

السنة الأولى ماستر - تخصص علم النفس التربوي - مادة القياس النفسي

والتربوي - أستاذ المادة: بوسالم عبد العزيز

عنوان المحاضرة: قياس التحصيل الدراسي

مدخل :

تزايد عدد الاختبارات التحصيلية المقننة المنشورة التي استخدمت البنود الموضوعية حتى وصل إلى المئات، وتضمنت الاختبارات مادة دراسية واحدة وبطاريات تشتمل مواد دراسية متعددة، واختبارات تشخيصية وتنبؤية، كما ازداد الاهتمام ببناء بنود موضوعية تقيس الفهم والاستنباط والتحليل، وغيرها من العمليات المعرفية العليا، وكذلك الأهداف التعليمية في مجالات مختلفة، وشهدت الثلاثينات من القرن الماضي، فتصميم آلة تصحيح الاختبارات الموضوعية، لتيسر تطبيق وتصحيح الاختبارات المقننة على نطاق واسع خاصة مع التضارب الذي كان يلاحظ في تقديرات المعلمين لتحصيل المتعلمين باستعمال الاختبارات المقالية، مما شجع على التوجه إلى الاختبارات الموضوعية خاصة عندما طوّرت بنودها لتقيس المستويات المعرفية العليا من القدرات العقلية.

لم يقتصر استخدام الاختبارات المقننة على الأغراض التربوية، وإنما استخدمت أيضا في انتقاء العاملين بالمصانع والمؤسسات الإنتاجية المختلفة والقوات المسلحة. بهدف اختيار الأفراد المناسبين لأداء مهام معينة بناء على نتائج تطبيق اختبارات تقيس مختلف جوانب الشخصية.

كما أن التوسع في استخدام الاختبارات ومشاركة علماء النفس في بناء الاختبارات المقننة ساهم في تطور حركة الاختبارات سواء في علم النفس أو التربية على حد سواء". فالاختبارات التحصيلية المقننة التي تقيس أهدافا تعليمية شاملة أو متسعة أصبحت تشبه اختبارات الذكاء العام في الخصائص الموضوعية التي تتوفر عليها.

1- مفهوم التحصيل الدراسي

يعد التحصيل الدراسي من المفاهيم القاعدية في الميدان التربوي على اعتبار أنه يبقى الهدف الأساس لكل الأنظمة التربوية وفي جميع المجالات، وحتى نقرب أكثر من معنى التحصيل الدراسي نتناول تعاريف بعض الباحثين لنقوم بتحليلها للوصول إلى معنى شامل يمكن تبنيه في هذه الدراسة.

عرّف التحصيل في موسوعة علم النفس (1994) بأنه "إنجاز عمل ما أو إحراز التفوق في مهارة أو مجموعة من المعلومات"، وعرّفه بأنه "متوسط ما يحصل عليه الطالب أو الطالبة من درجات في أحد الوحدات الدراسية، أو في مجموعة من الوحدات الدراسية والتي تقاس في هذه الحالة بالمعدل التراكمي.

كما عرف على أنه "المعلومات والمهارات المكتسبة من قبل المتعلمين كنتيجة لدراسة موضوع، أو وحدة دراسية محددة"

والتحصيل الدراسي هو تقديم معلومات وصفية تبين مدى ما حصله المتعلم بطريقة مباشرة من محتوى المادة الدراسية، وذلك من خلال الاختبارات التي يطبقها المعلم على المتعلمين على مدار العام الدراسي لقياس مدى استيعابهم للمعارف والمفاهيم والمهارات التي لها علاقة بالمادة الدراسية في وقت معين أو في نهاية مدة تعليمية معينة.

من خلال التعاريف السابقة للتحصيل الدراسي نستنتج أن عملية التعليم والتعلم المقصودة هدفها الأساسي هو تحقيق تحصيل دراسي لمجموعة من المعارف عند درجة مقبولة. فالتحصيل الدراسي نتيجة لعملية مقصودة يتميز بأنه:

- درجة الاكتساب التي يحققها المتعلم أو ما يصل إليه من نجاح في مواد دراسية معينة، هذه الدرجة يمكن تقديرها بواسطة اختبارات تحصيلية.

- مستوى من الإنجاز الفعلي يصل إليه المتعلم يظهر في شكل كفاءات يمكن الكشف عنها أو قياسها بواسطة اختبارات بنيت من طرف المعلم أو اختبارات مقننة من طرف خبراء في بناء الاختبارات.

ويمكن اعتبار أي اختبار أنه مقياسا للتحصيل إذا كان يقيس نواتج برنامج تعليمي أو تدريبي معين، أو خبرات تعليمية واسعة، لذلك يمكن القول بأن الاختبارات التحصيلية تقيس نتائج التعلم الذي يتم تحت شروط محددة بدرجة نسبية، وفي ظروف يمكن التحكم فيها داخل القسم، وأن التركيز يكون على حاضر المتعلم أو ماضيه أي ما تعلمه في إطار زمني معين.

فاختبارات التحصيل تستخدم في تقويم أداء الفرد عند انتهائه من تعلم مادة دراسية معينة، موضوعا دراسيا أو التدريب على مهارة معينة، وهي تبين ما يستطيع الفرد أدائه، أي أنها تركز على الحاضر أو تساعد في التنبؤ بالأداء في المستقبل (في حالة الاختبارات التنبؤية)، أي ما يمكن للفرد أدائه مستقبلا إذا ما أتاحت له الظروف المناسبة.

فقد تستخدم درجة تحصيل المتعلم في اختبار الإحصاء للتنبؤ بنجاحه اللاحق في القياس النفسي. أو تستخدم اختبارات التحصيل في بعض الوحدات مثل علم النفس المرضي أو الفيزيولوجي للتنبؤ بأداء الطالب في تخصص علم النفس المرضي مستقبلا. وهذا ما يتم عند تحديد معايير التوجيه للتخصصات الجامعية التي تسبق بجدع مشترك. حيث تعتمد نتائج التحصيل في بعض المواد للتنبؤ بإمكانية نجاح الطالب في تخصصات معينة.

الشيء نفسه يحدث في عملية التوجيه الجامعي بعد نجاح المتعلم في شهادة البكالوريا حيث يعتمد على نتائجه في مختلف المواد ليوّجه نحو تخصصات جامعية معينة، فالذي يتحصل على علامات جيدة في مادتي الرياضيات والفيزياء مثلاً يوجه لدراسة العلوم الدقيقة انطلاقاً من بعد تنبؤي وأنه يملك استعداداً يسمح له بالنجاح في هذا التخصص.

