



جامعة الجيلالي بونعامة-خميس مليانة -

2020-2019 السداسي الرابع

اسم المادة
القياس التربوي وبناء الاختبارات المدرسية

الأستاذ المحاضر			
الاسم واللقب	الرتبة	الكلية	البريد الالكتروني
لعزالي صليحة	MCA	العلوم الإنسانية والاجتماعية	saliha.lazali@univ-dbkm.dz

الطلبة المعنيين			
الكلية	القسم	السنة	التخصص
العلوم الإنسانية والاجتماعية	العلوم الاجتماعية	الثانية	علوم التربية : ارشاد وتوجيه

الأهداف التعليمية المرجو تحقيقها :

عزيزي الطالب بعد قراءة تلك لمحاضرات هذا المقياس ينبغي ان تكون قادر على :

1. التعرف على ماهية الاختبارات ، أنواعها واستخداماتها .
2. وصف الخطوات العملية التي يجب إتباعها لبناء الاختبارات سواء كانت نفسية او تربوية . .
3. إدراك أهمية التكوين الفرضي في بناء الاختبارات.
4. امتلاك القدرة لتحويل الأهداف العامة إلى أهداف إجرائية
5. التعرف جدول مواصفات، ماهو ؟ ما مكوناته ؟ والمعايير التي ينبغي توفرها اثناء بناءه ؟ وكذا متى وأين يوظف؟
6. القدرة على بناء اختبار تحصيلي موضوعي.
7. التعرف على اهم طرق التحقق من صلاحية أدوات القياس .

القياس التربوي وبناء الاختبارات المدرسية
المستوى: السنة الثانية
السداسي: الرابع
التخصص: علوم التربية
وحدة التعليم: المنهجية
المعامل: 02

مخطط المقياس: بعد ان تم التطرق في السداسي الأول الى القياس النفسي والتربوي، تعريفه ، مكوناته ، التطور التاريخي له واهم رواده وعلاقته بالعلوم الاخرى، ومسلماته ، ومستويات القياس ، سيكون في هذا السداسي مقياس بناء الاختبارات المدرسية الذي سوف نتطرق اليه معا وفق المراحل التالية.

المرحلة الأولى: التعرف على ماهية الاختبارات واستخداماتها.

المرحلة الثانية: التعرف على الخطوات الأساسية والضرورية لبناء اختبار في سياق عام ينطبق على معظم أنواع الاختبارات (النفسية والتربوية والمهنية)

المرحلة الثالثة: التعرف على جدول المواصفات وبناء اختبار تحصيلي موضوعي.

المرحلة الرابعة: الخصائص السيكو مترية بطريقة تفصيلية للتحقق من صلاحية أداة القياس التي تم بنائها.

المحاضرة رقم 01: ماهية الاختبارات، واستخداماتها.

1. تعرف الاختبار:

الاختبار النفسي بالضرورة هو مقياس موضوعي ومقنن 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 باجراءتهم 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 اختبار مقدار المفردات اللغوية لدى الطفل ، أو قدرة كاتب على أداء العمليات الحسابية أو تأزر 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

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 : تتنبأ بمدى قدرة الفرد على التعليم أو التدريب على مهنة معينة كالقدرة الميكانيكية و الموسيقية و الحسابة . 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 ل كودر .

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 ليكورت ، مقياس آلبرت 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 إحصائين مدرسين 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 تفكيراً ضمنياً 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 (02): 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 إنشائه 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 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 .

ومن شروط الهدف ان يتسم بالوضوح ، الشمولية ، القابلية للقياس ، ويمكن تقسيم الهدف الى ثلاثة مستويات ، وذلك حسب الغرض الذي يسعى الباحث اليه ، أثناء بناءه للمقياس ، فهل الهدف قياس التحصيل ؟ ام هدفه قياس السمات الشخصية ؟ ام هدفه قياس قدرة او كفاءة مهنية (تربوي ، نفسي، كفاءة مهنية).

إذا يمكن تلخيص أنواع او مستويات الأهداف في :

أ. تحديد المحتوى والمهارات المراد قياسها: (خاص بقياس التحصيل)

ان تحليل المنهج الدراسي يعرفنا بأهداف التربية وما يرمي الى تحقيقه، وهنا يقوم المُعد بتصنيف جميع المواد التي يشتملها المنهج ، بعدها يحلل كل مادة الى موضوعاتها الرئيسية والفرعية ثم يحدد أهمية كل موضوع بالنسبة للمادة او أهمية كل مادة بالنسبة للمنهج ثم يختار نوع الاختبار والفقرات (مقالية ، متعددة الاختيار ، التكملة ...).

ب. تحليل العمل (كفاءة مهنية):

يساعد تحليل العمل على معرفة القدرات والاستعدادات والسمات التي يجب ان تتوفر في الشخص للقيام بهذا العمل و تحليل العمل في خطواته النهائية يصل بيانات تلخص في قائمة تسمى "وصف العمل" ، تحدد العمليات المتضمنة في هذا العمل وطبيعتها وأهمية كل منها وكيف تؤدي ، بعد "وصف العمل" يتم تحديد "مواصفات العامل" أو القدرات التي يتطلبها العمل .

ج. تحليل السمات (من خلال التراث النظري):

يتم تحديد خصائص وسمات الهدف المراد قياسه بطريقة إجرائية فتخطيط الاختبار يجب ان يغطي معظم الخصائص للمفهوم، ويصلح هذا مع المستوى مع الاختبارات الشخصية التي تتكون من العديد من العوامل، والعوامل تحتوي على مجموعة من المؤشرات.

