الفصل الرابع

القياس المعرفى

١- طبيعة القياس المعرفى

٢ - خطوات بناء الاختبارات المعرفية

أولاً : التخطيط للاختبار

ثانيا : كتابة فقرات الاختبار

ثالثا : تطبيق الاختبار

رابعا : تحليل الفقرة

خامساً : مراجعة وتنقيح الفقرة

- معجم إنجليزى - عربى للمصطلحات
- مراجع الفصل الرابع

١ - طبيعة القياس المعرفى

يستهدف القياس بالنسبة للبعد المعرفى فى المجال الرياضى التعرف على التغيرات والتطورات التى تحدث فى المعلومات والفهم للعديد من الموضوعات المتعلقة بتاريخ التربية البدنية، وتاريخ الأنشطة الرياضية، وقواعد التحكيم، والمعلومات الأساسية المتعلقة بالجسم البشرى والتى تؤثر على كفاءته فى النشاط الرياضى، إلى غير ذلك من الموضوعات.

ويشير الفهم فى المجال الرياضى إلى القدرة على استخدام المعلومات والحقائق العلمية المختلفة فى مجال التطبيق العملى، فمثلا يمكننا أن نتوقع من اللاعب الذى يعرف الأوضاع والتحركات الصحيحة فى طريقة «رجل لرجل» الدفاعية فى كرة السلة، أن يقوم بتطبيق هذه الطريقة بشكل جيد أثناء المباريات فى كرة السلة.

مفهوم القياس المعرفى:

يشير القياس المعرفى إلى الاختبارات التى تستخدم لقياس عمليات التفكير thought processes ، حيث يلاحظ أن هذه الاختبارات تستخدم اللغة المكتوبة، بمعنى أنها اختبارات كتابية Written tests، وأنها تستخدم فى المجال الرياضى لقياس مستوى تحصيل الأغراض المعرفية.

وتأخذ الإجابة فى الاختبارات الكتابية نمطين رئيسيين هما :

النمط الأول : وفيه يطلب من المفحوص اختيار الإجابة (الاستجابة) المناسبة من بين عدد من الإجابات، ويتضمن هذا النمط وحدات الاختبار التالية: الصواب - الخطأ، الاختيار من متعدد، المزاوجة.

النمط الثانى : وهو يستلزم إمداد المفحوص بالمعلومات الضرورية التى تمكنه من الإجابة (الاستجابة)، ويتضمن هذا النمط الإجابات القصيرة كالتكميل وكتابة المقال.

متطلبات بناء الاختبارات المعرفية:

لكى يستطيع الباحث أو (القائم بإعداد الاختبار المعرفى) النجاح فى مهمته فإنه يلزم أن يتوافر له الآتى:

١ - أن يكون ملما بالأساليب الصحيحة والمناسبة لبناء الاختبارات.

٢ - أن يكون على دراية بمجال المادة (موضوع التخصص) الذى يعد له الاختبار.

٣ - أن تكون لدية المهارة فى التعبير المكتوب (التعبير كتابة) written expression عن الموضوعات التى يعدلها الاختبار.

٤ - أن يكون على وعى بمستويات فهم المفحوصين اللذين يعد لهم الاختبار.

٥ - الصبر والاستعداد للعمل لفترات زمنية طويلة.

٢ - خطوات بناء الاختبار المعرفى

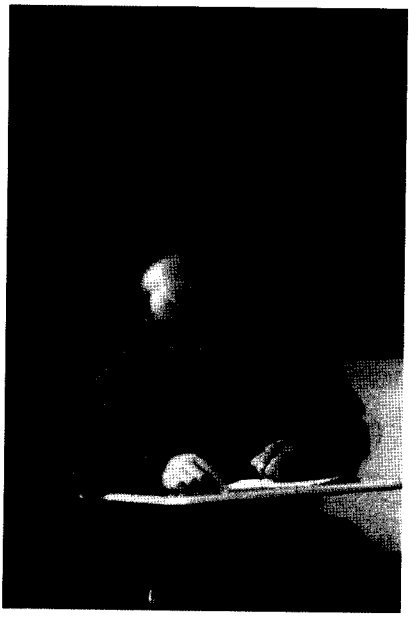
لبناء الاختبار المعرفى تتبع الخطوات الخمس التالية:

- التخطيط للاختبار Planning of the test
- كتابة فقرات الاختبار writing the items of the test
- تطبيق الاختبار test administration
- تحليل الفقرة item analysis
- مراجعة وتنقيح الفقرة item revision

وفيما يلى شرح موجز لكل خطوة من هذه الخطوات:

أولاً : التخطيط للاختبار:

إن أول خطوة يجب أن توضع فى الاعتبار عند بناء أى اختبار هى التخطيط للاختبار، حيث تتضمن هذه الخطوة الآتى:



القياس المعرفى

يمكن استخدام الاختبار لقياس معلومات الطالب عن الأساليب الفنية للعب

١ - أن يحدد القائم بوضع الاختبار مجال عمل (تطبيق) الاختبار والأغراض الخاصة به.

٢ - تحديد المصطلحات التى يمكن أن يتضمنها الاختبار تحديداً دقيقاً وواضحاً مثال ذلك (الذكاء الخطى - اللياقة البدنية، التحصيل الدراسى.. وغيرها).

٣ - أن يحدد المسئول عن إعداد الاختبار مسبقاً طبيعة محتوى الاختبار (الفقرات - العبارات) وطبيعة الاستجابة عليها، مثال ذلك (المقال - صح أو خطأ - التكميل - الاختيار من متعدد .. إلخ).

٤ - تحديد أسلوب العينة اللازمة لبناء الاختبار موضحا نوع العينة، وحجمها، وأسلوب اختيارها، والمجتمع الأصلى للعينة.

٥ - وضع التفاصيل والترتيبات الخاصة بالتطبيق الأولى والتطبيق النهائى للاختبار.

٦ - تقدير الوقت المتوقع أن يستغرقه الاختبار (طول الاختبار - والوقت اللازم لإكمال الاختبار).

٧ - تحديد الأساليب الإحصائية المتوقع التعامل معها لبناء الاختبار.

٨ - تحديد حجم العينة اللازم لإعداد (بناء) معايير الاختبار.

٩ - إعداد الدليل الخاص بالاختبار manual والذى يتضمن التعليمات وشروط تطبيق الاختبار.

وهناك وجهة نظر ترى أنه يجب أن يؤخذ فى الاعتبار صدق المحتوى (المضمون) Content validity بالنسبة لفقرات الاختبار test items كخطوة أولى عند التخطيط لبناء الاختبار، وأن يمثل الاختبار الأغراض التعليمية instructional objectives، التى تتضمنها الوحدة التدريسية أو التدريبية. ولاستكمال صدق المحتوى للفقرات يقترح علماء القياس استخدام ما يعرف بجدول مواصفات الاختبار table of test specifications، حيث يستفاد من جدول المواصفات فى الآتى:

- تحديد أنواع وعدد فقرات (أسئلة) الاختبار.
- تحديد أنواع المهام (الأعمال) tasks بالنسبة لعمليات التفكير المختلفة وأيضا تحديد عدد الفقرات الخاصة بكل عملية من تلك العمليات.
- تحديد العناصر أو المجالات الفرعية للموضوع الرئيسى الذى يعد له الاختبار وعدد الوحدات الخاصة بكل عنصر أو مجال فرعى. ويتضمن الجدول (٤ - ١) مثالاً لجدول المواصفات لاختبار للكرة الطائرة.

جدول (٤ - ١) : المواصفات الخاصة باختبار معرفى للكرة الطائرة

يتكون من ٥٠ فقرة

(المهام أو الأعمال - وعدد الأسئلة - والنسب المئوية)				
المجالات الفرعية	المعرفة	الفهم	التحليل	التطبيق
- التاريخ	٢ (٤٪)			
- قواعد اللعب	٥ (١٠٪)	٥(١٠٪)		٥(١٠٪)
- النواحى الفنية	٢ (٤٪)	٥(١٠٪)	٨(١٦٪)	
- استراتيجية الهجوم	٤ (٨٪)			٥(١٠٪)
- استراتيجية الدفاع	٤(٨٪)			٥(١٠٪)
المجموع	١٧(٣٤٪)	١٠(٢٠٪)	٨(١٦٪)	١٥(٣٠٪) = ٥٠ (١٠٠٪)

ماهيـة جدول المواصفات:

جدول المواصفات عبارة عن مخطط أو رسم أو تصميم يبين الأهمية النسبية لوحدات الاختبار فيما يتعلق بكل مجال من المجالات الفرعية وفقا للفئات الرئيسية للمجال المعرفى وكذا النسب المئوية لعدد الفقرات فى كل مجال من هذه المجالات الفرعية.

ويتكون جدول المواصفات من اتجاهين two-ways ، الأول ويمثل المحور الرأسى ويتضمن العناصر أو المجالات الفرعية للموضوع المطلوب قياسه، والثانى ويمثل المحور الأفقى ويتضمن الفئات الرئيسية للمجال المعرفى (الأغراض التربوية) ففى الجدول (٤ - ١) على سبيل المثال يمثل المحور الرأسى المجالات الفرعية للموضوعات التى سوف يتناولها القياس المعرفى بالنسبة للكرة الطائرة وهى خمسة مجالات فرعية، بينما يمثل المحور الأفقى الفئات الرئيسية لمستويات القياس المعرفى والتى تظهر فى أربعة مستويات هى: المعرفة والفهم والتحليل والتطبيق.

ويستخدم جدول المواصفات كوسيلة مهمة لتأكيد صدق المحتوى (المضمون) للاختبار المقترح حيث يشير هذا الصدق إلى المدى الذى تقيس به وحدات الاختبار عينات مناسبة من الموضوع الدراسى أو القدرات التى صمم الاختبار لقياسها.

فلو طلب منا - على سبيل المثال - القيام بإعداد جدول المواصفات لاختبار معرفى للريشة الطائرة (البادمنتون)، فى هذه الحالة يمكننا اتباع الخطوات التالية:

الخطوة ١: تحديد المجالات الفرعية (العناصر) التى تستوجب أن يشملها الاختبار: ويتم ذلك عن طريق الآتى:

- الرجوع إلى المراجع العلمية المتخصصة.
- تحليل الدراسات والبحوث السابقة المرتبطة.
- استطلاع آراء المتخصصين فى المجال.

الخطوة ٢ : عرض المجالات الفرعية (العناصر) على مجموعة من الخبراء فى مجال التخصص: ويتم كالتالى:

- إعداد استمارة لاستطلاع رأى الخبراء.

- وضع معايير اختيار الخبراء.

- تحديد الخبراء بحيث يكون عددهم مناسبا (توضع قائمة بأسمائهم فى المرفقات).

- عرض الاستمارة على الخبراء وذلك بغرض الآتى:

أ) التأكد من صحة المجالات (العناصر) الفرعية.

ب) اقتراح حذف أو إضافة مجالات (عناصر) أخرى.

جـ) تقدير الأهمية النسبية لكل مجال (عنصر) من المجالات الفرعية.

د) تقدير النسبة لكل مستوى من مستويات المجال المعرفى.

فلو افترضنا أن الباحث توصل نتيجة الخطوة الأولى إلى ثمانية مجالات (عناصر) فرعية هى: التاريخ، القيم، الأجهزة والأدوات، آداب وتقاليد اللعبة، الأمان، القواعد والقوانين المنظمة، الاستراتيجيات، الأساليب الفنية للعب. وعندما قام الباحث بعرض هذه المجالات الفرعية على مجموعة من الخبراء، جاءت الأهمية النسبية للمجالات (العناصر) الفرعية كالتالى :

١ - تاريخ اللعبة	٥٪
٢ - القيم	٥٪
٣ - الأجهزة والأدوات	١٠٪
٤ - آداب وتقاليد اللعبة	١٠٪
٥ - الأمان والسلامة	١٠٪
٦ - القواعد والقوانين	٢٠٪
٧ - استراتيجات اللعب	١٥٪
٨ - الأساليب الفنية للعب	٢٥٪
	١٠٠٪

كما جاءت الأهمية النسبية لكل مستوى من مستويات المجال المعرفى (ستة مستويات) كالتالى:

١ - تذكر المعلومات	٣٠٪
٢ - الفهم	١٠٪
٣ - التطبيق	٣٠٪
٤ - التحليل	٢٠٪
٥ - التركيب	٠٪
٦ - التقويم	١٠٪
	١٠٠٪

الخطوة ٣: يتم وضع النتائج السابقة وفقا لأهميتها النسبية فى جدول المواصفات بحيث تمثل المجالات الفرعية (الثمانية للعبة المحور الرأسى للجدول ، وتمثل مستويات المجال المعرفى (٦ مستويات) المحور الأفقى للجدول. وبناء على هذا يمكننا إعداد جدول للمواصفات (فى اتجاهين) يتكون من ٩ أعمدة، ١٠ صفوف جدول (٤ - ٢) كالتالى:

جدول (٤ - ٢): جدول المواصفات (من اتجاهين) لاختبار معرفى للريشة الطائرة يتكون من ٢٠٠ بند أو عبارة

المجالات (العناصر) الفرعية	الأهمية النسبية	تذكر المعلومات	الفهم	التطبيق	التحليل	التركيب	التقويم	المجموع
		٣٠٪	١٠٪	٣٠٪	٢٠٪	٠٪	١٠٪	١٠٠٪
التاريخ	٥٪	١,٥٪ ٣	٠,٥٪ ١	١,٥٪ ٣	١٪ ٢	٠٪ ٠	٠,٥٪ ١	١٠
القيم	٥٪	١,٥٪ ٣	٠,٥٪ ١	١,٥٪ ٣	١٪ ٢	٠٪ ٠	٠.٥٪ ١	١٠
الأجهزة والأدوات	١٠٪	٣٪ ٦	١٪ ٢	٣٪ ٦	٢٪ ٤	٠٪ ٠	١٪ ٢	٢٠
الآداب والتقاليد	١٠٪	٣٪ ٦	١٪ ٢	٣٪ ٦	٢٪ ٤	٠٪ ٠	١٪ ٢	٢٠
الأمان والسلامة	١٠٪	٣٪ ٦	١٪ ٢	٣٪ ٦	٢٪ ٤	٠٪ ٠	١٪ ٢	٢٠
القواعد والقوانين	٢٠٪	٦٪ ١٢	٢٪ ٤	٦٪ ١٢	٤٪ ٨	٠٪ ٠	٢٪ ٤	٤٠
الاستراتيجيات	١٥٪	٤,٥٪ ٩	١,٥٪ ٣	٤,٥٪ ٩	٣٪ ٦	٠٪ ٠	١,٥٪ ٣	٣٠
الأساليب الفنية	٢٥٪	٧,٥٪ ١٥	٢,٥٪ ٥	٧,٥٪ ١٥	٥٪ ١٠	٠٪ ٠	٢,٥٪ ٥	٥٠
المجموع	١٠٠٪	٦٠	٢٠	٦٠	٤٠	٠	٢٠	٢٠٠

الخطوة ٤ : تحديد الأوزان النسبية relative weights لكل خلية من خلايا الجدول والتى تمثل نقطة التقاء كل مجال فرعى (المحور الرأسى - ٨ مجالات) بكل مستوى من المستويات المعرفية (المحور الأفقى - ٦ مستويات)، فعلى سبيل المثال يمكن تحديد الوزن النسبى للخلية الأولى والخاصة (بتذكر المعلومات التاريخية للريشة الطائرة) كالتالى:

الخليـة الأولى= الوزن النسبى للتاريـخ × الوزن النسبى لتذكر المعلومات.