2- افتراضات يرتكز عليها قياس التحصيل الدراسي

يعد قياس التحصيل الدراسي كسمة معرفية افتراضية تعرف من خلال مؤشرات تظهر في استجابة المتعلم لبنود الاختبار، يعد بمثابة الخطوة الأساسية لتحقيق موضوعية القرارات التربوية بمختلف أنواعها ومستوياتها، وقياس التحصيل يستند إلى افتراضات أساسية هي:

- أن يقيس الاختبار المعد لقياس التحصيل نطاقاً سلوكياً يمكن تحديده بدقة، بمعنى أن ينطلق واضع الاختبار من أهداف تربوية تصاغ في عبارات إجرائية يمكن قياسها والتحقق منها، فبدون هذه الإجرائية لا يمكن معرفة هل تحققت الأهداف أم لم تتحقق، خاصة أن محتوى الاختبارات المقننة متسع بالمقارنة مع الاختبارات المقالية التي يعدها المعلمون عادة، مما يدفع إلى القول بضرورة الاعتماد على هذا النوع من الاختبارات.

- يوضع الاختبار في أساسه من أجل قياس أهداف تربوية واضحة في شكل مخرجات (يمكن أن تسمى في بعض الأحيان كفايات وقد يدخل قاموس التربية مفاهيم جديدة مستقبلاً ولكن في النهاية هي أهداف يسعى النظام التربوي إلى تحقيقها بغض النظر عن تسميتها)، وهذا يتطلب التحديد المسبق لهذه الأهداف لمعرفة ماذا استطاع المتعلم تحقيقه فعلاً. فالاختبار يقيس الأهداف المتعلقة بالمادة الدراسية، بمعنى أن يكون محتوى بنوده على درجة عالية من صدق محتواه، أي أن يقيس المهارات التي أعد من أجلها وليس مهارات عارضة.

- قبل بناء وتطبيق الاختبار لابد أن تتاح لجميع المتعلمين فرص تعلم المجال المعرفي الذي يقيسه الاختبار، خاصة عند استخدام الاختبارات المقننة، (صلاح الدين محمود علام 2003، ص 309)، فلا يمكن تطبيق اختبار على متعلم لم تمنح له فرصة تعلم نطاق سلوكي أو محتوى تربوي معين، فاختبار التحصيل يقيس أصلاً ما استطاع المتعلم اكتسابه بعد المرور بعملية تعليمية تعلمية معينة.

والهدف من بناء الاختبارات التحصيلية هو قياس مدى استيعاب المتعلمين لبعض المعارف والمفاهيم والمهارات المتعلقة بالمادة الدراسية في وقت معين أو في نهاية فترة تعليمية معينة.

فالتحصيل الدراسي يشير إلى الوضع الراهن لأداء الفرد أو ما تعلمه واكتسبه بالفعل مع اختلاف نمط السلوك المتعلم أو نوع المهارة المكتسبة من معارف ومهارات ترتبط ببرنامج تعليمي معين، ويمكن معرفة درجة الاكتساب عند كل متعلم من خلال الاختبارات التحصيلية.

3- اختبارات قياس التحصيل الدراسي

تعد اختبارات التحصيل الدراسي أكثر الاختبارات انتشاراً واستعمالاً ما بين الاختبارات التي تقيس مختلف جوانب الشخصية، وهذا نظراً لوجود عدد متزايد من المتعلمين يتطلب معرفة درجة تحصيلهم تطبيق مختلف الاختبارات التي يقوم ببنائها المدرسون أو خبراء القياس، ويمكن النظر إلى هذه الاختبارات وتصنيفها وفق معايير عدة. فيمكن تصنيفها حسب طريقة التطبيق إلى جماعية وفردية، وحسب طبيعة بنودها إلى مقالية وموضوعية، والتصنيف الأكثر استعمالاً وأهمية هو تصنيفها حسب طريقة تفسير النتائج إلى اختبارات مرجعية الجماعة واختبارات مرجعية المحك.

3-1- مفهوم الاختبار كوسيلة للقياس

تشير أدبيات القياس النفسي إلى وجود تعريفات متعددة قدمت للاختبار النفسي، إذ ترى أنستازي (Anastasi) "أن الاختبار هو أداة قياس موضوعية ومقننة للحكم بصورة كمية على عينة من السلوك".

ويرى براون (Brown,1978) أن الاختبار هو إجراء منظم لقياس عينة من السلوك الأفراد، أما كرونباك (Cronbach,1970) فيرى أن الاختبار النفسي أو التربوي هو أسلوب منظم لملاحظة ووصف خاصية أو أكثر من خصائص الفرد استناداً إلى ميزان عددي أو نظام تصنيفي متفق عليه.

يتضح مما سبق أن الاختبار التحصيلي هو أحد أدوات القياس والتي تقدم معلومات عن مدى اكتساب واستيعاب الفرد للخبرات المعرفية بمستوياتها المختلفة والمهارات التي تم دراستها، وفي ضوء ما تسفر عنه نتائج الاختبارات تتم عملية التقويم، وعليه فإن المشكلات المتعلقة بالاختبارات التحصيلية تصدق على التقويم التربوي.

3-2- الاختبارات المقالية كوسيلة لقياس التحصيل الدراسي

تعتبر الاختبارات التحصيلية المقالية الوسيلة الأكثر انتشاراً بين المعلمين في مختلف الأنظمة التربوية في العالم، كما تعتبر الأولى من حيث الظهور التاريخي فقد استعملت في تقويم المتعلمين قبل أن تبدأ في الظهور الاختبارات الموضوعية، وعلى الرغم من الانتقادات التي توجه لهذا النوع من البنود الاختبارية، والتي تمنح الفرصة للمجيب بأن ينشأ إجابة من إنتاجه وفق أسلوبه الخاص، مما يجعلها لا تتصف

بالموضوعية وأشار كل من ثورن دايك وهاجن إلى أن الاعتماد على البنود المقالية المفتوحة في الحكم على مستوى المتعلم من شأنه أن ينقص من قيمة القرارات نظرا لكون هذه الاختبارات لا تحقق المستوى المطلوب من معدلات الصدق والثبات وعدم شموليتها للمحتوى الدراسي، كما أنها تفتح المجال لتدخل ذاتية المصحح في منح الدرجات نظرا لغياب معيار دقيق وواضح يمكن الاعتماد عليه في تقييم إجابات المتعلمين.

وبالرغم من ذلك فإن الاختبار المقالي يمكن أن يقيس مستويات معرفية عليا قد لا يستطيع الاختبار الموضوعي الوصول إليها، وعليه لا يجب أن نتخلى عنها تماما وتختفي من ممارستنا التربوية، ولكن مطلوب أن نأخذ في عين الاعتبار هذه السلبيات عند استعمالها في قياس مختلف الخصائص والقدرات عند المتعلمين ونحاول تجنبها أو التقليل منها لتحقيق تنوع إيجابي لوسائل القياس (الاختبارات).

