المحاضرة 06: الخصائص السيكومترية للاختبار (صدق الاختبار)

يتم حساب الخصائص السيكومترية للاختبار في إطار عملية التقنين، حيث يبين علماء القياس معنيين للتقنين هما:

أولا: أن تكون تعليمات الاختبار وصياغة بنوده والزمن المخصص له، وشروط وإجراءات تطبيقه موحدة في كل المواقف، ومن هذه الناحية يعني التقنين التوحيد، توحيد الشروط المحيطة بالمفحوصين وضبط العوامل والمتغيرات جميعها التي يمكن أن تؤثر في أدائهم الاختباري وإبعاد أثرها إلى أقصى حد، وذلك بهدف قياس المتغير الذي صمم الاختبار لقياسه بحجمه الحقيقي لدى تلك العينة من الأفراد التي سيطبق عليها.

ثانيا: أن يخضع الاختبار للتقنين من خلال تطبيقه على عينة ممثلة للمجتمع الأصلي، بهدف استخراج معايير معينة تحدد معنى الدرجة الخام التي يحصل عليها الفرد في ضوء تمركز أو تشتت درجات أفراد عينة التقنين، والمؤكد منه أنه لا يمكن أن يكون الاختبار مقننا بالمعنى الثاني دون أن يكون مقننا بالمعنى الأول. (امطاينوس ميخائيل، 2016)

أما (صلاح الدين علام، 2000) فيرى أن التقنين يقصد به أن يكون بناء وتصحيح وتفسير نتائج الاختبار أو أداة القياس مستندا إلى قواعد محددة، يحيث تتوحد فيه وتتحدد بدقة مواد الاختبار وطريقة تطبيقه، وتعليمات إجابته وطريقة تصحيحه أو تسجيل درجاته، وبذلك يصبح الموقف الاختباري موحدا بقدر الإمكان لجميع الأفراد في مختلف الظروف.

كما يشمل التقنين حساب الخصائص السيكومترية للاختبار، وهي الخصائص الضرورية المتعلقة بالثبات والصدق، مع استخراج معايير لتفسير النتائج المتحصل عليها من تطبيق الاختبار.

يقوم مصمم الاختبار بحساب الصدق:

يعتبر الصدق من الخصائص السيكومترية للاختبارات النفسية، وهو يمثل إحدى الوسائل المهمة في الحكم على صلاحية هذه الاختبارات، وفيما يلي سوف نتطرق إلى مفهوم الصدق وأنواعه والعوامل المؤثرة فيه.

1-مفهوم الصدق:

تعرف (سوسن مجيد، 2014) الصدق أنه: قياس الاختبار فعلا أو حقيقة للسمة أو الظاهرة التي وضع لأجلها.

بمعنى أن الاختبار الصادق يقيس ما وضع لقياسه فقط فاختبار الذكاء يقيس الذكاء فقط ولا يقيس شيئا آخر مثل سمات الشخصية أو التحصيل أو جوانب انفعالية. (صلاح مراد وأمين سليمان، 2005)

فالصدق هو أن يقيس الاختبار ما وضع لقياسه، ويحقق الهدف الذي وضع من أجله.

يتميز الصدق بالخصائص التالية:

- صفة نسبية: فلا يوجد اختبار عديم الصدق أو تام الصدق، هذا إذا استعملنا الاختبار فيما وضع له، فقد يكون اختبار الذكاء الخاص بتلاميذ السنة الثانية ابتدائي نتائجه مرتفعة الصدق، إذا استعمل مع تلاميذ السنة الثانية ابتدائي، ومتناقصة الصدق إذا استعمل مع عينة أخرى، وعليه فالصدق ليس أمرا مطلقا بل يختلف من اختبار لآخر بحيث لا نستطيع أن نقول أن الاختبار صادق أو غير صادق بل نقول أنه صادق بدرجة ما.
- صفة نوعية: حيث يرتبط الصدق مباشرة بالغرض أو الاستعمال الخاص الذي صمم الاختبار من أجله، فالاختبار الذي أعد بغرض التنبؤ بأداء المفحوصين في فترة مقبلة لا يصلح لغرض التشخيص والكشف عن نقاط الضعف والقوة في أداء المفحوصين، كما يرتبط الصدق مباشرة بالفئة أو الجماعة التي صمم لها، فالاختبار الذي صمم لفئة أو فئات معينة من المعاقين لا يصلح للأسوياء.

