

القياس النفسى والتربوى

التعريف بالقياس ومفاهيمه وأدواته
بناء المقاييس ومميزاتها
القياس التربوى

. تأليف

الدكتور محمد عبد السيد أحمد

ملزمة الطبع والنشر
مكتبة النهضة المصرية
لأصحابها حسن محمد وأولاده
٩ شارع مصرى باشا بالقاهرة

القياس النفسى والتربوى

التعريف بالقياس ومفاهيمه وأدواته
بناء المقاييس وميزاتها
القياس التربوى

تأليف

الدكتور محمد عبد السيد محمد

ملقمة الطبع والنشر
مكتبة النهضة المصرية
لأصحابها حسن محمد وأولاده
٩ شارع مدني باب شرق القاهرة

الطبعة الأولى

١٩٦٠

القاهرة

مطبعة لجنة التأليف والترجمة والنشر

مقدمة

موضوع هذا الكتاب هو أهم موضوعات علم النفس بل هو العماد الذى يقوم عليه علم النفس كعلم Science فلم يتحول علم النفس من ميدان الفلسفة إلى ميدان العلوم إلا عند ما أخذ الباحثون فيه بالمنهج العلمى ، ذلك المنهج الذى يعتمد أساساً على الملاحظة الموضوعية . فكما قال كارل بيرسون Karl Pearson العالم الرياضى المشهور « إن الأساس الذى تقوم عليه العلوم ليس طبيعة مادتها أو موضوعها ولكن الأسلوب العلمى أو الطريقة العلمية التى تتبعها » . ويقاس تقدم أى علم من العلوم بدرجة الدقة التى يصل إليها فى القياس . على أن عملية القياس فى علم النفس والتربية أعقد وأشق كثيراً منها فى العلوم الأخرى ، نظراً لأن موضوع القياس وهو الإنسان ، سلوكه ونواحي حياته العقلية والوجدانية ، هو أرقى الكائنات وأعلاها تعقيداً وأكثرها مرونة وقابلية للتغير كما أنه أقلها خيماً قابلية للتحكم والتجريب . وبالرغم من ذلك كله فقد تقدم القياس فى علم النفس والتربية تقدماً كبيراً فى السنوات الأخيرة ؛

ومما ساعد على تقدم القياس فى علم النفس والتربية استخدامه علم الإحصاء الذى ساهم فى تطوير كثير من العلوم الأخرى . على أن المشتغلين بالقياس لا يقتصرون على استخدام ما يقدمه الإحصاء بل لقد استنبطوا لأنفسهم طرقاً إحصائية خاصة مثل طرق التحليل العاملى التى استعانوا بها فى التعرف على العوامل النفسية والعقلية . ثم شاع استخدام هذه الطرق فى كثير من العلوم الأخرى .

ويهدف هذا الكتاب إلى تزويد الطالب في علم النفس والتربية والمستغل بالقياس فيهما بالحقائق والنظريات الأساسية في موضوع القياس والتقييم . إذ يجيب عن تساؤلات من مثل : متى نلجأ إلى القياس ؟ ولماذا نستخدم المقاييس ؟ ومن هو مستخدمها ؟ ومتى يحق لنا أن نستخدمها ؟ وما هي شروط إجراء الاختبار ؟ وما هي شروط إنشاء المقياس ؟ وما هي خصائص المقياس الجيد ؟ وكيف نفاضل بين أدوات القياس المختلفة ؟ وكيف نفسر نتائج القياس ؟ وما معنى النتائج والأرقام التي يوصلنا إليها المقياس ؟ وإلى أى درجة يكون الاعتماد عليها ؟ بل وكيف ننشئ برنامجاً شاملاً للاختبارات في المدرسة ؟ وكيف نفسر درجات هذه الاختبارات المقننة ودرجات التلاميذ في الامتحانات المدرسية ؟ وما هي عيوب الطرق الحالية في الامتحانات المدرسية ؟ وكيف نتجاشاها إيجابياً بمعنى كيف نصمم طرقاً علمية موضوعية لتقييم الطالب والمدرسة والمدرس ؟

كل هذه التساؤلات وغيرها من المشاكل التي تدور بذهن من يود تقدير الأفراد والجماعات تقديراً موضوعياً كمياً وكيفياً دقيقاً سيجد لها القارئ إجابات وحلولاً . هذا وسنعرض لما حققه القياس من النتائج في بلادنا والخارج . وبالطبع لن نقدم دليلاً جامعاً لكل أدوات القياس في بلادنا والعالم . ولكننا سنقتصر على أهمها وأكثرها دلالة متناولين مشاكل ومناهج وأهداف القياس في التربية وقياس القدرات العقلية والمعرفية والقدرات الخاصة والاستعدادات الفارقة وقياس الشخصية ، عارضين في كل هذا النماذج من الاختبارات الشهيرة في كل ناحية مفسرين وناقدين ومقيمين لطرق إنشائها واستخدامها .

ونود هنا أن ننبه إلى أن للقياس أدوات أخرى فضلاً عن الاختبارات

سنعرض لأهمها ، ومنها طرق التقدير والملاحظة في الظروف غير العادية « الاختبارات الموقفية » وطرق التقرير الذاتي وبطاقات التقويم والطرق الإسقاطية والاستفتاء كما يستخدم في اختبارات الاتجاهات إلى آخر هذه الجوانب الهامة من الحياة الإنسانية ونواحي ملاحظتها وتقدير النمو والاختلال فيها .

وسنعرض ، فيما نعرض ، للبحوث العلمية والجهود المثمرة التي تقوم بها الهيئات والأفراد حتى نعكس للقارئ صورة مصغرة طبق الأصل من مناهج معالجة المشاكل المتضمنة في موضوعات هذا الكتاب عسى أن يزول تصور بحث مشاكلنا علمياً على أنه عمل مستحيل ، فتُحفز همهم وتُدفع كفاءات وتنشط جهود من تسمح له الفرص وتتوفر له الظروف فيقوم بمعالجة ميدان استشكل مهتدياً بهذه البحوث مستعيناً بأدوات القياس كما سيعرف عنها الكثير .

وفي القسم الأول علاج أصول القياس ونظرياته ومفاهيمه الأساسية ومصطلحاته العربية والأجنبية وأدواته واختباراته مصنفة معرفة . وفي القسم الثاني تناول لكيف تنشأ الاختبارات وفقراتها وكيف يبحث صدقها وثباتها وما هي طرق اشتقاق معاييرها ثم كيف تحلل فقراتها . أما القسم الثالث فهو يعرض للقياس التربوي مشاكله ومصطلحاته والتعريف به ، وتقييم المدرس ، وتقييم التلميذ ، ومشكلة الامتحانات ، وبرنامج الاختبارات في المدرسة ، وأصول ومميزات ومشاكل الاختبارات التحصيلية ، ثم إنشائها ، وتمثيل باختبارات عامة ، ثم أخرى خاصة . وينتهي القسم الثالث بعرض بحوث واختبارات محلية في ميادين القياس التربوي .

وهكذا ينتهي بنا المجلد الأول من الكتاب

(و)

الإلمام بالقياس في التربية يجب أن ندرس القياس ، النفسى ، أى أن المجلد الأول ينتهى ومن هنا يبدأ المجلد الثانى الذى سيغطى كل جوانب القياس وتطبيقاته وأهمها تطبيقاته التربوية . ففى المجلد الثانى عود على بدء المجلد الأول . والله أسأل أن يكون الكتاب قد أدى ما هدفنا له .

والى لقاء قريب فى المجلد الثانى بإذن الله :

محمد عبد السلام الصمير

فهرس المجلد الأول

صفحة

مقدمة ز

القسم الأول

للتعريف بالقياس ومفاهيمه وأدواته (١ - ١٢١)

للفصل الأول : التعريف بالقياس (٣ - ٣٦)

٥	مقدمة
٦	نشأة القياس في علم النفس والتربية
٩	المشكلات التربوية
١١	نشأة القياس في التربية
→ ١٢	لماذا نقيس
١٥	لماذا نقيس
(١٨ - ٢٢)	تطبيقات القياس النفسى والتربوى
٢٠ - ١٨	أولا : في التربية والتعليم
٢١ - ٢٠	ثانيا : في الصناعة
٢٢ - ٢١	ثالثا : في الجيش
٢٢	رابعا : في علم النفس الإكلينيكي
٢٢	خامسا : في الإدارة
(٢٤ - ٢٢)	أغراض القياس النفسى والتربوى
٢٢	أولا : المسح
٢٣	ثانيا : التنبؤ
٢٣	ثالثا : التشخيص
٢٤	رابعا : العلاج
(٢٦ - ٢٤)	تحديدات القياس
(٢٩ - ٢٦)	من يقيس ومتى يقيس
٢٦	١ - تأهيل الفاحص
٢٨	٢ - ألفة المفحوص بالاختبار

صفحة

ماذا يجب مراعاته في تطبيق الاختبارات (٢٩ - ٣٥)	
١ - النوافع والعلاقة الطيبة ٢٩	
٢ - الغش والتزييف ٣٢	
٣ - الإعداد لموقف الإختبار ٣٤	
المبادئ الخلقية في استخدام الاختبارات النفسية والتربوية ٣٦	

الفصل الثاني : المفاهيم الأساسية في الاختبارات

للفنسية والتربوية (٣٧ - ٨١)	

٣٩	مقدمة
٤١	الدرجة الخام وتفسيرها
٤٥	النزعة المركزية
٤٦	التشتت
٤٨	الدرجات المحولة وأنواعها
٥١	الدلالة
٥٢	العينات
٥٤	عود إلى الدلالة
٥٥	معامل الارتباط
٥٨	التحليل العامل
٥٩	تحليل التباين
٥٩	القياس والتقييم

قائمة المصطلحات والمفاهيم الأساسية في القياس

للفنسى والتربوى (٦١ - ٨١)	

الفصل الثالث : تصنيف وتعريف الاختبارات

للفنسية والتربوية وعرض للمشاكل المتعلقة بها (٨٣ - ١١٨)	

٨٥	مقدمة

(٨٦ - ٩٠)	أسس تصنيف الاختبارات النفسية والتربوية
(٩١ - ١١٨)	التعريف بالأنواع الكبرى من الاختبارات النفسية والتربوية

صفحة

أولا : التصنيف على أساس ما نقيسه	٩١ - ١٠٠
اختبارات الذكاء	٩١
اختبارات التصنيف العام	٩٢
اختبارات الاستعدادات الخاصة	٩٣
اختبارات التحصيل بأنواعها (تحصيل - كفاية - حرفة)	٩٤
اختبارات الاستعدادات الفارقة	٩٥
اختبارات الشخصية	٩٧
ثانيا : التصنيف على أساس شروط الإجراء	١٠٠ - ١١١
اختبارات فردية واختبارات جمعية	١٠١
اختبارات القوة واختبارات السرعة	١٠٨
ثالثا : التصنيف تبعا للمحتوى	١١١ - ١١٣
رابعا : التصنيف تبعا لطريقة التصحيح	١١٣ - ١١٤
خامسا : التصنيف على أساس تطبيق النتائج	١١٤
سادسا : أسس أخرى للتصنيف	١١٥ - ١١٨
اختبارات أقصى أداء ممكن واختبارات الأداء المعتاد	١١٥
اختبارات ذات بناء محدد واختبارات بناؤها غير محدد	١١٥
اختبارات الورقة والقلم	١١٦
الاختبارات الأدائية	١١٦
اختبارات لغوية واختبارات غير لغوية	١١٧
قائمة مراجع القسم الأول	١١٩ - ١٢١

القسم الثاني

بناء المقاييس النفسية والتربوية ومميزاتها (١٢٣-٣٣٥)

الفصل الرابع : بناء المقاييس (١٢٥-١٤٥)

أهمية الإلمام بخصائص المقاييس	١٢٧
تخطيط عام لمشروع بناء المقياس	(١٢٣ - ١٤٣)
أولا : المشكلة	١٣٣
ثانيا : تحديد الأهداف	١٣٦
ثالثا : الإعداد للمشروع	١٣٧

الصفحة

رابعا : تحليل المضمون	١٣٨-١٤٠
(أ) تحليل المنهج الدرامي	١٣٩
(ب) تحليل العمل	١٤٠
خامسا : مواصفات الاختبار	١٤١
خطوات إنشاء المقياس	١٤٣

الفصل الخامس : إعداد فقرات الاختيار (١٤٧-١٧٥)

مقدمة	١٤٩
أنواع فقرات الاختبار	١٥٢
مقارنة أنواع الفقرات	١٥٤
قواعد عامة في إعداد الفقرات	١٦٠
التحليل الإحصائي لفقرات	(١٦٢-١٧٣)
(أ) الكشف عن الفقرات غير الصالحة	(١٦٢-١٦٥)
(ب) النجاح بالصدقة	(١٦٥-١٧٣)
١ - فروض نظرية	١٦٥
٢ - نتائج تصحيح أثر الصدقة	١٦٨
٣ - نتائج طلب عدم التغمين من المفحوصين	١٦٩
٤ - تحاشي نتائج الصدقة	١٧١
أفضل طرق التصحيح	(١٧٣-١٧٥)
(أ) المفتاح الملقب	١٧٤
(ب) التصحيح الأتوماتيكي	١٧٤
(ج) التصحيح الآلي	١٧٥

الفصل السادس : الصدق (١٧٧-٢١٥)

مبادئ الصدق	١٧٩
معامل الصدق	١٨٤
أنواع الصدق	(١٨٧-١٩٣)
١ - الصدق الظاهري	١٨٨
٢ - صدق المضمون	١٨٩
٣ - الصدق التنبؤي	١٩٠
٤ - الصدق التلازمي	١٩١
٥ - الصدق التجريبي	١٩١
٦ - صدق المفهوم	١٩٢

(ك)

صفحة

١٩٣	٧ - الصدق التطابق
١٩٣	٨ - الصدق العامل
(١٩٨-١٩٤)	مشكلة المحكات
١٩٤	تعريف المحك
١٩٤	الحاجة إلى المحكات
١٩٥	خواص المحك الجيد
(١٩٨-١٩٦)	أنواع المحكات
١٩٦	أولاً في الصناعة
١٩٦	ثانياً في المدرسة
١٩٧	ثالثاً في المجالات الأخرى
(٢٠٧-١٩٨)	طرق حساب معامل الصدق
١٩٨	أولاً الصدق الظاهري
٢٠٠	ثانياً صدق المضمون
(٢٠٠-١٩٨)	ثالثاً الصدق التنبؤي
١٩٨	(أ) طريقة النسبة المئوية
٢٠٠	(ب) طريقة المتوسطات
٢٠٠	(ج) طريقة معامل الارتباط
٢٠٠	رابعاً الصدق التلازمي
٢٠١	خامساً الصدق التجريبي
٢٠٢	سادساً صدق المفهوم
٢٠٣	سابعاً الصدق التطابق
٢٠٥	ثامناً الصدق العامل
٢٠٧	معامل الصدق معامل ارتباط
(٢١٢-٢٠٨)	تصحيح معامل الصدق
(٢١٥-٢١٢)	استخدام الصدق في حالات أخرى
٢١٢	(أ) التنبؤ بالحالات الفردية
٢١٤	(ب) في التصفية
٢١٤	(ج) في تكوين البطاريات

الفصل السابع : الثبات (٢١٧-٢٥٢)

٢١٩	مقدمة
(٢٢٩-٢٢١)	اعتبارات منطقية في تقدير الثبات
٢٢١	الثبات وتحليل التباين

صفحة	
٢٢٣	أسباب تباين درجات الأفراد في الاختبار
٢٢٤	موضوعية القياس
(٢٤٢-٢٢٩)	طرق حساب معامل الثبات
٢٣٠	١ - معامل الثبات بطريقة العصور المتكافئة
٢٣٢	٢ - الثبات بطريقة إعادة الاختبار
(٢٣٨-٢٣٤)	٣ - الثبات بطريقة نصفي الاختبار
٢٣٦	(أ) معادلة سبيرمان - براون
٢٣٦	(ب) معادلة جتان
٢٣٧	(ج) معادلة رولون
٢٣٨	(د) معادلة هورست
(٢٤٢-٢٣٩)	٤ - معامل الثبات بطريقة تحليل التباين
٢٣٩	(أ) معادلة كودر - ريتشاردسون
٢٤٠	(ب) معادلة هويت
٢٤٠	(ج) معادلة چلكسون
٢٤١	٥ - معامل الثبات الحقيقي
٢٤٢	معادلة كرونباخ
(٢٤٦-٢٤٢)	تفسير معامل الثبات
(٢٥١-٢٤٦)	العوامل التي تؤثر على معامل الثبات
٢٤٦	(أ) طول الاختبار
٢٤٨	(ب) زمن الاختبار
٢٤٨	(ج) تجانس المجموعة
٢٤٩	(د) الصدق
٢٥١	(هـ) عوامل أخرى
٢٥١	بيانات الثبات

الفصل الثامن : تحليل الفقرات (٢٩٥-٢٥٣)

٢٥٥	مقدمة
(٢٦٢-٢٥٦)	اعتباران في تحليل الفقرات
٢٥٦	١ - مستوى الصعوبة
٢٥٨	٢ - القدرة على التمييز
٢٦٢	القياس الموضوعي للصعوبة
(٢٧٦-٢٦٤)	طرق تحليل الفقرات
٢٦٤	(أ) دلائل صعوبة الفقرة
٢٧٠	(ب) دلائل قدرة الفقرة على التمييز

صفحة

٢٧١	١ - تحليل الفقرات في حالة الدرجات المستمرة
٢٧٣	٢ - تحليل الفقرات في حالة الجماعات المتقطعة
٢٧٤	(ج) تحليل الفقرات في حالة الجماعات المتطرفة
٢٧٥	(د) تحليل الفقرات في حالة محك متقطع
٢٧٦	الصدق المتبادل
٢٨٠	ثبات الفقرات
٢٨١	وزن الإجابات عن الفقرات
٢٨٦	أهمية وزن الإجابات
(٢٩٣-٢٨٦)	استخدام صدق الفقرات
٢٨٧	الطريقة الأولى
٢٩٠	الطريقة الثانية
٢٩٢	الطريقة الثالثة
٢٩٣	استخدام الإتساق الداخلي

الفصل التاسع : المعايير (٢٩٧-٢٣٥)

٢٩٩	مقدمة
٢٩٩	الحاجة إلى المعايير
٣٠١	الصفر المطلق
٣٠١	المعايير والتقنين
٣٠٤	الدرجات المحولة
(٣١٩-٣٠٥)	أنواع المعايير
٣٠٦	معايير العمر
٣٠٩	العمر العقلي
٣١١	عدم ثبات وحدات العمر العقلي
٣١١	عيوب العمر العقلي
٣١٢	معايير الفقرة الدراسية
٣١٣	المعايير المثبتة
٣١٥	ميزة المثبتات
٣١٥	عيوب المثبتات
٣١٦	الدرجات المعيارية
٣١٨	عيوب الدرجات المعيارية
٣١٩	نسبة الانحراف

الصفحة

أنواع أخرى من الدرجات المعيارية المعدلة	...	...	...	...	(٣١٩ - ٣٢٢)
١ - الدرجة ت	...	...	...	...	٣١٩
٢ - معيار اختيار التصنيف في الجيش الأمريكي	...	...	...	...	٣٢١
٣ - معيار الجامعات الأمريكية	...	...	...	...	٣٢١
٤ - الدرجة ج	...	...	...	...	٣٢١
٥ - معيار التسع	...	...	...	...	٣٢٢
تحويل المعايير	...	...	...	...	٣٢٢
النسب	...	...	...	...	(٣٢٤ - ٣٢٦)
١ - نسبة الذكاء	...	...	...	...	٣٢٤
٢ - نسبة التعليم	...	...	...	...	٣٢٤
٣ - النسبة التحصيلية	...	...	...	...	٣٢٦
الصفحات النفسية	...	...	...	...	٣٢٦
اختلاف نتائج الاختبارات	...	...	...	...	٣٤١
تعليق	...	...	...	...	٣٣٣
كيف نستخدم المعايير	...	...	...	...	٣٣٤
(أ) مبادئ تراعى عند تفسير درجات مجموعة من الأفراد	...	...	...	...	٣٣٤
(ب) مبادئ تراعى عند تفسير أداء الفرد الواحد	...	...	...	...	٣٣٥

قائمة مراجع القسم الثاني	...	...	...	...	(٣٣٦ - ٣٤٥)

القسم الثالث

القياس التربوي	(٣٤٧ - ٧١٠)

مستهل العاشر : للتربية والقياس والتقييم	...	...	(٣٤٩ - ٣٧٥)

القياس النفسي والقياس التربوي	...	...	...	...	٣٥١
لماذا تتغير التربية - نظرة تاريخية	...	...	...	...	٣٥٣
القياس في التربية	...	...	...	...	٣٥٦
تعريف بالتربية	...	...	...	...	٣٥٨
رابط أهداف القياس بأهداف التعليم	...	...	...	...	٣٥٩
التقييم الرسمي والتقييم غير الرسمي	...	...	...	...	٣٥٩
مشكلة المحكات في القياس التربوي	...	...	...	...	٣٦٠

(مس)

صفحة

٣٦٢		معنى التحصيل
٣٦٣		أهداف التربية
٣٦٤		الأهداف الصريحة والأهداف الضمنية
٣٦٥		القيم المؤثرة في اختيار الأهداف

الفصل الحادى عشر : تقييم المدرس (٣٦٧-٤٠٦)

مميزات المدرس الناجح ٣٦٩

مقتضيات التدريس (٣٧١ - ٣٨١)

١ - دراسة عمل المدرسين ٣٧١

٢ - سؤال التلاميذ ٣٧٣

٣ - تحليل تواريخ حياة المدرسين المشهورين ٣٨٠

٤ - دراسة وظائف التربية ٣٨١

٥ - قوائم المقتضيات لرجال التربية ٣٨١

قائمة كاليفورنيا للقدرة على التدريس (٣٨٨ - ٣٨١)

١ - قادر على تعليم التلاميذ ٣٨١

٢ - يرشد ويوجه الطلبة بحكمة ٣٨٤

٣ - يساعد التلاميذ على فهم وتقدير الميراث الثقافي ٣٨٥

٤ - يساهم بفعالية في أوجه نشاط المدرسة ٣٨٦

٥ - يساعد على تكوين علاقات طيبة بين المدرسة وبقية المجتمع المحلي ٣٨٧

٦ - يعمل في مستوى مهني ٣٨٧

أهم ما يجب أن يتضمنه تخطيطك ٣٨٨

بطاقات يقوم بها المدرس قدرته ٣٩٠

مقياس التقدير الذاتي لقدرة التدريس ٣٩١

قائمة مراجعة العوامل الهامة في شخصية المدرس ٣٩٦

بطاقات محلية لتقييم رجال التربية (٣٩٧ - ٤٠٦)

بطاقة تقوم المدرس ٣٩٧

جهود وزارة التربية والتعليم ٣٩٩

اخبار الاستعداد للوظائف التعليمية ٤٠٤

الفصل الثاني عشر : تقييم التلميذ (٢٤٨-٤٠٩)

ما هي الدرجة { ١٩ }
وظائف الدرجات والتقدير (١١ - ١٧)

صفحة

أولاً : المعلومات التي يحتاجها التلميذ الذي ننتدبه	٤١١ - ٤١٣
١ - الدفع	٤١١
٢ - توجيه التعليم	٤١٢
٣ - توجيه الخطط التعليمية والمهنية التالية	٤١٢
٤ - توجيه ' والشخصي	٤١٤
٥ - العلاقة ' تقدير الذات وتقدير المدرسة	٤١٤
ثانياً : إبلاغ الآباء بقييم الأبناء	٤١٥
ثالثاً : تقديم التقديرات للمؤسسات التعليمية الأعلى	٤١٥
رابعاً : تقديم التقديرات لأصحاب الأعمال	٤١٦
خامساً : استفادة المدرسة	٤١٧
الأسس الفنية في إعطاء الدرجات والتقديرات	(٤٢٣ - ٤١٧)
أولاً : العوامل المؤثرة	٤١٧
ثانياً : أنواع الأدلة ووزن كل منها	(٤٢٣ - ٤١٨)
(أ) امتحان نهاية الفصل الدراسي	٤١٩
(ب) الامتحانات العامة	٤١٩
(ج) الامتحانات الدورية خلال الدراسة	٤٢١
(د) المقالات والأبحاث والتقارير	٤٢٢
(هـ) المشاركة في فواحي النشاط الجماعية	٤٢٣
(و) النشاط في العمل	٤٢٣
وسائل الاتصال بين المدرسة والبيت	(٤٢٤ - ٤٢٨)
١ - بطاقة التقرير التقليدية	٤٢٤
٢ - تعديل عدد ومعنى الدرجات	٤٢٥
٣ - إرفاق الدرجات برتب على عوامل أخرى	٤٢٥
٤ - القوائم التفصيلية	٤٢٥
٥ - المقابلة بين الأب والمدرس ومميزاتها	(٤٢٨ - ٤٢٦)
(أ) مرونة التواصل	٤٢٦
(ب) فرص التفسير	٤٢٦
(ج) التواصل متبادل وصعوباته هي :	٤٢٧
(أ) الوقت	٤٢٧
(ب) المهارة	٤٢٧
(ج) تعاون الآباء	٤٢٨
٦ - انعطافات غير الرسمية من المدرس إلى الأب	٤٢٨

للمعلم الثالث عشر : اختبارات المقال والاختبارات

الموضوعية (٤٢٩-٤٦٤)

أهمية الامتحانات المدرسية (٤٣٢-٤٣١)

١ - الدافع ٤٣١

٢ - التشخيص والتعليم ٤٣١

٣ - تحديد أهداف التربية ٤٣١

٤ - المقارنة بين التلاميذ والنقل ومنح الشهادات ٤٣٢

تخطيط الامتحان ٤٣٢

أنواع امتحانات المدرس (٤٣٤-٤٣٣)

اختبار المقال ٤٣٣

الاختبار الموضوعي ٤٣٤

شروط الامتحانات ٤٣٤

مزايا الامتحانات الموضوعية ٣٥

كيف نضع أسئلة الصواب والخطأ ٣٥

طريقة الاختيار من عدة إجابات ٣٦

قواعد عامة تراعى قبل إعداد الاختبار في صورته النهائية ٣٧

إعداد الأسئلة الموضوعية ٣٩

إعداد أسئلة الاختبار ٤١

أسئلة الاختيار من عدة إجابات ٤١

قواعد تراعى عند إعدادها ٤٢

أسئلة التكميل والأسئلة ذات الجواب القصير ٤٤

قواعد تراعى عند إعدادها ٤٥

أسئلة الصواب والخطأ ٤٦

قواعد تراعى عند إعدادها ٤٧

أسئلة التزاوج ٤٨

قواعد تراعى عند إعدادها ٤٩

اختبارات التصنيف ٥٠

قياس الفهم ٥٠

إعداد الاختبار الموضوعي ٥١

الاستخدام الفعال لامتحان المقال ٥٥

مَنى يستخدم اختبار المقال ٥٥

تصحيح اختبار المقال ٥٧

تحديد أهداف تدريس المادة ٥٨

صفحة:

الامتحانات شر لابد منه	٤٥٩
انتقاد نظام الامتحانات	٤٦٠

الفصل الرابع عشر : برنامج الاختبارات في المدرسة (٤٦٥-٥٠٠)

مقدمة	٤٦٧
ملخص وظائف البرنامج	٤٦٧
أولاً : وظائف في المدرسة	٤٦٩
ثانياً : الوظائف التوجيهية لبرنامج الاختبارات	٤٧١
ثالثاً : الوظائف التطبيقية لبرنامج الاختبارات	٤٧٢
خصائص برنامج الاختبارات الجيد	٤٧٥
برامج مقترحة للاختبارات في المدرسة	(٤٧٧-٤٨١)
١ - برنامج للمدرسة الابتدائية والإعدادية	٤٧٧
٢ - برنامج الاختبار في المدرسة الثانوية	٤٧٩
٣ - برنامج الاختبار في المعاهد	٤٨٠
السجل المجمع	٤٨١
اعتبارات في تخطيط البطاقة المجمة	٤٨٢
جهود وزارة التربية والتعليم	٤٨٤
صعوبات برنامج الاختبارات في المدرسة	(٤٨٨-٤٩٤)
مشاكل الصلح	٤٨٨
مشاكل المعايير	٤٩٣
خطورة برنامج الاختبارات على البرامج التعليمية	٤٩٥
الاختبارات التحصيلية المستخدمة في برامج الاختبارات	٤٩٧
الاختبارات النفسية في البرامج	٤٩٨

الفصل الخامس عشر : خصائص ومشكلات

الاختبارات التحصيلية (٥٠١-٥١٨)

خصائصها	٥٠٣
استخدامها	٥٠٤
مزايا الاختبارات التحصيلية	٥٠٨
عيوب الاختبارات التحصيلية	٥١٠
وظائف الاختبارات التحصيلية	٥١٥
أنواع الاختبارات التحصيلية	٥١٧

(ق)

صفحة

الفصل السادس عشر: اختبارات تحصيلية مقننة عامة (٥١٨-٥٤٧)

قياس التفكير	٥٢١ : ٥٢٨
اختبارات النمو التعليمي	٥٢٩ : ٥٣٥
قياس مهارات المذاكرة	٥٣٥ : ٥٣٩
بطاريات التحصيل العام	٥٣٩ : ٥٤٧
في المرحلتين الابتدائية والإعدادية	٥٣٩
في المرحلة الثانوية	٥٤٥

الفصل السابع عشر : اختبارات تحصيلية مقننة خاصة (٥٤٩-٥٩٦)

القراءة	٥٥١
اختبارات التهيؤ في القراءة	٥٥٣
اختبارات مسح القراءة	٥٦١
تشخيص القراءة	٥٦٤
الرياضيات	٥٧٠
التهيؤ والتنبؤ	٥٧١
المسح	٥٧١
التشخيص	٥٧٢
العلوم	٥٧٩
الجغرافية	٥٨٤
مجالات أخرى	٥٨٥
مقاييس الإنتاج	٥٨٨
أغراض تعليمية أوسع	٥٨٩
التفكير الناقد في الدراسات الاجتماعية	٥٩١
-الاختبارات التحصيلية المهنية	٥٩٣
اختبارات الحرف	٥٩٥

الفصل الثامن عشر : بحوث واختبارات تربوية محلية (٥٩٧-٦٧٩)

مقدمة	٥٩٩
بحث الأستاذ إسماعيل القباني	٦٠٠
بحوث وزارة التربية والتعليم	٦١٠
بحوث المؤلف	٦٢٣
القراءة : الدكتور رشدي خاطر	٦٥٦
الحساب ، المؤلف	٦٦٤
نحو المجلد الثاني	٦٧٩

المراجع (٦٨٠-٧٠٠)

القسم الأول

التعريف بالقياس ومفاهيمه وأدواته

الفصل الأول

التعريف بالقياس

- الحاجة إلى القياس النفسى والتربوى .
- نشأة القياس فى التربية وعلم النفس .
- ماذا نقيس .
- ماهو المقياس النفسى .
- من يقيس ومتى يقيس .
- تطبيقات القياس النفسى والتربوى .
- أغراض القياس .
- تحديدات القياس .

مقدمة :

الإنسان إنسان بما يشترك فيه مع غيره من صفات جسمية ونفسية ومن العيش في حضارة واحدة والتعرض لأنواع متقاربة من الخبرات العامة ويشترك مثلهم في المجتمع الإنساني الأكبر . فهو بذلك يشبه غيره من بني الإنسان والإنسان فرد بما ينفرد فيه عن غيره من تكوين نفسى معين وهو بذلك يختلف عن غيره من بني الإنسان . فمن الأمور الملحوظة ، في الإنسان وفي سائر الكائنات الحية ، أنه رغم ما بين أفراد الجنس الواحد من تشابه وتمائل إلا أن بين أفرادها فروقا كثيرة . وبعض هذه الفروق قد يظهر من مجرد الملاحظة العرضية العابرة ولكن كثيراً منها لا يتضح بسهولة بل يحتاج إلى التدقيق وإلى الاستعانة بوسائل مساعدة لإدراكها أو تمييزها .

على أن إدراك الفرق في أى صفة يتطلب ، أولاً ، التعرف على هذه الصفة وفهمها ثم اكتشاف واستنباط وسائل لقياس ما دق من هذه الفروق : فالشخص العادى يرى آلاف النجوم في السماء وتبدو كلها في نظره سواء ، لكنه كلما ازداد علماً وخبرة بها كلما زاد إدراكاً للفروق والمميزات بينها . والعالم الفلكى يحتاج إلى وسائل مساعدة لدراسة هذه الفروق كالتلسكوب ومقاييس الزمن و . . الخ . وبالمثل فمن لا يعرف صفات القطن وخصائصه لا يستطيع أن يميز بين أنواعه ورتبه المختلفة .

وقد دلت جميع الدراسات التجريبية على جماعات الحيوان أن بكل جماعة من الحيوان ، مهما بدت متساوية أو متقاربة ، يوجد بين أفرادها فروق كثيرة في النشاط التلقائى وفي درجات الحساسية وفي سرعة التعلم : فإنه ابتداء من أدنى الكائنات الحية ذات الخلية الواحدة وهي البروتوزوا وجد أن عدد المحاولات اللازمة لتكوين استجابة شرطية خاصة يرأوح في بعضها من

٧٩ إلى ٢٨٩ في البعض الآخر . هذا عن تجربة أجريت على ٨٢ من البروتوزوا : كما وجد في تجربة أخرى على عدد ١١ من الأغنام أن عدد المحاولات اللازمة يتراوح بين ٣ محاولات و ١٧ محاولة :

إذا كان هذا هو الحال في المملكة الحيوانية فحري أن تكون هذه الفروق أوسع مدى كفاً وكماً في الإنسان وهو ما أيدته بالإجماع جميع الدراسات والبحوث ونعرض لهذا فيما بعد عند الكلام عن الفروق النفسية وأنواعها :

نشأة القياس في علم النفس والتربية :

من الطريف أن نعرف أن أول من اكتشف الفروق النفسية هو عالم فلكي . إذ حدث في سنة ١٧٩٦ أن طرد أحد الفلكيين مساعده في المرصد لأنه أخطأ في رصد الزمن الذي يقضيه كوكب معين في مروره على لوحة التلسكوب أمام هذا المساعد . وكان مقدار هذا الخطأ هو ثانية واحدة . وربما لم تكن لهذا الحادث في ذاته أهمية لولا أن جاء العالم المشهور بسل Bessel سنة ١٨١٦ واهتم بجمع البيانات عن الأخطاء في تقديرات الفلكيين المختلفين . وخلص من دراسته لهذه الفروق إلى ما عرفه بالمعادلة الشخصية *personal Equation* . ومضمونها أن الأفراد يختلفون من حيث سرعة « زمن الرجوع » أي أن الزمن الذي يمر بين حدوث المثير وحدث الاستجابة يختلف طوله من فرد إلى آخر . وقد أدى هذا الحادث التاريخي إلى اهتمام الباحثين في النصف الأول من القرن التاسع عشر بقياس الفروق الفردية ولم يكن هدف علماء النفس التبريريون الأول قياس الفروق الفردية إذ كان الظن السائد أنها أخطاء ولذلك اهتموا بدراساتها للتخلص منها والوصول إلى صياغة أوصاف معممة للسلوك الإنساني أو قانون عام يصف السلوك الإنساني . كان هذا هو الاتجاه عند ثونت Wundt الذي أنشأ أول معمل لعلم النفس في لينزج سنة ١٨٧٩ :

كما انعكس تأثر من تعلموا على يديه بعلم وظائف الأعضاء والطبيعات خلال بحوثهم ، فاهتموا بالحساسية للمثيرات الحسية مثل البصرية ، والسمعية ، وبدراسة زمن الرجوع Reaction time . وكان لهذا الاتجاه صداه في الاختبارات النفسية الأولى التي ركزت على الظواهر الحسية ،

وكان لذلك الاتجاه تأثيراً آخرًا ؛ إذ نبه إلى أهمية الضبط الدقيق للظروف التي تسجل فيها الملاحظات . فمثلاً ، اتضح أن ألفاظ التعليمات المعطاة للمفحوص في تجربة زمن الرجوع قد تزيد أو تقلل من سرعته في الاستجابة وأن ، نصاعة أو لون المجال المحيط قد تغير مظهر المثير البصري . وهكذا لزم أن توحّد ظروف ملاحظة كل المفحوصين . ومن هنا أصبح تقنين الإجراء Standardization طابعاً يميز الاختبارات النفسية :

وجاء بعد فونت تلميذه سير فرانسيس جالتون Francis Galton البيولوجي الإنجليزي الشهير (١٨٢٢ - ١٩١١) فاهتم بدراسة الوراثة عند الإنسان وتحقق من أهمية قياس مميزات الأشخاص المرتبطين وغير المرتبطين بالمفحوص . وبهذا اكتشف بدقة درجة تشابه الذرية offspring ، أو الإخوة والأخوات ، وأولاد وبنات العم أو الخال ، أو التوائم twins . وفي سنة ١٨٨٢ أنشأ معملًا لعلم الإنسان القياسي anthropometry : وفيه أمكنه قياس حدة البصر والسمع ، والقوة العضلية ، وزمن الرجوع ، ووظائف حسية حركية sensorimotor بسيطة أخرى : ومن ثم جمع أول وأضحى ، مجموعة من البيانات عن الفروق الفردية في العمليات النفسية البسيطة ..

وصمم بنفسه في معمله اختبارات بسيطة طبقها ، ويطبقها البعض ، حتى الآن ، إما في صورتها الأصلية أو المعدلة : ومن مثلها نجد ، « قضيب جالتون » للتمييز البصري للأطوال ، « وصفارة جالتون » لتحديد أعلى مقام Pitch سمعى ، و . . الخ : وإن كان متأثراً بالفيلسوف لوك Locke ، بدليل اعتقاده بأن اختبارات التمييز الحسى تعد وسيلة لمعرفة المستوى العقلى

للشخص ، فضلاً عن كتاباته الفلسفية ، إلا أنه أسدى للحركة القياسية الكثير . فقد لاحظ ، وتأكد بعده ، أن المعتوهين idiots تنقسم في الأغلب القدرة على تمييز الحرارة ، والبرودة ، والألم :

كما كانت لجالتون الزيادة في تطبيق مناهج الاستبيان Questionnaire والمقياس المتدرج Rating Scale ، وكذا في استخدام منهج « التداعي الحر free association » . وله جهد مشكور في تطوير الطرق الإحصائية لتحليل البيانات عن الفروق الفردية بحيث يمكن أن يعالجها باحث غير مدرب رياضياً . وكذا سار من بعده تلميذه العملاق كارل بيرسون Karl Pearson .

كما أن جيمس ماكين كاتل James McKeen Cattell السيكولوجي الأمريكي له جهود كبيرة في تطوير حركة القياس النفسي وعلم النفس التجريبي . فقد أكل دراسة الفروق الفردية في زمن الرجوع وتعاون مع جالتون وإن لم يؤيدها قونت . وأنشأ وأدار معمل علم النفس التجريبي ، ونشر حركة القياس .

واستخدم اصطلاح « الاختبار العقلي mental test » سنة ١٨٩٠ لأول مرة إذ تعرض لبطارية اختبارات لقياس الكثير من السمات الحسية الحركية والنفسية . وأشار إلى أن الوظائف العقلية يمكن أن تقاس عن طريق اختبارات التمييز الحسي وزمن الرجوع : ويرجع تفضيله لمثل هذه الاختبارات إلى إيمانه بأن الوظائف البسيطة يمكن أن تقاس بتحديد ودقة ، بينما بدت المقاييس الموضوعية للوظائف الأكثر تعقيداً عملاً لا يبشر بالخير آنئذ .

وفي نهاية القرن التاسع عشر بذلت جهود لقياس وظائف أكثر تعقيداً في اختبارات القراءة ، والتداعي اللفظي ، والذاكرة ، والحساب البسيط . وفي سنة ١٨٩٣ دعى جاسترو Jastro إلى القيام باختبارات للعمليات الحركية ، والحسية ، والإدراكية البسيطة . وإلى الاهتمام بالمعايير norms :

وكان الارتباط ضعيفاً بين التقديرات على الاختبارات المختلفة وبين المستوى العقلي الذي كان يحدده المدرسون في تقديراتهم وفي الدرجات الأكاديمية .

المشكلات التربوية :

وفي القرن التاسع عشر أيضاً بدأ الاهتمام بتصنيف وتدريب ضعاف العقول بعد أن كانوا يعاملون بقسوة . وقد ظهرت الحاجة ماسة إلى اختبار لتشخيص وتصنيف حالات الضعف العقلي بعد أن انتشرت مؤسسات خاصة بضعاف العقول في أوروبا وأمريكا ، خاصة وأن الأمر تطلب الوصول إلى تحديد معايير القبول والوصول إلى نظام موضوعي للتصنيف . وكان أن فرق بين المجنون insane وضعيف العقل ، فالأول يبدى اضطرابات انفعالية قد ترتبط بالتدهور العقلي وقد لا ترتبط ؛ بينما يتميز الثاني أساساً بنقص العقل الذي وضح منذ الولادة أو الطفولة المبكرة .

بدأ هذا الجهد اسكيرول Esquirol ، طبيب فرنسي ، سنة ١٨٣٨ ، بتقسيم الضعف العقلي إلى مستويات ثلاثة تبدأ بالعتة idiocy . وقد أشار إلى أن المقاييس الجسمية ليست كافية وأن الاستخدام الفردي للغة هو محك موثوق به للمستوى العقلي . وعلى هذا الأساس ، ميز درجتين في البله imbecility وثلاث درجات في العتة . ففي أعلى درجات البله ، كما قال ، يكون استخدام الكلام سهلاً ميسوراً ؛ وفي الدرجة الأقل ، يكون الحديث أصعب والمفردات أقل .

ومن الجدير بالذكر أن نعلم أن المحكات الراهنة في الضعف العقلي لغوية في الغالب ، وأن اختبارات الذكاء اليوم مشبعة بالمحتوى اللفظي :

أما عن سيجون Seguin فهو طبيب فرنسي كان رائد تدريب ضعاف العقول ، إذ نادى ببقاء الفكرة السائدة وهي استحالة الشفاء من نقص العقل . وفي سنة ١٨٣٧ أنشأ أول مدرسة لضعاف العقول من الأطفال :

وله وسائل فنية لا زالت مستخدمة في « تدريب الحس » « وتدريب العضل » كما ساهم في اختبارات ذكاء الأداء Performance وصمم لوحة باسمه .

وقد أعد كراپلن Kraepelin ، الذى اهتم بالفحص الإكلينيكي لمرضى العقول ، سلسلة من الاختبارات لقياس ما كان يظنه عوامل أساسية في تشخيص الفرد . استخدمت هذه الاختبارات والعمليات الحسابية لقياس أثر التدريب ، والتذكر ، والقابلية للتعب والتشتيت . كما استخدم تلميذه أوهرن Oehrman اختبارات للإدراك ، والذاكرة ، والتداعي ، والوظائف الحركية في بحث عن العلاقات المتبادلة بين الوظائف النفسية . وتبعهم الألمانى إيبينجهوس والإيطالى فيرارى Ferrari في الاهتمام بالاختبارات في الحالات المرضية الأمر الذى انتقده ، في فرنسا ، بينيه Binet وهنرى هنرى Henri قائلين إن هذه الاختبارات غالباً حركية ، ومركزة دون حق على قدرات متخصصة . وقالوا إن قياس الوظائف المركبة لا يلزم أن يكون محدداً جداً ، ما دامت الفروق الفردية في هذه الوظائف كبيرة . واقترحا قائمة من اختبارات الذاكرة ، والتخيل ، والانتباه ، والفهم ، والقابلية للإيجاء ، والفهم الحسابى ، . . . الخ . تلك هى التى أدت إلى ظهور مقياس بينيه المعروفة .

لم يأل بينيه ومساعدوه جهداً في استخدام كل المناهج حتى قياس السمات الجسمية ، وتحليل الخطوط ، ودراسة الكف ، ونتيجة لذلك اقتنع بأنه لا مناص من قياس الوظائف العقلية المركبة . وفى سنة ١٩٠٤ كلف بينيه وسيمون Simon من قبل وزارة المعارف الفرنسية بدراسة خطوات تعليم الأطفال دون الأسوياء . وكانا ضمن لجنة فشرعا في مقياسهما الأول

Binet-Simon Scale

. وفى سنة ١٩٠٥ خرج المقياس ، مكوناً من ثلاثين بنداً item متبة تصاعدياً بحسب الصعوبة . وكان مستوى الصعوبة يحدد موضوعياً بتطبيق البنود على ٥٠ طفلاً سويّاً تتراوح أعمارهم من ٣ - ١١ سنة ، وعلى بعض

الأطفال ضعاف العقول . وكانت البنود مصممة بحيث تغطي وظائف متعددة مع التأكيد على الحكم ، والفهم ، والاستدلال ، إذ اعتبرها بينيه مكونات أساسية للذكاء . وكان الاختبار في الأغلب لفظياً .

وفي سنة ١٩٠٨ زاد عدد البنود ، وحذف ما لم تثبت صلاحيته ، وجمعت في مستويات السن . وعليه وضعت في مستوى ٣ سنوات كل البنود التي يستطيع الإجابة عنها الطفل العادي الذي سنه ٣ سنوات ، وهكذا في كل سن حتى ١٣ سنة . وهكذا أمكن حساب درجة الطفل في الاختبار خلال « العمر العقلي mental Age » ، أي خلال عمر الطفل العادي الذي يعادله في الأداء .

وفي سنة ١٩١١ عدل الاختبار وأصبح يقيس حتى سن الرشد . وترجم الاختبار إلى لغات كثيرة واهتم به العالم . وفي أمريكا كرس له الجهود I. M. Terman وراجعته في جامعة ستانفورد وعرف المقياس باسم Stanford-Binet وحسبت فيه نسبة الذكاء (I. Q) intelligence Quotient . وهي النسبة بين العمر العقلي والعمر الزمني .

نشأة القياس في التربية :

القياس في ميدان التربية أقدم كثيراً من القياس في علم النفس . ولقد شعر المربون قديماً بالحاجة إلى قياس التقدم والتأخر في تلاميذهم ، والحاجة إلى التعرف على نواحي الضعف فيهم . كما شعروا أيضاً بالحاجة إلى قياس مدى نجاح جهودهم وطرقهم في التدريس بقياس ما يظهره تلاميذهم من تقدم فيما يدرسونه . على أن هذا القياس كان يعتمد في أول الأمر على الملاحظة الذاتية والآراء الشخصية ثم مر بعد ذلك بثلاث مراحل مميزة هي :

١ - مرحلة الامتحانات الشفوية Oral Tests .

٢ - مرحلة الامتحانات التحريرية من نوع المقال essay type tests .

٣ - والمرحلة الثالثة وقد بدأت حديثاً ، نسياً ، عن طريق الامتحانات الموضوعية Objective tests أو ما يعرف «بالاختبارات الحديثة» . وسنعالج هذه الطرق في الفصول الخاصة بالامتحانات المدرسية والاختبارات التحصيلية :

وقد ذكرنا فيما سبق أن المشكلات التربوية المتعلقة بتعليم ضعاف العقل أدت إلى دفع حركة القياس النفسى قدماً إلى الأمام . ولقد كانت إجابة بعض القادة القدماء في التربية عن السؤال « لما ذا نقيس أو نقيم ؟ » ، أننا نقيس لتمكين الطلبة اللاتقيين من التعليم الصالح على أيدي مدرسين أكفاء : وهذا هو الهدف الاساسى لنظم التعليم الناجحة .

ما هو المقياس ؟

يقول K. Bean « إن المقياس ، في نظر التربية وعلم النفس ، هو مجموعة مرتبة من المثيرات أعدت لتقيس ، بطريقة كمية أو بطريقة كيفية ، بعض العمليات العقلية أو السمات أو الخصائص النفسية » . والمثيرات هنا قد تكون أسئلة شفوية أو أسئلة تحريرية مكتوبة وقد تكون سلسلة من الأعداد أو بعض الأشكال الهندسية أو النغمات الموسيقية أو صوراً أو رسوماً . وهذه كلها مثيرات تؤثر على الفرد وتستثير استجاباته .

والاختبار يجب أن يقيس شيئاً مقصوداً كأن يعطى درجة أو قيمة أو رتبة : الخ . وبعبارة أخرى يكون هدف الاختبار دائماً أن يقيس أو أن يقيم شيئاً مقصوداً . والمقياس يجب أن يعطى نوعاً من الدرجات أو أن يقدم تصنيفاً وصفيًا أو كليهما معاً ويعرف Cronbach كرونباخ الاختبار « بأنه طريقة منظمة لمقارنة سلوك شخصين أو أكثر » وهو يرى ضرورة هذا الشمول في التعريف حتى لا يُظن أن الاختبار قاصر على مجموعة من الأسئلة يجيب عليها المفحوص تحريراً أو شفويًا . فالاختبارات

النفسية واسعة التنوع وأصبحت أنواع المثيرات التي تستخدم لغرض الاختبار كثيرة وأخذت في النمو بسرعة كبيرة .

وعلى ضوء هذا التعريف للاختبار لا يعتبر المقابلة الشخصية interviewing ، وهي أداة هامة في دراسة الحالة ، اختباراً ، لأنها لا تقيس نوعاً واحداً من السلوك في جميع الأشخاص . فمن النادر أن يتبع نفس الأسلوب ونفس الأسئلة في مقابلة الأشخاص المختلفين . وهذا الاختلاف أمر يستلزمه الغرض من المقابلة وهو دراسة حالة الفرد لا مقارنته بالآخرين : على أن هناك نوعاً من المقابلة الشخصية يمكن أن تعتبر اختباراً حسب التعريف السابق وهي النوع المسمى « المقابلة الشخصية المقننة » Standardized interviews ذلك لأن هذا النوع يعتمد على أسئلة مقننة تستعمل مع جميع المفحوصين ولذلك يمكن استخدام نتائجها في المقارنة بين الأفراد .

وعلى أساس تعريف كرونباخ للاختبار بأنه طريقة منظمة لمقارنة السلوك ، يمكن اعتبار أى جهاز أو أى وسيلة للدراسة الإحباط Frustration مثلاً أو أى مجموعة من الأسئلة للدراسة الشخصية أو أى تقدير لأداء الفرد في موقف معين من مواقف الحياة يمكن اعتبار هذه جميعاً اختبارات فهي وإن كانت في كثير من الأحيان لا تعطينا قياساً أو تقديراً كمياً لما نختبره إلا أنها تعطى وصفاً لنواحي الشخصية أو السلوك يمكن مقارنتها في الأفراد . والمعروف أن جميع اختبارات الشخصية تحاول وصف شخصية الفرد ككل لا أن تحلل سلوكه أو شخصيته إلى أرقام وعلاقات كمية .

والقياس في علم النفس يقوم على ما نادى به ثورنديك Thorndike في قوله « إذا وجد شيء ، فإنه يوجد بمقدار فإذا كان موجوداً بمقدار فإنه يمكن قياسه » .

وبالرغم من أن هذه القضية التي نادى بها ثورنديك كانت حافزاً قوياً للقياس النفسى ، إلا أنها كثيراً ما أثارت الجدل حول قيمة المقاييس التي

لا تعطى تقديراً كمياً مثل كثير من اختبارات الشخصية التي تعطى تقديرات وصفية أو كيفية .

ولقد استقر الأمر أخيراً وأصبح من الأمور المسلم بها أن القياس الوصفي أو الكيفي — كما في حالة الشخصية — ضرورى في علم النفس حينما كان القياس الكمي متعذراً :

أما آنستازى Anne Anastasi فتقدم تعريفاً أوفى وأدق للاختبار النفسى فتقول « الاختبار النفسى ، أساساً ، هو مقياس موضوعى ومقنن لعينة من السلوك » والاختبارات النفسية ليست إلا عينة صغيرة ممثلة للسلوك المراد قياسه أو فحصه : فاختبار المفردات مثلاً الذى يتكون من ٥٠ كلمة هو عينة من الكلمات التى يفترض أن يعرف الفرد منها قدرًا معيناً ، هذه الكلمات قد تقدر بالآلاف : وبالمثل في اختبار للعمليات الحسابية أو التفكير أو . . . الخ .

وهكذا فالإحصائى النفسى يقوم بما يقوم به الكيميائى في معمله فهو يختبر عينة صغيرة من المادة المراد اختبارها . وقد يأخذ عينة أخرى من نفس المادة ليتأكد من ثبات نتائجه أو قد يحلل عينة ثالثة من نفس المادة بطرق وأدوات أخرى كى يتأكد من صحة نتائجه وهو في كل مرة لا يأخذ إلا عينة صغيرة جداً من المادة المراد فحصها ، وكذلك فإراز القطن الذى يستخرج عينات من بالات القطن ليفحصها ليحكم على جودة ورتبة القطن كله . ولكن يجب أن يختار الفاحص هذه العينة بكل دقة بحيث تكون نموذجاً تمثل جميع المادة في حالة الكيميائى والفراز أو تمثل جميع أنواع السلوك المراد قياسه في حالة القياس النفسى . ويتوقف مدى تمثيل الاختبار للسلوك موضوع الفحص على عدد وطبيعة الفقرات في العينة (أى الاختبار) .

والقيمة التشخيصية أو التنبؤية للاختبار النفسى تعتمد على المدى الذى يستخدمه من ذلك الميدان الواسع من السلوك ودلالة هذا المدى . فلا يهملنا أن فرداً معيناً أجاب بصحة على عدد معين من المسائل الحسابية في الاختبار

بل المهم هو مدى ارتباط هذا العدد بكل العمليات الحسابية التي يلزم أن يعرفها من نطبق عليه الاختبار .

وعلى ذلك يجب أن تؤخذ فقرات الاختبار من السلوك الذي نختبره أو نريد أن نتنبأ به ويجب أن يكون الاختبار عينة لكل السلوك المطلوب قياسه . وتختلف درجة ارتباط وتشابه فقرات الاختبار مع محتوى الميدان الواسع من السلوك المقاس من اختبار إلى آخر . فاختبار الحساب يتكون من ٢٠ مسألة مأخوذة من واقع ١٠٠ مسألة حسابية تعلمها فعلا من نجري عليهم الاختبار . كما أن هناك اختبار للسائقين يعقد في نفس الظروف الطبيعية التي سيعملون وسطها إذ يجري عليهم داخل السيارة وفي طريق غام . هذان الاختباران يمثلان السلوك المراد قياسه تمثيلا يزيد في الثاني عنه في الأول . ومن ناحية أخرى نجد اختبارات أخرى تقل درجة تمثيلها للسلوك المراد قياسه ، على الأقل ظاهريا ، إلى حد كبير مثل اختبار الرورشاخ Rorschach (بقع الجبر) ، حيث نريد معرفة استجابات المفحوص تجاه غيره من الناس بتفسيرنا لاستجاباته عن بقع الجبر — مادة الاختبار . فهنا تمثيل الاختبار تمثيل رمزي ليس واقعا .

ولكن أيا كان الاختبار ، بالرغم من الفروق الكبيرة أحيانا بين أنواع الاختبارات ، فهو عينة ممثلة للسلوك المراد قياسه والتنبؤ به . تتوقف قيمة الاختبار على مدى الارتباط الحقيقي بين أداء المفحوص له وبين أدائه في المواقف الأخرى المماثلة من حياته الواقعية :

لماذا نقيس :

الغرض الرئيسي من القياس هو الكشف عن الفروق بأنواعها المختلفة إذ أنه لولا وجود هذه الفروق لما كانت هناك حاجة إلى القياس . وتتلخص أنواع الفروق الرئيسية في أربعة طوائف نجملها فيما يلي :

١ - الفروق بين الأفراد inter-individual : ويهدف قياس هذا النوع من الفروق إلى مقارنة الفرد بغيره من أفراد فرقة الدراسة أو عمره أو بيئته في ناحية من النواحي النفسية أو التربوية أو المهنية . لتحديد مركزه النسبي فيها حتى يمكن تصنيف الأفراد إلى مستويات أو إلى جماعات متجانسة .

٢ - الفروق في - ذات الفرد intra individual : وقياس هذا النوع من الفروق يهدف إلى مقارنة النواحي المختلفة في الفرد نفسه لمعرفة نواحي القوة والضعف في الفرد بالنسبة لنفسه . بمعنى مقارنة قدراته المختلفة معاً ، للتعرف على أقصى إمكانياته في كل منها بغرض الوصول إلى تخطيط أفضل لبرامج تعليمه أو لتدريبه . كما تفيد في توجيهه مهنيًا وتربويًا حتى يحقق أكبر نجاح في حدود إمكانياته هو :

٣ - الفروق بين المهن inter-Occupational : فن المعروف أن المهن المختلفة تتطلب مستويات مختلفة من القدرات والاستعدادات والسمات . وقياس هذه الفروق يفيدنا في الانتقاء المهني وفي التوجيه المهني وفي إعداد الفرد عمومًا للمهن .

٤ - الفروق بين الجماعات inter-group : تختلف الجماعات في خصائصها ومميزاتها المختلفة . ذلك أن الدراسات العلمية المتعددة أثبتت فروقاً بين جوانب الحياة النفسية في كل من الجنسين ، وبين الجنسيات المختلفة ، وبين الأعمار المختلفة ، . الخ وقياس هذه الفروق يفيدنا في دراسة سيكولوجية الجماعات وخصائص النمو ، ودراسة العوامل التي قد تكون مسئولة عن هذه الفروق لإنماء الصالح منها والتغلب على العاقل . ولا يخفى أن قياس كل نوع من هذه الفروق ، له فائدته الكبرى في تطوير البحوث العلمية مما يؤدي بدوره إلى تطوير المقاييس ذاتها .

. والوظيفة الأساسية للاختبارات النفسية هي قياس الفروق الفردية . وإن كان تشخيص الضعف العقلي في مبدأ الأمر هدفاً للاختبارات النفسية

وما زال كذلك فى أنواع معينة منها . وهناك استخدامات إكلينيكية للاختبارات من مثل تشخيص الاضطراب الانفعالى ، والجناس : كما أن مشاكل التربية دفعت الاختبارات النفسية فى ماضيها دفعة قوية . وما تزال المدارس تفيد الكثير من الاختبارات فى حالات التخلف المدرسى والتفوق والتوجيه التربوى والمهنى فى التعليم الفنى والتعليم العالى .

وميدان الاختيار والتصنيف المهنيين ميدان خصب جديد حققت فيه بعض الاختبارات نجاحاً كبيراً وبخاصة تلك التى تفيد فى اختيار من يشغلون الوظائف العليا إذ أن المقابلة ليست وسيلة كافية للحكم بصلاحية طالب الوظيفة . هذا فضلاً عن أن الاختبارات تكون جزءاً هاماً من إعداد الأخصائيين فى شؤون الأفراد . كذلك الحال فى اختيار وتصنيف الجنود وأما الدليل ابتداء من الحرب العالمية الأولى إذ بدأت الاختبارات تغزو ميادين الخدمات العسكرية كلها .

كما أن الاختبارات النفسية تساهم الآن فى حل الكثير من المشاكل العملية ، إذ يستخدمها علم النفس الفارقى أداة فى البحث ، فى دراسة الفروق الفردية ، والتوحد بالسمات النفسية ، وقياس الفروق الجماعية ، وبحث العوامل البيولوجية والثقافية وارتباطهما بالفروق السلوكية :

وهكذا يتطلب الفهم الواعى بأى ميدان من ميادين علم النفس المعاصر أن نعرف شيئاً عن تلك الاختبارات التى لم تترك ميداناً منها إلا وحققت فيه نجاحاً . الأمر الذى يوجب الدراية ببعض الاختبارات النفسية الهامة وتقييمها وتفسير نتائجها بناء على علم بتكوينها ، ووظائفها ، وحدودها .

تطبيقات القياس النفسى والتربوى

أولاً : فى التربية والتعليم :

١ - التخلف الدراسى :

تعددت الاختبارات النفسية التى تكشف عن التخلف الدراسى وأسبابه ، وهذه الاختبارات لا تعتمد على تقييم التلميذ فى المواد التى درسها فحسب بل على تقديم صورة كاملة عن جوانب الحياة النفسية للتلميذ للكشف عن العوامل التى غالباً ما تتدخل أو تؤدى إلى فشله .

٢ - فى الاختيار التربوى :

تعدد فرص اختيار معاهد التعليم للطلبة الجدد كما تعدد الفرص أمام التلميذ فى مراحل التعليم المختلفة لاختيار مواد دراسية معينة وترك أخرى . وهنا تزود الاختبارات المعاهد والتلميذ بصورة دقيقة عما تتطلبه دراسة معينة من تفوق فى مواد أخرى . كما تعين الاختبارات النفسية التلميذ على فهم مستواه فى مواد وقدرات معينة تؤهله لدراسات تالية أو تجعله يحجم عن متابعة دراسة معينة لأنها تتطلب ما لا يتوفر لديه .

٣ - فى التوجيه التربوى :

على نفس الأسس السابقة تفيدنا الاختبارات النفسية والتربوية فى مساعدة التلميذ على اختيار دراسة تناسبه بعد أن يكون عن نفسه صورة موضوعية متكاملة تعتمد على معرفته بقدراته واستعداداته وميوله ، وسماته .

٤ - الإرشاد التربوى :

أما فى حالة اختيار التلميذ دراسة لا تناسبه أو توجيهه إليها أو قبوله فيها

بسبب عوامل مثل درجاته المدرسية ، هنا وبفضل استعدادات سابقة تمهد للاضطراب ، يكون التلميذ في حاجة إلى حل مشكلاته المتعلقة بالدراسة والمدرسة . وهنا أيضاً تفيدنا الاختبارات في كشف أسباب الاضطراب وفي تبصير التلميذ بها فيعمل على تلافي ما أدى إلى اضطرابه .

٥ - برامج الاختبار :

لكي تقوم أجهزة التربية والتعليم بوظائفها ، ولكي يقيم رجال التربية ما حققته نظم التعليم من الأهداف العامة للتربية في مجتمعنا ، ولكي نحسن تقدير التلاميذ ونقيم المدرسة والإدارة ، لكي نصل إلى هذا كله نحتاج إلى تخطيط وتنفيذ برامج شاملة للاختبارات النفسية والتربوية في المدرسة . وهذه البرامج ، كما سنرى فيما بعد ، هي تجميع ، بطريقة ما ، لبعض من الاختبارات النفسية والتربوية التي تتمشى مع الهدف المرجو من برنامج الاختبارات .

٦ - القراءة العلاجية :

تطبق الاختبارات النفسية والتربوية على التلاميذ في المدارس وفي العيادات ويقع التاميد أحياناً في أخطاء يعاقب عليها بالحصول على درجة منخفضة في الاختبار . ولكن استخدام الاختبارات في مجال القراءة العلاجية يكشف عن الأخطاء المعينة التي يقع فيها التلميذ أثناء القراءة فهو قد يخطئ دائماً في قراءة حروف معينة أو كلمات معينة أو يخطئ في فهم عبارة ترتبط بموقف مؤلم مرّ به . وهنا لا نهتم بدرجة الخطأ ولكن بنوعه ونهتم أكثر من هذا بتكشف الأسباب التي يرجع إليها فشله أحياناً في الاختبار .

٧ - تصنيف التلاميذ :

قد تُجرى الاختبارات في المدرسة لغرض تقسيم التلاميذ في الفرقة الدراسية الواحدة إلى فصول ، يكون كل فصل منها مجموعة متجانسة من

حيث المستوى العقلى أو التحصيلى أو الاقتصادى والاجتماعى حسب
فلسفة المدرسة .

ثانياً فى الصناعة :

١ - اختيار العمال :

أثبتت الاختبارات النفسية والمهنية نجاحاً فى الكشف عن الأفراد الصالحين
للعمل واستبعاد غير الصالحين : كما نجحت الاختبارات فى اختيار الأفراد
الصالحين للتدريب وإبعاد الذين لا نتوقع تقدمهم فى التدريب ، كذلك
ساعدت الاختبارات على ترشيح الأفراد الذين يستحقون الترقية أو المكافأة
أو تولى المراكز الرئيسية بعد الكشف عن كفايتهم فى العمل .

٢ - تصنيف العمال :

عندما يتقدم عدد قليل من الأفراد لشغل وظائف فى مؤسسة صناعية ،
تقوم المؤسسة بتصنيفهم حسب الأعمال . بعد الكشف عن قدراتهم
واستعداداتهم حتى يمكن وضع كل منهم فى العمل الذى يناسبه .

٣ - التقييم :

يمكن استخدام أنواع من الاختبارات النفسية (هى اختبارات الكفاية
Proficiency التى سنتحدث عنها فيما بعد) فى تقييم العمل وطرق أداء الأفراد
لوظائفهم وفى تقييم الآلات والأدوات ونظام الإشراف بل وجهاز
العمل كله .

٤ - التدريب :

قلنا إن الاختبارات نجحت فى الكشف عن يتوقع أن يفيدوا من التدريب
أكثر من غيرهم . وقد نجحت الاختبارات أيضاً فى تحديد أوجه النقص فى

أداء العمال وبخاصة ما يمكن منها علاجه بتنظيم برامج التدريب . بل استطاعت الاختبارات أن تقيس مدى تقدم التدريب وتقيم كفاية المدربين القائمين بالتدريب وبرامج التدريب ذاتها .

٥ - الحوادث :

يكاد يكون هناك اتفاق بين علماء النفس المهني على أن هناك أفراداً أكثر ميلاً للوقوع في الحوادث أو تسببها . وقد تمكنت الاختبارات من اكتشاف هؤلاء واستبعادهم فقل لذلك عدد الحوادث في كثير من الأعمال .

٦ - الموضوعية في الحكم :

الاختبارات الموضوعية أداة يمكن بواسطتها الحكم على الأفراد بطريقة لا تعرض القائمين بالحكم للأخطاء أو للتحيّز والتعصب . كما أن من نقيمتهم لا يعترضون على التقييم ما دامت الأداة موضوعية مقننة . وهكذا يحجم الأفراد من المستويات المهنية المنخفضة عن التقدم للعمل إذ أن الاختبارات ستستبعدهم .

٧ - التوجيه المهني :

يهدف التوجيه إلى دراسة الفرد وقدراته واستعداداته وميوله وقيمه وسماته ، ثم دراسة المهن ومقتضياتها ثم مساعدة الفرد على الالتحاق بأحد المهن . والاختبارات تفيد في تحليل الفرد وفي تحليل العمل ، وهكذا يمكن أن نوجه الفرد إلى عمل يلائمه أكثر .

ثالثاً : في الجليش :

كان للحربين العالميتين الفضل في دفع حركة القياس الجمعي بالذات والاختبارات النفسية عموماً دفعة قوية . فكانت الاختبارات تطبق على

الملايين الذين أتوا من آلاف المهن ليوجهوا إلى مئات الخدمات العسكرية : وبفضل اتساع نطاق الإجزاء والإمكانيات الضخمة أمكن تحديد مقتضيات مختلف الخدمات العسكرية والأسلحة ؛ وتحليل المتقدمين ، والمواءمة بين الواجبات المطلوبة والإمكانيات المتوفرة لدى المجندين والمتطوعين ، وقد حظى سلاح الطيران بكثير من الاختبارات التي قننت في مجاله لتطوير طرق الاختيار للطيارين والهابطين بالمظلات ولاكتشاف الأخطاء التي تتضمنها عمليات الطيران المختلفة .

رابعاً : في علم النفس الإكلينيكي :

توفر الكثير من الاختبارات على دراسة عوامل الاضطراب وتشخيصها وعلاجها ونجح في التحليل الكيفي والكمي للتكوين النفسي للأفراد المحتاجين إلى خدمات نفسية فردية .

خامساً : في الإدارة :

استخدام الاختبارات في مجال الإدارة تطبيق لاستخدامها في مجال الصناعة وإن اختلفت الأعمال في المجالين . ولكن الخبرة التي اشتقت من مجال الصناعة في الاختيار والتوجيه والتدريب والترفيه والتوظيف والنقل وتحديد الاختصاصات والمسؤوليات ، كلها كُيفت بحسب طبيعة العمليات المتضمنة في الإدارة وأهدافها .

كما نجحت الاختبارات ، على نفس الأسس ، عندما طبقت في مجال الوظائف الحكومية في نفس الوظائف تقريباً .

أغراض القياس النفسي والتربوي

أولاً : المسنج :

يقصد بالمسح Survey حصر الإمكانيات النفسية وتستخدم الاختبارات

النفسية والتربوية في تحديد المستويات العقلية والوجدانية والتحصيلية لمجموعات التلاميذ والعمال والمجندين . وهذا المسح لازم لتخطيط برامج التدريب والتعليم والعلاج بعد التشخيص . فالاختبارات هنا أداة لإحصاء المستويات الموجودة وبمقارنتها بالمستويات المطلوبة ، يمكن إنشاء برامج للعلاج والتعليم والتدريب ، في حدود الإمكانيات .

ثانيا : التنبؤ :

نقيس ونقيم الفرد والجماعة ، في وظائف معينة ، في وقت معين . وبافتراض ثبات سلوك الإنسان في حدود معينة ، ومرونته في حدود معينة أيضاً ، وخضوعه لكل نظريات علم النفس في حدود معينة ، كذلك يمكننا معرفة المستوى الحالي للفرد أن نقدر المستوى المتوقع أن يصله في نفس الوظائف التي قسناها .

ومن تطبيقات التنبؤ يمكننا معرفة مدى الاقتصاد الهائل الذي تحققه الاختبارات في كثير من المجالات وخاصة الصناعة . ففي دراسة لاختيار الأفراد الصالحين للعمل أدى استخدام الاختبارات إلى أن أصبحت نسبة الفاشلين في العمل ١٠٪ وكانت ٣٠٪ من قبل . وفي دراسة لاختيار الأفراد الصالحين في التدريب وفر أحد البنوك مبلغ ١٩٢٠٠٠ دولار .

ثالثا : التشخيص :

تستخدم الاختبارات النفسية والتربوية في تحديد نواحي القصور وتبيان الضعف والقوة في قدرات الفرد . وعند إجراء الاختبارات على ضوء هذا الهدف فإن معرفة الجوانب التي يعاني منها الفرد أو يحتمل أن تسبب اضطرابه مستقبلا . وتعتمد الاختبارات النفسية المختلفة على طرق تشخيصية مختلفة من مثل تحليل نموذج القدرات والاستعدادات ، وتحليل الجوانب المزاجية

والانفعالية ، وتحليل تشتت الاستجابات ، كما يقدم بعض الاختبارات معادلات للتناقص أو التدهور :

رابعاً : العلاج

بعد المسح والتشخيص نتعرف على نواحي الضعف وجوانب القصور ونبدأ في توليها بدراسة أعمق لمعرفة أسبابها ودينامياتها . وهنا تتكون لدينا صورة واضحة عن التكوين النفسى للفرد ، من حيث الوظائف التى قسناها . فإذا توفر لدينا كل هذا وتوفر لدينا المؤهلون للقيام بعمليات العلاج وإعادة التعليم والتكييف يسهل علينا أن نفيد من الاختبارات التى صممت بغرض العلاج أو التى يمكن استخدامها فى هذا الغرض .

تحديدات القياس :

لورجعنا إلى المقارنة بين الكيمائى وفراز القطن ومطبق الاختبارات النفسية لو جدنا أن السيكلولوجى وإن اتفق معهما فى طريقة الفحص ونوع الأداة إلا أنه يختلف عنهما فى نواحي أساسية . وهى أن تقديره ، لما يفحصه من سلوك المفحوصين ، لم يبلغ بعد دقة تقدير الكيمائى مثلاً . كما وأن درجة الموضوعية التى تتوفر فى المقاييس الطبيعية لا تتوفر فى المقاييس النفسية .^١ هذا فضلاً عن أن الميزان ، الذى يزن به أخصائى القطن ، يعطى نفس الأوزان دائماً لنفس الأشياء (بالات القطن) بينما يغلب ألا يكون تقدير السيكلولوجى ذا ثبات مماثل . هذا إلى أن الكيمائى ، أى كيمائى ،^٢ سيعطى نفس التقدير [بينما يختلف ، أحياناً ، السيكلوجيون فى تقدير سمات شخص واحد ولو اختلافاً قليلاً : ثم إن الكيمائى عندما يقوم بعملية معينة يعرف أكيداً أنه بصدد اختبار شق معين أو خاصية معينة بينما لا تتوفر هذه الدرجة من التأكيد فى القياس النفسى ، خاصة وأننا لا نقيس الخاصية

المعينة مباشرة إذ أننا نقيس السلوك الذى نستدل منه على هذه الخاصية .
 هذا إلى أن هناك فرق جوهري آخر ينبغي ألا يغيب عن الأذهان .
 وهو أننا فى حالة الاختبار الكيميائى نكون بصدد مشكلة نظرية هى فهم
 خاصية ما للمادة بغض النظر عن تأثير هذا الاختبار على المادة المختبرة فنحن
 مثلاً قد نتلفها أو نحولها كلها إلى غازات ولا نبقى عليها فى صورتها الأولى .
 بينما فى حالة الاختبار النفسى تكون المادة المختبرة هى الفرد نفسه الذى
 لا يتصور أننا نختل أن يلحق به ضرر . نتيجة للقياس ، فعلى العكس نحن
 نهدف أساساً إلى تحقيق مصلحته أو معاونته على تحقيق مصلحته ومصلحة
 جماعته أو الجنس البشرى كله .

ورغم ما حققته حركة القياس من نجاح فى ميادين القياس المختلفة
 وما قدمته من وسائل وأدوات بلغت درجة كبيرة من الموضوعية والدقة فى
 كثير من النواحي إلا أننا ما نزال ، حتى الآن ، ويبدو أننا سنظل هكذا
 فترة نرجو ألا تطول ، لا نزال نواجه الكثير من المشكلات الفنية فضلاً عما
 سبق أن عرضنا له . فما يزال أماننا العديد من الموضوعات التى لم نستطع أن
 نطور فيها مقاييس بلغت الموضوعية والدقة المرجوة . كما أن بعض المجالات
 لا يزال مبهماً غامضاً المعالم إلى حد ما . فلم نتمكن بعد من وضع تعريفات
 إجرائية لكثير من سمات الشخصية ، مثلاً ، بما يسهل علينا إنشاء أدوات
 لتقديرها . ونتيجة لهذا السبب نفسه ونتيجة لأن الإنسان ، موضوع
 القياس ، كائن عضوى كيميائى نفسى مركب بطريقة يستحيل معها تحليل
 ذلك الكل إلى جوانب بذاتها تصلح للقياس . نتيجة لكل هذا نجد أنفسنا
 أحياناً إزاء أدوات تقيس جوانب ليست نقية فيتشعب الاختبار ببعض ما كان
 يفضل ألا نقيسه ونحن بصدد جانب بذاته .

إلا أن كل هذه الاعتراضات وذلك القصور لا يجب أن يقعدنا عن

المزيد من الدراسات العلمية والعملية حتى نضل إلى قياس دقيق نقى لما نحتاج إلى التعرف عليه فقد أنجز العلم الكثير وليس ذلك على العلم بمستحيل .

من يقيس ومتى يقيس :

بعد أن تعرضنا لتحديدات الاختبارات ونواحي قصورها وما يكتنف عملية القياس النفسى والتربوى من تعقيد ، نجد أنه من اللازم أن نكون على حذر شديد عندما نستخدم الاختبارات النفسية والتربوية كأدوات للتقييم والتنبؤ والتشخيص . وتنحصر الأسباب الرئيسية التى تلزمنا بتقييد استخدام الاختبارات النفسية فيما يلى :

١ - تأهيل الفاحص : تختلف الشروط الواجب توافرها فيمن يقوم بإجراء الاختبار تبعا لنوع الاختبارات المستخدمة والغرض من تطبيقها . فمن الاختبارات ما يمكن أن يقوم بتطبيقه شخص على تدريب فى متوسط مثل اختبارات التحصيل يمكن أن يجريها مدرس الفصل الذى درب عليها ، وهذا التدريب لا يستغرق فى العادة إلا فترة قصيرة ضمن البرامج العامة فى معاهده إعداد المدرسين وعلى النقيض من هذا نجد أن بعض الاختبارات النفسية مثل اختبارات الذكاء الفردية والاختبارات الإسقاطية يتحتم أن يعد لها الإخصائى النفسى إعداداً فنياً طويلاً يقوم على أساس دراسة نظرية وتدريب عملى عميق تحت إشراف خبراء . هذا ويجب ملاحظة أن الطلبة الذين يجرون الاختبارات بفرض التدريب ليسوا أكفاء ولا يمكن الأخذ بنتائجهم أو تفسيراتهم .

على أن الحاجة إلى فاحص مؤهل تصبح ماسة فى ثلاث نواحي أساسية :
الناحية الأولى اختيار القياس الذى سيستخدمه : والناحية الثانية التطبيق
سليم . والناحية الثالثة هى تفسير نتائج الإجراء . هذا إلى أن الفاحص
الكفء يجب أن يضع فى اعتباره تكاليف إجراء الاختبار من حيث الزمن

والجهد والمال ، كذا يجب أن ينظر إلى الاختيار من حيث سهولة الاستخدام وإمكان نقل مواده من مكان إلى آخر . ومن الاعتبارات الهامة أيضاً في اختيار المقياس سهولة وسرعة التصحيح .

كل هذه المعلومات عادة نجدها في كتالوجات الاختبارات ويجب أن يحسب حسابها في تخطيط برامج الاختبار . على أن هناك مسألة أساسية في تقييم الاختبار نفسه وهى النظر إلى ثباته وصدقه ومعايره(*) . وهكذا يستطيع مستخدم الاختبار أن يحدد أى الاختبارات يناسب غرضاً معيناً وفئة معينة من المفحوصين .

كما يجب أن يتبع الفاحص التعليمات بدقة وأن يألف التعليمات المقننة ، وبخاصة إذا لزم أن نقارن بين نتائج تصحيح اختبار واحد أجره فاحصون مختلفون أو عند تصحيح الاختبار بناءً على معايير مختلفة . كما يلزم الضبط الدقيق لظروف الإجراء وإقامة علاقة طيبة بين الفاحص والمفحوص .

والتفسير الدقيق للدرجات يتطلب فهماً شاملاً للاختبار طبيعته ووظائفه ، وللمفحوصين طبيعتهم ومستواهم ، وظروف الإجراء والغرض منه ، وبطبيعة الجماعات التى طبق عليها الاختبار أثناء عمليات تقنيته كما تفيدنا كثيراً المعلومات الأساسية عن المفحوص . خاصة وأن المفحوصين المختلفين قد يحصلون على درجة واحدة . وليس معنى هذا بالطبع تساويهم تماماً في الوظائف المقاسة ، فإن تشابه النتائج لا يعنى بالضرورة تشابه الأسباب . فتفسير الدرجات المقاربة غالباً ما يختلف . ولكى نصل إلى تفسير ملائم في كل حالة يجب علينا أن نراعى العوامل السابقة وأن نضع في الاعتبار أيضاً العوامل الخاصة التى قد تؤثر على درجة الفرد مثل ظروف الإجراء غير العادية ، كحدوث ضجيج قرب مكان الاختبار ، والحالة الانفعالية والجسمية

(*) انظر الفصول الخاصة بالثبات والصدق والمعايير .

الراهنه للمفحوص ، وخبراته القريبة وطرقه فى الإستجابة .

٢ - ألفة المفحوص بالاختبار : إن المغالاة فى استخدام الاختبارات النفسية والتربوية كلما تراءى للفاحص أن يقدر المفحوص تجعل المفحوصين يحصلون على تدريب وخبرة بالاختبارات النفسية . وهنا لا نستطيع أن نؤكد أن درجة الفرد تمثل مستوى قدرته بل قد تمثل هنا قوة ذاكرته القريبة أو البعيدة وأثر التدريب والمران . فالمفحوصون ، أحياناً ، يحفظون الإجابات ويستجيبون بالتذكر . وهناك خطر آخر ، فقد يعطى المدرس تلاميذه أسئلة تشبه الموجودة فى اختبار الذكاء الذى سيجرى عليهم كى يعدهم جيداً للموقف . فزيف الاستجابات وتقل قيمة الاختبار كأداة تنبئية .

ويهمنا فى الحديث عن تأثير المران والتدريب أن نعرف مدى هذا التأثير وهل يشمل الاختبار كله أم فقرات منه أم يشمل نوعاً بأكمله من أنواع السلوك التى صمم الاختبار لقياسها . ومن الطبيعى أن خبرات التعلم التى يمر بها الفرد ، سواء كانت رسمية أم غير رسمية ، فى المدارس أم خارجها ، ستؤثر على أداء الفرد للاختبارات التى تمثل جوانب قريبة من تنظيمات السلوك التى تعلمها . ونوع التدريب ونوع العوامل التى دربت ونوع الاختبار أمور يتوقف عليها ما إذا كان المران والتدريب سيؤثران على أداء المفحوص لفقرات معينة من الاختبار المعين أم على كل الفقرات من هذا النوع المعين أم على مجال أوسع من أوجه النشاط . ولا تصبح درجة الاختيار غير صادقة بمعنى أنها لا تمثل مستوى القدرة الراهن لدى المفحوص إلا إذا كان مرور المفحوص بخبرات معينة يرفع درجته على الاختيار دون أن يزيد فى مستوى قدرته زيادة جوهرية . أما التدريب على فقرات تماثل فقرات الاختبار ، وبخاصة فى اختبارات الذكاء ، كما حدث فى اختبار ستانفورد بينيه للذكاء مثلاً ، فهو يرفع نسبة الذكاء ، ويزيد الارتفاع كلما زاد وجه الشبه بين الفقرات التى درب عليها المفحوص وفقرات الاختبار : والتكرار

أو المران عادة ما يؤثران على أداء الاختبار ، ولكن التأثير هنا أقل من الحالة السابقة . ويجب أن نلاحظ أن التدريب والمران قد يغيران من طبيعة الاختبار ما دام المفحوصون المختلفون يستخدمون مناهج مختلفة في حل المشاكل الواحدة التي يتضمنها الاختبار . إذ قد تصبح اختبارات الحساب مثلا اختبارات تذكر وسرعة نتيجة لأثر التدريب . وقد وجد أن الاختبارات التي يعتمد أداؤها على اكتشاف قاعدة عامة ترتفع الدرجة عليها بالتدريب إرتفاعاً يصل عند بعض المفحوصين إلى ٧٠٪ . ويصل عند البعض الآخر إلى ٢٠٠٪ كما في حالة اختبارات المتاهات ورسوم المكعبات .

كما أن أداء اختبار شبيهه بآخر يجعل التدريب على أحدهما يؤدي إلى زيادة الدرجة على الآخر . وكذلك إجراء صورة من الاختبار ثم إعادة الإجراء بصورة مكافئة من نفس الاختبار قد يرفع درجة المفحوص . ثم إن الشخص الذي أصبح ذا خبرة وتدريب على أغلب أنواع الاختبارات يمتاز على غيره عند أداء أى نوع من الاختبارات . فإن شيئاً شبيهاً بالاتجاه Attitude يكون قد نما لديه ، فهو مثلاً لا يشعر بالغرابة والحيرة في موقف الاختبار ، كما أن ثقته تكون أعلى ، كذا فهناك تداخل بين الكثير من الاختبارات من حيث الوظائف ونوع المحتوى . على أن النتائج تختلف ، واختلافها لا يقتصر على درجة التشابه بين الاختبارات المختلفة ، بل يتعدى ذلك إلى السن ، والتعليم ، والخبرة السابقة للمفحوص بالاختبارات .

ماذا يجب مراعاته في تطبيق الاختبارات :

- ١ - الدوافع والعلاقة الطيبة : كل اختبارات القدرة يجب أن تتوفر فيها الظروف التي تكفل للمفحوص أن يبذل أقصى جهده . ولهذا يجب أن توحّد الظروف وأن يدفع كل مفحوص إلى بذل أقصى جهده فعلاً ؛ وقد وجد مثلاً أن الحوافز ، أياً كان مستواها ، تدفع المفحوص إلى بذل كل

جهده ، فقد ارتفعت درجات تلاميذ إحدى المدارس في اختبارات الذكاء والحساب نتيجة استخدام التشجيع والطمأنينة .

كما قامت دراسات كثيرة على أثر الخوافز مثل الاستحسان والثناء والتشجيع ، وتدعيم مشاعر النجاح والفشل ، وتوفير ظروف المنافسة الفردية والجمعية ، ومعرفة النتائج ، وتأثير وجود ملاحظين ومعاونين في موقف الاختبار ، وأثر الجوائز والمكافآت المالية . ولكن نتائج هذه الدراسات كانت متضاربة ، فأحيانا لم تؤد بعض الخوافز إلى تغيير جوهري ، وأحيانا أدى بعضها إلى انخفاض درجة المفحوص . ولتوضيح ذلك يمكن أن نشير إلى التجربة التالية ، فقد اختبرت ثلاث مجموعات من طلبة المعاهد باختبارات لفظية وحسابية واختبارات شطب . وقد تم الإجراء في ثلاث فئات من الظروف هي : (١) ظروف مضبوطة عادية (ب) إعطاء تعليمات للمفحوص بأن يعمل بأقصى دقة ممكنة (ح) والمجموعة الثالثة من الظروف كان الفاحص فيها يقطع الإجراء كل ثلاث ثوان بأن يقول للمفحوصين إن المفروض أنهم أجابوا حتى الآن عن عدد معين من الفقرات . وكان هذا العدد دائما يستحيل على أغلبهم : والنش (ح) من الظروف تثير التوتر وتتضمن مشاعر الدونية والفشل ما دام المستوى الذي نطالب به المفحوصين يفوق إمكانياتهم ومن ثم كثرت أخطاؤهم .

كما أن عوامل الدفع تلعب دوراً كبيراً بالنسبة لنمط معين من المفحوصين أكثر مما يحدث بالنسبة لنمط آخر وبخاصة إذا كانوا مضطربين . وكذلك فإن مشاعر عدم الأمن والإحباط تؤثر على الأداء خاصة بالنسبة للأطفال . وكذلك تؤثر عوامل من مثل شعور المفحوص بأنه من جماعة أقلية عنصرية وأن الفاحص من الأغلبية التي تعادى جماعته : وهذا موجد ما نفيده من نتائج التجارب على الدوافع ؛ فهي أولا ترفع أو تنخفض درجات المفحوصين وهذا الرفع أو انخفاض يختلف باختلاف فئات المفحوصين . والدوافع ،

ثانياً ، يجب أن تؤخذ في الاعتبار عند تفسير درجة المفحوص ظروف الدفع غير العادية ، خاصة إذا كان المفحوص ذا خبرات أساسية تختلف عن خبرات أفراد عينة التقنين : وثالثاً ، يجب إقامة علاقة طيبة قبل إجراء الاختبار .

ويختلف الأسلوب الفني في إقامة العلاقة الطيبة Rapport باختلاف نوع الاختبار ونوع المفحوصين وإعداد الفاحص . فمثلاً ، عند اختبار أطفال ما قبل المدرسة نضع في الاعتبار ما يتميزون به من خجل مع الغرباء ، وسرعة تشتت انتباههم ، والسلبية والعناد . إزاء هذا يجب أن يكون الفاحص ودوداً ، مرحاً ، وأن يكون سلساً مطمئناً للطفل . فالطفل الخجول يلزمه وقت أطول ليألف الملاحظات :

كما يجب استبعاد عامل الدهشة من جانب الفاحص . فلا يصدر منه ما يتم عن استغرابه لأداء المفحوص . وكذا يحسن أن يعلم المفحوص ، مقدماً أي قبل الإجراء لمدة ، بأنه سيختبر حتى يوهل نفسه للموقف . وبأحبذا لو وزعت كتيبات تبين طبيعة الاختبار ووظائفه وتورد مساعدات على الاحتفاظ بالهدوء والألفة بموقف الاختبار . وأن يبذل جهد لإقناع المفحوص بأنه يجب أن يهتم بأن تكون درجته صادقة ، أي أن تمثل سلوكه العقلي في مواقف تشبه ما جاء بالاختبار فلا يهمه أن يبين ما يفضل أن يعمل به ، أو ما يعتبر أنه واجب عليه . وعلى ذلك فلا داعي للغش وعدم الصراحة ، حتى لا يترتب على هذا دخول المفحوص مدرسة أو مهنة ، مثلاً ، لا تناسبه وتسبب له متاعب يستطيع تجنبها إذا كان صادقاً ومتعاوناً :

ولهذا يراعى دائماً أن نبدأ الاختبار بفقرات من مستوى سهل نسبياً Preparatory Shock Absorbers كي يكتسب المفحوص الثقة بنفسه ويعتاد الإجابة عن نوع أسئلة الاختبار . هذا بالإضافة إلى إعطاء أسئلة للتمرين في تعليمات الاختبار وأسئلة تعدده للموقف .

٢ - الغش والتزييف :

هذه إحدى مشاكل الدفع وعلاجها إقناع المفحوص بأن صدق الدرجة (بالمعنى السابق ذكره) يجب أن يكون همه ، إذ أن تزييفها قد يسبب له المتاعب فيما بعد : وهذا الإقناع قد لا يكون جدي فعال مع أنماط معينة من المفحوصين : ولكن ، على أى حال ، نودى ظروف الدفع السليم وإقامة علاقة طيبة إلى تقليل فرص الغش والتزييف : على أن المفحوص أحياناً ما يكون شديد الاهتمام بالوصول إلى هدف معين يتوقف على درجته في الاختبار كأن يقبل بالمدرسة أو يختار في المهنة ، كما أن المفحوص أحياناً ما يكون شاعراً بعدم الأمن ، وفي هاتين الحالتين يكون هناك احتمال كبير بأن يبذل محاولات لتزوير الدرجات .

وقد دلت البحوث على أن تزييف درجات الاختبار النفسى والتربوى تصنف في الفئات الثلاث الآتية :

- (أ) محاولة المفحوص أن يرفع درجته في اختبارات القدرة .
 - (ب) محاولة الظهور بمظهر طيب في اختبارات الشخصية . (ج) التمويه باللونية أو اللاسواء في كل من اختبارات الشخصية واختبارات القدرة .
- ويعتمد المفحوص أثناء محاولته لرفع درجته في اختبارات القدرة على عدة وسائل مثل : (١) الاعتماد على المساعدات الخارجية . (٢) الغش من جاره في الاختبار الجمعى . (٣) أن يأخذ زمناً أطول من المقرر بأن يبدأ قبل غيره أو أن يستمر رغم انتهاء الزمن المحدد . (٤) كما قد يعتمد المفحوص على معرفته السابقة لفقرات الاختبار . (٥) التخمين أو الإجابات العشوائية على ما لا يعرفه من الأسئلة كى يستفيد من عامل الصدفة (ولكن هذا يمكن تلافيه عند التصحيح بطريقة خاصة سيأتى شرحها في فصل آخر) :

والمفحوص عند محاولته الظهور بمظهر طيب في اختبارات الشخصية

يعتمد على أحد خصائص هذه الاختبارات ونقصد أن لكل أسئلة اختبارات الشخصية إجابة واحدة مقبولة والفرد يميل إلى أن يوافق على هذه الإجابة الطيبة ، خاصة وأن بعض الاختبارات تتطلب من المفحوص أن يجيب بنعم أو لا عن فقرة مثل : هل تشعر بالضيق عندما تقابل الغرباء ؟ . واضح أن لمثل هذا السؤال إجابة واحدة طيبة ، وهى « لا » ، والمفحوص عادة أميل إلى أن ينفى عن نفسه مثل هذا الشعور . فهو لا يجيب بما يصدق عليه فعلا بل بما يستحسنه ، أو بما يشعر أنه مقبول لدى الآخرين ، خاصة إذا كان الاختبار لغرض الانتقاء المهنى .

وقد قامت دراسة طبق فيها اختبار الشخصية لبرنرويتز Bernreuter Personality Inventory ، مرتين على مجموعة واحدة من المفحوصين . وطلب إليهم فى المرة الأولى أن يتظاهروا فى إجاباتهم بالرغبة فى الالتحاق بوظيفة بائع . وعند إعادة الإجراء طلب إليهم أن يستجيبوا مدعين الرغبة فى شغل وظيفة أمين مكتبة . كان الفاحص يطلب من المفحوصين فى مرقى الإجراء أن يجيبوا وكأنهم فعلا ، برغم رغبتهم الحقيقية ، يرغبون فى الالتحاق بوظيفة بائع فى المرة الأولى ، وأمين مكتبة فى المرة الثانية . ونحن نعلم أن هاتين الوظيفتين تختلفان من حيث درجة ثقة من يشغل كل وظيفة منهما بنفسه ، ورغم هذا أمكن للمفحوصين أن يجيبوا فى المرة الأولى لإجابة البائع الناجح ، وفى الثانية لإجابة أمين المكتبة الناجح . رغم أن الأولى تختلف عن الثانية من حيث الدرجات على مفاتيح الانبساط ، الثقة بالنفس ، الاجتماعية ، والمفحوص عند ما يود الظهور بمظهر العاجز أو غير السوى فى اختبارات الشخصية واختبارات القدرة ، يدعى مثلا أنه ناقص العقل أو أنه يعانى من اضطراب انفعالى وذلك بغرض التهرب من التجنيد أو من اختياره لعمل لا يتفق مع رغباته أو يرى فيه خطراً عليه . وقد وضعت بعض اختبارات القدرة واختبارات الشخصية حلاً لهذا المشكل . إذ أعدت مفاتيح خاصة (٣ - المقياس)

لكشف التعميه بالدونية أو اللاسواء : والافتراض الأساسى الذى تعتمد عليه هذه المفاتيح يتلخص فى أن من يحاول التزييف يبالغ فيه . فعند تزييف اختبارات الذكاء قد ينجح المفحوص فى فقرات صعبة ثم يعتمد أن يجيب إجابة خاطئة عن فقرات أسهل ؛ مما يتنافى مع اتساق القدرة العقلية . وهنا يمكن اشتقاق المفتاح من إجابات ضعاف العقول ومقارنتها بإجابات مفحوصين طلبنا منهم أن يتصنعوا الغباء فى إجاباتهم بعد أن أعددناهم فنياً لهذا التزييف . وأحد حلول مشكلة التزييف فى اختبارات الشخصية هو إخفاء الهدف من الاختبار عن المفحوص . فعند إجراء اختبار يقيس « الذكورة - الأنوثة » يمكن مثلاً أن نقول للمفحوصين إننا نحصر الاهتمامات والاتجاهات ونقيسها . ولكن هذا الإجراء لا يتوفر فى أغلب اختبارات الشخصية إذ أن الكثير من الفقرات يكشف عن نفسه . وهناك طريقة لمعالجة الغش والتزييف ، وهى طريقة الاختيار الجبرى ، أو طريقة الأسئلة غير الشفافة . وذلك بأن تتضمن الفقرة استجابتين تبدوان متساويتين ، وعلى المفحوص أن يختار منهما واحدة تنطبق عليه . ونقصد بتساوى الاستجابتين أن المفحوص يتقبلهما بدرجة واحدة أو أن المفحوص يعتقد أن المجتمع يتقبلهما بدرجة واحدة . وهناك طرق أخرى لثلاثى التزييف أو كشفه ، وسنتحدث عنها بالتفصيل فى فصل آخر . أما فى اختبارات القدرات والاختيارات التحصيلية فتوجد فرصة الإجابة بالتخمين ، ولكن لهذا علاجه باستعمال معادلة التصحيح التى سيأتى الكلام عنها فى فصل آخر .

٣ - الإعداد لموقف الاختبار :

(١) يقوم الفاحص فى العادة بقراءة التعليمات من كراسة تعليمات الاختبار ، ولكى يكون الموقف طبيعياً ، غير متكلف ، يجب أن يألّف الفاحص التعليمات اللفظية . وتأتى هذه الألفة من دراسة التعليمات وقراءتها عدة مرات قبل الإجراء حتى لا تحدث أخطاء أثناء الإجراء .

(ب) ويجب أن يقوم الفاحص مقدماً ، أى قبل الإجراء ، بإعداد مواد الاختبار فيضعها على مائدة قريبة من مائدة الاختبار ، حتى تكون في متناوله فلا يبذل جهداً ولا يتعطل الإجراء ولا ينزعج المفحوص أو يتشتت انتباهه :

(ج) وعند استخدام جهاز آلى فى الاختبار ، يجب الكشف مبدئياً على سلامته والتدرب على استخدامه فى سهولة .

(د) وفى الاختبار الجمعى ، يقوم الفاحص بإعداد المكان ووضع موائد الاختبار على أبعاد مناسبة وأن يتأكد من أن الموائد خالية من كل ما يعطل المفحوص أو يعينه فى الإجابة وأن يتأكد من حسن إضاءة الغرفة وتهويتها وبعدها عن مشتتات الانتباه كالضوضاء .

(هـ) ويستطيع الفاحص أن يقوم بعمل « بروقه » على مفحوص بأن يحاول معه الإجراء بالطريقة التى سيتبعها مع من سيختبرهم حقاً .

(و) يقرأ الفاحص التعليمات بصوت واضح وبسرعة معتدلة ويهتم بالتوقيت ويتأكد من أن جميع المفحوصين يتبعون التعليمات ، ويجب عن الأسئلة الفردية للمفحوصين فى حدود ما تسمح به التعليمات ، ويمنع الغش ، ويجمع كراسات الإجابات من المفحوصين فى وقت واحد وإن لم يستطع ذلك بمفرده استعان بمن يعاونه ، ويحسن إضاءة ضوء أحمر على باب غرفة الإجراء حتى لا يدخل من يعطل الإجراء سواء كان مفحوصاً تأخر عن ميعاد الإجراء أو شخصاً آخر ، وفى الاختبارات الجمعية أيضاً يفضل أن يقف معاون للفاحص بالباب مانعاً الدخول من لحظة بدء الإجراء :

ومن المهم تحديد مدى تأثير ظروف الإجراء على درجات المفحوص ، فنشاط المفحوص يؤثر على إنتاجه خاصة لو أدى هذا النشاط إلى اضطراب انفعالى أو أى ظروف معطلة كاللعب والملل . وكذلك فالخبرة المباشرة قبل الإجراء قد تعطل أو تعرقل استجابات المفحوص .

المبادئ الخلقية في استخدام الاختبارات النفسية والتربوية :

١ - يجب أن يقتصر بيع وتوزيع الاختبارات النفسية والتربوية على من يحسن استخدامها وعلى الأشخاص المؤهلين في تطبيقها . وتختلف المؤهلات المطلوبة من اختبار إلى آخر :

٢ - يجب أن يقتصر تصحيح الاختبارات النفسية والتربوية على الأشخاص المؤهلين في التفسير . فعندما تقدم للفرد الدرجات التي حصل عليها في أحد الاختبارات ، يجب أن يقوم بتقديمها أخصائى مدرب في إرشاد من يضطربون إذا أعلنوا بدرجاتهم المنخفضة ، هذا بغض النظر عن صحة أو عدم صحة الدرجات :

٣ - ويجب أيضاً ألا يسمح الناشر لنفسه بإرشاد من يشتري الاختبار إلى نوع معين من الاختبارات حتى لو كان هذا الناشر أخصائياً نفسياً أو تربوياً :

٤ - ويجب ألا ينشر الاختبار قبل ثبات صلاحيته إلا إذا كان الهدف من النشر هو إجراء دراسات وبحوث عليه . وهذا الهدف يجب أن يوضح جيداً كأن ينص في كراسة التعليمات ، صراحة وبشكل واضح ، على أن الاختبار ما زال في دور التجربة :

٥ - كما يجب أن يرفق بالاختبار كراسة للتعليمات تذكر فيها طريقة تطبيق الاختبار وما يقيسه وتذكر فيها المعايير ووصف دقيق لعينات التقنين وظروف التطبيق أثناء التقنين والبحوث الهامة التي أجريت على الاختبار :

٦ - يجب عدم نشر أى جزء من اختبار في أغراض الدعاية أو الإذاعة أو الصحف لأن ذلك يؤدي إلى أضرار كثيرة أولها أن يفقد المقياس قيمته أو سرية وثانيها تكوين اتجاهات غير صحيحة نحو الاختبارات .

الفصل الثاني

المفاهيم الأساسية في الاختبارات النفسية والتربوية

مفاهيم الإحصاء والقياس

- مقدمة .
- الدرجة الخام وتفسيرها .
- الميل المركزي .
- التشتت .
- الدرجات المحولة وأنواعها .
- الدلالة والعينات .
- الارتباط .
- التحليل العاملي .
- تحليل التباين .
- القياس والتقييم .
- تعريفات ومفاهيم عن القياس والاختبارات .

مقدمة :

أصبح علم النفس علما science واستقل عن الأم الكبرى ، الفلسفة ، وبعد عن التصورات والمفاهيم الغامضة واقترب من العلوم التقديرية بعد أن استند إلى منهج البحث العلمى وعماده التجريب . والتجريب فى علم النفس ليس مستطاعا دائما ، خاصة فى صورته الكلاسيكية ، صورة التجارب العملية على الحيوان التى تقيس العلاقة بين ظروف يتحكم فيها المحرب ، ونسميها المتغيرات المستقلة Independent variables وبين عامل نتوقع أن يتأثر بتغير الظروف المضبوطة نسميه المتغير المعتمد dependent variable . قرب التجريب ، بهذه الصورة ، علم النفس إلى العلوم الطبيعية فالمنهج العلمى المتبع فيهما هو هو برغم اختلاف موضوع الدراسة . ولكن هذا التجريب هنا غير ممكن دائما .

ولنا معمل آخر ، معمل كبير هو تاريخ البشرية فلا يمكن لإحداثه من تغيير فى تكوين الإنسان موضوع الدراسة هنا يمكن دراسته وتتبعه خلال ظواهر حدثت من قبل أو تحدث الآن . هذه الظواهر لا يمكن أن تدخل إلى المعمل بل نخرج نحن إليها بالإحصاء . فالإحصاء هو الأداة التى نعالج بها الظواهر لنكشف عن العلاقات بينها وما إذا كانت هذه العلاقات من نوع علاقة السبب بالنتيجة ، أم من نوع علاقة التابع والارتباط ، ونحن نهتم بالعلاقات الوظيفية بين المتغيرات ، هذا النوع من العلاقات لا يمكن الجزم به من مجرد ارتباط مجموعتين من المتغيرات . ولهذا نلجأ إلى الإحصاء الذى يقدم لنا أدوات ومقاييس للتقدير والتقييم الكمي الموضوعى الدقيق فحسب ، بل وأيضاً يوصلنا إلى طبيعة تكوين الظاهرة موضوع الدراسة وعواملها ويمكننا فضلاً عن هذا من المقارنة بينها وبين غيرها ومن اكتشاف الفروق وجوهريتها ، بل وأبعد من هذا ، يذهب

بنا الإحصاء إلى التعرف على عوامل قد لا نستخلصها دون الأساليب الفنية الإحصائية .

بل إن المنهج التجريبي يصورته الكلاسيكية كي يعطينا نتائج نستطيع أن نعممها لا بد وأن يركن إلى الإحصاء . الإحصاء إذاً أداة لا غنى عنها للوصول إلى تقديرات رقمية تمكّننا من التفسير الأبعد . وهذا التفسير لا يعتمد على الإحصاء وحسب ، بل إنه مبني أساساً على الإعداد العلمى للقيام به . وعلى هذا فالإحصاء كما رأينا أداة ، بالنسبة لنا فى هذا المجال ، تساعدنا بحسب الأرضية العلمية من وضع فروض والتحقق منها ثم تعميم نتائجها بغرض الوصول إلى قوانين أو نظريات . وإن كان الإحصاء ، بالتالى ، منهجاً يساعدنا على تخطيط البحوث والتجارب ، وعلى معالجة نتائجها ، إلا أنه لا يستطيع دائماً أن يقطع بالنتائج ويؤكد التعميمات . وليس هذا راجعاً إلى قصور منطق الإحصاء بل إلى قصور إمكانياتنا فى قياس جوانب الحياة الإنسانية .

فالمقياس أداة موضوعية ولكنه فى الوقت نفسه عينة تمثل القدرة أو الوظيفة المراد قياسها . نحن إذاً لا نلجأ إلى كل القدرة بل إلى عينة منها ، وفى هذا ما فيه من مخاطرة . ثم إننا لا نأخذ من القدرة بل نأخذ استجابات نعتقد أنها تعكسها أو تمثلها . وهذا الاعتقاد وذلك التمثيل معرضان لكثير من الأخطاء . هذا إلى جانب صعوبات القياس الأخرى والى سبق عرضها ؛ وهكذا نرى أن الإحصاء وإن كان يقدم لنا مثلاً طريقة لحساب الخطأ الذى يحتمل أن نكون قد وقعنا فيه أثناء خطوات القياس المختلفة إلا أنه حتى فى هذا لا يقطع بالنتائج ، وإن كان الإحصاء قد أفاد القياس كما أوضحنا فإنه أيضاً قد أفاد منه ، فقد قدم المشتغلون بالقياس ، بفضل حاجتهم إليه ، طرقاً إحصائية أصبحت أساسية ومنها التحليل العاملى .

أمر آخر نريد أن ننبه إليه ثانية وهو أن دارس القياس بل ومستخدم

الاختبار لا يلزم أن يكون إحصائياً . وإن كان مصمم الاختبار يعتمد على الإحصاء في الكثير . إلا أننا هنا لسنا بصدد من يقومون بتصميم الاختبارات وإنما بصدد من يقومون بإجرائها . وهم في العادة لا يحتاجون إلى فهم الأسس الرياضية للعمليات الإحصائية وكيف استخلصت المعادلات والقوانين الإحصائية ، إذ نرى أنه يكفيهم أهمية فهم الإحصاء ونواحي قصوره ، ومتى يلجئون إليه ، وكيف يستفيدون منه ، وما هي الطرق والعمليات الإحصائية التي تعينهم في مختلف عمليات القياس على التقييم والتنبؤ . وباختصار نراهم في حاجة إلى الإلمام بالمفاهيم الإحصائية الأساسية للقياس . فضلاً عن أننا هنا لسنا ، بصدد كتاب عن الإحصاء .

الدرجة الخام وتفسيرها :

إذا قلنا إن فلاناً حصل على الدرجة ١٥ في اختبار للمفردات . ماذا نفهم من هذا ، وكيف نفسر الدرجة ؟ الحق أنها في ذاتها لا تعني شيئاً ولا يمكن تفسيرها على الإطلاق . فقد تكون الدرجة النهائية في اختبار المفردات ، هذا ، هي ١٥ فتصبح درجة فلان $\frac{15}{100}$ وهي أعلا درجة ممكنة . وقد تكون الدرجة النهائية هي ٥٠ فتصبح درجته $\frac{15}{50}$. ولنفرض أن الدرجة النهائية هي ٢٠ أي أن درجته $\frac{15}{20}$. و $\frac{15}{20}$ تبدو ، ظاهرياً ، درجة مرتفعة فهي تساوي ٧٥٪ . ولكن حتى لو علمنا هذا ، هل نستطيع أن نفسر الدرجة ؟ وماذا بعد ؟

قد يكون الاختبار صعباً جداً فنعتبر ٧٥٪ درجة عالية جداً . وقد يكون الاختبار سهلاً جداً فنعتبر ٧٥٪ درجة منخفضة جداً . وهكذا لا يكون للدرجة معنى إلا عندما يكون هناك مستوى نفهمها في إطاره .

ومدرس الفصل الذي يضع مثل هذا الاختبار يضمه بعض المفردات التي تعلمها تلاميذ فصله . أو بعض المفردات التي تقرب مما تعلموه ، أي التي تمثلها من حيث عددها ومن حيث مستوى صعوبتها . يفترض إذاً أن .

الاختبار يمثل ما تعلمه التلميذ ، وإنه بالتالى ، ذو مستوى صعوبة يناسبه .

ومدرس الفصل بعد أن يصحح الاختبار يعرف أن فلاناً هذا متقدم نسبياً عن زملائه . كيف توصل إلى ذلك ؟ وضع المدرس في ذهنه أن الاختبار يناسب قدرة المفحوص ويمثل ما تعلمه . ثم نظر إلى درجات الفصل عموماً فوجد أن نسبة كبيرة منها تقل عن درجة هذا التلميذ : أى أنه حسب متوسط درجات تلاميذ الفصل على هذا الاختبار ، أو هذا ما كان يجب عليه . ومن هنا عرف أن درجة فلان هذا عالية . فإن أراد أن يعرف مركز تلميذه بالنسبة لفصله في هذا الاختبار لجأ إلى حساب عدد التلاميذ الذين حصلوا على درجة أقل من درجته ثم قسم عدد هؤلاء ÷ عدد تلاميذ الفصل كله فحصل على نسبتهم المئوية . ومن هنا عرف أن ١٠٪ مثلاً من تلاميذ الفصل يفوقونه . أى أن رتبة فلان المئوية هي العاشرة فدرجته تزيد عن درجة ٩٠٪ من زملائه .

وإذا تصورنا درجات الفصل وأردنا أن نعرف مدى تفاوتها ومدى انسجامها : أى إذا أردنا أن نعرف درجة قرب أو بعد مختلف الأفراد عن متوسط درجة الفصل لجأنا إلى حساب التشتت .

ونحن لكي نقوم بهذه العمليات المختلفة من إيجاد النسبة المئوية والمتوسط والتشتت نحتاج إلى معرفة عدد تلاميذ الفصل والدرجة النهائية في الاختبار والدرجة الصغرى : وفي معالجتنا لدرجات تلاميذ الفصل نحتاج إلى تنظيمها بطريقة عملية تسهل علينا هذا .

ونحن نقوم بهذه العمليات بسبب عجز الدرجة الخام عن إعطاء صورة عن مستوى الفرد في بيان مركزه بالنسبة لمجموعته . فالدرجة الخام هي نتيجة إجراء المفحوص على الاختبار وتحسب عادة بعدد الأسئلة التي أجاب عنها إجابات صحيحة ، وقد تحسب أيضاً بأن تعطى درجة للإجابة الصحيحة عن مجموعة من الأسئلة ، وتختلف درجة مجموعة معينة من الأسئلة عن درجة

أخرى تبعاً لأهمية هذه الأسئلة في تقدير مستوى الفرد وتبعاً لدرجة صعوبة كل مجموعة من الأسئلة . وبعد أن استخلصنا الدرجة الخام بإحدى هذه الطرق أصبح لزاماً علينا أن نقارن درجة (أ) من التلاميذ بدرجة زميله (ب) في نفس الاختبار والأمر هنا سهل نسبياً فدرجة (ب) قد تعلو درجة (أ) أو تقل عنها . ولكن إذا أردنا أن نقارن درجة (أ) بدرجات زملائه في الفصل اختلف الأمر . كما أن مدرس الفصل عادة يقوم بتدريس نفس المادة في أكثر من فصل واحد ويحصل تلاميذ الفصول المختلفة على درجات مختلفة . فإذا أراد هذا المدرس أن يقارن بين درجات الفصل (س) مع بقية الفصول ، أو أن يقارن بين درجة التلميذ (أ) في الفصل س وبين درجة التلميذ (د) في الفصل ص اختلف الأمر أيضاً .

على أن ترتيب التلميذ في مجموعته يكون ذا معنى وأكثر قابلية للتفسير كلما زاد عدد أفراد المجموعة التي ينتمي إليها سواء كانت فصلاً أم عدداً من الفصول . وهذه الميزة تجعلنا أمام مشكلة عملية تلخص في طريقة تجميع درجات تلاميذ الفصل الواحد أو مجموعة الفصول بطريقة تسهل عمليات التفسير التالية .

وسيلتنا في هذا هي تبويب البيانات Tabulation of Data أى تجميعها في فئات على أساس مدى قرب درجات مجموعة من الأفراد أو مدى التشابه بين الأفراد الذين تشملهم الفئة . ولكي نقوم بهذه العملية نجرى الخطوات التالية :

١ - نطرح أصغر درجة حصل عليها تلميذ من أعلى درجة وصل إليها غيره ويسمى هذا بالمدى المطلق Absolute range :

٢ - نقسم هذا المدى إلى فئات لا يزيد عددها عادة عن ٢٠ ولا يقل عن ٨ لكي نحدد طول الفئة . أو مداها أى عدد الدرجات التي تشملها .

٣ - نقوم بحساب عدد الأفراد الذين نالوا درجات تنحصر في الفئة الأولى ثم الذين نالوا درجات تنحصر في الفئة الثانية ، وهكذا ويسمى هذا بالتركرار Frequency ، ومجموع التكرارات هو مجموع عدد الأفراد ، ويكون ذلك بوضع علامة تمثل درجة الفرد أمام الفئة ، وعدد العلامات أمام الفئة الواحدة هو تكرار هذه الفئة أى عدد المرات التى وردت فيها درجات تنحصر بين حدى الفئة .

ومن الطبيعى أن يختلف عدد الأفراد أمام الفئة باختلاف الدرجة التى تحدد أول الفئة والدرجة التى تحدد آخرها ، أى تبعاً لاختلاف طول الفئة وحدود الفئات .

وقد يلزم المشتغل بالقياس وخاصة عند تجريب مقياسه بعد تأليفه أن يعرف فكرة سريعة عن الصورة التركيبية لدرجات من طبق عليهم الاختبار ، وهؤلاء قد يكونون آلافاً . ولكى يصل إلى هذا يقوم بتمثيل درجات الأفراد فى رسوم بيانية . هذا العرض البيانى graphic presentation قد يكون فى صورة مضلع تكرارى أو مدرج تكرارى أو منحنى . فتمثل الفئات على المحور الأفقى والتكرارات على المحور الرأسى وتوضع نقطة تمثل مركز كل فئة وترتفع هذه النقطة على المحور الرأسى بحسب عدد الأفراد الذين نالوا درجات تشملها هذه الفئة . وتبعد هذه النقطة على المحور الأفقى على الأقل بقدر بعد مركز هذه الفئة عن أقل درجة حصل عليها فرد . وبتوصيل هذه النقط يتكون المضلع التكرارى ومنه يمكن عمل منحنى التوزيع . وفائدة هذا تلخص فى إعطاء فكرة سريعة وفى بيان مراكز تجمع الأفراد ومدى تفرقهم فتمثل المساحة المحددة بين العمودين عند نقطتين أو درجتين والمنحنى عدد الأفراد .

النزعة المركزية

ويقصد بالنزعة المركزية Central tendency ميل أفراد المجموعة إلى أن يتجمعوا في مركزها ، إذ نجد أن نسبة كبيرة من الأفراد يحصلون على درجات تشملها فئة . وأن نسباً متناقصة من الأفراد تحصل على درجات أقل وأن نسباً مماثلة متناقصة تحصل على درجات أعلى . والفئة أو الفئات ذات التكرار الأكبر غالباً ما تتوسط التوزيع التكرارى والرسم البياني . ونحن نلجأ إلى قياس هذه النزعة للتجمع أو التشتت عندما نريد أن نعطي رقماً يدل على درجات مجموعة من الأرقام ، فعندما نريد أن نصف المجموعة بدلاً من سرد درجات أفرادها يمكننا أن نقول إن متوسط درجاتهم هو كذا . وكذا هذه تأتي عادة ، في أبسط صورة ، من جمع درجات المجموعة كلها ثم قسمة هذا المجموع ÷ عدد الأفراد . أما في الجداول التكرارية فيحسب المتوسط بضرب تكرار الفئة في مركزها (مركز الفئة = $\frac{\text{مجموع بدايتها} + \text{نهايتها}}{2}$) ثم نجمع حواصل الضرب ونقسمها على عدد الأفراد (أى مجموع التكرارات) . وهناك طرق مختلفة لحساب النزعة المركزية ولكن المتوسط الحسابي (م) Mean أبسطها وأكثرها شيوعاً وأفضلها دقة في أغلب الحالات .

وإذا كان توزيع المجموعة متماثلاً أى أن عدد الأفراد الذين حصلوا على درجات أعلى من المتوسط يساوى عدد من حصلوا على درجات تقل عن المتوسط ، وكان التوزيع اعتدالياً بمعنى أن نسب من يزيدون عن المتوسط تتناقص تدريجياً بحيث أن أعلى الدرجات تحصل عليها نسبة صغيرة جداً من الأفراد ، وكذلك تحصل نسبة مماثلة على أصغر الدرجات ويحصل عدد معين من الأفراد على درجة تزيد عن المتوسط بمقدار ٢ مثلاً وفى نفس الوقت يحصل عدد مماثل على درجات تقل بمقدار ٢ عن المتوسط وهكذا .

نقول إن كان التوزيع متماثلاً واعتدالياً كان المتوسط الحسابي أفضل مقياس للنزعة المركزية ، بل وكان مجموع انحرافات من يزيدون عنه (والانحرافات هنا إشارتها +) مساوياً لمجموع انحرافات القيم التي تقل عن المتوسط (والانحرافات هنا إشارتها -) . أى أن المجموع الكلى للانحرافات يساوى صفراً . وهكذا نجد أنه لو كانت هناك قيم ترتفع كثيراً في المجموعة فإنها ترفع المتوسط الحسابي لهذه المجموعة ، وبالمثل يقل المتوسط إذا كانت هناك قيم منطرفة صغيرة . ولكن هذا لا يحدث إلا عندما يكون التوزيع ملتوياً أو مفرطحاً وبعبارة أخرى لا يكون متماثلاً اعتدالياً Normal .

ويفيد المتوسط ، كما أوضحنا في الكلام عن الدرجة الخام وتفسيرها ، في معرفة مدى انحراف درجات أفراد نبحتهم عن مستوى المجموعة . كما أننا عند مقارنة مجموعة بأخرى نقارن بين متوسطي المجموعتين لو توفرت ظروف المقارنة الأخرى مثل إجراء نفس الاختبار على مجموعتين متماثلتين .

التشتت :

عند مقارنة أفراد المجموعة الواحدة أو مجموعتين معاً ، نستخرج قيمة واحدة لكل مجموعة ، ثم نقارن الأفراد بهذه القيمة داخل الجماعة أو نقارن القيمة التي تمثل الجماعة الأولى (المتوسط) بالقيمة التي تمثل الجماعة الثانية . ولكن هذه القيمة وحدها لا تكفي لوصف المجموعة إذ نحتاج إلى معرفة التوزيع الداخلي للدرجات الأفراد ومدى قربهم أو بعدهم عن المتوسط أى مدى تشتتهم حول المتوسط . فقاييس التشتت إذاً هي مقاييس لقرب الأفراد من بعضهم ، أى لتجانسهم . وقد يتساوى متوسط درجات المجموعة (أ) مع متوسط المجموعة (ب) ولكن يختلف توزيع الدرجات في (أ) عنه في (ب) فيحصل عدد من أفراد المجموعة (أ) على درجة معينة وفي المجموعة (ب) يحصل عدد أكبر من السابق أو أصغر منه على

نفس الدرجة . وهكذا لا يفيد المتوسط وحده في وصف المجموعة ومقارنة أفرادها ومقارنتها بمجموعات أخرى . بل يجب الاستعانة بأحد مقاييس التشتت . وسنأخذ هنا الانحراف المعياري Standard Deviation لأنه أهم مقاييس التشتت وأكثرها استعمالاً ويرمز عادة للانحراف المعياري بالرمز ع :
ومن تعريفنا للتشتت نلاحظ أنه يعنى مدى تقارب أو تباعد الدرجات أو القيم عن المتوسط . ويعرف الانحراف المعياري بأنه « الجذر التربيعي لمتوسط مربعات انحرافات القيم عن متوسطها الحسابي » ويعتمد في حسابه على الخطوات التالية :

- ١ - حساب المتوسط الحسابي للقيم ويرمز له بالرمز م :
 - ٢ - حساب انحرافات القيم وذلك بطرح المتوسط من القيم المختلفة ،
أى أن الانحراف (ح) = الدرجة - المتوسط : (هذا في حالة البيانات غير المبوبة أما في البيانات المبوبة فيطرح المتوسط من مركز الفئة) :
 - ٣ - تربيع الانحرافات (ح^٢) :
 - ٤ - تجمع مربعات الانحرافات (مج ح^٢) . (في حالة البيانات المبوبة يضرب التكرار ك في مربع الانحراف المحسوب من مركز الفئة) :
 - ٥ - يقسم مجموع المربعات هذا (مج ح^٢) على عدد القيم (ن) .
 - ٦ - نستخرج الجذر التربيعي لخارج القسمة السابق فيكون هو الانحراف المعياري لهذه القيم أو الدرجات . .
- ∴ معادلة الانحراف المعياري هي :

$$\text{الانحراف المعياري ع} = \sqrt{\frac{\text{مجموع مربعات الانحرافات}}{\text{عدد الحالات}}} = \sqrt{\frac{\text{مج ح}^2}{\text{ن}}}$$

وهكذا نجد أن الانحراف المعياري ، بما أنه يعتمد على المتوسط ، يصلح

في حالات التوزيع الاعتدالي المتماثل : وهو يتأثر بالقيم المتطرفة ويتأثر لو استوفيت فئة من أفراد المجموعة . والانحراف المعياري يفيد في معرفة طبيعة توزيع أفراد المجموعة أي مدى انسجامها وهو يتأثر بالمتوسط وبتعدد المجموعة وتشتتها ومدى الاختبار المطبق لقدرة وتدريب أفرادها . كما يفيد في وصف المجموعة إذ يكفي ، مبدئياً ، ذكر متوسطها وانحرافها المعياري . وهذين المعلمين أيضاً يمكن مقارنة المجموعة بمجموعة أخرى . وأبعد من هذا فإننا بمعرفة توزيع درجات الأفراد ، نعرف مدى صعوبة الاختبار بالنسبة لهم ، فإذا كان صعباً نسبياً انخفض المتوسط وحصل عدد كبير نسبياً من الأفراد على درجات تنحرف عنه انحرافاً سلبياً (أي تقل عن المتوسط) . فإذا أضفنا فقرات أسهل ارتفع المتوسط وزاد عدد من يحصلون على درجات مرتفعة . بل ويمكن علاوة على هذا معرفة مستوى صعوبة الفقرات التي أضيفت أو التي حذفت أو التي عدلت بمقارنة توزيع الدرجات قبل وبعد الحذف أو الإضافة أو التعديل .

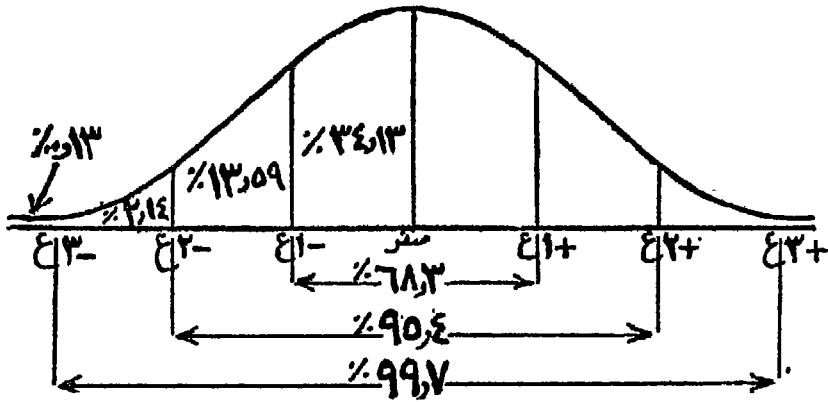
وعمليات إضافة وتعديل وتحليل واختصار فقرات الاختبار تجري عند تحليل فقراته item analysis وسنتحدث بالتفصيل عن هذه الطريقة الفنية فيما بعد .

الدرجات المحولة وأنواعها

في أول حديثنا عن الدرجة الخام وتفسيرها ذكرنا مثالا لو رجعنا إليه لوجدنا أن الدرجة التي يحصل عليها الفرد تفسر مبدئياً على أساس الدرجة النهائية للاختبار ومتوسط درجة المجموعة ، ثم عرفنا أن الانحراف المعياري وعدد المجموعة ومستوى صعوبة الاختبار تحدد هذا التفسير . وقلنا إن أبسط طريقة هي قسمة درجة الفرد على الدرجة النهائية للاختبار ثم حساب النسبة المئوية للأفراد الذين حصلوا على درجات أعلى منها . ويأتى

ترتيبه المئوية بعد هؤلاء الأفراد . وهذا يفترض نظرياً أن المجموعة يمكن تقسيمها إلى مائة وحدة كلا منها يسمى مئين Percentile . والرتبة المئينية للفرد هي موضعه النسبي في الجماعة الذي يحدد النسبة المئوية من الأفراد الذين حصلوا على درجات أعلى من درجته ، أو أقل منها .

وإذا أردنا أن نقارن درجة الفرد بدرجات المجموعة بفرض تفسير الأولى لزمنا أن نحولها إلى نوع من الدرجات المحولة ، والفكرة وراء هذه الدرجات المحولة هي أن تمثل المجموعة بمنحنى يوضح مراكز أعضاء هذه المجموعة . فالمنحنى يشمل بين خطه وقاعدته كل أفراد المجموعة (تقريباً) .



(شكل ١) - المنحنى الاعتدالي مقسم إلى وحدات هي الانحرافات المعيارية .

ويوضح الشكل النسب المئوية للدرجات الواقعة داخل حدود هذه الانحرافات

فلذا كان المنحنى اعتدالياً متماثلاً كان العمود الساقط من قمته على قاعدته يلتقي بها عند الدرجة المتوسطة ، وكان المماس للقمة يلتقي بالمحور الرأسى عند نقطة تمثل عدد من انحصرت درجاتهم في فئة المتوسط . ويقسم العمود المذكور المنحنى إلى قسمين متساويين .

وإذا فرضنا أن درجة الاختبار هي ٢٠٠ وأن المتوسط الحسابى (م)

(٤ - المقياس)

هو ١٠٠ وأن الانحراف المعياري (ع) هو ٢٥ ، وجدنا أن $م + ٣ ع$ أي الفرد الذي حصل على درجة تزيد عن المتوسط بمقدار ثلاثة أمثال الانحراف المعياري يحصل غالباً على أعلى درجة في المجموعة وهي $١٠٠ + ٣ \times ٢٥ = ١٧٥$ وهي تقع عند ذيل المنحنى من الناحية الموجبة بينما نجد أن $م - ٣ ع$ أي $١٠٠ - ٣ \times ٢٥ = ٢٥$ تقع عند ذيل المنحنى من الناحية السالبة . وغالباً ما تنحصر درجات هذه المجموعة بين $م + ٣ ع$ أي (١٧٥) و $م - ٣ ع$ أي (٢٥) ولكننا نجد أنه في حالة التمثيل الأمثل أي الاعتدالي ينحصر ٧٣ ، ٩٩ % بين $م + ع$ ، $م - ٣ ع$. أما إذا حصرنا عدد الأفراد الحاصلين على درجات تتراوح بين حد أعلى قدره $م + ٢ ع$ ($١٠٠ + ٢ \times ٢٥ = ١٥٠$) ، وحد أدنى $م - ٢ ع$ ($١٠٠ - ٢ \times ٢٥ = ٥٠$) فنجد نسبة هؤلاء الأفراد تعادل من مجموع الأفراد نسبة مئوية قدرها ٤٥ ، ٩٥ % وأخيراً فإننا لو حصرنا عدد الأفراد الحاصلين على $م + ع$ كحد أعلى ($١٠٠ + ٢٥ = ١٢٥$) وحد أدنى $م - ع$ ($١٠٠ - ٢٥ = ٧٥$) فنجد أنها تعادل من مجموع الأفراد نسبة مئوية قدرها ٢٧ ، ٦٨ % .

وهكذا يمكننا حساب النسبة المئوية لمن حصلوا على درجات تزيد عن المتوسط بنصف الانحراف المعياري أو تقل عنه بعشر الانحراف المعياري بالطريقة السابقة وهكذا . وهذه الطريقة نفيدنا في معرفة أين يقع الفرد بالنسبة لمجموعته أي ما معنى درجته في إطار درجات مجموعته . ونحن نلخص درجات مجموعته في معلمين هما $م$ ، $ع$ وتسمى الدرجة المحولة على هذا الأساس بالدرجة المعيارية ، وتحسب بإيجاد الفرق أي الانحراف بين درجة الفرد ومتوسط مجموعته مقسوماً على الانحراف المعياري للدرجات هذه المجموعة . وكأننا نعطي للدرجة معنى بمعرفة النسبة بين انحرافها عن المتوسط وبين انحراف درجات أفراد المجموعة عموماً عن هذا المتوسط ، فإذا رمزنا

إلى درجة الفرد بالرمز s والمتوسط الحسابي m والانحراف المعياري c فإن
الدرجة المعيارية تحسب بالطريقة التالية :

$$\frac{s - m}{c} = \text{الدرجة المعيارية}$$

وقيمة m ثابتة للاختبار الواحد وقيمة c أيضاً ثابتة ولكن القيمتين
تختلفان من اختبار إلى آخر وإن كانت أغلب الاختبارات تحاول الوصول
إلى $m = 100$ ، $c = 15$ وبعضها يعتمد على $m = 10$ و $c = 3$. بينما
يعتمد البعض على $m = 5$ ، $c = 2$ كما في الدرجة (ج) C. Score .

وبعض الاختبارات يستخرج متوسط درجات المجموعة التي جرب
عليها وانحرافها المعياري وتحول درجة كل فرد إلى درجة معيارية ، وتضرب
هذه الدرجة المعيارية في ١٠ ثم تجمع على ٥٠ والنتيجة هي الدرجة
النائية للفرد T score . وهنا تكون الدرجة النهائية ١٠٠ ، $m = 50$ ،
 $c = 10$. فإذا كانت درجة الفرد المعيارية هي ٢ تكون درجته النائية هي
 $70 = 50 + 10 \times 2$.

وهناك أنواع أخرى من الدرجات المحولة سنعرض لها تفصيلاً في فصل
عن المعايير .

الدلالة

إذا قارنا بين مجموعتين فإننا نقارن بينهما على أساس المتوسط والانحراف
المعياري ، وتأثير عملية المقارنة هذه بعدد أفراد المجموعتين : والدلالة
significance تعني الكشف عما إذا كان الفرق بين المجموعتين فرقاً جوهرياً
برغم تطبيق اختبار واحد عليهما وتوحيد الظروف في الحالتين . وكون الفرق
جوهرياً أو ذا دلالة يعني أن درجات أفراد المجموعتين من حيث تمركزها

وتشتها تختلف وأن الاختلاف راجع إلى طبيعة المجموعتين وليس راجعاً مثلاً إلى أخطاء في الاختبار أو في عمليات القياس . والفرق الجوهرى إذن لا يرجع إلى الصدفة ، ويعنى أنه لا يمكن اعتبار المجموعتين أجزاءً من مجموعة أكبر . وهناك عمليات مختلفة تجرى لقياس دلالة الفروق أهمها

$$\text{« اختبار ت » T. test ومعادلته هي : } t = \frac{\frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n-1}}{\sqrt{\frac{\sum x^2}{n} + \frac{(\sum x)^2}{n^2}}}$$

حيث $\bar{x}$ هي متوسط الدرجات لأفراد المجموعة الأولى ، $\bar{y}$ هي متوسط الدرجات للمجموعة الثانية ، $\sum x^2$ هي مربع الانحراف المياري لدرجات المجموعة الأولى ، أى التباين variance ، $\sum y^2$ هي مربع الانحراف المياري لدرجات المجموعة الثانية ، و n_1 عدد أفراد المجموعة الأولى ، n_2 عدد أفراد المجموعة الثانية .

وبالكشف في جداول « اختبار ت » يمكن معرفة ما إذا كان الفرق جوهرياً أم لا ، وما مدى جوهريه هذا الفرق :

ولكى يتضح لنا كل هذا لابد أن نعرض للعينات التى نأخذها سواء من المفحوصين أو الفقرات ثم نعود إلى الدلالة مستفيدين من عرضنا للعينات فى تفسير معنى الدلالة ومستوياتها .

العينات :

لا يستطيع المشتغل بالقياس أن يطبق الاختبار ، عند تقنيه على المجموعة الأصلية Population كلها ، أى لا يستطيع أن يطبقه على كل الأفراد . إذ أن عدة عقبات عملية تحول دون ذلك وخاصة إذا كان المجتمع الأصلى المراد قياسه كبير العدد . ولذا يقوم الباحث عادة باختيار عينة من الأفراد الممثلين لهذا المجتمع الأصلى . وهو فى اختياره هذا يحاول أن يجعل كل الصفات

والخصائص المميزة للمجتمع الأصلي متوفرة في العينة . وهو في هذا يحافظ على تمثيل العناصر المكونة للمجتمع الأصلي تمثيلاً تاماً من حيث عدد هذه العناصر . ونسبة كل عنصر منها في المجتمع الأصلي .

فإذا كان المجتمع الأصلي مدرسة عدد تلاميذها ٣٠٠ تلميذاً ؛ ولم تتوفر إمكانيات إجراء الاختبار عليهم جميعاً ، كان عليه أن يختار منهم عينة تمثلهم وليكن عدد أفرادها ١٠٠ . على أن أفراد هذه العينة يجب أن يمثلوا عدد الفصول والفرق الدراسية كأن يأخذ ٤٠ تلميذاً من السنة الأولى - ١٠ من كل فصل ، و ٣٠ تلميذاً من السنة الثانية - ١٠ من كل فصل ، و ٣٠ تلميذاً من السنة الثالثة - ١٠ من كل فصل . ذلك إن كان عدد فصول السنة الأولى ٤ وعدد فصول السنة الثانية ٣ ، وعدد فصول السنة الثالثة ٣ ، وكان عدد تلاميذ كل فصل منها هو ٣٠ تلميذاً ؛ على أن هذه العينة وعدد أفرادها ١٠٠ يجب أن تمثل أيضاً المستويات الاقتصادية - الاجتماعية ، فإن كان عدد التلاميذ من الطبقة الاقتصادية - الاجتماعية المرتفعة هو ٦٠ تلميذاً في المدرسة ، وجب أخذ ٢٠ منهم في العينة ما دامت العينة هنا $\frac{1}{3}$ المدرسة .

وبالمثل يراعى في اختيار العينة نسبة الطبقة المتوسطة والطبقة الدنيا ، كما تراعى العوامل الأخرى التي تؤثر فيما يقيسه الاختبار المطبق مثل السن ونسبة النجاح الدراسي و . . الخ :

وبالبحث في حساب العمليات الإحصائية والنفسية يعتمد على البيانات التي اشتقتها من العينة لكي يصل إلى تعميمات تنطبق على المجتمع الأصلي ؛ وعلى هذا العينة الجيدة يجب أن تتمثل فيها جميع صفات الأصل التي اشتقت منه حتى يصبح التعميم صادقاً . وعلى أي حال فيجب ألا يقل حجم العينة عن ٣٠ ٩

ولاختيار العينة طرق عدة أهمها :

أ - الطريقة العشوائية Random : وفيها تكون الفرض متكافئة لجميع أفراد المجتمع الأصلي كي يدخلوا ضمن العينة : وهنا قد نرجع إلى كشوف التلاميذ أو العمال أو المرضى ونختار بحسب تسلسل الأرقام فإذا كانت العينة $\frac{1}{10}$ من المجتمع الأصلي اخترنا الأرقام ١ ، ١١ ، ٢١ . الخ .

ب - الطريقة الطبقيّة Stratified : وقد تسمى الطريقة الحوضيّة وتتلخص الفكرة فيها في النقاط التالية :

١ - نصنف المجتمع الأصلي إلى فئات تبعاً للعوامل التي قد تؤثر فيها نقيسه :

٢ - نختار العينة العشوائية من كل فئة بحيث تمثل فئات المجتمع الأصلي كل حسب عدد أفرادها .

ج - الطريقة المضبوطة Controlled : ومثال المدرسة الذي ضربناها يوضح الفكرة والخطوات المتبعة فيها .

د - الطريقة العرضية Accidental : وفيها يقوم الباحث باختيار أقرب الأفراد إلى إمكانياته كأن يذهب إلى مدرسة قريبة منه ويطبق الاختبار على بعض التلاميذ في بعض الفصول دون تمثيل أو تقييد .

وهكذا نرى أن على الباحث أن يثبت صحة اختياره لعينته وذلك باختيار عينات أخرى ومقارنة النتائج .

عود إلى الدلالة

رأينا أن درجة تمثيل العينة للمجتمع الأصلي تؤثر على صحة النتائج التي نستخرجها من العينة لنعممها على الأصل . وعموماً يزيد صدق نتائجنا بزيادة عدد أفراد العينة : وتهدف الدلالة الإحصائية إلى الكشف عن مدى هذا

الصدق . ويزداد صدق التعميم كلما قربت النتائج في العينة من النتائج في الأصل أو في عينة أخرى مماثلة من نفس الأصل . ويقاس هذا القرب بإيجاد الانحراف المعياري للنتائج المستخرجة من العينة . ويسمى الانحراف المعياري لكل معلم من هذه المعالم بالخطأ المعياري . فمثلا الخطأ المعياري للمتوسط هو خارج قسمة المتوسط على الانحراف المعياري للعينة . ويتراوح المتوسط الحسابي الحقيقي ، أى مجموع درجات المجتمع الأصلي كله مقسوماً على عدد أفرادهِ ، يتراوح بين المتوسط الذى استخرجناه من العينة مضافاً إليه مقدار الخطأ المعياري هذا كحد أعلى ومطروحاً منه كحد أدنى وغالباً لا يزيد المتوسط الحسابي للمجتمع الأصلي كله عن (مجموع متوسط العينة + الخطأ المعياري لهذا المتوسط) ولا يقل عن (متوسط العينة مطروحاً منه الخطأ المعياري لهذا المتوسط) .

ولكل معلم إحصائي معادلة في حساب دلالة .

معامل الارتباط

معامل الارتباط Coefficient of Correlation هو قيمة رياضية (عددية) تحسب بإحدى طرق عدة : وهذه القيمة العددية تبين درجة العلاقة أو التمشي بين متغيرين ، والمتغير هنا هو أى صفة أو ظاهرة نفسية تختلف من فرد لآخر أو من ظرف لآخر ويمكن قياسها أو تقديرها كميّاً ، فالطول متغير والوزن متغير آخر والدرجات في اختبار للحساب أو للذكاء متغير . . . الخ .

وتراوح القيمة العددية لمعامل الارتباط بين + ١ ، - ١ وفيما يلي وصف لما تدل عليه بعض هذه القيم :

١ - معامل ارتباط قيمته (+ ١) : يدل هذا على علاقة تامة موجبة ، أى أن الزيادة في أحد المتغيرين يصحبها دائماً زيادة نسبية في التغير الثاني ،

وكذلك النقص في أحدهما يصحبه دائماً نقص نسبي في الثاني ، (وبعبارة أخرى يكون ترتيب الأفراد في كلا المتغيرين ثابتاً) .

من أمثلة ذلك العلاقة بين الحجم والوزن لأجسام من نفس المادة ونفس الكثافة . أو العلاقة بين درجة الحرارة وضغط الغاز إذا كان الحجم ثابتاً . انظر الحالة ١ في شكل (٢) .

٢- معامل ارتباط قيمته (- ١) : يدل هذا على علاقة تامة سالبة أو عكسية ، أى أن الزيادة في أحد المتغيرين يصحبها دائماً نقص نسبي في المتغير الثاني . أو العكس النقص في أحد المتغيرين ، يصحبه دائماً زيادة نسبية في المتغير الثاني ، (وبعبارة أخرى يكون ترتيب الأفراد في أحد المتغيرين عكسه في المتغير الثاني تماماً) ، مثل العلاقة بين حجم الغاز وضغطه إذا كانت درجة الحرارة ثابتة ، أو العلاقة بين الحجم والكثافة في الأجسام المتساوية الوزن . (انظر الحالة هـ في شكل ٢) .

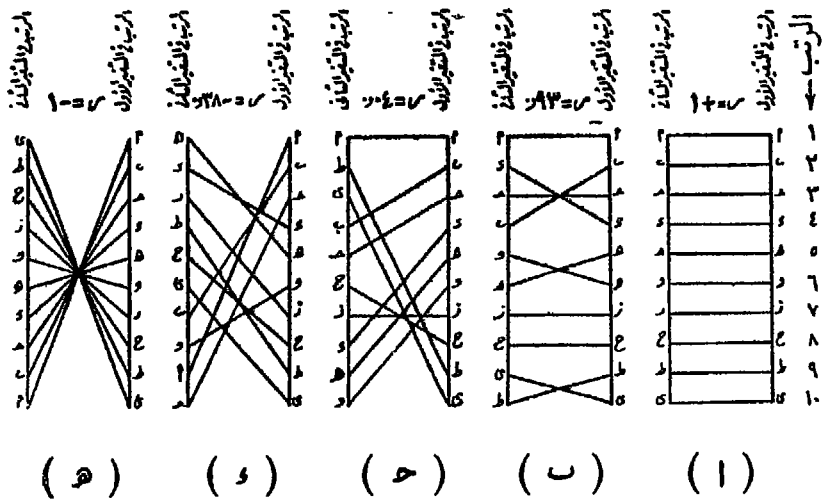
٣- معامل ارتباط بين + ١ ، صفر أى أكبر من الصفر وأقل من + ١ صحيح . يدل هذا على علاقة جزئية موجبة أى أن الزيادة في أحد المتغيرين كثيراً أو غالباً يصحبها زيادة نسبية في المتغير الثاني ، وكذلك النقص في المتغير الأول كثيراً أو غالباً ما يتبعه نقص نسبي في المتغير الثاني . من أمثلة ذلك العلاقة بين أطوال الأفراد وأوزانهم ، أو العلاقة بين أطوال الأفراد وأعمارهم فيما تحت سن ٢٠ سنة مثلاً . وكلما اقتربت قيمة معامل الارتباط من + ١ يوصف بأنه معامل ارتباط عالى موجب ومعنى ذلك أن المتغيرين يتمشيان مع بعضهما بدرجة كبيرة ، وبالعكس كلما اقتربت من الصفر دل ذلك على أن العلاقة ضعيفة . (انظر الحالة ب في شكل ٢) .

٤- معامل ارتباط يقع بين - ١ ، صفر أى أصغر من الصفر وأكبر من - ١ : يدل هذا على علاقة جزئية سالبة أو عكسية ، أى أن الزيادة في أحد

المتغيرين كثيراً أو غالباً ما يصحبهما نقص نسبي في المتغير الثاني وبالعكس النقص في المتغير الأول كثيراً أو غالباً ما يتبعه زيادة في المتغير الثاني ومن أمثلة ذلك العلاقة بين السن وسرعة الجرى . (انظر الحالة د في شكل ٢) :

٥ - معامل ارتباط يساوى صفر (أو صفر تقريباً) يدل هذا على عدم وجود علاقة أو ارتباط بين المتغيرين مثل العلاقة بين ذكاء الراشدين وأطوالهم . (انظر الحالة ح في شكل ٢) .

ويوضح الشكل (٢) رسوماً تخطيطية لبعض معاملات الارتباط محسوبة بطريقة سبيرمان لحساب معامل الارتباط بالترتيب Rank-Order Correlation ، وهى طريقة تعتمد على ترتيب درجات الأفراد في كل من المتغيرين وحساب الفروق بين رتبتي كل فرد من هؤلاء الأفراد .



شكل (٢)

تطبيقات على معامل الارتباط :

١ - معامل الثبات : هو معامل الارتباط بين نتائج الاختبار الواحد ،

ذلك بإيجاد معامل الارتباط بين مرات الإجراءات المختلفة لنفس الاختبار .
وقد تجرى صورة منه مرة ، ويعاد الإجراء بصورة أخرى مكافئة من نفس
الاختبار . وقد يستخرج معامل الارتباط بين نصفى الاختبار الواحد .
وهكذا فمعامل الثبات هو ارتباط الاختبار بنفسه : وسنناقش موضوع الثبات
في فصل مستقل آخر :

٢ - معامل الصدق : هو معامل الارتباط بين نتائج الاختبار وبين
ما يقيسه هذا الاختبار ، أى أن يوجد معامل الارتباط بين نتائج الاختبار
وبين نتائج اختبار آخر سبق أن ثبت صدقه ، والآخر يسمى المحك
Criterion . وقد يكون المحك هو مستوى النجاح فى المهنة ، أو تقديرات
الخبراء ، أو النجاح المدرسى : الخ . وسنناقش الصدق بالتفصيل فى
فصل تال :

التحليل العاملى

التحليل العاملى Factorial Analysis هو منهج إحصائى للبحث عن
العوامل التى تؤثر فى الظواهر المعقدة التى ترتبط ببعضها . ويهدف التحليل
العاملى إلى تفسير الظواهر الكثيرة التى يحللها إلى عدد قليل من العوامل .
أى أنه يهدف إلى التبسيط العلمى .

ويعتمد التحليل العاملى على المعالجة الإحصائية لمعاملات الارتباط بين
المتغيرات ، أو الاختبارات التى تقيس هذه المتغيرات ، فترصد معاملات
الارتباط فى جدول متناسق الخلايا (الخانات) يسمى مصفوفة الارتباط
Correlation Matrix . وفى مصفوفة الارتباط تكون خلايا العمود الرأسى
الأول ، وخلايا الصف الأفقى الأول هى أرقام الاختبارات ، أى الدليل ،
وترصد معاملات الارتباط فى الخلايا الداخلية وفقاً لأرقام الاختبارات فى
الدليلين الرأسى والأفقى ، وترك الخلايا القطرية خالية :

وبطريقة رياضية أو أخرى تشتق العوامل التي تلخص هذه الارتباطات أو بعضها ، ويفترض أن هذه العوامل تتعلق بسمات معينة :

ويساعد التحليل العاملي في تنقية الاختبارات بنية الوصول إلى اختبارات نقية Pure tests يقيس كل منها صفة معينة ، كما يساعد في حساب صدق الاختبارات ، وهو ما يعرف بالصدق العاملي Factorial validity ، وسنتناول تطبيقات طريقة التحليل العاملي في فصل تال :

تحليل التباين

التباين Variance هو مربع الانحراف المعياري أى أنه خارج قسمة مجموع مربعات الانحراف عن المتوسط ÷ عدد الأفراد : ونلجأ إلى تحليل التباين لمعرفة ما إذا كانت النتائج الجزئية ، التي حصلنا عليها من مصادر مختلفة متفرقة ، يمكن النظر إليها على أنها نتيجة واحدة من مصدر واحد أم لا . وأهم أداة في هذا التحليل هي اختبار « ت » t Test ذلك لمعرفة ما إذا كان الفرق بين متوسطي مجموعتين فرقا جوهريا لا يمكن أن يكون قد حدث بالصدفة أى لا يمكن اعتبار العينتين من مجتمع أصلي واحد . وبالرغم من أن التباين هو مربع الانحراف المعياري إلا أن استخدام التباين أعم فضلا عن أنه يخضع لعمليات إحصائية أخرى كثيرة : وهناك طرق أخرى لتحليل التباين لا يسمح المجال بذكرها .

القياس والتقييم

القياس Measurement والتقييم Evaluation اصطلاحان شائعان في ميادين علم النفس والتربية ويظن البعض أنهما مترادفان وهو ظن خاطئ إذ أن ثمة فروقا جوهرية تقوم بينهما . وأول هذه الفروق أننا نقصد بالقياس كل الملاحظات التي يمكن التعبير عنها بصورة كمية ، فنتيجة القياس مرة

نسبة الذكاء وأخرى درجة القدرة وهكذا : أى أن القياس يجب عن السؤال : ما مقدار كذا ؟ أما التقييم فهو أعم إذ أنه لا يقتصر على التحديد الكمي بل يذهب إلى أبعد من ذلك ، يذهب إلى الحكم على قيمة درجة الصفة المقاسة . فيقيم هذه الدرجة بأنها جيدة وتلك بأنها ضعيفة وهو في هذا يهدف إلى معرفة ما إذا كانت درجة معينة من الجوانب المقاس تنبئ بالنجاح أو تكشف عن حاجة إلى العلاج مثلا . أى أن التقييم يضع أهدافاً محددة يحكم على درجات الجوانب النفسية والتربوية من خلالها ، فإذا اقتربت درجة الجوانب من الهدف المحدد تنبأ بالنجاح :

وكان التقييم ، برغم اتصاله بالقياس واستخدامه له ، هو أعم منه لأنه يتضمن التقديرات والأحكام غير الرسمية ، الدالة مثلا على نجاح التلميذ ، وهذه الأحكام تتضمن تقدير الدرجات لمعرفة ما إذا كانت تؤهل لما يتمشى مع أهداف نرجوها : أم لا . وعلى أى حال فإن كل قياس ناجح يجب أن يسهم في تقدم عمليات التقييم :

قائمة المصطلحات

والمفاهيم الأساسية في القياس النفسي والتربوي

اتجاه Attitude : مصطلح متعلق بأحد جوانب الشخصية يصف الميل العام العاطفي المكتسب الذي يؤثر في الدوافع النوعية وفي سلوك الفرد وإليه يرجع السلوك المستمر المتسق نحو أو بعيداً عن مجموعة متقاربة من المواقف أو الأشياء :

أثر التدريب Practice effect : تأثير الخبرة السابقة المكتسبة من اختبار على الإجراء التالي لنفس الاختبار أو اختبار شبيه به . واصطلاحاً أثر التدريب يستخدم عندما يتدخل أو يستبعد أثر تدريب من موقف آخر غير الموقف المقاس .

اختبار الأداء Performance Test : اختبار يتطلب عادة استجابة يدوية أو استجابة حركية عموماً يقوم بها الفرد ، في مقابل الاستجابات التي يتطلبها اختبار من نوع الورقة – والقلم Paper-and-Pencil test :

اختبار استعداد Readiness : اختبار يقيس المدى الذي حصل به الفرد درجة من النضج maturity أو اكتسب به مهارات معينة أو معلومات يتطلبها البدء في نوع من التعلم الجديد . وهذا الاصطلاح غالباً ما يستخدم في حالة أطفال ما قبل المدرسة لتحديد استعدادهم readiness لدخول المدرسة .

اختبار التحصيل Achievement test : مقياس للدرجة التي بها حصل الشخص أهداف التعليم أو التدريب .

الاختبار التشخيصي Diagnostic test : اختبار يستخدم في التشخيص ،

أو في تبين نواحي القوة والضعف في مجال معين من الدراسة . وهو يتضمن مقاييس لمركبات ، أو أجزاء متفرعة ، من جهاز أكبر من المعلومات أو المهارات .

الاختبار التنبؤي Prognostic test : اختبار يستخدم في التنبؤ بالتحصيل المستقبلي في موضوع أو مجال معين .

اختبار التصفية Screening test : اختبار يستخدم بغرض تصفية المتقدمين للالتحاق بعمل أو دراسة تمهيداً للقيام بعمليات أخرى لاختيارهم .

اختبار التصنيف العام General Classification test : اختبار يجري بغرض تقسيم من يطبق عليهم إلى مجموعات تصلح كل مجموعة لمهنة أو دراسة معينة : والاصطلاح استخدم أساساً في اختبار حربي صمم في الحرب العالمية الثانية لتصنيف الجنود .

اختبار التمكن Mastery test اختبار غرضه الأساسي تحديد المدى الذي به تعلم أو تمكن الأفراد في الجماعات من وحدة معينة من التدريب . وهذا النوع من الاختبارات لا يقصد منه التفريق بين الأفراد على وجه الدقة ، ولكن لتحديد ما إذا كانت أو لم تكن مجموعة من الأفراد قد حصلت مستوى معيناً من الكفاية Proficiency ، وهو يستخدم لمعرفة ما إذا كانت أو لم تكن مجموعة من الأفراد مستعدة للتقدم في وحدة تالية من التدريب .

اختبار جمعي Group test : اختبار يستطيع الفاحص أن يطبقه على عدد من الأفراد مجتمعين في نفس الوقت .

اختبار حافل Omnibus test (وقد يسمى اختبار دشت) : اختبار يقيس تراته عديداً من العمليات العقلية المتنوعة ، و فقراته مرتبة معاً وتعطى عنها جميعاً درجة واحدة .

اختبار ذاتي Subjective test : اختبار يصحح على أساس اتجاهات ،
وتقدير ، وآراء ، وأفكار المصحح :

اختبار سرعة Speed test : اختبار له زمن محدد وفيه تعتبر سرعة
الأداء أو الاستجابة أهم متغير :

اختبار فردي Individual test : اختبار يطبق على فرد واحد فقط
في المرة الواحدة .

اختبار قدرة Ability test : اختبار يتطلب أقصى أداء ليعين مستوى
المركز الراهن أو القدرة الراهنة في الوظيفة :

اختبار قوة Power test : اختبار يقصد به قياس مستوى الأداء مع
تقليل أثر سرعة الاستجابة ؛ وعلى ذلك ، فهو اختبار إما لا تحديد للوقت
فيه وإما أن يكون رفته متسعاً .

اختبار عبر حضارى Cross-Cultural test (وقد يسمى اختبار خالص
من الثقافة Culture - free test) : وهو اختبار مصمم لاستبعاد آثار بيئة
الفرد الماضية على درجته . وليس في الإمكان عمل مثل هذا الاختبار فعلاً
وإن بذلت محاولات من هذا النوع ، فالاختبار الخالص من الثقافة لا يستبعد
هذه الآثار ولكنه فقط يجعلها متكافئة لدى كل الأشخاص الجارى مقارنتهم :

اختبار غير لفظي Nonverbal test : هو اختبار من نوع الورقة —
والقلم يستخدم مع الأطفال في الفرق الأولى ، حيث تكون فقرات الاختبار
رموزاً ، وأرقاماً وصوراً لا كلمات ، وتعطى التعليمات شفويّاً .

اختبار لفظي Verbal test : اختبار تعتمد نتائجه بدرجة ما على استخدام
وفهم الكلمات ، كما في أغلب اختبارات الورقة — والقلم .

اختبار المسح Survey test : اختبار يقيس التحصيل العام في موضوع

أو مجال معين . وهو عادة يعنى أكثر باتساع المدى ، منه بالتفاصيل النوعية أو اكتشاف العوامل المسببة . وهو غالباً ما يستخدم فى تصفية مجموعات كبيرة من الأشخاص .

اختبار مقنن Standardized test (وقد يسمى Objective test) :
اختبار أعطى من قبل لعديد من العينات Samples أو المجموعات تحت ظروف مقننة واشتقت له معايير ؛

اختبار موضوعى Objective test : اختبار يصصح بطريقة لا تتأثر
بشخص من يقوم بتصحيحه ، فالنتائج واحدة برغم اختلاف المصححين .
وتستخدم فى تصحيحه مفاتيح التصحيح أو المفاتيح المثقبة Stencils .

اختبار إجابة من عدة اختيارات Multiple-Choice item : فقرة
الاختبار التى يكون واجب الفرد فيها أن يختار الإجابة الصحيحة أو أحسن
إجابة من عدد من الاختيارات المقدمة أمامه .

اختبار استجابة من عدة استجابات Multiple-response item : نوع
خاص من الفقرات التى على المفحوص فيها أن يختار استجابة واحدة أو أكثر
من الاختيارات المعروضة أمامه ، وهى تشمل استجابتين صحيحتين
أو أكثر .

الاختيار المهني Vocational Selection : الاختيار المهني عمليات تهدف
إلى اختيار الفرد الأكثر ملاءمة ، من بين عدة أفراد متقدمين لشغل الوظيفة
المعينة ، على أن ينتج أحسن إنتاج ويكون أكثر رضا :

التوجيه المهني Vocational guidance : عمليات تهدف إلى مساعدة
الفرد على أن يختار بنفسه المهنة الأكثر ملاءمة له ، من بين عدة مهن قد يمكنه
الالتحاق بها ، وأن يلتحق بها ويتكيف معها ويتقدم فيها .

التوجيه التربوي Educational guidance : عمليات مساعدة التلميذ على

أن يختار بنفسه الدراسة الأكثر ملاءمة له ، من بين عدة دراسات قد يمكنه الالتحاق بها ، وأن يلتحق بها ويتكيف معها ويتقدم فيها :

الأداء المعتاد Habitual Performance : هو أداء الاختبار بحيث تمثل استجابة المفحوص سلوكه فعلاً في مواقف مشابهة لأن يعبر الفرد عما كان يجب أن يفعل ، في مثل هذه المواقف التي أتى بها الاختبار . وهذا النوع من الأداء نطلبه في اختبارات الشخصية :

العامل Factor : تكوين افراضى مستنتج من إجراء عمليات التحليل العاىلى لعلاقات الترابط بين عدد من المتغيرات أو الاختبارات المتعلقة بأحد الوظائف . وهو مفهوم لوصف أحد مجالات السلوك بعد تصميم اختبار يقيس عاملاً بذاته أو مجموعة من العوامل المتضمنة في إحدى القدرات العامة للذكاء .

القدرة Ability : هى المستوى الراهن من الوظيفة سواء كان متأثراً بالتدريب أم لم يكن . والقدرة تصنيف لأجهزة من الاستجابات ترتبط فيما بينها ارتباطاً كبيراً متميزة بذلك عن غيرها من الأجهزة .

الاستعداد Aptitude : احتمال الوصول إلى درجة معينة من التمكن بعد الاستفادة من التدريب . والتدريب هنا لا يقتصر على التعليم في المدارس والتدريب في المصانع وإنما يشمل أيضاً الخبرات التي مر بها الفرد التي تتعلق بجانب معين من السلوك .

المهارة Skill : مستوى القيام بالعمليات الحسية الحركية المعقدة قياماً تتكامل فيه عناصر الدقة والسرعة والتكيف مع الظروف المتغيرة .

الكفاية Proficiency : درجة النجاح في القيام بالوظائف والمقتنيات التي تتطلبها مهنة معينة :

السعة Capacity (وقد تسمى Capability) : أفضل مستوى يحتمل أن يصل إليه الفرد إذا حصل على أنسب تدريب أو تعليم .

استجابة حرة Free-response : فقرات في فئة من الاختبارات تطلب من الفرد أن يستجيب بكلماته الخاصة لا أن يختار من بين عدة متغيرات .

استعداد مدرسي Scholastic Aptitude (وقد يسمى Academic Aptitude) : وهو الخصائص الشخصية ، الفطرية والمكتسبة ، التي تمكن من الوصول إلى درجة معينة من النجاح الأكاديمي .

الالتواء Skewness : ميل التوزيع إلى البعد عن التماثل أو الاتزان حول المتوسط .

أقصى أداء Maximum Performance : هو أداء الاختبار يبذل أقصى سرعة ودقة وقدرة لدى الفرد في الإجابة عنه . وفي اختبار أقصى الأداء يقارن بين الأفراد على أساس أن كل فرد بذل كل ما يستطيعه إذا توفرت الظروف لذلك . وهذا الأداء مطلوب في كل الاختبارات ما عدا اختبارات الشخصية .

انحراف Deviation : الكمية التي بها تختلف الدرجة عن قيمة مرجعية معينة مثل المتوسط أو المعيار أو الدرجة على اختبار آخر .

الانحراف الربيعي Quartile Deviation (Q) (وقد يسمى نصف المدى الربيعي) : مقياس للتشتت يعرف على أنه نصف المسافة (بالدرجات) التي تقع بين الربع الأعلى Q_3 والربع الأدنى Q_1 .

الانحراف المعياري Standard Deviation (SD) : مقياس لتباين أو تشتت مجموعة الدرجات ، فكلما قربت الدرجات من المتوسط قل الانحراف المعياري . وفي التوزيع العادي (normal) تقع حوالي ٦٨٪ من

الدرجات في مدى ١ انحراف معياري فوق وأقل من المتوسط ، وحوالي ٩٥٪ في مدى ٢ انحراف معياري . ويرمز إليه عادة بالرمز ع :

بطارية Battery : مجموعة من الاختبارات المقتنة على نفس الأشخاص ، ومعاييرها مشتقة بطريقة تسمح بالمقارنة : وقد يقصد بالبطارية أحياناً اختبارين أو أكثر أعطيت لنفس الأشخاص ، سواء قننت معاً أم لم تقنن :

بطارية الاستعدادات الفارقة Differential Aptitude Battery : هي مجموعة من الاختبارات قننت معاً على نفس الأشخاص بطريقة تمكن من المقارنة ، وهي تقيس استعدادات يفترق فيها الأفراد فروقاً يمكن استغلالها في اختيارهم لمهنة أو دراسة .

التباين Variance : (انظر الانحراف المعياري) هو متوسط مربعات الانحرافات عن المتوسط الحقيقي أو الفرضي .

تحليل الوحدات Iten analysis : عملية تقييم فقرات الاختبار الواحد باستخدام أى من الطرق المختلفة . وتتضمن هذه العملية عادة تحديد قيمة الصعوبة difficulty value والقوة التفرقية discrimination power للفقرة وعلاقتها ببعض المحكات :

تحليل التباين Analysis of Variance : يقيس التباين الفروق الفردية والجماعية بحساب مدى انحراف كل فرد عن متوسط الأفراد ، أو مدى انحراف جماعة عن متوسط الجماعات ، أو مدى انحراف كل عينة عن الأصل الذي تنتمي إليه .