الخطوة الثالثة : تحديد اوزن الفقرات او المفردات

يساعد تحليل العمل او تحليل المنهج الدراسي او تحديد السمة ، على تحديد الاوزان المختلفة (عدد الفقرات اللازمة لكل مادة دراسية او لكل مهنة ما ، او لكل بعد من ابعاد خاصية ما) ، ان تحديد اوزان الفقرات يتطلب تصنيف المهام وترتيبها حسب أهميتها ، فكلما كان الموضوع مهم زاد عدد الفقرات او زاد وزنها وكلما كان قليل الأهمية قل عددها ، وهناك طريقة لتحديد اوزان الفقرات في الاختبارات التحصيلية تسمى جدول "التحديد ذي اتجاهين" ل بلوم Bloom، نسجل في العمودين الأول والثاني محتوى الموضوعات (العام ثم الخاص) وعلى الصف الافقي نسجل الأهداف التعليمية التي تتضمن عادة واحدة او اكثر من الأهداف التالية (الاسترجاع ، التعرف ، الفهم ، التطبيق ، التحليل ، التركيب ، التفسير ، التعميم ، التقويم) ، اما الخلايا الموجودة

في الوسط فهي لتحديد عدد الفقرات الضرورية مصنفة ، حسب المحتوى والهدف . كما هو موضح في المثال التالي :

جدول مواصفات

اسم المادة (المقياس): الإحصاء الوصفي
المستوى : السنة الأولى جذع مشترك
ملاحظة : عدد المفردات يحدد وفق للمحتوى و مستوى الأهداف – وسيتم في المحاضرة الثالثة التفصيل في جدول مواصفات وطريقة حساب الأهداف النسبية و اوزان الفقرات والتنقيط .

مجموع الأهداف	الأهداف التعليمية						عدد الحصص	المحتوى الخاص (موضوعات كل الوحدة)	المحتوى العام (الوحدة)
	التقويم	التركيب	التحليل	التطبيق	الفهم	التذكر			
10	/	/	2	4	2	2	2 حصص	-اسمي -رتبي - فئوي - نسبي	مستويات القياس
09	/	/	/	3	3	3	02 حصص	- المتوسط . - الوسيط - المنوال	مقاييس النزعة المركزية
09	/	/	/	3	3	3	02 حصص	- المدى - الانحراف المعياري - التباين	مقاييس التشتت
09	/	/	/	3	3	3	03 حصص	- الرتبة المئينية - معامل الاختلاف - الربيعيات	مقاييس المكانة النسبية
09	/	2	2		3	3	03 حصص	- معامل الارتباط بيرسون - معامل الارتباط سيرمان - معامل فاي	الارتباط

الخطوة الرابعة : الصورة الأولية للمقياس (اعداد الفقرات) (بناء المفردات الاختبارية)

بعد تحديد عدد الفقرات يشرع الباحث في اعداد وكتابة الفقرات ويجب عليه أن يكون ملما بالاشكال المختلفة التي تكتب بها صيغة الفقرة ، ويوجد عدة أشكال نختصرها فيمايلي :

1. اختيار إجابة من اجبايتين (صح/خطأ) ، (نعم ، لا).

2. اختيار متعدد.

3. المزوجة.

4. التكملة.

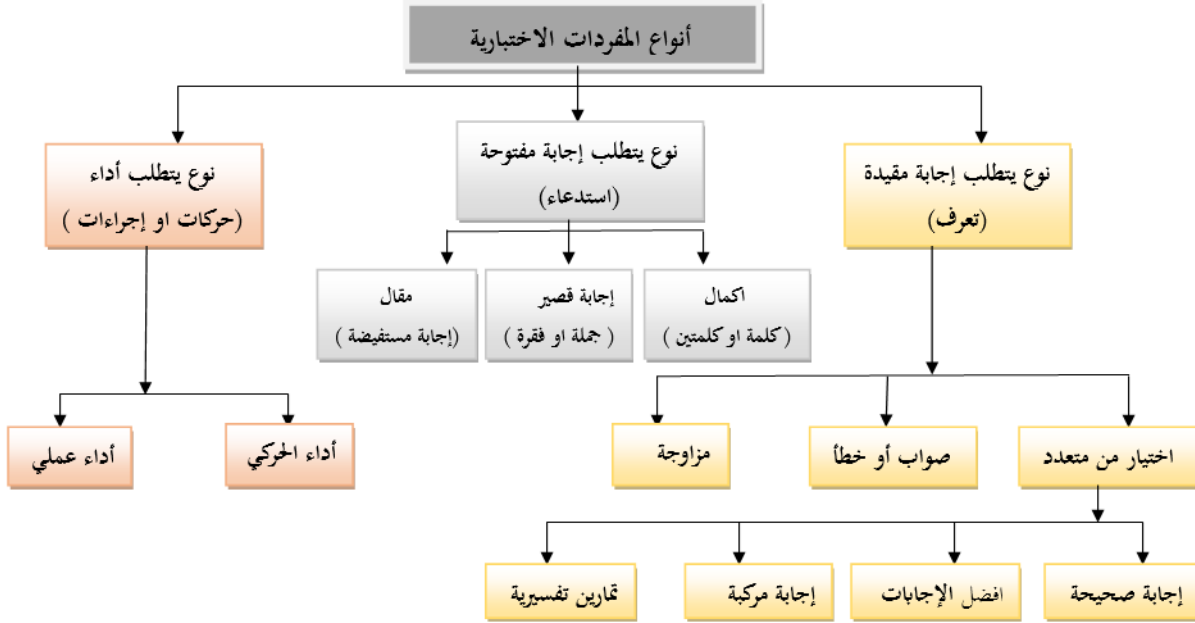
5. الاستدعاء البسيط.

6. التناظر الوظيفي.

7. إعادة الترتيب.

8. إختبار المقال.

ويضع (علام ، 1995 ، ص 64) شكلا تخطيطيا يوضح فيه أنواع المفردات الاختبارية التي يمكن استخدامها في قياس الاهداف السلوكية بمختلف مستوياتها



صياغة فقرات الاستبانات : Item formats for Inventories :

بعد التعرف على انواع الفقرات لا بد نتطرق الى اهم الصيغ الشائعة اثناء بناء فقرات الاستبانات ، هنالك صيغ ثلاثة شائعة لاستبانات الشخصية و الاتجاهات هي صيغة موافق -غير موافق ، صيغة ليكرت . الصيغة الوصفية ثنائية الأقطاب ، و فيما يأتي وصفا لكل منها :

تتألف صيغة موافق -غير موافق على الأغلب من جملة تقريرية متبوعة باختيار واحد من الاستجابتين . وهذه الصيغة موضحة بالفقرات الآتية :

