

جامعة لونيبي علي -البليدة 2_

كلية العلوم الإنسانية و الإجتماعية

قسم علم النفس و علوم التربية و الأرطوفونيا .

مطبوعة بيداغوجية لمجموعة محاضرات في :

وحدة القياس النفسي و التربوي -السداسي الأول

موجهة لطلبة السنة الثالثة ليسانس .

تخصص :إرشاد و توجيه .

إعداد: د. سامية سعدي

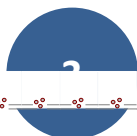
السنة الجامعية:2022_ 2023

3.....	مقدمة:
4.....	المحاضرة الأولى :مدخل لبعض المفاهيم ذات الصلة بنظرية القياس النفسى
26.....	المحاضرة الثانية: مسلمات و عناصر عملية القياس
30.....	المحاضرة الثالثة: مستويات القياس .
36.....	المحاضرة الرابعة: علاقة مستويات القياس بالمعالجات الإحصائية:
45.....	المحاضرة الخامسة:تكملة للمعالجات الإحصائية(المعالجات الإحصائية فى المستوى الرتبى و المسافات المتساوية
53.....	المحاضرة السادسة : صلاحية ادوات القياس (الخصائص السيكميتريية) :
59.....	المحاضرة السابعة: نظريات الصدق
64.....	المحاضرة الثامنة:مصادر أدلة الصدق:
70.....	المحاضرة التاسعة: ثبات الاختبار
86.....	المحاضرة العاشرة : ثبات التجزئة النصفية
97.....	المحاضرة الحادية عشرة : تابع لطرق حساب الثبات (ثبات كيودر ريتشاردسون, معامل ألفا كرونباخ
105.....	المحاضرة الثانية عشرة: المعايير و عملية التقنين
108.....	المحاضرة الثالثة عشرة: نماذج تفسير الدرجة على الاختبار او المقياس
118.....	المحاضرة الرابعة عشر :الخصائص السيكمومترية فى النظرية الجديدة للقياس النفسى والتربوى
121.....	المحاضرة الخامسة عشرة: بناء الاختبارات وفقا لنظرية السمات الكامنة
127.....	المحاضرة السادسة عشرة:مراحل بناء الاختبارات و المقاييس
132.....	قائمة المراجع

مقدمة:

يلعب القياس النفسي دورا بالغ الأهمية ، فهو إلى كونه يستهدف توفير الإختبارات النفسية التي تتوفر فيها الخصائص السيكومترية للإختبار الجيد لإثراء البحوث في علم النفس ، فإنه من زاوية أخرى يعطي أيضا للأخصائي ما يستفيد به في مجال تشخيص مختلف الإضطرابات النمائية . من هنا تأتي ضرورة وعي الأخصائي بكيفية تطبيقها على النحو الذي تعطينا فيما بعد الدقة و الموضوعية المطلوبة في القياس . من خلال المهارة التي تجمع بين النظرية العلمية و الممارسة عند تطبيقه للإختبارات و المقاييس و تصميمها و استخلاص نتائجها و تفسيرها و التعرف على دلالاتها و الدرجات التي يحصل عليها الفرد من أجل الوصول إلى اتخاذ القرارات الدقيقة و الموضوعية، و حتى إلى مرحلة القدرة على التنبؤ بالسلوك.و ذلك بعد الإعتماد على الإحصاء و الذي بدوره يلعب دورا مهما في العلوم النفسية و الإجتماعية ،حيث تطبق المقاييس و الإختبارات النفسية و تعالج نتائجها معالجة احصائية .فهو بالتالي إحدى الدعامات الرئيسية التي تقوم عليها الطريقة العلمية في البحوث النفسية والإجتماعية .كما للقياس جانب كبير من الأهمية ،حيث أصبح يعتبر من الأساليب الدقيقة التي تعتمد عليها البحوث النفسية و الإجتماعية التي تسعى لتطوير أساليب موضوعية دقيقة لقياس الظواهر و الوصول إلى فهمها و تفسيرها و التنبؤ بالعلاقات القائمة بين الظواهر و المتغيرات و محاولة ضبطها و التحكم فيها.

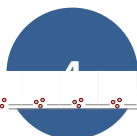
و حيث ان جمع البيانات عملية او مرحلة أساسية في أي بحث ،لذلك تعد أدوات القياس ضرورة لازمة مهما كان نوع البحث.و هناك صفتان أساسيتان لابد من توافرها في أي اداة قياس هما الثبات و الصدق .حيث لهما اهمية كبيرة في البحوث النفسية و التربوية لأن القياس



في هذين المجالين هو قياس غير مباشر . وبالتالي علينا التأكد من ان الأدوات التي نستعملها تقيس فعلا مانود قياسه و يمكن الثقة فيها . فعلى قدر صحتها و دقتها تتوقف دقة نتائج البحث و التحليل الإحصائي ،لأن ذلك يؤثر بالطبع في اهمية النتائج التي نتوصل إليها و صحة مانتخذه من قرارات .

المحاضرة الأولى :مدخل لبعض المفاهيم ذات الصلة بنظرية القياس النفسي

الأهداف : _ التعرف على بعض المفاهيم ذات الصلة بنظرية القياس



1_ مفهوم القياس : هو عملية و صف الظواهر و المتغيرات و صفا كميا، بناء على قواعد و قوانين معينة . باستعمال أدوات قياس . بمعنى استخدام الأرقام في وصف و تبويب و تنظيم المعلومات ، ومن ثم تفسيرها (سعد عبد الرحمان ، 1424هـ - 2003) .

كما يعرف القياس بأنه " عملية تحويل الأحداث الوصفية إلى أرقام بناء على قواعد وقوانين معينة إلى ما هو أسهل من حيث التعامل (سعد عبد الرحمان ، 1424هـ - 2003) .

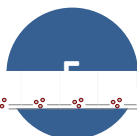
و هو أيضا العملية التي يتم من خلالها التعبير عن الخصائص المختلفة للأشياء أو الأشخاص أو الحوادث أو الظواهر يقيم كمية تحدد بناء على قواعد و شروط محددة (الصمادي، الداربيع _ 2004) .

مفهوم القياس النفسي

يُعرّف الموسوعيون القياس النفسي (Psychometrics) بأنه "مجال دراسة في علم النفس يهتم بنظرية وتقنية القياس . " وهو يهتم بالقياس الموضوعي للبنى الكامنة (المفاهيم) التي لا يمكن ملاحظتها مباشرة، مثل الذكاء والانطواء والتحصيل الدراسي . يتم استنتاج مستويات الأفراد في هذه المتغيرات غير الملاحظة من خلال النمذجة الرياضية بناءً على استجاباتهم على بنود الاختبارات والمقاييس .

وبالتالي، يمكن القول إن القياس النفسي هو:

- عملية منهجية تهدف إلى وصف الصفات النفسية كمياً.



• تقدير كمي لبعد من أبعاد السلوك، حيث نحصل على درجات تعبر عن مستوى الأفراد في التحصيل أو القدرات أو السمات الشخصية.

• تقييم نفسي يُحكم على جودته وصحته اعتماداً على مبادئ أساسية مثل الموثوقية (الثبات) والصدق.

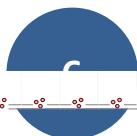
وتجدر الإشارة إلى أن للقياس النفسي جذوراً تاريخية تمتد من تيارين فكريين رئيسيين: التيار الأول الذي اهتم بقياس الفروق الفردية على يد داروين وغالتون وكاتل، والتيار الثاني الذي نشأ في مختبرات علم النفس التجريبي على يد هربارت.

و تتم عملية تحويل الظواهر و السمات و المتغيرات إلى أرقام بواسطة استخدام أدوات من مثل المقاييس و الإختبارات و التي تعرف بانها "أدوات قياس مقننة صممت من أجل الحصول على قياس موضوعي لعينة من السلوك بهدف موازنة اداء الفرد بمعيار اوبمحك اداء معين". تعريف انستازي(1997: anastasi).

بعض المفاهيم ذات الصلة و المستخرجة من التعريف السابق

1_ **التقنين**: و هو انتظام و توحيد الإجراءات عند تطبيق الإختبار و تقدير درجاته، واستخراج المعايير ،وهي إحدى الخطوات المهمة في تقنين إختبار معين.

هو العملية المنهجية التي تهدف الى توحيد اجراءات تطبيق الاختبار و تصحيحه و تفسير نتائجه ،بالاضافة الى تحديد المعايير المرجعية التي تقارن بها درجة الفرد باداء مجموعة



محددة. فهو بالتالي يجعل الاختبار عملية موحدة يمكن استخدامها لمقارنة الافراد ببعضهم البعض بشكل موضوعي، بغض النظر عنمتى او اين أو كيف تم تطبيق الاختبار .

المكونات الاساسية لعملية التقنين: تتضمن عملية التقنين اربعة مكونات رئيسية هي

_ توحيد الاجراءات: بمعنى يجب ان تكون جميع الظروف المحيطة بتطبيق الاختبار موحدة لجميع المفحوصين .

_ اختيار عينة التقنين: هي مجموعة من الافراد التي تطبق عليهم الاختبارات لتحديد المعايير المرجعية ، هذه الاخيرة يجب ان تتميز ب:

- حجم كبير من الافراد .

- تمثيل دقيق للمجتمع المستهدف

- اختيار عشوائي أو طبقي لتجنب التحيز .

_ تحديد المعايير المرجعية: هي جداول احصائية تحول الدرجات الخام الى درجات معيارية لها معنى تفسيري منها (المعايير العمرية،الصفية،المئينية،القياسية)

_ حساب الخصائص السيكومترية للعينة المعيارية (المتوسط و الانحراف ،الصدق و الثبات)

و عليه فإن عملية التقنين تكتسي اهمية بالغة من حيث :تحقيق الموضوعية ،القابلية للمقارنة ،التفسير العلمي للدرجات).

التقنين يعني توحيد الإجراءات. يمكن إضافة أن للتقنين معنيين مترابطين:

1. المعنى الأول (التوحيد): أن تكون تعليمات الاختبار وصياغة بنوده والزمن

المخصص له وشروط تطبيقه وطريقة تصحيحه موحدة في كل المواقف، بما يتيح الحصول على النتائج نفسها تقريباً في حال إعادة إجراء الاختبار. ويفقد الاختبار أساسه العلمي والموضوعي إن لم يكن مقنناً بهذا المعنى.

2. المعنى الثاني (استخراج المعايير): أن يخضع الاختبار للتقنين من خلال تطبيقه على

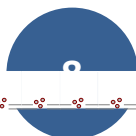
عينة كبيرة ممثلة للمجتمع الأصلي، بهدف استخراج معايير معينة تحدد معنى الدرجة الخام التي يحصل عليها الفرد في ضوء تركز وتشتت درجات أفراد عينة التقنين.

المعنى الأكثر دقة للتقنين هو الجمع بين هذين المعنيين: توحيد الإجراءات واستخراج المعايير من عينة تمثيلية

ملاحظة مهمة : التقنين لا يعني ثبات المعايير للأبد، فالمعايير المرجعية التي ينتجها التقنين لها صلاحية زمنية محدودة (عادة بين 10_15 سنة) و عليه يجب تحديث المعايير بشكل دوري من خلال إعادة عملية التقنين على عينة جديدة في المجتمع الحالي

2_المعايير: تكوين المعايير إحدى الخطوات المهمة في تقنين الإختبارات و المعايير حسب

ساكس هي توزيع الدرجات في اختبار من الإختبارات المطبق على مجموعة مرجعية تم اختيار أفرادها بصورة تمثل المجتمع الأصلي (شحاتة، 2007) .



ويعد مفهوم المعايير من المفاهيم الأساسية المتعلقة بتفسير الدرجات في الإختبارات مرجعية الجماعة، فالدرجة التي يحصل عليها فرد في اختبار ما و التي تسمى الدرجة الخام لا يكون لها معنى ويصعب تفسيرها ما لم يتم مقارنتها بإطار مرجعي. فهذا النظام هو الذي يسمح باستخلاص معلومات مفيدة من درجات الإختبار. كما ويعتمد هذا النظام على الجماعة المرجعية.

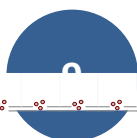
تكمن أهمية المعايير في القياس في انها تتيح لنا تفسير النتائج بدقة، ومن دونها لا يمكن الحكم على الاداء. وهي اساس كل قرار نفسي او تربوي .

يمكن تقسيم المعايير الى نوعين رئيسيين :

1- **المعايير مرجعية المحك** : هنا يفسر اداء الفرد بناءا على اتقانه امجموعة محددة من المهارات، بغض النظر عن اداء المجموعة التي ينتمي اليها. المعيار هنا هو الحد الأدنى (النقطة الفاصلة) الذي يحدد بناءا على حكم الخبراء أو تحليل المهام، يسمى غالبا "درجة القطع" (جلاس، 1963)

أمثلة : في الاختبارات التحصيلية، اختبار يقيس مهارة الضرب، المعيار هو أن يجيب التلميذ على 80% من الاسئلة بشكل صحيح ليعتبر متقنا للمهارة .

في المجال المهني ايضا، المعيار هو الحد الأدنى من الدرجة التي يثبت كفاءة الممارس في مهنة محددة.



اذن المعايير تركز على وصف "مايستطيع الفرد فعله" بدقة، وهي الأكثر عدالة في التعليم لأنها لاتعاقب الطالب بسبب وجود زملاء اقوياء في الفصل .

2- **معايير مرجعية الجماعة** : هنا يفسر أداء الفرد بناءا على مقارنته بأداء المجموعة التي ينتمي اليها (المجموعة المرجعية) .المعيار هنا ليس درجو مطلقة ،بل هو الموقع النسبي داخل المجموعة .

إذا التركيز هنا حول لفكرة " كيف هو أداء الفرد مقارنة بأقرانه" و المعيار هنا يتمثل في هنا، يُفسر أداء الفرد بناءً على مقارنته بأداء مجموعة أخرى (المجموعة المرجعية). المعيار هنا ليس درجة مطلقة، بل هو الموقع النسبي داخل المجموعة.

الفكرة : كيف أداء هذا الفرد مقارنة بأقرانه؟

المعيار : يتمثل في مؤشرات إحصائية مثل:

الرتب المئوية : (Percentile Ranks) إذا كان الطالب في المئين 85، فهذا يعني أنه تفوق على 85% من المجموعة التي قورن بها.

الدرجات المعيارية : (Standard Scores) مثل (Z) ، T ، IQ).

التطبيقات:

اختبارات الذكاء : معامل الذكاء (IQ) هو درجة معيارية متوسطها 100 وانحرافها المعياري

15، يتم حسابها بمقارنة الفرد بأقرانه من نفس العمر. (وكسلر دافبيد ، 2008)

اختبارات القبول الجامعي: مثل SAT أو GRE ، حيث يعتمد القبول على ترتيب الطالب بالنسبة للمتقدمين الآخرين. (college Board 2021)

المميزات: مفيدة لتمييز الأفراد عن بعضهم (الانتقاء) وتحديد الموهوبين أو ذوي صعوبات التعلم.

3_ .التحديات والأخلاقيات في المعايير النفسية والتربوية

التحيز الثقافي: (Cultural Bias) المعايير المبنية على مجموعة معينة (مثل مجموعة ثقافة غربية) قد لا تكون عادلة عند تطبيقها على مجموعات ثقافية أخرى. هذا هو أحد أكبر الانتقادات الموجهة لاختبارات الذكاء المعيارية. (Norm-referenced)

أشار ألفريد بينيه سيمون ، واضع أول اختبار ذكاء عملي ، إلى أن "المقياس لا يقيس الذكاء الفطري بقدر ما يقيس المعرفة المكتسبة في سياق ثقافي معين (Binet et Simon ; 1916)

تغير المعايير عبر الزمن: ظاهرة "تأثير فلين (Flynn Effect) « تشير إلى أن درجات الذكاء ترتفع عبر الأجيال، مما يعني أن المعايير التي وضعت قبل 20 عامًا أصبحت قديمة ويجب تحديثها (Flynn, 1984).

التفسير الخاطئ: هناك خطأ شائع وهو الخلط بين المعيار المرجعي إلى مجموعة (مئين 70) والمعيار المرجعي إلى محك (درجة 7 من 100). الأول يعني أن الطالب أفضل من 70% من زملائه، والثاني يعني أنه أتقن 70% من المادة.

خلاصة

في القياس النفسي والتربوي، المعايير هي الأطر المرجعية لتفسير الدرجات. إما أن تكون:

معيّارًا إلى محك (Criterion): يركز على إتقان المهارة (مطلوب في التعليم والترخيص المهني).

معيّارًا إلى مجموعة (Norms): يركز على ترتيب الفرد بين أقرانه (مطلوب في التصنيف والانتقاء).

اختيار نوع المعيار المناسب يعتمد على هدف القياس: هل نريد أن نعرف ماذا يتقن الفرد؟ أم نريد أن نعرف أين يقف الفرد مقارنة بغيره؟

تعريف المحك

المحك (Criterion) هو معيار خارجي مستقل يُستخدم للحكم على مدى صلاحية وفعالية أداة القياس النفسي. يعرف الإطار المفاهيمي للقياس النفسي المحك بأنه "المتغير التابع الذي نرغب في إصدار بيان بشأنه". (National Assessment Governing Board) "يمثل المحك "الحقيقة" أو "الواقع" الذي نسعى لقياسه أو التنبؤ به من خلال الاختبار.

بعبارة أخرى، عندما نطبق اختبارًا نفسيًا، نسأل: هل نتائج هذا الاختبار تعكس الواقع الفعلي للفرد؟ المحك هو ذلك الواقع الذي نقارن به نتائج الاختبار. وفقًا للمعايير المهنية المقبولة في جمعية علم النفس الأمريكية (APA)، فإن صدق المحك هو أحد الأنواع الرئيسية لأدلة

الصدق التي يجب تقديمها عند التحقق من صلاحية أداة القياس (Equal Employment Opportunity Commission) .

ثانياً: أنواع المحك

1. المحك الآني (Concurrent Criterion)

التعريف: هو المحك الذي يُجمع في نفس وقت تطبيق الاختبار أو في فترة زمنية متقاربة جداً. يعرف بأنه "نوع من أدلة الصدق المرتبط بالمحك حيث يتم جمع كلا المتغيرين في نقطة زمنية واحدة. (National Assessment Governing Board) "

الغرض: يقيس قدرة الاختبار على تقدير الوضع الحالي للفرد بدقة.

أمثلة تطبيقية:

تطبيق اختبار جديد للاكتئاب على مجموعة من المرضى، ومقارنة نتائجه مع تقييم أطباء نفسيين سريريين في نفس الوقت

مقارنة اختبار ذكاء جديد بمقياس وكسلر للذكاء يُطبق في نفس الجلسة

اختبار سريع للكشف عن اضطراب نقص الانتباه مقارنة بالتشخيص السريري الحالي

متى يُستخدم؟

عندما يكون هناك أداة قياس موثوقة متاحة بالفعل

عندما نريد بديلاً أسرع أو أقل تكلفة من الأداة الأصلية

في الدراسات التي تهدف إلى التحقق من صحة أداة جديدة

2. المحك التنبؤي (Predictive Criterion)

التعريف: هو المحك الذي يُجمع في وقت لاحق بعد تطبيق الاختبار، بفواصل زمني قد يتراوح بين أشهر إلى سنوات. يعرف بأنه "نوع من أدلة الصدق المرتبط بالمحك حيث يتم جمع البيانات من مصدرين في نقاط زمنية مختلفة (National Assessment Governing Board) (National Assessment Governing Board) .

يُعرف موسوعة Sage للأساليب البحثية (Sage Encyclopedia of Research Design) الصدق التنبؤي بأنه "المدى الذي يمكن من خلاله استخدام الدرجات من مقياس جديد للتنبؤ بالأداء على مقياس محك يُطبق في وقت لاحق. (Mislevy & Rupp) " **الغرض:** يقيس قدرة الاختبار على التنبؤ بأداء أو سلوك مستقبلي.

أمثلة تطبيقية:

اختبارات القبول في الكليات: مقارنة نتائج الطلاب في الاختبار بمعدلاتهم التراكمية بعد 4 سنوات

اختبارات الكفاءة الوظيفية: مقارنة نتائج المتقدمين بتقارير تقييم أدائهم بعد 6 أشهر من التوظيف

مقياس الميول المهنية: مقارنة نتائج المراهقين بتخصصاتهم الجامعية الفعلية بعد 3 سنوات

متى يُستخدم؟

في عمليات الانتقاء والتوظيف

في التوجيه التربوي والمهني

في الدراسات الطولية (Longitudinal Studies)

ثالثاً: شروط المحك الجيد

1. الموثوقية (Reliability)

يجب أن يكون المحك نفسه أداة قياس موثوقة وثابتة. إذا كان المحك غير موثوق، فإن حسابات الصدق ستكون مضللة. تشير المعايير المنهجية إلى أن معامل ارتباط المحك المقبول يجب أن يزيد عن 0.70 عند مقارنته بمحك معياري. (Terwee et al.)

مثال: استخدام تشخيص سريري كمعيار (محك) يتطلب أن يكون هذا التشخيص متسقاً بين أطباء مختلفين و عبر أوقات مختلفة.

2. الملاءمة (Relevance)

يجب أن يكون المحك ملائماً للبناء النفسي الذي يقيسه الاختبار. تشير الأدبيات إلى أن محكات القياس يجب أن تكون "ذات صلة بالقرارات المرغوبة. (Mislevy & Rupp)"

مثال:

اختبار الذكاء: المحك الملائم هو الأداء الأكاديمي أو القدرة على حل المشكلات

المحك غير الملائم: طول القامة (لا يرتبط بالذكاء)

3. الاستقلالية (Independence)

يجب ألا يكون المحك جزءاً من الاختبار نفسه أو مشتقاً منه. يحذر الباحثون من أن "استقلالية المحك عن المتنبي هي شرط أساسي للتحليل الإحصائي السليم (Psychological Scales) .

4. الخلو من التحيز (Unbiasedness)

يجب أن يكون المحك عادلاً لجميع المجموعات بغض النظر عن خصائصهم الديموغرافية. تؤكد المعايير الفيدرالية الأمريكية أن أدلة الصدق يجب أن تكون "خالية من التحيز" (Equal Employment Opportunity Commission) .

5. الجدوى العملية (Practicality)

يجب أن يكون المحك متاحاً وقابلًا للتطبيق عملياً من حيث التكلفة والوقت والجهد. يشير الباحثون إلى أن "القيود العملية المرتبطة بمحكات غير ملائمة أو مكلفة أو يصعب الحصول عليها قد تفوق الصفات المرغوبة الأخرى لهذه المحكات". (Mislevy & Rupp)

رابعاً : المشكلات والتحديات المرتبطة بالمحك

1. مشكلة نقص المحك المناسب (Criterion Deficiency)

يحدث عندما لا يغطي المحك جميع جوانب البناء النفسي الذي يقيسه الاختبار. تُعرّف الأدبيات البحثية هذه المشكلة بأنها "الفرق بين المعيار المفاهيمي والمعيار الفعلي" (Luong, Sprung, & Zickar) .

مثال: استخدام المعدل التراكمي كمحك للذكاء يفقر لقياس جوانب مثل الإبداع أو الذكاء العملي.

2. مشكلة تلوث المحك (Criterion Contamination)

يحدث عندما تتأثر قيم المحك بمعرفة نتائج الاختبار. يُعرّف الباحثون تلوث المحك بأنه "خلل منهجي خطير يحدث عندما تؤثر المعرفة بالمتغير المتنبئ على قياس المتغير المحك" (Psychological Scales) .

