

الفصل

1

الرفاهية الأساسية  
الخاصة بقباس  
ونفوسهم نمو الطفل

يتناول هذا الفصل مدخلاً في قياس وتقويم نمو الطفل يتمثل في أهداف القياس و التقويم، والخلفية التاريخية للقياس والتقويم، والفرق بين القياس السيكمومتري والإكلينيكي، مواصفات المقاييس والاختبارات المناسبة لقياس وتقويم نمو الطفل .

## ● أولاً - مدخل إلى قياس وتقويم نمو الطفل

يقصد بتقويم نمو الطفل إصدار الحكم على مدى مناسبة علامات ومظاهر النمو المختلفة لعمر الطفل الزمني من الجنين وحتى بداية مرحلة المراهقة .

وحيث إن النمو النفسي للطفل يقصد به تلك التغيرات الجسمية والفسولوجية والاجتماعية والعقلية والانفعالية التي يمر بها الفرد ، فإن التقويم بالنسبة لنمو الطفل يقصد به التعرف على مدى السواء في النمو أي مطابقة المظاهر السلوكية في الجوانب المختلفة لنمو الطفل مع معايير النمو لمعظم الأطفال الذين هم في السن نفسه ، ويفيد ذلك في معرفة مظاهر التقدم أو التأخر في النمو بالنسبة للوقت المناسب لعمر الطفل من أجل التخطيط لبرامج إثرائية مناسبة للتدخل المبكر في حالات الضعف والتأخر في النمو أو تقديم البرامج المناسبة للفائقين في النمو المعرفي أو الاجتماعي .

ولكي يكون الحكم صادقاً وموضوعياً على نمو الطفل في جوانبه المختلفة فإننا نستند إلى القياس النفسي في نمو الطفل وإلى أدوات القياس والاختبارات الموضوعية التي ترجع إليها الدقة في الحكم وفقاً لسلامة إجراءاتها في التقنين والتطبيق واستخراج الدرجات وكيفية تفسيرها .

## ● ثانياً -أهداف قياس وتقويم نمو الطفل

لا شك أن دراسة تقويم سيكولوجية الطفولة يعتبر مهماً في حد ذاته ومفيداً بالنسبة لفهم مرحلة الطفولة ونمو شخصية الفرد بصفة عامة .

وترجع أهمية دراسة تقويم النمو من الناحية النظرية ومن الناحية التطبيقية، وبالنسبة لعلماء النفس وبالنسبة للمربين والوالدين وبالنسبة للأفراد والمجتمع إلى ما يأتي :

### من الناحية النظرية

- تزيد من معرفتنا بالطبيعة الإنسانية ومدى التأثير بالبيئة التي يعيش فيها .
- تؤدي إلى تحديد معايير النمو في كافة مظاهره وخلال مراحلها المختلفة ، مثل معايير النمو الجسدي والعقلي والانفعالي والاجتماعي في مرحلة ما بعد الميلاد، ثم مرحلة الطفولة.

### من الناحية التطبيقية

- تزيد من قدرتنا على توجيه الأطفال والمراهقين والراشدين والشيوخ، وعلى التحكم في العوامل والمؤثرات المختلفة التي تؤثر في النمو بما يحقق التغيرات التي نفضلها على غيرها، ويقلل أو يوقف التغيرات التي لا نفضلها.
- يمكن قياس مظاهر النمو المختلفة بمقاييس علمية ، تساعدنا من الناحية النفسية والتربوية في مساعدة الأفراد إذا ما اتضح شذوذ النمو في أي من هذه النواحي عن المعيار العادي.

### بالنسبة لعلماء النفس

- تساعد الاختصاصيين النفسيين في جهودهم لمساعدة الأطفال والمراهقين والراشدين والشيوخ، خاصة في مجال علم النفس العلاجي والتوجيه والإرشاد النفسي والتربوي والمهني.
- تعين دراسة قوانين ومبادئ النمو وتحديد معاييرها في اكتشاف أي انحراف أو اضطراب أو شذوذ في سلوك الفرد ، وتتيح معرفة أسباب هذا الانحراف وتحديد طريقة علاجه .