1/ يجب أن يطيع الأطفال الوالدين دون أي سؤال
موافق موافق غير موافق

2/ يحتاج أطفال اليوم لضبط أقوى في البيت
موافق موافق غير موافق

3/ يثور معظم الأطفال الذين لديهم مشكلات سلوكية على الضبط الوالدي
موافق موافق غير موافق

والطريقة الأسهل لإعطاء الدرجات لمثل هذه الفقرات هو تقرير أي طرفي متصل الاتجاه سيعطى الدرجة العالية و أيها سيعطى الدرجة المنخفضة . فمثلا للفقرات المذكورة أعلاه ، يمكن وصف الاتجاه على متصل يتراوح بين

الإخضاع الكامل و التساهل من وجهة نظر والدية . و هنا يتوفر لمطور الاختبار فرصة لتخصيص المستجيب الأكثر سلطوية على أنه الذي يحصل على درجة أعلى على الأداء أو أن المستجيب الأكثر تساهلا يحصل على الدرجة الأعلى . دعنا نفترض أن مطور الاختبار يرغب أن تعكس الدرجات الأعلى للاتجاه الأكثر تساهلا . بعدها يتم تحديد الدرجة التي تستحقها كل فقرة (إيجابية أم سلبية) بالنسبة للبناء . و في مثالنا لأن المتصل معرف بلغة التساهل فالمتصل معرف بلغة التساهل فإن الفقرتين 1 ، 2 تعد سلبية و الفقرة 3 تعد إيجابية ، و تصحيح الفقرات بإعطاء درجة واحدة لكل استجابة بالموافقة للفقرة الإيجابية . و نقطة واحدة لكل استجابة بعدم الموافقة للفقرة السلبية ، و علامة المستجيب الكلية هي مجموع درجاته على الفقرات جميعها . و في المثال الحالي المستجيب الذي أشر غير موافق ، غير موافق ، موافق للفقرات (1 ، 2 ، 3) على التوالي فإنه سيحصل على الدرجات (1،1،1) و مجموعه 3 درجات . و مستجيب آخر أشر على غير موافق ، موافق ، غير موافق يكون نمط درجاته (1،0،0) و مجموعه درجة واحدة .

و طريقة بديلة لتصحيح فقرات صيغة موافق - غير موافق تستخدم أوزان الفقرات . فكل فقرة يؤشر لها بقيمة موزونة بحسب قوة التعبير الوجداني - الانفعالي - نحو البناء الذي تقيسه ، كذلك فإن هناك طرائق عدة تستخدم في الحصول على أوزان الفقرات ، و الطريقة الأكثر شيوعا هي طريقة الفترات المتساوية - القياس الفاصلي - الذي وصفه ثيرستون (Thurstone ; 1928) . و عادة يعد كاتب الفقرات عددا كبيرا من الفقرات (اقترح ثيرستون حوالي 100 فقرة) تتراوح بين موجب متطرف إلى سالب متطرف بالنسبة للبناء المقيس ، بحيث يكون بعضها ذو تأثير متعادل . و قد تكتب كل فقرة على بطاقة منفصلة ، و تعرض مجموعة الفقرات على مجموعة من المحكمين لتدريبها على متصل الفترات المتساوية ، و يطلب من كل محكم قراءة الفقرة و وضعها على متصل مجزأ إلى 7 أو 9 أو 11 فترة متساوية في المسافات الفاصلة بينها . و على المتصل المؤلف من 11 فترة تقع الفقرة المتعادلة في الفئة أو الفترة السادسة.

و تحقق البيانات المستخلصة من عملية التحكيم هذه هدفين - الأول اختيار فقرات الصيغة النهائية للمقياس و تأشير الأوزان لكل منها ، ووزن الفقرة هو وسيط تقديرات المحكمين . و علاوة على ذلك تختار فقرات المقياس من المجموعة الكلية للفقرات على أساس إحصائي مثل Q Semi- interquartile range :

$$Q = \frac{75م - 25م}{2}$$

حيث أن م75 هو القيمة العددية للرتبة المئينية 75

م25 هو القيمة العددية للرتبة المئينية 25 لمدى التوزيع .

وكلما كانت قيمة Q أقل كلما كانت الأحكام متفقة بدرجة أكبر في قوة الانفعال المعبر عنه بالفقرة ، و هذه الفقرات هي الأفضل في تكوين المقياس . و مع ذلك فإن مطور الاختبار يحاول تضمين المقياس فقرات من كل فئة تصنيف . و عندما يطبق المقياس على المستجيبين فيما بعد (و في كل تطبيق) يقر كل مستجيب فقرة تنطبق عليه أكثر من غيرها ، و يضاف وزن هذه الفقرة إلى درجته الكلية .

و الدرجة الكلية تقسم فيما بعد على عدد الفقرات التي أقرها المفحوص من أجل الحصول على متوسط وزن الفقرة . و يستخدم هذا المتوسط لحساب أو وصف إتجاهات المفحوصين أو المستجيبين .

و الصيغة الثانية الأكثر شيوعا للاستبانات اقترحها ليكارت (Likert , 1932) ، و تتطلب طريقته كتابة مجموعة من الفقرات تبين كل منها إتجاه إيجابي أو سلبي بالنسبة للبناء المقيس . و هنا لا يتضمن المقياس فقرات محايدة ، و عندما يقرأ المستجيب الجملة يختار استجابة على متصل مؤلف من خمس نقاط تتراوح بين موافق بشدة و غير موافق بشدة . و كما هو مبين على النحو الآتي :

يجب أن يطيع الأطفال والديهم بدون استفسار (دون معرفة الأسباب)

غير موافق بشدة غير موافق محايد موافق موافق بشدة

و الفقرات هنا لا يتم تدرجها قبل تطبيقها على المستجيبين ، و لكن يستخدم المستجيب متصل الإجابة (1-5) ليؤشر درجة اقراره للموقف المذكور في الفقرة . و لإعطاء الدرجات على متصل الإجابة المتدرج تعطى الدرجة (1) للاستجابة التي تظهر أقل مستوى من التعبير الانفعالي ، و الدرجة (2) للاستجابة التي تليها في درجة التعبير الانفعالي، و الدرجة (3) للاستجابة التي تليها في المستوى أو الشدة و هكذا . و إذا افترضنا ثانية أن مطور الاختبار يريد إعطاء درجات أعلى للمستجيب الأقل انفعالا أو أكثر تساهلا ، فإن الفقرات الثلاث المذكورة أعلاه تكون درجاتها على النحو الآتي :

