



جامعة الجيلالي بونعامة-خميس مليانة –

اسم المادة

القياس التربوي وبناء الاختبارات المدرسية

الطلبة المعنيين

الكلية	القسم	السنة	التخصص
العلوم الاجتماعية والإنسانية	علم النفس وعلوم التربية	الثانية	علوم التربية : ارشاد وتوجيه

معلومات عن المقياس

القياس التربوي وبناء الاختبارات المدرسية

المستوى: السنة الثانية

السداسيين الثالث والرابع

التخصص: علوم التربية

وحدة التعليم: المنهجية

المعامل: 02

المحاضرة رقم 07: ماهية الاختبارات، واستخداماتها.

1. تعرف الاختبار:

الاختبار النفسي بالضرورة هو مقياس موضوعي ومقتن لعينة من السلوك والاختبارات النفسية تماثل الاختبارات في أي علم آخر، من حيث إن الملاحظات تجري على عينة صغيرة ولكن يتم انتقاؤها بعناية ، وبهذا الشكل يقوم علماء النفس بأجراءتهم بطريقة مماثلة الى حد كبير لما يقوم به عالم الكيمياء الحيوي الذي يختبر دم مريض معين، أو مصدر مياه المجتمع المحلي، حيث يحلل عينة أو أكثر منها ، فإذا أراد عالم النفس

القياس التربوي وبناء الاختبارات المدرسية

اختبار مقدار المفردات اللغوية لدى الطفل ، أو قدرة كاتب على أداء العمليات الحسابية أو تآزر العين واليد لدى قائد الطائرة ، فانه يفحص أداءهم بمجموعة ممثلة من الكلمات ، أو المسائل الحسابية ، أو الاختبارات الحركية ، ويعتمد ما إذا كان الاختبار يغطي بدرجة كافية السلوك موضع الاعتبار على عدد و طبيعة المفردات في العينة ، فمثلا ، اختبار الحساب الذي يشتمل فقط على خمس مسائل أو يشتمل فقط على مفردات تتطلب عملية الضرب ، سوف يكون مقياسا ضعيفا للمهارة الحسابية لدى الفرد ، واختبار المفردات اللغوية الذي يشتمل فقط على مفردات تتعلق بكرة القاعدة سوف يصعب أن يقدم تقديرا للمدى الكلي لمفردات الطفل اللغوية .

وتعتمد القيمة التشخيصية أو التنبؤية للاختبار النفسي على درجة التي يساعد بها كمؤشر لمجال سلوك متسع نسبيا وجوهري . وقياس عينة السلوك التي يشتمل عليها الاختبار مباشرة يكون نادرا – ان لم يكن مطلقا – الغرض من الاختبارات النفسية. فمعرفة الطفل لقائمة معينة تشتمل على 50 كلمة لا تكون بذاتها موضع اهتمام كبير. وكذلك أداء المتقدم لعمل على مجموعة معينة تشتمل على 20 مسألة حسابية لا يكون له أهمية كبيرة ، غير انه اذا كان ممكن تبيان انه يوجد تناظر وثيق بين معرفة الطفل قائمة الكلمات واتقانه الكلي للمفردات اللغوية ، او يبين درجة المتقدم لعمل في مسائل حسابية وادائه الحسابي في العمل ، فان الاختبارات بذلك تخدم الغرض منها (انا ستازي او ريانا ، 2015 ، ترجمة علام)

يعرف الاختبار على انه : إجراء منظم لقياس سمة ما ، من خلال عينة من السلوك ، يشير هذا التعريف الى عدة أمور أهمها :

ان الاختبار اجراء منظم ، فالاختبار طريقة تتألف من خطوات معينة متتابعة ، وكل خطوة تتطلب مجموعة من القوانين .

السلوك نلاحظ أداء المفحوص في موقف اختباري معين.

العينة : لكل سمة من السمات عدد من السلوكيات الدالة عليها ، ولبناء اختبار وقياس أداء ما لا بد من اتباع مراحل معينة .

2. تصنيفاتها :

يمكن تصنيف الاختبارات إلى عدة تصنيفات و هذا حسب :

1.2. الوظيفة او ما تقيسه .

2.2. حسب شروط الإجراء .

3.2. كيفية الأداء .

4.2. مادة الاختبار (طبيعة المادة) .

1.2. التصنيف على أساس الوظيفة : يمكن أن تصنفها إلى خمسة أنواع :

1.1.2. اختبارات تحصيلية : و هي تقيس مدى اكتساب الفرد لمعلومات أو مهارات معينة في

موضوع ما نتيجة تعليم خاص (القراءة ، الحساب ، اختبارات الكفاءة على الكمبيوتر) .

2.1.2. اختبارات الذكاء : و هي تقيس القدرة العقلية العامة التي تنعكس في سرعة الفهم ، القدرة على التعلم، الكفاءة العامة ، سرعة الإدراك ، القدرة على التكيف . مثل اختبار : ستانفورد بنيه ، وكسلر .

3.1.2. اختبارات القدرات الخاصة : تنتبأ بمدى قدرة الفرد على التعليم أو التدريب على مهنة معينة كالقدرة الميكانيكية و الموسيقية و الحسابية . هذا النوع من الاختبارات في بطاريات بحيث تقيس كل بطارية عددا من القدرات الخاصة .

4.1.2. اختبارات الميول : و هي تقيس ميول الأفراد نحو الأنشطة أو مهن معينة و من أمثلة الاختبارات المهنية : الصفحة المهنية لسترونج ، اختبار التفضيل لكوذر .

5.1.2. اختبارات الاتجاهات و القيم : يقيس هذا النوع من الاختبارات طبيعة و أبعاد الاتجاهات و المعتقدات التي يتمسك بها الفرد إزاء أفراد آخرين أو إزاء مختلف القضايا الاجتماعية و الأنشطة من أشهرها : مقياس ليكرت ، مقياس ألبرت للقيم ، مقياس ثرمستوم .

2.2. تصنيف الاختبارات حسب شروط الإجراء : تصنف إلى نوعين :

أ. الاختبارات الفردية : يكون موقف القياس فردي (فاحص/مفحوص) (مرشد/مسترشد) و يتطلب هذا النوع إحصائيين مدربين ذوي مهارات كبيرة في التطبيق و التحليل و من أمثلتها اختبارات الذكاء و الاختبارات الإسقاطية.

ب. الاختبارات الجماعية : و هي اختبارات تطبق على الأفراد في وقت واحد لها فائدة كبيرة في ميدان التربية ، الجيش ، الصناعة . و من أمثلة ذلك : اختبار الجيش " ألفا " ، اختبار الجيش " بيتا " .

3.2. تصنيف الاختبارات حسب طبيعة مادتها :

إن الاختبارات النفسية التي تعتمد على اللغة يصعب تطبيقها على كل المجتمعات و ذلك لتأثرها بالثقافة ، حيث أن الاختبار الذي وضع في محيط اجتماعي ما لا يصلح للتطبيق في محيط آخر ، و للتخلص من هذه المشكلة وضع الباحثون بعض الاختبارات اعتبرت متحررة من الثقافة ، و من ذلك اختبار " بيتا للجيش " الذي وضع أثناء الحرب العالمية الأولى بهدف تطبيقه على الجنود الذي ينتمون إلى ثقافات أخرى غير الثقافة الأمريكية و يتألف هذا الاختبار من متاهات و رموز و أشكال مصورة ناقصة .

