القياس النفسي

إعداد :

د/ دعاء المدثر محمد حسان

مدرس بكلية الآداب

مراجعة :

أ.د/ محمد السيد عبد الوهاب

الفهرس

المحتويات	الصفحة
▪ **الفصل الأول**	**٣-٢١**
▪ **الفصل الثاني**	**٢٢-٦٦**
▪ **الفصل الثالث**	**٦٧-١٠٢**
▪ **الفصل الرابع**	**١٠٣-١٤٩**

تاريخ القياس النفسي

تاريخ القياس النفسي

مقدمة:

من العسير حصر العوامل والشروط الممهدة لظهور حركة القياس النفسي بصورتها الحالية، فقد اختلف الباحثون في تحديد تلك العوامل وفي الوزن النسبي لكل منها بقدر اختلافهم في تعيين نقطة زمنية محددة لنشوء هذه الحركة بيد أن الامر الذى يكاد يجمع عليه العلماء هو أن الدراسات الخاصة يكشف الفروق الفردية في زمن الاستجابة او ما يسمى بدراسات المعادلة الشخصية التي جرت خارج اطار علم النفس وقادها الفلكيون منذ القرن الثامن عشر كان مما مهد لظهور تباشير حركة الاختبارات النفسية او ما عرف لاحقاً بحركة القياس النفسي، ففي عام ١٧٩٦ حدث أن طرد أحد الفلكيين مساعده في المرصد لأنه أخطأ بمقدار ثانية واحدة في رصد الزمن الذي يستغرقه كوكب معين في مروره على لوحة التلسكوب (المرصاد الفلكي) وقد أدى هذا الحادث الى اثارة الاهتمام بدراسة الفروق في تقديرات الفلكيين وانتهت هذه الدراسة الى ان الزمن اللازم لحدوث الاستجابة نحو مثير معين (زمن الرجع) يختلف من فرد لآخر وهذا ما عرف بالمعادلة الشخصية.

العوامل التي أدت الى اتساع نطاق

يمكننا أن نتتبع تطور القياس العقلي بدراسة العوامل التي أدت الى تشجيع هذا النوع من القياس واتساع نطاقه وفيما يلي دراسة موجزة لهذه العوامل التي أدت الى اتساع نطاق القياس وتطوره:

أولاً: في النصف الأول من القرن التاسع عشر

بدأ خلال النصف الأول من القرن التاسع عشر الاهتمام بمسألة التخلف العقلي وبرزت الحاجة الى التمييز بين الافراد في القدرة العقلية، وقد قدم الطبيب " جان اسكيرول" اسهاماً كبيراً في هذا المجال وذلك عندما الح على ضرورة التمييز بين من أصيبوا باضطرابات عصبية وضعاف العقول مما مكنه ولأول مرة من التمييز والفصل بين مفهومي المرض العقلي والتخلف العقلي ولم يتوقف نشاط اسكيرول عند هذا الحد بل عمل على تصنيف مستويات التخلف العقلي، وميز بين درجتين من البله Imbecility وثلاثة درجات في العته Idiocy كما لاحظ أن القدرة اللفظية واستخدام اللغة هي المحك مباشر للقدرة العقلية العامة.

وقد سبق اسكيرول عصره بهذه الفكرة فاختبارات الذكاء المنتشرة حالياً مشبعة بالعامل اللفظي كما أن المحكات المعتمدة لتشخيص حالات الضعف العقلي هي

محكات لغوية في معظمها، وتابع اسكيرول في عمله هذا سيجان الذي انشأ اول معهد في فرنسا لتدريب ضعاف العقول عام ١٨٣٧ وتركزت جهوده على إيجاد محك للتمييز بين مستويات التخلف العقلي، ويعد سيجان بحق أحد الرواد الذين كرسوا اهتمامهم بضعاف العقول، وقد اورثنا سيجان لوحة الاشكال الشهيرة والمعروفة بلوحة سيجان التي تدخل ضمن الكثير من الاختبارات الادائية وتستعمل على نطاق واسع وتستعمل على نطاق واسع الى يومنا هذا في اختبار القدرة العقلية للصم والاميين ولاسيما ضعاف العقول.

لهذا يمكن القول بأن أهمية الاعمال التي قدمها كل من اسكيرول وسيجان تعود الى طرح مشكلة التخلف العقلي وتحديد مستوياته مما أدى الى اثارة مسألة الفروق الفردية ودعا الى الاهتمام بتلك الفروق.

ثانياً: في النصف الثاني من القرن التاسع عشر

العامل الحاسم في ظهور حركة القياس النفسي بوصفها حركة علمية تجريبية الى حيز الوجود يتبدى في الاتجاه الذي قاده فونت والذي كان بمثابة نقطة تحول حاسمة في تاريخ تطور علم النفس وأدى الى تجاوز المنهج الاستبطاني التقليدي ووضع حجر الأساس للمنهج التجريبي الكمي.

وقد أنشأ فونت أول معمل لعلم النفس التجريبي في مدينة لايبزغ في المانيا ١٨٧٩، ومع أن فونت وتلاميذه تجاهلوا مسألة الفروق الفردية وكانوا يهدفون من

وراء تجاربهم " السيكوفيزيقية" التي تركزت على دراسة الاحساسات المختلفة من سمعية وبصرية ولمسية الى الوصول أوصاف أو قوانين عامة للسلوك البشري، فقد أسهمت محاولاتهم اسهاماً كبيراً في تشجيع حركة القياس. ذلك أن تلك التجارب تطلبت استخدام مجموعة من الاختبارات التي ركزت على الظواهر الحسية البسيطة وكانت تلك الاختبارات بمثابة البنة الأولى لحركة القياس أو المؤشر الأول لظهورها كما أن شروط الضبط التجريبي وما أملته من ضرورة توحيد الظروف المحيطة بالمفحوصين كافة وهى ما تندرج ضمن إجراءات التقنين انعكست بصورة مباشرة على حركة القياس ومهدت السبيل لظهور الاختبارات المقننة التي تعد الإنجاز الأكبر من إنجازات هذه الحركة.

ثالثاً: دور العالم البيولوجي الإنجليزي " فرنسيس جالتون"

أولى فرانسيس جالتون وهو عالم البيولوجيا الإنجليزي الشهير (١٨٢٢- ١٩١١) اهتماماً كبيراً بمسألة الوراثة عند الانسان واجري دراسات موسعة حول الصفات المختلفة عند التوائم والاقارب والأشخاص الذين لا تربطهم صلة القرابة كما اشتهر بدراساته حول الموهوبين ويعد جالتون المكتشف الحقيقي لمجال الفروق الفردية ويمكن مقارنة دوره في حركة القياس بما انجزه وقدمه لعلم الفيزياء حيث تم بفضله التخلي نهائياً عن طرائق التقدير الحدسية والتخمينية التي استمرت لألاف السنين والانتقال الى طرائق علمية ترتكز على التجريب والاختبار.

وقد قدم جالتون مجموعة كبيرة من البيانات حول الفروق الفردية في العمليات النفسية المختلفة وصمم بنفسه مجموعة من الاختبارات التي لازالت تستخدم حتى الان بصورتها الأولى او المعدلة.

ونظر جالتون الى الذكاء على انه قدرة فطرية وليست مكتسبة بالتدريب والمران واكد على إمكانية قياس هذه القدرة عن طريق بعض الاعمال البسيطة، كما أكد على ان الذكاء يرتبط بالقدرة التمييز الحسي بين الاوزان المتقاربة جداً في الوزن وذلك عن طريق اليد وليس الميزان، وقد كان جالتون اول من استخدم الإحصاء في تحليل نتائج الاختبارات وأعد مجموعة من الطرائق والأساليب الإحصائية التي استخدمها وطورها فيما بعد كارل بيرسون في الولايات المتحدة الامريكية، ويظهر التأثير الأكبر لجالتون في حركة القياس النفسي انما يظهر في اعمال كارل بيرسون اللاحقة التي كانت استمراراً لأعمال جالتون وطرائقه الإحصائية وتطويراً لها والاساس الذي يقوم عليه المنهج الاحصائي برمته.

رابعاً: دور الولايات المتحدة في تطوير القياس العقلي

في الولايات المتحدة الامريكية أسهمت اعمال وجهود عالم النفس الأمريكي جيمس كاتل اسهاماً كبيراً في تطوير حركة القياس النفسي وعلم النفس التجريبي وقد تتلمذ كاتل على يد فونت في المانيا ولكنه قام ببحوث متميزة عن بحوث بقية تلاميذ فونت مهتماً بالفروق الفردية وزمن الرجع، وأعد رسالة عن الفروق الفردية في زمن

الرجع رغم معارضة استاذه، وقد أنشأ كاتل معملاً لعلم النفس التجريبي في أمريكا وأعد مجموعة من الاختبارات التي تصدت لبعض السمات الحسية الحركية مثل قوة السمع وحدة الابصار والتمييز بين الاوزان وسرعة الحركة والقوة العضلية وزمن الرجع، ومع أن كاتل اعد هذه الاختبارات ١٨٨٥ فإنه لم يتمكن من نشرها الا في عام ١٨٩٦ بسبب المعارضة التي لقيها من استاذه فونت وكان كاتل أول من استخدم مصطلح اختبار عقلي ١٨٩٠ .

وبعد كاتل بحق مؤسس التجريب والقياس النفسي في الولايات المتحدة الامريكية وزعيمها الأول وقد أخذ عنه تلميذه العملاق كارل بيرسون الشيء الكثير غير أن كاتل اعتقد كزميله الانجليزي جالتون ان اختبارات التمييز الحسي وزمن الرجع تصلح لقياس العمليات العقلية الأكثر تعقيداً حيث يمكن عن طريق بعض الاعمال البسيطة الكشف عن الفروق الفردية في القدرة العقلية بصورة غير مباشرة .

بينيه وقياس الذكاء

يعد ألفريد بينيه المؤسس الحقيقي لحركة قياس الذكاء بصورتها الحالية، والواقع أن بينيه انشغل منذ اواخر القرن التاسع عشر بمحاولة إيجاد أداة مناسبة لقياس الذكاء وقد أعد بالاشتراك مع هنرى مقالاً نشر في عام ١٨٩٥ وجه فيه نقداً شديداً الى الاتجاه السائد آنذاك في القياس العقلي، وتركزت أفكار بينيه على ان الفروق

الحقيقية في الذكاء وأن اختبارات التمييز الحسي وزمن الرجع لا تصلح لقياس العمليات والوظائف العقلية العليا كالتذكر والانتباه والتفكير ولا مناص بالتالي من قياس هذه الوظائف والعمليات العقلية المعقدة بصورة مباشرة ودون الاعتماد على المقاييس الحسية والحركية السابقة. وفي عام ١٩٠٤ اصدر وزير التربية والتعليم قراراً بتشكيل لجنة أوكلت اليها مهمة إيجاد الوسيلة او الأداة المناسبة لعزل الأطفال المتخلفين عقلياً من البرامج التعليمية المدرسية، واقترح بينيه وزميلة سيمون عضوا هذه اللجنة مقياسهما الأول للذكاء الذى عرف باسم مقياس بينيه – سيمون فظهر الى الوجود اول مقياس للذكاء بالمعنى المعروف حالياً ١٩٠٥.

ضم المقياس في صورته الأولى ثلاثين بنداً مرتباً تصاعدياً بحسب درجة صعوبتها وتم تحديد مستوى الصعوبة بتطبيق البنود على ٥٠ طفلاً سوياً تتراوح أعمارهم بين ٣ سنوات و ١١ سنه وعلى بعض الأطفال المتخلفين عقلياً، واجرى بينيه تعديلاً على هذا المقياس عام ١٩٠٨ فزاد عدد البنود وحذف ما لم تثبت صلاحيته منها وجمع هذه البنود وصنفها في مستويات عمرية فوضعت في مستوى ثلاث سنوات البنود التي يستطيع الطفل العادي (الطفل المتوسط) في سن الثالثة الإجابة عنها بنجاح وهكذا في بقية الاعمار حتى الثالثة عشر، وطرح بينيه مفهوم العمر العقلي الذى يعبر عن أداء الطفل المتوسط في عمر زمنى معين، وامكن عن طريق هذا المعيار مقارنة درجات اي طفل في الاختبار بمتوسط درجات أبناء عمره،

كما امكن مقارنة أداء الطفل بالمستويات العمرية الأدنى والاعلى من عمره فاذا استطاع طفل في الخامسة من عمره اجتياز الاختبارات الخاصة بالمستوى العمري ٦ سنوات فإن عمره العقلي ٦ سنوات واذا لم يتمكن طفل في الثامنة من اجتياز الاختبارات المخصصة من المستوى العمري ٧ سنوات فإن عمره العقلي هو ٧ سنوات وكان لمعيار العمر العقلي أهمية فائقة بوصفه محكاً للتمييز بين مستويات الذكاء وجذب الاهتمام لاختبارات الذكاء لسهولته ووضوح دلالته .

وفي عام ١٩١١ اجري بينيه تعديلاً اخر على مقياسه وأضاف اليه بنوداً جديدة لمستويات عمرية اعلى تصل الى عمر الراشدين وقد لاحظ شتيرن وايده في ذلك تيرمان تغير العمر العقلي للطفل مع تقدمه في العمر الزمني وأشار الى ان مفهوم العمر العقلي يعبر بصورة مطلقة عن تقدم او تخلف الفرد مقيساً بالسنوات دون نسبة هذا التقدم او التخلف الى عمره الزمنى، ونتيجة الى ذلك تم ادخال مفهوم جديد الى القياس العقلي وهو مفهوم حاصل الذكاء (نسبة الذكاء) والذي يحسب بنسبة العمر العقلي الى العمر الزمني ثم يضرب في (١٠٠) ويعبر عن التقدم او التخلف النسبي للفرد من خلال مقارنة أدائه الفعلي كما يعبر عنه العمر العقلي بعمره الزمني، وبخلاف العمر العقلي الذي قد يتغير بصورة ملحوظة مع التقدم في العمر الزمني فإن حاصل الذكاء (نسبة الذكاء) يبقي ثابتاً او يتغير قليلاً من سنه لأخرى بالنسبة لأغلب الأطفال.

اثار مقياس بينيه – سيمون للذكاء اهتماماً كبيراً لدى الأوساط المعنية في الولايات المتحدة الامريكية، وفي عام ١٩١٠ بدأ لويس تيرمان بإعداد دراسة موسعة حول هذا المقياس في جامعة ستانفورد وطبقه على عينات أمريكية تتألف من ٢٣٠٠ طفلاً ، وتضمنت هذه الدراسة تعديلات كبيرة على المقياس وانتهت الى نشر المقياس بصورته المعدلة عام ١٩١٦ باسم مقياس ستانفورد بينيه واصبح المقياس بذلك مقياساً أمريكياً من أصل فرنسي إن جاز التعبير وسرعان ما انتشر هذا المقياس على نطاق واسع في الولايات المتحدة وغيرها من البلدان وقد استطاع هذا المقياس أن يثبت وجوده على مر الزمن بوصفه مقياساً للذكاء على درجة عالية من الجودة واسهمت التعديلات العديدة التي تعرض لها فيما بعد في تحسينه وتعزيز مكانته وما زال يحتفظ بقيمته ويتبوأ مكانة خاصة حتى يومنا هذا.

التطور في الفترة بين الحربين العالميتين

كان للحرب العالمية الأولى تأثيرها الحاسم في دفع حركة قياس الذكاء فقد برزت الحاجة في اثناء الحرب الى استخدام مقاييس عقلية جمعية تطبق على اعداد كبيرة من الافراد في وقت واحد واستدعى ذلك تصميم اختبارات ذكاء جمعية سهلة التطبيق كان اشهرها اختبار الفا للجيش واختبار بيتا للأميين وغير الناطقين بالإنجليزية، وقد استخدمت نتائج هذه الاختبارات في فرز المجندين وانتقاء القادة ورجال المهمات الخاصة كما استخدمت في استبعاد المتخلفين عقلياً وتحويل منخفضي الذكاء الى

اعمال مناسبة، وهذا الحماس الذي شهدته حركة القياس العقلي في اثناء الحرب العالمية الأولى لم يخمد في فترة السلم، وبدءاً من ثلاثينات القرن الماضي انتشرت مقاييس الذكاء في بلدان عديدة واستخدمت لأغراض عديدة.

ففي فرنسا استخدمت مقاييس الذكاء لأغراض التوجيه المهني والتشخيص النفسي، وفي الولايات المتحدة وانجلترا دخلت الاختبارات الى مختلف مجالات الحياه واعتمدت نتائجها في دخول المدارس والمعاهد والجامعات وفي الحصول على عمل أو وظيفة ، وفي الاتحاد السوفيتي استخدمت اختبارات الذكاء في مجالي التعليم والتوجيه والاصطفاء المهني.

ادي هذا الانتشار الواسع للمقاييس العقلية وبفعل تأثيرها المباشر في حياة الملايين من الناس الى ظهور تيارين متعارضين الأول يناصر حركة القياس والثاني يعارضها، ففي عام ١٩٣٦ أصدرت الاتجاهات المسؤولة في الاتحاد السوفيتي قرار بحظر استخدام الاختبارات العقلية من منطلق انها ليست أدوات علمية صادقة كما انها يمكن ان تؤدى الى تعزيز الفروق الطبقية وتكريس التمايز الاجتماعي بين الناس وهذا ما أدى الى جمود حركة القياس وتعطيلها تماماً نحو ثلاثين عاماً في الاتحاد السوفيتي، كما وجهة الى الاختبارات العقلية انتقادات شديدة في ارجاء أخرى من العالم بما في ذلك الولايات المتحدة الامريكية.

غير أن الهجوم الشديد على اختبارات الذكاء لم يوقف مسار هذه الحركة ونموها ففي عام ١٩٣٧ حدث تطور مهم في حركة القياس حيث ظهر التعديل الجديد لمقياس ستانفورد – بينيه من قبل تيرمان وميريل فقد تم تقنين هذا المقياس على عينات واسعة من الافراد ورفع فيه سقف العمر العقلي الى ٢٢ سنه واصبح له صورتان هما الصورة L والصورة M ، وحظى هذا المقياس الجديد بانتشار واسع واستمر العمل به حتى أوائل السبعينيات حيث ظهر التعديل دمج الصورتين في صورة واحدة .

وفي عام ١٩٣٩ حدث تطور مهم اخر في حركة القياس فقد وضع ديفيد وكسلر المختص النفسي في مستشفى بلفيو في نيويورك مقياساً للذكاء عرف باسم مقياس وكسلر – بلفيو، وجاء هذا المقياس ليلبي الحاجة الى مقياس ذكاء للراشدين، كما استخدم على نطاق واسع للأغراض التشخيصية ويتضمن مقياس وكسلر – بلفيو مجموعتين من الاختبارات الفرعية الأولى هي مجموعة الاختبارات اللفظية والثانية هي مجموعة الاختبارات العملية ويفيد تطبيق هاتين المجموعتين في استخلاص حاصل الذكاء اللفظي وحاصل الذكاء العملي وحاصل الذكاء الكلي (العام)، وتبرز أهمية هذا المقياس في أنه يولى اهتماماً كبيراً للجوانب الادائية (غير اللفظية) للذكاء، ويسد بذلك فجوة كبيرة تعاني منها مقاييس الذكاء المشبعة بالعامل اللفظي في معظمها كما يتميز بخصائصه التمييزية الدقيقة وقدرته التشخيصية العالية، وقد

وضع وكسلر بالإضافة الى هذا المقياس مقياساً خاصاً بالأطفال من ٥ – ١٥ سنة كما طور مقياساً للأطفال من ٤– ٦ سنوات ونصف، وسار وكسلر في تصميمه تصميم مقياسه الأول.

التحليل العاملي وتطور القياس

كان لمنهج التحليل العاملي دور لا يستهان به في تطور حركة القياس العقلي فقد استخلص سبيرمان منذ أوائل القرن الماضي عامل الذكاء العام (العامل العام) بوساطة التحليل العاملي لنتائج مجموعة من الاختبارات، وجاء رفض ثرستون لنظرية سبيرمان في العامل العام الذي يتضمن أن نتائج الارتباطات بين مجموعة من الاختبارات لا تعبر عن عامل واحد عام بل تعبر عن عوامل متعددة، بمثابة قوة دافعة جديدة لحركة القياس. ووفقاً لنظرية العوامل المتعددة صمم ثرستون بطارية الاختبارات العقلية الأولية PMA وانتشرت هذه البطارية على نطاق واسع خلال الثلاثينيات وتضم هذه البطارية اختبارات لسبع قدرات أولية هي : الفهم اللفظي، والطلاقة اللفظية، والقدرة العددية، والادراك المكاني، والتذكر، والسرعة الادراكية، والاستدلال المنطقي. ويعطي تطبيق هذه البطارية من الاختبارات درجات منفصلة لكل قدرة على حدة مما يفسح المجال لتحليل قدرات الفرد وتشخيصها كما يقدم تشخيصاً اكثر تركيباً وتعقيداً لقدرات الفرد من الدرجة الكلية التي يقدمها بينيه أو التصنيف الثنائي الذي يتضمنه مقياس وكسلر، وقد عارض تومسون فكرة العامل

العام كما عارض فكرة العوامل المتعددة، غير أن بحوثه في مجال التحليل العاملي كان لها وزنها وأدت هذه البحوث كما أدت بحوث زملائه من قبله الى تطوير البحوث الخاصة بالقدرات العقلية وقياسها.

وعموماً فإن من الملامح البارزة لحركة القياس أثناء الحرب العالمية الثانية وبعدها الاهتمام باختبارات القدرات الخاصة التي بدأت تنافس اختبارات الذكاء العام وكانت بمثابة خطوة متقدمة في حركة القياس، وقد جاءت هذه الاختبارات لتلبى حاجات عملية ملحة برزت اثناء تلك الحرب الساخنة واشتدت في فترة الحرب الباردة

من الملامح البارزة لحركة القياس ظهور الحركة النشطة في مجال قياس الشخصية التي بدأت في أثناء الحرب العالمية الأولى وانتشرت بعدها على نطاق واسع. والواقع أن حركة قياس الشخصية لم تظهر بصورة مباشرة في أثناء الحرب العالمية الأولى بل كانت نتائج محاولات وجهود سابقة واستمراراً لها تماماً كحركة قياس الذكاء، ومما مهد لظهور هذه الحركة الجهود الذى بذلها كارل بيرسون وروزانوف في الولايات المتحدة الامريكية ويونغ في سويسرا وكريبلن في المانيا وكان الهدف من وراء تلك الجهود والمحاولات هو الكشف عن سمات الشخصية والتمييز بين الافراد الاسوياء وغير الاسوياء، من جهة أخرى أدى انتشار اختبارات الذكاء

العام والقدرات الخاصة واستخدامها على نطاق واسع في مجالات الحياه المختلفة الى اثارة العديد من الأسئلة بصدد فاعلية هذه الاختبارات وقدرتها على الكشف عن مستوى الأداء العقلي للفرد حيث تبين أن ثمة عدداً من المتغيرات " غير العقلية " التي يمكن أن تتدخل في الأداء العقلي للفرد وتتصل هذه المتغيرات بشخصية الفرد و الشخصية وعدم الاقتصار على الاختبارات العقليه وحدها في دراسة الفرد.

قياس الشخصية عن طريق التقرير

ومن الأسماء المهمة في حركة قياس الشخصية روبرت وودورث الذي وضع في اثناء الحرب العالمية الأولى قائمة لقياس الشخصية معتمداً على أسلوب التقرير الذاتي أي وصف الشخص لنفسه وتقديره لسمات شخصيته وانفعالاته الخاصة، وتضمن قائمة وودورث للشخصية ١١٦ سؤالاً وتتناول مجموعة من اضطرابات الشخصية كالمخاوف والوساوس وغيرها وقد طبقت هذه القائمة على اعداد كبيرة من المجندين بهدف التعرف على المضطربين نفسياً والذين يعانون من مشكلات انفعالية حادة تؤثر في أدائهم في الحرب، وفي عام ١٩٢٨ ظهر اختبار ألبورت للسيطرة والخضوع والذي اعتمد بدوره أسلوب التقرير الذاتي واتجه الى الكشف عن الميل الى السيطرة أو الخضوع من خلال دراسة استجابات الفرد للمواقف الاجتماعية التي تتأثر بذلك الميل.

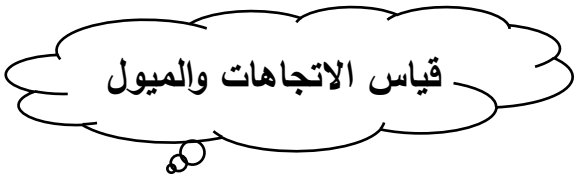
غير أن التطور الحاسم في حركة قياس الشخصية عن طريق التقرير الذاتي تأكد من ظهور اختبار مينسوتا المتعدد الأوجه للشخصية MMPI الذي وضعه هاتاواى وماكلينى عام ١٩٤٣ وتضمن هذا الاختبار ٥٥٠ بند تتناول الجوانب المختلفة للشخصية وقنن هذا الاختبار على عينة من الاسوياء والمرضى المقيمين في مستشفيات الامراض العقلية ويمكن من خلال تطبيق هذا الاختبار رسم الصفحة النفسية للفرد (البروفيل) التي تعبر بوضوح عن درجات في الاستواء او الاضطراب النفسي لدى الفرد في المقاييس المختلفة التي يضمها الاختبار ويعاب على اختبارات الشخصية عن طريق التقرير الذاتي أن البيانات التي تقدمها قد لا تصف الشخص بصورة موضوعية وصادقة طالما أنها تعتمد على وصف الشخص لنفسه بنفسه، وقد ظهر أن المفحوصين يميلون غالباً الى الرد بصورة إيجابية على الأسئلة التي تتضمن بعض الخصائص أو المواقف الاجتماعية المرغوبة كما أن ردودهم غالباً ما تكون سلبية اتجاه النواحي غير المرغوبة اجتماعياً.

والاتجاه الثاني في حركة قياس الشخصية تمثل في الطرق الاسقاطية التي بدأها الطبيب النفسي السويسري هرمان رورشاخ منذ عام ١٩١١ والذي لاحظ أن الاسوياء والمرضى من الناس عندما ينظرون الى اشكال غامضة لا معنى لها كبقع الحبر

مثلاً فإنهم يرون فيها اشكال عديدة ومتنوعة مما يشير الى انهم يسقطون اتجاهاتهم وعقدهم الشخصية على تلك المثيرات الغامضة، وقد وضع رورشاخ اختباره الشهير لبقع الحبر عام١٩٢١ ويحتوى هذا الاختبار على عشر بطاقات بيضاء تظهر على كل منها بقعة حبر كبيرة وتأخذ شكل نصفين متناظرين، وتكون مهمة المفحوص هي التعبير عما يراه في تلك البقع بحرية وتؤخذ اجاباته دليل على بعض سماته الشخصية الانفعالية، ويعد هذا الاختبار من أكثر الاختبارات الاسقاطية شيوعاً وانتشاراً ويستخدم بخاصة لأغراض التشخيص الاكلينيكي.

ومن الاختبارات المهمة التي اعتمدت على الطرائق الاسقاطية اختبار تفهم الموضوع الذي نشره موراى ومورجان عام ١٩٣٥ وتضم المجموعة الكاملة لهذا الاختبار ثلاثين بطاقة أو صورة تعبر عن مواقف مختلفة كالخطر والخوف والعدوان والميول الانتحارية والجنس وغيرها بالإضافة الى بطاقة واحدة بيضاء، ويطلب من المفحوص أن يؤلف قصة من خياله تناسب كل صورة من الصور المعروضة عليه، ويقوم هذا الاختبار على افتراض أن المفحوص سيسقط مشاعره الذاتية وقيمه واتجاهاته على تلك المثيرات كما سيعبر عن الضغوط التي يعاني منها، ويختلف هذا الاختبار عن اختبار بقع الحبر في أنه يتضمن مواقف شبيهه بمواقف الحياة الواقعية ولهذا فه ينتمي الى فئة الاختبارات المحددة البناء جزئياً ويستخدم هذا الاختبار في البحث وفي التشخيص الاكلينيكي.