التحليل العامل Factor Analysis : طريقة في الإحصاء لتحديد العوامل من بين عدة ارتباطات داخلية بين مجموعة من المتغيرات ، هذه المتغيرات عادة تكون اختبارات

تصحيح أثر التخمين Correction for Guessing : تقليل الدرجة

بسبب إجابات خاطئة كى تتكيف الدرجات على الاختبارات الموضوعية بحيث تلغى آثار التخمين الذى قد يقوم به الفرد فيصل إلى إجابات صحيحة . وتعتمد هذه العملية على افتراض خاطئ هو أن التخمين هو الصدفة : وقد يكون من المفيد فى تقييم الدرجة أن نحدد ما إذا كانت أعلى من المستوى الذى قد توصل إليه الصدفة وحدها ، كما فى اختبار الدقة فى التنبؤ باستجابات الآخرين .

التكرار المتجمع Cumulative Frequency : هو طريقة لحساب

عدد الأفراد أو الدرجات التى تزيد أو تقل عن حد معين بحيث يمكن من مجرد النظر فى عمود التكرار المتجمع ، أمام فئة معينة ، أن نعرف عدد من يقعون قبل هذه الفئة .

التقنين Standardization : يتضمن التقنين تحديد شروط تطبيق

الاختبار تحديداً دقيقاً تبعاً لمبدأ مراعاة ضبط جميع العوامل التى تؤثر فى الظاهرة التى تبحث . فتوضع تعليمات الإجراء والتصحيح وتذكر المعايير ويحسن تحديد ظروف الإجراء الأخرى .

التوزيع الاعتنالى Normal Distribution : توزيع الدرجات أو المقاييس

بطريقة بيانية لها شكل جرسى Bell shaped : مميز وفيها تماثل حول المتوسط ، فالحالات متركزة قرب المتوسط وتتناقص فى تكرارها كلما بعدت عن المتوسط ، متمشية مع معادلة رياضية محددة .

التوزيع التكرارى Frequency Distribution : تبويب الدرجات من

الأعلى إلى الأقل أو العكس ، لتبين عدد الأفراد الذين حصلوا على كل درجة أو الذين يقعون فى كل فئة من الدرجات .

توحيد الظروف Uniformity of Conditions : وضع كل الأفراد فى

ظروف موحدة لاستبعاد أثر العوامل المتدخلة التي نرغب قياس آثارها .
وتوحيد الظروف في عمليات القياس يشمل ظروف المكان والإضاءة والتهوية
والحرارة ... الخ .

الثبات Reliability : مدى اتساق الاختبار مع نفسه في قياس أى
جانب يقيسه . أى مدى استقرار الدرجات التي يحصل عليها نفس الأفراد
في مرات الإجراء سواء أعيد الإجراء بنفس الصورة أو بصورة مكافئة من
نفس الاختبار . أو كنا بصدد اتفاق درجتى الأفراد على نصفى الاختبار .

جدول الانتشار Scatter Diagram (وقد يسمى Scatter gram) :
جدول ثنائى يستخدم في توزيع الدرجات على اختبارين أو متغيرين معاً
بالنسبة لكل شخص في المجموعة . وهو يقدم دليلاً بيانياً على العلاقة
بين المتغيرين .

جدول الأعمار للفرق الدراسية Age-grade table : جدول يبين عدد
الأفراد أو نسبهم المئوية في كل سن داخل كل فرقة مدرسية :

اتجاه الاستجابة Response-set : ميل الفرد إلى أن يتبع نمطاً معيناً
في الاستجابة . مثل الاتجاه إلى الإجابة بـ « لا أعرف » في اختبارات
الشخصية .

الخطأ المحتمل Probable Error : دليل على تباين المقياس ، أى
المدى الذى به تنحرف قيمة معينة حصل عليها الفرد عن مستواه ، وهو
مقياس للخطأ في اختيار العينة . وفي التوزيع الاعتدالى تقع نصف الانحرافات
عن مقياس النزعة المركزية في مدى خطأ محتمل واحد فوق وتحت ذلك
المقياس : والخطأ المحتمل هو ٠.٦٧٤٥ من الخطأ المعياري .

الخطأ المعياري Standard Error : تقدير لكمية خطأ قياس الدرجة
أى الكمية التي بها تختلف درجة حصل عليها الفرد عن درجته الحقيقية التي

نفترضها . والخطأ المعياري مأخوذ بحيث أنه في حوالى ثلثي الحالات لا تختلف الدرجة التي حصل عليها الفرد عن درجته الحقيقية بأكثر من خطأ معياري واحد :

الدرجة التائية T. Score ، هي درجة معيارية متوسطها ٥٠ وانحرافها المعياري ١٠ تستخدم عادة لتحويل الدرجات الخام على اختبارين أو أكثر إلى درجات يمكن مقارنتها ، ذلك بغرض تسهيل التفسير .

الدرجة الجيمية C. Score : نوع من الدرجات المعيارية تستخدمه فئة من الاختبارات متوسطه الحسابي ٥ وانحرافه المعياري ٢ :

الدرجة الخام Raw Score : هي النتيجة الأصلية ، التي لم تعالج إحصائياً ، التي حصلنا عليها من الاختبار أو من أى أداة قياس أخرى : وهي عادة عدد الإجابات الصحيحة ، أو النقاط التي أجاب الفرد عنها بصحة في مقياس النقط .

الدرجة الحقيقية True Score : الدرجة التي قد نحصل عليها إذا استخدمنا أداة قياس ثابتة تماماً . فإذا أمكن قياس الفرد عدة مرات بنفس الاختبار ، دون أن يحدث تغير في الفرد ، قدرت درجته الحقيقية بمتوسط كل درجاته على الاختبار . ونحن دائماً لانصل إلى درجات حقيقية بل إلى درجات نعتبرها افتراضية :

الدرجة المئينية Percentile Score : الدرجة التي تمثل النسبة المئوية من الأفراد الذين يقعون تحت درجة خام معينة .

الدرجة المعيارية Standard Score : درجة يعبر فيها عن درجة كل فرد على أساس عدد وحدات الانحراف المعياري لدرجته عن المتوسط :

الآلة Significance : أى مدى جوهرية النتيجة التي نحصل عليها :

وتستخدم الدلالة عادة لتيين مدى أصالة فرض وضعناه أو مدى جوهرية فرق بين متوسطين :

الربيع Quartile : أحد ثلاث نقط تقسم توزيعاً متلجراً إلى أربعة أجزاء متساوية : والربيع الأدنى Q_1 أو المئين الخامس والعشرين ، والربيع الأعلى Q_3 أو المئين الخامس والسبعين ، والربيع المتوسط Q_2 أو المئين الخمسين وهو الوسيط Median :

العشير Decile : أحد النقط التسع التي تقسم توزيعاً مرتباً من الدرجات إلى عشر أجزاء متساوية :

السقف المنخفض Low Ceiling : اصطلاح يستخدم في حالة الاختبار الذي يكون سهلاً جداً لكثير من الأفراد ، بحيث يحصل الكثيرون على الدرجة النهائية أو أقصى درجة يسمح بها المقياس ، والعكس في حالة السقف المرتفع (High Ceiling) إذ لا يحصل أحد على الدرجة النهائية أو أقصى درجة يسمح بها المقياس :

السمات traits : الصفة التي يمكن أن نفرق على أساسها بين فرد وآخر ويعتمد مفهوم السمة على نظرية تقول إن لكل فرد سمات شخصية ثابتة يمكن أن تلاحظ فيه . والسمات بهذا المعنى ثابتة وعامة :

السوسيوجرام Sociogram : (وقد يسمى رسم أو تخطيط العلاقات الاجتماعية) : وهو خريطة أو رسم تخطيطي يوضح التفاعلات باستخدام الأسهم في تمثيل العلاقات والاختيارات المتبادلة بين الأفراد . وهو عادة يبين الأفراد المرغوب فيهم وغير المرغوب فيهم من أعضاء الجماعة :

السوسيومتري Sociometry (وقد يسمى مقياس العلاقات الاجتماعية) : وهو طريقة تحديد العلاقات الاجتماعية أو الشخصية بين أعضاء الجماعة .

الشخصية Personality : يقصد بهذا الاصطلاح المفهوم الذى به نصف التنظيم الدينامى الذى يتكون من تنظيمات مختلفة متكاملة تضم أجهزة السلوك والإدراك التى تعتمد على التفاعل النفسى الجسمى . على أن هذا التنظيم يحدد الطريقة الفريدة التى يتميز بها الشخص عن غيره فى استجاباته للبيئة ، عالم الأشخاص ، وعالم الأشياء ، ونفسه ، مع التأكيد على مدى التوافق فى المجال الاجتماعى .

الصدق Validity : هو أن يقيس الاختبار ما افترض أنه يقيسه . ويعرف الصدق على أساس أغراض مختلفة أى أن أنواعاً مختلفة من الأدلة تستخدم فى تحديد أنواع الصدق وأكثر أنواع الصدق شيوعاً هى : صدق المحتوى Content validity ، التى تصف مدى جودة تمثيل محتوى الاختبار لفئة من المواقف أو الموضوعات التى نقيسها .

الصدق التلازمى Concurrent validity : الذى يصف مدى ارتباط الدرجة على الاختبار بمحركات الأداء الراهنة أو مركز الفرد حالياً ، والصدق التنبؤى Predictive validity : ويدل على مدى مسابقة التنبؤات التى نصل إليها من الاختبار مع دليل نجمعه فى وقت تال على إجراء الاختبار ، وصدق البناء construct validity الذى يدل على الدرجة التى بها تعد تكوينات تفسيرية أو مفاهيم معينة ، تعد مسؤولة عن الأداء فى الاختبار .

الصدق الظاهرى Face valieity (قد يسمى الصدق السطحى) : الصدق الظاهر للاختبار هو كيف يبدو الاختبار مناسباً وملائماً للفرد الذى نقيسه . والمدى الذى به يتكون الاختبار من فقرات تبدو مرتبطة بالمتغير الذى يقاس .

النسبة الشخصية Profile : رسم يبانى لنتائج أداء الفرد لمجموعة من الاختبارات أو الاختبارات الفرعية .

صلب السؤال Stem : الجزء المقدم من سؤال يختار فيه الفرد أحداً من عدة اختيارات ، وهذا الجزء على الفرد أن يكمله باختيار أحد الاستجابات التي تتبع السؤال .

الصورة المصغرة للعمل Job Miniature : طريقة من طرق صياغة فقرات الاختبار بحيث يكون التشابه تام بين عينة السلوك الذي يقيسه الاختبار وبين السلوك المراد التنبؤ به .

الصورة المكافئة Equivalent Form (وقد تسمى الصورة الموازية Parallel form أو الصورة المساوية Equal form) : أحذى صورتين أو أكثر من الاختبار الواحد التي تتوازي كثيراً جداً في مضمون الفقرات وصعوبتها ، والتي تؤدي إلى متوسطات للدرجات قريبة جداً في الصور المختلفة ، والتي تقيس الاختلاف وتقدر الثبات في مجموعة معينة من الأفراد .

طابع القدرات للمهنة Occupational ability Patterns (وقد يسمى نموذج القدرات المهنية) : والمقصود به نمط تنظيم مجموعة القدرات التي يقتضيها النجاح في مهنة معينة . ويتم ذلك بتحديد متوسطات الدرجات لقدرات الناجحين في هذه المهنة . وتعد طريقة تساعد في عمليات الاختيار والتوجيه المهني والتربوي ، فمن يقترب الرسم البياني ، الذي يمثل قدراته ، من هذا الطابع يختار لهذه المهنة أو الدراسة أو يوجه إليها ، وبالمثل يحدد طابع الاستعدادات للمهنة وكذلك طابع الميول للمهنة .

الطريقة الإسقاطية Projective Technique : إحدى طرق الدراسة في الشخصية وفيها يكون الفرد حراً في استجاباته بسلسلة من المنبهات مثل بقع الحبر والصور والجمل الناقصة ... الخ . وفي موقف الاستجابة - الحرة هذا يسقط المفحوص في استجاباته جوانب شخصيته المميزة .

العلاقة المستقيمة Linear Relationship : هي العلاقة بين متغيرين أو أكثر والتي يمكن تمثيلها بخط مستقيم ، وعلى ذلك فإن تزايد أو تناقص أحد المتغيرين يرتبط بزيادة أو نقص الآخر . بنسبة ثابتة على طول التغير . سواء كانت زيادة المتغير الأول تتبعها زيادة مماثلة في الآخر أم نقصاً بنفس المعدل .

العلاقة المنحنية : Curvilinear Relationship : علاقة بين كيتين متغيرتين يمثلها منحني فزيادة المتغير الأول تؤدي إلى زيادة المتغير الثاني . أو نقصه ولكن بمعدل غير ثابت :

عدم التجانس Hetrogeneity : ميل الجماعة إلى أن تكون غير متماثلة .

العمر الزمني (C.A.) Chronological Age : الفترة منذ ميلاد الفرد حتى تاريخ إجراء الاختبار . وتلجأ بعض الاختبارات إلى تعديل هذه الفترة بضررها في معامل خاص وفقاً لظروف خاصة بالاختبار ويرمز إليه في الكتب العربية (ع ز) .

العمر التحصيلي Achievement Age : هو العمر الذي تكون درجة الاختبار التحصيلي تقديراً له أو تقديراً متوسطاً له . وقد يسمى العمر التعليمي : Educational Age

العمر العقلي (M.A.) Mental Age : العمر الذي تكون الدرجة في اختبار الذكاء أو القدرة المدرسية متوسطاً له . والعمر العقلي هو متوسط العمر للأفراد الذين يحصلون على الدرجة المتوسطة . ويرمز إليه في الكتب العربية بالرمز (ع ع) .

العمر القرائي Reading Age : وهو الدرجة المكافئة للعمر التي تعطى للدرجة المتوسطة على اختبار القراءة المطبق على أفراد من عمر زمني معين .

عينة عشوائية Rondon Sample : عينة مأخوذة بحيث أن كل فرد

من المجموعة الأصلية Population أعطيت له فرصة مكافئة لأن يؤخذ ضمن العينة .

عينة ممثلة Representative Sample : هي العينة التي تكافئ المجموعة الأصلية من حيث مكوناتها ومستوياتها ونسب كل فئة فيها :

فقرة الاسترجاع Recall Item : وهي فقرة الاختبار التي تتطلب من الشخص أن يقدم الإجابة الصحيحة من ذاكرته :

فقرة التعرف Recognition Item : فقرة الاختبار التي تتطلب من الفرد أن يختار أو يتعرف على الإجابة الصحيحة من بين اختياريين أو أكثر قدما له .

فقرة التكملة Completion Item : سؤال في الاختبار يتطلب التكملة أو ملء الفراغ في الكلمة أو العبارة أو الجملة أو : : : الخ . ففي الجملة أو الموضوع يحذف جزء أو أكثر وعلى المفحوص أن يأتي به :

قائمة Inventory : أداة تستخدم لسرد كل أو عينة من السلوك أو الاهتمامات أو الاتجاهات أو ... الخ ، تعد مفيدة أو متعلقة بغرض معين : وهي ليست في العادة مقياس وليست لها إجابات صحيحة أو خاطئة .

قائمة الشخصية Personality Inventory : أداة يقصد منها تقييم أو قياس واحد أو أكثر من الجوانب غير العقلية من التكوين النفسي للفرد :

قدرة عامة General Ability : تعبير عام يستخدم في حالة القدرات عندما نتعامل مع مشاكلها على نطاق واسع . ويستخدم كمرادف للذكاء : ويرادف في القياس التربوي القدرة المدرسية :

قدرة خاصة Special Ability : اصطلاح يستخدم مع قدرات من مثل الميكانيكية ، والكتابية ، والموسيقية ، والفنية ... الخ ، متميزة بذلك عن القدرة العامة :

القوة التفريقية Discriminating Power (وقد تسمى القدرة التمييزية) :

قدرة فقرة الاختبار على التفریق بين الأشخاص الذين يتمتعون بدرجة مرتفعة من السمة ، من ناحية ، وبين من يتمتعون بدرجة منخفضة من السمة من ناحية أخرى .

القيم Values : تنظييات معقدة لأحكام عقلية انفعالية نحو الأشخاص

أو الأشياء أو المعاني سواء كان التفضيل الناشئ عن هذه التنظيمات صريحاً أو ضمناً .

قيمة الصعوبة Difficulty value : النسبة المئوية من جماعة معينة ،

مثل تلاميذ من سن واحد أو فرقة دراسية واحدة ؛ التي أجابت عن الفقرة إجابة صحيحة . وكلما كانت الفقرة سهلة ، أو كلما زادت النسبة المئوية من التلاميذ الذين يجيبونها إجابة صحيحة ، زادت قيمة صعوبتها .

المئين Percentile : درجة نقطة في التوزيع تقع تحتها النسبة المئوية من

الدرجات التي يشير إليها المئين المعين . فمثلاً ، المئين الثلاثين هو النقطة التي تقع تحتها ٣٠٪ من الدرجات ،

المتوسط (م) Mean (M) : وسنقتصر في التعريف على المتوسط

الحسابي فهو أكثر طرق المتوسطات شيوعاً : والمتوسط الحسابي Arithmetic mean هو المجموع الكلي لمجموعة من الدرجات مقسوماً على عدد الدرجات . ويرمز إليه في كتب الإحصاء الأجنبية بالرمز $\bar{X}$.

الحكم Criterion : معيار نحكم به على الاختبار أو نقيمه ، وقد يكون

مجموعة من الدرجات أو المقاييس أو التقديرات أو الإنتاج أو ... الخ ؛ صمم الاختبار للتنبؤ بها ، أو للارتباط معها ، كقياس لصدقها . وهو أيضاً مجموعة من المفاهيم أو الأفكار المستخدمة في الحكم على محتوى الاختبار عند تقدير مضمونه أو صدقه المنطقي :

المدى Range : المسافة بين أعلى وأقل درجة في التوزيع .

الاستمرار Continuum : مستقيم ندرك على امتداده سمة أو مجموعة من الدرجات على أن توزيعها مستمر : فمتغير مثل هذا لا يهتمنا فيه أن نعرف مدى تقارب قيمتين مثلاً ، إذ أن من الممكن دائماً أن تكون قيمة ثالثة بينهما . وقد يسمى المتغير المستمر Continuous variable :

معادلة سبيرمان - براون Spearman-Brown Formula : معادلة تبين العلاقة بين طول الاختبار ومعامل ثباته . وبواسطتها يمكن التنبؤ بمعامل الثبات للاختبار إذا زاد أو نقص طوله بقدر ما . وهي تستعمل كثيراً في تقدير معامل ثبات الاختبار بأكمله من معامل الثبات بين نصفى الاختبار نفسه .

معادلة كودر - رتشرد سون Kuder-Richardson Formula : وهي معادلة لتقدير ثبات الاختبار من معلومات عن فقرات الاختبار ، أو معرفة الدرجة المتوسطة ، والانحراف المعياري ، وعدد الفقرات في الاختبار . وهي تدل على الاتساق الداخلى Internal Consistency للاختبار .

معامل الارتباط Coefficient of correlation : مقياس لدرجة العلاقة بين مجموعتين من القياسات مأخوذة إما من نفس المجموعة من الأفراد ، أو من ثنائيات من الأفراد .

معامل ارتباط بيرسون Pearson Product-moment correlation : معامل ارتباط يستخدم في حساب العلاقات المستقيمة على نطاق واسع . وهو يستخدم انحرافات القيم وحواصل ضرب الانحراف على المتغير الأول في الانحراف على المتغير الثانى بالنسبة لكل فرد .

معامل ارتباط الرتب Rank-Order Correlation : طريقة في حساب

معامل الارتباط بترتيب درجات كل الأفراد وتحديد العلاقة بينها . ويعتمد على إيجاد الفرق بين رتبتي درجات الأفراد على المتغيرين .

معامل الثبات Reliability Coefficient : معامل ارتباط نحصل عليه بين صورتين متكافئتين من الاختبار أو بين الدرجات في مرات الإجراء لنفس الاختبار (طريقة إعادة الاختبار Test-Retest) أو بين نصفي الاختبار (الطريقة النصفية Split-half R.) ، أو باستخدام معادلة كودر - ريتشر دسون .

المعايير Norms : قيم تصف أداء مجموعات متعددة على اختبار أو قائمة . والمعايير وصفية لأنماط موجودة من الأداء ولا يجب اعتبارها مستويات مثلى أو مستويات مرغوب في الوصول إليها .

معايير العمر Age Norms : قيم تمثل الأداء الفعلى أو المتوسط لأشخاص من أى مستوى عمر زمنى معين فى أى صفة مقاسة .

التسيع Stanine : درجة معيارية متوسطها ٥ وانحرافها المعياري تقريباً ٢ وتراوح بين ١ ، ٩ . وهى أحد الخطوات فى طريقة المقياس ذى النقط التسع ، نتخذها عند تسوية الدرجات المعيارية . وعلى ذلك فالاسم مأخوذ من المعيار التاسع .

معيار الفرقة الدراسية Grade Norm : الدرجة المتوسطة التى يحصل عليها تلاميذ فرقة دراسية معينة .

المعيار القومى National Norm : معيار مبنى على عينة مأخوذة على نطاق الدولة .

المعايير المحلية Local Norms : معايير حسبت بجمع البيانات من مدرسة معينة أو منطقة تعليمية معينة أو فئة حوضية واستخدامها

بدلاً من استخدام المعايير القومية أو المعايير الإقليمية ، ذلك كي نقدر أداء الأفراد .

Stencil Key : مفتاح أداء مبصم بحيث أنه عند وضعه فوق ورقة الإجابة Answer Sheet ، وهى إما كتيب الاختبار أو ورقة إجابة منفصلة ، يسمح بالتمييز السريع وحصر الإجابات الصحيحة ؛

Graphic Rating Scale : طريقة لتسجيل التقديرات ، تبعاً لقوة الصفة أو السمة على طول خط مستقيم (استمرار) ، وتحت هذا الخط تذكر العبارات الوصفية للسمة المقاسة ؛

Age Scale : مقياس تكون وحدات القياس فيه هى الفروق بين مكافئات الأعمار المتتابة ، ويؤخذ هذا الفرق كما لو كان مكافئاً لأى فرق آخر .

Rating Scales : طرق يقوم فيها الشخص المقدر بتسجيل درجة السمة المعينة أو الصفة التى نقيسها . وهى غالباً إما وصفية ، أو رقمية ، أو مقاييس للتدرج بالرسم .

Point Scale : اختبار يستجيب فيه الفرد عن كل فقرة فيحصل على نقطة أو أكثر إذا كانت إجابته صحيحة . وتتكون درجة الفرد من مجموع هذه النقاط .

Age Equivalent : هو العمر الزمنى الذى تقدره أو تقدر متوسطه الدرجة المعينة .

Grade Equivalent : مستوى الفرقه الدراسية الذى تقدره أو تقدر متوسطه الدرجة المعينة ؛

المنوال Mode (وقد يسمى النمط أو الشائع) : هو الدرجة أو القيمة التي لها أكبر تكرار في التوزيع .

الميول Interests : الميل هو استجابة الفرد استجابة إيجابية أو سلبية نحو شخص أو شيء أو فكرة معينة استجابة لها صبغة وجدانية .

النزعة المركزية Central Tendency : نقطة في التوزيع تميل أغلب الحالات إلى الوقوع حولها ، والاصطلاح يشير إلى مقاييس هذا الميل . كما يشير إلى القيمة الممثلة للتوزيع . وأهم طرق حسابها هي المتوسط ، والوسيط ، والمنوال .

نسبة التحصيل (AQ) Achievement Quotient : وهي نسبة بين المستوى الفعلي للأداء المدرسي وبين المستوى المتوقع ونحدها بنسبة العمر التعليمي إلى العمر العقلي . أي نسبة التحصيل
$$= \frac{\text{العمر التعليمي}}{\text{العمر العقلي}} \times 100$$
 وتسمى أيضاً بنسبة الإنجاز Accomplishment Quotient . وهذا المفهوم لا يزال مستخدماً وإن كان مضللاً ، والسبب في تضليله أنه يعتمد على افتراض خاطئ مؤداه أن العمر التعليمي والعمر العقلي يرتبطان معاً ارتباطاً تاماً عند التلميذ السوي أو المتوسط .

نسبة التعليم (EQ) Educational Quotient : وهي النسبة بين العمر التعليمي والعمر الزمني أي أن :

$$\text{النسبة التعليمية} = \frac{\text{العمر التعليمي}}{\text{العمر الزمني}} \times 100 .$$

نسبة الذكاء (IQ) Intelligence Quotient : هي النسبة التي نحصل عليها بقسمة العمر العقلي على العمر الزمني . أي أن :

$$\text{نسبة الذكاء} = \frac{\text{العمر العقلي}}{\text{العمر الزمني}} \times 100$$

وهي مقياس للقدرة العقلية العامة يأخذ في الاعتبار كلا من الدرجة على اختبار الذكاء والعمر الزمني لمن حصل على هذه الدرجة .

نسبة الذكاء الانحرافية Deviation IQ : درجة معيارية متوسطها ١٠٠ وانحرافها المعياري حوالى ١٥ تحول إليها نسبة الذكاء .

الوسيط Median (وقد يسمى الأوسط) : هو الشخص الذى حصل على الدرجة التى تقع فى منتصف الدرجات أى أن عدد من حصلوا على درجات أعلى منه يساوى عدد من حصلوا على درجات أقل منه . ودرجته ليست بالضرورة الدرجة المتوسطة .

الفصل الثالث

تصنيف وتعريف الاختبارات النفسية والتربوية
وعرض للمشاكل المتعلقة بها .

- مقدمة .
- أسس التصنيف .
- التصنيف على أساس ما نقيسه .
- التصنيف تبعاً للمحتوى .
- التصنيف تبعاً لطريقة التصحيح .
- التصنيف على أساس تطبيق النتائج .
- أسس أخرى للتصنيف .
- تعريف بالأنواع الكبرى من الاختبارات النفسية
وللتربوية .

الفصل الثالث

تصنيف وتعريف الاختبارات النفسية والتربوية
وعرض للمشاكل المتعلقة بها .

- مقدمة .
- أسس التصنيف .
- التصنيف على أساس ما نقيسه .
- التصنيف تبعاً للمحتوى .
- التصنيف تبعاً لطريقة التصحيح .
- التصنيف على أساس تطبيق النتائج .
- أسس أخرى للتصنيف .
- تعريف بالأنواع الكبرى من الاختبارات النفسية والتربوية .

٢ - الذكورة - والأنوثة

٣ - الترويحى

(و) الاتجاهات والقيم :

١ - الاجتماعى

٢ - العنصرى Racial

٣ - السياسى

٤ - الخلقى Moral

٥ - العقائدى religious

(ز) اختبارات حسية حركية :

١ - الإبصار

٢ - السمع

٣ - المهارة الحركية

٤ - التأثر الحسى الحركى

(ا) البسيط (عين - يد)

(ب) المركب (الأطراف والحواس)

٢ - تصنيف على أساس شروط الإجراء Administration :

(ا) فردية (الاختبار يعطى لشخص واحد فى المدة الواحدة)

(ب) جمعية

(ح) التعليمات والفقرات شفوية

(د) التعليمات والفقرات والاستجابات كتابية

(هـ) اختبارات سرعة Speed (تقدم فى زمن حدوده ضيقة لقياس

الكمية التى أداها الفرد أداء صحيحاً فى وقت قصير) .

(و) اختبارات قوة (الحدود الزمنية واسعة ولكن درجة الصعوبة هى

التي تمحدد أقصى ما يصل إليه المفحوص)

٣ - التصنيف تبعاً للمحتوى Contént

(١) الفقرات في صيغة مقال أو فقرات إجابتها حرة .

(ب) فقرات الاستجابة عليها موضوعية أو محددة :

١ - الاختيار من إجابتين :

(١) صحيح - غير صحيح

(ب) نعم - لا

(>) صحيح - خطأ

٢ - الاختيار من إجابات متعددة

٣ - المزاوجة Matching

٤ - التكميل

٥ - صيغ نادرة .

(>) اختبارات أدائية أو غير لفظية :

١ - عامة

٢ - المادة عينة من الشغل في أعمال معينة .

٤ - التصنيف تبعاً للتصحيح :

(١) اختبارات تتطلب تقييماً كيفياً من الخبراء .

(ب) التصحيح بالنسخ . Stencile أى التصحيح بمفتاح مثقب .

(>) التصحيح بالماكينات

٥ - التصنيف على أساس تطبيق النتائج :

(١) اختبارات غير رسمية (اختبارات يضعها المدرسون) .

(ب) اختبارات مقننة :

٦ .. أسس أخرى للتصنيف مثل التفريق بين الاختبارات التي :

- (أ) تتطلب أقصى أداء يستطيعه الفرد (كاختبارات الذكاء والتحصيل و . . . الخ) .
(ب) اختبارات تتطلب أداء عادى أو معتاد (وهى اختبارات الشخصية) .

وكذلك التصنيف حسب التصميم إلى :

- (أ) اختبارات ذات بناء محدد .
(ب) اختبارات غير محددة البناء unstructured (كالإسقاطية)
والتصنيف على أساس الفلسفة القائم عليها الاختبار .

والتصنيف على أساس نوع المواد فيه مثل :

- (أ) لفظية
(ب) أرقام ومسائل حسابية
(ج) رموز
(د) صور
(هـ) مواقف من الحياة

تصنيف على أساس جوانب السلوك التي تمثلها وتصنيف الاختبارات إلى :

- ١ - اختبارات العامل العام
٢ - اختبارات التصنيف العام أو التصنيفية
٣ - اختبارات الاستعدادات الخاصة
٤ - اختبارات القدرات الخاصة
٥ - اختبارات الاستعدادات الفارقة

وهناك تفرقة أخرى على أساس وسط Medium الاختبار إلى :

- ١ - اختبارات - ورقة - وقلم Paper-and-Pencil

٢ - اختبارات أداء Performance

وهناك تقسيم على أساس المحتوى السائد Content

١ - لفظي Verbal

٢ - عددي numerical

٣ - مكاني Spatial

وهناك تصنيف إلى اختبارات :

١ - لغوية language

٢ - غير لغوية non - language

وكذا تصنيف على أساس طبيعة الاختبار إلى :

١ - إسقاطية projective

٢ - مفصلة structured أو ذات بناء محدد :

إلى آخر هذه التصنيفات على أن بكل أساس بعض وجاهته ومبرراته وليس يهم في الكثير أن يقيم أساس التصنيف إذ أن لا طائل في هذا . ذلك فضلاً عن أن التصنيف الجامع المانع لا وجود له بين الاختبارات النفسية وربما كان هذا راجعاً أساساً إلى أنها نفسية . فمثلاً نجد أن الاختبار قد يستخدم في غير ما صمم من أجل قياسه . كما أن بعض الاختبارات يمكن استخدامها استخدامات مختلفة من حالة إلى أخرى ومن قائم بالإجراء إلى آخر .

والآن نتحدث بشيء من التفصيل عن هذه الأسس وتلك التصنيفات معرفين بالأنواع الكبرى من الاختبارات النفسية ، وبالأسس التي صنفت تبعاً لها هذه الاختبارات ومتاولين بعض المشاكل المتعلقة بها . على أننا سنعود للحديث المفصل في فصول قادمة عن أغلب هذه الاختبارات .

وكذلك الحال في اختبارات اختيار المستخدمين والعمال الذي أصبح بعضها يسمى اختبارات « التصنيف العام » أو « اختبارات التصفية » . والاسم مأخوذ من حقيقة أن اختبارات الذكاء ، وخاصة الجمعية منها ، تستخدم كأدوات للتصنيف أى لتقسيم الجماعة المختبرة إلى مستويات من يجتازها نطبق عليه بعد ذلك اختبارات أكثر تفصيلا للاستعدادات الخاصة ، كالاستعداد الميكانيكى ، والكتائى ، والموسيقى ، والحسابى . وأصبحت الاختبارات متوفرة في كل هذه الميادين ، تستخدم في التوجيه التربوى والمهنى ، والاختيار المهنى ، وفي مجالات أخرى كثيرة .

وقد أصبح هناك اعتقاد متزايد لتدعيم ما يسمى باختبارات الذكاء . ولا أدل على هذا من اجتماع ضم ٧٩ من أقطاب علم النفس سنة ١٩٤٤ . وافق ٥٥ منهم على الاعتقاد بأنه « من الممكن لإنجاز الكثير لوركر علماء النفس على عوامل عقلية منفصلة » وليس معنى هذا أن ذلك الفريق لا يعطى أهمية لاختبارات الذكاء ، فعلى النقيض من ذلك وافق ٧٥ ٪ منهم على أنها أصلح الاختبارات في أغراض تصنيف الناس تبعاً للقدرة العامة كما في الجيش ، وفي المدارس ، وفي الصناعة .

اختبارات الاستعدادات الخاصة :

ونتيجة لهذا ظهرت اختبارات الاستعدادات الخاصة Special Aptitude tests . ثم تلاها ظهور اختبارات الاستعدادات الفارقة Differential Aptitude tests وإن كان هدف النوعين هو ملء الفجوات التى تركتها اختبارات الذكاء العام . وميادين الاستعدادات الخاصة ميادين متخصصة جداً ، مثل الاستعداد الموسيقى ، والفنى ، والميكانيكى... الخ ، دعت إليها الحاجة الماسة للاختبار المهنى الذى شجع التنبؤ بهذه الاستعدادات . والاستعداد Aptitude هو إمكانية الوصول إلى درجة من الكفاية أو القدرة

عن طريق التدريب : وفي الاختبارات التي تقيسه يقصد به السرعة التي نتوقع بها أن يتعلم الفرد طرق الأداء المتعلقة بهذا الاستعداد : بمعنى أن الاستعداد للأعمال الكتابية ، مثلاً ، يشير إلى قدرة الفرد على تعلم طرق الأداء المكتنية إذا توفر له التدريب .

واختبارات الاستعدادات تتداخل في وظائفها مع بطاريات الاستعدادات الفارقة ، وإن كان هذا لا يحدث دائماً : وقد توفرت الدراسات المتعددة في مجال الاستعدادات الخاصة حتى أصبحت هناك مقاييس غاية في التخصص تقيس جوانب غاية في البساطة ، مثل حدة الحس أو سرعة حركة الإصبع . وعلى الجانب الآخر نجد اختبارات أخرى معقدة شاملة للاستعدادات مثل الاستعداد للمحاماة أو للهندسة أو للطب :

اختبارات التحصيل بأنواعها :

ويجب أن نشير هنا إلى التمييز الذي قال به البعض بين اختبارات الاستعدادات واختبارات التحصيل . فاختبار التحصيل يقيس أثر نوع معين من التدريب ، ومثله الاختبار التعليمي ، إذ يقنن الاختبار على برنامج دراسة مادة معينة . أما اختبار الاستعداد فيقيس السرعة المتوقعة لتعلم الفرد أو مدى استفادته من التدريب أو التعليم : واختبارات التحصيل في مجال الصناعة تسمى « اختبارات الكفاية » Proficiency test ، فهي تقيس مدى استفادة الفرد من برنامج التدريب :

أما « اختبارات الحرفة » Trade Tests فهي نوع من اختبارات التحصيل أيضاً في مجال الصناعة . وتستخدم في اختيار وتصنيف الموظفين والعمال في الصناعات والحرف . وعلى أي حال فالتفرقة بين الاختبار التحصيلي ، في المدرسة ، واختبارات الكفاية أو اختبارات الحرف ، في المصنع ، تفرقة نسبية أكثر منها مطلقة . إذ أن كل نوع من الاختبارات

وكذلك الحال في اختبارات اختيار المستخدمين والعمال الذي أصبح بعضها يسمى اختبارات « التصنيف العام » أو « اختبارات التصفية » . والاسم مأخوذ من حقيقة أن اختبارات الذكاء ، وخاصة الجمعية منها ، تستخدم كأدوات للتصنيف أى لتقسيم الجماعة المختبرة إلى مستويات من يجتازها نطبق عليه بعد ذلك اختبارات أكثر تفصيلا للاستعدادات الخاصة ، كالاستعداد الميكانيكى ، والكتائى ، والموسيقى ، والحسابى . وأصبحت الاختبارات متوفرة في كل هذه الميادين ، تستخدم في التوجيه التربوى والمهنى ، والاختيار المهنى ، وفي مجالات أخرى كثيرة .

وقد أصبح هناك اعتقاد متزايد لتدعيم ما يسمى باختبارات الذكاء . ولا أدل على هذا من اجتماع ضم ٧٩ من أقطاب علم النفس سنة ١٩٤٤ . وافق ٥٥ منهم على الاعتقاد بأنه « من الممكن إنجاز الكثير لو ركز علماء النفس على عوامل عقلية منفصلة » وليس معنى هذا أن ذلك الفريق لا يعطى أهمية لاختبارات الذكاء ، فعلى النقيض من ذلك وافق ٧٥ ٪ منهم على أنها أصلح الاختبارات في أغراض تصنيف الناس تبعاً للقدرة العامة كما في الجيش ، وفي المدارس ، وفي الصناعة .

اختبارات الاستعدادات الخاصة :

ونتيجة لهذا ظهرت اختبارات الاستعدادات الخاصة Special Aptitude tests . ثم تلاها ظهور اختبارات الاستعدادات الفارقة Differential Aptitude tests وإن كان هدف النوعين هو ملء الفجوات التى تركتها اختبارات الذكاء العام . وميادين الاستعدادات الخاصة ميادين متخصصة جداً ، مثل الاستعداد الموسيقى ، والفنى ، والميكانيكى... الخ ، دعت إليها الحاجة الماسة للاختبار المهنى الذى شجع التنبؤ بهذه الاستعدادات . والاستعداد Aptitude هو إمكانية الوصول إلى درجة من الكفاية أو القدرة

عن طريق التدريب : وفي الاختبارات التي تقيسه يقصد به السرعة التي نتوقع بها أن يتعلم الفرد طرق الأداء المتعلقة بهذا الاستعداد : بمعنى أن الاستعداد للأعمال الكتابية ، مثلاً ، يشير إلى قدرة الفرد على تعلم طرق الأداء المكتتية إذا توفر له التدريب .

واختبارات الاستعدادات تتداخل في وظائفها مع بطاريات الاستعدادات الفارقة ، وإن كان هذا لا يحدث دائماً : وقد توفرت الدراسات المتعددة في مجال الاستعدادات الخاصة حتى أصبحت هناك مقاييس غاية في التخصص تقيس جوانب غاية في البساطة ، مثل حدة الحس أو سرعة حركة الإصبع . وعلى الجانب الآخر نجد اختبارات أخرى معقدة شاملة للاستعدادات مثل الاستعداد للمحاضرة أو للهندسة أو للطب :

اختبارات التحصيل بأنواعها :

ويجب أن نشير هنا إلى التمييز الذي قال به البعض بين اختبارات الاستعدادات واختبارات التحصيل . فاختبار التحصيل يقيس أثر نوع معين من التدريب ، ومثله الاختبار التعليمي ، إذ يقن الاختبار على برنامج دراسة مادة معينة . أما اختبار الاستعداد فيقيس السرعة المتوقعة لتعلم الفرد أو مدى استفادته من التدريب أو التعليم : واختبارات التحصيل في مجال الصناعة تسمى « اختبارات الكفاية » Proficiency test ، فهي تقيس مدى استفادة الفرد من برنامج التدريب :

أما « اختبارات الحرفة » Trade Tests فهي نوع من اختبارات التحصيل أيضاً في مجال الصناعة . وتستخدم في اختيار وتصنيف الموظفين والعمال في الصناعات والحرف . وعلى أي حال فالفرقة بين الاختبار التحصيلي ، في المدرسة ، واختبارات الكفاية أو اختبارات الحرف ، في المصنع ، فرقة نسبية أكثر منها مطلقة . إذ أن كل نوع من الاختبارات

يمكن أن يستخدم ، في ظروف معينة ، لقياس أثر الخبرة الماضية والتنبؤ بالإنتاج المقبل .

اختبارات الاستعدادات الفارقة :

وقد أدى التقييم النقدي لاختبارات الذكاء الذى تبع انتشار استخدامها في كل الميادين إلى توضيح نقطة أخرى ، ألا وهى ، أن الأفراد يختلفون في أدائهم للأجزاء المختلفة من اختبار الذكاء الواحد . اتضح هذا خصيصاً في الاختبارات الجمعية حيث تغزل الفقرات عادة في اختبارات فرعية Subtests . وهذه الاختبارات الفرعية متناسقة المحتوى نسبياً . كما اتضح أن الشخص الواحد يختلف أدائه جوهرياً من قسم إلى آخر في الاختبار الواحد ، خاصة إذا كان أحد الأقسام من نوع لفظي وكان الآخر من نوع الرسوم الهندسية مثلاً .

وكثيراً ما استخدم الفاحصون ، وخاصة الأكاديميون ، تلك المقارنات الداخلية ليصلوا إلى استبصار أكثر بالتكوين النفسى للفرد . وعليه فلم تكن « نسبة الذكاء » أو غيرها من الدرجات الكلية فحسب موضع الدراسة ، بل إن الدرجات على الاختبارات الفرعية أيضاً تفحص في دراسة حالة الفرد ، ومثل هذا الإجراء لم يعمم ، إذ لم تكن هناك اختبارات ذكاء مصممة بغرض تحليل الاستعداد الفارقى . وإن غلب أن حوت فقرات قليلة على تحديد ثابت لقدرة معينة : وبناء عليه ، لو اختبر الفرد مرة ثانية بصورة أخرى من نفس الاختبار لتغيرت درجاته على الاختبارات الفرعية فيها . فإن أردنا القيام بمثل هذه المقارنات في ذات - الفرد ، لاحتجنا إلى اختبارات صممت خصيصاً لتوضيح الفروق في أداء الوظائف المتنوعة .

وقد غطت اختبارات الاستعدادات الفارقة تلك الحاجة ، وهى

نتيجة للتطور في دراسة تنظيم السمات دفعت إليها دراسات إحصائية عن طبيعة الذكاء والقيام بعمل الارتباطات الداخلية والتحليل العامل بفضل كيلى T. L. Kelly وثيرستون L. L. Thurstone . وقد أدت دراسات التحليل العامل Factor Analysis إلى استخلاص عدد كبير من العوامل المستقلة نسبياً أو السمات . وبعضها تمثل بدرجات متفاوتة في اختبارات الذكاء مثل عامل الفهم اللفظي وعامل الاستدلال الحسابي .

كما أدت تلك الدراسات أيضاً إلى تكوين بطاريات الاستعدادات الفارقة لقياس مستوى الفرد في كل من هذه السمات . فبدلاً من استخراج « نسبة الذكاء » تحسب درجة الفرد في سمات مثل الفهم اللفظي ، الاستعداد العددي ، الاستدلال المكاني ، الاستدلال الحسابي Arithmetic Reasoning وقد نجحت مثل هذه البطاريات في تحليل ذات الفرد أى في التشخيص الفارقى . وهى أيضاً تضيف إلى الميدان معرفة منظمة متعمقة لما لم تتضمنه اختبارات الذكاء فضلاً عن اختبارات الاستعدادات الخاصة .

وقد حلت بطاريات الإمتعداد الفارقة محل اختبارات التصنيف العام في عدد من الأغراض مثل توجيه المراهقين والراشدين . وهذه البطاريات تغطى مجالا أوسع من الجوانب النفسية وهى لا تعطى ، كما قلنا ، درجة كلية مثل نسبة الذكاء . وتسمح فضلاً عن هذا ، كما قلنا أيضاً ، بعمل مفارقات أو مقارنات بين القدرات الخاصة للفرد الواحد ، ذلك أنها تصل إلى « صفحة نفسية » Psychological Profile . أو « مبيان نفسى » Psychograph . ويعتمد تكوين بطاريات الاستعدادات الفارقة على آخر ما توصل إليه التحليل العامل ، ما دامت الاختبارات التى تكون هذه البطاريات قد اختيرت لتمثل السمات الأساسية التى حددها التحليل العامل ، وهى لهذا تضمن تغطية قدرات أكثر تنوعاً دون تداخل .

اختبارات الشخصية :

يرى البعض أن القياس النفسى والتربوى هو قياس للشخصية بكل جوانبها ، وهم فى هذا يستندون إلى فهمهم أو نظرتهم للشخصية باعتبار أنها تتضمن كل الجوانب العقلية ، أى المعرفية ، وغير العقلية التى تتضمن الجوانب المزاجية والانفعالية بل والاجتماعية أيضاً . بينما يرى آخرون أننا نكون أحياناً بصدد جوانب عقلية وأحياناً نكون بصدد جوانب غير عقلية ، وفى الأولى نكون بصدد اختبارات القدرات والاستعدادات والتحصيل ، وفى الثانية نكون بصدد اختبارات تقتصر على المميزات الشخصية مثل التوافق الانفعالى والعلاقات الاجتماعية والدوافع والاهتمامات . . . الخ . ولا يعنى هذا أن الفريق الثانى يغفل العلاقة الدينامية التكاملية بين الجوانب العقلية وغير العقلية ، بل لأنهم يرون أن ميزة القياس الأولى هى التحديد ، وهذا فإن الاختبارات التى تقيس الجوانب غير العقلية ، وإن كانت هذه الجوانب نفسها تتأثر بالجوانب العقلية وتؤثر فيها أيضاً ، هى ما يسمونه باختبارات الشخصية .

وعلى هذا سنقصر الحديث فى اختبارات الشخصية على ما يقيس منها الجوانب غير العقلية . وقياس الجوانب غير العقلية من السلوك ما زال فى المهد ، إذ أن اختبارات الشخصية هنا تتضمن ، مثلاً ، مقاييس للتوافق الانفعالى التى تعرف باختبارات عدم الثبات الانفعالى أو « قوائم الشخصية » Personality Inventories كما تتضمن مقاييس السمات الاجتماعية بما فى ذلك العلاقات الأساسية مع الأشخاص مثل الخضوع – السيطرة ، والانطواء – الانبساط ، وكذلك الكفاءة – الذاتية . وتتضمن أيضاً قياس « الذكاء الاجتماعى » ، وهو المعرفة والمهارات المطلوبة فى كل المواقف الاجتماعية .

والذكاء الاجتماعى مشكلة تقابلنا عند وضعه تحت أى من التصنيفات .
 فاختبار الذكاء الاجتماعى ، بما أنه يقيس المعرفة والمهارات التى يقتضها
 التوافق فى المواقف الاجتماعية ، هو قياس لمدى القدرة على تكوين العلاقات
 الاجتماعية من حيث الكم أو العدد ومن حيث العمق أى مدى التأثير المتبادل
 بين طرفى أو أطراف العلاقة الاجتماعية ، ومن حيث الشدة أى حرارة
 العلاقة وفتورها ، ومن حيث الدوام أو الثبات أى مدى بقاء هذه العلاقة فى
 حالة سوية . الذكاء الاجتماعى بهذا المفهوم التحليلى البنائى الوظيفى Functional
 Structural Analysis (F.S.A.) قد يضم إلى اختبارات الذكاء فهو قدرة
 أو إلى اختبارات الشخصية فهو توافق كما يضم إلى اختبارات الاستعدادات
 فهو يشير إلى مدى سرعة الفرد فى تكوين العلاقات خلال المواقف .

وتنمى إلى اختبارات الشخصية أيضاً مقاييس السمات الخلقية مثل
 الأمانة ، التحفظ ، والتعاونية . وتشمل اختبارات الشخصية أيضاً مقاييس
 الدفع ، والاهتمامات ، والاتجاهات . والطرق المستخدمة فى اختبارات
 الشخصية هى الاستبيان ، والطرق الإسقاطية ، وطرق التقدير الذاتى ،
 والمقاييس المتدرجة ، والاختبارات الموقفية :

وأول سابقة على قياس الشخصية نجدها فى أعمال كريبلين Kraepelin
 الذى استخدم طريق « التداعى الحر » Free Association مع مفحوصين
 غير أسوياء ، فكان المثير كلمة يلقيها . هو على المفحوص الذى يرد بأول
 استجابة ترد بخاطره . كذلك اختبارات عن الآثار النفسية للتعب ، والجوع ،
 والمخدرات . حيث انتهى إلى أن كل هذه الوسائل تزيد التشوش النسبى للتداعى
 السطحى . كما أن سومر Sommer دعم طريقة التداعى كوسيلة فنية فى
 القياس ويجب الاعتراف بفضل جالتون ، وپيرسون ، وكاتل ، فى تقنين
 الاستبيانات Questionnair والوسائل الفنية للتقدير المتدرج إذ أسهمت كل

هذه الوسائل في دراسة الشخصية ، وإن لم تهدف ، عندما صممت ، إلى قياس الشخصية .

ومن أمثلة استبيان الشخصية ، أو قائمة التقرير الذاتي Self Report Inventory ، صحيفة البيانات الشخصية Personal Data Sheet التي وضعها وودورث Woodworth في الحرب العالمية الأولى : وكان هدفه تشخيص الحالات العصبية (Neurotic) الخطيرة التي قد تلتحق بالخدمة العسكرية . شملت هذه الصحيفة عدداً من الأسئلة عن الأعراض العصبية العامة يجيب عنها الشخص بنفسه وفي بعضها قوائم فرعية للاضطراب الانفعالي ، ثم قوائم أكثر تخصصاً عن التوافق في البيت وفي المدرسة وفي العمل : هذا وإن ركزت قوائم تالية ، وضعها آخرون ، على استجابات أكثر تميزاً ، مثل السيطرة - الخضوع . ثم تلها استبيانات عن الاهتمامات والاتجاهات .

وهناك نوع آخر من مقاييس الشخصية هو الاختبارات الموقفية ، وهدف هذه الاختبارات لا يعلن للمفحوص ، إذ نلاحظ أرجاعه واستجاباته دون علمه ، قام بأول سلسلة من هذا النوع هارتشون وماي Hartshone and May وقتناها على تلاميذ المدارس ، فيما يتعلق بالغش ، والكذب ، والسرقة ، والتعاونية ، والمثابرة . كما صممت سلاسل أخرى في الطيران كان التقدير فيها ذاتياً .

أما الطرق الإسقاطية ، فهي أحد مناهج دراسة الشخصية التي يتبعها الإكلينيكيون . وبناء مادة الاختبار فيها « غير محدد » Unstructured كي يسمح لذاتية الفرد أن تعمل . والفكرة وراؤها هي أن الفرد « سيسقط » طريقه المميزة في الاستجابة عند التطبيق . وهدف هذه الاختبارات يخفى على المفحوص ، ومن ثم تقل الفرصة أمامه في أن يخلق متعمداً انطباعات مرغوباً فيه . وأول أنواعها اختبارات التداعي الحر ، تلها اختبارات تكوين الحمل :

والطرق الإسقاطية تلجأ إلى الرسم وتركيب اللعب لتكوين منظر ، وإلى الصور الغامضة ، ويقع الخبر ، . . الخ : وأكثرها انتشاراً اختبار « الرور شاخ » Rorschach Ink Plots واختبار رسم « المنزل والشجرة والشخص » House, Tree, Person (H.T.P.) واختبار « تفهم الموضوع » Thematic Apperception test (T.A.T.)

وأحدث تطور في إعداد اختبارات الشخصية يعتمد على نتائج التحليل العامل . وتختار فقرات هذا النوع من بطاريات الاستعدادات ، وتجمع على أساس الارتباطات ونتائج التحليل العامل . والنمط الغالب منها هو اختبارات الورقة والقلم ، ولكنها لا تزال في دور الاستطلاع .

وباختصار فإن الأنواع الموجودة من اختبارات الشخصية تعرض مشاكل خطيرة نظرية وعملية أكثر مما تحل ، فحلولا ليست تامة دائماً . ولكل منها مزاياه الخاصة وعيوبه :

والطرق المختلفة والنظرات المتعددة ونواحي القصور في اختبارات الشخصية عموماً لا ترجع إلى تقصير الجهود بقدر ما تعزى إلى طبيعة الشخصية ، فإن هذا المفهوم المركب المعقد الفردى الفريد هي التي تخلق أكثر مشاكل قياسها .

والتمييز بين اختبارات القدرة ، واختبارات الشخصية ، برغم ما قلنا ، تمييز مربك مصطنع ، فالفرد عند تطبيق أى اختبار عليه يتأثر بعوامل القدرة وخصائص الانفعال وبدوافعه واهتماماته وبخصائص غير عقلية أخرى . ولكن في بناء الاختبار عادة تعطى الأهمية إلى نواح في جوانب السلوك : وتظهر تلك الأهمية في التعليمات المعطاة للمفحوص وفي تكوين العلاقة الطيبة والتصحيح والتفسير إلى آخر ذلك من عمليات القياس .

ثانياً - التصنيف على أساس شروط الإجراء :

ويقصد عادة بشروط الإجراء ما إذا كان الاختبار قن ، وبالتالي ،

يمكن تطبيقه على المفحوصين أفراداً ، أم أنه قنن وبالتالي ، يطبق على المفحوصين جماعات . ومن شروط الإجراء أيضاً ما يتعلق بالإعداد لموقف الاختبار وبخطوات عملياته وما إذا كان له وقت محدد ، وهنا تصبح السرعة هي العامل الأساسي ، أم أن الوقت فيه متسع كي يجيب من يقدر على الإجابة وهنا تصبح القدرة هي العامل الأساسي : ومن شروط الإجراء أيضاً ما يجب أن يتوفر في الفاحص وما يجب توفره في المفحوص وفي الموقف .

وتقسم الاختبارات اعتماداً على هذا الأساس إلى « اختبارات جمعية » Group tests ، و « اختبارات فردية » Individual test . والاختبارات الجمعية تطبق على مجموعة أو مجموعات من الأفراد في وقت واحد . وهنا يكون موقف القياس جمعياً ويكون المفحوص عضواً في جماعة . أما الاختبارات الفردية فيطبق الواحد منها على فرد واحد في المرة الواحدة ، وهنا يكون موقف القياس موقفاً فردياً Individual Setting .

وأبسط فرق ظاهري بين هذين النوعين من الاختبارات أن التعليمات في الاختبارات الجمعية أبسط منها في الفردية . وتعطى التعليمات في الموقف الجمعي عند بدء الاختبار مرة واحدة في الجلسة الواحدة . في حين أن التعليمات ، في الاختبارات الفردية ، معقدة وكثيرة وهي تختلف من فقرة إلى أخرى في الاختبار الواحد ، وكل فقرة أو مجموعة فقرات تخصها مجموعة من التعليمات :

وللاختبارات الجمعية ، عادة ، ميزة أخرى هي أن تطبيقها ، من جانب القائم بها أي الفاحص ، لا يحتاج إلى مهارة فائقة وتدريب شاق طويل متخصص ، كما لا يشترط غالباً أن يكون الفاحص إخصائياً نفسياً ، وليس في هذا خروج على أخلاقيات مهنة القياس : بينما نجد ، في الوقت نفسه ، أن معظم الاختبارات الفردية يجب أن يتوافر في القائم بها تدريب شاق طويل متخصص ومهارة فائقة : وفي العادة ، يقوم بها إخصائى نفسى مدرب عليها : وأحياناً ينبغي أن يكون قد درب على الاختبار الذى يطبقه بالذات

فضلاً عن إعداداته السابق . فإن لم يتوفر هذا كان خروجاً على أخلاقيات المهنة : وإن كان هذا الخروج يحدث أحياناً في تدريب الإخصائيين النفسيين في الجامعات ومراكز البحوث .

والاختبارات الفردية تطبق كثيراً في البحوث الإكلينيكية ، وهنا لا يهتم الفاحص بالكم أى بمدى صحة الاستجابات وبدرجة المفحوص في الاختبار . بل إنه يهتم أساساً بالكيف أى بطريقة استجابة المفحوص فيلاحظ مدققاً أموراً عدة من مثل زمن الرجوع أى الوقت الذى مضى بين السؤال والجواب وبالاتفاعلات التى صاحبت الإجابة من تعبيرات الوجه وحركة الأطراف ، وبنوع الأخطاء ، فقد يخطئ الفرد في نوع معين من الأسئلة وينجح فى أخرى تتطلب مهارة أكثر ، وهنا يفترض أن ثمة أسباباً لخطئه . فقد تتعلق الأسئلة التى فشل فيها بخبرات غير سارة أو باضطراب فى القدرة أو غير ذلك من التفسيرات الإكلينيكية .

ولهذا السبب نجد أن الكثير من المؤسسات والهيئات العلمية ومراكز البحوث لا يأخذ بنتائج الاختبار وتفسيره إلا إذا طبقه سيكولوجى درب عليه نظرياً وعملياً ، بل ، وأحياناً اجتاز فيه امتحاناً للتأكد من صلاحيته للقيام بهذا العمل . وفى العادة يجرى الإخصائى النفسى الاختبار فى كراسة منفصلة ويصححه بنفسه ، وفى نفس الوقت ، يقوم أخصائى نفسى خبير أو مشرف عليه بتسجيل إجابة المفحوص فى كراسة أخرى مماثلة ثم يصححها ويفسرها . ولا يؤخذ بتصحيح الإخصائى النفسى المبتدئ إلا إذا اتفق تصحيحه مع تصحيح الإخصائى النفسى الخبير :

وعلى ذلك يجب التفريق بين نتائج كل من الاختبارات الفردية والاختبارات الجماعية . إذ أن الاختبارات الجماعية تطبق فى مجالات تمس الأفراد من حيث هم أعضاء فى مجموعات كبيرة ، فى حين أن الاختبارات الفردية تمس الفرد من حيث هو فرد . ونتيجة لطبيعة وظروف موقف الإجراء ، يرى

البعض أن الاعتماد على نتيجة الاختبار الجمعى أسلم منه على الفردى ، خاصة وأن الاختبار الجمعى فى تقنيته لا يسمح للقائم بإجرائه بالتصرف الشخصى ، فلا تعمل ذاتية الفاحص فى الإجراء ، كما لا تؤثر فى التصحيح الذى يتم بطريقة آلية دائماً ، ودور الفاحص فى تفسير نتائج الاختبار الجمعى محدود إذا قورن بدوره فى تفسير نتيجة الاختبار الفردى .

والمفحوص ، فى الموقف الجمعى ، يكون أقل شعوراً بالموقف منه فى الموقف الفردى . إذ يحس فى الأخير بالحرج ولا يألف الموقف سريعاً فى حين أنه فى الموقف الجمعى يدرك أن غيره يشاركه فيما يمر به ، كما أنه يكون منصرفاً عن الشعور بالموقف لأنه يعلم أن أحداً لا يركز الانتباه عليه . وقيمة هذا أن معايير الاختبار المقنن على مواقف جمعية لا تجعله صالحاً للإجراء فى المواقف الفردية . فإن المعايير المستخلصة بطرق إحصائية تمكنا من مقارنة الأفراد المنتظر قياسهم بنظرائهم من مجموعة التقنين . وعلى هذا نطبق المقياس فى مواقف من نفس نوع المواقف التى قن فيها . خاصة وأنه فى المواقف الجمعية تبرز عوامل تفتقر إليها المواقف الفردية مثل التعاون ، والتنافس ، والشعور بأن الخبرة عامة يمر بها أفراد مختلفون فى نفس الوقت ، وتوحيد الظروف ، وتوحيد طريقة الإجراء ، . . . الخ .

كما أن هناك عوامل تظهر فى الموقف الفردى يفتقر إليها الموقف الجمعى ، مثل إقامة علاقة طيبة ، وضمان تعاون المفحوص ، وكسب ثقته ، وإثارة اهتمامه ، وإمكان الحصول على استجابات قد لا يمكنه أن يدلى بها فى الموقف الجمعى . هذا وإن كانت هذه المميزات وتلك النقائص تختلف ، من موقف إلى آخر ، ومن اختبار إلى آخر ، ومن فاحص إلى آخر ، كما يرجع الكثير من الفروق إلى نوع الوظيفة أو الوظائف المقاسة وطبيعة المفحوصين . بدليل أن الاختبار الواحد قد يقن جمعياً مرة وفردياً مرة أخرى : وتختلف المعايير فى الحالتين كما حدث فى اختبار المصفوفات

المتابعة Progressive Matrices ، إذ اختلفت متوسطات درجات عينتي التقنين وكذلك الانحرافات المعيارية وبالتالي انخفضت الدرجة المتوسطة في التقنين الفردي عنها في الجمعي .

وتعد اختبارات الجيش ألفا والجيش بيتا Army Alpha & Army Beta التي وضعت في الحرب العالمية الأولى ، تأليف أرثر أوتس Arthur Otis ، من أوائل الاختبارات الجمعية للذكاء . وتطبق الاختبارات الجمعية الآن على نطاق واسع في الصناعة باسم اختبارات الكفاية ، وفي المدارس تسمى اختبارات التحصيل . أما عن الاختبارات الجمعية للشخصية فإن أشهرها « قائمة مينيسوتا المتعددة الأوجه للشخصية » Minnesota Multifasic Personality Inventory (M.M.P.I.) التي وضعها هاثاوي وما ككنلي Hathaway and MaKinley بجامعة مينيسوتا وتتكون هذه القائمة من عدد من المقاييس التي تقيس أوجهها مختلفة للشخصية . وهي تصلح جمعية كما تصلح فردية . وفي الحالة الجمعية يكون الاختبار على شكل كتيب Booklet form أما في الحالة الفردية فيكون الاختبار على شكل بطاقات Card-form على كل منها أحد أسئلة القائمة وعلى المفحوص أن يسجل في ورقة الإجابة ما إذا كان مضمون الفقرة ينطبق عليه فيملأ الخانة أمام (نعم) تحت رقم هذا السؤال . وإذا كانت الفقرة لا تنطبق على المفحوص أولاً يوافق عليها ، ملأ الخانة (لا) تحت رقم هذا السؤال بالذات : والأسئلة تقيس عشرة جوانب من الشخصية كل جانب ينحصر عدد من الأسئلة ، ولكن الأسئلة كلها مرتبة عشوائياً بحيث لا يتبين المفحوص أن مجموعة بلداتها تتناول جانباً بذاته من جوانب شخصيته . وبهذه الطريقة العشوائية أيضاً يمكن أن تعاد صياغة السؤال ويتكرر وردوه في مكان آخر ، وذلك للتحقق من مدى اتساق وثبات استجابات المفحوص :

وسنعود للحديث عن اختبار M.M.P.I. هذا تفصيلاً في فصل آخر :

أما جيلفورد وزوجته J.P. Guilford and Ruth E. Guilford فقد قاما بعمليات التحليل العاملى لمقاييس الشخصية، واستخلصا ثلاثة عشر عاملاً : ثم وضعت اختبارات ثلاثة تقيس هذه العوامل ، أولها اختبار S.T.D.C.R. والحروف مأخوذة من اختصارات العوامل الخمس التى يقيسها هذا الاختبار . وسنعرض لهذه الاختبار والاختبارتين المشتقتين من الثمانية عوامل الأخرى (الثانى يقيس خمس عوامل والثالث يقيس ثلاث عوامل) فيما بعد .

وكذلك اختبار الورشاخ لبقع الحبر الذى تبذل الآن محاولات لتطبيقه جمعياً بعرض بطاقات البقع بواسطة الفانوس السحرى . ومن اختبارات الشخصية اختبار T.A.T. وأغلب الطرق الإسقاطية ، على العموم ، فردية .

وقضية الاختبارات النفسية الفردية قضية تعرضت للكثير من الادعاءات والالتمامات وإن ثبت بطلان أغلب هذه الدعاوى . وكل هذه الدعاوى تقريباً تتعلق باستخدام الاختبارات النفسية استخداماً فردياً عن يد الإكلينيكين : وأصحاب هذه الإدعاءات أساساً هم الأطباء النفسيون وحجتهم فى هذا أن المفحوص فى الموقف الفردى يكون فى حرج وأنه يكون واعياً بذاته بدرجة تعطل استجاباته الطبيعية ، أى التى تمثل تكوينه النفسى . والحق أنهم فى إنكارهم هذا يستندون إلى إطار مرجعى معين فى التفكير يفضى بهم فى نهاية الأمر إلى إنكار قيمة القياس كله فهم ضعيفو الإيمان به : ودفع هذا الادعاء يأتى من أننا هنا نقارن المفحوص بمفحوصين كانوا ، فى جماعة التقنين ، يمرون بنفس الموقف تقريباً ويشعرون بالمعاناة من ذات مشاعر الوعى الشديد بالموقف وبالتحرج إلى آخر هذه المشاعر .

فهنا يصدق على الفرد ما يصدق على من وضع في موقف مشابه : وأساس المقارنة إذاً هو نفس أساس المقارنة في كل عمليات القياس ، ألا وهو توحيد الظروف . ومقياس المقارنة هو هو انحراف الشخص عن معيار معين انحرافاً ذا دلالة على أن موضعه النسبي بعيد عن متوسط موضع من مروا بالموقف نفسه ، وأن هذا البعد يقارن بمتوسط بعد أفراد عينة التقنين عن متوسط درجات أفرادها (ومقياس المقارنة إذاً هو المتوسط ، م ، والانحراف المعياري ، ع) .

ومن الأمثلة الواضحة على الاختبارات النفسية والتربوية من النوع الفردي فيما يتعلق بالذكاء والقدرة العقلية مقياس « ستنفرد - بينيه » مراجعة سنة ١٩٣٧ التي أعدها تيرمان ، ميريل . وكذلك مقياس « وكسلر بلقيو لذكاء الراشدين والمراهقين » الذي وضعه وكسلر وهما أفضل وأشهر المقاييس العالمية لقياس القدرة العامة . أما عن اختبار « المصفوفات المتتابعة » الذي وضعه رافن فهو فردي ، إذ قن في مواقف فردية ، وجمعي إذ قن في مواقف جمعية : وله صسورتان أولاهما للأطفال والأخرى للراشدين :

ومن الاختبارات التي استطاعت أن تفرق بين القدرات المعرفية ، اختبار « فيرنو » الذي وضعه الدكتور فيرنو Furneaux ، وهو اختبار جمعي ميزته التفريق بين القدرات المعرفية في ناحيتين هما : الأولى هي الدقة في حل المقياس ، والثانية هي السرعة في العمليات العقلية . وللعاملين أهميتهما في التقييم والتشخيص والاختيار والتوجيه .

ونعود إلى الحديث عن الاختبارات الفردية في مقابل الجمعية ، فالاختبارات الفردية مثل « ستنفرد - بينيه » تتطلب مهارة عالية من الفاحص وهي مصممة إكلينيكية أساساً للبحث المتعمق للحالات الفردية : وهي تمكن

الفاحص من الملاحظة الدقيقة . إذ يلاحظ ، وقد يسجل ، نوع الأداء ، والاستجابات الاجتماعية والانفعالية للمفحوص . والفاحص هكذا يعرف أكثر من مجرد « نسبة الذكاء » عن المفحوص ، خاصة إذا كان اختصاصياً نفسياً خبيراً . والاختبار الفردى زيادة على ذلك يمثل مقابلة مقننة فيوفر فرصة لإقامة علاقة طيبة وللتعلون والاحتفاظ باهتمام المفحوص بالاختبار والموقف . كما أن الفاحص يتحاشى أى ظروف معطلة للمفحوص .

وكثير من فقرات الاختبارات الفردية يحتاج من الفحوص إلى استجابة شفهية أو معالجة يدوية . ومعظمها يلزمه التوقيت الفردى عند الإجراء . كما أن هذا النوع من المقاييس نشأ عن اهتمام بحالات الاضطراب فى القدرة العقلية والتخلف المدرسى ، ولهذا فهى أدوات إكلينيكية . بينما نشأت وتطورت الاختبارات الجمعية بفضل الحاجة العملية المتزايدة إلى اختبارات للتصنيف والتصفية وحصر الإمكانيات العقلية وتخطيط القوى البشرية . وقد حدث مثلاً سنة ١٩١٧ أن طلب من علماء النفس الإسهام بجهودهم فى الحرب العالمية الأولى فأعد آرثر آوتس Arthur S. Otis اختبارين جَمِيعين هما Army Alpha و Army Beta . الأول يستخدم فى عمليات القياس العادية مع مفحوصين متعلمين يعرفون الإنجليزية ، أما الثانى فيطبق على الأميين والأجانب الذين لا يعرفون الإنجليزية . والاختباران جَمِيعان يمكن تطبيقهما على جماعات كبيرة . عدل الأول بعد الحرب العالمية الأولى كى يناسب الشئون المدنية :

والاختبارات الجمعية مبسطة التعليمات وخطوات الإجراء ، بحيث لا يصعب الإجراء على الفاحص المتوسط التدريب ؛ وإن كان انتشار استخدام الاختبارات الجمعية قد أدى إلى تطوير حركة القياس النفسى والتربوى وخاصة فى مجال الجيش والمدرسة ، إلا أن سهولتها النسبية فى الإجراء والتصحيح وشيوع استخدامها قد جلبا على حركة القياس التورط

والتعرض لكثير من النقد ، إذ أصبحت الاختبارات (موضحة) ، وأصبح في وسع من يريد قياس ابنه أو عائلته ؛ حتى لو لم يكن ذا علم بأى الأسس النفسية ، أصبح في إمكانه أن يقوم بهذا ، فالأمر بسيط . كما أصبح في إمكان بعض ذوي الأغراض الشخصية أن يجروا الاختبارات الجمعية فيثبت أحدهم مثلاً أن متوسط ذكاء الجزائريين منخفض عن متوسط ذكاء الفرنسيين ، كما قام إنجليزى آخر بإعلان أن فئة من شعوب أواسط أفريقيا تقل في ذكائها وفي مستوى عملياتها العقلية كثيراً عن أبناء أوروبا البيض ١١

أدى استخدام الاختبارات الجمعية على نطاق واسع إذاً إلى الوصول إلى ما رفضه العلم وأثبتت خطأه البحوث التجريبية لعلماء من نفس الجنسيات (الفرنسية والإنجليزية) . كما أدى هذا إلى أن يدعى بعض المشتغلين بحركة القياس ، وخاصة غير المتعمق منهم إلى أن يؤمل في الكثير من الاختبارات وإلى أن يعلن للعامة أن مختلف جوانب الحياة العقلية والنفسية يمكن تقديره بدقة فائقة : وهكذا ، توقع العامة من الاختبارات أن تحقق أكثر مما تستطيع ففشلت في مسابقة توقعاتهم ، بل وأصبحوا يعممون هذا الفشل على كل الاختبارات النفسية : فهي في نظرهم لا تستطيع شيئاً وتتكلم عن مفهومات غامضة لا تستطيع تقديرها تقديرأ ثابتأ دقيقأ .

اختبارات القوة واختبارات السرعة

يفرق بين الاختبارات النفسية والتربوية على أساس شروط الإجراء أيضاً تبعاً للعامل الأساسى الذى تتوقف عليه طبيعة استجابة المفحوص . فالاستجابة مرة تعتمد على عامل السرعة speed ومرة أخرى تعتمد على عامل القوة power . ففي اختبارات السرعة يفرق بين الأفراد تبعاً لسرعتهم في الأداء ، ولهذا فاختبار السرعة يتكون من فقرات قليلة الصعوبة ، وقد يتكون من فقرات متساوية في الصعوبة تقريباً بحيث يستطيع المفحوص أن

يحلها جميعها . ولكن عدد هذه الفقرات أكبر من أن يحلها المفحوص فعلا في حدود الزمن الذى وضع للاختبار: والزمن هنا محدد بحيث لا يكون في إمكان أحد أن يجيب عن الفقرات كلها . ومن هنا تكون درجة الفرد وهى عدد الأسئلة التى أجاب عليها إجابة صحيحة ممثلة لسرعته فى أداء ما يستطيعه ؛ وعلى هذا تكون اختبارات السرعة هى اختبارات فى الدقة والسرعة معاً فى أداء مستوى خاص من الأسئلة .

وعلى الطرف الآخر نجد اختبارات القوة Power tests وهى اختبارات أسئلتها متدرجة فى الصعوبة والغرض منها قياس أعلى مستوى يستطيع المفحوص أن يجيب عليه وتكون درجة المفحوص هى عدد الأسئلة التى أجاب عليها إجابة صحيحة . ويراعى فى اختبارات القوة أن يكون الزمن كافياً لكل مفحوص أن يصل إلى أعلى مستوى يستطيع الإجابة عليه إجابة صحيحة . لذلك فاختبارات القوة تقيس قوة المفحوص فى المادة موضوع الاختبار ولكنها فى نفس الوقت تقيس دقة المفحوص فى الإجابة عن هذه الأسئلة .

أى أن الدقة تدخل على أى حال فى كل من اختبارات السرعة واختبارات القوة . وبعبارة أخرى تكون الدرجة التى يحصل عليها المفحوص فى اختبار القوة دليلاً على أقصى مستوى صعوبة يستطيع الإجابة عليه إجابة صحيحة :

والمثال التالى يوضح الفرق بين اختبارات السرعة واختبارات القوة ؛ فاختبار السرعة يقابل سباقاً وضعت فيه موانع كثيرة جميعها من ارتفاع صغير يستطيع جميع المتسابقين أن يتخطوها ، وتقاس هنا سرعة المتسابقين بعدد الموانع التى استطاعوا أن يتخطوها فى زمن محدد ؛

أما اختبار القوة فهو كالسباق الذى وضعت فيه موانع متدرجة فى الارتفاع وقصص إلى ارتفاعات لا يستطيع أى واحد من المتسابقين أن يجتازها : ففى هذا السباق تقاس قوة المتسابق بأعلى ارتفاع من الموانع

استطاع اجتيازه بنجاح ويكون ارتفاع هذا الحاجز هو أقصى قدرته على القفز .

ويجب أن نلاحظ هنا أن غالبية الاختبارات الشائعة في التحصيل وفي الذكاء وفي القدرات تجمع بين خصائص اختبارات السرعة واختبارات القوة فهي تتكون من أسئلة متدرجة في الصعوبة ولكن الدرجة تمثل مقدار ما أجابه المفحوص إجابة صحيحة بصرف النظر عن صعوبة الأسئلة المحابة .

ونحن نجد في اختبارات القوة أن الزمن المحدد يسمح لأى مفحوص أن يحاول الإجابة عن كل الفقرات إن كان يقدر أى إن كان مستواه في الوظائف التى يقيسها الاختبار مرتفعاً بحيث ينهى الإجراء وتكون إجاباته صحيحة . والفقرات هنا متدرجة الصعوبة . ولكن غالباً ما يستحيل ، نسيباً ، حل بعضها بحيث لا يحصل المفحوص على الدرجة النهائية . فالمفحوص إذاً في اختبار السرعة لا يحصل على الدرجة النهائية وإن كان مستواه في الوظائف التى يقيسها الاختبار مرتفعاً بدرجة تمكنه من أن يحل كل الفقرات ، ولكنه لا يحلها جميعها فعلاً لأن الزمن المحدد يحول دون ذلك . والمفحوص في اختبارات القوة ، أيضاً ، لا يحصل على الدرجة النهائية ، ولكن ليست السرعة في حل الفقرات في حدود الزمن هى التى تحول دون ذلك كما في اختبارات السرعة ، ولكنه مستوى قدرته الذى يحول دون ذلك .

والفكرة في ذلك هى أن واضع اختبار السرعة لا يستطيع معرفة أقصى سرعة لمن سيطبق عليهم الاختبار ، ولهذا فهو يضع عدداً كبيراً من الأسئلة يستبعد أن يجيب عنه أى مفحوص إطلاقاً في حدود الزمن المعين للاختبار . وهنا يكون سقف الاختبار كافياً لتمكين كل فرد من أن يظهر كل ما يستطيع إنجازُه من الفقرات من حيث العدد : أما واضع اختبار القوة فهو لا يعرف بالتحديد مستوى قدرات من سيجرى عليهم الاختبار ، ولهذا فهو يضع أعلى

مستوى من الصعوبة . ويكون سقف الاختبار هنا كافياً لتمكين كل فرد من أن يظهر كل ما يستطيع لإنجازه من الفقرات من حيث الصعوبة .
هذه التفرقة بين اختبار السرعة واختبار القوة تفترض أن بعض الاختبارات خالص ونقى من القوة والبعض الآخر خالص ونقى من السرعة . فهى تفرقة كمية إذ أن أغلب الاختبارات يجمع بين الأساسين بدرجات متفاوتة . وتلزمنا معرفة السرعة والقوة ومدى تأثيرها في أداء المفحوص للاختبار المعين عندما نفسر درجاته . كما يلزمنا هذا عند التقييم الفنى للاختبار ، إذ أن تجاهل عامل السرعة قد يوقع في خطأ كبير عند حساب ثبات الاختبار reliability كما سنعلم فيما بعد .

ثالثاً — التصنيف تبعاً للمحتوى :

هناك اختبارات تكون فقراتها في صيغة مقال essaay وتكون الإجابة عنها حرة free كبعض اختبارات القراءة التى نطلب فيها من المفحوص أن يقرأ مقالاً ويخلص ما فهمه . وكذلك الاختبارات التى نطلب فيها من المفحوص أن يقرر أهم خصائص موقف نعرضه عليه كتابة . أو تلك التى يطلب فيها حل مشكلة . وقبل أن نستطرد في هذا الحديث ، نعرف المحتوى في هذا المجال بالذات ، بما يطلب من المفحوص أى بطريقة الاستجابة : وطريقة الاستجابة قد تكون ذاتية كما في اختبارات المقال ، أو موضوعية كما سنشير حالاً ، أو أدائية أو غير لفظية .

واختبارات حل المشاكل اختبارات تطلب الإجابة الذاتية في صورة حرة عن فقرات موضوعية في صيغة مقال . وهى تعقد للمرشحين لشغل الوظائف التى تتطلب تصرفاً شخصياً ، وهى عادة وظائف عليا من وظائف الإدارة ، كذلك عند اختيار أفراد يقومون بالمقابلة أو يتعاملون مع الجمهور ، أو إخصائين في شئون الأفراد . ويتمثل هذا النوع أيضاً في بعض اختبارات الإنشاء إذ يطالب المفحوص بالتلخيص والتركيز وأحياناً بعرض الرأى .

ونجد أمثلة لاختبارات المقال في بعض الاختبارات المهنية التي تستخدم في ميدان الإرشاد النفسى والتربوى والمهنى ، حيث يقرأ المفحوص بعض الفقرات ويكتب عن تأثيره بها أو عن تأثيره بنوع مماثل من خبراته الماضية . والنوع الآخر من الاختبارات النفسية المقابل لاختبارات المقال هو الاختبارات التي تكون فقراتها موضوعية وتكون الاستجابة عنها موضوعية ومحددة . كأن يجيب المفحوص بصحيح أو خطأ ، أو موافق أو غير موافق ، أو مؤيد جداً - مؤيد - محايد - معارض - معارض جداً كما في اختبارات الاتجاهات ، وقد يجيب المفحوص بصحيح أو خطأ في مقاييس الحكم . ومن قبيل التحديد في الاستجابة نجد اختبارات يجيب المفحوص فيها باختيار إجابة واحدة من بين عدة إجابات موضوعية أمام السؤال . وقد تكون الإجابة بالزوجة matching فتوضع مسائل متتالية ، مثلاً ، في صف وتوضع حلول هذه المسائل في صف آخر . والحلول ليست مرتبة بحسب ترتيب المسائل وعلى المفحوص أن يختار من بين الإجابات إجابة معينة يعتقد أنها صحيحة لتناسب أسئلة معينة وهكذا .

هذه الطرق من الاستجابة ، موضوعية كانت أم ذاتية ، كلها لفظية . وهناك استجابات أدائية . والاختبارات النفسية من نوع الأداء قد تكون عامة مثل اختبار المتاهات الذى وضعه پورتيوس Portius . وقد تكون خاصة فتكون مادة الاختبار الأدائى عينة من الشغل في أعمال معينة كاختبارات السائقين واختبارات عمال النسيج .

وهناك صيغ تخرج عن هذا ، مثل اختبارات اكتشاف السخافات . وفيها يطلب من المفحوص أن يقرر ما إذا كان هناك وجه للغرابة في أحد الصور أو القصص كما يسأل أن يحدد عن الشيء الذى يستبعد أن يكون معقولاً في الصورة أو القصة . ومن أمثلة الاختبارات التي تخرج عن هذه الأنواع ما يتبع في الطرق السيكوفيزيكية Psychophysique إذ تعرض على المفحوص

بطاقات على كل منها خط مستقيم ويطلب منه أن يقدر أطوالها وأن يقارن بينها .

رابعاً - التصنيف تبعاً لطريقة التصحيح :

بعض الاختبارات النفسية والتربوية يتضمن فقرات في كراسة التعليمات يعرف الفاحص الإجابة الصحيحة عنها دون مشقة من مفتاح التصحيح scoring key وأسهلها ، بالنسبة للفاحص اختبارات المسائل الحسابية ، ويتطلب البعض الآخر من الاختبارات أن يقوم الفاحص ، بعمليات حسابية لكي يفسر درجة الفرد الخام بناءً على ما تمده به قائمة المعايير . كما أن هناك اختبارات تسجل الإجابة عنها في ورقة منفصلة معدة لذلك تسمى استمارة الإجابة Answer sheet . وفي هذه الورقة مسافة تحت رقم كل سؤال يملؤها المفحوص ، إما بكتابة الجواب وإما بوضع علامة $\times$ على الحرف الدال على الجواب الصحيح من الإجابات المعطاة مع السؤال ، وقد يقوم المفحوص بملء أحد فراغين أمام أو تحت رقم السؤال أحدهما صحيح والآخر خطأ أو أحدهما نعم والآخر لا ، أو أحدهما موافق والآخر غير موافق . ومفتاح التصحيح هو الأداة التي يكشف بها الفاحص عن الإجابات الصحيحة أو نوع الإجابة . ويصمم هذا المفتاح بحيث تكون به ثقب إذا طبق المفتاح على ورقة الإجابة تطابق كل ثقب مع الإجابة الصحيحة عن كل سؤال ، فإذا لم يكن تحت الثقب في ورقة الإجابة علامة تدل إجابة المفحوص ، كانت الإجابة عن السؤال خاطئة أو غير مطلوبة . ويجمع عدد العلامات التي تظهر خلال الثقب نحصل على الدرجة الخام للمفحوص . وقد يكون للاختبار الواحد عدة مفاتيح ، كل مفتاح منها ينحصر عدداً من الأسئلة وضع لقياس أحد الجوانب . فإذا فرض أن الاختبار يقيس خمسة جوانب من الشخصية ، صممت له إذاً خمسة مفاتيح للتصحيح مفتاح لكل جانب . فالأسئلة رقم ١ ، ٥ ، ٩ ، ١٠ ، ١٨ . الخ مثلاً تقيس الانطواء .

(٨ - القياس)

ومفتاح الانطواء يشمل ثقباً تطابق مواضع إجابات هذه الأسئلة حسب أرقامها وتحتجب ما عداها من الأسئلة . ومجموع عدد الإجابات التي تظهر من الثقب هو الدرجة الدالة على الانطواء لهذا الفرد .

وهناك تطوير أكثر حداثة في تصحيح الاستجابات وفيه نلجأ إلى طرق آلية أو ماكينات . وتستخدم هذه الطريقة في الاختبارات الجمعية التي تطبق على نطاق واسع كما في اختبارات اختيار الجنود وتصنيفهم . وعلى النقيض من هذا تماماً نجد اختبارات يتطلب تصحيحها تقيماً كيفياً لنوع الأداء ودلالته ومصاحباته . وهذا التقييم يقوم به خبير في اختبار معين بذاته مثل اختبارى الرورشاخ واختبار H.T.P. وكل الاختبارات الإسقاطية من هذا النوع إذ يعتمد التصحيح والتفسير ، إلى حد كبير جداً ، على خبرة الإخصائى النفسى وتكوينه العلمى وذاتيته . وهى لهذا اختبارات أقل موضوعية مما سبق .

خامساً - التصنيف على أساس تطبيق النتائج :

وتبعاً لهذا الأساس من التصنيف تقسم الاختبارات النفسية إلى ، غير رسمية informal Tests واختبارات مقننة standardised T. والأولى اختبارات توضع عن برامج معينة كالتى يضعها المدرسون عن البرامج الدراسية والتى يضعها الإخصائيون المهنيون عن برامج تدريب مهنى معين . وهى اختبارات غير مقننة . أما الاختبارات المقننة ، وهى التى تهتمنا غالباً كإحصائيين نفسيين أو تربويين ، فقد جربت على مجموعات من الأفراد يمثلون نوع ومستوى ونسب الأفراد الذين ستجرى عليهم فيما بعد . وتفسر درجة الفحوص على أساس درجات العينة المستخدمة في التجريب أى عينة التقنين . وسيأتى المجال الذى نفضل فيه هذا الحديث .

سادسا - أسس أخرى للتصنيف :

لا نقصد هنا أن نحصر كل التصنيفات ولاحتي أن نسرد أسسها المختلفة ، ولكن لا بأس من أن نتحدث عن بعض تلك التصنيفات ولو بطريقة موجزة :

يفرق كرونباخ بين اختبارات القدرة Ability (وهي عنده اختبارات أقصى أداء ممكن Maximum Performance) وبين اختبارات الأداء المعتاد (وهي عنده اختبارات الخصائص العادية Habitual Characteristics) .
والأولى من نوع اختبارات القدرة كاختبارات الذكاء والاستعدادات الخاصة . أما اختبارات الأداء المعتاد فهي اختبارات الشخصية . واختبارات أقصى أداء تهدف إلى معرفة مستوى أداء الفرد في أفضل ظروف ممكنة إذ يبذل كل جهد ممكن في مثل هذه الاختبارات لضبط الدفع والاهتمام والظروف المحيطة والعوامل الأخرى بحيث تضمن أن الفرد سيبذل أقصى جهد في الاختبار . بينما يتوقع من الفرد ، في اختبارات الشخصية ، استجاباته المعتادة أو العادية . فاختبارات الشخصية لا تبحث عما يعتقد الفرد أنه أحسن حل ، أو الحل الذي يود أن يقوم به ، بل تبحث عن سلوكه العادي في الموقف المعين .

وكذلك التصنيف إلى اختبارات ذات بناء محدد structured tests واختبارات بناؤها غير محدد Unstructured T. وتراوح درجة البناء في الاختبارات النفسية بين مجرد تحديد معالم صورته تعرض على المفحوص ، من ناحية ، وبين موقف مصمم بدقة وعناية فائقة فيه عوامل مركبة ؛ ومواد ، وطرق ، وعقبات ، وأشخاص بل وانفعالات كما في الاختبارات الموقفية . والاختبارات الإسقاطية اختبارات بناؤها غير محدد ، واختبار T.A.T. عشرون بطاقة على كل منها صورة

غير واضحة وضوحا تاما : وهذه الدرجة من البناء أعلى مما نَجدها في اختبار الورشاخ إذ أن المادة فيه يقع خبر لم يقصد أنها تمثل شكلا ما ، وليس عدم تحديد بناء الاختبار بدقة عيبا في الاختبارات الإسقاطية بل إنه نوع من التصميم قصد منه إفساح الفرصة لذاتية الفرد وإسقاطاته وتنظيمه للموقف أو إدراكه له وفي هذا ما فيه من نواحي الإسقاط والتعبير الحر إلى آخر الجوانب اللاشعورية .

. والفرق الأساسي بين الاختبارات الإسقاطية والاختبارات المفصلة هو من حيث طبيعة المادة التي يحتويها الاختبار كمنبه يثير عمليات ووظائف نفسية معينة من ناحية ، والمنبه في الإسقاط إجمالي global من ناحية أخرى . فالمفروض فيها أنها تؤثر من حيث الكل لا من حيث تفاصيلها . وهناك فرق آخر بين النوعين ، هو أن الاختصاصي النفسي عندما يتناول نتائج تطبيق الاختبار الإسقاطي ينظر إلى الفرد ككل ، في حين أنه في الاختبارات المفصلة يؤكد على أجزاء من هذا الكل .

كما أن الاختبارات النفسية والتربوية تصنف على أساس الفلسفة التي قام عليها كل منها ، وكذلك على أساس نوع المواد التي تستخدمها ؛ فهناك فقرات لفظية ، وفقرات من أرقام ومسائل حسابية ، أو رموز ؛ وأخيرا هناك اختبارات تأخذ موادها من مواقف الحياة ؛

وهناك تفرقة بين الاختبارات النفسية على أساس وسط الاختبار medium ؛ إلى اختبارات ورقة وقلم واختبارات أداء ؛ بدأ هذا التقسيم في اختبارات الذكاء ثم انتقل إلى اختبارات الشخصية ، وفي اختبارات الورقة والقلم يعطى لكل مفحوص نسخة من الاختبار يجيب عنها بنفسه ، وترك للإجابة مسافات فارغة في كراسة الأسئلة أو في ورقة لإجابة منفصلة ؛ والمفحوص في الاختبارات الأدائية يقوم بمعالجة الأشياء

يدويا ، أو بمعالجة الصور ، أو المكعبات ، أو الأجهزة الميكانيكية ، أو قد يقوم بنشاط معقد يمثل بعض أوجه النشاط في مواقف الحياة اليومية . والاختبارات الأدائية فردية غالبا ، إذ يصعب توفير مجموعة من الأجهزة والمواد تكفي للإجراء الجمعي ، كما يصعب ضبط الموقف ويتكلف الكثير وتعطل القدرة ويحتمل الغش بالمحاكاة . هذا وإن كانت اختبارات الأداء تطبق أحيانا على مجموعات صغيرة من الأشخاص ، وهنا توضع حواجز على مائدة الاختبار كي لا يغش المفحوصون . وهذا الاحتياط لا لزوم له في حالة الاختبارات الحسية - الحركية المعقدة إذ يصعب الغش وكذا لا يتأثر المفحوصون ببعضهم البعض في الأداء .

وهناك تصنيف إلى اختبارات لغوية Language وغير لغوية Non Language . وفي الأخيرة لا تستخدم أى لغة كتابة أو كلاما لا في التعليمات أو الفقرات . وهي تستخدم في حالة الأميين والأجانب والصم والعجزة كما تستخدم في حالة الاختبارات التي تقيس السمات التي يفترض أنها موحدة بين مختلف الثقافات ويسمى هذا النوع من الاختبارات « الاختبارات عبر الحضارية Cross Cultural T. » . وقد تكون الاختبارات غير اللغوية من نوع الورقة والقلم بالصور ، والرسوم ، والرموز ، ورسوم المكعبات . وهنا يسجل المفحوص الإجابة برموز بسيطة والتعليمات تعطى بالإشارة والحركات وعرض الخرائط والرسوم . ولكن الاختبارات غير اللغوية في أغلب الأمر لا تستغنى عن اللغة ، فاللغة إن لم تستخدم في التعليمات أو في الفقرات استخدمت في الإجابة - فهي على أى حال مستخدمة وإن لم تكن عنصرا يحدد الإجابة أو يقف دون تعبير الفرد عن استجابات تمثل قدرته .

وهناك تقسيم بين الاختبارات على أساس المحتوى الذى يسود الاختبار المعين . وهذا المحتوى السائد قد يكون لفظى ، أو عددى ، أو مكانى و . . الخ . وفيها قد تستخدم الصور لقياس الفهم اللفظى كما فى اختبارات ما قبل المدرسة Preschool إذ ينطق الفاحص بكلمة أو عبارة أو جملة تعبر عن شيء صور فى الاختبار ويعلم المفحوص على الصورة التى تناسبه . وهذه الجمل التى يقولها الفاحص قد تصف الشيء أو تسرد أهميته أو تعرف وظائفه وعلى ذلك فاستجابة الفرد باختبار الصورة فى كل هذه الحالات تختلف بحسب نوع الفقرات والتعليمات . وهناك اعتراض على هذا النوع من الاختبارات هو أن المفحوصين يجرون نفس العمل بطرق مختلفة ، كما أوضحنا فى حالة الصور ، تعكس درجات مختلفة من نمو الوظائف النفسية المختلفة . ومع هذا ، ما دامت الإجابة واحدة ، يحصل المفحوصون على درجة واحدة مع أنهم اعتمدوا على عمليات عقلية بعضها معقد والآخر بسيط ، والنتيجة على الاختبار واحدة رغم هذه الاختلافات .

قائمة مراجع القسم الأول

- 1— Adkins, Dorothy C. The effects of Practice on intelligence test Scores. J. educ. Psy. 1937.
- 2— American Psychological association. Technical recommendations for Psychological tests and diagnostic techniques Washing. ton : Amer. Psychol- Assoc. 1954.
- 3— American Psychological Association. Ethical standards of Psychologists. Washington : Amer. Psychol. Assoc. 1953.
- 4 Anastasia, Anne. Psychological Testing N. Y. : Macmillan Co., 1959.
- 5— Anastasia, Anne. The Concept of Validity in the interpretation of test scores. Educ. Psych : Measurement, 1950.
- 6— Anastasia, Anne. Differential Psychology (3rd Ed.) N.Y. : Macmillan, 1958.
- 7— Baron, D. and Bernard, H.W. Evaluation Techniques for Classroom Teachers. Mc-Grow Hill Book Co. 1958.
- 8— Bean, K. Construction of Educational and Personnel. Mc. Graw-Hill Books Co. 1953.
- 9— Boring. E. G. A. History of Experimental Psychology. (Rev. Ed.) N.Y., Appleton—Century—Crops, 1950.
- 10— Cronbach, L. J. Essentials of Psychological - Testing. N.Y. : Harper 1949.
- 11— Crow, L. E. and Crow, A. Educational Psychology. American Book Co. (Revised ed.) 1958
- 12— Dixon, W. J. and Massey, F. J. Mc. Grow-Hill 1951.
- 13— Garrette, H. E. Statistics in Psychology and Education. (Third edition) Longmans, Green and Co. 1948.
- 14— Good, C. V. and Merkel, W. R. Dictionary of Education. (2nd edition) Mc. Graw-Hill 1959.

- 15— Goodenough Florence]. Mental testing. N.Y. Rinehart, 1949.
- 16— Greene, H.A. and Jorgenson, A.N. and Jerberich, J. R. Longmans, Green and Co. 1947.
- 17— Greene, E.B. Measurement of Human Behavior N.Y. Odyssey 1941.
- 18— Guilford, J.P. Fundamental Statistics in Psychology and Education. Mc Grow—Hill Book Co. 1942.
- 19— Hilderth, Gertrude H.A. Bibliography of Mental tests and rating (2 nd Ed) N.Y : Psychology Corp. 1939.
- 20— Hilderth, Gertrude H.A. Bibliography of mental tests and rating Scales, 1945 supplement N.Y. Psychology. Corp. 1946.
- 21— Johnson, V.O. Statistical Methods in Research. Prentice-Hall., Inc. 1949.
- 22— Mc. Nemar, Q. Psychological statistics. John Willy, Son. 1949.
- 23— Mursell, J.L. Psychological Testing. (2 nd edition) N.Y. Longmans, Green and Co. 1950.
- 24— Murphy, G. An historical introduction to modern Psych. Rev. Ed.) N.Y : Harcourt, Brace, 1949.
- 25— Nunnally, Jum C. Tests and Measurements Assessment and Prediction. Mc. Grow Hill Book Compay Inc. New York 1959.
- 26— Peatman, J.G. Descriptive and Sampling Statistics Harper Brothers 1947.
- 27— Peters, C.C. and Van Voorhis, W.R. Statistical Procedures and their Mathematical Basis Mc—Graw Hill Book Co. 1940.
- 28— Remmers, H.H. and Gage, N.L. and Runamel J.F. A Practical Introduction to Measurement and Evaluation. Harper Brothers, 1960.
- 29— Shaffer, L.F. information which should be provided by test publishers and testing agencies on the validity and

- use of their tests : Personality tests. Proe, 1949 invit
Conf. testing Service.
- 30— Skinner, C.E. Educational Psychology. Prentice Hall 1959.
- 31— Super, D. E. Appraising Vocational Fitness. Harper
Brothers N.Y. 1949.
- 32— Torndike, R.L. and Hagen, A. Measurement and Evalua-
tion in Psychology, and Eucation. John Willy, Sons Inc. 1957.
- 33— Travers, R. M. W. Eductional Measurement N.Y. Mac-
millan Co. 1955.
- 34— Walker, H. M. Elimentary statistical methods N.Y. Henry
Holf & Co. 1947.
- 35— Walker, H. M. and Lev, J. Statistical Inference. Teachers
Collage Columbia Univ., 1950.
- 36— Warren, H. C. Dictionary of Psychology Houghton Miffjin
Co., 1934.
- 37— Wesman, A.G. Faking Personality test scores in a simulated
employment situation. J. Appl. Psychol., 1952.
- ٣٩ - أحمد زكي صالح : علم النفس التربوى (الطبعة الخامسة)
القاهرة ، مكتبة النهضة ، ١٩٥٧
- ٤٠ - السيد محمد خيرى : الإحصاء فى البحوث النفسية والتربوية والاجتماعية
القاهرة ، دار الفكر العربى ، ١٩٥٦
- ٤١ - عطية محمود هنا : التوجيه التربوى والمهنى
القاهرة ، مكتبة النهضة المصرية ، ١٩٥٩
- ٤٢ - فؤاد البهي السيد : علم النفس الإحصائى وقياس العقل البشرى
القاهرة ، دار الفكر العربى ، ١٩٥٨
- ٤٣ - محمد عثمان نجاتى : علم النفس الصناعى
القاهرة ، مكتبة النهضة المصرية ، ١٩٦٠
- ٤٤ - نجيب اسكندر ولويس كامل ورشدى فام : الدراسة العلمية للسلوك الاجتماعى
القاهرة ، مؤسسة المطبوعات الحديثة ١٩٥٩

القِسمُ السَّانِي

بناء المقاييس النفسية والتربوية ومميزاتها

الفصل الرابع

بناء المقاييس

- أهمية الإلمام بخصائص المقاييس
- تخطيط عام لمشروع بناء المقاييس
- خطوات إنشاء المقياس

أهمية الإلمام بخصائص المقاييس :

١ - إن كل مشغل بمهنة يعرف الأدوات التي يستعين بها في طرق العمل . يعرف خصائص كل أداة ويستطيع أن يختار الأداة المناسبة في الوظيفة المناسبة . وهو يعرف ما هي الخصائص الرئيسية التي يجب أن تتوفر في الأداة ومن ثم يستطيع أن يقيمها وأن يفاضل بينها وغيرها . والمشغل بمهنة يعرف كيف يصون الأدوات التي يستخدمها ويعرف الخطوات المنظمة والاحتياطات الضرورية التي تجعله آمناً محافظاً على خلق المهنة في نفس الوقت . فالطبيب مثلاً يعرف الأداة التي يستعين بها في قياس درجة الحرارة ، ، ، أو ضغط الدم ، أو قياس اضطراب وظيفة عضو معين : وهو فضلاً عن هذا يعرف ، أو يجب أن يعرف ، أن بعض الأدوات يجب أن تكون من نوع معين عند استخدامها في معالجة فرد يتميز بخصائص معينة . وهذا هو الحال تماماً بالنسبة للاخصائي النفسي سواء كان في التربية أو الصناعة أو العيادة كما سرى .

٢ - ولكن الطبيب لا يعرف كيف يصمم أى أداة من هذه الأدوات التي يستعملها ، وهو لا يقوم فعلاً بهذا العمل ، وهذا لا ينتقص من قدره كطبيب ناجح أحياناً . فلا يقوم المشغل بمهنة إذن ، في العادة ، بتصميم أوثق الأدوات ارتباطاً بمهنته . ولكنه يعرف ، بالتحديد ، وظيفة كل أداة ، ومتى يجب أن يستخدمها ، ومتى يحسن أن يستخدم أداة من نوع معين ، وكيف يفاضل بين عدة أدوات تؤدي نفس الوظيفة . بل ويعرف ، فضلاً عن هذا ، ما هي المواصفات العامة التي راعاها من صمم هذه الأدوات . وأكثر من هذا فالمشغل بمهنة يستطيع أن يحدد مدى ثقته بالحكم الذي بناه مستنداً إلى أدوات معينة :

٣ - والمشتغل بمهنة أيضاً يعرف كيف يخطط وحدة العمل التي سيقوم بها . وقد يخطط الأسس العريضة العامة التي سيتبعها في حالات من نوع معين أو في حالة بذاتها . فهو ، قبل معالجته لموضوع معين ، يحدد الأسلوب الفني والطريقة المهنية التي ستجدي أكثر في التعامل مع عميل أو مشكلة من نوع معين . وهو إلى هذا يتوقع أن يصل إلى أحكام تتعلق بجوانب المشكلة موضوع الدراسة . فهو يعرف نوع التقييم الذي سيقوم به أو يصل إليه عند اتباع أسلوب معين واختبار أداة معينة . وهو في كل هذا ، عند ما يقوم بالعمل فعلاً ، يوفر ظروفاً معينة ويتحاشى أخرى ويتحكم في ثالثة : وباختصار فهو يعد نفسه ويعد أدواته ويعد مشكلته ، يعدها جميعاً للبدء في العمل . وهكذا الاختصاصي النفسي .

٤ - والمشتغل بمهنة يعرف ، بكل تأكيد ، الوظائف التي يخدمها واستخدامه لأداة معينة . فلكل أداة ، عنده ، وظائفها ، فما يقيس نسبة الرطوبة في الجو ، وما يقيس الضغط الجوي عند ارتفاع معين ، يقيسها فعلاً وبالتأكيد ولا يحتمل إطلاقاً أن يكون جهاز الضغط قد قاس كثافة المادة مثلاً : وهو في هذا يختلف عن الاختصاصي النفسي الذي غالباً ما تكون ثقته ، محددة بنسبة معينة ، في الوظائف التي تقيسها الأداة التي يستخدمها . بل ، وأحياناً ، يكتشف الاختصاصي النفسي أن الأداة التي أراد بها قياس ظاهرة ما قاست أخرى بجوارها أو قاست أخرى ولم تقس ما أراد .

٥ - والمشتغل بمهنة يكرر استخدام نفس الأدوات في مئات الحالات فتعطى له نتائج ثابتة أو متقاربة جداً . وهو في هذا يختلف عن الاختصاصي النفسي الذي لوكرر استخدام نفس الأدوات في نفس الوظائف لوجد اختلافاً في النتائج بين المرات المختلفة ، وإن كان هذا الاختلاف أحياناً كبيراً وأخرى مقبولا .

٦ - والمشتغل بمهنة يستطيع أن يحدد مدى تأكده من سلامة النتائج التي سيصل إليها . فالطبيب مثلاً قبل إجراء العملية المعينة قد يعلن أن نسبة الخطورة في العملية هي كذا في المائة ، وأن النجاح محتمل بنسبة معينة ، وأن المضاعفات ستتعلق بجوانب معينة ، وأن إجراء العملية والعود إلى الحالة الطبيعية سيستغرق مدة تتراوح في حدود معينة . وهو ، وإن كان قد حدد نسبة للثقة وفترة العلاج والجوانب التي ستأثر بالعملية إلا أنه حدد هذا كله بنسبة مثوية أو في مدى معين . ومع ذلك كثيراً ما تخرج النتائج الواقعة عن توقعاته . وهكذا الاختصائي النفسي ، تماماً ، يضع مستوى للثقة ، ويحدد جوهرية أو دلالة النتائج ، ويخطط البرنامج ، بل ويحدد الخطأ المحتمل والخطأ المعياري . وبرغم هذا قد تحدث نتائج واقعة تخرج عن هذه التوقعات . وهذا لا يعيبه فليس الضبط هنا أمراً سهلاً ميسوراً وليست الحياة النفسية أرقاماً وتقديرات ولا حتى تقييمات ، فهي كل متفاعل متكامل ، إلى جوار ثباتها النسبي وخضوعها لقوانين السلوك العامة ، لها مرونتها وتميزها الفردي . ونحن لا نجد هذا بالطبع في المادة التي يتعامل معها المهندس أو الكيميائي أو الطبيب فكلهم يتعاملون مع مواد ملموسة .

٧ - والطبيب ، لكي يصل إلى تشخيص صادق يستعين بمحكات أخرى ثبت صدقها أو محكات ترتبط تجريبياً بالوظائف التي يتناولها . فالصدق التلازمي عنده هو الاستعانة بفريق من الأطباء المبرزين في حالات معينة (كونسولتو) . أي أنه يستعين بأحكامهم لمعرفة مدى صدق حكمه . وصدق الوقائع الخارجية عنده هو الاستعانة بتأثير أو مضاعفات علاجات معينة فإذا تحسن العضو أو الوظيفة التي يعالجها كان حكمه صادقاً . وهكذا الاختصائي النفسي يستعين بفريق Team يتناول معه الجوانب المختلفة كل فيما تخصص فيه .

٨ - والاختصاصات النفسى لا يبنى حكماً ، أو هكذا يجب ، بناء على أدلة تبيينها عند استخدام أداة واحدة مهما بلغ ثباتها وصدقها وشمولها . فهو يستعين بأدوات مختلفة ، ويجمع بيانات عن عدة أوجه من الحياة النفسية . والبيانات والأدوات كل واحدة تدعم الأخرى وتكشف عن صدقها وثباتها ، ومثله فى ذلك مثل الطبيب تماماً ، إذ يستعين الأخير بالتحاليل المختلفة ، والكشف بالأشعة ، والقياسات المتعددة ، كل يبين صدق الآخر ويكشف عن ثباته .

٩ - والاختصاصات النفسى عند دراسة الفرد أو الجماعة ، يبدأ بتكوين فكرة كلية تقريبية ، ثم يشرع فى تحليل هذا الكل أو الجشطلت إلى عناصره ويدرس كلا منها . ثم يعيد بناء ما حله ليكون صورة كلية أدق من الأولى . فهو يطبق اختبارات المسح ، ليكون فكرة عامة عن الفرد أو الجماعة ، ثم يقيم ، ثم يضع علامات الاستفهام على نقاط الضعف والقوى فى التكوين النفسى لمن يدرسه . ويبدأ فى دراسة أكثر تفصيلاً لجوانب الإشكال . وبعد هذا يدرس ديناميات الحالة بطرق مثل الطرق الإسقاطية وأخيراً يكون صورة كلية متكاملة يعرف تفاصيلها وإبعادها ونسبها ، وماهى الزوايا التى ينظر منها إلى هذه الصورة وتلك الأبعاد .

١٠ - والطبيب والفلكى والمشتغل بالقياسات الألكترونية ، كل منهم يضع مستوى من الدقة إذا وصل إليه القياس كان قياساً صادقاً وثابتاً . وهم فى هذا يحددون ، إن تفصيلاً وإن ضمناً ، نسبة ثبات وصدق قياساتهم فإذا لم يتجاوز القياس حد الخطأ المعيارى هذا ، كان القياس والمقياس صالحين . هذا على حين أننا نجد أن الاختصاصات النفسى لم يتفقوا على مستويات يحددون بها صدق وثبات مقاييسهم .

١١ - عالم الطبيعة وعالم الكيمياء يجعلان مواداً يعرفان درجة نقائها . بل ويستطيع كل منهما أن يصل إلى عناصر نقية يقيس فاعليتها ، أما المشتغل بالقياس النفسى والتربوى فقد كان ، إلى حين قريب ، يقيس عوامل مستقلة ليبين علاقاتها بعوامل متغيرة فيجد أنه لا يقيس هذه ولا تلك بالتحديد ؛ وإنما يقيس عوامل متدخلة أو متوسطة ؛ وقد بدأت الطرق الفنية للوصول إلى عوامل ذات نقاء أكثر ، بدأت تتقدم حينئذ بفضل عدة مناهج علمية أهمها التحليل العاملى .

١٢ - وأخيراً نجد أن كل مشتغل بمهنة يعرف ، بدقة ، شروط إجراء عملياته وظروفها ولديه المعايير أى المستويات التى يرجع إليها عند الحكم ؛ فقد مرت الأدوات والطرق الفنية فى دور التجربة بعملية التقنين ،^١ وأصبح ، بالتالى ، استخدامه للأداة أو الطريقة قانونياً أو مشروعاً . والأخصائى النفسى بالمثل يجرب أدواته على عينة تسمى بعينة التقنين ، ويصل من ذلك إلى تحديد الشروط والظروف والمعايير والصدق والثبات ؛ وهكذا رأينا أن الأخصائى النفسى يحتاج إلى تحديد الوظيفة التى يقيسها بتحليلها ، وإلى تحديد الأغراض التى تخدمها أدوات معينة بتحليلها ، ثم يطابق بين هذه وتلك كى يختار الأدوات المناسبة . كما رأينا أنه يختار ويقيم ويحكم ويفاضل بين الأدوات وإن كان لا يقوم بتصميمها ، وهو فى هذا يتساوى مع غيره من الأخصائيين لا يلزم أن يكون مصمم الأداة التى يستخدمها . إذ أن تصميم الاختبارات ليس بالعمل السهل كما سئى . والمشتغل بالقياس النفسى عليه أن يضع تخطيطاً مستنداً إلى إعداد العلمى ، ومعمداً على دراسته للمشكلة ، ومتناسباً مع إمكانيات القياس لديه .

والأخصائى النفسى عليه ، كما رأينا أن يتأكد من صدق قياس الأداة لما زعمت أنها تقيسه . كما يلزمه أن يحدد مدى ثبات النتائج التى وصل

إليها . بل وعليه أن يحدد مدى صدق وثبات النتائج التي قد يصل إليها ونوع هذه النتائج ، وأن يهتم بنوع من المعلومات أكثر من غيره أهمية في تأثيره على ما يقيس ، فيحدد بذلك نوع المثيرات التي سيقدمها وطريقة تقديمها أى مسألة صياغتها في فقرات :

ورأينا أيضاً أن الحياة النفسية لا تلخص في صفحة نفسية تعبر عن مستويات كمية . وإن كان هذا أمر يعين الدراسة ، إلا أنه غير كاف إذ يجب أن يقوم بالتحليل الكيفي ويدرس الديناميات والعلاقات الوظيفية بين المتغيرات . وهو في دراسته هذه ، كغيره من المهنيين ، يعتمد على آراء غيره ممن يكونون معه فريقاً متكاملًا . ويلجأ إلى الإفادة من أدوات مختلفة كل يقيم جانباً أكثر مما يقيم جوانب أخرى . وهو في كل هذا يكون صورة كلية انطباعية ثم يحلل هذا الكل حتى يصل إلى تكوين صورة كلية أدق من الأولى . على أن هذه الصورة لا توضع بطريقة حتمية مؤكدة تماماً إذ يلزم اعتبار تدخل عوامل لم يستطع ضبطها . وفي النهاية يلزم الأخصائي النفسي أن يعرف حدوده وحدود الأدوات التي يستخدمها وظروف القياس المختلفة .

والأخصائي النفسي وإن كان لا يلزم أن يصمم الاختبارات بنفسه كما رأينا ، يحتاج لكي يعرف الوظائف والمعايير وشروط الإجراء وطريقة التصحيح ومعنى الدرجة ودلالاتها والتفسير الكيفي وثبات وصدق الأداة ودرجة ثقته فيها ، يحتاج إلى أن يعرف كيف صممت الأدوات التي يستخدمها . ونحن لا ننكر ، بالطبع ، على الأخصائي النفسي القيام بتصميم الاختبارات وإلا فن يصممها ؟ هذا برغم ما يكتنف هذا العمل الضخم من مشقات وما يلزمه من احتياطات وما يقتضيه من علم نظري وعملي . إلا أن الأخصائي النفسي قد يجد نفسه أمام مشكلة لم تعد تجدى

فيها اختبارات موجودة ، كما قد يجسد أن الاختبارات التي تقيس ما يريد قياسه قننت على عينات من نوع يخالف طبيعة المجتمع الذي يريد دراسته . وهنا يلزمه أن يقوم بتصميم اختبارات جديدة إن كان له ذلك .

تخطيط عام لمشروع بناء المقياس

أولاً - المشكلة :

برغم ما يكتنف إعداد المقياس من مشقة جمة إلا أن الأخصائي النفسي قد يجد نفسه إزاء موقف يلزمه بأن يقوم بهذا العمل أو أن يسهم فيه . فالأخصائي النفسي قد يقوم بعمليات الاختيار التربوي أو المهني ، أو التوجيه التربوي أو المهني ، أو الإرشاد النفسي ، أو عمل يدخل في نطاق الخدمات النفسية الفردية أو الجماعية ، وهو في كل هذا يحتاج إلى أداة موضوعية مقننة لتقدير الأفراد الذين يتعامل معهم بطبيعة عمله . وهو قد يجد تلك الأداة في مقاييس جيدة تناسب الوظيفة التي استخدمها ، إلا أن مثل هذه المقاييس ، لسعة انتشارها ، قد تكون فقراتها معروفة . وهنا يفقد الاختبار الموجود قيمته فلا يصبح مقياساً للقدرة أو السمة ، بل اختباراً للتذكر والاسترجاع . فيجد أن الأفراد يعرفون الإجابات الصحيحة غالباً ، ومن ثم لا يستطيع أن يحدد من هو أصلحهم .

والأخصائي النفسي قد يجد المقياس الذي ينشده ، وهو يتعرف عليه غالباً من اسمه ، ولكن هذا المقياس أحياناً يقيس شيئاً آخر : وأحياناً أخرى لا يمثل المقياس الوظيفة المراد قياسها فحسب ، بل وأيضاً وظائف أخرى قد تكون غير ذات دلالة بالنسبة لمهنة أو دراسة معينة ، وقد يلائم اختبار موجود بالفعل الغرض المنشود ، ولكنه كثيراً ما يكون قد جرب

وقن واشتقت معايير من جماعات تختلف في طبيعتها عن الجماعات المراد دراستها . ومن هنا يلزم إعداد اختبار يقن على الجماعات التي سيطبق عليها :

والأخصائى النفسى ، فى كثير من الأحيان ، يجد أن الاختبار الموجود يقيس وظيفة يتطلبها عمل معين . ولكنه لا يقيس كل أو أهم الوظائف التى يقتضيها النجاح فى هذا العمل . ولهذا فهو يقوم بإعداد اختبار أو سلسلة من الاختبارات تغطى هذه الوظائف معاً . وقد يحدث أن يجد الأخصائى اختباراً يقيس وظائف معينة ولكنه لا يرقى إلى تقييم الأفراد الذين وصلوا إلى مستوى علمى أو مهنى معين . كما أن الاختبار قد يقيس الوظيفة ولكنه يفعل هذا من وجهة نظر معينة . بينما يريد الأخصائى تقييم نفس الوظيفة من زاوية أخرى . وهنا يصمم اختباراً يصحح من وجهة نظر معينة .

وقد تكون هناك صعوبات عملية تحول دون استخدام اختبار موجود بالفعل . فالأخصائى النفسى محدود ، فى البرنامج الذى يستخدم فيه الاختبارات ، بتوقيت معين يخطط على أساسه الزمن اللازم لإجراء الاختبار . فإذا كان الاختبار الموجود يتطلب وقتاً يزيد على ما يسمح به التوقيت ، كان استخدام الاختبار هنا مستحيلاً ، ولزم إعداد اختبار يفى بالغرض فى حدود الزمن المحدد . كما أن بعض الاختبارات الموجودة تتطلب من القائم بها خبرة وإعداداً فنياً غير متوفر فيمن سيقومون بإجرائها فى مؤسسة أو مدرسة معينة . ولذا يلزم إعداد اختبار يسهل على الأخصائين إجراؤه . ثم إن الاختبار الموجود قد يتطلب فى إجرائه نفقات لا تحتملها ميزانية المؤسسة أو المدرسة التى يعمل بها الإخصائى ، وهنا يلزم إعداد اختبار أكثر اقتصاداً . وبالمثل قد تكون مواد الاختبار من نوع يتلف

عندما ينقل من جهة إلى أخرى أو يصعب نقلها بطريقة عملية ، وهنا يستعاض عنها بمواد يمكن نقلها بطريقة عملية .

وأخيراً فإن الأخصائى قد يجد الاختبار الملائم الذى لا يعيبه أى نقص مما سبق ذكره . ولكنه يعطى الاختبار لأخصائين مختلفين فيطبقونه على بعض الأفراد ويكتشف أن درجة الفرد الواحد تختلف باختلاف من قام بإجراء الاختبار عليه ، فالاختبار إذن ذاتى . وقد يقوم أخصائى واحد بإجراء الاختبار عدة مرات على نفس الفرد دون أن يحدث تغير جوهري في الفرد ، وبرغم هذا تختلف درجة الفرد من مرة إلى أخرى . وكأن الاختبار لا يعطى نتائج ثابتة . والعيان ، الذاتية وعدم ثبات النتائج ، يجعلان الاختبار الموجود أداة لا يمكن الاعتماد عليها . ومن ثم يصمم الأخصائى اختباراً يمكن الاعتماد عليه من هذه الناحية .

ولا يلزم في كل الأحوال التى سبق ذكرها ، أن يقوم الأخصائى بتأليف الاختبار من أوله إلى آخره ، فهو قد يكفى بتعديل اختبار موجود حتى يتلاءم مع هدف منشود . وأياً كان الجهد المبذول ، سواء في التأليف أو التعديل ، فهو جهد يهدف إلى إعداد اختبار جيد . والاختبار الجيد لا يوجه إليه نقد من الانتقادات السابقة . فهو اختبار يمثل وظيفة . وهى وظيفة نقصد قياسها لأنها تتعلق بالنجاح في عمل معين . وهو عمل يشمل جوانب في الفرد يمثلها الاختبار تمثيلاً صادقاً ويعطى عنها درجة . وهى درجة لها عندنا معايير ، تمكنا من تفسيرها . وهو تفسير ثابت لاختبار صادق فيما يريد قياسه ، وهو قياس عملي من حيث الوقت والتكاليف وسهولة النقل وتوفر من يستطيعون إجراؤه .

ثانيا - تحديد الأهداف :

عندما نتضح المشكلة بالطريقة السابقة يكون الاختصاصي قد اقترب من الوصول إلى تحديد أهداف إنشاء أو إعداد اختبار جديد : فيلاحظ الاختصاصي مثلاً أنه يجب أن يصمم مقياساً يعينه في اختيار أفضل المتقدمين للعمل أو الدراسة . ويكون الهدف هنا هو الاختيار وتكون المشكلة هنا هي أن عدد المتقدمين للاختيار أكثر من عدد الوظائف الشاغرة الجديدة . وقد يكون الهدف هو إعداد أدوات تساعد في التوجيه المهني أو التربوي وتكون المشكلة هنا هي أن أمام الفرد عدة فرص قد ينجح في أحدها ويتعين تحديد هذه الفرصة بالذات . وفي حالتى التوجيه المهني والتربوي والاختيار المهني والتربوي يهدف المقياس إلى مسح وتقييم الفرد .

وقد يهدف الاختصاصي النفسى إلى تحديد نقاط الضعف والقوى في أفراد يدرسه أى إلى تشخيصهم ويكون هدف الاختبار إذن هو التشخيص وتكون المشكلة هي التقييم الكيفى عن طريق استجابة المفحوص لموقف الاختبار . والاختصاصي قد يريد أن يتنبأ بالنجاح أو الفشل الذى سيحققه من يلتحق بمهنة أو دراسة . وقد يهدف الاختصاصي إلى معرفة ما يتوقع من تحصيل المفحوص لأهداف تالية بناءً على ما حصله من أهداف سابقة . وقد يريد الاختصاصي معرفة ما إذا كانت حالة مفحوص معين تنبئ بتخلفه من أعراض مرضية معينة بعد اتباع أسلوب تشخيصى وعلاجى معين . وفي كل هذه الحالات يكون هدف الاختبار هو التنبؤ . وتكون المشكلة هي معرفة مدى ارتباط السلوك الراهن بالسلوك المقبل أى معرفة دلالة نتيجة الاختبار على سلوك المفحوص مستقبلاً .

والأخصائي ، فضلاً عن كل هذا ، قد يريد تصنيف جماعة كبيرة إلى مجموعات متناسقة صغيرة تتلقى تدريباً مهنيّاً معيّنّاً أو منهجاً دراسياً متقدماً أو تلقى عناية أكبر في دراسة معينة . وهناك أهداف أخرى كثيرة قد يحدد الأخصائي نفسه بها في إنشاء المقياس كتحقيق العمل أو ترشيح من يستحقون الترقية أو العلاوة أو . . الخ . وأخيراً وليس آخراً قد يكون الهدف هو خدمة البحث العلمي في ميادين علم النفس والتربية :

ثالثاً - الإعداد للمشروع :

١ - بعد وضوح المشكلة وتحديد الهدف يقوم الأخصائي بعرض فكرة المشروع على الخبراء في القياس . ويناقش معهم أهمية المشكلة وقيمة الأهداف وما سيحققه الاختبار . محاولاً التحقق من صحة الخطوة أو الحصول على توجيهاتهم قبل المضي في المشروع .

٢ - يحصل الأخصائي على تصريح يطلب فيه المسؤولون من مختلف المدارس أو الهيئات تقديم العون للأخصائي لإجراء تجاربه ودراساته على أن لا يكون لذلك تأثير في سير العمل أو الدراسة .

٣ - يقوم الأخصائي بتقدير تكاليف مشروع الاختبار ويقدم تقريراً بذلك إلى المسؤولين إذا كان ينبغي إنفاقهم على المشروع أو الحصول على مساعدة مالية لتحقيق المشروع .

٤ - يعين المسؤولون هيئة أو فرداً يشرف على المشروع ، وهذا المشرف يجب أن يكون خبيراً وعالماً في الوقت نفسه . وهو غالباً ما يكون أستاذاً بالجامعة .

٥ - يقوم المشرف باختيار فريق من الإحصائيين يساعده في المشروع ويدرب هذا الفريق على عمليات القياس المختلفة شارحاً لهم طبيعة وأهداف المقياس الجديد :

- ٦- يضع المشرف والفريق معاً تخطيطاً بالحدود الزمنية للمشروع ويقلمونه للمسؤولين الذين يلائمونه مع اعتبارات أخرى هامة في نظرهم .
- ٧- يقوم المشرف والفريق بالاطلاع على الاختبارات ومراجع القياس الموجودة حتى يكونوا محيطين بالمشكلة ، ويناقشون الاختصائي فيها .
- ٨- يضع المشرف تخطيطاً لتوزيع أدوار وتحديد مسؤوليات أعضاء الفريق :

- ٩- يحدد المشرف والفريق الأعمال أو الدراسات التي سيطبق الاختبار عليها :
- ١٠- يجمع الفريق معلومات وملاحظات عن هذه الأعمال أو الدراسات :

- ١١- يقوم الفريق بتحليل هذه الأعمال أو الدراسات :

رابعا - تحليل المضمون :

نقصد بتحليل المضمون Content Analysis تحليل العمل Job Analysis إن كنا بصدد مهن أو تحليل برنامج الدراسة Curriculum Analysis إن كنا بصدد الدراسة : ولا نقصد هنا تعريف هذه المصطلحات التي سبق أن عرفناها ، بل معرفة ماذا تفيدنا فيه هذه الخطوة وما هي أهميتها . فلا شك في أن الاختبار ، لكي يكون مبنياً على بيانات واقعية ، يجب أن يؤسس بعد تحليل العمل . وتحليل العمل يعرفنا بالوظائف التي ترتبط بالنجاح في العمل ومدى أهمية كل وظيفة في هذا العمل . أما تحليل المنهج الدراسي فهو يعرفنا بأهداف التربية وأهداف تدريس منهج بذاته : وبناءً على هذه الأهداف يتبنى المنهج أو يتطلب وظائف معينة : هذه الوظائف يجب أن نحدد

٢٠١

في تحليل المنهج . وكأن تحليل المضمون ، سواء كان عملاً أو منهجاً دراسياً ، هو « دراسة علمية شاملة دقيقة تعتمد على عدة طرق ومصادر ، وتهدف إلى معرفة واجبات ومقتضيات المهنة أو الدراسة ومسؤولياتها ، وتحديد ما يجب توفره في القائم بها » .

- وهكذا نرى أنه يمكننا أن نتحدث في تحليل المضمون عن :
- (أ) تحليل المنهج الدراسي عند تصميم اختبارات التحصيل .
- (ب) تحليل العمل عند تصميم الاختبارات النفسية في المهن .

(أ) تحليل المنهج الدراسي : يتصفح المدرس الكتب المقررة ، ويراجع مذكرات الدروس التي ألقاها ، عندما يكون بصدد وضع امتحان نهاية العام . ولكنه إذا اعتمد على الكتب والمذكرات فقط لما تبين مدى صدق امتحانه . وهو قد يسأل نفسه هنا قائلاً : « أى هذه الموضوعات يصلح موضعاً للسؤال ؟ » . ولكنه إذا كان قد حدد أهدافاً للتربية بوجه عام وللمدرسة بوجه خاص ، بل وحدد أهداف تدريس منهج دراسي معين ، لأصبح في موقف أفضل . وهو هنا يسأل نفسه الأسئلة الآتية :

- ١- ما الذي أريد أن يفعله التلاميذ بالحقائق التي تعلموها ؟
 - ٢- ما الذي سيحتاجه التلميذ ، فيما بعد ، في الدراسة أو العمل ؟
 - ٣- ما هي العلاقات التي يجب أن يخرج بها التلميذ من المفاهيم التي شرحت له ؟
 - ٤- ما هي الآراء والأفكار التي يجب أن يضيفها التلميذ من واقع خبرته ؟
 - ٥- ما هي المهارات التي يجب أن يكون التلميذ قد تعلمها ؟
- ونعطي للموضوعات المتعددة وتفاصيلها أهمية نسبية في محتوى المنهج ،

وبالتالى فى محتوى الاختبار ، على أن يشملها الاختبار فى حدود الزمن المسموح به . وبدلا من وزن الدرجات حسب أهمية الأسئلة بالنسبة للأهداف ، يمكن وضع عدد من الأسئلة يتناسب مع جوهرية كل موضوع . ويقرر مصمم الاختبار أى نوع من الفقرات يناسب غرضه ؛ فقرات المقال ، أو التكميل ، أو المزاجية ، . . الخ . ويستطيع المصمم أن يقسم الاختبار إلى عدة جوانب أساسية وتحت كل جانب يحدد نوع الأسئلة الملائمة والعدد المقترح منها . على أن كل هذا التحديد مبدئى وغالباً ما يتغير تبعاً لما تكشف عنه البحوث ، ولكن هذا لا يقلل من قيمة عمل خطة واضحة .

(ب) تحليل العمل : هو الدراسة العلمية المنظمة للعمل لتحديدته ومعرفة طبيعته وعملياته وأهميته كل عملية منها ومعرفة ما هى الخصائص والشروط التى يجب أن تتوفر فىمن يقوم به ، وحصر الواجبات الخاصة بكل وظيفة . ويقول سوبر D. Super « يقدم تحليل العمل لمن يصمم الاختبار قائمة بالاستعدادات والسمات التى تعتبر هامة فى ذلك العمل . ولكن هذه القائمة تخضع لعينين خطيرين : وهما ذاتية الأدلة وعدم التأكد من أن عاملا Factor معيناً يميز هذا العمل عن غيره ، حتى لو برهن العامل على أهميته » :

وتحليل العمل فى خطواته النهائية يصل إلى بيانات تلخص فى قائمة تسمى وصف العمل Job Description تحدد العمليات المتضمنة فى هذا العمل وطبيعته وأهميته كل منها وكيف تُؤدى ، أى أنه يجب عن الأسئلة الثلاث الآتية :

١- ما هى العمليات والواجبات المتضمنة وما طبيعتها what

? the worker does

٢ - لماذا تُؤدى كل عملية منها وما هى الأهمية النسبية لكل عملية منها

? Why he does it

٣ - كيف تُؤدى هذه العمليات وما هى الطرق العامة المتبعة فى

كل منها ? how he does it

كان هذا فيما يتعلق بوصف العمل . أما فيما يتعلق بوصف العامل فان الجزء الخاص به يلخص فيما يسمى بمقتضيات العمل Job Requirements أو مواصفات العمل Job Specifications . وقائمة شروط العمل هذه تحدد الشروط والخصائص العقلية والبدنية والمؤهلات والخبرة الواجب توفرها فيمن يقوم بهذا العمل . أى أنها تجيب عن السؤال الآتى :

٤ - ما هى المهارة المتضمنة The Skill involved ?

وهكذا رأينا أن تحليل المحتوى ، والذي يشمل تحليل المنهج الدراسى كما يشمل تحليل العمل ، يمدنا من ناحية بخصائص العمل ومن أخرى بالخصائص التى يجب توفرها فيمن يقوم به . ومن ثم نستطيع أن نحدد : ما هى المجالات التى ستغطيها أسئلة الاختبار ؟ وما هى الأهمية النسبية لكل مجال منها ؟ وما نوع الأسئلة الذى يناسب كل مجال ؟ وما عدد هذه الأسئلة لكل ؟ فإذا حددنا هذا أمكننا أن نحدد طبيعة المجتمع الذى سيجرب عليه الاختبار ثم يقنن ثم يستخدم . وحددنا المعايير وطريقة تصحيح وتفسير كل سؤال .

خامسا - مواصفات الاختبار :

إلى هنا تكون صورة غامضة قد تكونت لدى مصمم الاختبار . صورة ناقصة تقريبية عما سيكون عليه الاختبار أخيرا . هذه الصورة تنمو وتتضح كلما تقدم المصمم فى تخطيطه ، ولكن معالمها لا تتضح إلا عند ما يبدأ فى إنشاء الاختبار فعلا . وقبل هذا الإنشاء على المصمم أن يحدد نفسه

كتابة بمواصفات الاختبار الذى ينوى إنشاؤه ، وهذا التحديد سيوضح له
ولمساعديه ما يحاول أن يعمل : وهذه المواصفات يجب أن تغطى
النقاط التالية :

١ - الغرض : وصف دقيق للغرض الرئيسى للاختبار والأغراض
الأخرى .

٢ - ماذا يقيس : وصف دقيق للخاصية أو الوظيفة التى يقصد
قياسها ، مع ذكر الظروف والعوامل التى قد تؤثر فى
نتائج القياس :

٣ - وصف الاختبار : مادة الاختبار ، عدد الأجزاء ونوع
وعدد الفقرات فى كل جزء ، تعليمات الإجراء ، طريقة
التصحيح :

٤ - العينة : تحديد الجماعات التى يجرى عليها الاختبار ، تحديد
حجم وطبيعة العينة التى تحسب منها خصائص الاختبار والمعايير :
٥ - منهج البحث : خطوات البحث ، الطرق الإحصائية لاختبار
الفقرات وترتيبها وحساب خصائص الاختبار والمعايير :

على أن كل هذه الخصائص تظل انسيابية فى المراحل الأولى من إنشاء
الاختبار ، بمعنى أنها تتطور وتتحدد أثناء عمليات التقنين التالية . ووصف
الاختبار إذا غطى النقاط السابقة كان ممثلاً لطريقة التفكير فى موضوع
الاختبار وقت إنشائه . وإن كان يجب أن يوضع بطريقة تقبل التعديل
والمراجعة على أساس الأفكار الحديثة المتعلقة بالموضوع عموماً
وفى ضوء البيانات الواقعية المشتقة من تجريب عينة من التعليمات والفقرات ،
أر الصورة الأولية للاختبار خصوصاً : وهكذا تنمو المواصفات
بنمو الاختبار .

ومواصفات الاختبار ، إذا وضعت هكذا ، تكون ذات قيمة فى

بلورة خطة الاختبار في ذهن مصممه . كما أن لها قيمة في توضيح الخطوة الظاهرة لمعاوني مصمم الاختبار ، بحيث يعملون فيه بأعلى كفاية فتتار وتوجه الاقتراحات الإيجابية وأوجه النقد الموضوعية . وتنحصر الفائدة الثالثة لتعيين خصائص الاختبار في كونها مرشداً يعين في تأليف صور مكافئة من الاختبار نفسه . فإذا صيغت عبارات واضحة عما يتضمنه الاختبار من حيث نوع ومحتوى وصعوبة الفقرات ، وإذا عينت العبارات الخطوات العملية لتأليف واختيار الفقرات ، تيسر إعداد صور تالية من الاختبار تكون مكافئة حقاً للصورة الأولى في الوظائف المقاسة والمدى المشمول ومستوى الصعوبة .

خطوات لإنشاء المقياس :

بعد أن تكون المشكلة قد اتضحت في ذهن واضعي الاختبار ، وبعد أن دفعتم الحاجة إلى إعدادده ، تعدد أمامهم الأهداف . وبعد أن تنضح المشكلة ويتحدد الهدف يقومون بالإعداد للمشروع ، وهنا يحللون العمل أو منهج الدراسة (حسب نوع الاختبار) ، ومن ثم يحددون المواصفات التي سيكون عليها الاختبار . وهنا يبدأ الفريق في الخطوات العملية لإنشاء الاختبار ، وهي تنحصر غالباً فيما يلي :

١ - اختيار نوع وحدات الاختبار :

٢ - اختيار نوع مواد الاختبار .

٣ - تأليف الفقرات و

٤ - تجربة الفقرات و

٥ - تحديد مستوى صعوبتها و

٦ - البحث عن المحكات و

- ٧ - جمع بيانات عن المحكات .
 - ٨ - تحديد درجات الأفراد على المحكات .
 - ٩ - مقارنة درجات الأفراد على الفقرات بدرجاتهم على المحكات .
 - ١٠ - تحديد طريقة التصحيح .
 - ١١ - تعيين أوزان الفقرات .
 - ١٢ - تحديد نوع الجماعات التي سيطبق عليها الاختبار .
 - ١٣ - تحديد طرق وعمليات الإجراء .
 - ١٤ - اختيار عينة التقنين .
 - ١٥ - إجراء الاختبار على عينة التقنين :
 - ١٦ - اشتقاق المعايير .
 - ١٧ - تحديد الدرجات الفاصلة .
 - ١٨ - حساب الثبات والصدق :
 - ١٩ - التأكد من صلاحية الاختبار في قياس ما أريد به على المجتمع الأصلي :
 - ٢٠ - نشر الاختبار وتحديثاته .
 - ٢١ - شروط إعداد الاختصاصيين الذين سيستخدمون الاختبار .
 - ٢٢ - متابعة إجراء بحوث على الاختبار ونشرها :
- وهكذا نرى أن أهم الخطوات العملية في إعداد الاختبار تنحصر في :
- (أ) إعداد الصورة الأولى للاختبار .
 - (ب) تجريبها .
 - (ج) تعديلها :
 - (د) الإجراء المبدئي .

- (هـ) عمل صورة أو صور مكافئة من نفس الاختبار .
 (و) التقنين .
 (ز) الصورة النهائية من الاختبار .
 (ح) نشر كراسة التعليقات ونماذج التصحيح والدراسات .
 لم نرد الاستطراد في كل نقطة من النقاط السابقة إذ ستكون موضعاً
 للبحث في فصول تالية هـ

الفصل الخامس

إعداد فقرات الاختبار

- مقدمة .
- ماذا يجب .
- كيف نصمم الفقرات .
- أنواع فقرات الاختبار .
- مقارنة الأنواع .
- قواعد عامة في إعداد الفقرات .
- التحليل الإحصائي للفقرات .
- تصحيح الفقرات .

مقدمة :

يتكون الاختبار عموماً من عدد من الفقرات : والفقرة item هي وحدة في الاختبار : هذه الوحدة قد تكون سؤالاً واحداً بسيطاً أو أكثر . كأن يوضع سؤال يحدد المطلوب في الإجابة عن عدة أسئلة تالية : والفقرة تسأل عن إحدى العمليات أو المتغيرات النفسية ، وتكون لها مادة من نوع معين : وتصصح الفقرات بطرق مختلفة .

وهكذا نرى أنه يجب :

- ١ - تحديد النوع المفضل من الفقرات .
 - ٢ - اختيار أنسب أنواع المواد :
 - ٣ - تحديد المستوى السائد أى ما هو النوع من المضمون الذى سيغلب على بقية الأنواع .
 - ٤ - تعيين أوزان الفقرات .
 - ٥ - اختيار أنسب طريقة للتصحيح .
- ونحن نجد أننا في هذه الأمور لسنا بصدد قواعد عامة جامدة تطبق في جميع الأحوال . فإن اختيار نوع الفقرات ، والمواد ، والمحتوى السائد ، والأوزان ، وطريقة التصحيح ، كل هذا يتحدد تبعاً لعدة اعتبارات أهمها :
- (أ) الوظيفة أو الوظائف التى يقيسها الاختبار .
 - (ب) الغرض الذى سيستخدم الاختبار فيه .
 - (ج) نوع المجتمع الأصلي المراد قياسه وخصائصه ومستواه في الوظائف المقاسة .
 - (د) إمكانيات الأخصائيين الذين سيصممون الاختبار ومن سيجرونه :
 - (هـ) التكوين العلمى لمصمم الاختبار وما يراه أساسياً :

(و) حدود الوقت والتكاليف .

(ز) مستوى الثقة الذى يراه مصمم الاختبار كافياً إذا توفر ، من حيث ثبات وصدق الاختبار إذ أنهما يتأثران بالفقرات كما سنرى .

وبالرغم من كل هذا فن المؤكد أن الاختبار ، باعتباره أداة موضوعية مقننة لقياس جانب أو جوانب من السلوك بطريقة ثابتة ، يجب أن يمثل الجوانب المراد قياسها . وهذا التمثيل لكى يكون كافياً يجب أن تتوفر فيه صفتى الكم والكيف . بمعنى أنه ما دام الاختبار عينة من السلوك ، فيجب أن يمثل أنواع ومستويات الجوانب المقاسة فيتضمن كل الأنواع والمستويات ويجب أن يمثلها من حيث أهمية كل نوع وكل مستوى فيتضمن فقرات يتناسب عددها مع أهمية الجانب الذى تمثله فى الوظيفة المقاسة . ولنفرض أننا بصدد عمل اختبار تحصيلى فى الحساب ، كيف نصمم فقراته ؟

أولاً : نحلل محتوى منهج الحساب الذى درسته الفرقة الدراسية التى نصمم الاختبار لقياس تحصيلها .

ثانياً : نحدد العمليات التى درست ولتكن الجمع ، الطرح ، الضرب ، والقسمة .

ثالثاً : نحدد نسبة كل عملية منها وأهميتها فى أهداف تدريس المادة كما درست بالفعل ولتكن هذه النسب هى : ٢٠٪ جمع ، ٢٠٪ طرح ، ٣٠٪ ضرب ، ٣٠٪ قسمة .

رابعاً : نمثل هذه الأنواع وتلك النسب فى الفقرات بإحدى الطريقتين الآتيتين :

(١) لنفرض أن عدد الأسئلة هو ٢٠ فيكون عدد الأسئلة موزعاً هكذا ، ٤ جمع ، ٤ طرح ، ٦ ضرب ، ٦ قسمة ولكل سؤال درجة واحدة .

(ب) يمكننا أن نوزع عدد الأسئلة بالتساوى بين العمليات الأربعة فيخصص كل من الجمع والطرح والضرب والقسمة ، ٤ فقرات أو أسئلة وهنا نعطي درجة واحدة عن الإجابة الصحيحة عن أى من أسئلة الجمع والطرح ، ودرجة ونصف عن الإجابة الصحيحة عن أى من أسئلة الضرب والقسمة .

وهكذا رأينا أننا حللنا المنهج الدراسى واخترنا مواد تناسبه وهى عمليات الحساب ، وحددنا ما تقيسه الفقرات وعددها وتوزيعها وأوزانها . حدث كل هذا ، بالطبع ، تبعاً لمستوى الفرقة الدراسية التى نقيسها فلا يصح مثلاً أن نعطي عمليات قسمة مطولة لمن تعلموا القسمة البسيطة فقط فإن مستوى الصعوبة يجب أن يتناسب مع طبيعة المجموعة .بقى شيء آخر وهو تحديد نوع الأسئلة ، فقد يفضل المدرس إعطاء أسئلة يطلب من التلاميذ الإجابة عنها . وقد يعطيهم السؤال والإجابة ويتطلب من التلميذ أن يضع علامة صح أو خطأ على الإجابة . وقد يضع المدرس السؤال متبوعاً بعدة إجابات يختار التلميذ إحداها . وقد يضع المدرس السؤال شكل عبارات يترك خلالها مسافات يجيب فيها التلميذ بالتكميل . وقد يضع عدداً من الأسئلة فى عمود ويضع إجابات هذه الأسئلة فى العمود المقابل ، بعد أن يخلط الإجابات خلطاً عشوائياً ، ثم يطلب من التلميذ أن يكتب أمام السؤال الحرف الموجود أمام الإجابة .

وهكذا نجد أن هناك أنواعاً متعددة من الفقرات يختلف واضعوا الاختبارات فى تفضيلهم لأحدها على الآخر . ولكن يجب أن نراعى أن تصاغ الفقرات بحسب الطريقة التى درس بها التلاميذ المادة على أنه من المؤكد أن الاختبار يكون أفضل عندما يستخدم عدداً كبيراً من الفقرات ، وعندما تتنوع المواد المستخدمة . فما لا شك فيه أن الاختبار الذى يشمل عدداً كبيراً من الفقرات هو اختبار أفضل عموماً من الاختبار القصير . وليس هذه المسألة أيضاً مطلقة فى كل الحالات ، فأحياناً لا يكون هناك

متسع من الوقت لإجراء اختبار طويل . كما أن الاختبار إذا كان طويلاً جداً تأثرت الإجابة عن أجزائه الأخيرة بعوامل التعب والملل ، وإذا كان على درجة كبيرة من القصر يزيد احتمال تدخل عوامل التذكر والحفظ ، على أنه بالرغم من عدم وجود قواعد جامدة في وضع فقرات الاختبار ، هناك ملاحظات يجب اعتبارها - سنتحدث قريباً عنها .

أنواع فقرات الاختبار

١- فقرات الصواب - الخطأ True-False items : وهى فقرات تقدم فيها للمفحوص جملة أو عبارة أو مشكلة أو مسألة ونطلب منه أن يحدد صحة الفقرة بوضع علامة ✓ أو خطأ (X) أو وضع خط تحت كلمة صحيح أو خطأ .

٢- فقرات التكميل Completion items وهى فقرات نطلب فيها من المفحوص أن يكمل عبارة أو مسألة بإضافة العنصر المناسب مثل : حاصل ضرب 6×5 هو ...

٣- فقرات الاختيار من عدة إجابات multiple-choice items وهنا نعطي المفحوص السؤال ونتبعه بعدة إجابات إحداها صحيحة وعليه أن يختار الإجابة التي يعتقد هو أنها صحيحة . مثل خارج قسمة $72 \div 6$ هو ١٣ ، ١١ ، ١٥ ، ١٢ ، ١٤ .

٤- فقرات المزاوجة matching items : وهنا نقدم للمفحوصين فقرات كل فقرة منها تتكون من عنصرين . ونضع الفقرات مصفوفة في عمودين ، العمود الأول يشمل أحد عنصرى كل فقرة والعمود الثانى يشمل العنصر الآخر : وكل عمود له ترقيم أو ترميز (ا ب ج د الخ) . وعلى

المفحوص أن يشير بسهم من الذئهر الأول إلى العنصر الذى يتسم الفقرة ،
أو أن يضع رمز العنصر المتمم أما العنصر الأول من الفقرة أمثلة :
المثال الأول : هذه عمليات ضرب والمطلوب منك أن تضع رمز الإجابة
أمام السؤال :

السؤال	الإجابة
١ - 6×5	(أ) ٢٨
٢ - 9×8	(ب) ٢٧
٣ - 4×7	(ج) ٤٨
٤ - 9×3	(د) ٣٠
٥ - 8×6	(هـ) ٧٢

المثال الثانى : هذه أسماء مخلوقات وأمامها أنواع الأماكن التى تعيش
فيها ، والمطلوب منك أن توصل المخلوقات بالأماكن بواسطة أسهم :

تعيش الدواجن فى	كباين
يعيش اللصوص فى	ثكنات
يعيش المصطافون فى	حظائر
يعيش الجنود فى	أوكار

لمثال الثالث : هذه أسماء بلاد وأمامها أسماء عواصم والمطلوب منك
وضع رقم العاصمة أمام اسم البلد الذى تقع فيه :

(أ) إنجلترا	١ - باريس
(ب) يوغوسلافيا	٢ - بودابست
(ج) إيطاليا	٣ - براغ

٤ - لندن	(د) تشيكوسلوفاكيا
٥ - بلغراد	(هـ) رومانيا
٦ - روما	(و) فرنسا
٧ - بوخارست	(ز) المجر

مقارنة أنواع الفقرات :

لا يمكن أن نقارن بين الأنواع المختلفة من الفقرات ، التي سبق ذكرها ، مقارنة مطلقة . فكل نوع استخدامه المختلف . كما أن لكل باحث وجهة نظر مختلفة في تفضيل أحد الأنواع . ولكن الأستاذ Edward Green عقد مقارنة بين هذه الأنواع . وقبل أن نعرض جدول المقارنة يجب أن نعرف بأسسها . لقد اعتمد على ملاحظاته وخبرته ، وعلى المزايا والصعوبات العملية لكل نوع ، وعلى نتائج البحوث التي أجريت في المفاضلة بين الأنواع الستة . وبناء على هذا أعطى لكل نوع تقديراً . وهذا التقدير يتراوح بين ٣ ، ١ . والرقم ١ يعبر عن أعلى رتبة بينما يعبر الرقم ٢ عن رتبة متوسطة ، والرقم ٣ يعبر عن أقل رتبة والرمز ؟ يدل على تشكك في قيمة نوع الفقرة في صفة معينة . وقد قارن بين الطرق الستة على أساس ١٠ مزايا . والطرق الستة هي الطرق الأربعة التي شرحناها مضافاً إليها طريقتين هما :

٥ - فقرات إعادة الترتيب rearrangement items : وفيها تقدم

للفحوص فقرات كلماتها غير مرتبة وعليه أن يرتبها لكي تصبح جملاً ذات معنى .

مثال : أعد كتابة العبارات التالية مرتباً كلماتها بحيث تصبح جملاً

ذات معنى :

١ - المطعم محمد العشاء تناول في أمس .

٢ - السرير الولد على الليل ينام في .

٣ - الدراجة المواصلات وسيلة السيارة من أسرع من وسائل :

٦ - فقرات المقال Essay items : نقدم للفحوصين أسئلة يجيبون عنها

في مقال ، أو نطلب مقالا بعنوان معين ، أو نعطي قضية أو مشكلة ونطلب رأي المفحوص فيها أو نسأله أن يناقشها .

مثال : ١ - ما رأيك في تحديد النسل .

٢ - اكتب مقالا في تزايد السكان .

٣ - تعدد الزوجات أمر موضع مناقشة ، ناقش .

ويحذر جرير من أخذ هذه المقارنة كما هي إذ أن الحكم يختلف من باحث إلى آخر ، ولكن لا بأس من سردها ما دامت مثالا يوضح رتبة كل من الأنواع الستة . على أساس نقاط عشر :

وجه المقارنة	س، خ،	التكامل	الاختبار من عدة إجابات	المزاوجة	إعادة الترتيب	المقال
سهولة الإعداد	٢	٢	٣	٢	٢	١
سهولة فهم التعليمات	١	١	١	٢	٢	١
قصر الزمن	١	٣	١	١	١	٢
صغر المسافة الخاصة بالفقرة	٢	٢	٢	١	٣	١
سهولة التصحيح ، لا عقاب جزائي	١	٢	١	١	١	٣
عدم تدخل عامل الصدقة في الإجابة	٣	١	١	١	١	٣
تعقد التفكير	٢	٢	١	١	١	١
وضوح السؤال الفقرة ليست مشكلة	٢	٣	١	١	١	٢
اعتماد الإجابة على الاسترجاع لا التعرف	٣	٢	٣	٢	٢	١
تحليل النتائج	٣	٢	١	٢	٢	١

الجدول (١) مقارنة بين أنواع من الفقرات

يوضح هذا الجدول أن الفقرات التي يجاب عنها بصحيح / أو خطأ (X) تفضل الأخرى من الفقرات في نواح ثلاث هي : سهولة الإجراء ، وسهولة التصحيح ، وقصر الزمن اللازم للإجابة عن كل فقرة . بينما وضعت في المرتبة الثانية من حيث : الاقتصاد في المساحة التي تشغلها الفقرة ، وسهولة إعدادها ، ووضوحها . ووضعت في المرتبة الثالثة من حيث : تدخل عامل الصدفة في الإجابة ، واعتمادها على الاسترجاع أكثر من اعتمادها على التعرف ومن حيث تحليل النتائج . وتحتاج الإجابة عن فقرات صحيح - خطأ إلى القليل من التفكير المركب عادة ، إلا أنها أحياناً تحتاج إلى تفكير مركب ومن ثم وضعت أمامها علامة استفهام لعدم تحديد رتبها . وفقرات الاختبار إذا وضعت بصيغة صحيح - خطأ تعتبر من أفضل أنواع الفقرات إذا توفرت فيها الشروط الآتية :

١ - عندما يكون أمام الباحث وقت قصير لإعداد الاختبار وإجرائه وتصحيحه .

٢ - عندما يقوم بتصحيح الاختبار كتبة مساعدون يجهلون موضوعه .

٣ - عندما يكون التفكير المركب واسترجاع المعلومات قليل الأهمية في معرفة مدى (البيانات المطلوبة للإجابة عن الفقرات) .

٤ - عندما لا تتدخل الصدفة في النجاح أو الفشل تدخلا خطيراً كما في اختبارات الشخصية .

كما يوضح الجدول أن فقرات الاختبار من نوع التكميل تمتاز عن غيرها من الأنواع : بسهولة الإجراء ، وخلوها من النجاح بالصدفة وكانت رتبها الثانية في سهولة الإعداد ، والمساحة التي تشغلها الفقرة ، وسهولة التصحيح ، واعتمادها على الاسترجاع أكثر من التعرف ، وتحليل النتائج : وكانت رتبها الثالثة في طول الزمن اللازم للإجابة ، ووضوح الأسئلة ، ولكن تركيب التفكير في الإجابة عنها غير محدد . ويعتقد بعض من يفضلون استخدام فقرات

التكميل عن فقرات صحيح - خطأ ، أن فقرات التكميل أقل عرضة للتخمين وهي لا تشجع التفكير السطحي :

وفقرات الاختبار إذا جمعت بين مزايا نوعي صحيح - خطأ ، والتكميل تصبح من نوع الاختيار من عدة إجابات multiple type . وهي تطبق وتصصح بنفس السرعة في تطبيق وتصحيح فقرات صحيح - خطأ ، وهي خالية من عامل الصدفة خلو فقرات التكميل . وفقرات الاختيار من عدة إجابات تصلح أكثر من غيرها في اختبارات الرياضيات والمهجاة والمفردات . وهي تستخدم في الاختبارات المقننة أكثر من غيرها من الأنواع :

أما فقرات المزاوجة Matching فهي تناسب الاختبارات التي تهتم بتصنيف الحقائق كما في اختبارات الكيمياء والطبيعة : وهي توفر فرصة متوسطة لتحليل النتائج وتحاشي تدخل عامل الصدفة . وقد أعطيت أعلى رتبة في ستة جوانب هي : صغر المساحة التي تشغلها الفقرة ، وسهولة التصحيح ، وخلوها من أثر الصدفة ، وتركيب التفكير ، ووضوح السؤال ، واعتمادها على الاسترجاع أكثر من التعرف : بينما أعطيت الرتبة الثانية في الجوانب الأربعة الأخرى .

وتتوفر في فقرات الاختبار التي تتطلب إعادة الترتيب نفس المزايا التي تتوفر في فقرات المزاوجة ، بل وتزيد عنها في دقة التصحيح . ولكن التعليقات في كل من فقرات المزاوجة وإعادة الترتيب لا يسهل فهمها بسهولة فهم تعليقات بعض الأنواع الأخرى ، ولكنها عموماً تناسب تلميذ الشهادة الإعدادية فما فوق ،

وفقرة المقال تتطلب توضيحاً شفوياً أو كتابياً أو عرضاً للحقائق . وإعدادها أسهل من إعداد أي نوع آخر من الفقرات . إلا أنها أصعب أنواع الفقرات جميعاً في تصحيحها ، ما دام هناك تبايناً كبيراً جداً في الإجابات

عن الفقرة الواحدة : وهى عادة تتطلب استرجاعاً وتنظيماً للأفكار أكثر من غيرها . ولهذا السبب تستخدم فقرات المقال على نطاق واسع فى قياس القدرة على حل المشاكل ، وهى أيضاً مستخدمة على أوسع نطاق فى تقدير التراكيب اللغوية . وهى الطريقة الوحيدة المستخدمة فى اختبار المفحوصين الذين يحتاجون فرصة للتعبير عن أسلوبهم الفردى وخيالهم . وفى مجال الأدب لا يمكن الاستغناء عن هذه الطريقة . وما دامت المقالات الطويلة أصعب فى تصحيحها من القصيرة ، فقد اتبع تقليد فى الاختبارات المدرسية هو تحديد الإجابة فى ١٥٠ كلمة أو بمساحة معينة . ولهذا التحديد ميزة دفع المفحوص إلى مناقشة إجابته بوضوح قبل أن يكتبها .

ولا يغيب عن أذهاننا ما لنوع فقرات الاختبارات وأسئلة الامتحانات من تأثير على استعداد التلاميذ لموقف الامتحان . فإن التلميذ يوجه مذاكرته بحيث يعد نفسه للإجابة عن نوع معين من الفقرات يشمله الاختبار غالباً . كما أن المدرسين يوجهون الطلبة فى شرحهم وإرشادهم إلى الكتب الإضافية ، وفى تمارينهم يراعون أنواع فقرات الاختبار . وهذا مثال يوضح هذا الرأى نسوقه من تجربة قام بها مايرز Meyers سنة ١٩٣٦ ، وتؤيدها دراسات أخرى كثيرة : طلب الباحث من مجموعات متكافئة من التلاميذ أن تعد كل مجموعة نفسها لأحد الأنواع التالية من الامتحانات : صحيح - خطأ ، الاختيار من عدة إجابات ، التكميل ، والمقال . وعند نهاية فترة المذاكرة أعطى لكل الجماعات كل الأنواع الأربعة من الامتحانات معاً . فوجد أن التلاميذ الذين أعدوا أنفسهم لاختبارات المقال حصلوا على درجات فى كل أنواع الاختبارات متوسطها أعلى من كل المجموعات الأخرى . وحصل من أعدوا أنفسهم لاختبارات التكميل على درجات تليها : بينما جاء ترتيب درجات من أعدوا أنفسهم لاختبارات صحيح

— خطأ ومن أعدوا أنفسهم لاختبارات الاختيار من عدة إجابات متساوياً وهو أسوأ من المجموعتين السابقتين .

وقد حاول بعض مصممي الاختبارات الجمع بين نوعين من الفقرات : فقام ماك كلاسكى Mc Clusky سنة ١٩٣٤ بالجمع بين فقرات صحيح — خطأ وفقرات التكميل . فطلب من التلاميذ عند وضع علامة (X) على الإجابة أن يكتبوا بجوارها الحقائق التي استندوا إليها في الحكم بخطأ الإجابة . جعلت هذه الطريقة زمن الأجراء والتصحيح أطول مما تستغرقه اختبارات صحيح — خطأ ، ولكنها أعطت صورة أوضح عن معرفة التلميذ بالموضوع .

وهناك اختبارات أخرى تجمع بين طريقتي الاختيار من عدة إجابات والتكميل . فقد أعطى كيرتس Curtis ١٩٢٨ تلاميذه فقرات وأمام كل فقرة أربعة اختيارات ومساحة فارغة يستطيع التلميذ أن يكتب فيها أدق إجابة يستطيعها . هذا الجمع أطال في زمن الأجراء والتصحيح كثيراً وخاصة في بعض الأسئلة بالذات وبالنسبة لبعض التلاميذ بالذات . ولها ميزة دفع التلاميذ إلى التفكير الأعمق ليصلوا إلى إجابة صحيحة ؛ وهذه الطريقة تفضل طريقة الاختيار من عدة إجابات . كما أنها أيضاً تمكن من مراجعة وتحسين فقرات الاختبار من نوع الاختيار من عدة إجابات ، لأن التلاميذ غالباً ما يكتبون إجابات خاطئة ، فإذا وضعت هذه الإجابات ضمن الاختيارات المرفقة بالفقرات ، فإنها تجعل فقرة الاختبار أكثر تمييزاً من ذي قبل . وهذا المصدر لمتغيرات الاختيارات هام لأنه يصعب عادة إيجاد ثلاثة اختيارات خاطئة تشبه الاختيار الصحيح للإجابة ، ذلك بغرض جعل تمييز الإجابة صعباً بدرجة كافية ، ذلك لأن الاختيارات الخاطئة يجب أن تكون قريبة من الإجابة الصحيحة بدرجة تجعل بعض المفحوصين يرونها مرادفة لها »

قواعد عامة في إعداد الفقرات :

- ١ - استعمل الصيغة الإيجابية ما أمكن : فمثلا تجنب الصياغة التالية :
(أنا لا أنام في الليل كثيراً) واستخدم (أنا أنام في الليل قليلا)
- ٢ - تجنب استعمال الكلمات التي تحمل أكثر من معنى واحد . مثل كلمة (ثقافة) وإذا استخدمتها فحدد ما تقصده بها .
- ٣ - اجعل السؤال يشتمل على فكرة واحدة مستقلة ما أمكن .
- ٤ - تجنب التعميمات المطلقة في الزمان أو المكان لأن هذا غالبا ما يكون خطأ والمفحوص يستغل ذلك في الإجابة . فلا تلجأ إلى عبارات مثل (دائماً) أو (في كل مكان)
- ٥ - تجنب العبارات الشديدة الإيجاز التي تجعل المفحوص شاكاً في معناها . فلا تضع عبارة مثل (إن معاني الكلمات يتوقف على ...) ، أو (إن مشروع الثورة هو ...) بل استخدم بدلا منها عبارة مثل (أهم فائدة للسد العلى هي ...) .
- ٦ - وضح في التعليمات الزمن المحدد للإجابة .
- ٧ - ضع تعليمات واضحة وتأكد من أنها ستلتصق .
- ٨ - بالنسبة لكل الفقرات ، قم بمراجعة صعوبتها إحصائياً وادرس أثر عوامل الصدفة والغموض وسجل النتائج على بطاقة منفصلة بالنسبة لكل فقرة . هذا سيساعدك على عمل سجل للفقرات التي تعرف صعوبتها وقيمتها :

وهناك طريقتان ناجحتان نستخدمهما في الكشف عن الفقرات الغامضة هما : (أ) ملاحظة السلوك (ب) والتحليل الإحصائي . ولطرق الملاحظة قيمة كبيرة في تقييم كل من اختبارات الأداء والاختبارات اللفظية . فقد أمكن الكشف عن غموض الدرجات بملاحظ المفحوصين المختلفين أثناء أدائهم

للاختبار . ففي اختبارات التجميع Assembly واختبارات المعالجة اليدوية manipulation ، قد نجد الأفراد المختلفين يقومون بالعمل المطلوب بطرق عدة ومع ذلك يحصاؤون على درجات واحدة . والدرجة الموحدة هنا تمثل مهارات مختلفة ، وهكذا لا يمكن مقارنة الأفراد المختلفين على أساس درجاتهم فالدرجات هنا غامضة الدلالة عن المستوى والوظيفة .

وأفضل الطرق ، في الحالة السابقة ، ليست الدرجات الرقمية ، ولكنها وصف طريقة قيام الأفراد المختلفين بالعمل الواحد . وكذلك لا تجدى الدرجة الدالة وحسب على الزمن المستغرق في الأداء . ولهذا يحسن أن نحدد ونرتب موقف الاختبار أو تعليماته بحيث يقوم كل المفحوصين بخطوات واحدة بطريقة موحدة .

وفي الاختبارات اللفظية ، يسهل ملاحظة غموض التفسير . وأول طرق هذه الملاحظة ، أن يقوم الباحث بنفسه بالإجابة عن الاختبار ويسجل خبراته المحيرة . كما قد يسأل بعض الخبراء أن يخبروه كيف يفسرون كل فقرة . والتحليل الذاتي ، من هذا النوع ، عملي واقتصادي جداً . وفي تقييم الاختبارات الموضوعية التي يصممها المدرس ، يجب ، إذا أمكن ، أن يترك المدرس الاختبار أسبوعاً أو أكثر بين الإنشاء والتقييم . وهذه الفترة الزمنية تسمح للمدرس بأن يعد جزئياً عن حالته النفسية التي تأثر بها أثناء الإنشاء أو تحيز لها والتي حددت صياغة واختيار فقرات معينة ، وهذه الفترة تمكنه من أن ينظر للأمر من زاوية أخرى . وبهذا قد يصل الباحث إلى نتائج تجعله يعدل من وجهة نظره . وبعد القيام بهذا التحليل الذاتي ، وبعد تطبيق الاختبار على الجماعة المقصودة ، يجب أن يقوم الباحث بالتحليل الإحصائي .

