الجامعة المستنصرية
كلية التربية
قسم الارشاد النفسي والتوجيه التربوي

مفردات مادة

الاختبارات والمقاييس النفسية

لطلبة الصف الثالث

م.د. ازهار محمود حسين

١

الفصل الاول

ـ مقدمة في القياس النفسي ـ

معنى القياس النفسي :

إن كلمة القياس مألوفة لدينا ولدى الانسان غير المختص ، ويرد ذكرها بكثرة في احاديثنا وحياتنا اليومية . وعمليات القياس مهمة في حياتنا اليومية . ورغم الاتفاق على أهمية القياس ، الا انه لا يوجد اتفاق تام بين العلماء على تعريفه . وفيما يلي بعض تلك التعاريف :

- يعني القياس وصف البيانات بأستخدام الارقام .
- القياس هو اعطاء ارقام الى اشياء او ظواهر حسب قواعد معينة .
- القياس هو ارقام تعطى بموجب قواعد لتمثل مقادير الصفات الموجودة في الاشياء .
- القياس هو جمع معلومات عن خاصية معينة في الاجسام او الكائنات او الظواهر لتحديد مقدارها .

ان التعاريف المذكورة تشير الى ان القياس عملية نحاول بواسطتها الحصول على صورة كمية رقمية للشيء او الظاهرة التي نريد دراستها . وبما ان الظواهر النفسية لا يمكن تلمسها بشكل مباشر ، لذلك نلجأ الى قياس السلوك الذي يدل عليها بعد أن نحدد ذلك السلوك بدقة . فالقلق مثلا ظاهرة نفسية موجودة لدى الناس بدرجات متفاوتة . وللحصول على قياس لها ، لابد من تعريف القلق وحصر السلوكيات او الاعراض التي تشير الى وجوده ، ثم تحديد مقدار وجود تلك الاعراض الدالة على القلق لدى الاشخاص الذين نريد قياس القلق لديهم . وهكذا بقية الظواهر النفسية الاخرى ، كالذكاء ، والأمانة ، ومستوى الطموح ، والانبساط والانطواء ، وغيرها .

ويلازم القياس مصطلحان آخران يرتبطان به ارتباطا وثيقا هما « الاختبار » Test* ، « والتقويم » Evaluation . ويعتبر الاختبار أضيقهما ، وهو غالبا ما

---

* يميز البعض بين مصطلحي « الاختبار » test ، « والمقياس » Scale . فهم يرون ان مصطلح الاختبار يجب أن يحصر استخدامه في الادوات المستخدمة لقياس الذكاء والتحصيل الدراسي . لأن نوع الاسئلة في هذه الادوات تكون الاجابة عنها اما بـ « صح » أو بـ « خطأ » .

اما « المقياس » فيرون إن تسميته ان تقتصر على الادوات المستخدمة في قياس الشخصية . وذلك لأن الاسئلة في هذا النوع من الادوات ليس فيها اجابة « صحيحة » او « خاطئة » . اما بالنسبة للكتاب الحالي فإن مصطلحي الاختبار والمقياس سوف يستخدمان بنفس المعنى .

يتضمن تقديم مجموعة قياسية موحّدة من الاسئلة لكي يجيب عنها الفرد المفحوص ونتيجة لأجابته عن هذه السلسلة من الاسئلة نحصل على « قياس » كمي رقمي يمثل مقدار وجود خاصية معينة لدى ذلك الشخص . غير ان القياس غالبا ما يتضمن معنى اوسع من ذلك . فنحن نستطيع أن نقيس بعض الخصائص دون أن نعطي اختبارات للافراد ، مستخدمين الملاحظة ومقاييس التقدير ، أو أية اداة تتيح لنا الحصول على معلومات بصيغة رقمية .

أما التقويم فإنه يعني تقدير الشيء او الظاهرة النفسية والحكم على قيمتها وهذا يعني ان التقويم يختلف عن القياس ، لأن القياس يحدد فقط مقدار ما يحتويه الشيء او الظاهرة النفسية من خاصية او صفة معينة ، بينما يذهب التقويم الى أبعد من ذلك ، اذ يتضمن اصدار حكم على قيمة الشيء او الظاهرة على اساس المقارنة بين واقع الشيء ومعيار او مستوى معين ، لنرى الى أي حد يقترب الشيء الذي قمنا بعملية قياسه من المعيار ، هل انه يساويه ، أم انه دونه أم أعلى منه .

وبناء على ذلك ، فإن التقويم يعتبر أشمل من القياس ، وان القياس هو جزء من التقويم ، لأنه يزودنا بالأدوات التي نستخدمها في عملية التقويم . ولكي يتوضح الفرق بينهما ، نضرب المثال الآتي : لنفرض ان أحد المعلمين أعطى لتلاميذه في الصف اختباراً في الاملاء يتضمن ١٠٠ كلمة . وقد وجد المعلم عندما صحح اجاباتهم ، ان عدد الكلمات المكتوبة بصورة صحيحة يتراوح بين ٢٠ ـ ١٠٠ كلمة . أي إن أضعف تلميذ قد كتب ٢٠ كلمة بصورة صحيحة فقط ، وان أقوى تلميذ قد كتب كل الكلمات الـ ١٠٠ بصورة صحيحة . اما بقية التلاميذ فقد تراوحت الكلمات التي كتبوها بصورة صحيحة بين ٢٠ ـ ١٠٠ كلمة . وهكذا فإن كل تلميذ حصل على رقم يمثل عدد الكلمات الصحيحة التي كتبها . فهناك من حصل على رقم ٤٠ ، وآخر على ٦٥ ، وهكذا .

ان هذه الارقام التي حصل عليها كل تلميذ تمثل عملية القياس . وهي لوحدها لا تعني الكثير ، لأننا لا نعرف هل ان الذي كتب ٧٠ كلمة بصورة صحيحة مثلا ، يعتبر متفوقا ام انه تلميذ اعتيادي ؟ وهنا يأتي دور التقويم لكي نتمكن من اصدار حكم على هذا التلميذ وغيره من زملائه . فماذا يفعل المعلم هنا ليتمكن من اصدار الحكم ؟ انه يلجأ الى ايجاد « معيار » يتمكن بواسطته من مقارنة رقم كل تلميذ به ، وبالتالي معرفة مدى تقدمه أو تأخره . وأبسط طريقة للحصول على « معيار » في هذا المثال هي استخراج (معدل) الكلمات المكتوبة بصورة صحيحة من قبل تلاميذ الصف . ولنفترض ان المعدل كان (٧٥) كلمة . فإن هذا يعني ان التلميذ المتوسط قد تمكن من كتابة ٧٥ كلمة بصورة صحيحة . فيصبح الرقم (٧٥) هو المعيار الذي نقارن بقية الارقام به في هذا المثال .

ولو عدنا الآن الى التلميذ الذي نجح في كتابة ٧٠ كلمة ، وقارناه بالمعيار . نجد إن رقمه يقع دون المعيار . فماذا يعني ذلك؟ انه يعني ان هذا التلميذ دون المتوسط بـ ٥ أرقام . فهو اذن بحاجة الى أن يبذل جهده ليصل الى مستوى صفه الذي يتمثل بالمعيار على الاقل . إن وصفنا لهذا التلميذ بأنه « دون المتوسط » يعني اصدارنا الحكم عليه . وهذه هي عملية التقويم .

أهميـــة القيـــاس النفســـي :

إن القياس بشكل عام موضوع مهم ليس في علم النفس فحسب . وانما في بقية العلوم الاخرى ايضا . ان العلوم على اختلاف انواعها تقوم بتوجيه اهتمامها نحو دراسة العلاقة بين المتغيرات variables في الميادين التي تبحثها . ولكي يتمكن أي علم من هذه العلوم من التنبؤ بطبيعة تلك العلاقة . لابد له من أن يقوم بقياس تلك المتغيرات . فلمعرفة العلاقة بين متغيري « ضغط الدم » و « عدد ضربات القلب » ، يقوم الباحث بقياس ضغط الدم لدى عينة كبيرة من الناس . وعدد ضربات القلب لديهم ، ثم يحاول أن يخرج باستنتاج عن طبيعة العلاقة بين هذين المتغيرين ، وقد يتوصل الى أن « عدد ضربات القلب تزداد بازدياد ضغط الدم » . اما لمعرفة العلاقة بين متغيري « القلق » و « التوافق المدرسي » . فإن على الباحث ايضا أن يقوم بقياس مقدار درجة القلق لدى الفرد . ثم درجة توافقه في المدرسة ، وقد يتوصل الى انه كلما ازدادت درجة القلق لدى الفرد هبط توافقه في المدرسة .

وهذا يعني ان المتغيرات التي يهتم بها العلم يجب أن تكون قابلة للقياس بشكل دقيق لنتمكن من دراسة العلاقات فيما بينها بصورة مفصلة .

ويتخذ القياس النفسي أهمية خاصة . لأنه يستخدم لاتخاذ قرارات محددة تتعلق بشخص معين أو جماعة معينة من الناس . فتتحكم بحريات حياتهم وتوجهه وجهة معينة . لذلك يتطلب الامر دقة عالية وخبرة واسعة في مجال القياس النفسي ـ التربوي . فالمعلم ومدير المدرسة والمرشد النفسي وعالم النفس في المعمل او في المؤسسات الصناعية . يقوم كل منهم بتطبيق الاختبارات المناسبة . ثم يوجهون الاشخاص بناء على نتائجهم في الاختبارات .

وسوف تتبين لنا أهمية القياس النفسي بصورة مفصلة لدى الحديث عن تطبيقاته في كل من التربية والصناعة والادارة والجيش والعلاج النفسي .

لمحة تأريخية عن تطور الاختبارات والمقاييس النفسية :

يعتبر الأطلاع على التطور التأريخي لأي فرع من فروع المعرفة أمرا على جانب كبير من الأهمية . ويعود السبب في ذلك الى اننا لكي نفهم الحاضر . علينا أن

نعرف كيف وصل هذا الحاضر الى صورته الراهنة . ومعرفة الماضي تعني مراجعة الادبيات التى كتبت فى ذلك الفرع من العلم ، والنظريات التي وضعت ، وأوجه الشبه والاختلاف فيما بينها ، ومن هم أبرز العلماء في ذلك العلم . وتمثل هذه المعرفة الأرضية الضرورية لمن يريد أن يتعرف على ذلك الميدان من ميادين العلم ، ولمن يريد أن يقدم اسهامات جديدة ويبدع فيه .

ومن الواضح ان طرق القياس والتقويم النفسي وأدواتها المستخدمة في هذه الايام كانت قد تطورت في الماضي ، وان فهم وظائفها وحدودها وقصورها يتم من خلال معرفة كيفية تطورها وظهورها الى الوجود . هذا بالأضافة الى ما يوضحه لنا استقراء التاريخ من حقيقة تشير الى ان العلم نتاج انساني مشترك تتضافر فيه جهود كثير من العلماء من مختلف الازمنة والأقطار ، وانه عمل تراكمي يعتمد فيه العالم على جهود من سبقوه . وسنحاول فيما يلي أن نتعرف على مسيرة القياس والتقويم منذ القديم والى وقتنا الحاضر .

يمكننا القول ان الانسان قد مارس القياس والتقويم بشكل ما من الاشكال ، وان كان في صورته البدائية ، منذ أن وجد على وجه الكرة الارضية . فعندما كان الانسان القديم يحاول اصطياد حيوان ، كان عليه أن يعرف مدى بعد ذلك الحيوان عنه ليستطيع تحديد حجم الصخرة التي يقذفه بها لأصطياده . فاذا نجح في ذلك ، فإنه يقدّر هل انه يكفيه هو واسرته وجماعته ، ام انه بحاجة الى اصطياد غيره .

كذلك عندما كان ينظر ذلك الانسان البدائي الى جماعته من حوله ، فإنه يعرف ان زميله الضخم أقوى من في الجماعة ، وان الآخر اسرعهم في العدو ، والثالث أبرعهم في الرقص . . . الخ . وهو ايضا نتيجة تناوله مختلف انواع لحوم الحيوانات يتوصل الى ان لحم الغزال مثلا هو الأشهى يليه لحم الضأن . . وهكذا .

كل هذه الامور الحياتية وغيرها تحتوي على عمليات قياس وتقويم مثل : اقوى ، واسرع ، وابرع ، واشهى . وهذا يعني ان القياس والتقويم شيء يتحسسه الأنسان ويقوم على ادراك الانسجام والتناسق في الشيء الذي نريد أن نصدر حكما عليه .

وننتقل الآن الى مرحلة اخرى في القياس والتقويم تعتبر اكثر تطوراً ، وتحتفظ لنا بها السجلات التاريخية ، ظهرت لدى بعض الامم الرائدة في مضمار الحضارة . فنلاحظ مثلا ان الصين القديمة قد استخدمت نظام الامتحانات التحريرية قبل ما يقرب من ثلاثة الآف عام كأساس للقبول والترقية في الخدمة المدنية . وكانت تلك الامتحانات تمتد لساعات طويلة . وقد استمر هذا النظام في الصين الى بداية القرن الحالي . اما لدى الاغريق فلقد كان القياس والتقويم ملازمان للعملية التربوية . وكانت الاختبارات تستخدم لتقويم مدى السيطرة على المهارات البدنيه والعقليه .

وتشير طريقة سقراط في التعليم القائمة على الاختبار ثم التعليم ، الى شبه بأسلوب حديث يعتبر من المبتكرات التربوية في الوقت الحاضر يطلق عليه « التعليم المبرمج » Programmed Learning .

اما العرب القدامى ، فلقد عرفوا معنى التقويم منذ القديم . وقد تمثل ذلك بتقويمهم للانتاج الفكري الذي كان يعبّر عنه شعرا او نثرا . فكانت الندوات تعقد في الاسواق كسوق عكاظ وفي مواسم الحج ، ويجلس أحد الخبراء المتمرسين حكما . ويأتي الشعراء والخطباء فينشدون اشعارهم ويلقون خطبهم . ثم يقوم الخبير باصدار حكمه معتمدا بذلك على معايير متفق عليها . فظهرت على هذا الاساس « المعلقات السبع » التي اعتبرت من القصائد القمم .

ثم جاء الأسلام ، والاسلام كما هو واضح ، نظام تربوي يشتمل على خطوط عريضة من المبادىء ، وعلى تفصيل شامل للأساليب التي على المرء أن يتبعها في مختلف المواقف الحياتية ، فأصبح العرب المسلمون يقوّمون سلوك الانسان بناء على مدى انطباق تعاليم الاسلام على سلوكه . ويوجد الكثير من الآيات الكريمة والاحاديث النبوية الشريفة التي تتضمن اشارات صريحة الى عملية التقويم وفق معايير الاسلام مثل : « هل يستوي الذين يعلمون والذين لا يعلمون » . و « ومن يعمل مثقال ذرة خيراً يره ، ومن يعمل مثقال ذرة شراً يره » . كما ان مراعاة الفروق الفردية امر ظاهر في الاسلام ، مثل : « لا يكلف الله نفسا الا وسعها » والفروق الفردية مبدأ اساس في عملية التقويم .

ظهور اختبارات الذكاء :

شهد القرن التاسع عشر يقظة قوية للاهتمام بالمتأخرين عقليا والمعتوهين ومعاملتهم معاملة انسانية بعد أن كانت معاملتهم تتميز بالاهمال او الاستهزاء او التعذيب في بعض الاحيان . وتبع هذا الاهتمام المتزايد بالمتخلفين عقليا الشعور بضرورة وجود معايير موحدة قياسية لتشخيص حالات التخلف وتصنيفها حسب شدتها . وكان من الضروري اولا التمييز بين المجنون insane والمتأخر عقليا mentaly retarded . وربما تكون اشارة الفيزيائي الفرنسي اسكرول Esquirol هي الاشارة الصريحة الاولى لمثل هذا التمييز . وقد وردت في مؤلف له يقع في مجلدين نشر عام ١٨٣٨ ، حيث خصص فيه لموضوع التخلف ما يزيد على مائتي صفحة . وقد جرّب اسكرول عدة اساليب لتصنيف التخلف الى درجات . وتوصل في النهاية الى ان استخدام اللغة في الاختبارات هو افضل معيار لتحديد المستوى العقلي للفرد . ومن الطريف اننا نلاحظ . حتى في ايامنا هذه . ان اختبارات الذكاء تعتمد بصورة رئيسة على المحتوى اللغوي .

اما سيكوين Seguin ، وهو الآخر فيزيائي فرنسي هاجر عام ١٨٤٨ الى اميركا ، فقد كان لاسهاماته دور مهم في تطور اختبارات الذكاء . اذ كان رائدا في تدريب المتأخرين عقليا . ويعود اليه الكثير من اساليب تدريب الحواس وتدريب

العضلات المستخدمة حاليا في مؤسسات التخلف العقلي كما ادخلت هذه الاساليب اخيرا في اختبارات الذكاء الأدائية غير اللفظية .

ولعالم النفس الاميركي كاتيل J.M. Cattell مكانة بارزة في تطور القياس النفسي . فقد انصب اهتمامه على قياس الفروق الفردية وهو أول من استخدم مصطلح « الاختبار العقلي » mental test في مقالة كتبها عام ١٨٩٠ وصف فيها عددا من الاختبارات التي كانت تطبق سنويا على طلاب الكليات لتحديد مستواهم العقلي وكانت هذه الاختبارات تطبق فرديا وتقيس : القوة العضلية ، وسرعة الحركة ، وتحسس الالم ، والدقة في الابصار والسمع ، وتمييز الاوزان المختلفة ، وقياس زمن رد الفعل ، والذاكرة وما اشبه ذلك .

وهناك اسماء اخرى قدمت اسهامات لمسيرة القياس مثل جاسترو Jastrow وكريبلين Kraepelin . وايبنكهاوس Ebbinghaus وفيراري Ferrari

وقد نشرت في فرنسا مقالة عام ١٨٩٥ كتبها بينيه Binet وهنري Henri انتقدا فيها الاختبارات الموجودة انذاك لأنها تعتمد فقط على قياس الحواس والقدرات العقلية البسيطة ، واشارا الى ان الاختبارات العقلية يجب أن تغطي وظائف اكثر تعقيدا مثل التذكر والتصور والانتباه والاستيعاب والايحاء والتذوق الجمالي . وقد أدت هذه الاتجاهات في القياس اخيرا الى ظهور مقاييس الذكاء الشهيرة لبينيه .

صرف بينيه وزملاؤه سنوات عديدة من البحث في طرق قياس الذكاء . وقد جربوا طرقا متعددة كان من بينها قياس ابعاد الجمجمة ، والوجه ، والكف ، وتحليل الكتابة . وفي عام ١٩٠٤ عين وزير التعليم الفرنسي بينيه في لجنة لدراسة طرق تعليم الاطفال المتأخرين . وقد ظهر نتيجة لذلك اول مقياس وهو المعروف بمقياس بينيه ـ سيمون Binet-Simon scale عام ١٩٠٥ ، لتحديد المستوى العقلي للاطفال .

تألف هذا المقياس من ٣٠ مشكلة (او سؤالا) تتدرج في مستوى الصعوبة . وكان مستوى الصعوبة يحدد عمليا بتطبيق الاختبار على ٥٠ تلميذا سويا تتراوح اعمارهم بين ٣ ـ ١١ سنة ، وعلى بعض التلاميذ المتأخرين عقليا والراشدين . وقد صممت الاسئلة لتغطي انواعا متعددة من الوظائف العقلية مع تأكيد خاص على الحكم judgment ، والاستيعاب comprehension والاستدلال reasoning التي اعتبرها بينيه المكونات الاساسية للذكاء . ورغم ان هذا المقياس قد احتوى على اسئلة حسية ، الا ان نسبة المحتوى اللفظي فيه كانت اكبر منها .

اما في المقياس الثاني عام ١٩٠٨ ، فقد زيد عدد الاسئلة ، وحذفت بعض الاسئلة الواردة في المقياس الاول بعد أن ثبت أنها غير مرضية ، وجمعت الاسئلة على شكل مجموعات بحيث تمثل كل مجموعة مستوى عمريا معينا . وعلى هذا الاساس وضع لمستوى عمر ثلاث سنوات كل الاسئلة التي اجتازها ٨٠ ـ ٦٠ بالمائة من الاطفال الاسوياء بعمر ٣ سنوات ، وفي مستوى عمر اربع سنوات ، وضع كذلك كل الاسئلة التي اجتازها الاطفال الاسوياء بعمر ٤ سنوات . . وهكذا الى عمر ١٣ سنة . وتقدم لنا الدرجة الكلية التي يحصل عليها الطفل في الاختبار « المستوى العقلي » mental level لذلك الطفل . اما في الترجمات والتعديلات التي اجريت على مقياس بينيه فقد استخدم مصطلح « العمر العقلي » mental age بدلا من المستوى العقلي . وظهرت مراجعة ثالثة لمقياس بينيه ـ سيمون عام ١٩١١ ، وهي السنة التي توفي فيها بينيه .

لقد جذب اختبار بينيه ـ سيمون اهتمام علماء النفس في العالم . وظهرت الترجمات والتكييفات بلغات عديدة . كما ظهرت في امريكا مراجعات عديدة له اشهرها ما قام به تيرمان L.M. Terman في جامعة ستانفورد عام ١٩١٦ وعرف بمقياس ستانفورد ـ بينيه ، واستخدمت فيه لأول مرة معادلة الذكاء Intelligence Quotient (IQ) وهي النسبة بين العمر العقلي mental age والعمر الزمني chronological age واجري تعديل ثان عام ١٩٣٧ ، وتعديل ثالث ١٩٦٠ (انظر الفصل الرابع) .

