

الجمهورية الجزائرية الديمقراطية الشعبية

وزارة التعليم العالي والبحث العلمي

جامعة الشهيد حمزة لخضر - الوادي

كلية العلوم الاجتماعية والإنسانية

قسم علم النفس وعلوم التربية

محاضرات في القياس النفسي

مقدمة لطلبة سنة ثانية ليسانس علم النفس

إعداد:

د/ السعيد نصرات

السنة الجامعية: 2024/2023

المحور الأول: القياس في العلوم التربوية والسلوكية

تمهيد	4
1-1 القياس أهميته وفوائده	5
1-1-1 تعريف القياس	5
1-2 القياس في مجال التربية وعلم النفس:	6
1-3 أهمية القياس التربوي:	7
1-4 نشأة القياس التربوي وتطوره.	8
1-4-3 مرحلة الامتحانات التحريرية الموضوعية Objective Types	9
II-أخطاء القياس	15

المحور الثاني: الخصائص السيكومترية للاختبارات

II- 1. الصدق Validity	19
II- 1-1. معنى الصدق:	20
II- 1-2. مصادر أدلة الصدق Sources of Validity Evidence	22
- 1-3. الدليل المتعلق بالمعيار: Criterion- Related Evidence	24
II- 1-4. الصدق التنبؤي Predictive Validity	26
II- 1-5. الصدق التلازمي Concurrent Validity	27

28	 أنواع المحكات: 6-1-11
29	 Construct- Related Evidence الدليل المتعلق بالمفهوم البنائي 7-1-11
33	 الثبات: 2-11
33	 نظرية الثبات: 1-2-11
36	 طرائق الثبات: 2-2-11
50	 العلاقة بين الصدق والثبات: 3-2-11
51	 العوامل المؤثرة في الثبات: 4-2-11
53	 البنية العاملية: 3-11
53	 النموذج العاملي: 1-3-11
54	 Exploratory factor models النموذج العاملي الاستكشافي: 2-3-11
54	 افتراضات التحليل العاملي الاستكشافي: 3-3-11
54	 حجم العينة: 1-3-3-11
55	 استقلال المفردات: 2-1-3-11
55	 القيم الشاذة والمتطرفة: 3-1-3-11
55	 مقياس ملائمة البيانات (KMO): 4-1-3-11
56	 تدوير العوامل: 5-1-3-11

56	 Confirmatory factor models : النموذج العاملى التوكيدي: 2-3-II
57	 أنواع النموذج العاملى التوكيدي: 3-3-II
60	 الفرق بين التحليل العاملى الاستكشافى والتحليل العاملى التوكيدي: 4-3-II
61	 أهداف التحليل العاملى التوكيدي: 5-3-II
62	 تقدير مؤشرات جودة المطابقة: 6-3-II
64	 قائمة المراجع:

المقدمة

ان الطلبة والباحثين في العلوم النفسية والتربوية يحتاجون بدرجة أساسية الى اتقان المفاهيم والمبادئ المتعلقة بالقياس النفسي وتطبيقاته المتنوعة في المجالات التربوية، والنفسية، والمهنية، وغيرها.

فالقياس النفسي يعد من العلوم المتطورة، ويستند الى نظرية ومنهجية وتقنيات متقدمة، ويستخدم في كل من هذه المجالات نطاقا متسعا من أدوات القياس لتقييم الجوانب المعرفية والوجدانية والمهارية والشخصية والسلوكية والنفسية والعصبية وغيرها. وانه من المهم معرفة كيفية بناء الادوات أو اختيارها على أسس علمية سليمة، واستخدامها في مجالات الاهتمام استخداما مستتبصرا ومناسبا.

وتهتم هذه المطبوعة بتنمية معارف ومهارات الطالب والقارئ فيما يتعلق بالأسس الاحصائية والسيكومترية الضرورية لبناء أدوات القياس بمختلف أنواعها. لذلك تناول بعض الاساليب الاحصائية الاساسية التي تستخدم في تقدير الخصائص السيكومترية للاختبارات والمقاييس المتعلقة بالثبات والصدق.

وتتناول في الفصل الأول القياس والتقويم النفسي من حيث المفهوم والاهمية والتطور التاريخي ومستويات القياس والعلاقة بين القياس والتقويم، كما أهتم الفصل الثاني بعرض تفصيلي عملي لكيفية حساب الصدق وطرائق الثبات، وأنتهت المطبوعة البيداغوجية بقائمة مراجع عربية وأجنبية

المحور الأول: القياس في العلوم التربوية والسلوكية

تمهيد

لا تلك أن تقدم أي علم من العلوم يقاس بدرجة الدقة التي يصل إليها في تحديد مفاهيمه وفي دقة الأدوات المستخدمة لقياسه، كما لا تلك أن عملية القياس في التربية وعلم النفس عملية شاقة ومحكمة بالمقارنة بالعلوم الطبيعية الأخرى نظراً لأن موضوع القياس يعتمد على الملوك البشري في مجالات أدائه المتحدة العقلية والاتصالية والنفس حركية

هذا السلوك المكث التي يتغير بتغير الموقف والأقل قابلية للتحكم قد تقدم كما كبيراً في إعداد أدوات القياس النفسي والتربوي يفضل استخدامه العلم الإحصاء والحزم الإحصائية والبحث الراهن يتعرض المعنى القياس وأهميته ونظرية في مجال التربية والتعليم بمعرفة خصائصه ومستوياته هذا من جهة ومعرفة ماهية التقويم من حيث أنواعه وخطواته .

ويهدف هذا البحث إلى تزويد الطالب بالحلق والمفاهيم والمهارات الأساسية الأولية في

موضوع القياس ومدى فهم المفاهيم والتفريق بينهم حتى تصل الفائدة المرجوة من العرض .

1-1 القياس أهميته وفوائده

1-1- تعريف القياس

القياس لغة من قاس بمعنى قدر تقول قاس الشيء بعيد أو على غير أي فرد على مثاله. أما اصطلاحاً يعرفه ستيفنز (1915) Stevens هو عملية تحديد أرقام الأشياء، أو أحداث وفق قوانين. (سبع، 2008: 15)

ويعرفه ننائي (1970) Nannaly فيرى أن القياس وصف للبيانات أو المعطيات بأرقام وفق قواعد معينة، بحيث تدل على الأشياء بطريقة تشير إلى كميات من الخاصية. (فؤاد أبو حطب، 1973: 36)

من خلال التعريفات السابقة نستنتج أن القياس عملية إعطاء رقمي (مقياس) للخصائص أو الصفات موضوع الاهتمام بوحدة معيارية متفق عليها.

من المعلوم أن القياس النفسي لا يختلف كثيراً عن القياس التربوي في شيء فهما يقتصران على موضوع واحد وهو سلوك الفرد باعتباره محصلة عدة قوى تعكس لنا صورة نشاطه العقلي/ المعرفي المتمثل في: القدرة العقلية العامة أو الذكاء، واستعداداته، وقدراته الطائفية وتحصيله، وأيضاً نشاطه الانفعالي/ الوجداني المتمثل في ميوله، واتجاهاته، وقيمه وسماته الشخصية، وأخيراً نشاطه المهاري الحركي أو ما يعرف باسم النفسحركي، إضافة إلى أن القياس النفسي نشأ ونما في أحضان التربية، فمعظم المقاييس النفسية جربت وقننت على عينة من التلاميذ في مختلف المراحل الدراسية وفي العديد من المؤسسات التربوية، ولا يزال علم النفس يأخذ من مشكلات التربية دفعات قوية يحسن بها أدواته.

فإذا كان القياس يستخدم الاختبارات المرجعية للمحك C.R.T (أو المرجعة للهدف (O.R.T (Objectiv Refrend Test) يسمى بالقياس التربوي.

أما إذا كان القياس يستخدم الاختبارات المرجعة إلى المعيار N.R.T (أو المرجعة إلى الجماعة G.R.T (Group Refrenced Test) يسمى بالقياس النفسي.

والغرض الرئيسي من القياس سواء كان نفسياً أو تربوياً هو الكشف عن الفروق الفردية بأنواعها سواء كانت فروقاً بين الأفراد (Inter Individual) أو بين الجماعات (Inter group) أو فروقاً في ذات الفرد (Inter Individual) مستخدمة أساليب وأدوات قياس متعددة، ولا يقتصر هدف القياس على التشخيص بل يمتد إلى التفسير والتنبؤ والتحكم.

1-2- القياس في مجال التربية وعلم النفس:

نحن نقيس عدداً كبيراً من الظواهر النفسية، ففي المجال العقلي/ المعرفي Cognitive Domain نقيس العمليات العقلية مثل التعلم، والتفكير، والتذكر وكذلك نقيس القدرات العقلية العامة (الذكاء)، والاستعدادات والقدرات الطائفية، مثل: القدرة الكتابية والقدرة المكانية والقدرة اللغوية والقدرة الرياضية والقدرة الميكانيكية... إلخ، وكذلك القدرات الخاصة. (صلاح وآخرون،

2005: 14)

وفي المجال الانفعالي الوجداني (Affective Domain) نقيس الميول والاتجاهات والقيم وبعض سمات الشخصية مثل الانبساطية مقابل الانطوائية، السيطرة في مقابل الخضوع ...إلخ، وفي المجال النفسحركي Psycctomotor Domain نقيس المهارات الحركية لدى الأفراد، مثل مهارة استخدام الإعلام الآلي ومهارة تشغيل واستخدام الأجهزة في المعمل ...إلخ.

1-3- أهمية القياس التربوي:

لا شك أن القياس عملية هامة وجوهرية في رصد التقدم العلمي للظواهر التربوية والنفسية وتطورها، ولذا نستخدم نتائج القياس في التربية للتعرف على مستويات الطلبة وقدراتهم، كما نستخدم للكشف عن الفئات الخاصة وفي التوجيه التربوي Education Guidance، أي التوجيه عام أو تقني، وأيضاً التوجيه لنوع الشعبة التي تناسب قدرات واستعدادات وميول الطالب، وتستخدم نتائج القياس أيضاً في الإرشاد النفسي Counselling حيث يستخدمها المرشد النفسي المدرسي School Consellor للتغلب على بعض مشكلات التكيف المدرسي المتمثلة في التخلف الدراسي، والسلوك العدواني، والغش في الامتحانات ...

1-4- نشأة القياس التربوي وتطوره.

بحكم تخصصنا في علوم التربية سنحاول التطرق إلى القياس في ميدان التربية لعله أقدم كثيراً من القياس في علم النفس، فقد يشعر العاملون في ميدان التربية بالحاجة الماسة إلى قياس كمية المعلومات والمهارات السابق دراستها، ومقدار التقدم أو التأخر الحادث في التحصيل الدراسي للطلبة في المواد المختلفة، وأيضاً الرغبة في التعرف على مدى نجاح جهودهم في عملية التدريس بهدف التعرف على أفضل الطرق الممكنة للتدريس مما جعلهم يهتمون بإعداد أدوات القياس تلك المهارات.

ومن الجدير بالذكر أن القياس التربوي في أول الأمر كان يعتمد على الملاحظة الذاتية، والآراء الشخصية للحكم على دقة سير العملية التعليمية، لذا يلاحظ أن القياس التربوي مر بثلاث مراحل نمو وتقدم يمكن إجمالها فيما يلي:

1-4-1- مرحلة الامتحانات الشفوية: **Oral Examination**، ولهذا جذورها القديمة عند

الصينيين منذ عام 2000 قبل الميلاد.

1-4-2- مرحلة الامتحانات التحريرية: **Written Examination**

وهي تستخدم الاختبارات من نوع الورقة والقلم Paper and Pencil Test غير الموضوعية Sybjective Types من نوع المقال Essay Type بصورة المختلفة (المقال القصير أو المقال المستفيض)، وقد يرجع تاريخها إلى جامعة كمبرج بإنجلترا عام 1800 وفي أمريكا 1845.

1-4-3- مرحلة الامتحانات التحريرية الموضوعية Objective Types

وهي وليدة القرن العشرين، ويعتبر ثور نديك Thorndike الرائد الأول لحركة الاختبارات التحصيلية المقننة، فتن نشر اختباراً للخط عام 1910، ونشر كورتس Courtis عام 1911 اختباراً للحساب، وتوالت بعد ذلك الأنواع المختلفة للاختبارات التحصيلية المقننة.

1-4-4- مرحلة الامتحانات العملية (اختبارات الأداء) Performance

والتي اهتمت بإعداد أدوات لقياس الأداء العملي.

1-5- خصائص القياس النفسي والتربوي:

يتميز القياس النفسي والتربوي بمجموعة من الخصائص العامة يمكن تحديدها في النقاط الآتية:

1-5-1- القياس هو عبارة عن تقدير كمي (رقمي) Quantitative لبعدها من أبعاد السلوك

الإنساني، وهذا التقدير يعبر عن مستوى الأداء في السمة موضع القياس.

1-5-2- أنه قياس غير مباشر Indirect.

1-5-3- أنه قياس نسبي Relative وليس مطلق Absolute.

1-5-4- الصفر في القياس النفسي والتربوي هو صفر افتراضي غير حقيقي.

1-5-5- توجد أخطاء في القياس النفسي والتربوي.

أ- أخطاء الملاحظة

ب- عدم حساسية الأدوات في قياس الظاهرة.

ج- عدم ثبوت الظاهرة المقاسة.

1-6- مستويات القياس:

يقصد بمستويات الطريقة التي تصنف بها الأشياء أو ترتب بها أو تقارن بينها تبعاً

للخصائص المشتركة، بهدف معرفة الفروق بينها وأيضاً مدى هذه الفروق، والمعالجات

الإحصائية الممكنة لهذا المدى.

لذا اقترح العالم الإنجليزي ستيفنز (1951) Stevens أربعة مستويات للقياس في ميدان

علم النفس والتربية نرتبه من البسيط إلى المعق كما هو موضح في الجدول:

المقاييس	الخواص المميزة له	أمثلة
الاسمي Nominal	- تدل الأعداد لأغراض التسمية أو التصنيف. - التصنيف يتم في فئات بحيث يتساوى أفراد الفئة في الصفة (ذكر، أنثى)، (ناجح، راسب). - المعالجة الإحصائية الممكنة هي عدد الحالات داخل الفئة حساب المنوال، الوسيط).	أرقام اللاعبين، الشوارع، التليفونات، المنازل، السيارات. التصنيف (ذكر، أنثى) الحالة الاجتماعية: أعزب، متزوج. الحالة النفسية: سوية، غ سوية.
الترتيبي Ordinal	- الفروق في الرتب لا تدل	- الطلاب في الفصل.