إن جميع هذه المثيرات لا تتأثر بالثقافة إلا تأثرا قليلا ، من أشهر اختبارات الذكاء المتحررة من الثقافة اختبار المصفوفات لرافين .

4.2. حسب كيفية الأداء :

المقصود بكيفية الأداء التنوع الذي يصدر عن المفحوص أثناء الإجابة عن مختلف فقرات الاختبار ، إذ تميز بين اختبار :

الورقة و القلم : يفكر المفحوص في المشكلات التي تعرض له تفكيرا ضمنيا ثم يسجل نتيجة أفكاره .

الاختبارات العملية : يقوم المفحوص بمعالجة المواد التي يتألف منها الاختبار معالجة صريحة ، كبناء المكعبات ، تفكيك الأجهزة ، وضع البراغي ، الكتابة على الكمبيوتر .

3. مجالات استعمال الاختبارات النفسية و التربوية :

1.3. في التربية و التعليم : تعتبر المؤسسات التربوية و التعليمية (المدارس و مراكز التكوين) من أكثر المؤسسات استعمالاً للاختبارات النفسية و التربوية ، فهي تستعمل للأغراض التالية :

- أ. الكشف عن المتأخرين دراسياً ، و معرفة أسباب هذا التأخر بمعرفة جوانب حياته النفسية .
- ب. اختبار الطلبة الجدد .
- ج. توجيه التلاميذ أو المتكولين إلى الدراسات و المهن التي تتناسب و قدراتهم .
- د. إرشاد التلاميذ و الطلبة الذين يجدون مشاكل في دراستهم .
- هـ. تقويم تحصيل التلاميذ و معرفة مدى فعالية برامج التعليم و مناهجه .
- و. تصنيف التلاميذ : تساعد الاختبارات في تصنيف التلاميذ حسب قدراتهم إلى أقسام أو مستويات ليكون التجانس بين الأقسام من حيث القدرات و الاستعدادات لضمان نجاح العملية التربوية .

2.3. في الصناعة و الإدارة : (المجال المهني)

من المجالات الأساسية التي يستخدم فيها القياس النفسي المجال المهني و الصناعي فالاختبارات و المقاييس النفسية تطبق في :

الكشف – الاختيار – التصنيف – التوجيه – التقويم – التدريب المهني .

و كل هذا بهدف وضع الفرد المناسب في (التخصص/ المهنة) المناسبة . فإذا كان التوجيه يقتضي معرفة خصائص الأفراد و توجيههم وفق هذه الخصائص و ذلك لاحتمال (التنبؤ) بأنهم سيبرزون فيها أعلى قدر من النجاح .

التقويم : إذ تستخدم الاختبارات في تقويم أداء العمال و مدى فعالية طرق العمل و نظم الإشراف .

تستخدم في الكشف الأفراد الذين يتعرضون لحوادث عمل أكثر من غيرهم فيتم استبعادهم

3. في الجيش : (في المجال العسكري)

شهدت حركة القياس النفسي تطوراً كبيراً بعد الحربين ، و قد استعملت الدول خلال الحربين اختبارات نفسية لاختبار الجنود و تصنيفهم على مختلف المهام حسب قدراتهم ، و قد كان لها دور كبير في زيادة أداء الجيوش في الميدان العسكري .

4. في المجال العيادي :

تزايد الاهتمام بالصحة النفسية للأفراد مع تطور مختلف ميادين علم النفس ، كما تزايد الوعي بإمكانية معالجة و مساعدة المرضى نفسياً و هنا ظهرت الحاجة الماسة إلى الاختبارات النفسية تمكن من الكشف عن مختلف أعراض الاضطرابات النفسية فعند تعاملنا مع أي فرد

يعاني من اضطراب نفسي فمن المفيد منذ البداية أن نعرف ما هو الاضطراب و كيفية الاستجابة للعلاج .

في حالة اختيارنا لاختبارات مناسبة يمكننا من تحديد الأعراض التي يعاني منها الفرد و من ثمة اقتراح الأسلوب الأنجح للعلاج إذن يستخدم القياس النفسي في مجال العلاج بصفة واسعة فعلى أساس نتائج تطبيق الاختبارات النفسية و العقلية ، يمكن تشخيص الاضطرابات التي يعاني منها الفرد و اقتراح خطة العلاج و لا يقتصر دور القياس على عمليتي التشخيص و العلاج بل يتضمن معرفة قدرات المريض و ذكائه العام لمعرفة مدى أثر هذه العوامل في اضطرابه و مدى توظيفها في إعادة تكييف مع الحالة العامة .

المحاضرة رقم (08): خطوات ومراحل بناء الاختبارات النفسية والتربوية

2. خطوات ومراحل بناء الاختبارات النفسية والتربوية :

يتم تلخيص هذه المراحل اعتمادا على مراجع عدة نذكر منها (مقدم عبد الحفيظ ،2003، عبد الهادي السيد عبده ،فاروق السيد عثمان(2002)، صلاح احمد مراد ، امين علي سليمان ، 2005، سامي محمد ملحم ، 2015).

الخطوة الأولى: تحديد الهدف بدقة (ما الذي أريد بناءه ؟ لماذا ؟ لمن ؟)

في هذه الخطوة يتم تحديد السمة او الخاصية المراد قياسها ، ان اول خطوة في بناء المقاييس والتي يجب على معد المقياس ان ينتبه لها هي ، تحديد الهدف الذي يرمي اليه من وراء بناءه أو إنشاء للمقياس

مثال 1: تصميم مقياس بغرض قياس "سمة الايثار"

مثال 2 : تصميم مقياس بغرض قياس "تحصيل الطلبة في مادة علم الاجتماع".

لا بد من تحديد مفهوم السمة المراد قياسها اصطلاحا بالرجوع الى التراث النظري المتوفر لدى الباحث (قواميس ، موسوعات علمية ، كتب ، مقاييس سابقة ...) (وهذه الخطوة تعبر عن تحديد الأهداف العامة)

بعد تحديد الأهداف العامة من خلال التعريف الاصطلاحي للسمة المراد قياسها ينتقل معد المقياس او الباحث الى الخطوة الثانية .

الخطوة الثانية : تحديد النطاق السلوكي للسمة او الخاصية

الخطوة الثالثة : تحديد الأهداف بصورة إجرائية، ماذا نعني بصورة إجرائية إمكانية

قياس الأهداف بصورة كمية ، وتأتي هذه الخطوة لتحويل الأهداف العامة الى اهداف يمكن قياسها بطريقة يسهل ترجمتها الى مجموعة من المفردات يقوم المبحوث بالإجابة عنها ، ويتم ذلك من خلال تقسيم الظاهرة او السمة الى عناصرها الأولية (الابعاد) حيث يمثل كل

عنصر (مستخرج من الاطار النظري طبعا) اطارا مرجعيا يمكن اشتقاق الفقرات منه ، ومن ثمة كتابتها .