مثال: إذا علم المشرفون بنتائج اختبار الموظفين قبل تقييم أدائهم، فقد يتأثر تقييمهم. تحدث الأدبيات البحثية عن تأثير هذه المشكلة: "نتيجة هذا التلوث هي إدخال خطأ منهجي أو تحيز، يؤدي إلى تضخم مصطنع في تقدير العلاقة بين المتنبئ والمحك (Psychological Scales) .

3. مشكلة التغير الزمني (Temporal Instability)

قد يتغير المحك نفسه بمرور الوقت. تشير الدراسات إلى أن "فترة زمنية معقولة تختلف باختلاف الوظائف وحالات التوظيف ولكنها نادرًا ما تتجاوز 5 سنوات (Equal Employment Opportunity Commission) .

مثال: النجاح الوظيفي بعد 10 سنوات قد يتأثر بعوامل لم تكن موجودة وقت تطبيق الاختبار.

4. مشكلة القيود الأخلاقية (Ethical Constraints)

بعض المحكات المثالية قد تكون غير متاحة لأسباب أخلاقية.

اختيار المحك المناسب هو الخطوة الأهم في تقدير صدق المحك. يجب أن تتوفر "حجج مقنعة بأن المعيار المعياري هو بالفعل 'معياري'. (Terwee et al.)"

استخدام محكات متعددة يعطي صورة أكثر شمولاً عن صدق الاختبار. (Hodge)

الاعتراف بحدود المحك وعدم اعتباره "حقيقة مطلقة". يشير الباحثون إلى "مشكلة المحك"

كأحد التحديات الأساسية في دراسات التعميم. (Luong, Sprung, & Zickar)

توثيق خصائص المحك (موثوقيته، ملاءمته، حدود استخدامه) عند نشر نتائج الدراسات

. (Equal Employment Opportunity Commission)

التمييز الواضح بين أنواع المحكات حسب الغرض من القياس (تشخيصي، تنبؤي، انتقائي)

. (National Assessment Governing Board)

الالتزام بالمعايير المهنية مثل المعايير المشتركة للاختبارات التربوية والنفسية الصادرة

عن APA و AERA و NCME (Cheng)

الموضوعية : المقصود بها تجنب الذاتية والإبتعاد عن الحكام الذاتية في تقدير و تفسير

درجات الإختبار من طرف القائمين على الإختبار. ويحصل ذلك إذا ما تم التقيد بكل القوانين

التي تحكم اداة القياس ، من مثل كيفية التنقيط و كذا كيفية تفسير الدرجات.

الموضوعية تعني ايضا درجة خلو عملية القيتاس من التحيز الشخصي ،بحيث تكون النتائج

مستقلة عن ذاتية الفاحص أو المصحح،وتعتمد فقط على استجابات بالمفحوص او الفرد

.الموضوعية في القياس النفسي تعني تحرر عملية القياس من تأثير الذاتية، أي من آراء

المفحص أو ميوله أو توقعاته، بحيث يؤدي القياس إلى نفس النتيجة بغض النظر عن الشخص الذي يقوم بالتقييم أو تصحيحه.

أبعاد الموضوعية

1. موضوعية التطبيق

- توحيد شروط وإجراءات تطبيق الاختبار على جميع المفحوصين
- استخدام تعليمات موحدة (لفظية أو مكتوبة) في تقديم الاختبار
- ضبط الظروف البيئية (الإضاءة، درجة الحرارة، الوقت)
- الالتزام بالوقت المحدد للإجابة دون تمديد أو تقصير

2. موضوعية التصحيح

- وجود مفتاح تصحيح محدد ومسبق الإعداد
- استخدام طرق تصحيح موضوعية (مثل التصحيح الآلي أو النماذج المقننة)
- تجنب التأويل الشخصي للإجابات، خاصة في الاختبارات المقالية
- الاتفاق بين المصححين (معامل ارتباط بين تقديرات مصححين مختلفين)

3. موضوعية التفسير

- الاعتماد على معايير ثابتة (أوزان معيارية، مئينات، درجات معيارية)

-تفسير النتائج في ضوء المعايير الموضوعية مسبقاً

-تجنب القفز إلى استنتاجات غير مدعومة بالبيانات

الموضوعية في القياس النفسي تعني تحرر عملية القياس من تأثير الذاتية، أي من آراء المفحص أو ميوله أو توقعاته، بحيث يؤدي القياس إلى نفس النتيجة بغض النظر عن الشخص الذي يقوم بالتقييم أو التصحيح.

طرق تحقيق الموضوعية

الاختبارات الموضوعية

- أسئلة الاختيار من متعدد

- أسئلة الصواب والخطأ

- أسئلة المزاوجة

-أسئلة إكمال الفراغ (إذا كانت الإجابة محددة)

توحيد الإجراءات

-إعداد دليل تعليمات موحد للتطبيق

-تدريب المطبقين على نفس الأسلوب

-توحيد زمن الاختبار

-توحيد طريقة توزيع أوراق الإجابة

التحقق من الموضوعية إحصائياً

-حساب معامل ارتباط بين تصحيحات مصححين مختلفين

-تحليل التباين لتقدير نسبة الخطأ الموضوعي

-استخدام نظرية القياس الكلاسيكية أو نظرية الاستجابة للمفردة

تحديات الموضوعية

في الاختبارات الإسقاطية

-صعوبة تحقيق الموضوعية في اختبارات مثل رورشاخ وتات

-الحاجة إلى مصححين مدربين تدريباً عالياً

-محاولات تطوير أنظمة تصحيح موضوعية (مثل نظام إكزرنر لرورشاخ)

في التقييمات الأدائية

-صعوبة وضع معايير موضوعية للأداءات المتنوعة

-الحاجة إلى مقاييس تقدير مفصلة (روبيريك)

-تدريب المقيمين على معايير موحدة

في المقابلات التشخيصية

-تأثير شخصية المقابل وطريقة طرح الأسئلة

-الحاجة إلى مقابلات منظمة) مثل(SCID)

-استخدام مفتاح تصحيح موضوعي للاستجابات

أهمية الموضوعية :

في مجال التقييم التربوي

-عدالة توزيع الدرجات بين الطلاب

-إمكانية مقارنة نتائج الطلاب عبر فصول ومدارس مختلفة

-اتخاذ قرارات موضوعية (قبول، نجاح، منح دراسية)

في مجال التوظيف والانتقاء

-تحقيق تكافؤ الفرص بين المتقدمين

-التنبؤ الموضوعي بالأداء الوظيفي

-تجنب التمييز العنصري أو الجندي

حدود الموضوعية

-لا تعني الموضوعية بالضرورة الدقة المطلقة

-قد يؤدي التطرف في الموضوعية إلى فقدان معلومات نوعية مهمة

-بعض الظواهر النفسية يصعب قياسها موضوعياً بالكامل (كالوعي، المشاعر المعقدة)

-الموضوعية العالية قد تأتي على حساب الصدق في بعض السياقات

2_خصائص القياس

يختلف القياس باختلاف الظاهرة موضوع القياس، ففي العلوم الفيزيائية يكون أكثر دقة عنه في العلوم السلوكية ،لأن هذا المجال الأخير يكون عرضة للخطأ.لهذا يقتضي الأمر أن نلم بخصائص القياس النفسي و هي كما يلي :

1-2 - القياس النفسي كمي

2-2 - القياس النفسي غير مباشر، من حيث أننا لانقيس السمة مباشرة ،و إنما نقيس المؤشرات الدالة على السمة .

3-2 _ الخطأ مكون أساسي من مكونات عملية القياس النفسي.

4-2 - القياس النفسي نسبي ، غير مطلق،وعليه فلا يمكن تفسير نتائج القياس إلا بمقارنتها بمعيار او بمستويات متسقة من أداء المفحوصين أو أداء عينة مماثلة لعينة المفحوصين.

5-2 - رغم أن وحدات القياس النفسي – الدرجات _ توصف بأنها غير متساوية حقيقة ،إلا أننا نستطيع إيجاد وحدات تتصف بالثبات و التساوي مثل الدرجات المعيارية.

2_6 - الصفر في القياس النفسي صفر اعتباطي ،بمعنى لايدل على انعدام الظاهرة .

(بشير معمريه،2007: 37-38).

ويمكن تلخيص أهم الفروقات بين القياس المادي و القياس النفسي في الجدول التالي :

الجدول (رقم 1). يوضح الفرق بين القياس المادي و القياس النفسي.

القياس النفسي	القياس المادي
غير مباشر	مباشر
نسبي	مطلق
مرتبط بغيره	مستقل
أقل ضبطاً و دقة	أكثر ضبطاً و دقة
غير مقنن او ممعير عالميا	مقنن، عالمي، ممعير
لا تتوفر فيه وحدات قياس محددة	تتوفر وحدات قياس محددة
صدق الأدوات خاضع للشك	صدق الأدوات لاشك فيها

المصدر: ملحم سامي محمد، 2012:32)

تستند خصائص القياس النفسي (القياس غير المباشر، النسبي، احتواؤه على الخطأ، إلخ)

إلى نظرية الاختبار التقليدية. (Classical Test Theory)

تنص هذه النظرية على أن الدرجة الملاحظة التي يحصل عليها الفرد في الاختبار

هي مجموع الدرجة الحقيقية (الدرجة الخالية من الأخطاء) ونتيجة الخطأ. ويمكن تمثيل

ذلك بالمعادلة:

الدرجة الكلية = الدرجة الحقيقية + _ درجة الخطأ

هذه المعادلة هي الأساس النظري الذي يفسر لماذا يعتبر الخطأ مكوناً أساسياً في القياس النفسي، ولماذا نسعى دوماً لتقدير مقدار هذا الخطأ من خلال حساب معاملات الثبات

المحاضرة الثانية: مسلمات و عناصر عملية القياس

الأهداف:- أن يتمكن الطالب من معرفة مسلمات القياس النفسي

1_ مسلمات القياس النفسي: تقوم مسلمات القياس النفسي على الآتي

1-1 قابلية أداء الفرد للقياس: وهذا يعني أنه بالإمكان تحويل أداء الفرد من صيغة وصفية نوعية إلى صيغة كمية باستخدام الأرقام حسب قواعد معينة . أي التسليم بأن كل الظواهر النفسية و التربوية مثلها مثل الظواهر الطبيعية يمكن قياسها و تقديرها . كما ان خبراء القياس يعتقدون أنه يمكن قياس المتغيرات النفسية و التربوية و هذا استادا إلى قبولنا بأن لهذه المتغيرات وجود حيث يقول كروباخ " إذا وجد شيء ما فإنه يوجد بمقدار معين، و إذا وجد بمقدار فإنه يمكن قياسه".د

1_2 اداء الفرد له خصائص مميزة: وهذا يعني أن كل اداء أو سلوك إنما يصدر عن خاصية واحدة او مجموعة خصائص يتميز بها الفرد عن غيره من بقية الأفراد . و ان الخاصية الواحدة تعطي لكثير من نمط او أداء، كما ان الأداء الواحد ينتج اكثر من خاصية واحدة . وهذه المسلمة تفترض ان أداء الفرد هو دالة خصائصه يدور حول محورين :

- علاقة الخاصية بالأداء من حيث النوع و الكم.

- تأثر اداة القياس بهذه العلاقة .

1-3 إختلاف العلاقة بين الخصائص و الأداء: تتركز هذه المسلمة على فكرة أن الخاصية و

الأداء والعلاقة بينهما تختلف من فرد إلى آخر، و أن هذا الإختلاف هو الذي قامت عليه عملية

القياس .حيث أشارت التجارب الأولى التي أجريت في مختبرات علم النفس في بداية تطوره وبالضبط في مختبر فونددت في ألمانيا ،حيث كانت تهدف التجارب إلى قوانين عامة مشتركة و قانون عام موحد لسلوك الإنسان و أدائه ،وعندما كان يسجل إختلاف اداء الفرد عند الإستجابة لنفس المثير ،كان يعتبر ذلك من باب الخطأ.أما الإتجاه الآخر و هو الإتجاه الذي يؤكد فكرة القياس العقلي واستخدام ادوات القياس فقد اعتبرت هذه الفروق و الإختلافات و التباين أساس عملية القياس بل مانهدف إلى قياسه فعلا.و بالتالي فكل عملية من عمليات القياس إنما تتم في إطار نسبي هو إطار الإختلاف و التباين الذي يوجد فعلا بين خصائص الأفراد و ادائهم .و بالتالي فغن عملية القياس في هذا الإطار نسبية و ليست مطلقة.

1_4 الخطأ مكون أساسي من مكونات عملية القياس : فكل درجة على مقياس ما إنما تتكون من درجتية هما : الدرجة الحقيقية و الدرجة التي تعود إلى الخطأ . وفقا للمعادلة التالية :الدرجة الكلية = الدرجة الحقيقية+ الدرجة التي تعود إلى الخطأ.

و على أساس ذلك يمكن أن نصنف أنواع الخطأ على النحو التالي :

أنواع الخطأ في القياس النفسي:

1- الخطأ الثابت: و هو النوع من الخطأ الذي يعود إلى المقياس في حد ذاته و يتكرر بصفة منتظمة و له نفس التأثير على كل درجة على هذا المقياس. و من ثم فإذا عرفت كمية الخطأ ،فإنه لايشكل مشكلة هامة بالنسبة لعملية القياس.

2- خطأ القياس: و هو الخطأ الناتج عن استخدام الدرجة الظاهرية في القياس بدلا من الدرجة الحقيقية ،وهو النوع من الخطأ الذي يحتاج إلى معالجة احصائية خاصة للتحكم فيه.

3_ **خطأ الصدفة او العشوائية:** و هذا الخطا لا يحتاج إلى توضيح ،إذ ان هذا النوع من الخطأ —بحكم تسميته— لايمكن ضبطه او السيطرة عليه ،لأنه لابد ان يكون عشوائيا . وهذه الأخطاء العشوائية هي التي يلغي بعضها البعض الآخر.و خاصة إذا كان حجم العينة كبيرا ،و على ذلك فإننا نلجأ إلى مجموعة من المسلمات الفرعية لتحديد العلاقة بين هذه الأخطاء العشوائية و الدرجة الظاهرية ،او الدرجة الكلية التي حصل عليها الفرد و درجته الحقيقية التاي تعبر عن قدرته الفعلية على البعد الذي يتم قياسه.

2_ **عناصر عملية القياس:** تركز عملية القياس على العناصر الأساسية التالية :

- **موضوع القياس :**بمعنى تحديد الظاهرة النفسية أو التربوية المراد قياسها . فعملية القياس تتناول ظاهرة محددة كموضوع للقياس و التي نريد قياسها لدى الأفراد .من منطلق المسلمة التي تقول بإمكانية قياس السمات و الخصائص النفسية و الظواهر .و في هذا الصدد يشير سعد جلال (1985) إلى أن القياس لا يوجه إلى الأشياء أو الأشخاص و إنما لخصائص معينة يتم تحديدها بصفة دقيقة ،بحيث يمكن في النهاية تحويل الظاهرة او السمة او الخاصية من الوصف الكيفي إلى الوصف الكمي الذي يسمح بإجراء مختلف الأساليب الإحصائية لفهم و تفسير الظاهرة موضوع القياس.(سعد جلال ، 1985).

فبعد ان يحدد الباحث الظاهرة ،يتم ضبطها إجرائيا ،مع محاولة عزلها عن بقية الظواهر التي يمكن ان تتداخل معها . و نعني بالتحديد الإجرائي هو التحديد الواضح لمؤشرات الظاهرة القابلة للقياس.و هي الخطوة الأساسية و المهمة في عملية القياس

تحديد الأداة المناسبة لعملية القياس

وهذه الخطوة تركز في الأساس على الخطوة السابقة و المتمثلة في تعريف الظاهرة تعريفا إجرائيا بحيث نعتمد عليه في اختيار او حتى في بناء الأداة المناسبة .من منطلق ان الإختبارات و المقاييس حسب صفوت فرج (1980) ماهي إلا مثيرات مرتبة أعدت لتقيس بطريقة كمية في أغلب الحالات و كيفية في بعضها لتقيس عينة من السمة أو الخاصية.(صفوت فرج،1980).

_العدد (الرقم) : القياس النفسي في أصله يقوم على التعامل مع السمات موضوع القياس بلغة الأرقام (الرقم هو دالة السمة أو الخاصية أو الظاهرة) .يقول صلاح أحمد مراد (2000) ،يمكن من خلال خصائص العدد مقارنة قدرات الأفراد فيما بينهم ،و معرفة درجة توفرها عند كل فرد .كما ان الرقم يتيح لنا تقديم التفسير الموضوعي لظاهرة معينة و إعطائها دلالة. الرقم هو نتيجة لعملية القياس ،ينطوي على معاني عدة ،لهذا من الضروري معرفة دلالة الرقم او الدرجة و قواعد استخدامها ،من منطلق ان الرقم كما يقول صلاح أحمد مراد(2000) هو نتيجة لسلسلة من العمليات و الإجراءات التي اتبعت للحصول عليه،مما يتطلب أخذها بعين الاعتبار عند التعامل مع أي قيمة عددية. بمعنى آخر معرفة مستوى القياس الذي ينسب إليه الرقم أو العدد أو الدرجة.

المحاضرة الثالثة: مستويات القياس .

الأهداف : بعد دراسة هذه المادة العلمية و ضمن هذه المحاضرة ،يتوقع من الطالب أن يكون قادرا على :

التمييز بين مستويات القياس الأربعة من حيث نوعية البيانات، و العمليات الحسابية

- التمييز بين المتغيرات الكمية و النوعية

يصنف المتغيرات حسب عدد من المعايير.

إن التعرف على مستويات القياس من الأساسيات التي ينبغي على الباحث الإطلاع عليها لأنها تمكنه من اختيار الطرق او المعالجات اختبار الفرضيات .وقد قام ستيفنس (1946) في مقال بعنوان " نحو نظرية مستويات القياس " بتقسيم مستويات القياس إلى أربع مستويات هي : 1. s s Steven ;

1-المستوى الإسمي (التصنيفي): وهو أدنى مستويات القياس ، لأنه يستخدم الأرقام للدلالة على الأشياء و يستخدم في معظم الأحوال مع المتغيرات النوعية ،حيث يتم تصنيف الأفراد مثلا حسب الجنس أو المستوى الإجتماعي و غيرها . أي أن الهدف هو مجرد التصنيف فقط . مثل (ناجح،راسب) ، (منبسط،منطوي) , واستخدام الأرقام هنا لايعكس مقادير كمية ،ومتغيرات هذا المستوى تكون دائما منفصلة.كما انه لايسمح بإجراء العمليات الحسابية الكلاسيكية ماعدا العد البسيط (التكرارات) .

2-المستوى الرتبي : يعد هذا النوع من المقاييس الأعلمن حيث مقارنته بالمستوى الإسمي ، فبالإضافة إلى تصنيف الأفراد وفق خصائص معينة ، فإنه يمكن ترتيب الأفراد تصاعديا او تنازليا بناء على خاصية أو صفة معينة. وهو بالتالي يتأثر ببداية العد ، كما ان المسافة بين رقم و آخر لا يشترط أن تكون متساوية . و عليه فإنه لا يمكن تحديد مقدار الفرق بين الأفراد المتجاورين في أي مجموعة .

3-مستوى المسافات المتساوية : هذا المستوى يقترب كثيرا إلى المعنى الكمي للقياس حيث تحمل الأرقام هنا معنى كمي . ولعل اهم ما يميز هذا المستوى من القياس وجود ما يسمى الصفر الافتراضي حيث لا يمثل الصفر هنا غياب الظاهرة او السمة نهائيا ، و إنما الصفر هنا نسبي و ليس مطلق .

و معظم الإختبارات النفسية تصمم للحصول على درجات تقع في المستوى الفئوي (**باترسون 2002**) . و نستطيع التعامل مع بيانات المستوى الفئوي باستخدام جميع العمليات الحسابية المعروفة ، كما نستطيع استخدام ومعظم المؤشرات الإحصائية من مثل مقاييس النزعة المركزية و مقاييس التشتت مع البيانات المحصل عليها .

4-المستوى النسبي : يعد هذا القياس اعلى و ادق من جميع المستويات الثلاثة السابقة ، من حيث انه يوجد فيه الصفر الحقيقي (المطلق) الذي يدل على انعدام الظاهرة أو السمة . ويصعب و من النادر استخدام هذا النوع من المقاييس في قياس الظواهر السلوكية و النفسية ، حيث يصعب افتراض انعدام الظاهرة نهائيا ، بينما يستخدم هذا النوع من القياس في العلوم الطبيعية و العلوم الدقيقة التي لا يوجد فيها احتمال الخطأ عند عملية القياس الإحصائية

المناسبة لتحليل البيانات ،و قبل هذا التعرف على نوعية البيانات المتحصل عليها و منه لكي يستخدم الطرق الإحصائية المناسبة لتحليل البيانات .

جدول يوضح : الفروق بين مستويات القياس الأربعة.

المعالجات الإحصائية	أمثلة توضيحية	الخصائص القياسية	العمليات الحسابية	اسم المستوى
التكرارات، المنوال، الوسيط	الجنس المهنة الجنسية الحالة العائلية	الرقم لايدل على كم أو مقدار السمة . الأرقام تحل محل الأسماء تميز الأرقام بين المجموعات لايمكن إجراء العمليات الحسابية .	العد البسيط	الإسمي أو التصنيفي

الوسيط ،المئينيات،معامل ارتباط الرتب لسبيرمان ،معامل ارتباط كندل للرتب	درجات الطلبة في وحدة القياس النفسي	الأرقام عبارة عن رتب الأرقام مرتبة تنازليا أو تصاعديا المسافات بين الرتب متساوية ظاهريا و غير متساوية فعليا يهتم بترتيب الأفراد وفق خاصية محددة .	التساوي الترتيب	الرتبي أو الترتيبي
المتوسط الحسابي الإنحراف المعياري معاملات الارتباط : بيرسون ،التنائي و التنائي الأصيل .	الدرجات في الإختبارات و المقاييس التحصيلية او النفسية	الرقم يدل على كم الظاهرة أو مقدار السمة القيم متصلة	كل العمليات الحسابية	المسافات المتساوية او الفئوي

مقاييس أو اختبارات الدلالة الإحصائية .		وضع الأفراد في مقياس متصل يتكون من مسافات متساوية الصفر اعتباطي يمكن إجراء المقارنة بين الدرجات		
جميع المعالجات الإحصائية الممكنة	زمن رد الفعل الطول الوزن .	العدد يدل على كم او مقدار (قيم متصلة) وضع الأفراد في مقياس متصل يتكون من مسافات متساوية الصفر حقيقي مطلق.	جميع العمليات الحسابية	النسبي

		يمكن استخدام النسب للمقارنة .		
--	--	----------------------------------	--	--

المحاضرة الرابعة: علاقة مستويات القياس بالمعالجات الإحصائية:

الأهداف :

- أن يعرف الطالب طبيعة العلاقة بين مستويات القياس و المعالجات الإحصائية.
- أن يميز الطالب بين الإحصاء الباراميتري و الإحصاء اللاباراميتري.
- أن يعرف الطالب مختلف المؤشرات الإحصائية الواسعة الاستخدام.
- أن يميز الطالب بين مختلف المعالجات الإحصائية في كل مستوى من مستويات القياس.