- صفة تتعلق بنتائج الاختبار وليس بالاختبار نفسه، ولكننا نربطها بالاختبار من قبيل الاختصار أو التسهيل، والأصوب أن نتكلم عن صدق النتائج، أو بالأحرى صدق تفسيراتنا للنتائج. (امطاينوس ميخائيل، 2016)

فالصدق إذن لا يتعلق بالاختبار نفسه بقدر تعلقه بتفسير الدرجات المستخرجة منه، وهذا يعني مدى فائدة أداة القياس في اتخاذ قرارات تتعلق بغرض أو أغراض معينة، فالصدق هنا هو صدق تفسير النتائج أو تأويلها وليس صدق الاختبار ولا صدق الدرجات.

2- أنواع الصدق:

يرى (بشير معمرية، 2012) أن الجمعية الأمريكية لعلم النفس (1954) صنفت الصدق إلى ثلاث أنواع أساسية هي: صدق المحتوى، الصدق المرتبط بمحك وأخيرا صدق التكوين الفرضي.

2-1-صدق المحتوى:

ويطلق عليه البعض الصدق المنطقي أو صدق المضمون أو الصدق بحكم التعريف، ويشير صدق المحتوى حسب (صلاح مراد وأمين سليمان، 2005) إلى مدى تمثيل بنود الاختبار أو المقياس لمحتوى السمة موضع القياس، ويرتبط صدق المحتوى بمفهوم الشمول حيث يمكن اعتبار هذا النوع من الصدق دليلا على شمول الأداة ودرجة تمثيلها للمحتوى.

بمعنى أن الاختبار يكون صادقا إذا كان المحتوى الذي نريد قياسه ممثلا تمثيلا صادقا من خلال بنود الاختبار، كما أن هذه الطريقة تدخل في المراحل الأولى لتكوين أي اختبار.

ويرى (بشير معمرية، 2012) أن هذه الطريقة تصلح خاصة للاختبارات التحصيلية وبرامج التدريب على الأداء والكفاءة المهنية، كما أنه من الصعب حساب صدق المحتوى للاختبارات النفسية، كسمات الشخصية والاتجاهات والقيم وغيرها.

ففي مجال قياس الشخصية قد يكون بند يقيس الانبساطية / الانطوائية مثلا: هل تفضل الذهاب إلى حفلة مع أصدقائك أو تبقى في المنزل لقراءة كتاب؟ فإذا أجاب قراءة كتاب فهو انطوائي ولكن ربما يجيب هكذا لميله للمطالعة، فهنا يظهر أن البند يقيس الميول، فبالتالي ففي اختبارات الشخصية لا يمكن إصدار حكم على أن بندا ما يقيس الخاصية التي وضع لقياسها.

يندرج ضمن صدق المحتوى:

-**الصدق الظاهري (السطحي):** والذي يتمثل في فحص محتوى الاختبار والتأكد من جودته فيما يقيسه دون فحص تجريبي.

-**صدق المحكمين:** نسبة إلى استخدام مجموعة من الخبراء والمختصين للحكم على جودة الاختبار ومدى تمثيل بنوده للمحتوى، فإذا كان الاتفاق بين آراء المحكمين مرتفعا دل ذلك على الصدق، كما يجب ألا يقل عدد المحكمين عن خمسة، ويجب أن يكونوا من المتخصصين في المجال مع الاستعانة بمتخصصين في القياس النفسي، ولا تقل درجة الاتفاق على كل بند من البنود عن 80%. حسب ما يشير إليه (صلاح مراد وأمين سليمان، 2005)

تتضح أهمية صدق التحكيم في قيمة ما قد يقترحه المحكمون بعد اطلاعهم على الأداة، فقد يزودون الباحث بملاحظات قيمة يكون قد أغفلها عند بناء أداة القياس، كاقتراح عبارات جديدة ذات نوعية جيدة، أو تعديل فيها أو إضافة أبعاد إضافية...لذلك يجب على الباحث أن يولي اهتماما كبيرا لاستمارة التحكيم او كما تعرف باستمارة الخبرة.