- المتسابقين حسب زمن وصولهم. - ترتيب المعادن. - ترتيب الأبناء داخل الأسرة.	على فروق متساوية في الصفة المقاسة. - الأرقام التي تعطى للترتيب لا تقدم معلومات عن كم أو مقدار الصفة المقاسة. - تستخدم فكرة الزيادة أو النقصان (أكبر، أقل) دون تحديد كمياتها. - المعالجة الإحصائية الممكنة: لا يمكن إجراء العمليات الحسابية الأساسية، ولكن يمكن حساب معامل ارتباط الرتب، الوسيط، المئينيات.	
---	--	--

اختبارات التحصيل، الذكاء، الاستعدادات، الميول، الاتجاهات. مقياس الحرارة، الرطوبة والضغط.	-تتمتع بوحدات متساوية من الصفة المراد قياسها، الفروق المتساوية في الأرقام تقابل فروقاً متساوية في الصفة المراد قياسها، يتدرج المقياس بحيث تتساوى وحداته لتبين الزيادة وكمياتها. - المعالجة الإحصائية الممكنة هي: جميع العمليات الحسابية (ما عدا القسمة)، المتوسط، الانحراف المعياري، معامل الارتباط بيرسون، اختبارات الدلالة: ت، ف.	الفتري Interval
قياس الطول، والوزن، وشدة	تتمتع بوجود صفر مطلق	النسبي

Ratio	(حقيقي)، وكذلك وحدات متساوية يتم قياس الصفات بطريقة مباشرة. - النسبة بين أي درجتين على المقياس لا تتأثر بوحدات القياس. - العمليات الإحصائية والحسابية الممكنة مسموح بها والمعالجات الإحصائية المختلفة ممكنة مثل: مقاييس النزعة المركزية، التشتت، الارتباط، دلالة الفروق (ت، ف).	الصوت. - قياس الخصائص العقلية والانفعالية بوحدات فيزيائية مثل: قياس زمن الرجوع، قياس التعلم بوحدات الزمن.
-------	---	--

II- أخطاء القياس

مهما حاول القياسيون النفسيون أو يكونوا دقيقين في القياس، إلا أنهم لا ينجون من الوقوع في أخطاء، وهذه الأخطاء بعضها خارج عن إرادتهم. ومن الضروري أن يكونوا على وعي بهذه الأخطاء قبل قيامهم بالقياس، كي يحدوا منها، ويخلصوا منها الدرجات التي يحصلون عليها بواسطة الطرق الإحصائية، أو يأخذونها بعين الاعتبار عند تفسير هذه الدرجات.

وتتجم هذه الأخطاء عن ما يلي :

1 - عدم حساسية أدوات القياس : بعض أدوات القياس أكثر دقة من بعضها الآخر. والاختبار الذي يصممه فاحص خبير مدرب أكثر حساسية لما يقيس من اختبار وضعه فاحص لا يدري عن القياس شيئاً. والامتحان الموضوعي أكثر دقة من الامتحان المقالّي .

2 - عدم ثبات الظاهرة السلوكية : معظم ما نقيس في ميدان علم النفس ديناميكي، أي متقلب ومتغير فالشخص الذي نقيس أداؤه أو سمة من سماته، كل ساعة هو في شأن. ففي البداية يكون في حالة نفسية معينة، وبعد ساعة تتغير حالته، وبعد قليل قد يشعر بالتعب أو الملل أو الضجر أو الخوف، أو يفقد اهتمامه وتتنخفض دافعيته. ويتأثر أيضاً بدرجة الحرارة والتهوية والإضاءة، وبقيّة المتغيرات البيئية. فالسلوك البشري في حالة تغير مستمر، وكثيراً ما يخدع القياسي النفسي، ومعقد تعقيداً كبيراً.

لذا فإن القياسيين النفسيين يحاولون أن يقتربوا من الحقيقة بأخذ عدة قياسات، ثم يستخرجون متوسطها واعتباره أقرب إلى الأداء الحقيقي. أما الفرق بين متوسط القياسات وأي قياس منها يعتبر خطأ في القياس.

3 - خطأ الملاحظة : تختلف دقة الملاحظة من شخص إلى آخر.

فلو طلبنا من عدة أفراد قياس طول فرد، فإنهم يعطوننا أطوالاً مختلفة مع أن أداة القياس واحدة لم تتغير، وكذلك الفرد موضوع القياس. فتأتي 150.45، الأطوال كما يلي : 150.00 سم، 150.50 سم، 150.40 سم سلم ... وهذه الفروق في القياسات ممكنة الحدوث لدى الشخص نفسه إذا قام بالقياس عدة مرات.

الفرق والعلاقة بين العد والتقدير والقياس :

هناك فرق جوهري بين العد والقياس وهو ما يلي : العد يكون دائماً بوحدات كاملة ويتصف بالدقة، فنقول : ثلاثون طالبا وعشرون فردا وأربعون شجرة والبيانات التي نحصل عليها بالعد نتعامل معها إحصائياً على شكل تكرارات أو نسب مئوية. أما القياس فهو تقريبي دائماً، تنتج عنه قيم قياسية تمثل الظاهرة المقاسة بدرجة تقريبية، وهذا التقريب يعتمد على دقة أداة القياس المستعملة. فلو قمنا بقياس طول بعد ارتفاع جدار مثلاً، فإن نتيجة القياس تكون تقريبية وتتجزأ إلى وحدات صغيرة. فلو كان طول هذا البعد عشرة أمتار، فإن طوله إما أن يكون أقل من عشرة أمتار بسنتيمترات أو بملترات، أو أكثر بسنتيمترات أو بملترات، ونفس الشيء إذا وزنت ثقبلاً، فإذا كان وزنه عشرين كيلو غراماً، فهو يزيد عنها ببضع غرامات أو يقل عنها ببضع غرامات وهكذا. فالقياس بما أنه تقريبي، فإنه يتضمن دائماً كسوراً. ولهذا فإذا قلنا أن طول الفرد هو 150 سنتمتر، فمعنى هذا أن طوله يتراوح بين 149.50 - 150.49 سنتمتر.

وهناك فرق آخر بينهما وهو ما يلي : وحدات العد تكون متقطعة، لأن الفرد يعد أشياء تكون منفصلة عن بعضها، ولهذا فالمعطيات التي يقدمها العد لا تعتبر ذات طابع كمي. أما وحدات القياس فهي متصلة ومستمرة فلا فجوات بينها، لأن الفرد يقيس خاصية واحدة، ولهذا تعتبر وحداته كمية.

أما العلاقة بينهما، فإن العد يرتبط بالقياس ارتباطا وثيقا. ففي أحيان كثيرة نتعامل مع وحدات العد كأنها وحدات قياس. فعدد الأطفال في الأسرة عدد صحيح، إلا أن الإحصاءات السكانية تشير إلى أن معدل عند الأطفال في الأسرة الجزائرية هو 5.6 طفل. وفي أحيان أخرى فإننا عندما نقيس، فإننا نعد أيضا. فعند قياس طول فرد، فإننا نعد الوحدات التي يساويها طوله وهي السنتيمترات، ويقوم بذلك شريط القياس (المتر). وكذلك عندما نمتحن متعلما باختبار مقالي، فإننا نقوم بقياس معلوماته، ونمثل مقدارها بعدد من الدرجات تكون عادة عددا صحيحا، والدرجات عبارة عن قياسات وهي تقريبية. فإذا كانت درجة المتعلم 15 فهي تتراوح بين 14.50-15.49

التقدير والقياس :

هناك علاقة بين التقدير والقياس لكونهما أدوات لجمع معلومات عن الفرد محل الدراسة، كما يشكلان أيضا أساسا للتقويم. إلا أنه توجد بينهما فروق نذكرها فيما يلي:

- 1 - أدوات التقدير تمدنا بأساس انطباعي ذاتي، بينما أدوات القياس موضوعية.
- 2 - أدوات التقدير كيفية في أغلبها، بينما أدوات القياس كمية.
- 3 - أدوات التقدير يستعملها المقدر (الفاحص غالبا، بينما أدوات القياس يستعملها المفحوص

نفسه. (معمرية، 2007، ص41-43)

المحور الثاني: الخصائص السيكومترية للاختبارات

إن الترجمة العربية للمصطلح السيكومتري هي القياس النفسي ويتألف من مقطعين هما: سيكو (Psycho) وهي تعني نفسي، وهي كلمة تأخذ إحدى الاستعمالات الآتية: كل ما يتعلق بالظواهر النفسية: اسم عام لكل الظواهر النفسية، ومتري (Metry) وتعني بالإغريقية القياس عموماً (ابراهيم، 2012، 49) (Jline, 1997, P.1).

ويقصد بالخصائص السيكومترية للاختبار تلك الصفات الضرورية والمتعلقة بمدى فعالية بنود الاختبار (بوسالم، 2014، 60)

وتشير هذه الخصائص إلى مفهومين أساسيين يتعلقان بالاختبارات والمقاييس التربوية والنفسية هما صدق الاختبار Test validity، مفهوم ثابت درجات الاختبار (علام، 2000، 130).

Test validity يعد مفهوماً الصدق والثبات من أهم المفاهيم التي تركز عليها نظرية القياس النفسي إن لم تكن أهمها على الإطلاق، ومما يعزز أهمية هذين المفهومين أنهما يمثلان الشرطان الأساسيان من شروط صلاح الاختبار النفسي أيًا كان نوعه، ومن المعلوم أن الاختبار النفسي يوضع أساساً لتلبية غرض أو مجموعة معينة من الأغراض، ويكون الأساس في اتخاذ

عدد من القرارات المهمة. ولكي يؤدي الاختبار دوره على النحو الأمثل، ويلي الغرض أو الأغراض المرسومة له، ويتيح بالتالي اتخاذ قرارات صائبة لابد أن يقدم معلومات واقعية وصحيحة حول السمة أو الخاصية التي توضع موضع القياس، أي يكون "صادقا"، وأن يكون حساسا للفروق الدقيقة بين الأفراد في هذه السمة أو الخاصية، أي يكون "ثابتا" (امطانيوس، 2016، 163).

II- 1. الصدق Validity

تقوم نظرية الصدق على فلسفة الأنماط، باعتباره خاصية وليس عملية ملتزمة بالدلالة النظرية للمفهوم. ولما كان الصدق شأنه شأن الثبات خاصية أو صفة في أداة القياس فمن المنطق أن تصنف إلى أنماط متميزة ليشكل الصدق مظلة لتتنوع أنماطه أو أنواعه. ولقد لخص الدليل الإرشادي للقياس التربوي والنفسي الأول، الذي تم نشره سنة 1954 من طرف الجمعية الأمريكية لعلم النفس تحت عنوان التوصيات أو الإرشادات الفنية المتعلقة بالمقاييس النفسية والأدوات الاسقاطية Recommendation for Psychological Testand Diagnostic Techniques أربع أنواع من الصدق: صدق المحتوى Centent Validity والصدق التنبؤي Predictive Validity، والصدق التلازمي Validity Concuvent وصدق المفهوم، وقلص الدليل الإرشادي الصادر سنة 1966 هذا التصنيف الرباعي إلى تصنيف ثلاثي والذي أشير له

من النقاد بنوع من التفكه بالنظرة الثالوثية للصدق، إذ أبقى على صدق المحتوى، وصدق المفهوم، أما الصدق التتبوي، الصدق التلازمي فأدرجا تحت مسمى الصدق المعياري Criterion Validity كما اتسمت أيضا بالتركيز على التمايز الوظيفي لأنواع الصدق، أي أن لكل نوع وظيفة خاصة به تختلف عن وظائف الأنواع الأخرى من الصدق، وترتب على ذلك أن كل نوع يناسب بعض أنواع الاختبارات والمقاييس ولا يناسب الأنواع الأخرى، فصدق المضمون يناسب اختبارات التحصيل، والصدق المحكي يناسب اختبارات القدرات والاستعدادات، وصدق المفهوم يناسب اختبارات الشخصية (تيعزة، 2008، 02).

II-1-1. معنى الصدق:

يعرف الصدق عادة بما " إذا كان الاختبار يقيس ما يهدف لقياسه". وهذا يتلاءم جدا مع التعريف القاموسي للمصطلح "صدق" الذي يعني ما إذا كان شيء ما متأسس جيدا أو منطقي، أو يمكن تبريره وينبغي مراعاة ما يأتي عند التحقق من صدق اختبار معين (Dennis, Duncan, 2016, 400).

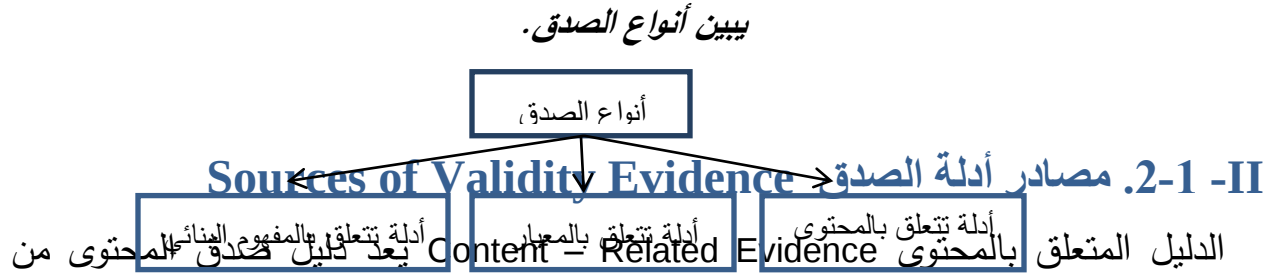
وقد أشار ن. ي جرونلند في سياق حديثه عن الصدق إلى أن أقرب العبارات إلى هذا المفهوم وأكثرها دلالة عليه تلك العبارة التي يدلي بها الشاهد في المحكمة عادة والتي تقول "إنني أتعهد أن أقول الحق، كل الحق ولا شيء غير الحق" (فرج، 1980، 276).