وعموماً توجه الى الاختبارات الاسقاطية انتقادات شديدة منها أن هذه الاختبارات تضم مثيرات غامضة وغير موحدة المعنى، كما أن "معايير" تفسير الدرجات فيها كثيراً ما تتأثر بالعوامل الذاتية للمختبر مما يطرح سؤالاً كبيراً حول تقنينها وصحة معاييرها هذا بالإضافة الى ان تلك الاختبارات تعاني من ضعف مستوى الصدق والثبات بصورة واضحة مما يدعو الى التساؤل حول جدواها كأدوات بحث وقياس والتشكيك بقيمتها العلمية والعملية.



بالإضافة الى الطرق الاسقاطية والتقرير الذاتي التي عملت على دراسة الشخصية بوصفها كلا متكاملاً ظهر اتجاه لدراسة جوانب ومتغيرات مهمة في الشخصية كالاتجاهات والميول والقيم والآراء بهدف القاء المزيد من الضوء على الشخصية بأبعادها المختلفة، والواقع ان الاتجاه الى دراسة الاتجاهات والميول وغيرها جاء متمماً للاتجاهات السابقة التي تركزت على دراسة السمات الأساسية للشخصية، واستطاع هذا الاتجاه ان يطور أدوات بحث وقياس على درجة عالية من الدقة متخطياً بذلك الكثير من الصعوبات المنهجية التي تعرضت لها تلك الاتجاهات.

ومن المحاولات المبكرة في هذا المجال المحاولة التي قام بها ثرستون في بداية العقد الثالث من هذا القرن والذى وضع مع مساعديه نحو ثلاثين سلماً للاتجاهات

تناولت موضوعات عديدة كالمواقف من الحرب وقد انتشرت مقاييس ثرستون على نطاق واسع. كما وضع ليكرت عدد من مقاييس الاتجاهات شملت العيد من القضايا وسلك طريقة جديدة في بناء هذه المقاييس مخالفاً طريقة ثرستون التي تعتمد على اراء المحكمين في تقويم البنود المختلفة، واعتمدت طريقة ليكرت على تقديم جمل محايدة وخمس بدائل للإجابة تعبر عن شدة الاتجاه، وهذه الطريقة ابسط من طريقة ثرستون وأكثر شيوعاً وانتشاراً.

وظهرت المحاولات الأولى لدراسة الميول وقياسها منذ أوائل القرن الماضي ففي عام ١٩٠٧ وضع استانلى هول استبانة الكشف عن ميول الأطفال نحو فعاليات الاستجمام وتبع هذه المحاولة محاولات أخرى كان ابرزها مقياس سترونغ للميول المهنية الذى ظهر عام ١٩٢٧ وتضمن عدداً كبير من البنود، واتجه هذا المقياس الى الكشف عن الميول الخاصة بكل من الجماعات المهنية المختلفة كالمعلمين والأطباء والمهندسين وغيرهم وقد حظى باهتمام كبير وانتشار على نطاق واسع وتم تعديله وتحسينه عام ١٩٧٤ حيث ظهر باسم اختبار سترونغ- كامبل للميول ولهذا المقياس أهمية خاصة في مجال التوجيه المهني واختيار المهنة، ومن الأدوات المهمة لقياس الميول قائمة كودر للميول والتفضيلات التي ظهرت بصورتها الأولى عام ١٩٣٩ واختبار لى- ثورب للميول.

وعموماً تقدم مقاييس الميول والاتجاهات اسهاماً كبيراً في مجال دراسة تلك الجوانب والمتغيرات المهمة في الشخصية التي لا تتعرض لها مقاييس الشخصية ذات الطبيعة الكلية والشاملة، وتقدم هذه المقاييس خدمات عديدة في مجال الحياة العملية ولا سيما في مجال التوجيه المهني واختيار المهنة.

مفاهيم نظرية في القياس

مقدمة:

يعتبر العلم الأداة النافعة للتعامل مع البيئة لتذليلها وتوجيهها لخدمة الإنسان ولإشباع رغباته ويرجع ذلك إلى أن العلم يساعد الإنسان على فهم ما حوله من ظواهر، وعلى التنبؤ بها قبل وقوعها، وعلى التحكم فيها لتسخيرها وتوجيهها لمنفعته وخدمته . ويتميز العلم بخصائص متعددة منها الإمبيريقية والتنظيم والقياس .فالقياس خاصية من خصائص العلم بل ووسيلة من وسائله التي يستخدمها لتحقيق أهدافه.

ويستخدم القياس للحصول على تقديرات دقيقة للمساحات والحجوم والأطوال والأوزان وأبعاد جسم الإنسان وغيره من الكائنات ولما لدى الفرد من سمات عقيلة كالذكاء والقدرات وسمات وجدانية كالخضوع والاتزان الانفعالي والاجتماعي والمسؤولية والخجل وتقدير الذات وغيرها ويستخدم القياس للتعبير عن الخصائص المختلفة للظواهر والأشياء ولسمات الإنسان تعبيراً كمياً، على نحو يمكن باحثين مختلفين أن يقوموا بقياسها دون وجود فروق جوهرية بينهم . وهذا معناه أن القياس يحقق للعلم المزيد من الموضوعية. ويستخدم العلم القياس لتحديد الفروق الكمية بين

خصائص الظواهر والأشياء والسمات. فبقياس ذكاء تلاميذ فصل معين يمكن تبين ما بينهم من فروق كمية في درجة الذكاء . وبقياس حدة أبصارهم نتبين ضعاف البصر والأقوياء ومتوسطي حدة الإبصار.

ويعتمد القياس النفسي في جوهر على مصدر واحد سواء مباشرة أو بطريقة غير مباشرة وهو رياضيات الاحتمال . ويمكن القول بصفة عامة أنه حتى قبل عام١٩٦٠ لم تكن توجد أي مفاهيم رياضية حول الاحتمال ، إلى أن استطاع علماء الإحصاء باستخدام ألعاب المصادفة أن يوجهوا الاهتمام إلى ما يمكن أن يسمى رياضيات المصادفة الموضوع . فقد نشر برنوللي (١٦٥٤-١٧٠٥) أول كتاب شهير حول الموضوع، ويعود الفصل إلى دي موافر (١٦٦٧ – ١٧٥٤) إلى اكتشاف منحنى التوزيع الاعتدالي حوالي١٧٣٣ ومنذ ذلك الحين زاد الاهتمام عند علماء الفلك وعلماء الرياضيات بهذا الموضوع . وفي عام ١٨١٢ كتب لابلاسر (١٧٤٩-١٨٧٢) ما يعد أعظم ما كتب حول نظرية الاحتمالات وفيه أعلى البراهين الرياضية على طريقة المربعات الصغرى وقد استطاع باوس (١٧٧٧- ١٨٥٥) أن يبرهن على الأهمية العظمى للمنحنى الاعتدالي .

يري مهرنز (Mehrens, 1975) أن القياس عملية هي تلك العملية التي تمكن الاخصائي من الحصول على معلومات كمية عن ظاهرة ما والتعبير عنها بلغة رقمية.

ماذا نقصد بقولنا " إننا قسنا ذكاء جميع التلاميذ ؟ أو أننا قسنا أطوال جميع التلاميذ ؟ " يمكن أن نقصد " بالقياس " الإشارة إلى العملية التي قمنا بها لتقدير ذكاء أو أطوال التلاميذ ، أو الإشارة إلى النتيجة التي توصلنا إليها أي القيم التي تعبر عن درجات ذكاء التلاميذ أو أطوالهم ، أو الإشارة إلى الأداة التي استخدمت في قياس الذكاء أو الأطوال.

المعاني السابقة للقياس نموذج للاختلاف في تحديد معناه ولكن القياس الذي نقصده هنا هو مقارنة الخاصية المراد قياسها بوحدة معينة ثابتة . مثل مقارنة طول كل تلميذ بالسنتيمترات أو مقارنة وزنه بالجرامات أو مقارنة ذكائه بالوحدة التي يستخدمها الاختبار . وتكون نتيجة المقارنة هي التعبير الكمي عن الخاصية المراد قياسها أي استخدام قيمة رقمية للتعبير عن عدد الوحدات التي تتضمنها الخاصية المقاسة كالقول بأن طول التلميذ هو 155 سنتيمتر أو أن عمره العقلي 15 سنة وثلاثة شهور. ومن التعريف السابق للقياس يمكن تحديد عناصره بأنها:

١- خصائص نريد التعبير عنها كمياً.

٢- وحدة معينة ذات قيمة رقمية ثابتة نستخدمها للتعبير الكمي عما نريد قياسه

٣- مقارنة ما يراد قياسه بالوحدة المستخدمة لمعرفة عدد الوحدات التي تعبر عن الخاصية المقاسة، وتكون هذه المقارنة وفقاً لقواعد معينة .

هو عملية مقارنة شيء ما بوحدات معينة او بكمية قياسية او بمقدار مقنن من نفس الشيء او الخاصية بهدف معرفة كم من الوحدات التي يتضمنها هذا الشيء، وهو التعبير الكمي عن الظاهرة موضوع الدراسة على نحو يمكن باحثين آخرين مدربين أن يقوموا بقياسها دون وجود فروق جوهرية إذا قيست الظاهرة تحت تأثير ظروف متماثلة . فإذا كان من الممكن قياس الحرارة وسرعة الضوء فإن كثيراً من المتغيرات النفسية يمكن يقاسها فالذكاء والاتجاهات والميول أبعاد للشخصية يمكن قياسها والتعبير الكمي عنها . والتعبير الكمي عن الظواهر يساعد على ترجمة المفاهيم في صورة رقمية تسهل تحقيق خاصية التنظيم وتسهل المعالجة وكشف التناقضات والأخطاء.

ومن الواضح أن القياس الكمي اكثر دقة وموضوعية من القياس الكيفي فإذا قلنا أن احمد طوله ١٨٥ سم بينما محمد طوله ١٧٢ سم يتضح معنى الدقة والموضوعية في القياس الكمي، ورغم دقة وموضوعية القياس الكمي الا انه به قدر من الخطأ ويعد ذلك من خصائص عملية القياس ويطلق عليه خطأ القياس Error of Measurement وترجع أخطاء القياس لعدة أسباب منها ما يتصل بالفاحص أو

بظروف الموقف الذي يتم فيه القياس أو ترجع الى طبيعة أداة القياس ذاتها او قد ترجع الى المفحوص الذى نقوم بتقدير الصفة لديه.

كما تختلف دقة القياس باختلاف المجال فهو اكثر دقة في مجال الفيزياء والكيمياء مقارنة بدقته في مجال العلوم البيولوجية (الحيوان والنبات) حيث توجد عوامل كثيرة لا يمكن التحكم فيها تؤثر على الظاهرات البيولوجية، ونجده اقل دقة في مجال العلوم الإنسانية والاجتماعية مثل علم النفس ويدفعنا ذلك الى قبول أخطاء القياس تصل الى ٥% في قياس الظاهرات الإنسانية .

وبهذا يعد القياس النفسي Psychological Measurement احد الوسائل الشائعة التي تستخدم في قياس الظاهرة النفسية التي تتميز بالتعقيد والتعددية المتغيرة .

ابرز تعريفات القياس

١- **القياس** عبارة عن جمع معلومات وملاحظات كمية عن الموضوع المراد قياسه.

٢- **القياس** هو تقدير الأشياء والمستويات تقديراً كمياً وفق اطار معين من المقاييس المدرجة وذلك اعتماداً على الفكرة السائدة بأن كل شيء يوجد بمقدار وكل مقدار يمكن قياسه وهذا التعريف يمثل التعريف الاجرائي للقياس.

٣- **القياس** هو مقارنة أشياء معينة بوحدة او مقدار معياري منه بهدف معرفة عدد الوحدات المعيارية التي توجد فيه.

٤- **القياس** عملية تعتمد على الرقم في التعبير عن الخاصية المقاسة ويتم التوصل الى الرقم عن طريق وحدة مقياس يتم الاتفاق عليها.

٥- **القياس** هو العملية التي يتم فيها تقدير شيء ما تقديراً كمياً في ضوء وحدة قياس معينة أو بالنسبة لأساس معين.

الفرق بين القياس والتقويم والتقييم

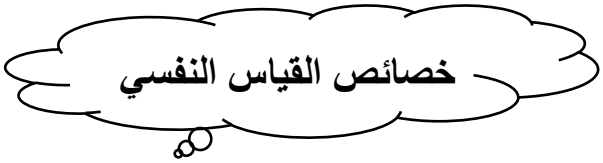
التقويم "هو إصدار حكم عن مدى تحقيق الأهداف المنشودة على النحو الذي تتحدد به تلك الأهداف بناء على الموضوعات والمعلومات التي يوفرها القياس " مثلاً هل حققت العملية التربوية الأهداف التي تسعى نحو تحقيقها وما هي الأهداف التي تحققت والأهداف التي لم تتحقق؟ ولماذا؟

"وللتمييز بين مفهومي التقويم والقياس يفضل البعض أن يقتصر التقويم على الحكم الكلي على الظاهرة، أما القياس فيعني الحكم التحليلي الذي يعتمد على استخدام الاختبارات وغيرها من المقاييس."

ويستخدم التقويم عادة معلومات وبيانات ذات طبيعة كمية وكيفية مع اً .أما القياس فإنه يقوم على جمع البيانات بطريقة منظمة، والتعبير عنها بأسلوب كمي ، وتنظيمها بطريقة تساعد على التواصل إلى نتيجة رقمية معينة .فلكي يصدر المدرس حكماً تقويمياً على التلميذ فإنه يجمع بيانات عن تحصيله وعلاقاته بزملائه باستخدام

الاختبارات والمقاييس المختلفة وبملاحظاته له أثناء تفاعله مع زملائه .أما إذا أراد أن يقيس ذكاء نفس التلميذ فإنه يستخدم اختباراً ذا مواصفات معينة ويطبقه عليه متبعاً قواعد خاصة ويصححه وفقاً لقوانين مرسومة ثم يعطى للدرجة معنى باستخدام أسلوب معين.

أما التقييم فهو عملية اصدار حكم على قيمة الأشياء او الموضوعات او الافراد في حين القياس هو تحيد ارقام للأشياء او الاحداث طبقاً لقواعد معينة.



١- القياس النفسي هو قياس غير مباشر:

فنحن لا نقيس الصفة على نحو مباشر انما نستدل عليها من مظاهرها السلوكية الدالة عليها.

٢- القياس النفسي هو قياس نسبي وليس مطلق :

فوحدات قياس التحصيل والذكاء والدافعية وغيرها من الصفات النفسية التي لا تؤسس على مقياس لا يتضمن صفر مطلق مثلما هو حاصل في قياس الطول او العرض.

٣- القياس النفسي قياس لعينه من السلوك:

فنحن في القياس النفسي نقيس عينة من السلوك وهذه العينة تكون ممثلة للسلوك ككل.

٤- القياس اجراء حيادي:

أي لا يتضمن احكاماً قيمية مثل ممتاز وجيد انما يعطى ارقام مثل حص الطالب " محمد" في اختبار الرياضيات ٣٠ من ٥٠ .

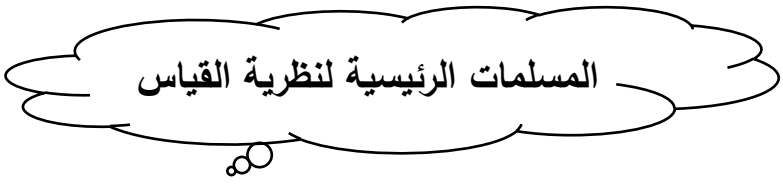
٥- القياس النفسي وسيلة وليس غاية في حد ذاته

١- الفروق الفردية بالمعنى السابق توضيحه أي الاختلاف أو التباينات بين الأفراد في درجة الصفة أو الخاصية المقاسة . وهذه الفروق تكون في التنظيم المزاجي والمعرفي للشخصية.

٢- إن وجود درجات متفاوتة من الخاصية أو السمة لدى الأفراد يوضح أنها توجد بمقادير متفاوتة " وأن ما يوجد بمقدار يمكن أن يقاس " . بناء على هذه العبارة التي وضعها ثورنديك يمكن أيضاً قياس التغيرات التي تطرأ على سمات الفرد أثناء مراحل النمو .

٣- تستخدم الاختبارات النفسية لقياس السمات المختلفة . والاختبار النفسي مقياس موضوعي مقنن لعينة من السلوك . وهذه " العينة يفترض أنها تمثل المظهر المراد قياسه للسمة أو الخاصية " . فاختبار الذكاء الذي يقوم على استخدام الأرقام يستخدم عينة من المثيرات (الأسئلة) العددية يستجيب لها الفرد استجابات تؤدي إلى قياس المظهر العددي للذكاء . ويفترض أن المظهر المقاس بذلك الاختبار يمثل المظهر العددي للذكاء فقط ولا يمثل الذكاء ككل . ومع هذا توجد اختبارات تقيس الذكاء كعامل عام وفي هذه الحالة فإنها تمثل الذكاء كعامل عام .

٤- العينة السلوكية التي يتضمنها الاختبار النفسي تمثل المظاهر المختلفة للسمة المراد قياسها ويعتمد القياس النفسي على تقديم بعض المثيرات (أو الأسئلة أو المواقف أو وحدات الاختبار) التي يرد عليها الفرد باستجابات معينة . من هذه الاستجابات أو الأداءات الظاهرة العملية الواقعية يمكن الاستدلال على سمات الفرد وخصائصه .



- **المسلم الأول : أداء الانسان يمكن قياسه وتقديره**

وهذا يعني أنه:

١- يمكن تحويل أداء الافراد من صيغة وصفية الى صيغة كمية باستخدام الأرقام حسب قواعد معينه.

٢- قابلية أداء الافراد للقياس والتقدير تمهد للعمليات المختلفة والمتتالية والمترتبة على هذه القابلية.

٣- أداء الفرد عندما يتم قياسه او تقديره في مرحله من مراحله يصبح الامر بعد ذلك ممكناً للتنبؤ بالمراحل التالية من هذا الأداء او الاداءات الأخرى.

٤- يتضمن مفهوم قابلية أداء الفرد للقياس والتقدير معنى اخضاع هذا الأداء لظروف وعوامل خارجية قد تؤثر بدرجة او بأخرى في عملية القياس والتقدير مثل ظروف التجريب التي يتعرض لها الانسان في موقف من مواقف البحث والدراسة اذ انه من الصعب جداً ان لم يكن من المستحيل عزل الأداء المطلوب قياسه عن بقية الكل الشامل للإنسان بأنماط سلوكه المختلفة.

٥- مواقف التجريب او مواقف القياس لابد ان تأخذ في اعتبارها هذا التداخل وهذه العلاقة الدينامية (علاقة اخذ وعطاء) او التبادلية بين الجوانب المختلفة للإنسان.

٦- أداة القياس والتقدير لابد ان تأخذ في اعتبارها ذلك ايضاً وهذا الامر ليس كذلك في القياس الطبيعي مثل الطول والوزن ودرجات الحرارة .

٧- يحتاج هذ القياس الى أدوات من نوع خاص وبالتالي فإن هذه الأدوات لابد ان تتميز عن بعضها البعض كما تتميز ايضاً عن الأدوات التي تستخدم في القياس الطبيعي او القياس الكيميائي او القياس الفيزيائي.

- **المسلم الثاني : أداء الانسان انما هو دالة خصائصه**

وهذا يعني أن:

١- كل أداء او سلوك انما يصدر عن خاصية واحدة او مجموعة من الخصائص يتميز بها الفرد عن غيره من بقية الافراد.

٢- ومن هذا يتضح تعقيد العلاقة بين الخصائص والأداء الامر الذى يؤثر بطبيعة الحال على الأداة المستخدمة في القياس من حيث البناء والتكوين ومن حيث الدلالة والتفسير.

٣- عند قياس الأداء الذى يرتبط بالتعبير اللغوي على سبيل المثال يجب ان نعلم ان هذا الأداء انما هو نتاج خاصية التعبير اللغوي بجانب خواص أخرى مثل الذكاء والقدرة الاجتماعية ومن هنا يتحتم علينا ان نأخذ ذلك في اعتبارنا عند فحص دلالة أداة القياس وتفسير نتائجها.

٤- وبالمثل فإنه عند بناء او تكوين أي أداة لقياس خاصية معينة مثل القدرة الرياضية فإنه يجب ان نأخذ في اعتبارنا ان هذه الخاصية تعطى اكثر من نوع واحد من الأداء.

٥- شدة العلاقة بين الخاصية والأداء فلو فرضنا ان الخاصية هي القدرة الرياضية وان الأداء هو حل المسائل الرياضية فإنه يصبح من الضروري ان تكون أداة القياس على درجة كبيرة من الحساسية لشدة العلاقة بين القدرة والأداء حتى نتمكن من قياس الأداء وارجاعه الى الخاصية الأولى او خصائص متعددة.

- **المسلم الثالث: الخاصية والأداء والعلاقة بينهما تختلف من فرد لآخر وان هذا الاختلاف هو ما قامت عليه عملية القياس**

وهذا يعني أن :

١- وجود الفروق الفردية والاعتراف بها ضمن مسلمات نظرية القياس يحدد موقف عملية القياس وأدوات القياس من وسائل المعالجة الرياضية والاحصائية.

٢- اثر مفهوم التباين والاختلاف والفروق الفردية على بناء أداة القياس في حد ذاتها واختيار وحداتها والتأكد من فاعلية هذه الوحدات فالأداة التي تبنى من اجل قياس الفروق تختلف عن الأداة التي تبنى من اجل قياس الكمية .

٣- عملية تحليل وتفسير الدرجات التي نحصل عليها عن طريق هذه الأدوات التي تبنى من اجل قياس الفروق او القياس النسبي، فعند التحليل او التفسير لابد ان نشير الى اطار مرجعي تنسب اليه هذه الدرجات وقد يكون هذا الاطار المرجعي هو جدول المعايير.

- **المسلم الرابع: كل درجة على مقياس ما انما تتكون من درجتين وهما الدرجة الحقيقية والدرجة التي تعود للخطأ**

وهذا اعتراف واضح وصريح بوجود الخطأ كمكون من مكونات الدرجة:

١- لتحديد العلاقة بين المكون الحقيقي ومكون الخطأ لدرجة ما فإننا نسلم ايضاً بأن الدرجة الكلية = الدرجة الحقيقية + الدرجة التي تعود للخطأ

وهنا يمكن ان نقول ان الخطأ يمكن تصنيفه على النحو التالي :

- **الخطأ الثابت Systematic Error**

وهو نوع من الخطأ يعود الى المقياس في حد ذاته ويتكرر بصفه منتظمة وله نفس التأثير على كل درجة على هذا المقياس.

- **خطأ المقياس Measurement Error**

وهو الخطأ الناتج عن استخدام الدرجة الظاهرية في القياس بدلاً من الدرجة الحقيقية وهو نوع من الخطأ يحتاج معالجة إحصائية خاصة للتحكم فيه.

- **خطأ الصدفة او العشوائية Random Error**

وهذا هو الخطأ الذي لا يحتاج الى شرح او توضيح اذ ان هذا النوع من الخطأ لا يمكن ضبطه او السيطرة عليه لأنه يكون عشوائياً.

٢- متوسط هذه الدرجات التي تعود الى الخطأ العشوائي لابد ان يساوي صفر وذلك ايضاً عندما يكون حجم العينة كبير جداً.

٣- معامل الارتباط بين الدرجات الحقيقية والدرجات التي تعود الى الخطأ العشوائي لابد ان تساوى صفر.

٤- درجات الخطأ العشوائي عند تطبيق مقياس ما على جماعة لا علاقة لها بدرجات الخطأ العشوائي عند تطبيق مقياس اخر تساوى صفر.

١- الوصف:

ويقصد بالوصف القيام بحصر جميع المعلومات والامكانيات المتعلقة بالموضوع المراد دراسته وهو يعد بمثابة تخطيط مسبق للموضوع المراد قياسه وتقويمه بهدف توفير كافة الظروف الملائمة لنجاح العملية التعليمية، ويساعد الوصف في وضع الرجل المناسب في المكان المناسب .

٢- التشخيص :

يقصد به تحديد جوانب القوة وجوانب الضعف لدى التلاميذ بناء على درجاتهم في الاختبارات المدرسية ويقصد بالتشخيص النفسي Psycho Diagnostics تحديد الصفات النفسية لشخص ما بمساعدة طرق معينة، كما يقصد بالتشخيص النفسي اية تقنية ترمى الى استكشاف السمات الأساسية خصوصاً تلك السمات التي تهيئ الفرد نحو الاضطرابات العقلية .

٣- التنبؤ:

وهو هدف علم النفس بصفته علماً فبناءً على فهم إمكانيات الفرد ومقتضيات العمل او مجال الدراسة ومقتضيات كل منهما يمكن توجيه الفرد الى مهنة معينة او مجال معين للدراسة ويكون احتمال نجاحه مرتفعاً ، على سبيل المثال بناءً على نتائج الاختبارات المزاجية للشخصية يمكن التنبؤ بكيفية تصرف الفرد في بعض المواقف.

مستويات القياس:

ان كلمة القياس كما يستخدمها علماء علم النفس تشمل مجموعة كبيرة من العمليات وأن الشيء الوحيد المشترك بين جميع هذه العمليات هو استخدام الأرقام، فالتعريف العام جدا الذي نضعه للقياس هو ببساطة أن القياس يعني تحديد ارقام حسب قواعد معينة.

وهذه القواعد ليست طبيعة ضيقة ومحددة كما كثيراً ما يظن البعض من ذوى المعلومات الرياضية المحددة فإذا لم نفكر جيداً في ماهية القياس فإنه من المحتمل جدا ان نفترض أن العمليات الحسابية الأولية وهى الجمع والضرب والطرح والقسمة يجب أن تطبق على جميع ان أنظمة القياس وبالتالي يكون رد الفعل الأول هو ان نستنتج ان القياس مستحيل ان لم يكن من الممكن تطبيق بعض العمليات الحسابية فلنأخذ ميدان قياس الذكاء على سبيل المثال فمنذ بداية استخدام هذه الاختبارات نجد أن ذوى الفكر الناضج كثيراً ما يقولون انه لا معنى للقول بأن الطفل اذا كانت نسبة ذكائه ١٥٠ فإنه يكون اذكى مرتين من طفل نسبة ذكائه ٧٥ حيث ان اختلاف الأطفال بعضهم البعض في الاستجابة لمواقف التفكير المجرد حقيقة واقعة ولكن ليس هناك أي مبرر في سلوكهم يدعو الى القول بان هذا الاختلاف يمكن التعبير عنه بكسر او نسبة او قسمة فالحقيقة انه لا معنى لقسمة نسبة ذكاء ما على ذكاء اخر.

وبمجرد ان تنبه الاخصائيون النفسيون الى هذه الخاصية في بعض الأرقام التي كانوا يستخدمونها ادركوا ان هناك مواقف مناظرة لذلك خارج علم النفس ومثال ذلك درجة الحرارة فانخفاض درجة الحرارة اثناء النهار لا يمكن القول بانه مساوي لانخفاض درجة الحرارة اثناء الليل ولذلك فان ارقام الدرجات التي تعلو الصفر لا

يمكن ان تعالج بنفس الطريقة التي تعالج بها الأرقام في الطول او الوزن، ومن الواضح ان قياس الذكاء اسبه بقياس درجة الحرارة من قياس الطول والوزن.

وقد تمكن ستينفز Stevens من تقديم تصور عام للأنواع او المستويات المختلفة للقياس التي كانت ذات فائدة كبرى لعلماء النفس، وبناء على هذا التصور او التنظيم فانه يمكن ان نقسم طرق استخدام الأرقام الى أربعة أنواع كل نوع من هذه الأنواع المختلفة من القياس له قواعده وحدوده الخاصة كما أن لكل منها طرقاً احصائياً معينه مناسبة له.