### بالنسبة للمربين

- تساعد في معرفة خصائص الأطفال وفي معرفة العوامل التي تؤثر في نموهم وفي أساليب سلوكهم ، وفي طرائق توافقهم في الحياة ، وفي بناء المناهج وطرائق التدريس وإعداد الوسائل المعينة في العملية التربوية .
- تؤدي إلى فهم النمو العقلي ونمو الذكاء والقدرات الخاصة والاستعدادات والتفكير والتذكر والتخيل والقدرة على التحصيل في العملية التربوية ، حيث توصل إلى أفضل طرائق التربية والتعليم التي تناسب المرحلة ومستوى النضج .
- تفيد في إدراك المعلمة للفروق الفردية بين الأطفال ، وأنهم يختلفون في قدراتهم وطاقاتهم العقلية والجسمية وميولهم ... وهكذا ، لا يكفي المعلم بالتربية الجماعية، بل يوجه انتباهه أيضاً إلى التربية الفردية حيث يراعي كل فرد حسب قدراته.

### بالنسبة للوالدين

- تساعد الوالدين في معرفة خصائص الأطفال مما يعينهم ، وينير لهم الطريق في عملية التنشئة والتطبيق الاجتماعي لأولادهم .
- تعين الوالدين على تفهم مراحل النمو والانتقال من مرحلة إلى أخرى من مراحل النمو، فلا يعتبرون الأطفال راشدين صغار، ولا يعتبرون المراهقين أطفالاً، وهكذا يعرفون أن لكل مرحلة

- من مراحل النمو خصائصها المميزة، حيث تنمو شخصية الفرد بمظاهرها المختلفة الجسمية والعقلية والانفعالية والاجتماعية .
- تتيج معرفة الفروق الفردية الشاسعة في معدلات النمو، فلا يكلف الوالدان الطفل إلا وسعه ولا يحملانه مالا طاقة له به ، ويكافئانه على مقدار جهده الذي يبذله وليس على مقدار مواهبه الفطرية .

### بالنسبة للأفراد

- تفيد بالنسبة للأطفال ، وهم راشدوا المستقبل ؛ فبفضل فهم أولياء الأمور والقائمين على التربية والرعاية النفسية والطبية والاجتماعية لمعايير تقويم النمو ، أصبح التوجيه على أساس دليل علمي ممكناً ، مما يحقق الخير للأفراد من الطفولة إلى الشيخوخة .
- تساعد على أن يفهم كل فرد - بقدر مستوى نموه - طبيعة مرحلة النمو التي يعيشها، ويعتبر أن عليه أن يحيها بأوسع وأكمل وأصح شكل ممكن ، باعتبارها غاية في حد ذاتها.

### بالنسبة للمجتمع :

- يفيد فهم الفرد ونموه النفسي وتطور مظاهر هذا النمو في المراحل المختلفة في تحديد أفضل الشروط الوراثية والبيئية الممكنة التي تؤدي إلى أفضل نمو ممكن ، وحتى لا يخطئ في تفسيره، تحقيقاً لخير الفرد وتقدم المجتمع.
- تعين على فهم المشكلات الاجتماعية والانفعالية وثيقة الصلة بتكوين ونمو شخصية الفرد والعوامل المحددة لها ، مثل : مشكلات التأخر العقلي ونقص الانتباه وفرط النشاط ، ومشكلات التوحد والعدوان وتقديم برامج التدخل المناسبة لها.
- تساعد في ضبط سلوك الفرد وتقويمه في الحاضر بهدف تحقيق أفضل مستوى ممكن من التوافق الشخصي والتربوي والاجتماعي والمهني ، بما يحقق صحته النفسية في الحاضر والمستقبل كإنسان منتج وفعال في المجتمع الذي يعيش فيه .