ولما كان الاختبار المعد لهذه الدراسة هو من نوع الاختبارات الموضوعية، وبالضبط من الاختيار من متعدد فلن نفصل كثيرا في الاختبارات المقالية، ونكتفي بالتركيز على اختبارات الاختيار من متعدد.

3-3- الاختبارات الموضوعية كوسيلة لقياس التحصيل الدراسي

لقد ظهرت الاختبارات المقالية كوسيلة من وسائل تقدير مستوى التحصيل الدراسي كرد فعل أو كنتيجة للانتقادات التي وجهت للاختبارات المقالية ولعدم موضوعيتها، بالرغم من وجود دعوة تربوية تشير إلى أن الهدف من الاختبار هو الذي يحدد طبيعة البنود بين المقالية والموضوعية.

والاختبارات أو البنود الموضوعية تتطلب من المتعلم تمييز الإجابة الصحيحة ووضع علامة عليها دون الحاجة إلى الكتابة أو الشرح، فالمتعلم يختار الإجابة الصحيحة مما يسلم له، وهذا يضيف على عملية تصحيح الاختبار موضوعية بحيث لا تتدخل ذاتية المصحح في التأثير بالزيادة أو النقصان على درجة الفرد، نظرا لكون الإجابة النموذجية معروفة مسبقا، وكل سؤال له إجابة صحيحة واحدة فقط إما يعرفها المتعلم أو لا يعرفها.

والاختبارات الموضوعية أنواع، سوف لن نفصل فيها جميعا يتم عرضها باختصار مع التركيز نمط الاختيار من متعدد لكون اختبار الدراسة الحالية من هذا النوع.

3-3-1- بنود الصواب و الخطأ (الاختبار الثنائي)

يطلب في هذه الاختبارات من نوع الصواب والخطأ من الفرد أن يحدد مفردة أو مفردات معطاة بصواب أو خطأ، وهذا النوع من مفردات الاختبار عادة ما تسبق بتوجيهات حول ما إذا كانت العبارة صحيحة كما هي معطاة، اكتب الكلمة "صواب" في المكان الخالي التالي للمفردة، وإذا كانت المفردة خطأ، أوجد بديلا للكلمة أو العبارة التي تحتها خطأ والتي تجعلها صوابا، ثم اكتب البديل في الفراغ الذي يلي المفردة.

ومن الأمثلة على بنود الصح والخطأ ما يلي: حدّد العبارات الصحيحة مما يلي:

خطأ	صحيح	- تتمدد الأجسام بالحرارة.
خطأ	صحيح	- بياجى هو مؤسس مدرسة التحليل النفسى.
خطأ	صحيح	- مفهوم الاختبار أعم من مفهوم المقياس.

وقد تتطلب توجيهات مفردات اختبار الصواب والخطأ من المتعلم استجابة تُعطى في ثلاثة أشكال، استجابة بوضع دائرة حول الإجابة الصحيحة، أو كتابة الاستجابة. وإما تعديل استجابة الصواب والخطأ، أي تعديل أو تصحيح الخطأ.

3-3-2- بنود الملاءمة

تتطلب بنود الملائمة من المتعلم أن يزاوج بين كلمة واحدة، أو شيء، أو رمز أو سلوك والاستجابة المرتبطة بها والمعطاة للمتعم. ومن الأمثلة على هذا النوع:

إليك مجموعة من المعطيات الكمية والمطلوب ربط كل معطى بمستوى القياس المناسب:

- عدد التلاميذ في قسم دراسى.	- التصنيفى
- درجات التلاميذ في مادة دراسية.	- الترتيبى
- ذكاء مجموعة من الأطفال.	- المسافات
- أرقام لقاءات الدراسة.	- النسبى

في هذه الحالة يتطلب من المتعلم أن يصل كل عبارة من عبارات العمود الأول مع العبارة المناسبة في العمود الثانى بدون أن يأتي بأي معطيات علمية من عنده بل يكفي فقط باستنتاج الجواب الصحيح . ويمكن أن نلاحظ أن بنود الملائمة يمكن اعتبارها مجموعة مختلفة من مفردة الاختيار من متعدد؛ أو بتعبير آخر، مجموعة جذور الاختيار من متعدد مرتبطة بمجموعة مفردة من الاستجابات الممكنة.

3-3-3- بنود التكملة

تتطلب مفردات اختبار التكملة جمل غير كاملة يتم تكملتها من طرف المتعلم الذي يطبق عليه الاختبار، وهناك حالتين للتكملة، فإما أن تعطى الكلمات التي يستعملها المتعلم في التكملة تعطى له وهو يختار المناسبة منها لوضعها في الفراغ المناسب (وهذا النوع هو الغالب)، وإما أن يطلب من المتعلم استحضار الكلمات أو المفاهيم المطلوب أن يملأ بها الفراغ من عنده، ويستحسن النوع الأول مع تجنب كثرة الفراغات حتى لا تفقد العبارات معناها بالنسبة للمتعم. ومن أمثلة هذا النوع من البنود.

- عندما يطبق الباحث الاختبار ويعيد تطبيقه على نفس المجموعة، فمعامل الثبات الذي يتحصل عليه يسمى ثبات..... وهو يصلح مع الاختبارات..... ولا يصلح مع اختبارات..... لأنها تتأثر بالتذكر.

3-3-4- بنود الإجابة القصيرة

تتطلب بنود الإجابة القصيرة من المتعلم أن يجيب على السؤال بكلمة أو عبارة أو عدد، أو حتى جملة قصيرة، وهذه البنود مفيدة بوجه خاص في قياس تذكر المعلومات والمعارف، ويتطلب أن يلتزم من يقوم ببناء الاختبار في جعل بنوده بطريقة تسمح بالحصول على إجابات يكون من السهل تقرير هل هي صواب أم خطأ. حتى تحافظ على موضوعيتها ولا تترك المجال لتدخل ذاتية المصحح في تقدير مدى صحة أو خطأ الإجابة. وإن لم تبذل عناية كبيرة لهذه النقطة فإننا سنحصل إما على إجابات غامضة أو سيتم استخدام مفردات غير مهمة، ولكنها سهلة كتابتها.

ومن الأمثلة على البنود التي تتطلب الإجابة القصيرة.

