبار معين أو بمعنى اتباع نفس الطريقة في تقدير استجابات اختبار. فتعتبر الاختبارات التي تشتمل على مفردات تتطلب اختيار الاستجابة (مثل: الاختيار من متعدد، أسئلة الصواب أو الخطأ، أو المزاوجة) والتي يتم تقدير درجاتها وفق مفتاح تصحيح محدد مسبقا، في أغلب الأحيان باختبارات موضوعية، في حين تعتبر اختبارات المقال باختبارات ذاتية باعتبارها تعتمد في تقدير الدرجات على الأحكام الذاتية، وللحد من الذاتية، على القائم بتطوير الاختبارات الذاتية بتصميم مقاييس تقدير وصفية.

3. **اختبارات السرعة واختبارات القوة:** تصنف اختبارات الأداء الأقصى إلى اختبارات السرعة التي تعكس السرعة في الأداء، تكون أسئلتها سهلة نسبيا ولكنها مقيدة بزمن إجابة محدود، مما يمنع المفحوصين من استكمال الإجابة على جميع الأسئلة في الوقت المحدد، واختبارات القوة التي أسئلتها تكون متدرجة الصعوبة، ويعطى الوقت الكافي للإجابة بعض الأسئلة تكون صعبة للغاية، يتوقع أن جميع المفحوصين لن يستطيعوا الإجابة على جميع مفردات الأسئلة إجابة صحيحة. وتصمم اختبارات السرعة والقوة بحيث لا يمكن لأي مفحوص أن يتحصل على درجة نهائية أو تامة، وإذا حدث ذلك، فإن الاختبار فشل في تقييم الحدود العليا للأداء.

- **اختبارات الأداء المميز:** تقيس الاختبارات الأداء المميز تكوينات فرضية مثل الشخصية، الاتجاهات والدافعية، والسلوك والانفعالات والميول، عادة ما نستعين بمصطلح الشخصية ويستعمل في سياق أوسع لتعكس مجموعة كبيرة من الخصائص غير معرفية.

وعند وصف الاختبارات الشخصية يمكن تصنيفها إلى:

1-**الاختبارات الموضوعية للشخصية:** وهي تلك الاختبارات التي تستخدم مفردات تتطلب الإجابة عليها الاختيار (مثل الصواب أو الخطأ) وتقدر درجاتها بطريقة موضوعية بالاستعانة بمفتاح التصحيح الذي تم اعداده مسبقا لتقدير الدرجات.

2- **الاختبارات الإسقاطية للشخصية:** تتضمن عادة الاختبارات الإسقاطية تقديم مواد غير محددة البنية أو غامضة، تثير استجابات لا نهائية عند المفحوصين وبدون قيود. يعاب على هذا النوع من الاختبارات القدر الكبير للذاتية عند تقدير درجات الاستجابة عن الاختبار (محمود علام، 2013، ص ص. 41-45).

أما أنستازي تصنف الاختبارات على النحو التالي:

- **اختبارات النمو العقلي:** تضم مقاييس ستانفورد للذكاء، والاختبارات الجمعية والانجاز، ومقياس وكسلر وبلفيو لذكاء الراشدين، ومقاييس الضعف العقلي.
- **اختبارات القدرات المنفصلة:** تضم مقاييس الاستعدادات والاختبارات المهنية.
- **اختبارات الشخصية:** وتضم استبيانات التقرير الذاتي ومقاييس الميول والاتجاهات والأساليب الإسقاطية وأساليب أخرى لتقدير الشخصية.

وإلى جانب هذه التصنيفات يمكن تقسيم الاختبارات إلى:

- **اختبارات فردية واختبارات جماعية:** تصنف الاختبارات في هذا التصنيف وفق طريقة الاجراء، حيث تطبق الاختبارات الفردية بشكل فردي في مقابلة بين الفاحص والمفحوص، مثل اختبار بينيه لذكاء، بينما تطبق الاختبارات الجماعية على مجموعة من المفحوصين دفعة واحدة، مثل اختبار الذكاء لسعد عبد الرحمن.
- **اختبارات الأداء واختبارات الورقة والقلم:** تصنف الاختبارات في هذه الحالة تبعا لطريقة الأداء، تتطلب الاختبارات الأدائية القيام بعمل ما او حل مشكلة محددة، مثل اختبارات الأداء في القدرة الميكانيكية، في حين تستدعي اختبارات الورقة والقلم تدوين الإجابة في صحيفة الإجابة.
- **الاختبارات اللفظية واختبارات غير لفظية:** تعتمد على استخدام الرموز اللفظية سواء الحروف أو الرموز الرقمية، أما الاختبارات غير اللفظية تقوم على استخدام الأشكال أو الصور، وتستخدم غالبا في حالة المفحوصين الذين لا يجيدون القراءة (بشرى إسماعيل، 2004).
- **اختبارات معيارية المرجع واختبارات محكية المرجع:** يتم تصنيف الاختبارات وفق طريقة تفسير الدرجات المتحصل عليها في الاختبار، حيث تهدف اختبارات المعيارية المرجع بتحديد مقدار المعلومات والمهارات التي تم تحصيلها التلميذ في موضوع ما، وتفسر الدرجات المتحصل عليها في الاختبار بمقارنة أداء التلميذ بأداء المجموعة التي ينتمي إليها

(مثل متوسط أداء المجموعة)، فيعاب على هذه الاختبارات المعيارية المرجع كونها لا تقيس الأهداف التعليمية التي يسعى لتحقيقها المعلم. بينما يقارن أداء الفرد في الاختبارات المحكية بمعايير أو محكات محددة مسبقا، مستقلة على أداء المفحوصين، بحيث تقيس الاختبارات المحكية المرجع مستوى التحكم والاتقان في الأداء (بشير معمرية، 2012؛ صلاح أحمد مراد وأمين علي سليمان، 2005).

ثانيا: الاختبارات التحصيلية:

تمهيد:

يعتبر تقييم التحصيل الدراسي للمتعلم من أهم المشكلات التربوية التي تعترض مهام المعلم والمختصين التربويين على سواء، حيث البحث عن أدوات أو أساليب القياس التي تضمن له الدقة والموضوعية في التقييم هي من أهم الرهانات التي تواجهه

لذا سنحاول في هذا العرض التعرض للاختبارات التحصيلية بشكل عام (الاختبارات التقليدية)، ثم تناول الاختبارات الموضوعية بالتفصيل وهذا بتقديم خطوات بناء اختبار تحصيلي موضوعي.

1-مفهوم الاختبار التحصيلي:

هو الأداة التي تستخدم في قياس المعرفة والفهم والمهارة في مادة دراسية، أو تدريبية معينة، أو مجموعة من المواد (فؤاد أبو حطب وأمال صادق، 2008، ص.467)

وهو أهم أدوات القياس شيوعا في مجال التقويم التربوي، وفي بلادنا يعتبر الوسيلة الوحيدة المستخدمة للمصادقة على المكتسبات من معارف ومهارات وسلوكيات التي تم تحصيلها من طرف المتعلم،

وبذلك تعتبر الوسيلة الوحيدة الحصرية التي تستخدم في اتخاذ القرارات التربوية في الانتقال من سنة دراسية إلى سنة دراسية تليها، وفي التوجيه إلى التخصصات المختلفة.

2-وظائف الاختبارات التحصيلية:

- للاختبارات التحصيلية وظيفة دافعية قوية، فبمعرفة المتعلم بمدى تقدمه في الدراسة، نقاط قوته وضعفه، فإنه سيوجه جهوده لتحسين مكتسباته.
- تزود المعلم بمعلومات حول مدى تقدم المتعلم في التعلم ونقاط قوته وضعفه لتصحيح أو معالجة المشكلات التي تحول دون تحقيق أهداف التعلم، كما تساعده على الحكم على الطرق والأساليب المستخدمة من طرفه في التدريس ومدى تحقق الأهداف التي قام بالتخطيط لها.
- تساعد على تحديد الفروق الفردية بين المتعلمين، وتحديد المستويات التي تتطلبها المراحل الأعلى، وما يتبعها من ترفيع وإعادة.
- تفيد في عملية التوجيه والإرشاد والتصنيف والانتقاء بين المتعلمين (محمود أحمد عمر وآخرون، 2010)

وفي نفس السياق، حدد كل من توكمان وجرولاند (Tuckman & Gronlund, 1985) ستة وظائف أساسية تدفع المعلم لاستخدام الاختبارات التحصيلية، نذكرها فيما يلي:

1 -وظيفة تشخيصيةDiagnostic function :

إن للاختبارات التحصيلية دور في تشخيص مواطن الضعف والقوة لدى الطالب واحتياجاته المختلفة في التعلم بحيث يستخدم المعلم هذه المعلومات والبيانات للتأكد من مدى توفر المكتسبات الأولية التي تسمح له بالبدء في عملية التدريس، فقد يحتاج الطالب لتدريب أولي لمعالجة النقائص لديه.

2-وظيفة وصفية (Descriptive function):

تستخدم الاختبارات للكشف ما تم تعلمه وما لم يتم تعلمه أثناء عملية التدريس، وبذلك فهي تزود المعلم بمعلومات تساعده على اتخاذ قرارات قد تتعلق بتعليم إضافي عند الضرورة، أو إعطاء واجبات منزلية إضافية،..الخ

فهي تشبه في ذلك الوظيفة التشخيصية غير أنها تختلف عنها في أن الوظيفة الوصفية تكون أثناء عملية التدريس وليس قبلها، وتمس جميع الطلبة.

3-وظيفة إعطاء الدرجات (Grading function) :

تسمى اختباراتها بالمسحية، بمعنى أن المعلم عند انتهاء عملية التدريس يقيس أداء الطالب ليقيمه ويعطيه درجة ما، تزود المعلم بمعلومات حول ما تم اكتسابه.

فبخلاف الوظيفة التشخيصية والوصفية واللتان لهما طابع تكويني، فإن وظيفة إعطاء الدرجات لها طابع تجميعي، تساعد فقط في اتخاذ قرارات بشأن نجاح الطالب أو رسوبه (محمود أحمد عمر وآخرون، 2010، ص ص. 382-384).

4-تقويم كيفية التدريس (Evaluation instruction) :

قد يرغب المعلم في تقييم طريقة تدريسه للوحدة التعليمية من خلال أداء الطلاب الجماعي، وهذا النوع من التقويم يعتبر تجميعي غير أن الفرق بينه وبين وظيفة إعطاء نقطة، أنه ينصب على طريقة التدريس، ومع ذلك فإن أداء الطلاب الجماعي أساس هذا التقويم (Tuckman, 1988, p. 2-3)

5-وظيفة تكوينية(Formative function) :

يقدم الاختبار التكويني بشكل دوري أثناء عملية التدريس بهدف مراقبة تقدم تعلم التلاميذ وتقديم التغذية الراجعة باستمرار للمعلم والطالب على السواء لتصحيح ومعالجة النقائص من أجل تحسين أو انجاح عملية التعلم. يتم بناء هذه الاختبارات بطريقة تسمح بتقديم توجيهات تصحيحية بالنسبة للمفردات التي تمت الإجابة عليها بشكل خاطئ.

والجدير بالذكر أن نتائجها نادرا ما تستخدم لإعطاء درجات للطلاب، وعادة ما تستخدم في التشخيص غير أن محور اهتمامها مدى تحقق الأهداف التعليمية (محمود أحمد عمر، 2010)

6-التعزيز الإيجابي:

قد تستخدم الاختبارات التحصيلية من أجل تعزيز ما تعلمه الطالب وتغيير اتجاهاته نحو المادة الدراسية، بحيث يختار المعلم أسئلة أو مفردات الاختبار التي يعرف مسبقا أن طلابه سوف يجيبون عليها إجابة صحيحة.