## ثالثاً- الخلفية التاريخية للقياس والتقويم

حتى الآن، لم ننظر فيما إذا كان هناك أي فرق بين مصطلح الاختبار ومصطلح المقياس. وعلى الرغم من أن هناك تداخلاً في المعنى إلا أنهما ليسا مترادفين تماماً، فمصطلح المقياس يستخدم في ميادين كثيرة من ميادين البحوث النفسية حيث لا يكون مصطلح الاختبار مناسباً للاستخدام. فعلى

سبيل المثال ، في التجارب التي تدرس الإحساس ، والإدراك الحسي، والحكم ، استخدم الباحثون كثيراً من الطرائق النفسية الفيزيائية (السيكوفيزيائية)، بمعنى قياس الحكم الفيزيائي المقابل للحكم النفسي (مثل درجة وضوح الضوء ، ودرجة ارتفاع النغمة). فإذا كان موضوع الدراسة هو البحث عن الحد الأدنى والحد الأعلى لقوة السمع عند الإنسان ، فإنهم يقيسون مقدار التردد . وعليه ، فإن القياس الفيزيائي يستخدم للإجابة عن سؤال سيكولوجي .

ومن المعتاد أن نقول عن القياس النفسي أنه اختبار إذا كنا نستخدمه أساساً لتقدير بعض خصائص الفرد أكثر مما نجيب عن سؤال عام ، مثل ذلك السؤال عن قوة السمع عند الإنسان في الفقرة السابقة . ومن الممكن - بالطبع - استخدام مقاييس درجات عتبات السمع كاختبارات بهذه الطريقة . غير أن الاختبار - في الأغلب - عبارة عن أسئلة أو مهام تقدم للشخص المطلوب اختباره ، كما أن الدرجات التي نحصل عليها لا يعبر عنها بوحدات فيزيائية من أي نوع .

وعلى ذلك ، فليست المقاييس اختبارات ، وكذلك ، فإن العكس صحيح أيضاً، فليست كل الاختبارات مقاييس. فهناك بعض اختبارات الشخصية - مثلاً- لا يحصل المفحوص فيها على درجات. وقد يستخدم الاختصاصي النفسي مثل هذا الاختبار كي يساعده على إعداد وصف لفظي للفرد . وهو هنا ليس في حاجة إلى أن يلجأ إلى القياس في أي مستوى من المستويات ، فالقياس هو عملية استخدام الأرقام في تحديد الأشياء حسب قواعد معينة. و «الاختبار» يمكن أن يعرف على أنه موقف مقنن صمم خصيصاً للحصول على عينة من سلوك الفرد. وعندما يعبر عن هذه العينة بالأرقام، فمن الممكن أن نستخدم كلمة اختبار أو مقياس .

وعلى الرغم من أن التداخل بين هذين المفهومين ليس كاملاً، فإنه لا يزال بإمكاننا أن نقول أن معظم الاختبارات هي وسائل قياس ، وأن معظم المقاييس النفسية يمكن أن تستخدم كاختبارات . وسوف نعالج في هذا الفصل بعض الأفكار التي تنطبق على الاختبارات على وجه خاص .

يشير سعد عبد الرحمن (1998) إلى أنه منذ أن بدأ علم النفس كعلم - أي منذ إنشاء «ولهم فونت» (Wilhelm Wundt) لمختبره في مدينة «لايبزج» (Leipzig) في عام (1879) - ركز الباحثون جهودهم ابتداءً على البحث عن مبادئ عامة يمكن أن تطبق على كل فرد . فقد حاول الاختصاصيون النفسيون اكتشاف كيفية عمل عين الإنسان ومخه معاً لإدراك اللون والشكل والحجم، وهي إدراكات حسية تحدث لنا جميعاً. كما صمموا التجارب لدراسة عملية التعلم ، ومن ثم، استنتجوا «قوانين التعلم». كما درسوا أيضاً نمو الإنسان، وذلك بمقارنة مجموعة ذات مرحلة عمرية معينة بمجموعة أخرى ذات مرحلة عمرية أخرى ، وقاموا بتحديد معايير السلوك لكل مرحلة عمرية، ووضعوا نظريات حول عملية النمو .

