

جامعة البليدة 2 علي لونيبي

كلية العلوم الاجتماعية والإنسانية

قسم العلوم الاجتماعية

شعبة علم النفس

السنة الأولى ماستر * تخصص علم النفس التربوي

مقياس القياس والتقويم التربوي – السداسي الأول

أ.د. بوسالم عبد العزيز

الدرس الأول : مفهوم القياس في علم النفس والتربية

يقول ثورنديك (Thorndike) إذا وجد شيء فإنه يوجد بمقدار وكل مقدار يمكن قياسه، هذه المقولة تشير إلى قابلية السمات النفسية للقياس قصد تحديد الفروق الكمية بينها أو تحديد كم هذه السمات ليكون الحكم عليها وتفسيرها أكثر موضوعية، بهدف فهم السلوك الإنساني والتنبؤ به مستقبلا وهو غاية العلم بصفة عامة.

والقياس لغة من الفعل قاس يقيس أي قدر الشيء، فنقول قاس الشيء بغيره أو على غيره أي قدره على مثله أو بمثله، ويستعمل القياس بصفة عامة وفق معاني منها.

التقدير: فيقال (قست الشيء بالشيء) أي قدرته به، وقايست بين الأمرين، أي قدرت بينهما.

التسوية: ويطلق على النوعين الحسي كأن يقال قست النعل بالنعل أي ساويتها بها، أما البعد المعنوي كقولنا فلان لا يقاس بفلان أي لا يساوى به.

أما اصطلاحا فيشير مفهوم القياس حسب كامبل (Campell) إلى العملية التي يتم من خلالها تمثيل الخصائص أو السمات النفسية بأرقام أو أعداد وفق قواعد معروفة ومحددة مسبقا، وعرفه كيرلنج ((Kerling على أنه نسبة الأرقام إلى السمات أو الخصائص وفق قواعد معينة، هذه الأرقام تعد بمثابة رموز لا تحمل معنى كليا في حد ذاتها إلا إذا حدد الباحث بنفسه هذا المعنى الكمي وفق الإجراءات التي اتبعها للحصول على الرقم، أو قواعد تحويل السمات من بعدها الكيفي إلى بعد كمي.

كما يعرف جيلفورد (Guilford) القياس النفسي على أنه وصف للبيانات أو السمات النفسية بواسطة الأرقام وفق قواعد متفق عليها.

ويشير القياس في علم النفس إلى "عملية تحديد قيم رقمية لأشياء أو موضوعات تبعا لقواعد معينة متفق عليها بين العاملين في ميدان القياس" كما يشير القياس في ميدان علم النفس إلى "عملية مقارنة سمة نفسية بوحدات معيارية معينة. أو بكلمة قياسية أو مقدار مقنن من السمة نفسها، بهدف معرفة كم من الوحدات يتضمنها هذا الشيء"

أما ستفنز ((Stevens فيشير أن القياس هو العملية التي يتم بواسطتها التعبير عن السمات والخصائص برموز وأعداد استنادا إلى قواعد معينة يحددها الباحث.

من خلال هذه التعاريف يمكننا القول أن القياس في علم النفس على غرار نظيره في العلوم الأخرى، يعتبر تقديرا كميًا للظواهر باستعمال قواعد محدّدة تسمح بإعطاء معنى معيّن للنتيجة الكميّة المتحصل عليها، فالأرقام التي نحصل عليها من خلال عملية القياس والتي ترتبط بقواعد معينة استخدمت من الحصول عليها، تبقى بدون معنى علمي إلا من خلال المعنى الذي يعطيه إياها الباحث.

والقياس بصفة عامة هو الاستعاضة بالأرقام عن السمات، أي بدلا من الوصف الكيفي للسمة يتم الوصف الكمي لها، وفق قواعد أو معايير والتي تعتبر القاسم المشترك بين مختلف التعاريف المقدمة للقياس، وهي التي تميز القياس في العلوم السلوكية عنه في ميدان العلوم الطبيعية، التي تمتاز بوضوح هذه القواعد وإجماع المختصين حولها، في حين يصعب تحديدها والاتفاق حولها في العلوم السلوكية نظرا لطبيعة السمات التي تدرسها والتميزة أساسا بالتداخل فيما بينها، النسبية في قياسها وافتراضية المؤشرات الدالة عليها، مما يجعل عملية قياسها بموضوعية أمرا معقدا. ولفهم عملية القياس في سياقها الواقعي، نتناول مكونات هذه العملية والتي تعتبر المدخل لتحقيق قياس موضوعي سليم.

-عناصر عملية القياس

تشير النظرة التحليلية لعملية القياس إلى وجود ثلاثة عناصر أساسية متضمنة في هذه العملية. وهي الخاصية أو السمة المراد قياسها، أي موضوع القياس، وأداة القياس التي تستخدم لإصدار حكم كمي عن السمة، وأخيرا العدد أو الرقم الذي يشير إلى درجة وجود السمة المدروسة بناء على قواعد مختلفة.

موضوع القياس

عملية القياس لا تتم في فراغ، بل تتناول سمة محددة يقرر الباحث قياسها، فعندما نريد أن نقوم بعملية قياس فإننا لا نتناول فردا ونقيسه من جميع جوانب شخصيته، بل نحدد سمة أو خاصية من خصائصه ونخضعها لعملية القياس من منطلق قابليتها لذلك، فالقياس لا يوجه إلى أشياء أو

أشخاص بل لخصائص معينة يتم تحديدها بصفة دقيقة، بحيث يمكن في النهاية تحويل الوصف الكيفي إلى وصف كمي يسمح بإجراء مختلف المعالجات الإحصائية لفهم السمة وتفسير السلوك والتنبؤ به مستقبلاً.