المحاضرة رقم 09 : الخصائص السيكمترية للأداة القياس

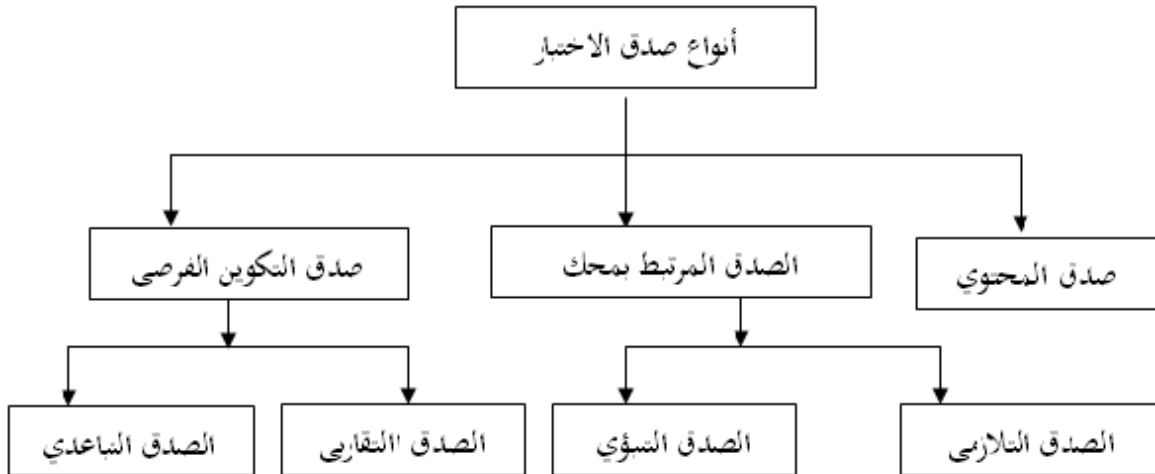
من خلال الخطوات بناء الاختبارات التي تم التطرق إليها في المحاضرة رقم (08) تعتبر الخطوة الثامنة ، وهي خطوة التحقق من صلاحية الأداة المبنية قبل إخراجها بشكل نهائي من أهم الخطوات، ويتم ذلك من خلال حساب الخصائص السيكمترية والتي سنذكر أنواعها فيما يلي:

1. الخصائص السيكمترية :

1.1. الصدق وأنواعه:

إن اللجنة القومية الأمريكية للقياس و التقويم (NCME) حددت ثلاث مصطلحات رئيسية يشير كل منها إلى نوع من أنواع الصدق ملخصة في الشكل التالي حسب علام (2006):

الشكل(09): يوضح أنواع صدق الاختبارات



الجدير بالذكر أن اللجنة أكدت هذه الأنواع الثلاثة مترابطة إجرائيا و منطقيا، فنادرا ما يكون أحد هذه الأنواع ذو أهمية بمعزل عن النوعين الآخرين.

وهذا ما يجعل الكثير من علماء القياس يرون أن الأدلة التي تسترشد بها للتحقق من الصدق وليس الاختبار هي ما ينبغي أن يطلق عليها مفهوم " الصدق " ، ولا يجب التفكير في الأنواع

الثلاثة على أنها مختلفة و إنما كأدلة متعددة عن مدى صلاحية الإختبار في ترشيد القرارات، حسب علام (2002).
إن تكامل هذه الأنواع و تداخلها، لا يمنع من محاولة التطرق إلى كل واحدة منها على إحدى بالشرح و التفصيل.

1.1.2. صدق المحتوى Content Validity:

يطلق عليه أيضا صدق المضمون أو الصدق المنطقي أو الصدق بحكم التعريف ،أو صدق عينة الاختبار ، حيث يرى فرج (2000) أن المعنى الأخير هو أقرب المعاني للمقصود فالاهتمام الأساسي هنا ينصب على ما إذا كان مجال سلوكي معين و محدد بشكل دقيق ممثلا في شكل مجموعة من البنود بصورة مناسبة.

يوضح علام (2006) أن صدق محتوى الاختبار يتعلق أساسا بالحكم على مدى كفاية مفرداته كعينة ممثلة لنطاق محتوى أو أهداف يفترض أن الاختبار يقيسها ،ويمكن أن ينطبق صدق المحتوى على جميع أنواع الاختبارات المعرفية و مقاييس الشخصية ،إلا أنه يناسب بدرجة أكبر الإختبارات التحصيلية المقننة .

و الحقيقة أنه لا يوجد مؤشر كمي لتقدير صدق المحتوى، وإنما يتم تقييم هذا النوع من الصدق بداية من بناء الاختبار حيث يبدأ التقييم حسب (anastasi & urbiaa 1997) بداية من إختيار الفقرات الملائمة ،بالنسبة للاختبارات التحصيلية مثلا: فإن إعداد الفقرات لابد أن يسبق ببحث شامل و متعمق، ومنتظم لخطط المادة و مقرراتها بالإضافة إلى استشارة الخبراء و المختصين و بناءا على المعلومات التي تم جمعها في هذه المرحلة تكون مواصفات الاختبار قد تم إعدادها بحيث تحدد هذه المواصفات جوانب المحتوى أو المواضيع التي يجب أن يشملها الإختبار، والأهداف السلوكية أو العمليات التي سيتم قياسها و الأهمية النسبية لكل من هذه المواضيع و العمليات، وبالتالي تحديد عدد الفقرات لكل موضوع .

كما يستند تقدير صدق المحتوى إلى ثلاث فروض صاغها لينون lenon و أوردها علام (2006) كالتالي:

- 1- ينبغي أن يكون المجال الذي يختبر فيه المفحوصين محددا و يمكن تعريفه تعريفا دقيقا.
- 2- يمكن اختبار عينة من الأسئلة من هذا المجال بطريقة هادفة و مناسبة .
- 3- يمكن تحديد عينة الأسئلة و طريقة المعاينات المستخدمة بدقة كافية لكي يتمكن مستخدم الأداة من الحكم على مدى تمثيل عينة الأسئلة للمجال الذي تقيسه.

فالمقصود بصدق المحتوى إذا مدى تمثيل بنود الاختبار و ملائمتها لما تقيسه.

2.1.2. الصدق المرتبط بمحك Criterion Related Validity :

يقصد بالمحك: "قياس أو إختبار مستقل خارجي لما تهدف الأداة إلى قياسه أو التنبؤ به أو تشخيصه، أو ميزان لتحديد صلاحية الأداة أو إختبار لإختبار " .(أبو حطب ،1997،ص148).

يشير عودة (1997) أن هذا النوع من الصدق أهم من غيره عندما يكون الغرض من الإختبار هو التنبؤ بسلوك مرتقب مثل التنبؤ بنجاح الطالب في التحصيل الجامعي من تحصيله في المرحلة الثانوية أو التنبؤ بنجاح موظف في عمله من معدله التراكمي الجامعي.