إلى جانب ذلك، كان هناك اتجاه ثانوي للاهتمام والجهد، فقد تنبه بعض الباحثين الأوائل في المختبرات الألمانية لعلم النفس التجريبي الحديثة النشأة، إلى الاختلاف الذي يحدونه بين الأفراد المفحوصين . فهؤلاء الباحثون الذين كانت لهم في كثير من الأحيان اتجاهات عملية، أدركوا أن قياس الفروق بين الناس قد تكون له تطبيقات مهمة في المدارس، والمصانع والمكاتب. وقد كان «جيمس ماكين كاتل» (James Makeen Cattell) - وهو أمريكي درس في مختبر «فونت» - ذا تأثير قوي في حركة استخدام المقاييس النفسية كاختبارات عقلية. وقد كان هو- في الحقيقة- أول من استخدم لفظ «الاختبارات العقلية» في عام (1890)، ولكن، كان هناك آخرون ظهوروا بوضوح في هذه الحركة خلال السنوات الأخيرة من القرن التاسع عشر.

إن ما كان يتطلع إليه هؤلاء العلماء كثيراً هو إيجاد طريقة كمية لتقدير الذكاء العام. وكانوا يعتقدون أنه يمكنهم الحصول على دليل للذكاء إذا أمكنهم قياس جميع الخصائص وهي متجمعة معاً في الأفراد ، تلك التي كان يقاس كل منها على حدة في المختبرات التجريبية (الإحساس ، والإدراك الحسي ، والانتباه ، والتمييز ، وسرعة رد الفعل )، وهكذا . والإنسان المتفوق - حسب تفكيرهم هذا - هو الشخص الذي يحصل على درجات عالية في كل هذه الخصائص. ورغم أنه ليس هناك خطأ واضح في هذا التفكير ، إلا أن محاولة قياس الذكاء بهذه الطريقة لم تنجح ، وذلك لأنه عندما قام الاختصاصيون النفسيون بتحليل مقاييسهم اكتشفوا أن هذه السمات لم يرتبط بعضها ببعض ارتباطاً وثيقاً . فضلاً عن ذلك، فإن مجموع الدرجات لم يبد أنه من الممكن اعتباره دليلاً على الذكاء العام ؛ فالتلاميذ الضعفاء - على سبيل المثال- حصلوا على درجات عالية شأنهم في ذلك شأن التلاميذ المتفوقين . لقد أصبح واضحاً أن مشكلة قياس الذكاء كانت بحاجة إلى أسلوب مختلف .

وسرعان ما حدث هذا الانطلاق الجديد في فرنسا ، عندما بدأ «ألفرد بينيه» (Alfred Binet) بافتراض أن الذكاء خاصية مركبة فطرية ، وليس مجرد مجموعة سمات بسيطة كثيرة. ولقياس الذكاء من وجهة نظر «بينيه» - لابد من إيجاد وسائل لتقييم كيفية تعامل الأفراد مع المواقف التي تتطلب التفكير ، والحكم ، وحل المشكلات . وعلى مدى سنوات عديدة قام «بينيه » بتجريب أنواع كثيرة من المهام كاختبارات محتملة لذكاء الأطفال على أطفاله وعلى أطفال من مختلف الأعمار في المدارس الفرنسية في باريس. وقام أخيراً في عام (1905)- بالتعاون مع «تيوفيل سيمون» (Theophile Simon) - بنشر أول اختبار حقيقي للذكاء، والذي يعتبر الأصل بالنسبة لكل الاختبارات الحالية .