مدخل

كما هو معلوم أن الأساليب الإحصائية تعتمد على حجم العينة و توزيع البيانات بالدرجة الأولى، بالإضافة إلى مستوى القياس ، ففي مستويات القياس الاسمية او الرتبية نستطيع استخدام الأساليب الإحصائية اللاباراميتريية (اللامعملية) التي لا تعتمد على التوزيع الطبيعي (المعتدل) .بينما في المقاييس في مستوى المسافات المتساوية او النسبية فيمكن استخدام الأساليب الباراميتريية (المعلمية) التي تعتمد على التوزيع الطبيعي (المعتدل)

الفرق بين الأساليب الباراميتريية و اللاباراميتريية

الجدول رقم (2) يوضح الفرق بين الأساليب الباراميتريية و الأساليب اللاباراميتريية

الأساليب الباراميتريية	الأساليب اللاباراميتريية

لاتوجد شروط معينة في تحديد حجم العينة أو شكل التوزيع للظاهرة المدروسة . تستخدم في مستوى القياس الإسمي و الرتبي . المتغيرات كيفية و رتبية . تستخدم في التحقق من صحة الفروض المتعلقة بمجموعات معالمها غير معروفة المعلومات . تستخدم في أغلب الأحيان للعينات الكبيرة . لا تشترط طريقة معينة في اختيار العينة . من أمثلتها : التكرارات ، النسب المئوية ، كاف مربع ، كروكسال والس ، سبيرمان ، المنوال .	تخضع لشروط من بينها : اعتدالية التوزيع تجانس العينات الإختيار العشوائي للعينات . تستخدم في مستوى القياس : المسافات (البيانات كمية) ، و النسبي . تستخدم في التحقق من صحة الفروض المتعلقة بمجموعات قيم بارامترات محددة و معروفة . تستعمل للعينات الكبيرة و الصغيرة من أمثلتها : المتوسط ، الانحراف ، التباين ، اختبار ت ، معامل بيرسون ، تحليل التباين .
--	---

سوف نستعرض المعالجات الإحصائية الممكنة و نرتبها وفقا لمستويات القياس ، و
لا يخفى علينا ان مستويات القياس مرتبة تنازليا من أبسط مستوى إلى أدق مستوى ،
و هذا تبعا لخصائص و اهداف كل مستوى و هذا ما سوف نوضحه في العناصر
التالية :

أولا : . المعالجات الإحصائية في المستوى الإسمي، الهدف هو التصنيف وفقا لمعايير
يحددها الباحث و البيانات عبارة عن تكرارات ، لن التصنيف يعتمد في أساسه على العد
البسيط ، من هنا يمكن التمييز بين المعالجات الإحصائية الممكنة في هذا المستوى وفقا لأنواع
التصنيف المعتمدة نجد:

التصنيف الثنائي الحقيقي، والثنائي الغير الحقيقي، والتصنيف المتعدد .و هذه الأنواع من
التصنيفات تتوافق مع نوع المتغيرات التي تم تصنيف الأفراد على أساسها

. التصنيف الثنائي الحقيقي، يتمثل في تصنيف الأفراد وفقا لمتغيرين، لكل واحد منهما
تقسيمان إثنين لا ثالث لهما و حقيقيين ، بمعنى ان المتغيرين موجودين في الواقع على تلك
الصفة و لم يتدخل الباحث في ذلك و لا تتغير ، ثابتة . و في هذه الحالة يمكن ان نجد حالات
قليلة من المتغيرات الثنائية الحقيقية (الجنس ، الأسئلة التي تحتل إجابتين فقط إما صح او
خطا) .و من اجل دراسة العلاقة بين هذه الأنواع من المتغيرات نجد :

$$\phi = \frac{ad - bc}{\sqrt{(a+b)(c+d)(b+d)(a+c)}}$$

معامل ارتباط فاي:

معامل فاي

$$P_1 = \frac{f_1}{n} \quad P_2 = \frac{f_2}{n} \quad Q_1 = 1 - P_1 \quad Q_2 = 1 - P_2$$

$$\chi^2 = n \cdot \phi^2$$

$$\phi_{max} = \sqrt{\frac{P_1}{Q_1} \cdot \frac{Q_2}{P_2}}$$

يستعمل لقياس العلاقة الارتباطية بين متغيرين من المستوى الإسمي لكلاهما تقسيم ثنائي حقيقي، بحيث البيانات تكون في شكل جدول توافق مكون من صفوف واعمدة. (2 + 2). -تتراوح قيم معامل ارتباط فاي بين (0 _ 1) ، بمعنى أنه لا يأخذ القيم السالبة ، فهو إذن لا يحدد اتجاه العلاقة.

الفرضيات في هذا النوع من المعالجات تصاغ كما يلي :

_ توجد علاقة بين إجابات الطلبة على سؤالين من أسئلة اختبار تحصيلي (حيث ان

$$\phi = \frac{ad - bc}{\sqrt{(a+b)(c+d)(a+c)(b+d)}}$$

السؤال (ان يحق) (فاجاب) (d فقط) (b صحيح) (a خطأ)

ملاحظة : معامل فاي حالة خاصة يستخدم من أجل دراسة صدق الأسئلة الثنائية التقسيم .

و من أجل دراسة الدلالة الإحصائية لمعامل فاي نعتمد على اختبار الدلالة الإحصائية كاف مربع ، إذن دلالة معامل فاي من نفس دلالة كاف مربع ..

الإرتباط الثنائي الأصيل ، و هنا يهتم الباحث بقياس العلاقة الإرتباطية بين متغيرين احدها كمي من مستوى المسافات المتساوية و الآخر كفي ثنائي حقيقي . و هو يصلح لدراسة صدق محتوى بنود الإختبارات التحصيلية التي تعتمد على الإجابات الصحيحة و الإجابات الخاطئة . فهو يستعمل في قياس صدق البنود من خلال دراسة العلاقة بين البند و الدرجة الكلية للإختبار بحيث تكون اجابة على البنود ثنائية (صح ، خطأ) وتعطى الدرجة 1 للإجابة الصحيحة و الدرجة 0 للإجابة الخاطئة . ويسمى صدق محتوى البند ، أي هل البند يتماشى ومحتوى الإختبار .

هذا المعامل لايشترط خطية العلاقة باعتباره مقياسا لباراميتريا . ولايمكن أن يأخذ القيم السالبة ، فهو إذن لايهتم باتجاه العلاقة . الفرضية التي تتضمن مثل هذه البيانات تصاغ كما يلي :توجد علاقة بين إجابات الطلبة على السؤال الأول (صح ، خطأ) و الدرجة الكلية في اختبار تحصيلي . و الهدف هو دراسة صدق السؤال ، بمعنى هل يقيس السؤال نفس ما يقيسه الإختبار ككل.

معامل الإرتباط الثنائي الأصيل:

معامل الارتباط
الثنائي الأصيل

$$r_{pb} = \frac{\bar{Y}_0 - \bar{Y}_1}{\sqrt{\frac{N}{n_1 \cdot n_0}} \sqrt{\sum y^2 - \frac{(\sum y)^2}{N}}}$$

$$t = \frac{r_{pb}}{1/\sqrt{n}} \quad n < 30 \quad df = n - 2$$

$$Z = \frac{r_{pb}}{1/\sqrt{n}} \quad n > 30 \quad df = n - 2$$

معامل الارتباط الثنائي:

حيث يهدف الباحث إلى دراسة العلاقة بين متغيرين أحدهما كمي و الآخر إسمي منفصل ،
أو بين متغيرين بحيث أن المتغير الإسمي ثنائي غير حقيقي و المتغير الثاني كمي . و تصاغ
الفرضية بالصيغة التالية:

توجد علاقة بين التوافق النفسي الإجتماعي(توافق (1) عدم التوافق (0)) و درجات
التحصيل لدى طلبة كلية العلوم الإجتماعية . في هذه الحالة فالمعالجة المناسبة هي معامل
الارتباط الثنائي ،والذي يعطى بالصيغة التالية:

معامل الارتباط الثنائي:

نستخدم أحد القانونين:

$$r_{bi} = \frac{\bar{x}_1 - \bar{x}_0}{S_x} \left(\frac{p \times q}{h} \right)$$

معامل الارتباط الثنائي

أو

$$r_{bi} = \frac{\bar{x} - \bar{x}_0}{S_x} \left(\frac{p}{h} \right)$$

المعالجات الإحصائية في حالة التقسيمات أو التصنيفات المتعددة

أما في حالة التصنيف المتعدد التقسيمات .حيث في هذه الحالة يتم تصنيف المجموعة الواحدة وفقا لثلاث متغيرات أو أكثر . و بالتالي فإن التصنيفات هنا غير حقيقية و الباحث هو الذي يتدخل باستخدام الأدوات المناسبة (ادوات القياس) حيث يتم تصنيف العينة وفقا لما تسفر عنه عملية القياس من نتائج .

وعليه فإن البيانات تكون في جدول توافق مكون من صفوف و اعمدة (2 + 3) ويعتمد على:

كاف مربع للإستقلالية .:وهذا الأخير (كاف مربع) هو اختبار استدلالي للفروق وتعميم النتائج ، وهو اختبار لاباراميتري _ ليس اعتدالي _ يعتمد على المتغيرات النوعية ويستعمل كما سبقت الإشارة في المستوى الإسمي .

1 _ باعتبار كاف مربع يعتمد في حسابه على التكرارات ، فلا يمكن تطبيقه إذا كان التكرار المتوقع أقل من 5 ، لذلك يلجأ الباحث إلى دمج الخانات او مضاعفة حجم العينة .

2 _ يطبق كاف مربع المصحح في حالة درجة الحرية تساوي 1 ، سواء كان اختبار التتابع أو اختبار الإستقلالية

ملاحظة:

يمكن التمييز بين معالجتين إحصائيتين تبعا دوما لنوع المتغيرات ، نجد متغيرات غير قابلة لا للقياس الكمي و لا للقياس الرتبي (المؤهل العلمي ، المستوى الإجتماعي

(....)، و متغيرات إسمية أيضا لكن يمكن تحويل بياناتها إلى رتب أو إلى درجات (التحصيل الدراسي ، التوافق النفسي) .

معامل التوافق c: هو نوع من المعالجات الذي يستخدم مع نوع من المتغيرات التي لا يمكن تحويلها لا إلى رتب و لا إلى درجات خام، مثل المتغيرات المدرجة في هذه الفرضية

1_ توجد علاقة بين **نوع التخصص في البكالوريا و التخصصات المرغوبة في الجامعة** .

هذا النوع من الفرضيات يتضمن متغيرين (نوع البكالوريا) و (التخصصات المرغوبة) ، هذين المتغيرين لا يمكن التعبير عنهما لا ببيانات رتبية و لا بدرجات . في هذه الحالة نستعمل معامل التوافق الذي يعتمد في حسابه على اختبار كاف مربع_

معامل التوافق:

$$c = \sqrt{\frac{G - 1}{G}}$$

اختبار كاف مربع:

في حالة ما إذا كانت المتغيرات في المستوى الإسمي، و لكن قابلة للتحويل إلى رتب أو إلى درجات

من بين المتغيرات من هذا النوع نجد تلك الموجودة في الفرضية التالية:

_توجد علاقة بين مستوى الذكاء (مرتفع ، متوسط ، ضعيف) و مستوى التحصيل الدراسي (جيد ، متوسط ، ضعيف) . المتغيرين المتضمنين في الفرضية من المستوى الإسمي حيث تم التعبير عنهما بصيغة المستوى ، غير أنه يمكن ترتيب الأفراد وفقا لهذين المتغيرين او التعبير عنهما بصيغة كمية . نلاحظ أيضا ان التقسيم غير حقيقي و المتغيرين لهما اكثر من تقسيمين . و تعالج البيانات من هذا النوع بواسطة معامل كاف مربع ، الذي يقيس الدلالة الإحصائية للفروق و عندما يكون هناك متغيرين يسمى باختبار الإستقلالية ، و هذا المعامل يسمح بالقول بان متغير مستقل عن متغير آخر، أي لا يؤثر أحدهما في الآخر ، و يكون التأثير عندما درجة الفروق كبيرة و دالة احصائيا

اختبار كاف تربيع:

$$\chi^2 = \sum \frac{(fo - fe)^2}{fe}$$

ليات و ،البيانات نوعية

المحاضرة الخامسة:تكملة للمعالجات الإحصائية (المعالجات الإحصائية في المستوى

الرتبي و المسافات المتساوية

اولا :المعالجات الإحصائية في المستوى الرتبي :المستوى الرتبي ، البيانات هنا

عبارة عن رتب و الهدف هنا هو ترتيب الأفراد وفقا لمعايير معينة ، و هنا يمكن أن

نجد الحالات التالية :

_معامل ارتباط سبيرمان: يستخدم في حالة ترتيب المجموعة الواحدة وفق درجاتهم

على متغيرين إثنين (الترتيب وفقا لدرجات الإحصاء و درجات المنهجية) في هذه

الحالة نستعمل معامل ارتباط سبيرمان . و تصاغ الفرضية كما يلي :

-توجد علاقة بين رتب الأفراد في وحدة الإحصاء و رتبهم في وحدة المنهجية

معامل ارتباط سبيرمان:

$R_s = 1 - \frac{6(\sum D)^2}{N(N^2 - 1)}$	معامل سبيرمان
--	------------------

$r_s = \frac{\sum x^2 + \sum y^2 - \sum d^2}{2 \sqrt{\sum x^2 \cdot \sum y^2}}$	معامل سبيرمان المعدل
---	----------------------------

معامل ارتباط كندل:يستعمل فيفي حالة ترتيب المجموعة الواحدة وفقا لدرجاتهم على ثلاث

معايير معينة .

تصاغ الفرضية كالتالي : توجد علاقة بين رتب الأفراد في الوحدات التالية : الإحصاء

، المنهجية ، القياس .

ملاحظة: في حالة وجود عدد الرتب المشتركة يزيد عن 25/ ، نستعمل معامل ارتباط سبيرمان المعدل ، ونستعمل معامل التعديل .

معامل ارتباط كندل:

$$kendall's = \frac{\Sigma D}{1/2n(n-1)}$$

ثانيا: المعالجات الإحصائية في. مستوى المسافات المتساوية :

البيانات في هذا المستوى كمية مع افتراض التساوي بين الوحدات كما تمت الإشارة إلى ذلك فيما سبق ، و مجمل المتغيرات محل الدراسة في العلوم الاجتماعية تنتمي إلى هذا المستوى و ذلك بفضل عملية القياس التي اكسبت الموضوعية للبحوث الاجتماعية حيث انه أصبح بالإمكان التعبير عن الظواهر النفسية بواسطة الأرقام او بعبارة اخرى تكميم الظواهر. (علما ان لغة الأرقام هي لغة الدقة و الموضوعية) .

و في هذا المستوى سوف نميز بين نوعين رئيسيين من المعالجات الإحصائية ، تلك التي تهتم بدراسة العلاقات الارتباطية البسيطة بين متغيرين و المعقدة بين أكثر من متغيرين ، و تلك التي تهتم بدراسة الفروق بين مجموعتين او بين عدة مجموعات .

المعالجات الإحصائية التي تهتم بدراسة العلاقات الارتباطية

1_ العلاقة الارتباطية بين متغيرين

يستعمل عندما يفترض الباحث ان أي تغير في المتغير الأول يتبعه تغير في المتغير الثاني .
البيانات عبارة عن درجات خام (كمية) و حجم العينة لا يقل عن 50 فردا و يجب التأكد من
خطية العلاقة باستخدام معامل إيتا ، وعليه فإن معامل إيتا يستخدم مع المتغيرات الكمية
المتصلة ، كما يصلح لقياس العلاقة الارتباطية غير الخطية ، كما يستعمل كما تم الإشارة من
قبل لمعرفة خطية العلاقة .

الفرضية بالنسبة لمعامل ارتباط بيرسون تصاغ على النحو التالي : توجد علاقة دالة احصائيا
بين درجات الطلبة على مقياس التوافق النفسي و درجاتهم على مقياس تقدير الذات .
دلالة قيمة معامل ارتباط بيرسون : تحسب بدلالة ت .

معامل ارتباط بيرسون:

$R_p = \frac{N \sum (x.y) - (\sum x) . (\sum y)}{\sqrt{[N \sum x^2 - (\sum x)^2][N \sum y^2 - (\sum y)^2]}}$	معامل برسون
--	----------------

معامل الارتباط بين اكثر من متغيرين (الارتباط المتعدد) :

معامل الارتباط المتعدد يستعمل لقياس العلاقة الارتباطية بين أكثر من متغيرين في
مستوى المسافات المتساوية ، له نفس شروط معامل ارتباط بيرسون . وهو يفيد في
معرفة العلاقة بين متغير مستقل واحد و متغيرين تابعين ، وبالتالي فهو ينطلق من
مصفوفة الارتباطات .

الفرضية تصاغ على النحو التالي : توجد علاقة ارتباطية بين دافعية الإنجاز و الإتجاه نحو التدريس ، وتقدير الذات .

دافعية الإنجاز هي المتغير المستقل و نعطيه الرقم 1. الإتجاه نحو التدريس هو المتغير التابع الأول و نعطيه الرقم 2 وتقدير الذات هو المتغير التابع الثاني و نعطيه الرقم 3 . تصبح بعد ذلك الصيغة الرياضية لمعامل الارتباط المتعدد كما يلي.

ملاحظة : إن معامل الارتباط الذي نحصل عليه باستخدام الارتباط المتعدد غالبا ما لا يدل على ان النتيجة المحصل عليه.

معامل الارتباط المتعدد:

$$R = \sqrt{\frac{r_{x1y}^2 + r_{x2y}^2 - 2(r_{x1y} \cdot r_{x2y} \cdot r_{x1x2})}{1 - r_{x1x2}^2}}$$

أو يكتب
معامل الارتباط المتعدد

$$R = \sqrt{\frac{r_{1y}^2 + r_{2y}^2 - 2(r_{1y} \cdot r_{2y} \cdot r_{12})}{1 - r_{12}^2}}$$

$$F = \frac{R^2/p}{(1 - R^2)/(n - p - 1)}$$

المعالجات الإحصائية التي تهتم بدراسة الفروق :

أولا : المعالجات التي تهتم بدراسة الفروق بين مجموعتين فقط :

يمكن التمييز بين مختلف المعالجات التي تدرج ضمن هذا المحور تبعا للتصميم التجريبي المعتمد و طبيعة المجموعتين محل الدراسة من حيث كونها مرتبطة أو غير مرتبطة ، او من حيث كونها متجانسة او غير متجانسة . .

__ عندما تكون للباحث عينة واحدة يطبق عليها اختبارا قليا و اختبارا بعديا .

__ عندما تكون للباحث مجموعة واحدة يلاحظها في وضعيتين مختلفتين .

__ عندما تكون لدى الباحث عینتان مختلفتان و لكنهما متشابهتين في بعض الخصائص ، غير انه في مثل هذه الحالة ، يجب على الباحث ان يتأكد من ان هذه الخصائص ترتبط ارتباطا وثيقا بالمتغير التابع.

في هذه الحالة تحسب دلالة الفروق بين المتوسطين بطريقة الفروق . في حالة العينة الواحدة ، يستعمل في التصميم التجريبي من خلال قياس واحد . أو في حالة تطبيق اختبار قبلي واختبار بعدي على نفس العينة .

حالة عينتين مستقلتين غير متجانستين : نقصد بالعینتين غير المتجانستين ، عندما يكون تباين العينة الأولى يفوق تباين العينة الثانية بأكثر من مرتين . في حالة عدم تساوي حجم العينتين يجب اولا اختبار التجانس بحساب قيمة ف التي هي النسبة بين التباين الكبير و التباين الصغير

ونقصد بالتجانس ، تجانس التباينات أي أن تشتت درجات العينة الأولى متقارب مع تشتت درجات العينة الثانية .

تصاغ الفرضية في هذا النوع من المعالجات كما يلي :

توجد فروق بين درجات الذكور و درجات الإناث في مادة الإحصاء.

حالة عينتين مستقلتين متجانستين : و العينتين المتجانستين هما عينتان مختلفتان من حيث الأفراد و كذلك من حيث الخصائص (خاصية أساسية مهمة لها علاقة بالمتغير الذي نقيسه.

و العينتان المتجانستان هما عينتان متساويتان من حيث العدد أو الحجم ،وكذلك من حيث التباين .وفي حالة ماختلف حجم العينتان يجب اختبار التجانس عن طريق اختبار ف ،نسبة إلى فيشر

إن قيمة ف لا توضح بالضبط ما هي المقارنات الدالة و المقارنات غير الدالة ، لذلك يجب عند اكتشاف دلالة ف لابد من اختيار المقارنات او ما يسمى بتحليل المقارنات و هو اختبار شيفر المحافظ لإختبار الدلالة الإحصائية لمختلف المقارنات .

ملاحظة : نسمي المتغير في تحليل التباين بالعامل ، و العامل يجب ان يكون لديه اكثر من مستويين

مثلا نقارن بين ثلاث مجموعات مختلفة في مستويات القلق و نتائج كل مجموعة في التحصيل الدراسي.

تصاغ الفرضية كما يلي : هل تؤثر مستويات القلق (منخفض / متوسط / مرتفع) على درجات التحصيل الدراسي .

او : هل هناك فروق ذات دلالة إحصائية في درجة التحصيل الدراسي باختلاف مستويات القلق (منخفض/ متوسط/

إختبار ف لتحليل التباين: هو عبارة عن نسبة التباينين :

1_ تباين يسمى تباين ما بين المجموعات او العينات المختلفة و الذي يعود أساسا إلى الفروق بسبب أثر المتغير المستقل 2_ تباين يسمى تباين داخل المجموعات و هو راجع إلى الفروق الفردية و اخطاء القياس .

3_ كلما كان التباين ما بين المجموعات اكبر من التباين داخل المجموعات ، نقول ان هناك تأثير للمتغير المستقل على المتغير التابع ، و منه نرفض الفرضية الصفرية .

اختلف حجم العينتان يحب اختبار التجانس عن طريق اختبار ف ، نسبة إلى فيشر .

تصاغ الفرضية في هذه الحالة كما يلي :

توجد فروق ذات دلالة احصائية بين درجات القلق لدى طلبة كلية الإقتصاد و طلبة كلية الحقوق .

المتجانستين هما عينتان مختلفتان من حيث الأفراد وكذلك من حيث الخصائص (خاصة أساسية مهمة لها علاقة بالمتغير الذي يقيسه الباحث

. إن قيمة ف لا توضح بالضبط ما هي المقارنات الدالة و المقارنات غير الدالة ، لذلك يجب عند اكتشاف دلالة ف لابد من اختيار المقارنات او ما يسمى بتحليل المقارنات و هو اختبار شيفر المحافظ لإختبار الدلالة الإحصائية لمختلف المقارنات .

ملاحظة : نسمي المتغير في تحليل التباين بالعامل ، و العامل يجب ان يكون لديه اكثر من مستويين

مثلا نقارن بين ثلاث مجموعات مختلفة في مستويات القلق و نتائج كل مجموعة في التحصيل الدراسي

تصاغ الفرضية كما يلي : هل تؤثر مستويات القلق (منخفض / متوسط / مرتفع) على درجات التحصيل الدراسي .

المحاضرة السادسة : صلاحية ادوات القياس (الخصائص السيكميتريية) :

أولا الصدق

الأهداف :

- التعرف على مفهوم الصدق ، خصائصه،انواعه و طرق التاكّد منه و العوامل المؤثرة فيه

مدخل :

تعتمد دقة اتخاذ القرار في أي مجال من المجالات الإنسانية على نوعين من البيانات ومدي تغطيتها للمتغيرات التي يتكون منها المجال موضوع البحث. وهذه البيانات هي نتيجة الأداء على ادوات القياس المقننة. وعند استخدام هذه البيانات في اتخاذ القرارات أو في صنع القرار علينا التاكّد من خصائصها من حيث كونها ثابتة ، متناسقة، موضوعية صادقة ، لأن ذلك يساعد في ان تكون الإستنتاجات التي نشقها من تلك البيانات صادقة .

يجب ان تتصف أدوات القياس بخصائص و تتوفر فيها شروط ، تؤخذ بعين الاعتبار خلال عملية بنائها و تطبيقها ، وتحليل نتائجها ، و إذا لم تتوفر مثل هذه الخصائص و الشروط في أداة القياس فإنه لا يمكن ان نتق فيما نتوصل إليه من بيانات و لا إذا كانت فعلا تقيس ما وضعت لقياسه . لذلك و قبل تطبيق أي أداة قياس علينا اختبار خصائصها السيكميتريية و منها التاكّد من صدق الأداة، بحيث يجب ان تكون أداة القياس صادقة في تمثيلها للمجتمع الأصلي ، شاملة لجميع مكونات القدرة او السمة أو الخاصية المراد قياسها.

