جامعة الجزائر 2 أبو القاسم سعد الله

كلية العلـوم الاجتماعيــة

قسـم علـم الـنفس

مطبوعة دروس في القياس النفسي

موجهة لطلبة السنة الثانية جذع مشترك علم النفس

اعـداد: د. بوشـــارب فوزيـــــة

أستـــاذة محـــــاضرة صنـــف أ

السنة الجامعية :2022-2023

الفهرس	الصفحة
الدرس الاول: المفاهيم النظرية للقياس	4
. مفهوم القياس	4
. عناصر عملية القياس	6
. العلاقة بين العدّ والقياس	7
. العلاقة بين التقدير والقياس	8
. علاقة بين التقويم والقياس	10
. أهداف القياس	14
. العوامل المؤثرة في القياس	14
. مصادر الخطأ في القياس	15
الدرس الثاني: القياس في العلوم الاجتماعية	16
. طبيعة وخصائص القياس في العلوم الاجتماعية	16
. طبيعة و خصائص القياس في علم النفس	17
الدرس الثـــــالث: نشأة القياس وتطوره في علم النفس	29
الدرس الرابع: عرض أنواع أدوات وبيانات ومتغيرات ومحكات القياس النفسي	44
الدرس الخـــــــامس: مستويـــــات القيـــــاس	52
. القياس الاسمي	52
. القياس الرتبي	54
. القياس الفتري	55
. القياس النسبي	57
الدرس السادس: مفهوم الخصائص السيكومترية للمقاييس	62
الدرس السابع: طرق حساب الصدق	89
الدرس الثامن: طرق حساب الثبات	107
المراجع	119

الملاحق

.الملحق الأول: نماذج من أدوات التقدير(مع الشرح)

. الملحق الثاني: جدول T test

الدرس الاول: المفاهيم النظرية للقياس

مفهوم القياس :

منذ أن وجد الإنسان على وجه الأرض، وهو يقوم بالعد والتقدير والقياس ثم التقويم، حتى أن الحياة دون هذه العمليات لا يمكن أن تستقيم. يعتبر العد Enumération/Counting أول مهارة مارسها الإنسان لتحديد عدد النقود والأشياء والأشخاص والحيوانات. و يكون العدّ دائما بوحدات كاملة، فيقال مثلا: ـ يوجد في الحجرة ثلاثون طاولة وليس ثلاثون طاولة ونصف طاولة، مما يعني أن وحدات العدّ منفصلة أو متقطعة Discontinuous، وليست متصلة أو مستمرة Continuous، إذ لا يوجد في العد جزء أو أجزاء من الوحدة في الحالة الطبيعية. ومن الأمثلة على ذلك، عدد أفراد الأسرة أو عدد النقود أو عدد التلاميذ في الحجرة الدراسية، أو عدد السيارات أو عدد الأشجار،... إلخ.(محمود أحمد عمر وآخرون ،2010: 79)

هو بأوسع معنى، عملية تحديد قيم رقمية لأشياء أو موضوعات تبعا لقواعد معنية متفق عليها، مثلا : تحديد الأوزان بوحدات قياس مثلا الغرام (غ)، (كغ) (طن) ...الخ؛ وتحديد الطول والمساحات والأحجام بوحدات قياس أخرى مثل (سم)، (م) ...الخ؛ وتحديد درجات الحرارة المئوية بوحدات.

والقياس اصطلاحا هو عملية يتم بواسطتها تحديد كمية ما يوجد في الشيء من الخاصية المقاسة بدلالة وحدة قياس مناسبة؛ و بالتالي يقوم على تحديد الخصائص باستخدام الأعداد حسب قواعد معينة.

القياس لغة هو التقدير: يقال قاس الشيء بغيره أو على غيره أي قدره. ويقال قاس الطبيب الجرح أي قدره وبيّن عمقه. وفي الفقه معناه حمل فرع على أصل لسمة مشتركة بينهما مثل تحريم المخدرات على أساس أن كل ما يذهب عقل الإنسان فهو حرام.(سامي محمد ملحم، 2015: 26)

والقياس في لغة الإحصاء هو تقدير أشياء ومستويات مجهولة الكم أو الكيف تقديرا كميا وفق إطار معين من المقاييس المندرجة وذلك باستعمال وحدات رقمية متفق عليها أو مقننة. و هو تعريف عام لأن الأرقام وحدها لا تعطينا علما، ولا يكفي أن تضع مقياس ونطبقه لكي ينمو العلم، ولكن لابد من أن ندخل نتائج المقياس في بناء نظري يجمع بينها وبين أجزاء أخرى من معلوماتنا بطريقة متسقة .

وقد ورد في قاموس webster أن القياس هو التحقق من المدى أو الدرجة أو الكمية أو الأبعاد أو السعة بواسطة معيار وأداة القياس هي الطريقة أو الأسلوب الذي تقاس به صفة أو موضوع ما.

ويستخدم تعيير قياس في الحياة اليومية بمعنى العملية التي تحصل بها على تقديرات دقيقة للأشياء مما يؤدي إلى ضبط التعامل بين الناس.

تعريف القياس إجرائيا [1]: هو العملية التي تحدد بواسطتها كمية ما يوجد في الشيء من الخاصية أو السمة. من الصعب التفكير في أي صفة لا يمكن تكميمها. والقياس عبارة عن تحويل وصف الظواهر إلى ما هو أسهل من حيث التعامل ألاَ وهو الرقم.

والأرقام أساس عملية القياس إذ أنها الوحدات البنائية التي عادة ما تستخدم في تكوين أي نظام قياسي من أجل التقدير الكمي لأي ظاهرة من الظواهر. والأرقام التي نتعامل معها دائما في القياس لها نزعتان هما التمركز أو التشتت وما نقيسه هو صفة أو سمة في الشيء وليس الشيء ذاته.

القياس في التربية و التعليم أقدم كثيرا من القياس في علم النفس، حيث يشير القياس في اختبارات التحصيل إلى الرقم أو الرمز الذي يتحصل عليه الطالب. والقياس إما أن يكون قياسا مباشرا مثلا طول قطعة قماش او قياس غير مباشر مثل قياس درجة حرارة الفرد أو التحصيل الدراسي أو **الا**ستعدادات و القدرات أو الاتجاهات. فالقياس هو تكميم مقدار ما

[1]- التعريف الإجرائي : هو ما يمكن نقله إلى حيز الواقع.

يملكه شخص معين من سمة[2] أو خاصية بعينها. و في كل الأحوال يبقى وسيلة وليس غاية.

عناصر عملية القياس :

1- فئات خصائص الأشياء :

عندما نقوم بقياس، فإننا نتعامل مع أشياء معينة بهدف تقديرها كميًّا. والأشياء ليست هي المعنية في حد ذاتها بالقياس، بل خصائص معينة فيها. فالفرد ليس هو المعني بالقياس، بل خصائص فيه كالذكاء أو التوتر أو التعاون وغيره، كما تكون خاصية الوزن أو الكثافة في المواد الفيزيائية، أو خصائص الهجرة أو العدوان أو الخصوبة في الحيوانات، وفي كل الحالات فإننا نتعامل مع فئات خصائص الأشياء. ويتعين أن تكون متجانسة حتى يمكن الحصول على قياس جيد لها.

2- فئات الأعداد :

العنصر الثاني في عملية القياس، هو فئة الأعداد التي نتعامل بها لنعيِّن كميات فئات خصائص الأشياء، لكن ينبغي أن نعرف أن استعمالنا للأعداد قد يختلف في معناه من حالة إلى أخرى. فهناك فرق بين استعمال رقم 30 مثلا بوصفه دالا على رقم منزل، وبين استعماله بوصفه دالا على عدد من جرامات تشير إلى وزن شيء معين. وهناك فرق بين استعمال الرمز 30 مرة بوصفه رقما ومرة أخرى بوصفه عددا ومرة بوصفه دالا على ترتيب طالب بين مجموعة الطلبة في امتحان معين ؛ ففي المرة الأولى استعمل كرمز لشيء، وفي المرة الثانية استعمل للإشارة إلى كمية شيء وفي المرة الثالثة استعمل للإشارة إلى ترتيب شيء.

[2] - السمة هو مفهوم افتراضي غير قابل للملاحظة مباشرة مثلا الذكاء أو الاضطرابات الشخصية وهي مجموعة من السلوكات المترابطة التي تمثيل إلى الحدوث معا.

3- قواعد المقابلة بين فئات الأشياء وفئات الأعداد :

بما أن استعمال الأعداد يتضمن معان متعددة كما بيّنّا، يصبح من الضروري أن تتوفر في القياس الجيد قواعد معينة للتقابل بين خصائص الأشياء والأعداد. وكما رأينا فإنه يمكن أن نقصر الأعداد على التعبير عن الكميات، إلا أننا نستطيع أن نجعلها تعبر عن رموز الأشياء أو تعبر عن ترتيب يبين تسلسل الزيادة أو النقصان في الكميات. فتشير الأرقام التالية : 1، 2، 3، 4، 5 إلى تدرجهم في الطول على التوالي، بحيث يكون الأول أولهم والخامس أقصرهم، دون أن يكون الفرق بين الأول والثاني مساويا للفرق بين الثالث أو الرابع والخامس. فالأعداد 1، 2، 3،4،5 تشير هنا إلى ترتيب وليس إلى كميات الخاصية (مقدار الطول) ؛ أما الأرقام 184، 181، 172، 165، 161 فتشير إلى كميات الخاصية (الطول)، وعندما نحدد قواعد استعمال الأعداد ونعيّن خصائص الأشياء بها، تصبح عملية القياس صحيحة. (بشير معمرية، 2007: 16)

العلاقة بين العدّ والقياس :

_هناك فرق جوهري بين العد والقياس، فالعد يكون دائما بوحدات كاملة ويتصف بالدقة. فنقول: ثلاثون طالبا وعشرون فردا وأربعون شجرة، والبيانات التي نحصل عليها بالعد نتعامل معها إحصائيا على شكل تكرارات أو نسب مئوية. أما القياس فهو تقريبي دائما، تنتج عنه قيم قياسية تمثل الظاهرة المقاسة بدرجة تقريبية، وهذا التقريب يعتمد على دقة أداة القياس المستعملة. فلو قمنا بقياس طول بعد(ارتفاع جدار مثلا)، فإن نتيجة القياس تكون تقريبية وتتجزأ إلى وحدات صغيرة. فلو كان طول هذا البعد عشرة أمتار، فإن طوله إما أن يكون أقل من عشرة أمتار بسنتمترات أو بملمترات، أو أكثر بسنتمترات أو بملمترات. ونفس الشيء إذا وزنت ثقلا فإذا كان وزنه عشرين كيلو غراما فهو يزيد عنها ببضع غرامات أو يقل عنها ببضع غرامات وهكذا فالقياس بما أنه تقريبي، فإنه يتضمن دائما كسورا. ولهذا فإذا

قلنا أن طول الفرد هو 150 سنتمتر فمعنى هذا أن طوله يتراوح بين 149.50 – 150.49 سنتمتر.

– هناك فرق آخر بينهما وهو ما يلي : وحدات العد تكون متقطعة لأن الفرد يعُدُّ أشياء تكون منفصلة عن بعضها، ولهذا فالمعطيات التي يقدمها العد لا تعتبر ذات طابع كمي. أما وحدات القياس فهي متصلة ومستمرة فلا فجوات بينها، لأن الفرد يقيس خاصية واحدة، ولهذا يعتبر وحداته كمية.

أما العلاقة بينهما، فإن العد يرتبط بالقياس ارتباطا وثيقا؛ ففي أحيان كثيرة نتعامل مع وحدات العد كأنها وحدات قياس، فعدد الأطفال في الأسرة عدد صحيح، إلا أن الإحصاءات السكانية تشير إلى أن معدل عدد الأطفال في الأسرة الجزائرية هو 5.6 طفل. وفي أحيان أخرى فإننا عندما نقيس، فإننا نعد أيضا، فعند قياس طول فرد، فإننا نعد الوحدات التي يساويها طوله وهي السنتمترات ويقوم بذلك شريط القياس (المتر). وكذلك عندما نمتحن متعلما باختبار مقالي فإننا نقوم بقياس معلوماته، ونمثل مقدارها بعدد من الدرجات تكون عادة عددا صحيحا والدرجات عبارة عن قياسات وهي تقريبية فإذا كانت درجة المتعلم 15 فهي تتراوح بين 14.50 – 15.49. (بشير معمرية، 2007: 16)

العلاقة بين التقدير والقياس:

مفهوم التقدير Rating

التقدير أيضا قديم قدم الإنسان، وهو أقدم من القياس ومرادفا للتخمين أي تحديد الشيء بالانطباع الشخصي أو بالحدس أو بالظن أو الوهم، فيقال : قدر الشيء أي بيّن مقداره، والمقدّر أو المخمّن هو الشخص الذي يبين مقدار الشيء وما يترتب عليه من تبعات.

وفي حياتنا اليومية يقوم الناس بتقدير الأشياء والظواهر، فيقدرون المسافات والأحجام والأطوال والأوزان وغيرها، فيرفع الفرد كيسا مملوءا فيقدر وزنه، من خلال الإحساس العضلي بأنه يساوي كذا. والفلاحون يقدرون الوقت بالنظر إلى موقع الشمس في السماء.

ويقوم الناس أيضا بتقدير الأشياء التي لم يتم تحديد خصائصها بالأرقام للتعرف على خصائصها الكيفية فيها، فيبحثون عن مستوى الجودة فيها، فيصفونها بأنها أكثر أو أقل جودة من شيء آخر، أو يبحثون عن مستوى الدقة أو الصحة، ليحكموا أن شيئا ما أدق من شيء آخر أو أصحّ من شيء آخر. وكما نرى فإن استخدام الناس لهذه التعبيرات النسبية، مثل أكثر من وأقلّ من، أو أدقّ من، أو أصح من، لا يتضمن تقديرا كميّا محددا، وهي تشبه وقوف الناس أمام لوحتين فنّيتين، فيقارنون بينهما ليصلوا إلى أن إحداهما أكثر جمالا من الأخرى، دون أن يضعوا أي تقدير كمّي لجمال كلٍّ منهما أو تقدير لفارق كمّي في الجمال بينهما.

مجالات استعمال التقدير:

يستعمل التقدير بأدواته المختلفة في مجالات عديدة، فهو يستعمل كمحكات لإثبات صدق أدوات قياس أخرى، كما يستعمل لجمع ملاحظات عن أداء أو منتوج أو طريقة في العمل لا نستطيع قياسها، لأننا لم ننجح في تصميم أدوات قياس أدق وأكثر موضوعية من أدوات التقدير. ونذكر فيما يلي بعض المجالات التي يستعمل فيها التقدير بأدواته المختلفة، وهي كما يلي:

- تقدير الأداء: في الدراسة، التدريس، العمل، الغناء، التمثيل، الرياضة وغيرها.
- تقدير الإنتاج: في الصناعة والزراعة والبناء والرسم والكتب والخياطة والتدبير المنزلي وغيرها.
- تقدير الطريقة : في القيام بالتجارب وتصليح السيارات وإدارة اجتماع وإنجاز بحث والتدريس وغيره.
- تقدير السلوك السوي/اللاسوي: التكيف مع مجالات الحياة، اكتساب عادات سوية، إتقان المهارات الاجتماعية ... إلخ.
- تقدير الذات: تقدير سمات الشخصية، تقدير الميول، الاتجاهات، المهارات، القدرات وغيرها.

- تقدير المناهج الدراسية: تقدير المحتوى الدراسي، تقدير الأهداف التربوية، تقدير الوسائل التعليمية تقدير طرق التدريس وغيرها.

- تقدير المؤسسة كالمصنع والمدرسة والإدارة وغيرها: من حيث الموقع، ومواصفات البناء والتجهيزات، والأبواب والنوافذ والكراسي والطاولات ولون الطلاء وغيره.

العلاقة بين التقدير والقياس:

هناك علاقة بين التقدير والقياس كونهما أدوات لجمع معلومات عن الفرد محل الدراسة، كما يشكلان أيضا أساسا للتقويم، إلا أنه توجد بينهما فروق نذكرها فيما يلي:

- إن التقدير لغة قد يعني القياس، إلا أنه يختلف عنه من الناحية العملية، فهو تحديد للصفة أو الخاصية بطرق الحدس أو الانطباع الذي يكون لدى الإنسان، وهو أقل موضوعية ودقة من القياس.

- أدوات التقدير تمدّنا بأساس انطباعي ذاتي، بينما أدوات القياس موضوعية.

- أدوات التقدير كيفية في أغلبها، بينما أدوات القياس كمية.

- أدوات التقدير يستعملها المقدر(الفاحص)غالبا؛ بينما أدوات القياس يستعملها المفحوص نفسه. (بشير معمرية، 2007: 33- 43)

علاقة بين التقويم والقياس:

-القياس وسيلة التقويم وهي سابق له وأساس له.

- مفهوم التقويم أعم وأشمل من القياس (القياس هو وصف الظاهرة والتقويم إصدار الحكم عليها من خلال الوصف).

- كل قياس لابد من أن يتضمن تقويما ولكن ليس من الضروري أن يتضمن كل تقويم قياسا.

- يقتصر القياس على التقدير الوصفي الكمي للسلوك أما التقييم يشمل التقدير الكمي والنوعي للسلوك إضافة الى الحكم.

– يوفر القياس وصفا أكثر موضوعية للسمات من التقويم ويسير المقارنات ولكنه أقل منه قيمة من الناحية التربوية والعلاجية.

– القياس أضيق في معناه من التقويم، لأنه مع دقته لا يعطينا سوى فكرة جزئية عن الشيء الذي يقاس.

– يتناول السلوك الملاحظ وليس كل السلوك فهو يتناول عينة محدودة.

–الأرقام هي لغة القياس الوحيدة ؛ بينما التقويم يسمح بلغة الكيف إلى جانب لغة الكم.

– يرتكز التقويم على مقارنة الفرد بنفسه وبغيره، بينما يعطينا القياس نتائج وصفية للشيء دون ربطه بأشياء أخرى.

–يعطي التقويم لدرجة القياس معنى، فالدرجة التي نحصل عليها من القياس ليس لها في حقيقة الأمر قيمة في ذاتها. (سامي محمد ملحم، 2015: 66–67)

الفرق بين القياس، التقويم ، التقييم و التقدير

القياس	التقويم Evaluation	التقييم Assessement	التقدير Jugement
– يقوم علم النفس القياسي على أساس إمكانية قياس مختلف الخصائص النفسية فكما يقول ثورندايك (Thorndike) أحد رواد علم النفس:"إذا وجد شيء يوجد بمقدار، وإذا كان يوجد بمقدار فإنه يمكن قياسه". – لها عدة معان تصب في اعتبارها اسما للتعبير عن عملية القياس وعلى نتائج القياس وعلى الأدوات المستخدمة ومن وحدات	لغويا: مشتق من فعل قوم ويعني إصلاح الاعوجاج بناء على نتائج القياس والتقدير. –هو أعم وأوسع فهو فحص، حكم و اتخاذ القرارات العلاجية الملائمة اصطلاحا: إنه فعل مقصود ومنظم يؤدي إلى إنتاج حكم قيمة. يقوم على جمع المعلومات، الحكم واتخاذ القرار. أي مقارنة بين حالتين ماهو كائن وما يجب أن يكون وتؤدي إلى اقتراح التعديلات التي يجب القيام بها أي تقيس الفارق الموجود بين النتائج المحققة والأهداف المسطرة والكشف عن أسباب هذا الفارق وبالتالي تسمح لنا بتصور التصحيحات الملائمة.	– التقييم مشتق من فعل قيّم وتعني لغويا إعطاء مقدار. – يهدف التقييم إلى إصدار حكم كمي أو كيفي عن شيء من أجل اتخاذ قرارات دون البحث في سبب حدوثه. هو مجموعة من العمليات التي تستخدم	– هو أقدم من مفهوم القياس– يقوم على التخمين الذاتي. نقول لغويا قدر الشيء أي بيّن مقداره بناءًا على تخمينات شخصية مثلا: أقل من – أكبر

القياس أو بوصفها فعلا للتعبير على عملية تقدير المدى أو الفترة أو البعد أو كمية الأشياء. –له عدة تعاريف مثلا: تعريف جيلفورد Guilford(1973): نسب مقادير عددية (أرقام) إلى الأشياء وأحداث وفق قواعد منطقية. هو تقدير الأشياء أو الظواهر تقديرا كميا وذلك باستخدام وسائل القياس المناسبة لخصائص الظاهرة المقاسة كقياس الوزن بواسطة (غ). نواتجه كمية – يسبق التقييم والتقويم.	هدفه: العمل على تدعيم النتائج تدارك النتائج السلبية عن طريق حصص علاجية مناسبة. هو الأنسب في الميدان التربوي. يرمي بالدرجة الأولى إلى تدارك النقائص ومعالجتها بعد الكشف عنها. التقويم أشمل يشملهم الأربعة. إنه فعل مقصود ومنظم يؤدي إلى إنتاج حكم قيمة. يقوم على جمع المعلومات، الحكم واتخاذ القرار. أي مقارنة بين حالتين ماهو كائن وما يجب أن يكون وتؤدي إلى اقتراح التعديلات التي يجب القيام بها أي تقيس الفارق الموجود بين النتائج المحققة والأهداف المسطرة والكشف عن أسباب هذا الفارق وبالتالي تسمح لنا بتصور التصحيحات الملائمة.	بواسطة أخصائيين متمكنين للوصول إلى وضع تصورات وانطباعات حول فرد معين واتخاذ قرارات واختبار فروضه تتعلق بنمط خصائص سلوكه وتفاعله مع بيئته. –يستخدم المختصون في عمليات التقييم أدوات قياس متعددة (الذكاء، التحصيل، الميول..).	من – أكثر من – ضخم – هزيل ..نواتجه كيفية يسبق التقييم والتقويم.

القياس و التقدير(نتائجه كمية) يسبقان التقييم و التقويم(نتائج كيفية)

أهداف عملية القياس:

1.التصنيف: هو تحديد أحد الاختبارات الأكثر مناسبة للفرد أو تحديد الفرد في وضع دون آخر.

2. التشخيص: وهو إبراز جوانب القوة والضعف لدى المفحوص من أجل توظيفها في برامج العلاج.

3.المعرفة الذاتية: حصول التغذية الراجحة نتيجة الاختبار يكون أساس تعديل المفحوص لسلوكه وتغييره.

4. التقييم و التقويم: تعد الاختبارات وسيلة من وسائل تقييم السمات السلوكية من أجل الوقوف على تشخيصها و دراسة فعاليتهاو علاج أي خلل على مستوى وظائفها.

5.التنبؤ.

6. المتابعة.

العوامل المؤثرة في القياس :

1.السمة أو الخاصية المراد قياسها : الخاصية التي تقاس بطريقة مباشرة تكون أكثر دقة من الطريقة غير المباشرة.

2.نوع المقياس ووحدة القياس المستخدمة فيه: مثلا الميزان الاليكتروني أكثر دقة من الميزان العادي، مثلا اختبار الذكاء أكثر دقة من مقاييس الشخصية لأن شخصية الفرد معقدة جدا.

فكلما تقدمت أساليب القياس في علم من العلوم كان ذلك مظهرا من مظاهر رقي هذا العلم وتقدمه.

3.أهداف القياس : إذا كان هدف القياس اختيار واحد من عدد كبير من المتقدمين لوظيفة ما، فإن بنود المقياس ستكون صعبة جدا.

4. مدى تدريب وخبرة القائمين على عملية القياس.

5. طبيعة المقياس وعلاقته بالظاهرة المقاسة: كلما كان المقياس أكثر ملائمة للظاهرة المقاسة كانت النتائج أكثر دقة، فالاختبارات الكتابية على سبيل المثال لا تقيس مهارة السباحة والخطابة.

مصادر الخطأ في القياس : كأن يكون القائم بالقياس دقيقا فلابد أن يقع في الخطأ وهي ناجمة عن :

1. عدم حساسية أو دقة القياس : كلما كانت السمات واضحة كانت أدوات القياس دقيقة ومناسب وكلما كانت الأداة دقيقة ووحداتها صغيرة كان حجم الخطأ أقل.

2. عدم ثبات الظاهرة المقاسة : مثلا قلق الاختبار : كأخطاء الفرد المستخدم لأداة القياس. خطا العينة.

الدرس الثاني: القياس في العلوم الاجتماعية

- طبيعة وخصائص القياس في العلوم الاجتماعية:

يختلف القياس باختلاف الظاهرة المقاسة، ففي الفلك والفيزياء والكيمياء مثلا، يكون القياس أدق ما يمكن ولا يقبل بوجود أخطاء مهما صغرت.

وإذا انتقلنا إلى ميدان علم الأحياء، فإن القياس يصبح أقل دقة، لأنّ هناك عوامل مؤثرة أكثر عددا وقابليتها للتغير أكثر حدوثا؛ أما إذا انتقلنا إلى العلوم الاجتماعية فإن القابلية للتغير تكون أكثر، مما يجعل القياس أقل دقة، وأكثر عرضة للخطأ وتسامحا في قبوله، ويمكن أن نتجاوز عنه إذا كان لا يزيد عن 5 %، إلا أنه يجب أن ننتبِه إلى هذه الأخطاء وأسبابها، كي نستعمل أدوات القياس استعمالا ذكيا، ونفسر نتائجه تفسيرا صحيحا، ولهذا يقتضي أن نلمّ بخصائص القياس في العلوم الاجتماعية وهي كما يلي :

-الأرقام أساس عملية القياس، إذ أنها الوحدات البنائية التي عادة ما تستخدم في تكوين أي نظام قياسي من أجل التقدير الكمي لأي ظاهرة من الظواهر، والأرقام التي نتعامل معها دائما في القياس لها نزعتان هنا التمركز أوالتشتت.

-ما نقيسه هو صفة أو سمة في الشيء وليس الشيء ذاته.

-الخاصية الاجتماعية لا يمكن قياسها قياسا مباشرا فإننا نلجأ إلى مؤشرات ليست هي السمات في حد ذاتها ولكنها تشير إلى مدى توفر السمة في الإنسان مثلا : معرفة الحرارة أو التيار الكهربائي عن طريق علامات أخرى ليست هي الكهرباء.

-القياس وسيلة وليست غاية.

-القياس في ميدان التربية أقدم بكثير من القياس في علم النفس و من قياس باقي الظواهر الاجتماعية.

-القياس في أول الأمر كان يعتمد على الملاحظة الذاتية والآراء الشخصية ثم مرّ بعد ذلك بثلاث مراحل:

أ- مرحلة الاختبارات الشفوية.

ب- مرحلة الاختبارات التحريرية المقالية.

جـ - مرحلة الاختبارات التحريرية الموضوعية.

- **طبيعة و خصائص القياس في علم النفس:**

تعريف القياس في علم النفس أو القياس النفسي: Psychométrie

إن تطور وتقدم أي علم من العلوم إنما يقاس بقدرة هذا العلم على تطوير واستخدام أدوات قياسه. الخاصية النفسية لا يمكن قياسها قياسا مباشرا، نحصل على وجودها بمؤشرات أخرى فمثلا، التيار الكهربائي نتعرف على وجوده من تشغيل الآلات الكهربائية الموجبة الترمومترية و قياس الحرارة نتعرف عليها من خلال ارتفاع الزئبق.

تختلف العلوم الاجتماعية عن العلوم الطبيعية في الخصائص والسمات محور الاهتمام حيث تكون خصائص الأشياء في العلوم الطبيعية واضحة محسوسة يمكن قياسها مباشرة على عكس العلوم الاجتماعية تشكل مفاهيم افتراضية مجردة لا يمكن ملاحظتها وقياسها بشكل مباشر.

بالاضافة إلى أن القدرات والسمات تختلف من فرد لآخر أي أن الأفراد يختلفون في درجة إمتلاكهم لسمة معينة وتختلف أيضا في الفرد الواحد، هذا أساس الفروق الفردية: التفرد، التميز والخصوصية.

ان القياس النفسي يقيس الظاهرة النفسية التي تتميز بالتعقيد والتعددية المتغيرة. و هو عبارة عن وصف البيانات في صورة رقمية. القياس عبارة عن الأداة التي تستخدم في رصد الظاهرة السلوكية بصورة إجرائية، و عليه يعرف بأنه الوسيلة أو الأداة التي تستخدم في قياس عينة ممثلة من السلوك.

هو ذلك الفرع من علم النفس الذي يهتم بقياس مظاهر السلوك والتوصل إلى تقدير كمي أو كيفي لهذه المظاهر بالتحديد الدقيق للظواهر والعمليات النفسية : الذكاء، القدرات، سمات الشخصية، الميول، الاتجاهات، المعتقدات والقيم؛ و تقديرها كميا أو كيفيا.

القياس النفسي والقياس الطبيعي

القياس النفسي	القياس الطبيعي
-يتميز بالتعقيد والغموض والتعددية. -يتم القياس بطريقة غير مباشرة. - يستخدم الفهم لتفسير الظاهرة النفسية. - يحتوي على العديد من المعاني.	-يتميز بالثبات والاستقرار. - سهل التعامل معه. -يتعامل مع الظاهرة الطبيعية من خارجها. - القياس بطريقة مباشرة. - يستخدم التفسير لأن الظاهرة الطبيعية تحتوي على العديد من الحقائق.

إن رواد علم النفس في القرن 19 عملوا على تطوير علم النفس الفيزياء تشبها بالفيزياء وذلك بتبني التجريب والقياس في دراسة بعض الظواهرالنفسية مثل عتبات الإحساس وزمن الرجع. (عبد الهادي السيد عبده و فاروق السيد عثمان، 2006: 22)

طبيعة وخصائص القياس في علم النفس:

- القياس هو عبارة عن تقدير كمي (رقمي) لبعد من أبعاد السلوك الإنساني وهذا التقدير يعبر عن مستوى الأداء في الصفة أو السمة موضع الاهتمام.

- القياس النفسي قياس غير مباشر، فعالم النفس لا يقيس الذكاء بعينه وإنما يقيس ما يدل عليه وهو السلوك الذي يؤدّى، يكما يستدل على التحصيل الدراسي من خلال نتائج الاختبارات؛ ومثلا : تقيس التعلم بحساب الزمن المستغرق لإنجاز العمل ونقص في عدد

الأخطاء ونقيس التذكر من خلال الزمن المستغرق للتعرف أو لاستدعاء المعلومات ونقيس الانتباه.

-تواجد أخطاء في القياس النفسي تعزى إلى أحد المصادر :

أ-أخطاء الملاحظة: من المعلوم أن الاشخاص يختلفون في دقة الملاحظة لظاهرة ما ويرجع ذلك إلى ظاهرة الفروق الفردية بينهم لذلك يفضل عند قياس أي ظاهرة أن يتم الرصد عدة مرات حتى تثبت قراءتان متتاليتان.

ب-عدم حساسية الأدوات في قياس الظاهرة: أي في قياس الصفات والسمات لعدم وجود الصفر المطلق اضافة الى أن أدوات القياس النفسي تحتاج إلى مهارة سواء في إعدادها أو في طريقة تطبيقها.

جـ- عدم ثبات الظاهرة المقاسة : السلوك البشري سلوك معقد يتضمن جوانب معرفية انفعالية وجدانية، مهارية، نفسية، حركية،...الخ. كما أنه يتغير بتغير الموقف سواء بسبب الظروف البيئية الفيزيقية أو الاجتماعية أو النفسية.

ملاحظة : من الجدير بالذكر أن نسبة الخطأ المسموح به عند قياس الظواهر السلوكية أو التربوية أو الاجتماعية تتراوح بين 1 % - 5 % أي أن 1%شك و 99 % ثقة أو 5 % شك و95 % ثقة والتجاوز عن هذه النسبة يشير إلى وجود خطأ في عملية القياس يجب تداركه ومراجعته.

-القياس النفسي نسبي أي غير مطلق، لا تتوفر في القياس النفسي وحدات محددة للقياس بينما تتوفر في القياس الفيزيائي. فلا توجد وحدات نفسية مماثلة للسنتمتر أو للصفر، وعليه لا يمكن تفسير نتائج القياس بمقارنتها بمعيار أو بمستويات مشتقة من أداء المفحوصين أو أداء عينات مماثلة لعيّنة المفحوصين، أي أن نتيجة القياس لا يمكن فهمها وتفسيرها إلا بعد نسبتها إلى نقاط مرجعية. فإذا حصل فرد على الدرجة 36 على استبيان يقيس سمة شخصية، فإن هذه الدرجة لا تعني شيئا. أما إذا عرفنا أن متوسط العينة التي أجابت على عبارات نفس الاستبيان هو: 29 فإن هذه المعرفة تكسبها معنى. وإذا عرفنا كذلك أن 20 % من الدرجات تقع فوقها، فإن هذه المعرفة تكسبها معنى آخر، وهذه المعاني

لم يمكن الحصول عليها إلا بإرجاع درجة الفرد إلى معايير وعلامات العينة التي كانت تنتمي إليها.

-رغم أن وحدات القياس النفسي (الدرجات) توصف بأنها غير متساوية حقيقة، إلا أننا نستطيع إيجاد وحدات تتصف بالثبات والتساوي مثل الدرجات المعيارية.

-الصفر في القياس النفسي صفر اعتباطي، فهو ليس بصفر حقيقي، أي لا يدل على عدم وجود الخاصية، فإذا حصل فرد على درجة صفر في استبيان للقلق مثلا، فلا يعني ذلك أنه خال تماما من القلق، وإنما هناك عوامل أخرى جعلته يحصل على هذه الدرجة، فقد يكون قد اختار أن يجيب بهذا الأسلوب (تزييف الاستجابة)، أو أن عبارات الاستبيان غير واضحة، أو أنها صيغت بطريقة جعلته يستجيب لها بالإنكار... إلخ، لأنه لو يتم استبدالها بعبارات أخرى فإنه من المحتمل أن نعتبر الدرجة، وبناء عليه ينبغي أن نعتبر الصفر في القياس النفسي كالصفر في مقياس الحرارة، فهو يدل على وجود درجة ما للخاصية، ولا يدل على عدم وجود الخاصية ولما كان هذا الصفر اعتباطيا يمثل نقطة أصل، أو استناد، أو نقطة مرجعية، فيستحسن أن نتخذ متوسط الدرجات كنقطة استناد بدلا منه، ونقيس درجة الفرد ببعدها عنه أو قربها منه عندما نريد تفسيرها.

