

القسم السابع جمع البيانات

الفصل الثامن عشر: جمع البيانات باستخدام

الاختبارات والمقاييس

الفصل التاسع عشر: جمع البيانات باستخدام

الاستبيان والمقابلة

الفصل العشرون: جمع البيانات باستخدام

الملاحظة وتحليل المحتوى

الفصل الحادي والعشرون: الصفات الواجب توافرها في

أدوات جمع البيانات

القسم السابع جمع البيانات

يستخدم الباحث العديد من الأدوات في جمع البيانات لاختبار الفروض أو الإجابة على الأسئلة، وقد يختار أدواته من بين الأدوات المتوفرة التي بناها باحثون آخرون، وقد يضطر في حالة عدم توفر أداة مناسبة لبحثه إلى بناء الأداة بنفسه. ولكل أداة مزاياها وعيوبها، إلا أن الأداة التي نختارها يجب أن تكون أفضل أداة في جمع البيانات الخاصة بالبحث. فإذا اخترنا الاستبيان مثلا كأداة لجمع البيانات يجب أن تكون أفضل وأكفأ وسيلة للإجابة على أسئلة البحث. وينطبق هذا على أية وسيلة أخرى، سواء كانت اختبارة أو ملاحظة، أو غير ذلك.

وبعض أدوات جمع البيانات مباشرة تستخدم مؤشرا واحدا ليمثل المتغير المراد قياسه، في حين أن بعض الأدوات الأخرى يسعى لقياس مفاهيم مركبة أو تكوينات فرضية، لا يمكن قياسها قياسا مباشرا، بل لابد من تعريفها تعريفا إجرائيا قابلا للملاحظة المباشرة حتى يمكن قياسها. ومن الأمثلة على المتغيرات التي يمكن قياسها بشكل مباشر، الخلفية التعليمية للفرد وذلك بسؤاله عن أعلى صف وصل إليه، ومن المتغيرات الأخرى التي يسهل قياسها المستوى الاقتصادي والاجتماعي، والحالة الاجتماعية، وعدد الأولاد، وغير ذلك من المتغيرات المشابهة. ففي كل حالة من هذه الحالات يمكن قياس المتغير بمؤشر واحد مقبول لأن هذه المتغيرات تشير إلى ظاهرات واضحة. أما النوع الثاني من المتغيرات التي تشير إلى تكوينات فرضية، فهي أكثر تعقيدا ومن الصعب قياسها، وفي مثل هذه الحالات من غير المناسب استخدام مؤشر واحد كوسيلة لقياس متغير من هذا النوع المركب.

ومن أصعب المهام أمام الباحث اختيار أو بناء المقاييس والأدوات التي يحتاجها لقياس المتغيرات ذات الخصائص المركبة، مثل الذكاء، والاسعدادات العقلية، والتحصيل الدراسي، والشخصية، والدافعية، ومفهوم الذات، وغير ذلك من المفاهيم. ويمكن استخدام أدوات مختلفة لقياس وتكميم الصفات المتنوعة. وهناك طريقتان للحصول على مثل هذه المقاييس: الأولى هي استخدام مقياس موجود فعلا وسبق استخدامه، والثانية هي بناء الأداة أو الأدوات التي يحتاجها الباحث لجمع بيانات بحثه. وفي الحالة الأولى لابد أن يرجع الباحث إلى المصادر الخاصة بالمقاييس الجاهزة، أو يرجع إلى البحوث السابقة ليستعين بأحد الأدوات التي استخدمها باحث قبله. وإذا وجد الباحث أداة مناسبة يجب أن يتأكد قبل استخدامها أن هذه الأداة يمكن أن تقيس متغيراته بكفاءة، كما يجب أن يتأكد أنها صادقة وثابتة، وأنها قننت على مجتمع مشابه لمجتمع بحثه. وإذا لم يجد أداة مناسبة لبحثه، فعليه بناء الأداة أو الأدوات التي يحتاجها، ويجب في هذه الحالة أن يتأكد من إجراء جميع الخطوات اللازمة لبناء وتقنين الأداة وتحقيق صدقها وثباتها. وسوف نتناول في

الفصول من السادس عشر إلى الثامن عشر أدوات جمع البيانات، على أن يعقبها في
الفصل التاسع عشر مناقشة لكيفية تحقيق صدق وثبات أدوات القياس.
وسوف نتناول في هذا الفصل والفصلين التاليين وسائل جمع البيانات
التالية:

- الاختبارات والمقاييس.
- الاستبيان والمقابلة.
- الملاحظة وتحليل المحتوى.

الفصل الثامن عشر

جمع البيانات باستخدام

الاختبارات والمقاييس

الاختبار مجموعة من المثيرات التي تقدم للفرد لاستثارة استجابات تكون أساساً لإعطاء الفرد درجة رقمية، وهذه الدرجة القائمة على عينة ممثلة لسلوك الفرد، تعتبر مؤشراً للقدر الذي يمتلكه الفرد من الخاصية التي يقاس بها الاختبار (Ary et al., 1996, p. 233).

وفائدة الدرجات التي نحصل عليها من الاختبارات والمقاييس كمؤشرات للمفهوم الذي نرغب في قياسه، هي إلى حد كبير دالة لموضوعية الاختبارات وصدقها وثباتها. وتشير الموضوعية إلى درجة الاتفاق بين مقياسي الدرجة. وبعض الاختبارات، مثل اختبارات الاختيار من متعدد، واختبارات الصواب والخطأ توصف بأنها موضوعية لأن وضع الدرجة يتوقف على مقارنات درجات الطلبة بمفتاح التصحيح، ولذلك لا يحتاج الأمر إلى اتخاذ قرار من جانب مقدري الدرجة. أما اختبارات المقال فهي أقل موضوعية لأن إعطاء الدرجة يتأثر بحكم ورأي مقدر الدرجة. وبالنسبة للصدق والثبات فسوف نناقشه في الفصل التاسع عشر. أما في هذا الفصل فنتناول أهم الاختبارات والمقاييس والتي منها الاختبارات التحصيلية، واختبارات الاستعدادات العقلية، واختبارات الشخصية، ومقاييس الاتجاهات، ومقاييس التقدير، وقوائم المراجعة، والأساليب السوسيومترية.

أولاً: الاختبارات التحصيلية:

وهي الاختبارات التي يراد بها قياس التحصيل الدراسي، ويطلق عليها أحياناً اختبارات القلم والورقة، وتعتبر من أهم وسائل تقويم التحصيل، وتحديد مستوى الطلبة التحصيلي. والاختبارات التحصيلية واسعة الاستخدام في البحوث

التربوية وبخاصة البحوث التي تتناول فاعلية طرق التدريس، إذ يكون المتغير التابع في هذه البحوث هو التحصيل الدراسي. وإذا لم يتوفر للباحث اختبار تحصيلي مقنن لقياس المتغير الذي يريده، فعليه أن يبني الاختبار بنفسه. وفي هذه الحالة يجب عليه اتباع المبادئ التي سوف نذكرها بعد قليل حتى تزوده بأساس راسخ لبناء الاختبار التحصيلي. والتقيد بهذه المبادئ يساعد على بناء اختبار صادق وثابت لقياس المعلومات والمهارات التي يريد الباحث قياسها (رجاء أبوعلام، ١٩٨٧).

ويمكن تصنيف الاختبارات إلى اختبارات تحريرية، واختبارات عملية، واختبارات شفوية.

١ - **الاختبارات التحريرية:** الاختبارات التحريرية هي التي يراد بها تقويم التحصيل الدراسي في نهاية الفترات، وفي امتحانات النقل، والشهادات العامة. ويطلق عليها أحيانا اختبارات القلم والورقة، وتعتبر من أهم وسائل تقويم التدصيل، وتحديد مستوى التلاميذ التحصيلي.

وهناك نوعان من الاختبارات المستخدمة في قياس وقت وديم التدصيل الدراسي، وهما:

- اختبارات الاستدعاء (وهي التي يعطي فيها الطالب الإجابة من عنده، أي يستدعيها ولا يتعرف عليها)، وتنقسم إلى:
 - أ- أسئلة المقال ذات الإجابة المستفيضة.
 - ب- أسئلة المقال ذات الإجابة المحدودة.
 - ج- الأسئلة ذات الإجابات القصيرة (عبارة أو جملة).
 - د- أسئلة الإكمال (ملء الفراغات).
- اختبارات التعرف (وهي التي يختار فيها الطالب الإجابة من عدة إجابات معطاة)، وتنقسم إلى:
 - أ- أسئلة الصواب والخطأ (الإجابة البديلة).
 - ب- أسئلة الاختيار من متعدد (اختيار إجابة من عدة إجابات).
 - ج- تمرينات المطابقة (مطابقة عدد من الإجابات مع عدد من الأسئلة).

واختبارات الاستدعاء التي يعطي فيها الطالب الإجابة من عنده تمثل عملاً أو مهمة أقل تحديداً من اختبارات التعرف، وبالتالي فمن الصعب ضبط استجابة الطالب فيها. وأعظم صعوبة نواجهها في النوع الأول من الأسئلة، وهو المقال ذو الإجابة المستفيضة، فهذا النوع من أنواع الأسئلة عام في طبيعته، مثال ذلك: 'صف

دور الصحافة كسلطة رابعة'. فهذا النوع من الأسئلة يتيح للطلاب حرية غير محدودة تقريبا في اختيار المعلومات التي يضمنها إجابته على السؤال، وفي تنظيم هذه المعلومات كما أن تقويم الإجابة غير محدد كذلك. وبذلك فإن هذا النوع يمكن استخدامه في قياس أهداف مثل القدرة على اختيار المادة المناسبة، والقدرة على تنظيم الأفكار، وتقويمها، والقدرة على الابتكار، ولكن قيمته محدودة للغاية في قياس إجابات قصيرة خاصة. وهو غير مناسب لقياس معرفة حقائق معينة أو بعض أنواع الفهم مثلا، لأن هذه الأهداف يمكن لها أن تظهر أو لا تظهر في الإجابة، نظرا للحرية الكبيرة المعطاة للطلاب. ورغم أنه يمكن وضع بعض الضوابط على إجابة الطالب خلال تعليمات معينة، إلا أنه من الضروري إذا كنا نريد قياس نواتج تعلم أكثر تحديدا أن نستخدم نوعا آخر من الأسئلة أكثر تحديدا في بنائه.

فاختبارات المقال ذات الإجابة المحدودة مثل (اذكر اثنين من مزايا إن شاء مترو الأنفاق) تحدد نوع إجابة الطالب على السؤال وطولها وتنظيمها. ويلاحظ أن هذا الضبط لإجابة الطالب يجعل السؤال أكثر فائدة كمقياس لفهم الحقائق الخاصة، ولكنه غير مناسب كمقياس للقدرة على تنظيم الأفكار. وهذا يوضح أهمية اختيار نوع السؤال الأكثر ملاءمة لقياس نواتج التعلم المرغوبة.

وتمدنا أسئلة الاستدعاء الأخرى بوسيلة أكثر تحديدا لاس تجابة التلميذ. ونظرا لأن أسئلة الإجابة القصيرة وأسئلة الإكمال تتطلب كتابة عبارة أو كلمة فقط، فإنها تكاد تكون قاصرة على قياس نواتج التعلم المتعلقة بالتذكر مثل تذكر الحقائق.

أما النوع الثاني من الأسئلة وهو نوع التعرف الذي يختار فيه الطالب إجابة من إجابتين أو أكثر، فإن مفردات الاختبار محددة تماما، ويمكنها قياس العديد من نواتج التعلم من البسيط إلى المعقد. فأسئلة الصواب والخطأ مثلا يمكن أن تحث على قضايا تتعلق بجوانب معينة مثل الحقائق، والمبادئ، والقوانين، والتطبيقات والتفسيرات، ويطلب من التلميذ أن يبين إذا ما كانت القضية المعروضة صوابا أم خطأ. وتمارين المطابقة عبارة عن مجموعة من المقدمات والإجابات لمطابقتها مع بعضها البعض. وقد تكون هذه مصطلحات، أو تعاريف، أو قواعد، أو أمثلة، أو مبادئ، أو أمثلة لمبادئ، وما شابه ذلك. وتمثل أسئلة الاختيار من متعدد مشكلة أو سؤالاً، وعدة حلول بديلة يختار الطالب منها الاستجابة الصحيحة. وهذه البدائل قد تكون عبارات تتعلق بحقائق، أو أمثلة لإجراءات، أو أمثلة لمبادئ، أو أي نوع آخر من الإجابة. ويلاحظ رغم الطبيعة الخاصة لكل نوع من هذه الأسئلة أن الاستجابة التي يقوم بها التلميذ لكل منها إجابة محددة، إذ يجب عليه التعرف على صحتها

على القضية المعروضة كصواب أو خطأ أو مطابقة المقدمة مع اسد تجاباتها الصحيحة، أو اختيار الإجابة الصحيحة من بين الإجابات المقترحة، أو الحل والمحتمة للمشكلة. فهو ليس حراً في إعادة تحديد المشكلة أو إعطاء إجابات صحيحة جزئياً، أو إعطاء معلومات لا تلائم المطلوب في السؤال. ولذلك فإن الضبط الكبير الذي تتصف به أسئلة التعرف يزيد احتمال مطابقة الاستجابات التي يطلبها السؤال لتلك التي تعكسها نواتج التعلم التي نريد قياسها.

ويطلق على اختبارات التعرف عادة الاختبارات الموضوعية، نظراً لأن إجاباتها محددة لا يختلف اثنان على تقدير درجاتها، وفي بعض الأحيان تعتبر أسئلة الإكمال أيضاً من النوع الموضوعي، ولكنها في الواقع تقع في موقع وسط بين أسئلة المقال من ناحية، والأسئلة الموضوعية من ناحية أخرى.

٢- الاختبارات العملية: (اختبارات الأداء) تتعلق الاختبارات العملية بالمجال النزوعي للأهداف، أي بالمنتجات النفسية الحركية (المهارات الحركية)، وكذلك ببعض جوانب الأهداف المعرفية. وهذا النوع من الاختبارات هام للغاية في بعض مجالات التدريس مثل العلوم (مهارات المعمل)، والرياضيات (المهارات العملية لحل المشكلات)، واللغة العربية واللغات الأجنبية (مهارات الاتصال والتفاهم والكتابة)، والاجتماعيات (مهارات رسم الخرائط والرسوم والتعامل بفاعلية في الجماعة). كما أن الأهداف النزوعية ترتبط ارتباطاً وثيقاً ببعض المقررات مثل الموسيقى والرسم والتعليم التجاري، والاقتصاد المنزلي والتربية الرياضية.

وقد يرغب الباحث في بعض الأحيان في قياس أداء الطالب أي قياس ما يستطيع عمله. وفي هذه الحالة لابد للباحث من بناء اختبار عملي. والاختبار العملي وسيلة لقياس قدرة الفرد على أداء أنواع معينة من السلوك، وتستخدم هذه الاختبارات استخداماً واسعاً في مجال قياس الاستعدادات العقلية حيث يفرض إجراءها على الأطفال الذين يعانون من مشكلات في النطق أو اللغة، وبهذا المعنى يعتبر الاختبار العملي مرادفاً للاختبار غير اللفظي، ويستخدم أساساً للتمييز بين الأداء اللفظي والأداء العملي اليدوي.

وللاختبارات العملية أهميتها في مجال التحصيل الدراسي، إذ أنها تهدف إلى توفير وسيلة موضوعية لقياس مخرجات التعلم المتعلقة بالمنتجات الحركية والعملية. كما هو الحال في مقررات العلوم حيث تتعلق بعض مخرجات التعلم بالمهارات العملية في المعمل، ومقررات الاجتماعيات التي تهتم بالمنتجات المتعلقة برسم الخرائط والرسوم البيانية والتعامل بفاعلية مع الجماعة. كما تهتم اللغة

العربية واللغات الأجنبية بمهارات النطق والكتابة. وللمهارات العملية لها أهميتها في المواد الدراسية الأخرى كذلك.

ويمكن تعريف الاختبار العملي بأنه الاختبار الذي يتم فيه محاكاة الموقف الطبيعي بدرجة أكبر مما تسمح به اختبارات القلم والورقة. ويعتبر الموقف الطبيعي في هذه الحالة محك الأداء. وطبقا لهذا التعريف فإن محاكاة الموقف الطبيعي يعتبر أمرا يتعلق بالدرجة، ولهذا تختلف الاختبارات العملية من حيث بعدها أو قربها من محاكاة الموقف الطبيعي. وكلما اقتربنا من الموقف الطبيعي كانت درجة الواقعية في موقف الاختبار أكبر.

ويركز الاختبار العملي على إجراءات العمل، أو على الإنتاج أو على الاثنين معا. وعادة ما تحدد طبيعة الموقف أين يكون التركيز. وبعض أنواع الأداء لا يترتب عليها نتائج محسوسة، مثال ذلك استخدام أجهزة المعلم، أو إلقاء حديث، أو عزف آلة موسيقية، ومختلف الألعاب البدنية مثل السباحة والرقص وإلقاء الكرة. وتتطلب مثل هذه الأنشطة قياس الأداء أثناء القيام به، مع توجيه الانتباه إلى ما يتضمنه الأداء من حركات ومدى تتابعها تتابعا سليما.

خطوات بناء الاختبار العملي:

- أ- تحديد مخرجات التعلم التي نريد قياسها.
- ب- اختيار المهمات التي يتضمنها الاختبار.
- ج- إعداد تعليمات الاختبار.
- د- إعداد أسلوب الملاحظة المستخدم في قياس الأداء.

٣- الاختبارات الشفوية: الاختبارات الشفوية أقدم الوسائل التي استخدمت لتقويم التحصيل، وما زالت تستخدم حتى الآن استخداما واسعا، وتعتبر أفضل وسيلة لتقويم بعض الأهداف التربوية، وخاصة ما يتعلق منها بقدرة التلميذ على التعبير عن نفسه لفظيا وشفويا، فإنه لا يوجد وسيلة أخرى لقياس هذه القدرة.

ويقصد بالامتحانات الشفوية أسئلة غير مكتوبة تعطى للتلاميذ، ويطلب منهم الإجابة عليها دون كتابة، والغرض منها معرفة مدى فهم التلميذ للمادة الدراسية، ومدى قدرته على التعبير عن نفسه، لهذا فهي تعطى كوسيلة تقدير أو تقويم بجانب وسائل التقويم التحريرية.

وللاختبارات الشفوية أهميتها الخاصة في مجال الدراسات اللغوية لأنها من أنسب الوسائل للتعرف على قدرة التلميذ على النطق والتعبير.

ومن أهم عيوب الاختبارات الشفوية ما يلي:

أ- تتأثر بعيوب التقدير الذاتي، إذ أن المدرس يحكم على مدى كفاية الإجابة على السؤال، ويصدر حكماً ذاتياً عليها، ومثل هذا التقدير قد يتأثر بالحالة النفسية للمدرس. وما قد يحكم عليه في موقف ما على أنه إجابة سليمة، قد يحكم عليه بسبب تغير اتجاهه في موقف آخر على أنه إجابة غير مرضية. كما أن حكم المدرس على الإجابة قد يتأثر بفكرته عن التلميذ نفسه، فقد يتغاضى عن خطأ بسيط للتلميذ المتفوق، على أساس فكرته أنه يعرف أكثر من ذلك، في حين أن نفس الإجابة من طالب ضعيف قد تؤخذ على أنها دليل واضح على ضعفه.

ب- اختلاف مستوى صعوبة الأسئلة التي توجه للطلبة المختلفين، فقد يبدو حصول أحد التلاميذ مرضياً لأن السؤال الذي وجه إليه كان سهلاً، وأمكنه الإجابة عليه، في حين أن تلميذاً آخر قد يوجه إليه سؤال صعب لا يستطيع الإجابة عليه، ويتم الحكم على تحصيله بأنه غير مرض، رغم معرفته بإجابة السؤال الأسهل.

ج- لا يمكن أن تكون الأسئلة الشفوية شاملة لجميع قدرات التلاميذ ففي فصل مكون من ثلاثين تلميذاً يكون من الصعب سؤال كل تلميذ أكثر من ثلاثة أو أربعة أسئلة خلال ساعة واحدة، وقد يكون متوسط عدد الأسئلة التي توجه للتلميذ أقل من هذا، ونتيجة لذلك يكون من الصعب الحصول على صورة صادقة وكاملة عن مدى ما يعرفه الطلاب في مادة الامتحان باستخدام هذه الوسيلة بمفردها.

د- قد يتأثر تقدير الدرجة في الاختبار الشفوي بعوامل أخرى بعيدة عن المادة الدراسية، وبالتالي لن يتم التقويم في ضوء الأهداف المحددة للتدريس، مثال ذلك القدرة اللغوية للتلميذ، وقدرته على مواجهة الموقف نفسه، أي أن العوامل النفسية، وعوامل القدرة الشخصية للطلاب قد يكون لها أثرها في تقدير إجابته، وهذه العوامل ليست مرتبطة بالأهداف الحقيقية التي يتم التقويم على أساسها.

ولهذه الأسباب يجب ألا يستخدم الباحث الامتحانات الشفوية كوسيلة لجمع البيانات، إلا إذا وجد أن طبيعة بحثه تتطلب هذا النوع من أدوات جمع البيانات، فإذا كان المتغير التابع في البحث يتعلق باللغة العربية، أو إحدى اللغات الأجنبية، وكان البحث يتطلب قياس التعبير اللغوي الشفوي، فقد يجد الباحث من الضروري لقياس التعبير الشفوي استخدام اختبار شفوي.

المبادئ الأساسية للاختبارات التحصيلية:

١- يجب أن تقيس الاختبارات التحصيلية نواتج محددة ومتنوعة للتعليم تنسجم مع أهداف التدريس، مثل معرفة الحقائق، ومعرفة المصطلحات وفهم المبادئ والمفاهيم، والقدرة على تطبيق الحقائق والمبادئ، ومختلف مهارات التفكير. ولذلك فإن الخطوة الأولى في بناء الاختبارات التحصيلية ليست كتابة الأسئلة وفقرات الاختبار، ولكنها تحديد نواتج أو مخرجات التعلم التي تقيسها. ويجب أن تتبع هذه النواتج من أهداف تدريس الوحدة أو الوحدات التي يرغب الباحث في وضع اختبار تحصيلي لها.

٢- يجب أن تقيس الاختبارات التحصيلية عينة ممثلة لنواتج التعلم للمادة التي تحتويها موضوعات الوحدة أو الوحدات التي يرغب الباحث في وضع اختبار تحصيلي لها. فالاختبارات والمقاييس موجهة دائماً نحو عينات من السلوك، ولا يمكن أن نسأل كل الأسئلة التي نريد أن نوجهها في الاختبار. ففي نهاية كل خبرة تعليمية توجد عادة مئات من الحقائق والمصطلحات التي نتوقع أن يعرفها الطالب، ولكن نظراً للوقت المسموح به لإجراء الاختبار، ولاعتبارات أخرى، لا تقيس أسئلة الاختبار إلا جزءاً محدوداً من هذه الحقائق والمعلومات. ومعنى هذا أن الاختبارات التحصيلية لا تغطي إلا عينة محدودة نسبياً لسلوك الذي نريد قياسه، ولذلك يجب على الباحث عند وضع الاختبار أن يتأكد من أن عينة الأسئلة التي يشملها الاختبار، عينة ممثلة لمجتمع المواقف التي يفترض فيها أن تتمثل.

ولذلك نلجأ عادة إلى استخدام ما يسمى بجدول المواصفات، وهو وجدول يحدد أهداف التدريس وموضوعات المقرر التي يشملها الاختبار، والعلاقة بينهما، كما يحدد عدد الأسئلة في الاختبار.

٣- يجب أن تحتوي الاختبارات التحصيلية على أنواع الأسئلة المناسبة لقياس نواتج التعلم المرغوبة. ذلك أن مخرجات التعلم لمقرر ما هي التي تحدد أنواع السلوك التي يمكن لنا قبولها كأدلة على تحقيق أهداف المقرر. والاختبار التحصيلي ليس إلا وسيلة لاستدعاء السلوك الخاص حتى يمكن الحكم على الدرجة التي حدث بها التعلم. ومفتاح الاختبار التحصيلي الصادق هو اختبار نوع الأسئلة المناسب وبناء الأسئلة بعناية حتى يمكن لها استدعاء الاستجابة المرغوبة، واستبعاد الاستجابات الأخرى غير الملائمة. ويتضمن استدعاء الاستجابة المرغوبة عن طريق سؤال معين ضبطاً لاستجابات الطلاب لمفردات

الاختبار. وتتغير درجة ضبط هذه الاستجابات حسب أنواع الأسئلة المستخدمة، فيكون الضبط أعلى ما يمكن في الأسئلة الموضوعية (أسئلة التعرف)، وأقل ما يمكن في أسئلة المقال المستفيض.

٤- يجب أن تكون الاختبارات التحصيلية ثابتة على قدر الإمكان كما يجب تفسيرها بحرص: يعتبر الاختبار ثابتاً إذا اتفقت الدرجات التي حصل عليها الطلبة في الاختبار مع الدرجات التي حصلوا عليها في إجراء سابق للاختبار أو في صورة مماثلة له. وتحتوي كل درجات الاختبار على بعض الخطأ (نتيجة لتباين ظروف إجراء الاختبار، واستجابات الطلبة وغير ذلك من العوامل)، ولكن نسبة الخطأ يمكن خفضها إلى أقل حد ممكن وذلك عن طريق إعداد اختبار سليم. وبشكل عام يمكن زيادة ثبات الاختبارات بزيادة طوله، وتحسين نوعية كل سؤال. فالاختبارات الطويلة تسمح بوجود عينة أكثر تمثيلاً للسلوك المراد قياسه كما أن الاختبار الذي يصمم بشكل سليم يعطينا وصفاً أدق لأداء كل فرد في الاختبار.

وعندما يكون الغرض من قياس التحصيل هو الحصول على الترتيب النسبي للطلاب (بغرض بناء اختبار معياري المرجع) فإنه يمكن الحصول على ثبات أكبر ببناء اختبار يمدنا بمدى واسع من درجات الاختبار، ويمكن تحقيق ذلك بكتابة أسئلة مستوى صعوبتها ٥٠٪ (أي أن ٥٠٪ من الطلبة يجيبون إجابة صحيحة على السؤال)، مع الاهتمام بوجه خاص بقدرة كل سؤال على التمييز بين مجموعات التحصيل العليا والدنيا من الطلبة.

وإذا كان الغرض من الاختبار هو تحديد مهمات التعلم التي يستطيع (أو لا يستطيع) الطالب أدائها (بناء اختبار محكي المرجع)، تكون صعوبة المهمة التعليمية التي نريد قياسها هي التي تحدد صعوبة السؤال. ولذلك يجب مطابقة محتوى السؤال مطابقة تامة للمهمة التعليمية التي نريد قياسها، وذلك باستخدام عدد وافر من الأسئلة لكل مهمة تعليمية. وكتابة عدد من الأسئلة يمكنها بوضوح استدعاء السلوك المرغوب.

ثانياً: اختبارات الاستعدادات العقلية:

يحتاج الباحث أحياناً إلى استخدام اختبارات الاستعداد العقلي لقياس بعض المتغيرات العقلية في بحثه. وقد يستخدمها الباحث كوسائل لضبط المتغيرات الخارجية أو المتغيرات الدخيلة. وليس من السهل بناء اختبارات الاستعدادات كما هو الحال في قياس التحصيل الدراسي أو قياس الاتجاهات. ولذلك يجب أن يلجأ

الباحث إلى الاختبارات المقننة في البيئة التي يجري فيها بحثه. ويجب أن يراعى الباحث عند اختيار اختبار معين لاستخدامه كأداة لجمع البيانات في البحث أن يكون تقنين الاختبار المستخدم قد تم على عينة شبيهة بعينة البحث. وأن الاختبار مستوف لجميع شروط الاختبار الجيد من حيث الصدق والثبات.

وتتميز معظم اختبارات الذكاء والاستعدادات والقدرات بأنها ذات ثبات عال (رجاء أبوعلام ونادية شريف، ١٩٩٥)، فمعاملات ثبات هذه الاختبارات (بإحدى طرق قياس الثبات)، تتراوح عادة بين ٨٥ و ٩٥، وهذه المعاملات مبنية على بيانات جمعية، وتشير هذه البيانات إلى أن الأفراد في المجموعات المختلفة يميلون إلى الاحتفاظ بمكانتهم النسبية من وقت لآخر، وأن معظم الأفراد لن تتغير درجاتهم بشكل ملحوظ من موقف اختبار إلى موقف اختبار آخر. إلا أن بعض الأفراد قد تختلف درجاتهم اختلافا كبيرا، إذ نرى أحيانا أن نسبة ذكاء فرد ما تختلف بمقدار ٢٠ درجة، من جلسة اختبار إلى جلسة أخرى. وهذا صدح بالذات نسبة للأطفال الصغار الذين يلاحظ أن ثبات الاختبارات بالنسبة لهم ليس بنفس درجة ثبات الاختبارات للأطفال الأكبر والراشدين.

ويمكن تصنيف اختبارات الاستعدادات العقلية في نوعين:

- ◆ اختبارات الاستعداد العقلي العام، أو اختبارات الاستعداد الدراسي العام، وهي ما اصطلح على تسميته باختبارات الذكاء.
 - ◆ اختبارات الاستعداد العقلي الخاص، وهي التي تقيس القدرات والاستعدادات العقلية الخاصة في المجال العقلي المعرفي.
- وسنتناول فيما يلي هذين النوعين باختصار.

أ- اختبارات الاستعداد العقلي العام: نظرا لتعدد النواحي المعرفية العقلية التي يظهر فيها أثر القدرة العقلية، فإن أي اختبار للذكاء ليس إلا مقياسا لعينة من مظاهر الحياة العقلية المعرفية العامة (رجاء أبوعلام ونادية شريف، ١٩٩٥). وتشمل هذه العينة السلوك العقلي كما يبدو في أداء الاختبارات اللفظية والعملية. وطبيعي أنه كلما شمل المقياس أنواعا متعددة من السلوك العقلي كان صدق الاختبار عاليا.

- ويطلق على اختبارات الذكاء أحيانا اختبارات الاستعداد الدراسي العام. ويفضل الكثيرون التسمية الأخيرة على تسمية الذكاء لاعتبارات كثيرة منها:
- ١- أن كثيرا من الناس أصبحوا يقرنون مصطلح الذكاء بالقدرة الموروثة.
 - ٢- هناك كثير من الجدل حول معنى الذكاء والعوامل التي ترتبط بهذا المفهوم.

٣- كثر استخدام الاختبارات في هذا المجال للتنبؤ بالتحصيل الدراسي ووصف قدرات التعلم. ولذلك بدلا من مصطلح /اختبار الذكاء/ ظهرت مصطلحات أخرى مثل /اختبارات القدرة على التعلم/، و/اختبارات القدرة المدرسية/، و/اختبارات القدرة المعرفية/، و/اختبارات الاستعدادات الدراسية/. وتؤكد كل هذه المصطلحات حقيقة أن هذه القدرات تقيس القدرات المتعلمة التي تقيس في التعلم، وليس القدرة الفطرية أو القدرة غير المتعلمة.

وأفضل تفسير لدرجات التلاميذ في اختبارات الاستعداد الدراسي أو اختبارات القدرات هو ذلك التفسير الذي يقوم على أنها مقاييس للقدرة على التعلم. ويتأثر أداء التلاميذ في هذه الاختبارات بعوامل كثيرة منها عوامل وراثية، ومنها عوامل تتعلق بالخبرة والخلفية الثقافية، والدافعية، وبعض المهارات الخاصة (مثل القراءة، وأداء الاختبار)، والانتباه، والإصرار، والثقة بالنفس، والتوافق الاجتماعي. وكل هذه العوامل جزء من قدرات الفرد الحالية على الأداء، ولذلك فلها تأثيرها على درجات الاختبارات وعلى التحصيل المدرسي. ويمكن تعديل كثير من هذه القدرات بالخبرات المتعلمة. ويترتب على ذلك تحسن كل من التحصيل المدرسي والقراءة على التعلم. ونخطئ في استخدام النتائج عندما نفسر درجات الاختبارات على أنها مقاييس لقدرات التعلم غير قابلة للتعديل.

ويلاحظ أن الأداء في اختبارات الاستعداد الدراسي يعتمد بدرجة معينة من الدرجات على خبرات الحياة، وبخاصة الخبرات المدرسية. فأي اختبار لفظي يتطلب من الشخص أن يكون قد تعلم أن يتكلم اللغة التي وضع بها الاختبار، وأي اختبار مصور يحتاج أن يكون لدى الفرد ألفة بالموضوعات المصورة، وأي اختبار مهما كان نوعه يتطلب أن يكون الفرد قد تعلم كيف يبذل جهدا، وأن يحاول أقصى جهده في موقف الاختبار.

ويمكن تصنيف اختبارات الاستعداد الدراسي في نوعين: النوع الأول طبقا لأسلوب الأداء، ويشمل الأداء الجمعي، والأداء الفردي. والنوع الثاني طبقا لمحتوى الاختبار، ويشمل الاختبار اللفظي، والاختبار غير اللفظي. وبناء على هذين التصنيفين يصبح لدينا أربعة تصنيفات فرعية يمكن توضيحها في الجدول رقم (١٨-١).

وبناء على الجدول (١٨-١) نجد أن أهم تصنيفات اختبارات الذكاء هي:

- ١- اختبارات فردية لفظية.
- ٢- اختبارات فردية غير لفظية.

٣- اختبارات جمعية لفظية.

٤- اختبارات جمعية غير لفظية.

جدول ١٨-١ تصنيف اختبارات الذكاء (اختبارات الاستعداد الدراسي العام)

محتوى الاختبار	لفظي	غير لفظي
أسلوب الأداء		
فردية	فردية لفظية	فردية غير لفظية
جمعية	جمعية لفظية	جمعية غير لفظية

ومن أمثلة اختبارات الاستعداد الدراسي العام مقياس ستانفورد بينيه للذكاء، ومقاييس وكسلر للذكاء وهي مقاييس فردية تجمع بين المحتوى اللفظي والمحتوى غير اللفظي. ومن الاختبارات اللفظية الجمعية اختبار الذكاء اللغوي واختبار القدرة العقلية العامة، وغير ذلك من الاختبارات، ومن أمثلة الاختبارات الجمعية غير اللفظية اختبار المصفوفات المتتابعة.

ب- اختبارات الاستعداد العقلي الخاصة: رأينا أن اختبارات الاستعداد الدراسي العام تقيس القدرة العقلية العامة، أو القدرة على الأداء في المدرسة بشكل عام. أما اختبارات الاستعدادات الخاصة فإنها تهدف إلى التنبؤ بمستوى أداء الفرد في مجال خاص. وتصلح اختبارات الاستعدادات الخاصة إلى قياس متغيرات كثيرة في المجال الأكاديمي وغير الأكاديمي، ومن أمثلة هذه المجالات الرياضيات واللغة، والاستدلال الميكانيكي، والاستدلال المجرد. ويكثر استخدامات هذه الاختبارات في البحوث كأساس لتكوين مجموعات متكافئة حسب متغيرات معينة يختارها الباحث، مثل القدرة الرياضية، أو القدرة اللغوية، أو غير ذلك من القدرات والاستعدادات. ومن أهم اختبارات الاستعدادات التي استخدمت في مصر بطارية الاستعدادات الفارقة، وتتألف هذه البطارية من ثماني اختبارات هي الاستعداد اللفظي، والقدرة الرياضية، والاستدلال المجرد والقدرة على استخدام اللغة، واختبار الإملاء، والاستدلال الميكانيكي، والعلاقات المكانية.

ثالثاً: اختبارات الشخصية:

من أهم المجالات التي يهتم بها الباحثون مجال قياس الشخصية. وهناك عدد من مقاييس الشخصية، يعكس كل منها اهتماماً بجانب نظري خاص، فبعضها يهتم بنظريات السمات، في حين يهتم بعضها الآخر بنظريات التحليل النفسي.

والدافعية. ولذلك يجب أن يحدد الباحث بالضبط المتغيرات التي يريد قياسها قبل أن يختار أداة القياس. وأكثر مقاييس الشخصية استخداما في البحوث هي اسد تنبيانات الشخصية والأساليب الإسقاطية.

أ- **استبيانات الشخصية:** تتكون استفتاءات الشخصية من مجموعة من العبارات تصف السلوك موجهة للمستجيب، وعليه أن يجيب على كل عبارة أو سؤال بأحد الاختيارات: 'نعم' - 'لا'، أو 'موافق' - 'غير موافق' - 'غير متأكد'، أو غير ذلك من الاختيارات المشابهة. ودرجة الفرد في الاستفتاء هي مجموع استجاباته في السمة التي يقيسها الاستفتاء. مثال ذلك أن الأفراد الذهانيين يستجيبون عادة بـ 'نعم' للعبارة التالية: 'يتكلم الناس دائما من وراء ظهري' ولا للعبارة التالية: 'أنتوقع من القاضي أن يكون عادلا'. وبالطبع فإن مثل هذه الاستجابات لعبارتين فقط لا يمكن أن تكون دليلا على نزعات ذهانية. ولكن مثل هذه الاستجابات على عدد كبير من العبارات أو الأسئلة قد تكون مؤشرا بوجود ميول ذهانية.

وبعض استفتاءات الشخصية تقيس بعدا واحدا (سمة واحدة) من أبعاد الشخصية وبعضها الآخر يقيس أكثر من بعد أو سمة. فاستفتاء الشخصية العاملي مثلا (رجاء أبوعلام ونادية شريف، ١٩٨٥) يقيس ١٦ بعدا من أبعاد الشخصية. وقد وضع الباحثان معايير جديدة لهذا الاختبار اختصرت الأبعاد الستة عشرة إلى سبعة أبعاد (نادية شريف ورجاء أبوعلام، ١٩٨٩). ومن الأمثلة الأخرى لاسد تنفتاءات الشخصية 'قائمة موني للمشكلات'.

وتستخدم استفتاءات الشخصية في البحوث التربوية والنفسية للحصول على أوصاف لسماات أفراد العينة في بعض المجموعات المددة في البحوث مثل 'المتأخرين دراسيا' و'المتسربين من المدرسة' وهكذا. كما تستخدم كذلك في البحوث التي تتناول العلاقة بين سمات الشخصية وبعض المتغيرات مثل الاستعدادات العقلية، والتحصيل الدراسي، والاتجاهات.

وتتميز اختبارات الشخصية بأنها اقتصادية، وبسيطة، وموضوعية. إلا أن لها عددا من العيوب التي كثيرا ما تؤدي إلى انخفاض صدقها، ومن أهم هذه العيوب ما يلي:

- ١- شفافية معاني الأسئلة مما يتيح للمستجيب أن يزيّف الصورة التي يعطيها عن نفسه، سواء كانت في الجانب الأفضل أو في الجانب الأسوأ.
- ٢- أن الاستجابات تعتمد على معرفة الفرد لنفسه، وكيف يتصرف في المواقف المختلفة، فإذا كانت صورة الفرد مشوشة عن نفسه، أو كان الفرد لا يعرف

- نفسه معرفة صحيحة، فلن تكون استجاباته معبرة عن سماته الواقعية.
- ٤- أن على المستجيب أن يختار استجابة من الاستجابات المعروضة (نعم، لا) مثلا، أو يختار (عبارة من بين عدد من العبارات) دون أن يستطيع أن يضيف إلى هذه الاستجابة أو يحددها أو أن يتحفظ إزاء بعض المواقف أو الظروف.
- ٥- ضرورة معرفة القراءة مع الفهم الدقيق لما يقرأ، وهذا لا يتوفر لعدد كبير من الأفراد العاديين.
- ٦- لا توضح هذه الاختبارات الأسس والدوافع التي تجعل المستجيب يختار إجابة دون غيرها، مما يخفي جانبا من تفسير سلوك المستجيب.

ب- الأساليب الإسقاطية: تؤكد استقراءات الشخصية الـ سابق ذكره لـ الجانب الموضوعي، ولذلك تتميز بسهولة إجراءاتها وتقدير درجاتها، كما تتميز بصدق وثبات مرتفع، إلا أنها ذات طبيعة محددة إذ تتكون من أسئلة أو عبارات معينة يستجيب لها الفرد باختيار أحد الاستجابات المقترحة له، مما يحد من حريته في التعبير. أم لـ الأساليب الإسقاطية فهي على العكس من ذلك تسمح للفرد بحرية كبيرة مما يمكنه من التعبير عن نفسه فيما يقدم من استجابات.

ويتكون الاختبار الإسقاطي من مثير غامض يستجيب له الفرد استجابة حرة بالطريقة التي يريدها. ونظر لأن المثير الغامض لا يتطلب استجابة معينة، فإن الفرد يسقط ذاته وشخصيته على هذا المثير، تماما كما يسقط جهاز العرض الصور على شاشة العرض. فالاختبارات الإسقاطية تتطرق إلى مخرطة الفرد، ومن المفروض، أنه نتيجة لعمليات التخيل يكشف الفرد بعضا من نفسه أثناء استجابته للاختبار. ومن أهم الأساليب الإسقاطية: اختبار رورشاخ الذي يستخدم بقع الحبر كمثيرات له، واختبار تفهم الموضوع الذي يتكون من مجموعة من الصور التي يبني المستجيب حولها قصة من القصص.