يستنتج من هذا أن صدق المحتوى هنا شرط غير لازم بالنسبة للصدق المرتبط بمحك، فالإختبار الذي يكون صادق المحتوى يعمل كمتنبئ أما الإختبار الذي يعمل كمتنبئ ليس بالضرورة أن يعمل كممثل فأحيانا يصعب حصر مهام العمل و العوامل المؤثرة في النجاح بهذا العمل فإن الحديث عن الصدق التنبؤي هذا هو الأهم.

و يمكن التميز بين نوعين من هذا الصدق: الصدق التلازمي Concurrent Validity و الصدق التنبؤي Predictive Validity، و الفرق الرئيسي بينهما حسب علام (2006) يتعلق بالمدة الزمنية الفاصلة بين تطبيق الإختبار المراد التحقق من صدقه و تطبيق المحك.

أ- الصدق التلازمي:

تتلخص الإجراءات في هذا النوع من الصدق، بجمع المعلومات على الإختبار الذي نبحث عن صدقه، وعلى المحك بنفس الفترة الزمنية، أو بفواصل زمني قصير جدا، بمعنى أن الأفراد الذين يمثلون عينة الصدق يخضعون للمحك عند جمع المعلومات على المتنبئ، عودة (1999)، أي أننا نقارن بين درجات الأفراد على الإختبار و درجاتهم على مقياس موضوعي آخر، لقد تم الإعتماد في هذه الدراسة على معدلات الطلاب في التحصيل الدراسي (كمحك).

حيث تم التحقق من صدق مقياس "مداخل الدراسة" عن طريقة حساب معامل الارتباط بين درجات الطلاب في مقياس مداخل الدراسة و معدلاتهم في التحصيل الدراسي و يمكن تبرير هذا الاختيار بأنه لا يمكن الفصل بين توجهات الطلبة نحو الدراسة أثناء معالجتهم للمعلومات الموجودة في المقررات التي يدرسونها و معدلاتهم في التحصيل الدراسي و هو ما جعل الباحث يعتمد على معدلات الطلبة كمحك لتأكد من صدق مقياس مداخل الدراسة، كما و أن الدراسات السابقة التي تم الإطلاع عليها استخدمت كذلك هذه المعدلات كمحك.

ب- الصدق التنبؤي:

يعني الصدق المحك التنبؤي بإيجاد علاقة بين النتائج الإختبار و نتائج محك نحصل عليها في المستقبل، و الغرض من الصدق التنبؤي هو تحديد مدى إمكانية استعمال علامات مقياس ما للتنبؤ عن علامات مقياس آخر يسمى المحك و يشرح عودة (1999) هذا التعريف بالمثال التوضيحي التالي :

معامل الارتباط بين المعدل التراكمي الجامعي، و النجاح مستقلا بمهنة التدريس مقدرا من قبل المشرفين التربويين، هو معامل الصدق التنبؤي للمعدل الجامعي.

إلا أنه لابد من توافر عدد من الخصائص في مقياس المحك هي:

- الصلة الوثيقة بالموضوع: أي أن يكون متعلقا بالوظيفة التي وضع الإختبار لقياسها.
- الخلو من التحيز: بمعنى المقياس كمحك يجب أن يهيئ لكل شخص نفس الفرصة لأخذ درجة عادلة.

أن يتوافر في المحك خاصية الثبات ملحم (2002).

3.1.2. صدق المفهوم (صدق التكوين الفرضي):

يفيد صدق المحتوى في تحديد مدى تمثيل درجات الإختبار لتحصيل الطلاب في مجال دراسي معين، أما الصدق المرتبط بمحك يفيد في تقدير أو التنبؤ بأداء حالي أو مستقبلي في مجال أكاديمي أو مهني معين، فيرى علام (2006) أنه في حالات أخرى ربما نود تفسير درجات الإختبارات في ضوء تكوينات فرضية Constructs، أو سمات نفسية Traits، أو خصائص عامة Qualities مستمدة من نظرية سيكولوجية معينة مثل: القلق، والإتجاه العلمي، والذكاء، والإبتكارية، والدفاعية و التفكير الإبداعي، وغير ذلك.

وصدق المفهوم يركز على ثلاث جوانب أساسية يذكرها عبد الفتاح الغامدي (2003) هي درجة المقياس ومدى علاقتها بالتكوين الفرضي إلى جانب تركيزه على السمات المراد قياسها، وهذا مرتبط بتفسير هذه الدرجة وأخيراً فهو يركز على ماذا يقيس المقياس من وجهة نظر القائم على إعداده أي أنها تركز على كل من التكوين الفرضي و التفسير و النظرية . و لتحقيق صدق التكوين الفرضي إقترح كرونباخ و مهيل (1955) خمسة أنواع من الأدلة ينبغي جمعها و هي كالتالي:

1. الفروق بين مجموعات من الأفراد: فعينات من الأفراد الذين يفترض اختلافهم في المتغير موضع البحث يمكن التنبؤ بتباين أدائهم.
2. التغير في الأداء و ذلك بإجراء دراسات طولية لتعرف التغيرات عبر الزمن.
3. الارتباط: أي أن تربط درجات اختبار معين بدرجات اختبار آخر بقياس المتغير نفسه.
4. الاتساق الداخلي أي الارتباط بين درجات مفردات الاختبار.
5. دراسات عملية لطريقة تناول الفرد للاختبار من خلال طرح التساؤلات التالية: ماذا يفعل الفرد عندما يطبق عليه الإختبار ؟ ما العمليات العقلية التي تتطوي عليها استجابات الفرد لمفردات الاختبار، وغيرها من الأسئلة التي ينبغي دراستها بأساليب تجريبية و إحصائية.

التكوينات الفرضية (المفاهيم) النفسية Psychological Constructs :

إن القياس النفسي عملية تقوم على أساس السلوكيات العيانية التي يمكن مشاهدتها. ومن ثم فإن الاختبار النفسي ليس أكثر من أن يكون عينة من السلوكيات. ولتحديد ما إذا كان أي اختبار يوفر قياساً جيداً عن تكوين فرضي معين يجب أن تترجم التكوين الفرضي التجريدي إلى صورة عيانية وعملية توفير وصف مفصل للعلاقة بين سلوكيات معينة. وتكوينات تجريدية يشار إليه على أنه تفسير التكوين الفرضي وهو المفتاح لصدق التكوين الفرضي لأي اختبار .

و عملية شرح التكوين تتكون من ثلاث خطوات :

- 1- حدد السلوكيات التي ترتبط بالتكوين الفرضي الذي نقيسه .
- 2- حدد التكوينات الفرضية الأخرى و قرر ما إذا كانت ترتبط أو لا ترتبط بالتكوين الذي نقيسه .
- 3- حدد السلوكيات التي ترتبط بكل من هذه التكوينات الإضافية ، و على أساس العلاقات بين التكوينات ، حدد ما إذا كان كل سلوك يرتبط بالتكوين الذي نقيسه . و كمثال لهذه العملية نأخذ اختبار أعد لقياس العدوانية بين تلاميذ المدارس .