3-الأسس التي يقوم عليها الاختبار التحصيلي:

- يقيس الاختبار التحصيلي نطاقا سلوكيا واسعا يمكن تحديده، ينبغي أن يستند إلى أهداف تربوية يمكن صياغتها صياغة سلوكية محددة بحيث يمكن قياسها.
- ينبغي أن تحتوي الاختبارات التحصيلية على أنواع الأسئلة والبنود التي تتناسب وقياس نواتج التعلم المرغوبة. فأسئلة المقال تقيس القدرات المعرفية، مثل القدرة على التعبير، تنظيم الأفكار، التمييز، التحليل وتطبيق القواعد على مواقف جديدة، تقيس الأسئلة الموضوعية بدورها مستويات مختلفة من الأهداف التعليمية، سنتناولها لاحقا.
- أن تشمل أسئلة الاختبار جميع المحتويات التي تم تدريسها في المقرر الدراسي، لذا يلجأ المعلم لاختيار عينة من الأسئلة التي تمثل محتويات المقرر بالاستعانة بما يسمى جدول المواصفات، والذي يحدد أهداف التدريس، مواضيع المقرر الدراسي، عدد الأسئلة في الاختبار وتوزيعها وفق المواضيع، وسنتطرق لهذا الموضوع بشكل من التفصيل لاحقا.
- لا تقتصر الاختبارات التحصيلية على قياس المعارف المكتسبة في المادة الدراسية بل تهتم بقياس قدرة الطالب على استخدام وتوظيف المعرفة في مواقف تتطلب التفكير وحل المشكلات، كما تقيس المهارات العلمية المختلفة، كمهارة الطالب في تطبيق ما تعلمه في مادة معينة، كاستعمال المجهر، تشريح حيوان، تشغيل جهاز ما،.....الخ.

- ينبغي أن يراعي الاختبار التحصيلي الفروق الفردية بين المتعلمين ويكون مناسبا للأفراد الذين سيطبق عليهم الاختبار، يظهر هذا المشكل في حالة قيام جهات أخرى لإعداد الاختبار بدلا من المعلم (محمود أحمد عمر وآخرون، 2010، ص ص. 97-98)

4-أنواع الاختبارات التحصيلية:

الاختبارات التحصيلية بشكل عام يمكن تصنيفها إلى نوعان، أولا الاختبارات التقليدية التي يعدها المعلم بنفسه، مثل الاختبارات الشفهية والمقالة، وثانيا الاختبارات التحصيلية المقننة والتي تعدها الجهات المختصة لتقويم تحصيل التلاميذ.

4-1. الاختبارات التقليدية:

4-1.1. الاختبارات الشفهية:

ربما تكون الاختبارات الشفهية أقدم طريقة استخدمت في تحديد مدى استيعاب المتعلمين للدروس، فيقال أن سقراط استعمل الاختبارات الشفوية منذ القرن الرابع قبل الميلاد للوقوف على مستوى مستمعيه لكي يبني تعليمه لهم على أساس خبرتهم الماضية.

تهدف هذه الاختبارات إلى التعرف على مدى اتقان التلاميذ للمادة الدراسية بمعزل من قدرات الكتابة والتعبير، وفي تقييم قدرة المتعلم على القراءة والنطق السليم والحكم على مدى استيعاب المفاهيم، كما تستخدم للكشف عن أخطاء المتعلمين وتصحيحها في الحال.

تعتبر الاختبارات الشفهية ضرورية في المراحل الدراسية الأولى من الدراسة الابتدائية، كما يمكن استخدامها في مناقشة رسائل الدكتوراه والماجستير.

لاقت الاختبارات الشفهية جملة من الانتقادات وخاصة فيما يتعلق بموضوعيتها، نذكر من عيوبها:

- تأثر المعلم بالفكرة المسبقة التي كونها عن المتعلم أو تأثر المعلم بحالته المزاجية والتي قد تختلف من فترة إلى أخرى.
- تأثر المتعلم بموقف الامتحان، فقد يرتبك المتعلم نتيجة الخوف أو الخجل الذي قد يعتريه، وليس لعدم معرفته بالمادة.
- عدد الأسئلة التي يمكن صياغته لمواضيع محددة يكون محدودا نوعا ما، قد يسمح للمتعلمين الاستفادة من إجابات زملائهم الذين تقدموا للإجابة قبلهم.
- تدخل الصدفة وعدم توفير نفس الحظوظ للمتعلمين في الإجابة، فقد يطرح على متعلم سؤالا سهلا قياسا بأسئلة توجه لمتعلمين آخرين.
- يتطلب وقتا طويلا لإجرائها مما لا يتيح للمعلم إخضاع جميع التلاميذ للاختبار في الوقت المحدد.
- صعوبة تحلي المعلم بالدقة والموضوعية طيلة فترة التقويم، حيث سرعان ما تتأثر أحكامه وتقديراته بالذاتية لعوامل متعددة.
- عدم شمولية هذه الاختبارات للمواضيع المختلفة التي يتضمنها المقرر الدراسي والتي يراد قياسها.

أما عن مزايا الاختبارات الشفهية نذكر:

- عدم إمكانية الغش فيها.

- توفر فرصة للتعلم من خلال المناقشات التي تسمح بها
- توفر فرصة للتفاعل المباشر بين المعلم والمتعلم، بدلا من أن يتعامل معه من خلال أسئلة مكتوبة.
- تكون مناسبة لتقويم بعض الصفات والأهداف التي لا يمكن تقويمها إلا من خلالها.

(سوسن الجلبي،2005، ص ص.215-217)

4-1.2. اختبار المقال (التحريرية):

يطلق على هذا النوع من الاختبارات عدة تسميات منها اختبار المفتوح، التحليلي، التفسيري والتحريري...الخ، وسميت بالتقليدية باعتبارها من أنواع الاختبارات التحصيلية الأولى ظهورا، وهذا لا يعني أنها لم تعد تصلح لقياس تحصيل التلاميذ، بل على العكس من ذلك تعتبر في بعض المواقف الوسيلة الوحيدة لقياس بعض القدرات أو السمات، مثل قياس القدرة على التحليل والتركيب والقدرة على التعبير وابداء الرأي.

وعرف لوجندر(Legendre) اختبارات المقال بأنها" عبارة عن سؤال يتناول مسائل معقدة تتطلب جهود تفكير وتنظيم" (غريب العربي، دون سنة، ص. 75).

هذا النوع من الاختبار التحصيلي يطلب فيه من الممتحن أن يبني اجابته لوحده بطريقة مستقلة أو شبه مقيدة ونصف محددة، فالمتعلم لديه الحرية النسبية في تقدير كيفية أو منهجية تناول المشكلة ومعالجتها واختيار المعارف والخبرات وكيفية تمييز وتنظيم اجابته (محمد رضا البغدادي، 1983، ص. 185).

وهذا النوع من الاختبارات أكثر استخداما من طرف المعلمين في منظومتنا التربوية لسهولة اعداده وبساطة اجراءاته، غير أنه عرف انتقادات كثيرة نتطرق لبعضها فيما يلي:

عيوب اختبارات المقال:

وجهت لهذا النوع من الاختبارات انتقادات عديدة، نذكر منها تلك التي اسفرت عليها الدراسة النقدية التي قام بها هنري بيرون(Henry Pieron) التي أوضح من خلالها ضعف هذه الاختبارات للاعتبارات التالية:

- عدم موضوعية التصحيح وتدخل عوامل الذاتية منها أثر الهالة، سوء التقدير، تدخل ذاتية المصحح وأفكاره واتجاهاته.
- لا تغطي كل البرنامج، حيث أن عدد أسئلتها محدود، وبالتالي لا تقيس كل الأهداف المحددة.
- صعبة التصحيح خاصة إذا كان الاختبار طويلا وعدد الممتحنين كبيرا والخط غير واضحا.
- صعبة التحقيق مع المتعلمين ضعاف اللغة وذوي المستويات الدنيا.
- إمكانية الخروج عن الموضوع من طرف الممتحنين وخاصة إذا لم يكن واضحا ودقيق اللغة.
- لا تمكن التلميذ من تقويم ذاته.

محاسن الاختبارات التحريرية:

رغم الانتقادات التي لاقتها غير أنها تبقى الوسيلة الوحيدة التي لا يمكن الاستغناء عنها لقياس بعض السمات لدى المتعلم، ومن محاسنها ما يلي:

- تنمية الثروة اللغوية لدى المتعلم وتشخيص القدرة التعبيرية لديه.
- تساعد على معرفة مدى قدرة المتعلم على حل المشكلات.
- تمكن من قياس قدرة المتعلم على الفهم، التحليل، التركيب، النقد والتقويم.
- سهلة الاعداد والتحضير (غريب العربي، دون سنة، ص. 76)

الاختبارات التحصيلية الموضوعية:

لديها تسميات كثيرة منها المحدودة، المغلقة، المقننة، وجاءت نتيجة الدراسات النقدية للاختبارات التقليدية وظهرت بإنشاء هنري بيرون لفرع جديد في علوم التربية أطلق عليه علم التباري (Docimologie) وهو علم يهتم بالامتحانات.

يتضمن هذا النوع من الاختبارات إجابات محددة تتقلص فيها احتمالات الأحكام الذاتية مثل اختيار إحدى الإجابات من بين بدائل عديدة للسؤال الواحد.

كما تكون إجابته على هذا النوع من الاختبارات صحيحة أو خاطئة ولا وجود للإجابة القريبة من الصحة، فالاختبار التحصيلي الموضوعي هو: "مجموعة أسئلة ذات الإجابات القصيرة تساعد على قياس الاستدعاء والتعرف، ويمكن الإجابة عليها في الوقت المخصص للدرس" (محي الدين توق وعبد الرحمان عدس، 1984، ص.337).