إن مفهوم الصدق في الاختبار يشير أن الاختبار يقيس السمة أو الصفة التي صمم الاختبار لأجل قياسها، فالصدق يتناول العلاقة الأساسية بين المفهوم والاختبار (عباس، 1996، 24). فمثلا الاختبار المقنن لمادة القياس قد يستخدم لقياس تحصيل نهاية السنة في مادة القياس في الكلية، ففي هذه الحالة قد يهتم المرء بالمدى الذي يمكن للاختبار أن يتنبأ فيه، بالتحصيل المستقبلي فالأهداف المختلفة للاختبارات تتطلب أنواعا مختلفة من الشواهد لدعم صدق ذلك الاستخدام المعين، ولذا عرفه مسيك (1989) Messik بأنه حكم تقييمي متكامل للدرجة التي بها تؤيد الأدلة الامبريقية والمنطق النظري كفاية وملائمة الاستدلالات والأفعال استنادا إلى درجات الاختبار وغيرها من أنواع التقييمات (Messick, 1989, 13).

كما عرفت جمعية البحوث التربوية الأمريكية (AERA et al) الصدق بأنه الدرجة التي تؤيد به الأدلة والنظرية وتفسير درجات الاختبارات الذي يستلزمه استخدامات مقترحة معينة للاختبارات. (AERA et al, 1999, 09) ورغم أن الصدق مفهوم موحد إلى أن هناك أنواعا من الأدلة يمكن جمعها لدعم الاستدلالات المستمدة من درجات أداة القياس، فالمعايير المعدة من لجنة مشتركة من الجمعية النفسية الأمريكية (APA) وجمعية البحوث التربوية الأمريكية (AERA) والمجلس القومي في التربية (NCME) تصنف أدوات جمع أدلة الصدق

في الأصناف التالية: دليل صدق المحتوى، دليل صدق المعيار، دليل صدق المفهوم البنائي (Donald , et al, 2013, 285).

شكل (05):



أفضل أنواع الأدلة صدق الاختبارات التحصيلية لكون المادة الدراسية نطاقها محدود (Adkins, 1974, P132).

وذكرت وثيقة المعايير (AERA et al, 1999) أنه يمكن الحصول على أدلة قيمة بفحص العلاقة بين محتوى الاختبار، والنطاق الذي صمم الاختبار من أجل قياسه، ومحتوى الاختبار في هذا السياق يشمل الأفكار الرئيسية التي اشتمل عليها البند، وصياغتها وصبغتها، ومهامها أو أسئلتها. وكذلك الارشادات المتعلقة بتطبيق الاختبار وتقدير درجته (AERA, et al, 1999, p.11).

فقد ذكر رينولدس (Reynolds 1986) أن أدلة الصدق التي تعتمد على محتوى الاختبار تركز على جودة معاينة مفردات الاختبار للسلوكيات أو المادة الدراسية التي يصمم الاختبار من أجل قياسها.

وبعبارة مماثلة ذكرت انستازي واريناي (1997) أن أدلة الصدق التي تعتمد على محتوى الاختبار تتضمن فحص محتوى الاختبار لتحديد ما إذا كان يقدم عينة ممثلة للنطاق موضع

القياس، وقد صاغ (Popham 2000) وذلك بدقة كالآتي: "هل يغطي الاختبار المحتوى الذي

يفترض أنه يغطيه؟" (Popham, 2000, P.96).

وإن الدليل المتعلق بالمحتوى هو قياس لمدى تمثيل الاختبار لنواحي الجانب المقاس، وهنا نقوم بتحليل منطقي لمواد الاختبار وبنوده لتحديد الوظائف والجوانب والمستويات الممثلة فيه ونسبة كل منها إلى الاختبار ككل (الأنصاري، 2000، 96).

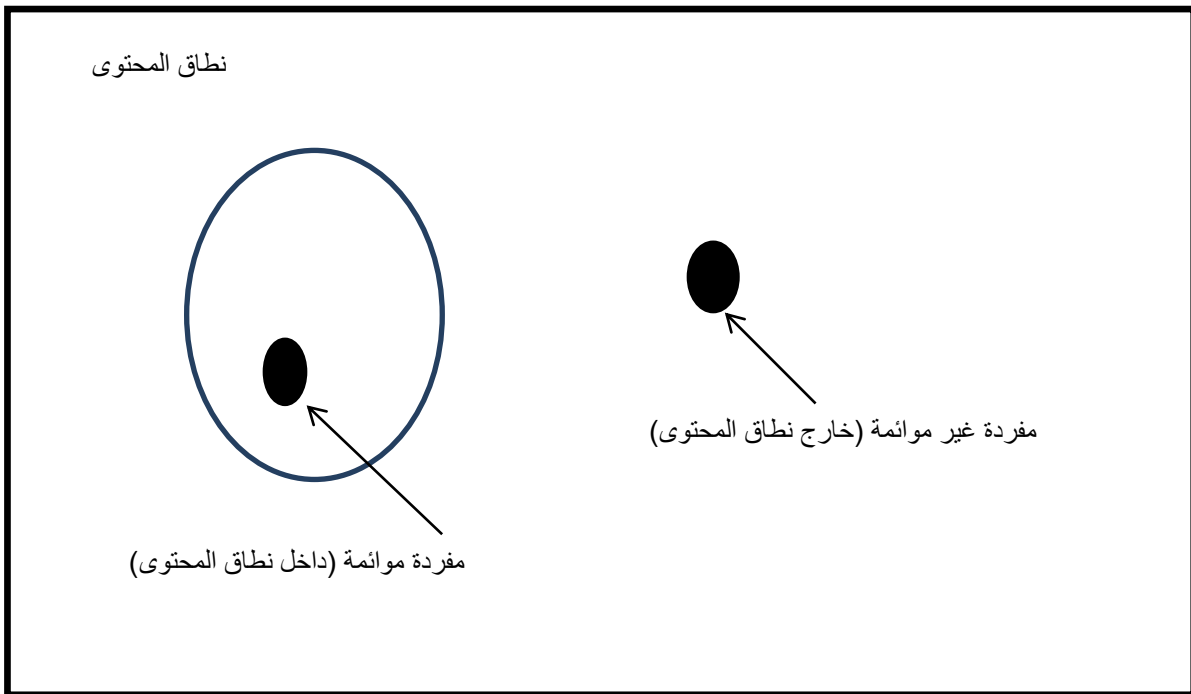
ويتناول الخبراء أو المحكمون قضيتين رئيسيتين هما: مواءمة المفردات وشمول المحتوى،

ولتقييم مواءمة المفردات يفحص الخبراء كل النطاق المعين، ولتقييم شمول المحتوى ينظر

الخبراء إلى الاختبار ككل ويقدرّون درجة اشتغال المفردات على النطاق المعين والأشكال الآتية

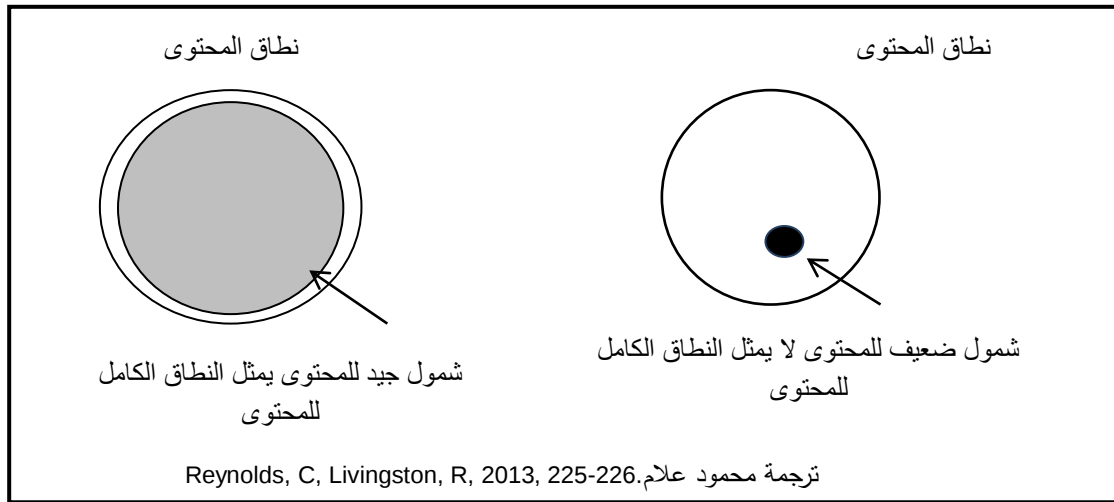
توضح ذلك:

شكل (06):



شكل (07):

تمثيل شمول المحتوى.



- 3-1. الدليل المتعلق بالمعيار: Criterion- Related Evidence

يبين الدليل المتعلق بالمعيار، المدى الذي ترتبط به درجات أداة القياس مع متغير خارجي

مستقل (معيار) (Donalad, et al, 2013, 287).

كما عرفه فؤاد أبو حطب (2003) بأنه يقوم على دراسة الارتباط بين درجات المقياس

ودرجات مقياس آخر يعتمد كمحك له (معيار)، ويقصد بالمعيار مقياس مباشر ومستقل لما

يهدف الاختبار إلى قيامه، والتنبؤ به أو تشخيصه أو هو ميزان لتحديد مدى صلاحية الاختبار، أي هو اختبار للاختبار (أبو حطب، 2003، 148).

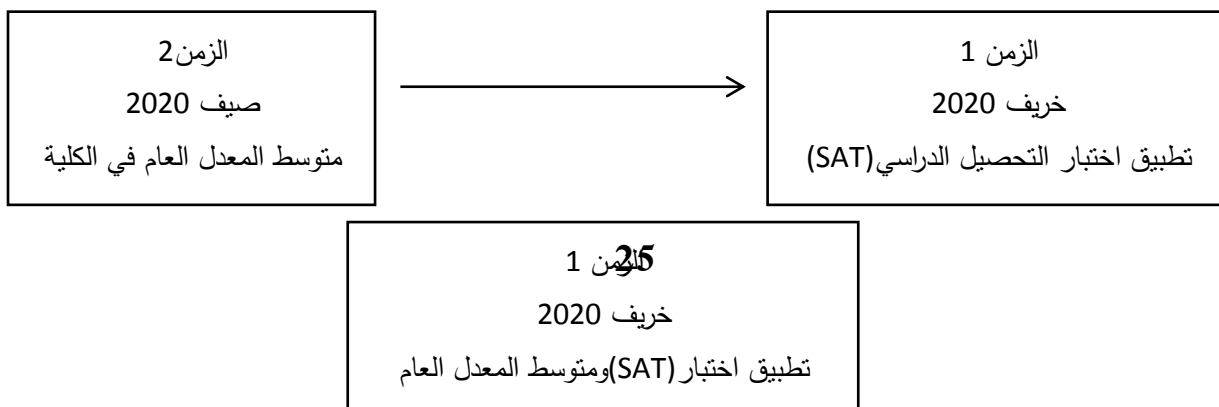
يصمم كثير من الاختبارات للتنبؤ بالأداء على متغير معين، يطلق عليه عادة "المحك"، وقد عرفت وثيقة المعايير (AREA et al, 1999) المحك بأنه "مقياس لخاصية أو ناتج معين يكون موضع اهتمام أساس" (AREA , et, 1999, P.14).

ومن الوجهة التاريخية أطلق على هذا النوع من أدلة الصدق التنبؤي أو صدق المحك، أو الصدق المتعلق بالمحك، ويوجد نوعان مختلفان من دراسة الصدق تستخدمان عادة لجمع أدلة عن "العلاقة بين الاختبار ومحك معين هما الدراسات التنبؤية، والدراسات التلازمية، وفي الدراسة التنبؤية" يطبق الاختبار، وتوجد مدة زمنية فاصلة، ثم يقاس المحك، وفي الدراسة التلازمية، يطبق الاختبار، ويقاس المحك في نفس الوقت تقريبا (Gronlund, 1981, P.83).

شكل (08):

تمثيل الدراسات التنبؤية والتلازمية.

تصميم تنبؤي



وعلى هذا الأساس، تحدد معايير القياس (APA,1985) نوعين من الصدق المرتبط بالمحك وهما: الصدق التنبؤي والتلازمي (عبد الخالق، 1993، 184).

II- 4-1. الصدق التنبؤي Predictive Validity

يشير الصدق التنبؤي إلى مدى الصحة التي نستطيع أن نقدم بها تخمينات أو توقعات عن خاصية معينة لدى الفرد من خلال وسيلة مستقلة ويقدم هذا النوع من الصدق معلومات ترتبط بالاختبارات التي تستخدم في انتقاء الأفراد وتصنيفهم وتوجيههم تربوياً أو مهنياً أو عسكرياً وكذلك في أغراض التنبؤ (اسماعيل، 2004، 87).

في الصدق التنبؤي Predictive لا يتوافر المحك في الحاضر بل في المستقبل، وتصلح هذه الطريقة لحساب الصدق في الاختبارات التي تصمم بهدف اختيار المستخدمين، وتصنيفهم، واختيار الطلاب للالتحاق بكلية معينة، أو تحديد برامج التدريب المهني للمستخدمين في الجيش، أو استخدام الاختبارات لفرز من يحتمل أن يطوروا الاضطرابات الانفعالية في البيئات الضاغطة، واستخدام الاختبارات لتحديد المرضى في المجال الطبي النفسي الذين يحتمل أن يستفيد من علاج معين (عبد الخالق، 1993، 185).

ويقاس معامل الصدق التنبؤي بإيجاد العلاقة بين الدرجات على الاختبار، والدرجات على مقياس المحك الذي يطبق بعد إجراء الاختبار، ويعد جمع البيانات عن المحك (ناجح، راسب) ولحساب الدرجات على الاختبار نقوم بإيجاد العلاقة بينهما بأحد الطرق الثلاث هي:

1-طريقة النسب المئوية.

2-طريقة المتوسطات.

3-طريقة معامل الارتباط (الأنصاري، 2000، 100).

إذ هذه الأخيرة أشهر الطرق وأكثرها استعمالاً من طرف الباحثين للتحقق من الصدق المحكي لأي اختبار حيث أن قيمة معامل الارتباط ما بين نتائج تطبيق الاختبار ونتائج تطبيق المحك بمثابة معامل صدق، تفسر على أساس اقترابها أو ابتعادها عن قيمة واحد الصحيح (+1) (بوسالم، 2014، 67).