-إن القياس النفسي لا يهدف إلى قياس الفرد برمته وإنما يهدف إلى قياس سمة من سمات الفرد أو خاصية من خصائصه كالذكاء والدوافع والانبساط والانطواء...الخ. والسمة ليست صفة منفردة وإنما تجمع العديد من الصفات المترابطة مثلا الذكاء يتضمن كل من: الفهم الاستدلال، التذكر واللغة ...الخ.

-إن قياس السمات والخصائص النفسية في غالب الأحيان يكون قياسا نسبيا وليس قياسا مطلقا فهو غير تام: لا يمكن حصر جميع السلوكات التي تدل على السمة والاتفاق حول تلك السلوكات غير ممكن، مثل ما هو موجود في العلوم الفيزيائية وذلك لعدة أسباب :

*/ يصل مستوى القياس في علم النفس في أحسن الحالات إلى مستوى المسافات المتساوية ومن النادر أن يصل إلى مستوى النسبة لأن الصفر في مستوى المسافات

المتساوية هو صفر اعتباطي وليس حقيقي مثلا: إذا حصل مفحوص على درجة (0) في الرياضيات فهذا لا يعني أن الطالب ليس لديه أي معارف في الرياضيات نفس المثال ينطبق على الذاكرة.

*/هناك عدد كبير من المتغيرات التي يمكن أن تؤثر على النتائج لذا يجب أن تضبط. على عكس علم الفيزياء يكفي فقط أن نحضر مسطرة لنقيس طول الطاولة.

*/مقارنة نتائج الطالب بما يسمى بجماعة المعيار مثلا إذا تحصل الطالب على 20/9 وكان مستوى جماعة قسمه يدور في معدل 20/10 فهي نتيجة حسنة وإذا كان أعلى من 20/16 فهي تحت المتوسط.

*/إن القياس في علم النفس معرض لوجود نسبة من احتمالات الخطأ في نتائجه فهو غير دقيق تماما، تعود إلى عدة أسباب منها:

**/التأثيرات السلبية لمتغيرات دخيلة (في حالة عدم ضبطها).

**/ تقنيات القياس الضعيفة التصميم.

**/حدوث تغير في السمة موضوع القياس ناتج عن النمو أو تغير في العوامل المحيطة.

**/ يقوم القياس على وجود العديد من الفروق الفردية في سلوك الإنسان وقدراته وظروفه وعليه فأساس القياس هو وجود الفروق.

–وجود خلل في صدق وثبات أدوات القياس النفسي: يتمثل صدق أداة القياس في أنها تقيس ما وضعت من أجله فقط، فالمتر لا يقيس إلا الطول. حيث لا نشك بصدق هذا المقياس كما أن هذا المقياس يتمتع بدرجة عالية من الثبات حيث تعطينا نفس القراءات إذا ما قمنا بعملية القياس عدة مرات والأمر مختلف في علم النفس.

– أدوات القياس النفسية والتربوية غير مقننة عالميا أي أنها غير معيرة بينما أدوات القياس الفيزيائية مقننة ومعيرة.

– القياس في العلوم النفسية وسيلة وليس غاية، تقوم لأعراض المسح، التنبؤ، التشخيص والعلاج. فمثلا يساعد المعلم والمرشد النفسي والتربوي والباحث على فهم السلوك الإنساني ويساعد صانع القرار عليه.

– يتمتع القياس النفسي بدرجة من الثبات هي أقل من الثبات في القياس الفيزيائي.

–يتضمن القياس النفسي عدة مجالات :

*–القياس في المجال التربوي.

*– القياس في المجال النفسي : الجانب العقلي، الجانب المعرفي، الانفعالي، الجانب الوجداني، الميول والاتجاهات والقيم والمعتقدات، والجوانب النفسية الحركية...الخ.

*– القياس في المجال الاجتماعي : التفاعل الاجتماعي، ديناميكية الجماعات، الاتجاهات...الخ.

*– القياس في المجال الصناعة : عملية الانتقاء والاختيار، القدرات والاستعدادات...الخ.

*– القياس في المجال الخدمة العسكرية : انتقاء الجنود لنوع السلاح، الجوي أو المشاة أو البحري، قياس الروح المعنوية للجنود وقياس الإشاعات...الخ.

العوامل المؤثرة في القياس :

–الخاصية التي نقيسها بطريقة مباشرة يكون قياسها أكثر دقة من الخاصية التي نقيسها بطريقة غير مباشرة.

–كلما كانت السمات واضحة ومعروفة، كانت أدوات القياس دقيقة ومناسبة وكانت النتيجة التي تخرج بها قابلة للفهم والتفسير بيسر وسهولة.

–نوع القياس ووحدة القياس المستخدمة فيه، لأن الشخصية غاية في التعقيد، فكلما تقدمت أساليب القياس في علم من العلوم كان ذلك مظهرا من مظاهر رقي ذلك العلم

وتقدمه. فالميزان الكتروني أكثر دقة من الميزان العادي و مقياس الذكاء أكثر دقة من مقياس الشخصية.

صعوبات قياس البنى النفسية (الافتراضية):

إن ما نهتم به هو مجموعة الصفات والمفاهيم المجردة التي لا يمكن قياسها مباشرة :

1-عدم الاتفاق على أسلوب قياس واحد لأي مفهوم (إن التعريفات الإجرائية للمفهوم الواحد سوف تتعد بتعدد وجهات النظر ومن هنا فإننا سنجد أنفسنا أمام عدد من الاختبارات التي تقيس مفهوما واحدا).

2-الاعتماد على قياس عينات من السلوك: نتعامل مع عينة من المجتمع الإحصائي.

3-عملية القياس النفسي دائما عرضة للخطأ : الاعتماد على عينات محددة في فترة زمنية محددة مثلا لنفرض أن طلبة امتحنوا في نفس الاختبار مرتين متتاليتين من المتوقع أن لا يكون نتائجهم في المرتين متطابقة، لعدة أسباب منها : التعب، الملل، النسيان والخطأ ..

4-عدم القدرة على تحديد وحدات القياس لأدوات القياس: لنفرض أن طالبا تحصل على علامة (0) هل يعني أن هذا الطالب ليس لديه أي معارف أو مهارات ؟ وهل الفرق بين علامات الطلبة، تفصل بينهم نفس المسافة على مقياس القدرة؟

5-تداخل المفاهيم النفسية : من المهم تعريف البناء النفسي(المفهوم) على مستويين:

-يجب أن يعرف في ضوء السلوكات الملاحظة.

-يجب أن يعرف في ضوء علاقاته الرياضية والتطبيقية مع المفاهيم الأخرى ضمن النظام النظري. فإذا لم يكن من الممكن توضيح هذه العلاقات تجريبيا فإن القياسات تصبح عديمة المعنى والقيمة.

ما الذي نقيسه في علم النفس؟

يتناول القياس في علم النفس عددا كبيرا من المتغيرات والعمليات والاستجابات السلوكية التي تقع في مجال تخصص هذا العلم ولتوضيح ذلك نتناولها فيما يلي :

1-**قياس العمليات المعرفية:** يقيس المتخصص في علم النفس الذكاء والقدرات العقلية، ويميز بينها كميا تمييزا دقيقا، فهو يقيس الذكاء ويميز فيه بين خصائصه المختلفة؛ فيميز بين الذكاء اللفظي والذكاء الأدائي والذكاء الاجتماعي، ويميز بين الذاكرة قصيرة المدى والذاكرة طويلة المدى، ويميز بين التفكير التقريري والتفكير التغييري كما يقيس قدرة الفرد على الاستدلال والتجريد والتعميم والتمييز وتكوين المفاهيم وغيره.

2-**قياس سمات الشخصية:** يقيس السيكولوجي كذلك السمات المزاجية للشخصية، ليحصل على مقادير كمية لوجود هذه السمات، فيقيس سمات مثل :التعاون، السيطرة، التوتر، الانطواء والعصبية الاجتماعية وغيرها. هناك افتراضات أساسية وراء القياس النفسي لسمات الشخصية، تقوم على اعتبار وجود سمات مشتركة متشابهة لدى جميع الأشخاص مثلا: الاجتماعية، الخجل، الاكتئاب، التفاؤل والقلق ...الخ. و يمكن تقديرها كميا. و الاهتمام بالطريقة التي يدرك الشخص بها نفسه وليس الطريقة التي يدركه الآخرون بها.

3-قياس المهارات الحركية: ويقيس السيكولوجي أيضا خاصية الدقة في أداء أعمال حركية، والسرعة في الحركة وغيرها.

4-قياس الاضطرابات النفسية والعقلية: و يقيس السيكولوجي أيضا مظاهر السلوك اللاسوي التي يبديها الأفراد، فيقيس القلق والاكتئاب والوسواس القهري والفصام والبارانويا وغيرها.

5-قياس الدوافع والاتجاهات والقيم والميول: يقيس السيكولوجي أيضا الدوافع عند الأفراد كالدافع إلى العمل والدافع للإنجاز والدافع للدراسة وغيرها. و يقيس السيكولوجي أيضا الاتجاهات نحو القضايا المختلفة، كالاتجاه نحو العمل والاتجاه نحو الالتحاق بالدراسة الجامعية والاتجاه نحو الإنجاب وغيره. و يقيس السيكولوجي أيضا الميول المهنية والميول الدراسية و يقيس السيكولوجي أيضا القيم ليتعرف على السائد منها والفروق بين الأفراد فيها كالقيم الدينية والقيم الخلقية والقيم الاقتصادية وغيرها.

6-قياس التفكير الإبداعي وسلوك حل المشكلات: يتولى السيكولوجي أيضا قياس القدرات الإبداعية وتحديد وجودها كميا، وكذا الفروق بين الأفراد فيها، كالطلاقة والمرونة والأصالة والحساسية للمشكلات وغيرها، و يقيس السيكولوجي أيضا استراتيجيات سلوك حل المشكلات، فيتعرف على الأساليب التي يسلكها مختلف الأفراد في سعيهم لحل المشكلات التي تعترضهم، والفروق بينهم في هذه الاستراتيجيات.

يتناول القياس النفسي كل هذه الموضوعات (و ما هي إلا نماذج من عديد من الموضوعات الخاصة بميادين علم النفس)، ويقدم لها تقديرات كمية تسهم في تحديدها بشكل دقيق إلى حد كبير ولا يستطيع الباحث السيكولوجي أو المعلم أو الموجه المدرسي والمهني أن يحصل على تقدير سليم لإمكانيات الأفراد وأدائهم وسماتهم ودوافعهم واتجاهاتهم دون قياس دقيق لها.

تصنيف كيلي لمجالات القياس النفسي (1974) وهي 08 أصناف :

أ-**السمات الجسمية** : الطول، الوزن، أبعاد الجسم، قوة القبض...

ب-**السمات الفيزيولوجية:** مختلف أنواع السلوك الداخلي لأعضاء الجسم الحية.

جـ- **الاستعدادات:** القدرات الكامنة عند الافراد: الاستعدادات الدراسية، الميكانيكية، الحسية الحركية، اللغوية والفنية .. ويمكن لها أن تتطور بالتكوين والتدريب.

د-**المهارات والتحصيل:** الأفراد يختلفون في المهارات المكتسبة من أجل أداء عمل وهي متصلة ومرتبطة بالاستعدادات وتبنى على ما هو متوفر من استعدادات.

هـ- **الميول** : تتصل بالجوانب الوجدانية للفرد وتعكس ما يحب وما يكره، والأنشطة التي يميل إلى تفضيلها كالمطالعة والموسيقى...

و- **الاتجاهات** : تتصل أيضا بالجوانب الوجدانية تعكس الآراء و الاتجاهات

ز- **القيـــم:** تتصل بالجوانب الـوجدانية الــعميقة كـالمعتقدات والـمبادئ والـمرجعيات الأخلاقية ...الخ.

ي- **السمات المزاجية:** السمات التي تدخل في فئة أنماط ونزعات وطبائع مثلا الانبساط، والانطواء، والتفاؤل، الثقة بالذات، والاندفاع، والدوافع...الخ.

و ضم تصنيف آخر أربعة (04) أصناف، و هي :

- السمات البدنية والفيزيولوجية.
- السمات الذهنية والمعرفية.
- السمات الشخصانية.
- السمات الميتامعرفية.

ملاحظة : لم يشمل تصنيف كيلي هذا البعد Métacognition لأنه لم يكن محل اهتمام الباحثين إلا مع بداية الثمانينات.

يقصد بمفهوم الميتامعرفي : المعارف التي نملكها حول سيرورة نشاطاتنا وقدراتنا الذهنية من جهة وقدراتنا على تقويم وتعديل وتنسيق هذه النشاطات أي معرفة المعرفية la cognition de la cognition

- **لماذا نقيس السلوك ؟**

رغم أن التعبير عن خصائص الأشياء تعبيرا كميا سمة من سمات العلم، إلا أنه لا ينبغي أن نقول أننا نقيس الخصائص السلوكية بهدف تحديدها كميا ثم نكتفي بذلك، فهذا التحديد الكمي ليس غاية في حد ذاته ولكنه وسيلة تستعمل لخدمة أهداف أخرى بعيدة.

ويمكن أن نحدد هدفين واسعين لعملية القياس، وما تتميز به من تكميم مقنن للقدرات والسمات والاتجاهات والدوافع والقيم وغيرها مما يدخل في مجال تخصيص علم النفس الذي نصل إليه في عملية القياس.

الهدف الاول: إن عملية القياس في جوهرها ملاحظة مضبوطة، نحصل من خلالها على معلومات مقننة معينة بالأرقام، ننتقل منها إلى تصنيف هذه المعلومات وتنظيمها وفق أنساق معينة تؤدي في نهاية الأمر إلى صياغة القانون العلمي، كما تؤدي من جانب آخر إلى الإيحاء بفروض مفسرة لجوانب جديدة أو التنبؤ وفقا لاحتمالات إحصائية بمواقف جزئية لم تدخل في إطار ونطاق الملاحظات الأولى التي صيغ القانون العلمي على أساسها.

ويمثل هذا الهدف مسار علم النفس العلمي في سعيه للتعرف على قوانين السلوك والتفكير من خلال منحى الفروق الفردية، وهو المنحى الذي بدأ مبكرا من خلال أعمال جميس ماكين كاتل وفرانسيس جالتون.

الهدف الثاني: استعمال نتائج القياس للحصول على معلومات محددة، تفيد في توظيف العلم لصالح المجتمع سواء على المستوى الجمعي أو على المستوى الفردي. فالسيكولوجي يواجَه بمطالب اجتماعية معينة، من الضروري أن يقوم بها لتأدية دوره، فهو مطالب بأن يقيس الاستعدادات والقدرات والدوافع والاتجاهات والميول والقيم وسمات الشخصية والاضطرابات النفسية والعقلية، لأنه مطالب بأن يساهم في وضع العامل المناسب في المهنة المناسبة، وأن يشكل أفواجا دراسية متجانسة في الميول والقدرات، وأن يشخص حالات سوء التوافق من أجل تقديم العلاج والرعاية المناسبين، وأن يضع برامج التدريب والتعلم. فالسيكولوجي عليه أن يتواجد في المدرسة وفي العيادة وفي المصنع وفي كل المؤسسات الاجتماعية.

ويمثل هذا الهدف في اتساعه عددا كبيرا من الخدمات التي يقدمها القياس النفسي، سواء في قياس الذكاء أو القدرات أو الاستعدادات النوعية، الذي أصبح ضروريا ومطلوبا سواء في المدرسة أو في المهنة أو في التدريب على المهارات المختلفة. والقياس النفسي يقدم خدمات جليلة من خلال استخدام الاختبارات النفسية من أجل التعرف على قدرات الأفراد واستعداداتهم، ووضع محكات تشخيصية، سواء للمتفوقين أو الموهوبين أو المتأخرين واستخدام اختبار القدرة العقلية العامة يفيد في تشخيص حالات التأخر العقلي والمرض العقلي وحتى السيكوباتية.

والهدف من استخدام اختبارات القدرات العقلية واستبيانات سمات الشخصية والدافعية والميول والاتجاهات، هو تقديم العون والإرشاد للأفراد، سواء الإرشاد التعليمي أو المهني أو التوافقي بصفة عامة وفي جميع المجالات سواء في المدرسة أو العمل أو العيادة يتجه استعمال الفرد، ليتمكن من استثمار قدراته أو لتمكين المجتمع من تقديم العون له سواء

بوضعه في المستوى والتخصص الدراسيين المناسبين أو في العمل المناسب أو في الظروف المناسبة.

ما الفرق بين القياس النفسي والاختبار النفسي و العقلي:

يتجاوز علم النفس القياسي الاختبارات بصورة كبيرة و هو أوسع بكثير من الاختبار النفسي الذي هو من مواضيع علم النفس القياسي وهو جزء فقط منه.

يعتبر العالم فرانسيس جالتون أول من استخدم مصطلح الاختبار النفسي والعالم جيمس ماكين كاتل (1980) أول من استخدم مصطلح الاختبار العقلي. يقومالفرق بينهما على اعتبار أن الاختبار النفسي يشمل الاختبارات العقلية وغير العقلية كالاختبارات التي تقيس الجانب الوجداني أو الجانب النفسي الحركي وأيضا سمات الشخصية. هناك فرق آخر بين الاختبار والمقياس حيث المقياس مصطلح أكثر عمومية لأنه يستخدم في كل ميادين البحث النفسي للحصول على أوصاف كمية /رقمية على عكس الاختبار.((بشير معمرية، 2007: 93)

أخطاء القياس :

مهما حاول القياسيون النفسيون أن يكونوا دقيقين في القياس، إلا أنهم لا ينجون من الوقوع في أخطاء، وهذه الأخطاء بعضها خارج عن إرادتهم. ومن الضروري أن يكونوا على وعي بها قبل قيامهم بالقياس، كي يحدوا منها، ويخلصوا منها الدرجات التي يحصلون عليها بوساطة الطرق الإحصائية، أو يأخذوها بعين الاعتبار عند تفسير هذه الدرجات. وتنجم هذه الأخطاء عن ما يلي :

1.**عدم حساسية أدوات القياس** : بعض أدوات القياس أكثر دقة من بعضها الآخر، والاختبار الذي يصممه فاحص خبير مدرب أكثر حساسية لما يقيس من اختبار وضعه فاحص لا يدري عن القياس شيئا. والامتحان الموضوعي أكثر دقة من الامتحان المقالي.

2.عدم ثبات الظاهرة السلوكية : معظم ما نقيس في ميدان علم النفس ديناميكي، أي متقلب ومتغير. فالشخص الذي نقيس أداءه أو سمة من سماته، كل ساعة هو في شأن ففي البداية يكون في حالة نفسية معينة، وبعد ساعة تتغير حالته، وبعد قليل قد يشعر بالتعب أو الملل أو الضجر أو الخوف، أو يفقد اهتمامه وتنخفض دافعيته، ويتأثر أيضا بدرجة الحرارة والتهوية والإضاءة وبقية المتغيرات البيئية. فالسلوك البشري في حالة تغير مستمر وكثيرا ما يخدع القياس النفسي ومعقد تعقيدا كبيرا.

لذا فإن القياسيين النفسيين يحاولون أن يقتربوا من الحقيقة بأخذ عدة قياسات ثم يستخرجون متوسطها واعتباره أقرب إلى الأداء الحقيقي؛ أما الفرق بين متوسط القياسات وأي قياس منها يعتبر خطأ في القياس.

3. خطأ الملاحظة :تختلف دقة الملاحظة من شخص إلى آخر، فلو طلبنا من عدة أفراد قياس طول فرد، فإنهم يعطوننا أطوالا مختلفة، مع أن أداة القياس واحدة لم تتغير، وكذلك الفرد موضوع القياس. فتأتي الأطوال كما يلي : 150.00 سم، 150.50 سم 150.40 سم، 150.45 سم ...الخ. وهذه الفروق في القياسات ممكنة الحدوث لدى الشخص نفسه إذا قام بالقياس عدة مرات.(فخري رشيد خضر، 2003: 55)

الدرس الثـــــالث: نشأة القياس وتطوره في علم النفس

نشأة القياس وتطوره في علم النفس:

استخدم الصينيون نظاما متطورا للامتحانات منذ أكثر من 3000 سنة من أجل اختبار وتثبيت الموظفين الساميين للدولة. كان المرشحون إلى الوظائف السامية في الصين يخضعون إلى امتحانات شفوية وكتابية يشرف عليها الامبراطور نفسه، وكانت على درجة عالية من الصعوبة تتم على عدة مراحل في الموسيقى والقانون والكتابة ومبادئ الكونفو وركوب الخيل هي الأولى من نوعها في تاريخ المجتمعات الإنسانية.

كانت الطريقة المتبعة في القرن 18 تعتمد على العين والأذن أي تآزر الإحساسات السمعية والبصرية وهي الطريقة المتبعة في رصد النجوم.

في القرن 19 اهتم علماء الفلك (وهم الأوائل) بقياس الفروق الفردية بين تقدير اثنين من الراصدين لحركة نجم من النجوم (الفرق بالثواني) بين تقدير اثنين في الراصدين : تقدير بالملاحظة والسمع و الترصد، وأصبحت تسمى فيما بعد بالمعادلة الشخصية.

امتدت حركة قياس الفروق الفردية إلى قياس الظواهر النفسية العليا كالذكاء والاستعدادات والقدرات والميول وسمات الشخصية...

تطور القياس النفسي بشكل عام ولكن تأخر في مجال الشخصية عن القياس في مجالات أخرى كالمجال المعرفي، الذكاء والقدرات. و قد مرّ بعدة مراحل :

1.**المرحلة الأولى:** الأحكام المبنية على الفراسة:

أ–**دراسة ملامح الوجه:** كان الانسان قديما يقارن بين قواه ومواهبه مع ما يحيط به من كائنات حية. كانت المقارنة ترتكز على ملامح الوجه والجمجمة من كان وجهه كوجه القرد فهو ذكي ومن كان وجهه يشبه الحمار فهو غبي وصبور ومن كان مثل الأسد فهو شجاع وجريئ ومن كانت عيونه ثاقبة مثل عيون الثعلب فهو ماكر.

أكدت كل من دراسة جالتون Galton (1886) وبيرسون Berson حول دراسة الوجه و دراسة جوران Gorane (1913) حول فراسة التشوهات الخلقية لدى المجرمين، على خطأها في قياس الذكاء.

ب-دراسة تضاريس الجمجمة: بدأ العلماء بين 1758-1828 أمثال فرانز جوزيف جول يعمل بخرائط الجمجمة ويعترف وجود خصائص في مناطق محددة بالقشرة المخية يفترض ويمكن الاستدلال على ذلك بوجود بروزات في الجمجمة.

2. المرحلة الثانية: بداية التجريب والقياس (قياس النواحي الحسية-الحركية).

بدأت في أواخر القرن18 وبدايات القرن19 قياس النواحي الحركية -الحسية الدنيا للنشاط العقلي بواسطة أجهزة وآلات معملية بسيطة مثلا استخدام كاتيل Cattali العلاقة بين النواحي الحسية الحركية في قوة قبضة اليد، سرعة الاستجابة للصوت، سرعة ذكر أسماء الألوان.

واستخدم جيلبرتGilbert (1897) لإيجاد العلاقة بين النواحي الحسية والحركية و الذكاء: أقل ألم يشعر به الفرد نتيجة الوخز، قوة الرفع بمعصم اليد، قوة الرفع باليد، تقدير الطول بالنظر، قوة الصدر وسعته، الوزن.

أثبتت هذه الاختبارات فشلها في القياس الدقيق للذكاء.

3. المرحلة الثالثة: قياس العمليات العقلية خلال القرن17 ونهاية القرن18: إن حركة القياس النفسي ظهرت على يد علماء الفلك وليس على يد علماء النفس منها :

دراسة فرانسيس جالتون Galton في 1869 عن الوراثة وتحديد النسل في انجلترا.

ظهور حركة القياس النفسي التجريبي في ألمانيا بإنشاء أقل معمل للقياس النفسي.

ظهور حركة العناية بضعاف العقول في فرنسا وظهور اختبار بينيه وسيمون- Binet Simon.

-انتشار حركة القياس العقلي في أمريكا على يد ماكين كاتل وتلميذه جالتون وغيرهم مثل وليام جيمس.

4. المرحلة الرابعة : قياس العلميات العقلية العليا: خلال القرن 19 مثل: التفكير، الانتباه، الإدراك...الخ. اتسمت بظهور أدوات قياس أكثر دقة وأكثر تنوعا. من أشهر أدوات القياس خلال هذه الفترة :

-مقياس استانفورد -بنييه الفرنسي Stanford-Binet

-مقياس وكسلر-بالفو Wechsler –Bellevue الذي وضعه وكسلر الإخصائي النفسي في مستشفى فرنسي للتخلص من عيوب اختبار بينييه.

-لوحة أشكال سيجان.

-اختبار رسم الرجل.

-اختبار المصفوفات المتتابعة لرافين.

-الاختبارين المشهورين للجيش الأمريكي المعروفين باسم بيتا Beta الذي يستخدم مع الجنود الأُمّيين وألفاAlpha الذي يستخدم لقياس ذكاء الجنود الذين يعرفون القراءة والكتابة.

-نظرية العاملين لسبير مان Spearman

-نظرية التنظيم الهرمي للقدرات مثل تنظيم Galford, Vernon, Burt

-نظرية الذكاء المتعدد لجاردنر.

5. المرحلة الخامسة : تنوع الاختبارات لقياس مختلف الظواهر النفسية في الشخصية :

لم يعد القياس محصورا على الجوانب العقلية المعرفية بل امتدت لتشمل الجوانب الانفعالية الوجدانية فظهر اختبار:

-استرونغ لقياس الميول المهنية.

-اختبار كيودر Kuder لقياس الميول اللامهنية.

- قياس ألبورت، فرنون ولندزي لقياس القيم.

-قائمة "موني" للمشكلات ومقاييس الشخصية.

-مقاييس الشخصية : مقياس كاتل، اختبار الشخصية العاملي، اختبار الشخصية المتعدد الأدلة – مقياس ايزنك للشخصية

-مقاييس سمات الشخصية مثل (مفهوم الذات، تقدير الذات، تأكيد الذات، مقاييس القلق والخوف).

تطور القياس و توجهاته في العصر الحديث:

شهد القرن 19 اهتماما متزايدا وقويا بالمتأخرين عقليا والذهانيين أو المرضى العقليين من أمثال Seguin, Esquirol ثم تلاهم Binet وجميعهم من فرنسا.

في عام 1879 أسس فونت Wundt أول معمل تجريبي في لايبزج في ألمانيا. وفي عام 1884 افتتح Galton الذي يعد مؤسس الدراسة العلمية للفروق الفردية معمل القياس الإنساني في انجلترا anthropométry .

في أمريكا استخدم جميس ماكين كاتل Cattel (1890) لأول مرة مصطلح الاختبار. وفي عام 1905 أعد بينيه سيمون أول مقياس للذكاء.

في عام 1892 استخدم Kraeplin اختبار تداعي المعاني مع المرضى في المجال الطبي النفسي. وضع Woodworth النموذج الأصلي لاستخبارات الشخصية عام 1919 وفي العشرينات و الثلاثينات بدأ منحى الاختبارات الموقفية الأدائية على يد Hartshorne و mary في 1928.

ظهر بعدها الاتجاه السيكومتري الذي حاول مطابقة الإحصاء على النفس (الثبات، الصدق و المعايير) مثل : Bell (الاختبار التوافق الذي أعده أواخر الثلاثينات) Bern Reuter(اختبار الشخصية في أواخر الثلاثينات) Hat Haway MC Kinely (اختبار الشخصية متعدد الأوجه في الأربعينات واسعة الشهرة).

في سنة 1905 ظهور اختبار بينيه وسيمون Binet – Simon لقياس الذكاء وضع بينيه وسيمون اختبار لقياس الذكاء سنة 1905.

–سمحت هذه الأداة بتقييم المستوى العقلي للأطفال. – يتكون هذا الاختبار من 30 بند (فقرة) مرتبة حسب تدرجها في الصعوبة وهذا بالنسبة إلى الفئات العمرية 3 و 5 و 7 و 9 و 11 سنة.

– إن بنود هذا الاختبار موجهة بالأساس إلى قياس الوظائف الذهنية العليا، وتتضمن عددا قليلا من الفقرات لتقدير التطور الحسي – الحركي.

–إن الضعف العقلي صنف إلى ثلاثة أصناف هي :

العتة (Débilité) والعبط (Idiotie) والبله (Imbécilité).

–عرف هذا الاختبار عدة تعديلات ظهرت النسخة الثانية سنة 1908 والنسخة الثالثة سنة 1911.

نالت النسخة الثانية والثالثة نجاحا عالميا كبيرا حيث خضعت إلى ترجمة في العديد من البلدان:

–السويد 1914 جادر هولم (Jaederholm) – بلجيكا 1909 ديكرولي وديقو (Decroly – Degaud).

_ ألمانيا 1914 مومان و بوبرتاغMoumann et Bobertag

_ ايطاليا 1911 سافيوتي Saffiatti

– بريطانيا 1921 بيرت (Burt)

–الولايات المتحدة جودار (Goddard) ثم تيرمان (Terman).

توصل زازو (Zazou) إلى إخراج نسخة جديدة سنة 1966 وهذا بعدما أدخل تعديلات جديدة مناسبة على الاختبار حيث حذف وحول أصناف بعض الفقرات.

و في سنة 1911 ظهر مفهوم البروفيل النفسي (البروفيل القطبي) Profil de polarité–قدم السيكولوجي الروسي الأصل روسوليمو (Rossolimo) سنة 1911 طريقة عرض جديدة للبيانات بواسطة أشكال تسمى البروفيل القطبي أو البروفيل النفسي.

–إن هذا البروفيل أصبح أداة واسعة الاستخدام في علم النفس الفارق مع العلم أن استخداماته يمكن أن تكون متعددة.

–إن مفهوم البروفيل القطبي لديه أهمية خاصة لدى الباحثين في العلوم الإنسانية والاجتماعية وذلك لأن الظواهر السيكولوجية والاجتماعية تكون في الغالب نتيجة تفاعلات عوامل متضادة ومتناقضة.

–عمل السيكولوجيون على جعل هذا الاتجاه في التفكير إلى تناول علمي حيث ينظرون في أغلب الأحيان إلى مواضيع الدراسة من خلال استقطاب متناقض مناسب وذلك بهدف الوصف والتحليل.

–إن هذا الاستقطاب المتناقض قد يكون ثنائي محدد بقطبين فقط مثلا يدوي – ذهني ، بدوي –حضري، رجل – امرأة .. أو متعدد يتكون من عدة نقاط موجودة على سلم. ومن الأمثلة التي تندرج في هذا النوع من الاستقطاب المتعدد، يمكننا أن نجد الفئات العمرية من الطفولة إلى الشيخوخة والمداخل الاقتصادية من الفاقة إلى الرخاء، ونتائج اختبار في الذكاء من العبط إلى العبقرية .. يعتبر الاستقطاب الثنائي أو المتعدد طريقة مناسبة للقيام بمختلف المقارنات.

و في عام 1912 ظهور معادلة الذكاء–إن معادلة الذكاء هي العلاقة بين العمر العقلي الذي يحصل عليه طفل في اختبار للذكاء وعمره الزمني. لقد سمح مفهوم هذه المعادلة بحدوث تقدم كبير في قياس الذكاء.

–إن أول خطوة مهدت لظهور هذا المفهوم تمت بفضل أعمال بينيه سنة 1911 عندما أقترح مفهوم العمر العقلي.

–فحسب بينيه إن العمر العقلي لطفل معين مثلا عُمَر، يوافق العمل الفعلي للطفل المتوسط الذي يستطيع أن يحل الفقرات التي توصل عمر إلى حلها، فإذا توصل عمر إلى تجاوز الفقرات التي يحلها في المتوسط طفل عمره 8 سنوات بنجاح فإن عمره العقلي هو 8 سنوات وهذا رغم أن عمره الزمني الفعلي يساوي 10 سنوات أو 6 سنوات.

–في سنة 1912 قدم استيرن (Stern) معادلة الذكاء والتي يمكن حسابها بتقسيم العمر العقلي على العمر الزمني أي :

$$\frac{6}{8} = 1{,}33 \quad \text{أو} \quad \frac{8}{10} = 0{,}8$$

فإذا كان الناتج أكبر من واحد (1) فإن المفحوص متفوق عن عمره الزمني وإذا كان أقل من 1 فإنه يعتبر أقل من عمره الزمني.

–يعتبر بوبرتاغ أول من استخدم هذه المعادلة باقتراح من أستيرن.

–أدخل تيرمان تعديل آخر على هذه المعادلة سنة1916 وذلك بضرب الناتج في مائة :

$$\frac{8}{10} \times 100 = 80$$

وبالتالي من الناحية النظرية العلامات التي تقع فوق 100 تدل على ذكاء فوق المتوسط والتي تقع تحت 100 تدل على ذكاء تحت المتوسط.

-إن المتخلفين عقليا لديهم ذكاء أقل من 70 درجة.

- إن معنى مفهوم معادلة الذكاء (العلاقة بين العمر العقلي والعمر الزمني) ليس لديها قيمة بالنسبة للكبار وذلك لأن قدرات الذكاء تتوقف عن النمو عند حوالي سن 16 ، ولهذا يجب القيام بحسابات إحصائية للوصول إلى علاقة لديها دلالة مشابهة لمعادلة الذكاء للأطفال.