= ٠,٠٥ × ٠,٣٠ = ٠,٠١٥

= ٠,٠١٥ ÷ ١٠٠ = ١,٥ أى ١,٥٪

ويـمكن تحديـد الوزن النسبى للخليـة الثانيـة كالتالى:

= ٠,٠٥ × ٠,١٠ = ٠,٠٠٥

= ٠,٠٠٥ ÷ ١٠٠ = ٠,٥ أى ٠,٥٪

وهكذا بالنسبة لبقيـة خلايـا الجدول والبالغ عددها ٤٨ خليـة (٦×٨ =٤٨).

ويـوضع الناتج فى داخل الخلايـا وفق ما هو موضح بالجدول (٤ - ٢).

الخطوة ٥ : تحديـد عدد العبارات (الأسئلة أو الفقرات) الخاصة بكل خليـة:

ويـسبـق هذه الخطوة تحديـد العـدد الكلى للعـبارات (الأسئلة) للاخـتبار ككل، فلو افترضـنا أن الباحث قرر أن يـكون عدد عبارات الاخـتبار المعرفى للريـشة الطائرة هـو ٢٠٠ عبارة، فى هـذه الحالة يـتـم تحديـد عـبارات كل خليـة من خلايـا جدول المواصفات (٤٨ خليـة) كالتالى:

عدد عبارات الخليـة الأولى = $\frac{\text{الوزن النسبى لـلخليـة} \times \text{العدد الكلى للعبارات}}{١٠٠}$

$= \frac{٢٠٠ \times ١,٥}{١٠٠}$

= ٣ عبارات

وبالنسبة للخليـة الثانيـة = $\frac{٢٠٠ \times ٠,٥}{١٠٠}$

= ١ عبارة واحدة

وهكذا بالنسبة لبقيـة الخلايـا.

ثانيا : كتابة فقرات الاختبار:

يمثل إعداد الفقرات items الخطوة الثانية من خطوات إعداد اختبارات الورقة والقلم paper and pencil tests ، ووفقا لتعريف بين Bean ١٩٥٣م يتبين أن الفقرة عبارة عن سؤال مفرد أو واجب أو عمل لا يمكن تحليله إلى ما هو أبسط.

تعريف الفقرة:

تعرَّف الفقرة item على أنها عمل أو واجب أو سؤال مفرد لا يمكن تحليله أو تجزئته إلى وحدات أصغر أو أبسط. وتعد الفقرة وحدة الاختبار الأساسية، فكل مقاييس أو اختبار يتكون من عدد من الفقرات (العبارات أو الأسئلة أو البنود).

متطلبات إعداد وصياغة الفقرات:

عندما ينتهى الباحث من التخطيط للاختبار، يبدأ فى إعداد وكتابة الفقرات معتمداً فى ذلك على المصادر التالية:

١ - المراجع العلمية فى مجال التخصص.

٢ - الكتب الدراسية المقررة والمذكرات التى يعدها المتخصصون.

٣ - خبراته السابقة فى المجال الذى يعد له الاختبار.

ولكى ينجح الباحث فى هذه المهمة فإنه يلزم توافر الآتى:

- أن يكون لدى الباحث معرفة وتمكن من الموضوع الذى يكتب عنه، وأن يكون ملما بكل الحقائق facts والمبادئ principles والمفاهيم والمدركات الخاطئة misconceptions المتعلقة بهذا الموضوع.
- أن يكون ملما بطبيعة الأشخاص الذين يعد لهم الاختبار، كأن يكون على علم بمستوى ذكائهم وقدراتهم وميولهم ومستوياتهم

الاقتصادية والاجتماعية، لأن ذلك يمكنه من الاختيار الصحيح للفقرة والصياغة الجيدة لها من حيث الألفاظ المستخدمة ودرجة الصعوبة.

- أن يكون ملماً بالعبارات والألفاظ المناسبة وغير المناسبة، الواضحة وغير الواضحة، المحددة والعامة إلى غير ذلك من العبارات والألفاظ.
- أن تكون لديه حصيلة لغوية (عامة - وخاصة) واسعة، وبحيث تشمل هذه الحصيلة: أصل الكلمة ومرادفاتها، وأن يكون الباحث ملماً بمختلف المعانى الخاصة بالكلمة word التى يختارها، وأن يكون قادراً على صياغة الفقرة بأبسط لغة ممكنة.
- أن يكون لدى الباحث القدرة على الاطلاع والبحث، وكذا القدرة على التصور والتخيل والابتكار حتى يتمكن من اختيار فقرات جيدة.
- بعد أن يقوم الباحث بتحديد وصياغة الفقرات، فإنه يفضل أن يقوم بعرضها على مجموعة من الخبراء لإبداء آراءهم فى تلك العبارات.

وقبل أن يبدأ الباحث فى كتابة الفقرة item writing يلزمه الإجابة على السؤالين التاليين:

السؤال الأول: ما هو نوع الاختبار المطلوب إعداده؟ وهل هذا الاختبار سيستخدم فقرات الصواب والخطأ، أو الاختيار من متعدد، أو المقال، أو التكميل أم كلها مجتمعة؟ وتكمن أهمية تحديد نوع الاختبار، فى كون جميع الاختبارات لا تستخدم نفس البنية أو الشكل، فاختبارات الشخصية والذكاء - على سبيل المثال - تتطلب كل أنواع الاستجابات.

السؤال الثانى: ما هو نمط الاستجابة على الاختبار؟ ففى الاختبارات التى تتطلب الإجابة بصح أو خطأ تستخدم أسئلة الصواب والخطأ، والاختيار من متعدد، والمزاوجة، والمقال.

وفى جميع الحالات يجب على المسئول عن إعداد الاختبار أن يقرر مبكراً عند التخطيط لبناء الاختبار أى الفقرات سوف يختار، ففى حالة اختيار فقرات المقال essay items يكون الاختبار «اختبار مقال - essay test»، وفى حالة اختيار فقرات: الصواب - الخطأ، والاختيار من متعدد، والمزاوجة يكون الاختبار «اختبار موضوعى - objective test».

فقرات اختبار المقال Essay test items

لعل من أهم ما يميز فقرات اختبار المقال أنها تعتمد على أسلوب التقويم الذاتى subjective evaluation، حيث تصمم الفقرات لقياس قدرة الأفراد على استخدام العمليات العقلية العليا higher mental processes فى التعبير عن أنفسهم عن طريق الكتابة بالنسبة لموضوعات تتطلب: تعيين هوية الشىء identifying ، وتفسير أو توضيح الشىء interpreting ، والتوحيد والدمج integrating ، والتنظيم organizing ، والتركيب synthesizing.

ولعل من أهم ما تتميز به فقرات اختبار المقال مايلى:

- سهولة الإعداد، بحيث يمكن وضع العديد من الفقرات فى دقائق معدودة.
- تستخدم لقياس المفاهيم المعقدة complex concepts، والقدرة على التفكير thinking ability ، ومهارات حل المشكلات problem - solving skills.
- تشجع المفحوصين: على كيف يتعلمون بفاعلية ونظام، وكيف يعبرون عن أفكاهم ideas.
- التخمين فى هذا النوع من الاختبارات يكون أقل ما يمكن.

ومن عيوب فقرات اختبار المقال مايلى:

- تتطلب وقتا طويلا عند تقويمها لتقدير الدرجات.
- تقدير الدرجة يتطلب اتخاذ بعض القرارات فيما يتعلق بتسجيل أجزاء الدرجة، مما يؤدى إلى تناقص ثبات هذه الاختبارات.
- نظراً لكونها تستغرق وقتا طويلا فى الإجابة، لذا فقد لا يتمكن المفحوص من الإجابة على كل الفقرات مما يجعلها محدودة كوسيلة لقياس المعرفة.

توجيهات عامة لكتابة فقرات المقال :

لبناء فقرات اختبار المقال بسهولة ويسر، يجب اتباع التوجيهات التالية:

١ - أن تعبر الفقرة عن المعرفة الرئيسية بالنسبة للموضوع الدراسى الذى درسه الطالب (الموضوع المطلوب قياسه)، حيث يلزم أن تستدعى الفقرة المادة الرئيسية فى الكتاب الدراسى الذى درسه الطالب.

٢ - أن يكون للفقرة إجابة واحدة صحيحة، لأنه فى حالة وجود أكثر من إجابة صحيحة للفقرة، فإنه يصبح من الصعب تقويم مستوى معرفة الفرد (المفحوص).

٣ - أن تتضمن الفقرة مجال واتجاهات الإجابة المطلوبة، لأن الفقرات (الأسئلة أو البنود) الغامضة غير الواضحة vague تؤدى إلى تباين واسع فى استجابات المفحوصين مما يجعل عملية التقويم أكثر صعوبة.

٤ - أن تكون جميع الفقرات (الأسئلة) موحدة بالنسبة لكل المفحوصين، بمعنى أن يطلب من كل المفحوصين الإجابة على نفس الفقرات (الأسئلة)، لأنه فى حالة الإجابة على فقرات مختلفة، فإنه يصعب مقارنة درجات المفحوصين بعضهم ببعض، لأنه عندما يختار المفحوص

الفقرات، فإن ذلك يعطيه ميزة تحقيق «أفضل إجابة the best - answer» مما قد يترتب عليه أن يصبح مدى الدرجات على الاختبار محدوداً وبالتالى تناقص ثبات درجات الاختبار.

٥ - إظهار الزمن (الوقت) التقريبى المخصص للإجابة على كل فقرة (سؤال)، وكذا الدرجات المخصصة لكل فقرة، لأن العديد من المفحوصين يحتاجون إلى دليل يرشدهم إلى الزمن الذى يستغرقونه للإجابة على كل سؤال.

٦ - وفى جميع الحالات يفضل استخدام عدد مناسب من فقرات المقال القصيرة short essay items فى مقابل عدد قليل من الفقرات الطويلة، لأن الفقرات القصيرة تكون أقل فى الغموض ambiguous بالمقارنة بالفقرات الطويلة مما يسهل على المفحوص فهمها.

كيفية تقدير درجات فقرات المقال:

لتقدير درجات فقرات اختبارات المقال، يلزم اتباع التوجيهات التالية:

- الاتفاق على طريقة موحدة لتقدير درجات الاختبار. لأن بعض المعلمين يركزون على النقاط الرئيسية التى يجب أن يتضمنها الاختبار، فى حين يركز البعض الآخر على ترتيب الفقرات وفقا لنوعية الاستجابة quality of response .

- يفضل تقويم نفس الفقرة same item بالنسبة لكل المفحوصين قبل الانتقال للفقرة التالية. كما يفضل أن يقوم مقدر الدرجات بمراجعة اتساق تقديره لدرجات الفقرة عن طريق الرجوع إلى أول مجموعة من الأوراق قام بتصحيحها.

- يفضل عدم ذكر اسم المفحوص على ورق الإجابة، لأن تقويم درجات المفحوص على الاختبار بدون أن يكون اسمه معروفا للقائم بعملية التقويم إجراء يكفل الحد من التحيز biases.

أمثلة لفقرات المقال:

اقرأ كل فقرة بعناية ثم أجب على كل واحدة فيها إجابة كاملة قدر الإمكان. الوقت التقريبى اللازم للإجابة على كل فقرة والدرجة المخصصة لها موضح بين الأقواس.

١ - عرِّف الارتباط، ثم وضح أوجه الاختلاف بين معامل ارتباط سبيرمان لفروق الرتب ومعامل الارتباط التتابعى لبيرسون (١٠ دقائق - ٢٠ درجة).

٢ - عرف الدرجة المعيارية، ثم اكتب المعادلات التى تستخدم لحساب كل من الدرجة التائية ت -T ، والدرجة ذ - Z ، والمئينيات Percentiles (١٠دقائق - ٢٠ درجة).

٣ - أراد أحد المعلمين أن يتعرف على فعالية طريقته الجديدة فى التدريس. اشرح كيف يستطيع هذا المعلم التحقق من فعالية هذه الطريقة، على أن يتضمن هذا الشرح الآتى:

- إجراءات القياس التى يجب أن يستخدمها المعلم.
- متى تتم عملية القياس.
- التحليل الإحصائى الذى يجب أن يستخدمه المعلم (١٥دقيقة - ٣٠ درجة).

فقرات الاختبار الموضوعى Objective test items

يعتقد بعد المتخصصين أن الاختبارات الموضوعية يمكن أن تستخدم لقياس المستويات المعرفية البسيطة، وأنها لا تصلح لقياس عمليات التفكير بما تتضمنه من مستويات الفهم والتحليل والتطبيق، وأن اختبارات المقال هى المقاييس المناسبة لعمليات التفكير. إلا أن مجموعة كبيرة من علماء القياس يقررون أن

الفقرات الموضوعية objective items يمكن أن تستخدم لقياس أنواع مختلفة من عمليات التفكير، إلا أن إعداد مثل هذا النمط من الفقرات يستغرق وقتاً وجهداً كبيرين.

أنماط فقرات الاختبارات الموضوعية:

تتضمن الاختبارات المعرفية الموضوعية مجموعة مختلفة من أنماط الفقرات هى : الصواب - الخطأ، الاختيار من متعدد، التكميل، المزاوجة، وفيما يلى شرح موجز لكل نمط من هذه الأنماط.