الفقرة موافق موافق محايد غير موافق غير موافق بشدة

1	2	3	4	5	1/ يجب أن يطيع الأطفال والديهم دون استفسار
1	2	3	4	5	2/ يحتاج أطفال اليوم ضبط أقوى في البيت
5	4	3	2	1	3/ يثور معظم الأطفال الذين لديهم مشكلات

- وعلاوة المستجيب الكلية هي مجموعة درجات الفقرات جميعها . و فيما يأتي بعض الإرشادات العامة التي تفيد في كتابة و مراجعة فقرات استبانة ليكرت أو موافق – غير موافق :
- 1/ اكتب الجمل أو الفقرات في صيغة الحاضر .
 - 2/ لا تستخدم جمل واقعية أو تلك التي تكون قابلة للتفسير الواقعي .
 - 3/ تجنب الجمل التي تحتمل أكثر من تفسير .
 - 4/ تجنب الجمل التي يجيب عنها الجميع ب غالبا (أو موافق بشدة) أو نادرا (أو غير موافق بشدة).
 - 5/ حاول أن تكون عدد الفقرات التي تعبر عن انفعالي إيجابي مساويا لعدد الفقرات التي تعبر عن انفعال سلبي .
 - 6/ أن تكون الجمل قصيرة و لا تتجاوز 20 كلمة .
 - 7/ أن تكون الجمل صحيحة من حيث قواعد اللغة .
 - 8/ يجب تجنب الجمل التي تتضمن كليات مثل لا شيء ، و أبدا لأنها تعطي الغموض .
 - 9/ تجنب استخدام متطلبات غير محددة (indefinite qualifiers) مثل فقط ، دائما ، ليس غير ، عديد ، قليل ، نادرا .
 - 10/ كلما أمكن ، حاول أن تكون الفقرات بسيطة ، و ليست جمل معقدة أو مركبة ، و تجنب الجمل الشرطية التي تستخدم (إذا) أو (لأن) .
 - 11/ استخدم الألفاظ التي تفهم بسهولة من قبل المستجيبين .
 - 12/ تجنب استخدام علامات النفي مثل (لا ، و لا واحدة ، أبدا) .
- و يمكن الرجوع إلى مصادر حديثة تفيد في بناء فقرات مقاييس الاتجاه مثل اندرسون (Anderson , 1981) ، و داويس (Dawes, 1972) و يودنسكي و زملاؤه (Undinsky , et.al , 1981) .
- و هناك صيغة أخرى شائعة في بناء الاستبانات هي استخدام أزواج الصفات القطبية . و يعود أصل هذه الصيغة إلى أوزجود و زملاؤه (Osgood , et.al , 1957) إذ اقترحوا استخدامها في دراسة المعاني اللغوية للأبنية النفسية، إذ يكتب في الأعلى اسم البناء النفسي المراد قياسه و يتبع بزواج من الصفات التي تمثل قطبين متعاكسين على المتصل نفسه ، و يفصل بينهما إما (5 أو 7) نقاط على المتصل ، و يوجه المفحوص ليضع علامة على المتصل في المكان الذي يعكس مشاعره . كمثال في محاولة تقويم اتجاهات المعلمين في العمل مع الطلبة المعاقين عقليا ، يمكن استخدام الصيغة الآتية :

جميل _____ قبيح

سعيد _____ حزين

قذر _____ نظيف

و بعد أن اختبر أوزجود و زملاؤه عددا كبيرا من أزواج الصفات المتنوعة وجدوا أن معظمها يتجمع ضمن واحد من الأبعاد الثلاثة الآتية : التقويم ، الفعالية ، و النشاط و التي تقع ضمن المعنى اللغوي للبناء اللفظي . و قد اقترحوا طريقة في تحليل استجابات هذه الفقرات و تفسيرها ، و أطلقوا على أدواتهم هذه اسم التفاضل اللغوي . و على مطور الاختبار الذي يعد مجموعة متميزة من الصفات ثنائية القطب الجديدة (أي دون استخدام الأزواج المختبرة من قبل أن يحدد كيفية إعطاء الدرجات و تفسيرها) . و في مثل هذه المواقف يجب وصف الفقرات في قائمة صفات ثنائية القطب بدلا من مقياس التفاضل اللغوي .

و من المهم عند تطوير فقرات المقياس الانتباه إلى مجموعة الاستجابات التي تؤثر على سلوك المفحوصين منفردين . و يمكن تحديد مجموعة الاستجابات على أنها نيل المفحوصين لأن يستجيب بطريقة معينة لصيغة معينة من الفقرات بغض النظر عن محتواها .

و قد قدم جيلفورد (Guilford ; 1954) تحديدا تقليديا لأنواع عديدة من الاستجابات المألوفة ، و قدم مقترحات لخفض أو ضبط أثر هذه التأثيرات . و مجموعتين من الاستجابات تؤثر في الدرجة على المقياس هي الميل للموافقة على جملة بغض النظر عن محتواها و التفسيرات التفاضلية للأفراد على الأشياء و المتطلبات غير المحددة مثل بعض، عادة .

و هناك قضايا أخرى تتعلق ببناء الاستبانات تتضمن أثر استخدام أنواع مختلفة من النقاط الرابطة (anchor points) على متصل الإجابة انظر (Frisbie & Branden bug , 1979) و (Lam & Klockers , 1982) أو عدد مختلف من بدائل الاستجابة (Masters , 1974 Velicer & Stevenson , 1978)

و مثل هذه الدراسات توضح أن القرارات حول كيفية قياس السلوك يمكن أن تثبت بصورة حتمية الشيء المقيس .

الخطوة الخامسة : مراجعة مفردات المقياس والتأكد من فعالية نوع المفردة وملائمتها للفئة المستهدفة ، فعملية المراجعة تؤدي إلى تحسين مفردات المقياس من حيث المعنى والفهم والغاية وبالنسبة لمقاييس الشخصية فإن معد الاختبار يجب عليه الاطلاع على المعارف المتنوعة، والنظريات المختلفة للشخصية ، والمفردة الجيدة يجب أن تمر بمراحل عديدة فهي تبدأ بالكتابة والتهديب والاختبار والمراجعة وهذه الخطوات مطلوبة حتى يتسنى لنا وجود مفردة جيدة ، تكون مفهومة ولها معنى واحد، يمكن كتابة العديد من المفردات وتخزينها في بنك الأسئلة ، لأنها الرصيد الذي يحتاجه المعلم أو المختص في الأسئلة حتى يسهل عليه تقديمها للمفحوصين .