و في سنة 1914 ظهراختبار المتاهات لبورثوسPorteus الذي وضعه بورثوس اختبار المتاهات بأستراليا سنة 1914.

و في 1916 ظهر نسخة اختبار استانفور-بينيهStanford - Binet-ترجم تيرمان اختبار بينيه - سيمون وكيّفه على عينة أمريكية واشتهرت هذه النسخة بالتسمية التالية : اختبار استانفورد-بينيه (نسبة إلى جامعة استانفورد التي كان يشتغل بها تيرمان).

-عرفت هذه النسخة عدة تعديلات وتحسينات فيما بعد، ففي سنة 1937 نشر تيرمان بالتعاون مع ميرل نسخة جديدة خاصة بالأطفال (بداية من سن السنتين) والمراهقين والكبار، تتضمن 122 فقرة من بينها بعض الفقرات لا تعتمد على اللغة. -ظهرت أثناء الحرب العالمية الأولى الحاجة الملحة إلى تصنيف سريع لعدد كبير من الأفراد المجندين مما أدى إلى ظهور بطاريتين من الاختبارات الجماعية، بطارية ألفا سنة 1917 وبطارية بيتا سنة 1918.

1917و1918 ظهور بطارية ألفا (1917) وبطارية بيتا (1918) -البطارية الأولى موجهة للأفراد الذين يحسنون اللغة الانجليزية والبطارية الثانية موجهة للأفراد غير المتعلمين، حيث تتضمن بنودا تتطلب أداء غير لغوي.

–يعود الفضل إلى السيكولوجي يركس (Yerkes) الذي جُنِّد في سنة 1917 في تطوير هاتين البطارتين.

–لقد تم فحص بواسطة هذين البطارتين أكثر من 1.700.000 من أفراد القوات المسلحة خلال 18 شهر وتم توجيههم إلى مناصب تناسب قدراتهم.

و في 1920 ظهور أول قائمة للشخصية لوود وورث Woodworth–قدم وود وورث في سنة 1920 أول قائمة للشخصية والتي يمكن تطبيقها بصورة فردية أو جماعية.

– تتكون النسخة النهائية من 116 سؤال والإجابات عنها عبارة عن اختبارات ذات طبيعة مؤكدة أو سلبية. –تسمح لنا النتائج المتوصل إليها بواسطة هذه الأداة بفرز الأفراد الذين لديهم تصورات انفعالية مضطربة.

1920 عقد أول ملتقى حول علم النفس التقني لقد أصبحت مسألة الاختبارات النفسية تحتل مكانة هامة في الملتقيات العلمية العالمية، حيث أدى هذا الاهتمام إلى عقد أول مؤتمر عالمي حول الموضوع في سنة 1920 بجنيف.

–شارك في هذا الملتقى باحثين من عدة دول، وعلى إثره تم إنشاء الجمعية الدولية لعلم النفس التقني.

1921 ظهور اختبار بقع الحبر رورشاخ Rorschach –أخرج رورشاخ المختص في الطب النفسي وفي علم النفس اختبار بقع الحبر لفحص الشخصية في الفترة الموالية 1918–1921. –يتكون هذا الاختبار من 10 لوحات (6 منها بالأبيض والأسود و4 بالألوان).

–كان رورشاخ مولع بفن الرسم (Peinture) وهذا بالإضافة إلى اهتمامه بالتحليل النفسي والأحلام وترابط الكلمات. توصل هذا الباحث إلى بناء هذا الاختبار من خلال دراساته الاختلافات بين الأفراد في الاستجابة إلى لوح أو بقع حبر، حيث كان يسجل إجابات

مرضاه ويقارنها مع إجابات أفراد عاديين، فاكتشف أن الإدراك البصري يتأثر بالشخصية والاضطرابات النفسية.

-إن الانتشار الواسع لهذا الاختبار لم يكن إلا بعد عشر سنوات من وفاة صاحبه ويعتبر حاليا الاختبار الأكثر استعمالا في فحص الشخصية.

1923 ظهور اختبار المكعبات لكوهس Kohs يعتبر اختبار كوهس من اختبارات الذكاء الخاصة بالأطفال (5 – 11سنة) التي نالت اعترافا واسعا فيما يخص صدق وثبات نتائجه حيث ينظر إليه على أنه اختبار جيد لتشخيص البله العقلي كما أنه يندرج ضمن الاختبارات المتميزة للنمو الذهني. تجدر الإشارة إلى أنه يعتبر من الاختبارات الخالية من تأثير العامل الثقافي.

-تعلن العديد من الدراسات على أنه يتمتع بميزات جيدة فيما يخص الخصائص السيكومترية (الصدق والثبات).

1926 ظهور اختبار رسم الرجل لقودينوفGoodenough-إن المبدأ الأساسي في هذا الاختبار هو أن قدرة الإنسان على معرفة الأشياء والأجسام من خلال رسومها وصورها ترتبط بالذكاء، حيث تدل القدرة على تمثيل الأشياء والأجسام بواسطة الرموز والرسم والصور على مزيد من الذكاء وذلك لأن مثل هذه القدرة تتطلب درجة عالية من التجريد والاختزال والإبداع.

-إن الرسم حسب قوديناف هو لغة للتعبير، ادواتها ليس الكلمات وإنما الخطوط والأشكال المرسومة. إن هذه الخطوط تعبر عن مفاهيم، والمفاهيم ترتبط مباشرة بالذكاء وبالتالي فإن الرسم يصلح لقياس الذكاء. تشير الدراسات حول رسم الرجل إلى أنه يعطي أفضل النتائج مع الأطفال الذين تتراوح أعمارهم فيما بين سن الرابعة والثالثة عشر، مع العلم أن بعض الدراسات تمدد هذه الفئة العمرية إلى سن الخامسة عشر في حين تقترح دراسات أخرى العكس أي تقليصها إلى سن الحادية عشر وذلك للحد من تأثير أثر التعلم.

–يصف البعض هذا الاختبار على أنه متحرر من تأثير الثقافة إلا أن بعض الباحثين مثلا البدري وغيره يشيرون إلى متغير ثقافي هام وهو اختلاف الزي التقليدي عن الزي الأوروبي الذي بنت على أساسه قودينوف لبروتوكولات التصحيح لرسم الرجل في هذا الاختبار (1926).

1927 ظهور اختبار قائمة الميول المهنية لسترونغ Strong –تساعدنا قائمة الميول المهنية لسترونغ في تعيين الأفراد في مناصب العمل المناسبة لميولهم المهنية والتي يمكن أن يحققوا فيها الرضا المهني. –إن الاستخدام الواسع لهذا الاختبار أدى بصاحبه إلى تحضير عدة مراجعات وبالتالي إلى ظهور نسخ معدلة في سنة 1933 و 1938 و 1946.

–مع العلم أنه بعد وفاة استرونغ واصل الباحثان كامبل وهانسن عملية تحديث الأداة وهذا في العديد من المرات أي سنة 1974 و 1981 و 1985 و 1992.

–يتمتع هذا الاختبار بقدرة على التنبؤ عالية وبخصائص سيكومترية جيدة وهذا حسب ما تشير إليه العديد من الدراسات.

1935 ظهور اختبار تفهم الموضوع لموراي Murray طور موري اختبار تفهم الموضوع لدراسة وفحص الشخصية ولقد عرف هذا الاختبار انتشارا واسعا فيما بين المختصين.

–يعتبر موري من أبرز السيكولوجيين الذين درسوا الشخصية، مع العلم أنه كان متأثرا جدا بتصورات فرويد فيما يخص الهو والأنا والأنا الأعلى وكذلك مفهوم الشعور واللاشعور والعقد ومراحل النمو (الشرجية والفمية والقضيبية). ويختلف مع فرويد فيما يخص الدوافع، حيث يرى أن بناء الشخصية يدور حول الحاجيات، ويشير إلى وجود عشرين منها ظاهرة مثل العدوان والسيطرة والاستقلالية ...الخ؛ و8 منها مستترة. وقوة هذه الحاجيات بالنسبة إليه تكون حسب كل فرد وتصطدم مع طاقات من المحيط.

– شهدت سنة 1938 ظهور اختبار المصفوفات المتدرجة لقياس الذكاء لرافن Raven–يعتبر اختبار المصفوفات المتدرجة لقياس الذكاء لرافن من أكثر اختبارات الذكاء المستخدمة في العديد من البلدان والمطبقة في مختلف ميادين القياس السيكولوجي (المهني– التربوي والعيادي). –تقيس هذه المصفوفات القدرة على الاستدلال وتتضمن ثلاث نسخ متدرجة في الصعوبة مع العلم أنها خالية من تأثير العامل الثقافي وذلك لعدم استخدام للغة في بناء بنودها.

1938 ظهور سلم مقاييس النضج لجيزلGesell، يهدف سلم مقاييس النضج لجيزل إلى تحديد العمر التطوري للأطفال (4 – 6 سنوات) وهذا بالاستناد إلى مستويات من النضج وعينات من السلوك.

–يقول جيزل 1964 "أن الطفل يكتسب معلومة ذهنية بنفس الصورة التي يكتب بها نموذج من السلوك، بمعنى بواسطة سيرورة النمو".

1939إلى 1997 ظهور اختبار الذكاء لويكسلير Weschslerومختلف التعديلات والنسخ–أولا – سلم ويكسلير –بلفيو الخاص بالكبار (Wais). maj

–ظهر سلم ويكسلير – بلفيو للذكاء في سنة 1939 والذي أصبح بنافس اختبار استانفورد – بينيه، خاصة وأن هذا الأخير لا يناسب الكبار ويعتمد كثيرا على اللغة.

–إن هدف هذا الاختبار (WAIS) هو تكميم ذكاء الكبار (16–74 سنة) مع العلم أن ويكسلر يستغني على مفهوم العمر العقلي في عملية التكميم وذلك لأنه ينظر إلى الذكاء على أنه مجموعة من القدرات الديناميكية المترابطة فيما بينها وليس قدرة واحدة استاتيكية ويعتبر الترابط بين هذه القدرات هام جدا حيث يسمح للفرد من أن يفهم العالم وأن يتكيف مع محيطه.

–لقد انتشر واشتهر اختبار(WAIS)بسرعة فيما بين المختصين في علم النفس، إن المراجعات الأولى لهذا الاختبار كانت في سنة 1977 و 1981. نالت نسخة 1981 اعترافا وانتشارا واسعا عبر العالم. تم تكييف هذا الاختبار إلى الفرنسية سنة 1989.

ثانيا – سلم ويكسلر الخاص بالأطفال (WISC) 1949. بناء على اختبار (WAIS) طور ويكسلر في سنة 1949 سلم لقياس ذكاء الأطفال. في سنة 1974 تم القيام بالتعديل الأول لهذا الاختبار الخاص بالأطفال. في سنة 1991 ظهرت لهذا الاختبار النسخة الثالثة المنقحة والتي تعرف بـ (WISC–III).

ثالثا – سلم الذكاء لويكسلر الخاص بمرحلة ما قبل التمدرس والابتدائي 3–7 سنوات (WPPI). طور ويكسلر لأطفال مرحلة ما قبل التمدرس والابتدائي (4 –6 سنوات ونصف) سلم مناسب لتطورهم الذهني. إن مراجعة هذا الاختبار سمحت بشمل مرحلة عمرية أوسع (3 – 7سنوات). يشير استون وقريدلي وجيورك إلى أن هذا الاختبار يصلح أكثر لصغار السن.

من 1942 إلى1990 ظهور قائمة الشخصية لمينيسوتا المتعدد الأوجه 1 و 2 MMPI–نشر هاتاوي (Hathaway) وماك كينلي (Mc Kinly) قائمة الشخصية لمينيسوتا المتعدد الأوجه لأول مرة سنة 1942. تم تصميم هذا الاختبار بالاعتماد على التناول الأمبيريقي. إن اختبار قائمة الشخصية لمينوسوتا المتعدد الأوجه يعتبر أكثر اختبارات الشخصية المنتشرة وأكثرها استخداما في مختلف أنحاء العالم، حيث ينظر إليه على أنه من الاختبارات العيادية المفيدة. تعود المراجعة الأساسية الأولى للاختبار إلى سنة 1990 حيث عرف عدة تحسينات وأدى إلى ظهور ما أصبح يعرف بـ MMPI–2.

في سنة 1959 ظهور نظرية التنظيم العقلي لجيلفورد Guilford. قدم جيلفورد في سنة 1959 نموذجا للتنظيم العقلي سماه نموذج تركيب العقل. يتضمن هذا النموذج ثلاثة أبعاد (العمليات العقلية– المحتويات – النواتج) إن التفاعل فيما بين مكونات العمليات العقلية والمحتويات والنواتج يدل على وجود 120 عامل. مع العلم ان كل عامل يحتوي على الأقل

على قدرة عقلية، وبالتالي يوجد على الأقل 120 قدرة عقلية مع العلم أنه تم اكتشاف وقياس أغلبية هذه القدرة.

ظهر في 1972 سلم الاستعدادات للأطفال لماك كارثيMC Carthy. طورت ماك كارثي سلم الاستعدادات لقياس مهارات الأطفال في سن ما قبل التمدرس 2-6 سنوات (وذلك من أجل التنبؤ بالأداء المدرسي لأطفال لهذه الفئة). تشير مختلف الدراسات إلى أن هذا الاختبار يتمتع بخصائص سيكومترية جيدة من صدق وثبات.(محمود بوسنة، 2007)

الدرس الرابع: عرض أنواع أدوات و بيانات و متغيرات ومحكات القياس النفسي

الذي يتم قياسه ليس الأشياء والظواهر بل خصائص الأشياء والظواهر كالطول الحجم والوزن...الخ. إن معظم الخصائص أو السمات التي تخضع للقياس في العلم قد لا تقاس مباشرة، فمشكلة عدم قابلية بعض الخصائص للقياس المباشر توجد أيضا في العلوم الأخرى كالبيولوجية والفيزياء فهناك خصائص لا تدرك حسيا ولا تقاس مباشرة مثل: ضغط الدم، والبلوغ، العمر البيولوجي، التوتر العضلي، المغناطيسية، الكهرباء، الجاذبية، المقاومة..

والمعرفة بها وقياسها تتم عن طريق مؤشراتها أي ما يدل على وجودها كميا وتغير هذا الكم.

الذي يقبل الملاحظة المباشرة والقياس في علم النفس هو الأداء (السلوك الظاهري) وهو على ثلاث مستويات.

1.الأداء اللغوي

2.الأداء الحركي.

3.الأداء الفيزيولوجي

يستخدم الباحثون عدة طرق لقياس الخصائص أو التكوينات الفرضية وأهمها :

–تكرار أو احتمال حدوث مؤشرا لخاصية (الأداء).

– شدة حدوث مؤشر الخاصية

– مدى حدوث مؤشر الخاصية.

أدوات القياس النفسي:

- الاستبيانات questionnaires
- مقاييس التقدير

- قوائم الصفات
- الطرق الاسقاطية
- اختبارات الشخصية الموضوعية
- المقاييس الفيزيولوجية

البيانات :

البيانات التي يتعامل معها الباحثون في علم النفس قد تتخذ عدة أشكال منها :

-**البيانات المتقطعة:** مثلا: سنوات العمر، مستويات التعليم، نسب الذكاء، درجات التحصيل وغيرها.

تعتمد على مستوى القياس الاسمي أو الرتبي أو المسافات المنتظمة، ويمكن وضع هذه البيانات على شكل جداول إحصائية وفي فئات معينة.

-**البيانات المتواصلة:** مثلا: قياس الضوء والصوت والجهد ونشاط الدماغ والعضلات والحرارة، ويمكن تحويلها إلى بيانات متقطعة لتسهيل إجراء التحليل الإحصائي عليها. تعتمد على استعمال مستوى قياس المسافات المنتظمة ومستوى النسب.

- **بيانات الدوام** : وهي مرتبطة بقياس الزمن. والزمن متغير متواصل إلا أن بياناته يمكن تحويلها إلى بيانات متقطعة مع قبول نسب معينة من الخطأ، فيقال أن الساعة تشير إلى العاشرة مثلا وذلك رغم استمرار سريان الوقت وعدم توقفه تماما. معالجة هذه البيانات يعتمد على المقياس النسبي.

ملاحظة : يصلح استخدام الإحصاء اللابارمتري مع جميع مستويات القياس إلا أن الإحصاء البرامتري فيمكن استخدامه فقط مع مستوى المسافات المتساوية ومستوى النسبة.

القياس إما أن يكون رقميا نحصل فيه على تقديرات رقمية Psychométrie وإما أن يكون وصفياDescriptive testing يعطي صورة وصفية عن الفرد ويسمى الوصفي بالاتجاه التأثريImpressionnistic .

تقسم السمة والخاصية إلى نوعين :

الخاصية البسيطة كالوزن والطول.

الخاصية المركبة كالذكاء والقدرات اللغوية.

ويكون القياس إما مباشرا أو غير مباشر.

أنواع المتغيرات :

1-**المتغيرات النوعية** : هي مجموعة من الخصائص والصفات التي لا تقاس بشكل كمي وتوجد على أشكال مختلفة وليس لها ترتيب منطقي، كأنواع الفواكه والخضر وأنواع السيارات والجنس والأسماء.

2- **المتغيرات الكمية** : وهي التي يمكن قياسها بشكل كمي وهي على نوعين :

أ-متغيرات منفصلة أو متقطعة وتأخذ قيم محددة ضمن المدى ولا يمكن التعبير عنها بكسور و انما نعبر عنها بالأعداد الصحيحة : عدد الطلبة، عدد المدارس...

ب- المتغيرات المتصلة المستمرة: وهي مجموع الخصائص التي يمكن أن تأخذ أي قيمة محتملة ضمن المدى ويمكن استخدام الكسور العشرية (الطول، الوزن والحجم ..).

3-**المتغيرات الرتبية:** وهي مرتبطة بالرتبة كترتيب الأشياء والأشخاص حسب درجة امتلاكهم لخاصية ميعنة (الأول، الثاني ..).

نتكلم عن مستويات القياس التي تسمح لنا بمعالجة المتغيرات من خلالها.(عبدالله الصمادي و ماهر الدرابيع ، 2004)

-أنواع المحكات:

نورد فيما يلي أنواع المحكات كما لخصتها أ. أنستازي (1961)

1)الفروق في العمر:

يستخدم هذا المحك عندما يكون المتغير موضع القياس من المتغيرات التي تتأثر بالنمو خاصة في مرحلة الطفولة وبداية المراهقة. فالذكاء مثلا يتأثر بالزيادة في العمر، فيمكن اتخاذ التغير في العمر كمحك خارجي لتقدير صدق اختبار يقيس الذكاء، فيحسب معامل الارتباط بين التغير في العمر والتغير في درجات اختبار الذكاء والقدرات العقلية.

ومعامل الارتباط المحصل عليه هو معامل الصدق، فإذا كان دالا إحصائيا دل ذلك على صدق الاختبار، وإذا لم يكن دالا إحصائيا دل ذلك على عدم صدق الاختبار. وقد حسب بيني صدق اختباره للذكاء بهذا الأسلوب، حين اعتبر ارتفاع معامل السهولة في الإجابة على بنود الاختبار لعمر عقلي معين مع الزيادة في العمر الزمني، مؤشر على صدق الاختبار لفئة هذا العمر. فعلى سبيل المثال هناك مجموعة أسئلة اختبارية لقياس الذكاء وضعت لمن عمرهم 5 سنوات عقلية، أجاب عنها 50 % من أصحاب هذا العمر، فإذا أجاب عنها 63 % من أصحاب عمر 6 سنوات و 4 % من ذوي العمر 7 سنوات و85 % من ذوي العمر 8 سنوات و 96 % من فئة العمر 9 سنوات، اعتبرت الأسئلة الاختبارية صادقة لقياس الذكاء لدى ذوي العمر 5 سنوات عقلية، لأن السهولة في الإجابة عن الاختبار التي زادت في الارتفاع مع الزيادة في العمر، دليل على صدق الاختبار. ونفس الإجراء نقوم به عند حساب صدق أي اختبار وضعناه لقياس المتغيرات المتعلقة بالنضج، كالسلوك الاجتماعي والسلوك الوجداني والسلوك اللغوي والسلوك المعرفي والسلوك الانفعالي وغيره.

2)التحصيل الدراسي :

ثبت من نتائج العديد من الدراسات أن الارتباط بين الذكاء والتحصيل الدراسي دال إحصائيا، ولذا من الممكن اتخاذ التحصيل الدراسي محكا لحساب صدق الاختبار في الذكاء. فنحسب معامل الارتباط بينهما، فإذا جاء دالا إحصائيا كان ذلك مؤشر على صدق اختبار الذكاء، كما أن مستوى التحصيل الدراسي يصلح أن يكون محكا لحساب صدق اختبار تحصيلي آخر، فلو نصمم اختبار لقياس الاستعداد الرياضي (دراسة الرياضيات) فإننا نقوم بتصميم هذا الاختبار ثم نطلب من أساتذة الرياضيات أن يضعوا تقديرات لتلاميذهم في تحصيلهم لهذه المادة. ويصنفونهم إلى مرتفعين ومتوسطين ومنخفضين في التحصيل. وتعتبر التقديرات التي يضعها الأساتذة لتلاميذهم في تحصيلهم لمادة الرياضيات، بمثابة محك خارجي، ثم نطبق الاختبار الذي وضع على التلاميذ أنفسهم. فإذا تبين أن من حصلوا

على درجات مرتفعة في الاختبار، هم الذين قدرهم أساتذتهم بالممتازين في مادة الرياضيات. وأن من حصلوا على درجات متوسطة في الاختبار، هم الذين قدرهم أساتذتهم بالمتوسطين في مادة الرياضيات، وأن من حصلوا على درجات منخفضة في الاختبار، هم الذين قدرهم أساتذتهم بالضعفاء في تحصيلهم في مادة الرياضيات، فإن هذا التوافق بين الاختبار المصمم وتقديرات الأساتذة، دليل على صدق الاختبار الذي يقيس الاستعداد لدراسة الرياضيات.

3)الأداء في المهنة :

يعتبر الأداء في المهنة مؤشرا جيدا لحساب صدق الاستبيان الذي صمم لقياس الأداء المهني. ولنأخذ الأداء المهني في التدريس كمثال، نقوم بتصميم استبيان لقياس أداء المعلمين في مهنة التدريس، ثم نطلب من المفتشين في مواد تعليمية مختلفة ولمرحلة تعليمية معينة، أن يضعوا تقديرات للمعلمين الذين يشرفون عليهم ويصنفوهم إلى أكفاء ومتوسطين وضعفاء في الاداء التدريس، وتعتبر هذه التقديرات التي يضعها المفتشون للمعلمين، بمثابة محك خارجي ثم نطبق الاستبيان الذي صممناه على نفس المعلمين، فإذا تبين أن مرتفعي الأداء في التدريس حسب تقديرات المفتشين حصلوا على درجات مرتفعة في الاستبيان. وأن متوسطي الأداء في التدريس حصلوا على درجات متوسطة في الاستبيان. وأن منخفضي الأداء في التدريس حصلوا على درجات منخفضة في الاستبيان، فإن هذا التوافق بين الاستبيان المصمم وتقديرات المفتشين، دليل على صدق الاستبيان الذي يقيس الأداء المهني في التدريس، أو نحسب معامل الارتباط بين درجات التقديرات التي منحها المفتشون للمعلمين والدرجات التي حصلوا عليها في الاستبيان، فإذا جاء معامل الارتباط دالا إحصائيا كان ذلك دليلا على صدق الاختبار.

ونقوم بنفس الإجراءات إذا أردنا أن نصمم اختبارا لقياس الأداء عند الموظفين والعمال في وظائف ومهن مختلفة.

4)الأداء في التدريب:

وهو أيضا نوع من الأداء الفعلي يستخدم كمحك لحساب صدق الاستبيانات، يصمم استبيان لقياس الاستعداد لدى الأفراد في أداء التدريب على مهنة معينة (مثل مهنة التمريض أو السكرتارية أو التعليم أو الشرطة وغيرها). ويطبق الاستبيان على هؤلاء الأفراد المختارين وتؤخذ درجاتهم عليه، ثم يتدرب هؤلاء الأفراد فعلا على أداء تلك المهنة، وتمنح لهم درجات على هذا التدريب (محكات). ثم يحسب معامل الارتباط بين درجاتهم على الاستبيان ودرجاتهم على أداء التدريب، فإذا كان دالا إحصائيا دل ذلك على صدق الاستبيان، وإذا لم يكن دالا إحصائيا دل على عدم صدق الاستبيان.

5)الارتباط بين اختبار جديد واختبارات أخرى:

هناك اختبارات واستبيانات ذات شهرة علمية واسعة، وذات سمعة جيدة ومعترف بها من حيث الصدق والثبات والتقنين، مثل اختبار ستانفورد – بيني للذكاء، وقائمة أرون بيك للاكتئاب، وقوائم أيزنك للشخصية وغيرها. إن هذه الاختبارات والقوائم تصلح كمحكات لحساب صدق اختبارات وقوائم أخرى، تم تصميمها حديثا لقياس نفس الخصائص، وذلك بحساب معامل الارتباط بينها، فإذا جاء دالا إحصائيا دل ذلك على صدق الاختبار، وإذا لم يكن دالا إحصائيا دل على عدم صدق الاختبار. وتدل نتائج الدراسات النفسية من ناحية أخرى على أن هناك متغيرات ترتبط فيما بينها ارتباطا دالا إحصائيا، مثل الاكتئاب واليأس، الاكتئاب والقلق، الاكتئاب وتوهم المرض، المثابرة والطموح، والعصابية والخجل.. وغيرها. يمكن كذلك في هذه الحالة تقدير صدق قائمة وضعت لقياس اليأس بحساب معامل الارتباط بينها وبين الاكتئاب. وتقدير قائمة وضعت لقياس الخجل بحساب معامل الارتباط بينها وبين العصابية وهكذا.

6)التقديرات :

المقصود بالتقديرات كما تناولناها سابقا، هو قيام شخص بتقدير خاصية عند شخص أو أشخاص آخرين. وينبغي أن يتوفر القائم بالتقدير على معرفة جيدة بالشخص الذي يقدره،

وخبيرا بالخاصية التي يطلب منه تقديرها، ومتّزنا انفعاليا وكفؤا في عمله وموضوعيا في تقديراته. ومن الأمثلة على ذلك تقدير المعلمين لتلاميذهم والآباء لأبنائهم والإداريين لموظفيهم والضباط لجنودهم وغيرهم. ويمنح المقدرون درجات (محكات) للأشخاص الذين قاموا بتقديرهم في الخاصية موضوع التقدير، ثم يطبق عليهم الاختبار أو الاستبيان الذي يراد حساب صدقه عن طريق ارتباطه بالمحك، ثم يحسب معامل الارتباط بين درجات التقدير (المحكات) والدرجات على الاختبار أو الاستبيان، فإذا جاء معامل الارتباط دالا إحصائيا دل ذلك على صدق الاختبار، وإذا لم يكن دالا إحصائيا دل على عدم صدق الاختبار.

ومن أحسن الأمثلة على الجهود الأمبريقية الناجحة التي قام بها علماء القياس النفسي لحساب صدق اختبار بتقديرات معينة، ما قام به الفرد بيني وتيودور سيمون في فرنسا، عندما قدرا صدق اختبارهما للذكاء بتقديرات المعلمين لتلاميذهم في التحصيل الدراسي.

في نهاية القرن التاسع عشر وبداية القرن العشرين، ظهرت حاجة عملية ملحة للتعرف على المتمدرسين الذين لا يستطيعون الاستفادة من البرامج التعليمية التي يستفيد منها عمل هذان الباحثان على هؤلاء الاطفال في مدارس باريس. فاختارا مجموعة من البنود التي تتطلب التخيل والتذكر والفهم والحكم والتفسير، وقدماها لأطفال المدارس للإجابة عنها، ثم طلبا من المعلمين أن يضعوا تقديرات لتلاميذهم في التحصيل الدراسي، فأوضحت النتائج ما يلي : أولا ،أن الأداء على هذه البنود يتحسن مع التقدم في العمل (محك). ثانيا ، أن الأداء على البنود يرتبط جوهريا بالتقديرات المستقلة للمعلمين عن ذكاء التلميذ ودرجاته التحصيلية (المحك).

شروط اختيار محك مناسب:

لكي يتم تقدير صدق الاختبارات والاستبيانات بشكل جيد، ويتوفر على مستوى عال من الثقة، ينبغي أن تتوفر في المحك المختار الشروط التالية :

- أن يتحدد في مؤشرات سلوكية بحيث يمكن ملاحظتها مباشرة وقياسها.
- أن تكون الطريقة التي يقاس بها المحك واضحة وممكنة.

- أن يكون غير مكلف ويسهل الحصول عليه.
- أن يتيسر جمع المعلومات حوله من حيث الوقت والجهد.
- أن يتوفر على مؤشرات الصدق والثبات.
- أن يكون مستقلا عن الاختبار ولا يتأثر به.
- أن تكون العينة ممثلة للمجتمع الأصلي.
- اختيار الأسلوب الإحصائي المناسب لحساب الصدق بين المحك والاختبار.

الدرس الخـــــــــامس: مستويـــــــات القيـــــــاس

مستويات القياس :

إذا كان القياس هو العملية التي يتم من خلالها التعبير عن الخصائص المختلفة للأشياء أو الحوادث أو الظواهر أو الأشخاص بقيم كمية تحدد بناءًا على قواعد وشروط محددة.

فيقصد بمستويات القياس تحديد أرقام طبقا لقواعد معينة تستخدم لقياس الظاهرة النفسية. (عبد الهادي السيد عبده و فاروق السيد عثمان، 2006: 22)

اقترح العالم الانجليزي Stevens (1951) أربعة مستويات للقياس في ميدان علم النفس والتربية مرتبة من البسيط إلى المعقد تبعا ل :

* مدى استخدامها.
* مدى تطبيق وتدارك العمليات الحسابية عليها.
* وتبعا لدقة الصيانة الكمية/الرقمية للمتغيرات التي تدرسها.

طور العالم كومبسCoombs(1964) هذا التصنيف وحدد العلاقة التي تربط المستويات المختلفة وأضاف فئات جديدة لها. والذي يصنفها كالتالي :

1. **مستوى القياس الاسمي (أو العدي أو التصنيفي):** Nominal measurement

وهو من أدنى و أبسط المستويات وفيه تستخدم الأعداد فقط كعناوين أو أقسام منفصلة للتمييز بين مختلف عناصر أو أعضاء الجسم و لا يتضمن معنى الأفضلية، فالأرقام في المستوى الاسمي ليس لها دلالات كمية وإنما دلالاتها تمييزية فقط، فالرقم 4 ليس أكبر من 3 وإنما يعني فقط أنهما مختلفان.

ونظرا لأن هذه المقاييس ليست كمية فإنها تسمى شبه مقياس (Pseudo measurement)، المستوى الاسمي يشير بالأرقام إلى خصائص الأشخاص والأشياء و في هذا المستوى لا نقيس بل يتم تسمية وتصنيف المشاهدات ويمكن استعمال العد والتكرار والنسبة المئوية و كاف تربيع و التحليل التمييزي.

مثلا: العناوين والتسمية: كأرقام الشوارع والمنازل، أرقام السيارات والهواتف والرقم الجامعي، أرقام الهوية الشخصية وأرقام الجلوس، أنواع السيارات أو لاعبوا كرة القدم...الخ.

و بصفة عامة يقوم التصنيف حسب :

– الديانة: (مسلم، مسيحي، يهودي، لا دين له).

–الحالة النفسية: (مضطرب، سوي...).

– تصنيف الكتب وتصنيف الصخور حسب درجة صلابتها والخامات حسب درجة نقاءها وكذا تصنيف سرعة الرياح...

–الجنس: 1 للذكور و2 للإناث

– الحالة العائلية :(1) أعزب، (2) متزوج، (3) مطلق، (4) أرمل.

مثلا : راسب/ناجح – ذكر / أنثى – علمي / أدبي – مواطن/غير مواطن أبيض/أسود ...

أو مثلا: أرقام الشوارع، أرقام اللاعبين وأرقام الرحلات الجوية وأرقام البطاقات الأحوال المدنية، وأرقام لوحات، السيارة وأرقام قاعات الدراسة.

هذه الأرقام تقوم مقام الأسماء وليس لها دلالة على الكم فالشارع (44) لا يدل رقمه على أية دلالة كمية.

الهدف من هذا المقياس هو مجرد التصنيف فالبيانات التصنيفية يمكن تصنيفها إلى أقسام متشابهة كالكتب في مقابل الصحف والمجلات أو الذكور في مقابل الإناث. لأن معظم أنشطة الانسان تتضمن هذه العمليات التصنيفية. مثلا أهل المدن يرمز لهم بـ 1 وأهل الريف 2.

لا يمكن إجراء العمليات الحسابية وكل ما يمكن إجراءه هو حساب عدد مرات وجود الشيء أو الصفة أي التكرار أو حساب النسبة المئوية أو العشرية و حساب المنوال. العد هو العملية الوحيدة المسموح بها.

2. مستوى القياس الرتبي : (ordinal measurement)

هو أعلى من المستوى الاسمي من حيث الدقة العلمية حيث يرتب الأشياء والأفراد في مجموعات تصاعديا أو تنازليا. ولا تدل الفروق في الترتيب على فروق متساوية في درجات الصفة أو السمة المقاسة أي أن الأرقام المستخدمة في الترتيب لا تقدم معلومات عن كمّ أو مقدار الصفة أو الخاصية المقاسة وإنما تحدد فقط فكرة الزيادة أو النقصان، او بعبارة أخرى فإن قيمة الأرقام في رتبتها وليس في الكم الذي تحتويه. تشمل على فئات تحمل معنى الترتيب ولا تحمل معنى التساوي فعند ترتيب أحمد هو (1) وماهر هو (2) وأيمن هو (3) .. فإن هذا لا يعني بالضرورة أن المسافة بين أحمد وماهر هي نفس المسافة بين الآخرين.