ففي الميدان الفيزيائي مثلاً، لا نقيس القاعة بل نقيس طولها أو عرضها أو ارتفاعها، وكذلك في العلوم السلوكية لا نقيس الفرد، بل نقيس ذكائه تحصيله، ميوله أو احتراقه النفسي، فبعد أن يحدد الباحث السمة، يتم ضبطها بدقة وإجرائية مع محاولة عزلها عن بقية السمات الأخرى، خاصة مع التداخل الموجود بين السمات النفسية، فيصبح التحديد الواضح لمؤشرات السمة الخطوة الأساسية الأولى والمهمة والأكثر تعقيداً في عملية القياس والتي توجه الخطوة الموالية وهي تحديد طبيعة الأداة المناسبة لقياسها.

-الأداة

للقيام بعملية القياس لا بد من توفر أداة قياس معينة تتناسب وطبيعة السمة أو الخاصية المراد قياسها، فالخطوة الثانية في عملية القياس بعد التحديد الدقيق والإجرائي للسمة المراد قياسها هي اختيار أو بناء الأداة المناسبة. فالاختبار في علم النفس يعتبر بمثابة مجموعة من المثيرات مرتبة أعدت لتقيس بطريقة كمية في أغلب الحالات، وكيفية في بعضها لتقيس عينة من السلوك أو الخصائص، والمثيرات قد تكون أسئلة شفوية، كتابية، سلسلة من الأعداد أو الأشكال الهندسية وغيرها من المثيرات التي من شأنها استثارة استجابة معينة من طرف الأفراد تخضع لعملية التكميم.

والاختبار النفسي يقيس سمة محددة ليعطي درجة أو قيمة كمية بمقدار السمة، والباحث إما أنه يجد الاختبار جاهزاً فيقرر مدى صلاحيته لقياس السمة عند أفراد يتميزون بخصائص معينة، فيطبق الاختبار عليهم، وقد لا يجد الباحث اختباراً جاهزاً يناسب طبيعة السمة التي يريد قياسها أو لا تتناسب وخصائص العينة، كما يمكن أن تكون الاختبارات المتوفرة لا يحقق أهداف البحث. في هذه الحالة على الباحث أن يقوم ببناء اختبار وفق منهجية وخطوات معينة ومتعارف عليها بين المختصين في بناء الاختبارات، هذا الاختبار يحرص الباحث على أن تكون بنوده عينة ممثلة فعلاً للمجال السلوكي المراد قياسه بطريقة غير مباشرة، بمعنى على درجة مقبولة من الصدق .

فالاختبارات في علم النفس والتربية أنماط مختلفة، فهناك اختبارات أدائية وأخرى تحريرية والبعض منها شفوي والآخر كتابي، وتطبيق هذه الاختبارات يعطينا وصفا كميًا في شكل أرقام تعتبر المكون أو العنصر الثاني في أي عملية القياس.

العدد

يعتبر العدد أو الرقم بمثابة اللغة المشتركة بين العلوم باختلاف طبيعتها ودقتها، وما التطورات التي عرفتتها مختلف العلوم إلا بعد أن أصبحت تتخذ من الرقم وسيلة للتعبير عن الظواهر التي تدرسها. والقياس النفسي في أصله يقوم على التعامل مع السمات موضوع القياس بلغة الأرقام، فالرقم هو الرمز أو التعبير الكمي الدال على هذه السمات، وهو يوفر قدرا من الدقة لا يوفره الحكم الكيفي. حيث يمكن من خلال خصائص العدد مقارنة قدرات الأفراد بعضها ببعض ومعرفة درجة توفر خاصية ما لدى فرد أو عدد من الأفراد، مما يسمح لنا باتخاذ قرارات مختلفة لا يمكن الوصول إليها بغير هذا الأسلوب، كما يصبح من الصعب تقديم تفسير موضوعي لسمة معينة وإعطائها دلالتها دون توفر هذا النوع من المعطيات.

ويمكن القول أن العدد كنتيجة لعملية القياس ينطوي على معاني عدة، فقد يدل الرقم (04) مثلا على عدد أفراد أسرة معينة أو عدد المواد التي نجح فيها طالب معين، أو طول قاعة دراسة بالأمتار. كما قد يدل على رقم من أرقام أحد الرياضيين أو رقم ولاية من الولايات، أو رتبة طالب بين زملائه في تحصيل مادة معينة. أو درجة ذكاءه. هذه الاستعمالات تبين اختلاف معاني الرقم (04) في كل حالة من الحالات السابقة. فقد يستعمل مرة بوصفه رقما أو عددا ومرة بوصفه رتبة ودرجة أو كمية شيء معين، فيمكن أن يقتصر استخدام العدد للتعبير عن الكميات كما يمكن استخدامه كرمز دال على الخصائص أو الأشياء وبدل عنها، دون أن يكون له دلالة كمية، كما يمكن أن يعبر عن ترتيب عدد من الأفراد وفق خصائص معينة بتسلسل الزيادة أو النقصان حسب امتلاك كل فرد لتلك الخصائص وغيرها من الاستعمالات التي يمكن أن نوظف الرقم فيها.