١) فقرات الصواب - الخطأ The true - false items

وهى فقرات تقدم للمفحوص فى شكل جملة أو عبارة أو مشكلة أو مسألة ويطلب منه أن يحدد صحة الفقرة بوضع علامة (صح-✓) أو (خطأ - ×) أو وضع خط أسفل كلمة صح أو خطأ.

وتستخدم فقرات الصواب - الخطأ فى الاختبارات المعرفية عند قياس تذكر المعلومات عن طريق اختيار الإجابة الصحيحة لحقائق أو مصطلحات أو مبادئ محددة فى مجال محدد.

معنى هذا أن فقرات الصواب - الخطأ عبارة عن صياغات (جمل) تقريرية صريحة تعطى للمفحوص لكى يقرر ما إذا كانت الجملة صحيحة أو غير صحيحة. ويلاحظ أن هذا النمط من الفقرات يستخدم على نطاق واسع فى الاختبارات التى يعدها معلم الفصل teacher-made tests لسرعة كتابة الفقرات وسهولة تقدير الدرجة.

مميزات فقرات الصواب - الخطأ:

- تستخدم لتغطية مدى كبير وواسع من المادة الدراسية عن طريق إجراء القياس لمرة واحدة فقط. فالوقت المطلوب لكى يستجيب المفحوص على فقرات الصواب - الخطأ أقل من الوقت المطلوب للاستجابة على فقرات

الاختيار من متعدد أو التكميل، لذا يمكن تضمين اختبارات الصواب - الخطأ عدد كبير من الفقرات يفوق أى نمط آخر من الاختبارات.

- سهولة التصحيح وتقدير الدرجات.
- سهولة إعداد الفقرات وتطبيقها وبخاصة عند استخدامها لقياس تذكر المعلومات، إلا أنها قد تستغرق وقتا أطول عندما تعد لقياس عمليات التفكير المركبة.

عيوب فقرات الصواب - الخطأ:

- تتأثر بالتخمين العشوائى random guessing لكونها تتضمن إجابتين محتملتين فقط، حيث ينتج عن هذا التخمين العشوائى أن يصبح لدى المفحوص فرصة ٥٠٪ للحصول على إجابات صحيحة نتيجة التخمين.
- نظراً لوجود فرصة للحصول على ٥٠٪ من الإجابات الصحيحة نتيجة التخمين، لذا يتجه ثبات درجات اختبارات الصواب - الخطأ نحو الانخفاض.
- غالبا ما تتوقف (تعتمد) الإجابة الصحيحة على كلمة واحدة فقط مما يزيد من احتمالات الوقوع فى الخطأ.

توجيهات عامة لكتابة فقرات الصواب - الخطأ:

يمكن استخدام فقرات الصواب - الخطأ بفاعلية إذا اتبع عند إعدادها الآتى:

- تجنب استخدام العبارات التى توحى بإجابات معينة، وتجنب استخدام التعميمات فى الزمان أو المكان مثل: دائما - أبدا - فى كل وقت - أحيانا - عادة إلى غير ذلك من الكلمات.
- يجب أن يكون عدد الفقرات الصواب مساويا لعدد الفقرات الخطأ، أو أن يكون عدد الفقرات الخطأ أكبر من عدد الفقرات الصواب، لأن للفقرات الخطأ قدرة على التمييز أكبر من الفقرات الصواب.

- تجنب استخدام لغة الكتاب المدرسى لكونها لغة جافة قد تتميز بالصرامة.
- تجنب استعمال الصيغ السلبية كأن تقول «أنا لا أنام فى الليل كثيراً..» ولكن الأفضل فى الصياغة هو « أنا أنام فى الليل قليلا...»، وذلك لكون الصياغات السلبية ربما تربك المفحوص، وقد لا تسهم بفاعلية عند قياس القدرات. وينصح فى حالة الحاجة إلى استخدام العبارات السلبية أن يوضع خط أسفل المصطح السلبى كالتالى «أنا لا أنام فى الليل كثيرا..».
- تجنب استخدام الكلمات الغامضة والتى قد تحمل أكثر من معنى مثل كلمة «ثقافة» لكون مثل هذه الكلمات تثير الشك والحيرة لدى المفحوص، فالأساس فى الصياغات (الجمل) بالنسبة لمثل هذا النوع من الفقرات هو الآتى:

«ألا تحمل الجملة أى شك فى أنها صحيحة تماماً أو خطأ تماماً».

- العمل على أن يكون لجميع الفقرات نفس الطول تقريبا، حيث يلاحظ فى بعض الاختبارات المعرفية أن الجمل (الفقرات) التى لها إجابات صحيحة تكون أطول من الجمل التى لها إجابات خاطئة.
- أن تسأل الفقرة (السؤال أو الجملة) عن مفهوم أو موضوع أو فكرة واحدة مفردة فقط. فالفقرات التى تشتمل على أكثر من مفهوم أو موضوع أو حقيقة علمية غالبا ما تؤدى إلى إرباك المفحوصين.

أمثلة لفقرات الصواب - الخطأ:

مثال (٤ ـ ١): اقرأ بعناية كل جملة من الجمل التالية، ثم ضع علامة صح (✓) أمام الجملة التى تعتقد أنها صحيحة، وعلامة خطأ (×) أمام الجملة التى تعتقد أنها خطأ (لكل إجابة صحيحة ٣ درجات).

- ١) إذا زاد الانحراف المعيارى (SD) لدرجات المجموعة، يزيد بالتالى التشتت.

- ٢) الدرجة المعيارية ت - T التى تساوى ٦٠ تقابل وحدة انحراف معيارى واحدة فوق المتوسط الحسابى.

- ٣) معامل الارتباط +٦٠,٠ يساوى مرتين معامل الارتباط +٣٠,٠.

- ٤) إذا كان منحنى توزيع البيانات متجها نحو اليمين كان الالتواء سالباً.

مثال (٤-٢): الجمل التالية تصف بعض المواقف فى رياضة التنس. والمطلوب: هو أن تقرأ الجمل جيداً، ثم تضع علامة صح(✓) أمام الجملة التى تعتقد أنها صحيحة، وعلامة خطأ أمام الجملة التى تعتقد أنها خطأ (×) (لكل إجابة صحيحة ٣ درجات):

«اللاعب (أ) قام بأداء إرسال جيد، فقام اللاعب (ب) برد الإرسال بنجاح، فأى اللاعبين يمكن أن يفقد (نقطة واحدة) إذا:

- ١) ارتدت الكرة مرتين فى ملعبه.

- ٢) اللاعب ضرب الكرة فلمست الشبكة قبل الهبوط فى ملعب خصمه.

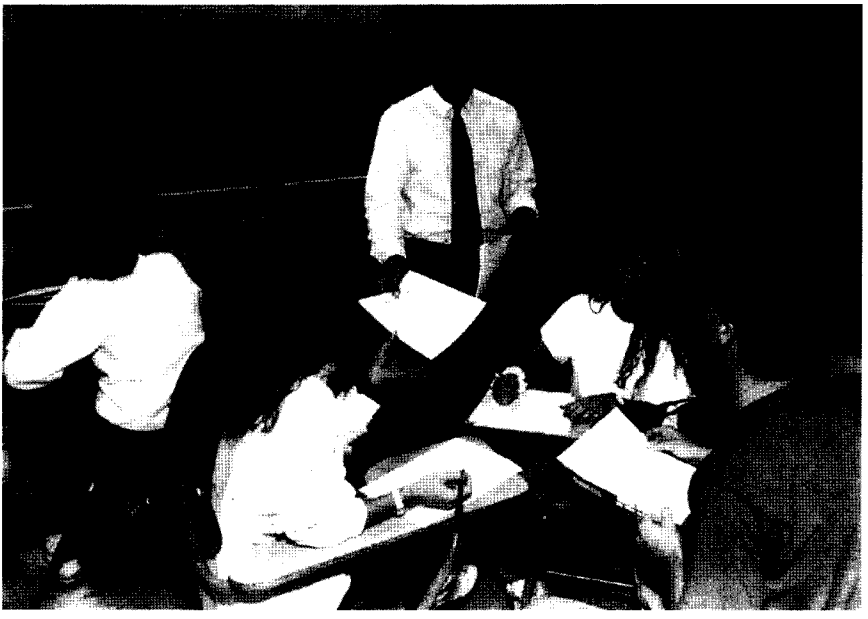
- ٣) بعد ضرب الكرة سقط المضرب من يد اللاعب.

- ٤) اللاعب ضرب الكرة قبل أن تعبر الشبكة فى اتجاه ملعبه.

٢) فقرات الاختيار من متعدد

The Multiple - Choice items (MCq)

وفيها يعطى المفحوص السؤال يتبعه عدة إجابات إحداها فقط صحيحة، وعليه أن يختار الإجابة التى يعتقد هو أنها صحيحة. فالإجابات فى هذا النمط من الاختبارات يكون فيها إجابة واحدة صحيحة وعدد من الإجابات



القياس المعرفى يستهدف قياس تحصيل الأهداف المعرفية

اختبارات الاختيار من متعدد يمكن تطبيقها على أعداد كبيرة من المفحوصين

غير صحيحة incorrect responses والتى تعرف بالإجابات «المضللة dis tractor» وهى إجابات تعد بحيث تغرى المفحوص باختيارها فى اختبارات الاختيار من متعدد أو اختبارات المزاوجة.

وتستخدم فقرات الاختيار من متعدد لقياس نواتج التعلم بالنسبة لمستويات متبانية من المعرفة بدءا بالمستويات البسيطة وحتى المستويات المركبة، كما تستخدم عند قياس التحصيل بالنسبة للعديد من الموضوعات الدراسية.

وتتطلب فقرات الاختيار من متعدد - فى بعض الأحيان - أن يختار المفحوص «أفضل إجابة - the best answer » من بين مجموعة من الإجابات الصحيحة التى تتطلب من المفحوص اختيار أفضلها، حيث يعنى ذلك أنه يطلب من المفحوص أن يختار من بين مجموعة من الاستجابات الصحيحة أفضل الاستجابات المناسبة للإجابة على السؤال الذى تتضمنه الفقرة.

ومن الملاحظ أن فقرات الاختيار من متعدد MCq تتضمن مكونين رئيسيين هما:

١ - الأرومة (البند الرئيسى) Stem : وهى عبارة عن سؤال مباشر أو جملة غير تامة (غير كاملة) تعرض لمشكلة تفاصيلها غير كافية ولا يوجد بها أى غموض فيما يتعلق بالشىء الذى تسأل عنه.

٢ - الاستجابات Responses: وهى تتراوح من ٣ إلى ٥ استجابات واحدة منها صحيحة وبقية الاستجابات الأخرى غير صحيحة، حيث يلاحظ أن الاستجابات غير الصحيحة تشير إلى إجابات خاطئة تغرى المفحوص باختيارها، حيث تعرف فى مجال القياس المعرفى باسم «المضلل أو المشتت - Distractor».

مميزات فقرات الاختيار من متعدد:

- تستخدم لقياس القدرة على الفهم.
- تصلح لقياس أنماط مختلفة من المواد الدراسية.
- لا تستغرق وقتا طويلا فى التطبيق وتصحيح النتائج.
- فرص التخمين بالنسبة للإجابات الصحيحة أقل بكثير من فقرات الصواب - الخطأ.
- تستخدم كثيراً فى الاختبارات المعرفية فى مجالات: الرياضيات والهجاية ومفردات اللغة وغيرها.

عيوب فقرات الاختيار من متعدد:

- صعوبة إعدادها بالمقارنة بالفقرات الأخرى، حيث يستغرق إعدادها وقتا أطول وذلك لكونها تتطلب من (٣ - ٥) استجابات لكل فقرة (سؤال أو جملة).
- تشجع - فى بعض الأحيان - على استظهار (الحفظ عن ظهر قلب) memorization الحقائق أكثر من فهم understanding تلك الحقائق.
- تمتاز اختبارات الاختيار من متعدد بأنه يمكن تضمينها عدد أقل من الفقرات بالمقارنة باختبارات الصواب - الخطأ، وذلك عند قياس نفس الموضوع أو المادة الدراسية، لذلك فهى تستغرق وقتا أقل.

توجيهات عامة لكتابة فقرات الاختيار من متعدد:

مع أن إعداد فقرات الاختيار من متعدد من الصعوبة بمكان، إلا أنه يمكن الاسترشاد بالتوجيهات التالية عند إعدادها:

أولاً : بالنسبة للأرومة (البند الرئيسى) Stem :

١- يجب أن يكون البند الرئيسى مختصرا وسهل القراءة والفهم، وأن يحتوى على المسألة أو القضية الرئيسية للفقرة، كما يشترط عدم

تكرار أية كلمات يتضمنها البند الرئيسى فى الاستجابات، كما ينصح بتجنب استخدام أية مواد (موضوعات) غير ذات صلة irrelevant بالأرومة (البند الرئيسى).

كما يفضل أن تصاغ الأرومة فى جملة أو سؤال غير كامل حتى يتمكن المفحوص من تكميله عن طريق الاستجابة.

٢- يجب تجنب التعميمات مثل: دائما - مطلقا - أبدا - كلية وغيرها.

٣- يجب تحديد الأرومة (البند الرئيسى) بكلمات وتعبيرات موجبة Positive words ، وفى حالة الكلمات السالبة فإنه يفضل وضع خط أسفل هذه الكلمات أو كتابتها ببنط كبير يميزها.

٤- يفضل أن تبدأ الأرومة (البند الرئيسى) بالحرف «W» أى بكلمات مثل: أى - أية which، لماذا why ، أين where ، ما - ماذا what، متى when، من who ، لأن مثل هذه الكلمات تقدم للمفحوص الموضوع الرئيسى أو النقاط الرئيسية للفقرة أو السؤال.

٥- لا يجب تضمين الأرومة (البند الرئيسى) أية كلمات تستهدف سؤال المفحوص عن رأيه opinion، وإنما المطلوب هو أن يحدد الإجابة الصحيحة.

ثانيا: بالنسبة للاستجابات:

١ - يجب أن تكون جميع الاستجابات معقولة وقابلة ظاهريا للتصديق plausible ، وأن توجد بين هذه الاستجابات استجابة واحدة فقط صحيحة، كما يجب عدم تضمين الاختبارات أية استجابات هزلية أو ضعيفة.

٢ - يجب ألا يقل عدد الاستجابات عن ثلاث استجابات وألا يزيد عن خمس استجابات بالنسبة لكل فقرة، وفى جميع الحالات يفضل ألا يقل عدد الاستجابات عن أربع استجابات لأن هذا العدد يجعل عامل التخمين guesing factor عند حدود ٢٠٪ .