ويستخدم الأساليب الإسقاطية على وجه الخصوص الأخصائيون النفسيون الإكلينيكيون لدراسة وتشخيص المشكلات الانفعالية للفرد. وتحتاج هذه الأساليب لإجرائها إلى مختصين على مستوى عال من التدريب في تطبيقها، ولذلك فإن استخدامها في البحوث النفسية والتربوية محدود للغاية نظرا للتكاليف العالية التي قد يتطلبها الحصول على مختصين مدربين على تطبيق هذه الأساليب، كما أن بعض الباحثين يعتقدون أن صدق هذه الأساليب لم يصل بعد إلى الدرجة التي يمكن بها الاعتماد على نتائجها.

رابعاً: مقاييس الاتجاهات:

تستخدم بعض المقاييس المدرجة لقياس الاتجاهات والقيم والآراء وغير ذلك من المتغيرات التي لا يمكن قياسها بالاختبارات أو غير ذلك من الأساليب. والمقياس المدرج مجموعة من الفئات أو القيم العددية التي تعطى للصفة أو السلوك وفقاً لاستجابات الفرد بغرض قياس بعض المتغيرات. ويطلق على عملية إعطاء الدرجات لهذه الصفات لقياس مفهوم من المفاهيم تحديد قيم المقياس، أو ترتيب المقاييس. وتختلف المقاييس المدرجة عن الاختبارات في أنها لا تهدف لنجاح الفرد أو رسوباً أو تبين نواحي قوة أو نواحي ضعف، ولكنها تقيس الدرجة التي يظهر بها الفرد خاصية من الخواص. فقد يستخدم باحث مقياس الاتجاه لقياس اتجاهات طلبة الجامعة نحو 'الديمقراطية' أو غير ذلك من الموضوعات. وتستخدم مقاييس الاتجاهات استجابات متعددة - عادة ما تكون استجابات لبعض الأسئلة أو العبارات - وتجمع هذه الاستجابات في درجة واحدة (Shaw & Wright, 1967).

ويقصد بالاتجاه نحو موضوع ما، الاعتقاد أو عدم الاعتقاد فيه، فقد نعتقد أن شيئاً ما على صواب، وأن شيئاً آخر على خطأ، ويتضمن الاتجاه ثلاثة جوانب رئيسية:

◆ **هدف** هو موضوع الاتجاه، وهذا الموضوع يرتبط بعوامل معرفية هي

ما يفهمه الفرد أو يعرفه عن الموضوع.

◆ **حالة انفعالية وجدانية** هي الشعور نحو الموضوع بشعور معين سواء

كان موجباً أو سالباً.

◆ **توجيه السلوك**، فبناء على الحالة الانفعالية لدى الفرد نحو موضوع

الاتجاه نجده ينزح إلى القيام بسلوك معين مؤيد أو معارض للموضوع.

وكثيراً ما يهتم الباحثون في التربية وعلم النفس باتجاهات الأفراد أو الجماعات نحو موضوعات معينة. ويقصد بالاتجاه استجابة موجبة أو سالبة للفرد نحو موضوع، أو مؤسسة، أو مفهوم أو قضية ذات صبغة اجتماعية غالباً. وتستخدم مقاييس الاتجاهات بأنه من الممكن توزيع الأفراد على طول متصل ذي قطبين: أحدهما يعبر عن الجانب الموافق أو الجانب المؤيد والثاني عن الجانب المعارض. وغالباً ما يضطر الباحثون إلى بناء مقاييس اتجاهات خاصة ببحثهم. وسوف نناقش فيما يلي نوعين من مقاييس الاتجاهات: مقياس ليكرت (Likert) (طريقة التقدير المجمع) والمقياس أوسجود (Osgood) (مقياس المعاني الفارق).

أ- **مقياس ليكرت (طريقة التقدير المجمع):** مقياس ليكرت

(Likert, 1932) من أكثر الأساليب استخداماً لقياس الاتجاهات. ويتكون مقياس

ليكرت من مجموعة من العبارات تقيس الاتجاهات نحو موضوع معين ويطلب من المستجيبين الاستجابة لكل عبارة بأحد الاستجابات التالية: 'موافق تماما' - 'موافق' - 'غير متأكد' - 'غير موافق' - 'غير موافق بالمرّة'، وتعطى كل استجابة من هذه الاستجابات قيمة عددية، ونحصل على درجة المقياس بجمع استجابات الفرد لعبارات المقياس ويعبر المجموع عن اتجاه الفرد نحو موضوع الاتجاه.

ويتم بناء مقياس ليكرت بتجميع عدد كبير من العبارات عن الموضوع. ومن المفروض أن تعبر نصف هذه العبارات عن اتجاه مؤيد، في حين أن نصف العبارات الآخر يعبر عن اتجاه سالب نحو موضوع المقياس. ولا تستخدم العبارات المحايدة في مقياس ليكرت. ويجب أن تكون هذه العبارات عينة ممثلة لكل الآراء أو الاتجاهات الممكنة نحو موضوع الاتجاه. ومن المهم أن يبدأ بناء المقياس بالتفكير في الموضوعات الفرعية التي يتكون منها الاتجاه، وبعدها نكتب عبارات تمثل كل موضوع فرعي. ويجب أن تتصف العبارات بالصفات التالية حتى يكون مقياس الاتجاه تعبيراً صادقاً عن الاتجاه (رجاء أبوعلام ونادية شريف، ١٩٩٥):

- ١- أن تكون العبارة قابلة للجدل أي تعبر عن رأي لا حقيقة.
- ٢- أن تعبر كل عبارة عن موضوع الاتجاه.
- ٣- أن تكون العبارة قابلة لتفسير واحد فقط.
- ٤- أن تكون العبارة بسيطة غير مركبة.
- ٥- أن تكون العبارة قصيرة.
- ٦- أن تكون العبارة كاملة في التعبير عن وجهة نظر معينة متعلقة بالاتجاه.
- ٧- أن تحتوي العبارة على فكرة واحدة فقط.
- ٨- أن تكون العبارة واضحة تماماً ومباشرة.
- ٩- أن تبنى العبارة في صيغة المعلوم وليس صيغة المجهول.
- ١٠- أن تحتوي العبارة على المصطلح موضوع الاتجاه أو النقاش كموضوع لها.

ويجب تحويل التقديرات اللفظية إلى تقديرات رقمية حتى يمكن جمع استجابات الأفراد لعبارات المقياس. وتعطى هذه التقديرات في مقياس ليكرت على النحو التالي: ٥ (موافق تماماً)، ٤ (موافق)، ٣ (غير متأكد)، ٢ (غير موافق)، ١ (غير موافق بالمرّة). وعندما تكون العبارات سالبة تعكس التقديرات بحيث تكون: ١ (موافق تماماً)، ٢ (موافق)، ٣ (غير متأكد)، ٤ (غير موافق)، ٥ (غير موافق بالمرّة). والمثال التالي (جدول ١٨-٢) يوضح ذلك:

جدول ١٨-٢ مثال لمقياس ليكرت

العبارة	موافق تماماً	موافق	غير متأكد	غير موافق	غير موافق بالمرة
الرياضيات هي موضوعي المفضل (+)	٥	٤	٣	٢	١
يجب أن تكون مادة الرياضيات اختيارية (-)	١	٢	٣	٤	٥

فإذا كان اتجاه الفرد مؤيداً للرياضيات فسوف يوافق على العبارة الأولى ولا يوافق على العبارة الثانية، ولذلك قد يختار الاستجابة الثانية (٤) في العبارة الأولى، والاستجابة الرابعة في العبارة الثانية (٤) أيضاً، ويكون مجموع هـ (٨) في هـ هذه الحالة، وهو ما يعبر عن اتجاه موجب. ويعتبر مجموع تقديرات الفرد في مقياس الاتجاه معبراً عن اتجاهه نحو موضوع الاتجاه. وتكون الدرجة العظمى لمقياس الاتجاه مساوية $5 \times n$ (عدد عبارات المقياس). ويلاحظ أنه في مقياس ليكرت لا بد من وجود خمس درجات لكل استجابة، اثنتان منهما للاتجاه الموجب، واثنان للاتجاه السالب، والوسطى للاتجاه المحايد. ولذلك من الخطأ وضع أربع درجات فقط في مقياس الاتجاه حيث إن هذا يخل بالأساس الذي بني عليه.

تحليل فقرات مقياس الاتجاه: بعد بناء مقياس الاتجاه يجب إجراء تطبيق أولي له الغرض منه تحليل عبارات المقياس للتعرف على أفضلها. ويجب أن يعطي التحليل ثلاث نتائج لكل عبارة (Henderson, Morris, & Fitz-Gibbon, 1987):

- ◆ مؤشر قدرة العبارة على التمييز.
- ◆ عدد أو نسبة الأفراد الذين اختاروا كل استجابة.
- ◆ متوسط العبارة وانحرافها المعياري.

ويبين مؤشر التمييز ما إذا كانت العبارة تميز بين الأفراد بنفس الدرجة التي تميز بها الدرجة الكلية للمقياس. ويعني هذا أن الحاصلين على درجات مرتفعة في المقياس ككل يحصلون على درجات مرتفعة في العبارات التي يتكون منها المقياس، وأن الحاصلين على درجات منخفضة في الدرجة الكلية يحصلون على درجات منخفضة في عبارات المقياس. ويمكن الحصول على مؤشرات التمييز بـ إجراء معاملات ارتباط بين درجة كل عبارة والدرجة الكلية للمقياس. وحتى يمكن الاستفادة من أي عبارة في المقياس يجب أن يكون مؤشر تمييزها ٢٥، على الأقل (معامل ارتباط العبارة بالمقياس). ويجب حذف العبارات التي تصل على معاملات ارتباط منخفضة، أو ارتباطات سالبة، لأنها لا تقيس ما تقيسه الدرجة الكلية، وبذلك فهي لا تضيف أي شيء للدرجة الكلية التي تعبر عن الاتجاه. ولذلك يجب على الباحث فحص العبارات غير المميزة وحذفها من المقياس، وعادة ما

تكون مثل هذه العبارات غامضة أو مزدوجة يوجد بها فكرتان أو اتجاهان في نفس الوقت. وقد تؤدي مراجعة هذه العبارات وإعادة صياغتها إلى تدوينها وجعلها صالحة، ولكن في هذه الحالة يجب إعادة تطبيق المقياس مرة أخرى وإعادة حساب مؤشرات التمييز.

أما العمليات الإحصائية الأخرى فتبين مدى استخدام الأف راد للاس تجابات المختلفة في كل عبارة، وتفضل العبارات التي تنتشر فيها استجابات الأف راد على المواقع المختلفة، على العبارات التي تتجمع في استجابة واحدة أو استجابتين.

وبعد اختيار العبارات السليمة الجيدة يجب على الباحث إعادة تطبيق المقياس الجديد على عينة أخرى، وحساب العمليات السابقة مرة أخرى، لمراجعة مؤشرات تمييز العبارات المختلفة.

ب- مقياس أوسجود (أسلوب المعاني الفارق): وهذه طريقة أخرى في قياس الاتجاهات، وإن كانت تستخدم أيضا في قياس المفاهيم. وقد وضع هذه الطريقة أوسجود وزملاؤه (Osgood, Suci, & Tannenbaum, 1967). ويقوم مقياس المعاني الفارق على مسلم أساسي هو أن لكل شيء معنيين بالنسبة للأفراد: رمزي وضمني، ويمكن تقدير كل من هذين المعنيين بشكل مستقل عن الآخر. والمعنى الرمزي هو المعنى الذي يعطيه القاموس للكلمة، أما المعنى الضمني فيشير إلى الارتباطات أو التداخيات التي تثيرها هذه الكلمة في الفرد. ومن الأسهل على الفرد أن يعطي المعنى الرمزي من إعطاء المعنى الضمني. إلا أنه من الأسهل قياس المعنى الضمني للشيء بشكل غير مباشر وذلك بسؤال الأفراد أن يقدروا هذا الشيء باستخدام عدد من الصفات ثنائية القطب. وبذلك يكون معنى شيء ما بالنسبة للفرد هو نمط تقديراته لهذا الشيء على مقاييس الصفات ثنائية القطب.

وقد وجد أوسجود وزملاؤه بعد استخدام التحليل العاملي لتقديرات المقاييس ثلاثة تجمعات للصفات، نتج عنها ثلاثة عوامل هي:

- ١- عامل التقويم ويتكون من صفات مثل 'جيد - سيئ' أو 'قيم - غير قيم'.
 - ٢- عامل القوة ويتكون من صفات مثل 'قوي - ضعيف' أو 'ثقل - خفيف'.
 - ٣- عامل النشاط ويتكون من صفات 'إيجابي - سلبي' أو 'سريع - بطيء'.
- ويبدو أن عامل التقويم هو أكثر العوامل الثلاثة دلالة في قياس الاتجاهات. ويمكن بناء مقاييس المعاني الفارقة باختيار أزواج الصفات التي تقيس عامل التقويم. وتوضع الصفات المتقابلتان عند طرفي المقياس المدرج المكون من سبع درجات، ويختار الفرد الفراغ الذي يحدد درجة الاتجاه بالنسبة له من بين الفراغات

السبع، ويضع علامة (x) في هذا الفراغ. فإذا افترضنا مثلاً أننا نريد قياس اتجاهات الطلبة نحو المدرسة، فإنه يمكن بناء المقياس المبين في الشكل (١٨-٣).

ويلاحظ في هذا المقياس تبادل وضع الصفات الموجبة وال سالبة في الاتجاهين للتقليل من عامل تثبيت الاستجابة، ويقصد 'تثبيت الاستجابة' النزعة إلى تفضيل أماكن معينة في قائمة الصفات. فقد يكون لدى الفرد نزعة لاختيار الطرف الأيمن فيضع كل علاماته في أقصى الطرف الأيمن في كل صفة. ولكن إذا تغير اتجاه المقياس بطريقة عشوائية بحيث لا يكون الطرف الأيمن هو المعبر دائماً عن الصفة الأفضل، يضطر المستجيب إلى قراءة كل صفة والاستجابة لها حسب الاتجاه الفعلي نحو موضوع الاتجاه، وليس طبقاً لموقع مفضل. وعند تقدير درجة مقبـاس المعاني الفارقة تقدر المواقع بتقديرات تتراوح بين ١ و ٧. بحيث تمثل القيمة ٧ أكثر المواقع إيجابية. ففي الفقرة الأولى في المقياس السابق تحصل 'جيد' على تقدير '٧' وهو أعلى تقدير، ويحصل الموقع المجاور على تقدير '٦'، وتـحصل 'سيئ' على تقدير '١' وهو أقل تقدير. وفي الفقرة الثانية يعكس التقدير فتـحصل 'سلبـي' على تقدير '١' وهو أقل تقدير، ويحصل الموقع المجاور على تقدير '٢'، وتحصل 'إيجابي' على تقدير '٧' وهو أعلى تقدير.

المدرسة	
جيد	_____ : _____ : _____ : _____ : _____ : _____ : _____ : _____
سلبـي	_____ : _____ : _____ : _____ : _____ : _____ : _____ : _____
نابه	_____ : _____ : _____ : _____ : _____ : _____ : _____ : _____
حزين	_____ : _____ : _____ : _____ : _____ : _____ : _____ : _____
قيم	_____ : _____ : _____ : _____ : _____ : _____ : _____ : _____
صلب	_____ : _____ : _____ : _____ : _____ : _____ : _____ : _____
ثقيل	_____ : _____ : _____ : _____ : _____ : _____ : _____ : _____
ضعيف	_____ : _____ : _____ : _____ : _____ : _____ : _____ : _____
سريع	_____ : _____ : _____ : _____ : _____ : _____ : _____ : _____

شكل ١٨-٣ نموذج لمقياس المعاني الفارق

وتجمع التقديرات في كل الفقرات وتكون الدرجة الكلية هي المعبرة عن اتجاه الفرد نحو الموضوع، ويمكن أيضا الحصول على متوسط التقديرات ويكون هذا المتوسط هو المعبر عن الاتجاه.

خامسا: مقاييس التقدير:

تستخدم مقاييس التقدير عندما نريد تحديد درجة حدوث سلوك ما. وبخاصة في المواقف التي يكون فيها للأداء أو الإنتاج جوانب متعددة يتطلب كل منها تقديرا خاصا. مثال ذلك عندما يريد الباحث قياس قدرة أفراد العينة على الخطابة، يمكنه إعداد مقياس تقدير يستخدمه عند ملاحظة كل فرد والاستماع إليه وهو يلقي خطبة ما. وفي هذه الحالة يمكن أن يشتمل المقياس على الأبعاد التالية: محتوى الخطبة، وتنظيمها، وسهولة العرض، وصحة اللغة والنحو، والقدرة على التعبير، واسد تخدام الإيماءات والإشارات، ويعطى لكل بعد من هذه الأبعاد تقديرات منفصلة.

ويتبين من هذا أن مقياس التقدير يتكون من مجموعة من الخصائص أو الصفات للحكم عليها، ومقياس مدرج لتحديد درجة تواجدها لصفة أو الخاصية. وعلى هذا الأساس فإن الاستمارة المستخدمة هي مجرد أداة لتسجيل الملاحظات، وتتوقف قيمتها في جمع البيانات على العناية في بنائها والدقة التي تنفذ بها. وكما هو الحال في أدوات جمع البيانات الأخرى يجب أن يتم بناء الأداء في ضوء أهداف البحث الذي تستخدم فيه الأداة. ويجب أن يقتصر استخدامها على تلك المجالات التي تتوفر فيها فرص القيام بالملاحظات الضرورية.

وإذا طبق هذان المبدآن تطبيقا سليما فإن مقياس التقدير يساعد على تحقيق وظائف مهمة مثل:

- ♦ توجيه ملاحظات الباحث نحو مظاهر سلوكية خاصة.
- ♦ توفير إطار مرجعي يمكن من مقارنة أداء أفراد العينة على نفس الخصائص.
- ♦ توفير أداة مناسبة لتسجيل ملاحظات الباحث.

ويختلف أسلوب التقدير من مقياس لآخر فقد يستخدم ثلاث درجات: 'جيد، متوسط، ضعيف' أو خمس درجات: 'ممتاز، جيد جدا، جيد، متوسط، ضعيف'. وقد يستخدم سبع درجات، وقد يستخدم الأرقام بدلا من الألفاظ بحيث يدل الرقم الأعلى على قوة الصفة، والرقم الأدنى على ضعف الصفة. ويلاحظ أن زيادة عدد التقديرات يزيد من حساسية المقياس المدرج، فسبع درجات أكثر حساسية من خمس

درجات، وهذه أكثر حساسية من ثلاث درجات. إلا أن زيادة حساسية المقياس تزيد من صعوبة التمييز بين الدرجات المتتالية، ولذلك يجب على الباحث أن يوازن بين العاملين قبل اختيار عدد الدرجات التي يضمنها مقياسه.

وتفيد مقاييس التقدير في قياس كثير من المتغيرات وبخاصة المتغيرات المركبة متعددة الأبعاد، مثال ذلك التربية العملية، والقيادة، والتمثيل، وفي مادة السيارات، وعزف الآلات الموسيقية، والقيام بالتجارب في المعمل. ويمكن استخدامها كذلك في قياس التوافق الشخصي والاجتماعي لأفراد العينة.

عيوب مقاييس التقدير:

تتلخص عيوب مقاييس التقدير فيما يلي:

١- من أهم عيوب مقاييس التقدير غموض معاني السمات التي تقدر إلى درجة تجعل المقدر يعطي معنى مخالفاً للسمة. مثال ذلك صفة 'العدوان' قد تعني معاني مختلفة بالنسبة للمقدين المختلفين. فقد يعتبر بعض الملاحظين 'العدوان' مرادفاً 'للعداوة' في حين أن بعض الملاحظين الآخرين قد يعتبرونه سمة إيجابية في الفرد.

٢- غموض التقديرات نفسها، فقد تعني 'جيد جداً' أو درجة '٥' عند البعض شيئاً مختلفاً عما تعنيه عند البعض الآخر، ولذلك يجب تحديد مدلول نقاط التقدير، بشكل واضح موضوعي.

٣- من عيوب مقاييس التقدير أيضاً تأثرها ببعض العوامل مثل:

أ- **تحيز المقدر:** مما يدفعه إلى تقدير السمات المختلفة أعلى أو أقل مما تستحق لوجود عوامل خارجية مؤثرة على أفكاره.

ب- **أثر الحالة:** وهو ما يترتب على انطباعات الملاحظ عن الفرد الذي يلاحظه بشكل قد يؤثر على تقدير كل سمة من السمات.

ج- **النزعة المركزية:** وهي ميل المقدر إلى إعطاء تقديرات غير متطرفة وإعطاء تقدير متوسط في معظم الحالات.

٤- قد تتأثر مقاييس التقدير أيضاً بما يسمى **'الخطأ المنطقي'**، والخطأ المنطقي لا يرجع إلى التحيز الشخصي أو أثر الحالة، ولكنه يرجع إلى الاعتقاد بارتباط السمات بعضها ببعض مثل الذكاء والتحصيل. فإذا قام المقدر بتقدير صفة على أساس اعتقاده بوجود ارتباط بين هذه الصفة وصفة أخرى، ويعطيها نفس تقدير الصفة الثانية، يكون قد ارتكب خطأ في التقدير يرجع إلى الخطأ

أنواع مقاييس التقدير:

هناك أنواع متعددة من مقاييس التقدير، إلا أنه يمكن تصنيفها في واحد من الأنواع الثلاثة التالية:

مقاييس التقدير الرقمية: من أبسط مقاييس التقدير ذلك النوع الذي يقوم فيه المقدر بوضع علامة (√) أو دائرة حول الرقم الذي يبين درجة وجود الـ صفة. ويمكن إعطاء كل رقم وصفا لفظيا يظل ثابتا من صفة لأخرى. وفي بعض الحالات تحدد التعليمات أن أعلى رقم يمثل أعلى تقدير، والرقم (١) يمثل أدنى تقدير، وأن الأرقام الوسيطة تمثل قيما متوسطة.

والمقاييس الرقمية مفيدة إذا أمكن تصنيف الصفات أو الخصائص في عدد قليل من التقديرات (الفئات) التي يمكن الاتفاق على معنى كل منها. إلا أن الغموض يشوب في معظم الأحيان تحديد معنى هذه الأرقام، وبالتالي يختلف تفسيرها من فرد لآخر. ويوضح الشكل (١٨-٤) مثلا على مقاييس التقدير الرقمية.

تعليمات: بين الدرجة التي يسهم بها الطالب في المناقشة في الفصل وذلك بوضع دائرة حول الرقم المناسب، وذلك على النحو التالي:

٥- ممتاز ٤- فوق المتوسط ٣- متوسط ٢- تحت المتوسط ١- غير مرض

١- ما مدى إسهام الطالب في المناقشة في الفصل؟

١ ٢ ٣ ٤ ٥

٢- ما مدى ارتباط تعليقاته بموضوع المناقشة؟

١ ٢ ٣ ٤ ٥

شكل ١٨-٤ مثال لمقاييس التقدير الرقمية

مقاييس التقدير البيانية: يميز مقاييس التقدير البيانية وجود خط أفقي بعد كل صفة. وتوضع علامة (√) على الخط. ولكل فئة موقع محدد على الخط، إلا أن المقدر حر في وضع علاماته بين فئتين لتدل على درجة وسيدة لـ صفة بين الفئتين. ويوضح الشكل (١٨-٥) مثلا على هذا النوع من المقاييس. ويلاحظ أن هذا المقياس يستخدم نفس فئات التقدير لكل صفة ويطلق على هذا النوع "المقياس ذو الب دائل

الثابتة' وإذا تغيرت البدائل من صفة لأخرى يظل عليه 'المقياس' نوباً دائل المتغيرة'.

ورغم أن الخط البياني المستخدم في هذا النوع من مقاييس التقدير يمكن من وضع العلامات عند نقط وسيطة بين الفئات، إلا أن هذا النوع ليس له ميزة كبيرة على المقياس الرقمي، فنادرًا ما يتفق المحكمون على معنى 'نادرًا' و 'أحيانًا' و 'كثيرًا'. ولذلك يحتاج الأمر إلى أوصاف للسلوك تحدد بدرجة أكبر نوع السلوك الذي يظهره التلاميذ في كل فئة من الفئات.

تعليمات: بين الدرجة التي يسهم بها الطالب في المناقشة في الفصل وذلك بوضع علامة (√) في المكان المناسب على الخط عقب كل سؤال.

١- ما مدى إسهام الطالب في المناقشة في الفصل؟

أبداً نادراً أحياناً كثيراً دائماً

٢- ما مدى ارتباط تعليقاته بموضوع المناقشة؟

أبداً نادراً أحياناً كثيراً دائماً

شكل ١٨-٥ مثال لمقاييس التقدير البيانية

مقاييس التقدير البيانية الوصفية: يستخدم هذا المقياس عبارات وصفية لتحديد النقاط على الخط البياني. وتحدد النقاط كيف يسلك التلاميذ عند كل نقطة من نقاط الخط. وفي بعض المقاييس تحدد نقاط البداية والوسط والنهاية فقط، وكثيرًا ما يضاف حيز يذكر فيه المقدر أية تعليقات يراها لتوضيح التقدير.

ومقاييس التقدير البيانية الوصفية أكثر مقاييس التقدير مناسبة للاستخدام في البحوث، حيث إنها تصف للمقدر نوع السلوك الذي يعبر عن كل فئة من فئات التقدير، مما يساعد على تحقيق موضوعية الملاحظات، لأن أوصاف السلوك المحددة تساعد على تحقيق الدقة والموضوعية أثناء عملية التقدير، ويمكن إضافة قيم رقمية إلى الصفات عند كل موقع للمساعدة على تحويل الصفات إلى درجات. ويوضح الشكل (١٨-٦) نموذجاً لمقاييس التقدير البيانية الوصفية.

قواعد إعداد مقاييس التقدير:

١- يجب تحديد مجال 'السلوك' الذي يشمل المقياس، سواء كان هذا المجال

- سمات وجدانية أو مهارات حركية، كما يجب التأكد من تعريفها تعريفًا إجرائيًا، لا يختلف اثنان على تفسيره.
- ٢- تحديد نقاط التقدير أو مراتبه بحيث يكون مفهوما المقصود بكل نقطة من النقاط.
- ٣- إعطاء عينة ممثلة من السمات لمجال السلوك بحيث يمكن الخروج بتعميمات سليمة حول السلوك.
- ٤- من الأفضل التعبير عن السمات التي تقدر بأسئلة بدلا من العبارات الخبرية.

تعليمات: بين الدرجة التي يسهم بها الطالب في المناقشة في الفصل، وذلك بوضع علامة (√) في أي مكان على الخط عقب كل سؤال. وفي المكان المخصص للتعليقات أضف أي تعليق تراه يمكن أن يساعد في توضيح سلوك الطالب.

١- ما مدى إسهام الطالب في المناقشة في الفصل؟

يشارك أكثر	يشارك بنفس	هادئ، سلبي
من أي طالب	الدرجة التي	لا يشارك
آخر في الفصل	يشارك بها زملاؤه	أبدا

تعليقات:

٣- ما مدى ارتباط تعليقاته بموضوع المناقشة؟

التعليقات مرتبطة	التعليقات لها علاقة	التعليقات مشوشة
دائما بالموضوع	بالموضوع ولكنه	ولا علاقة لها
	يخرج أحيانا عنه	بالموضوع

تعليقات:

شكل ١٨-٦ مثال على مقاييس التقدير البيانية الرقمية

- ٥- تحديد عدد النقاط التي يشملها المقياس المدرج، فالمقياس المكون من سبع نقاط أكثر حساسية من المقياس المكون من خمس نقاط، ولكن الأول يتطلب تمييزا دقيقا بين مراتب التقدير أكثر من الثاني.

قواعد تحسين التقدير:

- ١- تدريب الملاحظين على طريقة استخدام المقياس وكيفية إعطاء التقديرات، مع إبراز عوامل الخطأ التي تؤثر في التقديرات حتى يتجنبها الملاحظون.
- ٢- من الأفضل أخذ مجموع أو متوسط عدة تقديرات مستقلة لنفس المقياس لزيادة ثبات التقديرات، لأن تعدد المقدرين يلغي الأخطاء الناجمة عن كل مقدر.
- ٣- للإقلال من أثر الهالة يفضل تغيير اتجاه المقياس من مفردة لأخرى، بحيث تبدأ المفردة الأولى بالطرف الموجب للمقياس، وتبدأ الثانية بالطرف السالب، وهكذا.

استخدامات مقاييس التقدير:

يمكن استخدام مقاييس التقدير في ثلاثة مجالات من مجالات القياس:

- ◆ الإجراءات
- ◆ الإنتاج
- ◆ النمو الشخصي والاجتماعي

قياس الإجراءات:

قد يحتاج الباحث إلى قياس الأداء الذي يظهره أفراد العينة، وبخاصة عقب المعالجة التجريبية. مثال ذلك إلقاء خطبة، واستخدام أدوات المعمل، والعمل بفاعلية في الجماعة، وعزف آلة موسيقية، والقيام ببعض الحركات الرياضية. ولا تنتهي هذه الأنشطة عادة بمنتج معين يمكن قياسه، كما أن اختبارات القلم والورقة لا تصلح لقياس الأداء المطلوب. ولذلك لابد من ملاحظة الأداء نفسه والحكم عليه.

وترجع أهمية مقاييس التقدير في قياس الإجراءات إلى أنها تركز على نفس مظاهر الأداء بين جميع أفراد العينة، كما أن لها مقياساً واحداً تحدد عليه التقديرات. وإذا أمكن إعداد استمارة التقدير في ضوء التعريف الإجرائي للمتغير التابع أو المتغيرات التي يرغب الباحث في قياسها، تصبح وسيلة ممتازة لقياس المتغيرات المرغوبة، لأنها تبين الأبعاد وصفات السلوك المستخدمة ونوع الأداء المرغوب.

قياس الإنتاج:

إذا ترتب على أداء الطالب إنتاج معين فقد يكون من الأفضل قياس الإنتاج بدلاً من الأداء. مثل القدرة على إعداد خطة معينة، وهنا نجد من الأفضل الحكم على الخطة نفسها بدلاً من طريقة إعدادها. حيث إنه لا فائدة من ملاحظة الطالب أثناء إعداد الخطة. وفي بعض المجالات مثل الطباعة على الآلة الكاتبة، والطبخ، والنجارة، قد يكون من الأفضل تقدير الإجراءات في مرحلة مبكرة من المعالجة، والإنتاج بعد انتهاء المعالجة. ومن الأمثلة على قياس الإنتاج الرسم ورسم الخرائط

والرسوم البيانية والبحوث والتقارير وغيرها من أنواع الإنتاج التي يمكن أن تكون جزءا من متغيرات البحث.

قياس النمو الشخصي والاجتماعي:

من أهم استخدامات مقاييس التقدير في البحوث تقدير مظاهر النمو الشخصي والاجتماعي. إذ يمكن للباحث تصميم استمارة لتقدير بعض الصفات الشخصية والاجتماعية مثل القيادة والتوافق الاجتماعي والتوافق الشخصي وغيرها، وذلك في البحوث التي تسعى إلى تقدير مثل هذه المتغيرات سواء كمتغيرات مستقلة أو متغيرات تابعة أو متغيرات ضابطة.

ويختلف تقدير الخصائص الشخصية والاجتماعية عن تقدير الإجراءات أو الإنتاج. ففي الحالتين الأخيرتين يتم التقدير أثناء الأداء أو بعده مباشرة. أما تقدير الخصائص الشخصية والاجتماعية فيحتاج إلى مراحل فترية ممتدة زمنيا، قد تكون فصلا دراسيا أو عاما دراسيا كاملا. ونظرا لاعتماد التقديرات على ملاحظات المقدر التي قد تكون عارضة ومنتشرة على فترة طويلة نسبيا من الزمن، فمن المتوقع أن تكون هذه التقديرات أكثر تعبيراً عن مشاعر المقدر وتشيعاته من تلك التقديرات التي نحصل عليها من الملاحظة المباشرة.

سادساً: قوائم المراجعة

تتكون قوائم المراجعة من عدد من الخطوات أو الأنشطة أو السلوك التي يستعين بها الملاحظ لتسجيل ملاحظاته. وتتشابه قوائم المراجعة مع مقاييس التقدير في مظهرها إلا أنهما يختلفان في طريقة إصدار الأحكام. ففي مقاييس التقدير نحدد درجة ظهور أو تكرار السلوك. أما في قوائم المراجعة فإننا نحكم على السلوك (أو الصفة) بأنه موجود أو غير موجود. ولذلك فإن قوائم المراجعة وسيلة لتسجيل حدوث أو عدم حدوث عمل ما، أو وجود أو غياب خاصية من الخصائص. فإذا كنا نريد تحديد درجة السلوك فإننا لا نستخدم قوائم المراجعة، بل نستخدم مقاييس التقدير. وتتميز قوائم المراجعة بأنها تمكن الملاحظ من تحديد السلوك بسرعة.

ولقوائم المراجعة فائدتها على وجه الخصوص عندما يعتمد القياس على أسلوب الملاحظة وليس على الاختبار. وإذا أمكن تعريف المتغيرات التي نريد قياسها تعريفا إجرائيا في سلوك ظاهر محدد، فإنه يمكن استخدام قوائم المراجعة وذلك ببيان أي سلوك موجود وأيها غير موجود. ويحدد التعريف الإجرائي مظاهر

السلوك المراد قياسها.

ويمكن استخدام قوائم الملاحظة في قياس العديد من أنواع السلوك، إلا أن فائدتها الكبرى تكمن في قياس العمليات التي يمكن تقسيمها إلى مجموعة من الخطوات أو الأعمال المحددة المنفصلة، مثال ذلك نموذج قائمة المراجعة الموضح في الشكل (٧-١٨).

وإذا أمكن بناء قوائم المراجعة بعناية فإنها تساعد على:

- ١- حصر انتباه الملاحظ في الخصائص الموجودة بالقائمة.
- ٢- مقارنة أفراد العينة بعضهم ببعض بالنسبة لمجموعة من السمات أو الخصائص.
- ٣- توفير أداة بسيطة لتسجيل الملاحظات.

ويجب عند استخدام قوائم المراجعة مراعاة ما يلي:

- ١- تحديد السلوك المراد قياسه بوضوح.
- ٢- ملاحظة فرد واحد في نفس الوقت.
- ٣- تركيز الانتباه على النقاط المحددة بالقائمة.
- ٤- تخصيص قائمة لكل فرد، وعند الرغبة في الحصول على صورة عامة للمجموعة يمكن تفريغ ملاحظات الأفراد في قائمة تشمل المجموعة كلها.

ضع علامة (+) أمام السلوك الذي يظهره الفرد، و (-) أمام السلوك الذي لم يحدث. توضع العلامة داخل المربع أمام العبارة.	
١- يتعرف على الأعداد من ١ إلى ١٠.	☐
٢- يعد حتى ١٠.	☐
٣- يتعرف على الأشكال الهندسية.	☐
٤- يتعرف على أجزاء العملة.	☐
٥- يرتب الأشياء حسب قيمها الرقمية.	☐
٦- ينقل الأعداد من ٠ إلى ١٠.	☐
٧- يحدد الوقت إلى أقرب نصف ساعة.	☐
٨- يتعرف على ربع الدائرة.	☐

شكل ٧-١٨ قائمة مراجعة لقياس بعض مهارات الرياضيات

وإذا أراد الباحث تفريغ بيانات كل فرد في كشف واحد، فإنه يقيم وم به ذه العملية بعد الانتهاء من عملية الملاحظة. ويوضح الشكل (٧-١٨) كشف تفريغ لقائمة المراجعة المبينة في الشكل (٧-١٨).

خطوات بناء قائمة المراجعة:

تلخص الأسس التالية خطوات بناء قائمة المراجعة لقياس صدق الصفات.

- ١- تحديد الأعمال الخاصة المراد أن يشملها الأداء.
- ٢- إضافة الأعمال التي تمثل الأخطاء الشائعة إلى القائمة (إذا كان لها أهميتها في قياس المتغير، وقليلة العدد، ويمكن تحديدها بوضوح).
- ٣- ترتيب الأعمال المطلوبة (والأخطاء إذا استخدمت) في التتابع المتوقع أن تحدث فيه.
- ٤- تحديد طريقة بسيطة لبيان العمل كما يحدث، أو ترتيب الأعمال بالتتابع إذا كان ذلك مناسباً.

عدد	رقم العبارة	الاسم	١	٢	٣	٤	٥	٦	٧	٨
١										
٢										
٣										
٤										
٥										
٦										
٧										
٨										
٩										
١٠										
١١										
١٢										
١٣										
١٤										
١٥										

شكل ١٨-٨ كشف تفريغ لبيانات قائمة المراجعة في شكل ١٨-٧

وبالإضافة إلى استخدام قوائم المراجعة في قياس إجراءات العمل، يمكن استخدامها أيضاً لقياس الإنتاج. ولتحقيق ذلك تحتوي استمارة القياس على قائمة بالخصائص المفروضة أن تتوفر في المنتج. ولقياس الإنتاج يحدد الباحث إذا كانت كل خاصية موجودة أم لا. وقبل استخدام قوائم المراجعة لقياس الإنتاج، يجب أن يقرر الباحث هل يمكن قياس المنتج باستخدام قوائم المراجعة (وجد أو غيب) أم لا.

الصفة). أما إذا كانت الصفة مما يفضل قياسه ببيان درجة وجود الصفة، يكون من الأفضل استخدام مقاييس التقدير بدلا من قوائم المراجعة.

وفي مجال النمو الشخصي والاجتماعي يمكن استخدام قوائم المراجعة لقياس السلوك المطلوب. وفي هذه الحالة تحتوي قائمة المراجعة على أنواع السلوك التي تم تحديدها كعينة ممثلة للسلوك الذي نسعى إلى قياسه. وفي مجال أدوات العمل على سبيل المثال يمكن للباحث أن يُصنّف قائمة المراجعة أنوع السلوك التالية:

ينفذ التعليمات.	_____
يطلب العون عندما يحتاجه.	_____
يعمل في تعاون مع الآخرين.	_____
يشارك الآخرين عند الضرورة في الأدوات التي يستخدمونها.	_____
يحاول القيام بأنشطة جديدة.	_____
ينهي العمل الذي بدأه.	_____

سابعا: الأساليب السوسيومترية

تستخدم الأساليب السوسيومترية لدراسة التفاعل بين الأفراد داخل الجماعات الاجتماعية. والإجراء الأساسي في القياس السوسيومتري يتضمن سؤال أعضاء الجماعة لبيان اختياراتهم الأولى والثانية وكذلك الاختيارات التالية حول ما يرغبون أن يكونوا زملاءهم في نشاط من الأنشطة. مثال ذلك قد يُطلب من مجموعة من الأطفال تحديد اثنين أو أكثر من الأطفال في مجموعتهم للعمل معهم في نشاط من الأنشطة، أو للعب معهم.

والأسلوب السوسيومتري طريقة لقياس اختيارات كل فرد من أفراد مجموعة معينة، ويمكن رسم الاختيارات التي يقوم بها أعضاء الجماعة ويطلق على هذا الرسم 'السوسيوگرام' الذي يبين نمط العلاقات المتداخلة بين أعضاء الجماعة. والأسلوب السوسيومتري يستخدم على نطاق واسع في البحوث في علم النفس الاجتماعي، وفي التربية، حيث يمكن دراسة التركيز الاجتماعي للعينة وعلاقته بمجموعة أخرى من المتغيرات مثل القدرة العقلية العامة، والتدصيل الدراسي، ومفهوم الذات، والمعلمين الذين يفضلهم الأطفال (Ary et al., 1996). وفيما يلي مثال على استمارة سوسيومترية، والغرض من هذه الاستمارة اختيار التلاميذ لمن يشاركونهم مقاعد في الفصل، وفي النشاط المدرسي، وفي جماعات اللعب بالمدرسة.