يتضمن الاختبار الموضوعي التحصيلي عدة أنماط من الأسئلة، نذكر منها ما يلي:

- **اختبار الصواب والخطأ:** يعتبر هذا النوع من الأسئلة الأكثر شيوعا نظرا لسهولة إعداده وقدرته على تغطية المادة، وتشتمل الأسئلة الصحيح والخطأ على عبارة تقريرية يمكن أن

تتألف من جملة واحدة أو أكثر، بحيث يطلب من المتعلم أن يقرر ما إذا كانت الإجابة صحيحة او خاطئة، من مزايا هذا النوع من الأسئلة أن لها القدرة الكبيرة على تغطية المادة، كما أنها لا تحتاج إلى وقت طويل للإجابة، وهي سهلة التصحيح ويمكن تصحيحها بطريقة موضوعية وبسرعة. أما عن أهم عيوبها لا تصلح في الموضوعات المعقدة، نسبة التخمين فيها عالية (سامي محمد ملحم، 2017)

- **اختبار اختيار من المتعدد:** تعد أسئلة الاختيار من متعدد من أهم وأفضل أنواع الأسئلة الموضوعية لأنها تشتمل على معظم أنواع الأسئلة الموضوعية الأخرى (الاختيار من بديلين، المزاوجة، الترتيب، الإجابة القصيرة، التفسيرية) في صورة الاختيار المتعدد، كما يمكن قياس جميع المستويات العقلية إذا أحسن صياغتها. وتتألف أسئلة الاختيار من متعدد في أيسط صورة من مقدمة يقال لها (متن السؤال) قد تكون في صيغة عبارة استفهامية أو حتى شكل عبارة خبرية ناقصة يحدد فيها المعلم المشكلة، يلي المقدمة قائمة من الحلول المحتملة للمشكلة المقدمة وهذه الحلول تسمى بدائل، وأحد هذه البدائل هو الإجابة الصحيحة أو الأكثر صحة. وباقي البدائل عبارة عن مشتتات أو بدائل غير صحيحة (أمين علي محمد سليمان، 2010، ص.279)
- **اختبار المطابقة أو المزاوجة:** يستخدم هذا النوع من الأسئلة بشكل كبير في المرحلة الدراسة الأساسية، تتكون أسئلة المزاوجة من عمودين متوازيين، يمثل العمود الأول فقرات أو مثيرات تسمى المقدمة تتكون من مفردات أو رموز أو أعداد، بينما يمثل العمود الثاني الاستجابات

التي تتكون من كلمات أو رموز أو أعداد أو جمل قصيرة، ولكل مقدمة في العمود الأول إجابة صحيحة تقابلها في العمود الثاني. ومن أهم مزايا هذا النوع من الأسئلة تغطية مواضيع متعددة لاحتوائها على كمية كبيرة من الحقائق المتجانسة التي يمكن قياسها في فترة زمنية قصيرة، كما هي سهلة التطبيق، اما عن عيوبها فهي تقيس حقائق تعتمد على الحفظ، يقتصر استعمالها على المعلومات التي تساعد على بيان العلاقة (سامي محمد ملحم، 2005)

- **اختبار ملأ الفراغ أو التكملة:** تشبه بشكل كبير اختبارات الإجابة القصيرة، الفرق بينهما أن المتعلم يكمل الإجابة الناقصة في حالة الإجابة عن أسئلة التكملة بينما يستحضر إجابة قصيرة لسؤال يحتوي على عبارة كاملة. وما يميز هذا النوع من الأسئلة على بقية الأسئلة الموضوعية أنه يقل فيها التخمين، وتمتاز عن الأسئلة المقال بأنها أكثر موضوعية وأن تصحيحها أسهل. أما عن عيوبها تتطلب جهدا وقتا في تصحيحها نظرا لتعدد الإجابات
- **اختبار الترتيب:** تقيس هذا النوع من الأسئلة نواتج التعلم في المستوى البسيط: مستوى التذكر (مستوى المعرفة) سواء كانت معرفة المصطلحات والحقائق والمعلومات والقوانين والنظريات وأيضا مستوى الفهم (الاستيعاب) وقد تمتد لقياس مستوى التفكير التقاربي. ويشتمل هذا النوع من الأسئلة على مجموعة من وحدات المعلومات قد تكون في صورة كلمات أو جمل لها معنى أو صورة رموز او أشكال، جميعها غير مرتبة في صورة منطقية والمطلوب إعادة ترتيبها في شكل منطقي طبقا لأحداثها (أمين علي محمد سليمان، 2010، ص. 261).

والجدير بالذكر أنه ينصح عند اعداد اختبار تحصيلي موضوعي الجمع بين أنماط مختلفة من الأسئلة، وفي هذا السياق حددت أدناه سوسن الجلبي النسب التقريبية للأنواع المختلفة من الأسئلة التي يفضل أن يحتوي عليها الاختبار:

- ✓ الإجابة القصيرة (محدودة بسطور قليلة) بين 20-25%من الاختبار الكلي.
- ✓ الاختيار من البدائل المتعددة ما بين 20-25 %
- ✓ الخطأ والصواب ما بين 10-15 %
- ✓ تكملة الجمل أو ملأ الفراغات ما بين 10-15 %
- ✓ المزاوجة بين الكلمات والمصطلحات ما بين 10-15 %

(سوسن الجلبي،2005، ص.222)

مزايا الاختبارات التحصيلية الموضوعية:

تتميز هذا النوع من الاختبارات بجملة من الخصائص الجيدة تناولها عبد المجيد النشواتي فيما يلي:

- سهولة التطبيق واستخراج النتائج وتفسيرها.
- تخلو من العوامل الذاتية في المجال الإيجابية أو التصحيح، حيث لا يختلف اثنان في تقدير علامة الإجابة الواحدة.
- توظف مجموعة كبيرة من الأسئلة المختلفة في قياس القدرات العقلية، قد يتجاوز عدد الأسئلة مائة سؤال، وهذا يمكننا من تغطية البرنامج ككل.

- تستعمل مع الفئات التي تعاني نقصا في اللغة خاصة مع المستويات التعليمية الدنيا.
- سهلة التصحيح بالنسبة للمدرس كما أنها تمكن التلميذ من تقويم ذاته.
- ارتفاع نسبة مصداقية نتائجها وهذا ينعكس إيجابا على القرارات والأحكام المتخذة من خلال معطياتها (عبد المجيد نشواتي، 1985، ص.618).

عيوب الاختبارات التحصيلية الموضوعية:

رغم الميزات الإيجابية التي تتمتع بها غير أن الاختبارات التحصيلية الموضوعية بدورها لا تخلو من العيوب نذكر منها:

- تحتاج لوقت وجهد كبيرين لبنائها واعدادها.
- لا تسمح بقياس مختلف الجوانب المتعلقة بشخصية المتعلم كالقدرة على تنظيم الأفكار وتوظيفها، التحليل والتركيب والتفكير الناقد.
- لا تزود المعلم بالمعلومات حول كيفية أو طريقة الوصول إلى الاجابة.
- تمكن المتعلم من الاعتماد على التخمين أو الغش في الإجابة.
- تتطلب مهارات خاصة لبنائها قد لا نجدها عند المعلم (غريب العربي، دون سنة)

خطوات بناء الاختبار التحصيلي الموضوعي:

يجمع معظم المختصين على وجود قواعد (خطوات) أساسية يجب اتباعها في اعداد الاختبار التحصيلي لقياس نواتج التعلم نلخصها فيما يلي:

الخطوة الأولى: تحديد الغرض من الاختبار:

والمقصود به تحديد الغرض من استخدام الاختبار، حيث قد يستخدم لغرض تشخيص المكتسبات المسبقة لدى المتعلم، أو التعرف على الصعوبات التي يعاني منها للتخطيط لعملية التعليم والتعلم، أو قد يستخدم لأغراض تكوينية (تقويم بنائي أو تكويني) لمعرفة مدى تحقق الأهداف التعليمية بالنسبة للمعلم والتقدم في التعلم بالنسبة للمتعلم، كما يمكن استعماله لغرض المصادقة على المكتسبات التعلمية من خلال منح درجات أو شهادات للمتعلم (تقويم ختامي أو نهائي) (سوسن جلبي،2005).

الخطوة الثانية: تحديد الأهداف التعليمية من الاختبار:

يتطلب تحديد الأهداف التعليمية التي يرغب المعلم في تحقيقها لدى التلاميذ بعد الانتهاء من المقرر الدراسي أن يحدد أولا مجال الأهداف التعليمية (العقلي / المعرفي، الانفعالي/ الوجداني، والمهارى/النفس حركي) ثم يضعها بصورة إجرائية.

ومن أجل ذلك، لا بدا من توفر شرطين يتمثلان في:

- ✓ التعرف على أنواع (المجالات) الأهداف التعليمية ومستوياتها.
- ✓ صياغة الهداف التعليمية للمادة بعبارات سلوكية تعكس نواتج تعليمية يمكن ملاحظتها وقياسها.

الخطوة الثالثة: تحليل محتوى المقرر الدراسي وتعيين الوزن النسبي لمحتوياته (المواضيع التعليمية):

مسح لجميع المواضيع المتضمنة في المقرر الدراسي لغرض تحديد المواضيع والفقرات التي يجب أن يغطيها الاختبار حتى تتوفر فيه سمة الشمولية (صدق المحتوى).

كما ينبغي تحديد الوزن النسبي للأسئلة أو المفردات في الاختبار بشكل يعكس أهميتها في المقرر الدراسي من أجل تحقيق شرط تمثيل المفردات أو الأسئلة لمحتويات المقرر الدراسي (صدق المحتوى) (سوسن الجلبي،2005؛ صلاح أحمد مراد وآخرون، 2005).

الخطوة الرابعة: إعداد جدول المواصفات:

يعتبر جدول المواصفات المخطط التفصيلي الذي يربط العناصر الأساسية للمحتوى بمجالات التقييم، كما يعتبر الأداة التي تسمح بالتحقق من صدق المحتوى من خلال توفر شرطي الشمولية والتمثيل في الاختبار. "والشمولية تعني يجب أن تكون فقرات الاختبار تشتمل(تغطي) على جميع مكونات المحتوى من أهداف. أما التمثيل فيعني أن تكون فقرات الاختبار تمثل عينة المحتوى تمثيلا صادقا لجميع جوانب التحصيل" (سوسن الجلبي،2005، ص ص. 235-236).

ولبناء جدول المواصفات نتبع الخطوات التالية:

- تحديد المواضيع (الوحدات أو العناصر) التي يتضمنها المقرر الدراسي التي سيتم ترجمتها على شكل أسئلة قابلة للقياس والملاحظة، بإعداد قائمة للمواضيع الأساسية والفرعية التي يتضمنها المقرر الدراسي.

- تحديد الأهداف التعليمية للمادة الدراسية التي يسعى المعلم لمعرفة مدى تحققها (نواتج التعلم).
- تحديد نسبة التركيز (الأهمية أو الوزن النسبي) لكل جزء في المادة الدراسية وذلك من خلال تحديد عدد الحصص الخاصة بكل جزء ومقابلتها بعدد الحصص الشاملة للمادة الدراسية وهذا من خلال المعادلة التالية:

$$\text{نسبة التركيز} = \frac{\text{عدد حصص الوحدة الدراسية}}{\text{عدد حصص جميع المادة الدراسية}} \times 100\%$$

ومن المحكات الأخرى التي يمكن استخدامها في تحديد الأهمية النسبية لعناصر المحتوى عدد الصفحات التي يشغلها كل موضوع على حدا بالنسبة للعدد الإجمالي لصفحات الكتاب، أو الزمن المستغرق لتدريس الموضوع على حدا مقابل الوقت المخصص للمادة، أو آراء المتخصصين في المادة وهو أفضل المحكات.