وبعد ذلك ، حدث تقدم سريع ، فقد عدل اختبار «بينيه» و«سيمون» ليستخدم في كثير من الدول، كما صممت اختبارات للكبار بالإضافة إلى الاختبارات التي صممت خاصة للأطفال. وكذلك ، وضعت الاختبارات الجمعية أثناء الحرب العالمية الأولى ، ثم طورت سريعاً بعد ذلك لتستخدم في

المدارس والمؤسسات الصناعية . وصممت كذلك اختبارات لكثير من القدرات التي ليست في مثل شمول وعمومية الذكاء. وأخذت محاولات قياس خصائص الشخصية وكذلك القدرات تزداد انتشاراً . وفي منتصف القرن العشرين توفرت آلاف من الاختبارات على اختلاف أنواعها.

أثناء كل هذا النشاط في ميدان الاختبارات العقلية، ظهرت المعايير والمبادئ، وهي تستخدم الآن كإرشادات مفيدة لمن يريد أن يبني اختبارات جديدة ، أو لمن يحتاج أن يختار بحكمة من الاختبارات الموجودة. وفضلاً عن ذلك ، فقد أصبح من الضروري - في عالمنا الحديث الذي يشعر بأهمية الاختبارات - أن يعرف كل رجل مثقف أو امرأة مثقفة شيئاً عن هذه المعايير والمبادئ . فالمدرس الذي يريد أن يعدل طريقته التدريسية لتلائم الخصائص المتباينة لتلاميذه، والأب الذي يرغب في مساعدة ابنه أو ابنته على القيام بتخطيط حكيم يتعلق بالدراسة أو المهنة ، ورجل الأعمال الذي لا يريد أن يخدعه بعض من لا خلق لهم من ناشري الاختبارات وبائعها، كل هؤلاء في حاجة لمعرفة عملية عن كيفية بناء الاختبارات، وعن كيف يمكن الحكم عليها . وفي مجال النمو النفسي تفيد الاختبارات في معرفة معايير النمو ومدى التقدم أو التأخر (كما سبق ذكره في مقدمة هذا الفصل).

## رابعاً- مواصفات الاختبار المناسب لقياس وتقويم نمو الطفل

إن أهم ما يؤخذ بعين الاعتبار هو الصدق الذي يتصل مباشرة بالسؤال الذي يقول: «ماذا يقيس الاختبار؟» وإذا لم تكن لدينا إجابة كافية لهذا السؤال، فإن أي اختبار سوف يكون عديم الفائدة في محاولتنا للتعامل بدرجة سواء أو عدم سوية نمو الطفل. بل ربما كان الموقف أكثر سوءاً من عدم الفائدة ، لأننا لو اعتمدنا على الافتراض الخاطئ فيما يتعلق بمعنى درجات الطفل ، فقد نوجهه إلى قرارات قد تؤدي إلى تغيير مسار النمو أو زيادة درجة الإعاقة وال فشل في التعامل بإيجابية مع المجتمع إزاء حالة الإعاقة .

وهناك كثير من الأشخاص لا يستطيعون معرفة أن عنوان الاختبار لا يقدم لنا في الحقيقة شيئاً على الإطلاق عما يقيسه هذا الاختبار . فكل إنسان يستطيع أن يكتب مجموعة من الأسئلة ، ويكون واثقاً من أن هذه الأسئلة تدور حول عملية «التفكير» أو «الاستعداد الميكانيكي» أو «المرونة المعرفية». ولكن معرفة أي العمليات العقلية يجب أن يستخدمها المفحوص في الواقع للإجابة عن هذه الأسئلة - هي مهمة شاقة وطويلة . فقد يبني شخص ما اختباراً لقياس «القدرة على التفكير الاستدلالي» ، وربما يثبت أن هذا الاختبار إنما يقيس اتجاهات الطبقة المتوسطة . وقد يحدث أن اختباراً في «الاستعداد الميكانيكي» قد يقيس فعلاً الذكاء العام ، وأن اختباراً في «النضج الانفعالي» قد يقيس فقط ما يعرفه المفحوصون عن التقبل الاجتماعي لنمط معين من السلوك. إن المهارات الأولى

التي يجب أن يجيدها الفرد عند تقييم صدق الاختبارات هي عدم أخذ عناوين الاختبارات بعين الاعتبار . فالأمر المهم ليس التسمية التي يضعها المؤلف لهذا الاختبار ، ولكن ماذا تتعلق درجات هذا الاختبار .