اختبارات الذكاء الجمعية :

إن مقياس بينيه ومراجعاته هي مقاييس فردية تطبق على شخص واحد فقط في الوقت الواحد . وان العديد من الاسئلة في هذه المقاييس تتطلب استجابات لفظية من المفحوص ، او معالجة يدوية لبعض المواد ، او توقيت زمن الاستجابة . ولهذه الاسباب المذكورة لم يكن من الممكن استخدامها في التطبيق الجمعي . هذا بالاضافة الى ان تطبيق مثل هذه الاختبارات يتطلب شخصا مدربا تدريبا عاليا .

وقد ظهر أول اختبار جمعي للذكاء نتيجة الحاجة العملية التي برزت في الحرب العالمية الاولى عندما ارادت الولايات المتحدة الأمريكية الاستفادة من علم النفس في الجهود الحربية . وبدأت لجنة من علماء النفس برئاسة يركس R.M. Yerkes العمل ، فرأت إن الحاجة تتطلب تصنيف ملايين المجندين بناء على مستوى الذكاء العام ، لتوجيههم الى ما يلائمهم من انواع التخصصات العسكرية . وقد أخذ هؤلاء العلماء كل ما كان متوفرا من اختبارات الذكاء ، وتوصلوا الى وضع اختبارين جمعيين عرفا بـ الفا Army Alpha وبيتا Army Beta . وكان

الاول لأغراض الأختبارات العامة ، اما الثاني فقد كان اختباراً غير لغوي ليستخدم مع الاميين والأجانب الذين لا يعرفون اللغة الانكليزية .

وبعد انتهاء الحرب العالمية الاولى بفترة قصيرة ، عدلت هذه الاختبارات لغرض استخدامها في الحياة المدنية ، واجري عليها العديد من المراجعات ، وهي لا تزال تستخدم الى اليوم . وقد اصبحت هذه الاختبارات نماذج تحتذى في بناء الاختبارات الجمعية . وظهرت بسرعة اختبارات الذكاء الجمعية لكل الاعمار ، من الاطفال بأعمار ما قبل السن المدرسي الى طلاب الدراسات العليا . وقد تطلب تطبيق الاختبارات الجمعية على اعداد كبيرة من الناس أن تكون تعليماتها بسيطة وأن لا يحتاج من يطبقها الا الى القليل من التدريب .

اختبارات الاستعدادات Aptitude Tests :

على الرغم من إن اختبارات الذكاء كانت في الاصل قد صُممت لتقيس وظائف عقلية متنوعة واسعة نتمكن من خلالها تقدير مستوى الذكاء العام للفرد ، فقد ظهر ان مثل هذه الاختبارات لا تغطي إلا جوانب محدودة من الذكاء ، لأن معظمها يقيس بصورة رئيسية القدرة اللغوية ، يليها في ذلك القدرة العددية . وقد ادرك علماء القياس النفسي بالتدريج ان مصطلح « اختبار الذكاء » intelligence test تسمية غير دقيقة للسبب الذي ذكرناه . ولذلك فإن بعض اختبارات الذكاء اصبحت تعرف فيما بعد بأختبارات الاستعدادات المدرسية scholastic aptitude tests . ويدل هذا الانتقال في مصطلحات التسمية على الاعتراف بأن الكثير مما يدعى بأختبارات الذكاء تقيس القدرات التي تتطلبها الدراسة في المدرسة .

وأسهم التقييم الناقد لأختبارات الذكاء ، بعد أن انتشر استخدامها انتشارا واسعاً ، الى جذب الانتباه الى حقيقة اخرى . فقد ظهر ان اداء الفرد يتباين على الاقسام المختلفة في اختبار الذكاء . وكان هذا واضحا بشكل خاص في الاختبارات الجمعية التي تعزل فيها الفقرات* المتشابهة وتجمع في مجموعة واحدة يطلق عليها الاختبار الفرعي subtest . فقد يحصل الفرد مثلا على درجة عالية في اختبار فرعي لفظي ، ودرجة واطئة في اختبار فرعي حسابي ، او العكس . ويلاحظ مثل هذا التباين الداخلي الى حد ما حتى في اختبار ستانفورد – بينيه .

ويستفيد علماء النفس العياديون من مقارنة هذه التباينات الداخلية للحصول على تبصّر أفضل في التركيب النفسي للفرد . وبناء على ذلك فإن معرفة نسبة

---

* الفقرات items : مصطلح يستخدم في القياس النفسي ليشير الى الاسئلة او المشكلات التي يتألف منها الاختبار . كما أنه قد يستخدم احيانا مصطلح « العبارات » كمرادف له .

الذكاء لوحدها لا تكفي ، وإنما يتطلب الامر معرفة درجات الفرد على الاختبارات الفرعية التي يتألف منها الاختبار الأصلي بأكمله ، لغرض تقييم حالة الفرد بصورة أقرب الى الدقة .

غير ان اختبارات الذكاء لم تصمم في الاصل لقياس الاستعدادات الفارقية ، ولأن الأختبارات الفرعية فيها لا تحتوي في بعض الاحيان على عدد مناسب كاف من الفقرات ، وهذا يؤدي الى عدم ثبات درجة الفرد على الاختبارات الفرعية . لذلك، اذا أردنا اجراء مقارنات بين قدرات الفرد المختلفة ، فإننا بحاجة الى تصميم اختبارات خاصة للقدرات بحيث يمكنها قياس الفروق الموجودة لدى الفرد في تلك القدرات

وقد هيأت الدراسات التي قام بها بعض العلماء مثل، سبيرمان Ch. Spearman وكيلي T.L. Kelley وثرستون L.L. Thurstone عن طبيعة الذكاء ، الوسائل لبناء اختبارات الاستعدادات الفارقية ، وذلك باستخدام الدرجات الكلية التي حصل عليها المجيبون في انوع عديدة من اختبارات الذكاء ودرجاتهم في الاختبارات الفرعية فيها ، بأتباع اسلوب احصائي يسمى « التحليل االعاملي » factor analysis . وقد أشارت البيانات التي تم الحصول عليها بهذه الطريقة الى وجود عدد من العوامل (أي القدرات) المستقلة عن بعض . وبعض هذه القدرات كان موجودا وبنسب متباينة في اختبارات الذكاء التقليدية ومن الامثلة على تلك القدرات : الفهم اللغوي verbal comprehension ، والاستدلال الحسابي numerical reasoning . اما القدرات المكانية spatial ، والادراكية perceptual ، والميكانيكية mechanical فإنها توجد في اختبارات الاستعدادات الفارقية بدرجة اكبر مما هي عليه في اختبارات الذكاء العام .

ونتج عن استخدام التحليل العاملي تطوير « بطاريات » الاستعدادات المتعددة multiple aptitude batteries . وتعني البطرية مجموعة من الاختبارات . وقد صممت هذه البطريات لتقدم لنا قياسا عن مكانة الفرد في عدد من القدرات . وهكذا فبدلاً من الحصول على درجة كلية للذكاء ، فإنه يتم الحصول على درجة منفصلة لكل قدرة على حدة ، فنتمكن بذلك من القيام بعمل نوع من التحليل الداخلي للفرد ، او ما يسمى بالتشخيص الفارقي لمعرفة أي القدرات قوية لديه وأيها ضعيف وهذا ما كان يسعى اليه العاملون في العيادات النفسية .

إن بطريات الاستعدادت الفارقية تمثل تطورا متأخرا نسبيا في ميدان القياس النفسي ، حيث إن معظمها ظهر تقريبا منذ عام ١٩٤٥ .

ولغرض التوضيح نشير الى إن مصطلح « اختبار الذكاء » يشير الى اختبار تكون فقراته متباينة وغير متجانسة وتقدم لنا درجة كلية شاملة عن ذكاء الفرد العام . اما « اختبار الاستعداد » فإنه يعني اختباراً تكون فيه الفقرات متجانسة

نسبيا وتقيس قدرة واحدة محددة بوضوح . اما بطريات الاستعدادات الفارقية المتعددة فإنها تقيس عددا من القدرات ويحصل فيها الفرد المفحوص على درجة عن كل قدرة بحيث تشكل كلها صحيفة نفسية profile عامة لقدراته .

اختبارات التحصيل المقننة Standerdized Achievement Tests

في نفس الوقت الذي كان فيه علماء النفس منشغلين في تطوير اختبارات الذكاء واختبارات القدرات ، كان يطرأ على الاختبارات المدرسية التقليدية إعدد من التحسينات الفنية . وقد بدأت اولى هذه الاختبارات لقياس نتائج التعليم المدرسي في بداية القرن العشرين متأثرة بأعمال ثورندايك Thorndike ومستفيدة من مبادىء القياس النفسي . فظهرت اختبارات لتقدير جودة الخط والانشاء والاملاء والحساب . ثم ظهرت بطريات التحصيل ، مبتدئة بنشر الطبعة الاولى من اختبار ستانفورد التحصيلي Stanford Achievement Test عام ١٩٢٣ ، الذي يقيس اداء التلاميذ في موضوعات دراسية مختلفة ، وتُقيم درجاتهم في ضوء درجة معيارية . وقد استخدمت في الاختبارات المقننة الفقرات الموضوعية objective items وسعى العاملون الى تصميم فقرات من هذا النوع لقياس الفهم understanding والتطبيق application وغيرها من الاهداف التربوية . تم ظهرت في الثلاثينات مكائن تصحيح الاختبارات test-scoring machines .

ونتيجة لأزدياد اسهام علماء النفس في بناء الاختبارات التحصيلية المقننة ، فإن الجوانب الفنية في هذه الاختبارت اصبحت تماثل بشكل متزايد تلك الجوانب الفنية الموجودة في اختبارات الذكاء والقدرات . كما إن قياس اختبارات التحصيل للأهداف التربوية الواسعة جعلها تتشابه في المحتوى مع اختبارات الذكاء والقدرات ايضا . ولذلك فإن الفرق بين اختبارات الذكاء والتحصيل في ايامنا هذه قليل .

قياس الشخصية :

يهتم الجانب الاخر من القياس النفسي بالجوانب الانفعالية او غير العقلية من السلوك . وتعرف الأختبارات التي تصمم لهذا الغرض باختبارات الشخصية personality tests ، رغم وجود بعض علماء النفس الذين يرون إن مصطلح الشخصية مفهوم واسع يشير الى االفرد بأكمله من الناحيتين العقلية والانفعالية . غير أنه يستخدم في القياس النفسي مصطلح « اختبار الشخصية » ليشير الى مقياس يقيس خصائص مثل التوافق الانفعالي ، والدافعية ، والميول ، والاتجاهات .

نجد البداية المبكرة لقياس الشخصية في استخدام كريبلين Kraepelin عام ١٨٩٢ لأختبار التداعي الحر free association ، حيث كان يقدم للمفحوص كلمات تنبيهية منتقاة بشكل خاص ، ويطلب منه الاجابة عنها بأول كلمة ترد الى ذهنه . وقد استخدم التداعي الحر لأغراض متعددة ولا يزال يستخدم الى الآن . كما يمكن الاشارة هنا ايضا الى اعمال كالتون ، وبيرسون وكاتيل في تطوير الاستفتاءات ومقاييس التقدير .

وتعتبر قائمة البيانات الشخصية Personal Data Sheet التي وضعها ودورث Woodworth خلال الحرب العالمية الاولى نموذجا لاختبارات التوافق الانفعالي . وبعد انتهاء الحرب أعدت صيغة مدنية منها ثم مالت الاختبارات فيما بعد الى تجزئة التوافق الانفعالي الى مكونات اكثر تحديدا مثل : التوافق البيتي ، والتوافق المدرسي ، والتوافق المهني . وركز بعض الاختبارات على موضوعات اجتماعية مثل السيطرة ـ الخضوع . ثم ظهرت اختبارات الميول والاتجاهات .

واستخدم اسلوب آخر لقياس الشخصية يعتمد على تطبيق الاختبارات الموقفية situational tests التي غالبا ما يكون الغرض منها مخفيا عن المفحوص ويكون في مثل هذه الاختبارات ، لدى المفحوص مهمة يؤديها ، ومعظمها تشبه مواقف الحياة اليومية . وأول هذه الاختبارات صممت في نهاية العشرينات وبداية الثلاثينات من قبل هارتشورن Hartshorne ومي May وجماعتها . وقد اهتمت هذه الاختبارات بقياس الامانة والغش والكذب والسرقة والتعاون والمثابرة . كما وضعت اختبارات موقفية متطورة خلال الحرب الثانية .

اما الاساليب الاسقاطية projective techniques ، فإنها تمثل اسلوبا ثالثاً لدراسة الشخصية وقياسها ، وقد استخدمها العاملون في العيادات النفسية بشكل خاص . ويعطى فيها للمفحوص موضوع غير محدد يسمح بمدى واسع من الاستجابات ، بحيث يعكس كل فرد انماط خصائصه الشخصية . كما استخدمت اختبارات اكمال الجمل ، والرسم ، واللعب التمثيلي ، وتفسير الصور او بقع الحبر .

وتواجه كل انواع اختبارات الشخصية هذه صعوبات خطيرة . فلكل طريقة فوائدها وقصورها . وعلى العموم فإن قياس الشخصية قد تأخر كثيرا وراء مقاييس القدرات والذكاء ، ورغم ذلك فقد حقق قياس الشخصية تقدما ملحوظاً منذ عام ١٩٥٠ ، وإن الكثير من ادوات القياس المتطورة لا تزال تحت التجريب والاستقصاء .

تطبيقات القياس النفسي :

يستخدم القياس النفسي في مجالات عديدة كالتربية ، والصناعة والادارة ، والجيش ، والعلاج النفسي ، وفي الابحاث النفسية والتربوية ، وفي كل مجال يتطلب توفر مواصفات عقلية ونفسية انفعالية معينة لدى العاملين فيه . وسوف نتحدث بشيء من التفصيل عن تطبيق القياس في التربية والتعليم .

القياس في التربية :

إن وظائف القياس واسعة وكثيرة في التربية . ويمكن تصنيفها على النحو الاتي : ـ

(١) تقويم التحصيل الدراسي للتلاميذ : غالبا ما يقوم المدرس بهذه الوظيفة . وفيها يلجأ الى وسائل قياسية متعددة ، منها الاختبارات التحصيلية واستخدام الملاحظة الموضوعية والمناقشات الصفية والامتحانات الشفوية . ومن خلال المعلومات التي يحصل عليها المدرس من هذه الوسائل القياسية يستطيع أن يكيف طريقة تدريسه ونوعية معاملته لتلاميذه تبعاً للفروق الفردية فيما بينهم . كما أن تقويم التحصيل المدرسي يساعد التلاميذ في التعرف على مدى تقدمهم في التحصيل ، وبالتالي يحفزهم ذلك على المزيد منه .

(٢) تشخيص صعوبات التعلم لدى التلميذ الواحد او طلبة الصف جميعا : فقد يجد المدرس من خلال عمله التدريسي أن هناك بعض التلاميذ الذين يتخلفون عن بقية تلاميذ الصف في الوصول الى المستوى التحصيلي المناسب . ويتجه المدرس الى القياس والتقويم من أجل اكتشاف هذه الصعوبات ومعالجتها . اذ يقوم بدراسة استعدادات التلاميذ وقدراتهم الخاصة وصفاتهم الشخصية والمزاجية المؤثرة على تحصيلهم . ويستعين المدرس في الوصول الى ذلك باستخدام وسائل القياس المتعددة كالمقابلة الشخصية والاستفتاءات والمقاييس وكتابة التقارير . وعندما تتجمع لدى المدرس معلومات وافية عن سلوك التلميذ ، يستطيع عندئذ أن يتعرف على العوامل المؤثرة في تحصيله ، مما يساعده على معالجة هذه العوامل .

(٣) تشخيص التغيرات في سلوك التلميذ : هناك طرق عديدة لاكتشاف مدى ما تحقق من الاهداف التربوية لدى التلميذ نتيجة للخبرات التربوية التي اكتسبها . ويمكن تصنيف هذه الطرق الى مجموعتين : مجموعة تستخدم اساليب قياسية ومجموعة ثانية لا تستخدم ذلك . وتتضمن الأساليب القياسية استخدام اختبارات تحريرية وشفوية او اختبارات الاداء . وتشمل الاساليب

غير القياسية المقابلات بين المدرس والتلميذ ، والبطاقة الشخصية للتلميذ التي تحتوي على معلومات مفصلة عنه ، وغير ذلك .

وهناك عقبتان تمنعان من الوصول الى التقويم الشامل لمدى تحقيق الاهداف التربوية عند التلميذ . وتتضمن العقبة الاولى وجود اهداف تربوية تصعب ترجمتها الى سلوك ملحوظ لدى التلميذ كالاهداف المتعلقة بتنمية الاتجاهات والميول الايجابية . ولذلك فإن التشخيص الشامل والدقيق للتغيرات الحاصلة في سلوك التلميذ ، لا يمكن أن نتوصل اليه بسهولة .

اما العقبة الثانية ، فهي إن التغير المرغوب في سلوك التلميذ لا يمكن ملاحظته الا بعد مضي اشهر او سنوات على تعرض التلميذ للخبرات التربوية . فاعداد التلميذ لمهنة مناسبة يعتبر احدى الاهداف التربوية التي تسعى المدرسة الى تحقيقها . غير انه لا يمكن التعرف على مدى تحقيق هذا الهدف إلا بعد دخول التلميذ الى ميدان العمل وقضائه فترة من الوقت فيه .

(٤) مساعدة المدرس على معرفة تلاميذ صفه : لكي يستطيع المدرس التعرف على تلاميذه بصورة جيدة ، فلابد من تعرفه على بعض جوانب حياتهم العائلية ، والاجتماعية والمدرسية ، والنفسية . وعندما تتوفر للمعلم هذه المعلومات الشاملة عن التلميذ فإنه يكون قادراً على تكوين صورة واضحة عنه ، تمكنه بعد ذلك من تقييمه بصورة دقيقة .

(٥) زيادة دافعية التلميذ نحو التعلم : ان اهتمام المدرس بنتائج الاختبارات التحصيلية في المواد الدراسية يؤدي الى ارتفاع مستوى تحصيل التلميذ في هذه المواد . فعندما يشعر التلميذ ان مدرسه يقيّم ما قام به من انجاز بصورة موضوعية ، فإن ذلك يزيد من دافعيته نحو التعلم . فالتلميذ الذي ينال درجات عالية في موضوع دراسي معين ثم يثاب عليها فإن ذلك يحثه على بذل المزيد من الجهد ، لأن النجاح يقود الى النجاح . اما اذا حصل التلميذ على درجات واطئة ، فإنه يكون بحاجة الى تشخيص الاسباب والصعوبات التي أدت الى مثل هذه النتيجة الضعيفة ، ثم يقرر المدرس في ضوء ذلك أنجح الوسائل الكفيلة برفع مستواه .

(٦) تطوير الجانب الاداري المدرسي : ويمكن للادارة المدرسية أن تستفيد من نتائج التقويم في الامور المتعلقة بشؤون المدرسة عموما والتلاميذ خاصة . فمن خلال تدوين المعلومات المتعلقة بتحصيل التلاميذ واستعداداتهم وميولهم باستخدام الاختبارات او الاستفتاءات او المقابلات او السجلات ، يستطيع المدير الناجح أن يكون صورة واضحة عن التلاميذ تفيده في عملية تقويمهم وفي توزيعهم على شعب الصفوف ، وفي ارسال تقارير دورية الى الجهات

التربوية ذات العلاقة عن مدى ما حققته المدرسة من تقدم .

(٧) توجيه التلاميذ وارشادهم : ويفيد التقويم والقياس في الحصول على البيانات عن التلاميذ الذين يعانون من مشكلات التكيف ، وفي توجيه التلاميذ نحو التخصصات العلمية المختلفة ، وفي التوجيه المهني . فالمرشد او الموجه التربوي يستخدم اختبارات الاستعداد والتحصيل والشخصية ، ويجري العديد من المقابلات سواء كانت فردية ام جمعية من أجل جمع المعلومات المختلفة عن التلميذ . وكلما تعددت وتنوعت وسائل القياس المستخدمة في تقويم التلميذ ، كان ذلك أضمن للوصول الى تشخيص دقيق صادق . ولا يتوقف دور الموجه او المرشد عند التشخيص ، بل يمتد الى اتخاذ القرار المناسب لعلاج الظاهرة المشخصة . فإن كان الأمر يتناول سوء التكيف لدى التلميذ ، تطلب ذلك من المرشد او الموجه أن يشخص طبيعة سوء التكيف وايجاد الحلول المناسبة له . وعلى هذا الاساس يلعب القياس دوراً مهماً في تشخيص السلوك المرغوب وغير المرغوب لدى التلميذ . اما تقويم ذلك السلوك فلا يتم الا بعد الوصول الى تشخيصه . وعندما تتوفر مقاييس جيدة لبعض الصفات السلوكية ، فإن ذلك يجعلنا قادرين على تشخيصها بشكل دقيق .