II-5-1. الصدق التلازمي Concurrent Validity

يشير الصدق التلازمي إلى ذلك المدى من التطابق أو الارتباط بين نتائج الاختبار الحالي ونتائج محك في الوقت نفسه أو خلال فترة زمنية قصيرة، ويمكن التحقق من دلالات الصدق التلازمي للاختبار بإيجاد معامل الارتباط بين نتائج الاختبار الجديد ونتائج المحك، وهذا تعكس الارتباطات العالية بين الاختبارين درجة أعلى من دلالات الصدق التلازمي (أبو الديار، 2012، 31).

وترى انستازي أن الصدق التلازمي هو أنسب الأساليب، ويمكن تقديم العديد من التبريرات التي تدعمه، كما نرى أن التمييز المنطقي بين الصدق التلازمي والصدق التنبؤي لا يقوم على الفروق الزمنية بين الاختبار والمحك في الحالتين ولكن على الأهداف من الاختبار، فالفرق يتضح من الاختلاف بين صيغة هذين السؤالين: هل سميت عصيباً؟ (صدق تلازمي)، وهل سميت معرضاً لأن يكون عصيباً؟ (صدق تنبؤي) (مجدي، 1996، 303).

ويستخدم هذا النوع من أنواع الصدق عدة وسائل في تقدير حساب الصدق التلازمي:

1- المجموعات المتضادة.

2- التقديرات.

3- معاملات الارتباط بين الاختبار وغيره من الاختبارات (الأنصاري، 2000، 102-

(103)

II- 1-6. أنواع المحكات:

توجد عدة أنواع من بينها ما لخصته أناستازي (1961):

- الفروق في العمر: وهو إمكانية اتخاذ مدى تغير متوسط درجات الاختبار تبعاً لمتغير

متوسط أعمار العينة كمحك خارجي لصدق الاختبار (مقدم، 2003، 148).

- التحصيل الدراسي: الذي يستخدم في التحقق من صدق اختبارات التحصيل، كما يستخدم في

بعض الحالات في التحقق من صدق اختبارات الذكاء.

- **الأداء في التدريب:** ومحك الأداء في برنامج تعليمي أو تدريب متخصص كأداء الطيارين في مواقف التدريب الفعلية الذي يستخدم عادة محكا لاختبار انتقاء الطيارين.
- **مستوى الأداء في العمل:** والذي يركز على سجلات تتبعه للأداء الفعلي في العمل (أو عدد الأخطاء المرتكبة).
- **المجموعات المتناقضة:** كالمقابلة بين المرضى والعاديين في اختبارات الشخصية. (امطانيوس، 2015، 92).
- **التقدير الشخصي:** يجب أن يكون الشخص الذي يقوم بالتقدير ذو معرفة جيدة بأفراد العينة وأن يكون كفؤاً أو متزناً نفسياً وموضوعياً في أحكامه.
- **الارتباط بينه وبين اختبارات أخرى:** وهو إيجاد معامل الارتباط بين الاختبار الحالي واختبارات أخرى معترف بها، فإذا كان معامل الارتباط مرتفعاً دل ذلك على أن هذا الاختبار صادق (مقدم، 2003، 149)

II-7-1. الدليل المتعلق بالمفهوم البنائي Construct- Related Evidence

ظهر مصطلح صدق المفهوم لأول مرة في تقرير اللجنة القومية الأمريكية لمعايير الاختبارات التي شكلتها الجمعية النفسية الأمريكية لعلم النفس (APA) سنة 1954 ويذكر التقرير أنه "يقدر صدق المفهوم بتحديد المفاهيم التفسيرية والتكوينات النظرية المعينة المسؤولة عن الأداء على الاختبار" (معمرية، 2007، 148-149).

وبهذا المعنى، فإن صدق المفهوم يتعلق ببناء نظرية. وللكشف فإن صدق المفهوم لاختبار معين لا يختلف بدرجة أساسية عن الإجراءات العلمية العامة لبناء النظريات وتوكيدها" (Cronbach and Meehl, 1955, P.300).

ويعد صدق المفهوم من وجهة نظر الباحث عملية واسعة ومعقدة وتتعلق بدرجة أكبر بعملية تطوير العلم أكثر من كونه مؤشر صدق وحيد.

وأشار جرونلاند Gronland إلى ثلاث خطوات:

- التعريف بالإطار النظري للسمة التي يفترض ارتباطها بنتائج الاختبار .
- اشتقاق فرضيات حول نتائج الاختبار من خلال الإطار النظري.
- التحقق من صحة الفرضيات منطقياً أو تجريبياً .

(بشير، 2015، 85) عن (عودة، 1993، 384-385).

وفيما يلي بعض الأساليب والطرق المستخدمة في تأسيس صدق المفهوم:

1- أساليب تعتمد على التحليل المنطقي: تعتمد هذه الأساليب على الفحص الدقيق للاختبار

والأداء الذي يتطلبه، وإحداث تكامل بين نتائج هذا الفحص، وبين النظرية التي يستند إليها الاختبار، وآراء المفحوصين الذين سبق أن اختبروا باختبارات مشابهة، ويترتب على ذلك افتراض وتفسيرات، وربما تختلف عن التفسيرات التي اقترحها معد الاختبار.

ويرى كرونباخ Cronbach أن هذا التحليل المنطقي من أهم مصادر التوصل إلى فروض بديلة فيما يتعلق بالأداء في الاختبارات والمقاييس.

(بشير، 2015، 85) عن (علام، 2002، 227).

2- التحليل العاملي: التحليل العاملي من أهم الوسائل التي يتم اللجوء إليها في تحديد صدق

المفهوم، حيث نبحث في هذه الطريقة لإيجاد الصدق عن عوامل مشتركة تقيسها عدة اختبارات صادقة لنحدد اشتراك هذه الاختبارات في قياس تلك العوامل أي أن معامل الصدق العاملي هو عامل مشترك بين هذه الاختبارات وأي اختبار آخر أو مجموعة اختبارات أخرى تقيس الشيء نفسه، فإذا أردنا تحديد الصدق العاملي لاختبار نقيس مفهومًا ما ونحلل هذا المفهوم إلى عناصره الأولية، فالقدرة العقلية يمكن أن تحلل إلى عدة عوامل كالعامل اللغوي والعامل العددي، والعامل المكاني ثم يؤلف اختبار يقيس جميع العوامل التي تتألف منها ويعطي للتلاميذ فإذا كان هناك ارتباط بين نتائج هذا الفحص ونتائج فحوص أخرى تقيس هذه العوامل ومتصفة بالصدق، نقول أن الاختبار الجديد يتصف بالصدق العاملي

(مفلح، 2005، 118).

وتجدر الإشارة إلى أن الصدق العاملي يشير إلى ترابط الاختبار مع كل ما هو مشترك بين مجموعة من الاختبارات، أو غيرها من الأدوات "اللاختبارية". وهذا يعني أن التقديرات الرتبوية

والمقاييس المحكية الأخرى يمكن استخدامها، إضافة لاختبارات أخرى، في استطلاع البنية
العاملية لاختبار معين وتحديد السمات المشتركة التي تقيسها .

(امطانيوس، 2015، 194)، (Anastasi & Uvbina, 1997, P.128).

3- الصدق التقاربي والتباعدي (التمييزي):

ويقترح جيليكوس Gulliksen: إجراء عدة خطوات علمية موضوعية لقياس معامل صدق
المفهوم بالطريقة التقاربية أو التباعدية (التمييزية) يذكر ضمنها إجراء طريقة الاختبار الذي
يقيس قدرة معينة ثم تنظم برامج لتدريب هذه القدرات لدى من طبق عليهم الاختبار ثم إعادة
إجراء الاختبار بعد انتهاء التدريب فإذا ظهر فرق دال كان الاختبار صادقاً، ومن أهم طرق
حساب الصدق التقاربي والتمييزي إجراء عدة اختبارات تقيس منطقة واحدة من السلوك ثم عقد
المقارنات بين درجات الأفراد على هذه الاختبارات من بينها الاختبار الجديد، فإذا كان هناك
معامل ارتباط (بطريقة الارتباط المتعدد) كبير مع الاختبارات الأخرى الصادقة، كان الاختبار
الجديد صادقاً (مجدي حبيب، 1996، 310).

II- 2. الثبات:

مفهوم الثبات: هو أحد المفاهيم الأكثر أهمية في المنهجية السيكولوجية، إلى جانب مفهوم الصدق، وتطبيق هذه المفاهيم على درجات الاختبارات السيكولوجية وغيرها من المقاييس (علام، 2016، 438).

وتعرف معظم القواميس "الثبات" من حيث جدارة الثقة أو الاعتماد، أو أن تكون هناك درجة عالية من الثقة في شيء ما، والثبات في سياق القياس النفسي يهتم بدرجة ما بنفس هذه العوامل، ولكنه يمتد لمفاهيم مثل الاتساق والاستقرار والتكافؤ. وبعبارة بسيطة - في سياق الثبات - يشير الثبات إلى اتساق أو دقة، أو استقراء نتائج التقييم (Aera, et all, 1999, 31).

II- 1-2. نظرية الثبات:

يستخدم الثبات بالمعنى العام للدلالة على مدى اعتماد الفروق في درجات الاختبار على الفروق الحقيقية في السمة أو الخاصية المقيسة من جهة وأخطاء القياس العشوائية أو الناجمة عن الصدفة من جهة أخرى، ومن المعلوم أن الاختبار في العلوم النفسية والتربوية يستحيل عده مقياساً ثابتاً مطلقاً كالمقاييس الفيزيائية (المتر، كغ، م²، ... إلخ) (أمطانيوس، 2016، 203). وأن الدرجة التي يحصل عليها الفرد في مقياس معين تسمى الدرجة الملاحظة Obtained score، غير أن هذه الدرجة تكون في كثير من الأحيان مشوبة بأخطاء القياس،

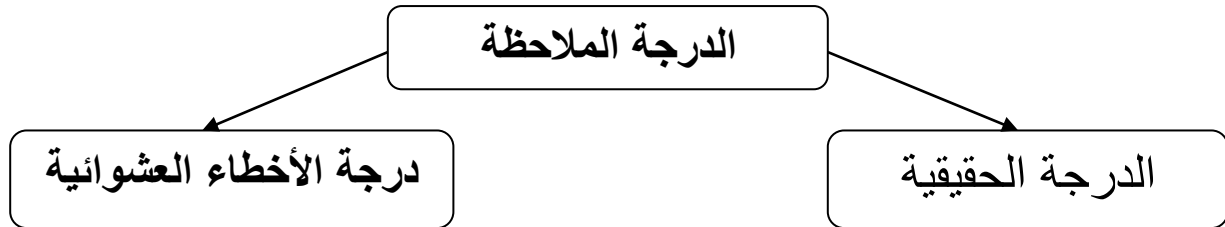
التي إذا أمكن تحديد مقدارها؛ فإننا نحصل على درجة الخطأ Error score، وإذا طرحنا درجة الخطأ من الدرجة الملاحظة فإننا نحصل على درجة الفرد خالية من الأخطاء العشوائية، وهذه تسمى الدرجة الحقيقية True score، أي أن:

الدرجة الحقيقية = الدرجة الملاحظة - درجة الأخطاء العشوائية (علام، 2000، 131).

شكل (09):

مكونات الدرجة الملاحظة.

مكونات الدرجة الملاحظة.



وهذا يعني أن الدرجة التي تشير إلى ثبات الاختبار لا تعبر عن الأداء الحقيقي للفرد، بل تمثل الأداء الحقيقي مضافا إليه الشوائب المؤثرة على الدرجة، أي أن درجة الفرد على الاختبار تعبر عن التباين الحقيقي للفرد وتباين الخطأ، ولذلك يؤكد جيلفورد على ضرورة حساب ثبات الاختبار كي تحدد الدرجة الحقيقية أو التباين الحقيقي للاختبار، فمعمل الثبات يوضح نسبة التباين الحقيقي في الدرجة المحسوبة على الاختبار (بوسالم، 2014، 80).

فالقائمة العددية لمعامل ثبات الاختبار تتوقف على تباين الدرجات في هذا الاختبار، على أن يرجع هذا التباين إلى الفروق الحقيقية بين الأفراد في الصفة التي يقيسها الاختبار. ولا يكون الاختبار ثابتاً عندما يكون خطأ التباين فيه كبيراً، ونعود فنقول أن المشكلة الأساسية في تحديد ثبات عمليات القياس هي مشكلة تحديد التباين الحقيقي بين الأفراد وتحديد مدى خطأ التباين الذي نقبله (أحمد، 1960، 223).

مصادر أخطاء القياس:

يرجع التباين في درجات الاختبار إلى أحد أو بعض أو كثير من العوامل والتي نذكر أهمها:

- 1- أخطاء معاينات المحتوى وتعد عادة أكبر مصادر الخطأ في درجات الاختبار.
- 2- أخطاء القياس التي ترجع إلى معاينات الوقت تعكس التذبذبات العشوائية في الأداء من وقت إلى آخر، وتحد من قدرتنا على تعميم درجات الاختبار عبر مواقف مختلفة (علام، 2013، 166-167).
- 3- الأخطاء الناجمة عن شروط التطبيق إذا لم تراعى هذه الشروط كما تنص عليها تعليمات الاختبار، أو روعيت بصورة غير كافية فقد تؤدي إلى نتائج غير صحيحة.
- 4- الأخطاء الناجمة عن المفحوص نفسه وهذه الأخطاء كثيرة ومتنوعة (أمطانيوس، 2016، 204).