٣- يجب أن تكون جميع الاستجابات متسقة فى المحتوى والصياغة وقواعد النحو ولها نفس الطول تقريبا. ويفضل تجنب تضمين الاستجابة الصحيحة معلومات أكثر من الاستجابة الخاطئة.

٤- إذا أعطيت الفقرات الأرقام (١, ٢, ٣, ٤، . . . إلخ)، فإنه يلزم استخدام الحروف (أ، ب، جـ، د، هـ. . .) كأسلوب لترقيم الاستجابات، كما يفضل أن توضع الاستجابات مرتبة رأسيا لأن هذا الأسلوب يساعد على قراءتها بسهولة أكثر من وضعه فى نظام الترتيب الأفقى.

٥- تجنب إعطاء الاستجابات الصحيحة نمط ثابت من الترقيم، كأن تعطى جميع الاستجابات الصحيحة فى الاختبار الحرف (جـ) ولكنه يفضل أن يتم توزيع جميع الاستجابات الصحيحة فى الاختبار على الحروف (أ، ب، جـ، د، هـ) بالتساوى تقريبا.

أمثلة لفقرات الاختيار من متعدد:

ضع الحرف الذى يدل على الإجابة الصحيحة على الخط الأسود الموجود على يمين رقم الفقرة (كل فقرة لها درجتان).

- ١) فى اختبار معرفى، كان للدرجة ٥٥ رتبة مئينية ٤٠٪ ما دلالة ذلك؟

أ - أن ٤٠ تليمذا أجابوا على ٥٥٪ من أسئلة الاختبار إجابات صحيحة.

ب- أن ٤٠ ٪ من التلاميذ لديهم الدرجة ٥٥ أو أقل.

جـ- أن ٥٥٪ من التلاميذ لديهم الدرجة ٤٠ أو أقل.

د - أن ٤٠٪ من التلاميذ لديهم الدرجة ٥٥ أو أكثر.

- ٢) إذا كان معامل الارتباط بين متغيرين هو - ٠,٨٠ فمعنى ذلك أن:

أ - العلاقة بين المتغيرين منخفضة.

ب - المتغيرين غير مترابطين.

جـ - النقص فى قيم أحد المتغيرات يتبعه نقص فى قيم المتغير الآخر.

د - الزيادة فى قيم أحد المتغيرات يتبعها نقص فى قيم المتغير الآخر.

- ٣) أى العضلات تعمل على مد الساق وقبض الفخذ ؟

أ - العضلة التوأمية (خلف الساق).

ب - العضلة الفخذية ذات الرأسين.

جـ - العضلة الإلية الكبيرة.

د - العضلة الفخذية ذات الأربع رؤوس.

-٤) ما هو بالتقريب الحد الأقصى لمعدل القلب لفرد عمره ٢٠ سنة؟

أ - ١٨٠ نبضة.

ب- ١٩٠ نبضة.

جـ- ٢٠٠ نبضة.

د - ٢١٠ نبضة.

٣) فقرات الإجابة القصيرة والتكميل

Short - answer and completion items

يكمن الفرق بين فقرات الإجابة القصيرة والتكميل فى نقطتين أساسيتين هما: الطول والشكل (البنية) format. فالفقرات فى الإجابة القصيرة تتطلب من المفحوص أن يجاوب على السؤال فى كلمة word أو جملة sentence أو بهما معاً. بينما فى فقرات التكميل تحذف كلمة أو عدد من الكلمات من الجملة، ويطلب من المفحوص استكمال الكلمات الناقصة. وتصلح فقرات الإجابة القصيرة وفقرات التكميل لقياس المعرفة الحقيقية factual knowledge والفهم والقدرة على تحديد وتعريف المفاهيم والمدركات.

ففى فقرات الإجابة القصيرة يطلب من المفحوص تحديد أو تسمية شىء ما فى كلمة أو جملة أو رقم، ومن أمثلة ذلك ما يأتى:

- من هو أول رئيس للجمهورية فى مصر؟
- $\frac{١}{٨} + \frac{١}{٤} + \frac{١}{١٦} =$
- ما هى الغدة التى تنتج هرمون ACTH فى جسم الإنسان؟

مميزات فقرات الإجابة القصيرة والتكميل:

- أقل تأثراً بالتخمين من الصواب - والخطأ أو الاختيار من متعدد.
- تصلح لتقويم تذكر المعلومات أكثر من أى نمط آخر من الفقرات.
- يسهل إعدادها.

عيوب فقرات الإجابات القصيرة والتكميل:

- يستغرق حساب الدرجات فيها وقتا أطول من فقرات الاختيار من متعدد، لكونها تتطلب تصحيح مدى دقة وسلامة اللغة المستخدمة فى الإجابة.
- قد تكون بعض الإجابات صحيحة ككل أو صحيحة جزئيا مما يضع مقدر الدرجة فى حيرة بالنسبة لتقدير درجات كل فقرة، وبناء على ذلك يرى علماء القياس أن الشخص القادر على تقدير الدرجات على اختبارات الإجابات القصيرة والتكميل هو الشخص الذى يقوم ببناء الاختبار.
- تشجع على التعلم الصمِّى rote Learning وهو تعلم يقوم على التكرار بقصد الاستظهار (الحفظ) دون الاهتمام بالفهم وغيره من المستويات المعرفية، ومع أن هذا النمط من الفقرات يتضمن العديد من الفرص الملائمة للاسترجاع والحفظ عن ظهر قلب memorization فإننا نجدها

مناسبة لإعداد الاختبارات بالنسبة لبعض المجالات مثل الإسعافات الأولية وبالنسبة لعمليات إنقاذ الغرقى والإنعاش فى حالات الأزمات القلبية.

توجيهات عامة لكتابة فقرات الإجابات القصيرة والتكميل:

تتميز فقرات اختبارات الإجابات القصيرة والتكميل بأنها أسهل فى إعدادها من غيرها من فقرات الاختبارات الموضوعية الأخرى، ومع ذلك فإنه يمكن الاسترشاد بالتوجيهات التالية عند إعداد مثل هذا النوع من الفقرات:

١- يجب العمل على أن تكون الإجابة على الفقرة بكلمة واحدة مفردة أو بعدد من الكلمات وأنه توجد إجابة واحدة صحيحة فقط.

٢- التأكد من أن جميع المفحوصين يعرفون ما هو نمط الاستجابة المطلوب، وكذا المؤشرات الخاصة بدقة الاستجابة وبخاصة فى الحالات التى تستخدم فيها أجزاء الدرجة أو الكسور العشرية.

٣- يجب على من يقوم بإعداد مثل هذا النمط من الاختبارات أن يفكر أولا فى الإجابة، ثم يحاول بعد ذلك كتابة الفقرة التى تمثل الإجابة التى أعدها، فعن طريق هذا الإجراء يمكن تجنب إعداد فقرات لها أكثر من إجابة صحيحة.

٤- فى فقرات التكميل، يفضل تخصيص منطقة فراغ خالية من الكتابة قرب نهاية الجملة، حيث يفيد هذا الإجراء فى جعل الجملة أكثر وضوحاً وتجنب وضع أكثر من إجابة للجملة.

٥- يجب عدم استخدام أكثر من منطقتين (فراغين) فى الجملة الواحدة، لأن وجود أكثر من فراغين فى الجملة يزيد من غموضها مما يؤدى إلى اضطراب المفحوص.

٦- عدم نقل الفقرات مباشرة من الكتاب المدرسى، لأن الجملة التى تؤخذ من فقرة فى كتاب قد لا تعبر بشكل مناسب عن المحتوى المراد التعبير عنه.

٧- أن تكون جميع الفراغات متساوية الطول بالنسبة لجميع الاستجابات، لأن الاختلاف فى أطوال الفراغات وفقا لأطوال الإجابات المتوقعة، قد يعطى المفحوص إيحاءات بأهمية الإجابة، ولتسهيل تقدير الدرجة والحد من الأخطاء، يجب وضع خطوط سوداء قصيرة فى الفراغات الخاصة بكل جملة لكى يسجل عليها المفحوص استجاباته.

أمثلة لفقرات الإجابة القصيرة:

١- ما هى المقاييس الثلاثة للنزعة المركزية؟

٢- أظهرت دراسة لقوة الرجل وقوة القبضة ارتباط مقدارة (-٠,٩٠) فسر هذا الارتباط؟

٣- فى توزيع للدرجات له متوسط حسابى ٢٢ وانحراف معيارى ٢، ما هى الدرجة المئوية للدرجة ١٦ والدرجة ٢٨؟

أمثلة لوحدات التكميل:

أكمل الفقرات التالية بكتابة الإجابة الصحيحة على الخط الأسود الموجود داخل كل فقرة (كل فقرة لها درجتان)

- ١) يستخدم ـــــــ كواحد من أهم مقاييس التشتت

- ٢) الدرجة التائية ت - T لها متوسط حسابى هو ـــــــ وانحراف معيارى هو ـــــــ .

- ٣) المقياس Z-Scale له متوسط حسابى هو ـــــــ وانحراف معيارى هو ـــــــ .

٤) فقرات المزاوجة (المواءمة) Matching items

تعتمد اختبارات المزاوجة (المواءمة أو التوفيق) على المواءمة بين عدة فقرات (بنود) فى قائمتين، حيث يطلب من المفحوص أن يربط بين كل فقرة

فى القائمة الأولى مع ما يناسبها من القائمة الثانية، ويحدد السؤال طبيعة المواءمة، وهل هى ترادفاً أو تضاداً أو علاقة سببية أو علاقة زمنية أو علاقة مكانية أو غيرها.

معنى ذلك أن اختبار المزاوجة matching test يتكون من عمودين (أيمن، وأيسر) بكل عمود مجموعة من الفقرات (الكلمات أو الجمل) التى تستخدم كمثيرات، حيث يلاحظ أن الفقرات فى العمود الأيمن متزاوجة paired مع الفقرات فى العمود الأيسر. وتمثل الفقرات فى الجهة اليمنى المقدمات المنطقية premises للسؤال، بينما تمثل الفقرات فى الجهة اليسرى الاستجابات على السؤال.

وتشبه فقرات المزاوجة (المواءمة) فقرات الاختيار من متعدد، فى كونهما يعطيان المفحوص حق الاختيار فيما بين مجموعة من البدائل alternatives كما أنها تشبه فقرات الإجابة القصيرة لكونها تتطلب استخدام معلومات حقيقية ومختصرة ومحددة، هذه المعلومات قد تكون أسماء أو تواريخ أو أرقاماً أو ألقاباً.

مثال (٤ - ٣):

أمامك قائمتان، القائمة اليمنى تتضمن بعض الأساليب الإحصائية، بينما تتضمن القائمة اليسرى أسماء العلماء، الذين ابتكروا هذه الأساليب الإحصائية.

والمطلوب: المزاوجة بين اسم العالم والأسلوب الإحصائى الذى ابتكره عن طريق توصيل خط بينهما.

(ب) معامل الارتباط التتابعى (ppm)	(أ) فيشر
(جـ) معامل ارتباط فروق الرتب (Rho)	(ب) بيرسون
(د) معامل الاتفاق المطابقة (W)	(جـ) سبيرمان
(أ) النسبة الفائية (F-R)	(د) كندال

يلاحظ من المثال السابق وجود قائمتين، الأولى تتضمن مجموعة من المقدمات، والثانية تتضمن أسماء مجموعة من العلماء، حيث يلاحظ أن القائمتين متساويتين فى العدد.

وعادة ما يطلب من المفحوصين المزاوجة بين الفقرات فى القائمتين، هذه الفقرات قد تكون أحداث وتواريخ، أو أحداث وأسماء ، أو كتب وأسماء مؤلفين، أو مصطلحات وتعاريف إلى غير ذلك. لهذا تعد فقرات المزاوجة (المواءمة) من الوسائل التى تقيس قدرة المفحوص على ربط شىء بآخر.

ومن مميزات فقرات المزاوجة مايأتى:

- يمكن الاستفادة منها فى قياس كمية كبيرة من المعلومات الحقيقية المرتبطة بالمادة الدراسية فى وقت قصير.
- يسهل إعدادها وتقدير درجات الاستجابة عليها بالمقارنة بفقرات الاختيار من متعدد.
- تثير دافعية المفحوصين نحو المعرفة المتكاملة عن طريق التفكير فى العلاقة فيما بين الفقرات.
- يظهر التخمين بالنسبة للإجابات الصحيحة أقل بالمقارنة بالأنماط الأخرى من الفقرات.

ومن عيوب فقرات المزاوجة ما يأتى:

- تستخدم فقط لقياس المعرفة الحقيقية المتعلقة بالحقائق والأسماء والأرقام الجافة.
- محدودة الاستخدام بالنسبة للموضوعات الدراسية، لكونها لا تصلح بالنسبة لكل الموضوعات.
- المغالاة فى التركيز على الحفظ عن ظهر قلب للعلاقات البسيطة والاستجابات الصحيحة لحقائق علمية.

توجيهات عامة لكتابة فقرات المزاوجة (المواءمة):

لبناء فقرات المزاوجة بكفاية ووضوح تتبع الإرشادات التالية:

١- يجب أن يكون عدد الإجابات البديلة (الاستجابات أو الاختيارات) أكبر من عدد المقدمات (فى العمود الأول) وذلك لضمان الحد من احتمالية تأثير المزاوجة وبخاصة بالنسبة للزوجين الأخيرين من الوحدات، فلو كان عدد الوحدات فى القائمة الأولى أربع وحدات، وعدد الاختيارات فى القائمة الثانية أربع اختيارات، واستطاع المفحوص أن يعرف الإجابات الصحيحة لأى ثلاثة أزواج من الفقرات، حينئذ يستطيع تلقائيا أن يجيب الإجابة الصحيحة على الزوج الأخير من الفقرات، أما فى حالة ما إذا كان عدد الفقرات فى القائمة الثانية (قائمة الاستجابات أو الاختيارات) خمس أو ست فقرات مثلا، فى هذه الحالة تقل احتمالية التصحيح الذاتى للفقرة.

٢- يجب أن تكون أسس المزاوجة والاختيارات الخاصة بكل فقرة واضحة فى التعليمات المعطاة، بمعنى أن توضح التعليمات المعطاة للمفحوص عما إذا كان لكل فقرة إجابة واحدة صحيحة، أو أكثر وأيضا كيفية عمل المزاوجة بين الفقرات.