(٨) تقويم كفاءة المدرس : وهناك وسائل عديدة تستخدم لقياس كفاءة المدرس . وأسهل هذه الوسائل هي الامتحانات . وقد اعتبرت نسب النجاح العالية دليلا على نجاح المدرس في وظيفته . واستخدمت الاختبارات التحصيلية المقننة في الدروس كوسيلة من وسائل تقويم المدرس ايضا . ويمكن ايضا الاستفادة ، بحذر ، من تقويم التلاميذ للمادة التدريسية واساليب التدريس وقدرة المدرس على ايصال المعلومات اليهم . كما تستخدم انواع متعددة من مقاييس التقدير لقياس كفاءة المدرسين . وهذه المقاييس عبارة عن معايير متدرجة لبعض الصفات العقلية والمزاجية والوجدانية والاجتماعية الضرورية لنجاح المدرس في مهنته . ويطلب من المدرس أن (يقوم) نفسه للتعرف على جوانب القوة والضعف في تدريسه .

ولكي يكون تقويم المدرس اكثر دقة ويمكن الاعتماد عليه ، فمن الأفضل أن تستخدم عدة وسائل قياسية في آن واحد ، اذ يمكن الجمع بين نتائج الامتحانات والاختبارات التحصيلية المقننة ، وتقويم التلاميذ لمدرسهم واستخدام مقاييس التقدير المختلفة من أجل تكوين صورة واضحة عن كفاءة المدرس .

١٦

القياس في العلاج النفسي :

مَر بنا الحديث عن استخدام القياس في توجيه الطلبة وارشادهم . ونضيف هنا الى ان هذا الاستخدام لا يقتصر على مجال التربية فقط ، وانما يشمل الانسان في مختلف الاعمال والاعمار . فلقد اصبحت الصحة النفسية والعلاقات الشخصية السليمة هدفاً بارزاً للعلاج النفسي . وهناك تأكيد متزايد على استخدام الاختبارات لزيادة فهم الذات والنمو الشخصي .

ان الاختبارات النفسية تستخدم في الوقت الحاضر لحل العديد من المشكلات العملية التي نواجهها في حياتنا ، كما أنها تستخدم كأدوات مهمة في الابحاث الاساسية للحصول على البيانات عن ظاهرة معينة . ومن الامثلة على ذلك الدراسات عن طبيعة الفروق الفردية وقياس الفروق بين الجماعات ، والبحث عن العوامل البيولوجية والحضارية المرتبطة بالفروق السلوكية . وكل هذه الابحاث تتطلب مقاييس دقيقة . وهناك ابحاث اخرى كثيرة تحتاج الى المقاييس مثل مدى تأثير الطرق التربوية المختلفة على التعلم او التغير في السلوك ، ونتائج العلاج النفسي وتأثير القلق على الاداء . . الى آخر ذلك من الابحاث .

الفصل الثاني

الاختبارات والمقاييس النفسية

تعرفنا في الفصل الاول على معنى الاختبار او المقياس النفسي ، والفرق بينه وبين التقويم . وفي هذا الفصل سوف نتعرف على خصائص الاختبار النفسي .

وقبل البدء بهذه الخصائص ، لابد من التعرف على ثلاث خطوات ضرورية لفهم الاجراءات التي ينبغي القيام بها عند بناء الاختبارات ، وعلى انواع الاختبارات والمقاييس النفسية .

وفيما يلي الخطوات الثلاثة ، وهي : —

١ ـ **تحديد الخاصية النفسية التي يراد قياسها** : نحن لا نقيس الشيء او الشخص ، بل نقيس خاصية ذلك الشيء او الشخص . فنحن نقيس طول الفرد او وزنه او ذكاءه او النضج الانفعالي لديه . وعندما نتعامل مع خصائص مادية كالطول مثلا ، لا نجد صعوبة في تحديد معنى او خاصية الطول ، فالكل يعرف ماذا يعني الطول . ومثل هذا التحديد الواضح للمفاهيم لا نجده في المجال النفسي ، اذ ما المقصود بالذكاء مثلا ؟ وما نوع السلوك الذي يمكن وصفه بالسلوك الذكي ؟ وهل يعالج الذكاء مفاهيم مجردة ام موضوعات مادية ؟ وهكذا نجد نفس الصعوبة في تحديد معظم المفاهيم النفسية . ولهذا فإن اولى المشكلات التي يواجهها الباحث النفسي عند محاولته قياس الظاهرة النفسية ، هي الوصول الى تحديد واضح ودقيق للخاصية التى يريد قياسها . فبدون هذا التحديد للظاهرة لا يمكن الوصول الى قياسها بدقة .

٢ ـ **تحديد الاجراءات التي يمكن بواسطتها توضيح الظاهرة النفسية التي نريد قياسها** :

إن الجانب الثاني للقياس يتمثل بايجاد مجموعة من الاجراءات تعزل الظاهرة عن غيرها من الظواهر وتُبرزها . وتختلف الظواهر النفسية عن الظواهر الطبيعية في هذه الناحية . فمن السهل ايجاد الوسائل وتحديد الاجراءات التي يمكن بواسطتها قياس طول المنضدة او الشخص اذ توضع المسطرة بمحاذاة المنضدة او الشخص ، وعندئذ تعتبر أفضل وسيلة لقياس الطول .

اما في مجال القياس النفسي ، فما إن الباحث لا يمكنه قياس الظاهرة النفسية بصورة مباشرة ، لذلك فإنه يلجأ الى وضع « تعريف » للظاهرة النفسية التي يريد قياسها . وهذا التعريف هو عبارة عن وصف للسلوك الظاهر الذي يمكن ملاحظته والذي تتميز به تلك الظاهرة عن غيرها من الظواهر النفسية الاخرى . فإذا أراد الباحث مثلا بناء مقياس « للتوافق النفسي » فإنه يلجأ الى تعريف التوافق ، مستندا في ذلك على ما كتب عن موضوع التوافق من نظريات ، وما اجري من ابحاث . وقد يتوصل مثلا في النهاية الى وضع التعريف التالي : ـ

التوافق النفسي : يتضمن التوافق النفسي العناصر والمكونات التالية : ـ

أ) الاعتماد على الذات وهو أن يكون الشخص مستقلا معتمدا على نفسه في اتخاذ قراراته وقادرا على توجيهها ، ومستقرا من الناحية الانفعالية ، ومتحملا للمسؤولية .

ب) الاحساس بقيمة الذات : وهو أن يمتلك الفرد احساسا بأنه محترم من قبل الآخرين ، وأنه شخصية جذابة ، وأنه واثق من نجاحه في المستقبل .

جـ) الشعور بالانتماء : وهو أن يكون الفرد قادرا على تكوين الصداقات ، وحب الآخرين ، والتعامل مع الناس .

إن وضع مثل هذا التعريف للتوافق النفسي يمكّن الباحث الذي يرغب في بناء اختبار لقياس هذه الظاهرة النفسية من وضع « الفقرات » اللازمة للاختبار .

إن أهم مشكلة يواجهها العاملون في القياس النفسي ، هي عدم وجود اجماع على تعريف الظواهر النفسية ، لأن هناك نظريات متنوعة كل منها يعرف هذه الظواهر من وجهة نظر معينة . ودور الباحث هنا الوصول الى أشمل تعريف ، لكي يصبح بمثابة الدليل او الاطار الذي توضع في ضوئه فقرات الاختبار لقياس تلك الظاهرة .

٣ ـ تحويل اجابات الفرد على المقياس النفسي الى وحدات كمية :

في هذه الخطوة يقوم الباحث بتحويل اجابات الفرد على فقرات المقياس الى وحدات كمية ، وذلك باعطاء ارقام تعبّر عن تلك الاجابات . فاذا كان لدينا مقياس لقياس الاتجاه نحو التعليم المختلط ، فإن بالامكان وضع « بدائل » امام كل فقرة تقيس درجة الاتجاه نحو مضمون تلك الفقرة . ثم يقوم الباحث باعطاء ارقام لكل « بديل » . وهكذا عندما يختار المجيب احدى هذه البدائل ، فإنه بالامكان التعويض عن اجابته برقم معين .

مثـــال : ـ

ـ « إن الاختلاط بين الجنسين في الدراسة يؤدي الى تهذيب شخصية كل من البنين والبنات » :

اوافق بشدة ـ اوافق ـ لا ادري ـ لا اوافق ـ ارفض بشدة

ومن المتعارف عليه في القياس النفسي اعطاء الارقام التالية لكل بديل ، وهذه الارقام تعبّر عن « وزن الاتجاه او شدته » :

البديل	الوزن (الرقم)
اوافق بشدة	+ ٢
اوافق	+ ١
لاادري	صفر
لا اوافق	ـ ١
ارفض بشدة	ـ ٢

فاذا اختار المجيب البديل « اوافق » فإنه يحصل على (+ ١) في هذه الفقرة اما اذا اختار البديل « ارفض بشدة » فإنه يحصل على (ـ ٢) ، ثم تجمع الدرجات التي حصل عليها في كل فقرة ، فنحصل على « درجة كلية » total score تعبّر عن اتجاهه نحو التعليم المختلط ، وهو اما أن يكون « مؤيدا » او « معارضا » او « لا رأي له » في الموضوع .

ان وحدات القياس في الظواهر الطبيعية يمكنها أن تقيس الظاهرة بصورة مباشرة . فلقياس الوزن مثلا ، يمكن استخدام الغرامات او الكيلوغرامات . اما في قياس الظواهر النفسية ، فإننا نستخدم ، كما رأينا ، بدائل مثل : « اوافق بشدة ـ اوافق ـ لا ادري . . الخ » او « ممتاز ـ جيد جداً ـ جيد ـ مقبول . . الخ » . وهذه كلها وسائل غير مباشرة لقياس الظاهرة النفسية .

وفي القياس النفسي لا نضمن أن تكون المسافة مثلا بين : (اوافق بشدة ـ واوافق) هي نفس المسافة بين : (اوافق ـ ولا ادري) . بينما في قياسنا للوزن مثلا نستطيع بكل ثقة أن نقول : إن كل غرام يساوي الآخر تماما . واذا أردنا أن نقارن بين وزني شخصين مثلا يمكننا أن نقول باطمئنان إن وزن هذا الشخص ضعف الآخر ، اذا كان احدهم يزن ١٠٠ كلغم والآخر ٥٠ كلغم . اما في قياس الاتجاه نحو التعليم المختلط فمن الصعب القول إن اتجاه الشخص الذي حصل على درجة (١٦٠) يبلغ ضعف اتجاه الشخص الذي حصل على درجة (٨٠) . كل الذي يمكن قوله ، هو إن الشخص الاول اكثر تأييدا للتعليم المختلط من الثاني

انواع الاختبارات والمقاييس النفسية :

تتنوع طرق تصنيف الاختبارات والمقاييس النفسية ، ولكنها تشترك في الاشارة الى الاختبارات : التحصيلية ، والذكاء ، والاستعدادات الخاصة (او الفارقية) ، والشخصية . وفي الحديث عن انواع الاختبارات والمقاييس النفسية هنا ، لابد من الاشارة الى الاختبارات التحصيلية وذلك لأن الاجراءات الرئيسية في عملية بناء هذه الاختبارات تكاد تتشابه مع اجراءات بناء بقية الاختبارات وخاصة اختبارات الذكاء والاستعدادات الخاصة .

وفيما يأتي انواع الاختبارات : —

اولا — الاختبارات التحصيلية : وتهدف الى تقييم مدى تقدم الطالب في التحصيل الدراسي . والاختبارات في هذا المجال على انواع هي : —

(أ) الاختبارات التي يعدها المعلم لغرض تقييم تقدم الطالب الدراسي ، وهي اما أن تكون اختبارات موضوعية او مقالية . وتتطلب الاختبارات الموضوعية أن توضع اهداف للمادة العلمية التي يراد بناء اختبار لها ، وتحدد نسبة ما يوضع من فقرات لكل مجال من مجالات المادة العلمية وذلك تبعا لأهمية المجال . وتوضع هذه النسبة من الفقرات على اساس العدد الكلي من الفقرات الذي يقرره مصمم الاختبار .

(ب) الاختبارات التحصيلية المقننة : لا يمثل هذا النوع من الاختبارات شيئا مختلفا عن الاختبارات الموضوعية التي يعدها المعلم في قياس التحصيل الدراسي . فتصميم فقراته يتخذ نفس النمط ، ويتناول نفس جوانب المعرفة العلمية التي يتناولها النوع الاول . ورغم هذا التشابه بين هذين النوعين من الاختبارات الا إن هناك بعض الفروق بينهما هي : —

١ . إن الاختبارات المقننة تبنى على اساس محتوى واهداف المواد الدراسية لعموم القطر ، في حين تبنى الاختبارات التي يعدها المعلم على اساس المحتوى والاهداف التي تحددها المدرسة المعينة او المعلم نفسه (في حالة لا مركزية التعليم) .

٢ . إن الاختبارات المقننة تصمم بمساعدة مصممي الاختبارات ، في حين إن الاختبارات التي يعدها المعلم تصمم من قبله ، وقد يستعين قليلا بغيره ، في بعض الاحيان .

٣ . تستخدم الاختبارات المقننة فقرات سبق تجريبها مبدئيا على الطلبة ونقحت وحللت قبل أن تصبح جزءً من الاختبار ، أما الاختبارات

التي يعدها المعلم فقلما تتطلب مثل هذه الأجراءات .

٤ . تتطلب الاختبارات التحصيلية المقننة صدقا وثباتا عاليين ولا تتطلب الاختبارات التي يعدها المعلم مثل هذا الارتفاع في الصدق والثبات .

٥ . تهيء الاختبارات التحصيلية المقننة معايير مشتقة من عينات ممثلة لجميع طلبة القطر ، في حين تقتصر الاختبارات التي يعدها المعلم على صف او مدرسة معينة .

(جـ) الاختبارات التشخيصية : وهي اختبارات تحصيلية تهدف الى اعطاء صورة تفصيلية لقوة الطالب او ضعفه في موضوع دراسي معين . ومثل هذا التحليل التفصيلي يمكن أن يبين اسباب ضعف الطالب ويهيء الاجراءات العلاجية له . فالطالب الذي يعطى له اختبار في القراءة للصف الرابع ويتبين من خلال ادائه في الاختبار أنه غير قادر على توليف الاصوات لتشكيل الكلمات ، وأنه يكرر تلفظ الكلمات بصورة خاطئة ، فإن مثل هذه النتائج وغيرها تهيء اساس تخطيط التدريس العلاجي للتلفظ الصحيح . ويهتم مصممو الاختبار التشخيصي عادة بأولئك الطلبة الذين يعانون من ضعف ظاهر في بعض المهارات الدراسية كالقراءة والحساب .

ثانيا ـ اختبارات الذكاء : وهي تختلف عن الأختبارات التحصيلية في أنها تقيس القدرة العقلية العامة ، وتتنبأ بأداء الفرد مستقبلا . كما أنها لا تتناول ما يتعلمه الطالب في المدرسة فقط ، بل يمكن أن تتناول مجالات حياتية غير تحصيلية . واختبارات الذكاء على انواع ، فمنها اختبارات الذكاء الفردية وهي التي تطبق على فرد معين ، ومنها الاختبارات الجمعية التي تعطى الى مجموعة من الافراد في آن واحد وتكون تعليماتها موحدة . ويعتبر اختبار ستانفورد ـ بينيه واختبار ويكسلر امثلة على الاختبارات الفردية للذكاء . ويعتبر اختبار أوتس ـ لينون من امثلة اختبارات الذكاء الجمعية .

وقد تكون اختبارات الذكاء لغوية ، أي إن الفرد يجيب عنها باستخدام اللغة ، او قد تكون مصورة ، أو قد تتناول معالجة يدوية للأشياء .

ثالثا ـ اختبارات الاستعدادات الخاصة : وهي تختلف عن اختبارات الذكاء في أنها تقيس قدرة الفرد في مجال معين ، في حين إن اختبارات الذكاء تتناول قدرة عقلية عامة وتقيس قدرة الفرد في تعامله مع الافكار المجردة والعلاقات . وتتناول اختبارات الاستعدادات الخاصة مجالات مختلفة ، كالمجال الفني والموسيقي والمجال الميكانيكي واللغوي والكتابي وغيرها .

رابعا ـ **اختبارات الشخصية** : هناك فرق اساس بين الانواع السابقة من الاختبارات واختبارات الشخصية . اذ أن الانواع الاولى تهدف الى قياس القدرة ، أي ما يستطيع الفرد عمله . اما اختبارات الشخصية فتهدف الى تقييم ما سوف يفعله الفرد في مواقف الحياة الاعتيادية .

ويمكن تقسيم الشخصية (لاغراض القياس) الى العناصر الآتية : الخصائص المزاجية Temperament ، وتتمثل بصفات كالحيوية ، والتسلطية ، والخضوع ، والانطوائية والانبساطيه ، والأستثارة ، والهدوء وما شابه ذلك . اما العنصر الثاني فهو التكيف ، وهو يشير الى قدرة الفرد للعيش بسلام مع نفسه ومع العالم الذي يحيط به . فكلما تقبل الانسان نفسه والعالم الذي يحيط به اعتبر متكيفا والعكس صحيح . وتعتبر الميول احدى عناصر الشخصية ، وتشير الى نزعة الفرد الى البحث عن نشاطات معينة والمشاركة فيها . اما الاتجاهات والتي تتصل عن كثب بالميول ، فهي حالة من النزوع عند الفرد نحو تقبل جماعات معينة او افكار معينة أو مؤسسات اجتماعية معينة او عدم تقبلها .

وهناك عدة اساليب لقياس الشخصية هي : ـ

١ . ما يذكره الفرد عن نفسه : ويتم ذلك من خلال مقابلته او من خلال اعطائه استفتاءاً او اختباراً للشخصية .
٢ . تقييم الفرد من وجهة نظر الآخرين . ويتم تقييم الفرد في هذا الاسلوب بواسطة استخدام مقاييس التقدير .
٣ . قياس السلوك الفعلي للفرد : ويتم ذلك من ملاحظة الفرد اثناء قيامه بسلوك معين وتسجيل استجاباته او تقييمها .
٤ . قياس الاستجابات الاسقاطية : ويتم تقويم الشخصية في هذا الاسلوب بناء على ما يسقطه الفرد من ميول واتجاهات وقيم عندما يعرض عليه اختبار للبقع الحبرية مثلا او عندما يقوم برواية قصص عن اختبار يحتوي على صور غامضة غير محددة المعالم ، وما الى ذلك من الاساليب الأسقاطية .

خصائص الاختبارات والمقاييس النفسية :

هناك بعض الخصائص الأساسية في بناء المقاييس النفسية وهي : التقنين ، والموضوعية ، والثبات ، والصدق ، والمعايير . والباحث العلمي او المربي بحاجة الى معرفة خصائص الاختبار ليتمكن من تفسير بياناته التي يحصل عليها من تطبيقه الاختبار ، وليكون على ثقة اكبر في القرارات الارشادية او التوجيهية او التربوية التي يتخذها .

ويمكن الحصول على هذه الخصائص من القيام بتطبيق الاختبار على عينة من الافراد . ثم استخدام النتائج التي نحصل عليها من هذا التطبيق لتحديد تلك الخصائص .

وفيما يلي عرض لهذه الخصائص : ـ

آ ـ التقنيـــــن standardization

إن الاختبار المقنن هو اختبار قد حُددّت اجراءات تطبيقه ، واجهزته ، وتصحيحه بحيث يصبح من الممكن اعطاء الاختبار نفسه في اوقات واماكن مختلفة

وتزود بعض الاختبارات بجداول للمعايير norms تحدد الدرجة التي يحصل عليها الافراد . والاختبارات التي تحتوي على مثل هذه المعايير تدعى في بعض الاحيان « الاختبارات المقننة » ، وتدعى عملية جمع البيانات لاشتقاق المعايير « التقنين » . وقد يكون للاختبار جدول للمعايير ، إلا إن اجراءات تطبيقه قد تكون غير محددة ، كما أنه قد يكون للاختبار اجراءات تطبيق مقننة تقنيناً جيداً ، إلا أنه لم تشتق له المعايير . ومن الواضح ان استخراج المعايير لا يكون دقيقا ما لم تقنن الاجراءات تقنيناً جيداً .

ان الخطوة الاساسية الاولى في تقنين الاختبارات النفسية قد جرت عام ١٩٠٥ . ومعظم الاختبارات النفسية المنشورة في هذه الايام مقننة تقنينا جيدا اما في اختبارات الشخصية ، فإنه لا يزال يستخدم عدد من اجراءات التطبيق غير المقننة .

وتختلف الأختبارات من حيث مدى اكتمال تقنينها . لذلك يجب تحديد كل الشروط المؤثرة في اداء المجيب على الاختبار ، اذا اريد للاختبار أن يكون مقننا بصدق . وعندما يكون التقنين في تمام الفاعلية . فإن المفحوص يحصل على نفس الدرجة تقريبا اذا اعيد عليه الاختبار بغض النظر عمن يجري الاختبار . وهذا غير ممكن تماما لظهور بعض الصعوبات التي تنشأ بسبب طبيعة العلاقة بين المجيب ومن يعطيه الأختبار .

ب ـ الموضوعية objectivity

تتباين الأختبارات بدرجة موضوعيتها . فالموضوعية الكاملة في الاختبار هي أن يتوصل كل ملاحِظ او حكمٍ يلاحظ اداء الفرد على الاختبار الى نفس النتيجة في تقريره الذي يقدمه عن ذلك الفرد . وللقيام بهذه المهمة ، عليه أن ينتبه الى نفس الجوانب من الاداء . وأن يدون ملاحظاته لكي يزيل الاخطاء التي يمكن أن يقع فيها فيما لو اعتمد على ذاكرته ، وأن يصحح الأجابة على الاختبار

بأستخدام نفس القواعد . إن موضوعية اجراءات التطبيق يجب أن يحكم عليها بواسطة درجة الاتفاق بين الدرجة النهائية التي يقدمها ملاحظان مستقلان عن بعض او اكثر . وكلما كانت الملاحظة والتقييم ذاتيين subjective ، كلما انخفضت درجة الاتفاق بين الحكمين .