إذا يجب توفر الشروط التالية في المفردة: البساطة، الوضوح، التسلسل، المنطقي، الاستقلالية، عدم التركيب، الأيجاز، الدقة .

الخطوة السادسة: مفتاح التصحيح أو سلم التنقيط من العوامل التي تؤثر في الاختبار طريقة تصحيحه ، و وزن كل فقرة من مفرداته ، ونوع المفتاح .

وتعتبر هذه الطريقة إحدى طرق استخراج مفتاح التصحيح، وهي تعتمد على المتوسط الفرضي أو المتوسط النظري

وهي قيمة تعادل (50%) من الدرجة الكلية للمقياس المراد بناءه إذا كان من نوع التقرير الذاتي، وتحسب من خلال القانون الآتي:

$$\text{المتوسط الفرضي} = \frac{\text{مجموع عدد البدائل} \times \text{عدد فقرات المقياس}}{\text{عدد البدائل}}$$

أو بطريقة أخرى

$$\text{المتوسط الفرضي} = \frac{\text{أقل درجة في المقياس} + \text{أعلى درجة في المقياس}}{2}$$

مثال: إذا افترضنا ان لدينا مقياس يتكون من :

20 بند / ولديه 5 بدائل / البديل الأول يعطي الرقم واحد أو الدرجة واحد والبديل الأخير يعطي الدرجة خمسة فانه إذا طبقنا القانون الأول كالتالي:

$$\text{المتوسط الفرضي} = 20 \times \frac{5+4+3+2+1}{5} = 60$$

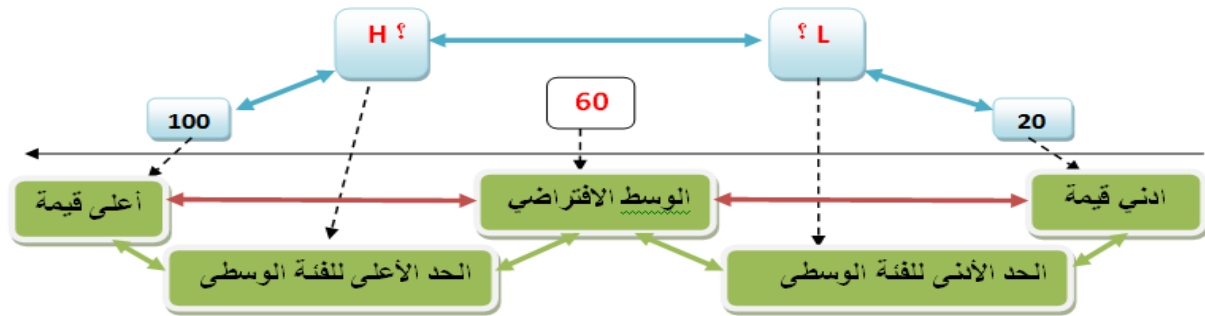
الطريقة الثانية: أدنى درجة هي $20=1 \times 20$ أما أعلى درجة فهي $100=5 \times 20$

$$100+20$$

$$\text{المتوسط الفرضي} = \frac{\quad}{2} = 60$$

إذا نلاحظ نفس القيمة باستخدام الطريقتين، هذا إذا ما أردنا استخراج مفتاح تصحيح ثنائي باستخدام المتوسط الافتراضي فيصبح المفتاح كالاتي : من 20 الى 60 لا يمتلك الخاصية / من 60 الى 100 يمتلك الخاصية

أما إذا أردناه ثلاثي أو رباعي فإننا نستخدم نفس الطريقة إلا أننا يجب ان نبحث عن حدود الفئة الوسطى وفق الطريقة الآتية:



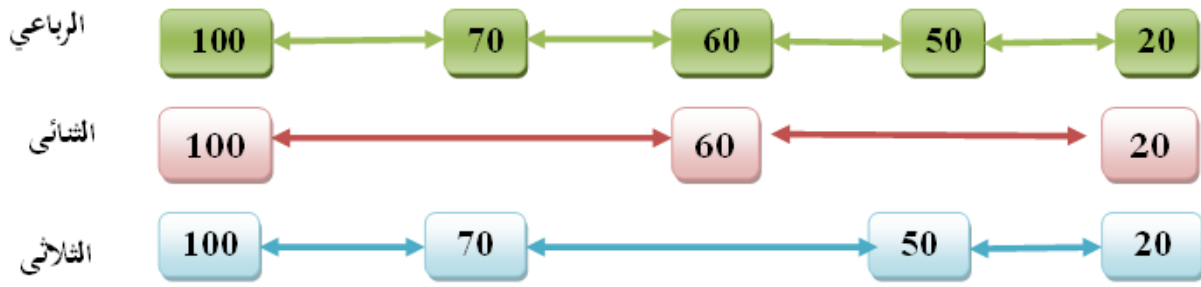
$$\text{الحد الأدنى للفئة الوسطى} = \text{المتوسط الفرضي} - \frac{\text{عدد البنود}}{2}$$

$$50 = \frac{20}{2} - 60 = \text{الحد الأدنى للفئة الوسطى}$$

$$\text{الحد الأعلى للفئة الوسطى} = \text{المتوسط الفرضي} + \frac{\text{عدد البنود}}{2}$$

$$70 = \frac{20}{2} + 60 = \text{الحد الأعلى للفئة الوسطى}$$

وعليه يصبح مفتاح التصحيح كالاتي :



الخطوة السابعة: التعليم

إعداد تعليمية المقياس بهدف تسهيل مهمة المبحوث وتتضمن مايلي : مقدمة استهلاكية يوضح فيها معد كيفية الإجابة على مقياس + تقديم مثال على ذلك ، يشير فيها أنه لا توجد إجابة صحيحة واجابة خاطئة ، وبطمئن المبحوث بأنه سيحافظ على سرية الإجابة .(الاستهلال + مثال + التطمين)

الخطوة الثامنة: تجريب الصورة الأولية، بهدف التحقق من الخصائص السيكومترية. تتأتى هذه الخطوة بعد عرضه على محكمين (يجب ان تتوفر في المحكمين بعض الشروط نذكر منها : ان يكون في التخصص + ان يمتلك خبرة في مجال الخاصية المراد قياسها)

الخطوة التاسعة: الصورة النهائية.