- تحديد نسبة أهمية الأهداف وفق مستوياتها.
- تحديد العدد الإجمالي للأسئلة المراد وضعها للاختبار.
- تحديد عدد الأسئلة التي ينبغي صياغتها لكل جزء (أو وحدة) تبعا لأهميتها وهذا من خلال المعادلة التالية:

عدد الأسئلة لكل جزء (أو وحدة): عدد الأسئلة الاجمالي xنسبة التركيزx نسبة الهدف (سوسن الجلبي، 2005؛ أحمد مراد وآخرون،2005)

مثال (01):

جدول (05): الأوزان النسبية لعناصر المحتوى ومستويات الأهداف في المادة

الأهداف ومستوياتها	المجال المعرفي %			المجال المهارى	الوزن النسبي
عناصر المحتوى	تذكر	فهم	تطبيق	النفس حركي %	لعناصر المحتوى%
الموضوع 1	20	15	10	5	50
الموضوع 2	12	8	5	5	30
الموضوع 3	8	7	5	–	20
الوزن النسبي لمستويات الأهداف	40	30	20	10	100

(صالح أحمد مراد وآخرون، 2005، ص. 147)

مثال (02):

جدول رقم (06) : يوضح كيفية تحديد عدد الأسئلة وفق أهمية المواضيع والأهداف

الأهداف ومستوياتها	المجال المعرفي %			المجال المهارى	الوزن النسبي
عناصر المحتوى	تذكر	فهم	تطبيق	النفس حركي %	لعناصر المحتوى%
الموضوع 1	20	15	10	5	50 (15 سؤال)
الموضوع 2	12	8	5	5	30 (09 أسئلة)
الموضوع 3	8	7	5	-	20 (06 أسئلة)
الوزن النسبي لمستويات الأهداف	40 12) سؤال)	30 09) أسئلة)	20 06) أسئلة)	10 (03 أسئلة)	100 (30 سؤال)

(صالح احمد مراد وآخرون، 2005، ص. 149)

الخطوة الخامسة: صياغة أسئلة الاختبارات التحصيلية

بعد تحديد المواضيع الأساسية والفرعية والأهداف التعليمية ومستوياتها والأوزان النسبية لها وتحديد عدد الأسئلة الإجمالي للاختبار والخاصة بكل جزء، كما أشرنا له في الخطوات السابقة، يقوم المختصون في مراكز البحوث التربوية أو المعلمون بصياغة مجموعة من الأسئلة للاختبار تزيد عن

الحاجة الفعلية، ولا تكون الأسئلة جميعها بالضرورة صالحة ومستوفية للشروط، بل يختار معدوا الاختبار عددا منها الأنسب وأكثر ملاءمة لقياس الأهداف المرجوة من تعلم المواضيع المقرر.

(سوسن الجلبي والجدير بالذكر على المختص أن يكون على دراية بعيوب ومزايا الأنواع المختلفة للأسئلة والتي سبق أن أشرنا إلى بعضها فيما سبق، ليختار أنواع الأسئلة التي تتوافق والأهداف المقاسة، كما عليه أن يتدرب على صياغة الأسئلة بأنواعها وعلى بعض الأساليب الإحصائية اللازمة لتحليل الإجابات واتخاذ قرارات فيما يخص أسئلة الاختبار من حذف وتعديل.

الخطوة السادسة: مراجعة أسئلة الاختبار وإخراج الصورة الأولية للاختبار:

في هذه المرحلة تتم مراجعة الأسئلة بعد تجميعها وكتابتها ووضعها في صورة اختبار، يتم تعديل حذف أو إضافة أسئلة عند الحاجة، ثم ترتيب الأسئلة بشكل يسهل قراءتها وفهمها من طرف المتعلمين لتفادي الاستفسارات المتعلمين حولها. وفي هذا السياق هناك عدة طرق لترتيب الأسئلة، منها حسب نوع المفردة، الصعوبة أو المحتوى.

وفي نهاية المطاف، اخراج الصورة الأولية للاختبار من حيث الطباعة (الشكل العام) ووضع التعليمات اللازمة والمحددة لكيفية الإجابة المطلوبة، وكذلك توزيع الدرجات بشكل مفصل تسهل عملية التصحيح بأكبر قدر ممكن من الدقة والموضوعية (سوسن الجلبي،2005؛ صالح أحمد مراد،2005).

الخطوة السابعة: التجريب على عينة صغيرة

التجريب على عينة صغيرة بغرض التأكد من (وضوح التعليمات، ملاءمة الزمن) ثم المعالجة الإحصائية لبنود الاختبار بغرض حساب:

- معامل السهولة والصعوبة لكل سؤال
- معامل التمييز
- فعالية المشتتات

الخطوة الثامنة: اخراج الاختبار في صورته النهائية

بعد التأكد من وضوح الأسئلة والتعليمات الخاصة بالاختبار والتأكد من جودة الأسئلة، يتم اعداد الصيغة النهائية للاختبار من أجل الاستخدام (صالح أحمد مراد وآخرون، 2005، ص. 143)

المحور الخامس: دراسة الخصائص السيكو مترية للاختبارات والمقاييس (الصدق والثبات)

مقدمة: يعتبر الاختبار من أهم أدوات جمع البيانات حول الظواهر النفسية والتربوية في علم النفس، ولا شك أن الفضل الكبير لانتشاره يرجع إلى جهود الباحثين في مجال القياس النفسي، حيث أولت أهمية كبيرة لدراسة الخصائص السيكو مترية للاختبارات. ففي هذا الفصل سنحاول التطرق إلى الشروط التي يجب أن يتمتع بها الاختبار من أجل الوثوق في نتائجه وسنخص تناول الصدق والثبات بشكل من التفصيل.

1-مدخل عام لشروط الاختبار الجيد:

الاختبار الجيد هو باختصار هو الذي يصلح للغرض الذي وضع من أجله وهذا يتم بالشروط الآتية:

يذكر بلام ونايلر (Blum & Nyler,1968) ثلاث شروط للاختبار الجيد تتمثل في:

- أن تكون عينة السلوك ممثلة بدرجة كافية للسلوك الذي يراد قياسه حتى تمكننا من التعميم والتنبؤ.
- أن يكون الاختبار مقننا ونقصد به حساب معاملات صدقه وثباته على عينة واسعة وتوحيد التعليمات وإجراءات التطبيق والتصحيح.
- أن تكون قيمة الصدق والثبات عالية وأن تكون له معايير خاصة. ونقصد بالصدق أن يقيس الاختبار ما وضع من أجله، أما الثبات فيعني استقرار الدرجات التي يحصل عليها نفس

الفرد إذا أعيد اجراء الاختبار، بينما يشير المعيار إلى قيم تصف أداء مجموعات متعددة على اختبار معين. (بشير معمرية، 2011، ص. 110)

أولا: ثبات الاختبار

مفهومه: يدل على ما إذا كانت الأداة جيدة وصالحة للقياس

وهو يعني الاستقرار، فلو كررت عملية القياس مرة أخرى لأظهرت شيئا من الاستقرار. (سامي محمد ملحم، 2000، ص. 280)

كما يقصد بالثبات حسب (Gregory, 1992 ; Mehrens& Lehman,1978) درجة الاتساق في عمليات القياس التي تجري لشيء من الأشياء، أما (Anastasie ; 1982) فإنها تعرف الثبات بأنه الاتساق في الدرجات التي يحصل عليها نفس الأفراد بتطبيق نفس الاختبار (عبد الرحمن بن سليمان الطريري، 1997، ص.171).

وهو مدى الاتساق بين البيانات التي تجمع عن طريق اعادة تطبيق نفس المقاييس على نفس الأفراد أو الظواهر وتحت نفس الظروف أو ظروف مشابهة إلى أكبر قدر ممكن. (سامي محمد ملحم، 2000، ص. 280).

كما هو عبارة عن درجة الاتساق أو التجانس بين نتائج مقياسين في تقدير صفة أو سلوك ما، في ضوء ذلك يتوقع أن تكون درجات الفرد ثابتة (موسى النبهان،2004، ص. 229)

وهو درجة الاتساق أو الاستقرار بين مقياسين أو اختبارين من نفس النوع. (قاسم علي الصراف، 2002، ص.189).

وعمليا يعرف (Crocker&Algina, 1986) الثبات بأنه درجة الاتساق التي يكون عليها درجات الإنحرافية أو المعيارية عندما يتم تطبيق الاختبار على نفس الفرد عدة مرات أو عندما تطبق صور متعددة من الاختبار.(عبد الرحمن بن سليمان الطريري، 1997، ص.171).

أنواعه وطرق تقديره:

الثبات بطريقة إعادة الاختبار:

تتلخص هذه الطريقة في تطبيق الاختبار على مجموعة من الأفراد ثم إعادة تطبيق نفس الاختبار على نفس الأفراد بعد مضى فترة زمنية تتراوح بين 4 إلى 8 أسابيع، ثم يحسب معامل الارتباط بيرسون بين درجات الأفراد في التطبيقين. ويحبذ أن يكون التطبيق الثاني في ظروف مشابهة لظروف التطبيق الأول من حيث المكان والزمان وشروط الاجراء.

يعاب على هذه الطريقة:

- تدخل عامل الذاكرة، التعلم والتدريب وتأثيره على النتائج عند اعادة الاختبار بعد فترة زمنية قصيرة.
- حصول تغيرات ناتجة عن النمو على الأفراد عندما يتم اعادة الاختبار بعد فترة زمنية طويلة.

- في بعض الأحيان قد يصعب اجراء التطبيق الثاني. (سعد عبد الرحمن، 1998، ص.166)

مثال تطبيقي: ثبات بطريقة إعادة تطبيق الاختبار

أجري اختبار مكون من 20 سؤالا لعينة حجمها 12 تلميذا وسجلت علاماتهم (س1) ، وبعد مدة زمنية قدرت بشهرين أعيد الاختبار نفسه على المجموعة ذاتها تحت ظروف مشابهة ، فتم تسجيل درجاتهم (س2)،أحسب معامل الاستقرار.

جدول رقم (07): مثال تطبيقي ثبات بطريقة إعادة تطبيق الاختبار

طلبة	1	2	3	4	5	6	7	8	9	10	11	12
تط1	13	10	9	15	14	10	11	12	8	9	10	11
تط 2	15	9	10	14	13	11	13	13	11	9	14	12

المرجع: (سامي محمد ملحم، 2017، ص.303)

1- حساب المجاميع التي نحتاج لها لحساب معامل الارتباط

جدول رقم (08): حساب المجاميع

X	Y	X. Y	X^2	Y^2
13	15	195	169	225
10	9	90	100	81
9	10	90	81	100
15	14	210	225	196
14	13	182	169	169

10	11	110	100	121
11	13	142	121	169
12	13	165	144	169
8	11	88	64	121
9	9	81	81	81
10	14	140	100	169
11	12	132	121	144
132	144	1617	1502	1772

$$\sum X = 132 \quad \sum Y = 144 \quad \sum X.Y = 1617 \quad \sum X^2 = 1502 \quad \sum Y^2 = 1772$$

نقوم بالتعويض في المعادلة:

$$r = \frac{N\sum XY - \sum x.\sum y}{\sqrt{\left(N\sum x^2 - (\sum X)^2\right)\left(N\sum Y^2 - (\sum Y)^2\right)}} = \frac{12X1617 - 132X144}{\sqrt{(12X1502 - 132^2)(12X1772 - 144^2)}}$$

R=0.70

وهي قيمة ثبات مقبولة

ثبات التكافؤ:

يتوجب في هذه الطريقة اعداد صورتين متكافئتين للاختبار وتطبيقهما على نفس الأفراد. ونقصد بالتكافؤ تساوي في عدد ونوع البنود وطريقة الاجابة والتصحيح والزمن المحدد للإجابة والتعليمات إضافة إلى تساوي معاملات الصعوبة والتمييز ومعاملات الارتباط بين البنود والتباين والمتوسط الحسابي والانحراف المعياري (ليلى السيد فرحات، 2012، ص.159) و(سعد عبد الرحمن،1998، ص. 167)

يسمى الثبات بهذه الطريقة بمعامل التكافؤ عندما يجرى التطبيقين للصورتين المتكافئتين للاختبار بعد فاصل زمني قصير أو في نفس الوقت ويسمى بمعامل الاستقرار عندما تطبق الصورتين بعد فاصل زمني مختلف. (موسى النبهان، 2004، ص.240)

يتمتع ثبات التكافؤ بالإيجابيات التالية:

- يصلح في اختبارات الذاكرة والتحصيل
- يلغي أثر التدريب والتذكر (أمين علي محمد سليمان، 2010، ص. 573)

أما عن عيوبه:

- في بعض الأحيان لا يتوفر الوقت والظروف لإجراء التطبيق الثاني (سامي محمد حلمي، 2000، ص.282؛ بشرى إسماعيل،2004)
- صعوبة تصميم اختبارين متكافئين

- التعب والملل اللذان قد يتعرض لهما التلميذ. (مصطفى القمش، 2008، ص. 114).