إن الاختبارات التي يختار فيها المفحوصون البديل الصحيح او البديل الافضل من بين عدة بدائل (مثل صواب ـ خطأ ، الاختيار من متعدد) يطلق عليها « الاختبارات الموضوعية » objective tests ، لأن بامكان المصححين كلهم استخدام مفتاح التصحيح scoring key والاتفاق على النتائج اتفاقا كاملا . وعلى النقيض من ذلك ، فإن اختبارات المقال الاعتيادية تفسح المجال امام الاختلاف الواسع بين المصححين . ويمكن بواسطة التعليمات الدقيقة للملاحظ او المصحح أن تصبح اختبارات الاجابة ـ الحرة free response tests ، واختبارات المقال ، والملاحظات ، موضوعية الى حد مقبول .

جـ ـ الثبات reliability

على الرغم من أن الصدق هو المفهوم الاكثر أهمية ، فإننا سنتعرف اولاً على الثبات ، وذلك لأن الصدق يتضمن الثبات الى حد ما .

يعرف الثبات بأنه الاتساق في النتائج ويعتبر الاختبار ثابتا إذا حصلنا منه على نفس النتائج لدى اعادة تطبيقه على نفس الافراد وفي ظل نفس الظروف . فلو افترضنا أن احد الباحثين يريد القيام ببحث لمعرفة متوسط اعمار المعلمين في العراق ، فإنه يعدّ استمارة لهذا الغرض ، ويختار عينة تمثل المعلمين ، ويوزع عليها تلك الاستمارة ، ويطلب منهم أن يدونوا فيها تاريخ الولادة . ثم يقوم بعد ذلك بحساب معدل العمر . ولكي يتأكد الباحث من ثبات اجابات المعلمين ، فإنه يقوم بارسال الاستمارة اليهم مرة اخرى بعد فترة من الزمن ، ويطلب منهم أن يدونوا نفس المعلومات التي طلبها منهم في المرة الاولى . وبذلك يصبح لدى الباحث اجابتان لكل معلم . فإذا اتفقت اجابات المعلمين في المرة الاولى مع اجاباتهم في المرة الثانية ، فإن ذلك يدل على ثبات اجاباتهم على تلك الاستمارة . ويستخدم في حساب معامل الثبات عادة ، معامل الارتباط بين الدرجات . وهو يستخرج من ايجاد علاقة الارتباط بين الدرجات التي حصل عليها الافراد في المرة الاولى والدرجات التي حصلوا عليها في المرة الثانية . ويطلق على النتيجة التي نحصل عليها مصطلح « معامل الثبات » . ويتراوح معامل الثبات بين الدرجتين (صفر ـ ١)* ، ويعتبر الصفر ادنى معامل الثبات ، اما الدرجة (١) فتمثل اعلى معامل للثبات .

* نود ان نشير الى ان معامل الارتباط يتراوح بين (-١) و (+١)

وعلى العموم ، تكون درجات الافراد في الاختبارات النفسية والتربوية اقل ثباتا من الاختبارات التي تتناول دراسة الظواهر الطبيعية وذلك للاسباب الثلاثة الاتية : ـ

١ ـ يمكن قياس الظواهر الطبيعية بصورة مباشرة ، بينما نستخدم في الاختبارات النفسية والتربوية مقاييس غير مباشرة لدراسة الظواهر . فقياس الطول مثلا يتم باستخدام المتر واجزائه ، وهي أداة تقيس الطول بصورة مباشرة . اما عندما نقوم بقياس الذكاء فإننا نلجأ الى استخدام بعض الادوات كالمكعبات والمثلثات والاشكال الهندسية الاخرى ، او استخدام اسئلة تمثل مشكلات تتطلب الحل ، وكل هذه تعتبر وسائل غير مباشرة في قياس الذكاء . وهذا الاختلاف بين نوعي القياس يؤثر على درجة الثبات ، فيكون في المقاييس والاختبارات النفسية والتربوية اوطأ مما هو عليه في مقاييس الظواهر الطبيعية .

٢ ـ ان المقاييس المستخدمة في دراسة الظواهر الطبيعية تتصف بدرجة عالية من الدقة ، على عكس المقاييس النفسية والتربوية التي تتصف بشيء من عدم الدقة . فالمتر واجزاؤه مقاييس دقيقة تماما لقياس الطول ، في حين أن قياس ذكاء الفرد باستخدام الاشكال الهندسية او الاسئلة اللفظية قد لا يكون دقيقا كالمتر . وتؤثر في كل الاحوال دقة المقياس على درجة الثبات .

٣ ـ ان خصائص الظواهر الطبيعية التي يراد قياسها تتصف بالثبات بدرجة اعلى من خصائص الظواهر النفسية . فالظواهر النفسية تختلف من فرد الى اخر ، وتختلف لدى الفرد الواحد نفسه بين فترة واخرة . فطول الفرد ، مثلا ، يكون اكثر استقرارا من ظاهرة القلق النفسي لدى لفرد ، وذلك لأن القلق يتذبذب تبعا للخبرات اليومية التي يمر بها الفرد . وهبوط درجة الاستقرار تؤثر كثيرا على معامل الثبات . ففي هذه الحالة قد يكون المقياس ثابتا الا أن الخاصية المقاسة غير مستقرة .

طرق حساب الثبات : هناك عدة طرق لحساب الثبات وأهم هذه الطرق هي : ـ

١) طريقة التجزئة النصفية : وهي كثيرة الاستخدام في البحوث التربوية والنفسية . وتعتمد اساسا على تقسيم عبارات المقياس الى قسمين ، وحساب معامل الارتباط بين اجابات الافراد عن هذين القسمين . وهناك عدة طرق لتقسيم الاختبار ، منها تحديد مستوى صعوبة كل سؤال ، وتوزيع الاسئلة بالتساوي على قسمي الاختبار حسب مستوى الصعوبة . وقد توزع الاسئلة بناء على مدى صلاحيتها للتمييز بين الافراد الذين حصلوا على درجات عالية فيها ، والافراد

الذين حصلوا على درجات واطئة فيها . او أنها قد توزع الى مجموعتين بحيث توضع الاسئلة الفردية في قسم والزوجية في القسم الآخر . والطريقة الاخيرة هي اكثر طرق التجزئة النصفية استخداما . والمثال الآتي يوضح كيفية تطبيقها : نفرض أن احد الباحثين قام باعداد اختبار يتألف من (١٢) عبارة او سؤالا . وقد وزع الاختبار على عينة من التلاميذ بلغ عدد افرادها (١٠) تلاميذ . وعند التصحيح ، أعطى (درجة واحدة) اذا كانت اجابة التلميذ على العبارة اجابة صحيحة ، واعطاه (صفرا) اذا كانت اجابته على العبارة خاطئة . وفي الجدول الآتي يكون الطالب الذي حصل على المرتبة الثانية قد اجاب عن العبارة رقم (١) ، (٢) ، (٣) ، (٦) ، اجابة صحيحة ، في حين كانت اجاباته عن بقية العبارات خاطئة . وفي هذه الحالة تعطى له درجتان لأجابته الصحيحة عن العبارات الفردية (١) ، (٣) ، وتعطى له درجتان ايضا لاجابته عن العبارات الزوجية (٢) ، (٦) اجابة صحيحة .

جدول يوضح كيفية حساب درجات الاسئلة الفردية والزوجية لاحدى الاختبارات

ترتيب الطلبة	الاسئلة الفقرات ١	٢	٣	٤	٥	٦	٧	٨	٩	١٠	١١	١٢	درجات الاسئلة الفردية	درجات الاسئلة الزوجية
١	١	٠	١	٠	٠	٠	٠	٠	٠	٠	٠	٠	٢	٠
٢	١	١	١	٠	٠	١	٠	٠	٠	٠	٠	٠	٢	٢
٣	١	١	١	١	٠	٠	٠	٠	٠	٠	٠	٠	٢	٢
٤	١	١	٠	١	١	٠	٠	١	٠	٠	٠	٠	٢	٣
٥	١	١	١	١	١	٠	٠	٠	٠	٠	٠	٠	٣	٢
٦	١	١	١	٠	١	١	١	٠	٠	٠	٠	٠	٤	٢
٧	١	١	١	١	١	١	١	٠	٠	٠	٠	٠	٤	٣
٨	١	١	١	١	٠	١	١	١	١	١	٠	٠	٤	٥
٩	١	١	١	١	١	١	١	١	١	١	١	٠	٦	٥
١٠	١	١	١	١	١	١	١	١	١	١	١	١	٦	٦

ثم يحسب الباحث معامل الارتباط بين الدرجات الفردية والزوجية للتلاميذ . وفي المثال السابق يبلغ معامل الارتباط (٠,٨٠٩) . وتوضح الدرجات الارتباطية الثبات بالنسبة الى نصف الاختبار فقط وليس الاختبار كله ، لأننا كنا قد قسمنا

الاختبار الى قسمين . ولاجل حساب ثبات الاختبار بصورة كاملة . نلجأ الى استخدام معادلة تسمى معادلة سبيرمان براون لتصحيح معامل الارتباط السابق . وفي المثال اعلاه يبلغ معامل الثبات بعد تصحيحه بهذه المعادلة (٠,٨٩٤) . وعلى العموم فإن هذه المعادلة ترفع من درجة معامل الارتباط .

إن طريقة التجزئة النصفية ذات فائدة في الاختبارات التي تكون فيها الاسئلة او العبارات متجانسة ، أي أنها عبارات تقيس خاصية نفسية واحدة . كالتكيف او العدوان ، او دافعية التلميذ نحو الدراسة . اما اذا قاس الاختبار عددا من الخصائص او الموضوعات فلا يكون من المناسب استخدام هذه الطريقة . لأنه لا يمكن من الناحية العملية المساواة بين فقرات الاختبار في الصعوبة ، والمساواة بين متوسطي قسمي الاختبار وانحرافاتها المعيارية . ومن ميزات هذه الطريقة أنها اقل كلفة في الجهد والوقت ، وذلك لأنها لا تتطلب اعادة تطبيق الاختبار .

٢) طريقة اعادة الاختبار :

وتتضمن هذه الطريقة تطبيق الأختبار على عينة ممثلة من الافراد ثم اعادة تطبيق الاختبار عليها مرة اخرى بعد مرور فترة مناسبة من الوقت . ثم يحسب معامل الارتباط بين الدرجات التي حصل عليها افراد العينة في المرة الاولى والثانية . والفرق بين طريقتي التجزئة النصفية واعادة الاختبار ، هو ان معامل الثبات في الطريقة الاولى يشير الى طريقة التجانس بين العبارات . ويقصد بالتجانس إن العبارات تقيس مفهوما واحدا ، بينما يشير معامل الثبات بطريقة اعادة الاختبار الى درجة استقرار الافراد في اجاباتهم على المقياس عبر فترة مناسبة من الزمن .

ويدلنا معامل الثبات العالي بطريقة اعادة الاختبار على وجود استقرار في اجابات الافراد عبر فترة معينة من الزمن . اما معامل الثبات الواطىء فقد يعني إن الظاهرة التي يقيسها الاختبار تتذبذب لدى الافراد الذين يطبق عليهم الاختبار ، بين وقت واخر . او ان المقياس قد يكون غير دقيق في قياس تلك الظاهرة ، او ان اجابات الافراد تتأثر بظروف لا تتعلق بالاختبار نفسه بقدر تعلقها بالمجيب نفسه او بالظروف المحيطة به .

ومن الضروري في حساب الثبات بطريقة اعادة الاختبار ، تقدير الوقت بين تطبيق الاختبار في المرة الاولى وتطبيقه في المرة الثانية بعناية . فاذا كانت الفترة قصيرة جدا بين التطبيقين ، فإن هناك احتمالا كبيرا بأن يتذكر المجيب الاجابة

التي ذكرها في المرة الاولى . وعندما تكون الفترة طويلة جدا فقد يحدث تغيير حقيقي في الصفة السلوكية التي يقيسها الاختبار (وذلك بسبب نمو الفرد خلال هذه الفترة) . ولهذا السبب فإنه ليس هناك اتفاق على الفترة الزمنية المناسبة بين تطبيقي الاختبار . ويمكن القول بصورة عامة ، ان ذلك يعتمد في غالب الاحيان على طبيعة الظاهرة التي يقيسها الاختبار . وفي كثير من الحالات يعتبر اعادة تطبيق الاختبار بعد فترة اسبوع او اسبوعين فترة مناسبة .

ومن المآخذ على هذه الطريقة ان الدرجات التي يحصل عليها الافراد في المرة الثانية من تطبيق الاختبار تكون اعلى بقليل من درجاتهم في التطبيق الاول ، وذلك بسبب الفة المجيب بالأختبار وتذكره لأجابته في التطبيق الاول . كما تتأثر اجابات الفرد في الاختبارين (التطبيق الاول والثاني) بعوامل عديدة كالصحة العامة والدافعية للاجابة عن الاختبار والحالة الانفعالية لديه ، وطريقة فهمه لتعليمات الاختبار واسئلته ، والظروف البيئية المحيطة به . لذلك فمن الضروري بالنسبة لمصمم الاختبار أن يهتم بهذه المتغيرات الدخيلة العرضية التي يسبب اهمالها انخفاض معامل الثبات الذي نحصل عليه من طريقة اعادة الاختبار .

٣) طريقة الصور المتكافئة :

تقوم هذه الطريقة على اساس اعطاء التلميذ – في نفس اليوم – صورتين من الاختبار نفسه ، بحيث تكون كل صورة مكافئة للاخرى وتساويها من حيث محتوى الاختبار والمتوسط والانحراف المعياري . كما يجب أن تكون عبارات الصورتين متساوية من حيث الصعوبة والعدد ، وأن تكون البدائل التي يختار منها المفحوص اجابته ، وتعليمات الاجابة والتطبيق واحدة في صورتي الاختبار .

ومن المآخذ على طريقة الصور المتكافئة هي صعوبة اعداد صورتين متكافئتين . وفي حالة تحقيق ذلك فعلا فإنه يضاعف الجهد والوقت والكلفة لبناء الاختبار .

٤) المزج بين طريقتي الصور المتكافئة واعادة الأختبار :

وهناك طريقة تحاول التقليل من مشكلات « اعادة الاختبار نفسه » . وذلك عن طريق اعداد صورتين من اختبار معين ، تُعطى الصورة الاولى منه في وقت معين وتُعطى الثانية بعد ذلك بفترة ، ثم يحسب معامل الارتباط بين درجتي الصورتين . وبذلك تجمع هذه الطريقة بين مزايا الطريقتين معا .

ثبات درجات المصححين :

يلاحظ في الاختبارات المقالية والاسقاطية بوجه عام إن تصحيح الاختبار يعتمد الى حد ما على بعض الجوانب الذاتية للمصحح ، رغم وجود قواعد يلتزم بها

(٤٨) ثبات بطريقة التجزئة النصفية

أمثلة

طبق اختبار ذكاء مكون من (٨) فقرات على (٦) افراد فكانت الدرجات كما في الجدول الآتي، احسب معامل ثبات درجات الاختبار بطريقة التجزئة النصفية؟

الافراد	١	٢	٣	٤	٥	٦	٧	٨
١	١	٠	١	١	١	١	٠	١
٢	١	٠	٠	١	١	٠	١	١
٣	٠	٠	٠	١	٠	٠	٠	١
٤	١	١	٠	٠	٠	٠	١	١
٥	١	١	١	١	١	٠	١	١
٦	١	١	١	١	١	١	١	١

الكل

س	ص	سَ	صَ	سَصَ
٢	٢	٩	٩	٩
٢	٢	٩	٤	٦
٠	٢	٠	٤	٠
٢	٢	٤	٤	٤
٤	٢	١٦	٩	١٢
٤	٤	١٦	١٦	١٦
١٦	١٦	٥٤	٤٦	٤٧

المجموع

$$ر = \frac{ن \text{مج} سص - (\text{مج} س)(\text{مج} ص)}{\sqrt{[ن \text{مج} سَ - (\text{مج} س)^{٢}][ن \text{مج} صَ - (\text{مج} ص)^{٢}]}}$$

$$= \frac{٦ \times ٤٧ - (١٦)(١٦)}{\sqrt{[٦ \times ٥٤ - (١٦)^{٢}][٦ \times ٤٦ - (١٦)^{٢}]}}$$

$$= \frac{٢٨٢ - ٢٥٦}{\sqrt{(٣٢٤ - ٢٥٦)(٢٧٦ - ٢٥٦)}} = \frac{٢٦}{\sqrt{٦٨ \times ٢٠}} = \frac{٢٦}{٣٧,٨٧١١٧}$$

= **٠,٧٠٥** معامل ثبات نصف الاختبار

معادلة سبيرمان - براون $= \frac{٢ ر}{١ + ر}$

$$= \frac{٢ \times ٠,٧٠٥}{١ + ٠,٧٠٥}$$

$$= \frac{١,٤١}{١,٧٠٥} =$$ **٠,٨٢٧** معامل ثبات كل الاختبار

وهو معامل ثبات جيد

الثبات بطريقة اعادة الاختبار

طبق احد الباحثين اختبار ذكاء مكونا من (٦) فقرات على (١٠) افراد فكانت النتائج كما في الجدول (أ) وبعد شهر اعاد تطبيق نفس الاختبار على نفس الافراد، فكانت النتائج كما في الجدول (ب) احسب معامل الاستقرار (اعادة الاختبار)

(أ)

الافراد / الفقرات	١	٢	٣	٤	٥	٦
١	١	١	٠	٠	٠	٠
٢	١	١	١	٠	٠	٠
٣	١	١	٠	٠	٠	٠
٤	١	١	١	١	٠	٠
٥	١	١	١	٠	٠	٠
٦	١	١	١	١	٠	٠
٧	١	١	١	١	١	٠
٨	١	١	١	١	١	٠
٩	١	١	١	١	١	١
١٠	١	١	١	١	١	١

(ب)

الافراد / الفقرات	١	٢	٣	٤	٥	٦
١	١	١	٠	٠	٠	٠
٢	١	١	١	٠	٠	٠
٣	١	١	٠	٠	٠	٠
٤	١	١	١	١	٠	٠
٥	١	١	١	٠	٠	٠
٦	١	١	١	١	٠	٠
٧	١	١	١	١	١	٠
٨	١	١	١	١	١	٠
٩	١	١	١	١	١	١
١٠	١	١	١	١	١	١

(الكل)

س	ص	س²	ص²	س ص
٢	٠	٤	٠	٠
٢	٢	٤	٤	٤
٢	٢	٤	٤	٤
٢	٤	٤	١٦	٨
٢	٤	٤	١٦	٨
٢	٤	٤	١٦	٨
٤	٤	١٦	١٦	١٦
٤	٥	١٦	٢٥	٢٠
٦	٥	٣٦	٢٥	٣٠
٦	٦	٣٦	٣٦	٣٦
٣٥	٤٠	١٤٥	٢٠٠	١٦٠

المجموع

$$= \frac{\text{ن مجـ س ص} - (\text{مجـ س})(\text{مجـ ص})}{\sqrt{[\text{ن مجـ س}^2 - (\text{مجـ س})^2][\text{ن مجـ ص}^2 - (\text{مجـ ص})^2]}}$$

$$= \frac{١٠ \times ١٦٠ - (٣٥)(٤٠)}{\sqrt{[١٠ \times ١٤٥ - (٣٥)^2][١٠ \times ٢٠٠ - (٤٠)^2]}}$$

$$= \boxed{٠٫٨٠٨} \text{ معامل ثبات الاختبار (معامل الاستقرار)}$$

وهو معامل ثبات جيد

د. أزهار عبود

[بناء الاختبارات التحصيلية]

↓

الخارطة الاختبارية (جدول المواصفات)

الخطوات ← ① تحديد الاهداف
② اعداد الخارطة الاختبارية

اولاً: تحديد الاهداف التعليمية (التربوية) :-

الهدف السلوكي: هو سلوك ايجابي يتوقع ان يكتسبه الطالب نتيجة تفاعله مع موقف تعليمي وتأثره بمحاضره.

تصنيف الاهداف السلوكية: صنف (بلوم) الاهداف السلوكية الى ثلاث مجالات رئيسية هي المجال المعرفي (العقلي)، والمجال الوجداني (الانفعالي)، والمجال المهاري (النفسحركي).
وصنف (بلوم) الاهداف السلوكية في المجال المعرفي (العقلي) الى ستة مستويات متدرجة من الادنى الى الاعلى ويتضمن دائماً المستوى الاعلى المستوى الادنى منه وهي :-

١- المعرفة (التذكر): هو تذكر المعلومات التي تعلمها الطالب سابقاً او التعرف عليها كتذكر او حفظ المصطلحات والحقائق والقوانين والقواعد والنظريات ومن امثلة الاسئلة في هذا المستوى (عدد انواع التقويم التربوي أو عدد شهداء المعلقات).

٢- الاستيعاب (التذكر): هو القدرة على ادراك معنى الموضوع الذي يدرسه الطالب او يعبر عن الكلمات الواردة في الموضوع بكلمات اخرى او بلغة اخرى، وتلخيصها الموضوع، والوصول الى استنتاجات او توقعات تعتمد على فهم الموضوع ومن امثلة الاسئلة في هذا المستوى على اسباب اجراء التقويم التمهيدي).