يتضح مما سبق أن لتمييز معاني الأعداد التي نحصل عليها في عملية القياس وتحديد قواعد استخدامها أهمية بالغة في ميدان الدراسة الموضوعية للسلوك الإنساني، فلا بد للباحث أن يراعي القواعد التي يقابل بها بين الأعداد والسمات. ويحاول إدراك المعنى الحقيقي للعدد بناء

على المنطق المستخدم في نسبته لسمة معينة نظرا لاختلاف معنى العدد، فعلامة الطالب في امتحان معين ودرجة نكائه ورتبته بين زملائه ورقم شهادته الدراسية، يعبر عنها بأعداد تختلف في معناها باختلاف القواعد المستخدمة فيها، على اعتبار أن العدد هو نتيجة لسلسلة من العمليات والإجراءات التي اتبعت للحصول عليه، مما يتطلب أخذها بعين الاعتبار عند التعامل مع أي قيمة عددية نحصل عليها من خلال عملية القياس، وبمعنى آخر يتطلب معرفة مستوى القياس الذي يمكن أن ينسب إليه العدد.

الدرس الثاني : -قياس التحصيل الدراسي من منظور تاريخي

إن الفهم الجيد للطرق و الأساليب التي تستخدم في وقتنا الحاضر في قياس التحصيل الدراسي فهما موضوعيا ينبغي أن يستند إلى رؤية واضحة للتطورات التاريخية التي مرت بها هذه الأساليب حتى أصبحت كما نراها اليوم.

فالامتحانات المدرسية كانت تعتمد اعتمادا أساسيا على الاختبارات الشفوية حتى جاء هوراس (Horace) سنة 1845 القادة التربويين البارزين في حركة تطوير التعليم في الولايات المتحدة الأمريكية، وأكد على ضرورة استخدام الامتحانات التحريرية التي تتسم بالموضوعية والاتساق بدلا من الاختبارات الشفوية في تقويم التحصيل الدراسي للمتعلمين، عبد الهادي نبيل (2002).

ففي عام (1864) أعد المربي الانجليزي جورج فيشر (G.Fisher) أول اختبار تحصيلي كتابي يتكون من عدة مقاييس متدرجة في الصعوبة وأمثلة ومواصفات متنوعة يمكن باستخدامها في تقييم تحصيل المتعلم في مواد مختلفة، وتقديم صورة تربوية تبين جوانب قوته وضعفه. وبذلك وضع هذا الاختبار الأساس لبناء الاختبارات التحريرية الموضوعية، سليمان أحمد عبيدات،(1988) ويعتبر هذا الاختبار الركيزة الأساسية التي أسندت إليها حركة بناء الاختبارات مرجعية المحك في وقتنا الحاضر.

التطور الموالي في قياس التحصيل يتمثل في إعداد ريس (J.Rice) لأول اختبار تحصيلي موضوعي في الولايات المتحدة الأمريكية، يقيس قدرة تلاميذ المدارس الابتدائية على الهجاء،

وطبقه على ما يقارب (16000) تلميذ في المرحلة الابتدائية، كما قام بوضع إختبارات تحصيلية أخرى في الحساب واللغة الانجليزية والتاريخ .

هذه الاختبارات تعد القاعدة للتطور الحاصل في بناء الاختبارات التحصيلية المقننة، حيث شهد أوائل القرن العشرين(1908) ظهور عدد من الاختبارات المقننة متمثلة في اختبارات الحساب و اللغة الانجليزية التي أعدها ستون(Stone) ، واختبار جودة الخط الذي أعده ثورنديك (Thorndike) عام 1909، وتزايد عدد هذه الاختبارات زيادة كبيرة بقرب نهاية عام (1920)، بما في ذلك بطاريات الاختبارات مثل بطارية إختبارات ستانفورد التحصيلية للمرحلة الابتدائية (1923)، وبطارية إمتحانات إيوا للمحتوى الدراسي للمرحلة الثانوية عام (1925).

وعدد 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 أزداد الاهتمام ببناء بنود موضوعية تقيس الفهم والاستنباط والتحليل، وغيرها من العمليات المعرفية العليا، وكذلك الأهداف التعليمية في مجالات مختلفة، وشهدت الثلاثينات من القرن الماضي حسب سبع محمد أبولبدة (1975) تصميم 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 التحصيلية المقننة على الأغراض التربوية وإنما استخدمت أيضا في إنتقاء العاملين بالمصانع والمؤسسات الإنتاجية المختلفة والقوات المسلحة، بهدف إختيار الأفراد المناسبين لأداء مهام معينة بناء على نتائج تطبيق إختبارات تقيس مختلف جوانب الشخصية.

ويشير(صلاح الدين محمود علام، 2003، ص25) إلى 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.

- مفهوم التحصيل الدراسي:

يعد التحصيل الدراسي من المفاهيم المفتاحية 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(جابر عبد الحميد جابر وطاهر محمد عبد الرزاق،1990، ص35) 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 بالمعدّل التراكمي.

وعرّف (محمد السيد علي، 2003، ص227) التحصيل الدراسي بأنه: "المعلومات والمهارات المكتسبة من قبل المتعلمين كنتيجة لدراسة موضوع، أو وحدة دراسية محددة "

والتحصيل الدراسي هو الحصول على معلومات وصفية تبين مدى ما حصّله الطلبة بطريقة مباشرة من محتوى المادة الدراسية، وذلك من خلال الاختبارات التي يطبقها المعلم على طلبته على مدار العام الدراسي لقياس مدى استيعابهم للمعارف والمفاهيم والمهارات التي لها علاقة بالمادة الدراسية في وقت معين أو في نهاية مدة تعليمية معينة. (آدم أحمد، 2003، ص 173).

وعرّف (محمد خليفة بركات،1982، ص21) التحصيل الدراسي بأنه درجة الاكتساب التي يحققها متعلم، أو هو مستوى النجاح الذي يحرزه أو يصل إليه في مادة دراسية أو مجال

تعليمي أو تدريبي معين، فالاختبارات التي تطبق على المتعلمين خلال فترة زمنية معينة تقيس تحصيلهم الدراسي."