مثال تطبيقي: تقدير الثبات بطريقة الصورة المكافئة

طبق اختبار تحصيلي على عينة من 10 طلبة، وبعد مرور خمسة أسابيع تم تطبيق على نفس العينة صورة مكافئة للاختبار، فكانت نتائج التطبيقين على النحو التالي:

جدول رقم (09): مثال تطبيقي في تقدير الثبات بطريقة الصورة المكافئة

رقم الطالب	1	2	3	4	5	6	7	8	9	10
التطبيق 1	7	4	9	8	7	6	6	5	8	6
التطبيق 2	5	6	8	8	6	7	6	5	7	7

المرجع: (سامي محمد ملحم، 2017، ص.308)

1-أحسب قيمة الثبات

2-فسر قيمة الثبات

3-كيف نسمي معامل الثبات في هذه الحالة؟

الجواب :

لتقدير قيمة الثبات باستخدام الصورة المكافئة للاختبار يتوجب حساب معامل الارتباط لبيرسون بين نتائج تطبيق الاختبار التحصيلي ونتائج الصورة المكافئة للاختبار بتطبيق المعادلة التالية:

$$R = \frac{N\sum XY - \sum x.\sum y}{\sqrt{\left(N\sum x^2 - (\sum X)^2\right)\left(N\sum Y^2 - (\sum Y)^2\right)}}$$

أولا: حساب المجاميع التي نحتاجها في المعادلة

الجدول رقم (10): حساب المجاميع

Y^2	X^2	X.Y	Y	X	N
25	49	35	5	7	1
36	16	32	6	4	2
64	81	72	8	9	3
64	64	64	8	8	4

36	49	42	6	7	5
49	36	42	7	6	6
36	49	42	6	7	7
25	25	25	5	5	8
49	64	56	7	8	9
49	36	42	7	6	10
439	469	444	65	67	Somme

نتحصل على:

$$\sum X = 67 \quad \sum Y = 65 \quad \sum X.Y = 444 \quad \sum X^2 = 433 \quad \sum Y^2 = 469$$

بالتعويض في المعادلة :

$$R = \frac{10X444 - 67X65}{\sqrt{(10X469 - 67^2)\ (10X433 - 65^2)}}$$

$$R = \frac{4440 - 4355}{\sqrt{(4690 - 4489)\ \ (4330 - 4225)}}$$

ثانيا: $R = \frac{85}{\sqrt{201X105}} = 0{,}58$

تفسير قيمة الثبات

قيمة الثبات مقبولة إلى حد ما ولكن يستحسن الرفع من قيمة الثبات بمعالجة البنود أو زيادة عدد البنود.

ثالثا: يسمى معامل الثبات في هذه الحالة بثبات الاستقرار والتكافؤ.

ثبات الاتساق الداخلي:

كثيرا ما تعترضنا عوائق في ميدان البحث تلزمنا بتبني اختيارات منهجية معينة فقد يصعب علينا مثلا التعامل مع العينة مرة أخرى نظرا لكون الاختيار كان عرضيا أو ايجاد صيغة مكافئة للاختبار. ففي هذه الحالة يتحتم علينا اجراء الاختبار مرة واحدة. فيقوم الباحث بإجراء المقارنة الداخلية لأجزاء الاختبار من خلال بنوده والتأكد من مدى وجود الاتساق والتجانس الداخلي ويتم ذلك من خلال اتباع الطرق الآتية:

✓ **الثبات بطريقة التجزئة النصفية:**

يتم تقدير الثبات من خلال تقسيم الاختبار إلى نصفين يفترض أنهما متكافئين، والطريقة البسيطة لتقسيم الجزأين تتمثل في اختيارالبنود بالتناوب بين العبارات الفردية والزوجية لتشكل كل منها اختبار فرعي مستقل بذاته. ثم حساب معامل ارتباط بيرسون (الذي تم تناول معادلته في المحاضرة السابقة) بين نتائج تطبيق الاختبارين الفرعيين (الجزء الفردي والزوجي) الذي يقيس الاتساق الداخلي أو التجانس فإذا كان الارتباط بين الجزأين مرتفع فهذا يعني أنهما يقيسان نفس الشيء.

والجدير بالذكر أنه هناك أربع طرق شائعة لتقسيم الجزأين وهي:

– الطريقة الأولى: اختيار البنود الفردية لتمثل الجزء الأول من الاختبار والبنود الزوجية لتمثل الجزء الثاني للاختبار

– الطريقة الثانية: ترتيب الأسئلة وفق مستوى الصعوبة والسهولة، ثم اختيار البنود الفردية لتشكل الاختبار الفرعي الأول والبنود الزوجية لتؤلف الجزء الفرعي الثاني للاختبار والعكس كذلك صحيح.

–الطريقة الثالثة: تجزئة الاختبار إلى جزأين بحيث تكون مزاوجة بين فقرات الجزأين من حيث المحتوى.

– الطريقة الرابعة: اختيار بطريقة عشوائية لجزأي (نصفي) الاختبار(ليندا كروكر وجيمز الجينا، 2009، ص. 185)

غير أن تقدير الثبات بهذه الطريقة يكون منخفضا، لأنه يتم من خلال نصف الاختبار، والمعلوم أن ثبات الاختبار يتأثر بطول الاختبار، فكلما زاد عدد البنود ارتفعت قيمة الثبات (ليندا كروكر وجيمز الجينا، 2009، ص.186؛ إبراهيم محمد المحاسنة، 2013، ص. 128).

ولتصحيح قيمة الثبات يعتمد الباحث على استخدام معادلة سبيرمان براون لتصحيح طول الاختبار حيث تسمح بتقدير ثبات الاختبار الكلي.

- تطبيق معادلة سبيرمان براون:

تقوم هذه الطريقة على افتراض أساسي وهو التوازي بين نصفي الاختبار، وعليه بعد تجزئة الاختبار إلى جزأين، البنود الفردية والبنود الزوجية مثلا، يتعين علينا التأكد من تكافؤ الجزأين وهذا من خلال حساب تباييني جزئي الاختبار والتأكد من أنهما: متساويين أو متجانسين (باستعمال اختبار التجانسF).

حالة التجانس نلجأ لها بعدما تبين لنا عدم تساوي تباييني نصفي الاختبار $S^2_2 \neq S^2_1$

وفي توفر أحد الشرطين السابقين يمكن تطبيق معادلة سبيرمان براون التالية:

$$r_{1.1} = \frac{2\, r_{\frac{1}{2}\frac{1}{2}}}{1 + r_{\frac{1}{2}\frac{1}{2}}}$$

بحيث $r_{1.1}$ هو قيمة ثبات الاختبار ككل و $r_{\frac{1}{2}\frac{1}{2}}$ معامل الارتباط بين نتائج نصفي الاختبار

(سيسيل رينولدسو رونالد ليفي نقستون، 2013، ص. 173)

وللتذكير لحساب تباين الجزأين (نصفي الاختبار) نستخدم المعادلتين التاليتين:

تباين النصف الأول للاختبار:

$$s^2{}_x = \frac{N\sum x^2 - (\sum x)^2}{N_1(N_1 - 1)}$$

تباين النصف الثاني للاختبار:

$$S_Y^2 = \frac{N\sum y^2 - (\sum y)^2}{N_2(N_2 - 1)}$$

لتأكد من تجانس نصفي الاختبار نقوم باستخدام اختبار F

أما إذا لم يتحقق شرط التكافؤ في نصفي الاختبار، في هذه الحالة أوجد رولون(Rulon,1939) طريقة بديلة لتقدير الثبات تعتمد على حساب الفرق بين درجات الأفراد في نصفي الاختبار، وتصاغ معادلتها على النحو الآتي:

$$r_{xy} = 1 - \frac{S^2{}_D}{S^2{}_X}$$

S^2_D : تباين الفرق في الدرجات بين نصفي الاختبار

S^2_X: التباين الكلي للاختبار(**Laveault, 2002 , p. 119**).

بالإضافة إلى هذه الطريقة، اقترح جوتمان Guttman معادلة لتقدير الثبات تأخذ في الحسبان اختلاف تباين نصفي الاختبار وتعطي نفس نتائج معادلة رولون والتي صيغت على النحو التالي:

$$R_{xx} = 1 - \frac{S^2_1 - S^2_2}{S^2_x}$$

والجدير بالذكر أن طريقة التجزئة النصفية يعاب عليها أنها تقدم لنا معلومات على ثبات نصف الاختبار وما يهمنا هو الاختبار ككل، فتكون قيمة الثبات منخفضة باعتبار أنها تحسب من نصف الاختبار، والمعلوم أنه كلما زاد طول الاختبار ارتفعت قيمة الثبات.

كما يعاب على الطريقة أن الثبات يتأثر بنوع التجزئة المختارة من طرف الباحث (بنود زوجية وفردية، تقسيم حسب تسلسل البنود إلى نصفين، اختيار عشوائي للبنود...الخ)، وتكون له قيم مختلفة باختلاف نوع التجزئة المستخدمة(**Laveault, 2002, p. 119**) حيث قد يتأثر النصف الثاني من الاختبار بعامل الملل والاجهاد وضيق الوقت.

ومن السلبيات تقدير الثبات بطريقة التجزئة أنها لا تصلح مع الاختبارات الموقوتة، صعوبة تكافؤ الأسئلة الفردية والزوجية من حيث مستوى الصعوبة، التمييز التباين،....الخ .

أما عن إيجابيات تقدير الثبات باستخدام التجزئة النصفية فإنه يسمح لنا بتقديم معلومات حول مدى التماسك الداخلي للاختبار فكلما كانت قيمة الثبات مرتفعة دل ذلك على أن النصفين يقيسان نفس الشيء، كما يكون ملائما للاختبارات الطويلة.

يقضي كذلك على عامل الألفة في الاختبارات، اقتصادي في الوقت والجهد والتكلفة، كما يجنبنا الأخطاء الراجعة لمعاينة الزمن (أمين علي محمد سليمان، 2010؛ بشرى إسماعيل، 2004)

مثال تطبيقي: التجزئة النصفية

الجدول رقم (11): مثال تطبيقي في موضوع التجزئة النصفية

	البنود										بنود ف	بنود ز	بنود 5-1	بنود 10-6
	1	2	3	4	5	6	7	8	9	10				
1	2	2	3	1	3	4	1	2	4	1	13	10	11	12
2	4	3	3	1	3	5	2	4	5	2	17	15	14	18
3	5	3	5	3	4	5	3	5	5	2	22	18	20	20
4	2	1	3	2	3	4	1	3	5	1	14	11	11	14
5	3	2	5	4	5	5	2	4	5	1	20	16	19	17
6	3	3	5	4	5	5	2	4	5	1	20	17	20	17
7	1	1	3	2	3	3	0	2	4	2	11	10	10	11
8	1	1	3	1	4	3	0	1	3	0	11	6	10	7
9	0	1	2	0	3	3	0	1	3	0	8	5	6	7
10	5	3	5	3	3	5	3	5	5	3	21	19	19	21
											157	110	140	144
	معامل الارتباط بين الجزء الفردي والزوجي من الاختبار										0,97			
	معامل الارتباط بين الجزء الأول (1-5) والجزء الثاني (6-10)										0,88			

المرجع: (D.LAVEAULT & J.GREGOIRE, 2002, p.118)

نلاحظ من خلال الجدول أن قيمة الارتباط بين جزأي الاختبار تختلف باختلاف نوع التجزئة.