- ما هو عدد مستويات القياس في علم النفس؟

- كيف يسمى نموذج التفسير الذي يعتمد على متوسط الجماعة؟

- ماذا نسمي القيم العددية التي نتحصل عليها من الجداول الإحصائية؟

3-3-5- بنود الاختيار من متعدد

يعتبر هذا النوع من أفضل أنواع الاختبارات وأكثرها صدقا وثباتا واستخداما في الاختبارات التحصيلية في وقتنا الحاضر، وبنود الاختيار من متعدد هي بنود يختار فيها الفرد الاستجابة الصحيحة أو الأفضل من عدة اختيارات، حيث تتكون من جذر في شكل عبارة غير كاملة أو سؤال أو رسم تخطيطي أو حساب أو تعليمات مطلوب تنفيذها، وعدد من الاستجابات (اختيارات أو بدائل)، وإحدى هذه الاستجابات المقترحة صحيحة، بينما باقي الاستجابات غير صحيحة أو متداخلة مع الاستجابة الصحيحة، ويطلب من المتعلم اختيار الإجابة الصحيحة، ويستخدم هذا النوع من البنود لقياس معرفة الحقائق العلمية واستيعابها وتطبيق المعرفة العلمية في مواقف جديدة، ويتصف هذا النوع من الأسئلة بخصائص هي:

تغطية أكبر نسبة ممكنة من المحتوى الدراسي أو التدريبي على اعتباره يشتمل عدد كبير من البنود التي من شأنها تغطية مختلف جوانب المحتوى مما يحقق لها درجة عالية من صدق المحتوى.

- سهولة التصحيح فهي لا تتطلب جهدا كبيرا ووقتا طويلا إذا ما قارناها بالاختبارات المقالية بحيث يكون مفتاح التصحيح دليل واضح يتبعه المصحح.

- صعوبة الإعداد والتصميم، فمن خلال المراحل التي يمر بها بناء هذا النوع من الاختبارات والخصائص التقنية التي يتطلب توفرها في البنود فإنه يحتاج إلى خبرة وتدريب قصد إعداد اختبار يتوفر على الشروط الموضوعية.

- تمتعها بدرجة عالية من الثبات، فالدرجات لا تتأثر بذاتية المصحح لأن الإجابات الصحيحة محددة مسبقاً وبشكل دقيق، فلو أعدنا تصحيح الاختبار عدة مرات من عدد من المصححين فسوف نتحصل على نفس النتيجة، بحيث الإجابة النموذجية محددة وواحدة.

ويمكن صياغة بند من نوع اختيار من متعدد كما يلي:

تدرج مستويات القياس النفسي من الاسمي إلى النسبي بناء على:

أ- المعالجات الإحصائية المستعملة في كل مستوى.

ب- الخصائص الشكلية لكل مستوى.

ج- العمليات الحسابية الممكنة في كل مستوى.

د- عقد طبيعة المعطيات في كل مستوى..

بنود اختبار الاختيار من متعدد هي حالياً الأكثر استخداماً في ميدان الاختبارات التحصيلية المقننة أكثر من الأنواع الأخرى من المفردات أو الاختبارات، والصيغة المذكورة لمفردة الاختيار من متعدد مع الجذر تسأل أو تدل ضمناً على سؤال مباشر وهي توفّر طريقة مناسبة وحقيقية وطبيعية لاختبار تحصيل الفرد، وهناك مجموعة من الخصائص أو الشروط الواجب توفرها في بنود اختبار الاختيار من متعدد حتى يحقق هذا النوع الغرض الذي وضع من أجله، وأهم هذه الشروط هي:

أ- استخدام جذر كافياً

يطلب كتابة الجذر بكلمات واضحة ومختصرة بقدر الإمكان، كما يجب أن يتضمن معلومات كافية لكي يكون منطقياً مع الإجابة المطلوبة، بحيث يعرف الفرد المتعلم قبل أن يقرأ الاختيارات أي نوع من المشكلة يجب أن يحلها. وبتعبير آخر يجب على الجذر أن يكون ذو معنى حتى وإن ذكر بمفرده. وعادة ما يكون الجذر أطول والاختيارات أقصر (كما هو موضح في المثال السابق)، وعندما يقرأ المتعلم الجذر مرة واحدة، فإنه سيعرف الإجابة التي يبحث عنها، بعدها يقرأ الاختيارات لمرات عديدة.

ب- تجنب التكرار غير الضروري في البدائل

عند القيام بصياغة الجذر يجب أن يساهم في جعل البدائل لا تحتوي على تكرار عبارات غير ضرورية، ففي المثال التالي نلاحظ أن عبارة "الصدق" تكررت في كل البدائل في حين كان بالإمكان تجنب ذلك وجعلها متضمنة في الجذر. نوع الصدق الذي يناسب بدرجة أفضل اختبارات التحصيل هو:

أ- الصدق التلازمي. / ب- الصدق التنبؤي.

ج- الصدق الظاهري. / د- الصدق العملي

و يمكن تحسين هذا البند بجعل عبارة "الصدق" التي تتكرر في البدائل الأربعة موجودة في الجذر وحذفها من البدائل ليصبح البند بالشكل التالي:

نوع الصدق الذي يناسب بدرجة أفضل اختبارات التحصيل هو الصدق:

أ- التلازمي. / ب- التنبؤي.

ج- الظاهري. / د- العملي

ج - الجذر يجب أن يكون صحيحا نحويا في ذاته وفي علاقته مع البدائل

الجذر الجيد يجب أن لا يخرج عن قواعد النحو مع أي من الاختيارات المقترحة ففعل الكينونة للمفرد - يكون- وفعل الكينونة للجمع- يكونوا- ففي حالة ورودها في الجذر يجب أن تكون بنفس الصيغة في البدائل.

د - الجذر المناسب به فاعل واحد

عند صياغة العبارات فلا نضع المبادئ والنظريات والقوانين في الجذر واحد فيتشتت ذهن المتعلم ولا يفهم محتوى السؤال جيد، ففي المثال:

في النظام العشري يستخدم العدد 10 للعد ، بينما يستخدم النظام الاثنا عشري وحدة أساسية مختلفة وهي: (لن نكتب الإجابات الأربعة)، لأن ما يهمنا هو الجذر والذي احتوى على أكثر من فاعل، وبالتالي يمكن تحسينه.

في النظام العشري للعد، العدد 10 هو الوحدة الأساسية للعد، وفي النظام الاثنا عشري، الوحدة الأساسية هي: (هناك أربع احتمالات).

هـ الصيغة الإيجابية للجذر : كقاعدة، يستحسن أن تكون البنود في صيغة إيجابية، وكلما كان ذلك ممكنا يجب كتابة الجذر في عبارة بحيث تسأل عن الأصح وليس مجرد إجابة صحيحة مفردة وبوضع خط تحت الكلمة التي تدل على الإجابة "الأصح"، نثير انتباه المتعلم مثلا ويجب التأكد من أنه توجد إجابة واحدة فقط هي "الأصح". (البدائل الأخرى قد تكون صحيحة ولكنها ليست الأفضل).