مثال (٤ - ٤):

أمامك قائمتان ، الأولى تتضمن بعض الدول ، والثانية تتضمن الألعاب الرياضية التى قامت بابتكارها هذه الدول. والمطلوب المزاوجة بين أسماء الأنشطة وأسماء الدول التى ابتكرتها وذلك عن طريق اختيار الحرف المناسب من العمود الثانى ووضعه كتابة على الخط الموجود أمام الإجابة المناسبة فى العمود الأول (كل إجابة صحيحة لها ٣ درجات).

(العمود الأول)		(العمود الثانى)
أ، ب	١) الصين	أ - المصارعة
جـ	٢) الأغريق القدماء	ب - الملاكمة
و	٣) الهند	جـ - الميدان والمضمار
د	٤) ألمانيا	د - الجمباز
هـ ، ز	٥) انجلترا	هـ - كرة القدم
		و - الريشة الطائرة
		ز - التنس

يلاحظ من هذا المثال الآتى:

- عدم تساوى الفقرات فى العمودين (٥ فقرات فى العمود الأول) فى مقابل (٧ اختيارات فى العمود الثانى).
- تتطلب طريقة المزاوجة فى هذا المثال كتابة الحرف المعبر عن الاختيار المناسب على الخط الموجود أمام الفقرة.
- يلاحظ أنه يوجد أكثر من اختيار واحد للفقرة، فالإجابات الصحيحة لهذا السؤال تبين أن الفقرة (١) لها إجابتان صحيحتان هما (أ، ب) ، وأن الفقرة (٥) لها أيضاً إجابتان صحيحتان هما (هـ، ز)، وأن لكل من الفقرات أرقام (٢, ٣, ٤) إجابة واحدة صحيحة.

٣ - يجب وضع كل الفقرات والاختيارات (الاستجابات) فى صفحة واحدة وفى عمودين متقابلين، ولا يجب بأى حال من الأحوال توزيعها على صفحتين، لأن تنظيمها فى عمودين فى صفحة واحدة يساعد المفحوص على الإجابة فى أقصر وقت، كما يحد من وقوعه فى الخطأ نتيجة التقليب فى الصفحات.

٤ - يجب استخدام الشكل المناسب لتقديم فقرات المزاوجة، حيث يفضل أن توضع الفقرات فى عمود على اليمين، وتوضع الاختيارات (الاستجابات) فى عمود على اليسار، ولتسهيل عملية الاستجابة، وتقدير الدرجة، يعمل خط أسود أمام الفقرة ليكتب عليه الحرف الذى يعبر عن الاختيار المناسب.

٥ - يتم ترتيب الاختيارات (الاستجابات) بإعطائها الحروف الأبجدية (أ، ب، جـ، . . إلخ). أو أى ترتيب منطقى آخر.

٦ - يفضل أن يكون هناك أكثر من اختيار (استجابة) لكل فقرة، فوجود اثنين أو ثلاثة أو أكثر من الاختيارات بالنسبة للفقرة الواحدة يحد من تأثير التخمين guessing.

٧ - يفضل أن يكون عدد الفقرات فى القائمتين مناسباً للمرحلة العمرية التى يعطى لها الاختبار، فيكون فى المرحلة الابتدائية من ٥ إلى ٦ فقرات، ويزيد هذا العدد مع تقدم السن ليصبح من ١٠ إلى ١٥ فقرة، وفى جميع الحالات يجب ألا يكون عدد الفقرات والاختيارات طويل جداً، وألا يكون قصير جداً، لأنه فى حالة ما إذا كان عدد الفقرات والاختيارات (الاستجابات) طويلاً جداً، فإنه يحتمل أن تفقد تجانسها بالإضافة إلى أنها سوف تستغرف وقتا طويلا فى التطبيق، وإذا كان عدد الفقرات قصيراً جداً، حينئذ يزيد تأثير عامل التخمين وينخفض صدق مزاوجة (مواءمة) الفقرة the validity of the matching item .

ثالثا : تطبيق الاختبار:

بعد الانتهاء من إعداد الصورة الأولية للاختبار المقترح، يبدأ الباحث فى تطبيق الاختبار وفق مجموعة من التدابير والإجراءات يمكن تحديدها على النحو التالى:

إجراءات تتخذ قبل تطبيق الاختبار : وتشمل الآتى:

١ - الإعداد لتنفيذ الاختبار : ويتضمن الآتى :

- تجهيز عدد كاف من بطاقات الإجابة (نسخ مكتوبة من الاختبار).
- توفير كافة الأدوات اللازمة لتطبيق الاختبار مثل: الأقلام، وساعة الإيقاف، وأوراق المسودات وغيرها من الأدوات.

ويفضل إعداد هذه الأدوات قبل تطبيق الاختبار بأسبوع على الأقل، لإتاحة الفرصة أمام المسئولين لمراجعة متطلبات التطبيق، والتغلب على أية مشكلات يتوقع حدوثها. كما يفضل تخزين مثل هذه الأدوات فى مكان أمين وقريب من مكان تطبيق الاختبار.

٢ - اختيار الموقع المناسب لتطبيق الاختبار:

ويفضل أن تكون غرفة من غرف الصف الدراسى المعتادة فى حالة تطبيق الاختبار على عينات من التلاميذ، وعموما يشترط فى الغرفة أو القاعة أو الصالة التى يتم فيها تطبيق الاختبار الآتى :

- أن تكون مناسبة لعدد المفحوصين بحيث يجلسون وبينهم مسافات كبيرة.
- أن تكون مناسبة من حيث المساحة والتهوية والإضاءة والحرارة حتى تحقق للمفحوصين الراحة البدنية والنفسية.
- يفضل أن يكون مكان تطبيق الاختبار (غرفة الصف أو غيرها) بعيداً عن الشوارع أو أية مصادر للضوضاء يمكن أن تشتت انتباه وتركيز المفحوصين.
- يتم جلوس المفحوصين بطريقة تسمح لهم رؤية وسماع المسئول عن إدارة الاختبار (الباحث)، وأن تسمح للباحث بالتحرك فيما بين المفحوصين بسهولة عند الرد على تساؤلاتهم واستفساراتهم.

٣ - توفير الاحتياطات المناسبة لتطبيق الاختبار، مثل:

- عمل لوحات إرشادية بموعد وأماكن تنفيذ الاختبار.

- عمل لوحات إرشادية تفيد بعدم دخول الطلاب (اللاعبين) أو الزوار إلى المنطقة التى يتم فيها تنفيذ الاختبار.

- التنبيه بعدم استخدام جرس المدرسة أو مكبرات الصوت أو إقامة المسابقات الرياضية أثناء تطبيق الاختبار.

إجراءات تتخذ أثناء تطبيق الاختبار: وتشمل الآتى:

١- قراءة تعليمات الاختبار قبل البدء فى الإجابة على الأسئلة.

٢- بدء الإجابة على الاختبار وحساب الزمن الذى يستغرقه التطبيق بدقة، حيث يتم تشغيل ساعة الإيقاف عند إعطاء اشارة البدء، وإيقاف الساعة عند انتهاء الزمن المقرر للاختبار كالتالى:

	الساعة	الدقيقة	ثانية
- زمن بدء الاختبار	٢	١٠	.
- الزمن المسموح به		٢٠	.
- ينتهى زمن الاختبار	٢	٣٠	.

٣- يجلس المسئول عن إدارة الاختبار (الباحث) فى المكان المعتاد الذى يجلس فيه معلم الفصل، ويفضل عدم جلوسه بجوار أى مفحوص، أو التحرك أثناء الإجابة على الاختبار.

٤- يقوم المسئول عن إدارة الاختبار بالرَّد على أية أسئلة أو استفسارات يقدمها المفحوصين. ولتجنب تشتت انتباه المفحوصين، يتم الرد على كل مفحوص على حدة وهو جالس فى مكانه، وبحيث يكون الرد بصوت منخفض لا يصل إلى بقية زملائه.

٥- حث المفحوصين على الأداء لأقصى ما تمكنهم قدراتهم لتحقيق أفضل نتائج ممكنة.

٦- تسجيل كل الأحداث والملاحظات المهمة التى تحدث خلال التطبيق مثل التوتر الزائد للمفحوصين، السرحان والنظر إلى النافذة لفترات زمنية طويلة، أو الإجابة بطريقة عشوائية على الأسئلة، إلى غير ذلك من الملاحظات التى يمكن الاستفادة منها عند إعادة تطبيق الاختبار.

إجراءات تتخذ بعد تطبيق الاختبار، وتشمل:

١- تصحيح الاختبار وتقدير درجات كل مفحوص على حدة.

٢- رصد الدرجات فى بطاقة تسجيل جمعية توطئة لإجراء المعالجات الإحصائية المناسبة، والتى تتضمن التحليل الإحصائى للفقرة، وهو ما سوف نعرض له فى الخطوة التالية (الخطوة الرابعة) من خطوات بناء الاختبار المعرفى.

رابعاً: تحليل الفقرة:

بعد كتابة فقرات الاختبار ومراجعتها بعناية وإعدادها مكتوبة وتطبيقها لتوفير البيانات الأولية اللازمة للتحقق من نوعية الفقرات quality of items وذلك من خلال ما يعرف بتحليل الفقرة item analysis.

ما هو المقصود بتحليل الفقرات؟

تحليل الفقرات هو الدراسة التى تعتمد على التحليل المنطقى الإحصائى والتجريبى لوحدات الاختبار لمعرفة خصائصها وحذف أو تعديل أو إبدال أو إضافة أو إعادة ترتيب هذه الفقرات حتى يتسنى الوصول إلى اختبار ثابت وصادق ومناسب من حيث الطول ومستوى الصعوبة.

الأهداف الرئيسية لتحليل الفقرة:

١- الكشف عما إذا كانت الفقرات صعبة جداً، أو سهلة جداً، أو متوسطة الصعوبة، حيث يكشف التحليل الإحصائى عن مستوى صعوبة الفقرة difficulty level أو ما يعرف بمؤشر الصعوبة index of difficulty.

٢- الكشف أيضاً عما إذا كان للفقرة قدرة على التمييز بين الأفراد المتميزين والأفراد غير المتميزين، حيث يكشف التحليل الإحصائى عن القوة التمييزية للفقرة discriminative power أو ما يعرف بمؤشر (دليل) التمييز the index of discimination وهو مؤشر يدل على صدق الفقرة item validity.

٣- إظهار لماذا فقرة معينة لا تؤدى وظيفتها بكفاية، وكيف يمكن تعديل هذه الفقرة لتصبح أكثر فعالية.

وتتوقف الإجراءات الإحصائية لتحليل الفقرة على أنماط الفقرات المستخدمة فى الاختبار، وعلى عدد المفحوصين، والغرض من التحليل، فالثقة تزداد فى الفقرة عندما يتم التحليل الإحصائى على مائة مفحوص أو أكثر، وتنخفض عندما يقل العد عن مائة.

ويهتم تحليل الفقرة فى الاختبارات المعرفية بثلاثة مصطلحات إحصائية هى بالترتيب:

- صعوبة الفقرة item difficulty
- القدرة على التمييز discrimination power
- نوعية الاستجابة response quality

خطوات تحليل الفقرة:

يتم تحليل الفقرة وفقا للخطوات التالية:

الخطوة ١: ترتيب الإجابات فى بطاقة التسجيل الجماعية من الأعلى إلى الأدنى.

الخطوة ٢: يتم تصنيف الإجابات فى بطاقة التسجيل الجماعية إلى ثلاثة مجموعات فرعية كالتالى:

- المجموعة الأعلى: وتمثل حوالى ٢٧٪ من عدد الإجابات.

- المجموعة الأدنى: وتمثل حوالى ٢٧٪ من عدد الإجابات.

- المجموعة المتوسطة: وتمثل حوالى ٤٦٪ من عدد الإجابات.

وتضم المجموعة الأعلى المفحوصين الذين حققوا أعلى درجات على الاختبار، وتضم المجموعة الأدنى المفحوصين الذين حققوا أقل درجات على الاختبار، وتستخدم المجموعة الأعلى والمجموعة الأدنى كمجموعات متضادة تعكس الحد الأقصى للفروق فى الأداء على الاختبار.

ملحوظة مهمة:

يتوقف حجم كل مجموعة من المجموعتين (الأعلى والأدنى) على العدد الملائم بالنسبة لعملية التحليل الإحصائى، كما يتوقف على العدد الكلى للإجابات. وقد ارتضى علماء القياس أن تتراوح هذه النسبة من ٢٥٪ وحتى ٣٣٪ بالنسبة لكل مجموعة.

الخطوة ٣: حساب عدد الذين أجابوا إجابات صحيحة على كل فقرة من فقرات الاختبار بالنسبة للمجموعة الأعلى upper group.

الخطوة ٤: حساب عدد الذين أجابوا إجابات صحيحة على كل فقرة من فقرات الاختبار بالنسبة للمجموعة الأدنى Lower group.

حساب مؤشر الصعوبة index of difficulty

يتم حساب مؤشر الصعوبة لكل فقرة من فقرات الاختبار باستخدام نتائج الخطوتين (٣، ٤) وتطبيق المعادلة (٤-١) التالية:

$$ب = \frac{س_1 + س_2}{ن_1 + ن_2} \times ١٠٠$$ معادلة (٤-١)

حيث إن:

ب = مؤشر الصعوبة.

$ر_1$ = عدد الإجابات الصحيحة فى المجموعة الأعلى.

$ر_2$ = عدد الإجابات الصحيحة فى المجموعة الأدنى.

$ن_1$ = عدد المفحوصين فى المجموعة الأعلى.

$ن_2$ = عدد المفحوصين فى المجموعة الأدنى.

مثال: (٤ - ٥):

طبق اختبار معرفى على ١٨٥ لاعب، والمطلوب حساب مؤشر الصعوبة للفقرات أ، ب، جـ، د. إذا كانت الاستجابات الصحيحة بالنسبة لكل فقرة موزعة بالنسبة للمجموعة الأعلى والمجموعة الأدنى كالتالى:

المجموعات	أ	ب	جـ	د	المجموع
- المجموعة الأعلى ٢٧٪	٢٨	٢	٦	١٤	٥٠
- المجموعة الأدنى ٢٧٪	٢٤	٨	١	١٧	٥٠

والمطلوب : حساب مؤشر صعوبة كل فقرة من الفقرات الأربع.