٢- التطبيق : هو القدرة على تطبيق المعرفة التي تعلمها في مواقف جديدة وتبدو أهمية هذا المستوى من أهمية انتقال التعلم لحل المشكلات والتكيف في حياة متغيرة باستمرار ، ومن امثلة الاسئلة في هذا المستوى (ميز بين التقويم المعياري والتقويم المحكي).

٤- التحليل : هو القدرة على تفكيك المادة العلمية الى اجزائها المكونة لها وادراك ما بينها من علاقات مما يساعد على فهم تركيبها وتحديد اوجه التشابه والاختلاف ومن امثلة الاسئلة في هذا المستوى (قارن بين الاختبارات المقالية والاختبارات الموضوعية).

٥- التركيب : هو القدرة على دمج اجزاء مختلفة مع بعضها لتكوين مركب او مادة جديدة وهو بذلك عكس التحليل. فمثلاً (ان يعد الطالب تقريراً عن القياس والتقويم).

٦- التقويم : هو القدرة على اعطاء حكم على قيمة المادة المتعلمة بموجب معايير محددة واضحة مثل الحكم الكمي والكيفي على موضوع في ضوء معايير

ثانياً: تحديد المحتوى (اعداد الخارطة الاختبارية) ٥-

الغرض من تحديد المحتوى هو اخذ عينة على شكل اختبار نفترض فيها ان تمثل محتوى المادة الدراسية ، وان هذه العينة من الاسئلة لابد ان تقيس اهداف معينة.

ولاعداد جدول المواصفات نتبع الخطوات الآتية :-

① تقسيم المادة الدراسية الى موضوعات او عناوين رئيسية لموضوعات فرعية.

② تحديد المجالات (معرفي ، وجداني ، ومهاري) والمستويات التي يمكن تمييزها ضمن كل مجال ، فالمستوى المعرفي يتضمن ست مستويات كما ذكر اعلاه.

(٣) تحديد وزن أو أهمية كل محتوى بالنسبة للمحتويات الاخرى في المادة الدراسية وقد يكون على اساس:

* عدد صفحات الموضوع الذي تم تدريسه.
* الزمن الذي استغرقه المدرس في تدريس الموضوع.
* رأي الخبراء

(٤) تحديد وزن أو اهمية كل مجال او مستويات المجال الواحد.
(فمثلاً لنفرض اننا سنعتمد في المجال المعرفي على المستويات الثلاث الاولى وهي التذكر والفهم والتطبيق وان الاوزان التي تناسب عدد واهمية الاهداف في كل مستوى هي ٤٥٪، ٣٠٪، ٢٥٪ على التوالي).

(٥) تحديد طول الاختبار التحصيلي اي عدد اسئلته مثلاً (٦٠) سؤال من نوع الصواب والخطأ.

عدد الاسئلة لكل وحدة = نسبة مستوى الهدف × نسبة المحتوى × عدد الاسئلة الكلي
٢٫٤ = ٦٠ × ٢٠٪ × ٢٠٪

ونقوم بتقريب عدد الاسئلة (٢٫٤) للموضوع الاول في مجال التذكر الى (٢) سؤال وهكذا.

مثال

صمم خارطة اختبارية في مادة علم النفس العام تتضمن اربعة مواضيع استغرق تدريسها للطلبة (٨، ٤، ١٢، ١٦) ساعة على التوالي موزعة على ثلاثة اهداف سلوكية في المجال المعرفي هي (٢٠٪)، (٣٠٪)، (٥٠٪) على التوالي، علماً ان عدد اسئلة الاختبار (٦٠) سؤال من نوع اختيار من متعدد.

٧- (٤٨)

المحتوى			الأهداف			المجموع
الموضوع	الزمن	نسبة المحتوى	التذكر ٤٠٪	الفهم ٢٠٪	التطبيق ٥٠٪	المجموع ١٠٠٪
① الرافعة	٨	$\frac{٨}{٤٠} \times ١٠٠ = ٢٠\%$	أ ٢	ب ٤	ج ٦	١٢
② التكيف	٤	$\frac{٤}{٤٠} \times ١٠٠ = ١٠\%$	د ١	هـ ٢	و ٢	٦
③ الذكاء	١٢	$\frac{١٢}{٤٠} \times ١٠٠ = ٣٠\%$	ز ٤	س ٥	ص ٩	١٨
④ الشخصية	١٦	$\frac{١٦}{٤٠} \times ١٠٠ = ٤٠\%$	ط ٥	ع ٧	ف ١٢	٢٤
المجموع	٤٠	١٠٠٪	١٢	١٨	٢٠	٦٠

عدد الأسئلة في كل خلية = نسبة مستوى الهدف × نسبة المحتوى × عدد الأسئلة الكلي

خلية أ = ٠٫٢٠ × ٠٫٢٠ × ٦٠ = ٢٫٤ ≃ ٢

" ب = ٠٫٢٠ × ٠٫٣٠ × ٦٠ = ٣٫٦ ≃ ٤

" ج = ٠٫٢٠ × ٠٫٥٠ × ٦٠ = ٦

" د = ٠٫١٠ × ٠٫٢٠ × ٦٠ = ١٫٢ ≃ ١

" هـ = ٠٫١٠ × ٠٫٣٠ × ٦٠ = ١٫٨ ≃ ٢

" و = ٠٫١٠ × ٠٫٥٠ × ٦٠ = ٣

" ز = ٠٫٣٠ × ٠٫٢٠ × ٦٠ = ٣٫٦ ≃ ٤

" س = ٠٫٣٠ × ٠٫٣٠ × ٦٠ = ٥٫٤ ≃ ٥

" ص = ٠٫٣٠ × ٠٫٥٠ × ٦٠ = ٩

" ط = ٠٫٤٠ × ٠٫٢٠ × ٦٠ = ٤٫٨ ≃ ٥

" ع = ٠٫٤٠ × ٠٫٣٠ × ٦٠ = ٧٫٢ ≃ ٧

" ف = ٠٫٤٠ × ٠٫٥٠ × ٦٠ = ١٢

د ـ الصـــــــدق validity

يعتبر الصدق من الخصائص المهمة التي يجب الاهتمام بها في بناء الاختبارات. فعندما يريد الباحث تصميم اختبار معين ، فلابد أن تكون هناك ظاهرة سلوكية معينه يقيسها الاختبار ، كأن يقيس ظاهرة التكيف الاجتماعي ، او الاتكالية ، او القلق ، او الذكاء ، او التحصيل في موضوع دراسي معين . ويقوم الباحث بتحويل هذه الظاهرة السلوكية الى عبارات يتألف منها الاختبار ، وعندئذ يسمى هذا الاختبار باختبار الذكاء او القلق او التكيف الاجتماعي . وعندما يتأكد الباحث بطريقة علمية إن الاختبار يقيس الظاهرة التي يريد دراستها او تشخيصها ، فعندئذ يعتبر الاختبار صادقا . فالاختبار الصادق اذن هو ذلك الاختبار القادر على قياس السمة او الظاهرة التي وضع لأجلها . فاختبار الاستعداد الدراسي يعتبر صادقا اذا كان قادرا على قياس الاستعداد الدراسي ، ويكون غير صادق اذا كان يقيس ظاهرة سلوكية اخرى . ويعتبر الاختبار صادقا بالنسبة لمجموعة الافراد الذين وضع الاختبار لأجلهم ، وقد يكون اقل صدقا اذا طبقناه على مجموعة اخرى من الافراد تختلف في خصائصها عن المجموعة الاولى . فالاختبار التحصيلي في موضوع التاريخ الذي اعد لطلبة الصف السادس الاعدادي ، قد لا يكون صادقا بالنسبة لطلبة الصف السادس الابتدائي ، وذلك لوجود فروق واضحة بين المجموعتين . كما أن الاختبار الصادق يكون في معظم الاحيان اختبارا ثابتا . ولكننا لا نضمن ان يكون الاختبار الثابت اختبارا صادقا في معظم الاحيان . فقد يكون أحد الاختبارات ثابتا يعطينا النتائج نفسها اذا أعدنا تطبيقه على العينة نفسها ، لكنه قد لا يكون صادقا في قياس الظاهرة السلوكية المقصودة .

انـــواع الصـــدق :

هناك انواع عديدة من الصدق هي : ـ

(١) **صدق المحتوى content validity** : ويهدف هذا النوع من الصدق الى معرفة مدى تمثيل الاختبار للظاهرة السلوكية او الموضوع الدراسي الذي يهدف الى قياسه . ويعتبر هذا النوع من الصدق أهم الانواع المستخدمة في الاختبارات التحصيلية . فالمعلم ، على سبيل المثال ، قد يرغب بالتعرف على مدى قدرة التلميذ في القراءة ، او تحديد ما يعرفه في الموضوعات الاجتماعية . وفي مثل هذه الحالة ، عليه ان يأخذ عينة ممثلة من محتويات الموضوع الدراسي المقصود ، ويقوم بصياغة فقرات الاختبار منها . وعندما لا تكون العينة ممثلة لكل المحتوى فإن الاختبار يعتبر غير صادق بالنسبة لذلك الموضوع الدراسي . والخطوة الاولى في الوصول الى صدق المحتوى هو القيام بتحليل الموضوع الذي نريد قياسه تحليلا منتظما لتحديد

مجالاته ، وتقدير الوزن لكل مجال بناء على أهميته بالنسبة لبقية المجالات الاخرى ، ثم اختيار عينة ممثلة لكل مجال من هذه المجالات . فعندما نقوم بوضع اختبار لموضوع الفيزياء ، مثلا ، فإنه يجب أن نختار عينة من كل مجال من مجالاته التي هي على سبيل المثال : الكهربائية ، والمغناطيسية ، والميكانيك ، والضوء ، وأن تكون العينة متناسبة مع أهمية المجال بالنسبة للموضوع . كما يجب أن تكون هذه العبارات ممثلة للاهداف السلوكية في موضوع الفيزياء . فاذا كانت هذه الاهداف هي : الاهداف المعرفية والاستيعاب والتطبيق ، مثلا ، فيجب أن تمثل العبارات ـ فقرات المقياس ـ هذه الاهداف بناء على الأهمية النسبية لكل هدف . ويطلب ، في غالب الاحيان ، من مجموعة من الخبراء في المناهج وتأليف الكتب ومن الذين لديهم خبرة تدريسية في الموضوع ، أن يقيّموا كل عبارة من عبارات الاختبار في ضوء مدى قياسها لمحتوى الموضوع الدراسي ، والاهداف السلوكية التي يسعى الى تحقيقها . وبناء على تقييم الخبراء ، تحذف الفقرات غير الصادقة ، وهي الفقرات التي لا تمثل ما يريد الاختبار أن يقيسه .

(٢) الصدق التجريبي : يختلف هذا النوع من الصدق عن صدق المحتوى ، لأننا نعتمد في استخراجه على اختبار آخر يسمى المعيار الخارجي ، بينما يعتمد صدق المحتوى على التحليل المنطقي للمحتوى دون اللجوء الى الوسائل التجريبية في اثبات ذلك .

والصدق التجريبي انواع هي : ـ

آ ـ الصدق التنبؤي predicrue validity : ويقصد به مدى قدرة الاختبار على التنبؤ بنتيجه معينة . ويستخدم هذا النوع من الصدق في اختبارات الاستعداد ، التي تهدف ، في الاصل ، الى التنبؤ بما يمكن أن ينجزه التلميذ في المجال الدراسي او المهني في المستقبل . ولكي يحصل الباحث على هذا النوع من الصدق فإنه يلجأ الى معايير مختلفة أهمها : ـ

١ . التحصيل الدراسي العام : يعتبر التحصيل الدراسي احدى المؤشرات المهمة للتنبؤ بصدق اختبارات الذكاء او الاستعداد الدراسي . ويمكن توضيح ذلك بالمثال الآتي : لنفرض إن أحد المعلمين أراد أن يعرف مدى قدرة احد أختبارات الاستعداد القرائي على التنبؤ بتحصيل التلاميذ في موضوع القراءة . وقد أعطى لهم اختبار الاستعداد في بدء السنة الدراسية وحصل منه على درجة لكل تلميذ . ثم أعطى لهم اختباراً تحصيليا في موضوع القراءة في نهاية السنة ، وحصل منه على درجة لكل تلميذ ايضا . فاذا كانت العلاقة الارتباطية عالية بين

درجات التلاميذ على اختبار الاستعداد القرائي ـ الذي أعطي لهم في بدء السنة الدراسية ، ودرجاتهم في الاختبار الذي أعطي لهم في نهاية السنة الدراسية ، فإن هذا يدل على قدرة اختبارات الاستعداد على التنبؤ .

٣ . انجاز الفرد في العمل نفسه : يتطلب هذا المعيار متابعة انجاز الفرد الفعلي في عمل ما . فمثلا يقوم الباحث . بعد اعطاء الاختبار . بملاحظة العامل او الموسيقي او المدرس في عمله الفعلي . ثم توضع له درجة عن ادائه للعمل . ويتحقق الصدق التنبؤي في اختبار الاستعداد او الذكاء او الشخصية عندما تكون العلاقة الارتباطية بين درجة اجابة الفرد في الاختبار وادائه في العمل عالية . وفي حالة تعذر ملاحظة الفرد اثناء قيامه بالعمل . فإنه يمكن الاعتماد على بعض المؤشرات الدالة على انجازه في العمل . منها مقدار ما ينتجه الفرد واستمراره في العمل والتقدم فيه . ورضاه عنه . او رضا رؤسائه عنه .

ب ـ الصدق التلازمي concurrent validity : في هذا النوع من الصدق لا تكون هناك فترة بين الاختبار الاول الذي صممه الباحث . والذي يريد التعرف على صدقه في قياس الظاهرة . وبين المعيار الخارجي . ففي الصدق التنبؤي نحصل على درجات المعيار الخارجي بعد فترة من تطبيق الاختبار الجديد قد تصل الى سنة او اكثر . اما درجات المعيار الخارجي في الصدق التلازمي فنحصل عليها اثناء تطبيق الاختبار الجديد . وهناك عدة طرق نستخدمها في الصدق التلازمي : ـ

١ . اعطاء اختبار يقيس نفس الظاهرة التي يقيسها الاختبار الاول . فعندما يصمم باحث اختبارا لقياس اتجاهات الطلبة نحو عمل المرأة ويريد أن يتعرف على صدقه ، فإنه يقوم باعطاء اختبار اخر يتناول نفس الظاهرة الى عينة الطلبة نفسها . ثم يستخرج العلاقة الارتباطية بين درجات الطلبة في الاختبارين . ويكون الاختبار صادقا صدقا تلازميا عندما تكون الدرجة الارتباطية عالية بين الاختبارين .

٢ . تقديرات المسؤولين عن العمل لانجاز العاملين لديهم . او قد يتضمن هذا مثلا تقديرات المعلمين لظاهرة سلوكية معينة لدى الطالب .فعندما يصمم باحث اختبارا لقياس دافعية الطلبة نحو الدراسة . ويريد التأكد من صدق الاختبار ، فإنه يطلب من مجموعة من المعلمين الذين يعرفون عينة الطلبة الذين يطبق عليهم الاختبار أن يقدروا دافعية كل طالب منهم وفق مقياس يتدرج من الدافعية العالية جدا الى الدافعية الواطئة

جدا نحو الدراسة . ثم يحسب العلاقة الارتباطية بين درجات التلاميذ على الاختبار وتقديرات المعلمين لهم . ويكون الاختبار صادقا اذا كانت العلاقة الارتباطية عالية .

٣ . المجموعات المتضادة : إن الاختبار الصادق هو الذي يميز بين الافراد فاختبار الذكاء يجب أن يميز بين الاذكياء والاقل ذكاء ، واختبار القلق يجب أن يميز بين الفرد القلق وغير القلق ، وهكذا . وعلى هذا الاساس يختار مصمم الاختبار مجموعتين من الافراد تختلفان في الظاهرة التي يقيسها الاختبار ، ويعطي لها الاختبار ، فاذا كانت هناك فروق واضحة بين اجابتي المجموعتين في الاختبار ، فإنه يكون قد تحقق فيه الصدق التلازمي .

ويمكن التمييز بين الصدق التنبؤي والصدق التلازمي بما يلي : ـ

((١)) وقت الحصول على درجات المعيار : ففي الصدق التنبؤي نحصل على درجات المعيار بعد فترة من تطبيق الاختبار الجديد ، وقد تصل هذه الفترة الى سنة او اكثر . اما درجات المعيار في الصدق التلازمي فيحصل عليها اثناء تطبيق الاختبار الجديد

((٢)) الفائدة من الاختبار : يستفاد من الاختبار الذي يتحقق فيه الصدق التنبؤي في التعرف على درجة انجاز الفرد في المستقبل . اما الصدق التلازمي فيستفاد منه في الاختبارات التي تقيم السلوك الحالي للفرد .

((٣)) يستخدم الصدق التنبؤي عادة في اختبارات الذكاء او في اختبارات الاستعداد الدراسي ، اما الصدق التلازمي فإنه يستخدم في اختبارات الشخصية والتحصيلية ايضا .

وهناك ثلاثة شروط اساسية يجب توفرها في المعيار الجيد (في كل من الصدق التنبؤي والتلازمي) هي : ـ

١)) أن لا يتأثر المعيار بالعوامل الذاتية : فالدرجات المدرسية ، مثلا ، تعتبر احدى المعايير المهمة للتوصل الى صدق الاختبار سواء كان تنبؤيا او تلازميا ، ولكن استخدامها يتطلب شيئا من الحذر عندما لا يعتمد تقديرها على اختبارات تحصيلية مقننة ، وذلك لأن درجة التلميذ قد تتأثر بالعوامل الذاتية لدى المصحح . كما يعتبر معيار تقديرات المعلمين والمشرفين على العمل والاقران مهما في الصدق التنبؤي والتلازمي لدى تقييم بعض جوانب الشخصية . ومثل هذا المعيار يتأثر ايضا بالاحكام الذاتية للمعلم او المسؤول عن عمل الفرد الذي يراد تقديره .

٢)) أن يكون المعيار ثابتا : ويقصد بثبات المعيار ، أن تبقى الدرجة التي يحصل عليها الفرد فيه ثابتة اذا اعيد تطبيقه عليه اكثر من مرة . ولكي يتأكد بعض الباحثين من ثبات معيار « تقديرات المعلمين والمسؤولين » . فإنهم يطلبون منهم أن يعيدوا تقدير الافراد مرة ثانية . فاذا تطابقت اجاباتهم في المرتين ، فإن ذلك يدل على ثبات المعيار .

٣)) أن يكون المعيار وثيق الصلة بما يقيسه الاختبار : فاذا كان الاختبار يقيس استعداد الطالب في موضوع الرياضيات ، او الاستقرار الانفعالي . فيجب أن يكون المعيار المستخدم (كالدرجات المدرسية او تقديرات المعلمين او الاقران) متعلقا بالموضوع نفسه ، دون غيره ، ليكون معيارا صادقا يمكن الاعتماد عليه في الحكم على الاختبار الجديد .

صدق البناء **construct validity** : يقصد بصدق البناء مدى قياس الاختبار لسمة او ظاهرة سلوكية معينة . فالباحث في هذا النوع من الصدق يحاول معرفة طبيعة الظاهرة السلوكية التي يسعى الاختبار الى قياسها . فاذا كان الاختبار اختباراً للذكاء او الأمانة ، فإن الباحث يحاول معرفة طبيعته والعوامل المكونة له . فهو يهتم اذن بالظاهرة التي يقيسها الاختبار ، ولا يهتم بقدرته على التنبؤ بالسلوك المستقبلي (صدق تنبؤي) للفرد ، او السلوك الحالي (صدق تلازمي) له . وهناك طرق عديدة يمكن أن نلجأ اليها للوصول الى هذا النوع من الصدق وأهمها : ـ

آ) معاملات الارتباط مع اختبارات اخرى . ويتم ذلك عن طريق حساب معامل الارتباط بين الاختبار الذي نريد الوصول الى صدق البناء فيه وبين الاختبارات التي تتشابه معه في الوظائف التي يقيسها ، والتي سبق التأكد من صدقها . فاذا افترضنا أن باحثا أراد بناء مقياس للطموح بين الطلبة ، فإنه لكي يتوصل الى صدقه ، يقوم بتطبيق اختبار آخر مستقل عنه يقيس « المثابرة » والتي تعتبر صفة من صفات الفرد الطموح . وعندما يحصل الباحث على معامل ارتباط عال بين الاختبارين ، يتوصل الى أن اختباره قد تحقق فيه صدق البناء .

ب) العلاقة بين درجات الفقرات والاختبار . وتفترض هذه الطريقة إن الدرجة الكلية للفرد تعتبر معيارا لصدق الاختبار . ويحاول الباحث عادة ايجاد العلاقة الارتباطية بين درجة كل فقرة من فقرات الاختبار والدرجة الكلية . وتحذف الفقرة عندما يكون معامل ارتباطها بالدرجة الكلية واطئا ، على اعتبار إن الفقرة لا تقيس الظاهرة التي يقيسها الاختبار

بأكمله . او أننا قد نختار مجموعتين متطرفتين من الافراد بناء على الدرجات الكلية التي حصلوا عليها في الاختبار . . ثم نستخرج الفرق بين اجابات افراد هاتين المجموعتين في كل عبارة من عباراته . وتحذف العبارة التي لا تظهر تمييزا واضحا بين هاتين المجموعتين . وقد نقوم باستخراج العلاقة بين درجات الاختبارات الفرعية (اذا كان الاختبار يتألف من عدة اختبارات فرعية) وبين الدرجة الكلية للاختبار

ج) تقديرات الخبراء : وفي هذه الطريقة نلجأ الى الخبراء لتقدير الافراد المبحوثين في الصفة السلوكية التي يقيسها الأختبار . فلو فرضنا أن أحد الباحثين صمم اختباراً لقياس التكيف الاجتماعي لدى طلبة المرحلة الاعداديّة وعرف الباحث التكيف الاجتماعي بأنه قدرة الفرد على ملاءمة سلوكه مع ما تتطلبه المواقف الاجتماعية . فإنه لغرض أن يعرف فيما اذا كان الاختبار يقيس التكيف الاجتماعي ام لا ، يختار مجموعة من الخبراء (المدرسون في المدرسة مثلا) الذين يجهلون اجابة الطلبة على المقياس ، ويطلب منهم تقدير سلوك كل طالب في بعض المواقف المدرسية التي تدل على التكيف . فاذا كانت درجة الارتباط بين تقديرات الخبراء واجابات الطلبة على المقياس عالية فإن ذلك يدل على أنه يتوفر في المقياس صدق البناء .