الجدول رقم (12): حساب المجاميع (مدخلات المعادلة)

عدد الأفراد	مجموع البنود الفردية x	مجموع البنود الزوجية y	X^2	Y^2	X . Y
1	13	10	169	100	130
2	17	15	289	225	255
3	22	18	484	324	396
4	14	11	196	121	154
5	20	16	400	256	320
6	20	17	400	289	340
7	11	10	121	100	110
8	11	6	121	36	66
9	8	5	64	25	40
10	21	19	441	361	399
مجموع ∑	157	127	2685	1837	2210

أولا: التحقق من عدم وجود تباين بين الجزأين وذلك بتطبيق معادلة التباين

$$s^2{}_x = \frac{N\sum x^2 - (\sum x)^2}{N_1(N_1 - 1)}$$

$$S_Y^2 = \frac{N\sum y^2 - (\sum y)^2}{N_2(N_2-1)}$$

تباين الجزء الأول :

بالتعويض في المعادلة:

$$S_X^2 = \frac{10\ X\ 2685 - 157^2}{10(10-1)} = \frac{2685 - 24649}{90} = 24.45$$

تباين الجزء الثاني:

بالتعويض في المعادلة :

$$S_Y^2 = \frac{10\ X\ 1837 - 127^2}{10\ X\ (10-1)} = \frac{18370 - 16129}{90} = \frac{2241}{90} = 24.9$$

نلاحظ أن تباين الجزأين غير متساويين $S_Y^2 \neq S_X^2$ لذا في خطوة ثانية لابدا التحقق من التجانس باستخدام اختبار F

$F = \frac{s_Y^2}{s_X^2} = \frac{24,9}{24.45} = 1,01$ يتم مقارنة هذه القيمة المحسوبة مع القيمة المجدولة لـ F التي تقدر بـ 3.18 عند درجة الحرية 9 والتي يظهر أنها أصغر منها وعليه يمكن استنتاج أن الاختبار يتميز بالتجانس، وبذلك المعادلة التي ينبغي استخدامها في تقدير الثبات هي سبيرمان براون.

تطبيق معادلة سبيرمان براون: $r_{1.1} = \frac{2\ r\ \frac{1}{2}\frac{1}{2}}{1+ r\ \frac{1}{2}\frac{1}{2}}$

$$R_{1.1} = \frac{2\ x\ 0,97}{1 + 0,97} = \frac{1.94}{1.97} = 0.98$$

نلاحظ أن قيمة الثبات المصححة المتحصل عليها أكبر من قيمة الثبات لنصف الاختبار، بحيث أنه كلما زاد عدد مفردات الاختبار ارتفعت قيمة الثبات.

✓ **طريقة ألفاكرونباخ (Cronbach) (α) لتقدير الثبات :**

تعتبر من أكثر الطرق استخدام التقدير ثبات الاتساق الداخلي، حيث تقوم على مدى اتساق إجابة المفحوصين على كل بند من بنود الاختبار مع البنود المتبقية للاختبار، وتعتمد الطريقة على حساب ارتباط البنود مع بعضها البعض وارتباط كل بند مع الاختبار ككل.تسمح هذه الطريقة بالتعرف على البنود التي لا تساهم في الاتساق الداخلي بحيث يكون معامل ارتباطها مع البنود ضعيفا، وحذف هذه البنود سيسمح برفع من تقدير الثبات.

وهي تصلح لأي نوع من أنواع الاختبارات الموضوعية التي تصحح بطريقة (0 و1) والاختبارات والمقاييس المتدرجة من النوع الثلاثي أو الرباعي أو الخماسي التدريج،...الخ مثل استبيانات أو مقاييس الاتجاهات، الميول والشخصية.

كما تصلح كذلك لاختبارات التحصيل م ننوع المقال التي يتم تقدير الدرجات تبع النوعية الإجابة.

حيث اشتق كرونباخ معادلته من الصيغة $KR20$ لكيودر ريتشاردسون، والتي تعتمد على تباينات بنود الاختبار والتباين الكليل لاختبار، وتصاغ معادلته على النحو التالي:

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

بحيث أن:

K: عدد الفقرات أو البنود

$\sum S_i^2$: مجموع تباين البنود المكونة للاختبار

S^2 : التباين الكلي للاختبار (أمين علي محمد سليمان، 2010؛ بشرى إسماعيل، 2004)

تحديد دلالة معامل ثبات ألفا كرونباخ (تفسير قيم α):

قيمة معامل α المحصورة بين (0.6 و 0.7) تدل على درجة مقبولة للثبات

قيمة معامل α تساوي 0.8 تدل على درجة أعلى من الاتساق الداخلي للبنود

قيمة معامل $\alpha \geq 0.95$ تدل على وجود تكرار في البنود وتداخل بين الفقرات نفسها ولا تدل على درجة عالية من الثبات.

كما يجدر الإشارة أنه إذا ما كانت تباينات الفقرات المكونة للاختبار أو الأجزاء المكونة للاختبار متكافئة وكانت تقيس سمة واحدة فإن أفضل تقدير لمعامل الثبات فسيكون باستخدام ألفا كرونباخ ،أما إذا لم تتوفر ظروف التكافؤ بالنسبة للفقرات فإن معامل الثبات الذي يعتمد تقديره على معادلة ألفا كرونباخ يعطي قيمة منخفضة للثبات أقل من القيمة الحقيقية لها.

وعليه فمعامل ألفا هو الحد الأدنى لمعامل الثبات النظري والذي نتحصل عليه من خلال حساب متوسط المعاملات الممكنة للتجزئة النصفية المحسوبة بطريقة رولون، ويسمى كذلك معامل الدقة.

✓ **طريقة كيودر ريتشاردسون (Kuder–Richardson):**

تعتمد هذه الطريقة على تحليل الإجابات على البنود وحساب تباينها، تطبق في حالة الاختبارات الموضوعية التي تكون الاستجابة على بنودها بـ (نعم، لا) أو (ص، خ) والتي تنقط بـ (1، 0) (أمين علي محمد سليمان، 2010، ص. 577)

تصاغ قوانينه على الشكل الآتي:

$$KR_{20} = \frac{K}{K-1}\left(1 - \frac{\sum PQ}{S^2}\right)$$

$$KR_{21} = \frac{K}{K-1}\left(1 - \frac{\bar{X}\ (K-\bar{X})}{KS^2_X}\right)$$

حيث أن :

K : عدد البنود أو الفقرات للاختبار

P: معامل الصعوبة ونتحصل عليه من نسبة (عدد الإجابات الخاطئة / عدد الأفراد الذين أجابوا على السؤال أو البند.

Q أو (P–1) : معامل السهولة و نتحصل عليه من نسبة (عدد الإجابات الصحيحة / عدد الأفراد الذين أجابوا على السؤال أو البند.

S^2_X: التباين الكلي للاختبار.

$\bar{X}$: المتوسط الحسابي لنتائج الاختبار (بشرى إسماعيل، 2004، ص ص. 78–79)

ملاحظة :

يجدر الإشارة أن كلتا المعادلتين KR_{20} ، KR_{21}تعطي نفس التقدير للثبات في حالة تساو مستوى الصعوبة والتباين لجميع البنود، بينما تعطي معادلة KR_{21} قيمة أقل للاتساق الداخلي عندما تختلف مستويات الصعوبة والتباين للبنود.

كما عادة ما تعوض معادلة K_{20} بمعادلة KR_{21} في حالة تساوي مستوى الصعوبة والتباين حيث تسمح بتقدير الاتساق من خلال المتوسط الحسابي للدرجات والتباين الكلي.

في حين نعتمد على معادلة KR_{20} في تقدير الثبات في حالة عدم تساوي مستويات الصعوبة والتباين للمفردات بحيث ينتج عن معادلة KR_{21} قيمة منخفضة للثبات (ليندا كروكر وجيمز الجينا،2009، ص. 190)

العوامل المؤثرة في الثبات

تمهيد:

هناك العديد من العوامل التي تؤثر في درجات الاختبار بعضها يعود للفرد نفسه مثل القدرة على أدائه نوعا معينا من المهارات تتصل بما يقيسه الاختبار وطريقته في هذا الأداء، وفهمه لتعليمات الاختبار، كذلك عوامل التعب أو الإجهاد أو الملل والتوتر الانفعالي والذاكرة وغير ذلك وما يتصل بالاختبار في حد ذاته مثل صياغة بنود الاختبار والتعليمات وعوامل الصدفة وطريقة الاجراء وغير ذلك. (سعد عبد الرحمن، 1998، ص.176)

غير أن العوامل المهمة التي يجب أن نشير إليها خاصة أن بعضها تحتاج إلى معالجة إحصائية يمكن أن نلخص أهمها فيما يلي:

عوامل مرتبطة بالاختبار:

- طول الاختبار: ويقصد به عدد بنود أو أسئلة الاختبار، وتجدر الإشارة إلى أنه تزداد قيمة الثبات بزيادة عدد البنود.

ان زيادة طول الاختبار بزيادة فقراته يجعل درجة ثباته أعلى، وذلك لأن الاختبار الأطول أكثر تمثيلا للخاصية أو السلوك المراد قياسه، مما يقلل من أخطاء القياس الناتجة عن الصدفة والتخمين إلى أدنى حد ممكن ويجعل الدرجة الملاحظة على الاختبار أكثر اعتمادا وارتباطا بخصائص الطالب المفحوص من ناحية وأكثر دلالة على المستوى الحقيقي لأدائه من ناحية أخرى. (بشرى إسماعيل، 2004، 82.)

إن هذه العلاقة التي تجمع طول الاختبار بالثبات، صاغ لها سبيرمان براون قانونا خاصا وضح فيه العلاقة الطردية بين قيمة الثبات وإضافة البنود وسماه بمعادلة تصحيح الطول.

$$ن = \frac{ر\ ن\ ن(1-ر\ ت\ ث)}{ر\ ت\ ث(1-ر\ ن\ ن)}$$

حيث

ن: مضاعفات الطول المطلوبة للاختبار

ر ن ن: معامل الثبات المطلوب للاختبار

رت ث: معامل الثبات الحالي للاختبار قبل زيادة عدد البنود

مثال تطبيقي:

إذا كان عدد البنود للاختبار يساوي 12 وقيمة معامل ثبات الاختبار تبلغ 0.60 وأردنا رفع قيمة الثبات إلى 0.80، فكم عدد البنود التي ينبغي أن تكون الاختبار.

بالتعويض في المعادلة:

$$ن = \frac{(0.60-1)0.80}{(0.80-1)0.60} = 2.7$$

فإذا كان عدد البنود يساوي 12 فبضربه في 2.7 نتحصل على عدد البنود التي يجب أن يتكون منها الاختبار وفي هذا المثال نتحصل على عدد 34 بند.