التحليل الإحصائي لل فقرات

ثبت من الكثير من التجارب أن فقرات الاختبار التي تتطلب إجابة مختصرة ، إذا أردنا بها أن تكون فقرات جيدة ، يصعب تصميمها . وقد يكون هدف الباحث هنا أن يبتعد عن التعليقات الكثيرة المعقدة وأن ينشئ اختباراً صادقاً يستطيع الكشف عن مدى معلومات المفحوصين أو مهارتهم . ولكن النتائج غالباً تأتي بعكس هذا . فقد وجد وود Wood سنة ١٩٢٧ أن حوالي ٣٠ ٪ من الفقرات المستخدمة في امتحان التلاميذ للصغار كانت غير فعالة في تمييز المستوى المنخفض من المستوى الأفضل . كما وجد Anderson أن ٣٣ ٪ من فقرات الاختبارات في ميدان القياس التربوي كانت أكفأ من الاختبارات الكلية ، التي تشمل هذه الفقرات ، في تقدير التحصيل أثناء الدراسة . بل لقد كانت بعض الفقرات لا تميز حتى بين الطلبة الممتازين جداً والمتخلفين جداً . والآن نناقش أسباب هذه النتائج المذهلة . على أن هم سبين لها هما (١) استخدام فقرات غير صالحة (ب) وتدخل عوامل الصدفة .

(١) الكشف عن الفقرات غير الصالحة :

بعد تطبيق الاختبار على جماعة التقنين ، تقيم كل فقرة بمعرفة قدرتها على التمييز بين المفحوصين تمييزاً يتفق مع تقسيمهم إلى فئات بحسب أحد أو بعض المحركات التي تقدر نفس الوظيفة التي يقيسها الاختبار . وهذا المحرك قد يكون الدرجة الكلية على الاختبار نفسه أو أى درجة أو تقدير آخر نعتبره دالاً . ويوضح الجدول (٢) تقدير السبعة فقرات تتطلب الإجابة عن كل منها وضح علامة / أو X . والجدول عن E.B.Green .

والمقارنة فيه بطريقتين تهدفان إلى الكشف عن مدى اتفاق النتيجة أو الدرجة عن فقرة مفردة مع مقياس المحك ، والمحك هنا هو الدرجة الكلية عن الاختبار . فقسم الأفراد إلى مجموعات ثلاثة بحسب الدرجة الكلية عن الاختبار ثم حسب النسب المئوية ، في كل قسم ، التي أجابت بصحة عن كل فقرة : والطريقة الثانية هي إيجاد معامل الارتباط بين نجاح الأفراد في الفقرة وبين نجاحهم في الاختبار ككل :

جدول (٢) تقييم فقرات صحيح - خطأ

النسبة المئوية الناجحة ، معاملات الارتباط

الفقرات	ا	ب	ج	د	هـ	و	ز
الثلث الأعلى	%٩٥	%٥٧	%٩٦	%٦٥	%٣٣	%٩٥	%٨٥
الثلث المتوسط	%٦٥	%٥٤	%٩٤	%٥٧	%٥٤	%٩٢	%٥٤
الثلث الأدنى	%٥٠	%٥٩	%٩٧	%٦٣	%٦٨	%٦٤	%٥٦
معامل الارتباط مع الدرجة الكلية	٠.٧٤	٠.١٤	٠.٠٧	٠.٠٦	٠.٣٢	٠.٣٠	٠.٣٢

فبإحدى الطريقتين قسم الأفراد كلهم إلى ثلاثة مجموعات : الثلث الأول هو الأعلى ، والثاني هو المتوسط والثالث هو الأدنى ، على أساس درجاتهم في الاختبار ككل . ثم حسبت النسبة المئوية للنجاحين في كل فقرة من بين أفراد كل مجموعة من المجموعات الثلاث . والطريقة الثانية التي استخدمت ، ونتائجها مدونة في الصف الأفقي الأخير ، هي إيجاد معاملات الارتباط بين النجاح في كل فقرة وبين النجاح في الاختبار ككل . فالفقرة (١) مثلاً نجح فيها

(٩٥٪) من الأفراد الذين حصلوا على أعلى الدرجات (وهم ثلث عدد المفحوصين كلهم) ، ونجح فيها (٦٠٪) من الثلث المتوسط و (٥٠٪) من الثلث الأدنى ، وكان معامل ارتباطها بالدرجة الكلية على الاختبار هو (٧٤ر) . مثل هذه الفقرة فقرة جيدة بمعنى أنها ميزت الأفراد إلى فئات تتفق كثيراً مع تميزهم إلى فئات بحسب درجاتهم على الاختبار ككل ؛ والاختبار الذي يتكون من مثل هذه الفقرة اختبار يتمتع باتساق داخلي عالي

High Internal consistency,

والفقرة (ب) نجح فيها أفراد الأقسام الثلاثة نجاحاً متساوياً ، فهي إذن لا تميز بين من حصلوا على درجات عالية في الاختبار ومن حصلوا على درجات متوسطة فيه ولا من حصلوا على درجات منخفضة ، كما أن معامل ارتباطها بالنجاح في الاختبار ككل منخفض مما يدل على أن العلاقة عشوائية . وهذه الفقرة في نفس الوقت ، كما يبدو من الأرقام ، صعبة على الأفراد عموماً .

أما الفقرة (جـ) فقد نجح فيها أفراد المجموعات الثلاثة بنفس النسبة تقريباً ومعامل ارتباطها قريب من الصفر . كما أنها سهلة جداً ولا تؤثر على المركز النسبي للأشخاص المفحوصين .

في حين أن الفقرة (د) وكذا الفقرة (هـ) أعطت نتائج غامضة . فالأرقام ، فيما يتعلق بالفقرة (د) تدل على أن المجموعات الثلاثة نجحت نفس النجاح ، وإن هذا النجاح فوق الصدفة (٦٣٪) فهو فوق (٥٠٪) ؛ وكان معامل ارتباطها مع الدرجة الكلية على الاختبار سالباً أي أنها لم تميز بين من حصلوا على درجات مرتفعة في الاختبار ككل وبين من حصلوا على درجات متوسطة أو منخفضة ؛ كذلك نجد الفقرة (هـ) غامضة والدلالة وغير مميزة ، فقد نجح فيها (٣٣٪) من أفراد أعلى مجموعة في الاختبار بينما نجح فيها (٥٤٪) من أفراد المجموعة المتوسطة

فى الاختبار . وكان معامل ارتباطها بالدرجة على الاختبار ككل .
(٣٢ ر ٠) ويمكن أن نفترض الفروض الآتية لتفسير نتائج الفقرتين
(د) و (هـ) :

١ - أن أفراد المجموعة العليا فى الاختبار فسروا الفقرة (د) تفسيراً
يختلف عن تفسير أفراد المجموعة الدنيا ، فيحتمل ألا يكونوا قد فهموا .
المطلوب فى الفقرة أو معناها .

٢ - نفس الفرض السابق محتمل فى الفقرة (هـ) .

٣ - أن الفقرتين (د) ، (هـ) تقيسان مهارة ترتبط ارتباطاً سالباً
بالمهارة التى يقيسها الاختبار ككل .

أما الفقرتين (و) ، (ز) فقد نجحنا فى التمييز بين المفحوصين نجاحاً
جزئياً . فالفقرة (و) ميزت بين الثلث الأدنى والثلث المتوسط بينما لم
تميز بين الثلث المتوسط والثلث الأعلى . أما الفقرة (ز) فقد ميزت بين
الثلث الأعلى والثلث المتوسط ، بينما لم تميز بين الثلث المتوسط والثلث
الأدنى . ولكن الفقرتين ارتبطتا بالحك ارتباطاً موجباً ولكنه منخفض .
والاختبار الذى يشمل فقرات من مثل الفقرة (و) والفقرة (ز) اختبار
مقبول ، وإن كان غير مميز مثل الاختبارات التى تشمل فقرات من نوع
الفقرة (ا) .

هذا النوع من التحليل الإحصائى لقدرة الفقرات على التمييز ولصعوبتها
تحليل سهل وهذا سبب شيوعه . وهو يبين متى تؤثر الصدفة والغموض
فى الدرجات .

النجاح بالصدفة :

١ - فروض نظرية .

فى الاختبارات التى تتطلب من المفحوص أن يختار إجابة من إجابتين

أو من عدة إجابات ترفق بالفقرة ، تكون هناك فرصاً لنجاح من لا يعرف الإجابة . فقد يقوم من يجهل الإجابة بوضع علامات أو إجابات مختاراً الإجابة الأولى عن كل الفقرات ، وهنا قد يحدث أن تكون الإجابة الصحيحة عن الفقرة الرابعة مثلاً هي الإجابة الأولى . وبالمثل قد يقوم بعض المفحوصين باختيار الإجابات بطريقة عشوائية أو بحسب أى ترتيب دون أن يكونوا عارفين ما هي الإجابة الصحيحة .

فإذا أجاب كل المفحوصين عن كل الفقرات في اختبار من نوع صحيح - خطأ أو اختبار يتكون من فقرات تكون الإجابة عنها باختيار إحدى الإجابات المدونة بعد كل فقرة ، لا يسهل تصحيح أثر الصدفة في إنجاح البعض تصحيحاً يعدل المكانة النسبية لهذا البعض . ولهذا السبب ، يطلب بعض الفاحصين من كل المفحوصين أن يخمنوا الإجابة عن الفقرات التي لا يعرفونها . ومن الناحية الأخرى ، قد يحذف الأفراد الحريصون الفقرات التي يشكون في معرفة إجابتها ، ولكن تصحيح أثر الصدفة في النجاح عادة ما يغير مراكزهم النسبية في الجماعة . وفي الكثير من الاختبارات المقتنة ، يطلب من المفحوصين ألا يخمنوا . ولكن مهما كانت التعليمات حازمة ، يحتمل دائماً أن يجيب بعض الأشخاص عن بعض الفقرات بالتخمين .

وعندما يكون على المفحوص أن يختار إحدى إجابتين ، كما في فقرات صحيح - خطأ ، وتكون الفقرات الصحيحة مساوية في عددها للفقرات الخاطئة ، يحتمل أن يخمن المفحوص متأثراً بالصدفة وحدها ، الإجابات الصحيحة عن هذه الفقرات . مختاراً الإجابة صحيح / في نصف الفقرات والإجابة خطأ (X) في النصف الآخر . فإذا افترضنا أن الفقرات التي يخطئ فيها المفحوص ضمن الفقرات التي تشكك في التخمين فيها ،

ربا فراض أنه تشكك في عدد من الإجابات الصحيحة يساوى عدد الإجابات الخاطئة التي تشكك فيها ، نستطيع هنا أن نستبعد أثر النجاح بالصدفة بطرح الإجابات الخاطئة من الإجابات الصحيحة فتكون درجة الفرد هي باقى عملية الطرح .

وعندما تكون الإجابات التابعة لكل فقرة هي ثلاثة اختيارات ، يكون متوسط النجاح بالصدفة في عدد من الفقرات هو ثلث هذا العدد ، وتكون الدرجة المصححة هي عدد الإجابات الصحيحة ناقصا نصف عدد الإجابات الخاطئة . وعندما يكون على المفحوص أن يختار إجابة من أربعة في كل فقرة تكون درجته المصححة عن الاختبار هي عدد الإجابات الصحيحة ناقصا ربع عدد الإجابات الخاطئة . وتكون المعادلة العامة لتصحيح الدرجة عن الاختبار ككل في حالة الإجابة بالاختيار من عدة احتمالات هي :

الدرجة المصححة = مجموع الإجابات الصحيحة - $\frac{\text{مجموع الإجابات الخاطئة}}{\text{عدد الاختيارات في كل فقرة} - 1}$

$$\text{أى أن د = ج - } \frac{\text{خ}}{\text{ن} - 1}$$

وهذا التصحيح ليس فعالا دائماً ، بالدرجة التي نفترضها فيه ، لأن إفتراض أن كل الإجابات الخاطئة أتت بمحض التخمين إفتراض لا يصح وخاصة في نسبة كبيرة من الحالات . بل وأكثر من هذا ففي فقرات الاختيار من عدة إجابات لا تكون الإجابات التابعة للفقرة عادة متكافئة الصعوبة - وهو الافتراض النظرى في تطبيق معادلة التصحيح . والشخص المحظوظ يزيد عن متوسط النجاح بالصدفة ويحصل على درجات لا يستحقها لأنه خمن الإجابة ، حتى لو طبقنا عليه معادلة التصحيح ، والشخص غير المحظوظ لا يحصل أحياناً على الدرجة التي يستحقها . ولكن الحالة الأخيرة نادرة لأن تخمين الفرد مهما كان صحيحاً لا يزيد عن الصدقة (أى لا يحصل الفرد عن أكثر من نصف الدرجة الكلية) غالباً . ولكن المفحوص عادة

ما يجد في المعلومات والأدلة التي يقرأها في الاختبار مرشحات توجهه في اختبار الإجابة ، وإن كانت هذه المرشحات لا تكن دائماً للاختيار الصحيح .
والباحث يستطيع أن يقرر ما إذا كان سيصحح أو لا يصحح أثر النجاح بالصدفة وكذا أثر الفشل بافتراض أن لها أثراً . والباحث هنا يستند إلى الأدلة العملية التي جمعها من إجراء الاختبار وإلى الأغراض التي ينوي استخدام النتائج في خدمتها . وإليك مثال يوضح هذا :

٢ - نتائج تصحيح أثر الصدفة :

قام رتش Ruch واستلرد Stoddoed سنة ١٩٢٥ ببحث صحافي فيه نتائج تدخل الصدفة . وقد استخدموا ٥ أنواع من الاختبارات التي تقيس التحصيل في مادة التاريخ الأمريكي . وهي :

فقرات إسترجاع : بدأت الثورة الأمريكية في سنة . . . :

الاختيار من خمس إجابات : بدأت الثورة الأمريكية سنة :

١٧٦٢ ١٧٧٥ ١٧٨٣ ١٧٨٩ ١٨١٢

٣ - الاختيار من ثلاثة إجابات :

بدأت الثورة الأمريكية سنة ١٧٩٢ ، ١٧٧٥ ، ١٧٨٩ .

٤ - الاختيار من إجابتين :

بدأت الثورة الأمريكية سنة ١٧٦٢ ١٧٧٥ :

٥ - صحيح - خطأ :

بدأت الثورة الأمريكية سنة ١٧٧٥ صحيح خطأ .

طبقت صورتان من كل نوع ، صورة حوت ١٠٠ فقرة ، على مجموعات ضخمة من تلاميذ المدارس الثانوية : فكانت فقرة الاسترجاع تعطى أولاً ، وتليها فقرة من أحد الأنواع الأخرى . وقد ظهر أن الإجابات

الصحيحة في أى نوع آخر ، على نفس الفقرة ، بلغت ضعف الإجابات الصحيحة عن فقرة الاسترجاع . وعند التصحيح النظرى للنجاح بالصدفة إقترَب متوسط الدرجة في فقرة صحيح - خطأ من متوسط الدرجة في فقرة الإسترجاع . ولكن ظهر أن النجاح بالصدفة نظريا يختلف كثيراً عن النجاح بالصدفة فعلا . كما ظهر أن الفقرات من نوع الاختيار من خمسة إجابات ارتفعت فيها درجة النجاح بالصدفة فعلا عن درجة النجاح بالصدفة نظريا . وربما ترجع هذه الزيادة إلى أن الفقرة التى حوت خمسة متغيرات قدمت معلومات إضافية هامة على هديها استطاع المفحوصون أن يختاروا الإجابة الصحيحة عنها .

كانت معاملات الثبات هى معاملات الارتباط بين نتيجة إجراء الصورة الأولى ثم الإعادة بالصورة المكافئة الثانية . وقد بلغت معاملات ثبات الدرجات التى صححت من أثر الصدفة مقادير تساوى معاملات ثبات الدرجات التى لم تصحح من أثر الصدفة .

٣ - نتائج طلب عدم التخمين من المفحوصين :

قام رتش وستودارد سنة ١٩٢٧ ببحث مشابه طلبا فيه من نصف عدد التلاميذ المفحوصين أن يخمنوا الإجابة التى لا يقطعون فيها بالتأكيد ، وطلبا من النصف الآخر ألا يخمنوا عند الشك . وعدد المفحوصين هو ٢٤٥٣ تلميذاً من أربعة فرق دراسية إعدادية وثانوية . وقد إرتفعت معاملات الارتباط بين الدرجات الخام فى أى من الصورتين المتكافئتين لدى التلاميذ الذين طلب منهم ألا يخمنوا وكان الارتفاع أكثر بقليل منه فى حالة التلاميذ الذين طلب إليهم أن يخمنوا . وهكذا تدل النتائج على أنه يحسن أن نطلب من المفحوصين ألا يخمنوا . وقد وضحت التجربة حقيقة هامة هى أن بعض المفحوصين لا يخمنوا الإجابة حتى لو طلب منهم أن يخمنوا ، كما أن بعض المفحوصين يخمنون الإجابة حتى لو طلب منهم عدم التخمين : وهذا يتضح

من حقيقة أن المفحوصين الذين خمنوا في الصورة الأولى لم يخمنوا في الصورة الثانية في نفس الفقرات ولا بنفس الدرجة ولهذا كان معامل الارتباط بين الصورتين منخفضاً نسبياً أى أن معامل الثبات كان منخفضاً نسبياً . فلو كان هؤلاء المفحوصون قد خمنوا بنفس الدرجة وبنفس الطريقة في نفس الفقرات في الصورة الثانية لما اختلفت مراكزهم النسبية في الصورة الثانية عنها في الصورة الأولى .

وكانت معاملات الارتباط بين الدرجات في الإختبارات التي تطلب الإختبار من عدة إجابات وبين الدرجات في الإختبارات التي تطلبت إجابة استرجاعية متقاربة بين المجموعتين من التلاميذ (المجموعة التي طلب إليها التخمين والمجموعة التي طلب إليها عدم التخمين) . وكان الارتباط بين الدرجات المصححة وبين درجات إختبار الإسترجاع أعلى من الارتباط بين الدرجات غير المصححة ودرجات إختبار الإسترجاع . فإذا افترضنا أن الإختبارات من نوع الاسترجاع أكثر صدقا ، لانتبهنا إلى أن التعليمات التي نصدرها للمفحوصين بالتخمين أو عدم التخمين لها تأثير صغير على صدق الإختبارات من نوع الإختبار من عدة إجابات .

وقد توصل توبس Toops سنة ١٩٢١ إلى نتائج مماثلة باستخدام إختبارات الحرفة Trade Tests على طلبة جامعيين ؛ كما توصل اندرو Andrew وبيرد Bird سنة ١٩٣٨ إلى نتائج مشابهة بإستخدام أسئلة من مقررات علم النفس التي درست .

ويمكن تلخيص نتائج هذه الدراسات فيما يلي :

- ١ - الأنواع الخمسة من فقرات الإختبار تختلف قليلا في قياسها للمعلومات والمهارات المختلفة ؛
- ٢ - كلى الأنواع لها معاملات ثبات مرتفعة ومتشابهة عندما تكون الفقرات بين القياس الأول والقياس الثاني مماثلة ؛

٣- التصحيح النظري للصدفة ينخفض الدرجات الكلية ، ويزيد الانحرافات المعيارية ، وهو عادة لا يغير في الثبات أو الصدق :

٤- التعليمات الصادرة إلى المفحوصين بعدم التخمين تقلل متوسط الدرجات الخام بنسبة من ١٠٪ إلى ٢٠٪ ، وتقلل متوسط الدرجات المصححة بنسبة من ٢٪ إلى ٣٪ وتزيد معامل الثبات ومعامل الصدق زيادة طفيفة .

٥- الفقرات من نوع صحيح - خطأ والفقرات من نوع اختيار لإجابة من إجابتين تأثرت بمعامل الصدفة أكثر من غيرها ، ومن ثم زاد تأثير الدرجة الخام بمعادلة التصحيح من أثر الصدفة زيادة أكبر من غيرها :

والتصحيحات النظرية للصدفة لا تنطبق كثيراً على فقرات الاختبار من عدة إجابات . ولكنها تنطبق عادة على فقرات صحيح - خطأ وفقرات الاختبار من إجابتين : والأسباب هي :

١- تصحيح الصدفة يقلل الدرجات الكلية ، ومن ثم يعطى صورة أكثر دقة عن القدرة الحقيقية للفرد .

٢- التصحيح عادة يزيد الفروق الفردية بين الدرجات الخام .

٣- يزيد صدق الاختبار بتصحيح الدرجة من أثر الصدفة :

٤- التصحيح من أثر الصدفة وقاية من النقد الموجه إلى الاختبار :

٤- تحاشي نتائج الصدفة .

تؤدي بنا مناقشة الصدفة إلى أن نقول إنها في الكثير من الاختبارات تكون ذات أثر صغير بحيث يمكن تجاهله ، ولكن ليست هناك طرق مؤكدة للتصحيح الكامل لأخطاء الصدفة في الدرجات الفردية . وهناك طرق ثلاثة لتحاشي النجاس بالصدفة تستخدم في إنشاء الفقرات والاختبارات وهي :

١- قد يتضمن الاختبار عدداً كبيراً من الفقرات وهذا يسمح لقانون

المتوسطات Law of Averages أن يقلل أخطاء الصدفة في الدرجة . وإذا كان الاختبار يتضمن ١٠٠ فقرة أو أكثر من نوع صحيح - خطأ ، كانت الأخطاء في الدرجات الراجعة إلى الصدفة أصغر نسبياً عادة منها في حالة اختبار من ١٠ فقرات من نفس النوع . فحظ الفرد أكثر عرضة للتغير خلال اختبار طويل منه خلال اختبار قصير .

٢- يمكن استخدام الفقرات التي يختار فيها المفحوص الإجابة من أربعة أو خمسة إختيارات ، أو فقرات التكميل وفيهما يكون لإحتمال الخطأ في الإجابة كبيراً . ففقرة الاختبار من نوع الإختيار من عدة إجابات إذا تبعت بأربعة إجابات تسمح فقط بتخمين ثلث الإجابات الصحيحة وثلاثة أمثال هذا العدد إجابات خاطئة تخمينية في المتوسط . والإختبارات المكونة من مثل هذا النوع يصغر تصحيح الصدفة فيها عادة بدرجة كبيرة بالنسبة للدرجة الكلية على الاختبار .

٣- قد يطلب من المفحوصين أن يسجلوا ما إذا كانوا متأكدين أم متشككين في إجاباتهم ، ثم تعطى درجات أكبر للإجابات المؤكدة . ونسبة الإجابات المؤكدة الصحيحة تدل على مدى الحرص في موقف الاختبار . وهذا الدليل على الحرص لم يدرس في بحوث كثيرة ، ولكن يبدو أن له دلالة عن إحدى سمات الشخصية . فقد وجد وست West سنة ١٩٢٣ ، و ترو Trow سنة ١٩٢٣ ، وجرين Greene سنة ١٩٢٩ وسنة ١٩٣٨ وجدوا أن الأشخاص يختلفون في ميلهم إلى الحرص تبعاً للعوامل الآتية :

(أ) نوع مادة الاختبار .

(ب) ألفة المفحوصين بهذه المواد .

(ج) أمزجة Temperaments المفحوصين ودقة تدريسهم :

إذا قلت النسبة المثوية للفقرات المؤكدة الإجابة عندما كانت المادة في الاختبار مجردة abstract بينما زادت نسبة التأكد من الإجابات الصحيحة

عن الفقرات في حالة المواد المحسوسة Concrete : والأشخاص الذين حصلوا على درجات عالية إرتفعت النسبة المئوية لتأكدهم من الإجابة الصحيحة عن الفقرات أكثر من تأكد من حصلوا على درجات منخفضة . وقد مال التلاميذ الضعاف إلى إدعاء التأكد من الإجابة كما فعل التلاميذ الأقوياء .

أفضل طرق التصحيح

تصميم الفقرة يحدد سرعة ودقة تصحيح الإجابة . وهناك عادة عمليتان تقوم بهما قبل حساب الدرجة . العملية الأولى هي مراجعة الإجابات بمفتاح Key التصحيح والثانية هي عد الإجابات الصحيحة لحساب الدرجة الكلية . كما تعد الاجابات الخاطئة والمحذوفة Omitted عند ما يجرى تصحيح الدرجة الكلية من النجاح بالصدفة .

وفي فقرات المقال أو فقرات التكميل ، يحوى المفتاح قائمة بالإجابات الصحيحة نقارن بها الإجابات التي أتى بها المفحوص ولا يبدو أن هناك طريقة آلية مختصرة لتصحيح مثل هذه الإختبارات . ويمكن زيادة سرعة تصحيح فقرات التكميل ، بكتابة كل الإجابات التي يأتي بها المفحوص في عامود رأسى حيث يسهل مقارنتها بالمفتاح .

أما فقرات صحيح - خطأ ، أو فقرات الإختيار من عدة إجابات ، أو إختبارات المزاجية ، هناك ثلاثة وسائل للتصحيح : المفتاح المثقب Stencil ، والتصحيح الأتوماتيكى Automatic ، والتصحيح الآلى Machine Scorers وهى توفر الوقت وتزيد الدقة . والإقتصاد يزيد في حالة البحوث التى تستخدم عدداً كبيراً من الإختبارات والتي تطبقها على عدد كبير من المفحوصين . ولا يمكن وضع قاعدة بسيطة يمكن على أساسها تحديد أكثر الطرق اقتصاداً في التصحيح لأننا يجب أن نحافظ على

ثلاثة متغيرات على الأقل وهذه المتغيرات هي : (١) أجر الكتبة القائمين بالتصحيح (٢) تكاليف تصميم وطبع مفاتيح التصحيح الأوتوماتيكية (٣) وتكاليف تشغيل الآلات .

(١) المفتاح المثقب :

تصحح فقرات الاختبار من من عدة إجابات وفقرات الاختبار صحيح - خطأ غالباً باستخدام مفاتيح مثقبة Stencil مطبوعة أو مفاتيح التصحيح التي توضع ، مطابقة تماماً ، على ورقة الإجابة أو موازية لها تماماً . والمفتاح المثقب يمكن المصحح من أن يرى الإجابات الصحيحة بسرعة ، وأن يعلم عليها بأي طريقة :

وفي الكثير من المؤسسات ، حيث يصحح عدد كبير من الإختبارات ويجرى إختبار أعداد كبيرة من الأفراد ، يطلب من المفحوص أن يجيب على ورقة إجابة منفصلة Answer sheet ثم تمر هذه الأوراق إلى ماكينة تطبع الإجابات الصحيحة بحبر ملون بجوار الإجابات التي كتبها المفحوص وهكذا يسهل مقارنة الإجابات بالمفتاح بسرعة .

(ب) التصحيح الأوتوماتيكي :

التصحيح الأوتوماتيكي يسهل عملية مراجعة الإجابات على المفتاح التي تلزمنا بها المفاتيح المثقبة . والكثير من الاختبارات طبع بطريقة تجعل وضع الإجابة يدل على مدى مطابقتها للمفتاح . وهذا المفتاح يوضع بحيث لا يستطيع أن يراه المفحوص أثناء إجراء الاختبار .

وقد نشر كلاب Clapp ويونج Young سنة ١٩٢٩ سلسلة من الإختبارات طبع فيها المفتاح على ورقة توضع تحت ورقة الإجابة : وظهر ورقة الأجابة مغطى جزئياً بطبقة رقيقة من الكربون الناسخ : ويتكون الإختبار من فقرات يختار فيها المفحوص الإجابة من أربعة إجابات بوضع

علامة X على إحدى مربعات أربعة صغيرة مرفقة تتبع الفقرة . والمفحوص يضع علامة X بقلم عادى فى المربع الذى يعتقد هو أنه يرمز إلى الإجابة الصحيحة . وفى نفس الوقت تطبع العلامة X على ورقة المفتاح التى تبين الإجابات الصحيحة بواسطة مربعات مطبوعة صغيرة . ولحساب الدرجة على الاختبار ككل ، يقوم الباحث بعد المربعات على ورقة المفتاح التى طبعت عليها علامة X . وكان هذا البحث أول بحث يستخدم طريقة التصحيح هذه . ولكن الطريقة الأوتوماتيكية لا تعطى الدرجة الكلية تلقائياً بل على الباحث أن يعد العلامات أو المربعات .

(ح) التصحيح الآلى :

ماكينة العمل العالمية (I.B.M.) International Business Machine المصححة صممت سنة ١٩٣٨ وفيها تستخدم ورقة إجابة مطبوعة مصممة بدقة حيث يعلم المفحوص إجاباته بقلم خاص . وورقة الإجابة مطبوع بها خطوط صغيرة متوازية أمام رقم كل سؤال ، وهذه الخطوط تحصر بينها مسافات صغيرة ويسود المفحوص بالقلم الخاص إحدى هذه المسافات أمام كل سؤال . ولتصحيح هذه الورقة ، توضع فى الماكينة ، ثم تحركها رافعة ، وتطبع الدرجة الكلية على الورقة : والتصحيح مبنى على إتصالات كهربائية فى المواضع التى بها علامات بالقلم الخاص . وتصحيح التخمين يتم فيها وذلك بوضع مفتاح مثقب خاص فى مكان معين بالماكينة بالضغط على موضع معين مرتين . وبهذه الطريقة يمكن تصحيح ٣٠٠ فقرة صح - خطأ فى نفس الوقت . وأوراق الإجابة تجرى خلال الماكينة بسرعة وبمجرد إدخالها تصحيح وتوضع الدرجة الكلية : وتحتاج هذه الماكينة إلى كاتب مدرب عليها .

الفصل السادس

الصدق

- معنى الصدق .
- أنواع الصدق .
- المحكات .
- حساب الصدق .
- الخطأ المعياري والقياس .
- جمع الاختبارات في بطارية .

الفصل السادس

الصدق

معانى الصدق :

الصدق Validity مفهوم واسع له عدة معانى تختلف بحسب استخدام الاختبار ، وقد يترجم المصطلح إلى « الصحة » أو « الصلاحية » ويرجع اختلاف الترجمة غالباً إلى اختلاف المسائل التى يهتم بها الباحثون عند التعرض لمعنى الصدق . وأول معانى الصدق هو أن يقيس الاختبار ما وضع لقياسه ، أى أن الاختبار الصادق اختبار يقيس الوظيفة التى يزعم أنه يقيسها ولا يقيس شيئاً آخر بدلاً منها أو بالإضافة إليها . فاختبار القدرة الميكانيكية مثلاً ، لكى يكون صادقاً يجب ، بقدر الإمكان أن يقيس هذه القدرة وحدها . فلا يقيس المهارة اليدوية ، مثلاً بدلاً منها ولا يقيس القدرة الميكانيكية والمهارة الحركية معاً .

والاختبار الصادق اختبار يصلح لقياس الجانب المقصود أى أنه اختبار يعطى درجة تعدد إنعكاساً أو تمثيلاً لقدرة الفرد . فاختبار الحساب الصادق اختبار يصلح لقياس مستوى التلميذ فى الحساب . فلا يكفى أن يكون اسمه اختبار حساب فالتسميات أحياناً تكون مضللة وهى أعم مما يقيسه الاختبار ، واسم الاختبار لا يدل على ما يقيسه بل على الغرض من القياس . واختبار الحساب الصادق يصلح لقياس الحساب من مستوى معين لا يجدى فى غيره كما فعل فى هذا المستوى وليكن الفرقة الدراسية المعينة . فاختبار الحساب المصمم للفرقة الثانية ، الذى ثبت صدقه بالنسبة لها ، لو طبق فى السنة الثالثة لتأثر بالذاكرة والسرعة ولو طبق فى الفرقة الأولى لبين الاستعداد

الحسابي والقدرة على القراءة والفهم . وهكذا فالاختبار يصدق بالنسبة لوظيفة معينة في مستوى معين ، أى أنه لا وجود للصدق المطلق .

وهكذا نجد أن الصدق نسبي Relative ، بمعنى أن الاختبار يكون صادقاً بالنسبة لجماعات معينة أجريت عمليات الصدق واشتقت معاملاته من درجات هذه الجماعات أو ممثلين لها . وهنا تتحدد نسبية الصدق بخصائص وطبيعة المجتمع الأصلي الذى اشتقت منه معاملات الصدق . فالاختبار الصادق بالنسبة للأطفال ليس صادقاً بالنسبة للمسنين ، والاختبار الذى صدق على الذكور قد لا يصدق على الإناث . والاختبار الذى ثبت صدقه فى أمريكا قد لا يثبت صدقه فى بلادنا . ونقصد أن الصدق قد لا يثبت بنفس الدرجة فى هذه الأحوال جميعا . ومن أجل هذا يجب عند ذكر معاملات الصدق لاختبار معين أن نذكر طبيعة وخصائص ومستوى العينة التى استخدمت فى حسابه .

والصدق نوعى Specific ، بمعنى أن الاختبار يكون صادقاً لأنه يقيس جانباً وضع لقياسه ، وهو لا يقيس جانباً آخر . فالاختبار بالنسبة لهذا الجانب الآخر غير صادق . وعند ذكر بيانات عن صدق الاختبار ، يجب أن نحدد الوظيفة أو الوظائف الممثلة فيه ، والتي يعتبر هو أداة لقياسها . وعند استخدام الاختبار نتحدد بهذه الوظائف ، فلا يستخدم اختبار H. T. P. ، مثلاً ، كاختبار فنى للرسم لأنه يعتمد على الرسم كاستجابة لموقف الاختبار ، بل يستخدم فى قياس الشخصية من حيث المتغيرات التى ثبت أنه يصلح لقياسها . وصدق الاختبار يعنى أنه يقيس ، فعلاً ، الجانب الذى وضع لقياسه ، فصدق الاختبار إذن يمدنا بدليل مباشر على مدى صلاحيته لقياس أحد المتغيرات ، وهل نجح فى تحقيق الأغراض التى وضع من أجلها فى حدود الوظيفة المقاسة ، والعينة التى نقيس داخلها هذه الوظيفة . وقد عرف Kurkon النسبة بأنه « إلى أى مدى يؤدي الاختبار عمله كما يجب » : وهو

تعريف شامل يتضمن تحقيق الغرض من عملية القياس وقياس الوظيفة المعينة ، وقياسها عند نوع معين من الجماعات . وهكذا يتوقف الحديث عن صدق الاختبار على الأمور الآتية :

١ - تحديد الوظيفة أو الجانب من الحياة النفسية الذى يقيسه الاختبار .

٢ - بيانات تتعلق بالعينة التى أجريت عليها بحوث الصدق ، فيما يتعلق بما يأتى :

(١) عدد الأفراد .

(ب) توزيعهم بحسب الجنس والسن والمهنة والتعليم . . . الخ .

(ج) تحديد خصائص وطبيعة العينة .

٣ - المعنى المحدد الذى نقصده بالصدق .

٤ - نوع الصدق .

٥ - معاملات الصدق المختلفة .

٦ - طرق حسابه .

٧ - الأخطاء المعيارية ، والأخطاء المحتملة .

٨ - تمثيل الاختبار ، باعتباره عينة من المثيرات ، للسلوك المراد قياسه :

٩ - طول الاختبار .

١٠ - معامل الثبات .

١١ - تجانس Homogeneity المجموعة .

وقد نعنى بصدق الاختبار قدرته التنبؤية ، فيعد الاختبار صادق أداة نقيس بها عينة من المثيرات تمثل المنطقة المطلوبة من السلوك ، وتعد الدرجة دليلاً تنبأ به على نجاح من يحصل عليها عند التحاقه بتدريب أو دراسة معينة يستخدم فيها قدرة أو استعداداً قاسه الاختبار . والحق أن هذا المعنى هو الأكثر شيوعاً ، وهو معنى يسهل على أساسه حساب الصدق إحصائياً

وتجريبياً . فالصدق بهذا المعنى هو صحة توقعنا لأن ينجح فلان في مهنة أو دراسة معينة ، لأنها تتطلب مستوى معين من القدرات والاستعدادات ، ما دام قد حصل على درجات معينة في اختبارات تقيس هذه الاستعدادات وتلك القدرات .

وهنا يتوقف صدق الاختبار على مدى صحة التوقعات التي نقوم بها عند استخدامه ، أى على مدى مطابقة سلوك المفحوص مستقبلاً لاستجاباته عن الاختبار . والمطابقة بين السلوك في موقف الاختبار ، والاستجابة لعينة المثيرات المتضمنة فيه ، وبين السلوك في مواقف مشابهة مستقبلاً ، أمر يتوقف على مدى هذا التشابه . أى على تمثيل الاختبار تمثيلاً كافياً لحالات وجوانب وأبعاد الوظيفة التي يقيسها . فالاختبار عينة للقدرة وهو يخضع لأخطاء اختيار العينة .

فيجب ، بهذا المعنى ، أن يتعلق الاختبار بالجانب المقاس وأن يتعلق الجانب بالدراسة أو المهنة التي نريد أن نتنبأ بالنجاح فيها أو الصلاحية لها . هذا التعلق relevance قد يكون كبيراً جداً كما في حالة اختبار السائقين الذي يجري في قيادة سيارة في الطريق ، فالسلوك في موقف الاختبار هو هو السلوك في مواقف تالية من نفس النوع . ويزيد التعلق في اختبارات تعتبر صوراً مصغرة Job miniature للعمل . والاختبارات التي صممت بتحليل أعمال ودراسات معينة ثم أخذ عينة صادقة من متعلقات العمل أو الدراسة ووضع أسئلة تكون هذه المتعلقات إجابة عنها ، اختبارات صادقة .

والصدق هنا على تنبئه بالمقارنة بين درجات الأفراد في الاختبار ونجاحهم في العمل أو الدراسة . وقد نقوم بتحليل منطقي للوظائف والعمليات الممثلة في الاختبار مقارنة بالعمليات والوظائف التي يعتمد عليها العمل

أو الدراسة : ويكون التحليل المنطقي للعمليات Operations قياساً كينافياً للتعليق : وهكذا يجب أن يتعلق الاختبار بالجانب المراد قياسه ، وأن يقيس هذا الجانب دائماً أى بثبات : والاتساق بين الاختبار ونفسه فى قياسه لما يتعلق به ، متمم للتعليق فى إعطاء معنى للصدق ؛

ومن هنا نتبين أن مفهوم الصدق يتضمن ، بمعنى ما ، مفهوم الثبات : بمعنى أن اختبار الذكاء لا يقيس الذكاء إلا إذا كان يعطى قياساً متقارباً فى المرات المختلفة ، أى أن يكون الاختبار متفقاً مع نفسه ، بينما يعنى الصدق أن يتفق الاختبار مع وظيفة يتعلق بها : ويستحيل ، بالطبع ، أن يتفق الاختبار مع ما يتعلق به أكثر مما يتفق مع نفسه : أى أنه لا يحدث أن يرتفع معامل الصدق بأكثر من معامل الثبات : وكذا فإن الاختبار الصادق اختبار ثابت فى قياساته فى المرات المختلفة ، ولكن الاختبار الثابت ، بهذا المعنى ، ليس بالضرورة اختباراً صادقاً . فإذا كان أمامنا اختبار ثابت يقيس جانباً لا نعرفه على وجه الدقة ويعطينا درجات متقاربة ، ما عرفنا أنه يقيس شيئاً ما بذاته . فالاختبار قد يزعم أنه يقيس الذكاء مثلاً ويعطى للفرد الواحد عند إعادة الإجراء درجة قريبة مما حصل عليه فى الإجراء الأول ، ويعطى درجات متقاربة لو استخدمه إخصائيون مختلفون ، ولكنه برغم هذا كله قد لا يكون اختباراً صادقاً فن أدرانا أنه يقيس الذكاء ولا يقيس الذكاء مثلاً . وقد يقيس الذكاء فى كل مرة ويكون ثابتاً فى هذا القياس وإن قاس ما لم يرد قياسه ، أى أن الاختبار لكى يكون صادقاً يجب أن يكون ثابتاً ، وليس العكس صحيحاً ، فالاختبار قد يكون ثابتاً وليس صادقاً . بل وأكثر من هذا ، فكما رأينا ، يتضمن الصدق التعليق والثبات أى أنه يكفى أن نقول إن هذا اختبار صادق فنستدل أنه ثابت ، إلى حد ما على الأقل . ومن هنا وجه الانتقادات إلى الكثير من الاختبارات الشهيرة ذات الثبات المرتفع

فقد تبين أن كثيراً منها لا يقيس كل ما كان يزعم قياسه ، وإن قاس بضربه ، أو أنها لا تقيس هذا فقط بل تقيس شيئاً آخر إلى جواره .

وبهذه الطريقة حدد استخدام تسميات مثل اختبار الذكاء ، واختبارات الشخصية . فقد قام التحليل العاملي بتحليل اختبارات زعمت أنها تقيس كل مفهوم منهما واشتق عوامل factors نقية إلى حد ما ، وضعت اختبارات تقيسها . هذه الاختبارات تقيس أحد العوامل التي ثبت ، باستخدام منهج علمي أنه متعلق بالعمل المعين ، وبالاختبارات المتعلقة بهذا العمل ، هذه الاختبارات إذن صادقة أكثر من غيرها خاصة وأنها تقيس ، غالباً ، عوامل نقية . وكلما حدد وحلل العامل الذي صمم الاختبار لقياسه ، كان الاختبار صادقاً ، وارتباط الاختبار بالعامل يحدد صدقه . وهذا هو الفهم العاملي factorial للصدق .

كما أصبح الاختبار الصادق هو الذي يرتبط ارتباطاً كبيراً ، في درجات الأفراد عليه ، بالعمل أو الدراسة ، في نجاح الأفراد فيها نجاحاً بدرجة معينة يطلبها واضع الاختبار أو مستخدمه ، أي أن المعنى التجريبي للصدق هو الذي يؤخذ به أكثر . وهو ، إلى جوار المعنى العاملي ، منهج علمي لدراسة الصدق بدلا من الاعتماد على ما ذكر مؤلف الاختبار من جوانب يقيسها ، وبدلا من الاكتفاء باسم الاختبار للتعرف على الوظائف التي يقيسها . فمن وجهة النظر العملية ، صدق الاختبار هو كفايته التنبؤية في أي جانب مقاس من الحياة اليومية .

معامل الصدق :

معامل الصدق Coefficient of Validity هو أحد تطبيقات معامل الارتباط ، ومع معامل للارتباط بين درجات الأفراد في الاختبار ودرجاتهم في الوظائف التي يتعلق بها الاختبار ونعرف درجات الأفراد في الوظيفة عن

دارين مقياس آخر غير الاختبار بقدر ما يقيسه الاختبار هو المحك criterion وسنتحدث عنه ، فيعطى الأفراد رتباً أو درجات على هذا المقياس (المحك) ونوجد بإحدى الطرق مدى ارتباط درجة الفرد في الاختبار بدرجة في المقياس الآخر . فإذا كان هذا الارتباط كبيراً كان معامل الصدق كبيراً وكان الاختبار صادقاً .

ومعنى هذا أن من يحصل على درجة مرتفعة في الاختبار يحصل على درجة مرتفعة في مقياس المحك ، وأن من يحصل على درجة منخفضة في الاختبار يكون مستواه منخفضاً في الوظيفة التي يقيسها الاختبار بعد أن عرفنا هذا المستوى من الدرجة على مقياس المحك . ويتراوح معامل الصدق نظرياً بين $- 1$ ، $+ 1$ وإن كان لا يحدث أن ترتبط الدرجة في الاختبار مع الدرجة في المحك ارتباطاً سالباً تاماً ولا سالباً على وجه العموم . فليست هناك أحوال يحصل فيها الأفراد على درجة عالية في الاختبار ودرجة منخفضة في مقياس المحك دائماً . أى أن معامل الصدق فعلاً ينحصر بين صفر ، الذى يعبر عن عدم وجود علاقة بين الدرجة على الاختبار ودرجة على مقياس المحك تمثل مستوى الوظيفة المقاسة ، وبين واحد صحيح ، الذى يعبر عن وجود علاقة تامة موجبة بين الدرجة على الاختبار والدرجة على مقياس المحك باعتبار الأخيرة ممثلة لوظيفة يقيسها الاختبار .

والمشكلة هي أن الارتباط الموجب التام لا يحدث في حالة الصدق ، بل نرى ارتباطاً موجباً جزئياً . فهل يكفي أى ارتباط جزئى موجب للدلالة على صدق الاختبار ؟ وما هي حدود الكسر أو النسبة المثوية لمعامل الصدق الكافى ؟ أى أن السؤال ينحصر في تحديد مستويات يصل إليها معامل الصدق فيكون كافياً أو لا يصل فلا يكون الاختبار صادقاً . والحق أن هذا السؤال لإجابة محددة له ولم يتفق علماء القياس على مستويات مرضية للصدق . ويمكن أن تخف حدة المشكلة وأن نحدد مستوى الصدق الذى يجب أن

يكون عليه الاختبار الذي نستخدمه ، إذا وضعنا في الاعتبار عدة ظروف أهمها الغرض الذي سنستخدم الاختبار لأجله . وبالنسبة لمستوى صدق اختبار يستخدم في حالات فردية تكون هناك مخاطرة في الحكم على شخص دون تحديد مستوى ثقتنا في هذا الحكم أى دون تحديد الخطأ المعياري للتقدير وسنتحدث عنه عما قريب . إلا أن هناك نسباً وضعت ، على أساس دراسة إحصائية ، تبين أن استخدام الاختبار يعطى نتائج أفضل من الاعتماد على عامل الصدفة بمقدار يتوقف على قيمة معامل الصدق . فعامل الصدق إذا بلغت قيمته ٩٠ ر أعطى الاختبار نتائج تفضل الصدفة بمقدار ٥٦٪ . أى لو كنا بصدد اختبار نختار به تلاميذ يلتحقون بدراسة معينة ، ولم نعتمد على الاختبار وكانت الصدفة هي العامل الذي تحدد به الاختبار ، وجدنا ، مثلاً ، ما يلي : إذا اخترنا معتمدين على الصدفة ١٠٠ تلميذ (بدون استخدام الاختبار) ، يكون عدد الناجحين ٥٠ بالصدفة أيضاً ، وإذا استخدمنا اختباراً بمعامل صدقه ٩٠ ر . استطعنا أن نختار ٧٨ تلميذاً ناجحين بعد استخدام هذا الاختبار .

ومعامل صدق الاختبار إذا بلغ ٨٠ ر دل على أن الاختبار أدق من الصدفة بمقدار ٤٠٪ بنفس الطريقة . وإذا بلغ معامل الصدق ٧٠ ر . كان الاختبار أفضل من الصدفة بمقدار ٢٩٪ أما إذا بلغ معامل الصدق ٤٠٪ تعطى الاختبار نتائج تفضل الصدفة بمقدار ٨٪ فقط . وفيما يلي (جدول ٢) بوضوح الكفاية التنبؤية تبعاً لقيمة معامل الصدق :

احتمال النجاح والفشل بين (١٠٠) مختارون بواسطة الاختبار		الزيادة في النسبة المئوية للكفاية التنبؤية	معامل الصدق
احتمال الفشل	احتمال النجاح		
٥٠	٥٠	صفر	صفر
٤٩.٧٥	٥٠.٢٥	٠.٥	٠.١٠
٤٩	٥١	٢.٠	٠.٢٠
٤٧.٥	٥٢.٥	٥	٠.٣٠
٤٦	٥٤	٨	٠.٤٠
٤٣.٥	٥٦.٥	١٣	٠.٥٠
٤٠	٦٠	٢٠	٠.٦٠
٣٥.٥	٦٤.٥	٢٩	٠.٧٠
٣٠	٧٠	٤٠	٠.٨٠
٢٢	٧٨	٥٦	٠.٩٠
١٥.٥	٨٤.٥	٦٩	٠.٩٥
١٠	٩٠	٨٠	٠.٩٨

(جدول ٢) الكفاية التنبؤية تبعاً لقيمة معامل الصدق

Clifford P. Froehlich and John G Darley, Studying Students —
Guidance Methods of Individual Analysis, Chicago : Science Research
Associates 1952

أنواع الصدق

للصدق أنواع تختلف بحسب التركيز على أحد معانيه التي سبق ذكرها .
وهناك أنواع متعددة اختلفت من مشغل بالقياس إلى آخر ، وسنقتصر على
الأنواع الأكثر شيوعاً واستخداماً وهي :

١ - الصدق الظاهري :

الصدق الظاهري face Validity هو البحث عما يبدو أن الاختبار يقيسه . فبالنظر إلى الاختبار قد يبدو صادقا لأن اسمه يتعلق بالوظيفة المراد قياسها ، مع العلم بأن الاسم يدل غالبا على الغرض من استخدام الاختبار ، فالصدق الظاهري هكذا يشير إلى كيف يبدو الاختبار مناسبا للغرض الذي وضع من أجله . ويتضح هذا النوع من الصدق بالفحص المبدئي لمحتويات الاختبار أى بالنظر إلى الفقرات ومعرفة ماذا يبدو أن تقيسه ثم مطابقة هذا الذى يبدو بالوظيفة المراد قياسها فإذا اقترب الاثنان كان الاختبار صادقا صدقا سطحيا أى بدا أن فقراته تتصل غالبا بجانب مطلوب .

وعند القيام بمثل هذا النوع من التحليل الصدقي تبدو بعض الفقرات ضعيفة أو عديمة الارتباط بالجانب المطلوب ، وهنا نستبعد هذه الفقرات . ولكن هذا الإجراء معرض لأخطاء كثيرة ، فقد يبدو اختباران ، من الناحية الظاهرية متساويين ، فى الصدق ، وعند إجرائهما يتضح الفرق . وقد حسب Guilford معامل الارتباط بين درجات مجموعة من الأفراد على اختبارين ، من الثابت أنهما صادقين ظاهريا ، وكان معامل الارتباط صفراً ، والاختباران من نوع اختبارات الترتيب بحسب الحروف الأبجدية Alphabetizing test وقد دعى هذا جيلفورد إلى أن يسمى هذا النوع من الصدق ، على سبيل التقليل من شأنه والتصغير من قيمة الاعتماد عليه ، صدق الإيمان أو العقيدة faith Validity . فمشكلة الصدق الظاهري تتوقف على تقبل الناس لهذا الاختبار وشيوع استخدامه فى وظيفة معينة ، اعتماداً على اسمه وعلى السوابق ، وهكذا يستطيع الباحث أن يكسب ثقة الناس فى الاختبار .

أما إذا كان الاختبار ، فى نظر غير المدقق ، يبدو غير مرتبط ظاهرياً بما يراد قياسه ، اعتبر الاختبار غير مناسب للغرض الذى يراد استخدامه فيه ، ومن ثم صعب كسب ثقة الجمهور وتعاونهم فى موقف الإجراء .

وليس هذا هو الحال دائماً ، فأحياناً لا يبدو الاختبار إطلاقاً ذا صلة بالوظائف المراد قياسها ، وهنا يخفى على المفحوصين الغرض من استخدام الاختبار كما هو الحال في الاختبارات الإسقاطية ، وبرغم هذا يكون صدق الاختبار مقبولا . أى أن الاختبار قد يصدق علمياً وعملياً ولا يصدق سطحياً وظاهرياً ، كما أننا لسنا بحاجة غالباً إلى مثل هذا النوع من الصدق ، حتى لكسب تأييد الجمهور المقاس . إلا أنه يلزم ، وخاصة عند اختبار المستويات التعليمية المتواضعة كالعمال ، أن يبدو أن نوع الفقرات والمواد المستخدمة ذات علاقة بالوظيفة المراد قياسها . بل ويحسن أن نستخدم في الاختبار مواداً تناسب طبيعة العينة ومهن أفرادها حتى لو كنا بصدد اختبار يقيس وظيفة بعيدة عن هذه المهنة ، فيحسن في اختبارات العلاقات المكانية إذا أجريت في المدارس أن يطلب من التلميذ تقدير أبعاد في الفصل أو المدرسة وهكذا . وبالمثل يفيد الصدق الظاهري في زيادة الدافع لموقف الاختبار .

وفيد استخدام الصدق الظاهري في اختبارات التلاميذ في الفصول ، والعمال في الاختيار المهني ، وفي المجال العسكري ، وفي أغراض أخرى كثيرة : ولا يفيد في اختبارات التشخيص أو في المجال الإكلينيكي .

٢ - صدق المضمون :

صدق المضمون Content Validity هو قياس مدى تمثيل الاختبار لنواحي الجانب المقاس . وقد يسمى بالصدق المنطقي logical Validity أو الصدق بالتعريف Validity by definition . وهنا نقوم بتحليل منطقي لمواد الاختبار وفقراته لتحديد الوظائف والجوانب والمستويات الممثلة فيه ونسبة كل منها إلى الاختبار ككل . ثم نقوم بمسح لمجال السلوك المطلوب

قياسه ، والتعرف على عواماه ، وأهمية كل عامل ، ونسبة تأثير هذا العامل على السلوك الذى يمثل الوظيفة ككل والأهمية النسبية لكل ناحية من النواحي المختلفة . ثم نطابق بين الاختبار والوظيفة التى يقيسها ، على هذه الأسس ، لتعرف على مدى تمثيل الاختبار للوظيفة المطلوبة وعواملها ومكوناتها ونسبها . وفى كل هذا نحتاج إلى التحليل المنطقى المنظم ، والتحليل اللغوى اللفظى للأسماء ، والتعريف الدقيق للمفاهيم .

فصدق المضمون إذن يعتمد على مدى تمثيل الاختبار للمواقف . أو الجوانب التى يقيسها ، مؤكداً على التمثيل الصادق . فإذا كان صدق مضمون الاختبار مرتفعاً كانت منطقة الساوك التى يفترض أن هذا الاختبار يقيسها ممثلة تمثيلاً جيداً فى فقرات الاختبار . ويفيد صدق المحتوى فى اختبارات التحصيل والكفاية . وكما رأينا . فاختبار التحصيل الناجح يرتفع صدق مضمونه كلما كان ممثلاً للموضوعات المدروسة بحسب نسبة تركيب كل موضوع منها . وسنتكلم عن الطرق المختلفة لحساب هذا النوع من الصدق .

٣ - الصدق التنبؤى :

الصدق التنبؤى Predictive Validity يقوم على أساس حساب القيمة التنبؤية للاختبار ، أى معرفة مدى صحة التنبؤات التى نبينها معتمدين على درجات الاختبار . والتنبؤ هنا يقوم على أساس أن السلوك يتمتع بدرجة من الثبات وأن الاتجاهات العامة والتنظيم فى الاستجابات يميز مجموعات من الأرجاع السلوكية لدى الشخص بصفة معينة أو بطابع خاص يظل يشكل استجاباته فى أغلب المواقف . هذا الطابع نادراً ما يتخلى الفرد عنه . وهكذا ، بقياس استجابات الفرد لمواقف معينة ، يفترض أنها تمثل قطاعاً

من المواقف التي سيعرض لها الفرد ، يمكن استنتاج كيف سيسلك المفحوص .
فيما بعد إذا عرفنا كيف يسلك الآن .

فإذا قسنا وظائف معينة ، ثبت أن النجاح في مهنة أو دراسة معينة يعتمد عليها ، ثم جمعنا معلومات عن مدى نجاح من حصلوا على درجات مرتفعة ومدى نجاح من حصلوا على درجات منخفضة ، أمكننا أن نعرف إلى أي مدى سينجح الفرد الذي يحصل على درجات معينة أو ما هو مقدار احتمال نجاح من يحصل على درجات معينة . وفي التنبؤ نحتاج إلى فترة بين الإجراء وجمع البيانات عن مقياس موضوعي آخر للنجاح المقبل في العمل أو الدراسة . أي أننا ننتظر إلى أن ينهي الطالب دراسته أو العامل تدريبه ، ثم نقارن درجاته في الاختبار بالنجاح الذي حصله مستقبلاً ونستدل على القيمة التنبؤية للدرجات المعطاه عن اختبار افترض أنه يقيس جوانب يعتمد النجاح في المهنة أو الدراسة عليها . وسنتحدث عن طريقة حساب هذا النوع من الصدق .

٤ - الصدق التلازمي :

الصدق التلازمي Concurrent Validity يعني كشف العلاقة بين الاختبار ومحك تجمع البيانات عليه وقت أو قبل إجراء الاختبار . أي أننا نقارن بين درجات الأفراد على الاختبار ودرجاتهم على مقياس موضوعي آخر يحسب مركزهم فيما يقيسه الاختبار ، على أن يعطى الأفراد درجات على المحك في نفس الوقت الذي نطبق فيه الاختبار ، وغالباً قبل الإجراء .

٥ - الصدق التجريبي :

الصدق التلازمي بالإضافة إلى الصدق التنبؤي قد يشار إليهما معا بالصدق التجريبي أو العلمي أو صدق الوقائع الخارجية Emperical Validity . فهما

معا يقيسان مدى اتفاق نتائج الاختبار مع الوقائع الخارجية المتعلقة بالسلوك الفعلي في جانب يقيسه الاختبار . والفرق بينهما ينحصر في أننا في الصدق التنبؤي نجمع البيانات عن المحك بعد إجراء الاختبار بفترة مستخدمين الطريقة التتبعية Follow-up method أى تتبع الأداء الواقعي للخاصية المختبرة ، وسنتحدث عن هذه الطريقة فيما بعد .

أما في الصدق التلازمي فإننا نوجد مدى الارتباط بين درجات الأفراد في الاختبار ورتبهم على مقياس المحك ، على أن يطبق المحك قبل أو أثناء إجراء الاختبار . وسنتحدث عن هذه الطريقة تفصيلاً فيما بعد . وهكذا نرى أن الصدق التجريبي ، بطريقتيه التنبؤية وفيها نتبع أفراداً قسناهم بالاختبار والتلازمية وفيها نقارن بين مجموعتين من الدرجات لإحدهما على الاختبار والأخرى على مقياس المحك المطبق حين إجراء الاختبار ، يعتمد على إجراء بحوث تجريبية أو القيام بعمليات إحصائية لمقارنة مدى دلالة درجة الفرد على اختبار في جانب ، وبين سلوكه الواقعي المعتمد على هذا الجانب

٦ - صدق المفهوم :

صدق المفهوم Construct Validity هو الارتباط بين الجوانب التي يقيسها الاختبار وبين مفهوم هذه الجوانب . أى أننا عند تحديد صدق المفهوم أو التكوين نقوم ، بطريقة أو بأخرى ، بتحديد ما نقصده بمصطلح يصف جانباً يقيسه الاختبار . ثم نفحص درجات الأفراد على الاختبار ونبين كيف تفسر هذه الدرجات باستخدام الجانب المقاس . فنبحث عما إذا كانت الفروق بين درجات الأفراد في الاختبار ترجع إلى الفروق بينهم في درجة الخاصية التي يقيسها الاختبار . فإذا كان الاختبار يقيس العصائية Neuroticism كانت الفروق بين درجات مجموعتين من الأشخاص ترجع إلى درجة قابلية كل منهما للإصابة بالعصاب (المرض النفسي) حسب تحديدنا لمفهوم العصائية .

فواضع الاختبار صاحب نظرية أو أنه صممه على أساس نظرية معينة تتعلق بجوانب افترض أن الاختبار يقيسها . وهو لكي يحدد ما إذا كان الاختبار يمثل هذه النظرية ، وما إذا كانت الفروق بين الدرجات على الاختبار تفسر بواسطة الفروق بين مستويات القدرة التي تعالجها النظرية وقياسها الاختبار . أى أن مصمم الاختبار يحاول أن يثبت صحة النظرية التي اعتمد عليها في إنشاء الاختبار ، وذلك بأن يحدد التنبؤات والتوقعات التي يفترضها بفحص الفروق في الدرجات مرجعاً إليها إلى النظرية التي صمم الاختبار على أساسها والتي غالباً ما تقضى بأن اختلاف الأفراد في أحد جوانب السلوك يعنى تفاوتهم في قدرة معينة . وهناك طريقتان لحساب صدق التكوين سنعرض لهما عما قريب .

٧ - الصدق التطابقي Congruent Validity :

هو أحد طريقتي حساب صدق المفهوم : ونحصل على معامل الصدق التطابقي بحساب مدى اتفاق درجات مجموعة من الأفراد في الاختبار مع درجاتهم على اختبار آخر ثبت أنه صادق في قياس نفس السمة التي يقيسها الاختبار الجديد . وسنتحدث عن هذه الطريقة فيما بعد .

٨ - الصدق العاملي :

في الصدق العاملي Factorial Validity نهتم بحساب درجة تشبع الاختبار بالجانب المطلوب قياسه . أى أننا نبحث عن عوامل مشتركة تقيسها عدة اختبارات لتحديد مدى اشتراك هذه الاختبارات في قياس تلك العوامل ، ومدى تقاء كل اختبار في قياسه لأحد أو قليل من هذه العوامل . والصدق العاملي هو الطريقة الثانية لإثبات صدق المفهوم . فنجد البعض يجمع بين الصدق التطابقي والصدق العاملي عند الحديث عن صدق المفهوم . وسنتحدث تفصيلاً عن طرق حساب الصدق العاملي عما قريب .

مشكلة المحكات

تعريف المحك :

المحك Criterion مقياس موضوعى مستقل ، عن الاختبار ، نقيس به صدق الاختبار . وأصل كلمة محك فى اللغة العربية يرجع إلى تقليد كان العرب يتبعونه . كانت هناك قطع من الأحجار الكريمة متفق على أنها أصيلة نقية (إذن فهي صادقة) يستخدمها الأعرابى فى فحص قطعة من الحجر ليتبين ما إذا كانت كريمة . وهو فى هذا يمسك بالقطعة الكريمة ، التى ثبت صدقها ، ويقوم بحكها (فهي محك) فى القطعة المراد اختبارها ، فإذا خدشت الأخيرة كانت غير أصيلة أى غير صادقة . ومن هنا نتبين أن المحك مقياس موضوعى ثبت صدقه نقارن به مقياسا جديدا لنبيين مدى صدقه .

الحاجة إلى المحكات :

وجدنا أننا فى حساب معامل صدق الاختبار نحتاج إلى إيجاد العلاقة بينه وبين المستوى الحقيقى للجانب الذى يقيسه الاختبار . ووجدنا أننا نصل إلى تقدير تقريبي للمستوى الحقيقى لهذا الجانب عن طريق مقياس موضوعى آخر هو المحك . ونحن إذن فى حاجة إلى المحك لمقارنة درجات الاختبار بالرتب الحقيقية للأفراد فى مجال واقعى يعتمد النجاح فيه على خاصية صمم الاختبار لقياسها . وفى الواقع نحن نعتمد على أكثر من محك واحد عند إثبات صدق الاختبار ، على أنه يجب أن تتوفر فى المحك خصائص سنعرض لها . ويجب ألا ننسى أن نواجد المحك المناسب أحيانا يكون أمراً عسيراً .

خواص المحك الجيد :

(أ) يجب أن يكون المحك متعلقاً بالوظيفة التي يقيسها الاختبار .

(ب) المحك الجيد عينة ممثلة لمنطقة السلوك المطلوب قياسها فتقدير المدرس لا يكون محكاً جيداً إذا اعتمد على مجرد إعطاء تقدير للتلميذ بناءً على موقف عارض . وكذا فتشخيص الطبيب النفسى غالباً ما يعد عينة لا تمثل كل سلوك المريض إذ أن الطبيب النفسى فى العادة أثناء التشخيص لا يعطى المريض وقتاً كافياً يلاحظ فيه سلوكه . وعلى هذا فنحن نشك فى قيمة تشخيص الطبيب النفسى كمحك .

(ج) يجب أن يكون المحك الجيد عملياً اقتصادياً . بمعنى أننا نختار محكاً ملائماً يسهل جمع البيانات عنه من حيث الوقت والجهد والمال .

(د) تتوفر فى المحك الجيد خاصية الثبات . أى أن المحك الجيد مقياس إذا قدرنا به فرداً واحداً عدة مرات كان التقدير متقارباً .

(هـ) المحك الجيد محك موضوعى . فدرجات أو رتب الأفراد على المحك يجب أن لا تتأثر بشخص من يقوم بإعطاء الدرجة أو التقدير . فلا يجب مثلاً أن نأخذ تقريراً عن عامل من صاحب عمل يقاضيه فى المحاكم .

(و) يجب أن يكون المحك مستقلاً . فلا تتأثر درجة الفرد فى الاختبار بدرجةه فى مقياس المحك . كما يجب ألا تتأثر رتبة الفرد فى مقياس المحك بدرجةه فى الاختبار . فن الخطأ مثلاً أن نطلب من المشرف على العامل أن يقدر العامل بعد أن نحيطه علماً بأن هذا العامل قد حصل على درجات مرتفعة فى اختبار الاستعداد لهذه المهنة .

أنواع المحكات :

أولاً : في الصناعة :

- ١ - كمية الإنتاج وسرعة الأداء .
- ٢ - جودة الإنتاج ودقة الأداء .
- ٣ - مقدار التلف الناشئ والفضلات المتبقية أثناء العمل .
- ٤ - عدد الحوادث والأخطاء .
- ٥ - المواظبة ودقة المواعيد .
- ٦ - الاستقرار في العمل .
- ٧ - الأجور والمكافآت والترقيات والعلاوات .
- ٨ - المسؤولية التي عهد بها إلى العامل وعضويته في جمعيات فنية أو نقابات مهنية .
- ٩ - مدى استفادة العامل ، أو تحصيله ، من العمل الذي يقوم به :
- ١٠ - السرعة التي ينهى بها فترة التدريب .
- ١١ - التقارير عن العمال أثناء فترة التدريب .
- ١٢ - تقديرات المشرفين والملاحظين للعامل .
- ١٣ - اختبارات الحرف .

ثانياً : في المدرسة :

- ١ - الامتحانات النهائية .
- ٢ - الامتحانات الشهرية والدورية وامتحانات الفترات .
- ٣ - سجلات المدرسة .
- ٤ - تقديرات المشرفين ومدرسي الفصول .
- ٥ - الاختبارات التحصيلية .

- ٦ - نشاط التلميذ خارج الفصل ودأبه : في العمل ، والملاعب ، والجمعيات المدرسية والهوايات والفرق الرياضية والنشاط الحر .
- ٧ - التاريخ المدرسي والتحصيل للتلميذ .
- ٨ - البحوث والتقارير التي يقوم بها التلاميذ .
- ٩ - درجات أعمال السنة ودرجات التمارين المعطاة داخل الفصل .
- ١٠ - الاشتراك في المناقشات والإجابة على الأسئلة وتوجيه أسئلة تدل على فهم المادة .
- ١١ - القراءات الصيفية والاشتراك في المكتبات .
- ١٢ - المحافظة على تقاليد ونظام الجو المدرسي .
- ١٣ - المعلومات العامة .

ثالثاً في المجالات الأخرى :

- ١ - تقديرات المفحوص ممن يعرفونه جيداً مثل الوالدين والإخوة والأقارب والأصدقاء والجيران و : الخ :
- ٢ - البيانات الرسمية المسجلة كعدد السوابق ونوعها بالنسبة للأحداث الجانحين .
- ٣ - التقارير من المستشفيات والعيادات ومراكز التوجيه والإرشاد :
- ٤ - تشخيص الطبيب النفسي أو تقرير الأخصائي النفسي أو الاجتماعي .
- ٥ - القروق في العمر : ففي اختبارات الذكاء والقدرات يعتمد مفهوم العمر العقلي على أن الفرد كلما تقدم في السن نما عقلياً .
- ٦ - مقارنة الفرد بالجماعة التي ينتمي إليها .
- ٧ - مقارنة الفرد أو الجماعة بجماعة معينة التقنين .

- ٨ - المجموعات المتقابلة Contrasted groups :وهنا نقارن أداء أفراد في مجموعة متفوقة في أحد الجوانب المقاسة بأفراد في مجموعة تفتقر إلى هذا الجانب . كالمقارنة بين درجات العصائين والدهانين في مقابل درجات الأسوياء عند بحث صدق اختبارات الشخصية .
- ٩ - مقارنة الاختبار الذى نبث صدقه باختبار آخر ثبت صدقه أجرى من قبل أو يجرى مع أو بعد الاختبار موضوع الدراسة .
- ١٠ - الإتساق الداخلى Internal Consistency أى معامل الارتباط بين نتيجة كل فقرة على حده مع نتيجة الاختبار ككل .

طرق حساب معامل الصدق

أولاً الصدق الظاهرى :

بالتحليل المبدئى لفقرات الاختبار لمعرفة ما إذا كانت تتعلق بالجانب المقاس . وهذا أمر يرجع إلى ذاتية الباحث وتقديره .

ثانياً صدق المضمون :

ويقاس بالتحليل المنطقى لمحتويات الاختبار ومطابقتها مع محتويات الجانب المقاس .

ثالثاً الصدق التنبؤى :

ويقاس بإيجاد العلاقة بين الدرجات على الاختبار والدرجات على مقياس المحك الذى يطبق بعد إجراء الاختبار . وتسمى هذه بالطريقة التنبؤية Follow-up method . وبعد جمع البيانات عن المحك وحساب الدرجات على الاختبار نقوم بإيجاد العلاقة بينهما بإحدى طرق ثلاث هى :

(١) طريقة النسبة المئوية Percentage method وذلك بتقسيم الأفراد على أساس رتبهم فى مقياس المحك إلى قسمين متقابلين ، كالناجحين فى مقابل

الفاشلين مثلاً ، وحساب النسب المئوية من الأفراد الذين حصلوا على درجات مرتفعة أو منخفضة في الاختبار في كل من المجموعتين . وهكذا يكون أمامنا جدول كالآتي :

المحك	الناجحون	فاشلون
صفر	صفر %	٥ %
١	٣ %	١٠ %
٢	٥ %	١٠ %
٣	٤ %	٢٤ %
٤	٦ %	٢٠ %
٥	٤٠ %	١٥ %
٦	٢٠ %	٨ %
٧	٨ %	٥ %
٨	٦ %	٢ %
٩	٦ %	١ %
١٠	٤ %	صفر %

ومن هذا الجدول يتضح لنا أن المجموعة الناجحة وعددها ١٠٠ شخص تميل إلى أن تحصل على درجات أعلى من المجموعة الفاشلة وعددها ١٠٠ شخص ومن هذا الجدول يمكن عمل خريطة توقعات expectancy chart بمجرد النظر إلى درجة الفرد نحدد احتمال نجاحه مستقبلاً ولكي نتأكد من هذا نقوم بحساب دلالة الفرق بين درجات المجموعتين من الأفراد فإذا كان الفرق جوهرياً دل هذا على أن الاختبار صادق إذ أن الاختبار نجح في أن

يعطينا درجات بنينا على أساسها توقعات أو تنبؤات بالنجاح أو الفشل ثبت صدقها فن حصل على درجات مرتفعة في الاختبار سيكون غالباً ناجحاً ومن يحصل على درجات منخفضة فيه غالباً ما يفشل مستقبلاً .

(ب) طريقة المتوسطات :

طريقة المتوسطات Average method تقوم على حساب دلالة الفرق بين درجات مجموعتين من الأفراد في الاختبار ، إحداهما أخذت تقديرأ مرتفعاً في مقياس المحك والأخرى أخذت تقديرأ منخفضاً في مقياس المحك أيضاً . فإذا ثبت أن هناك فرق جوهري بين درجات هاتين المجموعتين في الاختبار كان الاختبار صادقاً ، إذ أنه يقيس سمة يعتمد العمل مستقبلاً عليها .

(ج) طريقة معامل الارتباط :

طريقة معامل الارتباط Correlation method تعتمد على إيجاد معامل الارتباط بالطريقة الثنائية Biserial أو معاملاته بين درجات الأفراد على المحك ودرجاتهم على الاختبار . وهذه الطريقة أدق من الطريقتين السابقتين ، إذ أنها تعتمد على درجات كل الأفراد لا على متوسطهم أو نسبهم . كما أنها تمكنا من مقارنة درجات الأفراد في أكثر من اختبار خاصة إذا أردنا أن نحدد الأهمية التنبؤية لكل اختبار نستخدمه ضمن بطارية مثلاً ، غير أن طريقتي النسب المئوية والمتوسطات واضحتان في عرض نتائجهما ومقنعتان لغير المتخصص في الإحصاء والقياس من أمثال إدارة المدارس أو التعليم وأصحاب العمل .

رابعاً الصدق التلازمي : ويقاس هذا النوع من الصدق بإيجاد مدى ارتباط درجات الأفراد على الاختبار مع درجاتهم على مقياس المحك ، على

أن تجمع البيانات عن مقياس المحك وقت إجراء الاختبار . وتسمى هذه الطريقة بطريقة المستخدمين الحاليين The Present Employee method ، وفيها نطبق الاختبار على العمال أو التلاميذ الموجودين بالمصنع أو المدرسة ثم نجمع البيانات عن تقديراتهم على المحك وليكن كمية الإنتاج أو التحصيل ثم نحلل النتائج بطريقة النسب المئوية أو المتوسطات أو معامل الارتباط .

خامساً : الصدق التجريبي :

ويقاس بالطريقتين السابقتين معاً ، وهما الطريقة التتبعية وطريقة العمال الحاليين ، ما دام مفهوم الصدق التجريبي يتضمن مفهومي الصدق التنبؤي والصدق التلازمي . والحق أنه برغم أن حساب الصدق التجريبي يعتمد على الطريقتين معاً ، أى يحسن الجمع بينهما ، إلا أن ثمة مقارنة يمكن أن تعقد بينهما . على أن طريقتي التتبع والعمال الحاليين تستخدمان في مجالات كثيرة وليس في الصناعة وحدها ، ما دامت طريقة الإجراء هي هي وكذا ما دام تحليل النتائج ، بعد جمع البيانات على المحك ، تتبع الطرق الثلاث السابق ذكرها وهي طريقة النسبة المئوية وطريقة المتوسطات وطريقة الارتباطات .

فن حيث الدافع ؛ يكون المفحوص في الطريقة التتبعية مدفوعاً إلى جو العمل أو الدراسة الذي سيشارك فيه ، على عكس المفحوص في حالة المستخدمين الحاليين الذي يكون قد اطمأن على وضعه في العمل أو الدراسة والذي يعرف أن شيئاً من الاختبار لن يؤثر في مركزه .

ومن حيث كسب ثقة وتعاون المفحوصين في موقف الإجراء نجد أن الأفراد في حالة التتبع يحاولون الظهور أمام المختبر بمظهر المتعاون إذ أن هذا ، في نظرهم ، قد يؤثر على قبولهم بالمهنة أو الدراسة . بينما لا يهم المفحوص الذي يمارس العمل أو يواظب على الدراسة أن يبدو بمظهر حسن :

والطريقة التتبعية تمتاز ، فضلاً عن تقوية الدوافع وكسب ثقة وتعاون
المفحوصين ، بأنها تمكن الباحث من الملاحظة الدقيقة لسلوك الأفراد في
الفترة بين إجراء الاختبار وجمع البيانات عن المحك . فالباحث هنا يستطيع
أن يضع فروضاً وأن يقوم بتحقيقها على مهل وفي دقة خارجاً باستنتاجات
لا يستطيعها إذا كان بصدد عمال أو طلبة يقومون بالعمل فعلاً .

ولكن الطريقة التتبعية تستغرق وقتاً أطول . فقد يكون العمل
أو الدراسة في حاجة إلى الوصول إلى نتائج مباشرة وسريعة عن
يقومون بالعمل فعلاً ، وهنا يتعين على الباحث أن يتبع طريقة القائمين
بالعمل حالياً .

إن هناك صعوبات عملية تعترض طريقة العمال الحاليين . إذ أنه يلزم
لتطبيق الاختبار أن يتوقف العمل أو الدراسة فترة تطول بحسب زمن
الإجراء وقد يستحيل هذا في بعض الصناعات . بينما نجد أن إجراء
الاختبار في الطريقة التتبعية يكون أحد خطوات الاختيار أي أنه يتم قبل
الالتحاق بالعمل فعلاً .

كما أن الفرد الذي يقوم بالعمل فعلاً ، سواء في الصناعة أو التعليم ،
يكتسب خبرة ويحصل معلومات ويدرب قدرات وينمي مهارات تفسد
اختبار بعض جوانب السلوك كالاستعدادات والميول . وهنا يكون الباحث
بصدد قياس التحصيل أو الكفاية وليس الاستعداد .

وهكذا نجد أن لكل من الطريقتين عيوب ومزايا ، إلا أن الباحث
يتحدد بالفرض من إجراء الاختبار عندما يختار إحدى الطريقتين وإن كان
يحسن أن يجمع بينهما .

سادساً صدق المفهوم :

ويعتمد حسابه على تفسير نتائج الاختبار عن طريق النظرية التي وضع

المقياس بناءً عليها ، أى أنه فى نهاية الأمر يرجع الفروق بين درجات الأفراد على الاختبار إلى اختلاف مستوياتهم فى الجانب الذى تعالجه النظرية وقيسه الاختبار .

سابعاً الصدق التطابقى :

وهو أحد طريقتى حساب صدق المفهوم ويكون بحساب مدى اتفاق درجات الأفراد على الاختبار الجديد ودرجاتهم على اختبار آخر ثبت صدقه . ويقترح جيليكسون Gulliksen إجراء عدة خطوات عملية موضوعية لقياس صدق المفهوم بالطريقة التطابقية . يذكر ضمنها طريقة إجراء الاختبار الذى يقيس قدرة معينة ثم تنظم برامج لتدريب هذه القدرات لدى من أجرى عليهم الاختبار ثم إعادة إجراء الاختبار بعد انتهاء التدريب . فإذا ظهر فرق كبير ذو دلالة كان الاختبار يقيس القدرة التى دربت أى كان الاختبار يقيس ما وضع لقياسه فعلاً . ومن أهم طرق حساب الصدق التطابقى ، وقد ذكرها جيليكسون أيضاً ، إجراء عدة اختبارات تقيس منطقة واحدة من السلوك ثم عقد المقارنات بين درجات الأفراد على هذه الاختبارات ومن بينها الاختبار الجديد . فإذا كان هناك معامل ارتباط (بطريقة الارتباط المتعدد Multiple correlation) كبير مع الاختبارات الأخرى الصادرة كان الاختبار الجديد صادقاً .

وتفيد طريقة إجراء الاختبار قبل التدريب Pre-training performance وإعادة الإجراء بعد التدريب Post-training performance فى اختبارات القدرات والتحصيل والكفاية والمهارة . ولكن هناك انتقادات عدة توجه إلى استخدام هذه الطريقة فى إثبات صدق الاختبار ما دامت درجات القدرة التى دربت قد زادت ، وأهم هذه الانتقادات هى :

هل ترجع زيادة درجات الأفراد فى الإجراء بعد التدريب إلى التدريب

وحده ؟ صحيح أن التدريب نمي القدرة أو التحصيل أو الكفاية ولكن ألم ينمو الفرد نمواً معرفياً وعقلياً ونفسياً أثناء فترة التدريب بدرجة تؤثر على أداء الاختبار مرة ثانية وتتداخل مع التدريب ؟ بمعنى أنه لو لم يعطى الأفراد تدريباً معيناً فإنهم عند إعادة الإجراء يكونون قد نمواً في مجالات متعددة فترتفع الدرجات التي يحصلون عليها برغم عدم حصولهم على تدريب رسمي formal . وهنا يكون إرجاع الزيادة الجوهرية الكبيرة في الدرجات إلى ما أحدثه التدريب وحده أمر متعسف .

كما أن إعادة إجراء الاختبار ، في ذاتها وبدون افتراض حدوث نمو عقلي ونفسي ومعرفي ولوجدلا ، تتأثر بعوامل التذكر والحفظ والاعتقاد على موقف الاختبار ، الأمر الذي يرفع درجات الأفراد . وكأن ارتفاع درجات الأفراد ، عند إعادة الإجراء بعهد التدريب أو إعطاء منهج دراسي ينمي قدرة أو جانباً يقيسه الاختبار ، لا يعود إلى فعل التدريب وحده بل وأيضاً يرجع إلى النمو development والتذكر ، وهكذا لا نتأكد أن الاختبار يقيس هذه القدرة أو الجانب وحده .

كما أن هناك مجالات من السلوك لا تجدى فيها طريقة التدريب كإثبات للصدق ، فقياس الشخصية مثلاً لا يمكن أن يعتمد على هذه الطريقة . إذ أن الباحث لا يستطيع بسهولة أن ينمي « العصبية » أو « الانبساط والانبساط » مثلاً .

أما عن طريقة إيجاد معامل الارتباط بين درجات الاختبار الجديد ودرجات الاختبارات التي ثبت صدقها من قبل فإنها طريقة مضیعة ، إذ أنه لا يكون هناك مبرر لإنشاء اختبار جديد ما دام مرتبط باختبارات ثبت صدقها أي ما دام سيقاس ما قاسته هي من قبله بدقة قد يصعب عليه تصحيحها . فارتناح معامل الارتباط هنا يعني أن الاختبار الجديد صادق فهو يقيس ما تقيسه الاختبارات ، ولكنه يعني أيضاً أن الباحث

أحياناً لا تكون أهمه مشكلة في قياس الجانب المطلوب بحيث يكون مدفوعاً لإنشاء اختبار خاص وأن إنشاء الاختبارات وإثبات صدقها أمر غاية في المشقة

ثامناً الصدق العاملى :

هو الطريقة الثانية ، بعد طريقة التطابق ، في حساب صدق المفهوم ، وهو يعتمد على منهج التحليل العاملى Factor analysis . وبها نحدد مدى قياس عدة اختبارات لبعض العوامل المشتركة ، أى أننا نحدد مدى تشبع هذه الاختبارات بتلك العوامل . ونحتاج هذه الطريقة في حسابها إلى الحصول على عدد من الاختبارات الصادقة في قياسها لجوانب معينة ثم حساب معاملات الارتباط بين كل اختبارين من هذه الاختبارات . وهكذا يحصل الباحث على مصفوفة معاملات الارتباط Correlation matrix يمكن عن طريقها حساب درجة تشبع الاختبار بالعوامل المشتركة بين هذه المقاييس أو بين مجموعة منها . أى أن الصدق العاملى لأى اختبار هو معامل الارتباط بين هذا الاختبار وبين عامل أمكن استخلاصه بواسطة التحليل العاملى . أو هو العامل المشترك بين الاختبار وأى اختبار آخر أو مجموعة من الاختبارات تقيس نفس الوظيفة .

وتاريخ التحليل العاملى يدلنا على مدى تطور طرق الإحصائية . وليس هذا مجال التعرض لها . ولكن تطبيق منهج التحليل العاملى في إثبات صدق الاختبار يعتمد على طريقتين . أولاهما هي الطريقة التقليدية وتكون ، كما بينا ، بإدخال الاختبار الجديد مع مجموعة من الاختبارات . وهكذا يحصل الباحث على مصفوفة ارتباطات ، بين الاختبارات ، يحللها فيشتق عدة معاملات صدق عاملى .

ولكن لهذه الطريقة عيوب كثيرة أهمها أنه كلما كثر عدد معاملات

الصدق العاملي لاختبار واحد ، كان هذا الاختبار غير صادق إذ أنه يكون هنا مشبعاً بعدة عوامل أى أنه لا يقيس العامل الذى صمم لقياسه وحده بل يقيس إلى جواره عوامل أخرى . وهكذا يكون مضمونه موزعاً ولا يكون المقياس نقياً .

ويمكن معالجة هذا النقص بوضع تصميم يودى بالباحث إلى أن يكشف عن عدد من العوامل لا يزيد عن ثلث عدد الاختبارات المطبقة معا ، أى يجب أن تكون نسبة العوامل : الاختبارات = ١ : ٣ . وهكذا يتمثل كل عامل فى ثلاثة اختبارات على الأقل . كما أن من شروط تحسين هذه الطريقة أن لا يقل حجم العينة عن ٢٠٠ فرد بأى حال . وهناك مشكلة أخرى وهى تحديد مستوى الثقة فى معامل الارتباط . أى ما هو الحد الأدنى لمعامل الارتباط الذى يكفى لأن يكون جوهرياً أو ذا دلالة ؟ الواقع أن هذا التساؤل غير مقطوع له بإجابة . ولكن هناك شبه اتفاق على أن معامل الارتباط هنا إذا زاد عن ٠.٣ يكون ذا دلالة .

والطريقة الثانية فى حساب صدق الاختبار بالمفهوم العاملي تكون بإيجاد معاملات الارتباط بين فقرات الاختبار الواحد واشتقاق العوامل المشتركة بين هذه الارتباطات بين فقرات الاختبار الواحد (ويستخدم معامل الارتباط الرباعى Tetracoric Correlation) فإذا وصل الباحث إلى أن اختباره مشبع بعامل واحد فقط وأن ما دون هذا العامل بقايا ، يكون الاختبار صادقا . ولكن مفهوم الصدق هنا قريب جداً من مفهوم التجانس Homogeneity إذ أنه يصل إلى معرفة العوامل المشتركة بين الفقرات ولكنه لا يتأكد من طبيعة هذه العوامل ولا مما إذا كانت هذه العوامل هى التى قصد قياسها أم لا .

وهناك طريقة ثالثة فى حساب الصدق العاملي ، توصل إليها تياور Taylor يقوم فيها الباحث بإيجاد معاملات الارتباط بين الدرجة على كل فقرة على حدة وبين الدرجة على بقية الفقرات معاً . وهنا يقترب مفهوم الصدق العاملي

أيضاً من مفهوم التجانس . وهذه الطريقة تفضل كثيراً طريقة حساب معاملات الارتباط بين درجات الأفراد على كل فقرة مع كل فقرة أخرى على حدة التي تتطلب حساب عدد كبير جداً من معاملات الارتباط (فإذا فرضنا أن عدد فقرات الاختبار الواحد هي ١٠٠ يكون عدد معاملات الارتباط بالطريقة الثانية هو $\frac{99+100}{2} = 49.5$ في حين أننا بالطريقة الثالثة نوجد ١٠٠ معامل ارتباط) . كما أن الطريقة الثانية تعتمد على إيجاد معاملات الارتباط بين كل فقرة والأخرى ، ومن المقطوع به أن معامل ثبات الفقرة الواحدة غالباً منخفض بدرجة تجعل الاعتماد على هذه الطريقة مشكوكاً فيه .

وبالرغم من أننا بالصدق العاملي نستطيع الوصول إلى تنقية الاختبارات وإعداد اختبارات نقية ، إلا أن ثمة صعوبات منهجية تعترضنا . فأحياناً نوجد الارتباط بين الاختبار ومقياس محك من نوع كالتقديرات أو تشخيص الطبيب النفسى : وهنا يجب أن نحلل الاختبار تحليلًا عاملياً ثم نقوم بالتحليل العاملي لمقياس المحك وبعد ذلك نطابق بين العوامل التي كشفنا عنها في الاختبار والعوامل التي كشفنا عنها في المحك حتى نستطيع أن نعتمد على هذا المحك . فكأننا نجرى هنا عمليات التحليل العاملي في مرحلتين إحداهما عسيرة جداً وتخضع للذاتية في ذلك في حالة المحك . بينما في حالة الاختبارات التي تستخدم كمحك نقوم بعمليات التحليل العاملي مرة واحدة وبطريقة موضوعية .

معامل الصدق معامل ارتباط

قلنا إن معامل الصدق هو تطبيق لمعامل الارتباط الذي نوجده بين الاختبار والمحك . وإليك بيان بأنواع معاملات الارتباط المستخدمة في حساب الصدق :

(أ) معامل بيرسون Pearson : يستخدم في حالة المتغيرين المتصلين أي عندما تكون الدرجات على الاختبار وعلى مقياس المحك مستمرة لا انقطاع فيها continuous .

(ب) معامل الرتب Ranking : يستخدم معامل الرتب لسيرمان إذا كان الاختبار يعطى رتباً أو درجات نحوها إلى رتب وكذا المحك .

(ج) معامل الارتباط الثنائي Biserial : إذا كان المحك منفصلاً Dicotomus مثل محك التشخيص إلى سوى وعصبي على أن تكون درجات الاختبار متصلة :

(د) معامل الارتباط الرباعي Tetrachoric : إذا كانت الدرجات على الاختبار وعلى المحك متقطعة .

(هـ) معامل الارتباط المتعدد Multiple correlation : يستخدم في حالة إيجاد معامل الارتباط بين درجات الاختبار ودرجات عدة محكات أخرى . كما يستخدم في حالة تكوين البطاريات من عدد من الاختبارات .

تصحيح معامل الصدق :

قلنا إن معامل الصدق معامل ارتباط بين الاختبار والوظيفة التي يقيسها ، وأن دليلنا على مستوى هذه الوظيفة يكون غالباً مقياس المحك . وسنرى في الفصل السابع ، عند مناقشة الثبات ، أن أي قياس يشمل ، إلى جوار قياس القدرة المقصودة ، درجة من الخطأ التجريبي في القياس . وعلى هذا يأتي تباين (ع ٢) أو اختلاف درجات الأفراد في الاختبار من مصدرين ؛ أولهما هو الفروق الحقيقية بين مستويات قدرة الأفراد في الوظيفة المقاسة . وثانيهما التباين الناشئ عن الخطأ التجريبي في القياس . هذا الخطأ ، وبالتالي هذا التباين ، يكون في كل اختبار أو أداة قياس على حدة . فإذا حسبنا معامل الصدق بإيجاد معامل الارتباط بين

لإختبارنا وإختبار آخر يقيس نفس الوظيفة أو بينه ومقياس المحك لهذه الوظيفة ، كان لكل أداة خطأ فى القياس . وعلى هذا فنحن ندخل الخطأين معاً : وهكذا يزيد الخطأ التجريبي فى القياس ، وبالتالي فى حساب معامل الصدق ، بتداخل Overlapping الخطئين . وهذا لا يتحقق إذا كان معامل الارتباط بين كل من الاختبارين ونفسه (الثبات) كاملاً أى + ١ (وهذا لا يحدث عملياً) وهكذا يتأثر معامل الصدق بمعامل الثبات ، فضلاً عن تأثره بتباين الخطأ التجريبي . وعليه ، لا يكون معامل الارتباط بين الاختبار والمحك معامل صدق حقيقى true فهو يتضمن قدراً من الخطأ التجريبي . فإذا أردنا أن نحسب معامل الصدق الحقيقى أى أن نتقيته من الخطأ التجريبي ، ولذلك تسمى طريقة التصحيح للتقية Correction for Attenuation ، نطبق المعادلة التالية :

$$\frac{\text{معامل صدق س. ص}}{\sqrt{\text{معامل صدق س. ص} \times \text{معامل ثبات س. ص}}} = \text{معامل صدق حقيقى س. ص}$$

حيث (معامل صدق س. ص) هى معامل الصدق الحقيقى للاختبار (س) أى
هى معامل الارتباط بين الاختبار وبين المحك الخارجى (ص) الذى قد
يكون اختباراً :

معامل صدق س. ص هى معامل الصدق التجريبي الناتج من البحث وليكن (٠.٧٠)

معامل ثبات س. ص هى معامل ثبات الاختبار (س) وليكن (٠.٩٠)

معامل ثبات المحك (ص) وليكن (٠.٨٠)

ويكون معامل الصدق المصحح فى هذه الحالة هو :

$$\frac{0.7}{0.72\sqrt{}} = \frac{0.70}{0.80 \times 0.90\sqrt{}}$$

$$0.82 = \frac{0.7}{0.8487} =$$

ويتأثر معامل صدق الاختبار بطوله والمعادلة هي :

$$\frac{\text{م.س. ص}}{\sqrt{1 - \frac{\text{م.س. ص} + \text{م.س. ص}}{\text{ط}}}} = \text{م.س. ص}$$

حيث (ط) هي طول الاختبار الجديد وليكن (١٥٠ فقرة) بالنسبة
لطول الاختبار القديم وليكن ٥٠ فقرة . ∴ ط = ٣٠

، (س) هو الاختبار :

، (ص) هو المحك :

وتكون (م.س. ص) هي معامل الصدق بعد مضاعفة طوله ثلاثة أو أي

عدد من المرات وليكن (ط) :

، (م.س. ص) هي معامل ثبات الاختبار وليكن ٠٧٠

، (م.س. ص) هي معامل ارتباط الاختبار القديم مع المحك وليكن ١

(٠٦٤) :

ويكون معامل صدق الاختبار بعد جعل طوله ثلاثة أمثال الطول .

الأصلي :

$$\frac{0.64}{0.7 + \frac{0.7 - 1}{3}} \sqrt{\quad} =$$

$$\frac{0.64}{0.7 + \frac{0.3}{3}} \sqrt{\quad} =$$

$$\frac{0.64}{0.7 + 0.1} \sqrt{\quad} =$$

$$\frac{0.64}{0.8} \sqrt{\quad} =$$

$$0.72 = \frac{0.64}{0.89} =$$

وإذا أراد الباحث أن يحدد مدى طول الاختبار الجديد بالنسبة للاختبار القديم إذا كان يريد رفع معامل صدقه إلى قيمة معينة كان عليه أن يستعمل المعادلة التالية :

$$\frac{\text{مس. ط.}^2 (1 - \text{مس.})}{\text{مس. ط.}^2 - \text{مس.} \times \text{مس.}} = \text{ط.}$$

حيث ط طول الاختبار الجديد. بالنسبة للاختبار الأصلي (مس)

، مماس ص معامل الصديق المطلوب وأن مماس ص هي معامل صديق الاختبار الأصلي وأن مماس ص هي معامل ثبات الاختبار الأصلي وهي نفس الرموز المستعملة في المعادلة السابقة .

وبالتعويض في هذه المعادلة بالمقادير المستعملة في المثال السابق نجد أن

$$\text{ط} = \frac{(0.72)^2 (0.7 - 1)}{(0.7)^2 (0.72) - (0.764)^2}$$

$$= \frac{(0.3) (0.52)}{(0.7) (0.52) - (0.41)^2}$$

$$= \frac{0.156}{0.05}$$

$$= 3 \text{ تقريباً}$$

أى أن طول الاختبار المطلوب هو ٣ أمثال الاختبار الأصلي :

إستخدام الصديق في حالات أخرى

(١) التنبؤ بالحالات الفردية :

إذا قلنا إن معامل صديق الاختبار هو ٠.٧ وكنا نعى الصديق التنبؤي وكان هذا المعامل هو معامل الارتباط بين الاختبار الجديد واختبار صادق قديم فما معنى هذا ؟ إلى جوار كل المعاني التي سبق التعرض لها يدلنا هذا الرقم على أننا في ٧٠ ٪ من الحالات نصدر حكماً تنبؤياً سليماً باستخدام هذا

. اختبار . أى أنه لو كنا بصدد اختيار تلاميذ وتقديم لنا ١٢٠٠ تلميذ واخترنا باستخدام الاختبار ١٠٠٠ تلميذ كان تنبؤنا بنجاح ٧٠٠ تلميذ منهم تنبؤاً صحيحاً . ولكننا إذا كنا بصدد حالات فردية كان الأمر أشق وكانت هناك مخاطرة فى الحكم على الفرد بهذا الاختبار أو بأى اختبار آخر ، إذ أننا نكون معرضين للخطأ بإحتمال ٣٠ ٪ . هذا فضلاً عن أنه أثناء عمليات القياس المختلفة نكون قد تعرضنا لخطأ نسميه الخطأ المعيارى .

فإذا عرفنا هذا الخطأ المعيارى للدرجة التى أعطيناها للفرد أمكننا أن نضع حداً أعلى لا تزيد درجة الفرد الحقيقية عنه وحداً أدنى لا تقل عنه . هذا فى أغلب الأحوال خاصة وأن معامل الارتباط ، الذى نحسبه فى إيجاد أى معامل صدق ، يحسب من درجات عينة الثقلين وليس من درجة فرد واحد مما يجعل التنبؤ بدرجة الفرد عرضة للخطأ . فإذا فرضنا أن :

الخطأ المعيارى لدرجات المحك هو $\sigma_{\text{محك}}$

الخطأ المعيارى للدرجة الفرد هو $\sigma_{\text{فرد}}$

معامل الارتباط بين الاختبار والمحك هو r

باعتبار أن $\sigma_{\text{فرد}}$ ترمز إلى الاختبار ، $\sigma_{\text{محك}}$ ترمز إلى المحك .

إذن $\sigma_{\text{محك}}$ هى معامل الصدق .

وهكذا يكون الخطأ فى تقدير الدرجة ورمزه $\sigma_{\text{تقدير}}$ مساوياً للمقدار الآتى

$$\sigma_{\text{تقدير}} = \sigma_{\text{فرد}} \sqrt{1 - r^2}$$

مثال : فى الاختبار $\sigma_{\text{فرد}}$ حصل أحد المفحوصين على الدرجة ١٠ ووجدنا

أن الخطأ المعيارى لدرجات المحك هو ٠,٢٥ وكان معامل الارتباط بين المحك

والإختبار (أى معامل صدق الإختبار) هو ٠,٨ فتكون درجة الفرد الحقيقية فى هذا الإختبار مستبعدين بذلك كل أخطاء القياس والمقياس محصورة بين ما يلى :

$$ع_n = ع_v \sqrt{1 - r_{sv}^2}$$

$$= ٠,٢٥ \sqrt{1 - (٠,٨)^2}$$

$$= ٠,٢٥ \sqrt{1 - ٠,٦٤}$$

$$= ٠,٢٥ \sqrt{٠,٣٦}$$

$$= ٠,٢٥ \times ٠,٦$$

$$= ٠,١٥ \text{ من الدرجة}$$

وحيث أن درجة الفرد هى ١٠ فيكون الخطأ فى تقديرها $١٠ \times ٠,١٥ = ١,٥$ إذن درجة الفرد الحقيقية تنحصر بين $١٠ + ١,٥$ و $١٠ - ١,٥$ أى بين ٨,٥ ، ١١,٥

(ب) فى التصفية :

فى التصفية تستخدم جداول تيلور ورسيل Taylor and Russel التى تحدد النسبة المثوية التى يمكن إختيارها من بين المتقدمين على أساس صدق الإختبار ونسبة الناجحين من الأفراد الذين اخترناهم دون الاعتماد على الإختبار .

(ح) فى تكوين البطاريات :

تتكون البطاريات من عدد من الإختبارات التى تقيس عدة جوانب تتعلق بعمل أو مهنة أو دراسة أو . . الخ . ولكن يجب أن تقيس هذه

الإختبارات عدة جوانب لا تختلف حتى تتعارض ولا تقترب حتى تتطابق ؛
ويحصل الفرد على درجة كلية على البطارية أو على درجات في كل إختبار
منها . ويجب أن تحدد أوزان Weights أى مقادير بمعادلات التصحيح لكل
إختبار حسب أهمية الجانب الذى يقيسه بالنسبة لمجموع الجوانب التى تقيسها
البطارية . وهكذا تكون للفرد الواحد عدة درجات تمثل بياناً فى صفحة
نفسية Psychoprofile تعبر عن صورة كلية لأحد مجالات السلوك أو
الحياة النفسية .

الفصل السابع

الثبات

- مقدمة .
- إعتبارات منطقية في تقدير الثبات .
- الثبات وتحليل التباين .
- موضوعية القياس .
- عمليات حساب الثبات .
- تفسير معامل الثبات .
- العوامل المؤثرة على ثبات الاختبار ..
- معامل الثبات ومعامل الصدق ..
- بيانات الثبات .

الفصل السابع

الثبات

مقدمة :

يقوم التعريف بأى إختبار على ذكر معلمين أساسيين فيه هما ثباته وصدقه . وكلمة « إختبار ثابت » كترجمة للمصطلح « Reliable » ليست كافية تماماً . فالأصل الأجنبى للكلمة يدل على أن هذا الإختبار « موثوق فيه ويعتمد عليه » وهذا هو المقصود غالباً بكلمة ثابت هنا . وقد يكون هذا المعنى هو النتيجة التى تصل إليها باستخدام إختبار يعطى تقديرات ثابتة . أى أنه فى عملية القياس لو كررنا الإجراء لحصلنا على نتائج متسقة Consistant عن الفرد ، بمعنى أن درجته لا تتغير جوهرياً بتكرار إجراء المقياس عليه ، أو بمعنى أن موقف الفرد بالنسبة لجماعته لا يتغير جوهرياً بين مرات الإجراء المتعددة . وهنا يقرب معنى الثبات من مفهوم الإتساق ، لإتساق نتائج إجراء الإختبار مع نفسها بين مرات الإجراء المتعددة .

والثبات Reliability قد يعنى الإستقرار Stability ، بمعنى أنه لو كررت عمليات قياس الفرد الواحد لأظهرت درجته شيئاً من الإستقرار . كما أن الثبات قد يعنى الموضوعية objectivity بمعنى أن الفرد يحصل على نفس الدرجة كائناً من كان الأخصائى الذى يطبق عليه الإختبار أو الذى يصححه . وهنا يكون الإختبار الثابت مقياساً يقدر الفرد تقديرأ لا يختلف فى حسابه إثنان ، تماماً كالترموتر الذى لا يختلف فى قراءته إثنان :

ومعامل الثبات هو معامل الإرتباط بين درجات الأفراد على الإختبار بين مرات الإجراء المختلفة ، أو بين تقديرات من يقومون بتصحيح الإختبار فى المرات المختلفة ، أو بين نتائج إجراء الإختبار على مجموعة واحدة من

الأفراد على أن يقوم بالإجراء إحصائيون مختلفون . أى أن معامل الثبات هو معامل الارتباط بين الاختبار ونفسه . فنحن نحسب معامل الثبات من معامل الارتباط بين الدرجات على الاختبار في مرات الإجراء المختلفة سواء كنا نعيد الإجراء بنفس الاختبار بنفس الصورة Form أو بصورة مكافئة لها Equal form . ويتضح قولنا أن معامل الثبات هو معامل الارتباط بين الاختبار ونفسه في أحد طرق حساب الثبات ، ونقصد عند حساب الارتباط بين درجتى الفرد الواحد على نصفى الاختبار . وسيتضح هذا عند مناقشة طرق حساب معامل الثبات .

ومعامل الثبات ، بما أنه معامل ارتباط ، لا يكون كاملاً أو لا شىء . it is not all or none . أى أن معامل الثبات نسبي relative فهو يختلف تبعاً لمتغيرات عدة سنذكرها فيما بعد . كما أن معامل الثبات لا يصل إلى ١ صحيح كما لا يصل إلى صفر . والمشكلة هي نفس المشكلة التي صادفتنا في معامل الصدق : متى يكون معامل الثبات كافياً للإعتماد على نتائج الاختبار ؟ ليست هناك قاعدة تحدد قيمة معامل الثبات الكافى . ولكن لاشك في أن قيمة معامل الثبات مع غيرها من خصائص الاختبار وأهمها الصدق ، هي التي تحدد درجة دقة الحكم ومدى الاعتماد على الاختبار . ومن هنا نتبين أهمية معرفة قيمة معامل ثبات الاختبار الذى نستخدمه .

وأياً كان الجانب الذى نقيسه ، علماً أو لغة أو حساباً ، نكون معرضين لتدخل عامل الصدفة . ويحوى الاختبار خطأ الصدفة Chance error بدرجات متفاوتة ، ولكنه لا يخلو من خطأ الصدفة هذا أبداً . ويختلف خطأ الصدفة باختلاف الفاحص والمفحوص وموضوع القياس . وإذا كان خطأ الصدفة صغيراً زاد معامل ثبات القياس . وإذا زادت نسبة أخطاء الصدفة باختلاف الأفراد قل ثبات القياس .

قلنا إن الثبات نسبي وكنا نقصد بمدى اختلاف المفحوصين عن أنفسهم بين مرات الإجراء المختلفة . فنحن نحسب معامل الثبات بناءً على ارتباط

درجات مجموعة من الأفراد بنفسها أى درجاتهم فى المرة الأولى مقارنة بـدرجاتهم فى المرة الثانية بنفس الصورة أو بصورة مكافئة من الاختبار . وهنا نكون معرضين لخطأ . هذا الخطأ هو أن هؤلاء الأفراد أنفسهم يتغيرون . وعلينا أن نعرف مدى تغيرهم حتى نعرف سبب اختلاف درجاتهم فى المرات المختلفة فقد لا يكون التغير فى الاختبار بل فيمن يطبق عليهم الاختبار .

وهناك ثبات مطلق Absolute . وتحدد درجة هذا الثبات المطلق بمدى التغير الحقيقى فى نتائج إجراء الاختبار على نفس الفرد . وتسجل درجات الفرد فى كل مرة ثم يحسب الانحراف المعيارى لها ويسمى هذا الانحراف بالخطأ المعيارى للقياس Standard Error of Measurement .

إعتبرات منطقية فى تقدير الثبات

يتضمن تقدير ثبات أداة القياس نوعين من العمليات ، عمليات تجريبية وأخرى إحصائية . فمن ناحية يلزمنا أن نجرى المقياس على مجموعة محددة من الأفراد متبعين خطة تجريبية محددة ومحافظين على شروط التجريب المحددة . ومن ناحية أخرى ، تحال درجات الأفراد فى هذا الإجراء باستخدام أنسب الطرق الإحصائية التى تلائم مميزات ثبات الاختبار .

وتتضمن الخطوات التجريبية فى تقدير الثبات جوانب منطقية تتعلق بمشكلة القياس ، بحيث أنه يجب علينا أن نحلل الأغراض التى نحسب ثبات المقياس من أجلها ، فيجب أن تخطط العمليات التجريبية وأن تقيم من خلال هذه الأغراض :

الثبات وتحليل التباين:

عند إجراء أداة القياس على مجموعة من الأفراد وإعطاء أفرادها درجات ، نجد أن توزيع هذه الدرجات يشمل مدى كبيراً وسبب وجود

هذا المدى من الاختلاف ، أو انحراف الأفراد عن متوسطهم أى التباين Variance ، يرجع إلى عدة عوامل متداخلة . وبعض هذه العوامل يرجع إلى وجود فروق حقيقية بين أفراد الجماعة فيما يتعلق بالصفة التى نقيسها . والبعض الآخر من العوامل يشير إلى عدم الدقة فى قياس الأفراد كل على حدة . وتقدير ثبات أى مقياس يعتمد على تحديد مدى اختلاف درجات الأفراد للفروق الحقيقية بينهم ، ومدى دقة قياس الأفراد المعينين .

وقد إشتق الإحصائيون معاملات تفيد فى وصف مدى تغير أو اختلاف مجموعة من الدرجات . وأفضل هذه المعاملات هو معامل التباين ويرمز إليه بالرمز (ع^٢) . ويعرف بأنه متوسط مربعات انحرافات درجات جماعة من الأفراد عن متوسط هذه الجماعة وهو يساوى $\left(\frac{\sum x^2}{n}\right)$. حيث ح. هى الفرق بين درجة الفرد ومتوسط الجماعة ، $\bar{x}$ هى عدد أفراد الجماعة .

ومعامل التباين له ميزة إمكان جمع تباين الجماعات الصغيرة للحصول على تباين الجماعة الكلية وكذلك طرح تباين جماعه جزئية من الجماعة الكلية للحصول على تباين الجماعة الكلية الجديدة بعد حذف هذه الجماعة الجزئية منها ، كما يمكن جمع تباين نفس المجموعة من الأفراد بالنسبة لخاصية معينة على تباينهم بالنسبة لخواص أخرى . فمثلا يمكن حساب التباين فى المجموع الكلى للدرجات لفرقة من انتلاميذ بجمع تباينهم فى مواد مختلفة مثل الحساب والعلوم واللغة . . الخ فيكون التباين الكلى هو :

$$ع^2 = ع_1^2 + ع_2^2 + ع_3^2 + \dots$$

حيث ع^٢ هى تباين المجموعة بالنسبة للحساب ، ع^٢ تباينهم بالنسبة للعلوم ، ع^٢ تباينهم بالنسبة للغة . . الخ وحيث ع^٢ هى تباينهم بالنسبة للمجموع الكلى لهذه المواد .

وفى حالة القياس فى العلوم الطبيعية يمكن إعادة القياس والحصول على

نفس الدرجات وحساب التباين فتظل قيمته كما هي بالنسبة لنفس الخاصية .
 أى أنه يمكن إستعادة Reproducibility نفس التباين : أما بالنسبة
 للموضوعات النفسية والتربوية ، فإن النتائج تختلف باختلاف المرات .
 ويسمى الخطأ أو الفرق بين قيمة التباين فى المرات المختلفة بخطأ التباين
 Error variance . هذا الخطأ يشير إلى إنخفاض ثبات المقياس .

فالقيمة العددية لمعامل ثبات الإختبار تتوقف على تباين الدرجات على
 تباين الدرجات فى هذا الإختبار على أن يرجع هذا التباين إلى الفروق
 الحقيقية بين الأفراد فى الصفة التى يقيسها الإختبار . ولا يكون الإختبار
 ثابتا عندما يكون خطأ التباين فيه كبيراً . ونعود فنقول إن المشكلة
 الأساسية فى تحديد ثبات عمليات القياس هى مشكلة تحديد التباين الحقيقى
 بين الأفراد وتحديد مدى خطأ التباين الذى نقبله .

أسباب تباين درجات الأفراد فى الإختبار :

ههنا الآن أن نعرف لماذا تتباين أو تختلف درجات الأفراد فى
 الإختبار ، وهل يرجع التباين إلى الإختبار نفسه أم إلى المفحوصين
 أم إلى القاتم بالإجراء أو التصحيح . فإذا تعرفنا على مصادر التباين
 أمكننا أن نعالجها فيزيد ثبات الإختبار . وتختلف درجات الأفراد ،
 عن أدائهم للإختبارات ، بحسب عوامل ذات طبيعة مختلفة أتت من
 مصادر متعددة : وليست القائمة التالية حصراً لكل هذه العوامل وإنما
 تمثيلاً لأهمها : فالتباين فى درجات الأفراد قد يرجع إلى أحد أو بعض
 أو كثير من العوامل الآتية :

١- مستوى الفرد فى سمة عامة ثابتة تؤثر فى أدائه لعدد من
 الإختبارات ، مثل القدرة اللفظية :

٢- درجة قدرة الفرد على الإجابة تتأثر بطريقته فى فهم التعليمات .

وطريقة الأداء ومعالجة المواد . وقد تكون للفرد دراية خاصة بنوع من الخبرات والمهارات التي تعينه على الإجابة .

٣- قد يكون الفرد ملما ببعض الحقائق أو المعارف الصعبة وغير ملم بالأسهل منها . وقد يحدث أن يطلب الاختبار منه ذكر ما حدث . مصادفة إن كان يعرفه .

٤- مستوى دوافع الفرد للإجابة عن الاختبار .

٥- الحالة الجسمية والعقلية والوجدانية للفرد وقت الإجراء .

٦- قدرة المفحوص على التزييف والغش والخداع .

٧- الظروف الخارجية أثناء الإجراء كالضوضاء والتهوية والحرارة .

٨- ألفة الفرد بالإجابة عن نوع معين من الاختبارات أو بمعالجة الأجهزة والمواد المستخدمة فيها .

٩- اضطراب الذاكرة أو الانتباه .

١٠- تدخل عامل الصدفة عند ما يختار الفرد بين عدة إجابات فقد يخمن الإجابة الصحيحة فيعطى عنها درجة في مرة ثم يخمن إجابة مخالفة لنفس الفقرة ، في إجراء تالي ، فلا يحصل على نفس الدرجة .

١١- بعض السمات غير ثابتة بطبيعتها ، وخاصة سمات الشخصية .

وكل هذه المصادر ترجع إلى المفحوص نفسه : ولكن التباين قد يرجع إلى أسباب أخرى ، مثل الاختبار نفسه ، سنعرض لها فيما بعد .

موضوعية القياس :

لكي نوفز الدقة في تقديراتنا ، يجب أن نفترض أن كل درجة يحصل عليها المفحوص تعد تمثيلا لقدرته الحقيقية متأثرا ببعض الأخطاء . النتيجة في عمليات القياس فتكون :

الدرجة على الاختبار = الدرجة الحقيقية $\pm$ الدرجة الناتجة عن الخطأ التجريبي في القياس .

$$\text{أى أن } \text{مق} = \text{م} \pm \text{مخ}$$

وبالمثل فإن تباين الدرجات على الاختبار = تباين الدرجات الحقيقية $\pm$ تباين الدرجات الناتجة عن الخطأ التجريبي في القياس .

$$\text{أى أن } \text{ع}^2 = (\text{ع} \pm \text{عخ})^2$$

فإذا كان الخطأ التجريبي في القياس صغيراً جداً ، قاربت الدرجات على الاختبار الدرجات التي يجب أن يحصل عليها الأفراد الممثلة لمستواهم الحقيقي في القدرة المقاسة) . وهنا يرجع اختلاف درجات الأفراد في الاختبار (ع^2) وحسب إلى الفروق الفردية الحقيقة بينهم فيما يتعلق بالجانب المقاس (ع^2) ويصبح التباين في درجات الاختبار مساوياً للتباين في مستويات القدرة .

$$\text{أى أن } \text{ع}^2 = \text{ع}^2$$

ولكننا عادة لا نعرف مقدار المستوى الحقيقي لقدرات الأفراد . ولذا فنحن نطرح تباين الدرجات الناتجة عن الخطأ التجريبي في القياس من تباين الدرجات في الاختبار . وبالمثل يصبح التباين الحقيقي هو باقى طرح التباين بين الدرجات الناتجة عن الخطأ التجريبي في القياس من التباين بين درجات الأفراد في الاختبار .

وبما أننا عند إعطاء الفقرات درجات على الاختبار سلمنا بأن ثمة خطأ تجريبى يتدخل ، وبما أنه يهمننا أن نعرف ما هي نسبة الخطأ التجريبى في الدرجة على الاختبار ، فانه يلزمنا أن نعرف تباين الدرجات الناتجة عن الخطأ التجريبى في القياس حتى نعرف مدى الاختلاف أو التباين الحقيقي

(الذى يعبر عن فروق فردية حقيقية بين مستويات قدرة الأفراد) .
وهكذا نستطيع أن نحدد نسبة التباين الحقيقى المتضمن داخل تباين
الدرجات المعطاة على المقياس ، وهذا يوصلنا إلى تحديد قيمة معامل
الثبات . أى أن معامل الثبات هو ، بهذا المعنى ، جزء من التباين
الحقيقى الداخلى فى تباين الدرجات التجريبية ، وهو يساوى :

$$\frac{\sigma^2_{\text{ع}}}{\sigma^2_{\text{ق}}} = r_{\text{ق ق}}$$

حيث $r_{\text{ق ق}}$ هى معامل الثبات أى معامل الارتباط بين الدرجات على
الاختبار (وهى $r_{\text{ق}}$) فى مرتى الإجراء أو صورتى الاختبار أو نصفى المقياس .
وحيث $\sigma^2_{\text{ع}}$ هى التباين الحقيقى وهو مربع الانحراف المعيارى
لدرجات الأفراد التى تمثل قدرتهم فعلا . (وهى درجة افتراضية قد
نصل إليها عن طريق إجراء مقياس كامل الدقة ، أو يفترض أنها
متوسط درجات الفرد الواحد لو طبق عليه الاختبار عددا لا نهائيا من
المرات ، أو بتطبيق عدد كبير جداً من الاختبارات التى تقيس وظيفة
معينة على نفس الفرد) . أى أنها التباين الحقيقى .

وحيث $\sigma^2_{\text{ق}}$ هى مربع الانحراف المعيارى لدرجات الأفراد على
الاختبار أى أنها تباين درجات المقياس .

وهذا مثال يوضح طريقة الحساب :

إذا فرضنا أننا طبقنا الاختبار مرتين أو أننا طبقنا صورتين من
الاختبار أو أننا طبقنا اختباراً وقسمناه إلى قسمين ، وكان عدد
المتحوسبين ١٠ أفراد . وكانت درجاتهم فى مرتى القياس تحت العمود س
والعمود ص . فيكون معامل الارتباط بين هذه الدرجات ، والذى رمزنا
إليه بالرمز : $r_{\text{س ص}}$.

مربع الفرق بين الدرجات في القياس	الفرق بين الدرجات في القياسين	ح ص	ح ^٢ ص	انحراف القياس الثاني عن المتوسط ح ص	ح ^٢ ص	انحراف الأول عن المتوسط ح ص	القياس الأول القياس الثاني ص ص	القياس الأول س س	الأفراد
٤	٢ +	١٦٥ +	٩	٣ +	٣٠٢٥	٥٥ +	١١	١٣	محمد
٤	٢ -	٢٧	٣٦	٦ +	٢٠٢٥	٤٥ +	١٤	١٢	علي
١	١ -	٧٥ +	٩	٣ +	٦٢٥	٢٥ +	١١	١٠	حسن
٩	٣ +	٢٥ -	١	١ -	٦٢٥	٢٥ +	٧	١٠	زكي
صفر	صفر	٥٥ +	١	١ +	٢٥	٥٥ +	٩	٨	طله
٢٥	٥ -	٤٥ -	٩	٣ +	٢٢٥	١٥ -	١١	٦	فريد
٩	٣ +	٧٥ +	٢٥	٥ -	٢٢٥	١٥ -	٣	٦	موسى
٤	٢ -	٢٥ +	١	١ -	٦٢٥	٢٥ -	٧	٥	رضى
٩	٣ -	٩	٤	٢ -	٢٠٢٥	٤٥ -	٦	٣	كريم
١	١ +	٣٨٥ +	٤٩	٧ -	٣٠٢٥	٥٥ -	١	٢	رفعت
٦٦		١٠٢	١٤٤	صفر	١٢٤٥	صفر	٨٠	٧٥	المجموع
٦٦		مجموع ح ص	مجموع ح ^٢ ص		مجموع ح ^٢ ص		٨	٧٥	المتوسط

الانحراف المعياري لدرجات الأفراد في القياس الأول :

$$\sqrt{353} = 1245 \sqrt{V} = \frac{1245}{10} \sqrt{V} = \frac{1245^2}{10} \sqrt{V} = \frac{1550025}{10} \sqrt{V} = 155002.5 \sqrt{V}$$

الانحراف المعياري لدرجات الأفراد في القياس الثاني :

$$\sqrt{379} = 144 \sqrt{V} = \frac{144}{10} \sqrt{V} = \frac{144^2}{10} \sqrt{V} = \frac{20736}{10} \sqrt{V} = 2073.6 \sqrt{V}$$

معامل الثبات وهو معامل الارتباط بين درجات الأفراد في القياسين :

$$\frac{155002.5 \sqrt{V} \times 2073.6 \sqrt{V}}{155002.5 \times 2073.6} = \frac{155002.5 \times 2073.6}{155002.5 \times 2073.6} = 1$$

حيث n هي عدد الأفراد (١٠)

$$0.76 + = \frac{102}{133787} = \frac{102}{379 \times 353 \times 10} =$$

ويمكن حساب معامل الارتباط هذا بدون حساب الانحرافات المعيارية وذلك باستخدام المعادلة التالية :

$$\frac{102}{(144)(1245)} = \frac{102}{(144^2)(1245^2)} \sqrt{V} = \frac{102}{(144^2)(1245^2)} \sqrt{V}$$

$$0.76 + = \frac{102}{13379} = \frac{102}{17928} \sqrt{V} =$$

بقى أن نحسب التباين الحقيقي لدرجات الأفراد في القياسين من المعادلة التالية :

$$\frac{\infty^2 \text{ع}}{\text{ع}^2 \text{ق}} =$$

$$\text{حيث } \text{ع}^2 \text{ق} = ٠.٧٦ \quad \text{وحيث } \text{ع}^2 \text{ق} = ٦٦$$

$$\therefore \frac{\infty^2 \text{ع}}{٦٦} = ٠.٧٦ +$$

$$\therefore \infty^2 \text{ع} = ٦٦ \times ٠.٧٦ + = ٥٠.٢$$

وهكذا اتضح أن معامل الثبات هو ٠.٧٦ والتباين الحقيقي هو ٥٠.٢ والتباين التجريبي هو ٦٦ وسنوضح فيما يلي مزايا وعيوب كل طريقة من طرق حساب معامل الثبات والمعادلات التي تعالج هذه العيوب .

طرق حساب معامل للثبات

قلنا إن للثبات نوعان ؛ ثبات مطلق وثبات نسبي . وعند حساب معامل الثبات نحصل على درجات الأفراد في الاختبار بإجرائه أكثر من مرة ثم نحللها لنعرف مدى الاتساق بينها في المرات المختلفة . ويحسن بالطبع أن تتكرر مرات الإجراء ، ولكننا نعلم أن سلوك الأفراد في موقف الاختبار يتأثر بنجرتهم من هذا الموقف ، فهو يتأثر مثلاً بالتدريب والخبرة والتذكر والنمو . ولهذا فبدلاً من تكرار الإجراء نكتفي بالإجراء مرتين فقط على أن يكون عدد المفحوصين كبيراً وممثلاً .

وقد عرفنا معامل الثبات بأنه معامل الارتباط بين مجموعتين من القياسات المتكافئة الخاصة بمجموعة من الأفراد . وعلينا الآن أن نحدد هذه القياسات المتكافئة . وتنحصر الخطوات العملية في تحديد القياسات المتكافئة ، وبالتالي في تحديد معامل الثبات ، في الطرق الآتية :

- ١ - إجراء صورتين متكافئتين من الاختبار وحساب معامل الارتباط بين درجات نفس الأفراد في كل منهما *Equivalent forms* .
- ٢ - إجراء نفس الصورة من الاختبار على نفس الأفراد مرتين وحساب معامل الارتباط *Test-Retest* .

- ٣ - تقسيم الاختبار الواحد إلى مجموعتين متكافئتين من الفقرات ، وإعطاء درجة لكل فرد على كل من القسمين ، وحساب معامل الارتباط *split-half* .

- ٤ - تحليل التباين بين فقرات الاختبار ، وتحديد خطأ التباين *Analysis of variance* وهي أحياناً تسمى طريقة التكافؤ المنطقي *Rational-Equivalence*

٥ - معامل الثبات الحقيقي *index of reliability*

- ١ - معامل الثبات بطريقة الصور المتكافئة :

ما دمنا قد عرفنا معامل الثبات بأنه معامل الارتباط بين مجموعتين من القياسات المتكافئة ، فإن استخدام طريقة الصور المتكافئة يطابق هذا التعريف . ولكن علينا أن نحدد عمليات إعداد الصور المتكافئة حقاً من الاختبار . ويسمى جيليكسون هذه الطريقة بطريقة الصور المتوازية *Parallel Forms* ويسمى بيلفورد طريقة الصور المتبادلة *Alternate Forms* ويسمى البعض الآخر بطريقة الصور المتكافئة *Equivalent Forms* . ويفترض في هذه الطريقة أنه يمكن إنشاء صورتين متكافئتين من الاختبار

الواحد ، متكافئتان من حيث تمثيل جانب السلوك المطلوب قياسه . أى أن التكافؤ يجب أن يشمل الجوانب الآتية :

(أ) عدد مكونات الوظيفة التى يقيسها الاختبار .

(ب) نسبة الفقرات التى تخص كلامها .

(ج) مستوى صعوبة الفقرات .

(د) طريقة صياغة الفقرات .

(هـ) طول الاختبار وطريقة إجرائه وتصحيحه وتوقيته :

(ز) تساوى متوسط وتباين (ع^٢) درجات الأفراد على كل من

الصور .

ويمكن الوصول إلى الصور المتكافئة من الاختبار الواحد بتقسيم عدد الفقرات ، التى اتضح أن مستوى صعوبتها يلائم الغرض ويناسب العينة ، بحسب تدرجها فى الصعوبة وبحسب تمثيلها للوظيفة المناسبة . وهكذا نصل إلى مجموعتين أو أكثر من الفقرات ذات الأوزان المتكافئة بحيث أن الفرد الواحد ، لو وحدت شروط الإجراء وكل العوامل المؤثرة على أدائه ، يحصل على نفس الدرجة تقريباً أيا كانت الصورة المطبقة عليه / وقد يحتاج الباحث إلى عمل صورة مكافئة للاختبار الذى ثبتت صلاحيته . وهنا يقوم بصياغة فقرات كل فقرة منها تناظر فقرة من الاختبار الأصلي ، تناظرها من حيث مستوى الصعوبة ونوع الوظيفة المقاسة وطريقة الصياغة وطريقة الإجابة .

وبالرغم من أن فقرات الصور المتكافئة ليست متطابقة تماماً إلا أنه كلما اقتربت الصورة زادت فرصة انتقال أثر التدريب . فالمفحوص يمر بموقف الاختبار الأول على أنه موقف تعليمي وعلى هذا فهو يطبق ما اكتسبه من الإجراء الأول عندما يواجه موقفاً مشابهاً فى الإجراء الثانى فارتفع درجته

عند إعادة الإجراء بنفس الصورة أو بالصورة المكافئة . ولكن إذا أفاد كل المفحوصين من خبرتهم بالموقف الأول في تنمية قدراتهم بنفس القدر فإنهم جميعاً سيحصلون على درجات تزيد بنفس المعدل عن درجاتهم الأولى . فإذا تأكد الباحث من هذا كان معامل الثبات الذى يحسبه صحيحاً ، إذ أن المركز النسبي للفرد في جماعته لن يتأثر ما دام الكل قد نمي بنفس القدر . وتعرض هذه الطريقة لأخطاء تنشأ عن قصر أو طول المدة بين إجراء الصورتين المتكافئتين .

٢ - الثبات بطريقة إعادة الاختبار :

في طريقة إعادة الاختبار Test—Retest Method يكرر تطبيق نفس الاختبار بنفس الصورة على مجموعة من المفحوصين . ومعامل الثبات هذا معرض لكثير من الأخطاء التي تقلل من أهميته . فإذا تأكدنا من أن إجراء الاختبار في المرة الأولى لا يغير في المفحوصين إطلاقاً ، بما يجعلهم يستجيبون في الإجراء الثاني استجابات تمثل مستويات قدرتهم وحسب ، كنا محقين في استخراج معامل الثبات بحساب معامل الارتباط بين درجات الأفراد في المرتين . ولكن هذا لا يحدث ، إذ أن إجراء الاختبار على المفحوصين في المرة الثانية يتأثر بالعوامل الآتية :

(١) موقف الاختبار موقف تعميمي يفيد منه المفحوصون على الأقل في الألفة بمواقف الاختبار التالية . فلا يكونون عند إعادة الإجراء متأثرين بغرابة الموقف والتوتر الانفعالي عند بدء أى امتحان بنفس الدرجة التي أثرت في أداؤهم لنفس الصورة من الاختبار في المرة الأولى : وهذا النقص لا يؤثر وحسب على درجات الأفراد بل إنه يوجه إلى منهج البحث المتبع ، فلا نكون عند إعادة الإجراء قد وفرنا ووجدنا نفس ظروف الإجراء الأول ، وبالتالي فإن درجات الأفراد في مرة تتأثر بعوامل معينة (كالاضطراب الانفعالي

وصدمة الجدة وخطوات تكوين العلاقة الطيبة في الإجراء الأول) ، بينما تمثل درجاتهم في الإجراء الثاني محصلة عدة قوى منها الألفة بموقف الاختبار وتعلم الطريقة الاقتصادية في معالجته .

(ب) وبالمثل ، ولنفس السبب ، لا تمثل درجات المفحوصين في إعادة الإجراء مستويات قدرتهم وحسب بل وأيضاً تعد محصلة لعدة قوى منها انتقال أثر التدريب والمران Practice effect . كما أن هذه القوى تشمل أيضاً تعلم المفحوص طرقاً فنية في الإجابة عن الفقرات ومعالجة مواد وأجهزة الاختبار .

(ج) هذا إلى أن الفترة بين الإجراء الأول والإجراء الثاني interval تؤثر هي الأخرى على أداء نفس الصورة من الاختبار عند إعادة الإجراء . ويمكن توضيح هذا التأثير في النقاط التالية :

١ - إذا طالت هذه الفترة زاد احتمال تأثير النمو Development العقلي والجسمي والوجداني والاجتماعي الحادث فيها لدى الأفراد . وتزيد خطورة تأثير النمو كلما صغر سن المفحوصين . فنحن نعلم أن معدل سرعة النمو في الأعمار الصغيرة يكون كبيراً بما يجعل الاستجابة لنفس الصورة عند إعادة إجرائها بعد شهر من المرة الأولى انعكاساً لمقدار هذا النمو . ويقل معدل سرعة النمو كلما كبر سن المفحوصين ، وعلى هذا يزداد معامل الثبات بطريقة الإعادة في حالة عينة من الكبار سناً عنه في حالة الأصغر سناً .

٢ - ويزداد أثر عامل النمو كلما كان المفحوصون على مستوى عال من القدرة العقلية . فنحن نعلم أن معدل سرعة النمو يكون كبيراً عند الأذكى عنه عند الأقل ذكاء . وعلى هذا يزداد معامل الثبات كلما انخفض ذكاء المفحوصين إذ أن النمو الحادث أقل مقداراً منه في حالة الأذكى .

٣- إذا قصرت الفترة بين الإجراء وإعادة الإجراء كانت استجابات المفحوصين في إعادة محصلة لعدة قوى ، فضلاً عن تمثيلها لقدرتهم ، أهمها تذكر الأسئلة وحفظ الإجابات خاصة إذا كان الاختبار قصيراً أو كانت فقراته لا تتطلب أعمال وظائف معقدة .

وهكذا نجد أن طول الفترة بين الإجراء وإعادة الإجراء يجب أن يتحدد بحسب طبيعة العينة من حيث الذكاء والسن ، وبحسب طبيعة الاختبار من حيث الطول وتعدد الوظائف المقاسة . وعموماً نجد أنه ليست هناك قاعدة تحكم طول هذه الفترة ، وإن كانت غالباً لا تقل عن أسبوع ، في الاختبارات الطويلة أو التي تقيس وظيفة معقدة ولا تزيد عن شهر في الاختبارات القصيرة أو التي تقيس وظيفة بسيطة . ولكن هذه المدة قد تكون أياً ما معدودة في اختبارات الشخصية الطويلة مثل اختبار

M.M.P.I.

وعلى هذا يجب أن يعمل حساب أثر هذه المدة بين الإجراء وإعادة ، فإذا كان الأثر موحداً بين كل أفراد العينة ، كنا محقين في استخدام هذه الطريقة إذ أن مركز الفرد بالنسبة لجماعته في الوظيفة المقاسة لم يتغير كثيراً .

٣- الثبات بطريقة نصفى الاختبار

تستدعى طريقة تجزئة الاختبار إلى نصفين Split - half method تقسيم الاختبار الواحد إلى قسمين يحصل كل فرد على درجة عن كل قسم منهما . وهكذا يصبح كل نصف وكأنه صورة مكافئة . إلا أننا في الطريقة النصفية نطبق الاختبار كله (النصفين معا) ثم نصححه فيحصل الأفراد على مجموعتين من الدرجات إحداها عن القسم الأول

والأخرى عن الثانى . بينما فى طريقة الصور المتكافئة تمر فترة بين إجراء الصورتين . فالفترة الزمنية بين إجراء القسمين ، فى الطريقة النصفية : تصبح صفرا . ولكى نضمن أن التقسيم لا يخضع للتبادل المنتظم Systematic alternation بين كل فقرة من القسم الأول مع نظيرتها من القسم الآخر ، نقسم الاختبار بحيث يصبح أحد القسمين حاويا لل فقرات الفردية (الفقرة رقم ١ ، ٣ ، ٥ ، .. الخ) والقسم الآخر حاويا لل فقرات الزوجية (٢ ، ٤ ، ٦ ، .. الخ) . وهكذا يوزع المفحوص جهده ووقته ويستغل وظائفه بنفس القدر فى الجزئين . بينما لا يتوفر هذا فى تقسيم الاختبار إلى جزئين الأول (الفقرات من ١ إلى ٥٠ مثلا) والآخر (الفقرات من ٥١ إلى ١٠٠ مثلا) إذ تؤثر عوامل تراكم التعب والملل فى الإجابة عن الجزء الأخير . فضلا عن أن القسم الأول عادة ما يكون أسهل من القسم الأخير .

والميزة المنهجية Methodological فى طريقة التجزئة النصفية إلى قسمين : فردى وزوجى ، هى إجراء النصفين فى تآنى Simultaneously فتوحد ظروف الإجراء توحيدا تاما . ولكن هذه الطريقة يعيبها أننا نحسب معامل الثبات من درجات الأفراد على كل نصف من الاختبار باعتباره اختبارا أو صورة كاملة مكافئة ، ولا نحسبه من درجات الأفراد عن الاختبار ككل . ونحن نعرف أن الطول الأصلى للاختبار يعنى درجة تمثيل عينة فقراته لجوانب الوظيفة المقاسة ، وعلى ذلك لا يشمل القسم كل منطقة السلوك المطاوب قياسه والاختبار ككل يمكن الاعتماد عليه ، لهذا السبب ، أكثر من نصفه . وهناك معادلات إحصائية تعوض هذا النقص .

(أ) معادلة سبيرمان براون :

هذه الصورة العامة لمعادلة سبيرمان براون Spearman-Brown Formula لإجراء تصحيح إحصائي لمعامل الثبات المحسوب بالطريقة النصفية :

$$\frac{\frac{1}{3} \cdot \frac{1}{3}^{\sqrt{2}}}{\frac{1}{3} \cdot \frac{1}{3}^{\sqrt{2}} + 1} = 1.01^{\sqrt{2}}$$

حيث $1.01^{\sqrt{2}}$ هي معامل ثبات الإختبار كله .

، $\frac{1}{3} \cdot \frac{1}{3}^{\sqrt{2}}$ هي معامل الارتباط بين درجات الأفراد على نصف الإختبار أى أنها معامل ثبات كل نصف من الإختبار . فلو كان معامل الارتباط هذا هو 0.68 يكون معامل الثبات للإختبار كله هو :

$$0.81 = \frac{1.36}{1.68} = \frac{0.68 \times 2}{0.68 + 1} = 1.01^{\sqrt{2}}$$

ويعاب على هذه الطريقة أنها تتحدد بتساوى الانحراف المعياري لكل من النصفين . فإذا لم يتساوى التباين بين الجزئين لجئنا إلى طريقة أخرى مثل .

(ب) معادلة جتمان :

تفيد هذه المعادلة في حساب معامل الثبات للإختبار ككل ، من معامل ارتباط نصفية لو كانا غير متساويين في تباين درجات الأفراد عليهما .
والمعادلة هي :

$$s_{1.1}^2 = \left(\frac{1^2 E + 2^2 E}{2E} - 1 \right)$$

، $s_{1.1}$ هي معامل ثبات الاختبار ككل

، E^2 هي مربع الانحراف المعياري لدرجات الأفراد عن النصف ١
أي تباينها .

، E^2 هو تباين درجات النصف ٢ من نفس الاختبار -
فإذا كانت E^2 (في النصف ١) هي ٢ ، E^2 (في النصف ٢) ب
هي ١٥ ، E^2 للاختبار ككل هي ٦ كان معامل الثبات مساوياً :

$$s_{1.1}^2 = \left(\frac{1^2 + 2^2}{6} - 1 \right)$$

$$\left(\frac{5}{6} - 1 \right) =$$

$$= (0.83 - 1)$$

$$= 0.17 \times 2 =$$

$$= 0.34$$

(ج) معادلة رولون :

أما معادلة رولون Rulon Formula فهي تعتمد على ما افترضه صاحبها
من أن تباين درجات الأفراد على الاختبار ككل يرجع ، كما سبق القول ،

إلى تبين حقيقى فى مستويات القدرة المقاسة كما يرجع إلى أخطاء تجريبية فى القياس . والمعادلة هى :

$$\frac{E^2}{E^0} - 1 = 1.01^M$$

حيث 1.01^M هى معامل ثبات الاختبار ككل ، E^2 هى الخطأ التجريبى فى القياس وهو متوسط مربعات الفروق بين درجات الأفراد على نصفي الاختبار .

، E^2 تبين درجات المجموعة فى الاختبار ككل فيحصل كل فرد على درجتين إحداهما عن النصف الأول والأخرى عن النصف الثانى ، ونطرح الدرجتين من بعضهما بالنسبة لكل فرد ثم نربع الفروق ونجمعها ثم نقسمها على عدد الأفراد ثم نحسب متوسط درجات المجموعة على الاختبار ككل ثم نحسب انحراف الأفراد عن هذا المتوسط ثم نربع هذه الانحرافات ونطبق المعادلة .

(د) معادلة هورست :

وضع هورست Horest معادلته للتغلب على صعوبات عملية تواجه الباحث عادة . وأهمها أنه قد يستحيل تقسيم الاختبار إلى نصفين متكافئين تماماً . فإذا فرضنا أن معامل الثبات للاختبار كله هو (س) . وأن معامل الارتباط بين قسمي الاختبار هو (س_{١٢}) . وأن (س) هى نسبة عدد فقرات القسم الفردى إلى عدد فقرات الاختبار كله ، (س_{١٢}) هى نسبة عدد فقرات القسم الزوجى إلى عدد فقرات الاختبار كله . يكون معامل الثبات هو :

$$r = \frac{[\frac{7r^2}{n \cdot n} + 4 \cdot \frac{r}{n} (1 - r^2)]}{2 \cdot \frac{r}{n} (1 - r^2)}$$

٤ - معامل الثبات بطريقة تحليل التباين :

إن تقسيم الاختبار إلى نصفين إجراء مفتعل ، إذ أن الاختبار على قدر ما يحوى من الفقرات تكون هناك طرقاً لتقسيمه . كما أن حساب معامل الثبات من أحد قسمي الاختبار أو من القسمين معاً . فيه تجاهل لكثير من البيانات عن كل الفقرات التي يشملها الاختبار : فالثبات هنا هو اتساق الأداء من فقرة إلى أخرى في الاختبار ، وهذا المفهوم قريب من مفهوم الاتساق الداخلى Internal Consistency . ولكن تحليل اتساق فقرات الاختبار يتأثر بالحدود الزمنية للإجراء . ولهذا فسرى أن هناك معادلات لحساب ثبات اختبارات السرعة . وأهم معادلات حساب معامل الثبات بتحليل تباين (ع^٢) درجات الأفراد على فقرات الاختبار هي :

(١) معادلة كودر - ريتشاردسون :

تتماز معادلة كودر - ريتشاردسون Kuder-Richardson Formula :
بالبساطة العملية والسهولة والسرعة وصيغتها هي :

$$r_{KR} = \frac{n \cdot \sigma^2 (1 - r^2)}{n \cdot \sigma^2 - \sum (r_i^2)}$$

حيث r_{KR} هي معامل ثبات الاختبار ككل :

وحيث m هي عدد فقرات الاختبار : وليكن ٤٠ فقرة مثلاً .
 c ، هي تباين درجات الأفراد على الاختبار . وليكن ٣٦ مثلاً .
 m ، هي متوسط درجات الأفراد على الاختبار . وليكن ٢٥٧ مثلاً .
 وبالتعويض في المعادلة نجد أن :

$$\frac{(257 - 40) \times 36 - 36 \times 40}{36 \times 36} = 1.14$$

$$= 0.76$$

، (ب) معادلة هويت :

تفترض طريقة هويت Hoyt's Analysis of Variance Procedure

أن درجة الفرد على الاختبار قد تقسم إلى أربعة عناصر تكونها وهي :

١ - عوامل تتعلق بكل الفقرات عند كل الأفراد .

٢ - عوامل تتعلق بالفقرة .

٣ - عوامل تتعلق بالفرد .

٤ - عامل الخطأ .

$$\text{معاملة الثبات} = 1 - \frac{\text{خطأ التباين}}{\text{التباين بين الأفراد}}$$

، (ج) معادلة چلكسون :

تمتاز معادلة چلكسون Gulliksen عن معادلة سپيرمان براون بأنها
 تتصلح لحساب معامل ثبات إختبارات السرعة ، التي يحدد فيها الزمن بحيث

لا يستطيع أى مفحوص أن يجيب عن كل الفقرات ، إذ أن عامل الزمن له أثره فى معامل الثبات كما سنرى . والمعادلة هى :

$$\overline{r_{\frac{E}{X}}} - 10.13 = 10.13$$

حيث 10.13 هى معامل الثبات بعد تصحيح أثر السرعة .

، 10.13 هى معامل الثبات قبل تصحيح أثر السرعة محسوب بطريقة

سبيرمان براون وليكن 9 و 10 ، $r_{\frac{E}{X}}$ هى متوسط عدد الفقرات التى تركها الأفراد أى لم يسمح الوقت بمحاولة الإجابة عنها وليكن 1 فقرة .

، $r_{\frac{E}{X}}$ هى تباين عدد الفقرات التى يجيب عنها الأفراد إجابة خاطئة : وليكن 10 .

$$\therefore 10.13 = 0.9 - \frac{1}{3} = 0.8$$

٥ - معامل الثبات الحقيقى :

معامل الثبات هو معامل الارتباط بين الاختبار ونفسه أو بين الدرجات الحقيقية على الاختبار ونفسه . وطريقة الثبات الحقيقى *index of Reliability* يعتمد على استخراج الجذر التربيعى لمعامل الثبات المستخرج بأى من الطرق الثلاث أى أن :

معامل الثبات الحقيقى $\sqrt{V}$ معامل الثبات المحسوب بأى معادلة فإذا

كان معامل الثبات الذى حسبناه بأى معادلة هو 0.64 كان معامل الثبات

الحقيقى لهذا الاختبار هو $\sqrt{0.64} = 0.8$ وهذا الرقم هو الحد الأعلى الذى لا تزيد عنه معاملات الارتباط بين الاختبار وأى محك آخر .

معادلة كرونباخ :

إشتق كرونباخ Cronbach صورة عامة لمعادلة معامل الثبات (على أساس معادلة كودر - ريتشر دسن) وسماه المعامل ألفا (α) وهى :

$$\text{معامل ألفا } (\alpha) = \frac{n}{n-1} \left(1 - \frac{\sum E^2_k}{E^2_c} \right)$$

حيث n هى عدد أقسام الاختبار وهو غالباً ٢

، k هو أحد أقسام الاختبار

، E^2_c هو مجموع تباينات الأفراد فى هذا القسم

، E^2_k هو تباين الاختبار الكلى .

تفسير معامل الثبات :

كل معامل ثبات حسبه بإحدى الطرق أو المعادلات له معنى معين ولا يستخدم إلا فى حدود هذا المعنى . وكل معامل ثبات أوجدناه يقوم على شروط معينة يجب توفرها وإلا فلا يصح أن يستخدم . ومن هنا وجب عند ذكر معامل ثبات أى اختبار أن تذكر الطريقة والمعادلات المستخدمة والشروط التى توفرت قبل إيجاده . وبهذا فقط يمكننا أن نقارن بين معاملات الثبات بعد أن نعرف معنى كل معامل منها . ويقول كرونباخ Cronbach أن :

١ - معامل الارتباط بين مرتين الإجراء test-retest Method هو معامل

استقرار Stability .

٢ - معامل الارتباط بين صورتين أو صور الاختبار المتكافئة equivalent

forms method هو معامل تكافؤ equivalence وهو يصبح معامل تكافؤ ومعامل استقرار معاً إذا مرت فترة زمنية مناسبة بين إجراء الصورتين .
 ٣- أن معامل الارتباط بين درجات الأفراد على نصفي الاختبار split-half method هو معامل استقرار .

بمعنى أنه عند حساب معامل الثبات بطريقة الإعادة يهمننا أن نعرف مدى استقرار درجات الأفراد عن نفس الأسئلة بين مرات الإجراء . وعند حساب معامل الثبات بطريقة الصور المتكافئة يهمننا أن نعرف مدى ارتباط درجات الأفراد على صورتين أو صورة الاختبار أى مدى تكافؤ هذه الدرجات المعطاه عن صور متكافئة . بينما عند حساب معامل الثبات بالطريقة النصفية يهمننا أن نعرف التكافؤ ، بالمعنى السابق ، بافتراض أن منطق الصور المتكافئة قائم بين الأنصاف المتكافئة .

وعلى هذا ، ما دمنا نجمع بين الاستقرار والتكافؤ في طريقة تطبيق الصور المتكافئة بعد مرور فترة من الزمن بين الصورة والأخرى ، قد يتوقع البعض أن معامل الثبات بطريقة النصفين يكون أعلى منه بأى من الطريقتين الأخريتين . بينما نجد أن معامل الثبات النصفى دائماً أقل من معامل الاستقرار . وأقل من معامل التكافؤ أيضاً .

فعامل التكافؤ يدل على مدى دقة قياس الاختبار لأداء الفرد في لحظة معينة . فهو يقدر المدى الذى تتراوح فيه درجة الفرد إذا استخدمت عينات مختلفة من الأسئلة التى تقيس نفس القدرة . ومعامل التكافؤ يحسب عادة بالطريقة النصفية مصححة بمعادلة سبيرمان - براون .

وتعطى الطريقة النصفية نتائج مضللة إذا لم يكن النصفان متكافئين تماماً كالصور المتكافئة . فيتساوى المتوسط والانحراف المعياري لكل من النصفين تساويهما بين كل من الصورتين . ويجب أن تماثل الأنصاف

في محتواها فلا يقسم اختبار الذكاء الذي يشمل في نصفه أسئلة حساب وفي نصفه الآخر أسئلة أداء . وقد جرى تقليد أن يقسم الاختبار إلى نصفين أحدهما يشمل الفقرات الفردية والآخر يشمل الفقرات الزوجية . ولكن هذه الطريقة قد لا تؤدي إلى تكافؤ نصفي الاختبار . ويقترح كرونباخ أن تحلل أوراق الإجابة لتحديد مستوى صعوبة الفقرات عن طريق النسبة المئوية للنجاحين في كل فقرة ثم تقسم الفقرات إلى مجموعتين متكافئتين في الصعوبة وفي المحتوى . ثم يحسب معامل الارتباط بين النصفين مهما كان ترتيب الأسئلة التي اعتبرت ضمن أحد النصفين .

وتقسم الاختبار إلى نصفين متكافئين في الصعوبة والمحتوى بالطريقة التي اقترحها كرونباخ . يجعل معامل الثبات بالطريقة النصفية أكثر دقة . ولكن عدم الاهتمام بترتيب الفقرات التي يشملها كل نصف أمر يذهب بهذه الدقة . فقد يتضمن أحد النصفين ، بهذه الطريقة ، فقرات يجرى أغلبها في النصف الأول من زمن الاختبار في حين قد يتضمن النصف الآخر فقرات يجرى أغلبها في النصف الثاني من زمن الاختبار . وهنا لا تكون كفاءة المفحوص متسقة إذ يتأثر في النصف الآخر بتغير الظروف وبموامل التعب والملل وانتقال أثر التدريب . ويمكن التغلب على هذا العيب بأن توزع الأسئلة ذات المستوى الموحد من الصعوبة والتي تتكافأ في المحتوى توزيعاً متماثلاً في الاختبار بأن يجعل هذه الفقرات نصفين أحدهما فردي والآخر زوجي . ثم يحسب معامل الارتباط بين هذين النصفين فيكون معامل استقرار ومعامل تكافؤ مما وفي نفس الوقت .

ومعامل التكافؤ يمكن أن يقاس أيضاً بطريقة السور المتكافئة التي تشملنا مباشرة لها ، ما إذا كانت العينات المختلفة من الفقرات تعطي نفساً تاماً في طريقة هي الطريقة الفردية بين المفحوصين .

ومعامل الاستقرار يبين المدى الذى يتراوح فيه استقرار الدرجات ، عن فقرات اختبار معين ، فى مدى زمن معين هو الفترة بين الإجراء وإعادة الإجراء . فهو يدل على ما إذا كانت عينة السلوك المأخوذة فى وقت معين تطابق عينة السلوك المأخوذة فى أوقات أخرى : وهو يقدر بطريقة إعادة يأخذ نفس المفحوصين نفس الاختبار فى فرصتين مختلفتين ، بحيث لا تكون أمامهم فرصة للتدريب أو حفظ مواد الاختبار خلال الفترة الزمنية بين المراتين . ويعيب هذه الطريقة فضلاً عن العيوب المذكورة سابقاً ، أنها لا تفيد فى اختبارات الاستدلال إذ قد يتذكر المفحوص حل المسألة فلا تصبح المسألة إذن مقياساً للاستدلال بل للتذكر .

وتقرب الطريقة النصفية كثيراً من معنى الاتساق فنحن نحسب اتساق أداء الفرد مع نفسه فى موقف الاختبار وينطبق هذا ، على وجه الأخص ، عند تعديل معامل الارتباط بين نصفى الاختبار باستخدام معادلة كودر - ريتشر دسن التى تقيس الاتساق الداخلى بين الفقرات **Inter-Item Consistency** . وهكذا تقرب كثيراً من مفهوم التجانس **Homogeneity** . فيمكننا حساب معامل التجانس بطرح قيمة معامل كودر - ريتشر دسن من معامل الثبات النصفى بطريقة سبيرمان - براون ؛ ذلك أن معامل الثبات بطريقة كودر - ريتشر دسن يعتمد على تحليل التباين إذ يدخل فى اعتباره عدد الفقرات (n) والمتوسط الحسابى (M) والتباين (σ^2) . وعموماً نجد أن معامل الثبات بطريقة كودر - ريتشر دسن يعطى أصغر قيمة لقياس الثبات بينما تعطى معادلة سبيرمان - براون أعلى قيمة للثبات ويكون الفرق بينهما هو معامل التجانس .

وكما قلنا ليس هناك معامل ثبات كامل فالقليل من الاختبارات يقترب معامل ثباتها من واحد صحيح . وأحياناً تكون الاختبارات القصيرة ذات الثبات المنخفض ذات قيمة في أغراض خاصة مثل إصدار الأحكام السريعة وعموماً ترتفع معاملات الثبات في اختبارات القدرة والاستعداد عنها في اختبارات الشخصية . وتختلف قيمة الثبات المطلوب باختلاف الغرض من استخدام الاختبار . ولكن أياً كان الغرض يحسن أن نصل إلى أعلى معامل ثبات ممكن . والجدول التالى (ص ٢٤٧) عن كرونباخ يقارن بين خصائص معاملات الثبات .

العوامل التى تؤثر على معامل الثبات .

رأينا أن العوامل التى تؤثر على ثبات الاختبار هى :

١ - طول الاختبار : فمن المعلوم أن الاختبار عينة للجانب من السلوك المطلوب قياسه . ولهذا فكلما كبر حجم هذه العينة ، أى كلما طال الاختبار زاد تمثيلها للمجتمع الأسمى أى لمنطقة السلوك المراد قياسها . وذلك لأن درجة الفرد فى الاختبار الطويل تمثل قدرته . وقدرته ثابتة ولهذا فدرجته هنا أكثر ثباتاً . وعلى هذا يمكن رفع معامل الثبات بإطالة الاختبار Lengthening the test وذلك بالمعادلة الآتية :

$$\frac{r_{11}}{r_{11}(1 - r) + 1} = r$$

حيث r_{11} هو المعامل الذى نريد أن نصل إليه .

، r هو المعامل الحالى .

، r هى نسبة طول الاختبار الجديد إلى الاختبار الأسمى .

الاسم وطريقة الحساب	الفروض	الأخطاء المحتملة عند عدم توفر الفروض	تساؤلات يجب عنها المعامل
معامل التكافؤ الطريقة النصفية	تكافؤ الأنصاف	ترتفع معاملات ثبات اختبارات السرعة عند استخدام هذه الطريقة بدرجة كاذبة . وينخفض المعامل بدرجة شديدة إذا لم تتكافأ الأنصاف .	ما مدى دقة الاختبار في القياس ؟ ما مدى كفاءة تمثيل الاختبار لكل الفقرات التي كان يمكن أن يتضمنها ؟
كودر - ريتشر دسون	يجب أن يقيس الاختبار عاملاً واحداً	يرتفع معامل ثبات اختبارات السرعة بدرجة كاذبة عند استخدام هذا المعادلة بينما ينخفض هذا المعامل بدرجة كبيرة عندما تقيس الفقرات عوامل كثيرة .	
الصور المتكافئة ، بحيث لا تمر فترة طويلة بين إجراءاتها .	يجب أن تتكافأ	يبين المعامل درجة التكافؤ ، أكثر مما يبين مدى دقة قياس كل صورة للوظيفة المعينة .	
معامل الاستقرار	يجب أن لا تكون هناك فرصة لنفس القدرة بفعل التدريب خلال الفترة بين الإجراءين .	تدريب الوظيفة المقاسة يخفف هذا المعامل .	ما مدى استقرار القياس بهذا الاختبار ؟
إعادة الإجراء بعد مرور فترة			
معامل الاستقرار والتكافؤ معاً	يجب أن تتكافأ الصور ، وأن لا تكون هناك فرصة للتدريب	إذا لم يتوفر الفرضان انخفض تقدير هذا المعامل لدقة الاختبار في القياس .	ما مدى تعلق عينة السلوك في وقت بنتائج نفس العينة في وقت آخر ؟
إجراء الصور المكافئة بعد فترة من الزمن			

جدول (٣) يقارن بين خصائص وشروط ونقائص طرق حساب معاملات الثبات عن :

Cronbach, L. J., Essentials of psycholocial testing, N. Y. : Haper & Brothers Publishers, 1949.

فإذا أردنا رفع معامل ثبات الاختبار من ٠.٥ إلى ٠.٨

$$\therefore \frac{0.5 \times n}{0.5(1-n) + 1} = 0.8$$

$$0.5n = 0.8 + 0.4n - 0.5n$$

$$\therefore 0.5n = 0.8 - 0.4n$$

$$\therefore 0.9n = 0.8$$

$$\therefore n = 0.89$$

وهكذا يضاعف طول الاختبار إلى أربعة أمثال طوله الأصلي أى أن عدد الفقرات التى سنصممها أو نضيفها يساوى ٣ أمثال عدد فقرات الاختبار الأصلي : على أنه يجب ألا يطول الاختبار بحيث يصبح مملاً :

(ب) زمن الاختبار :

معامل الثبات يظل يرتفع بزيادة الوقت الذى يستغرقه إجراء الاختبار ، إلى درجة معينة ، ثم ينخفض الثبات بطول المدة . وهذه الدرجة تختلف من اختبار إلى آخر .

(ج) تجانس المجموعة :

يقل ثبات الاختبار بزيادة تجانس homogeneity المجموعة المطبق عليها ، ويرتفع الثبات بزيادة تباين درجات أفرادها عليه . وعلى هذا يمكن التنبؤ بمدى تغير معامل الثبات بتغير الانحراف المعياري لدرجات : عينة جديدة من الأفراد باستخدام المعادلة التالية :

$$ص.ص. = ١ - \frac{ع.ص.}{ع.ص.} (١ - ص.ص.)$$

حيث ص.ص. هي معامل الثبات عند إجراء الاختبار على المجموعة ص.

، ع.ص. هو تباين المجموعة ص (وليكن ٦٢٥)

، ع.ص. هو تباين المجموعة ص (وليكن ٤٠٠)

، ص.ص. هو ثبات المجموعة ص (وليكن ٠.٩)

$$\frac{٦٢٥}{٤٠٠} - ١ = (٠.٩ - ١) \frac{٦٢٥}{٤٠٠} - ١ = ص.ص.$$

$$٠.١٦ - ١ =$$

$$٨٤ =$$

(د) الصدق :

يتوقف معامل الصدق على معامل الثبات فيزيد بزيادته وينخفض بانخفاضه . وكما رأينا فإن معامل الصدق لا يمكن أن يزيد عن الجذر التربيعي لمعامل الثبات . فإذا كان معامل ثبات أحد الاختبارات هو ٠.٨١ ، فإنه لا يمكن إجراء أى تعديل يؤدي إلى أن يرتفع معامل الصدق ، في حدود هذا الثبات ، عن $\sqrt{٠.٨١} = ٠.٩$. ولكننا رأينا أيضاً أن معامل ثبات الاختبار يتوقف على طوله ، فيزداد ثبات الاختبار بزيادة طوله إلى درجة معينة . فإن طال الاختبار أكثر مما يجب كان الجزء الأخير فيه متأثراً بالتعب والملل وتغير ظروف الإجراء . وما دام معامل الثبات يرتفع بإطالة الاختبار ، فإن معامل الصدق أيضاً يرتفع .

وعلى ذلك يمكننا إيجاد العلاقة بين معامل الصدق ومعامل الثبات عند إطالة الاختبار العدد (n) من المرات (وليكن ٢) بفرض أن معامل ثبات الاختبار بطوله الأول (مس.س) هو (٠.٧٠) وأن معامل صدق الاختبار قبل الإطالة (مس.س) هو (٠.٥٦) فيمكن حساب معامل ثبات الاختبار بعد مضاعفة طوله ، بتطبيق المعادلة :

$$\frac{\text{مس.س}}{n} + \frac{\text{مس.س} - 1}{n} \sqrt{\frac{\text{مس.س}}{2}} =$$

$$\frac{٠.٥٦}{٠.٧٠ + \frac{٠.٧ - 1}{2}} \sqrt{\frac{٠.٥٦}{2}} =$$

$$\frac{٠.٥٦}{٠.٧٠ + \frac{٠.٣}{2}} \sqrt{\frac{٠.٥٦}{2}} =$$

$$\frac{٠.٥٦}{٠.٧٠ + ٠.١٥} \sqrt{\frac{٠.٥٦}{2}} =$$

$$\frac{٠.٥٦}{٠.٨٥} \sqrt{\frac{٠.٥٦}{2}} =$$

$$\frac{٠.٥٦}{٠.٩٢} =$$

$$= ٠.٦١ \text{ تقريباً .}$$

- (هـ) مستوى صعوبة الفقرات :
- (و) خصائص العينة .
- (ز) طريقة القياس
- (ح) خصائص المقياس .
- (ط) الخطأ التجريبي في القياس :
- (ي) أخطاء الطرق الإحصائية :
- (ك) التخمين والصدفة والغش والتزييف .
- (ل) الانتباه والتذكر والحفظ والتعلم وأثر التدريب والنمو وأثر التعب والملل والألفة .
- (م) المستوى الوظيفي لفرد وفت الإجراء أو إعادته ذلك من حيث الصحة والإنفعال ومستوى القدرة كما تتأثر بالشخصية :
- (ن) طريقة حساب الثبات ومعادلته .

بيانات الثبات :

يختلف معامل ثبات الاختبار من مجموعة إلى أخرى من الأفراد بل وبين المستويات المختلفة من العينة الواحدة . وهذه الفروق لا يمكن التنبؤ بها إحصائياً ولكن يمكن توقعها بالتجربة أى بتجريب الاختبار على عينات مختلفة لحساب ثباته . فالاختبار الواحد قد يقيس وظائف مختلفة إذا طبق على أفراد من مستويات مختلفة من حيث القدرة . كما أن السن والذكاء والتجانس كلها خواص في العينة تؤثر في الثبات وفي طريقة الإجابة عن الفقرات ومدى الاعتماد على التخمين وعامل الصدفة .

والخلاصة أنه عند ذكر معامل ثبات الاختبار يجب توضيح تفاصيل التجربة من حيث عدد وخصائص العينة والفترة بين القياسات المتتالية وأنواع معاملات الارتباط وطرق حساب الثبات . كما يجب أن يجرب الاختبار على عينات التقنين التي تنقسم بدورها إلى مجموعات فرعية كل

مجموعة منها متجانسة من حيث خصائص معينة . ويراعى عند استخدام الاختبار فيما بعد أن يطبق الاختبار على عينات مشابهة حتى يمكن وضع مستوى للثقة وتحديد نسبة متوقعة لثبات درجات الأفراد فيما بعد فإذا اختير الفرد بناء على درجته على الاختبار كان الاختيار سليماً لأن درجته ثابتة . وإلا فإذا لم تكن الدرجة ثابتة بما يكفى ، أو كان خطأ التجريب فى القياس كبيراً ولم تكن الدرجة بالتالى ممثلة لمستوى القدرة الحقيقى لدى الفرد ، كان حكمنا مبنياً على قطاع صغير من السلوك ، قطاع لا يتميز بالثبات ولا بالشمول والجوهرية . ومن هنا تأتى أهمية حساب الثبات لافى الاختبارات والمقاييس وحسب بل وأيضاً فى كل التقديرات والتقييمات التى تصدرها على الأفراد والجماعات .

الفصل الثالث

تحليل الفقرات

- مقدمة .
- اعتباران في تحليل الفقرات .
- طرق تحليل الفقرات .
- الصديق المتبادل .
- ثبات الفقرات .
- وزن الإجابات عن الفقرات .
- استخدام صديق الفقرات .
- استخدام الاتساق الداخلي .
- خاتمة .

الفصل الثامن

تحليل الفقرات

مقدمة :

درجة الفرد على المقياس هي مجموع أو محصلة درجاته على فقراته .
فالفقرات في الاختبار ، كالوحدات في أى مقياس ، هي الجزء الذي تؤدي خصائصه إلى أن يتميز الاختبار بخصائص معينة . فمن ناحية الكم تكون الدرجة الكلية على الاختبار نتيجة للدرجات على فقراته وخاصة إذا كان من نوع مقياس النقط Point Scale حيث تكون الدرجة هي مجموع عدد الفقرات التي أجاب عنها إجابة صحيحة . ومن ناحية الكيف تقوم دراسة نتائج تطبيق الاختبار ، على حالات فردية ، على تحليل طريقة الإجابة والأخطاء النوعية والأخطاء الثابتة في الإجابة عن فقرات بذاتها .