والصياغة التالية تبين توفر هذه الخاصية:

- ما هو التفسير الرئيسي لاختلاف قيمة الارتباط باختلاف القانون المطبق.

- التعريف الأفضل للاختبار النفسي هو.

ففي المثال الأول يمكن أن يوجد أكثر من تفسير لاختلاف قيمة معامل الارتباط باختلاف القانون المطبق حسب طبيعة المعطيات والهدف من الدراسة، ولكن المطلوب هو اختيار التفسير الرئيس من ضمن التفسير المقدمة، والشئ نفسه ينطبق على الجذر الثاني حيث يتطلب تحديد التعريف الأفضل للاختبار النفسي من ضمن مجموعة تعاريف قد تكون كلها صحيحة في جانب من جوانبها ولكنها ليست الأشمل والأكثر دقة. و- **الصيغة السلبية للجذر** : عندما نستخدم المعنى السلبى في الجذر، يتطلب أن نلفت نظر المتعلمين إلى السلبية بالحروف الكبيرة أو بوضع خط أسفلها أو بكليهما معاً، ومن أمثلة الكلمات السلبية، أقل، لا، فيما عدا، ليس غير صحيح وغيرها، مع العلم أن العبارات السلبية أصعب في الفهم من الايجابية. فعندما نقول أي مما يلي لا يُوصف بصدق الاختبار، أو جميع ما يلي هي صفات مميزة للنظرية المعرفية ما عدا، فهذه الصياغات السلبية تجعل البند غامضاً بالنسبة للمتعلم ويستحسن كتابة بنود تكون جذورها إيجابية الصياغة.

ز- **طول الجذر والبدائل** : يستحسن تجنب الجذر القصير عندما تكون الاستجابات طويلة، والخلل في ذلك أن المتعلم يحتاج إلى وقت أطول ليقرأ أربع أو خمس استجابات طويلة بدلاً من أن يقرأ جذراً واحداً طويلاً وبدائل أقصر وأكثر تركيزاً، ومن الأمثلة على ذلك:

أي العبارات التالية تعد صحيحة؟

- كلما كان عدد بنود الاختبار كبير كلما زاد احتمال أن ترتفع درجة ثباته.

- المعطيات النفسية في المستوى الاسمي أكثر تعقيداً منها في المستوى الترتيبي.

- معامل الارتباط التتابعي يستعمل لتقدير قيمة الارتباط للمعطيات الترتيبية.

هذا المثال يوضح أن البند يجب أن يعدل من حيث الصياغة بحيث يصبح الجذر أطول من البدائل ويشير بوضوح إلى فكرة معينة تتطلب تحديد البديل المناسب لها.

ح- **تجنب الإشارة الضمنية أو التلميح للإجابة الصحيحة**

يتطلب عند صياغة البنود تجنب التلميح الضمني للإجابة الصحيحة فعندما نقول:

- تصنف الاختبارات حسب شروط الإجراء إلى اختبارات فردية واختبارات:

- | | |
|-----------|----------|
| - الشخصية | - الذكاء |
| - جماعية | - مقالية |

ففي هذا المثال هناك إشارة واضحة من الناحية اللغوية إلى أن ما يقابل الاختبارات الفردية هي الاختبارات الجماعية، مما يجعل المتعلم يصل إلى الإجابة الصحيحة من دون أن يكون فعلا متمكن منها، وعليه يجب تجنب هذا النوع من الصياغات التي لا تسمح بالتحقق هل فعلا وصل المتعلم إلى الإجابة الصحيحة بفعل تمكنه أم أن التلميح المتضمن في السؤال هو الذي أوصله إلى الإجابة الصحيحة.

ط - الاستجابات يجب أن تكون متماثلة مع بعضها و منسقة نحويا

لا يفضل اقتراح استجابات يتم التعبير عنها بالكلمات التي تعني شيئا في الواقع وتستخدم فقط لزيادة عدد الاستجابات، فثلاث أو أربع استجابات كافية جدا لبناء مفردة اختيار من متعدد جيدة من دون البحث عن بدائل أكبر ولكن أقل فعالية. كما في المثال:

أي أنواع الصدق يجب استخدامها للتحقق من صدق اختبار تحصيلي؟

- الصدق العالمي - صدق المحتوى

- الصدق التلازمي - صدق الأمل

ففي هذه الحالة، واضح أن صدق الأمل، ليس نوعا من أنواع الصدق، وهذه الاستجابة أو البديل استخدم من أجل الوصول إلى أربع بدائل فقط ولا علاقة له بالجذر، في حين أن الاستجابات الخاطئة، المعروفة باسم البدائل أو المشتتات.

فيجب أن تبدو مقبولة ظاهريا بجعلها مألوفا كالإجابة الصحيحة، وتنتمي إلى نفس المفهوم ومنطقية بحيث يحتمل أن تكون هي الإجابة الصحيحة ولها تقريبا نفس طول البدائل الأخرى.

ك- عدم استعمال العشوائية عندما يكون هناك ترتيب طبيعي

يفضل عند صياغة بند يتضمن تسلسل منطقي معين أن نرتب البدائل وفق هذا المنطق، ففي حالة تواريخ معينة فمن الأفضل كتابتها وفق التسلسل الزمني وليس بطريقة عشوائية، فمثلا إذا قلنا ما هي السنة التي قدم فيها بنبيه اختباره للذكاء:

- 1902 / - 1903 / - 1904 / - 1905. وهو تسلسل زمني منطقي.

- تجنب وضع الإجابة الصحيحة في ترتيب معين، كأن تكون الثانية أو الرابعة أو موزعة وفق منطق معين يسمح للمتعلم باكتشافه، فتتضخم درجته بشكل غير حقيقي أي عن طريق التخمين، مثل توزيع الإجابات بالشكل (أ، أ، ب، ب، ج، ج، ج)، فهذه المنهجية تزيد من احتمال وصول المتعلم إلى الإجابة الصحيحة عن طريق التخمين. وعليه يمكن وضع الاختيارات بحيث تظهر الإجابة الصحيحة في مواضع متنوعة (مرات أو أخرى ب، وج، ود، وهكذا) لعدد متساوي تقريبا من المرات.