وتنظيم ستينفز يرتب مستويات القياس بناء على مدى تطبيق العمليات الحسابية المألوفة وهم كالتالي:

المستوى النسبي

المستوى الفتري

المستوى الرتبي

المستوى الاسمي

- **المستوى الاسمي Nominal Scale:**

وتسمى المقاييس المعتمدة عليه بالمقاييس الاسمية وهو المستوى الأول او الأدنى في هذه السلسلة، وفى هذا النوع من المقياس نستخدم الأرقام او الاعداد لتصنيف الاحداث او الأشياء او الافراد في فئة او أخرى وفى هذه الحالة فإن العدد المستخدم يشير فقط الى مسمى الفئة أي يشير الى اسم الفئة التي فيها التصنيف فهو مجرد عنوان للفئة للدلالة فقط على الفئات التي ينتمي اليها الافراد او الأشياء ويمكن ان يكون تصنيفا لأفراد نوع واحد مثل ترقيم ملابس لاعبي كرة القدم لتمييزهم او قد تكون تصنيفا لمجموعات من الافراد مثل إعطاء مجموعة من الرجال الرقم ١ ومجموعة من النساء مجموعة ٢، والعملية الحسابية التي يمكن ان تطبق في مقاييس التصنيف هي عملية العد اي مجرد تعداد الافراد في العمليات الحسابية الأربع : الجمع والطرح والضرب والقسمة.

عندما نستخدم العدد للتصنيف لفئات معينة نقوم بحساب عدد الحالات الموجودة في كل فئة أي تكرار كل فئة ومن المفترض ان جميع الحالات التي تقع في فئة معينة تشترك في صفة نوعية عامة دون ان تهتم بتفرد أي حدث او تميز كل فرد مقارنة بالأخر من نفس الفئة، اما عندما نستخدم العدد لتعيين عناوين للأشياء او الافراد فهو لا يحمل أي معنى تفضيلي بين الحالات ولا يحمل أي قيمة كمية تميز حالة عن أخرى فالسيارة رقم (١) في السباق لا يعنى انها الأولى او الأسرع او

الأقوى مقارنة بالسيارة رقم (١٠)، ويستخدم المقياس الاسمى حينما نكون بصدد الإجابة عن أسئلة خاصة بالمقارنات.

- **المستوى الرتبي Ordinal Scale:**

وتسمى المقاييس المعتمدة علية بمقاييس الرتبة وهو مستوى ارقى من مستويات القياس مقارنة بالمستوى الأدنى في المقاييس الاسمية وتعتمد مقاييس الرتبة على ترتيب مجموعة من الافراد وفقاً لدرجاتهم في صفة من الصفات في سلسلة تبدأ من الأقل وتنتهي بالأعلى بناء على الخاصية المطلوب قياسها دون أن تعطى أي أهمية لتساوى الفروق بين الرتب، فالأول يتميز بأنه صاحب اكبر درجة في الصفة المقاسة والثاني هو الذي يليه مباشرة في الدرجة وهكذا، والقياس الرتبي لا يسمح لنا ان نقول بدقة كم يكون الفرق بين اثنين من هؤلاء الافراد، فعندما تقوم لجنة بترتيب خمسة من المتقدمين للمنح الدراسية بناء على تقديرهم العام فإنها تستخدم في ذلك مقياس الترتيب، وكثيراً ما تستخدم الدرجات المئينية في التعبير عن نتائج اختبارات الطلاب وحيث ان هذه الدرجات هي نوع من الترتيب فإنها تمثل أيضا مقياس الترتيب.

والواضح ان القياس الرتبي هو اكثر صور القياس بدائية (باستثناء المقاييس الاسمية) لأن الرتبة تعطى معنى كمي عام غير دقيق يسمح لنا فقط بترتيب الافراد من الأعلى رتبة الى الأدنى رتبة دون ان تقدم معلومات دقيقة عن كم ومقدار الخاصية موضوع القياس، ويمكن ان نطلق على هذا الترتيب الوضع النسبي لأن

هذا الوضع يتغير بإضافة او حذف أي فرد ونظراً لعدم تساوى الفروق بين الرتب لا نستطيع اجراء أي من العمليات الحسابية الأربع على مثل هذه الرتب او الاعداد المناظرة لها كما لا يمكن استخراج متوسط لها او حساب الانحراف المعياري لها، ولكن الطرق الإحصائية التي تقوم على فكرة الترتيب تكون صالحة للاستخدام، حيث يمكن التعامل مع الرتب باستخدام بعض المقاييس الإحصائية مثل معامل ارتباط الرتب والذي يعتمد في جوهره على ترتيب الافراد في الصفة المقاسة، وتستخدم مقاييس الرتب في كثير من الحالات في علم النفس فمثلاً: حينما نرتب أسئلة الاختبار وفقاً لمستوى صعوبتها او حينما نطبق اختبار للاتجاهات نحو المدرسة ونستنتج ان اتجاهات "محمد" اكثر إيجابية نحو المدرسة بينما "محمود" لديه اتجاهات اقل إيجابية من "محمد" وأكثر إيجابية من "حسن" وهذا يعد ترتيب لدرجة إيجابية اتجاهاتهم لكن الفروق بين درجات ايجابيتهم لا تكون بالضرورة متساوية.

فمن الممكن على سبيل المثال ان نعين ما اذا كانت هناك علاقة بين رتب مجموعة من الأطفال على مقياس للشعبية ورتبهم للمقياس الاعتماد على الغير.

- **المستوى الفتري Interval Scale:**

يمكن أن نطلق عليه مقياس الوحدات المتساوية أو مقياس المسافة وتستخدم معظم البحوث في علم النفس مقاييس المسافة او النسبة وتمتد مقاييس المسافة بفكرة ترتيب الرتب لتشتمل على مفهوم المسافات المتساوية بين الاحداث التي تم ترتيبها

والبحوث التي تستخدم الاختبارات النفسية للشخصية والاتجاهات والقدرة تعد من اكثر الأمثلة شيوعاً للدراسات التي تستخدم مقياس المسافة إضافة الى مقاييس الذكاء والتحصيل.

ويعد مقياس المسافة اكثر دقة من المقياسين السابقين حتى تتساوى الفروق بين الأقسام المتتالية في الصفة موضوع القياس ذلك يعنى انه يتميز بوحدات متساوية تسمح لنا بتحديد مدى بعد شيئين أو شخصين بعضهما عن بعض في الخاصية التي نقيسها من خلال استخدام وحدات أو مسافات متساوية اى ان المسافات المتساوية على مقياس المسافة تدل على مقادير متساوية من الخاصية او السمة موضوع القياس بغض النظر عن موقع الوحدة على المقياس لذلك فالعمليات الحسابية المسموح بها في هذا المستوى من المقياس هي الجمع والطرح فقط دون القسمة وعلى الرغم من انه يمكننا ان نجمع ونطرح الدرجات على مقياس الوحدات المتساوية الا ان هناك عملية حسابية لا يمكن ان تجوز وهى عملية قسمة درجة على درجة أخرى اذ ان عملية القسمة ويرجع ذلك لعدم وجود صفر مطلق في هذا المقياس.

ونظراً لعدم وجود صفر مطلق ينعدم فيه وجود الصفة المقاسة لا يمكننا القول ان الطالب ل حصل على الدرجة ٨٠ والطالب ع حصل على الدرجة ٤٠ فإننا هنا يمكننا ان نقوم بعملية الجمع والطرح حيث يمكن للمعلم تزويد عدد الأسئلة او نقص عدد الأسئلة وبالتالي تكون فروق الدرجات متساوية عند الطالبين فمثلا لو زود

المعلم ١٠ أسئلة فنجد الدرجات التي يحصل عليها الطالب ل ٩٠ درجة في حين يحصل الطالب ع على ٥٠ درجة وهنا معناه ان الفرق كما هو ٤٠ درجة في الحالتين فلم تختلف الفروق بين الحالتين ولكن النسبة بين درجتيهما سوف لا تكون واحدة حيث تكون في المرة الأولى ٨٠/ ٤٠ = ٢ اما في المرة الثانية ٩٠/ ٥٠ = ١,٨ . ففي أي اختبار اذن لا يمكننا ان نعرف ما اذا كانت معلومات فرد ما هي ضعف معلومات الفرد الاخر او ثلاثة امثالها او تعادلها مرة ونصف ولكن اذا سلمنا بان كل سؤال من أسئلة الاختبار يعد مؤشراً جيداً للمعرفة في هذا الميدان او ذاك فإننا لا ننتهك قواعد الرياضيات او المنطق عندما نطرح درجة من درجة اخري او عندما نجمع مجموعة من الدرجات ثم نستخرج المتوسط فالسائل الإحصائية التي تستخدم لتحويل الدرجات الخام في أي اختبار الى أنواع مختلفة من الدرجات المشتقة تعتمد اساساً على عمليتي الجمع والطرح.

- **المستوى النسبي Ratio Scale:**

هو المستوى الأعلى من مستويات القياس ومن مميزات مقياس النسبة أن له صفر مطلق أي صفر حقيقي فالدرجة صفر في مقياس النسبة تعنى الغياب الكامل للصفة المقاسة وذلك بالإضافة الى تساوي الفروق بين المسافات، وهاتين الصفتين الصفر المطلق والوحدات المتساوية تجعل مقياس النسبة من ارقى المقاييس التي تسمح باستخدام جميع العمليات الحسابية الجمع والضرب والطرح والقسمة ومقاييس

النسبة لها جميع خصائص الوحدات المتساوية بالإضافة الى خاصة مميزه وهى وجود صفر حقيقي وربما كان هذا النوع من المقاييس مألفاً لدينا اكثر من الأنواع الأخرى لان كل الابعاد الطبيعية المعرفة تقاس بهذا المقياس.

وهذا النوع من المقاييس لا يوجد الا قليلاً في علم النفس مثل الدراسات التي تتطلب قياس عدد الأخطاء او عدد الكلمات التي تم استدعائها ودراسات زمن الرجع وقياس التعلم بوحدات زمنية (عدد الثواني) او الدراسات التي تتطلب قياس لأطوال التلاميذ او اوزانهم على سبيل المثال فنحن نستخدم وحدات الزمن المعروفة وهى الثواني وكسور الثواني وقد لا نرغب في استخراج نسب او قسمة عدد على عدد اخر.

ويسعى علماء القياس التربوي في الوقت الحاضر الى بناء نماذج رياضية تستخدم لبناء مقاييس الذكاء والتحصيل الدراسي والاتجاهات ويتوفر فيها الصفر المطلق الذي يناظر حقيقة نقطة انعدام الصفة او الظاهرة او السمة المقاسة مثل نماذج السمات الكامنة، ونظراً لتوافر جميع خصائص مقياس المسافة إضافة الى الصفر المطلق في هذه المقاييس لذلك يمكن استخدام كل العمليات الحسابية من جمع وضرب وقسمة في هذه المقاييس كما يمكن استخدام العمليات الرياضية العليا كالجبر والتفاضل والتكامل وكل الطرق الإحصائية الممكنة.

ولنعطي مثالاً يوضح لنا المقاييس الأربعة تخيل سباق للخيل في حلبة دائرية وعلى كل متسابق ان يجري خمس دورات متتالية دون توقف في هذه الحالة نجد أن :

- **الأرقام** المحددة للخيول هي المقياس الاسمي.
- **المراكز** التي وصلوا اليها بعد انتهاء دورات السباق هي مقياس رتبي.
- **عدد الكيلو مترات** التي تستغرقها الدورة الواحدة هي مقياس المسافة.
- **مقدار الزمن الكلي** الفائز في الدورات الخمسة هو مقياس النسبة.

الاختبار النفسي:

الاختبار هو طريقة منظمة لقياس سمة من خلال عينة من السلوك، والسمة هي مجموعة من السلوكيات المرتبطة التي يتصف بها .ويتطلب العمل بالنسبة لمهنة الأخصائي النفسي في كافة المجالات الاستخدام المناسب والأمثل لهذه الاختبارات، ولهذا هدفت دراستنا إلى توضيح أهم الاعتبارات العملية والفنية التي ينبغي مراعاتها عند اختيار الاختبارات النفسية، حتى يمكن التوصل إلى استنتاجات صائبة ،يمكن الاستناد إليها بدرجة عالية من الثقة في اتخاذ قرارات تتعلق بالأفراد، والتي من أهمها تحديد الغرض من القياس، ومراعاة طبيعة المجموعة المستهدفة من الأفراد، زمن وسهولة وظروف تطبيق الاختبار، وكذا طريقة تصحيحه، وسهولة تفسير درجاته،

هذا من الجانب العملي، أما فنيا فينبغي عند اختيار الاختبارات النفسية أن تتمتع بدلالات صدق وثبات مقبولة إحصائيا ،بالإضافة إلى توفر دليل الاختبار الذي يوفر لمستخدميه كل المعلومات المتعلقة به، خاصة ما يتعلق بمعايير الاختبار التي تسمح بتفسير درجاته تفسيرا مناسبا. وهو عبارة عن استشارة مقصودة تتطلب استجابة محددة وتتكون هذه الاستشارة من مثير أو عدة مثيرات وتكون هذه المثيرات في شكل أسئلة أو تساؤلات محسوبة.

تعريف الاختبار النفسي Psychological Test

عرف كرونباخ (١٩٨٤) الاختبار النفسي بأنه لا يوجد هناك تعريف مقنع للاختبار وكلمة اختبار عادة ما توحي في الذهن أنه عبارة على سلسلة من الأسئلة المقننة التي تعرض على شخص معين ويطلب منه الإجابة عنها كتابة أو شفهيا، و في ما يلي نستعرض بعض التعاريف التي تنطبق على مختلف أنواع الاختبارات و الاستبيانات:

عرف بين ١٩٥٣ Bean الاختبار بأنه مجموعة من المثيرات أعدت لتقيس بطريقة كمية أو بطريقة كيفية العمليات العقلية و السمات أو الخصائص النفسية، وقد يكون المثير هنا أسئلة شفاهية أو أسئلة كتابية أو قد تكون سلسلة من الأعداد أو الأشكال الهندسية أو النغمات الموسيقية أو صورا أو رسوما، و هذه كلها مثيرات تأثر عن الفرد و ستثير استجاباته. .

وعرف جون انيات Annette ١٩٧٤ الاختبار بأنه مهارة أو مجموعة من المهارات التي تقدم للفرد في شكل مقنن و التي تنتج درجة أو درجات وقيمة حول شيء تطلب من المفحوص لكي يحاول أداءه.

وعرف اناستازي ١٩٧٦ الاختبار النفسي بأنه مقياس موضوعي مقنن لعينة من السلوك، ويؤكد هذا التعريف على الأركان الأساسية التالية:

١- أن الاختبار مقياس موضوعي: بمعنى أن تقدير درجة الفرد على الاختبار يتم وفق نظام تقدير معين بعيداً عن عوامل التحيز الذاتي للقائم بعملية تصحيح الاختبار ، فعملية تقدير إجابات الفرد على الاختبار يتم وفق نظام تقدير معين وموحد يلتزم به جميع القائمين على عملية تقدير إجابات الفرد على الاختبار.

٢- أن الاختبار مقياس مقنن: ويقصد بالتقنين توحيد إجراءات تطبيق وتصحيح الاختبار وكذلك تفسير درجاته بحيث تكون مستقلة أيضا عن التفسيرات الذاتية فالاختبار المقنن هو ذلك الاختبار الذي يتطلب توحيد إجراءات تطبيقه وأسلوب تصحيحه وتفسير نتائجه وذلك حتى يتسنى استخدام نتائجه في اجراء مقارنه كبيرة بين الافراد الذين طبق عليهم الاختبار ومما لا شك فيه أن التقنين والموضوعية يمثلان شرطيين أساسيين من شروط بناء الاختبارات.

٣- ان الاختبار عينه ممثلة للسلوك المراد قياسه: فالاختبار لا يتألف من كل عناصر المجال السلوكي المراد قياسه وإنما يتألف من عينة ممثلة منه وبناء على

ذلك فإن فقرات أو أسئلة الاختبار يجب أن تكون ممثلة لعناصر المجال السلوكي المراد قياسه.

ويعرف جيلفورد الاختبار النفسي بأنه أداة لفحص عينة من سلوك الفرد في موقف مقتن. فالمفحوص حيث تعطي اختباراً نفسياً له يزود بوسيلة لكي يفصح عما يفعل إذا وجه بموقف معين معني تقنينه.

يعرفه سعد عبدالرحمن على أنه مجموعة من البنود أو الأسئلة أو المواقف التي تمثل القدرة أو السمة أو الخاصية المطلوب قياسها ، وعلى هذا فإنه يمكن القول بأن هذه الأداة إنما تمثل عينه من مكونات هذه القدرة أو الخاصية أو السمة وكلما كانت هذه العينة قادرة على تمثيل المجتمع الأصلي الذى أخذت منه (مكونات القدرة) كانت هذه الاداة جيدة وصالحة ويمكن الاعتماد على نتائجها.

الحاجة الى اختبارات

لقد برز القياس النفسي، أو «علم النفس القياسي»، Psychometric في بداية هذا القرن، فلم ينشأ القياس النفسي في المختبرات بل تحت ضغط الطلب الصادر عن المؤسسات الصناعية والتربوية والعيادية. فقد طرحت المؤسسات الصناعية السؤال التالي: كيف يمكن اختيار عدد كبير من الأفراد وتوجيههم بشكل

فعال؟ أما بالنسبة للمؤسسات التربوية: كيف نوجه الأفراد كيف نكتشف قدراتهم وإمكانياتهم؟ كيف نحدد لهم طريقة تربوية خاصة.

أما السؤال العيادي Clinique فكان: كيف يمكن تقييم مختلف عوامل الشخصية؟ كيف نقيس الشخصية بمختلف ديناميا‌تها؟ كيف نقيس التدهور العقلي؟ وطريقة الروائز هي التي تسمح بالإجابة على كل هذه الأسلة، ويمكن تلخيص الحاجة الى الاختبارات النفسية الى ما يلي:

١- **الحاجة لفهم نفسية الأطفال:** حيث يتم تطبيق الاختبارات في مدارس أمريكا على نطاق واسع، ولا يكاد يوجد طفل في أمريكا إلا وقد اجري عليه واحد من هذه الاختبارات. و من هنا نجد أن القياس النفسي في الوقت الحاضر يلعب دورا كبيرا في المجهود الذي يبذله المربون، لكي يفهموا نفسية الأطفال و لكي يمكنوهم من فهم أنفسهم.

٢- **أهداف جديدة للتربية:** كان اهتمام المدارس في السنوات الماضية متجها في اغلب الأمر نحو تلقين الحقائق للأطفال و تنمية مهاراتهم المختلفة وكان الغرض من الاختبارات التي كانت تستخدم و معرفة ما إذا كان التلميذ قد حفظ هجاء كلمة معينة أو حفظ درس في التاريخ... ثم في أخر السنة يرسلون تقريرا سنويا لأولياء التلميذ و الأولياء يشعرون بالألم لإخفاق ابنهم و في غالب الأحيان لم يكن الأبوين و المدرسين يعرفون سبب الإخفاق و لكن في السنوات الأخيرة التي قضاها الباحثون

في دراسة نمو الأطفال و حاجاتهم المختلفة في كل مرحلة من مراحل نموهم و أدت إلى أحداث تغيرات هامة في أهداف التربية، و أصبح علماؤها في الوقت الحاضر مقتنعين بأن الأطفال يختلفون في ميولهم و يتفاوتون في قدراتهم و يؤكد الدارسون بأنه لا بد من دراسة شخصية الطفل و العوامل التي شكلتها.

٣- **إشباع حاجات الأطفال:** وقد أدى الأخذ بهذه الأهداف إلى تحول اهتمام المدارس عن تلقين المعلومات لأطفالها إلى محاولة تعرف حاجات هؤلاء الأطفال و إشباعها، كما أدى الاهتمام المتزايد بتوجيه التلاميذ و إرشادهم إلى حث العلماء على وضع اختبارات جديدة أفضل من الاختبارات القديمة حتى نستطيع أن نعرف عن أطفالنا أكثر مما كنا نعرف من قبل و قد وضعت اختبارات لقياس كل الناحية من حياة الأطفال تقريبا و يمكن أن تكون لهذه الاختبارات قيمة عظيمة لو أحسنا استعمالها فعن طريقها نستطيع أن نعلم الكثير عن أطفالنا مثل التحصيل الدراسي، شخصية الطفل، الصحة الجسمية، القدرات الخاصة و نواحي الضعف، القدرة الدراسية.

دور الاختبارات النفسية

إنَّ الاختبارات النفسية هي أدوات أو وسائل هامة يستخدمها الأخصائي النفسي في عمليات تقدير امكانيات الفرد، وفي التشخيص والتنبؤ والتوجيه والإرشاد النفسي؟

ويمكن الإفادة منها، في دراسة مجال واسع من السلوك البشري والحصول على بيانات أو معلومات هامة عن شخصية الفرد، إذا أحسن استخدامها ووضعت لها الضوابط وأمكن معرفة معايير ثباتها وصدقها ودلالاتها الاكلينيكية وحدودها التي تقيس القدرة أو السمة المطلوب قياسها .

فالاختبار أداة هامة يتيح للسيكولوجي الحصول على طائفة منوعة من الدرجات، أو معلومات عن الفرد: اكتشاف قدراته وإمكانياته، وتحديد نقاط القوة او الضعف فيه. غير أنه لا ينبغي الوقوف على المعنى المجرد للأعداد الإحصائية الكم) في نتائج الاختبار، بل من الضروري أن نذهب إلى المعنى الكيفي لهذه الدرجات، أي فهم دلالاتها النفسية، وهي أن سلوك المفحوص على نحو معين هو الذي حدد حصوله على هذه الدرجات المعينة؛ أي ينبغي الوقوف على كيفية أداء المفحوص، وأبعاده الذاتية والموضوعية، وظروفه الخاصة في الموقف المعين.

وتستهدف الاختبارات، بالإضافة إلى اكتشاف قدرات الفرد واتجاهاته، التشخيص والتنبؤ. فمن خلال تشخيص الحالة يمكن التوقع لاتجاه سلوك الفرد في المواقف المختلفة. كما يوضّح التقييم في تقرير الوسائل الملائمة لتخفيف العوامل التي أدت إلى حدوث الحالة، ورسم خطة علاجية مناسبة، ومن ثم كانت الدقة في التشخيص والتقييم لها أهمية كبرى في عملية رسم استراتيجية التوجيه أو العلاج من أجل

تحقيق التوافق الذاتي للفرد. ومن هنا فالمقاييس تساعد السيكولوجي على اتخاذ القرار عما تعنية البحوث وعن حالة الأفراد.

١- يستهدف الاختبار النفسي التعرف على قدرات الفرد الخاصة وذكائه العام واستعداداته ومواهبه وميولة واتجاهاته وذلك بقصد تصنيفه الى مجموعات متجانسة أو من أجل توجيهه الى الاعمال التي تتناسب مع قدراته.

٢- يستخدم الاختبار النفسي في مجال التوجيه التربوي حيث يهدف الى توجيه الطلاب الى انواع التعليم المختلفة التي تتوافق مع ما لديهم من قدرات واستعدادات.

٣- يستخدم في عمليات التوجيه المهني حيث يتاح للفرد فرصة تحقيق اكبر قدر من الإنتاج

٤- تستخدم الاختبارات في عملية التنبؤ والتشخيص .

٥- تساعد الاختبارات النفسية في تقييم قدرات الفرد وامكانياته من حيث الذكاء العام وقدراته العقلية

٦- كشف الجوانب المختلفة من شخصية الفرد مشاعره وافكاره ورغباته واتجاهاته

مجالات استخدام الاختبارات النفسية

١- المجال التربوي:

تعتبر المؤسسات التعليمية والتربوية من أكثر المؤسسات استعمالاً للاختبارات النفسية فهي تستخدم الاختبارات النفسية للأغراض التالية:

١- الكشف عن المتأخرين دراسياً ومعرفة أسباب هذا التخلف بمعرفة جوانب حياته النفسية.

٢- تستخدم في توزيع الطلاب على أنواع التعليم التي تتناسب مع قدراتهم وميولهم واستعداداتهم وذكائهم العام وبذلك يمكن وضع الطالب المناسب في المكان المناسب حيث أن وضع الطالب في المكان الذي يميل اليه والتي تمكنه قدراته من النجاح فيه واحراز التقدم يؤدي ذلك الى حسن تكيفه وشعوره بالارتياح والهدف من ذلك تجنب شعوره بالإحباط والفشل.

٣- تستخدم الاختبارات في اختيار الطلاب الجدد او في اختيار المواد الدراسية التي تناسب اعماراً معينه.

٤- الارشاد التربوي ويعنى ارشاد الطلاب الذين يجدون مشاكل في دراستهم والذين يتعرضون لبعض الازمات فاستخدام الاختبارات يمكن الاخصائيين النفسين من معرفة بعض الاعراض المرضية او الأسباب التي أدت لهذه الازمات.

٥- برامج الاختبار وتستخدم في تقويم تحصيل الطلاب ومدى فاعلية برامج التعليم ومناهجه.

٦- تصنيف الطلاب تساعد الاختبارات المعلم على تقسيم الطلاب الى مجموعات متجانسة من حيث ما يملكونه من ذكاء وقدرات خاصة وكذلك يستخدم من أجل تقويم اعمال الطلاب وتحصيلهم ولمعرفة اثر أساليب التدريس وطرقه المختلفة التي يطبقها المدرس، فقد يسعى المعلم العوامل التي تؤثر في عملية التحصيل مثل الذكاء او التكيف النفسي أو الاتزان الانفعالي او الظروف الاسرية أو الظروف الصحية.

٢- في الصناعة والإدارة والمجال المهني:

تساعد الاختبارات النفسية على تطبيق مبدأ الرجل المناسب في المكان المناسب وهذا يعني توجيه الفرد لعدد من المهن التي يحتمل ان يحرز فيها اكبر قدر من النجاح والتقدم فينتج عنه تكيف الفرد مع عمله لأنه يتفق مع ميوله وذكائه وقدراته واستعداداته وحبه ومستوى طموحه ويمكن تحديد أهمية الاختبارات النفسية في مجال الصناعة والإدارة والمجال المهني من خلال التطبيق في :

١- **التوجيه المهني:** ويعنى توجيه الفرد الى نوع من المهن التي يحتمل أن يحرز فيها أكبر قدر من النجاح والتقدم وهذا التوجيه يقوم على أساس دراسة شخصية باستخدام الكثير من الوسائل كاختبارات الشخصية.

٢- **الاختيار المهني** : وفى الاختيار المهني نكون امام عدد كبير من الافراد المتقدمين لشغل وظيفه معينه ونختار لها من بين الافراد الشخص الذي يناسبها .

٣- **التدريب أو التكوين المهني:** نوع من التعليم أو اكتساب المهارات أو الخبرات والمعارف ويستخدم فيه القياس لتحديد الأشخاص لنوع معين من التدريب أي التنبؤ بنجاحهم واستفادتهم مما يقدم لهم من تدريب

٤- **التأهيل المهني:** ويصد به تدريب الافراد أو ذوى الاحتياجات الخاصة على الاعمال التي تتناسب مع ما لديهم من قدرات ومواهب واستعدادات وهذا يعنى نوع من التدريب والتأهيل.

٥- تستخدم الاختبارات النفسية في تقويم أداء العمال ومدى فعالية طرق العمل ونظم الاشراف.

٦- **في الحوادث:** تستخدم الاختبارات في الكشف عن الافراد الذين يتعرضون للحوادث اكثر من غيرهم فيتم استبعادهم من الاعمال الخطيرة .

٣- المجال العيادي:

تستخدم الاختبارات النفسية في المجالات الاكلينيكية أي العيادات والمستشفيات النفسية لمعرفة نوع الاضطرابات والامراض النفسية التي يعاني منها المريض فعلى أساس تطبيق الاختبارات يمكن تشخيص الاضطرابات ومن ثم يمكن برامج علاجية وتستخدم في التشخيص النفسي وفى تفسير سلوك المريض، وتستخدم الاختبارات النفسية في المجال العيادي من خلال الوظائف والتي منها :

١- التنبؤ بمدى فاعلية الدواء.