ومن الناحية الإحصائية ، يعتمد حساب مختلف المعالم الإحصائية للاختبار على المعالم الإحصائية لفقراته ، فيتأثر المتوسط ، والانحراف المعياري ، والتباين في الاختبار بخصائص فقراته . ومن ناحية القياس ، يؤثر عدد ونوع ومستوى الفقرات على ثبات وصدق الاختبار .

وطبيعي أن الاختبارات المختلفة تحوي أنواعا مختلفة من الفقرات سواء من حيث الصياغة أو المواد أو المستوى . وقد عرضنا لهذا في الفصل الخامس كما سنتناوله تطبيقيا في قسم القياس التربوي . وعلى أى حال فإن تحليل الفقرات يمكن أن يعرف هكذا :

« تحليل الفقرات هو الدراسة التي تعتمد على التحليل المنطقي الإحصائي والتجريبي لوحدات الاختبار بغرض معرفة خصائصها وحذف أو تعديل

أو إبدال أو إضافة أو إعادة ترتيب هذه الفقرات حتى يتسنى الوصول إلى اختبار ثابت صادق مناسب من حيث الطول والصعوبة .

اعتباران في تحليل الفقرات :

بناء على ما تقدم نجد أننا نهتم في تحليل الفقرات بناحيتين :

(١) هي مستوى صعوبة Difficulty Level الفقرة بالنسبة لمن ستطبق عليهم وبالنسبة لموقعها في الاختبار .

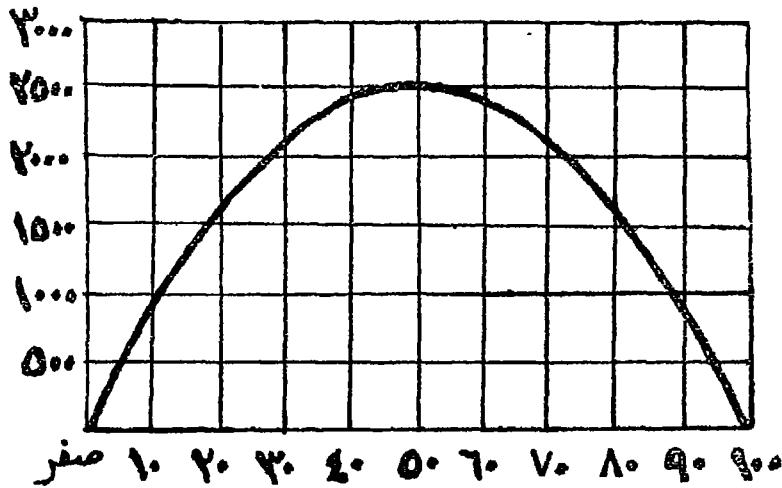
(ب) هي الدرجة التي تميز بها الفقرة بين مستويات قدرة المفحوصين ، روفى الناحية الثانية ، أى قوة التمييز Discriminative Power ، قد تقارن الأفراد على أساس أدائهم لمجموعة من الفقرات وهنا نهتم بالاتساق الداخلى internal consistancy بين الفقرات . وقد تقارن بين الأفراد على أساس محك خارجى هو أداء العمل وهنا نهتم بصدق كل فقرة Item Validity .

(١) مستوى الصعوبة :

إذا كان للفقرة أن تميز بين الدرجة العالية وبين الدرجة المنخفضة من السمة المقصودة ، ماوجب أن تكون سهلة بدرجة أن ينجح فيها الجميع أو صعبة يفشل فيها الجميع ؛ وعلى وجه العموم تحقق الفقرة الواحدة أقصى نجاح في التمييز بين الأفراد إذا كان مستوى صعوبتها يسمح بنجاح ٥٠٪ من أفراد المجموعة في الإجابة عليها ؛

فإذا أخذنا فقرة واحدة ، أمكننا أن نتصور احتمالات قدرتها على التمييز فيما يلى . فإذا أعطيت الفقرة لمفحوصين عددهم ١٠٠ ونجح فيها فرد واحد كانت تميزه عن الباقيين وعددهم ٩٩ . ويكون حاصل الضرب $1 \times 99 = 99$. أما إذا نجح فيها ١٠ أفراد وفشل ٩٠ كانت الفقرة تميز بين كل فرد من الناجحين العشرة وكل فرد آخر من جماعة الفاشلين التسعين . فإذا طبقت الفقرة على الأفراد فردين فردين كان عدد الحالات التي تنجح

فيها الفقرة في تمييز فردى الثنائيات هو $90 \times 10 = 900$ جالة : أما إذا نجح فيها ٥٠ فرداً وفشل ٥٠ فرداً فلإنها تميز بين كل عضو في مجموعة الحسنيين الناجحين وكل عضو آخر في مجموعة الحسنيين الفاشلين : فإذا أعطيت الفقرة لكل أفراد المجموعة اثنين فائتين مع التبدل نجحت في تمييز فردى الثنائي في $50 \times 50 = 2500$ حالة . ويبين الشكل التالي (شكل ٣) العلاقة بين مستوى الصعوبة ، معبراً عنه بالنسبة المثوية الناجحة في الإجابة على الفقرة وبين عدد المرات التي تنجح فيها الفقرة في التمييز بين الثنائيات .



(شكل ٣)

واضح إذن أن الفقرة التي ينجح فيها ٥٠٪ من الأفراد تميز ذا القدرة العالية عن ذي القدرة المنخفضة في ثنائيات الأفراد في عدد من هذه الثنائيات يفوق ما تستطيعه الفقرة الصعبة جداً أو السهلة جداً .

وتشير أغلب الدراسات النظرية والعملية إلى أنه يحسن الاعتماد على الاختبار الذي يميز بين أعضاء الجماعة المختبرة عندما يسمح متوسط مستوى صعوبة فقراته بنجاح ٥٠٪ تقريباً من الجماعة في كل فقرة : والتوزيع الصحيح لصعوبة الفقرات يضع في الاعتبار أيضاً العلاقة بينها : ففى حالة

ارتباط كل فقرة بكل فقرة أخرى ارتباطاً عالياً يتطلب الاختبار تشتتاً أكثر في مدى صعوبة فقراته حول مستوى الصعوبة المتوسط ٥٠ ٪. أما إذا انخفض معامل الارتباط بين كل فقرة والأخرى فيحسن أن يشمل توزيع صعوبة الفقرات مدى أصغر ، ولكننا لا نستطيع وضع توزيع أمثل لكل اختبار بحسب معاملات الارتباط الداخلية بين فقراته. إلا أن هناك اتفاقاً عاماً على أن الفقرة الصالحة متوسط مستوى صعوبتها يجب أن يسمح بنجاح ٥٠ ٪ من الأفراد عليها كما يجب أن توضح درجة من تشتت استجاباتهم عنها .

(ب) القدرة على التمييز :

نعني بالقدرة على التمييز قدرة الفقرة على أن تميز بين الأفراد الحاصلين على درجات مرتفعة وبين من يحصلون على درجات منخفضة في السمة التي تقيسها الفقرات كلها (أى الاختبار) . وهذا التمييز قد يكون على أساس الدرجة الكلية على الاختبار ، أو على أساس درجة على محك خارجي كالنجاح في الدراسة أو العمل . وهذا الأساس يختلف من حالة إلى أخرى تبعاً للغرض من إجراء الاختبار . وأحياناً نجمع الأساسين معاً . ولكي نوضح المواقف التي تدعو إلى تحليل العلاقة بين الفقرة وبين الدرجة على الاختبار ككل (وتسمى هذه العملية بتحليل الإتساق الداخلي internal consistency) analysis) والمواقف التي تدعو إلى تحليل العلاقة بين الفقرة وبين المحك الخارجي external Criterion (وتسمى هذه العملية بإيجاد صدق الفقرة item Validation) فإننا نعرض الآن لنوعين من أدوات القياس .

والنوع الأول من أدوات القياس نعطي فيه درجة واحدة كلية على الاختبار ، والدرجة الكلية على الاختبار لا تعني أن هناك وحدة أساسية في الفقرات نفسها ، ولكنها تعني الإعتبارات العملية في استخدام الاختبار .

وأوضح مثال على هذا النوع من أدوات القياس نجده في استمارة البيانات عن التاريخ الشخصي Bibliographical Data Blank التي تشمل أسئلة جد مختلفة عن التاريخ الماضي للفرد . ويشمل هذا النوع أيضاً استبيانات الشخصية Personality Questionnaires وقوائم الميول Interest Inventories . وعند إعداد هذه الأدوات يجعلها الباحث تشمل مدى واسعاً من الفقرات تكون العلاقات بينها سطحية . فهي ليست علاقة وظيفية أساسية ، ولكنها علاقة سطحية أتت من حقيقة أن كل الفقرات تشير إلى حقائق عن التاريخ الشخصي ، أو عبارات تتعلق بما يحبه الفرد وما لا يحبه ، أو عبارات عن المشاعر الفردية أو سلوك الفرد . فالسهولة العملية والتشابه السطحي في صياغة الفقرات يدفعان إلى جمع هذا العديد المتنوع من الفقرات في استمارة اختبار واحدة . وعلى هذا لا تكون هناك أسباب تجعلنا نعتقد أن فقرات هذا النوع تمثل جانباً موحداً مفرداً من النمو الفردي .

في مثل هذا الموقف يهمننا أن نتساءل عن صدق كل فقرة مفردة . فكل فقرة في ذاتها اختبار صغير مفرد . وفي الحكم على صلاحيتها يجب أن نهتم بصدقها . وليس هناك أساس مقدماً يحددنا في تصحيح الفقرات وجمعها في درجة واحدة . والأساس الوحيد في انتقاء الفقرات أو إعطاء درجة عنها هو الصدق العملي Emperical Validiyy لكل فقرة . ولهذا السبب يكون من المناسب ، في هذا النوع من الاختبارات ، أن نحدد قدرة كل فقرة (وقدرة كل استجابة محتملة ، بالنسبة لكل فقرة) على التمييز بين الأفراد على أساس محك خارجي كالنجاح في الدراسة أو العمل . وتكون الإستجابة على الفقرة صادقة إذا كان الأفراد الذين استجابوا بها أكثر نجاحاً فيما يتعلق بالمحك الخارجى ، أو كان أغلبهم أكثر نجاحاً ،

من الأفراد الذين لم يختاروا هذه الإستجابة لتلك الفقرة .
وعندما تمثل فقرات الاختبار المنفصلة مجالا غير متجانس ، من السلوك ، لا يكون هناك سبب لتوقع الارتباط بينها مقدماً . ولا يكون هناك ، بالتالى ، مبرر لتحليل علاقة الفقرة المفردة بالدرجة المعطاة عن مجموعة من الفقرات . وهنا لا نجد أساساً يمكننا من إعطاء درجة واحدة عن هذه المجموعة من الفقرات ، وبالتالى ، لا تكون للدرجة دلالة تدفعنا إلى بحثها بدراسة الإتساق الداخلى بغرض تحليل الفقرات .

أما النوع الثانى من أدوات القياس فيقوم على أساس قياس قطاع محدد معين من القدرة . هذا القطاع ، مثلاً ، قد يكون مدى المفردات . وفى مثل هذا الاختبار ، تهدف كل فقرة إلى الإسهام فى تقدير جانب موحد مفرد من القدرة . وفى هذا النوع من الاختبارات أيضاً يسهل التعرف مقدماً على الإجابة الصحيحة عن كل فقرة ، بل ويسهل أيضاً تصحيح الفقرات دون الرجوع إلى محك خارجى فإذا أعد الباحث ١٠٠ فقرة من هذا النوع ، أمكنه أن يعد مقدماً مفتاح التصحيح للاختبار دون أن ينتظر ما تأتى به نتائج التطبيق . وفى هذه الحالة علينا أن نتساءل عن درجة قياس الفقرة للوظيفة التى يقيسها الاختبار ككل . وقد يكون هذا ممكناً لأننا (١) نحصل على درجة عن الاختبار ككل (٢) ولأن هذه الدرجة لها دلالة لأن واضح الاختبار كان يهدف إلى أن تسهم كل فقرة فى قياس وظيفة واحدة متجانسة .

وعندما تكون فقرات الاختبار متجانسة بمعنى أن كل فقرة تقيس نفس العوامل أو الجوانب من قدرة الفرد بنفس تنظيمها يفقد ارتباط الفقرات المفردة بالمحك الخارجى دلالاته . فإما دامت الفقرة المفردة تقيس نفس الجزء من القدرة تماماً ، فإن صدقها يختلف وحسب باختلاف ثباتها : بمعنى أن

معامل ارتباطها بالحك الخارجي يعتمد فقط على معامل ثباتها أى ارتباطها بنفسها . وعندما تكون الفقرات متجانسة homogeneous حقاً ، ينعكس الفرق بين ثبات الفقرات على الفرق بين ارتباط هذه الفقرات بالدرجة الكلية على الاختبار . وفي هذه الحالة ، تجمع كل البيانات المتعلقة بالفقرات بواسطة القيام بتحليل الفقرات بالإتساق الداخلى ، ويكون تحليل الفقرات المفردة بربطها بحك خارجى أمراً مضيعاً .

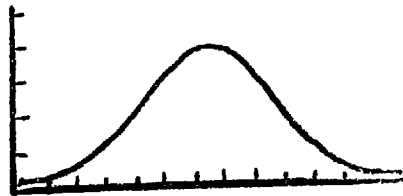
هذان النوعان من أدوات القياس طرفان لامتداد تقع على نقاطه المتفرقة أنواع أخرى متعددة . وأول الطرفين نجد الفقرة فيه يمكن أن تعد في ذاتها اختباراً كاملاً صغيراً ، وليست لها علاقة جوهرية بالفقرات الأخرى وبحسب صدقها بتقييم مزاياها . وعلى الطرف الآخر نرى اختباراً يتكون من فقرات كلها مقاييس متجانسة تماماً لنفس الوظائف في الفرد . وفي الاختبارات من النوع الأول تتم عملية تحليل الفقرات المفردة وحسب إيجاد صدق كل فقرة . وفي الاختبارات من النوع الثانى تتم عملية تقدير كل فقرة عند القيام بعملية تحليل الإتساق الداخلى . وكما قلنا ، لا نجد في القياس^{٢٦} حالتين متطرفتين تنطبق كل منهما على أحد النوعين ، فلكل اختبار درجة من التجانس ودرجة من التباين . وقد يسهل تصحيح الفقرة مقدماً كما قد يسهل الوصول إلى درجة كلية ، إلا أن الوظائف التى تقيسها الفقرات قد تكون متنوعة بدرجة تلزمنا بإيجاد صدق كل فقرة منها . ومن الجدير بالذكر أن نلاحظ أنه كلما زادت أهمية إيجاد صدق الفقرة المفردة قلت أهمية تحليل الإتساق الداخلى وكلما انخفضت دلالة صدق الفقرة المفردة زادت دلالة الإتساق الداخلى .

القياس الموضوعى للصعوبة :

قلنا إن الاختبار مقياس موضوعى . ولكن كيف يكون ذلك ؟ فى عرضنا السابق ، انضح ، أن التطبيق ، والتصحيح ، وتفسير الدرجات كل هذا موضوعى بمعنى أنها مستقلة عن الحكم الذاتى للفاحص نفسه . فدرجة المفحوص لا تتأثر بشخص الفاحص . هذا يحدث فى حدود معينة ، ما دما لم نتوصل إلى تقنين كامل ولا إلى موضوعية تامة : ولكن على الأقل ، هذه الموضوعية هى هدف تكوين الاختبار ، وقد حصلتها أغلب الاختبارات ، بدرجة كبيرة ، غالباً .

وهناك طرق هامة أخرى يمكن أن نحكم بواسطتها على موضوعية الاختبار بدقة فتحديد مستوى صعوبة الفقرة ، أو الاختبار كله ، وقياس ثبات الاختبار وصدقه ، مبنية على الموضوعية ، وهى خطوات عملية .

وليس ترتيب الفقرات وحسب ، بل وأيضاً اختيارها ، يمكن أن يعتمد على نسب من أجابوا عنها من العينات التجريبية . فإذا كانت عينة التقنين قطاعاً مستعرضاً ممثلاً من المجموع السكاني المراد قياسه ، كان من المتوقع أن ترسم الدرجات فى توزيعها منحنياً اعتدالياً . أى أن الأغلبية فى الوسط . وعلى الطرفين أقليتين . والانحدار من القمة إلى الذيل يميناً ويساراً تدريجى انظر الشكل (٤) ولا يمكن الحصول على المنحنى الاعتدالى إلا من عينات ضخمة ممثلة .



شكل (٤) المنحنى الاعتدالى

ولكن التواء المنحنى كما فى شكل (٥) ، وهو منحنى موجب الالتواء
Positively skewed فىعنى أن « أرضية » Floor الاختبار مرتفعة جداً بالنسبة
 لمن يراد تطبيقه عليهم ، وينقصه العدد الكافى من الفقرات السهلة التى تميز بدقة
 النهاية المنخفضة من المدى . وعلى ذلك فالأفراد المتوسطون الموزعون حول
 المتوسط سيحصلون على صفر أو درجات قريبة من الصفر :



(شكل ٥)

أما شكل (٦) وهو منحنى سالب الالتواء **Negatively skewed** فىعنى
 أن « سقف » Ceiling الاختبار منخفض فقد يحصل على الدرجات النهائية
 فيه بعض المفحوصين ولا يمكننا أن نعرف به أيهم أفضل أو أيهم يكون
 أداؤه أفضل لو كانت الفقرات أصعب :

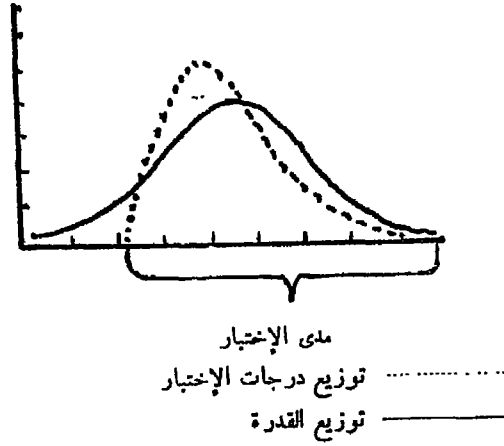


(شكل ٦)

وبعد رسم المنحنى يمكن أن نعدله بحيث يقرب من الاعتدال ذلك أنه لو كان ملتوياً نحو السهولة أمكننا إضافة بعض الفقرات الصعبة وحذفنا بعض السهولة وعدلنا البعض الآخر . ونفس الشيء ، بالعكس ، لو كان صعباً . على أنه أيضاً يمكن علاج الأمر بإعطاء وزن أكبر للفقرات الأصعب وأصغر للأسهل . وهكذا يمكن أن نجد نسبة كبرى من المفحوصين تستطيع الحصول على درجات عن ٥٠ ٪ من الفقرات . وهكذا يمكن الوصول إلى أقصى حد للتفريق بين الأفراد في كل مستويات القدرة

وشكل (٧) يدل على أن الفقرات لا تقيس كل مدى القدرة ولا تصلح لتقييم المتطرفين من ضعاف القدرة أو المتفوقين فيها . كما أن قمة الاختبار تعلقو قمة القدرة أى أن الاختبار يقيس مستوى بسيطاً من القدرة

(شكل ٧)



طرق تحليل الفقرات

(أ) دلائل صعوبة الفقرة :

نتعرف على مستوى صعوبة الفقرة من النسبة المئوية من الأفراد ، في مجموعة معينة ، التى تستطيع أن تجيب عن السؤال أو أن تحل المشكلة :

وكلما صغرت النسبة المئوية الناجحة في الفقرة ، زادت صعوبة الفقرة والعكس صحيح . وقد انتقد هذا الدليل على أساس أن طبيعة البيانات تتحدد بما تطلبه الإجابة عن الفقرة وليس تبعاً لتعقد الوظيفة العقلية التي تعمل في الإجابة عن هذه الفقرة . فليست هناك صعوبة في الوظيفة العقلية المستخدمة في معرفة أن $8 \times 9 = 72$ أكثر مما هو الحال في معرفة أن $3 \times 4 = 12$: كما أن الوظيفة العقلية الأساسية المستخدمة في الإجابة عن السؤال « ما هو الكنغر ؟ » بأنه حيوان ليست أصعب من الوظيفة المستخدمة في الإجابة عن السؤال « ما هو الجمل ؟ » بأنه حيوان . إذ أن الصعوبة في السؤال الأول متوقعة على ندرة وجود هذا الحيوان في بيئتنا لا على تعقد المفهوم الذي نقيسه في الحالتين . وهكذا نجد أن من يجيب على أحد السؤالين لا يقل في الوظيفة المقاسة عن من يجيب عن السؤالين إذ أن هذا يرجع إلى صغر محصول خبرة الفرد وهو أمر لا علاقة له بالذكاء مثلاً . وعلى هذا فإن نسبة النجاح في الفقرة لا تعبر تماماً عن مستوى صعوبتها ، إلا أننا مازلنا نقبل هذا المبدأ حتى نجد دليلاً أفضل .

فالاعتماد على النسبة المئوية للناجحين في الفقرة ، من بين المجموع السكاني الذي طبقت عليه ، كدليل على مستوى صعوبتها أمر تحوطه عقبتان ، فربما كان بعض الناجحين في إجابة الفقرة قد أجاب بالصدفة ، بالتخمين ، أو بطريقة لا تدل على زيادة معلومات الفرد أو على تفكيره في المشكلة المعطاة له . والعيب الثاني هو أننا لو تصورنا أن القدرة التي يقيسها الاختبار موزعة اعتدالياً بين الأفراد ، كانت الفروق المتساوية بين النسب ، المئوية للنجاح غير متعلقة بدرجات مكافئة من الصعوبة أي أنه لا يمكن تمثيل العلاقة بين نسب الناجحين ودرجات الصعوبة بخط مستقيم . وسنعالج التعديلات التي يمكن إدخالها على هذه الطريقة للتغلب على هذين النقصين .

وأول نقص سنعالجه هو تدخل عوامل الصدفة محاولين تقدير النسب المثوية لأفراد الجماعة الذين توصلوا إلى الإجابة الصحيحة خلال معرفة صحيحة أو استدلال صحيح وأن نستبعد الذين أجابوا بالتخمين **guessing** . وأى تصحيح للتخمين يتضمن شيئاً من التقريب . وهذا التقريب يعتمد على افتراضين . والافتراض الأول هو أن الفرد الذى يختار إجابة غير صحيحة عن الفقرة تنقصه المعلومات الصحيحة أو الفهم السليم وليس السبب هو سوء الفهم أو خلط المعلومات . والافتراض الثانى هو أن الشخص الذى لا يعرف الإجابة الصحيحة من بين عدة إجابات مرفقة بالفقرة تجذبه كل هذه الإجابات ليختار واحدة منها بنفس الدرجة من الجذب فهو لا يعرف أنها الصحيحة . إذا قبلنا هذين الفرضين ، توقعنا وجود نسبة معينة من الأفراد الذين لا يعرفون الإجابة الصحيحة فيختارون بمحض الصدفة إجابة يحدث أن تكون صحيحة . هذه النسبة المثوية من الأفراد هى متوسط النسب المثوية لمجموع الأفراد الذين يختارون كل إجابة من الإجابات الخاطئة . ونحن نستبعد هذه النسبة المثوية من النسبة المثوية للناجحين فى الفقرة لنصل إلى تقديرنا للنسبة المثوية التى عرفت حقاً الإجابة الصحيحة عن الفقرة . فتكون المعادلة هى :

$$\frac{\frac{خ}{1-هـ} - ع}{ع + خ + هـ} = \text{النسبة المثوية}$$

حيث النسبة المثوية هى النسبة المثوية للأفراد الذين يعرفون الإجابة الصحيحة عن الفقرة فعلاً :

، ح هى عدد من أجابوا إجابة صحيحة عن الفقرة .

، خ عدد من أجابوا إجابة خاطئة عن الفقرة نفسها .

، لك هو عدد من تركوا هذه الفقرة دون أى إجابة برغم أن الوقت سمح لهم بقراءتها .

، ن هو عدد متغيرات الاستجابة عن نفس الفقرة .

مثال :

أعطيت إحدى الفقرات لمفحوصين عددهم ٢٠ فرداً ، فأجاب ١١ فرداً منهم إجابة صحيحة وأجاب ٥ أفراد إجابة خاطئة ، وكان الوقت كافياً ولكن بعض المفحوصين قرأوا الفقرة ولم يعرفوا الإجابة إذ ترك ٤ أفراد الفقرة دون إجابة علماً بأن الفقرة متبوعة بعدد ٣ متغيرات يختار من بينها الفرد ، ما هي النسبة المئوية لمن يعرفون الإجابة الصحيحة فعلا عن هذه الفقرة بإعتبار النسبة دليل على مستوى سهولة الفقرة .

الإجابة :

النسبة المئوية لمن أجابوا بناء على معرفة صحيحة

$$\frac{٨٥}{٢٠} = \frac{\frac{٥}{٢} - ١١}{٢٠} = \frac{\frac{٥}{١-٣} - ١١}{٤ + ٥ + ١١} =$$

$$= ٠.٤٢٥ = ٤٣\% \text{ تقريباً .}$$

وبناء على هذا المفهوم لا يكون هناك غموض في المعادلة إلا الرمز كـ [

الدال على عدد من اختاروا أن يتركوا الفقرة : أى أن المفحوص هنا قرأ

الفقرة ثم تركها مختاراً. وفي اختبار السرعة أو الاختبار الموقوت Timed or speeded قد لا يفرق البعض بين الفقرات التي تركها المفحوصون مختارين لأنهم لم يعرفوا الإجابة عنها وبين الفقرات التي لم يصلوا إليها لضيق الوقت. وهناك أساس على التمييز بين الحالتين ، هو أن نفترض أن كل فرد وصل إلى فقرة معينة في إجابته عن الاختبار قد حاول كل الفقرات التي قبلها والتي تليها مباشرة . أما الفقرات ابتداء من التي تلي آخر فقرة أجاب عنها فنفترض أن الوقت لم يسمح بالإجابة عنها . وعلى هذا تكون الفقرات المتروكة قبل آخر فقرة أجاب عنها المفحوص هي فقرات لم يستطع إجابتها أو لم يعرف إجابتها . ونحن نرمز للفقرة الواحدة منها بالرمز ك .

فإذا كان الاختبار يتكون من ١٠٠ فقرة أجاب أحد الأفراد عنه حتى وصل إلى الفقرة رقم ٧٠. فإننا نفترض أن الفقرات من ٧٢ حتى ١٠٠ ، وهي ٢٩ فقرة ، لم يتسع الوقت لحلها وتكون الفقرات من ٧١ حتى ١ كلها فقرات حاول إجابتها فإذا وجدنا بينها ٣ فقرات متروكة إفتراضنا أنها أصعب من مستواه في الوظيفة المقاسة فإذا تبين لنا أن ٢٠ فرداً من إحدى المجموعات تركوا إحدى الفقرات بنفس الطريقة كانت قيمة ك في المعادلة هي ٢٠ .

يجب أن نلاحظ أن المعادلة السابقة تعتمد على الافتراضين معاً . فإذا كانت إحدى متغيرات الإجابة التي تتبع الفقرة المعنية متميزة عن بقية المتغيرات اختلف مقدار جذبها ، للأفراد الذين لا يعرفون الإجابة الصحيحة ، عن مقدار جذب غيرها . يحدث هذا سواء كانت واضحة الصواب إذ أن احتمال أن يختارها الفرد ، بطريق الصدفة أو التخمين ، الذي لا يعرف الإجابة الصحيحة يزيد عن احتمال أن يختار غيرها . ويحدث نفس الشيء لو كانت المتغيرة واضحة الخطأ إذ يقل احتمال أن يختارها نفس الفرد برغم أنه لا يعرف الإجابة الصحيحة . كما أن مقدار جذب كل متغيرة يتوقف على

عدد متغيرات الإجابة التي يجب على المفحوص أن يختار بينها : فإحتمال اختيار إحدى ٣ متغيرات يزيد عن احتمال اختيار إحدى ٥ متغيرات : وعلى هذا يجب أن نحدد عدد الاختيارات أمام الفرد في كل الفقرات بقدر الإمكان كما يجب أن يتساوى جذب كل متغيرة منها :

ومن ناحية أخرى ، نجد أن الأفراد الذين يجيبون عن الفقرة إجابات خاطئة لا يفعلون هذا نتيجة لجهلهم وخطئهم في التخمين بل نتيجة لعدم صحة معلوماتهم أو لخطأ العملية العقلية المستخدمة . فإذا كانت أسباب خطأ الأفراد في الإجابة كلها أسباب أصيلة لما تدخلت الصدفة أو التخمين . وفي هذه الحالة أيضاً يكون كل الأفراد الذين أجابوا إجابة صحيحة قد فعلوا هذا لأن معلوماتهم صحيحة ولأنهم استخدموا عمليات عقلية سليمة . هنا لا يكون محل للصدفة ، ولا يكون هناك مجال لهذه المعادلة . وفي الواقع نحن لا نعلم إلى أى مدى يتدخل هذان العاملان (العامل الأول هو تدخل عامل التخمين في كل من الإجابات الصحيحة وغير الصحيحة ، والعامل الثاني هو الجهل أو المعرفة وسوء استخدام العمليات العقلية أو سلامة التفكير كأسباب أصيلة تؤدي إلى الإجابات الخاطئة أو الصحيحة) . ولهذا فنحن نفترض أن العاملين يعوضان أثر بعضهما فتكون النتيجة هي أن النسبة المثوية الناجحة للفقرة هي النسبة المثوية من أفراد المجموعة التي عرفت الإجابة الحقة أو التي حلت المشكلة .

وبالرغم من أن القيمة المصححة بالنسبة المثوية التي عرفت الفقرة هي قيمة تقريبية وهي تتحدد بالافتراضات التي ناقشناها ، إلا أنها تعطي صورة أصدق من القيمة الخام . وهي طريقة تفيد في مقارنة مستوى صعوبة الفقرات التي (١) تختلف اختلافاً واضحاً في تكرار عدد المفحوصين الذين تركوها لصعوبتها . (٢) وبالنسبة للفقرات التي تختلف في عدد المتغيرات . وفي مقدار جذب كل منها : فالنسبة المثوية الخام لمن أجابوا إجابة صحيحة لا يمكن مقارنتها في هذه الحالات :

ويمكن التغلب على العيب الثاني من عيوب النسبة المئوية ، وهو عدم تساوى الوحدات ، (انظر الفصل الثانى والتاسع) للناجحين فى الفقرة بإعتبارها دليلا على مستوى الصعوبة إذا افترضنا بعض الفروض المتعلقة بشكل توزيع السمة المقاسة فى المجموعة المختبرة . فالافتراض العادى أن السمة موزعة اعتداليا . ومن هنا تكون النسب المئوية ممثلة فى قيم على المقياس على المحور الأفقى الذى رسمنا عليه المنحنى الاعتدالى . وعلى هذا ، تكون نسبة النجاح ٥٠٪ مقابلة للقيمة صفر على المقياس ، ونسبة النجاح ٧٥٪ مقابلة للقيمة - ٠.٦٧ ، وتكون نسبة النجاح ٢٥٪ مقابلة للقيمة + ٠.٦٧ ، وهكذا . . . وبهذا النوع من التدرج ، يمكننا أن نقارن صعوبة الفقرات التى حاولت الإجابة عنها جماعات مختلفة ، وهنا تتبع طريقة معينة فى تحديد الفرق فى المتوسط والانحراف المعياري للقدرة بين الجماعتين . ولمعرفة المزيد عن هذه الطريقة يمكن الرجوع إلى :

Davis F.B., Item analysis Data, Harvard Education Papers
No. 2, Graduate School of Education, Harvard University,
Cambridge, Mass , 1946.

(ب) دلائل قدرة الفقرة على التمييز :

اقترح الكثيرون عدة دلائل توضح الدرجة التى تكون بها الفقرة فعالة فى التمييز بين أصحاب القدرة العالية وأصحاب القدرة المنخفضة ، ونحن نعرف على مستويات القدرة هذه من الدرجة الكلية على الإختبار أو من محك خارجى آخر . وليس هذا مجال عرض هذه الدلائل جميعا إذ أنها إحصائية غالبا . ولكننا سنختار أبسطها وأميزها .

ونحن نهتم هنا بناحيتين : الأولى ، أننا قد نربط الأداء على الفقرة^٦ بأحد المقاييس المستمرة . وهذا المتغير المتصل أو المستمر غالبا هو الدرجة

على الاختبار ، الذى يشمل الفقرة المدروسة ، ككل . ولكن هذا المقياس المستمر قد يكون أحد المحكات الخارجية . وفى الناحية الثانية ، نربط الأداء على الفقرة بمتغير متقطع ، سواء أكان محكا أم كان الدرجة على الاختبار الكلى . والدرجة على الاختبار الكلى هى متغير مستمر ، ولكننا نستطيع تحويلها إلى متغير متقطع إذا حددنا درجة معينة من يفوقها يعتبر ذا قدرة عالية ومن يقل عنها يعتبر ذا قدرة منخفضة .

(١) تحليل الفقرات فى حالة الدرجة المستمرة :

عندما نحلل الفقرة على أساس متغير موزع إعتداليا ، تكون أماننا عدة طرق يمكن استخدامها . وأول طريقة هنا ، نستخدم فيها درجة موزعة توزيعا مستمرا ، ولكنها طريقة شاقة . وفى الطريقة الثانية نحدد نقطة على المقياس المستمر تقسم الدرجات إلى مجموعتين عليا ودنيا ، وتعتبر المجموعتان متغيراً متقطعا . وفى الطريقة الثالثة نقارن بين مجموعتين متطرفتين ، والتطرف هنا يكون فيما يتعلق بالمتغير المستمر . وذلك بأن نحدد درجة مرتفعة جداً كل من يعلوها يعتبر متطرفا موجبا ، ثم نحدد درجة منخفضة جداً من يقل عنها يكون مجموعة متطرفة سالبة (تبعا للذيل المنحنى الاعتدالى) ثم نقارن بين أداء أفراد هاتين المجموعتين على كل فقرة ندرسها .

وسنبداً بالحديث عن المتغير المستمر كدرجة كمية Quantitative Score فعندما نرغب فى تحديد العلاقة بين كل فقرة وبين الدرجة الموزعة توزيعا مستمرا ، نستخدم معامل الارتباط ذى الشعبتين (الثنائى Biserial أو الارتباط الثنائى الأصيل Point Biserial) . وفى هذه الحالة نعتبر أداء الفقرة متغيرا متقطعا بمعنى أن الفرد يحصل فى الفقرة على درجة أو لا يحصل على شئ (صفر) وتكون الدرجة الكلية على الاختبار متغيراً مستمرا . ونوجد درجة الارتباط بين الفقرة المفردة والدرجة الكلية على الاختبار .

وإذا رأينا أن النجاح أو الفشل على الفقرة يمثل قدرة مستمرة ، وأن التقسيم إلى أفراد ناجحين وأفراد فاشلين فيما يتعلق بهذه القدرة هو تقسيم ممكن ، كان اختيارنا للنقطة التي تفصل بين المجموعتين متوقفاً على صعوبة الفقرة المعنية . كان من المستحسن أن نستخدم معامل الارتباط الثنائي . يحدث هذا في حالة دلالة النجاح أو الفشل كعرض Symptom يخفى وراءه قدرة أو سمة معينة . ونحن نعلم أن قيمة معامل الارتباط الثنائي لا تتوقف على نسبة المجموعة الناجحة ونسبة المجموعة الفاشلة وهذا يعنى أن دليل صدق الفقرة أو تمييزها أمر منفصل تماماً عن صعوبتها .

ومعامل الارتباط الثنائي الأصيل يتميز عن معامل الارتباط الثنائي في طبيعة الفروض التي يقوم عليها وفي تأثير صعوبة الفقرة على القيم الناتجة (وهى درجات على الاختبار) . فمعامل الارتباط الثنائي الحقيقي يفترض أن من ينجحوا في الفقرة يكونوا جماعة متميزة عن جماعة من يفشلون . فهو يفترض أن الجماعتين تفرقان عن بعضهما إلا أن أعضاء كل جماعة يعتبرون متشابهين . وعلى هذا فقيمة الارتباط الثنائي الأصيل في تحليل الفقرات تعبر عن صعوبة الفقرة . ولكن افترض أن المجموعتين متميزتان ليس صحيحاً دائماً . فإن الناجحين والفاشلين كلهم يمثلون مدى من القدرة تستطيع الفقرة إلى حد ما أن تقيسه . ويجب هنا أن نفرق بين صدق الفقرة وصعوبة الفقرة . والليس بينهما يرجع إلى استخدام معاملات مثل الارتباط الثنائي الأصيل الذى يعطى قيمة تمثل الصعوبة والصدق معاً . ومن الأفضل أن نحصل على معلومات عن الصدق أو القدرة التمييزية على حدة وعن صعوبة الفقرة على حدة ثم نجمع بين هذه الحقائق بطريقة واضحة . وأخيراً فإن حساب معاملات الثبات الثنائية بالنسبة لكل فقرة عمل شاق جداً .

٢ - تحليل الفقرات في حالة الجماعات المتقطعة :

في عمليات تحليل الفقرات نضحي بالطبيعة المستمرة المتغير الذي نربط به الفقرة في سبيل تسهيل عمليات الحساب . فنعتبر أن من يحصلون على درجات عالية يمثلون مجموعة منفصلة عن مجموعة من حصلوا على درجات منخفضة برغم أن متغير الدرجة قد يكون مستمراً غالباً . ويحسن أن تقسيم الدرجة التي يختارها الأفراد إلى نصفين وتستخدم نفس الخطوات عندما يكون المتغير متقطعاً بالفعل مثل المحك (كالتقسيم إلى مريض وسوى أو ناجح وفاشل) :

فإذا قسمنا المجموعة بهذه الطريقة ، كانت أماننا عدة طرق لتقييم الفقرة . وأول هذه الطرق بسيط جداً ويخدم كثيراً من أغراض واضعي الاختبارات غير الرسمية informal Tests ويستعين بها المحللون في تحسين وتعديل الاختبارات . وفي هذه الطريقة نقوم بحساب النسبة المئوية الناجحة على الفقرة من أفراد المجموعة التي حصلت على أعلى الدرجات في الاختبار ككل ثم نحسب النسبة المئوية للأفراد الناجحين في نفس الفقرة من المجموعة التي حصلت على أدنى الدرجات في الاختبار . ثم نقعد مقارنة بين هاتين النسبتين المئويتين . وهذه الطريقة مفيدة في استبعاد الفقرات غير الصالحة . وخطوات هذه العملية هي : عزل أوراق إجابة المفحوصين الذين حصلوا على درجات عالية جداً عن أوراق إجابة من حصلوا على درجات منخفضة جداً . ثم جمع عدد الإجابات الصحيحة ، وغير الصحيحة والفقرات المتروكة لصعوبتها ، والفقرات التي لم يستطع الفرد الإجابة عنها لضييق الوقت . وسواء أكنّا نستخدم النسبة المئوية الخام للإجابات الصحيحة ، أو النسبة المئوية المعدلة لاستبعاد تدخل الصدفة ، فإننا نقارن بين المجموعتين في الإجابة عن كل فقرة على هذه الأسس .

وقد أجرى تعديل على هذه الطريقة بتقسيم الأفراد الذين طبق عليهم

الإختبار إلى مجموعة ناجحة وأخرى فاشلة ، ثم تحديد نسبة الناجحين في الإجابة عن كل فقرة ضمن المجموعتين الكبيرتين كل على حدة . ثم يستخدم عامل الارتباط الرباعي أو معامل phy لمعالجة هذه النسب المئوية . ويجب أن نضع في ذهننا أننا ندرس استجابة فقرة وليس أننا ندرس متغيراً متقطعاً منفصلاً . وعلى هذا تكون لدينا أربعة نسب مئوية توضع في جدول من أربعة خلايا وبحسب الارتباط بين هذه الخلايا إما بمعامل الارتباط الرباعي أو بمعامل فاي . وإختبار أحد معاملي الارتباط متوقف على افتراض استمرار توزيعي المتغيرين موضع الدراسة . فإذا افترضنا أنهما مستمران وموزعان اعتدالياً ، استخدمنا الارتباط الرباعي . وإذا افترضنا أن المتغيرين متقطعان استخدمنا ارتباط فاي . ومعامل الارتباط الرباعي لا يتأثر بصعوبة الفقرة ، بينما يمثل ارتباط فاي ، كما كان الحال ، في الارتباط الثنائي الأصيل ، مستوى صعوبة الفقرة المبنية .

(ج) تحليل الفقرات في حالة الجماعات المتطرفة :

إذا قررنا أن نترجم درجاتنا المستمرة إلى فئتين نقسم المجموعة الكبيرة تبعاً لهما إلى مجموعتين عليا ودنيا ، كنا بصدد إمكان تحليل هاتين المجموعتين المتطرفتين ، وإستقاط باقي الأفراد في المنطقة المتوسطة . فإذا كانت العلاقة بين درجة الاختبار ودرجة الفقرة علاقة مستقيمة Linear (أي أن زيادة أحدهما تؤدي دائماً إما إلى نقص الأخرى وإما إلى زيادتها) بحيث أن النسبة المئوية للناجحين في الفقرة تزايد بمعدل ثابت تزايد به الدرجة الكلية ، كان أخذ مجموعتين من الطرفين مؤدياً إلى تضخيم الفروق التي نلاحظها في الفترة المفردة . فالفترة تستطيع التمييز الواضح بين العشر الأعلى Top tenth والعشر الأدنى bottom tenth في المجموع السكاني المدروس أكثر مما تستطيع بين النصف ، الأعلى والنصف الأدنى . وبرغم أن لهذه الطريقة ميزة توضيح

الفروق وإبرازها إلا أنها تتجاهل الأفراد في الأقسام الثمانية الأخرى (المحصورة بين العشر الأعلى والعشر الأدنى) فبالرغم من أن الفقرة تميز بوضوح بين أعلى ١٠ أفراد وأدنى ١٠ أفراد في مجموعة من ١٠٠٠ فرد أكثر مما تميز بين أعلى ٥٠٠ وأدنى ٥٠٠ ، إلا أننا نتوقع أن مقارنة الفقرات المختلفة إذا اعتمدت على العشريين لا يمكن الاعتماد عليها على عكس المقارنة بين النصفين . وعلى ذلك فإن علينا أن نبحث عن التمييز مع الإبقاء على الثبات . فهل هناك نقطة فاصلة يكون فيها الاتزان أفضل بين حدة التمييز وبين ثبات القيم التي نحصل عليها ؟ بينت كيلي في :

Kelley T. L., "The Selection of the Upper and Lower Groups for the Validation of Test Items", J. Educ. Psychol., 30, 17-24 (1939).

إن النسبة بين الفروق والخطأ المعياري تبلغ أقصاها عندما تتضمن المجموعة العليا ٢٧٪ والمجموعة الدنيا ٢٧٪ تقريباً من كل المجموع السكاني المدروس . وأن زيادة أو قلة النسبة المثوية عن هذا الحد تؤدي إلى انخفاض الدقة التي نرتب بها الفقرات مبتدئين من أقل الفقرات قدرة على التمييز إلى أكثرها . وهكذا ندرس $27 + 27 = 54$ ٪ من الحالات المتطرفة ونرفض ٤٦٪ من الحالات المتوسطة ، ونسهل عمليات الحساب ونزيد دقة نتائجنا .

(هـ) تحليل الفقرات في حالة محك متقطع :

قد يتطلب الحال أن نربط النجاح على الفقرة بمحك متقطع مثل النجاح في الدراسة أو العمل (بالتقسيم إلى ناجح وفاشل مثلاً) . وهنا لا تقسم المجموعة المختبرة ، غالباً إلى قسمين متساويين ، ونوجد معامل الارتباط إما باستخدام الارتباط الرباعي أو ارتباط فاي .

الصدق المتبادل

يشير الصدق المتبادل Cross Validation إلى حقيقة أن صدق الاختبار يجب أن يتحدد بعينة من الأشخاص تختلف عن العينة التي جربت عليها فقراته عند اختيارها . فأى معامل صدق حسب على نفس العينة التي استخدمت في أغراض اختيار الفقرة يتضخم على حساب أخطاء الصدفة في حدود هذه العينة ويرتفع بشكل مبالغ فيه . والحق أن معامل الصدق المرتفع للاختبار في مثل هذه الحالات ، يخفى وراءه أن الاختبار ليس صادقا إطلاقاً في التنبؤ بمحك معين .

فإذا افترضنا أن لدينا ١٠٠ طالب طب ، اخترنا منهم أعلى ٢٧ وأدنى ٢٧ في درجاتهم الدراسية ، ليمثلوا محكاً من مجموعتين متقابلتين . فإذا قارنا بين هاتين المجموعتين في عدد من السمات التي لا تتعاق فعالاً بالنجاح في كلية الطب ، فسنجد بالتأكيد فروق الصدفة . فقد نجد مثلاً زيادة عدد الحضريين عن الريفيين داخل مجموعة المحك العليا . فإذا أعطينا درجة عن الإقامة أو الميلاد في إقليم حضري زادت درجات مجموعة المحك العليا عن مجموعة المحك الدنيا أكثر من ذى قبل . وليس في هذا دليل على صدق الاختبار كأداة للتنبؤ . فقد اخترنا عوامل تتغير بالصدفة وتميز هذه العينة بالذات ، بمعنى أننا لو اخترنا مجموعة أخرى فقد نجد أن الريفيين فيها أكثر .

وترفع معاملات الصدق ، بهذا الشكل المبالغ فيه ، في حالة اختبارات الشخصية حيث لا تصحح الدرجة أساساً على أنها صحيحة أو خاطئة . وفي هذه الحالة ، يتحدد اختيار وتصحيح الفقرات باستجابات العينة الأولى . وعلى هذا تحصل الاستجابات في مجموعة المحك العليا على الوزن + ١ ، وتشيع الإجابات التي تحصل على الوزن - ١ في المجموعة الدنيا . وفي مثل

هذه الأحوال ، تتدخل فروق الصدفة في أي من جماعتي المحك ، وتكرر :
العلاقة بين الاختبار والمحك عالية بلوحة مبالغ فيها . وفي :

Kurtz, A. K. A Research test of the Rorschach test.
Personnel Psychol. 1948.

نجد مثالا يدلنا على الحاجة إلى بحث الصدق المتبادل . فقد قام Kurtz في بحث على اختبار الورشاخ لتحديد ما إذا كان هذا الاختبار يفيد في انتقاء مديري المبيعات في إحدى شركات التأمين ، فطبق الاختبار على ٨٠ مديرا . وقد اختير هؤلاء المديرين من بين عدة مئات من موظفي ٨ شركات تأمين . وحددت جماعة المحك العليا بحيث كانت ٤٢ مديرا ناجحين في مقابل جماعة المحك الدنيا التي تكونت من ٣٨ مديرا اعتبروا فاشلين . وقام أحد الخبراء في اختبار الورشاخ بدراسة السجلات الثمانية ، وقد اختار مجموعة من العلامات تبلغ ٣٢ علامة هي مميزات في الاستجابة تكررت بنسبة أكبر في إحدى جماعتي المحك أكثر مما تكررت في الأخرى . والعلامات التي تكررت كثيراً في جماعة المحك العليا أعطيت الوزن + ١ عند من استجاب بها والوزن صفر إذا لم يستجب بها عضو الجماعة العليا ، والاستجابات التي تكررت أكثر في جماعة المحك الدنيا أعطيت الوزن - ١ إذا حدثت والوزن صفر إذا لم تحدث في استجابات عضو هذه الجماعة . وقد بلغ عدد العلامات التي تكررت في المجموعة العليا ١٦ علامة ووزنها + ١٦ كحد أقصى ، كما بلغ عدد العلامات التي تكررت في المجموعة الدنيا ١٦ علامة ومجموع وزنها كحد أدنى - ١٦ . أي أن الدرجة الكلية انحصرت نظرياً في المدى بين - ١٦ ، + ١٦ .

وعند إعادة تطبيق مفتاح التصحيح المبني على ٣٢ علامة على المجموعة الأصلية وعددها ٨٠ شخصاً وجدنا أن ٧٩ منهم قد وزعوا نفس التوزيع السابق إلى مجموعة عليا وأخرى دنيا . وعلى هذا كان معامل الارتباط بين

بين درجة الاختبار والحك قريباً جداً من ١ . وعند ما بحث الصدق المتبادل على عينة مقارنة ثانية مكونة من ٤١ مديراً ، ٢١ في المجموعة العليا و ٢٠ في المجموعة الدنيا فانخفض معامل الصدق إلى ٠.٢٠ . ومن هنا اتضح أن المفتاح الذي أنشئ في العينة الأولى كان غير صادق في اختيار هؤلاء المديرين .

أوضح البحث السابق مدى زيف النتائج متأثرة بعوامل الصدفة .
وهاكم بحثاً آخر يوثق هذا نجده في :

Cureton, E. E. Validity. Reliability and baloney. Educ, Psychol. Measnet 1950.

والمفحوصون في هذه الدراسة ٢٩ تلميذاً بالسنة النهائية في التعليم الثانوي وبالجامعة يدرسون المقاييس النفسية . والحك الذي طلب التنبؤ به كان متوسط درجة التلميذ . والفقرات تكونت من ٨٥ قصاصة من الورق مرققة من ١ حتى ٨٥ وكل الأرقام على وجه واحد ولإعطاء كل مفحوص درجة ، كانت القصاصات توضع في وعاء وتهرثم تسقط على المائدة . وكل القصاصات التي سقطت بحيث ظهر الوجه المرقم اعتبرت دالة على وجود تلك الفقرة المعينة في أداء التلميذ للاختبار . أدى إلقاء القصاصات (وعددها ٨٥) مرات عددها ٢٩ إلى سجلات كاملة لكل تلميذ ، يبين وجود أو غياب كل فقرة أو علامة في الاستجابة . هذه العملية ، بالطبع ، هي مجرد أداة سريعة للحصول على نتائج بمحض الصدفة . وقد أقيم تحليل لهذه الفقرات باعتبار متوسط الدرجة الحقيقية لكل تلميذ محكاً . وعلى هذا الأساس اختار الباحث ٢٤ فقرة من ٨٥ فقرة في الاختبار ، ومن هذه الفقرات الـ ٢٤ وجدنا ٩ فقرات تكررت أكثر بين التلاميذ الذين حصلوا على أعلى الدرجات ، و ١٥ فقرة شاعت فيمن حصلوا على أدنى الدرجات : وأعطى الوزن + ١ لكل فقرة من الفقرات التسعة ، وكان وزن كل

فقرة من الـ ١٥ فقرة هو - ٠١ مجموع أوزان هذه الفقرات بالنسبة لكل تلميذ يكون درجته الكلية . وبالزغم من معرفة أن الصدفة حرفت درجات هذا الاختبار إلا أن معامل ارتباطها بمحك الدرجة الحقيقية للتلاميذ في المجموعة الأصلية وعددها ٢٩ بلغ ٠.٨٢ . ومثل هذه النتائج تتفق مع بحث اختبار الرورشاخ . ففي كلتا الحالتين ، نشأ الارتباط الظاهر بين درجة الاختبار والمحك من استخدام نفس فروق الصدفة سواء عند تصميم مفتاح التصحيح أو عند مراجعة الصدق .

ويجب أن نلاحظ أن هذا الارتفاع الكبير لمعامل الصدق لا يحدث وحسب في اختيار الفقرات بل وأيضاً في اختيار الاختبارات لتكوين بطارية . فإن معامل الارتباط المتعدد بين البطارية والمحك ينخفض عادة عندما يعاد حسابه على عينة جديدة . وهذا الانخفاض المفاجئ سيكون صغيراً في حالة إنشاء البطارية ، ما دامت نسبة الاختبارات المحذوفة في العملية التالية صغيرة نسبياً ، وفي تحليل الفقرات ، من المعتاد أن نستبعد فقرات أكثر من تلك التي نبقى عليها . وعندما يكون عدد الفقرات الأصلية كبيراً ونسبة الفقرات التي أبقيناها صغيرة ، تزيد فرصة تضخم قيمة فروق الصدفة وبالتالي نحصل على معامل صدق مرتفع بشكل زائف . ولهذا السبب تزيد الحاجة إلى الصدق المتبادل عند تحليل الفقرات منها عند إنشاء البطارية .

وفي أي من الحالتين يزيد الزيف كلما صغرت العينة . فإدام معامل الصدق راجعاً إلى تجمع أخطاء العينة ، فإن زيادة هذه الأخطاء تؤدي إلى زيادة هذا الزيف ، ويحدث هذا بالطبع كلما صغر حجم العينة ، فعندما تتكون العينة الأصلية من عدة آلاف حالة ، قد يبلغ معامل الصدق بطريقة الصدق المتبادل رقماً صغيراً جداً . وعلى هذا فإن العوامل المحددة الأساسية هي حجم العينة ، وعدد الفقرات ، ونسبة الفقرات التي أبقيت وتصبح

الحاجة ماسة عندما تكون العينات صغيرة ، والفقرات الأصلية كثيرة ، ونسبة الفقرات التي أبقيناها قليلة . وفي مثل هذه الظروف ، نجد معامل الصدق المحسوب من نفس العينة المستخدمة في اختيار الفقرة معاملاً كاذباً .

ثبات الفقرات :

الاختبار مقياس وحداته الفقرات يعتمد صدقه على صدقها اعتماد ثباته على ثباتها . وعلى ذلك فكما بحثنا صدق الفقرة ، باعتباره مؤدياً إلى صدق الاختبار ، نبحث ثبات الفقرة بعدة طرق أهمها طريقة إعادة الاختبار وطريقة الاحتمال المتوالى Modal Probability .

وفي طريقة إعادة الاختبار نسجل نتائج الإجراء الأول في جدول كل فرد في عمود وكل فقرة في صف ويقسم الجدول إلى قسم للإجابات الصحيحة وآخر للإجابات الخاطئة . وترصد نتائج إعادة الاختبار بنفس الطريقة ثم يحسب معامل الارتباط الرباعي بين مرقى الإجراء :

وفي طريقة الاحتمال المتوالى ، نستخدم المعادلة التالية :

$$\text{معامل ثبات الفقرة} = \frac{n}{1-n} (1 - \frac{1}{n})$$

حيث n هي عدد متغيرات الإجابة التي يستطيع أو يجب أن يختار إحداها المفحوص بالنسبة للفقرة وليكن (٥) .

و 1 هي الاحتمال المتوالى أى النسبة بين أكبر عدد اختار إحدى متغيرات الإجابة وبين مجموع عدد الأفراد الذين طبقت عليهم هذه الفقرة وليكن عدد من اختاروا إحدى متغيرات الإجابة هو (٢٥) وعدد من طبقت عليهم هذه الفقرة هو (٥٠) فتكون النسبة $1 = \frac{25}{50} = ٥٠\%$.

إذن معامل ثبات هذه الفقرة $\frac{0}{1-0} = (0.50 - \frac{1}{2})$

$$\frac{2}{8} = (0.30) \frac{0}{4} = (0.20 - 0.50) \frac{0}{4} =$$

$$= 0.375 = 0.38 \text{ تقريباً } \bullet$$

وتفيد طريقة الاحتمال المتوالى كما رأينا في حالة إيجاد ثبات الفقرات التي تتطلب الإجابة عنها الاختيار من عدة إجابات أو الإجابة بإحدى الاحتمالات التي يستطيع الباحث أن يحصرها . ولمعرفة المزيد عن هذه الطرق ارجع إلى :

Hollzinger, K. J. Reliability of a single test item, J, Ed.
Vol 23. No 6.

وزن الإجابات عن الفقرات :

المشاكل التي صادفتنا عند وزن الاختبارات التي تتكون منها البطارية تعود لتواجهنا عند وزن الفقرات التي يتكون منها الاختبار . وتبعاً لإحدى المعادلات الإحصائية ، وهي معادلة الانحدار المتعدد Multiple Regression Equation ، يجب علينا أن نربط الفقرات بعضها ببعض الآخر كما نربط كل فقرة منها بالحكم ، ولكن هذا الأمر جد مرهق فهو يتطلب القيام ، عادة ، بحساب آلاف معاملات الارتباط . وللتغلب على هذه الصعوبة اشتقت عدة معادلات . ولكننا دائماً نراعى أن يرتفع معامل الارتباط بين كل فقرة وبين الحكم إلى أقصى حد ، في حين تنخفض معاملات الارتباط الداخلية بين

الفقرات وبعضها ، وإلا فإن الفقرة التي ترتبط ارتباطاً منخفضاً بالملحك يجب أن ترتبط ارتباطاً عالياً بفقرة أخرى أو بالفقرات الأخرى حتى تصبح ذات قيمة في التنبؤ بالملحك .

ولتبسيط معادلة الانحدار قامت محاولات كثيرة أهمها الطريقة التي نجدها في :

Richardson, M. W. and Adkins, D. C. : A Rapid Method of selecting test items J, educ. Psychol. 1938, 29.
وهذه الطريقة تحسب وزن كل فقرة بتطبيق المعادلة الآتية عليها :

$$W = \frac{r_{c \cdot s} - r_{c \cdot s} \cdot r_{c \cdot s}}{(r_{c \cdot s} - r_{c \cdot s} \cdot r_{c \cdot s})^2}$$

حيث (و) هو وزن الفقرة .

، (ماس ع) هو معامل الارتباط بين الفقرة والملحك وليكن (٠٩٠) ،
، (ماس ع) هو معامل الارتباط بين الفقرة ودرجة الاختبار الكلي
الشامل لهذه الفقرة وليكن (٠٣٠)

، (ماس ص) هو معامل الارتباط بين درجة الاختبار الكلي والملحك
ولتكن (٠٦٠)

، (ا) هي نسبة الناجحين في الفقرة ولتكن (٠٨٠)

، (ب) هي نسبة الفاشلين في الفقرة ولتكن (٠٢٠)

$$W = \frac{0.90 - 0.30 \times 0.60}{(0.90 \times 0.80 - 0.60 \times 0.20)^2}$$

$$\frac{0.18 - 0.90}{\sqrt{(0.27 - 0.60) \cdot 0.16}} =$$

$$\frac{0.72}{(0.33) \cdot 0.4} =$$

$$\frac{0.72}{0.132} =$$

$$\frac{6}{11} =$$

$$0.56 =$$

إذن وزن الفقرة هو ٠٥٦ ٪

أى أن من يجب عن هذه الفقرة إجابة صحيحة يحتمل له النجاح في العمل أو الدراسة التى نطبق الاختبار لانتقاء الملتحقين بها ويكون الاحتمال بمقدار ٥٦ ٪ واحتمال فشل نفس الفرد هو ٤٤ ٪ كما أن هذا يعنى أنه لو أجاب ١٠٠ فردا إجابة صحيحة عن هذه الفقرة توقعنا أن ينجح منهم (في المحك الذى نتنبأ له) ٥٦ فرداً وأن يفشل ٤٤ فرداً ٪

وتطبق هذه الطريقة في حالة الاهتمام بمعادلات الارتباط الداخلية بين الفقرات ، أما إذا أهملنا هذه المعاملات ، أو افترضنا أنها متسارية بين الفقرات المختلفة ، فإنه يمكننا أن نعطي وزناً للتصحیح بطريقة جيلفورد Guilford والمعادلة هي :

$$و = \frac{\text{نسبة جماعة المحك العليا} - \text{نسبة جماعة المحك الدنيا}}{\text{النسبة من الجماعة العليا} \times \text{النسبة من الدنيا}} + ٤$$

حيث (و) هي وزن طريقة معينة في الإجابة عن الفقرة .
 أ' ونسبة الجماعة العليا ، في المحك أو التي حصلت على درجات عالية في الاختبار ككل ، هي النسبة المثوية من الناجحين في المحك التي أجابت بها .
 الطريقة عن تلك الفقرة ولتكن (٠٧٠) أي أن (٠٣٠) من الناجحين أجابوا إجابات مختلفة .

ونسبة جماعة المحك الدنيا هي نسبة الأفراد من جماعة الفاشلين ، في المحك أو الاختبار ككل ، التي أجابت بالطريقة نفسها عن الفقرة نفسها ولتكن (٠٤٠) أي أن (٠٦٠) من الفاشلين أجابوا بطريقة مختلفة .

$$\therefore و = \frac{٠٣٠ - ٠٧٠}{٠٣٠ \times ٠٧٠} + ٤$$

$$٤ + \frac{٠٤}{٠٢١} =$$

$$= ٥٩$$

ونتيجة هذه المعادلة تتراوح بين صفر ، ٨ : وإذا كان الوزن الناتج هو ٤ كان ارتباط الفقرة بالمحك صفراً . وهذه المعادلة تطبق بوجه ضار في حالة تساوي مجموعتي المحك العليا والدنيا في عدد الأفراد في كل منهما . وإذا لم تتسوى المجموعتان أو كانت نسبة الناجحين في المحك الذين أجابوا إجابة معينة على فقرة معينة مساوية لنسبة الفاشلين الذين أجابوا بهذه

الطريقة ، في هذه الحالة سيكون الوزن ٤ والارتباط بين الفقرة والحك صفرأ
والخطأ المعياري كبيراً. أى في حالة فرض الصفر أو العدم null hypothesis
وهنا يحسن استخدام المعادلة التالية :

$$\sqrt{\frac{2}{n-1}} = \epsilon_r$$

حيث ϵ_r هي الوزن المعدل للفقرة

، n هي مجموع عدد الفاشلين وليكن (٢٠) والناجحين وليكن (٣٠)

، a هي نسبة الناجحين في المجموعة ولتكن (٠.٢٠) التي أجابت
بطريقة معينة :

، b هي نسبة الفاشلين في المجموعة ولتكن (٠.٣٠) التي أجابت
بنفس الطريقة .

$$\sqrt{\frac{2}{0.30 \times 0.20 \times 50}} = \text{الوزن المعدل} =$$

$$\sqrt{\frac{2}{3}} =$$

$$\sqrt{\frac{2}{17321}} =$$

$$= 0.0115$$

أهمية وزن الإجابات :

وتنحصر أهمية وزن الاستجابات عن الفقرات في أن اختبارات الشخصية عامة قد اتبعت طرقاً عقيمة في وزن الاستجابات على فقراتها . إلا أن بعض الاختبارات قد أفاد من وزن الإجابات ، مثل اختبار سترونج للميول المهنية (Strong, Vocational Intrests Blank (V.I.B.) واختبار برنروتر للشخصية Bernreuter Personality Test ، في رفع ثباتها . والتصحيح غير الموزون Unweighted Scoring هو إعطاء الاستجابة صفراً أو (١) . كما استفادت الاختبارات من وزن استجاباتها في رفع صدقها . وعلى هذا ، يكون من المهم أن نتأكد من أن الفقرات ترتبط ارتباطاً جوهرياً بالحركات : وبهذه الطريقة يمكننا ، بتطبيق الاختبار على عينات كبيرة ، أن نختار فقرات يمكن الاعتماد عليها وأن ننشأ اختباراً يمكن الاعتماد عليه مادامنا قد طبقناه على عينات صدق ممثلة أخرى غير العينات التي استخدمناها في اختيار الفقرات .

استخدام صدق الفقرات

قلنا إنه عند تحليل الفقرات نهتم بأمرين ، أولهما تحديد مستوى صعوبة الفقرات والعمليات المستخدمة في حسابه هي عمليات الاتساق الداخلي . والأمر الثاني هو القدرة التي بها تميز الفقرة بين المفحوصين ، تمييزاً يتفق مع تمييزهم في محك آخر يقيس ما تقيسه الفقرة وهو غالباً محك خارجي آخر كالنجاح في العمل أو الدراسة ، هذه العملية هي إيجاد معامل صدق الفقرة . وكما زاد عدد الفقرات وكانت غير متجانسة ، أي انخفضت معاملات الارتباط الداخلية بينها ، زاد احتمال أنها تقيس جوانب مختلفة من القدرة . وهنا لا يفيدنا إيجاد الارتباطات الداخلية بغرض تحديد الاتساق الداخلي ،

بل نهم بإيجاد صدق كل فقره . فنهال كل فقره بإيجاد معامل الارتباط بينها وبين الحك الخارجى . وهكذا يكون لدينا معامل صدق لكل فقره وكأن كل فقره أصبحت اختباراً منفصلاً . وأصبح لدينا إذن عدد كبير من معاملات صدق اختبارات عديدة مختصرة جداً . وتكون المشكلة ، إذن ، هى جمع هذه الاختبارات فى اختبار واحد كبير على أن يتمتع بأعلى معامل ارتباط مع الحك . ونحن هنا نقوم بنفس العمل الذى قنا به عند جمع عدة اختبارات فى بطارية واحدة فزاعى ارتفاع معامل الصدق وانخفاض درجة توحيد uniqueness الفقرات فى قياس الجانب المعين . بحيث تصبح الفقره صادقة إلى درجة كبيرة وتقيس شيئاً يختلف عن الذى تقيسه فقره أخرى .

فبعد الاختبار (فى حالة البطارية) أو تعد الفقره (فى حالة الاختبار) ذات قيمة تدفنا إلى فقرات أخرى أو اختبارات أخرى إذا توفر فيها الشرطان الآتيان : (١) ارتفاع معامل صدقها (٢) انخفاض معامل ارتباطها مع الفقرات الأخرى . فزيد قيمة الفقره كلما كانت تقيس جانباً لا يقيسه فقرات أخرى . فإذا أعطينا للفقرات أوزاناً وكانت متساوية ظهرت مشكلة اختيار أى هذه الفقرات ضمن الاختبار . على أن يكون العدد المختار من الفقرات اختباراً يحصل فيه الأفراد على درجات ترتبط أعلى ارتباط مع الحك ، أى تبلغ أقصى درجة ممكنة من الصدق . وحل هذه المشكلة يمكن أن يتم بالطرق الثلاثة الآتية :

الطريقة الأولى :

يمكننا أن نتجاهل معاملات الارتباط الداخلية بين الفقرات تماماً . وهكذا يكون الأساس الوحيد فى اختيار الفقرات هو صدق كل منها على حدة . وعلى ذلك يجب أن يحدد مستوى مناسباً للصدق إذا بلغته الفقره

كانت صالحة وإلا استبعدت من الاختبار . ويحسن ، في أنواع معينة من المقاييس كالاستفتاءات وتواريخ الحياة والاستبيانات ، أن تحلل كل إجابة محتملة عن الفقرة بالطريقة الآتية :

لنفرض أنه في أحد اختبارات التشخيص diagnosis وجدنا الفقرة التالية :

ما هو حجم الجماعة التي تميل إلى الانضمام إليها ؟

(أ) ٢ شخصاً

(ب) من ٢ - ٥ شخصاً

(ج) من ٥ - ٨ شخصاً

(د) من ٨ - ١٢ شخصاً

(هـ) أكثر من ١٢ شخصاً

فعلينا هنا أن نحلل هذه الإجابات المحتملة للفقرة الواحدة فإذا فرضنا أن ٥٠ ٪ من الأفراد الذين أجابوا مختارين الاختيار (أ) كانوا مصابين باضطراب معين ، وأن ٤٤ ٪ ممن اختاروا (ب) كانوا مرضى بأحد الأمراض . وأن ٣٢ ٪ ممن اختاروا (ج) كانوا مضطربين ، وأن ٢٣ ٪ من الأفراد الذين وافقوا على الاختيار (د) كانوا يعانون من أزمات نفسية . وأن ١٥ ٪ ممن أجابوا مختارين (هـ) كانوا مصابين بأحد النوبات . كل هذه الإجابات إذن ترتبط بالمحك الخارجي وهو هنا التشخيص . أى أن كل إختيار له درجة من الصدق ويعطى وزناً يحسب هذا الصدق . وتكون هذه الفقرة مناسبة لاختبار غرضة التشخيص .

فإذا كانت لدينا فقرات ذات معاملات صدق مختلفة كيف نفاضل :

بينها ؟ وما هو معامل الصدق الذى نتوقع أن يصل إليه الاختبار المكون من هذه الفقرات ؟ لنفرض أن لدينا ٥ فقرات (أ) معامل صدقها ٠.٢٥ ،

(ب) معامل صدقها (٠.٢٠) ، (ج) معامل صدقها ٠.١٥ ، (د) معامل صدقها ٠.١٠ ، (هـ) معامل صدقها ٠.٠٥ : فباستخدام إحدى المعادلات الإحصائية (ارجع إلى كتب الإحصاء) يمكننا أن نحسب معامل صدق الاختبار الذى يشمل الفقرات (أ) ، (ب) والاختبار الذى يتكون من (أ) ، (ب) ، (ج) ، والاختبار الذى يتكون من (أ) ، (ب) ، (ج) ، (د) . والاختبار الذى يتكون من (أ) ، (ب) ، (ج) ، (د) ، هـ : ذلك يفرض أن معاملات الارتباط بين كل فقرة والأخرى من هذه الفقرات الخمس هو ٠.٥٠ دائماً مرة و ٠.٢٠ دائماً فى أخرى وصفر من مائة فى الثالثة . بهذه نحصل على الجدول التالى :

معامل الارتباط بين كل الفقرات			الفقرات المتضمنة
٠.١٠	٠.٢٠	٠.٥٠	
٠.٢٥٠	٠.٢٥٠	٠.٢٥٠	(أ)
٠.٣١٨	٠.٢٩٠	٠.٢٦٠	(أ) ، (ب)
٠.٣٤٦	٠.٢٩٣	٠.٢٤٥	(أ) ، (ب) ، (ج)
٠.٣٥٠	٠.٢٧٧	٠.٢٢٢	(أ) ، (ب) ، (ج) ، (د)
٠.٣٣٥	٠.٢٥٠	٠.١٩٤	(أ) ، (ب) ، (ج) ، (د) ، (هـ)

جدول (هـ) يوضح صدق الدرجة الكلية عن خمسة فقرات لكل معامل صدقها . عن

Thorndike, R. L., Personnel selection N. y. : John wiley sons 1949..

ومن الجدول يتضح لنا :

١ - أنه كلما انخفض معامل الارتباط الداخلى بين الفقرات ، على أن تكون كل منها صادقة ، زاد معامل صدق الاختبار ككل .

٢ - أن زيادة الفقرات غير الصادقة كالفقرة (هـ) يخفض معامل صدق الاختبار ككل .

٣ - أن زيادة فقرة ذات معامل صدق صغير لا يزيد من صدق الاختبار بل يخفضه غالباً كما في إضافة الفقرة (د) .

٤ - أنه كلما زاد طول الاختبار زاد صدقه ، على أن تنخفض معاملات الارتباط الداخلية بين فقراته .

فالمشكلة إذن هي اختيار مجموعة الفقرات التي تكون اختباراً يحصل فيه أفراد العينة ، التي سيطبق عليها ، على درجات تتمتع بأعلى معامل صدق ممكن . وهذه المشكلة ، بدورها ، تتضمن مشكلة أخرى وهي تحديد القيمة المناسبة التي يجب أن يبلغها صدق الفقرة حتى يمكن إضافتها إلى عينة القدرة المقاسة أى إلى الفقرات الأخرى في الاختبار . وهنا يجب أن ندخل في الحساب انحدار معاملات صدق الفقرات من العينة الأولى (سواء كانت العينة هي فقرات الاختبار الأصلي الذي أخذت منه لتضاف إلى عينة جديدة هي فقرات الاختبار الذي نود إطالته أم كانت العينة هي جماعات المفحوصين التي كان الاختبار يطبق عليها ونريد تطبيقه على عينات أخرى كما يحدث عند نقل اختبار أجنبي إلى بيئة أخرى) ، ولا يتسع المقام لشرح هذه العمليات الإحصائية التي يمكن الرجوع فيها إلى كتب الإحصاء .

وإذا كانت لدينا مجموعة من الفقرات لا نعرف معاملات الارتباط الداخلية بينها ، علماً بأننا نسلم بأثر الفقرات على صدق الدرجة الكلية على الاختبار ، فلا يمكننا حل هذه المشكلة بطريقة تحليلية إذ أننا لا نستطيع بهذه الطريقة تحديد مدى هذا الأثر .

الطريقة الثانية :

يمكننا معالجة مشكلة انتقاء الفقرات لوزنها في مفتاح التصحيح بأن نضع في الاعتبار معاملات الارتباط الداخلية بين الفقرات بعد عمل سلسلة من

عمليات التقريب . والعملية العامة في هذا هي وضع مفتاح أولى للاختبار يستند إلى مجموعة صغيرة من الفقرات الأكثر صدقاً . ويحصل كل فرد على درجة باستخدام مفتاح التصحيح هذا . ويحدد صدق هذه الدرجات . ثم يحسب ارتباط كل فقرة بهذه الدرجة ، فباستخدام المفتاح الجديد لإعطاء الدرجات نكون قد استخدمنا الدرجة الكلية كمتغير متصل يفيد كحك في تحليل الفقرات . وقد وضع Flangan هذه الطريقة . وهكذا نحصل بالنسبة لكل فقرة على :

١ - معامل ارتباطها بالحك .

٢ - معامل ارتباطها بالمجموعة الأولى من الفقرات الموزونة .

٣ - ونعرف أيضاً صدق الدرجة عن المجموعة الأولى من الفقرات .

٤ - وهذه الطريقة نستطيع أن نحسب بالنسبة لكل فقرة معامل ارتباطها الجزئي بالحك عندما تصبح الدرجة عن الفقرات التي اخترناها أولاً مقداراً ثابتاً Constant .

٥ - وهذا يدلنا على فائدة كل فقرة بالنسبة لصدق المجموعة الأولى من الفقرات . أي أن تحليل كل من الفقرات التي اخترناها أولاً يدلنا على ما تضيفه الفقرة إلى باقي الفقرات من رفع معامل صدقها .

ونستطيع أن نختار الآن مجموعة ثانية من الفقرات لتصحيحها ، مختارين الفقرات ذوات معاملات الارتباط الجزئية الكبرى . هذا الاختبار لا يضع في الاعتبار صدق الفقرة الواحدة فحسب ، بل وأيضاً ارتباطها بالمجموعة الأولى من الفقرات المصححة . وهكذا تصبح المجموعة الثانية من الفقرات هي أحسن الفقرات . ويمكننا تكرار العملية الكلية ، إذا أردنا فنصحح المجموعة الثانية من الفقرات في الاختبارات ، ونحدد ما تزيده كل فقرة على صدق المجموعة الثانية :

والطريقة الثالثة :

لتقرير أفضل مجموعة من الفقرات يبدأ الباحث من المجموعة الكلية الموجودة من الفقرات ، مستبعداً الفقرات غير الصالحة . ذلك طبعاً بعد تحديد صدق كل فقرة . ويعد مفتاح تصحيح الاختبار الكامل . وهذا المفتاح يمكن أن يعد قبل إجراء الاختبار في حالة اختبارات المعرفة أو القدرة . أما في مقاييس الميول ، والتوافق ، وتاريخ الشخصية وما يشابهه ، فيعد المفتاح على أساس صدق الفقرات . وبعد إعداد المفتاح تصحح كل ورقة إجابة بواسطته . ثم نحصل على معاملات لإرتباط كل فقرة بالدرجة الكلية على الاختبار . كما نحسب معامل الارتباط بين الدرجة الكلية والمحك .

وبهذه الطريقة ، يمكننا أن نحسب الارتباط الجزئي لكل فقرة في المحك عندما تصبح الدرجة على الاختبار الكلي مقدراً ثابتاً . فإذا أردنا اختصار المقياس أمكننا حذف الفقرات ذوات معاملات الارتباط الجزئية الصغرى .

ويمكن استبعاد الفقرات باتباع أكثر من منهج . فمثلاً تستبعد الفقرات التي لا يبدو أن لها قيمة ويصحح الاختبار المختصر بالنسبة لكل فقرة . ويحدد صدق درجته ، وارتباط كل فقرة بالدرجة على الاختبار المختصر . ويجعل الدرجة الكلية على الاختبارات المختصرة مقداراً ثابتاً تحسب معاملات الارتباط الجزئية مع المحك وهي الأساس في استبعاد الفقرات الأقل صلاحية .

وتواجهنا من جديد مشكلة عدد الفقرات التي نحذفها من المجموعة الأصلية من الفقرات . وفي بعض الحالات يتحدد هذا العدد بالطول المرغوب أن يكون عليه الاختبار النهائي . وتفيد معاملات الارتباط الجزئية ، في الطريقة الثانية والثالثة ، في تعيين الفقرات التي ستستبعد . وفي حالات أخرى لا يكون هناك مثل هذا الاعتبار العملي ، وهنا يمكن إنقاص عدد الفقرات الكلية في خطوات متتابة . وفي كل خطوة ، يحسب معامل

! الارتباط بين الدرجة على الاختبار الكلى (بعد اختصاره) ودرجة المحك . ونظـل تستبعد فقرات ما دام معامل الارتباط بين الدرجة الكلية والمحك يـزايـد . ونـتوقـف فـى الـاختـصار عـند ما يتناقص هذا المعامل .

استخدام الإتساق الداخلى

نركن إلى فحص الإتساق الداخلى للفقرات عند ما نكون بصدد اختبار فقراته متجانسة ، بمعنى أن كل فقرة تهدف إلى قياس نفس الوظيفة التى تقيسها الفقرات الأخرى . فنحدد ارتباط كل فقرة بالدرجة الكلية عن مجموعة الفقرات التى يتكون منها الاختبار . وهذا النوع من التحليل ، والذي يمكن القيام به فور إجراء الصورة الأولية من الاختبار على مجموعة تجريبية ، يقدم لنا معلومات ذات قيمة فى اختيار فقرات الاختبار وفى مراجعة تأليف الفقرات . وسنعالج استخدام بيانات الاتساق الداخلى فى هذين المجالين .

نفيد من بيانات الاتساق الداخلى فى اختيار الفقرات بطريقة معقدة يحسن معها استخدام صدق الفقرات فى نفس الغرض . ويرتبط الاتساق الداخلى للفقرات المفردة بتجانس الاختبار الكلى . فكلما ارتفعت درجات الاتساق الداخلى زاد تجانس الاختبار ولكن ما هى درجة التجانس التى نرتضيها للاختبار ؟ قبل الإجابة عن هذا السؤال يجب أن ننبه إلى أن التجانس ليس هدفاً فى ذاته . فهو يعنى صدق درجة الاختبار ، وصدق درجات بطارية من الاختبارات ، أى مدى وضوح تفسير الدرجة على الاختبار من الناحية التحليلية .

فالفقرة ذات الارتباط المنخفض جداً بالدرجة الكلية عن الاختبار فقرة لا يمكن الاعتماد عليها وهى غالباً تقيس وظيفة تختلف تماماً عن تلك التى تقيسها بقية فقرات الاختبار . وعلى هذا فالفقرات التى ترتبط ارتباطاً

شديد الانخفاض أو سالباً مع الاختبار ككل ، فقرات لا يمكن الاعتماد عليها ويجب أن تستبعد تماماً أو على الأقل أن تعدل وتجرب من جديد . فإذا كنا نهم بتكوين اختبار نعطي عنه درجة مفردة لها معنى موحد ، أى درجة تتفق مع اسم وظيفة معينة ، كان علينا أن نستبعد الفقرات الأقل اتساقاً ، ولكن هل كل الفقرات الأكثر اتساقاً صالحة ؟

يجب أن يكون للفقرة ذات الاتساق مع الاختبار درجة عالية من الثبات . ويجب أن تتداخل مع الوظائف التي تقيسها الفقرات الأخرى بالاختبار ، مادام الاختبار متماسكاً فيه وحده . والاختبار الذي يقيس وظيفة واحدة ويعطي درجة واحدة يجب أن يغطي كل جوانب هذه الوظيفة .

فإذا افترضنا أن الصورة الأولية من اختبار افترض فيه قياس مدى المعرفة في العلوم الطبيعية ، تتكون من ٢٠٠ فقرة وكانت ٨٠ فقرة تتعلق بالطبيعة و ٨٠ تتعلق بالكيمياء ، و ٢٠ فقرة تتعلق بالجيولوجيا و ٢٠ بالفلك . فإذا نظرنا إلى أحد جوانب هذه المعرفة توقعنا أن معامل الارتباط الداخلى بين الفقرات في قسم الطبيعة مثلاً يزيد عن معامل الارتباط بين فقرات قسم الطبيعة وفقرات أى قسم آخر . ولكننا نتوقع أيضاً أن قسمي الطبيعة والكيمياء لسيادتهما في الاختبار (١٦٠ فقرة تخصهما من ٢٠٠ فقرة في الاختبار كله) ، أن معامل الارتباط بين الفقرة في أحدهما وبين الاختبار يرتفع . فإذا اخترنا ١٠٠ فقرة من فقرات الاختبار كله على أساس ارتفاع الاتساق الداخلى ، وجدنا أن هذه الفقرات محصورة في قسمي الطبيعة والكيمياء . وهذا يضيق المدى الذى يفترض أن يغطيه الاختبار الذى يحقق الغرض الأصيل من إنشائه . وهكذا نرى أنه عند بحث الاتساق الداخلى يلزمنا أن نتحدد بالإطار الذى يضعه التخطيط الأصيل والمواصفات الأولى للاختبار .

وفحص الفقرات على أساس الاتساق الداخلى يتعلق بثبات وتجانس الدرجة على الاختبار الكلى الناشئ . وهكذا نصل فى نهاية الأمر إلى أن بيانات الاتساق الداخلى تفيدنا كوسيلة لاكتشاف واستبعاد الفقرات ذات الارتباط المنخفض أكثر مما تفيدنا فى اختيار الفقرات الصالحة ؛

والمشكلة التى تواجهنا هنا هى تقرير مدى اتساع محتوى الاختبار والوظيفة التى يقيسها حتى يصبح متجانساً . ويفترض منطقياً أن درجة الاختبار الكلى عند ما تمثل وظيفة ومحتوى محدوداً معروفاً جيداً تكون أفضل . وعلى هذا عند تكوين اختبار يقيس وظيفة مركبة من عدة اختبارات فرعية كل منها يقيس ميداناً من ميادين هذه الوظيفة ، يحسن بنا أن نربط درجة الفقرة من نوع معين بالاختبار الفرعى الذى يشملها والذى يتكون من فقرات من نفس النوع . وهنا تعطى درجة واحدة عن كل اختبار فرعى .

فبيانات الاتساق الداخلى يجب أن تستخدم أولاً وقبل كل شئ فى استبعاد الفقرات غير الصالحة . ويمكن أن تفيد هذه البيانات فيما يتعلق بصعوبة الفقرة وبوضع تخطيط لمواصفات الاختبار ، وفى اختيار الفقرات الأصلح فى كل قطاع . متخصص من المحتوى أو الوظيفة المقاسة . وفضلاً عن هذا فإن بيانات الاتساق الداخلى فى تحليل الفقرات تعد مصدراً قيماً للمعلومات يرشد مؤلف الاختبار أثناء مراجعة فقرات اختبارهِ .

الفصل التاسع

المعايير

- مقدمة .
- الحاجة إلى المعايير .
- الصفر المطلق .
- المعايير والتقنين .
- الدرجات المحولة .
- أنواع المعايير .
- تحويل المعايير .
- النسب .
- الصفحات النفسية .
- كيف نستخدم المعايير .

الفصل التاسع

المعايير

مقدمة :

القياس والتقييم في التربية وعلم النفس علم تعرّضه عدة صعوبات أهمها أن الكثير من الصفات الإنسانية ليس خاضعاً للقياس بطريقة مباشرة ميسورة . ففي حين نجد أنه يسهل قياس الطول لأنه كمّي ولأنه يزيد بوحدات منتظمة نسميها السنتيمترات وكذا الوزن والعمر الزمّني . نجد أن الكثير من الصفات الإنسانية الأكثر تعقّداً مثل سمات الشخصية ، والخصائص العقلية ، والتحصيل ، ليست لها مقاييس متفق عليها موضوعية دقيقة دقة وموضوعية مقاييس الطول والوزن والسن .

وقد حدث منذ عصور طويلة إن كان قياس الطول والوزن والأيام والشهور يصادف مشاكل متعددة في القياس فلم تكن هناك مستويات للمقارنة ، ولا وحدات متفق عليها للقياس في هذه الميادين . وتواجهنا ، في أيامنا هذه نفس المشاكل ، من إقرار وحدات لقياس السلوك البشري في المجالات التي لم تحدد لها معايير مماثلة . وليست المشكلة هي إيجاد الوحدات بقدر ما هي البحث عن وحدات ذات دلالة ومعنى يمكن تفسيرها ومقارنتها . والمشكلة أيضاً هي وضع مستويات بها نقارن ونقيس ونقيم ونفسر ونعلل السلوك الإنساني . هذه الوحدات هي المعايير .

الحاجة إلى المعايير

إذا رجعنا إلى الفصل الثاني (مثال درجات التلاميذ في تفسير الدرجة الخام) وجدنا أن الدرجة الخام ومثيلاتها النسبة المئوية لعدد الناجحين إلى

العدد الكلى ، أو النسبة المئوية لعدد الفقرات الصحيحة إلى عدد فقرات الاختبار ، ليس لها فى ذاتها معنى أو دلالة . فهى لا تحدد مركز الشخص بالنسبة لمجموعته أو بالنسبة لعينة التقنين ، ولهذا لا بد من الرجوع إلى معيار يحدد معنى هذه الدرجة أو النسبة المئوية . فيدلنا على مركز الشخص بالنسبة للمجموعة وهل هو متوسط أو فوق المتوسط أو أقل من المتوسط . وما مدى بعده عن متوسط المجموعة التى ينتمى إليها ؟ وما هو وضعه بالنسبة لمن أيكافئونه من أفراد عينة التقنين ؟

ذلك أن الدرجة الخام أو النسبة المئوية للإجابات أو للأفراد تعجز عن إعطاء أى تفسير . فإذا كانت النسبة هى ٥٠ ٪ إجابات صحيحة من فقرات اختبار معين فإنها قد تساوى ٣٠ ٪ إجابات صحيحة من فقرات اختبار آخر أصعب منه ، وهى قد تساوى فى نفس الوقت ٨٠ ٪ فى اختبار آخر أسهل منه . وقد حدثت تجارب من هذا النوع أعطيت فيها للأفراد اختبارات تقيس وظيفة معينة ولكن مستوياتها مختلفة ، فكانت درجة الفرد تنخفض كلما ارتفع مستوى صعوبة الاختبار . وعلى هذا لا يمكننا أن نقول إن فلانا أعلى من فلان فى قدرة ما لأن الأول حصل على درجة فى أحد الاختبارات تعلو الدرجة التى حصل عليها الآخر فى اختبار آخر حتى لو كانا يقيسان نفس الوظيفة .

وهكذا لا يكون للدرجة الخام ولا للنسبة المئوية دلالة فى حد ذاتها بل تحتاج إلى معيار يكسبها معنى . يجب إذن أن نبحث عن نوع أفضل من الوحدات لتعبر بها عن نتائج الاختبار بطريقة أفضل من العد الخام للفقرات الصحيحة أو لوحات الدرجات والنسبة المئوية . ونحن نفضل أن تتمتع هذه الوحدات ، أى المعايير ، بالخصائص الآتية :

١ - أن يكون للدرجة الواحدة معنى موحد من اختبار إلى آخر . حتى

نوفر أساساً يمكن به أن نقارن درجات الأفراد في اختبارات مختلفة :

٢- أن تكون هذه الوحدات متساوية بحيث أن العدد المعين من الوحدات ، وليكن ١٠ درجات ، على جزء من اختبار يدل على نفس الشيء الذى تعنيه ١٠ درجات على جزء آخر من نفس الاختبار : بحيث يمكننا أن نقول إن الفرد الأول يزيد عن الفرد الثانى بـ ١٠ درجات مثلاً ، وهذه الزيادة تساوى زيادة الشخص الرابع عن الشخص الثالث .

٣- أن تكون هناك نقطة صفر حقيقية ، صفر مطلقة ، تعبر عن لا شيء من الصفة التى نقيسها ، بحيث نستطيع أن نقول إن درجة معينة ضعف درجة أخرى وأن ثلاثة تعبر ثلثى درجة رابعة .

الصفر المطلق :

والأنواع المختلفة من المعايير التى طورتها الاختبارات تمثل تقدماً ملحوظاً نحو الهدفين الأولين . ولكن النقطة الثالثة يحتمل أن لا نصل إليها أبداً فى السمات التى يخصصها المقياس النفسى والتربوى . فلسنا هنا بصدد مقاييس من نوع مقاييس الأوزان بحيث نضع سنجة هـ أرطال فى كفة من كفتى الميزان ونضع أمامها هـ أرطال من مادة معينة تزنهما تماماً . وهنا حينئذ نقول « لا يزن شيئاً » نعى فعلاً « أنه لا يزن شيئاً » ويمكننا أن نضيف وحدات الوزن معاً ، وليس لدينا هذا النوع من نقطة الصفر أو هذه الطريقة من الجمع معاً فى حالة القياس النفسى والتعليمى : فإذا جمعنا شخصين ضعيفي العقل معاً فلن يعادلا عبقرياً كما أن زوجاً من الضعاف فى الهجاية لا يفوق مجيد الهجاية .

المعايير والتقنين

قلنا إن عمليات التقنين هى الخطوات التجريبية التى يمر بها المقياس فى صورته النهائية عند إجرائه على عينات من الجماعات التى سيطبق عليها فيما بعد : ذلك بغرض اشتقاق المعايير . وعلى هذا فنحن نشق المعايير من عينة التقنين :

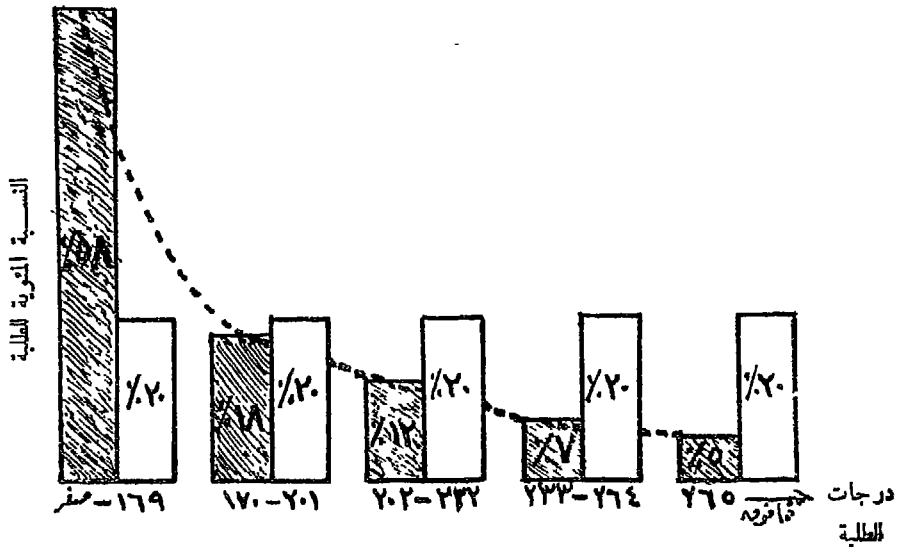
Standardization Sample التي تمثل المجتمع الأصلي المدروس Population ، فتكون الدرجات المحولة للدرجات الخام لأفراد هذه العينة مصدراً للمعايير . أى أننا عند إجراء الاختبار نقارن المفحوص بمن يكافئه من أفراد عينة التقنين .

فإذا تساوت درجة المفحوص مع درجة فرد في عينة التقنين ، أو مع متوسط إحدى فئات هذه العينة ، كان المركز النسبي للمفحوص في الجماعة التي ينتمي إليها معادلاً لمركز فرد معين في هذه العينة ، أو لمركز فئة معينة في هذه العينة . وهكذا فإننا نقارن درجة الفرد الخام بتوزيع الدرجات الخام لأفراد عينة التقنين ، فتتحول درجته إلى معيار من نفس نوع معيار الاختبار . وبذا تسهل مقارنة الفرد بالمستوى المتوقع له بما أنه عضو في جماعة معينة ، أى أننا نقارنه بمتوسط هذه الجماعة أو بمتوسطها وانحرافها المعياري كما سيتضح لنا فيما بعد .

وعلى ذلك فالمعايير ليست إلا نتائج إجراء الاختبار على عينات التقنين . فإذا كانت هذه العينات ممثلة للمجتمع الذي نود قياسه بالاختبار ، صلت هذه المعايير ما دام المجتمع المراد دراسته قد مثل تمثيلاً صادقاً عند اختيار عينة منه . وإلا أصبحت المعايير غير ذات قيمة أو ذات قيمة محدودة . ومن هنا وجب على كل مستخدم للاختبار أن يجربيه فقط على أفراد أو مجموعات مثلت في عينات التقنين . وعند نقل الاختبار من بيئة إلى أخرى يجب تقنيته على عينات تعد موازية لعينات التقنين الأصلية لاشتقاق المعايير الجديدة ، حتى يمكن قياس جماعات مقصودة تعد مكافئة للمجتمع الأصلي الذي كان يراد قياسه بالاختبار الأصلي قبل نقله . انظر الشكل (٨) .

وما دامت عينة التقنين عماد المعايير ، لذا يحسن أن نوضح الخصائص التي يجب أن تتوفر فيها حتى يصبح قياس وتقييم من نود قياسه وتقييمه : ويمكن تلخيص أهم الخصائص التي يجب أن تتوفر في عينة التقنين الحيدة فيما يلي :

١ - يجب أن تمثل عينة التقنين المجتمع الأصلي المراد دراسته تمثيلاً صادقاً من حيث الحجم والتركيب والنسب ومستوى القدرة أو الصفة المطلوب قياسها . ولا يتوفر هذا إلا بقيام الباحث ، قبل اختيار عينة التقنين ، بدراسة هذا المجتمع الأصلي دراسة مسحية Survey study كي يحدد ما سبق ذكره وكي يعرف بالتحديد حجم العينة التي يجب أن يختارها وخصائصها ، وحجم العينة التي سيختارها بالفعل وخصائصها . ذلك أنه دائماً ما تحول صعوبات عملية دون اختيار أمثل عينة . ونتيجة لهذا يتعين على واضع الاختبار أن يذكر صراحة نقائص عينته وحدود معاييرها .



(شكل ٨)

يوضح الشكل أهمية اشتقاق معايير محلية . والمستطيلات المظلة تمثل توزيع درجات ٢٠٣ طالب في جامعة بحسب المعايير القومية . ويلاحظ أنها تختلف دائماً عن المستطيلات غير المظلة التي تمثل توزيع ٦٤٦ طالبا بجامعة شيكاغو بحسب المعايير المحلية .

٢ - كلما كبر حجم العينة ، أى كلما زاد عدد أفرادها ، زاد الاعتماد على نتائج إجراء الاختبار عليها . وقد سبق أن بينا أن صدق الاختبار وثباته يتوقفان على حجم العينة وتمثيلها ؛ ولكن يجب أن ننبه إلى أنه ليست هناك أحجاماً معينة يجب أن تكون عليها عينة التقنين الجيدة ، فإن هذا يتوقف على عدة عوامل أهمها :

(أ) حجم المجتمع الأصلي المقصود دراسته .

(ب) طبيعة الاختبار ونوع الوظيفة المقاسة والغرض المنشود من عملية القياس ، فاختبار الفصل يقن على الفصل واختبار المدرسة يقن على المدرسة . بينما يقن اختبار الذكاء على أكبر عينة ممكنة ويقن الاختبار الإكلينيكي عادة على عينة أصغر ؛

(ح) إمكانيات الباحث ومدى تعاون أفراد العينة أو المسؤولين عنها مع الأشخاص الذى يجريه أو الهيئة المشرفة عليه .

٣ - يجب أن يتساوى متوسط وتشتت أفراد العينة مع متوسط وتشتت أفراد المجتمع الأصلي .

٤ - كلما صغر حجم القطاع المراد قياسه تمكن الباحث من تمثيله فى العينة وصدقت معاييرها .

ويمكن الرجوع إلى الفصل الثانى لمعرفة طرق اختيار العينات .

الدرجات المحولة

ولما كانت الدرجة الخام ومثلها النسبة المئوية ، لا تعنى فى حد ذاتها شيئاً كان لابد من تحويلها حتى تصبح ذات دلالة للحصول عليها .

وتلك الدرجات المحولة Transformed score هى التى نقصدها عند

الكلام عن المعايير . ونخدم المعايير غرضين :

١ - فهى تحدد مركز الفرد بالنسبة لعينة التقنين .

٢ - تمكنا من مقارنة مركز الفرد على مقياس بمركزه على غيره - فإذا

حصل فرد في مقياس مقدرات على ٤٠ درجة وفي مقياس قدرة ميكانيكية على ٢٢ درجة ، ففي أيهما يكون أقوى ؟ ولا يتيسر ذلك إلا بالرجوع إلى معايير كل مقياس لمعرفة مركز الفرد في كليهما ثم مقارنة مركزه على هذا بمركزه على ذلك .

ولا يمكن أن نعطي معنى لدرجة خام إلا بالرجوع إلى نوع معين من الجماعة أو الجماعات . فالدرجة لا تكون عالية أو منخفضة ولا جيدة أو رديئة ولكنها تكون أعلى من . . . أو أقل من . . . ، أو أحسن من . . . أو أسوأ من . . . وهناك طريقتان عامتان تربط بهما درجة الفرد بإطار عام . والطريقة الأولى هي أن نقارنه بسلسلة متدرجة من الجماعات . وأن نبحث عن الجماعات التي يتكافأ الفرد مع أفرادها بما أنه قد حصل على نفس متوسط درجاتهم . وكل جماعة في هذه السلسلة عادة تمثل فرقة دراسية معينة أو عمراً زمنياً معيناً . والطريقة الأخرى هي أن نحدد موضعه في جماعة معينة بالإشارة إلى النسبة المئوية من الجماعة التي يمتاز بها أو بالرجوع إلى متوسط الجماعة وانحرافها المعياري :

أنواع المعايير

وعلى هذا نجد ٤ أنماط أساسية في تفسير درجة الفرد :

نوع المعيار	نوع المقارنة	نوع الجماعة
١ - معايير السن	مقارنة الفرد بالجماعة التي يكافئها	جماعات متتابعة في العمر
٢ - الفرقة الدراسية	مقارنة الفرد بالجماعة التي يكافئها	جماعات متتابعة في الفرقة الدراسية
٣ - معايير المئين	النسبة المئوية من الجماعة التي يمتاز عنها الفرد	جماعات من السن أو الفرقة الدراسية التي ينتمي إليها الفرد
٤ - معايير الدرجة المعيارية	عددا الانحرافات المعيارية التي يزيد أو يقل بها الفرد عن متوسط الجماعة	جماعات من نفس السن أو الفرقة الدراسية التي ينتمي إليها الفرد

جدول (٥) المقارنة بين المعايير الشائعة .

معايير العمر :

أى سمة تنمو مع زيادة السن يمكننا أن نعد لها معايير عمر . ومعايير العمر هنا هو القيمة المتوسطة لهذه الصفة لأشخاص من نفس العمر . ولكى نوضح هذا نضرب المثال التالى :

لواخترنا عينة ممثلة للأطفال من عمر ٥ سنوات . ووزنا كل واحد منهم ثم حسبنا متوسط وزن كل أفراد هذه العينة . فإن هذا المتوسط يكون معيار الوزن لهذا السن . ونلاحظ هنا أن هذا المعيار ليس إلا مجرد قيمة متوسطة ، وليس نموذجاً . كما وأنه لا ينتظر أن يكون وزن كل فرد من هذا السن هو هذا الوزن المتوسط .

وبنفس الطريقة يمكننا تقدير الوزن المتوسط للأطفال من عمر ٣ سنوات ، ٤ سنوات ، ٦ سنوات . وبعد هذا يمكننا تمثيل كل متوسط بنقطة على رسم بيانى ونوصل النقاط فيصبح لدينا منحنى .

والنقط على هذا المنحنى محسوبة من أوزان السنوات الكاملة أى أن العينة فى كل سن كانت كلها تصل إلى حد هذا السن مثلاً كلهم سن ٣ سنوات تماماً ، ٤ ، ٥ ، ٦ ، . . . الخ . ولكن يمكننا اعتبار هذا المنحنى مستمراً Continuous أى يمكننا تقدير أوزان الأطفال الذين تقع أعمارهم بين حدى عمرين . مثلاً طفل وزنه ١١ كجم ووجدنا أن متوسط وزن طفل الثالثة هو ١٠ كجم ومتوسط وزن طفل رابع هو ١٢ كجم لتبين لنا أن طفلنا يعادل فى وزنه الطفل المتوسط من عينة التقنين الذى سنه ٣ سنوات و ٦ شهور .

ومن الناحية الأخرى ، إذا علمنا عمر الطفل بالتحديد وليكن ٥ سنوات ووزناه فوجدنا أنه يزن ١٢ كجم لعلمنا أن وزنه أقل من متوسط وزن عمره . وإذا كان طفل ثالث سنه ٣ سنوات ووزنه ١٢ كجم كان ، بالتالى ، بديناً .

هذه الطريقة في تقدير أوزان الأفراد أو أى سمة تنمو بتقدم العمر ، هى طريقة سهلة وعملية . ولكن لما عدة عيوب :

١ - وحدات معايير الأعمار غير متساوية . فإن الزيادة في الوزن بين ٣ سنوات و ٤ سنوات لا تعادل الزيادة في الوزن بين سن ١١ ، ١٢ سنة . وبالمثل بالنسبة لكل عمر على مقياسنا هذا . فإذا نظرنا إلى سن ١٥ سنة ، مثلاً ، وجدنا أن الزيادة في الوزن لا تتناسب مع الزيادة في العمر تماماً . وبعض الصفات الأخرى غير الوزن كثيراً ما نجد أن زيادة الصفة تتوقف عند سن معينة . فلا يزيد الطول بعد سن العشرين مثلاً وكلما اقتربنا من هذا السن قلت سرعة النمو . إن هذه الوحدات ، إذن ، ليست ذات معنى موحد .

والجدول (٦) جدول لتوزيع الأفراد حسب أعمارهم (العمود الرأسى) الذين حصلوا على درجات معينة في اختبار للمفردات (الصف الأفقى) . وعلى هذا فكل خانة ثقيلة الخطوط تقابل درجة هى معيار الأفراد من سن هو رقم العمود الذى يشمل هذه الخانة . ويكون معيار سن ٧ هو الدرجة ٢٨ وهكذا .

هذا النوع من المعايير لا يفيد في تقدير سمات الأفراد المتطرفين في التفوق أو التخلف .

كذلك فإن منحنيات النمو للسمات المختلفة لا يمكن مقارنتها مع بعضها فإن سرعة النمو تختلف من صفة لأخرى لدى الفرد كما تختلف بين الأفراد في الصفة الواحدة .

فهناك وظائف تستمر في النمو حتى سن متقدمة وأخرى يتوقف نموها في سن مبكر نسبياً .

(جدول ٦)

معايير العمر في اختبار للمفردات

العمر بالسنة	٦	٧	٨	٩	١٠	١١	١٢
٢٠	X						
٢١	XXX						
٢٢	X	XXXX					
٢٣	X	XX	XXXX				
٢٤	X	XX	XXXXXX				
٢٥	XX	XXX	XXXXXX				
٢٦		X	XXX	XXXX			
٢٧		X	XXXX	XXXX	XXXX		
٢٨		XX	XXXXXX	XXXXXX	XXXXXX		
٢٩		XX	XXXX	XXXX	XXXX	XXXX	
٣٠	X	XXX	XXXX	XXXX	XXXXXX	XXXX	X
٣١	XX	XXX	XXXX	XXXX	XXXXXX	XXXX	
٣٢	XXX	XXXX	XXXX	XXXXXX	XXXXXX	XX	
٣٣	XXX	XXXXXX	XXXXXX	XXXXXX	XXXXXX	XX	
٣٤	XXXX	XXXXXX	XXXXXX	XXXXXX	XXXX	X	
٣٥	XXXX	XXXXXX	XXXXXX	XXXX	XX		
٣٦	XXXXXX	XXXXXX	XXXXXX	XXXX	X		
٣٧	XXXXXX	XXXXXX	XXXXXX	XX			
٣٨	XXXXXX	XXXXXX	XXXX	XX			
٣٩	XXXXXX	XXXXXX	XXXX	XX			
٤٠	XXXXXX	XXXXXX	XXXX	X			
٤١	XXXXXX	XXXX	XX				
٤٢	XXXXXX	XXXX	XX				
٤٣	XXXXXX	XXXX	X				
٤٤	XXXXXX	XXXX	X				
٤٥	XXXX	XX					
٤٦	XXXX	X					
٤٧	XXX						
٤٨	XX						
٤٩	X						
٥٠							

علامة (X) تمثل تكرار الدرجة ، وألفانة المحددة بخط ثقیل (سميك)
تقابل معيار العمر أو متوسط درجة العمر الذي تتوسطه . فمثلا المعيار أو الدرجة
المتوسطة لعمر ٦ سنوات هي ٢٥ ولسن ٧ سنوات هي ٢٨ ولسن ٨ سنوات
هي ٣٠ الخ

معدل سرعة النمو في سن معين يختلف عنه في سن آخر فالطفل من عمر ٣ سنوات عندما يبلغ وزنه ١٢ كجم مثلاً تكون درجة بدائه واضحة أكبر من بدائه طفل ١١ سنة الذي يصل وزنه إلى متوسط وزن أطفال الثانية عشرة . هذه الملاحظة هي التي أدت إلى استخراج I.Q وغيرها من النسب التي ستعرض لها فيما بعد . هذه النسب تتلافى الفروق في معدل سرعة النمو . وعلى أي حال فإن مشكلة عدم تساوي وحدات العمر ستظل قائمة . والسمات التي لا يكون نموها مضطرباً مع تقدم العمر مثل حدة البصر لا يمكن تمثيلها بوحدات العمر . وكذلك المهارات التعليمية مثل التي تنمو مع تقدم الفرد في دراساته فيكون من الأنسب أن نقارنها بمعايير الفرقه الدراسية .

وقد استخدمت معايير العمر على نطاق واسع في كثير من اختبارات الذكاء العام والقدرات العقلية . وهنا كان يحدد عم عقل **Mental Age** وعمر زمني **Chronological Age** .

العمر العقلي

قلنا إن أي صفة تنمو بتقدم العمر يمكن اشتقاق معايير عمر لها . والنمو العقلي مضطرب ، إلى حد ما ، مع العمر الزمني . وقد كان هذا المفهوم دافعاً في مراجعة مقياس بينيه - سيمون سنة ١٩٠٨ ، إلى وضع فقرات يتناسب مستوى صعوبتها مع مستوى قدرة كل سن . وعلى هذا يطبق الاختبار ، الذي يستخدم معايير العمر العقلي ، على عينات من مختلف الأعمار تمثل المجتمع الأصلي بالأعمار المقصودة . وبحسب متوسط درجات أو مستوى إجابات كل سن على حدة ، ويعتبر هذا المتوسط هو معيار العمر العقلي لهذا السن .

فإذا حصل فرد على درجة معينة ، فإننا نبحث عن العمر الزمني الذي كان متوسط درجات أفراد العينة فيه مساوياً لهذه الدرجة . فإذا كانت

درجة الفرد الخام هي ٥٠ مثلاً فإننا نبحث في معايير الأعمار عن الفئة العمرية التي كان متوسط درجات أفرادها هو ٥٠ ، وليكن عمرها الزمني هو ١٥ ، فيصبح العمر العقلي لهذا الفرد هو ١٥ . أى أن العمر العقلي للفرد هو العمر الذي تساوى فيه درجة هذا الفرد مع متوسط درجات أفراد هذا العمر من عينة تقنين اختبار الذكاء المعين .

هذا المعيار يقوم ، كما سبق أن اتضح لنا ، على أساس أنه كلما زاد عمر الفرد نما عقلياً . أى أن الأفراد من سن معين يستطيعون الإجابة عن فقرات لا يستطيعها أفراد من سن أقل . ولكن هذا الافتراض غير صحيح ، إذ أن الفرد الواحد قد يستطيع إجابة أسئلة صعبة وضعت في سن ١٤ مثلاً ولا يستطيع في نفس الوقت أن يجيب عن فقرات أسهل وضعت في سن ١٣ . هذا التشتت في إجابات المفحوصين هو الذي دعى إلى فكرة العمر القاعدى ، وهو أعلى عمر يستطيع فرد معين أن يجيب إجابة صحيحة عن كل فقراته . وهو الذي دعى أيضاً إلى فكرة العمر الأقصى ، وهو العمر الذي لا يستطيع فيه المفحوص أن يجيب إجابة صحيحة عن أى فقرة .

وبين العمر القاعدى والعمر الأقصى يجب فيها المفحوص عن فقرات فيحصل على درجات تحول إلى شهور . وتجمع هذه الشهور إلى العمر القاعدى وتكون النتيجة هي العمر العقلي للفرد . وبقسمة هذا العمر العقلي على العمر الزمني لنفس المفحوص يخرج لنا كسر إذا ضرب في مائة حصلنا على نسبة ذكاء الفرد .

$$\text{أى أن نسبة الذكاء} = \frac{\text{العمر العقلي}}{\text{العمر الزمني}} \times 100$$

$$= \frac{M. A. \text{ ع. ع.}}{C. A. \text{ ع. ز.}} \times 100$$

ويمكن حساب العمر العقلي ونسبة الذكاء باختبارات غير مقسمة إلى مستويات عُمرية . ذلك بأن يحصل الفرد على درجة في الاختبار كله تقارنها بمتوسط درجة من يساوونه في العمر من أفراد عينة التقنين ، فإذا تساوت الدرجتان كان الفرد متوسطاً . وإذا زادت درجة الفرد عن متوسط درجات عمره كان متفوقاً ، وإذا حدث العكس كان متخلفاً (انظر فصل قياس الذكاء لمعرفة مزايا وعيوب العمر العقلي والطرق المختلفة في حساب نسبة الذكاء) .

عدم ثبوت وحدات العمر العقلي :

لا تظل وحدات العمر العقلي ثابتة بل تتناقص بتقدم العمر فالسنة تأخر في سن ٤ تعادل حوالى ٣ سنوات تأخر في سن ١٢ وهكذا فطفل عمره الزمنى ٤ والعقلي ٣ يكون متخلفاً عاماً واحداً عن متوسط سنه وتكون نسبة ذكائه $\frac{3}{4} \times 100 = 75$. ونسبة الذكاء ثابتة تقريباً فعندما يكون عمره ١٢ يكون عمره العقلي $= \frac{ع \cdot ع}{12} = 75$. ع . ع = ٩ سنوات أى أصبح متأخراً ثلاث سنوات عن متوسط سنه . وسبب هذا أن سرعة النمو العقلي تختلف فى الطفولة تزيد كثيراً جداً عنها فى الأعمار الأكبر .

ويقف نمو الذكاء عادة عند سن ١٥ - وإذا طبقنا تلك المعادلة فستقل نسبة ذكاء الشخص كلما كبر عن سن ١٥ . لهذا يكتفى بحساب عمر زمنى ١٥ سنة لأى شخص عمره أكثر من ١٥ سنة - ونعوض فى المعادلة كالمعتاد . وميزة العمر العقلي هى أنه واضح ومفهوم حتى لدى العامة .

عيوب العمر العقلي :

١ - عدم ثبات تباين الأفراد فى الأعمار المختلفة . فيزداد التباين كلما تقدم العمر ذلك لتناقص وحدات العمر إذ أن سنة فى عمر زمنى ٤ تعادل

حوالى ٣ سنة فى عمر زمنى ١٢ وعلى هذا يزيد التباين فى الأعمار الكبيرة ويقل فى الأعمار الصغيرة . وتتلافى نسبة الذكاء هذا العيب .

٢- يفقد العمر العقلى وضوحه ومدلوله فى حالة الراشد المتفوق والراشد المتقدم فى العمر . فالراشد المتفوق ذو العمر الزمنى ١٥ والعمر العقلى ٢٠ لا يعنى أنه يساوى ذكاء عمر عقلى لشخص متوسط فى عمر زمنى ٢٠ إذ أن هذا الأخير لا يزيد عمره العقلى عن ١٥ حيث لا ينمو الذكاء بعد سن ١٥ وهكذا يفقد العمر العقلى مدلوله فى هذه الحالات .

٣- لا نستخدم درجات العمر إلا فى السمات والقدرات التى تنمو بتقدم السن حيث تزداد مثل الذكاء أما تلك السمات التى لا تتأثر بتقدم العمر فيستحيل استخدام درجات العمر فيها ومن أمثال ذلك معظم سمات الشخصية كالانبساط والانطواء . الخ .

أما فى ميدان التربية نتيجة لطبيعة الاختبارات وللنقد الموجه إلى مفهوم العمر العقلى ، فيغلب استخدام معايير الفرقة الدراسية Grade Norms وهى قريبة فى أساسها من معايير العمر .

معايير الفرقة الدراسية

لا تختلف معايير الفرقة الدراسية Grade Norms كثيراً عن معايير العمر ، فمعايير الفرقة هو متوسط درجات أفراد الفرقة فى الاختبارات المعينة . وهنا نقارن الفرد المعين بمتوسط درجات الفرقة التى ينتمى إليها . أو نقارن الفرد بفرقة أخرى اقتربت درجته من متوسطها فيعد فى هذه الفرقة بالنسبة للمادة المقاسة . أو نقارن الفرد بمن فى سنه من أفراد فرقته . وهنا نستبعد المتطرفين من الفرقة سواء كانوا صغاراً فى السن أو كباراً عن العمر المتوالى modal age لهذه الفرقة .

ولمعايير الفرق الدراسية نفس عيوب معايير العمر والعمر العقلي .
 (وللمعرفة المزيد يمكن الرجوع إلى فصل خصائص اختبارات التحصيل) .
 وعموماً نجد أن هذه المعايير أسهل في حسابها ، أسهل لأنها مبنية على مجموعات
 مكونة بالفعل في النظام المدرسي أى أن الباحث لا يقوم بتقسيمها إلى
 مستويات كثيرة . ومفهوم « مستوى الفرق الدراسية » واضح عن مفهوم
 « مستوى العمر » . ولكن المفهوم الأول خارج الفرق الدراسية يكون
 محدود القيمة .

المعايير المئينية

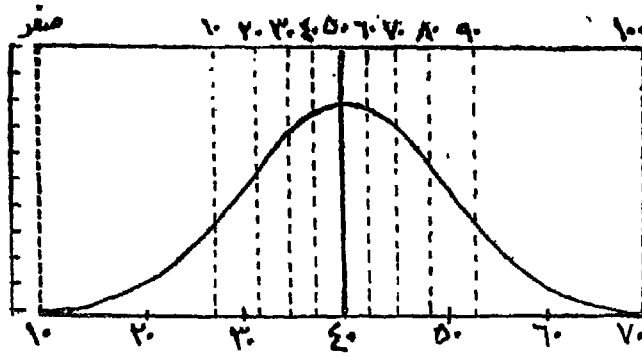
لاحظنا في كلامنا عن معايير العمر ومعايير الفرق الدراسية أننا نقارن
 الفرد بمتوسط عمر معين أو فرقة دراسية معينة . ولكن من الأجدى أن
 نقارن هذا الفرد بالآخرين من عمره أو الآخرين من فرقته ، حتى نحدد
 مركزه بالنسبة لهذه الجماعة . كأن نعرف مثلاً أنه يفوق ٧٥ ٪ مثلاً من
 هذه الجماعة في مادة معينة أو في اختبار يشمل عدداً من المواد .

وتتميز المعايير المئينية بأنها أكثر مرونة وأوسع مدى في تطبيقها عملياً .
 إذا يمكن تطبيقها على أى جماعة لتحديد موضع كل فرد بالنسبة للآخرين .
 وهى تطبق على أوسع نطاق في مجالات الصناعة . وعند تطبيق المعايير المئينية:
 لاختبار منشور يجب على الفاحص أن يكملها باستخراج معايير خاصة ،
 بالمجموعة التى يطبقه عليها .

كأن تستخرج معايير محلية خاصة بدرجة معينة من واقع درجات تلاميذها
 مما يوفر فرصة تفسير درجة التلميذ الفرد على ضوء درجات مجموعته المحلية .
 وهذا التفسير بالطبع أجدى من مقارنة التلميذ بالمعايير المئينية المشتقة من
 تلاميذ الدولة كلها .

والمشكلة الأولى هى مشكلة اختيار عينة التقنين التى يجب أن تمثل فرقاً وأعماراً

مختلفة من المجموع السكاني الذي تقيسه . وهنا نحتاج إلى عمل معايير مثيئية لكل فرقة دراسية على حدة أو لكل سن على حدة .



(شكل ٩)

يوضح الشكل تركيز توزيع الدرجات في الوسط وقلة التوزيع تدريجياً في الطرفين . لاحظ أن الخط القاعدي يمثل درجات خام وحداتها متساوية وأن الخط الموازي له يمثل مثنيات وحداتها غير متساوية

والمشكلة الثانية بالنسبة للمعايير المثيئية هي عدم تساوى الوحدات (التوزيع في الأطراف والتوزيع في المنطقة المتوسطة انظر الشكل (١٠) .) فالفرق بين رتبتين مثيئيتين في وسط المنحنى أصغر كثيراً من الفرق بين رتبتين مثيئيتين في طرفي المنحنى . وسيتضح هذا بعد أن نعرف الدرجة المثيئية .

الدرجة المثيئية هي الدرجة التي تقع تحتها نفس النسبة المئوية من الأفراد وهي النسبة المئوية للأشخاص الواقعين تحت درجة معينة . فمثلاً إذا حل ٢٨٪ من أفراد العينة أقل من ١٥ وحدة عن المقياس فإن الشخص الذي يحل ١٥ وحدة يقع في المئين الـ ٢٨ ويرمز لها بالإنجليزية 28 p. أى أن

المئين يحدد مباشرة ترتيب الفرد في المجموعة حيث المئين الـ ٥٠ هو المتوسط والمئين صفر يشير إلى درجة أقل من أى درجة في المجموعة والمئين ١٠٠ يحدد درجة أعلى من أى درجة في المجموعة . ويجب أن تفرق بين المئين وهو معيار وبين النسبة المئوية وهى درجة خام فى حاجة إلى أن تقارن بمعيار حتى يتضح مدلولها .

وفى حالة اختبارات ذكاء أقل من ١٥ سنة يصبح لكل مجموعة سن معينة مئينات خاصة بها . فهناك مثلاً مئينات لسن ٧ وأخرى لسن ٨ . . . الخ . وكذلك فى حالة ما إذا كان هناك خلاف كبير بين استجابات المجموعات الفرعية Subgroups داخل المجموعة الكلية مثلاً اختلاف قدرة الإناث . القدرة الميكانيكية - وضعفهم فيها عن الذكور . ففى أمثال هذه الحالات يحسن أن توجد معايير خاصة لكل من الإناث والذكور .

ميزة المئينات :

- ١ - تعطى صورة صادقة لترتيب الفرد بالنسبة للمجموعة .
- ٢ - مهلة فى حسابها ووضحه فى مدلولها .
- ٣ - أوسع فى انتشارها من درجات العمر إذ يمكن استخدامها بالنسبة للأطفال كما للراشدين إلى جانب أنها صالحة لأى نوع من المقاييس الشخصية أو الذكاء أو القدرات أو التحصيل .

عيوب المئينات :

- ١ - عدم تساوى الوحدات المئينية على منحنى التوزيع إذ تقل المسافات بين المئينات فى الوسط وتزيد كلما اتجهنا نحو الوسط . وبهذا فإن المئينات تتباعد فى تباین الأفراد فى الوسط وتقلل التباين فى الأطراف . فمثلاً نجد حوالى ١٠٪ من الحالات تقع على مسافة من الدرجات الخلف تقارب سبعة

أمثال المسافة التي تغطيها ١٠٪ من الحالات في الوسط حيث نجد المسافة بين المئين صفر ، ١٠ تساوى سبعة أمثال المسافة بين المئين ٤٠ و ٥٠ وهذا يعنى أنه ليس هناك اتفاق بين توزيع الدرجات الخام وتوزيع المئينات .

٢ - لا يعطينا المئين مدى اختلاف الدرجة الخام عن غيرها وكل ما يعطيه هو ترتيبها فقط .

الدرجات المعيارية :

الدرجة المعيارية هى المسافة التى تبعدنا الدرجة الخام عن المتوسط الحسابى معبرا عنها فى وحدات من الانحراف المعيارى Standard Deviation of the distribution والانحراف المعيارى هذا هو مقياس من مقاييس التشتت أو التباين ويعبر عن مدى الاختلافات الفردية فى درجات المقياس

استخراج الدرجة المعيارية :

وتستخرج الدرجة المعيارية بالتعويض عن المعادلة الآتية :

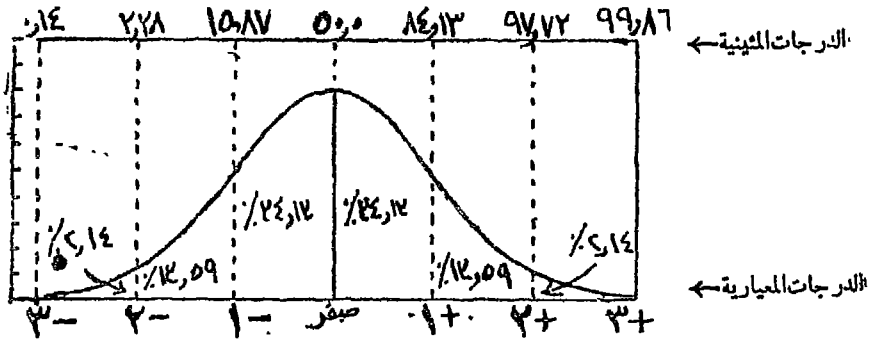
$$\text{الدرجة المعيارية} = \frac{\text{الدرجة الخام} - \text{المتوسط الحسابى للدرجات}}{\text{الانحراف المعيارى للدرجات}}$$

وواضح أن الدرجة الخام الأقل من المتوسط ستكون إشارتها عند تحويلها إلى درجة معيارية سالبة . أما الدرجة الخام الأعلى من المتوسط ستكون موجبة الإشارة والدرجة الخام التى تعادل المتوسط ستكون مساوية لصفر درجة معيارية .

ينبغي أن تحسب الدرجة المعيارية إلى كسر عشرى من رقم واحد على الأقل لكي تتمكن من التمييز بين الأفراد حيث غالباً ما تقع درجات التوزيع بين ٣+ درجة معيارية ، - ٣ درجة معيارية حسب التوزيع لاعتدالى الذى يغلب فى الاختبارات النفسية . (انظر شكل ١٠)

وهكذا يتضح لنا أن الدرجات المعيارية تحسب على أساس قيمتي المتوسط والانحراف المعياري بينما وجدنا من قبل أن معايير العمر والفرقة الدراسية والمئين لا تلتزم بانحراف الفرد عن متوسط جماعته إنحرافاً يعبر عنه عددياً بالانحراف المعياري . أما الدرجة المعيارية فهي تحسب نسبة انحراف الفرد عن المتوسط (س - م) إلى انحراف أفراد الجماعة كلهم عن المتوسط (ع) . وهذه النسبة تكون موجبة الإشارة في حالة من يعلو المتوسط وسالبة الإشارة في حالة من يقل عنه . وطبيعي أن يتساوى عدد من يزيدون عن المتوسط مع عدد من يقلون عنه ، وطبيعي أيضاً أن تتساوى الانحرافات السالبة مع الانحرافات الموجبة . وعلى هذا يكون مجموع الدرجات المعيارية لأي توزيع هو صفراً ($3+ - 3- = 0$ صفر) . وأهم خاصية في الدرجات المعيارية هي أن وحدة حسابها هي الانحراف المعياري . وإذا حسبنا الانحراف المعياري للدرجات المعيارية لأفراد أي مجموعة وجدناه مساوياً واحداً صحيحاً .

ومن شأن كل هذه الخصائص أن تسهل الدرجات المعيارية عمليات مقارنة المجموعات المختلفة حتى لو اختلفت متوسطاتها وانحرافات المعيارية ما دامت الواحدات ، في الدرجات معيارية وهي الانحرافات المعيارية ، متساوية . فبهذا يمكن مقارنة درجات الفرد الواحد في اختبارين مختلفين أو أفراد مختلفين في اختبار واحد .



(شكل ١٠) لاحظ أن المئينات المقابلة للدرجة المعيارية من صفر إلى + أضعاف المئينات المقابلة للدرجة المعيارية من ٢ + إلى ٣ + مثلاً

وفيما يلي توزيع مئوى للحالات فى المنحنى الاعتدالى Normal Curve بين الدرجات المعيارية المختلفة محسوب من الشكل (١٠) .

بين - ٣ ، - ٢	يقع ٢١٤	% من الحالات
بين - ٢ ، - ١	يقع ١٣٥٩	% من الحالات
بين - ١ ، صفر	يقع ٣٤١٣	% من الحالات
بين صفر ، + ١	يقع ٣٤١٣	% من الحالات
بين + ١ ، + ٢	يقع ١٣٥٩	% من الحالات
بين + ٢ ، + ٣	يقع ٢١٤	% من الحالات

وهذا يكون تقابل المئينات والدرجات المعيارية على الوجه الآتى من التوزيع الاعتدالى :

- ٣ درجة معيارية يقابلها المئين ١٤
- ٢ ، درجة معيارية يقابلها المئين ٢٢٨
- ١ درجة معيارية يقابلها المئين ٨٧
، صفر درجة معيارية يقابلها المئين ٥٠
+ ١ درجة معيارية يقابلها المئين ٨٤
، + ٢ درجة معيارية يقابلها المئين ٩٧
+ ٣ درجة معيارية يقابلها المئين ٩٩

عيوب الدرجات المعيارية .

١ - يقوم حساب الدرجات المعيارية على أساس حساب توزيع درجات أفراد العينة ، بإيجاد قيمتى المتوسط والانحراف المعيارى . وهذا التوزيع يختلف باختلاف العينة وظروف الإجراء ومواد الاختبار . ولهذا يحسن أن تشتق الدرجات المعيارية بعد تعديل التوزيع التجريبى للدرجات إلى توزيع

اعتدالى حتى يكون أساس الحساب ثابتاً . وتسمى هذه العملية « تسوية المنحنى Normalization » .

٢ - قد تكون إشارة الدرجة المعيارية سالبة ، ولهذا تهدف تعديلات الدرجة المعيارية إلى تسوية المنحنى وإلى التخلص من الإشارات السالبة . يجعل كل الدرجات المعيارية موجبة .

٣ - تكون الدرجة المعيارية غالباً كسراً عشرياً أو درجة وكسراً عشرياً . وعلى هذا يعبر عن الفروق بين الأفراد بوحدات صغيرة جداً لا تمثل مدى هذه الفروق . ولهذا أيضاً هدفت تعديلات الدرجة المعيارية إلى توضيح هذه الفروق والتخلص من الكسور .

نسبة الذكاء الإنحرافية I Q : Deviation

وهي في الحقيقة درجات معيارية معدلة ، إلا أن متوسطها الحسابي ١٠٠ وانحرافها المعياري ١٦ فإن ١ + درجة معيارية تقابل ١١٦ وهي بهذا تتخذ نفس متوسط وانحراف مقياس ستانفورد بينيه وذلك ، لشيوع استعمالها ودلالة نسب ذكائه ووضوحها . وبهذا فالشخص الذي تكون درجته المعيارية + ٢ فنسبة ذكائه الانحرافية = $(١٦ \times ٢ +)$ $١٣٢ = ١٠٠ +$.

وهذه النسب كما هو واضح لا تأتي من قسمة العمر العقلي على الزمني . مع ضرب النتائج في ١٠٠ كما هو معهود في نسب الذكاء المستخدمة في اختبار ستانفورد بينيه . ولكنها تأتي من استخدام الدرجات المعيارية بتحويلها إلى توزيع متوسط ١٠٠ وانحرافه ١٦ حتى تسهل مقارنة نسب الذكاء على اختبارات مختلفة .

أنواع أخرى من الدرجات المعيارية المعدلة

١ - الدرجة T :

الدرجة T : T. Score درجة معيارية معدلة تهدف إلى التخلص من

عيوب الدرجات المعيارية . والاسمات تؤخذ عن الحرف الأول من اسم
ثورنديك Thorndike . والفكرة الأساسية فيها هي تحويل الدرجات
المعيارية في أى اختبار إلى درجات معيارية اعتدالية متوسطها الحسابى ٥٠
والانحراف المعيارى لها ١٠ . وهكذا تتراوح درجات أفراد أى مجموعة بين
٥ - ، ٥ + درجة تائية ؛

فإذا فرضنا أن اختبارا طبق على مجموعة من الأفراد فحصل الفرد س
على الدرجة الخام ١٨ . وكان المتوسط الحسابى للمجموعة هو ١٥ والانحراف
المعيارى لها هو ٦ . فتكون درجته المعيارية هي :

$$٠.٥ = \frac{١٨ - ١٥}{٦}$$

وتكون درجته التائية هي :

$$٥٥ = ٥٠ + ١٠ \times ٠.٥$$

وإذا كانت الدرجة الخام للفرد ص في نفس المجموعة هي ٩ تكون
درجته المعيارية هي :

$$١ - = \frac{١٥ - ٩}{٦}$$

وتكون درجته التائية هي :

$$٤٠ = ٥٠ + ١٠ \times ١ -$$

وهكذا يرى أن الدرجة التائية هي درجة معيارية اعتدالية انحرافها
المعيارى ١٠ ومتوسطها الحسابى ٥٠ . وهي تحسب من الدرجات المعيارية
وليس من الدرجات الخام . ولكن يعيها أن الفروق بين درجات الأفراد

تمثلها وحدات صغيرة قد لا تفيد في بعض البحوث العملية التي تكون فيها الفروق بين الأفراد ، بطبيعة الحال صغيرة .

٢ - ولهذا نشأت درجات ثائية معدلة فمثلا في إحدى اختبارات الجيش الأمريكي ، واسمه اختبار التصنيف العام للجيش (A.O.C.T) Army General Classification test ، كان الانحراف المعياري ٢٠ والمتوسط الحسابي ١٠٠ فتكون الدرجة المعيارية لمن يحصل على ١٢ في اختبار متوسطه ١٠ وانحرافه المعياري ٤ هي :

$$٠.٥ = \frac{١٢ - ١٠}{٤}$$

وتكون درجته المعيارية ، بحسب هذا المعيار الحربي ، هي :

$$١٢٠ = ١٠٠ + ٢٠ \times ٠.٥$$

٣ - كما لجأت بعض الجامعات الأمريكية إلى تعديل الدرجة الثائية باستخدام متوسط حسابي ٥٠٠ وانحراف معياري ١٠٠. حتى تتضح الفروق أكثر . فإذا كانت الدرجة المعيارية لفرد ما هي - ٠.٣ فإن درجته المعيارية بحسب هذا المعيار الجديد هي :

$$٤٧٠ = ٥٠٠ + ١٠٠ \times ٠.٣ -$$

٤ - الدرجة ج :

الدرجة ج C Score هي تعديل للدرجات المعيارية بحيث تصبح درجات معيارية اعتدالية . اشتقها Guilford من توزيع متوسطه ٥ وانحرافه المعياري ٢ . وكان هدفه هو إيجاد درجات معيارية اعتدالية تناسب بعض الاختبارات التي تقيس وظائف تكون الفروق بين الأفراد فيها واضحة . كما أن مزايا هذا المعيار قلة عدد المستويات التي ينقسم إليها الأفراد .
نوطريقة حسابها هي :

إذا كانت الدرجة المعيارية للفرد هي ٠.٤

تصبح درجته بالمعيار ج هي :

$$٠.٤ \times ٢ + ٥ = ٥.٨$$

٤ - معيار التسيع :

معيار التسيع Standard Nine or Stanine يقوم على أساس الدرجة ج .
فبدلاً من تقسيم التوزيع إلى ١١ قسمًا أصبح سلاح الطيران الأمريكي يقسمه
إلى ٩ أقسام . وذلك بضم الدرجة ج (صفراً) والدرجة ج (١) معاً في
التسيع (١) وضم الدرجة ج (٩) والدرجة ج (١٠) معاً في التسيع (٩) .
وهكذا يوضح هذا المعيار الفروق بين الدرجات إذا لم يكن مداها
كبيراً ، خاصة إذا كان الأفراد المتطرفون سلبياً متقاربين من بعضهم وكان
الأفراد المتطرفون إيجابياً قريبين من بعضهم أيضاً . وطبيعي ألا يصلح
هذا المعيار إذا كان الأفراد موزعين في مستويات تشمل أقساماً دنيا
وأخرى عليا .

تحويل المعايير

يجب أن يقدم كل صاحب اختبار لمن يستخدمونه قائمة معايير Table
of Norms اختباره موضحاً في دراسته عدد ونوع ومستوى وطريقة اختيار
وتركيب العينة المستخدمة والعمليات الإحصائية وشكل التوزيع . ذلك حتى
يتمكن الأخصائيون في شؤون القياس من مقارنة الدرجات الخام أو المحولة
التي يحصل عليها الأفراد في اختبارات مختلفة في معاييرها . وهكذا يمكن
تحويل درجة الفرد سواء كانت خاماً أو معياراً إلى معيار آخر .

بل ويحسن أن توضح قائمة المعايير « المقابلات المعيارية المختلفة لكل
درجة خام » حتى أنه بمجرد النظر إلى القائمة نعرف ما تساويه الدرجة الخام
من معيار الفرقة الدراسية أو العمر أو الدرجة المعيارية أو التائية أو . الخ .

والشكل (١١) يوضح المنحنى الاعتدالى مبيناً عليه أنواع مختلفة من المعايير المتقابلة .

النسب

١ - أول هذه النسب Quotients هي نسبة الذكاء I. Q فإذا كان العمر العقلى للفرد هو ٩٠ شهراً وكان عمره الزمنى ٧٥ شهراً كانت :

$$\text{نسبة ذكاؤه} = 100 \times \frac{90}{75} = 120$$

وإذا كان العمر العقلى للفرد ١١ سنة و ٣ شهور والعمر الزمنى له ١٢ سنة و ٦ شهور كانت :

$$\text{نسبة ذكاؤه} = 100 \times \frac{11-3}{12-6} = 100 \times \frac{8}{6} = 133$$

$$\text{وبتحويل السنوات إلى شهور} = 100 \times \frac{135}{100} = 135$$

٢ - أما نسبة التعليم Educational Quotient (E. Q) فهي النسبة بين تحصيل الفرد وعمره الزمنى . ونحن نعلم أن التحصيل يقاس بحساب الفرقة الدراسية التى تتكافأ درجة الفرد مع متوسط درجات تلاميذها فى مادة الاختبار . فإذا فرضنا أن اختباراً للحساب قن على عدة فرق دراسية وأن التلميذ س حصل على الدرجة ٨ وأن التلميذ ص حصل على الدرجة ١١ وأن التلميذ ع حصل على الدرجة ١٥ . وكان متوسط درجات السنة الأولى الإعدادية هو ٦ والثانية هو ١٠ والثالثة هو ١٢ والأولى الثانوية هو ١٤ والثانية الثانوية هو ١٥ . فإذا كانت أعمار التلاميذ الثلاثة على الترتيب هى ١٠ سنوات ، ١١ سنة ، ١٢ سنة و ٦ شهور .

فإننا نستطيع أن نعد السنة الأولى الإعدادية هى السنة السابعة والثانية الإعدادية هى الثامنة والثالثة الإعدادية هى التاسعة والأولى الثانوية هى العاشرة

(ما دامت المرحلة الابتدائية تنتهى بالسنة السادسة) وهكذا يكون حساب العمر التحصيلي للتلميذ س بالطريقة التالية :

١٠ - ٦ = ٤ مدى الفئة أو الفرقة السابعة أى الأولى الإعدادية . قطع هو من هذا المدى ٨ - ٦ = ٢ أى أن فرقته هى السابعة مضافاً إليها نسبة ما قطعه من هذا المدى أى $\frac{2}{7}$ فتكون فرقته ٧ر٥ . وبفرض أن متوسط عمر هذا المستوى الدراسي هو ١٢ر٥ .

وتكون فرقة التلميذ ص = $8 + \frac{10 - 11}{10 - 12}$ حيث ٨ هى الفرقة الثانية ، ١٠ هى درجته ، (١٢ - ١٠) هى مدى الفرقة الثانية ، (١١ - ١٠) هو المدى الذى قطعه من هذه الفرقة . ونسبة ما قطعه من تحصيل إلى ما يجب أن يقطعه إلى آخر العام ، حتى ينقل إلى الفرقة الثالثة ، هى $\frac{1}{11}$ أى أن فرقته ٥ر٨ . ومتوسط عمر هذا المستوى الدراسي هو ١٣ر٥ .

وبالمثل يمكن حساب فرقة التلميذ ع = $10 + \frac{15 - 15}{14 - 15}$ = الفرقة ١٠ ومتوسط عمر هذا المستوى الدراسي هو ١٥ . فتكون أعمار الثلاثة التحصيلية على الترتيب هى ١٢ر٥ ، ١٣ر٥ ، ١٥ر٥ .

$$\frac{\text{العمر التحصيلي A.A}}{\text{العمر الزمني C.A}} = \text{النسبة التعليمية (E.Q.)}$$

$$\text{النسبة التعليمية للتلميذ س} = \frac{12.5}{10} \times 100 = 125$$

$$\text{النسبة التعليمية للتلميذ ص} = \frac{13.5}{11} \times 100 = 123$$

$$\text{النسبة التعليمية للتلميذ ع} = \frac{15}{12.5} \times 100 = 120$$

٣ - النسبة التحصيلية (A.Q.) Achievement Quotient هي النسبة بين تحصيل الفرد وعمره العقلي محسوبا عن اختبار للذكاء . وقد تسمى « نسبة الإنجاز » Accomplishment Q. (A.Q.) فإذا كانت نسبة ذكاء الفرد هي ١٢٥ ونسبته التعليمية هي ١٠٠ كان مستواه العقلي يسمح له بنجاح أكثر من الذي حصله وكانت نسبة تحصيله هي :

$$\text{نسبة التحصيل} = \frac{\text{نسبة التعليم}}{\text{نسبة الذكاء}}$$

$$٨٠ = ١٠٠ \times \frac{١٠٠}{١٢٥} =$$

وإذا كان العمر التحصيلي لفرد هو ١٢ سنة وعمره العقلي هو ١٠ سنوات كانت نسبة تحصيله هي :

$$\text{نسبة التحصيل} = \frac{\text{العمر التحصيلي}}{\text{العمر العقلي}} \times ١٠٠$$

$$١٢٠ = ١٠٠ \times \frac{١٢}{١٠} =$$

أى أنه يحصل أكثر مما يسمح له مستواه العقلي وعلى هذا نحس أنه يبذل مجهوداً أكثر من التلميذ المتوسط . وهكذا تفيدنا هذه المعايير في معرفة الكثير عن التكوين النفسى والتربوى للفرد .

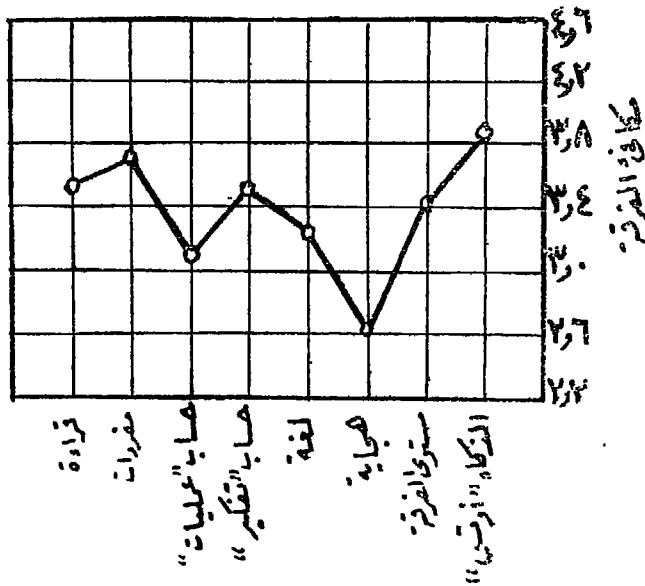
الصفحات النفسية

تمدنا المعايير السابقة بطريقة نعبّر خلالها عن درجات مجموعة من الأفراد في اختبارات مختلفة تماماً ، وهذا التعبير يكون بوحدات عامة بحيث يمكن

المقارنة مباشرة . وبالرغم من أن المعايير السابقة تسهل مقارنة درجات الفرد في اختبارات مختلفة تقيس نفس الشيء ، إلا أن أحدا من هذه المعايير لا يمكننا من مقارنة درجات الفرد في اختبارات تقيس وظائف مختلفة .

ولكن إذا حولت درجات أى اختبار يقيس أى وظيفة إلى نوع من المعايير ينسب الفرد إلى مستوى فرقة دراسية معينة في الحساب مثلا وينسبه اختبار آخر إلى فرقة أعلى في اللغة العربية وينسبه اختبار ثالث إلى فرقة أقل من الأولى ، سهلت عملية المقارنة . وهكذا يكون للفرد أكثر من فرقة دراسية حسب مستواه في المواد المختلفة ، إذ أن مستوى أحد تلاميذ الشهادة الإعدادية قد يفوق متوسط درجات تلاميذ السنة الثانية الثانوية وخاصة في المواد التي تدرس في الإعدادية ولا تدرس في الثانوية . انظر الشكل (١٢) وسيتضح هذا الكلام في قسم القياس التربوي .

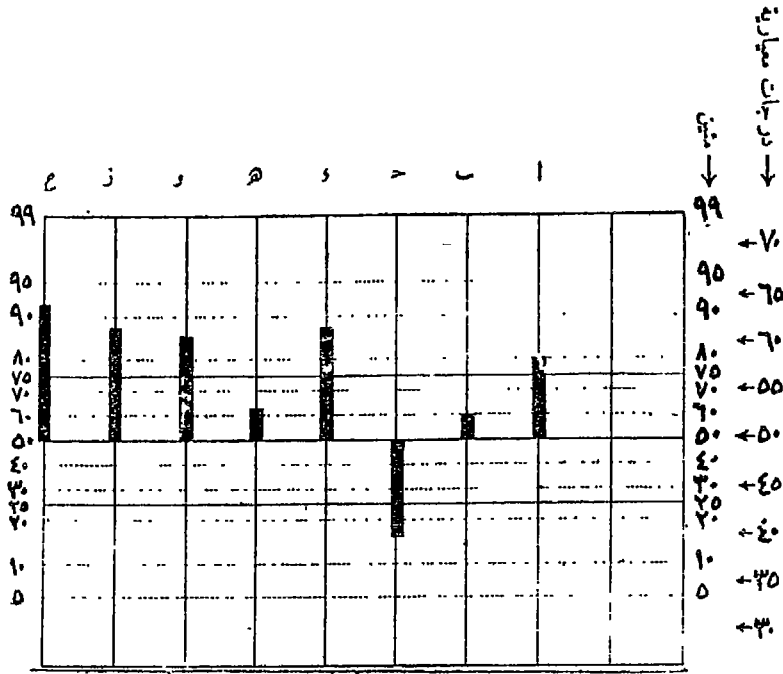
(شكل ١٢)



كما تسهل مقارنة مستويات الأفراد في وظائف مختلفة إذا عبرنا عن نتيجة كل اختبار في معيار هو النسبة المئوية لعدد من يفوقهم الفرد من أفراد الجماعة في الاستدلال اللفظي مثلاً ، والنسبة المئوية لعدد من يفوقهم الفرد في السرعة في الأعمال الكتابية مثلاً . هذا المعيار هو المئين . والمعيار السابق هو معيار الفرقة الدراسية .

ولست هذه المعايير جديدة علينا ، فقد سبق عرضها ، ولكن الجديد هو جمعها في صفحة واحدة هي الصفحة النفسية Psychopprofile وهي عرض بياني graphic presentation مجمع لدرجات الفرد في اختبارات مختلفة . وهناك صفحات تستخدم لتسجيل درجات مجموعة من الأفراد معاً وهي تسمى سجلات Records . في أحد فئاتها الأسماء وأمام كل اسم سطر ينقسم إلى عدة خلايا كل خلية يوضع بها رقم .

فمثلاً في صفحة سجل الفصل Class Record Sheet في اختبار كاليفورنيا للنضج العقلي California Test of Mental Maturity ، أمام خلية اسم التلميذ فئة لجنسه وأخرى لعمره الزمني وفئة ثالثة لعمره العقلي تنقسم إلى عمر عقلي في الاختبار ككل وعمر عقلي في القسم اللفظي من الاختبار وعمر عقلي في القسم غير اللفظي من الاختبار . ثم خلايا لنسبة الذكاء في الأقسام الثلاث السابقة . وآخر فئة من الخلايا تدرج فيها الرتبة المئينية للفرد في اختبارات التذكر ، والعلاقات المكانية ، والاستدلال المنطقي ، والاستدلال العددي ، والمفاهيم اللفظية ، وأخيراً خلية للرتبة المئينية للفرد في مجموعة الاختبارات اللفظية وأخرى لغير اللفظية .



(شكل ١٣)

صفحة نفسية لمفحوص بالفرقة التاسعة على بطارية اختبارات الاستعدادات

الفارسية عن D. A. T. Manual, The Psychological Corporation, 1947.

N.Y., 1947. والحروف تشير إلى الاختبارات وهي :

- (ا) استعمال اللغة - الجمل
- (ب) استعمال اللغة - المعجزة
- (ح) الدقة والسرعة للأعمال الكتابية
- (د) الاستدلال الميكانيكي
- (هـ) العلاقات المكانية
- (و) التفكير المجرد
- (ز) القدرة العددية
- (ع) التفكير اللفظي

وفي (الشكل ١٣) صفحة نفسية لمفحوص واحد على بطارية اختبارات الاستعدادات الفارقة Differential Aptitude Tests D.A.T. وهذه البطارية تقيس جوانب مختلفة من القدرة . وهذه الجوانب أساس برامج التوجيه في المدرسة الثانوية . ويلاحظ أن الاختبارات المختلفة تمثلها أعمدة منفصلة . والمعايير فيها هي الدرجة المعيارية محسوبة على أساس المتوسط الحسابي والانحراف المعياري لكل فرقة دراسية على حده . كما أنه ، بالنسبة لكل فرقة دراسية على حده تحسب الرتبة المئينية للفرد على أساس عدد من يفوقهم في الاختبارات المختلفة .

وواضح أن الاختبارات ، التي تتضمنها البطارية ، هي اختبار لفظي وآخر عددي وثالث للاستدلال ورابع للعلاقات المكانية وخامس للاستدلال الميكانيكي وسادس للأعمال الكتابية (وينقسم بدوره إلى قسمين) وأخيراً اختبار لغوي ينقسم إلى اختبار للهجاية وآخر للجمل . ومن الرسم يتضح لنا أيضاً أنه يمكن التعرف على مقابل الدرجة الخام سواء كان درجة معيارية داخل الفرقة الدراسة التي ينتمي إليها المفحوص ، أم كان الرتبة المئينية للمفحوص بالنسبة لنفس الفرقة . ويمكننا أن نستخرج من الشكل الجدول رقم (٧) :

الاختبار الدرجة	اللفظي	العددي	المجرد	المكاني	الكتابي	الهجاية	الحمل
الخام	٣٣	٢٦	٣٥	٣٩	٣٤	٤٣	٤٥
المئينية	٩٢	٨٨	٨٧	٦٣	٨٨	١٧	٨٠
المعيارية	٦٥	٦٣	٦٢	٥٤	٦٣	٤١	٦٩

الجدول (٧) الدرجات الخام ، للمفحوص في بطارية اختبارات الاستعدادات الفارقة ومقابلتها الدرجة المعيارية والرتبة المئينية داخل الفرقة الدراسية

والدرجة المعيارية Standard Score هنا متوسطها ٥٠ وانحرافها المعياري ١٠ . وواضح أن توزيع المئينات متركز في الوسط مشئت في الأطراف فوحداتها غير متساوية تساوى الدرجات المعيارية المقابلة . ونلاحظ أن توزيع المئينات يقترب من توزيعها في المنحنى الاعتدالى .

وهذه الصفحة النفسية تبين لنا نواحي القوة والضعف في المفحوص بالنسبة لمتوسط مستوى جماعته في كل اختبار على حدة . وهذا المتوسط كان المئين ٥٠ . وعند تفسير الصفحات النفسية يلزم أن نكون حريصين بحيث تمكن المقارنة وتكون صحيحة فتشتق معايير العمر أو الفرقة أو المئين من جماعات متكافئة ، ذلك في كل الاختبارات . ولضمان تكافؤ كل الجماعات يحسن تطبيق كل الاختبارات على جماعة واحدة .

ولكن يعيب الاعتماد على الصفحات النفسية الفردية أنها تجسم الفروق في ذات الفرد . ونحن نهتم عادة بالفروق بين الأفراد . وعلى هذا يستخدم هذا النوع من الصفحات في الخدمات النفسية الفردية : وهناك نوع آخر من الصفحات النفسية كسجل كاليفورنيا يمكن به مقارنة أفراد الجماعة معاً .

اختلاف نتائج الاختبارات

لا تعطى الاختبارات التي تقيس نفس الوظيفة نفس النتائج إذا طبقت عينة واحدة لهذه الأسباب :

١ - اختلاف عينة تقنين كل اختبار عن غيرها والتي عل أساسها خرجت معايير الاختبار . ولهذا يبدو الفرد الواحد متخلفا في قياس عندما في غيره تبعا لاختلاف عينة التقنين في كل منهما . فإذا كان لدينا تقياس مقنن على طلبة الجامعات وآخر مقنن على راشدين غير مختارين unselected من نفس السن فأى شخص سيبدو متخلفا في المقياس الأول عنه في المقياس الثانى .

٢ - اختلاف محتويات المقياس وما يقيسه من وظائف functions وما به من تعليمات وما له من خصائص وفي انحرافه المعيارى ومتوسطه الحسابى عن غيره .

وفىما يلى بيانات تثبت وجهة النظر هذه :

اختبرت ثلاث مجموعات متماثلة فى عمرها وفى فرقتهما ومستواها التحصيلى وطبق على كل مجموعة منها واحد من الاختبارات الآتية :

1 - Terman. Mc Nemer test of mental Ability

2 - Otis Quick Scoring Mental Ability Test

3 - Pintner general Ability Test verbal Series

ومن نتائج هذه الاختبارات الثلاث عمل جدول لنسب الذكاء المتوازية فى كل الاختبارات الثلاث . وذلك بإيجاد الدرجات التى توازى رتب المثبتات فى التوزيعات الثلاثة ، وهذا تحليل لاختلاف الاختبارات الثلاث فى نتائجها :

١ - فى اختبار پنتنر كانت نسب الذكاء أقل من ٢ إلى ٥ عنها فى اختبار تerman .

٢ - كان مقياس اختبار أوتيس ذا انحراف معيارى أقل من كلا مقياس تerman وپنتنر ، وهذا الاختلاف يجعل نسبة الذكاء قريبة من الوسط فى أوتيس عنه فى المقياسين الآخرين .

٣ - كانت نسبة ذكاء ٦٦ فى تerman تعادل ٧٦ فى أوتيس بينما نسبة ذكاء ١٤١ فى تerman تعادل ١٣٤ فى أوتيس

ولكى تكون المقارنة بين المقاييس صادقة لن يكفى أن يعبر كل مقياس فى معاييرهِ لوحدات متماثلة (عمر عقلى ، مئين ، درجة معيارية) بل يجب أيضاً أن تكون عينة التقنين متماثلة . وعند ذاك فقط يمكن مقارنة مختلف نتائج المقياس بغيرهِ .

وعلى كل حال فلا بد من نسبة الدرجة إلى المقياس الذى جاءت نتيجته فنقول مثلاً إن فلاناً نسبة ذكائه فى مقياس استانفورد بينيه كذا وإن فلاناً نسبة ذكاؤه الانحرافية فى المقياس كذا هى كذا : . . الخ .
وذلك لأن معيار المقياس ليس مطلقاً فى دلالته . فقد تعادل نسبة ذكاء ٩٤ على مقياس ما يعادل نسبة ذكاء ١٠٠ على مقياس آخر .

تعليق :

يمكن أن نحسب للاختبار معايير مختلفة ، وذلك لأن الدرجة الخام لا يمكن أن تفهم إلا على أساس هذه المعايير . ولكل من المعايير المستخدمة ميزاتها وعيوبها ولكن الشائع هو استخدام العمر العقلى ونسبة الذكاء فى اختبارات الأطفال واستخدام المئينات فى اختبارات البالغين .

ومهما كان المعيار المستخدم فيجب أن يفهم على أنه معيار خاص نوعى يتحدد بظروف التقنين وبالعينة التى استخدمت فى التقنين وخارج هذه الظروف وتلك العينة لا يكون من الصواب أن تنسب الدرجة الخام إلى معيار الاختبار .

فالاختبار الذى قنن فى أمريكا يحتاج إلى تقنين جديد واستخراج معايير جديدة إذا أعيد تطبيقه فى أى مجال آخر .

والاختبار الذى يقنن على أنه اختبار للسرعة للكبار قد يكون اختباراً للذكاء إذا استخدم للصغار . معنى هذا أن المعيار له صحة كما أن للاختبار صحة فالمعيار لاختبار السرعة مع الكبار يكون صحيحاً إذا اتخذ على أنه معيار سرعة مع الكبار ولكنه لا يكون صحيحاً إذا اتخذ على أنه معيار السرعة مع الصغار .

كيف نستخدم المعايير

رأينا أن المعايير تقدم لنا أساساً لتفسير درجة الفرد . فتحويل الدرجة الخام عن اختبار على حدة إلى مكافئ عمر أو فرقة دراسية ، أو إلى درجة مثبينة أو معيارية ، يسمح لنا بتفسير مستوى قدرة الفرد فيما يقيسه الاختبار المعين . فإذا جعلنا درجات الفرد عن مجموعة من الاختبارات تحول إلى معايير ذات وحدات موحدة ، وعرضناها في صفحة نفسية ، أمكننا أن نحصل على صورة وصفية للمستوى النسبي لأداء الفرد في مجالات الفصل ، أو مجموعة الفرقة الدراسية في المدرسة ، أو أطفال الفرقة الدراسية في النظام المدرسي كله . إذ بهذا نرى متوسط مستوى أداء المجموعة لوظيفة مفردة ، أو الأداء النسبي للمجموعة في كل من المجالات المتعددة . فالمعايير تقدم لنا إطاراً يمكن أن ننظر خلاله إلى هذه الصورة ، وأن نرجع كل أجزاء الصورة إلى ذلك الإطار العام . ولكن ما معنى هذه الصورة ، وماذا يجب علينا أن نفعل إزاءها ؟

لا نستطيع في هذا المجال أن نقدم تفسيراً كاملاً لأي مجموعة من الدرجات نحصل عليها من الإجراء في موقف الاختبار ، وسنقوم بهذا تبعاً في كل نوع من الاختبارات . إلا أنه بوسعنا أن نضع كل الخطوط العريضة والمبادئ العامة التي تساعدنا في تجنب الوقوع في أخطاء تفسير نتائج الاختبار . وسنبداً بتفسير نتائج الإجراء على جماعة . ثم تتبعه بتفسير درجات الفرد الواحد . ويجب أن ننبه إلى أن التفسيرين يتدخلان في كثير من النقاط .

(١) مبادئ تراعى عند تفسير درجات مجموعة من الأفراد :

١ - عند تقييم متوسط تحصيل مجموعة من الأفراد ، يجب مراعاة مستوى القدرة المتوسطة لهذه المجموعة . فإذا كان لدينا فصل من فصول

الفرقة الأولى الإعدادية ، وكان متوسط العمر العقلي لتلاميذه هو ١١ سنة ،
وجب أن لا نتوقع منه أداء اختبار حساب جودة أداء مجموعة أخرى
متوسط عمرها العقلي ١٢ سنة . وإن كان هذا لا ينطبق على كل المواد
الدراسية ، فايست العلاقة بين الذكاء والنجاح المدرسي في كل المواد
كبيرة ومتساوية .

٢- يجب أن نتوقع تأثير نوع ثقافة المفحوص على أدائه للاختبار .

٣- يجب عند أداء المجموعة أن يتم هذا على ضوء المنهج الدراسي ،
والنواحي التي أكدت فيه ، وأهداف التعليم العامة والخاصة .

٤- تستخدم نتائج الاختبارات بطريقة بنائية ، وليس لعقاب
المفحوصين .

(ب) مبادئ تراعى عند تفسير أداء الفرد الواحد :

١- يقيم التحصيل على ضوء معرفتنا باستعداد الفرد .

٢- يجب أن نضع في الاعتبار تأثير الأسرة والفروق الثقافية على الأداء

٣- يقيم أداء الفرد على أساس أنه اشترك في برنامج دراسي أو برنامج

تدريب مهني معين .

٤- في قياس الفرد الواحد يزيد احتمال تعرضنا للوقوع في أخطاء

التجريب .

قائمة مراجع القسم الثاني

- 1— Adkins, D. C. Construction and Analysis of Written tests for predicting Job performance. Educ. Psy. Measmt., 1946.
- 2— Adkins, D. C., et al. Construction and Analysis of achievement tests. Washington, D. C. : Government Printing office, 1947.
- 3— Adkins, D. C. Test Construction in Public Personnel administration Educ. Psy. Measmt., 1944.
- 4— American Psychological Association. Technical recommendations for Psychological tests and diagnostic techniques. Washington : Amer. Psychol—Assoc. 1954.
- 5— Anastasi, Anne, and Drake, J. An Imperical Comparison of Certain Tchniques for the Reliability of Speeded Tests. Educ. Psychol. Measmt., 1954.
- 6— Anastasia, Anne. The Concept of Validity in the interpretation of test scores. Educ. Psych. Measurements 1950.
- 7— Anastasi, Anne. The Influence of Practice Upon Test Reliability, J. Educ, Psychol , 1934.
- 8— Anastasi, Anne. Psychological Testing. (2. nd. edition), N.Y. : Macmillan Co., 1957.
- 9— Anderson, J.E. The Effect of Item Analysis Upon the Diseriminative Power of an Examination. J. Appl. Psychol., 1935.
- 10— Anonymous. Tables of Norms and Reliability Coefficients. New York : Cooperative Test Service, 1935.
- 11— Baron, D. and Bernard H. W. Evaluation Techniques for Classroom Teachers. Mc—Graw Hill Book Co. 1959.
- 12— Bean, K. Construction of Educational and Personnel. Mc — Graw — Hill Co N.Y. 1953.

- 13— Benett, G. K., Seashore, H. G., and Wesman, A. G. Differential Aptitude Test. N. Y. : Psychol. Corp., 1947.
- 14— Brown, W. Some Experimental Results in the Correlation of Mental Abilities. B. T. P., 1910.
- 15— Burt, C., The Reliability of Teachers Assessment of Their Pupils. B. T. Educ. P. Vol. XV, 1945.
- 16— Clapp, F. L., and Young, P. Self Marking Tests. Boston : Houghton Mifflin, 1929.
- 17— Conrad, H. S. Characteristics and Uses of Item—Analysis Data. Psychol. Monogr., 1948.
- 18— Cronbach, L. J. Coefficient Alpha and the Internal Structure of of Tests. Psychometrika, 1951.
- 19— Cronbach, L. J. Essentials of Psychological Testing. Harper Broths. N. Y., 1949.
- 20— Cronbach, L. J., and Warrington, W. G. Time — Limit Tests : Estimating their Reliability and Degree of Speeding. Psychometrika, 1951.
- 21— Cronbach, L. J. On Estimates of Test Reliability. J. of Educ. Psychol., 1943.
- 22— Cronbach, L. J. Test Reliability : Its Meaning and Determination. Psychometrika, 1947.
- 23— Grow, L. E. and Grow, A. Educational Psychology, American Book Co. (Revised ed) 1958.
- 24— Curtis, F. D., and Woods, G. G. A Study of A Modified Form of the Multiple — Response Test. J. Educ. Res., 1928.
- 25— Davis, A. The Measurement of Mental Systems (can intelligence be measured ?) Scient. Mon., 1948.
- 26— Davis, F. B. Item Analysis in Relation to Educational and Psychological Testing Psychol. Bull., 1952.
- 27— Davis, F. B. Item Analysis Data. Harvard Educ. Papers., 1946.
- 28— Davis, F. B. Item Selection Techniques, in Lindquist, E. F., (ed.) Educational Measurement. Washington : Amer. Council Educ., 1951.