وفي بداية حركة القياس كانت الطريقة المقبولة هي أن يعرف الفرد ماذا يريد أن يقيس، ثم يجمع بعد ذلك الأدلة والبراهين لإثبات مدى نجاح الاختبار في قياس ما يريده الباحث. وخلال هذه الفترة كان معظم الجهد موجهاً لقياس الذكاء العام. وقد كان واضحاً أنه على الرغم من تزايد عدد الباحثين ، إلا أنهم لم يتفوقوا جميعاً على تعريف محدد لهذا المفهوم المحير ، ولكنهم قد افترضوا جميعاً - من أجل الأغراض العملية- أن الذكاء هو السمة التي يعتمد عليها الحكم بالتفوق أو الغياب في الدراسة ، ومن ثم ، بحثوا في المواقف المدرسية عن البرهان على صدق اختبارات الذكاء. وكان الطريق هو تقدير مقدار العلاقة بين درجة النجاح في الدراسة - كما يمكن أن يتم التنبؤ بها من درجات الاختبار - ودرجة النجاح الحقيقية . وبناء على الأدلة التي حصلوا عليها ، تم استبعاد القياسات المخبرية مثل التمييز، وزمن الرجوع ، وما إلى ذلك ، واستبقيت الأسئلة من نوع اختبار «بينيه» ذات الاستجابات المركبة لتكون أساساً لاختبارات الذكاء .

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

إذاً، كان الصدق يعني مدى قياس الاختبار لما وضع لقياسه. ولهذا ، كان يحسب معامل الارتباط بين درجات الاختبار وبين مقياس المحك لمعرفة صدق الاختبار. ولكن، بمضي الوقت اتضح أن هناك تعقيدات في كل من الطريقة والمفهوم. ومما يؤسف له أنه لم يكن من الممكن في معظم الحالات أن نجد محكاً واحداً يمكن أن يعتبر دليلاً واضحاً على سمة عقلية ، فقد يقرر باحثان يدرسان القدرة نفسها - مثل الاستعداد الميكانيكي - استخدام محكات مختلفة ، ومن ثم، قد يحصل كل منهما على نتيجة تختلف عن الآخر . فالباحث الأول قد يرى أن المحك المناسب هو درجات القياس في مقرر الأعمال (الأشغال) اليدوية. والباحث الثاني قد يرى أن المحك المناسب هو طول الوقت الذي يستغرقه الأفراد الجدد لتعلم مهارة ميكانيكية بسيطة في المصنع . والآن ، ماذا نعمل عندما نجد أن معامل الارتباط بين درجات الاختبار والمحك هو في الحالة الأولى 0.06 وفي الحالة الثانية (0.59)؟ كيف يتسنى لنا أن نقرر ما إذا كان الاختبار صادقاً إذا أعطانا نتائج من هذا النوع ؟ هل هذا الاختبار يقيس الاستعداد الميكانيكي أم لا ؟

يتضح مما سبق أن عملية التحقق من صدق الاختبار هي عملية طويلة ، وهي لا تقوم على ملاحظة مفردة . ونحن نستطيع فقط أن نفهم ماذا يقيس الاختبار حقاً إذا قمنا بدراسة ارتباطات



الاختبار مجموعة مختلفة من المحكات. وإذا توفرت سلسلة من البحوث حول اختبار «القدرة الحركية» - على سبيل المثال - فقد توضح أنه يقيس في الحقيقة القدرة على القيام بالحركات الماهرة ذات الضبط الدقيق، ولا علاقة له بالقدرة على فهم العلاقات المعقدة للمفاهيم التبولوجية ومن ثم، فقد يرتبط هذا الاختبار ارتباطاً عالياً مع الأعمال الفنية ، ولكنه لا يرتبط ارتباطاً عالياً مع درجات المفاهيم التبولوجية (الخاصة بالوضع والمكان والاتجاهات). وقد يساعد هذا الاختبار على اختيار أنواع مختلفة من التعليم الفني عن التعليم الأكاديمي العادي .