و الخطوة الأولى في عملية شرح التكوين هي أن نصف السلوكيات ذات الصلة بالعدوانية . و الطفل الذي يؤدي الطلاب الآخرين أو الذي يتقدم إلى مقدمة الصفوف أو الذي يسيطر على معظم المباريات قد يلقب بأنه عدواني . و توجد سلوكيات أخرى يمكن اعتبارها كأمثلة لإظهار العدوانية . و في هذه المرحلة ينبغي أن نحاول أن تقدم أكبر عدد ممكن . وكلما زاد عدد السلوكيات التي نوردها كلما ازدادت الصورة وضوحا حول ما نعنيه عندما نقدم المصطلح "عدوانية".

أما الخطوة الثانية في تقديم وصف تفصيلي عن التكوين الفرضي "العدوانية" فهي أن نحدد التكوينات (المفاهيم) الأخرى التي يمكن أن تقاس في المجتمع المستهدف و تحديد ما إذا كان كل من هذه التكوينات له صلة أو ليس له صلة بالعدوانية . و على سبيل المثال ، فإنه من المفيد أن نحدد تكوينات أخرى مثل الحاجة للسلطة ، و التي لها صلة بالعدوانية . كذلك فمن المفيد تحديد التكوينات التي ليس لها صلة بالتكوين الذي نقيسه و على سبيل المثال ، إذا حددنا أن العدوانية لا ترتبط بالتكوين الفرضي "الصدق" أو الأمانة يساعدنا على تعريف حدود كل من العدوانية و الصدق و الأمانة أما الخطوة الثالثة فهي أن نحدد السلوكيات التي تتصل بكل من هذه التكوينات الإضافية – و على سبيل المثال فان الطفل الذي يتخذ القرارات دائما للمجموعة أو الذي يسيطر على المباريات قد يظهر حاجة عالية للسلطة (النفوذ) و الطفل الذي يمتنع عن الغش و يخبر الصدق لمدرسه يمكن أن نسميه أمينا . و لأننا استطعنا أن نصوغ بعض العبارات حول العلاقات بين التكوينات مثل الأمانة ، و الحاجة للسلطة ، و العدوانية ، فمن الممكن أن نقرر ما إذا كان أي من هذه السلوكيات له صلة بالعدوانية أو ليس له صلة بها . و على سبيل المثال ، إذا كانت العدوانية و الحاجة للسلطة متصلة ببعضها ، فمن الممكن أن بعض السلوكيات التي تدل على حاجة عالية للسلطة (مثلا يسيطر على المباريات) سوف تدل على مستويات عالية من العدوانية . و بالمثل إذا كانت العدوانية و الأمانة غير مرتبطتين (ليس بينها صلة) فإن علمنا بأن الطفل يمتنع عن الغش أو يقول الصدق لمدرسه لن يظهر شيئا حول مستوى عدوانيته .

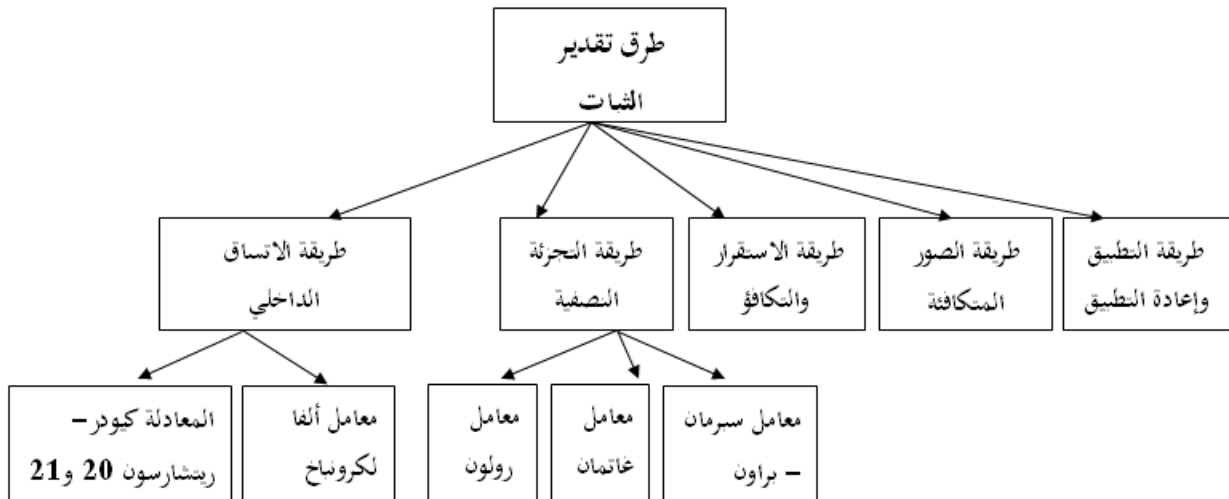
إن النتيجة النهائية لعملية تفضيل التكوين الفرضي هي وصف تفضيلي للعلاقات بين مجموعة من التكوينات الفرضية (المفاهيم النظرية) والسلوكيات. وهذه المنظومة من العلاقات والتي يشار إليها باسم شبكة معلومات Nomological Network يتيح لنا تعريفا لما نقصده بالعدوانية (كرونباخ ومهل 1955 , Cronbach & Meehl). ولأن التكوينات الفرضية ذات طبيعة تجريدية، فمن المستحيل أن تقدم تعريفا اجرائيا عيانيا لمصطلح مثل "العدوانية". وتتيح لنا شبكة المعلومات طريقة بديلة للوصف المنظم للتكوينات الفرضية . وفي مثالنا فإن العدوانية تعرف على أنها خاصية شخصية لها صلة بعدد كبير من السلوكيات (مثلا إيذاء الآخرين) ولكنها ليس لها صلة بسلوكيات أخرى (مثلا الامتناع عن الغش). ورغم أن التكوين الفرضي لا يمكن مشاهدته بشكل مباشر، فإنه يمكن الاستدلال عليه من سلوكيات مشاهدة – ويعني ذلك أننا لا نستطيع أن نقول ما هي العدوانية، و لكننا نستطيع أن نصف كيف يمكن أن يتصرف الشخص العدواني ويمكن أن نحدد عبارات ذات معنى عن مستوى العدوانية لدى الأطفال من خلال ملاحظة سلوكهم. وكلما زاد تفضيل وصف شبكة المعلومات، كلما زادت الدقة في وصف

القياس التربوي وبناء الاختبارات المدرسية

التكوين الفرضي. لكن كرونباخ، الذي وضع أسس الطريقة السابقة عام 1955، عاد ليقول عام 1988 أن تطبيقات هذه الطريقة أثبتت صعوبتها البالغة . وبذلك فقد أصبح لصدق التكوين الفرضي هدف أكثر تحديدا عما كان ينظر إليه من قبل و بدلا من إغراق كل تكوين فرضي في مجموعة كبيرة من العلاقات مع كثير من التكوينات الفرضية الأخرى ، فإن معظم الباحثين في صدق التكوين الفرضي في الوقت الحاضر يجعلون هدفهم تحديد ما هي الاستدلالات حول التكوينات الفرضية التي يمكن أو لا يمكن الوصول إليها على أساس من درجة الاختبار وما هي الظروف التي تكون هذه الاستدلالات في ظلها غير صادفه (Cronbach , 1988). (محمد إسماعيل محمد الفقي، 2005 ، ص93-96).