- 29 — Dunlap, J. W. Relationship Between Type of Questions and Scoring Errors, *J. Exp. Educ.*, 1938.
- 30 — Flangan J. C. General Considerations in the Selectional Test Items and A Short Method of Estimating the Product — Moment Coefficient From the Tails of the Distribution, *J. Educ. Psy.* 1939.
- 31 — Flanagan, J. C. Units, Scores and Norms. in E. F. Lindquist (ed.) *Educational Measurement*. Washington : Amer. Council. Educ., 1951.
- 32 — Freeman, F. S : *Theory and Practice of Psychological Testing*, New York : Henry Holt and Co. Inc. 1950.
- 33 — Froelich, Clifford P., and Arthur L. Benson : *Guidance Testing*. Chicago : Science Research Associates Inc. 1948.
- 34 — Garrette, H. E. *Statistics in Psychology*. (Third edition) Longmans Green and Co. 1948.
- 35 — Ghiselli, E. E., and Brown, C. W. *Personnel and Industrial Psychology*. N. Y. : Mc Graw — Hill, 1948.
- 36 — Goodenough, Florence L. A. Critical Note on the Use of the Term "Reliability" in Mental Measurement. *J. Educ. Psychol.*, 1936.
- 37 — Goodenough Florence L. *Mental Testing*, Rinehart, N. Y. 1949.
- 38 — Good, C. V. and Merkel W R. *Dictionary of Education*. (2 nd edition) Mc Graw—Hill, N. Y. 1959.
- 39 — Greene, E.B. *Measurement of Human Behavior*. Odessey N. Y. : 1941
- 40 — Greene, L. B. *Michigan Vocabulary Profile*. *J. Higher Educ.*, 1938.
- 41 — Greene, H. A. and Jorgenson A. N. and Jerberich, J. R. *Measurement and Evaluation in The Secondary School*. Longmans Green and Co. N. Y. 1947.
- 42 — Guilford, J. P. *Psychometric Methods*. N. Y. : Mc Graw-Hill, 1936.

- 43— Guilford, J. P. New Standards for test evaluation. Educ. Psy. Measmt., 1946.
- 44— Guilford, J. P. Fundamental Statistics in Psychology and Education, Mc Graw—Hill Book Co. 1942.
- 45— Guilford, J. P. Factor Analysis in a Test development. Psy. Rev., 1948.
- 46— Guillihsen, H. The Reliability of Speeded Tests. Psychometrika, 1950.
- 47— Guliksen, H., Theory of Mental Tests, 1950. N. Y. : John Wiley & Sons.
- 48— Guttman. L. A Basis for Analysing Test — Retest Reliability, Psychom., 1945.
- 49— Guttman, L. Prablems of Reliability in Studies in Social Psychology in World war II, 1950.
- 50— Holzinger, K. J. Reliability of Single Test Item, J. Educ., 1932.
- 51— Horst, P. Estimating Test Reliability from Parts of Unequal length. Educ. P. Meas. 1951.
- 52— Humm, D. O. Test Validation on Remote Criteria. J. appl. Psy., 1946.
- 53— Jackson, R. W. B. Note on the Relationship Between Internal Consistency and Test — Retest Estimates of the Reliability of a Test. Psychometrike, 1942.
- 54— Jackson, R. W. B., and Furgson, G. A. Studies on the Reliability of Test. Univer. Toronto, Dep. Educ. Res. Bull., 1941.
- 55— Jenkins, J. G. Validity for what? J. Consult. Psy., 1946.
- 56— Jones, G., and Galbrich, A. The Interpretation of Standardized tests, Sch. and Soc., 1941.
- 57— Kelley, T. L. The Reliability Coefficient. Psychometrika, 1942.
- 58— Kelley, T. L., The Selection of the Upper and Lower Groups for the Validation of Test Items. J. Educ. Psy. 1939.

- 59 — Kuder, G. F, and Richardson, M. W. The Theory of Estimation of Test Reliability. Psychometrika, 1937.
- 60 — Landis, C., and Kaby, S. E. The Validity of Certain Questions which Purport to measure Neurotic tendencies. J. Appl. Psy., 1934.
- 61 — Lowley, D. N. A method of Standardizing Group Tests, B. J. Psych. Stat. Sect., 1950. P. P. 86 – 89.
- 62 — Lawshe, G. H., Jr. A Nomograph For Estimating of Test Items. J. Appl. Psychol., 1942.
- 63 — Lindquist, E. F., and Cook, W. W., Experimental Procedures in Test Evaluation. J. Exp. Educ, 1933.
- 64 — Lindquist, E. L. (Ed.) Measurement Handbook. Washington, D. C : American Council on Education, Boston : Houghton Mifflin Co., 1951.
- 65 — Lorge, I. The Fundamental Nature of Measurement. In. E. F. Lindquist (ed.) Educational. Measurement. Washington : Amer. Council Educ., 1951.
- 66 — Mc Adory, M. The Construction and Validation of an Art Test. N. Y. : Teachers College Contributions to Education, 1929.
- 67 — Mandell, M., and Adkins, D. C. Validity of Written tests for Selection of administrative Personnel Educ. · Psy. Measmt., 1946.
- 68 — Mathews, L. H. An Item Analysis of Measures of teaching ability. J. Educ. Res., 1940.
- 69 — Mc Call, W. A., How to Measure in Education, 1922.
- 70 — Mc Clusky, H. Y. An Experimental Comparison of Two Methods of Correcting the Outcome of an Examination. Sch. and Soc., 1934.
- 71 — Mc Nemar, Q. Psychological Statistics. N. Y. : Willey, 1949.
- 72 — Meyer, G. The Effect of Recall and Recognition of the Examination Set in Classroom Situation. J. Educ. Psychol., 1936.

- 73 — Moffie, D. J. Validity of Self — estimated interest. J. appl. Psy., 1942.
- 74 — Mosier, C. I. A Short Cut in Estimating of Split—Halves Coefficients. Educ.
- 75 — Mosier, C. I. Factors Influencing the Validity of a Scholastic interest test. J. Educ. Psy., 1937.
- 76 — Mosier, C. I. A Critical Examination of the Concepts of face Validity. Educ. Psy. Measmt., 1947.
- 77 — Mosier, G. I. A Critical Examination of the Concepts of Face Validity. Educ. Psychol. Measmt, 1947.
- 78 — Mursell, J. L. Psychological Testing. (2 nd edition) Longmans Green and Co. 1950.
- 79 — Nunnally Jum C. Tests and Measurements Assessment and Prediction. Mc—Graw—Hill Book Mo. N. Y. 1959.
- 80 — Otis, S. A. Normal Percentile Chart : Manual of Directions. Yonkers—On - Hudson, N Y. : World Book Co., 1938.
- 81 — Patterson, C. H. On the Problem of the Criterion in Prediction Studies. J. Consult. Psy., 1946.
- 82 — Peatman, J. G. Descriptive and Sampling Statistics. Harper Brothers. 1947.
- 83 — Recktenwald, L. N. Homogeneity of Items in the Cleeton Vocational Interest Inventory. Occupations, 1946.
- 84 — Remmers, H. H. and Gage, N L. and Rummel, J. F, Practical Introduction to Measurement and Evaluation. Harper Brothers 1960.
- 85 — Richardson. M. W., and Ruder, G, F., The Calculation of test Reliability Coefficients based upon the Method of Rational Equivalence, V. Educ., 1939.
- 86 — Richardson, M. W. The Interpretation of Test Validity Coefficient of Increased Efficiency of A Selected Group of Personnel. Psychometrika, 1944.
- 87 — Richardson, M. W. The Relation Between the Difficulty and the Differential Validity of a Test. Psychometrika, 1936.

- 88— Ruch, G. M., and Stoddard, G. D. Comparative Reliabilities of Five Types of Objective Examinations. *J. Educ. Psychol.*, 1925.
- 89— Ruch, G. M., and Stoddard, G. D. *Test and Measurement in High School Instruction*, N. Y. : World Book Co., 1927.
- 90— Ruch, F. L. Report on the Validation of the Mac Quarrie Test of Mechanical Ability. and the California Short Form Test of Mental Maturity, Language Section, in the Section of Winch Operators. Washington, D. C. : Dept. of Commerce, 1946.
- 91— Rulon, P. J. A Simplified Procedure for Determining the Reliability of A Test by Split—Halves. *Harv. Educ. Rev.* 1936.
- 92— Seder, M. The Reliability and Validity of the A. C. E. Psychological Examination, 1938 Edition *J. Educ. Res.*, 1940.
- 93— Shaffer, L. F. Information Which Should be Provided by Test Publishers and testing agencies on the Validity and use of their tests : Personality tests. *Proc 1949 Invit Conf testing Service.*
- 94— Sharkey, V. J., and Dunlap, J. W. Study of Reliability and Validity of the Academic Preference Blank. *J. Educ. Psy.*, 1940.
- 95— Schmidt, H. O. Test Profiles as a Diagnostic aid : The Minnesota Multiphasic Inventory. *J. appl. psy.*, 1945.
- 96— Schmidt, H. O., and Billingslea F. Y. Test Profiles as a diagnostic Aid : the Bernreuter Inventory. *J. abnorm. Soc. Psy.*, 1945.
- 97— Schneider, G. G. Grade and age norms for the Minnesota Vocational Test for Clerical Workers. *Educ. Psy. Measmt.*, 1941.
- 98— Schultz, R. S. Standardized Tests and statistical Procedures in the Selection of life insurance Sales Personnel. *J. appl. Psy.* 1936.

- 99— Skinner, C. E. Educational Psychology. Prentice Hall 1950.
- 100— Spearman, C. Correlation Calculated from faulty Data. B. T. P. 1910.
- 101— Staff, Personnel Research Section, A. G. O. Validation of the General Clerical Abilities Test for Selection and Placement of War Dept., Civilian Personnel. Amer Psy., 1947.
- 102— Staff, Test and Research Section, Bureau of Naval Personell. Psychological test Construction and Research in the Bureau of Naval Personnel : The Basic Test Battery. Psy. Bull., 1945.
- 103— Stephens, E. W. A Comparison of New England norms with National Norms on the revised Minnesota Paper Form Board Test. Occupations, 1945.
- 104— Stevens, S. S. On The Theory of Scales of Measurement. Science, 1946.
- 105— Strong, E. K., JR. Weighted Vs. Unit, Scales. J. Educ. Psy., 1945.
- 106— Stuit, D. B. The effect of the nature of the criterion upon the Validity of aptitude tests. Educ. Psy. measmt., 1947.
- 107— Stuit, D. B., and Wilson, J. T. The effect of on increasingly well Defined Criterion on the Prediction of Success at Naval Training School (Tractical Radar). J. appl. Psy., 1946.
- 108— Super, D. E. The Validity of Standard and Custom-build Personality inventories in a pilot selection Program Educ. Psy. Measmt., 1947.
- 109— Super, D. E. Appraising Vocational Fitness by means of Psychological Tests. N. Y. : Harper Brothers, 1949.
- 110— Thomson, G. H. Standardization of group tests and the Scatter of intelligence Quotients B. J. Educ. Psych., 1932, esp, p. 91.

- 111.— Taylor, K. V. F., and Carter, H. D. Retest Consistency of Vocational Interest Patterns. J. Consult. psy., 1942.
- 112.— Taylor, K. V. F. Reliability and Performance of Vocational interests of Adolescents. J. exp. Educ, 1942.
- 113.— Taylor, H. G., and Russell, I. T., The Relationship of Validity Coefficients to the practical Effectiveness of Tests in Selection : Discussion and Tables, J. Ap. 1939.
- 114.— Thorndike R. L. and Hagen. A. Measurement and Evaluation in Psychology and Education. John Willy Sons Inc. N. Y. 1957.
- 115.— Thorndike, R. L. Reliability. in E. F. Lindquist (ed.) Educational Measurement. Washington : Amer. Council Educ, 1951.
- 116.— Thorndike, R. L. Personnel Selection. N. Y. : John Wiley, 1949.
- 117.— Thorndike R. L. (Ed.) Research Problems and Techniques, A A F Aviation Psychology Report, No. 3 Washington, D. C. : Government Printing Office, 1947.
- 118.— Thurstone, L. L. Current Issues in Factor Analysis. Psychol. Bull., 1940.
- 119.— Thurstone. L. L. A Method of Scaling Psychological and Educational Tests, J. Educ. Psychology, 1925.
- 120.— Thurstone, L. L. Scales Construction with Weighed observation, Psy., 1928.
- 121.— Thurstone, L. L., The Absolute Zero in Intelligence Measurement, Psy. Rev. 1928,
- 122.— Thurstone, L. L. The Unit of Measurement in Educational Scales, J. E. Psy., 1927.
- 123.— Thurstone, L. L. Psychological Implications of Factor Analysis, Amer. Psychologist, 1948.
- 124.— Tiffin, J. Industrial Psychology. (3 rd ed.) N. Y. : Prentice Hall, 1952.

- 125 -- Trabue, M. R. Graphic Representation of Measured Characteristics of Successful Workers. Occupations, 1934.
- 126 -- Travers, R. M. W. Educational Measurement. Mac-Millan Co. N. Y. 1955.
- 127 -- Vernon, P. E. Indices of Item Consistency and Validity. Brit. J. Psychol., Statis. Sect., 1948.
- 128 -- Wadsworth, G. W. Tests Prove their worth to a utility Pers. J., 1935.
- 129 -- Walker, Helen M. Item Selection by Sequential Sampling. Teachers Coll. Rec., 1949.
- 130 -- Warren H. C. Dictionnry of Psychology Houghton Mifflin Co. 1934.
- 131 -- Wilson' J. W., and Carpenter, E. K. The Need for resistandadization of altered test. Amer. Psy., 1948.

١٣٢ - السيد محمد خيرى : الإحصاء فى البحوث النفسية والتربوية والاجتماعية . القاهرة ، دار الفكر العربى سنة ١٩٥٦ .

١٣٣ - فؤاد البهى السيد : علم النفس الإحصائى وقياس العقل البشرى . القاهرة . دار الفكر العربى سنة ١٩٥٨ .

١٣٤ - محمد عثمان نجاني : علم النفس الصناعى . القاهرة ، مكتبة النهضة المصرية سنة ١٩٦٠ .

القِيمُ الثَّلَاثُ

القياس التربوي

الفصل العاشر

التربية والقياس والتقييم

- القياس النفسى والقياس التربوى .
- لماذا تتغير التربية – نظرة تاريخية .
- القياس فى التربية .
- تعريف بالتربية .
- معنى التحصيل .
- أهداف التربية .

الفصل العاشر

التريية والقياس

للقياس النفسى والقياس التربوى:

لا يختلف القياس التربوى عن القياس النفسى فى شىء ، بل ولا يختلف القياس فى ميدان عنه فى ميدان آخر فى شىء ، وكل ما نلاحظه هو خصائص نتحدد بها فى ميدان بذاته لأنها تحدد هذا الميدان ذاته . فالقياس ، أين استخدم ، يجب أن يكون موضوعيا مقننا ، والقياس ، متى أجرى ، يجب أن يكون ثابتا صادقا ، والتقييم ، وكيف أعد ، كفى كى يقدر ويفسر ، والدرجة ، لماذا أعطيت ، ذات معنى ودلالة . وأكثر من هذا فالقياس والتقييم يقتصران على موضوع واحد هو سلوك الإنسان باعتباره محصلة عدة قوى تعكس لنا صورة تمثل الإنسان ذكاه ، وقدراته ، واستعداداته ، وتحصيله ، شخصيته بميله ، وقيمه ، واتجاهاته ، ورأيه . وهكذا يشترك القياس والتقييم فى مميزات عامة وخطوط عريضة نبعث أساسا من طبيعة علم القياس وموضوعه الإنسان . ولكن هذا لا يعنى تطابق القياس فى مختلف المجالات والأغراض . فلكل مجال ولكل غرض ، فضلا عن الخطوط العريضة والخصائص العامة المشتركة ، مميزات تخصه . هذه المميزات أتت من الهدف من عمليات القياس فيه وتاريخه وإمكانات الباحث بالقياس ودوره وكيفية استخدام نتائج .

إلا أن القياس التربوى والقياس النفسى يقتربان فى الكثير . فإن الغالبية العظمى من الاختبارات والمقاييس جربت وقتنت وأجريت على التلاميذ فى مختلف الفرق الدراسية . وعلى هذا فإن هذه الاختبارات توثق ثمارها

عندما تستخدم في المدارس والمؤسسات التربوية . كما أن القياس نفسه نشأ ونمى في أحضان التربية وما يزال يأخذ من مشكلاتها دفعات قوية يحسن بها أدواته ويوائمها مع المشاكل الحالية . ولا يغيب عنا بالطبع أن دارس التربية درس ، أو يجب أن يدرس ، القياس كما أن المشتغل بالقياس درس التربية أو هو على الأقل قضى جزءاً كبيراً من حياته في مؤسسات التربية والتعليم أثناء إعدادة .

والمشكلات التي نجدها في المدرسة ، عموماً ، هي نفس المشكلات التي تعالجها المقاييس المختلفة فاختبارات الذكاء واختبارات الشخصية أدوات تعين المشتغل بالتربية على المسح ، والتصنيف والإعداد ، والتشخيص والعلاج والتنبؤ و . . الخ عندما يقوم ببحث يدرس فيه حالة تلميذ أو فصل أو النظام التعليمي كله . كما أن الاختبارات التي يضعها المدرس teacher-made وامتحانات النقل والعامية أنواع من المقاييس يمكن بإدخال بعض التعديلات عليها ، كما سنرى فيما بعد ، أن تصبح ، أدوات موضوعية تماماً كالاختبارات المقننة .

ونظرة سريعة إلى كليات التربية في بلادنا والخارج نجدها تضم الكثير من أساتذة علم النفس الذين أسهموا بالكثير في تقدم حركة القياس ، كما أن الكثير من المشتغلين بالقياس كانوا أساتذة في التربية أو مشتغلين بها . ونظرة إلى كتب القياس التربوي نجدها لا تختلف عن كتب القياس النفسي إلا في النذر اليسير بل إن الأولى تشمل فصولاً من الثانية والثانية تغطي كثيراً من موضوعات الأولى . ونظرة إلى المؤسسات التربوية في الخارج نجدها ألباب الواسع الكبير الذي يفتح للأخصائيين النفسيين بمختلف تخصصاتهم فيفيد منهم ويفيدون منه . ولقد بدأنا نسلك هذا السبيل ، وإن كانت هناك بعض التقيبات ، وإن كان التقدم بطيء ، ولكننا نأمل أن تكون هذه هي البداية ، وأول الغيث قطرة . وكذلك فإن البحوث التربوية إن لم

يقيم بها ، في بلادنا والخارج ، أخصائيون نفسيون ، فإنهم على الأقل يشتركون فيها .

وهكذا تضيق الهوة ، التي قد يتصورها البعض ، بين التربية وعلم النفس أو بين القياس التربوي والقياس النفسى . وهكذا أيضاً نستطيع أن نفيد من الأدوات التي نعرضها هنا في القسمين التاليين كما أننا سنفيد في القسم الحالى فعلا من الأدوات التي سنعرضها فيهما على أن يرجع القارىء إلى ما نشير به .

لماذا تتغير التربية - نظرة تاريخية

يعتقد رجال التربية أن تقدم العلوم التربوية رهين بالبحوث العلمية التي تهدف إلى اكتشاف الحقائق لتحديد قيمة العمليات والنظم والطرق التربوية ، وأن هذه البحوث يجب أن تتبعها محاولات لتحسين وتطوير نظم التربية والتعالم . ثم توضع هذه النظم الجديدة موضع البحث وتجمع الحقائق العلمية لمعرفة ما إذا كانت تفضل النظم القديمة أم تفتقر إلى بعض مزاياها ؟ وما مدى هذا التحسن ؟ وهل يستحق الجهود التي تبذل في البحث وفي التغيير أم لا ؟ وهل يمكن أن يطور النظام الجديد أم لا ؟ وإلى أى حد ؟ وما هى التأثيرات المتوقعة لهذا النظام الجديد ؟ وهل يؤدى إلى تغيرات جوهرية فعلا في إعداد المواطن الصالح وفي تنمية حصيلة التلاميذ أم لا ؟ وهذه الخطوة تستند إلى افتراض مؤداه أن الإنسان كائن منطقي مفكر إذا واجهناه بحقائق تكشف عدم كفاية الطرق التي يتبعها ، فإنه قطعاً سيحاول أن يغير هذه الطرق خاصة إذا كان كسبه ، من أى نوع نتيجة التغيير ، كبيراً . ولكن هذا الافتراض ليس صحيحاً تماماً . فإن تاريخ التربية يدلنا على أن الكثير من النظم التي كشف البحث العلمى عن جدواها لم تقلح في تغيير بعض النظم التربوية القديمة وخاصة منها ما كان له تاريخه كذا أن

تغيير النظم التربوية والتعليمية ، خاصة إذا كان كثيراً ومتقلبا ، مهما كانت فوائده مستقبلا ، يؤدي إلى بلبلة الأفكار واضطراب المتعلمين فضلا عن إعداد جيل جديد من المعلمين يجيد استخدام الطرق الجديدة ويؤمن بالنظام المزمع قيامه فلا فائدة لتشريع لا ينفذ ولا جدوى في إصدار أمر إلى من لا يستطيع تنفيذه .

يجب ، إذن ، أن نعد المعلمين لهذه التغييرات وندرّبهم على الطرق الجديدة ونقنعهم بالقيم الحافزة لهذا التغيير . وقد نادى بهذا الرأي أحد رواد التربية في القرن الماضي وهو جوزيف ماير ريس Joseph Mayer Rice الذي درس الكثير من النظم التربوية في أوروبا وأمريكا وخرج إلى أن التلميذ كان يتعلم القراءة والكتابة آليا دون أن يكتسب أى أفكار جديدة . والقراءة Reading ، والكتابة Writing ، والحساب Arithmetic (كل منها تبدأ في الإنجليزية نطقا بحرف R ولهذا تسمى three R's) عند ريس يجب أن نعلمها للطفل بطريقة مشوقة وأن يبدأ تعليمه إياها في الوقت الذي يسمح له نضجه بأكفاً تعلم ، وأن هذا السن صغير بدرجة . نجعلنا نستطيع أن نعلم الطفل ، ما دامت سنوات تعليمه ستزيد ، الكثير من المهارات التي سيحتاجها فيما بعد .

كانت آراء ريس صحيحة ، ولكن أحدا لم يعره اهتماما . ولكنه سرعان ما عرف أن رأيه لم تؤيده حقائق تعلن للناس ، ومن هنا بدأ يجمع ملاحظاته واستخدم برامج الاختبارات testing programs ليبين أن طرق التعليم المتبعة في ذلك الوقت غير كافية . وصمم اختبارات في الحساب ، والمجاية spelling ، واللغات ، ولكنه اكتشف أن جمع الأدلة عن تحقيق أحد الأهداف التربوية عمل صعب ، حتى لو كان الهدف سهلا نسبيا مثل إكساب التلاميذ مهارة المجاية . فبعد أن أجرى اختبار المجاية الأول الذي صممه على ١٦٠٠٠ تلميذا ، وجد أن تلاميذ مدارس معينة مستواهم

فى الاختبار منخفض وقد عرف من جولاته أن مرجع هذا إلى عيوب نظام الامتحانات الذى كان يتبع فى هذه المدارس .

ولهذا صمم ريس اختبارا ثانياً للهجاية خاليا من العيوب التى كانت تؤثر فى نتيجة الاختبار الأول ، وهى أن بعض المدرسين ، قلقا فى إعداد تلاميذهم للاختبار ، زدوهم بنطق بعض الكلمات نطقا واضحا ساعدهم فى هجائها ولهذا اختلفت نتائج هؤلاء التلاميذ عن نتائج غيرهم فى اختبار الهجاية . والنتيجة اللازمة لاستبعاد الكلمات التى يساعد نطقها على هجائها أن صدق اختبار الجديد ارتفع . وقد انتهى إلى أن مدة تعليم الهجاية فى المدرسة لا تؤثر كثيراً على مهارة التلميذ فى الهجاية إذ أن الأخيرة تتوقف على عوامل كثيرة أخرى ، ولهذا يحسن أن تختصر هذه المدة وأن تستغل فى تعليمه باقى نواحى المنهج الدراسى خاصة وأن هذا الاختصار لن يؤثر على مستواه فى الهجاية كثيراً .

هذا الأسلوب العلمى الذى اتبعه ريس كان صرخة فى وادى مهجور ، فلم تؤثر فى النظام التعليمى فى أى بلد بل إن الكثيرين من مؤرخى التربية يجدون فضله ما دام لم يغير نظاما رغم علمية طريقته وخطواته التجريبية على أعداد هائلة . والخطأ الوحيد الذى ارتكبه هو اعتقاده بأن عرض الحقائق على رجال التربية سيغير من طرقهم . وقد أفادت البحوث التربوية الحديثة من هذا الخطأ إذ لم تعد تهدف أساسا إلى تطوير البحث العلمى بل أصبحت تعنى أولا بتطوير التربية . ولهذا يجب أن يلم المعلمون بهذه النتائج وأن يفهمها القائمون بوضع برامج التعليم . ولكن إذا آمنت هيئة تقوم بالإشراف على المناهج الدراسية بأن الوقت قد حان لتغيير نظام تعليمى متبع فإنها بمجهودها الضخمة تستطيع أن تقوم ببحوث واسعة تجمع فيها الحقائق وتعرضها بطريقة تضمن التغيير إلى أفضل .

ونحن نعتقد أن المدرس يجب أن يحمل عبء المسؤولية الكبرى فى

تقدير نتائج التعليم . وإذا كنا نرى أن الآمال تعقد على المستشارين الذين يوجهون ويخططون هذه الدراسات ، إلا أن الجهد الأكبر يقع على كاهل المدرس نفسه . وهذا التقدير يجب ، كما سنرى ، أن يبعد ، قدر الإمكان ، عن الطرق غير الرسمية وأن يكون أكثر منهجية . ونأمل أن يفيد رجال التربية ، وعلى الأخص المدرسون ، من هذا القسم في إرشادهم إلى ميادين البحث التربوي وطرقه . ولكن هذا لن يتم إلا إذا اتسعت حركة القياس بما يجعل كل المدرسين مدربين على طرق القياس والتقييم العلمية الموضوعية . وهذا بالطبع لا يعنى أن يصبح كل مدرس خبيراً في القياس .

القياس في التربية

في القياس التربوي نعنى بالقيم التي حققها نظام معين . وهذا النظام قد يكون كل المؤسسات التربوية ، أو برنامجاً متضمناً فيها ، أو التخطيط الذي أدى إلى اتباع هذا البرنامج ، أو نتائج هذا البرنامج مثل ما يحصله التلاميذ الذين يعرضون له ، أو . . الخ . وما دام البرنامج التعليمي يقيم خلال ما أحدث من تغيرات في من اشتركوا فيه ، وهم المدرسون والتلاميذ ، فسنعرض للبرنامج في فصل ولتقييم المدرسين في فصل ولقياس التلاميذ في فصول .

وننتج البرنامج هي التغيرات الفعلية التي أحدثها في التلاميذ ، وأهداف البرنامج إذن هي توقع تحصيل هذه التغيرات مستقبلاً .

وليس القياس التربوي قاصراً على عملية تحديد النتائج الفعلية التي أحدثها البرنامج الدراسي ، بل إنه يشمل أيضاً إصدار أحكام تتعلق بماهية النتائج المرجو تحصيلها . ولكن القياس إذا تحول من تقدير النتائج خلال السلوك ، إلى معرفة ما إذا كانت هذه النتائج مرغوب فيها أم لا ، يصبح تقييماً evaluation . والحق أننا في نهاية عمليات القياس نصدر مثل هذه

الأحكام التقييمية حتى نحدد معالم الطريق الذى يجب أن تسلكه البرامج الدراسية .

فإذا حددت الأهداف وصممت المقاييس التى تعرفنا بمدى التقدم نحو هذه الأهداف ، أمكننا أن نعين مدى نمو الفرد فيما يتعلق بهذه الأهداف دون أن نصدر أحكاما تقييمية . وفى هذه الأثناء نصدر أحكام التقييم فقط عندما نقوم باختيار الأهداف التى نود تحصيلها . ومنذ أن نختار الأهداف ، بعد أن قيمناها ، لا نلجأ ثانية إلى التقييم ، بل نقيس المركز الحالى للفرد بالنسبة للمدى الذى قطعه من الطريق إليها . ولهذا يحسن أن نحدد استخدام اصطلاح «التقييم» باختيار الأهداف ، واصطلاح «القياس» *assessment* بعمليات تعيين المركز الراهن للفرد . وقد اصطلح على استخدام تعبير «التقييم» فى عمليات القياس التى تتضمن إصدار أحكام تقييمية إلى جوار القياس ، واستخدام تعبير «القياس» فى العمليات التى لا تتضمن إلى جوار القياس إصدار مثل هذه الأحكام .

فمحدد الطرق التى نرغب أن يسير عليها التغير الحادث ، فى التلاميذ بعد دخول البرنامج الدراسى ، هو صلب عمليات اختيار الأهداف المنشودة من البرنامج . وهذه العملية تتم قبل تنظيم البرنامج . وتوضع أهداف البرنامج فى صيغة يفهم منها أنها تغيرات مرجوة ، أو نتائج متوقعة جدير بالبرنامج أن يحققها فى التلاميذ .

ويمكن تحديد مدى فاعلية برنامج تعليمى بطريقة أفضل من تعيين التغيرات المطلوبة التى أحدثها . ويجب أن يتم هذا على ضوء الظروف التى تتم خلالها عمليات التعلم . ولكن شرح ووصف هذه الظروف ، دون معرفة محددة بكمية واتجاه عمليات التعلم ، لا يكفى لتقييم البرنامج ولا لقياس عمليات التعلم الحادثة . ولكن هذه المعرفة المحددة لا يستطيعها المدرس العادى فى فصله . بل إن التقديرات التى يقوم بها الملاحظون الخارجيون

الأقل خبرة من المدرس بالعوامل المتدخلة في مواقف معينة والتي تؤثر في التعلم ، قد تكون أفضل من تقديرات المدرسين حتى الخبراء منهم ما دامت تتم في الفصل : وواجب القياس التربوي هنا أن يعين المدرس على تطوير وسائل الحكم الدقيق وإصدار تنبؤات صحيحة عما سيحدثه التعلم من زيادة في المعارف بعد أن نكشف للمدرس عن طبيعة عمليات التعلم وظروفها وطبيعة من يقوم بتعليمه .

تعريف بالتربية

نعني بالتربية هنا كل العمليات التي تهدف إلى تغيير السلوك : وعلى أساس هذا التعريف يتم جزء كبير من التربية خارج المدرسة ويشمل محاولات الجرائد والمجلات في تغيير الاتجاهات ، ومحاولات الأشرطة السينمائية ، ومحاولات الإذاعة ، والمحاضرات والندوات ، هذا فضلا عن دور الآباء ، ومحاولات رجال الدين ، ورجال التربية الاجتماعية والرياضية ، ومحاولات كل الأفراد والهيئات الذين ييغون تنمية الفرد وإنمائه بإحداث تغيرات مرغوب فيها بحيث يتجه سلوكه اتجاهات معينة . كما أن تعريف التربية بهذا المعنى الشامل يدخل عوامل لا حصر لها كلها تهدف إلى إحداث تغيرات مرغوبة .

وقد حددت وزارة التربية والتعليم في الإقليم الجنوبي وظائف التربية في تقرير رسمي بما يأتي :

١ - اكتساب التلاميذ القدرة على فهم العلاقات الاجتماعية الصالحة وممارستها :

٢ - مساعدة التلميذ على كشف الصالح من استعداداته وقدراته .

٣ - تنمية عادات التفكير المنطقي السليم .

٤ - تعود النشاط المنتج .

٥ - تحصيل المعلومات وأدوات المعرفة .

٦ - تكوين جسم سليم .

٧ - تكوين اتجاهات عقلية صالحة .

ربط أهداف القياس بأهداف التعليم :

ترتبط عمليات القياس ارتباطاً وثيقاً بتحديد أهداف برنامج التدريس ، الذى على أساسه نقيس التحصيل . ومرجع الاختلاف الأساسى فى عمليات القياس والتقييم إلى الاختلافات فى تحديد الأهداف ، بل إن بعض المشتغلين بالقياس التربوى لا يضعون اعتباراً لهذه الأهداف . وإذا أردنا أن نستخدم الاختبارات التحصيلية المقننة وجب أن نتضح فى أذهاننا هذه الأهداف . وعلى ذلك يحسن أن يقوم واضع الاختبار بتحديد الأهداف التى يرى أن التعليم فى المادة المقاسة يهدف إليها ، حتى تصبح الدرجات فى مقياسه ذات دلالة على مدى تحصيل التلاميذ لها .

التقييم الرسمى والتقييم غير الرسمى :

نعالج هنا دراسات التقييم والقياس المنهجية الرسمية ، إلا أنه يجب ألا يغيب عنا أن أغلب عمليات التقييم والقياس فى التربية ليست منهجية ولا كذلك رسمية ، فإن كلا من المدرس والتلميذ يصدر أحكامه عن قيمة الجهد الذى بذله التلميذ ، والمواد المدروسة فى الفصل ، وطريقة التدريس والعلاقات الاجتماعية داخل الفصل و... الخ . وكذلك فإن الآباء وغيرهم من المواطنين خارج المدرسة يقيمون العمليات التربوية تقيماً غير رسمى وغير منهجى ، ولكن هذا التقييم يجب أن يشجع لأنه قد يحمل فى طياته إمارات الديمقراطية الصحيحة واشتراك كل المواطنين فى تقييم أمر يعد فى المكانة الأولى بالنسبة لهم ألا وهو التقدم التربوى .

وإن كان من الواجب أن تتم عمليات القياس والتقييم فى التربية على

أساس منهجى رسمى ، إلا أن هذا يتوقف على إعلان نتائج الكثير من البحوث التى تستغرق جهوداً مضنية وتحتاج وقتاً طويلاً ، ولهذا فإن التقييم غير الرسمى غير المنهجى يدفع هذه البحوث . فإذا وضعنا فى الاعتبار أن التقدير المنهجى لنتائج برنامج دراسى محدود يتكلف كثيراً إذ أنه لا يتضمن التعريف الدقيق للنتائج المرجوة وحسب ، بل إنه أيضاً غالباً ما يقتضى إنشاء مقاييس جديدة وجمع الوفير من البيانات لتحديد مدى صدق وثبات هذه المقاييس قبل استخدامها .

كما أننا نعلم أن القياس التربوى ميدان حديث نسبياً ، وأن عدد المقاييس التى يمكن استخدامها فيه عن ثقة عدد محدود . فإذا أضفنا إلى هذا حقيقة أن المدرسين فى تقديرهم لمدى نمو تلاميذهم يميلون إلى المبالغة ، انضحت لنا قيمة إنشاء اختبارات تربوية جديدة رغم صعوبة وضخامة هذا العمل . ومن الجدير بالذكر أن نشير إلى تعليق برتراند رسل Bertrand Russell عن تدريس الهندسة ، إذ يقول إن هذا العلم ظل إلى وقت قريب يدرس منفصلاً عن غيره من الميادين الرياضية مع العلم بأن هذا الفصل غير جائز إطلاقاً . وأهمية هذه الإشارة تنحصر فى وجوب رجوع مصمم القياس التربوى إلى تاريخ التربية كى يستمد منه حقائقاً تعينه على فهم طبيعة المادة .

مشكلة المحكات فى القياس التربوى :

فى الفصل السادس من هذا الكتاب تعرضنا لمشكلة المحكات ، وإذا رجعنا إلى وبستر Webster وجسدناه يعرف المحك بأنه معيار للحكم ؛ قاعدة كان أو اختباراً ، عن طريقة تقييم الحقائق ، والأسس ، والآراء ، والسلوك عند إصدار حكم صحيح يتعلق بها . وفى كل عمليات القياس والتقييم ، يبدأ الباحث بإقرار معيار معين به يحكم على صلاحية الأداة التى سيستخدمها . وفى برامج التعليم يجب أن توضع هذه المحكات على

أساس نريد به قياس أثر البرنامج . وهذه المحكات تقوم أولا على وصف ما يستطيعه التلاميذ في فرقة دراسية معينة بالنسبة لوظيفة يقيسها الاختبار الذى سيطبق عليهم .

وعلى هذا فالمحك في اختبار القراءة مثلا قد يكون وصفا مفصلا لما نتوقع أن تكون عليه قدرات القراءة لدى تلاميذ عمر معين وفرقة دراسية معينة وذكاء معين . هذه القدرات يجب تعريفها بدقة . ويحسن ألا يقتصر المحك على المستوى المتوقع فقط ، بل وأن تراعى فيه أيضاً كل العوامل المؤثرة في عمليات التعلم في موقف ندرسه . والمحك الذى يغلب استخدامه في القياس التربوى هو النجاح المدرسى . وهو محك يعتقد أنه ثابت طوال حياة الفرد ، بل وإنه يشمل أكثر من التحصيل المدرسى . فأحياناً يؤخذ هذا المحك على أنه دليل على نجاح الفرد أو توافقه في مجالات أخرى غير الدراسة . وهذا التعميم غير جائز إذ أنه يشمل جوانب لا يقيسها اختبار تحصيلي واحد ، بل ويشمل جوانب لا يستطيع قياسها بالدقة المنشودة باستخدام أى اختبار موجود .

وعلى هذا يحسن أن يكون المحك قاصرا على منطقة محددة من التحصيل ثبت أنها ترتبط بالموضوع الذى يقيسه الاختبار . فإذا لم تتوفر في المحك هذه الصفة ، كان غير مجدى . فإذا افترضنا أننا بصدد قياس نجاح التدريس ، وأننا اخترنا محكا عاما هو مدى ما حصله التلاميذ من أهداف هذا التدريس . رأينا أن قياس نمو التلاميذ يعيننا بطريقة غير مباشرة على قياس نجاح تدريسهم . والواقع أن قياسا من هذا النوع لا نقدر كفاءة المدرس ، إذ أن نمو التلاميذ في المادة المدروسة لا يرجع فقط إلى كفاءة المدرس بل إن ثمة عوامل أخرى كثيرة تتدخل مثل نموهم العقلى والانفعالى والجسمى وأثر خبراتهم خارج المدرسة . ولما كنا نفتقر إلى مقاييس موضوعية لتقدير كفاية المدرس ، فإننا لا نجد بدا من استخدام محك

تقدم التلاميذ ، على أن نضع في اعتبارنا دائماً أنه محك تقريبي . على أن هناك طرق أخرى لتقييم المدرس سنورد ذكرها في فصل آخر .

معنى التحصيل

نعني بالتحصيل حدوث عمليات التعلم التي نرغبها : وما دام التعريف يتضمن كلمة « نرغبها » فإنه بالتالي يتضمن حكماً تقييمياً بمعنى أن قيمة التحصيل تتوقف على ما إذا كانت عمليات التعلم تسير في الاتجاه الذي يعتبر أساسياً في نظر صاحب اختبار التحصيل . ونحن نقصد بالتعلم عمليات تعلم الحقائق والمعلومات بما في ذلك الاتجاهات الاهتمامات والقيم ما دامت ، ولو جزئياً ، مكتسبة أى متعلمة . كما أن التحصيل يتضمن أيضاً تعلم أنماط السلوك التي تصبح سمات تميز الشخصية . وبرغم اتساع هذا التعريف إلا أننا غالباً ما نقصره على تحصيل التلاميذ أو اكتسابهم لما تهدف إليه المدرسة أو المدرس أو نظام التعليم عموماً . إذ أن ثمة دلالات مؤكدة على تأثير المدرس ، في تنمية جوانب غير عقلية لدى التلميذ ، لم تتوفر بعد .

طبيعة التحصيل :

يجب أن نفرق بين الاستعداد Aptitude للتعلم وبين التحصيل ، أو التعلم فعلاً ، ويقصد دائماً بالاستعداد للتعلم الذكاء كما نقيسه اختبارات الذكاء . وعلى هذا يفترض أنه لو تساوت الفرص أمام كل المتعلمين ، فإن الفروق بين درجاتهم تعد نتيجة اختلاف درجاتهم للتحصيل أى اختلاف نسب ذكائهم .

وتختلف اختبارات التحصيل المدرسي عن اختبارات الاستعداد العام

فى أن النتائج على الاختبار المدرسى ترجع إلى اختلاف درجات المهارات والمعارف الناتجة عن اختلاف استفادة التلاميذ من التدريب الخاص . ويشد التدخل بين اختبارات الاستعدادات واختبارات التحصيل في أحد أنواع اختبارات التحصيل المسمى باختبار الاستعداد Readiness Test : والاستعداد بالمعنى الأخير هو القدرات الخاصة ، والمهارات ، والمعارف التى يتطلبها النجاح فى برنامج معين يمر به الفرد أثناء تعلمه :

والتحصيل هو نتيجة تفاعل عوامل ثلاثة هى :

الاستعداد aptitude للتعلم ، والاستعداد readiness or preparedness وفرصة التعلم .

أهداف التربية

إن الاهتمام المتزايد الذى يبدىه المدرسون بقياس وتقييم برامجهم بطريقة منهجية يدفعنا إلى إعادة مناقشة المشاكل المتعلقة بالخطوات الرئيسية فى أى عملية قياس . ونقصد خطوات تحديد الأهداف واختيار محكات القياس المتضمنة :

الأهداف الصريحة والأهداف الضمنية :

أهداف البرامج التعليمية غالباً ضمنية implicit بمعنى أن واضعى البرامج يضعون الأهداف فى أذهانهم ولا يشعرون بالحاجة إلى إعلانها سواء أكان الإعلان كتابة أم كان تحديداً لما يقصدونه . ولا يقلل هذا من شأن البرامج أو الأهداف ما دامت الأهداف واضحة فى أذهان القائمين بوضع البرامج الدراسية : ولكن القائم بالقياس أو التقييم قد لا يفسر أهداف البرنامج الذى يقيسه تأثيره تفسيراً يتفق مع ما كان يقصده واضع البرنامج ،

ولهذا يحسن دائماً أن تعلن الأهداف في فقرات يسهل تفسيرها . وهذا التحديد الصريح يكون إشكالا أمام مخطط البرنامج الدراسي .

القيم المؤثرة في إختيار الأهداف :

لا يمكننا هنا مناقشة كل العوامل التي يجب أن يضعها منظم البرامج المدرسية موضع الاعتبار عندما يحدد الأهداف التربوية العامة التي يرغب أن يصل إليها التلاميذ بعد دراسة هذا المنهج . وتزيد هذه العوامل عندما نكون بصدد وضع أهداف مرغوبة للتعليم العام ، وتقل بالنسبة للتعليم المهني أو الفني . ففي الحالة الأخيرة تتضح الأهداف إذ أن الغرض من التعليم المهني عادة هو إعداد أفراد يتميزون بخصائص معينة ، وتعدد الأهداف عندما نكون بصدد التعليم العام ولهذا يحسن أن نحدد الأسس التي بناءاً عليها يُختار هدف ويرفض آخر .

والموازنة بين أسس اختيار الأهداف ترجع ، على وجه العموم ، إلى الفلسفة التي يتبناها أو يتأثر بها واضع البرنامج . وقد يبدأ رجل التربية بافتراض أن القيم التي ينبغي أن يحصلها التلاميذ متمثلة في البرنامج ، وقد يعتقد أن التلميذ سيكتسب هذه القيم إذا نجح في البرنامج لأن البرنامج يخدم أغراضاً ترتبط بهذه القيم . وقد يفترض أن البرنامج يجب أن يوجه إلى تنمية قيم مطلقة مثل مواجهة الحاجات النفسية للفرد . وفي حالات أخرى يفكر واضع البرنامج الدراسي في تنمية النواحي الفكرية لدى الفرد وجعله وحدة متكاملة هي صورة للمواطن الصالح في مجتمع بناته وفي ظروف بذاتها يمر بها هذا المجتمع . أما واضع البرنامج الذي ينظر إلى الأمر نظرة علمية فإنه يفكر في إطار دوافع الأفراد وحاجاتهم التي لو أشبعت ووجهت توجيهاً سليماً يتفق مع الأساس البيولوجي والظروف الثقافية وعمليات التعلم فإنها تنمى

الفرد أكمل ما يسمح به تكوينه وأكمل ما يورده المجتمع وهذا هو الهدف النهائي للتربية .

وسنعرض في الفصل القادم لتقييم المدرس الذى يقوم بتنفيذ البرنامج الدراسى وقياس تقدم التلاميذ نحو أهدافه . وطبعى ، ما دام النصيب الأكبر من الجهد التنفيذى والتقييمى يقوم به المدرس ، أن نطمئن إليه وإلى كل القائمين على النظام التعليمى .

الفصل الحادى عشر

تقييم المدرس

- مميزات المدرس الناجح
- مقتضيات التدريس
- أهم ما يجب أن يتضمنه تخطيطك
- بطاقات يقوم بها المدرس قدرته
- مقياس للتقدير الذاتى لقدرات التدريس
- قائمة مراجعة العوامل الهامة فى شخصية المدرس
- بطاقات محلية لتقييم رجال التربية

الفصل الحادي عشر

مميزات المدرس وتقييمها

مميزات المدرس الناجح :

ماذا يلزمك حتى تصبح مدرساً ناجحاً ؟ يرى البعض أن كل ما تحتاجه هو التفوق في المادة التي تدرسها . بينما يرى البعض الآخر وجوب تحليك بالنظام والدقة . ويصر آخرون على أن بعض الناس ولد كي يصبح مدرساً ؛ وعلى هذا أففى رأيهم ينجح المدرس الذى خلق لهذا العمل . وبين هذه الآراء المتضاربة يحار المرء .

وربما كانت مشاعر الناس فيما يتعلق بالحكم على مميزات المدرس الناجح أعمق وأصدق من مشاعرهم عند الحكم على أى مهنة أخرى . وهذا يرجع بالطبع إلى أن كل الناس يتعلمون سنوات طويلة من أعمارهم حين قد لا يحتاج الفرد منهم إلى المحامى أو الجراح أو المهندس طيلة حياته . ومتوسط عدد المدرسين الذين يتصلون بالتلميذ المتوسط الذى يتخرج من المدرسة الثانوية هو ٦٠ مدرساً ، ومن هنا نلمس مدى كثرة واستمرار الاتصال بين الأفراد والمدرسين كما أن نجاح التلميذ في المدرسة يرتبط مباشرة بإمكانيات وقدرات مدرسية . وفي الوقت الذى يهتم التلميذ فيه بمميزات المدرس الناجح ، يهتم الآباء بتقديم أبنائهم مدرسياً .

والشخص الذى يعد نفسه كى يكون مدرساً يهتم أكثر من غيره بمقتضيات النجاح في التدريس . وهو قد يقوم باسترجاع كل المدرسين الذين اتصل بهم ، ثم يحللهم بعناية وخاصة منهم من يعتبره ناجحاً . وهذه طريقة ناجحة ، ولكن نجاحها يتوقف على صدق وواقعية الأسباب

التي من أجلها يعتبر الفرد مدرساً معيناً ناجحاً : ولهذا وجب عند تحليل المدرسين أن يسجل الفرد هذه الأسباب كتابة ، ثم يطلب من أصدقائه تقييم نفس المدرسين كتابة . وهكذا يستطيع أن يقارن أسبابهم بأسبابه :

وفي هذه المقارنة قد يجد الفرد فروقاً كبيرة في شخصيات المدرسين . فبعضهم هادئ ، وديع ، ودافئ . والبعض الآخر عدواني . وأياً كانت الفروق بينهم ، فالاستنتاج الذي يجب أن ينتهي إليه الفرد هو أنه لا يوجد نمط واحد تكون عليه شخصية المدرسين الناجحين . كما أنه من الحكمة أن نقول إن طريقة واحدة للتدريس تضمن النجاح لا وجود لها . فالمدرسون الناجحون شخصيات متميزة لهم طرق مختلفة في التدريس .

ولا يجب أن يزعم دارس التربية عندما يجد تبايناً كبيراً في خصائص المدرسين الناجحين ؛ ولا كذلك عندما يصرح بعض هؤلاء الناجحين بأنهم لا يتمتعون بكل الخصائص التي تؤكد نجاح أيهم في التدريس ، فصحيح أنهم يعرفون القدرات المطلوبة ، ولكنهم يترددون في تأكيد أن أحداً معيناً سيكون مدرساً ناجحاً جداً بلا شك ، وأن آخراً لن ينجح .

وأول أسباب صعوبة التنبؤ بالنجاح في التدريس أن المدرس يتعامل مع أكثر الأشياء تعقداً — الإنسان التلميذ . والمدرس الناجح يجب أن يكون قادراً على أن يؤثر في الناس بطريقة تحدث تغيرات مرغوب فيها في سلوكهم : ويجب أن يكون قادراً على تشجيع التلميذ في التفكير وفي اتخاذ قرارات ذكية ؛ فبوجه سلوكه بكفاية متزايدة نحو أهداف إيجابية بنائية ديمقراطية : كما أن تعقد وتركيب أوجه التفاعل بين المدرس والتلاميذ يصل إلى حد لا يمكن معه تحليل هذا التفاعل : وقد سبقنا إلى هذا القول هوراس مان Horace Mann سنة ١٨٥٣ عندما قال :

« التدريس أصعب الفنون قاطبة وأعمق العلوم جميعاً . وهو في كماله

المطلق قد يتضمن تعليم معرفة تامة بالوجود كله ، ومعرفة تامة بالطريقة المحددة التي يجب أن يسير عليها كل تطبيق ممكن .

وفي تخطيط العمل بالتعليم ، بعد تحليل المدرسين ، عليك تكشُّف الخصائص المرتبطة عادة بالنجاح في التدريس . ففي أثناء فحصك لكل مدرس ضع في ذهنك دائماً أن كل خاصية فيه متكاملة مع غيرها ، ويجب أن يؤدي تحليلك إلى تكوين إطار عام تستطيع خلاله أن تكون خططا معينة طويلة الأجل لمواجهة هذه المقتضيات requirements أفضل مواجهة ممكنة . وواضح أن كمال خططك سيعتمد على عمق فهمك ووضوح تمييزك أثناء دراسة هذه المقتضيات .

مقتضيات التدريس

قامت مئات الدراسات لتحليل الخصائص الشخصية والمهنية للمدرسين بغرض تعيين السمات التي يحسن أن يتحلَّى بها المدرسون ، وتخصيص نواحي القصور . وسنفحص نتائج بعض هذه الدراسات . وسيفيدك هذا في فهم الأساليب الفنية المتعددة التي استخدمت في جمع البيانات . وبهذه المعرفة قد تكون قادراً مستقبلاً على القيام ببحوث مماثلة . عليك إذن أن تدرس هذه النتائج بعناية ، وأن تحلل نفسك على ضوءها ، وبعد ذلك تستطيع أن تنشئ خططا معينة ومفصلة لمواجهة أكبر عدد من هذه المقتضيات أثناء إعدادك لمهنة التدريس . ولن تقدر على مواجهة كل هذه المقتضيات ، ولكن هذا لا يعجزك فبعضها يتطلب خبرة طويلة ، ودراسة ، وجهداً .

١ - دراسة عمل المدرسين :

قام شارترز Charters ، وويلز Waples سنة ١٩٢٩ ببحث مشكلة تحديد المؤهلات Qualifications الواجب توفرها في المدرس بدراسة شاملة لما كان يقوم به المدرسون وقتذاك . وبعد تبويب الواجبات التي يقوم بها

المدرسون ، درس الباحثان المهارات والقدرات ، والمعارف التي يتطلبها النجاح في القيام بهذه الواجبات . وخلص الباحثان إلى وضع قائمة بالسماات نقتطف منها ما يلي :

- ١ - القدرة على التكيف .
- ٢ - المظهر الشخصى الجذاب .
- ٣ - اتساع الميول - إلى المجتمع المحلى ، والمهنة ، والتلاميذ .
- ٤ - العناية - الدقة ، والتحديد ، والكمال .
- ٥ - الاعتبار ، التقدير ، العطف ، التعاطف ، الذوق ، عدم الأنانية .
- ٦ - التعاون ؛
- ٧ - الوثوق والاتساق .
- ٨ - الحماس - التيقظ ، . . الخ .
- ٩ - الطلاقة ؛
- ١٠ - القوة - الشجاعة ، الحزم ، الاستقلال .
- ١١ - الحكم الجيد - التبصر ، الاستبصار ، بعد النظر ؛
- ١٢ - الصبغة .
- ١٣ - الأمانة .
- ١٤ - المثابرة - الصبر ، التحمل .
- ١٥ - القيادة - المبادأة ، الثقة بالنفس .
- ١٦ - المغناطيسية - الابتهاج ، التودد إليه ، التفاؤل ، خفة الروح ، الاجتماعية ، لطافة الصوت .
- ١٧ - الدقة والنظافة ؛
- ١٨ - تفتح الذهن .
- ١٩ - الابتكار - سعة الخيال ، قوة الحجة ؛
- ٢٠ - التقديمية .

٢١ - دقة المواعيد .

٢٢ - التهذيب - سلامة الذوق ، التواضع ، الخلق ، البساطة :

٢٣ - التحصيل - الذكاء وحب الاستطلاع العقلي .

٢٤ - ضبط النفس - الهدوء ، الوقار ، . : الخ .

٢٥ - الاقتصاد .

٢ - سؤال التلاميذ :

وهذا أسلوب آخر شاع استخدامه ويكون بسؤال التلاميذ عن الصفات المتوفرة في المدرسين الذين يعجبون بهم . وقد وجد أن الأولاد والبنات الذين قضوا وقتاً طويلاً مع مدرس أقدر من غيرهم على إصدار أحكام أكثر صدقاً : وقد قام فرانك هارت Frank W. Hart سنة ١٩٣٤ بالتحدث إلى بعض التلاميذ والتلميذات بعد الأسبوع الأول من بدء الدراسة للتعرف على مدى تقبلهم للمدرسين . فسأل تلاميذ ٦٦ مدرسة في نهاية المرحلة الثانوية من أنماط وأحجام مختلفة عن المدرسين المفضلين لديهم ، وغير المرغوب فيهم ، وأقيمهم عند التلاميذ . طالباً منهم أن يذكروا أعظم المدرسين الذين قابلوهم طوال حياتهم المدرسية . كما طلب منهم أن يذكروا أسباب كل اختيار . وتدل النتائج التي توصل إليها هارت ، وهي مدونة في الجدول رقم (٨) ، على أن أهم هذه الأسباب عند البنين والبنات هي السمات الشخصية أو الإنسانية في المدرس .

(الجدول ٨)

أسباب الإعجاب بالمدرس ، مرتبة بحسب تكرار ذكرها ،
كما قال بها ٣٧٢٥ تلميذاً بالسنة النهائية من المرحلة الثانوية .

رتبتها	تكرار ذكرها	أسباب الإعجاب بالمدرس
١	١٩٥٠	متعاون في العمل المدرسي ، يشرح الدروس والواجبات بوضوح وشمول ، ويستخدم أمثلة في تدريسه .
٢	١٤٢٩	مبتهج ، سعيد ، طيب ، طروب ، مرج ، ابن نكتة .
٣	١٠٢٤	إنسان ، صديق ، رفيق ، « واحد منا » .
٤	٩٣٧	ميال ومتفهم لتلاميذه .
٥	٨٠٥	يجعل العمل مسلياً ، يخلق رغبة في العمل ، يجعل عمل الفصل مسراً .
٦	٧٥٣	حازم ، يضبط الفصل ، أوامره محترمة .
٧	٦٩٥	غير متحيز ، لا يميز أحداً ، (لا محاسيب له) .
٨	٦١٣	غير مبسط ، لا يكثر الشكوى ، لا يؤنب كثيراً ، غير ساخر .
٩	٥٣٨	يعلمنا الموضوع .
١٠	٥٠٤	شخصية ظريفة .
١١	٤٨٥	صبور ، عطوف ، متعاطف .
١٢	٤٧٥	عادل في درجاته وتقديراته ، عادل في امتحاناته واختباراتهِ .
١٣	٣٦٦	عادل ، واسع الحيلة في تعامله مع التلاميذ ، نظامه محكم .
١٤	٣٦٤	يطلب أداء العمل تماماً وبدقة ، ويجعلك تعمل .
١٥	٣٦٢	معتبر لمشاعره تلاميذه في حضور الفصل ، حكيم ، يجعلك تشعر بالراحة .

رتبتها	تكرار ذكرها	أسباب الإعجاب بالمدرس
١٦	٣١٧	يعرف الموضوع ويعرف كيف يشرحه :
١٧	٢٦٧	يحترم آراء التلاميذ ، ويدعوهم إلى المناقشة في الفصل :
١٨	٢١٦	ليس متعالياً ، لا يدعى معرفة كل شيء :
١٩	١٩٩	واجباته معقولة .
٢٠	١٩١	شخص معقول ، غير جامد أو متزمت :
٢٠	١٩١	يتعاون في المشاكل الشخصية للتلاميذ ، بما في ذلك ما يحدث خارج الفصل .
٢٢	١٤٦	ملابسه جذابه ، مناسبة ، منسقة ، ذوقها سليم .
٢٣	١٢١	صغير .
٢٤	١١٠	يعمل بخطة جيدة ، ويعلم ماذا يجب على الفصل أن يعمل :
٢٥	١٠٨	مهتم بحماس بالتعليم .
٢٦	٩٧	يعطي التلاميذ فرصة عادلة للتعويض .
٢٧	٩٦	واجبات المنزل عنده معقولة .
٢٨	٨٦	يتعرف على الفروق الفردية في القدرة .
٢٩	٧٨	صريح .
٢٩	٧٨	شخصية جذابة ، حسن المنظر .
٣١	٧٤	يدرس أكثر من الموضوع .
٣٢	٦٨	مهتم بنشاط المدرسة .
٣٣	٥٣	يدخل إلى الموضوع دون مقدمات .
٣٤	٥٢	حديث .
٣٥	٥٠	حلو ولطيف .
٣٥	٥٠	صوته جميل .
٣٧	٤٢	ذكي .
٣٨	٤١	دقيق ومحب للعمل .

رتبتها	تكرار ذكرها	أسباب الإعجاب بالمدرس
٣٩	٣٦	مخلص .
٤٠	٣٢	يعرف أكثر من الموضوع .
٤١	٣١	نشط .
٤٢	٢٢	يستخدم أحكام سليمة .
٤٣	٢٠	متقن ومهذب .

Witty, P. A. "The Teacher Who Has Helped Me Most" National Education Assoc. J., 1947

وقد وجد هارت أن ٨٠٪ من التلاميذ دللوا على أن المدرس الذي يعجبون به هو أيضاً أكفأ المدرسين وإن كان ٥٠٠ من التلاميذ قد قالوا بأن المدرس الذي يعجبهم أكثر (رأى أحسن مدرس) هو المدرس الذي يحبونه أقل من غيره .

وقد قام ویتی Witty بتحليل ١٢٠٠٠ خطاباً أرسلها تلاميذ التفرقة الثانية عشرة تقابل تلاميذ التوجيهية من كل أنحاء أمريكا إجابة على السؤال : « المدرس الذي ساعدني أكثر » . وكانت سمات المدرسين التي تكررت كثيراً في إجابات التلاميذ (خطاباتهم) ، مرتبة بحسب تكرارها هي :

١ - التعاون والاتجاه الديمقراطي .

٢ - العطف .

٣ - اعتبار التلاميذ ، والصبر عليهم .

٤ - تنوع الميول .

٥ - المظهر العام والطريقة اللطيفة .

٦ - العدل وعدم التحيز .

٧ - المرح .

٨ - الاستعداد الطيب والسلوك المتسق :

٩ - الاهتمام بمشاكل التلاميذ .

١٠ - المرونة .

١١ - استخدام المدح والتقريظ .

١٢ - تفوق وكفاية في تدريس موضوع معين .

وقد لخص ويني استنتاجاته بقوله إن « هؤلاء الأولاد والبنات يظهرون شاكرين للمدرسة بالدرجة التي توفر لهم بها الأمن ، والنجاح الفردي ، والخبرة المشتركة ، وفرص التوافق الشخصي والاجتماعي . وهذه بالتحديد هي العوامل التي تضمن تعلماً جيداً » .

وقد قام ويني بثلاث دراسات تخللت أمريكا كلها محاولاً أن يتحقق من نتائجها وأن يوسعها . وقد وجد أن السمات الاثني عشر المذكورة سابقاً كانت تتردد دائماً بالرغم من أن ترتيبها كان يختلف باختلاف الأعوام .

(جدول ٩)

الأسباب التي من أجلها يقل الإعجاب بالمدرس ، مرتبة بحسب
تكرار ذكرها ، كما وردت على لسان ٣٧٢٥ تلميذاً بالسنة النهائية
في التعليم الثانوى

رتبتها	تكرار ورودها	أسباب قلة الإعجاب
١	١٧٠٨	مُسبِط ، كثير الشكوى ، كثير التأنيب ، ساخر ، سريع الغضب غير متعاون في العمل المدرسى ، لا يفسر الدروس والواجبات ،
٢	١٠٢٥	غير واضح ، العمل غير مُخطط
٣	٨٥٩	متحيز له محاسيب أو تلاميذ مفضلون ، ويهاجم تلاميذ معينين
٤	٧٧٥	متعالى ، متكبر ، متعطرس ، لا يعرفك خارج الفصل وضيع ، غير معقول ، جامد ، لا يحتمل ، غير مهذب ، قاسى جداً ، يجعل الحياة تعيسة .
٥	٦٥٢	غير عادل في درجاته وتقديراته ، غير عادل في اختبارات وامتحاناته
٦	٦١٤	غير معتبر لمشاعر تلاميذه يهزأ بهم في حضور الفصل ، يجعل التلاميذ خائفين ، سريع الغضب
٧	٥٥١	غير مهتم بتلاميذه ولا يفهمهم
٨	٤٢٢	واجباته غير معقولة
٩	٣٥٠	لا نظام له ، لا يضبط الفصل ، لا يجبر أحداً على احترامه لا يدخل إلى الموضوع مباشرة ، يذكر كثيراً من الشؤون الشخصية غير المتعلقة بالموضوع ، كثير الكلام
١٠	٣١٣	لم نتعلم ما كان يجب
١١	٣٠١	كثير
١٢	٢٧٥	كثير ، غبي ، غير مسلي
١٣	٢٧٥	

رتبتها	تكرار ورودها	أسباب قلة الإعجاب
١٤	٢٢٤	رجعى ، عجوز
١٥	٢٠٣	غير عادل في معاملته مع التلاميذ
١٦	١٩٣	يعرف الموضوع ولكنه لا يستطيع توضيحه
١٧	١٩٠	لا يحافظ على المستويات ، مهمل في عمله
١٨	١٨٣	قاسى ، صعب جداً ، لا يترك فرصة للتعويض
١٩	١٧٠	لا يعرف الموضوع
٢٠	١٣٣	لا يحترم أحكام التلاميذ وآراءهم
٢١	١٢٢	كثير التغير ، غير متسق مع نفسه ، غير ثابت
٢٢	١١٥	كسول ، غير مبال للتدريس
٢٣	٩٨	ليس صديقاً ، ليس رفيقاً
٢٤	٩٥	يبدى تحيزاً للأولاد أو البنات
٢٥	٩٢	ملابسه غير جذابة أو ذوقها ردىء
٢٦	٨٥	ضعيف الشخصية
٢٧	٧٥	غير مخلص
٢٨	٦٥	شخصية غير جذابة
٢٩	٦٤	لا يعترف بفروق فردية بين التلاميذ
٣٠	٦٣	صوته غير سار .

Hart, F.W. Teachers and Teaching The Macmillan Co. N.Y., 1984.

وفي دراسة أخرى سئل ٣٩٠٥ تلميذ يمثلون المدارس الثانوية في مقاطعة إنديانا بأمريكا أن يذكروا المميزات المتوفرة في المدرس المفضل لدى كل . وقد أجاب أكثر من ٩٠ ٪ بما يأتي : القدرة على شرح الدرس بوضوح ، عدالة الدرجات ، الاستعداد لتقديم مساعدة إضافية لمن

يحتاجها ، خفة الروح ، القدرة على مسايرة المدرسين الآخرين . ثم أنت .
 خصائص أخرى تلى السابقة في الترتيب من مثل : معرفة الموضوع
 المدروس ، الاستعداد الطيب ، السعادة البادية مع ميل إلى الابتسام كثيراً ،
 وصوت الكلام السار .

كما علم من ٦٣٪ إلى ٧٨٪ على الخصائص التالية : فهم مشاكل
 الطلبة ، عدم التحيز ، جاذبية الملابس ، أن يكون أخا حقاً داخل
 وخارج المدرسة ، الاهتمام الشخصي بالتلميذ ، أن يشبه رجل الأعمال
 في الفصل وأن يصر على الدقة ، الإصرار على العمل بجد . وأجاب
 حوالي ٤٤٪ من التلاميذ مختارين الفقرات التالية : ساعدنى في الخارج
 عندما كنت في مأزق ، قال كلاماً لطيفاً عن عملى ، كان حسن
 المظهر ، لم يخرجنا في حضور تلاميذ آخرين ، لم يغضب منى أبداً ،
 ولم يجبرنى على العمل الشاق . بينما اختار ١٧٪ من التلاميذ باقى الفقرات
 التى سبق أن ذكرناها .

سوف تلاحظ في دراستنا ومقارنتنا لدراسات هارت ، ويقى ؛
 ريتسى Richey ؛ فوكس Fox اتفاقاً عاماً فيما يتعلق بخصائص المدرسين
 الذين يرغبهم التلاميذ . وقد ترغب في استعراض هذه القوائم وتحليل
 نفسك ل ترى ما إذا كنت تملك أو لا تملك المميزات المذكورة .

٣ - تحليل تواريخ حياة المدرسين المشهورين :

طريقة أخرى لكسب الفهم بخصائص التدريس هى الفهم الواضح
 لكثير من التواريخ الشخصية والكتب التى ألقت عن حياة بعض المدرسين
 المشهورين . ولكن هذه التواريخ تعنى خاصة بالصفات وبالعوامل
 الإنسانية وإن كانت هذه الصفات وتلك العوامل عظيمة الدلالة فى حياة
 التلاميذ والتلميذات .

٤ - دراسة وظائف التربية :

المدرس الناجح هو الذى يتمثل أهداف التربية ويكون فعالاً فى تحقيقها وهو الذى يقيس الخبرة التى مرَّ بها فى مدرسته بما يجب حسب أهداف التربية حتى يحدد إلى أى مدى حصل هو بعض هذه الأهداف .

٥ - قوائم المقتضيات لرجال التربية :

قام كثير من رجال التربية أفراداً وجماعات بوضع قوائم لمقتضيات التدريس مبنية على خبراتهم كمدرسين وملاحظين وكعاملين مع مدرسين آخرين ، وهذه القوائم تصاغ غالباً فى مقاييس للتقدير Rating Scales تستخدم فى تقييم عمل المدرس فى المدرسة وتقييم طلاب كلية التربية . وقد تكون راجعاً فى إستعراض مقياس ربما أتى الوقت الذى يقدر به عملك ، ويحسن هنا أن تدرس الخصائص المتضمنة . وبفهم هذه الصفات جيداً ستكون قادراً على تخطيط خبرات خاصة تحاول أن تكتسبها أثناء إعدادك لمهنة التدريس :

قام مجلس كلية التربية بكاليفورنيا California Council on Teacher Education بنشر قائمة لخصائص العاملين بالتدريس حتى تتضح وظائفهم لطلبة كلية التربية ولإعطاء تعليقات لتقييم طلبة كلية التربية ، وهذه هى القائمة :

قائمة كاليفورنيا للقدرة على التدريس

المدرس القادر :

١ - قادر على تعليم التلاميذ :

(١) يستخدم الأسس النفسية فى التعليم .

- ١ - يستخدم الدفع بطريقة فعالة ومستمرة .
- (أ) يتعرف على ميول التلاميذ ويستفيد منها ، وكذا قدراتهم وحاجاتهم :
- (ب) يفيد من خبرات التلاميذ وينظر إلى مواقف الحياة وإلى الاهتمامات التي يثيرها الموضوع .
- ٢ - يوفر خبرات تعلم متنوعة .
- ٣ - يستخدم طرقا متنوعة في التدريس ، مثل المناقشة والعرض ، : : الخ بطريقة فعالة .
- ٤ - يضع خططاً متعاوناً مع التلاميذ .
- (ب) يفيد من أسس نمو الطفل والنمو في مواقف التعلم
- ١ - يوفر أوجه نشاط متفرقة وواجبات لمواجهة حاجات وقدرات التلاميذ .
- ٢ - يعرف المكانة الصحية (العقلية والجسمية) لتلاميذه ويكيف أوجه النشاط مع حاجاتهم .
- (ح) يحافظ في الفصل على جو يساعد التعلم ويتميز بروح الانتران بين الحرية والأمن .
- ١ - يُبقى على موقف عمل فعّال .
- ٢ - يساعد التلاميذ لضطراباً على القيام بالقيادة والمسئولية .
- ٣ - يوفر للتلاميذ فرصاً للتعاون وممارسة القيادة في أوجه نشاط الجماعات الكبيرة والصغيرة .
- ٤ - يوفر فرصة للتعبير عن التفكير الناقد المستقل مؤكداً على حرية التعبير وتفتح الذهن .
- (د) يضع الخطط بطريقة فعالة .

١- يساعد الطلبة على تحديد الأهداف الجديرة بالتحقيق في أوجه النشاط الواسعة ، واليومية داخل الفصل ، وأوجه النشاط الخاصة بالفصل .

٢- ينظم تدريسه جيداً بأن يختار بحكمة خبرات تعليمية ، ومحتوى الدراسة ومواد التعليم .

٣- يختار ويستخدم مواد كثيرة التنوع للتعليم (مثل الكتب والكتيبات ، والأشرطة السينمائية ، والمجلات العلمية ، والإذاعة ، والصور والسجلات . . الخ) .

٤- يفيد من المصادر الموجودة في مكتبة المدرسة ومكتبة المجتمع المحلي .

(هـ) يستخدم عمليات التدريس المتنوعة

١- يستخدم عمليات التدريس (مثل التقارير الجماعية ، والمناقشة ، والتخطيط مع التلاميذ) المصححة لتحقيق الأهداف المرجوة من التعليم .

٢- يعتمد بطريقة فعالة على مساهمة الطلبة في أوجه نشاط الفصل •

٣- ينمى مهارات المذاكرة لدى التلاميذ :

٤- يستثير أوجه النشاط الخلاقية عند التلاميذ :

٥- يساعد الطلبة على تقييم تحصيلهم الخاص :

(و) يستخدم العمليات التشخيصية والعلاجية بطريقة فعالة :

١- يألّف الاختبارات التشخيصية العامة في ميدانه الخاص والميادين القريبة .

٢- ينشئ ؛ ويجرى ، ويفسر الاختبارات التشخيصية :

٣- يستخدم العمليات التشخيصية الأخرى الملائمة :

٤- يخطط ويفيد من العمليات العلاجية .

١- (ز) يستخدم عمليات صالحة لتقييم تحصيل الطلبة .

١- يستخدم عمليات التقييم غير الرسمية (سجل الوقائع ، استبيان المقابلة الشخصية) لجمع وتفسير البيانات المطلوبة .

٢- يستخدم الإختبارات التحصيلية المقننة .

(أ) يألف أكثر الإختبارات شيوعاً في ميدانه .

(ب) يختار ، ويجرى ، ويفسر نتائج الإختبارات ويستخدمها في التخطيط .

٣- يستخدم اختبارات المدرس .

(أ) ينشئ اختبارات مناسبة بمهارة .

(ب) يفسر النتائج ويستخدمها في التخطيط

٤- يحتفظ بسجلات دقيقة وكافية ، مثل ، الحالات المدروسة ، والسجلات المجمعة .

٥- يعمل تقارير وافية للطلبة والآباء فيما يتعلق بتقدم الطلبة في نموهم .

(ح) يدير الفصل بطريقة فعالة :

١- يخطط نظاماً جيداً ثابتاً لمعالجة الخامات ، والأجهزة والمواد .

٢- يستخدم وقته ووقت التلاميذ بفعالية .

٣- يراقب الراحة الجسمية لطلبته فيما يتعلق مثلاً بالحرارة ، والإضاءة ، والتهوية ، والجلسة .

٢- يرشد ويوجه الطلبة بحكمة :

(أ) يستخدم الأسس النفسية العميقة المتعلقة بتقدم ونمو الأطفال في توجيه الأفراد والجماعات .

١- يحافظ على الموضوعية عند معالجة السلوك العدواني أو غير السوي .

٢- يكون متعاطفا وحساسا في مشاكل التلاميذ الشخصية والاجتماعية وأيضاً في حاجاتهم التعليمية .

٣- يعمل تعديلات في المنهج الدراسي وفي الشروط الأخرى على ضوء حاجات التلاميذ .

٤- يضمن علاقة طيبة كافية مع الطلبة بحيث يأتون متطوعين للاستشارة .

(ب) يحافظ على علاقات فعالة مع الآباء .

١- يفسر حاجات ، وقدرات ، وميول ، ومشاكل التلاميذ لآبائهم .

٢- يحصل على تعاون الآباء في مساعدة الطلبة على مشاكلهم .

(ج) يجمع ويستخدم البيانات الجوهرية في الاستشارة .

١- يطبق اختبارات الاستعداد والذكاء .

٢- يفسر نتائج مثل هذه الاختبارات .

٣- يستخدم النتائج المتجمعة في الاستشارة مع تلاميذه .

٤- يجعل البحث مناسباً للتوجيه .

(د) يستخدم عمليات الاستشارة الملائمة .

(هـ) يحافظ على علاقات مناسبة مع أخصائيي التوجيه ، عارفاً دورهم وحدود مهارته وقدرته .

٣- يساعد التلاميذ على فهم وتقدير الميراث الثقافي :

(أ) ينظم الفصل لحياة ديمقراطية فعالة

(ب) يوجه الأفراد والجماعات إلى التطبيقات الجوهرية في شؤون الحياة أثناء عمليات التعليم في الفصل .

١ - يستخدم موضوعات ميادين معينة لإنماء فهم المشاكل الاجتماعية ، والاقتصادية ، والسياسية .

٢ - ينمى لدى التلاميذ فهما للدلالة موضوعات متعددة تدخل فى المواد الدراسية المقررة .

(>) يرسم على أساس حصيلته من الخبرة ليحقق النمو الثقافى للأفراد والجماعات .

٢ - يساعد الطلبة على أن يعرفوا ويطبقوا على حياتهم اليومية الأسس الديمقراطية التى تأصلت عميقاً فى نمونا التاريخى .

٤ - يساهم بفعالية فى أوجه نشاط المدرسة :

(١) يخطط تعاونيا وسائل تحصيل الأهداف التربوية .

١ - يشارك بفعالية فى مراجعة المنهج الدراسى ويكون قادرا على تقييم التقدم نحو الوصول إلى الأهداف التربوية .

(١) يحدد الأهداف بوضوح .

(ب) يجمع البيانات بكفاية ويشق منها نتائج مناسبة .

(>) يستخدم عمليات علاجية مناسبة .

(د) يبدى مرونة فى تعديل خططه وخطواته حتى تتلاءم مع خطط وخطوات المدرسة كلها .

(ب) يقوم بنصيبه من المسؤولية فى نشاط المدرسة .

١ - ينفذ بفعالية المسؤولية التنفيذية الموكولة إليه .

٢ - يساهم فى تخطيط وتنفيذ أوجه النشاط الخارجة عن المنهج الدراسى .

(>) يبقى على العلاقات الشخصية المتناسقة مع زملائه .

٥ - يساعد على تكوين علاقات طيبة بين المدرسة وبقية المجتمع المحلي :

(١) يدرب نفسه على المصادر المتوفرة في المجتمع ويستخدمها في أوجه النشاط داخل الفصل .

(ب) يحصل على تأييد الآباء لنواحي نشاط المدرسة .

(ج) يساعد في تحديد وحل مشاكل المجتمع المحلي :

١ - يساعد في تحديد مشاكل المجتمع المحلي وفي تنمية وعي التلاميذ والآباء بهذه المشاكل .

٢ - يستعين بالمصادر المتوفرة والملائمة لدى المدرسة في مواجهة مشاكل المجتمع المحلي .

(د) يشارك في شؤون المجتمع المحلي ومشروعاته .

(هـ) يلاحظ أخلاقيات المهنة في مناقشة المشاكل الخاصة بالمدرسة وخاصة مع رجل الشارع .

٦ - يعمل في مستوى مهني :

(١) يقدم أدلة على الأهمية الاجتماعية لمهنة الآباء ، والطلبة ، وغيرهم من أعضاء المهنة .

(ب) يسير بحسب ميثاق أخلاقي مهني .

(ج) يساهم في المهنة بالعضوية في المنظمات والمشاركة في أوجه نشاطها .

(د) يتحمل مسؤولية نمو المهني الخاص بتخطيط برنامج مناسب للتقدم المهني .

١ - يستمر في الدراسة المهنية عن طريق البرامج ، والمحاضرات ،

والمعاهد ، والقراءة المهنية وغير ذلك من نواحي النشاط .

(هـ) يساعد في الإشراف على طلبة كليات التربية وفي توجيه وإرشاد

المدرسين المبتدئين .

وهناك الكثير من القوائم عن الخصائص التي يجب توفرها في المدرس وضعها رجال التربية أفرادا وجماعات .

أهم ما يجب أن يتضمنه تخطيطك

أثناء قراءتك للقوائم السابقة ، قد يدهشك عدد مقتضيات التدريس . والمدرسة في يومنا هذا يلزمها عدد من المدرسين أكبر منه في الماضي . ويمكنك أن تلمس هذا في التوسع الكبير في نظمنا التعليمية . ففي الماضي كانت المدرسة في الريف تقتصر أساسا على تعليم الأولاد والبنات المهارات الأساسية في القراءة ، والكتابة ، والحساب . وكانت الأسرة ، والجامع والكنيسة ، والمجتمع المحلي يقدمون أغلب نواحي التعليم الأخرى الأساسية في حياة المواطن الصالح . وقد كانت الحياة بسيطة نسبياً . أما اليوم فالحياة أكثر تعقداً ، وهذا يرجع على وجه خاص إلى التقدم التكنولوجي الضخم الذي حصلناه . وبالمثل حدثت تغيرات جوهرية كبرى في حياتنا الاجتماعية ، والاقتصادية ، والسياسية . ووظيفة المدرسة يجب أن تتغير وأن تتوسع حتى تواجه حاجات شباب اليوم وشباب الغد . وهذه التغيرات تؤدي بدورها إلى زيادة وتوسع وحدائق مقتضيات مهنة التدريس .

فالمدرس اليوم يحتاج إلى فهم الجو الثقافي الذي تتم فيه كل عمليات التربية والتعليم . وهو يحتاج أن يرى جوهرية التصنيع والتحضر ، والتعاونية ، والتخطيط الاقتصادي ، والهوة الثقافية Cultural Lag ، والتفكك الثقافي ، ونمو الاتجاه العلمي نحو المشاكل ، إلى آخر كل هذه الخصائص المميزة لمجتمعنا المعاصر ما دامت تؤثر على حياة الناس . وعليه أن يكون قادراً على فهم المشاكل الحديثة في ضوء الظروف التاريخية ، وأن يتعرف على الطبيعة الراهنة لكثير من الشؤون التربوية ،

وأن يستخدم مصادر الماضي عندما يواجه مشاكل الحاضر ، وبناء على ذلك فعليه أن يميز بين الحلول التقليدية وبين الحلول التي آلت إلى الزوال . فهو يحتاج إلى فهم العمليات التي بها تتدخل القيم في توجيه سلوكنا وأحكامنا .

على المدرس أن يبين فهماً أساسياً يتضح خلال اتجاهاته ، ومهاراته ، وعاداته ، اتجاهات يتسم بالاشتراكية الديمقراطية التعاونية بالطريقة التي بها تؤثر القيم في ثقافتنا ، وبالطريقة التي بها ينمي كل فرد نفسه على أكمل وجه ، وبالطريقة التي نحل بها مشاكلنا ، وبالطريقة التي ننظم بها أنفسنا بحسب تغير حياتنا المشتركة . يجب أن نكون على قدر كبير من الفهم الناقد لوظيفة المجتمعات المحلية والهيئات الموجهة للنمو الإنساني ، وأن نفهم وظيفة المدرسة باعتبارها هيئة عن طريقها يبحث المجتمع عن إمداد وتوسع بغرض تقديم الفرد . يجب أن ينظر المدرس إلى المدرسة باعتبارها نظاماً عهد إليه بأكبر مسئولية في تحديد اتجاه وطابع نحو المجتمع . فهو يتعرف على الطريقة التي يحسن أن يتبعها في تدريس الطلبة والطالبات على أن يحددها إدراكه للوظيفة الحقيقية التي تؤدها التربية .

ألقى البحث التربوي ، خلال الأعوام العديدة الماضية ، الكثير من الأضواء الكاشفة على السلوك الإنساني وعمليات التعلم . فلم تعد التربية تلقيناً وتكديساً للمواد في أذهان التلاميذ حتى ينجحوا في الامتحانات . ومدرس اليوم يحتاج إلى فهم شامل لطبيعة النمو الفردي في كل جوانبه المتنوعة حتى يستطيع أن يتعرف بل وأن يسبق الحاجات المتزايدة لتلاميذه وأن يوفر الظروف التي تسمح بأفضل نمو . يجب أن يكون ملماً بالطرق التي يفضلها الناس وبالطرق التي يختلفون عليها : فيفهم كيف يتفق الفرد مع الآخرين وإن كان فريداً متميزاً أثناء نضجه . وهكذا يفهم فهماً شاملاً عميقاً طبيعة التعلم الإنساني ، خلال عادات التعلم ، واتجاهات ، ومهارات ، وقدرات

المتعلم ذلك في ضوء دوره كموجه فعال لعملية التعلم . وعلى هذا يجب أن تتوفر فيه مهارة الخبر بتوجيه تربية الآخرين نحو الأهداف المرجوة .

بطاقات يقوم بها المدرس قدرته

قد يعينك استعراض عدد من بطاقات التقويم المصممة بحيث يقيم الفرد نفسه بها . ونحن نعرض لك هنا بطاقتين . ويحسن بك أن تضع علامات على الخصائص المتوفرة فيك . ونأمل أن تكون آمينا مع نفسك فيها ، حتى تتضح لك نقاط القوة والضعف في نفسك .

وقائمة التقدير الذاتي Self Appraisal صممت في مدرسة التربية التابعة لجامعة أنديانا لتستخدم في منهج تمهيدى في التربية . وهى مبنية على قائمة من ١٥ خاصية يجب أن تتوفر في القائم بالتدريس ، أعدها أساتذة كلية التربية بجامعة ولاية أوهايو Ohio .

وتدلك القائمة التالية على أهمية تحليلك بسمات شخصية تساعدك على التأثير الصحيح المرغوب فيه على تلاميذك . فالتلاميذ بطبيعتهم ميالون إلى المحاكاة ، وعن طريق هذه المحاكاة يمكنك أن تكون فعالا .

وعليك أثناء استخدام البطاقة أن تكون دقيقا مسائلا نفسك حتى تعرف ما أنت عليه من قوة تساعدك في التدريس وما أنت عليه من ضعف قد يحول دونك والنجاح أحيانا . ولكن هذا لا يعنى أن كل الخصائص يجب أن تتوفر فيك فهذا مستحيل . عليك أن تتوقع مقدما أن بعض سمات الشخصية غير المرغوب فيها قد تغطى على البعض الآخر المرغوب فيه .

ولكى نساعدك أكثر من هذا في دراستك للتربية ، نقدم إليك قائمة للمراجعة تتضمن أهم عوامل الشخصية المؤثرة على نجاحك كمدرس . ويمكنك أن تلجأ إلى زملائك ، وخاصة منهم من يعد نفسه لمهنة التدريس فيقومون بتقييمك على نفس الأسس ثم تقارن تقييمك لنفسك بتقييمهم لك فتتضح لك صورة عن نفسك أكثر صدقا .

مقياس التقدير - الذاتى لقدرات التدريس

يحتاج المدرسون لكي ينجحوا ، أن يندو قدرات في ١٥ مجال على الأقل متضمنين في مقياس التقدير - الذاتى لقدرات التدريس . وراضح أن هذه القدرات ترتبط معاً تبادلياً . والخصائص المدونة تحت كل نوع من الفهم تفيد كنقط مرجعية لتسهيل عملية التقييم . فأنت تقييم نفسك على أساس قائمة القدرات بوضع علامة (س) في المكان الملائم . كن أميناً وعادلاً مع نفسك في تعرفك على نقاط القوة والضعف الأساسية حتى تقدم لك تقديراتك معلومات صحيحة عن مؤهلاتك الحالية .

القدرة	بني	أقل من المتوسط	متوسطة	فوق المتوسط	ممتازة
١ - تفهم فرص الخدمة وتقدر على الخدمة »					
١ - يعتمد عليك : دقيق ؛ مستعد لتقبل المسؤولية ؛ تقوم بالعمل عادة حتى تكمله بنجاح .					
٢ - ميل إلى الناس : تستمتع بتواجدك معهم ؛ وتعاون في النشاط الجماعي ؛ تحب الأطفال وتعمل معهم بفعالية .					
٣ - مبادئ : غالباً تقوم بالقيادة .					
٤ - الحساسية الإجتماعية : تضع المصلحة العامة فوق الرغبة الشخصية ؛ عارف بمشاكل وحاجات الآخرين .					
٥ - المثابرة : لا يسهل تثبيطك أو إغرائك بالتخلي .					
٦ - التقبل الإجتماعي : متقبل للناس من كلا الجنسين ، والمسنين والشباب ؛ والآخرين من سنك .					
٧ - شخصية متزنة : لست متزمتاً للعلم أو للتبريج أو للإجتماعيات ولكن متزناً في الصفات الطيبة .					
٨ - السلوك : متزن ؛ ومتحرر من القيود والتكلف .					
٩ - المظهر : نظيف ، ومنظم ، وحسن الملبس .					
١٠ - مستعد للتعاون مع الآخرين .					

القدرة	ديما	أقل من المتوسط	متوسطة	فوق المتوسط	ممتازة
١١ - مستعد وراغب في العمل مع من يختلفون عنك في الإطار الاجتماعي ، أو الاقتصادي ، أو الجنسي . ٢ - « تفهم طبيعة تخطيط واستخدام العمليات في الحياة اليومية والتعليم » ١ - تعبر عن القيم بوضوح . ٢ - تصيغ الأهداف ؛ وتعرف ما تريد أن تبحث عنه . ٣ - ترى الهوة بين ما تريده وما تجده . ٤ - تنفذ وتحقق الفروض أو البرامج الإجرائية لتحصل على ما تريد . ٥ - تستطلع المصادر التي عليك أن تعالجها . ٦ - تختار الفروض والمصادر . ٧ - تنظم الفروض ، والمصادر ، وأوجه النشاط الإجرائي لحل المشكلة . ٨ - تتنبأ بالمشاكل والصعوبات التي تقف في طريقك . ٩ - ترى الحلول البديلة خلال قيامك بحل المشكلة . ١٠ - تحس بالعلاقة بين كل جانب من الجوانب سابقة الذكر والجانب الآخر . ٣ - « تفهم طبيعة تقييم واستخدام العمليات في الحياة اليومية وفي التدريس » ١ - توضح معنى الاشتراكية الديمقراطية التعاونية . ٢ - تصيغ الأهداف واضحة ؛ وتفكر بعناية في أهداف الحياة ؛ وتضع اعتبارا للعقائد والقيم . ٣ - تتعرف على المشاكل وتستخدم العمليات التخطيطية في حلها . ٤ - « تفهم أسس الحياة الصحيحة ، عقليا وجسميا ، وتطبق هذه الأسس على حياتك الشخصية وعلى التدريس » ١ - مبرن انفعاليا ومتحرر من العاهات الحسية . ٢ - لك ميل تروبيجي عقلي وجسمي معا . ٣ - لك ميل كثيرة واتصالات شخصية مع أناس متنوعين ، شباب ومسنين ، ذكورا وإناثا . ٤ - ماهر في كثير من الميول التروبيجية . ٥ - تستمتع بالحياة .					

منازلة	فوق المتوسط	متوسطة	أقل من المتوسط	دينيا	القدرة
					٥ - « تفهم نمو الطفل وتعلمه وقدرته على تطبيق أسسها على التدريس » ١ - لديك حب أصيل للأطفال والشباب ورغبة حقيقية في العمل معهم . ٢ - كانت لك اتصالات كثيرة مع الأولاد والبنات . ٣ - لك معرفة علمية عريضة بعمليات التعلم . ٤ - لك معرفة كافية بالمشاكل الإنفعالية الأساسية ، والجسمية ، والمشاكل الأخرى الخاصة بنمو الأطفال ٦ - « تفهم الفروق الفردية والقدرة على التعامل الفعال مع هذه الفروق في التدريس » ١ - قادر على التعرف على الكثير من الطرق التي يختلف الناس فيها وقادر على التعامل الفعال مع الأنواع المختلفة . ٢ - تفهم الطرق الفنية في التعامل مع الفروق الفردية وقادر على تطبيقها في التدريس . ٧ - « تفهم عمليات الإرشاد الوظيفي وقادر على تطبيق هذه الأسس في التدريس » ١ - تفهم الطرق المتبعة في التوجيه ، والاختيار ، والإرشاد . ٢ - تفهم أهمية التوجيه في المدارس . ٣ - تعرف متى وكيف توفر التوجيه للآخرين . ٨ - « تفهم مواد وطرق التعليم والقدرة على استخدامها » ١ - تفهم سيكولوجية التعلم وفائدتها في التدريس . ٢ - تفهم ولديك المهارة في طرق التعليم . ٩ - « تفهم مكانة أوجه النشاط الخارجية عن المنهج الدراسي في المدرسة مع القدرة على توجيهها » ١ - كانت لك مشاركة في نواحي النشاط الخارجية عن المنهج المدرسي في المدرسة الثانوية ٢ - المشاركة في كثير من أوجه النشاط الخارجية عن المنهج الدراسي في الكلية . ٣ - فهم قيمة أوجه النشاط الخارجية عن المنهج الدراسي . ٤ - القدرة على توجيه واحد أو أكثر من أوجه النشاط الخارجية عن المنهج الدراسي .

القدرة	دنيا	أقل من المتوسط	متوسطة	فوق المتوسط	ممتازة
١٠ - « تفهم مكانة التقييم والقدرة على استخدام عمليات التقييم » . ١ - ماهر في تفسير نتائج اختبارات التحصيل ، والميول ، والاستعدادات ، والاختبارات العقلية والأنواع الأخرى من الاختبارات . ٢ - القدرة على تقييم نمو الشخصية على ضوء الانجازات وتغير الطابع . ١١ - « تفهم المجتمع المحلي والقدرة على ربط المدرسة بالمجتمع المحلي في التدريس » . ١ - تفهم أجواء المجتمع المحلي ، وميول ، وحاجات التلاميذ . ٢ - لديك المهارة في الطرق الفنية لزيادة فهمك لمصادر المجتمع المحلي وحاجاته . ٣ - المشاركة الإيجابية في أوجه نشاط المجتمع المحلي . ٤ - عضو من أعضاء المدرسة ومهنة التدريس فعال في المجتمع المحلي . ١٢ - « تفهم مكانة الوسائط المختلفة في التعبير الشفوي والكتابي والقدرة على استخدامها في التدريس » ١ - تلائم الأفكار مع الموقف ؛ وتعبر عن نفسه بطريقة مهذبة ، ورقيقة ، ومختصرة ، وكافية ، ومنطقية . ٢ - تعدد العرض بحسب الموقف ؛ وتستخدم طريقة مناسبة للتوافق مع مناسبة الحديث . ٣ - تنظم أفكارك وتعبر عنها بوضوح وفعالية ؛ وتتكلم أو يكتب بأسلوب طليق ، سلس ، بطريقة موحدة . ٤ - تستخدم قواعد اللغة المقبولة وبناء الجملة في الكلام والكتابة ٥ - تخرج الكلمات بطريقة صحيحة ومتميزة . ٦ - تستخدم صوتك بفعالية ؛ وله مدى كاف ، متحرر من الملل ، وله نغمة سارة ، ورنين وقوة كافيين . ٧ - يعرض الأفكار كتابة بتنسيق ، ووضوح ، وتنظيم منطقي مقبول . ٨ - صحيح الهجاء .					

ممتازة	فوق المتوسط	متوسطة	أقل من المتوسط	دنى	القدرة
					١٣ - « تفهم مكانة الوسائط الفنية المختلفة في التدريس والقدرة على استخدامها » ١ - القدرة على استخدام وإنشاء رسوم بسيطة ، وأشكال بيانية ، ورسومات للأغراض التوضيحية . ٢ - تفهم الفنون الجميلة والتطبيقية باعتبارها وسائط تدريس فعالة ١٤ - « تفهم مكانة التحصيل في العملية التعليمية والقدرة على جعله وظيفيا في التدريس » ١ - متمكن من ميدان تخصصه . ٢ - له معرفة واسعة بمبادئ المعرفة العامة القريبة . ٣ - على فهم بالأحداث الجارية في العالم . ٤ - ترتيبه متوسط أو متقدم في المدرسة الثانوية وفي الكلية . ٥ - نمت اهتماما حقيقيا بالتحصيل - وليس فقط بالدرجات . ١٥ - « تفهم فلسفة التربية والقدرة على جعلها وظيفية في التدريس » ١ - الاعتقاد في إمكانية التحسن عن طريق التربية . ٢ - توجهه فلسفة تربوية تهدف إلى إعداد التلاميذ حتى يصبحوا مواطنين صالحين . ٣ - تمارس الاشتراكية الديمقراطية التعاونية متسقاً مع نفسه .

Based on a list of 15 competencies for teaching in Student Planning in College, College of Education, Ohio State University, Columbus, Ohio, 1940.

قائمة مراجعة للعوامل الهامة شخصية المدرس

تعليمات :

اعتبر كلا من الأقسام العشرة في هذه القائمة منفصلة . واقرأ كل عبارة تحت العنوان الكبير بعناية وعلم على أى صفة أو سمة تفتقر إليها بوضوح الشخصية المقاسة . وقبل أن تنتقل إلى القسم الثانى راجع أى الفقرات التى خططت تحتها ثم ضع علامة (✓) في العمود الذى يعبر عن رأيك العام في الشخص فيما يتعلق بالجانب المقاس من الشخصية .

متابعة	تعليمات	توضيح	رقم الترتيب	المستويات المقترحة للسمة أو الصفة
				١ - <u>الاستقرار الانفعالى والصحة النفسية :</u> متحرر من المخاوف ، والمزعجات ، ومن مشاعر الذل ، والقلق عن الأشياء التافهة ؛ ومستعد لعمل تقديرات واقعية عن إمكانياته العقلية ؛ وغير حساس جدا للنقد ؛ وذو عقل غنى بالإمكانيات الكلية في الرضى عن النفس ؛ لا يسهل تهيجه ؛ متحرر من الحجل الزائد ومن التزامات المزاجية وأحلام اليقظة ؛ يحسن مواجهة المواقف غير المتوقعة ؛ يتكيف عن استعداد مع المواقف المتغيرة ؛ يمارس ضبط النفس ؛ متحرر من عقد الدونية والتفوق ؛ له ضبط على حالته الانفعالية ، لا يتقلب فجأة من حالة متطرفة إلى أخرى ؛ يواجه خيبة الأمل في الحياة بشجاعة تامة . ٢ - <u>المظهر الشخصى :</u> ملابسة تلائم المناسبة ؛ متيقظ ومترن ؛ يدل مظهره على الاتزان النفسى ؛ له ذوق جيد في اختيار ملابسه ؛ يعطى انطباعا بأنه مهذب ومثقف ؛ تركيبات الألوان متناسقة ؛ كمي الملابس وتنظيفها ؛ تبدو عليه صحة جيدة .

جزء مقتبس مع التعديل من :

Raleigh Schorling and Howard T. Batchelder, Students Teaching in Secondary Schools, Mc Graw-Hill Book Co. Inc., N.Y., 1956.

لم نرد عرض الكثير من الاستمارات من هذا النوع ، إذ أنها غالباً لا تناسب أغراض التقييم في بلادنا ، وإن كانت مفصلة شاملة قامت عليها دراسات أثبتت صلاحيتها . ومن الحدير بالذكر أن جهوداً من هذا النوع تقوم الآن في بلادنا . وإن كان بعضها لا يزال تحت التجربة . وطبعي أننا لن نعرض بطاقات التقييم هذه كاملة وعلى من يرغب في الاستزادة أن يرجع إلى ناشريها المدونة أسماؤهم في آخر هذا القسم .

بطاقات محلية لتقييم رجال التربية

لا يتسع هذا المقام لعرض فقرات هذه البطاقات ، وتناول تعليماتها وطرق إجرائها وتقديرها . وسنكتفى بتمثيل بعضها .

بطاقة تقويم المدرس

قام الأستاذ الدكتور أحمد زكى صالح بكلية التربية بإعداد بطاقة تقويم المدرس وكراسة تعليماتها سنة ١٩٥٩ .

خطوات بناء البطاقة :

١ - تحليل العمل إذ طلب من حوالى سبعين أخصائياً في التربية والتعليم ، أمضوا خمس عشرة سنة على الأقل في ممارسة هذه المهنة ، وصف العمل الذى يقوم به المدرس ثم أعدت بطاقة لوصف العمل .

٢ - وبعد حصر الصفات اللازمة للنجاح في عمل المدرس ، عُرِّفت الصفات المختلفة بطريقة إجرائية ، كما وردت في تقارير وصف عمل المدرس ثم عرضت هذه الصفات على مجموعة من الحكام المختصين في التربية والتعليم ، يبلغ عددهم حوالى ٨٠ مقلداً ، وطلب منهم تقدير درجة لكل صفة من هذه الصفات . وكان التقدير حسب أهمية الصفة في عمل المدرس ، يعطى في صورة رقمية من عشر درجات . وقد استبعدت أى صفة كان متوسطهما

أقل من خمسة . وهكذا بلغ عدد الصفات خمس وعشرون صفة . أضيفت إليها خمس عشرة صفة أخرى تبين أن البطاقة بوصفها الأول كانت تنقصها . وهكذا بلغ مجموع الصفات أربعين صفة .

٣ - تحديد الصفات : جمعت الصفات ووجد أنها تتميز إلى مجموعتين كل منها عشرين صفة ؛ المجموعة الأولى هي الصفات الشخصية والثانية هي الصفات المهنية . ووضعت أمام كل صفة خمس درجات تكون سلماً مصاعاً في أسلوب وصفي يمكن تحويله إلى أرقام . واستخرجت الدرجات المعيارية بطريقة ليكرت Likert .

وإليك بعض فقرات هذه البطاقة :

الدرجة	الصفة	الدرجة	الصفة
	٥ - الدقة في إعداد الدروس		١ - القدرة على اتباع التعليمات
	٦ - الدقة في تعبيراته الشفوية		٢ - العلاقة بالرؤساء
	٧ - الدقة في الأعمال التحريرية		٣ - الانتران الانفعالي
	٨ - اتجاهه إزاء الآراء الحديثة في التربية		٤ - إدراكه للبيئة المحلية
	مجموعة الدرجات في الصفات المهنية		مجموعة الدرجات في الصفات الشخصية

التوقيع

المجموع الكلي :

وإليك طريقة تقدير بعض هذه الصفات :

الصفات المهنية

٢١ - الدقة في إعداد الدروس :

وتتبدل في عناية المدرس بإعداد دروسه سواء في كراسة التحضير أو في

كراسة خارجية ، بحيث يضمن جميع احتمالات أسئلة الطلاب مع مراعاة الرجوع إلى مراجع غير الكتاب المقرر ، وشرح الطريقة التي تتبع .

- ٥ - ممتاز
- ٤ - جيد
- ٣ - متوسط
- ٢ - أقل من المتوسط
- ١ - مهمل .

٢٢ - الدقة في تعبيراته الشفوية :

وتتمثل في اختيار الألفاظ التي تعبر عن أفكاره في دروسه ومناقشاته .

٥ - دقيق جداً في تعبيراته الشفوية .

٤ - واضح في تعبيراته الشفوية .

٣ - يحتاج إلى شيء من التوجيه في تعبيره الشفوي لكي يحسنه .

٢ - لا يحسن التعبير الشفوي .

١ - يخطئ في تعبيراته الشفوية خطأ كبيراً .

٢٣ - الدقة في الأعمال التحريرية :

وتتمثل في الدقة في التعبير اللغوي بقصد تحديد المعنى بطريقة واضحة

مختصرة مع تجنب الأخطاء اللغوية .

٥ - دقيق فهماً ولغة .

٤ - دقيق فهماً وليس كاملاً لغوياً :

٣ - ينضج في الأعمال التحريرية بعد التوجيه :

٢ - ينقص المعنى .

١ - كثير الخطأ في الأعمال التحريرية .

جهود وزارة التربية والتعليم

قام فرع التقويم بإدارة البحوث الفنية بوزارة التربية والتعليم بالإقليم

الجنوبي عامي ١٩٥٨ ، ١٩٥٩ بتصميم :

- ١ - دليل تقويم المدرس .
 - ٢ - دليل تقويم المدرس الأول .
 - ٣ - دليل تقويم الناظر - الوكيل .
 - ٤ - دليل تقويم مفتش القسم .
 - ٥ - دليل تقويم العضو الفني .
 - ٦ - دليل تقويم الأخصائي الاجتماعي .
 - ٧ - دليل تقويم الموظف الإداري .
 - ٨ - دليل تقويم الموظف الكتابي .
 - ٩ - دليل تقويم الرئيس الإداري .
 - ١٠ - دليل تقويم رئيس القسم أو الوكيل .
- وقد أشرف الدكتور محمد خليفة بركات على إعدادها . وقام بالإعداد السيد محمد البسيوني الخطيب . واشترك في التجريب السيد رجاء محمود علام . واشترك في مناقشتها الخبراء من العاملين بكل ميدان منها . وكلها تحت التجربة . وهي ترسل بالبريد لعينات من الفئات العشر الجارى تقويمها .
- واليك بعض فقرات من دليل تقويم المدرس :

موضوع التقدير :	مراتب التقدير	الدرجة
أولا :	(أ) يتقن مادته إتقاناً تاماً ويحرص على الاستزادة منها باستمرار سواء بالاطلاع أو الاشتراك	
الإلمام بالمادة والاستزادة فيها	الفعال في الندوات والمؤتمرات ١٢	
مدى إلمامه بمقررات المواد التي يقوم بتدريسها ،	(ب) له إلمام كبير بالمواد التي يقوم بتدريسها ، ويطلع على كثير مما يجد منها ، ويذهب أحياناً إلى الندوات العلمية ٩	
المواد التي يقوم بتدريسها واستزادته فيها .	(ج) يلم بالمقررات التي يدرسها ، ويطلع أحياناً على ما يجد منها ٦	
	(د) تقتصر معلوماته على الكتب المقررة ، وقلما يطلع على الحديث من مادته ٣	

موضوع التقدير :	مراتب التقدير	الدرجة
ثانياً :	(أ) يفهم تماماً أغراض العملية التربوية في المجتمع ويحدد اختيار طرق التدريس بما يلائم مستويات التلاميذ وظروف بيئتهم ٨	الإلمام بالعملية التربوية وطريقة التدريس :
مدى إلمامه وفهمه لأهداف تربية التلميذ وانعكاس ذلك على سلامة اختياره وتطبيقه لطرق التدريس المناسبة لأحوال التلاميذ .	(ب) يدرك أهداف العملية التربوية ، ويحاول أحياناً تطبيق طرق التدريس المناسبة ٦	
	(ج) معلوماته عن العملية التربوية غير وافية ، وقليل التجديد في طرق تدريسه ٤	
	(د) معلوماته سطحية عن العملية التربوية وأهدافها ، ويخلط دون تبصر بين طرق التدريس ٢	
ثالثاً :	(أ) يلم عن طريق الممارسة والدراسة بأسس الشئون المالية والإدارية المدرسية . ويستخدم هذه الدراية متطوعاً لأداء الكثير من الأعمال المدرسية ٨	الإلمام بالإدارة المدرسية :
الإلمام بالشئون المالية والإدارية إلى الدرجة التي تعينه على أداء أعماله دون خطأ مالى أو إدارى . وإسهامه في الأعمال الإدارية المدرسية :	(ب) يلم بالشئون المالية والإدارية بدرجة تسمح له بالإسهام في الأعمال المدرسية ٦	
	(ج) يتطوع أحياناً في أداء الأعمال المدرسية مستعيناً في ذلك بما يعرفه عن الشئون المالية والإدارية ٤	
	(د) معرفته بالشئون المالية والإدارية قليلة مما يقلل إسهامه في الأعمال المدرسية ٢	

ويرسل الدليل إلى المدرس الأول للمادة ، مع ناظر المدرسة اللذان يقومان بتقدير المدرس . ولكل فقرة وزن كما هو مبين تحت فئة « الدرجة » وتجمع هذه الأوزان وتوضع ، بالطريقة الآتية :

موضوع التقدير ودرجات المراتب المناسبة للمدرس

المجموع بالأرقام

الدرجة بالحروف

اسم المقدر

التوقيع

الدرجة	الصفات	التسلسل
	الإلمام بالمادة والاستزادة فيها	أولاً :
	الإلمام بالعملية التربوية والطريقة	ثانياً :
	الإلمام بالإدارة المدرسية	ثالثاً :
	السلوك الشخصي والسمعة الخلقية	رابعاً :
	استعداده الذهني والقدرة على التصرف	خامساً :
	ضبط النفس والاتزان الانفعالي	سادساً :
	علاقة المدرس بالزملاء	سابعاً :
	علاقة المدرس بالتلاميذ	ثامناً :
	أثره في فواحي الحياة المدرسية	تاسعاً :
	نشاطه الاجتماعي خارج المدرسة	عاشرأ :
	المواظبة	حادي عشر :
	المظهر الخارجي	ثاني عشر :
	مجموع الدرجات	

- الوظيفة أو المرحلة التعليمية التي يرى ترقية المدرس إليها .
 - الوظيفة أو المدرسة التي يرى المقدر نقل المدرس إليها .
 - ملاحظات المقدر التفصيلية عن سلوك المدرس وسمعته الخلقية .
- ويقوم المقدر بكتابة ملاحظاته في مكان مخصص لهذا في الدليل . وكل هذه الخطوات متبعة في كل الأدلة بما في ذلك عدد الفقرات التي يقدر على أساسها المشتغل بالتربية ، وكذلك أوزان الفقرات وطريقة صياغتها وخطوات

الإجراء . . . الخ . فثلاثا نجد المقررات الثلاث الأولى في دليل تقويم الإحصائي الاجتماعي هي :

موضوعات التقدير	مراتب التقدير	الدرجة
أولاً :	(أ) ذو مقدرة كبيرة على تنظيم الحفلات والندوات وتوجيهها بحيث تؤدي إلى ما يقصد منها ١٢	
التأثير على توجيه النشاط الاجتماعي العام في المدرسة : وضع خطة سنوية للنشاط الاجتماعي	(ب) يحكم وضع الخطة السنوية للنشاط الاجتماعي العام بالمدرسة ، وبارع في تنفيذها حسب ظروف المدرسة ، ويحرص دوماً على أن تؤدي أغراضها - - - - - ٩	
و توجيهه نواحي النشاط الاجتماعي من حفلات وخدمات عامة بحيث تؤدي إلى الأغراض التي أقيمت من أجلها .	(ج) يعمل على تنظيم وتوجيه نواحي النشاط الاجتماعي العام بالمدرسة - - - - - ٦	
	(د) اهتمامه قليل بوضع خطة واضحة للنشاط الاجتماعي العام بالمدرسة بحيث يرتبك أمام ما يطرأ على المدرسة من ظروف - - - - - ٣	
ثانياً :	(أ) يتطوع بشكل فعال في نشاط الجماعات والمجالس المدرسية ، ويظهر أثر ذلك في نتائج هذا النشاط - - - - - ٨	
القدرة على العمل مع الجماعات المدرسية المختلفة .	(ب) يبث روح التعاون بين الجماعات والمجالس المدرسية باشتراكه في أعمال كل منها - - - - - ١	
تعاونه مع الجماعات المدرسية في النواحي الاجتماعية والثقافية والفنية والصحية واشتراكه في المجالس المختلفة بالمدرسة .	(ج) يتعاون مع الجماعات المدرسية المختلفة في أعمال المجالس المدرسية - - - - - ٤	
	(د) يؤثر بعض الجماعات المدرسية بنشاطه المحدود وقلما يشترك في أعمال المجالس المدرسية - - - - - ٢	

موضوعات التقدير	مراتب التقدير	الدرجة
ثالثاً :	(أ) يتقن معرفة البيئة المحيطة بالمدرسة مما يعينه على فهم حاجاتها وبذل جهده في خدمتها والاستفادة من إمكانياتها في خدمة	الإلام ببيئة المدرسة وخدمتها :
الإلام بالبيئة المحيطة بالمدرسة وإسهامه وتلاميذه في خدمتها والاستفادة من إمكانياتها في خدمة المدرسة .	البيئة ٨	(ب) لديه المعرفة الكافية بمشكلات البيئة ويعمل على حلها ويستعين ببعض إمكانياتها في خدمة المدرسة ٦
	(ج) يلم ببعض جوانب هذه البيئة ويسهم أحياناً في خدمتها وقلما يفيد المدرسة من إمكانياتها ٤	(د) معرفته بهذه البيئة قليلة وخدماته محدودة ٢
ومن المتوقع أن ينتهى إعداد هذه الأدلة بصورتها النهائية ، بعد التجريب ، قريباً .		

اختبار الاستعداد للوظائف التعليمية

قامت الدكتورة رمزية الغريب بكلية التربية باقتباس وتمصير هذا الاختبار عن اختبار ستانفورد للاستعدادات التعليمية Stanford Educational Aptitude Test تصميم ميلتون جنسون Milton B. Jensen سنة ١٩٢٨ .

والاختبار يتكون من :

القسم الأول : ترتيب المهنة المفضلة وزمنه المتوسط ٨ دقائق . ويقوم على أساس الاختيار من إجابتين وعدد فقرات . هذا القسم ٤ فقرات في كل يفترض أن أمام المفحوص فرصتان يستطيع اختيار إحدهما . مثل :

المقارنة الأولى : (علّم على واحدة) .

وظيفة رقم ١ : وظيفة في مدرسة خاصة في مدينة الاسكندرية بمرتب شهري قدره ٢٤٠ جنيتها مع زيادة سنوية لا تتعدى ٢٤ جنيتها . وهناك فصول ليلية بمرتب إضافي لمن يرغب في ذلك ولا تسمح جداول المعلمين بالدراسات العليا .



وظيفة رقم ٢ : وظيفة في مدرسة خاصة في مدينة الاسكندرية بمرتب سنوي قدره ١٨٠ جنيتها وتحدد الترقية إلى الوظيفة الإدارية في الغالب بالأقدمية ، كما أعدت جداول المعلمين بحيث تسمح لهم بالالتحاق بأقسام الدراسات العليا بالجامعة ، ويوجد قسم للأبحاث بمنطقة الاسكندرية التعليمية تشرف عليه جامعة الإسكندرية إشرافاً جزئياً ، ومعظم بحوثه متعلقة بالمناهج ومشاكلها ومستعد لبحث أية مشكلة أخرى متعلقة بالحياة المدرسية إذا توفرت الإمكانيات .



القسم الثاني : مشاكل النظام المدرسي وزمنه المتوسط ١٥ دقيقة .
ويقوم المفحوص باتباع التعليمات : تخير نوعاً من أنواع التصرفات المعطاة فيما بعد تطابق ما قد تفعله في كل موقف وارسم دائرة حول الحرف الذي يدل عليها (ا ، ب ، ح ، د ، هـ ، و) على أساس الاختيار من عدة إجابات هي :

(ا) أقول له (لهم) مدى خطئه فيما فعل .

(ب) لا أفعل شيئاً .

(ح) أنزل به عقاباً جسمانياً .

(د) ألومه (ألومهم) أمام جميع التلاميذ .

(هـ) أحرمه (أحرمهم) من بعض الحقوق .
 (و) أطلب منه (منهم) أن يعدنى بألا يعود إلى هذا العمل .
 وفي التعليمات يطلب تصور المفحوص أنه مدرس ويقرأ . المشاكل
 وعددها ٢٦ حالة . وبعد أن يختار متغير الإجابة من الإجابات الست
 السابقة يحدد درجة تأكد منها متبعا التعليمات الآتية :
 - بعد اختيارك مباشرة لنوع التصرفات المناسب لك في كل موقف ،
 بين مدى ثقتك بهذا الاختيار لكي يتبين إلى أى مدى أنت متأكد من
 اختيارك . ارسم دائرة حول الرقم المقابل للدرجة اختيارك (١ ،
 ٢ ، ٣ ، ٤) .

درجات التأكد :

- ١ - متأكد تماماً . ٢ - متأكد .
- ٣ - أكاد أشك . ٤ - غير متأكد تماماً .

القسم الثالث : نشاط المدرسة الثانوية ومتوسط زمنه ١٠ دقائق
 ويفترض أن المفحوص يتصور نفسه دائماً ناظر مدرسة ثانوية ، عند
 الإجابة باختيار واضحاً دائرة حول رمز إحدى خمس متغيرات وبيان
 التأكد بوضع دائرة حول رقم إحدى أربعة حالات بنفس الطريقة السابقة .
 وهنا نعرض له ٢٦ وجهاً للنشاط ونأخذ رأيه فيما يتوقع أن يفعله بالنسبة
 لكل نشاط منها .

وللاختبار مفتاح يحتاج إلى بعض التدريب ممن يستخدمه وهو يقيس
 استعداد المفحوص للعمل بالتدريس أو البحث أو الإدارة . ويجرى
 الآن إعداد الاختبار بالصورة النهائية بما يجعله مناسباً لبيئتنا وحاجتنا .

الفصل الثاني عشر

تقييم التليـذ

- ماهى الدرجة .
- وظائف الدرجات والتقديرات .
- وسائل الإتصال بين المدرسة والبيت .

الفصل الثاني عشر

تقييم التلميذ

ما هي الدرجة :

درجة التلميذ في المدرسة أو المعهد أو الجامعة ، في نهاية الأمر ، هي حكم يصدره عليه شخص آخر . وقد يكون الحكم حديسيا أو ذاتيا ، وقد يكون موضوعيا تماما عن طريق إجراء الاختبارات المقننة مثلا . وفي الحالة الثانية نجد أن :

١ - الاختبارات الموضوعية تقدم أسسا مناسبة للوصول إلى نظام للترتيب .

٢ - أن فقرات معينة قد دلت عن نواحي معينة في التحصيل :

٣ - أن الاستجابات المناسبة عن هذه الفقرات كانت هكذا :: :

(كما حدث فعلا) .

٤ - أن نسبة مئوية معينة من الجماعة يمكن أن تعتبر في المستوى ١

وهكذا نرى أن الطرق التي نصل بها إلى التقديرات النهائية مختلفة . فقد يدخل الحكم أساسا في تخطيط عمليات الاختبارات الموضوعية ، وقد يصدر على الفصل ككل ، وقد يدخل في تقدير أداء تلاميذ معينين . والصفات التي قدرت والأوزان التي حددت لكل تلميذ قد تختلف اختلافا كبيرا من مدرس إلى آخر ومن فصل إلى آخر . ولكن الدرجة ، على أي حال ، هي حكم يصدره المدرس على التلميذ .

والدرجة حكم نسبي Relative فإذا كانت درجة التلميذ مثلا ٧٥٪

فإنها قد تكون درجة عالية جدا ، أو ليست عالية ، والأمر يتوقف على

طبيعة الاختبار والعلاقة بين الاختبار والشخص أو الجماعة المختبرة . كما تختلف دلالة الدرجة من مدرس إلى آخر ومن مادة إلى أخرى .

والاختلاف من مدرس إلى آخر ومن مكان إلى آخر يشير إلى أنه ليس للدرجات نقطة مرجعية . وقد يكون المستوى المرجعى هو الجماعة نفسها كأن يقسم تلاميذ الفصل إلى ٢٠٪ ، ٣٠٪ ، ٤٠٪ ، ١٠٪ ، د مثلا . والمعيار Standard قد يكون انطباع المدرس عن أداء الفصول الأخرى وما يتوقعه منطقيا من إنجاز الجماعة المعنية على أساس أنها وصلت إلى مستوى معين من التقدم . وفى كلتا الحالتين المعيار Standard نسبي . فليس هناك أى مستوى ثابت نرجع إليه أداء الفرد .

ولا تستطيع المدرسة أن تهرب من إصدار مثل هذه الأحكام النسبية على التلاميذ . فهذه الأحكام ستكون هامة دائما فى إرشاد التلميذ فى عمله المدرسى وفى فهم عاداته الشخصية ونقائضه ، وفى مساعدته فى تخطيط مستقبله التعليمى والمهنى ، وفى التعاون مع المدارس الأعلى وأصحاب العمل على انتقاء الأفراد الأكثر صلاحية . ويجب هنا أن ننبج عن التساؤلات الآتية :

- ١- متى يجب إصدار الأحكام التقييمية ؟
- ٢- فى أى صيغة يجب أن تسجل هذه الأحكام ؟
- ٣- ما هى العوامل التى يجب تغطيتها ؟
- ٤- على أى أنواع الأدلة يجب أن تعتمد الأحكام ؟
- ٥- من يجب أن يكون مسئولاً عن تقدير الأحكام ؟
- ٦- لمن يجب أن تبلغ ؟ من هو المستهلك المناسب للمعلومات ؟

وظائف الدرجات والتقدير

أولاً : المعلومات التي يحتاجها التلميذ الذي نقدره

لماذا يحتاج التلميذ إلى أن يعلم تقدير المدرسة له ؟ وما هي المعلومات التي يحتاجها . يبدو أن المعلومات المرسله للتلميذ تخدم ٤ وظائف :

(١) دفع العمل المدرسي (٢) توجيه التعليم (٣) توجيه الخطط التعليمية والمهنية (٤) توجيه النمو الشخصي .

١ - الدفع : الدوافع المتضمنة في العمل المدرسي كثيرة مركبة وهي تتضمن الرضى عن الإشتراك في جماعات ، وعن الاستطلاع واستكشاف ميادين جديدة ، والرضى عن النجاح ، والرضى الناتج عن التقبل الاجتماعي ، والرضى عن مسايرة التوقعات و . . الخ .

كما أن هناك حوافز سلبية وهي حقيقية مثل الإحباط والفشل و . . الخ .

الاختبار والتقييم والدرجات ونظام التقارير كلها تدخل إلى دفع التلميذ عن أربع طرق على الأقل هي :

(١) خبرات النجاح والفشل في المحفوظات اليومية ، والإختبارات ، والتمارين المدرسية .

(ب) العلم بأن الآخرين يقدرون عمله ويكونون رأيهم عنه على أساس عمله .

(ح) بأن أدائه أيا كان تقدير المدرسة له سيصبح سجلاً دائماً وجزءاً من تاريخه الماضي الرسمي .

(د) توقعه لما ستجىء به التقارير المدرسية وأثرها الفعلي على عالمه ، عالم الأسرة وعالم الزملاء .

ولكن بعض رجال التربية ينتقدون دفع الدرجات المعطاة في المدرسة ، بطريقة التقليدية على أن الدرجات :

(١) فردية وتنافسية وتؤكد على التحصيل والتفوق الفردي على حساب التعاون والتحصيل المشترك .

(ب) غير جوهرية ولا ترتبط بأى أغراض أصيلة أو بالحاجة إلى تعلم الأشياء لذاتها .

(ج) وهى حائل دون تنظيم العمل فى المدرسة حول اهتمامات وحاجات التلاميذ الأصيلة .

٢- توجيه التعليم :

إن المعلومات عن نجاح وأخطاء التلميذ هامة أيضاً فى توجيه نواحى نشاط تعلمه . وقد بين البحث قيمة تعيين أخطاء بذاتها لدى التلميذ ، وأوجه نقصه وقيمة الإجراء الموجه نحو إحلال الأداء الصحيح محل تلك الأخطاء ، والطرق الفنية فى تشخيص نقاط الضعف الفردي لديه بحيث تعينه وهنا تصبح هذه المعلومات أدوات أساسية فى توجيه عمليات التعلم . والدرجات بالطرق الحالية لا تؤدى وظيفتها فى توجيه التعلم إذ أن درجات الفترات هى ملخص التقديرات بطريقة بعيدة عن التعلم الفعلى لأوجه النشاط وهى ليست كلية ، كما تعوزها القيمة التشخيصية : كما أن الدرجات لا تساعد فى توجيه نواحى النشاط اليومى فى المدرسة .

٣ - توجيه الخطط التعليمية والمهنية التالية :

يحتاج التلميذ كى يضع خططا ناجحة وواقعية عن مستقبله التعليمى المهنى ، بجوار حاجته إلى أشياء أخرى ، إلى فهم واقعى لمستواه فى التحصيل المهنى التعليمى . وهو يحتاج هذه المعلومات حاجة بالغة وقت اتخاذ القرارات ووقت تنفيذها . وعلى أى حال فإن القرارات والخطط تتبلور تدريجياً بحيث أنه فى كل وقت يكون من المناسب أن نعكس للتلميذ الفرد صورة واقعية عن نفسه فى علاقته بتوقعاته التعليمية والمهنية .

ودرجات الفترات ، إذا وضعت بانتظام وبمنهجية تكون أحد طرق إبلاغ الفرد بالأحكام عليه وتصبح ذات دلالة في تخطيطه التعليمي والمهني . فإن النجاح الأكاديمي في لحظة يبصر الفرد بتخطيطه في هذه اللحظة . والمشكلة التي تواجهنا الآن هي ، « ما هي أحسن الطرق » لتقدير التحصيل الأكاديمي وما هي أفضل الطرق لتوصيل هذا التقدير إلى التلميذ .

أما عن طريقة التقدير فيمكن مثلاً أن نعتمد على حكم المدرسين أفراداً ، لأنها مبنية على أي أنواع الأدلة يرون أن استخدامها يناسب الوصول إلى أحكامهم هذه . وهناك احتمال آخر هو إصدار حكم ضمنى يقدمه اختبار مقنن معين أو مجموعة من الاختبارات تقدم تقديراً عاماً عن احتمالات النجاح . وللاختبارات ميزة كبرى هي أن مقياس الدرجة له معنى موحد لا يعتمد على مدرس بذاته أو مدرسة بذاتها .

وأما عن إبلاغ التلميذ بتحصيله الأكاديمي عن طريق بطاقات التقاير الدورية فهي طريقة غير ناجحة إذ أنها لا توصل له دلالة هذا التحصيل في خطته الخاصة . فالأمر يازمه بعض الترجمة أو التفسير ونحن لا نتوقع أن يقوم التلميذ نفسه بذلك . فالطالب يهتم بدلالة سجله أكثر مما يهتم بالدرجات الخام نفسها سواء أكانت درجات مدرسية أو درجات اختبار مقنن . ويرى بعض المرشدين النفسيين إبلاغ الطالب بالتفسير المشتق من درجاته فقط إذ أن الدرجات نفسها لا دلالة لها عند الطالب . والدرجات في المدرسة هي مواد خام تحتاج إلى منظم ومفسر ذي خبرة يقوم بدراسة تتبعية غير رسمية لتلاميذه ليعرف النجاح المحتمل في الجامعة لمن حصل على درجة معينة ، ويقرر الأعمال المعينة التي نصح فيها من حصل على درجة معينة في المدرسة .

٤ - توجيه النمو الشخصي :

نحن نتعلم كيف نعيش مع الآخرين خلال الطريقة التي يستجيب بها الآخرون لأفعالنا . والجماعة التي نعيش فيها هي التي تحدد ما هو السلوك المقبول والمرغوب . والمدرسة أيضاً مسؤولة في أن تعكس للتلميذ الفرد تقديرها لطرقه في السلوك ، بحيث يطورها في اتجاهات مرغوب فيها اجتماعياً وهذا الانعكاس متضمن في كل اتصال يقوم به موظفو المدرسة مع التلميذ . وكل تعبير عن الاهتمام ، أو الموافقة ، أو التحمس ، أو الاعتراض هو مرشد يدل على الطريقة التي يعرف بها مجتمع المدرسة أنماط السلوك المرغوبة وبها يقدر التلاميذ . ونشك في أن التلميذ يفهم التقديرات الدورية وفي أن لها تأثير كبير عليه . ولكن الاتصال الشخصي المباشر في المدرسة والتفاعل اليومي وخاصة في المؤتمرات الخاصة ، هو الطريق المحتمل الذي بواسطته تمارس المدرسة تأثيرها على النمو الشخصي للتلميذ وتطوره .

٥ - العلاقة بين تقدير الذات وتقدير المدرسة :

بعض رجال التعليم التقديميين يحبذون أن يكون تقييم التلميذ ، أساساً ، تقييماً ذاتياً . وهم يفضلون أن يضع التلميذ أهدافه الخاصة لنفسه وأذنه تقيم تقدمه نحوها . ونحن نرى أن توجيه نواحي النشاط التعليمي يمكن أن يكون فعالاً لو أننا رجعنا إلى الأهداف التي وضعها التلميذ بنفسه .

هذا فضلاً عن أنه كثيراً ما يكون تقدير التلميذ لنفسه غير صحيح . والتلاميذ المراهقون تتسببهم الأرضية الأساسية للحكم على ما إذا كانت قدراتهم وإنتاجهم يتفقان له ما نخططه من دراسة علمياً أو عمل مهني . ونحن لا نتوقع من الطالب المتقدم للدراسة أو لعمل أن يحكم على صلاحيته حكماً سليماً ، ولهذا فإن تقدير الذات ليس بديلاً لتقدير المدرسة .

ثانياً : إبلاغ الآباء بتقييم الأبناء :

التواصل بين المدرسة والمنزل وبين المنزل والمدرسة هو الأساس في جعل هاتين الهيئتين تفهمان التلميذ فهما أفضل . وللتقارير التي ترسلها المدرسة إلى الآباء ، ضمن أهدافها ، هدف هو أن يعلم الآباء كيف يبدو التلميذ لموظفي المدرسة في الموقف المدرسي ، بحيث يستطيع الآباء أن :

- ١ - يتقبلوا ويؤيدوا ويدعموا الطفل أثناء مواجهته لمشاكل نموه .
- ٢ - أن يفهموا ويتعاونوا مع برنامج المدرسة من أجل الطفل .
- ٣ - أن يجعلوا الطفل يتبنى أهدافاً مهنية وتعليمية بنائية .

ثالثاً : تقديم التقديرات للمؤسسات التعليمية الأعلى :

تحتفظ المدارس بالسجلات من أجل إفادة المدارس الأعلى والمعاهد الفنية التي يحتمل أن يدخلها التلميذ مستقبلاً . والمدارس الأعلى والمعاهد عند قبول الطلبة الجدد تطلب هذه المعلومات حتى تضعها في الاعتبار عند قبول الطلبة الجدد . والمدارس السابقة عليها أن تقدم معلومات صادقة وتنبؤية حتى يمكن أن يقدر النجاح المحتمل للطلاب على أساسها .

وما دامت المدارس الأعلى تهتم بالقدرة على التفوق في الدراسات المجردة ، فيجب أن تهتم تقارير المدارس الأقل مستوى بهذه الناحية . ولكن ما هي أفيد المعلومات وما هي أحسن الطرق لإرسالها ؟

هي المعلومات عن الاستعداد والتحصيل المدرسين . والاختبارات المقننة تفوق تقارير المدرسين المبنية على عينة صغيرة من الأداء في أن الأولى ترجع إلى جدول عام للمعايير .

أما عن التحصيل فتقديره يحسن أن يجمع بين الناحيتين ؛ وتقديرات المدرسين غير الرسمية ، وتقديرات الاختبارات المقننة . فالأولى مرنة ويمكن

أن تكيف بحسب مدى واسع من أوجه النشاط : كالمهارات فى المشاركة الشفوية ، والنشاط فى المعامل ، ومهارات الإنشاء والتعبير ، . . . وتقديرات الاختبارات الموضوعية فى هذه المهارات الأساسية هامة ، على الأقل ، كدليل يؤيد تقديرات المدرسين . خاصة وأن الاختبارات المقننة فى مهارات القراءة ، والمهارات الحسائية والرياضية ، ومعرفة العلوم والدراسات الاجتماعية وجد أنها تفيد فى التنبؤ بالنجاح لنجاح فى التعليم الأعلى وأنها تفوق ، فى دقتها ، درجات المدرسين .

وعن التحصيل الأكاديمى فى المؤسسات التعليمية الأعلى نجد أن أغلب المقاييس الناجحة فيه هى التى تعطى تقديراً كلياً يتنبأ بالنجاح الأكاديمى . كما تفيد الاختبارات التى تركز على وظائف معينة إذا كان المطلوب معرفة النجاح المقبل فى ميادين ترتبط بهذه الوظائف . ولكن فى أغلب الحالات نحتاج تقديراً ملخصاً للأداء الكلى ؛ مثل الترتيب المئبى فى الفصل وما إليه من إحصائيات مشابهة .

رابعاً : تقديم التقديرات لأصحاب الأعمال :

كخدمة لكل من التلميذ والمجتمع يجب أن تكون المدارس مستعدة لتقديم المعلومات عن التلاميذ لأصحاب العمل المحليين . والمعلومات التى يحتاجها صاحب العمل تختلف من حالة إلى أخرى . ويقوم ناظر المدرسة أو عضو معين من هيئة التدريس بملاء الطلب الذى ترسله مؤسسة التوظيف بدقة وكمال بحسب ما تسمح به المعلومات . وهتا تفيد سجلات المدرسة إن كانت ، كافية وموضوعية . والمطلوب غالباً هو ملخص تقدير مستوى الأداء .

وصاحب العمل غالباً ما يحتاج معلومات عن نواحى غير أكاديمية للطلاب . وربما كانت أهم المعلومات تلك المبينة على حقائق عن النشاط

الخارجي للتلميذ وعضويته في الأندية والمكاتب التي اشترك فيها إلى آخر الأدلة التي تدل على اشتراكه أو مسؤوليته . كما يحتاج صاحب العمل إلى تقدير أعضاء هيئة التدريس للطلاب وخاصة إذا كانوا يتذكرونه . ولكن للخطوة الأخيرة عيوب هي التحيز والمبالغة والنسيان .

خامساً : استفادة المدرسة :

تساعد التقديرات المدرسة على تحسين عملية التعليم وتوجيه التلاميذ وتفيد الاختبارات المقننة في إعطاء تقديرات عن مهارات عامة وعن محتوى ميادين بذاتها . وهذه التقديرات تصبح جزءاً من السجل المجمع . والميزة هنا هي أن المدرس يعرف تلميذه جيداً وهو يحتك به في الفصل ويعرف مستواه من التمارين والتجارب والملاحظات . . الخ . وسجل المدرسة يفيد المدرسة في علاقتها بالتلميذ فيما يأتي :

- ١ - أنه مستمر ومجمع .
- ٢ - أنه وصفي وتشخيصي
- ٣ - أنه متفهم عام .
- ٤ - أنه منظم .

الأسس الفنية في إعطاء الدرجات والتقديرات

أولاً : العوامل المؤثرة :

تفيد الدرجات في خدمة الوظائف الآتية : -

- ١ - تساعد في توجيه الطالب ووالديه فيما يتعلق بخططة التعليمية المقبلة .
- ٢ - تساعد المدرسة في تقرير مدى استعداد التلميذ للدخول في برامج «نقائية معينة أو مناهج .
- ٣ - تساعد المستويات التعليمية الأعلى في تقدير احتمال نجاح الطالب في برامجها .

٤ - تساعد صاحب العمل المحلى فى تقرير مدى صلاحية الطالب لأعمال معينة تعتمد على مهارات أكاديمية .

والدرجة إذا كانت مقياساً خالصاً للجدارة كانت لها وحدة فى المعنى تجعل من الممكن تفسيرها والإفادة منها . ولكن الدرجة إذا اعتمدت على عوامل غير معروفة فى تقدير الجدارة والجهد والاتجاه كانت غير مجدية . والدرجة ليست أداة للنظام تستخدم فى الإثابة أو العقاب ولكنها سجل للتحصيل . والدرجة الجيدة هى التى تقدر شيئاً واحداً بحسن تقرير الجدارة فيه فى ميدان بذاته من التعليم .

ويعتمد تفسير وتحديد معنى الجدارة فى الميدان على المدرس أو مجموعة المدرسين . وأول خطوة فى تقرير ما يجب أن يمثل فى الدرجة على المقرر هى تحديد الأهداف المتعددة للمقرر وتقرير أهمها هنا . ولكن ما أهمية تقدير المعلومات ؟ وفهم المفاهيم والتعليقات ؟ والقدرة على تطبيق المعلومات فى حل المشاكل ؟ والقدرة على تنظيم الأفكار والتعبير عنها ؟ ومهارات العمل أو المعالجة اليدوية ؟ وما هو وزن كل منها ؟ أما عن تقدير الجدارة فى مقرر فيجب أن يعتمد على تعيين عوامل معينة تحدد الجدارة فى هذا المقرر بالذات .

ثانياً : أنواع الأدلة ووزن كل منها .

ما دام المعلم قد حدد أهدافه وقرر وزن كل هدف فى تقييمه الكلى للجدارة ، وجب عليه أن يحدد ما هى الأدلة التى يجب أن يجمعها لكى تمثل كل هدف وأن يحدد كيف أن الأنواع المختلفة من الأدلة ستوزن فى تقديره المركب . فيجب أن يجمع الأوزان التى يعطيها عن درجة امتحان الفترة والتمارين والامتحانات الدورية والشهرية والتقارير والمشاركة فى المناقشة داخل الفصل . . الخ .

(١) امتحان نهاية الفصل الدراسي :

يرى من يعطى الوزن الأكبر للامتحان النهائى الذى يعقد فى نهاية المقرر ، أن التقدير الجوهري لكفاءة الفرد فى ميدان من التعليم يجب أن يعطى لما استوعبه التلميذ فى نهاية هذا الفصل الدراسى . بل وربما كان من الأحسن أن يعطى بعد نهاية انتهاء هذا التعليم بفترة من زمن . وهم لا يعطون اعتبارا لكيف أو متى حصل هذا فالعبرة عندهم بما وصل إليه فى النهاية وما بقى من معلومات حتى نهاية المقرر أو بعده بفترة . وهم لا يرون أن للتعليم قيمة ما لم يستمر التلميذ مستوعبا أهدافه حتى نهاية الفصل الدراسى على الأقل .

(ب) الامتحانات العامة :

وإن كنا نوافق على أن تقييم الطالب هو تقدير جدارته فى النهاية إلا أننا لا نوافق على أنه التقدير الوحيد ولا على أنه التقدير الأساسى وهذه هى الاعتراضات :

١ - يستحيل تقرير أنواع معينة من الجدارة فى حدود الامتحان الكتابى . والقدرة على إيجاد وتنظيم المواد المتعلقة بالمشكلة ، والقدرة على عرض مهارات معينة - مثل استخدام الميكروسكوب - والقدرة على المشاركة الفعالة فى المناقشة الجماعية أو مشروع جماعى هى أمثلة للنتائج التى لا يلائمها التقدير عن امتحان كتابى .

٢ - عينة السلوك التى يمكن الحصول عليها فى امتحان هى عينة محدودة ، وعلى ذلك يكون التقدير أيضاً محدداً . ولكن إذا تضمن التقدير أدلة من مصادر أخرى كان أكثر ثباتا خاصة إذا كان الدليل الإضافى من مستوى عالى كالموجود فى الاختبار . ونحن هنا نهتم بالكم والكيف معا حتى نصل إلى تقدير ثابت بقدر الإمكان .

٣ - العينة المقاسة وقت الامتحان النهائى قد لا تكون عادلة بالنسبة لأفراد معينين . فقد يكون بعض الممتحنين مريضاً ، أو متعباً ، أو واقعة

تحت ضغط ظروف خارجية ، أو أن مستواه منخفض لأسباب أخرى وقت الامتحان . وأداء التلميذ في يوم معين وساعة معينة قد يفشل في تمثيل مستوى أدائه العادى .

٤- الأداء تحت ضغط الإمتحان قد يفشل في تمثيل جدارة الفرد تحت ظروف مريحة وسوية . والامتحان موقف ضغط . والضغط يزيد إذا كان الامتحان واحداً هاما ، وكانت نتيجته ذات آثار هامة على مستقبل الفرد . ويختلف الناس في مدى اضطرابهم تحت مثل هذا الضغط فبالرغم من أن تقدير القدرة على الأداء يؤخذ تحت ضغط وأنه يمكن الاعتماد جزئياً على هذا التقدير ، فإنه يصعب تبرير الاستناد عليه تماماً وعلى العكس نجد أن الاعتماد على تقديرات ومصادر مختلفة بالإضافة إلى الامتحان النهائى يخفف من ضغط الامتحان النهائى ورهبته .

٥- قد تكون لامتحان نهاية الفصل البالغ منتهى الدقة آثاراً ضارة على التعليم ونشاط التعلم خلال العام الدراسى . والعلاقة بين ما يمكن قياسه فى الامتحان وبين أهداف التعليم علاقة ناقصة . فإن بعض أهداف التعليم لا يمكن قياسها فى الامتحانات . وعند ما تضع هيئة خارجية الامتحان تكون العلاقة أكثر ضعفاً فإن نواحى الجدارة التى تبدو هامة من الأعمال التى وضعتها هيئة الإمتحان قد لا تتعلق إطلاقاً بالأهداف التى تبدو واقعية وجوهرية لدى جماعة التدريس . ويبدو هذا صحيحاً عند ما يقوم ممثلون من مستوى أكاديمى أعلى وأكثر تخصصاً ومنتقياً بعناية عند ما يقوم هؤلاء بوضع امتحانات نهاية الفصل الدراسى فى المستوى الأدنى من البناء التعليمى . مثل هذا الضبط للامتحانات من أعلى يشيع فى بعض البلدان الشرقية إذ قامت الجامعات ، التى لها أبسط وأصغر احتكاك بالسكان ، بضبط الامتحانات التى تمثل المحصلة النهائية للتخرج من المدرسة الثانوية وجعلتها فى صورة أكاديمية تشبه ما بالجامعة . وطبيعة امتحانات نهاية

الفصل حددت طابع امتحانات نهاية العام في المدرسة الثانوية ، وهذه بدورها حددت نمط امتحانات آخر العام في المدرسة الإعدادية . وعلى هذا فإن نمط التقييم وإلى حد كبير نمط برنامج التعليم لنسبة غالبية من تلاميذ المدارس الإعدادية ، وأغلبهم لن يذهب إلى الجامعة ، ضبطها نظام الامتحان الذي وضع لانتقاء الأفراد الذين سيتابعون الدراسة في مستوى الجامعة .

وأى برنامج للامتحانات في نهاية الفصل سيكون له تأثيره على التعليم الذى وضعت الامتحانات لتقييمه ، هذا التأثير قد يكون إيجابياً وبنائياً إذا كانت الامتحانات مبنية على تحليل متبصر ينظر إلى الأمام في أهداف التعليم أكثر مما تخص مدرسة بعينها سنطبق عليها طريقة التقييم هذه . وعلى أى حال فالأغلب أن المدارس تنظر إلى امتحانات نهاية الفصل التى تضعها هيئة خارجية على أن لها تأثيراً مقيداً ومحدداً يمنع المدرسة من العمل على تحقيق أكثر الأهداف التعليمية جوهرية .

(ج) الامتحانات الدورية خلال الدراسة :

ونقدم هذه الامتحانات دليلاً أفضل على تحقيق هذه الأهداف أكثر مما تستطيع الامتحانات النهائية . وللامتحانات الدورية خلال الدراسة ميزة تنحصر في استخدامها في الأغراض التعليمية والتقييمية . فإذا أظهر الامتحان الدورى عيوباً كيفية في تحصيل التلميذ ، كان هذا مرشداً يوجه التلميذ في إعادة مذاكرة مواد الدراسة بطريقة أفضل . وكل الامتحانات تحدث صدمة على التلميذ ولكن صدمة الامتحانات الدورية أقل حدة من صدمة امتحان نهائى واحد . فالتلميذ في الامتحان النهائى يتصور أن مصيره معلق على نتيجة الامتحان وحده . بينما في الامتحانات الدورية يعلم التلميذ أن أمامه فرصاً كثيرة لتعويض ما فاتته .

وقد وجد أن عقد الامتحان بعد الانتهاء مباشرة من الوحدات التي تدرب عليها التلميذ أفضل من عقد الامتحان بعد انتهاء مجموعة وحدات تدريبية مرفها التلميذ ببرامج مختلفة . ولكن الامتحانات الدورية تقيس الفهم والتذكر المباشر القريب بينما تقيس الامتحانات النهائية القدر الثابت من الفهم والتذكر للخطوط العريضة في المنهج . ويعيب الامتحانات النهائية أنها تكيف بحيث تقيم الإلمام بالمفاهيم العامة أو التكامل بين أجزاء المنهج في إطار يشمل البرنامج الدراسي كله . ولهذا النقائص روى أن تكون للامتحانات الدورية نسبة من الدرجة الكلية لتقييم التلميذ ، لا أن تكون هي المحك الوحيد .

(د) المقالات والأبحاث والتقارير :

تقيس هذه التقارير التي يعدها التلميذ بنفسه ، خارج الفصل ، عدداً من الأهداف لا تستطيع أن تغطيه الامتحانات الكتابية . فقد يطلب من التلميذ أن يجمع وينظم أدلة وشواهد على صحة فرض من الفروض : فيمنح فرصة للبحث المتعمق نسبياً في موضوع معين ويعطى وقتاً كافياً بطريقة مستقرة على قدر ما يستطيع . وموضوع البحث أو الجهد الخلاق يتكيف بحسب اهتمام الفرد .

وفي مستويات معينة وميادين معينة يكون تقدير هذه المهارات في اكتشاف الحقائق وانتقاء الأدلة وتنظيم العرض وجعله كاملاً ومنطقياً ، والتعبير عن الأفكار بطريقة واضحة يكون أمراً غاية في الأهمية : ولهذا فإن تقييم الأنواع المناسبة من التقارير المستقلة يجب أن يعطى وزناً جوهرياً في الدرجة النهائية .

ولإحساسنا بأهمية مثل هذه التقديرات يجب ألا يصرفنا عن جعلها ثابتة reliable . والصعوبات هنا هي صعوبات امتحان المقال وتزويد في أن لكل

شخص موضوع وليس كما هو الحال في امتحان المقال . وثبات تقدير موضوعات الإنشاء الفردية ، والأوراق ، والتقارير منخفض جداً . وهناك مشكلة أخرى في معرفة كيف أن العمل الذي يتم خارج الفصل يقيم جدارة التلميذ الفرد ، باعتباره عملاً متميزاً عن الأسرة ، والأصدقاء ، والعمليات الأكاديمية .

(هـ) المشاركة في نواحي النشاط الجماعية :

يجب أن يهتم المدرس في تقديره المتفهم للتلميذ بطبيعة ومدى اشتراك التلاميذ في أوجه نشاط الجماعات في الفصل أو المدرسة وأن يهتم بمدى تعاون التلميذ مع الآخرين ودرجة إيجابية الدور الذي يقوم به التلميذ في مشروعات الفصل أو مناقشاته ، ومدى حذق تعليقاته ومدى عمق فهمه الظاهر لما يشترك في مناقشته . وينكر البعض قيمة المشاركة في أوجه النشاط الجماعية عند تقدير التلميذ مدعين أن طبيعة ومدى مشاركة الفصل أمر يتعلق بالمزاج والانجاهات أكثر مما يتعلق بالمناقشة والتعاون ، خاصة وأنه ليس من السهل أن نقيم معرفة الفرد وفهمه بالاستماع إلى ما يقوله أثناء المناقشة ؛ ولكن تقدير المناقشة يكون أمراً جوهرياً عندما تتضمن أهداف التعليم الخاصة إنماء مهارات العمل في جماعة .

والنشاط في العمل :

يجب تقدير دراية التلميذ في استخدام أدوات المعامل وفي استعمال المواد وفي الاستفادة من الأسس النظرية في التطبيق العملي خاصة في الدراسات العلمية والعملية .

وسائل الاتصال بين المدرسة والبيت

١ - بطاقة التقرير التقليدية :

إن نظام الدرجات التقليدى فى المدرسة وبطاقة التقرير هى مقاييس ضعيفة لا توصل إلى المستويات الثلاث السابقة ، فالآباء لا يشجعون على الفهم والشعور بالألفة للدرجات مثل ٧٥٪ أو ٨٠ / أو أ ب أو ح فهى درجات باردة وهى تقييم غير شخصى وهى تفشل فى إثارة تعاطف الآباء بل وتحطم الروح المعنوية فى الأسرة ، وإن كنا لا نقبل الصورة الكاريكاتيرية كصورة للحياة ، إلا أن علينا أن نعرف على لب الحقيقة فيها ونقصد بهذه الصورة بطاقة التقرير .

فالتقرير ، الذى يحوى تقديرات عديدة أو رموزاً أبجدية ، قد يكون ذا قيمة فى مساعدة الآباء على أن يجعلوا أبناءهم يتبنون أهدافاً واقعية . فإذا كانت درجات الأبناء منخفضة ، تشجع الآباء على وضع أهداف محدودة لأبنائهم . ولكن الآباء غالباً لا يفهمون دلالة الطريقة المعينة لوضع الدرجات فى المدرسة المعينة وخاصة فى المستويات الأكاديمية العالية . كما أنه ينبغى أن يرفق بتقدير التحصيل الذى وضعه المدرس تقريراً آخر مشتق من مقاييس الاستعداد المدرسى ومن اختبارات التحصيل المقننة .

ويجب أن نعرف بأن ترجمة الدرجات فى المدرسة ودرجات الاختبار المقنن إلى تنبؤ عن التحصيل المقبل ليس أمراً يسهل القيام به ، حتى على الموثقين المربين المزودين بكل الأدلة الموجودة . فما بالنا إذن بالوالد غير المدرب الذى لا نعطيه إلا بطاقة التقرير المدرسية ؟ إنه قطعاً لن يضع حكماً سليماً . وإذا كان للمدرسة أن تساعد الوالدين فى وضع خططهم عن الطفل بطريقة واقعية ، وجب عليها أن تزودهما بتفسير لقدراته الأكاديمية والمهنية

مبنى على كل الأدلة الممكنة : وهذا التفسير لا يمكن إرساله في بطاقة التقرير .

٢ - تعديل عدد ومعنى الدرجات :

وفيه اختصرت فئات الترتيب إلى متوسط وحسن وضعيف مثلاً ولكن هذه الطريقة لم تفعل أكثر من أنها أمدت الآباء بمعلومات أقل عن أبنائهم .

٣ - إرفاق الدرجات برتب على عوامل أخرى :

والعوامل الأخرى هي عوامل غير تحصيلية مثل : يتقبل التغيرات - يتقبل قرار الجماعة - يتقبل المسؤولية - المظهر - الانتباه - العناية بالمواد - الحرص - النظافة - إنجاز العمل - السلوك - التعاون - يمكن الاعتماد عليه - المجهود . . . الخ .

ولكن يعيب هذه العوامل غموض دلالتها وذاتية مستوياتها وقلة الفرصة لملاحظتها ، وعدم ثبات الحكم . ونحن نجد أن الاختلاف في تفسير هذه العوامل لا يؤثر وحسب على المدرسين الذين يضعون الرتب عنها ، بل ويؤثر أيضاً في الآباء الذين ترسل لهم . كما أن المدرس الذي يقوم بهذه التقديرات لا يستطيع استيعاب رتب كل تلميذ على كل عامل . وقياس العوامل الشخصية بهذه الطريقة الذاتية دائماً يحيطه الغموض خاصة وأن السمات لا تعرف بطريقة صحيحة .

٤ - القوائم التفصيلية :

تحاول بعض المدارس أن تضع في تقريرها المرسل إلى الآباء مزيداً من المعلومات بأن تحلل العوامل والمهارات إلى عدد أكبر من العوامل

المحددة . ومثل هذا التحليل يعتمد على تحليل المدرسة لأهدافها في كل مستوى ومحاولتها وضع الأهداف في عبارات واقعية ونوعية . وهذه الطريقة مصممة بحيث تعطى للآباء معلومات عما تحاول المدرسة تحقيقه وعما وصل إليه التلميذ المعين من أهداف هذه المدرسة .

والتقارير من هذا النوع تفضل التقارير التقليدية إلا أن أمامها صعوبات عادة . وأول هذه الصعوبات هو التعقيد فقد يصل التقرير إلى خمسين حكماً منفصلاً عن كل تلميذ وبهذه الطريقة يكون التقرير مرهقاً للمدرس الذي يقوم به من ناحية ملئه كتابة ومن ناحية الدقة في تذكر المعلومات الكاملة عن كل تلميذ . كما أنه يقلق الآباء خاصة إذا كانوا غير متعلمين .

ولكن التقرير المبني على تحليل مفصل للأهداف والذي يغطي كثيراً من النقاط الخاصة ، تكون أقسام تقييم كل نقطة فيه عادة كافية وبسيطة وهذا التقرير إذا عني بالاستعداد والتحصيل والإنجاز معاً نجح في إمداد الآباء بكثير من البيانات المفهومة خاصة إذا كان أمام كل عبارة متغيرات للتقدير محددة لها معناها عند الآباء .

٥ - المقابلة بين الأب والمدرس :

عقد المؤتمرات الدورية بين الوالد والمدرس وكل موظفي المدرسة هي أفضل طريقة للاحتفاظ بتواصل فعال بين المدرسة والبيت وهذه المؤتمرات لها المزايا الآتية :

- (١) مرونة التواصل : يمكن أن يعدل المؤتمر بحسب حاجات الحالة المعينة . فيؤكد المدرس على ما يريد مناقشته من أمر هذا الطفل المعين .
- (ب) فرص التفسير : يمكن أن يدل التقرير ليس وحسب على حقائق منفصلة بل وعن العلاقات الداخلية بين الحقائق وتقييمها . وفيه تناقش الخطط الطويلة المدى على ضوء الأدلة المتجمعة حتى تاريخه :

(هـ) التواصل متبادل : لا تستطيع المدرسة وحسب أن تبلغ الآباء ولكنها أيضاً تتعلم منهم . فحقائق المنزل تلقى ضوءاً على حقائق المدرسة والعكس بالعكس .

(د) إزالة سوء الفهم : إذا بدى أن الوالد لم يفهم رسالة المدرسة أمكن توضيحها ومناقشته فيها . ولكن هناك عقبات تحول دون الاعتماد على المؤتمرات كطريقة أساسية أو وحيدة في التواصل بين المدرسة والبيت وهذه العقبات العملية تنحصر في :

(١) الوقت : تستغرق طريقة المؤتمرات الفردية ، إذا اتبعت كما ينبغي ، وقتاً طويلاً إذ أن على المدرس أن يقابل كل والد تلميذ عنده خلال العام الدراسي . وللطريقة ميزة هي أن المدرس يضع في ذهنه النقاط التي ناقشها كما أنه يسجل ملاحظاته في المقابلة وهذه تعتبر ضوءاً يلقي على السجل الدائم للطفل . فسجل النقاط التي نوقشت ، والمعلومات التي حصل عليها المدرس من الوالد ، والإنطباعات عن الوالد والبيت كل هذا يجب أن يكون مصدراً يمكن الرجوع إليه وهذا يعنى أن الزمن المخصص من وقت المدرس لكل تلميذ وهو أقل من ساعتين في العام الدراسي بأكمله لا يكفي إطلاقاً . ولهذا يجب أن يسمح وقت المدرس بهذه المؤتمرات بعد أن يخطط وقت كل المدرسين .

(ب) المهارة : يتطلب القيام بهذه المؤتمرات الفعالة مع الآباء مهارات خاصة من جانب المدرس . فهذا المؤتمر موقف إرشاد ، وهو يتضمن مهارات في إقامة علاقة طيبة ، وفي الحساسية لمشاعر ووجهة نظر الأب ، وفي القدرة على تكييف الاتصال حسب شخص الأب : هذه المهارات يمكن أن تحسّن عن طريق الممارسة والتدريب ولكن برغم هذا ستظل هناك فروقا فردية في المهارة : ونجاح نظام المؤتمر يعتمد على الأمن الشخصي

والحساسية والفهم من جانب المدرسين الذين يعتقدون هذه المؤتمرات مع الآباء .

(ح) تعاون الآباء : يتحمس بعض الآباء للذهاب إلى المدرسة ومقابلة مدرس الطفل أو المرشد النفسى وبعضهم لا يتحمس ، لأسباب مختلفة بالرغم من أن التعاون فى المقابلة الأولى أمر هام . ولكن إذا حدث وتمت المقابلة الأولى وأحيط الوالد علماً بمعلومات شعر بقيمتها ولم يُجرح شعوره زاد اهتمامه بتلك المقابلات . وبعض الآباء لا يأتى إلى المدرسة بعد المقابلة الثانية أو الثالثة . وهذا النوع من الآباء غالباً ما تكون المدرسة فى حاجة إليهم . وفى هذه الحالة لا تكون المقابلة أداة ناجحة ويمكن الاستعانة بطرق أخرى . أما فى الحالات التى يكون فيها الآباء متعاونين فإن المقابلة تكون أداة قيمة يعول عليها .

٦ - الخطابات غير الرسمية من المدرس إلى الأب :

تستغرق الخطابات وقتاً أقصر مما تستغرقه الطرق السابقة . ويمكن بها الاتصال بكل أنواع الآباء . والخطابات لا تخلو من المرونة ولكنها تفتقر إلى أثر الاتصال الشخصى وكذا فإنها لا توفر فرصة لتبادل الآراء أو إزالة سوء الفهم . وكتابة الخطابات الجيدة للآباء فن يحتاج إلى مهارة لا تتوفر فى كثير من المدرسين .

الفصل الثالث عشر

اختبارات المقال والاختبارات الموضوعية

- أهمية الامتحانات
- تخطيط الامتحان .
- أنواع امتحانات المدارس .
- قواعد عامة تراعى قبل إعداد الاختبار .
- رسم الخطة لعمل الاختبار .
- إعداد أسئلة الامتحان .
- إعداد الاختبار الموضوعي .
- المقارنة بين الاختبار الموضوعي وامتحان المقال .
- استخدام اختبار المقال وتصحيحه .
- تحديد أهداف تدريس المادة .
- انتقاد نظام الامتحانات .

الفصل الثالث عشر

اختبارات المقال والاختبارات الموضوعية

أهمية الامتحانات المدرسية :

تقيم تقدم التلميذ جانب هام من عمل المدرس ، فالصورة الصادقة عن مستوى التلميذ وكيفية تقدمه أساس في التدريس الناجح وتؤدي عمليات تقييم التلميذ إلى خدمة عدد من الوظائف تقتصر على أربع هامة منها وهي :

١ - الدافع :

يختلف الدافع إلى حد ما من تلميذ إلى آخر ومن فصل إلى آخر ، وتحدد الامتحانات متى يبدأ التلاميذ المذاكرة ، وماذا يجب أن يذاكروا ، وكيف يذاكروا . والامتحانات التي صممت بعناية وكفاية تستطيع أن تدفع التلاميذ إلى إنماء عادات المذاكرة الجيدة وتصحيح الأخطاء وأن توجه نشاطهم نحو تحصيل الأهداف المرغوب فيها . فعمليات الاختبار تضبط عملية التعلم إلى حد كبير .

٢ - التشخيص والتعليم :

تؤدي الامتحانات إلى تشخيص نواحي الضعف وإلى توفير الدليل عن المعارف والمهارات التي يجب توفرها ، فالأسئلة التي يفشل التلميذ في الإجابة عنها أو التي يفشل عدد من تلاميذ الفصل في إجابتها تساعد في تشخيص النقاط التي تتطلب دراسات تالية .

٣ - تحديد أهداف التربية :

إن ما يؤكد عليه المدرس في تقييمه لتلاميذه يدل على ما يعتبره المدرس هاماً . فالمدرس قد يؤكد في حديثه مراراً أنه يهتم باستخدام

الأدوات والأجهزة والتطبيقات العملية للموضوعات التي درست ، ثم يضع امتحاناً تعتمد الإجابة فيه أساساً على تذكر مادة كتبت تفصيلاً في محاضراته . وعلى هذا فحتى لو أقسم مراراً للتلاميذ باهتمامه بالتطبيق فإنهم سيعرفون أنه لا يهتم بها فعلاً كما ظهر من امتحانه فرجع إدراك التلاميذ هنا هو خبرة الامتحان نفسها . والمدرس قد لا يكون على علم بما يريده للتلاميذ كأن يعتبر الأمر يسيراً لئلا يناقش نفسه فيما ينبغي أن يحققه من أهداف في التدريس وبالتالي ما يضمنه من وظائف في الامتحان .

٤ - المقارنة بين التلاميذ والنقل ومنح الشهادات :

المدرس مسؤول عن تقرير مدى استحقاق التلميذ للنقل من فرقة دراسية إلى أخرى أعلى منها ، وعلى هذا فهو يقرر مدى النمو أو التقدم الذي أحرزه التلميذ بعد أن درس برنامجاً معيناً . بل ويبدئه أيضاً أمر تقرير مدى استحقاق التلميذ لاجتياز مرحلة من مراحل التعليم . وهو هنا أمام أمر أخطر ، إذ أنه عند إعطاء شهادة إتمام مرحلة دراسية يعطى جوازاً بالمرور إلى مرحلة تالية أعلى . هذه المرحلة الأعلى تتطلب بالأغلب لا مستوى أعلى من القدرات المستخدمة في المرحلة السابقة بل وأيضاً ، وهو المهم ، أنواعاً أخرى من القدرات والمهارات ، ولهذا فعليه أن يكون أكثر حرصاً .

تخطيط الامتحان

١ - تحديد الأهداف .

٢ - تخطيط المحتوى .

٣ - إعداد نسخة من الامتحان .

وتتكون خطة الامتحان عادة من إقرار النقاط التالية :

١ - تخطيط المحتوى والأهداف .

٢ - اقتراحات خاصة عما يمكن أن تغطيه الموائمة بين المحتوى والأهداف .

٣ - تحديد النسب المئوية التي سيشغلها كل جزء من محتوى المجال المقاس بالنسبة للامتحان على أساس المحتوى والهدف وتقدير العدد الكلى للأسئلة .

٤ - تحديد مدى صعوبة الأسئلة وطريقة تدريجها :

والخطوة التالية هى إعداد أسئلة الامتحان فعلا . وسناقش هنا اختيار أنواع الأسئلة وتوجيهات لتحسين صياغة استخدام أسئلة المقال . وقبل أن نستطرد فى مقارنة اختبارات المقال والاختبارات الموضوعية ننبه القارئ للعود إلى الفصل الثانى .

أنواع امتحانات المدرس

يمكن تقسيم الامتحانات التى يضعها المدرس إلى قسمين كبيرين هما امتحانات المقال أو الامتحانات ذات الإجابة الحرة والامتحانات الموضوعية أو التى تسمى الامتحانات الحديثة ؛ ولكل منها مزاياه وقصوره ؛ ولكل موضعه فى قياس التحصيل فى الفصل .

اختبار المقال :

الخصائص الرئيسية فى الإجابة التى يتطلبها امتحان المقال فى كل تلميذ هى :

- ١ - أن ينظم إجاباته ببذل أقل جهد .
- ٢ - أن يستخدم ألفاظه الخاصة .
- ٣ - أن يجيب عدداً صغيراً من الأسئلة .
- ٤ - أن يأتى بإجابات كاملة ودقيقة .

الاختبار الموضوعي :

يتضمن الاختبار الموضوعي أو الامتحان الحديث عدداً متنوعاً من طرق الإجابة تشترك جميعاً في أن الإجابة الصحيحة هي أحد الاحتمالات التي وردت في إجابات السؤال . وهنا يوضع السؤال ويتبعه الجواب الصحيح ضمن عدة إجابات غير صحيحة أو ناقصة .

وشروط الامتحانات هي :-

١- أن تكون صادقة أى أن تكون المعلومات التي يختبرها الامتحان هي نفس ما يراد منه اختبارها . فإذا كان الامتحان في مادة التاريخ وجب ألا تكون للعبارة المنمقة ، أو جودة الخط ورواءته ، أو الأخطاء الإملائية ، أو الرسم ، . . الخ نصيب في تقدير الدرجة .

٢- أن يكون الامتحان ثابت النتائج بمعنى أن لا يتأثر كثيراً بشخص المتحن أو المصحح أو وقت الإجراء فإذا طبق الامتحان مرة أخرى على نفس الفرد أو طبق امتحان يكافئه من حيث المدى الذي يغطيه ومن حيث صعوبة الأسئلة كانت النتائج متقاربة .