الخطوة العاشرة : تحديد زمن الاختبار ، ويتم اما عن طريق خبرة الباحث او حساب متوسط زمن المختبرين .

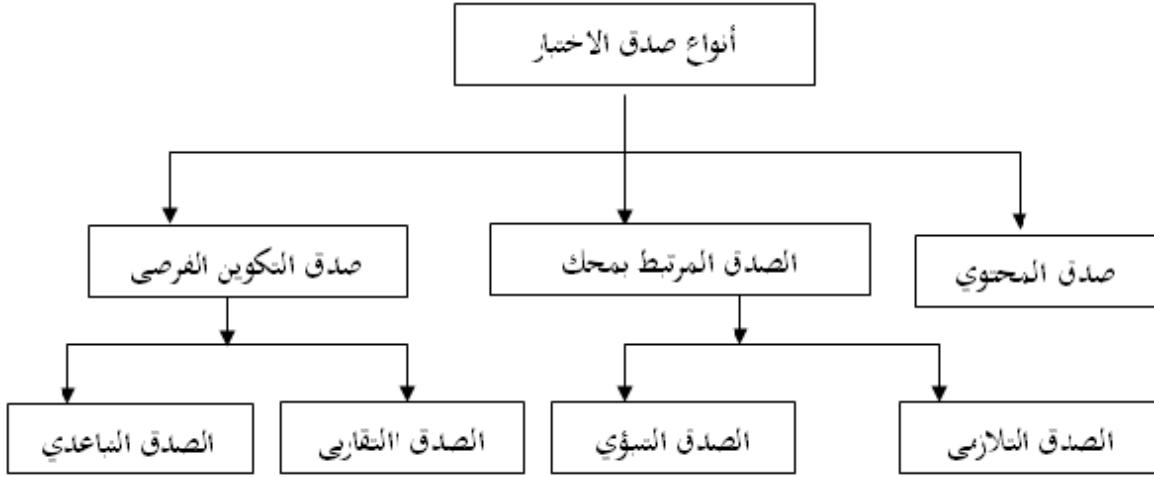
المحاضرة رقم (03): الخصائص السيكومترية للأداة القياس

من خلال الخطوات بناء الاختبارات التي تم التطرق إليها في المحاضرة رقم (01) تعتبر الخطوة الثامنة ، وهي خطوة التحقق من صلاحية الأداة المبنية قبل إخراجها بشكل نهائي من اهم الخطوات ، ويتم ذلك من خلال حساب الخصائص السيكومترية والتي سنذكر أنواعها فيما يلي :

1. الخصائص السيكومترية :

2. 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(01): يوضح أنواع صدق الاختبارات

الجدير بالذكر أن اللجنة أكدت هذه الأنواع الثلاثة مترابطة إجرائيا و منطقيًا، فنادرًا 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.

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 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.

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 بداية من إختيار الفقرات الملائمة، بالنسبة للاختبارات التحصيلية مثلا: فان 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 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" .

يشير 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 التمييز بين نوعين من هذا الصدق:الصدق التلازمي Concurrent Validity و الصدق التنبؤي Predictive Validity، و الفرق الرئيسي بينهما 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 نبحت 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، 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 الإرتباط 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 إستعمال علامات مقياس ما للتنبؤ عن علامات مقياس آخر يسمى المحك و يشرح هذا التعريف بالمثل 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 لا بد من توافر عدد من الخصائص في مقياس المحك هي:

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

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، فإنه 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، والإنجاء العلمي، والذكاء، والإبتكارية، والدفاعية و التفكير الإبداعي، وغير ذلك.

وصدق المفهوم يركز على ثلاث جوانب أساسية هي درجة المقياس ومدى علاقتها بالتكوين الفرضي إلى جانب تركيزه على السمات المراد قياسها، وهذا مرتبط بتفسير هذه الدرجة و أخيرا 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 إقتراح 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 الإرتباط 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 تنطوي عليها استجابات الفرد لمفردات الاختبار، وغيرها من الأسئلة التي ينبغي دراستها بأساليب تجريبية و إحصائية.

التكوينات الفرضية (المفاهيم) النفسية Psychological Constructs :

أن القياس النفسي عملية تقوم على أساس السلوكيات العيانية التي يمكن مشاهدتها . و من ثم فإن الاختبار النفسي ليس أكثر من أن يكون عينة من السلوكيات . و لتحديد ما إذا كان أي اختبار يوفر قياسا جيدا عن تكوين

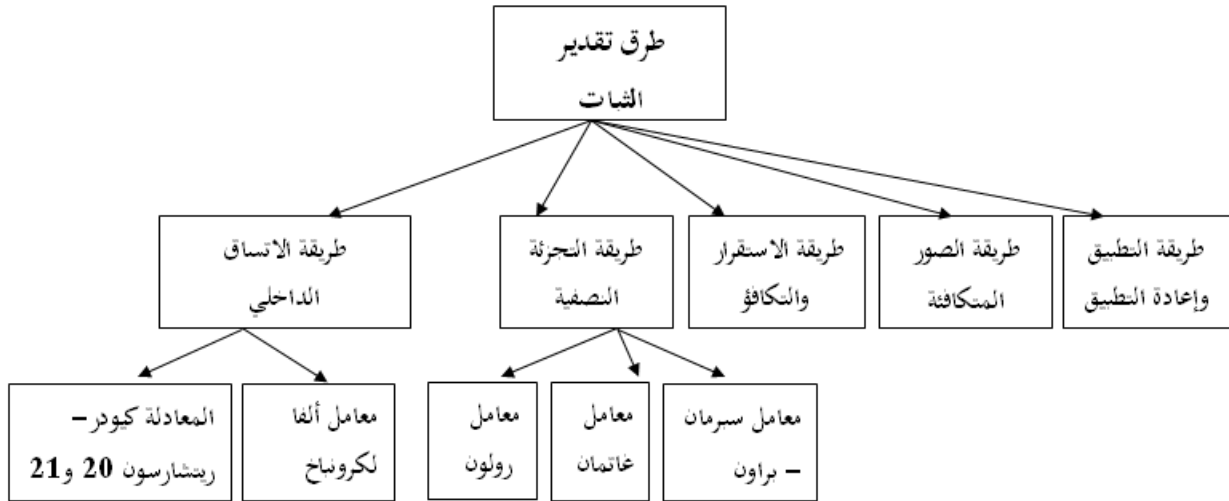
فرضي معين يجب أن تترجم التكوين الفرضي التجريدي إلى صورة عيانية و عملية توفير وصف مفصل للعلاقة بين سلوكيات معينة . و تكوينات تجريدية يشار إليه على أنه تفسير التكوين الفرضي و هو المفتاح لصدق التكوين الفرضي لأي اختبار .