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

إن فهم مضمون عملية التحقق من صدق الاختبارات يحمل ثلاثة معان واضحة لمستخدمي الاختبارات : أولها - أنه إذا كان الاختبار سوف يستخدم في اتخاذ قرارات حول الأفراد أو الجماعات، فلا بد من دراسة كل البيانات المتوفرة قبل أي محاولة لتفسير درجات الاختبار. وهذا يعني أن من يستخدم الاختبار يجب أن يكون قادراً على إجراء مثل هذه الدراسة. «برنامج تطبيق الاختبارات» في المدارس أو الروضات يمكن أن يؤدي وظيفته بفعالية في حالة ما إذا أشرف عليه شخص على درجة عالية من القدرة والكفاءة .

والمعنى الثاني هو أنه عندما يستخدم الاختبار من أجل التنبؤ أو الاختبار (مثل اختيار المتقدمين لعمل ما)، فإنه يجب أن يعين صدفة في الموقف النوعي الذي سوف يستخدم فيه، كلما كان ذلك ممكناً . وهذا يتطلب - بطبيعة الحال - دراسة خاصة في الموقف نفسه الذي يستخدم فيه الاختبار، وذلك قبل أن نبدأ باستخدام الاختبار بصورة روتينية في انتقاء المتقدمين. وفي حالة التخطيط لمثل هذه الدراسة فإن السؤال المناسب الذي علينا أن نسأله فيما يتعلق بصدق الاختبار لن يكون : «هل هذا الاختبار صادق؟» ، ولكنه سوف يكون : «هل يضيف هذا الاختبار جديداً لصدق الطرائق المستخدمة حالياً في عملية الاختبار؟».

فإذا كنا نستخدم الاختبارات في المواقف العملية التي نتعامل فيها مع الأفراد، أو في البحوث البحتة التي تهدف إلى زيادة معلوماتنا النظرية عن الفروق الفردية النمائية بين الأطفال العاديين وذوي الإعاقات النمائية فيجب أن نتذكر دائماً أن أفكارنا عن ماهية المعايير النمائية وكذلك أفكارنا عما تقيسه الاختبارات يجب أن تتعدل إذا أتيح لنا الحصول على أدلة جديدة .

فعندما لا ينجح الباحث في الحصول على معامل الارتباط المتوقع بين «فرط النشاط» و «ذوي النشاط الزائد»، فقد يحتاج إلى تغيير أفكاره عن الشيء الذي يقيسه اختبار فرط النشاط الذي استخدمه في بحثه. وفضلاً عن ذلك ، قد يكون عليه أيضاً أن يعدل من نظريته حول كيفية تأثير الشخصية بفرط الحركة . وعلى الذين يستخدمون الاختبارات الخاصة بكثرة- مثل المعلمين في المدارس العامة- أن يتابعوا عن كثب المعرفة الحديثة عن هذه الاختبارات، وكذلك ، عن الخصائص التي تقيسها هذه الاختبارات، إذ إن هناك كثيراً من المعلمين في المدارس حالياً- على سبيل المثال - يتخذون القرارات بالنسبة للأطفال بناء على مفاهيم قديمة للذكاء. والحقيقة أن هؤلاء المعلمين ليسوا في حاجة فقط إلى اختبارات في الذكاء على درجة أكبر من الصدق ، ولكنهم يحتاجون كذلك إلى معرفة أكثر دقة لمعنى كلمة الذكاء ، فنحن لا يمكن في الحقيقة أن نبحث في صدق الاختبار دون أن نخبر في الوقت نفسه صدق أفكارنا.