د) استخدام التحليل العاملي : ويهدف التحليل العاملي الى معرفة مدى قياس الاختبار للظاهرة التي صمم الاختبار لأجلها . وتعتمد فكرته على حساب معاملات الارتباط بين كل عبارة والعبارات الاخرى . وتتجمع نتيجة معاملات الارتباط « مصفوفة ارتباطية » تنقسم الى تجمعات ، ويجمع بين كل مجموعة عامل او اكثر . ونتيجة لهذه العملية ، فإن الاختبار يختزل الى عدد صغير من العوامل او السمات المشتركة والتي يطلق عليها « المكونات الاساسية » للظاهرة التي يقيسها الاختبار . ومن خصائص هذه الطريقة أنها تنقي الاختبار من العبارات الضعيفة التي يظهر أنها لا ترتبط بأي عامل او مكون اساسي من مكونات الاختبار . وهذا يعني أن هذه العبارة لا تقيس الظاهرة التي وضع الاختبار لقياسها .

(٤) **الصدق الظاهري** : وهو اقل انواع الصدق أهمية ، لأنه يحاول أن يتعرف على مدى قياس الاختبار للغرض الذي وضع من اجله ظاهريا . وقد يقوم واضع الاختبار نفسه بمراجعته من حيث نوع العبارات واسلوبها ووضوحها . ويتناول ايضا تعليمات الاختبار ومدى دقتها في توضيح كيفية الاجابة على عبارات الاختبار .

العوامل المؤثرة في الصدق :

١ ـ طول الاختبار : ان عدد عبارات الاختبار يؤثر في صدقه . وكلما زاد عدد العبارات ارتفع معامل الصدق ، وذلك لأن أية ظاهرة نفسية او سلوكية . كالقلق مثلا ، تتألف من مكونات اساسية ، وهذه المكونات بدورها تتألف من خصائص وصفات كثيرة . ويفترض من الناحية النظرية أن يكون المقياس الذي نبنيه لقياس هذه الظاهرة النفسيه (القلق مثلا) محتويا على كل تلك الخصائص والصفات ، مصاغة على شكل (فقرات) . وبما ان هذا غير ممكن من الناحية العملية ، لأن الاختبار في هذه الحالة يكون طويلا جدا تصعب الاجابة عليه ، لذلك تؤخذ عينة من هذه الفقرات . وبناء على ذلك فإن أي اختبار ليس الا عينة من الفقرات التي تمثل مجالا اوسع يحتوي على عدد كبير من الفقرات التي بمجموعها تمثل الظاهرة النفسية او السلوكية التي نريد قياسها . لذلك يعمد واضعو الاختبارات الى زيادة عدد الفقرات عندما يتبين لهم انخفاض معامل الصدق .

٢ ـ معامل الثبات : يؤثر ثبات المقياس في صدقه ، فانخفاض معامل الثبات قد يكون دليلا على وجود عيب في الاختبار ، فيهدم بذلك مؤشرا على انخفاض صدقه . وتجدر الاشارة في هذا المجال ، الى أن معامل الثبات العالي لا يكون دائما دليلا على صدق الاختبار ، حيث أنه من الممكن أن يكون الثبات عاليا والصدق واطئا في الوقت نفسه .

٣ ـ عينة الافراد : عندما يكون افراد العينة متجانسين في الخاصية التي يقيسها الاختبار ، فإن ذلك يؤدي الى انخفاض معامل الصدق . وعليه فإنه من الضروري أن يكون هنالك تباين بين افراد العينة في الظاهرة المقاسة .

٤ ـ ويحصل في بعض الاحيان أن يكون للمعيار في الصدق التجريبي علاقة واضحة بالاختبار الذي وضعه الباحث ، ولكن عدم الدقة في ضبط العوامل الدخيلة بين الاختبار والمعيار يؤثر على معامل الصدق . فالصعوبة او السهولة المتناهية للاختبار ، او عدم وجود رغبة لدى الطلبة في الاجابة على الاختبار ، او ضعف الفقرات في قياس الصفة السلوكية ، كل ذلك يعتبر من العوامل الدخيلة التي تؤثر كثيرا وتعمل على خفض معامل الصدق .

ولكي يمكن تفسير معامل الصدق تفسيرا دقيقا وخاصة عندما يكون المعامل منخفضا ، فإنه من الضروري أن تؤخذ بنظر الاعتبار كل العوامل التي أشرنا اليها ، وأن تدرس بعناية قبل الوصول الى حكم دقيق . ويمكن توضيح ذلك بالمثال الآتي : نفرض أن أحد الباحثين قد توصل الى معامل صدق واطيء مقداره (٠,٢٠) ، وقد حصل عليه من العلاقة بين معدل درجات الطلبة في معاهد المعلمين

ودرجاتهم على مقياس للتنبؤ بنجاحهم في مهنة التعليم . فكيف يفسر معامل الصدق الواطيء هذا الذي حصل عليه؟ يمكن أن يفسر هذه النتيجة على اساس أن المعلومات التي درسها الطلبة في المعهد ليست كافية للنجاح في مهنة التعليم . او ربما أن الدرجات الاكاديمية التي حصل عليها الطلبة لا تقيس من الناحية الفعلية معلومات الطلبة في الموضوعات الدراسية . او ربما إن الاختبار الذي صمم لتشخيص المعلم الناجح وتمييزه عن غير الناجح اختبار غير سليم ، او ربما أن العينة متحيزة وغير ممثلة لعموم الطلبة في معاهد المعلمين .

هـ ـ المعايير Norms

إن مصطلح المعيار غالبا ما يستخدم في القياس النفسي لكي يشير الى متوسط درجات جماعة معينة من الافراد على أحد الاختبارات النفسية . ويطلق على هذه الجماعة من الافراد اسم (الجماعة المعيارية) Norm group . والمعيار ـ عادة ـ ضروري في الاختبار النفسي او التحصيلي ، لأن درجة الفرد التي يحصل عليها في الاختبار (وتسمى الدرجة الخام) ليس لها معنى بحد ذاتها ، ولا تصلح للمقارنة مع درجته في اختبارات اخرى ، او مع درجة شخص آخر على نفس الاختبار ، أو على اختبارات اخرى ، الا بواسطة المعايير .

فالمعايير اذن مهمة ، لأنها تخبرنا عن كيفية اداء الآخرين على الاختبار ، فتوفر بذلك اساسا للمقارنة . فلو إن شخصا حصل على الدرجة (١٥) في اختبار المثابرة ، فإن هذه الدرجة لا تعني شيئا بحد ذاتها ، فهي لذلك غير قابلة للتفسير . ولكننا لو عرفنا إن الدرجة الكلية في الاختبار هي (٢٠) ، وهي اقصى درجة يمكن أن يحصل عليها أفضل المفحوصين ، فإن الدرجة ١٥ في هذه الحالة تمثل (٧٥ بالمائة) من الدرجة الكلية (٢٠) . ولكن تبقى الدرجة ١٥ غير واضحة ، لأنها تخبرنا فقط على ان هذا الفرد يمتلك قدرة على المثابرة تساوي ٧٥ بالمائة من المثابرة التي يقيسها الاختبار . لكن معناها يصبح أكبر اذا علمنا ان متوسط الجماعة التي ينتمي اليها ذلك الشخص يساوي (١٤) . فهنا علمنا ان هذا الشخص أفضل من متوسط جماعته بدرجة واحدة ، فمستواه في المثابرة اذن هو فوق الوسط بقليل . ان متوسط الجماعة (١٤) هنا هو (المعيار) الذي هيأ لنا اساسا للمقارنة .

ويجب أن لا نخلط بين المعيار والمستوى القياسي standard ان المعيار يخبرنا عن الاداء الحقيقي للافراد على الاختبار ، أي أنه يمثل الدرجات الفعلية التي حصل عليها الافراد في الاختبار . اما المستوى القياسي ، فإنه يشير الى المستوى الذي يجب أن يصله اولئك الافراد لكي تعتبر اجابتهم مقبولة . فمثلا ان الدرجة (١٠٠) تمثل المستوى القياسي للصحة النفسية في اختبار يقيس القلق ويتألف من ١٠٠ فقرة ، وهذا ممكن من الناحية النظرية . غير اننا في الحياة الواقعية لا نجد

شخصا ليس لديه شيء من القلق ، لذلك نجد إن المعيار قد يكون في مثل هذا الاختبار (٨٠) . وبناء على ذلك ، فإن الشخص الذي يحصل على درجة ٨٠ فأكثر نعتبره سويا لأنه يشبه اكثرية الجماعة التي ينتمي اليها .

ولكي تكون معايير الاختبارات دقيقة ، فإنها يجب أن تقوم على اساس درجات عينات كبيرة وممثلة من الاشخاص الذين بني الاختبار لأجلهم ، وأن تكون شروط تطبيق الاختبار عليهم موحدة ، وأن تكون اجابتهم على الاختبار اجابة جدية ، بنفس المستوى الذي يتوقع من الاشخاص الذين سوف يطبق الاختبار عليهم .

غير إن هناك بعض الحالات التي لا يكون فيها استعمال المعايير في الاختبارات ضروريا ، خاصة عندما يريد أحدهم تشخيص الفروق الفردية داخل جماعة معينة . فعلى سبيل المثال ، تكون المعايير ذات فائدة قليلة لمدير احدى المصانع الذي يرغب في استخدام أفضل عشرة من بين جماعة من المتقدمين الى العمل عنده ، اذ ان بامكانه ترتيبهم من الاعلى الى الادنى حسب الدرجات التي حصلوا عليها ، ثم اختبار العشرة الاوائل منهم .

٢٨

الفصل الثالث

بناء الاختبارات والمقاييس النفسية

مقدمة في بناء الاختبار :

لقد مضى على حركة بناء الاختبارات النفسية والتربوية مدة تزيد على النصف قرن . وقد أعدت خلال هذه الفترة الآف الاختبارات في مختلف انحاء العالم ، وأمكن التوصل الى طرق منهجية متعددة في بناء الاختبارات تتفق والهدف من الاختبار وطبيعة الموضوع المراد قياسه . فالتطور اذن لم يكن تطورا كميا يشمل الزيادة في عدد الاختبارات المتداولة فقط ، وانما شمل ذلك التطور الناحية النوعية ايضا . لذلك نجد اليوم اختبارات تختلف من حيث التصميم والاجراءات المستخدمة في الاعداد والاجابة . فنجد في هذا الباب اختبارات « الورقة والقلم » ، ومقاييس التقدير ، وقوائم الشخصية ، والاختبارات الاسقاطية . كما ان التنوع قد امتد الى الهدف من الاختبار ، حيث نجد ان هناك مقاييس الاتجاهات ، والميول ، والذكاء ، والاستعدادات الخاصة ، والتحصيل المدرسي . وقد شمل التنوع ايضا افراد العينة ، فهناك اختبارات للاطفال الصغار ، ولاطفال المدارس الابتدائية ، وللمراهقين ، والراشدين . كما يشمل هذا التنوع الافراد الاسوياء وغير الاسوياء .

ورغم وجود مثل هذا التنوع في الاختبارات ، وان لكل نوع منها خصوصيته في التصميم ، فإن هناك خطوات رئيسة ، يمكن بصورة عامة استخدامها في بناء الاختبارات النفسية ، باستثناء البعض منها . وفيما يلي خطوات بناء الاختبارات : —

١) تحديد الاهداف .
٢) تحديد المحتوى .
٣) جمع الفقرات .
٤) وضع تعليمات الاختبار
٥) شروط تطبيق الاختبار .
٦) اجراء التجارب الاستطلاعية لغرض تحليل الفقرات .
٧) اجراء التجربة الاساسية لقياس صدق الاختبار وثباته
٨) اشتقاق المعايير .

(اولا) : تحديد اهداف الاختبار :

يتطلب بناء الاختبار تحديد اهدافه ، والأهمية النسبية لكل هدف من تلك الاهداف . وبما إن الظواهر النفسية لا يوجد اتفاق كامل على تعريفها ، لذلك فإن الاهداف بدورها لا تكون واضحة ما لم يحدد مصمم الاختبار ماذا يقصد بالظاهرة النفسية التي يريد بناء اختبار لقياسها . ثم يقوم بعد ذلك بصياغة الهدف او الاهداف التي يريد للاختبار أن يحققها ، بشرط أن لا تكون تلك الاهداف متنافرة غير متجانسة ، كأن يلجأ الى قياس الدافع والذكاء في آن واحد .

اما بالنسبة للاختبارات التحصيلية ، فإن صياغة الاهداف فيها أيسر من الاختبارات النفسية ، لأن مصمم الاختبار التحصيلي لديه منهج دراسي مقرر يحتوي على مادة تعليمية لها ابعاد واضحة واهداف معروفة ، فيقوم بصياغة اهدافه وفق اهداف تدريس تلك المادة

(ثانيا) : تحديد محتوى الاختبار :

تعتبر هذه الخطوة من الخطوات المهمة في بناء الاختبارات ، لأنها الاساس الذي تبنى عليه الفقرات ، والمجال الذي تشتق منه . ولا شك إن لهذه الخطوة علاقة بالخطوة السابقة . فقد رأينا هناك إن من واجب مصمم الاختبار أن يحدد بدقة ماذا يقصد بالظاهرة النفسية او الخاصية التي يريد قياسها ، ليتمكن من صياغة اهداف الاختبار . وبعد تحديد الظاهرة النفسية واهداف الاختبار ، يقوم المصمم بتحديد ابعاد ومكونات تلك الظاهرة . وبمعنى آخر ، أنه يلجأ الى تجزئة تلك الظاهرة الى عناصرها الاولية ، فيمثل كل عنصر مجالا معينا ، او اطاراً مرجعيا لاشتقاق الفقرات منه ، ولتقييمها وايجاد صدقها الظاهري من قبل المحكمين في ضوء ذلك العنصر او المجال . ويضمن مصمم الاختبار في هذه الحالة أن يكون محتوى الاختبار قادرا على تحقيق الاهداف التي وضعها . وعلى مصمم الاختبار في هذه الخطوة أن يحاول تحديد الأهمية النسبية لكل عنصر او مجال في الظاهرة النفسية .

اما إذا كان الاختبار تحصيليا ، فإن على مصمم الاختبار في هذه الحالة أن يحدد الأهمية النسبية لكل وحدة من الوحدات التعليمية التي يتألف منها الموضوع . فاذا كان يريد بناء مقياس لموضوع العلوم مثلا ، فإن عليه أن يحدد الأهمية النسبية لكل وحدة من الوحدات التي يتألف منها كتاب العلوم ، لكي يعرف نسبة الفقرات التي ينبغي أن تخصص لكل وحدة .

(ثالثا) : جمع فقرات الاختبار :

يلجأ مصمم الاختبار النفسي الى عدة اساليب للحصول على الفقرات اللازمة لبناء اختباره . وتساعده الخطوة السابقة وهي « تحديد محتوى الاختبار » أن يقوم بنفسه في صياغة العديد من الفقرات التي تقيس العناصر المختلفة التي تكون الظاهرة النفسية . كما إن مراجعته ومسحه لما كتب عن الظاهرة او الخاصية النفسية التي يريد قياسها ، وما وضع لها من مقاييس ، يساعده كثيرا في الحصول على فقرات صالحة لأختباره .

ومن الوسائل المهمة الاخرى في الحصول على الفقرات ، اللجوء الى عينة صغيرة من الافراد الذين من المقرر أن يبنى الاختبار لهم ، والطلب اليهم أن يكتبوا بأجابات مفتوحة عن المجالات او العناصر التي تتألف منها الظاهرة او الخاصية النفسية التي يراد للاختبار أن يقيسها . فبأمكان المصمم أن يسألهم اذا أراد بناء مقياس لقياس « المثابرة » عن الخصائص التي يتميز بها الشخص المثابر ، فيحصل بذلك على الكثير من الفقرات الصالحة لقياس الظاهرة ، والنابعة من صميم مجتمعنا .

وقد تجمع فقرات مقاييس الشخصية عن طريق « دراسة الحالة » لمجموعة من الافراد الذين يحملون مواصفات يحاول الاختبار قياسها ، او مما يكتبه هؤلاء الافراد ، او من خلال مقترحات الاطباء النفسيين اذا كان الاختبار يتضمن نواحي مرضية ، او من الدراسات السابقة التي عالجت نفس الموضوع ، او من المصادر العلمية التي تناولته . وقد اتبعت مقاييس مشهورة في الشخصية كقائمة منيسوتا وقائمة كلفورد ـ زيمرمان اغلب هذه الاجراءات .

ويمكن في بعض اختبارات الشخصية الاعتماد على اطار نظري يعالج المفهوم الذي يتضمنه الاختبار . ويجب في هذه الحالة الالتزام عند جمع الفقرات بما يتفق وتلك النظرية . وقد اتبعت مقاييس عديدة مثل هذا الأجراء ، منها على سبيل المثال ، مقاييس « دافع الانجاز » التي اعتمدت في جمع الفقرات على نظرية مكليلاند واتكنسن .

اما بالنسبة للحصول على الفقرات في الاختبارات التحصيلية ، فكما رأينا في الخطوة الثانية ، فإن مصمم الاختبار يقوم بتحديد الأهمية النسبية لكل وحدة من وحدات الموضوع الدراسي ، ثم يقوم بوضع الفقرات لكل وحدة بحسب وزنها وأهميتها في الموضوع الدراسي .

ومن أجل اكمال بناء فقرات الاختبار ، على المصمم أن يتوصل الى اجابات عن الاسئلة الآتية قبل البدء بكتابة الفقرات . وهذه الاسئلة هي : ـ

١ . ما هو مقدار الأهمية النسبية لكل جانب من جوانب المحتوى ولكل هدف في الاختبار ؟ وبعبارة اخرى ما مقدار نسبة الفقرات التي توضع لكل جانب من المحتوى ولكل هدف ؟ وقد أشرنا الى ذلك عند الحديث عن تحديد اهداف ومحتوى الاختبار .

٢ . ما هو أفضل انواع الفقرات صلاحية للاستخدام في الاختبار ؟
٣ . ما طول الاختبار اللازم؟ أي كم عدد فقراته ؟
٤ . ما الاجراءات التي تتبع في تصحيح فقرات الاختبار ؟

أنواع الفقرات المستخدمة في الاختبار :

هناك انواع مختلفة من الفقرات التي يمكن استخدامها في الاختبارات . ففي الاختبارات التحصيلية ، هناك نوعان رئيسيان من الفقرات هما : (١) نمط من الفقرات يقدم المجيب فيه الاجابة بنفسه ، (٢) نمط يختار فيه التلميذ الاجابة من بين عدد من البدائل يضعها مصمم الاختبار . ومن امثلة النمط الاول ، « سؤال المقال » الذي يتطلب اجابة مفصلة ، والفقرات ذات الاجابة القصيرة التي لا يتطلب الاجابة عليها اكثر من جملة او جملتين ، والجمل الناقصة التي تتطلب اجابتها كلمة او عبارة . اما النمط الثاني فيتضمن فقرات الاختيار المتعدد ، او الصواب والخطأ ، وفقرات المطابقة او (المزاوجة) . ويعتمد اختيار نمط الفقرة ، الى حد كبير ، على طبيعة الهدف الذي يراد قياسه . وهناك عوامل اخرى تحدد الاختيار ، كطبيعة المحتوى ومهارة المصمم في بناء الفقرات .

اما الفقرات المتضمنة في اختبارات الشخصية والاتجاهات والميول فغالبا ما تكون من النوع الذي توضع له موازين تقدير متدرجة ثلاثية او خماسية ، او قد تكون البدائل ثنائية ، كأن تكون بشكل « نعم » « لا » مثلا . وقد ترتب الفقرات بشكل ازواج ، حيث يوضع كل زوج من الفقرات على حدة ويطلب من المجيب أن يختار الفقرة التي تنطبق عليه اكثر من الاخرى . او قد ترتب الفقرات بشكل ثلاثي ، أي أن توضع كل ثلاث فقرات على حدة ، ويطلب من المجيب أن يختار الفقرة التي تنطبق عليه اكثر من غيرها ، والفقرة التي لا تنطبق عليه اكثر من غيرها . وهناك نوع من اختبارات الشخصية يطلب من المجيب فيها أن يؤشر فقط على الفقرات التي تنطبق عليه في الاختبار ويهمل باقي الفقرات ، ثم تحسب الدرجة الكلية له على اساس الفقرات التي اختارها . كما قد تكون الفقرات من نوع « تكملة الجمل » وهي مألوفة في المقاييس الاسقاطية للشخصية .