II-2-2. طرق الثبات:

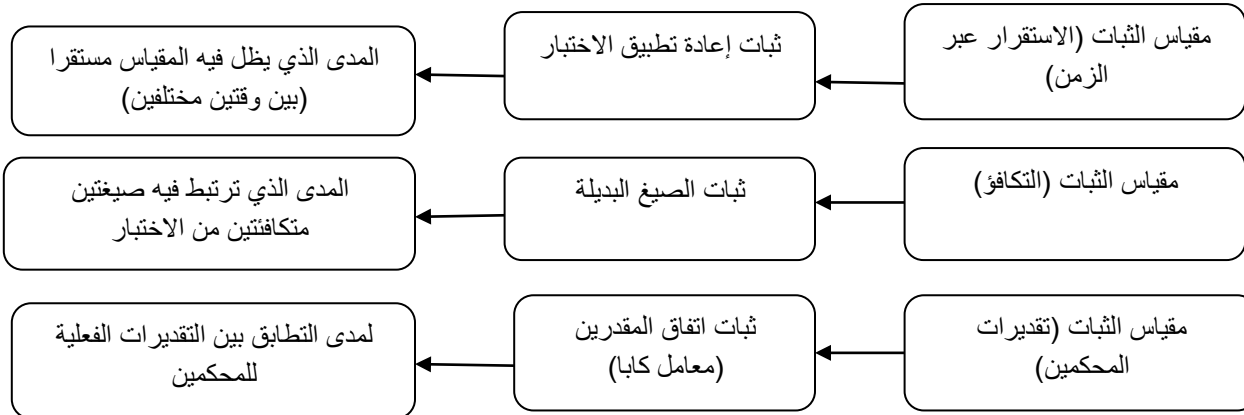
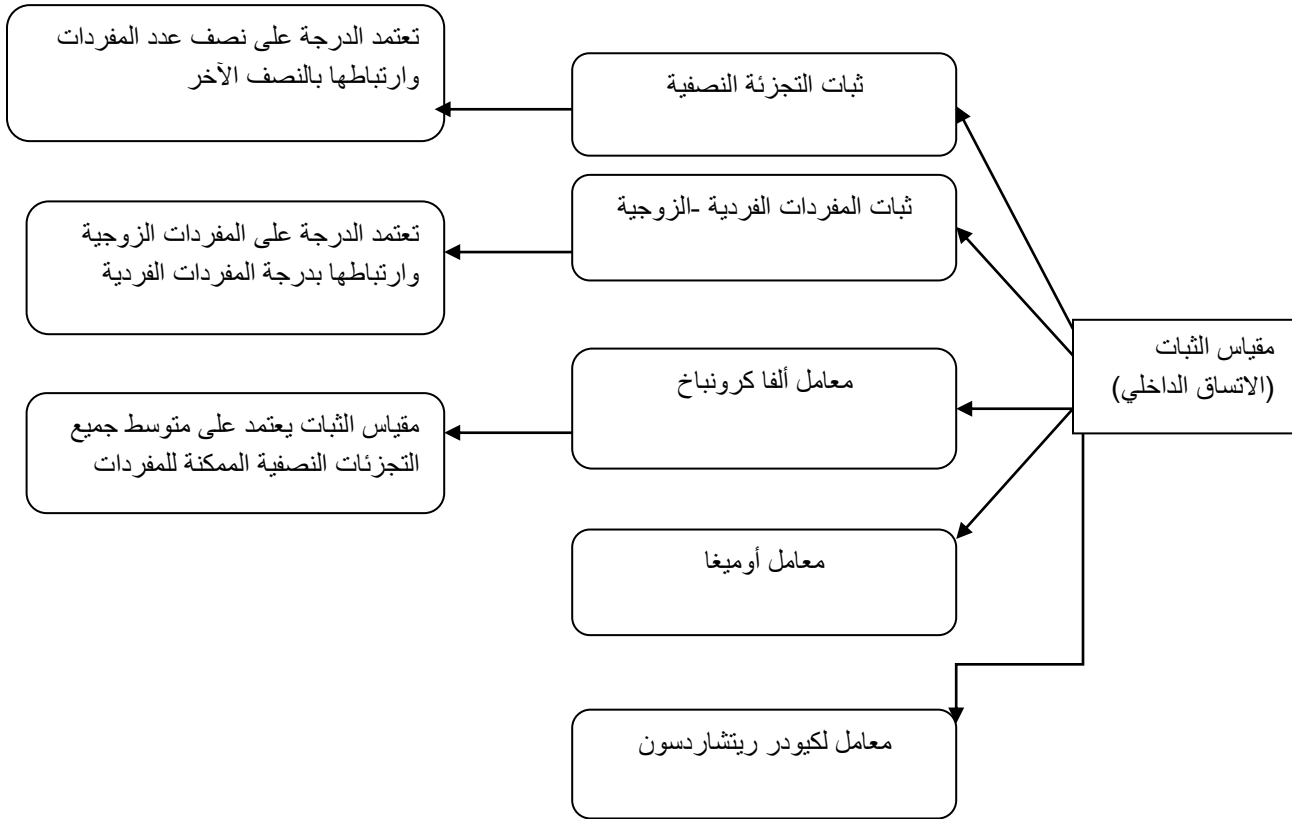
يمكن تصنيف معاملات الثبات في ثلاث أقسام عريضة (Aera, et all, 1999).

وهي تشمل:

- 1- ثبات الاتساق الداخلي: وهي معاملات اشتقت من تطبيق واحد للاختبار.
- 2- ثبات إعادة الاختبار: معاملات اشتقت من تطبيق نفس الاختبار في طرفين مختلفين.
- 3- ثبات الصيغ المتكافئة: معاملات اعتمدت على تطبيق صيغتين مختلفتين للاختبار.
- 4- ثبات تقدير المحكمين: يستخدم عندما يتضمن تقدير الدرجات قدرا كبيرا من الأحكام الذاتية (علام، 2013، 169).

ويمكن تمثيل هذا التصنيف في الشكل الآتي:

أنواع الثبات.



- الاتساق الداخلي للاختبار Internal reliability:

يشير الاتساق الداخلي إلى مدى اتساق جميع مفردات الاختبار في قياسها لمفهوم معين، فإذا كان الاختبار متسقاً فإن أية مجموعة من مفرداته يمكن اختيارها وسوف تقدم قياساً متسقاً كأية مجموعة أخرى من المفردات المأخوذة منه.

وأحدى الطرق التقليدية لحساب الاتساق الداخلي هي حساب الدرجات في نصف عدد المفردات في الاختبار وإيجاد ارتباطها بدرجات نفس الأفراد في بقية الاختبار، وتشكل هذه الإجراءات أساساً لمقاييس متعددة للاتساق الداخلي منها:

- ثبات التجزئة النصفية (Split-Half reliability):

وتعتمد هذه الطريقة على تطبيق الاختبار على عينة ممثلة مرة واحدة (محسن، 2011، 106). ومن ثم تقسيم الاختبار إلى نصفين متكافئين يتم تقدير درجات كل منهما على حدى (أبو هاشم، 2006، 7).

وتتضمن التجزئة النصفية حساب الدرجات لكل من نصفي عدد المفردات والارتباط بين هذين المجموعتين من الدرجات (علام، 2016، 705).

ويتم تقدير درجات كل من نصفي الاختبار كما لو كان كل منهما اختباراً منفصلاً (علام، 2002، 155)، ويطلق على الثبات المحسوب بهذه الطريقة اسم معامل الاتساق الداخلي (النمر، 2007، 84).

وهذا يعني أن الاختبار يعطى بكامله في جلسة واحدة وفي حدود زمنية واحدة، إلا أنه بعد التطبيق تشتق درجتان منفصلتان، واحدة من تصحيح الفقرات ذات الترتيب الفردي، والقسم الثاني من الفقرات ذات الترتيب الزوجي (أبو هاشم، 2006، 7). ولكن المشكلة تكمن في أن معامل الارتباط هو بين نصفي طول الاختبار، ولهذا سوف يكون أقل مما لو استخدمت على الاختبار بكامله (سعد، 1999، 169).

ولهذا يعاب على هذه الطريقة أنها تقيس ثبات نصف الاختبار ولا تقيس ثبات الاختبار كله (منسي، 2007، 84).

ولعلاج هذا العيب هناك طرق لمعالجة المشكلة والتحقق من ثبات الاختبار:

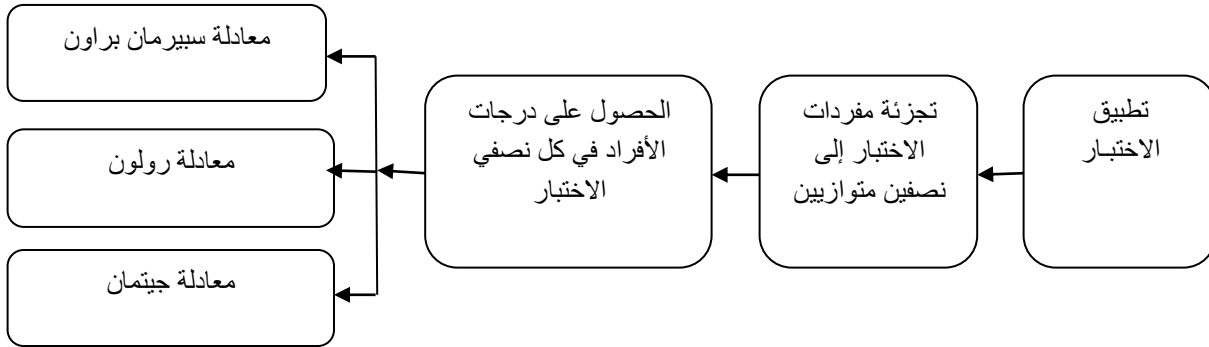
- ففي الحالة الأولى حيث عدم تجانس درجات النصفين نستعمل لحساب الثبات إما طريقة رولان أو طريقة جيتمان.

- أما إذا كانت درجات النصفين متجانسة فيتم تطبيق معادلة سبيرمان وبراون التصحيحية بعد حساب معامل الارتباط بيرسون بين درجات النصفين (بوسالم، 2012، 90).

ولهذه الطرق أو المعادلات ميزات وتطبيقات مختلفة وكذلك توجد لها نواحي القصور كما سنوضح لاحقاً.

شكل (11):

إجراءات تقدير قيمة معامل الثبات بطريقة التجزئة النصفية للاختبار.



(بوسالم، 2012، 91).

ونعرض هنا لأكثر المعادلات الإحصائية شيوعاً التي تقوم بحساب معامل ثبات القسمة النصفية على النحو الآتي:

أ- معادلة سبيرمان-براون: Spearman-Brown Formule

تستخدم هذه الطريقة في حالة النصفين المتكافئين (المتجانسين) للاختبار.

$$r = \frac{2r}{1+r} \text{ ، حيث } r$$

ويعاب على هذه الطريقة أنها تفترض تساوي الانحراف المعياري لكل من النصفين، فإذا لم

يتساوى التباين بين الجزئين لجأنا إلى طريقة أخرى، جيتمان أو رولون (الأنصاري، 2000،

(124).

ب- معادلة جيتمان: (فلانجان) (Guttman):

وتستخدم هذه الطريقة في حالة تساوي أو عدم تساوي تباين درجات الأفراد لجزئي الاختبار

وهي:

$$R_G = 2 \left[1 - \frac{S_1^2 + S_2^2}{S^2} \right]$$

(السيد، 1978، 387).

ومن عيوب هذه الطريقة أنها لا تصلح مع الاختبارات التي لا يمكن تقسيمها إلى نصفين

متجانسين.

ج- معادلة رولون: (Roulon):

يرى رولون أن تباين درجات الأفراد على الاختبار ككل يرجع إلى تباين حقيقي في

مستويات القدرة المقاسة، كما يرجع إلى أخطاء في القياس (ربيع، 2008، 92).

وفي ما يلي توضيح لمعادلة رولون:

$$R_r = 1 - \frac{S_{def}^2}{S_T^2}$$

حيث:

S^2_{def} : تباين فروق درجات النصفين

S_T^2 : تباين درجات الاختبار ككل.

الثبات بطريقة كيودر ريتشاردسون Kyder- Richardson:

فقد عرض كيودر وريتشاردسون في مقالتهما الأصلية (1937) صيغا متعددة لتقدير الثبات،

والصيغة الأكثر شيوعا واستخداما تعرف بأنها صيغة كيودر - ريتشاردسون (KR-20)، ويمكن

تطبيق هذه الصيغة عندما تكون درجات مفردات الاختبار ثنائية، أي إجابة (صح) أو (خطأ)

مثلا (Keith & Reynolds, 1990, 175)

$$K - R_{20} = \frac{K}{K - 1} \left[\frac{S_x^2 - \sum Pq}{S_x^2} \right]$$

حيث:

K: عدد البنود

S_x^2 : تباين درجات الاختبار الكلي (مربع الانحراف المعياري)

P: نسبة الاستجابات الصحيحة لبند واحد.

q: نسبة الاستجابات الخاطئة للبند ذاته.

وثمة معادلة أخرى (Kyder- Richardson 21) أسهل حسابيا لكنها تتطلب افتراض أن

جميع البنود في الاختبار ذات صعوبة متماثلة، وهذا الافتراض غير واقعي في الغالب

$$K - R_{21} = \frac{KS_x^2 - \bar{X}(K - \bar{X})}{S_x^2(K - 1)}$$

وتعد هذه الطريقة إلى حد بعيد، الأقل استهلاكاً للوقت من جميع أساليب تقدير الثبات (الحسيني، 2013، 308).

- الثبات بطريقة ألفا كرونباخ:

توجد مشكلة تتعلق بثبات التجزئة النصفية، فقيمتها تعتمد على أي المفردات تم اختيارها في كل من النصفين، فالثبات بتجزئة المفردات إلى فردية وزوجية لن يكون هو نفس الثبات الذي نحصل عليه بمقارنة نصف عدد المفردات بالنصف الآخر مثلاً. ويوجد حل واضح وهو أن نحسب ثبات كل تجزئة نصفية ممكنة، وذلك استناداً إلى اختيار جميع التوافيق الممكنة للمفردات في كل من النصفين، ومن ثم توجد متوسط لها جميعاً، ويعرف ذلك "بمعامل ألفا" "Coefficient Alpha" الذي اشتقه كرونباخ "Cronbach" صورة عامة لمعادلة الثبات من صيغة كيودر-ريتشاردسون وهي:

$$r_{\alpha k} = \frac{m}{m-1} \left(1 - \frac{\sum_{i=1}^m S_i^2}{S_T^2} \right)$$

حيث:

m: عدد المفردات.

S_T^2 : تباين الاختبار ككل.

S_1^2 : تباين درجات كل مفردة (علام، 2016، 707).

- الثبات بطريقة أوميغا:

يمكن حساب معامل الثبات لكل متغير كامن من المقياس باستخدام معامل الثبات المركب (CR) أو معامل رو (ρ) أو معامل أوميغا Ω (Omega) لكارل جورسكوج (K. Joreskog, 1970), ويعتبر معامل الثبات المركب أحد المعادلات التي تمكن من تقدير مدى دقة عدد المؤشرات في قياس بعد أو مفهوم (عامل كامن) (تيغزة، 2012، 37).