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

طرق تقدير ثبات الاختبارات:

هناك عدة طرق لتقدير ثبات الاختبارات نلخصها في الشكل التالي:



الشكل رقم (02) : طرق تقدير الثبات

1. طريقة التطبيق وإعادة التطبيق (ثبات الاستقرار):

من خلال الشكل أعلاه فان الطريقة الأولى تقوم على تطبيق الإختبار على عينة من الأفراد يطلق عليها عينة التقنين، ثم يعاد تطبيق الإختبار على نفس العينة بعد فترة زمنية لا تقل عن أسبوع ولا تزيد عن ستة أشهر و هذا ما يشير إليه ، ثم نحسب معامل الارتباط بين الإجراء الأول والإجراء الثاني وهو ما نسميه معامل الإستقرار .

2. طريقة الصور المتكافئة (ثبات التكافؤ) :

اما الطريقة الثانية لتقدير معامل الثبات فتقوم فكرة هذا المعامل على حساب الارتباط بين علامات عينة من الأفراد على الإختبار المرغوب حساب ثباته وعلاماتهم على إختبار آخر مكافئ للإختبار الأول وقياس نفس السمة

أو الخاصية ويقصد بالتكافؤ هنا، أنّ للاختبارين 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)

أما طريقة الاستقرار والتكافؤ تجمع هذه الطريقة بين الطريقتين السابقتين، فمعامل الاستقرار والتكافؤ، 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 تدقيقاً، وقيمه بمثابة الحد الأدنى لتقدير معامل الثبات.

أَنَّ بجمع بين طريقتي الإستقرار 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 وإّما بعد فترة زمنية تحدد من طرف الباحث وحسب الغرض من الإختبار كما هو الحال في حساب معامل الإستقرار.

ولإيجاد القيمة التقديرية لمعامل الإستقرار والتكافؤ نستخدم نتائج الإختبارين المتكافئين، ونوجد قيمة معامل الإرتباط بيرسون بين درجتهما.

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، 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 .

أ. طريقة التجزئة النصفية:

تشبه طريقة التجزئة النصفية لتقدير الثبات نظريا طريقة ثبات التكافؤ، مع ذلك تعتبر طريقة التجزئة النصفية مؤشرا لقياس الإتساق الداخلي، لأن الشكلين المتكافئين يؤلفان إختباراً واحداً.

وعند تقدير الثبات بطريقة التجزئة النصفية، يتم الحصول على على درجة فرعية لكل من النصفين، ثم يتم حساب معامل الإرتباط بين هذين النصفين، ويكون معامل الإرتباط هذا تقديرا لثبات إختبار طوله نصف طول

الإختبار الأصلي " ، ولتقدير ما سيكون عليه ثبات الإختبار ككل يتعين عليه تعديل هذا المعامل، و في مايلي طرق تصحيح معامل الثبات بطريقة التجزئة النصفية:

• **معادلة سبيرمان- براون:**

تقوم هذه المعادلة شحاتة (1999) على أساس أنه يمكن التنبؤ بمعامل ثبات الإختبار إذا علمنا معامل ثبات نصفه، حيث إشتق كل من سبيرمان و براون (Spearman&Brown) صيغة رياضية عامة لتقدير معامل ثبات الإختبار إذا زاد طوله أو نقص (ك) من مرات وذلك في إطار مبادئ النظرية الكلاسيكية للإختبارات، في حالة التجزئة النصفية فإن $K=2$ ، ومعامل الثبات التقديري

$$\frac{2r_{12} \times 2}{2r_{12} + 1} = r_{11}$$

وتفترض هذه الصيغة تساوي تباين درجات كل من نصفي الإختبار.

• **معادلة جاتمان (Gattman):**

يعاب على طريقة سبيرمان- براون أنها تفترض تساوي الإنحرافات المعيارية لكل من النصفين فإذا لم يتساوى التباين بين الجزئين لجأنا إلى معادلة جاتمان (فلانجان) 1954 و التي تعتبر صورة مبسطة لمعادلة رولون، تستخدم هذه الطريقة عند عدم تساوي تباين درجات الأفراد على النصفين .
معامل ثبات الإختبار ككل :

$$r_{ss} = \left[\frac{e_1^2 + e_2^2}{e_k^2} - 1 \right] \times 2$$

حيث: e_1^2 : تباين درجات الأسئلة الفردية .

e_2^2 : تباين درجات الأسئلة الزوجية.

e_k^2 : تباين درجات الأسئلة الكلية في الإختبار.

• **معادلة رولون:**

رولون (Rollon 1939) إقترح طريقة لحساب الثبات باستخدام طريقة التجزئة النصفية التي تتطلب حساب تباين الفرق بين درجات الأفراد على نصفي الإختبار (e_1^2) وتباين الدرجة الكلية (e_k^2) والتعويض بهما في نصف المعادلة التالية:

$$r_{ss} = 1 - \frac{E^2}{E^2}$$

• معادلة هورست Horest Formula :

أحيانا يواجه الباحث صعوبة وهي عدم تقسيم الإختبار إلى نصفين متكافئين تماما، في هذه الحالة يقترح ، تطبيق معادلة هورست التي تعتمد على معامل الإرتباط بين القسمين ونسبة عدد المفردات لكل قسم إلى المفردات في الإختبار كله.