دلالات الصدق : للصدق دلالات نذكر منها :

1_ أن تكون الأداة قادرة على قياس موضوع لقياسة ،بمعنى ان تكون وثيقة الصلة بالقدرة التي تقيسها .

2_ ان تكون الأداة محددة لقياس موضوع لأجله و لاتقيس شيئاً غيره .

3_ أن تكون اداة القياس قادرة على إظهار الفروق الفردية . (سعد عبد الرحمان، 1424هـ-

2003

خصائص الصدق: إن الصدق من المعالم الرئيسية و الهامة التي تقوم عليها الإختبارات و المقاييس في البحوث التربوية و النفسية ،كون ان القياس في هذا الميدان من البحوث هو قياس غير مباشر و يحتاج لكثير من الدراسة و التحليل ،ولذلك يجب التأكد من ان ماتقيسه أدوات البحث يمكن الوثوق بها و الإعتماد عليها في جمع البيانات .

و الصدق يرتبط بالهدف أو الأهداف المتوقع في أداة القياس تحقيقها ،وكذلك بمدى اتصاله بنوع و اهمية القرار الذي سيتم اتخاذه تبعا لذلك.

و الصدق مرتبط بمحتوى أو سلوك محدد،وموجه إلى فئة معينة من الأفراد.وهذا يعني ان الصدق مفهوم نسبي وليس مطلق.ومن جهة اخرى فإن الصدق لايرتبط باداة القياس نفسها ،بل بالطريقة التي سوف تفسر بها الدرجات المستخرجة من تلك الأداة.

وبالتالي فالتأكد من الأهداف التي توضع لقيسها الإختبار من حيث الصياغة السليمة لتلك الأهداف ،من ناحية وتحديد الهمية النسبية لكل منها ،ومدى مناسبتها لمستوى المستهدفين من

ناحية اخرى ،يعد بداية رئيسية في الحكم على صدق الإختبار. إضافة إلى دقة تحديد المجال السلوكي الذي يتوقع ان تمثله اداة القياس ،والصياغة الدقيقة لل فقرات .

ومنه يمكننا استخلاص خصائص الصدق فيما يلي :

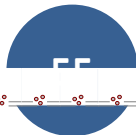
1 _ يتوقف الصدق على الغرض من الإختبار أو الوظيفة التي ينبغي أن يقوم بها .

2 - يتوقف على الفئة المستهدفة التي سيطبق عليها الإختبار.

3 - الصدق صفة نسبية و ليست مطلقة .

4 - يتوقف صدق الإختبار على ثباته ،وبالتالي فإن الثبات ضروري للمقياس و لكن لا يكون بديلا عن الصدق ،لأن انخفاض معامل الثبات يؤثر على صدق الإختبار .

إذن كانت البدايات الأولى لإستخدام الصدق تركزت على التأكد مما إذا كان الإختبار يقيس فعلا ماوضع لقياسه ،و هذا يعكس التعريف الكلاسيكي لمفهوم الصدق . غير ان التعريف الحديث يرتبط بالإستخدامات المختلفة لدرجات الإختبار و ليس بالإختبار نفسه.و عليه فإن مفهوم الصدق هو صدق استخدام الدرجات و ليس صدق الإختبار .و قد عرفه على أنه "حكم تقييمي شامل ،مبني على أدلة تجريبية و مبررات نظرية ،على مدى كفاية و ملائمة الإستدلالات و الإجراءات المستندة إلى درجات الإختبار" هذا المفهوم اعتمد فيما بعد من قبل وثيقة معايير العمليات الإختبارية النفسية و التربوية ، في اشتقاق التعريف التالي للصدق " على انه الدرجة التي تؤيد بها الأدلة و النظرية في تفسير درجات الإختبارات التي تتطلبها استخدامات مقترحة للإختبار"



يمكن توضيح الصدق في النقاط التالية:

- يشير الصدق إلى تفسير نتائج الإختبار و ليس الإختبار نفسه.

يستنتج الصدق من الأدلة المتوفرة وليست المناسبة.

-الصدق خاص باستخدام درجات الإختبار مثل استخدام الدرجات في اختيار عينة الأفراد أو توزيعهم أو تقويم أدائهم.

يعبر عن الصدق بدرجة وصفية مثل مرتفع ،منخفض، و قد تستخدم الأرقام مثل معاملات الارتباط أو الإتفاق لتوضيح الدرجة الوصفية (عبد السلام دعيدش،2018:

messick.1988.(32

أنواع الصدق : للصدق عدة انواع نذكر منها :

الصدق التجريبي،أو الصدق المرتبط بالمحك :وهو عبارة عن صدق الإختباركما يعين تجريبيا ،او كما يعبر عنه بمعامل الارتباط بين الإختبار وبين محك خارجي تاكدنا من صلاحيته (صادق و ثابت) .(سعد عبد الرحمان/1424هـ،2003) .

كما يقصد به فعالية اختبار ما في التنبؤ بأداء الفرد ،ويقاس في ضوء صلتة او ارتباطه بمحك ما و يصنف وفقا للغرض من استخدامه إلى :

أ_ الصدق التلازمي: ويقاس بالعلاقة بين الإختبار و محك خارجي في نفس الوقت

ب_ الصدق التنبؤي: وهم يعتمد على الإختبار في التنبؤ بانمط سلوك الفرد مستقبلا، وخاصة إذا كان هذا الموقف مستقبلي يتعلق بما يقسيه الإختبار. (سعد عبد الرحمان 1424هـ، 2003م)

الصدق العاملي: ويعتمد هذا النوع من الصدق على منهج التخليل العاملي الذي يقوم على تحليل مصفوفة معاملات الارتباط بين الإختبارات و المحكات المختلفة من اجل الوصول إلى العوامل التي أدت إلى إيجاد هذه الارتباطات. (سعد عبد الرحمان 1424هـ، 2003م)

الصدق الإفتراضي أو صدق التكوين الفرضي:و يقوم على افتراض من قام ببناء الإختبار، بأن هذا الإختبار يقيس قدرة معينة وذلك على ماورد فيه من بنود او وحدات او تعليمات. (سعد عبد الرحمان 1424هـ-2003م)

و يقصد به أيضا نجاح الإختبار في قياس مفهوم نظري لسمة معينة،و يتوقف على مقدار مانحصل عليه من معلومات في هذه السمة أو القدرة و خصائصها و مكوناتها، وذلك بإظهار القياسات و الدرجات التي يمكن أ تستخدم للوصول إلى استنباطات متسقة مع النظرية او الأبعاد النظرية التي وضع الإختبار على أساسها (بوسالم، 68:2012_69)

صدق المحتوى : و يقوم هذا النوع من الصدق على مدى تمثيل الإختبار او المقياس للأبعاد المختلفة للقدرة التي يقيسها. كما تم تحديدها في التعريف الإجرائي الذي تبناه صاحب الإختبار أثناء تعريفه للسمة المراد قياسها. (سعد عبد الرحمان، 1424هـ، 2003م). وكلما ارتفع مستوى تمثيل العينة من السلوك كلما ارتفع صدق المحتوى ويندرج ضمنه الصدق الظاهري الذي يقوم على فكرة مدى مناسبة الإختبار لما يقيس ولمن يطبق عليهم. ويبدوا مثل هذا النوع أيضا

في وضوح البنود ، ومدى علاقتها بالقدرة أو السمة أو الخاصية أو البعد الذي تقيسه .و غالبا مايقرر ذلك مجموعة من المتخصصين في المجال الذي يفترض أن يتمي إليه هذا الإختبار.

صدق البناء: و هو مدى قدرة بناء الإختبار ككل على قياس متغير ما بجميع جوانبه، و هو يعتمد أساسا على صدق المحتوى .

الصدق التمييزي: و يقصد به قدرة الإختبار على التمييز بين المجموعات المختلفة ،او حتى الأفراد الذين تقع درجاتهم على طرفي المنحنى .

الصدق الذاتي : يعرف بانه صدق الدرجات التجريبية للإختبار بالنسبة للدرجات الحقيقية التي خلصت من أخطاء القياس ،و بذلك تصبح الدرجات الحقيقية للإختبار هي المعيار الذي ينسب إليه صدق الإختبار (حسن علاوي،نصر الدين رضوان ،1999).

ملاحظة : التصنيفات او التسميات السابقة لأنواع الصدق تتدرج ضمن النظرية الكلاسيكية ،وان هذا التصنيف لايزال شائعا إلى يومنا هذا .وبعدها كانت هناك عدة مراجعات نظرية في هذا الشأن منها التي قام بها تيغزة (2008) . r

المحاضرة السابعة: نظريات الصدق

الأهداف:

- أن يعرف الطالب مراحل نظرية الصدق الكلاسيكية

__ أن يعرف الطالب الفرق بين النظرية الكلاسيكية و النظرية الحديثة في الصدق

أولاً: نظرية الصدق الكلاسيكية:

المرحلة الأولى: النصف الأول من القرن العشرين ، وتميزت بظهور أصناف عديدة و

تسميات متباينة لأنواع الصدق ، برزت على إثرها منظورين أساسيين في مجال الصدق :

- **الإتجاه الإمبيريقى (التجريبي)** : والذي روج للصدق المحكى أو صدق الإرتباط بالمحك، الصدق العاملي.

- **الإتجاه المنطقي**: يعتبر المنطق و الحكام الفردية (الخبرة) كاطر مرجعية لتقدير الصدق. وانبثق عن هذا الإتجاه صدق المحتوى.

المرحلة الثانية: تميزت بصدور الدليل الإرشادي و الفني للقياس النفسي و الإختبارات و الذي جاء نتيجة التغيرات التي طرات على مفهوم الصدق في اوساط الباحثين و كذا نتيجة تتويجا لإنضمام أعضاء الجمعية الأمريكية لعلم النفس التطبيقي إلى الجمعية الأمريكية لعلم النفس.

في عام 1954 صدر الدليل الأول (apa;aera) بعنوان " التوصيات الفنية للاختبارات النفسية وفنيات أدوات التشخيص " قدم هذا الدليل تصنيف رباعي لأنواع الصدق :

صدق المضمون.

الصدق التنبؤي.

الصدق التلازمي.

صدق التكوين الفرضي او المفهوم.

يعتبر صدق المفهوم إضافة جديدة في ميدان القياس في تلك الفترة ،وكان وراء هذا المفهوم وتوضيح دلالاته فيما بعد كل من cronbach et meeh (1954).

في عام 1966 صدور الدليل الإرشادي الثاني المنقح (apa ;aera) حيث قلص هذا الدليل التصنيف الرباعي للصدق إلى تصنيف ثلاثي :

صدق المضمون

صدق المحك

صدق المفهوم.

ظهر في هذا التصنيف تسمية صدق المحك كتسمية جديدة تضم الصدق التنبؤي و الصدق التلازمي.

في عام 1974 ظهر الدليل الإرشادي الثالث للقياس، هذا الأخير يشير إلى أنواع الصدق باستخدام لفظ "جوانب الصدق" و على هذا فالصدق ليس خاصية تقاس بطريقة مباشرة و

إنما يستنتج من دراسة جوانب أو أبعاده الثلاثة، و ان هذه الأبعاد مترابطة فيما بينها من الناحية الوظيفية و المنطقية. كما يمكن دراستها على انفراد. (عبد السلام دعيش ، 2018: 33).

في عام 1985 ظهر الدليل الإرشادي الرابع ، وأبرز ما نثيق عنه هو اعتبار الصدق مفهوم لا يمكن تجزئته ، و تبنى تعريف جديد للصدق يحمل تغييرا في المفهوم و الإستخدام " الصدق يدل على ملاءمة و دلالة وجدوى الإستنتاجات الخاصة المستمدة من الدرجات " واستعمال لفظ "أدلة الصدق" بدل أنواع الصدق". كما أبقى على التصنيف الثلاثي لأنواع الصدق (صدق المحتوى، صدق المحك، صدق المفهوم). (تيغزة، 2008).

إن نظرية الصدق الحديثة قد دعت في ظل المعايير الجديدة لأدلة الصدق إلى ضرورة توحيد مفهوم الصدق بعد ان كان منقسما إلى عدة أنواع من الصدق في النظرية الكلاسيكية حيث يصعب على الباحثين التحديد الصحيح لنوع الصدق الملائم. و بذلك تفقد معها النظرة الكلية و التكاملية للصدق و تضع الجهود في البحث بين انواع الصدق الملائمة .

كما ان من بين الإشكالات التي طرحتها النظرية الكلاسيكية للصدق ، الترويج للإعتقاد الخاطئ بأن الصدق هو خاصية ثابتة للأداة و ليس خاصية لعينة tracey et (1999: glidden) .

و عليه تخلص النظرية الحديثة في الصدق إلى اعتباره أي الصدق مفهوم وحدوي و أن صدق التكوين الفرضي هي في الواقع كل متكامل . و بالتالي ضرورة إعطاء نفس الأهمية لأنواع الصدق في جمع الأدلة لدعم الإستدلالات التي سيتم إجراؤها من نتائج الإختبارات.

ثانيا: نظرية الصدق الحديثة

إلى غاية الإصدار الخامس سنة 1999، و الذي حمل تطورات حدثت على مستوى مفهوم الصدق، وكذا اعتباره موحدا و لاينقسم إلى انواع (messck 1994)، و ان هذه الوحدة يتضمنها صدق المفهوم و الذي يعنى بدمج متكامل لمصادر أدلة الصدق التي تستهدف تفسير و تاويل درجات الإختبار. و هذه النظرة الوحدية لا تتنافى مع وجود أدلة متميزة في تصنيف مصادر أدلة الصدق و طبيعتها. والتي سوف نعرض لها في العنصر اللاحق.

العوامل التي تؤثر على صدق الإختبار

يتأثر صدق الإختبار بعدد من العوامل يؤثر بعضها بشكل مباشر على نوع معين من دلالات الصدق، و بشكل غير مباشر على انواع اخرى و قد صنفنا هذه العوامل على النحو التالي:

1-عوامل تتعلق بالإختبار نفسه : و هذه العوامل تتعلق بإعداد الإختبار من حيث لغته و إجراءات تطبيقه و تصحيحه و صياغة فقراته، و سهولة أو صعوبة الفقرات، و طول الإختبار. وكذلك مدى وضوح تعليماته .

2-عوامل تتعلق بتطبيق الإختبار و تصحيحه : و هي ترتبط بعدم إعطاء الوقت الكافي للإجابة و اختلاف معايير التصحيح أو عدم الإلتزام بها. و تدخل أيضا الشروط الخارجية لعملية تطبيق الإختبار .

3-عوامل تتعلق بإجابات الأفراد على الإختبار: وهي تتعلق بنمط الإستجابة على فقرات الإختبار. وكذلك المتعلقة بالناحية النفسية و الصحية. (أبو زينة، 1992).

إضافة إلى ذلك فإن الصدق يتأثر أيضا بعدة عوامل من أهمها : المحك الذي يستخدم لحساب
صدق الإختبار، العمر ، الجنس ، النضج ، خبرة التعلم للأفراد الذين يحسب عليهم الصدق ،
ويجب علينا أن نضع في الإعتبار مثل هذه العوامل عند حساب صدق أي اختبار من
الإختبارات (حسن علاوي ،نصر الدين رضوان،1979

إذن للصدق أهمية قصوى في بناء الإختبارات النفسية ،تتعدى المحتوى الداخلي للإختبار

المحاضرة الثامنة: مصادر أدلة الصدق:

الأهداف:

- أن يعرف الطالب مختلف الإجراءات أو الأدلة التي يعتمد عليها لتقدير صدق الإختبار. (محتوى الإختبار، سيرورات الإستجابة، البنية الداخلية، البنية الخارجية).
- أن يميز الطالب بين إجراءات تحليل البنية الداخلية (ارتباط البند بالدرجة الكلية، الارتباط بين الأبعاد، التحليل العاملي الإستكشافي، الإتساق الداخلي) و البنية الخارجية (الصدق التقاربي، الصدق التمايزي، الصدق المحكي: التلاومي و التنبؤي)

1 - أدلة تعتمد على محتوى الإختبار:

تعتمد هذه الطريقة على فحص العلاقة بين محتوى الإختبار و التكوين الفرضي (النطاق الذي صمم الإختبار من قياسه. فصدق المحتوى نحصل عليه من خلال التحليل المنطقي _ التقديري للبنود و صيغة المقياس. وحسب ; goodwin et leech 2003) يستند هذا النوع من أدلة الصدق على التحليلات المنطقية وتقييمات الخبراء لمحتوى الإختبار، والتي تشمل البنود من حيث المهام و الصياغة و العمليات المطلوبة من المختبرين. بشكل عام، فإنه يعالج الأسئلة حول المدى الذي يمثل فيه محتوى المقياس نطاق محتوى محدد.

وتهتم هذه الطريقة على الإجراءات التالية :

القيام بتحليلات منطقية ومراجعة الخبراء لفحص مدى تمثيل محتوى الإختبار للنطاق

(التناظر بين محتوى الإختبار و التكوين الفرضي)

التحليلات المنطقية و مراجعة الخبراء لفحص مدى ملاءمة البنود او المهام أو الأجزاء الفرعية من المقياس لتعريف التكوين الفرضي.

التحليلات المنطقية ومراجعة الخبراء لفحص نجاعة و اهمية ووضوح البنود وابتعادها عن التحيز.

التحليلات المنطقية و مراجعة الخبراء لفحص ماقد ينتج عنه نقص التمثيل أو الجوانب غير ذات الصلة من المقياس (goodwin et leech 2003).

وإضافة إلى كل الإجراءات السابقة التي كانت نوعية ،توزجد إجراءات أخرى لتقرير النتائج بطريقة أكثر كمية .ففي هذا السياق أشار (reynolds et livingston :2013) إلى أنه يمكن للباحثين تقرير عدد الخبراء و مؤهلاتهم ،عدد مرات مراجعة التقييم و تعليقاتهم عليه،و درجة اتفاقهم على قضايا تتعلق بالمحتوى .و في هذا السياق يمكن استعمال معادلة lawche 1975 التي تعكس درجة اتفاق بين الخبراء الذين يصدرون احكاما تتعلق بالمحتوى.

2- أدلة تعتمد على عمليات او سيرورات الإستجابة:

يذكر whiston 2009 أن أدلة الصدق التي تستند على عمليات الإستجابة " تركز على ما إذا كانت طريقة الإستجابة لبنود الإختبار تتوافق مع الهدف من عملية القياس .بمعنى التوافق بين الأنشطة العقلية و الإنفعالية التي يتكون منها المفهوم ، و بين الأنشطة العقلية و الإنفعالية التي يمارسها الفرد عند إجابته على فقرات الإختبار.و من الإجراءات التي يمكن أن تتيح هذا النوع من الأدلة عن الصدق مايلي:

- تحليل استجابات الأفراد بطريقة التفكير المسموع أثناء حل المشكلة او وصف
للمعاملات و الإستراتيجيات التي استخدموها في الوصول إلى الحل لمطابقتها مع
المعاملات التي يحددها التكوين الفرضي المراد قياسه.

- فحص المنطق الذي قامت عليه استجابات المفحوصين ،وكذا طرق إجاباتهم على
البنود من خلال تحليل أسبابها و دواعيها.

- بالنسبة للتقويم البديل الذي يتبنى فقرات أو مهمات أو أنشطة تقويمية واقعية ،فلا بد
من الإطمئنان بأن المصححين يستخدمون محكات التقدير بطريقة دقيقة ،وأنها لم تتأثر
بعوامل خارجية كالتحيز ،أو تاويل محكات التصحيح تاويلا خاطئا (تيغزة،2008).

_إجراء مقابلة مع المختبرين لمعرفة استراتيجيات استجاباتهم ،وتدوين مؤشرات
سلوكية :مثل زمن الإستجابة، حركات العين،أو تحليل أنواع الأخطاء التي يقعون فيها

- أدلة تعتمد على البنية الداخلية :

من خلالها يتم تحديد ماإذا كانت العلاقات بين مفردات الإختبار و عناصره متسقة مع
التكوين الفرضي المقترح تفسير درجاته(1999: et ncime :apa ;aera).وتتم هذه
المعاملة بعدة طرق لتحليل البنود كميا:

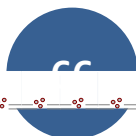
- تقدير الارتباط بين البنود و الدرجة الكلية للمقياس.

-تقدير الارتباط بين أبعاد التكوين الفرضي.

- التحليل العاملي الإستكشافي.

- التحليل العاملي التوكيدي.

-الإتساق الداخلي للبنية الداخلية للإختبار .



هذا النوع من أدلة الصدق تم تضمينه تقليديا تحت عنوان صدق التكوين
الفرضي. (reynolds et livingston :2013).

أدلة تعتمد على أدلة البنية الخارجية:

و يتم الحصول عليها من خلال فحص و دراسة العلاقة بين درجات الإختبار و
درجات مقاييس أخرى أو محكات ، ودراسة ما إذا كان مستوى هذه العلاقات و طبيعتها
أو اتجاهها ينسجم و دلالة المفهوم المقاس، ويندرج ضمن هذا النمط من الأدلة :
الصدق التقاربي، والتمايزي، الصدق المحكي سواء التلازمي أو التنبؤي). واهم
إجراءات الحصول على أدلة الصدق القائمة على البنية الخارجية مايلي goodwin
et leech (2003: و (تيغزة، 2008)، و هذه الأدلة وردت في التصنيفين الصادرين
عن (1999: 1985: ncime ; aera :

1_ الدراسة الارتباطية لقوة و اتجاه العلاقة بين درجات الإختبار المراد تقدير صدقه
و بين متغيرات خارجية تكون محكية ، ويعرف هذا النوع من الصدق بالصدق
التلازمي وفقا للنظرية الكلاسيكية.

2_ الدراسة الارتباطية لمعرفة مدى قدرة الدرجات على التنبؤ مع المحكات الخارجية
التي يتم قياسها لاحقا . ويعرف هذا النوع من الصدق بالصدق التنبؤي. فالدراسات
التلازمية تكون ملائمة بدرجة كبيرة عندما يكون الغرض من الإختبار تحديد الوضعية
الراهنة للفرد المختبر، في مقابل التنبؤ بالأداء المستقبلي anastasi et
(1997: uraina). و يستخدم في تقدير العلاقة بين الدرجات على الإختبار بمحك
خارجي .

3_ دراسة قوة العلاقة الارتباطية و اتجاهها بين درجات الإختبار بالمتغيرات الأخرى أو اختبارات أخرى تقيس نفس المفهوم ،و التي ينبغي ان ترتبط بها ارتباطا قويا بناء على المنطق النظري الذي قام عليه المقياس. ويسمى هذا النوع من الصدق بالصدق التقاربي.

4_ دراسة قوة العلاقة الارتباطية ،واتجاهها بين درجات المقياس بالإختبارات الأخرى تقيس مفاهيم أخرى مختلفة عن المفهوم الذي يقيسه الإختبارو التي ينبغي ان يرتبط بها ارتباطا ضعيفا،بناءا علنت المنطق النظري الذي قام عليه الإختبار أو المقياس.و يسمى هذا النوع من الصدق بالصدق التمايزي.

5_ الدراسة التقييمية لفعالية عملية الإنتقاء المهني وقرارات الوضع و التسكين .

6_ تقدير مدى مواءمة المحكات و جودتها ،بما في ذلك تقدير ثبات درجاتها و صدقها ،والصعوبات العملية عند استعمالها،وامكانية توفرها .وتتأتى اهمية هذه المحكات من امكانية الحكم على صدق الأداة بناءا على مدى ارتباطها بها .