المقاييس الرتبية تفترض أن الأشياء المختلفة تشترك في خاصية شائعة بينها، بينما تختلف في كمية هذه الخاصية دون تحديد كمية الاختلاف.

هذا المستوى الثاني يسمح بترتيب السمات والخصائص دون اعتبار تساوي الفروق بين أي رتبتين منها، فالشخص الذي يتصف أو يتميز بسمة معينة بدرجة أكبر من غيره يكون ترتيبه الأول والشخص الذي يليه في الدرجة يكون ترتيبه الثاني...الخ. حسب المثال الموضح في الجدزل أدناه.

زمن الوصول	30 د	45 د	120 د
الترتيب	1	2	3

مثلا: من المقياس الترتيبي مقياس ديكرت الثلاثي ومقياس ثرستون الخماسي التدرج في ترتيب الأفراد وترتيب الرياح حسب سرعتها (هادئة، أعاصير، نسمة)

يعطي مقياس الرتبة الخاصية أو السمة رموزا أو أرقاما مثلا: الطالب الذي تحصل على 10/10 يعطى له تقدير ممتاز جدا أو 10 نقاط. قد يتساوى إحصاء فرد أو أكثر في امتلاكهم للسمة ويكون لهم نفس الرتبة أو نفس الرقم الدال عليها.

قيمة الأرقام في رتبتها ليس في الكم الذي تحتويه. الوحدات لا تكون متساوية وإن بدت وكأنها متساوية مثلا :

(أ) له 10/10 الأول
(ب) له 10/6 الثاني
(ت) له 5,99 الثالث

مسافات غير متساوية

العمليات الحسابية المستخدمة هي عملية "أكبر من" أو "أصغر من" و لا يمكن إجراء العمليات الحسابية الاربعة (جمع، طرح، ضرب، قسمة) على البيانات الترتيبية ويمكن حساب النسبة المئوية، المنوال، الوسيط، كاف تربيع، اختبار حسن المطابقة وبعض الأساليب الإحصائية البرامترية مثلا: الوسيط، الرباعيات، المئينيات، معامل الارتباط، الرتب لسبيرمان، اختبار مان –وتيني، اختبار ولكوكسون، اختبار الإشارة، طريقة كروسكال – واليز، اختبار فريدمان ...الخ

المقياس الاسمي يناظر ما يسمى بالتصنيف الكيفي أو النوعي، أما المقياس الرتبي فيناظر ما يسمى بالتصنيف الكمي؛ ويحتوي المستوى الثاني على الأول.

شائع جدا في حقل التربية و علم النفس ، التربية الريضاية، الموسيقى، التمثيل، الخط، الغناء ...الخ؛ حيث يتعذر استخدام الاختبارات الموضوعية التي تقيس هذه القدرات بصورة دقيقة.

3. مستوى القياس الفتري أو المستوى الفئوي أو المسافات: (Interval measurement) و يسمى أيضا مستوى المسافات المتساوية و مستوى الوحدة المنظمة أو

المسافة: يعتبر أدق من الناحية الكمية من المقياسين السابقين، يتسم بالمسافات متساوية على المقياس. تهتم بإبراز الفروق في الدرجات بين شيئين أو فردين، اعتمادا على مصدر ثابت تبدأ منه الدرجات فرضيا وباعتبار المسافات المتساوية نسبية، تبدأ من مجرد ظهور أو ملاحظة تزايد في الخاصية، ثم تمتعه بمختلف الصفات السابقة المعلنة فيما يخص كل من المستوى الاسمي والرتبي. تدل على مقادير متساوية في الخاصية أو الصفة أو السمة المقاسة. تتساوى الفروق في هذا المستوى من القياس بين الأقسام المتتالية في السمة المقاسة. حيث يرتبط فيه الحديث عن وحدة القياس التي ترتبط بالرقم أو القيمة الدالة على السمة. فالترمومتر مقسم إلى وحدات متساوية فالفرق بين درجتي الحرارة (30° – 35°) مثلا يساوي الفرق بين (35°–45°). فالأرقام تحمل المعنى الكمي مما يبرر استخدام مفاهيم أكبر /أصغر – أكثر/أقل، وتكون وحدات القياس متساوية وتعكس مقادير متساوية على عكس الرتب مثلا : الفرق بين العلامة 85 و 80 هو 5 علامات. الصفر افتراض وليس حقيقي فإن الأرقام في هذا المستوى لا تخضع لعمليات الجمع والطرح والضرب والقسمة، فلا نقول مثلا : إن شخصا ذكاءه 120° = ذكاء شخصين كل منها 60 درجة.

كذلك إذا خلطنا كأسين من الماء درجة حرارة كل منهما =20 فإننا لا نستطيع القول أن درجة حرارة المزيج = 40.

– مثلا : نقول أن علامات الطالب الأول تزيد عن علامات الطالب الثاني بسبع (7) نقاط، أي أداء الفرد مقارنة بمتوسط أداء الجماعة، وتكون المسافات على شكل انحرافات معيارية.

إن هذا المستوى من القياس يستخدم كثيرا في القياس النفسي والتربوي حيث أن معظم الاختبارات النفسية واختبارات التحصيل هي من هذا النوع. فنحن مثلا لا نقيس سمات شخصية الفرد أو ميوله أو معارفه قياسا مطلقا وإنما نقيس الفروق الموجودة بين شخصين أو توازن بين مجموعتين من الأفراد أو تحدد مكانة فرد في مجموعته فيما يخص سمة من السمات.

من عيوبه عدم وجود صفر مطلق حقيقي يدل على عدم وجود الصفة (مثلا الصفر في الترمومتر لا يعني انعدام الحرارة بل يعني درجة التجمد). هو شائع في: اختبار الذكاء، اختبار القدرات العقلية والاستعدادات و الميول والاتجاهات والاختبارات التحصيلية، الترمومتر الحراري المئوي، مقاييس الرطوبة ومقياس الضغط...الخ.

يمكن إجراء جميع العلميات الحسابية من الجمع والطرح والضرب دون القسمة لأنه لا يوجد لصفر حقيقي. مثلا : عن عدم إمكانية إجراء عملية القسمة في هذا المقياس.

لنفرض أن عمر حصل على 60 درجة وعلي على 30 درجة أضاف الأستاذ لكل واحد 10 نقاط = يصبح لعمر 70 درجة و لعلي 40 درجة.

لكن 60 ÷ 30 = 2 و 70 ÷ 40 = 1,75

يمكن استخدام المتوسط الحسابي، الانحراف المعياري، التباين، معامل الارتباط الخطي لبيرسون، اختبار (T)، تحليل التباين، تحليل الانحدار، تحليل التغاير، التحليل العاملي،...الخ.

يشمل على خصائص القياس الترتيبي والاسمي وتعد المشكلة الرئيسية في القياس الفئوي عدم وجود الصفر كنقطة بداية وتصلح لهذا المستوى الاختبارات البرامترية.

تقع معظم المتغيرات في العلوم الإنسانية و الاجتماعية ضمن هذا المستوى.

مستوى القياس النسبي : (Ratio measurement)

هو المقياس الذي تقدر بواسطته كمية الخاصية في شيء ما بالنسبة إلى وحدات أصبحت موضع اتفاق، أي كل رقم يمكن النظر إليه باعتباره مقياسا لمسافة ما من نقطة الصفر. ومن خلال الاستناد إلى الصفر يمكن معرفة حجم الظاهرة وقياسها. هو أرقى مستويات القياس حيث تستخدم فيه جميع العمليات الحسابية وطرق الإحصاء البرامتري. يتوفر فيه الصفر المطلق إلى جانب تساوي الفروق بين الفترات المختلفة وهذا الصفر المطلق يناظر حقيقة انعدام الظاهرة أو السمة المقاسة، الصفر المطلق يناظر نقطة انعدام الخاصية المقاسة وتعتبر المسطرة العادية التي يستخدمها الطلاب مثلا للقياس النسبي.

إن تسمية هذا النوع من المقياس باسم النسبة جاءت من إمكانية التعبير عنها في صورة نسبة. القياس النسبي يناسب بشكل خاص المتغيرات الفيزيائية.

يسعى علماء علم النفس إلى بناء نماذج رياضية تستخدم لبناء مقاييس الذكاء والتحصيل والاتجاهات يتوفر فيها الصفر المطلق الذي يناظر حقيقة نقطة انعدام الظاهرة أو السمة المقاسة. نقول مثلا درجة الحرارة اليوم 28°م استنادا إلى صفر حقيقي مطلق. جميع مقاييس الطول، الوزن، المساحة والحجم هي من هذا النوع، لها صفرا مطلقا يدل على عدم وجود الصفة. وحداته متساوية. يمكن قياس الصفات بطرق مباشرة بواسطة وحدات قياس معيارية. النسبة بين أي درجتين على المقياس لا تتأثر بوحدات القياس فهي ثابتة، مثلا عندما نقيس الوزن بـ (غ) ومضاعفاته ثم نقيس بوحدات الأوقية فإن النسبة المئوية بين الوزنين تكون ثابتة. يمكن استخدام جميع العمليات الحسابية بما فيها القسمة. استخدام مختلف المعالجات الإحصائية. من أمثلة : قياس الوزن بوحدات النظام الفرنسي (الغرام ومضاعفاته).

في ميدان علم النفس، نادرا ما يتم استعمالها إلا عند قياس الخصائص العقلية أو الوجدانية بوحدات فيزيقية مثل قياس زمن الرجع، قياس التعلم بوحدات الزمن وقياس أطوال أو أوزان الأطفال.

في العلوم الطبيعية: قياس الأطوال (بالمتر ومضاعفاته)، الوزن (غ ومضاعفاته)، الزمن.

ويمكن إجراء جميع العمليات الحسابية الأساسية على الأعداد أو الأرقام الناتجة عن علميات القياس النسبي لوجود الصفر المطلق اضافة الى طرق الإحصاء الوصفي: مقاييس النزعة المركزية (الوسيط، المنوال..) مقاييس التشتت، التباين (المدى، الانحراف عن المتوسط، الانحراف المعياري، التباين)، مقاييس الارتباط.

ملاحظات :

- إن ترتيب مستويات القياس يعني أن كل مقياس من المقاييس الأربعة يمتلك خصائص القياس التي تسبقها بالإضافة إلى خاصية مميزة له وهنا تظهر هرمية القياس، كما يمكن تحويل الكميات في مقياس معين إلى مقياس آخر يسبقه في الترتيب وليس العكس.
- الأساليب الإحصائية تختلف من مستوى إلى آخر.
- إن ترتيب المقاييس السابقة الذكر يعود الى أن كل مقياس يمتلك خصائص المقاييس التي سبقتها خاصية مميزة له. كما يمكن تحويل الكميات في مقياس معين إلى مقياس آخر يسبقه في الترتيب وليس العكس.
- إن اختبار مستوى القياس يعتمد بالدرجة الأولى على طبيعة السمة المقاسة.
- فبعض السمات أو الخصائص لا يمكن أن تقاس إلا بمقياس الرتب، وهناك سمات أخرى يمكن قياسها بمختلف أنواع المقاييس.
- كثير من المقاييس في علم النفس لا يمكن اعتبارها مقاييس نسبة، وإن كان عدد قليل منها يمكن اعتباره كذلك كقياس خاصية عقلية عن طريق وحدات فيزيائية من نوع خاص: زمن الرجع، سرعة استجابة سائق السيارة بالضغط على الكوابح...
- نظريات القياس سواء في العلوم السلوكية أو العلوم التربوية تضخع لكثير من المفاهيم الرياضية والإحصائية في أساسها وفي إجراءاتها وفي طرق استنتاج النتائج والتأكد من مطابقتها لنماذج النظرية الشائعة.
- يكون القياس النسبي عادة منحصرا في العلوم الطبيعية و حيث أنه من النادر ما يصل القياس في العلوم النفسية إلى هذا المستوى.
- تعطينا معظم الاختبارات وصفا كميا مباشرا لأداء الشخص ويطلق على الوصف الكمي للأداء اسم الدرجة الخام وقد تكون الدرجة الخام:

*- عدد الأسئلة التي أجاب عليها المفحوص صحيحة و خاطئة.

*- الزمن الذي استغرقه لذلك.

*–النسبة المئوية لعدد الإجابات الصحيحة أو الخاطئة...الخ.

يفسر معظم الأفراد الدرجة الخام ويستخدمونها دون معرفة لجوانب القصور فيها، لأن هذه الدرجات تتميز بالسهولة في الحصول عليها، ولأننا ألفناها من خبرتنا الطويلة بالامتحانات المدرسية. مع ذلك فالدرجة الخام في الاختبار النفسي لا معنى لها إلا باللجوء إلى معالجة أخرى. نؤكد أننا لا نستطيع تفسير الدرجات في الاختبار النفسي كما نفعل بالقيم التي نحصل عليها من مقياس الظواهر الطبيعية كالطول مثلا، لأن مقياس الظواهر الطبيعية الطول هي من نوع مقاييس النسبة، أي لها صفر مطلق وتتألف من وحدات متساوية ولذلك فإن هذه المقاييس تسمح لنا بتفسير مباشر للقيم التي نحصل عليها منها.

أما في القياس النفسي فلا نستطيع ذلك لأن المقاييس في هذه الحالة لا تكون في أحسن صورتها إلا من نوع مقاييس المسافة. مثلا : حصل الطفل (س) على 10 % من 3 أسئلة في اختبار في معاني الكلمات، فهل يعني هذا أن الطفل لا يعرف إلا عشر كلمات ؟ لا طبعا. فقد لا يكون الفاحص مهتما إلا بالكلمات الصعبة غير المألوفة لدى الطفل ولهذا نجد أن محصوله اللفظي قد يكون أكبر في اختبار أقل صعوبة. ولو حصل الطفل على علامة (0) فهذا لا يعني أن الطفل ليس لديه أي محصول لغوي على الإطلاق. فالدرجة التي تدل على 65 % صواب في أحد الاختبارات قد تتساوى مع 30 % في اختبار آخر أصعب ومع 80 % في اختبار آخر أسهل.

فالفروق بين الأفراد في الدرجات الخام للاختبارات النفسية لا يدل بالضرورة على وجود مسافات حقيقية بين الأفراد لأن هذه الاختبارات لا يتحقق فيها الشرطان الأساسيان لمقاييس المسافة وهما الصفر المطلق والوحدات المتساوية مثلا :

الفرد (أ) حصل على 53.

الفرد (ب) حصل على 56.

الفرد (ج) حصل على 59.

قد تكون الفروق بين الدرجات الخام متساوية، مع ذلك لا يدل هذا على أن الفرق بين (أ و ب) يساوي تماما الفرق بين (ب و ج).

– من النادر اختبار كل فرد في مجتمع معين حيث أن عضوية الفرد في المجتمع تتغير بصورة مستمرة، فبعض الأطفال الذين ينتمون إلى مجتمع 6 سنوات من العمر اليوم سوف يصبحون من مجتمع 7 سنوات غدا.

– يتوقف اختيار مستوى القياس المناسب على:

*– نوع المشكلة وهي موضوع الدراسة.

*– طريقة صياغة هذه المشكلة.

*–التحديد الجيد للمشكلة في شكل تعاريف إجرائية.

*–فرضيات واضحة تساعد على اختيار مستوى القياس المناسب.

(فخري رشيد خضر، 2003: 56– 60)–(بشير معمرية، 2007: 78–87)–

(عبد الهادي السيد عبده و فاروق السيد عثمان، 2006: 22–23)

الدرس السادس: مفهوم الخصائص السيكومترية للمقاييس

-تعريف الخصائص السيكومترية:

هي عبارة عن مجموعة مؤشرات تعبر عن إمكانية الثقة في نتائج المقياس واستقرار نتائجه واتساقها كما أنها تعبر عن الأسس التي يعتمد عليها المقياس في تفسير نتائجه.

مفهوم الصدق validity :

يعد مفهوم الصدق Validity واحدا من أكثر المفاهيم الأساسية أهمية في مجال القياس النفسي. يعد أحد المؤشرات التي تدل على مصداقية وجودة المقياس، فالمقياس الجيد هو الذي يقيس السمة التي يهدف إلى قياسها وتختلف الاختبارات في درجات صدقها تبعا لاقترابها أو ابتعادها وتقدير تلك السمة التي تهدف إلى قياسها.

وإذا أردنا أن نعرف الصدق، فإننا نجد تعريفات كثيرة ومتنوعة. نورد فقط التعريفات الثلاثة التالية :

1-تعريف جيولكس Guliksen (1950): الصدق هو ارتباط الاختبار ببعض المحكات.

2- تعريف كيورتن Cureton (1950): الصدق هو تقدير للارتباطات بين الدرجة الخام للاختيار والحقيقة الثابتة ثباتا تاما".

3-تعريف جلفوردGuilford (1954): الصدق هو تحديد لمعامل الارتباط بين الاختبار وبعض مقاييس أو محكات الأداء في مواقف الحياة".

نلاحظ على التعريفات السابقة للصدق، أنها توجب توفر محك، ثم يتم ربط الاختبار به. أي أنها تركز على الصدق المرتبط بالمحك.

أما التعريف العام والشائع جدا للصدق، هو أن الاختبار يعتبر صادقا، إذا كان يقيس ما وضع لقياسه. وهذا يجعلنا نفكر في المتر والغرام واللتر كأدوات قياس، واختبارها لكل من الأطوال والأوزان والسوائل، فهي اختبارات صادقة، لأنها تقيس فعلا ما وضعت لقياسه، ولا تقيس أشياء غيرها.

ووفقا لهذا التعريف، هناك ثلاثة مفاهيم أساسية تتعلق بصدق الاختبار. بمعنى أنه لا يكون الاختبار صادقا إلا إذا توفر ما يلي :

1-أن يكون الاختبار قادرا على قياس ما وضع لقياسه، أي ان تكون بنود الاختبار على علاقة وثيقة بالخاصية التي تقيسها. فالاختبار الذي صمم لقياس القدرة الرياضية مثلا، يجب أن يكون واضحا أنه يقيس هذه القدرة من خلال علاقة بنوده بمكونات هذه القدرة وعناصرها. وكذلك الاستبيان الذي وضع لقياس قلق الامتحان، أو لقياس سمة المثابرة، أو لقياس مصدر الضبط وغيرها.

2-أن يكون الاختبار قادرا على قياس ما وضع لقياسه فقط. أي أن يكون قادرا على أن يميز بين الخاصية التي يقيسها، والخصائص الاخرى التي يمكن أن ترتبط بها. فاختبار القدرة الرياضية، ينبغي ألا يتأثر بالقدرة اللغوية مثلا، واختبار يقيس المثابرة، ينبغي أن يميز بينها وبين سمات أخرى قد ترتبط بها مثل الدافعية الداخلية والطموح والدافع للإنجاز.

3-أن يكون الاختبار قادرا على التمييز بين طرفي الخاصية التي يقيسها. أي أن يميز بين الأداء القوي والاداء الضعيف لدى أفراد العينة. أما إذا كانت درجات الاختبار جميعها متقاربة، دل ذلك على صدق ضعيف، لأن الاختبار لم يقم في حقيقة الأمر بالمهمة الأساسية في عملية القياس، وهي إظهار الفروق بين أفراد العينة. فالاستبيان الذي وضع

لقياس سمة القلق لدى عينة معينة، عليه أن يبين أن هناك مرتفعي القلق ومتوسطي القلق ومنخفضي القلق بين أفراد هذه العينة، وإلا كان استبيانا غير صادق.

هذه هي المفاهيم الأساسية الثلاثة لصدق الاختبار، وهي أيضا الأسس التي يمكن أن نشير منها إلى أنواع الصدق والطرق المختلفة لحسابه.

ويذكر أخصائيو القياس النفسي، أنه عند دراسة صدق الاختبار، ينبغي الاهتمام بالإجابة عن سؤالين :

الأول: ماذا يقيس الاختبار؟ الثاني : ما هي درجة الدقة التي يقيس بها الاختبار ما وضع لقياسه ؟

الصدق صفة نسبية أي أن الاختبار الصادق بالنسبة لهدف معين لا يكون كذلك بالنسبة لهدف آخر. والاختبارات توصف بالصدق فقط بالقياس إلى الغاية المقصودة من استخدامها، كذلك ما يصدق من الاختبارات على الأطفال لا يصدق على الكبار، وما يصدق على الذكور قد لا يصدق على الإناث، والاختبار الذي ثبت صدقه في بلد ما قد لا يثبت صدقه في بلد آخر كاختبار الذكاء والشخصية.

إن مفهوم الصدق ليس درجة كاملة أو صدقا تاما على أن هناك حدا أدنى لكي يعد الاختبار صادقا.

إن أهم ما يؤخذ في الاعتبار هو الصدق الذي يتصل مباشرة بالسؤال الذي يقول: "ماذا يقيس الاختبار ؟" وإذا لم يكن لدينا إجابة كافية بدرجة معقولة لهذا السؤال، فإن أي اختبار سوف يكون عديم الفائدة في محاولاتنا للتعامل بحكمة مع البشر، صغار أو كبار، نحن أو الآخرين. بل ربما كان الموقف أكثر سوءا من عدم الفائدة، لأننا لو اعتمدنا على الافتراض الخاطئ فيما يتعلق بمعنى درجات الفرد، فقد نوجهه إلى قرارات قد تؤدي إلى عدم التوافق والوقوع في أخطاء كبيرة.

وهناك كثير من الأشخاص لا يستطيعون معرفة أن عنوان الاختبار لا يقول لنا في الحقيقة شيئا على الإطلاق عما يقيسه هذا الاختبار. فكل إنسان يستطيع أن يكتب مجموعة من الأسئلة، ويكون واثقا من أن هذه الأسئلة تدور حول عملية "التفكير" أو "الاستعداد الميكانيكي" أو "المرونة المعرفية"، ولكن معرفة أي العمليات العقلية يجب أن يستخدمها المفحوص في الواقع للإجابة على هذه الأسئلة، فهي مهمة شاقة وطويلة. فقد يبني شخص ما اختبارا لقياس "القدرة على التفكير الاستدلالي" وربما يثبت أن هذا الاختبار إنما يقيس اتجاهات الطبقة المتوسطة. وقد يحدث أن اختبارا في "الاستعداد الميكانيكي" قد يقيس فعلا الذكاء العام، وأن اختبارا في "النضج الانفعالي" قد يقيس فقط ما يعرفه المفحوصون عن التقبل الاجتماعي لنمط معين من السلوك. إن المهارة الأولى التي يجب أن يجيدها الفرد عند تقييم صدق الاختبارات هي عدم أخذ عناوين الاختبارات في الاعتبار. فالأمر المهم هو ليس التسمية التي يضعها المؤلف لهذا الاختبار، ولكن بماذا تتعلق درجات هذا الاختبار.

كانت الطريقة المقبولة في حركة القياس هي أن يعرف الفرد ماذا يريد أن يقيس، ثم يجمع بعد ذلك الأدلة والبراهين لإثبات مدى نجاح الاختبار في قياس ما أراده الفرد، وخلال هذه الفترة، كما رأينا سابقا، كان معظم الجهد موجها لقياس الذكاء العام، وقد كان واضحا، أنه على الرغم من تزايد عدد الباحثين، إلا أنهم لم يتفقوا جميعا على تعريف محدد لهذا المفهوم المحيّر، لكنهم قد افترضوا جميعا، من أجل الأغراض العملية، أن الذكاء هو السمة التي يعتمد عليها الحكم بالتفوق أو الغباء في الدراسة، ومن ثم بحثوا في المواقف المدرسية عن البرهان على صدق اختبارات الذكاء، وكان الطريق هو تقدير مقدار العلاقة بين درجة النجاح في الدراسة، كما يمكن أن يتم التنبؤ بها من درجات الاختبار ودرجة النجاح الحقيقية. وبناء على الأدلة التي حصلوا عليها، تم استبعاد القياسات المختبرية مثل التمييز، وزمن الرجع، وما إلى ذلك، واستبقيت الأسئلة من نوع اختبار بينيه ذات الاستجابات المركبة لتكون أساسا لاختبارات الذكاء.

وقياسات إحدى السمات النفسية في موقف ما من مواقف الحياة الحقيقية، مثل الفصل الدراسي، تسمى قياسات المحك، والارتباطات بين درجات الاختبار وقياسات المحك هذه تسمى معاملات الصدق.

فمنذ البداية، إذن، كان الصدق يعني مدى قياس الاختبار لما وضع لقياسه، و كان يحسب معامل الارتباط بين درجات الاختبار وبين مقاييس المحك لمعرفة صدق الاختبار، ولكن بمضي الوقت اتضح أن هناك تعقيدات في كل من الطريقة والمفهوم. ومما يؤسف له، أنه لم يكن من الممكن في معظم الحالات أن نجد محكا واحدا يمكن أن يعتبر دليلا واضحا على سمة عقلية. فقد يقرر باحثان يدرسان نفس القدرة – مثل الاستعداد الميكانيكي – استخدام محكات مختلفة، وبالتالي قد يحصل كل منهما على نتيجة تختلف عن الآخر. فالباحث الأول قد يرى أن المحك المناسب هو درجات التلاميذ في مقرر الأعمال (الأشغال) اليدوية، والباحث الثاني قد يرى أن المحك المناسب هو طول الوقت الذي يستغرقه الأفراد الجدد لتعلم مهارة ميكانيكية بسيطة في المصنع. والآن، ماذا نعمل عندما نجد أن معامل الارتباط بين درجات الاختبار والمحك هو في الحالة الأولى 6، 0 وفي الحالة الثانية 59، كيف يتسنى لنا أن نقرر ما إذا كان الاختبار صادقا إذا أعطانا نتائج من هذا النوع ؟ فهل هذا الاختبار يقيس الاستعداد الميكانيكي أم لا ؟

لقد اتضح من واقع الخبرة، في مثل الحالة السابقة وأمثالها، أن عملية التحقق من صدق الاختبار إنما هي عملية طويلة وهي لا تقوم على حادثة مفردة. ونحن نستطيع فقط أن نفهم ماذا يقيس الاختبار حقا، إذا ما قمنا بدارسة ارتباطات الاختبار بمجموعة مختلفة من المحكات، وإذا توفرت سلسلة من البحوث حول اختبار "الاستعداد الميكانيكي" – على سبيل المثال– فقد توضح أنه يقيس في الحقيقة القدرة على القيام بالحركات الماهرة ذات الضبط الدقيق ولا علاقة له بالقدرة على فهم العلاقات المعقدة بين الأجزاء الميكانيكية.

ومن ثم، فقد يرتبط هذا الاختبار ارتباطا عاليا مع درجات أعمال النجارة ، ولكنه لا يرتبط ارتباطا عاليا مع درجات الأعمال الميكانيكية. وقد يساعد هذا الاختبار على اختيار عمال قادرين لمصنع دون آخر.

ولذلك، فبدلا من أن نسأل السؤال التقليدي عن الصدق وهو "إلى أي مدى يقيس هذا الاختبار ما وضع لقياسه ؟" فإننا نميل الآن إلى أن نسأل : ماذا يقيس هذا الاختبار فعلا ؟ ونحن نعلم الآن أنه يجب أن نقوم بتحليل محتوى الاختبار ونفحص كثيرا من معاملات الارتباط بين درجات الاختبار ومجموعة مختلفة من المحكات في جماعات مختلفة قبل أن نصل إلى إجابة لهذا السؤال. ومن خلال هذا العمل فإن معلوماتنا عن السمات العقلية الأساسية، وكيف يرتبط كل منها بالأخرى، إنما تزيد باستمرار وتتغير، كما أن معلوماتنا حول الاختبارات ذاتها تزيد أيضا وتتغير، فليس من الضروري في الواقع أن يحدد الأخصائي النفسي بدقة في بداية بحثه السمة التي يأمل أن يقيسها الاختبار، فإذا كانت لديه فكرة عامة عن هذه السمة وعلاقتها بالمفاهيم النظرية أو بالمواقف العملية، فإن دقته في تحديد معنى السمة سوف تزيد كلما جرب الاختبار في سلسلة من البحوث المنفصلة.

إن فهم مضمون عملية التحقق من صدق الاختبارات يحمل ثلاث معان واضحة لمستخدمي الاختبارات. أولها، أنه إذا كان الاختبار سوف يستخدم في اتخاذ قرارات حول الأفراد أو الجماعات، فلابد من دراسة كل البيانات المتوفرة قبل أي محاولة لتفسير درجات الاختبار. وهذا يعني أن من يستخدم الاختبار يجب أن يكون قادرا على إجراء مثل هذه الدراسة، "فبرنامج تطبيق الاختبارات" في المدارس أو الصناعة يمكن أن يؤدي وظيفته بفعالية في حالة ما إذا أشرف عليه شخص على درجة عالية من القدرة والكفاءة.

والمعنى الثاني هو أنه عندما يستخدم الاختبار من أجل التنبؤ أو الاختبار (مثل اختبار المتقدمين لعمل ما)، فإنه يجب أن يعين صدقه في الموقف النوعي الذي سوف يستخدم فيه، كلما كان ذلك ممكنا، وهذا يتطلب بطبيعة الحال، دراسة خاصة في نفس الموقف الذي سيستخدم فيه الاختبار، وذلك قبل أن نبدأ باستخدام الاختبار بصورة روتينية في انتقاء

المتقدمين. وفي حالة التخطيط لمثل هذه الدراسة فإن السؤال المناسب الذي علينا أن نسأله فيما يتعلق بصدق الاختبار لن يكون : "هل هذا الاختبار صادق ؟" ولكن سوف يكون السؤال : " هل يضيف هذا الاختبار جديدا لصدق الطرق المستخدمة حاليا في عملية الاختبار ؟". فقد ترى إحدى الكليات التي تطبق معايير محددة للاختبار على أساس درجات المدرسة الثانوية أن ما يسمى باختبار "الاستعداد الجامعي" لا علاقة له بالبرنامج المطبق في هذه الكلية. وقد ترى كلية أخرى ذات برنامج مختلف، وتستمد طلبتها من نوعيات مختلفة، أن اختبار "الاستعداد الجامعي" ذاته على درجة عالية من الصدق. غير أنه ليست جميع المواقف تسمح بإجراء صدق الاختبار بهذه الطريقة المباشرة السريعة، ولكننا نوصي بها إذا سمح الموقف بذلك.

يعد الصدق أكثر المفاهيم الأساسية أهمية في مجال القياس النفسي، إن لم يكن أهمها على الإطلاق، ورغم هذا فهو أكثرها إثارة للمناقشة والجدل، وتنوعا في التعريف. وهذا التنوع والاختلاف في تعريف مفهوم الصدق، يدور بعضها حول صلاحية المفهوم نفسه، ويدور بعضها الآخر حول فلسفة المفهوم في صلته بأغراض القياس، ويدور البعض الآخر حول إجرائية التعريفات وتطبيقاتها.

ولكن هذه الاختلافات أدت من جانب آخر إلى إثراء المفهوم وتعميق أسسه، بالإضافة إلى ما أسهمت به بعض المفاهيم الحديثة مثل "صدق التكوين" الذي يربط التنظير في علم النفس بالقياس وأغراضه، وهو الربط الذي ظل مفقودا لفترة طويلة، نتيجة الفلسفة العامة في القياس التي كان محور اهتمامها الكشف عن الفروق الفردية، وصياغة هذه الفروق وفقا لمعادلات رياضية وإحصائية، تحكمها قوانين الاحتمالات، دون تجاوز ذلك إلى مستوى التكوينات الفرضية ذات الأسس النظرية الواسعة.

يقاس الصدق بحساب معامل ارتباط الاختبار بالميزان (المحك) الذي يدل على المقياس الصادق للظاهرة ومن صوره الصدق التلازمي والصدق التنبؤي ويسمى هذا بمعامل

الصدق، كما يقاس الصدق من خلال ما يطلق عليه صدق التكوين الفرضي أو صدق البناء.

العوامل التي تؤثر على صدق الاختبار :

1-طول الاختبار: يزداد صدق الاختبار تبعا لزيادة عدد الأسئلة، طول الاختبار يقلل من أخطاء القياس، وبما أن الصدق يعتمد على الثبات والثبات يعتمد على طول الاختبار إذن فالصدق يعتمد على طول الاختبار.

2-ثبات الاختبار: يتأثر صدق الاختبار بالقيمة العددية لمعامل ثبات الاختبار تأثرا طرديا فيزداد الصدق تبعا لزيادة ثبات الاختبار.

3-تباين المحك كلما زاد تباين درجات المحك بين المجموعة العليا والمجموعة الدنيا أدى ذلك إلى زيادة قيمة معامل صدق الاختبار، بمعنى أن صدق الاختبار يتناسب طرديا مع تباين درجات الأفراد على المحك.

4-التباين (تباين الاختبار والمحك) بما أن الصدق هو صورة من صور الارتباط القائم بين الاختبار، والمحك، فالتباين الضعيف يقلل من أثر الصدق، والتباين القوي بين (الاختبار والمحك) يزيد من أثر الصدق. (فخري رشيد خضر، 2003: 76)

ملاحظة:

- يتأثر الصدق بثبات الاختبار وليس العكس وكذلك بثبات الميزان وجميع العوامل التي تؤثر في الثبات السابق عرضها.
- إن الاختبار الصادق ثابت بالضرورة ولكن الاختبار الثابت ليس صادقا بالضرورة.
- فيما يتعلق بمعامل الصدق، فهو أحد تطبيقات معامل الارتباط بين درجات الأفراد في الاختبار ودرجات في الوظائف، فإذا كان الارتباط كبير وكان معامل الصدق كبيرا كان الاختبار صادقا.