بتطبيق المعادلة (٤-١) على البيانات المذكورة سلفاً نحصل على مؤشر الصعوبة كالتالى:

$$\text{مؤشر صعوبة الفقرة (أ)} = \frac{٢٨ + ٢٤}{٥٠ + ٥٠} \times ١٠٠$$

$$= \frac{٥٢}{١٠٠} \times ١٠٠$$

$$= ٥٢٪$$

ومؤشر صعوبة الفقرة (ب) = $\frac{٨ + ٢}{٥٠ + ٥٠} \times ١٠٠$

$= \frac{١٠}{١٠٠} \times ١٠٠$

$= ١٠٪$

ومؤشر صعوبة الفقرة (جـ) = $\frac{١ + ٦}{٥٠ + ٥٠} \times ١٠٠$

$= \frac{٧}{١٠٠} \times ١٠٠$

$= ٧٪$

ومؤشر صعوبة الفقرة (د) = $\frac{١٧ + ١٤}{٥٠ + ٥٠} \times ١٠٠$

$= \frac{٣١}{١٠٠} \times ١٠٠$

$= ٣١٪$

وهناك طريقة أخرى لحساب مؤشر صعوبة الفقرة بدون تقسيم المفحوصين إلى مجموعتين (أعلى وأدنى)، حيث يتم تقدير عدد الإجابات الصحيحة على الفقرة بالنسبة لكل المفحوصين، ثم تطبق المعادلة (٤ - ٢) وهى:

$$ب = \frac{ل}{ن} \times ١٠٠ \quad \text{معادلة (٤-٢)}$$

حيث إن:

ب = مؤشر الصعوبة.

ل = عدد من أجابوا إجابات صحيحة على الفقرة.

ن = العدد الكلى للمفحوصين الذين طبق عليهم الاختبار.

مثال (٤-٦):

طبق اختبار معرفى على ٦٠٠ طالب، وقـد بلغ عدد الإجابات الصحيحة على إحدى فقرات الاختبار ٣٢٠، فما هو مؤشر صعوبة هذه الفقرة ؟

بالتعويض فى المعادلة (٤-٢) ينتج أن:

$$\text{مؤشر صعوبة الفقرة} = \frac{٣٢٠}{٦٠٠} \times ١٠٠$$

$$= ٥٣٪$$

دلالات مؤشر الصعوبة:

مما سبق يتضح أن مؤشر صعوبة الفقـرة يشير إلى النسبة المئوية للأفراد أو المفحوصـين الذين أجابوا إجابات صحيـحة على الفقرة، هذه النسـبة تعرف باسم مؤشر الصعوبة أو معامل الصعوبة، ويمكن تقويم هذه النسبة كالتالى:

١- إذا كانت هذه الـنسبة (مؤشر الـصعوبة) ٩٠٪ دل ذلك على أن الـفقرة سهلة جداً، ومن ثم تصبح هذه الفقرة غير قادرة على التمييز.

٢- إذا كانت هذه النسبة ١٠٪ فإن ذلك يدل على أن الفقرة صعبة جداً.

٣- إذا كانت هذه النسبة ١٠٠٪ أو صفراً ٪ فهـذا يعنى عدم صلاحية الفقرة لكونها لا تستطيع التمييز بأى حال من الأحوال بين المفحوصين.

٤- يتم تقويم مؤشر صعوبة الفقـرة وفقا للمستويات التى حددها ابل Ebel ١٩٧٩ كالتالى:

مؤشر الصعوبة	التقويم
٨٠٪ فأكثر	تستبعد الفقرة
من ٧١٪ الى ٧٩٪	يمكن قبول الفقرة إذا كان لها مؤشر تمييز مقبول
من ٣٠٪ الى ٧٠٪	الفقرة جيدة
من ٢٠٪ الى ٢٩٪	يمكن قبول الفقرة إذا كان لها مؤشر تمييز مقبول
أقل من ٢٠٪	تستبعد الفقرة

٥- تصنف الفقرات التى لها مؤشر صعوبة من ٤٠٪ الى ٦٠٪ كفقرات متوسطة الصعوبة. ويفضل وضع الفقرات التى لها مؤشر صعوبة يزيد عن ٦٠٪ فى أول الاختبار، والفقرات التى لها مؤشر صعوبة أقل من ٤٠٪ فى نهاية الاختبار، حتى نضمن أن يكون لتوزيع الفقرات خصائص التوزيع المعتدل normal distribution بالنسبة لمستوى صعوبة الفقرات.

مما سبق يتضح الآتى:

- يجب ألاَّ تكون الفقرة سهلة جدا بحيث ينجح فيها جميع المفحوصين، وألا تكون صعبة جداً بحيث يرسب فيها كل المفحوصين، لأن الفقرات المتطرفة فى مستويات صعوبتها أو سهولتها، لا تمدنا بالمعلومات الكافية عن مستوى تحصيل المفحوصين، وقد ارتضى علماء القياس أن يتراوح مستوى صعوبة الفقرة من ٣٠ إلى ٧٠٪.
- تحقق الفقرة أقصى قدرة على التمييز بين المفحوصين إذا كان مستوى صعوبتها يسمح بنجاح ٥٠٪ من المفحوصين الذين أجابوا عليها.
- كلما زادت النسبة المئوية للنجاح زادت سهولة الفقرة، وكلما زادت النسبة المئوية للرسوب، زادت صعوبة الفقرة.
- لتقدير مؤشر صعوبة الفقرة يمكن استخدام أى من المعادلتين (٤ : ١أو٢)، حيث يلاحظ أن المعادلة (٤-١) أسهل فى التطبيق لكونها تتعامل مع حوالى ٥٤٪ من عدد المفحوصين، وإن كان يعاب عليها أنها تفترض أن المجموعة الوسطى والتى تمثل ٤٦٪ من عدد المفحوصين تدخل فى إطار التوزيع المعتدل للبيانات وتكون المعادلة (٤-١) أفضل فى التطبيق فى حالات الأعداد الكبيرة، ففى حالة وجود ١١١٢ مفحوص مثلاً، يكون حجم المجموعة الأعلى (٢٧٪) يساوى ٣٠٠ مفحوص، وحجم المجموعة الأدنى (٢٧٪) يساوى ٣٠٠ مفحوص

أيضاً، فى هذه الحالة يتم التعامل مع درجات ٦٠٠ مفحوص فقط بدلاً من ١١١٢ مفحوص، بينما تكون المعادلة (٤-٢) أنسب فى حالات المجموعات صغيرة العدد، فإذا كان عدد المفحوصين ٥٠ مفحوصاً - على سبيل المثال - فإنه لا يفضل تقسيم هذا العدد إلى ثلاث مجموعات، المجموعة الأعلى (٢٧٪) وتضم حوالى ١٤ مفحوصاً، والمجموعة الأدنى (٢٧٪) وتضم ١٤ مفحوصاً أيضاً.

صعوبة الفقرة والاختبار المرجعى المحك (CR):

تتحدد صعوبة الفقرة فى اختبارات التمكن مرجعية المحك (CR) وفقا للطبيعة الخاصة للإنجاز المستهدف من عملية القياس، فإذا كانت المهمات المطلوب انجازها سهلة، كانت فقرات الاختبار هى الأخرى سهلة، وإذا كانت المهمات (الواجبات) متوسطة الصعوبة، كان من الضرورى أن تكون فقرات الاختبار هى الأخرى متوسطة الصعوبة، ولا يجوز بأى حال من الأحوال محاولة إجراء أى تعديل فى مستوى صعوبة الفقرة فى الاختبارات مرجعية المحك، أو حذف الفقرات السهلة من الاختبار لتوسيع مدى درجات الاختبار.

ولا يتأسس مستوى صعوبة الفقرة فى الاختبارات مرجعية المحك على قدرة الفقرة على التمييز بين المستوى المرتفع والمستوى المنخفض للمفحوصين - كما فى حالة الاختبارات معيارية المرجع (NR) - وإنما تتحدد صعوبة فقرة الاختبار فى ضوء طبيعة المهمة أو الواجب المطلوب القيام به، وبناء على ذلك فإنه يمكن تطبيق أى من المعادلتين السابقتين (٤ : ١ أو ٢) للتحقق من مؤشر صعوبة الفقرة فى الاختبارات مرجعية المحك (CR)، ومع ذلك فإن النتائج التى يتم الحصول عليها من تطبيق أى من المعادلتين قد لا تستخدم فى اختيار فقرات الاختبار أو فى معالجة صعوبة الفقرة، فمعظم الفقرات فى اختبارات التمكن مرجعية المحك (CR) يكون لها مؤشرات صعوبة مرتفعة (بمعنى أن النسبة المئوية المطلوبة للنجاح على الاختبار تكون مرتفعة).

صعوبة الفقرة والاختبار المرجعى المعيار (NR):

لأن الاختبارات مرجعية المعيار صممت لإعطاء رتب للمفحوصين وفقا لترتيبهم فى الإنجاز، وهى أيضا اختبارات صممت لإظهار الفروق الفردية بين المفحوصين فى مدى واسع من الدرجات، ولكى يتحقق لها ذلك فإنها تستبعد الوحدات السهلة التى يستطيع المفحوصين الإجابة عليها إجابات صحيحة لكونها لا تظهر الفروق فيما بين المفحوصين. فالفروق بين المفحوصين فى الاختبارات مرجعية المعيار (NR) فى حالة اختبار يتكون من ١٠٠ فقرة - على سبيل المثال - تظهر بأقصى قدر لها إذا كان مدى الدرجات يتراوح من ٥ إلى ٩٥ بمتوسط ٥٠ درجة.

وإذا كانت الصعوبة بالنسبة للاختبارات مرجعية المحك (CR) قد أعدت على أساس الحد الأدنى لمستوى الكفاءة the minimum proficiency Level، لذا فإنها تعد على أساس توقع نجاح ما لا يقل عن ٨٠ إلى ٨٥٪ من المفحوصين عندما تطبق عليهم هذه الاختبارات، وفى المقابل تعد الاختبارات مرجعية المعيار (NR) على أساس أن الفقرة متوسطة الصعوبة.

علاقة مؤشر الصعوبة بثبات الاختبار:

يتفق علماء القياس على أن مستويات صعوبة الفقرات يؤثر على ثبات الاختبار، وأن مستويات الصعوبة الأمثل التى تعطى أعلى ثبات للاختبار تتوقف على مجموعة من العوامل من أهمها عدد فقرات الاختبار، ونوع الفقرات التى يتكون منها الاختبار، حيث يوصى علماء القياس بالآتى:

- يزداد ثبات الاختبار عندما يكون عدد فقراته ٣٠ فقرة أو أقل، ويميل إلى الانخفاض كلما زاد عدد الفقرات عن ٣٠ فقرة.
- فى فقرات الصواب - الخطأ أو (موافق - غير موافق) يجب أن يكون مؤشر صعوبة الفقرة فوق ٧٠٪.

- وأن هذا المؤشر بالنسبة لفقرات الاختيار فيما بين (٢ من البدائل) يجب أن يكون ٨٥٪، وفيما بين (٣ من البدائل) ٧٧٪، وفيما بين (٤ من البدائل) ٧٤٪، وفيما بين (٥ من البدائل) ٦٩٪.

ويمكن حساب ثبات درجة الفقرة (البند) بتطبيق المعادلة (٤-٣) التالية:

$$\text{ر} = \frac{\text{ن}}{\text{ن} - ١} \times \left(\text{ك} - \frac{١}{\text{ن}} \right) \qquad \text{معادلة (٤-٣)}$$

حيث إن:

ر = معامل ثبات الفقرة (البند).

ن = عدد الاختيارات (احتمالات الإجابة).

ك = أعلى تكرار نسبى فى الاختيارات.

مثال: (٤ - ٧)

أجاب ٢٥٠ مفحوص على إحدى فقرات الاختبار، وكانت الإجابة على هذه الفقرة تتطلب اختيار إجابة من خمس إجابات هى: أ، ب، جـ، د، هـ. وقد جاءت الإجابات (الاستجابات) كالتالى:

الاختيارات	أ	ب	جـ	د	هـ	المجموع
- تكرارات الاستجابة	٢٥	٤٥	٤٠	١٠٠	٤٠	٢٥٠
- التكرار النسبى	,١٠	,١٨	,١٦	,٤٠	,١٦	١,٠٠

والمطلوب: حساب درجة ثبات هذه الفقرة.

ولتحقيق ذلك يتبع الآتى:

١- حساب تكرارات الـ ٢٥٠ مفحوص على كل اختيار من الاختيارات (الاستجابات) الخمسة فنجد أنها تساوى بالترتيب: ٢٥، ٤٥، ٤٠، ١٠٠، ٤٠ (على سبيل المثال).

٢- حساب التكرار النسبى لكل اختيار (استجابة) من الاستجابات الخمس، وذلك بقسمة تكرار كل اختيار على العدد الكلى للمفحوصين، فنجد أن هذه التكرارات النسبية كانت كالتالى:

(١,٠ + ١٨,٠ + ١٦,٠ + ٤,٠ + ١٦,٠ = ١,٠).

٣- تحديد أعلى تكرار نسبى، فنجد أنه يساوى ٤,٠ .

٤- نقوم بالتعويض فى المعادلة (٤-٣) كالتالى:

$$= \frac{٥}{١ - ٥} \times \left(\frac{١}{٥} - ٤,٠ \right)$$

$$= \frac{٥}{٤} \times (٤,٠ - ٢,٠)$$

$$= ٢٥,٠$$

إذاً معامل ثبات الفقرة يساوى ٢٥,٠

ومن ناحية أخرى يمكن استخدام مؤشر الصعوبة لحساب الانحراف المعيارى (ع - SD) للفقرة باستخدام المعادلة (٤-٤) التالية:

$$ع = \sqrt{ب \times (١٠٠ - ب)}$$ معادلة (٤-٤)

حيث إن:

ع = الانحراف المعيارى للفقرة (البند).

ب = مؤشر صعوبة الفقرة (البند).

١٠٠ = مقدار ثابت.

ففى المثال (٤-٥) كانت مؤشرات الصعوبة للفقرات أ، ب، جـ ، د كالتالى بالترتيب: ٥٢٪، ١٠٪، ٧٪، ٣١٪ .