حيث ترى (نادية الزقاي، 2017) أنها أداة لجمع معطيات يكون مصدرها مجموعة الخبراء ذوي الاختصاص يجب أن تخضع لشروط بناء مضبوطة، وشروط تطبيق موضوعية، كما أنها لا تحتمل حين تفريغ أو عرض نتائجها أية عشوائية، لذلك يجب تزويد المحكمين بمعطيات وافية عن الاختبار، حيث يجب أن تشتمل الاستمارة على الهدف من الاختبار موضوع التحكيم، المطلوب أو مجموعة المطالب الموجهة للسادة المحكمين، التعريف الاجرائي للخاصية موضع القياس، تحديد الأبعاد الفرعية للاختبار والبنود التي تنتمي إلى كل بعد مع ضرورة تضمين تعريفات إجرائية لأبعاد الاختبار في الاستمارة.

دون أن يهمل الباحث الواجهة التي تذكر فيها المعلومات الأساسية مثل الجهة التي صدرت عنها استمارة التحكيم (الجامعة، الكلية، القسم...) عنوان الاستمارة، السنة، والتعليمات الموجهة للخبراء، وبياناتهم الشخصية (الاسم واللقب، الدرجة العلمية والتخصص)

2-2- الصدق المرتبط بمحك: يعتمد هذا المفهوم على درجة علاقة درجات الاختبار بالأداء الفعلي على محك خارجي، ويقصد بالمحك الخارجي اختبارا آخرا جيدا أو نوعا من الأداء العملي تستخدم فيه السمة موضع الاهتمام، وهناك نوعين من الصدق المرتبط بمحك هما كالتالي:

-**الصدق التلازمي:** يدل الصدق التلازمي (المصاحب) أو التزامني على حجم العلاقة بين درجات الأفراد على الاختبار ودرجاتهم على محك أخر بحيث لا يكون هناك فاصل زمني (أو فاصل زمني قصير) بين أداء الأفراد على الاختبار وأدائهم على المحك، مثلا العلاقة بين درجات الأفراد على اختبار ويكسلر وبينيه.

-**الصدق التنبؤي:** يهتم الصدق التنبؤي باستخدام درجات الاختبار في التنبؤ بالأداء المستقبلي على مقاييس أخرى هي المحكات، فقد تستخدم درجات اختبار الاستعداد للنجاح في المدرسة للتنبؤ بدرجات الطلبة في مواد دراسية معينة، أو درجات التحصيل في الثانوية لمادة الرياضيات كمؤشر للنجاح في كلية الهندسة، ودرجات التحصيل في الثانوية لمادة العلوم الطبيعية كمؤشر للنجاح في كلية الطب مثلا أو علوم الطبيعة والحياة، فيدل هذا الصدق على مدى كفاءة درجات الاختبار في التنبؤ بسلوك المشارك في الاختبار في وقت لاحق، بمعنى أنه يوجد فاصل زمني لا يقل عن 6 أشهر بين تطبيق الاختبار والمحك. (صلاح مراد وأمين سليمان، 2005)

غير أن المشكلة الأساسية في تقدير الصدق المرتبط بمحك حسب ما يذكر (صلاح الدين علام، 2000) تتركز في إمكانية الحصول على محكات مناسبة، فجمع بيانات تتعلق بالمحكات تعد من مشكلات القياس النفسي والتربوي، فالمحك يجب أن يتميز بخصائص تمكننا من **الثقة** فيما يترتب عليه من بيانات، فإلى جانب الأحكام التقييمية للمحك ينبغي أيضا التحقق من ثبات درجاته وخلوها من التحيز ومواءمة المحك للأداء المطلوب (صدقه)، فإذا اعتبرنا مثلا مقياس تقديرات المشرفين على العمل محكا لأداء العامل ولدينا مقياس نريد التأكد من صدقه هو مقياس الأداء المهني للعامل فيجب التحقق من دقة تقديرات المشرفين وعدم تحيزها لكي تكون ذات فائدة في الحكم على صدق القرار الذي يتخذ في ضوء ذلك.