إن احترام المدرس أو القائم على بناء الاختبار، لهذه الشروط عند بنائه لاختبار من نوع الاختيار من متعدد، يجعل منه أداة سليمة يمكن أن تساهم فعلا في الكشف عن الجوانب التي تمكن منها المتعلم والتي لم يتمكن منها، فعندما يجعل لكل بند هدف تعليمي واحد يقيسه، ويتجنب تكرار تناول المضمون الواحد في أكثر من مرة في كلمات واضحة وسليمة لغويا، ومراجعة البنود قبل تطبيق الاختبار تطبيقا نهائيا كل هذه الإجراءات من شأنها أن تعطي قيمة علمية للبنود، وتجعل منها وسيلة موضوعية للقياس تمكن من الحصول على معلومات يمكن توظيفها في اتخاذ قرارات تربوية وإدارية مختلفة.

فاختبارات الاختيار من متعدد تصلح لقياس مهارات التعرف على المفاهيم والمصطلحات العلمية التي سبق للمتعلم وأن تناولها، ولكنها غير قاصرة على قياس مستوى معين من المعارف أو الأهداف، بل تشمل قياس جميع مستويات المجال المعرفي، كالفهم والتطبيق والتحليل والتركيب، وعليه تعتبر أفضل أشكال البنود الموضوعية إلا أنها صعبة الإعداد حيث تتطلب وقتا طويلا وجهدا مضاعفا ومهارة قد لا تتوفر عند الكثير من المعلمين.

4- قياس التحصيل الدراسي وفق النظرية الكلاسيكية

تقوم النظرية الكلاسيكية على مبدأ أساسي يقوم على تقدير الفروق بين الأفراد والتمييز بينهم في الأداء بصفة عامة، حيث تكتسب درجة الفرد على الاختبار معناها عندما تسند إلى معيار مستوى الجماعة التي ينتمي إليها الفرد، ويحدد مستوى الجماعة بمتوسط درجاتها على الاختبار، ويتغير هذا المعيار بتغير مستوى الجماعة، حيث تعتمد فلسفة القياس على التمييز بين الأفراد بهدف الانتقاء والاختيار، وتعد الاختبارات الجيدة هي الأقدر على التمييز بين الأفراد.

فالاختبارات في ظل النظرية الكلاسيكية تعتمد على الفروق الفردية، حيث تتوزع درجات الأفراد في إطارها توزيعاً اعتدالياً، يخدم الأغراض التالية:

- عندما يكون الموضوع الدراسي ليس بشكل متراكم أو سلسلة مترابطة يعتمد بعضها على البعض الآخر، وفي الحالات التي لا يطلب من المتعلمين الوصول إلى مستوى محدد من الكفاءة، وعندما تتطلب وضعية بيداغوجية معينة اختبار عدد من الأفراد المتميزين أو الأوائل للالتحاق بتخصص معين عدد الأماكن فيه محدود.

- عندما يكون الهدف هو التنبؤ بدرجة نجاح المتعلم، لأنها تعتمد على الفروق الفردية في الخاصية التي يفترض أنها ترتبط بأداء المتعلمين في المستقبل.

والنظرية الكلاسيكية للاختبارات النفسية والتربوية ارتكزت على مجموعة من القواعد شكلت مبادئ أساسية تتناول المعطيات الكمية المتحصل عليها في ضوءها.

هذه القواعد يمكن إجمالها فيما يلي:

- الخطأ المعياري للقياس ينطبق على جميع الدرجات عندما نطبق الاختبار على عينة محدودة، أي أن النظرية الكلاسيكية تفترض تساوي الخطأ المعياري للقياس عند جميع الأفراد بغض النظر عن مستوى قدراتهم.

- الاختبارات الطويلة تكون أكثر ثباتاً من الاختبارات القصيرة، عدد بنود الاختبار له علاقة مباشرة بدرجة ثباته، فيلجأ الباحث عندما يريد أن يرفع من قيمة معامل ثبات اختباره إلى زيادة عدد بنوده مما يزيد في درجة صدق محتواه ومنه الزيادة في قيمة ثباته.

- تكون مقارنة درجات الاختبار أفضل عندما تتوفر غدة صور متكافئة للاختبار نفسه، ويمكن أن تطبق عدة مرات في ظروف مشابهة على نفس الأفراد، ليتم مقارنة نتائجهم في التطبيقات المتتالية للحكم على مستواهم الفعلي.

- التقديرات غير المتحيزة لخصائص البنود تعتمد على مدى تمثيلها لمجتمع البنود الممكن أن يحتوي عليها الاختبار، فالبنود عبارة عن عينة ممثلة لمجموع البنود التي يمكن أن يحتوي عليها الاختبار.

- تكتسب درجات الاختبار مدلولها من خلال مقارنة وضعها بمعيار المجموعة الواحدة، أي أن الدرجات تفسر بناء على متوسط العينة التي يطبق عليها الاختبار فمعيار تفسير النتائج يستمد من النتائج الفعلية للأفراد على الاختبار.

- خصائص مقاييس المسافة تتحقق بالحصول على توزيع اعتدالي للدرجات، حيث هدف الاختبار هو الكشف عن الفروق الفردية في السمة المقاسة، والهدف هو رفع مستوى القياس إلى مستوى المسافات المتساوية حيث يسمح بالحصول على درجات من تطبيق الاختبار على أي مجموعة تتوزع وفق المنحنى الطبيعي.

- الصيغ المختلطة للبنود تؤدي إلى أثر غير متوازن على الدرجة الكلية للاختبار.

- التحليل العاملي للبنود الثنائية يؤدي إلى عوامل مصطنعة وليست عوامل حقيقية. بمعنى أن تحليل الدرجات التي يحصل عليها الأفراد على كل بند تكشف عن عوامل تؤثر في أداء الأفراد لا تعود إلى تباين مستوى القدرة الحقيقي.

وعلى الرغم من الفوائد التي يمكن أن تقدمها الاختبارات المبنية وفق أسس النظرية الكلاسيكية، لما تتميز به من خصائص الصدق، الثبات وقدرة على التمييز إلا أن عيوبها أو مشكلاتها تحد من فائدتها التربوية وتجعل الاعتماد على نتائجها منقوص الفائدة نظراً لاعتبارات هي:

- تقتصر هذه الاختبارات على التمييز بين الأفراد ولا يتجاوز ذلك إلى معرفة ما تحقق من أهداف العملية التعليمية، مما يؤدي إلى حذف عناصر ووحدات من الاختبار إذا كانت غير مميزة بين الأفراد بغض النظر عن قيمتها العلمية مما يؤثر في بناء الاختبار وصدقه.

- الاقتصار على اتخاذ الجماعة كمحك لتفسير درجة الفرد على الاختبار يؤدي إلى عدم التيقن من موقع الفرد وجماعته بالنسبة لأهداف التعلم ومدى تحقيقها.