و من خلال ما سبق نحوصل في الجدول التالي أهم أنواع الصدق و تعريفها:

نوع الصدق	تعريفه
صدق المحتوى	درجة تمثيل فقرات الاختبار أو محتويات لجونب المتغير او المفهوم المراد قياسه. أو الدرجة التي تقيس فقرات فقرات المقياس من خلالها جميع جوانب المفهوم موضوع القياس بشكل كاف.
الصدق التنبؤي	الدرجة التي يمكن ان تتنبأ(ترتبط) درجات الاختبار من خلالها بالسلوك ، او الدرجات التي يمكن ملاحظتها أو الحصول عليها في وقت من الأوقات في المستقبل أي يطبق في وقت لاحق.
الصدق التلازمي	مدى ارتباط درجات الاختبار بمجموعة من الدرجات المستمد من اختبار او سلوكات أخرى ذات صلة بها، لكنها مستقلة عنها أي يطبق في الوقت نفسه. أو هو أحد أنواع الصدق المرتبط بمحك، ويشير الى مدى ارتباط الدرجات بدرجات على مقاييس أخرى حول مواضيع ذات علاقة ثم تطبيقها في الوقت نفسه
صدق البناء (المفهوم)	مدى ارتباط درجات الاختبار بمقاييس أو سلوكات أخرى بطريقة متسقة منطقيا و نظريا. و يشمل هذا التحديد التقارب و الصدق التمييزي.
الصدق المرتبط بمحك	مدى تنبؤ ارتباط درجات المقياس بدرجات مقاييس أخرى ذات علاقة
ملاحظة: الصدق التنبؤي و الصدق التلازمي نوعيان فرعيان للصدق المرتبط بمحك	

(تيموثي ترول،2007: 300)

مفهوم الثبـــات:

يقصد بالثبات مدى استقرار ظاهرة معينة في مناسبات مختلفة. فالفرد مثلا الذي يتميز بالاستقرار في سلوكه، يعني أن السلوك الذي سيؤتيه في موقف معين (السيطرة مثلا) غدا أو

في الأسبوع القادم، هو نفس السلوك الذي صدر منه اليوم أو الأسبوع الماضي. والفرد الثابت في سلوكه يمكن التنبؤ بسلوكه، أما الفرد غير الثابت في سلوكه، فهو الذي يتغير سلوكه من يوم لآخر، وبالتالي يصعب التنبؤ بسلوكه في المواقف المختلفة.

من الشروط السيكومترية للاختبار الجيد الثبات وهو يدل على اتساق ترتيب الأفراد عندما يطبق عليهم الاختبار أكثر من مرة، وهو يدل على حصول الأفراد على نفس الدرجات عندما يطبق عليهم الاختبار في مرات متتالية ويعرف الثبات بأنه النسبة بين التباين الحقيقي والتباين الكلي لدرجات المفحوصين.

ومن هنا فإن مفهوم "الثبات" من المفاهيم الجوهرية في القياس النفسي، ويمثل مع مفهوم الصدق أهم الأسس التي تقوم عليها إجراءات إعداد المقاييس والاختبارات للاستخدام.

إلا أن بعض السيكولوجيين ينتقدون هذا المفهوم، سواء من حيث كونه مصطلحا أو من حيث فائدته في عملية إعداد الاختبارات.

فيشير أحمد محمد عبد الخالق(1991) إلى أن الثبات Reliability أصبح مفهوما كلاسيكيا، ويفضل كثير من علماء النفس عليه مصطلحا حديثا هو إمكان التعميم Generalizability.

ويذكر ن.د. سند بيرج N.D. Sundberg(1977) أن بعض المنظرين في الاختبارات، ينتقدون كلا من المصطلح القديم وبديله الأحدث، ويضعون تركيزا أكبر على الصدق والاستخدام المباشر لجداول التنبؤ، أكثر من التقديرات المعتمدة على حساب الأخطاء المعيارية للمقاييس (وهي إحدى الفوائد العملية لحساب الثبات).

وينبه أحمد محمد عبد الخالق أيضا إلى أن الترجمة العربية للمصطلح Reliability بكلمة الثبات، ليست دقيقة ولا شاملة لمختلف معاني المصطلح، إذ يشير اللفظ الانجليزي إلى ما يمكن أن نسميه "الاعتمادية" أو درجة الركون إلى ... والثقة في نتائج الاختبار، ويشتمل ذلك على ثبات النتائج وعدم تغيرها.

والنقطة الهامة التي ينبغي توضيحها، هي أن الاختبار نفسه بوصفه مجموعة من بنود أو أعمال معينة لا ثبات له، إن الثبات هو ثبات الإجابة على الاختبار، ثبات الأداء على الاختبار. بهذا فإن الذي يحسب ثباته، إنما هو عينة استجابات نحصل عليها من مجموعة من الأفراد. ولأن لكل مجموعة من الأفراد خصائصها، فمن الضروري أن نحصل على أكثر من معامل ثبات للاختبار الواحد لدى أكثر من عينة. ولا يعني حصولنا على معامل ثبات لأداء عينة من الراشدين على اختبار معين، إن ثبات أداء المراهقين والأطفال في المستوى نفسه، كما لا يجوز أن نقبل دون تحفظ ثبات أداء عينة الذكور، على أنه مطابق لثبات أداء عينة الإناث على الاختبار نفسه. وتلعب الفروق الحضارية والثقافية والجنسية والعمرية والتعليمية وغيرها، دورا هاما في التأثير على ثبات المقياس.

إذن، يعني الثبات مدى الدقة والاستقرار والاتساق في نتائج الأداة لو طبقت مرتين فأكثر على نفس الخاصية في مناسبات مختلفة. فمثلا، لو قام أحد الأفراد بقياس طول قطعة من القماش فوجد أن طولها 80 سم، ثم قام غيره بقياس نفس القطعة سيجدها أيضا 80 سم، ولو فعل شخص ثالث نفس الشيء لوجد أيضا نفس كمية الطول. لأن المقياس الذي تم به القياس ثابت، لا يتغير من شخص إلى آخر ومن وقت إلى آخر. ويحاول السيكولوجيون أن تكون الاختبارات النفسية ثابتة أيضا في النتائج التي تعطيها.

والثبات في القياس النفسي هو باختصار : ضمان الحصول على نفس النتائج تقريبا إذا أعيد تطبيق الاختبار على نفس الفرد أو نفس المجموعة من الأفراد. وهذا يعني قلة تأثير عوامل الصدفة على نتائج الاختبار، مثلا، لو يطبق اختبار ستانفورد – بيني للذكاء على طفل عمره 101 شهر في أول أيام الأسبوع، فيكون معامل ذكائه : 120. وفي آخر الأسبوع عندما يطبق نفس الاختبار على نفس الطفل، فيكون معامل ذكائه : 122. ففي هذه الحالة يعتبر الاختبار ثابتا، لأن النتيجتين (المعاملين) قريبتان جدا من بعضهما، وبالتالي يوجد استقرار واتساق في نتائج الاختبار. أما إذا وجد أن معامل ذكائه في المرة الثانية هو : 90، ففي هذه الحالة لا نثق في نتائج هذا الاختبار لأنها غير مستقرة ولا متسقة. و من هنا يمكن

أن نستنتج العلاقة القوية بين بنود الاختبار والأداء الحقيقي للفرد. والمقصود بالأداء أنه دالة للخاصية محل القياس.

إن الصدق هو أهم ما يجب أن يؤخذ في الاعتبار عند بناء الاختبارات بكافة أنواعها، وكذلك عند استخدامها، يلي ذلك في الأهمية الثبات، يمكننا أن نبدأ مرة أخرى بسؤال هو : "بأي قدر من الدقة، أو بأي قدر من الاتساق، يقيس هذا الاختبار السمة التي يقيسها، أيا كانت هذه السمة ؟

عندما يؤدي الفرد اختبارا ما فإن درجته على هذا الاختبار قد تتأثر بعوامل عديدة بعيدة كل البعد عن القدرة أو سمة الشخصية التي صمم الاختبار من أجل قياسها. فإذا تشتت انتباه المفحوص أو كان متمردا فقد لا يعطي إجابات صحيحة للأسئلة مثلما يعطي عندما يكون مركزا على الاختبار، وبالتالي فإنه يحصل على درجة أقل مما يجب. ومن ناحية أخرى، فإنه لو حدث وشاهد المفحوص – في الليلة السابقة للاختبار– برنامجا في التلفزيون أو اشترك في مناقشة عابرة والتقى ببعض الأسئلة الخاصة التي يحتويها الاختبار، فإنه في هذه الحالة يحتمل أن يحصل على درجة أعلى مما يؤهله له أداؤه في الظروف العادية. وهذان مثالان على مدى تأثر درجات الاختبار بعوامل الصدفة. وهناك عوامل أخرى مثل ردود الفعل الانفعالية بالنسبة للفاحص، ودرجة الحرارة والتهوية في غرفة الاختبار، والحظ الحسن أو السيء في تخمين إجابات الأسئلة التي لا يعرفها المفحوص، وهكذا. وهذه العوامل لا يمكن التنبؤ بها، وهي لا تؤثر في جميع المفحوصين بنفس الطريقة أو بنفس الدرجة.

إن ما يجب على الأخصائي الذي يبني الاختبار أن يكتشفه ويوصله إلى الأخصائي الذي يستخدم الاختبار هو إلى أي مدى سوف تتأثر دقة تقدير درجات الأفراد بعوامل الصدفة هذه، فإذا توفرت هذه المعلومات، استطاع الأخصائي الذي يفسر درجات الاختبار أن يأخذ ذلك في اعتباره، وسوف تكون خلاصة تقريره عن شخص ما على النحو التالي: " الفرد أ ذكاؤه فوق المتوسط بقليل، فهناك احتمال 50–50 إن نسبة ذكائه تقع بين 105 و 115 ولكنه لن يستنتج إن الشخص أ له نسبة ذكاء تساوي 110.

وهناك عدة مفاهيم لمعنى ثبات الاختبار يمكن أن نشير إليها بحيث لا يكون الاختبار ثابتا إلا إذا تحقق ما يلي :

1-أن يعطي الاختبار نفس النتائج تقريبا إذا أعيد تطبيقه على نفس المجموعة من الأفراد.

2-أن تكون هناك علاقة قانونية بين بنود الاختبار، مما يدل على التناسق في البناء الداخلي للاختبار، وهذا يعني أن معامل ثبات الاختبار سوف يتوقف على الارتباط بين كل بند وبند آخر (الارتباطات البينية). كما يتوقف أيضا على ارتباط كل بند بالدرجة الكلية للاختبار. ويتضح من هذا أن تماسك الاختبار أو تناسق بذاته بدل ثبات درجاته هو المؤشر على الثبات، بل يمكن أن نحسب معامل الثبات من هذه العلاقة القانونية القائمة بين بنود الاختبار.

3- يعني ثبات الاختبار دلالته على الأداء الفعلي أو الأداء الحقيقي للفرد الذي يعبر عنه بالدرجة الحقيقية التي يحصل عليها الفرد في اختبار ما، بمعنى أن هناك احتمالا بوجود أخطاء في عملية القياس التي تؤثر على استقرار الاختبار، لأن درجة الفرد على الاختبار تتكون في الواقع من درجة حقيقية ودرجة خاطئة، ويعبر عنها بالمعادلة التالية :

س ك = س ح + س خ، حيث أن س ك : الدرجة الكلية التي حصل عليها الفرد و س ح هي الدرجة الحقيقية و س خ هي الدرجة الخطأ. فالفرق بين الدرجة الكلية والحقيقية هو ما يسمى بالخطأ القياس أو الخطأ المعياري للاختبار.

وهناك طرق عديدة لتحليل ثبات الاختبار وإعداد نتائج التحليل، ولكن معظمها يبدأ من منطلق واحد : وهو تطبيق صورتين من نفس الاختبار على مجموعة من الأفراد لهم نفس الخصائص العمرية والاجتماعية لمجتمع الأفراد الذي يجهز له الاختبار. فإذا كنا نريد أن نعين درجة ثبات اختبار ذكاء خاص بأطفال سن الثامنة، فإننا نطبق الصورة أ والصورة ب على مجموعة أطفال في نفس السن وفي نفس الوقت. ولكن إذا لم تتوفر غير صورة واحدة،

فإننا نقسمها إلى نصفين – وعادة ما يكون النصف من الأرقام الفردية، والنصف الآخر من الأرقام الزوجية – لكل منهما له مفتاح خاص للتصحيح. وبالتالي فإننا نحصل على درجتين لكل فرد، بدلا من درجة واحدة، وإذا كنا نريد تعيين درجة ثبات اختبار في مهارة استخدام الأصابع لتلاميذ في مدرسة مهنية، فغالبا ما سوف نطبق الاختبار مرتين على نفس المجموعة من التلاميذ، مع مراعاة وجود فاصل زمني قصير بين التطبيقين. وهذه الطريقة أيضا تمدّ الفاحص بدرجتين لكل فرد، وفي كل مثال من الأمثلة السابقة، ليس هناك سبب لأن نفترض أن جميع عوامل الصدفة سوف تؤثر على الدرجتين بنفس الطريقة، لذلك، فإنه يمكن أن نقول أنه كلما كانت درجتا كل فرد في التطبيقين أكثر تشابها، كان الاختبار أكثر تناسقا، أو دقة، أو ثباتا في قياس قدرة أو سمة شخصية غير متأثرة بالصدفة. ولكن إذا كثرت حالات الاختلاف بين الدرجتين في مجموعة من الأفراد، فإننا نكون أكثر ريبة في أن درجات جميع المفحوصين تعتمد إلى درجة كبيرة على عوامل الصدفة.

إن الطريقة المعتمدة للتعبير عن التشابه (أو الاختلاف) بين مجموعتين من الدرجات هو معامل الارتباط، ولأن عوامل الصدفة دائما موجودة إلى حد ما فلن يستطيع الأخصائي الذي يبني الاختبار الحصول على معامل ارتباط تام بين مجموعتي الدرجات، أي لن يصل معامل الارتباط أبدا إلى (1,00) الواحد الصحيح، ولكن إذا بني الاختبار بطريقة جيدة وطبق بعناية تحت الظروف المناسبة، فإنه لن يكون من المستبعد أن نحصل على معامل ثبات يصل إلى حوالي 0,90 وفي بعض الأحيان ولأغراض خاصة، يمكن استخدام اختبار معامل ثباته أقل من ذلك بكثير، والمهم أن الأخصائي الذي سوف يقوم بعملية الاستدلال عن الأفراد أو الجماعات على أساس نتائج الاختبار لابد أن يعرف إلى أي مدى هذا الاختبار ثابت حتى يستطيع أن يأخذ في اعتباره القيام بالتصحيحات اللازمة لنواحي القصور الموجودة في دقة الاختبار.

ومن المؤسف أن وصف ثبات الاختبار عن طريق ذكر معامل الارتباط بين صورتين متكافئتين للاختبار له مساوئ عديدة، وربما كان أهمها أن معامل الارتباط، في حالة ما إذا كان مدى توزيع درجات المجموعة التي يحسب الثبات على أساسها كبيرا، يكون أعلى منه

مما لو كان هذا المدى ضيقا. ولذلك، فإنه من الضروري لمن يستخدم الاختبار، في حالة تقدير ثباته لغرض ما، أن يتأكد من أن معامل الثبات قد حسب في مجموعة لها من صفات التباين ما للمجموعة التي سيطبق عليها الاختبار.

وهناك، لحسن الحظ، طريقة أخرى يمكن بها التحقق من مدى دقة أو عدم دقة درجات الاختبار، وهي طريقة حساب الخطأ المعياري للقياس. وهذه الأداة الإحصائية تعطي لنا تقديرا لمدى الاختلاف في درجات الفرد إذا طبق عليه الاختبار عددا لا متناهيا من المرات أو بمعنى آخر، إذا استطعنا أن نحصل على عدد كبير من "العينات" من درجات هذا الاختبار بأن نتخذها من "المجتمع" الكلي لدرجات هذا الاختبار. هذا المدى هو منطقة عدم الدقة الواقعة على كلا جانبي الدرجة. أما عن الأسس المنطقية والإحصائية لهذه الطريقة فهي معقدة جدا بحيث لا يمكن التعرض لها هنا، ولكن التفسير النهائي لهذا بسيط نسبيا بحيث يمكن فهمه.

إذا كان الخطأ المعياري للاختبار في الذكاء، على سبيل المثال، هو 6، 3، فإن هناك احتمال بمقدار 3/2 بأن الدرجة الحقيقية للفرد تقع في مدى 6، 4 نقطة من الدرجة التي حصل عليها فعلا. فإذا كان الفرد أ حصل على درجة 154، فإنه يمكن القول بدرجة من الثقة أن "درجته الحقيقية " – خالصة من تأثير عوامل الصدفة – تقع بين 149,4 و158,6، وهذا يبنى على حقيقة أنه في أي توزيع اعتدالي نجد أن 3/2 الحالات تقع في مدى انحراف معياري واحد من المتوسط، كما أوضحناه سابقا، فإن الخطأ المعياري هو نوع خاص من الانحراف المعياري. ولكن ما زال هناك احتمال بمقدار 3/1 بأن "درجته الحقيقية"، قد تكون أعلى أو أقل من ذلك، ونستطيع إذا شئنا، باستخدام التفكير الذي يعتمد على مفاهيم الخطأ المعياري، أن نقول أن هناك احتمالا مقداره 5 % بأن تتراوح درجة الفرد بين 145 – 163 (ذلك لأنه في التوزيع الاعتدالي يقع 95 % من الحالات في مدى انحرافين معياريين من المتوسط). ولكننا عادة لا نقوم بمثل هذه التحديدات الدقيقة، إننا نستخدم الخطأ المعياري كدليل على مقدار عدم الدقة في تفسير الدرجات واتخاذ القرارات بالنسبة للأفراد. ففي مثالنا السابق إذا كنا سوف نستخدم اختبار الذكاء هذا في اختيار

الطلاب لدراسات متفوقة، وكان القرار أن كل طالب حصل على درجة 155 أو أعلى له الحق في هذه الدراسة، فإن الفرد أ، قد حصل على 154 يبدو للوهلة الأولى كأنه غير مؤهل لها، ولكن المعلم الذي يأخذ في اعتباره المقدار العام لعدم الدقة في درجات الاختبار، سوف يرى أن الفرد أ قد يكون من المجموعة المؤهلة لهذه الدراسة، إذ أن من المحتمل جدا أن تكون درجة أ على الأقل 155، وفي هذه الحالة سوف يبحث المعلم في كثير من الدقة عن أدلة أخرى كثيرة للتحقق من حقيقة حالة الطالب أ قبل أن يقرر ما إذا كان يوضع في هذا الفصل الدراسي الخاص أم لا. وعلى ذلك، فإن فهم المعنى الحقيقي لثبات الاختبار سوف يكون له أثر واضح في استخدام الاختبارات في عملية اتخاذ القرارات. (وعندما نؤكد أن الاختبارات ليست ثابتة ثباتا تاما، فيجب أن نتذكر دائما أن الطرق الأخرى في تقييم الناس ليست ثابتة كذلك. إن مؤلف الاختبار يخبرنا عادة عن مقدار عدم الدقة في اختباره ولكن ليست لدينا، عادة، أية طريقة لمعرفة مقدار عدم الدقة في "أحكام المعلمين").

إن الشخص الذي لا يفهم معنى "ثبات" الاختبار قد يتأثر تقييمه العام للاختبار بدرجة كبيرة بما يقال عموما عن المعنى الدارج لكلمة الثبات. فكلمة "ثابت أو موثوق به"، كما تستخدم في كلامنا الدارج، تحمل معاني حسنة – فيقال رجل موثوق به (بمعنى "طيب")، أو مؤسسة موثوق بها (بمعنى "مستقيمة")، وإنتاج موثوق به (بمعنى "قيم"). فمن المحتمل أن المعلمين، والعاملين في شؤون الأفراد، والأخصائيين الكلينكيين قد يستنتجون أن الاختبار الثابت هو، في الواقع، اختبار جيد صالح لأي غرض يفكرون فيه. وهذا خطأ يقود إلى أخطاء أكثر خطورة في الحكم على الناس. فإذا كانت هناك أدلة على أن اختبارا ما دقيق في قياسه لشيء ما، فإن هذا لا يعني كثيرا، إننا نحتاج إلى معلومات أكثر عن الاختبار. وقبل أن نعرف ماذا يقيس هذا الاختبار، لا نكون محقين في استنتاج نتائج عن الناس بناء على درجاتهم في الاختبار، وكما أوضحنا فإن عملية البحث عن ماذا يقيس الاختبار بالضبط، هي عملية طويلة وشاقة، فكلمة "ثابت" لا تعني "صالح" لكل شيء. وعلى ذلك، فإن صدق الاختبار يظل أهم من "ثباته".

المعايير والدرجات المشتقة

عند تقويم الاختبارات وتفسير الدرجات، يجب أن نأخذ في اعتبارنا، بالإضافة إلى صدق الاختبار وثباته، الوحدات المختلفة التي صيغت بها هذه الدرجات. فالدرجة التي تدل على عدد الأسئلة التي أجاب عنها الفرد إجابة صحيحة لا تقول لنا شيئا كثيرا عن الفرد ما لم يكن عندنا معيار خاص للمقارنة. فإذا حصل تلميذ ما على درجة 47 في الحساب، مثلا، فإنه من المحتمل أن يكون أول سؤال يسأله والد التلميذ هو :"ما هي الدرجات الأخرى التي حصل عليها بقية التلاميذ ؟" إن الدرجات المشتقة أو المحولة التي تستخدم في الاختبارات المقننة، سواء في الاستعدادات أو التحصيل أو الشخصية، إنما هي درجات صممت خصيصا لمقارنة درجات الفرد بمعايير المجموعة، ويمكن أن نقوم بعملية اشتقاق الدرجات، أو تحويلها بطرق عديدة.

إن أبسط هذه الطرق هي أن نجهز معايير المثينية للمجموعة، ثم نحول درجة كل فرد إلى الرتبة المثينية المقابلة – ولكي نقوم بتجهيز جدول معايير مثينية لمجموعة ما، نقوم بقسمة عدد الأفراد الذين حصلوا على درجات أقل من كل درجة خام معينة، بالإضافة إلى نصف عدد الذين حصلوا على نفس هذه الدرجة، على العدد الكلي للمجموعة. وفي مثالنا السابق عن اختبار الذكاء الذي حصل فيه التلميذ أ على درجة 154، فإذا وجدنا أن 91% من المجموعة حصلوا على درجة أقل من 154، فإن الرتبة المثينية للتلميذ أ في هذه المجموعة هي 91.

إن الطريقة الأخرى الشائعة لتحويل الدرجات الخام إلى درجات متكافئة تقوم بتوضيح مركز الفرد بالنسبة للمجموعة، تعتمد على استخدام المتوسط، والانحراف المعياري كأساس للمعايير. ففي التوزيع الاعتدالي هناك علاقة ثابتة بين البعد عن المتوسط والمساحة المحصورة تحت المنحنى. وبغض النظر عن كون الانحراف المعياري لأي مجموعة خاصة من الدرجات كبيرا أو صغيرا، فإنه يمكن استخدامه كوحدة ذات معنى لتحديد المسافات على قاعدة المنحنى. ففي التوزيع الاعتدالي، إذا ابتعدنا وحدة انحراف معياري واحد عن المتوسط، فإننا نجد أنفسنا مباشرة تحت النقطة التي يتغير فيها شكل المنحنى من التحدب إلى التقعّر.

وعندما نبتعد عن المتوسط بمقدار ثلاث وحدات انحراف معياري نصل تقريبا إلى نقطة قريبة جدا من نهاية التوزيع. وعلى ذلك، فإنه يمكن، إذا شئنا، أن نعبر بوحدات الانحراف المعياري عن مقدار بعد كل درجة عن المتوسط، وبالتالي نستطيع أن نحدد بهذه الطريقة مكان كل فرد من التوزيع العام، فبدلا من أن نذكر أن الدرجة 25 قد حصل عليها فرد ما في المجموعة 6، فإننا نستطيع أن نذكر درجته على أنها – 1,0 انحراف معياري (أي تقل عن المتوسط بقمدار وحدة انحراف معياري واحدة). وبمجرد رؤية ذلك، فإن أي شخص له معرفة بمنحنى التوزيع الاعتدالي يمكنه أن يعرف بالضبط كم تقل درجة هذا الفرد عن المتوسط.

إن هذا النوع من تحويل الدرجات مكن علماء النفس من القيام بمقرنات لم يكن من الممكن القيام بها باستخدام الدرجات الخام. ففي توزيع اعتدالي آخر لدرجات اختبار ما حيث المتوسط 178 والانحراف المعياري 31، نجد أن الأرقام الموجودة على قاعدة المنحنى ممثلة للدرجات الخام مختلفة عن الأرقام المناظرة لها ولكن الدرجات المعيارية متماثلة. فالدرجة 144 في هذا الاختبار الثاني (178 – 34) هي درجة تماثل في ضعفها الدرجة 25 في الاختبار الأول، حيث أن كلاهما يساوي – 0,1 انحراف معياري.

ورغبة في التبسيط. فإننا كثيرا ما نعبر عن الدرجات المعيارية الأساسية بطرق أخرى كثيرة. فمثلا، من أجل الحصول على درجات مقسمة إلى وحدات أصغر، يمكن أن نضرب كل الأعداد × 10 بحيث يتراوح مدى الدرجات من – 30 إلى +30 بدلا من 3 إلى +3 ، ثم من أجل إزالة الدرجات السالبة، يمكن إضافة 50 بحيث تتراوح الدرجات من 20 إلى 80، مما يجعل المتوسط 50 بدلا من الصفر. ومثل هذه التعديلات لا تغير أبدا في العلاقات الأساسية بين الدرجات. فالقيم على الخطوط الرأسية في الشكل والموجودة على مختلف الخطوط الأفقية تحت المنحنى الاعتدالي سوف تكون متساوية إذا كانت المجموعات المعيارية التي وضعت هذه القيم على أساسها متشابهة.

وبالإضافة إلى الرتب المئينية والدرجات المعيارية، هناك طرق أخرى عديدة لجعل الدرجات مفهومة: مثل معايير العمل الزمني و فرق الدراسة في اختبارات التحصيل المدرسية وهي لا تحتاج إلى شرح خاص، أو نسبة الذكاء QI .

عدم التحيز(الموضوعية)

لقد ثار جدل كثير، كما أجريت بحوث كثيرة، طوال السنوات الماضية، حول موضوعية الاختبارات العقلية بالنسبة للمفحوصين الذين هم من أجناس وجنسيات مختلفة، ومن مستويات اجتماعية متباينة. وفي منتصف الستينات من هذا القرن، عندما تركزت الأضواء على الأحوال السيئة للفقراء في الولايات المتحدة الأمريكية، أصبحت هذه القضية ذات أهمية كبيرة. وبالتالي فقد أصبح كل من يستخدم الاختبارات في مهنتهم – كالمعلمين والإداريين في المدارس، ومديري الأعمال، ورؤساء الوكالات الحكومية – أصبح كل هؤلاء يهتمون بموضوعية الاختبار المستخدم بالنسبة للمفحوصين، وذلك بالإضافة إلى الصدق والثبات والتقنين المناسب.

والموضوعية مفهوم معقد، مثل غيره من المفاهيم التي شرحناها في هذا الفصل، وذلك لأن كل سمات الإنسان قد تكونت من خلال الخبرة، وبالتالي فإن الأشخاص ذوي الخلفية المتباينة سوف تختلف استجاباتهم حتما للأسئلة التي يتضمنها أي اختبار ما. وليس هناك ما يسمى بالاختبار الخالي من العوامل الحضارية. إن ما يجب أن يفعله مؤلف الاختبار أو من يستخدمه، إذن، هو أن يطبق الاختبار على مجموعات مختلفة من الأفراد الذين سيطبق عليهم هذا الاختبار في نهاية الأمر، لكي يرى ما إذا كانت بعض الجماعات تحصل على درجات أقل من غيرها، بل وكذلك لكي يرى كيف ترتبط درجاتها بالمحكات ذات العلاقة.

فعلى سبيل المثال، لو أن معظم أفراد جماعة من جماعات الأقلية يميلون إلى الحصول على درجات منخفضة في اختيار للاستعدادات صمم للتنبؤ بمدى نجاح المتدربين في برنامج تدريبي لمهنة ما ولو أن نفس الأفراد يميلون للحصول على درجات منخفضة على مقياس المحك (بحيث يكون معامل الارتباط بين درجات الاختبار ودرجات مقياس

المحك 1,59، مثلا)، فإن الاختبار يمكن اعتباره موضوعيا حتى ولو كانت نتائجه ليست من صالح هذه الجماعة. وما يجب أن ينتبه إليه دائما من يستخدم الاختبار في أغراض اختيار الأفراد، هو عدم تشابه الصدق التنبئي لهذا الاختبار في مجموعة الأقلية مع الصدق التنبئي في المجموعات الأكثر تمثيلا لطالبي العمل الذين أجريت عليهم من قبل دراسات الصدق. فلو كان معامل الارتباط بين درجات الاختبار ودرجات المحك، في مثالنا السابق، ليس 1,59 بل 1,06 فإننا سوف نستنتج أن هذا الاختبار ليس موضوعيا بالنسبة لهذه الجماعة من طالبي العمل، بغض النظر عن مدى صدقه في جماعات أخرى. والطريقة الوحيدة للتأكد من أن الاختبار الذي سوف نستخدمه لتقدير قدرة ما عند جماعة من طالبي العمل الجدد هو اختبار موضوعي بالنسبة لأفراد هذه الجماعة، هي أن نجري مجموعة من الدراسات لبحث العلاقة بين درجات الاختبار والأداء في العمل.

وعندما تكون معاملات صدق الاختبار عالية في جماعة ما "مهضومة الحق أو محرومة"، فإن الاختبار يمكن أن يكون أداة مفيدة تساعد على التصدي للتمييز والتفرقة، ويفتح الفرص أمام الأفراد. وتفيد الاختبارات في كثير من الأحيان في اكتشاف المواهب في الحالات التي قد لا يلتفت إليها لو لم تستخدم الاختبارات، وذلك بسبب التعصب الشعوري أو اللاشعوري لدى القائمين بالمقابلة ضد طالبي العمل من السود أو السمر أو من ذوي الملابس الرثة.

ونعود ونقول أنه عندما نتحدث عن موضوعية الاختبار يجب أن نتذكر أن الوسائل الأخرى للحكم على قدرات الأفراد وإمكاناتهم هي أيضا عرضة لكي تكون غير موضوعية مثل الاختبارات – وربما تكون أقل موضوعية من الاختبارات.

ومن المهم ايضا في بعض الاحيان ان نعرف علاقة أي اختبار جديد بالاختبارات الأخرى في مجموعة مختلفة من المفحوصين، بحيث يمكن، عن طريق عملية التحليل العاملي، توضيح طبيعة بناء القدرة في هذه الجماعة. وباختصار، فإن الحكم على موضوعية الاختبار يجب أن يبنى على نتائج كل جماعة على حدة، وكذلك بالنسبة لكل غرض من أغراض القياس.

مصادر خطأ ثبات الاختبار

هناك عدة مصادر للخطأ تسبب تباينا في نتائج اختبار معين يمكن حصرها في أربعة عوامل هي :

1-الحالة التي يكون عليها المفحوص:

عندما يؤدي الفرد اختبارا ما، فإن درجته على هذا الاختبار قد تتأثر بعوامل عديدة بعيدة كل البعد عن الخاصية التي صمم الاختبار لقياسها، مثل تشتت الانتباه أو تغير المزاج فيحصل تبعا لذلك على درجة أقل أو أكثر مما يجب.

2-مطبق الاختبار:

الطريقة التي يباشر بها المطبق تطبيق الاختبار قد تؤثر في الدرجة التي يحصل عليها المفحوص، فقد يكون متغير المزاج أو عدوانيا فيشعر معه المفحوص بالتهديد أو قلقا أو غير مدرب جيدا على إجراءات تطبيق الاختبارات أو يسعى إلى تجسيد بعض اتجاهاته وقيمه فيوحي للمفحوصين بإجابات معينة.. إلخ.

3-الظروف الخارجية :

تؤثر الظروف الخارجية المحيطة بعملية إجراءات التطبيق على الدرجة التي يحصل عليها المفحوص، مثل المكان الذي يجري فيه التطبيق، قد يكون مألوفا أو غريبا فيشعر فيه المفحوص بالخوف.

والموقف الاجتماعي السائد أثناء التطبيق كأن يكون منفردا أو مع جماعة فيشعر المفحوص بعدم الأمن معها أو يشعر بأنه في موقف منافسة.

4-الاختبار:

تؤثر بعض المعطيات المتعلقة بالاختبار على الدرجة التي يحصل عليها المفحوص، مثل مدى وضوح التعليمات وأسلوب صياغة البنود والمدة الزمنية اللازمة للإجابة عن كل بنود الاختبار، فكلما طالت هذه المدة ارتفع ملل المفحوص، مثل مدى وضوح التلعيمات وأسلوب صياغة البنود والمدة الزمنية اللازمة للإجابة عن كل بنود الاختبار، فكلما طالت هذه المدة ارتفع ملل المفحوص وكذلك مدى تماثل البنود في حالة الاختبار المكون من صورتين.

العناصر المكونة للثبات

هناك عدة عناصر مكونة للثبات يتصل بعضها بتمثيل وثبات وعمومية السلوك الذي تؤخذ منه عينة ويتصل بعضها بدرجة الاتفاق بين الملاحظين.

1- تمثيل وثبات وعمومية السلوك الذي تؤخذ منه العينة :

عند القيام بقياس خاصية معينة كالقلق مثلا لا يمكن قياس كل السلوك الذي يدل عليها، بل تختبر عينة منه فقط. وتعرف العينة Sample على أنها جزء من كل. أما الكل الذي تؤخذ منه العينة يسمى الأصل الكليpopulation . وعند دراسة العينة تعمم النتائج على الكل ممثله له في كل الظروف وثابتة من موقف لآخر ويمكن تعميمها من موقف لآخر.