2-3- صدق التكوين الفرضي: ويسمى أيضا صدق البناء أو صدق المفهوم، ومصطلح صدق التكوين الفرضي اقترحه كل من كرونباخ وميهل عام 1950 الذي يتناول العلاقة بين نتائج الاختبارات والمقاييس وبين المفهوم النظري الذي يهدف الاختبار لقياسه مثل مفهوم الذكاء، القلق، الانطواء، الابتكارية...فصدق التكوين الفرضي يهدف لتحديد التكوينات الفرضية التي يعزى إليها تباين الأداء في الاختبارات وليس درجات المحك أو سلوك الفرد.

لكن السؤال المطروح: ما المقصود بالتكوينات الفرضية؟ التكوينات الفرضية هي أفضل ما يتوصل إليه الباحثون في وقت ما لتوضيح طبيعة السمات والقدرات الموجودة لدى الانسان.

فالتكوين الفرضي إذن هو صفة أو خاصة نفسية نفترض وجودها وذلك لتفسير جوانب معينة من سلوك الأفراد، ويتطلب تقدير هذا النوع من الصدق مجموعة من الطرق نذكر أهمها فيما يلي:

-المقارنات الطرفية (الصدق التمييزي):

يقوم الصدق التمييزي حسب ما يطلق عليه البعض على أحد مفاهيم الصدق وهو قدرة الاختبار على التمييز بين طرفي الخاصية التي يقيسها (بشير معمرية، 2012)، ويمكن تلخيص إجراءات تطبيق هذه الطريقة كما يلي:

- ✓ تطبيق الاختبار وإيجاد الدرجة الكلية التي يحصل عليها كل فرد في الاختبار
- ✓ ترتيب الدرجات من أدنى درجة إلى أعلى درجة أو العكس
- ✓ أخذ 27% من الدرجات العليا و 27% من الدرجات الدنيا ليكون المجموع 54% وما يتبقى من نسبة (46%) لا تدخل في عملية الحساب.
- ✓ ايجاد المتوسطات الحسابية والانحرافات المعيارية لدرجات أفراد المجموعتين العليا والدنيا

- ✓ تطبيق اختبار ت لعينتين مستقلتين (بعد التحقق من استيفائه لشروط التطبيق واختبار تجانس النصفين) ومعرفة دلالته الاحصائية ومن ثمة اتخاذ القرار الذي يتعلق بقدرة الاختبار على التمييز بين المجموعتين الطرفيتين وهو ما يعتبر دليلا على صدق الاختبار.

-الصدق العاملي:

إن التحليل العاملي يساعد مصمم الاختبار على تحديد المكونات الأساسية (العوامل) للسمة أو القدرة التي صمم الاختبار لقياسها، والتي يمكن اعتبارها أبعادا للاختبار، بالإضافة إلى تحديد درجة تشبع مفرداته بكل عامل من هذه العوامل، وهذه التشبعات تمثل معاملات الارتباط بين مفردات الاختبار والعوامل، كما يمكن من خلاله اقتراح نموذج يجسد أبعاد الاختبار وما يندرج تحتها من مفردات، والتأكد من مطابقتها لبيانات عينته.

-الاتساق الداخلي:

من خلال حساب معاملات الارتباط بين درجات الأفراد على البنود ودرجاتهم الكلية على الاختبار، أو حساب معاملات الارتباط بين درجات الأفراد على البنود ودرجاتهم على الأبعاد الفرعية للاختبار التي تنتمي إليها تلك البنود، أو حساب معاملات الارتباط بين درجات الأفراد على الأبعاد الفرعية ودرجاتهم الكلية على الاختبار.