كما عرّفه (دوكاتل 1982 De kettell ، 32) على أنه عبارة عن "مستوى معين من الإنجاز أو الكفاءة في العمل الدراسي كما يتم تقييمه من قبل المعلمين بواسطة الاختبارات المدرسية أو الاختبارات المقننة أو الاثنين معا."

وقدم (أحمد عودة ، 2000 ، ص10) تعريفاً للتحصيل يشير إلى مكتسبات المتعلم في مختلف جوانب الشخصية فيقول بأنه "مجموعة من المعارف والمهارات والميول الملاحظة لدى الدارسين نتيجة عملية التعليم، وهو عامل يتأثر بعوامل أهمها خصائص المتعلم، المعلم، المنهج، الإدارة المدرسية، الأسرة والتقنيات التربوية واللوائح التنظيمية وغيرها "

أما (فؤاد أبو حطب و سيد أحمد عثمان، 1970، ص172) فالتحصيل الدراسي بالنسبة إليهما " يرتبط ارتباطاً وثيقاً بمفهوم التعلم المدرسي، إلا أن مفهوم التعلم المدرسي أكثر شمولاً، فهو يشير إلى التغيرات في الأداء تحت ظروف التدريب والممارسة في المدرسة، كما يتمثل في أساليب التكيف، ويشمل النواتج المرغوبة وغير المرغوبة، أما التحصيل الدراسي، فهو أكثر اتصالاً بالنواتج المرغوبة للتعلم أو الأهداف التعليمية".

كما قدم (بيرون، Piéron، 1969، ص33) تعريفاً للتحصيل اعتبره " النتيجة المتوصل إليها من خلال تطبيق الاختبارات التحصيلية المعمول بها بالمدارس في الامتحانات في نهاية العام الدراسي، وهو ما يعبر عنه المجموع العام لنقاط المتعلمين في مختلف المواد الدراسية 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 الانجاز الفعلي يصل إليه المتعلم يظهر في شكل كفاءات يمكن الكشف عنها أو قياسها بواسطة إختبارات مبنية من طرف المعلم أو إختبارات مقننة مبنية من طرف خبراء في بناء الاختبارات .

ويمكن إعتبار أي إختبار 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 إختبارات 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 نتائج في مختلف المواد ليوجه نحو تخصصات جامعية معينة، فالذي يتحصل على علامات جيدة في مادتي الرياضيات و الفيزياء مثلا يوجه لدراسة العلوم الدقيقة إنطلاقا من بعد تنبؤي وأنه يملك إستعدادا يسمح له بالنجاح في هذا التخصص.

الدرس الثالث : 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 الاجرائية 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، و هذا يتطلب التحديد المسبق لهذه الأهداف لمعرفة ماذا أستطاع المتعلم تحقيقه فعلاً. فالاختبار يقيس الأهداف المتعلقة بالمادة الدراسية، بمعنى أن يكون محتوى بنوده على درجة عالية من الصدق، أي أن يقيس المهارات التي أعد من أجل قياسها وليس مهارات عارضة أخرى تتداخل معها أو لها مؤشرات مشتركة.

- 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، ص309)، فلا يمكن تطبيق اختبار على متعلم لم تمنح له فرصة تعلم نطاق سلوكي أو محتوى تربوي دراسي معين، فإختبار التحصيل فيس مادة الفيزياء مثلاً يقيس أصلاً ما أستطاع المتعلم إكتسابه 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 وأكتسبه بالفعل مع إختلاف 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.

إختبارات قياس التحصيل الدراسي :

تعد اختبارات التحصيل الدراسي أكثر الاختبارات إنتشارا وإستعمالا ما بين الاختبارات التي تقيس مختلف جوانب الشخصية، وهذا نظرا لوجود عدد متزايد من المتعلمين يتطلب معرفة درجة تحصيلهم تطبيق مختلف الاختبارات التي يقوم بنائها 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 استعمالا وأهمية هو تصنيفها حسب طريقة تفسير النتائج إلى إختبارات مرجعية الجماعة وإختبارات مرجعية المحك .

مفهوم الاختبار كوسيلة للقياس:

تشير أدبيات القياس النفسي إلى وجود تعاريف متعددة قدمت للإختبار النفسي. إذ ترى أنستازي (Anastasi) حسب موسى النبهان(2004) أن الاختبار هو أداة قياس موضوعية و مقننة للحكم بصورة كمية على عينة من السلوك. يرى 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 إستنادا إلى ميزان عددي أو نظام تصنيفي متفق عليه .

يتضح مما سبق أن الاختبار التحصيلي هو أحد أدوات القياس والتي تقدم معلومات عن مدى إكتساب وإستيعاب 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، وحتى يتحقق الهدف من الدراسة الحالية فإنه يتم تناول المشكلات المتعلقة باختبارات قياس التحصيل الدراسي والتي تعرضت لها الدراسات السابقة.