٢- تقويم اثر العلاج النفسي.

٣- بحث المشاكل المتعلقة بمعنى التشخيص وعملياته.

٤- في الجيش (القوات المسلحة):

لقد عرفنا ان حركة القياس النفسي شهدت تطوراً كبيراً بعد الحربيين العالميتين وقد استعملت الدول خلال الحربين الاختبارات النفسية لاختيار الجنود الصالحين وتصنيفهم على مختلف المهام حسب قدراتهم .

وتعمد الدول المتحضرة الى ادخال الاخصائيين النفسيين في مجالات الجيش والشرطة وتقريباً كل المصالح الحكومية لضرورة واهمية الاخصائي النفسي الاكلينيكي ولتقديم المساعدة لمن يحتاج المعالجة والمساعدة النفسية وانشاء ملف متكامل لكل حالة بتاريخها المرضي، وعند اختيار اختبارات القبول لمتقدمي الكليات العسكرية والشرطة يخضع لاختبارات نفسية تؤهله للانضمام للسلك العسكري من عدمه، وعليه يصبح تقرير الأخصائي النفسي في الاستمارة النفسية عامل حسم منذ البداية وقبل الدخول الى الكليات العسكرية .

ويمكن اعتبار الاختبارات والمقاييس النفسية احد الاليات واسعة الانتشار في العمليات التالية:

١- الاختيار:

للمفاضلة بين الناس للحصول على وظيفة أو ترقى أو قبول في كلية أو تخصص يعتمد على درجة الفرد في الاختبارات النفسية، وتستخدم في المؤسسات الكبرى لقبول موظفين جدد أو لترقية الموظفين في الخدمة .

٢- المركز:

تلعب الاختبارات النفسية دوراً تشخيصياً لمعرفة أوجه النقص لدى الفرد حتى يسهل تقديم برامج علاجية له وان تحديد المستويات للفرد تسهم بدرجة كبيرة في تطوير اليات تفكيره مع المهن المتغيرة حيث يقوم العديد من الأشخاص في جميع المؤسسات بأخذ مجموعة من الدراسات بهدف تنمية مهارات العمل وخاصة عند الترقي .

٣- التشخيص:

تستخدم لمعرفة نقاط القوة والضعف عند الفرد من خلال محكات أو معايير وبالتالي يسهل تقديم برامج علاجية لهم.

٤- اختبار الفروض:

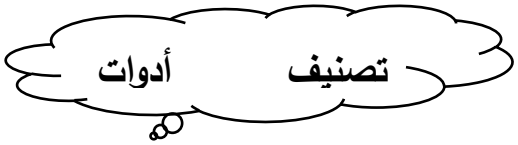
يمكن اختبار صحة الفروض من خلال الاختبارات النفسية .

٥- بناء الفروض:

يمكن بناء العديد من الفروض للتحقق من دراسة ظاهرة معينة، كما ان الاختبارات تفيد في بناء العديد من الفروض التي يحاول أن يختبرها في ميدان تخصصه من خلال المعلومات المتوفرة لديه .

٦- التقويم:

تستخدم الاختبارات في عملية التقويم وذلك باستخدام اختبارات التحصيل، فعملية التقويم في أساسها عملية قياس تهدف الى التعرف على مدى تحقيق الأهداف فالتقويم عملية تسبق اتخاذ القرار.



يمكن أن تصنف أدوات القياس بصورة أولية اختيارية الى نوعين رئيسيين هما:

١- الاختبار وهو عبارة عن مجموعة من الأسئلة أو البنود لكل منها إجابة واحدة صحيحة فقط اختبارات التحصيل ، اختبارات الذكاء، اختبارات القدرات العقلية وغير ذلك من الاختبارات التي تقيس مجموعة من الحقائق.

٢- الاستفتاء (الاستخبار) وهو عبارة عن مجموعة من الأسئلة او البنود التي تدور حول موضوع واحد أو عدة مواضيع وليس لها إجابات صحيحة أو إجابات خاطئة إذ أن المطلوب هو معرفة رأي الفرد أو نوعية استجابته في موقف من المواقف التي يمثلها ذلك السؤال أو البند .

تصنيف الاختبارات النفسية

توجد عدة طرق لتصنيف الاختبارات ، فقد تصف طبقاً لطريقة التطبيق أو محكات التصحيح أو الزمن أو المحتوى أو الأهداف ..... ونستعرض فيما يلي لهذه التصنيفات:

١-<u>التصنيف بحسب طريقة التطبيق:</u>

وهنا تصنف الاختبارات الى اختبارات فردية واختبارات جماعية :

١- الاختبارات الفردية Individual Test:

وهى الاختبارات التي تطبق على فرد واحد في المرة الواحدة مثل الاختبارات التي يتم تطبيقها عادة في مقابلة شخصية بين الفاحص والمفحوص وتحتاج بطبيعة الحال الى تعليمات من نوع خاص والى توضيح دائم لهذه التعليمات وقد يتطلب هذا النوع من الاختبارات الى ملاحظة الفاحص لأداء المفحوص في بعض المواقف والقيام بتسجيل هذه الملاحظة وتقييم هذا الأداء، ومن امثلة الاختبارات الفردية اختبار بينية في قياس الذكاء.

٢- الاختبارات الجماعية Group Test:

وهى الاختبارات التي يمكن تطبيقها على مجموعة من الافراد دفعة واحدة دون الحاجة الى جلسة خاصة في مقابلة شخصية وعلى ذلك فإن المتوقع أن تكون تعليمات هذا النوع من الاختبارات بسيطة وواضحة كما أن أداء الافراد ليس من

الداعي ملاحظته او تقييمه أثناء الاختبار بل يتم تقييم الأداء بعد الانتهاء من الاختبار ككل ومن أمثلة الاختبارات الجماعية اختبارات التحصيل المدرسي واختبار الذكاء العالي واختبار الذكاء المصور.

٢-<u>التصنيف بحسب نوع الاداء:</u>

وهنا تصنف الاختبارات أي اختبارات اقصي الأداء واختبارات الأداء النمطي (المميز):

١- اختبارات اقصي الأداء Test of Maximum Performance :

وهى الاختبارات التي يتم فيها اثارة دافعية المتعلم لتقديم أفضل ما عنده من إجابة والحصول على اعلى درجة مثل الاختبارات التحصيلية بأنواعها ونهيئ الفرصة له للدراسة والاستعداد وتحقيق اعلى مستوى للتحصيل.

٢- اختبارات الأداء النمطي Test of Typical Performance :

وهى الاختبارات التي تعكس سلوك المتعلم في الظروف العادية أو الطبيعية دون محاولة خارجية لتوجيه هذا السلوك حتى يكون التقدير الذي يحصل عليه منسجما مع السلوك الواقعي فإذا كان الاختبار يقيس قلق الطالب في الامتحان مثلا فإن القرارات التي يتم اتخاذها في ضوء النتائج تعتمد على صدق الإجابات ولذلك تعتبر مقاييس الاتجاهات والميول من نوع الأداء العادي أو النمطي.

٣-<u>التصنيف بحسب المحتوى:</u>

وهنا تصنف الاختبارات الى اختبارات اللفظية واختبارات غير لفظية واختبارات عملية (ادائية) :

١- الاختبارات اللفظية Language Test:

وتسمى الاختبارات الكتابية أو اختبارات الورقة والقلم وتقوم هذه الاختبارات التي تكون الإجابة فيها تحريرية (ورقة وقلم) أو شفوية وتعتمد هذه الاختبارات على استخدام اللغة سواء في عرض أو صياغة المفردات وكذلك كتابة التعليمات وفى الإجابة عن هذه الاختبارات، وهي الاختبارات التي لا تتطلب ولا يستدعي تنفيذها القيام بعمل يدوى ولكنها تحتاج الى تسجيل الاستجابات في صحيفة الاستجابة او الاختبار باستخدام القلم بمعنى الإشارة الى او كتابة الإجابة الصحيحة. وتتطلب هذه الاختبارات في الإجابة عليها القلم في الكتابة ، ومن مميزات الاختبارات أنها تؤدى الى تقنين موقف الأداء عليها بدرجة عالية كما يمكن استخدامها في الاختبارات الجماعية حيث يمكن اختبار عدد كبير في وقت واحد ، ومن عيوب هذه الاختبارات أنها يصعب استخدامها مع الأطفال الصغار ، كما يصعب استخدامها مع الاميين ، ويصعب استخدامها مع الافراد الذين لا يتحدثون اللغة المحلية ، وتتضمن هذه الاختبارات اختبارات القدرات العقلية واختبارات الشخصية.

٢- الاختبارات العملية (الادائية) Performance Test:

وهى الاختبارات التي تتطلب القيام بعمل ما أو اداء محدد لحل مشكلة معينة في شكل ترتيب او تصنيف او تركيب الاشكال او خامات او نماذج مختلفة ومن أمثلة هذه الاختبارات: الاختبارات الادائية بمقياس وكسلر للذكاء ومقياس ستانفورد بينيه للذكاء واختبار متاهة بورتيوس وأجهزة التأزر الحركي والتي تتطلب أداء معيناً لا يحتاج الى الورقة والقلم.

٤-<u>التصنيف حسب الجهة التي تعد الاختبار:</u>

وهنا تصنف الاختبارات الى اختبارات مقننه واختبارات من إعداد المعلم:

١- اختبارات مقننه Standardized Tests :

هي اختبارات رسمية أو منشورة حيث يعدها فريق من المتخصصين في مراكز القياس والاختبارات وتستخدم للمقارنة بين مجموعات معينه من الافراد ويقصد بالتقنين توحيد التعليمات وإجراءات التطبيق والتصحيح بالنسبة لجميع الافراد الذين يطبق عليهم الاختبار.

٢- اختبارات من إعداد المعلم Teacher-Made Tests:

وهى الاختبارات غير رسمية حيث يعدها المعلم نفسه لاستخدامها داخل الفصل المدرسي لقياس مدى تحقق أهداف المواد التي يقوم بتدريسها.

٥-<u>التصنيف بحسب نوع الاستجابة:</u>

وقد تصنف الاختبارات الى اختبارات تعرف واختبارات استجابة حرة :

١ – اختبارات التعرف Recognition :

وفى هذا النوع من الاختبارات يتم اختبار الإجابة الصحيحة بين مجموعة من البدائل وذلك كما في حالة الاختبارات الموضوعية مثل اختبارات الصواب والخطأ واختبارات الاختيار من متعدد واختبار المزاوجة.

٢ – اختبارات الاستجابة الحرة Free Response Tests :

وفى هذا النوع من الاختبارات يكون السؤال غير محدد المعالم بحيث يستثير استجابات متعددة لدى الافراد ومن أمثلة هذه الاختبارات : اختبار الرورشاخ (بقعة الحبر) اختبار تفهم الموضوع (التات) والاختبارات الإسقاطية الأخرى.

٦ –<u>التصنيف بحسب سرعة الاستجابة:</u>

وقد تصنف الاختبارات الى اختبارات سرعة واختبارات قوة:

١ – اختبارات السرعة Speed Tests:

وهى الاختبارات التي تكون فيها سرعة الاجابة هي العامل الحاسم في تحديد أداء المتعلم حيث يصعب على معظم الطلاب انهاء الاجابة على جميع الفقرات ضمن الزمن المحدد، وهى اختبارات تتألف من عدد كبير من المفردات وعلى المفحوص الإجابة عنها في إطار زمن محدد لا يسمح لجميع المفحوصين بالإجابة على جميع المقررات ومن ثم لا يتوقع لأي فرد أن يجيب على جميع المفردات في

اطار زمن محدد وتعكس درجة الفرد على الاختبار سرعة الاستجابة وتقدر درجة الفرد بعدد المفردات التي حاول الفرد الإجابة عليها.

٢- اختبارات القوة Power Tests :

وهى الاختبارات التي يعطى فيها زمن شبه مفتوح للإجابة بحيث يكون كل فرد قادر على محاولة الإجابة عن كل سؤال الا أن صعوبة الأسئلة وقدرة المتعلم على اجاباتها هي التي تحدد اداءه بمعنى أن المتعلم لا يحصل على الدرجة النهائية بسبب صعوبة الأسئلة وليس بسبب كثرتها .

٧-<u>التصنيف بحسب ميدان القياس:</u>

١- اختبارات الوظائف الذهنية Function Intellectuals Tests :

وتشمل هذه الوظائف القدرات اللفظية والقدرات الادائية والتي من أهم أنواعها :

- اختبارات التحصيل Achievement Tests :

التي تهدف لقياس التعلم الماضي للفرد أو الخبرة السابقة وتعتمد على تحديد الاعمار التحصيلية أو تحديد المستويات التحصيلية بمستويات الذكاء المختلفة.

- اختبارات القدرات Abilities Tests :

وهى التي تهدف لقياس القدرات العامة أي قياس النشاط العقلي كما هو قائم فعلا وكما يبدو في النشاط الذي يؤديه المفحوص

- **اختبارات الاستعدادات Aptitude Tests :**

والتي تهدف الى التنبؤ بما يستطيع الفرد أن يقوم به في المستقبل.

٢- مقاييس الشخصية والنواحى المزاجية Temperamental and Personality Tests :

من أهم أنواعها :

- **الاستبيان Questionnaire:**

أو ما يسمى بالتقرير الذاتي وهو يهدف الى معرفة رأي المفحوص في موضوع ما ويهدف الى جمع البيانات ويصلح لقياس الاتجاهات والميول والرأي العام ، ويستخدم فيها الورقة والقلم وتكون الإجابة بنعم أو لا.

- **المقاييس الاسقاطية Projective Tests :**

وهى تهدف للكشف عن النواحي المزاجية للحكم على مدى تكيف المفحوص لحياته القائمة وما يشوبها من جنوح وشذوذ، وهى تنظر الى الشخصية كعملية دينامية

- **المقابلة Interview :**

ويصلح هذا النوع لقياس النواحي التي لا تصلح لها المقاييس الأخرى للحكم العام على مدى صلاحية الفرد لعمل ما

- **الاختبارات الموقفية Situational Tests :**

وتصلح هذه الاختبارات لقياس القدرة على التصرف والكشف عن صفات الزعامة والاتزان الانفعالي وغير ذلك من الصفات الأخرى.

٨-<u>التصنيف حسب طريقة تفسير النتائج:</u>

١- اختبارات معيارية المرجع :

وهى الاختبارات التي تعتمد عند تفسيرها للدرجات على خصائص المجموعة التي طبقت عليها حيث يجب ان تكون هذه الخصائص مماثلة مع خصائص الطلاب من حيث النوع والجنس والمنطقة الجغرافية والمرحلة الدراسية وتهتم هذه الاختبارات بالكشف عن الفروق الفردية بين الطلاب وذلك من خلال مقارنة أداء الطالب بأداء اقرانه كجماعة معيارية بمعنى أن هذه الاختبارات تركز على وضع الطالب النسبي في علاقته بالأخرين .

٢- الاختبارات محكية المرجع

وهى الاختبارات التي لا تعتمد عند تفسيرها لدرجة الطالب على خصائص مجموعته التي طبق عليها الاختبار ولكنها تعتمد على مقارنة أداء الطالب بمحك او مستويات أداء متوقعة ومحددة مسبقة ويتم تحديد هذه المستويات في ضوء الأهداف السلوكية المراد قياسها وتقيس هذه الاختبارات مدى تحقيق الطالب لهذه الأهداف السلوكية وبذلك يمكن التحقق من كفاءة كل طالب من اكتسابه للمهارات والمعارف المرجوة ومن ثم تشخيص نواحي القوة والضعف.

الفصل الثالث

نظريات القياس

أولاً: النظرية التقليدية classical test theory

طور هذه النظرية شارل سبيرمان (١٩٠٤) ويطلق عليها نظرية الدرجة الحقيقية ودرجة الخطأ وتم تطوير هذه النظرية بواسطة عدد من السيكومترين المعاصرين مثل فيلدت وبرينان (١٩٨٩) ، جالكسون (١٩٥٠) ، كلاين (١٩٨٦) ولورد ونوفيك (١٩٨٦) . حيث أشار هامبلتون ٢٠٠٤ أن النظرية الكلاسيكية من أقدم النظريات التي استخدمت في تطوير الاختبارات وبنائها وقد استمر العمل بها لأكثر من ٨٠ عاماً حيث تم استخدامها في الآلاف من الدراسات والبحوث النفسية والتربوية وهى حتى الآن لازالت تستخدم في العديد من الدراسات .

تعرف هذه النظرية بنظرية الدرجة الحقيقية (True Score Theory) وذلك لاهتمامها بإيجاد الدرجة الحقيقية للمفحوص ، حيث تركز على تعزيز العلاقة بين الدرجة الملاحظة والدرجة الحقيقية ، وايضاً تسمى نظرية الثبات بسبب أن الهدف الرئيسي منها هو تقدير ثبات الدرجات الملاحظة على الاختبار حيث تستخدم لدراسة مجموعة من استجابات المفحوصين على الاختبار ، حيث إنه يقيس مدى نجاح المفحوصين على المفردة وعليه يتم تحليل المفردة لاستخراج معاملات الصعوبة والتمييز وكذلك تحديد فاعلية البدائل حتى نتمكن من الحكم على المفردة للكشف عن نواحي القوة والضعف ومن ثم حذف وتعديل وذلك حتى يمكن الحصول على مقياس جيد .

فوجد أن الفقرات المتوسطة الصعوبة تزيد من عامل الثبات عن المفردات الصعبة والسهلة حيث أنه إذا كانت القيمة تقترب من ١,٠ فإنها تشير إلى سهولة المفردة ، وإذا كانت القيمة تقترب من صفر فإنها تدل على صعوبة المفردة ، وكذلك بالنسبة للفقرات القادرة على التمييز بين الأفراد تعتمد بدرجة عالية عن طريق وضع الفقرات ونوعيتها وأن العمل على تحسين القدرة التمييزية للفقرات في الاختبارات الصعبة هو بحد ذاته عمل لتحسين ثبات الاختبار .

أيضاً استندت هذه النظرية إلى افتراضات فحواها أن لكل فرد قدراً من السلوك أو الاستجابات غير الملاحظة التي لا يمكن قياسها بصورة مباشرة ، ولكن يمكن تقديرها من خلال الاستجابات الملاحظة للمفحوص.

فقد سيطرت هذه النظرية فترة ليست بقصيرة وما زالت تستخدم أسس هذه النظرية في مواقف اختبارية متنوعة تتضمن بناء مختلف أنواع الاختبارات النفسية وكذلك تحليل البيانات المستمدة من هذه الاختبارات.

تعتمد النظرية الكلاسيكية على المجموع الكلى لدرجات الاختبار المكون من بنود متعددة حيث تفترض أن النتيجة عبارة عن الدرجة الحقيقة ودرجة الخطأ فمثلاً إذا قمنا بتطبيق اختبار ما لمجموعة من المفحوصين في ظروف مختلفة فإننا سنحصل على العديد من الدرجات المختلفة للمفحوصين وهنا تتضمن الدرجة الملاحظة خطأ عشوائي ويمكن أن تأخذ شكل المعادلة التالية :

$X = T + E$

X : تعنى الدرجة الملاحظة للمفحوصين

T : تعنى الدرجة الحقيقية (الصحيحة)

E : تعنى درجة الخطأ

وتقوم النظرية الكلاسيكية في القياس على أربع مسلمات رئيسية هى :

١-أداء الفرد يمكن قياسه وتقديره .

٢-أداء الفرد هو دالة لخصائصه .

٣-الخاصية والأداء والعلاقة بينهما تختلف من فرد لآخر "الفروق الفردية ".

٤-القياس الظاهري الكلى يتكون من قياس حقيقي وآخر يرجع إلى الخطأ .

<u>أساليب القياس التقليدية أو الكلاسيكية :</u>

انتشرت أنواع كثيرة من الاختبارات النفسية والتربوية، وتعددت نتيجة لتعدد الأغراض وتعدد السمات المقاسة ، لذلك فان تصنيفها يعتمد على طرق عدة و مختلفة وعلى حسب تفسير النتائج إلى قسمين هما:

١- الاختبارات المعيارية المرجع

٢- الاختبارات المحكية المرجع

الافتراضات التي تقوم عليها النظرية التقليدية في القياس :

كانت النظرية التقليدية الأساس لعملية القياس لأكثر من ٨٠ عام وهناك أسس وافتراضات تقوم عليها هذه النظرية حيث طورت بعض المفاهيم الأساسية لعملية القياس، وعليه نجد أن لهذه النظرية عدة افتراضات تقوم على أساسها والتي نوضحها في الأسطر التالية :

١-**الاستقلال الخطى (Linear independence)** : تفترض النظرية الكلاسيكية للاختبار أن انحدار الدرجة الحقيقية على الدرجة الخام خطياً أى أنه كلما زادت قدرة الفرد زادت درجته الملاحظة على الاختبار بشكل خطى ٢-الدرجة الكلية التي يحصل عليها الفرد هي نتاج نوعين من الدرجات : درجة حقيقية ودرجة خطأ وليس بينهما ارتباط .

٣-درجة الخطأ ليست محددة وثابتة في كل المواقف والظروف بل تتغير بتغير المواقف والظروف .

٤-التوزيع الاعتدالي لدرجات الطلاب على متصل القدرة التي يقيسها الاختبار يحقق خصائص مقياس المسافة (Interval Scale) لدرجات الاختبار ، حيث تعتمد النظرية الكلاسيكية في اختيار مفردات الاختبار على افتراض أن عدد الأفراد الواقع في مستوى معين من القدرة ، يناظر العدد المتوقع من دالة الكثافة الاحتمالية Normal Probability Density Function .

٥-إن الدرجات التي يحصل عليها الفرد في مفردات الاختبار يمكن جمعها كما لو كانت تمثل ميزاناً خطياً ، والمفردات المتعلقة بالمتغير المراد قياسه تحمل المعنى نفسه لدى جميع المختبرين .

٦-أخطاء القياس هي أخطاء عشوائية غير منتظمة تؤثر في دقة القياس ، حيث إن الأخطاء غير المنتظمة تختلف عن الأخطاء المنتظمة و التي تمثل سمة الاختبار التي لا تؤثر في دقة القياس و يكون متوسط خطأ القياس يساوى صفر .

٧- الخطأ المعياري للقياس لا يختلف من شخص لشخص فمثلاَ لو طبق اختبار معين على مجموعة من الأفراد فإن النتيجة يمكن تعميمها على أفراد آخرين مقبلين على الاختبار بغض النظر عن الدرجة الخام للاختبار حيث إن الخطأ المعياري لكل الدرجات واحد .

رغم الاهتمام الكبير بنظرية القياس التقليدية، والاعتماد عليها لفترةٍ طويلة من الزمن ، والتي مازالت مستمرةً إلى يومنا هذا ؛ إلا أنها تعرضت للكثير من الانتقادات التي تتهمها بالاعتماد على الدرجة الحقيقية، وإهمال تحقيق الموضوعية في القياس .

مشكلات النظرية الكلاسيكية :

على الرغم من استخدام الباحثين للنظرية التقليدية في بناء الاختبارات وتحليل وتفسير نتائجها ، إلا أن هناك بعض المشكلات التي تقلل من دقة وموضوعية هذا الاستخدام حيث يشار إلى أنها نماذج ضعيفة وهذه المشكلات هي :

١-تفترض النظرية التقليدية أن أداء الفرد على اختبار ما يأخذ شكل المنحنى الاعتدالي في الاختبارات جماعية المرجع وهو ما يخالف الواقع في كثير من الأحيان .

٢-صعوبة إثراء الاختبارات المعدة باستخدام أساليب القياس الكلاسيكية بمفردات جديدة ، حيث يتطلب ذلك إعادة تقنين جميع الاختبارات مرة أخرى ويعد هذا أمراً صعباً للغاية .

٣-تختلف قيمة معامل ثبات الاختبار من مجموعة لأخرى ، حيث إن ثبات الاختبار يعتمد إما على تطبيق الصورة الاختبارية مرتين على أفراد العينة ، أو على إعداد صورة متكافئة من الاختبار ويعد هذا أمراً صعباً حيث يحد من إمكانية تعميمه ، وذلك لأن بعض درجات الاختبار قد تتضمن أخطاء للقياس أكثر من الأخرى .

٤-انعدام خطية القياس : وتعنى خطية القياس أن تقدير الفرق بين أي قياسين متتالين على نفس التدريج يكون ثابتاً ولا يتغير هذا الفرق بتغير الأداة المستخدمة طالما كانت أداة مناسبة ولها وحدة قياس ثابتة ، حيث إن المقاييس الخطية توفر

وحدات قياس متساوية على مدى المستويات المختلفة من المتغير موضوع القياس ، أما في المجال السلوكي فإن درجات الاختبار لا تعطى قياسات خطية .

٥-لا تقدم هذه النظرية تفسيراً سيكولوجياً يوضح كيف يحاول الفرد الإجابة على إحدى مفردات الاختبار على الرغم من أن هذا التفسير يعد ضرورياً ولازماً إذا أردنا التنبؤ بخصائص الدرجات المستمدة من مجتمع معين أو مجتمعات مختلفة من الأفراد أو إذا أردنا تصميم وبناء اختبارات تتميز بخصائص سيكومترية معينة تناسب مجتمع ما من الأفراد. هذا بالإضافة إلى تكوين مفردات اختبار يتغير بتغير عامل الزمن أي بمضي الزمن بالنسبة لعينة الأفراد .

٦-جميع خصائص الاختبارات التي تستند في بنائها على أسس النظرية التقليدية مثل معاملات الصعوبة ومعاملات التمييز والثبات والصدق تختلف باختلاف طبيعة وخصائص عينة الأفراد التي يجرى عليها الاختبار ، وهذا لا يقدم تقدير دقيق للقدرة المنخفضة أو العالية وهذا بدوره يمكن أن يؤثر على صحة النتائج ، وأيضاً تستند على خصائص عينة المفردات التي يشتمل عليها الاختبار .

٧-تتطلب النظرية التقليدية نماذج متكافئة تماماً للاختبار الواحد وهذا مطلب صعب في التطبيق العملي ، مما يقلل من قيمة النتائج المبنية على نظرية القياس التقليدية.

٨-تساوى أخطاء القياس لجميع الأفراد الذين يطبق عليهم الاختبار ، ويُعد هذا شيء غير مقبول حيث نلاحظ في بعض الأحيان أن الأفراد يكون أداؤهم في الاختبار أكثر اتساقاً عن غيرهم من الأفراد ، وأن هذا الاتساق يختلف باختلاف مستوى قدرة الأفراد أو مستوى القدرة التي يقيسها الاختبار . ولهذا فإننا ربما نتوقع أن الأفراد ذوى القدرة المرتفعة يكون أداؤهم في صورة متوازية من صور الاختبار أكثر اتساقاً من أداء الأفراد ذوى القدرة المتوسطة .

٩-تأثر الدرجة الكلية للفرد في اختبار ما بمفرداته ، حيث تكون درجة الفرد عندما يختبر بمفردات سهلة أعلى من المفردات الصعبة فلا يمكن تقدير قدرته فيما تقيسه هذه المفردات تقديراً دقيقاً لذا تختلف نتيجة القياس باختلاف الاختبار المستخدم .

١٠-الدرجة الحقيقية أو النتيجة ليست سمة مطلقة للمفحوصين ، وإنما يُعبر عنها من خلال الدرجة الملاحظة للأداء على الاختبار . كما إن قدرة الفرد تعرف فقط كشرط من شروط الاختبار .

١١-هناك عيب رئيسي للنظرية الكلاسيكية في القياس وهو صعوبة المفردة وتمييز المفردة يعتمد على طبيعة وخصائص عينة الأفراد المستخدمة . بمعنى أن معاملات صعوبة المفردات تتغير بتغير مستوى درجات السمات أو القدرات ،

وبذلك يجعل الاستفادة من نتائج الاختبارات تقتصر على العينة المستخدمة التي استمدت منها هذه النتائج أو عينة مماثلة لها .