(عبد الله زيد الكيلاني،2009؛ بشير معمرية ، 2012)

- صعوبة أو سهولة الأسئلة: كلما كانت الأسئلة صعبة جدا أو سهلة جدا، كان معامل ثبات الاختبار منخفضا، والأفضل أن يكون الاختبار متوسط الصعوبة، ويعود ذلك إلى أن الأسئلة الصعبة أو السهلة تؤدي إلى انخفاض تباين الأسئلة وبالتالي انخفاض التباين الكلي للاختبار. (إبراهيم مصطفى علي حماد، 2012)

ويتعين أن يكون مستوى صعوبة الأسئلة يتراوح ما بين 40% و60% حتى تكون قيمة الثبات مقبولة

وللإشارة يحسب معامل الصعوبة للبند على النحو الآتي:

$$\text{معامل صعوبة البند} = \frac{\text{عدد الذين خطأأجابوا عن البند}}{\text{عدد أفراد العينة الكلية}} \times 100$$ (بشير معمرية، 2012)

- الاستقلال بين البنود: يعني الاستقلال هنا ألا تؤدي إجابة معينة على بند إجابة على بند آخر. (بشير معمرية، 2012)
- صياغة البنود: الأسئلة الغامضة والخادعة تقلل من قيمة الثبات (سامي محمد ملحم، 2000)
- زمن الاختبار: يؤثر الزمن المحدد للإجابة على الاختبار بشكل مباشر على الثبات. فيزداد معامل الثبات بزيادة الزمن المستغرق للإجابة.

ويشير كل من Lindguist& Cook إلى أن معامل الثبات يزداد تبعا لزيادة الزمن المستغرق للإجابة على فقرات الاختبار حتى يصل إلى الحد المناسب للإجابة على فقرات الاختبار ثم يقل الثبات تبعا لذلك كلما زاد الزمن عن ذلك الحد (سامي محمد ملحم، 2017)

- موضوعية التصحيح: ان مشكل موضوعية التصحيح في اختبارات الموضوعية غير مطروح لعدم تدخل الذاتية في التصحيح، بينما في اختبارات المقال قد يختلف المصححين في تقييم الاختبار ويؤثر ذلك على قيمة الثبات.
- صدق الاختبار: الاختبار الصادق يكون ثابتا، وليس الاختبار ليس بالضرورة صادقا، فكلما زاد صدق الاختبار زاد ثباته. (فاروق الروسان، 2012)

1-عوامل مرتبطة بالفرد:

- تجانس العينة: يؤدي التجانس الشديد لعينة الأفراد التي يحسب لها الثبات من خلال أدائها إلى انخفاض ملموس في معامل ثبات الاختبار لأن التباين داخل هذه العينة المتجانسة يكون منخفضا بقدر لا يسمح بتقدير التباين الحقيقي للاختبار. (بشير معمرية، 2012).
- أثر التخمين: ان زيادة التخمين تنقص ثبات أي اختبار. لأن الإجابة التي تعتمد على التخمين في المرة الأولى لإجراء الاختبار لا تعتمد على نفس هذا التخمين في المرات القادمة لإجراء ذلك الاختبار، تعتبر الاختبارات من المتعددة البدائل أكثر الاختبارات تأثرا بالتخمين.

حيث يرتفع أثر التخمين كلما قل عدد الاختيارات أو البدائل، فالاختبار الذي يتضمن اختيار إجابة واحدة من بديلين أو خيارين، يكون احتمال الإجابة عليه بالتخمين او الصدفة بـ 50%، بينما يكون احتمال الإجابة بالصدفة على اختبار يتضمن أسئلة لثلاث بدائل 33 % و 25 % للإجابة على أسئلة تحتوي على أربعة بدائل.

ولتصحيح الدرجة المتحصل عليها من أثر التخمين، اقترح الأخصائيون في القياس المعادلة التالية:

$$\text{د ص} = \text{ص} - \frac{\text{خ}}{\text{ب}-1}$$

حيث د ص: الدرجة المصححة من أثر التخمين

ص: عدد الإجابات الصحيحة

خ : عدد الإجابات الخاطئة

ب: عدد البدائل في الاختبار

(سامي محمد ملحم، 2000 ؛ بشير معمرية، 2012).

- حجم العينة: إن علاقة الثبات بحجم العينة علاقة طردية، فكلما زاد عدد أفراد العينة ارتفعت قيمة الثبات. (فاروق الروسان، 2012)

ثانيا: صدق الاختبار

مقدمة:

يعتبر الصدق من الشروط المهمة الواجب توافرها في الاختبار الجيد إلى جانب الثبات، حيث يوفر الدليل المادي والمبرر النظري اللازمين لإثبات كفاية وملاءمة ومعنى أي تأويل أو فعل يبني على درجة الاختبار.

فالصدق يعتمد على عدد من العوامل وفقا لظروف الاستخدام والقرارات المترتبة على النتائج والاستدلالات المبنية على الدرجة ومعناها.

وبذلك لا يمكن اقتصار على جمع نوع واحد من الأدلة الصدق الاختبار، بل يتعين جمع أكثر من دليل مع الأخذ في الاعتبار أن الأدلة التي يتم الحصول عليها متكاملة فيما بينها ومتداخلة ويدعم بعضها البعض.

يرى المختصون في مجال القياس النفسي أن الصدق هو الخاصية الوحيدة التي تحدد جوانب الاختبار وأن الاختبار الصادق هو الذي ترتبط درجاته بدرجة عالية لسلوك الفعلي الذي يراد قياسه.

التعريف بمفهوم الصدق:

عرف جلفورد(Guilford, 1954) الصدق هو تحديد لمعامل الارتباط بين الاختبار وبعض مقاييس أو محكات الأداء في مواقف الحياة. (بشير معمرية، 2011، ص.180)

بينما يعرف ساكس (1980) الصدق على أنه الدرجة التي تكون فيها المقاييس ذات فائدة في اتحاذ القرارات المتعلقة بهدف أو غرض من الأغراض.

أما ألن وين (Allen &Yen,1979) يعرفان الصدق بأنه القدرة على أن يقيس الاختبار ما وضع من أجله. بينما كرونباخ (1971) يصف الاختبار بأنه العملية التي يبحث من خلالها مصمم الاختبار عن دليل يدعم به الاستنتاجات التي يصل إليها من خلال درجات الاختبار. بينما مهرنس وليمان(1975) يعرفان الصدق بأنه الدرجة التي يتمتع بها الاختبار في تحقيق أهداف محددة(عبد الرحمن بن سليمان الطريري، 1997، ص ص. 218-219).

مفاهيمه الأساسية:

- أن يكون الاختبار قادرا على أن يقيس ما وضع لقياسه، بمعنى أن يكون الاختبار ذا صلة وثيقة بالقدرة المقاسة.
- أن يكون الاختبار قادرا على أن يقيس ما وضع لقياسه وفقط، بمعنى أن يكون هذا الاختبار قادرا أن يميز بين القدرة التي يقيسها وقدرات أخرى التي يحتمل أن تختلط بها أو تتداخل معها.
- أن يكون قادرا على أن يميز بين طرفي القدرة التي يقيسها. بمعنى أن يميز بين الأداء العالي، المتوسط والأداء الضعيف وبذلك يظهر الفروق الفردية بين أفراد العينة. (سعد عبد الرحمن، 1998، ص. 183)

أنواع الصدق وطرق تعيينه:

رغم تعدد مسميات الصدق مثل الصدق الظاهري، المحكمين، التمييزي العاملي، التنبئي، التلازمي، المحكي، التكويني... ..إلخ ، غير أنه اتفق المتخصصين على وجود ثلاث أنواع رئيسية للصدق ، سنتطرق لها بالتفصيل فيما يلي :

صدق المحتوى أو المضمون:

مفهومه: هو فحص مبدئي لمحتويات الاختبار لمعرفة مدى ملاءمة المحتوى في تمثيله للمجال المراد قياسه، ويتم عن طريق التحليل المنطقي لمحتوى أداة القياس اعتمادا على حكم فردي. (عبد الحفيظ مقدم، 146) و (موسى النبهان،2004، ص. 274).

ففي الاختبارات التحصيلية من السهل التحقق من صدق المحتوى بحيث يشير إلى جودة عينة الفقرات التي تتضمنها أداة القياس، بمعنى مدى تمثيلها للمادة التعليمية أو المنهج الدراسي. بينما قد يصعب ذلك، عندما نقيس خاصية غير واضحة التعريف، ففي هذه الحالة قد يحتكم إلى رأي متخصصين قد راجعوا المجال الذي يمثله الاختبار حول مفردات الاختبار.

ومما سبق يتضمن صدق المحتوى شكلين وهما:

- الصدق الظاهري أو الشكلي أو السطحي: يعكس مدى انسجام فقرات الاختبار ومفاهيمه وفقا لتعريف صاحب الاختبار، ويتم الحكم على الصدق الشكلي من خلال التبصر في مضمون فقرات الاختبار وتفحص مدى علاقتها بمحتوى المادة التعليمية. فيستدل على

صدق الاختبار من الشكل الخارجي له والعلاقة الظاهرة بين فقراته والموضوع الذي يستخدم فيه التقويم. (موسى النبهان، 2004، ص. 275).

- الصدق المنطقي أو العيني: يقاس هذا النوع بالتحليل المنطقي لمحتويات الاختبار ومطابقتها بمحتويات الجانب المقاس.

طرق تعيينه: لا يوجد مقياس احصائي دقيق لتقدير صدق المحتوى وهذا لصعوبة وصف مجال المحتوى. لذلك عادة ما نلجأ لتقدير صدق المحتوى بناء على محكمين من ذوي الخبرة والكفاءة في موضوع الاهتمام الذي يخصه الاختبار، يطلب منهم ابداء آرائهم حول ما مدى ملاءمة فقرات الاختبار من حيث بنيتها وصياغتها وتوافقها مع الأهداف التي وضعت لقياسها وهذا بملئهم لاستمارة خاصة. بعدها يتم تحليل البيانات احصائيا لمعرفة مدى الاتفاق بين المحكمين والذي يحدد عادة من 80% فما فوق. ويشترط ألا يقل عدد المحكمين على 05.كما يمكن تقديره بمعامل كندال. (بشير معمرية، 2011، ص ص. 184-186) (صلاح أحمد مراد، 2002، ص. 351)

صدق المرتبط بالمحك (التجريبي):

مفهومه: تتم هذه الطريقة بحساب معامل الارتباط بين الاختبار وبين اختبار آخر يطلق عليه المحك ثبت صدقه، وكلما كان معامل الارتباط مرتفعاً كان معامل الصدق مرتفعاً ،حيث يتم مقارنة درجات الاختبار المطلوب حساب صدقه بمقاييس أخرى للأداء سواء في وقت سابق أو وقت متلازم مع وقت أداء الاختبار أو في وقت لاحق لتطبيقه. (سامي محمد ملحم، 2000،

ص.288) ونقصد بالمحك هو مقياس موضوعي مستقل عن أداة القياس نقيس به صدق الاختبار من خصائصه:

- أن يكون متعلق بالهدف الذي تريد أن تقيسه الأداة (الاختبار)
- أن يكون موضوعي ويتميز بالثبات
- أن يكون مستقلا بمعنى لا يتأثر درجة الفرد في الاختبار بدرجته في المحك. (عبد الحفيظ مقدم، 1993، 147)

طرق تعيين صدق المحك (التجريبي):

الصدق التلازمي: ويمكن حساب معامل الصدق في هذه الطريقة بمعامل الارتباط بين درجات الأفراد على الاختبار ودرجاتهم في الأداء الفعلي في جوانب السلوك التي يقيسها الاختبار ٠ وذلك بشرط أن تكون درجات أداء الأفراد قد تم جمعها في نفس وقت إجراء الاختبار أو بفاصل زمني قصير. (عبد الله زيد الكيلاني، 2009، ص. 228)