الاختبارات المقالية كوسيلة لقياس التحصيل:

تعتبر الاختبارات التحصيلية المقالية الوسيلة الأكثر إنتشارا بين المعلمين في مختلف الأنظمة التربوية في العالم، كما تعتبر الأولى من حيث الظهور التاريخي فقد أستعملت في تقويم المتعلمين قبل أن تبدأ في الظهور الاختبارات الموضوعية، وعلى الرغم من الامتقادات التي توجه لهذا النوع من البنود الاختبارية، والتي تمنح الفرصة للمجيب بأن ينشأ إجابة من إنتاجه وفق أسلوبه الخاص، مما يجعلها لا تتصف بالموضوعية وهنا أشار كل من ثورندايك وهاجن (Thorndike, Hagen) 1989 إلى أن الاعتماد على البنود المقالية المفتوحة في الحكم على مستوى المتعلم من شأنه أن ينقص من قيمة القرارات نظرا لكون هذه الاختبارات لاتحقق 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 المقال يمكن أن يقيس مستويات معرفية عليا قد لا يستطيع الاختبار الموضوعي الوصول إليها، و عليه لا يجب أن نتخلى عنها تماما وتختفي من ممارستنا التربوية، ولكن مطلوب أن نأخذ في عين الاعتبار هذه السلبيات عند إستعمالها في قياس مختلف الخصائص والقدرات عند المتعلمين ونحاول تجنبها أو التقليل منها لتحقيق تنوع إيجابي لوسائل القياس

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 إختبارات الاختيار من متعدد.

الدرس الرابع : الاختبارات الموضوعية كوسيلة لقياس التحصيل

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

والاختبارات أو البنود الموضوعية تتطلب من المتعلم تمييز الاجابة الصحيحة ووضع علامة عليها دون الحاجة إلى الكتابة أو الشرح، فالمتعلم يختار الاجابة الصحيحة مما يسلم له، وهذا يظفي على عملية تصحيح الاختبار موضوعية بحيث لا تتدخل ذاتية المصحح في التأثير بالزيادة أو النقصان على درجة الفرد، نظرا لكون الاجابة 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 بإختصار مع التركيز نمط الاختيار من متعدد لكون إختبار الدراسة الحالية من هذا النوع .

بنود الصواب والخطأ

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، أكتب الكلمة "صواب" في المكان الخالي التالي للمفردة، وإذا كانت المفردة خطأ، أوجد بديلا للكلمة أو العبارة التي تحتها خطأ والتي تجعلها صوابا، ثم أكتب البديل في الفراغ الذي يلي المفردة .

ومن الأمثلة على بنود الصح و الخطأ : حدد العبارات الصحيحة مما يلي:

- تتمدد الأجسام بالحرارة . صحيح - خطأ

- فرويد هو مؤسس مدرسة التحليل النفسي. صحيح - خطأ

- مفهوم الاختبار أعم من مفهوم المقياس. صحيح - خطأ

وقد تتطلب توجيهات مفردات اختبار الصواب و الخطأ من المتعلم إستجابة تُعطى في ثلاثة اشكال، إستجابة بوضع دائرة حول الاجابة الصحيحة، أو كتابة الاستجابة. وإما تعديل عدل إستجابة الصواب و الخطأ، أي تعديل أو تصحيح الخطأ .

-بنود الملاءمة

تتطلب بنود الملءئة من المتعلم أن يزاوج بين كلمة واحدة، أو شيء، أو رمز أو سلوك والاستجابة المرتبطة بها و المعطاة للمتعلم.و من الأمثلة على هذا النوع:

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

- عدد التلاميذ في قسم دراسي. - التصنيفي

-درجات التلاميذ في مادة دراسية. - الترتيبي

-ذكاء مجموعة من الاطفال. - المسافات

-أرقام لقاءات الدراسة. - النسبي

في هذه الحالة يتطلب من المعلم أن يصل كل عبارة من عبارات العمود الأول مع العبارة المناسبة في العمود الثاني بدون أن يأتي بأي معطيات من عنده .

ويمكن أن نلاحظ أن بنود الملاءمة يمكن اعتبارها مجموعة مختلفة من مفردة الاختيار من متعدد؛ أو بتعبير آخر، مجموعة جذور الاختيار من متعدد مرتبطة بمجموعة مفردة من الاستجابات الممكنة.

-بنود التكملة

تتطلب مفردات إختبار 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 يملء بها الفراغ من

عنده، ويستحسن النوع الأول مع تجنب كثرة الفراغات حتى لا تفقد العبارات معناها بالنسبة للمتعلم. ومن أمثلة هذا النوع من البنود.

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

بنود الاجابة القصيرة

تتطلب بنود الاجابة 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 الاجابة. 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 الاجابة القصيرة.

- ماهو عدد مستويات القياس في علم النفس و التربية ؟
- كيف يسمى نموذج التفسير الذي يعتمد على متوسط الجماعة ؟
- ماذا نسمي القيم المتحصل العددية التي نتحصل عليها من الجداول الاحصائية ؟

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 (إختيارات أو بدائل)، وإحدى هذه الاستجابات المقترحة صحيحة، بينما باقي الاستجابات غير

صحيح أو متداخلة مع الاستجابة الصحيحة، ويطلب من المتعلم إختيار الإجابة الصحيحة، ويستخدم هذا النوع من البنود لقياس معرفة الحقائق العلمية وإستيعابها 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 إختبار يتوفر على الشروط الموضوعية.

- تمتعها بدرجة عالية من الثبات، فالدرجات لا تتأثر بذاتية المصحح لأن الاجابات الصحيحة محددة مسبقا وبشكل دقيق، فلو أعدنا تصحيح الاختبار عدة مرات من طرف مصححين آخرين فسوف نتحصل على نفس النتيجة، بحيث الاجابة النموذجية جاهزة و واحدة .

و يمكن صياغة بند من نوع إختيار من متعدد كما يلي:

تتدرج مستويات القياس النفسي من الاسمي إلى النسبي بناء على:

- المعالجات الإحصائية المستعملة في كل مستوى.

- الخصائص الشكلية لكل مستوى.

- العمليات الحسابية الممكنة في كل مستوى.

- تعقد طبيعة المعطيات في كل مستوى..