2- تمثيل عينة السلوك للكل:

يعني التمثيل أن يكون السلوك الدال عن القلق والذي تم اختباره بأداة مناسبة ممكن الظهور دون أن يتأثر بالظروف التي تحيط بالفرد، أو ببعض خصائصه كالتعب، وتغير المزاج والوقت الذي أجري فيه الاختبار، هل أجري فرديا أو جماعيا.. إلخ. أما إذا كانت العينة من سلوك القلق تبدي حساسية في حالات معينة دون أخرى فلا يمكن الاستدلال منها على الكل لأنها غير ممثلة.

3- ثبات السلوك :

ونعني بثبات السلوك أن يقيس الخاصية بنفس الدرجة تقريبا في كل المواقف التي يصدر فيها السلوك. فإذا كانت نتائج القياس تتغير في حالتي التعب والراحة أو عندما يكون الفرد في مواقف اجتماعية دون أخرى أو تتغير نتائج القياس في نفس الموقف، فإن الثبات يكون منخفضا. مثلا إذا كان سلوك السيطرة الذي يقيسه الاختبار يظهر عند الفرد في كل المواقف التي يتفاعل معها، فإنه يمكن تعميمه. أما إذا كان هذا السلوك كما يقيسه الاختبار لا يظهر إلا في مواقف معينة دون الأخرى، فإنه لا يمكن تعميمه إلى جميع المواقف لأنه ليس ممثلا لكل خاصية السيطرة وليس ثابتا من موقف لآخر. هناك العديد من إجراءات القياس وكل إجراء أعطى درجته الخاصة ولا يمكن تحديد أي درجة هي الحقيقية والأقرب تمثيلا لسمة السيطرة المفترضة عند الفرد. وكلما كانت الدرجات أكثر تنوعا قلت الثقة في الاختبار المستخدم لأن الفرضية التي يقوم عليها الثبات هي أن السيطرة مثلا أو أية خاصية أخرى هي هي، بغض النظر عن الحالات والظروف المحيطة بالفرد، رغم أن الواقع يبين أن بعض الخصائص يظهر فقط في مجال محدد من مواقف المثير أو في سياقات خاصة.

4- درجة الاتفاق بين الملاحظين:

والمقصود به الاتفاق الذي يحدث بين ملاحظي عينة السلوك من حيث الدقة في الوصف والتفسير. فإذا اتفق عدد من الملاحظين على عينة معينة من سلوك الفرد على أنه عدواني مثلا فهذه العينة من السلوك التي تمت ملاحظتها ثابتة في قياس خاصية العدوانية. أما إذا كان سلوك شخص ما يوصف أو يفسر بشكل متنوع ، فلا يعتبر ثابتا، لأنه لا يمكن عندئذ أي الأوصاف أو التفسيرات هو الصحيح، ويمكن رد هذا النقص في اتفاق الملاحظين إلى عدة أسباب، فقد تتم ملاحظة سلوك منتقى أو مظاهر مختلفة لخاصية سلوكية معقدة أو يتم الاعتماد على الذاكرة أو يلاحظ السلوك في مواقف مختلفة فعند الحكم على العدوان من استجابة شخص، قد يرى البعض الملاحظين في العبارة الغاضبة حالة دفاع لدى فرد يشعر بالتهديد ؛ بينما يرى آخرون في نفس العبارة دليلا على عدوان لا مبرر له.

أنواع الثبات

1-ثبات الاختبار

يشير ثبات الاختبار إلى الاتساق والدقة وإمكان استخراج نفس النتائج بعد إجراءات التطبيق لأكثر من مرة (مقدم عبد الحفيظ، 1993: 152)؛ أو هو ببساطة "مدى اتساق الدرجات عند تكرار التجربة". بينما يخبرنا الصدق عن العلاقة بين نتائج الاختبار وجوانب "خارجية" مرتبطة به، فإن الثبات يخبرنا عن علاقات "داخل الاختبار"، إذ يبين إلا أي مدى تتحرر الدرجات من تأثير الصدفة، ومن ثم يجيب عن السؤال: إلى أي مدى يمكن أن تكون الدرجة المستخرجة محل ثقة ؟ وتبعا لنظرية الاختبارات فإن الدرجة الواحدة على الاختبار تعد الدرجة الحقيقية للفرد مضاف إليها نوع من الخطأ.

فلو طبقنا استخبارا للعصبية على 200 شخص، فسوف تكون لهم درجات تختلف عن بعضها البعض، فلا يمكن أن يحصلوا جميعا على نفس الدرجة، وهذا ما نسميه بالفروق الفردية أو التباين أو التباين الكلي. وهناك نوعان من التباين :

1.1-التباين الحقيقي: هو التفاوت بين درجات الأفراد والذي يرجع إلى السمات الحقيقية أو الخصائص المستقرة لديهم أو هو التباين الراجع إلى الدرجة الحقيقية للعصبية في هذا المثال.

2.1-التباين الخطأ : وهو التفاوت بين درجات الأفراد والذي يرجع إلى كل ما عدا التباين الحقيقي، مثل عوامل الصدفة أو تجمع عدد من الظروف المحيطة بموقف الاختبار التي لم يفطن إليها الباحث، وتعد أغلبها مؤقتة تميل غالبا إلى عدم التكرار إذا ما تكرر تطبيق الاختبار وأمثلتها كثيرة يصعب حصرها كذلك ضبطها والتحكم فيها.

2-ثبات القائم بالتطبيق

ويشير إلى مدى استقرار النتائج رغم اختلاف القائمين بالتطبيق لأن خصائص القائم بالتطبيق وطريقته في إلقاء التعليمات وقدرته على ضبط موقف الاختبار وغير ذلك من المتغيرات قد تتبّه لدى المفحوصين دوافع شتى كالتعاون الصادق أو دوافع التزييف أو الإهمال أو الرغبة في أن يسبب السرور للمجرب أو إحباطه وإغاظته كل ذلك بتأثير من شخصية المجرب وسلوكه إبّان موقف القياس.

3-ثبات المصحح

ويسمى الثبات بين المصححين، ويشير إلى أي مدى تتغير النتائج في الدقة والثقة، إذا ما تغير القائم بالتقدير والتصحيح. ويعد هذا النوع من الثبات مشكلة في الاختبارات التي تترك جانبا من التقدير للمقدر أم المصحح، فيخشى عندئذ أن يصبح ذاتيا. ومثال ذلك الطرق الإسقاطية وبعض اختبارات القدرات الإبداعية وعدد قليل من اختبارات الذكاء كاختبار رسم الرجل الإسقاطي وبعض بنود اختبار ستانفورد – بيني وفي تشخيص الأمراض النفسية عندما يقوم أكثر من طبيب نفسي بتشخيص عدد من المرضى ويشير معدل الاتفاق إلى درجة ثبات ما بين المشخصين أو المقدّرين.

4-ثبات نظام التصحيح

يبين إلى أي مدى تتغير نتيجة فرد أو مجموعة من الأفراد طبق عليهم الاختبار مرة واحدة وصححت الاستجابات بأكثر من طريقة أو نظام، وعدم ثبات نظام التصحيح واحد من أهم أسباب انخفاض ثبات الطرق الإسقاطية ومن أبرز جوانب النقص فيها.

لا شك أن أدوات القياس والتقويم تعتمد في دقتها على مدى ثبات نتائجها وتعتمد على مدى تحقيقها للأهداف التي وضعت لها على صدقها ومقياس الثبات يعتمد في جوهره على حساب معامل الارتباط بين الاختبار ونفسه الذي يسمى في هذه الحالة بمعامل الثبات.

وتتلخص أهم الطرق العلمية لقياس الثبات في ما يلي :

1-إعادة التطبيق.

2- طريقة التجزيئية النصفية.

3- طريقة الصور المتكافئة.

4- طريقة التباين التي تعتمد على حساب الثبات من الفروق الفردية التي يسفر عنها الاختبار، وتسمى هذه الطريقة تحليل التباين مستعينا بمعادلة كل من : كيودر وريشاردستون سواء الصيغة KR20 أو KR21 أو معادلة ألفا (α) لكرونباك.

العوامل المؤثرة على الثبات

هناك العديد من العوامل التي تؤثر في ثبات الاختبار بعضها يعود إلى :

1-الفرد المفحوص من حيث قدرته على أداء المهارات التي يقيسها الاختبار وطريقته في الأداء، وفهمه لتعليمات الاختبار وعوامل الإجهاد والتعب والملل والتوتر والانفعالي والذاكرة وغيرها.

2-الاختبار من حيث صياغة بنوده، وتعليماته وطريقة الإجراء.

3-تباين درجات المجموعة.

4-طول الاختبار وهو من أهم العوامل المؤثرة على ثباته، المقصود بطول الاختبار عدد فقراته التي تقيس السمة التي يقيسها الاختبار، والعلاقة بين عدد فقرات الاختبار وثباته

علاقة طردية، بعنى أنه إذا زاد عدد فقرات الاختبار زاد معامل الثبات.(فخري رشيد خضر، 2003: 83)

و من خلال ما سبق نحوصل في الجدول التالي أهم أنواع الصدق و تعريفها:

نوع الثبات	تعريفه	طريقة قياسه احصائيا
ثبات اعادة الاختبار	مؤشر على درجة الاتساق او الترابط بين درجات الاختبار عبر فترة زمنية ما	–معامل ارتباط بيرسون –الارتباط الداخلي
ثبات الصور المتكافئة	مؤشر على اتساق درجات الاختبار عبر الزمن لا يكون عرضة لانتقال أثر التدريب (لا تتم إعادة الاختبار نفسه)	–معامل ارتباط بيرسون
ثبات التجزئة النصفية	مؤشر على الاتساق الداخلي للاختبار (هل يبدو ان الفقرات تقيس المتغير او المفهوم نفسه؟) يجزء الى فقرات فردية و أخرى زوجية	–معامل ارتباط بيرسون
ثبات الاتساق الداخلي	المؤشر المفضل للاتساق الداخلي، يتم من خلاله حساب معدل جميع الارتباطات الممكنة في التجزئة النصفية.	–ألفا كرونباخ –كودر ريتشاردسون
الثبات بين المقيمين او المحكمين	درجة الاتفاق بين اثنين او اكثر من المقيمين او المحكمين حول مستوى وجولا/اوغياب السمة في خاصية او تشخيص ما.	–معامل ارتباط بيرسون –الارتباط الداخلي –ألفا

(تيموثي ترول،2007: 299

الدرس السابع: طرق حساب الصدق

يعد أحد المؤشرات التي تدل على مصداقية وجود الاختبار، فالاختبار الجيد هو الذي يقيس السمة التي يهدف إلى قياسها، وتختلف الاختبارات في درجات صدقها تبعا لاقترابها أو ابتعادها وتقدير تلك السمة التي تهدف إلى قياسها.

طرق حساب الصدق :

يوجد العديد من الطرق لحساب معامل الصدق، وتتوقف هذه الطرق على هدف الاختبار والامكانيات المتاحة من حيث وجود ميزان خارجي ذا مصداقية كبيرة، وهذه الطرق هي :

1.صدق المقارنة الطرفية (التمييزي) : تستخدم هذه الطريقة للمقارنة بين 27 % من درجات المتفوقين و 27 % من درجات المتأخرين (و تم اختبار 27% كتقسيم الافراد) حتى يسهل التعامل مع جدول فلاجان التي تستخدم هذا التقسيم.

و الاختبار الصادق هو الذي يستطيع أن يميز بين المتفوقين والمتأخرين تحصيليا ويستخدم اختبار (t) للمقارنة بين متوسط أداء المجموعتين.

ومن أبسط الطرق التي تستخدم لتحقيق هذه الفكرة مقارنة متوسطات درجات الأقوياء بمتوسطات درجات الضعفاء ثم حساب دلالة الفروق بين هذه المتوسطات، وعندما تصبح لتلك الفروق دلالة إحصائية واضحة نستطيع أن نقرر أن الاختبار يميز بين الأقوياء والضعفاء في الميزان، وبذلك نطمئن إلى صدقه، وعندما لا تصبح لتلك الفروق دلالة إحصائية فإننا لا نستطيع الاطمئنان إلى صدق مثل هذا الاختبار أي أن هذه الطريقة تدل على صدق الاختبار ولا تدل بطريقة عددية أكيدة على مقدار هذا الصدق.

ويحسب من خلال اختبار " ت " لعيّنتين مستقلتين ومتجانستين أو متكافئتين وه و كما يلي:

$$t = \frac{\bar{x}1 - \bar{x}2}{\sqrt{\left(\frac{(n1-1)s^2 1 + (n2-1)s^2 2}{n1+n2-2}\right)\left(\frac{1}{n1}+\frac{1}{n2}\right)}}$$

حيث:s يمثل التباين ويمثل المتوسط الحسابي $\bar{x}$

أما درجة الحرية فتحسب على هذا النحو:

$$df = n1 + n2 - 2$$

ويمكن الإشارة إلى أنّ المقارنة الطرفية تتمّ بين (% 27) أعلى و (% 27) الأدنى ، كما تتم المقارنة بأسلوبين:

- **النوع الأول:** المقارنة الطرفية لنفس المقياس وهو الأكثر اعتمادا، كونه لا يحتاج إلى محك خارجي.
- **أما النوع الثاني:** يتم التطرق إليه عند تناول الصدق المرتبط بالمحك.

المقارنة الطرفية في الاختبار فقط: مقارنة درجات الثلث الأعلى بدرجات الثلث الأدنى في الاختبار

وتتم هذه المقارنة عن طريق حساب الدلالة الإحصائية للفرق بين المتوسطين . فإذا كانت هناك دلالة إحصائية واضحة للفرق بين متوسط الثلث الأعلى ومتوسط الثلث الأدنى

يمكن القول بأن الاختبار صادق والحقيقة أن هذه الطريقة عموما هي الطريقة السهلة والأقل دقة من طريقة التحليل العاملي أو المحك الخارجي . ولكنها تعطى مؤشرا سريعا عن صدق الاختبار.

تمرين1 :

لديك نتائج اختبار كما هي ممثلة في الجدول المطلوب :احسب صدق هذا الاختبار بطريقة المقارنة الطرفية:

االفراد	الدرجات	االفراد	الدرجات
1	6	9	6
2	8	10	6
3	8	11	7
4	7	12	9
5	6	13	5
6	9	14	9
7	3	15	9
8	2	16	9

– نقوم أوّلا بترتيب درجات الأفراد ترتيبا تصاعديا أو تنازليا.

–ثم نقوم بتحديد الطرف العلوي والسفلي للمقياس وذلك بتحديد (27 % العلوي والسفلي)

حسب معادلة " كيلي " وهي كالتالي:

$$\frac{16 \times 27}{100} = 4.32 \simeq 4$$

عدد أفراد العينة =4.32 ≈ 4

الترتيب التنازلي	9	9	9	9	8	8	8	7	7	6	6	6	6	5	3	2

9-9-9-9 طرف علوي

و 6-5-3-2 طرف سفلي

نقوم باستخراج المتوسط الحسابي والتباين لكل من الطرف العلوي والسفلي للمقياس.

درجات الطرف العلوي X	درجات الطرف السفلي Y	X^2	Y^2
9	6	81	36
9	5	81	25
9	3	81	9
9	2	81	4
$\sum 36$	$\sum 16$	$\sum 324$	$\sum 74$

الطرف العلوي :

$$\overline{x} = \frac{\Sigma x}{n} = \frac{36}{4} = 9$$

$$s^2 = \frac{432-(36)^2}{4(4-1)} = \frac{0}{12} = 0$$

نقوم بحساب الطرف السفلي بالطريقة نفسها

بعدها نحسب:$df = n1 + n2 - 2 = 4 + 4 - 2 = 6$

5.49=T

بالرجوع إلى جدول "T " نلاحظ أنّ:

$2.447 = to < 5.49 = tc$

عند مستوى الدلالة $\alpha = 0.05$ و درجة حرية(6)

هذا يعني أنّ الفرق بينهما دال أي يوجد اختلاف بين طرفي المقياس (العلوي والسفلي) وبهذا

فالمقياس الذي نبحث عن صدقه قادر على التمييز فهو بذلك صادق.

تمرين2 :

افترض أن لدينا مجموعتين من الطلاب المتفوقين تحصيليا والمتأخرين تحصيليا وكانت درجاتهم كالتالي :

الناجحون : $X_1 = 4,79$ $S_1^2 = 2,4$ $n_1 = 192$

الراسبون : $X_2 = 2,40$ $S_1^2 = 1,70$ $n_2 = 48$

يمكن تطبيق معادلة (t) للمقارنة بين الفروق وبين المجموعتين.

تعتبر قيمة 9,19 المحسوبة أكبر من القيمة المعدولة وبالتالي الفرق لا يرجع إلى عامل الصدفة، فالاختبار يعتبر مميزا بين الطلاب المتفوقين والمتأخرين.

النقد :

رغم أن طريقة المجموعات المنفصلة تستخدم في تقدير الصدق إلا أنها لا تعطي قيمة لمعامل الصدق لأنها تقوم على التمييز بين الناجحين والراسبين. ولكي نحصل على قيمة عددية لمعامل الصدق فإننا نقوم بترتيب جميع الأفراد في الميزان ثم نقسم هؤلاء الأفراد إلى قسمين : ناجح، راسب، وتحسب بعد ذلك معامل الارتباط الثنائي الأصيل وهذا الارتباط يدل على القيمة العددية لمعامل الصدق.

طريقة اتخاذ القرار :

يحاول متخذوا القرار أن تتسم قراراتهم بالدقة، وكلما كان القرار دقيقا سهل تنفيذه والاستفادة منه ويمكن الاستفادة منه في تعيين موظفين جدد أو في الترقية أو في قبول الطلاب بالجامعة ويمكن تصور اتخاذ القرار كما يلي :

الاختبار	ناجح (+)	راسب (−)
ناجح (+)	A (++)	B(+ −)
راسب (−)	C(− +)	D (− −)

نسبة الناجحين = $\frac{\text{عدد الناجحين}}{\text{عدد المختارين}} = \frac{A}{A+B}$

مثال توضيحي

1.نفترض أن لدينا اختبارا في الحساب ونود أن نحسب درجة صدقة فإنه يمكن استخدام ميزان اختبار الجبر لتقدير قيم الصدق ونفترض أن الطلاب الناجحين هم الذين يحصلون على تقديرات A,B,C أما الطلاب الراسبون فهم الذين يحصلون على تقديرات D.F

وكانت النتائج كما يلي :

درجات	درجات اختبار الذكاء (الجبر)		
اختبار الحساب	ناجح	راسب	العدد الكلي
8	9	0	9
7	17	0	17
6	37	0	37
5	35	4	49
4	35	7	52
3	37	9	36
2	8	12	20
1	4	16	20
المجموع	192	48	240

2.نفترض أننا قدرنا الناجحين هم الذين حصلوا على (5) درجات في اختبار الحساب فتكون نقطة القطع كما يلي :

اختبار الحساب	ميزان الأداء	
	ناجح	راسب
ناجح	108	4
راسب	84	44

نسبة الطلاب الناجحين = $\frac{A}{A+B}$ = $\frac{108}{112}$ = 0,97

هذا يعني أن الدرجة (5) يمكن أن تعطينا درجة من الصدق بنسبة 96 % عند اتخاذ القرار ويتوقف هذا القرار على المتاح من الإمكانيات فقط بتغيير قيمة الدرجة (5) فربما تكون (6) أو (7) في ضوء الأماكن الشاغرة للحصول على الوظيفة وزيادة الدرجة تعني زيادة جودة الملتحق بالوظيفة وتلعب نقطة القطع دورا في اتخاذ القرارات بدرجة من الدقة، وبالتالي يمكن أن نؤكد على أن الصدق يساعد على اتخاذ القرار بدقة.

3.الصدق الذاتي

يعرف بأنه صدق الدرجات للاختبار عندما يكون الخطأ صفرا وبذلك تكون درجات الاختبار الحقيقية مساوية للدرجة الكلية ويمكن توضيح ذلك في المعادلة التالية :

$$X = T + E$$

حيث : X = الدرجة الكلية المشاهدة

E = خطأ القياس

T = الدرجة الحقيقية

عند E = o تكون X = T

معامل الصدق الذاتي = $\sqrt{\text{معامل الثبات}}$

يسمى هذا الصدق الذاتي بالثبات القياسي وهو الذي يحدد النهاية العظمى لمعاملات الصدق، ولهذا لا يمكن أن تتجاور معاملات الصدق معامل الصدق الظاهري.

2- صدق المحتوى Content Validity

يطلق على هذا النوع من الصدق أحيانا اسم الصدق المنطقي، أو الصدق بحكم التعريف، أي تعريف المفهوم أو المتغير محل القياس أو صدق العينة، عينة السلوك، وهذا الاسم الأخير هو أقرب إلى المعنى المقصود.

-طريقة استشارة الخبراء

يقدّر صدق المحتوى بإجراء مراجعة منظمة لمجموع البنود التي يتضمنها الاختبار من قبل الخبراء لتقدير مدى تمثيلها للخاصية التي أعدّ الاختبار لقياسها. وبما أن أي خاصية تتحدد من خلال تعريفها، فإن صدق المحتوى يصبح دالة لتعريف الخاصية موضوع القياس. والطريقة الشائعة لتقدير صدق المحتوى للاختبار، هي استشارة مجموعة من الخبراء/المحكمين الذين يكونون من ذوي الخبرة والكفاءة في المادة، أو في مضمون المادة الدراسية التي صُمّم لها الاختبار. وعلى مُعدّ الاختبار أن يقدم اختباره في استمارة تتضمن ثلاثة أعمدة. يتضمن العمود الأول الأهداف. ويتضمن العمود الثاني الأسئلة التي تقيس كل هدف. ويتضمن العمود الثالث مكانا فارغا يضع الخبراء فيه ملاحظاتهم. يقرأ الخبير بأسلوب أفقي الهدف، ثم السؤال الذي وضع لقياسه، ثم يضع ملاحظته في المكان الفارغ الخاص بذلك أمام السؤال. والملاحظات (الأحكام) التي يضعها هي يقيس أو لا يقيس، أو تكون هاتان الكلمتان مكتوبتين في العمود الثالث بأعلى الاستمارة، ويقوم الخبير بوضع علامة معينة في المكان الخاص لذلك، وأمام كل هدف والسؤال. فيضع علامة × مثلا تحت كلمة يقيس أمام السؤال الذي يرى أنه يقيس الهدف، ويضع نفس العلامة تحت كلمة لا يقيس إذا كان يرى أن السؤال لا يقيس الهدف.

3-الصدق المرتبط بالمحك Criterion – Related Validity

يطلق على هذا النوع من الصدق اسم الصدق الواقعي أو الصدق العملي أو الصدق التجريبي. ويقصد به مجموع الاجراءات التي نتمكن من خلالها من حساب الارتباط بين درجات الاختبار ودرجات محك خارجي مستقل، هو السلوك نفسه أو النشاط الذي تناوله الاختبار بالقياس. والمحك هو مقياس موضوعي مستقل عن الاختبار، نقيس به صدق الاختبار، أو هو ميزان نحدد به مدى صدق الاختبار. والهدف من الصدق المرتبط بالمحك، هو معرفة ما إذا كان من الممكن إحلال الاختبار أو الاستبيان بإجراءاته البسيطة وغير

المكلفة، محل المحك الذي قد يصعب الحصول عليه أو يكون مكلفا، أو يتطلب وقتا وجهدا، أو يتطلب الحصول عليه الاعتماد على جهات أخرى.

ومن الامثلة على الصعوبات التي تعترض الفاحص للحصول على محكات، أنه عندما يكون الهدف هو الحصول على معلومات (محكات) من أجل معرفة مستويات الاداء في التدريس لدى المعلمين، فإما أن نحصل على هذه المعلومات من تقويم المعلمين لزمالئهم، أو من تقويم المفتشين للمعلمين، أو من تقويم مديري المؤسسات التعليمية للمعلمين، وكما نلاحظ فإن "المحك" المطلوب هنا نحصل عليه من أشخاص آخرين، وهنا يمكن أن تظهر صعوبات تحول دون الحصول على محك جيد وبطريقة بسيطة، وعندما يكون هدف أحد الباحثين أيضا هو الحصول على معلومات فيزيولوجية، تتعلق بنشاط الهرمونات أو بنشاط الجهاز العصبي للاستدلال بها على مدى وجود خاصية سلوكية معينة، فإن الحصول على هذه المعلومات (المحكات) يكون صعبا أو يكون مكلفا ماليا، أو ربما يحتاج الامر إلى توفر بعض الشروط التي لا تكون متاحة دائما كالعينات أو الاجهزة التي نحصل بها على معلومات فيزيولوجية.

ولذا فإنه بدلا من أن يسعى الاخصائي في القياس إلى جهات أخرى، في كل مرة يريد قياس الخاصية، للاستعانة بها من أجل الحصول على المؤشرات الضرورية لوجود الخاصية السلوكية محل البحث، يقوم بتصميم استبيان لقياس الخاصية ويحسب صدقه محكيا مرة واحدة ثم يحله محل المحك في عمليات القياس التي يقوم بها بعد ذلك. والصدق المرتبط بالمحك من أفضل وأقدم الطرق لتقدير صدق الاختبارات والاستبيانات، ذلك أن المحكات التي يمكن أن تمثل الخاصية في الواقع الخارجي، تعتبر أفضل ميزان لتقدير مدى صدق الاختبار أو الاستبيان.

فلو تم تطبيق اختبار لقياس الاستعداد الموسيقي على مجموعة من اللذين يمارسون نشاطا موسيقيا، (النشاط الموسيقي هو المحك) وتبين أن من حصلوا على درجات مرتفعة في الاختبار هم أنفسهم من يعتبرون ممتازين في أدائهم الموسيقي، وأن من حصلوا على

درجات متوسطة في الاختبار، يعتبرون متوسطين في أدائهم الموسيقي، وأن من حصلوا على درجات منخفضة في الاختبار، يعتبرون رديئين في أدائهم الموسيقي، فإن هذه النتيجة تعتبر أحسن دليل على صدق الاختبار في قياسه الاستعداد الموسيقي، ذلك لوجود هذا الارتباط الوثيق بين درجات الاختبار ومدى النجاح في محك واقعي خارجي عملي، هو الاداء الفعلي في الموسيقى، على اعتبار أن هذا الاداء يصلح لتمثيل الاستعداد الموسيقي.

والمحكات كثيرة ومتنوعة وما سبق ما هو إلّا نماذج منها فقط وسنتناول أنواعها في الفقرة التالية.

4-أنواع المحكات

نورد فيما يلي أنواع المحكات كما لخصتها أ. أنستازي 1961 في :

- **الفروق في العمر:**

يستخدم هذا المحك عندما يكون المتغير موضع القياس من المتغيرات التي تتأثر بالنمو خاصة في مرحلة الطفولة وبداية المراهقة. فالذكاء مثال يتأثر بالزيادة في العمر، فيمكن اتخاذ التغير في العمر كمحك خارجي لتقدير صدق اختبار يقيس الذكاء، فيحسب معامل االرتباط بين التغير في العمر. والتغير في درجات اختبار الذكاء والقدرات العقلية. ومعامل الارتباط المحصل عليه هو معامل الصدق، فإذا كان دالا إحصائيا دل ذلك على صدق الاختبار، وإذا لم يكن دالا إحصائيا دل ذلك على عدم صدق الاختبار. وقد حسب بيني صدق اختباره للذكاء بهذا الاسلوب، حين اعتبر ارتفاع معامل السهولة في الإجابة على بنود الاختبار لعمر عقلي معين مع الزيادة في العمر الزمني، مؤشر على صدق الاختبار لفئة هذا العمر.

-التحصيل الدراسي :

ثبت من نتائج العديد من الدراسات أن الارتباط بين الذكاء والتحصيل الدراسي دال إحصائيا، ولذا من الممكن اتخاذ التحصيل الدراسي محكا لحساب صدق اختبار في الذكاء، فنحسب معامل الارتباط بينهما، فإذا جاء دالا إحصائيا كان ذلك مؤشرا على صدق اختبار الذكاء، كما أن مستوى التحصيل الدراسي يصلح أن يكون محكا لحساب صدق اختبار تحصيلي آخر، فلو نصمم اختبار لقياس الاستعداد الرياضي (دراسة الرياضيات) فإننا نقوم بتصميم هذا الاختبار ثم نطلب من أساتذة الرياضيات أن يضعوا تقديرات لتلاميذهم في تحصيلهم لهذه المادة. ويصنفونهم إلى مرتفعين ومتوسطين ومنخفضين في التحصيل. وتعتبر التقديرات التي يضعها الاساتذة لتلاميذهم في تحصيلهم لمادة الرياضيات، بمثابة محك خارجي، ثم نطبق الاختبار الذي وضع على نفس التلاميذ، فإذا تبين أن من حصلوا على درجات مرتفعة في الاختبار، هم الذين قدرهم أساتذتهم بالممتازين في مادة الرياضيات. وأن من حصلوا على درجات متوسطة في الاختبار، هم الذين قدرهم أساتذتهم بالمتوسطين في مادة الرياضيات. وأن من حصلوا على درجات منخفضة في الاختبار، هم الذين قدرهم أساتذتهم بالضعفاء في تحصيلهم في مادة الرياضيات، فإن هذا التوافق بين الاختبار المصمم وتقديرات الاساتذة، دليل على صدق الاختبار الذي يقيس الاستعداد لدراسة الرياضيات. ويمكن حساب العلاقة بين نتائج الاختبار والمحك بأسلوب استخراج معامل الارتباط الخطي لبرسون من الدرجات الخام.

- **الاداء في المهنة :**

يعتبر الاداء في المهنة مؤشرا جيدا لحساب صدق الاستبيان الذي صمم لقياس الاداء المهني. ولنأخذ الاداء المهني في التدريس كمثال، نقوم بتصميم استبيان لقياس أداء المعلمين في مهنة التدريس، ثم نطلب من المفتشين في مواد تعليمية مختلفة ولمرحلة تعليمية معينة، أن يضعوا تقديرات للمعلمين الذين يشرفون عليهم ويصنفوهم إلى أكفاء ومتوسطين

وضعفاء في الاداء التدريس، وتعتبر هذه التقديرات التي يضعها المفتشون للمعلمين، بمثابة محك خارجي ثم نطبق الاستبيان الذي صممناه على نفس المعلمين، فإذا تبين أن مرتفعي الاداء في التدريس حسب تقديرات المفتشين حصلوا على درجات مرتفعة في الاستبيان وأن متوسطي الاداء في التدريس حصلوا على درجات متوسطة في الاستبيان. وأن منخفضي الاداء في التدريس حصلوا على درجات منخفضة في الاستبيان، فإن هذا التوافق بين الاستبيان المصمم وتقديرات المفتشين، دليل على صدق الاستبيان الذي يقيس الاداء المهني في التدريس، أو نحسب معامل الارتباط بين درجات التقديرات التي منحها المفتشون للمعلمين والدرجات التي حصلوا عليها في الاستبيان، فإذا جاء معامل الارتباط دالا إحصائيا كان ذلك دليلا على صدق الاختبار. ونقوم بنفس الاجراءات إذا أردنا أن نصمم اختبارا لقياس الاداء عند الموظفين والعمال في وظائف ومهن مختلفة.

- الاداء في التدريب:

وهو أيضا نوع من الاداء الفعلي يستخدم كمحك لحساب صدق الاستبيانات، يصمم استبيان لقياس الاستعداد لدى الافراد في أداء التدريب على مهنة معينة (مثل مهنة التمريض أو السكرتارية أو التعليم أو الشرطة وغيرها) ويطبق الاستبيان على هؤلاء الافراد المختارين وتؤخذ درجاتهم عليه، ثم يتدرب هؤلاء الافراد فعلا على أداء تلك المهنة، وتمنح لهم درجات على هذا التدريب (محكات) ثم يحسب معامل الارتباط بين درجاتهم على الاستبيان ودرجاتهم على أداء التدريب، فإذا كان دالا إحصائيا دل ذلك على صدق الاستبيان، وإذا لم يكن دالا إحصائيا دل على عدم صدق الاستبيان.

- **الارتباط بين اختبار جديد واختبارات أخرى:**

هناك اختبارات واستبيانات ذات شهرة علمية واسعة، وذات سمعة جيدة ومعترف بها من حيث الصدق والثبات والتقنين، مثل اختبار ستانفورد – بيني للذكاء، وقائمة أرون بيك لالكتئاب، وقوائم أيزنك للشخصية وغيرها. إن هذه الاختبارات والقوائم تصلح كمحكات لحساب صدق اختبارات وقوائم أخرى، تم تصميمها حديثا لقياس نفس الخصائص، وذلك بحساب معامل الارتباط بينها، فإذا جاء دالا إحصائيا دل ذلك على صدق الاختبار، وإذا لم يكن دالا إحصائيا دل على عدم صدق الاختبار. وتدل نتائج الدراسات النفسية من ناحية أخرى على أن هناك متغيرات ترتبط فيما بينها ارتباطا دالا إحصائيا، مثل الاكتئاب واليأس، والاكتئاب والقلق، والاكتئاب وتوهم المرض، والمثابرة والطموح، والعصابية والخجل.. وغيرها. يمكن كذلك في هذه الحالة تقدير صدق قائمة وضعت لقياس اليأس بحساب معامل الارتباط بينها وبين الاكتئاب. وتقدير قائمة وضعت لقياس الخجل بحساب معامل الارتباط بينها وبين العصابية وهكذا.