اما اختبارات الاستعداد فتتضمن انواعاً عديدة من الفقرات بعضها لفظي والبعض الآخر مصور او هندسي او عملي . ففي الفقرات اللفظية ، قد يطلب من المجيب أن يختار من بين بدائل الفقرة ما هو متشابه مع محتوى تلك الفقرة ، وقد يطلب ما هو مخالف مع محتواها . او قد تعطى جملة ناقصة وعدة بدائل ويطلب اكمالها من بين تلك البدائل . وقد تكون الفقرة شكلا هندسيا ناقصا يطلب اكماله من بين عدة بدائل هندسية ، احداها صحيح والباقي خاطىء . او قد تكون

الفقرة على هيئة اشكال مصورة لها معنى معين ، وقد رتبت الاشكال بطريقة غير سليمة . ويطلب من الفرد ترتيبها بشكلها الصحيح ليصبح لها معنى .

طـــول الاختبـــار :

من الضروري بالنسبة لمصمم الاختبار أن يقرر العدد الكلي لفقرات الاختبار ، بحيث يكون مناسبا لقياس الظاهرة المعينة ، وأن لا تحذف بعض الفقرات التي تشكل جانبا اساسيا في قياس تلك الظاهرة . وقد وجد من بعض الدراسات أنه عندما يكون تصميم الاختبار وانتقاء الفقرات جيدين ، يمكن الحصول على درجة مناسبة من الصدق والثبات بأختبار قصير نسبيا .

ويعتبر الوقت المتاح لتطبيق الاختبار عاملاً مهماً في تحديد عدد فقرات الاختبار . وغالبا ما يتم معرفة الوقت الكافي للاختبار من خلال التجربة المبدئية لتحليل الفقرات ، والتي سنشير اليها فيما بعد ، وذلك عن طريق حساب معدل الوقت الذي يقضيه الطلبة في الاجابة عن الاختبار . وهناك عوامل عديدة تؤثر في تحديد عدد فقرات الاختبار ، بالاضافة الى الوقت ، منها : ـ

١ . نمط الفقرات المستخدمة في الاختبار : فالفقرة التي يجاب عنها اجابة قصيرة من قبل الطالب short-answer item تتطلب وقتاً اطول من فقرات الصواب والخطأ ، او نعم ـ لا ، او فقرات التقدير المتدرج الثلاثي او الخماسي او فقرات الاختيار المتعدد التي يطلب فيها من التلميذ أن يؤشر فقط على احدى البدائل .

٢ . عمر المجيب ومستواه الثقافي : فالطلبة في مرحلة الدراسة الابتدائية الذين تكون قدراتهم في القراءة والكتابة لا تزال في دور النمو يحتاجون الى وقت اكثر للاجابة عن الفقرة من الطلبة في المراحل الدراسية الاخرى الذين تطورت قدراتهم القرائية والكتابية بشكل واضح .

٣ . مستوى قدرة المجيب : فالافراد الذين يكون مستواهم الادراكي العام اوسع ، يستطيعون الاجابة عن الفقرة بوقت أسرع ممن هم اقل منهم قدرة .

٤ . مدى قدرة المجيب على الاستمرار في الاجابة دون تعب : فبعض المجيبين لا يستطيعون الاستمرار في الاجابة لفترة طويلة وخاصة صغار السن ، في حين أن الافراد الاكبر سنا تكون لديهم القدرة على الاستمرار في الاجابة .

٥ . طول الفقرة : فالاختبار الذي يتضمن فقرات قصيرة لقلة عدد كلماتها ، يحتمل أن تكون واضحة غير معقدة . وفي مثل هذه الحالة تكون الاجابة عنها اسرع مما لو كانت الفقرة طويلة ومعقدة .

٦ . المفهوم الذي يقيسه الاختبار : فالاختبار التحصيلي الذي يقيس تذكر المعرفة التي درسها الطالب يمكن الاجابة عنه بصورة اسرع من الاختبار الذي يقيس الفهم والاستيعاب . كما إن اختبار الشخصية الذي يقيس سمة واحدة تكون الاجابة عنه بصورة اسرع مما لو تضمن قياس عدة سمات .

تصحيـــح الاختبـــار :

من الضروري ، بالنسبة لمصمم الاختبار ، أن يتخذ قراراً فيما اذا كانت فقرات الاختبار لها نفس الوزن بالنسبة للدرجة الكلية ، ام ان بعض الفقرات اكثر وزنا من البعض الآخر ، مما يجعل الدرجات التي تحصل عليها الفقرات تختلف بحسب اختلاف أهميتها وقوتها في قياس الظاهرة .

وفي الاختبار التحصيلي يتوجب على مصمم الاختبار أن يقرر ما اذا كان سيستخدم التصحيح من أثر التخمين ام لا ، قبل تطبيق الاختبار .

ويلاحظ عموما ان هناك فروقا ملحوظة بين المجيبين في ميلهم الى التخمين في الاختبار الموضوعي عندما لا يكونوا متأكدين من الجواب . وهذه الفروق في الرغبة للتخمين تؤدي الى تباين في الدرجات ، لكنها لا تعود الى الفروق الحقيقية في التحصيل الاكاديمي بين التلاميذ . ولذلك فإن الغرض من اعطاء التعليمات للمجيبين بضرورة عدم التخمين ، وفرض العقوبة على الاجابات الخاطئة ، هو التغلب على هذه المشكلة .

ويعتبر التخمين مشكلة بارزة تؤثر في اختبارات السرعة والأختبارات التي تتألف فقراتها من بديلين . ويمكن الاعتماد على درجات هذه الاختبارات ، اذا ما صححت من أثر التخمين . اما الاختبارات التي تتضمن فقراتها اربعة او خمسة بدائل ، والتي يعطى للمجيب فيها وقت كاف للاجابة عن كل فقرة ، فإن الدرجة الكلية الناتجة من الاجابات الصحيحة تعتبر مقبولة ، وليست هناك حاجة الى التصحيح من أثر التخمين .

كتابـة الفقـرات :

بعد تهيئة المعلومات عن السمة او الخاصية التي يقيسها الاختبار ، تأتي الخطوة التالية وهي تحويل هذه المعلومات الى فقرات تقيس كل منها سلوكا معينا يرتبط بتلك الخاصية . ويفضل كتابة كل فقرة على بطاقة منفصلة ، لأن ذلك يمكّن المصمم من اعادة تنقيحها او حذفها او ترتيبها في الاختبار .

وتتطلب كتابة الفقرات أن تكون لدى المصمم قدرة لغوية وفنية على صياغة الفقرات . فكثيراً ما يشوه محتوى الفقرة او يصعب فهمه بسبب ضعف صياغتها . وهناك عدد من القواعد التي يجب مراعاتها في صياغة الفقرة ، وهذه القواعد

هي : —

(١) أن يكون محتوى الفقرة واضحا وصريحا ومباشرا .

(٢) الابتعاد عن التعبير اللغوي المعقد والمربك في الفقرة .

(٣) أن تحتوي الفقرة على المتطلبات الضروريه التي تساعد المجيب على انتقاء الاجابة المناسبة منها .

(٤) أن تثير المجيب بحيث تدفعه الى الاجابة بشكل صريح .

(٥) يفضل عدم استخدام الفقرات الطويلة ، وأن لا تزيد على (٢٠) كلمة .

(٦) تجنب استخدام بعض الكلمات مثل « كل » « دائما » « ابدا » « في الغالب » .

(٧) تجنب نفي النفي في الفقرات مثل « السل ليس مرضا غير معد » .

(٨) يجب أن تحتوي الفقرة على فكرة واحدة فقط .

(٩) أن تكون بدائل الفقرة قصيرة قدر الامكان .

إن القواعد السابقة يمكن أن تستخدم في بناء فقرات أي اختبار ، سواء كان تحصيليا ام غيره . على ان هناك بعض القواعد التي تنطبق على بناء فقرات الاتجاهات فقط ، وهذه القواعد هي : —

١ — يجب صياغة الفقرات بلغة الحاضر ، والابتعاد عن صيغ الماضي .

٢ — تجنب الفقرات التي تشير الى الحقائق .

٣ — الابتعاد عن الفقرات الجدلية والتي يمكن أن تفسر بعدة تفسيرات .

٤ — انتقاء الفقرات التي تغطي المدى الكامل لموضوع الاتجاه .

٥ — من المفضل أن يعبر نصف الفقرات المصاغة عن اتجاه ايجابي ، والنصف الاخر عن اتجاه سلبي .

٦ — تجنب الفقرات التي يوافق (او لا يوافق) الجميع عليها .

(رابعا) : وضع تعليمات الاختبار :

الخطوة التالية في تصميم الاختبار هي كتابة تعليماته . وهناك نوعان من التعليمات : النوع الاول لتوجيه الافراد الذين يجيبون عن الاختبار ، والنوع الثاني لتوجيه القائم بتطبيق الاختبار . وتكتب التعليمات في صفحة مستقلة من صفحات الاختبار ، او تدون في كتيب صغير ملحق بالاختبار اذا كانت تفصيلية . وهناك بعض القواعد التي تتبع في وضع التعليمات ، وهي : —

١ . أن تكون التعليمات سهلة الفهم ، قادرة على ايصال ما هو مطلوب من المجيب .

٢ . أن تؤكد التعليمات على ضرورة اتباع ما يرد فيها بدقة .

٣ . يجب أن تعطى التعليمات بصورة مبسطة وواضحة وبالتتابع

٤ . اعطاء فرصة للمجيبين للاستفسار اذا كان هناك ضرورة لذلك . وعلى المجرب الا يترك امرا غامضا بالنسبة للمجيبين .

٥ . يفضل قبل تطبيق فقرات الاختبار ، أن يفسح الوقت الكافي امام المجيبين لقراءة التعليمات والانتباه الى الامثلة في الكتيب .

٦ . يفضل وضع امثلة من الاختبار تبين للمجيب كيفية الاجابة عنها قبل البدء بتطبيق الاختبار . وفي الاختبار التحصيلي يفضل أن تعطى ثلاثة امثلة ، ويكون السؤال الاول سهلا جدا ، ويكون الثاني معتدل الصعوبة ، اما الثالث فيكون صعبا . ويجب أن تجيب التعليمات على المثال الاول ، ويمكن أن تترك الاجابة على السؤال الثاني والثالث الى الافراد لكي تعطى لهم الثقة لحلها بأنفسهم ، اما اجوبتها فيمكن أن توضح شفويا . اما في اختبار الشخصية فيمكن اعطاء امثلة تكفي لتوضيح المجال الكلي للاختبار ، وخاصة اذا احتوى الاختبار على اختبارات فرعية .

٧ . يفضل الا يوضح الغرض من اختبار الشخصية ، لأن ذلك قد يؤدي الى أن يجيب الافراد عنه بالاتجاه المرغوب فيه اجتماعيا .

٨ . يجب أن تكون التعليمات مقننة . ويعني التقنين ضرورة أن تعطى التعليمات الى المجيبين كما دونت اصلا في كتيب التعليمات ، وأن يتقيد المجرب بالتعليمات المطبوعة وأن يقرأها لهم وألا يضيف او يغير فيها شيئا . ان تقنين التعليمات لا يعني أن يكون المجرب خشن المعاملة ، بل يفترض أن يكون ودودا .

٩ . يجب أن تسمح التعليمات للمجيب بتقديم الاسئلة ، وعندما يجيب مطبق الاختبار ، عليها ، فإن عليه الا يضيف شيئا الى الافكار الواردة في التعليمات ، لأن مثل هذه الاضافات قد ينتفع منها مجيب واحد دون البقية ، ويؤثر ذلك على نتائج الاختبار .

إن أكثر الاسئلة ازعاجا للمجرب هي تلك التي تتناول امورا لم تجب عنها التعليمات المطبوعة . وهذه الاسئلة قد تتناول امورا مثل « هل نخمن عندما لا نكون متأكدين من الاجابة ؟ » « كم يعطى للجواب الخاطىء ؟ » « هل هناك اسئلة كاشفة للكذب ؟ » وتعتبر تعليمات الاختبار غير جيدة اذا اهملت مثل هذه الاسئلة .

ان تعليمات بعض الاختبارات تحدد الوقت المسموح به لانهاء الاختبار . وعندما لا يكون هناك وقت محدد مذكور في التعليمات ، فإن على مطبق الاختبار دراسة وضع المجيب بعناية واختيارما يناسبذلك المجيب بحيث يكون له تأثيره الايجابي على انجازه .

إن من اولى مهمات مطبق الاختبار أن يكسب ود المجيب وخاصة اذا كان في عمر صغير . وما لم تحدث الألفة بينهما فقد يؤثر ذلك على نتائج الاختبار . وان الوقت والجهد الضروريان لتحقيق هذه العلاقة يعتمد على شخصية كل من مطبق الاختبار والمجيب .

(خامسا) : شروط تطبيق الاختبار :

تعتبر هذه المرحلة محاولة لنقل الاختبار من حالة التهيئة الى حالة الاجابة عنه . فبعد اعداد فقرات الاختبار ووضع تعليماته تبدأ مرحلة تطبيقه على عينة من الافراد لغرض تحليل الفقرات . ويهمنا في هذه المرحلة أن نتعرف على المشكلات العامة للتطبيق ، والتي تعتبر مشتركة بين كل الاختبارات . وهذه المشكلات هي : -

١) **الظروف الفيزياوية لدى اعطاء الاختبار :** فاذا كانت التهوية والاضاءة رديئة ، فإن ذلك يؤثر كثيرا على اجابات الافراد . وفي اختبارات السرعة ، على وجه الخصوص ، تهبط درجات الافراد اذا لم تكن اماكن جلوسهم مناسبة للكتابة . ويجب أن يكون مكان المجرب مناسبا بحيث يستطيع الافراد سماع التعليمات ورؤية كيفية الاجابة . وتعتبر القاعات الكبيرة جدا غير مناسبة للتطبيق الجمعي ، ما لم يكن هناك عدد كاف من المساعدين ، لتلبية ما يحتاجه الافراد .

وقد تؤثر الحالة المزاجية للفرد وقت تطبيق الاختبار على الدرجة الكلية للاختبار . فعندما يكون الافراد في وضع نشط بدنيا وذهنيا ، وخاصة وقت الصباح ، يحتمل أن يبذلوا في الاجابة جهدا اكثر من الافراد المتعبين ذهنيا وبدنيا . ومن الواضح ان التعب لا يؤثر على الدافعية اكثر من تأثيره على قدرة الفرد للاجابة .

وفي الحالات التي يكون فيها الاختبار طويلا ، يفضل تجزئة الاختبار ، بحيث تكون هناك فترة مناسبة لاراحة المجيب ، وذلك من اجل تجنب التعب المتراكم . وقد وجد من بعض الدراسات ، ان اهتمام المجيب بالفقرات الاخيرة من الاختبار الطويل اقل من اهتمامه بالفقرات الاولى منه .

ويفضل عند تطبيق الاختبار ، وخاصة في المدارس ، أن يتم اختيار الوقت الانسب للتطبيق بحيث لا يكون قبل وقت الامتحان بفترة قصيرة ، والا يكون قبل اعطائهم موضوعا مهما او صعبا بوقت قليل ، والا يطبق في الساعات الاخيرة من الدوام المدرسي ، ويفضل الا يتجاوز وقت التطبيق حصة دراسية كاملة .

٢) **اثارة دافعية المجيب :** عندما يريد باحث قياس شيء مادي . كأن يكون وزن الفرد ، فإنه لا يواجه مشكلة الدافعية ، بل ان كل ما يتطلبه الموقف هو أن

٤٧

يوضع الفرد على ميزان معين ونحصل على وزن دقيق له ، بغض النظر عن شعوره تجاه هذا الاجراء . لكننا في الاختبارات النفسية يجب أن نضع الجانب النفسي في الاعتبار ، وما لم يكن للفرد اهتمام بالاجابة على الاختبار ، فإن نتائج ذلك الاختبار لا تقدم لنا الحقيقة .

وقد وجد في كثير من الدراسات ان رغبة الفرد في الاجابة عن الاختبار عامل مهم في الوصول الى النتيجة الصحيحة . وتعتبر استثارة رغبة المجيب عاملا مهما ينبغي الانتباه لها ، اذ يحتمل جدا فشل الاختبار الجيد التصميم ، عندما لا تكون لدى المجيب الرغبة في الاجابة عنه .

وقد درس علماء النفس العوامل التي ترفع من رغبة الفرد في الاجابة عن الاختبار ، فأستخدموا الجوائز والعبارات التشجيعية والمكافآت المالية . وأظهرت نتائج بعض الدراسات ان المجيبين الذين لا يحصلون على حوافز تكون اجابتهم نمطية ، أي متشابهة في الكثير من فقرات الاختبار رغم اختلافها في المضمون .

وهناك بعض العوامل التي تؤدي الى عدم الجدية في الاجابة عن الاختبار ، منها إن المجيب قد يكون في بعض الاحيان غير راغب في الاجابة عن الاختبار بالصورة التي تظهر لنا قدراته او شخصيته الحقيقية . اذ نجد ان هناك بعض الحالات التي يحاول فيها التلاميذ الحصول على درجات واطئة في اختبارات القدرة العقلية لكي لا يكلفهم المعلمون بواجبات اضافية تتفق مع قدراتهم العالية . وقد وجد في نتائج بعض الدراسات الاجنبية ان قسما من الجنود يجيبون عن الاختبارات العقلية او اختبارات الشخصية اجابات ضعيفة او مغلوطة لكي يتهربوا من تكليفهم بمهات خطيرة او صعبة . كما قد تتدخل بعض الحالات النفسية في الاجابة حتى اذا توفرت الرغبة لدى المجيب ببذل اقصى جهد للاجابة عن الاختبار . ومن تلك الحالات القلق والتوتر التي قد تؤدي الى ارتكاب الاخطاء في الاجابة .

ان على مطبق الاختبار أن يقنع المجيب بأن نتائج الاختبار سوف تستخدم لصالحه لا ضده ، وبذلك يزداد الصدق في الاجابة . فمثلا اذا كان المتقدم الى عمل يخشى من أخذ الاختبار الذي يتيح الاستعداد لذلك العمل ، فإن على مطبق الاختبار أن يخبره بأن النتيجة قد ترشده الى العمل الذي يحتمل أن ينجح فيه لأنه يتوافق مع قدراته وخصائصه الشخصية . والمريض النفسي الذي يخاف من معرفة نتيجة الاختبار يفهمه مطبق الاختبار بأن معرفة النتيجة سوف تساعد على معرفة العلاج .

٣) **تهيئة المجيب للموقف الاختباري** : إن للمجرب دورا مهما في تهيئة المجيب للاجابة ، وكسب تعاونه بعد أن يشعر المجيب بالثقة والفائدة من أخذ الاختبار . ويفضل في هذا المجال أن يُوضح الهدف من تطبيق الاختبار بصورة شفوية حتى في حالة وجود ذلك في تعليمات الاختبار . وقد دلت الدراسات على ان قيام مطبق الاختبار باشعار المجيب ان اجابته سوف تحترم وتستخدم لغرض البحث العلمي فقط دون أن تقدم لأية جهة اخرى ، يساعد على تهيؤ الفرد للاجابة ، وتكون النتيجة أفضل عندما لا يذكر المجيب اسمه .

اما اذا كان مطبق الاختبار تسلطيا ، وأحس المجيب بأنه يريد أن يفرض عليه أخذ الاختبار ، فإنه يخسر تعاون المجيب .

٤) تقنين الموقف الاختباري : ويقصد بتقنين الموقف الاختباري محاولة مطبق الاختبار ضبط الموقف الذي تعطى فيه التعليمات واثارة الدافعية المناسبة وذلك عن طريق توحيد الموقف الاختباري لجميع الافراد .

(سادسا) : تجارب تحليل الفقرات :

١ . **التجربة الاولية** : عند الانتهاء من كتابة فقرات الاختبار ، تأتي الخطوة الرئيسة التالية ، وهي تجربتها بصورة مبدئية على مجموعة من الافراد يبلغ عدد افرادها (٣٠) مجيب . إن اغراض هذه التجربة هي التعرف على مدى وضوح التعليمات والكشف عن جوانب الضعف فيها ، من حيث الصياغة والمضمون ، ومعرفة الوقت الذي يستغرقه الاختبار ، والطول المناسب له ، وقيام المجيبين بتشخيص الفقرات الغامضة او الصعبة بهدف اعادة صياغتها . ويفضل أن يجري مصمم الاختبار هذه التجربة بنفسه . ويجب أن يستفيد الباحث من النقاط التي يثيرها المجيبون عن جوانب الضعف في فهم التعليمات ، بحيث يعدل فيها بما يجعلها مفهومة لديهم . اما بالنسبة للفقرات ، فهي اما أن تبدل تماما او أن تعدل . ويفضل أن يناقش الباحث هذه التعديلات التي اجريت على التعليمات والفقرات مع عينة اخرى صغيرة من الافراد للتأكد من وضوح التعديلات التي اجريت على التعليمات والفقرات .

٢ . **التجربة الثانية** : والغرض من هذه التجربة هو تحليل فقرات الاختبار ، ويطبق فيها الاختبار على مجموعة من الافراد يمثلون المجتمع الذي نعد له الاختبار . ومن المفضل هنا تطبيق الاختبار على عينة ممثلة يبلغ عددها (٤٠٠) فردا .

تحليل الفقرات item analysis

إن تحليل الفقرات هو عبارة عن عملية فحص أو اختبار استجابات الافراد عن كل فقرة من فقرات الاختبار . وتتضمن هذه العملية الكشف عن مستوى « صعوبة الفقرة » وقوة « تمييز الفقرة » . وفعالية البدائل في فقرات الاختبار .