2.1. طرق تقدير ثبات الاختبارات:

هناك عدة طرق لتقدير ثبات الاختبارات نلخصها في الشكل التالي:
الشكل رقم (10) : طرق تقدير الثبات



1. طريقة التطبيق وإعادة التطبيق (ثبات الاستقرار):

من خلال الشكل أعلاه فإن الطريقة الأولى تقوم على تطبيق الإختبار على عيّنة من الأفراد يطلق عليها عينة التقنين، ثم يعاد تطبيق الإختبار على نفس العيّنة بعد فترة زمنية لا تقل عن

أسبوع ولا تزيد عن ستة أشهر و هذا ما يشير إليه خطاب (2000)، ثم نحسب معامل الارتباط بين الإجراء الأول والإجراء الثاني وهو ما نسميه معامل الاستقرار.

2. طريقة الصور المتكافئة (ثبات التكافؤ) :

اما الطريقة الثانية لتقدير معامل الثبات فنقوم فكرة هذا المعامل على حساب الارتباط بين علامات عينة من الأفراد على الإختبار المرغوب حساب ثباته وعلاماتهم على إختبار آخر مكافئ للإختبار الأول وقيس نفس السمة أو الخاصية ويقصد بالتكافؤ هنا حسب عودة (1999) أن للإختبارين نفس الخصائص من حيث: صعوبة السؤال وقدرته التمييزية، وتشابه المحتوى الذي يقيسه السؤال وتشابهه في إنسجام السؤال مع الهدف الذي يقيسه، وكذا تشابه شكل التوزيع للعلامات على الاختبارين، وأن يكون لهما نفس الوسط الحسابي، ونفس التباين وليس المقصود هنا التطابق في القيم ولكن المقصود أن يكون الفرق غير دال.

3. طريقة الاستقرار والتكافؤ (ثبات الاستقرار والتكافؤ)

اما طريقة الاستقرار والتكافؤ تجمع هذه الطريقة بين الطريقتين السابقتين، فمعامل الاستقرار والتكافؤ حسب علام (2006) يعكس الأخطاء العشوائية الناجمة عن إختلاف مفردات صيغتي الإختبار، وكذا إختلاف نتيجة التغيرات التي تحدث للأفراد المختبرين أو تذبذب السمة التي يقيسها الإختبار وكذلك يعد هذا المعامل أكثر المعاملات تدقيقا، وقيمه بمثابة الحد الأدنى لتقدير معامل الثبات.

ويرى ملحم (2002) أنّ بجمع بين طريقتي الاستقرار والتكافؤ فيجري إختبار ما في فترة زمنية محدّدة، ثم تجرى صورة مكافئة لهذا الإختبار في فترة زمنية أخرى أي تطبيق الصورة المكافئة لا يكون كما و هو الحال في حساب معامل التكافؤ في نفس الوقت أو بفاصل زمني قصير وإنما بعد فترة زمنية تحدد من طرف الباحث وحسب الغرض من الإختبار كما هو الحال في حساب معامل الاستقرار.

ولإيجاد القيمة التقديرية لمعامل الاستقرار والتكافؤ نستخدم نتائج الإختبارين المتكافئين، ونوجد قيمة معامل الارتباط بيرسون بين درجاتهما.

4. معاملات الإتساق الداخلي أو ثبات التجانس:

الطرق السابقة لحساب معاملات الثبات يمكن أن تناسب الإختبارات المقننة أكثر من الإختبارات التي يعدها المعلم، لأنها تحتاج إلى تطبيق الإختبار أو صورة مكافئة له أكثر من مرة، غير أنّه في بعض الأحيان يصعب تطبيق الإختبار مرتين، كما قد يصعب بناء صيغتين متكافئتين (عودة، 1999، ص348) ويمكن تقسيم طرق الإتساق الداخلي إلى:

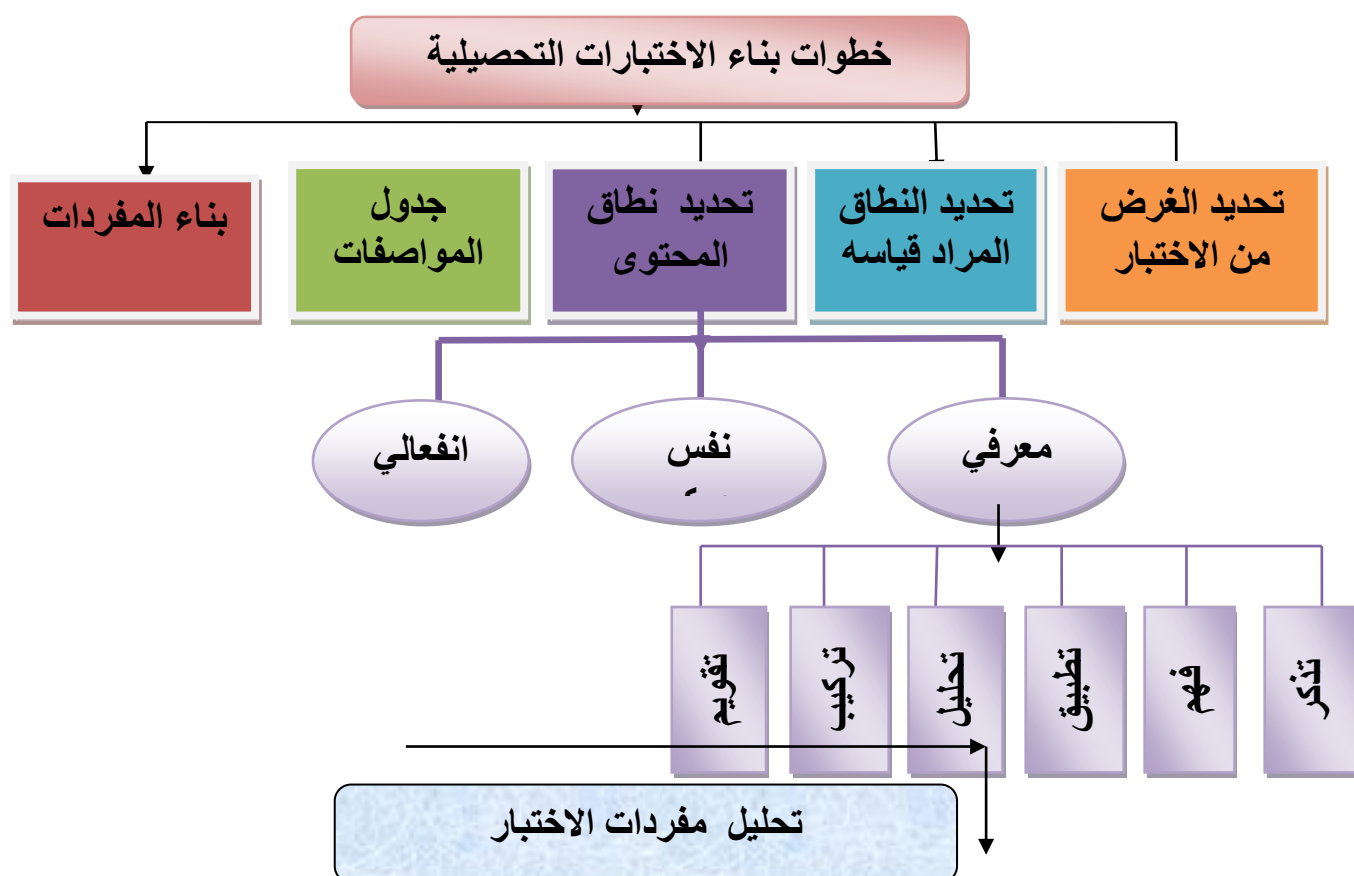
أ. طريقة التجزئة النصفية (معامل الإتساق الداخلي).