- **التقديرات :**

المقصود بالتقديرات قيام شخص بتقدير خاصية عند شخص أو أشخاص آخرين. وينبغي أن يتوفر القائم بالتقدير على معرفة جيدة بالشخص الذي يقدره، وخبيرا بالخاصية التي يطلب منه تقديرها، ومتّزنا انفعاليا وكفؤا في عمله وموضوعيا في تقديراته. ومن الامثلة على ذلك تقدير المعلمين لتلاميذهم والآباء لابنائهم والاداريين لموظفيهم والضباط لجنودهم وغيرهم. ويمنح المقدرون درجات (محكات) للاشخاص الذين قاموا بتقديرهم في الخاصية موضوع التقدير، ثم يطبق عليهم الاختبار أو الاستبيان الذي يراد حساب صدقه عن طريق ارتباطه بالمحك، ثم يحسب معامل الارتباط بين درجات التقدير (المحكات) والدرجات على الاختبار أو الاستبيان، فإذا جاء معامل الارتباط دالا إحصائيا دل ذلك على صدق الاختبار، وإذا لم يكن دالا إحصائيا دل على عدم صدق الاختبار.

ومن أحسن الامثلة على الجهود الامبريقية الناجحة التي قام بها علماء القياس النفسي لحساب صدق اختبار بتقديرات معينة، ما قام به الفريد بيني وتيودور سيمون في فرنسا، عندما قدرا صدق اختبارهما للذكاء بتقديرات المعلمين لتلاميذهم في التحصيل الدراسي. في نهاية القرن التاسع عشر وبداية القرن العشرين، ظهرت حاجة عملية ملحة للتعرف على المتمدرسين الذين لا يستطيعون الاستفادة من البرامج التعليمية التي يستفيد منها عمل الاثنان على هؤلاء الاطفال في مدارس باريس. فاختارا مجموعة من البنود التي تتطلب التخيل والتذكر والفهم والحكم والتفسير، وقدماها لاطفال المدارس للاجابة عنها، ثم طلبا من المعلمين أن يضعوا تقديرات لتلاميذهم في التحصيل الدراسي، فأوضحت النتائج ما يلي : أولا ،أن الاداء على هذه البنود يتحسن مع التقدم في العمل (محك) ثانيا ، أن الاداء على البنود يرتبط جوهريا بالتقديرات المستقلة للمعلمين عن ذكاء التلميذ ودرجاته التحصيلية (المحك).(مقدم عبد الحفيظ،1993:148–149)

5–شروط اختبار محك مناسب:

لكي يتم تقدير صدق الاختبارات والاستبيانات بشكل جيد، ويتوفر على مستوى عال من الثقة، ينبغي أن تتوفر في المحك المختار الشروط التالية :

1–أن يتحدد في مؤشرات سلوكية بحيث يمكن ملاحظتها مباشرة وقياسها.

2–أن تكون الطريقة التي يقاس بها المحك واضحة وممكنة.

3–أن يكون غير مكلف ويسهل الحصول عليه.

4–أن يتيسر جمع المعلومات حوله من حيث الوقت والجهد.

5– أن يتوفر على مؤشرات الصدق والثبات.

6– أن يكون مستقلا عن الاختبار والتأثر به.

7– أن تكون العينة ممثلة للمجتمع الأصلي.

8-اختيار الاسلوب الاحصائي المناسب لحساب الصدق بين المحك والاختبار (عبد الرحمن بن سليمان الطريري، 1997: 260 – 261 .)

6-أنواع الصدق المرتبط بالمحك :

- الصدق المرتبط بالمحك تلازميا :

يقوم هذا النوع من الصدق المرتبط بالمحك على كشف العلاقة بين درجات الاختبار ودرجات المحك في نفس الوقت (متلازمين) أي أن حساب الصدق التلازمي نقوم به حين نكون بصدد قياس خاصية سلوكية قائمة يتزامن فيها حصولنا على درجات الاختبار ودرجات المحك، مثل أن نقيس التحصيل الدراسي في مادة الرياضيات باختبار تحصيلي تم إعداده لهذا الغرض، ونحصل في نفس الوقت على محك خارجي لمستويات أدائهم في مادة الرياضيات من معلميهم مثلا، ثم نقوم بحساب معامل الارتباط الملائم بين المجموعتين من الدرجات، ومثل أن يقوم أخصائي نفسي بقياس الاكتئاب بقائمة أعدها لهذا الغرض، وفي نفس الوقت يقيس نفس الخاصية لدى نفس العينة بقائمة أرون بيك، فيحصل على مجموعتين من الدرجات، ويحسب معامل الارتباط المناسب بينهما لمعرفة قيمته واتجاهه.

- **الصدق المرتبط بالمحك تنبؤيا:** ويعني مدى التنبؤ بنجاح الفرد في مهنة معينة، أو في دراسة معينة، أو في أداء معين، كمؤشر على صدق الاختبار أو الاستبيان، مثل أن يقيس الاخصائي النفسي الاستعداد لممارسة مهنة معينة لدى مجموعة من الافراد و التي حصلوا عليها في الاختبار، ثم يلتحق هؤلاء الافراد فعلا بالمهنة التي تم توظيفهم فيها، وبعد 6 أشهر أو أكثر من ذلك يطلب من المشرفين المباشرين على عمل هؤلاء الافراد، أن يزودوه بمستويات الاداء المهني لهؤلاء الافراد، فإذا تبين أن الذين حصلوا على درجات مرتفعة على اختبار الاستعداد المهني كانوا أيضا مرتفعين في الاداء المهني، وأن المتوسطين في درجتهم على اختبار الاستعداد المهني كانوا متوسطين في الاداء المهني، وأن المخفضين في اختبار الاستعداد المهني كانوا أيضا منخفضين في الاداء المهني، دل ذلك على أن اختبار

الاستعداد المهني صادق لأنه تنبأ لديهما سوف تكون عليه الخاصية محل القياس. إن الصدق المرتبط بالمحك تنبئي كمبدأ عام، يتضمن توقعات على المدى البعيد، إذ يمكن الحصول على البيانات المتعلقة بالمحك في وقت لاحق على بيانات الاختبار. وتعتبر اختبارات الاستعداد بصفة عامة أكثر من غيرها تعلقا بهذا النوع من الصدق، أين يتم اختبار الافراد وتصنيفهم وتوجيههم تربويا ومهنيا وفق استعداداتهم لذلك. إلا أن التمييز المنطقي بين الصدق التلازمي والصدق التنبؤي الذي يقوم على الفروق الزمنية بين الاختبار والمحك، ولكن على الهدف من الاختبار. فالفرق يتضح من الاختلاف بين صيغة هذين السؤالين : هل السيدة (س) وسواسية ؟ (تلازمي) هل السيدة (س) معرضة لأن تصبح وسواسية؟ (تنبؤي) . (صفوت فرج، 275:1989)

الدرس الثامن: طرق حساب الثبات Test Reliability

أولا: مفاهيم الثبات

يوجد عدة مفاهيم لمعنى الثبات أو ثبات المقياس بحيث لا يكون الاختبار والمقياس ثابتا إلا إذا تحققت هذه المفاهيم. معناه أن الاختبار يعطي تقديرات ثابتة أي أننا لو كررنا الإجراء في عملية القياس لحصلنا على نتائج متسقة عن الفرد، إذن فدرجة الفرد لاتتغير جوهريا بتكرار إجراء المقياس عليه. والثبات قد يعني الاستقرار stability بمعنى أنه لو كررت عملية قياس الفرد الواحد لأظهرت درجته شيئا من الاستقرار.

المفهوم الأول:

أن يعطي الاختبار نفس النتائج تقريبا إذا أعيد تطبيقه على نفس المجموعة من الأفراد، أو يعطي نفس النتائج على اختبار آخر موازٍ. فالثبات يعني هنا الارتباط بين الدرجات على الاختبار وبين الدرجات على اختبار آخر مواز (صورة له).

فإذا طبق الاختبار "×" على المجموعة من الأفراد، طبق اختبار آخر مواز "×" بحيث يكون صورة له، يصبح لكل ممتحن علامتان الأولى على الاختبار × والثانية على صورته ×'.

ويكون معامل الثبات = معامل الارتباط

$$r_{\times\times'}=P_{\times\times'}$$

فإذا بقيت علامة الطالب في الاختبارين متساوية وبقي ترتيبه في الاختبار الأول نفس ترتيبه في الاختبار الثاني فإن معامل الثبات = +1. أما إذا تغيرت النتيجة فأصبح الطالب الذي ترتيبه الأول على الاختبار × ترتيبه الأخير على الاختبار ×' فإن معامل الثبات = −1.

المفهوم الثاني : يعرف معامل الثبات بأنه مقدار ما تفسره النتيجة على اختبار "×" من تباين اختبار آخر مواز ويرمز له $p^2\times\times1$ فعند معرفة قيمة الدرجات على الاختبار × نستطيع أن نتنبّأ بتباين الدرجات على الاختبار المواز '× فإذا كان معامل الثبات 7.0 نستطيع نتنبأ بـ$(0.7)^2$ = 0.49 من تباين الأفراد على الاختبار المواز.

وهذا المعنى أكثر من مفهوم الارتباط لأنه يحمل معنى الارتباط + التنبؤ.

المفهوم الثالث: معامل الثبات يعرف بأنه مربع معامل الارتباط بين الدرجات المشاهدة على الاختبار وبين الدرجات الحقيقية ويكتب على الصورة $p\times\times'=p\times T^2$.

حيث p×× معامل الثبات ، $p\times T^2$ مربع معامل الارتباط بين الدرجة المشاهدة والحقيقة.

$$PxT = \sqrt{pxx'}$$

فإذا كان معامل الثبات لاختبار = 0.81 فإن معامل الارتباط بين الدرجة الحقيقية والمشاهدة

$$= \sqrt{0.81} = 0.90$$

ومن هذا المفهوم نستطيع أن نستنتج أنه كلما زاد معامل الثبات كلما زاد ارتباط × مع T أي تكون الدرجة الظاهرية أقرب للحقيقة ويكون الاختبار أكثر دقة وأكثر تعبيرا.

كما يفهم من هذا المفهوم أن الدرجة المشاهدة في الاختبار سوف ترتبط مع درجتها الحقيقية أكثر من اربتاطها بأي درجة أخرى مشاهدة على اختبار آخر مواز، أي أن $p\times y<p\times\times'$ وبما أن p×y تمثل الارتباط بين الدرجة المشاهدة على الاختبار والدرجة المشاهدة على اختبار آخر مواز وهذا ما يسمى صدق الاختبار فإن معامل الثبات بين الدرجة الحقيقية والمشاهدة اكبر من معامل الصدق.

المفهوم الرابع:

معامل الثبات = 1− مربع ارتباط بين الدرجة المشاهدة والخطأ.

وبعبارة أخرى :

معامل الثبات = 1− نسبة ما يفسره الخطأ من التباين الظاهري من الدرجة المشاهدة.

$$p\times\times' = 1-p\times E$$

حيث p××' معامل الثبات

$p\times E^2$ التباين الظاهري أو نسبة ما يفسره الخطأ من التباين الظاهري.

فإذا كانت الخطأ E = صفر فإن التباين $p\times E^2$ = صفر ويكون معامل الثبات

p××' = 1

وكلما زاد الخطأ E يزيد التباين $p\times E^2$ ويقل معامل الثبات. فالعلاقة عكسية.

المفهوم الخامس :

الثبات = $1-$ $\frac{\text{نسبة تباين الخطأ}}{\text{التباين الظاهري}}$

وبالرموز $p\times\times' = 1-\frac{\sigma_E^2}{\sigma_x^2}$

وهذه العلاقة تدل على أنه إذا زاد تباين الخطأ يقل معامل الثبات وإذا قل تباين الخطأ زاد معامل الثبات

فعندما يصبح $\sigma E^2 =$ صفر فإن $-\frac{\sigma E^2}{\sigma X^2}$ = صفر ويصبح معامل الثبات = 1

أما إذا زاد الخطأ ليصبح $\sigma = {}^2\times\sigma_E{}^2$ فإن $1 = \frac{\sigma E^2}{\sigma X^2}$

ويصبح معامل الثبات = صفرا.

المفهوم السادس : معامل الثبات = $\frac{\text{معامل الثبات الحقيقي}}{\text{تباين الأداء الظاهري}}$

وبالرموز $p\times\times' = \frac{\sigma T^2}{\sigma X^2}$

فإذا كان $1 = \frac{\sigma T^2}{\sigma X^2}$

أي أن التباين في الدرجات الظاهرية يساوي التباين على الأداء الظاهري فإن معامل الثبات = 1 وإن قل التباين على الدرجات المشاهدة يعكس التباين في الدرجات الحقيقية ولا أثر

لتباين الخطأ كذلك إذا كان معامل الثبات = 1 فإن أي اختلاف في درجات الممتحنين المشاهدة تعكس الفروق من درجاتهم الحقيقية.

كذلك فإنه إذا كان معامل الثبات للمقياس = 1 فإن التباين الظاهري = التباين الحقيقي وأن تباين الخطأ = صفرا وأن المقياس قد صمم بدون أخطاء.

إذا كان معامل الثبات > أقل من 1 أي 1 > p××' فإن الخطأ موجود في القياس ويزداد هذا الخطأ كلما قل معامل الثبات وعندما يصبح معمل الثبات = صفرا p××' = 0 فإن التباين الحقيقي= التباين الظاهري أي $\sigma=\sigma T^{2\times^2}$ = ، وهذا يعني أن جميع الدرجات المشاهدة على الاختبار ناتجة عن الخطأ ولذلك فإن الفروق بين الممتحنين على الاختبار (ممثلة في درجاتهم المشاهدة) تمثل الخطأ العشوائي أكثر مما تمثل الفروق بين درجاتهم الحقيقية. (عبد الله الصمادي و اخرون، 2004: 188-192)

ثانيا : طرق تقدير معامل الثبات

يوجد عدة طرق لتعين معامل الثبات أهمها الطرق الرئيسية الثلاث الآتية :

1-طريقة إعادة الاختبار Test/retest

2-طريقة الصور المتكافئة Paralled or alternative forms

3-طريقة الاتساق الداخلي Internal consistency

وكل من هذه الطرق الثلاث سوف تعطي تقديرا مختلفا لمعامل الثبات وعلينا أن نتذكر أن هذه الطرق تخضع للتقدير في حساب معامل الثبات، وأن القيمة الحقيقة لمعامل الثبات غير ملاحظة.

أولا: طريقة إعادة الاختبار

وهي من أبسط وأسهل الطرق المستخدمة لتعيين معامل الثبات وتتلخص في تطبيق الاختبار على مجموعة من الافراد، ثم يعاد تطبيق الاختبار على نفس المجموعة بعد فاصل زمني قصير يتراوح بين ساعة – بضعة أيام وضمن ظروف متشابهة، تصحح نتيجة كل طالب على الاختبار في الحالتين بنفس الطريقة والأسس، ثم نحسب معامل ارتباط بيرسون

(كما مر سابقا) بين درجات التلاميذ في المرة الأولى ودرجاتهم في المرة الثانية لتحصل على معامل الثبات والذي يسمى معامل الاستقرار Coefficient of stability وهناك عدة عوامل تؤثر على نتيجة هذه الطريقة.

أ-عوامل يحتفظ بها المفحوصون وتنتقل معهم Carry Over Effect وأهم هذه العوامل:

1.أثر التدريب: Practice حيث يلجأ بعض الممتحنين إلى زيادة خبراتهم بعد الامتحان الأول لتحسين نتيجتهم خاصة في الاختبارات والقدرات العقلية والمهارات اليدوية ولكن هذا التدريب يختلف من شخص لآخر، إذا أن البعض لا يهتم ولا يكترث بتحسين نتيجته، فعندما يلجأ البعض إلى زيادة خبرتهم أو تحصيلهم في الاختبار الثاني سوف يكون الارتباط بين الاختبار الأول والثاني أقل تمثيلا لمعامل الثبات.

2.انتقال أثر التعليم : Transfer of Learningحيث يكتسب الممتحنون معلومات ومعارف بعد الامتحان الأول نتيجة مناقشتهم في الاختبار وإجابتهم للأسئلة، كما قد يبدي بعض الممتحنون نصائح لبعضهم البعض لتحسين نتيجتهم في الاختبار الثاني وعندما يحصل ذلك فإن معامل الثبات المحسوب يكون أقل دقة وتقديرا.

3.المزاج Mode : قد يتحسن مزاج الممتحن في الاختبار الثاني عنه في الاختبار الأول فتكون درجتهم أفضل فيتأثر معامل الثبات ويتأثر المزاج بعدة عوامل كالظروف الجوية أو مكان أداء الاختبار أو حالته النفسية.

4.عامل التذكر Memory يتذكر الطلاب كثيرا من المعلومات عند إعادة الاختبار خاصة إذا كانت الفترة الزمنية الفاصلة قصيرة مما يؤدي إلى حصول أفراد المجموعة على نتائج أفضل في الاختبار الثاني ويؤثر على معامل الثبات.

ب- عامل الزمن: يعتبر من أهم العوامل المؤثرة على طريقة إعادة الاختبار حيث :

1. إذا كانت الفترة الزمنية قصيرة فإنه يجعل أثر انتقال المعلومات يعود إلى التذكر وإلى التدريب والمزاج.

2. أما الفترة الزمنية الطويلة فإنه يجعل النتيجة متأثرة باكتساب المعلومات الجديدة وكلما زادت الفترة الفاصلة فإنه يتدخل عامل هام وهو عامل النضج Maturation وعامل الإدراك، حيث أن الفترة الزمنية الطويلة تزيد من نضج الأفراد ويكون أدائهم على الاختبار الثاني أفضل. لذلك يجب معرفة طبيعة السمة التي يقيس الشخص ثباتها فإذا كانت تتغير بسرعة فلا يجب أن تستخدم هذه الطريقة. كما يجب أن نعمل توازن بين عوامل النضج والعوامل التي يحتفظ بها المفحوصين لتحديد الفترة الزمنية الفاصلة المناسبة. إن على أن هذه الطريقة تستخدم كثيرا لقياس السمات التي لا تتأثر بالنضج وبالاستقرار كقياس الميول والقيم والاتجاهات الذكاء والإحساس والتمييز.

ثانيا : طريقة الصور المتكافئة Paralled forms

ويتم في هذه الطريقة إعداد صورتين متكافئتين من الاختبار والتكافؤ يشمل:

أ–تساوي عدد الأسئلة أو عدد الفقرات في كلتا الصورتين.

ب–أن تتساوى درجة سهولة وصعوبة كل فقرة من الفقرات الواردة في الصورة الأولى مع مثيلتها في الصورة الثانية، أي أن السؤال الأول يتكافأ مع السؤال الأول من المجموعة الثانية من حيث السهولة والصعوبة، وبالتالي تتساوى مقدرة الفقرات على التمييز بين المفحوصين على السمة المراد قياسها.

ج–أن تتساوى المستويات المعرفية التي تقيسها كل فقرة في الصورة الأولى مع مثيلتها في الصورة الثانية وأن يتساوى المتوسط الحسابي والانحراف المعياري لكلتا الصورتين.

فإذا تم إعداد الصورتين بعد تحليل محتوى المادة تحليلا كافيا فإننا نحصل على معامل ثبات عالي؛ أما إذا لم تتوفر شروط التكافؤ بين الصورتين فإننا نحصل على معامل ثبات منخفض وهذا ما يميز هذه الطريقة إذا أنه يصعب الحصول على صورتين متكافئتين تماما مما ينتج عنه أخطاء في القياس. أما المصادر الأخرى للخطأ في القياس في هذه الطريقة فإنها ترجع إلى :

1.عوامل تنتقل من المفحوصين carry over وهي نفسها التي تم شرحها في طريقة إعادة الاختبار كالتذكر، التدريب، المزاج انتقال أثر التعليم.

2.عامل الزمن حيث إذا كان الفاصل الزمني بين تطبيق الصورتين كبيرا أدى إلى ظهور عامل النضج والإدراك.

إن معامل الثبات المحسوب في هذه الطريقة يعتبر معامل تكافؤ + معامل استقرار

stability + equevalence

على أن قيمة معامل الثبات المحسوب بطريقة الصور المتكافئة يكون أقل من قيمة معامل الثبات المحسوب بطريقة إعادة الاختبار وذلك لاستحالة عدم وجود أخطاء في القياس.

ثالثا : طريقة الاتساق الداخلي Internal-Consistency

ويعتمد مبدأ هذه الطريقة على مدى اتساق وحدات أو بنود الاختبار مع بعضها البعض أو اتساق كل فقرة وارتباطها مع الاختبار ككل بحيث تكون جميع الفقرات تدور حول السمة المراد قياسها والتي صمم من أجلها الاختبار، والتناسق بين فقرات الاختبار يتأثر بمصدرين من مصادر الخطأ، أخطاء ناتجة عن محتويات الفقرة، وأخطاء ناتجة عن عدم تجانسها ولذلك فكلما كانت الفقرات متجانسة فيما تقيسه كان التناسق بينها عاليا، إذ أكثر الطرق المستخدمة لقياس معامل الثبات في هذه الطريقة هي :

1.**طريقة التجزئة النصفية Split-half**

2.**طريقة كروبناخ ألفا** (Cronbach Alpha) ومعادلة كودر - ريشاردسون (Kuder Richardson)

أولا: الطريقة النصفية

وتستخدم هذه الطريقة عندما يتعذر إعادة تطبيق الاختبار مرتين أو يجد الباحث صعوبة في إعداد وتطبيق اختبار من صورتين متكافئتين أو فحص الطلاب مرتين متتاليتين نتيجة عدم تعاون التلاميذ في المرة الثانية أو عدم اهتمامهم بالامتحان، كما تجنب هذه الطريقة الباحث الآثار السلبية الناتجة عن إعادة الاختبار.

وتعتمد هذه الطريقة على تجزئة الاختبار المطلوب تعيين معامل ثباته إلى نصفين متكافئين وذلك بعد تطبيق الاختبار ككل على المجموعة، وعند التصحيح نقسمه إلى قسمين متساويين بحيث يحتوي القسم الأول على نصف الأسئلة والقسم الثاني على النصف الثاني، ونجمع درجات كل قسم على حدا فيصبح لكل تلميذ درجتان ونحسب معامل ارتباط بيرسون بين هاتين الدرجتين.

يتم تجزئة الاختبار إلى قسمين متكافئين بعدة طرق فبعض الباحثين يأخذ الفقرات من (1-50) والنصف الثاني (51-100) وتعني في هذه الحالة قد تختلف إجابات الطلاب على النصف الأول عنه في النصف الثاني بسبب الملل والإجهاد وضيق الوقت.

وقد يلجأ البعض الآخر من الباحثين إلى أخذ الفقرات الفردية في قسم والفقرات الزوجية في القسم الآخر، ولكن قد يختلف تباين القسم الأول في الاختبار عن تباين القسم الثاني بسبب عدم تساوي بنود الاختبار في درجة صعوبتها.

ثانيا : طريقة كروبناخ ألفاCronbach Alpha

تستخدم هذه الطريقة إذ كان تباين الدرجات على نصفي الاختبار غير متساوي أو أن هناك مؤشرات تدل على أن نصفي الاختبار غير متوازنة وعند تطبيق هذه المعادلة والحصول على معامل ثبات للاختبار فإننا قد نحصل على إحدى النتائج التالية :

أ-الحصول على معامل ثبات كلي عالي بطريقة كرونباخ ألفا يدل على أن معامل ثبات الاختبار يجب أن يكون مرتفع.

ب-إذا كان معامل ثبات ألفا منخفض فهذا مؤشر على أن ثبات الاختبار الحقيقي منخفض.

المعادلة العامة لقياس معامل الثبات باستخدام طريقة كروبناخ ألفا والتي هي مؤشر على الاتساق الداخلي لفقرات الاختبار فهي : $\propto = \frac{k}{k-1}\left(1 - \frac{\sum si^2}{st^2}\right)$

حيث : K عدد فقرات الاختبار

Si^2 تباين الأداء على الفقرة

St^2 تباين الأداء الكلي على الاختبار

وبشكل عام فإن معامل الثبات الذي نحصل عليه من تطبيق هذه المعادلات يكون كبيرا إذا كانت الفقرات مترابطة معا ومتناسبة داخليا وتقيس نفس السمة. (عبد الله الصمادي و اخرون، 2004: 193-199)

العوامل المؤثرة على الثبات

هناك العديد من العوامل التي تؤثر في ثبات الاختبار بعضها يعود إلى :

1-الفرد المفحوص من حيث قدرته على أداء المهارات التي يقيسها الاختبار وطريقته في الأداء، وفهمه لتعليمات الاختبار وعوامل الإجهاد والتعب والملل والتوتر والانفعالي والذاكرة وغيرها.

2-الاختبار من حيث صياغة بنوده، وتعليماته وطريقة الإجراء.

3-تباين درجات المجموعة.

4-طول الاختبار وهو من أهم العوامل المؤثرة على ثباته، المقصود بطول الاختبار عدد فقراته التي تقيس السمة التي يقيسها الاختبار، والعلاقة بين عدد فقرات الاختبار وثباته علاقة طردية، يعنى أنه إذا زاد عدد فقرات الاختبار زاد معامل الثبات والطريقة المستخدمة لتحديد هذه العلاقة هي معادلة سبيرمان وبراون في صيغتها العامة $pxx' = \frac{2\times r}{1+r}$

حيث p××` معامل ثبات الاختبار بعد زيادة وحداته

pyy` معامل ثبات الاختبار قبل زيادة وحداته

N هي النسبة بين عدد وحدات الاختبار بعد الزيادة إلى عدد وحداته قبل الزيادة. وهذه المعادلة توضح العلاقة بين معامل ثبات الاختبار الأصلي ومعامل ثباته بعد إضافة عدد من الفقرات المتوازنة إليه.

أما إذا كنا نعرف معامل ثبات الاختبار p××' وأردنا تقصيره بأخذ عدد من الفقرات منه بنسبة $\frac{1}{N}$ من طوله فإن معادلة سبيرمان براون تصبح

$$pyy' = \frac{N(pxx')}{1 + \left(\frac{1}{N} - 1\right)pxx'}$$

حيث `pyy معامل الثبات الناتج بعد التقصير

`p××معامل الثبات الأصلي للاختبار

مثال اختبار معامل ثباته `pyy = 0.7 فإذا زدنا فقراته 3 مرات فكم يصبح معامل ثباته الجديد .

ويمكن استخدام هذه المعادلة لمعرفة قيمة "N" أو النسبة التي يجب أن تزيد فيها عدد فقرات الاختبار أو نقللها للحصول على معامل ثبات معين حيث أن :

$$N = \frac{pxx'(1 - pyy')}{pyy'(1 - pxx')}$$

حيث `p××معامل الثبات المطلوب

`pyy معامل الثبات الحالي

أي أن = $\frac{\text{معامل الثبات المطلوب (1- معامل الثبات الحالي)}}{\text{معامل الثبات الحالي(1- معامل الثبات المطلوب)}}$

مثال (1) : اختبار معامل ثباته 0.7 وعدد فقراته 50 فقرة، فإذا أردنا زيادة فقراته ليصبح معامل ثباته 0.9 ما هو عدد الفقرات اللازمة إضافتها ؟

الحل : $\frac{N}{4} = \frac{0.90(0.7-1)\ 27}{0.70(0.9-1)\ \ 7}$ =

إذن عدد فقرات الاختبار الجديد = 50 × 4 = 200 فقرة.

مثال (2): اختبار معامل ثباته 0.9 سحبنا منه 10 فقرات لنكون منها اختبار جديد ما هو معامل الثبات لهذا الاختبار ؟ $N = \frac{\text{نسبة عدد فقرات الاختبار الجديد}}{\text{عدد فقرات الاختبار الأصلي}}$

$$N = \frac{10}{50} = 0.20$$

$$pyy = \frac{.2(.9)}{1+(.2-1).9} \qquad = \frac{18}{.28} = .64$$

معامل ثبات الاختبار المكون من 10 فقرات = 0.64

ويجب أن نلاحظ إن افتراضات سبيرمان وبراون وضعت على اعتبار إن الفقرات التي يمكن إضافتها أو سحبها من الاختبار متوازنة ومتسقة داخليا مع فقراته أما إذا زاد طول الاختبار بإهمال ودون مراعاة أن تكون الفقرات متوازنة فإن معامل الثبات قد لا يتأثر بالزيادة وقد يقل.

مثال: اختبار مكون من 10 فقرات معامل ثباته 0.6 فإذا ضاعفنا عدد الفقرات بإضافة 10 فقرات متوازية فإن معامل ثباته تصبح :$pxx` = \frac{(0.6)2}{0.6+1} = 0.70$

أما إذا أضفنا لهذا الاختبار عشرة فقرات غير متسقة ومتباينة فإن تطويل الاختبار يمثل زيادة مقدار ثابت لعلامة طالب وليس زيادة في معامل الثبات ويبقى معامل الثبات = 6.0

كما أن استخدام معادلة سبيرمان وبراون يمكن أن تعطي نتائج أقل تقديرا من معامل الثبات.

فإذا كان لدينا اختبار من 10 فقرات معامل ثباته = صفرا، وأضفنا إليه 10 فقرات موازية فإن معامل الثبات الجديد يصبح :

$$pxx` = \frac{0.0 \times 2}{0.0 + 1} = 0.0$$

أما إذا أضفنا 10 فقرات غير متوازية فإن معامل الثبات في هذه الحالة لا يساوي صفرا.

وبذا يصبح تقدير معامل الثبات بهذه المعادلة ليس دقيقا.

5-أثر تباين درجات المجموعة على معامل الثبات

يتأثر معامل الثبات بالتباين الحقيقي للمجموعة التي تقيس ثباتها، فعندما يكون التباين بين أفراد المجموعة الواحدة قليلا يقل معامل الثبات، وعندما يزيد التباين الحقيقي يزيد معامل الثبات، أي أن العلاقة بين الثبات والتباين الحقيقي علاقة طردية، إذا كان التباين العام يعود إلى تباين حقيقي وليس تباين في الخطأ.

ويمكن تطبيق المعادلة الآتية لتحديد العلاقة بين التباين ومعامل الثبات.

$\frac{\sigma 1^2}{\sigma 2^2} = \frac{1-pxx`}{1-pyy`}$ $\sigma 1^2$تباين المجموعة القليل

$\sigma 2^2$ تباين المجموعة الأكثر

p××` معامل ثبات الاختبار للمجموعة الكبيرة التباين.

pyy` معامل ثبات الاختبار للمجموعة الصغيرة التباين.

مثال: اختبار معامل ثباته 0.70 عندما طبق على مجموعة س تباينها = 16 فكم يكون معامل ثباته لو طبق على مجموعة تباينها 25 ؟

الحل : $\frac{1-pxx`}{1-p\ 0.70} = \frac{16}{20}$

p××` =1.25 =(1−)

$= .76pxx` = \frac{3.8}{5}$

أي أن معامل ثبات الاختبار على المجموعة الجديدة التي تباينها كبير ازداد وأصبح يساوي 0.76 (عبد الله الصمادي و آخرون، 2004: 188−203)

المراجع:

- أمين علي محمد سليمان، (2010)، القياس و التقويم في العلوم الإنسانية: أسسه و أدواته و تطبيقاته، الطبعة الأولى، مصر: دار الكتاب الحديث.
- بدر محمد الانصاري، (2009)، قياس الشخصية، مصر: دار الكتاب الحديث.
- بشير معمرية، (2007)، القياس النفسي و تصميم أدواته للطلاب و الباحثين، الطبعة الثانية، الجزائر: منشورات الخبر.
- تيموثي ترول،(2007)،علم النفس لاكلينيكي، ترجمة فوزي طعيمة داود و حنان لطفي زين الدين، عمان: دار الشروق.
- حسن مصطفى عبد المعطي، (2009)، المقاييس النفسية المقننة، الطبعة الأولى، مصر: مكتبة زهراء الشرق.
- سامي محمد ملحم، (2015)، القياس و التقويم في التربية و علم النفس، الطبعة السابعة، عمان : دار المسيرة.
- سمارة، محمد عبد القادر إبراهيم، عصام النمر، (2000)، مبادئ القياس و التقويم في التربية، الطبعة الثانية،
- شين لوبيز، ك.ر. سنايدر، (2013)، القياس في علم النفس الإيجابي: نماذج و مقاييس، ترجمة يوسف الاعسر و آخرون، الطبعة الأولى، مصر: المركز القومي للترجمة.
- صفوت فرج، (1989)، القياس النفسي، مصر: مكتبة الانجلو المصرية.
- عبد الرحمن بن سليمان الطريري،(2014)، القياس لنفسي و التربوي، نظريته، أسسه و تطبيقاته، الرياض: مكتبة الرشد.
- عبد الله الصمادي، ماهر الدرابيع، (2004)، القياس و التقويم النفسي و التربوي بين النظرية و التطبيق، الطبعة الأولى، عمان: دار وائل للنشر.

- عبد الهادي السيد عبده، فاروق السيد عثمان، (2006)، القياس و الاختبارات النفسية (أسس و أدوات)، الطبعة الأولى، القاهرة: دار الفكر العربي.
- فخري رشيد خضر، (2003)، الاختبارات و المقاييس في التربية و علم النفس، الامارات العربية المتحدة: دار القلم.
- محمود أحمد عمر، حصه عبد الرحمن فخرو، تركي السبيعي، آمنه عبد الله تركي، (2010)، القياس النفسي و التربوي، الطبعة الأولى، عمان: دار المسيرة.
- محمود بوسنة، (2007)، علم النفس القياسي: المبادئ الأساسية، الجزائر: ديوان المطبوعات الجامعية.
- مقدم عبد الحفيظ، (1993)، الإحصاء و القياس النفسي و التربوي، الجزائر: ديوان المطبوعات الجامعية.
- موريس انجرس، (2006)، منهجية البحث العلمي في العلوم الإنسانية ترجمة بوزيد صحراوي و آخرون، الطبعة الثانية، الجزائر: دار القصبة للنشر.
- Dany Laveault. Jacque Grégoire , (2005), introduction aux théories des tests en psychologie et en sciences de l'education, 2éme édition, Bruxelles :édition de Boeck université.