وبتطبيق المعادلة (٤-٤)، يمكن حساب ثبات كل من هذه الفقرات كالتالى:

بالنسبة للفقرة (أ)

$$ع = \sqrt{٠,٥٢ \times (١,٠٠ - ٠,٥٢)}$$

$$= \sqrt{٠,٥٢ \times ٠,٤٨}$$

$$= \sqrt{,٢٤٩٦}$$

إذاً ع = ٠,٥٠ تقريباً

بالنسبة للفقرة (ب)

$$ع = \sqrt{٠,١٠ \times (١,٠ - ٠,١)}$$

$$= \sqrt{٠,١ \times ٠,٩}$$

$$= \sqrt{,٠٩}$$

إذاً ع = ٠,٣

كما يمكن حساب تباين الفقرة حيث (التباين = ع٢) وبالتالى يكون تباين الفقرة (أ) = (٠,٥)^{٢} = ٠,٢٥ ، وتباين الفقرة (ب) = (٠,٣)^{٢} = ٠,٠٩ ، ومن الملاحظ أن أعلى قيمة لتباين الفقرة هى ٠,٢٥ ، وهى القيمة التى يساوى فيها مؤشر الثبات حوالى ٥٠٪.

حساب مؤشر التمييز (I D) : Index of discrimination

يشير مؤشر (دليل) التمييز إلى الدرجة التى تميز بها الفقرة (البند) بين مستويات المفحوصين، بمعنى قدرة الفقرة على أن تميز بين الأفراد الذين يحصلون على درجات مرتفعة، والأفراد الذين يحصلون على درجات منخفضة فى الصفة أو السمة التى تقيسها كل فقرات المقياس أو الاختبار. فعندما يحصل عدد كبير من المفحوصين فى المجموعة الأعلى upper group على إجابات صحيحة على الفقرة، نقول: إن القدرة التمييزية للفقرة إيجابية،

بمعنى أن لـلفقرة مؤشـر تمييز إيـجابى positive discrimination، أى أن اتجاه التمييز للفقرة هو نفس اتجاه الدرجة الكلية للاختبار.

فإذا كان لفقرة الاختبار القدرة على التمييز بصورة جيدة، فإن ذلك يعنى أن المفحوصـين الذين يحققون درجات عـالية (مرتفعة) على الاخـتبار سوف يجيـبون على هذه الـفقرة إجـابة صحيـحة، عكس المفـحوصين الـذين لهم درجات منخفضـة على الاختبار، وتعد الفقـرة غير القادرة على التميـيز فقرة ضعيفة البناء poorly Construction .

بينما يشـير التمييز السلـبى للفقرة negative discrimination إلى أن أكثر المفحوصين فى المجموعة الأدنى Low group يحصلون على إجابات صحيحة للفقرة.

ويعرف مـؤشر (دليل) الـتميـيز (ID) بمؤشر صدق الـفقرة item validity index والذى يعد أحد الإجراءات المهمة بالنسبـة لتحليل الفقرة. حيث يمكننا من تصنيف كل فقرات الاختبار إلى ثلاثة أنواع من الفقرات هى:

١ - فقرات لها مؤشر تمييز إيجابى.

٢ - فقرات لها مؤشر تمييز سلبى.

٣ - فقرات ليس لها قدرة تمييزية.

ويمكن حساب مؤشر تمييز الفقرة بإحدى الطريقتين التاليتين:

١ - طريقة النسب المئوية Percentages

وتتم هذه الطريقة بتقسيم المفحوصين إلى مـجموعتين متساويتين فى العدد وفقـا للدرجة الـكلية لـلاختبـار (درجات جمـيع الفـقرات)، بحيـث تضم المجموعة الأولى (ن١) المتفوقين، وتضم المجموعة الثانية (ن٢) غير المتفوقين.

وتستخدم المجموعتين (ن١، ن٢) فى المقارنة كمجموعتين متطرفتين. ويمكن استخدام أى من النسب المئوية التالية لتحديد حجم كل مجموعة.

٣٣٪ كقيم طرفية لكل مجموعة - كيرتون Cureton ١٩٥٧م.

٢٧٪ كقيم طرفية لكل مجموعة - كيلى Kelly ١٩٣٩م.

٢٥٪ كقيم طرفية لكل مجموعة - ايبل Ibel ١٩٧٩م.

٩ ٪ كقيم طرفية لكل مجموعة - ايبل Ibel ١٩٧٩م.

٢ - حساب عدد المفحوصين الذين أجابوا إجابات صحيحة على الفقرة المستهدفة بالنسبة للمجموعة المتميزة (ن١) حيث نرمز لذلك بالرمز ($\checkmark_1$).

٣ - بالمثل يتم حساب عدد من أجابوا إجابات صحيحة على نفس الفقرة بالنسبة للمجموعة التى حصلت على أقل درجات على الاختبار (ن٢) ، حيث نرمز لذلك بالرمز ($\checkmark_2$).

٤ - يتم طرح عدد الاستجابات الصحيحة للمجموعة الضعيفة ($\checkmark_2$) من عدد الاستجابات الصحيحة للمجموعة المتميزة ($\checkmark_1$) ، ثم يقسم الناتج على العدد الكلى للمجموعة المتميزة (ن١) باستخدام معادلة مارشال وهاليس marshall and hales ١٩٧٢م، وصورتها الرياضية كالتالى:

$$د = \frac{\checkmark_1 - \checkmark_2}{ن_1}$$ معادلة (٤ - ٥)

حيث إن:

د = مؤشر التمييز للفقرة (صدق الفقرة).

$\checkmark_1$ = عدد الإجابات الصحيحة فى المجموعة المتميزة.

$\checkmark_2$ = عدد الإجابات الصحيحة فى المجموعة الضعيفة.

ن١ = عدد المفحوصين فى المجموعة المتميزة.

مثـــال:

فى المثال : (٤-٥) كانت بيانات الإجابة على الفقرة (أ) كالتالى:

$ل_{١} = ٢٨$

$ل_{٢} = ٢٤$

$ن_{١} = ٥٠$

إذاً يمكن حساب مؤشر التمييز لهذه الفقرة بتطبيق المعادلة (٤-٥) كالتالى:

$$د = \frac{٢٨ - ٢٤}{٥٠}$$

$$= \frac{٤}{٥٠}$$

= ٠,٠٨ إذاً مؤشر التمييز لهذه الفقرة يساوى (١,٠) تقريبا.

دلالات مؤشر التمييز (ID):

١- يتراوح مؤشر التمييز من (+١,٠٠ إلى -١,٠٠) ولكنه نادراً ما يصل إلى أى من القيمتين.

٢- عندما يكون مؤشر التمييز سلبيا، فإن ذلك يعنى أن عدد المفحوصين الذين أجابوا إجابات صحيحة على الفقرة فى المجموعة الأدنى ($ن_{٢}$) أكثر ممن أجابوا على نفس الفقرة فى المجموعة الأعلى ($ن_{١}$)، ومن ثم يحذف مثل هذا النوع من الفقرات تماماً.

٣- بالنسبة للاختبارات مرجعية المعيار (NR) يقبل مؤشر تمييز الفقرة عندما يساوى (٠,٤) فأكثر. وقد حدد ايبل Ibel ١٩٧٩ مؤشرات التمييز للفقرات فى الاختبارات مرجعية المعيار كالتالى:

مؤشر التمييز	التقويم
٠,٤٠ فما فوق	لها قدرة على التمييز جيدة

من ٠,٣٠ إلى ٠,٣٩	لها قدرة على التمييز مناسبة، وربما تحتاج إلى تعديل وبخاصة فيما يتعلق بصعوبة الفقرة
من ٠,٢٠ إلى ٠,٢٩	تحتاج إلى مراجعة
أقل من ٠,٢٠	تستبعد الفقرة

ملحوظة مهمة:

لاتطبق المستويات السابقة الخاصة بمؤشرات التمييز على الاختبارات مرجعية المحك (CR).

٢- الطرق الارتباطية Correlational techniques

قدم ليندكويست Lindquist ١٩٥١م الدرجة الكلية للاختبار كمحك داخلى لحساب صدق الفقرات بالرجوع إلى هذا المحك الداخلى، حيث أطلق ليندكويست على هذه الطريقة اسم (مؤشر تمييز الفقرة عن طريق الاتساق الداخلى Internal Consistency). وتعتمد هذه الطريقة على تحليل العلاقة (الارتباط) بين درجة المفحوص على الفقرة ودرجته الكلية على الاختبار ككل والتى تستخدم كمحك (أو ميزان داخلى) لتقويم صدق كل فقرة من فقرات الاختبار.

فالعلاقة بين الفقرة والدرجة الكلية للاختبار تظهر كيف تقيس الفقرة بشكل جيد الوظائف التى يقيسها الاختبار نفسه، حيث يبين مؤشر الصدق المدى الذى تمتد إليه فقرة محددة فى التمييز فيما بين المفحوصين الذين يظهرون فروقاً حادة فى الوظائف التى يقيسها الاختبار ككل.

ويرى بعض خبراء القياس أنه يمكن النظر إلى ارتباط الدرجة الكلية للفقرة على أنه مؤشر جيد للتمييز، حتى أنه يمكن اختيار الفقرات على أساس الارتباطات الكلية للفقرة، إذا كانت نسبة الارتباطات الموجبة لا تقل عن ٧٥٪ وتزيد قيمها عن ١٥,٠.

ويستخدم لقياس الارتباط بالنسبة لهذا الموضوع خمسة مقاييس شائعة هى: الارتباط التتابعى (PPM) والارتباط الثنائى biserial (r) والارتباط الثنائى الحقيقى Point biserial (r) ، والارتباط الرباعى tetrachoric ومعامل فاى Phi Coefficient - . ويعتبر فلانجان Flanagan أول من استخدم معاملات الارتباط الثنائى لتقدير نسبة المفحوصين الناجحين على الفقرة بالنسبة لـ ٢٧٪ العليا كمجموعة مرجعية Criterion group ونسبة المفحوصين الناجحين على نفس الفقرة فى الـ ٢٧٪ الدنيا، وذلك على أساس ألا يقل حجم كل مجموعة عن ١٠٠ مفحوص وأن يكون العدد الكلى للمفحوصين ٣٧٠ مفحوصا.

التخمين وعوامل الصدفة:

من العوامل التى تؤثر سلبيا على صعوبة الفقرة - عند استخدام النسبة المئوية للإجابات الصحيحة - هو لجوء بعض المفحوصين إلى التخمين guessing فيما بين الاختيارات المعروضة عليهم، بمعنى أن يختار المفحوص جواباً من بين عدة أجوبة دون دراية بسبب الاختيار ثم يحدث أن تكون هذه الإجابة صحيحة عن طريق الصدفة by chance. فالتخمين يعنى أن نسبة معينة من المفحوصين الذين لا يعرفون الإجابة الصحيحة أجابوا بمحض الصدفة إجابات صحيحة على الفقرة.

تصحيح أثر التخمين:

للتغلب على هذه المشكلة يتم ما يعرف بتصحيح التخمين أو أثر التخمين (تصحيح الصدفة) Correction for guessing والتى تستهدف استبعاد النسبة المئوية للتخمين من النسبة المئوية للناجحين فى الفقرة للوصول إلى النسبة المئوية الحقيقية للإجابات الصحيحة للفقرة. ويتم تصحيح أثر التخمين كالتالى:

أولاً: بالنسبة لاختبارات الصواب - الخطأ

تستخدم المعادلة التالية:

س = ل - خ معادلة (٤-٦)

حيث إن :

س = الدرجة بعد التصحيح.

ل = عدد الإجابات الصحيحة.

خ = عدد الإجابات الخاطئة.

وتتأسس هذه المعادلة على مسلمة مؤداها أن الإجابات الخاطئة جاءت نتيجة التخمين.

مثال (٤-٨):

طبق اختبار الصواب - الخطأ يتكون من ٧٥ فقرة على أحد التلاميذ، وقد تمكن التلميذ من الإجابة على ٦٠ فقرة إجابات صحيحة، والإجابة على ١٥ فقرة إجابات خاطئة. والمطلوب حساب درجة هذا التلميذ بعد تصحيح أثر التخمين.

لتحقيق المطلوب يتبع الآتى:

- حساب الدرجة بعد التصحيح بتطبيق المعادلة (٤-٦) كالتالى:

س = ٦٠ - ١٥

= ٤٥ درجة

أى أن درجة التلميذ أصبحت ٤٥ بدلا من ٦٠

ثانيا : بالنسبة لاختبارات الاختيار من متعدد (MCq)

تستخدم المعادلة التالية :

$$س = ل - \frac{خ}{ن - ١}$$ معادلة (٤-٧)

حيث إن :

س = الدرجة بعد التصحيح.

ل = عدد الإجابات الصحيحة.

خ = عدد الإجابات الخاطئة.

ن = عدد البدائل (الاحتمالات أو الاختيارات) لكل فقرة.

ملحوظة مهمة :

المعادلة (٤-٦) هى نفسها المعادلة (٤-٧) وذلك فى حالة ما إذا كانت (ن=٢) كالتالى :

$$س = ل - \frac{خ}{٢ - ١} \quad \text{معادلة (٤-٧)}$$

$$= ل - \frac{خ}{١}$$

$$س = ل - خ \quad \text{معادلة (٤-٦)}$$

وإذا كان التلميذ قد أجاب إجابات صحيحة على ٦٠ فقرة (المثال : ٤-٨) فى اختبار الاختيار من متعدد (mcq) احتمالات الإجابة فيه (٣, ٤, ٥) فى هذه الحالة يمكن حساب درجة التلميذ بعد التصحيح بتطبيق المعادلة (٤-٧) كالتالى :

فى حالة (٣ اختيارات) = ٦٠ - $\frac{١٥}{٣ - ١}$

$$= ٦٠ - \frac{١٥}{٢}$$

$$= ٦٠ - ٧,٥$$

$$= ٥٢,٥$$

٥٢,٥

فى حالة (٤ اختيارات) = ٦٠ - $\frac{١٥}{٤ - ١}$

$$= ٦٠ - \frac{١٥}{٣}$$

$$= ٥٥$$ ٥٥

فى حالة (٥ اختيارات) $= ٦٠ - \frac{١٥}{٥ - ١}$

$$= ٦٠ - \frac{١٥}{٤}$$

$$= ٦٠ - ٣,٧٥$$

$$= ٥٦,٢٥$$ ٥٦,٢٥

مما سبق يلاحظ الآتى:

- فى حالة الفقرات التى تعتمد على إجابة واحدة من إجابتين، فإن إجابات المفحوص تتأثر بعوامل التخمين أكثر منه بالنسبة لأى نمط آخر من الاختيارات، فقد بلغ عدد الإجابات الصحيحة ٤٥ إجابة بعد التصحيح فى مقابل ٦٠ إجابة قبل التصحيح.
- أنه فى حالة فقرات الاختيار من متعدد يقل تأثير التخمين فى مقابل الفقرات ثنائية الاختيار (صح - خطأ). وأنه كلما زاد عدد الاختيارات (الاحتمالات) قل تأثير التخمين. ففى حالة (٥ اختيارات) كان تأثير التخمين = ٣,٧٥ درجة وفى حالة (٤ اختيارات) كان تأثير التخمين = ٥ درجات، وفى حالة (٣ اختيارات) كان تأثيره = ٧,٥ درجة، وفى حالة (اختيارين فقط) كان تأثيره = ١٥ درجة.