وأشار كل من (Raykov, 2004) و (Hancock & Mueller, 2011) أن معامل الثبات المركب أو معامل "Ro" أفضل من معامل (ألفا) في تقدير ثبات النماذج العاملية، وأنه لا يقيس ثبات المؤشرات إذا كانت لا تتكون من عامل واحد بمعنى أنه حساس لأحادية البعد (Kline, 2016, 313). ويعتمد في تقدير الثبات المركب على مساهمات كل متغير مقياس في تفسير عامله مع الأخذ في الحسبان قيم الأخطاء، ويتم حساب معامل الثبات المركب عن طريق المعادلة المذكورة في (Bacon et al, 1995, 396):

$$\rho = \frac{(\sum_i^n \lambda)^2}{(\sum_i^n \lambda)^2 + \sum Var(\varepsilon_i)}$$

حيث:

ρ : معامل الثبات المركب (CR) للعامل.

λ : التشبع المعياري للفقرة على بعدها.

$Var(\varepsilon_i)$: تباين خطأ الفقرات المعياري، حيث $(\varepsilon_i = 1 - \lambda_i^2)$.

كما يمكن حساب معامل الثبات المركب كذلك باستخدام إضافة برمجية لبرنامج (SPSS) مقترحة

من طرف (Hayes, 2020) ويتم تحميلها من الموقع:

(<https://www.afhayes.com/public/omega.zip>)

وتجدر الإشارة إلى أن القيم العالية لمعامل الثبات المركب تدل على درجات عالية من الثبات وتعتبر مقبولة إذا كانت أكبر من 0.60 (Diamantopoulos & Siguaw, 2000). (Hair & al, 2017, 92).

وأخيراً، من المناسب مراعاة الحيلة، فكثير من المراجع يؤكد أن الثبات بطريقة الاتساق الداخلي يعد مظهراً مهماً وأساسياً للمقياس الجيد. ويعد أمراً صحيحاً حتى حينه، وإذا حاولنا قياس أحد المفاهيم أو التكوينات الفرضية المعرفة جيداً (الذكاء، الانبساط، الحاجة للإنجاز)، فإن الثبات بالاتساق الداخلي عندئذ ينبغي أن يكون مرتفعاً قدر المستطاع، ومن جهة أخرى، ربما لا يكون قياس تلك المفاهيم السيكولوجية المنقحة هو الهدف الأساسي للمقياس (علام، 2016، 397-398).

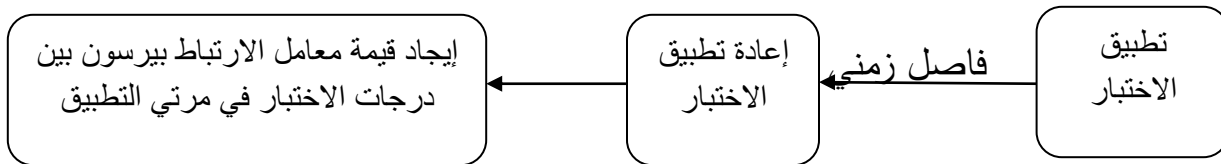
- الاستقرار عبر الزمن: Stability over time

تقوم الفكرة الأساسية في هذه الطريقة على إجراء الاختبار على نفس المجموعة من الأشخاص وفي ظروف مماثلة بعد مضي فترة زمنية بين الإجراء أو القياس الأول والإجراء أو القياس الثاني، ثم نحسب معامل الارتباط بين الإجراء الأول والثاني وهو ما نسميه معامل الثبات (ربيع، 1993، 71).

والشكل التخطيطي يوضح ذلك: شكل (12):

إجراءات تقدير قيمة معامل الاستقرار.

إجراءات تقدير قيمة معامل الاستقرار.



(علام، 2000، 149).

كما أن هذا المعامل يتأثر بدرجة أساسية لخطأ القياس الذي يرجع إلى معايينات الزمن (الوقت)، أي إذا كانت الفترة قصيرة بين التطبيق وإعادة التطبيق فإن الذاكرة تلعب دورها وهذا يرفع من معدل الثبات (سبع، 2008، 229).

أما إذا كانت الفترة الزمنية الفاصلة كبيرة بحيث تبلغ عدة أشهر، فإننا نتوقع تدخل كثير من العوامل غير المرغوبة التي ربما تفسر اختلاف الدرجات في مرتتي التطبيق فإننا سوف نلاحظ أن قيمة معامل الاستقرار تقل بزيادة هذه المدة غير أن هذه المدة ليست سببا في اختلاف الدرجات في مرتتي التطبيق، إنما عامل المران أو الخبرة التي يكتسبها الفرد خلال هذه المدة (علام، 2000، 149).

ويشير الخبراء إلى أن الفترة الزمنية بين الأداتين ينبغي أن لا تقل عن أسبوعين ولا تزيد عن الستة شهور، ويحدد مضمون الاختبار وخصائص العينات المستخدمة في الدراسة مدة الفترة الزمنية. ن تطول الفترة الزمنية بين الأداءين، أما من حيث خصائص العينات فقد تكون في مرحلة عمرية سريعة التأثير بمتغيرات النمو أو التعلم مما يؤثر في ثبات الاختبار إذا أعيد بعد فترة طويلة (معمرية، 2007، 173-174).

ونتيجة لذلك فإن الاختبارات التي لا تتأثر تأثيراً كبيراً بهذه التأثيرات الانتقالية تكون هي فقط المناسبة لهذه الطريقة من طرق تقدير الثبات.

- ثبات الصيغ المتكافئة: Alternate-from reliability

للتخلص من أثر التدريب والتذكر أو النسيان لدى العينة جراء الطريقة السابقة (إعادة التطبيق) فإننا نلجأ إلى بناء صيغتين متكافئتين أو متوازيتين للاختبار من حيث المحتوى والصعوبة وغير ذلك من خصائص المواءمة والتطبيق، وتطبق الصيغتان بعد ذلك على نفس مجموعة الأفراد وبعد فاصل زمني قصير (آني) أو أيام (مؤجل) وبعد التصحيح نجد معامل الارتباط بين نتائج هذين الفحصين المتكافئين ويسمى معامل الارتباط هذا بمعامل الثبات أو معامل التكافؤ (Norman, 1976, 110) coefficient of equivalence.

وثبات الصيغ المتكافئة الذي يعتمد على التطبيق الآني يكون حساساً بدرجة أساسية لأخطاء القياس التي تعزى إلى معايير المحتوى، وثبات الصيغ المتكافئة الذي يعتمد على

التطبيق المؤجل يكون حساسا لأخطاء القياس التي تعزى إلى كل من معايير المحتوى ومعايير الوقت (علام، 2016، 172).

وتمتاز هذه الطريقة عن الطريقة السابقة بأنها تلغي أثر التدريب والتذكر وكذلك النسيان، ومن عيوبها هو صعوبة إيجاد صورتين متكافئتين أو متوازيتين للاختبار، لذا فإن بناء صيغ متكافئة بحيث تكون متكافئة بالفعل يستغرق وقتا طويلا، وكثير من مطوري هذا الاختبار لا يلجؤون لهذا الخيار، غير أنه من المرغوب فيه أحيانا أن يكون أكثر من صيغة واحدة لاختبار معين، وعندما تتوفر صيغ متعددة، يكون ثبات الصيغ المتكافئة موضع اهتمام (Dawine, 1967, 84).

- مقياس الثبات (تقدير المحكمين): Inter-Rater Reliability :

إذا اعتمد تقدير درجات اختبار على أحكام ذاتية، فمن المهم تقويم درجة الاتفاق عندما يقوم أكثر من فرد بتقدير الدرجات، ويطلق على ذلك "ثبات تقدير المحكمين"، وحساب ثبات تقدير المحكمين يعد عملية مباشرة إلى حد ما، إذ أن الاختبار يطبق في مرة واحدة، ويقدر فردان كل على حدى درجات الاختبار ومن ثم يتم حساب الارتباط بين الدرجات التي نحصل عليها منها. ويعكس هذا التقدير للثبات بدرجة أساسية الفروق التي ترجع إلى الأفراد الذين يقدرون درجات الاختبار، ويغفل بدرجة كبيرة الأخطاء الناجمة عن معايير المحتوى أو الوقت، بالإضافة إلى المدخل الارتباطي التي يعين فيها فردان (محكمان) نفس الدرجات لأداء الطلاب،

ويطلق عادة على هذا المدخل "اتفاق المحكمين"، أو "النسبة المئوية للاتفاق"، وطريقة حساب ذلك باستخدام الصيغة الآتية:

$$\text{الاتفاق بين المحكمين} = \frac{\text{عدد الحالات التي لها نفس الدرجات} \times 100}{\text{العدد الكلي للحالات}}$$

ويفضل كثير من المؤلفين "معامل كابا كوهين" على النسبة المئوية المعتادة للاتفاق عند تحليل البيانات الوصفية، فمعامل كابا يعد مقياساً أكثر حساسية للاتفاق، لأنه يأخذ درجة الاتفاق المتوقعة بالصدفة بعين الاعتبار (Hays, 1994, 178).

ويتوافر أيضاً معامل كابا الموزون، وهو يناسب البيانات من المستوى الرتبي، ويأخذ اختلاف التقديرات بعين الاعتبار، ويستخدم أيضاً معامل كابا في أي حالة يكون فيها الاتفاق في التصنيف موضع اهتمام، والحالات التي ربما يحدث فيها ذلك يمكن أن تشمل ما إذا كان الاختبار يطبق في مرتين مختلفتين تفصل بينهما مدة زمنية (علام، 2016، 179).

ويستخدم في تصنيف الأفراد في مجموعات تشخيصية (أو مجموعات أخرى، مثل توظيف أو رفض أفراد). وفي هذه الحالة سوف يصنف كل فرد أو يعين في مجموعة باستخدام درجات الاختبار في كل من المرتين، وتقارن درجة الاتفاق بين المرتين باستخدام كابا، ويمكننا أيضاً استخدام اختبارين مختلفين بتطبيقهما على نفس المجموعة من الأفراد في نفس الوقت، وتصنيف الأفراد باستخدام كل مجموعة من الدرجات على حدى، ومن ثم حساب الاتفاق بين

درجات الاختبارين في التصنيف، وفي بعض الاختبارات يكون ثبات تقدير المحكمين موضع اهتمام قليل، فمثلا في اختبار على اختبارات من متعدد، أو صواب أو خطأ، يكون تقدير الدرجات مباشرا والتقدير اليقظ للدرجات ينبغي أن ينتج عنه درجات ثابتة ودقيقة (علام، 2016، 180).

II- 3-2. العلاقة بين الصدق والثبات:

يعد الصدق خاصية أكثر أهمية وشمولا من الثبات كما أنه أكثر صعوبة في التحديد. وتذكر أغلب الدراسات البحثية المنشورة بيانات الثبات أكثر بكثير من بيانات الصدق. ويعزى هذا إلى كون أن الصدق لا يمكن الحصول عليه بشكل مباشر كالثبات، وتقدير أداة القياس يتطلب حكما ذاتيا غير موضوعي بدرجة كبيرة، وعلى المرء أن يجيب على أسئلة حول ملاءمة محتوى الاختبار، وكفاية المعايير، وتعريفات السمات الإنسانية ومواصفات المجال السلوكي، والنظرية المتعلقة بمحتوى الاختبار، وما إلى ذلك. فجميع هذه القضايا تشمل الحكم وجمع البيانات من عديد المصادر من الناحية الأخرى، يمكن التحقق من الثبات بشكل مباشر من بيانات الاختبار، ولا حاجة لبيانات خارجية. إن القضايا الأساسية للثبات ثلاث بسهولة التحليل الرياضي والاستنتاجات المنطقية حول مقدار الخطأ الذي يمكن التعبير عنه بصورة رياضية (الحسيني، 2013، 318).

ولهذا يعتمد الصدق جزئياً على الثبات، ويمثل الثبات شرطاً ضرورياً وغير كاف للصدق. ومن الواضح أن المقياس الذي يعطي نتائج غير متسقة لا يمكن أن يعطي نتائج صادقة في حين أن المقياس الذي يعطي نتائج صادقة لا بد أن يعطي نتائج متسقة ويتمتع بصفة الثبات، وبطبيعة الحال فإن الزيادة في معامل الثبات قد تؤدي إلى الزيادة في معامل الصدق ولكنها لا تحتم هذه الزيادة (أمطانيوس، 2016، 207).

- 2-4. العوامل المؤثرة في الثبات:

بعد أن تعرضنا للطرق المختلفة التي يحتاج إليها الباحثون، ومجال استخدام كل الطرق لحساب ثبات الاختبار، ينبغي أن نبين كذلك العوامل التي تؤثر في ثبات الاختبار حتى يتمكن الباحث أو أخصائي القياس من التحكم فيها، وخفض تدخلها في الثبات.

1- طول الاختبار (عدد البنود): فكلما ازداد عدد البنود، كلما ارتفع معامل الثبات، وهذا يعود إلى أن الاختبار الأطول يزيد تمثيله للمجتمع الأصلي، أي لمنطقة السلوك المراد قياسه، لذلك يمكن رفع معامل الثبات بإطالة الاختبار بالمعادلة التالية:

$$\text{معامل الثبات المراد الوصول إليه} = \text{ن} \times \text{معامل الثبات الحالي}$$

$$1 + (1 - \text{ن}) \times \text{معامل الثبات الحالي}$$

(الانصاري، 2000، 132).

- 2- زمن الاختبار: يضل معامل الثبات مرتفعا بزيادة الوقت الذي يستغرقه إجراء الاختبار إلى درجة معينة، ثم ينخفض الثبات بطول المدة، وهذه الدرجة تختلف من اختبار إلى آخر.
- 3- صعوبة البنود: تعطي الاختبارات الصعبة جدا أو السهلة جدا، درجات ضعيفة في ثباتها، تؤدي إلى انتشار محدود في الدرجات، وتظهر فروقا طفيفة بين الأفراد، وعموما إذا شاء واضع المقياس أن يزيد من ثباته عليه أن يضع أسئلة يتراوح مدى صعوبتها بين 0.30 و0.70، بحيث يقترب معامل صعوبة الاختبار ككل من 0.50.
- 4- تجانس العينة: يقل ثبات الاختبار بزيادة تجانس العينة المطبق عليها، ويرتفع الثبات بزيادة تباين درجات أفراد العينة عليه (الانصاري، 2000، 133).
- 5- الموضوعية في التصحيح: لكي يعطي المقياس نتائج ثابتة يجب أن لا تختلف الدرجات من مصحح إلى آخر ومن وقت إلى آخر، ومن المعلوم أن الاختبارات الموضوعية عموما تحقق قدرا عاليا من الثبات، بخلاف الاختبارات المقالية التي تفسح المجال لتدخل العوامل الذاتية للمصحح (أطانيوس، 2016، 233).