الملحق الأول: نماذج من أدوات التقدير

التعريف بأدوات التقدير :

هناك نوعان من أدوات التقدير، هما: قوائم التقدير وسلالم التقدير. وأدوات التقدير ليست أدوات قياس، وإنما هي أدوات لتسجيل الملاحظات عن جوانب معينة من السلوك، إما على شكل موجود أو غير موجود، كما هو الحال في قوائم التقدير، أو أن المقدر يسجل انطباعه عن مدى وجود ما يقدره بشكل كيفي، كما هو الحال في سلالم التقدير، وهذا يتم بطرق مختلفة مثل: قليل، متوسط، كثير، أو ضعيف، متوسط، جيد، جيد جدا، ممتاز أو أوافق بشدة، أوافق، محايد، أعارض، أعارض بشدة، وذلك عندما يقدّر اتجاه شخص إزاء قضية ما كالتدخين مثلا وفي هذه الحالة لا يقوم المقدِّر بوضع علامات التقدير على السلم، وإنما الشخص المفحوص نفسه هو الذي يقوم بذلك، إذا كان يقرأ ويكتب، ويسمّي هذا بالتقدير الذاتي Self Rating.

وأدوات التقدير رغم أنها أقدم من أدوات القياس، إلاّ أنها ما زالت تستعمل إلى اليوم بجانب أدوات القياس. فالناس كانوا يقدرون وما زالوا درجة حرارة المريض بوضع اليد على جبهته، رغم اختراع مقياس الحرارة الطبّي وما زالوا يقدرون الذكاء على أساس الانطباع الشخصي، كما يفعل المعلمون عندما يقدرون ذكاء تلميذ، إلى جانب قياسه باختبارات صممت لهذا الغرض. وما زالوا يقدرون سمات الشخصية بالإضافة إلى قياسها باستِبيانات الشخصية.

وأدوات التقدير نوعان، وهما : قوائم التقدير وسلالم التقدير، نتناولهما فيما يلي مع ذكر أمثلة عنهما :

1-قوائم التقدير :

يطلق البعض على قوائم التقدير Rating Cheeclists أسماء أخرى، مثل : قوائم التدقيق أو المراجعة أو الجرد، لأنها تستعمل لهذه الأغراض، إلا أن الاسم الأول هو الشائع، وقوائم التقدير هي قوائم تحتوي على كلمات أو عبارات أو جمل، تصف سمة سلوكية ما أو أداء معينا أو إنتاجا أو طريقة وغيرها.

وكما سبق أن بيّنّا، فإن قوائم التقدير تستعمل لتسجيل الملاحظات لتتخذ كأساس لتقويم سمة سلوكية أو أداء أو إنتاج... إلخ. وفيما يلي بعض النماذج من قوائم التقدير :

1-قائمة تقدير نظافة طفل الروضة :

تعليمات : ضع علامة X أمام البنود التالية إذا كان مضمونها ينطبق على نظافة الطفل.

1. يدان نظيفتان، (.....)
2. أظافر مقلمة، (.....)
3. وجه مغسول، (.....)
4. شعر ممشوط، (.....)
5. عينان نظيفتان،(.....)
6. أسنان نظيفة، (.....)

وهكذا توضع قائمة تقدير مثلها وينبغي أن تشمل كل أو معظم جوانب النظافة عند الطفل.

2- قائمة تقدير روس ل. موني R.L Mooney لمشكلات الشباب :

وهي من أشهر قوائم التقدير، تستعمل لدراسة مشكلات الطلبة في مختلف المراحل التعليمية موزعة على الميادين التالية: الصحة، المال، الحياة الاجتماعية، العلاقات الإنسانية، الدين، الدراسة، الجنس، المنهاج وطرق التدريس، المستقبل التعليمي والمهني، الأسرة وتستعمل هذه القائمة من قبل الطالب نفسه وليس من قبل المقدِّر، فيعطينا الطالب من خلالها تقديرا عن مشاكله الذاتية. وفيما يلي تعليمات استعمالها وأمثلة عن بنودها.

تعليمات :

الخطوة الأولى: اقرأ القائمة بتأنٍّ، وإذا وجدت عبارة تتحدث عن مشكلة تضايقك فضع تحتها خطًا.

الخطوة الثانية: أعد النظر في المشكلات التي وضعت تحتها خطأ، وضع دوائر حول أرقام المشكلات التي تضايقك أكثر من غيرها.

1. أشعر بتعب في معظم الأوقات.
2. لا أحصل على نوم كافٍ.
3. لا أشعر بارتياح عندما أكون مع جماعةٍ.
4. أشعر بالنقص.
5. الامتحانات غير عادلة.
6. لا أصدقاء حميمين لي في الجامعة.

3- قائمة تقدير الإنشاء :

قلنا سابقا أن المفحوص هو الذي يسجل تقديرا لذاته إذا كان يعرف القراءة والكتابة، وإذا كان بالإمكان أن يسجل التقديرات بصورة موضوعية، وإلاّ تسجل التقديرات من قبل المقدِّر كما هو الحال في هذه القائمة لأن البنود الواردة فيها الهدف منها هو تقدير إنجاز قام به المتعلم، ولذا فإن المعلم هو الذي يتولى تسجيل التقديرات.

تعليمات :

ضع علامة X أمام البنود التالية، إذا كانت تشير إلى أخطاء ارتكبها المتعلم أثناء إنجازه لموضوع في الإنشاء الكتابي.

1. أخطاء إملائية (.....)
2. أخطاء في الإعراب (.....)
3. مفردات عامية (.....)
4. لا يعرف كيف يبدأ الفقرات ويُنهيها (.....)
5. خط سيء وغير واضح (.....)
6. استعمال كلمات في غير معناها (.....)

4- قائمة كونرز لتقدير سلوك الطفل (تقدير المعلم):

إن البنود الواردة في هذه القائمة لا يستطيع الطفل قراءتها وفهمها، فالمعلم هو الذي يتولى تسجيل التقديرات.

تعليمات : فيما يلي قائمة من البنود تصف الأنماط السلوكية التي يكون عليها الطفل في حجرة الدراسة أو خارجها، والمطلوب منك قراءتها ووضع علامة X أمام البنود التي تصف سلوكا يوجد لدى المتعلم.

1. يكون عصبيا لأتفه الأسباب (.....)
2. يحدث انزعاجا وضوضاء (.....)
3. يضايق زملائه (.....)
4. يسرق الأشياء (.....)
5. ينعزل عن زملائه (.....)
6. من السهل.....(.....)
7. غير متعاون (.....)

قدمنا فيما سبق نماذج عن قوائم التقدير في مجالات سلوكية مختلفة، وبيّنا فيها أن المفحوص (المقدَّر) في بعض الأحيان، والفاحص (المقدِّر) في أحيان أخرى، يسجل التقدير بطريقة معينة إذا كان السلوك أو الوضع أو المهارة أو الخطأ المقصود بالتقدير، موجودا أم ليس موجودا. وتستخدم قوائم التقدير لأهداف ومجالات مختلفة لتقدير مدى نظافة طفل الروضة، أو مدى وجود مشكلات لدى الطلبة أو مدى مهارة المتعلم في إنجاز تعبير كتابي، أو مدى اتصاف تلميذ بسمات السلوك اللاسوي... وهكذا. والآن ننتقل إلى الأسلوب الثاني وهو سلالم التقدير.

2- سلالم التقدير :

إذا كانت قوائم التقدير تستعمل لمعرفة وجود أو عدم وجود سلوك ما أو خاصية ما، فإن سلالم التقدير Rating Scales تستعمل لتُبيّن على أي درجة من السُّلّم تقع الخاصية التي يجري تقديرها، أي أنّ السُّلّم يمتاز عن القائمة بأنه يطلب من المقدر أن يحدد مقدار وجود الخاصية. لكن الأدائين (القائمة والسلم) يشتركان في أنهما يستعملان لتسجيل

الملاحظات حول سلوك الأفراد في مجالات الحياة المختلفة، قدراتهم واستعداداتهم وإنجازاتهم وسماتهم وطرقهم في الأداء.. إلخ.

ويمكن تحويل قائمة تقدير إلى سلم تقدير، وذلك بإضافة سلم ومدرج على خط يمثل أحد طرفيه انعدام وجود الخاصية التي يجري تقديرها، أو وجود مقدار ضئيل منها، ويمثل الطرف الآخر تمام أو كمال وجودها، وما بين الطرفين يمثل درجات متفاوتة من وجودها، كلما اتجهنا من الطرف الصفري إلى طرف التمام والكمال، ويتبع المقدرون استعمال سلالم التقدير طرقا مختلفة هي أنواع من سلالم التقدير، نتناولها فيما يلي :

1.سُلّم التقدير العددي :

يستعمل هذا الأسلوب لتقدير مقدار وجود خاصية ما لدى جماعة من الأفراد في نفس الوقت، قد يكونون عمالا أو تلاميذ أو معلمين او مضطربين نفسيا.. إلخ.

تصمم خانتان على ورقة، يتم وضع قائمة بأسماء الأفراد المعنيين بتقدير خاصية سلوكية لديهم في الخانة التي على يمين الورقة، ويكتب في أعلاها عبارة أسماء المفحوصين ويكتب في أعلى الخانة اليسرى اسم الخاصية السلوكية المراد معرفة مقدار وجودها لدى كل مفحوص. ويكتب أمام اسم كل مفحوص مدرج من سلسلة أرقام تمتد من 1 إلى 9. وعند الاستعمال يبيّن المقدِّر مقدار وجود الخاصية السلوكية لدى كل مفحوص بوضع دائرة حول رقم من سلسلة الأرقام الموجودة أمام اسمه. والجدول رقم 1 يبيّن سلم التقدير العددي.

تعليمات : ضع دائرة حول الرقم الذي تعتقد أنه يمثل السمة السلوكية للفرد (الدافعية نحو العمل). علما أن مقدار وجود هذه السمة السلوكية يزداد بازدياد الرقم، بحيث يمثل الرقم 1 أسوأ دافعية ويمثل الرقم 9 أفضل دافعية، بينما يمثل الرقم 5 درجة متوسطة من الدافعية.

جدول رقم 01 يبين سلم التقدير العددي

المفحصون	الدافعية نحو العمل								
أ	1	2	3	4	5	6	7	8	9
ب	1	2	3	4	5	6	7	8	9
ج	1	2	3	4	5	6	7	8	9
د	1	2	3	4	5	6	7	8	9

يمكن للمقدرِّ أن يحتفظ بالتقديرات السابقة ويلخصها في جدول عام مع ذكر التواريخ ليقارن تقديرات كل مفحوص بنفسه عبر الزمن ليرى هل حصل تقدم لدى المفحوص أم لا. ويبين الجدول رقم 02 ذلك.

جدول رقم 02 يمثل ملخصا لسلّم التقدير العددي

	التقديرات بالأسابيع						ملاحظات
الأفراد	الأول	الثاني	الثالث	الرابع	الخامس	السادس	
			2				لا تحسن، تدرس حالته
			6				تحسن واضح
			8				ممتاز
			6				تقهقر، ابحث عن السبب

وبالنظر إلى الجدول رقم 02 نلاحظ أن المفحوص "أ" لم تتحسّن دافعيته نحو العمل طيلة الأسابيع الستة، ولذا ينبغي الاتصال به وبحث حالته لمعرفة السبب. بينما تحسن المفحوص "ب" ؛ أما المفحوص "ج" فقد كان طيلة الأسابيع الستة ممتاز. وتقهقر المفحوص "د" وينبغي البحث أيضا عن السبب.

وفي بعض الأحيان يحتاج الفاحص (المقدِّر) إلى القيام بتقدير مجموعة من الخصائص السلوكية أو كميات الإنتاج، لدى كل مفحوص على حده، فيصمم سلّما يحقق له هدفه. وفيما يلي يبين الجدول رقم 03 سلما من هذا النوع.

جدول رقم 03 يبين تقدير كل مفحوص على حده في عدة سمات

الخاصية السلوكية	ضعيف	متوسط	ممتاز
الانتباه	1 2 3	4 5 6	7 8 9
الاجتهاد	1 2 3	4 5 6	7 8 9
المثابرة	1 2 3	4 5 6	7 8 9
الدقة	1 2 3	4 5 6	7 8 9
الاعتماد على الذات	1 2 3	4 5 6	7 8 9
الهدوء في العمل	1 2 3	4 5 6	7 8 9
التعاون	1 2 3	4 5 6	7 8 9

وعيب هذا السلم أنه لم يحدد بالضبط معنى الدرجات المختلفة من الخاصية؛ فالمستوى ضعيف تقابله ثلاث درجات هي : 1، 2، 3، فأي من هذه تمثل مستوى ضعيف تمثيلا دقيقا. ونفس الملاحظة تقال للمستويين : متوسط وممتاز. ولهذا يستحسن أن تختصر سلسلة الأرقام إلى ثلاثة فقط كما يلي : ضعيف يقابله 1، متوسط يقابله 2 وممتاز يقابله 3، أو تستخدم الحروف : أ، ب، ج، تقابل كل منها ضعيف، متوسط وممتاز على التوالي. وبهذه الطريقة يتخلص المقدر من تعدد الدرجات لمستوى تقديري واحد.

2- سلم التقدير العددي الوصفي :

في هذا النوع من سلالم التقدير، توجد أرقام تمثل درجات متفاوتة من الأداء أو السمة المقدرة. كما يوجد وصف (الجمل الاستفهامية) يوجه المقدر. ومثاله السلم الآتي الذي يمثله الجدول رقم 4.

جدول رقم 04 يبين سلما لتقدير أداء الطفل لأنماط سلوكية يومية.

النمـط السلوكـي		درجات أدائه
ارتداء الملابس	هل يرتدي ملابسه بنفسه ؟	5 4 3 2 1
	هل يمكنه لبس جواربه بمفرده ؟	5 4 3 2 1
	هل يلبس حذاءه بمفرده ؟	5 4 3 2 1
النظافة	هل يحافظ على نظافة ملابسه ؟	5 4 3 2 1
	هل يغسل يديه بمفرده ؟	5 4 3 2 1
	هل يغسل وجهه بمفرده ؟	5 4 3 2 1

اللعب	هل يجيد ركوب الدراجة ؟	5 4 3 2 1
	هل يستطيع القفز على الحبل ؟	5 4 3 2 1
	هل يستطيع مسك الكرة والجري بها ؟	5 4 3 2 1
السلوك الاجتماعي	هل يتوازن مع الآخرين ؟	5 4 3 2 1
	هي يتعاون مع أترابه ؟	5 4 3 2 1
	هل ينسجم مع قواعد الانضباط ؟	5 4 3 2 1

هذا النوع من سلالم التقدير أفضل من سابقه، لأنه يتميز بوجود أوصاف وتعريفات تحدد الجوانب التي يتم تقديرها، فهناك ارتداء الملابس، النظافة، اللعب والسلوك الاجتماعي، وكذلك أسئلة التقدير التي تم وضعها على أوصاف السلوك الذي يؤدي، وهذا ما يساهم في تسهيل عملية التقدير ويزيد من فحوص اتفاق المقدرين مما يؤدَّى إلى أن تكون النتائج أكثر ثباتا وأكثر فائدة.

من مزاياه كذلك أن المقدّر يمكن أن يحلل أي سؤال من الأسئلة السابقة إلى أسئلة فرعية تدور حول سلوك واحد.

كما أن هذا النوع شائع الاستعمال في مختلف نواحي الحياة في المعمل والعيادة والمدرسة والأسرة وغيرها.

3- سلم التقدير البياني اللفظي :

في هذا النوع من سلالم التقدير تُذكَر أنماط سلوك الأداء أو الخصائص في عمود على هامش الصفحة، ثم توضع أمامه أعمدة تحمل درجات متفاوتة من وجود الأداء أو الخاصية، ويقوم المقدر بملاحظة السلوك أثناء الأداء، ثم يضع إشارات في الأعمدة، مبيِّنًا

درجة انطباق الخاصية على الفرد. وفي أحيان أخرى يقوم الفرد نفسه بوضع الإشارات في العمود المناسب، مبينا مقدار توفر الخاصية لديه، ثم يقوم على أساس درجة تكرار وجود الأداء أو الخاصية في كل عمود. وفيما يلي نموذج يمثل ذلك في الجدول رقم 5.

جدول رقم 05 يبين سلم تقدير بياني لفظي لتقدير اشتراك الفرد في مناقشة جماعية.

الأداء السلوكي أو الخاصية	درجات وجود السلوك أو السمة		
		قليلا	كثيرا
1.يطلب المشاركة في المناقشة برفع يده. 2. يحترم حق الآخرين في الكلام. 3. يطرق موضوع المناقشة مباشرة. 4.لا ينزعج إذا لم تقبل وجهة نظره. 5. يقدم اقتراحات بناءة.	... ...	... .	....

يطلق البعض على هذا السلم اسم السلم المدرج الوصفي. إن كلمة مدرج تفيد المعنى نفسه الذي تفيده كلمة بياني، إلاّ أن كلمة وصفي استعملت في غير محلها، إذ لا توجد أوصاف تحدد مستويات السلوك المختلفة، وإنما توجد ألفاظ مثل قليلا، كثيرا وهذه غامضة لا تعني الشيء نفسه لدى عدة مقدرين.

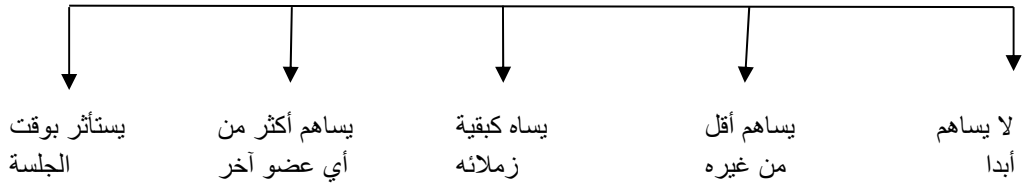
وللخروج من هذا الغموض يقوم المقدر بإعطاء رقم لكل لفظ يندرج من أقل درجة وجود إلى أكبر درجة وجود الأداء أو الخاصية، فتعطى مثلا لـ " لا" صفر وقليلا "1" وكثيرا "2" فيصبح الوصف عندئذ كميًّا.

4- سلم التقدير البياني الوصفي :

في هذا السلم يستعمل المقدر أوصافا محددة تبين وجود مقادير مختلفة من الخاصية أو الأداء أو الإنتاج وهو أكثر تحديدا ودقة من السلم السابق.

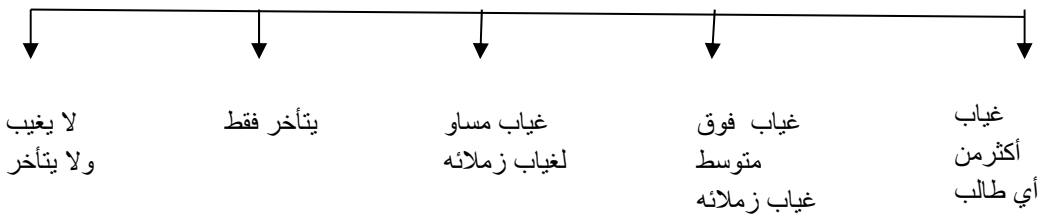
فالألفاظ مثل : " قليلا، كثيرا، جيد جدا، متوسط، ضعيف ". قد لا تعني الشيء نفسه بالنسبة للمقدرين الذين يقدرون الخاصية نفسها لدى الفرد الواحد في الوقت نفسه ؛ فقد يعني قليلا نسبة 10 % لدى أحدهم، 20 % لدى ثان وأقل من 50 % لدى ثالث. كما أن ممتاز تعني في بعض الأحيان 85 % فما فوق، وفي بعضها تعني 90 % فما فوق، ولذا يفضل بعض المقدرين أن تتحول بنود السلم اللفظي إلى بنود تحتوي على عبارات تدل على درجات مختلفة من الخاصية أو الأداء، كما يتبين في الشكلين رقم 1 و 2.

1 – ما مدى مساهمة الفرد في المناقشة ؟



شكل رقم 1 يبين مدى مساهمة الفرد في المناقشة

2 – ما مدى مواظبة الطالب على الحضور إلى المحاضرات ؟



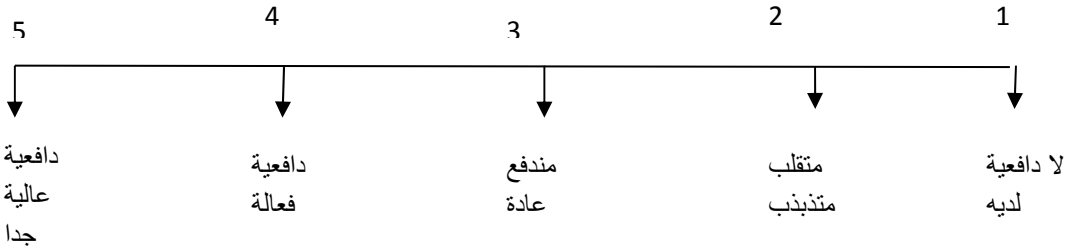
شكل رقم 2 يبين تدرج مواظبة الطالب على الحضور إلى المحاضرات

5– سلم التقدير البياني الوصفي العددي :

ميزة بنود هذا السلم تتمثل في وجود أرقام فوق درجات السلم الفرعي أو البند، كما في الشكل رقم 3.

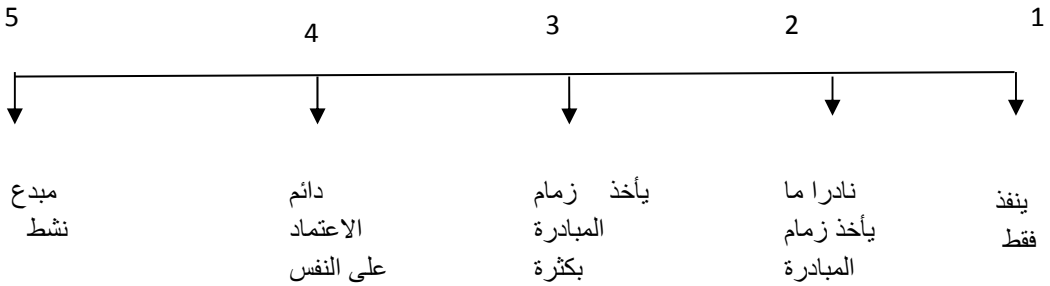
تعليمات : ضع دائرة على ما ينطبق على الطالب.

1– الدافعية :



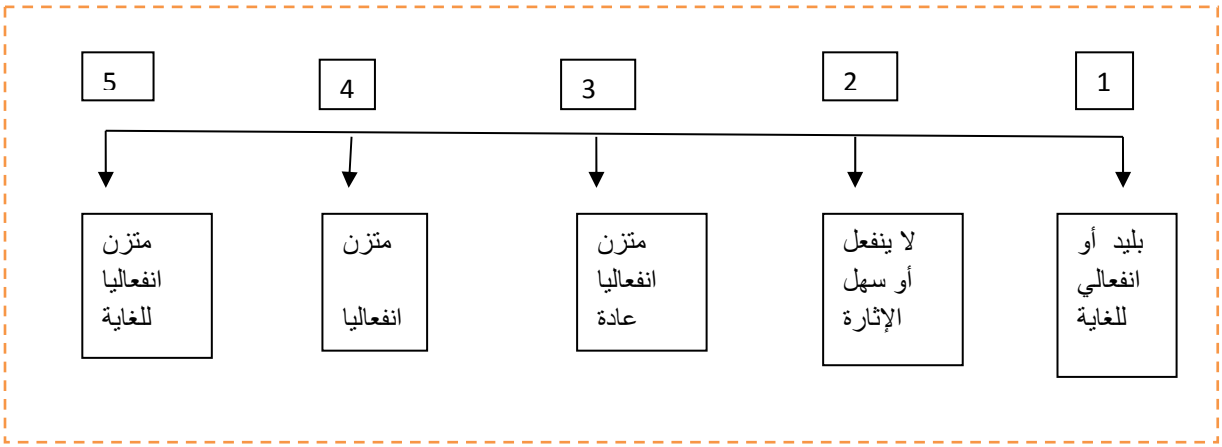
شكل رقم 3 يبين تدرّج دافعية الطالب نحو الدراسة

2- المبادأة :



شكل رقم 4 يبين تدرج مبادأة الطالب

3- الاتزان الانفعالي :



شكل رقم 5 يبين تدرج الاتزان الانفعالي

ميزة هذا السلم أنه يمكن المقدر من التقدير الكمي. كما يمكنه من تقدير الوضع بين درجتين، ويتوقف ذلك على رغبة المقدر أو مصمم السلم، إذ يمكنه أن يدرج البند درجات عديدة أو قليلة.

كيفية وضع سلالم تقدير :

عندما ينوي المقدر/الباحث وضع سلم تقدير، فإنه يفكر في ثلاثة محاور :

الأول هو : الأداء السلوكي أو الخاصية التي يسعى إلى تقديرها.

الثاني هو : سلم التقدير المناسب.

الثالث هو : كيفية استعمال السلم.

وفيما يلي خطوات يمكن أن يسترشد بها المقدر/الباحث عند وضع سلم تقدير

1-ألا يضع وألا يصمم أي سلم لتقدير نمط سلوك أو خاصية توجد له أداة قياس تقيسه بدقة إلا إذا أراد أن يضع محكا لإثبات صدق تلك الأداة، ولا توجد محكات أخرى أكثر صدقا من أدوات التقدير.

2- ألا يضع وألا يصمم سلما لتقدير أداء أو خاصية أو سلوك ويمكن ملاحظته بصورة مباشرة بمعنى أنه ينبغي أن يكون هناك دليل أو مؤشر على وجود ما يتم تقديره، لأن التقدير يستلزم وجود الخاصية أو السلوك أو الأداء، ثم يتم تقدير درجة وجوده، فمثلا من السهل على المقدر أن يقدر مدى مواظبة الطالب على حضور المحاضرات، بينما يجد من الصعب أن يقدر مدى وطنيته والسبب في ذلك هو أن المواظبة توجد أدلة ومؤشرات تدل عليها مثل عدد الغيابات وعدد مرات التأخر وغيرها. بينما في حالة الوطنية يعد المفهوم

غامضا. ومن الصعب أن يجد المقدر أدلة ومؤشرات تدل عليه يمكن ملاحظتها والاستنتاج منها أن الطالب وطني أو غير وطني.

3- على المقدر أن يحدد أبعاد السلوك أو الخاصية أو الأداء الذي يريد ملاحظته وتقديره.

4- أن يقوم المقدر بتعريف الأبعاد التي تتكون منها الخاصية أو السلوك "الأداء" بحيث لا يتناول بعدا واحدا فقط، لأن السلوك البشري ليس بعدا واحدا، بل متعدد الأبعاد. ولتقدير السلوك تقديرا شاملا وسليما، ينبغي تحليله إلى أبعاد ثم يتم تعريفها والقيام بعد ذلك بملاحظتها وتقديرها.

5- أن يحدد المقدر معاني الألفاظ الدالة على درجة وجود الخاصية أو الأداء أو السلوك. إن ألفاظا مثل : أحيانا، عادة، غالبا، بكثرة ... ينبغي تحديد معانيها بدقة، وهناك من يحددها على شكل نسب مئوية. لقد حدد براون W.F.Brown وهولتزمان W.H.Holtzman معانيها في سلّمهم لدراسة اتجاهات وعادات المطالعة لدى الطلاب على النحو التالي : نادرا من صفر إلى 15%؛ أحيانا من 16% إلى 35 %؛ بكثرة من36 % إلى 65 % ؛ عموما من 66 % إلى 85 % ؛ دائما من 86 % إلى 100 %.

وفي سلم التقدير البياني الوصفي نحدد كل نقطة أو درجة على البند بأوصاف كافية تمنع الالتباس وتسهل الاتفاق بين المقدرين.

6- أن يبسط لغة السلّم حيث تكون في المستوى العقلي والتعليمي للفرد الذي يستعمله.

7- يفضل بعض المقدرين أن يكتبوا بنود السلّم على قصاصات ورق، فيكتب كل بند على قصاصة أو صفحة، حتى لا يتأثر المقدر عند تقديره لخاصية جديدة بتقديره لخاصية سابقة. كما يقترح بعضهم أيضا أن تتم مخالفة نهايات السّلم، فتكون أفضل مستويات السلم في بند على اليمين، بينما تكون في بند آخر على اليسار. ويرى البعض أن نحدد ذلك

بطريقة عشوائية كطريقة القرعة، وليس حسب مزاج المقدر. وتقيد هذه البنود ومن ثم تجبره على التفكير في كل خاصية على حدة.

8- بالنسبة لتدريج السلم أو تحديد مستوى وجود الخاصية أو السلوك في كل نقطة أو درجة، هناك أساليب تناولناها سابقا، وتحدثنا عن السلم العددي والبياني والوصفي، وتعتبر طريقة السلم البياني الوصفي العددي أكثرها شيوعا، إذ توجد أوصاف محددة تبين كل نقطة على السلم، مما يساعد المقدر في عملية التقدير ويقلل من الاختلاف بين المقدرين.

9- تعوّد مصمِّمو سلالم التقدير على تحديد عدد الدرجات على البند أو السلم الفرعي، بما يتراوح بين ثلاث إلى سبع درجات، والشائع استعمال خمس أو سبع. وطبيعة الأداء أو الخاصية التي يتم تقديرها هي التي تحدد عدد الدرجات، فإذا كان الأمر يتطلب تقديرا عاما، فإن عدد الدرجات لا يتعدى ثلاث، أما إذا كان الهدف من التقدير يتطلب تفصيلا دقيقا، فإن عدد الدرجات يصل إلى سبع ولا يتعداها، لأنّ المفحوص أو المقدر لا يستطيع أن يميز بدقة بين أكثر من سبع درجات.

وهذه الإجراءات كلها توضح في تعليمات تطبيق سلم التقدير.

10- عند تطبيق السلم أو عندما يريد الأخصائي النفسي تقدير الخاصية أو السلوك أو الأداء عند أحد الأفراد، يمكنه أن يطلب من عدة مقدرين (معلمون أو مديرون أو آباء أو رؤساء ورشات عمل وغيرهم)، أن يقدروا سلوكه كل على انفراد ثم يلخص تقديراتهم باستخراج متوسطها. إن هذا الأسلوب يجعل التطرف أو التحيّز أو التعصّب للمقدر أو ضده، يلغي بعضه بعضا ويزيد من ثبات التقدير.

11- لا يمكن أن يقوم مقدر أو مقدرون بتقدير خاصية أو سلوك أو أداء لا يعرفونه أو لا خبرة لهم به. إن تقدير سلوك فرد مضطرب نفسيا، يقوم به الأخصائي الذي يقوم بتشخيص اضطرابه وعلاجه؛ والتلميذ يقوم المعلم بتقدير سلوكه؛ والعامل في المصنع يقوم

رئيسه أو رؤساؤه المباشرون بتقدير أدائه .. وهكذا، وإذا لم يستطع المقدر ملاحظة جانب من السلوك أو الخاصية أو الأداء فلا يقوم بتقديره.

استعمال سلالم التقدير :

كما سبق و اوضحناه، أن أدوات التقدير ومنها سلالم التقدير، تستعمل حيث لا يمكن استعمال أدوات القياس المعروفة كالاختبارات والاستخبارات الموضوعية، وهي تستعمل بشكل خاص للتقدير، ومن ثم تقويم الأداء وطرق العمل كطريقة العمل في المخبر أو في حجرة الدراسة أو في ورشة إنتاج، وفق قواعد معينة يعني ذلك وجود ثلاث عناصر في عملية القياس. (بشير معمرية، 2007: 18-30)

الملحق الثاني: جدول T test:

t Table

cum. prob	$t_{.50}$	$t_{.75}$	$t_{.80}$	$t_{.85}$	$t_{.90}$	$t_{.95}$	$t_{.975}$	$t_{.99}$	$t_{.995}$	$t_{.999}$	$t_{.9995}$
one-tail	**0.50**	**0.25**	**0.20**	**0.15**	**0.10**	**0.05**	**0.025**	**0.01**	**0.005**	**0.001**	**0.0005**
two-tails	**1.00**	**0.50**	**0.40**	**0.30**	**0.20**	**0.10**	**0.05**	**0.02**	**0.01**	**0.002**	**0.001**
df											
1	0.000	1.000	1.376	1.963	3.078	6.314	12.71	31.82	63.66	318.31	636.62
2	0.000	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	0.000	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	0.000	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	0.000	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	0.000	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	0.000	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	0.000	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	0.000	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	0.000	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	0.000	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	0.000	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	0.000	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	0.000	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	0.000	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	0.000	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	0.000	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	0.000	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	0.000	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	0.000	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	0.000	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	0.000	0.686	0.858	1.061	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	0.000	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807	3.485	3.768
24	0.000	0.685	0.857	1.059	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	0.000	0.684	0.856	1.058	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	0.000	0.684	0.856	1.058	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	0.000	0.684	0.855	1.057	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	0.000	0.683	0.855	1.056	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	0.000	0.683	0.854	1.055	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	0.000	0.683	0.854	1.055	1.310	1.697	2.042	2.457	2.750	3.385	3.646
40	0.000	0.681	0.851	1.050	1.303	1.684	2.021	2.423	2.704	3.307	3.551
60	0.000	0.679	0.848	1.045	1.296	1.671	2.000	2.390	2.660	3.232	3.460
80	0.000	0.678	0.846	1.043	1.292	1.664	1.990	2.374	2.639	3.195	3.416
100	0.000	0.677	0.845	1.042	1.290	1.660	1.984	2.364	2.626	3.174	3.390
1000	0.000	0.675	0.842	1.037	1.282	1.646	1.962	2.330	2.581	3.098	3.300
z	0.000	0.674	0.842	1.036	1.282	1.645	1.960	2.326	2.576	3.090	3.291
	0%	50%	60%	70%	80%	90%	95%	98%	99%	99.8%	99.9%
	Confidence Level										