- طغيان المنحنى الاعتدالي المعياري باعتباره النموذج الأمثل لتوزيع الدرجات وهو الذي يتحقق عندما يكون مستوى البنود يتلاءم مع مستوى الأفراد المطبق عليهم الاختبار، وعندما تكون البنود صعبة أو سهلة نحصل على منحنى ملتو التواءً سالباً أو موجباً، مما يدعو إلى التركيز على مستويات المتعلمين بدلاً من الأهداف التعليمية.

ومن الاختبارات التي تستخدم معيار الجماعة كنظام مرجعي لتفسير نتائج الاختبارات التحصيلية المقننة، والتي تعتمد على موازنة أداء المتعلم بأداء أقرانه وفقاً للمجموع الكلي للدرجات، والتي تخضع لخصائص عينة التقنين، كما تخضع لخصائص عينة الأسئلة أو البنود التي يشتمل عليها الاختبار، مما يؤثر في صدق الموازنة وبالتالي صعوبة تعميم النتائج المتوصل إليها، ففي ظل النظرية الكلاسيكية يختلف تدريج صعوبة البنود باختلاف قدرات الأفراد. كما أن قياس قدرات الأفراد يتباين بتباين صعوبة البنود.

5- مشكلات ناتجة عن النظرية الكلاسيكية في قياس التحصيل الدراسي

تواجه أساليب القياس في مجال الظواهر السلوكية بصفة عامة وفي مجال التحصيل الدراسي بصفة خاصة بعض المشاكل تؤثر في موضوعيتها ومصداقية نتائجها والتفسيرات التي تتبع هذه النتائج، وهناك بعض المشاكل التي تواجه الأساليب التقليدية في القياس النفسي والتربوي، وفي الوقت نفسه تعد بمثابة نقائص في هذه الأساليب تجعل من تطبيقها في مختلف الممارسات والاستفادة العلمية والعملية من نتائجها محدوداً، كما أن العامل في الميدان التربوي يجب أن يكون مطلعاً على هذه المشكلات حتى يأخذها بعين الاعتبار ولا يفسر نتائج اختباراته بعيداً عنها، ولكن في إطار نسبي مرتبط بها، ويمكن تلخيصها كما يلي:

5-1- تقيد الدرجة الكلية للاختبار ببنوده

تتلخص هذه المشكلة في أن الدرجة التي يتحصل عليها الفرد في اختبار معين تعتمد على بنود هذا الاختبار، فإذا حصل فرد على الدرجة (15) في اختبار تحصيلي معين فإننا لا نستطيع أن نجزم بأنه سوف يحصل على نفس الدرجة إذا ما استبدل هذا الاختبار باختبار آخر يقيس نفس المتغير (التحصيل)، أي أن درجة الفرد في الاختبار تتغير بتغير بنوده.

هذه المشكلة لا تطرح في مجال القياس الموضوعي حيث أن الدلالة الكمية لطول أي قاعة مثلا لا تتغير بتغير أداة القياس المستخدمة لقياس الطول، وهذا يقلل من درجة موضوعية النتائج المتحصل عليها. وتصبح عملية المقارنة بين أداء مجموعة من الأفراد خضعوا لاختبارات مختلفة لقياس الظاهرة نفسها في أوقات مختلفة.

5-2- انعدام خطية القياس

يقصد بخطية القياس وجود معدل ثابت لتدرج القياس على متصل المتغير أو السمة موضوع القياس، أي أن الدرجات التي يحصل عليها الفرد اختبار معين يمكن جمعها كما لو كانت تمثل مقياسا خطيا، والمعدل الثابت يتمثل بوحدة قياس ثابتة، وهو ما تتسم به المقاييس في المجال الفيزيائي، حيث أن تقدير الفروق بين أي قياسين متتالين على نفس التدرج يكون ثابتا، ولا يتغير هذا الفرق بتغير الأداة المستخدمة طالما كانت أداة مناسبة ولها وحدة قياس ثابتة. ومثال ذلك الفرق بين طولين لقاعتين دراسيتين 5 متر و 8 متر هو نفس الفرق بين الطولين 10 و 13 أي ثلاثة أمتار.

أما في مجال القياس السيكومتري، فإن درجات الاختبار لا تعطي قياسات خطية، فعلى الرغم من قدرة الدرجات على تصنيف مستويات الأفراد إلا أنها غير قادرة على تقدير المسافات بين هذه المستويات بصورة ثابتة، وذلك لأن المسافات بين درجات الاختبار غير متساوية واقعا حتى لو كانت تبدو ظاهريا متساوية.

5-3- القياس في أكثر من بعد

تتميز أداة القياس الفيزيقي بالدقة أكثر من المقاييس النفسية والتربوية، كما أن المقاييس الفيزيقيّة تقيس شيئا واحدا وقت واحد كالطول مثلا، إلا أن اختبارا للذكاء قد يقيس متغيرات أخرى مع قياسه للذكاء كالثروة اللغوية، أو الخصائص الاجتماعية وغيرها، وبالتالي فإن الدرجة التي يحصل عليها الفرد على هذا الاختبار لا تشير إلى درجة ذكاءه فقط، بل هناك متغيرات أخرى تختلف عن حقيقة ما يقيسه الاختبار، وبالتالي يصعب تقرير أحادية البعد في هذا النوع من الاختبارات، أي يصعب تفسير النتيجة المتوصل إليها باعتبارها انعكاسا لقدرة واحدة هي موضوع القياس، والتي يقيسها الاختبار .

5-4- تزايد درجات الاختبار مع مستوى المتغير المقاس

يشير القياس الموضوعي إلى أنه كلما زادت درجة الفرد في الاختبار دل ذلك على زيادة قدرته في الظاهرة المقاسة، ولكن قد يحصل بعض الأفراد مرتفعي القدرة على درجات منخفضة كما يحدث العكس أحيانا لذوي القدرات المنخفضة حين يتحصلون على درجات مرتفعة، فتفترض الاتجاهات التقليدية للقياس أن درجات الاختبار دالة خطية متزايدة للسمة المقاسة، أي كلما زادت درجة التحصيل زادت درجة القدرة، ولكن هذا غير صحيح في جميع الحالات نظرا لتدخل عوامل كثيرة مؤثرة في عملية القياس المبني على

أسس تقليدية لا تفسر كيفية حدوث التفاعل بين الأفراد والبنود وصولاً إلى استجابات قابلة للقياس وتعبّر عن السمة الكامنة.