ب. طريقة كودر - ريتشاردسون (KR-20)، . (KR-21)

ج. طريقة ألفا الكرونباخ .

المحاضرة رقم 10: خطوات إعداد وبناء اختبارات تحصيلية :

الشكل رقم 11: خطوات بناء الاختبارات التحصيلية



المحاضرة رقم 11: تصنيف بلوم Bloom للأهداف التعليمية / التدريسية في المجال المعرفي .

يلخص كل من (صلاح احمد مراد ، أمين علي سليمان .(2005) ، سسيل ورونالد ترجمة صلاح الدين محمود علام (2013)، احمد محمد عبد الرحمان (2010) الاهداف التعليمية كمايلي :

طور Englehart,Furst,Hill,and Krathwohl Bloom تصنيفا للأهداف المعرفية الذي يشار اليه عادة بأنه تصنيف بلوم ، ويُقدّم هذا التصنيف طريقة مفيدة لوصف تعقد هدف معين ، وذلك بتصنيفه في واحد من ستة اقسام هرمية تتراوح بين الأكثر بساطة والأكثر تعقيدا ، وفيمايلي يأتي وصف لكل منها .

1. المعرفة / التذكر Knowledge :

المعرفة / التذكر هو ابط مستوي في التصنيف، والأهداف في هذا المستوى تتضمن تعلم أو تذكر حقائق نوعية، ومصطلحات محددة ، وأسماء وتواريخ ، وغير ذلك ، ومن أمثلة : الأهداف التعليمية في هذا المستوى :

أ- سوف يكون المتعلم قادر على تسمية مستويات القياس

سوف يكون المتعلم قادر على كتابة قائمة بأسماء رواد القياس النفسي حسب المراحل تطور القياس التي درسها

ب-معرفة /تذكر طرق ووسائل تناول الخصوصيات

وتشير هذه المعرفة إلى الطرق والوسائل التي يستخدمها الفرد في عملية جمع التفاصيل وتنظيمها . وتنقسم هذه الفئة إلى خمسة أقسام فرعية هي :

ب-1 معرفة /تذكر التقاليد الشائعة أي معرفة / تذكر طرق مميزة لمعالجة وتقديم أفكار و ظواهر معينة يجب على المتعلم إكتسابها مثل معرفة قواعد المرور ، معرفة قواعد الإعراب في اللغة .

ب-2 معرفة /تذكر التوجهات و النواتج Knowledge of trends and sequences أي معرفة عمليات ، ووجهات ، وحركات الظواهر بالنسبة للموقف مثل حركات الإصلاح الاجتماعي المختلفة ذات التتابع الزمني .

ب-3 معرفة / تذكر التصنيفات ، والفئات التي تنقسم إليها الظواهر Knowledge of classification and categories.

ب-4 معرفة / تذكر المحكات Knowledge of criteria التي تستخدم في الحكم على صحة الآراء والمبادئ و الحقائق أو صور النشاط التي تحتويها المادة الدراسية مثل المحكات الخاصة بالحكم على عمل فني (موسيقى ، رسم، باليه... الخ) .

ب-5 معرفة / تذكر مناهج البحث في ميدان معين أي بطريقة وتقنيات البحث التي يجب على المتعلم معرفتها لدى معالجته لموضوع ما مثل طريقة جمع و تبويب البيانات ، وطرق الملاحظة .

(ج) معرفة / تذكر العموميات و أهم الأفكار في مجال معين .

ويقصد بذلك معرفة / تذكر الأفكار الرئيسية التي تنتظم حولها الظواهر التي يتناولها ميدان معين و تنقسم هذه الفئة إلى :

القياس التربوي وبناء الاختبارات المدرسية

- ج-1 معرفة / تذكر الأساسيات و العموميات المتعلقة بمادة دراسية معينة و يجب على المتعلم حفظها و استرجاعها مثل تذكر مبادئ التعلم ، و مبادئ الوراثة ، و مبادئ اكتساب اللغة .
- ج-2 معرفة / تذكر التكوينات أي معرفة / تذكر النظريات و البنى الفكرية العامة وما يترتب عنها من نتائج وتعميمات .

2. الفهم Comprehension

تتطلب الاهداف في مستوى الفهم الاستيعاب – وليس فقط التذكر الصم وتستخدم الاهداف في هذا المستوى في اغلب الأحيان أفعالا مثل يفسر ، او يترجم ، او يوضح ، او يلخص ، ومن امثلة الاهداف التعليمية في هذا المستوى :

- سوف يكون المتعلم قادر على توضيح كيفية تأثير جدول المواصفات على بناء اختبار موضوعي

و يقصد به قدرة المتعلم على إدراك معنى المادة التي يدرسها ، و قدرته على إستيعابها ، و إستعادة المادة السابق دراستها بلغته . و هي أكثر المهارات العقلية شيوعا في التربية ، و يمكن الإستدلال على هذه القدرة من خلال ثلاث عمليات هي :

(أ) الترجمة Translation :

هي تحويل المعلومات من صورة إلى أخرى على نحو دقيق ، بحيث تحافظ على المادة و الأفكار التي تنطوي عليها صيغة المعلومات الأصلية ، وللترجمة عدة صور منها التحويل من أشكال Figural أو Symbol أو ألفاظ إلى صورة أخرى ، وكذلك من مستوى تجريد إلى مستوى أعلى مثل تحويل الرموز الرياضية إلى رموز خاصة بالحاسب الآلي .

(ب) التفسير Interpretation :

تعني قدرة الفرد على شرح أو تلخيص أو إعادة تنظيم الأفكار التي تشملها مادة دراسية معينة ، كتفسير قاعدة أو تفسير نظرية ، و تستلزم عملية التفسير عادة التعامل مع المحتوى كوحدة كلية من المعاني و الأفكار ، و ادراك العلاقات القائمة بينها ، و التعرف على الأفكار الرئيسية ، والتمييز بين الأفكار الرئيسية و الثانوية .