II-3. البنية العاملية:

تمهيد:

يسعى أسلوب التحليل العامل إلى استخلاص العوامل من المؤشرات، بحيث يكون العامل الأول هو أكثرها ارتباطاً بالمؤشرات أو أكثرها تفسيراً للتباين المشترك يليه العامل الثاني وهكذا، ويجب أن يكون في كل عامل عدد كبير من المؤشرات، ومن أهم معايير تحديد عدد العوامل المستخلصة هو كمية التباين التي يساهم بها العامل.

II-3-1. النموذج العامل:

في التحليل العامل يفترض الباحث أن للمفهوم بنية، ويفترض أن هذه البنية تتألف من مكون واحد أو مكونين، أو عدة مكونات، فإذا كان المفهوم يتألف من مكون واحد يدعى مفهوم متجانس، أو مفهوم وحيد البعد أو العامل، كأن يفترض الباحث مثلاً أن مفهوم الذكاء العاطفي ينطوي على بعد واحد، بمعنى أن المؤشرات أو المتغيرات التي تساهم في تركيب مفهوم الذكاء العاطفي تؤلف فيما بينها مساحة مشتركة من الدلالة تمثل مفهوم الذكاء العاطفي، وإذا تألف المفهوم موضوع التحلي من مكونين أو أكثر فيدعى مفهوم متعدد الأبعاد أو العوامل، ويوجد نوعين من النماذج العاملية هما:

- النموذج العامل الاستكشافي.
- النموذج العامل التوكيدي (غير الهرمي، الهرمي) (تيغزة، 2012، 155).

II-3-2. النموذج العاملي الاستكشافي: Exploratory factor models

يسعى أسلوب التحليل العاملي الاستكشافي إلى تقليص المتغيرات أو العوامل في عدد أقل، بحيث يكون لكل عامل من هذه العوامل مؤشرات تشبعت على هذا العامل، ومجموع مربعات ارتباط المؤشرات بالعامل تمثل مقدار التباين المستخلص، ويمكن من خلال هذه المؤشرات إعطاء تفسير لهذا العامل، وقد ظهر هذا الأسلوب في دراسة المقاييس النفسية بحيث يمكن إرجاع مجموعة معينة من المؤشرات إلى عامل معين وأخرى إلى عامل آخر وهكذا، وترتكز فكرة التحليل العاملي الاستكشافي على استخلاص مجموعة من العوامل مرتبطة بالمؤشرات الأصلية، بحيث تفسر هذه العوامل أكبر نسبة ممكنة من التباين في هذه المؤشرات (Begg, 2002, 167).

II-3-3. افتراضات التحليل العاملي الاستكشافي:

II-3-3-1. حجم العينة:

إذا كانت العينة تحتوي على 50 مشاهدة أو أقل لا يمكن إجراء التحليل العاملي، ويجب أن لا يقل حجم العينة عن 100 فرد، والأفضل أن يكون حجم العينة عشرة أضعاف عدد الفترات (Cohen, 1990, 198). ويبرر هذا الإجراء إلى أنه كلما كبر حجم العينة قلت

الأخطاء العشوائية، وازدادت الثقة في النتائج، ومن الضروري توفر ذلك في التحليل العاملي نظرا لحساسية الموضوع المراد دراسته وتعميم نتائج الدراسة التي قد يعتمد عليها الكثير من الباحثين وأنه مسلم بها (القهوجي، 2018، 168).

II-3-1-2. استقلال المفردات:

حيث يتطلب من المفحوص الإجابة الفردية على البيانات وليس الجماعية أو التشاورية، وأي انتهاك لهذا الفرض يؤدي إلى زيادة مقدار خطأ القياس مما يؤثر على اختبار الفرضية الصفرية بحيث ترفض بشكل أكبر أو أقل مما تم تحديده من قبل الباحث (القهوجي، 2018، 168-169).

II-3-1-3. القيم الشاذة والمتطرفة:

يجب التأكد من خلو البيانات من القيم الشاذة أو القيم المتطرفة، حيث تعمل على التقليل من قيم معاملات الارتباط بين المتغيرات.

II-3-1-4. مقياس ملائمة البيانات (KMO).

ويشير إلى مدى ملائمة البيانات للتحليل العاملي، ويجب أن يكون أكبر من 0.6 وبديل هذا المفهوم على مقدار الارتباط الجزئي بين المتغيرات (Bourderau, 1999, 169).

II-3-1-5. تدوير العوامل:

يتم إعادة تحديد مواضع المؤشرات وصولاً إلى أقصى درجة من الثبات، ويهدف إلى تحسين وضع العوامل المستخرجة مما يقلل من التشبعات القليلة ويكبر من التشبعات الكبيرة للعوامل وبالتالي الحصول على أكبر قدر ممكن من التباين المفسر (القهوجي، 2018، 169).

II-3-2. النموذج العاملي التوكيدي: Confirmatory factor models

النموذج العاملي التوكيدي على النقيض من النموذج العاملي الاستكشافي يقوم على إطار نظري واضح يستطيع الباحث بموجبه أن يقترح نموذجاً عاملياً ينطوي على المعلومات التالية قبل الانتقال إلى اختبار نموده باستعمال التحليل العاملي التوكيدي:

- أ- عدد العوامل.
- ب- طبيعة العوامل ومسمياتها.
- ج- لكل عامل مجموعة من المؤشرات المقاسة الخاصة به التي تقيسه، وتشبع عليه دون العوامل الأخرى.
- د- وجود علاقات ارتباطية بين العوامل عند تصور عدم استقلالها، أو عدم وجودها.
- هـ- تبيان أخطاء القياس (خطأ التباين لكل مؤشر مقاس) ما إذا كانت مستقلة، أو غير مستقلة بحيث تكون بعض الأخطاء مرتبطة (تغزة، 2012، 158).

II-3-3. أنواع النموذج العاملي التوكيدي:

يوجد نوعان من النموذج العاملي التوكيدي حسب عدد المراحل وهما:

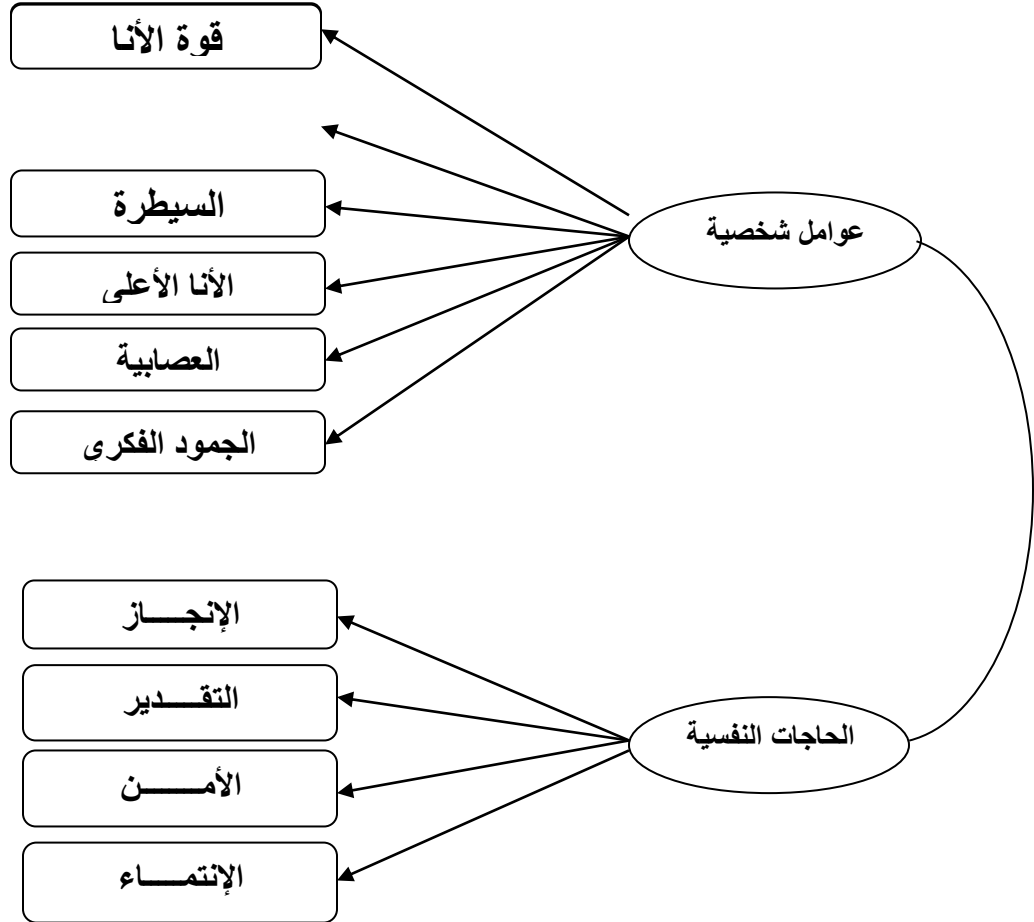
- النموذج العاملي التوكيدي من الدرجة الأولى:

سمي بذلك لوجود عامل كامن واحد، أو عدة عوامل كامنة ترتبط بالمؤشرات الخاصة بها، وعدم احتوائها على عامل عام تتدرج العوامل الأخرى تحته، وقد يكون أحادي البعد، ذو مفهوم واحد مشترك المؤشرات أو الفقرات مع هذا المفهوم، بحيث تلتقي هذه المؤشرات عند ذلك المفهوم، ولذلك يسمى تكوين فرضي وحيد البعد أو متجانس، ويجب أن لا تقل عدد المؤشرات للعامل عن ثلاث كما أشار لذلك (Daire, 2008)، ويجب أن نشير إلى أن العامل أو المتغير الكامن هو الذي يؤثر في المؤشرات سواء أكانت فقرات أم اختبارات أو مقاييس وخلافا للنموذج الأحادي العامل، يقوم النموذج العاملي التوكيدي متعدد العوامل على افتراض وجود أكثر من عامل واحد، ليصبح قادرا على استيعاب بنية مفهوم موضوع الدراسة، ولكل عامل متغيراته المقاسة، ويجب على الباحث تكوين بنية واضحة للنموذج المفترض من حيث عدد العوامل وطبيعتها، ومدى وجود علاقات ارتباطية من عدمها، والكشف عن أخطاء القياس لكل مؤشر (القهوجي، 2018، 59).

شكل رقم (13).

يمثل نموذج عاملي توكيدي من الدرجة الأولى (غير الهرمي).

يمثل نموذج عاملي توكيدي من الدرجة الأولى (غير الهرمي).

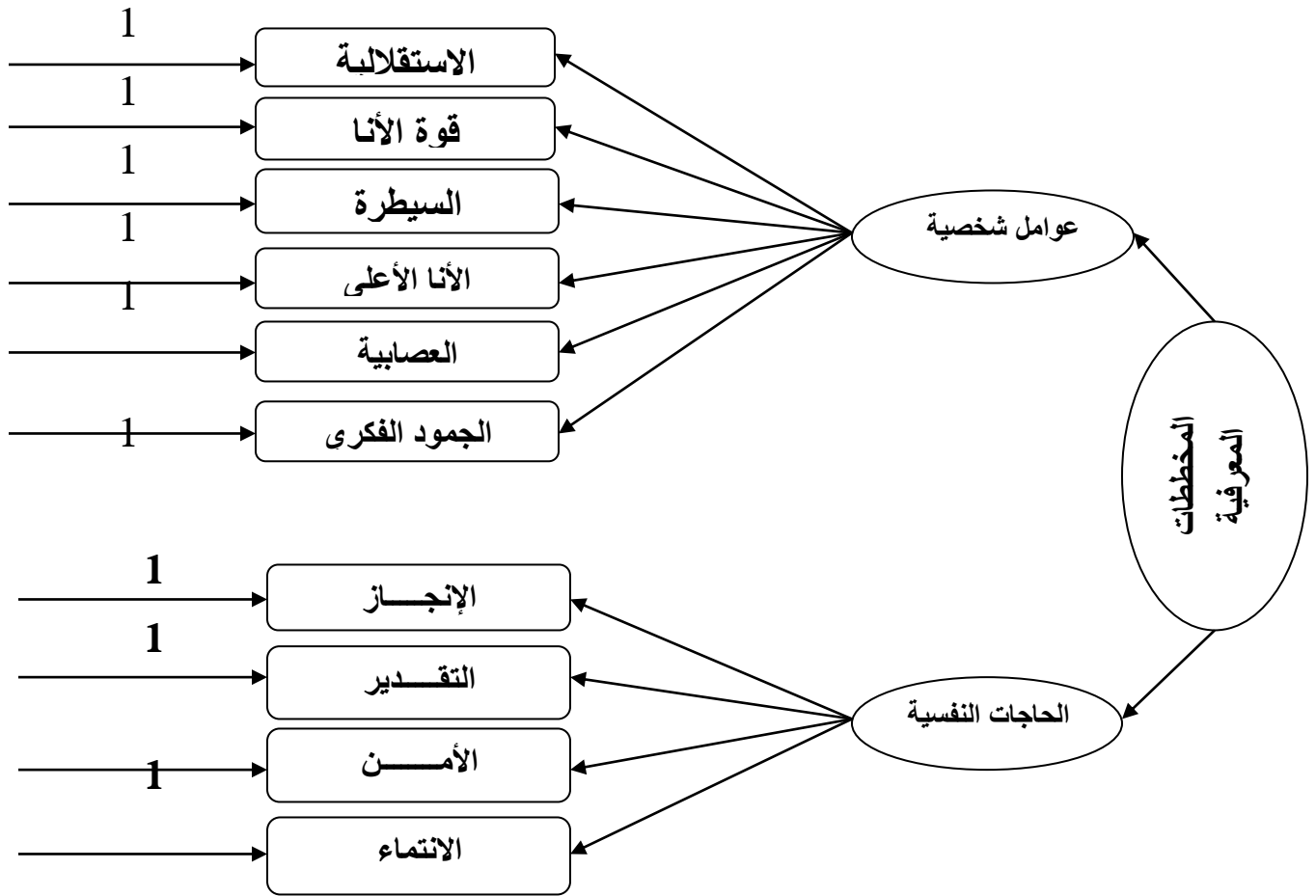


- النموذج العاملي التوكيدي من الدرجة الثانية (الهرمي):

أهم ما يميز التحليل العاملي التوكيدي من الدرجة الثانية وجود عامل هرمي عام ويسمى القائد، ويتم فيه استبدال العلاقات الارتباطية بين العوامل بمسارات تدل على تأثير العامل الهرمي العام، ويلاحظ في النموذج العاملي من الدرجة الثانية أو الهرمي وجود نسبة للتباين غير المفسر (البواقي). وبالتالي يصبح لكل عامل فرعي مصدران للتباين أحدهما يفسر العامل العام ويسمى التباين المفسر، والآخر تفسره عوامل أخرى غير معروفة ويسمى التباين غير المفسر، وعلى الباحث في الاختصار بعدد العوامل الفرعية من خلال جمعها بعامل يسمى الهرمي.