جدول ١٨-٣ استمارة سوسيومترية

الاسم: _____ التاريخ: ---/---/----							
سوف تعدل في الأسابيع القادمة أماكن الجلوس على المقاعد في الفصل، وتكوين جماعات النشاط الحر وجماعات النشاط الرياضي. وبما أنه قد مضى وقت كاف من السنة لكي يعرف التلاميذ بعضهم البعض يمكن مساعدة المدرسة في تنظيم جماعات النشاط المختلفة. ويمكنك ذلك بكتابة أسماء التلاميذ الذين ترغب أن يكونوا معك في هذه الجماعات. ويمكنك اختيار أي تلميذ من هذا الفصل حتى ولو كان غائبا. ولن يطلع أحد على الاختيارات التي تقوم بها، بما في ذلك التلاميذ الذين اخترتهم. دقق في اختيارائك حتى يكون زملاؤك في أي جماعة هم من ترغب في أن يكونوا معك. وسوف نحاول تنظيم الجماعات بحيث يكون مع كل تلميذ في جماعته اثنين على الأقل ممن اختارهم. ومن الصعب أحيانا تحقيق الرغبات الأولى ولذلك من الأفضل أن تكتب خمسة اختيارات.							
تذكر ١- أن تكون جميع اختيارائك من تلاميذ في هذا الفصل حتى ولو كانوا متغيبين. ٢- أن تكتب الاختيارات الخمسة كلها بعد كل سؤال. ٣- يمكنك اختيار نفس التلميذ في أكثر من جماعة إذا رغبت. ٤- اختيارائك سرية ولن يطلع عليها أحد.							
أرغب أن أجلس بالقرب من <table border="0"> <tr> <td>١- _____</td> <td>٢- _____</td> </tr> <tr> <td>٣- _____</td> <td>٤- _____</td> </tr> <tr> <td>٥- _____</td> <td></td> </tr> </table>		١- _____	٢- _____	٣- _____	٤- _____	٥- _____	
١- _____	٢- _____						
٣- _____	٤- _____						
٥- _____							
أرغب أن أشارك في النشاط مع <table border="0"> <tr> <td>١- _____</td> <td>٢- _____</td> </tr> <tr> <td>٣- _____</td> <td>٤- _____</td> </tr> <tr> <td>٥- _____</td> <td></td> </tr> </table>		١- _____	٢- _____	٣- _____	٤- _____	٥- _____	
١- _____	٢- _____						
٣- _____	٤- _____						
٥- _____							
أرغب أن أشارك في اللعب مع <table border="0"> <tr> <td>١- _____</td> <td>٢- _____</td> </tr> <tr> <td>٣- _____</td> <td>٤- _____</td> </tr> <tr> <td>٥- _____</td> <td></td> </tr> </table>		١- _____	٢- _____	٣- _____	٤- _____	٥- _____	
١- _____	٢- _____						
٣- _____	٤- _____						
٥- _____							

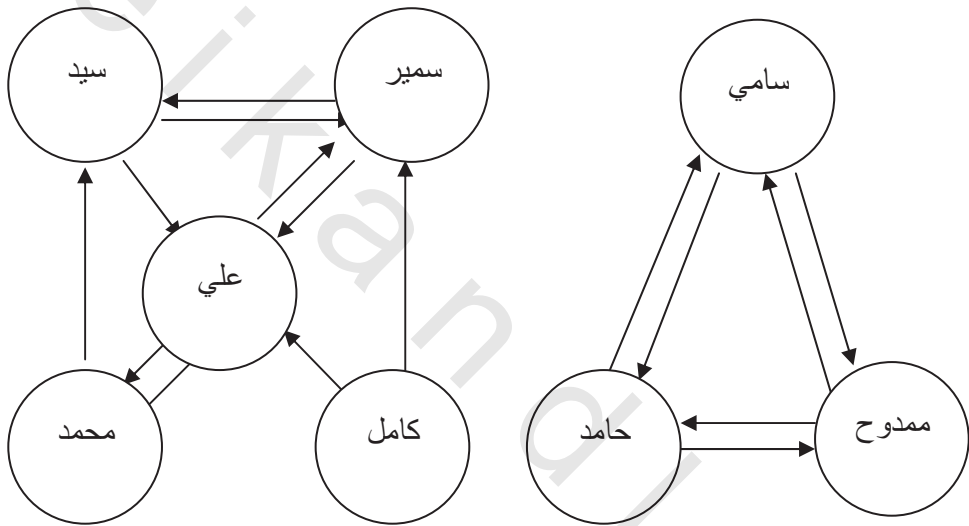
- وتشرح التعليمات عدة مبادئ مهمة في الاختيار السوسيومتري:
- ١- يجب أن تكون الاختيارات حقيقية وجزءا طبيعيا من أنشطة الفصل.
- ٢- يجب توضيح أساس الاختيار والقيود التي توضع عليه.

- ٣- حرية جميع تلاميذ الفصل أحراراً في الاشتراك في الأنشطة المطلوبة.
- ٤- يجب الاحتفاظ بسرية جميع الاختيارات التي يقوم بها التلاميذ.
- ٥- يجب استخدام هذه الاختيارات لتنظيم أو إعادة ترتيب الجماعة المدرسية. فالتلميذ سوف يقبل على مثل هذه الأساليب في المستقبل إذا علم أن رأيه يحترم.

ومعروف أن الأنشطة المدرسية تزخر بالمواقف التي يمكن فيها استخدام الأسلوب السوسيومترى. فيمكن للتلاميذ اختيار زملائهم في المعامل وفي جماعات النشاط، وفي المشروعات المختلفة التي يطالب بها التلاميذ وغيرها من المواقف والأنشطة الصفية واللاصفية. ورغم أنه من المتوقع أن تحدث اختلافات من موقف لآخر إلا أن عنصراً كبيراً من التقبل الاجتماعي يتخلل جميع الاختيارات السوسيومترية. فالتلميذ الذي يُختار في نشاط ما يقع عليه الاختيار أيضاً في الأنشطة الأخرى. وتحدث الاختلافات الكبيرة في عملية الاختيار عندما يكون الغرض من الاختيار الاشتراك في جماعات خاصة يكون أساس الأداء الناجح فيها المهارة والقدرة. فحتى التلميذ غير المحبوبين يختارون زملاء في فريق كرة القدم مثلاً عندما يكون من المعروف عنهم أنهم لاعبون ممتازون. إلا أنه من غير المحتمل اختيارهم كزملاء في الأنشطة الصفية مثلاً، لأن هذا النوع من الاختيار يعتمد كلية على التقبل الاجتماعي.

ويختلف الخبراء حول جدوى أو مرغوبة مطالبة التلاميذ تسمية أولئك الذي لا يودون أن يكونوا زملاءهم. ويذكر أولئك الذين ينادون بهذا الرأي أن مثل هذه الاختيارات السلبية تمكن من التعرف على التلاميذ المنبوذين حتى يمكن مساعدتهم، كما أن مثل هذه الاختيارات تمكن من تجنب أية احتكاكات قد تحدث عند تنظيم جماعة من الجماعات. أما الرأي المعارض للرأي السابق فيذكر أن مثل هذه النوع من الاختيارات يجعل التلميذ أكثر إحساساً بمشاعر النبذ وقد يؤثر هذا على أخلاقيات الجماعة وعلى النمو الانفعالي للتلميذ. ويبدو أن أسلم إجراء هو تجنب الاختيارات السلبية إلا إذا كانت ضرورية تماماً للغرض الذي يستخدم من أجله الأسلوب السوسيومترى. وإذا كان هذا الاستخدام ضرورياً يجب أن تبذل طريقة السؤال عرضية، ويسمح للتلميذ أن يذكر اختياراته السلبية، ولكن لا يكون ملزماً بإعطاء مثل هذا الاختيار. ويكفي عادة ذكر عبارة كالعبارة التالية ليتحقق ذلك: 'إذا كان هناك تلميذ لا ترغب في وجودهم معك في الجماعة، يمكنك ذكر اسمهم'.

ومن الأفضل تحديد عدد الاختيارات التي يقوم بها التلاميذ في كل سدؤال
سوسيومتري. وفي معظم الحالات يكون من الكافي عمل خمسة اختيارات. فقد
دلت الدراسات التي تمت على نتائج القياس السوسيومتري أن ثباته يزداد حتى خمس
اختيارات، وأنه لا يزيد بعد ذلك. كما أن خمسة اختيارات تجعل من السهل تنظيم
الجماعات لأنه من الصعب أحيانا تحقيق الاختيارات الأولى لجميع التلاميذ. وفي
الصفين الأول والثاني الابتدائي يفضل ألا تزيد الاختيارات عن اثنين أو ثلاثة. فمن
الصعب على التلاميذ الصغار التمييز بين الاختيارات التي تزيد عن هذالعدد
(Linn & Gronlund, 2000, p. 328).



شكل ١٨-٩ سوسيوجرام يبين مجموعتين من الاختيارات السوسيومترية

يلاحظ في هذا الشكل أن الاختيارات التي وقعت على 'علي' كانت
أكثر من غيرها، ولذلك يطلق عليه 'النجم'، ويلاحظ أن سامي
وممدوح وحامد اختار كل منهم الآخر. ولذلك يمكن اعتبارهم
مجموعة واحدة، ويطلق على مثل هذه المجموعة 'عشيرة'.
ويطلق هذا المصطلح عندما يشترك ثلاثة على الأقل في اختيارات
بعضهم البعض. ويلاحظ كذلك أن كامل لم يوجه إليه أي اختيارات،
ولذلك فهو يعتبر 'منبوذا' من الجماعة. وتكمم اختيارات
السوسيوجرام عادة، لتحليلها ضمن نتائج البحث.

مثال ذلك أن إبراهيم ع. قام بالاختيارات التالية:

اختار	نبت
١- حسن م.	⊗ السيد أ.
٢- حسين س.	⊗ حسن ب.
٣- أحمد س.	
٤- كامل ع.	
٥- راشد س.	

جدولة النتائج السوسيومترية:

يجب تنظيم الاختيارات السوسيومترية التي قام بها التلاميذ حتى يمكن تفسيرها واستخدام النتائج استخداما سليما. وتبين العلامات التكرارية البسيطة لعدد الاختيارات التي حصل عليها كل تلميذ درجة تقبله الاجتماعي. وقد يكون هذا كل ما يحتاجه الباحث لتحقيق الغرض الذي من أجله استخدم الأسلوب السوسيومتري، إلا أن ذلك لن يكشف عن التلاميذ الذين قاموا بالاختيار، وهل اختار تلميذان معينان بعضهما البعض، وما التركيب الاجتماعي للجماعة. ويمكن تحقيق ذلك إذا سجلت جميع النتائج السوسيومترية وتنظيمها في جدول كالمبين في الجدول (١٨-٤)، أو بعمل 'سوسيوگرام' كالموضح في شكل (١٨-٩). وتسجل أسماء التلاميذ في هذا الجدول وترقم من ١ إلى ١٥. وتكتب هذه الأرقام في أعلى الجدول بنفس تتابعها حتى يمكن تسجيل اختيار كل تلميذ في الأعمدة. وقد سجلت هذه الاختيارات في الجدول أمام اسم إبراهيم ع. وذلك بوضع رقم ١ في العمود الذي يشير إلى حسن م. ووضع رقم ٢ في العمود الذي يمثل حسين س.، وهكذا. ويمثل الرمز ⊗ الأفراد المنبوذين، فنجد أن هذا الرمز وضع أمام اسم إبراهيم ع. في العمود الذي يشير إلى السيد أ.، والعمود الذي يشير إلى حسن ب. والأرقام التي تحتها خط تمثل اختيارات مشتركة. مثال ذلك أن حسن ب. اختار السيد أ.، كما أن السيد أ. اختار حسن ب. ويلاحظ أن عدد الاختيارات المشتركة متساوي على جانبي القطر بين الصفوف والأعمدة.

وعند جمع عدد الاختيارات أعطي كل اختيار قيمة تساوي واحدا، بغض النظر عن مستوى الاختيار. ويفضل البعض إعطاء أوزان للاختيارات بحيث يعطى الاختيار الأول وزنا أكبر من الاختيار الثاني، وهكذا، إلا أن هذه الأوزان أوزان اعتبارية، ولا تعبر عن قيم حقيقية، ورغم أن الاختيار الأول له أهمية أكبر من

الاختيار الثاني، إلا أنه يحدث في بعض الأحيان أن يجد التلميذ صعوبة في التمييز بين اختياره الثالث والرابع، مما يجعل هذين الاختيارين في واقع أمرهما قريبين جدا من بعضهما البعض. ويفترض عند إعطاء أوزان تساوي المسافات بين الاختيارات المتتالية، ولكن هذا في واقع أمره ليس صحيحا. ونظرا لعدم وجود نظام وزن يمكن استخدامه بكفاءة يجب استخدام الإجراء الأبسط وهو حساب كل اختيار بواحد، مع تسجيل مستوى الاختيار في الجدول لأن له فائدته عند تنظيم الجماعات.

والعدد الذي يحصل عليه التلميذ من الاختيارات دليل على مدى تقبله الاجتماعي بين أقرانه. وعند استخدام خمسة اختيارات كما هو الحال في هذا الجدول، فإن التلاميذ الذين يحصلون على ٩ اختيارات أو أكثر يطلق عليهم 'نجوم' في حين أن الذين لا يحصلون على أي اختيار يطلق عليهم 'منعزلين'، أما الذين يحصلون على اختيار واحد يطلق عليهم 'مهملين'، أما الذين يقعون بين هذين الطرفين فلا يُعطون تسمية خاصة، والتلاميذ الذين يحصلون على نبذة فقط يطلق عليهم 'منبوذون'، والذين يتبادلون الاختيار يطلق عليهم الاختيارات المشتركة.

جدول ١٨-٤ تسجيل جميع النتائج السوسيومترية

عدد التنبؤ	اختيارات	رقم	الاسم	رقم	١	٢	٣	٤	٥	٦	٧	٨	٩	١٠	١١	١٢	١٣	١٤	١٥
٢	٥	١	ابراهيم ع.	١				⊗	⊗	١	٢			٥		٤			
١	٥	٢	احمد س.	٢		٣		١	⊗	٤		٣		٥					
١	٥	٣	احمد ص.	٣		٢			⊗	١	٢		٤	٥					
	٥	٤	السيد ا.	٤					١	٢		٣	٤			٥			
	٥	٥	حسن ب.	٥				١		٢	٣	٤						٥	
١	٥	٦	حسن م.	٦						٢	١				٣		⊗	٤	٥
١	٥	٧	حسين س.	٧		٢		٥		١			٣			⊗			
١	٥	٨	حسين ع.	٨	٤				⊗	١	٣			٧		٥			
	٥	٩	خالد ا.	٩		٢		٣	٥	٤	١								
	٥	١٠	رائد س.	١٠					٥	٢	١	٣						٤	
	٥	١١	رشد م.	١١					٥		٤					١	٣		٢
	٥	١٢	كامل ع.	١٢							٤				١		٥	٧	٢
١	٥	١٣	ممد ر.	١٣		٣		٢		٤	٥						⊗		
١	٥	١٤	محمد ض.	١٤					٥		١		٤			٢			
	٥	١٥	ممد ي.	١٥				٥				١		٣	٢		⊗		
عدد الاختيارات					١	٥	٥	٥	٤	١١	٩	٧	٥	٣	٥	٣	٣	٤	
عدد التنبؤ					٠	٠	١	٠	٤	٠	٠	٠	٠	٠	١	١	١	٠	
الاختيارات المشتركة					٠	١	٢	١	١	٣	١	٢	١	١	٣	٧	١	١	

- الأرقام التي تحتها خط تمثل اختيارات مشتركة

- رمز التنبؤ ⊗

oboiyadi.com

الفصل التاسع عشر

جمع البيانات باستخدام

الاستبيان والمقابلة

يستخدم الاستبيان والمقابلة بشكل مكثف في البحوث التربوية والنفسية لجمع البيانات عن الظواهر التي لا يمكن ملاحظتها مباشرة. الخبرات الذاتية للأفراد، والآراء، والقيم، والميول، والاتجاهات وغيرها. ويمكن استخدامها كذلك لجمع البيانات عن الظواهر التي يمكن ملاحظتها ولكن بسهولة أكبر من استخدام الملاحظة المباشرة. فمن الأفضل مثلاً أن يستخدم الباحث الاستبيان أو المقابلة لتوجيه عدة أسئلة لناظر المدرسة عن عدد المدرسين في مدرسته، وكم منهم لديه جهاز حاسب آلي، بدلاً من التجول في المدرسة بنفسه وعمل عملية حصر لما يريد الحصول عليه من معلومات. وبالطبع تصحيح هذه الاستفادة عديمة الجدوى إذا كانت البيانات التي نحصل عليها بهذه الطريقة ذات صدق ضعيف.

والاستبيانات عبارة عن وثائق توجه نفس الأسئلة إلى جميع الأفراد في العينة. (ويلاحظ أنه إذا كانت هناك مجموعات فرعية في العينة فقد توجد لها هذه المجموعات أسئلة مختلفة). ويسجل المستجيبون إجابات مكتوبة لكل مفردة من مفردات الاستبيان. ولذلك نجد أن المستجيبين عملياً يتحكمون في جمع البيانات: لأنهم يملأون الاستبيان بالطريقة التي يرونها، وقد يستجيبون لبعض المفردات دون البعض الآخر، وبالترتيب الذي يرونه.

وتتكون المقابلات من أسئلة شفوية يسألها المقابل ويحصل على إجابات شفوية من المشاركين في المقابلة. وتتضمن المقابلة عادة مقابلات فردية، ولكن يمكن أن تحدث مقابلات جماعية. ويتكلم المستجيب مستخدماً الكلمات التي يراها، ويسجل المقابل استجاباتهم أثناء المقابلة إما حرفياً أو باستخدام التسجيل الصوتي أو

التسجيل المرئي. أو قد يكتب مذكرات قصيرة يفصلها بعد المقابلة. وتتميز المقابلة بأن المقابل هو الذي يتحكم في سير المقابلة وفي طريقة الاستجابة، وقد يتفق مع المشارك في موعد معين للمقابلة يناسب الاثنين، ولكن المقابل بعد ذلك هو الذي يتحكم في المقابلة وفي طريقة تسجيل الاستجابات.

الاختيار بين الاستبيان والمقابلة:

للاستبيان ميزتان على المقابلة كوسيلة لجمع البيانات. الأولى أن تكلفة المعاينة على رقعة جغرافية كبيرة من الأرض أقل، كما أن الزمن المستغرق في جمع البيانات من العينة أقل بكثير. إلا أن الاستبيان لا يستطيع التوغل بعمق في معتقدات المستجيب واتجاهاته، وخبراته الشخصية. كما أنه بمجرد توزيع الاستبيان فلا يمكن تعديل الأسئلة الموجودة به حتى إذا كانت غير واضحة للمستجيب.

والميزة الأساسية للمقابلات هي قابليتها للمواءمة مع موقف المقابل. إذ يستطيع المقابلون المهرة متابعة استجابات المشارك للحصول على معلومات أكثر وتوضيح العبارة الغامضة التي أدلى بها. ويمكنهم تكوين علاقة من الألفة والثقة بينهم وبين المشاركين، مما يتيح لهم الحصول على بيانات قد لا يكشف عنها المستجيب باستخدام أي وسيلة أخرى من طرق جمع البيانات.

وهذه الميزة التي تتصف بها المقابلة قد تنتفي نتيجة لبعض نواحي القصور. أولها أنه من الصعب تقنين موقف المقابلة حتى لا يؤثر المقابل في طريقة استجابة المشاركين للأسئلة ويجعلهم يستجيبون بطريقة معينة على الأسئلة. وثانيها أن المقابلة لا تستطيع أن توفر السرية للمستجيب. وبمعنى آخر لا بد من أن يكشف المستجيبون عن هوياتهم للمقابل. إلا أن المقابل يستطيع تحليل بيانات المقابلة ويكتب تقريراً عنها دون الكشف عن شخصية المستجيب.

واستخدام الاستبيان أكثر شيوعاً من استخدام المقابلة في جمع البيانات في البحوث الكمية، لأنه مقنن، وطبيعة بنائه المحددة تتفق مع هذا النوع من البحوث. أما المقابلة فأكثر استخداماً في البحوث الكيفية، لأنها تسمح بالاستكشافات ذات الطبيعة المفتوحة، كما أنها تسمح للمستجيبين بحرية غير محدودة في الإدلاء بما يريدون من استجابات. ولا يعني هذا أن استخدام الاستبيان قاصر على البحوث الكمية، وأن استخدام المقابلة قاصر على البحوث الكيفية، فمن الممكن استخدام أي من الطريقتين في البحوث الكمية أو البحوث الكيفية.

الاستبيان

الاستبيان أكثر الوسائل استخداما للحصول على معلومات وبيانات عن الأفراد، ويرجع ذلك لأسباب عديدة منها أن الاستبيان اقتصادي نسبيا، ويمكن إرساله إلى أشخاص في مناطق بعيدة، كما أن الأسئلة أو المفردات مقننة من ف رد لآخر، ويمكن ضمان سرية الاستجابات، كما أنه يمكن صياغة الأسئلة لتناسب أغراضا محددة. ويمكن أن تستخدم الاستبيانات الصيغة الاستفهامية، أو الصيغة الإخبارية، دون أن يؤثر ذلك على مضمون السؤال أو الفقرة، وفي أي من الحالتين فإن الأفراد يستجيبون كتابة لما هو مطلوب منهم.

يجب قبل البدء في بناء الاستبيان أن يحدد الباحث مبررات استخدام الاستبيان كوسيلة لجمع البيانات كما يحدد الأهداف من الاستبيان.

مبررات الاستبيان: الاستبيان أحد الوسائل العديدة للحصول على البيانات، والباحث الذي يريد استخدام الاستبيان يجب أن يكون متأكدا من أنه لا توجد وسيلة أخرى أكثر صدقا وثباتا يمكن استخدامها في بحثه. ويعتمد هذا القرار على معرفة بنواحي قوة وضعف كل وسيلة. ولابد من ذكر مبررات استخدام الاستبيان وبخاصة إذا كان استبيانا جديدا. ويمكن في كثير من الأحيان استخدام استبيانات موجودة فعلا، أو تطويرها لاستخدامها بدلا من بناء استبيان جديد. وإذا استطاع الباحث الحصول على استبيان سبق استخدامه، فإنه يوفر الجهد والمال اللذين ينفقهما في بناء استبيان جديد وتحقيق صدقه وثباته.

أهداف الاستبيان: بعد تحديد مبررات الاستبيان تأتي الخطوة التالية وهي إعداد قائمة بالأهداف الخاصة التي سوف تحققها البيانات التي نحصل عليها من الاستبيان. ويجب تحديد هذه الأهداف في ضوء أسئلة البحث ومشكلته، مع توضيح كيف تستخدم كل جزئية من البيانات. وليس من الضروري أن تكون الأهداف المصاغة أهدافا سلوكية، ولكن يجب أن توضح الأهداف كيف أن الاستجابات التي نحصل عليها من كل سؤال لها وظيفتها في البحث. ويؤدي تحديد الأهداف إلى تحديد نوع المعلومات التي يرغب الباحث في الحصول عليها. وإذا لم يستطع الباحث التقييم بهذه الخطوة فهذا يعني أن مشكلة البحث ليست واضحة. وكثير من الباحثين يضعون أسئلة لم يفكروا فيها تفكيراً جيداً ويترتب على عدم استخدامها، ضياع

الوقت والجهد اللذين بذلا في إعدادها.

كتابة الأسئلة أو العبارات: إذا تم تحديد الأهداف والتأكد من أنه لا توجد أداة جاهزة يمكن استخدامها، يبدأ الباحث في كتابة مفردات الاستبيان مع مراعاة جانبين هامين هما:

١- أن تكون الأسئلة مرتبطة بأهداف الاستبيان.

٢- أسلوب تحليل البيانات بعد جمعها، إذ يتوقف شكل الاستبيان وصورته على هذا الأمر. وسوف نتناول هذه النقطة بعد قليل عند مناقشة أنواع أسئلة الاستبيان.

ويذكر شوماكر وماكميلان (Schumacher & McMillan, 1984) مجموعة من الإرشادات لكتابة أسئلة الاستبيان، هي:

١- أن تكون العبارات واضحة ذات معنى مدد، بحيث يفهمها جميع المستجيبين بنفس الطريقة. ويجب استخدام لغة عربية سهلة وسليمة، وأن توجه إلى أقل المستويات التعليمية، حتى نضمن أن تفهمها جميع المستويات. ولا يجب الافتراض أبداً أن القارئ سوف يفهم ما هو مكتوب، وكثيراً ما تكون المفاهيم والعبارات والكلمات التي يستخدمها الباحث غير واضحة للمستجيبين. ولا يجب أن يكون السؤال عاماً أكثر من اللازم فتكون تفسيراته كثيرة ومختلفة، فإذا سألت سؤالاً مثل 'ما رأيك في المنهج الجديد؟' فإن ردود الفعل بالنسبة له تكون متعددة، فقد يسأل شخص مثلاً 'أي منهج يقصد؟'. كذلك يجب تجنب الكلمات الغامضة غير الواضحة مثل 'عادة' و'أحياناً'، و'قليلاً'.

٢- تجنب كتابة العبارات أو الأسئلة المزدوجة، إذ يجب ألا يحتوي السؤال إلا على فكرة واحدة فقط. فالأسئلة المزدوجة التي تحتوي على أكثر من فكرة قد تربك المستجيب وبخاصة إذا كان يريد الاستجابة لكل فكرة بطريقة مختلفة عن الأخرى. فإذا طلب من فرد أن يوافق أو لا يوافق على عبارة مثل 'ينفق المرشدون التربويون وقتاً كبيراً في إعداد السجلات ولا ينفقون وقتاً كافياً في إرشاد الطلبة في مشكلاتهم الخاصة'، فقد يوافق الفرد على الجزء الأول من السؤال، ولا يوافق على الجزء الثاني أو العكس.

٣- يجب أن تستثير الأسئلة إجابات غير غامضة، فإذا سألنا مثلاً طلبة الجامعة عن درجة استخدامهم للمكتبة الجامعية، فمن الأفضل أن نسأل عن عدد

مرات زيارة المكتبة أسبوعيا، إذا كان الاستبيان من النوع المفتوح، أما إذا كانت الاستجابة من النوع المقيد، يجب أن تكون الاستجابات كمية مثلاً: (يومية - ٥ مرات في الأسبوع - مرتان في الأسبوع - وهكذا).

٤- يجب أن تكون الأسئلة في مستوى المستجيبين حتى يمكن إعطاء معلومات صحيحة موثوق بها. فلو طلبنا من المدرسين مثلاً ذكر أحد دأث معينة ماضية، أو أن يذكروا ما فعلوه منذ أسابيع فإن ما يذكرونه من بيانات لن يكون موثقاً بها غالباً. كذلك من غير المناسب أن تسأل مدرس الرياضيات أن يدلي برأيه في اختبار اللغة العربية. وهذاك مواقف لا يستطيع المستجيب أن يكون رأياً محدداً نحوها ولذلك فمن الأفضل أن تكون إحدى بدائل الإجابات 'لا أدري' أو 'غير متأكد'، وبذلك يكون لديه الفرصة للتعبير بحرية عن اعتقاده.

٥- يجب أن تكون الأسئلة متعلقة بالموضوع ومهمة للمستجيب، فلو سألنا أسئلة تافهة وغير مهمة بالنسبة له، فلن يعطي المستجيب اهتماماً كافياً للاستبيان وبذلك تكون النتائج مضللة. فمثلاً لو سألنا المدرس عن رأيه في أي اختبارات الذكاء أفضل، فإنهم لن يهتموا بالإجابة، حيث إنهم غالباً لا يهتمون بمثل هذا الموضوع. وقد تكون استجاباتهم قائمة على فكرة عامة وليس على دراسة متعمقة للموضوع.

٦- الأسئلة البسيطة هي أفضل الأسئلة، لذلك يجب تجنب المقارنات الطويلة المعقدة لأن فهمها صعب. افترض دائماً أن المستجيبين يقرأون الأسئلة بسرعة ويستجيبون لها بسرعة، ولذلك يجب كتابة عبارات بسيطة مفهومة يسهل الإجابة عليها.

٧- يجب تجنب العبارات المنفية لأن من السهل إساءة تفسيرها، فكثيراً ما يغفل المستجيبون أداة النفي، ويفهمون العبارة عكس المقصود منها. وإذا اضطر الباحث إلى استخدام أداة النفي يجب أن يضع تحتها خطاً حتى يلتفت الانتباه إليها.

٨- يجب تجنب المفردات أو المصطلحات المتحيزة فإن صياغة السؤال بطريقة خاصة أو إضافة مصطلح ما إليه قد يوحي بإجابات معينة. فلو سألنا المدرسين في مدرسة معينة مثلاً 'هل توافق على الرأي الذي ذكره الناظر عن ٠٠٠' قد يستثير استجابات معينة متأثرة باتجاه المستجيب نحو الناظر المدرسة، وليست تعبيراً عن رأيه حول الموضوع. كذلك كثيراً ما يلمح

الباحث إلى نوع الإجابة التي يرغبها، مما يجعل المستجيب يعطي الاستجابة التي يريدها الباحث بغض النظر عن رأيه الفعلي.

الصورة العامة للاستبيان:

يعطي تنظيم الاستبيان وإعداده في صورة مقبولة مريحة للعين انطبعا جيا دا عن الاستبيان، وتجعل المستجيب يقبل عليه ويعطي إجابات صادقة موثوقة ل به ل. ولتحقيق ذلك يجب الالتزام بالقواعد التالية التزاما تاما.

- ١- التأكد من خلو الاستبيان من الأخطاء اللغوية.
- ٢- يجب أن تكون الطباعة واضحة سهلة القراءة.
- ٣- يجب عدم استخدام مصطلحات فنية معقدة لا يستطيع المستجيبون فهمها.
- ٤- تجنب استخدام كلمة "استبيان" أو "مقياس تقدير" كجزء من عنوان استمارة الاستبيان، لأن كثيرا من الناس يتحيزون ضد مثل هذه الألفاظ.
- ٥- يجب كتابة تعليمات واضحة ومطبوعة ببنط واضح ت سهل قراءته. وأن تكون التعليمات مختصرة ومفهومة. ويفضل إعطاء أمثلة على كيفية الاستجابة للاستبيان وبخاصة إذا كانت طريقة الاس تجابة غير واضحة للمستجيب.
- ٦- لتكن الأسئلة مختصرة على قدر الأمكان لأن الاختصار يزيد من الوضوح.
- ٧- تجنب استخدام بعض الألفاظ والعبارات غير محددة المعنى مثل: عدية، ومعظم، وعادة، وما شابه ذلك من الألفاظ التي لا تحمل معنى محددا دقيقا.
- ٨- تجنب ازدحام صفحات الاستبيان بالأسئلة وذلك عن طريق ترك مساحات بيضاء كافية على جانبي الصفحة، مع طباعة الأسئلة على مسافات كافية بين كل سطر والذي يليه.
- ٩- يجب أن يكون عدد الأسئلة في الاستبيان معقولا فلا يكون الاستبيان أطول من اللازم، حتى لا يمل المستجيب وبالتالي يعطي إجابات غير صحيحة. وبمعنى آخر يجب أن يكون الاستبيان قصيرا على قدر الإمكان.
- ١٠- عند الانتقال إلى موضوع جديد داخل الاستبيان يفضل وضع عبارة انتقالية تمهد للمستجيب للموضوع الجديد.
- ١١- يفضل أن يبدأ الاستبيان بأسئلة سهلة جذابة ومهمة للمستجيب حتى لا يشعر بتهديد من الاستبيان.

- ١٢- يفضل وضع الأسئلة التي قد تشعر المستجيب بنوع من التهديد في نهاية الاستبيان، وكذلك الأسئلة الصعبة.
- ١٣- تجنب الأسئلة المنفية على قدر الأمكان إذ أن المحتمل إساءة فهم المقصود منها، وإعطاء إجابة عكس المقصودة من السؤال.
- ١٤- يجب تجنب الأسئلة المزدوجة التي تتطلب من المستجيب أن يعطي إجابة منفصلة لكل جزء من السؤال.
- ١٥- إذا كان هناك سؤال عام وسؤال خاص يتناولان موضوعا واحدا يجب البدء بالسؤال العام، لأن البدء بالسؤال الخاص في مثل هذه الحالة قد يؤدي من تفكير المستجيب في الأجوبة على السؤال العام.
- ١٦- إذا كان بالاستبيان أسئلة مفتوحة يجب ترك مساحات كافية للكتابة فيها.
- ١٧- يجب تنظيم أسئلة الاستبيان بحيث يكون تتابع الأسئلة تتابعاً منطقياً، وأن تجمع الأسئلة التي تتناول موضوعاً واحداً مع بعضها البعض.
- ١٨- يجب ترقيم الأسئلة والصفحات.
- ١٩- إذا كان الباحث يعتقد أن بعض الأسئلة يصعب فهمها يجب إضافة أمثلة عنها حتى يتضح المقصود منها للمستجيب.
- ٢٠- إذا اضطر الباحث إلى بناء استبيان طويل يجب وضع الأسئلة المهمة في بداية الاستبيان.
- ٢١- يؤثر موضع بعض الأسئلة وتتابعها على الاستجابات، ولذلك يجب أن يراعي واضع الاستبيان ألا تؤثر الإجابة على سؤال في الإجابة على سؤال آخر.
- ٢٢- يجب أن تكون استمارة الاستبيان جذابة على قدر الإمكان بحيث يقبل عليها المستجيب، وتزداد دافعيته للاستجابة.

أنواع الأسئلة:

يمكن تصنيف أسئلة الاستبيان في نوعين رئيسيين هما:

- ١- الأسئلة المفتوحة حيث يعطي الفرد حرية كتابة الاستجابة التي يريدها.
- ٢- الأسئلة المقيدة حيث يختار الفرد الاستجابة من بين إجابات معطاة.

ولكل نوع مزاياه وعيوبه (انظر الجدول ١٩-١)، ولذلك فإن اختيار أي من النوعين لصياغة أسئلة الاستبيان يجب أن يكون قائماً على معرفة بهذه المزايا

والعيوب، بالإضافة إلى طبيعة الاستبيان وأهدافه.

ففي الاستبيان ذي الأسئلة المفتوحة يسمح للأفراد بالاستجابة في حرية كاملة وفي عباراتهم الخاصة بدلا من إجبارهم على اختيار إجابة محددة تحديدًا قاطعا. ولذلك فإن هذا النوع يعطي الفرد الفرصة لأن يكشف عن دوافعه واتجاهاته، كما يسمح له بتحديد أسباب ما ذكر من استجابات. ولكن يعيب الأسئلة المفتوحة أنها أحيانا لا توجه الفرد، مما يجعله يحذف عن غير قصد معلومات مهمة. ويحدث هذا عندما يكون السؤال عاما ليس به مؤشرات توجه تفكير الفرد، فلا يستطيع إعطاء تفاصيل كافية، وبخاصة إذا لم يكن المستجيب على درجة كبيرة من الوعي والثقافة، أو إذا لم يكن لديه الوقت الكافي للإجابة على الاستبيان. ومن ناحية أخرى إذا كان الفرد قادرا على إعطاء بيانات كثيرة، فقد تخلق إجاباته مشكلة أخرى هي صعوبة تصنيف البيانات وتلخيصها.

أما في الأسئلة المقيدة فإن الفرد يختار إجابته من قائمة محددة مدد من الاستجابات. وهنا يضع المستجيب علامة أمام الإجابة التي تناسبه من بين اختارين أو أكثر. مثال ذلك:

كم ساعة تستذكر في اليوم عادة؟

٣-٢ ☐

٥-٤ ☐

٧-٦ ☐

٨ فأكثر ☐

ففي هذا المثال يضع المستجيب علامة في المربع المقابل للاستجابة التي يختارها.

ومن السهل تطبيق الاستبيانات ذات الأسئلة المقيدة، كما أنها تساعد الفرد على الاحتفاظ بذهنه مرتبطا بالموضوع. وتيسر عملية تبويب البيانات وتحليلها. إلا أنها غالبا لا تكشف عن دوافع المستجيب وسبب اختياره استجابة معينة، كما أنها لا تعطي في معظم الأحيان معلومات كافية وعميقة. وقد تتركز الأسئلة المقيدة المستجيب باتخاذ موقف محدد لم يكن قد تبلور في ذهنه، كما قد تجبره على إعطاء إجابات لا تعبر عن أفكاره تعبيرا دقيقا. وقد توضع الاستجابات في نظام يشجع المستجيب على أن يختار استجابات تتفق ورأي الباحث. إلا أنه من الممكن التغلب على نقاط الضعف المذكورة إذا اتخذ الباحث احتياطات المناسبة عند بناء الاستبيان.

جدول ١٩ - ١ مقارنة بين الأسئلة المقيدة والأسئلة المفتوحة

مزايا الأسئلة المقيدة	عيوب الأسئلة المقيدة
- أسهل للمستجيبين وأسرع في الإجابة عليها. - يسهل مقارنة إجابات المستجيبين. - يسهل ترميز الاستجابات وتحليلها إحصائياً. - يزيد احتمال استجابة أفراد العينة للأسئلة. - يقل عدد الأسئلة الغامضة والمحيرة. - يسهل إعادة إجراء البحث.	- تعطي المستجيب فرصة إعطاء استجابات لم يفكر فيها. - يستطيع المستجيبون الذين ليس لديهم فكرة عن الموضوع الإجابة على الاستبيان. - قد يصاب المستجيبون بالإحباط لعدم توفر الإجابة التي يرغبون فيها. - إذا زاد عدد الإجابات عن عشرة فقد يحير هذا المستجيب ويصعب عليه الإجابة. - قد يمر عدم فهم السؤال دون ملاحظة. - قد يصعب التمييز بين الإجابات المختلفة. - قد يخطئ المستجيب في وضع العلامة. - تضطر المستجيب إلى إعطاء استجابة بسيطة لقضية معقدة. - قد تضطر المستجيب إلى إعطاء استجابة لا يقوم بها في الحياة الواقعية.
مزايا الأسئلة المفتوحة	عيوب الأسئلة المفتوحة
- للمستجيب حرية إعطاء أي عدد من الاستجابات. - يستطيع المستجيب إعطاء إجابة مفصلة وتوضيح وتبرير إجاباته. - يمكن الحصول على نتائج غير متوقعة. - تسمح بإعطاء استجابات كافية للقضايا المعقدة. - تسمح بحرية الابتكار والتعبير عن الذات وبثراء في التفاصيل. - تكشف عن التفكير المنطقي للمستجيب والإطار المرجعي لعملياته العقلية.	- يختلف المستجيبون فيما بينهم في درجة التفصيلات التي يعطونها. - قد تكون الاستجابات غير مرتبطة بالموضوع أو ضائعة في تفاصيل غير ضرورية. - يصعب تحليلها وعمل مقارنات بين الإجابات المختلفة. - يصعب ترميز الإجابات. - ذوو التعليم الأعلى لهم ميزة على ذوي التعليم البسيط. - قد تكون الأسئلة عامة جداً يصعب الإجابة عليها. - تحتاج إلى وقت طويل حتى يمكن التفكير في الإجابات المختلفة وتدوينها. - قد تحبط بعض الأسئلة المستجيبين. - تحثل الإجابات مساحة كبيرة في الاستبيان.

ومن الممكن مزج النوع المفتوح والنوع المقيد فيترك الباحث الإجابة للفرد في الأسئلة التي تتطلب ذكر سبب معين مثلا، أو عندما يطلب إضافة إجابة أخرى، أو عندما يكون من غير المناسب إعطاء إجابات محددة (Foddy, 1993).

ومن الأفضل بناء الاستبيان على مرحلتين، تكون الأسئلة في المرحلة الأولى من النوع المفتوح، ويطبق الباحث هذه الصورة الأولية على عينة محدودة من الأفراد. ثم يقوم بتحليل الاستجابات واستخدام أكثر الاستجابات شيوعا في كتابة أسئلة مقيدة، وعلى هذا النحو يقلل الباحث من التحيز الذي قد يؤثر على صياغة الأسئلة إذا قام بكتابتها في صيغة أسئلة مقيدة من البداية. وبعد إعداد الاستبيان في صورته النهائية يقوم بتطبيقه على عينة كبيرة من الأفراد. ويمكن صياغة الأسئلة المقيدة بعدة طرق منها:

١- مقاييس التقدير:

في هذا النوع من الأسئلة تعطى قيم أو درجات أو مستويات للإجابة، ويقوم المستجيب باختيار المستوى الذي يتفق ورأيه. ويستخدم هذا النوع بكثرة وبخاصة في مقاييس الاتجاهات كما سبق ذكره. وذلك لأن كثيرا من معتقداتنا وآرائنا يمكن التعبير عنها بدرجة معينة من القوة. فنحن مثلا نوافق تماما على فكرة معينة، أو نعارضها معارضة شديدة. أو قد نكون على حياد بالنسبة لها. وبذلك يسمح هذا الميزان المدرج بالتعبير عن الرأي سواء كان موجبا أو سالبا، أو كان محايدا (راجع مقاييس التقدير في الفصل السابق).

٢- مقاييس الترتيب:

إحدى المشكلات التي ترتبط بمقاييس التقدير هي أن الإجابات قد تكون كلها متشابهة، مما يجعل من الصعب التمييز بين سؤال وآخر. فإذا استخدمنا مثلا طريقة ليكرت في تقدير أهمية مجموعة من المواد لدى الطلبة فقد نحصل على تقديرات تضع جميع المواد في نفس الأهمية. ولكن إذا طلبنا من المستجيبين أن يرتبوا نفس المواد حسب أهميتها فإننا نحصل على قيم مختلفة لكل مادة مثال ذلك:

رتب المواد الدراسية التالية من حيث أفضليتها بالنسبة لك. ضع رقم (١) أمام المادة الأكثر أهمية، ورقم (٢) أمام المادة التي تليها في الأهمية، وهكذا بحيث يعبر رقم (٥) عن أقل المواد أهمية بالنسبة لك.