معامل ثبات الإختبار ككل

$$r_{ss} = \frac{\sqrt{r^2 + 4b \times c - (r-1)^2}}{2b(r^2)}$$

• معادلة جالكسون:

تمتاز معادلة جالكسون عن معادلة سبيرمان - براون "أنّها تصلح لحساب معامل ثبات إختبارات السرعة، لأنّه في حالة الإختبارات الموقوتة لا يكفي الزمن المعطى لحل أسئلة الإختبار ،وكلّما قل الوقت المخصص للإختبار زاد تبعا لذلك نسبة الأسئلة المتروكة و قد يؤدي ذلك إلى ازدياد التشابه القائم بين نصفى الإختبار، ويزداد تبعا لذلك معامل ثبات الإختبار".

ولذا ينصح بتصحيح قيمة العددية لهذا الثبات وفق المعادلة التي إقترحها جالكسون :

$$r_{ss} = \frac{c}{E^2}$$

رأ.أ: هي معامل الثبات قبل تصحيح أثر السرعة.

م ن : متوسط عدد الفقرات المتروكة

ع² خ : تباين عدد الفقرات التي أجاب عنها الأفراد إجابة خاطئة.

أنّ إستخدام طريقة التجزئة النصفية في تعيين معامل ثبات الإختبار يثير عدّة ملاحظات:

أولاً: اختلاف النصف الأول عن النصف الثاني، وخاصة إذا أخذت البنود من (1-50) ثم من (51 - 100) وهذا يعني أن إجابات الأفراد في النصف الثاني سوف تتأثر بعوامل الإجهاد والملل وضيق الوقت أكثر من إجابات الأفراد في النصف الأول وهذا لا يعطي نتائج يمكن الوثوق بها بدرجة كبيرة.

ثانياً: تعدّد طرق تجزئة الإختبار ،يعني أنّه من المحتمل أن نحصل على معامل إرتباط بين نصفى الإختبار في الحالة الأولى يختلف عن المعامل الذي نحصل عليه في الحالة الثانية والثالثة وهكذا.

ب. طريقة كيودر- ريتشارسون: Kuder & Richardson

" تعتمد هذه الطريقة لقياس الثبات على الإتساق في أداء الأفراد على الإختبار من فقرة إلى أخرى، ولذلك فهي ليست بحاجة لتطبيق الإختبار أكثر من مرة، أو تقسيمه إلى نصفين متكافئين ثم تصحيح طوله، وإنما يقسم الإختبار إلى عدد كبير من الأجزاء، بحيث يتكون كل جزء من فقرة واحدة من فقرات الإختبار، وكما زاد الإتساق بين الفقرات زاد ثبات الإختبار ككل "

معايير إستخدام طريقتي كيودر- ريتشارسون (20-21).

أشار كل من صفوت فرج (2000)، أبو حطب وآخرون (1997)، علام (2006) إلى أهم المعايير التي يجب توفرها لإستخدام معامل كيودر- ريتشارسون لحساب الثبات فكانت كما يلي:

1. تستخدم لتقدير ثبات إختبارات القوة فقط.
2. تستخدم في الأدوات التي يكون نظام التصحيح فيها ثنائي أي إما واحد صحيح أو صفر.
3. تستخدم في الأدوات المتجانسة المحتوى التي تقيس سمة أحادية البعد.
4. يجب أن لا يكون عدد الأسئلة المتروكة كبيراً.
5. يجب أن تتقارب مستويات صعوبة الأسئلة، ويشترط في معادلة كيودر- ريتشارسون (K.R-21) أن تتساوى مستوى صعوبة المفردات حوالي (0,50)، وتكون معادلة كودر- ريتشارسون (20) وفق الصورة التالية:

$$R = \frac{N}{1-N} \left(\frac{MS_{\text{صح}}}{E^2} - 1 \right)$$

ن : عدد مفردات الإختبار.

ع² : تباين الدرجات الكلية في الإختبار (مربع الانحراف المعياري).

س: عدد الأفراد الذين أجابوا على أي مفردة إجابة صحيحة (درجة صعوبة المفردة).

ص: نسبة عدد الأفراد الذين أجابوا على أي مفردة إجابة خطأ. (علام، 2006، 99)

أما معادلة كودر- ريتشارسون (K.R-21) فتكون كالآتي:

$$R = \frac{N}{1-N} \left(\frac{S(S-N)}{E^2 \times N} - 1 \right)$$

حيث (س) ترمز إلى متوسط الدرجات الكلية في الإختبار، وكما ذكر في المعايير لتطبيق هذه المعادلة لا بد أن يتراوح متوسط درجة صعوبة جميع المفردات (0,50)، ولكن نظراً لصعوبة تحقق هذا الشرط في كثير من

الإختبارات أو المقاييس فإن المعادلة (20) ربما تكون أكثر ملائمة لحساب معامل ثبات التجانس حسب علام (2002).

ج. طريقة كرونباخ :

يمكن النظر إلى معادلة ألفا لكرونباخ باعتبارها متوسطاً لمعاملات ثبات كل الأنصاف الممكنة للإختبار وتعني الدرجة العالية لمعامل الثبات أن هذه الأنصاف ترتبط فيما بينها إرتباطاً عالياً وهذا يتحقق إذا كانت بنود الإختبار تتمتع بدرجة عالية من التجانس، بمعنى أنها تتجه إلى أن تقيس سمة واحدة مشتركة ويعتبر معامل ألفا مؤشراً للدرجة التي يقيس بها الإختبار عاملاً واحداً حيث تعطينا هذه الإختبارات قيماً عالية لمعامل " α " ولهذا السبب يشار إليه على أنه مؤشر للإتساق والصورة الأساسية لمعامل ألفا حسب (الفقي إسماعيل، 2005، ص48) و هي:

$$r = \frac{n}{n-1} \left(\frac{\text{مج } ع^2}{ع^2} - 1 \right)$$

ع²_ن: تباين درجات كل مفردة من مفردات الإختبار .

مج ع²_ن: مجموع تباين درجات جميع المفردات.

ن: العدد الكلي لمفردات الاختبار.

خطوات إعداد وبناء اختبارات تحصيلية :