١٢-اقتصار الاختبارات التقليدية على هدف واحد فقط من أهداف القياس (التمييز بين الأفراد في حالة جماعي المرجع والتحصيل أو الاكتساب في حالة محكي المرجع) وإغفال أهداف أخرى للقياس كتقدير مدى النمو خلال فترة زمنية معينة وتقدير التغير الحادث في مستوى القدرة .

١٣-أظهرت أساليب القياس الكلاسيكية عجزاً في مواجهة كثير من المشكلات مثل تصميم الاختبارات باستخدام الحاسوب وتفسير نتائجها ، وبناء الاختبارات المفصلة ، والاختبارات المحكية ، وكذلك تصميم بنوك الأسئلة ومعادلة صور اختبارية مختلفة تقيس المحتوى .

١٤-عدم إمكانية المقارنة بين تحصيل الأفراد في حالة استخدام اختبارات تحصيلية مختلفة .

١٥-تعتمد معالم المفردة (الصعوبة والتمييز) على عينة الأفراد التي طبق عليها الاختبار وعليه يكون مستوى الصعوبة منخفض إذا كانت قدرة الأفراد منخفضة والعكس ، وكذلك يكون مؤشر تمييز الاختبار منخفض إذا كانت العينة متقاربة في مستوى القدرة .

١٦-عندما يتم تطبيق اختبارين يقيسان سمة ما على فرد ما نجد أن درجة هذا الفرد تختلف على الاختبارين، وهذا يعنى أن درجة الفرد على الاختبار تعتمد على مفردات الاختبار وليس على قدراتهم على السمة المقاسة.

١٧-يقوم الثبات في النظرية الكلاسيكية على أساس حساب خارج قسمة التباين الحقيقي مقسوماً على التباين الكلى للأفراد على مفردات الاختبار لذلك فهو يتعلق بالأفراد أكثر منه بالاختبار .

كل هذا النقد وأوجه القصور موجه لنظرية القياس التقليدية ، وحيث إن النظرية التقليدية للقياس تؤدى لبناء اختبارات غير مرنة ، فقد وجه المتخصصون في القياس جهودهم لوجود نظام قياس أكثر موضوعية يركز على انتقاء المفردات الاختبارية بشكل أفضل ، ويسمح بإضافة أو حذف مفردات إلى الاختبار دون أن يتأثر الاختبار ككل ، وقد أدت هذه الجهود إلى ظهور ما يطلق عليه نماذج السمات الكامنة.

ثانياَ : نظرية الاستجابة للمفردة (نظرية السمات الكامنة) Item Response Theory (IRT)

يرجع الفضل في وضع الأسس النظرية لنظرية الاستجابة للمفردة إلى عالم القياس النفسي والتربوي لورد (1968) الذى قدم بعض ملامح التطوير النظري المبدئي للنظرية حيث ظهرت هذه النظرية للتغلب على العديد من المشاكل التي

تعانى منها النظرية الكلاسيكية ومن ثم بدأ الاهتمام بنظرية الاستجابة للمفردة وازدادت شعبيتها . وظهور نظرية الاستجابة للمفردة في تصميم وبناء وتقييم الاختبارات جعلت من الممكن الحصول على سمات المفردة بند ودرجات القدرة ، كما أنها تعتبر مهمة جداً في مجال التعليم والاختبارات التعليمية كما يتم من خلالها إجراء العديد من البحوث لحل العديد من مشاكل التعليم ، كما يطلق عليها نظرية السمات الكامنة .

تعد نظرية الاستجابة للمفردة تطوراً حديثاً وهاماً في مجال القياس النفسي والتربوي، وذلك لأنها قدمت العديد من الحلول المقنعة للكثير من قضايا القياس مثل بنوك الأسئلة والاختبارات التكيفية ومعادلة الاختبارات وغيرها. وفى الوقت الحاضر تحظى هذه النظرية باهتمام كبير في تقييم الاختبار واختيار المفردات ومعالجة مسألة التحيز وكذلك من أهم مميزاتها ثبات معلمة المفردة ، إلا انها تحتاج لعينة كبيرة لتقدير المعلمات .

تقوم هذه النظرية على فكرة رئيسية وهى استخدام نموذج رياضي للتنبؤ باحتمال نجاح الشخص على المفردة اعتماداً على قدرة الشخص وصعوبة المفردة فهي تدرس العلاقة بين سمات الأفراد وتأثيرها على بنود الاختبار وكذلك تدرس العلاقة بين تقديرات قدرة الشخص ومستوى صعوبة المفردة ، حيث أن هناك سمات كامنة لا يمكن أن تقاس مباشرة لكنها تؤثر مباشرةً على استجابة المفردات

والتي يفترض أنها تحدد هذه الاستجابات . والسمة بعد كمي يمكن أن يحدد عليه مواضع الأفراد. وتهدف هذه النظرية إلى وصف العلاقة بين خصائص الأفراد (مثل القدرة) وخصائص المفردات (مثل الصعوبة) وبين أنماط الاستجابة التي يقدمها الأفراد على هذه المفردات .

تقوم هذه النظرية على افتراض أن تقدير القدرة الكامنة للمفحوصين مستقلة تماماَ عن محتوى الاختبار والعلاقة بين احتمال الإجابة بشكل صحيح وقدرة المفحوصين يمكن أن تصاغ بطريقة مختلفة اعتماداً على طبيعة الاختبار، وأن المفردات تقيس قدرة كامنة . كما أنها تفترض أن كل الأفراد يستجيبون لمفردات الاختبار تبعاً لقدر القدرة التي يمتلكها وبسبب ذلك سميت بنظرية السمات الكامنة ، ويعنى هذا أن الفرد ذا المستوى المنخفض يزداد احتمال فشله في الإجابة عن المفردة ، والفرد ذا المستوى المرتفع يزداد احتمال نجاحه في الإجابة عن المفردة .

وتفترض هذه النظرية أيضاً أنه يمكن التنبؤ بأداء الأفراد أو يمكن تفسير أدائهم في اختبار نفسى أو تربوي معين بناءً على خاصية أو خصائص معينة مميزة لهذا الأداء تسمى السمات . وتفترض أيضاً نماذج رياضية للسمات الكامنة، حيث تفترض أن القيمة الاحتمالية لاستجابة مفحوص لمفردة اختبارية دالة لكل من قدرة الفرد التي يفترض الاختبار قياسها، وخصائص المفردة التي يحاول الفرد الإجابة

عنها. ومن ثمّ، فإن الدالة المميزة لخصائص المفردة تحدد العلاقة بين مقدار هذا الاحتمال وبين البارامترات المناظرة لقدرة المفحوص وخصائص الفقرة .

وتوضح هذه الدالة المنحنى المميز للمفردة الذى يعد البنية الأساسية لنظرية الاستجابة للمفردة ويعد مدى قدرة الفرد من سالب ما لا نهاية إلى موجب ما لا نهاية ، أو نحدها من −٣ إلى +٣ ومن الممكن أن تتجاوز القيم هذا النطاق .وهناك نوعان من الخصائص الفنية أو التقنية لمنحنى الاستجابة للمفردة : الأول يطلق عليه صعوبة المفردة وتعنى أن هناك مستويات مختلفة من القدرة ، فمثلاً إذا مثَّل المدى مستوى مرتفع من القدرة كانت المفردة أكثر صعوبة حيث إنها تقيس مستوى عالياً من القدرة ، أما عندما يمثل المدى مستوى منخفض من القدرة فإن المفردة تكون سهلة . أما الخاصية الثانية فهي التمييز ويعرف بأنه ميل المنحنى المميز للمفردة عند مركزه فنجد أن انحدار المنحنى يبدأ من أفضل المفردات قدرة على التمييز إلى الأقل قدرة على التمييز .

أسس لنظرية الاستجابة للمفردة :

١−وظيفة الاستجابة للمفردة Item Response Function:

هي دالة رياضية تصف العلاقة بين موقع الفرد على متصل القدرة واحتمال استجابته على مفردة معينة وتقوم على سمة تؤثر تأثير مباشر على الاستجابة للمفردة ، والمفردات الأكثر قدرة على التمييز تكون أكثر قدرة على التفرقة بين

الأفراد حول صعوبة المفردة . حيث إن الهدف من دالة الاستجابة للمفردة هو تحديد كل مفردة من حيث مستوى الصعوبة للمفردة وقدرة الفرد وكذلك من حيث معامل التمييز وأيضاً معامل التخمين.

٢- دوال المعلومات Information Function:

ويعتبر مفهوم دوال المعلومات من المفاهيم الأساسية فى نظرية الاستجابة للمفردة فمن خلاله يتم تحديد أي المفردات أكثر دقة قياسياً عند مستويات محددة من القدرة . ويمكن للمفردات المختلفة أن توفر معلومات عن سمة كامنة معينة حيث إن المفردات السهلة هي الأفضل للتمييز بين الأفراد منخفضي السمة في حين المفردات الصعبة نسبياً هي الأفضل للتمييز بين الأفراد مرتفعي القدرة .

ويوضح هامبلتون وآخرون ١٩٩١ و فيكميروفتش ٢٠٠٩ أن مفهوم دالة المعلومات من المفاهيم التي تساعد المهتمين بالقياس التربوي في تحديد قدر المعلومات التي تقدمها مفردة ما عن فرد ما ، وبهذا يمكننا أن نحدد أي المفردات الاختبارية تقيس المتغير المراد قياسه (السمة) بدرجة أفضل عند مستويات محددة للقدرة. في حين ترى هند عبد المجيد ،٢٠١١ أن دالة المعلومات هي دليل يشير إلى مدى مستوى القدرة التي يكون فيها الاختبار أو الفقرة أكثر فائدة للتمييز بين الأفراد ، وتؤدى دالة معلومات الاختبار دوراً مهماً في هذه النظرية، إذ يمكن من خلالها تحديد الخطأ المعياري في التقدير.

بشكل عام تختلف كمية المعلومات بشكل كبير تبعاً لاختلاف مستوى القدرة التي يمتلكها الفرد ، فكلما كان المنحنى المميز للمفردة سريعًا زادت قيمة المعلومات التي تقدمها المفردة.

وفقاً لهامبلتون وآخرين هناك أثنين من المسلمات الأساسية لنظرية الاستجابة للمفردة وأضاف إليهم شوقي أحمد واحدة :

١- أداء المفحوصين على الاختبار يجعل من الممكن التنبؤ به بواسطة السمة أو القدرة الكامنة.

٢- العلاقة بين أداء المفحوصين على المفردة والقدرة التي يمكن أن تؤثر في أدائه يمكن تسميتها أو وصفها ب " دالة طردية " يطلق عليها دالة خصائص المفردة أو منحنى سمة المفردة ، ويمكن تمثيلها بدالة رياضية

٣- يختلف أداء الأفراد طبقاً لاختلاف كمية القدرة التي يمتلكونها .

<u>تطبيقات نظرية الاستجابة للمفردة :</u>

١-تحليل وتقنين الاختبارات: أصبح من الممكن مقارنة قيم القدرات الفردية عند التقنين بدلاً من الاعتماد على التوزيع التكراري للعينة التي قُنن عليها الاختبار . أي أنه أصبح من الممكن إعداد بنوك من الأسئلة التي حددت مؤشراتها والاستفادة منها في تقنين الاختبارات.

٢-الموازنة أو المعادلة بين الاختبارات : يحدث في كثير من الأحيان مقارنة شخصين أو أكثر بناء على درجاتهم الخام في نماذج مختلفة من اختبار معين ، وهذا فيه نوع من الظلم ولهذا فإن النظرية الحديثة توفر أسلوب أفضل للمقارنة فهي لا توازن الدرجات الخام وإنما يكون الهدف هو معرفة موقع القيمة الرقمية لقدرة الفرد على مقياس القدرة .

٣-بناء الاختبار : يمكن تحديد نوع الاختبار الذى نريده لقياس مستوى محدد من القدرة تم تحديده سلفاً.

٤-الاختبارات التكيفية أو المفصلة : وهى اختبارات فردية يعطى فيها الفرد الأسئلة المناسبة لقدراته .

٥-الكشف عن التحيز في الاختبار : فقد وفرت نظرية الاستجابة للمفردة أفضل طريقة للكشف عن تحيز الاختبار .

بنك الأسئلة ITEM BANK:

وعند الحديث عن الاختبارات التحصيلية لابد من طرح فكرة بنوك الأسئلة ، وقد استخلصت فكرة استخدام بنوك الأسئلة من قبل العديد من المدارس بغرض تقييم

الامتحانات حيث أن بنوك الأسئلة تغطى مدى واسع من المحتوى موضوع الدراسة بحيث نجد أن جميع البنود تفى معايير الموضوعية حتى يمكن تطوير البنك في السنوات التالية .

والفكرة التي تقوم عليها بنوك الأسئلة تتمثل في تخزين كم هائل من الأسئلة في مختلف مجالات المعرفة ، مع تصنيف هذه الأسئلة في ذاكرة الحاسب طبقاً لخصائص عديدة ، بحيث يكون لكل سؤال خصائصه الإحصائية الناتجة عن تطبيقه في الميدان لحساب هذه الخصائص ، ويتم عمل تدريج لهذه الأسئلة طبقاً لمستويات الصعوبة متحررة من طبيعة خصائص أفراد العينة ، وكذلك معادلة هذه الأسئلة لإمكانية عمل المقارنات بين نتائجها من حيث المستوى ، ويعتمد ذلك على برامج إحصائية وبرامج تخزين واستدعاء باستخدام أحد نماذج الاستجابة للمفردة .

- **مفهوم بنك الأسئلة :**

أصبح مفهوم بنك الأسئلة من المفاهيم الرائجة في الميادين التربوية وفي أوساط المهتمين في القطاعين الحكومي والخاص ولهذا سوف نعرض عدة تعريفات لبنك الاسئلة من منظور تاريخي .

يعرف ويلموت بنك الاسئلة بأنه عبارة عن مجموعة من البنود التي يتم إيداعها في البنك وذلك حتي يمكن بسهولة عملية السحب من خلاله . وهذه البنود التي يتم إيداعها ليس مجرد مجموعة بنود يتم إيداعها بل لابد من تحديد جودة هذه البنود من

حيث الهدف والمحتوى والأبعاد الإحصائية، وعرفه ديفيد وآخرون أنه مخزون الاختبارات المقننة التي بنيت وفق مواصفات ومعايير قياسية ، وتم ترتيب وتصنيف الأسئلة في صورة تمكن من معرفة الهدف السلوكي الذي يقيسه والمستوى المعرفي الذي يتضمنه.

اطلق جيفورى قديماً على بنك الاسئلة العديد من المصطلحات منها مستودع البنود ، تجمعات البنود ،مجموعات من البنود ، مكتبات من البنود وغيرها إلا أن مفهوم بنك الأسئلة أعم وأشمل حيث أنه يعد بمثابة بنك يمكن من خلاله القيام بعمليتي السحب والإيداع من خلاله وهكذا تتكون بنوك الاسئلة من عشرات الآلاف من البنود والتي تستخدم لأغراض التقييم والممارسة والتشخيص. كما يعرفه ملمان و جاسون على أنه مجموعة كبيرة نسبياً من الأسئلة الاختيارية التي يمكن الوصول إليها بسهوله وينتج منها مجموعة من الأسئلة غير المقيدة نسبياً حيث يتم توفير عدد من البنود يفوق العدد المستخدم في الاختبار بأضعاف كثيرة

بينما تعرفه شادية عبد العزيز ،١٩٩٦ بأنه رصيد كبير نسبياً من البنود الموضوعية الجيدة لمحتوى دراسي معين في تدريج واحد مشترك يشكل شبكة تكُون فيما بينها نسيجاً من الاختبارات المدرجة على ميزان تدرج واحد وتغطى مدى واسعاً من المتغير موضوع الدراسة ويمكن استخدام هذه البنود مرات متعددة ، كما يمكن إضافة مفردات جديدة يتم معايرتها وتدريجها بنفس وحدة القياس وعلى نفس تدريج

مفردات البنك، ويتفق معها. ويعرفه صلاح علام على أنه مجموعة من بنود الاختبار يكون لها خصائص سيكومترية مميزة ومعلومة ، ويجب أن تودع هذه البنود في البنك بطريقة تيسر عملية السحب.

بينما يعرف محمد حسين أن بنوك الأسئلة مجموعة من البنود المفهرسة والمنظمة بطريقة تشبه فهرسة وتنظيم الكتب مع الاخذ بالاعتبار الخصائص السيكومترية لهذه البنود (مثل الصعوبة والتمييز والثبات والصدق)، ويتفق معه السيد محمد أبو هاشم بأنه توفير عدد كبير من البنود في محتوى دراسي معين وهذه البنود لها خصائص سيكومترية محددة وتستخدم في بناء اختبارات تحصيلية متنوعة طبقاً لما تستهدفه عملية التقويم حيث يعد نوع من بنوك المعلومات.

ويوضحه عادل يوسف، ٢٠٠٧ بأنه مجموعة ملفات تضم الأنواع الملائمة من البنود والتي يتم تصنيفها وترميزها تبعاً للمواصفات الدراسية والمستوى التعليمي والأهداف التعليمية التي يتم قياسها .

وتعرفه كاثرين وآخرون بأنه مجموعة من البنود المصممة لقياس منهج معين ويجب أن تكون البنود دقيقة وصالحة للتقييم ، و بنوك الاسئلة وسيلة لتحسين تحصيل الطالب من المعارف والمهارات وهناك معايير لبنك الأسئلة مستمدة من معايير الاختبارات النفسية والتربوية وجمعية البحوث التربوية الأمريكية AERA والرابطة الأمريكية لعلم النفس APA .

يعرفه محمود أحمد عمر وآخرون، ٢٠١٠ بأنه المكان الذى تتجمع فيه مجموعة كبيرة من الأسئلة الجيدة المصنفة والمفهرسة في مختلف مجالات المعرفة التي تقدم للطلاب في المراحل التعليمية المختلفة وهذه الأسئلة يتم تقييمها وتدريجها وفقاً لمستويات صعوبتها كما يتم تحديد دقيق للهدف الذى وضع كل منها أجل قياسه إضافة إلى حساب خصائصها الإحصائية وهذه الأسئلة يتم معايرتها أو وضعها على تدريج واحد للقدرة بحيث تسمح بالسحب منها لتشكيل اختبارات في مجال معين تتميز بالصدق والثبات ، ويمكن القول مجملاً أن بنك الأسئلة هو مستودع لأسئلة الاختبار يضم العديد من البنود التي تستهدف قياس متغير ما والتي درجت على أعداد كبيرة من الأفراد ، ويعتمد تكوين البنك على دمج مجموعات مختلفة من الأسئلة متفاوتة الصعوبة ويجب أن تغطى كل مستويات القدرة والموضوع المستهدف

.

- **<u>أنواع بنوك الأسئلة :</u>**

١- **بنك الأسئلة المفتوح:** ويستخدم في الولايات المتحدة الأمريكية والهدف منه في عملية التقويم التكويني والتقويم التشخيصي حيث يمكن للمعلمين الاعتماد عليه في انتقاء مفردات اختبارية تناسب مختلف المواقف التعليمية

٢- **بنك الأسئلة المؤمن أو المغلق:** ويستخدم في انجلترا واستراليا ومصر حيث يكون الهدف من استخدامه لإعداده التقويم النهائي في نهاية العام أو نصف العام

ولذلك تعتبر الأسئلة سرية تماماً لا يمكن لغير المسئولين من إعداد الامتحانات الاطلاع عليها.

- **بناء بنوك الأسئلة :**

يتشكل فريق إدارة بناء البنك من مجموعة من الخبراء في مجالات المقرر المستهدف بناء بنك أسئلة لنواتج تعلمه ، والتحليل الإحصائي للبيانات في ضوء النظريات الحديثة ونظم المعلومات ، ويعتمد منهج معالجة البيانات على تقدير قدرة الفرد وليس درجته على الاختبار بحيث لا تختلف قدرة الفرد باختلاف الاختبار ولا تختلف مستوى صعوبة الاختبار ومفرداته باختلاف العينة ، ويتم استخدام ميزان تدريج له وحدة قياس ثابتة معرفة رياضياً LOGIT واختيار مفردات البنك في ضوء خصائصها السيكومترية ويتم استبعاد المفردات بالغة السهولة والصعوبة وغير المميزة بين الطلاب مختلفي القدرات ويلى ذلك تخزين مفردات البنك مع خصائصها السيكومترية .

- **مراحل إنشاء بنك الأسئلة:**

بناء بنك الاسئلة يعد نظام هندسي معقد في البداية حيث يتم تحديد خصائص البنود لإنشاء بنك للأسئلة في تطوير أساليب قياس وتقويم التحصيل الدراسي ، اقترح صلاح الدين علام تصميماً لبناء بنكي للأسئلة :

١-مرحلة التخطيط والتنسيق :

فنظام بنك الأسئلة يتطلب وضع خطة متكاملة ومحكمة قبل البدء في تنفيذ المشروع ، كما يتطلب تخطيطاً وتنسيقاً لتحديد المسئوليات والمهام ، كذلك يتطلب تهيئة المسئولين الإداريين والمربين في نظم وأساليب التقويم والامتحانات من خلال اللقاءات والمشاركات . في خلال هذه المرحلة يتم تحديد الأهداف السلوكية المراد قياسها .

٢-مرحلة التدريب والإعداد :

في هذه المرحلة ينبغي وضع خطة تنظيمية لتدريب الأطر الفنية التي ستوكل إليها مهام تحديد الأهداف والكفايات التعليمية وصياغتها بطريقة تسمح بقياس نواتج التعلم وكتابة مختلف أنواع الأسئلة والمفردات الاختبارية التي تقيس الأهداف التي يتم تحديدها ، وينبغي التأكد من امتلاكهم القدرة والمهارات الفائقة في صياغة هذه الأسئلة والمفردات .

٣-مرحلة تصميم وبناء الأسئلة :

هي خطوة تواجه صعوبات كثيرة من قبل غير المتدربين على ذلك حيث يجدون صعوبات بالغة في التفرقة بين أنواع الأسئلة الموضوعية وفنيات صياغتها ، والمستويات المعرفية التي تقيسها تلك الأسئلة وغير ذلك من مشكلات . وللتغلب على ذلك يستلزم عمل دورات تدريبية لأعضاء هيئة التدريس للتدريب على كتابة مفردات الأسئلة في كل مقرر دراسي وقد يستلزم ذلك وجود متخصص في التقويم ضمن هيئة تدريس المقرر الدراسي .

فالأسئلة والمفردات الاختبارية هي المادة الخام التي تستخدم في بناء الاختبارات ولذلك فهي من أدق المراحل وأهمها ولذلك ينبغي أن يقوم بهذه المرحلة خبراء متمرسون وعلى أعلى مستوى من الكفاية.

هناك بعض الشروط يجب مراعاتها عند تصميم الأسئلة ومنها :

١-يجب أن يكون كل سؤال مستقل عن غيره .

٢-يجب أن تكون الأسئلة واضحة وصريحة .

٣-يجب أن تتجنب الأسئلة الغموض والتعقيد .

٤-يجب ألا تكون صيغة أحد الأسئلة تحمل في طياتها الإجابة على سؤال آخر .

٤-مرحلة التجريب والتعبير :

فأهم ما يميز بنك الأسئلة احتواؤه على رصيد متجدد من الأسئلة ومفردات اختبارية ذات خصائص قياسية معلومة . فالانتهاء من كتابة الأسئلة في ضوء أهداف المحتوى الدراسي ومراجعتها لا يعنى أنها اصبحت صالحة للإيداع في البنك ، إذ ينبغي أن تخضع لعمليات مستمرة من التجريب الميداني على عينات مناسبة من الطلاب واستخدام الأساليب السيكومترية والإحصائية لتحليل البيانات المتعلقة بكل سؤال أو مفردة لتقييم فاعليتها من حيث مدى ملاءمتها للمجتمع الذى تمثله هذه العينات من المتعلمين وقدرتها على التمييز بين المستويات التحصيلية المتباينة وذلك وفق نموذج سيكومترى مناسب .

٥-مرحلة التصنيف والتخزين :

فتنظيم بنك الأسئلة يعد من الأمور المهمة إذ من الصعب أن يحقق البنك الفائدة المرجوة منه دون تنظيم دقيق وإدارة محكمة . فأساليب الإيداع والسحب من رصيد الأسئلة والمفردات الاختبارية والتطوير المستمر لمخزون البنك وعمليات الإضافة والحذف من رصيده وإجراءات مراقبة هذه العمليات ومتابعتها والإشراف عليها والتحكم في حركتها تتطلب جميعها إجراءات تنظيمية دقيقة ومحكمات وضوابط فنية كما تتطلب أن تكون الهيئة الإدارية المسئولة عن بنك الأسئلة ذات كفاءة عالية .

٦-مرحلة الاستخدام والتقييم للبنك :

فمحك فاعلية نظام بنك الأسئلة يتحدد في ضوء الاستخدام الميداني الفعلي للأسئلة والمفردات الاختبارية المودعة فيه . ومما لا شك فيه أن نتائج الاستخدام وما يترتب عليه تكون بمثابة تغذية مرتدة تعين تقييم فاعليته وإعادة النظر في مكوناته وأسلوب تنظيمه بصفة مستمرة . إذ يمكن تشكيل اختبارات كاملة بمواصفات محددة من رصيد البنك يمكن على أساسها تقويم تحصيل الطلاب في المجالات الدراسية المختلفة والإفادة من التطبيق الميداني في المراجعة المستمرة لأسئلة البنك ومفرداته سواء بالإضافة أو الحذف أو إعادة التعبير بل وربما يؤدى لإعادة النظر في المنهج المدرسي وفقاً لما تسفر عنه نتائج الاستخدام .

- **إيجابيات بنك الأسئلة :**

١-**تخزين العديد من البنود غير متوفرة** : حيث أن بنود واسئلة الاختبار المحدودة يمكن تقاسمها واستخدامها في الممارسة من قبل الطلاب كما أن انتاج العديد من البنود يقلل من الغش حيث أن تكرار التجارب في التعليم يؤدى لإتقانها ومع كفاية البنود لكل موضوع يمكن تغطية محتويات المادة.

٢-**يتم الحصول على تصنيف للبنود** : حيث يتم الاهتمام ببناء وتصنيف نظام للبنود وهذا يوضح الاهداف من البنود لتحسين الجودة.

٣-**تكوين البنود بمساعدة الكمبيوتر** : وذلك لتوفير الوقت والجهد والمال لتجميع الاختبارات وخلوها من الاخطاء حيث أن استخدام الكمبيوتر يجعل من الممكن حساب وحفظ احصاءات البند الذى بدوره يحسن الاختبارات ويتم حذف البنود الضعيفة وكذلك لعدم تكرار نفس البنود.

٤-**تحقيق الموضوعية في التقويم** : وهذا هو الهدف الأساسي لبنوك الأسئلة ، ويعتمد تحقيق هذا الهدف على مجموعة من العوامل التي يمكن أن تؤثر على بناء وتكوين بنك الأسئلة مثل فلسفة القياس ، ومدى الدقة في صياغة الأسئلة ، وتحليلها وتحديد مواصفاتها الإحصائية .

٥-**إتاحة الفرصة لقياس العمليات العقلية العليا** وذلك نظراً لاتساع المدى الذى يمكن أن يغطيه البنك من حيث المحتوى أو المستويات المعرفية المختلفة

٦-**المفردات والأسئلة المخزنة بها متقنة** : لذا يمكن الاعتماد عليها في بناء اختبارات جيدة لنظم الامتحانات .

٧-**الاقتصاد:** حيث يسمح بالاستخدام المتكرر للمفردات التي تخضع للفحص المستمر والتي تظل فاعلة لفترة طويلة .

٨-**المرونة** :حيث يسمح بنك الأسئلة بتكوين اختبارات تناسب اغراضاً محددة ، ومدى واسعاً أو ضيقاً من قدرة الأفراد وهذا يجعل إمكانية إضافة مفردات جديدة

واستبعاد مفردات قديمة خلال المراجعة المستمرة لمخزون البنك من الأسئلة والمفردات الاختبارية.