والصدق التلازمي هو أكثر أنواع الصدق ملاءمة للاختبارات التي تستخدم لأغراض التشخيص، يستعمل بكثرة في المجال الصناعي لانتقاء الأشخاص للوظائف وفي التشخيص الاكلينيكي. (صلاح الدين، 209)

الصدق التنبئي: يقوم على أساس استخدام درجات الاختبار في التنبؤ بالأداء في المستقبل على مقاييس أخرى مشابهة في للموضوع، ويتم الحصول عليه عن طريق حساب معامل الارتباط بيرسون

بين نتائج الاختبار والمحك الخارجي الذي جمعت المعلومات عنه بعد فاصل زمني لا يقل عن 6 أشهر ٠

يستخدم هذا النوع من الصدق في انتقاء الأفراد وتصنيفهم وتوجيههم تربوياً أو مهنيا ٠(سامي محمد ملحم، 2000، ص. 288) و (عبد الله زيد الكيلاني، 2009، ص. 288)

ملاحظة:

يمكن التمييز بين " الصدق التلازمي والصدق التنبئي " باستخدام متغير الزمن الذي تجمع فيه البيانات الخاصة بالمحك، فإذا جمعت هذه البيانات تقريباً في نفس الوقت الذي نحصل فيه على بيانات الاختبار فإننا نكون بصدد الحديث عن الصدق التلازمي، أما إذا جمعت بيانات المحك في وقت لاحق فإننا نكون بصدد الحديث عن الصدق التنبئي ٠

وثمة طريقة أخرى للتفرقة بين هذين النوعين من صدق المحك وهو تمييز منطقي أكثر منه تمييز إجرائي ولا يهتم بالزمن وإنما يركز على الغرض من الاختبار فإذا كان الغرض منه تشخيصي فإننا نخص بالاهتمام الصدق التلازمي وإذا كان الغرض منه التنبؤ بالسلوك فنولي الاهتمام بالصدق التنبئي (صلاح أحمد مراد وآخرون،2002، ص. 354)

صدق التكوين الفرضي أو صدق المفهوم أو البناء:

مفهومه:

يقصد به إلى أي حد يقيس الاختبار السمة التي لها وجود فعلي وإلى أي حد يكون هناك تناظر بين التفسير المقترح للسمة وما يقيسه الاختبار فعلا. (موسى النبهان، 2004، ص. 295)

طرق تعيين معامل الصدق:

- صدق الاتساق الداخلي: يتم تعيينه بحساب بمعاملات الارتباط بين الدرجة الكلية للمقياس ودرجات كل بند من بنود الاختبار وبين الدرجة الكلية للاختبار ودرجة المقاييس الفرعية أو الأبعاد. (صلاح أحمد مراد وآخرون، 2002، ص. 357)
- التحليل العاملي: يعتمد على تحليل الارتباطات بين المتغيرات (فقرات المقياس) ومن ثم تفسير هذه الارتباطات واختزالها في عدد أقل من المتغيرات تدعى عوامل. (موسى النبهان، 2004، ص. 300)

ويقصد به مدى تشبع بنود الاختبار بالعامل الذي نفترض أن يقيسه الاختبار، فكلما كان التشبع كبير كلما دل ذلك على ارتفاع مستوى صدق الاختبار. (نعمان شحادة، 2009، ص. 259)

العوامل المؤثرة في الصدق:

لخص بشير معمرية (2012) العوامل التي تؤثر في الصدق فيما يلي:

- طبيعة عينة التقنين: إن الاختبار الواحد قد يقيس عمليات نفسية مختلفة إذا طبق على عينات مختلفة في العمر أو الجنس أو المهنة أ والمستوى التعليمي أو السواء أو اللاسواء أو أي متغير آخر يرتبط بمفهوم الصدق. مثلا اختبار في العمليات الحسابية الذي يكون اختبار صادقا للاستدلال الحسابي عند أطفال المدرسة الابتدائية نجده يقيس بصدق السرعة العددية لدى تلاميذ المرحلة الثانوية.
- العلاقة الخطية بين الاختبار والمحك: من المعروف أن حساب معامل الارتباط بين الاختبار والمحك بطريقة كارل بيرسون يشترط افتراضا جوهريا هو أن تكون العلاقة بين المتغيرين (الاختبار والمحك) علاقة خطية على امتداد مدى الفروق الفردية بينهما. فإذا لم تتحقق فرضية العلاقة الخطية تنخفض قيمة معامل الارتباط بين درجات الأفراد على الاختبار ودرجاتهم على المحك.
- تجانس عينة التقنين: كلما زاد التجانس بين أفراد العينة في الخاصية محل القياس أدى ذلك إلى انخفاض محسوس في مدى درجات أفراد عينة التقنين على الاختبار الأمر الذي يترتب عليه انخفاض واضح في قيمة معامل الارتباط بين درجاتهم على الاختبار ودرجاتهم على المحك.

- الفترة الزمنية بين تطبيق الاختبار وتطبيق المحك: تعد الفترة الزمنية الفاصلة بين تطبيق الاختبار وجمع المعلومات عن المحك من العوامل التي تؤثر على الصدق. فعندما تطول هذه الفترة الزمنية ينخفض معامل الصدق. ويعزى ذلك إلى عامل **الصدفة** التي تؤثر في درجات الأفراد على المحك وإلى عوامل أخرى مثل نقص أفراد العينة بسبب التسرب او الغياب أو انتقالهم إلى أماكن أخرى أو عدم تعاملهم. (بشير معمرية، 2012، ص ص. 244- 245)

خلاصة

إن القياس من الأدوات الأساسية التي ينبغي الاستناد لها لتقصي وفهم الظواهر في كل المجالات، ولا سيما في المجال التربوي التي يعد الموجه والمحرك للعملية التعلمية التعليمية، فيجب أن ينظر إلى القياس واستخدام الاختبارات في مجال التربية كعملية مكملة لعملية التدريس، كما سبق لبرنر ودايفيز (Purnel Davis, 1939) مناقشة هذه الفكرة عند تناولهما لاستخدام اختبارات الفصل المدرسي من أجل توجيه عملية التعلم، حيث ذكرا" أن القياس التربوي (Educational Measurement) يمكن أن يكون له أثر عميق في تحسين التعليم، وليؤدي هذا الدور يجب أن ينظر له جنبا إلى جنب مع التخطيط للتعليم، ونتائجه يجب أن تستخدم بصفة مستمرة لتوجيه كل من تخطيط وتطوير البرامج".

المراجع:

1- أحمد مراد صلاح وأمين علي سليمان (2005)، الاختبارات والمقاييس في العلوم النفسية والتربوية، خطوات اعدادها وخصائصها، دار الكتاب الحديث، القاهرة.

2- أبو حطب فؤاد، سيد أحمد عثمان، أمال صادق (2008)، التقويم النفسي، مكتبة الانجلو المصرية، القاهرة.

3- الأنصاري بدر محمد (2009)، قياس الشخصية، دار الكتاب الحديث، القاهرة.

4- البغدادي محمد رضا (1983)، الهداف والاختبارات بين النظرية والتطبيق، دار المعارف، القاهرة.

5- بشرى إسماعيل (2004)، المرجع في القياس النفسي، مكتبة الانجلو المصرية، القاهرة.

6- بن سي مسعود لبنى (2008)، واقع التقويم في التعليم الابتدائي في ظل المقاربة بالكفاءات، دراسة ميدانية بولاية ميلة، رسالة مقدمة لنيل شهادة ماجستير في العلوم التربوية، جامعة قسنطينة

7- بوسنة محمود (2007)، علم النفس القياسي المبادئ الأساسية، ديوان المطبوعات الجامعية، الجزائر.

8- بو علاق محمد (2012)، الموجه في الإحصاء الوصفي والاستدلالي في العلوم النفسية والتربوية والاجتماعية،

9- توق محي الدين وعبد الرحمان عدس (1984)، أساسيات علم النفس التربوي، الجامعة الأردنية.

10- الخوري توما جورج (2008)، القياس والتقويم في التربية والتعليم، المؤسسة الجامعية للدراسات والنشر والتوزيع، بيروت، لبنان.

11- سعد عبد الرحمن (1998)، القياس النفسي: النظرية والتطبيق، دار الفكر العربي، القاهرة، مصر.

12-سوسن شاكر الجلبي (2005)، أساسيات بناء الاختبارات والمقاييس النفسية والتربوية، مؤسسة علاء الدين للطباعة والتوزيع، دمشق، سوريا.

13- سيسيل رينولدس و رونالد ليفينقستون، ترجمة صلاح الدين محمود علام (2013)، اتقان القياس النفسي الحديث : النظريات والطرق، دار الفكر، الأردن.

14- الصراف قاسم علي (2002)، القياس والتقويم في التربية والتعليم، دار الكتاب الحديث، الكويت.

13- الطريري عبد الرحمن بن سليمان (1997)، القياس النفسي والتربوي، نظرياته، أسسه، تطبيقاته، مكتبة الرشد، الرياض.

14- عبد الخالق أحمد محمد (1993)، استخبارات الشخصية، دار المعرفة الجامعية، القاهرة.

15- عوض عباس محمود (1998)، القياس النفسي بين النظرية والتطبيق، دار المعرفة الجامعية.

16- غريب العربي (بدون سنة)، التقويم التربوي، مفهومه، أنواعه وأدواته، دار الغرب لنشر والتوزيع، وهران

17- فرحات ليلى السيد (2012)، القياس والاختبار في التربية الرياضية، مركز الكتاب للنشر.

18- فرج صفوت (2007)، القياس النفسي، مكتبة الانجلو المصرية، القاهرة.

19- قمش مصطفى وخليل المعايطة ومحمد البواليز(2008)، القياس والتقويم في التربية الخاصة، مطبعة دار الفكر، الأردن.

20- محاسنة إبراهيم محمد (2013)، القياس النفسي، دار الجرير للنشر والتوزيع،

21- معمرية بشير (2012)، أساسيات القياس النفسي وتصميم أدواته للطلاب والباحثين في علم النفس والتربية، دار الخلدونية، الجزائر.

22- ملحم سامي محمد (2017)، القياس والتقويم في التربية وعلم النفس، دار المسيرة للنشر والتوزيع، الأردن.

23- مقدم عبد الحفيظ (1993)، الإحصاء والقياس النفسي والتربوي، ديوان المطبوعات الجامعية، الجزائر.

24- موسى النبهان (2004)، أساسيات القياس في العلوم السلوكية، دار الشروق للنشر والتوزيع، مصر.

25- كامل محمد مغربي (2017) ، أساليب البحث العلمي في العلوم الإنسانية والاجتماعية، دار الثقافة للنشر ، عمان.

26- كروكر ليندا وجيمز الجينا، ترجمة زينات يوسف دعنا (2009)، مدخل إلى نظرية القياس، دار الفكر، الأردن .

27- عبد الله زيد الكيلاني عبد الله زيد وآخرون، 2002، القياس والتقويم في التعلم والتعليم ، الشركة العربية المتحدة التسويق والتوريدات، القاهرة، مصر.

28- Cronbach, lee J (1984), Essentials of Psychological Testing, New York : Harper & Row.

29- D. Laveault & J. Gregoire (2005), Introduction aux théories des tests en psychologie et en sciences de l'éducation, éditions de boeck, Bruxelles.

30- Tuckman, Burce, W (1988), Testing for Teachers, New York : Harcourt Brace Jovanovich, inc.