-الصدق التقاربي والتمايزي:

تذكر أدبيات القياس النفسي والتربوي ما يعرف بالصدق التقاربي في مقابل الصدق التمايزي، أو الصدق التوافقي في مقابل الصدق التعارضي، حيث أدرجت ضمن (صدق المفهوم الفرضي) حسب عدة باحثين من بينهم (بشير معمرية، 2012) و(امطاينوس ميخائيل، 2016) رغم أنها تعتمد على حجم العلاقة الارتباطية بين الاختبار واختبارات أخرى قد نعتبرها محكات خارجية.

فالصدق التقاربي يعني دراسة العلاقة بين الاختبار الجديد واختبار آخر مشابه، ثبت صدقه في قياس نفس التكوين الفرضي أو السمة التي يهدف الاختبار الجديد إلى قياسها، فإذا كانت قيمة معامل الارتباط بين الاختبارين مرتفعة، فإن ذلك يعد دليلا على أن الاختبار الجديد يقيس ما يقيسه الاختبار الذي ثبت صدقه، أما الصدق التمايزي نعني به حساب معاملات الارتباط بينه وبين الاختبارات الأخرى التي تقيس تكوينات فرضية أخرى مستقلة ومختلفة عما يقيسه الاختبار، ويتحدد صدق التكوين الفرضي من خلال فحص معاملات الارتباط، فإذا كانت قيم هذه المعاملات منخفضة كان ذلك دليلا على أن الاختبار يتمتع بصدق تمايزي عال. (امطاينوس ميخائيل، 2016)

ويطلق عليه (عبد الرحمن الطريري، 1997) أسلوب التشابه والاختلاف مع خصائص أخرى لتقدير الصدق، حيث يتمثل في العلاقة الممكن حدوثها أو عدم حدوثها مع مقاييس يحددها الباحث على أنها تقيس نفس الخاصية أو لا تقيسها.

ويشير (بشير معمرية، 2012) إلى ما يعرف بالصدق التوافقي أو الاتفاقي في مقابل الصدق التعارضي أو الاختلافي حيث يتحدد الصدق بوضوح بواسطة اكتشاف التوافق والتعارض بين الاختبار وغيره من الاختبارات، وفقا للإطار النظري الذي تنتمي إليه المتغيرات محل القياس، مثل ارتباط الاكتئاب باليأس وبالتشاؤم وبالميل للانتحار، وارتباط المثابرة بالدافعية وارتباط القلق بالخوف توافقيا، وكذلك ارتباط الاكتئاب بالسعادة وبالتفاؤل وارتباط الانبساطية بالخجل تعارضيا.

3-العوامل المؤثرة في صدق الاختبارات النفسية:

هناك عدد من العوامل التي تؤثر في مدى صدق الاختبار تتمثل فيما يلي:

- عوامل تتعلق بالاختبار نفسه: مثل لغة الاختبار إذا كانت فوق مستوى المفحوصين، فإن بعضهم سوف يعجز عن فهم السؤال وبالتالي لا يستطيع الإجابة عنه مما يقلل من مستواه في الاختبار، كذلك غموض أسئلة الاختبار تجعل المفحوص يفسرها تفسيرات متباينة وبجيب عنها إجابات خاطئة، تقلل هي أيضا من مستواه الفعلي في الاختبار.

كما أن سهولة أسئلة الاختبار أو صعوبتها أو طريقة صياغتها تجعل المفحوص يحصل على درجات لا يستحقها، فيتأثر حكمنا عليه بأنه مثلا ممتاز أو وسط أو ضعيف وهي لا تمثل في الحقيقة قدرة المفحوص الفعلية في الاختبار.