٣- يجب أن لا يتأثر السؤال بسواه من الأسئلة وأن تكون الأسئلة من الصعوبة بحيث تظهر الفروق الفردية بين الأقوياء ، ومن السهولة بحيث تظهر الفروق الفردية بين الضعفاء . كما يجب أن لا تكون الأسئلة من الصعوبة بحيث ينال فيها تلميذ صفرأ ولا من السهولة بحيث ينال البعض فيها الدرجة النهائية . وتقاس صعوبة السؤال بتوزيع درجات التلاميذ عنه ، فإن كانت نسبة من أجابوا عليه إجابة صحيحة نسبة كبيرة كان السؤال سهلاً والعكس بالعكس كذلك .

٤- أن يكون الامتحان شاملاً ممثلاً للمادة التي يراد تقييم تحصيل التلاميذ فيها .

- ٥ - أن تكون الإجابة موضوعية أى أن تعتمد على الحقائق لا أن تعتمد على ميول ووجدانات واتجاهات التلاميذ .
- ٦ - سهولة الإجراء والتصحيح وقلة التكاليف المالية والإدارية والوقت وإمكان نقل الاختبار من مكان لآخر .
- ٧ - أن يواجه الممتحن نفسه بالسؤال عما يريد تحقيقه من الإجابة عنه بالتحديد وبحيث يتمشى هدف الامتحان مع أهداف التدريس .
- ٨ - ألا تكون الأسئلة ذات أثر صادم كبير وأن يسبق الإمتحان علاقة طيبة لتوفير ظروف الدفع وضمان التعاون ومنع الغش والتزييف .

مزايا الامتحانات الموضوعية :

- ١ - تمنع التقدير الذاتي .
- ٢ - تنفادى غموض الإجابة .
- ٣ - تمنع الإجابات الخارجة عن الموضوع .
- ٤ - تشمل مقداراً كبيراً من المادة المراد الامتحان فيها :
- ٥ - تتمتع بتوزيع سليم لأنواع المعلومات والأسئلة عنها ووزن كل منها في مجموع الدرجات .
- ٦ - إذا افترضنا وجود خطأ في بعض هذه الوحدات فإنه يكون خطأ يسيراً جداً .
- ٧ - سهولة للطالب والمدرسة والإدارة المدرسية .
- ٨ - تنفادى تأثير النظافة والاهتمام بالشكليات في الدرجة .
- ٩ - يسهل بواسطتها عمل معايير .

كيف نضع أسئلة الصواب والخطأ :

- ١ - يجب أن يكون عدد الأسئلة المطلوب الإجابة عنها كافياً .

- ٢- أن تحتل العبارة معنى واحداً .
- ٣- أن لا تكون صياغة السؤال موحية بالجواب .
- ٤- أن لا يزيد عدد الصواب عن الخطأ زيادة كبيرة .
- ٥- أن لا يكون ترتيب الأسئلة وتسلسلها مؤدياً إلى معرفة الإجابة :
- ٦- ينبغي تفادى التفصيل في مواد قاستها مواد أو امتحانات أخرى .
- ٧- يراعى الاختصار في عدد الكلمات التي تتركب منها العبارة ما أمكن .
- ٨- يجب أن تكون العبارة غير معقدة فلا تتضمن أكثر من سؤال واحد أو فكرة واحدة .
- ٩- يحسن الابتعاد عن العبارات المعممة .
- ١٠- يعمل تصحيح للدرجة النهائية ضد التخمين وذلك بطرح عدد الإجابات الخاطئة من عدد الإجابات الصحيحة للحصول على الدرجة المصححة .
- ١١- يحسن استخدام دليل للتصحيح أو مفتاح للتصحيح .

طريقة الاختيار من عدة إجابات :

وتتميز هذه الطريقة بعدة مزايا يدركها من يراجع الفصل الخاص بإنشاء الاختبار ، وهذه المزايا هي :

- ١- أنها تلفت الطالب إلى ضرورة التمييز ومعرفة الحقائق معرفة دقيقة .
- ٢- أنها تعود الطالب على الحكم الصائب والموازنة وتمييز أفضل الأشياء .
- ٣- من السهل تصحيحها بعد إعداد مفتاح لإجاباتها .
- ٤- قريبة إلى طابع تفكير الطالب .
- ٥- يمكن استعمالها في مواد كثيرة خصوصاً المواد الاجتماعية والعلمية .
- ٦- لا يكاد يدخل فيها عنصر التخمين :

٧- يستطيع المدرس أن يتعرف منها على الأنواع التي يشيع الخطأ فيها فيستغلها في تدريسه لمادته وفي وضعه للامتحان التالي .

قواعد عامة تراعى قبل إعداد الاختبار في صورته النهائية :

١- يجب إعداد أسئلة أكثر من العدد المحدد للاختبار في صورته النهائية . كأن نعد ضعف العدد المقدر للاختبار أو ما يقرب من ذلك . لأنه في كثير من الحالات يتضح ضعف بعض الأسئلة أو عدم صلاحيتها كما يتضح أن بعض الأسئلة لا تتفق مع الأهداف الحقيقية التي أعد لها الاختبار .

٢- يجب اختبار وفحص الأسئلة جيداً في فترات متباعدة متعددة فكثيراً ما يبدو السؤال جيداً أو ممتازاً ولكن بعد تركه فترة من الزمن ثم فحصه ثانياً يظهر أن فيه غموضاً أو التباساً أو أنه أصعب مما بدا أولاً وهكذا .

كما يجب تجريب الاختبار عدة مرات وتحليل النتائج ومعرفة انجاهات وآراء المفحوصين وبعض من يلمون بنفس المادة من المدرسين .

٣- يجب طبع الاختبار بحيث يكون واضحاً (بالرونيو أو البلوطة) وتعد منه نسخ أكثر من المطلوب .

٤- يجب ترتيب أسئلة الاختبار بحيث يسير فيه المفحوص دون اضطراب أو تعثر ويراعى في ذلك النقاط التالية :

(أ) ترتب الأسئلة بحسب صعوبتها الأسهل أولاً ثم الذي يليه في الصعوبة وهكذا .

(ب) الأسئلة ذات الصيغة الواحدة توضع مع بعضها كأسئلة الصواب والخطأ ثم أسئلة التكميل وهكذا .

(ج) الأسئلة التي تتناول موضوعاً مشتركاً أو تتقارب في مادتها نوضع مع بعضها .

٥ - يجب أن تعد للاختبار تعليمات صريحة واضحة ومحددة وأن تحضر استمارة للإجابة بصورة تبسط عملية الإجابة وأن يعد للاختبار مفتاح .

٦ - يشار في تعليمات الاختبارات إلى التصحيح ضد الصدفة والتخمين وأن يفهم المفحوص ذلك كي لا يؤخذ على غرة ويكون التصحيح حسب المعادلة الآتية :

$$\text{الدرجة} = \text{عدد الإجابات الصحيحة} - \frac{\text{عدد الإجابات الخاطئة}}{\text{عدد احتمالات إجابة كل سؤال} - 1}$$

رسم الخطة لعمل الاختبار

إن أهم وظيفة لعملية التقييم في التربية هي أن نعرف إلى أي حد حقق التلاميذ أهداف وأغراض الدراسة . ولكي يقيس الاختبار هذه الناحية يجب أن يوضع له تخطيط دقيق . أما أن يجلس المدرس ويضع عن المقرر عدة أسئلة يضمها في اختبار فإن ذلك في الغالب يجعل الأسئلة تقيس المعلومات وتهمل الأهداف . كما أن بعض الموضوعات يسهل كتابة أسئلة عنها بينما يصعب ذلك في موضوعات أخرى . وفوق هذا كله فإن الأهداف تضع في زحمة المعلومات وتواردها وفي اهتمام المفحوص بأيا يختار .

وهذا الكلام ينطبق على اختبارات المقال كما ينطبق على الاختبارات الموضوعية . لذلك كان تحديد أهداف تدريس المادة والمدرسة عموماً هو أهم خطوة يجب العناية بها قبل البدء في عمل الاختبار وكتابة الأسئلة أو اختيارها .

لذلك كان تحديد أهداف تدريس المادة والدراسة عموماً أهم خطوة يجب العناية بها قبل البدء في عمل الاختبار وكتابة الأسئلة أو اختيارها .

ويجب أن يستعين المدرس على تحديد الأهداف بالبحوث والكتب التربوية عن مادته كما يستعين بمدرسين آخرين وفوق هذا يجب بذل جهد من الدراسة والتفكير لمعرفة هذه الأهداف وتحديد ها وترتيبها تبعاً لأهميتها ومعرفة نواحي النشاط والعمليات والجهد الذى تتضمنه هذه الأهداف .

إعداد الأسئلة الموضوعية

فيما يلى بعض القواعد الهامة التى يجب مراعاتها عند إعداد الأسئلة الموضوعية . ولكن ينبغى ألا نعتقد أن مجرد الإلمام بهذه القواعد واتباعها حرفياً كفى بأن يجعل الأسئلة موضوعية جيدة . فإن إعداد الأسئلة عملية فنية تحتاج إلى ذوق ومران وخبرة بالإضافة إلى مراعاة هذه القواعد : ولكنها تساعدنا على السير فى الاتجاه الصحيح وتجنب الأخطاء .

أولاً : سهولة المستوى اللغوى لصيغة الأسئلة : فى جميع الاختبارات التى لا يقصد بها قياس القدرة والفهم والتحصيل اللغوى ينبغى أن تكون الصيغة اللغوية للسؤال سهلة وواضحة بالنسبة للجماعة التى يعد لها الاختبار : والقصد من ذلك هو أن لا تكون قدرة الطالب على القراءة والفهم عائقاً يحول بينه وبين ذكر المعلومات التى يتطلبها السؤال :

أما فى اختبارات القدرة على القراءة والفهم فإن الصيغة اللغوية للسؤال تتدرج فى الصعوبة حتى تصل إلى مستوى أعلى من جميع الطلبة اللذين أعد لهم الاختبار .

ثانياً : لا تضع فى السؤال عبارة بنصها من الكتب المقررة أو نصوص المحاضرات : لأن وضع العبارات الواردة فى الكتب بنصها يشجع الاعتماد على الذاكرة الصماء ويقلل من الاهتمام بالفهم . ويمكن تحقيق ذلك (١) إما بتغيير مبنى العبارة دون معناها أو (ب) استعمال عبارة أخرى تحمل

نفس المعنى بالضبط . أو (ح) ذكر إحدى الاستعمالات أو التطبيقات التي تتضمن المعلومات التي يتطلبها السؤال :

ثالثاً : الأسئلة المبنية على الآراء الجدلية ضعيفة : عندما يكون السؤال مبني على رأى أو قول مأثور عن مصدر خاص يجب ذكر صاحب الرأى أو القول . ذلك لأن الآراء والأقوال القابلة للجدل يختلف الأفراد في فهمها والحكم عليها . ولكن ذكر صاحب الرأى يكسب السؤال شيئاً من التحديد . على أن الآراء الجدلية غالباً لا تعطى أسئلة جيدة ويجب تجنبها إلا في بعض الحالات الخاصة التي تقتضى ذلك .

رابعاً : يجب أن لا تكون صيغة أحد الأسئلة تحمل في طياتها الإجابة على سؤال آخر في نفس الاختبار . لأن ذلك يجعل الإجابة على السؤال الأخير تعتمد على الحيلة أو نوع من الملاحظة وذلك بالطبع يجعل المقياس غير صالح في قياس ما أعد له .

خامساً : أن يكون كل سؤال مستقلاً عن غيره : أى لا تكون الإجابة على أحد الأسئلة متوقفة على معرفة الجواب الصحيح للسؤال السابق : لأن استقلال الأسئلة يجعل الفرصة متساوية للجميع أمام كل سؤال .

سادساً : أن تكون الإجابات الصحيحة على الأسئلة تسير على نظام عشوائى بحيث لا يستطيع الطالب أن يكتشف أو يعتقد أن الإجابات الصحيحة تسير على ترتيب خاص ثم يحاول أن يبنى إجاباته وفقاً لهذا النظام مثلاً :

١ خطأ ٢ صواب ٣ خطأ ٤ صواب ٥ خطأ ٦ صواب
أو ١ صواب ٢ صواب ٣ خطأ ٤ خطأ ٥ صواب ٦ صواب

سابعاً : يجب أن تكون الأسئلة واضحة وصریحة : وخالية من أى خدعة أو فخ : ويجب أن لا يكون السؤال فخاً أو مصيدة هدفها إيقاع الطالب أو خداعه فإن مثل هذا النوع كثيراً ما ينخدع فيه

الطالب الأقوى والأكثر تحصيلاً . لأن وفرة معلوماته وثقته بنفسه يجعلانه أقل حذراً فيكون أكثر قابلية للوقوع في الفخ أما الآخر قليل التحصيل والأضعف ثقة بنفسه يكون متردداً ومتشككاً وبذلك لا يقع في الفخ وبالطبع تكون النتيجة مخالفة للغرض الأصلي من الاختبار .

لذلك يجب أن تكون صيغة السؤال واضحة وصريحة بعيدة عن العبارات الخداعة والحيل المضللة .

ثامناً : تجنب الغموض والتعقيد : سواء من حيث العبارة اللغوية أو من حيث المعنى الذى يقصد من السؤال .

يجب أن تكون عبارة السؤال واضحة خالية من أى غموض أو تعقيد أو إبهام كما يجب أن لا تحمل العبارة إلامعنى واحداً لا لبس فيه ولا احتمالات .

تاسعاً : تجنب الأسئلة فى نقط تافهة : يجب أن تتناول الأسئلة النقط والموضوعات القيمة وأن تنصب على الأجزاء ذات الأهمية . ويمكن التحقق من ذلك بالتأكد من أن معرفة الإجابة الصحيحة عن هذا السؤال تميز بين الطالب الضعيف والطالب القوى فى المادة التى يقيسها الاختبار .

إعداد أسئلة الاختبار

أسئلة الاختيار من عدة إجابات :

فى هذا النوع من الأسئلة توضع أمام السؤال عدة إجابات ليس بينها إلا إجابة واحدة صحيحة وتكون إجابة المفحوص باختيار أحد هذه الإجابات بوضع علامة أمامه أو تحته . وأسئلة الاختيار من عدة إجابات هى أحسن نوع من الأسئلة الموضوعية ذلك لأنها أكثر مرونة إذ يمكن أن تصاغ بطرق كثيرة ولأنها تصلح لقياس نواحي كثيرة كالتحصيل . اللغوى والمفردات ،

والمعلومات العامة والفهم والتطبيقات العملية فهى تصلح لقياس ناحية من نواحي التعليم ما عدا تنسيق وترتيب المعلومات وعرضها وهى أكثر دقة وتقل فيها الصدفة بدرجة كبيرة . ولكن هى فى نفس الوقت تحتاج إلى جهد ومهارة .

ويتكون سؤال الاختيار من إجابات متعددة من جزئين الجزء الأول هو صلب السؤال أو جوهر السؤال وهذا الجزء هو العبارة الأساسية التى تعرض المشكلة المطلوب حلها أو الإجابة عليها وهو قد يكون عبارة ناقصة وقد يكون سؤالا كاملا .

والمبتدئون فى وضع الاختبارات يجدون أن صيغة السؤال الكامل أسهل وهى تجنبهم الكثير من الأخطاء . على أن العبارة الناقصة إذا أعدت بمهارة تكون أقصر وأدل . والجزء الثانى من سؤال « الاختيار من إجابات متعددة » هو قائمة الإجابات التى يختار منها الجواب الصحيح . وكلما زاد عددها فى السؤال كلما ضعف أثر الصدفة والتخمين ويفضل أن تكون هذه القائمة مكونة من أربعة أو خمسة احتمالات . إذ ليس المهم كثرة عدد الإجابات المعروضة إذ فقد يكون عددها خمسة ولكن منها اثنين أو ثلاثة يبدو خطأها واضحا فيصير الاختبار الحقيقى من إجابتين أو ثلاثة وتتوقف درجة صعوبة السؤال (الاختيار من إجابات متعددة) على دقة الفروق بين الإجابات المعروضة للاختيار ولكن يكون بينها واحد فقط هو الصحيح .

قواعد تراعى عند عمل أسئلة الاختيار من عدة إجابات :

أولا : يجب أن يكون صلب السؤال واضحا وأن تبدو جميع الإجابات المعطاة من الجائز أن تكون إجابات على المشكلة الواردة فى صلب السؤال . كذلك يجب أن تكون الإجابات قصيرة وأن لا يكون كل منها شبيهاً بأسئلة الصواب والخطأ .

ثانياً : لجعل صلب السؤال يتضمن أكبر جزء من السؤال لأن هذا يحقق الإنجاز في الإجابات فيقلل من طول السؤال كله وبالتالي يوفر الوقت والجهد في القراءة . معنى ذلك أن يكون صلب السؤال طويلاً نسبياً بينما تكون الإجابات قصيرة ما أمكن .

ثالثاً : تجنب شحن السؤال بأشياء غير ضرورية لأن المطلوب أن تكون المشكلة التي يطوى عليها السؤال واضحة وبارزة .

إستثناء : يستثنى من هذه القاعدة الأسئلة الموضوعة لقياس القدرة على استخلاص النقط المهمة من وسط مادة كثيرة ؛ ففي هذه الحالة توضع في السؤال مادة غير أساسية مختلطة بالنقط الأساسية أو النقط الهامة .

رابعاً : تأكد تماماً أنه لا يوجد بين الإجابات المعطاة إلا جواب واحد فقط هو الصحيح تماماً أو هو الأفضل من بين الإجابات المعطاة .

خامساً : الأسئلة الموضوعة للاختبار والفهم والقدرة على التطبيق (تطبيق المبادئ والمعلومات في مواقف خاصة) يجب أن تصاغ في عبارات غير العبارات التي درست بها أو وردت بها في الكتب والمحاضرات .

سادساً : تجنب الأسئلة التي يكون فيها الجواب الصحيح بينه وبين صلب السؤال ترابط سمعى أو لفظى لأن في هذه يعتمد المفحوص على هذه الإرتباطات السطحية . وعلى العكس يستفاد كثيراً من هذه الترابطات بين صلب السؤال وبعض الإجابات الخاطئة ولكن يجب استعمال ذلك بحذر شديد وعدم الإسراف فيها حتى لا يتحول السؤال إلى فخ مضلل

سابعاً : تجنب وجود إشارات نحوية أحياناً : ترد في صلب السؤال بعض إشارات نحوية يستعين بها المفحوص لمعرفة الجواب الصحيح . كأن ترد في السؤال صيغة الجمع وتكون الإجابات الموجودة بينها واحد فقط في صيغة .

الجمع . أو كأن يرد في السؤال ضمير وتكون الإجابات بينها واحد فقط يمكن أن يعود إليه هذا الضمير وهكذا .

ثامناً : تجنب العبارة الطويلة في صلب السؤال : هناك احتمال كبير أن العبارات الصحيحة تكون أطول وذلك لأن صحة العبارة تقتضى ذكر كثير من الخصائص والتحديدات لقيام صحتها على عكس العبارات الخاطئة فإنها غالباً تكون قصيرة وذلك لأنه لا حاجة إلى الاهتمام بأى تخصيصات أو تحديدات .

Short Answer & أسئلة التكميل والأسئلة ذات الجواب القصير

Completion items

الأسئلة التى تكون الإجابة عليها بكتابة إجابة قصيرة (كلمة أو جملة صغيرة) ، تشبه بدرجة كبيرة الأسئلة التى يطلب فيها تكميل العبارة بملء الفراغ المتروك فى العبارة . والفرق الوحيد بين هذه الأنواع من الأسئلة هو أن النوع الأول تكون فيه صبغة السؤال عبارة كاملة ويطلب الإجابة عليها ، أما فى النوع الثانى فتكون العبارة الواردة فى السؤال ناقصة ولا يتم المعنى إلا بوضع الكلمة المكمل لها :

ويصلح هذان النوعان لاختبار تحصيل المفردات واختبار ومعرفة الأسماء والتواريخ والتعاريف وتحديد المفاهيم identification of concepts - وكل المسائل الحسابية والرياضية كالجبر التى تنتهى بجواب عددى أو جبرى محدود .

ولكن هذين النوعين من الأسئلة يتطلبان جهداً كبيراً فى التصحيح أو إعداد مفتاح الإجابة نظراً لتعدد الإجابات وتعدد الكلمات التى ترد فى تكميل الأفراد للأجمل وهذا يحتاج إلى دراسة هذه الإجابات المختلفة وتقدير ما يعتبر منها صحيحاً .

قواعد تراعى عند عمل أسئلة التكميل والأسئلة ذات الإجابات القصيرة :

أولاً : تجنب الأسئلة غير المحددة والأسئلة المفتوحة : الأسئلة المفتوحة والأسئلة غير المحددة تكون لها إجابات كثيرة صحيحة وهذا بالطبع يجعلها سهلة جداً :

ثانياً : تجنب ترك فراغات كثيرة في العبارة المطلوب تكملتها : يحسن أن يكون الفراغ المتروك مكان كلمة أساسية يعتمد عليها معنى الجملة Key word أما ترك فراغات كثيرة فإنه يجعل الإجابة أسهل كما يجعل الإجابات الصحيحة تتعدد . والأسئلة ذات الفراغات الكثيرة تشجع على التخمين .

ثالثاً : لجعل الفراغ المطلوب ملئه قرب نهاية الجملة : عند ما يكون الفراغ المطلوب ملئه قريباً من نهاية الجملة تصبح العبارة ذات معنى محدد نسبياً قبل أن يصل المفحوص إلى الفراغ .
أما جعل الفراغ في بداية الجملة فإنه يجعل المفحوص يقرأ إلى آخر العبارة بفهم ناقص ويحمله على معاودة القراءة من البداية أو التردد أثناء القراءة وهذه عادات سيئة في القراءة يحرص المربون على استئصالها :

رابعاً : في أسئلة الحساب والجبر يحسن تحديد نوع الوحدات المطلوبة في الجواب يحسن تحديد نوع الوحدات المطلوبة في الجواب على المسائل الرياضية لأن ذلك يجعل عملية التصحيح أسهل كثيراً وأكثر موضوعية كما أنه يبعد عن المفحوص كثيراً من الحيرة والتخمين عن نوع الوحدة التي يكتب بها الجواب . وأحياناً يترك المفحوص كتابة الوحدات أمام الجواب ولا يستطيع المصحح التأكد هل هذا سهواً أم جهلاً من المفحوص .

أسئلة الصواب - والخطأ :

كثير إقبال المدرسين على عمل اختبارات الصواب والخطأ ، وذلك يرجع إلى سهولة صياغتها . مما أدى إلى ظهور الأسئلة الموضوعية الضعيفة من هذا النوع . مع أن كثيراً من الأسئلة الجيدة من هذا النوع يعتبر عملية دقيقة ويحتاج إلى دراية وخبرة . ولعل شيوع الأسئلة الضعيفة هذه كان العامل الأول في كثير من الحملات الموجهة إلى الاختبارات الموضوعية .

والحقيقة أن اختبارات الصواب والخطأ عموماً وحتى الجيد منها ، لا تصلح إلا في العبارات التي لا جدال في صحتها أو التي يكون خطؤها واضح . ولذلك فهي تستعمل في الأسئلة التي تتناول معلومات جزئية خاصة أو حقائق قليلة الأهمية . كذلك يمكن استعمالها بنجاح في اختبار معاني المصطلحات وصيغها ولكن من الصعب جداً استعمال هذا النوع لاختبار الفهم الدقيق والاستنتاج السليم أو لاختبار التطبيقات العلمية .

لذلك كانت اختبارات الصواب والخطأ تتناول جزئيات قليلة الأهمية وتشجع على الحفظ الآلى والدراسة السطحية وتشجع على التخمين وتأثر كثيراً بعامل الصدفة . كل هذا جعلها هدفاً للنقد الشديد . ذلك النقد الذي امتد أثره إلى غيرها من الاختبارات الموضوعية على غير حق .

والصغية الشائعة لاختبارات الصواب والخطأ هي أن تذكر عبارة . ويوضع أمامها صواب ، خطأ ويطلب من المفحوص أن يضع علامة خطأ أو علامة صواب على الكلمة التي تمثل حكمه على هذه العبارة . مثال :

في سنة (١٩٥٨) تم اتحاد سوريا ومصر لتكوين الجمهورية العربية المتحدة
صواب خطأ

وقد أدخلت بعض تعديلات في صياغة هذا النوع من ذلك وضع خطأً تحت الكلمة أو الجزء الأسامي من السؤال هكذا :

في سنة ()
 ومن تعديلات هذه الطريقة أن يطلب من التلميذ أن يكتب الإجابة الصحيحة عن الأسئلة التي يرى أن إجاباتها المرفقة خاطئة : ويكون التصحيح هنا بكتابة رقم أو اسم أو شيء مجرد بدلا من كتابة العبارة الأساسية التي تحتها خط في السؤال .

وصورة أخرى من أسئلة الصواب والخطأ هي أن ينص على أن العبارة خاطئة في الجزء أو الكلمة التي تحتها خط ويطلب من المفحوص أن يصحح. هذه العبارة بكتابة الكلمة أو الرقم الصحيح في المكان المخصص لذلك .

قواعد تراعى عند عمل أسئلة الصواب والخطأ :

أولا : تجنب استعمال الكلمات الشاملة لأن المفحوص يستعين بها على معرفة الجواب مثل دائماً ، وأبداً ، في جميع الأحيان ، لأن المفحوص يرى في مثل هذه العبارات الشاملة تلميحاً إلى خطأ العبارة وكذلك يرى المفحوص تلميحاً إلى صحة العبارة في الأوصاف الجزئية مثل في غالب الأحيان أو أحياناً وذلك التلميح يتجه إليه الطالب الذي اعتاد هذا النوع من الأسئلة وبدلاً يكون لنفسه حيلة يستعين بها .

ثانيا : تجنب الصيغ الوصفية غير المحددة : إن الصيغ الوصفية غير المحددة. مثل غالباً أو كثيراً أو عظيماً أو بلرجة شديدة . يختلف الأفراد في تفسير مدلولها . فمثلاً عندما نقول هذا الأمر يحدث كثيراً فإن تفسير ذلك يختلف فيه الأفراد فمنهم من يعتبره ٩٠٪ ومنهم من يعتبره ٨٠٪ أو ٧٠٪ أو ٦٠٪ أو حتى ٥٥٪ .

ثالثا : تجنب صيغة النفي وبصفة خاصة نفي النفي : إن صيغة النفي كثيراً ما تخطئها القارئ السريع الذي يعنى بقراءة الفكرة لأنه يغفل عن

أداة النفي (لا) ومثيلاتها لصغرها ولأنها ليست من صلب العبارة الأصلية . أى أن المعنى الإيجابي يكون وحدة فكرية كاملة بالرغم من إغفال أداة النفي . أما نفي النفي فهو معقد وكثيراً ما يثير الالتباس عند المفحوص .

رابعا : تجنب العبارات التي تشمل أكثر من فكرة واحدة : العبارات التي تشمل أكثر من فكرة تقيس القراءة والفهم أكثر من قياسها للمعلومات والتحصيل ولذلك نجدها تصلح جداً لإختبار القراءة والفهم .

على أن هذا العيب يكون أشد إفسادا للسؤال إذا كانت إحدى الفكرتين صحيحة والأخرى خاطئة .

خامسا : تجنب العبارات الطويلة : العبارات الطويلة تحمل في طياتها تلميحا للإجابة . ذلك لأن العبارات الصحيحة غالبا ما تكون أطول من العبارات الخاطئة . وهذا يرجع إلى ضرورة إشتغال العبارة الصحيحة على بعض التحديدات والقيود كي تصبح صحيحة إطلاقا .

أسئلة الزواج Matching : هي حالة خاصة من أسئلة الإختيار من عدة إجابات والفرق بينهما أنه في أسئلة الإختيار من إجابات متعددة يكون صلب السؤال عبارة عن مشكلة واحدة ويختار المفحوص الإجابة من قائمة الإجابات المعطاة . أما في أسئلة الزواج فهناك قائمتان : القائمة الأولى فيها عدد من المشاكل والقائمة الثانية فيها إجابات هذه المشاكل ولكن بترتيب مخالف . والمطلوب من المفحوص أن يربط كل مشكلة من القائمة الأولى مع جوابها من القائمة الثانية ويكون ذلك بأن يكتب أمام كل مشكلة رقم الإجابة التي وردت في قائمة الإجابات .

وأسئلة الزواج تعتبر من الأنواع الممتازة وتفيد كثيراً لاختبار معاني المفردات ، وتواريخ الحوادث ، ونسبة الكتب إلى مؤلفيها ، والأحداث إلى ظروفها أو عواملها ، والنظريات إلى أصحابها والاكتشافات إلى مكتشفها ،

والرموز مع ما تدل عليه كالرموز الكيميائية وأسماء المركبات التي تدل عليها :
ومن مميزات أسئلة التزاوج أنه يمكن استبدال قائمة الإجابات برسم
تخطيطي لخريطة أو رسم بياني أو رسم لحيوان أو جزء من تشريح حيوان
وتوضع أرقام على بعض أجزائه ويكتب المفحوص رقم كل جزء أمام
الاسم أو الوصف الخاص به وهكذا .

قواعد تراعى عند عمل أسئلة الربط أو المزاوجة :

أولاً : يجب أن تكون جميع المشكلات الواردة بالسؤال متجانسة أى من
موضوع واحد أو من طائفة واحدة كأن تكون جميعها مركبات كيميائية أو أن
تكون جميعها أسماء نظريات علمية أو جميعها معلومات جغرافية وهكذا .
وذلك لأن الاختلاف الكبير في مشكلات السؤال الواحد كثيراً ما تساعد
على اكتشاف الإجابة الصحيحة أو على الأقل تخفض كثيراً من صعوبة
السؤال وتسمح بالتخمين والحيلة .

ثانياً : يجب أن تكون قائمة الإجابات في السؤال أكبر من عدد
المشاكل الواردة به . لأنه إذا كانا متساويين فإن الإجابة على المشكلة
الأخيرة تكون محلولة بنفسها إذ تربط الإجابة الباقية بالمشكلة الباقية .
ولكن يمكن مخالفة هذه القاعدة عندما تكون الإجابة من قائمة
الإجابات ستستعمل أكثر من مرة كجواب لأكثر من مشكلة .

ثالثاً : يجب أن يكون سؤال الربط قصيراً نسبياً وذلك لأن اختبار
مشكلات متجانسة يكون ممكناً وأيسر في القوائم الصغيرة . كما أن القوائم
الطويلة تربك وتقع فيها أخطاء كتابية أثناء عملية الربط وهى أخطاء لا علاقة
لها بموضوع الاختبار .

رابعاً : يجب أن تكون قائمة الإجابات مرتبة ترتيباً منطقياً فمثلاً في قوائم
التواريخ أو الأعداد توضع سلسلة أو في قوائم الأسماء تكون مرتبة أبجدياً

وذلك لتلافى الجهد والأخطاء الكتابية (في عملية الكتابة) وهى أخطاء لا علاقة لها بالناحية التى يقيسها الاختبار .

خامسا : يجب أن تكون التعليقات واضحة فى شرح أساس عملية الربط كأن يذكر أنه يمكن استعمال الإجابة الواحدة أكثر من مرة .

اختبارات التصنيف : هى صورة من صور اختبار المزاوجة (الربط والتوفيق) matching وهى مفيدة فى اختبار المفاهيم خصوصا ما كان منها متقاربا . ويتكون اختبار التصنيف من القائمة الرئيسية وهى قائمة تشمل مجموعة من التعاريف والأوصاف أو الخصائص أو التفسيرات وأمام كل واحد منها حرف أو رقم يرمز له . ثم يعطى المفحوص مجموعة أخرى من المفاهيم أو المشاكل ويطلب منه أن يكتب أمام كل منها الرمز الدال على الوصف أو الشرح الذى ينطبق عليها .

قياس الفهم :

الصيغة الغالبة على الاختبارات الموضوعية هى قياس المعلومات ودرجة التحصيل أو حالته : ولذلك كثيراً ما تهاجم بأنها لا تقيس الفهم : والواقع أن الفهم متضمن فيها من ناحيتين : أولاً قدرة الفرد على تحصيل الحقائق والمفاهيم تطلب قدرته على فهم ما يقرأ أو ما يسمع . وثانياً أن الإجابة على الأسئلة وعمل المقارنات وإدراك الفروق الدقيقة بين الإجابات المختلفة كلها عمليات تطلب الفهم .

على أن الفهم فى الأسئلة المهمة بالمعلومات والتحصيل يقاس بدرجة ضعيفة نسبياً . لذلك عملت اختبارات أخرى موضوعية لقياس الفهم .

يتكون السؤال فى اختبار الفهم من فقرة فى أول السؤال يقرأها المفحوص ثم يجب على أسئلة تتناول المعانى والأفكار والنقد الخاص بتلك الفقرة .

وتقاس درجة المفحوص بعدد الإجابات الصحيحة وهي تتوقف على درجة الفهم كما تتوقف أيضاً على سرعة القراءة والفهم في الاختبارات المعدة لهذا الغرض . ويجب ألا تكون القائمة الرئيسية طويلة وذلك لتعاشي عيوب التذكر والكتابة ويستعمل مثل ذلك أيضاً في نواحي أخرى غير القراءة اللغوية مثل إعطاء رسم بياني أو تخطيطي لقياس القدرة على قراءة ما فيه من معلومات وحقائق وكذلك رسم خريطة جغرافية أو رسم منحنى رياضي .

إعداد الاختبار الموضوعي

اختيار نوع وحدات الاختبار : والأنواع هي :

- (أ) فقرات تعتمد على الاسترجاع البسيط :
- (ب) فقرات التكميل :
- (ح) اختبار الاستجابة من عدة استجابات :
- (د) فقرات المزاوجة .

اختيار نوع المواد : والمواد غالباً هي :

- (أ) لفظية .
- (ب) عددية .
- (ح) رسوم :
- (د) رموز :
- (هـ) مكعبات :
- (و) مواد عملية :
- (ز) صور أخرى نادرة مثل بقع الحبر .

تأليف الفقرات : وتراعى عادة عند إنشاء فقرات موضوعية
الأمور الآتية :

- (أ) مراعاة قواعد التعبير اللغوى الجيد :
- (ب) استبعاد الكلمات الصعبة .
- (ح) تجنب استخدام ألفاظ وردت بالكتب المدرسية المقررة .
- (د) البعد عن الغموض .
- (هـ) عدم استخدام فقرات لها إجابات واضحة .
- (و) على المؤلف ألا يشير إلى الإجابة أو يوحى بها .
- (ز) استبعاد الفقرات التى تعتمد إجابتها على الذكاء وحده .
- (ح) استخدام كلمات كمية والبعد عن الكلمات الكيفية .
- (ط) البعد عن العبارات المضللة .
- (ى) عدم تداخل الفقرات .
- (ك) تخصيص مكان للإجابة عن الفقرات فى شكل عمود .

وعند تأليف فقرات الاسترجاع البسيط وفقرات التكميل تراعى
الأمور الآتية :

- (أ) يكون طول الخطوط المتروكة للإجابة كافيا وموحدا فى كل الفقرات .
- (ب) تحديد الاستجابات المرغوب فيها ، كأن تكون كلمة أو تاريخ أو عدد أو رمز أو معادلة أو . . الخ .
- (ح) يجب أن تكون الإجابات المرغوب فيها هامة .
- (د) كل إجابة صحيحة تأخذ درجة .
- (هـ) غالباً لا يعاقب المفحوص على أخطاء الهجاية .

- (و) تنتظم مواضع الإجابة في نهاية كل جملة .
 (ز) تكلمة الموضوعات تتبع في الموضوعات التي تكون كلا موحدا .
 (ح) تكلمة الموضوعات يجب ألا تعتمد على الإجابة في مسافات كثيرة
 وعند إنشاء فقرات اختبار كاختبار إحدى استجابتين تراعى
الأمور الآتية :

- (أ) تجنب استعمال جملتين سلبيتين .
 (ب) عدم استخدام جملة تتكون من شقين أحدهما صحيح والآخر خطأ .
 (ج) المحددات التي تعين الاستجابة مثل كلمة دائماً وغالباً وأبداً ، تستخدم بحرص شديد .
 (د) تطلب الإجابات بطريقة موضوعية جداً .
 (هـ) يتساوى عدد الجمل الصحيحة مع عدد الجمل الخطأ .
 (و) تخلط العبارات الخاطئة مع العبارات الصحيحة بطريقة عشوائية :
وعند إنشاء فقرات الاختيار من عدة إجابات تراعى الأمور الآتية :
 (أ) توضع الكلمات التقديمية في أول الأسئلة ولا داعي لتكرارها .
 (ب) متغيرات الإجابة تصاغ تبعاً لقواعد اللغة والأسلوب .
 (ج) عدم وضع اختبارات واضحة الخطأ .
 (د) يجب أن تكون لكل فقرة أربع أو خمس متغيرات (إجابات) :
 (هـ) توضع المتغيرات دائماً في نهاية كل جملة .
 (و) تطلب الإجابات بطريقة موضوعية .
 (ز) توزع الإجابات الصحيحة بطريقة متكافئة بالنسبة لترتيب متغيرات الإجابة .
 (ح) توزع الإجابات الصحيحة بطريقة عشوائية بين المتغيرات .

وعند إنشاء فقرات المزاوجة تراعى الأمور التالية :

- (أ) لكل فقرة إجابة واحدة صحيحة .
- (ب) اتساق قواعد اللغة .
- (ح) اتساق التصنيفات .
- (د) فقرات المزاوجة لا تكون طويلة جداً ولا قصيرة جداً .
- (هـ) توضع الفقرات بطريقة عشوائية فى كل قائمة .
- (و) كل مجموعة فقرات مزاوجة توضع فى صفحة واحدة :
- (ز) تطلب إجابات موضوعية .

المقارنة بين الاختبارات الموضوعية واختبارات المقالة :

المقال الموضوعية	العامل
	١ - توفر فرصة لاختبار قدرة التلميذ على الانتقاء ، والتنظيم ، أو التكامل +
	٢ - تطلب من التلميذ أن يقدم الإجابة لا أن يتعرف عليها وحسب +
+	٣ - خالية من عوامل مهارة التعبير وحسن الخط
+	٤ - خالية من فرص المراوغة
+	٥ - خالية من فرص التخمين
+	٦ - توفر عينة ممثلة كافية من المواد المطلوب قياسها
+	٧ - يمكن إعدادها بسرعة
+	٨ - يمكن تصحيحها بسرعة
+	٩ - يمكن أن يصحها الكاتب آلياً
+	١٠ - يكون تصحيحها متسقاً بين المصححين المختلفين

والموازنة بين أهمية هذه العوامل تختلف من موقف إلى آخر : ومن الواضح أن أياً من النوعين لا يشمل كل المزايا : والمدرس ، في تقديره لعمل فصله ، يحسن به أن يستخدم طريقتي القياس معاً .

الاستخدام الفعال لامتحان المقال

سبقي اختبارات المقال مستخدمة في تقييم أداء التلميذ من أجل المزايا الآتية :

(١) سهولة إعدادها .

(ب) مزاياها في تقييم القدرات على تنظيم إجابة عن سؤال .

(ج) وفي تقييم استرجاع recall وانتقاء المعلومات المناسبة .

(د) » » عرض المعلومات عرضاً منطقياً فعالاً .

وإذ كان للمدرس أن يستخدمها كان علينا تزويده ببعض المبادئ التوجيهية من مثل ، متى يستخدمها ، وماذا عليه أن يعمل كي يتغلب على نقاط الضعف العام فيها . وهذا الضعف يتضح له جزئياً في صيغ الأسئلة ، كما يتضح في عملية تقييم الإجابات التي أجابها التلاميذ .

متى يستخدم اختبار المقال

تعد العوامل التي من أجلها نستخدم اختبار المقال عوامل عملية من ناحية ونظرية أساسية من ناحية أخرى .

اعتبارات عملية مباشرة :

١ - يكون اختبار المقال حل وسط عملي في فصل عدد تلاميذه صغير أو عندما يكون وقت إعداد الاختبار محدوداً . فمن السهل إعدادها ، ومشقة الإعداد بالنسبة للاختبار الموضوعي تأخذ وقتاً وجهداً يزيد عما تأخذه قراءة إجابات التلاميذ في حالة اختبار المقال .

- ٢ - قلة عدد نسخ الاختبار الموضوعي وخاصة إذا كان جهازاً .
- ٣ - عدم وضوح فقرات الموضوعية أحياناً ووجود ما يلتبس على التلميذ فهمه في أسئلتها . وهذا العيب ليس في طبيعة الاختبارات الموضوعية ولكنه راجع إلى ضعف إلمام واضعها بفن إعداد هذا النوع . ولكن هذا لا يدفعنا إلى رفضها واللجوء إلى اختبارات المقال بل يدفعنا ، وهذا هو الأفضل ، إلى توضيحها خاصة وأن جهد المدرس هنا سيكون بسيطاً .

إعتبرات نظرية أساسية :

هناك وظائف يجيد اختبار المقال قياسها بأحسن من الاختبار الموضوعي هذه الوظائف هي ؛ القدرة على انتقاء الأفكار ، وربطها ، وتنظيمها ، وخلق أنماط أساسية جديدة من الأفكار ، واستخدام الفرد لغته للتعبير عن أفكاره . ولكن يجدر استخدام الاختبارات الموضوعية في حالة قياس إعادة صياغة الحقائق العلمية .

يمكن تلخيص عيوب اختبار المقال في :

١ - الفروق في معرفة الحقائق الأساسية تخفي الفروق في القدرة على تنظيم هذه الحقائق .

٢ - ضغوط الزمن تخفي -بتودة تعبير الفرد كتابة .

هناك استعمالان آخران لاختبارات المقال ؛ الأول اختبار الكتاب المفتوح open -book T. وفيه يسمح للمفحوص بالإستعانة بكتبه ومذكراته وكل ما يمكن أن يعينه من جداول . . الخ . وهنا تقل أهمية التذكر ، وتبرز قدرة المفحوص على استخراج المعلومات المطلوبة وقدرته على الانتقاء والاستفادة من الحقائق .

والثاني ، أن يعطى السؤال كامتحان خارج الفصل ودون تحديد وقت

مما يقلل ضغط الشعور بضيق الوقت ويجعله أقرب ما يكون إلى اختبارات القوة من حيث تنظيم المعلومات ومن حيث حسن التعبير كتابة .

تصحیح اختبارات المقال

نذكر فيما يلي بعض خطوات تعين على تقليل أثر الذاتية في التصحيح وهي تعتمد على تجزئة عملية التقدير إلى نواحي محدده :

١ - يقرر المصحح ، مقدماً ، ما هي العوامل التي ينبغي قياسها : وإذا كان هناك أكثر من صفة أو خاصية يعمل لكل منها تقديراً مستقلاً . فإذا كانت المعلومات مهمة يعمل تقدير منفصل على أساس المعلومات وحدها ، وإذا كان التنظيم مهما يعمل تقدير منفصل على أساس التنظيم : وإذا كانت استعمالات اللغة يعمل تقدير على تقدير الجمل والهجاء : وينبغي أن يحرص المصحح على أن يجعل تقدير أى صفة لا يتأثر بالصفات الأخرى :

٢ - أن يعد المصحح ، مقدماً ، نموذجاً للإجابة يبين النقاط الهامة ثم يخصص درجة لكل نقطة حسب أهميتها . وبعد ذلك يقارن النموذج بعينة من إجابات التلاميذ ، ثم يعدل النموذج والدرجات المخصصة للنقاط الفرعية على ضوء هذه الإجابات . بعد هذا يصبح النموذج معياراً يستعين به عند تقدير كل ورقة .

٣ - أن يقرأ المصحح سؤالاً واحداً في جميع الأوراق حتى يتمكن من وضع تقديراته على هذا السؤال على أساس موحد . ثم ينتقل إلى سؤال آخر وهكذا .

٤ - يقدر المصحح أوراق الإجابة دون أن يلتفت إلى أسماء أصحابها حتى لا يتأثر بفكرته عنهم ، لتكون تقديراته موضوعية ما أمكن .

٥ - يراجع المصحح تقديراته السابقة على أساس الإجابات التي حصلت على تقدير متوسط ، ويعدل فيها على ضوء هذا المتوسط كميّار .

تحديد أهداف تدريس المادة كخطوة أساسية

لوضع خطة عمل اختبار

نذكر فيما يلي أمثلة لبعض الأهداف الهامة :

- ١ - يفهم التلميذ معانى المصطلحات والمفاهيم الهامة :
- ٢ - أن يكون لدى التلميذ محصول قاعدياً من المعلومات والحقائق .
- ٣ - أن يفهم التلميذ عمليات التطور والعلاقات السببية :
- ٤ - « » « الأسس والقواعد العامة :
- ٥ - أن يطبق هذه « » « على مواقف جديدة .
- ٦ - أن يكتسب القدرة على التعرف على النقط غير الأساسية وتفسيرها .
- ٧ - أن يطبق العادات التي اكتسبها في النقد التقييمي للأدلة .
- ٨ - أن يظهر مشاعر وتفهمات مرغوبا فيها .

كما أن هناك أهدافاً أخرى من تدريس بعض المواد مثل :

- ١ - القدرة على استعمال بطاقات الفهارس من المكتبة للوصول إلى مصادر معلومات خاصة :
- ٢ - استعمال الفهرست لتحديد مواضع المعلومات في الكتب وفي دائرة المعارف .
- ٣ - أن يقرأ ويحصل على المعلومات من الرسوم البيانية والجداول والنشرات الدورية والصور :
- ٤ - أن يكتشف العلاقات ويعمل الاستنتاجات بناء عليها من العرض البياني أو الأرقام :
- ٥ - أن يدرك تحديدات مجموعة خاصة من البيانات .

الامتحانات شر لا بد منه

الامتحانات شر لا بد منه للأسباب الآتية :

- ١ - لأنها تختبر وعى المعلومات المطلوبة ؛
- ٢ - لأنها ترفع حماس الطالب ؛
- ٣ - لأنها تدفع التلميذ إلى الاستعداد بالمذاكرة ؛
- ٤ - تحدد الغرض من دراسة المادة .
- ٥ - تقيس كفاية المدرس .
- ٦ - عن طريقها نقيم التلميذ :
- ٧ - منها يمكن تشخيص نواحى الضعف فى برنامج الدراسة ؛
- ٨ - خلالها نستطيع أن نقرر مدى صلاحية المدرس لتدريس مادة بعينها مع تكشف النقاط التى لا يجيد توضيحها .
- ٩ - تعيننا على اختيار أفضل الطريق لتدريس كل مادة :
- ١٠ - تحدد لنا نواحى الضعف والقوة فى المجموعة .
- ١١ - بعد هذا التحديد نستطيع تقسيم المجموعة إلى فصول متقدمة وأخرى متخلفة نتولاها بعناية فى تدريس المادة .
- ١٢ - وتبعاً لهذا يمكننا تخطيط برنامج تربوى علاجى يتناول الصعوبات التى يجدها التلاميذ فى المادة تناولا علمياً مدروساً .
- ١٣ - نستطيع أن نقرر عن طريقها مدى تحقق برنامج الدراسة لأهداف التعليم .
- ١٤ - لا بد من إجراء الامتحانات لغرض تقييم التلميذ والمدرس والبرنامج والمدرسة وتشخيص نواحى الضعف والقوى فى كل من هؤلاء وتناولها بالعلاج .

١٥ - عن طريقها يمكن التنبؤ بمدى النجاح المحتمل للتلميذ المنقول من فرقة إلى فرقة أخرى .

١٦ - بواسطتها نستطيع تقرير مدى صلاحية التلاميذ للالتحاق بدراسة أعلى .

١٧ - تعين الطالب على اختيار مواد التخصص بعد أن يعرف درجاته في هذه المواد والمواد القريبة منها :

١٨ - تعتبر أداة لتصفية المتقدمين للالتحاق بنوع معين من الدراسات .

انتقاد نظام الامتحانات

١ - يوجه أنصار مدرسة الجشطالت إلى الامتحانات الموضوعية نقداً يتلخص في أن الممتحن لا يهتم بتكوين صورة كاملة عن الطالب ، إذ أنه يقدر الأجزاء كل على حدة . على أن هذا النقد مردود ، إذ أن تكوين فكرة كلية عن الطالب يعتبر انطباعاً *inpression* ونحن نعلم أن الانطباعات يتأثر بذاتية من يكونه وبالمرات الأولى التي التقى فيها مع تلميذه .

إذا أردنا إلغاء الامتحانات التقليدية وجب علينا أن نراعى كثرة الأسئلة ودقتها وكل الشروط التي تتعلق بوضع الأسئلة الموضوعية مع مراعات تدرجها في الصعوبة فيأني الأسهل في الأول ، على أن نوزع أوزاناً نسبية على هذه الأسئلة بحسب صعوبتها وبحسب أهميتها بالنسبة لأهداف تدريس هذه المادة .

على أن هناك انتقادات أخرى كثيرة توجه إلى نظام الامتحانات نجملها ونرد عليها فيما يلي :

١ - أصبحت الامتحانات غاية في ذاتها بعد أن كانت أداة توصلنا إلى الأهداف . ولهذا يجب ألا ننسى أن الامتحان وسيلة وليس غاية في ذاتها :

٢ - لم تعد الامتحانات أداة تعليمية والمفروض أنها خبرة يجب أن يفيد منها كل من التلميذ والمدرس والمنهج الدراسي وطريقة التدريس والمدرسة والنظام التعليمي .

٣ - يجب الاعتماد في كل المستويات عند تقدير الطالب لآعلى درجة في الامتحان النهائي وحده بل أيضاً على درجات أخرى أهمها :

(أ) الامتحانات الدورية ويحسن أن تكون شهرية حتى يكون الطالب قد تعلم وحدة كاملة من كل منهج .

(ب) درجات المجهود الشخصي الذي يقوم به التلميذ في المنزل :

(ج) درجات عن اشتراك التلميذ في المناقشات .

(د) درجات عن النشاط العملي والمعملي للتلميذ :

(هـ) درجات عن هوايات التلميذ واهتماماته وأوجه نشاطه الأخرى في المدرسة وخارج المدرسة :

(د) تكليف التلميذ بعمل تمارين على كل درس وإعطاءه درجات عنها توضع في الاعتبار عند تقييمها :

(ز) يقوم التلميذ باختيار بحث أو موضوع في كل مادة ينخصص له جزء من الحصص لمناقشة التلاميذ في تخطيطه وتوضع عنه درجة لها وزنها في تقييم التلميذ .

٤ - الامتحانات بصورتها الحالية غالباً ما تؤكد على عامل الحفظ والاسترجاع وتغفل عامل الفهم العام .

٥ - الامتحان النهائي باعتباره الفرصة الوحيدة أمام الطالب وباعتباره أيضاً الحكم الفيصل الذي يقرر أمراً خطيراً في تاريخ التلميذ المدرسي ، الامتحان بهذا الاعتبار يحدث صدمة تؤثر على استعداد التلميذ له ، ولكي نعود التلميذ على مواجهة مثل هذا الموقف يجب إعداد له بأن نعقد

امتحانات تجريبية تماثل موقف الامتحان النهائى من كل الوجوه حتى لا يكون الامتحان النهائى موقفاً غريباً .

٦ - قد تؤثر على إجابة التلميذ فى الامتحان النهائى عوامل مختلفة مثل مرضه أو اضطرابه أو حدوث ظروف فى الأسرة مما يجعل اعتبار درجة هذا الامتحان محكاً وحيداً أمراً غير عادل .

٧ - أسئلة المقال غالباً ما تشمل عدداً محدوداً جداً من نقاط المنهج ، وهنا يتدخل عامل الصدفة فقد يكون التلميذ من مستوى تحصيلى منخفض عموماً ولكنه أستعد فى نقاط معينة قليلة حدث إن أنت هى أو بعضها فى الامتحان وهكذا لا يكون من العدل أن نحكم بصلاحية مثل هذا التلميذ بينما نحكم بعدم صلاحية زميله الذى ألم بكل المنهج فيما عدا نقط قليلة أتى منها الامتحان ولم يجب عنها بالتالى إجابة كافية . ولعلاج هذا الضعف يجب ، إذا لزم أن تتبع طريقة المقال ، أن نراعى ما يأتى :

(أ) مراعاة أهداف تدريس المادة .

(ب) مراعاة أهداف المدرسة .

(ج) مراعاة أهداف نظام التعليم .

(د) مراعاة طرق التدريس التى اتبعت :

(هـ) مراعاة النقاط التى أولاها المدرس عناية خاصة والنقاط التى كان لها نصيب أقل من اهتمامه .

(و) مراعاة مستوى التلاميذ .

(ز) تحليل الأهداف التى من أجلها وضع الامتحان تحليلاً منطقياً

فإن طريقة الامتحان تختلف بحسب أهدافه باعتباره وسيلة تمكنا

مرة من نقل التلميذ إلى فرقة تالية ، وفى أخرى تمكنا من الكشف

عن الموهوبين . فإذا كان الهدفان طرفى تقيض فيجب بالتالى أن

تصاغ الأسئلة بطريقة تبرز مدى استحقاق التلميذ للنقل إلى
فرقة تالية مرة ومدى كفاءته في الحصول على مكافأة تشجيعية
في أخرى .

(ح) يجب أن نزن كل هذه الأمور في وضعنا للأسئلة بحيث تتضمن
الموضوعات التي يتوقع التلاميذ أن يسألوا فيها بناء على إعدادهم
بطريقة معينة :

(ط) يجب أن تغطي الأسئلة الموضوعات الرئيسية لا التفاصيل الدقيقة
التي تتطلب ذاكرة رقية والنقط الهامشية التي قد تخفى على
أغلب التلاميذ .

(ى) يجب الإكثار من عدد الأسئلة حتى تكون هناك فرصة للاختيار
أمام الطلبة .

(ك) يجب إعداد الطلبة للإجابة عن هذا النوع من الامتحانات
كطريقة ترتيب النقاط وعرضها والاستعانة بالأمثلة حيث أن
كتابة المقال حتى في العلم هي فن .

(ل) يجب أن لا يتأثر المصحح بلغة الطالب وتنميته وخطه خاصة في
المواد العلمية والاجتماعية ،

(م) يدرب المصححون على طريقة موحدة في تصحيح هذا النوع
من الأسئلة :

(ن) تعد نماذج للإجابة عن الأسئلة يراعى فيها مستوى الطلبة والنسبة
المطلوب نقلها أو نجاحها في حدود التخطيط العام التعليمي والمهني :

(س) تصاغ الأسئلة بطريقة لا لبس فيها مع البعد عن الكلمات
الغامضة بحيث يكون مضمون السؤال الذى قصده واضعوه
مفهوما تماما لدى كل التلاميذ .

(ع) يحسن أن يلحق بالسؤال قائمة بالنقاط التي يجب أن يبرزها التلميذ في إجابته (مثل مع التأكيد على : : مع الإشارة إلى : : أو موضحا في إجابتك كذا وكذا : . الخ) :

ولعلاج نقاط الامتحانات والاختبارات التي يضعها المدرس نرى
 فضلا عما ذكرنا من الأساليب الفنية للتغلب على المشكلات السابقة
 أن تقوم المدرسة ببرنامج للاختبارات : وستناول هذا بالتفصيل في
 الفصل القادم :

الفصل الرابع عشر

برنامج الاختبارات فى المدرسة

- مقدمة .
- وظائف البرنامج .
- اشتراك جهاز المدرسة فى تخطيط البرنامج .
- خصائص برنامج الاختبارات الجيدة فى المدرسة .
- برامج مقترحة للاختبارات فى المراحل التعليمية المختلفة .
- السجل المجمع .
- جهود وزارة التربية والتعليم .
- عرض نتائج برامج الاختبارات على الجمهور .
- صعوبات برامج الاختبارات فى المدرسة .
- خطورة برنامج الاختبارات على البرامج التعليمية .
- الاختبارات التحصيلية المستخدمة فى برامج الاختبارات
- الاختبارات النفسية فى البرامج .

الفصل الرابع عشر

برنامج الاختبارات في المدرسة

مقدمة :

الخطوة الأولى في تخطيط برنامج للاختبارات هي اكتشاف أى أنواع المعلومات عن التلاميذ ترى هيئة التدريس أنها تحتاجها واكتشاف كيف تستخدم نتائج الاختبار . فقبل أن نسأل : « أى الاختبارات يجب أن نعطي ؟ » علينا أن نسأل : « ما هي المعلومات ، التي ليست لدينا ، ولا تحتاج إليها ؟ » ، « ومتى نحتاج لها ؟ » ، « ومن سيستخدمها ؟ » . فنقطة البدء هي المدرسة ومناهجها ، وهيئة التدريس وحاجاتها . على أن الذى يهمنا ، قبل معرفة أى الاختبارات نطبق ، هو أن نعرف كيف ستستخدم الاختبارات . فالسؤال الأول إذن : لأى الأغراض ستستخدم الاختبارات في مدرستك ؟ إن تحديد الوظائف والأغراض هو الحصان الذى يجر عربة برنامج الاختبارات .

ملخص وظائف البرنامج

وهكذا نجد أن برنامج الاختبارات في المدرسة له عدة وظائف يمكن تلخيصها فيما يأتى :

وظائفه فى الفصل	وظائفه التوجيهية	وظائفه التطبيقية
١ - تضيف التلاميذ لتلقى	١ - إعداد الدليل	١ - تكوين وتعيين
التعليم داخل الفصل .	لتوجيه المناقشات	جماعات الفصول :
	مع الآباء عن أبنائهم	

وظائفه في الفصل	وظائفه التوجيهية	وظائفه التطبيقية
٢ - توجيه تخطيط نشاط تلامذة معينين .	٢ - أن يبني التلاميذ صورة الذات .	٢ - وضع التلاميذ الجدد .
٣ - تعيين التلاميذ الذين يحتاجون إلى دراسة تشخيصية خاصة وتعليم علاجي .	٣ - إمداد التلاميذ بفرص الاختيار الحالية .	٣ - المساعدة في تحديد اللياقة للجماعات الخاصة .
٤ - تحديد المستويات التحصيلية المعقولة لكل تلميذ وتقييم التفاوت بين القدرة الكامنة والتحصيل .	٤ - مساعدة المدرس المرشد وتطوير فهم الآباء للحالات المشكلة .	٤ - المساعدة في تحديد أى التلاميذ سينقلون إلى فرق أعلى .
٥ - تعيين الدرجات في المنهج .	٥ - تقييم المنهج الدراسي ونقط التأكيد والتجارب على المنهج الدراسي .	٥ - تقييم المدرسين .
		٦ - تحسين العلاقات العامة .
		٧ - تحسين العلاقات العامة .
		٨ - إمداد الهيئات الخارجية بالمعلومات .

أولاً : وظائفه في المدرسة

١ - تقسيم التلاميذ إلى جماعات متجانسة للتعليم داخل الفصل : تنفيذ

الاختبارات المقننة في معرفة الأفراد ذوى المستوى المتقارب من المهارة الذين يوضعون ، بالتالى ، فى مجموعات تعمل معا على نفس المواد وبنفس السرعة . والاختبارات المقننة تقدم المعلومات بسرعة وبطريقة موضوعية فى بدء العام الدراسى وتمكن من الوصول إلى أقصر الطارق وأسرعها وإلى التعرف على قدرات كل تلميذ ومهاراته . واختبارات القراءة تؤدى هذه الوظيفة فى الفرقة الأولى ، والاختبارات التحصيلية تؤدىها فى المستويات الأعلى .

٢ - توجيه تخطيط نشاط تلاميذ معينين : بهذه الطريقة نجعل العمل بموضوعات المهارة يناسب المستوى الراهن لكل تلميذ . فيشجع التلميذ الموهوب على التحرك قدما إلى الأمام بمعدل سرعته هو ، ويزود بأوجه النشاط والإمكانيات التى تعينه على هذا . ويسمح للتلميذ ذى التحصيل المحدود بأن يتحرك ببطء نحو أهداف أكثر تحديداً . وفى هذا النوع من التخطيط نستعين بمقاييس الاستعداد المدرسى ومقاييس التحصيل التعليمى .

٣ - تعين التلاميذ الذين يحتاجون إلى دراسة تشخيصية خاصة وإلى

تعليم علاجي : عندما تكون هناك أسبابا تجعل النظام المدرسى محتاجا إلى

نوع من الدراسة الشخصية الخاصة لأفراد من التلاميذ وإلى مدرسين معينين كى يوفرأ علاجا تعليميا ، هنا يستطيع برنامج الاختبار أن يقدم بيانات هامة تساعد فى تعين التلاميذ الذين يحتمل أن يفيدوا أكثر من غيرهم من ذلك التعليم . وإذا ترك لمدرس الفصل أن يختار التلاميذ لمثل

هذه الخدمات الخاصة ، تعرض للوقوع في خطأ اختيار التلاميذ الذين يعتبر مستوى تحصيلهم منخفضاً جداً . فإنه يصعب عليه أن يفرق بين القدرة العامة المنخفضة وبين نقص خاص في مهارة محددة معينة . وبرنامج الإختبار الذى يقدر كلا من التحصيل والاستعداد يساعد في انتقاء هؤلاء التلاميذ ذوى النقص الذين يمكن علاجهم .

٤ - تقييم التفاوت بين القدرة الكامنة والتحصيل : اختيار تلاميذ للعلاج التعليمى هو جزء من مشكلة عامة هى تحديد مدى التفاوت بين التحصيل الممكن والتحصيل الفعلى . هذا التفاوت يساعد المدرس على أن يركز جهوده على تلاميذ معينين في جماعته الخاصة ، ويساعد في توجيه المناقشات في مؤتمر مع الآباء ، أو قد يؤثر على التقدير في الفترات أو الدرجات الشهرية المرسلة إلى البيوت . والهدف الأخير يهتم المدرسة إذ أن بطاقة التقرير تحاول أن تقيم تحصيل التلميذ الفرد في علاقته باستعداده الكامن . وهنا نحتاج إلى مقاييس جيدة لكل من القدرة الكامنة والتحصيل .

٥ - تقدير الدرجات في الفرقة الدراسية : يجب أن يعتمد الكثير من مدرسى المدارس الثانوية على نتائج الاختبارات المقتنة كجزء من تقييم أداء الطالب . وقد يلائم هذا مناهج معينة حيث الاختبارات المقتنة الموجودة توازى في محتواها أهداف تعليم المنهج . وعلى أى حال فبرنامج الاختبار على نطاق المدرسة أو المدينة لا يناسب هذا الغرض تماما . فمحتوى الاختبار المقتن هنا هو محتوى عام لا يرتبط بتدريس مدرسة معينة أو فرقة دراسية معينة ، فهو يغطى مدى كلياً من المعارف والمهارات . ولهذا السبب ، فالإختبار المقتن لم يعد خصيصاً لقياس أشياء معينة تعلمها التلميذ في فصل معين أو في فترة تعلمية معينة :

ثانيا : الوظائف التوجيهية لبرنامج الاختبارات

١ - التقرير المقدم للآباء عن تقدم أبنائهم : على المدرسة مسؤولية إحاطة الآباء علما بتقدم أبنائهم : وهذه المعلومات المرسله يجب أن تتضمن ما إذا كان عمل التلميذ جيداً أم رديئاً : وهذا يساعد المدرس على تزويد تقريره للآباء بالأدلة العملية الموضوعية عن أداء أبنائهم ، مما يبعده عن أن ينظر الآباء إليه نظرهم إلى المتحيز المتعصب ذى الرأى الفردى الذاتى فقد أصبح مزوداً بالحقائق :

٢ - تكوين صورة - الذات بطريقة واقعية : لبرنامج التوجيه وظيفة عامة هى تحسين فهم كل فرد لقدراته ونواحى نقصه : فبرنامج الاختبار يقدم الدليل الموضوعى عن هذه النقاط من الضعف والتفوق حتى تفسر للتلميذ الفرد :

٣ - مساعدة التلميذ بفرص الاختيار الحالية : أمام التلميذ عادة عدة فرص يجب أن يختار بينها ، فعليه أن يقرر أن يدخل نوعا معينا من التعليم كالثانوى والفنى و . . : وبرنامج الاختبار يوفر دليلا يساعد التلميذ والمرشد النفسى فى إتخاذ هذه القرارات :

٤ - مساعدة التلميذ على تحديد أهدافه التعليمية والمهنية : يساعد برنامج الاختبار التلميذ فى تخطيطه البعيد للتعليم أو العمل : وصورة - الذات التى ساعد فى بنائها تفسير المرشد النفسى لنتائج الإختبار ستؤثر على الخطط .

٥ - تحسين فهم الحالات المشككة : فى كل مدرسة تلاميذ مشكلين من الناحية التعليمية أو من حيث التوائى الاجتماعى : : : الخ وبرنامج

الاختبارات المنتظم ، بالإضافة إلى النظام الجيد في التسجيل ، يقدمان بعض الأسس التاريخية لفهم المشكلة الراهنة فلكل مشكلة حالة ماضية .
كما يساعد البرنامج في توفير تسهيلات أكثر تعمقا للطفل المشكل .
وهنا نجد العديد من الاختبارات التي تناسب حالات فردية وإن وجب أن تتفق مع البرنامج الموحد المطبق على كل التلاميذ .

ثالثا الوظائف التطبيقية لبرنامج الاختبارات

١ - تكوين وتعيين جماعات الفصول : إذا كانت المدرسة واسعة بما يكفي لتكوين عدة جماعات في مستوى الفرقة الدراسية الواحدة أو عدة فصول للمنهج المعين ، وجب أن يتخذ قرار على أسس عميقة تتعلق بمن *who* يذهب إلى أى *which* جماعة . والتجميع يصبح فعالا في حالة المناهج الخاصة المفردة مثل ، الرياضة والإنجليزى ، ... الخ . ويجب أن يتخذ القرار على أساس الاستعداد أو المستوى الأولى *initial* للتحصيل .

٢ - وضع التلاميذ المحولين من مدارس أخرى : عندما يحول تلميذ من مدرسة إلى أخرى ، يجب أن يتخذ قرار بشأن مستوى الفرقة الدراسية والجماعة التي سيلتحق بها . والاختبار هنا يجب أن يضع في الاعتبار المستوى التحصيلي في تحديد الفرقة أو يعدل التعلم في جماعة الفصل حسب الحاجات المعينة للتلميذ الجديد .

٣ - المساعدة في تحديد اللياقة لمجموعات خاصة : قد تكون هناك

فصول للتلاميذ بطيئ - التعلم وهنا قد يكون القبول في هذه الفصول على أساس نقص نسبة الذكاء عن مستوى معين . وقد تكون هناك فصول تدرس مواد كالجبر تتطلب نسبة ذكاء فوق مستوى معين أو تحصيل حسابي معين أو درجة معينة على اختبار تنبؤي معين .

٤ - المساعدة في تحديد أى التلاميذ يجب أن يتقنوا : التقدير الموضوعى

للتحصيل والاستعداد فى اختبارات مقننة يقدم عوناً للمدرس كى يحكم على تقدم الطفل مقارناً بتقدم المدرسة حتى يتمكن من نقل الطفل إلى فصل أعلى .

٥ - تقييم المنهج ونواحي التأكيد فيه والتجارب عليه : إن تطبيق

نتائج الاختبارات لتقييم منهج دراسى محلى أو التجارب فى منهج دراسى معين يجب أن يكون عادلاً . والاختبارات المقننة تقيم فقط جزءاً من الأهداف التى توصل إليها نتائج تدريس المنهج . وتغطية بقية الأهداف تتطلب طرقاً غير رسمية فى التقييم فيجب أن تفسر نتائج الاختبار فى ضوء الصورة الكلية للأهداف الواضحة وغير الواضحة . وعلى أى حال ففى أى برنامج ذى منطق تجريبي يكون من المهم تحديد مدى الكفاية التى استوعبت بها المهارات العامة الأساسية .

٦ - تقييم المدرسين : يضع بعض النظار تقييمهم لجدارة المدرسين

أفراداً على أساس نتائج الاختبار المقنن ولكن هذا التقييم وحده لا يعتمد عليه ويعبر التربويون الأمريكيون عن هذا بقولهم objectives are more than what the test measures.

٧ - تقييم المدرسة كوحدة : قد يمد برنامج الاختبار ناظر المدرسة

بنوع من محك الجودة التعليمية . فإن الجداول الملخصة للمدارس والفصول ببرنامج الاختبارات ننظر لها فى ضوء الاستعداد والخصائص الرئيسية لكل جماعة مما يوضح نقاط القوى والضعف للمدارس والفصول فالمعلومات قد تدل على أن أماكن معينة أو أشخاص معينين يحتاجون إلى مساعدة خاصة أو إلى تغيير فى النقاط التى يركزون عليها . وياحبذا لو كانت الاختبارات

تهدف إلى المساعدة لا الحكم سواء على المدرسين أم المدارس وإلا أصبحت أداة هدم

٨ - تحسين العلاقات العامة : قد يرى الآباء أن المدارس تهدف إلى تعليم أبنائهم القراءة والكتابة والحساب ولكن هذه النظرة قاصرة ونتائج الاختبارات الموضوعية تصحيحها فهي تعرض المهارات والوظائف القريبة والبعيدة التي خدمتها البرامج . فإذا أعلن أن المدارس نجحت في هذه الأهداف أمكن أن تتقدم في نفس الاتجاه . وتزيد أهمية وظيفة الاختبار في العلاقات العامة عندما تلجأ المدارس إلى البعد عن الطريق المألوف المعتاد وإن كانت تصل إلى نفس الهدف بطريقة أثبت العلم أنها أفضل . فهذه الابتكارات تجعل الناس يعترضون أن الأولاد لم يتعلموا المهارات الأساسية . وعلى أى حال فالاختبارات توضح مدى صدق هذه المزاعم ونتائج الدراسات عن تحصيل المدارس في المجتمع المحلي ، أو تقييم تغيرات المنهج ، يجب أن توضع في أيدي المدرسين حتى يمكنهم أن يواجهوا انتقادات المجتمع وأن يقدموا صورة صادقة عن التحصيل في المدارس .

(١) إمداد الهيئات الخارجية بالمعلومات : هذه الهيئات قد تكون اجتماعية أو توظيفية أو مدارس أخرى ومعاهد وبرامج الاختبار معضداً بنظام كاف في السجلات يقدم المعلومات المطلوبة بطريقة مقننة . هذه المعلومات الموضوعية تتضمن أيضاً الاستعداد والتحصيل .

اشتراك جهاز المدرسة في تخطيط البرنامج

نحن نعتقد أنه في النظام المدرسي يجب أن تكون هناك مشاركة واسعة النطاق في التخطيط الأصلي لبرنامج الاختبارات والمراجعة بين الحين والحين . فإدات الفئات الأساسية من مستخدمي الاختبار هي : (١) الإدارة والإشراف . (٢) التوجيه ، (٣) مدرسي الفصول . لهذا نعتقد أنه يجب أن يشترك

يمثلون من الفئات الثلاث في التخطيط فيجب أن نضع في الاعتبار حاجات كل جماعة منها والزمن اللازم للإجراء والطريقة التي ستعالج بها نتائج الاختبار . ولا يقتصر الاشتراك على تطبيقه وحسب بل والمراجعة الدورية لمعرفة أحسن الطرق لاستخدام نتائج الاختبار :

خصائص برنامج الاختبارات الجيد في المدرسة

١ - اعتبار الوظائف : يجب في تخطيطنا أن نضع في الاعتبار الوظائف السابقة والوظائف المحتملة الأخرى فعلى هذا الأساس نختار الاختبارات ونحدد مواعيد تطبيقها بحيث نجتمع المعلومات المطلوبة حديثة بقدر الإمكان . فمدرس رياض الأطفال مثلاً يحتاج إلى تقسيم الفصول إلى جماعات فرعية على أساس تفاوت سرعات القراءة وهنا يجب أن يطبق اختبار الاستعداد للقراءة في أكتوبر للسنة الأولى ثم يعاد في مايو . كما أن التلاميذ والمرشدين النفسانيين في السنة النهائية في المدارس الثانوية يضعون قراراتهم التي يجب أن تبنى على المعلومات الموفرة عن استعدادات التلاميذ وتحصيلهم ومن المهم هنا أن نحدد ما هي الاختبارات التي سنستخدمها ومتى سنستخدمها .

٢ - التكامل : يجب أن ننظر إلى برنامج الاختبار ككل فربط نتائج الاختبار في السنة الخامسة بنتائجه في السادسة بها في السابعة كذلك يجب أن يكون كل نوع من المعلومات المجموعة قد جمع بطريقة وفي وقت يجعله يخدم هذا الكل إلى أقصى حد . وبإيجاز لو استخدمت مجموعة من الاختبارات تصلح لقياس التقدم في المهارات الأساسية ذلك في أكبر مدى ممكن وهكذا يكون من الممكن مقارنة الدرجات في الفرقة المعينة بالدرجات في الفرقة التالية على أساس أن المحتوى عام . وهكذا تكون الصورة عن نمو التلميذ أكثر صدقاً . كما يجب أن نضع في الاعتبار أثناء قياس المدارس الإعدادية

ما قسناه فعلا في الابتدائية وما سنقيسه في الثانوية . فإذا طبقنا اختبار ذكاء في الفرق السادسة وجب تطبيق اختبار ذكاء مماثل في السابعة . وإذا طبقت اختبارات استعداد في التاسعة وجب تطبيق بطارية مماثلة في العاشرة .

كما يجب أن يتبع السجل الطالب في تحركاته من المدرسة الابتدائية إلى الإعدادية فالثانوية ، وتسجل المعلومات في السجل بطريقة فيها استمرار حتى يكون كاملا بقدر الإمكان . والتكامل أخيرا يعني تحديد الأوقات التي ستجرى فيها الاختبارات بحيث يغطي كل الأغراض ، فاختبار الاستعداد المطبق في السنة العاشرة مثلا يفيد في التنبؤ بالنجاح في المعهد .

٣- الإستمرار : الاستمرار في البرنامج له ميزتان . فمن ناحية تتجمع البيانات في السجلات عن التلميذ الفرد فالبيانات عن نتائج الاختبارات التي طبقت على نفس التلميذ منذ سنة أو أكثر تلقى ضوءاً على مركزه الراهن فهكذا نرى ما إذا كان الولد المشكل قد عرض من قبل لانتجائه مماثل أو أن المشكلة ليست جذرية والفائدة الأخرى هي أن استمرار استخدام اختبار أو سلسلة من الاختبارات محليا يسمح بإنماء مستويات توقع محلية ، وقد يأخذ هذا اتجاها غير رسمي للمعايير القومية في تفسير الأداء المحلي كما قد يأخذ صورة خاصة من المعايير المحلية وعلى ذلك فإذا كان أحد أحياء المجتمع المحلي يتمتع بنسبة مثوية عالية من تلاميذ المعاهد فقد نجد أن المئينات المبينة على المجموع السكاني بمدارس الحى تقدم إطارا أكثر ملاءمة للحكم على المكانة الأكاديمية لتلميذ من تلاميذ الحى أكثر مما يلائمه العمر القومى أو معايير الفرقه .

ومعرفة الاختبارات تتضمن معرفة ما نقيسه كما تتضمن معرفة المستويات المحلية للتوقع . فإذا استخدم المدرسون اختبارات وأفادوا من نتائجها فإنهم سيعرفون ماذا تغطيه هذه الاختبارات وما هي المفاتيح التي

يمكن أن تشتق منها لتشخيص نواحي القوى والضعف في الجماعة . كما يعرفون ما هي أوجه نقص هذه الاختبارات : ولهذا يحسن أن يستخدم النظام المدرسي في استمرار نفس هذه الاختبارات لفترة من السنين ويتحول عنها فقط عند ما تصبح غير صالحة أو عند ما تثبت الدراسات أن هناك اختبارات أفضل .

برامج مقترحة للاختبارات في المدرسة

١ - برنامج للمدرسة الابتدائية والإعدادية : يتركز اهتمامنا في هاتين المدرستين على مساعدة الفرد على التمكن من وسائل التعليم والاتصال أثناء تعلمه كيف يعيش ويعمل في جماعة من رفاقه . وفي نفس الوقت تتكون للتلميذ فردية أساسية من القدرة والمعرفة والفهم في مستواه الخاص . فالاختبارات يجب أن تغطي هذه الوظيفة في هذا المستوى .

ويرى البعض أن من اللازم اختبار الأطفال في الإبصار والسمع ونحن نعتقد أن هذه المقاييس جزءاً من الفحص البدني وليست جزءاً من برنامج الاختبار المدرسي . ولكننا لا نشك في قيمة الإبصار والسمع وأثرهما في التحصيل خاصة في حالة من لديهم عيب في أحد الحاستين .

(١) اختبارات القراءة : تعتبر القراءة الطريق الأساسي لكسب كل أنواع المعرفة المنظمة وإن قلت قيمتها في عالمنا اليوم بسبب انتشار السينما والإذاعة والإذاعة المرئية فالتعلم من الكتب بالرغم من كل هذا سيصبح لب التربية خاصة في المستويات العليا ولهذا يجب أن نشخص منذ البداية العجز القرائي . هذا ويجب أن يطبق اختبار القراءة كل عام أو عامين ابتداءً من السنة الأولى في المدرسة الابتدائية حتى نهاية المرحلة الإعدادية .

(ب) اختبار الذكاء الجمعي : للمساعدة في تفسير نتائج اختبار القراءة والجوانب الأخرى من التحصيل الأكاديمي ، وللمساعدة في وضع مستويات للتلاميذ ، وللمساعدة في فهم الحالات المشككة ، تستخدم نتائج قياس الذكاء : ولاعتبارى الوقت والتكاليف يفضل الاختبار الجمعي وإذا كان سيطبق على السنوات الأولى الابتدائية يجب ألا يكون من نوع الذكاء القرائي ولكن الاختبارات في هذا المستوى نادراً ما تعطى نتائج ثابتة :

(ج) بطارية المهارات الأساسية : وهي تتضمن القراءة وتحل محل اختبار القراءة المستقل : وهي تفيد في تخطيط البرامج الفردية في التعليم وفي توزيع التلاميذ على البرامج الخاصة في التعليم العلاجي .

ن (د) اختبار استعداد القراءة : قد ننظر إلى هذا الاختبار على أنه دليل لمدرس الفرقة الأولى في تنظيمه لفصله إلى جماعات فرعية لتعليمهم القراءة وباعتباره أساساً يساعد المدرس على تقييم تقدم التلاميذ أفراداً .

الأنواع الأربعة السابقة من الاختبارات هي أكثرها شيوعاً في هذين المستويين ولكن هناك اختبارات أخرى تستخدم أحياناً . وترجع قلة استخدامها دائماً إلى التكاليف (الوقت ، والجهد ، والمال) فضلاً عن أنها تخدم وظائف تبدو أقل أهمية : ومن هذه الاختبارات نذكر :

(أ) اختبارات الذكاء الفردية : إذا كانت الإمكانيات كافية والفريق مدرباً كانت أصح من الجمعية خاصة في السنين الأولى .

(ب) اختبارات التحصيل في محتوى المواد الدراسية : بعض الاختبارات التحصيلية في السنوات الإعدادية النهائية تعالج محتوى مجالات العلوم والأدب والدراسات الاجتماعية : وبعض الاختبارات التحصيلية تنطبق على مدارس معينة ولا تصلح كجزء من برنامج الاختبار في المدارس الإعدادية :

(ج) أنواع أخرى من المقاييس : وهى قد تستخدم فى حالة تلاميذ معينين كالاختبار التشخيصى إذا كان الطفل متأخرا فى مادة معينة والاختبارات الإكائينية كالإسقاطية التى يستخدمها الأخصائى النفسى الإكلينيكى فى دراسة الحالات المشككة .

ونحن لم نوص باستخدام استبيانات الشخصية من نوع الورقة والقلم بسبب عيوبها فى : (١) صدق المعلومات (٢) عمق التفسير . فإننا نرى أن فهم الطفل كشخص يجب أن يعتمد على الملاحظات المباشرة وغير الرسمية التى يقوم بها المدرس وبقيّة هيئة المدرسة :

٢ - برنامج الاختبار فى المدرسة الثانوية : فى المدرسة الثانوية يصل التلميذ إلى نقطة يواجه فيها عدة فرص تعليمية وقرارات يختار بينها فقد يختار مواد معينة أو مناهج مختلفة ويبدأ فى إعداد خطط المستقبل : هل سيبقى فى المدرسة ؟ أم سيذهب إلى معهد أو ما هى الخطط فى عالم العمل خاصة وأن المنهج الدراسى أصبح فى هذه المرحلة يغطى المهارات الأساسية فضلا عن العديد من المحتويات الخاصة ونحن نعطي الأولوية فى هذا المستوى للاختبارات الآتية :

(١) اختبار الاستعداد المدرسى : يجب أن يتخذ كل فرد قرارا خطيرا يتعلق بالمسافة التى سيقطعها من السلم التعليمى وما إذا كان سيتحول إلى التعليم التجارى أم الزراعى أم الصناعى أم أنه سيواصل الدراسة فى معاهد . واختبار الاستعداد المدرسى يساعد فى الوصول إلى هذه القرارات وله قيمته فى تفسير مدى تقسّم التلميذ مدرسيا وفى موضع مستوى توقع التلميذ .

(ب) اختبار قراءة : لا فائدة من اختبار القراءة فى هذا المستوى ما لم تستفد المدرسة من نتائجه فى توجيه النشاط العلاجى : وإذا كان الأمر كذلك وجب تطبيقه فى أول هذه المرحلة :

(هـ) اختبارات الاستعدادات الخاصة : إذ كنا ننظر إلى التوجيه المهني

خاصة لمن لن يذهبوا إلى المعاهد العليا فإن برنامج الاختبار في المدرسة الثانوية يجب أن يتضمن مقاييساً للاستعدادات الهامة في النجاح في المهن مثل بطارية D.A.T. واختبارات الاستعدادات الخاصة لا تلائم السنوات قبل الثامنة أو التاسعة ويحدد وقت إعطائها بالبرنامج الكلي للتوجيه في المدرسة ويجب أن يطبق الاختبار في الفرقة التي سيتخذ فيها الطالب قرارات هامة .

(د) الميل Interests : الاستعدادات والميول أرضية أساسية

للتخطيط المهني ويمكن قياس الاثنين معاً باختبار مثل اختبار كودر .

(هـ) الاختبارات التحصيلية في محتوى المجال : وهي أقل من سابقها

في مجال التوجيه وإن كانت ذات قيمة في تقدير إمكانية مواصلة التعلم .

(و) الاختبارات التنبؤية : تفيد في مواد معينة مثل الجبر واللغات

الأجنبية والاختزال . وتستخدم في تصفية التلاميذ قبل أخذ هذه المواد فعلاً .

(ز) قوائم الشخصية والتوافق : تفيد في أن يكون فريق الإرشاد

أرضية أساسية جيدة وتكون هذه القوائم في أيديهم أدوات تصفية .

٣- برنامج الاختبار في المعاهد : تصبح البرامج هنا متنوعة

ومتخصصة بعد أن تكونت المهارات الأساسية . وليس المهم هنا هو التعليم بقدر التوافق والاختيار المهني كمشاكل أساسية لدى التلاميذ .

(أ) الاستعداد المدرسي : مفيدة جداً خاصة بالنسبة للطلبة الذين

يعانون من تحصيل غير مرض لإرشادهم على أساسها . فالتقدير المنفصل للقدرات اللفظية والعديدية له قيمة كتقدير تشخيصي . ويجب أن يعطى هذا الاختبار عند بدء المرحلة . ويحسن استخدام بطارية لا اختبار .

ولسنا هنا بحاجة إلى اختبارات الاستعدادات الخاصة لأن المجال المهني أصبح أضيق من ذي قبل بعد أن قرر التلميذ دخول معهد معين فهو ليس إذن بحاجة إلى اختبارات ميكانيكية ومكانية وكتابية .

(ب) القراءة : هنا نقيس بهدف فهم مشاكل الفصل الأكاديمي :
ففي المعاهد أحياناً عيادات قراءة حيث القراءة علاجية . ولهذا فالاختبار يجب أن يطبق في أول المرحلة لتحديد من سيقون عوناً خاصاً في تعليمهم :
(ح) الميول : هنا ثبتت الميول إلى حد ما واقترب الفرد كثيراً من الاختبارات المهنية النهائية ولذا يحسن تطبيق اختبار الميل في مهن معينة مثل سترونج .

(د) اختبارات الوضع في مواد خاصة : تقوم معاهد معينة بعمل دراسات إعدادية على أساس قدرة التلاميذ الجدد . ومثل هذا التقسيم في الإعداد يوصل المحتويات والتعليم إلى الجماعات المتطرفة . وهنا تطبق اختبارات تقيس مجالا واحداً .

(هـ) قائمة التوافق : تفيد هنا أكثر مما في المستويين السابقين الأقل تضجاً وتعلباً خاصة إذا كانت هناك خدمات إرشادية جيدة . ولكن هناك اعتراض على القوائم والطرق الإسقاطية إذا طبقت جميعاً . فاختبارات الشخصية تفيد أكثر لو أتي العميل متطوعاً ، كما أنها تطلب تدريباً خاصاً في من يستخدمها . .

السجل المجمع

يجب أن تكون المعلومات المتجمعة عن كل طالب في نهاية الأمر بصورة منظمة بحيث تناسب السجل المجمع Cumulative Record وهو يتضمن كل أنواع المعلومات عن التلميذ وهو أهم ما يفيد المدرسة في نشاطها التوجيهي للتلميذ وفيه مسئوليتها عن تزويد المعاهد أو أصحاب العمل

بالمعلومات متى يطلبونها . والأغلب أن كل مدرسة تحتفظ ببعض السجلات التي تجمع المعلومات عن الطالب طوال سنى بقائه بها ولكن النظام الحالي لهذه السجلات ليس فعالاً بقدر ما نود فهو لا يسمح بإمداد المعلومات النافعة في توجيه التلميذ في كل مراحل عمله المدرسي .

وسجلات التجميع شأنها شأن برنامج الاختبار هي أمر محلي يعتمد على الحاجات والإمكانات المحلية ، فأنواع السجلات التي تحتفظ وطريقة ومكان حفظها ، والأشخاص الذين سيحفظونها كل هذه أمور تتقرر بناء على أغراض وتوجيه المدرسة المحلية . وعلى أنواع المعلومات التي جمعت عن التلاميذ وعلى طرق استخدام هذه المعلومات على إمكانيات التوجيه والمستخدمين المكتبة . . وليس المهم في هذه السجلات ما نضعه فيها ولكن ما يستنتج منها . كما يهمننا معرفة من سيستخدمها وما هي المعلومات المطلوبة وفي أى الأغراض تستخدم كما يهمننا أن نعرف الأسس العميقة عن محتواها وتكوينها ووضعها وخدماتها .

اعتبارات في تخطيط البطاقة المجمعة : (أ) ما هي المعلومات التي تتضمنها (ب) في أى صورة تحفظ السجلات (ج) أين ستحفظ السجلات (د) ما هي وظائف الأفراد المختلفين الذين سيعيدون ويستخدمون هذه المعلومات .

(أ) محتوى السجل المجمع : أهم أنواع المعلومات فيه ما يأتى :

- ١ - نسبة الحضور والمعلومات المتعلقة بهذا ؟
- ٢ - تقديرات المدرسين للتحصيل الأكاديمي والدرجات التي أعلنت للتلاميذ أو لآبائهم .
- ٣ - سجلات كل الاختبارات المقننة .
- ٤ - سجلات متسلسلة وتقديرات غير رسمية أخرى عن التلميذ .

- ٥ - الفحص البدني وبيانات عن الصحة .
- ٦ - سجلات عن الاشتراك في الرياضات والأندية وأوجه النشاط خارج المنهج والعمل .
- ٧ - معلومات عن البيت والأسرة .
- ٨ - معلومات عن اهتمامات التلميذ وخططه وأهدافه .
- ٩ - المكانة المالية .

(ب) صياغة السجل المجمع : يجب أن نوفق بين عدة أهداف

مختلفة وأحيانا متضاربة أثناء صياغة السجلات :

- ١ - نريد أن يتضمن السجل من المعلومات أكبر قدر ممكن .
- ٢ - وأن يتضمنها بطريقة يسهل قراءتها .
- ٣ - نريد مكانا لكثير من الفقرات .
- ٤ - نريد مكانا كافيا لكل فقرة منها .
- ٥ - نريد سجلا خفيفا يسهل ملؤه .
- ٦ - نريد نظاما من السجلات يمكن أن يستخدمه أشخاص مختلفون في أماكن مختلفة .
- ٧ - ولكنتا نريد المعلومات مركزه .
- ٨ - وهناك من يريد المعلومات مطولة وهناك من يريد لها مقتصرة :
- ٩ - والبعض يريد أن تختصر السجلات حسب الطريقة المكتوبة إلى أقصى حد .

والمعلومات الآتية يجب عادة أن تسجل فيما يتعلق بنتائج الاختبار المقنن :

- ١ - اسم الاختبار .
- ٢ - الصورة المعينة من الاختبار .
- ٣ - تاريخ تطبيق الاختبار .

٤ - الدرجة الخام .

٥ - نوع الدرجة المحولة .

٦ - الدرجة المحولة نفسها .

(ج) وضع السجلات المجمعة : وجدت السجلات لكي تستخدم ولهذا يجب أن توضع حيث يكون للجميع الفرصة المشروعة لاستخدامها والإفادة منها . في نفس الوقت يجب أن تحمي السجلات من شر أن يستخدمها أشخاص غير مسئولين . ومشكلة المعلومات السرية تحل بأن توضع في سجل خاص منفصل ويشار إليها في السجل الكبير برقم وهكذا فمن يريد الرجوع إليها يستدل عنها بواسطة الرقم . وتوضع السجلات الأصلية في مكان بحيث :

١ - يسهل أن تستخدم في أى وقت طوال اليوم الدراسي .

٢ - يكون هناك مكان للاطلاع عليها .

٣ - أن تكون تحت إشراف شخص يستطيع أن يتحمل مسؤوليتها .

(د) وظائف أعضاء الهيئة المختلفين : يجب أن يكون نظام بطاقات

التجميع عملاً تعاونياً يعتمد على كل مجتمع المدرسة . فيعلم التلاميذ أن البطاقات خططت لصالحهم وأنها جزء من برنامج التوجيه . فيقدم التلميذ جزءاً من المعلومات : عن البيت ، ونشاطه خارج المدرسة ، واهتماماته ، وخططه ، وأهدافه . ومدرس الفصل يجب أن يكون مساهماً ومستخدماً لبطاقات التلاميذ . فهو يساهم فيها بوضع بيانات تقييمه عن نمو الطالب ودرجاته في المواد الدراسية ويستخدمها ليكون فكرة أفضل عن تلاميذه أفراداً .

جهود وزارة التربية والتعليم

قامت إدارة البحوث الفنية بوزارة التربية والتعليم بالإقليم الجنوبي

سنة ١٩٥٧ بإعداد ونشر :

- ١ - البطاقة المدرسية لتلاميذ المرحلة الابتدائية .
- ٢ - البطاقة المدرسية لتلاميذ المدارس الإعدادية .
- ٣ - البطاقة المدرسية لتلاميذ المدارس الثانوية .
- ٤ - البطاقات المدرسية : وهو بمثابة كراسة تعليمات للبطاقات المدرسية الثلاث . والنواحي التي تشملها البطاقة هي :
 - (أ) البيانات الأولية : الاسم والسن والفرقة والفصل و . . الخ .
 - (ب) التحصيل المدرسي : المواد وتقديراتها في سنوات الدراسة المختلفة بهذه المرحلة من التعليم .
 - (ج) المواظبة في كل السنوات : بعذر مقبول أم بدون عذر أو بعذر غير مقبول .
 - (د) الناحية الصحية : قوة النظر وقوة السمع وسلامة النطق والحالة الصحية العامة .
 - (هـ) صفات الشخصية : الجدية في العمل ، الاعتماد على النفس ، تحمل المسؤولية ، القيادة ، إبداء الرأي ، روح التعاون ، الانطواء والانبساط ، والاتزان الانفعالي ، . . الخ .
 - (و) النواحي الاجتماعية : تكوين الأسرة ، عدد غرف المسكن ، ووسائل شغل الفراغ .
 - (ز) الاختبارات : نوع الاختبار ، اسمه ، التاريخ ، الفرقة ، النتيجة .
 - (ح) أوجه النشاط الحر : رياضي ، فني ، ثقافي ، واجتماعي .
 - (ط) مشكلات التلميذ .
 - (ي) الملاحظات العامة .
 - (ك) رأى المدرسة في توجيه التلميذ .

وهكذا نرى أن هذه البطاقات هي سجلات مجمعة مصممة لمواجهة الحاجة إلى جمع بيانات عن التلميذ الفرد تتوفر فيها الخصائص التي عرضناها .
 إلا أنها لا تستخدم في كثير من الحالات ومرجع عدم استخدامها أساساً إلى :
 ١ - وقت المدرس لا يتسع لهذا العمل .

٢ - عدد تلاميذ الفصل الواحد كبير .

٣ - ليس لها مكان ولا شخص مسؤول عنها في المدرسة .

٤ - لا يمكن - حالياً - تغطية الكثير من بياناتها .

عرض نتائج برامج الاختبارات على الجمهور

في عرضنا لنتائج الاختبار على الجمهور ، نهتم بعدة أنواع من المقارنات :
 ١ - مقارنة أداء الجماعة المحلية بالمعايير القومية أو بمعايير أكثر تخصصاً
 إن وجدت .

٣ - مقارنة التحصيل في المواد المختلفة .

٢ - مقارنة أداء الجماعة المحلية بمستوى القارة في الجماعة .

٤ - مقارنة المدارس المختلفة في النظام المدرسي الواحد أو بين فصول مختلفة من نفس المدرسة إن أمكن .

٥ - المقارنة بين الجماعات التي تعلمت بطرق مختلفة - التي استخدمت طرقاً مختلفة أو مواداً مختلفة .

٦ - المقارنة بين الجماعات في مستويات الفرق الدراسية المختلفة .-

٧ - مقارنة نفس الجماعة في أوقات مختلفة لتبين نمو التلميذ .

ثانياً : هناك ثلاث صفات يحسن أن نجدها في تفسير الدرجات للطلاب أو أييه :

١ - يجب أن يوضع التفسير في الإطار المرجعي للتلميذ المعين . وعليه ،

فإن نتائج الاختبار التحصيلي الممنون يجب أن تفسر على أساس ما نعرفه عن استعداد التلميذ وعن أهدافه التعليمية والمهنية .

٢ - يجب أن يوجه التفسير نحو الفعل الإيجابي البناء . فيؤكد على خواص الصفحة النفسية للاستعداد ، أو يوجه إلى الفعل العلاجي عند ما يكون التحصيل أقل مما نتوقعه على أساس الاستعداد . فيشير التفسير هنا إلى العلاج التعليمي أو المهني . ؛

٣ - يجب أن يعتمد التفسير على الحقائق ولا يكون عاطفياً كما يجب ألا يبدو أحكاماً تصدر على الفرد . فيجب أن تذكر نتائج الاختبارات والأدلة الأخرى بصدق ودقة ، ولكن بطريقة ودية مقبولة .

على أنه يجب دائماً أن يعلم التلاميذ نتائج الاختبارات التي طبقت عليهم . فإن الاختبار يكون ذا معنى ودلالة لدى التلاميذ إذا رأوا بعض نتائجه وفوائده واستخداماته . وأى تقرير يجب أن يكون معدلاً بحسب درجة نضج الجماعة . وعلى أى حال هناك عدة طرق أخرى لتقديم هذه البيانات نذكر منها :

١ - مدرس الفصل قد يقدم تقريراً ملخصاً بسيطاً وتفسيراً عن أداء الجماعة ككل . ويشير فيه إلى نواحي القوة والضعف . وتقام المقارنات مع المعايير العامة لمستوى الفرقة الدراسية . ويشار فيه إلى الطريقة التي تكون بها النتائج مفيدة لعمل الجماعة وللتخطيط لأفراد معينين في الجماعة :

٢ - ترسل الدعوة للتلاميذ لمناقشة الاختبارات مع المدرس . وفي المؤتمرات الفردية ، قد يعطى للتلاميذ صورة عامة عن دلالة نتائج الاختبار وكيف يرتبط بخطة التلاميذ .

٣ - وفي حالة كون نتائج الاختبارات ذات مضمونات هامة ، مثلاً وجد تفاوت كبير بين الاستعداد والتحصيل ، وتفاوت كبير بين أداء الاختبار

وما يعتقده المدرس عن التلميذ من صورة - الذات ، وقد يوجه المدرس الدعوة إلى التلاميذ ويناقشهم في نتائج الاختبار أفراداً وربما كان مع كل تلميذ والداه :

صعوبات برامج الاختبارات في المدرسة

مشاكل الصدق

١ - تمثيل المحتوى Sampling of Content : لا يكون الاختبار عادلاً في المعرفة الحقيقة في مجال كلي إلا إذا كان يضع أوزاناً متكافئة لكل المعالم الجوهرية في المجال . فاختبار التاريخ ، هذا النوع من التاريخ ، هو الذي تنوى المدراس تدريسه . والاختبار ، باعتباره عينة ، قد يمثل وحسب نواحي من المواد الدراسية ، وهنا تكون هناك فرص أمام بعض التلاميذ أو بعض المدراس أن تتفوق لأنه حدث ، مصادفة ، أن كانت متفوقة في مادة يتضمنها عدد من فقرات الاختبار .

وأحد المزايا الهامة للاختبار الموضوعي على اختبار المقال هي أن الأول يقلل احتمال أن يحصل التلميذ الضعيف على درجة عالية لأن اختبار المقالة يختار الأسئلة بطريقة مبتورة فالفرص فيه قليلة . فإذا حدث مثلاً أن الاختبار تضمن سؤالين يمثلان مجالين صغيرين من المجال الدراسي وحدث أن كان هذا التلميذ الضعيف يتقن الاثنين ، برغم ضعفه الشديد في بقية المجالات الأساسية ، نال التلميذ الضعيف درجات عالية في هذا الاختبار : بينما في الاختبار الموضوعي ، بسبب شمول تمثيله للمجال المقاس ، يحوى فرصاً متعددة إذا عرف التلميذ الضعيف إحداها أو بعضها فلن يعرفها كلها وعلى هذا يقل احتمال حصوله على درجة لا تمثل مستواه فعلاً . فإذا سألت مدرساً عن الغاية من تدريس مادة ما . فإنه سيعدد لك

أهدافاً كثيرة . فمثلاً مدرس الهندسة ، قد ينكر أن الغرض من تدريس الهندسة هو توصيل عدد معين من النظريات إلى أذهان التلاميذ وتوصيل بعض الأسس التطبيقية وبعض المهارات في استعمال الفرجار والمنقلة . قد ينكر كل ذلك ويؤكد بدلاً عنه أن الغرض هو تكوين عادات فكرية ومهارات في تحقيق الفروض ، والحذر في تقبل استنتاجات غير محققة . ولكن لو فحصنا الاختبارات الموضوعية التي يضعها هذا المدرس نفسه فإننا نجد أنها تؤكد الأهداف الأولى التي أنكرها وتهمل الأهداف التي ذكرها .

وقد عالجنا الاختبارات التحصيلية الحديثة هذا النقص الخطير بدراسة المناهج الدراسية وحصرها بدقة واعتبارها عند تصميم الاختبار عينة لها ؛ إذ أن الاختبار ، باعتباره عينة ، يجب ، وهذا ما حدث في الناجح منها فعلاً ، يمثل المجال المقاس بكل محتواه وأبعاده بحسب نسبها وأوزانها ؛ ذلك إما بوضع أسئلة كثيرة تغطي كل العوامل الأساسية في المجال بحيث يخصص عدد أكبر من الفقرات للسؤال عن العوامل الأكثر أهمية . أو يخصص عدد متساوٍ من الفقرات للسؤال عن مختلف العوامل التي تكون المجال المقاس ، ولكن تعطى أوزان وقيم أكبر للعوامل الأهم .

ومثل هذا الاختبار سيعطي درجة منخفضة للفصول التي تركز على موضوع أو اثنين وإن تفوقت فيها تماماً . ولكن تلاميذ هذه الفصول أقل تفوقاً في المجال الكلي كما يدرس في أغلب المدارس ؛

٢- صدق المنهج Curricular Validity الاختبار إذا كان عينة

ممثلة لمحتوى المجال يكون اختباراً صادقاً من وجهة واحدة . وقد يكون

غير صادق من وجهة أخرى . فعندما يستخدم الاختبار للحكم على فعالية الخبرة التعليمية ، يكون صدق المنهج فيه ضرورياً . والاختبار الذي يتمتع بصدق المنهج يقيس ، بكفاءة ، المدى الذي تعلمه من الوظائف التي يهدف المنهج الدراسي أن يعلمهم إياها .

٣ - قياس جوانب السلوك المناسبة Measurement of Appropriate

Behaviors لا يصبح أن تقتصر على قياس كل التحصيل عن طريق قياس المعرفة بالحقائق العلمية أو المعلومات أو مجموعة محددة من المهارات . إذ أن المدارس تختلف من حيث عدد وأنواع المناهج وحتى لو حدد العدد والنوع فإنها تختلف أيضاً من حيث الأهمية التي تعطيها للأنواع المختلفة . هذا فضلاً عن أن المدرسين أنفسهم يختلفون في وضعهم لفلسفة وأهداف تدريس مادة معينة . وعلى أي حال فإن أهداف التعليم المدرسي متعددة جداً وقد جمعها Tyler وزملاؤه في الآتي :

- ١ - المعلومات الوظيفية . أي ليست مجرد معرفة جامدة ، ولكنها معرفة يمكن تطبيقها على المواقف الجديدة التي تربط بها .
- ٢ - مهارات التفكير وعاداته .
- ٣ - الاتجاهات والحساسية الاجتماعية (التحمل Tolerance ، والروح العلمية في الاستقصاء ، وتفهم الموسيقى ، و . . الخ) .
- ٤ - الاهتمامات والأهداف والأغراض (الأهداف المهنية ، والاستخدام الحكيم لوقت الفراغ ، . . الخ) :
- ٥ - دراسة المهارات وعادات العمل .
- ٦ - التوافق الاجتماعي والشخصي .
- ٧ - الإبداع

٨ - الصحة الجسمية .

٩ - فلسفة وظيفية للحياة .

وللحصول على دليل عن نمو التلميذ متقدماً نحو تحصيل هذه الأهداف ، يجب أن نعرفها تعريفاً إجرائياً Operational Definition خلال أنماط السلوك التي تتعلق بها أو تتضمنها .

فمثلاً إن أردنا أن نقيس سلوك الطالب فيما يتعلق بقدرته على الاستنتاج من الحقائق العلمية فإننا قد نسأل أنفسنا عدة أسئلة منها أن نسأل : كيف يسلك شخص ذو قدرة عالية على الاستنتاج في مثل هذا الموقف ؟ أو أن نسأل : كيف يمكننا التعرف على التلميذ الذي له فلسفة وظيفية في الحياة ؟ ومن أهداف التعليم ما يسهل تحويله إلى صور سلوكية ومنها ما يصعب تحويله . ولكن إذا نجحنا في ترجمة الأهداف الأخيرة إلى صور سلوكية سهل علينا قياسها عن طريق الاختبارات وطرق الملاحظة . فمثلاً ، المهارة في تفسير البيانات تظهر في تصرفات خاصة . فلو أعطينا شخصاً ماهراً جدولاً أو رسماً بيانياً في موضوع لم يدرسه وجدنا أنه يميز الأفكار الرئيسية ويصرف نظره عن الاختلافات الجزئية التي قد تكون راجعة إلى فروق في العينة التي أتت منها هذه الرسوم . فيرى مثلاً أن العامل (أ) والعامل (ب) يتغيران معاً ولكن أحدهما لا يؤثر في الآخر . فمثل هذه المهارات يمكن قياسها مادماً قد استطعنا تعريفها عن طريق نواحي السلوك التي تنتج عنها .

وبعض اختبارات التحصيل تسمح لتغيرات لا تتعلق بجوهر ما نقيسه irrelevant أن تؤثر على درجة الفرد . فقدرة الفرد على القراءة تؤثر ، غالباً ، على مستوى إجاباته عن اختبارات التحصيل ، وهذه المشكلة يصعب تحاشيها ، ولكن لا يمكن الحصول ، مثلاً ، على مقياس صادق للمعرفة

إذا كان الشخص الذى يعرف حقيقة معينة لا يفهم الفقرة الخاصة بها بسبب صعوبات لفظية .

هناك عامل آخر يؤثر على صدق الاختبارات التحصيلية هو عامل السرعة . والإجراء الدقيق عادة يتطلب وضع حدود زمنية ، ولكن عندما تصبح السرعة عاملاً غالباً فى تحديد درجة الشخص ، يزداد احتمال كون الدرجة لا تعبر عن قدرته بدقة . والسرعة تكون عاملاً مشروعاً فى الاختبارات التحصيلية فقط إذا كان هدف المنهج الدراسى فعلاً إعداد أشخاص يعتمدون على السرعة فى عمل معين . وهذا يصدق مثلاً فى حالة إعداد مهرة فى الحساب أو موظفى خزانة ولكنه لا يصدق عند قياس صحة الاستدلال أو سعة المفردات أو المعرفة العلمية لدى المفحوص .

وفى أغلب الاختبارات التحصيلية يكون توقف الاختبار على السرعة مشروعاً وحسب إذا استخدم الاختبار عملياً فى التنبؤ بالنجاح فى عمل يعتمد على السرعة ، أو إذا أيدت البيانات أن الدرجات على اختبارات السرعة ذات معامل ارتباط عالى جداً مع الدرجات على اختبار القدرة . واختبارات السرعة ذات صدق محدد عند وصف معرفة التلاميذ أفراداً ، ما دام هناك عدد قليل من التلاميذ المتفوقين يؤدى عمله ببطء .

٤- التعرف والاسترجاع Recognition and Recall لأن الاختبارات الموضوعية وحدها هى الاختبارات التى يمكن أن يعتمد عليها فى مقارنة التلاميذ المختبرين فى أماكن مختلفة ، أصبح من الضرورى استخدام طريقة الاختيار من عدة إجابات وغيرها من أسئلة التعرف . كان هذا مصدرأ لاهتمام مدرسين كثيرين شعروا بأن الاختبارات التى تتطلب استجابات حرة وحدها هى التى نستطيع أن نقيس ، بكفاية ، ما يقومون بتدريسه ؛ ويسهل قياس التعرف على الاستجابات الصحيحة عن طريقة

الفقرات الموضوعية . ولكن المدرسين يميلون إلى تفضيل اختبارات الاستجابة - الحرة عندما يكون هدف التعليم هو إنماء القدرة على الاسترجاع أو اختراع حلول جديدة . وقد أيدت البحوث دائماً الاختبارات الموضوعية ذات البناء المحدد جيداً :

هـ - الدقة الفنية Technical Accuracy يتضمن الكثير من مفاتيح الاختبارات المنشورة أخطاء وهذا عيب خطير يذهب بصدقها . هذه الأخطاء تنشأ عن عدم دقة الصياغة ، والغموض ، والتبسيط المحل للمشاكل الفنية حتى تناسب صيغة الفقرة - الموضوعية . وقد جمع دياموند Diamond نقداً لـ ١٦ اختباراً نشرت عن العلوم Sciences والنظريات . وصلت نسبة الخطأ في بعضها إلى ١٨ ٪ من الفقرات التي حوت حقائق ونظريات ليست صحيحة .

مشاكل المعايير

تستخدم الاختبارات التعليمية عادة في تقنينها معايير الفرق الدراسية Grade Morns ، فتحول الدرجة الخام للتلميذ إلى فرقة مكافئة grade equivalent . وعلى هذا الأساس يقال ، مثلاً ، أن فلاناً من التلاميذ المتفوقين في الفرقة السادسة تقع درجته في الحساب في « الفرقة التاسعة » أى أن درجته الخام تساوى الدرجة الخام للتلميذ المتوسط في الفرقة التاسعة في مادة الحساب ،

ولكن هذا التقرير مضلل ما لم نخط علماً بالعيوب الثلاثة الآتية :

١ - تنشأ معايير الفرق على عينات تختلف درجة تمثيلها للتلاميذ في الأمة بأكملها . فبعض قطاعات الأمة متفوق كثيراً على البعض الآخر ، والسبب اختلاف قدرة التلاميذ واختلاف كفاءة المدرسين ، واختلاف توسع الإقليم في الاهتمام بالتعليم واختلاف البرامج : والأساس الوحيد العادل لمقارنة

المدارس هو الحكم على كل مدرسة مقارنة بالمدارس من نمط مشابه وذات إمكانات مماثلة .

٢ - ليست المعايير norms مستويات standards . فمن الخطأ الشائع أن نفترض أن كل التلاميذ في الفرقة السادسة يجب أن يصلوا إلى معيار norm - الفرقة السادسة . ففي عينة التقنين يقع ٥٠٪ من التلاميذ دون المعيار . كما أن الاختبار يجب أن يصمم لقياس المستويات الراهنة في المدرسة وليس ما يجب أن يحصله التلاميذ تبعاً لأقصى جهد للمدارس وبأحسن طرق التدريس .

٣ - معايير الفرقة مبنية على وحدات صناعية غير متساوية من القياس . فقد يبدو أن تلميذاً معيناً ممتاز جداً لأنه وصل إلى « الفرقة التاسعة » في الجغرافيا بينما هو فعلاً في الفرقة السادسة . وفي الكثير من معايير الاختبارات لا يعد هذا الفرق كسباً ضخماً ، إذ أن التلميذ المتوسط تنمو قدرته بمقدار صغير بين الفرقة السادسة والفرقة التاسعة .

وليس الوصول إلى مستوى « الفرقة التاسعة » صعباً بنفس الدرجة على كل تلميذ الفرقة السادسة . فالتلميذ الذي يفوق مستوى فرقته في مادة معينة بمقدار « فرقتين » أحياناً يعد ممتازاً جداً ، وأحياناً أيضاً لا يعد ممتازاً جداً لأن نسبة كبيرة من فصله تستطيع الوصول ، أو وصلت فعلاً ، إلى ما وصل إليه في هذه المادة . كما أن التفوق بفرقتين في مادة معينة لا يعادل التفوق بفرقتين في مادة أخرى . وبروفيل Profile درجات الفرق يعطى صورة جد مختلفة من حيث نقاط القوة والضعف أكثر مما يفعل بروفيل الدرجات التي تقارن تلميذ الفرقة المعينة بزملائه في نفس الفرقة .

ويقترح كرونباخ Cronbach علاجاً لذلك العيب بأن توضع معايير الفرقة الدراسية في شكل درجات معيارية . وبعض الاختبارات المقننة الآن تستخدم الدرجات المئينية داخل كل فرقة دراسية .

خطورة برنامج الاختبارات على البرامج التعليمية

يقال إن القياس المقنن سلاح ذو حدين يحطم البرنامج التعليمي بقدر ما يساعده . والصعوبة الأساسية هي أن الاختبارات التحصيلية تحدد المنهج الدراسي . فيكرس كل من تلاميذ المعاهد والمدارس الكثير من وقتهم لمحاولة تحديد ما سيختبرون فيه حتى يتمكنوا من مذاكرة هذه المواد وإهمال ما عداها . ويستخدم التلاميذ طرقاً مختلفة في المذاكرة ويركزون على مواد مختلفة فيصلون إلى مستويات مختلفة من الإنجاز ذلك عندما تستبدل الاختبارات الموضوعية باختبارات المقال .

كما أن المدرس مدفوع بالاختبار إلى توجيه طاقاته إلى الطرق التي تملئها عليه الاختبارات . فمدرس الكيمياء مثلاً يعدل منهجه ليضغط التطبيقات والاستدلال ويتوسع في المعادلات الفنية إذا خشى أن يفصله ، وبالتالي تدريسه ، سيعبد فاشلاً عند ما يطبق اختبار مقنن في آخر العام يعتمد أسئلته على المعادلات . وهكذا نجد أن الاختبارات تحدد المنهج الدراسي بطريقة أو بأخرى . ولا تشجع الاختبارات التعليم السليم واتجاهات المذاكرة السليمة إلا عندما تقيس كل مظاهر السلوك الذي تهدف التربية إلى إنمائه .

وعندما استخدمت الاختبارات لتحديد من في أعضاء الجماعة أضعف من الآخرين أدت إلى الامتناع . فبرغم زيادة الدفع ، زاد التوتر . هذا فضلاً عن أن المدرسين كثيراً ما اتهموا لأن الاختبارات المقننة أظهرت أن تدريسهم رديء . ومن العيوب أيضاً أن النقل من فرقة إلى فرقة أحياناً يعتمد على نتيجة اختبار مقنن ، دون اعتبار لسجل التلميذ بالمدرسة :

ونأظر المدرسة الذى يدخل الاختبارات يجب أن يجعلها مساعدة للمدرسين والتلاميذ معا ، وطريقة تساعدهم على معالجة نواحي الضعف . ولكن إذا لم يكن حريصا ، فقد تثير القلق الذى يخفض الروح المعنوية فى المدرسة ويقلل فعاليتها .

هناك خطأ شائع فى تطبيق الاختبارات المقننة هو استخدام المعايير دون اعتبار ما نتوقه استدلالا من قدرة التلميذ . فإدام الذكاء يختلف من شخص إلى آخر ومن جماعة إلى أخرى فلا يمكننا أن نحكم على ما إذا كان التقدم كافيا إلا بمقارنة الشخص أو الجماعة بغيرهم ممن لهم قدرات مشابهة . وهناك طريقة شائعة تتم بها هذه المقارنة هى طريقة نسبة الإنجاز (Accomplishment Quotient. A. Q.) ونسبة الإنجاز هى النسبة المئوية بين « العمر التعليمى » إلى العمر العقلى . فإذا كان لدينا مثلا تلميذ سنه ١١ سنة وعمره العقلى ١٢ سنة وصل إلى معيار تحصيل تلميذ الحادية عشرة ، هذا التلميذ تكون نسبة إنجازه ٩٢ (هذا الرقم هو خارج قسمة 11×100)

هذه العملية لتحديد من هم « المتخلفون تحصيليا » و « المتفوقون تحصيليا » لها عيوبها التى نشأت عن تحديدات معايير الفرقة وعن الفروض الإحصائية غير السليمة هذا فضلا عن أن هذه التحديدات تفسر نسبة التحصيل وتؤيد فكرة أن المقارنة يجب أن تكون بين ما تعلمه التلميذ مقارنا بما تعلمه التلاميذ الآخرون الذين لهم استعداد أو قدرة أولية مشابهة . ومثل هذه المقارنات تستند عادة إلى المعايير المحلية . ولكن فى الإمكان إنشاء معايير على الأمة بأكملها من هذا النوع وقد قدم ترمبل وكرومباخ للمدرسين الخطوات التى بها يستطيع أن يضع مثل هذا التحليل .

الاختبارات للتحصيلية المستخدمة في برامج الاختبارات

يجب أن توصلنا هذه الاختبارات إلى إجابات حاسمة عن التساؤلات الآتية :

١ - متى يجب أن تستخدم الاختبارات التحصيلية المقننة ؟

٢ - ماذا يجب أن تقيس ؟

٣ - أى الاختبارات يجب أن يختار ؟

٤ - كيف يجب أن تستخدم النتائج ؟

وأهم هذه الأسئلة هو رابعها ؛ فالوظيفة الدقيقة للاختبار في المدرسة هي تحسين البرنامج التعليمي وهو قد يؤدي هذه الوظيفة بإحدى الطرق الآتية :

(أ) بتحديد الخبرات التي يحتاجها التلميذ •

(ب) بالإشارة إلى طرق أفضل للتدريس :

(ج) بتعديل وإنشاء اتجاهات لدى التلاميذ والمدرسين تسمح

بتحسين التدريس .

ويقول مك كون McConn « إن كل أهداف القياس بالمدارس هي

توجيه التلاميذ pupil guidance » وليس التوجيه عنده محدداً بالمهن ، بل يتسع ويصبح هادفاً إلى اكتشاف كل ما يتعلق بالتلميذ حتى يوفر له نمواً أفضل :

إذا اتفقنا على هذا استتبع الأمر عدة نقاط ، فإذا كانت الاختبارات بهدف الإرشاد أصبحت جزءاً من العملية التعليمية ووجب أن تكون طبيعية لا صناعية ، ولهذا كان اختيار وقت القياس أمراً هاماً ، فالقياس بعد وقته

الأنسب لا يفيد كثيراً ولهذا تلجأ مدارس كثيرة إلى تطبيق الاختبارات التحصيلية المسيحية في أول العام الدراسي .

وعندما تنطبق اختبارات مناسبة وتوضع النتائج في يدى كل مدرس ، تصبح أساساً سليماً لتخطيط أعمال السنة . وهكذا يصبح من الضروري أيضاً أن يعاد الاختبار في نهاية العام الدراسي لتقدير مدى التقدم الذى أحرزه التلاميذ . وبهنا أن يكون الاختبار تشخيصياً توجيهياً ولا يكفى أن يقدم لنا درجات فقط ، وإذا كانت الاختبارات مستخدمة بغرض التوجيه فهى إذن تستخدم للتلميذ وليس عليه . فتصبح وسيلة تعينه على معالجة نقاط ضعفه ، وتصبح سلاحاً فعالاً له يقنعه بالاشتراك فى برامج معينة أو تغيير عادات معينة والاختبارات هنا تقنعه أفضل من ضغط المدرس عليه . وفى مثل هذا القياس يجب أن يأخذ التلميذ الاختبارات لأنه يريد أن يعرف النتائج .

ويجب أن نهتم أكثر بنواحى معينة من الأداء – وبالنجاح فى فقرات معينة ، أو مهارات فرعية معينة فى الاختبارات المركبة . ويجب أن نهتم بنوع الخطأ وليس بمقداره :

كان استخدام الاختبارات التحصيلية فيما مضى على غير هدى لأنها أجبرت المدارس على تدريب التلاميذ لا تعليمهم . ولكن إذا أعطيت الاختبارات على ضوء فكرة كاملة عن ماضى التلميذ ، وإذا أريد بالاختبارات أن ترشده فى تقدمه المقبل ، فإنها ستكون خير معين .

الاختبارات النفسية فى البرامج

من العرض السابق يتضح لنا أن الاختبارات النفسية والتربوية معا هى وحدات برامج الاختبارات فى المدرسة ، وفى البرامج المقترحة يلاحظ القارئ أن الاختبارات التى حوتها تقيس وظائف ، سواء كانت تربوية

أو نفسية ، تتسق مع مراحل نمو التلميذ باعتباره فرداً أولاً ثم مستفيداً من برامج محددة مدرسية ثانياً . ومما يدعم استخدام الاختبارات النفسية في المدارس أن :

١ - التلميذ فرد وهو مفحوص كغيره ممن يجري قياسهم :

٢ - أغلب جماعات التقنين حوت تلاميذاً .

٣ - أننا عند ما نقيس التلميذ لا نحتاج وحسب إلى تقدير أثر التدريب الرسمي الذي ناله في المدرسة ، بل وأيضاً إلى تقدير :

(أ) عضويته في جماعات المدرسة (وأداة هذا القياس هي مقاييس العلاقات الاجتماعية والذكاء الاجتماعي والنزج الاجتماعي والسوسيوجرام)
(ب) صحته النفسية وتوافقه (وأداة هذا القياس هي اختبارات الشخصية ومقاييس الصحة النفسية)

٤ - أن نمو التلميذ يلزمنا بقياس :

(أ) قياس الذكاء والمهارات الأساسية والاستعدادات في سن المدرسة الابتدائية والإعدادية .

(ب) قياس الاستعدادات الخاصة والميول والشخصية والتوافق في سن المدرسة الثانوية .

(ح) قياس الميول والتوافق في سن المعاهد العليا والجامعة .

٥ - أن استعدادات تلميذ المدرسة الابتدائية تتضح في قدراته في نهاية سن المدرسة الإعدادية تقريباً . ومن هنا وجب قياس أثر التلمذة فيهما على هذا النمو وأثر هذا النمو على النجاح في المدرستين :

٦ - أن هذه القدرات تخصص كلما تقدم التلميذ في التعليم وتتضح الاستعدادات الخاصة والفارقة و . . الخ مما يدفعنا إلى قياس أثر الحياة في المدرسة على النمو وأثر النمو على النجاح المدرسي .

٧- أن المدارس تعد الأفراد للمهنة وهنا نستخدم اختبارات الكفاية واختبارات الحرف لقياس أثر التدريب الفني على كفاية الفرد :

٨- أن المدارس تعد الأفراد للحياة ، وإذا أضفنا إلى هذا أن أثر خبرات الطفولة والمراهقة على التكوين النفسى للفرد وتوافقه المقبل يزيد عن أثر الخبرات التالية كثيراً ، رأينا ضرورة قياس التوافق والشخصية في مجال المدرسة .

ولهذه الأسباب يتعين علينا ، بعد دراسة إنشاء الاختبارات الموضوعية واختبارات المقال ، أن نعرض لخصائص اختبارات التحصيل في الفصل القادم :

الفصل الخامس عشر

خصائص ومشكلات الاختبارات التحصيلية

- خصائصها .
- استخدامها .
- تحديداتها .
- صيغتها .
- معاييرها .
- وظائفها .
- أنواعها .

الفصل الخامس عشر

خصائص ومشكلات الاختبارات التحصيلية

خصائصها :

تعتبر الاختبارات التحصيلية أكثر أنواع الاختبارات المقننة شيوعاً ، والهدف الأساسي للاختبارات التحصيلية هو إظهار أثر الدراسة أو التدريب الخاص . وبالرغم من أنه يغلب تطبيقها في المدارس إلا أنها تستخدم في ميادين أخرى ، كأن تقيس التدريب المهني والتخصص والخبرة في كثير من الأعمال .

ومن المعتاد أن نقارن أيضاً بين الاختبارات التحصيلية واختبارات الاستعدادات . وتختلف الاختبارات التحصيلية عن اختبارات الاستعدادات في درجة توحيد الخبرة السابقة المقاسة . إذ تقيس الاختبارات التحصيلية أثر مجموعات من الخبرات المقننة هي المناهج الدراسية . بينما نجد أن اختبارات الاستعدادات تقيس الأثر المتجمع للخبرات المتعددة في الحياة اليومية . ويمكننا القول بأن اختبارات الاستعدادات تقيس أثر التعلم تحت ظروف غير معروفة وغير مضبوطة بينما تقيس الاختبارات التحصيلية أثر التعلم تحت ظروف معروفة ومضبوطة إلى حد ما ،

واختبارات التحصيل تختلف أيضاً عن اختبارات الاستعدادات من حيث استخدام كل منها . فاختبارات الاستعدادات تفيد في التنبؤ بنوع الأداء المنتظر ، فهي تستخدم في تقدير المدى الذي سيفيد به الفرد من التدريب كما تستخدم في التنبؤ بنوع التحصيل في موقف جديد . في حين نجد أن الاختبارات التحصيلية تبين المكانة الحالية للفرد ومدى إفادته من

التدريب إذ يدلنا الاختبار على ما يمكن أن يعمله الفرد في زمن محدود بعد قياس كفايته أو تحصيله في زمن محدد .

ولكن لا يمكن أن تميز بين هذين النوعين من الاختبارات بصورة قاطعة فبعض اختبارات الاستعدادات تعتمد على تدريب خاص سابق موحد بين الأفراد ، وهنا تقرب كثيراً من الاختبارات التحصيلية . وكذلك فإن بعض الاختبارات التحصيلية تغطي مدى واسعاً من الخبرات التعليمية غير المقننة وهنا تقرب كثيراً من اختبارات الاستعدادات التي تقيس أثر خبرات غير مقننة . وبالمثل فإن أى اختبار تحصيلي يمكن أن يستخدم في التنبؤ بالتعليم المستقبل ، هذا فضلاً عن استخدامه في ميادين التحصيل الأخرى ، فمثلاً التقدم الذي أحرزه طالب في مادة معينة ، حسباً تدل درجاته في اختبار تحصيل في المدرسة الإعدادية ، دليل نعتد عليه في التنبؤ بنجاح نفس الفرد في هذه المادة في المعاهد العليا . وعلى هذا فدرجات الأفراد في الاختبارات التحصيلية التي تغطي المنهج الدراسي ، تفيدنا في التنبؤ بنجاحهم مستقبلاً :

وفي التمييز بين اختبارات الاستعدادات والاختبارات التحصيلية يجب أن نقول ببساطة أن الاختبارات التحصيلية تقيس ، أثر التعلم المقنن وقد أصبح من الواضح أن كل الاختبارات النفسية والتربوية تقيس سلوك الفرد الحالي الذي يعكس أثر تعلمه السابق في مختلف المجالات وقياس أثر التعلم السابق يصلح للتنبؤ بالتعلم المستقبل .

استخدامها :

تستخدم الاختبارات التحصيلية في التعليم والخدمات المختلفة والجيش كما أنها تستخدم أيضاً في مجال علم النفس الإكلينيكي وهي تخدم أغراضاً عديدة أهمها :

١ - الاختيار أو الانتقاء Selection : كانتقاء الصالحين لنوع من التعليم، أو المدارس الفنية . . الخ .

٢ - التصنيف Classification : كتصنيف الجنود بعد معرفة خبراتهم وتدريبهم العمل السابق أو تصنيف أطفال المدارس الابتدائية والإعدادية حتى يكونوا جماعات متناسقة يضم كل جماعة منها فصل أو جماعة مدرسية .

٣ - في الحكم على مستوى مهارات الفرد ومعرفته الحالية وهذا أمر لازم لوضع خطة تعليمية أو تدريب مهني حتى يمكن الحكم عليه .

٤ - التحقق مما إذا كان الفرد قد وصل أم لم يصل إلى مستوى معين، كما في حالة استخدامها في الجيش أو في الصناعة .

٥ - في الأغراض الإكلينيكية ، كتشخيص قدرة الفرد أو عجزه ومدى هذا العجز في ناحية ما مثل صعوبات القراءة أو أى ضعف تعليمي آخر ومعرفة الأفراد المتمردين والجائحين . خاصة وأنا أحيانا نجد القشل المدرسي وسوء التوافق الدراسي عند بعض الأطفال الموهوبين .

٦ - تقييم البرامج الدراسية وكذلك تعديلها ، فن طريق الاختبارات التحصيلية نفيس فعالية البرامج ومدى تحقيقها لأهداف التربية وما ينبغي إدخاله على تلك البرامج من تعديلات .

٧ - تقييم التلميذ ، لأن موضوعية الاختبارات التحصيلية وشيوعها يساعدان المدرس على وضع تقييم أدق للتلميذ .

٨ - تعيين البرامج العلاجية (الهوايات مثلا) فهي تعين في تحديد نوع البرامج العلاجية الملائمة للفرد ، وتحدد درجة استفادته منها حاليا ومستقبلا .

٩ - وعلى ضوء ما سبق تساعد الاختبارات التحصيلية وتيسر عمليات.

التعلم وذلك بمعرفة نقاط الضعف في التعلم السابق وتحديد طريقة الدراسة والمذاكرة التي يتبعها الطالب حالياً فتدفعه إلى التحصيل الدراسي والتعلم بطريقة أفضل .

وقد ظهرت فائدة استخدام الاختبارات التحصيلية وأثرها في رفع ودفع التحصيل والتعلم في تجارب عديدة نذكر منها المثال التالي الذي ضربته أنستازى :

أخذت ستة (٦) فصول كمجاعة ضابطة وأعطيت اختباراً في أول البرنامج الدراسي واختباراً آخر في منتصفه واختباراً ثالثاً في نهايته . وأما الجامعتين التجريبتين فكانت كل جامعة تكون من ٤ فصول أعطيت ١٢ اختباراً أثناء الدراسة . وفي إحدى الجامعتين التجريبتين استخدم جهاز كانت لإجابات المفحوص عن فقراته تسجل بالضغط على أحد أزرار خمسة فإذا كانت الإجابة صحيحة ترك الزرار علامة حمراء وأظهر له الجهاز الفقرة التالية أما إذا ضغط على زرار يقابل إجابة خاطئة فإن الفقرة تظل أمامه ولا تظهر الفقرة التالية وهنا يحاول المفحوص الأزرار الأخرى حتى يصل إلى الإجابة الصحيحة ، وتظهر له الفقرة التالية وهكذا يعرف المفحوص أخطائه . والدرجة على الاختبار تقل تبعاً لعدد محاولات الضغط على الأزرار أما المجموعة التجريبية الأخرى فكانت نتيجة الاختبار الكلي تعلن لأفرادها فور انتهاءهم منه .

وفي نهاية مدة التدريب ظهر أن درجات تلاميذ المجموعتين التجريبتين تزيد كثيراً عن درجات تلاميذ المجموعة الضابطة . وزادت درجات تلاميذ المجموعة التجريبية الثانية ، التي كان المفحوص فيها يعرف بنفسه أخطائه وصوابه في كل فقرة ، عن درجات تلاميذ المجموعة الأولى ، التي كان المفحوص فيها يعرف النتيجة النهائية للاختبار الذي أجرى عليه ، زيادة كبيرة .

١٠ - وتفيد الاختبارات التحصيلية عموماً في تحسين طرق التعلم ومراجعة محتويات المناهج والوقوف على كفاية كل محتوى والمواقف التي تتطلب توحيد التدريب ومدى ونوع الأخطاء الشائعة في فهم موضوع معين .
تحديداتها :

إننا بدون شك نعرف بفوائد الاختبارات التحصيلية غير أن البعض يغالى في هذه الفوائد ويظهر هذا عند التطبيق . ومن أكثر الأخطاء شيوعاً تقنين التعليمات فإذا كان التقنين مصحوباً بتعلم المهارات الأولية كما في بعض المواقف المهنية والعسكرية التي قن فيها الاختبار التحصيلي إلا أن التحصيل لا يقوم على تعلم نفس المهارات التي قن عليها في مجالات أخرى من التدريب . وعلى أى حال فإن اضطراد توحيد المهارات يقضى على التلقائية والإبداع والأصالة في التفكير ، وأكثر من هذا فإن التقنين يتجاهل الفروق الفردية في المعلم والمتعلم ويتجاهل الحاجات والظروف المحلية الخاصة لكل من الفاحص والمفحوص .

وما زالت أغلب الاختبارات التحصيلية تهتم ببعض التعليمات أكثر من غيرها ، فمن السهل مثلاً أن نضع أسئلة تتعلق بكل تفاصيل المنهج الدراسي المعين بينما يصعب ذلك في حالة السؤال عن الفهم والإدراك النقدي وتطبيق المبادئ العامة . إلا أنه يغلب نسيان التفاصيل الواقعية المتعلمة خلال الدراسة بينما نجد تفسير المبادئ المتعلمة وتطبيقها في مواقف جديدة يمثل القدرة على الاسترجاع والكسب الحقيقي بعد مرور زمن من إتمام المنهج ، ولهذا يعاب على الاختبارات التحصيلية أنها غالباً ما تشحن بمواد لها قيمتها الحالية في الدراسة ، وتقل قيمتها فيما بعد عند مواجهة الحياة الواقعية .

وبالرغم من كل هذه التحديدات للاختبارات التحصيلية فإن التلاميذ والقائمين بالاختبارات يكونون بفضلها مدفوعين دائماً إلى الاهتمام بالتركيز على المواد التي ترفع درجاتهم في هذه الاختبارات .

ويجب أن ينظر إلى الاختبارات التحصيلية على أنها أدوات وليست أهدافاً . وأكثر من هذا فعلى أن نذكر أنها تعطينا بيانات جزئية ليست كافية لفهم الفرد ولهذا نحتاج إلى ملاحظات أخرى إضافية . وهذا وجه من أوجه الشبه بينها وغيرها من أنواع الاختبارات الأخرى . والواقع أن الاختبارات التحصيلية في تحسن دائماً فهي الآن تقيس مظاهر التعلم المعقدة. وتتنبأ من خلال هذه الدرجات بالنجاح في الدراسة أو المهنة مستقبلاً . إلا أن بعض ميادين التربية يصعب فيها تصميم اختبار جيد . ومن أهم هذه الميادين قياس القدرة الناقدة ، وقياس استخدام المعارف القديمة في مواقف جديدة . ومع هذا فالاختبارات التحصيلية تخطو الآن سريعا نحو هذه الميادين :

صدق الاختبارات التحصيلية

لا تختلف طرق حساب معاملات الصدق ومعانيه في الاختبارات التحصيلية عنها فيما عداها من الاختبارات ، إلا من حيث تفضيل استخدام طرق معينة تناسب مع طبيعة المادة ، واستخدام محكات معينة تتوفر في أكثر من غيرها في هذا المجال . ففي الاختبارات التحصيلية يوجد الصدق بطريقتي تحليل المحتوى Content analysis والتحليل المنطقي logical analysis . وهذا لا يعني أنهما الطريقتان الوحيدتان في ميدان التربية إذ تستخدم كل الطرق الأخرى في إيجاد الصدق .

وأهم أسباب شيوع استخدام طريقتي التحليل المنطقي وتحليل المحتوى أن واضع الاختبار التحصيلي لا يجد وقتاً كافياً لدراسات مستفيضة إذ أن الاختبار التحصيلي الجيد هو الاختبار الذي يساير المنهج الدراسي الذي يقيسه . وهذا المنهج يختلف من بيئة إلى أخرى ومن نوع معين أو مرحلة معينة من التعليم إلى آخر ، ناهيك عن اختلافات بعض المناهج بين

مدارس المنطقة الواحدة واختلافات طرق تدريس المقرر الواحد في المادة المعينة من مدرس إلى آخر ومن فصل دراسي ذي مستوى معين إلى آخر ، [كل هذا يوجب عمل اختبارات تتحدد بمجال معين ضيق ، وبالتالي تتعدد اختبارات التحصيل التي تقيس نفس المادة في نفس الفرقة الدراسية . وهذا التعدد من شأنه أن يجعل اختبارات التحصيل عرضة دائماً للتغير والتجديد مما لايسمح بفرصة كافية لقيام الباحث بدراسات الصدق والثبات وغيرها من البحوث التي تلتزم بها الاختبارات في المجالات الأخرى .

فإذا أضفنا إلى هذا أن طبيعة المادة المقاسة ، وهي غالباً مقرر موضوع معين ، تحول بين الباحث وبين القيام بدراسات تجريبية تتبعية وبينه وبين القيام بأنواع أخرى من التحليل مثل التحليل العاملي ، وجدنا أن أنسب طرق إيجاد الصدق في الاختبارات التحصيلية هما طريقتي تحليل المنهج والتحليل المنطقي .

وفي تحليل المنهج يقوم واضع الاختبار بدراسة الأهداف ودراسة المحتوى أي المنهج الدراسي وتحديد أجزائه ونسبها . ثم يقوم بوضع فقرات تناسب كل هذا ، واضعاً في اعتباره أن يمثل عدد من الفقرات جزءاً من المحتوى بحسب وزن هذا الجزء بالنسبة للمنهج المقاس كله . أما في التحليل المنطقي فيقوم الباحث بوضع المفاهيم التي يود أن يكون التلاميذ قد حصلوها بعد أن درسوا المنهج ، ثم يعالج الفقرات منطقياً لتبين درجة تغطيتها لهذه المفاهيم متمسكاً بتحصيل الأهداف الذي يميز بين التلميذ الضعيف والتلميذ المتفوق . وفي الحالتين يهتم الباحث بأن يكون الاختبار كافياً في تمثيل المنهج الدراسي .

والحك الذي يغلب استخدامه في الاختبارات التربوية هو النجاح الدراسي . فإذا كان أغلب الحاصلين على درجات مرتفعة في الاختبار ناجحين دراسياً ،

كان الاختبار صادقاً . وإذا تزايدت نسبة الناجحين في الفقرة بارتفاع مستواهم ، كان الاختبار صادقاً : وهنا تختار الفقرات التي تميز بين الأفراد بما يتفق مع تميزهم إلى مستويات دراسية مختلفة . ولكن استخدام تقدم الدرجات إذا ارتبط بالنجاح الدراسي ، وتزايد نسب الناجحين في الفقرة إذا تدرجت مستوياتهم الأكاديمية ، استخدام هذين الدليلين كمحك في الاختبارات التحصيلية ليس مشروعاً دائماً .

وأول أسباب عدم كفاية هذا المحك ، الاختلافات البيئية والإقليمية بين مستويات ومستويات المنهج الدراسي ، وقد سبق الإشارة إليه . كما أن بعض المواد الدراسية يتوقف تدريسها في مراحل معينة ، وهنا لا يحصل تلميذ الفقرة الأعلى على درجة أكبر بل وقد يحصل على درجة أقل من تلك التي يحصل عليها من يدرس هذه المادة . وليس هذا ، بالطبع ، داليل على انخفاض مستوى تحصيل طالب الفقرة الأعلى ، إذ أن طالب الفقرة الأقل لو اختبر بنفس الاختبار بعد انتهاء تدريس هذه المادة بمدة ، لحصل على درجة أقل من درجته الأولى .

وأخيراً ، فإن الاختبارات التحصيلية قد تستخدم في التنبؤ بالأداء المقبل بافتراض أنه متسق مع الأداء الحالي الذي تقيسه . وفي هذا الغرض يجب بحث الصديق التجريبي بالطريقة التبعية ، وهنا ينهج الأخصائي النفسي التربوي منهج الأخصائي النفسي في مجالات القياس الأخرى . (انظر الفصل الرابع عشر) .

معايير الاختبارات التحصيلية

المعيار أساس عام مشترك به تصبح الدرجة الخام ذات معنى فتفسر درجة الفرد بمقارنتها بنقطة تنسب إليها . وبينما يطبق هذا المفهوم في

اختبارات الذكاء باستخدام معيار العمر العقلي ، نجده يطبق في الاختبارات التحصيلية باستخدام معيار الفرقة الدراسية Grade Norm .

وسبب استخدام معيار الفرقة الدراسية هو أن التلاميذ مقسمون إلى جماعات متجانسة مكونة بالفعل هي جماعات الفصول والفرق . فلا يحتاج الباحث إلى القيام بتصنيف المفحوصين إلى جماعات يحسب لكل منها معيار يناسبها . كما أن معيار الفرقة الدراسية واضح يمكن تفسيره مباشرة . ونحصل على هذا المعيار بحساب متوسط درجة كل فرقة دراسية بالنسبة للمادة التي يقيسها الاختبار . وتحول الدرجة الخام للفرد إلى مكافئ الفرقة Grade Equivalent إذا كانت تساوى متوسط درجات تلاميذ هذه الفرقة : وهكذا يصبح التلميذ في مستوى فرقة متقدمة في إحدى المواد لأنه متفوق فيها ، وفي مستوى فرقة متأخرة في مادة أخرى لأنه ضعيف في هذه المادة .

وبهذه الطريقة نعتبر أن التلميذ في عدة فرق دراسية ما دامت درجاته في المواد المختلفة تصل إلى مستوى هذه الفرق ، أى تساوى متوسط درجات التلاميذ الذين أخذناهم من هذه الفرق في عينات تقنين تلك الاختبارات . وهذا يمكننا من رسم صفحة نفسية للتلميذ (انظر شكل (١١) .
الفصل التاسع) .

وقد اقترح كيلي T.L. Kelly معيار العمر المنوالى modal-age ، بمعرفة العمر الشائع للفرقة الدراسية المعنية واستبعاد من يزيدون عنه أو يقلون . ثم يحسب متوسط درجات أفراد هذه الفرقة الدراسية بعد أن أصبحوا من عمر واحد . وتتراوح أعمار أفراد الفرقة الدراسية بهذه الطريقة في مدى سنة كاملة . وما دام مدى كبير أو صغر سن التلميذ عن منوال فرقته يختلف باختلاف المدراس وباختلاف الفرق الدراسية داخل المدرسة .

الواحدة ، فإن المعيار الذي يحسب من درجات مجموعة أفراد الفرقة الدراسية لا يمكننا من مقارنة الفرق الدراسية المختلفة .

وبرغم مزايا معايير الفرق الدراسية وسهولتها العملية ، إلا أن ثمة اعتراضات كثيرة تقف دونها . وأول هذه الاعتراضات هو أن محتوى ومستوى التعليم يختلف من فرقة إلى أخرى . كما أن حصول التلميذ على درجة تعادل متوسط فرقة دراسية أعلى من فرقته لا يعنى وصوله فعلاً إلى مستواها في مادة الاختبار نتيجة لتحصيله الرسمي . بل قد يعود ارتفاع الدرجة إلى القراءات والاطلاع ، وهذا يعنى اتفاق المادة مع ميول الطالب واهتماماته بالوظائف المقاسة في الفرقتين .

والنقد الذي وجهناه إلى مفهوم العمر العقلي ينطبق أيضاً على مفهوم معيار الفرقة ونقصه أن الوحدات غير متساوية . بمعنى أن نمو الفرد في الأعمار الصغيرة يكون بسرعة أكبر منه في الأعمار الكبيرة ، والفرق بين مستوى الفرقة الدراسية الثانية والثالثة قد يكون أكبر من الفرق بين مستوى الفرقة الثالثة والرابعة إذ أن قدرة التلميذ على التحصيل في الفرق الأدنى تزيد عنها في الفرق الأعلى .

نأً ومن عيوب معيار الفرقة أيضاً ، أنه خلال أي فرقة دراسية واحدة ينمو التلاميذ في معرفة المواد بسرعات تختلف من مادة إلى أخرى بحسب المنهج الدراسي . فالتلميذ الذي يعلو فرقته الدراسية بفرقتين في مادة معينة ويعلوها بفرقتين أيضاً في مادة أخرى قد يزيد عن متوسط فرقته في المادة الثانية بدرجة أكبر من زيادته عن متوسط الفرقة في المادة الأولى لأن الفرقة في المادة الأولى تتقدم علمياً بسرعة أقل منها في المادة الثانية . وفي هذه الحالة نجد أن بعض تلاميذ هذه الفرقة يفوقها بفرقتين ولا يلزم أن قدراته فيها عالية .

وتدريس مقررات خاصة في مادة معينة قد يجعل الفرد متفوقاً في هذه المادة ولكن تفوقه لا يصل إلى مستوى فرقة دراسية أعلى من فرقته ؛ وإذا كان الاختبار التحصيلي مصمماً لقياس مدى تفوق التلميذ بالنسبة لفرقته ، وحصل التلميذ على درجة عالية ، كان من الخطأ أن نعتبر أنه وصل إلى مستوى فرقة دراسية أعلى لأن الاختبار يقيس التفوق داخل الفرقة بالنسبة لمجموعتها ولا يقيس الوصول إلى مستويات فرق أعلى من هذه الفرقة .

فإذا حصل تلميذ في السنة الأولى الإعدادية عن اختبار حساب على الدرجة ٣٥ . وكان متوسط درجات تلاميذ السنة الثالثة الإعدادية في نفس الاختبار هو ٣٥ ، فيجب أن نعتبر هذا التلميذ متفوقاً في فرقته . ولا يصح اعتباره في مستوى الفرقة الثالثة ، وإلا أصبح معيار الفرقة مضللاً إذ يحصل التلميذ المتفوق في السنة الأولى على درجة التلميذ المتوسط في الثالثة ؛ هذا بينما تختلف في المادة الثانية وأن عدداً أصغر يفوقها بفرقتين أيضاً ؛

وآخر عيب نذكره ، لمعيار الفرقة الدراسية هو أنها لا تشمل أفراداً متجانسين بدرجة كبيرة فهناك فروق فردية بين تلاميذ الفرقة الواحدة تجعل بعضهم يفوقها بعدة فرق والبعض الآخر يقل عنها بعدة فرق خاصة في المواد التي يرتبط فيها التحصيل بالميل أو الاهتمام أو يتأثر فيها الفرد بظروفه أو أسرته . وبناءً عليه لا يصح اعتبار الفرقة الدراسية كلها من مستوى أداء واحد فهناك فروق فردية قد تكون واسعة المدى أحياناً ؛

وعلاجاً لكل هذه العيوب لجأت الاختبارات التحصيلية إلى أنواع أخرى من المعايير وإلى تعديلات على معيار الفرقة الدراسية كان أولها معيار العمر المنوال . وقد تعرضنا في الفصل التاسع والفصل الرابع عشر لأغلب هذه المعايير وهي :

١ - معايير العمر التعليمي .

٢ - النسبة التعليمية :

٣ - نسبة التحصيل أو الإنجاز .

٤ - المئين داخل الفرقة الدراسية .

وقد بدأت الاختبارات التحصيلية تعتمد على معيار تستخدم فيه توزيعاً مركباً واحداً لكل الدرجات المعيارية في كل الفرق الدراسية ؛ فاختار فرقة دراسية معينة كجاعة مرجعية ، ويستخدم متوسط هذه الفرقة وانحرافها المعياري في حساب الدرجات المعيارية . فبغض النظر عن العمر أو الفرقة الدراسية تحول الدرجة الخام إلى مكافئها الدرجة المعيارية ؛ ومن هنا يفيد هذا المعيار في المقارنة وفي رسم وتفسير الصفحات النفسية :

ويرى البعض عمل توزيع بحسب السن ، ويرى آخرون عمل توزيع بحسب الفرق الدراسية المقاسة ذلك داخل الفرقة المرجعية باختبار يعتمد على هذا النوع من الدرجات المعيارية الكلية . وهكذا تحول الدرجة الخام للتلميذ إلى درجة معيارية عالية إذا كان سنه صغيراً أو من فرقة دراسية منخفضة ، ويحصل التلميذ الأكبر عمراً أو من فرقة دراسية أعلى على درجة معيارية كلية أقل لأن عمره أو فرقته تسمح له بتحصيل أفضل . بينما يحسب البعض الآخر المقابلات المئينية لهذه الدرجة المعيارية الكلية فيعد التلميذ في المئين ٧٠ مثلاً لأن درجته المعيارية تقابل هذا المئين ؛ ولكن كل هذه التعديلات تواجه النقد الذي سبق أن وجه إلى معايير العمر والمعايير المئينية ومعايير الفرقة الدراسية ؛ وأحدث المعايير استخداماً في اختبارات التحصيل يعتمد على تحديد مستوى يفترض أن التلميذ وصل إليه ما دام قد أنهى برنامجاً دراسياً معيناً :

وظائف الاختبارات التحصيلية

وقد يحس القارئ تداخلا بين الاستخدامات Uses والوظائف Functions ونحن نوافقه إلى حد ما . وبيان علاقة الجزء بالكل بين الاستخدامات والوظائف يقتضى مناقشة فلسفية ليس هذا مجالها . وأفضل طريقة لحسم هذا النقاش هى عرض تقرير مكتب شئون الأفراد بالبحرية الأمريكية . وفى هذا التقرير عرض للأغراض التى نجحت فى تحقيقها الاختبارات التحصيلية المتقنة وأهمها :

١ - ساعدت الاختبارات فى تقنين التعليم لكل المدارس التى تعد الناس لنوع معين من الوظائف . فبالرغم من أن أهداف التعليم والمناهج كانت مقننة ، إلا أن بعض المدارس قد انحرفت بإهمال مواضيع معينة أو المغالاة فى الاهتمام بها . وأى انحراف سينعكس فى الدرجات على الاختبار فتتخفص درجات المدرسة عن درجات بقية المدارس والاختبارات لهذا أجبرت المدرسين على أن يسيروا حسب ما اعتمد واضعو المنهج .

٢ - تعتبر الاختبارات أساساً يوحى بمراجعة المنهج الدراسى وتحسين طرق التعليم . فإذا وجد من الاختبار أن الوقت لا يسمح بإتقان مواد دراسية معينة ، كان العلاج هو إعطاء وقت إضافى فى المنهج لهذه المواد . فقد كان المدرسون ، قبل تطوير الاختبارات التحصيلية ، يعتقدون أن لإنهاء جزء من المنهج يعنى أنهم درّسوه على خير وجه . ولكن نتائج الاختبارات أصبحت معيناً للمدرسين يلجئهم إلى المشرفين والمتخصصين ليحسنوا طرقهم فى التدريس . وكأن الاختبارات ساعدت المدرسين على أن ينموا نقد الذات

Self- Criticism

٣ - تطلبت بعض الاختبارات أن يظهر التلاميذ مهارتهم فى العمل أكثر من مجرد الإجابات اللفظية . وهكذا وجهت الاختبارات الانتباه إلى

أهمية تعليم السلوك الذى يهدف المنهج إلى إنمائه لدى التلاميذ . فانخفضت أهمية المناقشة فى المحاضرة كطريقة فى التدريس . ووجد أن فعالية التدريب قد زادت .

٤ - كانت درجات التلاميذ تعطى على أساس الانطباع الذاتى فى حالة عدم وجود الاختبارات التحصيلية أو ، على الأقل ، فى حالة عدم إجراء اختبارات يضعها المدرسون . ومثل هذه الدرجات لا يعتمد عليها ، وقد فشلت فى لفت النظر إلى كل الجوانب الجوهرية من المنهج ، وخضعت للتعصبات . والاختبارات الموضوعية ، نفسها ، عندما يضعها مدرسون لم يأخذوا تدريباً خاصاً ، كانت دائماً أدوات ذات درجة منخفضة من الموضوعية . والإعداد الدقيق للاختبارات يزيد احتمال حصول كل فرد على درجته بدقة تامة .

٥ - الاختبارات التى أكدت على كل الجوانب الجوهرية من المقرر أثبتت أن التحصيل بصفة عامة أصبح موضع اعتبار وأن النقائص فيه أصبحت ملحوظة ففى كلية الهندسة مثلاً ، قبل الأخذ بالاختبارات المقننة ، كانت الدرجات تتأثر كثيراً بأداء النواحي الرياضية من المنهج ، وربما كان هذا لأن هذه القدرة يسهل قياسها . فتغيرت الامتحانات لتضع التأكيد على الفهم الميكانيكى . ونتيجة لهذا ، وجه الانتباه إلى الأشخاص الذين برغم مهارتهم الحسابية ضعاف فى مواد أساسية أخرى .

٦ - سمحت الاختبارات المقننة بتحليل الدرجات ومستوياتها وقللت الفروق الراجعة إلى كرم « بحجة » أو قسوة المصححين .

٧ - تحسن الدافع لدى التلاميذ والمعلمين بقيام المنافسة على أساس عادل . فأحسن طريقة لدفع الدارس هى أن نوضح له الطرق المعينة التى يحتاج فيها إلى إنماء مهاراته .

٨ - فى إنشاء اختبارات انتقاء التلاميذ أصبحت الدرجات النهائية الدقيقة ^أ محكاً أفضل على أساسه يشتق صدق هذه الاختبارات .

٩ - حطمت الاختبارات المقننة اعتبار طول مدة التعليم مقياساً لقدرة المتعلم . فقد وجد أنه لا يمكن اعتبار عدد السنوات ولا عدد المناهج أساساً للحكم على الشخص . ففى دراسة أجريت سنة ١٩٣٨ أخذ آلاف الطلبة الجامعيون اختبارات مقننة للمعرفة فى ميادين متعددة ؛ ووجد أن كثيراً من طلبة السنوات النهائية فى الجامعة يعرفون أقل مما يعرف التلميذ المتوسط فى السنة النهائية بالمدرسة الثانوية .

١٠ - وللاختبارات التحصيلية أيضاً فوائدها الإكلينيكية والتنبؤية ، فضلاً عن وظيفتها العامة فى تقييم التعليم . واختبارات التحصيل غالباً ما تكون أداة ممتازة فى التنبؤ بنجاح التلميذ باختيار مناهج أعلى من التى أخذها التلميذ فعلاً . وهى قد تستخدم فى اختيار الموظفين - كما أن الفوائد الإكلينيكية لاختبارات التحصيل تنضح فى حالة الجانح الذى قد يظهر تفوقاً فى مواد الدراسة وهنا يتحول العلاج إلى استغلال نواحي تعليمية . واختبارات التحصيل مفيدة أيضاً فى تقدير وتشخيص التدهور العقلى الناشئ عن أسباب عضوية أو انفعالية .

أنواع الاختبارات التحصيلية

بعض الاختبارات تقيس التحصيل فى مادة واحدة كالتاريخ أو العلوم : وهناك بطاريات من اختبارات التحصيل أعدت للفرق الإعدادية هذه البطاريات تتكون من عدة اختبارات فرعية وكل اختبار يقيس إحدى مجالات التعلم وما دامت الاختبارات الفرعية مقننة معاً ، فإن الاختبار يجبرنا بمعرفة الطفل لمادة بالمقارنة بقدرته فى المجالات الأخرى . وأكثر مجالات القياس شيوعاً المستوى الإعدادى هى القراءة والهجاية والقواعد والتاريخ ، والحساب . أما اختبارات الثانوى والمعاهد فقد أعدت لإعطاء

ضرورة عن النمو في كل المجالات ، ولكن على وجه العموم ، زاد شيوع استخدام اختبارات المادة الواحدة عن استخدام بطاريات من اختبارات المواد .

وبالإضافة إلى اختبارات المسح التي تقدر التحصيل العام للشخص في مجال واحد ، هناك اختبارات تشخيصية سببا في القراءة والحساب . والاختبار التشخيصي لا يحاول أن يحدد ، المستوى الذي وصل إليه الشخص ، بل يحدد نقاط القوة والضعف لدى التلميذ . فالاختبارات التشخيصية تركز على العملية التي يستجيب خلالها المفحوص ، أكثر مما تركز على نتيجة هذه العملية .

والعمليات التشخيصية أحيانا تستخدم في تدريس أي موضوع . فالمدرس لا يستطيع أن يدرس بفعالية ما لم يعرف عيوب التلميذ . ولكن الاختبارات التشخيصية نادرا ما تكون مقننة لأنها تهتم أساسا بالمقارنة في ذات الفرد وليس بمقارنة الفرد بالآخرين

و أغلب الاختبارات المنشورة تقيس المعرفة ، وتهمل جوانب السلوك الأخرى . وقد حدث تطور حديث في القياس التعليمي هو إنشاء مقاييس عديدة جديدة لأهداف مثل : القدرة على تطبيق المبادئ ، ومهارات الاستدلال ، والقدرة على تفسير البيانات . هذا فضلا عن أن بعض المدارس قد استخدمت اختبارات الاهتمامات واختبارات الاتجاهات لدراسة تقدم التلميذ .

الفصل السادس عشر

اختبارات تحصيلية مقننة عامة

- قياس التفكير .
- اختبارات النمو التعليمي .
- قياس مهارات الذاكرة .
- بطاريات التحصيل العام .
- خاتمة .

الفصل السادس عشر

اختبارات تحصيلية مقننة

عامـة

قياس التفكير :

برغم أن تدريب التفكير من أهم الأهداف العامة للتربية ، إلا أن الدراسات التربوية لم تكن به كثيراً . فقل عدد الاختبارات التحصيلية التي تقيس مدى التقدم في عمليات التفكير نتيجة الاشتراك في برنامج دراسي ، ومهارات التفكير بعد أن دربتها مختلف المؤسسات التربوية . وأكثر من هذا أن تعريف التفكير Thinking كان يتسع أحياناً فيشمل الكثير من المهارات والعمليات ، ويقتصر في أحيان أخرى على مفهوم كالدكاء ، أو العامل العام ، أو النجاح الأكاديمي .

وقد قامت إحدى الدراسات ، التي استغرقت ثمان سنوات ومن ثم أصبح اسمها دراسة الثمان سنوات Eight Years Study ، بمحاولة لتصنيف عمليات التفكير إلى فئات قليلة محددة جيداً ومن ثم إلى إنشاء اختبارات موضوعية لكل فئة منها . وإن كان التصنيف غير شامل ، إلا أنه تضمن مظاهر من التفكير يمكن وصفها بالعبارات التالية :

١ - القدرة على تفسير البيانات .

٢ - القدرة على تطبيق المبادئ العلمية .

٣ - القدرة على تطبيق المبادئ المنطقية .

٤ - فهم طبيعة البرهان .

وكل فئة من الفئات الأربع تشمل منطقة واسعة من السلوك ، الأمر الذى دفع إلى تحديدها فى جوانب معرفة تبرز فى كل فئة . ففى حالة القدرة على تفسير البيانات ، قيست عناصر القدرات التالية :

(أ) القدرة على إدراك العلاقات بين البيانات :

١ - القدرة على عقد المقارنات :

٢ - القدرة على استخلاص عناصر مشتركة بين البيانات :

٣ - القدرة على قراءة البيانات .

٤ - القدرة على القيام ببعض العمليات الحسابية البسيطة بين هذه البيانات .

٥ - القدرة على فهم الرموز التى استخدمت وطرق عرض البيانات .

(ب) القدرة على معرفة حدود البيانات :

١ - القدرة على تحديد بيانات إضافية نحتاجها فى اشتقاق نتيجة معينة .

٢ - القدرة على التوقف عن إصدار الحكم حتى تتوفر كل البيانات اللازمة .

٣ - القدرة على الخروج بتعميمات من البيانات تكون قابلة للتحقيق Justifiable .

٤ - القدرة على تعيين التفسيرات التى لا يمكن تحقيقها .

٥ - القدرة على الخروج باستنتاجات جيدة تتعلق بالاتجاهات الواضحة بين البيانات .

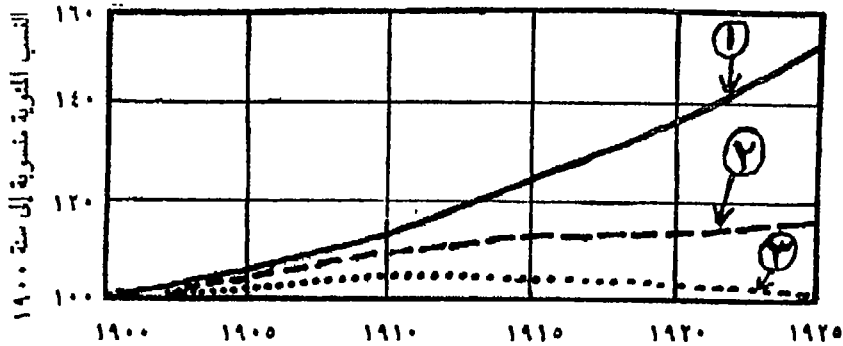
ولم تكن الفئات الفرعية السابقة كل فئات القدرة على تفسير البيانات ، فقد شملت مثلاً القدرة على تقدير البيانات التى لا يمكن الاعتماد عليها : أو القدرة على التقييم الناقد للاستنتاجات التى يقوم بها الآخرون : ولكن

مثل هذه القدرات لم تُختَر للدراسة ، لأنه ليس عمليا أن يُدرس كل جانب من القدرة على تفسير البيانات ؟

أما الطرق التي صممت لقياس هذه القدرة فقد حوت بيانات عرضت بطرق مختلفة - بالرسم البياني ، والجداول ، والألفاظ ، . . الخ . اختيرت هذه البيانات من مواد مقررة ، تشمل مجالات متعددة من الحياة ، وتمثل العديد من العلاقات بينها .

والشكل (١٤) يمثل طريقة عرض بعض البيانات التي تقيس القدرة على تفسير البيانات ؟ والشكل البياني متبوع بسلسلة من الجمل التي يطلب من المفحوص أن يحكم على صحتها مستعينا بالبيانات المعروضة بيانيا . والحكم

(شكل ١٤)



نموذج لأحد الأسئلة التي تقيس المهارة في تفسير البيانات

الخط (١) مقدار الإنتاج الزراعي

الخط (٢) عدد السكان الزراعيين الذين في سن العمل (الاستخدام)

الخط (٣) عدد العمال الزراعيين الذين يعملون (المستخدمين)

يكون على أساس مدى كفاية البيانات المعروضة وحدها في تمكيننا من إصدار أحد الأحكام التالية :

- (أ) هذه الجملة صحيحة .
- (ب) يحتمل أن تكون صحيحة .
- (ج) هذه الجملة غير صحيحة .
- (د) يحتمل أن تكون غير صحيحة .
- (هـ) ليست هناك بيانات تمكننا من أن نصنف الجملة في فئة من الفئات الأربع السابقة .

المشكلة الأولى : هذا الرسم البياني يمثل الإنتاج ، والسكان ، وعدد المستخدمين في مزارع الولايات المتحدة في كل خمس سنوات ابتداء من سنة ١٩٠٠ حتى سنة ١٩٢٥ :

الجميل :

- ١ - النسبة بين الإنتاج الزراعي وبين العمال الزراعيين تزايد كل خمس سنوات بين سنة ١٩٠٠ وسنة ١٩٢٥ .
- ٢ - زيادة الإنتاج الزراعي بين سنة ١٩١٠ وسنة ١٩٢٥ ترجع إلى اتساع استخدام الآلات الزراعية .
- ٣ - متوسط عدد العمال الزراعيين المستخدمين خلال المدة بين سنة ١٩٢٠ وسنة ١٩٢٥ أعلى من عددهم خلال المدة بين سنة ١٩١٥ وسنة ١٩٢٠ .
- ٤ - يجب أن تنقذ الحكومة العمال الزراعيين العاطلين .
- ٥ - بين سنة ١٩٠٠ وسنة ١٩٢٥ ، زادت كمية الفواكه المنتجة في مزارع الولايات المتحدة بمقدار ٥٠٪ .
- ٦ - خلال المدة الكلية بين سنة ١٩٠٥ وسنة ١٩٢٥ كان عدد

السكان الزراعيين من أعمار يمكن تشغيلها أكبر من عدد الأفراد المطلوبين لتشغيل المزارع .