ويوجد العديد من الطرق للقيام بتحليل الفقرات . والطريقة التي سوف نعرضها ، تعتبر من الطرق المبسطة في تحليل الفقرات . كما إن النتائج التي تقدمها لا تختلف كثيرا عن النتائج التي تقدمها الطرق المعقدة ، وخاصة اذا أردنا بناء اختبارات او مقاييس لمجموعات ليست بالكبيرة ، كما هو الحال في بناء اختبار لأحد الصفوف المدرسية الذي يبلغ عدد طلبته بحدود (٤٠) طالبا . لذلك فإن مسألة اختيار احدى الطرق هي مسألة ذوق ، او مسألة تتعلق بالكادر المتوفر ، او ما يتوفر من امكانات للقيام بالعمليات الاحصائية .

١) ايجاد صعوبة الفقرة item difficulty :

لغرض ايجاد صعوبة الفقرة نتبع الخطوات الآتية : -

١ . ترتيب الدرجات التي حصل عليها الطلبة في الاختبار من اعلى درجة الى ادنى درجة .

٢ . أخذ مجموعتين من الدرجات ، تمثل احداها الافراد الذين حصلوا على اعلى الدرجات في الاختبار ، وتمثل الثانية الافراد الذين حصلوا على اوطأ الدرجات . وقد وجد ان نسبة الـ ٢٧ ٪ العليا والدنيا من الدرجات تمثل أفضل نسبة يمكن أخذها في ايجاد صعوبة الفقرة ، وذلك لأنها تقدم لنا مجموعتين بأقصى ما يمكن من حجم وتمايز ، عندما يكون توزيع الدرجات على الاختبار على صورة منحني التوزيع الاعتدالي . وهنالك من يأخذ نسبة ٢٥٪ او ٣٠٪ العليا والدنيا من الدرجات .

اما في الاختبارات الصفية ، وهي الاختبارات التي تبنى لتطبيقها على أحد الصفوف المدرسية ، او الاختبارات التي تبنى لتطبق على مجموعة صغيرة لا يزيد عدد افرادها على عدد صف من الصفوف المدرسية ، فإنه ليس من المهم جدا التقيد بهذه النسب المئوية في اختيار المجموعتين العليا والدنيا فمثلا اذا كان عدد طلاب الصف (٤٠) ، فإن بامكان الباحث أن يقسم الى مجموعتين بالتساوي بعد ترتيب درجاتهم على الاختبار من الاعلى الى الادنى ، بحيث يكون في المجموعة الاولى ٢٠ طالبا حصلوا على الدرجات العليا ، ويكون في المجموعة الثانية ٢٠ طالبا حصلوا على الدرجات الواطئة أي إن الباحث في هذه الحالة يقسم طلبة الصف الى نصفين : النصف الاعلى والنصف الادنى .

٣ . احصاء عدد التلاميذ الذين اجابوا عن الفقرة بصورة صحيحة ، في كل من المجموعتين العليا والدنيا .

٤ . اضافة عدد الاجابات الصحيحة في المجموعة العليا الى عدد الاجابات الصحيحة في المجموعة الدنيا ، وتقسيم الناتج على مجموع عدد الافراد في كل من المجموعتين العليا والدنيا ، والقيمة التي تحصل عليها هي صعوبة الفقرة .

ومعادلة صعوبة الفقرة هي :

$$ص = \frac{م}{ك}$$

حيث ان :

ص = صعوبة الفقرة .

م = مجموع الافراد الذين اجابوا عن الفقرة بصورة صحيحة في كل من المجموعتين العليا والدنيا .

ك = مجموع عدد الافراد في كل من المجموعتين العليا والدنيا .

ولتوضيح ما نحتاج اليه من بيانات عند تحليل الفقرة ، نعرض فيما يلي فقرات أخذت من اختبار لنرى كيفية تحليل كل منها : ان اعلى درجة حصل عليها أحد الافراد في الاختبار كانت (٨٥) درجة ، وان ادنى درجة حصل عليها أحد الافراد كانت (١٤) . ثم نقوم بترتيب الدرجات التي حصل عليها الافراد في الاختبار من اعلى درجة الى ادناها ، أي أننا نبدأ بالدرجة (٨٥) وننتهي بالدرجة (١٤) . ثم نأخذ الـ ٢٧٪ العليا من الدرجات والـ ٢٧٪ الدنيا من الدرجات . ولنفرض إن مجموع الطلبة الذين طبق عليهم الاختبار بلغ (١٠٠) طالب ، فعندئذ يبلغ عدد افراد المجموعة العليا (٢٧) طالبا ، وعدد افراد المجموعة الدنيا (٢٧) طالبا ايضا . وتخصص في العادة بطاقة منفصلة لكل فقرة ، سواء كانت من نوع الاختيار من متعدد ام من نوع الخطأ والصواب تدون فيها اجابات الطلبة من المجموعتين العليا والدنيا على كل بديل من بدائل الفقرة ، ويوضع خط تحت البديل الصحيح .

وفي ادناه امثلة لكيفية حساب صعوبة الفقرة ومعرفة مدى جاذبية بدائلها ، على افتراض إن عدد الطلبة الذين اجابوا على الاختبار (١٠٠) .

مثـــــال (١) :

محتوى الفقرة : أي من المحاصيل الآتية لا تزرع في جنوب العراق ؟

	المجموعة العليا ٪٢٧	المجموعة الدنيا ٪٢٧
آ . التمر	٢	٦
ب . التبغ	١٦	١١
جـ . الرز	٥	٥
د . الحنطة	٤	٥

حساب صعوبة الفقرة : عدد الاجابات الصحيحة في المجموعة العليا + عدد الاجابات الصحيحة في المجموعة الدنيا .

١٦ + ١١ = ٢٧

يقسم هذا المجموع (٢٧) الذي حصلنا عليه على مجموع عدد الافراد في المجموعتين العليا والدنيا (٥٤) :

$$\frac{٢٧}{٥٤} = ٠,٥٠ .$$

وتعتبر هذه فقرة متوسطة الصعوبة لأن (١٦) من افراد المجموعة العليا و (١١) من افراد المجموعة الواطئة أي (٥٠٪) من افراد المجموعتين قد أجابوا عن الفقرة بصورة صحيحة .

مثـــــال (٢) :

محتوى الفقرة : ما هو الشيء الذي يعتبر اكثر ضرورة لحياة الانسان من بين الاشياء الآتية : -

	المجموعة العليا ٪٢٧	المجموعة الدنيا ٪٢٧
آ . البروتين	٤	٧
ب . الفيتامينات	٨	١٥
جـ . الماء	١٥	صفر
د . المعدنية	صفر	٥

تعتبر هذه الفقرة صعبة ، لأن عدد الذين أجابوا عنها بصورة صحيحة يبلغ (١٥) من اصل (٥٤) فردا . أي إن نسبة الصعوبة تصل الى حوالي ٠,٢٨ . كما إن بعض البدائل الخاطئة قد جذبت بعض افراد المجموعة العليا . كما توضح الفقرة ان غالبية افراد المجموعة الدنيا قد ركزت على اختيار خاطىء معين وهو البديل (ب) الذي يعتبر اكثر جاذبية من غيره .

مثـــــال (٣) :

محتوى الفقرة : إن أفضل مصدر للحصول على فيتامين سي C هو : -

	المجموعة العليا	المجموعة الدنيا
	٢٧ ٪	٢٧ ٪
آ . الكبد	١	٤
ب . زيت كبد الحوت	صفر	صفر
جـ . عصير البرتقال	٢٦	٢٣
د . الرز غير المقشر	صفر	صفر

تعتبر هذه الفقرة سهلة جداً لأن عدد الذين اجتازوها بنجاح بلغ (٤٩) فردا من اصل (٥٤) ، أي أن ٩١ ٪ تقريبا أجابوا عنها بصورة صحيحة . وهذا يعني أن مستوى صعوبتها يساوي ٠,٩١ . كما أن بديلين من البدائل الاربعة لم يخترها أحد من بين الطلبة . واذا اريد تحسين هذه الفقرة وتعديلها بصورة أفضل فلابد من استبدال البديلين (ب) و (د) ببديلين آخرين اكثر جاذبية للطلبة فتزداد صعوبة الفقرة .

ان الغاية من حساب صعوبة الفقرة هو أختيار الفقرات ذات الصعوبة المناسبة ، وحذف الفقرات السهلة جدا والصعبة جدا . فمن الواضح أنه اذا لم يتمكن أحد من الاجابة عن الفقرة بصورة صحيحة فإنه ليس من المنطقي ابقاؤها في الاختبار . وينطبق الشيء نفسه على الفقرات التي يجيب عنها كل الافراد . إن كلا من هذين النوعين من الفقرات (السهلة جدا والصعبة جدا) لا تتيح لنا فرصة للتعرف على الفروق بين الافراد . وطالما إن هذه الفقرات لا تؤثر على التباين في درجات الاختبار فإنها لا تسهم في ثباته او صدقه . فكلما اقترب مستوى صعوبة الفقرة من (١) او (صفر) فإن قدرتها على التمييز بين الافراد تصبح قليلة جدا . وعلى النقيض من ذلك ، كلما اقترب مستوى الصعوبة من (٠,٥٠) كانت الفقرة اكثر قدرة على التمييز . ويكون الاختبار جيدا اذا تراوحت صعوبة فقراته بين ٠,٢٠ الى ٠,٨٠ ، وبمعدل يتراوح بين ٥٠ الى ٠,٦٠ . كما إن حساب صعوبة

(٥٢)

الفقرة يحقق غرضين آخرين ، اضافة الى ما ذكر اعلاه ، فنتائج صعوبة الفقرات يمكن أن تستخدم كوسيلة تشخيصية لمعرفة ما تعلمه الطلبة من موضوعات وما فشلوا في تعلمه ، فالفقرات التي يثبت أنها صعبة على جميع المجيبين ، تشير الى أن الموضوع الذي تنتمي اليه تلك الفقرة لم يكن المجيبون قد فهموه ، وإن الامر يتطلب من المدرس مزيدا من التقصي لمعرفة سبب الصعوبة .

العلاقة بين نوع الفقرة وصعوبتها : إن الطريقة المتبعة في حساب صعوبة الفقرة التي شرحناها اعلاه ، لا تنطبق على كل انواع الفقرات . اذ أنها تختلف في اختبار المقال عنه في الاختبار الموضوعي . فاختبار المقال يتطلب جوابا أوسع مما يتطلبه الاختبار الموضوعي . وتعتبر الصعوبة مناسبة في اختبار المقال عندما يكون ذلك الاختبار قادرا على الحصول على اجابات تختلف من حيث مقدار الافكار واكتمالها وتنظيمها . بينما نجد إن الصعوبة في الفقرة الموضوعية يمكن تحديدها بصورة اوضح من سؤال المقال .

معيار صعوبة الفقرة : رغم إن مستوى الصعوبة (٥٠٪) هو المعيار المفضل للوصول الى اقصى تمييز بين الافراد ، الا أنه في بعض الاختبارات يوضع حد معين ليكون المعيار الذي تنتقى على اساسه الفقرات . فعندما يراد استخدام اختبار لغربلة المرشحين المتقدمين لبعثات دراسية مثلا ، فإن مستوى صعوبة الفقرات يجب أن يكون أكثر من المعدل الاعتيادي للصعوبة وهو ٥٠٪ . كما أنه اذا أردنا انتقاء الطلبة الضعاف في احدى الموضوعات الدراسية لغرض وضعهم في صفوف تقوية فإنه تؤخذ الفقرات التي تقع صعوبتها دون المتوسط . وهكذا فإن اختبار مستوى الصعوبة المناسب يعتمد على الغرض من تصميم الاختبار .

ويتطلب تحقيق مستوى معتدل من الصعوبة في الاختبار ، أن يقوم مصمم الاختبار باعداد بعض الفقرات الصعبة جدا والتي لا يتمكن من الاجابة عنها الا عدد قليل من الطلبة تبلغ نسبتهم ٣٠٪ او ٤٠٪ ، كما يقوم ايضا باعداد فقرات سهلة يتمكن من الاجابة عنها ما بين ٨٥٪ – ٩٠٪ من الطلبة .

تنظيم فقرات الاختبار حسب الصعوبة : بعد أن يتم حساب صعوبة كل فقرة من الفقرات ، يقوم مصمم الاختبار بترتيب الفقرات في الاختبار حسب صعوبتها ، من السهلة الى الصعبة ، لكي يثق المجيب بنفسه وبقدرته على الاجابة عن فقرات الاختبار ، ولكي لا يضيّع الكثير من الوقت المخصص على فقرات قليلة تقع فوق قدرته ، فيما لو عرضت هذه الفقرات دون تدرج في الصعوبة .

ايجاد تمييز الفقرة item disicrimination

يقصد بـ « تمييز الفقرة » مدى قدرتها على التمييز بين الافراد الممتازين في الصفة التي يقيسها الاختبار ، وبين الافراد الضعاف في تلك الصفة . ولغرض حساب تمييز الفقرة نتبع ما يأتي : —

(١) اجراء الخطوات الثلاث الاولى التي وردت في ايجاد صعوبة الفقرة .

(٢) طرح عدد الافراد في المجموعة الدنيا الذين أجابوا عن الفقرة بصورة صحيحة ، من عدد الافراد في المجموعة العليا الذين أجابوا عن الفقرة بصورة صحيحة ، ثم تقسيم الناتج على نصف مجموع عدد الافراد في المجموعتين العليا والدنيا .

والقيمة الناتجة هي قوة « تمييز الفقرة » . إن هذه الطريقة ، التي وردت في هذه الخطوة ، هي احدى الطرق في ايجاد تمييز الفقرات وأبسطها . ومعادلة ايجاد قوة تمييز الفقرة هي : —

$$\text{ت} = \frac{\text{م}_{ع} - \text{م}_{د}}{\frac{١}{٢}\text{ك}}$$

حيث إن :

ت = قوة تمييز الفقرة

$\text{م}_{ع}$ = مجموع الاجابات الصحيحة للمجموعة العليا

$\text{م}_{د}$ = مجموع الاجابات الصحيحة للمجموعة الدنيا

$\frac{١}{٢}$ ك = نصف مجموع عدد الافراد في كل من المجموعتين العليا والدنيا .

مثال لحساب تمييز الفقرة :

لو عدنا الى المثال رقم (١) الذي ورد في ايجاد صعوبة الفقرة ، واستخدمنا معادلة تمييز الفقرة المذكورة اعلاه ، يتبين ما يلي : —

$$\text{قوة التمييز} = \frac{١٦ - ١١}{\frac{١}{٢}(٥٤)}$$

٥
= ——
٢٧

= ٠,١٩ ويعتبر هذا التمييز اقل من الحد الادنى المطلوب .

واذا كانت القيمة التمييزية للفقرة موجبة ، فإن ذلك يعني إن نسبة الطلبة الممتازين الذين أجابوا عنها بصورة صحيحة اكبر من نسبة الطلبة الضعاف الذين أجابوا عن هذه الفقرة بصورة صحيحة . واذا كانت القيمة تساوي (صفرا) ، فإن تمييز الفقرة يكون صفرا . وهذا يحدث في الحالات الآتية : —

(١) اذا كانت الفقرة سهلة جدا او صعبة جدا (اذا أجاب عنها كل افراد العينة بصورة صحيحة ، او اذا أخطأ كلهم في الاجابة عليها ، فإن التمييز يكون صفرا) .

(٢) اذا كانت الفقرة غامضة .

اما اذا كان عدد الطلبة الضعاف الذين أجابوا عن الفقرة بصورة صحيحة اكثر من عدد الطلبة الممتازين الذين أجابوا عنها بصورة صحيحة ، فإننا نحصل على تمييز سالب .

وبصورة عامة ، كلما كان تمييز الفقرة أعلى ، كلما كانت أفضل . وبالنسبة للاختبارات الصفية او المخصصة لمجموعات صغيرة بحجم الصف فإن المرء يتوقع أن يكون التمييز لمعظم الفقرات اكثر من (٠,٢٠) . اما اذا استخدمنا نسبة الـ ٢٧ بالمائة العليا والدنيا ، فإن المرء يتوقع الحصول على قيم تمييزية اعلى ، لأن الفرق بين المجموعتين العليا والدنيا يكون اكبر .

ويرى البعض إن الفقرة الجيدة يجب أن تكون قوة تمييزها ٠,٣٠ فأكثر . والجدول التالي الذي قدمه ايبل Ebel المختص في القياس النفسي والتربوي ، يستخدم كمعيار لمقارنة القوة التمييزية للفقرات التي يحصل عليها مصمم الاختبار .

دليل التمييز	تقييم الفقرات
٠,٤٠ فأعلى	فقرات جيدة جدا
٠,٣٠ الى ٠,٣٩	جيدة الى حد مقبول ولكنها يمكن أن تخضع للتحسين .
٠,٢٠ الى ٠,٢٩	فقرات حدية . تخضع عادة الى التحسين .
اقل من ٠,١٩	فقرات ضعيفة ، تحذف او يتم تحسينها .

فعالية البدائل الخاطئة effectivness of distractors :

في الاختبارات التي تحتوي على فقرات من نوع لاختيار من متعدد ، يحتاج مصمم الاختبار أن يقوم بفحص اجابات الطلبة على كل بديل من بدائل الفقرة . وهنا تتبع نفس الطريقة التي وضحناها في ايجاد قوة تمييز الفقرة ، وباستخدام نفس المعادلة المذكورة هناك .

ويسعى مصمم الاختبار الى الحصول على قيم سالبة للبدائل الخاطئة ، لكي تكون الفقرة جيدة . أي أنه يجب أن يكون عدد الطلبة الضعاف (المجموعة الدنيا) الذين يختارون البدائل او الاجابات الخاطئة اكثر من عدد الطلبة الممتازين (المجموعة العليا) الذين يختارون البدائل الخاطئة .

ويعتبر البديل الخاطىء جذابا عادة وصالحا ، اذا اختاره عدد من طلبة المجموعة الدنيا . وليس هناك اتفاق بين المختصين في الاختبارات والمقاييس على هذا العدد .

ان ما تقدم من موضوع يتعلق بتحليل الفقرات ، أي ايجاد كل من صعوبة الفقرات ، وقوتها التمييزية ، ينطبق على الاختبارات التحصيلية والذكاء والاستعدادات الخاصة . اما بالنسبة لاختبارات الاتجاهات والميول الشخصية ، فإنه يستخرج لفقراتها عادة « تمييز الفقرات » فقط . وذلك لأن الاجابة على فقراتها ليست فيها صحيح او خطأ .

ان تحليل الفقرات ، بما يتضمن من ايجاد صعوبة الفقرات وتمييزها وفعالية بدائلها الخاطئة ، يبين لنا في بعض الاحيان ، سبب عدم كفاءة بعض الفقرات في الحصول على الاجابة المطلوبة من الافراد ، وكيفية تطوير هذه الفقرات بشكل يجعلها صادقة في قياس ما يهدف اليه الاختبار .

كما إن تحليل الفقرات يهيء الفرصة امام واضع الفقرات لتقييم قدرته على تصميم الفقرات وتطوير هذه القدرة لديه . ويفضل أن يبتدىء الباحث فى تجربة التحليل بعدد كبير من الفقرات ، بحيث يستطيع استبقاء العدد الكافي الجيد منها بعد سقوط بعضها في العمليات الاحصائية .

(سابعا) : اجراء التجربة الاساسية لقياس صدق الاختبار وثباته :

تناولنا في الصفحات السابقة من هذا الفصل الاجراءات التي يجب أن يتبعها مصمم الاختبار للوصول الى تصميم جيد للاختبار . اما المرحلة التالية فهي تتناول الحكم على مدى فائدة الاختبار للاغراض التي وضع لأجلها ، وهذا ما نطلق عليه مفهوم « الصدق » . وعند ايجاد صدق الاختبار ، فإن بعض الباحثين يميلون الى ايجاد « صدق الفقرة » Item Validity لكل فقرة من فقرات الاختبار ، بالاضافة الى ايجاد « صدق الاختبار » Test Validity باكمله .

ومن بين الاساليب المتبعة لايجاد « صدق الفقرة » هو : ايجاد العلاقة الارتباطية بين درجة الفقرة ودرجة اختبار آخر (معيار خارجي) يقيس نفس الظاهرة النفسية التي تقيسها الفقرة . او ايجاد قوتها التمييزية بين مجموعات متطرفة تختلف في الظاهرة النفسية التي تقيسها الفقرة ، وقد اشرنا الى ذلك بشيء من التفصيل بهذا الفصل وتحت موضوع (تحليل الفقرات) . ومن الوسائل الاحصائية المستخدمة فى ايجاد التمييز ، بالاضافة الى ما ذكر هناك ، استخدام الاختبار التائي t-test او تحليل التباين ، وغيرها من الطرق .

اما بالنسبة لايجاد صدق الاختبار وثباته ، فقد فصلنا ذلك في الفصل الثاني من هذا الكتاب .

(ثامنا) : اشتقاق المعايير :

بعد انتهاء الباحث من ايجاد صدق الاختبار وثباته ، فإنه يقوم عادة باشتقاق المعايير له ، ليتمكن في ضوئها من التعرف على قيمة الدرجة التي يحصل عليها الافراد الذين يطبق عليهم الاختبار . وقد تحدثنا عن المعايير وانواعها وكيفية اشتقاقها في الفصل الثاني .