- طول الاختبار: إذا كان الاختبار قصيرا يضعف تمثيله للمحتوى وبالتالي يضعف صدقه، وإذا كان طويلا جدا فيدخل هنا عامل الملل وعدم الجدية من طرف المفحوص في الإجابة. (امطاينوس ميخائيل، 2016)

وكذلك ترتيب البنود بطريقة غير ملائمة، حيث من المفروض أن يرتب الاختبار وفق مبدأ التدرج في الصعوبة، فإذا وضعت البنود الصعبة في البداية يقضي المفحوص الكثير من الوقت للاجابة عنها، وقد لا يتسع الوقت للإجابة على البنود السهلة، ثم إن هذا الترتيب قد يضعف دافعيته مما ينعكس سلبا على أدائه.

-عوامل تتعلق بتطبيق الاختبار وتصحيحه: فالعوامل البيئية تؤثر على أداء المفحوص تأثيرا يقلل أو يزيد من مستوى صدق الاختبار، مثل الحرارة، البرودة، الضوضاء ....وغيرها.

كذلك طباعة الأسئلة وعدم وضوحها، التعليمات الغير واضحة، استعمال الاختبار في غير ما وضع له وعدم استعماله مع الفئة التي وضع لها، الوقت المخصص للإجابة غير كاف أو أكثر من اللزوم، فيتأثر الأداء الاختباري بالضغط على المفحوصين في الحالة الأولى، وإرباكهم وإتاحة الفرصة للغش والفوضى في الحالة الثانية، كذلك أخطاء التصحيح كلها عوامل تضعف معامل الصدق.

-عوامل تتعلق بشخصية المفحوص المؤثرة في إجاباته: فالتخمين أو الغش أو اضطراب المفحوص في الاختبار تجعله يحصل على نتيجة لا تمثل قدراته الفعلية، وهذا ما يؤثر على صدق الاختبار. (عبد الرحمن الطريري، 1997)

المحاضرة 07: الخصائص السيكومترية للاختبار (ثبات الاختبار)

1- مفهوم الثبات:

يرى (محمد غنيم، 2004، ص61) أن مفهوم الثبات يعبر عن مدى تماسك بنية الاختبار، كما يعبر أيضا عن دقة الاختبار فيما يزودنا به من بيانات عن الخاصية المراد قياسها، فثبات الاختبار يعبر عن علاقة الاختبار بنفسه والتي يمكن تقديرها بإعادة تطبيق نفس الاختبار على نفس العينة أو باستخدام اختبارات متكافئة معه.

أما (صلاح الدين علام، 2000، ص131) فيرى أن ثبات درجات الأفراد على الاختبار هو مدى خلوها من الأخطاء غير المنتظمة (العشوائية) التي تشوب عملية القياس، أي مدى قياس الاختبار للمقدار الحقيقي للسمة التي يهدف لقياسها، فدرجات الأفراد على الاختبار تكون ثابتة إذا كان الاختبار يقيس سمة معينة قياسا متسقا في الظروف المتباينة التي قد تؤدي إلى أخطاء القياس، فالثبات بهذا المعنى يعني الاتساق أو الدقة في القياس.

ويشير (سعد عبد الرحمن، 2008) إلى أن ثبات الاختبار يعني دلالة الاختبار على الأداء الفعلي أو الأداء الحقيقي للفرد، هذا الأداء يعبر عنه بالدرجة الحقيقية التي حصل عليها الفرد في اختبار ما (دح)، والأداء الحقيقي هو جزء من الأداء العام أو الأداء الكلي الذي يعبر عنه بالدرجة الكلية (دك)، وهي الدرجة الملاحظة أو المسجلة على الاختبار والتي حصل عليها الفرد، أما الجزء الآخر فهو الأداء الذي يعود إلى أخطاء الصدفة أو الظروف الخارجية ويعبر عنه بدرجة الخطأ (دخ).

ولأن: الدرجة الكلية = الدرجة الحقيقية + درجة الخطأ

وتبين نظرية الثبات حسب (محمد غنيم، 2004، ص65) أن مفهوم ثبات الاختبار هو خاصية مجموعة من درجات الاختبار وليس خاصية درجة شخص واحد فقط، ولما كان الهدف الأساسي لأي