اللغة العربية ☐

اللغة الإنجليزية ☐

الاجتماعيات	☐
الرياضيات	☐
العلوم	☐

الشكل العام للاستبيان:

هناك عدة طرق لعرض أسئلة الاستبيان. وأفضل هذه الطرق وأوضحها هي كتابة السؤال في سطر والإجابة في السطر الذي يليه. ومن الأفضل استخدام المربعات (الصناديق) أو الأقواس لوضع العلامة بها. مثال ذلك:

- هل تمارس رياضة كرة القدم؟

نعم	☐
لا	☐

وهذه الصيغة أفضل من الصيغة التالية:

- هل تمارس رياضة كرة القدم؟ . . نعم . . لا

وإذا استخدم الباحث صياغة واحدة للأسئلة في الاستبيان أو في أجزاء منه، يمكن بناء الأسئلة والاستجابات على شكل مصفوفة كما في الجدول (١٩-٢).
ويجب أن يكونه تتابع الأسئلة في الاستبيان تتابعا منطقيا. فبدأ الاستبيان بالأسئلة العامة التي تطلب معلومات معينة عن المستجيب، ثم يتبعها أسئلة الاستبيان، ويجب أن تكون هذه الأخيرة متتابعة بشكل منظم ومنطقي.
ويجب أن يشتمل الاستبيان على تعليمات كاملة وواضحة، تبين كيفية الاستجابة للأسئلة، وبحيث تجعل الأمر سهلا على المستجيب فيقبل على الاس تجابة للاستبيان بشكل يخدم البحث ونتائجه.

جدول ١٩-٢ نموذج لجزء من استبيان على شكل مصفوفة

العبارة	موافق تماما	موافق	غير متأكد	غير موافق	غير موافق بالمرّة
١ يسود التعاون بين أفراد الجماعة.					
٢ أنشطة الجماعة متنوعة.					
٣ تخطط الجماعة لأنشطتها مسبقا.					

التطبيق الأولي للاستبيان:

من المرغوب فيه أن يقوم الباحث بإجراء الاستبيان على عينة محدودة من الأفراد كاختبار قبلي. ومن الأفضل أن تكون هذه العينة مشابهة في خصائصها لعينة البحث. ويجب ألا تقل عينة الاختبار القبلي عن ٢٠ فرداً. ويجب إجراء الاستبيان بنفس الطريقة التي سوف يجري بها في البحث، على أن يطلب الباحث من عينة الإجراء القبلي كتابة أية تعليقات يرونها على الأسئلة وعلى الاستبيان ككل. والغرض من هذا الإجراء أن يتبين الباحث مدى وضوح الأسئلة، والوقت الذي يستغرقه الاستبيان بشكل عام. وعما إذا كان هناك تباين كاف بين الاستجابات، ودرجة ثبات الاستبيان. ويقوم الباحث بإجراء أية تعديلات يراها ضرورية قبل تطبيق الاستبيان في صورته النهائية.

المقابلة

المقابلة استبيان منطوق. والخطوات الرئيسية لإعداد المقابلة هي نفسها خطوات بناء الاستبيان، وهي:

- تحديد المبررات.
- تحديد الأهداف.
- كتابة الأسئلة.
- الاختبار القبلي.

والفرق الأساسي بينهما أن المقابلة تتضمن التفاعل المباشر بين الباحث والمستجيب. ولهذا التفاعل مزاياه وعيوبه، وذلك مقارنة بالاستبيان. فالمقابلة مرنة ويمكن تعديلها حسب الموقف، ويمكن استخدامها مع أنواع مختلفة من المشكلات والأشخاص، فيمكن مثلا استخدامها مع الأميين والأطفال الذي لا يستطيعون القراءة أو الكتابة. ويمكن الاستطراد في الأسئلة ومتابعتها وتوضيحها والإطباب فيها حتى يحصل الباحث على استجابات دقيقة مرتبطة بالموضوع. كما يمكن ملاحظة السلوك اللفظي وغير اللفظي للمستجيب أثناء المقابلة، كما أن لدى الباحث فرصة تشجيع المستجيب. ويترتب على هذه العوامل أن نسبة الردود في المقابلة أعلى من نسبة الردود على الاستبيان، وبخاصة في الموضوعات التي تتعلق بالقدرة الشخصية والمشاعر الخاصة. ولذلك تفضل المقابلة على الاستبيان في الموضوعات الشخصية في حين يفضل الاستبيان في الموضوعات العامة.

وأهم عيوب المقابلة هي أنها معرضة للذاتية والتحيز من جانب الباحث، كما أن تكاليفها مرتفعة، وتستغرق وقتا طويلا. ويتوقف نجاح المقابلة على مهارة وتدريب الباحث ومساعدته، فالباحث المحنك يستطيع التأثير بسهولة على الفرد ويكون معه علاقة ودية تشعره بالارتياح وتدفعه للتعبير عن مشاعره الحقيقية. وقد يسأل الباحث أسئلة تمهيدية تهدف إلى تكوين علاقة آلفة بينه وبين المستجيب.

وللتغلب على عيوب المقابلة يجب أن يكون الباحث وسيطا محايدا لتبادل المعلومات. وإذا أمكن تحقيق ذلك فإن وجود الباحث لن يؤثر على المستجيب ويعطيه الفرصة للاستجابة في حرية. وبمعنى آخر فإنه إذا أمكن إجراء المقابلة

بشكل صحيح، فليس من المهم من يكون المقابل، ومهما اختلف المقابلون فإنهم سوف يحصلون على نتائج موضوعية لا تؤثر على نتائج البحث. وإذا أمكن تحقيق ذلك تكون المقابلة ذات ثبات مرتفع. وإذا اتفق مقابلان أو أكثر على طريقة تصنيف معظم الإجابات فإن عملية المقابلة تكون ثابتة، وذلك لاتفاق تقديرات المقابلين. ومن الممكن كذلك تقدير ثبات المقابلة بإجراء معامل ارتباط بين نتائج المقابلة الأصدلية والنتائج التي تحصل عليها مرة أخرى بواسطة نفس المقابل. وهناك أسلوب آخر يمكن استخدامه لزيادة دقة المقابلة، وهو السماح للمستجيب بمراجعة البادئ فيمادونه من استجابات، ومن مزايا هذه الطريقة تكوين علاقة إيجابية بين البادئ والمستجيب، وتقيد هذه العلاقة في إقبال المستجيب على المقابلة، وبذلك فهو في إعطاء المعلومات السليمة، كما أنها مفيدة إذا تبع المقابلة الأولى عدة مقابلات أخرى تتبعية تقويمية.

الإعداد للمقابلة:

بعد تحديد مبررات المقابلة وأهدافها يقوم الباحث بإعداد دليل المقابلة. ويحتوي الدليل على جميع الأسئلة التي توجه خلال المقابلة، مع ترك فراغات كافية يكتب فيها الباحث الاستجابات. وترتبط هذه الأسئلة ارتباطاً مباشراً بأهداف الدراسة، وتتبع ترتيباً معيناً يجب الالتزام به في كل مقابلة. وفي معظم الحالات تلقى الأسئلة المكتوبة شفها، مع إضافة أسئلة توضيحية مناسبة. ويمكن تصنيف أسئلة المقابلة في ثلاثة أنواع:

١- **مقننة**، وفيها تكون أسئلة المقابلة محددة ويتبع كل سؤال مجموعة من الاختيارات أو الإجابات يختار من بينها المستجيب الإجابة التي تتفق مع رأيه. مثال ذلك:

- هل تعتقد أن هذا البرنامج فعال؟
- فعال جداً
- فعال إلى حد ما.
- غير فعال.

وتمتاز المقابلة المقننة بأن الاختلافات بين المقابلين تكون محدودة، مما يرفع من ثبات المقابلة، كما أنه لا مجال أمام الباحث لتفسير أو تعديل الاستجابات مما يزيد في صدق المقابلة. إلا أن المقابلة المقننة قد تصبح جامدة إذا كانت العلاقة بين الباحث والمستجيب مجرد علاقة سؤال وجواب.

٢-شبه المقننة، وهنا نجد أن الأسئلة لا يتبعها اختيارات محددة ولكن تصاغ الأسئلة بحث تسمح بالإجابات الفردية، فالسؤال مفتوح ولكنه محدد للغاية في محتواه، مثال ذلك:

- ما أهم فائدة لبرنامج التربية العملية؟

وتمتاز المقابلة شبه المقننة بأنها تسمح بوجود علاقة تفاعلية بين الباحث والمستجيب، مما يجعله يقبل على المقابلة ويبدل جهده في إعطاء المعلومات الضرورية. ولكن يعيبها أنها أقل موضوعية من المقابلة المقننة.

٣-غير المقننة، وفي هذا النوع من المقابلة يقوم الباحث بتوجيه أسئلة عريضة في أي ترتيب يراه مناسباً، وهنا نجد أن المقابلة تركز على المستجيب، ورغم أن هذا الأسلوب يتميز بوجود علاقة ألفة كبيرة بين الباحث والمستجيب، إلا أن درجة صدقها وثباتها محدودة مقارنة بالنوعين السابقين.

ومن الأفضل في البحوث التربوية والنفسية أن نتبع مزيجاً من المقابلة المقننة وشبه المقننة، مما يحقق درجة عالية من الموضوعية والتجانس. وتسمح في نفس الوقت بالأسئلة البينية والتوضيحية. وبعد إعداد أسئلة المقابلة يجب إجراؤها في مقابلات أولية تكون اختباراً قلياً للمقابلة. ويجب عند صياغة الأسئلة البعد عن الأسئلة الإيحائية. مثال ذلك السؤال التالي:

- في ضوء سياسة ترشيد الإنفاق، هل تعتقد بضرورة البدء في إنشاء مبان جديدة لكلية التربية؟

ويذكر شوماكر وم إكميلان (Schumacher & McMillan, 1984, pp. 243-244) نقلاً عن (Sax, 1969) أن هناك متغيرات شخصية تتعلق بالباحث لها أثرها على المقابلة، ويجب أخذ هذه العوامل في الاعتبار. ويوضح الجدول (١٩-٣) هذه المتغيرات وأثرها على المقابلة.

إجراء المقابلة:

مظهر المقابل في غاية الأهمية، ولذلك يجب أن يكون مظهره عادياً لا يوحي بأنه ينتمي لفئة معينة. ويجب أن يكون الباحث (المقابل) ودوداً، ومسترخياً، وبشوشاً، ويظهر اهتماماً بالمستجيب. ويجب أن يشعر المستجيب بالاسترخاء وذلك هام لصدق المعلومات التي يدلي بها. ويجب أن توجه الدقائق الأولى من المقابلة نحو تكوين علاقات الألفة بين الباحث والمستجيب، لأن ذلك يساعد على شعور

المستجيب بالاسترخاء وبالتالي يزيد إقباله على المقابلة. ويجب على الباحث تفسير الغرض من المقابلة في اختصار، كما يعطي للمستجيب الفرصة لأية استفسارات حول المقابلة قبل البدء في توجيه أسئلة المقابلة.

ويجب توجيه الأسئلة بالنص كما هي مكتوبة في دليل المقابلة، ويجب أن يتجنب الباحث إعادة صياغة السؤال بعبارات من عنده لأن ذلك قد يؤثر على معنى السؤال. ويجب قراءة الأسئلة بلغة صحيحة وبدون أخطاء. ولذلك يجب أن يكون الباحث على ألفة كبيرة بالأسئلة قبل بدء المقابلة.

تسجيل الإجابات:

أفضل طريقة لتسجيل الإجابات هي تدوينها حرفياً كما أعطاه المستجيب. ونظراً لأنه من غير المستحب الانصراف عن المستجيب أثناء المقابلة لتدوين استجاباته، فمن الأفضل أن يقوم الباحث بكتابة نقاط مختصرة أثناء المقابلة ثم يوسع هذه النقاط عقب المقابلة، ومعنى ذلك ألا يحاول الباحث استغراق وقت أكثر مما هو اللازم أثناء المقابلة في تدوين الاستجابات حرفياً، ولا أن يترك تسجيل الإجابات إلى نهاية المقابلة حتى لا ينسى الاستجابات التي ذكرها المستجيب. ويلجأ بعض الباحثين إلى استخدام أجهزة التسجيل لتسجيل وقائع المقابلة، وفي حال استخدام جهاز تسجيل يجب أولاً استئذان المستجيب. وإن كان من الأفضل البعد عن مثل هذه الممارسة، لأن استخدام أجهزة التسجيل قد تؤدي إلى تحفظ المستجيب في إعطاء استجابات صريحة، لعلهم أن ما يقوله يسجل عليه، وبخاصة عندما تكون الاستجابات ذات طابع انفعالي، إذ يتحفظ المستجيب في إعطاء الاستجابة بحرية لأنه يعلم أن استجاباته تسجل. ولذلك يجب أن يشرح الباحث بعناية الغرض من التسجيل ويحاول كسب ثقة المستجيب، وذلك للإقلال من الآثار غير المرغوبة لتسجيل المقابلة.

ويجب أن يستفسر الباحث من المستجيب إذا وجد أن بعض إجاباته غير واضحة، أو غامضة. ويلاحظ أن الاستفسار عن الإجابات الغامضة مائة، إذا افتقدها الباحث قد تؤدي إلى الحصول على استجابات غير دقيقة أو غير كاملة. ويجب على المقابل أن يتيح للمستجيب وقتاً كافياً لإعطاء استجاباته، كما يجب أن يتجنب الإيحاء بأنه يتوقع إجابة معينة. ويجب أن تكون الاستفسارات محايدة تماماً، بحيث تكون توضيحاً لبعض نقاط الاستجابة عندما يجد الباحث ذلك ضرورياً، كما أن يقول مثلاً: 'ماذا تقصد؟' أو 'أرجو أن توضح لي هذه النقطة'، على أن يسجل

الإجابات الإضافية (عن الاستفسارات) في الهامش إلى جانب الاستجابة الأصلية. ويجب أن يسمح الباحث للمستجيب بعمل ما يريد من تعليقات حول سير المقابلة. وبعد انتهاء المقابلة يجب على الباحث أن يشكر المستجيب على وقته وعلى تعاونه مع الباحث أثناء المقابلة.

جدول ١٩-٣ المتغيرات التي تؤثر على المقابلة

عدد	المتغير	أثره على المقابلة
١	عمر الباحث	♦ تكون الألفة عالية إذا كان المقابل شابا يقابل مستجيبين من أواسط العمر، عمر المستجيب غير هام بالنسبة للباحثين الأكبر. ♦ أثر كفاية الاستجابة محدود إذا كان الباحث والمستجيب من الشباب ومن نفس النوع. ♦ يزداد الكفاية إذا كان الباحث والمستجيب من نفس العمر ولكن من جنسين مختلفين. ♦ الباحثون الذين تتراوح أعمارهم بين ٢٦ و ٥٠ سنة أفضل من الباحثين الأصغر أو الأكبر.
٢	التخصص الجامعي	المقابلون المتخصصون في العلوم السلوكية أكثر دقة من المتخصصين في العلوم الطبيعية. أقل المقابلين هم المتخصصون في الفنون الجميلة والتجارة والحقوق والإنسانيات.
٣	المستوى التعليمي	خريجو الجامعة أفضل من غير الخريجين إلا أن الفروق ضئيلة بينهما.
٤	الخبرة في المقابلة	تزداد دقة المقابلين مع زيادة خبراتهم في المقابلة.
٥	جنس المقابل	يحصل الباحثون الذكور على استجابات أقل من الإناث الباحثات.

obeikandi.com

الفصل العشرون

جمع البيانات باستخدام

الملاحظة وتحليل المحتوى

ناقشنا في الفصل السادس عشر جمع البيانات باستخدام الاختبارات والمقاييس، وفي الفصل السابع عشر جمع البيانات باستخدام الاستبيان والمقابلة، ولاحظ أنه في هذه الوسائل يقوم أفراد العينة بملاحظة أنفسهم ملاحظة ذاتية داخلية، ثم يعطون الاستجابة التي يعتقدون أنها الاستجابة المناسبة. ونظراً لأن الفرد يعطي الاستجابة بنفسه عن نفسه فإن هناك شبهة التحيز من جانب المستجيب فيما يعطي من استجابات. كما أن عامل النسيان قد يلعب دوره هو الآخر إذ قد لا يتذكر المستجيب بدقة بعض الأحداث الماضية المتعلقة بالبحث.

وبمعنى آخر يمكن القول إن جميع أساليب جمع البيانات هي نوع من أنواع الملاحظة، فالمستجيب للاستبيان مثلاً لابد أن يلاحظ نفسه ملاحظة ذاتية ليقرر كيف يتصرف في موقف معين أو ما رأيته بالنسبة لقضية معينة، ونفس الشيء يحدث في المقابلة، حيث تتوقف استجابة الفرد في المقابلة على ملاحظته لذاته. فالملاحظة بمعناها العام تشمل جمع البيانات بغض النظر عن الأداة المستخدمة في جمعها. إلا أن طريقة الملاحظة كأسلوب من أساليب جمع البيانات تختلف تماماً عن المقابلة أو الاستبيان. فطريقة الملاحظة تعتمد على رؤية الباحث أو سماعه للأشياء وتسجيل ما يلاحظه، ولا يعتمد على استجابات أفراد العينة لأسئلة أو عبارات يقرأونها في الاختبار أو الاستبيان، أو تلقى عليهم في المقابلة، أي أن الباحث لا يصل على الاستجابات من المستجيب (التقرير الذاتي)، ولكن يحصل عليها بنفسه عن طريق ملاحظة سلوك أفراد العينة في مواقف طبيعية.

فملاحظة سلوك أفراد العينة وبيئاتهم المادية والاجتماعية ملاحظة مباشرة في المواقف الطبيعية هي البديل للتقرير الذاتي. ولذلك قام الباحثون في كل من البحوث الكمية والبحوث الكيفية بتصميم وسائل منظمة لهذا الغرض. وإذا استخدمت

هذه الوسائل استخداما سليما فإنها كفيلة بتجنب التحيز وعدم الدقة اللذين يشوبان الوسائل الأخرى التي تستخدم تقرير الفرد عن نفسه. وحتى إذا لم تكن بيانات التقرير الذاتي بيانات متحيزة فإن وسائل الملاحظة يمكن أن تعطي بيانات أكثر دقة. مثال ذلك أن بعض التربويين قد بينوا أن المدرسين يتكلمون معظم الوقت في قاعات الدراسة على حساب مشاركة الطلاب. ولكن ما هي النسبة الفعلية لحديث كل من المدرسين والطلاب في الفصول؟ إن التقارير الذاتية للمدرسين أو الطلاب غير كفيلة بالكشف عن هذه النسبة، ولكن الملاحظات التي تسجل باستخدام التسجيلات السمعية أو البصرية يمكن أن تعطي هذه البيانات (Gall et al., 2003).

وبالرغم من أن الملاحظة تتفوق على التقرير الذاتي بالنسبة لأغراض بعض البحوث، إلا أنها أكثر استهلاكا للوقت. إذ يجب ملاحظة الأفراد على مدى فترة زمنية طويلة للحصول على بيانات ثابتة، بينما نجد أن استخدام الاختبارات أو الاستبيانات أو المقابلات يمكن أن تعطي بيانات ثابتة حتى ولو استخدمنا هذه الوسائل في زمن قصير. كذلك نجد أنه عند استخدام أساليب الملاحظة في الدراسات الكمية لابد من الحصول على ثبات تقديرات الملاحظين وذلك عن طريق استخدام ملاحظين مستقلين عن بعضهم البعض في الموقف الذي تتم فيه الملاحظة. وهذا مطلب صعب إذا كان الباحث يعتمد كلية في جمع البيانات على استخدام أسلوب الملاحظة.

ويتناول هذا الفصل أسلوب الملاحظة الذي يستخدمه الباحثون الكميون والكيفيون في جمع البيانات من الأفراد في بعض المواقف. وبعض أنواع الملاحظة تتطلب عدم معرفة أفراد العينة بأنهم في موقف ملاحظة. كما نتناول أيضا في هذا الفصل أساليب تحليل محتوى الوثائق التي نحصل عليها من بعض المواقف الطبيعية.

الملاحظة

أساليب الملاحظة في البحوث الكمية

الملاحظة المنظمة في المواقف الطبيعية وسيلة مفيدة للغاية في جمع البيانات عن أداء الأفراد وسلوكهم. والخطوة الأولى في الملاحظة في البحوث الكمية تعريف المتغيرات التي نلاحظها. ويرى جول وزملاؤه أن هناك ثلاثة أنواع لهذه المتغيرات (Gall et al., 2003):

١- **متغيرات الملاحظة الوصفية:** وهي المتغيرات التي لا تتطلب تدخلًا من جانب الملاحظ. ويطلق عليها أحيانا المتغيرات الظاهرة، وأهم ميزة لها هذه المتغيرات أنها تعطينا معلومات موضوعية.

٢- **متغيرات الملاحظة الاستدلالية:** وهي المتغيرات التي تتطلب استدلالًا من جانب الملاحظ من السلوك الملاحظ إلى التكوين الذي ينتمي إليه السلوك. فقد يطلب مثلا من الملاحظ أن يسجل الثقة بالنفس التي يظهرها المدرس عندما يشرح مفهوما رياضيا. فبعض المدرسين يتكلمون بدرجة عالية من الثقة بالنفس، في حين يبدو الآخرون غير متأكدين أو مرتبكين أو قلقين لأن فهمهم للموضوع ضعيف. والثقة وعدم التأكد والارتباك والقلق ليست سلوكا ولكنها تكوينات نفسية نستدل عليها من السلوك الظاهر. ولذلك يطلق عليها أحيانا **متغيرات استدلالية عالية**. ومن الصعب جمع بيانات عن المتغيرات الاستدلالية من المتغيرات الوصفية.

٣- **متغيرات الملاحظة التقويمية:** وهي المتغيرات التي لا تتطلب فقط الاستدلال من جانب الملاحظ ولكنها تتطلب أيضا حكما تقويميا. مثال ذلك أننا قد نطلب من الملاحظ أن يعطي تقديرات عن مستوى جودة شرح المدرس لمفهوم رياضي. والجودة ليست سلوكا ولكنها تكوين نستنتجه من السلوك، وهو تكوين تقويمي بطبيعته. ونظرا لأنه من الصعب ملاحظة المتغيرات التقويمية ملاحظة ثابتة، فإننا نحتاج إلى جمع أمثلة تحدد نقاطا على متصل يتدرج من الشرح الممتاز إلى الشرح الضعيف ونستخدم هذه النقاط لتدريب الملاحظين.

وتتراوح الملاحظة من الملاحظة القصيرة المضبوطة ضبطا عاليا يستخدم فيها مقاييس التقدير، وقوائم المراجعة والعينات الزمنية إلى الملاحظات الطويلة التي تستخدم السجلات القصصية. ويلاحظ أن مصطلحي 'أدوات الملاحظة' و'أسلوب

الملاحظة؛ ليسا مترادفين، والمقصود بأدوات الملاحظة الأدوات التي نستخدمها أثناء الملاحظة لتسجيل الملاحظات مثل قوائم المراجعة، ومقاييس التقدير، والسجلات القصصية. أما المقصود بأسلوب الملاحظة عملية ملاحظة السلوك ذاته لا تمهيدا لتسجيلها (Cartwright & Cartwright, 1984).

ويختلف دور الباحث أو الملاحظ في البحث بناء على درجة مشاركته وانهماكه في النشاط الذي يلاحظه. فمن ناحية يمكن أن يصبح الملاحظ عضوا في الجماعة التي تلاحظ، أي يكون في هذه الحالة مشاركا كاملا. ومن ناحية أخرى قد يظل الملاحظ منفصلا عن الجماعة التي يلاحظها، وبذلك يصبح دوره ملاحظا كاملا. فقد يرغب الباحث مثالا في دراسة توافق المتحقيين الجدد بالجامعة مع الحياة الجامعية وذلك عن طريق المشاركة كطالب مستجد في الجامعة، فيتفاعل مع الطلبة الآخرين كعضو من أعضاء الجماعة. ومن ناحية أخرى قد يلاحظ الباحث الطلبة المستجدين في مواقف مختلفة أثناء تواجدهم في الحرم الجامعي كشخص خارجي على الموقف، لا يشارك ولكنه يسجل ملاحظاته عما يرى أو يسمع. وبين هذين الأسلوبين المتناقضين هناك درجات مختلفة من مشاركة الملاحظ في البحث. وفي البحوث التربوية والنفسية نجد أن أغلب الدراسات التي تعتمد على الملاحظة تعرف دور الملاحظ بأنه ملاحظ كامل. إلا أن هناك توجها واضحا نحو استخدام الملاحظة بالمشاركة في البحوث التربوية، وهذا ما ناقشناه في القسم السابق عند مناقشة البحوث الكيفية. أما في هذا الفصل فإن المقصود بطريقة الملاحظة هو أن الباحث يقوم بدور الملاحظ الكامل.

مبررات الملاحظة:

من المزايا الأساسية لطرق الملاحظة هي أنها تقضي على عيوب الأساليب التي تعتمد على الملاحظة الذاتية مثل الاستبيان والمقابلة. فالملاحظة تقلل من المشكلات التي ترجع إلى تحيز التقارير الذاتية، والمرغوبة الاجتماعية، كما أن المعلومات التي يحصل عليها الباحث لا تتأثر بعوامل التذكر التي تتأثر بها الاستجابات التي يعطيها المستجيب للاستبيان أو في المقابلة. فالباحث يسجل السلوك كما يحدث أمامه في موقف طبيعي. وهذه الميزة مهمة للغاية للبحث الذي يصمم لدراسة ما يحدث في الحياة الواقعية الحقيقية. إلا أن طريقة الملاحظة باهظة التكاليف، وصعب إجراؤها إجراء ثابتا وذلك بسبب ملاحظة السلوك الإنساني المركب. ومن السهل تسجيل السلوك البسيط بموضوعية، إلا أن معظم الدراسات تركز على السلوك الأكثر تعقيدا الذي يصعب تعريفه وتقويمه باستخدام الملاحظة.

كما أن وجود الملاحظ في الموقف له أيضا تأثيره على سلوك الأفراد الذين يلاحظهم. ولأسلوب الملاحظة بعض نواحي القصور الأخرى، فهو كثيرا ما يتأثر بذاتية الملاحظ، إذا حاد عن الموضوعية، ولذلك يجب أن يقترب تسجيل الملاحظات من وسائل التسجيل الآلية، ليسجل السلوك الفعلي الذي تمت ملاحظته دون أن يحاول الملاحظ تفسير هذا السلوك أو تأويله (رجاء أبوعلام، ١٩٨٩).

تسجيل الملاحظات وتحليلها:

لضمان دقة تسجيل الملاحظات يجب أن نطلب من الملاحظين تسجيل البيانات عن متغير واحد فقط في المرة الواحدة. إذ يجد معظم الملاحظين من الصعب جدا تسجيل عدة مظاهر لسلوك المدرس أثناء تسجيلهم لنسبة الأطفال الذين ينتبهون للمدرس. وفي مثل هذه الحالة يكون ثبات هاتين المجموعتين من الملاحظة منخفضا. ولذلك يمكن تعيين ملاحظين مختلفين لتسجيل كل نوع من السلوك، أو قد يقوم ملاحظ واحد بالتبديل بين تسجيل سلوك المدرس لفترة محددة، ثم ينتقل إلى تسجيل سلوك الأطفال في الفترة التالية.

ويمكن تصنيف إجراءات تسجيل الملاحظات في أربعة أنواع رئيسية (Gall et al., 2003):

١- **تسجيل الفترة الزمنية:** في هذا النوع من التسجيل يقيس الملاحظ الفترة الزمنية التي يحدث فيها السلوك الملاحظ. ويستخدم الملاحظ ساعة توقيت عادة لهذا الغرض. ومن السهل تسجيل الفترة الزمنية عند ملاحظة سلوك واحد فقط، مثل قياس الفترة الزمنية التي يترك فيها طالب معين مقعده. ويمكن للملاحظ كذلك تسجيل عدة أنواع من السلوك إذا لم تحدث في نفس الوقت. إذ يمكن أن يسجل الملاحظ الفترة الزمنية التي يقضيها طالب ما في العمل أو الراحة. ولكنه لا يستطيع أن يسجل إذا ما أخذ الطالب بالنظام بدرجة كبيرة أو متوسطة أو صغيرة.

٢- **تسجيل عدد مرات حدوث السلوك:** وفي هذا النوع من التسجيل يسجل الملاحظ كل مرة يحدث فيها السلوك. ويستخدم الملاحظ في هذه الحالة أوراقا خاصة بالعلامات التكرارية. وأنسب أنواع السلوك الذي يمكن تسجيل تكراره السلوك ذو الفترة الزمنية القصيرة، وكذلك السلوك الذي لا يهم مدى استمراره. مثال ذلك الدراسة التي قام بها بورج (Borg, 1977) ودرب فيها كل ملاحظ على قياس تكرار ١٣ نوعا من سلوك المدرس المرتبط بإدارة الفصل، مثل إعطاء التوجيهات للتلاميذ، وتوجيه المديح،

وإعطاء الإشارات التحذيرية. وكانت فترة استمرار السلوك قصيرة، ولم يحدث أكثر من سلوك واحد في نفس الوقت. ولذلك كان ثبات النقط ديرات بين الملاحظين مرضيا إذ تراوح بين ٧١، و ٩٦، لأنواع السلوك الثلاثة عشرة.

٣- **التسجيل الفتري:** يتضمن هذا النوع من التسجيل ملاحظة سلوك الفرد عند فترات محددة. ويمكن اعتبار تعديل ساكاجوتشي لنظام تحليل فلاندرز للتفاعل مثالا لهذا النوع من التسجيل. وكان سلوك المدرس أو الطالب يرقم بأحد الأرقام الدالة على الفئات من ١ إلى ٢٥ كل ثلاث ثوان. وبمجرد تسجيل عينة من السلوك بهذه الطريقة يجب تلخيص البيانات وإعطاء تقرير عنها للحصول على أوصاف ذات معنى لما حدث. وكان الهدف الرئيسي من دراسة ساكاجوتشي تحديد كيفية اختلاف المدرسين الستة في العينة عن بعضهم البعض في أسلوب تدريسهم. وكانت البيانات الأولية في كل درس تمت ملاحظته عبارة عن عدد الفترات ذات الثلاث ثواني التي حدث فيها السلوك الملاحظ. وقد تم قسمة عدد الفترات التي حدث فيها السلوك على مجموع عدد الفترات التي سجلت وذلك بالنسبة لكل متغير. وكانت نتيجة هذه العملية الحصول على النسبة المئوية لمجموع زمن الحصص التي حدث خلالها كل متغير. وكانت الخطوة الأخيرة لهذا التحليل الحصول على متوسط النسب المئوية لكل متغير على مدى الحصص الست التي لوحظ فيها المدرس.

٤- **التسجيل المستمر:** يتطلب هذا النوع من التسجيل تسجيل جميع أنواع السلوك الصادرة عن الفرد الملاحظ أو الأفراد الملاحظين لفترة زمنية محددة. ولا يركز هذا النوع عادة على مجموعة معينة من السلوك. وبدلاً من ذلك يكتب الملاحظ "بروتوكولا" يروي فيه بشكل قصصي التتابع الزمني للأحداث لكل شيء يقوم به الفرد أو كل شيء يحدث في موقف معين، مثل قاعة الدرس. وتستخدم هذه الطريقة في الدراسات الاستطلاعية لمساعدة الباحث على التعرف على الأنماط السلوكية المهمة، والتي تدرس فيما بعد باستخدام طرق التسجيل الأخرى.

ونظراً لأنه من المستحيل تسجيل كل شيء في البروتوكول يجب أن يركز الملاحظ على الأحداث والمظاهر الأكثر ارتباطاً بمشكلة البحث. ولتحليل البروتوكولات يقرأها الباحث ويجري نظاماً لتحليل المحتوى المناسب

للبيانات. ثم يعيد قراءة البروتوكولات ويصنف السلوك الملاحظ ويدخله في النظام الذي وضعه. وتتوافق هذه العملية مع خطوات تحليل محتوى الوثائق وغيرها من وسائل الاتصال المستخدمة في البحوث الكمية. وسوف نعود إلى هذه النقطة في مكان لاحق من هذا الفصل عند مناقشة موضوع تحليل المحتوى.

اختيار الملاحظين وتدريبهم:

يمكن للباحث أن يقوم بملاحظة أفراد العينة بنفسه، أو يدرب آخرين للقيام بالملاحظات، أو يشارك آخرين هذا العمل. وميزة استخدام الأفراد الآخرين أن ذلك يسمح بضبط تحيز الملاحظ الذي يمكن أن يحدث إذا كان الملاحظ هو نفسه الباحث الذي صمم منهج البحث ووضع فروضه. وبالإضافة إلى ذلك إذا كان هذا ملاحظان أو أكثر حصلوا على نفس التدريب ويقومان بملاحظات مستقلة يستطيع الباحث تحديد درجة ثبات الملاحظات. وأكثر الملاحظين دقة يكونون عادة أذكاء وذوي طلاقة لفظية ولديهم دافع قوي للقيام بالمهمة (Harter, 1982).

والخطوة الأولى في تدريب الملاحظين هي مناقشة أدوات الملاحظة معهم. مع وصف كل مفردة في الأداة وصفا كافيا حتى يفهموا بعمق ما الذي يلاحظونه، وكيف يسجلون ملاحظاتهم. ويمكن عرض تسجيلات فيديو أثناء التدريب للمواقف التي سوف تلاحظ في الدراسة، وتمكن هذه التسجيلات الملاحظ من إعطاء أمثلة واقعية وربطها بتعريفات كل سلوك.

والخطوة التالية هي إقامة ملاحظات تدريبية يشارك فيها الملاحظون المتدربون. ويمكن استخدام تسجيلات الفيديو السابق الإشارة إليها في الملاحظات التدريبية. ويمكن عرض جزء بسيط من التسجيل ثم يطلب من الملاحظين المتدربين تسجيل السلوك على استمارة تسجيل البيانات كما يحدث السلوك في شريط الفيديو بالضبط. وبعد ذلك تراجع التسجيلات التي قام بها الملاحظون المتدربون للتأكد من أنهم سجلوا السلوك بشكل صحيح. وإذا اختلف الملاحظون المتدربون مع بعضهم البعض أو مع محكات التسجيل يعاد عرض شريط الفيديو مع التوقف عند كل سلوك لمناقشة أنسب وسيلة لتسجيله ولماذا. ويجب مراجعة استمارة تسجيل الملاحظات أثناء هذه المناقشات لإضافة أية توضيحات قد تثار أثناء جلسة التدريب. ويتطلب الأمر إعطاء مجموعة من القواعد لمساعدة الملاحظين على ملاحظة السلوك وتسجيله. وفيما يلي أهم القواعد التي تساعد الملاحظين على ملاحظة وتسجيل مواقف الملاحظة التي يشاركون فيها:

قواعد الملاحظة:

لكي تكون الملاحظات دقيقة وصادقة يجب اتباع القواعد التالية:

١- التخطيط مسبقا لما نلاحظه، وذلك بناء على أهداف الملاحظة التي ندرسها. ويجب الفصل تماما بين ما نلاحظه من سلوك وأية تفسيرات يمكن استنتاجها من الملاحظة. على أن يجيب الباحث قبل البدء في جمع البيانات على السؤالين التاليين عند التخطيط للملاحظة:

أ- ما الغرض من الملاحظة؟

ب- ما السلوك الذي يجب ملاحظته؟

والسؤال الثاني هو الأهم وهو المرتبط مباشرة بما نلاحظه، ولذلك يجب تحديد أنواع السلوك المراد ملاحظته، وتعريف هذا السلوك تعريفًا إجرائيًا ليس سهلاً تسجيله. ولهذا من الأفضل استخدام بعض الأدوات التي تساعدنا على تسجيل الملاحظات. ورغم أن أدوات الملاحظة ليست اختبارات بالمعنى التقليدي إلا أنها يمكن أن تعطينا بيانات تتعلق بما يلي:

أ- الفروق الفردية بين الأفراد.

ب- عمل تنبؤات عن السلوك في المستقبل.

ولذلك فمن الواجب أن نعرف مسبقا ماذا نلاحظ وكيف نلاحظه.

٢- التركيز على نوع أو نوعين من السلوك فقط. ففي فصل قد يزيد عدد الطلبة فيه على ٣٠ طالبا يستحيل ملاحظة كل أنواع السلوك الحادث في وقت ما. ولكن يستطيع الباحث أن يحصل على بيانات دقيقة إذا ركز على نوع واحد أو نوعين من السلوك. وقد يعتقد البعض أن التركيز على نوع واحد من السلوك يترتب عليه عدم الانتباه لأنواع أخرى مهمة من السلوك. إلا أن السلوك إذا كان يمثل نمطا فإنه سوف يتكرر. وإذا لم يحدث إلا مرة واحدة فلماذا أن نفترض أنه غير هام. ولكن في نفس الوقت يجب أن يكون الباحث على حذر من بعض الأحداث غير المتوقعة التي قد تكون مهمة في فهم سلوك الفرد.

٣- استخدام صفات واضحة غير غامضة حتى تكون الملاحظة مدددة صف السلوك وصفا سليما. ومن هنا أهمية تعريف السلوك المرغوب ملاحظته تعريفًا إجرائيًا حتى يمكن تحويله إلى سلوك ظاهر قابل للملاحظة المباشرة.

٤- أن يكون كل سلوك ملاحظ مختلفا عما عداه من أنواع السلوك الأخرى، ذلك أن تداخل أنواع السلوك يجعل من الصعب وصفها وتصنيفها. والتعريف

الإجرائي للسلوك يساعد على تجنب التداخل بين أنواع السلوك المختلفة.

٥- أن يكون الباحث واعيا بما يحدث من أخطاء الملاحظة التي تحدث نتيجة لاختيار أوقات معينة نلاحظ فيها السلوك. ولتقادي هذا الخطأ أو الإقلال منه لابد من ملاحظة نفس الفرد في أوقات وأيام مختلفة فنلاحظه في اليوم الأول مثلا الساعة التاسعة صباحا وفي اليوم الثاني الساعة العاشرة صباحا، وبعد يومين آخرين نلاحظه الساعة الثانية ظهرا وهكذا. أي نقوم بعمل جدول لملاحظة فرد، وجدول آخر مختلف لملاحظة فرد ثان، وهكذا. ومن الأفضل أن تتم الملاحظات على فترات متكررة وقصيرة موزعة على عدة أسابيع وفي أوقات مختلفة من اليوم، حتى نحصل على عينة زمنية ممثلة لسلوك الأفراد.

٦- تسجيل وتلخيص الملاحظات عقب حدوثها مباشرة، إذا كان من المتعذر تسجيل السلوك أثناء حدوثه.

٧- أن يختار الباحث من يلاحظه في كل مرة، فالإقتصار على عدد محدود من الأفراد كل مرة يجعل الملاحظة أيسر وأسهل في تسجيلها.

٨- تأجيل تفسير السلوك إلى ما بعد جمع البيانات، فكثيرا ما يترتب على محاولة الباحث تفسير السلوك أثناء حدوثه الإخلال بموضوعية الملاحظة.

٩- ألا يظهر الباحث أنه يلاحظ سلوكا ما أو فردا ما، فعندما يعلم الناس أنهم تحت المراقبة يتصرفون بطريقة غير طبيعية.

أدوات الملاحظة:

تجمع البيانات أثناء الملاحظة باستخدام عدد من الأدوات. وأهمية هذه الأدوات هي ضمان موضوعية التسجيل، لأن الأداة التي يستخدمها الملاحظ تجعله يركز ملاحظاته على السلوك المطلوب والمحدد في الأداة. واللجوء إلى الملاحظة دون استخدام أداة ما يتيح لعوامل التحيز أن تتسلل إلى الملاحظات.

وأكثر الأدوات استخداما لتسجيل الملاحظات هي مقاييس التفرقة والمراجعة. وقد سبق لنا دراسة هاتين الأداتين في الفصل السادس عشر. ويجب الالتزام بكل ما جاء في الكلام عن هاتين الأداتين حتى يمكن الحصول على بيانات سليمة أقرب إلى الموضوعية، كما يمكن استخدامها في تسجيل الملاحظات بشكل موضوعي، مما يسمح بالحصول على بيانات البحث المرغوبة.