(ت) (ج) التوقعات (التنبؤات و الإستكمالات) Extrapolation :

ويقصد بها قدرة الفرد على تجاوز المعلومات المعطاة و استنتاج و تنبؤ ما قد يترتب عليها من آثار مثال التنبؤ بالآثار التي قد تنجم عن زيادة ثقب الأوزون ، أو عن الانفجار السكاني ، أو عن البطالة أو إنتشار النزاعات الدولية .

3. التطبيق Application

تتضمن الاهداف في مستوى التطبيق استخدام قواعد أو مبادئ عامة أو مفاهيم مجردة لحل مشكلة معينة لم يصادفها من قبل ، ومن امثلة الاهداف في هذا المستوى يقصد بمستوى التطبيق قدرة المتعلم على إستخدام ما تعلمه في مواقف جديدة ، و يمكن أن يشمل ذلك إستخدام المفاهيم ، و القواعد ، و القوانين ، و النظريات في مواقف جديدة . وموقف التطبيق غالبا ما يتسم بالجدية (أي عدم التكرار بالمفهوم الإحصائي) عن السياق الذي تم فيه تعلم المعلومات المرغوب في إستخدامها .

4. التحليل Analysis :

يقصد به تجزئة المحتوى إلى عناصره أو أجزائه التي يتألف منها بحيث يتضح الترتيب الهرمي (التنظيم أو مستويات البنية المعرفية) للمحتوى ، ويشمل التحليل ثلاثة جوانب هي :

(أ) تحليل العناصر Analysis of Elements

أي تجزئة المادة المتعلمة إلى عناصرها وتحديد الأجزاء الرئيسية للمادة ، و مثال ذلك التمييز بين العبارات الدالة على الحقائق و الأخرى الدالة على الآراء الشخصية ووجهات النظر .

(ب) تحليل العلاقات Analysis of Relationships

تحليل العلاقات بين الأجزاء الرئيسية المكونة للمادة كالعلاقة بين الفروض و الأدلة أو الفروض و النتائج أو بين مختلف أنواع الأدلة .

(ج) تحليل مبادئ التنظيم Analysis of Organizationl Principales

ويشمل تحليل الأسس، والحجج ، ووجهات النظر ، والقواعد والمفاهيم التي تجعل من المادة بنية كلية منتظمة أي وحدة واحدة .

5. التركيب Synthesis :

(أ) يقصد به وضع العناصر و الأجزاء معا بحيث تكون نمطا Pattern أو تركيبا جديدا لم يكن موجودا من قبل ، و يؤكد هذا المستوى على الإنتاج الابتكاري لأداء المتعلم و الذي يظهر في ثلاثة أنواع هي :

(ب) إنتاج محتوى فريد Production of Unique Communication

و يعني إنتاج الأفكار و نقلها (توصيلها) إلى الآخرين مثال ذلك كتابة قصة قصيرة ، تأليف قطعة شعرية أو قطعة موسيقية ... الخ .

(ت) إنتاج خطة أو مشروع مقترح : Production of plan , or Proposal

و يقصد به إنتاج خطة أو مشروع يحقق شروط تنفيذ عمل أو واجب معين مثال ، ذلك : إعداد مشروع بحث ، و إعداد خطة لإدارة مؤسسة .

(ج) اشتقاق العلاقات المجردة : Derivation of Set of Abstract Relations

هي عملية استنباط العلاقات من مجموعة من القضايا الأساسية و التي تساعد على تفسير البيانات أو الظواهر أو تؤدي إلى حلول جديدة للمشكلات ، و مثال ذلك القدرة على تكوين فروض مناسبة مبنية على تحليل العوامل .

6. التقويم Evaluation :

هو عملية عقلية يقصد بها قدرة الفرد على إصدار حكم كمي (رقمي) أو كيفي أو كليهما معا على قيمة المحتوى الذي يدرسه لتحديد مدى ملاءمة المحتوى لمحكات معينة بغرض التحسين و التعديل و التطوير . و تنقسم هذه الفئة إلى :

(أ) الحكم في ضوء معايير داخلية :

و يقصد به تقويم مدى دقة المحتوى في ضوء شواهد داخلية مثل : الدقة المنطقية ، و الترتيب المنطقي للنص ، و الاتساق الداخلي كالكشف الأخطاء المنطقية في البراهين أو

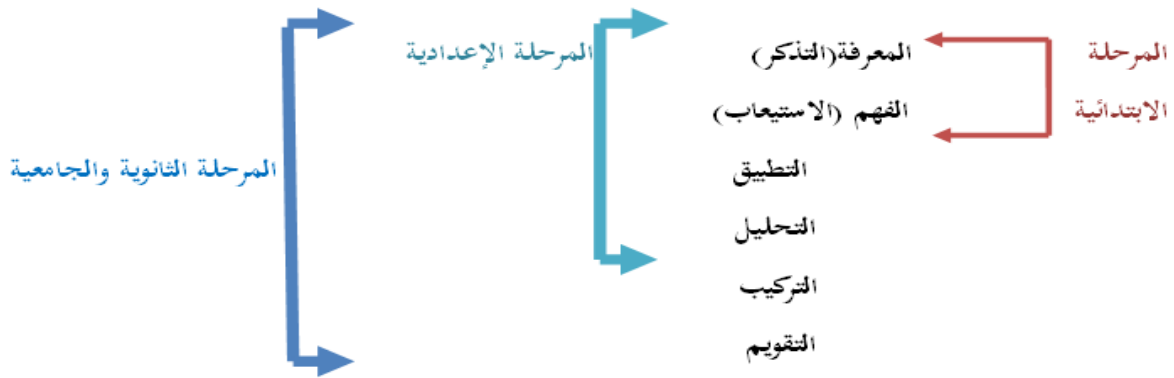
القياس التربوي وبناء الاختبارات المدرسية

استخدام المفهوم الواحد في موضوع واحد بأكثر من معنى ، و غير ذلك من المحكات الداخلية .

(ب) الحكم في ضوء معايير خارجية :

و يقصد به تقويم المحتوى في ضوء محكات خارجية محددة مسبقا و متفق عليها مثل الحكم على دور العقاب البدني في التربية المعاصرة في ضوء المعايير الأخلاقية السائدة في المجتمع . أو نقد عمل مسرحي أو روائي سائد بالمقارنة بينه و بين غيره من الأعمال المماثلة .

- يمكن تقديم تصور مقترح يبين العلاقة بين مستويات الأهداف التدريسية التعليمية الضمنية كما اقترحها بلوم و المراحل التعليمية كما يلي حسب (صلاح احمد مراد، أمين علي سليمان ،2005، ص 109-110) :



الشكل التالي يوضح الترتيب الهرمي لمستويات المجال المعرفي حسب تصنيف بلوم

القياس التربوي وبناء الاختبارات المدرسية



(صلاح احمد مراد، أمين علي سليمان. 2005، سسيل ورونالد ترجمة صلاح الدين محمود
علام، 2013)