مثال (٤ - ٩):

طبق اختبار معرفى الإجابة عليه بطريقة الاختيار من متعدد (mcq) لعدد أربعة اختيارات هى (أ، ب، جـ ، د) وبعد ترتيب استجابات المفحوصين وفقا للدرجة الكلية للاختبار من الأعلى إلى الأدنى، أخذت المجموعة الأعلى

(ن١) والمجموعة الأدنى (ن٢) كمجموعتين طرفيتين، وكانت استجابات كل مجموعة وفقا للبدائل (الاختيارات) المطروحة كالتالى:

المجموعات/ الفقرات	أ	ب	جـ	د	المجموع
المجموعة الأعلى(ن١)	١	٤	١٧	٥	٢٧
المجموعة الأدنى (ن٢)	٣	٧	٨	٦	٢٧

والمطلوب: تقويم هذه المجموعة من الفقرات وفقا لنتائج الاستجابة عليها.

لتحقيق ذلك نقوم بحساب مؤشر صعوبة كل فقرة بتطبيق المعادلة (٤-١) كالتالى:

بالنسبة للفقرة (أ) = $\frac{١+١}{٢٧+٢٧} \times ١٠٠ = \frac{٢}{٥٤} \times ١٠٠ = ٤٪$

وبالنسبة للفقرة (ب) = $\frac{٧+٤}{٥٤} \times ١٠٠ = \frac{١١}{٥٤} \times ١٠٠ = ٢٠٪$

وبالنسبة للفقرة (جـ) = $\frac{٨+١٧}{٥٤} \times ١٠٠ = \frac{٢٥}{٥٤} \times ١٠٠ = ٤٦٪$

وبالنسبة للفقرة (د) = $\frac{٦+٥}{٥٤} \times ١٠٠ = \frac{١١}{٥٤} \times ١٠٠ = ٢٠٪$

ووفقا للمعايير التى حددها إيبل ١٩٧٩م لمؤشر الصعوبة، فإنه يمكن استبعاد الفقرات (أ، ب، د) والابقاء فقط على الفقرة (جـ) لأن لها معامل صعوبة ينحصر فيما بين (٣٠ إلى ٧٠٪).

وللابقاء على هذه الفقرة أو استبعادها، يتم تصحيح مؤشر الصعوبة من أثر التخمين وذلك باستخدام المعادلة (٤-٨) كالتالى:

مؤشر الصعوبة بعد التصحيح = ب - $\frac{(١٠٠ - ب)}{ن - ١}$ معادلة (٤-٨)

حيث إن:

ب = مؤشر الصعوبة قبل التصحيح.

ن = عدد الاختيارات (احتمالات الإجابة).

١٠٠ = مقدار ثابت.

وبالتعويض فى المعادلة (٤-٨) ينتج أن:

مؤشر صعوبة الفقرة (جـ) بعد التصحيح

$$= ٤٦ - \frac{١٠٠ - ٤٦}{١ - ٤}$$

$$= ٤٦ - \frac{٥٤}{٣}$$

$$= ٤٦ - ١٨$$

$$= ٢٨$$

إذاً مؤشر صعوبة الفقرة (جـ) يساوى ٢٨ ٪.

ووفق المستويات التى حددها ايبل ١٩٧٩م لمؤشرات الصعوبة، فإنه يلاحظ أن حذف هذه الفقرة أو الابقاء عليها يتحدد على أساس قدرة الفقرة على التمييز وذلك بتطبيق المعادلة (٤-٥) كالتالى:

$$\text{مؤشر التمييز للفقرة (جـ)} = \frac{١٧ - ٨}{٢٧}$$

$$= \frac{٩}{٢٧}$$

= ٠,٣٣ (وهو معامل تمييز يجيز قبول الفقرة)

متى نلجأ إلى تصحيح أثر التخمين؟

نلجأ إلى تصحيح أثر التخمين فى الحالات التالية:

١ - فى الاختبارات المعرفية التى تتطلب السرعة فى الأداء، والتى تعرف باسم: اختبارات السرعة speed tests ، لكونها اختبارات تعتبر السرعة فى الأداء عاملا مهما، لذلك يتوقع أن يحدث التخمين عند الاستجابة على مثل هذه الاختبارات.

٢ - فى الاختبارات المعرفية التى تتميز بدرجة عالية من الصعوبة وتكون محدودة بزمن.

٣ - قد لا تكون هناك حاجة لتصحيح أثر التخمين عندما ينتهى جميع المفحوصين من الاختبار قبل انتهاء الزمن المخصص له، لأن هذا يعنى أن أثر التخمين فى الإجابات كان ضعيفا للغاية لعدم وجود أية ضغوط مرتبطة بالزمن قد تؤدى إلى التخمين.

٤ - يلزم تصحيح أثر التخمين فى الاختبارات التى تصمم بحيث لا يستطيع المفحوص الإجابة على جميع وحداتها خلال الزمن المقرر للاختبار.

نوعية الاستجابة Response Quality

يطلق على اختيار الإجابات على الفقرة فى حالة اختبارات الاختيار من متعدد (mcq) اسم: الاستجابات responses، حيث تشير الاستجابات غير الصحيحة - فى هذا النمط من الاختبارات - إلى ما يعرف بالمضللات (المشتتات) distractors والتى تتمثل فى إجابات خاطئة تغرى المفحوص باختيارها فى اختبارات الاختيار من متعدد أو اختبارات المزاوجة (المواءمة).

ففى اختبارات الاختيار من متعدد (mcq) يمكن أن تختار كل استجابة عن طريق بعض المفحوصين، فإذا لم تختار الاستجابة بواسطة أى مفحوص، فإن ذلك يعنى أن هذه الاستجابة ليس لها إسهامات تذكر بالنسبة للاختبار، ومن ثم فإنه يلزم حذف هذه الاستجابة من استجابات الاختيار من متعدد أو المزاوجه.

وقد اتفق علماء القياس على أن يتم الإبقاء على الاستجابات إذا اختارها ما لا يقل عن (٢ إلى ٣٪) من عدد المفحوصين الذين أعطى لهم الاختبار.

ويتفق علماء القياس على أن المضللات (المشتتات) تكون فعالة إذا كان عدد المفحوصين الذين اختاروها فى المجموعة الأدنى أكثر من عدد المفحوصين الذين اختاروها فى المجموعة الأعلى، لأن ذلك يعنى أن لهذه المضللات قدرة على التمييز بين المجموعة الأدنى والمجموعة الأعلى، وأن التمييز فى الاتجاه الإيجابى لكونه فى صالح المجموعة الأعلى صاحبة الاستجابات الأفضل على الاختبار.

مثال : (٤-١٠):

البيانات التالية تبين استجابات المجموعة الأعلى (ن١ = ٢٧) ، والمجموعة الأدنى (ن٢ = ٢٧) على فقرة فى اختبار للاختيار من متعدد (mcq) عدد الاختيارات فيه أربع احتمالات هى: أ ، ب، (جـ) ، [د] ، حيث يلاحظ الاستجابة (جـ) هى الاستجابة الصحيحة، وأن الاستجابة د هى الاستجابة المضللة.

المجموعة / الاختيارات	أ	ب	(جـ)	[د]
المجموعة الأعلى (ن١=٢٧)	٤	٠	١٢	١١
المجموعة الأدنى (ن٢=٢٧)	٩	٠	٨	١٠

ويمكن تقويم الاستجابات على هذه الفقرة كالتالى:

- تستبعد الاستجابة (ب) لعدم فعاليتها، كاستجابة بديلة لكونها لم يتم اختيارها من قِبل أى مفحوص فى المجموعتين.
- الاستجابة (جـ): هى الاستجابة الصحيحة لقدرتها على التمييز الإيجابى بين المجموعتين الأعلى والأدنى، لأن عدد المفحوصين الذين استخدموها فى المجموعة الأعلى (١٢ مفحوصا) وهو أكثر من العدد الذى اختارها فى المجموعة الأدنى (٨ مفحوصين).
- الاستجابة (د) لا تصلح كاستجابة مضللة لأن عدد الذين اختاروها من المجموعة الأعلى (١١ مفحوص) وهو أكبر من عدد الذين اختاروها فى المجموعة الأدنى (١٠ مفحوصين)
- الاستجابة (أ) تظهر كاستجابة مضللة، لأن المفحوصين الذين اختاروها فى المجموعة الأدنى (٩ مفحوصين) وهو أكبر من عدد المفحوصين الذين اختاروها فى المجموعة الأعلى (٤ مفحوصين).

مما سبق يتضح أن الفقرة فى اختبارات الاختيار من متعدد أو اختبارات المزاوجة يكون لها مجموعة من الاستجابات، يطلق على الاستجابة الصحيحة اسم: الاختيار الصحيح، وتعرف الإستجابات غير الصحيحة باسم: المضللات أو المشتتات، هذه المضللات (المشتتات) عبارة عن اختيارات (بدائل) غير صحيحة incorrect alternatives تستمد اسمها من وظيفتها لكى تضلل المفحوصين المتشككين فى الإجابة الصحيحة للفقرة.

خامساً: مراجعة وتنقيح الفقرة:

بعد الانتهاء من تحليل الفقرة - وهى الخطوة الرابعة من خطوات بناء الاختبارات المعرفية - يبدأ الباحث بعمل المراجعات والتعديلات الضرورية لتنقيح الاختبار. حيث تشتمل هذه المراجعات على الآتى:

- تغيير بعض الاستجابات.
- حذف بعض الفقرات (البنود) وإعادة صياغة البعض الآخر منها.
- نقل الفقرات من اختيار لآخر (كأن تنقل الفقرة من صيغة الاختيار من متعدد إلى صيغة الصواب - الخطأ).
- استبعاد العبارات (البنود) التى لا تحقق مؤشر التمييز المقبول وفقا للمعايير الخاصة بذلك.

معجم إنجليزى - عربى للمصطلحات

A

Ambiguous	غامض
Alternatives	بدائل

B

Biases	تحيز

C

Correction for guessing	تصحيحح أثر التخمين
Completion	التكميل
Completion items	فقرات التكميل
Complex concepts	مفاهيم معقدة
Criterion group	مجموعة مرجعية

D

distrator	مشتت - مضلل
difficulty level	مستوى الصعوبة
discriminative power	القوة التمييزية للفقرة

E

Essay	كتابة المقال
Essay test	اختبار مقال
Essay items	فقرات المقال
Essay test items	فقرات اختبار المقال

F

Facts	حقائق
Factual Knowledge	المعرفة الحقيقية

G

Guessing factor	عامل التخمين

H

Higher mental processes	العمليات العقلية العليا

I

Interpreting	توضيح - تفسير
Integrating	التوحيد والدمج
Item analysis	تحليل الفقرة
Item revision	مراجعة وتنقيح الفقرة
Item writing	كتابة الفقرة
Item validity	صدق الفقرة
Irrelevant	غير ذات صلة
Item difficulty	صعوبة الفقرة
Index of the difficulty	مؤشر الصعوبة

M

Misconceptions	مدركات خاطئة
Mcq	الاختيار من متعدد
Memorization	الحفظ عن ظهر قلب

Matching test اختبار المزاوجة

Matching items فقرات المزاوجة

N

NR اختبارات مرجعية المعيار

Negative discrimination

O

objective test اختبار موضوعى

organizing تنظيم

objective test item
فقرات الاختبار الموضوعى

opinion رأى

P

paper and pencil tests
اختبارات الورقة والقلم

Principles مبادئ

Positive words تعبيرات موجبة

plausible قابل ظاهريا للتصديق

positive discrimination
التمييز الإيجابى للفقرة

Poorly constructions فقرة ضعيفة البناء

Q

Quality of items نوعية الفقرات

R

Relative weights الأوزان النسبية

Random guessing التخمين العشوائى

Responses استجابات

Rote learning التعلم الصمى (الحفظ)

Response quality نوعية الاستجابة

S

Subjective evaluation تقويم ذاتى

Synthesizing تركيب - تأليف

Short essay items فقرات المقال القصيرة

Stem الأرومة (البند الرئيسى)

T

thinking ability القدرة على التفكير

true - false items
فقرات الصوات - الخطأ

V

Vague غير واضح

W

Written expression التعبير كتابة

word كلمة

مراجـع الفصل الرابـع

١ - أمين أنور الخولى، وجمال الدين الشـافعى (٢٠٠٠م) : مناهج التربية البدنية المعاصرة. دار الفكر العربى، القاهرة.

٢ - أمين أنور الخولى، ومحمود عبدالفتاح عنان (١٩٩٩م): المعرفة الرياضية (الإطار المفاهيمى - اختبارات المعرفة الرياضية - أسس بنائها ونماذج كاملة منها) دار الفكر العربى، القاهرة.

٣ - أنور محمد الشرقاوى (٢٠٠٣م) : علم النفس المعرفى المعاصر. مكتبة الأنجلو المصرية، القاهرة.

٤ - ليلى السيد فرحات (٢٠٠١م) : القياس المعرفى الرياضى. مركز الكتاب للنشر، القاهرة.

٥ - نورمان جرونلند (١٩٧٦م): الأهداف التعليمية - تحديدها السلوكى وتطبيقاته. ترجمة الدكتور أحمد خيرى كاظم، دار النهضة العربية، القاهرة.

6- Bloom, B.S. (Ed.) (1956): Taxonomy of educational objectives: Cognitive domain. Mckay : New Nork.

7 - Gronlund, N.E. (1981): Measurement and evaluation in teaching. Macmillan publishing Co., Inc. New york.

8 - Johnson, B.L., and Nelson, J.K (1986): Practical measurements for evaluation in physical education. Minneapolis: Burgess publishing company.

9 - Miller,D.K. (1994): Measurement by the physical educator : why and How, WCB Brown and Benchmark Publishers.

10 - Morrow, J.R., Jackson, A.W. Disch, J.G., and Mood, D.P. (1995): Measurement and evaluation in human performance. Human Kinetics.