٩-**الاتساق:** حيث يتم تدريج مفردات بنك الأسئلة على مقياس واحد مشترك بصفر مشترك يكون درجة عالية من الاتساق والتماسك وبالتالي سحب اختبارات متوازية ومتكافئة تناسب قدرات الأفراد المختلفة .

١٠-**السرية** :حيث حاولت بنوك الأسئلة التغلب على مشكلة سرية الاختبارات .

١١-**السرعة والسهولة:** التي يمكن بها استخدام الحاسب في تكوين اختبارات موضوعية جيدة مسحوبة من بنك الأسئلة وهذا يتيح فرصة للتقويم المستمر

١٢-**إنشاء بنوك الأسئلة** قد يكون خطوة هامة في طريق الارتفاع بجودة التعليم الجامعي ، فإن ما توفره بنوك الأسئلة من أدوات موضوعية متعادلة القياس يتيح المقارنة العادلة بين مستوى أداء الطلاب بالأقسام والكليات والجامعات .

- **<u>أهداف بنوك الأسئلة :</u>**

الهدف استخدام بنوك الأسئلة في المؤسسات التعليمية تتمثل في ثلاثة نقاط فيما يلى :

أولاً: بالنسبة للمعلم

١- تستخدم بنوك الأسئلة في متابعة التحصيل الدراسي لذلك يتجه المعلم لاستخدامه في عملية التعليم حتى يقوم بتقويم التحصيل الدراسي للطلاب ومتابعة تقدمهم.

٢- ويستخدم المعلم بنوك الأسئلة بتقويم تحصيل كل طالب على حدة وموضع هذا الأداء على ميزان القدرة ونمط إجاباته على مفردات الاختبار . وهذا يساعد المعلم على معرفة مستوى تحصيل كل طالب في المواد الدراسية المختلفة ، وكذلك تحديد مفردات الاختبار التي لم يستطيع كل منهم الإجابة عنها ، وكذلك تستخدم في تقرير النتائج لأولياء أمور الطلاب حتى يتمكنوا من متابعة التقدم الدراسي لأبنائهم وذلك بملاحظة تحرك موضع قدرات الأبناء على ميزان القدرة نحو إتقان المادة الدراسية .

٣- أن تدرج مفردات البنك على نفس التدريج بوحدة قياس واحدة يتيح للمعلم تطبيق مجموعات مختلفة من المفردات على الطلاب بما يتناسب مع مستوى نموهم ، وذلك دون أن يفقد إمكانية المقارنة ، وبشرط أن يكون الأداء صادق .

٤- احتواء بنك الأسئلة على مفردات جاهزة ومصنفة ييسر للمعلم تكوين العديد من الاختبارات التي تناسب المستويات الدراسية المختلفة وتناسب أهداف القياس المطلوبة.

٥- تشجيع المعلمين على فحص المفردات التي يتم تضمينها في الاختبار وكذلك تشجيعهم على تبادل وجهات النظر بشأنها وإجراء التحديثات الدورية للمفردات الاختبارية .

٦- إجراء تكافؤ للاختبارات متعددة المستويات وبناء صيغ متوازية من الاختبارات التي يكون لدرجاته نفس الدلالة ، وإجراء المقارنة بين مستويات الطلاب عبر المدارس وعلى المستوى الوطني .

٧- إمكانية التقويم المستمر أثناء العام الدراسي بسهولة بالاستعانة ببنوك الأسئلة . كذلك خفض الكلفة والجهد في بناء الاختبارات وتقليل رهبة الامتحانات .

ثانياً : بالنسبة للطالب

١- تعطى الطالب فرصة أكبر للنجاح والتفوق وإبراز المواهب من حيث شمول الأسئلة لمعظم جوانب المقرر .

٢- تعرف الطالب على مستواه العلمي ومواطن قوته وضعفه أول بأول يمكن أن يولد لديه دافعية داخلية نحو مزيد من التحسن والتعلم .

ثالثاً : بالنسبة للمقررات وأساليب تقويمها

١- تسمح لقياس العمليات العقلية العليا والبعد عن أسئلة الحفظ التي يغلب وجودها في الأسئلة التقليدية .

٢- لها فائدة كبيرة في مجال تطوير المناهج من خلال تقييمها والتعرف على سلبياتها ومشكلاتها ومحاولة علاجها .

٣- توفر فرصة لعدالة التقويم في أساليبه ومستواه مما يسمح بقدر من الاطمئنان لدى الطلاب.

١- توفر مرونة كبيرة في عملية القياس حيث تسمح بتشكيل أي اختبار في أي وقت بسرعة وسهولة وتكوين عدة صور مختلفة من الاختبار الواحد دون أن يؤثر على موضوعية ودقة عملية القياس لأن الأسئلة سبق تدريجها في ضوء قدرات الطلاب ومعايرتها قبل وضعها في البنك.

المميزات

١-بافتراض وجود مجموعة كبيرة نسبياً من فقرات المقياس التي تقيس السمة يكون تقدير قدرة الفرد مستقلاً عن عينة الفقرات التي تطبق عليه ، أي أن تقدير قدرة الفرد متحررة عن الفقرات Item Free .

٢-بافتراض وجود مجتمع كبير من الأفراد تكون تقديرات الخصائص السيكومترية للفقرات مستقلة عن عينة الأفراد التي استخدمت في تقدير هذه الخصائص أى لا تتأثر معالم الفقرات بالعينة المختارة من المجتمع Person Free .

٣-توفير المصداقية لتقديرات الدقة بحيث يمكننا الحصول على تقدير إحصائي (مثل الخطأ المعياري) لقياس القدرة لكل فرد وتقدير الدقة في خصائص الفقرات أيضاً وربما يختلف التقدير الإحصائي من فرد لآخر.

٤-ثبات معلمة المفردة وهذا يعنى أنه لا يهم إذا كنا قد قدرنا معلمة المفردة باستخدام مجموعات مختلفة من العينة سواء كانت ذات قدرات عالية أو منخفضة .

٥-تقديم معلومات بشأن كل بند عن السمة الكامنة .

على الرغم من أن هذه النظرية موجودة منذ أكثر من نصف قرن إلا أنها في الآونة الأخيرة بدأت فقط في تحقيق شعبية واسعة في القياس النفسي حيث كانت البرامج الإحصائية المبرمجة أو برامج الحاسوب صعبة جداً ومكلفة بالنسبة لمعظم المستخدمين حيث إن هذه النظرية بها العديد من المزايا عن النظرية الكلاسيكية فهي مألوفة من كل الجوانب النظرية والعملية.

الأساس المنطقي لنماذج الاستجابة للمفردة الاختبارية :

يمكن إيجاز الفكرة الأساسية لنماذج الاستجابة للمفردة الاختبارية في أن هذه النماذج تحاول اشتقاق قياسات أو قيم تقديرية للسمة أو السمات التي تنطوي عليها مجموعة من الاستجابات لمجموعة من المفردات الاختبارية ، وهذه القياسات تتميز

بخصائص تفسيرية تتخطى حدود المجموع الموزون لدرجات الاختبار أي أن هذه النماذج تفسر استجابة الفرد للمفردة الاختبارية التي تنطوي على السمة التي تقيسها هذه المفردة .

الأنواع الرئيسية لنماذج الاستجابة للمفردة الاختبارية :

وقد انبثقت عن هذه النظرية مجموعة من النماذج تعرف باسم نماذج السمات الكامنة Latent Trait Models التي تهدف إلى تحديد العلاقة بين أداء الفرد في اختبار معين وبين السمات أو القدرات التي تكمن وراء هذا الأداء و تفسره ، بمعنى آخر فهي تحاول تحديد العلاقة بين قدرة الممتحنين و احتمال استجابة الممتحن بشكل صحيح .

وتعد هذه النماذج دوال رياضية احتمالية تختلف باختلاف عدد بارامتراتها أو مكوناتها ، وتنقسم هذه النماذج إلى مجموعتين رئيسيتين كما يلى :

النماذج الديناميكية Dynamic Models :

وتهتم هذه النماذج بقياس التحسن أو التغير الذى يحدث في السمات النامية المختلفة في فترات زمنية متباعدة . وبعض هذه النماذج تعتبر التغير النمائي عملية تدريجية أي أن التغير متصل ، بينما يعتبر البعض الآخر أن هذا التغير

انتقالي من حالة لآخري أي أن التغير منفصل . ومن هذه النماذج نموذج بوك (BM) ونموذج فيشر (FM) .

النماذج الاستاتيكية Static Model :

وتهتم هذه النماذج بالقياس في وقت معين ، كما تهتم بتحديد العمليات التي ينطوي عليها الأداء في الاختبارات التربوية والنفسية ، فمثلاً يمكن نمذجة الأداء في اختبار للاستعداد الدراسي أو التحصيل بواسطة هذا النوع من النماذج ومنها :

١- نماذج تحديدية DMS :

ومنها نموذج جتمان التراكمى GCM ونموذج لازرسفلد للفئات الكامنة LLCM .

٢- نماذج احتمالية PMS :

تكون هذه النماذج أكثر شيوعاً في تصميم وبناء المقاييس والاختبارات التربوية والنفسية وهم ثلاثة نماذج :

١-نموذج راش(RM) أحادي المعلم (معلم صعوبة المفردة)

يعتمد هذا النموذج على افتراض أن كل مفردة من مفردات الاختبار صممت لقياس سمة كامنة أو قدرة كامنة ، وعليه لو وجد مثلاً اختبار يضم ثلاث قدرات القراءة والكتابة والاستماع هنا يكون من الصعب استخدام نموذج راش الأحادي المعلم ، وعليه يفضل استخدام النموذج المتعدد الأبعاد.

٢-نموذج لورد ثنائي المعلم (LTPM)(معلمي صعوبة المفردة والتمييز)

يعد من النماذج الأوسع انتشاراً واستخداماً وهو مناسب للاستجابات ثنائية الدرجة (١،٠) ويعتمد بالإضافة على معلم الصعوبة معلم التمييز

٣-نموذج بيرنبوم ثلاثي المعلم (BTHPM) (الصعوبة والتخمين والتمييز)

يعتمد على احتمال إجابة الفرد إجابة صحيحة على المفردة ويعتمد على عامل التخمين ، حيث إن هذا النموذج يعتمد على ثلاث معلمات وهما معلم الصعوبة ومعلم التمييز ومعلم التخمين ، ونجد طبقاً لبكر أن المعدل الطبيعى لمعلمة التخمين ٠,٣٥ والقيم التى أعلى من تلك القيمة لا تعتبر مقبولة .

الفصل الثالث

الخصائص السيكومترية

مقدمة

عندما نتحدث عن الصدق والثبات في الاختبارات فلابد أن نتذكر بأن الاختبار أداة قياس وأن أي أداة قياس لابد أن تكون معدة لقياس خاصة معينة أو سمة معينة . فالمسطرة مثلاً معدة لقياس الطول بوحدة معينة كالسنتمتر أو جزاء منه أو مضاعفاته ، وذلك حسب الدقة المطلوبة في القياس ، فالدقة المطلوبة في قياس أبعاد القطع المختلفة المكونة لجهاز معين ربما تصل إلى جزء بالمئة من المليمتر ، وبالتالي فإن الخطأ الذي يمكن احتماله يعتبر صغيراً جداً إذا قورن بالخطأ الذي يمكن احتماله عند قياس طول الفعل في الروضة ضمن برنامج مقابلة النمو الجسمي للطفل . إلا أن الصفة المقاسة في جميع الحالات بالمسطرة حتى صفة الطول مهما تغير شكل جهاز القياس الذي يتم تشكيله عادة ليتناسب شكل الشيء المقاس ، فقد يكون قطر كره أو قطر أسطوانة من الداخل ، وبالتالي قد يستخدم الميكروميتر أو الورقية (وهى أدوات قياس يمكن أن يجدها المعلم في مختبر العلوم) ليس الهدف من هذه المقدمة التعريف بأدوات قياس الطول ، بل للتعريف بأنه نوع أداة القياس يتحدد في ضوء الغرض من القياس ، وفيما يلي شرح للخصائص السيكومترية والصدق والثبات وانواعهم والعوامل المؤثرة فيهم:

أولا :الصدق Validity

ويعد الصدق من أهم خواص القياس، ويشير مفهوم الصدق إلى الاستدلالات الخاصة التي نخرج بها من درجات المقياس من حيث مناسبتها ومعناها وفائدتها وتحقيق صدق القياس معناه تجميع الأدلة التي تؤيد مثل هذه الاستدلالات .

ولذلك يشير الصدق إلى مدى صلاحية استخدم درجات المقياس في القياس بتفسيرات معينة . فإذا كان المقياس أو الأداة اختياراً يستخدم لقياس تحصيل أفراد العينة وجب أن يكون تفسير الدرجات تمثل مجال التحصيل الدراسي الذي يقيسه الاختبار .

وإذا استخدمت نتائج الاختبار للتنبؤ بأداء أفراد العينة في نشاط قادم . وجب أن تكون تفسيراتنا قائمة على تقدير وفيق ما أمكن للأداء في المستقبل . وإذا استخدمت النتائج لقياس مستوى الفهم في القرار فإننا نود أن تكون تفسيراتنا قائمة على أدلة تبين أن الدرجات تعكس بالفعل الفهم دون أن يتداخل معها عوامل أخرى خارجية . فالصورة يرتبط أساساً بالاستخدام الخاص لنتائج القياس ، وبمدى صحة التفسيرات المفترضة لهذه النتائج .

إذن نستطيع أن نقول أن الاختبار الصادق هو الاختبار الذي يقيس ما وضع لقياسه ولهذا يشير الصدق الى الدرجة التي يمكن فيها للاختبار ان يقدم معلومات

ذات صلة بالقرار الذي سيبنى عليها فإذا أعد المعلم اختباراً لقياس قدرة الأطفال على جميع الأعداد فإن الاختبار يكون صادقاً إذا قاس القدرة على الجمع ، أما إذا قاس قدرة أخرى مثل الكتابة مثلاً فإنه يكون غير صادقاً . فمن الضروري أن يكون الاختبار صادقاً لأننا نريد أن نقيس ظاهرة معينة وليس ظاهرة أخرى غيرها . وسوف نتحدث فيما بعد عن أنواع الصدق .

وأنه مقدرته على قياس ما وضع من أجله أو السمة المراد قياسها أو قدرته على قياس ما يدعي قياسه من جوانب سلوك الافراد او صلاحية الاختبار لقياس ما وضع لقياسه ويعرف الصدق احصائيا بانه نسبة التباين الحقيقي المرتبط او المنسوب للسمة المقاسة الى التباين الكلي ويشار الى هذه النسبة بمعامل الصدق

ويورد صفوت فرج عدد من التعريفات منها :

تعريف جيولكسون: الصدق هو ارتباط الاختبار ببعض المحكات محددا بذلك ان الارتباط بمحك خارجي في شكل معامل هو مؤشر الصدق

تعريف كورتن : هو تقدير الارتباط بين الدرجات الخام للاختبار والحقيقة الثابتة ثباتا تاما ويعتمد هذا التعريف على فكرة معامل الارتباط والتباين المشترك بين الدرجات الخاصة بالاختبار وهذه الحقيقة وهو يتطلب بالتالي تصحيح قياسنا للحقيقة الخارجية من عوامل عدم الثبات.

تعريف كرونباخ : يقدر اكتمال درجة الاختبار للسمة المعينة والثقة من هذا التفسير بقدر صدق الاختبار وهو يربط بذلك بين الدرجة على الاختبار وقدرتها التفسيرية

تعريف كاتل : الصدق بمعناه الواسع باعتباره قدرة الاختبار على التنبؤ ببعض الوظائف او اشكال السلوك المحددة والمستقلة عن الاختبار والتي تعد محكا لصدق الدرجة .

ويتضح من التعريفات السابقة التكامل والتشابه بينهما فبعضها إجرائية وبعضها نظرية يجعل مؤشر الصدق هو الاتقان مع المحك وبعض التفسيرات والسمة ولكن **هذه التعريفات تلخص في أن :**

١- ان يكون قادرا على قياس ما وضع لقياسه اى ان يكون الاختبار ممثلا تمثيلا حقيقيا للقدرة التي صمم لقياسها.

٢-ان يكون الاختبار قادر على ما وضع لقياسه فقط اى ان الاختبار يقيس القدرة التي صمم لقياسها دون ان يخلط معها قياس قدرة أخرى غير مطلوب قياسها فاختبار الذكاء مثلا يجب ان يقيس الذكاء ككل ولا يقيس شيئا اخر مع الذكاء.

وقد تختلف المصطلحات مثل صلاحية الاختبار او صحته او صدقه الا انها كلها تعني معنى واحد وان كان المصطلح المشهور هو صدق الاختبار.

كما ان مفهوم الصدق مرتبط بنسبة التباين الحقيقي الخاص بالصفة التي نريد قياسها فمن المعلوم ان التباين الكلي للدرجة الملاحظة يتكون من تباين الدرجة الحقيقية وتباين أخطاء القياس ومفهوم الصدق مرتبط بالتباين الحقيقي ويحدده نسبة التباين الحقيقي المناسب والناتج عن الصفة التي نريد قياسها فقد يكون الاختبار رصد تباينا حقيقيا ولكن هذا التباين لا يعود جميعه للصفة التي نريد قياسها فقد يكون الاختبار يقيس قدرة أخرى واستطاع ان يرصد التباين الحقيقي لها أيضا وبالتالي يصبح التباين الحقيقي للاختبار مكونا من جزأين جزء يعود الى القدرة او الصفة التي نريد قياسها وهو الذي يمثل صدق الاختبار وجزء يعود الى قدرة او صفة أخرى استطاع ان يقيسها الاختبار حيث انها اختلطت وتداخلت مع الصفة التي نريد قياسها.

وهذا يعنى ان النتائج الخام لأى اختبار ينتج عنها تباين ينقسم الى تباين حقيقي مناسب يعود للصفة التي نريد قياسها وهو الذي يمثل صدق الاختبار بالإضافة الى التباين الحقيقي غير المناسب الذي يعود الى صفة اخري تداخلت مع الصفة التي نريد قياسها وهو تباين مشوش لعملية القياس ويحد موضوعيته وهو نوع من التباين الثابت والذي يعود للأداة نفسها ولا تتحكم فيه مؤشرات الثبات بالإضافة الى أخطاء القياس غير الثابتة والتي تتحكم فيها مؤشرات الثبات وبالتالي لكى

يكون الاختبار صادقا فلابد ان يقيس القدرة التي صمم من اجل قياسها فقط ولا يقيس قدرات اخري قد تتداخل معها.

٣- ان يكون الاختبار قادر على التمييز بين طرفي القدرة التي يقيسها اى ان للاختبار القدرة التمييزية التي تسمح له بالتمييز بين الأداء المرتفع والأداء المتوسط والأداء المنخفض للأفراد.

أهمية الصدق:

الصدق له أهمية كبرى في تحديد قيمة الاختبار ومغزاه فعنوان الاختبار قد لا يدل على ما يقيسه الاختبار فهو عادة ما يكون عبارة قصيرة تعرف بالاختبار بشكل عام والصدق هو الذي يقرر صلاحية أي فرض علمي او في تحديد معنى مفهوم معين . وللصدق أهمية كبرى في الكشف عن المحتوى الداخلي للاختبارات النفسية وفي الإفادة منه في الاختبار التعليمي المهني وفى التنبؤ في حياة الفرد التعليمية او المهنية توفيرا للجهد والمال والتدريب فيطمئن الفرد بانه يعمل في ميدان يتفق في ميدان مواهبه واستعداداته ومهاراته.

ونتيجة للاستخدام الضيق والمحدود لمفهوم الصدق من قبل الباحثين ولان الاختبارات والمقاييس تستخدم في مجملها لاتخاذ قرارات وإصدار احكام متنوعة كلا منه يتطلب نوعا معينا من الدراسات للتحقق من صدقه فقد اصدر علماء

القياس التابعين للجمعية الامريكية لعلم النفس (APA) والجمعية الامريكية للبحث التربوي (AERA) كتيبا ضمنوه اهم معايير الاختبارات والمقاييس النفسية والتربوية وتصنيف للقرارات والاحكام المتنوعة التي يهدف اليها القياس بصفة عامة وهى :

١- تحديد كيفية أداء الفرد في الوقت الحاضر في نطاق شامل لمواقف سلوكية تمثلها مفردات القياس تمثيلا جيداً

٢- التنبؤ بالأداء المستقبلي للفرد او تقدير مكانه في احد المتغيرات ذات الأهمية

٣- الاستدلال على درجة تملك الفرد سمة او خاصية معينة من حيث هي تكوين فرضي يبدو اثرها في سلوك الفرد او ادائه

خصائص الصدق:

١- يرتكز الصدق على عاملين هامين هما الغرض من الاختبار الذي ينبغي ان يحققه والفئة التي سيطبق عليها فإذا كان الاختبار وضع ليقيس القدرة اللغوية يجب ان يقيس القدرة اللغوية قياساً دقيقاً شاملاً ولا يقيس شيئاً غير ذلك وإذا كان الاختبار قد وضع لفئة معينة فيجب أن يطبق عليها وليس لها وليس على غيرها لأن كل فئة تختلف عن فئة أخرى من حيث النضج والخبرة وغير ذلك.

٢- الصدق صفة نوعية اى خاص بما وضع لقياسه فإذا استخدمنا اختبار ليقيس الميل والاتجاه نحو اللغة فسيكون الصدق معتدل او متوسط وإذا استخدمناه للتنبؤ بنجاح الطلبة في الجغرافيا للعام القادم فإن النتائج ستكون ضعيفة

٣- الصدق صفة نسبية او متدرجة فصفة الصدق ليست صفة مطلقة ولهذا لا يصح القول ان الفحص يتصف بالصدق التام او لا يتصف بالصدق ابدا ولكن تحدد درجة اتصاقه اما بالارتفاع او الاعتدال او الانخفاض

٤- الصدق صفة تتعلق بنتائج البحث فمن الشائع القول ان هذا الاختبار صادق والمعنى الادق هو القول بان نتائج الاختبار هي صادقة أي ان احكامنا ستكون صادقة متى ما خلت من الخطأ

٥- يتوقف صدق الاختبار على ثباته أي انه اذا اعيد اجراء الاختبار في نفس الظروف فإنه سيحقق نفس النتائج اما اذا اختلفت النتائج في كل مرة فهذا يعنى ان الاختبار غير ثابت وغير صادق أي ان الصدق يتوقف على ثبات الاختبار والعكس غير صحيح اى لا يتوقف الثبات على الصدق فعلى سبيل المثال قد يكون مؤشر الميزان غير سليم فيعطى نفس الوزن لاوزان مختلفة اذاً فالنتيجة ثابته ولكنها غير صادقة.

طرائق التحقق من صدق الاختبار Validity

يمكن القول بوجه عام أن جميع طرائف تحديد صدق الاختبارات تبحث عن العلاقة التي تربط أداء الأفراد في الاختبار وبعض الحقائق والمعلومات التي تتصل بالعملية النفسية موضوع الاختبار، والتي (أي هذه المعلومات والحقائق) يمكن ملاحظتها مستقلة، وهناك كثير من الطرائق التي تستخدم في هذا الصدد ولها مسميات كثيرة، إلا أننا سوف نتحدث عن الأنواع الأربعة من الصدق التي وصفها كتيب التوصيات الفنية للجمعية الأمريكية لعلم النفس.

١- صدق المحتويات Content validity:

يمسى هذا النوع من الصدق بالصدق المنطقي، ويحسب هذا الصدق على فحص الاختبار وتحليل أسئلته لمعرفة مدى تمثيلها للسلوك الذي يقيسه الاختبار، وللتأكد من أن الأسئلة تغطي جميع جوانب السلوك. وغالبًا ما يستخدم هذا النوع من الصدق في الاختبارات التحصيلية، ويمكن أن نقول أنص دق المحتوى يعتمد على مدى فقرات المقياس تمثيلاً سليمًا للمجال الذي نريد قياسه ولذا نرعى أن يحقق هدف المقياس بهذه الطريقة يتطلب القيام بالخطوات التالية:

- تحديد المجال الذي نريد قياسه تحديدًا دقيقًا واضحًا، مع تحديد عناصره.
- بناء مجموعة من الفقرات أو الأسئلة الممثلة لهذا المجال.

فإذا أردنا مثلاً إعداد اختبار تحصيلي في اللغة العربية، فإننا لتحقيق الصدق بالطريقة المرتبطة بالمحتوى علينا لابد من تحديد المهارات اللغوية والمعرفية والفهم اللازم، بمعنى آخر تحديد أهداف تدريس اللغة العربية، ثم نقوم بتحديد محتوى منهج اللغة العربية، ونطاق أسئلة الاختبار مع الأهداف والمحتوى نرى مدى تمثيل محتوى الاختبار لمحتوى المنهج وأهدافه تمثيلاً صادقًا. وهذه الطريقة في تحديد الصدق هامة بوجه خاص في الاختبارات التحصيلية، بل أنها تعتبر أساسية بالنسبة لهذا النوع من المقاييس، نظرًا لأن تفسير درجات الاختبار ينعكس على المقرر بأكمله، أو الوحدة التي يقيسها الاختبار.

وهنا نشير أن الحكم على صدق المحتوى لا يعبر عنه بصيغة رقمية كما هو الحال في الطرف الأخرى لتحديد الصدق فأساسي الحكم على محتوى الاختبار يقوم على مدى احتواء الاختبار للجوانب النفسية التي يقيسها هذا الاختبار.

خطوات التحقق من صدق المحتوى:

يحدد الباحث أو الفاحص ما يريد قياسه تحديدا دقيقا من موضوعات المحتوى والمادة الدراسية وأهدافها ويتعرف على النامطلوب من التلاميذ من هذه المادة ثم يحلل الاختبار تحليلا شاملا بين اقسامه ويرتبها حسب أهميتها ويحدد وزن كل قسم وكل ذلك تمهيدا لصياغة فقرات الاختبار بحيث يتفق عددها مع هذه الاوزان

. فما تعرضه فقرات الاختبار هي التي تحدد حقيقة ما يقيسه الاختبار ولذا لابد من نظرة فاحصة الى كل فقرة فاذا وضع اختبار يقيس الكفاية في استخدام اللغة فكيف يمكن ان نتعرف على الدرجة التي يقيس بها هذا الاختبار ؟ والاجابة على ذلك نسأل انفسنا أولا ما المهارات والمعارف التي تؤلف الاستعمال السليم والفعال للغة والتي كانت أهدافا لتدريس اللغة بعد ذلك يتم فحص الاختبار وما يتطلبه من مهارة ومعرفة وفهم ثم لابد من اجراء المطابقة بين محتوى الاختبار وبين تحليل محتوى المادة واهداف تدريسها فترى بأية درجة تختلف نتائج هذا ومحتوى الاختبار.

أهميته وخصائصه :

يعتبر صدق المحتوى أهم أنواع الصدق عامة عندما يكون المجال محدودا او معروفا فالذكاء على سبيل المثال اقل تحديدا في تحصيل الطالب في مادة اللغة العربية وهو اجراء أولى بسيط يفيد في توليد الإحساس بالصدق ويعتبر ابسط معايير الصدق وأكثرها أولوية وتوليدا لليقين المباشر او التلقائي بصدق في الأداة وقياسها لما تنوى لها أن تقيسه وهو عملية سهلة بإمكان اى معلم تطبيقها بهدف تصميم اختبارا لمادته وهذا يقود الباحث للإشارة والتنويه بان وضع اختبارا قوميا دون مراعاة خصائص الثبات المختلفة لم يتصف بالصدق.