تحديد الاتفاق بين الملاحظين:

البيانات التي نجمعها باستخدام الملاحظة في المواقف الطبيعية لا قيمة لها إذا لم تجمع البيانات بواسطة ملاحظين موثوق بهم. فماذا الذي نقصد به بثبات الملاحظين؟ يمكن القول إن هناك ثلاثة أنواع من ثبات الملاحظة:

١- **ثبات الملاحظ المرتبط بمحك:** درجة الاتفاق بين تقديرات الملاحظ المدرب وتقديرات ملاحظ خبير مثل الباحث الذي صمم أداة الملاحظة. وهذا النوع من الثبات هام لأنه يوفر لنا الأدلة على أن فهم الملاحظ المدرب للمتغيرات التي لاحظها باستخدام أداة الملاحظة يشابه فهم الخبير أو الباحث. ويمكن تحقيق هذا النوع من الثبات بالحصول أولاً على تقديرات الخبير على جميع المتغيرات التي نقيسها باستخدام أدوات الملاحظة. وبعد قيام الملاحظ بإعطاء تقديراته للسلوك الذي لاحظته تحسب درجة الاتفاق بينه وبين تقديرات الخبير.

٢- **ثبات تقديرات الملاحظ:** وتعني هذه الطريقة الحصول على درجة الاتفاق بين تقديرات الملاحظ. ويمكن تحقيق هذا النوع بالحصول على تقديرات الملاحظ مرتين للمتغيرات التي لاحظها خلال بضعة أيام. ولا يتم تحقيق هذا النوع من الثبات عادة، ولكن يجب على الباحث تحقيقه قبل البدء في جمع البيانات الفعلية، ومن الأفضل الحصول عليه كذلك أثناء جمع البيانات.

٣- **ثبات التقديرات بين الملاحظين:** وهذا النوع من الثبات يعزى درجة الاتفاق بين الملاحظين بعضهم البعض أثناء الجمع الفعلي للبيانات. ولتحقيق ثبات التقديرات بين الملاحظين نحدد صل على تقديرات أزواج الملاحظين الذين يجمعون البيانات لنفس المتغيرات بشكل فردي. لنفرض أننا دربنا خمسة ملاحظين على جمع بيانات عن تدريس مجموعة من المدرسين خلال الفصل الدراسي. ويقوم كل ملاحظ بجمع البيانات بمفرده وبشكل مستقل عن الملاحظين الآخرين، ولكن في نفس الفصول وباستخدام نفس المتغيرات. وبعد ذلك نقوم بوضع تقديرات الملاحظين في جدول واحد، ونقوم بحساب درجة الاتفاق بين كل زوج من الملاحظين.

وعند حساب ثبات التقديرات نحسب النسبة المئوية للاتفاق بين كل ملاحظين، أو نحسب معامل الارتباط بين التقديرات إذا كان معامل الارتباط يناسب هذه التقديرات.

خفض أثر الملاحظ:

يقصد بأثر الملاحظ سلوك من جانب الملاحظ يترتب عليه أثر سلبي على صدق أو ثبات البيانات التي يجمعها. ويجب أن ينتبه الباحثون إلى احتمال وجود آثار للملاحظ، ويقومون باتخاذ الإجراءات الضرورية للإقلال منها أو تجنبها. وقد قامت كارولين إفرتسون وجوديث جرين (Evertson & Green, 1986) بالتعرف على الأنواع المختلفة لآثار الملاحظين، وفيما يلي وصف لكل أثر وكيفية يمكن ضبطه.

١- أثر الملاحظ على الملاحظين: إذا لم تكن الملاحظة غير مرئية فممن المحتمل أن يؤثر وجود الملاحظ على أفراد العينة الواقعيين تحت الملاحظة. مثال ذلك أن مجرد دخول الملاحظ إلى قاعة الفصل للمرة الأولى سوف يستثير فضول الطلبة والمدرسين. وقد يترتب على عدم انتباههم للدرس ما يؤثر على عينة السلوك التي يجري ملاحظتها. وللقضاء على هذه الظاهرة أو للإقلال منها يمكن للملاحظ القيام بعدة زيارات لنفس الفصل قبل جمع البيانات، وبذلك يصبح وجوده في الفصل أمراً عادياً.

٢- التحيز الشخصي للملاحظ: وينتج هذا النوع من الخطأ في جمع البيانات نتيجة للصفات الشخصية للملاحظ. ويمكن القول بشكل عام أن أي ملاحظة يقوم بها شخص ما عرضة لبعض التحيز من جانب الملاحظ بسبب معتقداته وآرائه الشخصية، وقد يكون هذا صحيحاً إلا أن تصميم بعض البحوث يزيد من احتمال تحيز الملاحظ. ويجب البحث عن مصادر التحيز والقضاء عليه. ومن أهم مصادر التحيز الاتجاهات السلبية ضد بعض المؤسسات أو المعتقدات. مثال ذلك أن يجمع ملاحظ بيانات عن سلوك بعض سكان المناطق العشوائية وهو نفسه يشعر باتجاهات سلبية نحو هذه المناطق. وإذا أطلقت بعض المسميات على الأطفال الذين تم ملاحظتهم، مثل "متخلف عقلياً"، أو "بطيء التعلم"، فإن مثل هذه الأوصاف قد تكون مصدراً للتحيز في البيانات التي يجمعها الملاحظون.

٣- أخطاء التقديرات: يؤدي استخدام مقاييس التقدير في تسجيل الملاحظات إلى حدوث أخطاء تتأثر بطريقة الملاحظ في تسجيل السلوك. ومن أهم مصادر هذا النوع من الخطأ ما يلي:

أ- خطأ التساهل: وهي النزعة إلى إعطاء تقديرات مرتفعة لمعظم المشاركين، رغم اختلافهم بالنسبة للمتغير الذي يلاحظونه.

ب- خطأ النزعة المركزية: وهي النزعة إلى إعطاء تقديرات قريبة من النقطة الوسطى في مقياس التقدير المستخدم في تسجيل الملاحظات. ويحدث هذا الخطأ عندما يريد الملاحظ تجنب التقديرات المتطرفة.

ج- أثر الهالة: وهو انتشار صفة من صفات الفرد على صفاته الأخرى، ونزعة الملاحظ إلى التأثير بفكرة مسبقة عن الفرد الذي يلاحظه. وفي هذه الحالة تنتشر فكرة الملاحظ عن صفة من صفات من يلاحظه إلى جميع صفاته الأخرى. ولذلك إذا كان الملاحظ انطباعاً معيناً عن أحد الأفراد في بداية جلسات الملاحظة فقد ينتشر هذا الانطباع على الصفات الأخرى الذي يلاحظها في هذا الفرد.

٤- فشل الملاحظ في تسجيل بعض الاستجابات: أحياناً لا ينتبه الملاحظ إلى تسجيل بعض أنواع السلوك عند حدوثها، ويمكن أن يرجع مثل هذا الفشل إلى تحيز الملاحظ الذي قد يتسبب في تجاهل بعض الأحداث التي يجب عليه تسجيلها، أو إلى أن بعض أنواع السلوك تحدث متأنية مع سلوك آخر يلاحظه الملاحظ، أو قد يحدث السلوك بسرعة كبيرة يصعب معها تسجيله، أو قد يكون حدوث السلوك نادراً من غير الممكن ملاحظته.

٥- تشتت الملاحظ: بعد تدريب الملاحظين على الملاحظة وكيفية تسجيل الملاحظات يجب أن يبدأ جمع البيانات فوراً، لأن التأخر قد يفقد الملاحظ بعض المهارات التي اكتسبها. وكذلك عندما تستمر فترات الملاحظة لأكثر من أسبوع يجب على الباحث أن يعقد جلسات تنشيطية للملاحظين، لأنه في مثل هذه الحالات قد تشتت ذهن الملاحظ ولا يركز فيما اكتسبه أثناء التدريب، ويترتب على ذلك تغير في تعريف المفاهيم التي درّب عليها بالنسبة للسلوك الذي يلاحظه.

أساليب الملاحظة في البحوث الكيفية

توجد ثلاثة اختلافات رئيسية بين الملاحظة في البحوث الكيفية والملاحظة في البحوث الكمية. ويمكن تلخيص هذه الاختلافات الثلاثة فيما يلي:

١- لا يسعى الملاحظون في البحوث الكيفية إلى التمسك "بالموضوعية" عند

ملاحظتهم للظواهر التي يلاحظونها. فقد يضمنون مشاعرهم وخبراتهم الخاصة في تفسير ملاحظاتهم. فالباحث الكيفي رغم أنه قد يتجنب أن ينحاز لفكرة أو رأي معين يلاحظه، إلا أنه يدّول أن يظهر مشاعره وردود فعله الخاصة للظاهرة التي يلاحظها، ويعتقد أن هذا حق له وجزء قانوني وشرعي من بحثه، ومن المهم أن يضمن تقريره هذه المشاعر التي يعتبرها جزءا من البحث.

٢- الاختلاف الثاني بين الملاحظة في كل من البحوث الكمية والبحوث الكيفية أن محور وبؤرة الملاحظة في البحوث الكيفية نتيجة طبيعية ومنطقية تتبثق أثناء البحث. أما في البحوث الكمية فإن الباحث يضع فروضا أو أسئلة أو أهدافا مسبقة توجه عملية البحث. ولذلك فإن الملاحظين في البحوث الكيفية أحرار في تحويل انتباههم إلى ظواهر جديدة في أي وقت أثناء عملية البحث، كلما انبثقت أمام الباحث أسئلة بحث جديدة.

٣- الاختلاف الثالث أن مدى الملاحظات عادة ما يكون أوسع في البحوث الكيفية. أما الملاحظون في البحوث الكمية فإنهم يميلون إلى التركيز على مظاهر محددة من السلوك، وفي سبيل ذلك قد لا يهتمون بالمضمون. ولذلك نجد الملاحظين في البحوث الكيفية ينظرون إلى السلوك ومحيطه البيئي من منظور شمولي كلي.

الغرض من الملاحظة في البحوث الكيفية

هناك وسيلتان لجمع البيانات في البحوث الكيفية وهما المقابلة وتحليل الوثائق وتتضمن هاتان الوسيلتان كلمات منطوقة أو مكتوبة بواسطة المشاركين في المواقف الطبيعية. وتتحدد هذه البيانات بمعلومات المشاركين وذاكرتهم وقد درتهم على توصيل المعلومات بوضوح ودقة، وكذلك برغبتهم في كيفية إدراك الأشخاص الخارجيين مثل الباحثين لهم. وتمكن الملاحظة الباحثين من تشكيل رؤيتهم الخاصة بما يحدث، بشكل مستقل عن المشاركين. وإضافة ملاحظات مختارة في تقرير الباحث يوفر وصفا أكثر اكتمالا للظواهر مما لو اقتصر الأمر على الإشارة لبعض العبارات التي وردت في المقابلات أو الوثائق. كما أن الملاحظات توفر مصادرا بديلة لتحقيق البيانات التي تم الحصول عليها بطرق أخرى.

اختيار الملاحظين و تحديد أدوارهم:

يختلف دور الملاحظ في البحوث الكيفية على طول متصل من الملاحظ الكامل إلى المشارك الكامل. وفي أحد طرفي هذا المتصل يكون الباحث الذي يقوم بدور الملاحظ الكامل منفصلاً تماماً عن الموقف الذي يلاحظه. وفي الطرف الآخر حيث المشارك الكامل يدرس الباحث موقفاً يكون فيه أحد أعضاء العينة أو يتدخل إلى عضو كامل خلال البحث.

ويوجد بين هذين الطرفين الملاحظ المشارك والمشارك الملاحظ. وفي حالة الملاحظ المشارك يقوم الباحث أساساً بدور الملاحظ، ولا يندمج للموقف إلا لجمع البيانات، ولا يتفاعل مع الأفراد أو المجموعات أثناء الملاحظة إلا بشكل عرضي غير مباشر. أما في دور المشارك الملاحظ يقوم الباحث بالملاحظة ويتفاعل تفاعلاً وثيقاً مع الأفراد ليكون شخصية ذات معنى داخل الجماعة، إلا أن الباحث لا يقوم بالنشاط الذي هو محور هوية الجماعة. مثال ذلك أن يدير أدر (Adler & Adler, 1991)، وهو عالم اجتماعي وأستاذ، لعب دور المشارك الملاحظ في دراسة عن الرياضيين في جامعتهم (جامعة كولومبيا). واحتفظ بهويته كباحث أثناء مشاركته النشطة مع فريق البيسبول، كما درب أعضاء هيئة التدريس على مختلف الأدوار بما فيها دور العضو والخبير ولكن لم يدرّبهم على الدور المحوري للمدرب الرياضي أو اللاعب.

وعندما يقوم الباحث بتصميم بحث كفي فلاحتمال الأكبر أنه يقوم بنفسه بدور الملاحظ أثناء مرحلة جمع البيانات. وقد يسند جمع البيانات إلى فريق، إلا أن الباحث يكون عضواً في هذا الفريق عند القيام بالملاحظات. واستخدام مجموعة من الملاحظين يخفف من العبء الواقع على الملاحظين ويسمح بوقت أطول للملاحظة. وإذا كان الملاحظون متنوعين من حيث العوامل المتعلقة بالظاهرة موضوع الدراسة (أي من حيث النوع والعمر وغير ذلك من العوامل) فإنه من الممكن زيادة صدق الملاحظات وذلك بمراجعتهم نتائج بعضهم البعض، واستبعاد التفسيرات غير الدقيقة.

الإعداد للملاحظة:

يمكن إعداد الملاحظ في مشروع بحث كفي بأن يتعلم على يد أحد الخبراء في نوع الملاحظة المطلوب للبحث. وقضاء فترة في هذه التلمذة شيء مرغوب فيه لأن مهارات الملاحظة الكيفية معقدة ودقيقة، ولا تصلح لها إلا إجراءات التدريب المستخدمة في البحوث الكمية التي وصفناها في جزء سابق من هذا الفصل. وعندما يعمل الملاحظ الجديد تحت التدريب إلى جانب خبير يستطيع أن يكتسب بالتدريج

فهما لكيفية بأورة ملاحظاته، وكيف ينتقل بين المراحل الثلاث (الوصف، والبأورة، والاختيار) التي سوف نناقشها فيما بعد.

وتقدم بعض الجامعات مقررات يمكن فيها تنمية مهارات الملاحظة المطلوبة في البحوث الكيفية، إلا أنها ليست بديلا عن التملذة. وتتضمن هذه المهارات القدرة على كتابة أوصاف الأحداث التي تلاحظ، ووصف تفاصيل الأحداث وتدوينها في مذكرات ميدانية، وتحويل المسودات المكتوبة بخط اليد إلى تقرير مطبوع أي ق. ويمكن أن تتضمن هذه المحاضرات كذلك تدريباً على استخدام الشرائط المسجلة السمعية والبصرية، وغيرها من أساليب العمل الميداني.

وعندما تبدأ دراسة كيفية فقد يصادف الباحث مشكلات وقضايا خاصة بالوضع الميداني للبحث. ومن المفيد في مثل هذه المواقف الاتصال بأحد الخبراء طلباً للمساعدة. وعملية الاستشارة هذه لا تنتهك تماسك الدراسة لأنه من المسموح به في الدراسات الكيفية أن تتغير طريقة الملاحظة كلما تعمق الباحث في الظاهرة التي يدرسها.

تحديد بؤرة الملاحظة:

كثيراً ما يتغير محور ملاحظات الباحث في البحوث الكيفية من المراحل الأولى للبحث إلى المراحل الأخيرة. ويذكر جيمس سبرادلي (Spradley, 1980) أن هذا التغير أو التحول يتضمن ثلاثة مراحل:

١- المرحلة الوصفية: وفيها تميل الملاحظة أن تكون عامة لا تتبأور على نقطة معينة، وتمتد هذه المرحلة الباحث بقاعدة يتفرع منها الملاحظون في اتجاهات متعددة.

٢- مرحلة التباور: وفيها يحدد الملاحظون ملامح الظاهرة التي يدرسونها والتي يهتمون بها أكثر من غيرها، ويبدأون في توجيه انتباههم إلى جمعة بيانات أعمق حول هذا المدى الأضيّق من الملامح.

٣- المرحلة المختارة: وتبرز في هذه المرحلة مشكلة البحث ويتحول محور البحث إلى تنقيّة وتعميق فهمهم للعناصر الأساسية التي بدرت كأهم العناصر من الناحية النظرية أو الأمبيريقية. وتجمع بيانات الملاحظة ويصل الباحثون إلى نقطة **التشبع النظري** وذلك عندما يرى الباحث أنه لم يعد يحصل على بيانات جديدة وأن البيانات التي يحصل عليها تكفي للبيانات السابق الحصول عليها.

وتتضمن مسألة محور البحث القرار بأي الأمور تلاحظ في وقت محدد، وكيف نتأكد أن كل شيء نهتم به أصبح في بؤرة الاهتمام. وحتى نتأكد من تغطية الظاهرة التي نهتم بها تغطية دقيقة يجب أن تتضمن المذكرات الميدانية إشارة صريحة إلى العناصر التالية (Gall et al., 2003):

- المشاركون.
- التفاعلات.
- الأشياء الروتينية.
- الطقوس.
- العناصر الزمنية.
- التفسيرات.
- التنظيم الاجتماعي للمشاركين.

ويؤدي تضيق وتعميق بؤرة الملاحظة إلى ظهور التفسيرات الكيفية والتي تشكل النتائج الأولية للدراسة.

الانتقال إلى الموقف الميداني:

كما هو متوقع من طبيعة الانبثاق التي تميز البحوث الكيفية لا توجد قواعد معينة لكيفية الانتقال إلى الميدان للقيام بالملاحظات. وهنا يقوم الباحث بوضع إجراءات قائمة على طبيعة الميدان وأعضائه وأين ينوي الباحث أن يحدد مكانه على طول متصل المشارك الكامل إلى الملاحظ الكامل.

ويمكن للباحث الحصول على معلومات عن تحديد إجراءات العمل في الميدان باستشارة خبراء البحوث الكيفية وقراءة تقارير بحوثهم. مثال ذلك أن إيفانز وإدرز (Evans & Eders, 1993, pp. 145-146) في دراسة حول المنع زلين اجتماعيا في المدارس المتوسطة وصفا إجراءاتهما للحصول على إذن بالدخول إلى قاعة الطعام في المدرسة كما يلي:

كانت استراتيجيتنا الدخول إلى المطعم كأقران وليس كرموز سلطة. لذلك ارتدبنا ملابس جينز غير رسمية تشبهها بمعظم الطلاب. ولم نحاول الاتصال بالمدرسين أو غيرهم من الكبار في الموقف. ومع أننا أخبرنا كل شخص بأننا كنا من الجامعة المحلية وأننا كنا معنيين بدراسة الصداقة بين المراهقين، إلا أننا لم نكتب مذكراتنا علنا. ولم نخبر الطلاب بأعمارنا رغم أن الفرق بين أعمار الطلاب وأعمار الباحثين يتراوح بين ١٠ و ٢٠ عاما. وكانت تعليقات الطلاب تشير

باستمرار إلى أنهم يعتبروننا طلابا في المرحلة الجامعية وأن أعمارنا تقترب من العشرين عاما.

ويظهر هذا الوصف أفكار الباحثين حول الحاجة إلى تقديم أنفسهم إلى المراهقين الذين يدرسونهم كشخصيات غير مهذبة وودودة. وإذا أراد الباحثون تغيير بؤرة البحث إلى المدرسين أو الإداريين في نفس المدرسة فإننا نتوقع منهم تعديل مظهرهم وسلوكهم إلى مظهر وسلوك آخر قريب من مظهر وسلوك العاملين بالمدرسة، مع الاحتفاظ بنفس الدرجة من الحياد والتحفظ مع هذه الجماعة كما كان وضعهم مع الطلبة، ليظهروا أحقيتهم أن يكونوا محل ثقة.

تسجيل الملاحظات:

يستخدم الباحثون الكيفيون نفس الأدوات التي يستخدمها الباحثون الكميون لعمل سجلات دائمة بملاحظاتهم. مثال ذلك يمكنهم عمل مذكرات ميدانية مكتوبة باستخدام حاسب آلي دفتري. ومن الخيارات الأخرى إملأ مذكرات إلى جهة تسجيل صوتي باستخدام ميكروفون مغطى متصل بآلة تسجيل ولها حد زام تمكن الملاحظ من حملها على كتفه كما يحمل آلة التصوير. ويمكن للملاحظ التحدث في الميكروفون أثناء حدوث نشاط دون أن يسمعه الناس القريبون.

وفي بعض المواقف الميدانية قد يتشتت المشاركون أو قد يفتعلون عداوة الملاحظ بعض الأحداث المهمة إذا قام بكتابة ملاحظاته في جهاز حاسب آلي دفتري أو في مذكرة إلكترونية، وفي مثل هذه المواقف يمكن للملاحظ أخذ مذكراته بطريقة سرية مثل كتابتها على غطاء مشط كبريت ليخفي دوره كملاحظ.

وإذا لم يتمكن الملاحظ من تسجيل ملاحظاته في الميدان فملا عليه إلا أن يحاول تذكر الأحداث المهمة التي رآها عقب انتهاء فترة الملاحظة مباشرة. وإذا لم يتمكن من كتابة ملاحظاته تفصيلا يمكنه كتابة ملخص بتتابع الأحداث مع التركيز على أهمها أثناء الملاحظة. ويمكن استخدام هذا الملخص فيما بعد كمثير لكتابة مذكرات أكثر تفصيلا.

ويجب أن تكون المذكرات الميدانية التي يكتبها الملاحظون تفصيلية وواضحة وغير غامضة، ولا تميل إلى العمومية المفرطة. ويذكر جول وزملائه (Gall et al., 2003, pp. 272-273) المثال الموضح في الجدول (٢٠-١) لتدوين المذكرات الميدانية:

ويلحظ أن الجزء المكتوب على اليسار أكثر وضوحا وتحديدا من الجزء المكتوب

على اليمين. وتظهر اللغة التي تكلمت بها العميلة أن الملاحظة لم تكن ظاهرة لها، وهذا النوع من الملاحظة أفضل من النوع المكشوف الذي يحتاج إليه الملاحظ أحيانا عندما يريد أن يبحث عن أنماط وأفكار في موقف الملاحظ أو أي م صادر أذ رى للبيانات.

جدول ٢٠-١ مثال لتدوين المذكرات الميدانية

مذكرات غامضة ومفرطة في العمومية	مذكرات تفصيلية ومحددة
كانت العميلة عدوانية نحو الموظف.	عندما أخبرت جودي، رئيسة الموظفين، العميلة أنها لا تستطيع عمل ما تريد، بدأت العميلة في الصراخ في جودي وقالت لها أنها لا تستطيع أن تسير حياتها كما تريد، وما عليها إلا أن تذهب إلى الجحيم، وهزت رأسها في وجه جودي واندفعت خارجة من الغرفة. تاركة جودي في حالة شديدة من الذهول والدهشة.

ويجب أن تتضمن مذكرات الملاحظة تفاصيل بصرية عندما يكون ذلك مناسباً، ولا يجب قصر هذه المذكرات على الكلمات المكتوبة. مثال ذلك أنه من الممكن أن يرسم الملاحظ تخطيطاً للمكان الذي تمت فيه ملاحظة الأحداث والأنشطة التي تمت في الموقف. وقد يحتاج الملاحظ في بعض الأحيان تصوير المكان باستخدام التسجيل المرئي إذا كان الموقف يستحق ذلك.

التعامل مع آثار الملاحظ:

يسلم البحث الكمي أن الملاحظة مستقلة عن الفرد الذي يقوم بالملاحظة. ولذلك يوجه الجهد نحو الإقلال من تحيز الملاحظ وضبط الآثار المحتملة له على من يلاحظ. أما الباحثون الكيفيون فيتخذون موقفاً مغايراً تجاه نفوذهم على الظاهرة التي يلاحظونها. فهم كملاحظين يعتبرون أن تحيزهم وردود أفعالهم الشخصية جزء من موقف الملاحظة. ولذلك فهم لا يستخدمون المحكات الموضوعية للحكم على مدى ارتفاع مستوى ملاحظاتهم، وبدلاً من ذلك يستخدمون الإجراءات التالية لمعالجة قضايا الصدق التي تظهر في الملاحظة الكيفية.

١- استجابات المشاركين والموظفين في البرنامج لوجود الملاحظ: الإجراء المتبع هو أن يعمل الملاحظون الكيفيون على عدم المبالغة في آثارهم على ما يلاحظون، كما يجب عدم الإقلال من هذه الآثار، ولكن يجب وصف وتحليل هذه الآثار كجزء من مشروع البحث.

٢- تأثير الملاحظ خلال الدراسة: الإجراء المتبع للتعامل مع هذه القضية يشبه

الإجراء المقترح للقضية السابقة، أي أن لدى الملاحظ ردود فعل وسدوف يسجلها.

٣- توجهات الملاحظ وتحيزاته الشخصية: يجب أن يتعامل الملاحظ مع هذه القضية عن طريق استخدام الإجراءات المتبعة لتحقيق صدق البيانات وتحليلها، حتى يقلل من المشتتات التي ربما أضافتها توجهات الملاحظ. ومن بين هذه الإجراءات الجهود الفعالة لاختبار التفسيرات المنافسة لنتائج البحث، واستخدام طرق البحث الكمي والكيفي لاختبار الظاهرة التي يدرسونها، واستخدام ملاحظين وباحثين متعددين، وفحص النتائج من أكثر من منظور نظري، وكتابة تقرير مشروع البحث بشكل مفصل تفصيلا كافيا بحيث يمكن للقارئ التدقيق في النتائج التي يقرأها.

٤- عدم كفاءة الملاحظ: تكون بيانات الملاحظات الكيفية عديمة الفائدة إذا لم يعد الملاحظون إعدادا كافيا للقيام بجمع البيانات التي تتطلبها مشكلة البحث ومنهجه. وحل هذه المشكلة واضح، وهو تدريب الملاحظين تدريباً دقيقاً وإعدادهم إعداداً كافياً قبل جمع البيانات.

تحليل البيانات الكيفية:

يستكمل البحث الكيفي عند انتهاء مرحلة جمع البيانات. ويكون لدى الباحثين كميات كبيرة من المذكرات والبيانات السمعية والبصرية التي تصلح سجلاً لملاحظاتهم. ويجب تحليل كل هذه البيانات وتفسيرها وكتابة تقرير عنها. وإجراءات تحليل هذه البيانات هي نفس إجراءات تحليل البيانات التي نحصل عليها من بعض الوسائل مثل المقابلة، والوثائق التي نجدها في ميدان البحث.

تحليل المحتوى

ناقشنا في هذا الفصل طرق ملاحظة السلوك الإنساني وبعض عناصر البيئة التي يحدث فيها السلوك. ومن أهم خصائص البيانات البشرية الرسائل التي يتلقاها الناس بشتى الأشكال مثل:

الوثائق المكتوبة: تتضمن المواد المكتوبة الكتب الدراسية والواجبات التي يؤديها الطلبة في المنازل، والاختبارات، ومخرجات الحاسب الآلي من مواد مطبوعة، والصحف والمجلات، وغيرها من المواد المكتوبة.

الوسائط البصرية: الصور الفوتوغرافية، والملصقات، والرسوم أمثلة للمواد البصرية التي يمكن تحليلها.

الوسائط السمعية: يستطيع الباحث أن يحلل التسجيلات الصوتية، وأسطوانات الليزر المدمجة، وكذلك المواد المذاعة.

الوسائط المدمجة: قد يكون من المرغوب فيه تحليل أنواع متعددة من الوسائط مثل برامج التلفزيون، والأسطوانات المدمجة التي تجمع بين الصوت والصورة (أسطوانات الفيديو).

وتتكون هذه الوسائط من مواد ورسائل من فرد أو مجموعة من الأفراد إلى فرد آخر أو جماعة. ويمكن أن تكون هذه المواد موضوعا للدراسة في بعض مشاريع البحوث. وللكتب المدرسية أهمية خاصة حيث إنها تعبير عن المذاهب الدراسية. وعادة ما يطلق الباحثون على الكتب المدرسية وغيرها من المواد المكتوبة وثائق. وسوف نستخدم هذا المصطلح هنا كما نستخدم مصطلح وسائط الاتصال للإشارة إلى الوثائق والمواد الأخرى التي تتكون أساسا من رسائل صوتية وسمعية.

وإذا أراد الباحث أن تكون وسائط الاتصال مصدرا لبيانات بحثية، فإنه يحتاج إلى أن ينظم التعرف عليها وتحليلها. وسوف نناقش في الأجزاء التالية من هذا الفصل بعض الإجراءات المناسبة لكل من التحليل الكمي والتحليل الكيفي.

تحليل المحتوى في البحوث الكمية:

يتضمن تحليل الوثائق في البحوث الكمية تحليلا للمحتوى. ويعرف برلسون (Berlson, 1952, p. 18) تحليل المحتوى بأنه "أسلوب في البحث يصف

بشكل موضوعي منظم وكمي محتوى الاتصالات". ويمكن أن تكون المادة الخام لتحليل المحتوى أي نوع من الوثائق أو أي وسيط آخر من وسائط الاتصال.

ويتكون معظم تحليل المحتوى في التربية وعلم النفس من جمع بيانات عن مختلف مظاهر الرسائل التي تنتج عن وسائل الاتصال. ويتضمن هذا التحليل عادة تصنيفات بسيطة أو جدولة بيانات معينة. ويمكن أن ينتج عن تحليل محتوى الموضوعات الإنشائية للطلبة في المقررات اللغوية، حرصاً للأخطاء الإملائية والنحوية، بالإضافة إلى دراسة لتكرار أنواع مختلفة من الأخطاء. ويمكن استخدام هذه البيانات لمراجعة المقررات اللغوية، أو وضع برامج لغوية علاجية.

خطوات تحليل المحتوى:

سوف نتبع خطوات تحليل محتوى الوثائق التي اتبعتها هارمون وهـ دريك وفوكس (Harmon, Hedrick & Fox, 2000) ووردت لدى جـ ول وزملائه (Gall et al., 2003). ورغم أن هذا التحليل ليس تخدم مواد مطبوعة فإن الإجراءات المذكورة تناسب تماماً تحليل الوسائط الأخرى مثل صفحات شبكة الإنترنت، والأفلام، والتسجيلات الصوتية:

١- تحديد أسئلة البحث أو فروضه أو أهدافه: انبثقت دراسة الباحثين من تقديرهم للدور المهم الذي تلعبه التعليمات اللفظية باستخدام المفردات في قدرة الطلاب على فهم مفاهيم الدراسات الاجتماعية التي تقدم في الصفوف العليا من المرحلة الابتدائية و صفوف المرحلة الإعدادية. وكان الباحثون على وعي كذلك بفاعلية الممارسات القائمة على البحث في زيادة نمو المفردات وقد تساءلوا عما إذا كانت هذه الممارسات ممثلة في كتب الدراسات الاجتماعية التي يستخدمها المدرسون. وقد أدى هذا الاهتمام من هارمون وهـ دريك وفوكس إلى صياغة ثلاثة تساؤلات بحثية توجه دراستهم:

أ- ما طبيعة الكلمات أو المصطلحات الأساسية التي تختارها الكتب المقررة للدراسات الاجتماعية؟

ب- إلى أي حد يتم تمثيل هذه المفردات في كل صف من الصفوف من الرابع إلى الثامن وكيف؟

ج- ما هي المعينات التي يقدمها الناشرون للمدرسين بشأن هذه المفردات؟

وتحتوي الكتب المقررة وبخاصة كتب المعلم على كثير من العناصر المرتبطة بالمحتوى وبصياغة أسئلة البحث المبينة من قبل، حدد الباحثون مجال

بحثهم كما وضعوا له محورا واضحا حتى يمكن القيام بدراسة قابلة للبحث.

٢- اختيار عينة من الوثائق لتحليلها: اختار الباحثون طبعات كتب المعلم لمقرر الدراسات الاجتماعية عام ١٩٩٧-١٩٩٨ التي تبنتها ولاية تكساس للصفوف من الرابع إلى الثامن. وقد اختيرت كتب المعلم لأنها ما تحتوي على الكتب المقررة على الطلبة، وتوصيات للمعلمين بكيفية تدريس الكتاب المقرر. وكان من الضروري الحصول على الوثائق التي تحتوي على هذه التوصيات حتى يمكن معالجة السؤال الثالث من الأسئلة السابقة.

٣- إجراء عملية لترميز الفئات: إن جوهر تحليل المحتوى هو وضع رموز للفئات التي تقسم إليها الوثائق. ولابد أن تمثل كل فئة متغيرا منفصلا يرتبط بأهداف البحث. ويجب أن تكون هذه الفئات حصرية بحيث يصنف كل جزء من المعلومات في الفئة الخاصة به.

ومن الأفضل استخدام نظام ترميز سبق استخدامه في بحث سابق إذا كان يتناسب مع طبيعة البحث الذي تقوم به، لأن هذا الأمر يوفر عليك الجهد والوقت الذي تبذله في وضع نظام جديد لنفسك. كما أن استخدام ترميز قياسي للفئات يسمح بمقارنة نتائج البحث بنتائج بحوث أخرى تستخدم نفس النظام، مما يساعد على أن تضيف نتائج البحث معرفة وأساسا نظرية في مجال البحث.

وفي بحث المفردات الذي نعرضه هنا، وضع الباحثون نظام ترميز خاص بهم، إلا أنها تتفق مع الفئات القياسية للكتب المقررة، بما في ذلك كتب المعلم. ومن هذه الفئات: أهداف التدريس، وعناصر الدرس، وبعض المقترحات بمهام يمكن للمدرسين والطلبة إكمالها لتأكيد التدريس والتعلم، والمصطلحات الأساسية (وهي عبارة عن كلمات أو عبارات أبرزها الناشرون)، مع فئات فرعية هي: (١) المصطلحات العامة التي توجد في أي مقرر (٢) المصطلحات الفنية ذات المعاني المتعددة (٣) الكلمات الفنية الخاصة بمجال الدراسات الاجتماعية (٤) المصطلحات الخاصة المرتبطة بأشخاص وأحداث وأماكن.

وبعد التحديد المبدئي لنظام تصنيف المحتوى، يجب على الباحث أن يحدد ما إذا كان من الممكن لمقربين متعددين تقدير استخدام هذا النظام بدرجة عالية من الثبات. ويمكن للباحث أن يحدد ذلك بحساب معامل الارتباط لحساب الثبات بين المقربين للتصنيفات المختلفة للثقتين. وإذا كان معامل الثبات بين المقربين منخفضا يجب تحديد نقاط الغموض في تصنيف

المحتوى وتوضيحها. ولزيادة الثبات الذي يمكن استخدامه مع نظام التصنيف من المفيد وضع مجموعة واضحة من قواعد تقدير الدرجات. ولم يذكر الباحثون الذين قاموا ببحث المفردات في تقريرهم معاملات ثبات. كما أن هناك إشارة متكررة لكلمة "نحن" في التقرير ولذلك فمن المحتمل أن الباحثين راجعوا بعضهم البعض باستخدام الرموز التي وضعوها أثناء المضي في عملية تحليل المحتوى. ومع ذلك فإن ثقتنا في النتائج كان من الممكن أن تقوى إذا حقق الباحثون الثبات بشكل واضح.

٤- إجراء تحليل المحتوى: يتكون تحليل المحتوى من عمل علامات تكرارية لكل فئة ترميزية في كل وثيقة في العينة. والإجراء العام هو عمل ملف في الحاسب الآلي يمكن عن طريقه استرجاع نص الوثيقة، (يمكن إدخال الوثيقة في الحاسب الآلي أو استخدام ماسح ضوئي ينقل الوثيقة مباشرة في ملف بالحاسب الآلي). وبعد ذلك نقرأ ملف الحاسب ونضع رمزا لكل عبارة أو فكرة تناسب فئة معينة من فئات تحليل المحتوى. ويستخدم بعد ذلك برنامج للحاسب لحساب تكرار كل رمز أو عمل قائمة بجميع أجزاء النص الذي يناسب رمزا معيناً.

ويمكن عرض جداول بالتكرارات في الجزء الخاص بالنتائج في تقرير البحث. ويمكن كذلك إعطاء الإحصاء الوصفي في التقرير، مثال ذلك متوسط العدد والانحراف المعياري للأحداث لكل فئة ترميزية وذلك في جميع عينة الوثائق. ويمكن كذلك تحليل العلاقات بين المتغيرات التي تمثلها الفئات المختلفة.

وفي بحث المفردات استخدم الباحثون العديد من هذه الأساليب الإحصائية. مثال ذلك أنهم حسبوا متوسط النسبة المئوية للمصطلحات الأساسية التي تناسب كل فئة من فئاتهم على طول الكتب المقررة التي كانت جزءاً من تحليل المحتوى. وقد وجدوا أن ٧٨٪ من المصطلحات الأساسية كانت في فئة الكلمات الخاصة بالمجال، وكان هناك ١٣٪ في فئة المصطلحات المرتبطة بأشخاص معينين، وأماكن وأحداث محددة. وفي تحليل مرتبط بهذا التحليل طلب الباحثون ثلاثة مدرسين للتعرف على ما يعتبرونه الكلمات الأساسية في العينات الممثلة للمادة في الكتب المقررة (وقد أعيد كتابة العينات على الآلات الكاتبة حتى لا يستطيع المدرسون التعرف على أي المصطلحات أشر عليها الناشر على أنها مصطلحات أساسية. ومن المدهش أن نسبة الاتفاق بين الناشرين والمدرسين لم تتعد ٤٨٪.

٥- تفسير النتائج: الخطوة الأخير في تحليل المحتوى هي تفسير معنى النتائج. وتتوقف عملية التفسير على الغرض من الدراسة وإطارها النظري. وفي دراسة المفردات التي نحن بصدد استخدامها استخدم هارمون وهديك وفوكس معلومات البحث عن مفردات التدريس الفعالة كأساس لتقويم ممارسات الناشرين كما كشف تحليل المحتوى. وتوضح هذه المعلومات البحثية قيمة بعض الأنشطة التدريسية مثل فرز الكلمات، وعملاً على رائط للمعاني، والمنظمات الصورية، في مساعدة الطلبة على الربط بين معرفتهم السابقة والمصطلحات الجديدة. ولم يجد هارمون وزملاؤه أي استخدام لهذه المعلومات البحثية في الكتب المقررة التي حللوها، إذ يقولون:

أشارت نتائجنا إلى أنه بالرغم من أن الناشرين اهتموا للمفردات في برامجهم، إلا أن أنشطة كثيرة ما زالت مؤسسة في أنشطة تدريس وتعلم المفردات التي لا تساندها شواهد أمبيرقية. ولذلك يجب أن يتخذ الناشر موقفًا عدوانيًا للعمل على تكامل المعرفة الحالية عن المفردات في إجراءاتهم التدريسية.

ويشير الباحثون أيضًا سؤالاً هاماً للبحوث المستقبلية بناءً على تحليل بيانات المدرسين والناشرين المتعلقة بالمصطلحات الرئيسية في الكتب المقررة للدراسات الاجتماعية. إذ يقولون أن الخلاف بين المدرسين والناشرين حول أي المصطلحات أكثر أهمية يثير تساؤلات حول كيفية اختيار الناشرين للمصطلحات الرئيسية. وهذا مثال جيد على المبدأ الذي يقول إن قيمة الدراسة ليست في نتائجها بل في الأسئلة التي تثيرها.

تحليل الوثائق والسجلات في الملاحظة الكيفية

كثيراً ما يدرس الباحثون الكيفيون وسائل الاتصال المكتوبة والتي توجد في المواقف الطبيعية. ولهذه الوسائل المكتوبة أنواع متعددة. ويمكن تعريف الوثائق بأنها وسائل اتصال مكتوبة تعد لأغراض شخصية أكثر منها لأغراض رسمية (Lincoln & Guba, 1985). أما السجلات فهي وسائل الاتصال المكتوبة ذات الأغراض الرسمية. ومن أمثلة الوثائق الخطابات الشخصية، والمذكرات الشخصية، ومسودات المقالات. بينما نجد أن من أمثلة السجلات العقود القانونية، وتقارير اللجان المتداولة، وتقديرات الضرائب، ومقالات الصحف.

وتعتمد الوثائق والسجلات اعتماداً أولياً على اللغة في توصيل المعاني التي

تريدها. ووسائل الاتصال الأخرى مثل الرياضيات والموسيقى ولافتات المروور تعتمد على نوع آخر من النظم الإشارية لإيصال المعنى. أي أنها تتعلق بالعلامات ودلالاتها.

وفي البحوث الكمية التقليدية يعتبر معنى النص المكتوب ثابتاً من قارئ لآخر ومن وقت لآخر. أي أن المعنى يوجد في النص نفسه، ويمكن تمثيل المعنى كمتغيرات ذات محتوى منفصل متميز يمكن دراستها باستخدام تحليل المحتوى. وعلى العكس من ذلك يعتقد الباحثون الكيفيون أن معنى النص يكمن في عقول كاتبه وقارئه. ولذلك يمكن أن يتغير معنى وثيقة ما أو سجل معين من قارئ لآخر، ومن فترة تاريخية لأخرى. كما يمكن أن يكون للوثيقة أو السجل معاني مختلفة في المستويات المختلفة للتحليل. مثال ذلك أنه يمكن تحليل الكتاب لتحديد الموضوعات التي يشتمل عليها. ويمكن كذلك أن نعتبره مصدراً ثانوياً مستمداً من مصدر أولي. ولذلك فإن العلاقة بين الكتاب المقرر ومصادره الأولية يمكن أن تكون مدوراً لدراسة كيفية.