بنود اختبار الاختيار من متعدد هي حاليا الأكثر استخداما في ميدان الاختبارات التحصيلية المقننة أكثر من الأنواع الأخرى من المفردات أو الاختبارات، والصيغة المذكورة لمفردة

الاختيار من متعدد مع الجذر تسأل أو تدل ضمنا على سؤال مباشر وهي توفر طريقة مناسبة و حقيقية وطبيعية لاختبار تحصيل الفرد، وقد اقترح مجموعة من الخصائص أو الشروط الواجب توفرها في بنود إختبار الإختيار من متعدد حتى يحقق هذا النوع الغرض الذي وضع من أجله، و أهم هذه الشروط هي :

أ- إستخدم جذرا كافيا

يطلب كتابة الجذر بكلمات واضحة ومختصرة بقدر الامكان، كما يجب أن يتضمّن معلومات كافية لكي يكون منطقيا 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 .

تجنب التكرار غير الضروري في البدائل

عندما 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 إختبارات التحصيل هو:

أ- الصدق التلازمي. / ب- الصدق التنبؤي.

ج- الصدق الظاهري. / ء- الصدق العاملي

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

نوع الصدق الذي يناسب بدرجة أفضل إختبارات التحصيل هو الصدق:

أ- التلازمي. / ب- التنبؤي.

ج- الظاهري. / ء- العاملي

الجزر يجب أن يكون صحيحا نحويا في ذاته وفي علاقته مع البدائل

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

د - الجزر المناسب به فاعل واحد :

عند صياغة العبارات فلا نضع المبادئ و النظريات و القوانين في جذر واحد فينشئت ذهن المتعلم و لايفهم 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 يهمننا هو الجذر والذي إحتوى على أكثر من فاعل، وبالتالي يمكن تحسينه .

في النظام العشري للعد، العدد 10 هو الوحدة الأساسية للعد، وفي النظام الاثنا عشري، الوحدة الأساسية هي :

الصيغة الايجابية للجذر

كقاعدة، يستحسن أن تكون العبارات أو البنود في صيغة إيجابية، و كلما 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 إنتباه المتعلم مثلا. ويجب التأكد من أنه توجد إجابة واحدة فقط هي " الأصح ". 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 لإختلاف قيمة معامل الارتباط بإختلاف القانون المطبق للحصول عليه، ولكن المطلوب هو إختيار التفسير الرئيسي 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 و لكنها ليس الأشمل.

الصيغة السلبية للجذر

عندما نستخدم المعنى السلبي في الجذر، يتطلب أن نلفت نظر المتعلمين إلى السلبية بالحروف الكبيرة أو بوضع خط أسفلها أو بكليهما معا، 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 ماعدا، فهذه الصياغات السلبية تجعل البند غامضا بالنسبة للمتعلم ويستحسن كتابة بنود تكون جذورها إيجابية الصياغة.

طول الجذر والبدائل :

يستحسن تجنب الجذر القصير عندما تكون الاستجابات طويلة، والخلل في ذلك أن المتعلم يحتاج إلى وقت أطول ليقراً أربع أو خمس إستجابات طويلة بدلا من أن يقرأ جذرا واحد طويلا و بدائل أقصر وأكثر تركيز، 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 الإسمي أكثر تعقيدا 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 و يشير 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 إختبارات فردية و إختبارات :

- 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 الاجابة الصحيحة من دون أن يكون فعلا متمكن منها، وعليه يجب تجنب هذا النوع من الصياغات التي لا تسمح بالتحقق هل فعلا وصل المتعلم إلى الاجابة الصحيحة بفعل تمكنه أم أن التلميح المتضمن في السؤال هو الذي أوصله إلى الاجابة الصحيحة.

الاستجابات يجب أن تكون متماثلة مع بعضها و منسقة نحويا

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

أي أنواع الصدق يجب إستخدامها للتحقق من صدق إختيار تحصيلي ؟

- الصدق العاملي . - صدق المحتوى.

- الصدق التلازمي . - صدق الأمل .

ففي هذه الحالة، واضح أن صدق الأمل، ليس نوعا من أنواع الصدق، وهذه الاستجابة أو البديل أستخدم من أجل الوصول إلى أربع بدائل فقط ولا علاقة له بالجدر، في حين أن الاستجابات الخاطئة، المعروفة بإسم البدائل أو المشتتات. فيجب أن تبدو مقبولة ظاهريا بجعلها مألوفة كإجابة الصحيحة، و تنتمي إلى نفس المفهوم و منطقية بحيث يحتمل أن تكون هي الاجابة الصحيحة ولها تقريبا نفس طول البدائل الأخرى.

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 و ليس بطريقة عشوائية، فمثلا إذا قولنا ما هي السنة التي قدم فيها بنييه إختباره للذكاء:

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

-تجنب وضع الاجابة 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،أ، ب،ب،ب، ج،ج،ج) فهذه المنهجية تزيد من إحتمال وصول المتعلم إلى الاجابة الصحيحة عن طريق التخمين.وعليه يمكن وضع الاختيارات بحيث تظهر الإجابة الصحيحة في كل مواضع متنوعة (مرات أ و أخرى ب، و ج، و د، وهكذا) لعدد متساوي تقريبا من المرات .

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

إذن إختبارات الاختيار من متعدد تصلح لقياس مهارات التعرف على المفاهيم والمصطلحات العلمية التي سبق للمتعلم و أن تناولها، ولكنها غير قاصرة على قياس مستوى معين من المعارف أو الأهداف، بل تشمل قياس جميع مستويات المجال

المعرفي، كالفهم و التطبيق 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 .