5-5- تغيير بنية الاختبار و معاني بنوده مع الزمن

تعاني أساليب القياس التقليدية من تغيير معاني البنود المستخدمة مع مرور الزمن ولذلك فإن التكوين العامي للاختبار غير ثابت، ويرجع ذلك إلى تغيير الظروف البيئية والثقافية وتغير قدرات وخصائص الأفراد، بالظواهر النفسية من ميزاتها الأساسية أنها غير ثابتة ومتغيرة في فترات زمنية متقاربة، هذه العوامل تؤدي إلى تغيير درجات الفرد بصورة يصعب التنبؤ بها مستقبلاً في ظل الأساليب التقليدية لقياس الأداء بصفة عامة والتحصيل بصفة خاصة.

5-6- اعتماد خصائص البنود على خصائص الأفراد

الخصائص السيكومترية للاختبارات الكلاسيكية تعتمد على خصائص عينة الأفراد التي يجري عليها الاختبار، وعلى خصائص عينة البنود التي يشتمل عليها الاختبار، مما يؤدي إلى نظام قياس نفسي وتربوي غير مستقر، بمعنى تغيير معاملات الصدق والثبات وصعوبة البنود بتغيير قدرات الأفراد، مما يجعل تطبيق الاختبار على عينات واسعة تملك خصائص مختلفة غير ممكن عملياً.

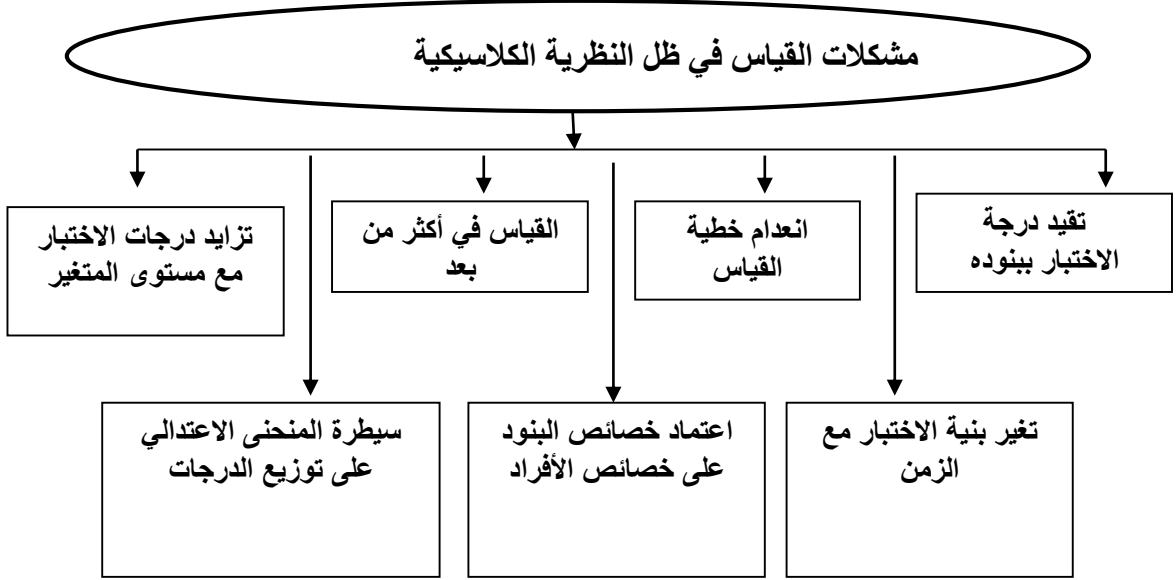
5-7- ارتباط توزيع الدرجات بالمنحنى الاعتدالي

لما كانت بنود الاختبارات المبنية على أسس النظرية الكلاسيكية تعتمد بالدرجة الأولى على معامل التمييز، أي قدرتها على التمييز بين المختبرين بناء على المستويات المختلفة للقدرة أو السمة المقاسة، فالاختبار الجيد هو الذي يعطينا درجات تتوزع طبيعياً، حيث يتم حذف البنود التي لا تميز مهما كانت درجة صدقها وتمثيلها للنطاق السلوكي أو للأهداف المراد قياسها، مع الإبقاء فقط على البنود غير المتطرفة في معاملات الصعوبة والسهولة قصد الإبقاء على خاصية التمييز، وهذه الفرضية الأساسية للنظرية السيكومترية الكلاسيكية، حيث توزيع الدرجات التي يتحصل عليها الأفراد على بعد القدرة التي يقيسها الاختبار يتخذ شكل المنحنى الاعتدالي، فيمكن تقدير النسب المئوية للأفراد في كل مستوى من مستويات هذه القدرة باستخدام دالة الكثافة الاحتمالية، لذلك يهتم باختيار البنود التي تجعل توزيع درجات الاختبار تقترب من التوزيع الاعتدالي، وهذا يعتمد على سهولة وصعوبة البنود بالنسبة لعينة الأفراد وعلى مدى تمييز كل بند منها بين المستويات المختلفة للقدرة.

هذا الافتراض يجعل من عملية التعليم والتعلم خاضعة للصدفة، حيث من لديه قدرات عقلية مرتفعة يتحصل على درجات مرتفعة ومن لديه قدرات منخفضة يتحصل على درجات منخفضة، وفق علاقة طردية بين القدرات والتحصيل أو الأداء، ولكن مع ظهور التعلم من أجل التمكن وبيداغوجية التحكم، أصبح من الممكن أن يصل أغلبية المتعلمين إلى تحقيق الأهداف المرجوة من التعلم، إذا منح لهم الوقت

الكافي وكيّفت طرق التدريس وإجراءاته حسب قدراتهم وخصائصهم، بمعنى عندما تراعي عملية التعليم والتعلم الفروق الفردية فإن نتائج التحصيل لن تتوزع اعتدالياً، وهذا ما يجعل النظرية الكلاسيكية قاصرة عن تفسير أداء الأفراد في هذه الحالة.

ويمكن إجمال أهم النفاص التي تعاني منها النظرية الكلاسيكية للقياس في الشكل التالي:



أمام هذه المشكلات التي تطرحها الاختبارات مرجعية الجماعة، أو ما اصطلح على تسميته بالنظرية الكلاسيكية في القياس النفسي، لم يقف علماء القياس بدون اقتراح الحلول، بل قاموا بجهود متواصلة للتغلب على هذه المشكلات والتي تؤثر في الخصائص السيكومترية للاختبارات التحصيلية، مما يؤثر في دقة وصحة النتائج، فقد تم التوصل إلى نماذج جديدة تجعل عملية قياس السمات أكثر موضوعية، بمعنى عدم تأثر درجة الفرد بخصائص عينة الأفراد الذين يجيبون على الاختبار ولا بخصائص عينة البنود المكونة لأداة القياس.