ولذلك يجب على الباحث الكيفي لكي يفهم الوثيقة أو السجل دراسة المضمون الذي صدرت فيه، وغرض الكاتب من كتابته، والظروف التي يعمل فيها الكاتب، والقراء المقصودون والفاعلون، وغرض القراء من قراءته. ويجب على الباحث الكيفي كذلك أن يعرف أن قراءته نصاً ما تولد له معنى خاصاً به.

وتشكل الوثائق والسجلات التي كتبت في الماضي مشكلات خاصة للباحث الكيفي، لأنه من غير الممكن مقابلة الكاتب أو القراء. كما أن الباحث لا يستطيع أن يلاحظ المواقف التي استخدمت فيها وسائل الاتصال المكتوبة أو كيف أمكن توفيرها للقراء المختلفين. وقد وضع المؤرخون طرق بحث متنوعة لمعالجة هذه المشكلة، وقد سبق أن أشرنا إليها عند مناقشة البحوث التاريخية.

ويتبع الباحثون الكيفيون نفس الخطوات التي يتبعها الباحثون الكميون الذين يستخدمون النصوص وغيرها من وسائل الاتصال كمصادر للبيانات. إذ يبدأون بتحديد الوثائق والسجلات التي هي جزء من الموقف الذي ينوون دراسته. وبعد أن يتعرفون على المادة ينتقلون إلى الخطوة الثانية لتحديد أي هذه المواد ذات علاقة بالبحث. ثم يحددون كيف يجمعون هذه البيانات لتحليلها مع اتباع دليل السلوك الأخلاقي للبحث. وإذا لم يكن من الممكن إزالة المواد من الموقف الطبيعي، فقد يكون من الممكن تصويرها وأخذ صور للموقف الطبيعي التي وجدت فيه لتحليلها فيما بعد. وإذا لم يتمكن الباحث من تحقيق ذلك فما عليه إلا أن يفكر في طريقة

لتحليلها في الموقع. وأخيرا عليهم أن يحققوا صدق المواد. وهناك عدة طرق لتحقيق صدق الوثائق. وقد سبقت الإشارة إليها عند مناقشة النقد الخارجي والنقد الداخلي للوثائق في الفصل الخاص بالبحوث التاريخية.

واستخدام الوثائق والسجلات كمصادر للبيانات في البحوث الكيفية يختلف اختلافا كبيرا عن البحوث الكمية وغالبا ما يكون ذلك في مرحلة التحليل. ففي البحوث الكمية يتم تحديد مجموعة من المتغيرات ثم تطبق بنفس الطريقة على جميع وسائل الاتصال في العينة. كما تقاس المتغيرات بحيث تعطي بيانات كمية يمكن تحليلها بالأساليب الإحصائية التقليدية. أما في البحوث الكيفية فمن المحتمل أن تتبع إجراءات التحليل مع تقدم البحث. ويمكن تحليل نفس الوثيقة أو السجل أكثر من مرة خلال مراحل زمنية مختلفة من البحث، ويعطي كل تحليل تكوينات، وفروض وأفكارا جديدة. وليس من الضروري التعبير عن التحليل كمياً. كما أنه من الممكن تحليل الوثيقة أو السجل من منظور مختلف، ولتحقيق أغراض مختلفة.

وتأخذ نتائج التحليل الكيفي للباحث شكل التفسيرات والفروض. ويقتدر هودر (Hodder, 2000) أن يتم وزن هذه الفروض والتفسيرات باستخدام سياقين مختلفين: السياق الذي نشأت فيه الوثيقة أو السجل والسياق الذي يتم فيه الآن تفسيره لأغراض البحث. ويجب على الباحث أن يأخذ في حسابه التغيرات في المعنى التي تحدث عند دراسة الوثائق والسجلات عبر المكان والزمان والثقافات. ويقترح هودر خمس معايير لتأكيد التفسيرات المبنية على البيانات التي جمعت من الوثائق والسجلات:

- ١- التماسك الداخلي بمعنى أن الأجزاء المختلفة من الجاذب النظري لا تتعارض مع بعضها البعض كما يجب أن تكون الخلاصات نتيجة للمقدمات المنطقية.
- ٢- التماسك الخارجي بمعنى أن تناسب التفسيرات النظريات المقبولة داخليا وخارجا مجال الدراسة.
- ٣- التوافق بين النظرية والبيانات.
- ٤- أن تقوم النظرية على افتراضات صحيحة، تبين عدد التوجهات الجديدة للنظرية، ومسارات الاستقصاء التي تتحقق النظرية من خلالها.
- ٥- مصداقية الباحث ومؤهلاته العلمية ومنزلته.

الفصل الحادي والعشرون

الصفات الواجب توافرها في أدوات جمع البيانات

جمع البيانات عملية أساسية في أي بحث. ولذلك فإن أدوات القياس ضرورة لازمة مهما كان نوع البحث. وهناك صفتان أساسيتان لا بد من توافرها في أدوات جمع البيانات، وهما الصدق والثبات. وكلاهما ضروري لأي وسيلة قادرة على جمع البيانات بفاعلية. ويجب على كل مستخدم لوسائل جمع البيانات أن يكون ملماً بطرق تحقيق الصدق والثبات، حتى يستطيع تحقيق صدق وثبات الأدوات التي يبنيتها لجمع البيانات، بحيث تكون أدواته صالحة لقياس متغيرات البحث قياساً سليماً.

وللصدق والثبات أهمية خاصة في البحوث التربوية والنفسية، لأن القياس في هذين المجالين قياس غير مباشر. ولذلك يجب التأكد من أن ما نقيسه أدوات البحث يمكن الثقة فيه والاعتماد عليه في جمع البيانات.

الصدق

الصدق أهم خاصية من خواص القياس، ويشير مفهوم الصدق إلى الاستدلالات الخاصة التي نخرج بها من درجات المقياس من حيث مناسبتها ومعناها وفائدتها. وتحقيق صدق القياس معناه تجميع الأدلة التي تؤيد مثل هذه الاستدلالات (APA, 1985; AERA, 1999).

ولذلك يشير الصدق إلى مدى صلاحية استخدام درجات المقياس للقيام بتفسيرات معينة. فإذا كان المقياس أو الأداة اختباراً يستخدم لوصف تحصيل أفراد العينة يجب أن تفسر الدرجات على أنها تمثل مجال التحصيل الذي يقيسه الاختبار.

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

أما مصطلح الثبات فهو يشير إلى مدى اتساق نتائج المقياس، فإذا حصلنا على درجات مشابهة عند تطبيق نفس المقياس على نفس مجموعة الأفراد مرتين مختلفتين، فإننا نستدل من ذلك على أن نتائجنا لها درجة عالية من الثبات من إجراء لآخر. وبالمثل إذا قام عدد من جامعي البيانات بشكل مستقل بتقدير بعض الصفات في أفراد العينة باستخدام أداة معينة، وحصلوا على تقديرات متشابهة فإننا نستخلص من ذلك أن النتائج لها درجة عالية من الثبات من مقدر لآخر. وكما هو الحال في الصدق، فإن الثبات يرتبط ارتباطاً وثيقاً بنوع التفسيرات التي نقوم بها بالدرجات الاختبار أو أداة جمع البيانات. فقد نهتم مثلاً بالنسبة لبعض الاستخدامات بالسؤال عن مدى ثبات هذه النتائج عبر عينات مختلفة لنفس السلوك. إلا أننا في جميع الحالات التي نتطرق فيها إلى الثبات إنما نهتم بمدى اتساق النتائج أكثر من اهتمامنا بمدى صحة التفسيرات التي نستقيها من النتائج (الصدق).

وأحياناً ما تكون العلاقة بين الصدق والثبات محدودة للأشياء خاصاً الذين يصادفون هذين المصطلحين للمرة الأولى. فثبات المقياس لازم للحصول على نتائج صادقة، إلا أننا نستطيع الحصول على نتائج ثابتة وليست صادقة، فقد نحصل على قياس ثابت ومتسق إلا أنه يعطينا معلومات خاطئة ويفسر تفسيراً غير سليم. مثلاً ذلك إذا افترضنا أننا نستطيع قياس ذكاء الفرد من معرفة عرض جبهته، فإننا سوف نحصل على قياس ثابت لعرض الجبهة كلما أعدنا قياس جبهة نفس المجموعة من الأفراد، إلا أننا لا نستطيع الإدعاء بأن الفرد الأعرض جبهة أكثر ذكاءً، فهذا تفسير خاطئ لعملية القياس أي أنه تفسير غير صادق رغم ثباته. فالثبات إذن ضروري لأداة القياس ولكنه ليس كافياً لتحديد الصدق (رجاء أبوعلام، ١٩٨٧).

طبيعة الصدق:

١- رغم أننا ننسب الصدق إلى أدوات القياس إلا أن الصدق في الواقع يشير إلى الدرجة التي يمكن بها تفسير نتائج المقياس، أي أننا في الواقع نتكلم عن صدق التفسيرات التي نخرج بها من النتائج.

٢- الصدق نسبي، لأن هناك درجات للصدق، ولذلك لا يجب التفكير في نتائج القياس على أنها صادقة أو غير صادقة، ولكن من الأفضل التعبير عن الصدق في فئات تحدد درجته كأن نقول صدق مرتفع، أو صدق معتدل، أو صدق متوسط، أو صدق منخفض.

٣- يرتبط الصدق دائما بالاستخدام الخاص للمقياس، ولا يمكن أن تكون نتائج المقياس صادقة لجميع الأغراض، مثال ذلك أن النتائج التي نحصل عليها من اختبار في الحساب قد يكون لها درجة عالية من الصدق بالنسبة للدلالة على الفهم الحسابي، ودرجة متوسطة من الصدق للتنبؤ بالنجاح في مقررات مقبلة في الحساب. وليس صادقا بالمرّة فيما يتعلق بالتنبؤ بالنجاح في الموسم يقي. وعلى هذا فعندما نصف الصدق فمن الواجب أن نحدد الاستخدام الذي نقيده نتائج فيه. فنتائج القياس لا يمكن أن تكون صادقة بالمعنى المطلق للصدق، إذ أن لكل تفسير من التفسيرات درجات مختلفة من الصدق.

٤- الصدق مفهوم واحد (APA, 1985, p.9)، فليس هناك أنواع مختلفة من الصدق، وما اصطلح عليه في السابق بأنه أنواع للصدق ليس إلا وسائل أو طرقا لجمع الأدلة على هذا المفهوم الواحد. وهذه الطرق المتعددة هي:

♦ ما يرتبط بالمحتوى.

♦ ما يرتبط بمحك.

♦ ما يرتبط بتكوين الأداة.

وهذه الأقسام الثلاثة هي بمعنى آخر طرق لوصف الصدق، ورغم أن طريقة منها قد تكون أفضل من الأخرى في إعطاء دلائل على صدق استخدام معين، إلا أن استخدام جميع هذه الطرق يساعد على الوصول إلى أفضل تفسير ممكن لدرجات الأداة.

طرق تحديد الصدق:

تتعلق النقطة الأولى التي يجب مناقشتها عند الكلام على طرق تحديد الصدق بعملية القياس. كيف نعرف أن اختبارا ما اختبار صادق؟ إن هناك العديد من المشكلات التي تبرز عندما نحاول تحقيق صدق الاختبار، ويجب أن يكون واضح الاختبار واعيا بهذه المشكلات حتى يستطيع أن يعالجها. والمثال التالي يمكن أن يوضح ذلك: لنفرض أننا نريد بناء اختبار يقيس صحة الضمير، كمتغير له أهمية في دراسة الشخصية، وفي بداية الأمر يبدو من غير الواضح كيفية تحقيق صدق مثل هذا الاختبار. وقد يكون الأمر سهلا إذا وجدنا اختبارا مستقلا يقيس

صحة الضمير، ولكن إذا كان ذلك ممكنا فما الداعي لبناء مثل هذا الاختبار. إن الأمر لن يتطلب أكثر من الحصول على هذا الاختبار واستخدامه. ولكن الأمر الأكثر احتمالا هو أننا لن نجد مثل هذا الاختبار ولذلك فإننا نحتاج إلى اللجوء إلى مصدر آخر للحصول على أدلة لصدق الاختبار الذي نحن بصدده. وهنا قد نجد أن من الأمور الشائعة اهتمام الناس بتقويم بعضهم البعض باعتباره يمتلكون أو لا يمتلكون ضميرا حيا، وهذا أمر وارد في الحياة اليومية، إذ يقوم الكثيرون بتقويم غيرهم من حيث صحة الضمير، فإذا استطعنا الحصول على تقديرات بعض الأفراد الذين يعرفون أفراد العينة التي نريد تحقيق الصدق عليها معرفة جيدة، فإن هذا قد يحل المشكلة، حيث إنه يمكننا في هذه الحالة إجراء معامل ارتباط بين تقديرات المحكمين ودرجات أفراد العينة في الاختبار الذي وضعناه. إلا أن هذا الأمر يسلم بأن التقديرات التي حصلنا عليها تقديرات صادقة، وهو ما يسلم به صعب إعطاؤه. ويدخل بنا في متاهة جديدة.

هذا المثال ليس إلا وسيلة لتوضيح مدى صعوبة تحقيق صدق الاختبارات في المجالين النفسي والتربوي. فتحقيق الصدق ليس عملية سهلة مباشرة، بل يحتاج من الباحث أو واضع الاختبار البحث عن مصادر لتحقيق صدق الأداة التي وضعها. وهي مصادر لا تكون دائما سهلة المنال. وهذا بعكس الثبات، حيث نجد أن تحقيق الثبات هو بالضرورة أمر فني، يمكن إجراؤه بشكل مباشر، وبأس تخادم بعض العمليات السهلة والتي سوف نناقشها بعد الانتهاء من الكلام على موضوع الصدق.

ويمكن القول إن هناك ثلاث طرق لتحديد الصدق (الجدول ٢١-١). ويصل الصدق إلى أقوى درجاته عندما يكون لدينا أدلة على الصدق بالطرق الثلاث، أي أن تفسير درجات المقياس يمكن أن يكون ذا صدق أكبر إذا كان لدينا معلومات كافية عما يلي:

- ١- محتوى المقياس والمواصفات التي بني على أساسها.
- ٢- علاقة درجات المقياس بدرجات مقاييس أخرى مهمة.
- ٣- طبيعة الخاصية التي نقيسها.

وسوف نتناول فيما يلي الطرق الثلاث لتحقيق صدق المقاييس:

- ١- **الطريقة المرتبطة بالمحتوى:** أدلة الصدق المرتبطة بالمحتوى مهمة بوجه خاص عندما نريد أن نصف كيفية أداء الفرد في مجال من المجالات التي يمثلها المقياس. مثال ذلك إذا كنا نتوقع أن يعرف التلاميذ كيفية كتابة ٢٠٠ كلمة، فلا يمكن لنا أن نتوقع عمل اختبار إملاء مكون من ٢٠٠ كلمة، نظرا للوقت الذي

يستهلكه هذا الاختبار، ولذلك فإننا نختار عينة مكونة من ٢٠ كلمة مثلاً، لتمثيل المجال كله، والمكون من ٢٠٠ كلمة. فإذا كتب أحد التلاميذ ٨٠٪ من الكلمات المملأة عليه بشكل صحيح، فإننا نود القول إنه يستطيع كتابة ٨٠٪ من كلمات المجال (٢٠٠ كلمة) كتابة صحيحة. أي أننا نريد أن نعمم من أداء التلميذ في الاختبار المكون من ٢٠ كلمة على المجال الذي يمثل الاختبار. ومثل هذا التفسير يكون صادقاً بالدرجة التي يكون بها الاختبار المكون من ٢٠ كلمة ممثلاً تمثيلاً سليماً للمجال المكون من ٢٠٠ كلمة.

ويمكن على هذا الأساس القول أن الطريقة المرتبطة بالمحتوى تعتمد على مدى تمثيل مفردات المقياس تمثيلاً سليماً للمجال الذي نريد قياسه، ولذلك فإن تحقيق صدق المقياس بطريقة المحتوى يتطلب القيام بخطوتين هما:

- ١- تحديد المجال الذي نريد قياسه تحديداً واضحاً، مع تحديد عناصره.
- ٢- بناء مجموعة ممثلة من المفردات أو الأسئلة لهذا المجال.

جدول ٢١-١ طرق تحديد الصدق

الطريقة	الإجراء	التفسير
المرتبطة بالمحتوى	مقارنة مفردات أو أسئلة المقياس بالمواصفات التي تحدد المجال الذي نريد قياسه.	مدى تمثيل مفردات المقياس للمجال الذي يقيسه.
المرتبطة بمحك	مقارنة درجات المقياس بدرجات مقياس آخر (المحك) نحصل عليه فيما بعد (الصدق التنبؤي)، أو بدرجات مقياس آخر يطبق في نفس الوقت (الصدق التلازمي).	الدرجة التي يتنبأ بها الاختبار بالأداء في المستقبل أو يقدر الأداء الحالي في مقياس هام غير المقياس نفسه، ويطلق على المقياس الثاني المحك.
المرتبطة بالتكوين	تحديد معنى درجات المقياس وذلك بدراسة تكوين المقياس والتحديد الميداني العملي للعوامل التي تؤثر في الأداء.	الدرجة التي يمكن بها تفسير الأداء كمقياس له معنى خاص أو صفة محددة.

أي أن تحديد الصدق بالطريقة المرتبطة بالمحتوى تعكس مدى تمثيل مفردات المقياس للمجال الذي نريد قياسه. فإذا أردنا مثلاً إعداد اختبار تحصيلي في اللغة العربية، فإننا لتحقيق الصدق بالطريقة المرتبطة بالمحتوى لابد لنا من تحديد المهارات اللغوية والمعرفة والفهم اللازم، وبمعنى آخر تحديد أهداف

تدريس اللغة العربية، ثم نقوم بتحديد محتوى منهج اللغة العربية، ونطابق أسئلة الاختبار مع الأهداف والمحتوى لنرى مدى تمثيل محتوى الاختبار لمحتوى المنهج وأهدافه تمثيلاً صادقاً.

وهذه الطريقة في تحديد الصدق مهمة بوجه خاص في الاختبارات التحصيلية، بل إنها تعتبر أساسية بالنسبة لهذا النوع من المقاييس، نظراً لأن تفسير درجات الاختبار ينعكس على المقرر بأكمله، أو الوحدة التي يقيسها الاختبار. وهذا يعني أن الذي يحصل على درجة منخفضة سوف تفسر درجته على أساس أنه لم يحقق إلا جزءاً محدوداً من مخرجات التعلم لهذا المقرر أو تلك الوحدة.

ولا يعبر عن الأدلة على الصدق المرتبط بالمحتوى في صيغة رقمية عادة، بعكس الحال في الطرق الأخرى لتحديد الصدق. فأساس الأدلة هو الحكم على محتوى الاختبار من الأسئلة أو المفردات مقارنة بمحتوى المجال الذي تغطيها الوحدة، وبالأهداف أو مخرجات التعلم التي وضع الاختبار لقياسها. وهذا يتطلب تحليلاً دقيقاً لأسئلة الاختبار أو مفرداته لتحديد ما إذا كانت فعلاً تمثل محتوى المقرر وأهدافه. ومن أهم مصادر تلك الأدلة المدرسون الذين يدرسون المقرر أو الوحدة التي يقيسها الاختبار أو خبراء المناهج. ولذلك يجب على واضع الاختبار استطلاع رأي المدرسين والخبراء في أسئلة الاختبار، ومدى تمثيلها لمحتوى المقرر وأهدافه. وإذا وافقت نسبة لا تقل عن ٩٠٪ من المحكمين على أن الاختبار عينة ممثلة لعناصر المقرر وأهدافه، يمكن اعتبار محتوى الاختبار صادقاً. ويجب في مثل هذه الحالة بناء جدول مواصفات الاختبار، وهو جدول ذو بعدين يحددان توزيع مفردات الاختبار بالنسبة لأهدافه ولعناصر المقرر الذي وضع له الاختبار (رجاء أبو علام، ١٩٨٧).

ويحتاج الأمر بعد تطبيق الاختبار تطبيقاً أولياً إلى تحليل مفردات الاختبار للحصول على مؤشرات صعوبة المفردات (نسبة أفراد العينة الذين أجابوا إجابة صحيحة على كل مفردة)، ومؤشرات صدق المفردات (معامل ارتباط المفردة بالدرجة الكلية المصححة للاختبار)، وبناء على هذه النتائج يمكن تعديل المفردات التي تحتاج إلى تعديل أو حذف المفردات الضعيفة غير القابلة للتعديل (رجاء أبو علام، ٢٠٠٣). ومعنى ذلك أن تحقيق صدق المحتوى يتطلب جانبين: الأول يقوم على بناء جدول المواصفات للحصول على عينة ممثلة من المفردات للمجال الذي نضع له الاختبار. ويقوم الثاني على الحصول على بيانات عن استجابات أفراد العينة للمفردات وللاختبار ككل. وهاتان الطريقتان

أساسيتان للحصول على اختبار يتصف بدرجة عالية من صدق المحتوى ويمكن الاطمئنان إلى أن درجات الطلبة تعكس فعلا مستواهم في المجال المراد قياسه.

٢- **الطريقة المرتبطة بمحك:** تستخدم في هذه الطريقة معامل الارتباط بين المقياس ومقياس آخر يطلق عليه "المحك"، وكلما كان معامل الارتباط مرتفعا (موجبا أو سالبا) كان معامل الصدق مرتفعا. والطريقة المرتبطة بمحك على نوعين:

أ- **الصدق التنبؤي:** ويقوم على أساس حساب القيمة التنبؤية للمقياس، أي معرفة مدى صحة التنبؤات التي تقوم بها من درجات المقياس. والتنبؤ هذا يقوم على التنبؤ من استجابات الفرد بالنسبة لسمة معينة (كالقدرة اللفظية، أو التحصيل الدراسي) بأدائه في مجال معين بعد فترة من الوقت قد تكون سنة أو عدة سنوات. أي أننا بقياس استجابات الفرد في مواقف معينة يمكن استنتاج كيف يسلك في مواقف مشابهة فيما بعد.

فإذا قسنا وظائف معينة، تبين 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 زمنية بين الإجراء وجمع البيانات عن مقياس موضوعي آخر (المحك) للنجاح في العمل أو الدراسة. أي أن علينا أن ننتظر إلى أن ينهي الطالب دراسته أو العامل تدريبه مثلا، ثم نقيس درجة الارتباط بين درجات الفرد في الاختبار ودرجاته في المحك، ويكون معامل الارتباط مرتفعا كلما كانت القيمة التنبؤية للاختبار عالية، مما يدل على أن الصدق التنبؤي للاختبار مرتفع.

ب- **الصدق التلازمي:** تعني هذه الطريقة في تحقيق الصدق بمدى ارتباط درجات المقياس بالدرجات أو البيانات التي تجمع من محك آخر وقت إجراء الاختبار. أي أننا نقارن بين درجات الأفراد في المقياس ودرجاتهم على مقياس موضوعي آخر يقيس نفس السمة تقريبا التي يقيسها المقياس، وذلك بأن يعطي الأفراد درجات على المحك في نفس الوقت الذي نطبق فيه المقياس أو قبلها بقليل. فإذا كان لدينا اختبار ذكاء فإننا نختار المحك من بين اختبارات الذكاء المقننة الثابتة الصادقة وهكذا.

الشروط الواجب توافرها في المحك: من أهم الشروط الواجب توافرها في

المحك:

- ١- أن يكون المحك متعلقا بالصفة أو الخاصية التي يقيسها المقياس.
- ٢- أن يكون المحك صادقا.
- ٣- أن يكون المحك ثابتا.
- ٤- يتصف المحك الجيد بأنه عملي اقتصادي. بمعنى أنه يجب اختيار محك يسهل استخدامه لجمع البيانات بحيث يوفر الوقت والجهد والمال.
- ٥- أن يكون المحك موضوعيا لا تتأثر درجات الأفراد فيه بشخص من يقوم بإعطاء الدرجة أو التقدير.
- ٦- أن يكون المحك مستقلا فلا تتأثر درجة الفرد في المقياس بدرجة في المحك. وألا تتأثر رتبة الفرد في المحك بدرجة أو رتبته في المقياس. فلا يجوز مثلا أن يطلب باحث من مدرس أن يقدر تحصيل التلميذ في المادة بعد أن نحيطه علما بأنه حصل على درجات مرتفعة في الاختبار التحصيلي.

المشكلات التي تنشأ عن عدم وجود محك: الغالبية العظمى من الاختبارات النفسية والتربوية التي يضعها الباحثون تقيس متغيرات لا يوجد لها مذكور صدق مقبول. وقد يتطلب الأمر لتحقيق الصدق التلازمي في هذه الحالة اختيار محك مشكوك في صدقه بالنسبة للمتغير الذي نرغب في قياسه، وهذا موقف يجب تجنبه على قدر الإمكان. إلا أنه إذا اضطر الباحث إلى استخدام مثل هذه المحكات فيجب أن يستخدم معاملات الارتباط التي يحصل عليها كجزء من تحقيق صدق التكوين الذي سوف نناقشه بعد قليل.

ويمكن القول إن قيمة الصدق التلازمي تكمن في الحصول على محكات ذات صدق مقبول، وإلا فإن البيانات التي حصلنا عليها لتحقيق الصدق التلازمي يجب اعتبارها مظاهر لصدق التكوين.

المشكلات التي تنشأ عن وجود محك مقبول: لعل المجال الوحيد الذي توجد به اختبارات يمكن استخدامها كمحكات تعطينا ارتباطات عالية تكون مؤشرا جيدا بصدق الاختبار هو مجال الذكاء (Kline, 2000). حيث نجد دوافع من الاختبارات الفردية والجمعية التي يمكن استخدامها كمحكات لتحقيق صدق الاختبار الذي وضعناه. ولكن حتى في مثل هذه الحالة فإن كثيرا من الاختبارات الجيدة التي قد نجدها هي اختبارات مضي عليها فترة طويلة من الزمن ويحتاج

الأمر إلى إعادة حساب صدقها وثباتها، وهو أمر قد لا يكون متيسراً للباحث دائماً. وحتى مع وجود محك جيد يمكن اعتباره نموذجاً للمحكات المقبولة، فلا بد من وجود اختلافات بينه وبين الاختبار الذي وضعناه، وهذا في حد ذاته يقلل من قيمة المحك. وإذا كان الاختبار مشابهاً تماماً في محتواه للمحك فليس هناك داع أصلاً لاستخدام اختبار جديد ويمكن استخدام المحك في جميع البيانات التي نريدها.

٣- الطريقة المرتبطة بالتكوين: كان واضحاً من المناقشة السابقة حول محكات الاختبارات أنه يوجد عدد كبير من الاختبارات لا يناسبها الطرق التي ذكرناها حتى الآن في تحقيق صدق الاختبارات. ولذلك فكر كل من كرونباك وميدل (Cronback & Meehl, 1955) فيما أطلقوا عليه "صدق التكوين". ومصطلح "تكوين" مرادف تماماً لمصطلح "مفهوم". إلا أن "التكوينات" في عملية البحث هي التي تخضع للاستقصاء العلمي، ولكنها لا تكون مفيدة إلا إذا أمكن تعريفها تعريفاً دقيقاً (راجع الفصل الثاني).

ويعتمد تحقيق هذه الطريقة على الارتباط بين الجوانب التي يقيسها المقياس وبين نظرية معينة أو فروض تتعلق بهذه الجوانب، أي أننا عند تحديد صدق التكوين نقوم بتحديد ما نقصده بمصطلح يصف الجانب أو الجوانب التي نقيسها. ثم نفحص درجات الأفراد ونبين كيف نفسر هذه الدرجات باستخدام الجانب المقاس. فنبحث عما إذا كانت الفروق بين درجات الأفراد في المقياس ترجع إلى الفروق بينهم في درجة الخاصية التي نقيسها. فإذا كنا نقيس التحصيل في مقرر ما كانت الفروق بين درجات مجموعتين من الأشخاص ترجع إلى درجة إتقان كل منهما لأهداف التحصيل في المقرر.

فواضع المقياس صاحب نظرية أو أنه صممه على أساس نظرية معينة تتعلق بجوانب افترض أن المقياس يقيسها. وهو لذلك يحدد ما إذا كان المقياس يمثل هذه النظرية، وما إذا كانت الفروق بين الدرجات يمكن تفسيرها على أساس الفروق بين مستويات القدرة أو السمة التي تعالجها النظرية وقيسها الاختبار. أي أن مصمم المقياس يحاول أن يحقق صحة النظرية التي اعتمدها عليها في تصميم المقياس، وذلك بأن يحدد التنبؤات والتوقعات التي يفترضها بفدص الفروق في الدرجات، مرجعاً إياها إلى النظرية التي صمم المقياس على أساسها، والتي غالباً ما تقضي بأن اختلاف الأفراد في أحد جوانب السلوك يعني تفاوتهم في قدرة أو سمة معينة.

وتعتمد البيانات التي تجمع من أجل تحقيق صدق التكوين على طرق منطقية وطرق أمبيريقية في نفس الوقت. وأحد مظاهر الطريقة المنطقية هو التأكد من أن العناصر التي يتكون منها المفهوم الذي نقيسه، هي نفسها العناصر التي يقيسها الاختبار. ويذكر آري (Ary et al., 1996, p.268) مثالا على ذلك أن دول Doll عندما أراد أن يضع مقياس فاينلاند للنضج الاجتماعي Vineland Social Maturity Scale عرف مفهوم 'النضج الاجتماعي' كتكوين فرضي بأنه عدد من العناصر المترابطة مثل المعونة الذاتية، والتوجيه الذاتي، والحركة، والمهنة، والتواصل، والعلاقات الاجتماعية. وهذه العناصر مظاهر للمفهوم يجب أن يشملها مقياس للنضج الاجتماعي. وأحيانا ما يحدث خلاف على المقصود بعناصر المفهوم. مثال ذلك إذا فكرنا في التفكير الفرضي 'الذكاء' على أنه مجموعة المهارات التي تمكن الفرد من التوافق مع البيئة الأكاديمية، فإننا نتوقع أن يشتمل اختبار الذكاء على مثل هذه المهارات. وإذا عرفنا الذكاء كمجموعة من المهارات التي لا ترتبط بالبيئة المدرسية أكثر مما ترتبط بالبيئات الأخرى، فإننا لن نهتم بوضع مهارات خاصة بالمدرسة في هذا الاختبار.

ومن المظاهر الأخرى للطريقة المنطقية هي فحص مفردات الاختبار، للتحقق من صلاحيتها لتقييم عناصر التكوين الفرضي المراد قياسه. ففي مقياس فاينلاند يسأل والد طفل في السادسة من عمره، إذا كان ابنه يستخدم الزلاقات، أو العربة (عمل)، ويذهب لسريه دون مساعدة (المعونة الذاتية)، ويكتب كلمات بسيطة (التواصل)، ويلعب ألعابا بسيطة (الحركة)، ويمكن الثقة في قدرته على التعامل بالنقود (توجيه الذات). ويبدو أن هذه الأسئلة مناسبة لقياس عناصر النضج الاجتماعي. وإذا تضمن الاختبار الأصلي أسئلة تتعلق بتفضيل الطفل لأطعمة معينة، أو عن اليد المفضلة لدى الطفل (اليمنى أم اليسرى)، فإن مثل هذه الأسئلة سوف تحذف، لأنها لا ترتبط بعناصر التفكير الفرضي.

وتجمع أيضا البيانات الأمبيريقية كأدلة على الطريقة المرتبطة بتكوين الاختبار. فمن الناحية الداخلية، يجب أن تكون الارتباطات الداخلية بين مفردات المقياس، كما هو متوقع من التكوين الفرضي. ومن الناحية الخارجية، يجب أن تكون معاملات الارتباط بين درجات المقياس وغيرها من الملاحظات الخارجية منسجمة مع المفهوم الذي يقيسه الاختبار. فإذا تكون الاختبار من

عدة مقاييس فرعية، يجب أن تكون معاملات الارتباط بين المقاييس الفرعية مرتفعة، وذلك كدليل على أن العناصر التي يتكون منها المفهوم عناصر مرتبطة فعلا. وإذا كانت هذه الارتباطات أقل مما هو متوقع فمعنى ذلك أن الاختبار لا يقيس التكوين الفرضي، أو أن تعريف هذا التكوين غير سليم.

ويجب أن ترتبط درجات الاختبار بالمقاييس الخارجية بطريقة تتسجم مع التكوين الفرضي الذي يقيسه الاختبار. فاختبار الذكاء الذي يركز على الجانب اللفظي يجب أن يكون ارتباطه عاليا بغيره من اختبارات الذكاء اللفظية، وأقل ارتباطا بالاختبارات غير اللفظية.

ويجب أن يكون مقياس مفهوم معين مستقلا عن غيره من مقاييس المفاهيم الأخرى على قدر الإمكان. فإذا وضعنا اختبارا لقياس القدرة على تطبيق قواعد النحو، ووجدنا أن هذا الاختبار ذو ارتباط مرتفع باختبارات الإملاء، وأقل ارتباطا باختبارات القواعد، فلنا أن نستنتج أن الاختبار الذي وضعه لا يقيس التكوين الفرضي المقصود، وعلينا تعديله، أو بناء اختبار آخر أكثر قدرة على قياس عناصر التكوين الفرضي الذي نريد قياسه. ومن ناحية أخرى إذا وضعنا اختبارين لقياس تطبيق قواعد النحو، ووجدنا أحدهما ذا ارتباط مرتفع باختبار الإملاء، بعكس الاختبار الثاني، فإننا في هذه الحالة نعتبر الاختبار الثاني أعلى صدقا من الاختبار الأول، لأنه أكثر استقلالا عن اختبار الإملاء من الاختبار الأول.

طرق تحديد الصدق المرتبط بالتكوين: ليس هناك طريقة واحدة لجمع أدلة على صدق التفسيرات الخاصة بتكوين الاختبار. ويمكن استخدام أية طريقة مادامت لها فائدتها في الحصول على أدلة على الصدق المرتبط بالتكوين. وفيما يلي بعض الوسائل التي يمكن استخدامها للحصول على أدلة من هذا النوع.

١- **الارتباط بالمقاييس الأخرى:** ذكر ميسيك (Messick, 1989) أنه يمكن استخدام درجة التقارب بين مؤشرات التكوين الفرضي، وكذلك القدرة على تمييزها عن التكوينات الفرضية الأخرى، كوسائل لتحقيق الصدق المرتبط بالتكوين. ويقصد 'بالتقارب بين المؤشرات' أن المقياس يرتبط بمقاييس أخرى يفترض أنها مؤشرات صادقة لنفس التكوين الفرضي. ويمكن الحصول على هذا التقارب بين المؤشرات، بالبحث عن المقاييس التي يجب أن ترتبط بالمفهوم من الناحية النظرية، ثم يبين كيف تكون هذه الارتباطات. فمقياس الاستدلال الرياضي يجب أن يرتبط بدرجات الرياضيات. فإذا كان الاختبار يرتبط فعلا

بدرجات الرياضيات فهذا دليل على التقارب.

على أن التقارب بين المؤشرات ليس كافياً فقد أشد مارمسيك (Messick, 1989) إلى وجود الحاجة إلى إظهار إمكانية التمييز أمبيريقيا بين المفهوم الذي نحاول قياسه وغيره من المفاهيم. وللحصول على أدلة على إمكانية التمييز بين المفاهيم نحصل على بيانات تدل على أن المفهوم لا يرتبط بالمقاييس المعروفة عنها أنها تقيس مفاهيم مختلفة. أي أن واضع الاختبار يحصل على بيانات عن عدم ارتباط المقاييس التي تقيس مفاهيم أخرى بالمفهوم الذي نريد قياسه. فاختبار الاستدلال الرياضي يجب أن يكون ارتباطه ضعيفاً بالاختبارات التي تقيس القدرة على القراءة، لأن القراءة مفهوم ليس له ارتباط بالرياضيات. وإذا وجدنا أن الارتباط ضعيف بين اختبار الرياضيات واختبار القراءة كان هذا دليلاً على تمايز المفاهيم. وبالطبع فإن اختيارنا للمفاهيم التي نحدد أنه يمكن تمييزها عن المفهوم الذي نرغب في قياسه، يجب أن يكون محكوماً بأسس منطقية، وعلمية. فمن المعقول مثلاً أن نستخدم مفهوم القراءة لتحقيق أدلة متميزة عن مفهوم الرياضيات، حيث إن كلا من مجال الرياضيات ومجال القراءة، جانب هام من القدرة المعرفية. ولكن من غير المعقول مثلاً أن نستخدم اختباراً في مهارة إلقاء الكرة لتحقيق التمايز بين اختبار الرياضيات فهذا النوع من التمايز لا معنى له. ولعله من الأفضل أن نفكر في مقياس التكوين الفرضي باعتباره نقطة على طول متصل من المفاهيم المرتبطة، وأن التقارب أو التمايز بين المفاهيم هي نقاط أخرى على نفس المتصل. فمن المفروض مثلاً أن يرتبط مقياس 'الميل الاجتماعي' ارتباطاً موجباً بمتغير الانبساط (تقارب)، ويرتبط ارتباطاً سالباً بمقياس الانطواء (تمايز).

ومن الأساليب الأخرى لجمع الأدلة الارتباطية التحليل العاملي. فالتحليل العاملي أسلوب إحصائي لدراسة الارتباط بين مجموعة من درجات الاختبارات لتحديد عدد العوامل (المفاهيم) التي يمكن أن تفسر هذه الارتباطات. ويمدنا هذا الأسلوب أيضاً بمعلومات عن أي العوامل تحدد الأداء في كل اختبار، وكذلك النسبة المئوية للتباين في درجات الاختبار التي تفسرها هذه العوامل. ويبدأ الباحث عادة بمجموعة كبيرة من الاختبارات المختلفة، ودراسة الارتباطات بينها ومعرفة أي المقاييس ترتبط ببعضها البعض، يمكن أن نخفض العدد الكبير من الدرجات إلى عدد أقل من العوامل. وتشير الارتباطات ليس فقط إلى أي الاختبارات تقيس نفس العامل، ولكن إلى أي درجة تقيس العامل. وبفحص

محتوى الاختبارات ذات التشبع المرتفع على نفس العامل، يمكن استنتاج طبيعة التكوين الفرضي الذي نقيسه.

٢- **الدراسات التجريبية:** يمكن اختبار الفرض بأن درجات الاختبار تتغير عند تقديم أنواع معينة من المعالجات. مثال ذلك، عند تقنين مقياس في القلق، فإذا وضعنا فرضاً بأن درجات المقياس يمكن أن تتغير عندما يوضع الأف راد في موقف مثير للقلق. فإذا تم معالجة القلق في موقف تجريبي مضبوط، وتغيرت درجات القلق في مقياس القلق بالطريقة التي يتنبأ بها الفرض، ي صبح لدينا بعض الأدلة على أن المقياس يقيس القلق.

٣- **مقارنة درجات مجموعات محددة:** يمكن استخدام مجموعات مع روف أنها مختلفة، ونفترض أن درجات الأداة التي نبنيها تميز بين كل مجموعة والأخرى. فيمكن مثلاً أن نفترض أن درجات اختبار استعداد موسيقي تميز بين الطلبة المقيدين في كلية التربية الموسيقية، وعينة عشوائية من طلبة الجامعة. وبالمثل إذا أمكن التمييز بين الميكانيكيين وغير الميكانيكيين باستخدام درجات اختبار استعداد ميكانيكي، فإن هذا يعطي سنداً لصدق الاختبار كمقياس للاستعداد الميكانيكي. وإذا وضعنا استفتاء لقياس التوافق النفسي، فمن المتوقع أن تميز درجات هذا الاستفتاء بين مجموعة من الأفراد معروف أنها متوافقة نفسياً ومجموعة أخرى من المرضى النفسيين، والمجموعات المختلفة التي تستخدم في مقارنات الصدق، يمكن أن تكون مجموعات عمرية، أو مجموعات جنسية، أو مجموعات تختلف من حيث كمية التدريب في بعض المجالات المرتبطة بالتكوين الفرضي الذي نقيسه، أو مجموعات معروف أنها طبيعية في مقابل مجموعات معروفة بسوء التوافق، وهكذا. وإذا تأكدت الفروق التي نتنبأ بها، فإن هذا يكون دليلاً على صدق الاختبار كمقياس للتكوين الفرضي الذي نسعى إلى وضع اختبار له.