الدرس الخامس: النظرية الكلاسيكية في قياس التحصيل الدراسي

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 للإلتحاق 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 إرتكزت 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 في 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، فليجأ الباحث عندما يريد أن يرفع من قيمة معامل ثبات إختباره إلى زيادة عدد بنود الاختبار مما يزيد في درجة صدق محتواه ومنه الزيادة في قيمة ثباته.

-تكون مقارنة درجات الاختبار أفضل عندما تتوفر عدة صورة متكافئة للإختبار نفسه و يمكن أن تطبق عدة مرات في ظروف مشلبهة 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 بناءاً على متوسط العينة التي طبق عليها الاختبار. فمعيار تفسير النتائج يستمد من النتائج الفعلية للأفراد على الاختبار.

-خصائص مقاييس المسافة تتحقق بالحصول على توزيع إعتدالي 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.

و على 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 و تجعل الاعتماد على نتائجها حسب صلاح الدين محمود
علام (2006) منقوص الفائدة نظرا 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 و وحدات من الاختبار إذا كانت غير
مميزة بين الأفراد بغض النظر عن قيمتها العلمية مما يؤثر في بناء الاختبار.

-الاقتصار على إتخاذ 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 يتلائم 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 النتائج 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 تدرج صعوبة البنود باختلاف قدرات
الأفراد، كما أن قياس قدرات الأفراد يتباين بتباين صعوبة البنود (ناهدة سكر، 2003).

6-مشكلات تواجه 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 و التفسيرات
التي تتبع هذه النتائج، وقد أشار كل من (أمينة كاظم 1988. صلاح الدين علام، 2003 .
Wright، 1998)إلى بعض المشاكل التي تواجه الأساليب التقليدية في القياس النفسي
والتربوي، وفي الوقت نفسه تعد بمثابة نقائص في هذه الأساليب تجعل من تطبيق في مختلف
الممارسات والاستفادة العلمية والعملية من نتائجها محدود، كما أن العامل في الميدان التربوي

يجب أن يكون مطلعاً على هذه المشكلات حتى يأخذها بعين الاعتبار ويفسر نتائج إختباراته بعيداً عنها، و لكن في إطار نسبي مرتبط بها، و يمكن تلخيصها كما يلي:

تقيد الدرجة الكلية للاختبار ببنوده:

تتلخص هذه المشكلة في أن الدرجة التي يتحصل عليها الفرد في إختبار 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 أستبدل 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 حسب صلاح الدين محمود علام (2006)، 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.

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 مقياساً خطياً (عبد الوارث عبده الرازي 2001)، 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 ثابتاً، 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 واقعياً حتى لو كانت تبدو ظاهرياً متساوية.

القياس في أكثر من بعد

تتميز أداة القياس الفيزيقي بالدقة أكثر من المقاييس النفسية و التربوية، كما أن المقاييس الفيزيكية تقيس شيئاً واحداً وقت واحد كالطول مثلاً، إلا أن إختباراً 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 ذكائه 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 إنعكاس لقدرة واحدة فقط هي موضوع القياس، والتي وضع الاختبار من أجلها.

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، و لكن 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 نظراً لتدخل عوامل كثيرة مؤثرة في عملية القياس المبني على أسس تقليدية لتفسير كيفية حدوث التفاعل بين الأفراد و البنود وصولاً إلى إستجابات قابلة للقياس و تعبر عن السمة الكامنة.

تغير بنية الاختبار ومعاني بنوده مع الزمن

تعاني أساليب القياس التقليدية من تغير معاني البنود المستخدمة مع مرور الزمن ولذلك فإن التكوين العامي للاختبار غير ثابت، ويرجع ذلك إلى تغير الظروف البيئية.

والثقافية وتغير قدرات وخصائص الأفراد، بالظواهر النفسية من ميزاتها الأساسية أنها غير ثابتة ومتغيرة في فترات زمنية متقاربة، هذه العوامل تؤدي إلى تغير درجات الفرد بصورة يصعب التنبؤ بها مستقبلاً في ظل الأساليب التقليدية لقياس الأداء بصفة عامة و التحصيل بصفة خاصة.

اعتماد خصائص البنود على خصائص الأفراد

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

أمام هذه المشكلات التي طرحها الاختبارات مرجعية الجماعة، أو ما أصطلح على تسميته بالنظرية الكلاسيكية في القياس النفسي، لم يقف علماء القياس أمام هذه المشكلات بدون إقترح الحلول، بل قاموا بجهودا 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 الأسئلة المكونة لأداة القياس، وهو ما يتم عرضه في الفصل الخاص بنظرية الاستجابة لبنود الاختبار.

المقاربة الحديثة في قياس التحصيل الدراسي:

نظرا للمشكلات التي طرحها تطبيق أسس ومبادئ النظرية الكلاسيكية في قياس التحصيل الدراسي نقام علماء القياس النفسي بجهود متعاظمة للتغلب على هذه المشكلات والتي تؤثر في الخصائص السيكومترية للاختبارات التحصيلية، مما يؤثر في دقة وصحة النتائج، فقد تم التوصل إلى نماذج جديدة تجعل القياس موضوعي بمعنى عدم تأثر درجة الفرد بخصائص عينة الأفراد، ولا خصائص عينة البنود المكونة لأداة القياس، هذه الخصائص تحققها نظرية السمات الكامنة التي خصصنا لنا فصل نظرا لكونها الخلفية التي بنيت عليها الدراسة.