شكل رقم (14).

يمثل نموذج عاملي توكيدي هرمي.



4-3- الفرق بين التحليل العاملي الاستكشافي والتحليل العاملي التوكيدي.

الفارق الجوهرى بينهما هو أن التحليل العاملي التوكيدي يستخدم لاختبار النموذج

النظري على أساس توكيدي للنتيجة من صحة النموذج وصلاحيته، في حين أن التحليل العاملي

الاستكشافي يستخدم لاستخراج العوامل الكامنة للمتغيرات المقاسة بطريقة استكشافية، أي يتم

التعرف على العوامل الكامنة للمتغيرات المقاسة بعد التحليل، وعلى النقيض من التحليل العامل الاستكشافي، فإن التحليل العامل التوكيدي يتطلب بالضرورة أن يحدد الباحث نموذج النظرية العامل بدقة، أي يحدد قبل إجراء التحليل العامل اعتمادا على تأصيله النظري للموضوع، الأبعاد التالية للنموذج العامل:

- نوع النموذج العامل (أحادي، ثنائي، أو متعدد).
- المتغيرات المقاسة (فقرات، مقاييس فرعية، اختبارات ... وغيرها).
- يحدد ما إذا كانت العوامل التي حددها مرتبطة فيما بينها أم أنها مستقلة.
- أخطاء القياس وهو باقي التباين الذي لم يقدر العامل على تفسيره بالنسبة لكل مؤشر من مؤشرات المقاسة (أخطاء عشوائية، أخطاء منتظمة) (تيغزة، 2012، 179-180).

II-3-5. أهداف التحليل العامل التوكيدي:

- أ- دراسة العلاقات بين العوامل الكامنة والمؤشرات التي تمثلها والعوامل فيما بينها.
- ب- تقديم أدلة الصدق البنائي للمقياس وهي ثمان مؤشرات يجب أن تتحقق جميعا، وفي حالة عدم تحققها جميعا نستنتج من الدراسة العلاقة بين العوامل.

II-3-6. تقدير مؤشرات جودة المطابقة:

قبل تفسير النتائج لابد من دراسة مؤشرات التطابق:

مربع كاي (كا²):

يعبر كا² عن مدى حسن المطابقة بين مصفوفة التباين الملاحظة والمتوقعة، فهو يهدف إلى اختبار الفرضية البديلة، أي أن مستوى الدلالة يجب أن تكون غير دالة إحصائياً أي أنه لا يوجد فروق بين النموذج النظري والبيانات ($\alpha \leq 0.05$)، بينما إذا كان مستوى الدلالة ($\alpha > 0.05$)، فهذا يعني وجود فروق وعندها يجب إعادة النظر بحجم العينة أو أن نأخذ كا² المعياري، وهو ظاهر أصلاً بالتحليل ويجب أن يكون أقل من (5)، ويجب أن نشير إلى أن قيمة كا² يساوي صفراً، وتخلو من درجات الحرية، ويعني ذلك أن النموذج المفترض يطابق البيانات المتوقعة تماماً، وكلما ازدادت قيمته، فإن مطابقة النموذج تزداد سوءاً، ولكن على الباحث الاعتماد على مؤشرات أخرى لحسن المطابقة بمعيار كا² لحساسيته لمعاملات الارتباط، فكلما كانت معاملات الارتباط مرتفعة، ترتفع قيمة كا²، وكذا حساسيته لحجم العينة، فهو يزداد طردياً مع زيادة حجم العينة، وأيضاً لعدم واقعيته بين افتراض تطابق النموذج الملاحظ والمتوقع (القهوجي، 2018، 123).

- مؤشر المطابقة المقارنة: Comparative Fit Index (CFI)

تتراوح قيمته بين (0.9 – 1.00) وكلما اقترب من الواحد يكون أفضل، وعندها نقول أن هناك ارتباطات وعلاقات تجمع بين متغيرات النموذج، وأن النموذج يبتعد عن النموذج الصفري، أما إذا كان أقل من 0.90 يتم تعديل النموذج وإعادة توصيفه (المرجع السابق).

- مؤشر تاكر - لويس Tucker-Lewis:

ويسمى بمؤشر المطابقة غير المعياري (NNFI) وهو من عائلة (CFI) وتتنطبق عليه نفس الشروط.

- مؤشر رمسي (RMSEA) Root Mean square error of

Approximation.

يقيس الاختلاف بين مصفوفة التباين للنموذج الذي أخذت منه العينة والمجتمع ويفضل أن أقل من (0.08)، وكلما كان أقل كان أفضل واقترب من الصفر زادت نسبة التطابق، بينما إذا كان بين (0.08 – 0.10) يعتبر غير مقبول وأكثر من ذلك يجب رفض النموذج (القهوجي، 2018، 126).

قائمة المراجع:

اسماعيل، ميمي السيد أحمد (2007). *الخصائص السيكومترية لاختبار القدرة العقلية باستخدام*

نموذج راش لدى طلبة المرحلة الثانوية. كلية التربية، جامعة الزقازيق. مصر.

اسماعيل، بشرى. (2004). *المرجع في القياس النفسي*. ط1. مكتبة الأنجلو المصرية. مصر.

عبد الخالق، أحمد محمد. (1993). *استخبارات الشخصية، الاسكندرية، دار المعرفة الجامعية،*

ط2.

مقدم، عبد الحفيظ (2003). *الإحصاء والقياس النفسي والتربوي*، ط2، الجزائر، ديوان

المطبوعات الجامعية للنشر والتوزيع.

معمرية، بشير (2007). *القياس النفسي وتصميم أدواته، الجزائر، منشورات الحبر.*

أبو حطب، فؤاد. (1996). *القدرات العقلية، الطبعة 5، القاهرة، مكتبة الأنجلو المصرية.*

أبو حطب، فؤاد، وعثمان، سيد أحمد (2003). *التقويم النفسي، الطبعة 3، القاهرة، مكتبة*

الأنجلو المصرية.

أمطانيوس، نايف ميخائيل (2015). *دليل استخدام مجموعة من المقاييس النفسية العالمية في*

البيئة العربية: الجزء الأول، دار الإعصار العالمي للنشر والتوزيع، الأردن.

أمطانيوس، نايف ميخائيل (2015). *القياس والتقويم النفسي والتربوي للأسوياء وذوي الحاجات*

الخاصة، ط1، دار الإعصار العالمي للنشر والتوزيع، الأردن.

أمطانيوس، نايف ميخائيل (2017). الدليل الموسع لاستخدام مقاييس المهارات الاجتماعية والذكاء الانفعالي والاجتماعي في البيئة العربية، ط1، دار الإعصار العالمي للنشر والتوزيع، الأردن.

أمطانيوس، نايف ميخائيل (2016). بناء الاختبارات والمقاييس النفسية والتربوي وتقنيها، دار الإعصار العالمي للنشر والتوزيع، ط1، عمان، الأردن.

باهي، مصطفى حسين ابراهيم، وفياض، ناهد خيري عبد اللهم (2008). التحليل السيكمومي لأدوات القياس الثبات والصدق: دراسة تحليلية. المجلة العلمية للتربية البدنية والرياضية. مصر. 54(1).

بورزق، يوسف، رياش، السعيد (2018). الخصائص السيكمومترية والتحليل العاملي التوكيدي لمقياس جودة الحياة في العمل، دراسة ميدانية ببعض ثانويات مستغانم. مجلة العلوم النفسية والتربية، مجلد7، العدد1.

تايلر، ليونا (1983). الاختبارات والمقاييس. ترجمة سعد عبد الرحمان. بيروت: دار الشروق.

تيغزة، امحمد بوزيان. (2012). التحليل العاملي الاستكشافي والتوكيدي، دار المسيرة، عمان

السيد، أبو هاشم (2006). الخصائص السيكمومترية لأدوات القياس في البحوث النفسية والتربية باستخدام *Spss*. كلية التربية، جامعة الملك سعود.

أبو هاشم، السيد محمد. (2008). مكونات الذكاء الاجتماعي والوجداني والنموذج العلائقي

بينهما لدى طلاب الجامعة المصريين والسعوديين، دراسة مقارنة، مجلة كلية التربية بنها،

المجلد 18، العدد 76، ص ص 157-230.

علام، صلاح الدين محمود (2000). القياس و التقويم النفسي والتربوي، القاهرة. دار الفكر

العربي.

علام، محمود صلاح الدين (2002). القياس والتقويم التربوي والنفسي، القاهرة، دار الفكر

العربي.

مجدي، عبد الكريم حبيب. (1996). التقويم والقياس في التربية وعلم النفس، القاهرة، النهضة

المصرية، ط1.

الأنصاري، بدر محمد (2000). قياس الشخصية، دار الكتاب الحديث

بوسالم، عبد العزيز (2014). القياس في علم النفس والتربية، الجزائر، ط1، قرطبة للنشر

والتوزيع.

تيغزة، امحمد بوزيان (2008). نظرية الصدق الحديثة ومتضمناتها التطويرية لواقع القياس، ندوة

علم النفس: علم النفس والتنمية الفردية والمجتمعية، السعودية.

تيغزة، امحمد بوزيان (2017). توجهات حديثة في تقدير صدق وثبات درجات أدوات القياس-

تحليل نظري تقويمي وتطبيقي، مجلة العلوم النفسية والتربوية، جامعة الوادي.

كوافحة، تيسير مفلح (2005). القياس والتقييم وأساليب القياس والتشخيص في التربية الخاصة.

عمان، دار المسيرة للنشر والتوزيع، ط2.

علام، صلاح الدين محمود (2013). *إتقان القياس النفسي الحديث*. ط 1. عمان: دار الفكر.

القفاص، وليد كمال عفيفي. (2011). *التقويم والقياس النفسي والتربوي*، ط1، المكتب الجامعي

الحديث.

علام، محمود صلاح الدين (2013). *إتقان القياس النفسي الحديث*. ط1. دار الفكر ناشرون

وموزعون. عمان - الأردن.

علام، محمود صلاح الدين (2016). *مقدمة لحزمة البرامج الإحصائية SPSS في علم*

النفس. ط1. دار الفكر ناشرون وموزعون. عمان - الأردن.

القهوجي، أيمن سليمان، أبو عواد، فريال محمد. (2018). *النمذجة بالمعادلات البنائية باستخدام*

برنامج أموس، دار وائل للنشر والتوزيع، الأردن.

القهوجي، أيمن سليمان، وأبو عواد، فريال محمد. (2018). *النمذجة بالمعادلات البنائية باستخدام*

برنامج أموس. دار وائل للنشر والتوزيع. عمان - ط1

مراد، صلاح، وسليمان، أمين (2002). *الاختبارات والمقاييس في العلوم النفسية*

والتربوية. القاهرة، دار الكتاب الحديث.

سبع، محمد أبو لبدة (2008). *مبادئ القياس النفسي والتقييم التربوي*، ط1، دار الفكر، الأردن.

السيد، فؤاد البهي. (1978). علم النفس الإحصائي، قياس العقل البشري، دار الفكر العربي، مصر.

عباس، فيصل. (1996). الاختبارات النفسية تقنياتها وإجراءاتها، ط1، دار الفكر العربي.

ربيع، محمد شحاتة (2008). قياس الشخصية، دار المسيرة للنشر والتوزيع والطباعة.

بام، روبنزين سكوت. (2000). الذكاء العاطفي، ترجمة صفاء الأعسر وعلاء الدين كفاقي، القاهرة، ط1، دار قباء للنشر والتوزيع.

المراجع الاجنبية:

Anastasi, A, Urbina, S. (1997). Psychological testing (7th ed). Prentice-Hall, Inc.

Bacon, Donald.R.& sauer, paul L. & Young, Murray (1995). Composite reliability in structural equations modiling. Educational and psychological Measurment, 55(3).

Begg; David, Stanley fisher & Rudiger Dorbusch: Macroéconomie. Ed. Dunod, Paris, 2002 (01).

Bourdreau, Andre and Guy Couillard, systems Integration and khnowledge Management. Information systems Management 1999: (02).

Cohen Wesly M, Levinthal D, "Absorptive capacity: A new perspective on Learning and Innovation", Administrative science quartley, vol.

- 35, 1990.t Teenagers: parenting with love laughter and limits. New York: Homony/ Randon house.
- Cronbach, L.J, Mehl, P.E, (1995). Construct validity in psychological tests. Psychological bulletin, 52, 281-302.
- Kline, D. (1979). Psychometric and psychology, New York Holt.
- Kline,Rex, B. (2016). Principales and practice of structural equation modeling,3rd ed New York the Guilford press.
- Messiek, S. (1989). Validity, in R. Linn (ed), Educational Measurement 3rd ed. pp 13-103.. Upper sdd le River, NJ: Merill Prentice Hall.
- Ammerican educational research association American Psychological association, National council on measurement in education (1999). Standards for educational and psychological testing, Washington, DC: AERA.
- N.M. Dowine. (1967). Fundementals of Measurment (New York): Oxford university press.1967 p. 84.
- Norman E. Cronlund. (1976). Measurment and Evaluation in teatching. 3rd edition. London : collier Mac millain.1976 p 110
- Nunally, J (1970). Introduction to psychological. New York : Mac Graw-Hill.