٤- **التحليل الداخلي للاختبار:** يهتم التحليل الداخلي للاختبار بدراسة الاختبار نفسه، فيجمع معلومات عن محتوى الاختبار، والعمليات المستخدمة في الاستجابة لمفرداته، والارتباطات بين مفردات الاختبار. ويمكن أن توفر البيانات التي نحصل عليها من الدراسات المرتبطة بالمحتوى معلومات عن التكوين الفرضي الذي يقيسه الاختبار. وعندما يحدد المجال السلوكي الذي يمثل الاختبار، فإن هذا يعطينا فهماً لطبيعة التكوين الذي يقيسه الاختبار. مثال ذلك إذا حددنا

المجال السلوكي الذي يقيسه اختبار استدلال بوصف القدرات التي يحتوي الاختبار على عينة منها (مثل القدرة على الاستدلال باستخدام المتشابهات الكمية واللفظية)، فإن هذا يزيد من فهمنا لصدق الاختبار.

ويمكن كذلك استقصاء العمليات والمهارات العقلية التي يستخدمها الفرد في الاستجابة لمفردات الاختبار. مثال ذلك، يمكن سؤال الطلبة أن يفكروا بصوت عالٍ، وهم يجيبون على أسئلة اختبار الاستدلال. ومثل هذه الطريقة يمكن أن تكشف أن الاختبار يقيس فعلاً القدرة على الاستدلال التي يزعم قياسها، أو قد تكشف عن أن الاختبار يقيس عوامل أخرى ترتبط بالمفردات المستخدمة، أو القدرة على فهم المادة المقروءة.

ويجب كذلك دراسة درجة تجانس محتوى الاختبار لمعرفة ما إذا كان الاختبار يقيس سمة أو خاصية واحدة. فمقاييس التناسق الداخلي مثل معامل ألفا، أو معامل كودر ريتشاردسون، قد توفر لنا معلومات عن التناسق الداخلي للاختبار. وسوف ندرس هاتين المعادلتين بعد قليل أثناء مناقشة ثبات المقاييس. وتوفر لنا مقاييس التجانس أدلة ذات علاقة بالتكوين الفرضي الذي نقيسه، لأنها تصف لنا السمة أو السمات التي يقيسها الاختبار. ومع ذلك فإن التحليل الداخلي للاختبار ليس كافياً كدليل وحيد على صدق الاختبار، ولكنه يوفر لنا معلومات مساعدة، يمكن أن تلقي ضوءاً على صدق الاختبار.

صدق الاختبارات محكية المرجع:

تهتم الاختبارات محكية المرجع بقياس أهداف تدريس ضيقة عادة. ولذلك فإن الاهتمام الرئيسي لهذا النوع من الاختبارات يتركز على الصدق المرتبط بمحتوى الاختبار. ويشير هدف التدريس إلى عنصرين (Thorndike, 1997):

١- معلومات المحتوى المتعلقة بالمجال الذي يقيسه الاختبار.

٢- ما يمكن للطلاب أن يفعله بهذا المحتوى.

ويجب أن يظهر هذان العنصران في مفردة الاختبار بالطريقة التي يدعو إليها الهدف إذا أردنا أن تكون المفردة مؤشراً صادقاً بالتحصيل لهذا الهدف.

ومن الجوانب التي قد تهمل من محتوى الاختبار الظروف التي يختبر فيها الطلاب. فإذا كان الزمن المخصص للاختبار أقل من أن يسمح لمعظم الطلاب بالإجابة على جميع الأسئلة تصبح سرعة الاستجابة عاملاً في محتوى الاختبار. وإذا كانت ظروف إجراء الاختبار غير مواتية، تصبح هذه الظروف جزءاً من

محتوى الاختبار بالنسبة لهؤلاء الطلاب. ونظرا لأن مثل هذه العوامل تختلف من مجموعة لأخرى من الطلاب، يختلف محتوى الاختبار حسب المجموعة التي يطبق عليها. ويرتبط صدق المحتوى بدرجة الاختبار وجميع العوامل التي تؤثر على هذه الدرجة، بما في ذلك وضوح التعليمات وسلامة إجراءات تقدير الدرجة. وعندما يكون محتوى الاختبار عاملا مؤثرا في استخدام درجة الاختبار بجنب الانتباه تماما إلى جميع العوامل التي تؤثر على محتوى الاختبار وتصبح جزءا منه.

ويحدد صدق المحتوى كما سبق أن رأينا رأي محكمين من المدرسين وخبراء المنهج. وهناك طريقة أخرى إحصائية أو أمبيريقية للحصول على أدلة لصدق المحتوى بالنسبة للاختبارات محكية المرجع. وهي أن الاختبارات محكية المرجع يجب أن تزيد الفروق بين المجموعات التي أتقنت المقرر والمجموعات التي لم تتقن هذا المقرر، كما يجب أن تقلل في نفس الوقت من الفروق داخل المجموعة الواحدة (رجاء أبو علام، ١٩٨٧). ولذلك فإن الاختبار محكي المرجع المثالي يجب أن يقسم المفحوصين إلى مجموعتين متجانستين: مجموعة المتقنين ومجموعة غير المتقنين.

والمشكلة بالطبع هي كيفية تعريف الإتقان. وأحد الطرق هي سؤال المدرسين عن الطلبة المؤكد أنهم طلبة متقنون، والطلبة المؤكد أنهم غير متقنين. وهذه الطريقة يمكن أن تعطينا مجموعتين مختلفتين اختلافا واسعا، وفي هذه الحالة يجب حذف الطلبة الذين يقعون في الوسط. ومن الطرق الأخرى مقارنة بين الطلبة الذين تلقوا تعليما في الهدف والطلبة الذين لم يتعلموا هذا الهدف بعد. وهناك طريقة ثالثة وهي مقارنة أداء الطلبة قبل التعليم بأدائهم عقب التعليم. ولكل طريقة من هذه الطرق نواحي قصورها، وكلها تتأثر بمشكلة عدم ثبات فروق الدرجات أو تغييرها. إلا أن أي طريقة منها توفر لنا بعض الأدلة عن صدق الأنشطة الموجهة نحو أهداف التدريس، وصدق الاختبار كمقياس لإتقان تلك الأهداف.

وتشبه هذه الطريقة في تحقيق الصدق الطريقة التي سبق ذكرها عند مناقشة الصدق المرتبط بالمحتوى في جزء سابق من هذا الفصل، والتي تقوم على تحليل مفردات الاختبار. وعند تطبيق هذه الطريقة على الاختبارات محكية المرجع فإننا نقسم المفحوصين إلى مجموعتين: المتقنين في مقابل غير المتقنين، أو الذين درسوا الهدف في مقابل أولئك الذين لم يدرسوه، ونحصل بذلك على نسبة الذين أجابوا إجابة صحيحة عن المفردة في كل مجموعة. والمفردة الصادقة هي المفردة التي يزيد عدد من أجابوا عليها إجابة صحيحة في المجموعة المتقدمة أو التي درست

الهدف زيادة كبيرة على نسبة النجاح في المجموعة الأخرى. وفي أقصى الحالات يجب أن تكون نسبة نجاح المتقنين ١٠٠٪، في حين تكون نسبة نجاح غير المتقنين راجعة إلى عامل الصدفة. ويجب أن تقسم الدرجة الكلية للاختبار المفحوصين إلى متقنين حصلوا جميعا على الدرجة النهائية أو درجة قريبة من النهائية، وغير المتقنين الذين حصلوا على درجات كلية قريبة جدا من الصفر. ويجب ألا يكون هناك أية أخطاء راجعة إلى عملية التصنيف.

ومن النادر أن يكون الإتقان مطلقا. ففي كل صف نجد درجات متباينة من الإتقان. والواقع أن أحد مشكلات مفهوم الاتقان ترتبط بتحديد تعريف مقبول للإتقان (Shrock & Coscarelli, 2000). وميزة التقويم القائم على الاختبارات محكية المرجع هي أننا نعرف نظريا على الأقل ما يمكن أن يقيوم به الطلاب المتقنون، ونوع المشكلات التي يستطيعون حلها أو المهام التي يمكن أن يقوموا بها. ويرتبط تعريف مفهوم الاتقان بما يعرف بدرجة القطع، وهي الدرجة التي تفصل بين المتقنين وغير المتقنين.

ويرى شروك وكوسكاريلي (Shrock & Coscarelli, 2000, p. 41) أن تحديد درجة القطع عملية صعبة تماما رغم أهميتها، فهي متطلب ضروري للاختبارات محكية المرجع، لأن على أساسها نحدد الطلبة المتقنين وغير المتقنين. وعند الانتهاء من بناء الاختبارات بالطريقة السابق ذكرها لابد من تحديد درجة القطع حتى يمكن تحديد مستوى الاتقان. وهناك عدة أساليب لتحديد هذه الدرجة أهمها ثلاث طرق هي:

١- طريقة المحكمين الخبراء.

٢- طريقة أنجوف.

٣- طريقة المجموعات المتقابلة.

ومهما كانت الطريقة التي يستخدمها واضع الاختبار لتحديد درجة القطع إلا أن المحكمين يجب أن يكون لهم دور في مرحلة ما من مراحل تحديد تلك الدرجة. ورغم أن بعض الطرق تعتمد اعتمادا كبيرا على التكميم والعمليات الإحصائية، إلا أننا يجب ألا ننساق إلى الاعتقاد بأن الأرقام يمكن أن تكون دليلا عن رأي المحكمين.

الثبات

يقصد بالثبات دقة المقياس أو اتساقه، فإذا حصل نفس الفرد على نفس الدرجة (أو درجة قريبة منها) في نفس الاختبار (أو مجموعات من الأسئلة المتكافئة أو المتماثلة) عند تطبيقه أكثر من مرة فإننا نصف الاختبار أو المقياس في هذه الحالة بأنه على درجة عالية من الثبات. وبهذا المعنى يرتبط مفهوم الثبات بما يسمى إحصائيا أخطاء القياس المتضمن في كل درجة من درجات المقياس.

ومعامل الثبات هو معامل ارتباط بين درجات الأفراد في المقياس في مرات الإجراء المختلفة، أو بين تقديرات من يقومون بتقدير الدرجات في المرات المختلفة، أو بين نتائج إجراء المقياس على مجموعة واحدة من الأفراد يقوم بالإجراء فيها أخصائيون مختلفون. أي أن معامل الثبات هو معامل ارتباط بين المقياس ونفسه، أو فنحن نحسب معامل الثبات بحساب معامل الارتباط بين درجات المقياس ونفسه، أو بين درجات المقياس وصورة أخرى مكافئة له. ولذلك فإننا عندما نتكلم عن الثبات إنما نتكلم عن طريقتين مختلفتين: هما استقرار درجات الاختبار، أو التماسك الداخلي لمفردات الاختبار.

ولذلك فإن تقدير ثبات الاختبار أو أي مقياس آخر إنما يشير إلى نوع معين من اتساق الدرجات. فدرجات المقياس ليست ثابتة بشكل مطلق، فهذه الدرجات قد تكون ثابتة عبر فترة زمنية (استقرار الدرجات)، أو عبر عينة مماثلة من البنود (التماسك الداخلي للاختبار)، أو عبر مقدرين مختلفين، وهكذا. ومن الممكن أن يكون الاختبار ثابتا في جانب من هذه الجوانب وغير ثابت في جانب آخر. ويحدد استخدام المقياس والمعنى الذي نستقيه من درجاته الأسلوب المناسب من الثبات. فإذا أردنا معرفة حالة الأفراد في المستقبل فلا بد من معرفة مدى استقرار الدرجات عبر فترة زمنية. ومن ناحية أخرى إذا أردنا أن نعرف مدى التحول في فرد من الأفراد في حالة القلق التي لديه من لحظة لأخرى يكون عامل استقرار الدرجات أمرا غير مرغوب فيه حتى نحصل على المعلومات التي نريدها. وعلى هذا فإننا نحتاج لكل تفسير من التفسيرات نوعا مختلفا من الثبات.

والثبات ضروري للمقياس، ولكنه ليس بديلا عن الصدق. والمقياس الذي يعطينا درجة غير مستقرة أو غير متسقة، لا يمكن أن يكون صادقا. ومن الناحية الأخرى إذا كانت درجة ثبات الاختبار عالية جدا، فهذا ليس دليلا على صدق

المقياس. ولذا فإن انخفاض معامل الثبات يحد من الصدق، ولكن ارتفاع معامل الثبات، لا يضمن درجة مرضية من الصدق.

نظرية الثبات:

من المفيد للتمييز بين ثبات الاختبار وصدقه، أن نفرق بين الأخطاء العشوائية للقياس، والأخطاء المنتظمة للقياس. والأخطاء العشوائية هي أخطاء ترجع إلى الصدفة البحتة. وهذه قد تؤدي إلى تضخم الدرجات أو انكماشها بطريقة لا يمكن التنبؤ بها. مثال ذلك نفرض أن أحد اختبارات التربية الرياضية يتضمن ركل كرة القدم لأبعد مسافة ممكنة. ويعطي المدرب تعليمات بذلك إلى التلاميذ. ويتم قياس المسافة التي يركل التلميذ فيها الكرة. وهدف الاختبار هو الحصول على درجة تعبر عن قدرة التلميذ على ركل الكرة. ومن المؤكد أنه إذا طلبنا من نفس التلميذ ركل الكرة في عدة مناسبات مختلفة، فإننا لن نحصل على نفس الدرجة (المسافة) في كل مرة.

نفرض أننا طلبنا من التلاميذ ركل كرة القدم في يومين متتاليين. فإذا قارنا بين درجات (مسافة ركل الكرة) كل تلميذ في اليوم الأول بدرجاته في اليوم الثاني، فإننا غالباً ما نحصل على درجات مختلفة لكل تلميذ. ومعظم الفروق تكون صغيرة، إلا أن بعضها يكون كبيراً، مع وجود نسبة محدودة من الفرق الكبيرة جداً. والنتائج غير متجانسة من يوم لآخر. أي أنها غير ثابتة.

وهناك ثلاثة أنواع من أخطاء الصدفة، أو الأخطاء العشوائية، التي تؤدي إلى عدم الاتساق بين الدرجات في اليومين:

١- **قد يتغير التلميذ بالفعل** من اليوم الأول إلى اليوم الثاني. فقد يشعر في يوم أنه أفضل منه في اليوم الآخر. فقد يشعر في يوم بأنه متعب، أو أقل اهتماماً بالنشاط.

٢- **قد يتغير العمل نفسه** في اليومين. فقد تكون الكرة المستخدمة في اليوم الأول غير الكرة المستخدمة في اليوم الثاني، فقد تكون أخف، أو أقل امتلاءً بالهواء. وقد يسمح المدرب للتلميذ في اليوم الأول بالجري حول الملعب قبل بدء نشاط ركل الكرة. وقد يطلب منهم في اليوم الثاني ركل الكرة فقط. وقد تساعد هذه التغيرات بعض التلاميذ دون البعض الآخر.

٣- **عينة السلوك المحدودة المستخدمة** تؤدي إلى عدم استقرار الدرجة.

فالعينة الصغيرة للسلوك معرضة لأنواع كثيرة من مؤثرات الصدفة، فقد يكون هناك نشاط للرياح عند ركل الكرة في يوم دون اليوم الآخر. وقد يفقد التلميذ توازنه عند ركل الكرة، أو قد تخونه قدمه عند تصويبها نحو الكرة. ويتعلق ثبات القياس بهذه الأنواع من التأثيرات العشوائية في الدرجات.

ونجد من ناحية أخرى أن بعض الأخطاء أخطاء منتظمة يمكن التنبؤ بها. ففي المثال السابق عن ركل كرة القدم، لنفرض أن المدرب يعطي التعليمات باللغة العربية الفصحى التي لا يستطيع بعض الأطفال فهمها. وهنا نجد أن الأطفال الذين لا يفهمون اللغة العربية الفصحى فهم جيداً سوف تنخفض درجاتهم مقارنة بالتلاميذ القادرين على فهم التعليمات باللغة الفصحى. وهذا الانخفاض في الدرجات منظم في عملية ركل الكرة في كل مرة، ويمكن التنبؤ بأنه سوف يؤدي إلى خفض درجة هؤلاء الأطفال. وهذه الأخطاء المنتظمة هي أخطاء تتعلق بصدق القياس. حيث إننا نجد أن صدق الاختبار ينخفض كلما حدث تغير منتظم في الدرجات لا يرجع إلى عملية القياس نفسها. ففي هذه الحالة لا نقيس فقط سلوك ركل الكرة، ولكننا نقيس بدرجة ما قدرة التلميذ على فهم اللغة العربية الفصحى.

وللتأكد من نوع المشكلة التي أمامنا وهل هي تتعلق بالثبات أم بالصدق، فإننا نحدد أولاً ما إذا كانت المشكلة تتعلق بالأخطاء العشوائية أم الأخطاء المنتظمة. فلو أعطينا مثلاً أحد الفصول اختبار ركل كرة القدم السابق، واستخدمت فيه كرتين إحداهما قوية، والثانية رخوة، وكان استخدام التلاميذ لأي من الكرتين يرجع للصدفة البحتة، فإن التباين الذي يرجع إلى الكرة يكون مشكلة تتعلق بالثبات. ذلك أن التباين الذي يرجع للكرة خطأ عشوائي ويؤثر على تناسق القياس. وإذا طلب من التلاميذ ركل كرة القدم حسب الحروف الأبجدية، وكان اليوم ممطراً، وازدادت تأثيرات المطر مع تتابع ركل الكرة، فإن التباين الذي يرجع إلى تأثير الكرة بمياه الأمطار هو تباين يتعلق بالصدق. ذلك أن درجات التلاميذ الذين كانوا في أول القائمة تصبح أعلى من درجات التلاميذ الموجودين في نهاية قائمة الفصل، بسبب تأثير الكرة بالمطر. أي أن صدق درجات ركل الكرة ينخفض لأن الدرجات لا تعكس قوة الركلة فقط، بل إنها تعكس كذلك الترتيب الأبجدي لتلاميذ الفصل. وهذا مثال على الخطأ المنتظم الذي يؤثر على صدق القياس.

ويهتم الثبات بمدى تناسق ما نقيسه. ولا يهتم إن كنا نقيس ما نريد قياسه، فهذا أمر يخص الصدق. ومن الممكن أن تكون الأداة ثابتة وليست صادقة، إلا أنها

لا يمكن أن تكون صادقة إلا إذا كانت ثابتة أولاً. وقد سبق أن ذكرنا مثال عرض الجبهة كمثال على قياس الذكاء، فهذا دليل على ثبات القياس، دون صدقه. ويمكن أن يكون القياس ثابتاً مهما تعددت مرات قياس عرض الجبهة، ولكن هذه الطريقة ليست صادقة في قياس الذكاء، لأن عرض الجبهة لا يرتبط بأي مقياس آخر للذكاء، ولا تسانده أي نظرية من نظريات قياس الذكاء.

ويتأثر الثبات بالأخطاء العشوائية، وهذه ترجع إلى العوامل التي تؤدي إلى الاختلافات بين الدرجات من تطبيق نفس الأداة مرات متتالية، أو في تطبيق أداة مكافئة لها.

وترجع الأخطاء العشوائية إلى مصادر متعددة. فقد تكون الأخطاء متأصلة في الأداة ذاتها، كأن يكون الاختبار قصيراً جداً، مما يؤدي إلى حدوث التلاميذ الذين يعرفون الإجابة على هذا العدد المحدود من الأسئلة على درجات أعلى مما يستحقون، في حين أن التلاميذ الذين لا يعرفون الإجابة على هذه الأسئلة سوف يحصلون على درجات أقل مما يستحقون. فإذا صمم اختبار م ثلاثي لقياس درجة معرفة التلاميذ بعواصم المحافظات المصرية. وسألنا خمسة أسئلة فقط، فمن المحتمل أن يحصل التلميذ الذي يعرف أسماء العواصم الخمس فقط على درجة كاملة، في حين يحصل التلميذ الذي لا يعرف أسماء هذه العواصم الخمس ويعرف أسماء باقي العواصم، على صفر. ففي مثل هذه الاختبارات القصيرة جداً يلاحظ دوراً كبيراً منه في الاختبارات الطويلة. وإذا كان الاختبار سهلاً جداً لدرجة أن معظم التلاميذ يعرفون الإجابة الصحيحة، فإن الدرجة النسبية للتلاميذ سوف تعتمد على عدد قليل من الأسئلة، ويدخل الحظ أيضاً كعامل رئيسي في الدرجة. وإذا كانت الأسئلة في معظمها غامضة غير واضحة، فإن الذين يستجيبون بالطريقة التي يريدها المدرس، يحصلون على درجات عالية، في حين أن باقي التلاميذ الذين أجابوا إجابات تعتمد على الصدفة أيضاً سوف يحصلون على درجات منخفضة. أي أن طريقة تقدير الدرجات تؤثر في الثبات، ولذلك إذا كان تقدير الدرجات دقيقاً أدى ذلك إلى ارتفاع الثبات، وإذا كان غير دقيق أدى ذلك إلى انخفاض الثبات.

وقد ترجع الأخطاء إلى طريقة إجراء الاختبار، فإذا أجرى فاحص غير متمرس اختباراً فردياً في الذكاء مثلاً، ابتعد أدائه عن الطريقة المقننة لإجراء الاختبار، وترتب على ذلك أخطاء في الإجراء وفي تقدير الدرجات. كما أن ظروف إجراء الاختبار مثل الإضاءة والتهوية والحرارة، تؤثر على أداء التلاميذ. وإذا كانت تعليمات إجراء الاختبار غامضة أثر ذلك على أداء التلاميذ أيضاً.

وهناك الأخطاء التي ترجع إلى التلميذ نفسه، فالتقلبات التي ترجع إلى عوامل مثل الدافعية، والميل، والتعب، والحالة الصحية، والقلق، وغير ذلك من العوامل الانفعالية والعقلية تؤثر على نتائج الاختبار.

طرق حساب معامل الثبات:

عند حساب معامل الثبات نحصل على درجات الأفراد في المقياس بإجرائه أكثر من مرة أو بالحصول على درجات متكافئة في نفس الإجراء، أو أكثر من إجراء وعلى هذا الأساس يمكن تحديد طرق حساب الثبات (ج دول ٢١-٢)، فيما يلي:

١- إجراء صورتين متكافئتين من المقياس وحساب معامل الارتباط بين درجات نفس الأفراد في كل منهما (تكافؤ الدرجات).

٢- إجراء نفس الصورة من المقياس على نفس الأفراد مرتين وحساب معامل الارتباط (استقرار الدرجات).

٣- تجزئة المقياس الواحد إلى مجموعتين متكافئتين من المفردات، وإعطاء درجة لكل فرد على كل من القسمين وحساب معامل الارتباط (الاتساق الداخلي).

٤- تباين المقياس وتباين المفردات (الاتساق الداخلي).

١- طريقة الصور المتكافئة: يفترض في هذه الطريقة أنه يمكن تصميم صورتين

متكافئتين من المقياس الواحد، ويشترط لتكافؤ الصورتين أن يتحقق ما يلي:

أ- أن تكون المجالات والموضوعات التي تقيسها الصورتان واحدة.

ب- تساوي نسبة المفردات التي تخص كل موضوع من الموضوعات.

ج- تماثل مستوى صعوبة المفردات أو انحرافها المعياري.

د- تشابه طريقة صياغة المفردات.

هـ - تساوي طول الصورتين وطريقة إجرائهما وتقدير درجاتهما وتوقيتهما.

و- تساوي متوسط وتباين درجات الأفراد في الصور المتكافئة.

ومن المفيد في كثير من المواقف التطبيقية أن يكون لدينا صور مختلفة من الاختبار. مثال ذلك إذا كنا نتابع تطور حالة فرد العلاجية، فقد نحتاج إلى إجراء اختبار لقياس القلق أو الاكتئاب على فترات منتظمة. وتطبيق نفس الاختبار عدة مرات أمر غير مناسب، وفي هذه الحالة من الأفضل أن يكون لدينا أكثر من صورة للاختبار. ولكن لا نستطيع الاعتماد على استخدام الصور المتكافئة إلا إذا كان

الارتباط بين درجاتهما مرتفعاً (وكذلك المتوسطات، والانحرافات المعيارية، وتوزيع الدرجات). وهذا النوع من الارتباط هو ثبات الصور المتكافئة. وإذا انخفض معامل الثبات عن ٩٠، فمن الصعب التسليم بتكافؤ الدرجات في صورتين. ومن النادر أن نجد اختبارات تعطي معامل ثبات الصور المتكافئة مرتفعاً إلى هذا الحد. ولكن مع التطور الهائل في استخدام الحاسب الآلي في تقنين الاختبارات أصبح من الأسهل الحصول على اختبارات لها صور متكافئة.

جدول ٢١-٢ طرق تقدير الثبات

الطريقة	نوع الثبات	الإجراء
الصور المتكافئة	قياس التكافؤ	إعطاء صورتين من المقياس لنفس المجموعة في فترتين متتاليتين.
إعادة الإجراء	قياس الاستقرار	إعطاء نفس المقياس مرتين لنفس المجموعة في فترتين تفصلهما فترة قصيرة (يوم) أو فترة طويلة (عدة سنوات).
إعادة الإجراء بصورة مكافئة	قياس الاستقرار والتكافؤ	إعطاء صورتين من المقياس لنفس المجموعة تفصلهما فترة زمنية طويلة نسبياً.
التجزئة النصفية	قياس التناسق الداخلي	إعطاء الاختبار مرة واحدة، فصل درجات المقياس في نصفين متكافئين (المفردات الفردية والزوجية)، مع تصحيح معامل الثبات باستخدام معادلة سبيرمان براون.
تباين المفردات	قياس التناسق الداخلي	إعطاء المقياس مرة واحدة، تقدير درجات المقياس كاملاً ثم تطبيق معادلة كودر ريتشاردسون، أو معادلة ألفا.

٢- طريقة إعادة الإجراء: في هذه الطريقة يعاد تطبيق نفس المقياس على مجموعة من الأفراد. وحساب معامل الثبات بهذه الطريقة يتأثر ببعض العوامل، وقد سبق أن تناولنا هذه العوامل عند مناقشة نظرية الثبات والعوامل المؤثرة في الثبات بشكل عام. وسوف نناقش هذه العوامل هنا بتفصيل أكثر حيث إنها ترتبط ارتباطاً وثيقاً بطريقة إعادة الإجراء. وفيما يلي نتناول هذه العوامل:

أ- التغيرات الحادثة في الأفراد: إذا أعطينا مجموعة من الأفراد اختباراً في الذكاء، ثم أعدنا إجراء نفس الاختبار عليهم عقب انتهائهم منه مباشرة، فلا شك في أننا سوف نحصل على معامل ثبات مرتفع للغاية، مع التسليم بأن الأفراد لم يكونوا متعبين ولم يصبهم الملل. إلا أن الحصول على معامل

ثبات بهذه الطريقة يعطينا معاملا زائفا بالغ الارتفاع حيث إن معظم الأفراد سوف يتذكرون إجاباتهم. وللتغلب على مثل هذه الصعوبة فمن الطبيعي أن يفصل بين الإجراءين فترة زمنية معقولة لا تقل عن أسبوعين ويمكن أن تزيد عن ذلك. فإذا طبقنا الاختبار على مجموعة من الأطفال وكانت الفترة الفاصلة بين الإجراءين ثلاثة أشهر مثلا، سوف تحدث تغيرات حقيقية في هؤلاء الأطفال، وحيث إن الأطفال ينمون بسرعات متفاوتة فإن الارتباط بين مجموعتي الدرجات سوف "يبدو" منخفضا، وكلمة يبدو هنا مقصودة لأنها تعكس حقيقة أن التغيرات في الدرجات هي تغيرات ترجع إلى أخطاء القياس وليست إلى تغيرات حقيقية في المتغير. وإذا كنا نقوم بدساب معامل الثبات لاختبار يقيس حالات مزاجية عابرة مثل الخوف أو الغضب، فإن معامل الثبات المحسوب بطريقة إعادة الإجراء سوف يكون منخفضا جدا، لأن هناك اختلافات حقيقية سوف تحدث بين الإجراءين. ولذلك فإن استخدام طريقة إعادة الإجراء مع الاختبارات التي تقيس متغيرات حالات وقتية مثل المتغير السابق ذكره يستلزم محاولة رفع مستوى الخوف والغضب في الإجراءين. ومن هذا يتضح أنه بالرغم من أننا نرغب في الحصول على معامل ثبات مرتفع بطريقة إعادة الإجراء إلا أن اختلاف الدرجات في المرتين يدعونا إلى عدم الثقة في أي من الإجراءين، وعلى هذا فإننا عند اختيار طريقة لحساب معامل الثبات يجب أن نأخذ في الاعتبار نوع المتغير الذي نقيسه، وطبيعة العينة التي نستخدمها. أضف إلى هذا أن موقف المقياس في الإجراء الأول كثيرا ما يختلف عن موقف المقياس في الإجراء الثاني، حيث لا يكون المفحوصون متأثرين في المرة الثانية بغرابة الموقف، وما قد يصحبه من توتر انفعالي. وبذلك تتأثر درجات الإجراء الأول بعوامل نفسية معينة بينما تتأثر درجات الإجراء الثاني بعوامل أخرى مثل الألفة بموقف الاختبار.

ب- العوامل التي تزيد من أخطاء القياس: ننتقل الآن إلى العوامل التي تضعف معامل الثبات بطريقة إعادة الإجراء وبالتالي تزيد من أخطاء القياس. وبعض هذه العوامل واضح ولا يحتاج إلى مناقشة. فبعض الأفراد لا يشعرون أنهم في أحسن أحوالهم أثناء اختبارهم مما يضعف من أدائهم في الاختبار. وأهم أسباب ذلك الحالة الصحية للبارد والصداع، والمشكلات المنزلية أو الانفعالية التي تؤثر في درجات الاختبار. والشعور بالإرهاق من الليلة السابقة، إلى آخر ذلك من العوامل التي لا نهاية لها.

وهناك عوامل أخرى مشابهة تؤثر في درجات الاختبار، مثل كرسي غير مريح، أو قلم غير سليم، أو ترك مجموعة من الأسئلة دون انتباه، أو قلب صفحات من صفحات الاختبار معاً. وكل هذه العوامل قادرة على خفض معامل ثبات الاختبار. وهذه العوامل يصعب تجنبها لأنها جزء من الممارسات الحياتية اليومية، ولذلك فإن معامل ثبات أفضل الاختبارات لا يمكن أن يكون تاماً. كما يجب ألا ننسى انتقال أثر التدريب من الإجراء الأول إلى الإجراء الثاني مما قد يؤثر على درجات الأفراد خاصة وأن الفروق الفردية تلعب دورها في اختلاف درجة الانتقال.

والعوامل التي ناقشناها حتى الآن ترجع إلى الأفراد وليس إلى الاختبار نفسه. إلا أن هناك بعض خصائص الاختبارات التي قد تؤدي إلى ضعف معامل الثبات بطريقة إعادة الإجراء. ومن المفروض أن يتم استبعاد هذه العوامل أثناء بناء الاختبار، إلا أنه من الصعب في بعض الاختبارات تجنب مصادر الخطأ تجنباً تاماً، وتتمثل مصادر الخطأ هذه فيما يلي:

- **احتواء الاختبار على تعليمات رديئة:** وهذه تعتبر من المصادر الشائعة في الاختبارات مما يؤدي إلى صعوبة فهم المختبرين بما يجب عليهم عمله. وقد يترتب على هذا أن يكون أداء الأفراد أفضل في مرة دون أخرى. وهذا مصدر آخر مشابه للخطأ ينشأ عندما يطلب الاختبار شيئاً كالمقابلة الاستجابية. فعندما يطلب الاختبار مثلاً وضع علامة على ورقة الإجابة على 'أ'، إذا كان السؤال صحيحاً، وعلى 'ب' إذا كان خاطئاً فإن مجال الخطأ هنا يكون محدوداً، إلا أن بعض الاختبارات تتطلب استجابات معقدة جداً تصعب من الأمر على المستجيب، وبذلك تصبح مصدراً من مصادر الخطأ.
- **التقدير الذاتي للاستجابات:** وهذا مصدر آخر من مصادر الخطأ إذ يترتب عليها اختلافات كبيرة بين المقدرين في المرة الواحدة أو بين تقديرات المقدر الواحد في المرات. ولاشك في أن التقديرات الذاتية تضعف ثبات الاختبار بشكل كبير، ولهذا السبب تستخدم معظم الاختبارات التبروية والنفسية الطريقة الموضوعية في تقدير الاستجابات.
- **التخمين:** والتخمين مصدر آخر من مصادر عدم الثبات. ومن الواضح أن هذا الأمر يؤثر بشكل رئيسي في اختبارات القدرات. إلا أن استخدام الاختيار من متعدد، وعدد كبير من الأسئلة يقلل من أثر هذا العامل.

ج- العوامل التي تضخم أو تشّتت معامل الثبات بطريقة إعادة الإجراء: سبق أن ذكرنا أننا نقيس الثبات باستخدام معامل الارتباط. وهذا عوالم ل تؤدي إلى زيادة معامل الارتباط زيادة كبيرة ولا تمثل الثبات الحقيقي بطريقة إعادة الإجراء بمعنى أنه قياس خال من الأخطاء لاسد تقرار درجات الاختبار. وأهم هذه العوامل ما يلي:

- **الفاصل الزمني بين الإجراءين:** وهذه أحد مصادر الخطأ التي سبق ذكرها. وهي ترك فترة صغيرة بين الإجراءين. ففي هذه الحالة يرتفع معامل الارتباط لأن أفراد العينة يتذكرون بعض استجاباتهم السابقة. ويرى كلاين (Kline, 2000) أن الفاصل الزمني بين الإجراءين لا يجب أن يقل عن ثلاثة أشهر. وإن كنا نرى أن هذه الفترة طويلة جداً وبخاصة بالنسبة للاختبارات التي لا تتأثر كثيراً بتذكر الاستجابات مثل اختبارات الشخصية ومقاييس الاتجاهات. أما الاختبارات التحصيلية واختبارات القدرات فإن عامل التذكر فيها مرتفع، ولذلك نرى أن الفترة بين الإجراءين يجب ألا تقل عن شهر في الاختبارات التي لا تتأثر كثيراً بعامل التذكر. أما الاختبارات التي يخشى من تأثرها بعامل التذكر فيجب ألا تقل عن ثلاثة أشهر كما ذكر كلاين.

- **مستوى صعوبة مفردات الاختبار:** ولهذا العامل أهمية خاصة في اختبارات القدرات واختبارات التحصيل، حيث يلعب مستوى صعوبة المفردة دوراً هاماً. فإذا كانت الأسئلة سهلة جداً بالنسبة للمستجيبين ارتفع معامل الثبات بطريقة إعادة الإجراء ارتفاعاً كبيراً، لأن المستجيبين يعطون الإجابة الصحيحة باستمرار. وكمثال على ذلك إذا احتوى الاختبار مثلاً على سؤال مثل $2 + 2 = ؟$ فإن مثل هذا السؤال لن يتأثر بالعوامل السابقة التي ذكرنا أنها تخفض معامل الثبات. كما أن الأسئلة الصعبة جداً سوف يكون لها نفس الأثر على معامل الثبات لأنه يستحيل على المستجيبين إعطاء الإجابات الصحيحة عليها.

- **المستجيبون:** المعاينة من العوامل الأخرى المؤثرة على معامل الثبات بطريقة إعادة الإجراء. حيث إن لطبيعة العينة التي يطبق عليها الاختبار أثرها على معامل الثبات. مثال ذلك إذا كان لدينا اختبار مصمم للمجموعات الإكلينيكية فمن الضروري الحصول على بيانات معامل الثبات بطريقة إعادة الإجراء من نفس المجموعات. ومن المعروف أن بعض

المجموعات الإكلينيكية من الصعب اختبارها مثل الذهانيين، وبالتالي فإن معامل الثبات سوف يكون منخفضاً. ويجب على الإكلينيكي أن ينتبه إلى هذه المشكلة عند تفسير نتائج الاختبارات النفسية التي يطبقها على مثل هذه المجموعات. وإذا حصلنا على بيانات معامل الثبات من مجموعة من الأفراد العاديين لاختبار سوف يستخدم مع المجموعات الإكلينيكية فإن معامل الثبات سوف يكون مضللاً.

• **حجم العينة:** معاملات الارتباط كغيرها من المؤشرات الإحصائية لها خطأ معياري يبين مدى الثقة في النتائج. ولقد رأينا من قبل عند مناقشة أثر حجم العينة على الخطأ المعياري كيف أن كبر حجم العينة يؤدي إلى انخفاض الخطأ المعياري (الفصل السادس). ويترتب على ذلك أنه يجب أن يكون حجم العينة التي نحصل منها على بيانات عن معامل الثبات بطريقة إحصائية إجراء كبيراً بدرجة كافية حتى نخفض من قيمة الخطأ المعياري. وهذا يعني من الناحية العملية أنه يجب ألا يقل حجم العينة عن ١٠٠ فرد. ولكن حجم العينة وحده ليس كافياً لضمان حسن اختيار العينة، صحيح أن الحجم الكبير للعينة يقلل من الخطأ المعياري، ولكن إذا كانت العينة غير ممثلة للمجتمع فإن البيانات التي نحصل عليها من هذه العينة لن تكون ذات قيمة كبيرة. مثال ذلك إذا أردت حساب معامل الثبات لأحد اختبارات الذكاء وحصلت على العينة من مدرسة معروفة بانتقائها لتلاميذها فإن معامل الثبات الذي نحصل عليه من هذه العينة لن يكون ذي قيمة كبيرة مهما انخفضت قيمة الخطأ المعياري.

وخلاصة القول أننا نرى، في ضوء العوامل السابق ذكرها، أن معامل الثبات بطريقة إعادة الإجراء يجب أن يكون مرتفعاً بدرجة كافية (لا يقل عن ٨٠)، مع فاصل زمني كافٍ بين مررتي الإجراء، إذا كنا نريد أن يكون لهذا المعامل فائدة عملية أو نظرية. كما يجب أن يكون حجم العينة كبيراً (لا يقل عن ١٠٠ فرد) وتكون العينة ممثلة للمجتمع الذي ننوي تطبيق الاختبار عليه.

٣- **طريقة التناقص الداخلي:** وتضم هذه الطريقة طريقتين هما التجزئة النصفية وحساب تباين مفردات الاختبار. ويلاحظ أنه كما كان الحال في قياس الثبات بطريقة إعادة الإجراء فإن جميع المعاملات يجب أن تحسب على عينات ممثلة للمجتمع الذي صمم الاختبار من أجله، ويجب أن يكون حجم العينة كبيراً لا يقل عن ١٠٠ فرد. ونناقش فيما يلي كلامنا عن الطريقتين اللتين

تقيسان التناسق الداخلي للاختبار.

أ- **التجزئة النصفية:** في هذه الطريقة تتم تجزئة المقياس إلى نصفين، ويعطى كل فرد درجة في كل نصف، أي أننا بعد تطبيق المقياس نقسمه إلى صورتين متكافئتين. وأفضل أساس للتقسيم في هذه الطريقة هي أن يحتوي القسم الأول على المفردات الفردية (١، ٣، ٥، ..)، والقسم الثاني على المفردات الزوجية (٢، ٤، ٦، ..) حتى نقل ما أمكن من العوامل المؤثرة في أداء الأفراد، مثل الوقت، والجهد، والتعب، والملل، وغيرها، وميزة هذه الطريقة هو توحيد ظروف الإجراء توحيدا تاما، ونظرا لأن معامل الارتباط في هذه الحالة يكون بين نصفي المقياس، فقد ظهرت عدة معادلات تعدل معامل الارتباط بين النصفين، بحيث تأخذ في اعتبارها مضاعفة طول الاختبار، ومن أهم هذه المعادلات معادلة سبيرمان براون:

$$\frac{r_{2v}}{r_{1v} + 1} = r_{11}$$

حيث

r_{11} معامل ثبات الاختبار كله.
 r_{1v} معامل الارتباط بين نصفي الاختبار.

فإذا كان معامل الارتباط بين نصفي الاختبار = ٠,٦٠، فإن

$$r_{11} = \frac{(0,60)^2}{0,60 + 1} = \frac{0,36}{1,60} = 0,225$$

ويلاحظ أن طريقة التجزئة النصفية تؤكد حقيقة مهمة فيما يتعلق بالثبات، وهي أن معامل ثبات الاختبار يرتبط بطوله. فيزيد معامل الثبات عندما يزداد طول الاختبار. ويندر أن يكون معامل ثبات الاختبارات التي تقل عن عشر مفردات مرتفعا. ولذلك يعتبر عشر مفردات هو الحد الأدنى للاختبار الثابت (Kline, 2000).