أدلة تعتمد على عواقب العملية الإختبارية :

هذا النوع من أدلة الصدقيتعلق بالنتاج المتوقعة و غير المتوقعة للقياس على حد سواء الإيجابية او السلبية، وقد ظهرت فقط في الإصدار الخامس (1999) من دليل معايير العمليات الإختبارية النفسية و التربوية.منبثقة من اعمال (messik :1989). وتدعى أحيانا بصدق المترتبات أو النتائج.

"عادة ما يتم بناء المقاييس من أجل مثلا اختيار الأفراد في المهن المختلفة بمعنى افنتقاء ،أو تحسين الممارسات التدريسية الصفية ...الهدف الأساسي لتقدير صدق

درجات المقياس تحديد ماإذا تحقيق هذه الأهداف المنتظرة من تطبيق أدوات القياس
يمكن تحقيقها فعلا" (1999: aera et apa et ncime).

إن نظرية الصدق الحديثة قد دعت في ظل المعايير الجديدة لأدلة الصدق إلى ضرورة
توحيد مفهوم الصدق بعد أن كان منقسما إلى عدة أنواع من الصدق في النظرية
الكلاسيكية حيث يصعب على الباحثين التحديد الصحيح لنوع الصدق الملائم. و بذلك
تفقد معها النظرة الكلية و التكاملية للصدق و تضع الجهد في البحث بين أنواع
الصدق الملائمة .

كما أن من بين الإشكالات التي طرحتها النظرية الكلاسيكية للصدق ، الترويج للإعتقاد
الخاطئ بأن الصدق هو خاصية ثابتة للأداة و ليس خاصية لعينة tracey et
(1999: glidden) .

و عليه تخلص النظرية الحديثة في الصدق إلى اعتباره أي الصدق مفهوم وحدوي و
أن صدق التكوين الفرضي هي في الواقع كل متكامل . و بالتالي ضرورة إعطاء
نفس الأهمية لأنواع الصدق في جمع الأدلة لدعم الاستدلالات التي سيتم إجراؤها
من نتائج الاختبارات.

المحاضرة التاسعة: ثبات الاختبار

الأهداف:

ان يعرف الطالب مفهوم الصدق (نظريا و احصائيا)

أن يعرف أنواع معاملات الثبات و ان يميز بينها.

التمييز بين طرق التأكد من معامل الثبات

مفهوم الثبات نظريا

يشير الثبات بالمعنى الشامل إلى أنه ضمان الحصول على نفس النتائج تقريبا إذا أعيد تطبيقه على نفس المجموعة و في نفس الظروف. و هو في ذات السياق يمتد لمفاهيم مثل: الإستقرار و الإتساق في النتائج او الأداء . و أفضل طريقة لمقارنة هذه الدرجات هي حساب معامل الارتباط بين درجات الاختبار في المرة الأولى و درجاتهم في المرة الثانية . ونصل إلى بأعلى قيمة لمعامل الارتباط إذا كانت درجات التطبيق الأول مطابقة و هي نفسها في التطبيق الثاني . و يصبح بالتالي قيمة معامل الارتباط يساوي الواحد الصحيح. لكن المقاييس النفسية و

التربوية لاتصل إلى الدقة مثل هذه (أبو علام ، 2006، البهي السيد، 1978)

كما يشير الثبات إلى الدرجة الحقيقية التي تعبر عن أداء الفرد على اختبار ما ، ومعنى ثبات الدرجة أن المفحوص يحصل عليها في مرات التطبيق سواء بالاختبار نفسه أو بصورة مكافئة له تقيس الخاصية نفسها وسواء اختبر في نفس الظروف أو في ظروف مختلفة لاتتدخل فيها عوامل عشوائية. و يوضح ملحم (2002) بأن الثبات ليس صفة الاختبار بحد ذاته ، بل نتحدث

عن ثبات الدرجات التي نتحصل عليها ، فكلما كان الاختبار ملائماً أو مناسباً للسمة المراد قياسها زاد ثبات الدرجات المتحصل عليها . (بوسالم ، 2012: 78).

إن المقياس الثابت هو المقياس الذي يعطي نفس النتائج إذا أعيد تطبيقه عدة مرات على نفس العينة و في نفس الظروف . و لكن لايمكن الوصول إلى الدقة التامة لوجود احتمال الخطأ مادمنّا نتعامل أو نقيس متغيرات . و باستعمال ادوات قياس يعني بطريقة غير مباشرة . فهو يعني الإستقرار بمعنى انه لو تكررت عملية القياس للفرد الواحد لأظهرت النتائج شيئاً من الإستقرار (غنيم صبري، 2000) .

مفهوم الثبات إحصائياً : من الناحية الإحصائية يشير الثبات إلى النسبة من تباين الدرجة على المقياس التي تشير إلى الأداء الفعلي للفرد ، ويتضمن هذا المعنى إلى امكانية تقسيم الدرجة على المقياس إلى مكونين رئيسيين هما : الدرجة الحقيقية و الدرجة الناتجة عن عوامل الصدفة العشوائية (صفوت، 2007: 197).

يشير علام (2000) إلى انه على الرغم من أن الدرجة الحقيقية و درجات الخطأ العشوائية تعد من المفاهيم الإفتراضية ، إلا أنها تفيد بدرجة كبيرة في التوصل إلى المفهوم الإحصائي للثبات .

فقد تبين انه إذا كانت درجة الخطأ غير مرتبطة بالدرجة الحقيقية فإن تباين الدرجات الملاحظة هي تباين الدرجة الحقيقية مضاف إليها تباين الخطأ .

وبما أننا لانعرف مطلقاً الدرجة الحقيقية حسب عودة (1999) فلا يمكن حساب الثبات.

العوامل التي تؤثر على ثبات الاختبار.

_ **ظروف إجراء الاختبار:** حيث يجب توحيد تعليمات و ظروف إجراء الاختبار من تطبيق لآخر .وإلا فسوف يتأثر معامل الثبات .

_ **تجانس الأفراد الذين يطبق عليهم الاختبار :** فإن ثبات الاختبار يقل بزيادة تجانس المجموعة التي يطبق عليها الاختبار .لأن التباين داخل هذه العينة المتجانسة يكون منخفضا ،بقدر لايسمح بتقدير التباين الحقيقي للاختبار أي ثباته .

فتجانس درجات المجموعة يؤدي إلى انخفاض درجة تشتتها و انحرافاتها عن المتوسط ،وهذا بدوره يقلل من قيمة معامل الثبات الذي يعتمد على قيمة الإنحراف عن المتوسط .

_ **طول الاختبار (عدد بنود الاختبار) :** بحيث انه كلما زاد عدد البنود كلما زاد احتمال تمثيل العينة السلوكية بالاختبار و بالتالي يزداد الثبات .و قد نستنتج ذلك من خلال معادلة سبيرمان و براون لتصحيح الطول ،أن العامل الأول الذي يؤثر في ثبات الاختبار هو طوله ،أي عدد بنوده ،فكلما زاد عدد البنود ارتفع معامل الثبات .و تفسير ذلك ،أن العدد الكثير من البنود يؤدي إلى الحصول على عينة أكبر من السلوك ،وكلما حصلنا على عينة اكبر من السلوك كلما كان من المتوقع أن يمثل السمة محل القياس .هذا من جهة ،ومن جهة اخرى فإن طول الاختبار يؤثر بالتأكيد على تباين الدرجة الحقيقية و الدرجة الملاحظة .فالاختبارات ذات العدد المحدود من البنود تكون اكثر تأثرا بعامل الصدفة (كامل،1996) .لكن إذا زاد عدد البنود فإن ليس هناك تأثير لعامل الصدفة في الإجابة على البنود على الدرجة الكلية

للإختبار. و في هذا الشأن طور سبيرمان _ براون معادلته و التي تسمح بالتحكم في ثبات الإختبار عن طريق زيادة عدد البنود .(وسوف يتم التفصيل في هذه النقطة لاحقا).

إن حساب العلاقة بين ثبات الإختبار و عدد البنود يمكننا من التعرف على قيمة معامل الثبات الذي يمكن الوصول إلى تحقيقه و هذا بعد زيادة عدد البنود ، مع الحرص أن تكون البنود المضافة متسقة مع محتوى بنود الإختبار الأصلي.

_ استقلال بنود الإختبار عن بعضها البعض: لأن الارتباط المرتفع بين بنود الإختبار يقلل من معامل ثباته . ويعني الإستقلال ، ألا تؤدي إجابة معينة على بند إلى إجابة على بند آخر ، لأن هذا يؤدي ضمنيا إلى انخفاض عدد البنود ،ويقلل من الفروق بين الأفراد و هذا ما يؤدي إلى انخفاض معامل الثبات .

فالإختبارات التي تتكون من أسئلة صعبة جدا أو سهلة جدا تكون درجات ثباتها ضعيفة ،لأنها تؤدي إلى انتشار محدود للدرجات و تظهر فروقا طفيفة ،كون أن معامل الثبات يعتمد في حسابه على تباين درجات الإختبار و مدى انتشارها.

_ التخمين في الإجابة : فإن عدم الوضوح في صياغة البنود يبعث على التخمين في الإجابة ،وهذا ما من شأنه أن يقلل من ثبات الإختبار .(خاصة في الإختبارات التحصيلية) .

مستوى صعوبة او سهولة الفقرات : بحيث أن مستوى الصعوبة المرتفع أو مستوى السهولة المرتفع للبنود يؤدي إلى عدم القدرة على التمييز بين الأفراد وهذا يزيد من تجانس الأفراد .وبالتالي يجب ان تكون البنود متدرجة من حيث الصعوبة و السهولة .

موضوعية التصحيح : بحيث كلما كان تصحيح الإختبار موضوعيا كلما زاد ذلك من قيمة معامل الثبات. كما أن درجة الموضوعية بين المصححين تنخفض بالنسبة لاختبارات الأداء أو مقاييس الشخصية لأن التصحيح يتطلب أحكاما ذاتية حول استجابات المختبرين مما يؤثر تأثيرا كبيرا على ثبات التقديرات بحيث تنخفض قيمة معامل الثبات تبعا لذلك.

زمن الإختبار: يزداد معامل الثبات بزيادة الوقت الذي يستغرقه في إجراء الإختبار إلى درجة معينة يرتفع ثم ينخفض الثبات إذا كان الزمن أكبر من اللازم. و بالتالي فإن الزمن المحدد للإجابة على الإختبار يؤثر بشكل مباشر على الثبات. فاختبارات السرعة أو الإختبارات الموقوتة تكون معاملات ثباتها مرتفعة، مقارنة بالإختبارات التي تمنح متسعا من الوقت. ولذلك ينبغي على مصمم الإختبار أن يحدد الوقت المناسب للإجابة، دون أن يعطي متسعا من الوقت للضعفاء في الإجابة حتى ولو كان الإختبار ليس موقوتا .

و هذا ماكدته دراسات و أبحاث (السيد، 1989) من حيث أن إختبارات السرعة تظهر فيها نسبة كبيرة من البنود في آخر الإختبار يتركها المختبرون دون الإجابة عنها ، و هذا ليس بسبب صعوبتها لكن بسبب عامل الزمن وتصحح عموما بدرجة الصفر وينجر عن ذلك ارتفاع في الارتباط بين البنود الخيرة ، ما يمنح اتساق أكبر لهذه البنود . وكما هو معروف أنه كلما كان الإتساق بين البنود مرتفعا كان معامل الثبات مرتفع. cook et linguist.

بالرغم من أن الثبات يمثل جانبا مهما في تحديد نوعية اداة القياس ، إلا انه صفة نسبية و موقفية من خصائص تلك الأدوات ، بمعنى أن اداة القياس ثابتة في ضوء ظروف و معطيات معينة ، و تتغير هذه الصفة ببتغير العينة المستهدفة بالقياس و نوعية العوامل المحيطة بهم .

و عليه يمكن انه من غير الممكن ان نجد مقياسا ثابتا تماما و دائما ، و ان اتباع اجراءات بناء الإختبار التي تم تناولها سابقا تؤدي بدرجة كبيرة إلى التوصل إلى إختبار ثابت .

الثبات واستخدام الدرجات

كثيرا ما يثار تساؤل يتعلق بدرجة الثبات التي يجب ان تتمتع بها أداة القياس حتى تكون جيدة . الأمر هنا يتعلق بالهدف الذي سوف تستخدم فيها الدرجات على تلك الأداة

فإذات كانت القرارات التي يراد استخدام الإختبار في صنعها غاية في الأهمية او لايمكن اصلاحها فيما بعد ، فإن مقدار معامل الثبات عليه تعد مسالة غاية في الأهمية أكثر مما هو الحال في حالة ما إذا كانت القرارات عادية او غير مهمة .

من ناحية اخرى ، إذا كان استخدام الدرجات في صنع قرارات تتعلق بالفرد ، فإنه يجب ان تكون الدرجات أكثر ثباتا مما لو كانت تستخدم في صنع قرارات تتعلق بالجماعات .

تعتمد بيانات الثبات على نوع الإختبار ، و كيفية استخدامه ، ففي اختبارات الإستعدادات ، تكون تقديرات الثبات الأكثر أهمية هي تقدير ثبات الإستقرار ، لن الهدف هنا متعلق بلقدرة على التنبؤ . و هكذا بالنسبة لأنواع الإختبارات الأخرى .مثل الإختبارات غير المعرفية ، فإن انواع الثبات المطلوبة تختلف نسبيا.

معايير الصدق و الثبات في البحوث الكيفية .

-الصدق الكيفي

بصفة عامة، يشير الصدق في البحوث الكيفية، على غرار البحوث الكمية إلى درجة استقلالية الإجابات عن الظروف العرضية للبحث، ولكنه قد يتأتى كذلك من أمانة وعمق وثراء وسعة البيانات وتعددية المصادر؛ أما الثبات فيشير إلى الحد الذي يتم فيه فهمها فهما صحيحا.

ويعتمد صدق البحوث الكيفية، عموما، على حل مشاكلها المتعلقة بالتصميمين الداخلي والخارجي، فالصدق، كما مر معنا، يرتبط أصلا، باحتمال أن يعيد باحثون آخرون بناء استراتيجيات تحليلية أصلية. ومن ثمة، يرتبط الصدق بقابلية تكرار التجارب والاكتشافات العلمية.

ومع ذلك، فإن استعمال الصدق في التصميمات الكيفية (ومنها الاثنوغرافية) يكون معقدا من حيث مؤشر قابلية التكرار، لأن إجراء الدراسة يكون في بيئة طبيعية فريدة عادة، ولا يمكن إعادة إنتاج بعض الحالات المدروسة نظرا لدينامية (حركية) السلوك البشري. ولذلك، تكون الدراسات الكيفية، على العموم، معرضة بشكل خاص لمشكلة صعوبة، أو استحالة، تكرارها و في هذا الصدد يؤكد كل من غوتز و لوكونت عدم إمكانية إعادة بناء ظروف الدراسة .

في البحوث الاثنوغرافية بالدقة الكافية، بل يدعيان أيضا أن تكرار استعمال المناهج بدقة قد لا يُسفر هو الآخر عن نتائج مماثلة (درجة الثبات).

ومع ذلك، يمكن للباحث، وعلى طول عملية بحثه الإجرائية أن يلتزم بتطبيق التوجيهات الرئيسية الآتية في مجال البحوث الكيفية الاثنوغرافية التي اقترحها هذان المؤلفان، بغية زيادة صدق النتائج::

-موضوعية الدور الذي يؤديه الباحث في مجال البحث

تحديد هادف لمجالات جمع البيانات ووصف وضعية أعوان البحث: وحدهم أو مع الباحث وفي سياق المجموعة.

-عرض شفاف للتعريفات أو لوحدات التحليل المكونة لطبيعة البحث. مع ضرورة أن تكون هذه التعريفات واضحة و"متعدية ذاتيا (Intersubjective) "بما فيه الكفاية لكي تكون مفهومة لدى الغير.

فتحقيق الصدق والثبات ليس حصرا على البحوث الكمية، فالبحوث النوعية تهتم بتحقيق الموثوقية المصادقية بطرق متعددة، وتحت مسميات مختلفة، تهدف في مجملها إلى تحقيق جوانب القوة في البحث؛ كمسمى المصادقية، وقابلية النقل، والاعتمادية، والتأكيدية، والتي سيشار إليها بالتفصيل -كمعايير عامة تشترك فيها البحوث النوعية -قبل تناول المؤشرات والإستراتيجيات التي تحقق تلك المعايير.

إن البحوث النوعية تتبنى مجموعة من الطرق المختلفة؛ للتأكد من الصدق والثبات الذي يعبر عنه بمصطلح الموثوقية Trustworthiness ، فيورد سارانتيكوس (2012) , Sarantakos بأن الصدق التراكمي validity Cumulative هو الذي تقارن فيه نتائج الدراسة الحالة بنتائج الدراسات السابقة؛ ليصدر الباحث حكمه على مدى اتساق أو اختالف نتائج الدراسة مع نتائج الدراسات السابقة. وبهذا فإن الدراسات السابقة ذات الصلة المباشرة بالدراسة الحالية تعد مقاييس معيارية لنتائج الدراسة الحالية. كما يضيف نوع آخر من سبل قياس الصدق في البحث النوعي، تحت مسمى الصدق التواصل validity

Communicative حيث يعد التواصل مع المشاركين، وتتبع البيانات، والتأكد من صحتها، وتقييم عمليات الحصول عليها، وتحول الأهداف كما في نموذج دلفاي واستخدام المصادر المتعددة Triangulation ، وتقارير الخبراء والمتخصصين حول البحث؛ يسهم في تحقيق المصدقية والواقعية لنتائج الدراسات النوعية

وكذلك يقترح نوعا ثالثا لتحقيق الجودة في البحث النوعي؛ وهو مناقشة الصدق

Argumentative validity حيث يعد شكلا من أشكال تحقيق الصدق، يتمثل في تقديم نتائج الدراسة بطريقة نهائية قابلة للاختبار، والتتبع، ويختم بالشكل الرابع من أشكال الصدق في البحوث النوعية الذي يدعى بالصدق البيئي

(الصدق السياقي Validity Ecological) الذي تكون فيه نتائج الدراسة النوعية صالحة حينما تجرى في البيئة الطبيعية للظاهرة، والسياق الطبيعي للمشاركين، وباستخدام الطرق المناسبة لجمع البيانات، وأخذ ظروف وحياة المبحوثين بعين الاعتبار، حتى تحقق الدراسة أهدافها، وتثمر في الاستفادة من بيانات المشاركين قدر المستطاع

العلاقة بين الصدق و الثبات :

من المنطقي أن نتوقع وجود علاقة بين الصدق و الثبات باعتبارهما مجتمعين من خصائص الإختبار الجيد ،ويكشفان عن صلاحية الإختبار في أن يقيس ماوضع لقياسه و في اعطائه نتائج متماثلة ،إذ يفترض في الإختبار أن يكون صادقا وثابتا ولذلك يفترض أن تكون العلاقة بين كل منهما علاقة ارتباطية عالية (سعد عبد الرحمان ، 2003).

إن الصدق يوجب الثبات ،،بمعنى أن الإختبار الصادق يكون ثابتا بالضرورة ،و العكس غير صحيح ،بمعنى أن الإختبار الثابت قد يكون صادقا و قد لا يكون صادقا .و ترجع هذه العلاقة إلى مسلمة نظرية تقوم عليها العلوم الطبيعية و الإنسانية جميعا ،و هي التسليم بثبات الكون أو ثبات العالم ،فبدون التسليم بثبات الكون ابتداءا بصحة هذه المسلمة لا يصبح هناك مجال للعلم أيا كان نوعه .

حيث لايمكن أن تتصف درجات الإختبار بالصدق في استخدامات معينة دون ان تكون درجاته متسقة ، كما ان مؤشر الثبات يعد الحد الأعلى لقيمة معامل الصدق .وتؤثر قيم معائل الثبات في درجة العلاقة بين الإختبار و المحك الخارجي (التنبؤي) ،كما تسهم قيمة معامل استقرار درجات الإختبار في زيادة الثقة بصدق التكوين الفرضي للإختبار لأنه يهتم بالثبات النسبي للسمات.

كما يقدم معامل الإتساق الداخلي دليل آخر على صدق التكوين الفرضي ،لذلك يرى كرونباخ و جلاسر أن الصدق و الثبات مفهومان مترابطان ويمكن أن يندرجا تحت اسم "مقاييس إمكانية التعميم" فالفرق الأساسي بينهما يكمن في الأبعاد التي نريد التعميم عليها .لذلك تعد نظرية امكانية التعميم التي قام بصياغتها كرونباخ و زملائه من التطورات المعاصرة المهمة التي ساهمت في إبراز التكامل بين مفهومي الصدق و الثبات.

أنواع معاملات الثبات: هناك نوعين من الثبات في مجال أداة القياس هما :

الثبات الداخلي: هو مدى اتصاف عبارات المقياس بالتناسق الداخلي ، وتوجد عدة مقاييس لإختبار الثبات الداخلي للأداة منها معامل كرونباخ، اختبار كيودر_ريتشاردسون ، التجزئة النصفية .

الثبات الخارجي : هذا النوع يتعلق بدرجة ثبات أداة القياس مع مرور الوقت ،ويمكن قياس الثبات الخارجي من تطبيق أداة القياس مرتين و على فترتين متقاربتين ،ثم قياس معامل الارتباط بين نتائج المرة الأولى ونتائج المرة الثانية و من هذه المقاييس طريقة التطبيق و عادة التطبيق و طريقة الصور المتكافئة(محفوظ جودة،2008) .

طرق التأكد من ثبات المقياس أو الإختبار.

1ثبات الإستقرار عبر الزمن(الثبات بإعادة التطبيق)

من اهم طرق حساب الثبات ، و يعتمد على معامل الارتباط بين مرات التطبيق عبر فترات زمنية معينة , و يحدد مضمون الإختبار و خصائص العينات المستهدفة في الدراسة مدة الفترة الزمنية . فمن حيث المضمون هناك اختبارات تتأثر كثيرا بالتذكر ، فمن الأفضل ان تطول الفترة الزمنية ، أما من حيث خصائص العينات ، فقد تكون مرحلة عمرية سريعة التأثير بمتغيرات النمو او التعلم مما يؤثر في ثبات الإختبار إذا أعيد بعد فترة زمنية طويلة . في هذا النوع من طرق تقدير الثبات ، نستطيع ان نحدد مدى وثوقنا من امكانية التعميم من الدرجة التي يحصل عليها الفرد في مرة معينة على الدرجة التي سوف يحصل عليها إذا ما اجري عليه نفس الإختبار في وقت لاحق .

كثيرا ما يكون تقدير الثبات صعب ، لأن العديد من الإختبارات النفسية تفاعلية ، بمعنى أن أثر القياس نفسه يدفع الشخص على تغيير اجابته على المتغير المراد قياسه . حيث يحتمل ان تكون تأثيرات الإستجابة في المرة الأولى مختلفة لدى المستجيبين ، مما يؤدي إلى تقليل تقدير الثبات .

إن هذه الطريقة لحساب الثبات تصلح للإختبارات ذات الزمن المحدد و التي تعتمد إلى حد كبير على السرعة وتصلح أيضا للإختبارات التي لاتخضع إلى الزمن ، وتعتمد في جوهرها على قوة استجابة الأفراد أكثر مما تعتمد على قياس سرعة تلك الإستجابات . كما انها لاتصلح لحساب ثبات الإختبارات التي تهدف إلى قياس التذكر لأنها تتأثر بالزمن .

عند تطبيق هذه الطريقة في حساب الثبات يجب الأخذ بعين الإعتبار الإجراءات التالية

_ التأكد من ملائمة طريقة تقدير الثبات لطبيعة الإختبار .