٢- الصدق الظاهري Face Validity

ويعنى البحث عما يبدو ان المقياس يقيسه وهو المظهر العام للاختبار او الصورة الخارجية له من حيث نوع المفردات وكيفية صياغتها ومدى وضوح هذه المفردات وكذلك يتناول تعليمات الاختبار ومدى دقتها ودرجة ما تتمتع به من موضوعية ودقة تحديد الزمن المناسب للاختبارات الموقوتة التي تعتمد على السرعة وعلى تحديد مستويات الصعوبة للاختبارات غير الموقوتة التي تعتمد على القوة وعلى نوع الأسئلة ومدى صلاحيتها للإشارة والاستجابة المناسبة من المختبرين ، وهو يعنى ما إذا كان الاختبار يبدو صادقا في نظر المفحوصين من ناحية ومن نظر الفاحصين سواء أعد أعدادا فنيا او لم يعد من ناحية أخرى . ويكون الفحص كذلك في رأي المفحوص إذا كان يقيس الغرض الذي من أجله قد تقدم للفحص اى إذا بدأ للمفحوص ان الأسئلة المستخدمة ذات علاقة بالوظيفية التي يريد قياسها

ولتوضيح الفكرة فان اختبار الصف الأول الابتدائي مثلا يمكن ان يكون صادقا ظاهريا اذا كانت جميع فقراته ذات صلة بمادة الرياضيات للصف الأول الابتدائي ولا يقيس قدرته اللغوية مثلا فإذا كانت صياغة الفقرات ذات طابع لفظى فإن هذا يخفض من الصدق الظاهري ويثور الطلبة أحيانا في بعض الامتحانات لان الفحص يبدو لهم ضعيف الصلة بما يدرسونه.

أهميته:

يجب على واضع الاختبار ان يراجع اختباره ليتأكد من توفر وجود شروط الصدق الظاهري. والصدق الظاهري يكشف عن المفردات التافهة او الضعيفة او التي لا ترتبط كثيرا بالوظيفة المراد قياسها وتأتى أهميته أيضا بانه يدخل في باب العلاقة الطيبة التي يكونها الفاحص مع المفحوص وهو من الأمور الرغوب بها فإذا بدأ الاختبار للفاحصين انه لا معنى له ولا علاقة له بالظاهرة المراد قياسها او غير مناسب او ساذج او سخيفا فيكون تعاون هؤلاء ضعيفا وان كان الصدق الحقيقي للاختبار عاليا فينطوي هذا الصدق على سهولة الإمكانيات العملية لطبعه وتصحيحه وتفسير نتائجه.

تقييمه:

ان هذا النوع من الصدق يعد اقل أنواع الصدق جودة حيث انه من الملاحظ ان هذا النوع ليس الا صدقا ظاهريا لا يلمس الا سطح المقياس فان استخدام الباحث لهذا الصدق في اختبار صدق أداة معينة تقيس شيء ما وصممها وهى غير مقننه فان الأساس المدعم لقبول المعلومات التي حصل عليها عن طريق هذه الأداة سيكون ضعيفا هذا ويعتبر البعض انه لا يكاد يكون هذا النوع صدقا على الاطلاق وان بدا اعين الناس انه صادقا هذا ويمكن ان يتحسن الصدق الظاهري

للاختبار عن طريق إعادة صياغة اسئلته بحيث تبدو اكثر ارتباطا بموقف القياس الفعلي ومع ذلك فانه لا يمكن ان يكون بديلا لأنواع الصدق الأخرى وتحسينه لا يؤدي الى تحسين الصدق الموضوعي.

٣- صدق المفهوم:

هذا النوع من الصدق مرتبط بالاختبارات النفسية بصفة عامة بحيث ان معظم السمات في علم النفس افتراضية حيث لا يوجد اتفاق على تعريف اجرائي موحد لها وصدق المفهوم يهتم بجميع الدلائل الضرورية والتي بواسطتها التعرف على مدى قدرة درجة الاختبار في ان تكون مؤشرا على السمة التي يفترض ان يقيسها الاختبار أي يتناول العلاقة بين نتائج الاختبارات والمقاييس وبين المفهوم النظري والذي يهدف الاختبار لقياسه.

وهذا النوع من الصدق يشكل المرحلة النظرية او التمهيدية في تطوير الاختبارات والمقاييس وهو موجه لخدمة الاختبار نفسه وذلك بمحاولة الانتقال من الشك في ان الاختبار يقيس السمة التي اعد لقياسها.

وكلمة المفهوم تدل على شيء قد يكون محسوس كالشجرة او مجردة كصفة الذكاء والقدرة الموسيقية وهو يتمثل في الارتباط بين الجوانب التي يقيسها المقياس وبين مفهوم هذه الجوانب ويسمى صدق المفهوم لأنه يقوم على تحديد المفاهيم او

البنى المقدمة للظاهرة المقاسة وبعبارة أخرى فهو نجاح الاختبار في قياس سمة او قدرة معينة كالذكاء او الانطواء او الانبساط ويتوقف هذا النوع من الصدق على مقدار ما نحصل عليه من معلومات عن هذه السمة وخصائصها ومكوناتها ويسمى صدق المفهوم لأنه يقوم على تحديد المفاهيم او البنى المقدمة للظاهرة المقاسة واستنتاج المفاهيم والبنى التي تتولد عن منظومة البنى المقدمة من القدرات لدى الافراد ولتصمم مقياسا يقيس هذه الصفة او القدرة يقوم الفاحص بتعريف الصفة التي يريد قياسها منطلقا في اطار نظري ثم يبدأ بتحليل المجال او الناحية التي يريد قياسها تحليلا يكشف عن عناصرها المختلفة الرئيسية ثم يحلل كل قسم ويجزئه الى اجزائه ثم يقدر النسب المئوية لكل جزء من كل قسم من هذه الأقسام وهذا ما فعله الفريد بينيه عامي ١٨٦٧ و ١٩١١ وهو اول من وضع اختبار الذكاء وقد سار على الخطوات السابقة

أهميته:

وتتمثل قيمته في كونه يتناول بالدراسة أمورا مجردة لا تطالها الأدوات القياسية المباشرة كما هو الحال في قياس الشخصية والتذوق الادبي والاتجاه العلمي والعادات وقد يستخدم لقياس مظهر من مظاهر القدرة نفسها او جانب منها وقد تكون مهارة حسابية وقد تكون قدرة رياضية كما انه يربط طريقا بين القياس وبين التجريب في علم النفس ويهتم علم النفس المعاصر كثيرا بهذا النوع من الصدق

لأنه يمكنه من التخطيط التجريبي التحليلي لقياس الظواهر كان حرام قياسها مباشرة لأنها تسبب الأذى للإنسان او بسبب عدم توفر أدوات القياس المباشرة ، ان هذا النوع من الصدق يتخطى كل الصعوبات والعقبات.

٤- الصدق التجريبي او الاحصائي Empirical Validity

وهذا النوع من الصدق يعتمد على طريقة مقارنة نتائج الاختبار الحالي اى الذي طبقه الفاحص مع نتائج اختبار اخر لنفس العينة وبنفس المواصفات ومن هذه المقارنة يوضح مدى صدق الاختبار الأول ويسمى الاختبار الاخر بالمحك وقد يكون فحص المدرس او معدل الطالب الفصلي السنوي او التراكمي او معدل التخرج.

ويشترط في المحك اتصاله بالموضوعية والصدق والثبات وهناك طريقة أخرى وهى ان تتأكد من ان الاختبار قادر على التمييز بين الافراد الذي طبق عليهم الاختبار وهم من ذوى القدرات المتفاوتة وذلك عن طريق تحليل بنود الاختبار وإزالة القنوات غير المميزة وذلك عن طريق حساب ارتباط درجة كل فقرة بالدرجة الكلية للاختبار.

وهناك طريقة اخري كأن يطبق الاختبار على جماعتين متناقضتين كجماعة الاذكياء وجماعة الاغبياء فإن لم يميز بينهما فهو غير صادق وهو يعتبر من اهم

أنواع الصدق بالنسبة للاختبار ويقيس مدى قدرة الاختبار لقياس الوقائع الخارجية والتجريبية .

وهذا النوع من الصدق يحتوى على جميع أنواع الصدق التي تستخدم فيها الإحصاء والتي يحسب لها معامل ارتباط .

٥- صدق المحك:

يرتبط هذا النوع بالأدوات التي نريد من خلالها تقدير مدى قدرتها على التنبؤ بأداء لاحق او مستقبلي وينقسم هذا النوع من الصدق الى صدق تنبؤي وصدق تلازمي ويعتمد على معامل ارتباط بيرسون لتقدير معامل الصدق في كلا النوعين.

١- الصدق التنبئوي Preditive validity

يقوم على أساس حساب القيمة التنبؤية للمقياس، أي معرفة مدى صحة التنبؤات التي نقوم بها من درجات المقياس ويقصد بهذا النوع من الصدق قدرة الاختبار على التنبؤ بنتيجة معينة في المستقبل . يطبق الباحث الاختبار ثم يتابع سلوك المفحوصين فيما بعد فإذا اتفق مستوى أدائهم على الاختبار مع سلوك المفحوصين في مجال آخر يتصل بما قاسه الاختبار، فإن لهذا الاختبار قدرة تنبئوية، فالباحث الذي يريد أن يقيس القدرة اللغوية للاتصال، فإنه يطبق الاختبار ثم يتابع ملاحظة كلام الأطفال وقدراتهم اللغوية في أثناء حديثهم وألعابهم، فإذا كانت نتاج الاختبار

متفقة مع ملاحظات الباحث لحديث الأطفال فإن الاختبار يتمتع بالصدق التنبئوي أي أنه يستطيع أن يتنبأ بسلوك الأطفال في المستقبل.

- الصدق التنبئوي مفيد في كثير من المجالات العملية حيث يطبق في التربية والصناعة والإدارة وغيرها من الميادين.

- ففي مجال التربية يمكن توزيع الطلاب إلى مستويات دراسية معينة بعد تطبيق اختبار له قدرة تنبئوية مرتفعة. وبذلك نقلل من الاهدار والرسوب حين نوجه الطلاب وفق نتائجهم على الاختبار.

- وفي مجال الصناعة والإدارة يمكن اختبار العاملين بناء على درجاتهم في الاختبار لن الاختبار الذي يمتلك قدرة تنبئوية مساعدة الإدارة على اختبار من يمكن أن يصلحوا للعمل.

٢- الصدق التلازمي: concurrent validity

يقصد بالصدق التلازمي مدى ارتباط درجات المقياس بالدرجات التي تجمع من محل آخر وقت أجزاء الاختبار وهو نوع من الصدق الذي يدل على وجود علاقة بين درجات الاختبار ومؤشر المحك التي نحصل عليها في نفس الوقت تقريباً وهو يعتبر من أوسع أنواع الصدق انتشارا والأكثر استخداما في المرحلة الوقتية الراهنة وخاصة علم النفس. أي أننا نقارن بين درجات الأفراد في المقياس ودرجاتهم على

مقياس موضوعي آخر يقيس نفس السمة تقريبًا التي يقيسها المقياس. وذلك بأن يعطي الأفراد درجات على المحل في نفس الوقت تطبيق فيه المقياس أو قبلها بقليل. فإذا كان لدينا اختبار ذكاء فإننا نختار المحل من بين اختبارات الذكاء المقننة الثابتة الصادقة وهكذا.

الشروط الواجب توافرها في المحك:

يمكن أن نشير إلى أهم الشروط الواجب توافرها في المحك فيما يلي:

- يجب أن يكون المحك صادقًا.

- يجب أن يكون المحك ثابتًا.

- يجب أن يكون موضوعيًا أي لا تتأثر درجات الأفراد فيه بشخص من يقوم بإعطاء الدرجة أو التقدير.

- يجب أن يكون المحك مستقلاً فلا تتأثر درجة الفرد في المقياس بدرجته في المحك. كما يجب ألا تتأثر رتبة الأفراد في المحك بدرجته أو رتبته في المقياس.

وهنا يجب أن ننوه أن الصدق التلازمي يشبه الصدق التنبئوي إلا أن الصدق التنبئوي يتطلب وقتًا طويلاً في حين أن الصدق التلازمي لا يتطلب مثل هذا الوقت.

مثال على الصدق التلازمي. لدينا عمال متفوقون في مصنع وعمال غير متفوقين. نطبق عليهم اختبارًا لقياس القدرة على العمل. فإذا حصل العمال المتفوقون على درجات عالية وحصل العمال الآخرون على درجات منخفضة فإننا نستطيع القول بأن الاختبار صادق.

وهناك فرق ثاني بين الصدق التلازمي والصدق التنبؤي وهو فرق زمني فالمحك في الصدق التنبؤي يكون في زمن لاحق بينما في الصدق التلازمي هو في وقت مقارب ومصاحب للاختبار الذي وضعنا له هذا المحك

يتضح من الصدق التنبئوي والصدق التلازمي أنهما يعتمدان على التجريب وبذلك يطلق عليهما الصدق التجريبي Emprical validity.

تقييمه:

وللصدق التلازمي مشكلات أهمها يتمثل في صعوبة تحديد علاقات السبب بالاثر لانه يدل على علاقة (اقتران) اكثر مما يدل على علاقة (علية) لهذا فإننا عندما نستخدمه كبديل للصدق التنبؤي فإنه سوف يتضمن صعوبات منطقية ظاهرة لان التنبؤ يحتوى ويتضمن في جوهره هذه العلاقة السببية.

٦- صدق المحكمين Trustees:

يهدف هذا الصدق إلى الحكم على مدى تمثيل الاستبانة للميدان الذي نقيسه، أي فكرة هذا الصدق نقوم في جوهرها على اختبار اسئلة الاستبانة بالطريقة الطبقية العشوائية التي تمثل مجال الاستبانة.

ويمكن حساب هذا الصدق بعرض الاختبار على عدد من المتخصصين والخبراء في المجال الذي يقيسه الاختبار، فإذا قال الخبراء أن هذا الاختبار يقيس السلوك الذي وضع لقياسه.

٧- الصدق العاملي Factorial validity:

يمتد الصدق العاملي على استخدام منهج التحليل العاملي Factor Analysis وهو منهج إحصائي لقياس العلاقة بين مجموعة من العوامل، ونبحث في هذه الطريقة لإيجاد الصدق من عوامل مشتركة تقيسها عدة اختبارات صادقة لتحدد اشتراك هذه الاختبارات في قياس تلك العوامل ، فإذا اردنا معرفة مدى صدق اختبار يقيس صفة من الصفات فإننا نقوم بتحليل هذه الصفة الى عواملها الجزئية فالقدرة العقلية يمكننا أن نحللها الى عوامل تتجزأ منها القدرة اللغوية والقدرة المكانية والقدرة العددية فعندما نصمم اختبارا فإننا نصممه لكي يقيس جميع هذه العوامل الجزئية.

ويمكن حساب الصدق العاملي وفقًا لما يلي:

يطبق الباحث مجموعة من الاختبارات على عدد من المفحوصين، ثم يحسب معامل الارتباط بين كل اختبار وسائر الاختبارات الأخرى، فإذا وجد الباحث أن هناك معامل ارتباط عال بين اختبارين منا. فإن ذلك يعني أن هناك سمات مشتركة بين الاختبارين ويمكن وضعهما تحت عامل مشترك واحد يشملها معًا.

ويمكن حساب الصدق العاملي عن طريق حساب معامل الارتباط بين فقرات الاختبار الواحد، كما يمكن حساب الصدق العاملي عن طريق حساب معامل الارتباط بين كل فقرة وبين الاختبار ككل، ومن الواضح أن الفقرة تكون صادقة إذا كان معامل الارتباط بينها وبين الاختبار الكلي عاليًا.

العوامل المؤثرة في الصدق

يمكن أن نشير إلى العوامل التي تؤثر في الصدق تأثيرًا خاصًا ولابد من أن يوضع في الاعتبار.

١- طبيعة عينة التقنية

ومن المهم لفهم طبيعة معامل الصدق. وصف عينة التقنية وصفًا مفصلاً، لأن الاختبار قد يقيس عمليات نفسية مختلفة إذا طبق على عينات تختلف في العمر أو الجنس أو المهنة أو المستوى التعليمي أو أي متغير آخر يرتبط بمفهوم الصدق. وعلى ذلك فإن الاختبار الواحد قد يكون له هدف مرتفع في التنبؤ .

بمحك معين في بعض العينات بينما ينخفض صدقه بالنسبة إلى عينات أخرى وربما يكون مقياسًا صادقًا لعمليات نفسية مختلفة في عينتين مختلفتين، ومن ذلك مثلاً اختبار العمليات الحسابية الذي قد يكون مقياسًا صادقًا للاستدلال الحسابي عند أطفال الصفوف الدنيا في المرحلة الابتدائية، بينما يقيس السرعة العددية بصدق عند طلاب المرحلة الثانوية.

٢- طبيعة العلاقة بين الاختبار والمحك:

من المعروف أن حساب معامل الارتباط بين الاختبار والمحك، إذا استخدمت فيه طريقة بيرسون أو تقريباتها المختلفة، فإن ذلك يتضمن افتراضًا جوهريًا وهو أن العلاقة بين المتغيرين خطية Linear على امتداد مدى الفروق الفردية فيهما، وبالطبع فإن الافتراض قد يكون صحيحًا في عدد كبير من الحالات، إلا أنه في بعض الحالات قد لا تكون العلاقة كذلك، ومن ذلك مثلاً حين يتطلب أحد الأعمال حد أدنى من فهم القراءة يعين العاملين فيه على قراءة بعض التعليمات وفي هذه الحالة نجد أن حالما توافر هذا الحد الأدنى فإن أي زيادة عليه في القدرة على فهم القراءة يصبح غير مرتبط بدرجة النجاح في العمل، وتوصف العلاقة بين الاختبار والمحك في هذه الحالة بأنها غير خطية non- Linear ومن أمثلة ذلك العلاقة المنحنية Curvilinear والمثلثة triangle أو العلاقة التي تدل على اختلاف التباين ومن أمثل النوع الأخير من العلاقات حين نجد أن الأداء في

اختبار الذكاء يكون شرطًا ضروريًا ولكنه ليس كافيًا للنجاح المدرسي، وفي هذه الحالة نجد التلاميذ ذوي الذكاء المنخفض يؤدون أداء سيئًا في المدرسة، ولكنا حين نفحص أداء التلاميذ من ذوي الذكاء المرتفع سوف نجد بعضهم يؤدون أداء جيد في المدرسة وبعضهم الآخر قد يؤدون أداء سيئًا لأسباب تتصل بنقص الدافعية أو فقدان الرغبة أو غيرها. وفي هذه الحالة نجد تباينًا في التحصيل بين ذوي الدرجات المرتفعة في الذكاء أكثر اتساعًا مما نجده بين ذوي الدرجات المنخفضة في الذكاء.

٣- أثر ثبات الاختبار في صدقه:

من البديهي في القياس النفسي أن الاختبار غير الثابت وغير الدقيق لا يمكن أن يكون صادقًا. لأن معنى انخفاض ثبات الاختبار أن تباينت الخطأ فيه يكون أكبر من التباين الحقيقي، وبالطبع فإن تباين الخطأ في درجات المفحوصين لا يمكن أن يرتبط ارتباطًا دالاً (أي يتجاوز مستوى الصدقة) مع أي محك، وعلى هذا كلما زاد تباين الخطأ يقل معامل الصدق، لأن الصدق هو التباين الحقيقي للاختبار .

<u>ثانياً الثبات</u>

وفيما يتعلق بمصطلح الثبات فهو يشير إلى مدى إتساق نتائج المقياس ، فإذا حصلنا على درجات مشابهه عند تطبيق نفس المقياس على نفس مجموعة الأفراد

مرتين بفاصل زمني ليس بقريب وليس ببعيد ، فإنا نستدل من ذلك على أن المقياس يتمتع بدرجة عالية من الثبات .

فإن الثبات يرتبط ارتباطاً وثيقاً بنوع التفسيرات التي نقوم بها لدرجات الاختبار أو أداة جمع البيانات ، وفى جميع الحالات التي نتطرق فيها إلى الثبات . نهتم أولاً وأخيراً بمدى اتساق النتائج أكثر من اهتمامنا بمدى صحة التفسيرات التى نستقيها من النتائج (الصدق) وأحياناً ما تكون العلاقة بين الصدق والثبات محيره للأشخاص الذين يصادقون هذين المصطلحين للمرة الأولى . فثبات المقياس لازم للحصول على نتائج صادقة ، إلا أننا نستطيع الحصول على نتائج وليست صادقة ، فقد نحصل على قياس ثابت ومتسق إلا أنه يعطينا معلومات خاطئة ويفسر تفسيراً غير سليم . مثال على ذلك إذا افترضنا أننا نستطيع قياس ذكاء الفرد من معرفة عرض جبهته ، فإننا سوف نحصل على قياس ثابت لعرض الجبهة كلما أعدنا قياس جبهة نفس المجموعة من الأفراد ، إلا أننا لا نستطيع الادعاء بأن الفرد الاعرض جبهة أكثر ذكاء . فهذا تفسير خاطئ لعملية القياس أي أنه تفسير غير صادق رغم ثباته . فالثبات إذا ضروري لأداء القياس ولكثير ليس كافياً لتحديد الصدق . أخيراً نستطيع أن نقول أن الاختبار الجيد يتصف بالثبات والاختبار الثابت هو الاختبار الذي يعطي نتائج متقاربة أو نفس النتائج إذا طبق أكثر من مرة في ظروف متماثلة . مثال على ذلك فإذا طبق اختبار الذكاء على

طالبان وحصل على درجة ١٢٠ فأن هذا الطالب يجب أن يحصل على نفس الدرجة تقريباً لو تقدم لنفس الاختبار بعد أسبوعين أو ثلاثة أسابيع .

طرائق حساب ثبات الاختبار :-

يمكن حساب الاختبار بالطرائق التالية

١- إعادة الاختبار Test Retest method

يسمى معامل الثبات الناتج من هذه الطريقة بمعامل الاستقرار أي استقرار نتائج الاختبار خلال الفترة بين التطبيق الأول والثاني للاختبار ، حيث أن هذا النوع في حقيقته هو معامل الارتباط بين مجموعتي درجات الاختبار في التطبيقين ، وذلك بإعطاء نفس مجموعة الافراد الاختبار نفسه مرتين " يطبق الباحث اختبار على عدد محدود من المفحوصين . ثم يكرر تطبيق نفس الاختبار على نفس المفحوصين بعد فترة زمنية محددة . وتحسب درجات المفحوصين على الاختبار في المرة الأولى ودرجاتهم في المرة الثانية . ثم يحسب معامل الارتباط بين درجاتهم في المرتين" .

ولا يستخدم هذا الأسلوب في استخراج الاتساق الداخلي إذ قد يكون معامل الارتباط عالياً بينما يكون معامل الثبات الداخلي منخفضاً ويفضل عادة عند حساب معامل الثبات بهذا الأسلوب ألا يكتفى الباحث بحسابه على مدى فترة زمنية واحدة بل على اكثر من فترة إذ من الممكن إعادة التطبيق بعد أسبوعين .

عيوب طريقة إعادة التطبيق:

١- تمثل الفقرات التي استخدمت في الاختبارين عينة واحدة فقط ، مما لا يكون درجات إعادة الاختبار دليلاً واضحاً عن كمية تغيير مقدار الدرجات فيمل اذا استعملنا عينة مختلفة من الفقرات او الأسئلة.

٢- إجابات المفحوص في الاختبار الثاني لا تكون مستقلة عن اجاباته في الاختبار الأول إذ قد تتأثر اجاباتهم في المرة الثانية في تذكرهم لاستجاباتهم في التطبيق الأول او من خلال المناقشات التي تجرى بين المفحوصين بعد الانتهاء من التطبيق الأول ، أو قد تتأثر بأنواع أخرى من الجهود التي يبذلها بعض المفحوصين عادة لمعرفة الإجابات الصحيحة بعد الانتهاء من التطبيق الأول مما

تساعدهم هذه الجهود في معالجة الصعوبات التي واجهوها وبالتالي ستتحسن اجاباتهم في التطبيق الثاني.

٣- قد تتداخل أخطاء القياس مع التغييرات الحقيقية في قابليات وقدرات المفحوصين عند اعادة الاختبار نتيجة للتعلم او النمو.

٤- في أحيان كثيرة بعض المفحوصين لا سيما من الطلبة لا يروق لهم استخدام هذا الأسلوب باعتباره مضيعة لوقت التعلم وارباك للنظام المدرسي لا سيما في التطبيق الثاني ، كما ان المعلمين لا يحبذون ضياع وقت الطلبة في مثل هذه الأمور، مما يؤدى ذلك الى ضعف رغبة قسم من الطلبة في الإجابة على الاختبار الثاني اكثر من الأول وبذلك يقل ثبات الاختبار.

٥- التذبذب العشوائي في أداء الفرد الواحد في الموقفين لان أداء الفرد لا يتم من خلال قوالب صارمة في كل الأوقات وبالإضافة الى التغيرات التي قد تحدث في حالته الصحية او النفسية او تغيرات الجو وما الى ذلك من مؤثرات على الفرد اثناء الاختبار مما تؤثر بشكل ما على معامل الثبات.

خطوات حساب معامل الثبات بهذه الطريقة:

١- يقدم الاختبار او المقياس الى افراد العينة ثم يصحح وتدون نتيجة او درجة كل فرد .

٢- يعاد تطبيق الاختبار على افراد العينة بعد الفاصل الزمنى المناسب وضمن ظروف مشابهة في التطبيق الأول ثم يصحح حسب قواعد تصحيح التطبيق الأول ثم تدون النتائج او الدرجات لكل فرد امام درجته في التطبيق الأول.

٣- يتم حساب معامل الارتباط بين مجموعتي الدرجات درجات التطبيق الأول ودرجات التطبيق الثاني ومعامل الارتباط هذا يمثل معامل الثبات وعلى الباحث ان يراعي طبيعة البيانات عند استخدام معامل الارتباط ، فإذا كانت متصلة يمكن استخدام " معامل ارتباط بيرسون" أما اذا كانت البيانات منفصلة فيمكن استخدام " معامل ارتباط سبيرمان للرتب " ، إلا انه في حالة وجود درجات او قيم مماثلة كثيرة أي ان نفس الدرجة تتكرر لدى عدد من الافراد فإنه لا يفضل استخدام " معامل ارتباط بيرسون " بل يفضل استخدام " معامل ارتباط كاندل " .

العوامل التى تؤثر على الاختبار في هذه الطريقة :

- **الفترة الزمنية :**

إذا كانت الفترة بين تطبيق الاختبار في المرة الأولى والمرة الثانية قصيرة فإن المفحوصين سيتذكرون بعضاً من أجزاء الاختبار وذلك تتحسن نتائجهم في المرة الثانية ، وإذا كانت الفترة بين تطبيق الأول والتطبيق الثاني طويلة وخاصة إذا كانت العينة صغيرة السن فقد يتأثر التطبيق الثاني بعوامل ترجع إلى النضج

العقلي أو الانفعالي أو الاجتماعي أو الجسمي ، مما يؤثر على النتائج ولذا يجب أن تكون الفترة الزمنية لا تقل عن أسبوعين .

- **موقف القياس :**

إن المفحوصين قد يتعلمون من الاختبار في المرة الأولى ويألفونه وبذلك يستفيدون في أثناء إعادتهم للاختبار في المرة الثانية . كما أن درجات المفحوصين في المرة الثانية تتأثر بعوامل متعددة مثل ألفة الاختبار وانخفاض التوتر وانتقال أثر التدريب مما يؤثر على النتائج .

يفضل ان تكون الفترة الزمنية قصيرة في الحالات التالية:

١- اذا كانت الظاهرة تتعلق بالنمو الجسمي والعقلي لأفراد العينة مثلاً وفي الممارسة العملية يتطلب إعادة التطبيق في إيجاد الثبات لها فترة قصيرة بعد التطبيق الأول للاختبار.

٢- عند اختيار الافراد الصغار وبسبب سرعة معدل النمو لديهم لان التغيرات التي تحدث في الأداء النسبي بعد فترة زمنية طويلة يمكن اعتبارها ضرباً من النمو والتقدم وليس مجرد تغير عشوائي ناتج الظروف التي تطرأ على الموقف الاختباري.