٧ - انخفضت الأجور المدفوعة للعمال الزراعيين سنة ١٩٢٥ لأن عدد العمال الموجودين كان أكثر من العدد الممكن تشغيله :

٨ - عدد العمال الذين استخدموا سنة ١٩٢٥ زاد عن عدد المستخدمين سنة ١٩٠٠ :

٩ - منذ سنة ١٩٠٠ ، كانت الزيادة في إنتاج كل عامل صناعي مساوية لزيادة إنتاج كل عامل زراعي .

١٠ - بين سنة ١٩٠٠ وسنة ١٩٢٥ ، زاد الإنتاج الزراعي بأكثر من ٥٠ ٪ .

١١ - زاد العمال الزراعيون إنتاجهم بعد سنة ١٩١٠ حتى يستفيدوا من الارتفاع السريع في الأثمان .

١٢ - متوسط كمية الإنتاج الزراعي في الفترة بين سنة ١٩٢٦ وسنة ١٩٣٠ كان أكثر منه في الفترة بين سنة ١٩٢٠ وسنة ١٩٢٥ ،

١٣ - بين سنة ١٩٠٠ وسنة ١٩٢٥ ، كانت هناك زيادة في عدد السكان الزراعيين من أعمار يمكن تشغيلهم في المنطقة المتوسطة الغربية وهي أكبر المناطق الزراعية في الولايات المتحدة :

١٤ - كان عدد السكان الزراعيين من أعمار يمكن تشغيلها سنة ١٩٣٠ أقل من عددهم سنة ١٩٠٠ .

١٥ - إنتاج القمح ، وهو أكثر المحاصيل الزراعية في الولايات المتحدة ، كان في سنة ١٩١٥ مساوياً له في سنة ١٩٢٥ :

وطريقة تصحيح هذا الاختبار معقدة بعض الشيء . وقد شملت

صحيحة التسجيل المستخدمة في دراسة السنوات الثمانية أما كن لعدة درجات :
وهذه الدرجات تتبع الفئات الأربع التالية :

١ - الدقة العامة : هذه الدرجة الكلية over-all تقيس مدى دقة استجابات التلميذ في إنفاقها مع مفتاح تصحيح الاختبار . وقد اشتق المفتاح من أحكام مجموعة من الخبراء لكي يبين مدى صحة أو احتمال صحة الحكم في كل حالة . ودرجة الدقة العامة يمكن أن تنقسم بدورها إلى الفئات الخمس التي يختار إحداها المفحوص حكماً يصدره على مدى كفاية البيانات في الرسم في دلالتها على الجملة المعطاة .

٢ - الحرص : هذه الدرجة تقيس ميل التلميذ إلى شدة الحرص في تفسير البيانات . فإذا صنف جملة في فئة « البيانات غير كافية » وكان الواجب أن تصنف في فئة « تحتل الصحة » ، كان المفحوص شديد الحرص . ودرجة الحرص Caution هي عدد الاستجابات التي بالغ فيها التلميذ في الحرص على الاختبار .

٣ - ما وراء البيانات : هذه الدرجة تقيس الميل إلى زيادة التعميم من البيانات . أي أنها كانت عكس زيادة الحرص ، وبالتالي تقابل درجة ما وراء البيانات Beyond Data درجة الحرص . وعلى أي حال فإن التلميذ يستطيع أن يحصل على درجات في المقياسين إذا عرف أن يتوسط بين الحرص الزائد وقلة الحرص ما دام التلميذ يحصل على درجة هنا ودرجة هناك .

٤ - الخطأ الفاحش : وهو عدد الأخطاء الواضحة ، وهي الأخطاء التي تحدث عندما يحكم المفحوص على الجملة بأنها غير صحيحة في حين تكون صحيحة أو العكس .

وبالإضافة إلى الفئات السابقة ، تعطى درجات عن الإجابات المتروكة .

وهدف عملية التصحيح بمختلف مراحلها هو عرض الطريقة المميزة التي يفكر بها التلميذ عادة . فليس الاختبار إذن مقياسا ، له معيار التحصيل الكلى ، يقيم التلميذ في هذا الميدان . ويبين الجدول (١٢) أنواع التفسير التي يمكن أن تعطى للصفحات النفسية في هذا الاختبار . وهو يعرض درجات اختبار واحد والتفسير الذي يقترحه واضعو الاختبار لهذه الصفحة . وهم سميث Smith وتيلر Tyler والفريق الذي ساعدهما .

الفئة	درجات التلميذ س	متوسط الفصل
الدقة العامة	٥٤	٥٧
الدقة في الحكم بـاحتلال الصحة ويحتل عدم الصحة	٣٥	٣٨
الدقة في الحكم بعدم كفاية البيانات	٥١	٥٨
الدقة في الحكم بصحيح وغير صحيح	٧٦	٧٣
الحرص الزائد	٤٨	٢١
التميم الزائد (ما وراء البيانات)	٤٣	٣٦
الأخطاء الفاحشة	١١	٨

(جدول ١٢)

يبين صفحة نفسية لتفسير بيانات اختبار دراسة السنوات الثماني

ولكى نفسر هذه البيانات الكمية يلزمنا تفكير محدد لمعرفة مدى قرب التلميذ س من متوسط فصله . فدرجاته أقل من المتوسط بدرجة صغيرة . وهذا يدل على نقص القدرة على التعرف على حدود البيانات . وبفحص أنواع الخطأ تبين زيادة كبيرة عن المتوسط في الميل إلى الخروج عن

البيانات ، أو تقبل تعميمات متجاهلا حدود البيانات . وليست هذه الدرجة من أعلى الدرجات في الفصل (ولها دلالة ، ما دامت أغلب درجات التلميذ قريبة من متوسط الفصل) ، ولكنها أيضاً تدل على أن نسبة الأخطاء في زيادة التعميم مقارنة بالأخطاء في زيادة الحرص تزيد أيضاً عن متوسط الفصل (النسبة بين زيادة التعميم إلى زيادة الحرص لدى التلميذ = ٤٣ : ٤٨ ، ونفس النسبة بين متوسطي الفصل = ٣٦ : ٢١) ودرجة الأخطاء الفاحشة من أعلى الدرجات في الفصل .

وأحد صعوبات مثل هذا النوع من الاختبارات هو أنه يحاول أن يقيس الكثير جداً في وقت قصير جداً . والنتيجة هي أن الدرجات الجزئية ، لأنها تستند إلى فقرات قليلة ، لا يمكن الاعتماد عليها بدرجة كافية في توجيه أو قياس التلميذ الفرد .

وقد أعد الفريق الذي قام بدراسة السنوات الثمانية عدة اختبارات أخرى وأشهرها الاختبار الذي يحاول قياس القدرة على تطبيق المبادئ العلمية . وهو اختبار شامل جيد إلا أن نظام تصحيحه أكثر تعقيداً من نظام تصحيح اختبار القدرة على تفسير البيانات . إذ تحسب للتلميذ ٢٠ درجة وبدراسة هذه الدرجات يمكن تحديد نوع الخطأ الذي يقع فيه التلميذ والعادات الفكرية غير المرغوب فيها .

ومن الجدير بالذكر أن هذه الدراسة التي بدأت سنة ١٩٣٢ اقترحت ، إعداد بعض الاختبارات ولكنها لم تستخدم إلا في سنة ١٩٥٢ . وتمكيننا للدراسات التي تستخدم هذه الاختبارات بسمح الناشر بإعطاء أصولها للقائمين بالدراسة ويطيعونها ولكن ليس من حقهم بيعها كما لا يقوم الناشر نفسه ببيعها . وهذا يدلنا على مدى حرص العلماء في قصر استخدام الأدوات التي تحت التجربة على القائمين بالبحوث لا أن تباع فتفسد وتتحول إلى تجارة .

اختبارات النمو التعليمي

« اختبارات أيوا للنمو التعليمي » Iwa Tests of Educational Development يمكن أن تعد امتداداً « لاختبارات أيوا لكل تلميذ في المهارات الأساسية » Iwa Every pupil Tests of Basic Skills الذي يقيس الفرق الدراسية ابتداءً من الفرقة المقابلة للسنة الثالثة الابتدائية حتى الفرقة المقابلة للسنة الثانية الإعدادية . وإن كانت هذه السلسلة من الاختبارات تشمل اختبارين للمعلومات العامة وليست لمهارات التفكير . والاختبارات حسب ما تدلنا أسماؤها تقيس النمو التعليمي العام ، وإن كان الاعتماد على الأسماء هو من قبيل الصدق الظاهري إلا أنه يعطينا فكرة عما نقيس ، نحو الأهداف المرغوبة من التربية . إلا أن معنى هذه الأهداف هنا غير واضح تماماً .

وفي كراسة تعليمات الاختبار نجد مواصفات الاختبارات التي يشملها بالطريقة التالية :

الاختبار الأول : فهم المفاهيم الاجتماعية الأساسية — وهو اختبار يقيس

معرفة التلميذ العامة وفهمه العام لموقف المجتمع الحديث . وكل تلميذ يحتاج إلى هذه المعرفة حتى يتمكن من التفكير الجيد في المشاكل الحالية .

الاختبار الثاني : صحة الكتابة — وهو اختبار يحدد تمكن التلميذ من

العناصر الأربع الأساسية في الكتابة الصحيحة — التنقيط ، الاستخدام ، العنوان ، والمهجاية . وهذا الاختبار لا يقيس القدرة على حفظ القواعد ، بل القدرة على استخدامها .

الاختبار الثالث : المعلومات العامة في العلوم الطبيعية — وهو اختبار يحدد

معرفة التلميذ العامة وفهمه للظواهر العامة ، ولأهم إسهامات المدنية الحديثة .

وهو كالاختبار الأول ، لا يتحدد بـيتمقرر السنوات الدراسية في المدرسة الثانوية فحسب . فقد اختيرت الفقرات على أساس ما يجب أن يعرفه الشخص الذي نال تعليما عاما .

الاختبار الرابع : القدرة على التفكير الكيفي - يقيس الاختبار القدرة على حل المشاكل العامة العملية . وهذه المشاكل تمثل فهم التأمين على الحياة ، ونظام التقسيط في الشراء ، والقدرة على تقدير تكاليف بيت صغير وإصلاحه ومشاكل مشابهة :

الاختبار الخامس : القدرة على تفسير البيانات المقروءة في الدراسات الاجتماعية - يقيس الاختبار قدرة التلميذ على تفسير وتقييم قراءات مختارة من الكتب المقررة ، والمجلات ، والجرائد . والقدرة على تقييم ، وفهم المتضمنات ، واستدلال المضمون الرئيسي للقطعة المختارة كلها جوانب يقيسها الاختبار .

الاختبار السادس : القدرة على تفسير المواد المقروءة في العلوم الطبيعية - وهو كالاختبار الخامس ، يقيس قدرة التلميذ على تفسير وتقييم مختارات مقروءة . وهذا الاختبار يختلف عن الاختبار الخامس فقط في المادة الدراسية موضوع الاختبار .

الاختبار السابع : القدرة على تفسير مواد أدبية - وهذا الاختبار يعرض مقتطفات مختارة من الأدب العالمي - فكل الأنواع الكبرى من الأدب ممثلة : وقيس التعرف على جو الكاتب وهدفه من بين ما يقيس .

الاختبار الثامن : المفردات العامة - وهو اختبار يقدم مقياسا للقدرة التلميذ على معالجة الكلمات فقط ، بل وأيضا لاستعداده العام للتعلم .

الاختبار التاسع : استخدام مصادر للعلوم - وهو اختبار يقيس قدرة التلميذ على الاستفادة من مصادر المعلومات التي يغلب أن تشمل حل مشكلة معينة .

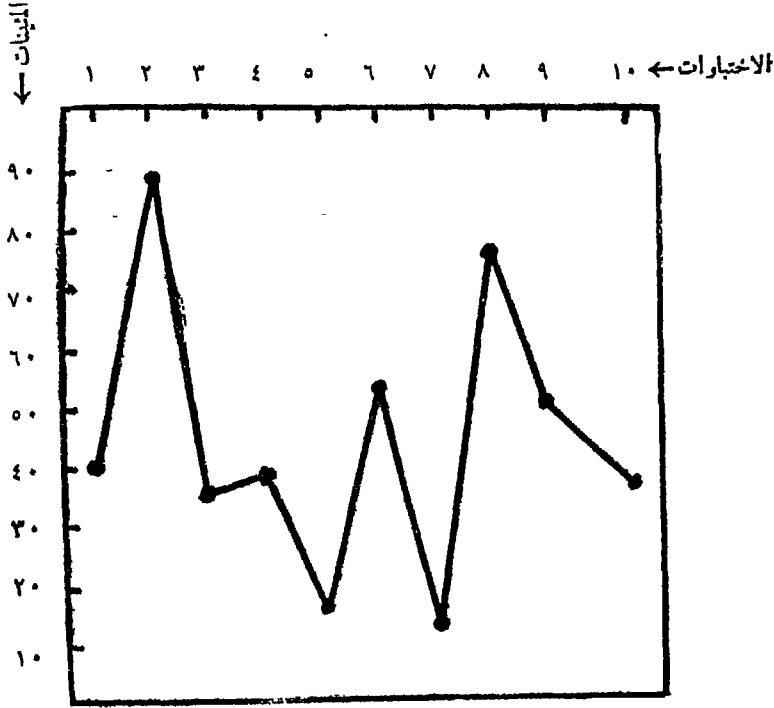
الاختبارات الخامس والسادس والسابع في هذه السلسلة هي الاختبارات التي تعيننا أكثر من غيرها في هذا المقام . ويجب أن نلاحظ أنها ليست اختبارات للمعلومات ، إذ كما تدل أسماؤها ، تؤكد على التفسير . والتفسير أحد جوانب قدرة التفكير التي تقاس بتقديم قطعة للتلميذ تتعلق بالمادة الدراسية موضوع القياس ويطلب منه أن يفسر معنى هذه القطعة . والأسئلة كلها مصاغة بطريقة الاختيار من عدة إجابات أو اختيار أفضل الإجابات Best Answer Form من بين الإجابات المرفقة بالقطعة . ولكن من الصعب أن نحدد بدقة ماذا تقيس مثل هذه الاختبارات . فهي قد تقيس المعرفة بموضوع المادة الدراسية ، وإن كان هذا الاحتمال صغيراً إذ أنها تطلب من المقحوص أن يستخدم أرضيته من المعلومات العامة . تفسير القطع .

وبالرغم من أن هذه البطارية تؤكد على تفسير المادة في مختلف ميادين الموضوعات الدراسية ، إلا أنها أيضاً تؤكد على قياس المعلومات الأساسية أو القاعدية إذ أن الأخيرة تعنى من يقيس النمو التعليمي . وليس هذا عيباً في الاختبارات ما دام واضعوها يرون في هذه الجوانب نقاطاً هامة يجب تقديرها .

وقد حدث تطوير لاختبار « أبوا » كي يلائم أغراض الاختيار والتوجيه في سلاح الطيران الأمريكي وعلى هذا أصبح اسمه « اختبار النمو التعليمي العام لمعهد سلاح الطيران بالولايات المتحدة » .

The United States Armed Forces Institute of General Educational Development. (U.S.A.F.I.G.E.D.).

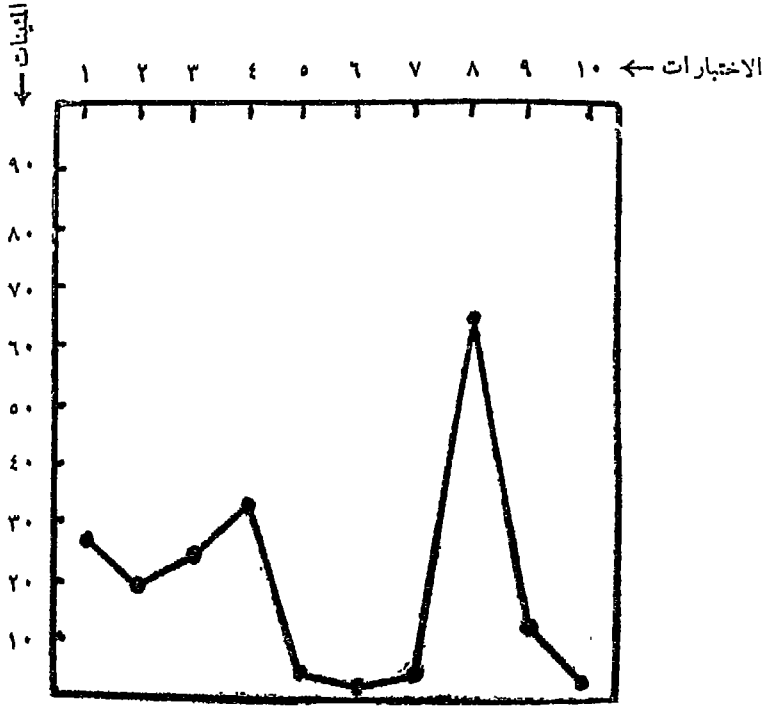
صفحة نفسية من اختبارات أيوا للنمو التعليمي .



(شكل ١٥)

- (١) اختبار فهم المفاهيم الاجتماعية الأساسية
- (٢) » المعلومات العامة في العلوم الطبيعية
- (٣) » صحة الكتابة
- (٤) » القدرة على التفكير الكيفي
- (٥) » تفسير البيانات المقروءة في الدراسات الاجتماعية
- (٦) » تفسير المواد المقروءة في العلوم الطبيعية
- (٧) » القدرة على تفسير مواد أدبية
- (٨) » المفردات العامة
- (٩) » الدرجة المركبة » الاختبارات من ١ - ٨
- (١٠) استخدام مصادر المعلومات

صفحة نفسية من اختبارات أيوا للنمو التعليمي



(شكل ١٦)

- (١) اختبار فهم المفاهيم الاجتماعية الأساسية
- (٢) » المعلومات العلمية في العلوم الطبيعية
- (٣) » صحة الكتابة
- (٤) » القدرة على التفكير الكيفي
- (٥) » تفسير البيانات المقروءة في الدراسات الاجتماعية
- (٦) » تفسير المواد المقروءة في العلوم الطبيعية
- (٧) » القدرة على تفسير مواد أدبية
- (٨) » المفردات العامة
- (٩) الدرجة المركبة « الاختبارات من ١ - ٨ »
- (١٠) استخدام مصادر المعلومات

وهو يستخدم لاختيار المتقدمين للإلتحاق بكلية الطيران من الحاصلين على دبلومات من المدارس الثانوية العسكرية ويتكون من أربعة اختبارات يقيس الأول صحة وسلامة التعبير ، و يقيس الثاني تفسير المواد المقروءة في الدراسات الاجتماعية، و يقيس الثالث تفسير المواد المقروءة في العلوم الطبيعية ، و يقيس الرابع تفسير المواد الأدبية المقروءة . ولم تذكر فيه معاملات ثباته ، أما عن بحوث الصدق عليه فقد استخدم محك النجاح في الكلية وكان عدد عينة الصدق ١٣٥ طالبا ومعامل الصدق ٠.٧٢ . ويصحح الاختبار بواسطة ما كينة التصحيح العالمية IBM والمعايير قومية . وعينة التقنين تمثل ستة مناطق جغرافية .

وهكذا نرى أن الاتجاه نحو قياس مهارات التفكير في التربية أصبح مدعما بفضل دراسات واسعة النطاق مثل دراسة السنوات الثمانية واستخدام مقاييس النمو التعليمي في عينات كبيرة كما حدث في سلاح الطيران الأمريكي . والافتراض ورائها هو أن تعلم طرق التفكير الصحيحة الفعالة يمكن الفرد من التكيف والتوافق في أكبر مدى من المواقف .

على أن المدرس الذي يستخدم أداة من هذه الأدوات يجب أن يحلل ما يملو أن الاختبار يقيسه فيقرأ المدرس الاختبار ، ثم يجيب عن أسئلته ، ثم يسأل نفسه قائلا :

١ - ما عدد فقرات الاختبار التي تقيس المعرفة بمفردات موضوع القياس ؟

٢ - ما عدد الفقرات التي تقيس القدرة على اشتقاق الأدلة من البيانات ؟

٣ - ما هو عدد الفقرات التي تسأل التلميذ أن يخرج بتعميمات من البيانات ؟

فإذا ساءل المدرس نفسه بهذه الطريقة كان في موقف يمكنه من الحكم على صلاحية الأداة في قياس الأغراض التي يضعها في ذهنه :

قياس مهارات المذاكرة

يؤكد المدرسون في كل مستويات التعليم أهمية تحسين مهارات المذاكرة ، ويعتبرها البعض ضمن جوانب مهارات التفكير . وتكثر الشكوى ، في [أ] أيامنا هذه ، من عدم كفاية الطرق التي يتبعها التلاميذ في المذاكرة ، وخاصة في مراحل التعليم الأولى . ولا يهمننا أن نلقى تبعة هذا على عواهل المدرس أو التلميذ قدر ما يهمننا أن تقوم أبحاث منظمة علمية لتحسين مهارات المذاكرة يقوم بها المدرسون والإخصائيون داخل وخارج الفصل . وسنتحدث عن مهارات المذاكرة بتقسيمها إلى الفئات الرئيسية التالية :

١ - مهارات القراءة : وهذه المهارات متضمنة غالباً في المهارات اللغوية ، ولكن المتخصصين في مهارات المذاكرة يهتمون عند التعرض لمهارات القراءة بتشخيص أسباب صعوبات القراءة .

٢ - مهارات الكتابة : بما في ذلك إعداد التقارير . وهي تتضمن مهارات التنظيم والعنونة وتنظيم عرض المفاهيم المتضمنة .

٣ - مهارات الإصغاء : وتتضمن مهارات الإصغاء Listening إلى جوار حسن الاستماع والانتباه مهارات كتابة المذكرات والمحاضرات والملخصات .

٤ - مهارات تنظيم الأفكار : وهي متضمنة في مهارات اللغة . ولكنها تبرز أثناء القراءة .

٥ - مهارات التخطيط : ونعني بها عمل الجداول وتخطيط عمليات المذاكرة .

٦- مهارات إجابة الامتحانات : وهي تتضمن مهارات الكتابة ومهارات استخدام الأدلة والنقط الهامة في الإجابة عن الأسئلة الموضوعية :

٧- مهارات جمع البيانات : ومهارات الإفادة من المكتبات .

وقد قام عدد من العلماء بتقديم قوائم يسجل فيها التلميذ نوع الصعوبات التي يواجهها في المدرسة . وهذه القوائم تفيد في القيام بالمقابلات الشخصية Interviews مع التلاميذ لتوجيههم وإرشادهم ، إذ أن التلميذ لا يكون عارفاً ما هي أسباب مشاكله ومصادرها :

وأفضل قائمة من هذا النوع هي قائمة مراجعة المشاكل التي أعدها موني Mooney : وهي تستخدم على نطاق واسع الآن خاصة في المدارس الثانوية والإعدادية في الخارج . وقد أعدت من هذه القائمة صور كثيرة أولاهها للمدارس الإعدادية ، وأخرى للثانوية ، وثالثة لطالبات التمريض ، ورابعة للشباب من القرويين وإن كانت الصورتان الأخيرتان لا تستخدمان الآن . كما أعدت صورة للراشدين وأخرى للجامعيين .

تم إنشاء هذه القوائم بناءً على الكثير من المواد ذات الدلالة التي جمعت عن مشاكل فعلية : قال تلاميذ المدارس الثانوية إنهم يواجهونها : وقد تم جمع هذه المواد عن طريق المقابلات الشخصية مع التلاميذ ، وتحليل سجلات الحالات ، وتحليل فقرات كتبها التلاميذ عن مشاكلهم ، ومن التعبيرات الكتابية للتلاميذ عن حاجاتهم التربوية ، ومن نتائج بحوث سابقة مماثلة .

وتتضمن القائمة الكثير من المشاكل مثل الخجل - الخشوع - التمرکز حول الذات - ضخامة المسؤولية العائلية : ومجموع فقرات القائمة ٣٠٠ فقرة . وبعضها يعالج عدم كفاية مهارات الذاكرة ، والعادات الخاطئة

في المذاكرة ، ومصادر الفشل الدراسي . ويطلب من الطالب أن يقرأ القائمة ببطء وأن يضع خطاً تحت الفقرات التي تشير إلى مشاكل يعاني منها . وبعد أن يكمل المفحوص هذه الخطوة الأولى ، يراجع الفقرات ويضع خطوطاً ودوائر حول أرقام بعضها الذي يعده « أكثر ما يتدخل في حياته » . وأخيراً يسأل أن يكتب ملخصاً لمشاكله . ولا تصحح القائمة بل تعد كشفاً لمشاكل التلميذ .

وقد اقترح موني عدة مناهج لبحث صدق قائمة مراجعة المشاكل :
والمناهج هي :

١ - يسأل الطالب الذي أجاب عن القائمة السؤال التالي :

« هل تشعر أن فقرات القائمة التي وضعت علامة عليها تعطي صورة كاملة عن مشاكلك ؟ » .

وقد وافق ٩٢٪ ممن أجابوا عن القائمة على أن هذه القائمة تعطي صورة كاملة تقريبية عن مشاكلهم واعتبروا ذلك دليلاً على صدقها ، والصدق هنا يعني أن إجابة المفحوص تعكس صورة عما يعتقد هو أنها مشكلات تعترضه . وهذا لا يعني بالضرورة أن إجاباته ترتبط بالتشخيص الإكلينيكي ارتباطاً كاملاً . والسبب أن استبصار التلميذ بطبيعته

الخاصة محدود . وإذا نمي تبصر التلميذ قد يرى في إدراكه لمشكلاته سابقاً رأياً آخر فيكتشف مثلاً أن ما كان يعاني منه فعلاً هو مشكلات أخرى غير التي وافق عليها في القائمة . وعلى هذا فالصدق بالمعنى الأول ليس الصدق الذي يعيننا كشتغلين بالقياس . ولكنه صدق رجل الشارع ، أي صحة أو عدم كذب الفقرات واتفاقها مع إدراك التلميذ لنفسه .

٢ - كما اقترح موني دراسة صدق قائمته بملاحظة مدى اتفاق الإجابات مع المشاكل الشخصية للمجيبين وكيف أن الإجابات تتغير بتغير المشاكل . ولبحث هذا الدليل قام موني بالتجربة الآتية :

أتى بمجموعتين تكونت الأولى من ٧٠ طالباً والأخرى من ٤٦ طالبة وأعطيت لهم قائمة المراجعة . وبعد ٩ أيام طلب منهم أن يضعوا علامات على المشاكل التي حلوها وعلى المشاكل التي جدد عليهم وكانت هذه العلامات توضع في إحدى الاستمارات المصغرة عن هذه القائمة . ثم أعيد إجراء القائمة وعندئذ ظهر أن ٨٣٪ من المشاكل التي وضعت عليها علامات في الاستمارات المصغرة وضعوا عليها أيضاً علامات في القائمة : ولكن هذا الاتفاق قد لا يعنى صدق القائمة بقدر ما يعنى درجة عالية من الإتساق بين الفرد ونفسه في مرتى الإجابة عن أسئلة تذهب إلى نفس المحرب .

٣ - أما الطريقة الثالثة فهي أفضل من سابقتها إذ تغلبت على النقد الموجه إليهما من أن التلميذ لا يعرف بدقة مصادر مشكلاته فاتفاقه في الإجابات أو اختلافه بعد فترة لا يعنى أن مشكلاته تغيرت ولكنه يعنى أن إدراكه لهذه المشاكل قد تغير . فإذا كان هذا الإدراك نفسه غير صادق تماماً فالأحرى أن تكون عمليات بحث صدق القائمة المعتمدة عليه غير مؤكدة أيضاً .

والطريقة الثالثة الفضلى هي اختيار مجموعات ثبت من دليل آخر صادق أن لها مشكلات محددة ثم تطبيق قائمة المراجعة عليها لتبين ما إذا كانت الفقرات تعكس المشكلات أم لا . وقد قام بدراسة من هذا النوع ستودجيل Stodgill وبينتون Benton حيث طبقا قائمه المراجعة على مجموعتين تتكون كل منهما من ٣٥ طالباً جامعياً . وإحدى المجموعتين فصل ينال تلاميذه علاجاً في مهارات المذاكرة والأخرى فصل ينال تلاميذه علاجاً في الصحة النفسية .

وقد زاد تكرار إجابات تلامذة المجموعة الأولى عن فقرات مثل « لا أعرف كيف أذاكر بطريقة جيدة » ، « أخشى الرسوب في الكلية » ، « لا أعمل شيئاً بكفاءة » ، « المدرسون لا يهتمون بالتلاميذ » ، « احتاج

إلى أن أعرف القدرات المهنية » ، « غير قادر على التركيز » ، بطيء القراءة » ، . الخ . أما تلامذة فصل الصحة النفسية فقد زاد تكرار إجابات تلاميذه عن فقرات مثل « أشعر بالاضطهاد » ، « مضطرب في عقائدي الدينية » ، « إن الآباء يرهقون الأبناء بكثرة مطالبهم » ، « لا أجد وقتاً كافياً لنفسى » ، « أحتاج إلى برامج علاجية أخرى » :

بينت هذه الدراسة أن المشاكل التي وضع الطلبة تحتها خطوطاً في قائمة المراجعة هي المشاكل التي يعانون منها . ومع ذلك يجب أن نكون حذرين جداً عند تفسير القوائم الفردية فنحن نعلم أن المريض لا يكون متبصراً بسبب مشكلاته . فهو يربط القلق بهذا الموضوع أو ذاك ويتلمس في الأسئلة أو الفقرات إشارة إلى هذا الموضوع . وعلى ذلك نجد أنه يعاني من القلق ويحجب عن فقرات القلق بالإيجاب أى الموافقة . ولكنه لا يتعرف على سبب هذا القلق ، حين أننا يجب أن نبحث دائماً عن مصادر هذه الأعراض أو المشاكل وليس عن الموضوعات التي ارتبطت بها : فالقائمة إذن تعين على تشخيص المشكلات وليس على تحديد أسبابها ومصادرها .

بطاريات التحصيل العام

في المرحلتين الابتدائية والإعدادية

تستخدم بطاريات التحصيل العام general Achievement ابتداء من المستوى الابتدائي حتى الدراسات العليا . إلا أن أفضل استخدام لها في المرحلة الإعدادية . وأغلب هذه البطاريات تعطى صفحات نفسية تمثل درجات مختلف المواد الدراسية كما أنها تعطى درجة كلية عن البطارية ككل . وهذه البطاريات تختلف فيما بينها في عدة خصائص سواء من حيث المواد التي تقيسها أم المستويات المقاسة . كما تختلف في معاييرها

ومعاملات صدقها وثباتها . ولذا يصعب أن نأخذ بطارية واحدة لنمثل بها خصائص هذا النوع من البطاريات .

وقبل أن نعرض هذه البطاريات ننبه إلى أنها لا تصلح للاستخدام في بلادنا إلا بعد الكثير من التعديلات . وأولى بوضعي الاختبارات أن يصمموا بطاريات واختبارات فيها أصالة Originality تمثل دراسة شاملة لحاجتنا وإمكانياتنا التربوية وتناسب المواد الدراسية في مختلف مستويات التعليم العام والخاص . كما ننبه إلى أن أغلب ما سنعرض يستخدم في أمريكا ونحن نقدمه مع بعض التعديلات حتى يمكن أن يفهم خلال أوضاعنا فيحضر البعض على استخدامه والآخر على تصميم اختبارات من هذا النوع .

أولا : اختبارات متروبوليتان التحصيلية : إن اختبارات متروبوليتان التحصيلية Metropolitan Achievement Tests من أشيع بطاريات التحصيل استخداما في مستوى المرحلة الإعدادية وإن كانت تستخدم في السنوات الدراسية ابتداء من الأولى الابتدائية حتى نهاية المرحلة الإعدادية . وهي تحتوي ٥ بطاريات ولكل بطارية ٤ أو ٥ صور متكافئة . وهي اختبارات قوة . ويحتاج تطبيق كل البطارية إلى حوالى ٤ ساعات . وكل الاختبارات التي تشملها البطارية الواحدة مطبوعة في كتيب واحد .

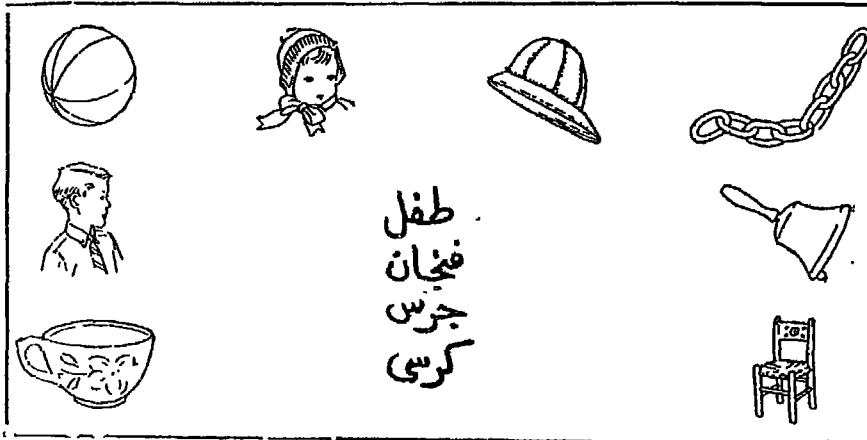
طبقت الاختبارات على نطاق واسع واختيرت عينه قومية وحسبت الأعمار المنوالية لكل فرقة دراسية . والمعايير التي استخدمت هي معايير العمر المنوالى والعمر العقلى والفرقة الدراسية والتوزيع المثبني للفرقة الدراسية . وقد حسبت معاملات ثبات كل بطارية على حدة بالطريقة النصفية وكانت معاملات الثبات تراوح بين ٨ وإلى ٠٠ر٩٧ وكان المحك في بحث الصدق هو صدق المنهج الدراسي وارتفاع درجات المفحوصين في الاختبارات بتقديم فرقهم الدراسية . وعلى نفس الأساس اختبرت الفقرات إذ قورنت

النسب المئوية للناجحين بفرقهم الدراسية كما قورنت لدرجات الأفراد في الفقرات واحدة بواحدة بدرجاتهم في الاختبارات .

وكى يتضح مدى ومحتوى هذه السلسلة من البطاريات نعرض أول بطارية منها وهى تقيس السنة الأولى الابتدائية . وتتكون من الاختبارات التالية :

الاختبار الأول : القراءة - صورة الكلمة Word Picture ويتكون من ١٤ مجموعة من الصور وكل مجموعة بها ٨ صور . ووسط كل مجموعة منها عدة كلمات أو عبارات مكتوبة يطلب من المفحوصين أن يضعوا علامات توصل كل كلمة أو عبارة مكتوبة بالصورة التى تناسبها . ويقوم الفاحص بمساعدة المفحوصين فى نطق الكلمات المطبوعة .

الاختبار الثانى : القراءة - التعرف على الكلمة Word Recognition ويتكون من ٢٦ مجموعة من الكلمات كل مجموعة من ٤ كلمات تشابه فى النطق و الهجاية . وينطق الفاحص بكلمة منها فيضع المفحوصون



(شكل ١٧)

فقرة من اختبار صورة الكلمة فى بطارية مَروبوليتان التحصيلية

إشارة على هذه الكلمة فمثلا الكلمات في الصف هي Pig big bid dig وينطق الفاحص كلمة big ويترك للتلاميذ فرصة ١٥ ثانية لوضع علامة على هذه الكلمة هـ هكذا في كل صف .

الاختبار الثالث : القراءة - معنى الكلمة Word meaning ويتكون

الاختبار من ١٢ مجموعة من الكلمات كل في صف وتتكون المجموعة من ٦ كلمات تشترك ثلاثة منها في صفة . ويقول الفاحص هذه الصفة ويطلب من التلاميذ وضع علامات على الكلمات التي تتوفر فيها هذه الصفة فمثلا نجد الصف الأول : منزل - ولد - حجر - فأر - رجل - أب ، ويطلب الفاحص أن يضع التلميذ علامة على أى كلمة في هذا الصف تدل على شيء متكلم ويساعد التلاميذ في المحاولتين الأولى والثانية بقوله مثلا « صحيح ؛ ولد - رجل - وأب » .

الاختبار الرابع : الأرقام numbers - ويتكون من ٣٦ صفا من الصور

أو الأرقام ويطلب من المفحوص أن يضع علامة حسب التعليقات . مثلا صف به ٩ رجال أمام باب ويطلب من المفحوص أن يضع علامة على الشخص الخامس . أو عدة أرقام ويطلب من المفحوص وضع علامة على الرقم ٦ مثلا . ثم ثلاثة صفوف لعمليات الجمع البسيطة عددها ١٥ عملية مثل $(\frac{1}{2})$ وصف لعمليات طرح عددها ٥ مثل $(\frac{2}{3})$.

وهذه الاختبارات تتطلب القراءة . وهي شبيهة باختبارات الذكاء العام



(شكل ١٨)

فقرة من اختبار الأعداد في بطارية متروبوليتان التحصيلية

فى المرحلة الابتدائية وخاصة فى اختبارات الأعداد وإن كانت أصعب فى البطارية منها فى اختبارات الذكاء إذ أنها فى الأولى تطلب من التلميذ قراءة وطرح وجمع كسور عشرية . والبطارية تقيس حتى نهاية المرحلة الإعدادية وهنا نجد اختبارات فى القراءة والمفردات وأسس الحساب ومسائله واللغة وأدبها والمواد الاجتماعية (مثل التاريخ والجغرافية) والعلوم والمهجاية .

وفى اختبارات القراءة ، فى نهاية المرحلة الإعدادية ، يقرأ التلميذ قطعة صغيرة ثم يجيب عن أسئلة تتعلق بمضمونها ومعناها . هذا الاختبار واختبار المفردات قريبان من اختبارات الذكاء . أما بقية الاختبارات فتختلف قليلا وتقرب كثيراً من محتوى المنهج الدراسى .

ثانياً اختبار ستانفورد للتحصيل : Stanford Achievement .

يعتبر اختبارات ستانفورد للتحصيل من أقدم البطاريات التعليمية المقننة . وقد نشر لأول مرة سنة ١٩٢٣ وروجع سنوات ١٩٢٩ ، ١٩٤٠ ، ١٩٥٣ وفى كل مراجعة حدثت تعديلات كثيرة . والطبعة الحالية (نقصد طبعة سنة ١٩٥٣ التى تصبح فعالة حتى ١ يناير سنة ١٩٥٤) تتكون من ٤ بطاريات وكل بطارية لها ٥ صور متكافئة . وتقيس البطاريات الأربعة الفرق الدراسية من الثانية الابتدائية حتى نهاية المرحلة الإعدادية . وفى المراجعة التى نتكلم عنها ، أصبح محتوى الاختبارات مسائرا لأحدث تعديلات المناهج الدراسية كما سهل التصحيح وبات موضوعيا .

وقد اشتقت معايير جديدة بعد تطبيق البطاريات على ٥٠٠٠٠ تلميذاً وتلميذة يمثلون ٣٦٣ مجموعة . واستخدمت الدرجة K Score وهى تشبه الدرجة المعيارية المستخدمة فى بطاريات المتروبوليتان . وبدلاً من أن تشتق هذه الدرجات من المنحنى الاعتدالى اشتقت من منحنيات ملتوية

تمثل توزيع تلاميذ كل فرقة دراسية على حدة بالنسبة للنجاح المدرسي .
والمعايير من نوع العمر المتوالى ، ومعايير الفرق الدراسية ، والتوزيع
المثبني لدرجات تلاميذ كل فرقة دراسية على حدة .

حسبت معاملات ثبات اختبارات هذه البطارية بالطريقة النصفية
وتراوح من ٠,٦٦ إلى ٠,٩٦ . وعند وجود معامل ثبات أقل من
٨٠ . وكانت الفقرات تراجع على أساس أن نسبة الناجحين في الفقرة يجب
أن تزيد بتقديم المفحوصين تعليميا .

والواقع أن هناك تشابه بين بطاريتي المتروبوليتان وستانفورد للتحصيل
ينحصر أساسا في :

- ١ - أنهما تقيسان نفس الفرق الدراسية تقريبا .
- ٢ - صعوبة الأسئلة متقاربة بين الاثنين .
- ٣ - التعليمات متشابهة إلى حد ما .
- ٤ - كل من البطاريتين به بطارية جزئية للمواد التي تتطلب مهارات
مثل الحساب والقراءة .
- ٥ - نشرتهما شركة الكتاب العالمى World book Co .
إلا أننا نلمس أيضاً أوجه الخلاف الآتية :
- ١ - بطارية متروبوليتان أحدث ولكن مراجعتها أقل .
- ٢ - بطارية ستانفورد تمثل أحدث الاتجاهات التربوية على عكس
بطارية المتروبوليتان .
- ٣ - بطارية المتروبوليتان بها اختبار منفصل للجغرافية وآخر للتاريخ
بينما في ستانفورد اختبار واحد للاثنين معاً .
- ٤ - اختبارات ستانفورد تمتاز بوجود اختبارات لدراسات مهارات
التعليم والمذاكرة الحديثة مثل قراءة الخرائط والرسوم البيانية والجداول
واستخدام المعاجم وغير ذلك من المصادر المرجعية .

فى المرحلة الثانوية

أولا : اختبارات كاليفورنيا التحصيلية : California Achievement

تقيس من السنة الأولى الابتدائية حتى نهاية المرحلة الثانوية وتشمل ٤ بطاريات لكل منها ٢ أو ٣ صورة متكافئة . ويرى واضعوها أن المواد الدراسية مثل العلوم والجغرافية والتاريخ يختلف محتواها من مدرسة إلى أخرى ومن منطقة إلى أخرى بينما يعم تدريس المواد التى تطلب مهارات . وكل البطاريات تشترك فى قياس المجالات الخمسة الآتية :

١ - مفردات القراءة .

٢ - فهم المادة المقروءة :

٣ - الاستدلال الحسابى :

٤ - الأسس الحسابية :

٥ - اللغة .

وتفسر الدرجات عن هذه الاختبارات حسب معايير الفرق الدراسية ، أو العمر ، أو التوزيع المثنى للدرجات داخل الفرق الدراسية الواحدة : وقد جمعت البيانات المعيارية من عدد كبير من المناطق الجغرافية . وهى اختبارات تقيس القوة أكثر مما تقيس السرعة . وكل بطارية يمكن أن تعطى فى أكثر من جلسة واحدة .

ولهذه البطاريات ميزة استخدامها كأدوات تشخيصية . ولهذا تحسب لكل تلميذ ٢٠ درجة فرعية فى أجزاء الاختبار الواحد ، فضلا عن ٥ درجات كل عن اختبار . وتمثل الدرجات فى صفحة نفسية . وهذه الطريقة مثلا ينقسم اختبار مفردات القراءة إلى أقسام مثل : مفردات حسابية ، وعلمية ، ومفردات الدراسات الاجتماعية والأدبية . كما ينقسم اختبار اللغة إلى التنقيط ،

ووضع الفواصل ، واختيار الحمل ، والكلمات واختبار لأجزاء الحديث
وآخر للهجاية وآخر للخط و : الخ :

أما اختبار الاستدلال الحسابي فينقسم إلى مفاهيم الأرقام ، والرموز ،
والقواعد ، والأرقام والمعادلات ، والمسائل : واختبار العمليات الحسابية
ينقسم إلى عمليات الجمع والطرح ، والضرب ، والقسمة وقد ينقسم أكثر
من هذا إلى جمع بسيط ومركب وإيجاد العوامل المشتركة ، : الخ .

ونلاحظ ، بناء على هذا العرض ، أن هذه الاختبارات تحاول أن
تقيس الكثير جداً بالقليل جداً من الفقرات التي لا تغطي هذا العدد الضخم
من القدرات : ولكن الاختبارات الخمس ثابتة ويعتمد عليها في الغرض
من إنشائها في كل المستويات الدراسية حتى نهاية المرحلة الثانوية . وقد
أوجدت معاملات ثبات كل اختبار من الخمسة ولم توجد معاملات
الاختبارات الفرعية إذ أن بعضها قصير جداً .

ثانياً : بطارية المقرر الأساسي للمدرسة الثانوية : Essential High

School Content Battery.

ثالثاً : الاختبارات التعاونية للتحصيل : The Cooperative General

Achievement Tests (C.G.A.T.)

رابعاً : بطاريات النمو التعليمي : وقد سبقت الإشارة إليها .

خاتمة :

ومن عرضنا السابق لكل بطاريات واختبارات التحصيل العام والنمو
التعليمي يتضح لنا :

١- أن الكثير من هذه الاختبارات لا يقدم بعض معالم القياس الهامة
فلا نجد مثلاً معاملات الصدق التجريبي لاختبارات أيوا واختبار :
U.S.A.F.I.O.E.D. كما لا نجد في الاختبارين أيضاً معاملات الارتباط

الداخلية بين الاختبارات الفرعية . ومعاملات ارتباط هذه الاختبارات :
مع اختبارات تحصيلية أخرى مقننة .

٢ - أن هذه الاختبارات ، عموماً ، تقيس وظيفة قريبة مما تقيسه
اختبارات الاستعدادات المدرسية . وكلا النوعين يأخذ محكاً له ، من
النجاح الأكاديمي .

٣ - أن بعض الاختبارات التحصيلية تعنى بقياس المهارات المعينة التي
تدخل في نطاق الاستعدادات الخاصة والاستعدادات الفارقة .

وباختصار فإن الاتجاه الحديث في إنشاء وتطوير الاختبارات التحصيلية
قد أفاد كثيراً من تقدم النظريات التربوية . كما أفادت منه المناهج
الدراسية .

إلا أنه ، كما أسلفنا ، لا تتوحد الخبرات التعليمية بين التلاميذ غالباً
نتيجة لتغيرات النظم التعليمية ولاختلاف امتحانات المناطق وتفاوت النجاح
الأكاديمي بين التلاميذ أفراداً . ولهذا فعند تطبيق اختبارات النمو التعليمي العام
على أفراد مختلفين في خبراتهم التربوية يجب أن يركز الاهتمام على وظيفتين
نقيم بهما التلاميذ موضوع القياس هنا وهما :

١ - القدرة على إدراك وفهم المعلومات :

٢ - مدى توحد خبراتهم وأثر تعليمهم السابق :

الفصل السابع عشر

اختبارات تحصيلية مقننة خاصة

- القراءة :
- الرياضيات .
- العلوم .
- الجغرافية .
- مجالات أخرى .
- مقاييس الإنتاج .
- أغراض تعليمية أوسع .
- الاختبارات التحصيلية المهنية .

الفصل السابع عشر

اختبارات تحصيلية مقننة

خاصة

القراءة :

تستحق دراسة القراءة اهتماماً خاصاً لعدة أسباب : أولها أن القراءة هي الأداة الرئيسية التي عن طريقها يتعلم الفرد كل المواد : وثانيها أن عدد اختبارات القراءة التي أعدت يزيد كثيراً من حيث العدد ومن حيث التنوع أيضاً عن غيرها من الاختبارات . وثانيها أنها أداة هامة شائعة الاستخدام في الإرشاد والتوجيه النفسي وفي ميادين الخدمات الإكلينيكية المختلفة . هذا فضلاً عن أن اختبارات القراءة تعرض لنا الكثير من المشاكل التي تصادفنا في عمليات إنشاء اختبارات التحصيل المقننة الخاصة .

تعريف القراءة :

قد يبدو لنا من أول وهلة أن القراءة مهارة بسيطة واضحة التعريف ميسورة القياس . ولكننا إذا ما استعرضنا اختبارات القراءة المشهورة وجدناها تختلف فيما تقيس من جوانب السلوك اختلافاً يزيد عما نجده في قياس أي مهارة أخرى . ولم يختلف مصممو اختبارات القراءة على جوانب السلوك التي يجب أن تتضمن خلال قياس القراءة فحسب ، بل وإنه أيضاً اختلفوا في تعريفاتها . اختلفوا في تحديد مدى تأثير عوامل السرعة ، والفهم ، والمفردات . . . إلخ : كما اختلفوا في تحديد أهداف قياسها .

وقد قام تراكسلر Traxler بفحص ٢٤ اختباراً للقراءة ووجد أنها تقيس ٤٨ مهارة مختلفة .. ولا يعني هذا أن القراءة قدرة أو مهارة

أو استعداد مركب من عدد ضخم من العوامل النوعية أو الخاصة أو الفرعية . إذ ادعى أحد الاختبارات أنه يقيس عدداً من مهارات القراءة المختلفة تماماً ، ولكن فحص معاملات الارتباط بين درجات المفحوصين في هذه المهارات بين أنها في الواقع وظيفة واحدة تقاس تحت أسماء متعددة .

وأحد طرق توضيح ما إذا كانت قدرات القراءة مختلفة حقاً هو التحليل العامل . فقد وجدت العوامل الآتية شائعة في ٢٥ اختباراً للقراءة ومهارات المذاكرة : الميل إلى القراءة بعناية (وهو اتجاه أو جهاز استجابات) ، والاستدلال الاستقرائي ، وسرعة القراءة ، والقدرة اللفظية ، والمفردات ، وقراءة الخرائط . ولهذا يجب عند الحاجة إلى استخدام اختبار للقراءة أن نعرف مضمون الاختبار ، وأن نحدد أى قدرة قرائية نريد أن نقيس ، وأن نرجع إلى مراجع يعتمد عليها في تحليل الاختبارات التي سنستخدمها .

اختبارات القراءة إذن من أهم الاختبارات التحصيلية التعليمية . وهي تنقسم ، كما تنقسم كل اختبارات التحصيل المقننة الخاصة ، إلى ثلاثة أقسام هي :-

- ١ - اختبارات التهيؤ readiness .
- ٢ - اختبارات المسح survey .
- ٣ - اختبارات التشخيص diagnostic Tests .

وتطبق اختبارات التأهب عند قبول التلاميذ بالمدارس ، لتحديد ما إذا كان الطفل ذا مهارات ومعلومات تمكنه من تعلم القراءة . أما اختبارات المسح فهي تستخدم المقاييس الكلية Overall measures في معرفة مستوى القدرة القرائية . بينما تستخدم اختبارات التشخيص في تعيين نقاط الضعف والقوة التي يعاني بسببها الفرد من مشكلات تعليمية أو نفسية . ولهذا نجد اختبارات تشخيص القراءة تهتم بالتحليل الكيفي لجوانب قدرات القراءة

ودراسة الأخطاء النوعية المعينة التي غالبا ما يقع فيها تلاميذ بذاته و٥١٠٠، تكرار الوقوع في هذه الأخطاء . ومن ثم يمكن بالاستقصاء التالى معرفة مصادر هذه العيوب وأسبابها وربطها غالبا بنوع من الخبرات المؤهلة أو بإرجاعها إلى عدة عوامل متفاعلة متكاملة كالأسرة والمدرسة و : . الخ .

« اختبارات التهيؤ فى القراءة »

برغم أن اختبارات التهيؤ تنتمى إلى اختبارات الاستعداد ، إلا أن استخدامها فى المواقف التعليمية وارتباطها بأنواع اختبارات القراءة الأخرى خاصة وبأنواع اختبارات التحصيل فى الميادين الأخرى عامة ، يحتم الحديث عنها هنا . هذا برغم أن اختبارات القراءة ، خاصة فى مستوى المدرسة الابتدائية ، والمدرسة الإعدادية ، تقرب من اختبارات الذكاء العام من حيث الوظيفة المقاسة ، ولكن اختبارات التهيؤ فى القراءة نهتم بقياس القدرات التى يعتقد أن لها أثر فى تعلم القراءة وفى النجاح المدرسى . واختبارات التهيؤ فى القراءة تقيس الوظائف الآتية :

١ - التمييز البصرى والسمعى .

٢ - الضبط الحركى .

٣ - الاستدلال اللفظى :

٤ - المفردات .

٥ - المعلومات العامة .

ومن أهم اختبارات التهيؤ فى القراءة وأكثرها استخداما نذكر الاختبارات الآتية :

١ - اختبارات جيئس للتهيؤ فى القراءة .

٢ - اختبارات المتروبوليتان للتهيؤ في القراءة :

Metropolitan Reading Readiness Tests

وهما اختباران جماعيان . أما الاختبارات المشهورة التي تستخدم في المواقف الجماعية والفردية أيضاً فأهمها :

اختبارات مونرو للاستعداد للقراءة : The Monroe Reading

A ptitude Tests

كما أن هناك اختبارات فردية للتهيؤ تهتم بالمقاييس النفسية والجسمية مثل بطارية بيتس للاستعداد للقراءة :

Betts Ready to Read Battery.

وتختلف هذه الاختبارات عن بعضها ، فضلاً عما ذكرنا من أوجه ، في المدى والوظائف المقاسة . كما تختلف أيضاً في بعض الخصائص الإحصائية والقياسية الأخرى . فمثلاً نجد أن معاملات صدقها تتراوح بين ٠.٥ إلى ٠.٨ ، وكذلك تتراوح معاملات ثباتها في نفس المدى . كما أن معاملات الارتباط بين كل منها وبين اختبار صادق للذكاء تختلف في مدى كبير . وسنختار أكثر هذه الاختبارات شيوعاً :

« اختبارات المتروبوليتان للتهيؤ في القراءة »

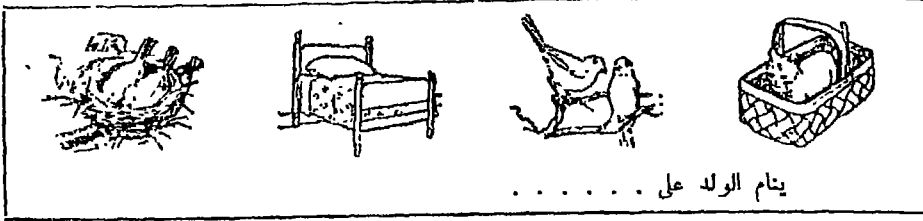
اختبارات المتروبوليتان للتهيؤ في القراءة : Metropolitan Reading

Readiness Test تتضمن ٦ اختبارات هي :

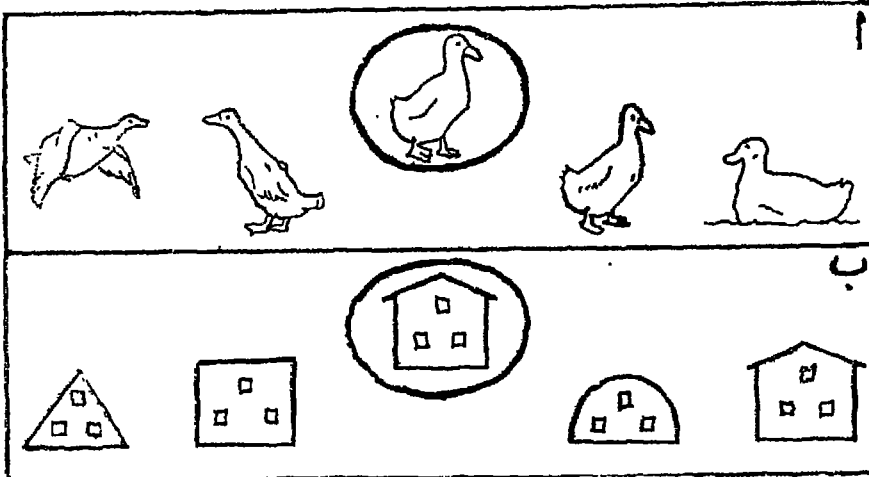
الاختبار الأول : معنى الكلمة word meaning وفي هذا الاختبار

ينطق القاصص بكلمة وعلى المفحوص أن يختار الصورة التي تمثل هذه الكلمة من بين ٤ صور في كل صف (انظر شكل ١٨) .

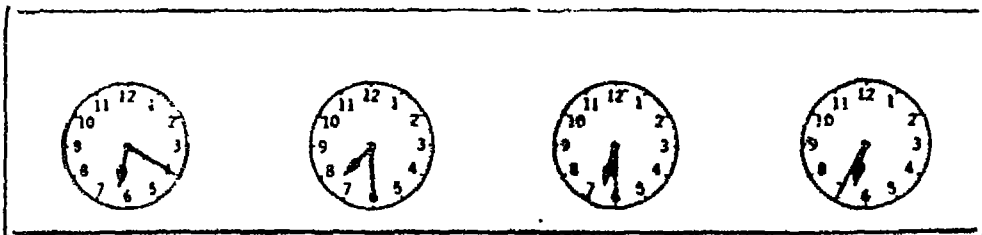
(شكل ١٩) عينة من اختبار المتروبوليتان للتمييز في القراءة



يقرأ الفاحص الجملة فيشير إلى صورة الشيء الذي يكمل اسمه الجملة



والمفروض هنا أن المفحوص سيضع دائرة حول أحد الأشكال الأربعة التي يطابق الشكل المحاط بدائرة



علم على الساعة التي تشير إلى السادسة والنصف

الاختبار الثاني : الجمل Sentences : وفيه يختار المفحوص جملة مصورة في أحد صور من بين الصف الذي يتكون من ٤ صور . وينطق الفاحص الجملة المكتوبة وسط هذا الصف ثم ينتقل إلى جملة تالية من بين ٤ جمل في الصف الثاني وعلى المفحوص أن يتعرف على الصورة التي تمثلها فيوصل بينهما خطأ .

الاختبار الثالث : المعلومات info motion : وفي هذا الاختبار يختار المفحوص صورة من كل صف تؤدي الوظيفة التي يقوم بها الاسم الذي ينطق بها الفاحص . فمثلا يقول الفاحص « اختر الصورة التي تقوم بوظيفة التقاط الصورة فيعلم المفحوص على آلة التصوير ويوصل بينها والجملة خطأ .

الاختبار الرابع : المزاوجة matching : وفيه يقوم المفحوص بوضع خط يوصل بين صورة في صف يتكون من ٤ صور وبين صورة أخرى تماثل إحداها . والصورة التي عليه أن يجد مثلتها محاطة بخط (انظر شكل ١٩ الأوزة) والصور الأربع في الصف تقرب من بعضها في خصائص وتختلف في أخرى . فمثلا في أحد الصفوف بطة وأوزة ودجاجة راقدة وأحد الطيور المنزلية الأخرى ووسط هذه الصور توجد صورة أوزة محاطة بدائرة . ويطلب من المفحوص أن يوصل هذه المحاطة بمثلتها . وهذا يتطلب منه معرفة أوجه الشبه والاختلاف بين الصور ومعرفة الخصائص الهندسية لكل منها أحيانا .

الاختبار الخامس : الأعداد Numbers : وهنا يطلب من المفحوص اقيام ببعض العمليات الحسابية البسيطة ومعرفة بعض المفاهيم الكمية . وهذا الاختبار يشبه ما تأتي به اختبارات الاستعدادات الفارقة واختبارات الذكاء العام .

الاختبار السادس : النسخ copying ويقوم المفحوص فيه بنقل أى إعادة رسم بعض الأشكال الهندسية البسيطة والحروف والأرقام . وهذا يتطلب درجة من النمو الجسمى وأخرى من النمو العقلى .

ولهذه البطارية صورتان متكافئتان . ومعايرها هى الدرجات المثيية وتستخدم كاختبار للتهيؤ فى القدرة وفى الأرقام كما تستخدم كمقياس للاستعداد العام للتعليم .

تكونت عينة تقنينها من ١٥٠٠٠ تلميذاً وكانت تطبق فى أول شهر يدخل فيه التلميذ المدرسة الابتدائية فى السنة الأولى . ويتبع هذا تطبيق اختبار « رسم رجل » Draw a man Test . ويعتبر هذا الاختبار إضافياً . ويصحح اختبار رسم الرجل فيحصل التلميذ على درجة نضعه حسبها فى فئة من خمسة فئات . ثم يوضع التلميذ حسب درجته فى البطارية فى فئة . ويوجد معامل الارتباط بين الفئة فى البطارية والفئة فى الاختبار ويكون هذا المعامل دليلاً على صدق البطارية .

اختبارات جيتس للتهيؤ فى القراءة

صممت هذه الاختبارات لقياس التهيؤ لتعلم القراءة ، وللتنبؤ بسرعة نمو القدرة القرائية ، ولتشخيص مكانة التلميذ وبيان حاجاته فى كل قدرة هامة يحتاجها عند تعلم القراءة .

وتعطى درجة مركبة أو متوسط درجات التلميذ على كل اختبار فرعى لتحديد مرحلة التهيؤ التى وصل إليها هذا التلميذ . وهذه الدرجة المركبة قد تستخدم فى تقرير ما إذا كان التلميذ مستعداً للبدء فى البرنامج المنتظم للدراسة ، ولتقدير نجاحه المحتمل إذا كان قد بدأ القراءة فعلاً ، ولتقدير ما هى الأعمال التى يحسن أن يبدأ بها قبل تعلمه القراءة ، وهذه الدرجة

المركبة ، على أى حال تغطيها فقط معلومات عامة عن استعداد التلميذ للقراءة .

وفيقيد المدرس من هذه الاختبارات فى مقارنة درجات التلميذ الواحد فى مختلف القدرات القرائية وفى مقارنته بغيره فى قدرة معينة . وقد تم إعداد هذه الاختبارات بعد القيام بسلسلة من البحوث التى استمرت أعواماً طويلة . وفى إحدى هذه الدراسات استخدم حوالى ١٠٠ اختبار مختلف ، وامتحانات ، وتقديرات . الخ . ذلك لتحديد أفضل طرق قياس هذا التهيؤ . والقدرات التى كشفت عنها هذه الدراسات وغيرها من الدراسات المماثلة وضعت لها اختبارات من النوع المألوف وجربت على كل تلاميذ إحدى البلدان التى اتبعت طرفاً مختلفة فى التعليم . وبعد دراسة نتائج هذه الاختبارات ، أعدت اختبارات مراجعة وجربت على مجموعات أخرى من التلاميذ بين سنة ١٩٣٨ وسنة ١٩٣٩ . ويمثل الاختبار الحالى مدى الاستفادة من هذه الدراسات . وهو يشمل خمس اختبارات فرعية هى : -

الاختبار الأول : تعليمات الصور Picture Directions فى هذا الاختبار صور ثلاث الأولى صفحة كاملة عن منظر مزرعة ، والثانية نصف صفحة عن منظر مدينة ، والثالثة نصف صفحة عن المنظر الداخلى لمحل تجارى عام .

يقوم الفاحص بالتحدث إلى التلاميذ عن المواقف أو الأشياء الممثلة فى الصورة ويطلب منهم تنفيذ تعليمات معينة بوضع علامات على الصورة وهكذا يمكن قياس القدرات الآتية : -

(أ) القدرة على الإصغاء لما يقوله الفاحص .

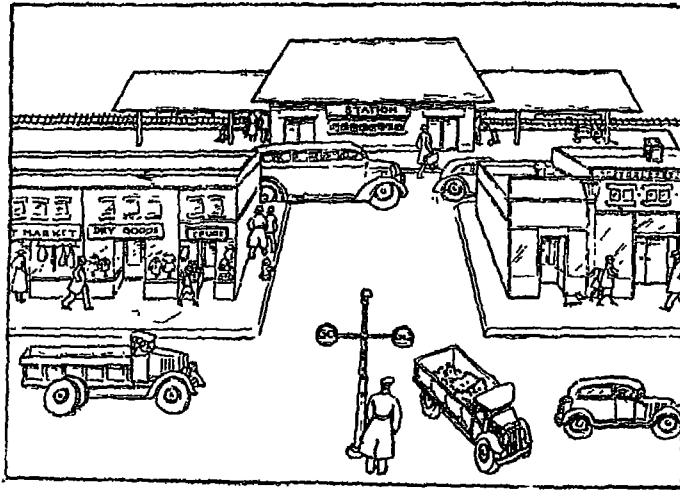
(ب) القدرة على فهم ما قاله .

(ج) القدرة على تذكر هذا القول لمدة قصيرة ٥

(د) القدرة على تمثيل واستخدام كلمات ومفاهيم الحياة اليومية الهامة المتنوعة التي تتعلق بالريف ، والمدن ، والمحلات ٥

(هـ) القدرة على تفسير صور يشيع وجودها في الكتب الابتدائية ٥ وهذا يتضمن فهم الصورة الكلية ، ووضع الأشياء فيها ، : : الخ .

(و) القدرة على استخدام كل القدرات السابقة في تنفيذ التعليمات ٥ ومادام هذا الاختبار يتطلب أن يصنع التلميذ إلى ما يقوله المدرس وأن ينفذ التعليمات لمدة زمنية متوسطة الطول ، فالاختبار إذن يقيس الانتباه في نشاط يشابه ما يحدث في أثناء الدرس الشفوي المعتاد في المدرسة ٥



(الشكل ٢٠)

فقرة من اختبار جينس للتهيؤ في القراءة وهي الفقرة الثانية من اختبار تعليمات الصور

وهو يقيس ، بالتالى ، عدداً من القدرات القاعدية على تعلم القراءة فى موقف الفصل . والقدرة التى يقيسها هذا الاختبار هى محصلة أوجه النشاط التى يجربها التلميذ قبل أن يدخل المدرسة إطلافاً . وهذه الخبرات تتعلق بكل أنواع الصور ، والإصغاء إلى كل أنواع التعليقات ، والجمل ، والقصص ، والتعليقات ، وتنفيذ كل أنواع التعليقات فى البيت أو مدرسة رياض الأطفال .

الاختبار الثانى : مزاججة الكلمات Word Matching : هذا الاختبار

مصمم لقياس مدى تعرف التلميذ على الكلمة المطبوعة سواء كان يستطيع قراءتها أم لا يستطيع . والكلمات كلها شائعة فى كتب الأطفال ، وعناوين الجرائد ، والمجلات ، والمجلات .

وفى الاختبار ١٨ مربعا فى كل ٣ كلمات منها كلمة مكررة أى أن المربع يشمل ٤ كلمات . ويضع المفحوص علامة بين الكلمة المكررة مرتين .

الاختبار الثالث : المزاججة بين الكلمة والبطاقة word-Card Matchiny

وفى هذا الاختبار تقدم للتلميذ عدة بطاقات، على كل منها كلمة يكتبها المدرس بخط معتدل كبير جداً والحروف ذات حجم مقنن وعدد البطاقات ٢٠ بطاقة . ويطلب الفاحص من التلميذ أن يضع علامة على الكلمة التى تشابه الكلمة المطبوعة فى البطاقة من بين ٤ كلمات فى كل مستطيل مطبوع فى كراسة التعليقات وعدد المستطيلات ٢٠ مستطيلاً .

الاختبار الرابع : السجع Rhyiming وفى هذا الاختبار صفحتان كل

منهما تشمل ٧ مجموعات من الكلمات . وكل مجموعة تتكون من ٤ كلمات . وفيه يساعد الفاحص التلاميذ على نطق كلمات كل مجموعة ثم يطلب منهم وضع علامة على كلمة ، فى كل مجموعة ، تتفق مع كلمة أخرى مكتوبة أمام المجموعة ينطق بها الفاحص مثل نام ، لعب ، يمشى ، أصبح . وأمام

هذه الكلمات كلمة مثل قام وعلى المفحوص أن يضع علامة تحت أى من الأربعة التى تتفق معها (والإجابة الصحيحة هى « نام ») .

الاختبار الخامس : قراءة الحروف والأرقام Reading Letter and

Number : وهذا الاختبار يتكون من ثلاثة أجزاء . الجزء الأول حروف مكتوبة بخط كبير capital والثانى مكتوبة بحروف صغيرة small والجزء الثالث الأرقام من صفر إلى ٩ . والحروف ، كبيرة كانت أو صغيرة ، والأرقام غير مرتبة . وينطق المدرس بالحرف أو بالرقم وعلى المفحوص أن يشير إليه فى كراسة التعليمات :

ومن الطبيعى عند نقل هذه الاختبارات إلى العربية أو إعدادها كى تناسب ظروفنا المختلفة يجب ، فضلاً عن عمليات التقين والصدق والثبات أن تراعى قواعد اللغة العربية كأن تستبدل حروف النسخ والرقعة بالحروف الكبيرة والصغيرة وأن تحل صور مألوفة فى بلادنا محل صور غريبة فى ثقافتنا مألوفة فى بلادها الأصلية :

والمعيار المستخدم هو الدرجات المئانية ، ومعاملات الثبات مرتفعه جداً ، ومعاملات الصدق بالارتباط مع اختبارات صادقة عالية أيضاً : إذ أن متوسطها ٠.٧٦ :

اختبارات مسح القراءة

تقيس هذه الاختبارات فى كل مستويات التعليم المستوى العام الذى وصل إليه المفحوص فى نموه القرائى وهى تستخدم لتصنيف التلاميذ الذين يمكن مساعدتهم بواسطة القراءة العلاجية ، وللتنبؤ بالنجاح فى المناهج الدراسية . ولتأكد مما إذا كانت القدرة القرائية لدى المفحوص ضعيفة وأن هذا هو السبب الذى من أجله كان أدائه ضعيف فى المدرسة أو فى اختبار عقلى جمعى : تشمل القدرة القرائية جانبين . هما السرعة والفهم وبالتالى ، ينجح

اختبار مسح القراءة إذا وضع في الاعتبار تدخل كل من هذين العاملين معاً . وقد حاول بعض علماء القياس اختبار كل عامل على حدة ولكنهم لم ينجحوا في هذا . وليس ذلك بغريب علينا إذ أنه يرجع مبدأ عام متفق عليه في علم النفس عموماً والقياس خصوصاً وهو أن الكل المركب لا يحسن تفكيته إلى مكوناته وقياس كل على حدة :

ويمكن ، من حيث النظرية ، عزل عاملي السرعة والفهم Comprehension وقياس كل منهما على حدة بتثبيت أحدهما وتقدير الآخر . فيمكن مثلاً إلغاء أثر السرعة أو تقليل أهميتها ذلك بإعطاء وقت كاف للإجابة عن الاختبار . وبالتالي تمثل درجة المفحوص قدرته على الفهم . أما عند تثبيت عامل الفهم بوضع اختبار يتوقع أن جميع المفحوصين يفهمون مواده ، لقياس عامل السرعة ، فهو أمر أصعب من سابقه . إذ أن لكل فرد عدد كبير من السرعات القرائية ، وتختلف سرعة القراءة باختلاف الهدف من القراءة ونوع ومستوى وصعوبة وطبيعة المادة المتروكة . ويميل مجيدو القراءة إلى تخفيض سرعتها كلما زادت صعوبة المادة التي يقرأونها .

المشكلة بالنسبة للفاحص إذن هي : كيف يمكن وضع قطعة يقرأها كل التلاميذ بنفس الدرجة من التركيز وب نفس الجهد للفهم ؟ هذه المشكلة يواجهها المدرس عندما يريد قياس سرعة القراءة . مما يزيد الأمر تعقيداً أنه يمكن أن تزيد السرعة ويقل التركيز والفهم فيقل التحصيل . وإذا لم يضبط الجهد المبذول في الفهم ، بأن يطلب من المفحوص أن يجيب عن أسئلة تتعلق بما قرأه ، أو أن يكتشف السخافات أثناء القراءة ، كانت النتيجة هي نمونين المادة التي يقرأها التلميذ بسرعة حتى يحصل على درجة عالية في أحد اختبارات سرعة القراءة .

ومن هنا يتبين لنا أنه يصعب فصل سرعة القراءة عن مستوى الفهم وهذه المشكلة تتعرض لها اختبارات المسح والتشخيص . ووجه الصعوبة ، كما قلنا ، هو أن سرعة القراءة تعتمد على صعوبة المادة المقروءة وسبب قراءة المفحوص لها . ومن هنا أيضاً تكون للمفحوص عدة سرعات في القراءة تختلف من اختبار إلى آخر وقد تختلف بين فقرات الاختبار الواحد . كما أن سرعة القراءة تتوقف على خبرات المفحوص الخاصة .

وعلى هذا فاختبارات مسح القراءة ، وخاصة عند استخدامها في المدرسة الابتدائية والإعدادية ، تتبع اختبارات التصنيف العام . ولكنها ، عند استخدامها في مستوى المدرسة الثانوية والجامعة ، تصبح اختبارات خاصة لأنها تقيس مجالات خاصة مثل العلوم والجغرافيا والتاريخ والأدب . وبهذه الطريقة يمكننا أن نتبين أن اختبارات المسح تواجه الصعوبات التالية :

١ - تتداخل الدرجات على السرعة مع الدرجات على الفهم بحيث يستطيع التلميذ أن يرفع إحداها على حساب الأخرى ، وهو قد يفعل هذا دون قصد .

٢ - تكون الاختبارات التي يفترض أنها تقيس الفهم ، غالباً ، مشحونة بعامل السرعة : ومثل هذه الاختبارات محدودة القيمة إذا أردنا بها التشخيص ، ولكنها تعتبر أدوات تنبؤية جيدة .

٣ - يغطي اختبار القراءة مجالا محدوداً مختاراً من محتوى المواد الدراسية ، وبما أن القدرة القرائية تختلف بعض الشيء باختلاف المادة ، فإننا نجد نتيجة لهذا السبب ونتيجة للطريقة التي اتبعت في اختيار فقرات الاختبار ، أن القيمة التنبؤية باختبار مسح القراءة تتحدد بمجال المادة الدراسية التي اختيرت فقراته منها . وعلى هذا لا يصح استخدام اختبار .

لمسح القراءة في مادة كالتاريخ كأداة للتنبؤ أو التشخيص في مادة أخرى كالحساب ولا حتى في مادة قريبة من التاريخ مثل الجغرافيا .

٤ - تختلف اختبارات مسح القراءة في مستويات الفهم التي تقيسها . فبعض الاختبارات يتطلب فهما بسيطاً عاماً ، بينما يتطلب في البعض الآخر تفسيراً عميقاً تاماً شاملاً . ولكننا إذا أردنا أن نتعرف على القارئ الماهر لا نكتفي بفهمه الجمل ، أو هذا ما يجب ، بل وأيضاً باستخلاصه الفكرة العامة في القطعة التي قرأها ، وجميعه الأفكار من الجمل المتفرقة ، وبتابعته مناقشة منطقية تستغرق أحياناً صفحات . ولكن اختبار مسح القراءة الواحد لا يمكنه أن يقيس كل أنواع ومستويات الفهم منفصلة بنفس الكفاءة .

تشخيص القراءة

الاختبار التحصيلي التشخيصي ، في أحسن صورة ، هو أداة تأثيرية . فعلى المدرسين ، مستعينين بهذه الاختبارات أو بغيرها أو مستخدمي طرقهم الأخرى الخاصة ، أن يدرسوا كل تلاميذهم ليحددوا أسباب مشكلاتهم . والاختبار التشخيصي يحاول أن يزيد كفاءة هذه الدراسة فيضع أمام ناظرى المدرس كل فهم يمكن الوصول إليه من بحث أسباب ومصادر القشل في عمليات التعلم . والاختبار التحصيلي التشخيصي الجيد يوجه انتباه المدرس إلى كل أو أغلب وأهم نواحي العملية القرائية التي يجد فيها التلميذ صعوبة . ومثل هذه الجهود يجب أن تستند إلى بحوث شاملة لتحديد أنماط الأخطاء التي يغلب أن يقع فيها التلاميذ ، والتي يغلب أن يعاني منها معظم أفراد كل فئة مرضية . ثم توضع هذه الأخطاء في قوائم تشخيصية .

وهناك اختبارات تشخيصية أبسط من الأنواع السابقة . وبعضها يجرى جمعياً . وهي تحوى اختبارات فرعية لقياس مختلف أنماط القدرة القرائية .

ويمثل الأداء بصفحات نفسية تبين المستويات النسبية لقدرات الفرد القرائية . ومنها يمكننا كشف نواحي الضعف والقوة لديه . وأشهر هذه الاختبارات :

اختبار أيوا للقراءة الصامتة

يشمل اختبار أيوا للقراءة الصامتة The Iwa Silent Reading Test الاختبارات الفرعية التسع الآتية :

الاختبار الأول والثاني : سرعة القراءة : يقرأ التلميذ مقاليتين قصيرتين ، ثم يجيب عن أسئلة بسيطة عن القطعتين . وتحسب للفرد درجتان إحداهما للسرعة والأخرى للدقة . ولأنه تحسب درجتان ، عد الاختبار اختبارين .

الاختبار الثالث : القراءة الموجهة : directed Reading يجيب التلميذ عن أسئلة عن الفقرات بالإشارة إلى مكان الإجابة عن السؤال بين كلمات الفقرة بوضع خط تحتها Underlining . وهذا الاختبار ، كالاختبارات الفرعية التالية ، له توقيت منفصل لا يكفي لإنهاء الاختبار غالباً .

الاختبار الرابع : الشعر : يقرأ التلميذ قطعاً من الشعر ويجيب عن أسئلة عن معناها .

الاختبار الخامس : معنى الكلمة : تعرض ٤ قوائم من الكلمات في العلوم الاجتماعية ، والعلوم ، والرياضيات ، واللغة . ويجوار كل كلمة خمسة متغيرات يختار الصحيحة منها كمعنى للكلمة . ويتم إجراء هذا الاختبار في ٤ مراحل كل مرحلة من ٩٠ ثانية . وعند انتهاء هذه المدة ينتقل التلميذ إلى المرحلة التالية . وتوضع درجة عن أداء كل جزء .

الاختبار السادس : معنى الجملة : يقرأ التلميذ جملاً تعالج معنى كلمات وتتضمن معلومات بسيطة . وكل جملة هي حقيقة أو معنى تفسري لكلمة ويحجب المفحوص بصحيح أو خطأ عن كل .

الاختبار السابع : معنى الفقرة : Paragraph meaning يقرأ التلميذ فقرات ويحجب أسئلة عن كل . وهذه الإجابة تطلب منه إعمال الذاكرة المباشرة والتعرف على الفكرة الرئيسية .

الاختبار الثامن : استخدام الفهرست : Use of index يعرض على التلميذ فهرست موجز ويطلب منه أن يحدد صفحات الكتاب التي يعتقد أنها تتضمن إجابات عن أسئلة معينة .

الاختبار التاسع : الكلمات الرئيسية Key Words : يشير التلميذ إلى الكلمات التي يجب أن يبحث تحتها في الفهرست كي يجد الإجابة عن سؤال مثل : « ما هي قيمة إنتاجنا السنوي من المواد المعدنية » ؟

وتختلف اختبارات التشخيص في القراءة من حيث درجة تحليل كل اختبار للقدرة القرائية إلى عدد من القدرات الفرعية . فمن جهة نجد الاختبارات الجمعية وفيها يحصل التلميذ على ٣ أو ٤ درجات منفصلة كل عن إحدى الجوانب المقاسة ، ومن أخرى نجد الاختبارات الفردية التي توضح الأخطاء النوعية التي يقع فيها بعض المفحوصين . وتشمل الاختبارات الأخيرة قوائم مقننة لتلك الأخطاء . ويستخدم بعضها جهاز التاكستسكوب لتتبع سرعة حركات العين أثناء القراءة . وهذا الجهاز ، ببساطة ، يمكن الحزب من وضع البطاقة فيه ويجذب أو يدار محرك فيستمر عرض البطاقة مدة يتحكم فيها الفاحص . وأثناء هذه المدة يستطيع المفحوص أن يقرأ المكتوب على البطاقة إن كان يقرأ بالطريقة الكلية . وهناك أزمنة

مقننة لتوسط المدة التي تستغرقها قراءة الكلمة . فإذا زاد عنها التلاميذ كان يقرأ بالطريقة الجزئية .

ومن أهم أمثلة اختبارات القراءة التشخيصية الجمعية في المدارس الابتدائية والإعدادية اختبارات جيتس للقراءة الأساسية Oates Basic Reading Tests وهي تقيس السنوات الدراسية من الثالثة الابتدائية حتى قرب نهاية المدرسة الإعدادية ؛ وهو جزء من سلسلة اختبارات لها شهرتها العالمية . ويشمل ٤ مقاييس منفصلة هي : اختبار فهم المعنى العام - اختبار التنبؤ بنتائج بعض الأحداث المعينة - اختبار تحديد اتجاهات معينة - اختبار ملاحظة التفاصيل . ولهذه السلسلة من الاختبارات ٤ صور متكافئة .

ويمكن اعتبار اختباري « أيوا » و « جيتس » اختبارات للمسح ، بل إنها أيضاً تشمل اختبارات مسح للقراءة . ومعاملات الارتباط بين الأجزاء الأربعة في اختبار « جيتس » مرتفعة . ومعاملات ثبات كل منها مرتفعة أيضاً . أما معاملات ثبات اختبارات « أيوا » فهي منخفضة ، وإن كانت معاملات الارتباط الداخلية بينها منخفضة أيضاً ، مما يجعلها أداة تشخيصية لا يمكن الاعتماد عليها تماماً . كما أن درجات الأفراد عن اختباري السرعة في « أيوا » قليلة الدلالة ، للأسباب التي ذكرناها .

ولكن هذه الانتقادات توجه إلى كل اختبارات القراءة ، فضلاً عن نقد أساسي آخر هو أنها لا تقيس المهارات الأولية المطلوبة في الأعمال المكتبية . وإن كانت الدرجات عن الاختبارات الفرعية منخفضة الثبات ، إلا أن معاملات ثبات الدرجات الكلية عن الاختبارات المتضمنة في كل بطارية منها معاملات مرتفعة وهذه ميزة فضلاً عن ميزة أخرى هامة هي أن معايير هذه الاختبارات حسبت من عينات كبيرة ومثلة من التلاميذ .

وأما بطارية امتحانات مونرو للقراءة التشخيصية Monroe Diagnostic Reading Examination فهي تعتبر برنامجاً للاختبار أكثر منها بطارية

مفردة ؛ وإكمال هذا البرنامج يحتاج الأمر عدة ساعات وتتطلب تطبيق اختبار للذكاء ، وعدة اختبارات مقننة للمسح في القراءة ، وبعض الاختبارات التحصيلية الأخرى ، وبعض اختبارات القراءة الخاصة . والبطارية جزء إضافي للتحليل الكيفي لصعوبات القراءة والبيانات المفصلة عن الأخطاء :

والبطارية الرابعة هنا هي بطارية دوريل لتحليل صعوبات القراءة Durrell Analysis of Reading Difficulty Battery فهي أقصر من السابقة وتطبق من السنوات الأولى حتى السادسة الابتدائية . وقد تم تقنين ما أسفرت عنه بحوث التحليل المفصل لصعوبات القراءة الشفوية والتذكر والاسترجاع والتعرف على الكلمات والنطق والكتابة والحجاية . وهي تستخدم جهاز التاكستسكوب لمقارنة سرعة التعرف على الكلمات بقراءة الكلمة في ظروف معينة ثم قراءتها في أخرى بسرعة أخرى وهكذا . وتعطي درجة جزئية عن تصحيح المفحوص لقائمة من الكلمات الخاطئة بعد أن يكون قد فهم قائمة بالكلمات الصحيحة . وقننت هذه البطارية على ٤٠٠٠ تلميذ .

ويجب أن نخص بالذكر اختبارات القراءة التشخيصية Diagnostic Reading Test التي قامت بها جمعية الاختبارات التشخيصية الأمريكية . وهي من جزئين أحدهما مسحي والآخر تشخيصي . واختبارات الجزئين كلها جمعية فيما عدا اختبار واحد في الجزء التشخيصي هو اختبار القراءة الشفوية . وتستخدم هذه البطارية ابتداء من أول المدرسة الإعدادية حتى السنة الأولى بالجامعة .

وللجزء المسحي ثمانية صور مكافئة . ويحصل الفرد فيه على ٣ درجات منفصلة لأن البطارية تقيس ٣ مجالات هي : سرعة القراءة مع الفهم المعقول للمادة المقروءة - المفردات العامة - القدرة على فهم الكتب .

المدرسية . وقد أوجدت معايير للفرقة الدراسية بناء على الدرجة المركبة عن البطارية ككل . كما يعد التلميذ في فرقة دراسية معينة بناء على الدرجات عن الاختبارات الفرعية الثلاث . والعينة التقنية أخذت من عدة مناطق جغرافية ومدن ومدارس عامة وخاصة وكثير من المعاهد والكليات .

والقسم التشخيصي يتطلب وقتاً أطول من المسحى ، ويمكن أن يطبق بصورة جماعية في الفصول . ويمكن الاكتفاء بتطبيق الأجزاء من البطارية التشخيصية التي ظهرت صعوبات القراءة فيها أثناء تطبيق الجزء المسحى . ويطبق الجزء المسحى التشخيصي في المراحل الثلاث السابقة بالإضافة إلى اختبار لمهارة التعرف على الكلمات . وتتضمن البطارية التشخيصية مفردات في قواعد اللغة ، والأدب ، والعلوم ، والدراسات الاجتماعية ، والرياضيات ، وفهم مواد من الكتب الدراسية يقرأها المفحوص أو تقرأ له . وتحسب سرعة القراءة للمواد المختلفة وعن ربط الكلمات سواء كان ذلك شفويّاً أم في صمت .

وآخر هذه الاختبارات يتضمن قوائم للكلمات ، وأخرى لل فقرات paragraph . والجزء الشفوي يتأثر بعوامل مثل مهارة استخدام المفردات ، والأخطاء في التعرف على الكلمات ، وطريقة ربط كلمات غير معروفة . وهناك جزء للملاحظات الكيفية ، وقائمة بالأخطاء الشائعة ، وربط الكلمات في صمت . وهذا الجزء يعطى درجة للقدرة على تقييم الكلمة ، وأخرى للقدرة على سماع وتمييز الكلمات .

والمعيار هو الدرجات المئينية . ومعاملات ثبات القسمين أكثر من ٨٠٪ ، ويجب ، عند الحديث عن الاختبارات التشخيصية في القراءة ، أن نتحدث عن :

أولاً : بناء هذه الاختبارات .

ثانياً : تفسيرها واستخدامها .

والجزء الثاني من الحديث موجه إلى الأخصائيين النفسيين الذين يستخدمونها فردياً أو في الفصول . فإن تشخيص صعوبات القراءة ، وطريقة علاجها مسألتان تحتاجان إلى أخصائيين مدربين . إذ أن تشخيص عجز القدرة القرائية يعتمد على الطريقة الإكلينيكية في دراسة الحالات ، وإلى جمع معلومات إضافية عن الإمكانيات الحسية ، والنمو الحركي ، ومعلومات عن الصحة ، والعلاج ، والتعليم ، والمنزل والأسرة ، والنواحي الانفعالية عن الفرد . ذلك أن تأخر القدرة القرائية قد يكون عرضاً لاضطراب الشخصية . وبما أن الاختبارات الجمعية للمسح والتشخيص تتطلب الاهتمام بالأفراد ، مما يجعلنا ننتهي إلى القول بأن تقييم نقاط الضعف والقوة في القدرة القرائية ، يدخل في صميم الخدمات الإكلينيكية .

الرياضيات

اهتم الكثيرون بتشخيص العجز والتأخر في القدرات الرياضية اهتمامهم باختبارات القراءة ، ابتداء من مستوى المدرسة الابتدائية حتى أعلى مستويات التعليم . ومع أن اختبارات الرياضيات ليست بنفس الكثرة والشيوع ، إلا أنها تنقسم أيضاً إلى الأنماط الثلاث السابقة وهي :

أولاً : اختبارات التهيؤ في الرياضيات .

ثانياً : اختبارات مسح الرياضيات .

ثالثاً : اختبارات التشخيص في الرياضيات .

ويغلب أن تجمع الاختبارات بين الفئتين الأولى والثانية في بطارية واحدة .

التهيو والتنبؤ

نجد أن بطارية المروبوليتان ، التي سبق عرضها ، تعطى درجة للاستعداد للقراءة عن اختبار واحد للأرقام . ومن ثم فهي ليست ذات دلالة كبيرة . وإن كانت تشير إلى الاستعداد للنجاح المدرسي خاصة إذا طبقت على أطفال السنة الأولى الابتدائية . وتقوم اختبارات التهيو بنفس الوظيفة في المستويات الأعلى حين تقدم اختبارات في الجبر والهندسة .

ومن أهم هذه الاختبارات ، اختبار « لى » لقدرة الجبر Lee Test of Algebra Ability وهي تتكون من ٤ اختبارات فرعية للمسائل الحسابية - ومقارنة الأرقام - وسلاسل الأعداد - والمعادلات . أما اختبار أورليانز للتنبؤ بالجبر Orleans Algebra Prognosis Test فهو يتكون من ١٢ جزءاً منها ١١ جزءاً عن دروس بسيطة في الجبر متبوعة بأسئلة عن محتواها . والجزء الأخير هو تلخيص لما تعلمه التلاميذ من المواد الجديدة .

اختبار « لى » إذن يقيس التهيو أكثر مما يركز على التنبؤ ؛ بينما اختبار « أورليانز » يقوم المفحوص فيه بعمليات بسيطة تهتم بالتنبؤ أكثر . وقد حسبت معاملات صدق كل من الاختبارين بإيجاد معاملات الارتباط بين الاختبار ودرجات الفترات وبين درجات الاختبار ودرجات نفس التلاميذ في اختبارات تحصيلية أخرى صادقة .

المسح

اختبارات مسح الرياضة في المدرسة الابتدائية تقتصر على الحساب وفي المستويات الدراسية الأعلى تشمل أغلب فروع الرياضة . وهي تكون جزءاً هاماً من البطاريات التعليمية . وتطبق هذه الاختبارات في مناهج الرياضة المختلفة مثل الحساب والجبر والهندسة . ومن أمثلتها اختبار أبوا للمهارات

الأساسية عند كل تلميذ Iwa Every Pupil Test of Basic skills ونجد
في المراجعة الحديثة له عدة اختبارات عن المهارات الحسابية تقيس :

١ - المفردات الفنية والمعلومات الكمية الأساسية .

٢ - المهارات الحسابية .

٣ - حل مسائل تقرأ للمفحوص .

ويعطى للتلاميذ في السنوات من الثالثة الابتدائية حتى الشهادة الإعدادية

التشخيص

من أمثلة هذه الاختبارات ، اختبار تشخيص العمليات الرئيسية في الحساب

The Diagnostic Test for fundamental processes in Arithmetic

وهو فردي . والمسائل تحل شفويا . ولهذا يسهل فيه أن يتناول الطفل
المشاكل والعمليات المختلفة . وهناك قائمة لحساب الأخطاء تتضمن الأخطاء
الحسابية الشائعة . اختيرت مسائلها بحيث تتضمن صعوبات توقع بالمفحوص
في هذه الأخطاء . ويطبق هذا الاختبار من السنة الثانية الابتدائية حتى أواخر
المرحلة الإعدادية تقريبا . ولم يعد له توقيت ولا معايير ولا تعطى درجة
كلية . والمسائل ، وقائمة مراجعة الأخطاء يعتمدان في تصحيحهما على
التحليل الكيفي أكثر من اعتمادهما على التحليل الكمي .

اختبارات التشخيص في الحساب

ومن أشهر الاختبارات هنا ، اختبارات التشخيص في الحساب

Diagnostic Tests in Arithmetic التي أعدها مونرو Monroe ويتكون
من أربعة سلاسل من الاختبارات هي :

السلسلة الأولى : وتشمل عمليات أرقام صحيحة integers وتحوي
الاختبارات الآتية :

١ - اختبار جمع ٣ أعداد . وكل عدد يتكون من رقم واحد . ويشمل
٢٤ عملية . وزمنه نصف دقيقة .

٢ - اختبار طرح عدد من رقم واحد من أعداد من رقمين ويشمل
٢٤ عملية . وزمنه نصف دقيقة .

٣ - اختبار ضرب عدد من رقم واحد في عدد من أربعة أرقام ويشمل
٢٠ عملية . وزمنه دقيقة واحدة .

٤ - اختبار قسمة عدد من ٤ أرقام على عدد من رقم واحد ويشمل
٢٤ عملية . وزمنه دقيقة واحدة .

٥ - اختبار جمع ٥ أعداد ، كل عدد من ٤ أرقام ، رأسياً . ويشمل
١٤ عملية . وزمنه ٣ دقيقة .

٦ - اختبار قسمة عدد من ٤ أرقام على عدد من رقمين ويشمل
٢٤ عملية . وزمنه ٢ دقيقة .

السلسلة الثانية : وتشمل عمليات أرقام صحيحة . وتحوى الاختبارات
الآتية .

٧ - اختبار جمع ١٣ عدداً . وكل عدد من رقم واحد ويشمل ١٥ عملية
وزمنه ٢ دقيقة .

٨ - اختبار ضرب عدد من رقمين في عدد من ٤ أرقام ويشمل ١٢
عملية وزمنه ٣ دقيقة .

٩ - اختبار طرح عدد من ٣ أرقام من عدد من ٣ أو ٤ أرقام
ويشمل ٢٤ عملية وزمنه دقيقة واحدة .

١٠ - اختبار ضرب عدد من رقمين أو ثلاثة أرقام في عدد من ثلاثة
أرقام ويشمل ٢٤ عملية وزمنه ٢ دقيقة .

١١ - اختبار قسمة عدد من ٥ أرقام على عدد من رقمين . ويشمل ١٢ عملية . وزمنه ٤ دقائق .

السلسلة الثالثة : وتشمل عمليات على الكسور الاعتيادية وهى تحوى الاختبارات الآتية :

١٢ - اختبار جمع كسرين اعتياديين البسط رقم واحد والمقام رقم أو اثنين . ويحوى ١٥ عملية وزمنه دقيقة ونصف .

١٣ - اختبار طرح ، بنفس الطريقة السابقة يحوى ١٥ عملية وزمنه ٢ دقيقة .

١٤ - اختبار ضرب ، بنفس الطريقة السابقة يحوى ١٥ عملية وزمنه دقيقة واحدة .

١٥ - اختبار جمع كسرين اعتياديين ، بنفس الطريقة السابقة . ولكن فى الاختبارين الماضيين يسهل على التلميذ إيجاد مقام هو المضاعف البسيط للمقامات المختلفة ، بينما فى هذا الاختبار والاختبار المشترك التالى لا يسهل هذا . ويتكون هذا الاختبار من ١٥ عملية وزمنه ٢ دقيقة .

١٦ - اختبار قسمة كسر اعتيادى بسيط على آخر ويشمل ١٥ عملية وزمنه ٢ دقيقة .

السلسلة الرابعة : وتشمل عمليات ضرب وقسمة كسور عشرية . وهى الاختبارات الآتية :

١٧ - اختبار قسمة رقم صحيح وكسر عشرى على كسر عشرى وأمام عملية القسمة الإجابة الصحيحة دون وضع العلامة العشرية المطلوب وضع هذه العلامة حسب ما يراه التلميذ صحيحاً . ويشمل الاختبار ٢٤ عملية وزمنه نصف دقيقة .

١٨ - اختبار ضرب رقم وكسر عشرى فى كسر عشرى . ونتيجة عملية.

الضرب مرصودة وعلى التلميذ أن يضع العلامة العشرية في حاصل
الضرب . ويتكون من ٣٠ عملية وزمنه نصف دقيقة .

١٩ - اختبار قسمة ، بنفس طريقة اختبار ١٧ ، ولكن المقسوم كسراً
من ثلاثة أرقام عشرية . ويشمل ٢٤ عملية . وزمنه نصف دقيقة .

٢٠ - اختبار ضرب رقم وكسر عشري في رقم وكسر عشري ، بنفس
طريقة اختبار ١٨ ولكن اختبار ٢٠ أصعب ، ويحوى ٣٠
عملية . وزمنه نصف دقيقة .

٢١ - اختبار قسمة عدد صحيح أو عدد وكسر عشري على كسر عشري
والنتيجة مكتوبة وعلى المفحوص أن يضع العلامة العشرية في
خارج القسمة . ويشمل ٢٠ عملية . وزمنه نصف دقيقة .
وهي اختبارات سرعة غالباً ، كما يتضح من الزمن المحدد لإجرائها .
ويحسب عدد المسائل التي حاول المفحوص الإجابة عنها وعدد الصحيح
منها . وهكذا تحسب للتلميذ درجة .

اختبار ستانفورد الجديد في الحساب

قام الأستاذ إسماعيل القباني والمؤلف بإعداد اختبارات ستانفورد الجديدة
في الحساب ، وقد حدث ذلك سنة ١٩٣٦ . ومنذ ذلك الوقت طبع الاختبار
خمس طبعات وأعدت له معايير محلية . ولكن لم يتم إعداد معايير جديدة
حسب التغييرات التي طرأت أخيراً على تنظيم التعليم في بلادنا .

والاختبار يتكون من جزئين واحد للعمليات الحسابية ، والآخر للتفكير
الحسابي . وله صورتان متكافئتان هما الصورة م ، والصورة ن . وإليك
تعليمات الاختبار :

كلف التلاميذ باستبعاد كل ما قد يوجد أمامهم . وتأكد أن مع كل
منهم قلماً رصاصاً مبرياً برياً جيداً وبعد ذلك قل للتلاميذ « أنا النهارده .

حاططكم اختبار فى الحساب : حاوزع علىكم كراسات للاختبار ولكن ما حاءش منكم يفتحها إلا لما أقول لكم ، علشان تبتدوا كلكم مع بعض » .

ثم كلف تلميذاً أو تلميذين بتوزيع الكراسات تحت إشرافك وبعد ذلك . قل « اكتبوا البيانات اللى على الصفحة الأولى . . . قدام اسم التلميذ اكتب اسمك . (وقفة) وقدام عمره اكتب عمره بالسنة والشهر إذا كنت تعرفه . (وقفة) وقدام المدرسة اكتب اسم مدرستك . . . مدرسة كذا . . (وقفة) وقدام السنة الدراسية اكتب السنة والفصل اللى انت فيه . . سنة كذا . فصل كذا . . (وقفة) وقدام التاريخ اكتب تاريخ النهارده . . كذا . . » .

وبعد أن يفرغوا من ذلك اقرأ عليهم التعليمات الموضوعه فى أسفل الصفحة الأولى من الكراسة بتأن وبصوت واضح بينا هم يتبعونها معك فى كراساتهم ثم قل « دلوقت انتهوا كويس واتبعوا التعليمات بالضبط . . ساعة ما أقول ابتدى تفتحوا الكرايس وتبتدوا تجاوبوا . . وساعة ما أقول ضع القلم لازم تبطلوا الكتابة طوالى . . وبعد ما تبتدوا مش عاوز حد يسأل أى سؤال أو يتكلم مع اللى جنبه . . وإذا حد انكسر قلمه يرفع صباعه من غير كلام وأنا أديله قلم غيره . مفهوم . . . لاحظوا أن الاختبار فيه عدة صفحات وكلما تخلصوا صفحة اقلبوها وكمّلوا على الصفحة اللى بعدها لغاية ما أقول لكم ضعوا القلم (وقفة) . . . بالله دلوقتى . . . ابتدى . . . وهنا لاحظ الوقت فى ساعتك بالضبط وسجله . . . وفى نهاية الزمن المقرر قل « ضع القلم » وتحقق من أنهم جميعاً أطاعوا فوراً . . . ثم اجمع الكراسات .

أما الزمن المقرر فهو :

للفرقتين الأولى والثانية الابتدائيتين : ١٠ دقائق لاختبار العمليات
ومثلها لاختبار المسائل .

للفرق التي تلى ذلك^(١) : ٣٠ دقيقة لاختبار العمليات ، ٢٠ دقيقة لاختبار المسائل .

تعليمات لتصحيح الاختبار

١ - ضع أمام كل سؤال علامة ✓ إذا كان صحيحاً أو علامة X إذا كان خاطئاً ولا تضع علامة أمام السؤال الذي تركه التلميذ بدون إجابة .

ودرجة التلميذ هي عدد المسائل التي أجاب عليها إجابة صحيحة .

٢ - لا تعطى كسوراً في الدرجات وإذا لم يكتب التلميذ الجواب الصحيح بالضبط فتعتبر الإجابة خطأ .

٣ - إذا أعطى التلميذ جوابين لسؤال واحد تعتبر الإجابة خطأ . ولكن إذا شطب الجواب الأول أو صححه بشكل واضح فيقبل الجواب الثاني .

٤ - إذا وضع التلميذ الجواب في غير محله فأقبله إذا كان واضحاً أنه يقصد أن هذا هو جواب المسألة المطلوبة .

٥ - لا يؤخذ التلميذ على عدم كتابة التمييز في الجواب .

٦ - إذا كتب التلميذ الجواب على صورة كسر عشرين بدله كسر اعتيادي يعتبر صحيحاً .

٧ - إذا كتب التلميذ الجواب على صورة كسر اعتيادي فيجب أن يكون مختصراً إلى أبسط صورة وإلا يحسب خطأ .

أولاً : اختبار التفكير الحسابي الصورة م .

يتكون من ٤٠ فقرة مثل :

١ - محمد عنده ٦ أرانب بيضاء و ٥ أرانب سوداء . فكم أرنباً عنده ؟

(الجواب :)

(١) توازي الفرق الأولى والثانية (نظام قديم) الفرق الرابعة والخامسة (نظام حديث) .

٤٠ - اشترى رجل بلحاً بسعر كل ٦ أرطال ثمانية قروش . وباعه
بسعر كل ٣ أرطال ثمانية قروش . وكان ربحه من ذلك ١٦ ر من الجنيهات .
فكم رطلاً من البلح باعها ؟ (الجواب —————)

وتناظر هاتين الفقرتين الفقرتان التاليتان في الصورة : ن

١ - إذا وضعت ٥ طيور و ٤ طيوراً معاً ، فكم طيراً تصير ؟

(الجواب —————)

٤٠ - منزل ثمنه ٥٠٠٠ جنيه ، ويدفع صاحبه عليه ضريبة سنوية
قدرها ٦٠ جنيهاً ، ويصرف عليه ٢٠٠ جنيه سنوياً للتصليحات . فبكم
جنيهاً يجب أن يؤجره شهرياً حتى يحصل على ربح صاف قدره ٨٪ من
ثمن البيت ؟ (الجواب —————)

ثانياً : اختبار العمليات الحسابية الصورة م :

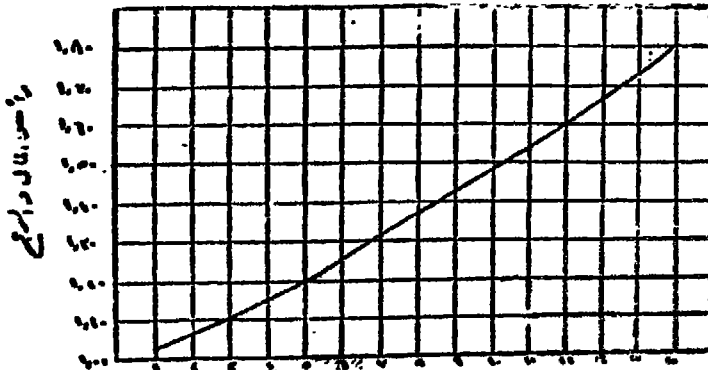
وتشمل ٦٠ قرة ومثلها الصورة ن .

ولذلك مثلاً من كل منهما .

الفقرة ٤٩ الصورة (م)

الرسم البياني الموجود أدناه يبين الجملة في كل ستة لمبلغ جنيه واحد
يربح ربحاً مركباً بسعر ٤٪ في السنة

(شكل ٢١)



فما جملة جنيه واحد بعد ١٢ سنة ؟ (الجواب)

الفقرة (٣٨) و(٣٩) السورة (ن)

أبعاد بعض المدن الشهيرة عن القاهرة بالكيلو متر

(شکل ۲۲)

[illegible]

من هذا الرسم البياني أوجد بالتقريب بعد المدن الآتية عن القاهرة
بالكيلومتر :

..... دمایط اخترطوم

العلوم

نكتفي في هذا القسم بذكر بعض الأمثلة من أهم الاختبارات الشائعة أو التي كانت شائعة . وأغلب أمثلة الكيمياء والحيوان والنبات والجغرافية مأخوذة عن الأستاذ إسماعيل القباني في كتابه مشكلة الامتحانات.

اختبار رتش وبوينو في العلوم العامة

وله ثلاثة صور :

« الصورة أ » :

وفہا قسمین نأخذ أمثلتنا من :

القسم الأول :

التعليقات : « ضع خطأً تحت كلمة واحدة أو عبارة واحدة تعتبرها أحسن جواب للسؤال ، كما في المثال الآتي :

« من أمثلة السوائل : الحديد - الزجاج - الهواء - الماء - الرصاص -
الخاصين - الأكسجين » .

وبلى ذلك ٥٠ سؤالاً من هذا النوع ، يخصص للإجابة عنها ١٥ دقيقة ،
نكتفى بإيراد خمسة منها على سبيل التمثيل ، وهى :

٢ - الأقرب إلى الاحتمال أن توجد البراكين فى : الصحارى - الوديان
- السهول الساحلية - الجبال - دالات الأنهار - المناطق الداخلية -
الجزر .

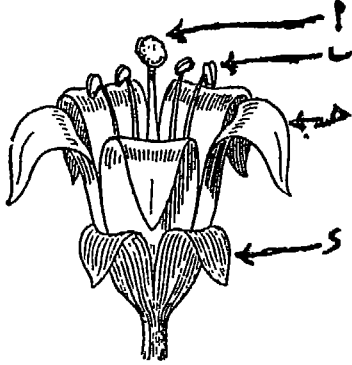
٧ - من أمثلة التأكسد : صدأ الحديد - تحليل الماء بالكهرباء - انصهار
الجليد - ذوبان الملح - تأثير حامض فى الرخام - تسخين كلورات
البوتاسيوم - غليان الماء .

١٠ - عندما يحتوى السائل على أكبر كمية يمكنه إذابتها من مادة
ما تسمى هذه الحالة : الأنشطار الغشائى - المسامية - الانصهار - التعادل
- الاختزال - الإفراز - التشبع .

القسم الثانى :

التعليمات : يوجد فى هذه الصفحة والصفحات التالية ٢٠ شكلاً بجوار
كل منها عبارات متعلقة به ، ولكن قد تركت فى كل من هذه العبارات
مسافة خالية ، عليك أن تملأها بحيث تجعل العبارة صحيحة ، وافية ، محدودة
المعنى . فى معظم الحالات لا تحتاج إلا إلى كلمة واحدة لملء المسافة .
وستعطى لك ٢٥ دقيقة للإجابة عن هذا القسم من الاختبار .

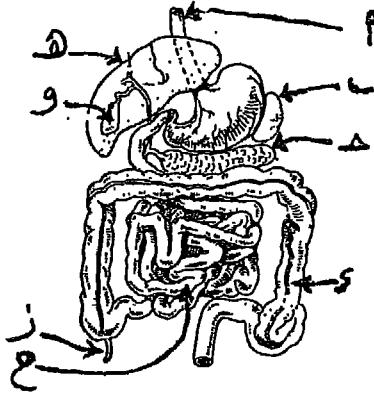
في هذا الشكل الذي يبين تركيب الزهرة النموذجية :



- (١) قد رمز للبلمات (التويج) بالحرف
 (٢) ورمز لأعضاء التذكير بالحرف
 (٣) ورمز للسبلات (الكأس) بالحرف
 (٤) ورمز لعضو التأنيث بالحرف

(شكل ٢٣)

في هذا الشكل الذي يبين القناة الهضمية :



- (١) قد رمز للأمعاء الدقيقة بالحرف
 (٢) ورمز للمريء بالحرف
 (٣) ورمز للكبد بالحرف
 (٤) ورمز للمعدة بالحرف
 (٥) ورمز البنكرياس بالحرف

(شكل ٢٤)

اختبار رتثش وكرسمان فى علم الحىاة

وللاختبار صورتان نأخذ أمثلتنا من الصورة ب . وكل صورة تتكون من ٥ أقسام :

القسم الأول : ٤٠ سؤالاً زمنه ١٢ دقيقة . والإجابة بالاختيار من سبعة إجابات عن كل سؤال .

القسم الثانى : ١٥ عبارة زمنها ٨ دقائق . والإجابة بتكميل كل عبارة .

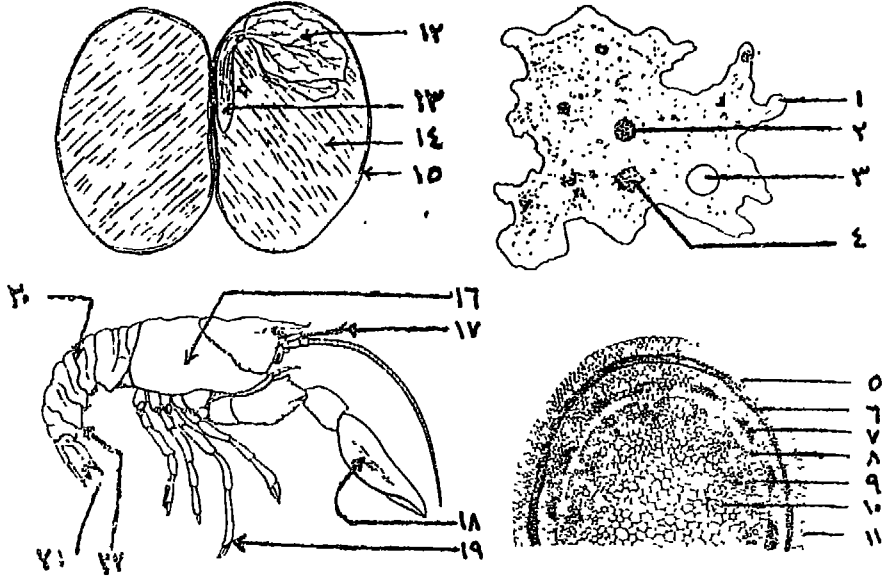
القسم الثالث : التعليمات : افحص الأشكال الأربعة المرسومة على الصفحة المقابلة لهذا جيداً . وقد وضعت على بعض الأجزاء فيها أرقام لخاصة ، وفيما يلى أسماء بعض هذه الأجزاء وعليك أن تضع أمام كل اسم الرقم الموضوع عليه فى الرسم الآتى (صفحة ٥٨٢) :

القسم الرابع : سؤال واحد عن قانون مندل فى الوراثة والإجابة بتكميل جدول .

القسم الخامس : التعليمات : اكتب الكلمات المتروكة فى الفقرات التالية ، بحيث تختار لكل مسافة بيضاء الكلمة التى تعطى أصح معنى ، كما فى المثال الآتى :

الشرب من إناء واحد يسبب انتشار . . . الأمراض . . . وبلى ذلك خمس فقرات يخصص لإكمالها عشر دقائق ، ونكتفى بإيراد الفقرة الآتية منها على سبيل التمثيل :

١ - للكشف عن وجود البروتين فى مادة ما تسحق هذه المادة جيداً ، ثم تغلى فى ماء قد أضيفت إليه قطرات من : . . . فإذا ظهر لون . . . دل ذلك على وجود البروتين . وإذا أضفنا كمية كبيرة من . . . يتحول هذا اللون إلى . . .



(شكل ٢٥)

مثال البطن ٢٠ :

..... الأرجل القابضة	 النواة
..... النفير	 الفلقة
..... القشرة الداخلية	 الزيليم
..... الفجوات القابضة	 الفجوات الغذائية
..... شعيرات الاستشعار	 القشرة الخارجية
..... الزعانف الذنبية	 اللحاء
..... الألياف	 الصدر الرأسى
	 الكامبيوم

اختبار جلن وولتون فى الكيمياء

التعليمات ثلاث من الكلمات الواردة فى كل سؤال من الأسئلة الآتية بينها علاقة متينة من حيث المعنى أو من حيث الاستعمال أو كليهما . المطلوب منك أن تكتب بين القوسين الموضوعين فى نهاية كل سؤال رقى الكلمتين أو الجملتين اللتين لا توجد بينهما وبين تلك الكلمات الثلاث أى علاقة .

وبلى ذلك ٣٠ سؤالاً نورد منها الأمثلة الآتية :

- (١) ١- يد ا ص ٢- يد ا بو ٣- (يد ا) كا
 ٤- ماء النار ٥- ا ك (.....)
 (٢٦) ١- أزوتات ٢- يد ز ٣- ا ك
 ٤- ملح شيلى ٥- الحرير الصخرى (.....)
 (١٣) ١- التحليل الكهربائى ٢- الكبريت ٣- الرخام
 ٤- السكر ٥- حامض الكبريتيك (.....)
 (٢٧) ١- آحادى التكافؤ ٢- أكسجين ٣- صديوم
 ٤- كلسيوم ٥- كلور (.....)
 (٣٠) ١- الاختزال ٢- جرام ٣- أكسيدات
 ٤- محلول ٥- فلزات (.....)

اختبار ويديفليد ووالتر فى الجغرافية

التعليمات : فى العمود الأيمن عبارات تدل على أشياء يعملها الناس فى مختلف الأقاليم ، وفى العمود الأيسر عبارات تدل على بعض أسباب ذلك . فمثلا العبارة رقم ٣ من العمود الأيسر تدل على سبب اشتغال أهالى

الزرويج بصيد السمك . فعليك أن تكتب رقم (٣) بين القوسين
الموضوعين بعد عبارة « صيد السمك في الزرويج » الواردة في النموذج .
اعمل مثل هذا مع باقي العبارات الآتية : -

الأسباب

النتائج

- نموذج : صيد السمك في الزرويج . () (١) جبال عالية مكسوة بالثلوج .
! (١) حياة التجول عند الأكسيمو . () (٢) أناث يعتمدون في غذائهم على
الحيوانات البرية :
(٢) استعمال الجمال كواسطة () (٣) أرض مجذبة صخرية مغطاة
للاتنقال .
(٣) بيوت من الخشب وقشور () (٤) جو بارد رطب وكثير من
الأشجار، لها سقوف مائلة من الأراضي المنخفضة الرطبة .
أوراق الشجر .
(٤) لبس الناس أحذية خشبية في () (٥) جوحار جاف وكثير من الرمال
أثناء تأدية أعمالهم .
(٦) جوحار ممطر وكثير من أشجار
« الرافيا » .
(٧) سهول مسطحة وأرض خصبة
ومطر معتدل في الصيف .

مجالات أخرى

أصبحت الاختبارات التحصيلية تطبق في أغلب الميادين التعليمية . ففي
المدرسة الابتدائية والإعدادية نجد الاختبارات العامة أما في الثانوية فنجد
اختبارات أكثر تخصصاً بينما في الجامعة تكون اختبارات متقدمة متخصصة

شاملة ويكون بعضها عاماً . وهى هنا أو هناك قد تدخل فى عداد فئات أخرى من الاختبارات سيأتى ذكرها فى القسمين التاليين .

وتعتبر سلسلة الاختبارات التحصيلية المتأذرة Coordinated Series of Achievement Test من السلاسل التى تصلح للاستخدام فى مختلف المناهج الدراسية . ومن أفضلها ثلاث اختبارات تطبق من أجل الأغراض التالية :

١ - الاختيار وقبول طلبة الجامعة الجدد .

٢ - تقييم المناهج التعليمية .

٣ - تقييم الاختبارات المنشورة فى مصادر أخرى والسمة الأساسية المميزة لهذه السلسلة هى استخدامها لمعايير موحدة مقارنة فى كل الاختبارات . وهذا بسهل المقارنة المباشرة بين الدرجات المختلفة المعطاة من مواد دراسية مختلفة إلا أن هذه السلسلة لم تقن كلها على نفس التلاميذ ، ذلك أن هناك تخصص فى مثل هذه المستويات التعليمية . فبينما يدرس طالب اللاتينية عامين ، يدرس آخر الأسبانية عامين أيضاً ، ومن هنا لا يمكن إعطاء الطالبين معا اختباراً فى اللاتينية أو فى الأسبانية .

وكانت الدرجات الخام فى كل اختبارات هذه السلاسل الثلاث تحول إلى درجات معيارية متوسطها ٥٠٠ وانحرافها المعيارى ١٠٠ : والجامعة المرجعية التى أخذ المتوسط والانحراف المعيارى من درجات أفرادها تكونت من ١٠٦٥١ تلميذاً طبق عليهم اختبار الاستعداد المدرسى Scholastic Aptitude Test وكانت درجات التلاميذ ، بعد إجراء هذا الاختبار ، التى حصلوا عليها فى السنوات التالية تقارن بدرجات الجامعة المرجعية . ثم يحسب الفرق بين متوسط التلاميذ بالجدد ومتوسط الجامعة المرجعية ، ومن دلالة هذا الفرق يقيم المنهج التعليمى الذى شاركوا فيه .

وقد أصبحت هذه السلاسل تقليدا منتظما ، إذ تطبق على الطلبة الجدد المتقدمين لكليات معينة . بل وأصبح بعضها صالحاً للاستخدام في أغراض التوزيع والتصنيف العام . ويجرى تعديلها بصفة دورية حتى تناسب المناهج الدراسية بالمدارس الثانوية ، وحتى تتفق مع التغيرات التي تطرأ على تعليم المواد الدراسية الآتية :

١ - اللغة الرسمية في البلاد .

٢ - اللغات الأجنبية .

٣ - العلوم الطبيعية .

٤ - الرياضيات .

٥ - التاريخ .

٦ - المواد الاجتماعية الأخرى .

وقد اختيرت الجماعة المرجعية التي حسب منها المتوسط والانحراف المعيارى بحيث كان أفرادها تلاميذا متوسطى القدرة والتحصيل . ذلك بالاعتصار على من حصلوا على درجات متوسطة في اختبار ذكاء صادق ، باعتبار هذا محكا لتوسط قدراتهم . كما كان يقتصر على الطلبة من أعمار يغلب أن يكون عليها تلاميذ الفرق الدراسية المتضمنة في الجماعة ، باعتبار هذا محكا لتوسط تحصيلهم .

كما أخذ محك آخر لتوسط تحصيل جماعات التقنين ، وهو حصول أفرادها على درجات متوسطة في بطارية تحصيل عام أخرى صادقة . وكذلك اتبع تقليد جديد وهو وضع الفقرات على أساس المدة التي قضها التلميذ المتوسط في تعلم المادة المقاسة . فثلا كانت الجماعة المرجعية في اختبارات اللغة الإنجليزية قد درست اللغة أربع سنوات بالمدارس الثانوية . بينما كانت الجماعة المرجعية بالنسبة لاختبارات اللغات الأجنبية قد درست هذه اللغات مدة سنتين .

مقاييس الإنتاج

الاتجاه إلى استخدام مقاييس الإنتاج Product Scales يعكس لنا تحولاً في طرق تقييم التلميذ . تحول إلى الإجراء ، وقرب إلى واقع الحياة المدرسية المعتادة ، وتأهيل للطلبة حتى يستطيعوا مواجهة الحياة العملية بعد الانتهاء من الدراسة . وأغلب موادها تعتمد على تقييم الخط الكتابي والإنشاء والقصص والوصف وفك وتركيب آلات واستخدام أجهزة وأدوات .

مقياس فان واجن للإنشاء الإنجليزية

يعتبر مقياس فان واجن للإنشاء والإنجليزية Van Wagnen English Composition Scale مثلاً جيداً لهذا النوع من المقاييس . وفيه اختبارات منفصلة للوصف والقصص . وتعطى للطلبة درجات عن مستوى المادة المعروضة ، وبنائها ، والأسلوب الفني في الكتابة . ثم تؤخذ متوسطات هذه الدرجات ، وتكون درجة التلميذ الكلية في المقياس هي متوسط درجاته الثلاث .

وقد أعدّ لويس Lewis مقياساً تحسب درجة الطالب فيه بناءً على درجاته في المجالات الخمس التالية :

- ١ - ترتيب الكلمات .
- ٢ - ترتيب ورود ومواضع العبارات .
- ٣ - الجانب الاجتماعي في القصة .
- ٤ - وصف الظواهر الاجتماعية .
- ٥ - بساطة العرض .

والعيب الأساسى فى أغلب مقاييس الإنتاج أنها لا تراجع حتى تتفق مع
تغير محتوى المناهج الدراسية والأحداث القومية والعالمية الهامة . ولكنها وسيلة
جيدة يمكن الاعتماد عليها ضمن اختبارات التحصيل المقننة ، باعتبارها
جزءاً موضوعياً لتقييم مواد كيفية .

أغراض تعليمية أوسع

ونقصد بها أهداف التربية البعيدة مثل : إنماء التفكير الناقد - وتكوين
عادات واتجاهات مرغوبة فى تذوق الفن والأدب - والفهم الذكى للأحداث
المعاصرة - وتكامل ما حصله المتعلم من مختلف المناهج الدراسية التى
اشترك فيها .

الاختبار التعاونى للشئون المعاصرة

الاختبار التعاونى للشئون المعاصرة Cooperative Contemporary
affairs Test يمثل هذا النوع من الاختبارات أصدق تمثيل . وهو يطبق
فى ساعة على طلبة الجامعة لقياس الميل والمعلومات فى ميادين ثلاث هى :

١ - الشئون العامة ..

٢ - العلوم والطب .

٣ - الأدب والفن .

ويعادل الاختبار كل سنة حتى يكون عصبياً .

الاختبار التعاونى للثقافة العامة

الاختبار التعاونى للثقافة العامة Cooperative General Culture Test

يعطى لطلبة الجامعة لقياس المعلومات العامة فى أهم خمسة ميادين لمناهج
الدراسة الجامعية وهى :

١ - التاريخ والدراسات الاجتماعية .

٢ - الأدب .

٣ - العلم .

٤ - الفن .

٥ - الرياضيات .

ولكل ميدان جزء منه من الاختبار يتناسب ٣٠ دقيقة . وقد اختبرت فقرات الاختبار بحيث يستطيع الإجابة عن أكثر من نصفها من نسيج في دراسة المواد المقررة في الجامعة . وكل اختبارات السلسلة التعاونية تستخدم المعيار المئتي . والعينة ، والإجراء جمعي . وتعطى درجات جزئية وأخرى كلية .

وقد قامت هذه السلسلة بنشر اختبارات تناسب تلامذة المدارس الثانوية والجامعة . وهي تقيس الميول والاتجاهات والسمات الاجتماعية . وهي تتضمن إلى جوار هذا اختبارات صممت لقياس المجالات التالية :

١ - الاستدلال المنطقي .

٢ - تفسير البيانات .

٣ - تطبيق المعلومات الأساسية في العلوم التالية :

(أ) الطبيعة .

(ب) الأحياء .

(ج) الدراسات الاجتماعية .

(د) الأدب .

فمثلاً في تفسير البيانات تقدم مجموعة منها في شكل جداول أو خرائط لـ أ و رسوم متبوعة بسلاسل من الجمل يقوم المفحوص بوضع علامة صحيح أو خطأ عليها بناءً على مدى كفاية البيانات في تأكيد ما أتت به هذه الجمل .

أما اختبارات تطبيق المعلومات الأساسية فهي تقتضي استخدام الأسس العلمية في وصف مواقف الحياة اليومية . وموادها غير موجودة في الكتب المدرسية : وكل مجموعة من الفقرات متبوعة بنقطة تصنف حسب رأى الطالب إلى : صحيحة ، تختمل الصحة والخطأ ، خاطئة . وعلى المفحوص أن يبرر رأيه باختيار أسباب منطقية تدعمه من القطع المكونه من اجل تفسر موقف الحياة اليومية المعروض له .

التفكير الناقد في الدراسات الاجتماعية

اختبار التفكير الناقد في الدراسات الاجتماعية Critical thinking in the Social Studies ، قلمه ريتستون Wrightsone . ويستخدم في المدارس الثانوية والمعاهد . وقد انتقد هذا الاختبار لأن عنوانه مضلل فهو لا يقيس التفكير الناقد فعلا ، بل ولم يقدم بيانات عن صدقه وثباته . والقيمة الأساسية لهذا الاختبار في أنه يقترح طرقا يمكن استخدامها في عمل اختبارات مماثلة . وهو يتكون من الأجزاء الثلاث التالية :

١ - الحصول على الحقائق .

٢ - استخلاص استنتاجات من هذه الحقائق .

٣ - تطبيق الحقائق العامة .

والجزء الأول هو اختبار لمهارة مذاكرة الدراسات الاجتماعية باستخدام الخرائط ، والرسوم البيانية ، والقوائم ونقلها من مصادر الحقائق . والجزء الثاني ، يشبه اختبارات فهم القراءة ، ويتطلب من المفحوص أن يختار ما يمكن أن يستفيد به من كل قطعة على حدة .

وفي الجزء الثالث يعاد تقديم تلك القطع مختلطة بقطع أخرى ويطلب من المفحوص أن يسترجع من كل قطعة من الأولى جملتين ويربطهما معاً

اختبار واطسن - جلاسز لتقدير التفكير الناقد

لم يعن اختبار واطسن - جلاسز بتقدير التفكير الناقد - Watson
Glasser Critical thinking appraisal test في ميدان بذاته من ميادين
المعرفة . وتتضمن آخر صورة معدلة لهذا الاختبار الأجزاء الخمسة الآتية :

١ - الاستدلال .

٢ - تحقيق الفروض .

٣ - الاستقراء .

٤ - التفسير .

٥ - التقييم .

والجزء الأول تعرض فيه قطع يتبعها استدلال منها ويصنف المفحوص
هذا الاستدلال إلى صحيح أو خطأ . والجزء الثاني به فروض وجمل والمطلوب
معرفة ما إذا كانت هذه الجمل تؤيد تلك الفروض أم لا ، وما درجة تأكيدها .
والجزء الثالث يتطلب القيام بعمليات الاستقراء العقلي . أما الجزء
الرابع فيطلب فيه من المفحوص أن يدلل على ما إذا كانت الاستنتاجات
المدونة مشتقة من قطعة معينة أم لا . والجزء الخامس والأخير يطلب فيه
من المفحوص أن يقرر ما إذا كان بناء القطعة قويا أم ضعيفا .

والمشكلة في تقدير إجابات المفحوصين هي هل يعد الخطأ خطأ فعلا
وكليا ؟ وهل يعد الصحيح صحيحا فعلا وكليا ؟ ولهذا قام واضعو الاختبار
بدراسة فقراته والاقتصار على أصدقها وأكثرها صلاحية . وكانت طريقة
الدراسة بعرض الفقرات على مجموعة من المحكمين عددهم ٣٥ قاض
[مختارين ممن نالوا تدريبا عالياً وتفوقوا في المنطق ، وفقه اللغة ، والكيمياء ،
والأحياء ، والطبيعة ، وعلم النفس ، والتربية ، والهندسة التطبيقية . فإذا

نالت الفقرة أغلبية أبقى عليها . وقد تم اختبار الفقرات على أساس آخر أيضاً هو الاتساق الداخلى . وفى بحث الصدق استخدمت طريقة الجماعات المتقابلة Contrasted groups وهى جماعة من الطلبة الجامعيين فى مقابل جماعة من العمال والموظفين .

والمعيار المستخدم هو الدرجات المئينية . ومعاملات الثبات منخفضة بالنسبة لبعض الدرجات الفرعية . أما ثبات الاختبار ككل فيتراوح بين ٠.٨٠ إلى ٠.٩٠ .

الاختبارات للتحصيلية المهنية

يجب ألا ننس أن الكثير من الاختبارات التحصيلية كان يستخدم فى الاختيار ، والتصنيف فى الصناعة ، والحكومة ، والقوات المسلحة . واختبارات التحصيل المهني Vocational Achievement Tests فى ميدان الصناعة تسمى اختبارات الحرف trade tests . وكثير من الاختبارات التحصيلية المهنية وضعت لتؤدى أغراض خاصة ، أى أنها « تفصيل tailor made » . وتحوى مواداً تختلف فى طبيعتها ، فقد يكون المحتوى كله لفظياً ، وقد يكون رسومات بيانية أو صوراً ، أو أسئلة تقدم شفويا أو كتابيا .

وهى تستخدم عينات تقنينية من العمال ، ويطلب من المفحوص القيام بعمل يشبه العمل المعتاد الذى يؤديه . وهذا يعنى أن الفقرات عينات من العمل Work Samples . وقرب السلوك المقاس من العمل اليومي ، وتمثيل عينة الفقرات لوحداث السلوك المراد قياسه ، أمران مميّزان لهذا النوع من الاختبارات .

ويسجل أداء المفحوص فى اختبارات عينة العمل بطريقة موضوعية . فمثلاً ، فى اختبار الآلة الكاتبة ، يحسب مجموع الأخطاء فى النسخة المكتوبة ،

والزمن ، والتنسيق . ولكن بعض هذه الاختبارات يلجأ ، فضلاً عن التصحيح الكمي ، إلى التفسير الكيفي الذي يقوم به الخبراء .

ومن أفضل الاختبارات التحصيلية المهنية تلك الطبقة في المهن المعروفة وخاصة الآلة الكاتبة ، والعنونة ، والسكرتارية . ويتضمن بعض هذه الاختبارات استخدام اللغة الرسمية ، واللغات الأجنبية ، والمعلومات العامة عن المهن ، ومقاييس للمهارة في المهنة .

اختبار بلاكستون للكفاية في السكرتارية

اختبار بلاكستون للكفاية في الأعمال المكتبية Blackstone Stenographic proficiency Test . يطلب من المفحوص تسجيل ملاحظات مختزلة هي تلخيص واختزال وترجمة لخطابين . وهناك جزء لقياس الأخطاء في القواعد ، والآخر في الهجاية ، وثالث في الضبط ، ورابع في الربط ، وخامس في التدريب على الأعمال المكتبية وتنظيم المكاتب .

اختبار سيشور - بينيت للكفاية في السكرتارية

يتطلب اختبار سيشور - بينيت للكفاية في السكرتارية Seashore Bennett Stenographic Proficiency Test . أن يكون المفحوص ملاحظات مختزلة عن ٥ خطابات تتدرج في الطول والتعقيد . ويكون إملاء الخطابات بسرعات تتدرج في الزيادة أيضاً . ويطلب من المفحوص أن يسجل ملاحظاته هذه بطريقة تسهل تدوينها بالآلة الكاتبة . والمعايير غير واضحة ، ولكن الصدق مرتفع باتباع طريقة الجماعات المتقابلة ، وبإيجاد معاملات الارتباط بين الدرجات على الاختبار وتقديرات المشرفين على العمل لنفس الأفراد المختبرين ؛ والثبات مرتفع بالصور المتكافئة .

اختبارات الشفوية

اختبارات الحرف الشفوية Oral trade Tests هي اختبارات تعيين في القيام بالمقابلات في مكاتب العمل . وقد استخدمت هذه الاختبارات على نطاق واسع في الحربين العالميتين : وقد وصف تشامان Chapman أول اختبار حرفة شفوية في الحرب العالمية الأولى ، في كتاب بأكمله . وقام طومبسون Thompson بمراجعة فقرات هذا الاختبار ووضعها في قالب جديد .

كما قام قسم تحليل العمل بمكتب العمل بالولايات المتحدة الأمريكية ، حديثاً ، بتأليف اختبار مماثل مقنن يصلح لقياس الكفاية في ١٢٦ عملاً . ولكل عمل ١٥ سؤالاً . ووضعت صور مكافئة بالنسبة لبعض هذه الأعمال وعددها ٤٩ عملاً . كما حسبت معاملات الثبات بطريقة الصور المتكافئة وتراوح بين ٠,٧٩ إلى ٠,٩٣ .

وجمعت البيانات عن صدق هذه الاختبارات بملاحظة الأعمال والاختلاط بالعمال والملاحظين والفنيين في كل عمل . والأسئلة موضوعة بلغة سهلة ومبسطة بحيث تناسب مستوى العمال . وإجابة بعض الأسئلة تقتضي السرعة ، ويقتضي البعض الآخر إصدار أحكام نافذة . وقد أدى الاختلاط بالعمال إلى مراجعة فقرات الاختبارات وتعديل بعضها وإضافة البعض الآخر . وروعت الفروق بين المناطق في مواد العمل وطرق أدائه .

وقد قسم العمال إلى هذه الأقسام :

(أ) الأسطوات : الذين قال رؤسائهم إنهم أعلى مستوى وهم أصحاب خبرة ٤ سنوات على الأقل في عمل بذاته .

(ب) المبتدئون : كالصبيان المساعدين .

(ج) من يعملون في أعمال متشابهة ويقرب مستواهم من الأسطوات .

وقد قننت الأسئلة كل سؤال على عمل من ٥٠ إلى ١٠٠ عامل من
الجامعة ١ ومن ٢٥ إلى ٥٠ من الجامعات ب ، ج .
وقد اختبرت الأسئلة على أساس القيمة التقريبية .
ونظراً للتغير الدائم في العمليات والمواد المهنية فمن الواجب دائماً أن
تراجع اختبارات الحرف ولكنها على أى حال لا تقوم مقام الأداء في
اختبارات عينة العمل . ولكنها تستخدم وحسب كقياس سريع لمعرفة
الخبرة المهنية للباحث عن عمل .

الفصل الثامن عشر

بحوث واختبارات تربوية محلية

- مقدمة .

- بحث الأستاذ للقباني .

- بحوث وزارة التربية والتعليم .

- بحوث المؤلف .

- دليل الاختبارات .

- القراءة .

- الحساب .

- نحو المجلد الثاني .

الفصل الثامن عشر

بحوث واختبارات تربوية محلية

مقدمة :

قد يعتقد البعض أننا لن نلحق ركب العلوم الموضوعية عامة والقياس خاصة ، ولكننا نعتقد أن هذا البعض عندما ينتهى من قراءة هذا الفصل سيكون أكثر تفاؤلاً بمستقبل البحوث والدراسات في بلدنا . ونحن لا نهدف إلى تدعيم هذا التشاؤم أو ذاك التفاؤل ، بل وإننا نحاربه ، إذ أن اعتقاداً متطرفاً من أحد النوعين لن يقدم لنا جديداً فضلاً عن تثبيطه هم البجاعة وعدم إعطاء الحق لأصحابه .

وإن كنا نثق بأنفسنا ونفخر بجهود روادنا ، إلا أننا لم نقصد باختيار وتقديم بعضها تقيماً أو تقسديراً ، بل تمثيلاً يعطى فكرة للقارئ ويستثير الكفاءات لتنهج نهجه . على أن البحوث التربوية التي استخدمت الاختبارات أداة لها طرقت مجالات لها في بلادنا تاريخ طويل ، وإن كانت قد بدأت متعثرة بعض الشيء إلا أن هذا هو شأن الجهود العلمية في ميدان جديد نسبياً كما أن إمكانيات هذه البحوث ، مهما كانت الهيات القائمة بها : فقيرة إذا ما قورنت بمثيلاتها في الخارج . إلا أنها تخطو خطوات واسعة وتتقدم تقدماً سريعاً .

ونأمل أن يحين الوقت الذى فيه تكون ثروتنا من البحوث والدراسات في المجال الذى يمر به كل فرد ويعيش فيه الكثير من الناس وهو ميدان التربية .

بحث الأستاذ إسماعيل القباني

قدم الأستاذ إسماعيل القباني بعض أسئلة امتحانات الشهادات العامة حتى سنة ١٩٣٨ . ثم عقب عليها مقدماً بحثه وملخصاً لإياه في كتاب مشكلة الامتحانات في مصر التي أصدرته رابطة التربية الحديثة ١٩٣٩ قائلا :

إن أمثال هذه الأسئلة يجدها الإنسان في كل ورقة من أوراق أسئلة المواد التي تعتبر « مواد معلومات » . وظاهر أن الإجابة عنها لا تتطلب من التلميذ أكثر من أن يحفظ بضع صفحات من الكتاب المقرر عليه ، ويستبقها في ذهنه إلى أن يحين وقت الامتحان ، فيفرغها على ورقة الإجابة كما هي من غير تصرف . وهذا يجعله أقرب إلى قياس الملكة اللفظية وذاكرة الألفاظ .

ولا يقتصر الأمر على هذا ، بل هناك ما هو أهم وأدهى : فإن مثل هذا الامتحان لا يمكن أن يكون مقياساً صحيحاً للتحصيل والقدرات العقلية المكتسبة من التعليم ، مع أن هذا هو الغرض المباشر من إجراءاته ؛ فما كان المقصود بالتحصيل العلمي يوماً من الأيام شحن الذاكرة بمحتائق غير مهضومة وغير مترابطة ، تدخل الذهن من ناحية ثم تخرج من الناحية الأخرى ولم تغد فيه قوة نافعة .

إن أكبر ما يؤخذ على التعليم في مدارسنا هو أنه لا يثقف عقل الطالب ولا ينمي فكره ، ولا يربي فيه قوة الابتكار وسلامة الحكم ، وروح النقد ، والاستقلال في الرأي ؛ ولا يبعث في نفسه حب العلم ، وتقديس الحق ، والشعور بالواجب ، وبالإجمال لا يهيئ الفرص لتكوين الشخصية ، وبناء الخلق المتين ، الذي يؤهل الطالب للكفاح في الحياة ، ولخدمة المجتمع . ومن أهم أسباب ذلك أننا أخطأنا غاية التربية ، فصرنا نعني بتكديس المعلومات ، ولا نبحث على أثرها في النفس . ولا شك في أن

الامتحانات بصورتها الحالية عامل قوى ، بل هى أقوى عامل ، يوجه التعليم فى ذلك الاتجاه .

إن الصورة التى يكونها المرء فى ذهنه عن الحياة العقلية للطالب المصرى عند ما يفحص أوراق الأسئلة يتمثل فيها العقل فى شكل صندوق مستقبل تصب فيه المعلومات كما يصب الماء فى مستودعه ، وتبقى فيه إلى أن يحين وقت تفريغها منه كما يسحب الماء من المستودع . ولكن الماء يبقى فى المستودع مستقلا عنه ، لا يؤثر فى تكوينه ، فهل هذه هى العلاقة التى نريدها بين المعلومات والعقل ؟ أم نريدها علاقة الغذاء بالجسم ينميه ويقويه ؟ والغذاء لا يبقى فى الجسم على حالته الأصلية ، بل يتفاعل مع إفرازاته ويتحول إلى عصارة مهضومة ، تختلف فى تركيبها عن الطعام الأصل ، ويمتصها الجسم فيكسب منها القوة ، أما إذا بقي الطعام فى المعدة على حالته الأولى فلا تنتج عنه إلا تخمة . فكذلك المعلومات إذا بقيت فى الذهن عائمة منفصلة ، مستعدة للخروج منه على نفس الصورة التى دخلت بها ، فإنه لا تنتج عنها إلا تخمة عقلية . فالمعلومات التى يحصلها الإنسان لا تصبح جزءاً من نسيجه ، فينمو بما اكتسبه منها ، ويصير بعد امتصاصها أقدر على التصرف فى المشاكل التى تعرض له مما كان قبلاً ، وبذلك تؤثر فى اتجاهات الإنسان النفسية ، وتصبغ نظراته العامة إلى أمور الحياة من غير أن يشعر . وفى هذه الحالة لا يصبح من المتيسر اختبار قيمة هذه المعلومات باستعادتها من العقل كما كانت ، بل بقياس القوة التى اكتسبها العقل منها .

وقد أحسن التعبير عن هذا المعنى الفيلسوف الإنجليزى الرياضى الأستاذ هوايتهد (Whitehead) عند ما حذر المربين فى كتاب ممتع له عن التربية مما سماه « الأفكار الخاملة » (inert ideas) وقد عرفها بأنها « الأفكار التى يقتصر العقل على استخدامها ، من غير أن يستخدمها ، أو يختبر صحتها ، أو يحولها إلى صور جديدة » . ثم فسر ما يقصده باستخدام الأفكار بقوله :

« أقصد باستخدام الفكرة إيجاد الصلة بينها وبين ذلك التيار الذى يتألف من إحساسات ووجدانات، وآمال ورغبات ومن نشاط عقلى يلائم بين نواحي التفكير المختلفة ، وهو الذى تتكون منه حياة الإنسان » .

إن أهم انقلاب فى أساليب التربية الحديثة ، وأهم غرض يقصد من الطرق التى ابتكرها الناهيون من المربين ، كطريقة المشروعات ، وطريقة دكرولى ، وغيرهما ، هو العمل على تحرير التعليم من عبادة الأفكار الخاملة ، وتحويله إلى التثقيف العقلى وتنمية الشخصية . ولكن الامتحانات التى هى من قبيلى امتحاناتنا عقبة كؤود فى سبيل تطبيق هذه الأساليب ، لأنه ما دام الامتحان يتطلب مرد حقائق مستظهرة من غير تصرف ، ومادام المعلمون والطلاب وأولياء أمورهم والمشفون على التعليم بوزارة المعارف (فى هذا الوقت) يحرصون - وطبيعى أن يحرصوا - على نجاح الطلاب فيها ، فالمدارس لا تتخرج من أن تصرف الوقت فى حشو أذهانهم بالأفكار الخاملة وتضن بهذا الوقت أن يصرف فى مطالعة حرة ، أو بحث شخصى ، أو نشاط اجتماعى ، أو هواية عملية ، أو رياضة بدنية ، أو غير ذلك من نواحي التربية الصحيحة ، التى تنمى الشخصية وتوسع الفكر . ولو أن أسئلة الامتحان كانت تتطلب إعمال الفكر ، وسلامة الحكم ، وحسن التصرف فى استخدام المعلومات ، لاتيجه التعليم إلى العناية بهذه الصفات ووجوه النشاط التى تنمىها . فأهم ما تمس إليه الحاجة إذن من وجوه الإصلاح فى الامتحانات عندنا هو إصلاح نوع الأسئلة .

وليس أثر الامتحانات السيى " فى توجيه التعليم مقصوراً على ما سبق ، إن الامتحانات العمومية من أقوى عوامل الجمود فيه . لأنها تتطلب وحيداً : المناهج والأساليب ، فتحول دون مراعاة المدارس لحاجات تلميذها مع الحوادث التى تثير اهتمام المجتمع ، وتمنعها من مساهمة

ميل تلاميذها وإطلاق الحرية المعقولة لهم في الدراسة . كما أنها تحدد من حرية المعلمين في اختيار مادة الدروس وطرق التدريس ، لأن المدرس لا يعرف مزاج الممتحن الخارجى ، وطريقته في وضع الأسئلة ، ووجهة نظره فيما يتعلق بالأهمية النسبية لموضوعات المنهج ، والمدى الذى يذنبى الذهاب إليه في تدريس كل منها ، ولذلك يخشى أن يعتمد على حكمه الشخصى في العناية بنقط معينة يراها هامة ، أو إغفال نقط أخرى يراها ثانوية أو غير ملائمة لمستوى تفكير تلاميذه وميولهم ، لئلا تنجى أسئلة الامتحان فيما أغفله . ونتيجة ذلك أن أكثر المدرسين يعنون بدرجة واحدة بكل كبيرة وصغيرة ، ويصرفون الوقت في تدريب تلاميذهم على أنواع معقدة من الأسئلة والتمارين لا يؤمنون بفائدتها ، ولا يقصدون منها إلا سد الأبواب على الممتحن الخارجى .

ثم إن شبح الامتحانات المائل دائما للأذهان يدفع المدرسين إلى إرهاق التلاميذ بالعمل ، والإكثار من التكاليف والواجبات المنزلية ، ويدفع الكثيرين من الآباء إلى الضغط على أبنائهم ضغطا مستمرا لحملهم على الاستدكار . غافلين عما يلزم لصحتهم من راحة ورياضة ، ويدفع التلاميذ عند اقتراب موعد الامتحان إلى وصل الليل بالنهار في مجهود مضن ، ينتهى في كثير من الأحيان بإضعاف أجسامهم وإتلاف أعصابهم .

وهناك ناحية أخرى لموضوع الامتحانات . وهى البحث عن مبلغ دقتها من حيث هى أداة للقياس . وبعبارة أخرى البحث عما إذا كانت الدرجة التى يحصل عليها طالب فى مادة من المواد تدل بالضبط على حقيقة قوته فى تلك المادة . إن أقل خبرة بالامتحانات المألوفة لتشعر بأن درجاتها بعيدة عن الدقة فى قياس قوى التلاميذ ، ولكن مدى الأخطاء التى تتعرض لها هذه الدرجات لم يظهر بجلاء إلا نتيجة للأبحاث التجريبية التى قام بها العلماء . وقد أجرى معهد التربية بإشراف الأستاذ القباني فى شهرى يناير

وفبراير سنة ١٩٣٠ تجربة من هذا القبيل على امتحانات المدارس المصرية ، قد يكون من الطريف أن تعرض نتائجها ، لنلمس مقدار الخطأ الذى تقع فيه بتعليق أهمية كبيرة على فروق يسيرة فى الدرجات بين تلميذين أو بين امتحانين للتلميذ الواحد .

والفكرة التى قامت عليها التجربة هى أن من علامات أداة القياس المضبوطة أننا إذا قسنا بها الشيء الواحد مرتين فلا بد أن تتفق نتائج القياسين ولو تغير الشخص الذى يقوم بالقياس ، أو الظروف التى أجرى فيها . فإذا قسنا طول مستقيم مرتين بمسطرة واحدة أو بمسطرتين متشابهتين . فلا يختلف طوله فى المرة الثانية عنه فى المرة الأولى مادام الشخصان اللذان قاما بالقياس فى المرتين يعرفان كيفية استعمال المسطرة . فبالمثل إذا كان الامتحان مقياسا مضبوطا لما يحاول أن يقيسه ، فإننا إذا أعطينا فصلا من التلاميذ امتحانين متماثلين فى مادة واحدة فى زمنين متقاربين ، وجب أن تتفق نتائجهما . فلتحقيق ذلك انتهز الأستاذ القباني فرصة امتحان وسط السنة فى شهر يناير سنة ١٩٣٠ ، فاتفق مع نظار بعض مدارس القاهرة الابتدائية والثانوية على إعادة الامتحان لبعض الفصول فى مواد معينة .

وقد أعطى الامتحان الثانى فى الأسبوع الأول أو الثانى عقب العودة من أجازة وسط السنة ، واتخذت كل الاحتياطات الممكنة لضبطه وجعله مكافئا لامتحان وسط السنة ومماثلا له فى جميع الوجوه . فاختر لتأليفه وتصحيحه أقل مدرسى المادة المختارة فى كل مدرسة . وأعطيت لهم نسخة من امتحان وسط السنة فى تلك المادة وطلب منهم أن يضعوا امتحانهم على نمطه ، وأن يجعلوه مماثلا له فى الصعوبة . وألا تخرج الأسئلة عن الموضوعات التى امتحن فيها التلاميذ فى وسط السنة ، بحيث أن هذه الأسئلة كان يصح أن تكون هى التى أعطيت فى امتحان وسط السنة لو أن المدرس الذى وضعها كان هو الذى اختير لوضع أسئلته .

وقد أخطر التلاميذ عن الامتحان مقدما في بعض الفصول قبل إجرائه بمدة تسمح لهم بالمذاكرة والاستعداد له ، وأعطى للبعض الآخر من غير إخطار سابق ، للمقارنة بين النتائج في الحاليتين . وبعد إجراء الامتحان صححت الإجابات بكل عناية ودقة ، ثم قورن بين نتائجه ونتائج امتحان وسط السنة في كل فصل . وفي الصفحة التالية جدول يحتوي على نتيجة المقارنة ، وقد وضعت العلامة * فوق المواد التي أخطر التلاميذ عن امتحانها مقدما .

وبالنظر إلى هذا الجدول نرى أن المقارنة قد أجريت بثلاث طرق مختلفة وهي :

أولا : حسب عدد التلاميذ الذين نجحوا في الامتحان معاً ، وعدد الذين رسبوا في الامتحان معاً ، وعدد الذين نجحوا في أحد الامتحانين ورسبوا في الآخر . وقد ظهر أن ٤٠ تلميذا من جملة تلاميذ المدارس الثانوية الذين أعيد امتحانهم - وعددهم ١٥٦ - قد نجحوا في أحد الامتحانين ورسبوا في الآخر ، أى أن نجاحهم أو رسوبهم في أى من الامتحانين مطعون فيه . ومثلهم ١٤ من تلاميذ مدرسة العباسية الابتدائية ، وعددهم ٦٥ .

ثانياً : قسمت النهاية العظمى للدرجات في كل مادة إلى خمس أقسام متساوية وبحث عن القسم الذى تقع فيه درجة كل طالب في كل من الامتحانين . ومن ذلك أمكن معرفة ما إذا كان الطالب في الامتحان الثانى قد بقى في القسم الذى كان فيه في الأول أو تقدم أو تأخر ، وما عدد الأقسام - أو الخطوات - إلى تقدمها أو تأخرها . فمثلا إذا كان الطالب قد حصل في امتحان وسط السنة في مادة ما على ٧ درجات من ٢٠ فدرجته تقع في القسم الثانى ، وإذا كان قد حصل في نفس المادة على ١٣ من ٢٠ في الامتحان الثانى فدرجته هذه تقع في القسم الرابع ، وبذلك يكون قد تقدم خطوتين .

الموازنة بين نتائج امتحانين متكافئين في مادة معينة للفصل الواحد

معدل الارتباط بين نتائج الامتحانين	عدد من تقدمت درجاتهم أو تأخرت بقدر :				عدد من مجموع الامتحانين درست في الآخر	عدد من مجموع الامتحانين درست في الآخر	عدد من مجموع الامتحانين درست في الآخر	عدد من مجموع الامتحانين درست في الآخر	العدد الكلي المتكافئين	الفرقة	المدرسة	المادة
	أربع خطوات	ثلاث خطوات	خطوتين	خطوة واحدة								
مدرسة ثانوي لمارس												
٠.٩٠	-	١	٣	١٤	٥	٠	٢	٢٦	٣٢	الرابعة - الثانوية	الثانوية	الرياضة
٠.٨٥	-	-	٣	٦	٧	٦	١٤	٤	٢٥	الثانوية	الثانوية	الرياضة
٠.٨٢	-	٢	٥	١٤	١٤	١٢	١٢	٦	٣٢	الثانوية	الثانوية	الرياضة
٠.٦٨	-	-	٣	١٠	٣	٥	٥	٢٠	٢٨	الثانوية	الثانوية	الرياضة
٠.٣٩	١	٢	٣	٢٢	١١	١٤	١٤	١٣	٣٨	الثانوية	الثانوية	الرياضة
	١	٥	١٧	٦٦	٤٠	٤٧	٤٧	٦٩	١٥٦	الثانوية	الثانوية	الرياضة
مدرسة ابتدائي												
٠.٦٨	-	-	٣	١٧	٨	٦	٦	٢٢	٣٦	الثانوية	الثانوية	الرياضة
٠.٧٥	-	-	٢	١٢	٦	٣	٣	٢٠	٢٩	الثانوية	الثانوية	الرياضة
	-	-	٥	٣٠	١٤	٩	٩	٤٢	٦٥	الثانوية	الثانوية	الرياضة

٥ في هذه الحالات أخطر التلاميذ عن الامتحان الثاني مقدما للامتحان الأول

(جدول ١٣)

ونرى من الجدول أن ٨٩ من تلاميذ المدراس الثانوية الذين أعيد امتحانهم - أى أكثر من نصفهم - قد انتقلت درجاتهم في الامتحان الثانى إلى قسم غير الذى كانت فيه في وسط السنة ، ومنهم ٦ انتقلوا ثلاث خطوات أو أكثر ، أى أن درجاتهم كانت في أحد الامتحانين أقل من ٢٠٪ من النهاية العظمى وفي الامتحان الثانى أكثر من ٦٠٪ منها ، أو كانت في أحد الامتحانين أقل من ٤٠٪ من النهاية العظمى وفي الامتحان الثانى أكبر من ٨٠٪ منها !

وفي مدرسة العباسية الابتدائية نجد أن ٣٥ تلميذاً من ٦٥ انتقلت درجاتهم إلى قسم غير الذى كانت فيه ، ومنهم ٥ انتقلت درجاتهم خطوتين .

ثالثاً : حسب « معامل الارتباط » بين درجات التلاميذ في الامتحانين في كل مادة .

وظاهر من الخانة الأخيرة في الجدول أن الاتفاق بين درجات الامتحانين في مجموعها ليس كبيراً ، بل هو ضعيف جداً في بعض المواد ، إذ المقرر في القياس العقلى أنه لا يصبح التعويل على نتائج أى اختبار إذا قل معامل الارتباط بين الدرجات التى نحصل عليها من صورتين متكافئتين منه عن نحو ٧٥ ر .

ولا يجوز أن ننسب اختلاف الدرجات في الامتحانين إلى عامل ثابت ، مثل قلة إهتمام التلاميذ بالامتحان الثانى أو كثرة استعدادهم لامتحان وسط السنة . لأن هذا لوصح لترتب عليه نقص غام في درجات سائر التلاميذ في الامتحان الثانى عنها في امتحان وسط السنة ، والواقع غير ذلك ، كما يتبين من النظر إلى الجدول التالى ، وفيه الدرجات التى حصل عليها التلاميذ في كل من الامتحانين ، في الحالات التى كان فيها اختلاف كبير . فمن هذا الجدول نرى أن بعض التلاميذ زادت درجاتهم في الامتحان الثانى عليها في

التلاميذ الذين تقدمت أو تأخرت درجاتهم في الامتحان الثاني عنها
في الامتحان الأول بقدر خطوتين أو أكثر
« الخطوة تعادل ٢٠٪ من النهاية المعطى لكل مادة »

المادة	المدرسة	المفحوص	الدرجة في امتحان وسط السنة	الدرجة في الامتحان الثاني
اللغة الإنجليزية النهاية المعطى ٣٠	التوفيقية الثانوية	١	١٠	٢٠
		٢	$١٦\frac{1}{4}$	٢٤
		٣	$١٧\frac{1}{4}$	٢٤
		٤	$١٨\frac{1}{4}$	$٥\frac{1}{4}$
الهندسة النهاية المعطى ٢٠	شبرا الثانوية	٥	٦	١٣
		٦	١٦	٩
		٧	٤	١٣
الهندسة النهاية المعطى ٢٠	التوفيقية الثانوية	٨	٧	١٤
		٩	٩	٢٠
		١٠	١٣	٧
		١١	٦	٤
		١٢	١٤	٧
		١٣	٢٠	٧
		١٤	$١٩\frac{1}{4}$	٧
		١٥	٢٠	٦
الطبيعة والكيمياء النهاية المعطى ٤٠	شبرا الثانوية	١٦	$١٠\frac{1}{4}$	٢٤
		١٧	$١٤\frac{1}{4}$	$٢٤\frac{1}{4}$
		١٨	١٤	٢٦
الجغرافية النهاية المعطى ١٥	الابراهيمية الثانوية	١٩	٦	$٢\frac{1}{4}$
		٢٠	$٥\frac{1}{4}$	٩
		٢١	$٢\frac{1}{4}$	٦
		٢٢	$١\frac{1}{4}$	١٢
		٢٣	٤	١١
اللغة العربية النهاية المعطى ٤٠	العباسية الابتدائية	٢٤	٣٤	٢٠
		٢٥	٣٥	٢٠
		٢٦	١٦	٦
الجغرافية النهاية المعطى ١٥	العباسية الابتدائية	٢٧	٩	٤
		٢٨	١٢	٨

امتحان وسط السنة ، والبعض الآخر نقصت درجاتهم فيه ، مما يدل على أن الاختلاف يرجع إلى أسباب تتصل بطبيعة الامتحان .

وأهم هذه الأسباب ما يأتي :

أولاً : الامتحان العادى فى أية مادة لا يقيس شيئاً واحداً ، بل أشياء كثيرة مختلطة بعضها ببعض ، ومن القواعد المقررة عند علماء القياس العقلى أن المقياس الذى يستعمل لقياس شيئين فى آن واحد لا يقيس شيئاً ما . فامتحان الجغرافية مثلاً يختبر تحصيل التلاميذ فى بعض موضوعات هذه المادة ، ويختبر فى الوقت نفسه قوة تفكيرهم ، وقدرتهم على التعبير عن أفكارهم ، وقوتهم فى رسم الخرائط ، وسرعتهم فى قراءة الأسئلة وفى كتابة الإجابة عنها ، غير ذلك من العوامل التى تتوقف عليها الدرجة التى تعطى للتلميذ . وهذه العوامل تختلط بعضها ببعض بنسب تختلف من امتحان لامتحان ، مما يؤدى إلى اختلاف نتيجتها مجتمعة .

ثانياً : ورقة الامتحان العادى تحتوى على عدد قليل من الأسئلة ، لا يمكن أن يتناول جميع موضوعات المنهج الذى يمتحن فيه التلميذ ونقطه . ولذلك قد يصادف الخط تلميذاً فى أحد الامتحانين فتأتى الأسئلة مما يتقنه ، وقد يعاكسه الخط فى الامتحان الثانى فتأتى بعض الأسئلة مما لا يعرفه ، ويتسبب من ذلك اختلاف درجته فى الامتحانين . فالدرجة لا تمثل قوة التلميذ فى المادة بأجمعها ، بل فى ذلك الجزء منها الذى اشتملت عليه ورقة الأسئلة .

ثالثاً : الامتحان مقياس ذاتى ، ومعنى هذا أن النتائج التى يعطيها تتأثر برأى الشخص الذى يقوم به وتقديره . وتتجلى هذه الذاتية فى جميع خطوات الامتحان : فالحكم على درجة صعوبة الأسئلة ومقدار ملاءمتها لقوة التلاميذ هو حكم ذاتى يتوقف على رأى واضعها الشخصى الذى

لا ضابط له . وتقدير الدرجة التي تستحقها إجابة كل تلميذ هو تقدير ذاتي يتوقف على رأى المصحح ، وقد أثبتت التجارب العلمية الكثيرة اختلاف المصححين اختلافاً بيناً في تقدير الدرجة للإجابة الواحدة ، بل واختلاف المصحح الواحد في تقدير الدرجة التي تعطى للإجابة إذا عرضت عليه مرات تفصلها فترات طويلة أو قصيرة من الزمن . ثم إن تحديد مستوى النجاح أيضاً هو تحديد ذاتي يتوقف على رأى المشرح ، وهو رأى اعتباطي لا يستند إلى سند معقول ، ولذا نجد الحد الأدنى للنجاح في بعض الأحيان ٤٠٪ ، وفي بعضها ٥٠٪ ، وفي بعضها ٦٠٪ ، وكل هذه النسب المثوية ليس لها في الحقيقة مدلول ثابت ، فدلولها يتوقف على ظروف كثيرة ، من أظهرها مقدار صعوبة الامتحان أو سهولته :

وقد انتهى الأستاذ القباني من هذا البحث إلى أن نظام الامتحانات بما فيه من عيوب ، كشف عن بعضها البحث ، لا تقل شأنًا عن الاختبارات الموضوعية المقتنة التي يمكن أن تفيد في نفس الأغراض . ذلك أن الاعتماد على الامتحانات وحدها لم يعد يفي الغرض ولا كذلك الاعتماد على الاختبارات وحدها .

بحوث وزارة التربية والتعليم

قامت وزارة التربية والتعليم بالإقليم الجنوبي بعدة بحوث ودراسات كما أنشأت الكثير من الاختبارات أدوات لهذه البحوث ، والبحوث ما زالت جارية ، وأخرى يعزم القيام بها . وقد اشترك المؤلف في عدد من هذه البحوث والاختبارات وهذا سرد سريع لأهم هذه البحوث :

(١) البحوث التي قاربت الانتهاء :

١ - بحث التقدير الرقي للموظفين .

٢ - بحث تشخيص الأخطاء الإملائية .

٣ - بحث تيسير الكتابة والطباعة في اللغة العربية .

(ب) البحوث الجارية :

١ - بحث كيفية الانتفاع بوقت الفراغ .

٢ - بحث مستويات الدراسة في العلوم والرياضة في المرحلة الإعدادية والثانوية .

٣ - بحث المستوى التحصيلي والتأخر الدراسي في المرحلة الابتدائية .

٤ - بحث مشكلة نظام القبول بالجامعات .

٥ - بحث تطوير مناهج التعليم في الأزهر .

٦ - بحث الطلبة المتفوقين في المدارس الثانوية .

٧ - بحث مقارنة ما ينتج من تعليم لغة أجنبية واحدة أو لغتين .

٨ - بحث النهوض بالتعليم في مناطق الحدود .

(ح) بحوث لا زالت في المراحل الأولى للتنفيذ :

١ - بحث مدى ملائمة التعليم الابتدائي للبيئة .

٢ - بحث تقويم التعليم الإعدادي الفني .

٣ - بحث دور المدرسة في تكوين الوعي القومي وتنميته .

٤ - بحث تعليم الخط العربي في المدارس الابتدائية .

(د) بحوث لا زالت في مرحلة التخطيط :

١ - برنامج وحدة بحوث المعاهد الزراعية العليا .

٢ - برنامج وحدة البحوث التربوية بكلية المعلمين .

٣ - برنامج وحدة بحوث العلوم والرياضة بكلية المعلمين .

٤ - برنامج وحدة البحوث الإحصائية .

وهذا بيان بالاختبارات التي أعدها الوزارة ، والتي تهتمنا في

هذا القسم :

اختبارات المستوى التحصيلي للمرحلة الابتدائية في اللغة العربية :

اختبار القدرة على الفهم في القراءة الصامتة للصف الثالث .

(صورتان ١ ، ب) .