_ الحرص على توحيد ظروف تطبيق الإختبار .

_ تفادي أن تكون المدة الزمنية الفاصلة بين التطبيقين طويلة حتى لاتتأثر بعامل النضج .

ملاحظة : جدير بالذكر أنه من النادر ماتستخدم هذه الطريقة من قبل المعلمين ، أي في حالة الإختبارات التحصيلية ، في حين يشيع استخدامها في حالة الإختبارات النفسية و العقلية و خاصة عند تقييم الإتجاهات و الميول و القيم و حتى القدرات و الإستعدادات . في هذا النوع من الثبات يمكن أن نحدد مدى وثوقنا من إمكانية التعميم من الدرجة التي يحصل عليها الفرد في مرة معينة على الدرجة التي سوف يحصل عليها في مرات لاحقة .

2_ ثبات التكافؤ (يعتمد على الارتباط بين الإختبار و الصورة المكافئة)

يتم في هذه الحالة الحصول على معامل الثبات بطريقة التكافؤ من خلال إعطاء شكلين متكافئين لإختبار معين لنفس المجموعة و في نفس الوقت .

و الصورة المكافئة هي عبارة عن اختبار مكافئ للإختبار الذي نريد تعيين معامل ثباته و يعني التكافؤ :

_ الإختباران يقيسان نفس السمة أو الخاصية .

_ تساوي عدد البنود في الصورتين .

_ تساوي معامل الصعوبة و السهولة .

_ تساوي معاملات الثبات بين البنود .

_ تساوي المتوسط الحسابي و الإنحراف المعياري .

_ يحملان نفس التعليمة الخاصة بالإختبار .

_ تشابه تام في امثلة التطبيق .

_ نفس الوقت المخصص للإجابة .

يعني أن تكون الصيغتان متكافئتان من الناحية الإحصائية ،بمتوسطات و تباينات و ارتباطات متساوية بين الفقرات ،إضافة إلى ضرورة تكافؤ محتوى الإختبارين ، و في حالات تسمى الإختبارات المتكافئة بالإختبارات المتوازية .

وبالتالي فإن الحصول على صيغتين متكافئتين يتطلب أن يكون محتوَاهما متماثلا من حيث عدد البنود و طريقة صياغتها و مستوى صعوبتها ولها نفس نطاق المحتوى (اسماعيل، 2004) .

يتم تطبيق هذه العملية كما يلي : يتم تطبيق الإختبار المراد تعيين معامل ثباته و تطبيق الصورة المكافئة له في نفس الوقت تقريبا (بدون فاصل زمني) و بحساب معامل الارتباط بين درجات الأفراد في التطبيقين او في الصورتين يعطينا معامل التكافؤ.

تجدر الإشارة إلى ان استخدام طريقة التكافؤ في تقدير الثبات يساعد في تحديد مدى الثقة في تعميم درجة فرد ما عندما يحصل عليها إذا خضع لإختبار مشابه و لكن بمحتوى مختلف (صياغة مختلفة من الأسئلة) . تبقى المشكلة القائمة في هذا النوع من الثبات هو في الحصول على الصور المتكافئة و تسمى أيضا بالإختبارات المتوازية ، لأن الأمر يتطلب ضرورة جعل المقاييس متكافئين من الناحية الإحصائية ، يمتوسطات و تباينات و ارتباطات متساوية بين الفقرات ، إضافة إلى ضرورة تكافؤ محتوى الإختبارين .وبالتالي إذا تم مراعاة التكافؤ في النواحي التي تمت الإشارة إليها فإن معامل الثبات يكون عاليا جدا و العكس صحيح .

ثبات الإستقرار و التكافؤ: الفرق الأساسي بين ثبات التكافؤ و ثبات الإستقرار و التكافؤ هو أنه في طريقة التكافؤ و الإستقرار تكون هناك فترة زمنية بين تطبيق الإختبار المراد تعيين معامل ثباته و الصورة المكافئة ، طبعاً مع التأكيد على ضرورة توحيد ظروف و إجراءات التطبيق في المرتين .ومن ثم يتم حساب معامل الارتباط بين درجات الإختبار و درجات الصورة المكافئة و كلما كان هناك تطابق بين الدرجتين كلما كان معامل الثبات مرتفع .

ملاحظة هامة: إن قيمة معامل الثبات المتحصل عليه بطريقة الإستقرار و التكافؤ تكون أقل بمقارنتها مع طريقة الإستقرار وطريقة التكافؤ ، وهذا راجع إلى اختلاف في مصادر الخطأ في كل طريقة . فطريقة الإستقرار و التكافؤ حقيقة تمون أقل قيمة لكنها الدق لأنها تعالج مصدرين من مصادر الخطأ و هما المدة الزمنية و الصورة المكافئة . هذه الملاحظة تشير ضمناً إلى ضرورة النظر إلى طرق تقدير الثبات التي تستخدم عند بناء أو تقنين المقاييس ، وليس الإكتفاء فقط بالقيمة المتحصل عليها لمعامل الثبات .

ثبات الإتساق الداخلي :

يمكننا الحصول على ثبات الإتساق الداخلي من خلال تطبيق الإختبار مرة واحدة على مجموعة واحدة من الأفراد ، وهي تعبر فعلاً معاملات تجانس لفقرات الإختبار ، أو انها تعكس مدى ترابط الإستجابات على الفقرة الواحدة مع درجة الإختبار الكلية . إذن فإن هذه الطريقة تعتمد على مدى اتساق وحدات مع بعضها البعض أو اتساق كل فقرة وارتباطها مع الإختبار ككل بحيث تكون جميع فقرات الإختبار تدور حول السمة المراد قياسها و التي صمم من أجلها المقياس .

و بالتالي فإن تقديرات الإتساق الداخلي للإختبار هي فعلاً معاملات تجانس فقرات الإختبار أو انها تعكس مدى ترابط الإستجابات على الفقرة الواحدة مع درجة الإختبار الكلية .

مقارنة بين طرق تقدير الثبات :

تشير البيانات الواردة في الجدول التالي إلى اختصار لطرق تقدير الثبات و مصادر الخطأ المحتملة في كل منها

جدول رقم(3).يبيّن مصادر الخطأ في طرق تقدير الثبات

ثبات التصحيح	الاتساق الداخلي	الإستقرار و التكافؤ	معامل التكافؤ	معامل الإستقرار	مصادر الخطأ
		X		X	عدم استقرار السمة
	X	X	X		خطا المعاينة
		X	X	X	إدارة الإختبار
X	X	X	X	X	خطا عشوائي
X					خطا التصحيح

المصدر: موسى النبهان،(2013: 307)

المحاضرة العاشرة .: ثبات التجزئة النصفية

الأهداف :

- ان يعرف الطالب مختلف طرق التجزئة النصفية.
- أن يميز بين معادلات التجزئة النصفية
- أن يطبق معادلات سبيرمان براون، معادلة رولون ، معادلة جوتمان.

الثبات بالتجزئة النصفية :

تعتبر طريقة التجزئة النصفية مؤشرا لقياس الإتساق الداخلي ، لأن الشكليين المتكافئين يؤلفان اختبارا واحدا ، بمعنى انه بدلا عن إجراء شكل بديل للاختبار نقوم بالتعامل مع اختبار واحد فقط .

وعند تقدير الثبات بطريقة التجزئة النصفية ، يتم الحصول على درجة فرعية لكل من النصفين ، ثم يتم حساب معامل الارتباط بين النصفين وهو يعبر عن ثبات اختبار طوله نصف طول الإختبار الاصلي و معامل ثبات الإختبار الكلي يتم الحصول عليه من خلال معادلة سبيرمان براون .

و احد اهم المشاكل التي يمكن ان تواجهنا في هذه الطريقة هي كيفية قسمة الإختبار إلى نصفين.وتجدر الإشارة إلى أنه لدينا شكليين للتقسيم النصفى هما

أ _ النصف الأول مقابل النصف الثاني :

حيث يتم تطبيق الإختبار ، ثم يقسم إلى جزأين متكافئين ، ثم يحسب معامل الارتباط بين درجات الأفراد في النصف الأول و درجاتهم في النصف الثاني ، و القيمة المتحصل عليها تعبر عن معامل ثبات نصف الاختبار ، و لتقدير معامل ثبات كل الإختبار نستعمل معادلة تصحيح الطول بمعادلة سبيرمان براون و التي تعطى بالصيغة التالية:

معادلة سبيرمان براون:

$$r = \frac{2r}{1 + r}$$

و كنتيجة فإن قيمة معامل ثبات كل الإختبار اكبر من قيمة معامل ثبات نصف الإختبار ، مما يعني انه كلما زاد عدد البنود ارتفع معمل الثبات .

ب_ البنود الفردية مقابل البنود الزوجية :

حيث يتم تطبيق الإختبار ثم يقسم إلى جزأين متكافئين ، بنود فردية و بنود زوجية ، ثم يحسب معامل الارتباط بين درجات الأفراد في البنود الفردية و درجاتهم في البنود الزوجية ، ثم يتم استخدام معادلة سبيرمان براون لإيجاد معامل الثبات لكل الإختبار.

و في غالب الأحيان لا يكون تباين نصفي الإختبار متساويين ، الأمر الذي أدى إلى استبدال معادلة سبيرمان براون بمعادلة جوتمان و التي تعطى بالصيغة التالية

معادلة جوتمان:

$$r_{xx'} = 2 \left[1 - \left(\frac{\sigma^2_x + \sigma^2_{x'}}{\sigma^2_T} \right) \right]$$

تباين درجات الاختبار الفردي $\sigma^2 \mathbf{x}$:

: تباين درجات الاختبار الزوجي $\sigma^2 \mathbf{x}'$

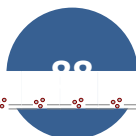
: تباين درجات الاختبار الكلي $\sigma^2 \mathbf{T}$

تجدر الإشارة إلى ان معادلتى سبيرمان براون و جوتمان تعطينا نفس قيمة معامل ثبات الاختبار عندما يكون تباين النصفين متساو . و من ناحية أخرى فإن مقدار معامل ثبات الاختبار يزداد كلما ازداد معامل الارتباط بين نصفيه .

تطبيقات لمعادلات التجزئة النصفية

معادلة سبيرمان و براون :قامت على أساس أنه يمكن التنبؤ بمعامل ثبات أي اختبار إذا عرف معامل ثبات نصفه ،بشرط ان تكون الأجزاء متكافئة .وتعتمد فكرة التكافؤ بين الأجزاء على تساوي القيم العددية لمقاييسهما الإحصائية ،المتوسط و الانحراف المعياري .

هذه الطريقة لاتصلح لحساب ثبات الاختبار عندما تترك فيه عدد كبير من الأسئلة دون إجابات (فقد في البيانات) .لأن كثرة الأسئلة المتروكة دون إجابة تؤثر على قيمة معامل الارتباط بين النصفين و بالتالي على قيمة معامل الثبات الكلي .



للتعرف على فكرة عمل هذه الطريقة ،إليك المثل التالي :

جدول يوضح نتائج خمسة طلبة في مقياس الدافعية، مكون من ستة بنود

البنود	1	2	3	4	5	6
أ	11	12	12	13	12	13
ب	15	14	15	15	16	15
ج	13	13	13	14	13	14
د	10	09	09	10	10	09
هـ	08	10	09	09	10	10

ثم نقوم بتقسيم الجدول أعلاه إلى نصفين عن طريق تجميع درجات الأفراد على البنود التي تحمل الرقم الزوجي ،و تجميع درجات الأفراد على البنود التي تحمل الرقم الفردي .ويصبح الجدول على الشكل التالي:

جدول يوضح تجزئة درجات مقياس الدافعية ،إلى بنود زوجية و بنود فردية .

الأفراد	درجات البنود الفردية	درجات البنود الزوجية
	5+3+1	6+4+2
1	3	4
2	5	6
3	9	7
4	8	4
5	2	3

و يحسب معامل الارتباط بين درجات الأفراد في النصفين و هو معامل ثبات نصف الاختبار، وبتطبيق معادلة سبيرمان براون نحصل على معامل ثبات كل الاختبار .

نتحصل على النتائج التالية:

معامل الارتباط بين نصفي الاختبار = 0,96

معامل الثبات الكلي باستخدام معادلة التصحيح = 0,98

معادلة رولون : تهدف هذه الطريقة إلى تبسيط معادلة سبيرمان براون ، وذلك بحساب تباين فروق درجات النصفين و حساب تباين درجات الاختبار ككل.

مثال تطبيقي .

نود حساب ثبات اختبار مكون من خمسة أسئلة على مجموعة مكونة من خمسة أفراد . من خلال البيانات التالية و الموضحة في الجدول التالي :

طريقة حساب معامل الثبات بطريقة التجزئة النصفية باستخدام معادلة رولون .

الأفراد	1	2	3	4	5	6
1	1	2	1	1	1	1
2	2	2	2	2	1	2
3	3	2	3	2	3	3
4	2	2	4	1	2	1
5	1	1	0	1	1	1

يتم تجميع درجات الأسئلة الفردية على حدى و كذلك بالنسبة لدرجات الأسئلة الزوجية كما

هو موضح في الجدول التالي :

درجات الأسئلة الفردية و الزوجية و و الفروق بينهما كذلك الدرجة الكلية.

الدرجات الكلية	فرق الدرجات	درجات الزوجية	الأسئلة	درجات الفردية	الأسئلة	الأفراد
7	1	4		3		1
11	1	6		5		2
16	2	7		9		3
12	4	4		8		4
5	1	3		2		5

قيمة معامل الثبات باستخدام معادلة رولون كانت 0,72

معادلة جيتمان : توصل جيتمان إلى معادلة عامة تصلح لحساب الثبات بطريقة التجزئة النصفية في حالة عدم تجانس النصفان. ويعتمد في حسابه على تباين درجات الأفراد على البنود الزوجية و تباين درجاتهم على البنود الفردية و كذا التباين الكلي _ تباين الدرجات الكلية _ .

تستعمل هذه المعادلة في حالة التجزئة النصفية عندما يكون تباين النصفان غير متساو. بمعنى النصفان غير متجانسان. تستعمل هذه المعادلة خاصة إذا كانت جميع بنود الاختبار على درجة

واحدة من الصعوبة أو إذا كانت البنود واردة من دزن ترتيب معين أو في حالة اختبارات السرعة .

أما بالنسبة للصيغة الثانية لهذه المعادلة – الصيغة 21- فهي تصلح للمقاييس الثنائية الإجابة و تستعمل في حالة توفر متوسطات البنود (عدم توفر إجابات صحيحة و أخرى خاطئة) .

تطبيق لتوضيح الفرق في تقديرات الثبات باختلاف معادلات التجزئة النصفية .

يبين الجدول التالي توزيعا لدرجات 10 طلاب على اختبار يتكون من 40 فقرة .و المطلوب حساب معامل الثبات باستخدام كل من معادلة بيرمان براون و معادلة جوتمان و معادلة رولون و من ثم المقارنة بينها .

التطبيق الأول :

الفرق بين الدرجتين	الدرجة الكلية	درجات النصف الثاني	درجات النصف الأول	الأفراد
1-	39	20	19	1
2	32	15	17	2
2-	34	13	11	3
9	9	0	9	4

5	8	7	15	1
6	19	17	36	2
7	13	11	24	2
8	9	9	18	0
9	15	18	33	3-
10	10	11	21	1-
التباين	18	35	102,3	11,21

النتائج

1- قيمة معامل الثبات باستخدام معادلة سبيرمان براون كانت 0,94

2- قيمة معامل الثبات باستخدام معادلة جوتمان كانت 0,96

3- قيمة معامل الثبات باستخدام معادلة رولون كانت 0,89

التفسير : يعزى الفرق بين قيم معاملات الثبات المتحصل عليها بعدد تطبيق مختلف معادلات التجزئة النصفية إلى اختلاف تباين نصفي الإختبار .الذي كان على التوالي : (18 و 35) .

التطبيق الثاني :

إليك البيانات التالية و الخاصة بدرجات مجموعة مكونة من 10 أفراد بعد تجزئة الإختبار

إلى نصفين:

الافراد	درجات النصف الاول	درجات النصف الثاني	الدرجة الكلية	الفرق
1	10	10	20	0
2	8	9	17	1-
3	5	4	9	1
4	4	5	9	1-
5	4	7	11	3-
6	9	8	17	1
7	7	4	11	3
8	4	4	8	0
9	8	8	16	0
10	3	7	10	4-

2.44	22.3	6.2	6.2	التباين
------	------	-----	-----	---------

بتطبيق معادلة سبيرمان براون و معادلة جوتمان و كذا معادلة رولون في حساب الثبات ، نجد أن : قيمة الثبات باستخدام كل المعادلات اعلاه كانت نفسها 0.88 . و تفسير ذلك ان تباين النصفان كان متساو .

المحاضرة الحادية عشرة : تابع لطرق حساب الثبات (ثبات كيودر ريتشاردسون, معامل

ألفا كرونباخ

تقدير ثبات كيودر _ ريتشاردسون :

إذا تم تصحيح الفقرات بشكل ثنائي (صح او خطأ) ، فإن إحدى طرق تجنب مشكلة كيفية تقسيم الاختبار هي استخدام معادلات كيودر _ ريتشاردسون ، و يمكن اعتبار هذه المعادلات ممثلاً لمعدل معامل الارتباط الحاصل من جميع التقديرات الممكنة للثبات المقسمة إلى نصفين . وهي بالتالي بديل لطرق التجزئة النصفية التي عجزت عن إعطاء نتيجة واحدة لإختبار معين .

هناك صيغتين من معادلات كيودر _ ريتشاردسون الصيغة رقم 20 و الصيغة رقم 21 . الفرق بينهما هو ان هذه الأخيرة تفترض أن جميع الفقرات ذات مستوى صعوبة واحد . و إن لم يتحقق هذا الافتراض فإن الصيغة رقم 21 ستعطي تقديراً للثبات أقل بقليل .

معادلة كيودر ريتشاردسون الصيغة 20

$$KR20 = \frac{k}{k-1} \left(1 - \frac{\sum p_i q_i}{\sigma^2_x} \right)$$

الصيغة رقم 21 لكيودر_ريتشاردسون تم اعتمادها كطريقة سهلة ولانحتاج فيها لحساب معاملات اسهولة و الصعوبة بكل بند ، وإنما تعتمد على حساب مايلي :

: عدد بنود الاختبار k

: مجموع نسب ضرب معامل الصعوبة في معامل السهولة $\Sigma p_i q_i$

: تباين الاختبار الكلي $\sigma^2 x$

معادلة كيودر ريتشاردسون الصيغة 21

$$KR21 = \frac{k}{k-1} \left(1 - \frac{\mu(k-\mu)}{k \sigma^2_x} \right)$$

: متوسط الدرجة الكلية μ

: عدد بنود الاختبار k

: تباين الدرجة الكلية $\sigma^2 x$

تجدر الإشارة إلى أنه في حالة تساوي معاملات صعوبة بنود الاختبار فإن قيم الثبات في كلتا الصيغتين (20 و 21) تكون متساوية، في حين أنه عندما تختلف معاملات صعوبة البنود فإن قيم معامل الثبات عند استخدام الصيغة 21 تكون أقل من القيمة المحسوبة بالصيغة 20.

و عند استخدام معادلة كيودر_ريتشاردسون يجب مراعاة الجوانب التالية :

- أن يكون جميع المفحوصون قد أجابو على كل أسئلة الاختبار.

- أن تتقارب الأسئلة في درجة صعوبتها و سهولتها.

- أن يكون الاختبار أحادي السمة.

- لا يمكن استخدام هذه المعادلة إلا في المقياس الثنائي.

ملاحظة: إن الفرق بين صيغتي معادلة كيودر ريتشاردسون (20_21) يكمن في أن الصيغة 21 تفترض أن جميع الفقرات ذات مستوى صعوبة واحد، وإلا فسوف تكون قيمة معامل الثبات المتحصل عليها منخفضة

تطبيق لطريقة استخدام معادلة كيودر ريتشاردسون

طبق اختبار مكون من 50 بنداً ، على 20 فرداً ، فكان عدد الذين اجابو عن كل سؤال كما هو موضح في الجدول التالي . علما بان الإجابة على البنود تكون ب صح (1) خطأ(0).
إذا علمت بان الإنحراف المعياري للدرجات = 8,5 . المطلوب تقدير معامل الثبات باستخدام معادلة كيودر ريتشاردسون .

نسبة ص.نسبة خ	نسبة الإجابات الخاطئة	نسبة الإجابات الصحيحة	الإجابات الخاطئة	الإجابات الصحيحة	البنود
0	0	1	0	50	1
0,03	0,04	0,96	2	48	2
0,12	0,14	0,86	7	43	3
0,09	0,1	0,9	5	45	4

5	35	15	0,7	0,3	0,21
6	37	13	0,74	0,26	0,19
7	30	20	0,6	0,4	0,24
8	32	18	0,64	0,36	0,23
9	28	22	0,56	0,44	0,23
10	27	23	0,54	0,46	0,24
11	28	22	0,56	0,44	0,24
12	23	27	0,46	0,54	0,24
13	20	30	0,4	0,6	0,24
14	22	28	0,44	0,56	0,24
15	34	16	0,68	0,32	0,28
16	15	35	0,3	0,7	0,21
17	18	32	0,36	0,64	0,23
18	08	42	0,16	0,84	0,13

19	05	45	0,1	0,9	0,09
20	03	47	0,06	0,94	0,05

نسبة الإجابات الصحيحة : عدد الإجابات الصحيحة على 50 (عدد الأفراد) يسمى أيضا معامل السهولة

نسبة الإجابات الخاطئة = عدد الإجابات الخاطئة على 50 (عدد الأفراد) يسمى أيضا معامل الصعوبة

و بالتعويض في المعادلة النتائج التالية : نسبة ص. نسبة خ = 3,19.

قيمة التباين = 72,25

نجد قيمة معامل الثبات بطريقة كيودر ريتشاردسون = 0,99

و باستخدام معادلة كيودر -ريتشاردسون (21) نجد قيمة معامل الثبات = 0,53.

طريقة معامل ألفا

وهي الطريقة التي اقترحها و طورها كرونباخ عام 1951 لتقدير ثبات الاتساق الداخلي للاختبار، و هي تعميم لمعادلة كرونباخ عندما لا يتم تصحيح الفقرات بالشكل الثنائي ، و يشيع استخدام هذه الطريقة في تقدير ثبات مقاييس اتجاهات و استطلاع الرأي و في مقاييس

الشخصية ، و في حالة اختبارات التحصيلية الصياغية (الإنشائية) . وجدير بالذكر أن طريقة ألفا تعطي الحد الأدنى للقيمة التقديرية لمعامل ثبات درجات الاختبارات .

إذن معادلة ألفا كرونباخ تصلح مع المقاييس ذات الميزان المترج أو استبيانات قياس الإتجاهات أو استبيانات الشخصية التي يستجيب فيها الفرد على ميزان ثلاثي أو خماسي التدرج مثل : موافق بشدة، موافق، محايد، غير موافق، غير موافق بشدة.

تعطى صيغة معامل ألفا كرونباخ بالصيغة التالية

$$\alpha = \frac{k}{k-1} \left(1 - \frac{\sum \sigma^2_i}{\sigma^2_x} \right)$$

عدد بنود الاختبار. k

مجموع تباين بنود الاختبار. $\sum \sigma^2_i$

تباين الاختبار الكلي. σ^2_x

و من مزايا هذه الطريقة مايلي:

- تعتمد على الإتساق الداخلي في أداء الفرد من فقرة إلى فقرة.

- تعتمد على الارتباطات الداخلية بين درجات الأفراد لكل فقرة و بين درجاتهم على الاختبار ككل .