٣- في الاختبارات التي لا تتأثر الاستجابات في التطبيق الثاني بتذكر استجابات في التطبيق الأول لها.

يفضل ان تكون الفترة الزمنية طويلة في الحالات التالية:

١- اذا كانت الظاهرة المدروسة لا تتأثر بالنمو والتغيير السريع.

٢- في حالة كون افراد العينة كبار السن ولديهم نوعاً من استقرار النمو.

٣- اذا كانت الاستجابات تتأثر بالتذكر من التطبيق الأول للاختبار.

٢-الثبات بطريقة التجزئة النصفية :-

يعتمد هذا الأسلوب اساساً على تقسيم فقرات المقياس او الاختبار الى قسمين متكافئين ثم حساب معامل الارتباط بين درجات هذين القسمين ولذلك فان هذا الأسلوب يصلح في الاختبارات التي تكون فيها فقرات المقياس او الاختبار متجانسة اى انها تقيس خاصية او سمة واحدة وعليه فانه يقيس التجانس بين الفقرات .

ومن الواضح ان طريقة التجزئة النصفية لا تقيس التجانس الكل للمقياس لانه يقسم الفقرات الى قسمين لذلك فهو معامل ثبات لنصف الاختبار او المقياس، ولتلافي ذلك ينبغي تصحيح معامل الثبات المستخرج بطريقة حساب الارتباط بين درجاتها

باستخدام بعض الطرق الإحصائية لإيجاد معامل الثبات لكل اختبار ومنها معادلة " سبيرمان – براون " .

كما أن هذا الأسلوب يجنب الباحث بعض الصعوبات التي تواجهه في إيجاد معامل الثبات بأسلوب إعادة الاختبار وأسلوب الصورة المتكافئة ويمكن اعتباره من اكثر الأساليب عملية وسهولة كما ينبغي ان نؤكد ان تقسيم الاختبار الى جزأين لا يتم قبل تطبيقه ولا يؤثر عليه بل يقدم على شكل متكامل ومن ثم يجرى تقسيمه بعد ذلك ، وهناك عدة طرق لتقسيم الاختبار او المقياس الى قسمين أهمها:

١- القسمة النصفية : وذلك بتقسيم فقرات الاختبار الى نصفين متساويان يمثل النصف الأول المجموعة الأولى من الفقرات ويمثل النصف الثاني المجموعة الثانية من الفقرات ، بيد انه لا يفضل استخدام هذه الطريقة لاعتبارات عديدة منها ان المجيب قد لا يكون بنفس الدافعية او القدرة عند الإجابة على النصف الثاني من الاختبار.

٢- الفقرات الفردية والفقرات الزوجية: وهذه من الطرق الشائعة في التجزئة النصفية.

٣- جزءا الاختبار: رغم ان طريقة التصنيف الى فردي او زوجي من افضل الطرق التي قد تؤدي الى نصفين متعادلين في خصائصهما ومتساويان في تعرضهما لظروف الأداء الا ان بعض المقاييس والاختبارات لا تصلح هذه

الطريقة في ثباتها لا سيما اذا كانت تقدم كل مجموعة من الفقرات في وحدة زمنية معينة او يكون المطلوب من المبحوثين تقديم عدد كبير من الاستجابات لفترة واحدة في فترة زمنية معينة ثم تقدم الفقرة الأخرى وهكذا ولذلك يفضل الفقرات الى اربع أجزاء مستقلة كل جزأين في قسم واحد وقد يتكون القسم الأول من الجزأين (١-٣) والقسم الثاني (٢-٤).

٤- وهناك من يقسم الفقرات الى قسمين بعد تحديد مستوى صعوبة كل فقرة ومن ثم توزيعها بالتساوي على قسمي الاختبار كما يمكن توزيعها بناءً على مدى صلاحيتها للتمييز بين الافراد وهذا الاجراء يتطلب عادة تطبيق الاختبار على عينة استطلاعية بهدف تحديد مستوى صعوبة وتمييز كل فقرة .

٣-طريقة الصورة المتكافئة أو البديلة:

إذا اعدت صورتان متماثلتان او اكثر لاختبار واحد بنفس الطريقة فمن المحتمل ان درجات هذه الصور البديلة سوف تظهر متكافئة وحينها يتم إعطاء كل فرد من افراد العينة صورتي الاختبار، فإن الاختبار بين الصورتين سيكون معاملاً للثبات ، ويسمى معامل الثبات بهذا الأسلوب بمعامل التكافؤ حيث يدل على مدى الارتباط بين درجات الاختبار واختبار اخر متكافئ معه يعطى في الوقت نفسه او من وقت متقارب جداً وتباين الخطأ في هذا الأسلوب يدل على أخطاء عينة

المحتوى ويعد هذا الأسلوب من افضل الأساليب في حساب الثبات في الاختبارات التحصيلية في حين لا يصلح للقياس في الاختبارات او المقاييس الموقوتة .

ويمتاز أسلوب الصور المتكافئة ببعض المميزات التي تجعله افضل الطرق لحساب معامل الثبات ومنها:

١- يساعد هذا الأسلوب في خفض احتمال تأثير التدريب او التعلم والتذكر على النتائج كما في أسلوب إعادة التطبيق.

٢- اذ تم استخدام الصورتين المتكافئتين في وقتين مختلفين فإننا نضمن تلاقي بعض العيوب (إعادة التطبيق) كما يمكن ان تقيس به معامل التكافؤ والاستقرار معاً.

٣- معامل الثبات باستخدام الصور والمتكافئة فيه تشابه كبير مع معامل الاتساق الداخلي الذي يتم حسابه باستخدام التجزئة النصفية لا سيما في تطبيق الصورتين في الوقت نفسه او متقاربة جداً.

٤- تعد الصور المتكافئة من الأساليب المناسبة لتقدير مدى تمثيل عينة الفقرات للمجال الذي يقيسه الاختبار ، فكلما كان هذا التمثيل جيداً فإننا نتوقع ان ينعكس على الصورة المتكافئة ، وهنا يظهر ايضاً التجانس في أداء الفرد على الصورتين في شكل معامل الثبات.

الا ان هذه المميزات لا تمنع من وجود بعض الصعوبات في هذا الأسلوب منها، صعوبة اعداد صورتين متكافئتين تماما لها نفس المحتوى ونفس الهدف فيما تقيسه من الظاهرة المدروسة وما تحتاج من جهد ووقت وتكلفة عند اعدادهما.

ويشترط لتكافؤ الصورتين أن يتحقق ما يلي:

- أن تكون الجوانب أو الموضوعات التي يقيسها المقياس واحدة.
- أن تكون الفقرات التي تخص كل موضوع من الموضوعات متساوية.
- تماثل مستوى صعوبة الفقرات أو انحرافها المعياري.
- أن تكون صياغة الفقرات متماثلة – أن تتعادل الأسئلة في مستويات الصعوبة.
- تساوي طول الصورتين وطريقة إجرائهما وتصحيحيهما وتوقيتهما.
- تساوي متوسط وتباين درجات الأفراد في الصورة المتكافئة.

ويقدر معامل الثبات في هذه الحالة بتطبيق صورتي الاختبار في جلسة واحدة أو جلستين مختلفتين على نفس الأفراد، ثم يحسب معامل الارتباط من الدرجات في الاختبارين ، والمهم في جميع الحالات ان تكون صور الاختبار متكافئة فعلاً. أي تقيسان نفس العمليات النفسية في الأداء في الاختبار بنفس الطريقة.

مما لا شك فيه أن حاجتنا إلى الصورة المتكافئة من الاختبارات تتعدى مجرد تمديد ثبات الاختبار، فالصورة المتكافئة أو البديلة تفيد كثيرًا في الدراسات التي تجريها حول آثار عامل تجريبي معين في الأداء الاختباري وخاصة إذا تطلب الأمر قياسًا قبليًا وبعديًا.

٤-تحليل التباين:

ان طريقة التجزئة النصفية تصلح لحساب معامل الاتساق الداخلي (تجانس نصفي الاختبار) لكنها لا تقيس تجانس الأسئلة ذاتها في الاختبار ، الا ان بعض المختصين يرون ان تجانس الأسئلة لا يدخل في مجال ثبات الاختبار وانما هو خاصية من خصائص الاختبارات الجيدة حيث انه يتناول استقرار استجابات المفحوص على أسئلة او فقرات الاختبار واحداً بعد الاخر وعادة ما يسمى هذا النوع من الثبات بمعامل الاتساق داخل أسئلة او فقرات الاختبار ويختلف عن معامل الاتساق الداخلي للاختبار حيث ان معامل الاتساق داخل الأسئلة يشير الى التجانس الكلي للاختبار بالمعنى المباشر في حين يكون معامل الاتساق الداخلي للاختبار اقرب الى معنى التكافؤ كما يتمثل في الصور المتكافئة ، ومن اشهر الطرق لحساب الاتساق الداخلي للاسئلة هي طريقة (كيودر- ريتشاردسون).

العوامل التي تؤثر على ثبات الاختبار:

هناك عدة عوامل تؤثر في ثبات الاختبار أو المقياس ولذا يجب على الباحث النفسي أن يأخذها في الاعتبار عند تفسير معامل الثبات:

١- طول الاختبار (عدد الفقرات او الأسئلة):

يؤثر طول الاختبار او عدد فقراته في درجة ثباته إذ كلما تزايد عدد الفقرات كلما ارتفع وذلك لأن العدد الأكبر من الفقرات يؤدي الى الحصول على عينة اكبر من السلوك وبالتالي يكون من المتوقع أن يمثل هذه العينة بشكل مستقر العدد الأكبر من عينات السلوك أو السمة او القدرة المقاسة والقابلة للظهور في مرتي التطبيق أو في نصفي أو جزئي الاختبار ، في حين أن الفقرات القليلة في المقياس تقلل من احتمال استقرار مكونات السلوك أو الصفات المراد قياسها وبالتالي انخفاض معامل الثبات، وكلما كان الاختبار طويلاً كلما زادت فرصة أن الخطأ يلغي بعضه بعضاً وبالتالي زادت قيمة التباين مما يؤدي الى زيادة معامل الثبات .

لكن نشير هنا أن تطويل الاختبار ليس هو فقط الطرقة الوحيدة المفضلة لتحسين ثباته إذ أن المعرفة الصحيحة والتكنيك الدقيق للاختبار سيعطى معامل ثبات عالي.

٢- مدي تقارب صعوبة الفقرات ودقتها:

أن وجود فقرات شديدة الصعوبة في الاختبار لا يستطيع الجميع الافراد أو اغلبهم الإجابة عنها لا تضيف ميزة للاختبار ولا يؤثر حذفها في الدرجة التي يحصل عليها أي فرد من افراد العينة ما دامت لا تميز بين فرداً وآخر ، وبالمثل فإن وجود فقرات شديدة السهولة التي يستطيع الإجابة عنها كل افراد العينة فإنها لا تميز بين فرد وآخر ، لذلك ينبغي حذف أو تعديل الفقرات الشديدة السهولة أو الصعوبة إذ أردنا أن نرفع ثبات الاختبار لأنها لا تضيف شيئاً لقيمة ثبات الاختبار لأن معامل الثبات مبنى في الأساس على التباين الحقيقي بين درجات افراد العينة.

وعليه فإن افضل فقرات في الاختبار هي التي يكون احتمال الإجابة عليها ٥٠% من الافراد ، كما أن الفقرات التي قد تحتمل تفسيرات عديدة ومتشعبة فإنها ستؤدى الى انخفاض تباين الأسئلة وبالتالي انخفاض التباين الكلي للاختبار.

٣- تعلق الإجابة على فقرة بالإجابة على فقرات أخرى:

عند استخدام اختبار يتضمن عدد من الفقرات ترتبط الإجابة عليها بعضها ببعضها الاخر يؤدى الى انخفاض معامل ثبات الاختبار لان هذه الخاصية ستؤدي الى خفض عدد الفقرات مما ينبغي أن تكون كل فقرة مستقلة في نوعية الإجابة عليها حتى وإن كانت تقيس سمة واحدة أو قدرة واحدة .

٤- موضوعية التصحيح:

تعتبر موضوعية التصحيح من العوامل المؤثرة في ثبات الاختبار ولا سيما في الاختبارات التي تعتمد على تقدير المصحح كاختبار المقال التحصيلية واختبارات الابداع او الاختبارات الاسقاطية ، حيث أن تباين التصحيح سيؤدي الى زيادة تباين الخطأ وبالتالي الى نقصان معامل الثبات ، مما ينبغي اللجوء الى تقدير معامل ثبات المصححين ، كما ينبغي أن يتضمن الاختبار وصفاً دقيقاً لمحكات الإجابة والتصحيح وطريقة وضع الدرجة عليها.

٥- أثر تخمين المفحوص:

في الاختبارات التي تعتمد على تقدير الفقرة أو السؤال بعدد من البدائل يلجأ بعض المفحوصين عادة في حالة عدم تأكدهم من الإجابة الصحيحة الى التخمين واختيار اى إجابة منها ، وعادة ما يكون التخمين في مثل ذلك عند إعادة الاختبار او في جزء اخر من الاختبار مما يؤدي الى خفض ثبات الاختبار ويمكن معالجة اثر التخمين او الكذب لا سيما في اختبارات الشخصية .

٦- زمن الاختبار :

يؤثر الزمن المحدد للإجابة على الاختبار على ثباته فيزداد الثبات تبعاً لزيادة الزمن حتى يصل الى الحد المناسب للاختبار- الزمن – فيصل الثبات الى اعلى درجة ممكنة من جراء تأثير هذا العامل ، أما إذا كان الاختبار يتطلب وقتاً كبيراً

لانجازه فإنه ذلك سيؤدى ذلك لخفض معامل الثبات لان ذلك قد يؤدي الى الاجهاد والتعب وضعف الدافعية على الإجابة وبالتالي قد تتعرض الفقرات الأخيرة من الاختبار الى أخطاء في الإجابة ومن ثم تؤدى الى زيادة تباين الخطأ.

٧- تجانس العينة:

يؤدى التجانس الشديد في العينة من حيث الظاهرة المدروسة الى انخفاض واضح في معامل الثبات لان التباين داخل هذه العينة المتجانسة يكون منخفضاً بقدر لا يسمح بتقدير التباين الحقيقي ، في حين كلما كبر حجم العينة وتزايد عدم تجانسها كلما كانت اقرب الى التوزيع الاعتدالي وبالتالي سيصبح تباين الخطأ صغيراً مما يؤدي لارتفاع ثبات الاختبار.

٨- بعض العوامل الأخرى التي تؤثر بمقادير مختلفة في ثبات الاختبار:

كضبط موقف التطبيق ودافعية المفحوص والمؤثرات الفيزيقية والمشتتات المتعددة في موقف الاختبار والحالة الصحية والانفعالية للمفحوص مما ينبغي ان يقوم الباحث بضبط دقيق لكل هذه العوامل حتى يتمكن من الوصول الى معاملات ثبات مرتفعة .

٣- المعايير:

الدرجات الخام التي يحصل عليها الطلاب في اختبار ما وكذلك نسبة الطلاب الناجحين في الاختبار الى المجموع الكلي ليس لها في ذاتها معنى او دلالة فهى

لا تحدد مركز الشخص بالنسبة لمجموعته او بالنسبة لعينة التقنين ولهذا لابد من الرجوع الى معيار يحدد معنى هذه الدرجة او هذه النسبة المئوية وهكذا فان هذا المعيار سيدلنا على مركز الشخص بالنسبة للمجموعة هل هو متوسط ، فوق المتوسط ، اقل من المتوسط .

خصائص المعايير:

١- ان يكون للدرجة الواحدة معنى موحد من اختبار لاختبار.

٢- ان تكون هذه الوحدات متساوية بحيث ان (١٠) درجات على جزء من الاختبار تدل على نفس الشيء الذي تعنيه (١٠) درجات على جزء اخر من نفس الاختبار.

أغراض المعايير :

١- تحدد مركز الفرد بالنسبة لعينة التقنين .

٢- يمكن من خلالها مقارنة الفرد على مقياس بمركزه على مقياس اخر .

أنواع المعايير:

١- المعايير القومية

٢- المعايير المحلية

٣- معايير العمر

٤ – <u>الموضوعية :</u>

تفضل الاختبارات الموضوعية على تلك التي يلعب فيها رأى المصحح دوراً في النتائج ، وللموضوعية اكثر من معنى منها:

المعنى الأول للموضوعية : هو عدم اختلاف المصححين في تقدير الإجابات عن أسئلة الاختبار.

المعنى الثاني: هو وجوب ان يكون هناك تفسير واحد للسؤال يتيح للتلميذ الذي يعرف الإجابة ان يفهمه من السؤال ولا يصح ان يوضع السؤال بحيث يفهم بمعاني مختلفة أي يقبل التأويل غير المقصود من السؤال ومن الضروري في مرحلة اعداد الاختبار ان تجرب اسئلة على عينة من الافراد ان ندرس اجاباتهم لتحليل ما فهمه كل واحد منهم حتى تستبعد الأسئلة الغامضة التي تقبل التأويل.

٥ – <u>التقنين:</u>

ويستلزم تقنين الاختبار وضع شروط موحدة لتطبيق الاختبار على جميع الافراد ويتطلب توافر طريقة موحدة لتصحيح الإجابات وتقديرها ، فالتقنين ما هو الا توفير شروط واحدة مضبوطة بالنسبة لجميع الافراد بحيث لا يكون هناك الا متغير واحد في موقف الاختبار هو المفحوص ، وبهذه الطريقة نستطيع حين نطبق الاختبار على شخصين مثلاً ويحصلان على درجتين مختلفتين أن نقول

أن هذا الاختلاف يرجع الى اختلاف قدرة الشخصين او تحصيلهما وليس الى اختلاف الزمن المعطى لهم للإجابة او التدريب على أسئلة الاختبار الذي أتيح لاحدهما اكثر مما أتيح للآخر او لاختلاف تعليمات التطبيق من شخص لآخر ، ولهذا يجب القاء التعليمات على المفحوص بطريقة موحدة في السرعة والنغمة الصوتية ووضع مواد الاختبار بنفس الطريقة مع كل مفحوص وتوفير جو ملائم للاختبار يساعد على إيجاد التعاون بين الفاحص والمفحوص.

٦- **جودة الأسئلة**:

يحاول صانع الاختبار ان يختار عناصر او أسئلة تعطى نتائج ثابتة وصادقة ومن المبادئ لوضع الأسئلة ما يلي:

١- **السؤال الجيد غير الغامض** : أي أنه ينبغي ان يكون للسؤال تفسير واحد واذا كانت إجابة السؤال في صورة اختبار من متعدد فينبغي ان تكون الإجابات في السؤال تتضمن إجابة واحدة مقبولة.

٢- **ان يكون السؤال مميز** : وهو يعنى ان تتفق نتيجته من النتيجة العامة للاختبار ككل أي ان نسبة من يجيبون عنه إجابة صحيحة من المجموعات ذات المستوى المرتفع اكثر بكثير من نسبة من يفعلون ذلك من المجموعة الدنيا.

٣- **السؤال الجيد ذو الصعوبة الملائمة للمجموعة موضع الاختبار:** وافضل الاختبارات لقياس جميع المستويات داخل المجموعة هي تلك التي يكون متوسط الصعوبة السؤال فيها ٥٠%

المراجع:

١- أحمد أبو أسعد. (٢٠١١) . ***دليل المقاييس والاختبارات النفسية والتربوية*** ، ط٢، مركز ديبونو لتعليم التفكير.

٢- أسماء عدنان الحسن . (٢٠١٦). ***بناء اختبار محكي المرجع لقياس درجة اتقان المعلمين لقواعد بناء الاختبارات التحصيلية***، رسالة ماجستير غير منشورة ، كلية التربية ، جامعة دمشق.

٣- اعتدال غازي . (٢٠٠٣) . ***استخدام نموذج راش في تدريج مقياس للقدرة العقلية لدراسة بعض العوامل المؤثرة على دقة القياس*** ، رسالة دكتوراه غير منشورة ، كلية البنات ، جامعة عين شمس.

٤- امطانيوس نايل ميخائل . (٢٠١٦) . ***بناء الاختبارات والمقاييس النفسية والتربوية وتقنينها*** ، ط١، دار الاعصار للنشر والتوزيع.

٥- السيد محمد أبو هاشم (٢٠٠٥) . ***المشاركة في ندوة " التربية الخاصة في المملكة العربية السعودية : مواكبة التحديث والتحديات المستقبلية*** " والتي نظمتها كلية التربية بجامعة الملك سعود في الفترة من ٢٩- ٣٠ نوفمبر ٢٠٠٥ م ، وذلك بورقة عمل موضوعها " التوجهات المستقبلية للتقويم النفسي والتربوي وتطبيقاتها في مجال التربية الخاصة "٠

٦-السيد محمد أبو هاشم (٢٠٠٦) . ***دراسة مقارنة بين النظرية التقليدية ونموذج راش في اختيار فقرات مقياس مداخل الدراسة لدى طلاب الجامعة*** ، مجلة كلية التربية، جامعة الزقازيق، ٥٢ ، ١-٥٢ .

٧-أمينة محمد كاظم . (١٩٨٨أ) . ***دراسة نظرية نقدية حول القياس الموضوعي للسلوك " نموذج راش "*** ، سلسلة الكتب المتخصصة للنشر والتوزيع .

٨-إيمان عبد الكريم زيب . (٢٠١١). ***أثر طول الاختبار على تحديد درجة القطع لاختبار تحصيلي، مجلة مداد الآداب*** ، ٦ ، ٢٣٠ – ٤٠٨.

٩-بركات حمزة حسن . (٢٠٠٨) . ***مبادئ القياس النفسي*** ، الدار الدولية للاستثمارات الثقافية القاهرة – مصر .

١٠- حجاج غانم . (٢٠٠٧) . ***بحوث معاصرة في القياس النفسي وعلم النفس التربوي*** . القاهرة . عالم الكتاب للنشر والتوزيع .

١١- خماد محمد . (٢٠١٤) . ***واقع استخدام الاختبارات النفسية في عملية الارشاد*** ، مجلة الدراسات والبحوث الاجتماعية ، ٦ ، ١٨٦-٢٠١.

١٢- سالم محمد مجاهد . (٢٠١٣) . ***نحو رؤية جديدة لإصلاح نظام القياس والتقويم التعليمي في ليبيا*** ، المجلة الجامعة ، ١٥ (٢) ،٢٣٣ -٢٦٠ .

١٣- سعد عبد الرحمن . (٢٠٠٨) . ***القياس النفسي (النظرية والتطبيق)*** ،ط٥، هبة النيل العربية للنشر والتوزيع.

١٤- سوسن شاكر مجيد . (٢٠١٤) . ***أسس بناء الاختبارات والمقاييس النفسية والتربوية***، ط٤، مركز ديبونو لتعليم التفكير.

١٥- _____ . (٢٠١٤) .***الاختبارات النفسية " نماذج "*** ، ط٢ ، دار صفاء للنشر والتوزيع .

١٦- سوميه شكري محمد محمود . (٢٠١٢) . ***استخدام نموذج راش في بناء بنك أسئلة لقياس التحصيل في مقرر سيكولوجية التعلم لدى طلاب كلية التربية بالمنيا*** ، رسالة دكتوراه غير منشورة ، كلية التربية ، جامعة المنيا .

١٧- شادية عبد العزيز . (١٩٩٦) . ***استخدام نموذج راش في بناء بنك أسئلة في مقرر علم النفس التعليمي وتحديد الدرجات الفاصلة المقابلة للتقديرات الجامعية*** ، رسالة ماجستير غير منشورة ، كلية البنات ، جامعة عين شمس.

١٨- صالح حسن أحمد . (٢٠١١) . ***أسياسيات القياس النفسي في الارشاد والصحة النفسية*** ، ط١، دار الحامد للنشر والتوزيع.

١٩- صلاح الدين محمود علام . (٢٠٠٠) . ***القياس والتقويم التربوي والنفسي " أساسياته وتطبيقاته وتوجهاته المعاصرة "*** ، القاهرة ، دار الفكر العربي.

٢٠- عادل يوسف محمد خضر . (٢٠٠٧) . ***تطورات معاصرة في التقويم التربوي (بنوك الأسئلة بين النظرية والتطبيق)*** ، دار السحاب للنشر والتوزيع.

٢١- فيصل عباس . (١٩٩٦) . ***الاختبارات النفسية وتقنياتها واجراءاتها***، ط١، دار الفكر العربي.

٢٢- ليونا تايلر " ترجمة" سعد عبد الرحمن؛ محمد عثمان نجاتي . (١٩٩٨) . ***الاختبارات والمقاييس*** ، القاهرة ، ط٢، دار الشروق.

٢٣- محمد أنور محمود السامرائي ؛ أحمد محمد شاكر الخفاجي . (٢٠١٢) . ***بناء اختبار تحصيلي محكي المرجع في مادة علم نفس الخواص لطلبة أقسام العلوم التربوية والنفسية*** ، مجلة الاستاذ ، ٢٠٣ ، ٩٦٤ -١٠٠٢ .

٢٤- محمد جاسم العبيدي. (٢٠١١) . ***القياس النفسي والاختبارات*** ، ط١، دار الثقافة للنشر والتوزيع.

٢٥- محمد خليفة بركات. (١٩٥٤) . ***الاختبارات والمقاييس العقلية*** ، دار مصر للطباعة.

٢٦- محمود أحمد عمر ؛ حصة عبد الرحمن فخرو ؛ تركى السبيعي ؛ آمنة عبدالله تركى . (٢٠١٠) . ***القياس النفسي والتربوي*** ، ط١، دار السيرة للنشر والتوزيع والطباعة.

٢٧- مجدى عبد الكريم . (١٩٩٦) . ***التقويم والقياس في التربية وعلم النفس*** . القاهرة ، بدون دار ناشر .

٢٨- نادية خليفي . (٢٠١٧) . ***اختيار وتقييم الاختبارات والمقاييس النفسية*** ، مجلة السراج في التربية وقضايا المجتمع، ١ ، ١٨٦ – ١٩٨.

ثانياً: المراجع الأجنبية:

1- David ,J. (2010) . ***Likelihood Ratio Tests for Special Rasch Models*** , Journal of Educational and Behavioral Statistics , 35(6) , 611-628 .

2- David ,M & Peter ,F . (2016) . ***Quality of life : the Assessment Analysis and Reporting of Patient – reported Outcomes*** , first published 2016 , John Wiley & Sons .

3-David ,F & Thomas ,O . (1983) . ***The Development of an Item Bank In Mathematics Using the Rasch Model , Canadienne de leducation*** , 8(1) , 57-70.

4-Geofferey ,N. (1984) **. *Constructing an Item Bank using Partial Credit Scoring*** , Journal of Educational Measurement , 21(1) , 19-32.

5-Kathryn ,E & Come ,B & Esi ,M & Kevin ,A & Keuin ,P. (2008) . ***Using Item Bank to Construct Measures of Patient Reported Outcomes in Clinical Trials*** : Investigator Perceptions , Clinical Trials , 15, 23-35.

6-Kubiszyn T . & Borich G . (2000) **. *Educational Testing and Measurement*** . 6 edition , New York , John Wiley & Sons .

7-Mary ,E. (2002) . ***Systematic Assessment of Learning outcomes : Developing Multi-Choice Exams*** , Jones and Bartlett Publishers , United State of American .

8-Mohd ,N & Pisol ,M & Fattawi ,M . (2014) . ***Simplified Reliable Procedure for Producing Accurate Students Ability Grade Using Rasch Model*** . Procedia – Social and Behavioral Sciences , 112 , 1077- 1082 .

9-Nordin ,B & Ahmed ,Z & Lei ,M. (2012) . ***Examining quality of Mathematics Preminarly Analysis*** , Procedia –social and Behavioral Sciences , 69 , 2205-2214 .

10-Ronald , K & Jac ,N . (1991) **. *Advances in Educational and Psychological Testing : Theory and Applicayions*** . Khrwer Academic Publishers , Boston , 1 st.