اختبار القدرة على الفهم في القراءة الصامتة للصف الرابع .

(صورتان ١ ، ب) .

اختبار القدرة على الفهم في القراءة الصامتة للصف الخامس .

(صورتان ١ ، ب) .

اختبار القدرة على الفهم في القراءة الصامتة للصف السادس .

(صورتان ١ ، ب) .

سلسلة الاختبارات التحصيلية للمرحلة الابتدائية :

في اللغة العربية : اختبار لكل صف من الأول إلى الخامس .

في الحساب : للصف الأول والثاني والثالث (عمليات ومسائل)

والرابع (عمليات ومسائل) والخامس (عمليات ومسائل وهندسة) .

في المواد الاجتماعية : اختبار لكل صف من الثالث إلى الخامس .

في مبادئ العلوم والتربية الصحية : اختبار لكل صف من الثالث

إلى الخامس .

اختبارات الحساب للمدارس الابتدائية :

الصف الأول .

الصف الثاني .

الصف الثالث (عمليات ومسائل) .

بحث التعليم الابتدائي

من أهم البحوث التي قامت بها وزارة التربية والتعليم بالإقليم الجنوبي والتي أسّرك فيها المؤلف ، بحث التعليم الابتدائي . ومرجع هذه الأهمية

لا إلى ضخامة الجهود المتضمنة فيه وحسب بل وأيضاً إلى أن التعليم الابتدائي هو القاعدة الكبيرة لكل مراحل التعليم التالية . وهو ، إلى جوار هذا ، يكلف الدولة نفقات طويلة خاصة وأنه بسبيل أن يكون إلزامياً فضلاً عن كونه مجانياً . كما أن السنوات التي يقضيها الفرد في هذا التعليم من أخطر سنوات تشكيل تكوينه النفسى والانفعالى والجسمى والعقلى . ومن ثم يجب ، بفرض أنه ممكن ، استغلال هذه المرحلة في إعداد المواطن الصالح .

وقد بدأ البحث في ٢٧/٣/١٩٥٥ وانتهى التقرير في سنة ١٩٥٧ وهذه خطواته كما وردت في تقرير لجنة التعليم الابتدائي .

أولا

الخطوة الأولى : البحث الاستطلاعى .

الخطوة الثانية : البحث التجريبي .

- ١ - استمارة يملؤها المدرس (الصورة التجريبية) .
- ٢ - استمارة يملؤها ناظر المدرسة (الصورة التجريبية) .
- ٣ - استمارة يملؤها مفتش القسم (الصورة التجريبية) .

دراسة تقارير اللجنة الفرعية :

أولاً : بيانات إحصائية عن البحث التجريبي في المناطق الثلاث (القاهرة

الشمالية - طنطا - المنيا) التلاميذ والفصول - مستويات المدارس - مجموع تلاميذ الفرق المختلفة - المباني - نفقات التلميذ .

ثانياً : تفريغ استمارات المدرسين والنظار والمفتشين .

تقرير عن جدول تفريغ استمارات النظار .

ثالثاً : دراسة الاستمارات الثلاث .

١ - تقرير عن استمارة المدرس .

٢ - تقرير عن استمارة الناظر .

٣ - تقرير عن استمارة مفتش القسم .

استفتاء المدرس (الصورة النهائية) .

استفتاء الناظر (الصورة النهائية) .

استفتاء المفتش (الصورة النهائية) .

كراسة مفتش القسم .

رابعاً : استفتاء الأهالى .

(أ) بطريقة فردية .

(ب) بطريقة الندوات .

البحث فى آراء الأهالى والمواطنين فى مسائل التعليم الابتدائى .

أولاً : وجهة نظر الأهالى فى شخصية تلميذ المدرسة الابتدائية :

(عاداته وهواياته - الإقبال على الاطلاع والقراءة - معرفة التلميذ لحقوقه وواجباته - الإلمام بالأحداث العامة الجارية . شخصية الطفل وثقته بنفسه) :

ثانياً : وجهة نظر الأهالى فى المدرسة والتعليم الابتدائى :

(مدى استفادة التلاميذ منه من حيث عدد سنوات الدراسة - اليوم المدرسى والجمع بين التعليم فى نصف اليوم والعمل فى النصف الآخر - استفادة التلاميذ من نظام التغذية - التعليم المختلط « اشترك البنات مع أولادهن ») .

ثالثاً : وجهة نظر الأهالى فى أثر المدرسة فى البيئة :

(هل تبيحت المدرسة فى إزالة الفوارق بين الطبقات - هل يشترك

المدرسون مع الأهالى فى النهوض بالبيئة - هل امتد أثر المدرسة إلى تثقيف الآباء والأمهات .

بيانات عامة عن نتائج الاستفتاء الفروى :

أولاً : أثر التعليم الابتدائى على التلميذ :

ثانياً : نوع المدرسة المفضل .

ثالثاً : أثر التعليم فى البيئة .

رابعاً : المباني .

خامساً : المدارس الخاصة بمصروفات .

سادساً : توفير السكن للمعلمين والمعلمات .

سابعاً : الإشراف الصحى .

ثانياً

حالة التعليم الابتدائى :

حالة التعليم الابتدائى فى صورة بيانية :

المدارس الابتدائية .

مدارس إعداد معلمى المرحلة الأولى :

المباني :

الفصول والتلاميذ .

المدارس المستكملة وغير المستكملة .

المدرسون .

الميزانية .

حالة التعليم الابتدائى كما تصورها الاستفتاءات .

أولاً : استفتاء مفتش القسم .

ثانيا : استفتاء ناظر المدرسة .

ثالثا : استفتاء المدرس .

حالة التعليم الابتدائي كما تصورها الاختبارات .

أولا : امتحان التحصيل .

ثانيا : امتحان القبول بالمدارس الإعدادية .

التوصيات

أولا : أوجه الإصلاح ووسائله .

ثانيا : علاقة التعليم الابتدائي بالمراحل الأخرى .

ثالثا : الاقتراحات بشأن عمل تخطيط لتعميم التعليم الابتدائي وتحسينه .

رابعا : خاتمة بشأن الاقتراحات الخاصة بالتشريعات والقرارات

اللازمة لتنفيذ الإصلاحات التي تقترحها اللجنة .

نماذج الاختبارات المستخدمة في هذا البحث

الشكل (٢٦)

اختبار العربي للسنة الثانية الابتدائية

الصورة الأولى الفقرة (٤)

يحمل الصياد حمامة

صاد الرجل سمكة

الرجل يأكل السمكة



وهنا يوثق المفحوص على الجملة التي يصورها هذا الرسم .

الفقرة المكافئة من الصورة الثانية الفقرة (٤)



(شكل ٢٧)

وطبيعي أن يطلب نفس الشيء من المفحوص ؟

الفقرة المكافئة من الصورة الثالثة الفقرة (٤)



(شكل ٢٨)

الفقرة الرابعة من الصورة المكافئة الرابعة

(شكل ٢٩)



اختبار المستوى التحصيلي للسنة الخامسة الابتدائية ؟

الصحة :

١ - ضع علامة / أمام العادة الصحية ، وعلامة X أمام العادة غير

الصحية فيما يأتي :

تلميذ يكثر من شرب الماء بعد الأكل .

تلميذ يستعمل فرشاة الأسنان بعد الأكل .

تلميذ يجد يديه نظيفة قبل الأكل فيجلس إلى المائدة ويأكل دون غسلهما .

تلميذ يسهر بالليل كثيراً وينام بالنهار .

تلميذ يكثر من أكل المكرونة ويقلل من أكل الخضروات .

تلميذ يحذ أظافره طويلة فيقضمها بأسنانه .

تلميذ يفتح شبابيك حجراته وبابها للتهوية ثم يجلس وسطها .
تلميذ يقرأ في ضوء خافت .

للتاريخ :

١ - اختر من المستطيل التالى اسم البناء الذى أقامه كل ملك من هؤلاء الملوك ؟

الملك خوفو شيد	مدينة الاسكندرية
الملك مينا بنى	منارة الاسكندرية
الملك بطليموس الثانى أقام	مدينة الجدار الأبيض
	الحرم الأكبر بالجيزة
	معبد الدير البحرى

- ٢ - ضع خطأً تحت الكلمة الصحيحة بين كل قوسين فى الجمل الآتية :
- (أ) موقع مدينة منف الآن هو (القاهرة - البدرشين - الأقصر) .
- (ب) صنع المصريون القدماء الورق من (قش الأرز - نشارة الخشب - نبات البردى) .
- (ج) اعتقد المصريون القدماء أن الإله الذى سيحاسبهم على أعمالهم يوم البعث هو (رع - أمون - أوزيريس) ؟
- (د) يرجع الفضل إلى الملك أحسن فى طرد (بدو الصحراء - الهكسوس - الليبيين) من مصر .
- (هـ) انتشرت المسيحية فى مصر فى عهد (الفراعنة - اليونان - الرومان) .

الجغرافيا :

١- أكمل الجمل الآتية :

- (أ) من أهم منتجات السودان و
 (ب) لون الهولنديين أما الهنود فلونهم
 (ج) أكبر قارات العالم هي
 (د) الجوفى بلاد الاسكيمو

٢- (أ) معك بوصلة ، نظرت إليها فوجدت الشمال أمامك ، فأين يقع الشرق بالنسبة لك ؟

(ب) كيف يعرف المصريون ليلة أول رمضان من كل عام ؟

وقد وزع تلاميذ كل فرقة حسب درجاتهم التي حصلوا عليها في امتحان كل مادة إلى أربع فئات .

- (أ) تلاميذ درجاتهم أقل من ٢٥٪ من النهاية الكبرى للدرجات .
 (ب) تلاميذ درجاتهم أقل من ٢٥٪ إلى ٤٩٪ من النهاية الكبرى للدرجات
 (ج) تلاميذ درجاتهم أقل من ٥٠٪ إلى ٧٤٪ من النهاية الكبرى للدرجات
 (د) تلاميذ درجاتهم ٧٥٪ فما فوق

فكانت نتائج هذه الامتحانات بصفة إجمالية كالآتي :

نتائج الاختبارات التحصيلية

الفرقة	المادة	حصة عدد المتحدين	عدد التلاميذ في كل من الفئات				النسبة المئوية لعدد التلاميذ في كل فئة			
			أ	ب	ج	د	أ	ب	ج	د
الثانوية	اللغة العربية	١٧١٤	٢٥٣	٣٨٦	٥١٧	٥٥٨	١٤٠٧٦	٢٢٠٥٢	٣٠٠١٦	٣٢٠٥٦
	حساب	٩١٢	١٠٢	١٩٥	٢٩٧	٣١٨	١١٠١٨	٢١٠٣٨	٣٢٠٥٧	٣٤٠٨٧
	اللغة العربية	١٦١٠	٢٤٤	٣٠١	٤٧٢	٥٩٣	١٤٠١٥	١٨٠٧٠	٢٩٠٣٢	٣٦٠٨٣
	حساب	١٠٥٤	٣٠٤	٢٤٨	٣١٠	١٩٢	٢٨٠٨٤	٢٣٠٥٣	٢٩٠٤١	١٨٠٢٢
السادسة	اللغة العربية	٥٨٥	٢٣	١١١	٢٥٩	١٩٢	٣٠٣٩	١٨٠٩٧	٤٤٠٢٧	٢٢٠٨٢
	حساب	٨١٧	٢٤٢	٢٢٢	٢٢٦	١٢٧	٢٩٠٦٢	٢٧٠١٧	٢٧٠٦٦	١٥٠٥٤
	معلومات	٧٢٥	١٢٩	١٦٣	٢٧٨	١٥٥	١٧٠٧٩	٢٢٠٤٨	٢٨٠٣٤	٢١٠٣٨

(جدول ١٥)

كذلك عمل الجدول التالى لتوضيح النسب المئوية لعدد التلاميذ الذين
امتحنوا فى كل فرقة ، وفى كل مستوى من المستويات الأربعة .

النسبة المئوية لعدد التلاميذ موزعين حسب مستوى

درجاتهم فى الامتحان التحصيلى لكل مادة

الفرقة	المادة	النسبة المئوية لعدد التلاميذ الحاصلين على درجات			
		أقل من* ٢٥٪	٢٥٪ إلى أقل من ٥٠٪	٥٠٪ أقل من ٧٥٪	٧٥٪ فأكثر
الثانية	اللغة العربية	١٥	٢٢	٣٠	٣٣
	حساب	١١	٢١	٣٣	٣٥
الرابعة	اللغة العربية	١٥	١٩	٢٩	٣٧
	الحساب	٢٩	٢٤	٢٩	١٨
السادسة	اللغة العربية	٤	١٩	٤٤	٣٣
	الحساب	٣٠	٢٧	٢٨	١٥
	معلومات	١٨	٢٣	٣٨	٢١

(جدول ١٦)

ومن هذا نتبين أن من الممتحنين فى مادة اللغة العربية من تلاميذ
السنة الثانية .

١٤,٧٦٪ لم يصلوا إلى مستوى الفرقة السابقة (الفرقة الأولى) .

٢٢,٥٢٪ لم يصلوا لأكثر من مستوى الفرقة السابقة (الفرقة الأولى) .

وكذا ٣٠,١٦٪ نجحوا فى تحصيل مستوى الفرقتين (الأولى والثانية) .

٣٢,٥٦٪ منهم ٣٠,١٦٪ بدرجة جيدة ، ٢,٤٠٪ بدرجة ممتازة .

ولم نرد أن نستطرد في استعراض نماذج من الاختبارات أو النتائج .
إذ أنها ، برغم انتهاء البحث ، ما زالت تحت التجربة في ضوء التطورات
الأخيرة في نظام التعليم . ويمكن الرجوع إلى مطبوعات الوزارة لمعرفة المزيد :

بحوث المؤلف

قام واشترك وأشرف المؤلف على عدد من البحوث التربوية والاختبارات
التحصيلية وأهمها .

- ١ - بحث تقييم التحصيل في المدرسة الابتدائية .
 - ٢ - بحث تقييم اختبارات القبول بالمدرسة الإعدادية .
 - ٣ - بحث تتبع خريجي المدارس الصناعية الثانوية .
 - ٤ - بحث تقييم بطاقات الحساب التدريسية التشخيصية .
 - ٥ - إعداد اختبار استانفورد الجديد في الحساب للاستخدام في المدارس
المصرية بالاشتراك مع الأستاذ إسماعيل القباني .
 - ٦ - إعداد بطاقات الحساب التدريسية التشخيصية بالاشتراك مع
الأستاذ إسماعيل القباني .
- وسنبداً بعرض بحث التعليم الابتدائي لنفس الأسباب التي سبق ذكرها :

بحث تقييم التحصيل بالمرحلة الابتدائية

في سنة ١٩٥٥ أشرف المؤلف على بحث لتقييم التحصيل بالمرحلة
الابتدائية كان الغرض الأساس منه مقارنة التحصيل في المناطق والمدارس
المختلفة ، إذ يرى البعض أن المناطق القريبة من المدن الكبرى تحظى بإمكانات
وكفايات وإشراف وتوجيه لا يتوفر في المناطق الأخرى النائية بينما يرى
البعض الآخر أن تحصيل التلاميذ خارج المدن الكبرى لا يقل عن تحصيل
نظرائهم فيها . ومن هذا كان لا بد من القيام بدراسة أو دراسات موضوعية
للإجابة على هذه المشكلة .

خطة البحث

أولاً : إعداد الاختبارات : كان من الضروري إعداد اختبارات موضوعية موحدة تطبق على عينات من تلاميذ المرحلة الابتدائية . تمثل جميع المناطق والمستويات . وتم إعداد أربعة اختبارات موضوعية وضعت لكل منها صورة تجريبية أولى طبقت على عينة تمهيدية . ثم عدلت ووضعت في صورة تجريبية ثانية وطبقت مرة ثانية على عينة تمهيدية أخرى من تلاميذ المدارس الابتدائية . وعلى ضوء التجربة الثانية ، وما ظهر من تحليل النتائج إحصائياً ، عدلت الاختبارات ووضعت في صورتها النهائية . وفيما يلي وصفها :

١ - اختبار في اللغة العربية : يتكون الاختبار من ٧ أسئلة « فقرات » ينقسم بعضها إلى ثلاث أو أربع أجزاء كما يحتوى الاختبار قطعة للإملاء والزمن المحدد للاختبار كله ٦٠ دقيقة خصص منه ٢٠ دقيقة للإملاء .

٢ - اختبار في الحساب : يتكون من ٢٠ فقرة والزمن المحدد له ٤٥ دقيقة .

٣ - اختبار في المعلومات العامة : يتكون من ٥٠ فقرة ، من نوع فقرات الاختيار من عدة إجابات ، والزمن المحدد له ٣٠ دقيقة .

٤ - اختبار للذكاء العام : يتكون من ٤٠ فقرة ، من نوع فقرات الاختيار من عدة إجابات والزمن المحدد له ٢٥ دقيقة .

ثانياً : جمع المعلومات عن التلميذ : أعدت لهذا الغرض بطاقة روعي فيها ذكر المعلومات التي تلقى ضوءاً على نتائج التلميذ في الاختبارات الموضوعية السالفة الذكر . وقد قام المدرسون بملء البيانات في هذه البطاقات عن كل تلميذ بناءً على معرفتهم للتلميذ وتقديرهم الذاتي ، كل مدرس في الناحية التي تخصه . كما خصصت في البطاقة مسافات لتسجيل

نتائج التلميذ في إختبارات البحث كذلك خصصت مسافات لتسجيل نتائج امتحان القبول بالمدرسة الإعدادية لمن يتقدم لها من هؤلاء التلاميذ .

ثالثاً : اختيار العينة : تم اختيار ١٤٢٩ تلميذاً (٩٨٥ ولداً ، ٤٣٤ بنتاً ، ١٠ لم ترد عن جنسهم بيانات) يمثلون ٢٠ مدرسة تمثل ٩ مناطق تعليمية على الأسس التالية :

(١) تمثيل المناطق التعليمية المختلفة بالإقليم المصرى كى يمكن تمثيل تلاميذ السنة الرابعة بالمدارس الابتدائية (كانت المرحلة الابتدائية وقتذاك خمسة سنوات) .

(ب) أن تمثل هذه المدارس جميع المستويات الاجتماعية والاقتصادية بحيث تمثل العينة الطبقات العليا والوسطى والدنيا .

(ج) أن تمثل العينة المدارس الخاصة بالبنين والمدارس الخاصة بالبنات والمدارس المشتركة .

رابعاً : فحص سجلات التلاميذ : لتبين تاريخهم التعليمى : وقد ركزنا على معرفة عدد مرات رسوب كل منهم ، وفى أى عمر رسب ، وما الفرق بين نسب الرسوب فى كل عمر إلى مجموع الراسبين ، وهل هناك فرق فى نسب الرسوب بين الذكور والإناث ، وهل هناك فرق فى نسب الرسوب بين المناطق التعليمية المختلفة . كما كنا نجمع عدد مرات الرسوب على الفرد فى الفرق الدراسية حتى الرابعة الابتدائية سواء فى ذلك من رسب مرة أو مرات ونقارنها بمن لم يرسب إطلاقاً . وأيضاً قارننا بين من لم يرسب ومن رسب مرة واحدة فى أى فرقة وبين من رسب أكثر من مرة طول تاريخه التعليمى :

خامساً : إجراء الاختبارات : أعدت كراسة تعاليمات مفصلة عن طرق إجراء الاختبارات ، وتصحيحها ، وطريقة ملء البيانات فى بطاقة التلميذ .

وأرسلت كراسات التعليمات والاختبارات والبطاقات إلى مفتشى المناطق ليتولوا توزيعها بأنفسهم على نظار المدارس التي اختيرت ضمن عينة هذا البحث . وكان لكل اختبار كراسة تعليمات منفصلة ، كما كان كل اختبار على حدة موضوعاً في كتيب منفصل .

سادساً : وصف العينة : تكونت العينة ، كما سبق القول ، من ١٤٢٩ تلميذاً يمثلون ٢٠ مدرسة تمثل ٩ مناطق تعليمية . ولم يتمكن الباحث من تحديد جنس ١٠ أفراد في العينة وكان الباقيون : ٩٨٥ تلميذاً ، ٤٣٤ تلميذة وكانت أعمار ٦٩٪ من البنين تتراوح بين ٩ ، ١١ سنة بينما كانت أعمار ٧٩٪ من البنات تتراوح بين ٨ ، ٩ سنوات . وكانت النسبة المئوية لمن تزيد أعمارهم عن ١٣ سنة في العينة الكلية ٥٪ .

سابعاً : النتائج : نقتصر هنا على تقديم وصف إجمالي لنتائج البحث فيما يتعلق بالمستوى التحصيلي . وقد جمعت عن كل تلميذ معلومات عن تقديرات المدرسين له في اللغة العربية ، والحساب ، والذكاء العام . ولكي نتأكد من فهم المدرسين لما نقصده بالذكاء العام عقدت ندوات اشترك فيها ممثلون لهم كما اشترك فيها مفتشو الأقسام وناقش الباحث معهم هذا المفهوم وكيف يوضع تقرير للتلميذ عنه . كما طبع كتيب منفصل فيه تحديد للمفهوم وبيان لطريقة تقديره بتقدير أقرب إلى الدقة وأرسلت الكتيبات لهم . وبرغم هذا لم نأخذ بتقديرات المدرسين للذكاء إذ أنها لم تتبع تعليماتنا كما لم تكن موضوعاً في مستويات متميزة .

كما حصلنا على درجات أفراد العينة في اختبارات البحث ، في اللغة العربية ، والحساب ، والمعلومات العامة ، والذكاء العام . وقد تقدم بعض أفراد العينة لامتحان القبول بالإعدادية في كل من اللغة العربية والحساب وجمعنا معلومات عن درجاتهم في هاتين المادتين . وهكذا ، في تحليل

النتائج ، قسمنا العينة إلى مجموعتين إحداهما تقدمت لامتحان القبول والأخرى لم تتقدم في حين جمعنا بطاقات المعلومات عن تلاميذ كل من القسمين وكذلك درجاتهم في اختبارات البحث .

ونظراً للاختلاف بين النهايات العظمى لتقديرات المدرسين واختبارات البحث ودرجات امتحان القبول . ولكي يتسنى لنا مقارنة هذه الدرجات حولت جميعها بحيث أصبحت النهاية العظمى لكل منها ١٠٠ . ولما كانت تقديرات المدرسين في اللغة العربية توضع من درجة عظمى هي ٤٠ فقد ضربت في ٢.٥ . ولما كانت درجات اختبارات البحث توضع من نهاية عظمى هي ٨٠ : فإنها كانت تضرب في ١.٢٥ : وكذلك ضربت درجات الطلبة في امتحان القبول للمدارس الإعدادية في اللغة العربية في ٢ إذ أن درجتها العظمى كانت ٥٠ . وبنفس الطريقة استعملت معاملات لتحويل باقي الدرجات بحيث تصبح نهاياتها العظمى ١٠٠ وبذلك أمكن مقارنة كل الدرجات معاً :

ثامناً : تبويب الدرجات : بوبت درجات العينة حسب تصنيفها إلى :

١ - الفئة التي حسبت لها درجات عن اختبارات البحث كما حصلنا على تقديرات المدرسين لأفرادها .

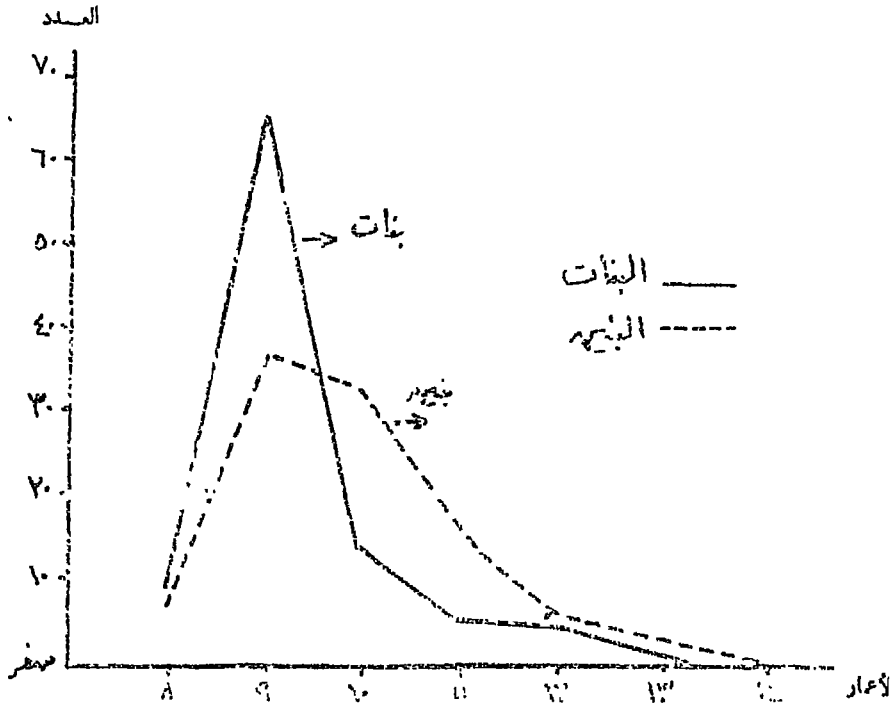
٢ - الفئة التي حسبت لها درجات عن اختبارات البحث ورصدت درجات أفرادها في امتحان القبول للإعدادية ، إذ أن كل أفرادها تقدموا له ؛

٣ - الفئة التي لم يرسب أفرادها أبداً في سنى الدراسة الابتدائية ٥ وكان بعضهم ينتمى إلى الفئة الأولى والبعض الآخر ينتمى إلى الفئة الثانية :

٤ - الفئة التي رسب أفرادها مرة واحدة أو أكثر خلال سنى الدراسة الابتدائية . وقد انتمى بعض أفرادها إلى الفئة الأولى ، وانتمى البعض الآخر إلى الفئة الثانية .

تاسعاً : نتائج البحث :

- ١ - ارتفعت تقديرات المدرسين ، في أغلب المدارس ، لأغلب التلاميذ ، كثيراً عن درجات هؤلاء التلاميذ في كل من اختبارات البحث ودرجات امتحان القبول للإعدادية ؛ وقد تراوحت التقديرات بين ٦٠ ، ٧٠ في أغلب الحالات ، بينما تراوحت متوسطات الدرجات في اختبارات البحث و امتحان القبول للإعدادية بين ٤٠ ، ٦٠ . ومن هذا يتضح لنا أن المدرسين ، في جميع المناطق التعليمية ، بالغوا في تقدير تحصيل تلاميذهم .
- ٢ - كان مستوى التلاميذ عموماً في اللغة العربية ، سواء اعتبرنا اختبارات البحث أو نظرنا إلى درجات امتحان القبول في هذه المادة ، أعلى من درجاتهم في الحساب .
- ٣ - كان هناك اتفاق كبير بين درجات التلاميذ في اختبارات البحث و امتحان القبول للمرحلة الإعدادية في كل اللغة العربية والحساب .
- ٤ - كانت المتوسطات في اختبار المعلومات العامة أعلى من المتوسطات في اختبار الذكاء العام . ولكن الفروق غير ذات دلالة .
- ٥ - فاق تحصيل البنات في اختبار المعلومات العامة تحصيل البنين فيه بدرجة كبيرة .
- ٦ - ظهر أن متوسطات الدرجات تنخفض بتقدم أعمار التلاميذ باستمرار . وكان اطراد الفروق واضحاً ، وإن لم تكن هذه الفروق نفسها كبيرة وخاصة في اختباري المعلومات العامة والذكاء العام .
- ٧ - متوسط أعمار البنات كان أقل كثيراً من متوسط أعمار البنين .
- ٨ - نسبة الناجحين في امتحانات النقل انخفضت بتقدم أعمارهم . وقد اتضح هذا ، على وجه أخص ، فيما يتعلق بمن زادت أعمارهم عن ١٢ سنة إذ قات ، نسبة نجاحهم في المتوسط عن ٥٠٪ .



(الشكل ٣٠) توزيع أعمار البنين والبنات

من الشكل يتضح أن أعداداً أكبر من البنات في الأعمار الصغيرة بينما نجد أن أعداداً أكبر من البنين في الأعمار الأكبر

عاشراً : التفسير :

لن يكون تفسيرنا هنا شاملاً ، إذ لا يتسع المجال لهذا ، وسنقتصر على ما يمكن أن يعد فروضاً توضع تحت البحث . وسنناقش النتائج مفسرين إياها حتى نخلص إلى الإجابة عن المشكلة الأولى التي دفعت إلى القيام بهذا البحث وهي معرفة ما إذا كان مستوى تحصيل وإعداد تلاميذ المناطق البعيدة عن المدن الكبرى أقل أم أكثر من تحصيل تلاميذ المدارس في المناطق القريبة من المدن الكبرى أو المدن الكبرى نفسها . ونرى :

١ - أن تقديرات المدرسين تزيد كثيراً عن درجات اختبارات البحث ودرجات امتحان القبول لأن المدرسين في هذا المستوى يفتقرون إلى معرفة

طرق التقدير الدقيقة الموضوعية . كما أنهم ، شعورياً أو لا شعورياً ، يرون أن تقدير أنفسهم وكفائتهم في التدريس يعتمد على تقدير تلاميذهم الذي إذا ارتفع رفع معه تقدير الآخرين للمدرسين وتقديرهم لأنفسهم .

٢ - ارتفعت درجات التلاميذ في الحساب عنها في اللغة العربية لأسباب نقترح أنها قد ترجع إلى عوامل معينة ، ولكن هذا الاقتراح لا يقوم إلا بدراسة لتحقيقه نرجو أن يحفز هذا البحث البعض لتعهدوا . ومن بين هذه العوامل أن اللغة المألوفة المستخدمة في الحياة اليومية هي اللغة العامية وليست العربية كما تدرس . ومن هنا كان احتمال تمكن التلميذ من اللغة المستخدمة في الحياة اليومية وميدان العمل فيما بعد برغم ضعفه في اللغة كما تقيسها الامتحانات أو تصاغ بها الاختبارات ، احتمالاً جائزاً . فإذا أضفنا إلى هذا أن مدرسي المدرسة الابتدائية ، وخاصة وقت القيام بالبحث ، غالباً من غير المؤهلين لتدريس العربية تأهيلهم لتدريس الحساب الذي لا يكلفهم خروجاً عن لغة الحياة اليومية الأرقام . ولاندعى أن هؤلاء المدرسين معدون لمادة أو لأخرى ، وهم يدرسون كل المواد ، ولكننا نرجح أنهم درسوا أثناء إعدادهم مواد أقرب في طبيعتها إلى الحساب بل وقد يعتمد بعضها عليه . وما من شك أن تطوير تدريس اللغة في هذه المستويات الابتدائية يرفع من درجاتها ، ولكن هذه المسألة بدورها تخضع للبحث . ولا ينكر أحد أن التلميذ يتعلم جملاً وتراكيباً لغوية لا ترتبط بواقع الاستخدام اليومي ارتباط تدريس الحساب بها ، ونفوق أن استخدام اللغة العامية مع صور مأخوذة من واقع بيئتنا سيغير الأمور إلى أفضل ، والجزء بقول كهذا غير مشروع دون أن تأيده بحوث تأمل القيام بها .

٣ - اتفاق درجات الاختبارات مع درجات امتحان القبول للمرحلة الإعدادية يؤكد لصدق هذه الاختبارات . وقد يرى البعض أنها ما اتفقت

مع الامتحانات إلا لأنها صيغت صياغتها ، ولكننا ندفع بأنها ما دامت اختبارات تحصيل واختبارات تطبق على التلاميذ فقط وأن البحث تربوي وجب أن تتفق موادها مع محتويات المنهج الدراسي .

٤ - قد يرجع ارتفاع درجات البنات عن درجات البنين إلى أثر البيئة فمن المعلوم أن الأسر التي تذهب بناتها إلى المدرسة أسر من مستوى ثقافي أعلى ، غالباً من الأسر التي تمتنع عن هذا .

٥ - من الطبيعي أن تنخفض درجات التلاميذ بتقدم أعمارهم وقد يرجع هذا إلى أحد الأسباب التالية :

(أ) أن كبار السن من التلاميذ غالباً ما التحقوا بالمدرسة في سن متأخر لعدم اهتمام أسرهم بالتعليم ومن ثم لا يجدون في ظروفهم ما يدفع إلا المذاكرة ويدعمها .

(ب) كبار السن من التلاميذ في السنة الرابعة الابتدائية بينهم نسبة كبيرة ممن رسبوا خلال سني تعليمهم وهذا الأمر يرجع انخفاض ذكائهم أو قلة قدرتهم على التحصيل : هذا فضلاً عن أن الرسوب ذاته ، خاصة إذا تكرر ، يثبط همهم في التحصيل .

٦ - البنات ، وخاصة بعيداً عن العاصمة ، إما أن نجد تشجيعاً للذهاب إلى المدرسة في سن مبكرة ، وإما أن نجد رفضاً لمبدأ الذهاب إلى المدرسة . ولهذا فهي إذا ذهبت إلى المدرسة كانت غالباً أصغر سناً من الولد . فإذا أضفنا إلى هذا أن البنات ، وخاصة بعيداً عن العاصمة ، عندما تكبر أعمارهم عن الثانية عشر لا يجدن تشجيعاً من الأسرة للاستمرار في المدرسة ، علمنا أن من تبقى في المدرسة غالباً ما تكون صغيرة لم تصل بعد إلى السن الذي تعتبره فيها الأسرة أثراً .

٧ - ظهر أن المدى الذي تراوحت فيه الانحرافات المعيارية لدرجات التلاميذ في كل مادة على حدة أكبر كثيراً من المدى الذي تراوحت فيه الفروق بين متوسطات هذه الدرجات . وبلغت الإحصاء ، كان الانحراف المعياري للانحرافات المعيارية لدرجات كل مادة أكبر من الانحراف المعياري لمتوسطات درجات التلاميذ في هذه المادة .

وقد حدث هذا في كل المدارس في كل المناطق القريبة منها من المدن الكبرى وكذا البعيد مما يجعلنا ننتهي إلى أن هناك تفاوتاً كبيراً في مستويات التحصيل برغم أنها قد تكون ، في المتوسط ، جيدة أو متوسطة . كما أن حدوث هذا في كل المدارس في كل المناطق الأمر الذي يجعلنا نصل إلى أنه لا فرق جوهري بين تحصيل وإعداد التلميذ في المناطق القريبة من المدن الكبرى وبينه في البعيدة عنها وهذا هو جواب مشكلة البحث .

بحث تقييم إمتحانات القبول بالمدارس الإعدادية

قام المؤلف ببحث لتقييم إمتحانات القبول بالمدارس الإعدادية . وقد بدأ البحث سنة ١٩٥٤ وانتهى بتقرير عن نتائجه سنة ١٩٥٦ . وضممته دليلاً باختبارات القبول لهذه المرحلة وما يجب أن يتوفر فيها من خصائص تجعلها جيدة . وإليك مقتطف عن البحث ونتائجه .

دراسة تحليلية لامتحانات القبول بالمدارس الإعدادية

في مناطق القطر لسنة ١٩٥٤ - ١٩٥٥

مقدمة :

تتوزع كل منطقة من المناطق التعليمية في فصل الربيع من كل عام بعقد امتحانين تحريرين في مادتي اللغة العربية والحساب وذلك للتلاميذ المتقدمين للالتحاق بالمدارس الإعدادية الذين أتموا السنوات الأربع بالمدرسة الابتدائية (هذا ما كان متبعاً وقت القيام بالبحث) .

ويقبل بالسنة الأولى بالمدرسة الإعدادية التلاميذ الحاصلين على ٦٥ درجة أو أكثر في مجموع المادتين بشرط ألا تقل درجة كل مادة على حدة عن ٢٥ وينقل التلاميذ الحاصلون على مجموع يتراوح بين ٥٠ ، ٦٥ درجة في المادتين ، بشرط ألا تقل درجة كل مادة على حدة عن ٢٥ ، وينقل هؤلاء إلى فرقة أعلى بالمدرسة الابتدائية (السنة الخامسة) بعد أن يؤدوا امتحانات باقى المواد الأخرى بالمدرسة . أما التلاميذ الحاصلون على أقل من ٥٠ درجة فى مجموع المادتين أو على أقل من ٢٥ درجة فى أية واحدة منهما فيعقد لهم امتحان فى مدارسهم مع بقية التلاميذ الذين لم يتقدموا لهذا الامتحان :

ويتولى المفتشون فى كل منطقة تعليمية وضع امتحانات القبول بالمدرسة الإعدادية والإشراف عليها . ونظراً إلى تباين تقدير المفتشين للمستوى الدراسى ، من ناحية ، وإلى تباين القيم الفنية للامتحانات نفسها ، من ناحية أخرى ، يتضح لنا أن توزيع التلاميذ على الفرق لا يخلو من بعض العيوب . فمن الجائز أن بعض التلاميذ قد نقلوا خطأ إلى فرق أعلى بينما حرم من ذلك تلاميذ آخرون ممن هم جديرون بالانتقال . وقد تكون مثل هذه الأخطاء ليست ذات بال بالنسبة إلى موظفى المدرسة ولكنها تعد بالنسبة إلى التلميذ من الخطورة بمكان . أضف إلى هذا أنه إذا كان ينبغى أن يستند توزيع التلاميذ . على نتيجة الامتحانات فإن القياس العلمى يتطلب أن تراعى الدقة فى هذا التوزيع بقدر الإمكان .

لذلك اهتم الباحث بهذه الامتحانات وطريقة إعدادها . فأعد دليلاً باختبارات القبول بالمرحلة الإعدادية وضمته توضيحاً للأغراض التى ينبغى أن يحققها هذا الامتحان والنواحي التى ينبغى مراعاتها عند وضع الامتحان وطريقة التصحيح وتوزيع الدرجات الخ . : ويتضمن هذا الكتيب أيضاً امتحانين كاملين فى اللغة العربية وامتحانين كاملين فى الحساب ، كأمثلة يهتدى بها عند وضع امتحانات المناطق . كذلك شمل الكتيب طريقة توزيع الدرجات

عن كل سؤال وعن أجزاء السؤال الواحد . وقد وزع هذا الكتيب في الوقت المناسب على القائمين بإعداد امتحانات المناطق وعقد لهم اجتماع عام بالوزارة نوقشت فيه جميع النواحي التي ينبغي مراعاتها عند وضع امتحانات القبول للمدارس الإعدادية .

واستكمالاً لما سبق قام الباحث بالدراسة الحالية التي تهدف إلى الأغراض التالية :

- (أ) تحديد مدى الاختلافات الكيفية ، إن وجدت ، بين امتحانات القبول بالمدرسة الإعدادية بوجه عام من منطقة إلى أخرى .
- (ب) تحديد مدى الاختلافات الكيفية بين مفردات الامتحان الواحد .
- (ج) تحديد الفروق الحقيقية ، إن وجدت ، بين مجموع الدرجات الكمية للتلاميذ من منطقة إلى أخرى ، في كل اختبار على حدة وفي كلا الاختبارين معاً .

منهج عام

رغبة في تقييم الامتحانات من الناحية الكيفية ، طلب إلى كل منطقة أن ترسل نسخة من أسئلة الامتحان في كل من مادتي اللغة العربية والحساب لعام ١٩٥٥ واختير عدد من المحكمين لتقييم هذه الامتحانات . وقد وصلت إلينا نماذج هذه الامتحانات من المناطق الخمسة عشر التالية :

أسوان	الفيوم	طنطا وكفر الشيخ
قنا	الجيزة	دمهور
سوهاج	القاهرة	الإسكندرية
أسيوط	شبين الكوم	المنصورة
بنى سويف	الزقازيق	القناطر

واختارت وزارة التربية والتعليم بعض الأساتذة المختصين لتقويم امتحانات مادة الحساب، كما اختير لتقويم امتحانات اللغة العربية بعض الأساتذة المتخصصين . وقد طلب إلى كل واحد من المحكمين أن يرتب مضمون الاختبارات في كل منطقة طبقاً للمعايير الأربعة الآتية :-

١ - هل السؤال صالح من حيث أنه يقيس عناصر أساسية في مادة الامتحان ؟

٢ - هل السؤال واضح ومحدد ؟

٣ - هل السؤال محدود بحيث أن التلميذ يستطيع الإجابة عنه إجابة وافية وقت مناسب ؟

٤ - هل السؤال مناسب للمستوى المفروض في تلاميذ السنة الرابعة ؟
وقد أعطى كل واحد من المحكمين بطاقة يمكنه أن يسجل بها تقديرًا مبدئيًا لكل بند من بنود امتحانات المنطقة .

ويبدأ هذا التقدير من ١ - ٩ وفقاً لفتاح التقويم التالي :-

(١ و ٢ و ٣) ضعيف

(٤ و ٥ و ٦) متوسط

(٧ و ٨ و ٩) جيد

وليك نموذج من استمارة التقويم هذه :

اسم الفاحص

اسم الفاحص

المنطقة

مادة الامتحان

رقم السؤال										
المجموع	١٠	٩	٨	٧	٦	٥	٤	٣	٢	١
مفتاح التقييم : ضعيف ١ و ٢ و ٣ - متوسط ٤ و ٥ و ٦ - جيد ٧ و ٨ و ٩										
١ - حل السؤال صالح من حيث أنه يقيس عناصر أساسية في مادة الامتحان .										
٢ - حل السؤال واضح ومحدد										
٣ - حل السؤال محدود بحيث أن التلميذ يستطيع الإجابة عنه إجابة وافية في وقت مناسب .										
٤ - حل السؤال مناسب للمستوى المفروض في تلاميذ السنة الرابعة										
(١) ما هو تقديرك العام للاختبار في مجموعه من حيث ملائمته لمستوى التقييم كما في مفتاح التقييم)										
(ب) اكتب في الفراغ التالى ما تراه من ملاحظات على الامتحان في ضوء النقط السابقة وغيرها مما تعتقد أن له أهمية بالنسبة للفرص من الامتحان .										

(جدول ١٧)

وبعد انتهاء المحكمين من تقدير مفردات الاختبار ، قام كل منهم بتقدير الاختبار ككل تقديراً عاماً مستخدماً في ذلك مفتاح التقييم السابق .

وقد عمل التحليل الكمي للدرجات التلاميذ في هذه الاختبارات على عينة تبلغ ١٠ في المائة من التلاميذ الذين أدوا الاختبار في كل منطقة .

وعمل توزيع تكرارى للمجموع الكلى للدرجات في الاختبار ، وللدرجات الاختبار في مادة الحساب ودرجات الاختبار في اللغة العربية لكل منطقة :

تقويم الاختبارات تقويمياً موضوعياً

أخذ متوسط التقديرات التي وضعها المحكمون (وهم خمسة محكمين لاختبار الحساب ، وسبعة لاختبار اللغة العربية) وذلك للحصول على تقدير مركب لكل بند من بنود الاختبار في كل منطقة . ثم حسب متوسط هذه التقديرات المركبة للحصول على تقدير جزئي لكل اختبار « بندا بندا » ثم حسب متوسط الانحراف للتقديرات الجزئية هذه للحصول على تقدير مدى تغير المفردات في كل اختبار . وأخذ أيضاً متوسط التقديرات التي وضعها المحكمون للاختبار ككل للحصول على تقدير عام لكل اختبار . ثم حسب الانحراف في هذه التقديرات لمعرفة مدى تغير تقديرات المحكمين في كل اختبار .

ويبين لنا الجدولان (١٨ و ١٩) التقديرات الجزئية « بندا بندا » والتقديرات الشاملة للامتحان ، والانحرافات المتوسطة ، والمتوسط والانحراف المعياري في كل المناطق لاختباري اللغة العربية والحساب ، وكان معامل الارتباط بين التقديرات الجزئية (بندا بندا) والتقديرات الشاملة هو ٠,٩٥ لاختبار الحساب و ٠,٨١ لاختبار اللغة العربية . وهذان المعاملان كبيران بحيث يمكن اعتبار أن هذه التقديرات ربما بلغت نفس الدقة التي تسمح بها عمليات التقدير المبينة على أسس موضوعية .

وبين لنا الرسم البياني رقم (٣١) المتوسط والانحراف المتوسط ومدى التقديرات الذاتية لكل منطقة في هذين الاختبارين ويتضح من هذا الرسم أن هناك تبايناً واسعاً بين تقييم المفردات في بعض المناطق بينما تتقارب كثيراً في مناطق أخرى .

ولعمل مقارنة عددية بين اختبارات المناطق حسب الدرجات التائية (T-scores) للتقديرات الجزئية (بنداً بنداً) وللتقديرات الشاملة لكل منطقة :
ويوجد متوسط هذه التقديرات في العامود الأخير الأيمن بالجدولين رقم (١٨ و ١٩)
وفي الرسم البياني رقم (٣١) .

وبين الرسم البياني رقم (٣١) أنه ينبغي مراجعة وتحسين الاختبارات في مادة الحساب للمناطق الأخيرة الثلاث ولمادة اللغة العربية في المنطقتين الأخيرتين على الأقل . كما هو الرسم أن المحكمين يرون أن اختبار اللغة العربية كان جيداً بنوع خاص في اثنتين من المناطق :

ولما كانت امتحانات القبول بالمدارس الإعدادية تلعب هذا الدور الهام في برنامج التربية ولما كانت هذه الدراسة قد أوضحت تبايناً كبيراً في هذه الاختبارات ، لذا يتضح وجوب العمل على تهيئة برنامج تدريبي في وضع الامتحانات لكل من توكل إليهم هذه المهمة من المدرسين وغيرهم حتى يكون هناك تناسق وتكافؤ بين امتحانات المناطق المختلفة هـ

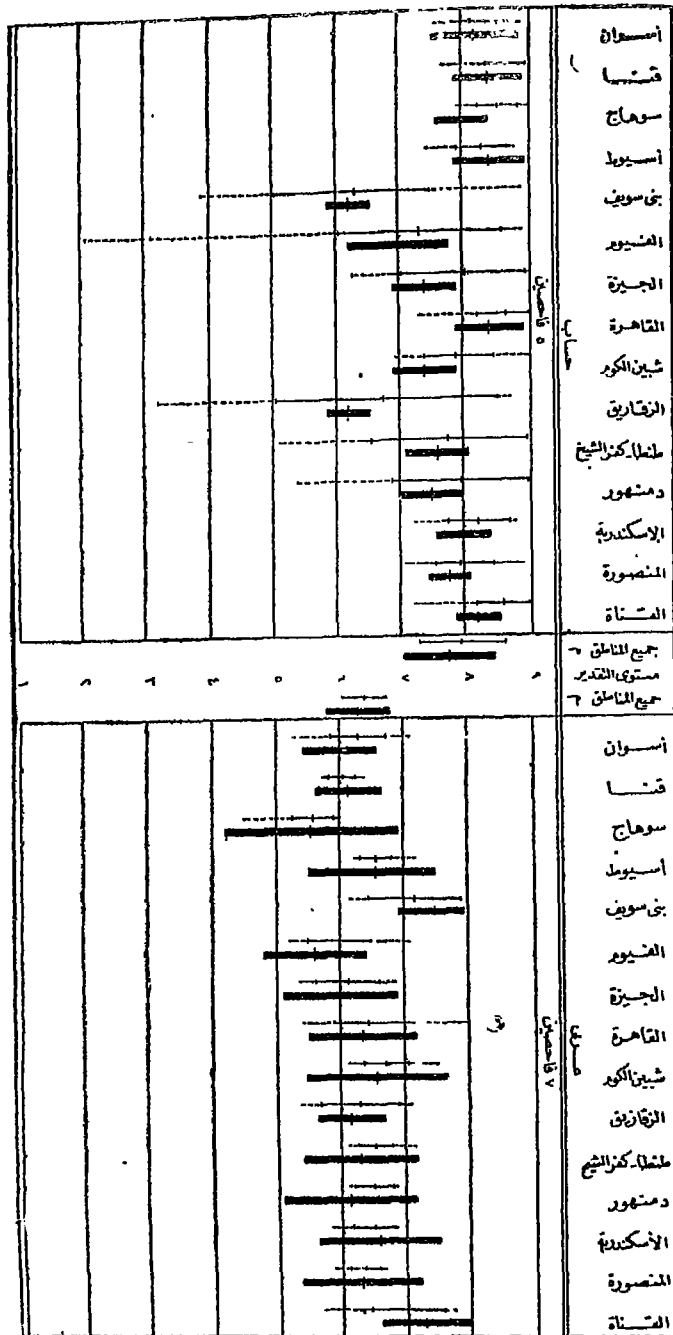
وإليك بعض نتائج البحث موضحة في جدولين ورسم بياني :

جدول رقم (١٨)
ملخص التقديرات الموضوعية لاختبارات الحساب

الدرجات النائية	التقدير الشامل للاختبار ككل	التقدير الجزئي (بنداً فبند) لكل سؤال في الاختبار	المنطقة
(T. Scores)	الانحراف المتوسط	الانحراف المتوسط	المتوسط
٥٥	٠.٦٤	٨.٢٠	٠.٤٠
٦١	٠.٤٨	٨.٤٠	٠.٣٠
٥٧	٠.٤٠	٨	٠.٣١
٥٨	٠.٤٦	٨.٤٠	٠.٣٥
٢٧.٥	٠.٣٢	٦.٢٠	١.١٥
٤٠.٥	٠.٨٠	٧	١.٢٦
٤٩	٠.٤٨	٧.٤٠	٠.٨٤
٦١	٠.٤٨	٨.٤٠	٠.٤١
٤٨	٠.٤٨	٧.٤٠	٠.٥٥
٣٠.٥	٠.٣٢	٦.٢٠	١.٧٣
٤٨.٥	٠.٤٨	٧.٦٠	١.٢٠
٤٩	٠.٥٠	٧.٥٠	١.٠٤
٥٤.٥	٠.٤٠	٨	٠.٤٦
٥١	٠.٣٢	٧.٨٠	٠.٤٧
٥٩	٠.٣٢	٨.٣٠	٠.٣٩
جميع المناطق			
٧.٩٤ (S. D.) ٠.٦٦ ٧.٦٥ (S. D.) ٠.٧٠			
R. = ٠.٦٥٤ -			

الجدول رقم (١٩)
ملخص التقديرات الموضوعية لاختبارات اللغة العربية

الدرجات النائية	التقدير الشامل (للاختبار ككل)		التقدير الجزئي بنداً فبند		المنطقة
	(T.scores)	الانحراف المتوسط	المتوسط	الانحراف المتوسط	المتوسط
٤٥٢	٠٥٧	٦٠٠	٠٤٣	٦٣٠	أسوان
٤٣٨	٠٤٩	٦١٤	٠١٩	٦٠٦	قنا
٣١٦	١٣٥	٥٥٧	٠٣٣	٦٦١	سوهاج
٤٥٥	٠٩٤	٦٥٧	٠٢٢	٦٥٧	أسيوط
٧١	٠٤٩	٧٤٣	٠٧٢	٧١٤	بنى سويف
٣٧	٠٧٨	٥٥٧	٠٥٢	٦	الفيوم
٤٢٨	٠٨٦	٦	٠٥١	٦١٢	الجيزة
٥٢٧	٠٨١	٦٣٤	٠٥٦	٦٤٣	القاهرة
٥٦٣	١٠٦	٦٥٧	٠٣٣	٦٧٠	شبين الكوم
٤٦٢	٠٤٩	٦١٤	٠٦٠	٦٢٧	الزقازيق
٤٨٤	٠٨٨	٦٢٩	٠٢٧	٦٥٦	طنطا - كفر الشيخ
٤٨	١٠٢	٦١٤	٠٣٣	٦٤٨	دمهور
٥٣٥	٠٩٤	٦٥٧	٠٣٤	٦٥٠	الاسكندرية
٤٩	٠٩٠	٦٢٩	٠١٧	٦٨٠	المنصورة
٦٧٥	٠٧٦	٧٢٩	٠٥٧	٧	القناطر
جميع المناطق					
(S. D.) ٠٥٠		٦٣٣	(S. D.) ٠٣٧	٦٤١	R = ٠٨٠٦



(شكل ٣١)

تقدير المحكمين لامتحانات القبول بالمدرسة الإعدادية ويوضح انخفاض مستوى العري عموماً عن الحساب ؛ وإن كانت الأحكام متقاربة عن العري أكثر منها عن الحساب إذ نجد أنهم يضعون امتحان الزقازيق في الحساب في الرتبة السادسة بينما يضعون امتحان القاهرة مثلاً في التاسعة . والخط الثقيل يمثل الانحراف المعياري للتقدير العام للاختبار ككل ، والنقطة مدهاء ، والنقط المنصرفة هي المتوسطات . والخط الرفيع يمثل تقدير الاختبار من بنوده

دليل الاختبارات

للقبول بالمرحلة الإعدادية

أعد المؤلف هذا الدليل سنة ١٩٥٥ وأرفق به نماذج لامتحان القبول المقترحة والخصائص التي يجب أن تتوفر فيها . وقد حذت وزارة التربية والتعليم حذوه وأخرجت أدلة مماثلة متتابعة كل عام . ونحن نقطف لك منه ما يلي :

مقدمة عامة

تنص المادة رقم ٣ من القانون رقم ٢١١ لسنة ١٩٥٣ بشأن تنظيم التعليم الثانوى على أنه يشترط فيمن يقبل بالسنة الأولى من المرحلة الإعدادية أن يؤدي بنجاح امتحان القبول الذي يعقد لهذا الغرض فى مادتي اللغة العربية والحساب .

ومن الواضح أن الغرض من هذا الامتحان هو انتقاء التلاميذ ذوى الاستعداد المناسب لمتابعة الدراسة فى مراحل التعليم الإعدادى . ولعل أصلح الطرق للانتقاء - لو أمكن الاطمئنان لتوافر الدقة والموضوعية فيها - هى طريقة الاعتماد على الأحكام التى تستخلص من تتبع أحوال التلاميذ طول مدة الدراسة وتسجيل ملاحظات المدرسين عنها . . . ولكن الاطمئنان إلى هذه الأحكام لا يتوافر قبل مضي زمن تتحسن فيه أحوال المدارس وتتهيأ للمدرسين أسباب الدراسة الشخصية لكل تلميذ ويكتمل تدريبهم عليه . ولذلك جعل الاختيار فى الوقت الحاضر على أساس امتحان القبول يدخله كل من أنس فى نفسه حسن الاستعداد له من تلاميذ المدارس الابتدائية مع وضع حدود معقولة للسن حتى يصبح الامتحان سابقة عادلة تتساوى فيها الفرص للحصول على الأماكن الموجودة .

ومهما قيل عن عيوب الامتحان فليس لدينا في الوقت الحاضر وسيلة عامة أصلح منها للاختيار وهي لاتزال الوسيلة العملية المعترف بها في العالم أجمع لاختيار الطلاب في مثل هذه الظروف على أن يراعى في إعداد أسئلة الامتحانات وطريقة إجرائها وتصحيحها ما يضمن حسن الاستفادة من نتائجها .

ولاشك أن من الخير أن يوجه التلميذ إلى نوع المدرسة المناسبة لاستعداداته ومن الظلم أن نقحم طفلاً إلى نوع من التعليم لا يستطيع أن يتابعه في يسر ونجاح ، إننا بذلك نجنى على هذا الطفل ونضيع وقته وجهده ونبتعد به عن التوجيه السليم ، كما نعرضه لمحنة نفسية من الشعور بالضالة والعجز ، وفي الوقت نفسه نضيع على أولياء الأمور وعلى الدولة مالا وجهدا ونشحن المدارس بمن لا يصلح وبمن يسد الطريق في وجه من هو خير منه .

لذلك يجب أن يعد هذا الامتحان على نحو يستطاع به تمييز من يستطيع متابعة الدراسة ممن تقف بهم قدراتهم عن المتابعة .

فيجب أن يعد الامتحان بحيث يقيس القدرة على التفكير إلى جانب قياسه القدرة على التحصيل الدراسي ، أعني أن يقيس الاستعداد كما يقيس التحصيل .

ولكن يجب أن يتم كل ذلك في بعد عن التكلف والتعقيد حتى يكون الامتحان مقياساً عادلاً يقيس في غير تحييف أو إعنات من يستطيع أن يواصل الدراسة ممن يحسن - لصالحهم ولصالح الوطن - أن يتجهوا وجهة أخرى .

أولا - اللغة العربية

الغرض من الامتحان :

نسعى في هذا الامتحان إلى قياس :

١ - قدرة التلميذ على الفهم .

٢ - قدرته على التعبير .

٣ - قدرته على رسم الكلمات رسماً مطابقاً لقواعد الهجاء :

٤ - قدرته على الاستعمال اللغوى الصحيح :

٥ - قدرته على الكتابة بنخط واضح حسن :

والمقصود بقياس قدرته على الفهم - قياس سرعة القراءة ، وسرعة الفهم وصحته ، مع ما يدخل تحت ذلك من قياس القدرة على الاستنباط ، والربط ، وإصدار الأحكام على المقروء بما يناسب التلاميذ في مثل هذه السن .

كيف نعد مقياساً لاختبار هذه القدرات ؟

أولاً - قياس القدرة على الفهم :

الطريقة الأولى :

عرض موضوع مناسب ومناقشته بأسئلة تكشف عن مدى فهم التلميذ لما يقرأ ، وتبدو مناسبة الموضوع للتلاميذ في النواحي الآتية : -

(أ) الأفكار التى يشتمل عليها الموضوع

(ب) الأسلوب الذى تعرض فيه هذه الأفكار .

(ج) طول الموضوع .

أما من حيث الأفكار :

فإراءى ألا تكون أفكاراً معنوية بحتة ، كما يجب أن يدور الموضوع حول خبرات وتجارب يستطيع التلميذ أن يمثّلها ويتصورها ويحكم عليها ، ويجب أن تكون هذه الأفكار واضحة مرتبة ترتيباً مفهوماً .

وأما من حيث الأسلوب :

فيجب تحاشي الغريب من الألفاظ والعبارات ، كما يجب تحاشي الإسراف في تطويل الجمل ، واستخدام أساليب التقسيم والتأخير ، والجمل الاعتراضية ، وكل ما من شأنه أن يعجز التلميذ عن متابعة القراءة في يسر وفهم .

وأما من حيث الطول :

فيراعى أن تكون القطعة في الحدود التي تعود التلميذ أن يراها في كتب المطالعة التي بين يديه ، وتراوح عدد كلمات القطعة المناسبة لهذا الامتحان بين مئة كلمة ومئة وعشرين .

صياغة الأسئلة :

ومن حيث صياغة الأسئلة التي يناقش بها الموضوع ، وتختبر بها القدرة على الفهم فيحسن أن يكون لكل سؤال غرض ، وأن يجس كل سؤال ناحية تكشف عن جانب من الفهم وتناقش القطعة ببضعة أسئلة :

١ - بعضها مباشر ، إجابته مذكورة بنصها في الموضوع .

٢ - وبعضها غير مباشر ، تستنبط إجابته من القطعة .

٣ - وبعضها أسئلة الفرض منها قياس قدرة التلميذ على إصدار حكم على المقروء ، كالحكم على شخص أو شيء بصفة معينة ، أو قياس قدرته على فهم الموضوع فهما كلياً .

الطريقة الثانية :

لـوتكون على نحو مما يأتي :

(١) بعرض قصة قصيرة مناسبة تنقصها بضع كلمات ، ويطلب من

التلميذ أن يقرأ القصة وأن يضع في كل مكان خال الكلمة المناسبة له من

بين الكلمات المكتوبة له في أعلى القصة .

(ب) أو عرض قصة قصيرة في عبارات غير مرتبة ويطلب من التلميذ أن يرتب هذه العبارات ترتيباً يؤلف منها قصة متصلة .
وينبغي أن تكون هذه العبارات مناسبة في طولها وعددها .

الطريقة الثالثة :

ويقصد بها اختبار مدى قدرة التلميذ على فهم مدلول اللفظ ، وما يقابله أو يماثله أو قدرته على اختيار لفظ بعينه لا تتم الجملة بدونه ، كأن يطلب منه ذكر عكس معنى كلمة مذكورة له أو اختيار كلمة من بين ما يماثلها ، أو تكميل جملة باللفظ من نوع معين كأن يكون ضميراً ، أو اسماً موصولاً ، أو اسم إشارة .

ثانياً - قياس القدرة على التعبير :

والأنواع السابقة من اختبار قدرة التلميذ على الفهم تقيس كذلك قدرته على التعبير إلى جانب قياس قدرته على الفهم ، لكنه تعبير قصير ومقيد بالإجابة عما هو مطلوب منه .

أما قدرته على التعبير المنطلق فيمكن أن تقاس بمطالبتة بالكتابة في موضوع أو رسالة .

(أ) الكتابة في موضوع :

يطلب إلى التلميذ أن يتحدث في موضوع مناسب ، في حدود قدرته وخبرته ومشاهداته ويراعى أن يكون هذا الموضوع منتزعا من البيئة التي يعيش فيها :

١ - فابن القرية الذى يعيش بين الحقول ويرى الفلاح فى عمله ويشاهد الماشية والدواب ويرى عن كتب الطيور والدواجن غير ابن المدينة :

٢ - وابن المدينة غير ابن العاصمة التى تعج من حوله بالحياة ،
ووسائل المواصلات والراديو ودور السينما والحدائق العامة والآثار
والمعارض والحفلات .

٣ - وهناك الحياة المدرسية التى قد تزدحم بمظاهر النشاط من رحلات
وحفلات وألعاب وتربية طيور فى حظائر ...

من هذه الأودية وما يشبهها ومن غيرها مما يتصل بحياة التلميذ ويتناسب
مع بيئته وظروفه يمكن أن تشتق الموضوعات التى يطلب إلى التلميذ
الكتابة فيها .

كيف يُصاغ عنوان الموضوع :

ويحسن أن تترك للتلميذ حرية الانطلاق فى التعبير ، فلا يقيد بعناصر
أو تفصيلات يتضمنها رأس الموضوع مما قد يقيد انطلاقه ، إلا إذا كان ذلك
ضروريا لمساعدته على الكتابة .

وكذلك يحسن أن يترك التلميذ حرا يكتب ما شاء ، قل ذلك أو كثر ،
فلا يقيد بسطور محدودة كما جرى على ذلك بعض الممتحنين ، إذ ليست
العبرة بكثرة ما يكتبه التلميذ أو بقلته إنما العبرة بالكيف لا بالكم .

على أن تقييد التلميذ بعدد محدود من الأسطر ليس تيسيرا له إنما هو
تقييد له وتضييق عليه .

(ب) الكتابة فى رسالة :

ويمكن أن يطلب إلى التلميذ الكتابة فى موضوع مما أشير إليه سابقا ،
ولكن فى صورة رسالة كأن يطلب إلى تلميذ من أبناء القرية أن يكتب

خطابا إلى صديق له بالقاهرة يدعو إلى زيارة القرية وسيدكر له ما سيراه فيها .

ويراعى كذلك أن يترك التلميذ - في كتابة الرسالة - حرا ينطلق في تعبيره كما شاء دون أن يقيد بعناصر معينة ، أو بعدد محدود من الأسطر ، إلا إذا كان ذلك ضروريا لمساعدته على كتابة الرسالة كما قدمنا .

تقدير إجابة التلاميذ في التعبير :

التلميذ الذى يتقدم لهذا الامتحان طفل صغير السن ، محدود الثقافة ، قليل الخبرة والتجربة .

وهذه أمور يجب أن نضعها نصب أعيننا ، ونحن نتخير له موضوع التعبير ، كما ينبغي أن نتنبه إليه ونحن نراجع عمله وإجابته ، ولكن ليس معنى ذلك أن نتقبل منه الغث الثافه ، ولكن نكتفى :

١ - بأن يكون تعبيره واضحا مفهوما .

٢ - وأن تكون أفكاره ، متنوعة ، ونقصد بالتنوع هنا أن يشتمل ما يكتبه على فكرتين أو ثلاث ، ولا يقف عند فكرة واحدة غامضة .

٣ - وأن يكون تعبيره سليما في حدود طاقة مثله ، فلا تتطلب منه أن يكتب بأسلوب عربى فصيح ، بل سنراه يمزج باللغة العربية اللغة الدارجة ، ولكن يجب أن يكون نصيب العربية أغلب وأن يكون طابعها أوفى ، فأدوات النتن والاستفهام والأمر والنهى وأسماء الإشارة والموصولات وما إليها مما درسه فى المرحلة الابتدائية - يجب أن تكون عربية ولكن يتسامح معه حينما يذكر كلمة عامية مكان كلمة عربية كأن يقول « شنطة » بدلا من « حقيبة » مثلا :

ثالثاً - قياس القدرة على الاستعمال اللغوى :

الذى يعيننا هنا هو قياس قدرة التلميذ على استعمال ما درس من أدوات وأساليب ولذلك يجب على واضع الأسئلة أن يراجع بدقة المنهج الذى درسه التلاميذ فى مادة التدريب اللغوى وأن تدور أسئلته حول ما درسوه .

ويجب أن نقيس هذه القدرة فى مواقف تشبه تلك المواقف التى تستخدم فيها اللغة فى الحياة . فمن أجل ذلك يحسن أن تدور الأسئلة حول موقف فى قصة أو موضوع سهل قصير حتى لا تبدو الأسئلة فى جو صناعى متكلف .

رابعاً - قياس القدرة على الهجاء :

وتختبر هذه القدرة بإملاء موضوع أو قصة ، ويجب أن يكون هذا الموضوع من حيث أفكاره وأسلوبه ومستوى صعوبة كلماته معنى وهجاء بحيث يستطيع التلميذ أن يتابعه فى يسر .

ويجب ألا تعتمد اختيار كلمات غريبة المعنى أو تزدحم بأنواع الهمزات وألوان الصعوبات الإملائية بل ينبغى أن تكون القطعة فى مستوى ما نجده فى كتب القراءة التى بين أيدي التلاميذ وكلماتها مستوحاة من قاموسهم وفى مستوى ما مر عليهم من قواعد الهجاء فى السنوات الماضية .

أما من حيث طول القطعة فينبغى أن تراوح كلماتها بين ٦٠ و ٧٠ كلمة^٣، وتقسم القطعة حين الإملاء إلى مقاطع مناسبة ، تملى على التلميذ مقطعاً مقطعاً .

خامساً - قياس القدرة على الخط :

جرت العادة أن يخصص سؤال للخط . يطالب فيه التلاميذ بكتابة عبارة معينة ، مرة بخط النسخ ، ومرة أخرى بخط الرقعة . ونرى أن التلاميذ فى

هذه السن لا يستطيعون أن يميزوا تمييزاً دقيقاً بين خط النسخ وخط الرقعة حين يكتبون أو أن يقدروا الاختبار في مادة الخط تقديراً ، ولذلك فهم غالباً يكتبون العبارة التي جاءت لقياس مدى قدرتهم على حسن الخط بمثل الخط الذى يكتبون به الإجابة عن سائر الأسئلة .

ولذلك نرى ألا يخصص سؤال للخط ولكن يخصص له عشر درجات من المئة وتقدر درجة التلميذ على أساس ما يلاحظ من حسن خطه ووضوحه أو رداءته وغموضه في ورقة الإجابة .

ثانياً - الحساب

أولاً - تتلخص أغراض هذا الاختبار فيما يلي :

١ - المهدف التحصيلي : يهدف هذا الاختبار إلى معرفة مستوى تحصيل التلميذ في مادة الحساب فيما درسه في المرحلة الأولى ، أى مدى إتقانه في العمليات الحسابية الأساسية وفروعها المختلفة ومدى فهمه للمسائل التي تحتاج إلى خطوات مختلفة ، يزداد عدد هذه الخطوات تبعاً لازدياد صعوبة المسألة . وتختلف أنواعها تبعاً للدرجة ارتباط أجزاء المسألة بهدفها النهائي أى أن للاختبار هدف معيارى يقيس المستوى الذى وصل إليه الطفل .

٢ - المهدف التنبؤي : ويهدف الاختبار أيضاً إلى معرفة المستوى العقلى للتلميذ في الحساب وفي القدرات المختلفة التي ترتبط ارتباطاً وثيقاً بالاستعدادات التي تساعد على التحصيل الرياضى للمراحل التعليمية المقبلة أى الكشف عن المواهب المختلفة في باكورة تفتحها والتي لها علاقة وثيقة بدراسة الرياضة في المرحلة الإعدادية والمراحل التي تليها .

٣ - النزعة التعليمية : ويجب أن يكون للاختبار قيمة في تحسين طرق التدريس بطريق مباشر أو غير مباشر .

ثانياً - أهم النواحي التي يقيسها الاختبار ما يلي :

١ - المهارة في إجراء العمليات الحسابية : ويقصد بها السرعة في إجراء هذه العمليات فالتلميذ قد يجرى العمليات بسرعة ولكنه يخطئ فيها أو قد يحلها حلاً صحيحاً ولكنه يأخذ وقتاً أطول مما يجب . ومهمتنا التأكد من اكتساب التلاميذ للمهارة التي تقوم على السرعة والدقة معاً في إجراء العمليات ، وهذه المهارة من الأحداث الهامة لتدريس الحساب لأنها الأساس الذي يقوم عليه حل المسائل وتقوم عليه دراسة الرياضيات فيما بعد . وتحتوي العمليات على الأعداد الصحيحة ، والكسور الاعتيادية والكسور العشرية . ولا شك في أن تحديد زمن الاختبار يعطى فكرة عن المهارة في إجراء تلك العمليات .

٢ - معرفة العلاقات الخاصة بالمقاييس والموازن وغيرها : إن التلميذ لا يستطيع أن يحل مسألة على المقاييس أو الموازين أو المكايل أو غيرها إلا إذا فهم المسألة أولاً وعرف العلاقة بين الوحدات المختلفة في المسألة ومعرفة هذه الحقائق جزء من تحصيله في الحساب ولذلك وجب علينا أن نقيس مدى معرفته بهذه العلاقات .

٣ - القدرة على الاستنتاج : ليست دراسة الحساب مجرد اكتساب مهارة في إجراء العمليات الأصلية لكنها تعمل على زيادة تدريب التلاميذ على الاستنتاج وهذا يتصل بالذكاء العام ولذلك فإن قياسها يعطينا فكرة عن مدى صلاحية التلميذ لمتابعة نوع من الدراسة التالية كما يميز بين التلاميذ في قدرتهم على الاستنتاج .

٤ - القدرة على الفهم : يلعب الدور الرئيسي في كل خطوة من خطوات التدريس سواء أكان في دراسة الحقائق أو غيرها أو في حل المسائل وإجراء

العمليات ولذلك يجب التمييز بين التلاميذ في قدرتهم على الفهم . ومن أمثلة ذلك فهم النظام العددي بإدراك العلاقة بين الأعداد ومعرفة الخانات التي تشغلها الأرقام في العدد ، وفهم معنى الكسور بأنواعها وفهم معنى الاصطلاحات الحسابية .

٥ - القدرة على حل المسائل : الملاحظ في دراسة الحساب بوجه خاص أن التلاميذ قد يستطيعون إجراء العمليات الأصلية ومع ذلك يتعثرون في المسائل ويرجع ذلك إلى النظر إلى الحساب على أنه علم آلى يهتم بالمهارة في إجراء العمليات فقط وتأتي المسائل كالتطبيقات على هذه العمليات وهنا يجب التمييز بين التلاميذ في قدرتهم على الفهم والاستنتاج وحل المسائل واستخدام المهارة في هذا الحل . وقد سبق التحدث عن الفهم والاستنتاج ، أما القدرة على حل المسائل فإنها تقاس بإعطاء مسائل ليحلها التلاميذ .

ثالثاً - طريقة عمل الاختبار :

إن طريقة عمل اختبار هي التطبيق المباشر لما سبق التحدث عنه وهي مرحلة في غاية من الأهمية وغاية الخطورة ولذلك يجب على واضع الاختبار أن يكون يقظاً وأن يختار العناصر بغاية الدقة لأن المسألة ليست مجرد إعطاء عناصر إنما يجب أن يكون لكل عنصر غرض واضح وأن يكون مستواه من حيث الصعوبة ومن حيث التعقيد مناسباً لمستوى التلاميذ من حيث تحصيلهم الدراسي ونموهم العقلي ، ولهذا يجب أن نسترشد بالنقط التالية في عمل الاختبار .

١ - كثرة الأسئلة وتدرجها : يراعى أن تكون الأسئلة كثيرة وقصيرة ومتدرجة في الصعوبة والغرض من كثرة الأسئلة أن تتنوع بحيث يشمل الاختبار النواحي السابقة التي يجب أن يقيسها الاختبار فلا يقتصر على حل المسائل فقط ، وإنما يشمل أشياء ذكرت فيما سبق كالمهارة ومعرفة العلاقات . د الخ ،

ولا ننسى أن يكون طول الاختبار مناسباً للزمن المحدد له والمستوى التلاميذ من حيث التحصيل والنمو العقلي ولهذا فإن المهارة في إجراء العمليات على الأعداد الصحيحة مثلاً تتطلب إعطاء أمثلة على الجمع والطرح والضرب والقسمة . وإذا طبقنا ذلك على الكسور الاعتيادية والكسور العشرية لكان عدد عناصر الاختبار كثيراً جداً ولذلك يحسن أن نكتفي بمثلين فقط في الأعداد الصحيحة فنأخذ مثلاً الجمع والقسمة في حالة الأعداد الصحيحة لأن الضرب عبارة عن جمع متكرر ، والقسمة عبارة عن طرح متكرر لهذا يحسن أن نأخذ مثلاً من الجمع وآخر من القسمة أو من الطرح وثالث من الضرب ؛ وأخذ عنصرين في كل من الأعداد الصحيحة والكسور الاعتيادية والكسور العشرية يعطينا ست عناصر وهي ملائمة لقياس المهارة . وليس هذا فحسب بل يجب العناية باختيار العناصر من حيث درجة الصعوبة فإن مثلاً بسيطاً جداً على الجمع يكون عديم الفائدة أو قد يفقد الاختبار قيمته لأن التلميذ سبق أن درس ذلك في المراحل المتقدمة جداً والمفروض أن يكون قد بلغ مرحلة أعلى من هذا المستوى البسيط . وكذلك يحسن اختيار نوع الجمع في جمل فيها ثلاثة أعداد مع عدم تساوى عدد الأرقام في الأعداد ، هذا يعطى سوئالاً صعباً نوعاً ما من أسئلة الجمع .

أما في القسمة يحسن اختيار نوع ليس سهلاً كل السهولة بل فيه صعوبة نسبية . مثل وجود الضفر في خارج القسمة . والسبب في ذلك أن تناسب أسئلة الاختبار مع المستوى الذى نتوقع أن يصل إليه التلاميذ .

أما في الكسور الاعتيادية فيحسن اختيار سوئال لجمع الأعداد الكسرية مع عدم تساوى المقامات حتى لا يكون السؤال بسيطاً جداً ولا صعباً جداً فيخطئ فيه أغلب التلاميذ . وينطبق نفس العمل على الكسور العشرية فنجعل الخانات العشرية غير متساوية . . . الخ :

٢ - تنوع الأسئلة : يلاحظ دائماً التنوع في وضع الأسئلة لقياس النواحي المختلفة في المهارة فلا تقتصر على قياس عملية واحدة أو عمليتين فقط بل تشمل على العمليات الأربع بصورها . لذلك يجب ألا يكون عدد الأسئلة قليلاً بحيث نجد أن السؤال الواحد يشمل أفكاراً كثيرة تغطي المنهج كله في مسألتين أو ثلاثة ، بل يجب أن تتدرج الأسئلة في صعوبتها بحيث تشمل أغلب أبواب المنهج ملمة بالأشياء الرئيسية التي نريد قياسها . وفي نفس الوقت يؤدي هذا التدرج في الصعوبة إلى التمييز بين التلاميذ المختلفين . فالتدرج في الصعوبة مهم حتى لا يفاجأ التلاميذ بمسائل صعبة قد تسبب ارتباكهم وخطأهم مما قد ينعكس على قدرتهم في حل المسائل التالية فيجب أن نبدأ بالأسئلة التي يحتمل أن يحلوها ليكتسبوا شيئاً من الثقة في نفوسهم ويعملوا على حل الاختبار .

٣ - اتصال المسائل بالبيئة : يجب أن نراعى كل منطقة اختيار المسائل المتصلة بالحياة العملية وبالبيئة المحلية وخاصة في حفظ علاقات بين وحدات القياس فمثلاً تهتم البيئات الريفية بالمكاييل والموازين أكثر من اهتمام البيئات المدنية وهكذا .

٤ - النوع المطلوب قياسه في الأسئلة : يجب أن يشمل الاختبار على أسئلة لقياس النواحي الهامة في تدريس الرياضيات وقد سبق ذكرها . وبهنا هنا أن نوضح نوع الأسئلة وكيفية اختيارها فمثلاً لإختبار المثال الخاص بالقدرة على الاستنتاج ، نختار $(36 \div 4 = 9 + 100)$ وهي أعداد بسيطة . أما اختيار أعداد كبيرة فإنه يجعل المسألة صعبة صعبة تدخل فيها المهارة في إجراء العمليات فمثلاً يجب عدم (استخدام $65 \div 15 = 200 - 100$) أو ما يشابه ذلك ، هذا لأن الحد الأول هنا مشكلة في ذاته وكذلك الحد الثاني وليس الغرض من هذا السؤال قياس المهارة وإنما هو قياس القدرة على الاستنتاج من العلاقات لأن للمهارة أسئلة أخرى سبق التدريب عليها .

ولذلك يحسن أن تكون الأعداد في حدود الجداول المحفوظة .

وكذلك الأسئلة التي تطلب اختيار الأعداد التي تتم سلسلة من الأعداد بعد معرفة العلاقة بينها . ومن الأمثلة في هذا المجال (اشطب العدد الذي لا يتمشى مع سلسلة الأعداد ٣ ، ٦ ، ٩ ، ١١ ، ١٢ ، ١٠٠٠) مثلاً .

ومن الأسئلة ما يقيس الفهم لمعنى النظام العددي والأعداد الصحيحة لمعرفة القدرة المكانية للرقم في العدد وفي هذا المجال يمكن أن تعطى أمثلة كما يلي :

١ - اكتب أكبر عدد مكون من رقمين

٢ - اكتب أصغر عدد مكون من رقمين دون استخدام الصفر

٣ - اكتب أكبر عدد مكون من رقمين مختلفين

٤ - اكتب أصغر عدد مكون من رقمين مختلفين دون استخدام الصفر .

وهكذا في حدود معقولة تناسب التلاميذ فلا تطلب منهم مثلاً أكبر عدد مكون من خمسة أرقام مختلفة أو أربعة أرقام يكون أصغرهما في خانة كذا . . . أى يجب أن نتجنب محاولة الإيقاع بالتلميذ وإعطائه مشاكل لا داعي لها .

٥ - تقدير الدرجات :

من عوامل نجاح الاختبار في قياس مدى تحصيل التلاميذ لمادة الحساب أن توضع درجات مناسبة لكل نوع من أنواع الأسئلة الموضوعه ، فمثلاً :

العمليات الحسابية : يصبح أن ينحصر لها ٣٠٪ تقريباً من الدرجة النهائية

من ورقة الامتحان .

(أ) يحتاج معلم القراءة إلى معرفة المستوى الذى بلغه كل تلميذ من تلاميذه حتى يستطيع أن يساعد الضعفاء ، ويرشد المتوسطين والمتفوقين . وهذا الاختبار يمدد بالمعلومات الدقيقة التى تمكنه من توجيه جهوده توجيهاً رشيداً فى هذه النواحي .

(ب) يحتاج معلم القراءة إلى معرفة ما حصله تلاميذه حتى يطمئن إلى سلامة الطريقة التى اتبعها فى تعليمهم أو عدم سلامتها : وهذا الاختبار أداة دقيقة تحقق له هذا الغرض .

(ج) يحتاج معلم القراءة أو المشرف على تعليمها إلى تجرب طريقة جديدة للتعليم ، أو مادة جديدة للقراءة ، ويريد أن يعرف أثر هذه الطريقة أو المادة فى تحصيل التلاميذ . وهذا الاختبار أداة تجريبية دقيقة يستطيع أن يقيس بها هذا الأثر .

(د) يحتاج المشرفون على التعام ، سواء كانوا نظاراً أو مفتشين أو مديري مناطق ، إلى معرفة التقدم الذى أحرزته مدارسهم أو أقسامهم أو مناطقهم فى هذه الناحية الأساسية من نواحي التعليم فى المرحلة الابتدائية ، وهى ناحية القراءة : وفى هذه الحالة يلزمهم أداة موضعية تمدهم بنتائج يمكن الاعتماد عليها فى تقويم العمل وتوجيهه . وهذا الاختبار يحقق لهم ذلك .

(هـ) يحتاج الباحث فى ميدان التربية عامة ، وفى ميدان تعليم اللغة خاصة ، إلى أداة لقياس القدرة على القراءة ، سواء كان ذلك لتجريب طريقة جديدة فى التعليم ، أو للمقارنة بين عدة طرق ، أو لإدخال تنظيم معين فى المنهج ، أو لتحديد العلاقة بين قدرات دراسية مختلفة ، أو لغير ذلك من الأغراض : وهذا الاختبار أداة دقيقة تعين الباحث على تحقيق هذه الأغراض :

ثانياً : تكوين الاختبار ونواحي القراءة التي يقيسها :

ينقسم هذا الاختبار قسمين : القسم الأول ، ويتكون من ثلاثة أجزاء :
 (١) اختبار القدرة على تعرف المفردات ، (٢) اختبار القدرة على
 تعرف الجمل ، (٣) اختبار القدرة على فهم معنى الجملة ؛ القسم الثاني ،
 ويقتصر على اختبار القدرة على فهم معاني الفقرة : والنواحي الأربع التي
 يتكون منها القسمان نواح أساسية في القدرة على القراءة ؛ إلا أن الناحيتين
 الأوليين تبرز أهميتهما في الصفوف الأولى ثم تأخذ في التلاشي بعد ذلك ،
 بينما تبقى الناحيتان الأخيرتان طوال جميع مراحل التعليم .

والاختباران الأولان يشتملان على أسئلة مصورة ، ويبدأن بمواد يسهل
 على التلاميذ المتوسطين في الصف الأول أن يميزوها ، وتنتهي بمواد تتطلب
 درجة كبيرة من القدرة على التمييز . والاختباران الأخيران يشتملان على
 عدد من الجمل والفقرات ، ويبدأن بجمل وفقرات يسهل على التلميذ المتوسط
 في الصف الأول أن يفهما ، وينتهيان بجمل وفقرات يصعب فهمها على
 التلميذ المتوسط في الصف الرابع .

وهذه الاختبارات كلها اختبارات قدرة ، وليست اختبارات سرعة .
 ومعنى ذلك أنها تقيس ما يستطيع التلميذ أن يعمل في كل ناحية من هذه
 النواحي إذا أُتيح له الوقت الكافي للقراءة .

ثالثاً : حدود الاختبار :

هذا الاختبار يقيس مدى من القدرة يمتد من نهاية الصف الأول إلى
 نهاية الصف الرابع الابتدائي . وهو مناسب لقياس المتفوقين من تلاميذ الصف
 الأول ابتداء من منتصف العام الدراسي ، ثم إنه يميز تمييزاً دقيقاً بين تلاميذ

الصفين الثاني والثالث ، ولكنه لا يميز بنفس الدرجة بين تلاميذ الصف الرابع وبخاصة في أواخر العام الدراسي .

(١) تعليمات عامة :

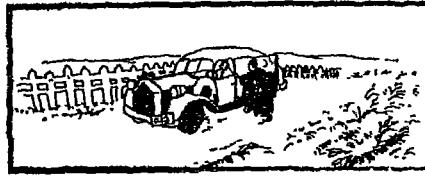
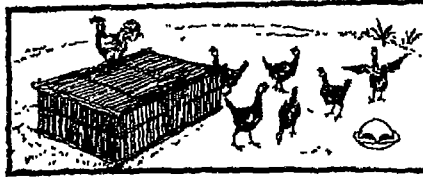
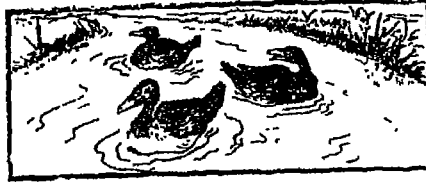
١ - هذا الاختبار مقسم إلى قسمين ، لكل منهما كراسة خاصة ، فعند إجرائه يجب تخصيص جلسة لكل قسم حتى يجد التلاميذ الوقت الكافي للإجابة ، ويبدلوا أقصى ما في وسعهم ، ولكن يجب أن تكون الجلستان متقاربتين في الزمن : ويحسن أن يجري الاختبار بقسميه في يوم واحد لضمان عدم تغيب بعض التلاميذ في القسم الثاني .

ويمكن الرجوع إلى كراسة تعليمات الاختبار لمعرفة باقيها . وإليك جزء من تعليمات الجزء الثاني من القسم الثاني .

(ب) الجزء الثاني : تعرف الجمل :

١ - ابدأ هذا القسم بأن تقول : « بص في صفحة ٩ . الصفحة دى فيها ثلاث صور ، وتحته جملة مكتوبة . اقرأ الجملة وافهمها : (انتظر قليلا حتى ينتهوا من القراءة) . بص في الثلاث صور الى فوق الكتابة ، ونقى منهم الصورة الى توافق الكلام المكتوب في الآخر ، وقول لى أنهى صورة فيهم : الأولى ، ولا الثانية ، ولا الثالثة ؟ (اترك لهم فرصة للإجابة ، ثم ناقشهم حتى يصلوا إلى أنها الصورة الأولى) . بعد ذلك قل : « وعشان كده حنط جنبها على اليمين علامة صح . (ارسم على السبورة بروازاً يماثل برواز الصورة ، ثم اكتب على يمينه علامة صح) . بعد ذلك قل : « كل واحد منكم يمسك القلم ، ويخط جنب صورة البط على اليمين علامة صح : (مر بينهم لتأكد من أنهم عملوا المطلوب » .

(الشكل ٢٢)



البط في الماء

فقرة من اختبار سرس الليان في القراءة الصامتة

تفسير نتائج الاختبار :

يعطى المفحوص درجات عن فقرات الاختبار تبعاً لمفتاح التصحيح ثم تحول هذه الدرجات إلى درجات جيمية تقابلها مستويات . وإليك الجدول رقم (٢٠) .

ثم تحسب الدرجة الثائية المقابلة للدرجة الخام باستخدام الجدول رقم ٢١ الذي نأخذ منه القطاع التالي .

(جدول رقم ٢٠) الدرجات الختام وما يقابلها من الدرجات الجسمية والمستويات

الصف الرابع			الصف الثالث			الصف الثاني			الصف الأول		
المستوى	الدرجة الجسمية	الدرجة الختام	المستوى	الدرجة الجسمية	الدرجة الختام	المستوى	الدرجة الجسمية	الدرجة الختام	المستوى	الدرجة الجسمية	الدرجة الختام
عاجز جدا	صفر	صفر - ٢١	عاجز جدا	صفر	صفر - ٩	ضعيف جدا	٢	صفر - ٥	ضعيف جدا	٢	صفر
عاجز	١	٢٩ - ٢٢	عاجز	١	١٤ - ١٠	ضعيف	٣	١٠ - ٦	ضعيف	٣	١ - ٢
ضعيف جدا	٢	٣٥ - ٣٠	ضعيف جدا	٢	١٨ - ١٥	تحت المتوسط	٤	١٥ - ١١	تحت المتوسط	٤	٥ - ٣
ضعيف	٣	٣٨ - ٣٦	ضعيف	٣	٢٣ - ١٩	متوسط	٥	٢٠ - ١٦	متوسط	٥	٨ - ٦
تحت المتوسط	٤	٤١ - ٣٩	تحت المتوسط	٤	٢٨ - ٢٤	فوق المتوسط	٦	٢٥ - ٢١	فوق المتوسط	٦	١١ - ٩
متوسط	٥	٤٣ - ٤٢	متوسط	٥	٣٣ - ٢٩	جيد	٧	٣٠ - ٢٦	جيد	٧	١٥ - ١٢
فوق المتوسط	٦	٤٥ - ٤٤	فوق المتوسط	٦	٣٧ - ٣٤	جيد جدا	٨	٣٦ - ٣١	جيد جدا	٨	١٩ - ١٦
جيد	٧	٤٧ - ٤٦	جيد	٧	٤٢ - ٣٨	ممتاز	٩	٤١ - ٣٧	ممتاز	٩	٢٥ - ٢٠
جيد جدا	٨	٥٠ - ٤٩	جيد جدا	٨	٤٧ - ٤٣	ممتاز جدا	١٠	٤٤ فأكثر	ممتاز جدا	١٠	٢٦ فأكثر
			ممتاز	٩	٥٠ - ٤٨						

جدول رقم (٢١) للدرجات الخام للاختبار والدرجات التائية المقابلة لها في كل صف من الصفوف الأربعة

الدرجة التائية				الدرجة الخام	الدرجة التائية				الدرجة الخام
الصف الرابع	الصف الثالث	الصف الثاني	الصف الأول		الصف الرابع	الصف الثالث	الصف الثاني	الصف الأول	
٣٠	٤٥	٥٨	٧٢.٥	٢٦	-	١٨	٢٣	٣٦.٥	صفر
٣٠.٥	٤٦	٥٩	٧٣.٥	٢٧	-	١٩	٢٤	٣٨.٥	١
٦٤	٦٩	٨٠	-	٤٩	٢٨	٤٢	٥٥	٧٠.٥	٢٣
٦٦.٥	٧٠	٨١	-	٥٠	٢٨.٥	٤٣	٥٦	٧١	٢٤
					٢٩.٥	٤٤	٥٧	٧٢	٢٥

ثبات الاختبار وصدقه :

حسب معامل الثبات بتنصيف الاختبار إلى نصفين وتطبيق معادلة سبيرمان - براول . وكانت معاملات الثبات كما يلي :

جدول رقم (٢٢) معاملات ثبات الاختبار في الصفوف الثلاثة الأخيرة من المرحلة الابتدائية

معاملات الثبات					الصف
المجموع	تعرف المفردات	تعرف الجمل	فهم الجملة	فهم الفترة	
٠.٩٢	٠.٧٨	٠.٥٨	٠.٧١	٠.٨٢	الثاني
٠.٩٢	٠.٧٧	٠.٨١	٠.٧٢	٠.٨٤	الثالث
٠.٩٢	٠.٧٥	٠.٨٦	٠.٨٠	٠.٨٤	الرابع

أما عن صدق الاختبار فيمكن أن يستدل عليه بارتباطه بمقياس خارجي . والمقياس الذي اختير لمعرفة ارتباطه بالاختبار هو تقدير المدرسين ، فقد أجرى الاختبار على ٥٠٠ تلميذ وتلميذة موزعين على ١٣ مدرسة ابتدائية من مدارس مديرية المنوفية ، وطلب من المعلمين في هذه المدارس تقدير درجات لهؤلاء التلاميذ والتلميذات في القدرة على القراءة . ثم حسب معامل الارتباط بين نتائج الاختبار وتقديرات المعلمين فكان ٠٫٧٥ وهو ارتباط عال نسبياً يدل على أن الاختبار يقيس فعلاً قدرة اتفق جمهور المعلمين على تسميتها بالقدرة على القراءة .

ويمكن الاستدلال على صدق الاختبار أيضاً بمعرفة العلاقة بين زيادة درجات الاختبار ونمو التلاميذ من صف إلى صف . والجدول رقم (٢١) يشتمل على الدرجات النائية للاختبار في كل صف من الصفوف الأربعة . وبمقارنة الدرجات النائية في كل صف من الصفوف بعضها ببعض ، وبمقارنة الدرجات الخام المقابلة للدرجة النائية ٥٠ في كل صف من الصفوف يتضح أن هذا الاختبار يقيس قدرة تزداد بانتظام من صف إلى صف :

الحساب

من الخطأ أن يعتقد البعض أن تدريس مادة بعينها أو منهج بذاته هدف من أهداف التربية والتعليم . والأمر على عكس ذلك تماماً لأن تدريس أى مادة أو منهج هو وسيلة أو أداة لتحقيق الأهداف العامة للتربية :

ومن أهم الأهداف العامة للتربية تقديم خدمات للفرد في حياته الخاصة . وأهمها المهارات التي تمكنه من أداء مطالب الحياة وأداء الواجبات التي يفرضها عليه المجتمع الذي يعيش فيه . والفرد الناجح في حياته هو الذي يستطيع التعامل مع الآخرين في سهولة ودقة والتعامل في المجتمع الحديث يقوم على إتقان العمليات الحسابية أى الدقة في حساباته وكذلك على السرعة في إجراء

هذه العمليات . لذلك كان اكتساب المهارة في إجراء العمليات الحسابية لتحقيق السرعة والدقة في مقدمة أهداف تدريس الحساب .

وكبرى المشكلات التي تعترض تحقيق هذه الأهداف هي أنها تحتاج إلى مران وتدريب طويلين مع العناية بناحيتي الدقة والسرعة معا .

وكثيراً ما ساد بين المشتغلين بالتربية اعتقاد خاطئ مؤداه أن هدف تعليم مادة الحساب هو القدرة على حل المسائل الحسابية المعقدة . ولذلك كثر ابتكار مشاكل ومسائل غاية في التعقيد . هذا الاعتقاد هو امتداد لما كان يراه المشتغلون بالتربية قديماً من أن المران على حل مشاكل ومسائل من هذا النوع يقوى التفكير عند الفرد ويساعده على حل المشاكل المعقدة التي تقابله في شتى ميادين الحياة .

ولكن الدراسات التربوية التجريبية الحديثة أثبتت بطلان هذا . وقد ارتكز إثباتها على الأسباب التالية :

أولاً : أنه لا يحدث انتقال لأثر التدريب من حل المشكلات الحسابية كما تدرس إلى حل ما عداها من المشكلات الأخرى .

ثانياً : أن هذا النوع من المسائل المعقدة لا يصادف الفرد في حياته العملية بهذه الصورة : على أن الهدف من تدريس الحساب لا يقتصر على مجرد المعرفة بالطرق والقواعد ، ولكن الهدف الأهم أن يكتسب الفرد مهارة في أدائها في سهولة ويسر أى تحقيق أكبر مستوى من الدقة والسرعة . وقد ثبت تجريبياً أن المران، والتدريب هو الطريق الوحيد إلى تكوين عادات الدقة والسرعة . والمران والتدريب المستمر على عمليات الحساب شاق وممل لكل من التلميذ والمدرس ، فالتلميذ يسأم التكرار ولا يجد الدافع القوي الذي يتغلب على أثر الملل والسآمة ، وكذلك المدرس لا يتسع وقته ولا ترتاح نفسه

لأن يراجع ويصحح كل ما يؤديه تلاميذ فرقة أو فصل دراسي من التدريب .

ولكن هذا التدريب يجب أن يكون في العمليات التي لا يجيد أداءها التلميذ بدرجة عالية من الدقة أو التي لا تصل سرعته فيها إلى حد مقبول . لذلك يكون من الضروري تحديد العمليات التي يحتاج فيها التلميذ إلى التدريب سواء لتحقيق الدقة أو لتحقيق السرعة . أى أنه يجب أولاً أن نقوم بتشخيص أنواع العمليات التي لا يتقنها التلميذ والتي لم يصل في أدائها إلى سرعة مناسبة .

وقد أعد الأستاذ إسماعيل القباني مع المؤلف مجموعة من البطاقات سميت (بطاقات الحساب التدريبية التشخيصية) تتكون من عدة مجموعات وتتكون كل مجموعة من عدد من البطاقات من الورق المقوى . والمجموعة الواحدة تخص نوعاً من العمليات الحسابية . والبطاقة الأولى من كل مجموعة هي اختبار تشخيصي يتناول جميع الاحتمالات الرئيسية المتعلقة بنوع العمليات التي تتناولها هذه المجموعة . وتسهيلاً للاستعمال خصص لكل مجموعة من البطاقات لون يميزها عن المجموعات الأخرى . وفيما يلي جدول يبين المجموعات والعمليات التي تتناولها وعدد البطاقات في كل منها .

وعلى أحد وجهي كل بطاقة طبعت عمليات حسابية وأمام كل منها مسافة يكتب التلميذ فيها الإجابة . وعلى الوجه الآخر طبعت الإجابات الصحيحة بحيث ننطبق تماماً على المسافات المخصصة للإجابة في الوجه الأول . (في شكل ٣٣) صورة لوجهي البطاقة ز - ٢

وقد أعدت كراسة للتدريب من الورق الشفاف ، بحيث يضع التلميذ البطاقة التي سيعمل فيها تحت الورقة الشفافة الأولى من كراسة التدريب يسجل إجاباته عليها بقلمه الحبر . وعند انتهاء الزمن المحدد للتدريب يقلب الوجه الآخر للبطاقة (الإجابة) ويدخله تحت الورقة الشفافة التي كتب

اجرة الاختار (٢-)

الاسقف

الاستاذ

الاستاذ

[illegible]

ظهر البطاقة (مفتاح التصحيح)

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ
 الْحَمْدُ لِلَّهِ الَّذِي
 (طرح)

$\frac{1}{2} = \frac{2}{4}$	$\frac{1}{2} = \frac{3}{6}$	$\frac{1}{2} = \frac{4}{8}$	$\frac{1}{2} = \frac{5}{10}$	$\frac{1}{2} = \frac{6}{12}$	$\frac{1}{2} = \frac{7}{14}$	$\frac{1}{2} = \frac{8}{16}$	$\frac{1}{2} = \frac{9}{18}$
$\frac{1}{3} = \frac{2}{6}$	$\frac{1}{3} = \frac{3}{9}$	$\frac{1}{3} = \frac{4}{12}$	$\frac{1}{3} = \frac{5}{15}$	$\frac{1}{3} = \frac{6}{18}$	$\frac{1}{3} = \frac{7}{21}$	$\frac{1}{3} = \frac{8}{24}$	$\frac{1}{3} = \frac{9}{27}$
$\frac{1}{4} = \frac{2}{8}$	$\frac{1}{4} = \frac{3}{12}$	$\frac{1}{4} = \frac{4}{16}$	$\frac{1}{4} = \frac{5}{20}$	$\frac{1}{4} = \frac{6}{24}$	$\frac{1}{4} = \frac{7}{28}$	$\frac{1}{4} = \frac{8}{32}$	$\frac{1}{4} = \frac{9}{36}$
$\frac{1}{5} = \frac{2}{10}$	$\frac{1}{5} = \frac{3}{15}$	$\frac{1}{5} = \frac{4}{20}$	$\frac{1}{5} = \frac{5}{25}$	$\frac{1}{5} = \frac{6}{30}$	$\frac{1}{5} = \frac{7}{35}$	$\frac{1}{5} = \frac{8}{40}$	$\frac{1}{5} = \frac{9}{45}$
$\frac{1}{6} = \frac{2}{12}$	$\frac{1}{6} = \frac{3}{18}$	$\frac{1}{6} = \frac{4}{24}$	$\frac{1}{6} = \frac{5}{30}$	$\frac{1}{6} = \frac{6}{36}$	$\frac{1}{6} = \frac{7}{42}$	$\frac{1}{6} = \frac{8}{48}$	$\frac{1}{6} = \frac{9}{54}$
$\frac{1}{7} = \frac{2}{14}$	$\frac{1}{7} = \frac{3}{21}$	$\frac{1}{7} = \frac{4}{28}$	$\frac{1}{7} = \frac{5}{35}$	$\frac{1}{7} = \frac{6}{42}$	$\frac{1}{7} = \frac{7}{49}$	$\frac{1}{7} = \frac{8}{56}$	$\frac{1}{7} = \frac{9}{63}$
$\frac{1}{8} = \frac{2}{16}$	$\frac{1}{8} = \frac{3}{24}$	$\frac{1}{8} = \frac{4}{32}$	$\frac{1}{8} = \frac{5}{40}$	$\frac{1}{8} = \frac{6}{48}$	$\frac{1}{8} = \frac{7}{56}$	$\frac{1}{8} = \frac{8}{64}$	$\frac{1}{8} = \frac{9}{72}$
$\frac{1}{9} = \frac{2}{18}$	$\frac{1}{9} = \frac{3}{27}$	$\frac{1}{9} = \frac{4}{36}$	$\frac{1}{9} = \frac{5}{45}$	$\frac{1}{9} = \frac{6}{54}$	$\frac{1}{9} = \frac{7}{63}$	$\frac{1}{9} = \frac{8}{72}$	$\frac{1}{9} = \frac{9}{81}$
$\frac{1}{10} = \frac{2}{20}$	$\frac{1}{10} = \frac{3}{30}$	$\frac{1}{10} = \frac{4}{40}$	$\frac{1}{10} = \frac{5}{50}$	$\frac{1}{10} = \frac{6}{60}$	$\frac{1}{10} = \frac{7}{70}$	$\frac{1}{10} = \frac{8}{80}$	$\frac{1}{10} = \frac{9}{90}$

(شکل ۳۳) وجه البطاقة (الامانات)

جدول رقم (٢٣)

المجموعة	لون المجموعة	الموضوع	عدد الاختبارات	عدد البطاقات
أ	أزرق	تغيير صورة الكسر	٥	٩
و	رمادي	جمع الكسور الاعتيادية	٦	٦
ز	أصفر	طرح الكسور الاعتيادية	٤	٤
ح	أخضر	ضرب الكسور الاعتيادية	٦	٦
ط	أحمر	الكسور العشرية	٧	٧
ي	رمادي	ضرب الأعداد العشرية وقسمتها	٥	٧
ك	أخضر	تحويل النسب المئوية إلى صور كسرية وبالعكس	٦	٦
ل	أصفر	تحويل النسب المئوية إلى صور كسرية وبالعكس	٧	٧
م	رمادي	المتوسط الحسابي	٥	٥
ن	أزرق	المقاييس المألوفة	٩	٩
س	أصفر	الربح	٧	٧
الفحص العام	أخضر	اختبار عام فيما سبق	٣	٦
		المجموع	٧٠	٧٩

محتويات بطاقة الحساب التدريبية التشخيصية

عليها حلو له فيجد الإجابات الصحيحة ظاهرة بجوار إجاباته التي سجلها . وبذلك يقوم التلميذ بنفسه بتصحيح إجاباته ثم يقوم بجمع عدد الصحيح منها ويسجل هذا العدد في المكان المخصص لذلك بنفس الصفحة . وفي نهاية كل بطاقة يوجد بيان للمستويات (ممتاز - جيد - ناجح) ليقم التلميذ نفسه على أساسها . فإذا كان ناجحاً أو جيداً أو ممتازاً ينتقل في اليوم التالي إلى البطاقة التالية . أما إذا لم يحصل على درجة النجاح كان معنى ذلك أنه لا زال محتاجاً إلى التدريب على هذا النوع من العمليات . ولذلك يستمر في التدريب على نفس البطاقة مرة أو مرات ولا ينتقل من بطاقة إلى أخرى إلا إذا نجح في الأولى .

وفي كراسة التدريب صفحات لتسجيل العمل اليومي وصفحات مبسطة لعمل رسم بياني « بالأرقام » يساعد التلميذ على معرفة مدى تحسنه أو تقدمه . وقد أعدت البطاقات جميعها بحيث تعطى زمناً واحداً ثابتاً هو ٤ دقائق للإجابة عن أى بطاقة . وعلى هذا ينحصر المدرس ٧ دقائق تقريباً من أول كل حصة حساب يمارس التلاميذ فيها جميعاً تدريبهم وتسجيلهم للنتائج . وكل يعمل بمفرده وفقاً لمستواه وسرعته . فقد نجد التلميذ الأول يعمل على الاختبار (و - ٢) من المجموعة (و) عن جمع الكسور الاعتيادية بينما يعمل التلميذ الثاني على الاختبار (ز - ٤) عن طرح الكسور الاعتيادية وهكذا .

التدريب بهذه الطريقة يبسط عملية التدريب ويختصر وقتها بالنسبة للمدرس والتلميذ على حد سواء . كما أنه يتميز بالدفع الطيب إذ يتبارى كل فرد أساساً مع نفسه لأنه يقارن تقدمه يوماً بيوم ، وتأتي منافسته مع غيره في المرتبة الثانية .

وفيما يلي وصف لطريقة استعمال هذه البطاقات كما وردت في كراسة التدريب :

طريقة استعمال للبطاقات

اتبع هذه التعليمات بدقة

١ - عندما تسمع التنبيه بالاستعداد للتدريب الحسابي ، أزل كل ما على درجك من كتب وأدوات ، ثم أخرج كراسة التدريب وقلماً ونشافة .

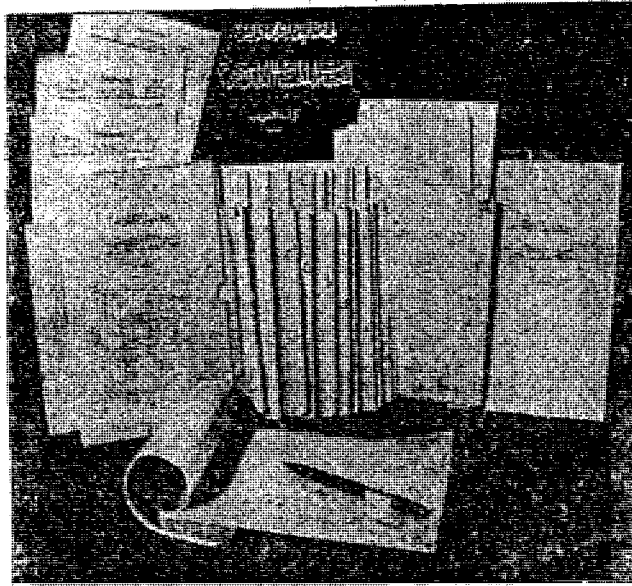
٢ - افتح كراسة التدريب وانظر في سجل العمل اليومي (من ص ٥ إلى ص ١٠) لتعرف رقم الاختبار الذي ستقوم بإجرائه :

٣ - خذ بطاقة هذا الاختبار ، ولاحظ أن كل بطاقة مكونة من وجهين ، الأول : يحتوي على الأسئلة وتجد دائماً في ركنه العلوي من اليمين مستطيلاً فيه حرف المجموعة ورقم الاختبار ، والوجه الثاني فيه الأجوبة ، ولا تحتاج إليه إلا بعد الانتهاء من الاختبار عندما يأذن لك المدرس بذلك :

٤ - ضع البطاقة تحت أول ورقة شفافة في الكراسة ، بحيث يكون الوجه الذي به الأسئلة أمامك ، وتكون الحافتان السفلى واليمنى منطقتين تماماً على حافتي الكراسة : ثم اقل الكراسة وانتظر الأمر ببدء الاختبار .

٥ - في بعض الأحيان يكون الاختبار شاغلاً وجهي البطاقة ، وفي هذه الحالة ستجد في نهاية الصفحة من أسفل كلمتي « له بقية » ، وعند الوصول إليهما اقلب الوجه الثاني ، ثم ضعه تحت الورقة الشفافة الثانية من كراستك ، واستمر في الإجابة عن الأسئلة بدون توقف . أما أجوبة هذا الاختبار فتجدها في بطاقة ثانية ، وعليك عند التصحيح أن تضع صفحة الأجوبة الأولى تحت الورقة الأولى ، والصفحة الثانية (تابع أجوبة ...) تحت الورقة الثانية .

شكل (٣٤)



مجموعة كاملة من بطاقات الحساب التدريبية التشخيصية
وأمامها كراسة التدريب مفتوحة

٦ - عندما يقول المدرس « ابتدئ » ، ابدأ في الإجابة على الاختبار ، واكتب الإجابات بالخبر في المواضيع المخصصة لها بالضبط مع جعل الأرقام واضحة ، اجتهد في أن تشتغل بسرعة ، ولكن مع مراعاة الدقة التامة حتى لا تقع في خطأ . استعمل هامش الورقة الشفافة للتسويد إذا لزم :

٧ - عندما يقول المدرس « ضع القلم » ، ارسم خطاً رأسياً بعد آخر عملية كنت تجربها ، ثم ضع القلم من يدك مباشرة .

٨ - عندما يقول المدرس « ابتدئ » في التصحيح ، اقلب البطاقة (أوخذ بطاقة الأجوبة في الاختبار الذي من صفحتين) ثم ضعها تحت الورقة الشفافة ، وبذلك تصبح الإجابات الصحيحة بجوار إجابتك (على عينيها غالباً) :

٩ - املأ البيانات الموجودة في رأس هذه الصفحة (الاسم - التاريخ - الفرقة - رقم الاختبار) .

١٠ - ضع علامة X على الإجابات الخاطئة فقط ، واطرك الصحيحة بدون علامة . وتعتبر الإجابات خاطئة في الحالات الآتية :

(أ) إذا كانت الأرقام غير واضحة .

(ب) إذا كان أى رقم قد صحح أو كشط .

(ج) إذا كان أى رقم مختلفاً عن الجواب المطبوع .

١١ - اكتب رقم المحاولة أمام كلمة (المحاولة رقم . . .) التي في أسفل الصفحة . فمثلاً إذا كانت هذه ثالث مرة تتدرب فيها على اختبار معين فاكُتب رقم ٣ .

١٢ - احسب عدد الإجابات الصحيحة واكتبه في المكان المخصص له :

١٣ - قارن بين هذا العدد والأعداد المقابلة للرتب الموجودة في أسفل الصفحة ، ثم اكتب الحرف الدال على ربتك في هذه المحاولة أمام كلمة (الرتبة . . .) .

١٤ - دوّن هذه النتيجة بالخبر في جدول العمل اليومي بهذه الكراسة ، وكذلك اكتب عدد جميع الأسئلة التي أجبت عنها (الخطأ والصواب معاً) في الخانة الخاصة بذلك في الجدول المذكور .

١٥ - انزع الورقة الشفافة التي كتبت عليها أجوبة الاختبار وسلمها للمدرس .

١٦ - إذا كنت غائباً في أحد أيام التمرين فاكتب كلمة غائب أمام التاريخ بالجدول اليومي . وإن كنت حاضراً ولكنك أعفيت من التدريب ، أو قضيت الوقت في التمرين الحر ، فاكتب « حاضر » . وفي كلتا الحالتين لا تكتب رقماً لليوم .

١٧ - لا تنتقل من اختبار قبل أن تنجح فيه تماماً : لأن الاختبارات متدرجة في الصعوبة ، ويتوقف كل اختبار على ما قبله . ويكفي للنجاح أن تحصل على رتبة ناجح (ب) المبنية في كل اختبار ، ولكن لا تقنع بذلك إذا كان في مقدورك أن تحصل على رتبة ممتاز أو جيد فإذا لم تحصل على الرتبة المطلوبة في أول مرة فكرر التدريب على الاختبار نفسه إلى أن تصل إليها .

١٨ - إذا كانت أخطائك كثيرة يمكنك أن تنقل رموس العمليات التي أخطأت فيها أو التي لم تستطع حلها ، ثم تتمرّن عليها أو على عمليات أخرى من نوعها في ورقة خارجية بالمنزل ، حتى تتقن معرفتها ويتسنى لك حلها بسرعة .

١٩ - إذا أجريت اختباراً ثلاث مرات متتالية دون أن تتقدم فيه بالمرة ، فانتقل إلى الاختبار الذي بعده ، وعد للاختبار الأول بعد مدة .

٢٠- إذا لزم لك في اختبار أكثر من ثلاث محاولات ، ولاحظت أنك تتقدم ، فابدأ بعمل رسم يبين لذلك الاختبار على الصفحة المخصصة له (انظر نموذج الرسم البياني في صفحة ٦٧٧ (شكل ٣٥) .

٢١- إذا نجحت في الاختبار التشخيصي الموجود في أول أى مجموعة ، دل ذلك على أنك لا تحتاج إلى إجراء بقية اختبارات هذه المجموعة ، وبذلك توفر الوقت وتنقل إلى المجموعة التي بعدها . وللنجاح في الاختبار التشخيصي يجب ألا يزيد الخطأ عن عملية واحدة في جميع الاختبار . وإذا لم تنجح فيه لأول مرة فإنه يجوز لك أن تجربه مرة ثانية إذا كان أملك في النجاح كبيراً .

سجل العمل اليومي

رقم اليوم	التاريخ	رقم الاختبار	رقم المحاولة	عدد عمليات الاختبار	عدد العمليات المحلولة	عدد العمليات الصحيحة	الرتبة

الفحص العام :

والمجموعة الأخيرة من البطاقات التدريبية التشخيصية مخصصة للفحص العام . فقد أعدت بحيث تعطى تشخيصاً دقيقاً لجميع العمليات الحسابية التي تناولتها البطاقات التدريبية جميعها . فبعد أن يجيب التلميذ على اختبارات الفحص العام ، وعددها ثلاث ، يحدد العمليات التي أخطأ فيها ثم يرجع إلى جدول الفحص العام ، المطبوع على غلاف هذه المجموعة ، فيجد فيه أمام كل سؤال (عملية) أرقام البطاقات التي يجب التدريب عليها إذا كان التلميذ أخطأ في هذا السؤال . وإليك صورة لجدول الفحص العام :

ما يدلني عليه الفحص العام

إذا أخطأت في العملية فإنه يجب عليّ أن أتدرب على الاختبارات رقم	إذا أخطأت في العملية فإنه يجب عليّ أن أتدرب على الاختبارات رقم	إذا أخطأت في العملية فإنه يجب عليّ أن أتدرب على الاختبارات رقم
١٧ : ٥ ، ٤ ، ٣ ، ١ : ل	١٧ : ٤ ، ٣ ، ٢ ، ١ : و	١ : ٤ ، ٣ ، ٢ ، ١ : و
١٨ : ٤ ، ٣ ، ٢ ، ١ : ل	١٨ : ٥ ، ٤ ، ٣ ، ٢ ، ١ : و	٢ : ٤ ، ٣ ، ٢ ، ١ : و
١٩ : ٤ ، ٣ : ل - ٣ : ل	١٩ : ٢ ، ١ : و - ٢ ، ١ : و	٣ : ٤ ، ٣ ، ٢ ، ١ : و
٢٠ : ٦ ، ٥ ، ٤ : ل - ٢ : و	٢٠ : ٣ ، ٢ ، ١ : و - ٤ ، ٣ ، ٢ ، ١ : و	٤ : ٦ ، ٥ ، ٤ : ل - ٢ : و
٢١ : ٦ ، ٥ ، ٤ : ل - ٢ : ل	٢١ : ١ : و - ٢ : و	٥ : ٦ ، ٥ ، ٤ : ل - ٢ : و
٢٢ : ٦ ، ٥ ، ٢ ، ١ : ل - ٢ : ل	٢٢ : ٢ : و - ٢ : و	٦ : ٦ ، ٥ ، ٢ ، ١ : ل - ٢ : و
٢٣ : ٦ ، ٥ ، ٢ ، ١ : و	٢٣ : ٢ : و - ٢ : و	٧ : ٦ ، ٥ ، ٢ ، ١ : و
٢٤ : ٥ ، ٤ ، ٢ : و	٢٤ : ٢ : و - ٢ : و	٨ : ٦ ، ٥ ، ٢ ، ١ : و
٢٥ : ٤ ، ٢ ، ١ : و	٢٥ : ٢ : و - ٢ : و	٩ : ٦ ، ٥ ، ٢ ، ١ : و
٢٦ : ٦ ، ٥ ، ٢ : و	٢٦ : ٤ ، ٣ ، ٢ ، ١ : و - ٢ : و	١٠ : ٦ ، ٥ ، ٢ ، ١ : و
٢٧ : ٥ ، ٤ ، ٣ ، ٢ : و	٢٧ : ٥ ، ٤ ، ٣ ، ٢ ، ١ : و	١١ : ٦ ، ٥ ، ٢ ، ١ : و
٢٨ : ٦ : و - ١ : ل	٢٨ : ٦ ، ٥ ، ٢ ، ١ : و	١٢ : ٦ ، ٥ ، ٢ ، ١ : و
٢٩ : ٦ : و - ١ : ل	٢٩ : ٢ ، ١ : و - ١ : ل	١٣ : ٦ ، ٥ ، ٢ ، ١ : و
٣٠ : ٩ : و - ٢ : ل	٣٠ : ٤ ، ٣ : و - ١ : ل	١٤ : ٦ ، ٥ ، ٢ ، ١ : و
٣١ : ٥ : م - ٨ ، ٧ : و	٣١ : ٤ : ل - ٣ : هـ	١٥ : ٥ ، ٤ ، ٣ ، ٢ ، ١ : و
		١٦ : ٥ ، ٢ ، ١ : و

نموذج للـ
سجل العمل اليومي

رقم اليوم	التاريخ	رقم الاختبار	رقم المحاولة	عدد عمليات الاختبار	عدد العمليات المحلولة	عدد العمليات الصحيحة	الرتبة
١	٥٩-١٠-٨	١ - ٥	١	٣٥	١٩	١٩	-
٢	٩		٢		٢١	١٩	--
٣	١٠		٣		٢٥	١٨	--
٤	١١		٤		٢٥	٢١	-
٥	١٢		٥		٢٨	٢٦	١٧
١٨	٥٩-١١-٤	٤ - ٥	١	٣٣	٢٣	٢٢	
١٩	٥		٢		٢٤	٢١	
٢٠	٦		٣		٢٩	٢٦	
٢١	٧		٤		٢٠	٢٩	٢٥

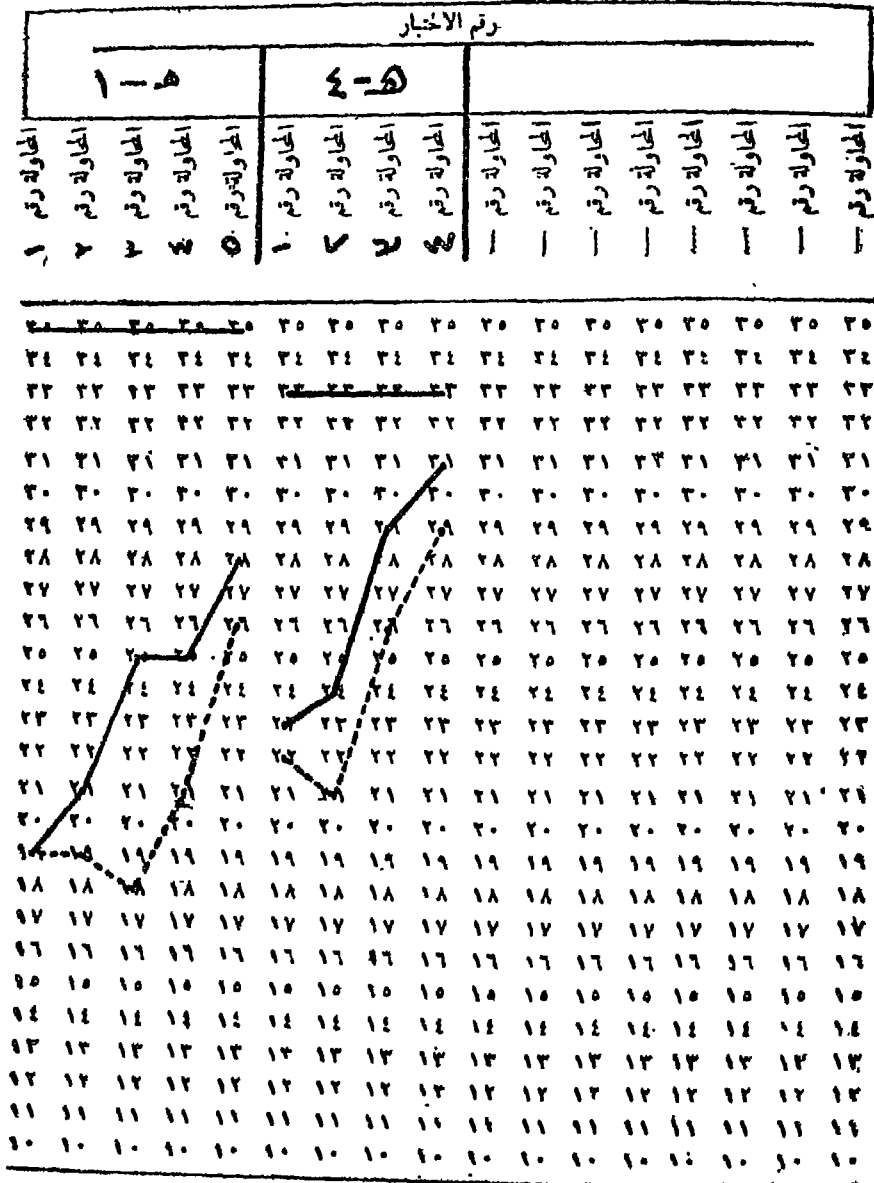
كيفية عمل الرسم البياني

لعمل الرسم البياني لأي اختبار اعرف من الفهرس الصفحة المخصصة لذلك ، ثم سر في العملية بالطريقة الآتية :

١- تملأ محاولة ضع نقطة على عدد الأسئلة التي أجبت عنها في العمود الخاص بالمحاولة . ثم ارسم (بالمسطرة) مستقيماً بين هذه النقطة ونقطة عدد الإجابات في العمود الذي قبله مباشرة . وب نفس الطريقة ضع نقطة على عدد العمليات التي أجبت عنها إجابة صحيحة في العمود الخاص بالمحاولة ثم ارسم

(بالمسطرة) مستقيماً منقطاً بين هذه النقطة والنقطة المقابلة لها في العمود السابق . وأخيراً ارسم مستقيماً أفقياً على العدد الكلي للعمليات في الاختبار ومده في كل محاولة .

واليك مثال لكيفية عمل الرسم البياني :



(شكل ٣٥) نموذج لإحدى صفحات الرسم البياني في كراسة التدريب

وقد قام المؤلف بتجربة لتقييم بطاقات الحساب التدريبية التشخيصية هذه . وفى التجربة ، طبق اختبار ستانفورد الجديد فى الحساب فى إحدى المدارس الابتدائية : وبعد معرفة مستويات فصول المدرسة ، فى الحساب ، اخترنا ثلاث فصول من السنة السادسة الابتدائية متجانسة المستوى . وكان فصلان . منها مجموعة ضابطة والثالث المجموعة التجريبية . ودرب تلاميذ المجموعة التجريبية على استخدام بطاقات الحساب . فاستغرق ذلك من مدرسة الحساب حصتين للشرح . وبعد ذلك كل حصّة ٧ دقائق منها ٤ دقائق للإجابة عن بطاقة اليوم . أما تلاميذ المجموعة الضابطة ، وهى فصلان ، فكانت حصص الحساب لديهم تسير دون هذا التدريب . وتوجيهاً للظروف ، كانت مدرسة الحساب واحدة بين الفصل التجريبي وأحد الفصلين الضابطين .

وبعد شهرين من الدراسة طبق اختبار ستانفورد الجديد فى الحساب مرة ثانية على المجموعتين فتمّين أن متوسط المجموعة التى نالت تدريباً على البطاقات أعلى من متوسط فصلى المجموعة الضابطة . والفرق بين المتوسطين جوهري ، ولكنه غير كبير . وبافتراض أن عدم كبر الفرق الجوهري يرجع إلى قصر مدة التدريب ، أعيدت التجربة على ثلاثة فصول ، فى السنة السادسة الابتدائية أيضاً ، ذلك عند بدء العام الدراسى التالى . وبنفس الطريقة سارت خطوات التجربة . وبعد تدريب ٤ شهور أعيد قياس مستوى الفصول الثلاثة فى الحساب بإعادة تطبيق اختبار ستانفورد ، فتمّين أن المجموعة التى دربت على البطاقات فاقت المجموعتين الأخرتين . وكان الفرق جوهرياً وأكبر من الفرق السابق ، مما يدلنا على قيمة هذه البطاقات .

إلا أننا وجدنا تفاوتاً كبيراً بين مستويات تلاميذ الفصل التجريبي ، إذ أن مستوى البعض فى بعض العمليات دل على أنها أعلى من مستوى قدراتهم وتحصيلهم : وأنهم لم يتقنوا بعض العمليات برغم استمرار تدريبهم شهراً . ومن هنا اتضح لنا أن هناك من يعانون مشاكل القيام بنوع معين

من العمليات وهذه هي الفائدة التشخيصية للبطاقات : وقد أفادت مدرسة الفصل من هذا الكشف بالتركيز على هذه العمليات . ولكن عدم فهم التعليمات أو عدم الارتفاع إلى المستوى المتوسط دل ، كما سبق القول : على احتمال عدم مناسبة مستوى بعض العمليات لهذا المستوى الدراسي .

ولهذا يجرى الآن بحث لتقييمها في السنة الأولى الإعدادية بإحدى المدارس بنفس الطريقة . وتدل النتائج الأولية على صلاحيتها في هذا المستوى سواء من حيث القيمة التشخيصية أم التدريبية .

نحو المجلد الثاني

سنتناول في المجلد الثاني الذكاء والقدرات والاستعدادات والسمات والميول والاتجاهات والقيم و . . الخ ، ولكل منها أثره ، كما سبق القول ، على التحصيل وعلى التلميذ بما هو فرد . وعلى هذا فكل أدوات تقدير ومساعدة التحصيل والعامل بالتربية والتلميذ ، لم تكتمل بعد : وسنتعرف ، من ثم ، على تكوين الفرد وجوانبه وكيف تقاس . ويمكن للمهتم بالتحصيل وغيره من شئون التربية ، بهذه الطريقة ، أن يفيدوا من المجلد الثاني لاستكمال الأدوات والتطبيقات حتى نصل إلى فهم شامل متكامل هادف أفضل .

قائمة مراجع القسم الثالث

- 1— Adkins, D.C., et al Construction and Analysis of achievement tests. Washington, D.C. : Government Printing office, 1947.
- 2— Allen, Wendell C., Cumulative Pupil records, New York, Bureau of Publications, Teachers College, Columbia University, 1943.
- 3— American Psychological Association, Technical recommendations for Psychological tests and diagnostic techniques. Washington : Amer. Psychol-Assoc. 1954.,
- 4— Anastasi, Anne. Psychological Testing. (2. nd. edition), N.Y. : Macmillan Co., 1957.
- 5— Anonymous. Tests and Testing programs. Princeton, N.Y. : Educational Testing Service, 1948.
- 6— Anonymous. Results of Orientation Test in Schools of business of fifty Colleges American of Accountants, College Testing Program, Bulletin No. 4. New York Committee on Selection of Personnel, 1947.
- 7— Arsenian, S. Informing College Freshmen of their Test Scores. Sch. and Sco., 1942.
- 8— Barnes, M.W. Gains in The A.C.E. Psychological Examination during The Freshman-Sophomore years. Sch. and Soc., 1943.
- 9— Baron, L. and Bernard H.W. H.W. Evaluation Techniques for Classroom Teachers, Mc-Graw Hill Book Co., 1958.
- 10— Barr, A.S., et al. The Measurement and Prediction of Teaching efficiency : A Summery, Inverstigation, J. exp. Educ., 1948.
- 11— Barrett, D. W. Prediction of achievement. in Typewriting

and Stenography in a liberal arts College. *J. appl. Psychol.*, 1946.

- 12— Bean, Kenneth L., *Construction of educational and Personnel Tests*, New York, McGraw-Hill, 1953.
- 13— Beatty, J.D., and Cleeton, G.U. *Predicting achievement in College and after Graduation.* *Pers. J.*, 1928.
- 14— Bennett, G.K., and Cruikshank, Ruth, M.A. *Summary of Clerical Tests* N.Y. : Psychol. Corp., 1949.
- 15— Betts, E.A. *Foundations of Reading Instruction.* N.Y. : American Book Co., 1946.
- 16— Betts, E.A. *Betts Ready to Read Tests : Manual* Meadville, Pa. : Keystone View Co., 1934.
- 17— Berdie, R.F. *Prediction of College achievement and Satisfaction.* *J. appl. Psychol.*, 1944.
- 18— Blackstone, E.G., and McLoughlin, M.W. *Blackstone Stenographic Proficiency Tests. Stenography Test : Manual of Directions.* Yonkers-on-Hudson, N. Y. : World Book Co., 1932.
- 19— Blair, G.M., and Kammam, F.F. *Do intelligence Tests requiring reading ability give spuriously low Scores To Poor readers at the College level ?* *J. educ. Res.*, 1942.
- 20— Blommer, Paul, and Dingist, E.F. "Rate of Comprehension of reading ; its measurement and its relation to comprehension." *J. educ. Psychol.*, 1944.
- 21— Brdin, E.S., and Bixler, R.H. *Test Selection : a Process of Counseling.* *Educ. Psychol Measmt.*, 1946.
- 22— Brewer, J.M. *Education as guidance* New York, : Macmillan, 1932.
- 23— Brown, Clara M., et al *Minnesota Food Score Cards.* Minneapolis : Univer. Minnesota Press, 1946.
- 24— Buros, O.C, (Ed) *The Nineteen-Forty Mental Measurement-*

- ments Yearbook. Highland Park, N.J. : Mental Measmt. Yearbook, 1941.
- 25— Buros, O.K. (Ed) The Third Mental Measurements Yearbook. New Bruns - wick, N. J. : Rutgers Univ. Press, 1949.
- 26— Buros, O.K. (Ed.) Mental Measurement Yearbook, New Brunswick, N.J. : Rutgers Univ. Press, 1938, 1939, 1940, 1949.
- 27— Burt, C. The Reliability of Teachers Assesment of Their Pupils, B.T. Educ. P. Vol. XV, 1945.
- 28— Buswell, G.T., and John, Lenore, Diagnostic Studies in arithmetic, Suppl. educ. Monoge., 1926.
- 29— Buswell, G.T., and John, Lonore, Diagnostic Test For Fundamental Processes in Arithmetic : Manual, Bloomington, Ill. : Public School Publ. Co., 1925.
- 30— Chamberlin, D., et al., Did they Succeed in College ?, New York Harper, 1942.
- 31— Chapman, J.C. Trade Tests, N.Y. : Henry Holt, 1921.
- 32— Charters, W.W., and Douglas Waples, The Commonwealth Teacher-Training Study, University of Chicago Press, Chicago ; 1929.
- 33— Columbia Research Bureau Algebra Test 2 : Manual, Yonkers-on-Hadson, N.Y. : World Book Co., 1933.
- 34— Conrad, H.S. Summary report on research and development of the Navy's applitude testing program. Final Report on Contract OEM sr 705, OSRD, 1945, Publ. Bd. No. B 289, Washington, D.C., Dept. of Commerce, 1946.
- 35— Cooperative Algebra Test : Elementary ; Intermediate. Princeton, N.J. : Coop. Test Div., Educ. Testing Service, 1951.
- 36— Cooperative Contemporary Affairs Test For College Students, Princeton, N.J. : Coop, Test Div., Educ. Testing Service, 1953.

- 37— Cooperative English Test : Manual. Princeton, N.Y. : Copp. Test Div., Educ. Testing Service, 1953.
- 38— Cooperative General Culture Test, Princeton, N. Y. : Coop Test Div., Educ. Testing Service, 1951.
- 39— Cooperative General Achievement Tests : Manual. Princeton, N.Y. Coop. Test Dev., Educ. Testing Service, 1954.
- 40— Cooperative Mathematics Tests for Grades 3,8, and 9. Princeton, N.Y. : Coop. Test Div., Educ. Testing Service, 1951.
- 41— Cooperative Plane Geometry Test, Princeton, N.Y. : Coop. Test Div., Educ. Testing Service, 1951.
- 42— Cottrell, Donald P. (ed.) Teacher Education for & Free People, The American Association of Colleges for Teacher Education, oneonta, N.Y., 1956.
- 43— Coxe, W.W, and Cornell, E.L. Prognosis of teaching ability of Student1 in New York State normal Schools, Univ. of State of N.G. Bull, 1934.
- 44— Cronbach, L.J. Eesentials of Psychological Testing. Harper Broths. N.Y., 1949.
- 45— Crow, L.E. and Crow, A. Educational Psychology, American Book Co. (Revised ed.) 1958.
- 46— Darley, J.G., Testing and Counseling in the high-School Guidance Program, Chicago : Science Research Associates, 1943.
- 47— Davis, F.B. Item Analysis in Relation to Educational and Psychological Testing. Psychol. Bull., 1952.
- 48— Diam and, Leon N. "Testing the Test makers" Sch. Sci. and math, 1932.
- 49— Diederich, P.B. The Measurement of Skill in Writing, School Review, 1946.
- 50— Doob, Leonard W., Public Opinion and Propag. N.Y. : Henry Habt Co., 1948.

- 51— Douglass, H.R., and Margaret Tallmadge, How university Students prepare for now Types of examinations, Sch. Soc, 1934.
- 52— Dougss, H R. Different Levels, and patterns of ability necessary For success in College Occupations, 1943.
- 53— Drought, N.E. An analysis of eight Measures of Personality adjustment in relation to relative Scholastic achievement. J. appl. Psychol., 1938.
- 54— Dubois, P.H. (Ed.) The Classification Program AAF Aviation Psychology Report, No, 2. Washington, D.C. : Government Printing Office, 1947.
- 55— Dunlap, J.W. The Predictive value of items for achievement in Various School Subjects, J. appl Psychol., 1935.
- 56— Durost, Walter N., what constitutes a Minimal testing Program ?, Educ. Psychol. Meas., Vol. 7, Spring, 1947.
- 57— Durost, W.N., et al. Evaluation and Adjustment Series Handbook, Yonkers-on-Hudson, N.Y. : World Book Co. 1937.
- 58— Durrell. D.D. Durrell Analysis of Reading Difficulty : Manual of Directions, Yonkers-on-Hudson, N.Y. : World book Co., 1937.
- 59— Dyer, H.S. College Board Scores : Their Use and Interpretation, No. 1. N.Y. : College Entrance Examination Board, 1953.
- 60— Dyer, H.S. Evidence on the Validity of the U.S.A.F.I. Tests of General Educational Development, Educ. Psychol. Measmt., 1945.
- 61— Dysinger, W.S. Test reports in educational Counseling. Educ. Psychol. Measmt., 1943.
- 62— Ebel, Robert L, Writing the test item, E.F. Lindquist, Editor. Educational Measurement, Washington, D.C., American Council on Education, 1951.

- 63— Eckert, R., and Marshall, T.O. when youth leave School.
New York : Mc Graw-Hill, 1938.
- 64— Eells, W.O. Scholastic ability of Secondary School Pupils
Educ. Rec., 1937.
- 65— Elsbrea, Willard S., Pupil progress in the elementary
School, New York, Bureau of Publications, Teachers
College, Columbia University, 1943.
- 66— Findley, W.J. Agroup testing Program for the modern
School. Educ. Psychol. Measmt., 1945.
- 67— Flanagan, J.C.A Bulletin Reporting the Basic Principles
and Procedures Used in the Development of Their
System of Scaled Scores. N.Y : Coop. Test Service. 1939.
- 68— Flanagan, J.C.A Preliminary Study of the 1940 edition of
the National Teachers Examination, Sch. and Soc., 1941.
- 69— Frank W. Hart, Teachers and Teaching, The Macmillan
Co., N.Y., 1934.
- 70— Freeman, F.S : Theory and Practice of Psychological
Testing, New York : Henry Holt and Co. Inc. 1950.
- 71— Froelich, Clifford P., and Arthur L. Benson : Guidance
Testing. Chicago : Science Research Associates Inc. 1948.
- 72— Gardner, E.F. The importance of reference groups in
Scaling Procedure. Proc. 1952.
- 73 Gardner, E.F. Value of norms based on a new type of
Scale unit. Proc., 1948.
- 74— Garretson, O.K. Relationships between expressed preferen-
ces and Curriculum abilities of ninth grade boys. New
York : Teachers College Contributions to Education, No.
396, 1930.
- 75— Gates, A.I. Gates Basic Reading Tests : Manual of
Directions. N.Y. : Teachers coll., Columbia Univer., Bur.
Publ., 1943.
- 76— Gaudet, F.J., and Marryott, E.S. Predictive value of the

- Ferson-Stoddard Law Aptitude Examination. An. Law Sch. Rev., 1930,
- 77— Gerberich, J.R. Validation of a State-wide educational guidance Program for high-School Seniors. Sch. and Soc., 1931.
- 78— Ghiselli, E.E., and Brown, C.M. Personnel and Industrial Psychology. N.Y. : Mc Graw-Hill, 1948.
- 79— Good, C.V. and Merkel W.R. Dictionary of Education. (2 nd edition) Mc Graw-Hill, N.Y. 1959.
- 80— Goodenough, Florence L. Mental Testing, Rinehart, N.Y. 1949.
- 81— Gotham, A.E. Personality and teaching efficienncy. J. exp. Educ. 1945.
- 82— Gray, W.S. Standardized Oral Reading Paragraphs : Manual. Bloomington, Ill. : Public School Publ. Co., 1915.
- 83— Greene, E.B.A. Sampling of Vocabularies of Superior adults. J. higher Educ., 1933.
- 84— Greene, E.B. Michigan Vocabulary Profile Test : Manual of Directions. Yonkers-on-Hudson, N.Y. : World Book Co., 1949.
- 85— Greene, E.B. The Michigan Vocabulary Profile Test, after ten years. Educ. Psychol. Measmt., 1951.
- 86— Greene, H.A., and Bince, H.W. Iowa Plane Geometry Aptitude Test. : Iowa City : Bur. Educ. Res. Service, Univer. Iowa, 1942.
- 87— Greene, H.A., Jorgensen, A.N., and Kelley V.H. Iowa Silent Reading Tests, New Edition (Revised). Advanced Test : Manual of Directions. Yonkers-on-Hudson, N.Y. : World Book Co., 1943.
- 88— Greene, H.A., and Kelley, V.H. Iowa Silent Reading Tests New Edition (Revised). Elementary Test : Manual of

- Directions. Yonkers-on-Hudson, N.Y. : World Book Co., 1943.
- 89— Greene, H.A., Wright, F.B., Ruch, G.M., and Studebaker, J.W. Compass Survey Tests in Arithmetic : Manual. Chicago : Scott, Foresmae, 1927.
- 90— Greenc, H.A. and Jorgenson A.N. and Jerberich, J.R. Measurement and Evaluation in The Secondary School. Longmans Grean and Co. N.Y. 1947.
- 91— Greene, E.B. Measurement of Human Behavoir. Odessey N.Y. : 1941.
- 92— Gregory, W.S. From high School to College. Occupations, 1939.
- 93— Gregory, W.S. Data regarding The Academic Interest Inventory. Educ. Psychol. Measmt., 1946.
- 94— Guilford, J.P. New Standards For Test evaluation. Educ. Psychol. Measmt, 1946.
- 95— Gulliksen, H. Intrinsic Validity. Amer. Psychologist 1950.
- 96— Guliksen, H., Theory of Mental Tests, 1950. & N.Y. : John Wiley & Sons.
- 97— Guilford, J.P. Psychometric Methods. N.Y. : Mc Caw-Hill, 1936.
- 98— Guilford. J. P. Factor analysis in a Test-development Program. Psychol. Res., 1948.
- 99— Harry, D.P., Jr., and Durost, W.N. Essential High School Content Battery : Manual of Directions. Yonkers-on-Hudson, N.Y. : World Book Co., 1951.
- 100— Hartmann, G.W. Measuring Teaching efficiency among College instructors. Arch. Psychol., No. 154, 1933.
- 101— Hawkes, Herbert, et al (eds.) The Construction and Use of Achievement examinations. Boston : Houghton. Mifflin, 1936.

- 102— Hessemer, M. The Thurstone PMA Tests in a study of academic success With School of engineering chemistry and physics. Unp-ubl. thesis, penn. State College, 1942.
- 103— Hildreth, Gertrude H., Metropolitan achievement tests : manual for interpreting , Yonkers . New York , world Book., 1948.
- 104 — Hilolreth, Gertrude H., and Griffiths, N. L. Metropolitan Readiness Tests : Directions for Administrating and Scoring. Yorkers-on-Hudson, N. Y. : world Book Co., 1949.
- 105— Hillegas, M. B. A. Scale for the measurement of quality in English Composition by young People. Teachers coll. Rec., 1912.
- 106-- Jones. G., and Galbraith, A. The interpretation of stand-ardired tests. Sch. and Soc., 1941.
- 107-- Jones, R. D. Prediction of teaching efficiency From objective measures. J. exper. Educ., 1946.
- 108-- Jordan, A. M. Educational psychology. New York: Henry Holt, 1928.
- 109— Jordan, A. M. : *Measurement in Education*, N. Y. : Mc. Grow Hill Book Co., 1953.
- 110— Kavin, Ethel, Records and reports : obeservations, tests, and measurements, Early Childhood education, Forty-Sixth Yearbook, National Society for the study of Education, 1947.
- 111— Kelley, T. L., and Ruch, G. M. Stanford Achievement Test : Manuals for primory, Elementary, Intermediate, and Advanced Batteries. Yonkerson - Hudson, N. Y. : World Book Co., 1953
- 112 — Kelley, T. L. Ridye-route norms. Harvard educ. Rev., 1940
- 113— Kelley, T. L. and Kery, A. C. Tests and Measeremens in the Social Sciences. N. Y. : Scribner, 1934.
- 114— Kohn, H. A. Achievement and intelligence examinations

Correlated with each other and with teacher's rankings, J. genet. psychol., 1938.

- 115— Lee, J. M. Lee Test of Algebra Ability : Manual. Bloomington, Ill. : Public School publ. Co., 1930.
- 116— Lee, Doris M., and Lee, J. M. Lee Test of Geometric Aptitude : Manual. Los Angeles : California Test Bur., 1931.
- 117— Lewis, E. E. Lewis Scale for Measuring Special Types of English Composition. Yonkers-on-Hudson, N. Y. : Wsrld Book Co., 1921.
- 118— Lindquist, E. F., Iowa tests of General Educational Development : Manual, Chicago, Ill., Science Research Associates, 1951.
- 119— Lindquist, E. F. (Ed.) Educational Measurement. Washington : Amer. Council Educ., 1951
- 120— Lindquisb, E.L. (Ed.) Meesurement Handbook. Washington, D. C. American Council on Education, Boston : Houghton Mifflin Co., 1951
- 121— Lindquist, E. F., Preliminary Considerations in objective test Construction, Chapter 5 in E. F. Lindquist, Editor, Educational measurement, Washington, D. C., - American Council on Educatism, 1951.
- 122— Little, J. K. Result of use of machines For testing and For drill, upon Learning in educotional psychology. J. exp. Educ., 1934.
- 123— Lord, A B A. Survey of Four hundred Forty-nine special Class Pupils. J. educ. Res., 1933.
- 124— Lorge, I. The Fundamental Nature of Measurement, In. E. F. Lindquist (ed.) Educational Measurement. Washington : Amer. Council Educ., 1951.
- 125— Mathews, L. H. An Item Analysis of Maasures of teaching ability. J. Educ. Res., 1940.

- 125— Mc. Call, W. A., How to Measure in Education, 1922
- 126— Meyer, G. An experimental study of old and new types of examination, J. educ. Psychol., 1934,
- 127— Meyer (george- "The effect on Recall and Recognition) of the examination set in classroom Situations" G. Educ. Psychol., 1937.
- 128 — Micheels, William J., and M. Ray Karnes. Measuring educational achievement, New York, Mc. Graw-Hill, 1950
- 129 — Monroe, Walter S., Editor Encyclopedia of educational research, reved., New York, Macmillan, 1950
- 130 — Monroe, Marion. Children who Can not Read. Chicago : Univer. Chicago Press, 1932.
- 131— Monroe, Marion. Reading Aptitude Tests : Manual. Boston : Houghton Mifflin, 1935
- 132— Mook E, H. Psychology For business and industry. New York : Mc Graw-Hill, 1942.
- 133— Mozzai, R. Desirable Traits of Successful Teacher, J. Teach. educ., 1951.
- 134 — Mursell, J.L. Psychological Testing. (2 nd edition) Longmans Green and Co. 1950.
- 135— National Society for The Study of Education, The Measurement of understanding, The Forty-fifth Year book, Part I, Chicago, Illinois, University of Chicago Press, 1946.
- 136— Nunally Jum C. Tests and Mearurements Assessment and Prediction. Mc-Graw-Hill Book Co. N.Y. 1959.
- 137— Odell, C. W., How to improve classroom Testing, Dubuque, Iowa. William C. Brawn, 1953.
- 138— Orleans, J.B., and Orleens, J.S. Orleans Geometry Prognosis Test, Revised Edition : Manual of Directions. Yonkers-on-Hudson, N.Y. : World Book Co., 1951.
- 139 — Orleans, J.B., and Orleans, J.S. Orleans Algebra Prog-

nosis Test, Revised Edition : Manual of Directions.
Yonkers-on-Hudson, N.Y. : World Book Co. 1951.

- 140— Patterson, C.H. on The Problem of the Criterion in Prediction Studies. J. Consult. Psychol., 1946.
- 141— Paterson, D.G., Schnelidser, G. and Williamson, E.G. Student guidance Techniques. New York : Mc Graw-Hill, 1938.
- 142— Paul A. Witly, The Teacher Who Has Helped the Most, National Educ. Assoc. J. 1947.
- 143— Payne, A.F. Sentence Completions New York Guidance Clinic, 1928.
- 144— Pond. M. Occupations, intelligence, age and Schooling. Pers. J., 1933.
- 145— Pressey, Luella G. An investigation of The Technical Vocabularies of the School Subjects, Educ. Res. Bull., Ohio State Univer., 1924.
- 146— Raths, Iovis. "Basis four Comprehensive evaluation" Educ. Res. Bull., 1936.
- 147— Remmers, H.H. and Gage, N.L. and Rummel, J.F. Practical Introduction to Measurement and Evaluation. Harper Brothers 1960.
- 148— Rice, J.M. Futility of Spelling Grind, Forum, 1897.
- 149— Rice, J.M. The Public School System of the United States, N.Y. : The Century Co., 1893.
- 150— Richey, R.W., and William H. Fox. "A Study of Some Opinions of High School Students with Regard to Teachers and Teaching" Bulletin of the School of Educ Indiana University Bloomington Ind., 1951.
- 151— Richey, R.W. Planning for Teaching (2nd.) N.Y. : Mc Graw Hill Co, 1958.
- 152— Rosenfeld, M.A., and Nemzek, C.L. Long-range Prediction of College marks. Sch. and Soc., 1938.

- 153— Ross, C.C. Should Low-rating College Freshmen be told their Scores in intelligence Tests? Sch and Soc., 1938.
- 154— Ross, C. C. Measurement in Today's Schools. N.Y. : Prentice-Hall, 1947.
- 155— Ruch, G.M., et al, Compass Diagnostic Tests in Arithmetic : Manual Chicago : Scott, Foresman, 1925,
- 156— Ruch, N.M., and Stoddard, G.D. Comparative Reliabilities of Five Types of Objective Examinations, J. Educ, Psychol., 1915,
- 157— Ruch, G.M., and Stoddard, G.D, Tests and Measurement in High School Instruction, N.Y. : World Book Co., 1927.
- 158— Russell, B, History of Western Philosophy, London, George Allen and Unwin, 1945,
- 159— Schneidder, G.G., and Berdie, R.F, Educational ability Patterns, J, educ, Psychol., 1942,
- 160— Schneider, G.G, Grade and age norms For the Minnesota Vocational Test For Clerical Workers. Educ, Psychol, Measmt., 1941.
- 161— Seagoe, M.V, Prognostic tests and Teaching Success, J, educ, Res., 1945,
- 162— Seagoe, M.V, Permanence of interest in Teaching, J, educ. Res., 1945.
- 163— Seashore, H., and Bennett, G, K, Seashore-Bennett Stenographic Proficiency Test : Manual, N.Y. : Psychol, Corp., 1946,
- 164— Seder, M, The reliability and Validity of the A.C.E, Psychological Examination, 1938,
- 165— Segel, David, State Testing and evaluation Programs, Federal Security Agency, Office of Education, Circular No, 320, Washington, D.C, U.S, Office of Education, 1951,

- 166— Shaffer, L.F, Information Which Should be Provided by Test Publishers and testing agencies on the Validity and use of their tests : Personality tests, Proc 1949 Invtit Coni testing Services,
- 167—Skinner, C. E. Educational Psychology. Prentice Hall 1950
- 168— Smith, E. R., Tyler, R. W., et al Appraising and Recording Student progress. N.Y. : Harper, 1942
- 169— Sorenson, H. Psychalagy in Education N. Y., 1954
- 170— Spitrer, H. F., et al. Iowa Every-pupil Tests of Basic Skills : Manual Boston : Houghton Mifflin, 1947
- 171— Staff, Personnel Research Section, A.O. O. Validation of the General Clerical Abilities Test for Selection and Placement of War Dept., Civilian Personnel. Amer. Psy, 1947
- 172— Stalnaker, John M., The essay type of examination, Chapter 13 in E. F. Lindquint, Editor, Educational measurement, Washington, D. C. American Council on Dducation, 1951
- 173— Stead, W. H., Shartle, C. L., et al - Occupational Covn-selinh Techniques. N. Y. : American Book Co., 1940
- 174— Ttephens, J.M. Educotional Psychalagy N. Y, 1951
- 175— Stevens, S. S. On The Theory of Scales of Measurement Science, 1946
- 176— Stopher, E. C. The freshmen testing program. J. higher Educ., 1941
- 177— Strang, R. Behavior and backgronnd of students in college and Secondary School New York : Harper & Brothers, 1937
- 178— Strang, R. Personal development and guidance in college and secodary School. New York : Harper & Brothers,1934
- 179-- Strang, Ruth, Reporting to parents, New York Bureau of publications, Teachers college, columbia University, 1947

- 180 — Stuit, D. B. A study of the vocational interests of a group of teachers, college freshmen. J. appl. Psychol., 1938
- 181 — Stuit, D. B. (Ed.) Personnel Research and Test Development in the Bureau of Naval Personnel. Princeton, N. J. : Princeton Univer. Press, 1947
- 182 — Super, D. E. Appraising Vocational Fitness by means of Psychological Tests. N. Y. : Harper Brothers, 1949
- 183 — Tiegs E. W., and Clork, W. W. California Achievement Tests : Manuals of Directions For Primary, Elementary, Intermediate, and Advanced Batteries . Los Angeles : California Test Bvr., 1951
- 184 — Terry, P. W., How students study for three types of objective tests, J. educ. Res., 1934
- 185 — The Evaluation of Student Teaching, Twenty-eighth Annual Yearbook of the Association for Student Teaching State Teachers College, Lock Haven, 1949
- 186 — The Purposes of Education in American Democracy, National Education Association, Educational Policies Commission. Washington, 1938
- 187 — Thomlinson, Loren R. Recent Studies in the evaluation of teaching, educ. Res - Bull., 1955
- 188 — Thorndike R. L. and Hagen, A. Measurement and Evaluation in Psychology and Education. John Wiley & Inc. N. Y. 1957
- 189 — Thorndike, R. L., and Lindquist, E. F. Educational Measurement, 1951
- 190 — Thorndike R. L. (Ed.) Research Problems and Techniques. AAF Aviation Psychology Report, No. 3 Washington D. C. : Government Printing Office, 1947
- 191 — Thurstone. L. L, A, Method of Scaling Psychological and Educational Tests, J, Educ, Psychology, 1925
- 192 — Thurstone, L, L, Scales Construction With Weighted Observation, Psy., 1928,

- 193— Thurstone, L.L. The Unit of Measurement in Education Scales, J.E. Psy., 1927.
- 194 — Travers, Robert M.W., How to make achievement test New York, odyssey Press, 1950.
- 195— Travers, R.M.W., Craig, H. and Fenchel, G. Objectiv of Teacher Education, J. Educ Res., 1951.
- 196— Travers, R.M.W. Educational Measurement, N.Y. ma millan Co., 1955.
- 197— Traxler, Arthur E., Administering and Scoring TI objective Test, Chapter 10 in E.F. Lindquist, Editc Educational measurement, Washington, D.C., Americ: Council on Education. 1951.
- 198— Traxler, A. E. Cumulative Test records : Their natu and uses. Educ. Psychol. Measmt., 1941.
- 199— Traxler, A. E. Evaluation of aptitude and achievement a guidance Program. Educ. Psychol. Measmt., 1946,
- 200— Traxler, A,E., et al, Introduction to Testing and the U of Test Results in Public Schools, N,N, : Harper, 1953
- 201— Traxler, Arthur E,, Techniques of guidance, New Yor Harper, 1945,
- 202— Traxler, Arthur E, The nature and use of reading Test Chicago : Science Research assoc 1941,
- 203— Traxler, Arthur E, The use of Test results in diagnos and instruction in the tool Subjects, rev, ed,, Education Records Bulletin No, 18, New York, Education Records Bureau, 1949,
- 204— Triggs, Frances O,, et al, Diagnostic Reading Tes A History of Their Con-Struction and Validation, N,Y, Common Diagnostic Reading Tests, Inc., 1952,
- 205 — Triggs, Frances O,, et al, Diagnostic Reading Tests Their Interpretation and Use in the Teacing of Rea ing, N,Y, : Comm on Diagnostic Readieg Tests, Inc 1952,

- 206— Trimble, H.C., and Cronbach, Lee g. A, practical procedure for the Rigorous interpretation, of test - Retest scores in terms of pupil growth,, J Educe. Res,, 1943,
- 207— Troyer, M, E,, and Angell, G, W, Manual for the SRA Self-Scorer, Chicago : Science Research Associates, 1949
- 208— Trump, J,D,, and Haggerty, Helen R, Basic principles in achievement test item construction PRS Report 979, Personnel Res, Sect, AGO, 1952,
- 209— Tyler, Ralph W,, The function, of measurement in improving instruction, Chapter 2 in E, F, Liudquist, Editor, Educational measurement, Washington, D, C,, American Council on Education, 1951,
- 210— Tyler, R,W, Permanency of Learning, J, higher Educ,, 1933
- 211— Tyler, R,W,, Constructing acheivement Tests, columbus, Ohio, Bur Educ Res,, Ohio Seote Univ., 1934,
- 212— United States Armed Forces Institute, Tests of General Educational Developement : Examiner's Manual, Washin-gton : Amer, Council Educ,, 1944,
- 213— United States office of Education, National Committee on Cumulative Records, Hand book of Cumulative Records, Federal Security Agency, office of Education, Bulletin 1944. No 5, Washington, United States Govern-ment Printing office, 1945
- 214— Van Wagenen, M, J, van Wagenen English Composition Scales, Yonkers-on-Hudson, N.Y. : World Book Co., 1923,
- 215— Voughn, K.W, Planning the objective test, chapter 6 in E,F, Lindquist, Editor, Educational meosurement, Washin-gton, D,C, American Council on Education, 1951,
- 216— Vernon, P.E, Educational abilities of training College Students, Brit J, educ, Psychol,, 1939
- 217— Warren H,C. Dictionary of Psychology Hangleton Mifflin Co. 1934

- 218— Watson. G., and Glaser, E.M. Watson-Gloser Critical Thinking Appraisal : Manual. yonkerson Hudson, N. Y. : World Book Co., 1952.
- 219— Webster, New International Dictionary, 1945.
- 220— Weitzman, E., and Monamara, W. J. Constructing Classroom Examinations Chicago : Science Research Associates, 1949.
- 221— Weitzman, Ellis. and Walter J. McNamare, Constructing classroom examinations : a guide for teachers, Chicago, Science Research Associates, 1950.
- 222— Wert, J.F. Twin examination assumptions. J. higher Educ., 1937.
- 223— Williamson, E.G. How to counsel Students. New York : Mc Grow-Hill, 1939.
- 224— Wilson, M.H. The Self-appraisal Program in the Philadelphia Junioc High Schools, Educ. Psychol. Measmt., 1946.
- 225— Wilson, J.W., and Carpenter, E.K. The need for restandardization of altered tests, Amer. Psychol., 1948.
- 226— Witty, Paul A, Mental Hygiene in modern Education, Fifty-fourth Yearbook of the National Society for the Study of Education, University of Chicaho Press Chicago, 1955.
- 227— Wolf, R.R., Jr. Differential forecasts of achievement and Their Use in educational counseling, Psychol. Monog., No. 1, 1941.
- 228— Wrightstone J.W. Test of Critical Thinking in the Social Studies : Manual, N. Y. : Teachers Coll., Columbia Unive., Buc. Publ., 1939.
- 229— Wrinkle, William L., Improving making and reporting Practices in elementary and Secondary Schools, New York, Rinehart, 1947.
- 230— Yeager, T.C. An analysis of certain of Selected high School Seniors interested in Teaching, New York : Teachers College Contribution, No. 660, 1935.

٢٣١ - أحمد زكى صالح : بطاقة تقويم المدرس : مكتبة النهضة المصرية
سنة ١٩٥٩ .

٢٣٢ - أحمد زكى صالح : علم النفس التربوى « الطبعة الخامسة » مكتبة
النهضة المصرية سنة ١٩٥٧ .

٢٣٣ - ارثر جيتس وآرثر جيرسلد ، ت . ر . ماكونل ، وروبرت س .
تشالمان : علم النفس التربوى « ٣ أجزاء » ترجمة إبراهيم حافظ ،
والسيد محمد عثمان ، ومحمد عبد الحميد أبو العزم بأشراف
الدكتور عبد الله القوصى .

٢٣٤ - إسماعيل . . . القباني ، دراسات فى مسائل التعليم : مكتبة
النهضة سنة ١٩٥١ .

٢٣٥ - إسماعيل محمود القباني : دراسات فى تنظيم التعليم : مكتبة
النهضة المصرية سنة ١٩٥٨ .

٢٣٦ - إسماعيل محمود القباني : التربية عن طريق النشاط : مكتبة
النهضة المصرية سنة ١٩٥٨ .

٢٣٧ - إسماعيل القباني ومحمد عبد السلام أحمد : اختبار استانفورد الجديد
فى الحساب (الطبعة الخامسة) لجنة التأليف والترجمة والنشر
سنة ١٩٦٠ .

٢٣٨ - إسماعيل القباني ومحمد عبد السلام أحمد : البطاقات التدريبية
والتشخيصية فى الحساب : لجنة التأليف والترجمة والنشر .

٢٣٩ - الأهداف والمستويات فى الترييه والتعليم : وزارة الترييه والتعليم
سنة ١٩٥٥ .

٢٤٠ - جلينفورد وآخرون : ميادين علم النفس : أشرف على جمعه
جلينفورد وعلى ترجمته الدكتور يوسف مراد - دار المعارف
سنة ١٩٥٦ .

- ٢٤١ - رابطة التربية الحديثة : مشكاة الإمتحانات فى مصر : القاهرة : مطبعة دار الأليف والترجمة والنشر سنة ١٩٣٨ .
- ٢٤٢ - رمزية الغرب : لإختبار الاستعداد للمهن التعليمية : كاية البنات سنة ١٩٥٧ .
- ٢٤٣ - سعد جلال : التوجيه النفسى والتربوى والمهنى : مكتبة النهضة المصرية سنة ١٩٥٧ .
- ٢٤٤ - عطيه محمود هنا : التوجيه التربوى والمهنى مكتبة النهضة المصريه سنة ١٩٥٩ .
- ٢٤٥ - محمد عبد السلام أحمد : تقرير بحث تقييم التلميذ بالمرحلة الابتدائية فى المناطق المختلفه لجنة التأليف والترجمة والنشر سنة ١٩٥٦ .
- ٢٤٦ - محمد عبد السلام أحمد : تقرير بحث تقييم امتحانات القبول بالمرحلة الاعداديه سنة ١٩٥٥ .
- ٢٤٧ - محمد عبد السلام أحمد : تقرير بحث تتبع خريجي التعليم الصناعى لجنة التأليف والترجمة والنشر سنة ١٩٦٠ .
- ٢٤٨ - محمد عبد السلام أحمد : تقرير بحث أثر البطاقات التدريبية والتشخيصية فى الحساب : لجنة التأليف والترجمة والنشر سنة ١٩٦٠ .
- ٢٤٩ - محمود رشدى خاطر : سلسلة اختبارات سرس اللبان فى القراءه للكبار سنة ١٩٦٠ .
- ٢٥٠ - محمود رشدى خاطر : سلسلة اختبارات سرس اللبان للقراءة الصامته سنة ١٩٦٠ .
- ٢٥١ - مصطفى فهمى : سيكولوجية التعلم (الطبعة الثانية) لجنة النشر للجامعين سنة ١٩٥٣ .

- ٢٥٢ - وزارة التربية والتعليم : دليل الإختبارات للقبول بالمرحلة الاعدادية بإشراف محمد عبد السلام أحمد سنة ١٩٥٥ .
- ٢٥٣ - وزارة التربية والتعليم : تقرير لجنة بحث التعليم الابتدائي . سنة ١٩٥٧ .
- ٢٥٤ - وزارة التربية والتعليم : التقرير السنوى لإدارة البحوث الفنية والمشروعات ١٩٥٩ - ١٩٦٠ .
- ٢٥٥ - وزارة التربية والتعليم : دليل تقويم التلميذ فى المرحلة الابتدائية سنة ١٩٦٠ .