أما عن القيم المقبولة لمعامل ثبات ألفا كرونباخ في تتراوح ما بين (0,70 و 0,80) وكلما زادت عن ذلك يكون أفضل.(محفوظ جودة،2008) .

ملاحظة : إن معامل الثبات المحسوب بطريقة ألفا كرونباخ يكون أقل مقارنة مع القيمة المتحصل عليها باستخدام باقي الطرق في حساب الثبات المشار إليها سابقا . (أبو زينة، 1992،

تطبيق لطريقة معامل ألفا كرونباخ

افترض أنه لدينا مقياسا للاتجاهات يتضمن 6 عبارات ،وتتدرج الإجابة عنها وفق مايلي (موافق=3)،(محايد=2)،(غير موافق=1)، ويتم تطبيقه على عينة مكونة من 5 أفراد. و المطلوب تقدير قيمة ثبات هذا المقياس اعتمادا على البيانات الموضحة في الجدول التالي :

الأفراد	بند1	بند2	بند3	بند4	بند5	بند6
1	2	1	3	2	3	1
2	3	2	3	3	2	3
3	2	1	2	3	1	2
4	2	1	1	1	1	1

5	3	2	3	2	3	2
التباين	3	3	3	7	1	7

بما ان المقياس متعدد البدائل فإن المعادلة المناسبة لحساب الثبات هي معادلة ألفا كرونباخ

و من اجل تقدير قيمة ألفا كرونباخ قمنا بحساب ما يلي :

$$\text{متوسط الدرجات الكلية} = 12,2$$

$$\text{تباين الدرجات الكلية} = 3,8$$

و بتطبيق المعادلة تم الحصول على ثبات ألفا كرونباخ بقيمة 0,84 .

المحاضرة الثانية عشرة: المعايير و عملية التقنين

الأهداف :

- أن يعرف الطالب خطوات تقنين الإختبار

مفهوم المعايير.

المعايير عبارة عن قيم تصف لنا أداء الأفراد على اختبار معين. نقول باختيار مجموعة من الأفراد يمثلون المجتمع الأصلي، وتسمى حين إذن عينة التقنين، وذلك لوضع معايير لهذا الإختبار التي يتم اشتقاقها من أداء أفراد عينة التقنين على الإختبار. و من ثم يتم مقارنة أداء الفرد بمن يكافئه من عينة التقنين. لذلك كان من الضروري توضيح خصائص عينة التقنين التي اعتمدنا عليها في استخراج المعايير. لأنه من الخطورة الإعتماد على معايير استخرجت من عينات لا تمثل المجتمع الذي ينتمي إليه الفرد من حيث كافة الخصائص التي تميزه.

عملية التقنين

ذكرت مجيد (2007) أنه يمكن القيام بعملية التقنين باتباع الخطوات الآتية:

الخطوة الأولى: تحديد المجتمع الذي سيقنن عليه الإختبار تحديدا إجرائيا دقيقا

هذه الخطوة هي الأساس لضمان صحة الخطوات اللاحقة في عملية التقنين، وتتضمن تحديد اهم خصائص و سمات المجتمع الديموغرافية من حيث الخصائص الجغرافية و السكانية و الإقتصادية، وتوزيع الفئات العمرية فيه، ونوعية التعليم، و التركيبة الإجتماعية، حيث يتم

من خلال هذه المعلومات تحديد العينة التي تمثل المجتمع تمثيلا جيدا ،كما أن هذه المعلومات تمثل خصائص المجتمع التي على ضوءها يتم تعميم نتائج الإختبار على المجتمعات الأخرى.

الخطوة الثانية : تحديد العينة

من حيث تمثيلها و طريقة اختيارها ،و هي تعتمد وبشكل كبير على المعلومات المتوفرة عن مجتمع الدراسة ،كما تعتمد على الإمكانات المادية و البشرية المتاحة ،وبصفة عامة كلما كان حجم العينة كبيرا كلما كان أفضل و أقرب إلى تمثيل المجتمع الأصلي .

الخطوة الثالثة:التخطيط الجيد و المسبق لتطبيق الإختبار :

وتتمثل هذه الخطوة في وضع خطة شاملة لتطبيق الإختبار تتضمن تحديد الإجراءات و الخطوات التي سوف تتبع ،وتجهيز جميع أدوات ومستلزمات الإختبار،مع وضع قوائم بأسماء الماكن التي سوف يتم تطبيق الإختبار فيها ،مع وضع برنامج زمني للتنفيذ.

الخطوة الرابعة :تطبيق الإختبار.

وتستلزم توحيد ظروف إجراء و تطبيق الإختبار لجميع أفراد العينة ،وذلك لتحقيق مبدأ تكافؤ الفرص امام الجميع لضمان ان الفروق التي رصدها تعود للفروق في اداء الأفراد فقط .

الخطوة الخامسة :تحليل فقرات الإختبار

وذلك للتعرف على مدى فعالية فقرات الإختبار،ومدى إسهامها في الحصول على خصائص سيكومترية تتفق مع خصائص الإختبار الجيد.وتشمل هذه الخطوة التعرف على القدرة

التمييزية لكل فقرة من الفقرات .ويقصد بالقدرة التمييزية للبنود حساسيتها للفروق الفردية بين الأفراد في السمة او القدرة المقيسة (إمطانيوس،2006) .

الخطوة السادسة: الخصائص السيكومترية

يشير محمود علام (2000) إلى أن أداة القياس الجيدة يجب ان تقيس ما وضعت لقياسه ،فالإستخدام الموضوعي لهذه الأدوات يفرض التأكد من خصائصها السيكومترية التي نفيدها من نتائجها في اتخاذ القرارات فيما يتعلق بالفرد أو الجماعة سواء في الإنتقاء أو التصنيف او التشخيص او العلاج ، وتشير- كما سبق الذكر- هذه الخصائص إلى مفهومي من المفاهيم الأساسية التي تتعلق بالإختبارات و المقاييس التربوية و النفسية هما مفهوم ثبات درجات الإختبار و صدق الإختبار

و إذا ما توفر هذين الشرطين في أداة القياس فغنها تسمح بتقديم "معلومات واقعية وصحيحة حول السمة او الخاصية التي توضع موضع القياس أي تكون صادقة ،وان تكون حساسة للفروق الدقيقة بين الأفراد في هذه السمة او الخاصية ،أي تكون ثابتة" (

إمطانيوس،2006:141

المحاضرة الثالثة عشرة: نماذج تفسير الدرجة على الإختبار او المقياس

الأهداف:

- ان يميز الطالب بين مختلف نماذج تفسير الدرجة
- ان يقارن الطالب بين الإتجاه الإيديوميثري و الإتجاه اتلسيكوميتري.
- ان يقارن الطالب بين الإختبار المحكي و الإختبار المعياري.
- _ أن يتعرف على نقاط قصور النظرية الكلاسيكية في القياس النفسي

مدخل.

إن الدرجة التي نتحصل عليها بعد عملية تطبيق الإختبارات و المقاييس لأمعنى لها في حد ذاتها ، إلا إذا تم تفسيرها وفق مرجعية نظرية يعتمد عليها الباحث بداية . هذه المرجعية النظرية في تفسير الدرجات يوضحها (صلاح الدين محمود علام ، 1995 الأساس الذي يبنى عليه تفسير النتائج.

أولاً: النموذج الذي يفسر أداء الفرد بالنسبة إلى الجماعة . (الإتجاه السيكوميتري) .

ويتم هذا النموذج بالإعتماد على معايير لمقارنة درجات كل فرد بدرجات المجموعة التي ينتمي إليها و تكون خصائصها معلومة . هذا النموذج يقوم على نظرية الفروق الفردية في السمات و الخصائص التي نقيسها . بمعنى تفسر النتائج بناء على متوسط الجماعة الذي بدوره يعطينا درجات **تتوزع توزيعاً طبيعياً** . الدرجات المستمدة من أدوات القياس تعتمد على خصائص الجماعة المرجعية التي نستخدم في الحصول على المعيار . و لا تكون للدرجة

تفسيرا إلى في ضوء المعيار المعتمد و المبتخر من الجماعة المرجعية . و بالإضافة إلى أنه حتي يتم تحديد موقع درجة الفرد بالنسبة للجماعة علينا تحويل الدرجة الخام إلى **الدرجة المعيارية**. و من امثلة الإختبارات مرجعية المعيار نجد الإختبارات التحصيلية المقننة التي تتمناز بتوحيد ظروف و طريقة تطبيق الإختبار و تعليمات الإجابة و طريقة التصحيح و التفسير . بحيث يصبح الموقف الإختباري موحدا قدر الإمكان لجميع الأفراد قصد تفسير الدرجة بإرجاعها إلى قدرات الفرد في حد ذاته و ليس لتأثير الظروف او المتغيرات الدخيلة . مما يسمح بإجراء المقارنو بصورة أكثر موضوعية .

ثانيا: النموذج الذي يفسر أداء الفرد بالنسبة إلى نفسه .

و يسمى بالقياس مرجعي الفرد ، بحيث يمكننا مقارنة أداء الفرد الواحد في مادتين مختلفتين . و مايميز هذا النموذج عن النموذج المعياري المرجع هو انه يساعد في مواقف التوجيه و الإرشاد النفسي و التربوي و العلاجي . و تقويم مدى فعالية برنامج علاجي معين .

النموذج الذي يفسر أداء الفرد بالنسبة إلى محك خارجي . (الإلتجاه الإيديوميتري) .

و يقوم هذا النموذج على فكرة مقارنة أداء الفرد بمحك أداء خارجي محدد و متوقع ، و يصاغ هذا الأداء في شكل كفايات أو نواتج متوقعة أو اهداف سلوكية تصف مختلف مستويات الأداء . و التركيز على ما يستطيع الفرد القيام به فعلا عند تطبيق الإختبار دونما مقارنته بأداء الآخرين .

هذه الإختبارات تقيس مدى تحقيق المتعلم لهذه الأهداف في ضوء أداء محدد مسبقا ، و بذلك يتحول الإهتمام من الموازنة بين المتعلمين على أساس نتائج مرجعية الجماعة ، إلى التحقق من

كفاءة كل منهم و اكتسابه للمعارف و المهارات و الإتجاهات الموجودة ،وتشخيص الصعوبات التي تعوقهم على تحقيق ذلك إستنادا إلى نتائج الإختبارات مرجعية المحك و من تصميم برامج علاجية مناسبة للتغلب على هذه الصعوبات .

تعريف المحك :يمكن تعريفه من خلال النقاط التالية :

1_ نطاق شامل من المعارف و المهارات المحددة تحديدا جيدا ،بحيث يمكن نتيجة لموازنة أداء الفرد في الإختبارات أن نعرف مايستطيع أن يؤديه هذا الفرد ومالايستطيع .فكل مفردة تعتبر معلومة او مهارة ينبغي ان يكتسبها الفرد المختبر لكي ينجح في أداء عمل محدد .

2_ يشير إلى مستوى أداء أو درجة قطع بحيث يجب ان يجيب الأفراد عن نسبة مئوية من المفردات يتم تحديدها مسبقا.

ويرجع اختلاف الإختبارات مرجعية المحك إلى اختلاف طبيعة النطاق السلوكي الذي تنسب إليه درجة الفرد وتفسر في إطاره .وهناك ثلاث تسميات لهذه الإختبارات اعتمادا على درجة تحديد النطاق السلوكي

- **إختبارات مرجعية الهدف :** تبنى على أساس مجموعة من الأهداف مصاغة صياغة سلوكية و تكون هناك مزوجة بين مفردات الإختبار وهذه الأهداف ،غير ان النطاق السلوكي لا يكون محددا ،وبالتالي عدد المفردات يكون قليل .لأنها لاتمثل النطاق الشامل للمفردات الممكنة التي تقيس مجموعة أهداف .وتطبق هذه الإختبارات عادة عند الإنتهاء من وحدة تعليمية ،وذلك بهدف تصنيف المختبرين إلى مجموعتين ،مجموعة حققت الأهداف و مجموعة لم تحقق الأهداف ،في ضوء نسبة مئوية محددة من المفردات أجيب عنها إجابة صحيحة

اختبارات مرجعية النطاق : تبنى بتحديد نطاق سلوكي شامل من المهارات تحديدا واضحا دقيقا ،وتنتقى المفردات من هذا النطاق انتقاء عشوائيا .

اختبارات التمكن أو الإتقان : تساعد في اتخاذ قرارات تتعلق باتقان الفرد لهدف تعليمي معين، وبذلك يمكن أن يكون مرجعي الهدف او مرجعي النطاق .وهذا يتطلب تحديد مستويات أداء في الإختبار تناظر المهام المختلفة.

ويفترض في هذا النموذج ان هناك متصل لإكتساب المعارف و المهارات ،يمثل أحد طرفيه عدم الكفاءة ،ويمثل الطرف الآخر الكفاءة التامة .و بالتالي يعطينا النموذج مرجعي المحك معلومات محددة عن درجة الكفاءة التي حققها الفرد مستقلة عن اداء الأقران .

المقارنة بين الإختبار المعياري و الإختبار المحكي

يمكن تلخيص الفرق بين الإختبارات مرجعية المعيار و الإختبارات محكية المرجع في الجدول التالي :

جدول مقارنة بين الإختبار المحكي و الإختبار المعياري

اختبار محكي	اختبار معياري	وجه المقارنة
يتم على ضوء أهداف محددة و مفصلة ومصاغة سلوكيا)	يقوم على أساس وصف عام للمهارات و الموضوعات التعليمية التي يتضمنها المحتوى	بناء الإختبار

		إجرائياً) بحيث يتم اشتقاق فقرة او أكثر لكل هدف
الهدف من الاختبار	يهدف إلى المقارنة و التمييز بين الأفراد	يهدف إلى قياس مدى تمكن الفرد بالنسبة لمستوى أداء محدد من قبل
نوع الفقرات	كل انواع الفقرات	كل انواع الفقرات
مستوى سهولة و صعوبة الفقرات	تكون متوسطة السهولة و الصعوبة ($0,30$ _ $0,70$)، بمعنى 30 إلى 70 بالمئة من الأفراد يستطيعون الإجابة عنها	تكون الفقرات على مدى واسع من الصعوبة ،أي أن معظم المتحنيين يستطيعون الإجابة عنها.
محك التمكن	لا يوجد محك تمكن	الأهداف السلوكية هي محكات الأداء المطلوب
إستخداماته	يساعد في تحديد مستوى تحصيل المتحن في مجال تعلم محدد. بعد مقارنته بمتوسط المجموعة المرجعية التي ينتمي إليها	يساعد في تشخيص الصعوبات و مواطن القوة و الضعف لدى الفرد بعد مقارنة أدائه بالهدف المحددة مسبقا و المتمثلة في

		محكات الأداء المطلوب و المتوقع تحقيقه
نوع الإختبارات المستخدمة	يمكن أن تتدرج ضمن هذا النوع من الإختبارات أصناف كثيرة من مثل الإختبارات التحصيلية و التربوية و النفسية و العقلية	هي أقل عددا ،ومعظمها تركز على قياس مهارات أساسية و تصنف إلى اختبارات مرجعية الهدف و اختبارات المجال ،واختبارات تمكن.
تفسير الدرجات	يتم تفسير الدرجة على ضوء موقعها بالنسبة لدرجات المجموعة التي ينتمي إليها الفرد ،فنقول ان " درجة الممتحن تنخفض او تفوق درجة متوسط المجموعة بمقدار محدد من الإنحرافات المعيارية	يتم تفسير الدرجة على أساس وصولها إلى مستوى تمكن معين أو على أساس مدى تحقيقها لهدف محدد.

النموذج الذي يفسر أداء الفرد تفسيراً موضوعياً

ارتبط هذا النموذج بمدخل جديد في القياس النفسي و هو مدخل السمات الكامنة. بما يشتمل من نظريات ونماذج سيكومترية مستحدثة (صلاح ادين أبو علام ، 1998) .وقد برز هذا اتجاه ردا على النقد الموجه للنموذج الذي يعتمد على معيار الجماعة في تفسير درجة الفرد

، ومقارنتها بأداء المجموعة ككل ، وبالتالي من الصعب تعميم نتائج الاختبارات أو الاستفادة العملية من نتائجها .

من حيث أن بنود الاختبار في صعوبتها و كذا ثباتها مرتبطة بتباين قدرات الأفراد الذين يختبرون هذه البنود. لهذه الأسباب ،اهتم علماء القياس النفسي و التربوي بمواجهة هذه المشكلات و التوصل بالتالي إلى نماذج جديدة تجعل القياس موضوعيا . و النقض بالموضوعية ان لا تتأثر درجة الفرد في الاختبار بعينة الجماعة التي ينتمي إليها . ولا تعتمد بالتالي على المعايير المستخرجة من هذه المجموعة – عينة التقنين _ بحكم ان النموذج المعياري الذي يعتمد على الجماعة يجعل الاختبار محكوما بالعينة . ومن جهة أخرى الموضوعية تعني أيضا ان يحصل الفرد الواحد على نفس الدرجة في اختبار الذي يقيس السمة بغض النظر عن النموذج الذي اعتمدناه في تفسير الدرجة – المعياري او المحكي _

لذلك فالنموذج الموضوعي أما تحديين هما :

- 1- إستقلالية أدوات القياس عن خصائص الأفراد الذين تطبق عليهم .
- 2- استقلالية قياس قدرات و سمات الأفراد من خصائص عينة البنود التي يشتمل عليها الاختبار.

وإن تم الفصل بين النماذج المختلفة المفسرة للدرجات على المقاييس ،إلا أن هذا لايعني أفضلية الواحد عن الآخر ،وإنما على الباحث ان يعتمد في اختياره للنموذج المناسب طبقا لطبيعة الأهداف الموضوعية و التي يسعى إلى تحقيقها من خلال تطبيق المقاييس

قصور النظرية الكلاسيكية في القياس النفسي

على الرغم من سيطرة النظرية الكلاسيكية في ميدان القياس النفسي ، واعتمادها في بناء مختلف انواع الاختبارات النفسية ، وكذلك اعتمادها في تحليل البيانات المستمدة من هذه الاختبارات . و على الرغم أيضا من سيطرة وانتشار تطبيق هذه النظرية وما ارتبط بها من مقاييس احصائية خاصة بتحليل مفردات الاختبار، إلا أنه تبين قصور هذه النظرية في مواجهة كثير من المشكلات المتعلقة بالخصائص السيكميترية منها:

_ طرق تحليل الاختبارات المبنية على الطريقة الكلاسيكية و المفاهيم السيكميترية المرتبطة بها من مثل معاملات الصعوبة و التمييز و الثبات كلها تتغير بتغير العينة التي حسبت منها هذه المؤشرات فهي مرتبطة بالعينة ارتباطا مباشرا . نأخذ كمثال معامل التمييز و هو يشير في مفهومه إلى ارتباط البند بالدرجة الكلية على الاختبار بمعنى صدق البند ، نجده يتأثر بباين درجات العينة ، فمدى تباين الدرجات يعكس مدى تباين العينة ، و منه فإن معامل التمييز يرتبط بخصائص العينة ، لذلك يصعب الفصل بين قيمة معامل التمييز عن طبيعة العينة . وبذلك من الممكن أن نحصل على معاملات تمييز و صعوبة مختلفة من عينة لأخرى تبعا لاختلاف العينة .

_ إن الأداء على الاختبار يختلف باختلاف بنود الاختبار الذي تم سحبها من مجتمع البنود الكبير .

_ تفترض النظرية الكلاسيكية تساوي تباين أخطاء القياس لجميع الأفراد الذين يطبق عليهم الاختبار . ولكن يمكن ان نجد أو أن نلاحظ في بعض الأحيان ان بعض الأفراد يكون أدائهم

في الإختبار أكثر اتساقا من غيرهم من الأفراد ، و أن هذا الإتساق يختلف باختلاف مستوى قدراتهم .فمن الطبيعي ان يزيد الخطأ في الإختبار الصعب المطبق على مجموعة من ذوي القدرات المنخفضة ، و يقل الخطأ إذا طبق الإختبار على مجموعة من ذوي القدرات المرتفعة _ هناك مواقف نصادفها في أثناء اختبار الخصائص السيكومترية من مثل الصدق و الثبات في النظرية الكلاسيكية تتطلب الحصول على صور متكافئة تماما لنفس الإختبار و هذا الأمر جد صعب التطبيق من الناحية العملية .

كل هذه الجوانب المشار إليها أنفا كانت المنطلق لمشكلات تناولها المختصون في ميدان القياس النفسي بهدف الوصول إلى تحقيق مستوى من الدقة و الموضوعية من زاوية ان تصبح نتائج القياس لاتتأثر بأداة القياس هذا من جهة و من جهة اخرى أن يصبح أداء الفرد و استجاباته على بنود الإختبار لايعتمد على عينة الأفراد . فكما هو متداول في الإختبارات النفسية في النظرية الكلاسيكية أثناء عملية التقنين ، يتم اشتقاق الدرجات الخام من عينة التقنين ثم يوازن أداء الفرد الذي يطبق عليه الإختبار فيما بعد بمعايير مستمدة من عينة التقنين ، و منه يتوقع انه إذا ما تغيرت عينة التقنين نفقد هذه المعايير دلالتها . و يبقى المشكل دائما مطروح هو ارتباط الإختبارات النفسية المبنية وفقا للنظرية الكلاسيكية بالعينة .

و القياس الموضوعي لايعتمد في نظامه المرجعي على مقارنة أداء الفرد بأداء الجماعة ، او بأداء الفرد الواحد في وقت لاحق او في اختبار آخر ، و إنما الموضوعية تكمن إذا ما توصلنا إلى الربط بين الأداء الملاحظ لدى الفرد في المقياس وبين السمة الكامنة التي تفسر هذا النوع

من الاداء . وقد أسفرت جهود المختصين في هذا المجال عن ما يسمى نظرية السمات الكامنة
او نظرية الإستجابة للمفردة و هي النظرية الحديثة في القياس النفسي و التربوي .

المحاضرة الرابعة عشر: الخصائص السيكمترية في النظرية الجديدة للقياس النفسي

والتربوي

مدخل

شهد القياس النفسي والتربوي في العقود الأخيرة تحولاً جذرياً من النموذج الكلاسيكي (CTT) إلى نماذج نظرية أكثر تطوراً، أبرزها نظرية الاستجابة للمفردة (IRT) ونظرية التعميم (GT). هذا التحول لم يأت فقط كتطوير تقني، بل كإعادة تفكير في مفهوم "القياس" ذاته في العلوم النفسية والتربوية، وما يعنيه أن نقيس "شيئاً" غير قابل للملاحظة المباشرة. تمثل النظريات الجديدة استجابة لأزمة أوسع في علم النفس، حيث أظهرت أبحاث أن العديد من المقاييس النفسية كانت تعاني من مشاكل جوهرية في الصدق والثبات وقابلية التكرار. وقد كشفت التحليلات النقدية أن الإفراط في الاعتماد على النمذجة الإحصائية جاء على حساب العمليات القياسية المؤسسة معرفياً.

أولاً: نظرية الاستجابة للمفردة (IRT) وإعادة تعريف الخصائص السيكمترية

1.1 ثبات المفردة وليس ثبات الاختبار فقط. في النظرية الكلاسيكية، كان التركيز على ثبات الدرجة الكلية للاختبار. أما في IRT، فالاهتمام ينتقل إلى مستوى المفردات نفسها:

دالة معلومات المفردة: (Item Information Function) تحل محل معامل الصعوبة والتمييز التقليديين، وتمثل مقدار الدقة التي تقدمها كل مفردة في قياس سمة معينة. كل مفردة

لها "نافذة" خاصة بها تقيس فيها بدقة - فالمفردات السهلة تقيس بدقة الأفراد ذوي القدرة المنخفضة، والمفردات الصعبة تقيس بدقة الأفراد ذوي القدرة المرتفعة.