### معنى ثبات الاختبار :

لتوضيح معنى الثبات، يمكننا أن نبدأ مرة أخرى بسؤال وهو : «بأي قدر من الدقة ، أو بأي قدر من الاتساق ، يقيس هذا الاختبار السمة التي يقيسها ، أيًا كانت هذه السمة ؟

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

إن ما يجب على الاختصاصي الذي يبني الاختبار أن يكتشفه ويوصله إلى الاختصاصي الذي يستخدم الاختبار هو إلى أي مدى سوف تتأثر دقة تقدير درجات الأفراد بعوامل الصدفة هذه. فإذا توفرت هذه المعلومات، استطاع الاختصاصي الذي يفسر درجات الاختبار أن يأخذ ذلك في اعتباره، وسوف تكون خلاصة تقريره عن شخص ما على النحو الآتي : الفرد (أ) ذكاؤه فوق المتوسط بقليل ، هناك احتمال (50 - 50) أن نسبة ذكائه تقع بين (105 و 115). ولكنه لن يستنتج : أن الشخص (أ) له نسبة ذكاء تساوي (110). (سعد عبد الرحمن ، 1998).

وهناك طرائق عديدة لتقدير ثبات الاختبار وإعداد نتائج التحليل ، ولكن معظمها يبدأ من منطلق واحد : وهو تطبيق صورتين من الاختبار نفسه على مجموعة من الأفراد لهم الخصائص العمرية والاجتماعية نفسها لمجتمع الأفراد الذي يصمم له الاختبار . فإذا كنا نريد أن نعين درجة ثبات اختبار ذكاء خاص بأطفال سن الثامنة ، فإننا نطبق الصورة (أ) والصورة (ب) على مجموعة أطفال في السن نفسه وفي الوقت نفسه. ولكن ، إذا لم تتوفر غير صورة واحدة، فإننا نقسمها إلى نصفين - وعادة ما يكون النصف من الأرقام الفردية ، والنصف الآخر من الأرقام الزوجية - لكل منهما مفتاح خاص للتصحيح . ومن ثم، فإننا نحصل على درجتين لكل فرد ، بدلاً من درجة واحدة . وإذا كنا نريد تعيين درجة ثبات اختبار في مهارة استخدام الأصابع لتلاميذ في مدرسة مهنية ، فغالباً، سوف نطبق الاختبار مرتين على المجموعة نفسها من التلاميذ ، مع مراعاة وجود فاصل زمني قصير بين التطبيقين (أسبوعين).

ولحساب مدى دقة درجات الاختبار تستخدم طرائق حساب الخطأ المعياري للقياس. إذا استطعنا أن نحصل على عدد كبير من «العينات» من درجات هذا الاختبار بأن نأخذها من «المجتمع» الكلي لدرجات هذا الاختبار . هذا المدى هو منطقة عدم الدقة الواقعة على كلا جانبي الدرجة . أما عن الأسس المنطقية والإحصائية لهذه الطريقة فهي معقدة جداً بحيث لا يمكن التعرض لها هنا ، ولكن التفسير النهائي لها بسيط نسبياً بحيث يمكن فهمه .

إذا كان الخطأ المعياري لاختبار في الذكاء - على سبيل المثال - هو (4.6) ، فإن هناك احتمالاً بمقدار (2/3) بأن الدرجة الحقيقية للفرد تقع في مدى (4.6) نقطة من الدرجة التي حصل عليها فعلاً . فإذا كان الفرد (أ) قد حصل على درجة (154) ، فإنه يمكن القول بدرجة من الثقة أن «درجته الحقيقية» - خالصة من تأثير عوامل الصدفة - تقع بين (149.4 و 158.6) وهذا يبني على حقيقة أنه في أي توزيع اعتدالي نجد أن (2/3) من الحالات تقع في مدى انحراف معياري واحد من المتوسط. وكما أوضحنا سابقاً ، فإن الخطأ المعياري هو نوع خاص من الانحراف المعياري . ولكن ، مازال هناك احتمال بمقدار (1/3) بأن «درجته الحقيقية» قد تكون أعلى أو أقل من ذلك . ونستطيع إذا شئنا -