ب- حساب تباين مفردات الاختبار:

- طريقة كودر- ريتشاردسون: تفترض هذه الطريقة تساوي المفردات مع بعضها البعض في الأداة. وتستخدم عندما تقدر مفردات المقياس صفر أو ١. وتستخدم هذه الطريقة معادلتين هما: معادلة كودر ريتشاردسون ٢٠:

$$r_{11} = \frac{\sum_{k=1}^n \frac{E_k^2}{n-1} - \frac{(\sum_{k=1}^n E_k)^2}{n(n-1)}}{\sum_{k=1}^n \frac{E_k^2}{n-1}}$$

حيث:

$$r_{11} = \text{معامل ثبات الاختبار.}$$

$$n = \text{عدد مفردات الاختبار.}$$

$$E_k = \text{تباين الاختبار.}$$

$$\frac{E_k^2}{n-1} = \text{تباين المفردة.}$$

ويعتمد حساب معادلة كودر - ريتشاردسون ٢٠ على توفر البيانات عن تباين كل مفردة من مفردات المقياس. وفي حالة عدم توفر هذه البيانات يمكن استخدام معادلة كودر-ريتشاردسون ٢١، التي تتميز بالسهولة والسرعة في حسابها حيث إنها لا تحتاج إلى معرفة تباين المفردات، ولكن يعيبها أنها أقل دقة من المعادلة السابقة.

معادلة كودر - ريتشاردسون ٢١

$$r_{11} = \frac{\sum_{k=1}^n \frac{M_k^2}{n-1} - \frac{(\sum_{k=1}^n M_k)^2}{n(n-1)}}{\sum_{k=1}^n \frac{M_k^2}{n-1}}$$

- طريقة معامل ألفا: يعادل هذا المعامل طريقة كودر- ريتشاردسون ٢٠، ولذلك فهو يفترض تساوي المفردات في المقياس مع بعضها البعض. وتستخدم عندما يكون تقدير المفردات صفر، ١، ٢، ...، ويعتبر معامل ألفا أنسب طريقة لحساب ثبات الأوزان المستخدمة في

البحوث المسحية كالاستبيانات أو مقاييس الاتجاه حيث يوجد م دى
من الدرجات المحتملة لكل مفردة.

$$r_{11} = \frac{0.000000 - 1}{n - 1} \left(\frac{0.000000}{n} - \frac{0.000000}{n} \right)$$

العلاقة بين الطرق الثلاث لحساب معامل الثبات:

رغم أننا ناقشنا كلا من الطرق الثلاث لحساب معامل الثبات بمفرده، إلا أنها في الواقع مرتبطة. فمعامل الثبات بطريقة التناسق الداخلي يتضمن العلاقة بين مفردات الاختبار وبعضها البعض. وتعتبر هذه المفردات عينة عشوائية من المجال الذي نقيسه. وثبات الصور المتكافئة مناظر بالضرورة ولكن المفردات وضعت في صورتين بدلا من صورة واحدة. ومعامل الثبات بطريقة إعادة الإجراء يشبه التناسق الداخلي، إلا أن المفردات في هذه الحالة أجريت على فترتين. ورغم أننا في الحالة الأخيرة نواجه بالتعقيدات الناجمة عن تطبيق الاختبار مرتين، إلا أنه مع ذلك يجب أن تتشابه معاملات الثبات التي نحصل عليها بالطرق الثلاث، وإن كان معامل الثبات بطريقة إعادة الإجراء غالبا ما يكون الأقل.

الخطأ المعياري للقياس:

كما سبق أن ذكرنا في مكان آخر من هذا الفصل، يمكن التعبير عن ثبات الاختبار بالخطأ المعياري للقياس، الذي يوفر لنا تقديرا لمدى التباين الذي تقع فيه الدرجة الحقيقية للتلميذ في حالة تطبيق الاختبار عددا لانهايا من المرات. وفي مثالنا السابق عن ركل كرة القدم فإننا نتوقع أن نحصل على درجة مختلفة لنفس الفرد في كل مرة نطلب منه ركل كرة القدم، ويرجع هذا الاختلاف في الدرجات إلى الصدفة. ومع تكرار ركل الكرة عددا كبيرا من المرات فإننا نحصل على توزيع تكراري للدرجات. ولهذا التوزيع متوسط هو أفضل تقدير لدرجة التلميذ الحقيقية. ولهذا التوزيع أيضا انحراف معياري، يدل على درجة انتشار الدرجات حول متوسطها. ونظرا لأن هذا الانحراف المعياري هو الانحراف المعياري لأخطاء القياس يطلق عليه الخطأ المعياري للقياس. وإذا رسمنا مضلعا تكراريا لهذا التوزيع فإن شكل التوزيع يقارب شكل المنحنى الاعتدالي. فالأخطاء المعيارية موزعة اعتداليا إذ نجد بعض الأخطاء صغيرة للغاية، وبعضها الآخر كبير جدا. والانحراف المعياري لهذا التوزيع (الخطأ المعياري للقياس) يعطينا تقديرا لحد من

الخطأ الذي نتوقعه عند استخدام الاختبار في أي عملية قياسية. وليس لدينا في معظم الحالات تكرار لعمليات القياس، ولكننا نستطيع تقدير الخطأ المعياري من معامل ثبات الاختبار:

$$\text{خم} = \sqrt{1 - 11.9}$$

حيث	خم	الخطأ المعياري للقياس
	ع	الانحراف المعياري للاختبار
	11.9	معامل ثبات الاختبار

فإذا بلغ معامل ثبات اختبار ما ٩٦، وانحرافه المعياري ١٥ كان الخطأ المعياري لهذا الاختبار:

$$\text{خم} = \sqrt{1 - 96} = 15 = (2, 3, 0)$$

والخطأ المعياري للقياس مؤشر للتباين المتوقع للدرجات حول الدرجة الحقيقية للفرد. ويمكن لنا تفسير الخطأ المعياري كما نفسر أي انحراف معياري حول متوسط الدرجات. فالمعروف أن للانحرافات المعيارية حول المتوسط في المنحنى الاعتدالي توزيعاً محدداً يصف كيف تنتشر الدرجات حول المتوسط. ويمكن بنفس الطريقة تفسير الخطأ المعياري حول الدرجة باعتبار أن الخطأ المعياري يحدد المدى الذي يحتمل أن تقع فيه الدرجة الحقيقية للفرد. ويطلق على هذا المدى من الدرجات حدود الثقة. ويمكن التسليم بأن أخطاء القياس موزعة توزيعاً اعتدالياً حول درجة معينة وعلى طول قاعدة المنحنى. وبتطبيق خصائص توزيع المنحنى الاعتدالي يمكن القول أن الدرجة الحقيقية للفرد تقع بين $\pm \text{خم}$ وذلك في حدود ثقة قدرها ٦٨٪، وأنها تقع بين $\pm (2) \text{خم}$ ، وذلك في حدود ثقة قدرها ٩٥٪، وأنها تقع بين $\pm (3) \text{خم}$ ، وذلك في حدود ثقة قدرها ٩٩٪. فإذا حصل طفل على درجة قدرها ٩٠ في الاختبار وكان الخطأ المعياري ٣,٠، يمكن القول إن الدرجة الحقيقية لهذا الطفل تقع بين:

وذلك في حدود ثقة قدرها ٦٨٪	٨٧ و ٩٣
وذلك في حدود ثقة قدرها ٩٥٪	٨٤ و ٩٦
وذلك في حدود ثقة قدرها ٩٩٪	٨١ و ٩٩

والخطأ المعياري للقياس ومعامل الثبات طريقتان تكملان بعضهما البعض في

التعبير عن ثبات المقياس. والدقة التي يمكن بها تحديد الدرجة الحقيقية للفرد ترجع إلى حجم الخطأ المعياري، الذي يتأثر بالتالي بمعامل ثبات المقياس. وعندما نبحث عن اختبار لاستخدامه في جمع البيانات يجب أن يكون معامل الثبات مرتفعاً له ذا الاختبار مع انخفاض الخطأ المعياري له.

ثبات الاختبارات محكية المرجع:

ذكرنا من قبل أن الطرق المستخدمة في حسب معامل الثبات تتطلب انتشار درجات المقياس. ولذلك فهي لا تصلح إلا للاختبارات معيارية المرجع، ولذلك فإن استخدام هذه الطرق مع الاختبارات محكية المرجع، قد تؤدي إلى نتائج مضللة. فالاختبارات محكية المرجع تتصف بانخفاض انتشار درجاتها.

وعند استخدام الاختبارات محكية المرجع لقياس درجة الإتقان، وهو استخدام الأهم لهذا النوع من الاختبارات فإن اهتمامنا يتركز على مدى التناسق الذي يستطيع به الاختبار تصنيف الأفراد في مجموعتين: مجموعة أتقنت مخرجات التعلم ومجموعة لم تتقن تلك المخرجات. فإذا أجرينا صدورتين متكافئتين من الاختبار على مجموعة من الطلبة مثلاً فإننا نود أن نرى أن نتائج الصورتين تصنفان نفس الأفراد وفقاً للمعيار السابق (الإتقان وعدم الإتقان). إلا أن الاتفاق الكامل بين الصورتين مستحيل نظراً لعوامل النسيان وعوامل التخمين. ولكن إذا كان عدد الطلبة في مجموعة الإتقان بعد تطبيق الصورة الأولى أكبر بكثير من عددهم بعد تطبيق الصورة الثانية، فإننا لا نستطيع أن نحدد من هو الطالب الذي أتقن مخرجات التعلم. ولذلك يمكن تحديد ثبات الاختبارات محكية المرجع بدسابة نسبة من ظهر في مجموعتي الإتقان وعدم الإتقان بعد إجراء الصورتين المتكافئتين من الاختبار.

وهذه الطريقة في مقارنة الأداء في صورتين متكافئتين من الاختبار سببها نسياناً. فبعد إجراء كل من الصورتين على مجموعة من الطلبة نضع النتائج في جدول على النحو المبين في الجدول (٢١-٣).

وهذه الطريقة بسيطة ولكن يشوبها بعض العيوب، من أهمها:

- ١- يتطلب الأمر صورتين من الاختبار، إلا أن واقع الأمر أن هذا ليس عيباً حيث إن معظم الاختبارات محكية المرجع تتطلب وجود صورتين أو أكثر، للحاجة أحياناً في إعادة اختبار الطلبة الذين يفشلون في الاختبار مرة أو أكثر.

جدول ٢١-٣ تصنيف درجات ٣٠ طالبا بعد تطبيق صورتين الاختبار

الصورة أ		الصورة ب	
إتقان		عدم إتقان	
٢٣	١	٥	١
إتقان		عدم إتقان	

ويمكن حساب نسبة الاتساق بالمعادلة التالية:

عدد مجموعة الإتقان (في الصورتين) + عدد مجموعة عدم الإتقان (في الصورتين)

العدد الكلي للأفراد ١٠٠ ×

$$\text{نسبة الاتساق} = ١٠٠ \times \frac{٥ + ٢٣}{٣٠} = ٩٣\%$$

عن:

Linn & Gronlund, 2000, p. 130.

٢- يصعب تحديد نسبة الاتساق اللازمة لموقف معين، وكما هو الحال في معامل الثبات فكلما زاد المعامل كان ذلك أفضل، ولكن ما هو الحد الأدنى المقبول للثبات؟ ليس هناك إجابة محددة على هذا السؤال، ولكن يمكن لواضع الاختبار استخدام المنطق في اتخاذ قرار بذلك، وخاصة إذا كان الأمر يتعلق بمجرد إعادة الاختبار مرة أخرى إذا لم يكن المتعلم قد حقق الإتقان المطلوب.

ثبات أدوات الملاحظة:

معامل الثبات له أهميته في الأدوات التي تستخدم في الملاحظة مثل مقاييس التقدير أو قوائم المراجعة، وهي الأدوات التي تتطلب إعطاء تقديرات للأفراد أثناء ملاحظتهم بواسطة أفراد آخرين. فيجب في هذه الأحوال أن يحدد الباحث ثبات التقديرات، ودرجة الاتفاق في الأحكام بين الملاحظين المختلفين، وأبسط طريقة لتحديد ثبات التقديرات هو أن يقوم ملاحظان أو أكثر بتقدير نفس السلوك تقديرا

مستقلا عن بعضهم البعض، ثم نقوم بدساب معامل الارتباط بين تقديرات الملاحظين. ويطلق على معامل الارتباط الناتج ثبات تقديرات الملاحظين. وإذا كان السلوك الملاحظ محددًا تحديدًا إجرائيًا، وكان الملاحظون مدربين تدريبًا عاليًا على الملاحظة، فإننا نتوقع أن يكون معامل الثبات مرتفعًا جدًا.

فإذا قام مثلاً ملاحظان بتقويم أداء عدد من الطلبة باستخدام مقياس تقديرات تتراوح درجاته بين ١ (ضعيف جدًا) و ١٠ (ممتاز)، يمكن تقييم الثبات باستخدام معامل الارتباط بنفس الطريقة التي يستخدم فيها إعادة الإجراء أو تطبيق الصور المتكافئة. فالملاحظ الثاني يؤدي وظيفة الإجراء الثاني، أو الصورة المتكافئة للاختبار.

وإذا كانت التقديرات ١ أو صفراً (حدوث السلوك أو عدم حدوثه)، كما هو الحال في قوائم المراجعة، فإننا نستخدم معادلة معامل الاتفاق.

معامل الاتفاق:

يتم باستخدام هذه المعادلة تقويم درجة اتساق تقديرات الملاحظين، فإذا قام الملاحظان بتقدير أداء ٥٠ طالباً فإننا يمكن أن نحصل على نتيجة كالمبينة في الجدول ٢١-٤.

جدول ٢١-٤ ثبات تقديرات ملاحظين باستخدام معامل الاتفاق

الملاحظ الأول

التقديرات	١	صفر	المجموع
١	٣٥ (أ)	٥ (ب)	٤٠
صفر	٣ (ج)	٧ (د)	١٠
المجموع	٣٨	١٢	٥٠ (ن)

الملاحظ الثاني

يلاحظ في الجدول (٢١-٤) أن الملاحظين أعطيا ٣٥ طالباً تقدير ١، كما أنهما أعطيا ٧ طلاب تقدير صفر. في حين أن الملاحظ الأول أعطى ثلاثة طلاب تقدير ١، وهؤلاء أعطاهم الملاحظ الثاني تقدير صفر، كما أن الملاحظ الأول أعطى خمسة طلاب تقدير صفر، وهؤلاء أعطاهم الملاحظ الثاني تقدير ١. ومعامل الاتفاق هو نسبة الطلاب الذين حصلوا على نفس التقدير من الملاحظين، أي أن معامل الاتفاق

$$M = \frac{A + D}{N} = \frac{7 + 35}{50} = \frac{42}{50} = 84\%$$

حيث	م	معامل الاتفاق
أ	عدد الأفراد الذين حصلوا على تقدير ١ من الملاحظين	
د	عدد الأفراد الذين حصلوا على تقدير صفر من الملاحظين	
ن	مجموع عدد الأفراد الذين تمت ملاحظتهم	

وتشير هذه النتيجة إلى أن نسبة الاتفاق ٨٤، (أي ٨٤٪)، وهي في هذه الحالة مؤشر ثبات الملاحظات. وإذا اتفق الملاحظان في جميع الحالات تكون نسبة الاتفاق ١٠٠، (١٠٠٪)، وهو الحد الأقصى لمعامل الاتفاق.

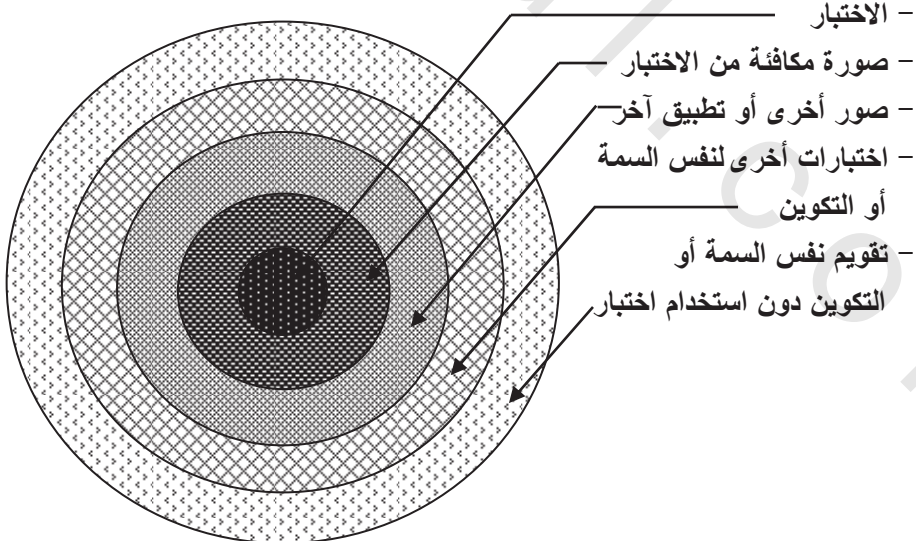
مقارنة بين الصدق والثبات:

الصدق صفة أكثر أهمية وأكثر شمولاً من الثبات، كما أنه أصعب في التحديد. ولذلك كثيراً ما نجد تقارير البحوث تنشر بيانات عن الثبات أكثر من البيانات عن الصدق. ويرجع ذلك إلى أن الحصول على بيانات الصدق ليس مباشراً كما هو الحال في بيانات الثبات. ويتطلب تقويم صدق أدوات القياس أحكاماً ذاتية أكثر من الثبات. ولابد من حسم بعض الأمور التي تتعلق بصلاحية محتوى الأداة، وكفاية المعايير، وتعريف الصفات الإنسانية، وتحديد المجال السلوكي الذي يقيسه الاختبار، والنظرية التي تساند محتواه، وغير ذلك من الأمور المتعلقة بصدق الاختبار. وتتطلب كل هذه الأمور أحكاماً من جانب واضع الأداة، والحصول على بيانات من مصادر متعددة (Ary et al., 1996).

أما الثبات فيمكن دراسته بشكل مباشر من بيانات الاختبار، فلا يتطلب الأمر الحصول على بيانات خارجية. كما أن القضايا الأساسية للثبات يمكن معالجتها بسهولة بالتحليل الرياضي، ويمكن صياغة نتائج معقولة عن كمية الخطأ في عبارات رياضية.

والمقياس الصادق ثابت أيضاً، ويحدد ثبات الأداة الحد الأعلى للصدق. وإذا كان معامل ثبات الاختبار صفراً، فإن ما نحصل عليه من درجات من هذه الاختبار درجات عشوائية تماماً ولا يمكن لها أن ترتبط بأي مقياس. ويزداد معامل الارتباط بين الاختبار والمحك كلما زاد معامل ثبات هذا الاختبار. إلا أن الاختبار كما ذكرنا سابقاً قد يكون ثابتاً دون أن يكون صادقا، ولذلك فإن الاختبار غير الصادق لا يستحق الجهد المبذول فيه لتحقيق ثباته.

وإذا نظرنا للصدق والثبات من زاوية مدى الاستفادة من درجة الاختبار فإننا نواجه في هذه الحالة بسؤال يتعلق بمدى إمكانية التعميم من درجة الاختبار. إننا إذا أجرينا صورتين من الاختبار في نفس اليوم فإن الارتباط بينهما يدلنا على الدرجة التي يمكن بها الوثوق في إمكانية التعميم من صورة من الاختبار للصور الأخرى. ويمكن استخدام مضمون هذه الفكرة لمعرفة إمكانية التعميم من درجة الاختبار إلى عينة المجال الذي يقيسه الاختبار. وإذا أجرينا الاختبار في يومين مختلفين فإن مصادر معلوماتنا عن الاختبار تأتي من هاتين المناسبتين. وإذا تم هذان الإجراءان بصورتين مختلفتين من الاختبار فإننا نحصل على معلوماتنا عن درجة التعميم من هاتين الصورتين بالإضافة إلى مناسبتتي الإجراء. وإذا أجرينا اختبارين مختلفين ولكنهما متشابهان في أن كلا منهما يقيس نفس السمة أو التكوين، فإن دائرة المعلومات التي نحصل منها عن درجة التعميم من اختبار لآخر تتسع باتساع مصادر التعميم. فنحن في هذه الحالة ننظر إلى مجال أوسع لا يتضمن فقط عينات من المهام المختارة طبقاً لجدول مواصفات مشترك، بل إننا ننظر إلى عدة جداول للمواصفات التي أعدها مؤلفون مختلفون. ونحن بهذه الفكرة الأخيرة ننقل من المفهوم العام للثبات إلى مفهوم الصدق. ونستطيع بالطبع التعميم من درجة الاختبار إلى مؤشرات أخرى لنفس الصفة نستخدم فيها وسائل غير اختبارية، مثل وصف الذات أو تقديرات الأقران، أو المدرسين. فهل يسمح لنا الاختبار بالخروج باستدلالات من هذه التقديرات؟ يقودنا هذا إلى التفكير في الثبات والصدق بطريقة تساعد على التعميم منهما للحصول على الاستدلالات المطلوبة.



شكل ٢١-١ مستويات التعميم من درجة الاختبار

ومعنى هذا أن مفهوم التعميم يشمل كلا من الثبات والصدق، مما يدل على أن الاختلاف بين المفهومين يرجع إلى مدى اتساع مجال التعميم. والإمكانية التي يحملها هذا المفهوم تقود إلى الاعتقاد بأن مفهوم التعميم قد يحل يومًا ما محل المفاهيم المختلفة المرتبطة بالصدق والثبات (Thorndike, 1997, p. 363). ويوضح شكل (٢١-١) فكرة التعميم. ويمكننا أن نعتبر موقف الاختبار مع ثلاثة مستويات التعميم ومنها يمكن أن نسأل أنفسنا إلى أي حد يمكن أن نعمم نتائج الاختبار على مستويات متتابعة تبتعد عن البؤرة المركزية التي تمثل الاختبار الأصلي.

الحد المقبول للصدق والثبات؟

درج بعض الباحثين على اعتبار الأداة صادقة وثابتة إذا كان معامل الصدق أو معامل الثبات (باعتبارهما معامل ارتباط) دالًا إحصائيًا. إلا أن ذلك لا يجب وز. ويرجع ذلك إلى مفهوم أساسي يرتبط بمذلول معامل الارتباط. فعندما يكون الارتباط بين متغيرين ٠.٦، مثلاً، فهذا يعني أن التباين المشترك بين المتغيرين هو ٣٦٪ (مربع معامل الارتباط). أي أن نسبة التباين التي يمكن أن نعزوها للمتغير الآخر هي ٣٦٪. لنفرض الآن أننا حصلنا من بيانات الاختبار على معامل ثبات قدره ٠.٢، وكان عدد أفراد العينة التي حسب عليها الثبات هي ١٠٠ فرد. هذه القيمة دالة إحصائياً عند مستوى ٠.٠١، فهل يمكن القول إن الاختبار ثابت لأن الارتباط دال إحصائياً عند مستوى ٠.٠١؟ بالطبع لا، لأن نسبة التباين المشترك التي ترجع إلى ارتباط الاختبار في نفسه لا تزيد على ٤٪، وهي نسبة ضئيلة للغاية، وبالتالي لا يمكن اتخاذها قرينة على ثبات الاختبار.

لهذا يجب البعد تماماً عن ذكر الدلالات الإحصائية لمعاملات الارتباط التي نستخدمها كمعاملات للصدق والثبات. إذن ما الحد المقبول لمعاملات الصدق والثبات؟ ليس هناك إجابة سهلة على هذا السؤال، إلا أنه يمكن أن نضع قائمتين للاسترشاد بهما عند تقويم معاملات صدق وثبات الاختبارات:

- ١- ألا تقل نسبة التباين المشترك عند حساب معامل ارتباط الاختبار في نفسه (معامل الثبات) عن ٥٠٪، ومعنى هذا أن الحد الأدنى للثبات يجب أن يكون ٧٠١٪ (الجذر التربيعي لنسبة التباين المشترك). ويرى جيلفورد (Guilford, 1956) ونانلي (Nannally, 1978) أن معاملات ألفا يجب ألا تقل عن ٧٠٪. إلا أن هذا الحد غير مقبول في كثير من الاختبارات كما سنرى في النقطة التالية. أما بالنسبة لمعامل الصدق، فنظراً لطبيعة المحك

فمن الصعب تحقيق مثل هذا المعيار. إلا أنه في جميع الأحوال لا يمكن قبول مستوى الدلالة كمؤشر لقبول معامل الصدق ويجب أن تكون نسبة التباين المشترك، كبيرة بدرجة تكفي للقول إن هناك عوامل مشتركة بين الاختبار والمحك.

٢- أن تكون معاملات الثبات أعلى ما يمكن، فإذا زادت عن ٨٠٪، فإن هذا مفضلاً عندما يتعلق الأمر بالمقاييس التي تتناول الاتجاهات والميول والنواحي الانفعالية. أما إذا كان معامل الثبات يتعلق باختبارات معرفية كاختبارات التحصيل والاستعدادات فهذه يجب أن تكون قريبة جداً من ٩٠٪، ويفضل أن تزيد على هذه القيمة، ويرجع هذا إلى اعتبارات تتعلق بطبيعة الأسئلة في مثل هذا النوع من الاختبارات. وبالنسبة لمعاملات الصدق فنظراً لأننا نحاول الربط بين الاختبار ومحك لا يمكن اعتباره اختباراً مكافئاً، فإن الحصول على معاملات عالية تزيد على ٦٠٪ أو ٥٠٪ قد يكون أمراً صعباً، ولكن في أي الحالات يجب الحصول على معاملات تعطينا نسبة لا بأس بها من التباين المشترك بين الاختبار والمحك.

نظرية الاستجابة للمفردة

كانت مناقشتنا للصفات الواجب توافرها في أدوات القياس موجهة بالدرجة الأولى نحو النظرية الكلاسيكية في القياس، وتقوم هذه النظرية على التسليم بأن أداء الفرد في المقياس يعكس درجة الفرد الحقيقية في الصفة التي نقيسها، والخطأ العشوائي الذي ينتشر حول الدرجة الحقيقية. ولقد ظهر اتجاه آخر في القياس يتصف بدرجة عالية من الثبات ويخلو في نفس الوقت من خطأ القياس، كما أنه أنسب لقدرة الفرد فمفرداته ليس صعبة جداً وليست سهلة جداً.

ولقد كان معظم القياس في علم النفس والتربية قائماً على النظرية الكلاسيكية، وهذه اختبارات جيدة إلا أنها معرضة لما يلي:

١- تعتمد معاملات الثبات وكذلك مؤشرات الصعوبة والتمييز للمفردات على العينة التي حصلنا منها على خصائص المقياس. ولذلك إذا استخدمنا باحث آخر نفس المقياس على عينة أخرى تمثل مجتمع آخر، فقد يتغير نتائج المقياس.

٢- قد يكون المقياس سهلاً جداً أو صعباً جداً بالنسبة لبعض الأفراد. وتحدث

مثل هذه الظروف فقد يكون المقياس تقديرا ضعيفا للدرجة الحقيقية لقدره الفرد التي نقيسها.

٣- تسلم النظرية الكلاسيكية في القياس أن كمية خطأ القياس واحدة من فرد لآخر من الأفراد الذين طبق عليهم المقياس. ولكن الواقع العملي يبين أن اختبارا ما قد يكون معامل ثباته أعلى (أي خطأ قياس أقل) للأفراد في أحد مستويات القدرة التي يقيسها الاختبار من أفراد آخرين في مستوى آخر من مستويات القدرة ذاتها.

٤- طبقا للنظرية الكلاسيكية في القياس تتحدد كمية خطأ القياس في مفردات الاختبار (في مقابل خطأ القياس الذي يمكن أن ننسبه لظروف الاختبار، وتقدير درجاته، وللأفراد الذين أخذوا الاختبار) بإجراء معاملات ارتباط بين أداء الأفراد في صورة مكافئة من الاختبار. ومن الناحية العملية فإنه من الصعب جدا بناء صورة مكافئة تماما من الاختبار، ولذلك فإن معامل الثبات الذي نحصل عليه بهذه الطريقة قد لا يكون فعلا ممثلا لمعامل ثبات الاختبار.

مسلمات نظرية الاستجابة للمفردة:

يمكن تجنب العيوب السابقة لبناء المقاييس في النظرية الكلاسيكية للقياس بناء مقاييس يقوم على نظرية الاستجابة للمفردة. ولذلك فإن تأثيرها على استخدام أدوات جمع البيانات في البحوث في تعاضد مستمر.

ويقوم استخدام نظرية الاستجابة للمفردة في بناء الاختبارات على المسلمات التالية:

- ١- يمثل أداء الفرد في الاختبار قدرة وحيدة.
 - ٢- يختلف أداء الأفراد في الاختبار طبقا لاختلاف كمية القدرة التي يمتلكونها.
 - ٣- يمكن تمثيل العلاقة بين متغيرات القدرة والأداء في المفردة بدالة رياضية.
- ويمكن وصف نظرية الاستجابة للمفردة باستخدام المثال التالي: لنفترض أننا نريد قياس القدرة على القراءة. ولنفترض أيضا أننا نختبر خمسة طلاب (أ و ب و ج و د و هـ) وتزيد قدرات كل منهم في تتابع بحيث يمتلك الطالب أ أقل قدرة في القراءة، أي أن جميع الطلاب الأربعة الباقيين أعلى قدرة من الطالب أ، وأن الطالب ب أعلى قدرة من الطالب أ، ولكنه أقل قدرة من الطلاب الثلاثة الباقيين، وهكذا.

لنفترض الآن أننا أعطينا المفردة '١' في إحدى اختبارات القراءة، وهي

مفردة سهلة جدا للطلاب الخمسة، ولذلك أجابوا جميعا عليها إجابة صحيحة. وتدلنا هذه النتيجة على أن الطلاب الخمسة يمتلكون حدا أدنى من القدرة على القراءة. بعد ذلك أعطينا المفردة '٢' التي أجاب عليها الطلاب ج ود وه إجابة صدحيحة إلا أن الطالبين أ وب أجابا إجابة خاطئة. وهذا يعني أن المفردة رقم '٢' أصعب من المفردة '١' ويمكن استخدامها للتمييز بين الطالبين أ وب، والطلاب ج ود وه. ثم نعطي للطلاب الخمسة بعد ذلك المفردة رقم '٣' التي يجيب عليها إجابة صدحيحة الطالب ه فقط. إننا الآن نعلم أن هذه المفردة الأخيرة ذات مستوى أعلى في القدرة على القراءة من المفردتين ١ و ٢، وأنها تميز قدرة القراءة للطلاب ه عن الطلاب الأربعة الآخرين.

مزايا نظرية الاستجابة للمفردة:

لهذه الطريقة في بناء المفردات وتحليلها ميزتان مهمتان:

- ١- أنها تمدنا بمعلومات عن كمية القدرة على القراءة التي تقيسها المفردة.
- ٢- أن أداء الطلاب للمفردة يمدنا بمعلومات عن كمية القدرة على القراءة التي لديهم.

إننا حتى الآن وضعنا ثلاث مفردات لكل منها مستوى مختلفا في القدرة على القراءة. لنفرض أننا وضعنا عشر مفردات أخرى تمثل المستويات في كل مستوى من مستويات القدرة الثلاث (١، ٢، ٣) بنفس الطريقة، أي أننا نضع عشر مفردات في المستوى الأول (١)، الذي يستطيع أن يجيب عليه جميع الطلاب. وكذلك بالنسبة للمستويين ٢، ٣. وبذلك نحصل على بنك للمفردات يتميز بما يلي:

- ١- نستطيع أن نفصل عملية القياس لنتناسب مع مستويات القدرة المختلفة للطلاب. مثال ذلك أننا نعطي الطالب عدة مفردات من المستوى الأول (١) ليجيب عليها، وإذا لم يستطع أن يجيب على أي منها فلا حاجة بنا إلى إحباطه بإطالة زمن الاختبار، بأن نعطي مفردات أخرى من المستوى الثاني والثالث.

- ٢- نستطيع بناء اختبارات متكافئة كثيرة لها نفس مستوى الصعوبة. مثال ذلك أن يمكن أن نحصل من بنك الأسئلة على عينة عشوائية تتكون من مفردتين من كل مستوى من مستويات القراءة الثلاثة لبناء اختبار مكون من ست مفردات. ونستطيع أن نكرر هذه العملية ونحصل على اختبار مكافئ له نفس مستوى الصعوبة.

٣- نستطيع أن نقلل من أخطاء القياس لفرد معين بإعطائه مجموعة من الأسئلة في المدى الذي يستطيع الإجابة عليه إجابة صحيحة. مثال ذلك إذا كان لدينا طالب يشبه الطالب ج في مثالنا السابق فإننا نستطيع أن نعطيه أسئلة كثيرة من المستوى الثاني لنحدد قدرة الطالب بدقة (لأن زيادة عدد المفردات في الاختبار يقلل من خطأ القياس). ولا داعي لتقديم مفردات من المستوى الأول التي سوف تكون سهلة جداً على الطالب أو المستوى الثالث الذي تعتبر مفرداته صعبة جداً عليه.

الأساس الرياضي لنظرية الاستجابة للمفردة

من الناحية العملية لا نستطيع أن نعرف القدرة الحقيقية للقراءة لدى الأفراد. فهي سمة كامنة أي سمة غير ظاهرة ولا يمكن ملاحظتها مباشرة، بل نفترض وجودها لنشرح السلوك الذي نلاحظه. مثال ذلك أننا إذا لاحظنا فرداً يقرأ صحيفة وبجيب أسئلة عليها بفهم جيد، فإننا نستخلص أن هذا الشخص له قدرة قرائية جيدة، لأننا نستطيع أن نلاحظ سلوك القراءة والإجابة على الأسئلة، وليس القدرة التي تقوم عليها.

وتستخدم نظرية الاستجابة للمفردة نماذج رياضية للتعرف على العلاقة بين السلوك الملاحظ (الأداء في الاختبار) والقدرة التي تمثل هذا الأداء. ومنحنى خاصة المفردة دالة رياضية تدلنا على العلاقة بين الأداء في مفردة الاختبار والقدرة التي تمثلها. ويبين الشكل (٢١-٢) منحنى خاصة مفردتين. ويمكن أن نذكر عدة ملامح لهذه المنحنيين:

١- يمثل الإحداثي الأفقي (السيني) القدرة الكامنة التي تقيسها المفردة ويمثلها الرمز (θ) . ويحدد مستوى القدرة وحدات من الدرجة المعيارية التي تكون قيمها سالبة بالنسبة للأفراد الأقل قدرة، وموجبة للأفراد الأعلى قدرة (انظر الفصل الحادي والعشرين).

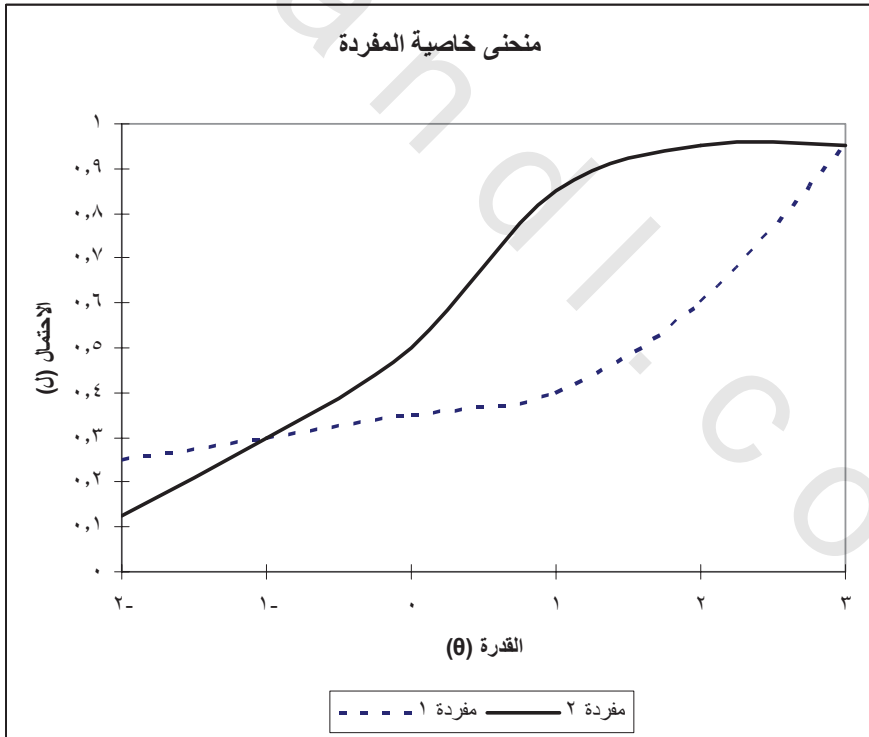
٢- يمثل الإحداثي الرأسي (الصادي) احتمال (ل) الإجابة على المفردة إجابة صحيحة. وبمعنى أدق يمكن تفسير (ل) على أنها احتمال أن الفرد عند مستوى معين من القدرة سوف يجيب على المفردة إجابة صحيحة عند اختياره عشوائياً من بين عينة من الأفراد الذين لهم نفس المستوى.

٣- يلاحظ أن المفردة رقم ٢ أكثر صعوبة بالنسبة للطلاب في معظم مستويات القدرة لأن احتمال الإجابة عليها إجابة صحيحة (ل) أقل عند معظم نقاط

الإحداثي السيني الذي يمثل القدرة (θ) .

٤- للمفردة رقم ١ قدرة أكبر على التمييز، وهذا يعني أن هناك فرقاً أكبر في الأداء في هذه المفردة بين الطلاب ذوي القدرة العالية والطلاب ذوي القدرة المنخفضة. فإذا نظرنا مثلاً للمفردة رقم ١ نجد أن الاحتمال قدره ٤٥، أن الأفراد عند مستوى قدرة (صفر) يجيبون على المفردة إجابة صحيحة، وأن ٨٠، عند مستوى قدرة ١ يجيبون عليه إجابة صحيحة. أما بالنسبة للمفردة رقم ٢ فإن هناك احتمالاً قدره ٣٥، أن الأفراد عند مستوى قدرة (صفر) سوف يجيبون عليه إجابة صحيحة، إلا أن هناك احتمالاً أعلى بقليل (٤٥)، أن الأفراد عند مستوى قدرة ١ سوف يجيبون عليه إجابة صحيحة. ونستطيع التعبير عن هذه النتائج بالقول إن المفردة رقم ١ تعطينا معلومات ذات قيمة أعلى من المفردة ٢ وذلك بالنسبة لمدى القدرة الذي نحن بصدد.

٥- وحتى بالنسبة للأفراد عند أقل مستويات القدرة (أقصى نقطة على يسار الإحداثي السيني) يمتلكون احتمالاً أعلى من الصفر في إجابة المفردتين إجابة صحيحة. وهذا يعكس حقيقة أن المفردتين تسمحان بالتخمين.



شكل ٢١-٢ منحنى خاصية المفردة لمفردتين

ويستخدم واضعو الاختبارات عددا من نماذج منحنيات خاصة المفردة لتناسب احتمالات الأداء في المفردات التي يتكون منها الاختبار للتمييز بين مستويات القدرة المختلفة. فهناك عدد من النماذج المستخدمة التي بتوقف على طريقة تقدير الدرجات (صح/خطأ، أو درجة جزئية، أو الاختيار من متعدد) وإذا ما كانت المفردات تسمح بالتخمين أو إذا ما كانت قوة تمييزها تختلف (كما هو الحال في المفردتين ١ و ٢ بالشكل ٢١-٢). ومن أهم النماذج المستخدمة نموذج راش (المنسوب إلى العالم الذي وضعه وهو جورج راش) نظرا لسهولة النسبية وقوته. ويميل معظم مستخدمي نموذج راش إلى حذف بعض المفردات القوية لأنها لا تناسب النموذج، في حين أن بعض الباحثين الآخرين الذين لا يتبعون نموذج راش كثيرا ما يبقون على مثل هذه المفردات وذلك بالبحث عن منحنيات خاصة المفردة التي تصف البيانات رياضيا.

والمغيرات التي يتم إدخالها في المعادلات الرياضية التي تصف منحنيات خصائص المفردة يمكن استخدامها في معادلات أخرى لتصف الكمية التي يمكن أن تضيفها كل مفردة والاختبار ككل لدقة القياس وللثبات. (ويطلق على هذه المعادلات دالة معلومات المفردة، ودالة معلومات الاختبار). وبمعنى أبسط إذا احتوى الاختبار على مفردات معظمها صعب على الأفراد الذين يأخذون الاختبار عند مستوى معين من القدرة، فإن دقة القياس تنخفض بالنسبة لهؤلاء الأفراد. وإذا احتوى الاختبار على عدد قليل من المفردات ذات مستوى مناسب من الصعوبة (أي مفردات تبلغ قيمة 'ل' لها حوالي ٥٠، بشكل عام)، فإن ثبات الاختبار يكون ضعيفا. والاختبار المثالي من منظور نظرية الاستجابة للمفردة هو الاختبار الذي تناسب مفرداته قدرة الفرد: أي لا تكون مفرداته سهلة جدا أو صعبة جدا، مع وجود عدد من هذه المفردات للإقلال من خطأ القياس.

وإجراءات بناء الاختبار باستخدام نظرية الاستجابة للمفردة إجراءات معقدة ولذلك لا يستخدمها سوى عدد قليل من الباحثين لقياس متغيرات بدوهم. إلا أن عدد هؤلاء في تزايد مستمر، وإذا رغب الباحث في استخدام أو بناء اختبار من هذا النوع، فمن الضروري أن يتأكد أنه قد ألم بمبادئ نظرية الاستجابة للمفردة على الأقل.