والمقاربة الحديثة في قياس التحصيل الدراسي بصفة عامة تقوم على أسس القياس المحكي الذي يهدف إلى تقويم أداء الفرد وتحديد مدى تقدمه بالنسبة لمجموعة من المهارات أو الأهداف، ومعرفة جودة الأداء في ضوء محكات محددة مسبقاً، أمينة محمد كاظم وآخرون، 1989 والاختبار محكي المرجع هو الاختبار الذي صمم عن قصد لإعطاء قياسات تفسر

مباشرة بالنسبة إلى معايير أداء محددة. بمعنى أنه يهدف إلى تقدير أداء الفرد بصرف النظر عن مستوى أداء أقرانه، حيث تفسر درجة الفرد على الاختبار في ضوء محك يحدد حسب الأهداف الموضوعية للقياس.

والاختبار مرجعي المحك يسهم في تقويم كل تلميذ بناءً على مستواه وقدراته حيث يقوم أداء التلميذ على ما يحتاجه هو بالنسبة لقدرته الفردية، وليس بالنسبة لمقارنته بمتوسط أداء مجموعته، أي أن التفسيرات مرجعه المحك تساعد على شرح ماذا يستطيع عمله الفرد بدون أداء الآخرين، والاختبارات المرجعة إلى المحك تكون مناسبة في الحالات:

- عندما تكون الموضوعات الدراسية تراكمية ومرتبطة مع بعضها البعض.

- عندما يكون الهدف الوصول إلى مستوى تمكن محدد، واكتساب مهارات، قصد شهادات أو درجات علمية .

- عندما يكون الهدف هو تحديد نواحي القوة والضعف لدى الفرد فقد توصلت نتائج دراسة عصام الدسوقي (1998) إلى فعالية نموذج (أنجوف) في تحديد مستوى الأداء المقبول لاختبار محكي المرجع، في وحدة من مقرر الجبر للمستوى الأول ثانوي، حيث كانت درجة القطع تساوي (74%)، بينما بلغت درجة القطع لنفس الاختبار باستخدام نموذج (هوفستي) (68%)، و أرجعت الدراسة هذا الاختلاف إلى أن نموذج (أنجوف) يعتمد على أحكام المتخصصين في مجال الاختبار لتحديد درجة القطع، بينما يعتمد نموذج (هوفستي) على أحكام المتخصصين بالإضافة إلى الأداء الفعلي للأفراد، مما يعني أن عينة الدراسة لا تلعب دوراً في تحديد درجة القطع للاختبار المبني وفق النظرية الحديثة في القياس .

وفي دراسة مصطفى كامل (1999) تم إعداد اختبار محكي المرجع وقد استخدم محك الإتيان (80%) من الدرجة ليمثل الحد الأدنى الذي ينبغي على المتعلم الوصول إليه لكي يعتبر متمكناً، وقد أسفرت النتائج عن فعالية استراتيجية التعلم حتى التمكن في الحكم على مستوى تمكن الطلاب من المهارات التدريسية لمقرر التقويم التربوي.

ونظراً لأهمية المقاربة الحديثة في القياس النفسي والتي تقوم على الاختبارات المحكية في قياس التحصيل الدراسي للأفراد، إعتقاداً على نظرية السمات الكامنة أو ما يسمى بنظرية

الاستجابة لمفردات الاختبار، حيث يمكن للباحث أن يعتمد على واحد من نماذجها والتي يرى الباحث أنه يلائم طبيعة المعطيات ويحقق أهدافه .

و نظراً لهذه الأهمية فقد خصصنا فصل لنظرية الاستجابة لمفردات الاختبار كمقاربة حديثة في قياس التحصيل، ونظرية الاستجابة للمفردة تتضمن عدة نماذج تسمح بالقياس الموضوعي، حيث يتم التركيز على المفردة الاختبارية مما يسمح بإضافة أو حذف أو تعديل مفردات دون أن يتأثر الاختبار ككل، مما يساعد في التغلب على مشكلات القياس التي التي عجزت عنها النظرية الكلاسيكية.

ومن المتطلبات التي يلزم توافرها لكي يكون القياس النفسي أكثر موضوعية حسب نظرية السمات الكامنة:

- توفر بنود صادقة يمكنها تعريف المتغير موضوع القياس إجرائياً.
- صدق التدرج لهذه البنود، بحيث يمكنها تمثيل هذا المتغير في صورة خطية.
- أنماط استجابات الأفراد يمكنها تحديد وضعهم على متصل المتغير.
- توافق بين درجات الأفراد وخصائص المفردات بحيث تؤدي إلى تقديرات لمستويات الأفراد لا تعتمد على اختبار في حد ذاته أو بنود معينة.
- قياسات خطية يمكن استخدامها لدراسة النمو أو للمقارنة بين الأفراد.

هكذا قدمت المقاربة الحديثة حلولاً نظرية و تطبيقية إحصائية لكثير من المشاكل التي واجهت النظرية الكلاسيكية في القياس، حيث تعتمد التقديرات 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 تغير معاملات الصدق والثبات وصعوبة المفردات بتغيير قدرات الأفراد. كما أن الدرجة التي يحصل عليها فرد على مفردات الاختبار يمكن جمعها، كما لو كانت تمثل ميزاناً خطياً، والحقيقة أن هذا الميزان منحنيًا، فالفرق الثابت بين درجتين من درجات الاختبار يختلف معناه بالنسبة لدرجة القدرة المقاسة بحسب موقع الدرجات على متصل القدرة .

ففي إطار نظرية الاستجابة لمفردات الاختبار، تكون المتغيرات الكامنة هي التي تحدد طبيعة ودرجة السلوك، حيث يتم الاستدلال عليها من ملاحظة أداء الفرد عندما يستجيب لبنود الاختبار، ومن ثم فإن خصائص المفردات والأفراد يستدل عليها من السلوك، مما يجعل النظرية الحديثة في القياس النفسي تؤدي إلى تقديم تفسيرات موضوعية لمختلف الأنماط السلوكية للأفراد.