دالة معلومات الاختبار: (Test Information Function) مجموع دوال معلومات المفردات، وهي توضح بدقة أين يكمن "قوة" الاختبار على امتداد سلم القدرة. هذا يتناقض مع النموذج الكلاسيكي حيث كان الثبات قيمة واحدة تنطبق على جميع المستويات.

1.2 معاملات المفردة المستقلة عن العينة

من الإنجازات الرئيسية لـ IRT تحقيق استقلالية معاملات المفردة عن عينة التقنين. فمعاملات الصعوبة والتمييز المحسوبة بنماذج IRT تكون (نظرياً) ثابتة بغض النظر عن خصائص العينة التي استخدمت في التقدير. هذا يسمح بمقارنة نتائج الأفراد من خلفيات مختلفة.

ثانياً: الصدق في النظرية الجديدة - ما وراء النماذج الإحصائية

2.1 الصدق كتقييم نظري وليس مجرد معامل

شهد مفهوم الصدق تطوراً عميقاً في النظريات الجديدة. فبدلاً من كونه مجرد معامل ارتباط، أصبح يُفهم الصدق على أنه "تقييم متكامل للاستدلالات النظرية" المستخلصة من الدرجات. هذا التحول جعل الصدق قضية نظرية قبل أن تكون إحصائية.

2.2 ثبات القياس عبر المجموعات (Measurement Invariance)

من الإسهامات الرئيسية للنظرية الجديدة هو التركيز على "ثبات القياس عبر المجموعات المختلفة" (ثقافياً، جندياً، عمرياً). السؤال المحوري: هل تعني الدرجة نفسها الشيء نفسه لدى الذكور والإناث؟ لدى ثقافات مختلفة؟.

2.3 نقد الصدق التقليدي: نحو منظور واقعي

برزت في النظرية الجديدة أصوات ناقدة للمفهوم التقليدي للصدق. يرى باحثون مثل Borsboom أن الصدق لا يمكن أن يكون مجرد اتفاق على تفسير الدرجات، بل يجب أن يقوم على "علاقات سببية" حقيقية بين القياس والسمة المقاسة. بمعنى آخر، الاختبار صادق إذا كانت التغيرات في السمة تسبب تغيرات في الدرجات.

الرسالة الأهم من النظريات الجديدة هي أن "القياس النفسي ليس مجرد تقنية إحصائية"، بل هو ممارسة معرفية تتطلب فهماً عميقاً للظاهرة المقاسة، وللأداة المستخدمة، وللسياق الذي يتم فيه القياس، وللهويات والثقافات التي يُمثلها المفحوصون. النظريات الجديدة لا تقدم إجابات نهائية، بل تفتح أسئلة أكثر عمقاً حول "ماذا" و"كيف" نقيس حقاً في العلوم النفسية والتربوية.

المحاضرة الخامسة عشرة: بناء الإختبارات وفقا لنظرية السمات الكامنة

الأهداف:

- أن يتعرف الطالب على نظرية السمات الكامنة
- أن يتعرف على أهم الفروقات بين النظرية الكلاسيكية و نظرية السمات الكامنة
- _ أن يعرف تطبيقات نظرية السمات الكامنة

مدخل.

تقوم الإختبارات النفسية بعامة على افتراض مفاده ان هناك سمات أو خصائص معينة يشترك فيها جميع الأفراد و لكنهم يختلفون في مقدارها و أنه يمكن الإستدلال عليها و على مقدارها من السلوك الملاحظ للفرد و المتمثل في استجاباته على مفردات الإختبار ، و من هذا جاءت التسمية بالسمات الكامنة . فالسمة التي تكمن وراء الإستجابة على مفردات اختبار لفظي ، تختلف عن السمة التي تكمن وراء استجاباته على مفردات اختبار من نوع آخر. ولكن يمكن ان تكمن سمة واحدة وراء استجاباته على مفردات اختبارين مختلفين متعلقين بنفس المحتوى لذلك فإن الهدف الأساسي من النظرية الكلاسيكية و النظرية الحديثة في القياس _ نظرية السمات الكامنة _ هو تحديد العلاقة بين استجابات الأفراد على اختبار معين و السمة الكامنة وراء هذه الإستجابات . من أجل الوصول إلى تفسير و التنبؤ بسلوك الأفراد في مواقف مماثلة ، وتتلخص افتراضات نظرية الإستجابة للمفردة في النقاط التالية :

_ إن استجابة الفرد لمفردات اختبار واحد تعبر عن سمة واحدة كامنة يمكن قياسها في بعد واحد .ويمكن الوصول إلى هذا التحديد خاصة إذا كانت بنود الإختبار متجانسة لقياس هذه القدرة او السمة الكامنة . و عليه فإن الإختبار المتعدد الأبعاد لايمكن ان يكون مناسباً لنماذج السمات الكامنة .

_ استقلالية البنود عن بعضها البعض بمعنى ان الإستجابة على بند من البنود لا تؤثر على الإستجابة على بند آخر .

_ إن العلاقة بين أداء الفرد في كل بند من بنود الإختبار وبين القدرة التي يقيسها يمكن وصفها بدالة رياضية تصاعدية ، بمعنى امكانية الإستجابة الصحيحة للبند في مستويات مختلفة من القدرة ، فاصحاب القدرات العالية يحتمل توصلهم إلى الإجابة الصحيحة على السؤال بشكل يفوق اصحاب القدرة المنخفضة و بالتالي فالنظرية الحديثة لاتهتم بالدرجة الكلية بقدر اهتمامها بالإستجابة على كل مفردة من مفردات الإختبار . و هذا الفرق جوهري يميز بين النظرية الكلاسيكية في القياس و النظرية الحديثة او نظرية السمات الكامنة .

من هذه الإفتراضات التي قامت عليها النظرية الحديثة يمكننا استنتاج أن مؤشرات البنود من مثل معاملات الصعوبة و السهولة و التمييز و كذا الصدق و الثبات لم تعد متعلقة بالعينة التي استخرجت منها هذه المؤشرات .

تقييم قدرات الأفراد لاتعتمد على الدرجة الكلية في الإختبار. فإن كانت النظرية الكلاسيكية تعتمد على الدرجة الكلية و التي ترتبط بأسئلة الإختبار ، فإن النظرية الحديثة تتحدث عن القدرة التي تكون هي الأخرى مستقلة عن أسئلة الإختبار.

يمكن تحديد مدى دقة التقديرات التي توفرها النظرية سواء فيما يخص البند (مؤشرات البند) أو فيما يخص الفرد (قدرات الفرد) .

و بالتالي الإستجابة على بنود الإختبار من قبل افراد تقع قدراتهم في مستوى معين ، يعتمد على ذلك المستوى و ليس على عدد الأفراد ، و عليه فإن مؤشرات البند (صعوبة ، سهولة، تمييز) و إن كانت في النظرية الكلاسيكية تتغير يتغير المجموعة التي حسبت منها هذه المؤشرات (عينة التقنين) ، فإنها تبقى ثابتة في النظرية الحديثة حتى لو تغيرت المجموعة أو زاد عددها .

نماذج نظرية السمات الكامنة:

نظرا لإعتماد نظرية السمات الكامنة فرضية أساسية مؤداها أن استجابة الفرد لمفردة اخیبارية تكون دالة لكل القدرة التي يفترض أن الإختبار يقيسها لدى الفرد ، وخصائص المفردة التي يحاول الإجابة عنها ، و بالتالي فالأمر يحتاج إلى معلومات من مصدرين : احدهما يتعلق بالفرد ، و الآخر يتعلق بالمفردة الإختبارية , و عادة ما نحتاج إلى قيمة عددية واحدة تتعلق بالفرد و هي بارامتر القدرة المقاسة لدى الفرد و قيمة عددية او اكثر تتعلق بالمفردة الإختبارية بارامتر او بارامترات المفردة . و في ضوء عدد بارامترات المفردات ، فإنه يندرج تحت نظرية السمات الكامنة مجموعة من النماذج السيكوميتريية من اهمها :

1_ النماذج ثلاثية البرامتر ، التي تفترض ان المفردة تختلف في صعوبتها و تمييزها ، و كذلك في بارامتر التخمين الذي يمثل احتمال توصل الأفراد ذوي القدرة المنخفضة إلى الإجابة الصحيحة عن المفردة ، و خاصة في الفقرات من نوع الإختيار من متعدد .

2_ النماذج ثنائية البرامتر ، و التي تفترض أن المفردات تختلف في صعوبتها و تمييزها بين المستويات مختلفة القدرة .

3_ النموذج احادي البرامتر ، و يطلق عليه نموذج راش ، ويفترض هذا النموذج بان مفردات الاختبار تختلف فقط في صعوبتها و تتساوى في تمييزها .

أهم الفروقات بين النظرية الكلاسيكية و نظرية السمات :

أولاً : قدمت نظرية السمات مفهوما جديدا للثبات يدعى (دالة المعلومة) التي نحصل عليها من تطبيق بند او اختبار . و هذه المعلومة تعني مقدار الثقة في اننا حصلنا على معلومات تقودنا إلى تقدير قدرة فرد معين أو مجموعة من الأفراد في مستوى واحد من القدرة . وتمثل دالة المعلومة عادة بمنحنى يبين أعلى و أقل قيمة تحصلنا عليها عند مستويات القدرة المختلفة بخلاف الثبات الذي يعطينا قيمة واحدة او مؤشر واحد فقط للاختبار بالنسبة لكل الأفراد الذين طبق عليهم الاختبار .

ثانياً: إن دالة المعلومة تعتمد فقط على مقياس القدرة و على منحنى استجابة البند ، كذلك فإن الثبات مؤشر للخطأ المعياري في المقياس وهو ثابت على جميع مستويات الدرجة الحقيقية ، بينما دالة المعلومة مؤشر للخطأ المعياري في التقدير و هذا الخطأ يحسب عند كل مستوى من مستويات قدرة الطالب وتعرف بالتالي دقة تقدير قدرة الفرد على كل مستوى من المستويات .

ثالثا : مؤشرات الإختبار في النظرية الكلاسيكية تتغير بتغير المجموعة التي طبق عليها الإختبار . اما مؤشرات الإختبار في نظرية السمات الكامنة فهي ثابتة مهما كانت المجموعة كما ان قدرة الفرد ثابتة مهما كانت الأسئلة التي طبقت عليها .

تطبيقات نظرية السمات :

إن نظرية السمات الكامنة أو نظرية الإستجابة للبند وفرت مجالا خصبا لتحسين تطبيقات القياس النفسي بالإضافة إلى تطبيقات أخرى :

_ أصبح بالإمكان مقارنة قيم القدرات الفردية عند التقنين بدلا من الإعتماد على التوزيع التكراري للعينة التي طبق الإختبار ، أي أنه أصبح من الممكن إعداد بنوك من الأسئلة التي حددت مؤشراتها و الإستفادة منها في تقنين الإختبار.

_ نظرية السمات الكامنة توفر طريقة للمقارنة بعيدا عن تلك التي تعتمد الموازنة بين الدرجات الخام فهي تعتمد و تهدف إلى معرفة موقع الدرجة التي يحصل عليها الفرد على مقياس القدرة .

_ أصبح بالإمكان تحديد نوع الإختبار الذي نريده بمعرفتنا لمستوى القدرة التي تم تحديدها مسبقا .

المحاضرة السادسة عشرة..مراحل بناء الإختبارات و المقاييس

الأهداف :

- أن يعرف الطالب مراحل بناء الإختبار النفسي

مدخل

لكل أداة خصوصية في التصميم و الأداء (التطبيق) الأمر الذي لابد قبل الشروع في بناء أي أداة قياس في الإنتباه لمجموعة من المبادئ و الشروط هي :

1_ الهدف المتوقع تحقيقه من أداة القياس:

فهناك ادوات تهتم بتقويم المهارات و اخرى بقياس الإستعدادات ، و هناك أيضا أدوات قياس التحصيل ... الخ .

2_المجال السلوكي موضوع الإهتمام :

فعلى الباحث قبل الشروع في بناء أداة القياس ان يحدد المجال السلوكي تحديدا دقيقا .و المجال السلوكي قد يكون :

_ مجال محدد و له مرجعية معروفة مثل ان يكون منهج دراسي تكون عندها أداة القياس إختبار تحصيلي.

مجال سلوكي غير محدد و ليس له مرجعية بحيث يمكن التحكم فيها تكون عندها أداة القياس إختبار

قدرة أو استعداد

_ الإختبارات النفسية المرجعية السلوكية فيها غير محددة .

3_ تحديد خصائص المستهدفين بالقياس :

بمعنى تحديد الفئة العمرية المستهدفة بالقياس إذ يترتب على ذلك :

_ تحديد الزمن الذي سوف يخصص للإختبار .

_ تحديد نوع الفقرات

_ تحديد شكل الأداء المطلوب.(غالبا ما تكون فقرات المقاييس التي تطبق على الأطفال على

شكل صور او أشكال)

4_ تحديد مدى توفر الإمكانيات المتعلقة بتطبيق و تصحيح أداة القياس .

من هنا تبرز اهمية المعلومات و المهارات اللازمة لتطوير ادوات القياس عموما و الإختبارات المتعلقة بالجوانب المعرفية و التحصيلية على وجه الخصوص .

و على الرغم من اتساع الإهتمام ببناء ادوات القياس ، فإن هناك خطوات رئيسية يمكن ان تتبع معظمها في بناء ادوات القياس النفسي و هي على النحو التالي :

1_ تحديد الهدف من عملية القياس :

تعتبر عملية تحديد أهداف أداة القياس المراد بناؤها خطوة اساسية في بنائها ، ثم ان عملية تحديد الأهداف تختلف من حيث سهولة الصياغة و درجة الوضوح حسب نوع الظاهرة أو

السمة اتي من المتوقع ان يتناولها المقياس ، من زاوية ان عملية اشتقاق الأهداف مرتبطة بالمجال السلوكي موضوع الإهتمام . فعلى سبيل المثال صياغة الأهداف بالنسبة للإختبارات التحصيلية اكثر سهولة ووضوحا من تلك المتعلقة بالمقاييس النفسية ، لن المجال السلوكي بالنسبة للإختبارات التحصيلية واضح و هو المنهاج الدراسي المحدد الذي يحتوي على محتوى مادة تعليمية لها ابعاد محددة وفقا للأهداف المتوخاة من تدريس هذه المادة التعليمية . بينما يتطلب الامر غير ذلك عند الحديث عن المقاييس النفسية ، إذ ان المجال السلوكي غير محدد ، و بالتالي فإن تحديد الأهداف يستدعي الرجوع إلى النظريات المتخصصة ذات العلاقة من اجل تعريف الظاهرة تعريفا اجرائيا كما يتطلب موضوع القياس بصورة تفصيلية قابلة للقياس و الملاحظة و ترتبط ارتباطا وثيقا بالسلوك المتوقع تحققه .

و عليه يجب ان تتم عملية بناء المقاييس في ضوء الأهداف الأساسية التي تشتق من طريقة تفسير و استخدام الدرجات على الإختبارات و المقاييس . بمعنى ان الإختبار الذي تستخدم نتائجه في عمليات التصنيف أو الإختيار يختلف حتما في محتواه و خصائصه عن ذلك الذي يستخدم في التشخيص و غيره .

2_ تحديد جوانب السلوك التي تمثل السمة :

تبرز أهمية هذه الخطوة في كونها الأساس الذي تبنى عليه الفقرات ، و المجال الذي تشتق منه ، فبعد تعريف الظاهرة (السمة) ، يتم تحديد أهداف المقياس بتوضيح و تحديد مكونات تلك السمة ، أو يلجأ إلى تقسيم الظاهرة أو السمة إلى عناصرها الأولية ، حيث يمثل كل عنصر مجالا معيناً ، او إطارا مرجعيا ، يتم اشتقاق الفقرات و صياغتها منه . و عليه فإن

مصمم الإختبار يبدأ بافتراض عدد من انواع السلوك الدالة على السمة المراد قياسها ، و التي يعتقد أنها تمثل تلك السمة ، من أجل صياغة أو بناء عدد من الفقرات لهذا الغرض . و قد يقود هذا الإجراء إلى حذف عدد من المجالات المهمة او التركيز على مجالات لا تتعلق بالسمة بدرجة كبيرة ، ذلك لأن هامشا كبيرا من ذاتية أو شخصية أو ثقافة أو خلفية الشخص القائم على بناء الإختبار يؤثر في تلك العملية . و للتحكم في العملية على أفضل وجه ، لابد لمصمم الإختبار من الإهتمام بالنقاط المجورية التالية :

أ_ تحليل الإستجابات على أسئلة مفتوحة.

فقد يبدأ مصمم الإختبار بصياغة أسئلة مفتوحة على العينة المستهدفة ، ثم يقوم بتصنيف الإستجابات و تحليلها ، و من ثم يشتق الأبعاد الرئيسية المكونة للسمة . و من امثلة المقاييس المعروفة و التي صممت بطريقة الأسئلة المفتوحة نجد مقياس مفهوم الذات الخاص بالأطفال حيث تم الإنطلاق من تحليل كتاباتهم في مواضيع التعبير للوصول إلى مقياس مفهوم الذات لديهم عن طريق اشتقاق الأبعاد التي تكونت منها الأداة .

ب_ مراجعة البحوث و الدراسات في موضوع السمة :

و هذا يساعد في تكوين فهم أوسع لمكونات السمة و علاقتها بسمات أخرى إضافة إلى المساعدة في تعريف شامل مفصل لها .

ج_ استشارة الخبراء:

إذ يمكن لمصمم الإختبار الحصول على مدخلات مهمة من الخبراء المشتغلين بالسمة ، فقد يمكن جمع البيانات حول السمة ، مفهومها و مكوناتها و غير ذلك .

د_ تحديد شكل الفقرات:

بعد أن يتم تحديد الهدف من الإختبار، يجرى العمل نحو تحديد شكل الفقرات التي سوف يتم صياغتها . هنا ينبغي استعراض عدد من البنود الأساسية التالية المتعلقة بفقرات الإختبار و إخراجها :

_ طول الإختبار او عدد فقراته .

_ مستوى الصعوبة التي يجب ان يكون عليها .

_ توقيت و عدد مرات التطبيق .

_ طبيعة المرفقات أو المعينات : (صورية ، كلامية ، أم غيرها)

_ طول الإختبار : ليس هناك معيار او معادلة تحدد طول الإختبار لكن يجب ان يكون العدد الإجمالي للفقرات كافيا لتمثيل شامل لمجالات السمة أو المحتوى و يغطي عددا معقولا من الأهداف أو المخرجات المراد تحقيقها .

قائمة المراجع

صلاح الدين أبو علام (2009) القياس و التقويم التربوي و النفسي أساسياته و تطبيقاته توجهاته المعاصرة، دار الفكر العربي ،القاهرة.

محمود علام (2013) اتقان القياس النفسي الحديث، دار الفكر موزعون و ناشرون ،عمان.

سعد عبد الرحمان (2003) القياس النفسي النظرية و التطبيق ،دار الفكر العربي،

عبد الله الصمادي، الدار ربيع (2004) القياس و التقويم النفسي و التربوي بين النظرية و التطبيق ،ط1، دار وائل عمان

محمد ربيع شحاتة (2007) قياس الشخصية ، عمان ،دار الميسرة للنشر و التوزيع

ملحم سامي محمد (2012) القياس و التقويم في التربية و علم النفس ،ط6،الأردن، دار الميسرة للنشر و التوزيع.

صفوت فرج (1980)،القياس النفسي ،مكتبة النجلو المصرية

صلاح احمد مراد (2000)، الأساليب الإحصائية في العلوم النفسية و التربوية و الإجتماعية،مكتبة النجلو المصرية .

حسن علاوي،نصر الدين رضوان (1999)،القياس في التربية الرياضية و علم النفس الرياضي ،دار الفكر العربي ،القاهرة.

أحمد تيغزة (2008) نظرية الصدق الحديثة و متضمناتها لواقع القياس ،ندوة علم النفس (الصفحات 1- 35) الرياض،قسم علم النفس،جامعة الملك سعود.

أبو علام (2006)، البهي السيد (1978) علم النفس الاحصائي و قياس العقل البشري ،القاهرة ،دار الفكر العربي

محمد عويضة كامل (1996) القدرات العقلية في علم النفس ،بيروت،لبنان،دار الكتب العلمية

السيد ،ف، (1989) علم النفس افحصائي و قياس العقل البشري،مصر،دار الفكر العربي

محفوظ جودة (2008)، التحليل الإحصائي الأساسي باستخدام برنامج spss ،دار دانة للنشر و التوزيع ،ط1.

إسماعيل بشري (2004). المرجع في القياس النفسي .ط4.مصر.مكتبة الانجلو المصرية

سعد عبد الرحمان (2003).القياس النفسي .النظرية و التطبيق .دار الفكر العربي

مجيد.دوسن شاكر (2007).الاختبارات و المقاييس النفسية و التربوية .ط1.عمان.دار ديبونو للنشر و التوزيع

إمطانيوس ميخائيل (2006).القياس النفسي .ج1.دمشق.منشورات جامعة دمشق

فريد كامل أبو زينة (1992)، أساسيات القياس و التقويم في التربية،مكتبة الفلاح للنشر.

موسى النبهان (2013) أساسيات القياس في العلوم السلوكية ،الطبعة العربية الثانية ،دار الشروق للنشر و التوزيع ،عمان الأردن

Anastasi ;A ;et uraina;s ,(1997) psychological testing ,new york

AERA ,APA,et NCIME(1999) standards for educational and psychological testing ,washington DC :Author ;

Pettersen ,n,(2002) ,Evaluation du potentiel humain dans les organisations
,canada,presses de l universite du quebec ,

Terwee, C. B., et al. (2007). Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology*, 60(1), 34-42.

National Assessment Governing Board. (n.d.). Evidence based on relations to other variables. U.S. Department of Education.

Psychological Scales. (2025). Criterion contamination: Definition & meaning.

Equal Employment Opportunity Commission. (1978). Uniform Guidelines on Employee Selection Procedures. 29 C.F.R. § 1607.5.

Terwee, C. B., et al. (2016). Quality criteria for psychometric properties. *Systematic Reviews*, 5(1).

Mislevy, J. L., & Rupp, A. A. (2012). Predictive validity. In *Encyclopedia of Research Design*. Sage Publications.

Hodge, R. (2010). *Issues in validity generalization: The criterion problem* (Master's thesis). University of Central Florida.

Cheng, W. (2019). Validation: An ongoing process of accumulating evidence about a test. *Credentialing Insights*.

Anastasi, A., & Urbina, S. (1997). *Psychological testing* (7th ed.). Upper Saddle River, NJ: Prentice Hall.

Angoff, W. H. (1971). Scales, norms, and equivalent scores. In R. L. Thorndike (Ed.), *Educational measurement* (2nd ed., pp. 508–600). Washington, DC: American Council on Education.

Bloom, B. S. (1976). *Human characteristics and school learning*. New York: McGraw-Hill.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52(4), 281–302.

Ebel, R. L. (1972). *Essentials of educational measurement*. Englewood Cliffs, NJ: Prentice-Hall.

Glaser, R. (1963). Instructional technology and the measurement of learning outcomes. *American Psychologist*, 18(10), 519–52

· Shultz, K. S., Whitney, D. J., & Zickar, M. J. (2014). *Measurement Theory in Action: Case Studies and Exercises* (2nd ed., Module 2: The Cornerstones of Measurement: Reliability, Validity, and Objectivity). Routledge.

· Geisinger, K. F. (Ed.). (2013). *APA Handbook of Testing and Assessment in Psychology* (Vol. 1: Test Theory and Testing and Assessment in Industrial and

Organizational Psychology, Chapter 4: Reliability). American Psychological Association.

Luong, A., Sprung, J., & Zickar, M. (2023). Criterion terms, concepts, and measures. In *Personnel Psychology*. Taylor & Francis.